



Stereo Microphone Techniques Explained

Part 1

Technique : Theory + Technical

Published in SOS February 1997
 [Print article : Close window](#)

PART 1: HUGH ROBJOHNS takes a historical look at stereo miking techniques and explains the whys and wherefores of the various methods available.

The first documented stereo microphone system was used (entirely by accident, in fact) at the great Electrical Exhibition in Paris in 1881. A French designer by the name of Clement Ader was demonstrating some improvements to an early telephone system, and stumbled across what we would now call the spaced-microphone stereo technique! Unfortunately, no one realised the significance of Ader's discovery and he went on to invent the inflatable bicycle tyre before playing with aeroplanes, calling his first plane 'A vion', which became the generic name for aeroplanes in the French language.

Most of the development of stereo recording as we know it today happened in the very early '30s, and almost simultaneously in America and the UK. In the USA, Bell Laboratories were working on systems using spaced microphones under the direction of Dr Harvey Fletcher. Meanwhile, in the UK, a very clever man called Alan Blumlein, working for EMI, was developing an alternative system which relied on coincident microphones.

Both methods were years ahead of their time and both had advantages and disadvantages. It was not until the invention of PVC in the '50s (which allowed micro-groove vinyl records to be produced) that either of these techniques were adopted commercially, but today both formats are alive and well, and are often used in concert with each other.

In this article, I'll be looking at what stereo microphone systems are trying to achieve, also taking a closer look at the coincident stereo ideas which have become the mainstay of many practical recording techniques. Next month, I'll talk about spaced microphone systems and combinatorial techniques.

WHAT IS STEREO?

The word 'stereophonic' is actually derived from Greek, and means 'solid sound', referring to the construction of believable, solid, stable sound images, regardless of how many loudspeakers are used. It can be applied to surround-sound systems as well as to simple two-channel techniques -- indeed, in the cinema, the original Dolby Surround system was called Dolby Stereo, even though it was a four-channel system! However, most people are conditioned to think of stereo as a two-channel system, and this is the definition I'll adopt in these articles.

There are basically three ways of creating stereo sound images over a pair of loudspeakers:

- * The first is an entirely artificial technique based on Alan Blumlein's work, and uses pan pots to position the sound images from individual microphones by sending different proportions of each microphone to the two channels.
- * The second technique (and one we will look at in more detail next month) is the use of two or more identical but spaced microphones. These microphones capture sounds at differing times because of their physical separation, and so record time-of-arrival information in the two channels.
- * The third system is that of coincident microphones, and this has become the backbone of all radio, television, and a lot of commercial stereo recordings. This technique uses a pair of identical directional microphones, each feeding one channel. The microphones capture sound sources in differing levels between the two channels, much like the pan-pot system, but this time the signal amplitudes vary in direct relation to the physical angle between microphones and sound sources.

COINCIDENT MICROPHONES

Blumlein developed coincident techniques to overcome the inherent deficiencies (as he saw them) of the spaced microphone systems being developed in America. Since our hearing mechanism relies heavily on timing information (see 'The Human Hearing Process' box), Dr Harvey Fletcher thought it reasonable to use microphones to capture similar timing differences, and that is exactly what the spaced microphone system does.

However, when sound is replayed over loudspeakers, both ears hear both speakers, so we actually receive a very complex pattern of timing differences, involving the real timing differences from each speaker to both ears, plus the recorded timing differences from the microphones. This arrangement tends to produce rather vague positional information, and if the two channels are combined to produce a mono signal, comb-filtering effects can often be heard.

Blumlein demonstrated that by using only the amplitude differences between the two loudspeakers, it was possible to fool the human hearing system into translating these into perceived timing differences, and hence stable and accurate image positions. We all take this entirely for granted now, and are quite happy with the notion that moving a pan-pot or balance control to alter the relative amplitudes of a signal in the two channels will alter its position in the stereo image in an entirely predictable and repeatable way.

This process is used every day to create artificial stereo images from multi-miked recordings, but contrary to popular belief, the level difference between the two channels which is necessary to move a sound image all the way to one loudspeaker is not very much. Typically, a 12 to 16dB difference between channels is sufficient to produce a full left or right image, and about 6dB will produce a half-left or right image -- although the exact figures vary with individual listeners, the monitoring equipment and the listening environment.

To create stereo images directly from real life, Blumlein needed to develop a microphone technique which captured level differences between the two channels, but no timing differences. To avoid timing differences, the two microphones must be placed as close together as is physically possible -- hence the term 'Coincident Stereo'. The normal technique is to place the capsule of one microphone immediately above the other, so that they are coincident in the horizontal plane, which is the dimension from which we are trying to recreate image positions (despite hi-fi magazines' claims to the contrary, conventional stereo recording does not encode meaningful height information!). Amplitude differences between the two channels are created through the microphone's own polar patterns, making them more or less sensitive to sounds from various directions. The

choice of polar pattern is the main tool we have for governing the nature of the recorded sound stage.

If you read books on stereo techniques, you'll find a variety of alternative terms used to describe the various methods in use. The kind of coincident stereo discussed here is also known as 'XY' recording (in America and parts of Europe), 'AB' recording (in the BBC and most other European broadcasters), 'crossed pairs', or just plain 'normal stereo'. The term 'AB stereo' takes on a different meaning in the USA, where it is often used to describe spaced microphone arrays -- beware of the potential for confusion!

PRACTICAL TECHNIQUES

In general, we aim to place sound sources around stereo microphones such that they occupy the complete stereo image. If you consider an orchestra, for example, it's usual to have the back row of the violins fully to the left, and the back row of the cellos or basses fully to the right.

To create this spread of sound using crossed cardioids to record the orchestra, it would be necessary to place them directly above the conductor in order to achieve the desired stereo image width. To take another example, crossed figure-of-eights would have to be positioned a long way down the hall to achieve the same stereo width (see Figure 1).

It should be obvious from these comments that in choosing the polar patterns for the microphones, you also determine the physical separation between sound sources and microphones for a given stereo width, and therefore the perspective of the recording. In the example above, the cardioids would give a very close-perspective sound, with little room acoustic and a distorted orchestral balance favouring the close string players over the musicians towards the rear and sides of the orchestra. In contrast, the figure-of-eights would give a much more natural and balanced perspective to the orchestra, but would also capture a great deal of the hall's acoustic, which might make the recording rather more distant than anticipated.

It's quite possible that neither of these basic techniques would produce an entirely satisfactory result, and a compromise might be to use crossed hypercardioid mics (with an acceptance angle of about 150 degrees). More likely, a combination of the two original techniques, plus a scattering of close 'spot' mics to reinforce the weaker sections of the orchestra (using pan-pots to match their stereo images to the main crossed pairs), would have to be used. The crucial point is that there is no absolutely correct technique, only an array of tools which you must choose and use to obtain the results you want.

COMBINING CROSSED PAIRS AND SPOT MICROPHONES

A very commonly-used technique is combining a crossed pair (to form the basis of a stereo image) with a number of close microphones (to give particular instruments more presence and definition in the mix). This applies equally whether we're talking about recording a philharmonic orchestra or a drum kit -- only the scale of the job changes; the techniques do not.

There are three things to consider with this combination technique: image position, perspective and timing.

The main stereo pair will establish image positions for each instrument and the close microphones must not contradict this, if we're to avoid confused and messy stereo images. The best technique I know for setting the panning for the close microphones is to concentrate on a particular instrument's image position in the main pair, then slowly fade up the corresponding spot mic and hear how the image changes. If it pulls to the right, fade the spot mic down, adjust the pan-pot to the left (or vice versa) and try again. With practice, you should be able to match image positions in three or four cycles, such that fading the spot mic up only changes the instrument's perspective, not its position.

Clearly, a microphone close to an instrument will have a completely different perspective to one further away, and this contrast is usually undesirable, as it draws undue attention to the instrument in question. The relative balance between the 'spot' mic and the main pair is critical, and it's surprising how little a contribution is required from the close mic in order to sharpen the instrument's definition, which is normally all you're trying to achieve. Remember, if you're aware of the close mic, it's too high in the mix.

"In general, we aim to place sound sources around stereo mics such as that they occupy the complete stereo image."

The last point is relative timing, but this is usually only a problem with large recording venues. Consider an orchestral recording again, where the main stereo pair of, say, hypercardioids, may be 50 or 60 feet away from the orchestra. As sound travels at about one foot every millisecond, the sound from the stereo pair will be about 60ms behind that from any close spot mics. The human hearing system is geared up to analyse the first arriving sounds, which means we naturally tend to be aware of sound from the spot mics before the main pair -- almost irrespective of how low they are in the mix. This is not the situation we want -- the spot mics are supposed to assist the main stereo pair, not the other way around!

The solution is to route all the spot mics to a stereo group (having balanced and panned them appropriately) and send the combined signal to a stereo delay line. Dial in a suitable delay (one millisecond per foot for the distance between the main pair and the most distant spot mic, and then add five to ten milliseconds for good measure). The output of the delay line is returned to the desk and mixed in with the main stereo pair to produce the final mix. By delaying the spot mics, you can cause their signals to be heard after the main stereo pair (by the five or ten milliseconds that were added), and they'll consequently be much harder to perceive as separate entities. In fact, delaying the close mics makes their level in the mix slightly less critical, as the hearing process takes less notice of them, although their panning is still crucial, of course.

This technique is extremely effective, but is rather time-consuming, and few people would bother with it if the main stereo pair was less than about 20 feet from any spot mic.

M&S COINCIDENT TECHNIQUES

There is an alternative coincident stereo technique, again developed originally by Alan Blumlein. This is the M&S, or Mid & Side, technique, mainly used by television sound recordists, but definitely worth knowing about, whatever you record.

M&S is a coincident technique in exactly the same way as the conventional systems already described. Instead of having directional microphones facing partially left and right, the M&S technique uses a pair of microphones, one of any polar pattern you like facing forwards and the other, a figure-of-eight, facing sideways. These two signals have to go through a conversion process before being auditioned on loudspeakers or headphones as normal left-right stereo.

The M&S system offers a number of practical advantages for television sound recordists (which are outside the scope of this article), but the single most useful aspect of the system for everyday recording tasks is that the perceived spread of sound sources across the stereo image can be controlled very easily from the desk.

The most common arrangement is to use a cardioid microphone facing forwards (the 'M' mic), together with a figure-of-eight microphone (the 'S' mic) facing sideways, and when these are converted into normal left-right stereo, they produce an identical acceptance angle to conventional crossed cardioids (see Figure 2). One important point to note: the polarity of the S lobe facing left should be the same as the polarity of the M mic. If this is not the case, the stereo image will be reversed.

As the balance between the M and S microphones is altered, so is the apparent distance between sound sources, as heard on the speakers (the effect is similar to adjusting the mutual angle between a conventional crossed pair of mics; see 'Terminology' box for more on this). This can be used to great effect, and it also allows the image width to be pushed outside the speakers by introducing an out-of-phase element to the signal, although this should be used with great care, as it will affect mono compatibility.

SOUNDFIELD MICROPHONE

This concept of M&S was extended in the design of the Soundfield microphone and its baby brother, the ST250. These microphones were originally developed for Ambisonic recording -- a technique which captures and reproduces true surround sound, with height information as well as 360-degree horizontal imaging (as opposed to the entirely artificial spatial positioning of the various cinema surround systems).

Unfortunately, Ambisonics has never really caught on and although a few companies are producing material suitably encoded (such as classical recordings from Nimbus), most people use the soundfield microphones as glorified, but stunningly accurate, stereo mics.

The soundfield microphones have an array of four cardioid capsules, arranged as the sides of a tetrahedron (two pyramids joined base-to-base), and these are combined electronically to produce four 'virtual microphones' called W, X, Y and Z. The first output (W) is designed to have an omnidirectional polar pattern, while the other three are figure-of-eights facing left-right, front-back and up-down. The way in which the W, X, Y and Z virtual microphones are created simulates extremely close spacing between capsules, so the stereo imaging is phenomenally accurate.

These four signals are combined together to produce a stereo output according to the settings on the control unit, in much the same way as the basic M&S arrangement described earlier. The omni (W) signal can be thought of as equating to the M microphone in a simple M&S pair, and the X, Y and Z signals equate to the S microphone, albeit with separate microphones for each direction (up/down, left/right and front/back).

The control unit allows the user to manipulate the Soundfield mic's characteristics to unprecedented degrees. The effective polar patterns of the simulated stereo pair can be selected, as can their mutual angle, and then this virtual stereo array can be pointed and tilted in any direction, simply by manipulating the way in which the four signals are combined. One of the most amazing aspects of the soundfield microphone is that by changing the balance between the W signal and all of the others, the mic can be made to appear to 'zoom in' to the sound source! It is even possible to record the four base signals individually (called the B-format) and then use the control unit to manipulate the microphone's characteristics on playback.

Next month, we'll look at spaced microphone arrays such as the Decca Tree and Binaural recording, as well as some of the more popular combinatorial techniques.

THE HUMAN HEARING PROCESS

The whole idea of stereo recording is to try to fool our auditory system into believing that a sound source occupies a specific position in space. So how does our hearing determine the positions of sounds around us in real life?

Without getting bogged down in the psychology and biology of the subject, we use three principal mechanisms to identify the positions of sounds around us. The first and probably most important one is that of differing arrival times of sounds at each ear, followed by level differences between the ears for high-frequency sounds, and finally, independent comb-filtering effects from the outer ear (the pinnae).

Since our ears are spaced apart on opposite sides of the head, any sound source off to one side will be heard by one ear fractionally before the other. Also, because there's a large solid object between the ears (the rest of the head), a 'sound shadow' will be created at high frequencies (above about 2kHz) for the distant ear.

Both these mechanisms highlight the possibility of confusion between the direction of a sound source at any given angle in front or behind the listener, since both the timing and level differences would produce the same results for both directions. To overcome this ambiguity, an automatic reflex action causes us to instinctively turn or tilt our heads slightly and the resulting changes in timing and level immediately resolve the confusion.

The third mechanism was discovered relatively recently, and is the reason for the bizarre shape of the pinnae. (I always knew they had to be there for something other than supporting glasses and earrings!) As sounds arrive at the outer ear, some of the sound enters the ear canal directly, while some is reflected off the curved surfaces of the outer ear and into the ear canal. Since the reflected sound has to travel fractionally further, it is delayed, and in combining with the original sound, produces a comb-filter effect, resulting in characteristic peaks and notches in the frequency response. These frequency-response anomalies depend on the particular direction of sound arrival, and it is thought that we build a 'library' memory of the comb-filter characteristics which can be used to help provide crude directional cues.

This whole concept of directional perception is the foundation of the sophisticated signal processing used in systems like QSound and RSS, which try to create surround sound information from a conventional two-channel stereo system. Modifying the frequency response of recognisable sounds to simulate the effects of the pinnae can trick us into perceiving sounds from locations outside the normal stereo spread between the loudspeakers.

TERMINOLOGY, RIGGING AND CALIBRATION

Blumlein performed all his experiments using microphones with figure-of-eight polar patterns (only these and omnidirectional mics were available at the time). Most of the time, the figure-of-eight microphones were arranged at 90 degrees to each other, such that one faced 45 degrees left, and the other 45 degrees right. The angle between microphones is called the 'Mutual Angle', and 90 degrees is the most commonly used. It is possible to change the mutual angle over a small range, to adjust the precise relationship between the physical sound source positions in front of the microphones and their perceived positions in the stereo image, although the effect is often very subtle and few people find it necessary to make such adjustments.

The usable working area in front of the microphone is defined by the polar patterns of the microphones, and is called the 'Acceptance Angle'. The diagrams below show the typical acceptance angles for figure-of-eights and cardioids crossed at 90 degrees. Note that because the figure-of-eights are bi-directional, with opposite polarity lobes, they have two acceptance areas and two out-of-phase areas at the sides.

It is essential to calibrate the microphones and their channels at the desk before attempting to record anything in stereo. Even nominally identical microphones will have slightly differing sensitivities, and the input channels in the desk could be set up completely differently -- so it is important to run through a line-up procedure (which is far quicker to do than to read -- honest!)

What we need to achieve is identical signal levels in the left and right desk channels for a given sound pressure level in front of the microphones. The easiest and most accurate technique starts with setting the microphones' polar patterns to the desired response (if using switchable mics) and connecting them to two desk channels (or a stereo channel, if available). Turn the pan pots on paired mono channels fully left and right and use a fader clip (or some other means, such as a large bulldog clip) to mechanically fix the two faders together so they track accurately. Rig the microphones one above the other with their capsules as close together as possible, and turn them to face in the same direction while someone speaks in front of them (about two feet away and at their mid-height, if possible, to ensure minimal level differences).

In the control room, switch the loudspeaker monitoring to mono (do not use the channel pan pots, because their centre positions may not be accurate), and adjust one mic channel for the typical operating gain you expect to need, with the fader in its normal operating position. Check that

there is no EQ in circuit in either channel and switch a phase reverse into the second channel. Adjust the second channel's gain until the combined output from the microphones is as quiet as possible -- there should be a very obvious null point (it will never completely cancel, because of inaccuracies in the microphones and desk channels, but it should get extremely quiet).

Next, remove the phase reversal and loudspeaker mono-ing, and with the two mics still facing forward, have your talking assistant wander in a complete circle all the way around the microphone array. If the stereo image moves away from the centre, the mics have incompatible polar patterns and will not produce accurate stereo images. Select another pair of microphones and start over.

Finally, rotate the microphones to face 45 degrees left and right (make sure the microphone connected to the panned-left channel is turned to face the left of the sound stage) and have your assistant confirm the image boundaries and left-right orientation. Having completed the line-up, do not re-plug the microphones, or adjust the channel gains, as the calibration will be destroyed and you'll have to go through the entire process all over again! In practice, this whole procedure should take about a minute and should become routine.

A lot of engineers use a 'stereo bar' as a more convenient way to mount a pair of mics from a single mic-stand. Although this technique introduces small timing differences into the recording, it is a perfectly acceptable technique, provided the microphones face outwards rather than inwards after the line-up process. The reason for this is that each microphone casts a sound shadow at high frequencies across the other, and if they face inwards this is likely to degrade the stereo image (particularly if the mics in question are physically large, such as C414s, or U87s). If the mics face outwards, the sound shadow will fall on the rear of each microphone, where it is relatively insensitive anyway (assuming cardioid or hypercardioid patterns) and will not cause imaging problems.

DECODING M&S PAIRS

To decode the M&S signals to normal left and right, pan the M microphone to the centre and split the S microphone to feed a pair of adjacent channels (or a single stereo channel). Gang the two S channel faders together, pan them hard left and right, and switch in the phase reverse on the right channel.

Listening with the monitoring switched to mono, balance the gains of the two S channels for minimal output (make sure there is no EQ switched into either channel). Once the two S channels have been aligned, revert to stereo monitoring, fade up the M channel and adjust the balance between the M and S signals for the desired image spread.

Putting a phase reverse in the M channel will swap the stereo image over -- left going to the right and vice versa -- and the image width can be varied from mono, through normal stereo, up to extra wide, simply by moving the S fader up and down.

Published in SOS February 1997

Sound On Sound Ltd is registered in England and Wales. Company number: 3015516. VAT number: GB 638 5307 26.
Registered office: Media House, Trafalgar Way, Bar Hill, Cambridge, CB23 8SQ, United Kingdom.
Email: Contact SOS | Telephone: +44 (0)1954 789888 | Fax: +44 (0)1954 789895

All contents copyright © SOS Publications Group and/or its licensors, 1985-2009. All rights reserved.
The contents of this article are subject to worldwide copyright protection and reproduction in whole or part, whether mechanical or electronic, is expressly forbidden without the prior written consent of the Publishers. Great care has been taken to ensure accuracy in the preparation of this article but neither Sound On Sound Limited nor the publishers can be held responsible for its contents. The views expressed are those of the contributors and not necessarily those of the publishers.

Web site designed & maintained by PB Associates | SOS | Relative Media