

Experiences from video-controlled sound installations

Anders Friberg
Speech, Music and Hearing, KTH
Lindstedtsvägen 24
10044 Stockholm, Sweden
afriberg@kth.se

Anna Källblad
Studio 323 wip:sthlm
Strahlenbergsgatan 17
12145 Johanneshov, Sweden
anna.kallblad@bredband.net

ABSTRACT

This is an overview of the three installations Hoppsa Universum, CLOSE and Flying Carpet. They were all designed as choreographed sound and music installations controlled by the visitors movements. The perspective is from an artistic goal/vision intention in combination with the technical challenges and possibilities. All three installations were realized with video cameras in the ceiling registering the users' position or movement. The video analysis was then controlling different types of interactive software audio players. Different aspects like narrativity, user control, and technical limitations are discussed.

Keywords

Gestures, dance, choreography, music installation, interactive music.

1. INTRODUCTION

At the core of this work is the collaboration and flow between artistic and technical ideas and knowledge. In particular, we are interested in the relation between gestures and sound both in terms of the resulting sound and choreography. A classic example of such a collaboration is the New York 69th Regiment Armory in 1966 where 10 New York artists and choreographers worked with 30 engineers and scientists from Bell Telephone Laboratories to create groundbreaking performances that incorporated new technology [6]. Since then there have been a number of productions using dancers to control the music, i.e. [11][12].

Previously, we explored the possibility of using video cameras for analyzing gestures in conjunction with investigations of musicians' gestures as well as in previous interactive models including an interactive, collaborative game, Ghost in the cave [10] and a gesture controlled conducting system [5].

This is an overview of the authors' recent experiences with the three installations Hoppsa Universum, CLOSE, and Flying Carpet. We will here present both an artistic point-of-view, some technical methods and descriptions, as well as discussing the interaction between artistic goals and technical possibilities.

2. USER INTERACTION

The intention of these installations has been that the user will explore it without help or guiding. Also to let the user focus on her/his body and movements rather than to focus elsewhere, for example on a screen or device, which tends to steer away the perception of one's own body. There is also a conceptual artistic reason for this, namely the nature of the work in that it

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

NIME'11, 30 May–1 June 2011, Oslo, Norway.

Copyright remains with the author(s).

only exists when someone uses it. This might be the essence of dance: it only exists in the moment, nothing is left, only the experience and memory in the users/dancers' body. The purpose is the sharing and the experience in the moment, and the memory of that experience.

This self-exploration concept needs careful consideration in the technical design. Also, since this type of sensing does not involve any artifact or physical contact with an object, the natural ecological connection between sound source and object is broken. Although this is common in contemporary society with electronic devices it makes the interaction less intuitive. In particular in this case where the object is missing. One implication is that the response time of the system must be fast so that the user easily can associate a gesture with a certain change of the resulting sound.

There is a narrative created by the visitors own body movements in time. Through the interaction design one can suggest and/or control this narrative by facilitating certain choices of body movements and locations in the room. We were looking for an interaction that inspire movements that evoke certain feelings, and tell stories by how users movements develop individually and as a group. By changing the user interaction and the sound material over time (for example in a time cycle) a further enhancement of the narrative dimension is obtained. This will avoid the often static sound environment otherwise often resulting from installation work.

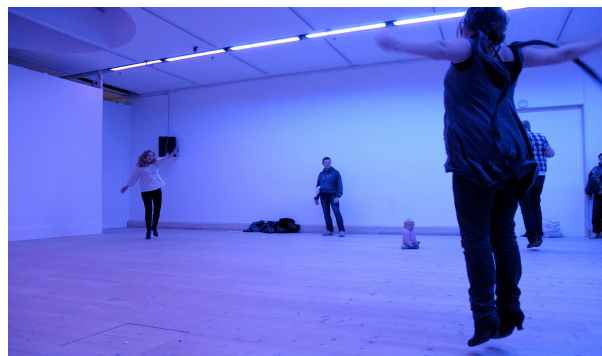


Figure 1. Hoppsa Universum during installation.

3. INSTALLATIONS

3.1 Hoppsa Universum

The artistic idea was to create a "magic room", where music and the light changed when you moved around. From looking at children's free movement, a choreography for five dancers was made for which the composer wrote and recorded the music. The music was then decomposed into five different scenes each with its own interaction scheme. Each scene featured different trigger areas that enabled certain choreographed movement patterns in the room. Both the sound and the light were controlled by the visitors movements. It was set up at Botkyrka Konsthall, see Figure 1.

The technical setup of Hoppsa consisted of 4 analogue video cameras in the ceiling covering the whole floor and 1

camera on the side for vertical sensing. All cameras were using infrared light. The length of each scene changed automatically providing a 20 min long narrative in total. A more detailed description is found in [9].

3.2 CLOSE – your body her voice

In this installation users trigger Palestinian women's recorded voices, see figure 2. CLOSE premiered 2009 at Sakakini Art Centre Ramallah and was then exhibited at Södra teatern Stockholm, Gottsunda dans och teater, Uppsala and Haninge konsthall 2010. We searched for a form of interaction that would evoke feelings of being limited, controlled and place users "involuntary" on intimate distance of each other in order to create a visual image of people bunched together. The process and the final artwork can be described as a physical interaction between ideas and context.

The background for CLOSE was the art project "Open Workshop for Culture and Arts" by ten Palestinian artists living in Gaza, Palestine and Israel with focus on the Palestinian women's public position from a feminist point of view, which resulted in a 20m long mural of which half stands in Ramallah and half in East Jerusalem. It shows women's full bodies in shapes inspired by literature references, an unusual sight in this context. Co-author and choreographer Källblad got the idea of a transient mural to connect live bodies and recorded voices. A mural is an ancient form of storytelling and usually through its construction fixed in one place and shape. The CLOSE mural would be the opposite. Its substance and meaning would appear when activated and stories would be retold in different orders and places depending on peoples movements.

In CLOSE the group had to cooperate and stay very close to each other in order to hear certain sounds. The trigger areas of these sounds were vaguely indicated in order to make people have the experience of searching and cooperating with strangers. One reviewer wrote "a practice in group dynamics" [3]. The language, Arabic, (translations were available) contributed to the context as well as available information of where the voices were recorded.

We used a laptop and webcams to facilitate travelling in restricted areas, thus the political situation dictated what technique we used. When travelling and collecting material we were daily affected by the occupation of the West Bank through roadblocks, shortage of water, questioning by soldiers etc. Long waiting for extra microphones due to difficulty in passing roadblocks etc all effected the working environment practically and emotionally. The feeling of constraint was overwhelming. We experienced how outside factors were filtered its way into the work. The outer context gave the form. Back in Stockholm building CLOSE we looked for an activation that would trigger the same emotion and image. The shape of the work giving its content, and vice versa.

The technical setup consisted of two webcams, laptop, 4 loudspeakers, and a carpet. The webcams were put in the ceiling making an approximate analysis of the number of people in each of 24 different zones using background subtraction (see below). A relative measure estimating the number of people in each zone was computed from the size of the blob, see Figure 4. When a visitor entered a zone, a single short sound was played. These sounds were short processed excerpts from the original recordings. When a group of visitors strolled along the carpet randomly, these sounds made a rather annoying and chaotic sound landscape. When the visitors cooperated and gathered in groups in a certain position, a single song or story was presented. This was determined by a threshold value of the estimated number of people.

Each group of visitors (10-30) was led into the exhibition area for 45 minutes. They were told that there was a carpet on

which they could activate a soundscape where 15 single voices were hidden. To hear a single voice a number of people had to stand very close to each other. A guide was available to answer questions and could on demand hand out a map indicating where each voice was located. A typical action during the 45 minutes was that people first moved around quite fast and individually, then stepped out for a while to watch others, some then entered again and gradually started to make small groups in different areas and then finally forming one group that moved around with small steps and stopping to hear the single voices.

After each presentation the visitors could fill in a questionnaire with 15 questions consisting mostly of dichotomous Likert scales coded from 1 to 6. The overall impression (bad-good) was rated very positive with a mean of 5.5. Despite some comments relating to a frustration in the beginning they seem to have understand rather well how it worked and rated it (scale easy-hard) with a mean of 3.0. Three questions relating to cooperation issues and getting close to each other were all positively rated (being close: negative-positive, $m=5.1$; cooperation: difficult-easy, $m=4.3$, frustrating-rewarding, $m=4.7$). The visitors in Ramallah and at Södra teatern responded in a similar way with only small differences. Thus, in the end they got a positive experience of being close although there was some initial frustration in the beginning.



Figure 2. CLOSE at Södra teatern. A group of visitors are gathered in a group to hear a particular voice.

3.3 Flying Carpet

This is a dance/DJ installation that was commissioned for the Art's Birthday Party at Södra teatern, Stockholm, 2011. It was installed in a room next to a bar and was intended as a replacement for a typical dance floor in a club, see figure 3. In a club setting it is the DJ that is trying to play music that will enable people to dance. Here we turned that concept upside down, thus, it is the people that control the music instead, with the effect that the people had to make some effort to hear the music. It is an extra challenge to change such intuitive and well-known paradigms so we thought about how to help people to anyway be enabled to move more and attract more people to interact. Following quote was posted on the wall as instruction and inspiration:

Remember:

"There are only two types of people in the world
the ones that entertain and the ones that observe"

Britney Spears

The video analysis was rather simple. Two webcams (Logitech Pro 9000) were placed in the ceiling above a carpet about 3.5X5 meter in size. Each camera extracted the quantity of motion (QoM, see below) and the centroid position of the motion in terms of x and y coordinates. The QoM value was controlling the overall state of the system divided into four

different interaction states using overlapping fuzzy functions (see i.e. [4]). Each camera controlled independently one channel of the resulting stereo audio output.

The first system state is when nobody is present on the carpet. Then a soft simulation of a wind sound is played.

The second state is when one person enters the carpet. This is a scratch mode using different samples about 10 s long taken from the record *Movies for your ears* by William Brunson [1]. The current sample is scrubbed (or scratched) with the x coordinate while the y coordinate is controlling the pitch. The sound is processed using a phase vocoder implemented by Ben Saylor and available in pd-extended version 0.41.4, [7].

The third state is the “disco” or dance mode. DJ Nico compiled a complete set consisting of three hours of electronic dance music. This music was continuously played but was not heard unless more people entered the carpet. The music was filtered with a bandpass filter using a variable q-value and centre frequency. When the QoM was relatively low the q-value was high resulting in a narrow band of music in which the filter frequency was controlled by the x-position. This resulted in a kind of phaser effect controlled by the movements. When QoM was relatively high the q-value was reduced resulting in the original full version of the music.

In the fourth state we attempted to further intensify the dance music by adding isolated voice samples taken from *Movies For Your Ears*. These were selected randomly and triggered on peak QoM values.

Considering the challenge of inspiring people to get involved we filtered the QoM signal so that its raise time was as short as possible and the decay time was fairly long, about 2 seconds. The fast response made the music easy to trigger and to start and the slow decay made it easier to maintain a high musical energy without too much physical effort. The intention was to do something similar to “spinning plates” tricks, thus, it was only necessary to make short “injections” of energy to get it running.

The informal feedback from the audience was very positive and many of the visitors that overcome the hesitation to try danced with much energy and big movements for a long time. Several people even suggested that we should apply for a patent.



Figure 3. A picture from the realization of Flying carpet at Södra teatern, Stockholm

4. TECHNICAL SETUP

4.1 Camera position and light

These installations were all made such that the users/visitors were free to move around and with a varying number of people. In this context, our preferred position is to place the video cameras in the ceiling pointing straight down. This makes it possible to register individual movements and to divide the floor in different areas, thus, facilitating using the room and

space as part of the installation concept. It is obviously better the higher the camera is mounted since it reduces the differences in the video projection of people due to the view angle in the picture. Therefore the ceiling height is an important parameter and often a limitation when choosing installation space. It also affects the user experience. When the cameras are low the people recognize them fairly easily and can try to “cheat” by directing movement toward the cameras. If the cameras are located rather high the users might not understand how it is working, thus, creating a more magical experience.

The light will interact with the cameras if not infrared filters are used. Preferably, the light in the room should be either diffused i.e. evenly spread out like a normal office space or the light sources should be placed in the ceiling close to the cameras. In this way the influence of shadows is minimized.

4.2 Cameras

We used either analog cameras designed for surveillance (Hoppsa) or webcams. The analog setup has many advantages. It is well-known reliable technology that has been used for a long time. It has fast response time, demands little computational power and there is not any problem with cable length. Since the cameras were also designed to automatically switch to black and white night vision with led lights we used infrared filters so that we could change the light in the room without too much interference with the motion recognition. Thus, in the Hoppsa Universum installation we used 5 such cameras connected to two capture cards in one desktop Windows computer. All the processing including zone division and QoM computation as well as all audio computation with four-channel output was running on the same computer without problems.

For the other two installations we used a lighter setup with two webcams connected to a laptop. This makes the whole setup very portable and actually fits in a backpack. Active USB cables makes it possible to have cables with sufficient length (about 20 m). The cameras used (Logitech Pro 9000) have a zoom controlled by software. This facilitates the adjustment of the active sensing area during installation. However, using USB webcams with their software drivers can give unpleasant surprises. For example, it is often difficult to turn off the automatic light adjustments and sometimes this is not even possible (i.e. the built-in camera in Mac Book Pro). Cheaper webcams are not recommended since the background noise is often too high. The resolution of the camera is less important due to the inaccuracy in the analysis. Thus a common VGA resolution or smaller is sufficient. A third possibility is to use firewire for the camera communication and there are currently many small such semi-professional cameras available.

4.3 Video analysis methods

An overall design goal has been to keep the video analysis as simple as possible. Thus, only consolidated techniques have been utilized that are easy to understand, easy to interface and demands rather small computational power. We have used the EyesWeb program, v. 3, for all video analysis using existing analysis blocks, [2]. Two basic methods have been used for analyzing the visitors in the video picture.

Frame difference. The simplest and most robust technique is to compute the difference between consecutive frames. In this way all movement is registered but not a person standing still. A further improvement can be made by adding a second movement analysis with the black and white video inverted. From that initial extraction the total area is divided in different zones and for each one the quantity of motion (QoM) and the position is computed (centre of gravity for the motion blob). QoM is defined as the size of area of the frame difference after

some thresholding, see [2]. This was used in Hoppsa and Flying carpet.

Background subtraction. An alternative method is to subtract a still picture of the background from the incoming video. In this way the whole area of each person is visible independent of if they move or not. This was used in CLOSE, see Figure 4. It feasible to use in an installation even during a longer time but it is more sensitive to changes in light and has to be more carefully calibrated.

Obviously, using these techniques it will only work if there is a difference in brightness between the user and the background. Usually, a light carpet color is in many cases better than a dark one. Often people (at least Swedish art-oriented ones) tend to have dark or often only black clothes. Alternatively the carpet could have a pattern of both dark and light colors. This would make the analysis independent of clothing or skin color.

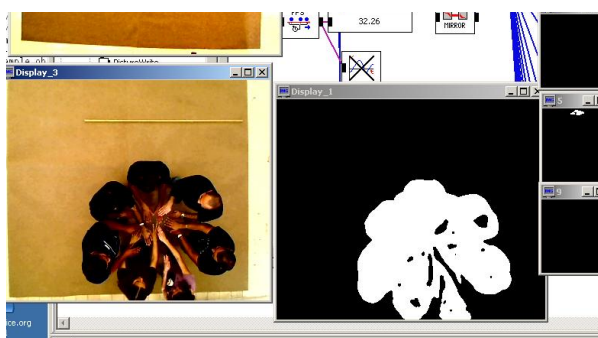


Figure 4. A screenshot of the video analysis during the testing of CLOSE. Left display is the raw video and the right display shows the detected area in white.

5. DISCUSSION

It has been challenging to find workarounds for the artistic ideas due to the technical constraints. However, these constraints can also be considered in the same way as budget and time schedule limitations - what is possible to do within these parameters? Then one has to try to be as clever as possible and to find a form that is interesting in itself so that the constraints serves as and can be used for artistic purposes. There is also joy from “beating the odds” that feeds energy to a project, as well as when artistic ideas can spur technical solutions.

Was the technical level enough for realizing the artistic goal? The answer depends partly on how open-ended the artist intended the work to be. Due to the self-exploratory design of the installations, the coupling between gesture and sound had to be immediate and intuitive. This made it necessary to constrain the interaction to rather simple models. Our experience is that even very simple actions like triggering a sound on a certain position is not grasped by every user, in particular, if there are several users active at the same time. In this view it is hard to see that more advanced video analysis would substantially improve the user experience. For example, in Flying Carpet even a simple division of QoM can make the system behave in four different states depending on the number of active users. There are from an artistic point of view pros and cons to the users’ fully understanding how the triggering works, a certain “mystique” can be intriguing.

We found that the fine-tuning of the different parameters was a crucial step that needs to be carefully considered both during testing and development and, in particular, at the final location, considering both the type of room and audience.

Overall the visitors seemed quite willing to lend their own bodies to fill or activate these installations. Their level of awareness or opinion that they themselves by this activity in fact make the artwork or performance remains unsaid.

A different approach is to use the video gesture analysis material directly and after further video processing project it as a part of the background or scenography. This has been explored by Frieder Weiss in a number of applications including dance performances [8]. An interesting future extension would be to combine the music interaction with light control using video processing and data projectors.

6. ACKNOWLEDGEMENTS

Hoppsa Universum by Anna Källblad in collaboration with Niko Rölke, music; Karl Svensson, light design; Anders Friberg, audio-video interaction; Tove Axelsson, set design; Linda Adami, Kerstin Abrahamsson, Johanna Klint, Maryam Nikandish, Stina Nyberg, dance. Supported by Swedish Arts Council, City of Stockholm, Stockholm County Council, City of Botkyrka, University of Dance and Circus, and Moderna dansteatern.

CLOSE by Anna Källblad and Annette Taranto in collaboration with Amal, Asmaa, Oraib, Rasha, Rula, Ruba, Majdal, Juma, Nadia, Zeynab, Leila, Kamelia, Mohah, Fatima, Rowan, El-Funoun Dance Troupe, voice; Jaime Fawcus sound-engineering, Anders Friberg audio-video interaction; Karl Svensson light design; Mervi Junkkonen, documentation. Supported by Swedish Arts Council, The Swedish Arts Grants Committee, and The Swedish Institute, City of Stockholm.

Flying Carpet by Anders Friberg and Anna Källblad in collaboration with Bill Brunson, music samples, DJ Nico, music compilation. Commissioned by Södra Teatern.

7. REFERENCES

- [1] Brunson, B. Movies for your ears, Electron Records, 2004.
- [2] Camurri A., Mazzarino B, Volpe G. Analysis of Expressive Gesture: The EyesWeb Expressive Gesture Processing Library. In Gesture-based Communication in Human-Computer Interaction, LNAI 2915, Springer Verlag, 2004.
- [3] Claeson, A. Uppsala Nya Tidning, May 6, 2010.
- [4] Friberg, A. A fuzzy analyzer of emotional expression in music performance and body motion. Proc. of Music and Music Science, Stockholm 2004, 2005.
- [5] Friberg, A. Home conducting: Control the overall musical expression with gestures. Proc. International Computer Music Conference, (pp. 479-482), 2005.
- [6] <http://www.9evenings.org>
- [7] <http://puredata.info/downloads>
- [8] <http://www.frieder-weiss.de/>
- [9] Källblad, A., Friberg, A., Svensson, K., & Sjöstedt Edholm, E. Hoppsa Universum – An interactive dance installation for children. Proc. of New Interfaces for Musical Expression - NIME, Genova, 2008.
- [10] Rinman, M-L., Friberg, A., Bendiksen, B., Cirotteau, D., Dahl, S., Kjellmo, I., Mazzarino, B., & Camurri, A. Ghost in the Cave - an interactive collaborative game using non-verbal communication. In Gesture-based Communication in Human-Computer Interaction (pp. 549-556), Springer Verlag, Berlin, 2004.
- [11] Siegel, W. and Jacobsen, J. The Challenges of Interactive Dance: An Overview and Case Study. Computer Music Journal, 22(4), 29-43, 1998.
- [12] Winkler, T. Motion-sensing music: Artistic and technical challenges in two works for dance. Proceedings of the International Computer Music Conference, 1998.