Surround sound, yesterday, today and tomorrow

By Dave Malham

For almost as long as we have been capable of reproducing sounds remotely from their original source, people have been attempting to include spatial information in the reproduction of an acoustic event. There was little, of course, that we could do with the original acoustic gramophones, but with the appearance of electronic technology and, in particular, good quality microphones, we could start to think about doing much more.

For a long time, the problem was the distribution media that were available, but with the appearance of the new generation of multichannel consumer sound carriers, such as DVD, systems for presenting full audio images have become truly practicable. Of course, even the best of the so-called surround sound systems, such as Ambisonics, are severely limited compared to our ears, but that should not stop us trying.

Microphones, of course, have a vital part to play in this search for the missing spatial elements in recordings and broadcasts – for these elements are missing and we do need to get them back. When we think of a beautiful moorland scene, it is the shape of the hills and the purple of the heather, with the white dots of sheep that we pay most attention to. The sound of the wind moving through that heather, the calling of the sheep and the plaintive cry of the curlew tend to be in the background. Yet imagine that scene without the sounds. Without them, it ceases to be a living, breathing part of nature, losing much of its depth and becoming just another picture. When we experience a concert in a wonderful concert hall, we are bathed in a sea of sound coming from all directions, but this enhances, rather than compromises, our enjoyment since our ears are truly amazing at sorting it all out.

Our hearing is the only one of our five senses which is truly capable of providing us with three-dimensional information about remote events. We are able to perceive where acoustic sources are in the space around us, left or right, front or back, above or below, whether they are moving or stationary. We can even estimate the distance of a sound source as well as get some idea of its nature and size. Not only that, but for us, as for the other life forms that are blessed with directional hearing, the 3-D sonic environment is one from which we cannot escape since, unlike sight, hearing cannot be cut off by anything as simple as night or a blindfold.

Historical overview

Over a hundred years ago, Clément Ader used multiple sets of telephone transmitters (microphones) and receivers (earphones) in the first multichannel audio transmission system. Although it is unclear whether he originally intended to give some sense of sound source position via the use of spatially separated microphones, the effect was certainly noted and remarked upon by many visitors to his exhibit at the 1881 Paris Exhibition of Electricity.

The lack of a suitable recording technology, or indeed of any practicable way to allow more than one person to listen to the output of any particular pair of transmitters, meant that no further work was done on multichannel audio for almost half a century. The limitations imposed by the lack of an adequate technology in some area or another of the implementation of what might be termed ‘spatial sound’ is a theme that runs throughout the whole history of audio technology, although it seems possible that we may now be nearing its end.

In the 1920s, the first attempts at binaural, headphone-based spatial presentation of sounds were made by Dr Harvey Fletcher and his team at Bell Labs. This system was not pursued further at the time because of the lack of a two-

channel recording medium. Of course, there was also the unsuitability of binaural for presenting a spatial sound image over loudspeakers, at least to audiences of any size, though it can be very useful for the domestic situation.

By the 1930s, Bell Labs were developing a system based on a curtain of microphones suspended in front of the original sound source, each microphone being used to feed a similarly placed loudspeaker. This probably represents the first attempt at spatial representation based on wavefront reconstruction. Once again, the lack of a suitable multichannel carrier hampered the practical application of this system. However, a simplified version using three spaced microphones feeding matching loudspeakers set left, centre and right proved to be capable of providing acceptable results, particularly for cinema systems. This was fortuitous since the only editable multichannel medium available at the time was based on the optical technology developed for recording film sound. This is essentially the same system as is used for cinema and home theatre surround today, albeit with the addition of one, two or more surround channels.

The Bell Laboratory technology was stretched to the limit (and beyond) by Walt Disney's 1939 film, *Fantasi*a. Very few cinemas ever showed the film with the sound projected as it was originally intended. Nine optical recorders, synced together, were used for the original recording. The production version of the film used a four-track optical playback system separate from, but synced to, the picture. A system involving the use of notches on the edge of film stock controlled the distribution of the sound to multiple loudspeakers arranged around the auditorium. This must surely have been one of the first uses of automation in audio. Unfortunately, the cost and complexity of the system prevented it being used any further. In any case, it was never likely to become a standard for home use!

Stereo

At the same time, in the early 1930s, spatial audio systems based on creating an illusion capable of mimicking the original audio image had been developed in the UK, principally by Alan Blumlein and his team at EMI. Blumlein's system relied on the way that human ears receive signals from a pair of loudspeakers to produce a stereo image. Differing amounts of a sound fed to loudspeakers mounted to the left and right of the listener, at an angle of about 60°, add at the ears in such a way that the difference between what the left ear hears and what the right ear hears mimics the signals that would occur for a real sound source placed in the space between the two speakers – one of the few times that crosstalk is a good, rather than a bad, thing. The position in which the sound source is heard is dependent not only on the amplitude difference between the two speaker signals, but also on the frequencies involved, though fortunately fairly weakly. This addition at the ears only produces stable images when the listener is facing the speakers.

The simplicity of using two channels to convey the image had (and still has) considerable appeal, especially since progress had also been made at the same time on two-channel versions of standard disc systems. Furthermore, the technology was available to build up an image from scratch, using amplitude panning to place multiple sources into an artificial image, as well as to capture real images in a simple way, using a crossed pair of directional microphones. As a bonus the speaker layout was compatible with the two-channel version of the spaced microphone technique developed by Bell Labs.

The term 'stereophonic' has become synonymous with this form of two-channel material. The term, of course, actually implies that a 'solid sound' is being reproduced. This is hardly an accurate description of a sound image where the whole of a 3-D acoustic environment has to be mapped onto the space between two speakers which, for the system to work, should themselves be not more than 60° apart when viewed from the listener's position. However, since it did not overburden the technology of the available media, it became the *de facto* standard for delivery of high quality audio into the home. Moreover, the recoding techniques used can be almost as simple or as complex as you like.

If you like to think of yourself as a 'purist', looking for a clean, simple, uncluttered sound image, then the crossed pair of directional microphones might be what you would use. These days, it is normal to use cardioids to avoid the image sounding excessively reverberant, but Blumlein originally used figure-of-eight microphones. If you prefer a sweeter, more 'spacious' image, though perhaps one which is a little less clear, then a spaced pair is a more likely choice, with the microphones commonly being omnidirectional. In both cases, the likelihood is that the microphones chosen will tend to be neutral in character. On the other hand, if you like to be in total control, you may prefer to have separate microphones for most or all the major sound sources in your image, using the directional characteristics of the microphones and their spacings to separate the sound sources rather than to capture the existing spatial information (see Figure 1). In this case, the tonal characters of the microphones used may well be chosen to enhance the sounds of the instruments they are associated with. Of course, positions anywhere between these two approaches are common, perhaps even more common than the two extremes themselves.

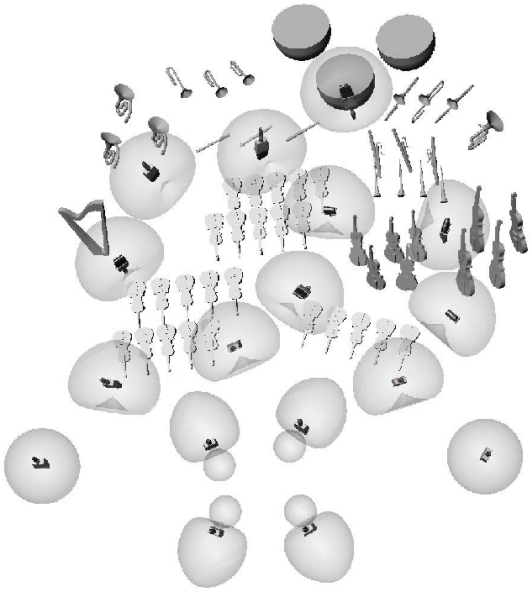

Figure 1

With the rise in surround sound in the 1990s, driven by the cinema, home theatre and computer games market, as well as the needs of virtual reality systems, the distinction is becoming even more blurred. Except for binaural systems where two microphones are used in the 'ears' of a dummy head, current approaches to recording for surround systems all use more than two microphones, frequently many more, even when individual sound sources are not being miked.

Surround sound

For surround sound, the equivalent to crossed-pair purist stereo is a surround version of the Blumlein approach. This is extended to three crossed figure-of-eight microphones at right angles to each other, plus an omnidirectional microphone, all mounted as close to each other as possible (see Figure 2) or perhaps effectively totally coincident as the outputs of the electronically compensated array in the Soundfield microphone (see Figure 3).

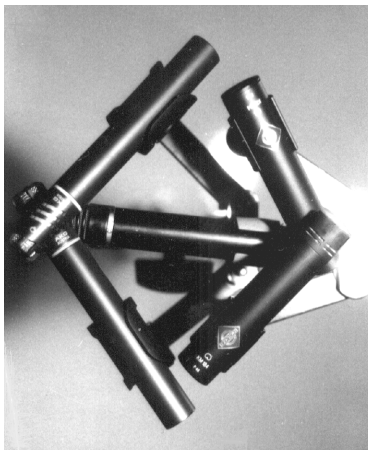

Figure 2

Figure 3

The Ambisonic approach can give a central listener an almost holographic sense of 'being

there' in the original recording location. It has not, however, achieved widespread acceptance. This is at least partly because in its current form it does not match the needs of film surround sound. In film and video use, it is imperative that the visual image is paramount so, for the most part, cinema surround sound systems concentrate on getting the stereo image in the screen area right and largely keep the surrounds decoupled from the front which avoids taking the audience's attention away from the screen. This has led to loudspeaker layouts which concentrate resources on the screen area and only barely cover the rest of the horizontal plane and do not address the vertical plane (see Figure 4).

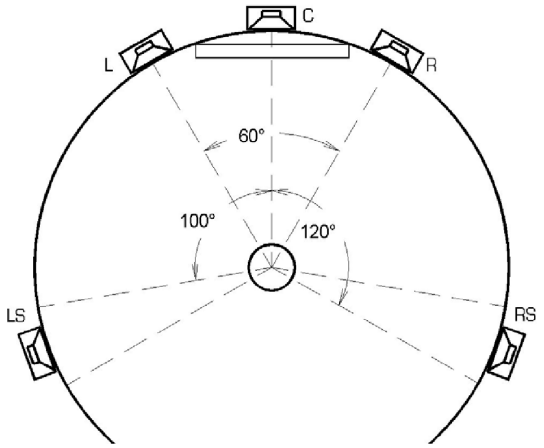

Figure 4

The microphone techniques used to accomplish this are generally more in line with the original Bell Labs approach than the Blumlein/ Ambisonics one. It is much more common to see large numbers of spaced microphones in use than, say, front and rear crossed pairs. This system is, perhaps regrettably, not ideal for use for music recording because of the concentration on a very small part (the front 60°) of the image, at the expense of the rest, but is still being pressed into use for this purpose. In skilled hands, this can be made to serve well for music recordings cast in the western tradition of presenting music on a small stage in front of an audience, but it cannot wholly immerse the listener in another acoustic. It is, however, at least some sort of window, no matter how murky, into what might become possible in future when, as available bandwidth goes up, we can use more and more microphones driving more and more loudspeakers in an ever closer approach to what happens in the real, acoustical world.

Dave Malham is an Experimental Officer at the Music Department, University of York. He has a broad practical and theoretical knowledge of 2- and 3-D sound systems, has published many articles in this field, and has acted as a consultant for multichannel audio sound systems.

First published Microphone Data Book 2001