# Sound People – Sonic Environments Reflecting Intersubjective Spatial Relations

Sune Hede
Information Science, Aarhus University
Helsingforsgade 8
Aarhus, Denmark
sunehede@hotmail.com

Jakob H. Andersen
Audiodesign, Aarhus University
Helsingforsgade 8
Aarhus, Denmark
jakobhandersen@hotmail.com

## ABSTRACT
In this paper, we describe the development of the interactive installations Sound People and our experience with its implementation. The installation seeks to amplify spatial relations in the public space by reflecting them closely in a sonic environment. Central to our work is the question of how to couple spatial relations to the sonic environment in a both intuitive and challenging way. As we work with camera tracking, this requires exploration of spatial relations that are perceivable and meaningful to both people and visual tracking technology. We experiment with and evaluate couplings of space and sound, and implement them in different sonic programs for the installation. Finally we evaluate the concrete implementation of the installation and consider questions of dependency and relations of installations to their surroundings.

## Keywords
Public installation, sonic feedback, responsive environments, amplified reality, HCI

## 1. INTRODUCTION
Despite the permanent near physical presence of other people in the urban environment, we observe that interaction and communication between strangers in public spaces often is confined to situations of consumption or at institutionalized meeting places such as night time in bars and discos. The constant physical presence is at the centre of Bourriaud's [7] notion of modern urban life, which he argues forces the arts to have the presence of the others and the intersubjective exchanges as its main theme. Art must according to Bourriaud create intermezzos or durations with rhythms in contrast to those communication practices forced upon our everyday life.

We wanted to make visible – and challenge – this practice of being physically close without any exchange or interaction, and we made the installation Sound People, which seeks to amplify the existing spatial relations between people by making the sonic environment reflect these relations very closely. Thus Bourriaud's diagnosis of postmodern urban life serves a twofold purpose in our endeavour. First it is very much in line with our normative motivation of focusing on the dualities of private and public, exchange and isolation, physical presence and communicative absence, and second it also provides us with a useful general definition of art's role in society.



## 2. RELATED RESEARCH AND INSTALLATIONS
The research on public interactive installations shows many examples that seek to make people interact with each other, either reciprocally by playing a game for two or more players (Finke et al. [4]), or indirectly where one person acts as user and others as spectators (Reeves et al. [2], Dalsgaard & Hansen [5], Bilda et al. [3]). Common to the above is that the interaction is mediated trough a system that is somehow localized in the physical space in the form of a screen or other in/out devices and artefacts. This places the system as a medium or a window for interaction, and direct contact or attention between the participants' physical presence is thus often a kind of secondary interaction. Because we focus on the physical presence of others we have tried to minimize the systems distraction from this. Our installation is of course still functioning as a medium, but as sound is more of an omnipresent medium, you can still look at, move towards and listen to a person while receiving the sonic output from the system. These properties and their related possibilities and problems in public installations are treated by Schloss & Stammen [1], who point to the fact that sound is inescapable as you cannot close your 'earlids' or *hear away* if you desire to. This has impact on the interaction as well as on the relationship between the installation and its surroundings.

Of installations using video tracking to produce a sonic environment it is worth mentioning David Rokeby's *Very Nervous System*, and Marie Sester's *Access*. Very Nervous system very closely reflects the smallest movement of the user. This work has developed over time, and appeared in different concrete implementations. In a 1991-version [9], the sonic environment consists of a lot of different instruments playing different tones and harmonies depending on how much and where the user moves. This musicality of the work is heavily countered in a 2009 version (experienced in Herning, Denmark), where all tonal material is gone and the sonic environment reflects movement by producing sounds as if you walked through an underwood of nails, screws and other metal parts revealing each tiny move, with the sound of metal falling to the floor or being crushed under your feet. This experimentation with degrees of musicality has also inspired the development of Sound People.

The way we seek to give roles to unknowing passers by is also seen in Marie Sesters work *Access* [8], where a spotlight follows randomly chosen person around in the public space, not

letting them escape the spotlight, while they increasingly attract other people's observing attention.

## 3. TECHNICAL IMPLEMENTATION

The installation was part of Quonga-Fest: a festival for theatre, music and art in the city of Aarhus [10] and the installation was made outside the main theatre building with the camera view spanning the whole breadth of the walking street.

A camera was placed 7 meters above the ground looking vertically down at the street. The live video signal was then fed to a computer-system inside on the top floor running a tracking algorithm and audio synthesis in Max (Max/MSP/Jitter), which roughly consists of parts illustrated in figure 1.

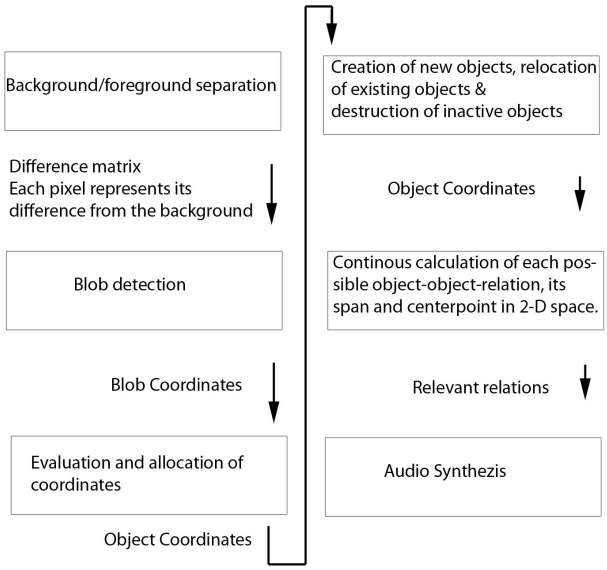


**Figure 1. Rough schema of video/audio flow.**

The audio was played through two ordinary home-stereo speakers, which were both placed in the same location, but each

pointing to a different end of the street (see figure 2). This was not to create a perceived stereo effect, but to be able to pan the sonic output of each relation to the area where the relation is situated.

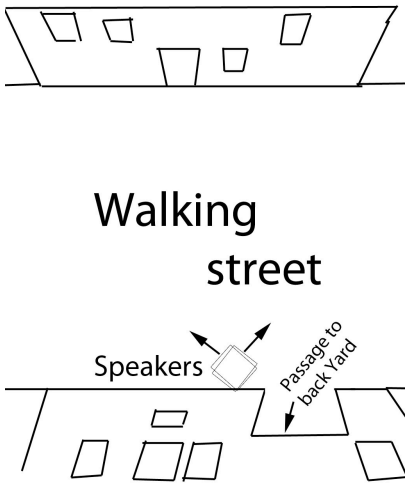


**Figure 2. Sketch of street view – vertical from suspended camera.**

## 4. DESIGNING INTERACTION

This section of the paper concerns our approach to designing concrete sonic implementations. In what way should the presence (or actions) of people be linked to a sonic expression?

In the development process, it was essential to us that we were able to test and experience different connections between human behaviour and sound. To this purpose we developed a testing tool in Adobe Flash. This testing tool essentially simulates the camera's point of view – a square in which you are able to place and move imaginary people (represented by coloured circles) around.

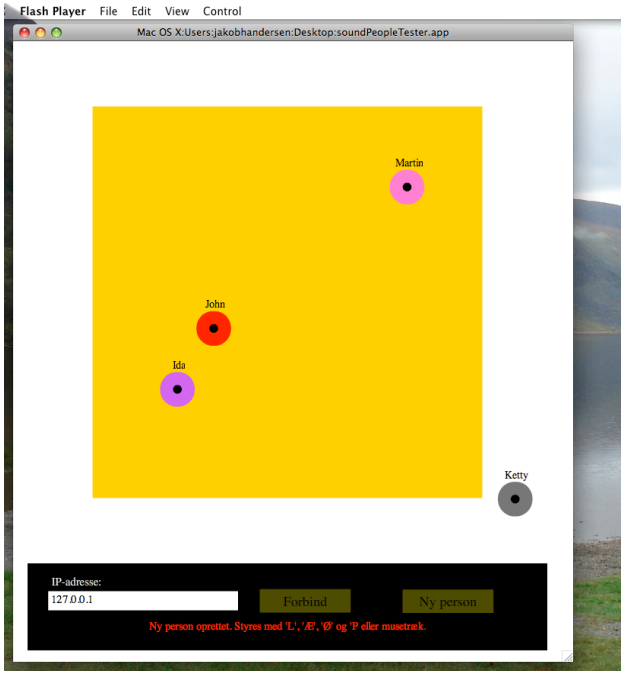


**Image 1. Screenshot of the testing tool in use.**

The data concerning the persons' positions are sent to Max where we are able to experiment with different sonic 'reactions' to this incoming data.

Obviously, it is not the same experience as it is to be present with your whole body, interacting with the actual physical environment. Nevertheless, this prototyping tool has been very useful to us in the process of developing concrete 'sonic programs' for the installation.

### 4.1 Human relations as input

To minimize the feeling of interacting through a medium, the installation has no distinct physical control surface, but only takes peoples position on the 2-D surface (the ground) as its input. Since our main motivation is to comment on, or even challenge, the way people relate to each other in the public space, we decided that the sound generation should reflect relations more than individuals. Thus the interaction with the installation was not to be dependent on the participant's absolute placement in the 2-D space, but only on his spatial position relative to other participants, and the only interactional reference points for the participant will thus be the other participants. We therefore needed to find and define suitable relations that were decodable for the people interacting as well as for the computer.

### *4.1.1 Relation types*

As human beings, we are capable of recognizing a truly impressive amount of relations in any encounter between people. Some of these relations are recognized by the sensing of very subtle variations in physical patterns and rely on interpretations informed by social, psychological and cultural experiences, not to mention empathy. For instance, we are often capable of recognizing different status relations between people at just a short glance. Our installation is not able to recognize such relations, since all we know about each person is the person's position. However we are capable of decoding different spatial relations.

The relation type with which we have been working the most is the distance between two people. It is a relation that is quite easily decodable for the computer and it is a relation that has a strong social and psychological significance for most people.

Another relation we have been working with is the angular orientation of the imaginary line between two people. This relation is greatly affected when people pass each other or if they rotate around each other.

Whereas the two above-mentioned relations can be categorized as one-to-one relations, we have also been working with one-to-many and many-to-many –relations. Partly because of the fact that the number of one-to-one relations quickly becomes very large as more people enter the installation.

An example of a many-to-many relation we have been working with is the degree to which all participating people are standing on a line.

The distance from one person to the closest other person is an example of a one-to-many relation. Though it could also be thought of as a one-to-one relation, it is based on one person's relation to all of the other participants, who potentially can receive the role of being nearest.

### *4.1.2 Known couplings of space and sound*

As we wanted the sonic environment to reflect the spatial relations in a way that is as easy to perceive as possible, we wanted to take advantage of existing couplings of space and sound, which we experience in our everyday. As a sound source comes nearer, the sound is experienced louder, with a brighter timbre and lower reverb-to-source-ratio. We also experience Doppler effect, when a sound source is moving relative to us[1]. This general experience of sounds (in general most experiences) being richer and more intensive when we approach them establishes a spectrum from poor/un-nuanced at a distance to rich and intense when close. We can use this abstract categorization to couple more unfamiliar sound phenomena to this scale. We can for example couple rhythmic patterns to this scale by making them more intense when close, which isn't something we normally experience in the real world, but which suits the close-is-intense logic. This way we have used both concrete familiar couplings of space and sound and couplings derived from more abstract relations.

## 4.2 Concrete sonic programs

We do not think of the installation as a closed and finished artefact, but rather as an ongoing process of exploring and experimenting. In this spirit, we have been working on different concrete sonic programs for the installation. By *sonic program* we mean a concrete setup of mapping between input (people's relations) and the resulting output (the sonic response). In the following subsections we will briefly present the two sonic programs we used at our case setup at Quonga-Fest.

### *4.2.1 The 'City Hall' program*

In this sonic program we have sought to create a mapping that, whilst sonically commenting people's relations, maintains a musically 'meaningful' whole.

The basic input of the *City Hall* program is the one-to-one-distance relation between each participating person. The basic musical fundament is a melody. Each time two persons get within a certain threshold of distance (almost touching each other), the next note in the melody will be triggered. That is, if two people get close enough, they will hear a bell-like synthesizer playing the next note in the melody and thus we move forward through the melodic sequence. The melody we chose for this program is the melody[2] that is also played by the bells of the city hall in Aarhus. In addition to this triggering of melody notes, each one-to-one-distance relation is controlling the expression of a specific continuous 'harmonic voice' following the sequence of the melody. These harmonic voices might have a clear feeling of tone/pitch or they might be sounds possessing more rhythmic or noisy characteristics. However, they are all composed to complement and move along with the sequence of the melody – like different voices (soprano, alto, tenor, bass…) in a western choral song. In concordance to a principle of interactional consistence, we tried to let the individual harmonic voices change (note-wise) as little as possible throughout the sequence.

So as two people approach each other, 'their' continuous harmonic voice will become increasingly more audible and it might also intensify on other sonic parameters (timbre, tempo etc.) . If they get close enough they might trigger the next note in the melody and hereby also the corresponding step in the harmonic sequence.

### *4.2.2 The 'Krueger Tribute' program*

Whereas the *City Hall* program is focused on maintaining a somewhat predetermined/composed musical expression, the *Krueger Tribute* program is not in this way concerned about any 'musically meaningful'[3] whole.

In developing this sonic program we have been inspired by Myron Krueger and his thoughts expressed in [6]. Krueger argues that interactive systems call for a new field of aesthetics concerning interaction itself, where *response is the medium* and locus of the aesthetics and thus *visual and auditory beauty is secondary*. In this program, we wanted to direct all response to emphasizing the relations between people with very concise coupling of spatial relations to the sound. Therefore we build the sound sources entirely from scratch within the Max environment, not involving any musicality-infected plugin-instruments (except for reverb) so that we were able to control every parameter of the sound. The tones were not selected at total random, but were restricted to have a ground frequency equivalent to a piano note.

---

[1] Build on the physical model of Doppler effect, but with an unnatural low speed of sound factor, to exaggerate the effect.

[2] A very old melody of unknown origin and date. First printed in *Piae cantiones,* 1582 Finland, as *Carmen Vernale*. A rather commonly used melody in Finnish, Swedish and Danish song traditions.

[3] Musical meaning is here understood in a traditional and conservative manner. Indeed it could be argued that, from a more modern perspective, the very lack of compositional control could constitute a musical meaning.

We used some of the above-mentioned known couplings of space and sound, and let distance between to people affect volume, timbre and reverberation of the sound representing their relation. In moments of change in the distance between people, the affected relation's sound would suffer from Doppler effect. Another continuous coupling was from the degree to which people were standing on a line, to a very significant FM-synth sound. When people were in situations close to being on a line, the FM-synth sound would appear and evolve as a sweep to a wildly roaring sound if the overall deference from the line further decreased. If only two people are present, they are of course always perfectly on a line, and would thus generate no response; just a static – and therefore confusing and non-informing – sound. Instead, when only two people were present, the being-on-a-line-relation was reformulated into whether the two people constituted a line *straight across*[4] the street. This would sonically result in an FM-sweep if two people rotated around each other, or if people passed each other along the orientation of the street. To allow for further exploration and creative interaction, we wanted to make it possible for this installation to evolve triggered by the users. When the constellation of people approached the state of being on a line, the roaring sound gets very significant and dominant over the other sounds, and we used this as a signal for something bigger to come. When the roaring sound is at its highest, all existing relation-sounds are shuffling through the possible pitch-values they can acquire, which causes all relations to be represented by new sound, and suddenly a new overall tonal harmony of the installation is established.

# 5. EXPERIENCES

It was very difficult to retrieve stable tracking data from the very uncontrolled visual environment, so the feedback was not always clear and consistent. We experienced, that a little inconsistency could cause much confusion. When for example 4 people were interacting with the installation and one of them wasn't consistently tracked by the computer, this would generate much uncertainty among the participants regarding how the system actually worked.

This recurring inconsistency made it hard for especially newcomers to perceive the actual mechanisms and essence of the work. On the other hand, when somebody figured out the system, we observed that they often tried to attract other people into experimenting with it. In these cases it really became clear that interaction in public spaces has an important performance-aspect as described by Dalsgaard & Kofoed, as people would often create attention to the connections between people, exaggerating them by playing bull and bullfighter or throwing themselves into each other's arms.

We also observed people walking determinately towards complete strangers to create attention to their sonic connection. It apparently seemed legitimate to do so, and thus the social communicative norms were temporarily different from those normally effective in this particular space. In moments like these the installation functioned exactly like we strived for, creating the durations and intermezzos from the prevailing norms of communication and social intercourse. To Bourriaud the desired intermezzo is not so much in the direct experience of art as in the reflection on the artwork. We also witnessed discussions among participants, who figured out the functioning of the system, and they often discussed the installation as a comment to society. So even though the installation often was working far from satisfactory, it also succeeded in periods to make the two kinds of intermezzos: the interaction experience and the more reflective one.

Using sound as a medium proved to let people focus on each others presence while still perceiving the sonic feedback. Though we have not been able to circumvent the fact that our speakers still are a perceptible *source* located in space, we did not experience that users devoted any attention to the sound source itself after figuring out how the interaction worked.

The ambient properties of sound also proved to pose problems. When the system was running the interaction oriented program, the people not interacting but still within hearing range were exposed to sound of annoying character, as the aesthetic quality wasn't designed to be in the sound, but in the interaction.

Along the same lines we found that the interaction-oriented program didn't work well when feedback wasn't very stable and concise, because the aesthetic qualities was utterly dependent on this.

The fact that the area outside the passage (see figure 1) was often used for smoking and chatting during breaks between shows also proved to be a tricky social environment for the installation to interact with. The large amount of people crowded together on little space made it hard for participants to discern the parts of the sonic environment and to perceive which physical relations they were coupled to. In chaotic situations like this, the installation was often not perceived as anything else than just background noise, some people even thought it was totally randomly generated sound, and one person even expressed that he liked it like that as a work of art.

This person has evidently perceived it as a work of art, but just not the artwork we intended. This leads us to talk about what makes the installation successful. Our installation is of course very much focused on the experience of making sound together with other people, but as we also have a motivation about making visible – and challenging – culture, we also want to present the experience as some entity invested with intention: an *installation*. In museums the experiences are expected represent intention-laden entities, because everything on and between their white walls per definition is such entities. In the street environment though it is necessary to design the installation to represent such an entity.

Our goal was to make the coupling of physical relations to sonic relations sufficiently intuitive, so that unknowing passers by would easily realize this. The intention behind this new necessarily crafted connection (the sound) is by the newcomer attributed to either the person approaching him or (if he realizes that there is a crafted entity behind) the installation. To incite reflection we wanted to show that the experience is coupled to the crafted installation, which is a product of our reflection. After the first day, we realized, that few people attributed this sound experience to anything other than a stereo playing a psychedelic music piece. The interaction was not intuitive enough for people to experience it as something else. To attract more focus to the installation, we set up a plate saying "Sound People – Interactive Installation – Camera, Computer, Stereo". Not everyone saw this plate, but when somebody started to see this as an installation, it often spread as people tried to find the camera, lure its functioning and involve others as mentioned above. The effect of this was that many more people participated and discovered this as a crafted entity: as an installation.

The lesson we learned regarding the notion of installations (or artworks in general) is that if we want our work to function as a *piece*, we must provide the necessary means for the audience to

[4] Orthogonal to the street orientation

see it as such. This counts especially for contexts were other modes of focus or behaviour is present as in the case of a public street.

## 6. FUTURE WORK

As previously mentioned we see the installation as part of an experimental process, rather than a closed, finished artefact. Therefore we also have many ideas for further development and exploration.

One of the future challenges is, of course, making the tracking of people more stable. This could be achieved through experimenting with changes in our tracking algorithms or by setting up the installation at other (perhaps more tracking friendly) locations. In addition to this, we are currently working on a solution involving infrared light and filtering.

Another aspect of future work is the development and testing of different sonic programs. We have already been trying out some prototypes of new sonic programs. One is the *Relation Radio*, which is a program where two participants can 'tune in' on different signals (emerging from the radio noise) by rotating around each other. Another program is the *Role Factory*, in which we have been experimenting with giving participants different roles in an industrial soundscape. One participant is, for instance, given the role of *Geiger Man* – as other participants get closer to *Geiger Man*, *Geiger Man*'s (imaginary) Geiger counter will emit more clicks. Another participant is given the role of forklift, making hydraulic and engine sounds when moving around.

Regarding the recent future, Sound People is to be exhibited on June 20th in Copenhagen at an event for performance art (*Risk Reclaim Entertain* - http://www.myspace.com/1rre).

## 7. ACKNOWLEDGMENTS

We would like to acknowledge the assistance of Helga Rosenfeldt-Olsen in exploring spatial relations and thank Quonga-Fest and theatre Katapult for providing occasion and location for our project.

## 8. REFERENCES

[1] W. Andrew Schloss and Dale Stammen, "Ambient media in public spaces," in *Proceeding of the 1st ACM international workshop on Semantic ambient media experiences* (Vancouver, British Columbia, Canada: ACM, 2008).

[2] Stuart Reeves et al., "Designing the spectator experience," in *Proceedings of the SIGCHI conference on Human factors in computing systems* (Portland, Oregon, USA: ACM, 2005), 741-750.

[3] Zafer Bilda, Chris Bowman, and Ernest Edmonds, "Experience evaluation of interactive art: study of GEO landscapes," in *Proceedings of the 5th Australasian Conference on Interactive Entertainment* (Brisbane, Queensland, Australia: ACM, 2008), 1-10.

[4] Matthias Finke et al., "Lessons learned: game design for large public displays," in *Proceedings of the 3rd international conference on Digital Interactive Media in Entertainment and Arts* (Athens, Greece: ACM, 2008), 26-33.

[5] Dalsgaard, P. and Hansen, L. K. 2008. Performing perception—staging aesthetics of interaction. *ACM Trans. Comput.-Hum. Interact.*, 15, 3, Article 13 (November 2008)

[6] Myron W. Krueger, "Responsive environments," in *Proceedings of the June 13-16, 1977, national computer conference* (Dallas, Texas: ACM, 1977), 423-433.

[7] Bourriaud, Nicolas, "Den relationelle æstetik", in *Exogen*, udstillingskatalog, Copenhagen 1997.

[8] *Access*: http://www.accessproject.net/archive.html

[9] *Very Nervous System*: http://homepage.mac.com/davidrokeby/vns.html

[10] Quonga-Fest: http://www.quonga.dk/Oversigt.html