4DSOUND: A New Approach to Spatial Sound Reproduction and Synthesis

PAUL OOMEN, POUL HOLLEMAN & LEO DE KLERK
Spatial Sound Institute

We argue that experiencing sound is inherently spatial and that the spatial properties of sound form a realm of musical expression that is as yet only partly discovered. To further the exploration of spatial sound as a medium, we argue that the generally applied methodologies to produce spatial sound are inherently limited by the sound source, i.e. loudspeaker boxes or headphones, and that reproducing or synthesising sound spatially asks for a new approach with regards to the medium. It is possible to create an innovative and non-conventional sound system that removes the localisability of the sound source from the equation, regardless of the listener's individual hearing properties. The system provides for a social listening area and improved loudness at equal acoustic power. The system is backwards compatible with existing audio reproduction formats, allows integration with a wide variety of control interfaces and encourages new approaches in design.

Introduction

Early in 2007, Paul Oomen conceptualised the design of a sound system that would be able to produce dimensional sound sources in an unlimited spatial continuum. Since this time Oomen has worked with Poul Holleman, Luc van Weelden and Salvador Breed on the development of object-based processing software and control interfaces for the system.

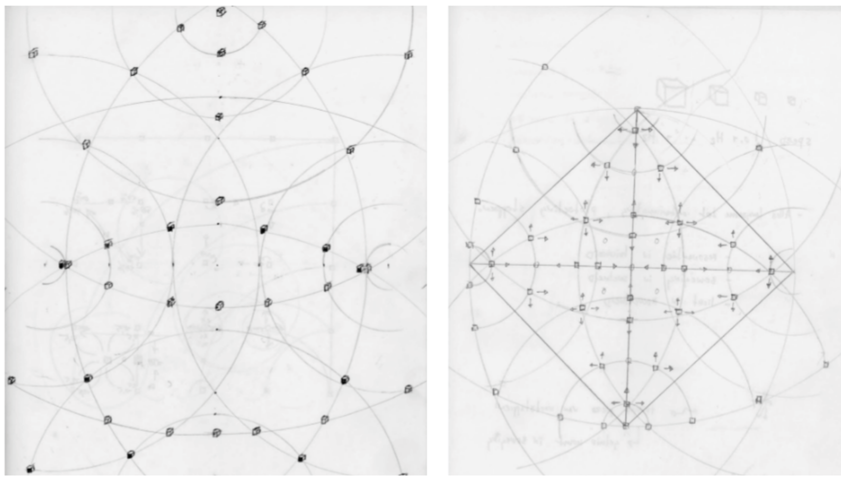

Figure 1 Dimensional sound source, sketches, P. Oomen, 2007.

Independent from Oomen, but in the same year, Leo de Klerk made public his patent application of an omnidirectional loudspeaker that in configuration produces coherent phantom images in both the vertical and horizontal plane (int. Patent no WO2006100250 A transducer arrangement improving naturalness of sound, 2007). Since 2010, Oomen, Holleman and de Klerk work together on the development of the system.

The system has been firstly presented to the public and in the presence of the authors et al. during Amsterdam Dance Event at the Muziekgebouw aan het IJ, October 2012, and has since been showcased in and outside Europe in association with a diverse range of artists from the field of electronic music, sound art and immersive technologies. The system has been subject of multiple lecture presentations, written publications and film documentaries (www.4dsound.net).

The article will at first outline the fundamental principles of the system by looking at a fallacy of spatial sound reproduction as can be derived from the theories by scientists Von Bekésy and Moore, and present a new approach to this problem by the authors. We will then continue with an overview of the technical implementation of the system, which incorporates the design of new hardware, software and interactive interfaces. Finally, we will present our conclusions with regards to the current status of artistic and technical achievements of the system.

Fundamental Principles

A. Fallacy of spatial sound reproduction

We are familiar with Von Bekésy's problem (Experiments in Hearing, 1960): the 'in the box' sound effect seems to increase with the decrease of the loudspeaker's dimensions. In an experimental research on the relation between acoustic power, spectral balance and perceived spatial dimensions and loudness, Von Bekésy's test subjects were unable to correctly indicate the relative dimensional shape of a reproduced sound source as soon as the source's dimensions exceeded the actual shape of the reproducing loudspeaker box. One may conclude that the loudspeaker's spatio-spectral properties introduce a message-media conflict when transmitting sound information. We cannot recognize the size of the sound source in the reproduced sound. Instead, we listen to the properties of the loudspeaker.

Why does the ear lock so easily to the loudspeaker characteristics? Based on the hypotheses of Brian Moore (Interference effects and phase sensitivity in hearing, 2002) et al., we may conclude that this is because the ear, by nature, produces two dimensional nerve signals to the brain that reflect the three-dimensional wave interference due to direct interaction of both the ear's and the physical sound source's spatio-spectral properties. Therefore any three-dimensional spatio-spectral message information embedded in the signal to be transmitted is masked by the physically present media information related to the loudspeaker's properties. One could say that in prior art solutions the input signal of the loudspeaker system merely functions as a carrier signal that modulates the loudspeaker's characteristics.

We experience a phantom sound source reproduced by a stereo-system to appear more realistic than the same sound reproduced from one box.

However, spatial sound reproduction by means of stereophonic virtual sources is only a partial solution because it fails for any stereophonic information that does not meet the L=R requirement for perfect phantom imaging. In fact stereophonics is no more than improved mono with the listening area restricted to a sweet spot.

B. A phantom loudspeaker

For a new approach of this problem we considered that one better could overrule the monaural spatio-spectral coding than trying to manipulate it (as f.i. in crosstalk suppression), thus preventing inter-aural cross correlation which is the root mechanism for the listener's ability to localize any non-virtual sound sources i.e. loudspeakers.

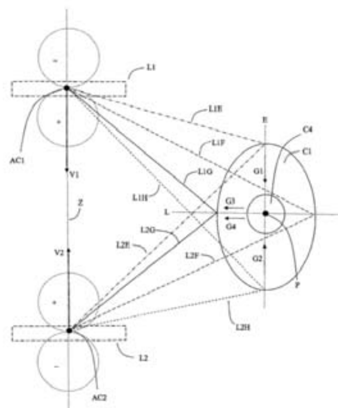

Figure 2 A phantom loudspeaker: two loudspeakers that are present at the same side of an ear are directed towards each other's acoustic centre make a non-localizable, virtual loudspeaker

The resulting application describes a sound system that produces coherent vertical phantom images that cannot monaurally nor binaurally be decomposed to their root sources i.e. the actual loudspeaker drivers. The loudspeaker becomes audibly non-localizable and any multi-channel horizontal configuration of these virtual speakers is possible without adding masking media properties. As a consequence the listening area is not anymore restricted to a sweetspot. Loudness perception is dramatically improved because the 'out of the box' sound screen will now fit to even the largest sound source shapes.

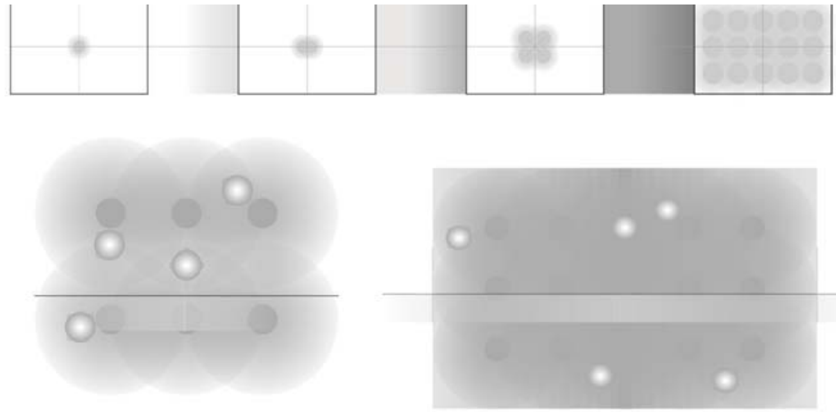

Figure 3 Multi-channel configurations of virtual loudspeakers

Technical Implementation

A. A rotation symmetrical response

For the production of a coherent vertical phantom image that also must be able to interact coherently in the horizontal plane in order to allow multi-channel configuration, a driver structure with a rotation symmetrical off-axis response is a main requirement. The loudspeakers are constructed from modified co-axial drivers, where the modification consists of an altered implementation of existing driver parts and does basically not need re-engineering of the chassis. However, the modification offers the advantage to implement further refining features, f.i. mass-less motional feed-back sensing, that were difficult to apply in the original design.

B. An unlimited spatial continuum

From this application derives the choice for a right-angled and equal spaced speaker configuration expanding in the vertical and horizontal plane to enable balanced sound pressure throughout the entire listening area. In this configuration we define the 'inside speaker field' and 'outside speaker field' marked by the physical borders of the speaker configuration, that in conjunction provide an unlimited spatial continuum, both incorporating the active listening area and expanding beyond it.

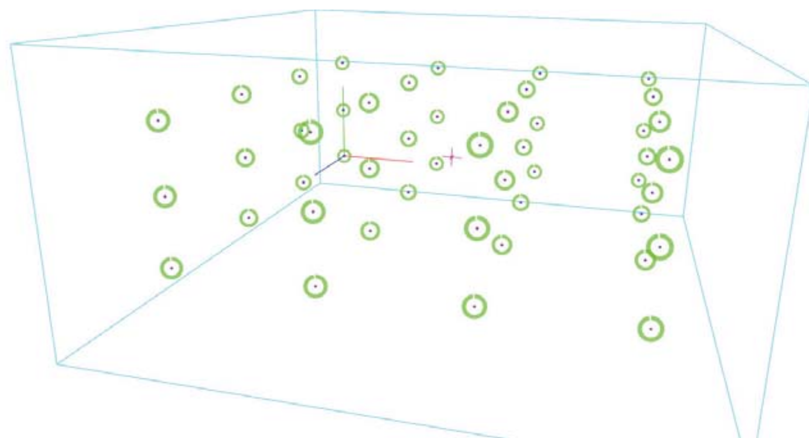

Figure 4 A right-angled and equally spaced speaker configuration to form an omnidirectional listening area.

The spatial continuum is defined in Cartesian coordinates X) left/right Y) above/below, and Z) front/back in which coordinate [0,0,0] becomes automatically in the center of the listening area on floor level. This infinite and omnidirectional sound screen can now be used to position dimensional sound sources of various geometrical shapes such as points, lines, planes and blocks, both inside the speaker field, in between and next to to the listeners, and outside the speaker field, above, beneath and around the listeners at unlimited distances.

C. Object-based processing software

The object-based processing software is written in C++ and accepts a virtually unlimited amount of discrete sound inputs, which are processed according to spatial data. More than 200 of its data parameters are dynamic and can be addressed via Open Sound Control.

The software defines sound sources, reflecting walls and global transformations of the spatial field such as translation, rotation and plodding. The spatial synthesis can be divided in two parts: i) mono processing on the input signal according to virtual spatial properties, that is consecutively converted to ii) multi-channel processing which treats each speaker output independently.

For initialization the software needs the speaker configuration properties, such as dimensions and spacing of the speakers, and the amount of active sources, whereas a source is a serial circuit of mono modules that synthesize spatial phenomena like distance, angle, and doppler. The multi-channel conversion happens at the end of this chain and translates the sound source's spatial properties like position and dimensions to speaker amplitudes and delays. Its algorithms are based on matrix transformations and vector math. Next to amplitudes and phase delay times is also processes interactive data, such as whether a source is located behind or in front of a virtual reflecting wall or not.

In the multi-channel processing, two panning algorithms are active, one for sources inside the speaker field and the other for outside the field. The inside speaker field panning is based on the shortest distance between the dimensional sound source and the speakers. The outside speaker field panning is based on the shortest distance between the projection of the sound source on the speakers defining the edge of the speaker field. The outside speaker field projection knows two types: 'perspective-panning', which takes a vantage point, typically the centre but can be any other co-ordinate inside speaker field and 'right-angled panning', which takes a vantage 'line' through any point, typically the centre, resulting in wider projections as it not diminishes with a longer distance.

Conclusion

In everyday life people process complex spatial sound information to interact with their surrounding, mostly involuntarily. We are exposed to continuous movement in the environment around, above and beneath us. And in this environment we move ourselves in complex patterns, sometimes fast and straight, then slow and hesitant, turning back and forth, possibly laying, standing, bending or sitting. We are continuously affected by changes of movement surrounding us, and by our patterns of behaviour we influence our perception of this environment and stimulate movement in the environment itself. To be able to further our explorations with spatial sound as a medium, it is essential that the reciprocality between the listener the sounding space is restored to achieve lifelike and meaningful experiences of spatiality.

Conclusively we may state that it is possible to create an innovative and non-conventional sound system that removes the localizability of the sound source from the equation, regardless of the listener's individual hearing properties, and improves loudness perception at equal acoustic power. The system provides for a social listening area, is backwards compatible with existing audio reproduction formats, allows integration with a wide variety of control interfaces and encourages new approaches in design.

Figure 5 4DSOUND System at the Spatial Sound Institute

Leo de Klerk, 1958 (NL) Tonmeister, composer, founder of Bloomline Acoustics and inventor of Omniwave, the inaudible loudspeaker.

Poul Holleman, 1984 (NL) Sound technologist, software developer, co-founder of 4DSOUND

Paul Oomen, 1983 (NL) Composer, founder of 4DSOUND, head of development at the Spatial Sound Institute, Budapest