Speaker-Herd: A Multi-channel Loudspeaker Project for Miscellaneous Spaces, Loudspeaker Architectures and Compositional Approaches

F. Bierlein*, E. Farchmin*, L. Fütterer*, A. Kerschkewitz*, D. Loscher*, P. Modler*, T. Möhrmann*, T. Myatt†, A. Rafinski*, R. Räpple*, S. Rosinski*, L. Schwarz*, A. Unger*, M. Zielke*,

* University of Media, Arts and Design, Department of Media Art - Sound, Karlsruhe, Germany

† University of York, Department of Music, York, UK

Abstract — **Strong interest in spatial sound has existed at all times and on various levels of aesthetic music and sound production and perception. In recent years the availability of high-quality loudspeakers and digital multi-channel audio systems has paved the way to incorporate spatial acoustics into musical composition. In this paper we describe a project which is aimed at providing flexible possibilities to experiment with miscellaneous loudspeaker architectures and multi-channel distribution systems. The system allows the use of up to 96 audio channels in real time which can be fed to loudspeakers setup according to varying spatial designs. As examples a number of realized architectures and compositions will be described.**

I. Introduction

Although historically music rarely used spatial parameters as part of their composition it was often composed to be played and perceived in spaces with tangible acoustic features, e.g. medieval cathedrals support chants by adding sophisticated reverberation to provide extension and prolongation of singing voices [1]. In later periods compositions considered acoustic aspects more directly e.g. by formulating needs for acoustic reproduction in spaces as well as through compositional instructions considering acoustic and spatial parameters such as position of the orchestra, type of concert hall, movements of players.

The availability of loudspeakers and multi-channel techniques boosted the interest to incorporate aspects of spatial listening into compositions.

In *Poem Electronique* precise spatial movements of acoustic events are part of the composition and were realized through loudspeaker arrays [2].

Reproduction of spatial acoustics evolved from stereo systems to surround sound systems such as *Dolby DTS*, *Ambisonics* or *Wavefront-synthesis*. These approaches aim to reproduce or create certain acoustic spaces, such as recreating the sound image of a classical symphony recorded in a concert hall through a living-room stereo speaker setup, or more sophisticatedly to recreate entire acoustic wave fields through large speaker arrays [3].

A different approach follows the concept of the loudspeaker orchestra (Diffusion, Acousmonium) . The aim to recreate spatial acoustic images from recorded material is waived in favour of exploring the characteristics of miscellaneous loudspeakers, different speaker and audience positions, different spaces and, especially for the diffusion approach, the interpretation by the performer who runs the mixing desk as integral part of the musical concert [4], [5].

In our approach we aim to provide a flexible environment which enables creative prototyping of experimental loudspeaker setups which are not bound to certain rooms or spaces and are open to applications of all kinds of multi-channel sound distribution strategies and software. Instead of relying on large scale, and thus fixed and inflexible, setups or built-in speaker systems in concert halls we aim to provide tools to set up speaker and multi-channel software for miscellaneous architectures and all kinds of buildings or spaces.

II. Hardware Components

Based on the aim of realising a flexible low-cost system but at the same time to provide as far as possible high-quality multi-channel audio we decided on five basic components:

- A large but still extendable number of high-quality loudspeakers
- A multi-channel digital audio interface with at least 64 audio channels
- A powerful computer for multi-channel mixing and spatialisation algorithms
- Necessary cabling, e.g. from the audio-interface to the speakers
- Rigging or stands for the speakers or additional gear for certain speaker setups

A. Speakers

For the loudspeakers we choose Genelec 8040 and 8240for the following considerations:

- Flat free-field frequency response of 48Hz-20kHz (± 2 dB) at 105 dB SPL@1m
- Compact physical dimensions of 350 x 237 x 223 mm and 8.6 kg weight
- Very low channel crosstalk
- High audio quality of amplifier and speaker

We accepted the drawback of additional cabling for the power supply of active speakers in favour of the high audio-reproduction features and the compact dimensions [6].

B. Processing Computer

Based on the aim of providing the main operating system platforms used for audio production, an Apple Intel machine (MacPro) with a quad-core kernel was chosen. The machine can be booted into OS-X, Windows and Linux, and is equipped with a second graphics card enabling the running of up to 4 separate screens.

C. Digital Multichannel Audio Interface

To keep the costs of the project in a certain range we looked for a solution offering a large number of high-quality audio channels at a low cost per channel rate. Due to the ePCI bus of the MacPro a range of audio interfaces were possible.

A Digidesign/Protools solution would have provided the highest audio quality but at a high price per audio channel.

The RME's MADI solution gives 64 audio channels with one ePCI card and the possibility of synchronising a second card with an additional 64 channels as a tandem solution. The RME-MADI system provides easy distribution of audio through digital MADI connections (BNC) or after a conversion through ADAT or AES/EBU connections. In combination with digital speakers this enables flexible and comfortable cabling between the audio interface and the speakers even in large and complex setups.

A third alternative is MOTU's 444 ePCI card combined with up to four 24IO converter units giving a total of 96 input and 96 output channels. In contrast to RME's MADI system the 24IO box is a straight digital-to-analogue and analogue-to-digital converter box with 24 balanced analogue inputs and outputs. An interesting feature of the 444 card is the possibility of connecting most of MOTU's audio interfaces, including older 2408 interfaces, which provide up to 24 ADAT I/Os. With this option it is possible to connect digital drains or sources to the 444 in combination with analogue 24IO units.

In short the RME-MADI is a fully digital system, providing a large total number of audio channels (128 in tandem mode) at a reasonable low cost per channel rate using digital speakers. The MOTU solution has a lower cost per channel rate with the trade-off of a maximum number of 96 channels. Due to the analogue I/Os of the 24IO more effort has to be put into cabling, e.g. multi-cores for larger distances.

Based on our aim of using only a certain budget and existing material, such as multi-cores, balanced leads or additional passive speaker units, we choose the MOTU 444 ePCI card with 24IO analogue interfaces and a 2408 digital interface which was already at hand.

III. SOFTWARE

A. Operating Systems

The versatile multi-boot option of the Intel MacPro machine allows us to work with most commonly used operating systems such as OSX, Windows and Linux.

B. Software for Audio, Spatialisation and Distribution

Most of the compositions realised on our system were based on Max/Msp/Jitter running on OSX. Pieces which had been projected in Max/Msp on Windows were converted to OSX, which enabled a faster loading of different composition projects during a performance. A smaller number of pieces were realised in PD running on Windows.

For the spatialisation and sound distribution the following software packages or tools are accessible:

- ICST Ambisonics Spatialisation Tools [7]
- SPAT IRCAM [8]
- vbab~ [9]
- Nuendo/Cubase [10]
- Zirkonium [11]

Besides these spatialisation tools a number of pieces were realised with non-standard audio distribution approaches which had been programmed and realised by the composers.

IV. SPEAKER ARCHITECTURES / SPEAKER SCULPTURES

All setups and compositions on our system were realised in multi-channel format, with a current number of channels of about 40-50.

A. Dome

As a first test configuration we set the speakers up in three circular rings forming a dome: a base ring with 12 speakers on the floor, a middle ring with 8 speakers at a height of about 6m, a top ring with 8 units at a height of about 12m.

B. Stereo Corridor or Pathway:

Fig. 1 and Fig. 2 show a corridor configuration where speakers are set up in two parallel lines at ear level.

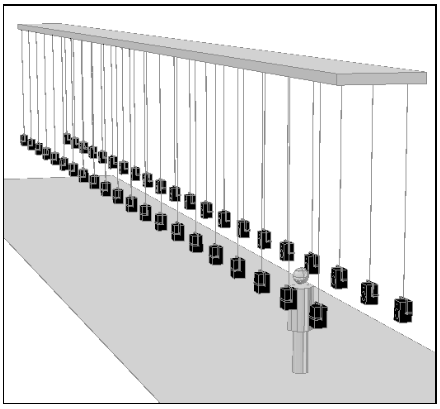

Fig. 1: Stereo corridor with 42 speakers (CAD)

Fig. 2: *Auditory-Channel*: stereo corridor setup

The distance between the speakers is about 1.5m. The audience is tracked through a camera/computer system which may be used to control the sound compositions. In the current setup the pathway has a length of about 30m.

C. Cone or Tree

42 Loudspeakers suspended from a rigging are arranged in the form of a cone with a height of about 6m and a base diameter of about 5m. The cabling of the speakers gives the impression of the root, trunk and branches of a tree (Fig. 3, Fig. 4).

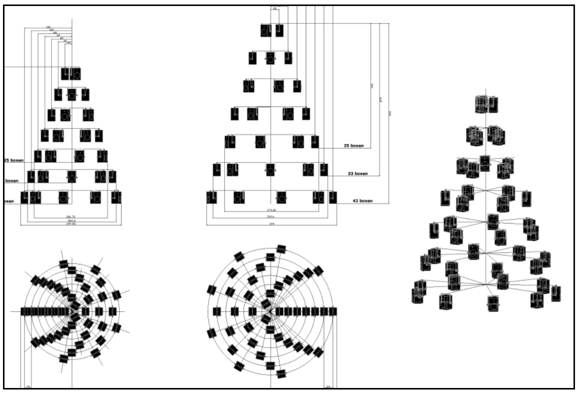

Fig. 3: Two cone/tree setups (CAD)

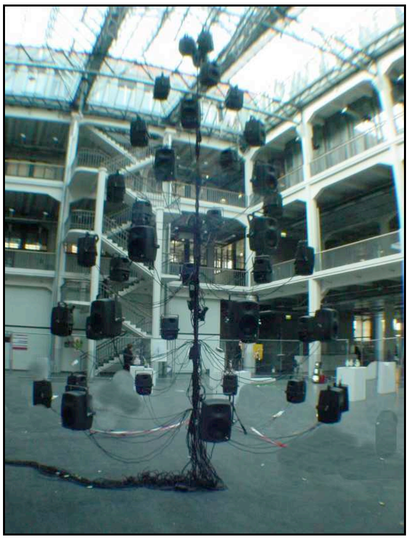

Fig. 4: *Sound-Tree*: Cone/Tree setup

D. Complete Sphere

In Fig. 5 to Fig. 7 the speakers are configured as a complete sphere. The level of the listeners is about the centre of the sphere.

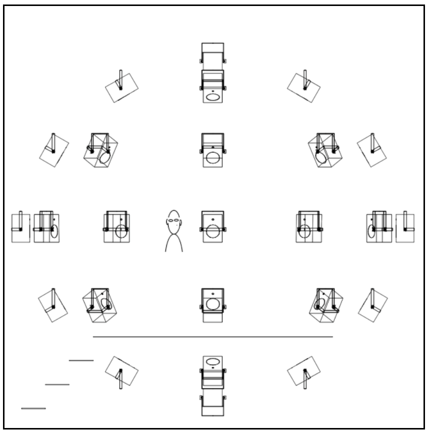

Fig. 5: Complete Sphere: sideview (CAD)

Listeners access the sphere through a stairway onto a platform of about 3x3 sqm size and 1m height. Various sound-spatialisation approaches are explored through different compositions (Fig. 5 to Fig. 7).

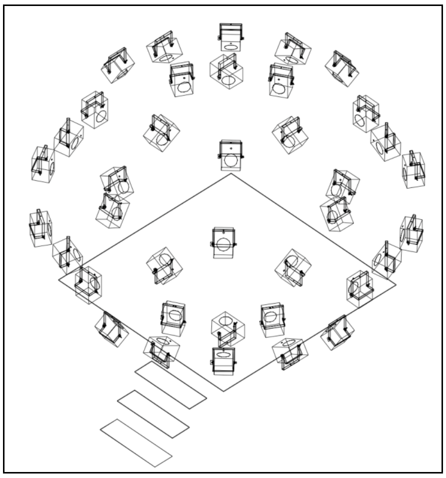

Fig. 6: Complete Sphere: 3D-view (CAD)

Fig. 7: *Sound-Sphere*: Full Sphere setup

E. Mixed Setups

Mixed speaker setups combine architectures, e.g. a dome and a circle, to achieve a combination of spatialisation results. Tony Myatt designed a special mixed speaker setup for his piece *untitled-3*, which was premiered on our system. The architecture is based on a geodesic dome with 13 speakers, a supporting inner circle close to the audience and one large rectangular cube configuration of 18x24x10m of 16 speakers. (Fig. 8. to Fig. 10).

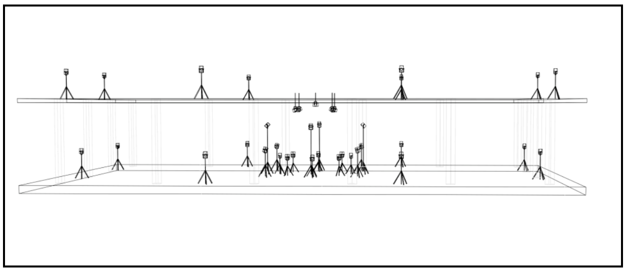

Fig. 8: Mixed setup side view (CAD)

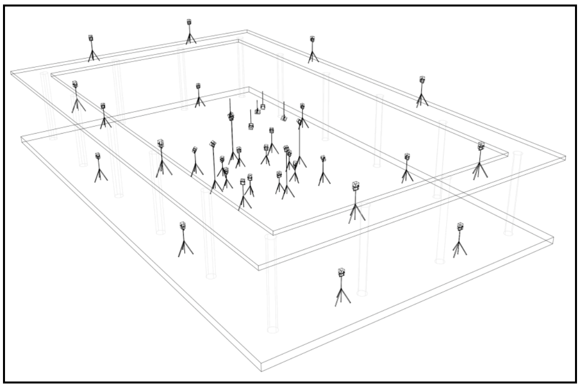

Fig. 9: Mixed setup 3d view (CAD)

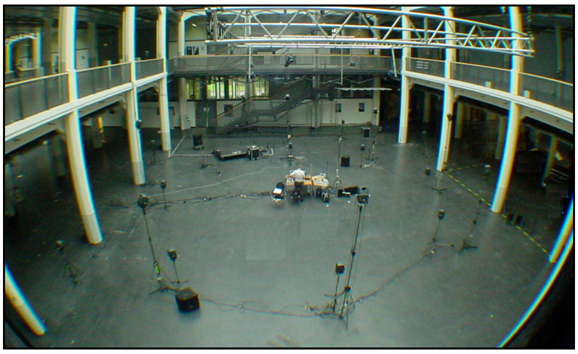

Fig. 10: *Untitled 3 by T. Myatt,* Mixed setup

F. Studio

In Fig. 11 and Fig. 12 a rectangular setup is shown which was realised in a studio space. The speakers are suspended with standard theatre hooks and can be arranged freely. The setup allows fast rearrangement of the speakers, as well as extensions to different geometries, which is necessary for a rehearsal and test situation in a studio space.

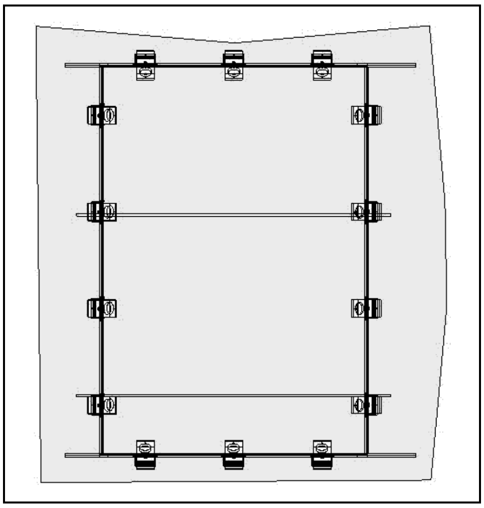

Fig. 11: Rectangular studio setup top view (CAD)

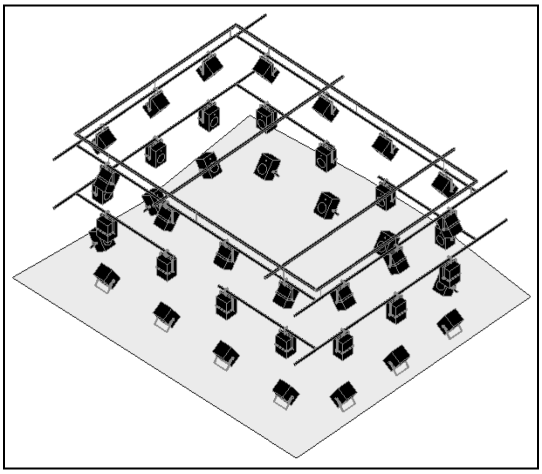

Fig. 12: Rectangular studio setup 3d view (CAD)

V. Approaches to Composition, Spatialisation and Diffusion

In this section a number of compositional approaches are described as well as the corresponding spatialisation strategies for the above-described loudspeaker configurations. The spatialisation approaches range from vector panning and ambisonics to non-standard algorithms which were part of individual pieces. In the following a number of compositions are briefly described.

A. Dome:

1) Schwarm: (Collective Composition)

In the initial state of this installation bird-like sounds concentrate on a random position on the floor of the audience space like animals having a rest.

The visitor may walk around freely but if he comes too close to the resting birds, he shoos them up into flock. After flying around in the three-dimensional space of the audience hall for a certain time the flock settles down at a new location on the floor. The position of the visitors is obtained through a video tracking system running on a separate CPU [16] .

B. Complete-Sphere:

1) Fetzenfische (L. Fuetterer)

Fetzenfische is an algorithmic composition realised with Reaktor (Native Instruments) and Max/Msp. The spatialisation for 42 audio outputs is based on two identical algorithms. The signal of one algorithm runs over 2 channels (a, b). Each position of a speaker in the sphere is assigned to a matrix which gives adjacent positions. After the completion of a movement from a to b the source moves on according to the values of the matrix.

2) Tension II (S. Rosinski)

Tension II is dominated by white noise. The motion of the noise sources is bound to a vertical plane in a three-dimensional sound space. Further sounds are distributed on randomly generated orbits around the vertical axis

3) One is missing (A. Rafinski)

One is missing is a 38-channel composition based on sine waves. The volume of the sine waves is altered casually. Additional noise sources circle around the listener.

4) radial (P. Modler)

radial is based on combinations of sounds, spatialisation tools and abstract-motion objects, e.g. a repetitive object moves a sound on the skin of the sphere from bottom to top. This vertical motion is contrasted by spiral movements of clicks. An outer ring of 6 speakers plays back variations of frequency and intensity patterns of sine waves which are intended to create spatial distance for the installation.

C. C one/Tree:

1) 42-Channel Plunderphonic Christmas Song (M. Zielke, D. Loscher)

From recordings of 42 famous Christmas songs only relevant phrases were selected such as, Christmas, tree, a child is born, Jesus, holy night, star, Bethlehem. Each recorded phrase is played back on a certain speaker. The simultaneously played song phrases interfere with each other, creating the impression of a holy cacophony of western consumption habits.

2) Gloria Exorbitante

Gloria Exorbitante is based on recordings of solo voice and trombone. The spatialisation of the piece is realised through a Cubase arrangement that maps single audio tracks to certain audio channels. The channels are ordered as 7 rings, which is equivalent to the conic speaker structure. During the piece audio material moves between vertical rings

D. Stereo Corridor/Pathway:

1) Karlsruhe klingt (D. Loscher)

Every city has its own soundscape: by walking through the corridor of speakers the visitor will hear 21 field recordings characterizing a certain city (Karlsruhe). At the beginning of the row the visitor will here a voice asking how the city sounds. By walking forward the visitor will listen to field recordings, each played back on a certain stereo pair in the corridor. The visitor walks through the city's sounds, such as those of pedestrians, crowds, children playing, skaters in a park, birds, trams, cars, current generators, New Year firecrackers.

2) Ge-h-ör-gang (F. Bierlein)

For this installation the recording of F. Nietzsches poem *Das trunkene Lied* spoken by K. Kinski is divided into 20 samples of similar length. Each sample is played on one loudspeaker pair in loop mode. If the visitor walks through the speaker pathway in the appropriate speed, the looped samples expand to the original poem until the end of the corridor is reached where the words *in Ewigkeit* (in eternity) are continually repeated. Outside the pathway the samples superimpose and morph into a rhythmical mantra of sound and words.

3) Angebot & Nachfrage (A. Kerschkewicz)

Fairground barkers are heard through the loudspeakers. From outside the atmosphere of a fairground is clearly observable. If the visitor enters the speaker corridor he experiences that the barkers don't linger at fixed positions, but instead hurry through the corridor unpredictably. A random generator controls their direction and speed.

4) Katze (A. Kerschkewicz)

If the visitor listens to the speaker corridor from outside it sounds as if a huge purring cat is lying in the pathway. If the visitor enters the corridor the purring of single cats can be heard from of the speakers. The sounds move in different speeds through the corridor and create a rhythm of the purring

E. Mixed Setups

1) Mundspülung (E. Farchmin)

This piece is based on recordings of a beat-box performance in which a beat detection (maxmsp) maps a beat to a certain speaker. Different audio buses of a Cubase arrangement reflect the mixed speaker setup. In the arrangement sound files are mapped by hand to the different speaker architectures to create circular movements which turn into spiral motions.

2) Untitled 3 (T. Myatt)

The speaker configuration of *untitled 3* was based on theories of spatial perception developed by Myatt and Lennox [13] and was designed to experiment with nested fields of loudspeakers to create positional cues calculated to increase amount of spatial audio information presented to the listener in the playback environment, e.g. by exploiting parallax displacement as listeners move, and through the presentation of simplified simultaneous coherent soundfields [14] and sound movements within spatial audio environments [12]. Audio spatialisation was realised with combinations of 3rd order Ambisonic and VBAP decoders [9] (Fig. 8. to Fig. 10).

VI. RESULTS

The speaker herd project provides the possibility to experiment with a larger number of similar loudspeakers. The audio distribution system offers sufficient CPU power for running spatialisation tools in real time and to a large extent for composition and sound generation approaches.

For the realisation of a certain speaker architecture quite a large amount of time was spent planning the physical setup of the speakers, such as precise rigging, suspension, etc. The use of a CAD system supported the design processby providing both fast access to ways and means and also a visual feedback of the projected speaker setup.

The routing of loudspeaker channels was done on a software level (such as the Max/MSP I/O Mappings page) and checked by assigning noise to every channel.

Line-level adjustment of the outputs of the MOTU 24IO interface in combination with the Genelec 8040A speakers turned out to be difficult, as the 24IO outputs do not provide a level reduction. For most cases the input level of the 8040s had to be adapted to the high audio level the MOTU interface delivers.

The pieces so far realised on the speaker herd may be categorised into pieces which try to approach audio distribution through a source/space model (e.g. ambisonics) or a non-standard diffusion approach (e.g. random distributions) or a combination of both.

Quite a number of pieces tried to leave the source/space model and define their own way of distributing audio onto a large number of speakers. F. Bierlein's piece, for example, followed the paradigm of a granulated spoken sentence where each grain of the sentence is sent to a certain speaker at a certain time. The listener may then perceive the sentence sequentially by walking through the whole line of speakers. L. Fuetterer's piece for the sound sphere also does not use a standard spatialisation approach but instead distributes grains of sound material to speakers depending on algorithmic computation. In contrast to this approaches the swarm uses the ICST tools for ambisonic spatialisation with quite a large number of virtual audio sources in space and controlled random movements to create the impression of a swarm of flying animals.

Although a detailed evaluation of the spatialisation software is beyond the scope of this paper, the following aspects were realised during the project. The ICST ambisonics Max/Msp tools provide a sophisticated and elaborate interface for ambisonics in Max/Msp. Although the tools were used in several compositions we are far from having explored all possibilities. An interesting point in the ambisonics realisation was that height worked well in a straight-plane setup but not in a speaker setup distributed over three dimensions. We experienced this with a classical dome setup and with the full sphere setup where in both the height of audio sources especially below the ear were difficult to reproduce.

VII. CONCLUSIONS

The speaker herd provided exceptional possibilities for experimenting with loudspeaker architectures and sound distribution and spatialisation algorithms. The selected hardware and software configuration provided a robust and extendable basis for various speaker architectures, audio distribution approaches and compositions. Although for certain compositions and situations speakers with higher power output may be desirable, the speakers of our system provided a very high audio quality combined with compact dimensions, which enabled flexible and non-standard speaker setups.

The complete sphere setup explores sound sources from below ear level which is not reflected in most of the standard approaches, e.g. domes or wave-field synthesis. The stereo pathway setup animates listeners to move along a corridor of speakers, whereas the cone/tree setup explores the acoustics of the installation space due to its circular acoustic radiation.

A number of further speaker architectures are planned to be realised based on this system and a collaborative group of artists, musicians and composers.

ACKNOWLEDGMENTS

This work would not have been possible without the advice and help of T. Myatt. The presented material emerged from rAUMkLANGbEWEGUNG a course directed by P. Modler at the University of Media, Arts and Design, Department of Media Art – Sound, Karlsruhe. We would like to thank T. Troge, Hochschule für Musik, Karlsruhe for his support. L. Fütterer provided the CAD-drawings.
This work was realised with the support of a grant from the Ministry of Education and Research of Baden-Württemberg, Germany.

REFERENCES

[1] B. Shield, T., Cox, Concert Hall Acoustics: Art and Science, http://www.acoustics.salford.ac.uk/acoustics_info/concert_hall_acoustics/?content=musical_acoustics, 1999/2000

[2] http://www.music.columbia.edu/masterpieces/notes/varese/notes.html, 2008

[3] R. Pellegrini, M. Rosenthal, and C. Kuhn, "Wave field synthesis: open system architecture using distributed processing," in Forum Acusticum, Budapest, Hungary, September 2005.

[4] Austin, Larry (2001). "Sound Diffusion in Composition and Performance Practice II: An Interview with Ambrose Field". Computer Music Journal 25 (4): 24. MIT Press.

[5] Harrison, Jonty (1988), "Space and the BEAST concert diffusion system", written at Ohain, in Francis Dhomont, L'espace du son, Musiques et Recherches

[6] http://www.genelec.com/technical-documents/, 2008

[7] Jan C. Schacher, Philippe Kocher, Ambisonics Spatialization Tools for Max/MSP, Proceedings of the 2006 International Computer Music Conference, New Orleans, 2006

[8] Jot, J.M., Warusfel, O. 1995a "Spat: a spatial processor for musicians and sound engineers", Proc. CIARM'95 Conference, Ferrara (Italie), Mai 1995.

[9] Pulkki, V. Creating generic soundscapes in multi-channel panning in Csound synthesis software. Org. Sound 3, 2 (Aug. 1998), 129-134.

[10] http://www.steinberg.net/, 2008

[11] http://on1.zkm.de/zkm/stories/storyReader$5970

[12] Lennox, P. Myatt, A. and Vaughan, J., 3D audio as an information environment, Proc. of AES 19th International Conference on Surround Sound Techniques, Schloss Elmau, Germany, 2001

[13] Myatt, A. Lennox, P. A Perceptual Approach to the Composition of Meaning in Artificial Spatial Audio Proc. AES Conference 30th International Conference: Intelligent Audio, Saarielkä, Finland, 2007

[14] Myatt, A. Lennox, P. Vaughan, J. From Surround Sound to True 3D, Proceedings of Audio Engineering Society 16th International Conference, AES, Rovanearme, Finland, 1999

[15] Myatt, T, Malham, DG, Three Dimensional Sound Spatialisation Using Ambisonic Techniques, Computer Music Journal, Vol.19 No.4, MIT Press, Cambridge MA,USA, pp58-70, 1995,

[16] Modler P, Myatt A., Image features based on 2-dimensional FFT for gesture analysis and recognition, Proceedings of the 4th Sound Music Computing Conf. Lefkada, 2007