Progress in the Application of 3-Dimensional Ambisonic Sound Systems to Computer Music

Dave Malham, Richard Grton
Music Technology Group
University of York
Heslington, York
YO1 5DD
United Kingdom

Email: dgm2, ro1 @ uk.ac.york.vaxa

ABSTRACT: Since the report on Ambisonics at the ICMC Glasgow 1990, considerable progress has been made in the application of Ambisonic 3-dimensional sound positioning techniques to computer music. Large-scale public multi-media concerts have been held and experience from these has led to refinements in decoder technology. This is reported upon, as is the development of a real-time fully digital MIDI-controlled Ambisonic processor for sound positioning and diffusion.

Ambisonics

To recap briefly, Ambisonics is a method of recording information about a soundfield and reproducing it over some form of loudspeaker array so as to produce the illusion of hearing a true three-dimensional sound image. We deliberately stress ‘illusion’ to emphasise that if one truly wished to reproduce the soundfield present in a two metre sphere – up to say 20 kHz – then from information theory it is possible to show that one would need 400,000 channels and loudspeakers.

The Ambisonic surround sound system is essentially a two part technological solution to the problems of encoding sound directions (and amplitudes) and reproducing them over practical loudspeaker systems in such a way as to fool the ears of listeners into thinking that they are hearing the original sounds correctly located. Ambisonic systems use four channels to carry sounds encoded in three dimensions. These so-called "B" format signals which carry the information necessary to convey the position of a sound within a three-dimensional soundfield may be generated from the following four equations:

1. $X = \cos P \cos QA$ (front–back)
2. $Y = \sin P \cos QA$ (left–right)
3. $Z = \sin QA$ (up–down)
4. $W = 0.707A$ (reference signal)

where $P$ is the anti-clockwise angle from centre front, $Q$ is the elevation and $A$ is the sound.

Ambisonic reproduction requires six or more loudspeakers depending on size of area and other factors. A practical minimum for anything other than a domestic size room is eight. There is no need to consider the actual details of the reproduction system when doing the original recording or synthesis, since if the B format specifications are followed and compatible loudspeaker/decoder setups are used then all will be well. In all other respects the two parts of the system, encoding and decoding, are completely separate.

Since an essentially unlimited number of sounds, at any number of different positions, can be encoded within the same four channels, complex manipulations of large numbers of signals can be accomplished with relatively simple algorithms, a feature which is impractical or unavailable in other systems.

Such features led us to mount a large research effort into the application of Ambisonic technology to the composition and performance of computer music. In the last eighteen months, this research has been moving rapidly from theoretical studies and small scale prototype systems to full scale practical usage. In the next course of this work, a considerable amount of practical experience has built up. Needless to say, this process continues and will do so for the foreseeable future.

ICMC 467
Although at the time of ICMC 1990 we had used the system in a number of smaller concerts with some success our first major use of the system came soon after the conference, with the Music Departments major multi-media production, Electric Zodiac. This was based on the year 2000 edition of an annual 'Practical Project' series and the first to be run with the involvement of the Music Technology Group. It was intended from the start to make Ambisonic sound projection a major feature of the production so work was undertaken to upgrade the programmable speaker decoder (demonstrated at ICMC 1990) in order to double its maximum capability of four loudspeaker pairs to eight. This enabled us to use twelve main speaker clusters in our concert hall, together with bass bins, and two 'fill-in' speakers not included in the figure (see Fig. 1.)

Speaker decoders for wide–height (3–D) Ambisonic sound projection currently rely on a decoding theory which requires the speakers to be arranged in diametrically opposed pairs, at equal distances from a central point and reasonably evenly disposed on the surface of the sphere. We soon discovered what happens if the disposition is too uneven. With the practical limitations we had in our concert hall, there was an insufficient height component. The software which drives the decoder attempts to compensate for this by driving the high speakers somewhat harder. Unfortunately, when there is a greater than a two–to–one spread between the RMS values of the three dimensions of all the speaker pairs, this simply results in the sounds being louder in some directions than others. A temporary solution was provided by exaggerating the height of the speaker pairs in the input to the decoder program. This corrected the loudness variations, at the expense of a more diffuse image when sounds moved above or below the audience.

In the short term this approach will be replaced by a more subtle one, where the shape of the soundfield will be altered prior to applying the decoding equations. In the longer term, a more flexible approach to the design of the decoding equations is being sought, without the limitation of requiring diametrically opposed pairs of speakers. However, this is not a trivial task, since one of the main reasons for introducing the concept of diametrically opposed pairs into decoder theory was to simplify the problem to the point where it can be handled by relatively low-level matrix manipulations. We suspect that rather than develop a mathematical formula which would attempt to simply calculate the required speaker feed signals, the answer may lie in the use of some form of iterative optimisation algorithm. Our success in the use of low amplitude fill-in speakers to correct an anomaly in the front image (caused by the difference in acoustic response of the forward speaker positions) lends credence to this approach, since this was essentially a human driver optimisation, not allowed for in the existing theory.

The mention of differences in acoustic response brings us to another lesson learnt during the course of this and other productions. It is very important that listeners get a similar response from all speakers, but particularly from the members of any one pair. Whilst the variation in loudness for differing listening positions translates into a relatively benign warping of the perceived soundfield, which reasonably closely follows what would happen in a natural soundfield, differences in frequency responses and patterns of early reflections can cause unnatural changes and disruptions of perceived directionality. This dictates a need to carefully align the speakers so that, for instance, as large a proportion as possible of the audience is sitting in matching portions of their radiation patterns.

**Fig 1: Speaker layout for the 'Electric Zodiac' concert**
Compositional Applications of Ambisonics

Ambisonics may be employed either as an integral part of the compositional system, or it may be used in diffusion at performance time. We focus here on the former aspect.

By utilising B Format, sonic location can be made integral to the compositional process, rather than submitting it to the vagaries of on-the-spot interpretation. The larger issue of what precisely constitutes "interpretation" we do not consider here: the main point is that the spatial positioning and movement of sounds can be an assured part of the structure of the composition. The composer has control over the whole 3-dimensional listening space, and a control that allows the location of sound to interact with other sonic parameters in a precise and intricate manner.

The provision of large-scale Ambisonic concert facilities is, so far, in its infancy, but we have every confidence that the practical experience of the use of Ambisonics for electroacoustic music, and especially the compositional use of Ambisonics either in the digital domain on the CDP system, or in the analogue domain with studio devices will provide increasingly accurate comparative judgements with more conventional diffusion techniques.

The number of composers who are using Ambisonics, either within the CDP system with the "SurroundSound" program, or using more conventional, analogue equipment within our studio, is growing all the time. Apart from the authors and Tony Myatt, we should mention Cathy Lane, an ex-Music Technology student, Tony Hood, who is doing a doctorate in composition at York and who has put much effort into extending and refining Sile O'Modhrain's SurroundSound, and Ambrose Field, currently a music undergraduate at York, who has been experimenting with the inclusion of dummy head recordings into Ambisonic soundfields, as part of the composition process.

Future Work

On the soundfield manipulation side, we are currently putting much effort into developing a real-time, fully digital soundfield processing unit with MIDI control capability (Fig. 2). This uses a low cost DSP board, originally developed by our Computer Science department, which is based on a 40MHz TMS320C25. The built in MIDI links are used for dumping program code to the DSP during development work and for controlling the board, if necessary, during operation. We have added four channels of A to D and D to A conversion, which can run up to 48kHz sampling rate (16 bit) and the ability to transfer audio data directly to and from an Atari ST computer so that it can be linked to the CDP environment. The implementation of the simpler manipulations (rotate, tumble, tilt, etc.) are proceeding well and by the time of the conference we hope also to be able to report progress on some of the more esoteric frequency-dependent effects such as
spectral spreading and true distance cueing. The development of this device is a vital component in our progress towards an integrated sound performance environment in which Ambisonic sound diffusion is used either on its own or in combination with other, more conventional diffusion techniques such as those used by Dennis Smalley at the University of East Anglia or the members of BEAST (Birmingham Electro-Acoustic Sound Theatre), with both of whom we are arranging exchange concerts.

Useful Reading

People planning to produce devices based on Ambisonic should note that commercial exploitation is covered by a comprehensive set of patents which are held by NIMBUS RECORDS. All enquiries should be directed to:
Stuart H. Garman, Nimbus Records
Wyaston Leys, MONMOUTH NP4 3SR
Great Britain Tel. (0600) 490962

2) GASKELL, P.S. "Spherical Harmonic Analysis and Some Applications to Surround Sound" BBC RESEARCH DEPARTMENT REPORT NO. BBC RD 1979/25 (1979)
4) GERZON, M.A. "Panpot and Soundfield Controls" NRDC AMBISONIC TECHNOLOGY REPORT NO. 9, August 1975.
7) FELLOGETT, P.B. "From Quadro to Surround Sound" ELEKTOR
GERZON, M.A. May 1977 pp 5/19 – 5/25
8) GERZON, M.A. "Surround Sound Decoders" in 7 parts. WIRELESS WORLD 1977
15) MALHAM, D.G. "Computer Control of Ambisonic Soundfields" Preprint No. 2460 (H2) presented at the 82nd AES convention 1877 10–13 March, London
Other relevant (British) patent numbers include 2073556 (Periphonic decoder), 1411994 (unequal speaker arrays), 155627, 1494761, 1447652, 1512514

ICMC 470