DJ SPAT: SPATIALIZED INTERACTIONS FOR DJS

Georgios Marentakis, Nils Peters and Stephen McAdams
CIRMMT, Schulich School of Music, McGill University
555 Sherbrooke St. West, Montreal, Québec, Canada H3A 1E3
{georgios.marentakis, nils.peters, stephen.mcadams}@mcgill.ca

ABSTRACT

A novel interactive design for spatialization in halls and its real-time control is presented. The DJ interaction metaphor is augmented to achieve control of spatialization as a bi-product of musical performance using motion-tracking technology. A system is specified and realized using commonly available hardware technology. The integration of the system in performance spaces is discussed. Based on the design we achieve seamless integration of spatialization and its real-time control in musical venues.

1. INTRODUCTION

In 20th century music, there has been considerable attention in the use of the realistic space of performance and perception. A considerable number of composers have entertained this idea. To name a few, Brant and Ives used simultaneous spatial layers, Varese used the concepts of sound mass and geometry, the Darmstadt school of Stockhausen and Boulez employed geometrical manipulations and simultaneity of sounds in space, as did Xenakis and, more recently, Roger Reynolds. In the context of this paper, we are interested in experimenting with the use of physical space as the space of performance and perception from the point of view of perception and action.

Physical space in music can be interpreted both as the space music is performed in and in connection to the spatio-temporal parameters of sounds themselves. Both factors influence the percept. The space the music is performed in interacts with acoustic, as well as electro-acoustic, material and shapes the perceived result. Variations in reverberation time, clarity, the ratio of direct to reverberant energy and other measures of the acoustical quality of space, substantially influence the perception of the acoustic material, directly affecting the immersion of the listener in the acoustic reality and the readability of the music. In addition, the performance space can be used in a theatrical way, by providing a frame of reference for the projection of acoustic as well as electro-acoustic geometries.

Such geometries are defined by the spatiotemporal parameters of the sounds and their evolution. As in numerous compositions, acoustic or electro-acoustic sounds are positioned and moved physically through the placement and movement of the musicians or by programming in auditory virtual environments. The audience can also be thought of as static vs. mobile and in this way interesting combinations can be imagined (Harley [1]). The design space for contemporary audio technology and art can thus associate static and moving sounds interacting with a dynamic audience reacting in static or mobile ways.

In any case, the successful realization of such attempts is strongly dependent on two additional aspects, other than the conception and performance of the musical material itself. The first is the communication of auditory cues to the audience and the second the control of space during the performance of music. In this paper, we make an initial attempt to investigate these aspects from the point of view of designing interaction and entertainment experience.

To answer the first question, we look into psychoacoustics and auditory scene analysis and synthesis and their interaction with the space within which music is realized. For the second, we investigate motor control and human-computer interaction as bi-products of musical performance.

2. PERCEPTUAL CONSIDERATIONS

As is mentioned in [2], auditory perception of space is not as accurate as visual perception of space. Two major metrics of localization ability are found in the literature: localization error and localization blur or minimum audible angle. The first corresponds to the mean error in localization judgments. The second corresponds to the minimum audible displacement of an auditory source. They both depend primarily on the direction, the spectral content, and the duration of the auditory source. The literature on spatial hearing is vast. Due to lack of space we point the reader to [2-4] for estimates of the aforementioned quantities. Roughly speaking, auditory spatial acuity is by far less accurate than visual acuity, and sometimes displacements of more than 10˚ are necessary to perceive that a sound, e.g. at the side, has moved [4]. Nevertheless, research indicates that the perception of sound direction and distance in the real world is consistent across people and can therefore be expected to be uniform within the audience as long as real sounds or sounds from single loudspeakers are concerned.

Contemporary performances, however, will inevitably use virtual spatialization systems. Spatialization systems work primarily in two ways: they either take advantage of psychoacoustic phenomena, such as summing localization to present a sound in a phantom position between two loudspeakers, or they try to recreate a sound field as it would be if the sounds were physically present in the space. The most well-known example of the first case is stereo reproduction, which has been extended by Pulkki [5] to apply to arbitrary horizontal or elevated geometrical arrangements of loudspeakers. Methods of sound field reproduction are the Ambisonics technique [6] and the WaveField synthesis technique [7].

Without going into detail, it is worth mentioning that there are considerable limitations associated with each method, as well as a lack of psychoacoustic evaluation. Firstly, it is commonly accepted that all virtual auditory environments are limited with respect to the area within which they can faithfully reproduce the spatialization effect. The limitation is either frequency dependent as in wavefield synthesis or space dependent as in pan-pot or Ambisonics systems where the illusion is only valid for people in the so-called sweet spot. The area of the sweet spot, however, is usually relatively small. Secondly, it has been found that sound localization in virtual environments is not as accurate as in the real world [8]. It improves however, as the number of loudspeakers increases.

It is therefore expected that in a given space, although physical sounds will be perceived consistently, sounds spatialized by a virtual audio system will not. This poses serious limitations for both the audience and the performers. Therefore special consideration has to be made and perhaps certain tradeoffs will have to be accepted. We come back to this point when we consider the performance space where our system can be used.

3. CONTROL OF SPATIALIZED SOUND

Control of spatialized sound is a major issue in the development of the use of space in a musical context. Here, we focus on the spatialization action, that is the way in which a sound will be assigned a location in space, moved, steered or displaced away from the audience.

Spatialization actions can be preprogrammed and triggered automatically based on elapsed time. Such an option is particularly appealing and has been used extensively in the context of electro-acoustic music, given that a lot of the material is preprogrammed. There are some major disadvantages, however.

Using this method, space is not integrated into the music performance process. Performers are therefore alienated from controlling space. In addition, the audience itself looses contact with the source of the spatialization action and has no access to multimodal information to connect it to. More than that, the preprogramming of the spatialization action (and its performance by a computer) has the disadvantage that it is not suitable for improvisation or real-time performance. In this sense, its application is limited in certain domains. We therefore seek ways to control spatialization that are visible and meaningful for the performers and audience.

We consider the integration of spatialization control in the music creation process a key but understudied aspect. This paper presents a design example that achieves seamless integration of spatialization in the music creation process and considers performers and the problems they might face when dealing with virtual spatial sound.

The design intends to exemplify an approach for spatializing and controlling space in performance spaces. It integrates physical and virtual sounds and their seamless control by augmenting an existing interface. By using existing functionality, it does not require any additional learning on the performer's side.

4. SPATIALIZATION AND LIVE REAL-TIME CONTROL

This study investigates and augments DJ performance in relation to spatialization. Although DJ performance has been studied [9, 10], this has mostly been done in the context of DJ augmentation and gesture analysis. Software and hardware such as Ms Pinky (www.cycling74.com), Final Scratch (stanton.dj.com) or Mixx (mixxx.sourceforge.net) track DJ gestures for controlling digital samples. They do not consider the augmentation of DJ performance with analogue records. In addition, they do not consider spatialization. Gestural control of spatialization is explored in [11], however it is not linked with DJ performance or performance in general in a direct way. Our paper is novel because it considers real-time control of spatialization using direct mappings from performers that offer multimodal information. In addition, it considers the placement of the performance participants so that they can perform their role without perceptual problems. Our point of departure is to use the angular displacement of the DJ's hand to control spatialization of the sound played by the turntable.

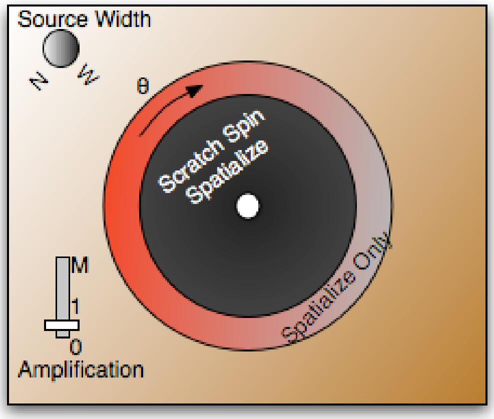

Figure 1. Turntable Design. The source spread knob controls source width (wide to narrow). User movement amplification is controlled by the slider. Two tracked areas are used: one for spatialization only and the other for scratch, spin and spatialize.

4.1. Specification

Here, we define the extra functionality we want to achieve. We then seek to augment the DJ's capabilities and compensate for localization problems in a non-intrusive way using the affordances and constraints of the designed system. The system should adhere to the following specifications. In this sense it should be possible to:

1 Perform (scratch/spin) without moving the sound, perform and move the sound, and move the sound without performing and perform and spatialized independently;

2 Adjust (amplify/attenuate) the effect of human movement on sound movement;

3 Adjust the distance/ perceived width of the sound;

4 Provide seamless user engagement/disengagement;

5 Provide scalability of the system (more turntables);

6 Achieve perceptual fidelity.

The above requirements are integrated into the system as depicted in **Figure 1**.

4.2. Design and Mappings

The design uses the angular displacement of the DJ's hand, two areas of distinct functionality, labeled 'Scratch Spin Spatialize' and 'Spatialize Only' in **Figure 1**, and linear and rotary potentiometers. In addition, we assume in this section that the angle θ of the user's hand angular displacement is known. At any moment sound position is increased/decreased by an amount $\Delta\Theta = \alpha\ \Delta\theta^{\circ}$ (1), where $\Delta\theta^{\circ}$ is the angular hand displacement.

The 'spatialize only' area is outside the record area. Therefore, movement in this area has an effect only on spatialization. On the other hand, movement in the 'scratch, spin and spatialize' area, will result in simultaneous scratch, spin and spatialization. Bi-manual interaction can be used to scratch and spatialize independently. In both cases, the magnitude of the effect is determined by parameter α. α is controlled by the linear potentiometer. Setting α to 0 results in interaction with the record (spinning & scratching) to be disengaged from spatialization. By setting α to 1 user displacement maps directly to sound movement. By setting the parameter in the intervals (1,M] or (0,1) the effect of user movement is amplified/attenuated. In this way, even small hand movements can result in perceptible sound displacements and the effect of spinning actions can be attenuated.

Intensity, a major cue for distance, is controlled by the fader; a DJ would use anyway for mixing the sound. No further controls need to be introduced. Optionally, if supported by the mixing equipment, the fader data could be forwarded to the spatialization system and mapped to the distance parameter of the spatialization algorithm. Alternatively, echoes or reverberation can be added using mixer resources to further improve the effect. Source width [narrow/wide], is controlled by the rotary potentiometer. When set to W the sound is played back at the same level by all loudspeakers.

User engagement/disengagement is achieved by placing the hand on (or removing it from) one of the active areas. In our current implementation, the user can also call the sound to a desired position by tapping their hand on it. The amplification control is used to determine the speed of movement in this case.

Perceptual fidelity (in accordance with Section 2) is achieved by placing the performer in the middle of the venue exactly on the sweet spot as in **Figure 2**, with the decks surrounding him/her. We entertain the possibility of extra musicians on the border of the deck area that perform live. They are placed there because they also need to have the best possible fidelity of the spatial experience. The audience, however, is left to freely move or sit inside the loudspeaker array that is used for sound spatialization. Finally, the system is scalable, more than one turntable can be used seamlessly as is shown in **Figure 2**.

4.3. Implementation Discussion

The overall implementation of the system is shown in **Figure 2** and can be subdivided into audio capture and reproduction, motion tracking and the software implementation of the spatialization and mapping.

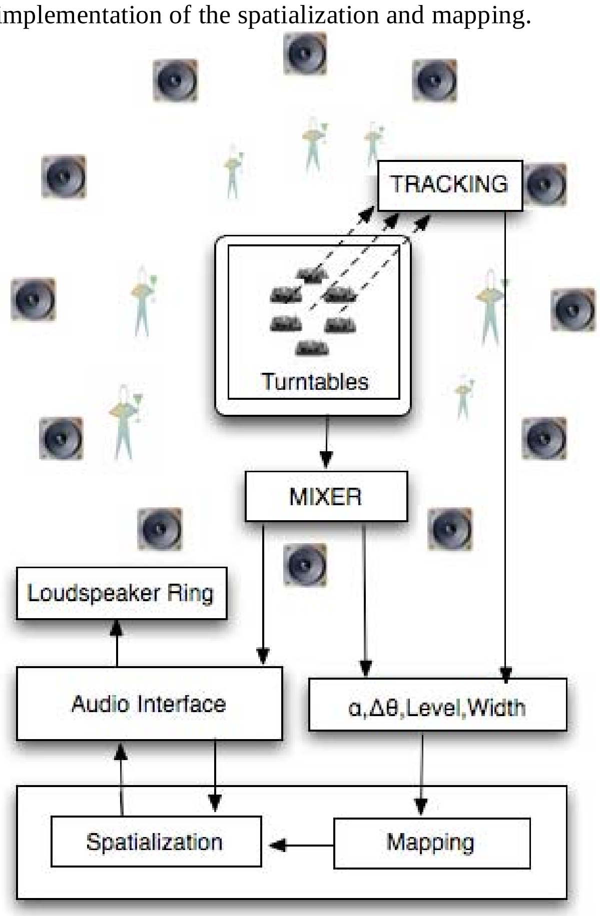

Figure 2. Performance Space and System Implementation. DJ in the sweet spot possibly surrounded by musicians, while the audience explores within the loudspeaker array.

Turntable signals are pre-amplified and passed through the mixing console and through the A/D interface into the computer running the software. Then, they are fed into the spatialization algorithm and sent back to the speaker array using the D/A interface.

Motion tracking can be performed in a number of ways. Camera, inertial, pressure or electromagnetic tracking can be considered depending on the desired application. The data from the tracking system are preprocessed to obtain angular hand displacement. Together with source level and width, they are mapped to the control parameters of the spatialization system. The output is then converted to analogue and sent back to the loudspeaker array.

An example implementation was realized using two turntables and a mixer, an RME Fireface audio interface, the VBAP [5] system implemented for Max/MSP and an 8-speaker array. The system worked smoothly with no major problems due to latency. A Powermate (www.griffintechnology.com) was used to adjust source width and the amplification parameter. The source width parameter was directly fed into source width employed by the VBAP system.

Simple camera tracking was used for gesture recognition [iSight / TapTools, (www.electrotap.com) for Max/MSP], tracking a colored marker (ring) worn on the hand by the user. A prior calibration of the system is performed so that we can recognize when the marker is inside each of the active areas based on distance to the centre. The update rate of the camera system was at 30 f/s. We could track medium speed movements at two decks reliably using a single camera and doing gesture recognition in the appropriate parts of the captured picture. The tracker size was used to determine when the marker was near to the record area and enable/disable the system. The same method was used to recognize the tap gesture.

A major advantage of such an implementation is that the DJ is completely free from any sort of cabling or tracking equipment. The disadvantage is that s/he has to keep her/his body and head away from the tracked surface of the vinyl, and that recognition performance will deteriorate in dim light conditions. This could be alleviated by using an illuminant marker.

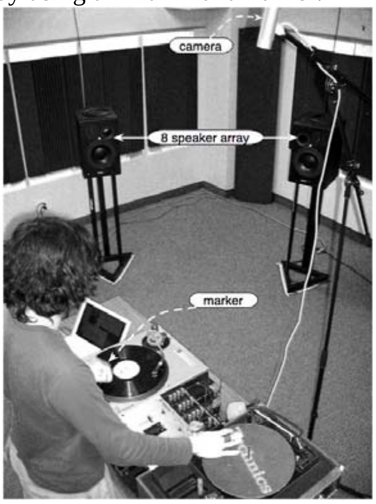

Figure 3. A DJ using the developed system.

The system was tested by a DJ and found to be working according to specifications and to be quite enjoyable to use (see **Figure 3**).

5. FUTURE WORK AND CONCLUSION

As future work, we are investigating possibilities for a big user study in a large venue with a larger version of our system. We will also be looking into ways the sound of performing musicians can be spatialized by the DJ in real time, complementary to his/her own performance and in accordance with formal composer directions. Simultaneous DJ performance is also planned.

In conclusion, we presented a novel approach for including spatialization and its real-time control in performance. We showed how DJ interaction can be augmented to enable control of spatialization using motion tracking technology. We specified a system and realized it using commonly available hardware technology and placed it in a perceptually relevant way in the venue space. Spatialization and its control were integrated seamlessly with spinning, scratching and performance.

6. REFERENCES

1. Harley, M.A., *Space and Spatialization in Contemporary Music: History and Analysis, Ideas and Implementation*, PhD thesis, 1994, McGill University: Montreal.
2. Blauert, J., *Spatial Hearing: The psychophysics of human sound localization*. 1999: The MIT Press.
3. Middlebrooks, J. and D. Green, *Sound Localization by Human Listeners*. Annual Psychology Review, 1991. **42**: p. 135-159.
4. Mills, A., *On the Minimum Audible Angle*. The Journal of the Acoustical Society of America, 1958. **30**(4): p. 237-246.
5. Pulkki, V., *Spatial Sound Generation and Perception By Amplitude Panning Techniques*, in *Helsinki University of Technology*, 2001.
6. Gerzon, M. *Panpot Laws for Multispeaker Stereo*. in *92nd Convention of the Audio Engineering Society*. 1992. Vienna, Preprint No. 3309.
7. Berkhout, A.J., D. de Vries, and P. Vogel, *Acoustic Control by Wavefield Synthesis*. Journal of the Acoustical Society of America, 1993. **93**(5): p. 2754-2778.
8. Pulkki, V. and T. Hirvonen, *Localization of Virtual Sources in Multichannel Audio Reproduction*. IEEE Transactions on Speech and Audio Processing, 2005. **13**(1): p. 105-119.
9. Beamish, T., K. MacLean, and S. Fels. *Manipulating Music: Multimodal Interaction for DJs*. in *ACM CHI*. 2004. Vienna.
10. Hansen, K. and R. Bresin. *Mapping Strategies in DJ Scratching*. in *NIME*. 2006. Paris.
11. Marshall, M., et al. *On the Development of a System for the Gesture Control of Spatialization*. in *ICMC*. 2006. New Orleans, U.S.A.