Another Thrill in the Wall: an Affective Eco-System Interface for Gestural Expressivity

Miguel Jerónimo
INESC-ID Lisboa and Instituto Superior Técnico,
Technical University of Lisbon,
Av. Prof. Dr. Anibal Cavaco Silva,
2744-016 Porto Salvo, Portugal
miguel.jeronimo@ist.utl.pt

Carlos Martinho
INESC-ID Lisboa and Instituto Superior Técnico,
Technical University of Lisbon,
Av. Prof. Dr. Anibal Cavaco Silva,
2744-016 Porto Salvo, Portugal
carlos.martinho@ist.utl.pt

Ana Paiva
INESC-ID Lisboa and Instituto Superior Técnico,
Technical University of Lisbon,
Av. Prof. Dr. Anibal Cavaco Silva,
2744-016 Porto Salvo, Portugal
ana.paiva@inesc-id.pt

ACE 2011, November 8–11, 2011, Lisbon, Portugal.

Abstract
The present work explores the paradigm of expressing emotions through the body on a physical medium. Our goal is to develop the AffectiveWall, an intermedia system that translates affective states, manifested by whole body interaction with a multi-touch surface, to music and digital painting generated in real-time.

Keywords
Affective body expression, intermedia instrument, multi-touch interface, gestures recognition

ACM Classification Keywords
H.1.2 [User/Machine Systems]: Human information processing. H.5.2 [User Interfaces]: Input devices and strategies. J.5 [Arts and Humanities]: Performing arts

General Terms
Human Factors, Experimentation, Performance

Introduction
According to Collingwood, Art is the expression of emotions in a given language [5]. And if body language can be a universal way to communicate, a system that uses affective expression through gestures as its interface emerges as a way of bringing Art closer to people, facilitating the process of communicating

emotions to an audience. In addition, if one of the most effective tools for creativity is multidisciplinarity, then using emotions to connect two different types of media can be a drive for a fully connected and intuitive intermedia instrument for live performances. As such, we chose "Another Thrill in the Wall" as the *leitmotif* for the present work, in which the word "thrill" refers to the physical manifestation of an emotion through the whole body. In this paper we present our system by focusing on the interaction with users, including the recognition of affective expressivity associated.

AffectiveWall

We are developing a system in which the user interacts by "painting on a canvas", i.e. touching a vertical multi-touch surface, using its whole body. These movements are then captured and translated into emotions, based on the "affective ecosystem interface" concept. Through these emotions, the system composes music and generates abstract paintings in real-time that are projected on the canvas, always according to the affective features it reads. These contents consist on performer's "augmented affective states", related to the "extended expressive gestures" described in [3], which are perceived by spectators as multimodal stimuli composed by physical movements, audio and video contents. The name of this project is AffectiveWall and, in the end, our aim is developing a musical and visual instrument (and consequently a tool for performance) to be played in an organic and intuitive way (with the lowest possible learning curve), enabling its use not only in Arts, but also for educational and therapeutic purposes. For this, we built a LLP multitouch surface, as described on Nui Group website[1], in order to maintain the lowest costs as possible, preserving the mindset of bring artistic performance closer to people.

[1] Info on: http://nuicode.com/attachments/download/115/Multi-Touch_Technologies_v1.01.pdf (accessed on 13-09-2011)

Affective ecosystem as interface

In this concept of "ecosystem interface", system has the main role on the interaction, responding to a complex environment (which can be altered by users, but also by space conditions, noise, etc.) while users can indirectly interact with the system by actions on the environment. One example is AESI (Audible Eco-Systemic Interface) by Di Scipio [6], which is set as an automatic feedback cycle: sound from the environment is captured by microphones, passing on to controller, processing and synthesis steps, producing sound that comes back to the environment through speakers. So, in this case, sound is the interface and users can only interact with the system by acting on the environment. Therefore, AffectiveWall relates to this concept by having affective states as the interface, where the system has the role of interpreting the whole "affective ecosystem" and the user has the chance to change it (and therefore indirectly change the system output). This interaction is also related to gestural control of music, interactive music systems and the concept of digital or virtual instrument. Thus, it is mandatory to consider some characteristics about this topic [17]: immediate response to user's movements, the non-limitation of interaction options (e.g. possible choices on a menu) leaving the interaction to a continuous sequence of controls, possibility of any audible sound as an output and, at last, the separation between the interface and the sound synthesis (both modules are independent and have a mapping of parameters between them). Regarding to this last point, Knapp and Cook [13] describe the potencial of this abstraction, where this separation between musician and the object

figure 1. User performing affective states on the wall.

responsible for the sound generation increases the creation of an emotional interface instead of a physical interface (like guitar strings or piano keys) – one of the main concepts of AffectiveWall's creative process. This way, the user doesn't need to worry about using the instrument to produce low-level contents, like chords, but to convey a specific affective state, a high-level content. Moreover, it is necessary to connect this to the metaphor of painting. Its relevance can not be overlooked because is the main help available to users, simplifying and providing them directives on how to use the system [7, 8], avoiding the use of manuals or other external supports. According to Buxton, the most natural interaction language consists on non-verbal dialogues, enhancing gestures as phrases with their own meaning [1]. Therefore, the interface should be replaced by actions that naturaly derive from the supported metaphor, rejecting the use of menus, buttons and windows (in the case of AffectiveWall there is only canvas, just like in painting). In terms of interface hardware, to maintain this metaphor is mandatory to preserve the traditional way of painting, i.e. the relationship between the artist and a physical and vertical medium. After talking about the users' interaction, it is time to explore the meanings of their expression: the emotions within the gestures.

Affective model for gestural expressivity

In the process of studying the link between emotions and gestures while painting, we found that is still a need for research in this area, mainly when considering the new opportunities that matter could provide on the emerging scene of multi-touch interaction. Some experiments have been made by Hiraga et al. [9] on the connection between emotions and drawings but mainly testing the recognition of emotions, not the production and interpretation themselves. Regarding gestures, Kipp and Martin [12] studied the expression of hands, while Höök et al. [10], Camurri et al. [3], Castellano et al. [4] and Wallbott and Scherer [16] researched full body movements. Most researchers, however, do not consider the constraints of a medium on the gesture expression, such as when the subject is interacting with a surface, as in the case in our work. One of the researched systems is more closely related to our approach: EyesWeb by Camurri et al. [3], a framework for analysis of dance performances, producing audio-visual output related with the emotions conveyed by performer. However, in this system the gestures are not performed on a physical interface.

In terms of expression analysis, Izard claimed that are emotions with patterns that convey particular meaning or information [11]. In Laban Movement Analysis [18], human movement is studied and decomposed in body, space, effort, shape and relationship, defining a language for interpreting, describing, visualizing and notating all kinds of movements. In the same way, Wallbott and Scherer [16] defined six dimensions, namely overall activation (quantity of movements), spatial extent (amount of space occupied), temporal extent (duration of movements), fluidity (smoothness of the movements), power (dynamics of the movement) and repetition. Camurri et al. [2] reported how they measured the emotional expression of drawings made by users with a laser pointed to a wall (when listening to musical excerpts), and identify a collection of relevant features: angularity, rarefaction, spatial occupation, vertical symmetry, horizontal symmetry, central symmetry, compactness, lateral location, vertical location, angular tendency, and spatial extension. Although similar to our work in terms of

classification, the expression here is performed using a single point in the wall, rather than the whole body.

Therefore, we adapted the model proposed in [16] to the performance of gestures on a surface. For instance, spatial extent on their model is represented by occupation of the canvas in our case. Also we use some descriptors proposed in [2], namely vertical location, compactness and spatial occupation. In addition, we considered the most frequent and distinguishable features we perceived in our evaluation, joining all together in the following metrics and measures: gesture length (punctual, short, medium, or long), area of touch (one fingertip, one fingertip of each hand, all fingertips, all fingertips of both hands, one hand, both hands, one arm, both arms, hands/arms and head, or whole body), quantity of gestures (one, some (two or three), or many), gesture speed (static, slow, medium, or fast), direction of the movement (none, downward, upward, inward, outward, sidewards, or random), shapes drawn (blob, straight line, curve, circle, or chaos), duration (sudden or sustained), location (low, medium or high height), and occupation of the canvas (confined, medium or expansive). Note that the purpose is to use the painting metaphor, thus gestures are only considered when touching the surface itself.

Relatively to the emotion range considered, we based ourselves on models from psychology, such as Russel's Circumplex Model [15] and OCC Model [14], and added aesthetic and artistic expression concerns, creating a group of, not only emotions, but more general affective states that are relevant when the matter is the expression on a canvas. For this work, the chosen affective states are: sadness, shame, anger, confusion, joy, freedom, melancholy, pride, pleasure, exaltation, tenderness, shyness, satisfaction, loneliness, hate, fear, relief, hope and disappointment. This selection was made with São Nunes[2], an action painting and performer expert, who works on painting in real-time accompanied by live music, exploring the emotional expression of her body and developing awareness to the problematic of interaction with a canvas.

Gestures evaluation and pattern recognition

Guided towards an affective model suitable to reality, we proceed with an experiment to find patterns of gestures related to each affective state, using the evaluation metrics presented on the last section. Is important to note that we are not aiming to find the perfect expression that will work for everyone (as we believe that would be an impossible task). We cannot forget that, as any other instrument, performers will have to discover their personal way of playing it (even in this case where the adaptation will be a lot lesser and looser than conventional instruments). In this way, we asked fifteen individuals to perform each affective state on a wall while we were recording video. Afterwards, we measured the gesture features from the video, aiming to find patterns in the subjects' body language. We asked the subject to perform one expression for one affective state at a time, first using the most spontaneous gestures that came to their mind and then, after a pause to think about the best expression, using the subsequent reflective gestures. To remove any bias, affective states were presented in a random and different order for each subject, and tests were performed isolated without the user having previously seen other expressions. After each gesture, the subject was asked about her expression difficulty

[2] Info on: http://www.saonunes.com/ (accessed on 14-08-2011)

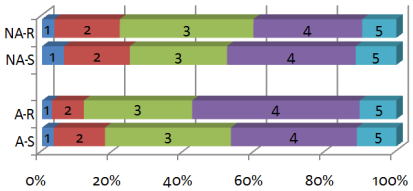

figure 2. Satisfaction/confidence ratings by background ("NA" stands for non-artistic and "A" for artistic) and type of gesture ("R" stands for reflective and "S" for spontaneous), allocated by a 1 to 5 scale (being 1 the lower confidence and 5 the higher). In this test, we only considered subjects with artistic background those who have studies in acting, performance or dance, due to their potentially developed skills in expressing their emotions through the body (and awareness of expressive features involved).

and her satisfaction/confidence, regarding the accuracy with which gesture represents the requested affective state. Due to the long process of this experiment (which includes several screenings of each video to make the feature analysis) and the duration of a single test (where we ask the user to perform nineteen different affective states twice), we decided that fifteen individuals would be a reasonable starting number to understand which affective states are likely to be "decomposed" into recognizable patterns and define how. To enrich and diversify the test group, the choice of individuals was made to cover a large range of ages (from 18 to 66) and backgrounds: with and without artistic background (and among artists, people from music, painting and dance), psychology, etc.

The first conclusions that we took from the experiment, was the similarity on the ease/difficulty of expression between users with or without an artistic background (values on figure 2), and the fact that the majority of the group, independently on the background, revealed acceptable confidence rates in the expressed affective states. All of these suggest that it may be possible to create a universal instrument that everybody could play regardless of the existence of artistic background. Regarding pattern recognition for proceeding with the model implementation, we focused for now on four complementary affective states in terms of valence and arousal. Thus we found the most common movements for each state addressed, and came up with the results shown above. Is important to note that these are the reflective expression results, with the number of subjects in the test group who performed each feature represented in brackets. Also, fractions may appear when the subject uses more than one type of feature and, in this case, the value allocated to one metric is distributed among all measures used. The results are:

- Sadness - One$^{(13)}$ slow$^{(11)}$ and sustained$^{(15)}$ gesture drawing a downward$^{(9.33)}$ line$^{(10.83)}$ at medium height$^{(8.66)}$ and with medium occupation$^{(8)}$;
- Anger - Punctual$^{(8)}$, static$^{(8)}$ and sustained$^{(9.5)}$ gestures (like punches) made with both hands$^{(8)}$, making blobs$^{(8.5)}$ at medium height$^{(11)}$ (fast movements as if they are ripping the canvas are also used$^{(6)}$);
- Joy - Many$^{(9)}$, long$^{(7.5)}$ and sustained$^{(13)}$ gestures at medium or high speed$^{(11)}$, performed at medium or high height$^{(14.32)}$ and with expansive occupation$^{(11)}$;
- Tenderness - One$^{(11)}$ long$^{(8)}$, sustained$^{(15)}$ and slow$^{(10)}$ gesture, drawing a line$^{(8)}$ at medium height$^{(11.83)}$ with medium occupation$^{(8)}$.

Due to being an issue with a large degree of subjectivity, and also because we are not aiming to find absolute expressions, our approach was to considered gestures that were performed by at least half of the group. In such manner, the model is appropriate to reality but also providing enought freedom and subjectivity to feed this kind of interaction interpreted by the system.

Conclusions and future work

This experiment and its results support some of our initial hypotheses regarding the AffectiveWall system, an interface that reads the affective expression of a performer through body language on a physical medium. Our next step will be comparing the recognition rate of the whole body expressivity with and without the "augmented affective states" composed by audio-visual outputs. Other future work is to incorporate new input methods and features, like

sensitivity/power of the touch (e.g. using piezos to capture the sound made when hitting the surface), calculation of area/speed changes along a gesture, and a more complete set of affective states. Nevertheless, we are following J. Han[3] words: "Multi-touch-sensing was designed to allow non-techies to do masterful things while allowing power users to be even more virtuosistic". With AffectiveWall, we aim to apply this to Arts, in order to simplify them to one of the most universal languages known to everyone: emotions. Because, all in all, it's just another thrill in the wall.

References

[1] Buxton, B. The "natural" language of interaction: a perspective on non-verbal dialogues. In INFOR Vol. 27 No. 2 (1989), 221-229.

[2] Camurri, A., Castellano, G., Ricchetti, M., Volpe, G. Subject interfaces: measuring bodily activation during an emotional experience of music. In Proc. Gesture in HCI and Simulation, Springer (2006), 268-279.

[3] Camurri, A., Mazzarino, B., Ricchetti, M., Timmers, R., Volpe, G. Multimodal analysis of expressive gesture in music and dance performances. In Proc. Gesture-Based Comm. in HCI, Springer (2004), 357-358.

[4] Castellano, G., Villalba, S., Camurri, A. Recognising human emotions from body movement and gesture dynamics. In J. ACII, Springer (2007), 71-82.

[5] Collingwood, R.: The principles of art. Oxford University Press (1938).

[6] Di Scipio, A. Sound is the interface: Sketches of a Constructivistic Ecosystemic View of Interactive Signal Processing. In Proc. of the Colloquium on Musical Informatics (2003).

[3] Quoted by King, J. in "The 2008 TIME 100", http://www.time.com/time/specials/2007/article/0,28804,1733748_1733754_1735325,00.html

[7] Fishkin, K. A taxonomy for and analysis of tangible interfaces. In Proc. of the Personal Ubiquitous Computing (2004).

[8] Dourish, P. Where the action is, the foundations of embodied interaction. MIT Press (2001).

[9] Hiraga, R., Kato, N., Yamasaki, T. Understanding emotion through drawings: comparison between hearing-impaired people and people with normal hearing abilities. In Proc. Inter. Conf. on Systems, Man, and Cybernetics, IEEE Press (2006), 103-108.

[10] Höök, K., Fagerberg, P., Ståhl, A. Designing Gestures for Affective Input: An Analysis of Shape, Effort and Valence. In Proc. Mobile Ubiquitous and Multimedia, ACM Press (2003).

[11] Izard, C. The Face of Emotion. Appleton-Century-Crofts, New York (1969).

[12] Kipp, M., Martin, J.C. Gesture and Emotion: Can basic gestural form features discriminate emotions? In Proc. Inter. Conf. on Affective Computing and Intelligent Interactions, IEEE Press (2009), 1-8.

[13] Knapp, R., Cook, P. The Integral Music Controller: Introducing a Direct Emotional Interface to Gestural Control of Sound Synthesis. In Proc. of Inter. Computer Music Conference (2005).

[14] Ortony, A., Clore, G.L., Collins, A. The cognitive structure of emotions. Cambridge U. Press (1990).

[15] Russell, J. A. A Circumplex Model of Affect. In J. of Personality and Social Psychology Vol. 39 (1980).

[16] Wallbott, H.G., Scherer, K.R. Cues and channels in emotion recognition. In J. of personality and social psychology Vol. 51 No. 4. APA (1986), 690-699.

[17] Wanderley, M. Gestural control of music. In Inter. Workshop Human Supervision and Control in Engineering and Music (2001).

[18] Zhao, L., Badler, N.I. Synthesis and acquisition of Laban Movement Analysis qualitative parameters for communicative gestures (2001).