CaDaReMi – an educational interactive music game

R Gehlhaar[1], P M Rodrigues[2] and L M Girao[3]

[1]Coventry School of Art & Design, Coventry University, UK

[2]Fundação Casa da Música, Porto, PORTUGAL

[3]Artshare Lda., Aveiro, PORTUGAL

[3]Planetary Collegium – CAiiA Hub,
Faculty of Technology, University of Plymouth, UK

Rolf.Gehlhaar@gmail.com, *prodrigues@casadamusica.com*, *Luis.Girao@gmail.com*

www.gehlhaar.org, *www.casadamusica.com*

ABSTRACT

This new multi-user interactive sound installation (≤ 8 persons simultaneously) implements proprietary glob-recognition and tracking software in order to allow visitors to a large empty space (~5m × 7m) to move an avatar – projected on a screen at the end of the space – simply by moving about the space, with the objective of taking it to specific, recognizable locations. Success in this endeavor causes sounds to be triggered.

1. BACKGROUND

In 1984 Rolf Gehlhaar developed SOUND=SPACE (1), an interactive multi-user musical environment in which visitors trigger and influence the production of sounds merely by moving about an empty space surveyed by an ultrasonic echolocation system. Since its development it has been displayed publicly worldwide, becoming a particular favourite with special needs groups primarily because it makes creative musical expression accessible to persons generally excluded. SOUND=SPACE is still being explored by visitors and participants in creative workshops for special needs groups. At the time of writing one of the systems is installed 'permanently' in Casa da Musica, Porto, Portugal, where workshops for special needs groups are taking place weekly.

Some of the many lessons we have learnt from SOUND=SPACE installations and workshops with special needs groups is that, if a multi-user installation is to make justifiable claims to being educational, several criteria must be met: How to use the installation must become clear upon only a brief exposure to it and gradually 'intuitive', i.e. with very little explanations required. Users must be able to understand the 'sound topology' (the spatial distribution of the sounds and the locations of the functions of control). The installation must be so fashioned that it is possible to get better at using it. Ideally, it should possible for 5−8 people to use it simultaneously with each user still being able to understand what it is they are doing individually, how they are contributing to the global sound.

Because in SOUND=SPACE there is practically no visual information about where the sounds are located, one of the problems we have experienced is that it is quite difficult for users to 'anchor' their activities spatially. In order to achieve some mastery, SOUND=SPACE demands fairly intense listening skills with considerable attention paid to the sounds, so that the users can relate different sounds to different locations in the space, different gestures to different effects.

2. AIMS & OBJECTIVES

CaDaReMi addresses this problem by providing a number of visual clues designed to help the user understand 'how things work', to use the 'spectacle' of the installations to 'explain' it to new users and to make the sound topologies visible. Beyond this, the following aims and objectives remain:

- To be entertaining and provide a stimulus for users to explore expressively a wide range of different sounds;

- To be collaborative: it may be used by several persons (up to as many as 8) at the same time;
- To be sonically challenging and interesting, providing a palette of both familiar and strange sounds;
- To be visually engaging, enhancing the users' experiences and promoting their ability to locate themselves and to decode the events of others in the space at the same time;
- To be socially engaging by promoting user-user interaction, thus strengthening the sense of community of activity and place;
- To be intuitive in its functionality and use, with no explanations required and no expertise in order to get 'first results'.
- To be learnable and masterable: sufficiently complex that users may, in time, enjoy the experience of "getting better" at using it; but, at the same time, sufficiently easy that beginners may quickly experience success.

3. THE STRUCTURE AND PROCESS OF *CaDaReMi*

An overhead digital infrared camera, whose images are processed by a proprietary glob-recognition and tracking program, monitors the empty horizontal space of the installation, approximately 5m × 7m. The space is semi-dark, illuminated only by the reflected light from a large screen, c. 4m × 4m, at the long end of the room. When the space is empty a projection consisting of distinctive opaque black, rounded, angular and sharp profiles against a light grey background (see Figure 1). These profiles never change shape.

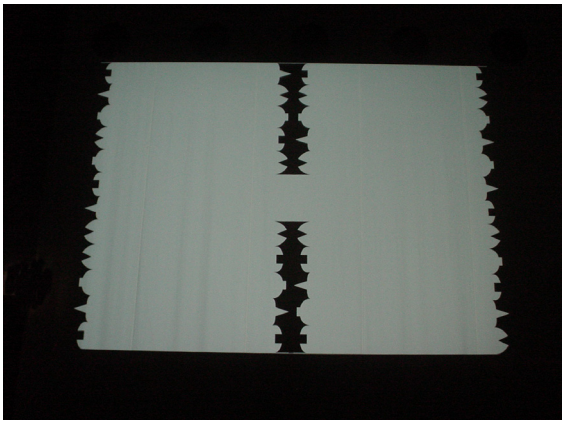

Figure 1. *The empty space.*

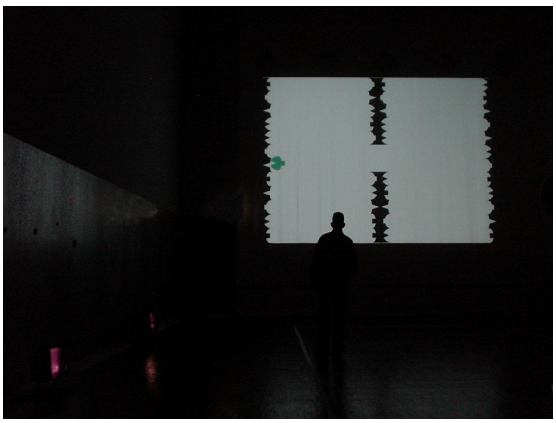

Figure 2. *A single player.*

The mere presence of a visitor in the space surveyed by the infrared camera causes the recognition program to generate and project a virtual 'object' into this landscape of profiles, an avatar of the located user that follows her about as she moves (see Figure 2). Simultaneous with the appearance of the distinctively shaped avatar appears, a specific sound is triggered. This individual sound, which we will refer to as the *name* of the avatar, is attached to every avatar and is only triggered only once, when the avatar first becomes visible. An avatar never changes shape, only its position on the screen and its colour. Whenever a user leaves the space, his/her avatar disappears. When a user enters the space again, a new avatar is generated, a new name is 'called'.

The left and right profiles of the avatar clearly suggest that they bear some relation to the landscape of profiles, demanding to be explored with the avatar. By moving in the space the visitors may move their avatar: moving right/left moves the avatar right/left, moving towards the screen moves the avatar down, moving away from the screen moves the avatar up. Access to the space is restricted to the far and of the space; hence all avatars enter the landscape of profiles at its top. Whenever a visitor moves her avatar (see Figures 3 & 4) to any one of these projected profiles whose shape will 'fit' the avatar, the avatar blinks and a sound specific to that 'object' is emitted. The metaphor here is that of *lock* (profile) and *key* (avatar). For the sake of simplicity, they will be referred to as such in the following.

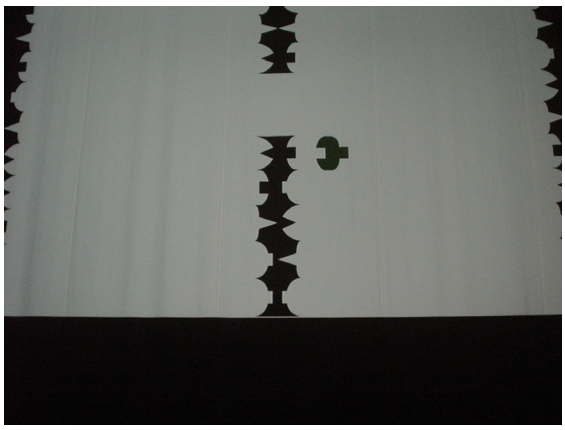

Figure 3. *A key approaches a lock.*

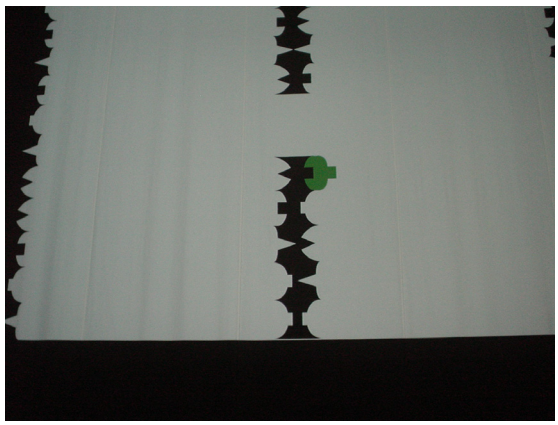

Figure 4. *The key fits & blinks.*

A further important function of the installation is that players may also 'fit' their keys to each other (see Figures 5 & 6). The profiles of the keys may also function as locks for other keys. With only 2−3 players in the space, there is no guarantee that there will be a key that fits another but as the number of players increases the probability steadily increases. When two keys are fit a sequence of sounds is emitted: the names of the avatars mentioned above. The two names are 'called' repeatedly, one after the other, as long as the fit is maintained. This function, extrapolated from experience gathered by one of the authors (Gehlhaar) in workshops with *SOUND=SPACE, is* primarily designed to engender a sense of community, to encourage cooperation and promote the spontaneous creation of 'close' encounters within the space, which can often result in many quite humorous 'entanglements'.

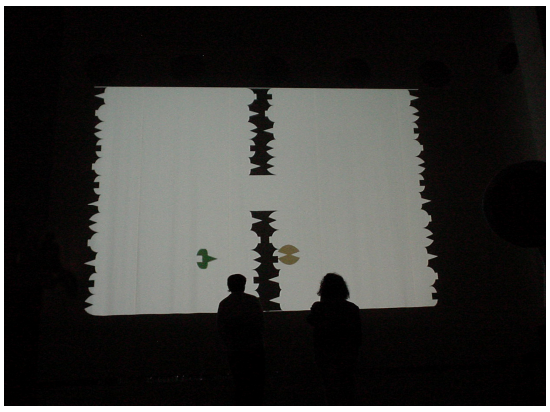

Figure 5. *Two players.*

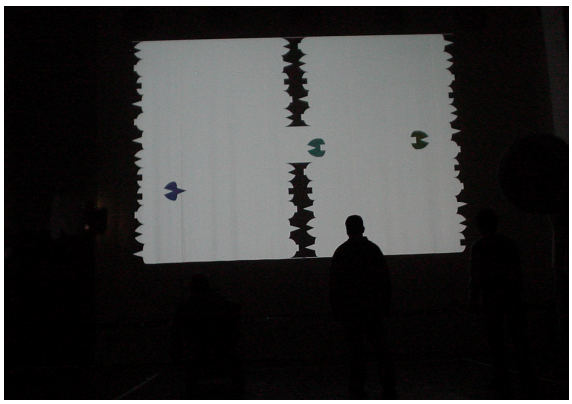

Figure 6. *Three players.*

The maximum number of different locks and keys possible (and required for multi-user applications) is determined by several factors: the focal length of the camera, the height of the camera, the resolution of the projector, the size of the person whose glob is being recognized and converted into an avatar, the desirable minimum number of different keys and the desirable minimum number of locks available to every key. In this version these factors interacted to allow for a total of 44 locks, 22 of four different left-hand profiles and 22 of four different right-hand profiles, with a maximum number of 16 different keys (four different *left* profiles × four different *right* profiles). Every time the program is executed a different combination of locks is generated. Similarly, every time a person enters (or re-enters) the space a different key is generated for them. Each key fits and average of 5−7 left hand locks and 5−7 right-hand locks, resulting in an average total of 10−14 different sounds available to a player.

The structure of the software developed for *CaDaReMi* is composed of two programs that run simultaneously in the same computer and communicate between them via network packets. One of them is dedicated to video capture and position analysis based on glob distinction techniques. The other is dedicated to the audiovisual outputs – graphic symbols and associated sounds. This solution allows future applications that only use one of the programs and makes the development and implementation of newer versions of each one of them easier to. Furthermore, this solution already opens doors for the development of multi-post and multi-location versions of *CaDaReMi*. In this phase, all software was written in Processing, an open source multi-platform programming environment. Therefore, the software of *CaDaReMi* can run in several different types of computers, running different operating systems.

The process of calibration of the system, which is normally one of the biggest problems in this type of installation, is very simple and reduced. All image analysis is done after a background subtraction process which simplifies immensely the distinction processes. This, combined with the fact that all images analyzed are in the infra-red range, also reduces the possibility of interference from other light sources. The only parameters to be adjusted for each iteration of this installation – all in the program that generates the audiovisual output − are the ones concerning the size of the space being used, the camera and the projection system. That is, scale factors need to be determined in order to establish a even relationship between the size of the image captured by the camera and the size of the projected 'avatar' (key).

This location and glob-recognition interface developed by us (Girao & Gehlhaar) for *CaDaReMi* is generic and will be used for any number of different activities and 'games', some of which may not necessarily be directed towards the creation and manipulation of sound.

4. FIRST INSTALLATION IN CASA DA MÚSICA, APRIL 2008

The first installation of *CaDaReMi* took place during the month of April, 2008 as a part of the programme 'Ao Alcance de Todos' (*Accessible to Everyone*) devised and programmed by one of the authors (Rodrigues). The Education Service at Casa da Música (CdM) organizes a broad range of regular activities and throughout the year many people with disability come to CdM to participate in workshops or to attend concerts. A number of medium term creative projects with disabled people at Casa da Música are also organized and there is a special program (A Casa vai casa) that allows institutions to request CdM educators to come to their facilities and develop short term projects on site. It is, however, during one week in April that the issue of music and disability receives a particular focus with 'Ao Alcance de Todos'. Specific workshops, conferences, concerts, showcases and debates focusing on Music, Technology and Special Needs are presented at this time, offering an opportunity that institutions dealing with special needs usually rapidly subscribe to.

In the case of *CaDaReMi,* five schools for special needs students responded to an invitation to send a group to participate in a 1½ hour-long workshop. The workshops, held in the morning and afternoon of three consecutive days, were facilitated by two of the authors (Gehlhaar & Girao), as well as accompanied by at least two teachers from the respective school and one or two observers. Dr. Eva Petersson of Aalborg University, Denmark, was invited by Casa da Musica to carry out a professional evaluation. The profile of the groups attending the workshops was varied: each consisted of 12−20 students of similar ages (11−20 yrs), all primarily with pronounced learning difficulties. Each group also encompassed two or three wheelchair-bound participants. Overall, their social skills within their group were, however, typical of what one would expect of any young persons of that age**.**

All of the workshops were organised along similar lines: each began with a short verbal introduction, explaining only that the workshop was about playing with sounds. This was followed by a quick demonstration: one of the facilitators enters the space, a *name* sound is triggered, and then he gradually moves his key to one of the locks, producing another sound. Usually this was repeated once or twice. And then participants of the workshop were invited one by one to try it out for themselves. Once everyone had tried it out alone, they were invited to 'play' together, two, three or four at the same time.

5. EVALUATION

The evaluation of the installation and the workshops was carried out in two ways: observation and the questionnaire circulated among the participants after the workshop. In some instances the questionnaires were completed for the student in consultation with the teacher. These questionnaires were taken for an evaluation by Dr Eva Petersson **(**Assistant Professor, Aalborg University, Department 7, 6700 Esbjerg, Denmark), who will be publishing her findings next year.

Beyond our aims and objectives for the installation that it should be a pleasurable, social experience, as outlined above, we also expected a variety of learning outcomes. The first of these were that the players of *CaDaReMi* would readily, without any explanation, understand the affordances of the installation, that they would develop an awareness of the concept of sound topologies and that they would improve their performance within a short time.

The influence of the accompanying teachers on the success of the activities in the space was notable. Some of the teachers just sit back and relax, seemingly delighted to be relieved of their usual responsibilities for a short time; others seek to maintain control, sometimes discouraging exploration thereby; others join in

and act as a fabulous catalysts for play and learning. In this way, the teachers at special schools seem to be very little different from teachers at other schools.

It was quite clear that, after the first demonstration, everyone completely understood the main metaphor and the affordances of the installation, as well as the directional correspondences between movement of self and movement of the key on the screen. Most of them were able quickly to identify a lock that would fit their key and then promptly decide to manoeuvre it there.

The main activity in the space – moving the key to its intended lock – is not without its difficulties. Control over its movement must be practiced a little. The tendency for the beginner is to move too fast or too far in the intended direction, resulting in the key crashing into the sides and getting stuck. As there is no haptic feedback, the player must develop a sense of the time it takes the computer to react to the movement of the player, how long it takes to move the key as well the relationship between the amount of movement of the body in real space and the amount of movement of the key in the projected space.

Because of this level of difficulty, success sometimes took more than a minute. Some participants were not even successful after several attempts, notwithstanding the frequent shouts of directions and encouragement from their fellow students all the while. However, no one finished their turn completely frustrated; it just took a bit longer.

As a result, the sound arising from a fit tended to be considered a reward for success rather than the main objective. Although the participants always applauded a success, demonstrating their participation in this social activity, this level of difficulty was not intended. Our intention was that, because of the visual referent, it would be relatively easy to trigger sounds. We clearly need to find a different compromise between accuracy of a person's control over their key and the accuracy of creating a match. However, after several attempts, many participants were able to improve their performance, to trigger a sound fairly quickly and repeatedly, a significant measure of the effectiveness of the installation as a creative and learning environment. Responses to questionnaire clearly indicated that the vast majority of the students wanted to do it again and have more time at it.

There is no question as to whether the 'lock and key' feature contributes to the development of spatial awareness and abstract conception of space: this is primarily evidenced by the fact that all players took great pleasure in simply moving their key about the space, even if they (only very few) were a little annoyed by not being able to fit a lock at the first attempt). But this annoyance did not stop them from continuing. As the players developed an increasing awareness of the correspondences between the real and the virtual space − something that is, we believe, becoming of increasing importance for everyone, able-bodied as well as disabled − they developed greater sensitivity of control and greater accuracy of movement. Thus, as long as there is no frustration, the hope for success remains alive. In any case, players are likely to have seen others succeed before taking their own turn.

6. CONCLUSIONS

6.1 How well did we meet our aims?

It was among our intentions:

- to create an entertaining environment which will stimulate users to explore expressively a wide range of sonically challenging and interesting sounds.

This aim was met with varying degrees of success: The visitors enjoyed the workshops immensely. The installation was accessible to all, everyone quickly understood the metaphor and the affordances and could immediately attempt to trigger sounds. Everyone was thoroughly stimulated, equally by being an actor in the space and a viewer. However, the degree of difficulty of successfully navigating the key into a position that would trigger a sound was higher than expected. It demanded almost total concentration, especially when there were two or more players in the space. As a result, the density of sound was less infrequent than expected. Even when there were 3−4 players in the space simultaneously, the density of sound was still quite low. However, they loved the sounds. These consisted mainly of slightly modified and extended percussion samples, each with its own distinct timbral and rhythmic 'personality'.

- to create a multi-user collaborative creative space and to provide a socially engaging experience that promotes user-user interaction strengthening the sense of community of activity and place.

This aim was also met with partial success: the exertions of the players in the space always received positive vocal support from the 'audience' (players-in-waiting). This is, however, only one aspect of a collaborative space; the other is players in the space actually interacting to create something they could not do alone. Due

to the level of difficulty mentioned above, this type of interaction was limited to the occasional rare moment. Furthermore, the difficulty experienced in the maneuvering one's key was increased by the presence of the players in the space as some players became aware of being crowded; they actually preferred to be the sole user of the space, a solo player as it were.

- to create an activity visually engaging, enhancing the players' experiences and promoting their ability to locate themselves and to decode the events of others in the space at the same time.

In this we were entirely successful. Many of the visitors expressed joy in being able almost effortlessly to move their key about the screen even if precise control was a little difficult. Some players liked to move across the 'threshold' in to the space, see their key appear, hear their sound, only to step back out and to repeat the process again and again.

6.2 What improvements are required?

It is, at first attempt, quite difficult to manoeuvre one's key into a lock; it is even more difficult for first time users with combined learning and mild physical difficulties. In the first instance, we want the installation to provide and incentive for people to play and create with sounds. Therefore it needs to be learnable and masterable: to be sufficiently complex that players may, in time, enjoy the experience of "getting better" at using it, but, at the same time, sufficiently easy that beginners may quickly experience some success. In order achieve this, we clearly need to find a way to adjust its level of difficulty. Perhaps a way forward here is to define an 'automatic acceptance zone' of a fixed number of pixels around each lock. Once a key is within this zone, it automatically 'pops' into place and triggers the sound. This will facilitate the triggering of sounds.

Another improvement that must be considered is a diminution of the difficulty encountered by a wheelchair player, alone or being pushed around the space. There are two main problems: 1. The wheelchair appears as a rather large glob, something which interferes with the accuracy of movement, particularly at the moment of the fit. 2. If a person is pushing the wheelchair their glob often separates itself from the player, spontaneously creating a new key; this confuses the system. Again, some means of adjusting the level of difficulty could solve this problem., for example by making the correspondences between glob size, key size and lock size easily – in real time – adjustable. We believe with some small adjustments in this domain of the software, the environment will succeed in fully delivering the wealth of highly enjoyable creative resources that it promises.

Due to other inadequacies of the software, related to the one discussed above, which we were not able to resolve due to lack of time, we were not able properly to evaluate the contribution of the key-to-key feature to the strengthening of social interaction in the space. The key-to-key interaction presents a particularly knotty software problem because the key (a virtual object) is a representation of a real object (the body) in real space. In order to fit the keys to one another, the bodies – the globs from which they derive their existence − must collide. When the globs collide they lose their identity. In order to solve this problem, we must carry out a fundamental revision of the recognition/avatar generation process. One way forward would be to make the avatars larger than the blob in the vertical dimension and to use the top and bottom profiles for defining a fit. We are currently working on this.

7. REFERENCE

R Gehlhaar, "SOUND=SPACE: An Interactive Musical Environment," *Contemporary Music Review*, Vol. **6**, No. 1, 59–721991) and *www.gehlhaar.org*