

Surround Sound Imaging

by William Sommerwerck

Reprinted from *The QUAD Quarterly*, Vol.2, No.1, June 1980.

What's with *imaging*? It goes ping-pang-pong-pung around the room, nice-like. Who cares?

Well, *I* care! If a surround-sound system does not produce crisp, stable imaging at the sides and rear, the performers and producer are limited in where they can place the sounds. And surprisingly, the brain depends heavily upon the sounds coming from the sides to determine the acoustic character of the space in which the listener finds himself. Correct side imaging is as important for accurate ambience reproduction as it is for pleasing, exciting surround sound.

Unfortunately, proper side and back imaging are not possible with any of the commercially available discrete or matrix systems! A serious mistake was made when surround sound was first introduced. To understand this mistake, we have to look at conventional stereo recording.

In stereo, a voice or instrument is positioned between the speakers by adjusting its relative loudness; the sound moves toward the louder speaker. The same effect can be achieved with two directional mikes (cardioid or figure-8) placed close to each other at an angle of 90 to 120°. Since the microphones' outputs will *not* be equal (except for sounds directly in front), there will be a different level on each channel, creating the desired stereo effect.

With quad, record producers have four *pairs* of speakers to work with: one at the front, one at the rear, and two at the sides. "Common sense" suggested that we could treat each pair of speakers like the front pair: simply adjust the relative loudness of the signal feeding that pair to position the sound. With microphones, we just put four in a square, at 90° to one another. This technique is called *Pair-Wise Mixing* (PWM). The problem is that hardly anybody ever bothered to question if this was the way the brain perceived directionality. Unfortunately, it isn't. PWM doesn't work.

Try listening to conventional stereo and mono recordings with the speakers directly in front, behind, and to the sides. You should notice most of the following things. Imaging behind you is fairly good, but not as precise as when you face the speakers. Sounds appear to be too close, even to the point of "brushing" the back of your neck, or "drilling" into your head. Side imaging is much worse, with oversized, unstable images. Slight movements may cause the sound to jump to the front or rear speaker. There may be an uncomfortable "pressure in the ear" sensation. All this occurs because PWM does not create the proper cues the brain needs for correct imaging. And virtually all four-channel sources, both discrete and matrix, are created using Pair-Wise Mixing techniques. Therefore, none of the popular surround-sound systems can correctly reproduce the full ambience of the recording site, or allow a performer to be placed anywhere desired around the listener.

Don't be confused by claims that the Tate System SQ decoder can produce results indistinguishable from the discrete master tape. It comes very close. But that isn't the point. The master tape is usually derived from Pair-Wise Mixing sources, so it is not correct. Of course, if PWM material is all we'll ever have, we might as well go with Tate; the improvement from full discrete *does not* justify the extra two channels.

But we have two extra channels available to us on tape, and we can add a third channel to stereo FM without any increase in bandwidth. Is there any way to use these channels to improve imaging over Pair-Wise Mixing? To find out, we have to know more about human hearing.

It is known that the ear and the brain determine directions of sounds by computing the difference between the left/right and front/back sound waves around us and comparing these with the overall sound. (Yes, this is an over-simplification.) In recording, these same three basic signals can be obtained using an omni-directional pickup (for the overall signal) combined with directional mikes (for the left/right and front/back figure-8 patterns). And at the mixing console, the same signals can, of course, be generated electronically. (By the way, the mathematical theory of spherical harmonics tells us that only these three pieces of information are needed to accurately position any signal around the listener. Is it so surprising that the process works identically in the human brain?)

I know what some of you are thinking. If we have four speakers and three signals, isn't there some loss of information? Of course not. We are simply creating signals which convey the precise information that the brain needs for exact imaging. This is neither a matrix nor a discrete system. We could care less about the exact speaker feeds. Rather, we simply transmit the minimum correct information to accurately locate the sound, and then allow a simple processing device and the brain to take care of the rest.

Are there any systems using this technique? Yes, there are three: UMX, Ambisonics, and the RCA proposal for surround-sound broadcasts. Since these systems are compatible with one another, let's look at RCA's. The omnidirectional signal is equivalent to a mono pickup, so it gets transmitted as the main channel signal. The left/right figure-8 signal goes on the subcarrier, and provides stereo reception. The front/back figure-8 also goes on the subcarrier, but at 90° to the L/R signal. Conventional receivers will ignore this F/B signal, which gives us *perfect* mono, stereo and surround compatibility, with *no* increase in bandwidth, and only a 3 dB loss in signal-to-noise ratio. No special decoding chip is required to remove the front/back signal; only an extra FM stereo chip and a simple 90° phase-shift network.

Remember, these signals are not the speaker feeds themselves. We can use them to create feeds for four to 4000(!) speakers. The more speakers, the more accurate the side images and the more stable everything is as the listener moves away from the center. However, the practical limit is considered to be six, and a signal processor can provide outputs for four or six speakers, as well as allowing adjustments for their position and for separation.

This isn't just theoretical. I've made recordings with this technique. Since the British have done the most research on this technology, their name for it is most used: Ambisonics. Compared with recordings made with four microphones in a square, Ambisonics is the clear winner. The size of instruments, their distance from the listener, and their precise location are far more accurate with Ambisonics. Front/back separation is virtually perfect, and rear-channel performers don't sound as if they are "drilling" into your head. Ambience reproduction is more natural, and very close to what one hears "live". Four-mike recordings tend to sound too reverberant, and somewhat "spacey". The reverberation is not as coherently related to the main sound, and is a bit thin to the sides. (This isn't surprising, since we have already seen that Pair-Wise Mixing produces unstable side images that tend to "hug" the front or rear speaker.)

We are now in a position to draw some pretty firm conclusions. Only *three* channels of information are needed to unambiguously, precisely and stably locate the sound anywhere around the listener.

These signals can be processed by a very simple device for "display" over four or more speakers. They provide perfect mono and stereo compatibility. Four "discrete" channels, representing direct speaker feeds, waste space and give *worse* imaging. These considerations are *critical* to FM. A fourth channel would seriously degrade signal-to-noise ratio, increase adjacent-channel interference, and make FM reception much more susceptible to multipath interference. The Ambisonics/UMX/RCA system would not be so severely affected, nor put so heavy a premium on state-of-the-art receiver design.

Another conclusion can be drawn. If we are content with less-than-optimum Pair-Wise Mixing approaches, we then might as well adopt the Tate System SQ as the standard. Four "separate" channels would not give that much more separation to justify the tremendous investment required by stations.

Oh yes. What about that fourth track now left vacant on open-reel and Q8 tapes, CD-4 records, and 4-4-4 FM. Well, the Ambisonics people have developed the technology for it to carry up/down information. It just requires four more speakers on your ceiling.

Reprinted from *The QUAD Quarterly*, Vol.2, No.1, June 1980.
Copyright © 1980, 2000 by Laurence A. Clifton

Last updated: July 16, 2000