An Interactive Support Tool to Convey the Intended Message in Asynchronous Presentations

Andrés Lucero
User Experience and Design
Nokia Research Tampere
FI-33721 Tampere, Finland
andres.lucero@nokia.com

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

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

ABSTRACT
In this paper we introduce an interactive wall-mounted display tool that supports conveying the intended message or ideas in asynchronous presentations. The tool allows to easily record presentations while capturing the richness of the presenter's individual presentation skills and style. The tool 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 to control the playback of audiovisual presentations. Once recorded, the presentation can be played back, explored and commented using a flexible and intuitive interaction based on hand movements and body position (i.e. proximity). We focus our work on a case study where designers present their mood boards (i.e. collages) to their clients. Evaluations with professional designers show that they are able to use the tool with no prior training, see a practical use of the proposed tool in their design studios, and see gesture trails as a creative tool for expression and aesthetics.

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

General Terms
Design, Human Factors, Performance.

Keywords
Gesture-Based Interaction, Distributed Collaboration.

1. INTRODUCTION
Knowledge workers [10] (e.g. designers, researchers, marketers, etc.) make widespread use of computer-generated overhead slides as means to share and present their work. The most popular software tools to create these sets of slides include Microsoft's PowerPoint and Apple's Keynote. The presentations produced by these tools tend to consist of a collection of single static slides that are changed sequentially over time [13]. Common difficulties that presenters and audience experience while going through sets of slides are the lack of overview (e.g. current position within the presentation, what are the previous and upcoming topics) and breaking the linear structure (e.g. skipping slides, jumping back and forth). Distributed and asynchronous presentations (e.g. webcasts) introduce new challenges, as the author of the slides is no longer available to make sense of the contents. Moreover, other modalities of the presentation such as the presenter's body language and hand gestures are missing, making it more difficult for the audience to understand the intended message [8, 23].

Desktop and digital systems provide ways of creating and replaying both co-located and distributed asynchronous presentations. However, these systems do not provide the conditions to simply stand and give an oral presentation, and make full use of their individual presentation skills and style. Moreover, these systems fail to provide flexible ways of browsing presentations, other than in a linear manner.

In this paper we present the design and evaluation of the Funky Wall, an interactive tool that supports conveying the intended message or ideas in asynchronous presentations. Our key contribution is a tool that easily records different aspects of a presentation (i.e. gesture, sound and visuals) and later provides the possibility to navigate through the presentation based on the recorded gestures and sounds. We believe the ideas presented here on gesture trails, image visibility and sound can apply more generally to any audiovisual presentations thus providing an alternative to linear navigation and the PowerPoint slide paradigm.

The rest of this paper is structured as follows. We first provide background information on the user studies that were conducted to gain understanding of how augmented reality could support professional users in their work, and on the resulting vision for a future work environment (i.e. funky-design-spaces). Next, we review the relevant related work. Then, we present the design principles and interaction techniques of the Funky Wall tool. Finally, we report the results of evaluations, followed by discussion and conclusions.

2. BACKGROUND
The ID-MIX research project [15] explores the relevance and impact of providing augmented reality support tools for professional users in their work. It does so by exploring why and how designers use mood boards in the early stages of the design

Ace 2009, Oct 29-Oct 31, 2009, Athens, Greece

process, and how augmented reality can support mood boarding by following a user-centered design approach.

To gain a better understanding of design practice, we have conducted several user studies using diverse methods. By using cultural probes (n=10) in a professional context we were able to identify a relevant task for designers: making mood boards. Designers commonly use mood boards in the early stages of the design process [15], to explore, communicate, and discuss ideas together with their clients. These boards can be created with different types of media although designers use most frequently images to say something or tell a story about the target audience, product, and/or company they are designing for.

After the cultural probes study, we have conducted contextual inquiries with Dutch industrial designers (n=4), as well as mood-board interviews with Finnish fashion and textile designers (n=10) to get a better understanding of why designers use mood boards and how they create and use them.

The data from these user studies was fed into co-design sessions with Finnish (n=8) and Dutch (n=6) designers where researchers and people (i.e. designers) collaboratively came up with new concrete ideas that support mood-board making with augmented reality. In these co-design sessions, different types of locations, tasks, and materials spark conversation and an exchange of ideas between researchers and participants (i.e. end users). By engaging in activities that rely on visual and tangible materials the complete design team involving researchers and participants is able to approach a given design problem from different entry points or perspectives and thus come up with novel design concepts. The ideas for the Funky Wall originated in these co-design sessions.

The basic underlying principles for the Funky Wall together with a short exploratory study of the tool have been previously presented in a short format elsewhere [16]. In this paper, we provide a detailed explanation of the interaction techniques as well as present the results of a larger user evaluation of the tool.

2.1 The Mood-Board Making Process

From our previous studies, we have identified the following five stages in the mood-board making process: 1) 'collecting', 2) 'browsing', 3) 'connecting', 4) 'building', and 5) 'presenting'. In the final stage, 'presenting', designers usually meet their clients face-to-face to share and discuss the intended story behind the mood board. Designers create a single large mood board or a series of smaller booklets for their clients to keep and share with other stakeholders. However, in large companies mood boards are made available on the company's Intranet for different departments to look at and experience them (e.g. marketing, sales, advertising). 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 explaining the mood board. In these two 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? In this paper we present the Funky Wall that explores how 'presenting' mood boards asynchronously can be supported in new ways.

2.2 The funky-design-spaces Vision

The previous user studies with Dutch and Finnish mood-board makers have shown that the process of making mood boards takes place in different contexts both in and outside of the design studio. For example, in the beginning of the process, designers can spend a considerable amount of time looking for images. Designers prefer going through their large collections of magazines in a comfortable place where they can freely start creating ad hoc piles of magazines and pictures. The process of making mood boards also goes beyond the activities and the time spent collecting and arranging images on a table. Mood-board making is a dynamic and iterative process in which designers constantly switch between searching and making (e.g. layout, gluing), then going back again to find the missing image that fits. Mood-board makers must also go out and meet their clients at different stages of the process to discuss ideas and present their results.

Based on these findings, we have come up with a vision for a new holistic design studio, a comfortable space that facilitates creative thinking in designers. Within this larger context, different funky-design-spaces or tools that are interconnected and stimulate designers to move around their design studios would support the process of making mood boards.

Designers present their mood boards to clients in face-to-face meetings by putting the mood boards up on the wall and giving an explanation in a stand up position. With this in mind, we have decided to design the interaction on a large wall screen to encourage presenting mood boards in a more natural setting within the design studio.

3. RELATED WORK

3.1 Gesture and Speech-Based Systems

Clark and Brennan [4] and McNeil [17] 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 systems, vertical displays, multi-device environments, and 3D virtual environments [1, 2, 7, 24]. Several projects have studied the application of hand gestures and movements to support human-computer interaction. Bekker et al. [2] looked at gestures that people use when engaged in design activities and, classified them into 4 groups: kinetic, spatial, pointing and others. They made two interesting observations that are relevant to this work: they observed that gestures are carefully synchronized with speech and that gestures occur in relation to the spatial organization of participants and work artifacts [2]. This is in line with the work of McNeil [17] who argues that gestures are an integral component of language. Hardenberg & Bérard [7] studied the usability of bare-hand human computer interaction. The study focused on using static hand postures for issuing a command, and fingers for pointing. They also proposed a number of application areas, one of which is a wall projection system. From the user's perspective, the complete system consists of an interface projected on the wall. The study demonstrated that the proposed prototypes could indeed be controlled using hands-only interaction [7].

An example of a public display system that is controlled by gestures was presented in [24]. 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. They also proposed four interaction phases that are based on the distance between the user and the display: ambient display (furthest distance), implicit interaction, subtle interaction and personal interaction (closest distance).

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 [1, 22, 23]. 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. Tivoli [18], an electronic whiteboard, is another example of a system where a free-form gesture based interface was employed to enrich presentations and discussions during meetings. VideoWhiteboard (VideoDraw) [22] enables remote collaborators to work together much as if they were sharing a whiteboard. The key aspect of the system is that collaborators not only see drawings but also the shadows of the gestures made by the collaborator at the remote site. The authors argue that the gestures' shadows provide a stronger sense of co-presence. Another system that employs gesture shadows is Mixed Presence Groupware [23]. Kirk et al. [11] studied different ways to represent gesture shadows (hands, hands and sketch, sketch only). They found that unmediated video representations of hands speed up performance without affecting accuracy.

Most of the previously presented research looks at real-time communication, where collaborators interact simultaneously (they can be either physically co-located or in different locations). It is, however, unclear to what extent their research findings can be applied in situations where communication happens offline rather than in real time.

3.2 Capturing and Browsing Meeting Content

There is a large area of research that looks at optimal meeting content capturing and browsing [6]. 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 [12]. The main difference with our tool is that most of these systems only look at speech and handwriting notes and not at hand gestures, as means to segment the meeting and to identify bookmarks [6, 3].

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. [9] use a motion estimation technique to detect key frames and segment the video (recorded presentation). The proposed method is robust with respect to slide motions, occlusions and gestures. In addition it enriches the slides by indicating where the speaker is pointing. Another example is the Active Multimodal Presentations [5] concept.

VideoPassage [21] is an interactive asynchronous video messaging system that allows both spatial citing to correct information presented by the other party, as well as temporal citing to insert a message within that of the other party.

The main difference is that the Funky Wall attempts to create a structure using only implicit information. Plus it uses gestures and segmentation (together with speech and vision) as an additional information channel (that is presented to the viewer through gesture traces).

4. DESIGNING THE FUNKY WALL

Based on the requirements we gathered from designers, we have decided to design support for the final part of this process, '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 movements and body position (i.e. proximity), 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 provided both sides own the same tool.

4.1 Design Principles

4.1.1 Proximity-Based Interaction

The Funky Wall employs four different ranges of interaction depending on the designer's proximity to the mood board: 'showing', 'contemplating', 'replaying', and 'exploring'.

Different interaction modalities and functionalities are made available to the users (i.e. designers or clients) based on the distance from the screen. Gesturing close to the screen is used to record a presentation or comment on an existing presentation (less than 0.5 meters). When the presentation has been created, designers or clients can then contemplate the mood board from a distance (more than 2 meters, no gestures), they can replay the entire presentation (gesturing between 1.5 and 2 meters), or they can also explore specific parts of the recorded presentation (gesturing between 0.5 and 1.5 meters). Our four ranges of interaction resemble the ranges proposed by Vogel and Balakrishnan [24] and in Hello.Wall [19].

4.1.2 Intuitive and Flexible: Hand movements and 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). We have defined intuitive interaction as tools that allow designers to simply walk up to them and start performing tasks using the designers' current skills and knowledge on the task that is being supported. There is no need to read manuals or learn new skills to master the functions provided by the tools.

To keep interaction simple, designers do not require any specific control over the tool, they only 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 as they would do in case of an actual presentation.

When a presentation is given, the tool automatically records and keeps essential aspects of the presentation at three main information layers and in combination with the mood board itself: gesture, sound (speech) and vision. These information layers are analyzed in order to split the presentation into a number of meaningful segments. Each segment is associated with a specific time, interval and specific area on the mood board (not every

segment has this property since part of the presentation can cover general aspects of the mood board without relating to any specific part). There are three important attributes that are used for splitting: the location of hands, the acceleration of gestures, and pauses in speech. Preliminary observations show that location and speed of the gesture can be used as means to create meaningful indices, i.e., to associate the speech layer with a particular area of interest. These attributes allow segmenting the audio file and associating every segment with a particular area in the mood board.

4.1.3 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 one for the client.

4.2 Interaction Techniques

4.2.1 Showing

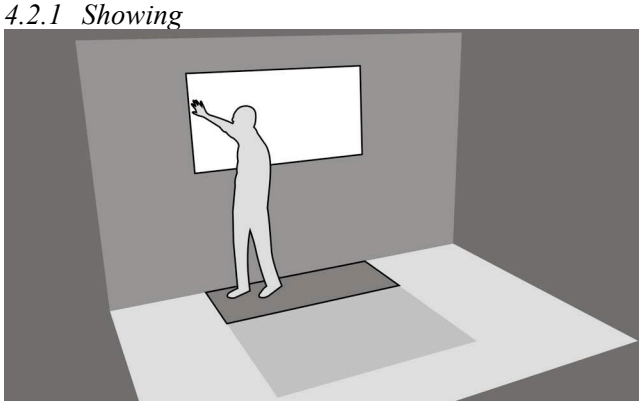

Figure 1. Showing by gesturing next to the screen (<0.5m).

To begin recording their presentation, designers simply need to gesture and speak in front of the mood board at close range (less than 0.5 meters from the screen) (Figure 1). As designers are using both hands to gesture in front of the screen, the tool 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% (Figure 2). Additionally, ten seconds after the gesture has been overlaid on top of the mood board it gracefully fades out to 25% opacity. In this way, recent traces made by the designer are made more prominent than old ones.

The tool captures and segments both the speech and the natural hand movements made by the designer, hence creating associations between audio segments and gestures. Wang Freestyle [14] showed that such automatic segmentation is possible. Wang Freestyle is a system that allows users to annotate a document (TIF file) using a stylus, and speak while they are annotating. The file is then emailed back to the intended recipient who can play back the audio synchronized with the marks made.

Figure 2. White traces of the gestures made by the designer are displayed by the tool.

4.2.2 Contemplating

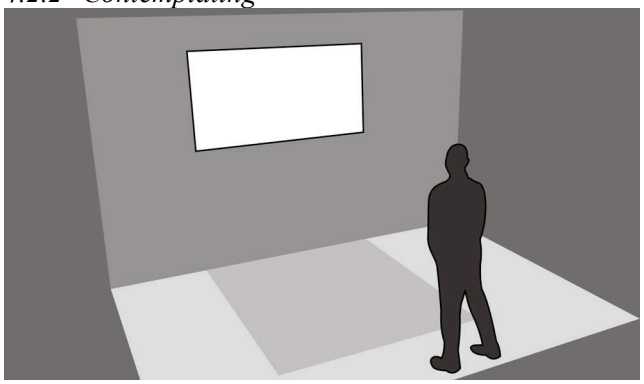

Figure 3. Contemplating the mood board (no gestures >2m).

Once a presentation has been completed, spectators (i.e. designers or clients) can review the recorded presentation (Figure 3). Users contemplate the mood board from a distance greater than 2 meters, allowing for a more comfortable and clean overview of the mood board. No gesturing is possible at this range.

4.2.3 Replaying

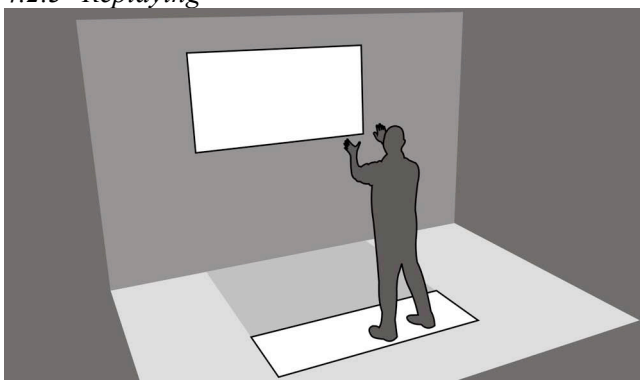

Figure 4. Replaying the entire presentation (gesturing 1.5-2m).

Spectators can replay the entire presentation by approaching the screen at a distance between 1.5 and 2 meters from the screen (Figure 4). At this range, users can have an overview of the associated recorded content created by the designer while showing the mood board, and which the tool recorded. Raising the

dominant hand results in displaying a static representation of all gestures made during the presentation semitransparent on top of the mood board (Figure 5). Raising the non-dominant hand triggers the complete speech or audio explanation. By putting both hands together, the recorded speech is played and the transparent dynamic gestures unfold as the presentation progresses.

Figure 5. Replaying: visual feedback displaying all gestures made in the presentation semitransparent on the mood board.

Having an overview of all gestures by displaying them as a static representation allows spectators to quickly see areas of high interest where gestures concentrate (Figure 5). This might be helpful for example if the spectator wonders whether or not the designer has addressed specific parts during the presentation.

4.2.4 Exploring

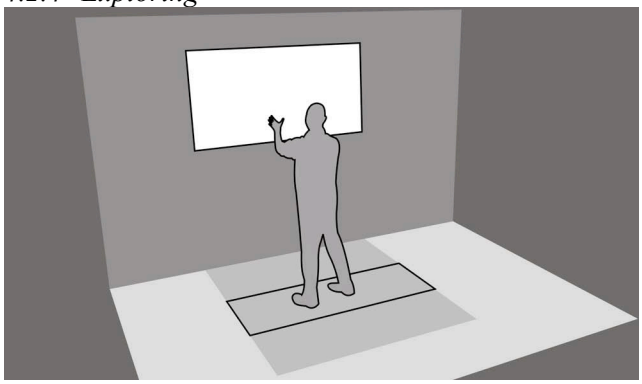

Figure 6. Exploring specific parts of the presentation (gesturing 0.5m-1.5m).

If spectators want to explore specific parts of the mood board, they can take one step closer towards the screen (between 0.5 and 1.5 meters) (Figure 6). 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 (Figure 7) while the rest of the traces remain hidden. These overlaid traces of gestures serve as guides for retrieval. We provide visual contextual feedback so that spectators are able to identify time-based connections between associated explanations (i.e. gesture and speech) within the presentation. The tool highlights both the explanations made by the designer just before and immediately after the currently selected gesture. The currently selected gesture is displayed in white, while 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.

Figure 7. Exploring a static representation of the traces made on top of the mood board using the dominant hand.

Putting both hands together display the dynamic gestures together with the corresponding spoken explanation. The tool allows the entire mood board to remain visible while individual areas are highlighted, and gesture trails are triggered.

If after reviewing the presentation the designer is unsatisfied with the results, they can go back and make the presentation once again by following the procedure described for 'showing'. The assumption here is that mood-board presentations usually last between 5 and 8 minutes. Therefore, instead of providing a tool that allows editing specific parts of the presentation, we propose that they make the entire presentation over.

The distances to the mood board that define the ranges of interaction were derived experimentally. The general guidelines were: 1) the size of each interaction zone should be large enough to avoid unintentional switching between different interaction zones; 2) the interaction zones that require more precision should be located closer to the mood board (showing and exploring); 3) the distance for the overview zones (replaying and contemplating) need to be large enough to have full overview of the mood board (i.e. a larger mood board would require larger distances).

4.2.5 Supporting Two-Way Communication

When designers are satisfied with the recorded presentation, they can share it with their clients who are located at a remote location. By having a similar Funky Wall in their offices, the clients can hear and see the associated explanation or story that the designer originally wanted to convey. The clients can explore the entire presentation or specific parts of it by following the procedures described in 'contemplating', 'replaying', and 'exploring'. But more importantly, to truly support two-way communication, clients must be able to give designers feedback based on their perception and interpretation of the mood board.

Clients can reply and add their own comments to the mood board using the same interaction modality used by designers to record their presentations described in 'showing' (gesturing next to the screen). In this way, designer and client can have several iterations throughout the mood-board making process.

5. EVALUATION

We conducted user evaluations of our prototype in order to test the usefulness and usability of the Funky Wall (Figure 8). 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.

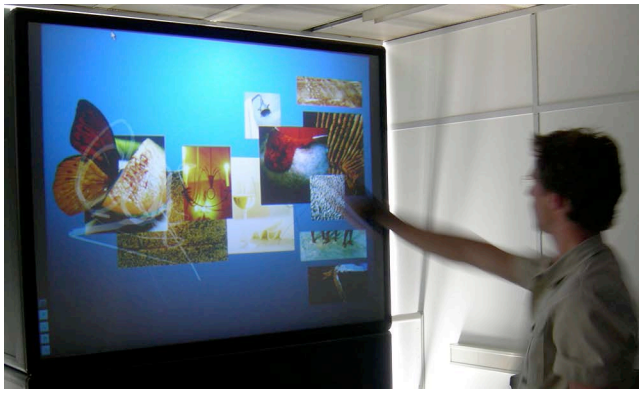

Figure 8. Experiment setup. A designer exploring different parts of the presentation wearing the interactive gloves.

We recruited 12 practicing designers with at least 5 years of experience in design practice (13 years of experience on average). The participants varied in their education (university/academy), age (between 30 and 46), gender (9 male, 3 male), and dominant hand (10 right, 2 left). The evaluations were conducted individually. All sessions were recorded on video.

5.1 Tasks

In the first part of the study, we asked participants to create their own story around a mood board that we gave them and present it to us. Each participant was told that during their presentation, they would be using a tool that tracked their hand movements and record their speech, while the tool would display traces of their hand movements. This part lasted 5 minutes.

In the second part of the study we asked them to play the role of the client by letting them discover the story behind a mood board that was provided to them. Each participant explored a presentation we had prepared in advance using the tool. Following a brief description of the interaction we allowed them to freely explore the functionality and get acquainted with the application. This part took 10 minutes.

In the third part of the study, we asked them to walk us through their experiences while creating their own presentation and then while exploring the given presentation. They shared their thoughts on their interpretation of the different stages of the interaction and the feedback provided. In this final discussion, we wanted to assess the relevance of our tool for presenting mood boards to their clients. We also tried to see if they were able to perform the hand gestures. This discussion lasted for 30 minutes, on average per participant.

5.2 Implementation

A tool was set up using a desktop PC connected to a back-projection screen with resolution 1024x768 pixels and physical size 2.0 x 1.5 meters, 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 gloves were made in Lycra to allow a comfortable fit for different sizes of hands. Participants stood in front of the screen (Figure 8). The application was written in C# and used OpenGL for visualization purposes. Both the presentation and replay parts of the prototype were fully functional. The analysis phase, where the presentation is segmented, was done manually.

5.3 Findings

5.3.1 Participants Agreed with the Principles

Designers were positive about the general underlying principles behind the tool. In general terms, they agreed with the way the tool provided support for presenting mood boards:

× "I think that with this system you can easily identify different parts of the presentation and play them back again. In a way, you have removed the presenter but you have kept the gestures and the impact the gestures make, which is nice." [P2]

× "It's beautiful! I like the idea a lot. It's very stimulating. In case of long and complex presentations this allows you to have reminders of where certain parts of the presentation were, something like chapters. You really get the feeling like you see the entire mood board and you can focus on the subjects you like by zooming into some parts. The feeling is good, you feel in control of the presentation." [P3]

× "I do like the gestures and maybe you could recognize people by (their gesturing) after some time. The way you tag automatically the spots with text that's really beautiful I think. So that's very, very rich." [P10]

5.3.2 Hand Gestures

In the first part of the study where participants were asked to present a mood board using the tool, designers were able to interact with the tool with no prior training. They especially liked the naturalness and simplicity of the interaction through hand movements. However, in the second part of the study participants experienced some difficulties when exploring a presentation by triggering sounds. They specifically expressed some concerns about the amount of gestures they would have to perform and fatigue. They also found some of the gestures awkward or uncomfortable to make (e.g. putting both hands together in mid-air to trigger a sound):

× "It can be a bit heavy in terms of all the gestures you have to make, but it is easy to step into (and move between the different proximity areas). But it's nice that it is physical. It's refreshing!" [P3]

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

5.3.3 Proximity-Based Interaction

We also wanted to hear from designers on the idea of using proximity-based interaction. Designers welcomed the introduction of this interaction style to reveal different parts of the tool to support the presentation of mood boards. They told us that the coupling between stages or locations and the available activities was natural as can be seen in these comments:

× "I think it's good because if you step back you immediately make it clear that I want to have the overview. If you go closer you want to have some detail or talk to somebody. And you go back again and you listen and you want to see the whole picture. So this is quite logic for me." [P8]

× "It totally makes sense that you say the place where you are standing at the back is an overview and (you increasingly) get more and more interactive with the system until the point where you are actually touching (the screen) and sending back your comments." [P10]

A few participants mentioned some minor difficulties in knowing the exact location they were in (i.e. showing, contemplating, replaying, exploring) due to a lack of feedback. This happened especially when designers would move between different parts of the tool (i.e. closer or further away from the display) without performing any hand gestures (i.e. keeping their hands in a resting position next to their body):

× "I miss some kind of feedback to know where I am standing. There should be dynamic transitions between locations." [P3]

5.3.4 Visual Feedback

Regarding the visual feedback provided by the tool by showing the traces on top of the mood board, the discussions were centered around three main topics: amount of visual information, dynamic gestures, and feedback for previous and next speech segment. In relation to the amount of visual clutter, participants had different opinions. Some participants commented on the amount of visual clutter that the gestures created while others suggested filtering out or grouping some of the gestures:

× "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]

× "The mood board disappears behind the gesturing. So it might be good to somehow filter it or simplify it towards blobs because I think it's a bit sad that the mood board fades." [P9]

Participants also commented on the helpfulness of being able to play back the dynamic gestures on top of the mood board as they heard the explanation. They especially reflected on the positive impact the dynamic gestures have on the overall presentation. Participants said the dynamic gestures made it richer, more alive, and more human than other types of standalone presentations (i.e. PowerPoint):

× "It helps to better explain the picture. (Having dynamic gestures played back) is really helpful. It enriches the experience. It gives a touch of sensibility as well; you are more connected. Although you are not there present anymore as a designer, 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. I think that is nice, it makes it more human." [P2]

× "(Seeing the dynamic gestures) is funny; you really get the feeling that the designer said this and was pointing while he was doing it." [P3]

The final aspect of visual feedback that designers reflected on was showing the previous and next speech segment together with the currently selected one. Designers saw this aspect as a bonus as it helped them get into the context of the presentation:

× "I really liked being able to explore temporally, going back and forth. There is a real nice coupling. In traditional presentations you have no cues about what is happening, where am I and where am I going to go next. That is a really nice aspect of this system. 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]

5.3.5 Rehearsing Presentations

Finally, participants reflected on how the Funky Wall could potentially become a support tool to improve presentation skills. Displaying the gestures in visual form made them more aware of how they use their hands during presentations, which could influence the way they present:

× "I also see it as a rehearsal tool so you can really put up a presentation, see how you did, what you forgot, and improve it." [P2]

× "It's a good idea for the presenter. It forces you to think of the whole structure of the presentation. This system could help (you) become a better presenter." [P5]

6. DISCUSSION

6.1 Tool Feasibility

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 tool. However, based on the results reported in the literature and the analysis of gesture-speech synchronization automation, our tool seems feasible [20].

For segmentation our tool 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) [25]. Stifelman [20] 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 have observed that speed can be used to separate between explanations of specific parts (slow movements near the surface of the display), connections between different parts (fast long movements), and the general discussion of the mood board (often fast short movements further away from the display).

6.2 Applying our Approach to Presentations

A similar approach can be used in webcasting, such as ePresence (http://epresence.tv/) or Microsoft webcasts (http://www.microsoft.com/events/). Webcasts are archived and can be accessed many times. By adding a gesture layer we can enrich the presentations, and improve the efficiency and understandability of the presentation. We can also use gestures to create indices or bookmarks that would help to browse through presentations. This is of course more applicable to highly visual presentations that would naturally lead to many pointing and outlining gestures.

6.3 Using other Media to Record and Play

We believe that the use of gestures allows designers to more clearly express the feelings and ideas for a mood board and therefore can enrich the presentation and improve the way that the client can later perceive the mood board. 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. Part of the richness will be lost, but nevertheless the message can be conveyed. In principle, the presentation could also be done on a desktop but we fear that the added richness will be lost.

7. CONCLUSION

The Funky Wall is an interactive wall-mounted display tool that supports asynchronous presentations while capturing the richness of the presenter's individual presentation skills and style in situations when face-to-face communication is not possible. The tool allows to easily recording a presentation, as well as to play back, explore and comment on it using a flexible and intuitive interaction based on hand movements and body position (i.e. proximity).

We have evaluated the tool 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 tool in their design studios. Participants felt that the tool 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. Gestures could be used as a creative tool for expression and aesthetics. Regarding the hand movements, participants also liked the naturalness and simplicity of the interaction.

We believe the three-layer approach we harnessed can potentially allow more flexible presentation browsing while also reducing the required bandwidth. Our findings can be generalized to other kinds of audiovisual presentations where face-to-face communication is not available, and may provide a welcome alternative to the PowerPoint slide paradigm.

8. REFERENCES

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

[2] Bekker, M, Olson, J. and Olson, G. 1995. Analysis of gestures in face-to-face design teams provides guidance for how to use groupware in design. In *Proc. DIS '95,* ACM Press, 157-166.

[3] Chiu, P., Kapuskar, A., Reitmeir, S. and Wilcox, L. 1999. NoteLook: Taking notes in meetings with digital video and ink. In *Multimedia '99,* ACM Press, 149-158.

[4] Clark, H.H. and Brennan, S.E. 1991. Grounding in Communication. In: Resnick, L.B., Levine, R.M., Teasley, S.D. (eds.) Perspectives on socially shared cognition, 127-149. Washington, DC: APA.

[5] Elsayed, A. 2006. Machine-Mediated Communication: The Techology. In *Proc. ICALT'06.* IEEE.

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

[7] von Hardenberg, C. and Bérard, F. 2001. Bare-hand human-computer interaction. In *Proc. PUI '01,* vol. 15, ACM Press, 1-8.

[8] Hartley, R., Elsayed, A. and Pesheva, M. 2005. The Anatomy of an Active Multimodal Presentation in Educational Contexts. In *Proc. of ICALT'05,* IEEE, 762-763.

[9] 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, 595.

[10] Kidd, A. 1994. The marks are on the knowledge worker. In *Proc. CHI '94.* ACM Press, 186-191.

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

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

[13] Lanir, J. and Booth, K. S. 2008. Presentation tools for high-resolution and multiple displays. In Proc. of HCC '08. ACM Press, 61-68.

[14] Levine, S.R. and Ehrlich, S.F. 1995. The Freestyle system: a design perspective. In: Baecker, R.M., Grudin, J., Buxton, W.A., Greenberg, S. (eds.) *Human-Computer interaction: Toward the Year 2000,* Morgan Kaufmann, 871-880.

[15] Lucero, A. 2009. *Co-Designing Interactive Spaces for and with Designers: Supporting Mood-Board Making.* Ph.D. thesis, Eindhoven University of Technology.

[16] Lucero, A., Aliakseyeu, D., and Martens, J. 2008. Funky wall: presenting mood boards using gesture, speech and visuals. In *Proc. AVI '08.* ACM Press, 425-428.

[17] McNeil, D. 2005. *Gesture and thought.* Chicago: University of Chicago Press.

[18] Pedersen, E., McCall, K., Moran, T. and Halasz F. 1993. Tivoli: an electronic whiteboard for informal workgroup meetings. In *Proc. CHI'93,* ACM Press, 391-398.

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

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

[21] Takada, T. and Harada, Y. 2000. Citation-Capable Video Messages: Overcoming The Time Differences without Losing Interaction. In *Proc. i3 Annual Conference,* 31-38.

[22] Tang, J. C. and Minneman, S. 1991. VideoWhiteboard: video shadows to support remote collaboration. In *Proc. CHI '91,* ACM Press, 315-322.

[23] Tang, A., Boyle, M. and Greenberg, S. 2004. Display and presence disparity in Mixed Presence Groupware. In *Australasian User interface - Volume 28,* ACM Press, 73-82.

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

[25] Wang, M.Q. and Hirschberg, J. 1992. Automatic Classification of Intonational Phrase Boundaries. *Computer, Speech, and Language,* 6, 175--196.