INSTITUTO SUPERIOR TÉCNICO
Universidade Técnica de Lisboa

AffectiveWall

An intermedia instrument for affective generation of music and paintings through gestural expressivity

Miguel Lopes Jerónimo

Dissertação para obtenção do Grau de Mestre em

Engenharia Informática e de Computadores

Júri

Presidente: Mário Rui Fonseca dos Santos Gomes

Orientador: Carlos Antonio Roque Martinho

Co-orientador: Ana Maria Severino de Almeida e Paiva

Vogal: Ricardo Jorge Jota Costa

Outubro 2011

Acknowledgements

Let me start by thanking my advisers Carlos Martinho and Ana Paiva, and also artist São Nunes for all their help along the way.

I would also like to thank all the people involved in some way with this work, namely and in a random order, Virgínia, Cláudia, Miguel, Bárbara, Joaninha, Guaka, João's, Lucila, Rui, Rute's, Tânia, Jesus, Adriana, Fox, Félix, Iolanda, Lara, Pedro, Jaime, David, Bernardo, Francisco, Tiago's, Carlos, Nuno's, André, Ana Isabel, Isa and a lot more wonderful people that keep me going every day. Not forgetting the coffee, for the good job of awakening me every night, and Bairro Alto, for maintaining my sanity in a relatively normal level.

Looking at the big picture (and surely an abstract one), for the past months as much as the past 24 years, I deeply thank my family(specially Mother, Father and David, my brother) for all the support, patience and incredible role models that make me always try to do better.

Emerging from the bubble of noise of all the rest,
They were the strong chord that supported me in completing this quest.

Abstract

English

In this thesis we present AffectiveWall, an intermedia instrument that connects music and painting through emotions/affective states expressed by a performer's body-language. This gestural expression happens on a wall/canvas (multi-touch interface) that analyzes features of gestures performed on its surface, feeding an "affective ecosystem" that decides the affective states expressed and, according to them, composes music and abstract paintings projected on the wall in real-time. Therefore, artistic creation is based on universal languages - corporeal and emotional - enabling the free expression of people regardless of their background, through an instrument easy to use, in order to bring Art closer to people. We reviewed the state-of-the-art of artistic trends and technologies related to AffectiveWall, and the classification of affective states, after all, the main interface between music and painting, and between users and system. We defined an aesthetic/artistic approach to an "affective model for gestural expressivity" by evaluating performers while they express themselves on a wall, in order to find gestural patterns. We present an architecture and implementation, evaluating the recognition by spectators of expressions using the system. We found that our "augmented affective states" as output (a multimodal stimulus composed by sounds, visuals and gestures) are universal and can help an audience to better perceive the affective expression if the person performs a focused/intimate interaction with AffectiveWall. We also found that, in this kind of interaction for affective expression, joy is the easiest affective state to recognize, and arousal the most perceived feature of interaction.

Keywords: affective interface, gestural expressivity, augmented affective states, music and painting.

Portuguese

Nesta tese apresentamos a AffectiveWall, um instrumento intermedia que interliga música e pintura através de emoções/estados afectivos expressos pela linguagem corporal dum performer. Esta expressão gestual dá-se numa parede/tela (uma interface multi-toque) que analisa os gestos na superfície e alimenta um "ecosistema afectivo" que decide quais os estados afectivos a ser expressados e, de acordo com eles, compõe música e pintura abstracta que é projectada na superfície em tempo real. Assim, a criação artística é baseada em linguagens universais – corporal e emocional – permitindo uma expressão livre independentemente da formação (artística ou não), num instrumento com curva de aprendizagem mínima de forma a aproximar Arte e pessoas. Revimos o estado-da-arte das correntes artísticas e tecnologias associadas à AffectiveWall, e a classificação de estados afectivos, no fundo, a interface que interliga música e pintura, e utilizadores e sistema. Definimos uma abordagem estética/artística de "modelo afectivo para expressividade gestual" através da avaliação de performers a exprimir-se numa parede para encontrarmos padrões de gestos. Desta forma, apresentámos uma arquitectura e uma implementação, testando o reconhecimento por espectadores das expressões efectuadas no sistema. Descobrimos que os "estados afectivos aumentados" como saída do sistema (um estímulo multi-modal composto por sons, visuais e gestos) são universais e podem ajudar uma audiência a perceber melhor a expressão se o performer estiver concentrado numa interacção íntima com a AffectiveWall. Concluímos também que, para este tipo de interacção para expressão afectiva, a alegria é o estado afectivo mais reconhecido e a excitação/energia do gesto a característica mais percepcionada.

Palavras-chave: interface afectiva, expressividade gestual, estados afectivos aumentados, música e pintura.

Table of Contents

List of Figures

List of Tables

"Truly fertile Music, the only kind that will move us (...)
will be a Music conducive to Dream (...)

One must not wish first to understand and then to feel.

Art does not tolerate Reason."

Albert Camus

1 Introduction

Many people question the role of Art in society. But if there is something that no one has any doubt about, it is the ability of Art to generate and promote emotions. It can be an image that accompanies us through life or a music that elevates us to a cathartic state, Art does have a primal role in our existence. And if no one is a one-dimensional person, and if every human is rich in thoughts and emotions that need to be expressed and explored, why should Art be confined to only a chosen few? Thus arises the need to bring Art, music and painting in particular, to a place where theories and musical notations are not mandatory, and artistic trends or preconceived ideas have no value. Thus emerges the concept of an instrument in which the creative process is based on body language and the rule of composition is the emotional expression – an almost universal language to everyone. And from this goal of universality, emerges another main purpose of this work: multidisciplinarity. If the conductor Leopold Stokowski said that "a painter paints his pictures on canvas, but musicians paint their pictures on silence", this dissertation aspires to merge these two fields, creating a way of composing music and digital painting through the act of interacting with a canvas. So, AffectiveWall – the final goal of this work - will be a system that integrates the following four elements:

- Affective states[1] (mood, emotions): why we want to communicate, i.e., the need of emotional expression - after all, the content that feeds the system and connects all its parts;
- Music/sound and digital painting: what will be created - the final result of the instrument after the analysis of its user's expressivity;
- Gestures/body-language: how it will be created, that is, the user interacts touching with the whole body on a surface, always having in mind the notion of experimentation, improvisation and freedom;
- Wall/canvas: a surface where the creative process and the visual/sound production takes place.

[1] Affective states are a more general definition that stands for the observable behaviors that represent emotional expressions (this issue is discussed in section 2.2.2).

This name – AffectiveWall – besides including the affective component of artistic creation on the system (ultimately an ubiquitous presence in whole thesis), has the presence of the “wall”. Not only due to the widespread use of this word when naming this kind of tangible and interactive surfaces, but also, and more importantly, due to the subversion of the concept of Art as an activity for elites. The painting will no longer live in an exhibition context as a canvas in a gallery. Instead, it will lie down on a simple “wall”, a paradigm related to Street Art where artistic contents stay vivid in the natural habitat of people, an urban context where everyday life happens.

In practical terms, the performer’s gestures are read by a video camera placed behind the surface (which acts as a canvas), and visual feedback is given through a projection of shapes and colors on the spots where the surface was touched. Data captured by the camera is processed by software, analyzing the expressed emotions, and the shapes and colors projected are produced according to these characteristics. Furthermore, also according to the emotions examined, the system synthesizes and plays music in real-time. Therefore, emotions, or more generally affective states, are in fact what puts together the whole system and can be considered as the interface between music and painting, and also between user and system.

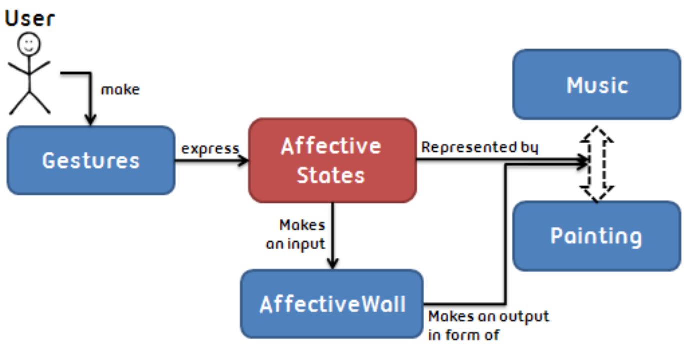


Fig. 1. Affective states as the interface between music and painting, and also between user and system.

1.1 Goals and motivations

As motivations for this work, it is important to mention some aspects related with the type of instrument that is presented here, which stands as the main goals that AffectiveWall undertakes to achieve, such as: ease of use, tangible and organic character of sound creation, multidisciplinary approach to arts where one can produce contents from an artistic field through processes of another field, the non-limitation of artistic creation opening it to a freer and more liberating creative process, a process supported by a more transversal artistic language, including, for example, aesthetical paradigms from performing or dance acts.

It is also relevant to note that it is important to create an instrument that somehow offers a dichotomy, due to the fact that it should be interesting for the two targets it wants to reach: performers and artists, so an aesthetically value and multiple potentials of creation are mandatory, in order not to limit their creativity; and amateurs and curious people, so it’s required to be easy and fun to use.

Therefore, it is possible to name two main motivations in the present thesis, the first being the easiness of use, i.e., the absence of manuals or any need of introductory classes (note that the interaction is based on the contact with the body itself, without the use of mouse, keyboard or other devices). For this purpose, AffectiveWall uses the visual and sound feedback for a better understanding of the musical creation and instrument control (therefore reducing the frustration of the typical first contact with, for example, a musical instrument, which in most cases leads to meaningless sounds or no sound at all).

The second motivation, and the one that stands for the main theme developed in our work, is that affective states are the link between the two artistic fields - music and painting – and acts as the interface to the user. We want to explore this and understand whether there are some affective states that are easier to recognize than others, and the influence that the person who is expressing and the person who is watching/listening have in the recognition of affective states. We want to study how different people from different backgrounds (artistic, non-artistic, etc.) establish a relationship with our system, in a attempt to define what we need to create a universal instrument.

1.2 Main contributions and publications

In the process of developing the present dissertation, we explored mainly the paradigm of connecting affective states with gestures, being this one of the main contributions of our work. Therefore, a paper was submitted and accepted to the workshop "Whole Body Interaction" of the 8th International Conference on Advances in Computer Entertainment Technology (ACE 2011) organized in cooperation with ACM and ACM SIGCHI. We named the paper "Another Thrill in the Wall: an Affective Ecosystem Interface for Gestural Expressivity", in which the word "thrill" refers to the physical manifestation of an emotion through the whole body. In this paper we described our system and our studies to reach an affective model for gestural expressivity, presenting other two main concepts of our work: affective ecosystem as interface and augmented affective states as output.

Affective Ecosystem as Interface

In the concept of "ecosystem interface", the 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. 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 theme is described in section 3.1, including the review of Di Scipio's work [1], our main inspiration for this concept.

Affective Model for Gestural Expressivity

This is the main framework that supports AffectiveWall, a theory that maps affective characteristics of body language (when painting) in sounds, shapes and colors that represent those affective states. The reason for this research is, as already stated, enabling people to use universal languages like gestures and affective states to express themselves. This potentiates the communication between the performer (who can be anyone, not only artists) and the spectator, in a sincere and organic performance. To achieve this, we performed an experiment (on section 4.3) with various users from different backgrounds (both artistic and non-artistic) to understand what patterns of gestures people use when expressing various affective states.

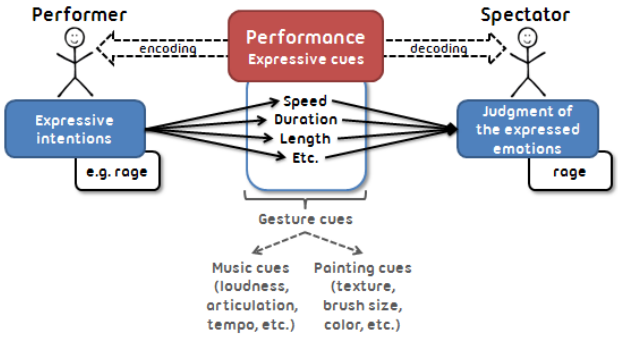


Fig. 2. Diagram of the communication between performer and spectator, supported by expressive cues related to the affective states expressed (inspired on [2]).

Augmented Affective States as output

The visual and sound contents produced by AffectiveWall, related to the affective states expressed by performers, consist of their "augmented affective states", (related to the "extended expressive gestures" described in [3]), in which they are perceived by spectators as multimodal stimuli composed by physical movements, audio and visuals (therefore the performer is always included in the system output). Our work aims to explore the aggregation of all these elements and their influence on connecting with an audience (for this last matter we conduct a user evaluation in chapter 5).

1.3 Application areas

Bearing in mind the way of interacting with AffectiveWall previously mentioned, one of the application areas which can be considered is the use in mixed-art contexts like installations and performances, or simply as an innovative tool for composing music in experimental and electronic music fields. Besides, there is the possibility of using it in a therapeutic context, due to the "tactile" contact with sound. This method can work as a way to explore creativity and communication skills of children and people with disabilities (as, for example, Art therapy - painting with the hands - and music therapy). Despite its potential, this issue falls outside the scope of this thesis and, therefore, is considered as a possible and interesting future work of the present thesis.

2 Background

This section is intended to specify the context of the present work, namely the artistic fields that influenced the conception of AffectiveWall (music, painting, intermedia, etc.), understanding the themes involved and the work already done on each one of them. Then, a survey about studies in emotions is presented and, above all, we will explore the connection between them and Art. In addition, we thought it would be interesting to introduce and develop some of the concepts related to the artistic process of AffectiveWall that will underpin this work (such as improvisation, creativity and synesthesia).

2.1 Artistic background

2.1.1 Music

One of the musical genres we can enumerate as an inspiration to AffectiveWall is Romanticism, which occurred between eighteenth and early twentieth century. This movement had Beethoven, Wagner, Brahms, Chopin and Tchaikovsky as leading figures, and had, as a major feature, the personal and emotional ingredients of composition, assuming that reality must be perceived mainly through emotions and feelings.

In the early twentieth century the so-called Modernist (and later the Post-modernist) movements appeared, comprising various styles, namely Impressionism, Expressionism, Dodecaphonism, Atonality, Minimalism, Musique Concrète and, more recently, electronic, computer music, ambient, glitch, etc. These new methods of composition emerged following the advance of technological, industrial and science progress, and therefore include some new creative approaches: field recordings, sound processing and computers. Thus, a profound revolution was initiated in the process of making music, giving great emphasis to creativity, innovation and sonic qualities of sound, with a gradual emancipation from conventional musical instruments [4,5].

To conclude, and without more developments about this issue which will be addressed later in the present document, the AffectiveWall proposes to reuse the emotional expression typical of Romanticism with the sound concerns of the avant-garde movements, in an experimental electronic music paradigm.

2.1.2 Painting

Although the act of "painting" on AffectiveWall does not involve any previous knowledge of fine arts, it is important to talk about some artistic trends that affect the composition method and the visual feedback of the system. Thus, the first movement that needs to be referenced is Art Brut (or Raw Art), created by Jean Dubuffet at the time of World War II rejecting every imposition or artistic style, considering as the purest Art the one that was produced with no artistic knowledge [6] - an interesting analogy with the act of composing on the AffectiveWall.

Continuing with the relationship between the artist and his/her painting, also Expressionism, born in Germany in the turn to the twentieth century, is relevant due to its emphasis on the artist's personal and impulsive view, instead of the actual representation of the world and its formal expression ("The painter shall not paint what he sees in front of him, but what he sees inside himself" [7]). So this artist's emotional attitude towards himself and the world around him, something analogous to the Romanticism in a musical context, is considered as a fundamental paradigm of AffectiveWall. An example of expressionist painter was Rothko, which claimed that he painted emotions. He was always worried about human expression and how to transpose those subjective and spiritual feelings to a canvas. Devoid of landscapes or human figures, his paintings only contained symbolic representations of emotions through color and blurred blocks, possessing the canvas with their own life force.

Regarding the aesthetics of the visual output, the Impressionism movement, from the late nineteenth century and led by Claude Monet, is a great inspiration due to its use of light and motion to give a strong dynamic feeling to the paintings (normally applied at that time to landscape paintings and mainly for water, clouds, etc.) [6]. Therefore, the visual feedback of AffectiveWall will simulate this dynamism of Impressionist works, focusing on the motions of the drawn shapes to represent the fluctuations of music produced.

In addition, Abstract Art (or Abstractionism) has an important role in AffectiveWall paintings, with its impulsive and intuitive kind of painting that does not represent anything from our physical reality, replacing them with compositions of colors and shapes that represent reality in an abstract and subjective way. Despite the above, the connection of Abstractionism with AffectiveWall is not fully applicable, since the act of abstracting always involves a simplification of reality, something that does not define truly the concept of AffectiveWall: the aim is to "paint" symbolic representations of affective states rather than representations of real objects or situations. So perhaps we can better define our kind of "painting" as Non-figurative Painting, a movement started in 1910 by Kandinsky who begins to use free, dynamic and, above all, completely unintentional strokes, to apply a purely spiritual meaning to shapes and colors. This also applies to Action Painting (a movement mainly associated with Jackson Pollock). This movement puts a great emphasis on the act of painting, where the artist employs all his/her energy to the work with the paints themselves as a plastic material, with their specific textures and viscosities. Therefore, resulting in extremely active and sensory rich paintings, claiming themselves as affective representations built through motions and gestures, not just with brushes, but especially with hands and whole body: hence their major influence on the creative process employed in AffectiveWall.

Fig. 3. Examples of artistic trends that influence AffectiveWall. From left to right: the Abstractionism of Kandinsky, the Expressionism of Rothko, and the Action Painting of Jackson Pollock.

2.1.3 Intermedia and the connection between music and visuals

The quest for the fusion of different artistic fields is not new: we only need to consider the cinema as a combination of literature (the argument), theatre (the acting), music (the soundtrack) and photography (the capture of image itself). Also in fine arts there has been some research with the intention of bringing together different Art forms, such as Bauhaus (which had the ambition of joining all arts together) or the development of Intermedia concept in the sixties [8,9]. This concept was employed in Fluxus movement, founded by George Maciunas and developed by artists from fields as distinct as John Cage (composer of contemporary music) and Yoko Ono (plastic artist). Even in the twenties there were developments in this field through the Lumia movement, created by Thomas Wilfred. This movement believed in manipulation of light as music [10] through shape, color and motion. Additionally, since the eighteenth century interest exists in instruments that could connect music and light [8,11], namely the Clavecin Oculaire, the Color Organ (nineteenth century) and the Laser Harp (twentieth century). Therefore, we can consider that the AffectiveWall fits in this artistic trend, where the production of artistic content is the composition of various different fields, like painting and music as the system output itself, but also the performing way of interacting with AffectiveWall (that we can resemble to the act of painting, dancing, etc.).

2.2 Emotions and affective states

2.2.1 Introduction

"Your intellect may be confused, but your emotions will never lie to you."
Roger Ebert

Emotions are perhaps the hardest human reaction to explain, something that has never discouraged the scientific community. Since at least 1884, when William James formalized a definition of the term, many other scientific definitions have emerged over time [12,13] (over 100 are reviewed in [14]), but still we cannot set a universal answer to this issue. Nevertheless, it is possible to summarize emotions as the synchronized and interrelated changes in the states of the five sub-systems of human body's functioning - cognitive, physiological, action tendencies, motor expression and subjective feeling [15,16,17,18,19] - in response to the evaluation of an external or internal stimulus [12,20,21].

Starting from this definition, it becomes necessary to outline a classification of emotions and affective states, in order to understand how they link with body-language (the AffectiveWall input) and different Art forms (the output produced).

2.2.2 Classification of emotions and affective states

Defining a consensus set of emotions is not an easy task. Tomkins suggests eight basic emotions: fear, anger, anguish, joy, disgust, surprise, interest and shame [22,23]. Plutchik considered fear, anger, sorrow, joy, disgust, surprise, acceptance and anticipation [22,24]. On the other hand, Johnson-Laird and Oatley [22,25] take a different approach, analyzing words for emotions and deducing a set of only five emotions - fear, anger, sadness, happiness and disgust – being other emotions a derivation or conjugation of more than one from this set. So, these theories represent one of the methods used by psychologists to measure human emotions – the discreet approach – that uses language to describe clearly separated states. Therefore it is easy to conclude that this method leads to incoherencies between the different studies and to serious difficulties when analyzing and interpreting the countless possible combinations of emotions [12,22].

Thereby a new method was introduced to represent emotions in a spatial way instead of verbal [26] – the dimensional approach – describing them by their position on a tridimensional space with the axes valence (positive-negative), arousal (calm-excited) and tension (tense-relaxed) [12,22]. This new way of classifying emotions, introduced by Wilhelm Wundt in *"Grundzüge der physiologischen Psychologie"* (1905), minimizes the cultural influences and other language-dependent factors, enabling also the possibility of measuring distances between emotions [26,27]. Despite that, due to the difficulty of consistently identifying the third dimension of the Wundt theory, many theorists focused on only valence and arousal, approximating to circle spaces [28] like Russell's Circumplex Model. This model is composed by a large number of emotions with equal distances between them [26,29], allowing the organization of their characteristics through angles: pleasure (0º), excitement (45º), arousal (90º), distress (135º), displeasure (180º), depression (225º), sleepiness (270º) and relaxation (315º) [27], as shown on Figure 4.

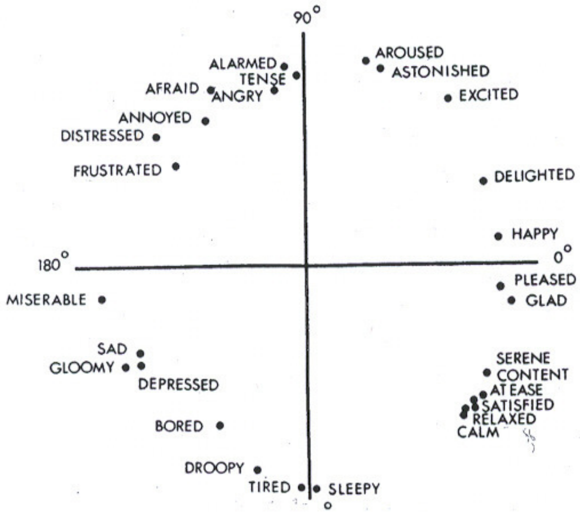


Fig. 4. Russell Circumplex Model [29].

Other methods can be mentioned, such as Whissel, Thayer, Tellegen-Watson-Clark, OCC (Ortony & Clore & Collins) and Izzard Model of Mood, but for this dissertation is rather interesting to reflect on the role of emotions in terms of their aesthetical value (due to the fact that the outcome of this work is an instrument for Art production). Considering our aesthetical purposes and mainly gesture expressivity (a topic that will be addressed on the next section), we find it quite limiting to focus only on emotions rather than on a more relative and embracing field called affective states[2]. These states can be described as observable behaviors representing the expression of a subjectively experienced feeling state (emotion). Therefore, in the next section we present some reviews of literature about emotional expressivity with gestures and the role of emotion in Arts, but later on, when speaking of implementation, we will consider affective states as the primary interface between all the components of the system.

2.2.3 Gestures and gestural expressivity

A gesture is a form of non-verbal communication [30]. It seems to be a universal, natural and spontaneous form of expression, even become a form of language all by itself. All the bodily actions communicate particular messages and, especially in our case, it is a way of communicate emotions. And this emotional expression through gestural messages is, of course, widely used in Arts (with the purpose of convey emotions to an audience), being mainly studied in a dance paradigm. However, irrespective of the context, no one doubts that a gesture can be considered expressive depending on its ability to communicate some expressive information or content [3]. So we are interested in studying the link between emotions and gestures when painting, an area where there is still a need for research, mainly when considering the new opportunities this matter could provide to the emerging scene of multi-touch interaction. Kipp and Martin [31] have studied the expression of hand gestures, while Fagerberg et al. [32], Camurri et al. [3], Castellano et al. [33] and Wallbott and Scherer [34] have

[2] This process of choosing which affective states will be addressed on this work was perform in collaboration with an expert on performance and action painting in real-time, as described later on section 4.3.1.

researched whole 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 canvas, as is the case in our work. Thus, we will conduct some experiments aimed at finding patterns of movements for each affective state addressed, being it important at this point to introduce some insights about expression analysis and body language.

Izard claims there are emotions with characteristic expressions or patterns that convey particular meaning or information for the expresser and the perceiver [35]. In Laban Movement Analysis [36], 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 [34] define six dimensions to analyze body language, 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. In [37], Camurri et al. report 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, expression is performed using a single point in the wall, rather than the whole body. Thus, for AffectiveWall we adapted the model proposed in [34] to the performance of gestures on a surface. For instance, spatial extent in their model is represented by occupation of the canvas in our case. Also we use some descriptors proposed in [37], namely vertical location, compactness and spatial occupation.

In addition, as stated in [3], technology enables the creation of an “extended expressive gesture” that is only partially related to explicit body movement. The authors consider that is the result of a juxtaposition of several dance, music, and visual gestures, and it is perceived as a whole multimodal stimulus by spectators. In a performance using AffectiveWall we also have a similar situation, where physical movements, music and digital paintings join together in a whole experience that we call performer’s “augmented affective states”. In the end, this is the output of the system, a way of showing what the performer wants to express, by including him/her in the system output and augmenting the expression using visual and sound stimuli.

We want to explore this kind of affective expression for artistic creation, so, we will now focus on how emotions influence music and painting production and perception.

2.2.4 Emotions in Art

Emotions result from the subconscious and therefore, the emotional reactions of each one often becomes a har to solve puzzle [38]. So, Art arises as a help to realize them and, according to Collingwood [39], is the expression of emotions in a given language - words, sculpture, painting, etc. - therefore contributing to their decoding [40].

Music: affective expression and perception

"Music is the shorthand of emotion."
Leo Tolstoy

Tolstoy, as probably everybody on Earth, had no doubts about the role of music on the emotional state of each one of us. Nietzsche stated that life without music would be a mistake and, in fact, emotional reaction to music is so strong that it is commonly used as a stimulus to emotions research [41,42]. Even more relevant for these studies is the fact that music is universal, rising above any language or culture, accessible to any untrained ear [2,43,44,45] and, as emotional expression, to any musician (even amateurs) [46].

Underlying the present dissertation, a living paradox should be mentioned that will always be present: the "music expresses emotions" versus "music induces emotions" issue. There is a big difference between classifying a music as sad and the listener actually feeling sad when listening to it [47]. And this process can be influenced by intrinsic or extrinsic causes [47,48]. Regarding the extrinsic factors, they also can be divided in emergent icons from the similarities between certain musical structures and agents/events that suggest some emotion (like a violin solo that mimes a child crying), or even casual associations with intimate events, people or contexts (like emotions derived from childhood memories).

Regarding the intrinsic factors, these relate with certain features of musical structures, like scale, tempo variations, dynamics/volumes, articulation[3], envelope[4], pitch, richness of the harmony, timbre, etc. [2] [49,50,51,52,53,54,55,56,57,58], as stated by Seashore [59] and later proved experimentally in [60,61]. In these experiments, a musician played a musical piece with ten different emotional nuances and the majority of listeners were able to recognize the emotional intention of each performance. Juslin concludes in a similar experiment that 70% of listeners can understand emotions that a musician wants to express [114]. Hevner [62,63] and Farnsworth [64] define and organize a set of possible emotions that are able to be transmitted by music and, through [47,54,65,57,66], we summarize these information about mapping emotions to musical features on Table 1.

.

[3] Articulation of a rhythm corresponds to the velocity of its decay (the duration between the beginning and the end of a sound), and it can be fast (staccato) turning the notes perceived as isolated, or slow (legato) turning the notes connected between them with fluid passages and without silences.

[4] Envelope represents the volume changes of a sound along time. This feature relates to the attack and decay of sound.

Table 1. Summary of the mapping between emotions and musical features.

Affective States	Sound Parameters
Satisfaction, pleasure	Fast tempo, few harmonics[5], low frequencies with many variations, sharp envelope, little variation of amplitude.
Excitation	Fast tempo, many harmonics, high frequencies with many variations, sharp envelope, little variation of amplitude.
Power, strength	Fast tempo, many harmonics, high frequencies, soft envelope.
Anger, rage	Fast tempo, many harmonics, high frequencies with few variations, high volume, staccato articulation, distorted timbre, sudden attacks, noise.
Boredom	Slow tempo, low frequencies with few variations, few harmonics, soft envelope.
Disgust	Slow tempo, many harmonics, many variations of frequencies, soft envelope.
Fear	Fast and irregular tempo, many harmonics, high frequencies with few variations, soft envelope, large variation of amplitude, low volume with many variations, staccato articulation.
Joy, happiness	Fast or moderate tempo, few harmonics, many variations of frequencies, sharp envelope, moderate variation of amplitude, moderate or high volume with few variations, staccato articulation, bright timbre.
Sadness	Slow tempo, few harmonics, low frequencies, soft envelope, legato articulation, melancholic timbre.
Surprise	Fast tempo, many harmonics, high frequencies with many variations, sharp envelope, high volume, staccato articulation, bright timbre.
Solemnity	Slow or moderate tempo, few harmonics, few low frequencies with few variations, soft envelope, moderate or high volume, legato articulation.
Tenderness	Slow tempo, few harmonics, soft envelope, low volume with few variations, legato articulation, soft timbre.

Lastly, resuming the discussion about cognitivism and emotivism, respectively, how music expresses an emotion that can be perceived by the listener and how music actually can induce an emotional response on the listener [47,67,68,69], we can loosely establish as a long-term challenge to AffectiveWall: try to transform "simulated" expressions (with meaningful gestures that can express emotions but not be truly felt) in more candid expressions where users have their affective state changed through the output they created (creating a feedback loop). These questions will be addressed later on but they mainly rely on the rise of awareness of own gestures, the conscience of one's own expressivity derived from the emancipation from the typically "focused on the instrument" way of playing.

[5] Sound is typically composed by many frequencies, the fundamental (the most audible) and other harmonics that are multiples of the fundamental, giving complexity to the sound and defining its timbre.

Painting: shapes and colors as affective expression

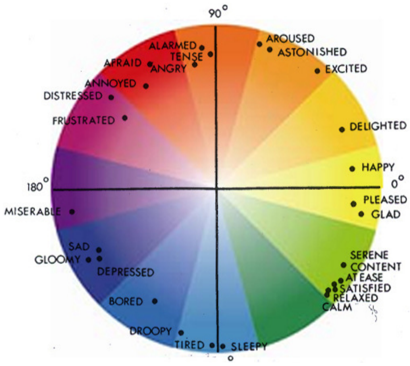
"Colors, like features, follow the changes of the emotions."
Pablo Picasso

Since Romanticism at least, painting flows more than just inks. Michelangelo claimed that Man paints with his brain and not with his hands. Picasso considered that he didn't paint what he sees but what he thinks about what he sees. After all, an artist with something to say doesn't limit himself to doing a photograph of reality, performing instead a personal interpretation that inscribes his emotional state into a timeless support.

And if the canvas is the vehicle, colors are the language. Representing one of the strongest stimuli that human brain can perceive, colors have the capability of affect cardiac rhythm, time perception, weight estimation, size and temperature [70]. By inducing emotions through direct effect on sensorial perception, colors stand for one of the most effective way of controlling the state of mind of an artistic work. This was widely explored by abstract artists, e.g. Kandinsky, who formulated almost an alphabet of meanings of colors to communicate thoughts and emotions through his paintings [71]. For example, for him, yellow represents an aggressive and earthly color (mainly related to madness) and blue is celestial, divine and calm.

In order to understand communication through color, it is mandatory to review some of the theorizing, classification and quantification of color made in the past. Goethe, in "Zur Farbenlehre" (1810), talks about how blue or blue mixed with red represent negative feelings and, on the contrary, yellow or yellow mixed with red represent positive feelings [72]. On the other hand, Ryberg [72,73] considers a color scale based on energy transmitted (red as the most energetic and blue as the least energetic color). Wright e Rainwater [74,75] categorize 48 emotions derived from color, scaled by happiness, ostentation, strength, heat (warm or cool colors), elegance and calmness; while Hogg [74,76] considers only 12 emotions, identified by impact (related to the chromatic component[6] of color), frequency with which color appears, evaluation (pleasant or unpleasant) and heat (related to the color tonality). In a latest study, Hogg and other authors [74,77] determine other factors, namely dynamism (related to the chromatic component), spatial quality and complexity (both related to the luminosity), emotional tone (related to the color tonality) e evaluation. Also Kobayashi [74,78] described a scale based on the three dimensions - warm/cool, soft/hard and clear/grayish - which can be linked to the dimensions invented by Sato et al. [74,79] – heat, weight and activity – which in turn relate, respectively, with tonality, luminosity and chroma. Still from [74], we can conclude that color classification in general is culture-independent, being these results of all studies cited above and the experiments documented in [80,81] shown on Figure 5.

[6] Chromatic component corresponds to the color "pureness", i.e., the higher the chroma of a color, less black, white and gray has.



Emotion	Hue
Happiness	0.167 (Yellow)
Love	0.667 (Blue), 0.75 (Violet)
Pride	0.167 (Yellow)
Tenderness	0.75 (Violet)
Curiosity	0.5 (Cyan)
Contentment	0.083 (Orange)
Anger	0 (Red)
Sadness	0.75 (Violet)
Fear	0.667 (Blue)
Disgust	0.75 (Violet)
Shame	0.083 (Orange)
Jealousy	0 (Red)

Fig. 5. Summary of different studies about the relation between emotions and colors (the table on the right was taken from the Bresin study presented on [81]).

Particularly the results from the experiment of [81] are very interesting for our work, due to the fact that they are related to colors that can help the affective expression of a musical performance. In terms of shapes and drawings related to emotions, we found little research in this area, but some studies can me mentioned, namely [82,83] exemplified on Figure 6.

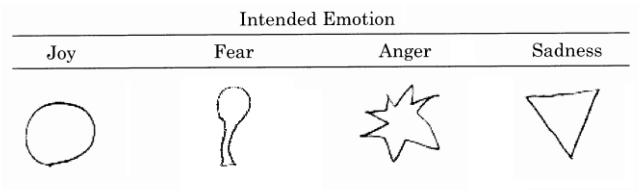


Fig. 6. Some results from experiments about emotional expression with drawings [82,83].

To conclude, another work should be mentioned where empirical data prove that emotions expressed by drawings can be conveyed through musical performances [84] - a statement that supports one of the main goals of the present thesis.

2.3 Inherent concepts in the creative process of AffectiveWall

2.3.1 Creativity

"Some men see things as they are and ask why.
Others dream things that never were and ask why not."
George Bernard Shaw

Some - like director Frank Capra - associated it with hunches, others - like writer William Plomer - with the power of connecting the seemingly unconnected. In fact, the manifestations of inspiration on artists have always been a discussed theme. According to [8,85,86,87], creativity is the act of creating something original and appropriate (useful or with an accentuated aesthetic value), and multidisciplinarity stands for a fundamental factor to its development. From here one of the premises of AffectiveWall arises: the stimuli of linking different artistic processes (specifically music and painting) to promote novel artistic contents.

2.3.2 Synesthesia

Synesthesia is a neurological phenomenon in which stimulation of one sense leads to the stimulation of another [8,88], consisting of a situation that enables numerous creative possibilities in an artistic context. An example of this is composer Olivier Messiaen who, when reading or listening to music, could visualize moving colors related to specific chords. On the other hand, painter Wassily Kandinsky, when seeing colors, could experience an audible version of them. In fact, there are so many cases of this kind of phenomenon in the music and painting fields that a new artistic movement was created: the Synchromism. Founded by painters Morgan Russell and Stanton Macdonald-Wright with the assumption that sound and color are intimately correlated, this movement allows the creation of visual works that evoke musical sensations [8,89] – an approach quite similar to AffectiveWall's.

2.3.3 Improvisation

Improvisation can be defined as a musical creation that takes place on its own performance [90], emancipating from the sheet music to create something new, that balances novelty and coherence, halfway between the expected and the unexpected [91,92]. From Bach in the baroque period, Mozart in classical and Beethoven in Romantism [90], to any current jazz player, improvisation has always had an important role in musical creativity. It is a vehicle for consciousness expansion and the tapping of deep intuitions [93], a powerful way of developing individual expressivity and what one can create when focused on Art through emotional impulses – a rather important leitmotif to AffectiveWall.

3 Related work

3.1 Virtual instruments and interactive musical systems - the new approaches to musical instrument

Built for so long as there is civilization, musical instruments have suffered radical changes over time. In fact, technological evolution in the 20th century accelerated more that ever new approaches of designing and thinking an instrument: electro-acoustic, electronic and even virtual instruments are just a few examples [94]. While conventional instruments have limited timbre and way of playing, as result of its own shape and setup (strings, percussion, etc.), electro-acoustic and electronic instruments began to enlarge the range of possible sounds (in electro-acoustic ones by manipulation of recorded acoustic sounds and, in the case of electronics, by synthesizing new sounds or sounds from other contexts (e.g. a synthesizer with wind sounds)). Extending these concepts to a whole new level is the virtual instrument [94], allowing not only any audible sound but also any human gesture as an input. Through, multi-touch interfaces, sensors or video cameras, for example, it is possible to adapt musical creation to body language and, therefore, to have a total customization to a musician and a reduction of the learning curve.

Also important for the present dissertation is to introduce the concept of interactive music system, which explores the notion of virtual instrument and studies the possible ways of interaction. It can be between the musician and the system, as a typical user interface where a user has the creative role and the system only reacts to his input, or between the system and an environment. In the first case, interaction can be divided into information reading (gestural, audio, etc.), interpretation or processing of data, and audible output production [95]. An example of this kind of systems is the Bloom[7] app, where users can click or tap on the entire interface and different sounds are produced depending on the spots touched (axes are related to sound frequencies and tones). Regarding to system-environment interaction, the system has the main role on the interaction, responding to a complex environment (that can be composed by one or more users, but also by space conditions,

[7] Bloom Multi-touch Music Application, by NUI Group, available on: www.nuigroup.com/bloom/bloom.swf (accessed on 11-06-2010)

environment noise, etc.) while users can indirectly interact with the system though actions on the environment. One example is the AESI interface (Audible Ecosystemic Interface), invented by Augusto Di Scipio [1], which is set as an ecosystem with an automatic feedback cycle: the environment has sound that is captured by microphones, passing to controller, processing and synthesis steps, producing again sound that comes back to environment through speakers. So, in this case, sound is the interface and user can only interact with the system by prompting actions on the environment that will be read and interpreted by the system.

Therefore, AffectiveWall relates to these two kinds of interaction because, on one hand, the user relates to the artistic creation by performing gestures that are read by the system. On the other hand, affective states are the real interface in this interaction, 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). Thus, we need to study the different ways in which the user can interact with this "affective ecosystem" using his own body language, i.e., to review the current solutions on gestural control of music, tangible/multi-touch interfaces and systems about visual production of music.

3.1.1 Gestural control of music as an interface

As said by Antonio Camurri [96], in order that a system interacts in an intelligent and spontaneous way with humans its communication capabilities should deal with human cognitive, affective/emotive, and motor capabilities. Thus, sometimes it is more productive for users to interact with a system by non-verbal communication like gestures for controlling parameters of a musical creation. As a real-time digital instrument it is mandatory to consider some features about this topic [94,97,98,99]: 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 and, finally, the separation between the interface and the sound synthesis (both modules are independent and have a mapping of parameters between them). Regarding this last point, Knapp e Cook [41] describe the potential of this abstraction, where this separation between musician and the object responsible for the generation of sound 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. I.e., the user does not need to worry about using the instrument to produce low-level contents, like chords, but to convey a specific emotion or affective state, a high-level content.

Moreover, it is mandatory to connect this gestural control with the metaphor of painting. Its relevance cannot be overlooked, being the main help to users by simplifying and providing them directives on how to use the system [100,101,102] (which avoids the use of manuals or other external supports). According to Bill Buxton, the most natural interaction language consists of non-verbal dialogues, enhancing gestures as phrases with their own meaning [103]. Therefore, an interface should be replaced by actions that naturally derive from the supported metaphor, rejecting the use of menus, buttons and windows (in the case of AffectiveWall there is only a canvas, as in painting).

Examples of this kind of interaction are EyesWeb [104,105,106] regarding the video component, Hands [107] and Conductor's Jacket [108] regarding haptic interfaces[8], and various tangible surfaces that will be presented in the next section. In the case of the Hands project, it consists of a glove full of sensors that reads the hands' movements performed by the user. On the other hand, Conductor's Jacket [108] monitors all the physiologic signals and interprets them according to the emotional expression. As for the EyesWeb platform [104,105,106], developed by Camurri, it is a multimedia system applied to dance that receives input from a video camera, processing these data about the body movements of a performer and translating them to visual and musical outputs on real-time according to the expressivity shown. For a more extensive list on this kind of interfaces, please refer to [94] and [109], this last one regarding to multi-touch surfaces, as presented in the next section.

3.1.2 Tangible and multi-touch interfaces

"Multi-touch-sensing was designed to allow non-techies to do masterful things while allowing power users to be even more virtuosic."
Jefferson Han

The history of touch technologies dates back to the transition from the 60's to the 70's, when Hugh Caine and Bob Moog (builders of electronic instruments such as synthesizers) performed some experiences with surfaces covered with capacitive touch sensors [109]. At the same time, IBM built the first touch screen and also the PLATO IV pedagogic platform was released to the general public. Also in the 70's, composer Iannis Xenakis (one of the pioneers in computer music) proposes the UPIC tablet which produces sound through drawings of its own spectrogram [110], and Bill Buxton develops some seminal studies about multi-touch interaction on the Dynamics Graphics Project. Still from [109], reliable solutions based on video started to appear in the early 80's, namely the VideoPlace/VideoDesk project by Myron Krueger and able to monitor more than one finger or hand. As for a truly multi-touch screen, it appeared in 1984 with the work of Bob Boie, based on a CRT monitor covered by capacitive sensors that enabled manipulation of graphical objects with an acceptable response time for a real-time system. Despite that, this technology only becomes widely spread since 2000, with various interesting solutions that are presented below:

Magic Wall, Perceptive Pixel, etc.

One of the main figures in this field is Jeff Han, a researcher who, in 2006, presented a groundbreaking concept of multi-touch interfaces [111,112], consisting of a camera and video-projector placed behind the surface itself that, because of its low-cost nature, became an inspiration to many other projects presented below. He is the creator of Magic Wall, an interactive wall used by CNN in news programs, and leader of Perceptive Pixel, one of the companies with avant-garde researches on this topic.

[8] Haptic interfaces consist of devices that explore touch as a way of interacting with the system, capturing the user's gestures with sensors and other similar tools.

waveTable

Multi-touch interface for music composition specialized on manipulation of sound waves [100].

Scrapple

Artistic installation with a behavior similar to a sequencer [110], where music is created by physical objects arranged on a table. These objects are then read by a video camera scanning their positions and shapes, which will lead to a sound production according to a matrix of instances in time and sound frequencies (respectively x and y on the table).

Xenakis

Collaborative musical instrument with a rather simplistic composition process, where users can play with physical objects to insert data about rhythm and notes of virtual instruments in a probabilistic model that generates the sound [113].

AudioPad

Modular interface for composition in real-time through objects that interact with some performance parameters – controlling of effects, definition of repetitions, sound mix, etc. – in an approach related to samples manipulation [114].

reactTable

Collaborative system for composition in real-time, based on physical objects disposed on a surface, where their positions and orientations control the sounds produced [115,116]. Each object is a sound (sample, sound wave, etc.) or an effect, and various objects can interact between them depending on their distance. In addition, it allows the manipulation with the hands and fingers of various parameters of each object, using metaphors of existing professional music devices like knobs to control volume or pitch.

There are many other systems that could be mentioned, like Microsoft Surface[9], Lemur by JazzMutant[10], SmartSkin [117], ConDio, Orai/Kalos, etc., but an extensive analysis of this kind of interfaces is out of the scope of this dissertation. Thus, further information about this topic can be found on Natural User Interface website

[9] Interactive table released in 2007 and capable of monitoring fingers and objects in order to control a large number of tasks (photo manipulation, videogames, etc.). More information on its website: www.surface.com (accessed on 11-06-2010)

[10] Controller of musical instruments and electronic devices (such as synthesizers, sound mixers, video-performance systems, etc.) through a highly modular and customizable multi-touch interface. More information on its website: www.jazzmutant.com (accessed on 11-06-2010)

(NUI Group)[11], and a brief chronology of research projects in this field can be found on Bill Buxton website [109], one of the main figures on this area.

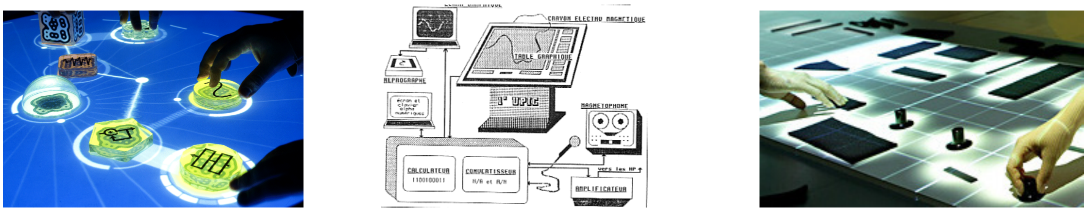


Fig. 7. Examples of tangible interfaces. From left to right: reacTable, UPIC and Scrapple.

3.1.3 Systems of visual production of music

Having presented and explained various kinds of interaction with the user, it's time to explore some concepts about the visual representation of sound in other systems that create music. The visual component is a relevant point for musical composition due to the fact that sketches and drafts are a fundamental part of the creative process [118]. According to the experiments of [119], 75% of composers use drawings at the initial steps of their compositions. Visualization of the sound parameters in real-time is a highly useful feedback for a composer/performer and even for an audience, not only in helping to understand complex structures and aspects on the composition, but also in generation new ways of communication with the audience (for example, at an emotional level).

The already mentioned UPIC system, by composer Iannis Xenakis, has represented a valuable inspiration for many solutions in this field, such as:

IanniX

Based on the assumption that definition of time as a fixed axis with linear progression (as is typical in this kind of applications) restricts the creation of various dynamic events that evolve at different speeds, this application uses an abstract interface composed by circles and lines enabling the existence of many objects (with different dimensions and times) that parameterize various concurrent events [120].

Metasynth

Application dedicated to additive sound synthesis through drawing and graphic manipulation of its spectral component[12]. The main advantages of this application in comparison with UPIC are the existence of sound feedback in real-time and the usage of color and luminosity of each stroke to define spatialization and volume. Despite that, it shares with the system idealized by Xenakis the creative process and the representation of axes –

[11] wiki.nuigroup.com/Papers (accessed on 11-06-2010)

[12] Available on: uisoftware.com/MetaSynth (accessed on 13-06-2010)

the horizontal meaning time and the vertical sound frequency – therefore restricting musical production to a specific kind of synthesis.

AudioSculpt

Software developed by IRCAM[13] with the same kind of sound synthesis as Metasynth, i.e., drawing and manipulation of sound spectrum [121].

Sonos

Application of real-time sound manipulation that uses color to handle a spectrogram [122]. Like Metasynth and AudioSculpt, its sound outputs are limited to the graphical organization of mathematical functions, not providing, for instance, tools for timbre control.

Hyperscore

Software that makes musical composition so simple that it can be done by everyone, including children and people with mental disabilities [123]. This is done by dividing the process in two different phases: in the first one it allows user to compose sound motives and associate them to colors, and the second phase consists of “painting” a sheet music with the pre-defined colors (where the horizontal axis represents time and the vertical one represents variations of sound parameters defined in the motives).

Music Sketcher

Is an application that does not restrict the axes to pre-defined parameters (such as time or frequency, as usual in this kind of software), taking as fundamental assumptions the ambiguity and imprecision typical of sketches, in order to allow for vague drawings that afterwards will be associated with sounds [118].

Granular Cloud Generator

Based on granular synthesis, this application allows the creation of events that will be afterwards distributed on time, showing them as clouds of sounds[14].

Chromatic Bricks

This platform uses the correlation between music and color to create compositions that respect a metaphor of a wall (as a sheet music) composed by organizations of colored bricks [124]. Each brick have a musical phrase (that corresponds to a specific color) consisting on sequences of notes in a melody, with a length associated (that defines the phrase duration).

[13] Institut de Recherche et Coordination Acoustique/Musique
[14] Available on: www.hkbu.edu.hk/~lamer (accessed on 13-06-2010)

Phonogramme

This graphical editor transforms images into sounds associating, for instance, lines with sound envelopes and color levels with amplitudes [125].

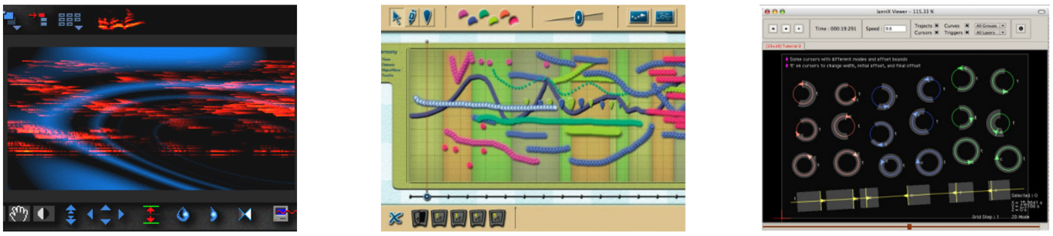


Fig. 8. Examples of visual representations of music. From left to right: Metasynth, Hyperscore e IanniX.

Sense^2

Lastly, and in a totally different field compared to the other solutions presented, there is Sense^2 developed by Pintado [8]. It is a system that generates rhythms according to paintings, applying features of the shapes painted (mainly their stability coefficient) using the "Just in Time" theory [126]. This stability (or instability) is calculated with the angles formed by all edges of each shape, represented on the output by the regularity (or irregularity) of the rhythm. For example, a circle or a line produces calm and regular rhythms, but a stranger and sharper shape (mainly composed by acute angles) will produce more nervous and odd rhythms.

3.1.4 Review

So, to conclude this section, we have selected some ideas from the virtual instruments and interactive musical systems concepts, such as the way of interaction from the AESI from Di Scipio [1] and the requirement that any gesture on the surface can be a valid input to the system. We took also some inspiration from the systems of gestural control of music that we studied, namely EyesWeb from Camurri [104,105,106], which relates to AffectiveWall because of its final purpose: give sound and visual feedback of the user's movements. But, instead of approaching a dance paradigm (thus considering a 3D space) like on the EyesWeb system, in the present thesis we focused on a 2D interaction using a painting paradigm – a novel approach that we wanted to explored in our work. About this kind of interaction with a physical medium (the "touching" canvas or wall), and due to the fact that we wanted to use a multitouch surface (reasons explained later in section 4.2), we have reviewed some interfaces mainly related with sound production. An obvious inspiration is ReacTable [115,116] for the completeness and creative approach but, as other solutions, supports itself on the manipulation of objects or sound waves in a horizontal table. Relating to this, we clearly distance ourselves by rejecting any kind of external accessories or paradigms for sound creation (like sound waves), following the sayings of Bill Buxton [103]: an interaction should be a non-verbal dialogue where gestures are phrases with their own meaning. In our case, with a meaning related to affective expression when painting. About visual production of music, all of the systems we reviewed were design having in mind different paradigms from ours: spectrograms, interfaces with geometrical components for controlling the sound, image-to-sound conversions, etc. But one in particular – the Hyperscore software – has two aspects in common with AffectiveWall that we aim to achieve. Firstly, the goal of giving the opportunity to everyone (children and people with disabilities included) to compose music. And, on the other hand, having the composition process divided in two stages: one of them in a micro-composition level (in Hyperscore, users create sounds associating them with colors and, in AffectiveWall, we create sounds associating them with affective states) and the other one in a macro-composition level (in Hyperscore, users paint with the predefined colors to create a music and, in AffectiveWall, users paint with their affective expression, that relates to our synths associated with each affective states addressed).

3.2 Affective Computing

As previously said, emotions have a central role in the rational and intellectual behavior of human beings [127,128,129]. And given that technology is moving towards an increasing adaptation to humans ([130] suggests that 84% of users feel frustrated with current human-machine interaction), it becomes mandatory to explore emotional sensitivity of people to a better communication, adapted to their idiosyncrasies, reactions and requirement changes [131]. Thus arises the concept of Affective Computing, a field that studies the influence of emotions on computational systems [22]. Picard, a pioneer researcher in this area, argues that this can no longer be considered as science fiction [132,133], showing that it is possible for a system to recognize with a success rate of 80% a set of emotions personified by an actress (namely neutral, anger, hate, grief, platonic love, romantic love, joy and reverence) [127]. Information about emotional interaction with computers can be consulted in [134,135,136,137] and, as examples of applications of Affective Computing, we can mention:

SenToy

A doll that works as an interface to manipulate the emotional state of a video-game character [138]. It has sensors that capture its user's gestures and movements, interpreting according to various emotions: anger, fear, surprise, sadness, disgust and happiness.

eMoto

SMS messaging system that adds a background image to the message according to the user's emotional state, being emotions expressed through gestures captured by an accelerometer and a pressure sensor [72,32,139,140]. In this way, the system commits cognitively and physically with the user, giving a more intimate nature to communication.

MOR2ART

Interactive music system that, like AffectiveWall, aims to give the opportunity to non-musicians to express their emotions through music [65]. Besides the expression of emotions, it enables the control of their intensity, converting them to an output on MIDI according to an estimation of notes, timbres and velocities[15] (requiring a synthesizer to transform the MIDI notes into sound). Unlike the gestural approach of AffectiveWall, in this system the input of emotions is based on a direct interface where the user clicks on the emotion that he wants to express.

I-Sounds

An automatic composition system, developed by Cruz [47], that generates music for virtual environments based on emotions. It is divided in three main layers – affective, composition and output – using an algorithm named AMADEUS that works on three steps: base (where tempo, clef and beat are defined), rhythm (where rhythm is composed according to the "Just in Time" theory [126]), and mode (where the rhythm created is associated with sound frequencies).

3.2.1 Review

In terms of summarizing the section about affective computing, the most relevant system to be mentioned is MOR2ART, due to being a system that controls music through emotions. This is performed by controlling the level of each emotion expressed on a music in a MIDI file, by clicking on the desired spot on a dimensional space similar to the Russel Circumplex Model [29]. The main issue of this software, the one that we want to solve, is the kind of interaction that it allows. In MOR2ART, the emotion is inserted in a direct manner, a way of interaction that, despite providing more accuracy, we do not believe that will be the most sincere. So, our approach will be providing this kind of intensity control for each affective state but in a gestural (hence more natural) way of input, as already stated in section 3.1.4.

3.3 Technological context

3.3.1 Tracking software and multi-touch libraries

There are many available options to perform the gestures tracking (i.e., the analysis of the data captured by the camera), all of them open source, multiplatform and adaptable to a large range of cameras and kinds of multi-touch setups. In addition, all of them work with the common communication protocols (introduced in the next section): OSC and TUIO. The main software's are OpenCV[16] (developed by Intel), Community Core Vision (CCV) or tbeta[17] (by NUI Group), Touchlib[18] (also developed by NUI Group) and reacTIVision[19] (developed by reacTable team).

[15] Amount of "strength" applied to play a note.
[16] Available on: http://sourceforge.net/projects/opencvlibrary (accessed on 15-06-2010)
[17] Available on: http://ccv.nuigroup.com (accessed on 15-06-2010)
[18] Available on: http://nuigroup.com/touchlib (accessed on 15-06-2010)
[19] Available on: http://reactivision.sourceforge.net (accessed on 15-06-2010), described in [149].

3.3.2 Communication protocols

One of the most common protocols to send and receive messages from tangible interfaces is Open Sound Control (OSC), offering great precision, interoperability, flexibility and large documentation [141,142]. Moreover, it has many online resources, such as the Oscpack, a simple, complete and portable implementation of OSC based on C++. On the other hand, the reacTable team [143] developed a framework based on OSC that enables the sending of messages of higher level information about multi-touch events, objects and gesture features (like position, angle, area, volume, speed, rotation, acceleration, rotational acceleration, etc.). The name of this protocol is TUIO and it has the advantage of providing a large number of open source software online, developed on many different languages like C++, Java, C#, Processing, Pure Data, Max/MSP, Quartz Composer and Flash AS3. Also, many implementations of the framework are available for OpenFrameworks, Python, Pure Data, Processing and Max/MSP, and implementations for the trackers reacTIVision, CCV and Touchlib. Another important feature is the existence of different kinds of multi-touch event simulators to help debugging the code, e.g. TUIO Simulator (based on Java), Windows MultiPoint TUIO and Multi-Platform C++ SimpleSimulator.

3.3.3 Visual synthesis software

When looking for tools that are able to create visual contents for AffectiveWall, we found the most efficient ones to be ActionScript[20], OpenFrameworks[21] or Processing[22]. Regarding ActionScript, the language used in Adobe Flash, its ease of implementation and its extensibility should be mentioned. On the other hand, it is not an open source software and has a critical lack of performance, a problem to a real-time instrument like AffectiveWall. On contrary, OpenFrameworks has a great performance, due to being based on C++, and is a simple platform focused on production of creative contents. It combines OpenGL for graphics and rtAudio for sound, being multi-compiler and multiplatform. Lastly, Processing is a simplification of Java specialized in artistic contents (widely used by artists and non-technologic people), also open source and multiplatform, but with major issues about performance (is based on Java).

3.3.4 Sound synthesis software

The two main softwares for real-time composition are SuperCollider[23] and Pure Data (Pd)[24]. The environment of SuperCollider is divided in "language" and "server", the former, developed on a combination of object-oriented and functional languages, has the responsibility of being the client, calling processes on the second module – the "server" – that consists of several plug-ins implemented in C language for the creation of composition algorithms [144,145]. Communication between the two modules is performed with OSC, potentiating the connection with other softwares (e.g. external control of the sound synthesis in the server, for instance, by OpenFrameworks sending direct OSC messages). In addition, the efficiency of its implementation

[20] Available at: http://www.actionscript.org (accessed on 15-06-2010)
[21] Available at: http://www.openframeworks.cc (accessed on 15-06-2010)
[22] Available at: http://processing.org (accessed on 15-06-2010)
[23] Available at: http://supercollider.sourceforge.net (accessed on 15-06-2010)
[24] Available at: http://puredata.info (accessed on 15-06-2010)

and the dynamic nature of its language, enabling the live composition of sound without rigid constraints, should be emphasized. On the other hand, Pure Data is a graphical language based on "boxes" and connections between them, enabling the development of objects (externals) and modules (patches) [146]. Created by Ircam[25], it is an open source version of the commercial software Max/MSP, keeping the programming style of patches based on analog synthesizers.

3.3.5 Review

The review of the technological context is performed in the next section, more precisely, in the section 4.2 about the choices that we made in the development process.

[25] Institut de Recherche et Coordination Acoustique/Musique

4 Solution

Having described the state of the art that supports the creation of AffectiveWall, it is time to organize and compile all the information in a system that fulfills the proposed goals. Therefore, we will present a list of requirements to the system, an architecture to materialize it, descriptions of the implementation of all its modules and a brief explanation about the AffectiveWall development process.

4.1 Requirements

First of all, we came to a set of requirements that represent the *modus operandi* of AffectiveWall and guided us to the kind of instrument that we wanted to create. In terms of high-level requirements, we based ourselves on the inspirational project reactTable [115,147,148,149], adapting some of their requirements:

— Be easy and intuitive, enabling its use by any person regardless of his/her background;
— Make use of the visual feedback to help users compose and, at the same time, to allow and audience to better understand the contents produced;
— Be controllable with the hands or other body parts, without the need of a mouse, keyboard, controller, etc.;
— Always generate feedback, decreasing user's frustration;
— Be always ready to receive input and create output, without distinguishing edit mode and execution mode;
— Enable an use that combines freedom, expression and creativity;
— Have a pedagogic nature but, at the same time, providing a fun experience.

Other more specific requirements can be enumerated that influence the implementation of all system:

— Modularity and extensibility, to allow the change of some part of the system;
— Real-time performance, i.e., low latency to prevent ruining user experience;
— Use of standards, particularly in terms of communication protocols, to obtain a wide interconnectivity;
— According to Human-Computer Interaction standards, always respect the chosen metaphor – painting – never using external controllers, interfaces with menus, options, buttons, etc.

4.2 AffectiveWall development process and the choices made

After being inspired with all the artistic background (section 2.1) related to AffectiveWall and reviewing all the related work (chapter 3) of this kind of system, we began to think of the path we need to take. In fact, if in the beginning the main idea was just a musical instrument with a visual and interactive way of composing, this concept evolved and matured a lot in the process. The desire of creating a natural interface to the user led us to research affective and gestural expression, aiming to find a way to create the most universal instrument possible. The ecosystem concept by Di Scipio [1] was one of the steps taken to approach that goal, and joining two different types of feedback – music and paintings – to create a fully intermedia instrument that helps and augments user's expression is another. But, to completely achieve our goal, we had to find and explore correlations between gestures and affective states considering our painting metaphor. This study of how we can express ourselves through body language, mainly in our approach of using a physical medium, began to reveals itself as one of our main contributions considering that this is a poorly researched area. Therefore we conducted an experiment to define an affective model for gestural expressivity, a tool that we later implemented on our system.

So, in terms of software, programming languages and protocols that we chose to implement our system, we have to name OpenFrameworks (OF), SuperCollider (SC), Community Core Vision (CCV) and TUIO. The first one, OF, performs the core tasks of the application: receiving data from the surface, analyzing the movements and deciding what affective states are correlated to them, all the communication between the different modules of the project and, in the end, generating visuals that are projected on the surface. It was the chosen programming language mainly due to real-time performance concerns: one of the major requirements for the system. In the case of SC, it performs the conversion to sound, i.e., it is the programming language used to code all the synthesizers of AffectiveWall. We elected this language for its flexibility and openness to creativity (it is kind of an empty sheet where we can create any possible sound). For the tracking software we chose CCV (mainly due to its stability and completeness of image tweaking options), having the responsibility of reading the user's gestures on the surface in real-time. Finally, TUIO protocol for the communication between CCV and OF was chosen due to having all the advantages of OSC and also some specific data about multi-touch events. In addition, it has many implementations and documentations online, a characteristic it shares with all the other tools that we chose.

In terms of the hardware to support the system and feed it with user's gestures, we had to think about interfaces capable of reading multi-touch input. After discarding LCDs and other kinds of devices with touch screen capabilities, mainly due to size and cost constraints (two main characteristics of AffectiveWall are the possibility of being implemented on surfaces with large dimensions to simulate actual painting canvas and the fact that one of the major goals is to bring Art closer to people, requiring the lowest possible costs), we began to dig the vast world of possibilities of low-cost multi-touch surfaces brought to the general public by Jeff Han [111,112].

4.3 Affective ecosystem as interface and affective model for gestural expressivity

Having explained how the development process grew, and mainly the importance of the correlation between gestures and affective states in the whole project, we will begin to detail this component of our work, namely the experiments performed and their results.

4.3.1 Introduction

In this concept of "affective ecosystem interface", inspired on AESI by Di Scipio [1], the system has the main role in the interaction, by interpreting the whole ecosystem composed by the affective states that users tried to express (this kind of interaction is explored in section 3.1 about interactive musical systems). Adapting it to accept gestural input for understanding the affective expressivity of users implies defining a model that correlates gestures and affective states. For this we made an experiment to understand, not only the better recognizable affective states that people can express on a canvas, but also the most used features and patterns of gestures of each one of them.

In what concerns the emotion range considered, we based ourselves on models from psychology, such as Russell's Circumplex Model [29] and OCC Model [150] (the first one being one of the most recognized affective models in the scientific community and the second being an important model for artificial intelligence), and added aesthetic and artistic expression concerns, creating a group of, not only emotions, but more general affective states that are relevant when the name of the game is 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[26], an action painting and performance 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. This choice of bringing an artist to help us in this stage of our work seemed natural due to the artistic purpose of AffectiveWall. In addition, the connection between affective states and gestural expressivity when painting is the part of the present thesis that is less explored by scientific community (being an opinion of an expert very important to compensate the scarce literature in this area).

4.3.2 Experiment

Despite our goal of developing an affective model suitable to reality, it is important to note that we are not aiming to find the perfect expression that will work for everyone (which we believe to be an impossible task). We cannot forget that, as with 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 the one for conventional

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

instruments). In this way, we asked fifteen individuals to perform each affective state on a wall or similar vertical surface while we were recording them on video. Afterwards, we measured the gesture features from the video, aiming to find patterns in the subjects' body language. In addition, we took several annotations during the live performance about, for instance, the difficulties of expression felt by users. We asked the subject to perform one expression for one affective state at a time, firstly 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. In order todiscard any bias, affective states were presented in a random and different order for each subject, and tests were performed in isolation without the user having previously seen other expressions. After each gesture, the subject was asked about his/her satisfaction/confidence, regarding the accuracy with which each gesture represents the requested affective state (an example of questionnaire for this test is presented in Appendix B). Due to the long process of this experiment (which includes several screenings of each video to make the feature analysis) and the actual 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 try 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 to perform the test was made to cover a large range of ages (from 18 to 66) and a large range of backgrounds: with and without artistic background (and among artists, people from music, painting and dance), psychology, etc.

Fig. 9. User performing a series of affective states on a wall.

In terms of evaluation, we had to set a list of metrics and measures to analyze the videos. As stated in section 2.2.3 about gestural expressivity, we adapted the model proposed in [34] to the performance of gestures on a surface (e.g. spatial extent is represented in our model by occupation of the canvas) and also used some descriptors proposed in [37] (e.g. 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:

Table 2. Evaluation metrics and measures for the experiment

Evaluation Metrics	Measures
Gesture length (the length of what the user drew on the surface)	Punctual (a gesture with no associated movement when the user is in touch with the surface); Short; Medium; Long (comparatively categorized according to the videos)
Area of touch on the surface	One fingertip; One fingertip of each hand; All fingertips; All fingertips of both hands; One hand; Both hands; One arm (it can be the forearm or the entire arm); Both arms; Hands/arms and head; Whole body
Quantity of gestures	One; Some (two or three); Many (more than three)
Gesture speed (only of the touch with the surface)	Static (no movement on the surface); Slow; Medium; Fast (comparatively categorized according to the videos)
Direction of the movement drawn on the surface by the gesture	None (no movement on the surface); Downward (descending movements); Upward (ascending movements); Inward (movements enclosing to a point between the two or more parts of the user's body touching the surface); Outward (movements spreading in all directions on the surface); Sideward's (movements spreading in the horizontal axis on the surface; Random (movements with no rational or fore-thought direction)
Shapes drawn on the surface	Blob (a dot or closed region of any size); Line (straight lines and harsh movements); Curve (undulating movements); Circle; Chaos (movements with no intentional shape)
Duration of the gesture (only when the user is in touch with the surface)	Sudden (when the user hit the surface without leaving the hand or another body part that made the gesture on the surface); Sustained (gestures with a prolonged contact, when some part of the user's body slides on the surface)
Location in the surface where gestures take place	Low (same height as the user's legs and feet); Medium (same height as the users' torso); High (higher than the users' head)
Occupation of the canvas	Confined (gesture that only needs wrist movements); Medium (gesture that needs the movement of entire arm(s)); Expansive (gesture that needs the movement of performer's body)

Note that the purpose is to use the painting metaphor, thus gestures are only considered when touching the surface itself.

4.3.3 Results and discussion

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 Fig. 10), and the fact that the majority of the group, independently of the background, revealed acceptable confidence rates in the expressed affective states. 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). Also we made a comparison between a spontaneous expression and another one made after reflecting about the affective states addressed.

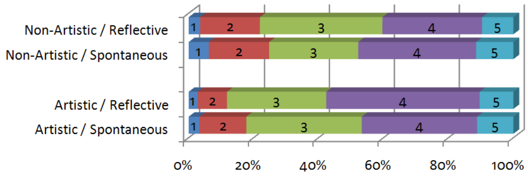


Fig. 10. Satisfaction/confidence ratings by background and kind of gestural expression, allocated by a 1 to 5 scale (being 1 the lower confidence and 5 the higher).

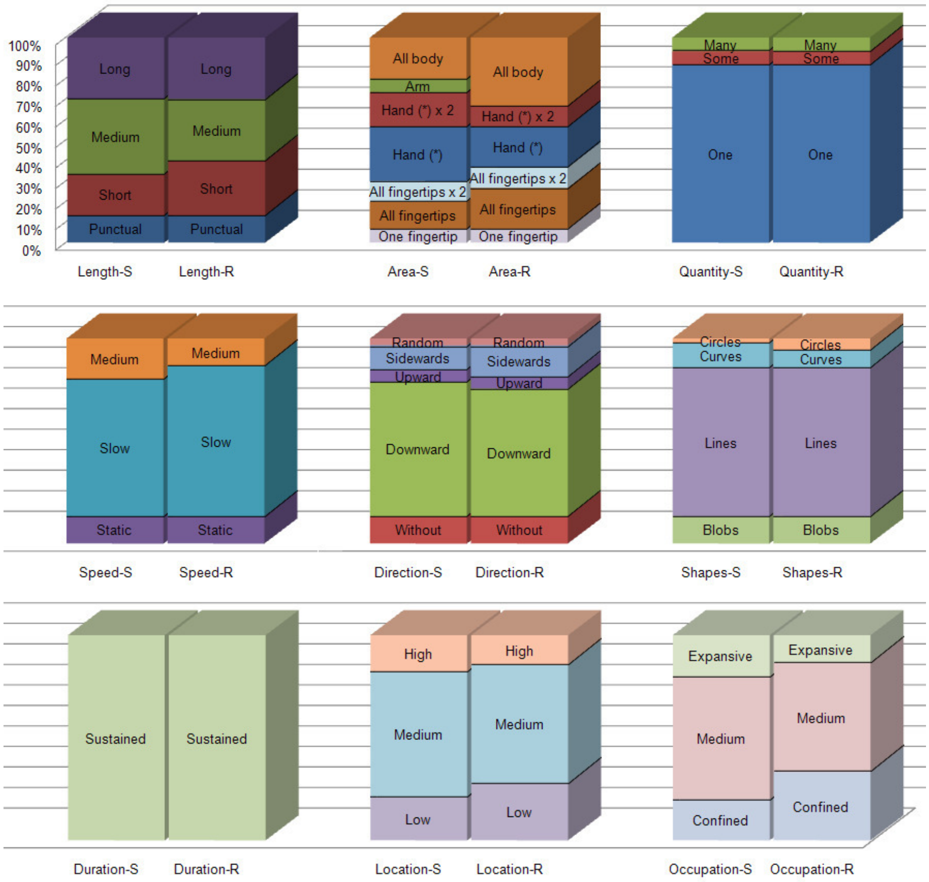


Fig. 11. Experimental results for sadness, where each of the graphs (representing the expression features) has two columns: one for spontaneous movement (the 'S' suffix on the label) and another for reflectivemovement (the 'R' suffix on the label).

As in the example depicted in Fig. 11, some features appear to change between spontaneous and reflective expression (e.g. area of touch) while others remain similar (e.g. quantity). We used a related-samples marginal homogeneity test for each emotion and each feature, in order to determine whether the difference between these two kinds of expression is significant. The null hypothesis is that distributions of different values across the two types of expression are equally likely, considering a significance level of 0,05. We also associated ordinal values with all the features except the ones concerning direction and types of shape (for instance, in terms of feature speed, we defined: 1=static, 2=slow, 3=medium and 4=fast). For that reason, direction and shape were not subjected to this statistical test. Four examples are presented on Table 3, namely sadness, tenderness, joy and anger.

Table 3. Examples of related-samples marginal homogeneity test for each emotion and each feature (p=0.05). The cells without value are the ones that the test could not compute, giving asymptotic significance as result.

Affective State	Length	Area	Quantity	Speed	Duration	Location	Occupation
Sadness	0,414	-	1	0,157	1	-	0,18
Tenderness	0,763	0,866	0,739	1	1	1	0,083
Joy	0,157	0,705	0,317	-	0,317	0,317	0,157
Anger	-	1	-	0,527	0,317	0,317	0,564

As exemplified by Table 3, although some differences exist between spontaneous and reflective expression, for the majority of features, the difference between these two kinds of expression is not statistically significant. These results could be related with the nature of the experiment: even users with an artistic background did not expect the need to express such specific affective states and showed some difficulties in expressing them in a totally instinctive way. Additionally, some features (for instance area of touch) need more data for us to reach a conclusion, a result of the large variety of measures used. Nevertheless, we consider these results to support the hypothesis of an intuitive system for affective expression. And, in addition to users' satisfaction/confidence shown on Fig. 10, they suggest that it may be possible to create a universal instrument that everybody could play regardless of their background (and that the metrics we defined could be used to uniformly recognize the affective expression of this wide diversity of users).

In terms of affective states addressed, most of the subjects agreed with the range we chose but some of them made some suggestions, namely the need to add pain/suffering, boredom, delirium, self-esteem (and lack of it), love, desire, faith, contempt and nausea. Some subjects also pointed some affective states as "redundant", not claiming that the affective states are similar, but finding it very hard to differentiate their respective expressions on a canvas, namely hate/anger and shame/shyness.

However, regarding pattern recognition for proceeding with the model implementation, we have focused for now on four complementary affective states in terms of valence and arousal: sadness, anger, joy (related to happiness) and tenderness [151,152]. AffectiveWall is a work-in-progress and the final goal, of course, it will be addressing a wider range of affective states, but we considered as a good start a set of four states, one for each quadrant of Russel's Circumplex Model [29] (we do not addressed more states to have a equal distribution around the Circumplex Model). We chose these affective states due to the fact that all of them are covered in studies of the relation between emotions and colors/sound parameters, and also because all of them are basic affective states to be conveyed in musical performances [151,152]. Thus we found the most common movements for each affective state addressed, and came up with the results shown above which we later implemented in our system. It 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 a 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).

Results of other affective states (not used in the system implementation) are presented in Appendix C.

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 consider gestures that were performed by at least half of the group. In this manner, the affective model is appropriate to reality while also providing enough freedom and subjectivity to feed the interaction interpreted by the system: the affective ecosystem as interface. So, in the next sections, we shall focus on the architecture and implementation of the system, and how all its modules work.

4.4 Architecture

Regarding the requirements of modularity and extensibility, we divided the project into modules correlated with the different main tasks of AffectiveWall:

- Interaction Module (consisting on *gesturesReader* file in OF, and CCV software for tracking purposes) that updates an *events* list with the blobs on the surface (or simulated blobs when using TuioSimulator, a software that simulates the interaction without the actual multi-touch surface, using only the mouse for input);
- Interpretation Module (consisting on *expressionInterpreter* file in OF) that analyses the information on an *events* list and updates the current *affectiveStates* and *mood* vectors;
- Composition Module (consisting on *musicComposer* file in OF and *server* file in SC) that translates values from the current *affectiveStates* and *mood* vectors to sound parameters and send to all synths in SC;
- Painting Module (consisting on *paintingsGenerator* file in OF) that converts values from the current *affectiveStates* and *mood* vectors to colors, and apply them to the interaction and manipulation of a particle system to create visuals correlated to the users movements.

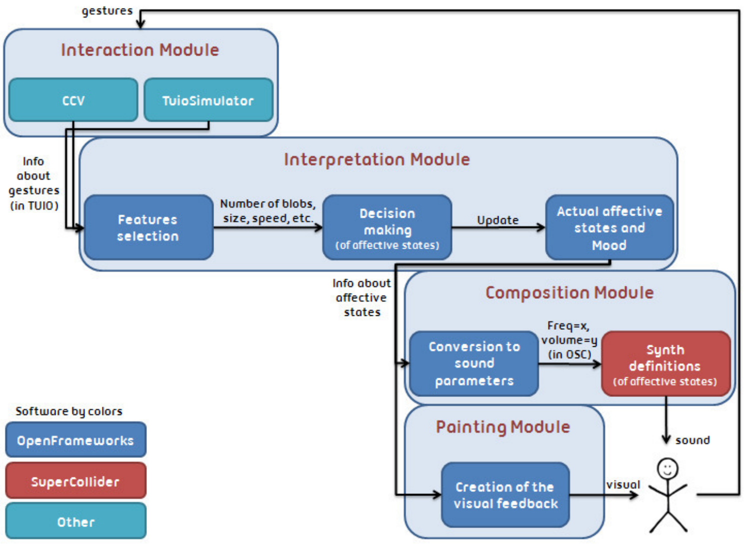


Fig. 12. General architecture of AffectiveWall.

In terms of the core of AffectiveWall, the code in OF, it is divided in different files as stated above, and in two distinct threads. One to perform just the creation of visual feedback[27] (that costs a large execution time) and the other one to all the other tasks (it is important to notice that real-time performance is one of the main requirements of this system).

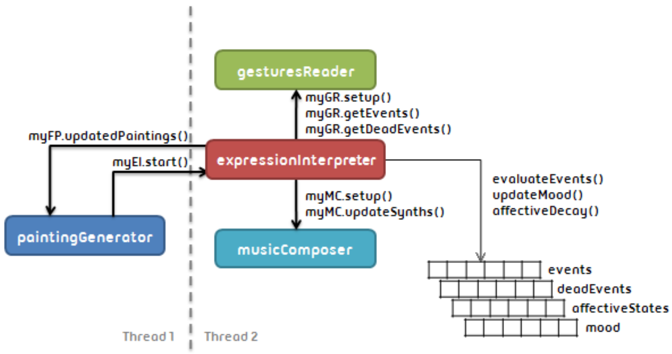


Fig. 13. Detailed architecture inside OpenFrameworks.

[27] Being this thread the one that uses OpenGL to create the visuals, it was better to maintain it as the initial thread.

Lastly, the communication between different modules is mainly supported by the four vectors that maintain all the information that circulates on the system:

— *events* (list with information about the blobs that come from CCV, being each event filled with information about blob's speed, position, direction, movement length, occupation on the canvas, etc.);
— *deadEvents* (list with basic information of events recently removed from the *events* list, namely removal time and movement length, that will be used later in feature analysis explained in section 4.5.3);
— *affectiveStates* (vector with the information of how much of each four affective states is being expressed in the movement, from a scale of 0 to 50);
— *mood* (vector with the same purpose of the last one but with a general perspective of the past affective states, from a scale of 0 to 1).

4.5 System layout and implementation

After defining the main architecture of the system, and the studies done to implement it, it is time to specify how it works and how it is implemented – from the moment the user touches the surface of the canvas until the appearance of visual and sound feedback.

4.5.1 User interaction - hardware

The hardware of AffectiveWall works by having a surface where the users perform their gestures, and behind it a camera is located that captures all movements and sends them to a computer that hosts the system. Later, it outputs visual contents that are projected again on the surface and sounds that are played due to a pair of speakers connected to the computer.

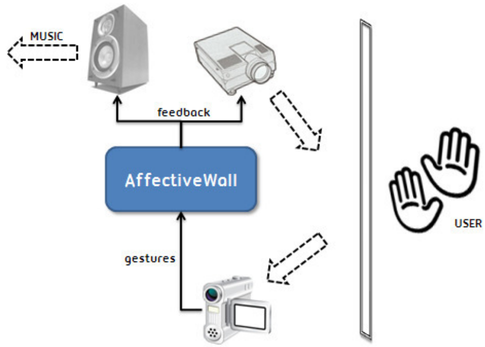


Fig. 14. Layout of AffectiveWall from the user viewpoint

In the process of planning and assembling the hardware of AffectiveWall, we had to take into account some concerns like simplicity, lowest time possible for assembling and testing, costs, etc. So, among other possible techniques such as DSI, DI or FTIR, we choose LLP (Laser Light Plane) for our setup [153]. This choice is mainly related to the low price of all the material involved (therefore fulfilling our goal of bringing Art closer to people), and also because it seemed an approach with reasonable good results of tracking for the different body parts that we want to addressed (for example we tested a DSI approach but we cannot manage to tweak the tracking parameters to have good results with fingers and hands simultaneously). So, LLP consists of an acrylic panel lined with a film paper to enable a high quality video-projection. Infrared lasers with line generators are placed on the corners to provide a thin layer of infrared light on the surface. Behind this, besides the video-projector, a video camera with an infrared filter is set to capture all the surface area. Thus, when the user touches the surface, infrared light reflects on the body and appears to the camera as a bright blob. The infrared filter on the camera has the function of limiting as much as possible all the color range except the one that is used by the laser (preventing uncontrollable light conditions of the environment), so in the computer we read the user interaction as shown in Fig. 16.

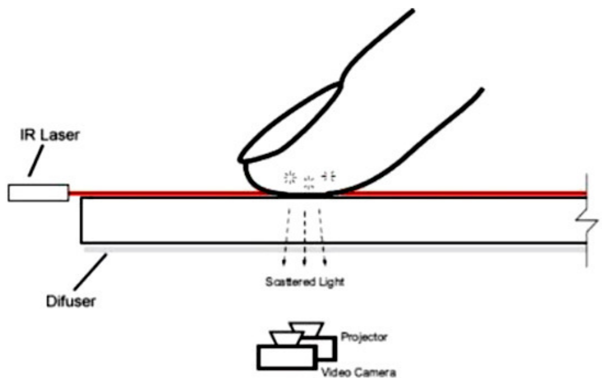


Fig. 15. Diagram of an LLP setup (Laser Light Plane) of a multi-touch surface.

The advantages of this setup are mainly its simplicity and its relatively low cost. Unlike other possible options, in LLP there is no need of a LED frame around the surface, a special kind of acrylic, a closed box behind the surface, silicone compliant surface or welding skills. It performs quite well what we are looking for: the capture of fingers, hands and other body parts touches (in other setups also objects and fiducials can be tracked but they are not part of the purposes of this project).

4.5.2 Events (gestures) reading and communication

So, in terms of hardware we have a camera prepared to capture infrared light reflected by user's body when touching the surface. In this manner, in terms of software we have CCV (Community Core Vision) which executes some image tweaks (like background subtraction, high pass filter, image threshold, etc.) and sends the information about the blobs to OF.

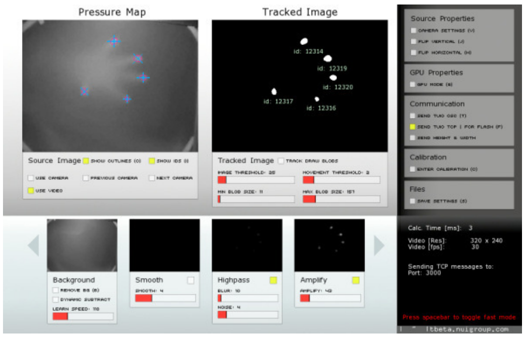


Fig. 16. CCV screenshot, which shows tracking of user's fingers.

This communication is done by a specific port to a specific IP and supported by Tuio protocol, developed by ReacTable team [143], which stands for an improved version of OSC [141,142] with some more information related to multi-touch events. Therefore, each event (i.e. each blob made by a user's touch) is a structure with the following data: session identifier (*sid*), blob identifier (*bid*), info about position (*xpos, ypos, xinit, yinit, lastx, lasty, flagHigh*[28], *flagMedium, flagLow*), speed (*xspeed, yspeed, motionspeed*), *area*[29], *length*[30], canvas occupation (*distance2source*[31]*, flagExpansive, flagMedOccup)* and movement direction (*xdir, ydir, flagUp*[32]*, flagDown, flagSide, flagUpSide, flagDownSide*).

In OF, reception of events' information is performed by function *getEvents* and *getDeadEvents* of *gesturesReader* file, which call namely *getBlobs* and *getDeadBlobs* (in the Tuio client) to update the *events* and *deadEvents* lists that will be later used to analyze gestures' features. Regarding the client of Tuio in our application, provided by *ofxTuioClient* of *ofxTuio* add-on, it stands for the class that manages all the Tuio communication, mainly by having a *receiver* that works as a listener of all the events' activities that came from CCV (namely additions, removals and updates of blobs).

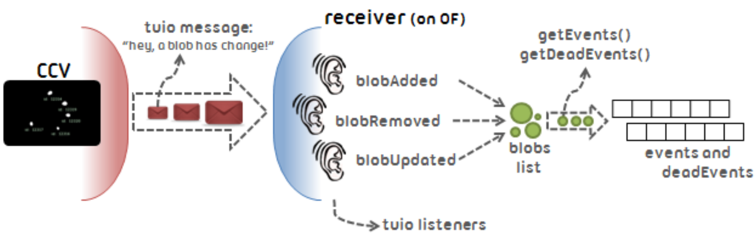


Fig. 17. Diagram of the whole process of reading and communicating events between CCV and *gesturesReading* on OF.

[28] These flags are triggered when user touches the matching zone on the canvas, having the three zones (high, medium and low height) equal sizes.
[29] In the end, we did not have the possibility of working with area information due to a problem with CCV that will only be resolved on the next version of the software.
[30] That is calculated in each evaluation by adding the length already calculated with the distance between the actual position and the last position (lastx, lasty).
[31] Used for compute the occupation, i.e., in each evaluation, depending on the distance between the actual position and the initial position, one of the occupation flags may be triggered.
[32] These flags are triggered depending on the calculation of motion angles in relation to the x and y axes.

Having described the whole communication process, we will now focus on the core of AffectiveWall, which controls all the phases of the application and connects all the modules.

4.5.3 Interpretation of the affective expression

Being the core of the whole system, interpretation of the affective expression is done on *expressionInterpreter* file and consists of a never ending cycle that performs the following tasks:

a) Update the *events* and *deadEvents*, the lists of current and recently finished events (by calling the functions *getEvents* and *getDeadEvents* of *gesturesReader* file);
b) Evaluate all the current events (updating the *affectiveStates* vector);
c) Update the *mood* vector using the data on *affectiveStates* vector;
d) Decay the values of current affective states (performed one time per second, to make all the non-expressed affective states on the present moment tend to zero);
e) Update the sound feedback (calling the function *updateSynths* of *musicComposer* file);
f) Update the visual feedback (calling the function *updatePaintings* of *paintingGenerator* file).

In terms of evaluation (step b in the above algorithm), it consists of a cycle covering all events five times per second and, for each one of them, analyzing its features (and measures), increasing or decreasing the affective states according to the gestures' patterns taken from the experiment of 4.3.2. This update is performed in each time a specific interval of measures was recognized on the gestures' analysis bearing in mind some important factors, for instance, the number of times each event will be evaluated along its life time. One example is the fact that the weight of a long gesture (evaluated many times) will be much higher than that of a short one (and correspondingly less evaluated), so the value of both updates needs to be aligned. The whole equalization process was performed by simulating the expected gestures for each affective state on TuioSimulator (a software for simulating multi-touch events with a common computer mouse), tuning the update values in order to have an uniform update of all states in the affective states vector (as summarized in Table 4).

Table 4. Update of affective states vector according to the features' evaluation of all events.

The values are the amount of increase or decrease for each occurrence of the feature on the user interaction.

Feature	**Measure (and values)[33]**	**Update Sadness**	**Update Anger**	**Update Joy**	**Update Tenderness**
Length	Long (>0.6)	+0.05	-0.7	+0.4	+0.5
	Medium (0.3-0.6)	+0.1	-0.025	+0.2	-0.025
	Short (0.05-0.3)	+0.1	+0.2	-0.1	+0.1
	Punctual[34] (<0.05)	-0.000005	+0.00001	-0.000001	-0.000005
Quantity	Many(>3)	-0.01	+0.2	+0.2	-0.01
	Some (2, 3)	-0.01	-0.01	-0.01	-0.01
	One (1)	+0.3	-0.2	-0.2	+0.3
Speed	Fast (>3.0)	-0.1	+0.2	+0.5	-0.1
	Medium (0.5-3.0)	+0.01	-0.1	+0.5	+0.01
	Slow (0.01-0.5)	+0.5	-0.7	-0.7	+0.4
	Static (<0.01)	-0.1	+0.5	-0.1	-0.1
Direction	Downward (abs(xdir)<0.1, abs(ydir)>0.0)	+0.5	+0.3	-0.7	+0.05
	Sidewards (abs(xdir)<0.1)	-0.2	-0.05	-0.2	+0.5
	Upward (abs(xdir)<0.1, abs(ydir)<0.0)	-0.7	-0.7	+1.0	-0.05
	Random (more than 3 directions)	-0.05	+0.1	+0.1	+0.1
	Other	-0.01	-0.01	-0.01	-0.01
Location	Low (<0.33)	+0.025	-0.025	-0.025	-0.025
	Medium (0.33-0.66)	+0.05	+0.1	+0.2	+0.2
	High (>0.66)	-0.2	-0.2	+1.0	-0.2
Occupation	Confined (distance2source<0.3)	+0.01	+0.01	-0.2	+0.01
	Medium (same but: 0.3-0.5)	+0.2	+0.025	+0.025	+0.2
	Expansive (same but: >0.5)	-0.025	+0.01	+0.5	-0.025

Is also worth noting that three features have a particular way of being evaluated. In terms of quantity, the number of events in a gesture does not only correspond to the current ones but also to the events that were finished and are recorded in *deadEvents* for a certain amount of time (e.g. a gesture for anger can be represented by consecutive punches on the surface and, in this case, the gesture is composed by all the events as a whole). In terms of area, the evaluation is done with the total area of all blobs at a certain moment, so the calculation of the amount to increase/decrease is only done after the cycle that covers the events vector[35]. Lastly, in the case of punctual gestures, the evaluation is done considering the events already finished and not the current ones. This is

[33] These values are the ones returned from CCV and calculated, among others, by the size of the screen.

[34] The amounts associated with punctual gestures are quite smaller comparing to the rest due to being processed in a different way (based on blobs recently removed and not on current blobs).

[35] As already stated, we did not have the possibility of working with area information due to a problem with CCV that will only be resolved in the next version of the software, but *expressionInterpreter* is already prepared to accept area info to perform the calculation.

done due to the fact that if we analyzed it on the current events we would mistake gestures of any length in the beginning by punctual gestures (causing, for instance, that affective state rage is always very active). This problem is minimized in the case of the other measures due to misjudgment being compensated along the gesture. Also, the use of this approach in the other measures will cause a strong delay on the response (calculation being done only when the user takes off his/her hand after a long gesture), a situation that does not occur in the case of punctual gestures because, for the affective states addressed, punctual and prolonged gestures were not noticed.

Regarding mood, its vector is updated when the affective states are updated, with the detail that, when being updated with the new values, past values gain extra weight with time (becoming harder to change it). Nevertheless, in the beginning this fact doesn't strangle the update (to prevent the mood from taking too much time to grow). The formula to do this is the following (being "*i*" each affective state):

$$\textit{New mood (i)} = \textit{minimum}(0.95,\ 0.5 \times \textit{elapsed time}) \times \textit{actual mood (i)} + 0.05 \times (\textit{actual affective state (i)} / \textit{total affectiveness})$$

Summarizing, the mood vector stands for a kind of history of the most expressed affective states. It also suffers decay, being the decay value variable depending on the arousal of each affective state (sadness and tenderness are less momentary than anger and joy so they take longer to vanish).

4.5.4 Affective composition of music

The musical composition can be divided in two parts: the synthesizers on SC and the control/conversion of sound parameters on OF. Regarding SC synths, it is a process that can be phased on three steps, being the first one the launch of a SC server to an address (with a port number associated) to be able to receive all the communication (with messages on OSC format).

Another phase is the synths definition, i.e., the creation of independent "modules" of code (called *SynthDef*'s) that receive some arguments, use them to control the production of sound and send the result to an output channel. In AffectiveWall this is done, for instance, receiving the mood value of a specific affective state and use it to control de volume of one synth, while at the same time it receives the current affective state and links to the frequency of a sound wave or the speed that a sound event is triggered. Our approach was to create various synths associated with the different affective states and one for the general rhythm. Implementation was done according to the conversion between affective states and sound parameters presented on Table 1 based and inspired on some processes of Cottle's book of Computer Music in SuperCollider [154] and SC140 Project[36].

[36] Available on: http://supercollider.sourceforge.net/sc140/ (accessed on 01-09-2011)

Once defined, synths need to be launched on the Server, each one with a different node associated that will be used by OF to send the sound parameters, as in the example below:

$$\underbrace{\textit{s}}_{\text{server name}}\underbrace{\textit{.sendMsg(}}_{\text{creation and sending of an OSC message}}\underbrace{\textit{“/s\_new”}}_{\text{instantiation of a new SynthDef}}\textit{, }\underbrace{\textit{“electric\_convolution”}}_{\text{synth name (already defined on the server)}}\textit{, }\underbrace{\textit{1003);}}_{\text{unique node number to be assigned to this synth}}$$

In relation to the control and conversion of sound parameters on OF, this is mainly performed by *updateSynths* function which, by receiving the *affectiveStates* and *mood* vectors, covers all synths and their parameters calling the function *sendToSc*. This function converts the numeric values in the vectors to OSC messages, associating them to synth arguments and sending to the address/port of the server and specific node of each synth, as shown in the example below:

$$\textit{sendToSc(}\underbrace{\textit{1003}}\textit{, }\underbrace{\textit{“volume”}}\textit{, }\underbrace{\textit{mood[1]}}\textit{);} \qquad \text{|| on OF}$$

$$\underbrace{\textit{s}}_{\text{server name}}\underbrace{\textit{.sendMsg(}}_{\text{creation and sending of an OSC message}}\underbrace{\textit{“/n\_set”}}_{\text{command for setting arguments}}\textit{, }\underbrace{\textit{1003}}_{\text{synth node}}\textit{, }\underbrace{\textit{“volume”}}_{\text{argument to change}}\textit{, }\underbrace{\textit{0.5}}_{\text{new value}}\textit{);} \qquad \text{|| on SC}$$

The conversion of the values’ range is performed on the SynthDef (the mood and current affective state values are always between 0.0 and, respectively, 1.0 and 50.0, being the values in SC of various different ranges). Is also important to note that different types of updates are used, depending on the situation:

- Simple: when, for instance, the mood of a specific affective state is directly correlated to an argument of a synth of that affective state;
- Double: when using values from two different affective states (for instance their addition or the bigger value from the two);
- Inverted: when values from some affective states are inversely proportional to an argument of a synth of another affective state.

Explained the whole process of creating and updating the synths, the list of all synths created for AffectiveWall (and the correspondent arguments) is shown on Table 5 being their source code on Appendix A.

Table 5. List of all synths and corresponding arguments.

Synth name	Affective state associated	Arguments	Parameters associated
Slow Sadness	Sadness	Sound events rate Volume	*as[0]+as[1]+as[2]+as[3]* *mood[0]*
Sad Circus	Sadness	Freq. change speed Volume	*2*(as[1]+as[2])/50* *mood[0]*
Indian Void Trip	Sadness	Filter openness Volume	*2*mood[3]* *mood[0]*
Radiohead Drone Wannabe	Sadness	Sound events rate Volume	*as[0]* *mood[0]*
Electric Convolution	Anger	Sound distortion Sound roughness Volume	*mood[1]-((mood[0]+ mood[2]+mood[3])/3)* *as[1]* *mood[1]*
Acid Straw	Joy	Sound pitch Sound variation Volume	*mood[2]*50* *affectiveStates[2]/50* *mood[2]-(mood[0]+2*mood[3])*
Gentle Drops On Your Head	Joy	Pitch Sound events rate Volume	*as[2]/30* *5*(mood[0]+mood[1]+mood[3])* *mood[2]-(mood[0]+2*mood[1])*
Dreamy Bells	Tenderness	Sound events rate Pitch Volume	*(as[1]+as[2])/2* *as[3]* *mood[3]*
Nostalgic Wall of Sound	Tenderness	Sound softness Volume	*as [3]* *mood[3]*(1 - 0.0067 * (as[1]+as[2]))*
Stress Rhythm	Anger	Sound variation Volume	*mood[1]* *2*(mood[1]-mood[2])*
Positive Rhythm	Joy	Speed Pitch Volume	*mood[2]* *as[2]* *2*(mood[2]-mood[1])*
Background Beat	All	Speed Sound openness Volume	*max(mood[1], mood[2])* *(as[2] + as[3])/2* *(mood[0] + mood[1] + mood[2] + mood[3])/4*

The chosen parameters associated and their tuning were done through a large amount of experimentations to reach a solution that:

— Respects the aesthetical concerns typically associated with experimental electronic music;
— Shows an innovative approach to the affective states addressed;
— Does not result in a huge amount of noise with the levels of all synths overlapping each other;
— Does not produce unintentional glitches or noises, providing an overall harmony between all sounds.

Just a final remark about the kind of composition that AffectiveWall delivers to users: a collaborative experience between our work (which can evolve and reinvent itself through time) and the actual performative act of using the instrument. Music composition can be seen as the product of two distinct but complementary processes: the design of individual sounds (a micro-compositional level), which in our case is conducted by us when we programming the synthesizers; and the arrangement of the designed sounds into a musical score (a macro-compositional level), done in real time by users when they interpret the system's sounds and interact accordingly to that.

4.5.5 Affective generation of paintings

The generation of paintings is performed by using a particle system that calculates the movements of the "digital paint". We use MSAFluid[37], a library for solving and displaying real-time fluid simulations based on Navier-Stokes equations and Stam's work [155]. So, our *paintingGenerator* interacts with the *fluidSolver* from MSAFluid by giving data about the gestures and, after the calculation of fluid dynamics, we use the information returned by fluidSolver to draw the visual output of AffectiveWall. To support this, *fluidSolver* maintains a structure with information about the color and the force for each position of the visualization area. In this manner, the first step is done by the *updatePaintings* function (on *paintingGenerator* file and called by *expressionInterpreter*) and the *addToFluid* function (also on *paintingGenerator*) that decide what colors represent the affective states expressed by each gesture. The list of *events* and *affectiveStates/mood* vectors are received from *expressionInterpreter*, and a loop is executed in which, for each event, color and force for each position affected (x,y) are calculated according to the following aspects:

— Color decision was inspired by a study by Bresin [81] that presents the hues for various affective states depending on how they can help the expression of musical performances: sadness is violet (hue = 0.75), anger is red (hue = 0), joy is yellow (hue = 0.167) and tenderness is also violet (hue = 0.75). Saturation and luminance are managed through the fluid dynamics and the forces calculated for each position. Therefore, an update of a RGB structure is performed, where all the affective states influence the Red component, only joy influences the Green component, and sadness and tenderness together influence the Blue component.

— Force is equivalent to the velocity that gestures were performed (in order to visual content follow the user's touch) and the entropy given by the affective states represented. This entropy is estimated by the arousal of the gesture (i.e. the visual shape will be more twisted depending on how large are the current anger and joy values, and more softer depending on how large are the current sadness and tenderness values), according to the formula below:

$$entropy = 0.001 + max(1.0,$$
$$anger\ value + joy\ value - sadness\ value - tenderness\ value)$$

[37] Available on: http://www.msavisuals.com/msafluid (accessed on 01-09-2011)

After this, data is inserted on the fluidSolver structure by the functions *addColorAtIndex* and *addForceAtIndex*, enabling the calculation of fluid dynamics. The second step is drawing the visual output through this information after the fluid computation. This is performed by the *draw* function (on *paintingGenerator* file) that calls the *drawMotion* function (provided by MSAFluid lib). This function runs a cycle with all positions of the visualization area and, for each one, uses the *getInfoAtCell* function (that returns the calculated data for each position) and updates a OpenGL texture that will be drawn and after projected on the canvas.

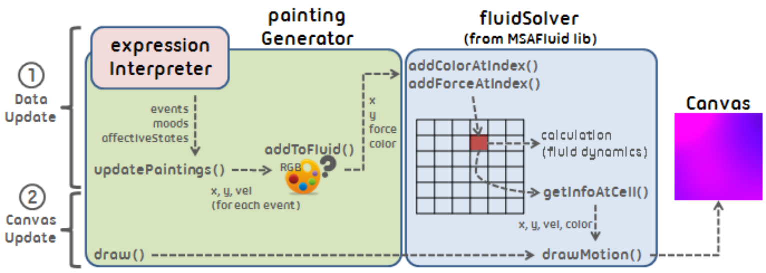


Fig. 18. Diagram of the whole process of paintings generation, divided by data update on fluidSolver and the use of that information to draw the paintings on the canvas.

5 Evaluation

5.1 Augmented affective states as output, an evaluation of recognition

5.1.1 Hypothesis

Have finished the first version of AffectiveWall (as we believe the whole system is a work-in-progress), we conducted an experiment to understand the current status of expression and recognition of affective states. Despite that, our concerns were not to simply evaluate how good the system is at converting gestures to music and sound so that an audience can perceive all the affective states expressed. To achieve that, another approach should be followed, where many people would performed the expression on the system, testing the effectiveness of the system's feedback by how many spectators recognized each expression. This leads to a very time consuming test, because a large number of participants expressing themselves in the system would be needed, and all the videos recorded from these expressions would be viewed by many people as well, in order to avoid that some outliers in terms of gesture's features used would disturb the whole experiment's outcome. So, due to the complexity of such a test, and regarding our initial motivations for this work, our approach was to try answer the following questions:

- To what extent does the performer influence the recognition of expressed affective states by an audience?
- Is it possible to have an instrument with a universal output that an audience, regardless of their age, gender or kind of background, can perceive the performer's expression?
- To what extent does their confidence of recognition vary when answering, depending on their having an artistic background or not?
- Are there some affective states easier to recognize than others?
- How different dimensions of affective expression, such as arousal or valence, are perceived by an audience?

5.1.2 Apparatus of the experiment

To answer the questions defined above, we divided our experiment in two phases: the expression and the recognition. The first one consists in recording the expression of performers (using the system) on video, and the other one consists in showing these videos to other participants. In the first stage, we mounted the AffectiveWall on a quiet room (to allow the maximum concentration of users) and a video camera in order to capture all the AffectiveWall outputs: music, sound and physical gestures of the subject (as stated on section 1.2 about the "augmented affective states as output" concept). In terms of the second part of the experiment – the recognition – it consists on an online questionnaire distributed by email and social networks, in order to reach a wider audience as much as possible.

5.1.3 Procedure and design

For the first stage of the experiment, the recording of expressions in video, the procedure was, after briefly explaining the context of the test and the whole project, ask the subject to use the system to express four affective states: anger, sadness, joy and tenderness, one at each time. To remove any bias, affective states were asked in a random and different order for each subject, and tests were performed in isolation without the user having previously seen other subject's expression.

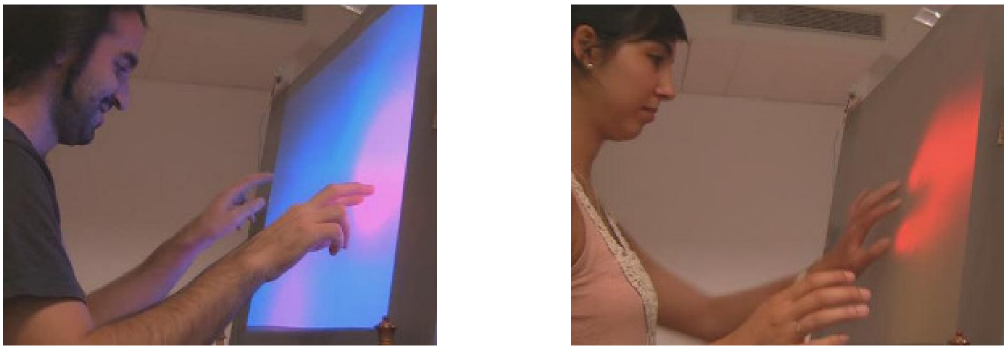


Fig. 19. Subjects using the system to express affective states while being recorded on video to be later showed to other participants for recognition.

In the end of the tests, we got sixteen videos of the four subjects performing the four affective states, so the second phase of the experiment was to show them to spectators through online questionnaires. In this questionnaire we requested the following data:

- Gender, age and country (where the subject spent most of his/her life);
- Kind of background (non-artistic, paintings/drawings, music, theatre/performance or other);
- For each video, which affective state is being expressed (from one of the following possible answers: anger, disgust, fear, joy, sadness, surprise, tenderness or don't know);
- The confidence in the given answer;
- Comments, suggestions, etc. (optional).

Again, to remove any bias or restriction to the participant's answer, in the "don't know" answer we gave the opportunity to participants to freely write the affective state that best suits the video. In order to enlarge the diversity of the test group, this test was performed by people in a wide range of ages, countries and backgrounds (being an example of the questionnaire presented in Appendix B).

5.1.4 Participants

The first part of the evaluation, the one about expression, was conducted with four subjects, two males and two females, two of them having artistic background and the other two not, and with ages around 26. In terms of the second part of the experiment, the online questionnaire, we had a participation of 84 subjects from many different countries (namely Portugal, Spain, Brazil, United Kingdom, United States, Canada, Colombia, Romania, Germany, Italy and Thailand), and distributed over the different variables of gender, age and background, as presented on Fig. 20.

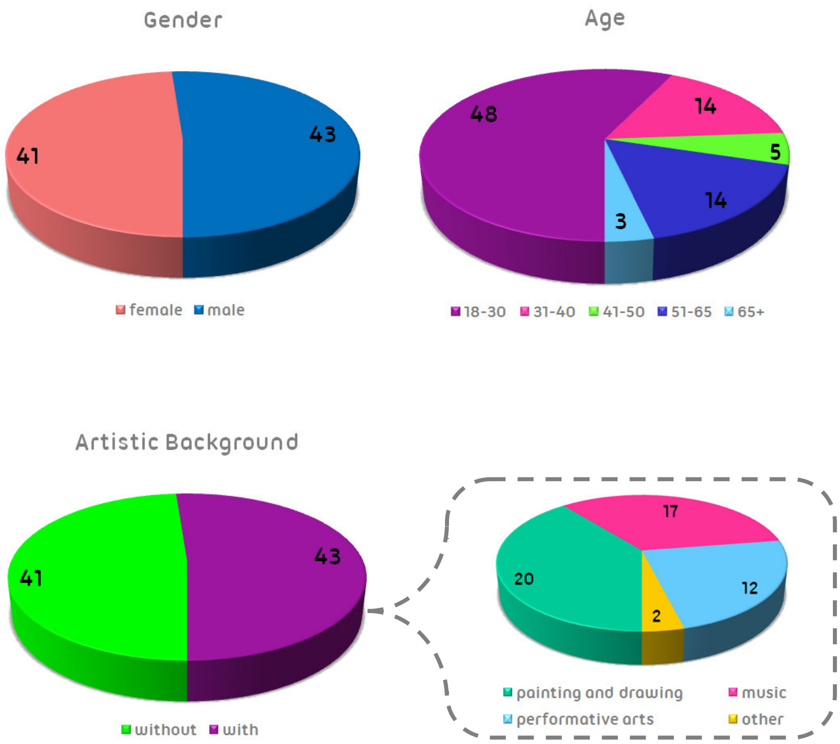


Fig. 20. Information about the subjects' profile (note that, in the case of artistic background, one can have more than one different kind of background).

5.1.5 Analysis method

What we extracted from the test was mainly the number of recognitions of the affective state expressed on each video. In addition, we also measured if the viewer, when failing to recognize the exact affective state, at least answered another state from the same quadrant from the Russel Circumplex Model [29] (i.e. same valence and arousal as the exact affective state), or answered a state that matches the arousal or the valence. Thus, we compared these results through non-parametric tests (such as Wilcoxon or Friedman) between the different metrics that we knew about the participants – age, gender, background – and also between the performer that are expressing in the video and the kind of affective states expressed.

5.2 Results and discussion

The first results we achieved were about the differences of recognition rates between the different kinds of participants that we had. We present the graphical results on Fig. 21 and Fig. 22, and the non-parametric tests comparing the results from various groups on Table 6 and Table 7.

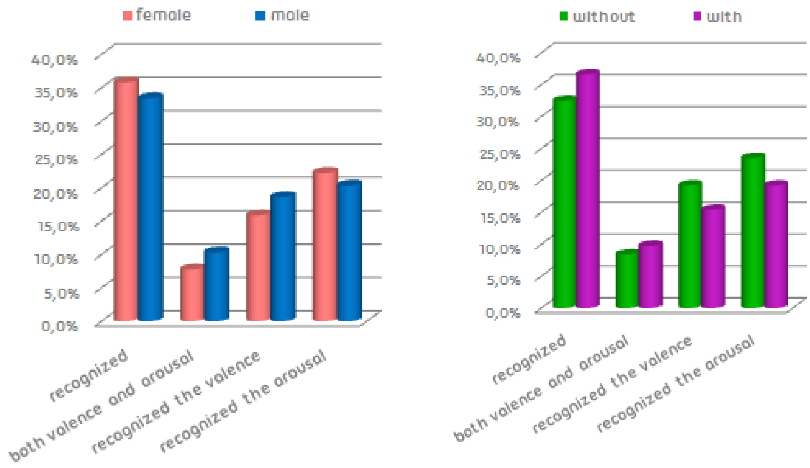


Fig. 21. Expression recognition rate by (a) gender and (b) kind of background of the spectator (note that the bars related to recognition of valence and arousal correspond only to situations where participants do not recognized successfully the affective state).

Table 6. Difference in the recognition rate between participants with and without artistic background (with *P* = 0,05).

	Recognized
Mann-Whitney U	216280,000
Wilcoxon W	431776,000
Z	-1,601
Asymp. Sig. (2-tailed)	**,109**

Table 7. Difference in the recognition rate between participants from the both genders (with *P* = 0,05).

	Recognized
Mann-Whitney U	220264,000
Wilcoxon W	457280,000
Z	-,921
Asymp. Sig. (2-tailed)	**,357**

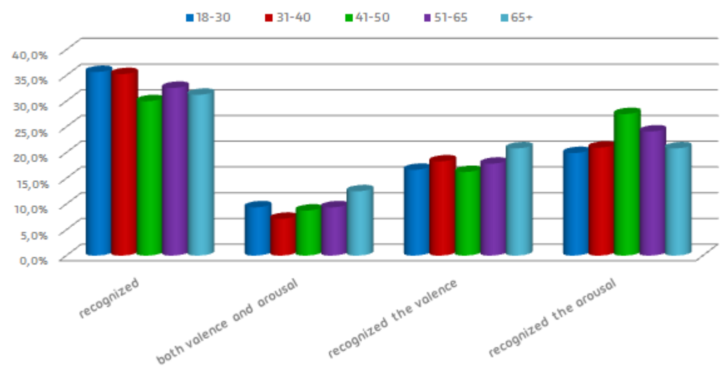


Fig. 22. Expression recognition rate by age of the spectator (note that the bars related to recognition of valence and arousal correspond only to situations where participants do not recognize successfully the affective state).

As we can see from the variety of participants and the results shown above, the output of AffectiveWall can be considered as universal. I.e. regardless of gender, age and, mainly, the existence or not of an artistic background, an audience can understand the system outcome in a statistically similar way. Also the confidence that participants felt when giving an answer for each video did not show a statistically significant difference between people with and without an artistic background (as presented on Table 8).

Table 8. Difference of confidence rates from participants with and without artistic background.

	Confidence
Mann-Whitney U	222488,500
Wilcoxon W	437984,500
Z	-,460
Asymp. Sig. (2-tailed)	**,646**

It is also interesting to assess whether the performer had influence on the overall recognition of the affective states expressed, being the results presented on Fig. 23 and Table 9.

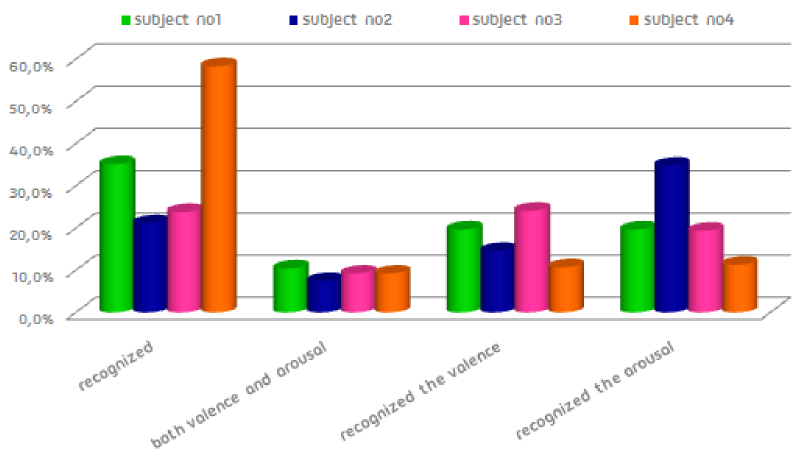


Fig. 23. Expression recognition rate by performer on the video (note that the bars related to recognition of valence and arousal correspond only to situations where participants do not recognized successfully the affective state).

In the process of testing whether the differences of the recognition rate between the expressions of the four performers are statistically significant, i.e. how much the person who uses the system influences the overall recognition of an audience, we ran the non-parametric Friedman test. As a result we get a $\chi^2(3) = 83{,}654$ with a significance level of $P = 0{,}05$. As there was overall statistically significant difference, we had to perform a *post-hoc* analysis with a Wilcoxon signed-rank test to discover in which groups this difference happened. This test evaluates the possible combinations between the videos of the four performers, as we can see on Table 9.

Table 9. Wilcoxon signed-rank test for the recognition between different combinations of performers expressing the affective states.

	Recognized SubjectNo2 - Recognized SubjectNo1	Recognized SubjectNo3 - Recognized SubjectNo1	Recognized SubjectNo4 - Recognized SubjectNo1	Recognized SubjectNo3 - Recognized SubjectNo2	Recognized SubjectNo4 - Recognized SubjectNo2	Recognized SubjectNo4 - Recognized SubjectNo3
Z	-3,419	-2,701	-4,402	-,732	-7,650	-6,982
Asymp. Sig. (2-tailed)	**,001**	**,007**	**,000**	**,464**	**,000**	**,000**

In this case it is necessary to use a Bonferroni adjustment [156] on the results (due to multiple comparisons being performed): *significance level (P = 0,05) / number of tests ran.*

The Bonferroni correction is a safeguard against multiple tests of statistical significance on the same data falsely giving the appearance of significance, as 1 out of every 20 hypothesis-tests will appear to be significant at the $P = 0.05$ level purely due to chance. Therefore, from the formula above the new significance level is $P = 0{,}0083$, resulting that in almost all cases (the exception occurred when comparing subject number two and three) the performer had an huge influence on how the audience perceives and decodes the affective states addressed. As we can see in Fig. 23, the recognition of each subject performance goes from almost twenty to sixty percent of successful recognitions. In fact, what we did notice was a deep disruption between the gestures' features that we are expecting from our affective model (which were an average of the behaviors of fifteen people) and what we notice in this experiment. This explains the clear difference between the performance of subject no.4 and the other ones. So, a more extensive evaluation should be conducted in order to statistically prove this statement, including the viewing and annotation of features of these sixteen videos to later compute the statistical difference between them and the results from the evaluation conducted for the affective model definition. This could be a seminal future work, due to it being a too overwhelming task to perform at this stage of our work. Besides, we consider that the results obtained this time where not unsatisfying at all considering these conditions, suggesting us that it will be possible to create a system where the "augmented affective states" really help an audience to better understand the performer's expression even if a wide variety of expressive cues (that the system was not expecting) were used.

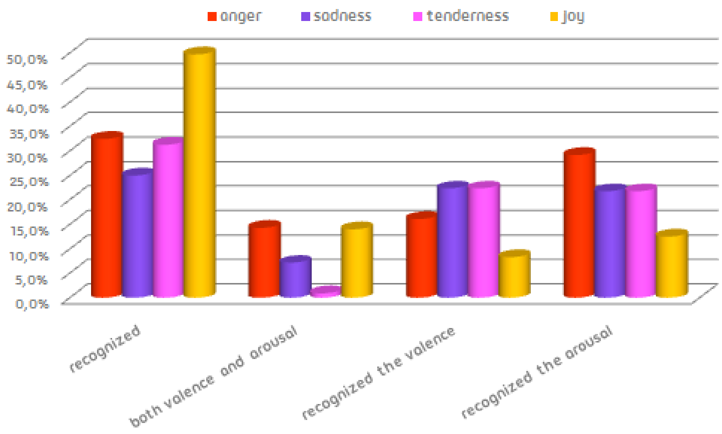


Fig. 24. Expression recognition rate by affective state expressed.

As in the case above, in testing the recognition rate between the four affective states addressed, we ran the non-parametric test Friedman test. As a result we get a $\chi^2(3) = 32{,}643$ with a significance level of $P = 0{,}05$. Again, due to the overall statistically significant difference, we performed a *post-hoc* analysis with a Wilcoxon signed-rank test to discover in which groups this difference happened. This test evaluates the possible combinations between the videos of the four affective states, as we can see on Table 10.

Table 10. Wilcoxon signed-rank test for the recognition between different combinations of the affective states expressed.

	Sadness - Anger	Joy - Anger	Tenderness - Anger	Joy - Sadness	Tenderness - Sadness	Tenderness - Joy
Z	-1,800	-3,491	-,273	-5,239	-1,528	-3,759
Asymp. Sig. (2-tailed)	**,072**	**,000**	**,785**	**,000**	**,127**	**,000**

We again used the Bonferroni adjustment [156] and came up with a significance level of $P = 0{,}0083$, resulting that, together with the values shown on Fig. 24, Joy was in fact the most perceived affective state in this evaluation.

Some conclusions about the recognition in general can be drawn, namely the fact that the arousal of affective states is the most perceived characteristic perceived by an audience (as shown on Fig. 25).

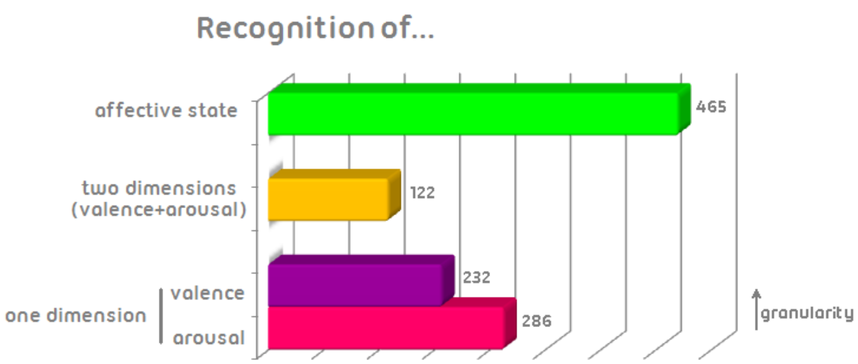


Fig. 25. Number and type of recognitions (in a total of 1344 visualizations).

6 Conclusions

"Music is what feelings sound like."
Anonymous author

This dissertation aims to make this statement more literal and proposes a system that materializes it. Thus, the AffectiveWall, an intermedia instrument where music is created without the paradigm of typical instruments like vibrating a string or pressing a key, and painting does not need a brush to come alive. Instead, all Art created results from the emotional expression using body language, opening artistic creation to all people regardless of their background.

We began by presenting the project and enumerating all the goals, motivations, contributions and application areas. Then, a study about the state of the art was conducted, including music and painting (presenting the artistic trends that had an influence on the creative process of AffectiveWall), emotions and affective states (classification, connection with gestural expressivity, etc.) and the way all these areas can link together. A research of the technologies was also carried out, regarding interactive music systems, tangible interfaces, affective computing, etc. Thereafter, we defined a solution to our system, which started with user tests to find patterns from gestures to each one of the affective states addressed. Then, with an affective model already decided, we move forward to an architecture of the system and its implementation. In this section, we explained step by step the whole procedure, from the moment when the performer touches the canvas, through the gestures' evaluation and decision making of affective expression, to the moment when the system produces the sound and visual feedback. After this, we introduced an evaluation methodology regarding the expression and recognition of performers' expressivity, describing the whole experiment and its results and discussions. It was showed that the system output is universal for any person regardless the respective background and profile info, and even in terms of confidence when answering the questionnaire, no significant difference was perceived between participants with and without artistic background. We also found that "augmented affective states" can help an audience to better perceive the affective expression even if a wide variety of expressive cues were used (we found huge differences between the performance of the four subjects in this evaluation and the results from our previous test that outcomes our model, although a more extensive evaluation should be performed to statistically prove this fact). These great differences (noticed on the behavior of subjects when expressing themselves on AffectiveWall) leading us to conclude that a strong relationship and commitment between the performer and the system must be built. AffectiveWall can help people to express themselves but asks for a focused and sincere performance on their part in order to fulfill the supposed requirements. Regarding the quality of recognition depending on the affective states expressed, we conclude that joy as the easiest to recognize on this kind of interaction, and also that arousal is the most perceived feature of the affective expression.

In the end, we can take from this work three main concepts: the “affective model for gestural expressivity”, after all, the framework that supports AffectiveWall by mapping gestures into affective states; the “affective ecosystem as interface” concept, in which the system has the main role in the interaction and the user can indirectly change the input by providing actions to the environment; and the “augmented affective states as the output”, a multimodal stimulus that is perceived by the audience as a composition of gestures, audio and video.

6.1 Some reflections

This kind of instrument, primarily the “affective ecosystem as interface” concept, raises some interesting questions about human expression and its relation with technology. In pursuit of the goal of bringing Art closer to people, rejecting rules and artistic trends, this project takes two totally different and paradoxical paths:

— On one hand, it adapts artistic creation to gestural and emotional expression, two intuitive and natural languages for everyone, making unnecessary the learning period normal of any instrument or the knowledge about musical notations, for example. Humans emancipate themselves from theory and even from the material (as much as the brush, in the case of painting, as the physicality of the instrument which everyone must adapt to, in the case of music), towards a way of creation through their own body (both interior, by emotions and affective states, and exterior, by the corresponding body language). This replacement of the artistic object for the performer’s body happens also in terms of the painting outcome, due to the fact that when the performer is not touching the canvas, the whole painting begins to fade out gradually, without recording the result of creative process on a timeless medium like a painting on a conventional canvas (therefore approaching the paradigm of real-time performance, like live music performances, where, without external media (e.g. audio recorder), the performance only records itself on the memory of the ones who saw it, being the memory a kind of medium that also suffers from a gradual decay);

— On the other hand, it adds a new element to the creative process: a kind of alive canvas. Whereas before the artist thought that he was in possession of the necessary free will to create something (an invalid presupposition because he is self-conditioned by his past life, theories, and artistic trends; being his expression always non-innocent and totally independent from external stimuli), with the AffectiveWall his “freedom” is conditioned by the system, the technology. This “technological entity”, who regulates herself by an own definition of human expression (the average of the behavior of all users that performed the user evaluation). From this arises the question: in which situation is the artist freer? When he leaves the final decision-making to the system (which mechanically analyzes its expression according to the "human average"), or when he gives his “freedom” to the social experience as an artist and human being who is part of a society?

6.2 Future work

In terms of future work, it would be rather important to evaluate how the quality of expression and the confidence of the performer change with the system (comparing to the kind of expression done in the first experiment, related in section 4.3). Also, to test the recognition rate in a correct manner, a longer evaluation and with a larger number of participants interacting with the system could be conducted (as said in chapter 5 about evaluation). In addition, many developments can be made to the visual feedback, including new kinds of shapes and movement effects that better express the affective states addressed, and to the sound feedback, adding more synthesizers and extending it to other kinds of music. Also in terms of gesture reading, other input methods and new features can be incorporated, like sensitivity/power of the touch (e.g. using piezos to capture the sound made by performer when hitting the surface), analysis of blob's shape, measurement of the time interval between two punctual hits on the surface, calculation of area and speed changes along a touch or an additional video camera for a better understanding of some gestures with large amounts of information happening outside the surface or of movements that can be misunderstood in terms of the body part used to touch the surface (an approach more similar to the work of Camurri et al. [3] and Castellano et al. [33]).

In a different approach, it would be interesting to perform tests with children and patients with mental disabilities to study their reaction to the AffectiveWall way of interacting and outputs produced. Adapting the affective model to the specificities of each possible user group, positive outcomes can emerge from the use of the system in a therapeutical context, mainly due to its direct contact with the music and visual creation (joining together two seminal fields in this area: therapy with music and therapy with paintings).

6.3 A final remark

So we talked about emotions and affective states, and the way we can expressed them, especially linking different artistic fields in a creative process that involves a physical medium: a wall or canvas. That is why, as in the paper which resulted from this dissertation, we chose "Another Thrill in the Wall" as the *leitmotif* for our work, in which the word "thrill" refers to the physical manifestation of an emotion through the whole body. We want to continue to explore this kind of intermedia instrument, using the vast and growing world of multi-touch interfaces. Jefferson Han[38] said "Multi-touch-sensing was designed to allow non-techies to do masterful things while allowing power users to be even more virtuosistic". We want to subvert his words and apply them 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.

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

Bibliography

[1] A Di Scipio, "Sound is the interface. Sketches of a Constructivistic Ecosystemic view of Interactive Signal Processing," , 2003.

[2] P N Juslin, "Emotional communication in music performance: A functionalist perspective and some data," *Music Perception*, pp. 383--418, 1997.

[3] A Camurri, B Mazzarino, M Ricchetti, R Timmers, and G Volpe, "Multimodal analysis of expressive gesture in music and dance performances," *Gesture-based communication in human-computer interaction*, pp. 357--358, 2004.

[4] D Albright, *Untwisting the serpent: modernism in music, literature, and other arts.*: University of Chicago Press, 2000.

[5] C Dahlhaus, *Nineteenth-century music.*: Univ of California Pr, 1991, vol. 5.

[6] H W Janson and D J Janson, *History of art: a survey of the major visual arts from the dawn of history to the present day.*: HN Abrams, 1962.

[7] P Selz, *German expressionist painting.*: Univ of California Pr, 1974.

[8] A Pintado, Sense2: A music system based on paintings, 2009, Master Thesis. Instituto Superior Técnico - Universidade Técnica de Lisboa.

[9] H Higgins, *Fluxus experience.*: Univ of California Pr, 2002.

[10] T Wilfred, "Light and the Artist," *The Journal of Aesthetics and Art Criticism*, vol. 5, pp. 247--255, 1947.

[11] K Peacock, "Instruments to Perform Color-Music: Two Centuries of Technological Exploration. Leonardo," *Journal of the International Society for The Arts, Sciences and Technology*, vol. 21, pp. 397--406, 1988.

[12] K R Scherer, "What are emotions? And how can they be measured?," *Social Science Information*, vol. 44, p. 695, 2005.

[13] P M Niedenthal, L W Barsalou, P Winkielman, S Krauth-Gruber, and F Ric, "Embodiment in attitudes, social perception, and emotion," *Personality and Social Psychology Review*, vol. 9, p. 184, 2005.

[14] Paul R Kleinginna and Anne M Kleinginna, "A categorized list of motivation definitions, with a suggestion for a consensual definition," *Motivation and Emotion*, vol. 5, no. 3, pp. 263-291, 1981.

[15] V Tran, "The influence of emotions on decision-making processes in management teams," *Unpublished doctoral dissertation, University of Geneva, Switzerland*, 2004.

[16] N H Frijda, "Varieties of affect: Emotions and episodes, moods, and sentiments.," *The nature of emotions: Fundamental questions*, pp. 197--202, 1994.

[17] C E Izard, *The psychology of emotions.*: Springer Us, 1991.

[18] I J Roseman, "Cognitive determinants of emotion: A structural theory.," *Review of Personality & Social Psychology*, 1984.

[19] K R Scherer,.: CA: Sage, 1984, pp. 37-63.

[20] K R Scherer, "Toward a dynamic theory of emotion: The component process model of affective states," *Geneva Studies in Emotion and Communication*, vol. 1, pp. 1-98, 1987.

[21] K R Scherer, "Appraisal considered as a process of multilevel sequential checking," *Appraisal processes in emotion: Theory, methods, research*, vol. 92, p. 120, 2001.

[22] R W Picard, *Affective computing.*: The MIT press, 2000.

[23] S S Tomkins, "Affect, imagery, consciousness: Vol. I. The positive affects.," 1962.

[24] R Plutchik and H Kellerman, *Emotion, Theory, Research, and Experience: Emotion, psychopathology, and psycho-therapy.*: Academic press, 1980.

[25] P N Johnson-Laird and K Oatley, "The language of emotions: An analysis of a semantic field," *Cognition and emotion*, vol. 3, pp. 81--123, 1989.

[26] N Mirenkov, K Kanev, and H Takezawa, "Quality of Life Supporters Employing Music Therapy," , 2008, pp. 1051--1056.

[27] J Altarriba, D M Basnight, and T M Canary, "Emotion representation and perception across cultures," *Online Readings in Psychology and Culture*, vol. 4, p. 4, 2003.

[28] J A Russell, "Pancultural aspects of the human conceptual organization of emotions.," *Journal of Personality and Social Psychology*, vol. 45, p. 1281, 1983.

[29] J A Russell, "A circumplex model of affect.," *Journal of personality and social psychology*, vol. 39, p. 1161, 1980.

[30] A. Kendon, *Gesture: Visible action as utterance*.: Cambridge Univ Pr, 2004.

[31] M Kipp and J C Martin, "Gesture and Emotion: Can basic gestural form features discriminate emotions?," , 2009, pp. 1--8.

[32] P Fagerberg, A Stahl, and K Höök, "Designing gestures for affective input: an analysis of shape, effort and valence," *Proceedings of mobile ubiquitous and multimedia, MUM*, 2003.

[33] G Castellano, S Villalba, and A Camurri, "Recognising human emotions from body movement and gesture dynamics," *Affective computing and intelligent interaction*, pp. 71--82, 2007.

[34] H G Wallbott and K R Scherer, "Cues and channels in emotion recognition.," *Journal of personality and social psychology*, vol. 51, p. 690, 1986.

[35] C E Izard, "The face of emotion.," 1971.

[36] L Zhao and N I Badler, "Synthesis and acquisition of laban movement analysis qualitative parameters for communicative gestures," 2001.

[37] A Camurri, G Castellano, M Ricchetti, and G Volpe, "Subject interfaces: measuring bodily activation during an emotional experience of music," *Gesture in human-computer interaction and simulation*, pp. 268--279, 2006.

[38] M Lewis, J M Haviland-Jones, and L F Barrett, "Handbook of emotions," 2000.

[39] R G Collingwood, *The principles of art*.: Oxford University Press, USA, 1958, vol. 11.

[40] K Oatley, "Creative expression and communication of emotions in the visual and narrative arts.," 2003.

[41] R B Knapp and P R Cook, "The integral music controller: introducing a direct emotional interface to gestural control of sound synthesis," , 2005, pp. 4--9.

[42] R Steinberg, *Music and the mind machine*.: Springer, 1995.

[43] T Yamasaki, "Emotional communication through music performance played by young Children," *Proc. of ICMPC8*, pp. 204--206, 2004.

[44] GM Kotlyar and VP Morozov, "Acoustical correlates of the emotional content of vocalized speech," *Sov Phys Acoust*, vol. 22, pp. 208--11, 1976.

[45] A Gabrielsson and E Lindström, "Emotional expression in synthesizer and sentograph performance," *Psychomusicology: Music, Mind and Brain*, vol. 14, pp. 94--116, 2010.

[46] T Yamasaki, "Emotional communication in improvised performance by musically untrained players," , 2002, pp. 521--524.

[47] R Cruz, I-sounds: Emotion-based music composition for virtual environments, 2008, Master Thesis. Instituto Superior Técnico - Universidade Técnica de Lisboa.

[48] J A Sloboda and P N Juslin, "Psychological perspectives on music and emotion.," 2001.

[49] C E Williams and K N Stevens, "Emotions and speech: Some acoustical correlates," *The Journal of the Acoustical Society of America*, vol. 52, p. 1238, 1972.

[50] P N Juslin, "Cue utilization in communication of emotion in music performance: Relating performance to perception.," *Journal of Experimental Psychology: Human perception and performance*, vol. 26, p. 1797, 2000.

[51] C Palmer, "Music performance," *Annual Review of Psychology*, vol. 48, pp. 115--138, 1997.

[52] A Gabrielsson, "Expressive intention and performance," *Music and the mind machine*, pp. 35--47, 1995.

[53] L H Shaffer, "How to interpret music.," 1992.

[54] K R Scherer and J S Oshinsky, "Cue utilization in emotion attribution from auditory stimuli," *Motivation and Emotion*, vol. 1, pp. 331--346, 1977.

[55] J A Sloboda, "Individual differences in music performance," *Trends in Cognitive Sciences*, vol. 4, pp. 397--403, 2000.

[56] Bee Y Chua and Guojun Lu, "Perceptual rhythm determination of music signal for emotion-based classification," in *Multi-Media Modelling*, 2006, p. 8 pp.

[57] K R Scherer, "Expression of emotion in voice and music*," *Journal of Voice*, vol. 9, pp. 235--248, 1995.

[58] P Lieberman and S B Michaels, "Some aspects of fundamental frequency and envelope amplitudes as related to the emotional content of speech.," *Journal of the Acoustical Society of America*, 1962.

[59] H G Seashore, An objective analysis of artistic singing, 1932.

[60] R Hiraga, N Kato, and N Matsuda, "Effect of visual representation in recognizing emotion expressed in a musical performance," , pp. 131--136.

[61] M Senju and K Ohgushi, "How are the players ideas conveyed to the audience?," *Music Perception*, pp. 311--323, 1987.

[62] K Hevner, "Experimental studies of the elements of expression in music," *The American Journal of Psychology*, pp. 246--268, 1936.

[63] K Hevner, "Expression in music: a discussion of experimental studies and theories.," *Psychological Review*, vol. 42, p. 186, 1935.

[64] P R Farnsworth, "The social psychology of music.," 1958.

[65] N Moriguchi and M Miura, "Emotion control system for music performance: MOR2ART," , 2009, p. 68.

[66] R Bresin and A Friberg, "Synthesis and decoding of emotionally expressive music performance," , vol. 4, 1999, pp. 317--322.

[67] V Konecni, "Does music induce emotion? A theoretical and methodological analysis.," *Psychology of Aesthetics, Creativity, and the Arts*, vol. 2, p. 115, 2008.

[68] P Kivy, *Sound Sentiment: An Essay on the Musical Emotions, Including the Complete Text of The Gorded Shell.*: Temple University Press, 1989.

[69] K R Scherer and K R Zentner,.: Oxford University Press, 2001, pp. 361-392.

[70] F Collopy, "Color, form, and motion: Dimensions of a musical art of light," *Leonardo*, vol. 33, pp. 355--360, 2000.

[71] C Butler, *Early modernism: literature, music, and painting in Europe, 1900-1916.*: Oxford University Press, USA, 1994.

[72] A Stahl, P Sundström, and K Höök, "A foundation for emotional expressivity," , 2005, pp. 33--es.

[73] K Ryberg, "Levande färger," *ICA Bokförlag, Västeras* , 1991.

[74] L C Ou, M R Luo, A Woodcock, and A Wright, "A study of colour emotion and colour preference. Part I: Colour emotions for single colours," *Color Research & Application*, vol. 29, pp. 232--240, 2004.

[75] B Wright and L Rainwater, "The meanings of color," *Journal of General Psychology*, 1962.

[76] J Hogg, "A principal component analysis of semantic differential judgements of single colors and color pairs," *Journal of General Psychology*, vol. 80, 1969.

[77] J Hogg, S Goodman, T Porter, B Mikellides, and DE Preddy, "Dimensions and determinants of judgements of colour samples and a simulated interior space by architects and non-architects," *British Journal of Psychology*, vol. 70, pp. 231--242, 1979.

[78] S Kobayashi, "The aim and method of the color image scale," *Color Research & Application*, vol. 6, pp. 93--107, 1981.

[79] T Sato, K Kajiwara, H Hoshino, and T Nakamura, "Quantitative Evaluation and Categorisation of Human Emotion Induced by Colour," *Advances in Colour Science and Technology*, vol. 3, pp. 53--59, 2000.

[80] N Kaya and H H Epps, "Relationship between color and emotion: a study of college students," *College student journal*, vol. 38, pp. 396--405, 2004.

[81] R Bresin, "What is the color of that music performance," , 2005, pp. 367--370.

[82] R Hiraga, N Kato, and T Yamasaki, "Understanding emotion through drawings comparison between hearing-impaired people and people with normal hearing abilities," , vol. 1, pp. 103--108.

[83] R, Kato, N Hiraga, "First steps toward determining the role of visual information in music communication," *Ubiquitous Computing and Communication Journal*, 2010.

[84] T Yamasaki, "Emotional communication mediated by two different expression forms: Drawings and music performances," *Proc. of ICMPC*, pp. 153--154, 2006.

[85] M A Boden, "State of the art: Computer models of creativity," *PSYCHOLOGIST-LEICESTER-*, vol. 13, pp. 72--77, 2000.

[86] M A Boden, *The creative mind: Myths and mechanisms.*: Psychology Press, 2004.

[87] R J Sternberg and T I Lubart, *The concept of creativity: Prospects and paradigms.*: Cambridge University Press, 1999.

[88] V S Ramachandran and E M Hubbard, "Hearing colors, tasting shapes," *SCIENTIFIC AMERICAN-AMERICAN EDITION-*, vol. 288, pp. 52--59, 2003.

[89] K Kennedy, Painting music: Rhythm and movement in art, 2006, 20th Annual Sheldon Statewide Exhibition prospectus.

[90] D M Randel, *The Harvard concise dictionary of music and musicians.*: Belknap Pr, 1999.

[91] P Kivy, *Introduction to a Philosophy of Music.*: Oxford University Press, USA, 2002.

[92] T M Gifford and A R Brown, "The Ambidrum: Automated Rhythmic Improvisation," 2006.

[93] J Pressing, "Improvisation: methods and models," *John A. Sloboda (Hg.): Generative processes in music, Oxford*, pp. 129--178, 1988.

[94] A Mulder, "Virtual musical instruments: Accessing the sound synthesis universe as a performer," , 1994, pp. 243--250.

[95] R Rowe, "Incrementally improving interactive music systems," *Contemporary Music Review*, vol. 13, pp. 47--62, 1996.

[96] M Leman and A Camurri, "Musical content processing for expressive gesture applications in interactive multimedia," in *Conference on Interdisciplinary Musicology* , 2004.

[97] A Hunt and R Kirk, "Mapping strategies for musical performance," *Trends in Gestural Control of Music*, vol. 21, 2000.

[98] J B Rovan, M M Wanderley, S Dubnov, and P Depalle, "Instrumental gestural mapping strategies as expressivity determinants in computer music performance," , 1997, pp. 3--4.

[99] M Wanderley, "Gestural control of music," in *International Workshop Human Supervision and Control in Engineering and Music*, 2001.

[100] G Roma and A Xambó, "A tabletop waveform editor for live performance," *Proceesings of New Interfaces for Music Expression*, 2008.

[101] P Dourish, *Where the action is: the foundations of embodied interaction.*: The MIT Press, 2004.

[102] K P Fishkin, "A taxonomy for and analysis of tangible interfaces," *Personal and Ubiquitous Computing*, vol. 8, pp. 347--358, 2004.

[103] B Buxton, "The natural language of interaction: A perspective on non-verbal dialogues.," *INFOR.*, vol. 27, pp. 221--229, 1989.

[104] G Castellano, R Bresin, A Camurri, and G Volpe, "User-centered control of audio and visual expressive feedback by full-body movements," *Affective Computing and Intelligent Interaction*, pp. 501--510, 2007.

[105] A Camurri et al., "Toward real-time multimodal processing: EyesWeb 4.0," in *AISB 2004 Convention: Motion, Emotion and Cognition*, University of Leeds, 2004.

[106] A Camurri, M Ricchetti, and R Trocca, "EyesWeb-toward gesture and affect recognition in dance/music interactive systems," , vol. 1, 1999, pp. 643--648.

[107] M Waisvisz, "The hands, a set of remote midi-controllers," , 1985, pp. 313--318.

[108] T Marrin and R Picard, "Analysis of Affective Musical Expression with the Conductors Jacket," , 1998, pp. 61--64.

[109] B Buxton, "Multi-touch systems that i have known and loved," *Microsoft Research*, 2007.

[110] G Levin, "The table is the score: An augmented-reality interface for real-time, tangible, spectrographic performance," , 2006.

[111] J Han, "TEDTalks Jeff Han: Unveiling the genius of multi-touch interface design," *Retrieved May*, vol. 4, p. 2008, 2006.

[112] J Y Han, "Low-cost multi-touch sensing through frustrated total internal reflection," , 2005, pp. 115--118.

[113] M Bischof et al., "Xenakis: combining tangible interaction with probability-based musical composition," , 2008, pp. 121--124.

[114] J Patten, B Recht, and H Ishii, "Audiopad: a tag-based interface for musical performance," , 2002, pp. 1--6.

[115] S Jorda , G Geiger, M Alonso, and M Kaltenbrunner, "The reacTable: exploring the synergy between live music performance and tabletop tangible interfaces," , 2007, pp. 139--146.

[116] S Jorda, M Kaltenbrunner, G Geiger, and R Bencina, "The reactable," , 2005, pp. 579--582.

[117] J Rekimoto, "SmartSkin: an infrastructure for freehand manipulation on interactive surfaces," , 2002, pp. 113--120.

[118] J B Thiebaut, P G Healey, N B Kinns, and Q Mary, "Drawing Electroacoustic Music," , vol. 8, 2008.

[119] P G Healey and J B Thiebaut, "Sketching Musical Compositions," , vol. 7, 2007, pp. 1079--1084.

[120] T Coduys and G Ferry, "IanniX aesthetical/symbolic visualisations for hypermedia composition," , 2004.

[121] N Bogaards, A Röbel, and X Rodet, "Sound analysis and processing with audiosculpt 2," , 2004.

[122] J B Thiebaut, "Visualization and Reversibility of Sound: The Software Sonos," *Proc. of JIM*, vol. 5, 2005.

[123] M M Farbood, E Pasztor, and K Jennings, "Hyperscore: a graphical sketchpad for novice composers," *Computer Graphics and Applications, IEEE*, vol. 24, pp. 50--54, 2004.

[124] D Margounakis and D Politis, "Converting Images To Music Using Their Colour Properties," , 2006.

[125] K Giannakis and M Smith, "Auditory-visual associations for music compositional processes: A Survey," , 2000.

[126] E Lopes, "Just in time Towards a theory of rhythm and metre," 2003.

[127] Ana Paiva, "Affective Interactions: Toward a New Generation of Computer Interfaces?," vol. 1814, pp. 1-8, 2000.

[128] A Damasio, *Descartes error.*: Papermac London, 1996.

[129] R W Picard, E Vyzas, and J Healey, "Toward machine emotional intelligence: Analysis of affective physiological state," | *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE*, pp. 1175--1191, 2001.

[130] R W Picard, "Affective computing for HCI," , vol. 1, 1999, pp. 829--833.

[131] E Hudlicka, "To feel or not to feel: the role of affect in human-computer interaction," *International Journal of Human-Computer Studies*, vol. 59, pp. 1--32, 2003.

[132] R W Picard, A Wexelblat, I N Clifford, and I N Clifford, "Future interfaces: social and emotional," , 2002, pp. 698--699.

[133] E Oliveira and T Chambel, "Emotional Video Album: getting emotions into the picture," *Emotion in HCI--Designing for People*, vol. 351, p. 16, 2010.

[134] E A Crane, N S Shami, and C Peter, "Lets get emotional: emotion research in human computer interaction," , 2007, pp. 2101--2104.

[135] A Herbon, A Oehme, and E Zentsch, "Emotions in ambient intelligence-An experiment on how to measure affective states," , 2006.

[136] AG Money and H Agius, "Are Affective Video Summaries Feasible," *Emotion in HCI: Joint Proceedings of the 2005*, pp. 142--149, 2006.

[137] AG Money and H Agius, "Video playing with our emotions," *Emotion in HCI: Joint Proceedings of the 2005*, pp. 168--171, 2006.

[138] A Paiva et al., "SenToy: an affective sympathetic interface," *International Journal of Human-Computer Studies*, vol. 59, pp. 227--235, 2003.

[139] P Fagerberg, A Stahl, and K Höök, "eMoto: emotionally engaging interaction," *Personal and Ubiquitous Computing*, vol. 8, pp. 377--381, 2004.

[140] P Sundström, A Stahl, and K Höök, "A user-centered approach to affective interaction," *Affective Computing and Intelligent Interaction*, pp. 931--938, 2005.

[141] M Wright et al., "Panel: Standards from the computer music community," , 2004, pp. 711--714.

[142] M Wright, A Freed, and A Momeni, "OpenSound Control: state of the art 2003," , 2003, pp. 153--160.

[143] M Kaltenbrunner, T Bovermann, R Bencina, and E Costanza, "Tuio-a protocol for table based tangible user interfaces," 2005.

[144] J McCartney, "Rethinking the Computer Music Language: Super Collider," *Computer Music Journal*, vol. 26, pp. 61--68, 2002.

[145] J McCartney, "SuperCollider, a new real time synthesis language," , 1996, pp. 257--258.

[146] M Puckette, "Pure Data: another integrated computer music environment" *Proceedings of the Second Intercollege Computer Music Concerts*, pp. 37--41, 1996.

[147] S Jorda, "Interactive Music Systems For Everyone: Exploring Visual Feedback As a Way for Creating More Intuitive, Efficient And Learnable Instruments," , 2003.

[148] S Jorda, "Sonigraphical instruments: from FMOL to the reacTable," , 2003, pp. 70--76.

[149] M Kaltenbrunner and R Bencina, "reacTIVision: a computer-vision framework for table-based tangible interaction," , 2007, pp. 69--74.

[150] B R Steunebrink, M Dastani, and J J Meyer, "The OCC model revisited," , 2009.

[151] P N Juslin, "Communicating emotion in music performance: A review and a theoretical framework," 2001.

[152] P N Juslin and P Laukka, "Communication of emotions in vocal expression and music performance: Different channels, same code?," *Psychological Bulletin*, vol. 129, p. 770, 2003.

[153] A Teiche et al., "Multi-touch technologies," *NUI Group*, 2009.

[154] D M Cottle, "Computer Music with examples in SuperCollider 3," 2005.

[155] J Stam, "Real-time fluid dynamics for games," , vol. 18, 2003.

[156] S Holm, "A simple sequentially rejective multiple test procedure," *Scandinavian Journal of Statistics*, pp. 65--70, 1979.

Appendix A

Source code dos synths

```
// SERVER INITIALIZATION
(s = Server.new("/ServerASSW", NetAddr("127.0.0.1", 57110));
s.quit;
s.options.memSize = 262144;
s.boot;
//o.memSize;)

// SYNTHS DEFINITION
(
// SYNTHS ABOUT SADNESS

// ----------------- Slow Sadness
(SynthDef("slow_sadness",
{arg rate = 0.02 /*0.02 - 0.1*/, volume = 0.0 /*0.0 - 2.0*/;
var new_rate = max(0, 0.1 - (0.008 * rate));
var new_volume = max(0, (2.222 * volume) - 0.222);
var result = Mix.fill(10,
                  {CombN.ar(
                      LinPan2.ar(
                        SinOsc.ar(/*frequency*/rrand(300, 1000),
                              mul: EnvGen.kr(
                                    Env.perc(rrand(0.1, 1), rrand(0.0001,
0.01), 1, 4),
                                    Dust.kr(new_rate), 0.4))*new_volume,
                              rrand(-1.0, 1.0) //spatialization
                       ),
                       0.5, 0.5, 4)}
                  );
Out.ar(0, result);}).send(s););


// ----------------- Sad Circus
(SynthDef("sad_circus",
{arg speed = 0.002 /*0.001 - 0.3*/, volume = 0.0 /*0.001 - 0.3*/;
var new_speed = min(0.3, 0.001 + (0.1495 * speed));
var new_volume = max(0, (0.332 * volume) - 0.032);
var result = CombN.ar(
                  VarSaw.ar((
                        Hasher.ar(
                              Latch.ar(
                                    SinOsc.ar((1..4)!2),
                                    Impulse.ar([5/2,5]*new_speed)
                              )
                        )*300+300
                  ).round(60),0,LFNoise2.ar(2,1/3,0.5))/5,
                  2,2,20)*new_volume;
Out.ar(0, result);}).send(s););
// adapted from SC140, Micromoog
```

```
// ----------------- Indian Void Trip
(SynthDef("indian_void_trip",
{arg openness = 0.1 /*100.0 - 400.0*/, volume = 0.0 /*0.0, 3.0*/;
var new_openness = max(0, 400 - (300 * openness));
var new_volume = max(0, (3.333 * volume) - 0.333);
var resultL = RLPF.ar(
                  Pulse.ar([50, 70],0.1,0.05),
                  new_openness, 0.1)*new_volume;
var resultH =     RLPF.ar(
                  Pulse.ar([200, 280],0.1,0.05),
                  new_openness*2, 0.1)*new_volume;
Out.ar(0, resultL+resultH);}).send(s););


// ----------------- Radiohead Drone Wannabe
(SynthDef("radiohead_drone",
{arg rate = 1.0 /*0.1 - 3.0*/, volume = 0.0 /*0.0 - 1.0*/;
var new_rate = 0.1 + (0.058 * rate);
var new_volume = max(0, (1.111 * volume) - 0.111);
var result = Mix.ar({LinPan2.ar(PMOsc.ar(2000.0.rand, 800.0.rand,
SinOsc.kr(3000), 0,
                  /*mul*/ EnvGen.kr(
                  Env.linen(2.0, 0.05, 4.0, 1.0, 'sine'),
                  Dust.kr(rrand(0.05, 0.1)*new_rate), 0.4)),
1.0.rand2)}.dup(10))*new_volume;
Out.ar(0, result);}).send(s););


// SYNTHS ABOUT ANGER

// ----------------- Electric Convolution
(SynthDef("electric_convolution",
{arg distortion = 75 /*0.1 - 150*/, roughness = 5000 /*0.0 - 16650*/,
volume = 0.0 /*0.0 - 1.0*/;
var new_distortion = 0.1 + (149.9 * distortion);
var new_roughness = max(0, (555 * roughness) - 11100); //333 * roughness;
var new_volume = max(0, (1.111 * volume) - 0.111);
var env = EnvGen.kr(Env.new([0.001,1,0.3,0.8,0.5,0.8,0],[2,1,1,2,3,1] *
0.1),
                              gate: Impulse.kr(2),
                              levelScale: LFNoise1.kr([5, 5], 2, 1));
var result = SinOsc.ar(LFNoise0.kr([10, 14],
      (SinOsc.kr(new_distortion, 0.5))*LFNoise0.kr(1, new_roughness, 70),
      500), 0,
      /*mul:*/env;)*new_volume;
Out.ar(0, result);}).send(s););


// RHYTHM SYNTHS

// ----------------- Background Beat
(SynthDef("background_beat",
{arg speed = 1.0 /*0.6 - 1.2*/, openness = 30.0 /*10.0 - 100.0*/, volume =
0.0 /*0.0 - 1.0*/;
var new_speed = 0.3 + speed;
var new_openness = 10 + (1.8 * openness);
var new_volume = max(0, (1.111 * volume) - 0.111);
```

```
var result =
LinPan2.ar(SinOsc.ar(Pulse.ar(new_speed)*new_openness))*new_volume;
Out.ar(0, result);}).send(s););
// adapted from SC140, Thor Magnusson


// ----------------- Stress Rhythm
(SynthDef("stress_rhythm",
{arg variation = 0.5 /*0.5 - 3*/, volume = 0.0 /*0.0 - 10.0*/;
var new_variation = 0.5 + (2.5 * variation);
var new_volume = max(0, (volume * 11.111) - 1.111);
var result = CombN.ar(
                  BPF.ar(LocalIn.ar(2)*7.5+Saw.ar([32,33],0.2),
                  2**LFNoise0.kr(new_variation*4/3,4)*300,
                  0.1).distort,2,2,40)*new_volume;

Out.ar(0, result);}).send(s););
// adapted from SC140, Nathaniel Virgo


// ----------------- Positive Rhythm
(SynthDef("positive_rhythm",
{arg speed = 1.0 /*0.0001 - 8.0*/, pitch = 1.0 /*0.1 - 2.0*/, volume = 0.0
/*0.0 - 1.0*/;
var new_speed = 8 * speed;
var new_pitch = 0.1 + (0.038 * pitch);
var new_volume = max(0, (1.111 * volume) - 0.111);
var result = AllpassC.ar(
                  SinOsc.ar(new_pitch*55).tanh,0.4,
                  TExpRand.ar(2e-4, 0.4, Impulse.ar(new_speed)).round([2e-
3,4e-3]),2)*new_volume;
Out.ar(0, result);}).send(s););
// adapted from SC140, Batuhan Bozkurt


// ----------------- Broken Rhythms
(SynthDef("broken_rhythms",
{arg sharpness = 20 /*2 - 200*/, showoff = 0.0001 /*0.00001 - 0.1*/, volume
= 0.0 /*0.0 - 1.0*/;
var result;
var new_volume = max(0, (1.111 * volume) - 0.111);
n={|r,f,n=0,d=1|round(r**LFNoise0.ar([4,1,8,2]!d)*f,n)};
result = Splay.ar(d=n.(3,0.6);
(Ringz.ar(d*showoff,n.(2,n.(sharpness,400),40,20),d).mean.tanh));
Out.ar(0, result*new_volume);}).send(s););


// SYNTHS ABOUT JOY

// ----------------- Acid Straw
(SynthDef("acid_straw",
{arg pitch = 1.0 /*0.0 - 2.0*/, variation = 0.01 /*0.001 - 2.0*/, volume =
0.0 /*0.1 - 1.0*/;
var freq, result;
var new_pitch = 0.01 + (0.0398 * pitch);
var new_variation = 0.001 + (1.999 * variation);
var new_volume = max(0, (1.111 * volume) - 0.111);
freq = Duty.kr(
Drand([0.01, 0.2, 0.4*new_variation], inf), // demand ugen as durations
```

```
0,
Dseq(new_pitch*[204*new_variation, 400, 201*new_variation, 502,
300*new_variation, 200], inf) // frequencies
);
result = SinOsc.ar(freq * [1, 1.4]) * new_volume;
Out.ar(0, result);}).send(s););


// ----------------- Gentle Drops On Your Head
(SynthDef("gentle_drops",
{arg pitch = 1.0 /*0.0 - 200.0*/, rate = 1.0 /*0.0 - 5.0*/, volume = 0.0
/*0.0 - 2.0*/;
var new_pitch = 200 * pitch;
var new_rate = 5 - (2.5 * rate);
var new_volume = max(0, 2 * volume);
var result =
Splay.ar({|i|RLPF.ar(0.6**i*40*Impulse.ar(new_rate*(2**i/32),1/2),
                        4**LFNoise0.kr(1/16)*300+new_pitch,5e-
3).sin}!8)*new_volume;
2.do{result = FreeVerb2.ar(*result++[0.1,1,1])};
Out.ar(0, result);}).send(s););


// SYNTHS ABOUT TENDERNESS

// ----------------- Dreamy Bells
(SynthDef("dreamy_bells",
{arg rate = 0.2 /*0.1 - 0.5*/, pitch = 1.0 /*0.5 - 2.0*/, volume = 0.0
/*0.0 - 1.0*/;
var new_rate = 0.1 + (0.2 * rate);
var new_pitch = 0.5 + (0.04 * pitch);
var new_volume = max(0, (1.111 * volume) - 0.111);
var result = CombN.ar(
      a=[0.02,0.1,1,2,3,4]*new_rate;
      k=LFPar.kr(a+0.5).sum;
      f=Latch.kr(k,Impulse.kr(a))*new_pitch;
      Splay.ar(SinOsc.ar(f*100+300*rrand(0.5, 1.5))/5),
0.2, 0.2, 3)*new_volume;
Out.ar(0, result);}).send(s););


// ----------------- Nostalgic Wall of Sound
(SynthDef("nostalgic",
{arg softness = 0.1 /*100 - 0.1*/, volume = 0.0 /*0.0 - 1.0*/;
var result;
var new_softness = 100 - (1.998 * softness);
var new_volume = max(0, (1.111 * volume) - 0.111);
r=({|i|Blip.ar((i+2).sqrt/256,3,[-1,1].wrapAt(i))*
Gendy1.ar(10,6,1,Blip.ar(i*i+1*[0.5,0.3],9),i+1*60,i+1*new_softness)}!8).su
m/4;
result = (BPF.ar(r,Array.geom(8,160,2),0.005,2).sum.dup) * new_volume;
Out.ar(0, result);}).send(s););
```

Appendix B

Example of evaluation questionnaires

Experiment No.1 – Correlation gestures - affective states (expression without the system)

Preparation

- Conduct the test in a quiet environment without external disturbances, allowing the maximum concentration of the users;
- Use a wall or other vertical surface to simulate as possible the system (delimit some reasonable limits for what will be the screen size of the system);
- Set up a video camera behind the user to reduce the Hawthorne effect (ie, the change of user behavior when is being studied);
- Have some writing material prepared to make the necessary annotations.

Introduction

Firstly, we thank you for accepting the invitation to this test. This is a academic project concerning the developing of a new musical instrument that merge various artistic fields, including music, painting and performance. Moreover, this connection is made through emotions, being the body language of the performer when painting on a canvas the main way to communicate with the system. Therefore, the purpose of this test is to simulate the interaction with this instrument in order to understand the relationship between emotions and gestures.

Thus, we would ask you to run a series of movements touching a vertical surface (just as if you were painting) in response to certain emotions that we will ask you. Note that you can use your entire body, not only your hands (and always without any acessory like brushs). In the first instance, please do the most spontaneous gesture possible, then, reflect briefly on how is the best way to express the requested emotion and repeat the gesture. It is important to note that you have full freedom of movement and there is no right or wrong answers. Finally, you can comment freely throughout the entire experience, and we assure you that the camcorder is only used to study more accurately movements (please try to forget as possible its presence).

Questionnaire

Age: _____ Sex (M/F): _____
Qualifications:

____ < forth grade	____ high school
____ forth grade	____ baccalaureate
____ sixth grade	____ graduation
____ ninth grade	____ post-graduation

Have you already had any type of artistic education? If yes, please indicate which fields (music, painting, performance, etc.) and specify which types (e.g. painting workshop, graduation in music, amateur musician in band, etc.) on the next page.

Field: Type:

______________________ ____
______________________ ____
______________________ ____
______________________ ____
______________________ ____

Emotion-gesture test

What are the gestures that you would choose to represent the following emotions? And after the gesture please indicate your confidence / satisfaction on the accuracy of the gesture to express the emotion required. (Note: mark with an S and an R on each position of the table, respectively the value of the spontaneous and the reflected gestures for each emotion)

	I partially disagree	I totally disagree	Indifferent	I partially agree	I totally agree
Sad					
Shame					
Anger					
Confusion					
Happiness					
Freedom					
Melancholy					
Pride					
Pleasure					
Exaltation					
Tenderness					
Shyness					
Satisfaction					
Loneliness					
Hate					
Fear					
Relief					
Hope					
Disappointment					

Do you agree with this range of emotions? Would you like to add or remove some?

____
____
____

Thanks for your help!

Experiment No.2 – Expression and recognition of affective states with the system

AffectiveWall User Evaluation

This questionnaire is part of a user evaluation, carried out by GAIPS research group (Portugal), integrated in the master thesis of Miguel Jerónimo called **AffectiveWall - An intermedia instrument for affective generation of music and paintings through body-language expressivity**.

AffectiveWall is an intermedia system that translates affective states, manifested by whole body interaction with a multi-touch surface, to music and digital paintings generated in real-time. This project aims to explore the paradigm of a "thrill in the wall", in which thrill stands for the manifestation of emotions through the body language.
Because, all in all, it's just another thrill in the wall!

Please make sure you read this introduction before starting.

The purpose of this questionnaire is to test the recognition of affective states in some interactions with the system. These affective states are:
- **Anger**
- **Disgust**
- **Fear**
- **Joy**
- **Sadness**
- **Surprise**
- **Tenderness**

We ask you to watch sixteen short videos, and for each of them, select the affective state you consider that the performer is expressing . If you don't know what to answer, you can select a "Don't know" option.
We also ask your confidence level about your answers. This should reflect how easy you felt it was to make your decision.

We're sorry for the sound quality and ask you to raise a little bit the volume.
And if you have any problem or suggestion, feel free to spam me: **mljeronimo@gmail.com**!

Thank you for your collaboration!

We will first start on some information about your profile.

Please fill out the following fields:

1. Gender: *
   - ○ Male
   - ◉ Female
2. Year of Birth: * [▾]
3. Country (where you spent most of your life): * [▾]
4. Do you have any artistic background? *
   - ☐ No
   - ☐ Yes, in theatre, performative arts, etc.
   - ☐ Yes, in painting, drawing, etc.
   - ☐ Yes, in music

Yes, in

Let's begin...

Please select the emotion that you think the performer is expressing:

*

- Anger
- Disgust
- Fear
- Joy
- Sadness
- Surprise
- Tenderness
- Don't know, but I would describe as:

Please select how confident you are about your answer: (being 1 the lowest confidence and 5 the highest confidence) *

(the same for the sixteen videos…)

Finally, if you have any comment, suggestion or feeling about the AffectiveWall that you want to share, here is the right place for it!

We'll be very pleased if you can help this research by sharing the questionnaire by email, twitter, facebook, etc!

Many thanks for your time!

* = Input is required

Appendix C

List of affective states with recognizable patterns in reflective expression

(the ones that were not used in the system)

Note: in brackets are the number of subjects in the test group who have fulfilled each pattern. Fractions appear when the subject uses more than one type of measure. In this case, the value allocated to one metric is distributed among all measures used.

Affective State	Patterns Description
Confusion	Many(10) fast(8), long(7), sustained(12) and random(9) gestures performed at medium height(10,83) and with medium(7) or expansive(7) occupation
Freedom	Long(10), sustained(14) and fast(7,5) drawings of lines(7) at medium(7,66) or high(5,66) location and with expansive(11) occupation
Pleasure	One(13) punctual(7,5), static(7) and sustained(14) gesture with all body(10), making a blob(10) at medium(6,33) or high(6,83) location
Exaltation	Many(11) punctual(9) and static(9) blobs(9,5), and so without direction(9,33), made at high(7,83) location and with expansive(9) occupation
Shyness	One(13) punctual(8,5), static(8,5) and sustained(10) blob(9,5) at medium(9,66) height with confined(9) occupation
Loneliness	One(13) punctual(8,5), static(8) and sustained(15) blob(8,5) at low(6,5) or medium(7,5) height and with confined(7) occupation
Hate	Static(8), punctual(8,5) and sustained(12) blobs(8,5) performed by one/both hands(10,16) at medium(9,16) height with medium(7,5) occupation
Fear	One(11) punctual(9,5), static(9,5) and sustained(12) blob(9,5) at medium(10,66) height with confined(9) occupation
Relief	One(13) medium/long(12) line(10,5), made with both hands(6) or all body(5) at low/medium(10,5) speed, at medium(9,66) height with medium(8) occupation
Hope	One(9) sustained(14), punctual(6,5) and static(7,5) (or long(6,5) at medium speed(6)) blob(7,5) at medium/high(13,66) location, with one/both hands(8,5)
Disappointment	One(10) medium/long(9) descending(6) line(8) (or punctual(5) blob(7)) sustained(14) and performed by one/both hands(8,5) at medium(9,83) height

Appendix D

M Jerónimo, C Martinho, A Paiva. Another Thrill in the Wall: an Affective Eco-System Interface for Gestural Expressivity. In proceedings with Whole Body Interaction Workshop of 8th International Conference on Advances in Computer Entertainment Technology – ACE (2011)

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)

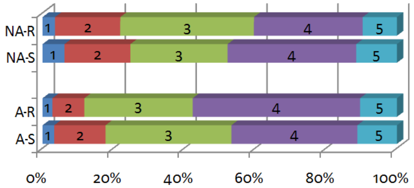


figure 2. Satisfaction/confidence ratings by background ("NA" stands for non-artistic and "A" for artistic) and type of gesture ("R" stands for reflective and "S" for spontaneous), allocated by a 1 to 5 scale (being 1 the lower confidence and 5 the higher). In this test, we only considered subjects with 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).

[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).

[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