



Collaborative creativity: The Music Room

Morreale, F; De Angeli, A; Masu, R; Rota, P; Conci, N

“The final publication is available at <http://link.springer.com/article/10.1007/s00779-013-0728-1>”

For additional information about this publication click this link.

<http://qmro.qmul.ac.uk/xmlui/handle/123456789/12002>

Information about this research object was correct at the time of download; we occasionally make corrections to records, please therefore check the published record when citing. For more information contact scholarlycommunications@qmul.ac.uk

Collaborative creativity: The Music Room

Fabio Morreale · Antonella De Angeli ·
Raul Masu · Paolo Rota · Nicola Conci

Received: date / Accepted: date

Abstract In this paper we reflect on our experience of designing, developing, and evaluating interactive spaces for collaborative creativity. In particular, we are interested in designing spaces which allow everybody to compose and play original music. The Music Room is an interactive installation where couples can compose original music by moving in the space. Following the metaphor of love, the music is automatically generated and modulated in terms of pleasantness and intensity, according to the proxemics cues extracted from the visual tracking algorithm. The Music Room was exhibited during the EU Researchers' Night in Trento, Italy.

Keywords Musical Interfaces · User-Experience · Performing Art · Active Listening · Proxemics

1 Introduction

The fostering of creativity has recently become a key educational, social and economic priority in many parts of the world [45]. There is a growing consensus that creativity, arts and wealth are deeply interrelated and that creativity stems from a collaborative context, rather than from the mind of an isolated individual. Contemporary research aims at a psychological and sociological understanding of the contexts which favour creativity and of the analysis of the meditational role of cultural artifacts, such as signs, tools and technologies. This recent and growing interest offers new challenging arenas for interactive spaces. Innovative metaphors and interfaces are needed to push the boundaries of interactive spaces out of the pragmatic context where they emerged as tools

Experiential Music Lab , MMLab
Department of Information Engineering and Computer Science, University Of Trento,
Trento, Italy
Tel.: +39 0461 283 792
E-mail: {morreale,deangeli,rota,conci} @disi.unitn.it, raul.masu@studenti.unitn.it

for the support of activities related to task accomplishment (borrowing Hasenzhals words, we might say for the support of do-goals). Spaces for collaborative creativity can satisfy higher human needs beyond the instrumental helping users to achieve ‘be goals’. According to Maslows theory, self-actualization is at the top of these needs, being the desire to fulfill a persons full potential [33]. Creativity falls into this complex domain.

This paper describes our experience with The Music Room, an exhibition designed to allow people with little or no musical education to compose original music while moving in an interactive space. In this exhibition, the space becomes the instrument and the stage of an act of creativity that has so far been confined to a small elite of educated musicians. Indeed, due to the complexity of playing and composing, most people can only experience music by dancing or listening to compositions created by somebody else. Only professional musicians can master the intrinsic issues of creating quality music. New technologies, such as ubiquitous computers, touchscreen devices and tracking systems have been designed to deal with the issues related to the creation of original music [4,13,25] by replacing traditional instruments with more intuitive devices. These technological advancements provide novel possibilities by leading a new set of design issues: finding new metaphors to ease the art of creating music and thus opening this creative area to a wider and untrained audience. Current research is focusing on the development of methods to assist music composition by hiding part of the complexity of creation to the user [14]. Such methods partly fail to reach their goals because they rely on traditional sonic and musical metaphors (e.g. scales, sequencers and filters) that are unlikely to convey any meaning to non-musicians. To overcome this limitation, a few works in artificial music composition have been recently trying to search for interaction paradigms into different domains [12,41]. This approach is consistent with the Blended Interaction conceptual framework suggesting that computer-supported collaboration should map elements through different domains by combining real world concepts with digital analogies [24].

With this purpose in mind, we aim at exploring new interaction metaphors, which are supposed to match a series of requirements: they have to be available to everybody, social, intuitive and naturally connected with music. Emotion seems to be the element that best meets these requirements. In each culture, music is one of the arts that can most effectively elicit emotions [16,2]. Significantly, it has always been connoted as emotional [27]. Bodily movements, which, in the different declinations of dancing and conducting, are traditionally associated to music [11,27], are the most appropriate medium through which emotions can be conveyed. The Music Room combines these metaphors by providing a space where people can compose music expressing their emotions through movements. It is designed to be experienced by couples according to the metaphor of love as closeness between people. Two users direct an original piano music composition by providing information about the emotionality and the intensity of the music they wish to create. In order to communicate these factors, users interact with each other by moving in the space: the proximity between them influences the pleasantness of the music, while their speed

influences the intensity. These intuitive compositional rules open endless possibilities of creative musical expression to everybody. The music is generated by Robin, an algorithmic composition system that generates original tonal music in piano style by following a rule-based approach. Proxemics information is acquired by a vision tracking system. It is then converted into emotional cues and finally passed to Robin.

The remainder of the paper is organized as follows. Section 2 presents the related research in the different areas involved in the project. Section 3 introduces The Music Room focusing on the conceptual design and the architecture. Section 4 reports some results from a user evaluation performed at the first public exhibition of the installation. Section 5 features an insight into future developments and discusses our conclusions on the potential of collaborative interactive spaces for fostering creativity.

2 Related works

The design of The Music Room spans several research areas. The rules of the compositional system are inspired by research on the Psychology of Music, while Robin is founded on existing approaches for algorithmic composition. The idea of exploiting the metaphor of gestures and emotions is partially influenced by previous collaborative interactive systems for music expression. Computer vision solutions have been adopted and preferred to other technologies such as wearable sensors, in order to capture the motion of the participants and analyze their level of interaction, so that they could move without any constraint in their behaviors.

2.1 Music and emotion

Music is probably the most widespread art form in the world, pervading the most diverse contexts of our lives, from intimate to public spaces [37,27]. Music has the power to touch peoples souls and influence their moods, as well as stirring emotions, evoking memories and bringing comfort [27,1,34]. Furthermore, music is an important part of social life. While making or listening to music together, people share feelings on a deep level, thus experiencing a bond of mutual empathy [27]. Research on the psychology of music suggests that the interpretation of its emotional connotation depends on acoustic cues embedded in composition (e.g. tempo, rhythm, mode) [23,49,18] and performance behaviors (forms of musical articulations and tempo dynamics such as staccato and accelerando, and other different manners to perform a note) [18,28], whose combinations stir different emotional responses in the listener. There is a significant difference in the emotional reaction that can be elicited by different combinations of musical factors. In the last movement of the Symphony n.6 Pathétique, Tchaikovsky expressed his deep sorrow and anguish through

minor mode and slow *andante lamentoso* tempo. By contrast, Vivaldi reproduced the concept of happiness with the major mode and fast *allegro* tempo in his *Primavera* from *Le Quattro Stagioni*.

For the purposes of measuring and classifying emotions in music, two approaches are available [47]. The categorical approach assumes a finite number of basic emotions which are used to describe unique and innate emotional reactions. Following this viewpoint, the basic emotions elicited by music (i.e., anger, happiness, fear and solemnity [7]) are mapped with a set of musical variables, such as tempo, sound level, timbre, vibrato and consonance [26, 17]. The dimensional approach, instead, describes emotions as the result of the interrelation of a small number (usually two or three) of independent dimensions. Many studies in the music domain [27, 17, 53, 38] are based on Russell's Circumplex model [44]. It suggests that emotions could be described as a continuum along two dimensions: valence, that refers to the positive vs. negative affective state, and arousal, that refers to the rest vs. activation difference. In 1937, Hevren identified the main compositional factors that determine changes in emotional expression in music [21]. She categorized those factors by labeling them on the music's expressive character (e.g. sad, happy, dreamy music). Juslin and Sloboda later elaborated this categorization, representing the emotions along the valence/arousal dimensions [28]. There is a consensus that at the compositional level, mode and rhythm are responsible for valence, while tempo and dynamics are responsible for arousal [28, 17]. At performance time, changes in articulations, loudness, timbre and timing cause a variation in the arousal level [28, 53].

2.2 Algorithmic music composition

Two approaches have so far been adopted to address the challenge of algorithmic music composition: one is based on the editing of existing musical material, the other on algorithmic generation of new melodies. In the editing approach, some compositional parameters (e.g. tempo, note duration, mode, legato phrasing) are modified in a series of excerpts of music [38, 51]. For example, the work by Oliveira et al. composed music by playing-back existing melodies where a number of parameters were dynamically changed according to emotion variations, described in terms of valence and arousal [38]. This is a simple and elegant solution, yet the limited number of musical parameters that can be used to edit existing tracks represents a major deficiency. The variation of pre-composed melodies is restricted, and outcomes are often repetitive. As a result, users are confronted with a limited range of scarcely original songs, which is likely to make their experience less appealing.

The generative approach has been widely explored in the last decade, often supported by a dimensional description of emotion in terms of valence and arousal [53, 30, 52]. Three basic methods have been proposed: rule-based, learning-based and evolutionary composition [50]. The first two methods differ in the procedure used to define the compositional rules [53, 48, 40]. In rule-

based methods, knowledge of music theory is manually coded into the system. The diversity of possible outcomes depends on the amount of taught rules. In learning-based methods, the system is trained with existing musical excerpts and rules are automatically added [8,22,46]. Whilst this method has the clear advantage of decreasing the reliance on designer knowledge, the quality of music is heavily dependent on the training set. Evolutionary algorithms allow the creation of original and unique music by means of mechanisms inspired by biological evolution. This approach ensures an original and unpredictable output [35,19,55] but the music might sound unnatural and structureless if compared to rule-based systems that are generally superior by virtue of the context-sensitive nature of tonal music [36].

2.3 Collaborative musical systems

Research on the design of collaborative systems for playing music has been growing in the last years [3]. Some of these systems target users that have at least a minimum musical training. In this category, we can identify tangible musical interfaces such as the Reactable [25] , Jam-O-Drum [4] , AudioPad [43] and GarageBand for the iPad. Despite the appealing interface and collaborative interaction of these systems, the profuse amount of sonic and musical inputs (e.g. envelopes, sequencers, source generators) tends to make their use difficult for non-musicians. Another category focuses on the concept of active listening, a novel generation of musical systems which are addressed to inexperienced users [43]. In these systems, users can interactively control the music content by modifying it in real-time, while they are actually listening to it [10]. In this area, several works have been trying to enable groups of people to shape musical contents through collaborative interaction [13,12]. In the Urban Musical Game, installation visitors interact with augmented foam balls to manipulate the pre-composed music material [41]. TouchMeDare aims to encourage strangers to collaborate. In fact, two or more people can compose music by interacting through a canvas: the pre-composed music samples, though, are only triggered when the canvas is simultaneously touched by both users [5]. Sync'n'Move allows users to experience music in a social interaction context [51]. Two users freely move their mobiles and the complexity of music orchestration is directly proportional to the synchronization of their own movements. Accordingly, if synchronization fails, there will be no orchestration at all; if synchronization is only partially achieved, orchestration will be quite elementary; in the case of perfect synchronization, the level of orchestration will reach its peak. In *Mappe per Affetti Erranti*[12], a group of people can experience active listening by exploring pre-composed music and navigating a physical and emotional space. The installation encourages collaboration, as music pieces can only be listened to in their full complexity if the participants cooperate with each other by moving through the space. Other studies endeavor to exploit expressiveness and emotions to influence the status of the system [13,12] , through body gestures [32] and dance movements [11].

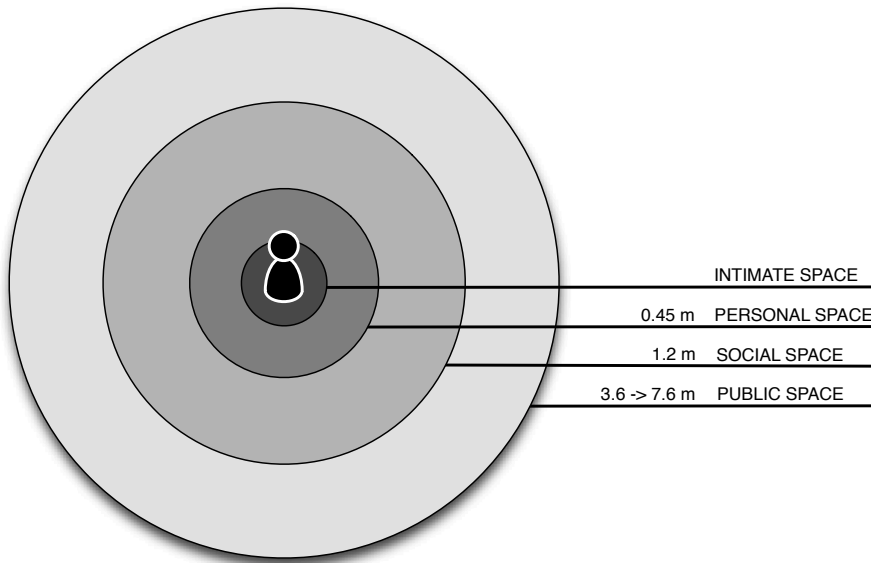


Fig. 1 Definition of the proxemics space as defined by Hall in [20].

2.4 Social interactions via proxemics cues

Proxemics is the study of how humans unconsciously structure spaces in their everyday interaction with others. It reflects the distance at which they are comfortable talking with each other in a specific setting. Proxemics is a form of non-verbal communication which provides important cues about interpersonal relationships. It depends on many factors such as situation, crowd density, racial identity and social relation. Figure 1 represents a high-level overview of the interpersonal distance as proposed by Hall [20]. Different distances reflect different kinds of interpersonal relations taking place. The innermost circle identifies the intimate space, which people share only for very personal relationships, as lovers or family members. While moving towards the outer circles, the distance increases, and the level of intimacy decreases accordingly.

Proxemics can be used as the basic information to understand the level of interaction among people and as such it has gained a significant attention in the computer vision community [9,42]. Proxemic cues are analyzed, as to infer the nature of social interaction among people. Eriksson et al. [15] have proposed a conceptual framework, which emerged from the observation of movement-based interactions; they explain the concepts of space, relations and feedback using four different applications. Motion also represents important information for the acquisition of new visual features, allowing the distinction between ongoing and completed activities [31].

3 The Music Room

New technologies and the increase in computational power are changing live performances with sensors-augmented instruments and widening musical expression with interactive devices such as tabletops, tablet applications and tangible and haptic devices [39,25]. Technological advancement, however, needs to develop new design concepts in order to simplify the access to playing and composition, if it is to make the experience of musical creativity available to everybody. The design of interactive systems for non-musicians introduces a new conceptual issue. The language in which one communicates their artistic intentions cannot rely on musical paradigms that are unknown to most of the people. The avoidance of musical input in the process of music creation has two main effects on the design process. First, in traditional musical instruments (e.g. piano, brasses, synths, percussions) a mechanical input is usually directly associated with a musical output. This does not hold for interactive musical systems, where the input can be separated from the music generation system, as there is no need to force the user to direct the music note by note. An intermediate layer gathers input information, which is then fed into a lower-level module represented by the actual music engine. Second, the purposes of music creation in contexts of collaborative interactive spaces should be reconsidered, since users do not expect to create masterpiece music. The design needs to focus on the experience of users and on the social aspect of interaction and not only on the musical outcome itself. The search for new input metaphors should be consistent with these changes of purpose. The language of emotions and body gestures can be a proper bridge between users' intention and the musical engine. This pervasive language allows humans to convey rich information which is clearly linked with the musical domain. The Music Room is based on the fusion of movements and emotions. It is designed to be experienced by a couple, according to the paradigm of love: two users direct the musical output by supplying the system with information about the emotionality and the intensity of the music they wish to create. To communicate these factors they interact with each other by moving through the space. Their behavior expressed by proxemic cues in terms of relative distance and speed is analyzed to map the pleasantness of music and its intensity respectively. This tool enables users to create music with the desired emotional characterization, merely relying on their own movements and emotions.

3.1 Design process

As discussed by Wiethoff et al.[54] , the design of interactive spaces for collective creativity presents some unique challenges. The Music Room evolved through three months of user-centered iterative design, involving a conceptual stage enriched by early evaluations of scenarios and storyboards and continuous testing of an evolving experience prototype. The ultimate goal of the process was to enable everyone to experience musical creativity through social

interaction. To compose their music, users needed to collaborate with each other in order to get the desired level of emotionality and intensity. During the conceptual design phase, we envisaged two basic scenarios, allowed by the manipulation of the proxemic cues of proximity and speed.

- *Acting scenario.* Users perform a drama and create its soundtrack at the same time. They can act out and perform a storyline, for instance a tragic event with a happy ending. Thus, at the beginning they could stand far from each other, eventually moving to opposite corners of the room. After this tense and tragic stage, they would gradually get closer. As they did so, the music would assume an increasingly cheerful character.
- *Dancing scenario.* Users enjoy the installation by dancing to the music played in the room. Two interaction styles were envisaged: either users passively feel the music, by synchronizing their movements with the beat, or they actively influence it, as to map the dancing style they desire. If they stand far from each other they will achieve a jazzy, improvisational style; on the contrary, if they stand close together they will get a more romantic music.

These scenarios were enriched with graphics and storyboards and used as probes in a design workshop involving 12 user experience researchers, and in 5 contextual interviews with students. Then the basic technological architecture of The Music Room was assembled by interfacing Robin, the music composition system and a visual tracking system. This architecture was tested and fine-tuned off-line through video analysis of scenes specifically recorded for the project and showcasing the two driving scenarios. As the architecture developed into a stable system, the evaluation moved into the laboratory and, only the final week before the Researchers' Night (due to some organizational constraints), into the actual display area. That week was an intense exercise of choreography, implementation and evaluation with a diverse sample of visitors. Design decisions were verified and eventually modified. For instance, the idea of introducing speed as the main interaction channel for controlling music volume emerged from behavioural observations and user comments. In particular, an eight year-old child testing the room with his mother used it as an augmented playground engaging in a musical chase and tag competition. The same requirement was also clearly requested by some student visitors.

3.2 Architecture

The Music Room is composed of two main building blocks: a visual tracking algorithm and Robin, the music composition system (Figure 2). The first module consists of video analysis tool aimed at detecting the moving objects captured by a camera installed in the room, in order to extract proxemic cues from the participants behavior. These values are sent via OSC protocol (Open Sound Control) to Robin, that, in turn, transforms the proxemic information

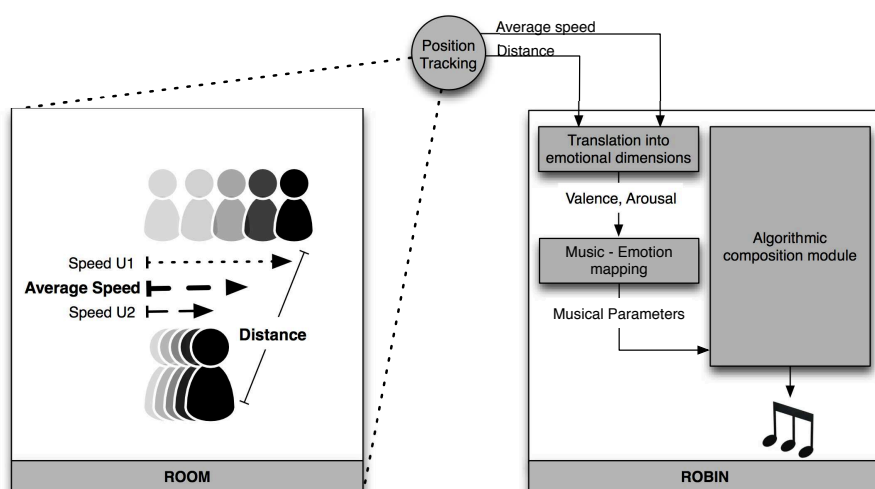


Fig. 2 Architecture of The Music Room.

into the emotional cues of valence and arousal, and then into musical variables. Finally, these musical variables are fed to the Algorithmic composition module, thus determining a change in the generated music.

3.3 Extraction of the proxemics cues

Motion of the participants in The Music Room was captured through a bird-eye video camera installed on the ceiling of the room which looked downwards. This configuration of the camera allowed us to minimize the risk of occlusions, which is intrinsic to any motion tracking application, thus limiting the occurrences of false and missed detections. The detection of the moving subjects has been implemented by applying a standard background subtraction algorithm [29]. The obtained foreground information was then processed by the CamShift tracking algorithm [6]. CamShift has been chosen mainly because of the reduced computational burden, which enabled us to keep up with real-time constraints. The proxemic cues which have been considered in our scenario are distance and speed, as they provide a reliable picture of the ongoing interpersonal relationships. The position of the participants returned by the tracking algorithm was progressively updated over time, and the extracted proxemic cues were supplied to the system, as to provide information about the level of intimacy between them. Figure 3 displays a view of the room as seen by the camera. Two different instances of the interaction are shown on the left, while the output of the detection and the tracking module are portrayed on the right. In particular, the algorithm capability of dealing with partial occlusions, occurring when the two subjects approach, can be observed in the bottom-right image.

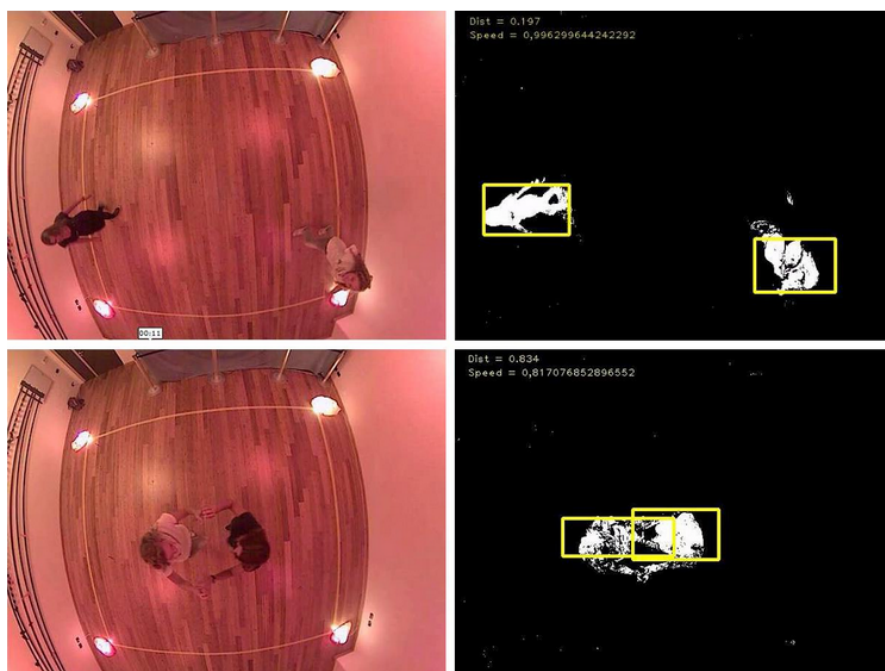


Fig. 3 Two different instances (a,c) of the same experiment with the respective foreground maps portrayed (b,d). The second situation (c,d) shows how the algorithm is effective even when the two participants are close to each other.

3.4 Robin

Robin is the result of several years of research carried out at the Experiential Music Lab of the University of Trento. It is based on a rule-based composition approach: the algorithm is taught a series of basic compositional rules of tonal music that are used to create an original composition in Western classical music. A set of compositional constraints guarantees musical coherence and a pleasant output. The major novelty of the algorithm, compared to related works [46,48,53], lies in its potential to adapt in real-time to users' intention. Depending on the user input and on the internal state of the system, which is determined by previously generated music, the best possible choices of harmonic progression, chords, rhythm and melody are computed at each bar following a statistical approach. Music quality is also improved by the iteration of short themes simulating features such as chorus and verses, that are typically found in every musical genre. At each new step, the system statistically decides whether to repeat the main theme according to the current state of the system or to progress to the generation of a new one. The information coming from the proxemic cues computed by the tracking system conveys the intended emotionality. To map distance and speed with emotions, the analogy of love was used. High proximity is mapped with positive emotions and low proximity

Table 1 How movements are mapped into music.

	INTIMATE SPACE	PERSONAL SPACE	SOCIAL SPACE
SLOW MOVEMENTS	Major mode High tempo Low volume Theme recurrence	Minor mode Mid tempo Low volume	Dissonant Very low tempo Low volume
Emotional Outcome	Romantic, serene	Sad, boring	Tragic, unpleasant
FAST MOVEMENTS	Major mode High tempo High volume Theme recurrence	Minor mode Mid tempo High volume	Dissonant Very low tempo High volume
Emotional Outcome	Happy, joyful	Frightening, depressive	Intense, unpleasant

with negative emotions; high speed with intense emotion and low speed with mild emotions. By matching the values of speed and proximity to emotion, Robin adapts the musical flow. The mapping between emotion and musical parameters is based on the related literature fine-tuned by a couple of ad hoc experiments run by some of the authors. The proximity value influences the emotionality of the song by means of mode, harmonic direction, consonance and tempo (BPM). High proximity is mapped into major mode, rising melody and fast tempo; medium proximity is mapped into minor mode, descending melody and medium tempo; low proximity is mapped into dissonant melody (a number of out-of-scale notes) and unnaturally changing tempo (very fast notes followed by long ones). In addition, in order to increase the difference between positive and negative situations, the recurrence of the theme was treated as an expressive quality of music.

Furthermore, emotion was controlled by theme repetition. In his pioneering work, *Emotion and Meaning in Music*, Leonard Meyer explained how fulfillment of expectations can elicit positive emotions while listening to a piece of music [34]. We operationalized the concept of fulfillment of expectations by presenting a theme numerous times. In contrast, we avoided any repetition to induce negative emotions. Theme recurrence was triggered by users through proximity: the closer they were, the higher was the probability of theme repetition. Lastly, the speed of users determined the dynamics of the song by means of the volume: fast movements were linked to high volume, while slow movements were linked to low volume. Table 1 summarizes the main differences in music parameters and emotional output in 6 conditions, according to the two proxemic cues manipulated: proximity and speed.

The algorithm was implemented in SuperCollider and it communicated via OSC (Open Sound Control) with the tracking system. The output of the SuperCollider patch was a score in MIDI format. Logic Pro, a Digital Audio Workstation, transforms the MIDI flow into piano music.

4 Evaluation

The first version of The Music Room was presented at the EU Researchers' Night on September 28th 2012 in the city center of Trento (Italy). The EU Researchers' Night is a European event which involves 300 venues and is aimed at showcasing research results to a broader population. In Trento, the Researchers' Night took place from 5pm to 2am and it hosted almost 90 demos and installations. The audience attending the event was quite varied: among others, it included students, researchers, families with children, and bystanders.

4.1 Settings

The installation took place in an empty 25 sq m room. The room setting had to be as static and minimal as possible, as to draw the attention of people to the musical cue only. The only visual decorations in the room were a sticky tape square delimiting the walking area on the floor and eight spotlights placed in the corners, four on the floor and four on the ceiling. All spotlights were hidden by yellow and orange decorative papers. The display area was separated from the control area (i.e., a table equipped with some computers and a mixer) by a thick curtain. This increased the level of intimacy of the experience. Each couple was explained the rules of The Music Room before starting the experience: music would be influenced by their own movements.

4.2 Procedure

Before entering the room, people signed an informed consent providing a brief introduction to The Music Room, explaining that their behaviour was videotaped, and asking them to express a preference on how they wanted the video to be treated (immediately deleted, used only for research purposes, published to showcase the installation). Then, they received a short verbal introduction to the installation and its features. Visitors were invited to experience the installation for as long as they wished. However, several couples were invited to leave because of the long queue which gathered around the installation. During the first hour the waiting time was relatively short (less than 15 minutes) but it gradually increased to almost two hours at around 9 PM. Hence, we resorted to a queuing system: the names of those who wished to try the installation were written on a list, so that they could get on with their visit and only come back to the room when their turn would actually arrive.

After they had left the room, couples were asked some questions about their experience. Three researchers in parallel performed the interviews of the couples. A qualitative approach was chosen, as this was the first demo and there were no clear expectations on how the audience would react to this novel experience.

1. What do you think about the installation?
2. Did you feel as if you were controlling or following the music?
3. What would you change? Is there anything that you did not like in particular?

Finally, people were given a card with a URL address and a personal code, which could be used to download their own music composition as an mp3 file.

4.3 Data analysis

A hidden camera (Figure 4) was mounted in the room to audiotape the visitor's performance from a perspective which allowed to see their main emotional reactions. A total of 87 videos were analyzed by 3 researchers and coded thematically. Coding was entirely data-driven. Themes were derived in a number of ways, such as their prevalence in the data and also their importance. Double coding was conducted on around 22% of the data, yielding an almost 88% interrater reliability.

Answers to interviews were loosely transcribed with pen and paper during the event, as no recording devices could be used due to the chaotic contexts where the interviews were performed. The transcripts were then transcribed on a digital support. The coding process was guided by the three main questions.

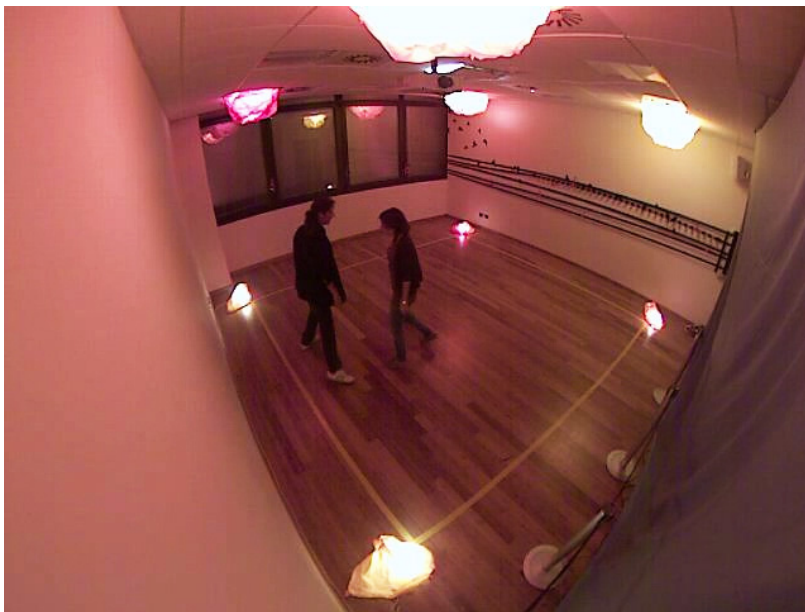


Fig. 4 The Music Room as seen from the hidden camera

4.4 Results

From the opening, The Music Room was constantly busy. Both attendee reviews and the long line suggested that The Music Room was extremely successful. In total, 87 couples (174 people) experienced the installation. The age of visitors was the most variegated: 57 couples of young people (roughly between 16 and 30 years of age), 19 couple of adults (30-60), 5 couples of children (younger than 16) and 6 couples composed of a child and an adult. Session lasted on average 5 minutes (from a minimum of 1m30s up to a maximum of 10 minutes), although these data are biased by the long queue that forced many participants to limit their session.

A common behavioural pattern emerged over time in the performance of the majority of the couples (75%). When entering the room, they initially explored both interactive dimensions of the installation. In this phase, people moved in the space to perceive how proximity affected music, and they changed their speed often ending up running. During this time, they appeared to be more interested in testing the interactive space, rather than enjoying the musical experience per se. However, they eventually found themselves dancing to the music they were creating. The remaining 25% of the couples immersed directly into the experience, dancing or moving closely, and skipping the first exploratory phase.

The most common behaviours in the room were dancing, walking and running. As predicted by the dancing scenario developed in the conceptual design phase, most couples (72%) spent at least 30 seconds dancing. The majority of them (75%) danced closely, often following the steps of traditional couples dancing in western culture. The level of intimacy varied within and between couples, and different dancing patterns emerged. At times, people engaged in slow dances: the lead partner held their hands on the following partner's waist, who draped their arms on the lead's shoulder. They also engaged in waltz style dancing: touching the partner's right hand and hugging their waist with the left. Sometimes, instead, they danced independently, often paying little attention to the partner. In these cases, they often closed their eyes. A significant number of couples (45%) kept moving the whole body (arms, hands, legs), jumping, running after each other and even twirling around. This behaviour was common among male children, who completely disregarded the dancing possibilities but tended to transform the room in a tag playground. Female children engaged in both behaviour, naturally passing from dancing to playing.

The acting scenario we envisaged during the conceptual design phase was less common. Only 25% of the couples spent some time acting as if they were performing a drama or playing an instrument. In around 20% of these cases we could identify a clear leader in the couple, as if one person was taking charge of directing the scene. In general, the room allowed a considerable amount of communication and collaboration. Almost 59% of the visitors were constantly talking to each other in order to define a joint composition strategy (e.g., going to the opposite sides of the room, getting close, running, doing sudden movements). The remainder of the sample preferred a more intimate

and less planned experience, though almost everybody had some form of social interaction mediated by words and/or physical contact.

By observing people's behaviour it was evident that everybody deeply enjoyed the experience. People were often laughing, smiling or openly commenting on it. These impressions were confirmed by the short debriefing interview. All users but two regarded their experience as extremely exciting. The most common words used to describe it were 'cool', 'interesting', 'unique', 'intimate', 'pleasant', and 'relaxing'. Several people also commented on the quality of the music, which was often perceived as barely distinguishable from that produced by a human musician. The two people who did not like the experience complained about a lack of interactivity in The Music Room as if they wanted something physical to be explicitly manipulated. The question concerning the perceived level of control on the musical outcome highlighted a dichotomy. Nearly half of the users reported that they felt as if they were actively controlling the music, while the other half declared that they were mainly following the music, only having the impression of playing an active role in a few situations.

Most people just replied 'nothing' to the question concerning possible improvement and desirable changes. However, some interesting suggestions were collected. The most important complain stressed an annoying latency between the users' movements and the sound they generated. Several interviewees argued that the response should have been prompter. Other people would have appreciated a wider range of musical genres and instruments, or suggested the possibility of associating a different instrument to each person. Some visitors also suggested extending the set of actions through which users can influence the music, including hand and arm gestures. Finally, several participants provided some interesting observations concerning the layout of the room. They suggested setting up The Music Room in a larger and darker space, as to foster intimacy. Other people suggested using lights to improve the general choreography of the room: they pointed out that light might be used to recreate a nightclub context, and perhaps change according to the movements of users.

4.5 Discussion

The video analysis and the interviews confirmed the suitability of the two scenarios we had envisaged at the conceptual design stage, with a strong predominance of the dancing scenario. Some interviews, supported by data gathered by the video analysis, put the emphasis on the expectations that people had toward the installation. Even though they had all been informed on how they could interact with the systems, almost half of them repeatedly moved their whole body, expecting their hands and arms movements to produce a change in music. A relevant number of people frequently expressed their enthusiasm towards the experience by twirling, jumping and playing with the partner even though they were aware that those movements were not influencing the music. Almost 40% of the couples did not collaborate as expected, preferring



Fig. 5 In The Music Room the users get influenced by the music they are actively directing.

an isolated experience, sometimes accompanied by solitary dancing. While in some cases this could be explained by lack of intimacy with the partner or shyness, in other cases users had assumed that they could individually change the music and instruments, as pointed in some interviews. Moreover, several other unexpected themes emerged. Two couples of skilled dancers reported that directing the same music they were dancing to was a completely new experience. Some users declared that for the first time they were not worried about dancing awkwardly, because it was music that was actually following their own movements. Their movements, the music they generated, their dancing in reaction to the music etc. originated a feedback circuit (see Figure 5).

5 Conclusion and Future works

Overall, The Music Room experience provided strong support to the potential of interactive spaces as a social-technological tools for fostering creativity. The exhibition definitely aroused people curiosity - so much so that they were willing to wait for a long time in an extremely crowded corridor.. We need to reflect on the reasons of such a success. One of the possible reasons may be due to the fact that The Music Room allowed them to do something they could have never done: they engaged in an act of collaborative musical creativity. In the introduction to this paper, we cited Maslows theory [33] of the motivational forces driving human behaviour. The Music Room demonstrated that technology can appeal to several highest human needs, such as the need for love, for esteem, for self-actualization. The experience also stressed the value of both summative and formative evaluation. The feedback from observation and interviews suggested new ideas for the improvement of the user experience. The most critical issue of the installation seems to be represented by the slow response of the music to user movements. Even though the musical system is specifically designed so that the music generation synchronizes with the user input, we purposely decided to avoid sudden changes in music. This choice was mainly dictated by aesthetic reasons: the phraseological structure of music should be preserved even in case of sudden change in the emotional input. For this purpose, the computation of the following musical phrase happens at

the last beat of the playing bar (the shortest meaningful musical element) that, for this demo, always measures $2/2$. The worst-case scenario, occurring with a 60 BPM tempo and the bar at the first beat, results in an approximately 2-second delay. In order to reduce the latency, yet keep the musical evolution coherent, two solutions have been adopted:

1. In order to shorten the duration of the bar at worst-case, we changed the tempo range from 60-140 to 90-150 BPM. In the worst-case scenario the bar lasts 1.50 seconds and 90 BPM is still slow enough to be perceived as *andante* (slow).
2. Every half a beat, a new input from the user is checked. If they rank above or below a specific threshold, therefore qualifying as relevant, a new bar immediately starts.

As a result, the latency time drops to 0.75 seconds in the worst-case scenario: this figure should be sufficiently low for these kinds of interactive situations. Furthermore, we are experimenting with more musical genres and instruments, as to enable users to choose their favorite musical style before trying the installation. Some changes in the room setting are also being planned: among them, a larger space and a more reliable tracking system that is not sensitive to light variations. As a consequence, the room could be darker and lights would change according to the music, thus enhancing the sensorial experience. The new version of The Music Room will be exhibited in the Museum of Sciences by March.

Acknowledgements First and foremost, we wish to thank all the visitors of the 2012 EU Researchers' Night in Trento, who often faced a long wait in order to experience The Music Room. We sincerely hope you enjoyed the installation and we are working on an improved version of it, basing on your valuable suggestions. We are much obliged to the members in our team, who helped us to set up and run The Music Room in two weeks of hard voluntary work. Special thanks to Patrizio Fava, Andrea Pecchielan, Maria Menendez, Chiara Gadotti, Zeno Menestrina and Michele Bianchi. Finally, we wish to express our gratitude to Costanza Vettori, who helped us to edit this paper.

References

1. Alpert, J., Alpert, M.: Music influences on mood and purchase intentions. *Psychology and Marketing* **7**(2), 109–133 (1990)
2. Balkwill, L., Thompson, W., Matsunaga, R.: Recognition of emotion in japanese, western, and hindustani music by japanese listeners¹. *Japanese Psychological Research* **46**(4), 337–349 (2004)
3. Blaine, T., Fels, S.: Contexts of collaborative musical experiences. In: *Proceedings of the 2003 conference on New interfaces for musical expression*, pp. 129–134. National University of Singapore (2003)
4. Blaine, T., Perkis, T.: The jam-o-drum interactive music system: a study in interaction design. In: *Proceedings of the 3rd conference on Designing interactive systems: processes, practices, methods, and techniques*, pp. 165–173. ACM (2000)
5. van Boerdonk, K., Tieben, R., Klooster, S., van den Hoven, E.: Contact through canvas: an entertaining encounter. *Personal and Ubiquitous Computing* **13**(8), 551–567 (2009)
6. Bradski, G.: Computer vision face tracking for use in a perceptual user interface (1998)

7. Bresin, R., Friberg, A.: Emotional coloring of computer-controlled music performances. *Computer Music Journal* **24**(4), 44–63 (2000)
8. Brooks, F., Hopkins, A., Neumann, P., Wright, W.: An experiment in musical composition. *Electronic Computers, IRE Transactions on* (3), 175–182 (1957)
9. Calderara, S., Cucchiara, R.: Understanding dyadic interactions applying proxemic theory on videosurveillance trajectories. In: *Computer Vision and Pattern Recognition Workshops (CVPRW), 2012 IEEE Computer Society Conference on*, pp. 20–27. IEEE (2012)
10. Camurri, A., Canepa, C., Volpe, G.: Active listening to a virtual orchestra through an expressive gestural interface: The Orchestra Explorer. In: *Proceedings of the 7th international conference on New interfaces for musical expression*, pp. 56–61. ACM (2007)
11. Camurri, A., Lagerlöf, I., Volpe, G.: Recognizing emotion from dance movement: comparison of spectator recognition and automated techniques. *International journal of human-computer studies* **59**(1), 213–225 (2003)
12. Camurri, A., Varni, G., Volpe, G.: Towards analysis of expressive gesture in groups of users: computational models of expressive social interaction. *Gesture in Embodied Communication and Human-Computer Interaction* pp. 122–133 (2010)
13. Camurri, A., Volpe, G., De Poli, G., Leman, M.: Communicating expressiveness and affect in multimodal interactive systems. *Multimedia, IEEE* **12**(1), 43–53 (2005)
14. Cope, D.: *Computer models of musical creativity* (2005)
15. Eriksson, E., Hansen, T., Lykke-Olesen, A.: Movement-based interaction in camera spaces: a conceptual framework. *Personal and Ubiquitous Computing* **11**(8), 621–632 (2007)
16. Fritz, T., Jentschke, S., Gosselin, N., Sammler, D., Peretz, I., Turner, R., Friederici, A., Koelsch, S.: Universal recognition of three basic emotions in music. *Current Biology* **19**(7), 573–576 (2009)
17. Gabrielsson, A., Juslin, P.: Emotional expression in music performance: Between the performer’s intention and the listener’s experience. *Psychology of music* **24**(1), 68–91 (1996)
18. Gabrielsson, A., Lindström, E.: *The influence of musical structure on emotional expression*. (2001)
19. Gartland-Jones, A., Copley, P.: The suitability of genetic algorithms for musical composition. *Contemporary Music Review* **22**(3), 43–55 (2003)
20. Hall, E.: *The silent language*. Anchor (1973)
21. Hevner, K.: The affective value of pitch and tempo in music. *The American Journal of Psychology* **49**(4), 621–630 (1937)
22. Hiller, L., Isaacson, L.: *Musical composition with a high-speed digital computer, machine models of music* (1992)
23. Ilie, G., Thompson, W.: A comparison of acoustic cues in music and speech for three dimensions of affect. *Music Perception* **23**(4), 319–330 (2006)
24. Jetter, H., Geyer, F., Schwarz, T., Reiterer, H.: *Blended Interaction Toward a Framework for the Design of Interactive Spaces*. In: *Proceedings of Workshop on Designing Collaborative Interactive Spaces AVI2012* (2012)
25. Jordà, S., Geiger, G., Alonso, M., Kaltenbrunner, M.: The reactable: exploring the synergy between live music performance and tabletop tangible interfaces. In: *Proceedings of the 1st international conference on Tangible and embedded interaction*, pp. 139–146. ACM (2007)
26. Juslin, P., Laukka, P.: Expression, perception, and induction of musical emotions: A review and a questionnaire study of everyday listening. *Journal of New Music Research* **33**(3), 217–238 (2004)
27. Juslin, P., Sloboda, J.: *Handbook of Music and Emotion: Theory, Research, Applications: Theory, Research, Applications*. OUP Oxford (2009)
28. Juslin, P., Sloboda, J., et al.: *Music and emotion*, vol. 315. Oxford University Press Oxford (2001)
29. KaewTraKulPong, P., Bowden, R.: An improved adaptive background mixture model for real-time tracking with shadow detection. In: *Proc. 2nd European Workshop on Advanced Video Based Surveillance Systems*, vol. 25, pp. 1–5 (2001)

30. Legaspi, R., Hashimoto, Y., Moriyama, K., Kurihara, S., Numao, M.: Music compositional intelligence with an affective flavor. In: Proceedings of the 12th international conference on Intelligent user interfaces, pp. 216–224. ACM (2007)
31. Lepri, B., Mana, N., Cappelletti, A., Pianesi, F., Zancanaro, M.: What is happening now? detection of activities of daily living from simple visual features. *Personal and Ubiquitous Computing* **14**(8), 749–766 (2010)
32. Mancini, M., Castellano, G., Peters, C., McOwan, P.: Evaluating the communication of emotion via expressive gesture copying behaviour in an embodied humanoid agent. *Affective Computing and Intelligent Interaction* pp. 215–224 (2011)
33. Maslow, A.H.: A theory of human motivation. *Psychological Review* **50**(4), 370–396 (1947)
34. Meyer, L.: *Emotion and meaning in music*. University of Chicago Press (1956)
35. Miranda, E., Biles, J.: *Evolutionary computer music* (2007)
36. Nierhaus, G.: *Algorithmic composition: paradigms of automated music generation*. Springer Verlag Wien (2009)
37. O’Hara, K., Lipson, M., Jansen, M., Unger, A., Jeffries, H., Macer, P.: Jukola: democratic music choice in a public space. In: Symposium on Designing Interactive Systems: Proceedings of the 5 th conference on Designing interactive systems: processes, practices, methods, and techniques, 04, pp. 145–154 (2004)
38. Oliveira, A., Cardoso, A.: A musical system for emotional expression. *Knowledge-Based Systems* **23**(8), 901–913 (2010)
39. O’Modhrain, M., Adviser-Chafe, C.: *Playing by feel: incorporating haptic feedback into computer-based musical instruments*. Stanford University (2001)
40. Rader, G.: A method for composing simple traditional music by computer. *Communications of the ACM* **17**(11), 631–638 (1974)
41. Rasamimanana, N.H., Bevilacqua, F., Bloit, J., Schnell, N., Flty, E., Cera, A., Petrevski, U., Frechin, J.L.: The urban musical game: using sport balls as musical interfaces. In: J.A. Konstan, E.H. Chi, K. Hk (eds.) *CHI Extended Abstracts*, pp. 1027–1030. ACM (2012)
42. Rota, P., Conci, N., Sebe, N.: Real time detection of social interactions in surveillance video. *Computer Vision–ECCV 2012. Workshops and Demonstrations* pp. 111–120 (2012)
43. Rowe, R.: *Interactive music systems: Machine listening and composition* (1993)
44. Russell, J.: A circumplex model of affect. *Journal of personality and social psychology* **39**(6), 1161 (1980)
45. Sawyer, R.: *Explaining Creativity: The Science of Human Innovation: The Science of Human Innovation*. Oxford University Press, USA (2012)
46. Simon, I., Morris, D., Basu, S.: Mysong: automatic accompaniment generation for vocal melodies. In: Proceedings of the twenty-sixth annual SIGCHI conference on Human factors in computing systems, pp. 725–734. ACM (2008)
47. Sloboda, J., Juslin, P.: *At the interface between the inner and outer world: Psychological perspectives* (2010)
48. Steedman, M.: A generative grammar for jazz chord sequences. *Music Perception* pp. 52–77 (1984)
49. Temperley, D.: *The cognition of basic musical structures*. MIT press (2004)
50. Todd, P., Werner, G.: Frankensteinian methods for evolutionary music. *Musical networks: parallel distributed perception and performance* pp. 313–340 (1999)
51. Varni, G., Mancini, M., Volpe, G., Camurri, A.: Synchmove: social interaction based on music and gesture. *User Centric Media* pp. 31–38 (2010)
52. Wallis, I., Ingalls, T., Campana, E.: Computer-generating emotional music: The design of an affective music algorithm. *DAFx-08, Espoo, Finland* pp. 7–12 (2008)
53. Wallis, I., Ingalls, T., Campana, E., Goodman, J.: A rule-based generative music system controlled by desired valence and arousal (2011)
54. Wiethoff, A., Gehring, S.: Designing interaction with media façades: a case study. In: Proceedings of the Designing Interactive Systems Conference, pp. 308–317. ACM (2012)
55. Wiggins, G., Papadopoulos, G., Phon-Amnuaisuk, S., Tuson, A.: Evolutionary methods for musical composition. *DAI RESEARCH PAPER* (1998)