Enchantment under the sea*: An intelligent enviroment*

Norma Elva Chávez, Rodrigo Savage
Universidad Nacional Autónoma de México (UNAM)
Departamento de Ingeniería en Computación,
Facultad de Ingeniería, Ciudad Universitaria
México City, México
normaelva@fi.b.unam.mx

Saiph Savage
University of California, Santa Barbara,
Computer Science Department,
USA
saiph@cs.ucsb.edu

Abstract— **Disc Jokeys (DJs) generally mix music in a confined isolated space. This can make the DJ have depression sentiments and it can also difficult the DJ's understanding of his public. We present Enchantment Under The Sea: a new intelligent environment that allows the disc jokey to roam freely, interact directly with his audience, receive informative feedback about the public's social interactions, while also respecting the DJs privacy concerns. The music interface is controlled using Microsoft's wireless touch mouse with ubiquitous gestures that resemble dance moves. The music mixing interface is displayed on the walls of the event, where two different display modalities are enabled: open interface, in which the public can observe all of the DJ's decisions with the music mixing interface and also actively give music suggestions to the DJ. And a closed interface, where all the music controllers are mapped to sea animals, that only the DJ knows the mapping to, thus providing privacy to the DJ's work. The public's social interactions are measured with sonar sensors whose data is provided to the DJ through the musical interface. We report results of a controlled usability inspection.**

Keywords: Multimodal interaction, Music Mixing, Intelligent, Environments, Musical Environments, Social Signals, Reality Mining

I. INTRODUCTION

In night clubs there is a celebration for life through dance and music. It is the disc jokey (DJ) who makes the crowd enter a festive state of mind. A DJ's job is not only to select songs, it is also to create shared moods; Their goal is to comprehend the sentiment of the audience and transport them to a better place. The problem is that in current club environments the DJ has to understand the crowd's feelings from a far: the DJ is generally locked away in an elevated somber booth with their eyes, hands and ears preoccupied with the audio hardware [4] [5]. Most DJs express that the solitude they experiment due to this isolation is often one of the most difficult aspects of their job [5]. In this project, we sought to create an intelligent environment that functioned as an outlet to some of the conflicts DJs encounter while performing. There has recently been many mobile music mixing applications in the market that allow the DJ to have mobility and be freed from being in a confined restricted space (see e.g., Nokia's DJ Mixer, or DJ Rig for IPhone) . The drawback of these applications is that they offered no eyes-free interaction, therefore the DJ was required to observe the screen at all times, and attention toward the audience was lost. Savage et. al, offered an eyes free multi-modal mobile music mixing application, by offering feedback in the form of vibrations and audio signals. But using vibrations and sound as the main feedback channels limited the system's responses. We believe that to provide a better user experience the DJ's environment can be used for visualizing the system's feedback, therefore allowing the system to give the DJ a boarder range of information. We envisioned displaying the music mixing interface and all inherent feedback visualizations on the walls of the event. Allowing the DJ to easily have access to it, independent of where they were situated. The DJ with a wireless controller would then be able to interact anywhere and at any time with the interface. Now, in this approach every person in the party would know exactly what the DJ is doing. This in times can be positive, as it allows for transparency. But in many cases, upon seeing the mixing interface people might wish to influence the music selection, possibly disturbing the DJ's job. To overcome these social problems, two different music mixing interfaces are implemented; one that follows a traditional music mixing interface and another that maps the controllers to figures that only the DJ understands.

Therefore, while the crowd enjoys whimsical figures on the wall, the DJ receives privacy to their work as well as visual feedback of their music interactions. Our system is named "Enchantment under the sea" because the whimsical figures the music controls are mapped to are all sea creatures. Walls provide a large visualization space, with which more information can be given to the DJ, such as feedback about the social interactions occurring in the event, which is key to the DJ's performance. In this work we sought to automatically detect social interactions and provide this information to the DJ as informative clues about their performance. For machines to autonomously understand social interactions, "social signals" are utilized. A social signal (According to Poggi and D'Errico [6]) is a communicative or informative signal that conveys information about social actions, social interactions, social emotions and attitudes.

Through social signal processing, i.e., by applying traditional signal processing techniques to social signals, machines can analyze human social behavior [7]. Mazalek et.al. [2], offered a form of analyzing a party's social environment and based on this analysis certain music was played automatically. Contrary to [2], our work leaves the music selection to the human expert, but does focus on autonomously measuring the party's overall social interaction . To measure social interaction, we use sonar sensors, with which the percentage of people that are sitting is measured. The value from each sonar sensor is then transferred to a server that controls what the interface displays. The music mixing interface then informs the DJ of the levels of social interaction found in the room. Our contributions in this paper are as follows: we present a novel intelligent environment that allows for new types of interactions between a music performer and his public. We enable the music performer to receive informative feedback about the social interactions of his crowd, and also allow the music performer to decide just how open or closed they wish to be about their work. It is to our best knowledge that no other intelligent environment has been built to remedy the privacy concerns encountered in modern party settings involving a DJ., as well as to provide the DJ such relevant informative feedback. In the following sections we explain in more detail how our intelligent environment was constructed, the usability inspection that was performed with found insights, and finish with conclusions.

II. TECHNICAL DETAIL

As mentioned in the introduction, it was very important for our project to allow the DJ to roam the party environment and engage with their crowd. A way for enabling this, is by integrating Bluetooth stereo headphones and a mechanism that allows the DJ to always be able to control the music mixing application, independent of where they are situated. To achieve the second task, we choose for the DJ to interact with our music mixing interface through Microsoft's wireless touch mouse. The touch mouse is a wireless mouse, coupled with a multi-touch sensor. The mouse takes a snapshot of the sensor's values approximately every 8 milliseconds. The mouse then cleans the data and delivers it to a host computer, where by utilizing machine learning techniques, fingertips can be tracked and gestures recognized. Touch mouse was selected primarily because it allows for a wireless connexion and also because of its promising touch capabilities. The gestures selected were all movements that could easily be confused with dancing. Gesture integration is a key component to our system because it is a medium through which the DJ can do their job and not feel intrusive or abnormal to the party. The gestures are performed all on the touch mouse, which sends the signal to the server. The served infers the gestures using a decision tree along with a hidden Markov model (HMM). Windows of sensor values are recollected. Each of these windows holds groupings of sensor values at different time frames. Each of these groupings is passed through a decision tree that returns a certain state. (this state represents a particular hand movement) Several states (or hand movements) are then fed into the HMM, that determines what gestures that series of hand movements represent. Our interface also considers that a DJ is similar to a musician that plays different instruments. Musicians want to have the feel that they are interacting with different devices when making music. We enable this, by recognizing different types of gloves the DJ wears when interacting with the wireless controller. The different cloths invoke different sound effects. The fact the user has to wear different clothes to interact with the device makes the user feel as if they are interacting with different devices, turning music mixing into a ritual. The detection of different cloths is done by measuring the intensities of the mouse sensors. Different thresholds for different types of clothes were manually found. These thresholds did not overlap each other, allowing for a fine grained division of cloth types. The clothes detected were: leather, silk, and wool. It is important to note, that our system allows the DJ to enable and disable the use of the mouse, so the DJ can in fact dance freely sometimes!

As mentioned in the introduction, two different interfaces were designed and built: an open interface, providing transparency to the DJ's actions and a "closed" interface whose control was hidden from the audience, by the display of an oceanic scene. The open interface permits the DJ to perform the typical music mixing tasks: such as play two sound tracks simultaneously, cross fade between the tracks, open and select songs to play, control the tempo and rate of each track independently, seek in a track to a specific cue point, and add special effects (the special effects are related with the cloth recognition mentioned previously). The open interface is presented in Figure 1. The open interface can also actively motivate the crowd to participate in the music selection process. The interface is designed so that the DJ can stroll over songs and the crowd can applaud or boo the song, showing their support or displeasure (the interface presents a dialog requesting feedback from the audience for this music selection, although it can be enabled or disable according to the DJ's criteria).

Figure1. Open Interface of the mixing interface, showing how the songs selected are decomposed into their wave forms, the tempo, rate, volume controllers, special effects controllers and other controllers relevant to music mixing. The monitor buttons, when selected requests the crowd to participate in the music selection.

The "closed" interface also allows the DJ to control the same tasks as in the open interface, but it differs in the provided visualization and feedback. The new mapping is as follows: all of the wave forms are translated to "blowfishes" where the amplitude of the original signal is now translated into the length of the "spines" of the blowfish. A series of blowfishes are grouped together to emulate the entire digital signal. The DJ can interact with this new representation the same way he or she would in the open interface. The groups of blowfish are placed in the same place as the waves in the open interface were placed. The tempo and rate are mapped to water bubbles, whose size the DJ can increase or decrease, manipulating the tempo and rate accordingly. Both were placed also in the same spot the closed interface had them, with an intent of reducing the DJ's cognitive workload when using our application. The volume controller was also placed in the same locality, but it now had the form of a sand castle, whose size could be modified, resulting in a change in the volume. A bigger castle size results in an increment in the volume. The DJ controls the different music mixing elements by making the sea creatures bigger or smaller. A bigger size increases the value (e.g. increases the rate value), while a smaller size decrease the value (e.g., decreases the rate value). The mapping decisions taken were based on van Gulik et.al. proposals in [8].

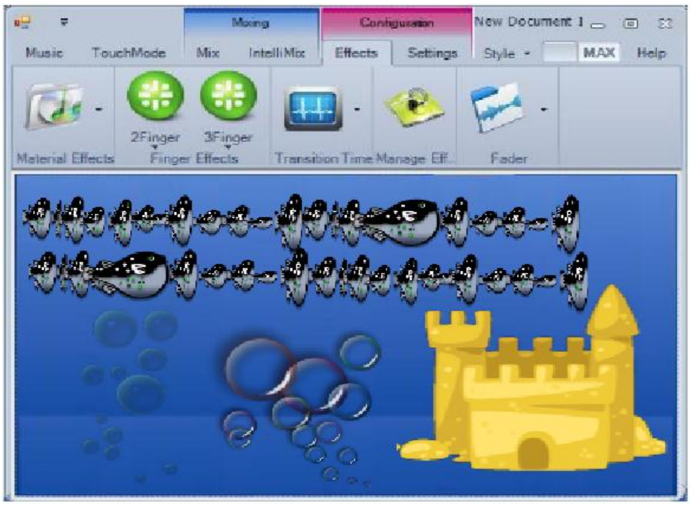

Figure2 Closed interface in which the wave forms are mapped to blowfish, the tempo and rate controllers are mapped to bubbles, and the volume is mapped to sand castle. By increasing or decreasing the size of these figures, the DJ controls the interface.

Our intelligent environment also infers social interactions. The social interactions are measured by placing sonar sensors on the stools of the party. Sonar sensors allow the detection of motion or whether a human has moved outside of the sensor's range. Therefore the sonar sensors provide information about whether the people at a table are sitting or on the dance floor engaging with the music. Next to the mixing interface in both the open and closed interface, we show to the DJ a map of the chairs in the party, and show for different areas of the map pie graphs, indicating how many guests are sitting in that particular area and how many are not. For testing the viability of our approach, all stools with their attached sonar sensors were hooked to the floor. The sonar sensors were all cabled to the floor, the floor was wired, and the cables were lead to the server that controlled the music interface. This server then utilized this information to provide informative feedback to the DJ.

III. USABILITY INSPECTION OF OUR INTELLIGENT ENVIRONMENT

In this section, we inspect the usability of our intelligent environment using the cognitive walkthrough methodology. The usability inspection was done inside a college laboratory. The laboratory was prepared to represent our designed intelligent environment: the seat sensor wiring was done and the chairs arranged, the projectors and speakers were also setup, and the participants for the study were recruited.

A. Users

Sixteen individuals were selected for our study. Four of these individuals were selected to play the role of the DJ (two of these four selected had been DJs previously for events, and where familiar with music mixing interfaces), the twelve others, were selected to play the role of people in the crowd. Three different individuals of the twelve were selected to be in one DJ trial.

B. Tasks

Certain tasks were selected for the DJs, and certain tasks were selected for the people in the crowd. Each DJ was asked to select five times two songs and mix them with the open interface and another time with the closed interface. The DJ was also asked to add effects using the gloves made of different materials. The DJ was asked, while on the open interface, to request feedback from the crowd for the music selection. After their performance the DJ was interviewed about his/her experience in the intelligent environment and was asked to describe the people in their crowd. The people playing the crowd on the other hand, were requested to dance and sporadically sit on the available chairs in the room. All of the participants were interviewed afterwards and questioned about their experience being in the intelligent environment.

Figure3 This figure presents the sonar sensors that are placed on the chairs of the event detecting if a person is sitting. The sonar sensors are placed in the back upper portion of the seats and are connected to the ground. There is cabling in the ground leading to the server that makes sense of this data and then displays it to the DJ.

C. *Results*

A few of the interesting points encountered in the cognitive walkthrough are listed: all four DJs expressed they enjoyed the ability of moving freely in the environment and controlling the music. The experts DJs, perhaps due to their previous experience, were the ones that emphasized their enjoyment of this property the most. Additionally the expert DJ users expressed in the interview that they felt very uncomfortable using the closed interface, (the interface that mapped the controllers to sea figures) because they felt a loss of control.

One of the DJ experts mentioned that in the open interface he did not mind openly expressing how the music was being manipulated, and felt it was another outlet for letting the public admire his work, as they could view how fast he was switching songs and sliding the controllers. All novice DJ users expressed they felt much more comfortable with the closed interface. Both novice users felt this interface was more intuitive and “friendlier” to use. One of the users expressed he felt overwhelmed in the open interface and felt embarrassed when the audience could see he was having trouble controlling the interface. We also questioned the DJs about the map layout and pie graph showing the percentage of people sitting.

All four DJs found that tool useful and helped them identify where they should perhaps move to, to animate the party. In the task, where the DJs had to request participation from the audience, one of the DJs expressed they would have liked to have multiple mouse controllers in the audience, enabling multiple collaborations. Two of the other DJs expressed they enjoyed this modality, as it allowed them to engage with their audience more. One of the DJs expressed she felt stressed in this task, because it required what she considered unwanted solicitation. Four of the participants expressed they liked the gestures and found them easy to remember. All participants were able to describe with detail their public. The eight of the individuals that acted as the guests expressed they enjoyed the projecting of the music mixing on the wall. Two individuals expressed they preferred the open interface over the closed interface, because it appeared “more modern”. Seven expressed a preference for the closed interface, because of the ambient it created. The rest did not express preference for one over another. A few expressed it made the “part environment” more interesting. All participants expressed they enjoyed being able to interact with the DJ.

IV. CONCLUSIONS

In this work we presented a novel intelligent environment that provides an outlet for some of the problems DJs encounter while performing. Our cognitive walkthrough appears to show, that this environment creates a better experience both for the DJ as for the public, but further user studies need to be performed. It would be especially interesting to analyze if in fact, novice users significantly preferred working with an interface where the controllers were mapped to whimsical figures than with a traditional music mixing interface. It might also be interesting to study if the choice of figures has an impact on the DJ's efficiency.

Also more social interactions need to be automatically detected, since there are cases where the public could be engaged and enjoying the music the DJ is playing but they are sitting. Detecting more social interactions would also allow the intelligent environments to become more responsive to the social context.

In this work we presented a novel intelligent environment that provides an outlet for some of the problems DJs encounter while performing. Our cognitive walkthrough appears to show, that this environment creates a better experience both for the DJ as for the public, but further user studies need to be performed. It would be especially interesting to analyze if in fact, novice users significantly preferred working with an interface where the controllers were mapped to whimsical figures than with a traditional music mixing interface. It might also be interesting to study if the choice of figures has an impact on the DJ's efficiency. Also more social interactions need to be automatically detected, since there are cases where the public could be engaged and enjoying the music the DJ is playing but they are sitting. Detecting more social interactions would also allow the intelligent environments to become more responsive to the social context.

REFERENCES

[1] Norma Saiph Savage et.al. Mmmmm:A multi-modal mobile music mixer. In Proceedings of the 2010 conference on New Interfaces for Musical Expression, NIME '10, pages 395-398, 2010

[2] Ali Mazalek and Tristan Jehan. 2000. Interacting with music in a social setting. In CHI '00 extended abstracts on Human factors in computing systems

[3] David Sprague, et.al. . 2008. Music selection using the PartyVote democratic jukebox. In Proceedings of the working conference on Advanced visual interfaces (AVI '08). ACM, New York, NY, USA, 433-436.

[4] *White J. Tim, “What Does it Take to Become a DJ?”* http://ezinearticles.com. http://ezinearticles.com/?What-Does-it-Take-to-Become-a-DJ?&id=1240816 *(2 April. 2012)*

[5] *"*What is a disc-jokey?”,
http://www.wisegeek.com.
http://www.wisegeek.com/what-is-a-disc-jockey.htm (2 April.2012)

[6] I. Poggi and F. DÉrrico, "Social signals: A psychological perspective.". Springer Verlag’s Advances in Pattern Recognition series, 2011, pp. 185-225.

[7] Pentland, A.: Social signal processing. IEEE Signal Process. Mag. 24(4), 108¨C111 (2007)

[8] Van Gulik, R., Vignoli, F., van de Wetering, H.: Mapping music in the palm of your hand, explore and discover your collection. In: ISMIR 2004 5th International Conference on Music Information Retrieval (2004) Available online