Proceedings of the Institute of Acoustics

MULTIDIMENSIONAL SOUND IN SOUND REINFORCEMENT FOR THEATRE

Robin Whittaker Out Board Electronics Ltd., Cambridge, England. www.outboard.co.uk

1. INTRODUCTION

Significant improvements in the delivery of amplified sound in a large listening area are possible using Source Oriented Reinforcement (SOR) techniques from a distributed rather than two channel loudspeaker system, resulting in a widening of the 'sweet spot' from typically 10% to 90% of the audiences listening positions. The 'sweet spot' is the area in which panoramic or spatial information in the mix is perceptible.

An appreciation of the interrelationships between electro-acoustics, architectural-acoustics and psycho-acoustics can help in preempting and solving problems encountered in specifying and setting up a sound system, and can also offer new techniques for achieving the best possible results.

Recent developments in digital signal processing have made SOR implementations a practical reality and success has been proved several times over at a number of top international musical, theatrical, operatic and orchestral productions.

2. WHAT IS SOURCE ORIENTED REINFORCEMENT

Source Oriented Reinforcement is the technique of amplification of a sound while localising sound back to its source. In a theatrical context this means back to the mouth of an actor on the stage or to the musical instrument making the original sound, or to the point on stage or in the auditorium from which a sound effect should emanate. Well implemented SOR sound designs result in most of the audience hearing correct localisation of the sound, so, for example, when the actor is stage left he is heard at stage left and his voice can be positionally differentiated from other actors at other stage locations. SOR is achieved by exploiting the psychoacoustic phenomena known as the precedence effect whereby the reinforced sound from a particular source is suitably delayed and leveled individually to each loudspeaker so as to fool the ear into not hearing the closest reinforcement loudspeaker to any listening position.

In an SOR design there are no overall system delays, each input has its own unique level and delay relationship to the speakers. In the simplistic example of a two channel proscenium left, right speaker system, when an actor is stage left, the delays on his / her microphone signal are set so that the delay to the proscenium left speaker is short and the delay to the proscenium right speaker is long. At the same time another actor is stage right, so the delays from the microphone need to be set with a short delay to proscenium right and a long delay to proscenium left. With this arrangement all the audience, even those seated directly in front of the proscenium speakers will hear different localisation's for the two actors. Taking this example further, as the actors move around the stage, the levels and delays to the loudspeakers are continuously altered for each voice so the sound image appears to follows the actor.

Expanding this simplified example to a real world case, a more typical sound reinforcement system may have speakers at proscenium left, center, and right as well as a number of front fills and under balcony delay speakers and perhaps some balcony fills as well, making a total of say 16 loudspeaker locations or groups. There may be a number inputs that need localisation treatment, so each of these feeds to the level delay matrix, whose outputs in turn drive the amplifiers and loudspeakers.

3. SPEAKERS AND SPACES

Consider the environment in which the sound reinforcement system is to operate and what sort of problems are likely to be encountered. Reflections of reinforced sound off architectural features may cause long disturbing echoes in audience areas, exciting reverberant spaces such as a roof dome or glass atrium will result in spectral changes and loss of intelligibility, an acoustically over damped environment will result in a very dead or lifeless sound in the reverberant field away from the direct field of the loudspeaker system. Often a concert hall or theatre needs to be considered as a number of different spaces, for example, the differing acoustic characteristics of the dress circle compared with the upper balcony, or of the central stalls area with a 20 meter domed roof as opposed to the stalls under-balcony area with a height to the ceiling of 3 meters.

In almost all applications, using a distributed sound reinforcement system will yield improvements over a large point source system because loudspeakers can be placed and targeted so as to largely avoid reflective surfaces and reverberant voids. More of the audience areas can be in direct field coverage, it is arguably easier to achieve a good sounding 100dBa at 5 meters distance than at both 5 and 50 meters distance from the source. The effect of unavoidable echoes can be minimised by filling in the gaps between first arrival and the echo. Smaller speakers are easier to rig and conceal, it is more practical to give specific treatment to specific areas of the space from a distributed system and to achieve better control of the distribution of SPL.

There are many different criteria for the selection of loudspeakers for any application, some of which are outlined below. Physical characteristics such as size and rigging features must be considered, while acoustic characteristics such as dispersion angle and coverage, horn type, efficiency, maximum spl, need careful consideration in conjunction with the choice of using acoustically coupled arrays or several individual distributed radiators.

3.1 VOCAL REINFORCEMENT

Front fills are usually mounted every few feet along front of stage, for coverage of first few rows of seats, and to hold auditory image down for balcony seating. Must be as small as possible with wide horizontal and narrow vertical dispersion with no on axis beaminess as these speakers are likely to be in the faces of front row audience.

Down fills - mounted at positions above the stage to cover main stalls seating area. Normally 8 - 10 meters high so around 15 meters distant from target. The dispersion angle would be selected subject to distance from target.

Under balcony - mounted on a drop off the balcony lighting rail or screwed to the underside of the balcony. Normally 2 - 4 meters from target with wide horizontal and narrow vertical dispersion. Should be small to avoid blocking sight lines.

A/B vocal systems - it is quite common to find two independent vocal reinforcement systems installed to prevent adjacent microphones on stage from phasing in the electronic mixing stage by maintaining two signal paths to two separate speakers.

3.2 BAND / ORCHESTRA / MUSIC REINFORCEMENT OR REPRODUCTION

Normally this is treated as a separate system because bandwidth and localisation requirements differ from those of the vocal and effects reinforcement systems. Speakers are normally mounted on the proscenium arch at low positions for stalls and high positions for balcony coverage.

3.3 SOUND EFFECTS

The auditorium surround system may require speakers to be mounted on the proscenium and side and rear walls of auditorium for surround sound effects and auditorium ambiance as well as for effect returns for reverberation. The boxes need to be wide angle full range and diffuse in character, a shallow front to back dimension is desirable. On stage effects are often delivered from very small loudspeakers mounted into set scenery, RF linked battery powered boxes are often used when mounted on moving scenery trucks. Where it is not possible to mount speakers into the set, scenery bouncers with a beamy narrow horn mounted on a down-stage lighting bar pointing at scenery to bounce sound back into the auditorium serve as a good alternative, as do floor bouncers mounted on an up-stage lighting bar pointing at the floor at an angle so the sound will bounce into the auditorium, and masked with drapes or a baffle so direct sound to the auditorium from the speaker is attenuated as much as possible. Up stage effects speakers for on stage ambiance are usually larger full range boxes mounted on lighting bars or behind and to the sides of the stage radiating directly into the auditorium.

3.4 SUB BASS

Required for band, orchestra and sound effects. Normally located either side of the stage floor mounted or under seating.

When reviewing the list, it is apparent that there is scope for integrating elements of these essentially separate sound systems to reduce redundancy, furthermore bringing the benefits of SOR into play will improve coverage, realism and intelligibility.

4. SENSES

The human auditory system has trained since birth to recognise patterns of electronic impulses from the vibration of hair cells caused by variations in sound pressure level. Torso and head shadowing as well as effects of the pinna and outer ear canal play a part in the cognition processes of localisation, intelligibility and voice recognition. The ear has a dynamic range of 120dB through the action of the cochlea amplifier extending the nominal 40dB dynamic range of the individual hair cells; the ear and can generate and suppress echoes; it has a level dependent frequency response characteristic and can be almost totally deaf to large instantaneous distortions while sensitive to the most insignificant steady state nuances in sound. This sophisticated and only partially understood organ is both more accurate and discerning than most measurement instrumentation while at the same time being highly subjective and open to influence by psycho-acoustic manipulation.

We are used to hearing sounds in the natural world from several different directions simultaneously and have learned to use differences in the level, spectrum and relative timing of sounds as they arrive at our ears to sense direction and distance. We have developed skills in echo management and have the ability to integrate several arrivals of sound together or ignore early reflections which plays an important part in auditory direction finding. We have the ability to suppress long echoes to minimise loss of intelligibility by a process of learning the character of a space and then ignoring or suppressing those spatial characteristics. These and other remarkable and as yet little understood facets of our hearing allow us to "focus" onto what is of interest to us and to filter out what is not. Particularly relevant to SOR is our ability, aptly named the Cocktail Party Effect, to discriminate between sounds with different localisations by cross correlating information received at two separate vantage points (our ears separated by our head).

If all of the sound that we hear from an amplification system comes from one direction we cannot use these auditory abilities and skills resulting in a reduction in intelligibility, reduced ability to differentiate between actors speaking, increased listener stress due to loss or correlation between visual and auditory perspective and from a theatrical point of view, an undermining of the state of suspension of disbelief that an audience must enter into for maximum enjoyment.

4.1 PRECEDENCE AND COMB FILTERING

Precedence describes how the auditory system gets much of its directional information from the first arrival of an audio wave front. No matter how loud a distant loudspeaker is reproducing a signal, if a small speaker in close proximity to the listener is also reproducing the same signal at very low level, it is localised. This is because the low level wave front from the small speaker arrives at the listeners ear before the high level wave front from the distant loudspeaker, i.e. it is precedent.

A listener hearing a centralised speech signal at a position mid way between two loudspeakers with a 60 degree separation will observe that introducing time delays incrementally from 1 - 10mS into the right hand speaker signal causes a hard image shift to the left speaker as well as tonal changes due to comb filtering each time the delay is increased. The comb filtering effect is due to the delayed and un-delayed signals combining and cancelling at the ear of the observer at frequencies harmonically related to the difference in arrival time.

Increasing the delay in 1mS steps from 10 - 30mS does not effect the image shift to the left, but no further comb filtering is heard because there is little correlation between the first and second arrival due to the magnitude of the time difference and the fundamental phasing frequency being below speech bandwidth.

Further increasing the delay from 30 - 50mS and beyond introduces an increasing echo from the right incrementally, causing disturbance to the listener and reducing intelligibility.

4.2 THE HAAS EFFECT AND BEYOND

In 1947 Helmut Haas submitted a paper entitled "The Influence of a Single Echo on the Audibility of Speech" as a dissertation towards his doctors degree from Gottingen University. Haas demonstrated the ability of the ear to integrate two arrivals of sound of equal level together and ignores the second arrival providing the time between arrivals is less than the threshold of echo perception (~25mS for speech). Haas comments on the shift in localisation to the un-delayed source almost in passing so it is somewhat ironic that the Haas Effect is the name given to the phenomenon of delay panning.

Other interesting research to students of Source Oriented Reinforcement was conducted by Blauert who researched directional sensitivity of the ear to different frequencies in the lateral and temporal planes. In addition Kutruff has shown that the ear can integrate several arrivals of sound together providing the times from the first arrival to second and the second to the third etc. are all within the threshold of echo perception. The threshold of echo perception is not only time related but also related to the relative levels and tonality of the two signals at the listening position and most importantly to the type of source material, for example, short impulsive or percussive sounds have a much shorter threshold of echo perception than long slow notes played on a string or wind instrument.

Research has shown that the human auditory system can pinpoint the position of an isolated audio event to an accuracy of 1-2 degrees if the event is in the frontal hemisphere, and provided the event is of a percussive nature. For steady state signals, especially high frequency sine waves, pinpointing the source is much less precise.

In order to spatially separate two signals such as speech, the sources need around 25 degree separation from the perspective of the listener, suggesting that the maximum number of sources that can be separately localised in the frontal hemisphere is seven. In addition the human auditory system is capable of localising two separate sources behind and one overhead.

In the natural world we are used to sounds arriving at our ears from all directions all the time, we learn to discriminate between primary arrivals and secondary arrivals and we learn to use reverberant secondary arrivals to aid intelligibility. We also learn to use our ability to sense direction to differentiate one sound from another and filter out background and foreground noise in order to focus on one sound source.

4.3 THE COCKTAIL PARTY EFFECT

The Cocktail Party Effect describes the ability to detect a sound even though it may be quieter than the background noise level. The trick depends on having two functional ears. The principle involved in cocktail party effect is that the significant item of information that might attract one's attention requires 12 dB greater SPL if the listener only has a single functioning ear. It follows that if all of the acoustic information is concentrated in a single direction, the acoustic energy field is more or less homogenous and we lose that ability to pick out a single element in the noise field. On the flip side, if multiple channels of sound from various directions are provided, a form of cocktail party effect goes to work for you. The result is a spatial spread of information that the auditory cortex can use to promote resolution of a clearer and more defined auditory ambiance. That includes the ability to hear dialogue when underscored sound or music is being used.

5. ABSOLUTE PANORAMA BEYOND STEREO

Most sounds in the natural world are monophonic, that is to say most sounds emanate from a single point in space and radiate spherically away from the center with energy dissipating following the inverse square law rule which dictates that every doubling of distance from the source results in a loss of 6dB level.

The human hearing system is binaural which gives us the ability to hear a mono sound from two separate vantage points and thus draw some conclusions about to position of the source of the sound from the cues discussed above. Part of the process of making these judgments is interaction of the listener with the environment by movement of the head resulting in changes in auditory perception of the environment.

We could then say that the human auditory system is not monophonic or stereophonic but omniphonic and can simultaneously determine the directions of multiple audio events in the x, y and z planes.

Delivering sound to an auditorium from a large stereo loudspeaker system has the disadvantage that for most listening positions it is not possible to hear any panoramic information in the mix because the sound from the closer loudspeaker will arrive at the listeners ear before the sound from the more distant loudspeaker. The law of Precedence dictates that even if the level of the later arrival of sound is significantly higher, the ear will perceive that the sound is coming from the closer loudspeaker, effectively masking one side of the PA.

5.1 A GROWING DEMAND

Surround sound and virtual reality, natural and unobtrusive reinforcement of acoustic instruments such as orchestral or operatic productions and enhanced reinforcement of rock or contemporary music shows fuels development of understanding, products and techniques. The growing utilisation of S.O.R. techniques result in a reduction of listener fatigue by breaking up the one dimensional wall of sound that is driven from a conventional loudspeaker system each side of the stage, by replacing it with a wide spatial panoramic sound field driven from a distributed sound system with more of the audience benefiting from an even spread of sound pressure over the entire listening area.

5.2 PRE-RECORDED STEREO AND MULTI-CHANNEL SOUND

For stereo music reproduction from a distributed loudspeaker system, by employing a delay matrix, it is possible for all loudspeakers to reproduce both channels of the stereo signal while maintaining the left and right perspective of the two signals for most listening positions.

Multi-channel film sound tracks are normally mixed in a studio that may bear some resemblance to the typical film theatre where the public go to watch a movie. In truth, it is near impossible to mix a multi-channel surround sound track for any place other than where it is being mixed. This is due to the several variances between one film theatre and another, let alone the differences between a recording studio seating 5 engineers and a film theatre seating an audience of anything between 200 - 1000 people.

Differences include the different sizes and acoustics of the rooms as well as differences in position, tonality, dispersion and directivity of the loudspeakers. It is also reasonably common practice to show movies in venues that are not film theater's at all, for example at out door events. In these cases there is obviously no correlation between intended typical replay environment and the actual.

S.O.R. techniques can be applied to improve the inequality dependent on seating position. By taking each of the 5 signals intended for the left, center, right, left surround and right surround loudspeakers and feeding them into a delay matrix, the signals can be post processed in such a way as to enable more speakers to be used and for all of the audience to get a much better delivery of spatial information.

All of the loudspeakers are delivering the sound from all 5 signal sources via 5 simultaneous time alignments with the result that the 'sweet spot' is enlarged to include most of the listening positions in the sound field. This technique can be adapted and used for any of the multi-channel recording standards currently in use.

6. SUMMARY

Source Oriented Reinforcement presents a superior solution for many sound reinforcement applications. It satisfies the multiplicity of requirements achieved by zoning and distribution of the loudspeaker system in terms of even distribution of SPL and tonality as well as minimisation of room effects. More seats benefit from improved direct to reverberant energy ratios and delivery of spatial information, aiding cognition and delivering realistic three dimensional soundscapes and impressive special effects.