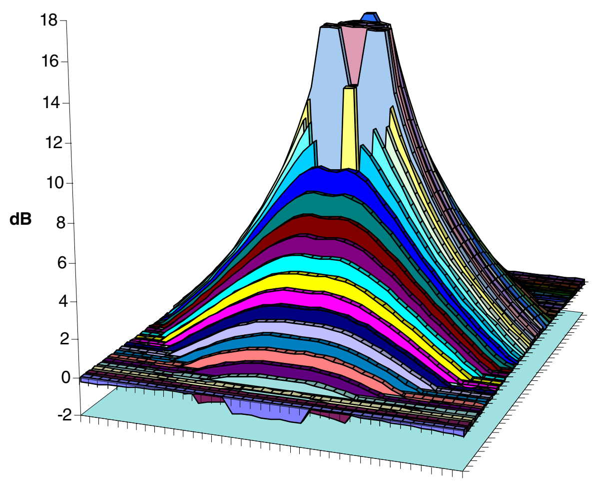
MacroFX™ algorithms

by Alastair Sibbald



In many PC games situations, it is desirable to create near-field proximity effects such that sounds seem to be very close, appearing to move from the loudspeakers up close to the head of the listener and, ultimately, even whispering into the ear. Consequently, a new MacroFX algorithm has been developed for Sensaura™ 3D-sound technology that provides the means of creating near-field effects. This has been achieved by careful modelling of the sound energy distribution in three dimensions around the head from all spatial positions (a single 2D section is shown left) and then transforming this data into a highly efficient algorithm. The algorithm is integrated directly with the Sensaura processing and controlled by DirectX™ so that it is transparent to applications writers, who now can create many new effects. For example, in an adventure game, a guide can whisper instructions into one of the listener's ears or, alternatively, in a flight-simulator, one could create the effect that the listener is a pilot hearing air-traffic information via headphones. In a combat game, it might be required to make bullets and missiles appear to fly close by the listener's head. These new effects are unique to Sensaura technology.

1 Synthesising sounds in three dimensions

A mono sound source can be processed digitally via a pair of Head-Related Transfer Functions (HRTFs) such that the resultant stereo-pair signal contains natural 3D-sound cues. These sound cues are introduced naturally by the head and ears when we listen to sounds in real life and they include the inter-aural amplitude difference (IAD), inter-aural time difference (ITD) and spectral shaping by the outer ear. When this stereo signal pair is introduced efficiently into the appropriate ears of the listener, by headphones say, then he or she perceives the original sound to be at a position in space in accordance with the spatial location of the particular HRTF pair which was used for the signal processing.

Sensaura™
3D POSITIONAL AUDIO

When one listens through loudspeakers, instead of headphones, then the signals are not conveyed efficiently into the ears because there is transaural acoustic crosstalk present that inhibits the 3D-sound cues.

This means that the left ear hears a little of what the right ear is hearing (after a small, additional time delay of around 0.2 ms) and vice versa. In order to prevent this happening, it is possible to create appropriate crosstalk cancellation signals from the opposite loudspeaker. These signals are equal in magnitude and inverted (opposite in phase) with respect to the crosstalk signals and designed to cancel them out. There are also some more advanced schemes which anticipate the secondary (and higher order) effects of the cancellation signals themselves contributing to further crosstalk and which make the appropriate correction.

2 3D sounds

When the HRTF processing and crosstalk cancellation are carried out correctly and when using high quality HRTF source data, the effects can be quite remarkable. For example, it is possible to move the image of a sound source around the listener in a complete horizontal circle, beginning in front, moving around the right-hand side of the listener, behind the listener and back around the left-hand side to the front again. It is also possible to make the sound source move in a vertical circle around the listener and, indeed, make the sound appear to come from any selected position in space. However, some particular positions are more difficult to synthesise than others, some for psychoacoustic reasons, we believe, and some for practical reasons.

For example, the effectiveness of sound sources moving directly upwards and downwards is greater at the sides of the listener (azimuth = 90°) than directly in front (azimuth = 0°). This is probably because there is more left-right difference information for the brain to work with. Similarly, it is sometimes difficult to differentiate between a sound source directly in front of the listener (azimuth = 0°) and a source directly behind the listener (azimuth = 180°). This is because there is no time-domain information present for the brain to operate with (ITD = 0) and the only other information available to the brain, spectral data, is similar in both of these positions. In practise, there is more HF energy perceived when the source is in front of the listener, because the high frequencies from frontal sources are reflected into the auditory canal from the rear wall of the *concha*, whereas from a rearward source, they cannot diffract around the *pinna* efficiently.

3 Measuring HRTFs

In practise, measurements can be made on an artificial head in order to derive a library of HRTF data such that 3D-sound effects can be synthesised. It is common practise to make these measurements at distances of one metre or thereabouts. This is done for several reasons, including the physical size of the sound source that must be used and for experimental convenience. In addition, for games, it is often required to create sound-effects positioned at distances of several metres or more and so, because there is little difference between HRTFs measured at one metre and those measured at much greater distances, the one metre measurement is used for practical convenience.

4 Distant sounds

The effect of a sound source appearing to be in the mid-distance (1 to 5 m, say) or far-distance (>5 m) can be created relatively easily by the addition of a reverberation signal to the primary signal. This simulates the effects of reflected sound waves from the floor and walls of the environment. A reduction of the high frequency (HF) components of the sound source can also help create the effect of a distant source, simulating the selective absorption of HF by air, although this is a more subtle effect. In summary, the effects of controlling the apparent distance of a sound source beyond several metres are well known and can be readily simulated.

5 Close sounds

In many games situations, it is desirable to create near-field distance effects such that sounds appear to be very close, appearing to move from loudspeaker distance, say, up close to the head of the listener and, ultimately, even whispering into one of the ears of the listener. For example, in an adventure game, it might be required for a guide to whisper instructions into one of the listener's ears. Alternatively, in a flight-simulator, it might be required to create the effect that the listener is a pilot hearing air-traffic information via headphones and in a combat game, it might be required to make bullets and missiles appear to fly close by the listener's head. These effects are not possible with HRTFs measured at 1 m distance.

6 Creating proximity effects

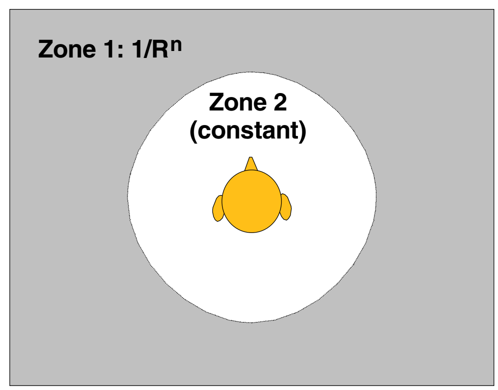
How might it be possible to create near-field proximity effects? In principle, it might be possible to make a full set of HRTF measurements at differing distances, say 1 m, 0.9 m, 0.8 m and so on, and switch between these different libraries for near-field effects. Such measurements are compromised by the physical dimensions of the associated loudspeaker diaphragm, which departs significantly from point-source properties at these closer distances. Also, considerably more memory space would be required to store each of the required additional HRTF libraries on the PC. A further problem would be that such an approach would result in quantised-distance effects: the sound source would not move smoothly to the listener's head, but would appear to jump when switching between the different HRTF sets. Ideally, what is required is a means of creating near-field proximity effects using a single, standard, 1 m HRTF set.

7 Sensaura MacroFX

A new algorithm has been developed for Sensaura technology, providing effective near-field proximity effects for 3D-sound synthesis using a single HRTF set. This has been achieved by careful modelling of the sound energy distribution in three dimensions around the head from all spatial positions into each ear and then transforming this data into a highly-efficient algorithm that is used in conjunction with the HRTF processing. The algorithm introduces the appropriate differential gain factors into the left and right channels after the HRTF processing stage (but prior to crosstalk cancellation), representative of a sound source at the chosen distance. The algorithm is integrated directly with the Sensaura processing and controlled by the DirectSound3D API so that it is transparent to applications writers: there are no additional rules to learn or other complications.

8 DirectX distance implementation

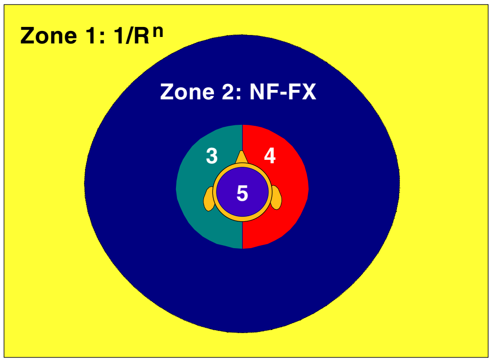
There are three operating zones defined by DirectX; the two proximal ones are shown below.



There is an outer, far-distant zone beyond which sounds are defined to have a constant intensity, but this is not relevant here. Inside this, there is a conventional outer zone, in which the sound intensity is distance dependent on the usual $1/R^n$ basis, and a circle of closest approach, within which the sound intensity remains constant. This is a way of preventing the intensity reaching extremely high levels when the distance is close to zero and thus constrains the required bandwidth of the system to practical levels. However, this also forces the distance of the virtual sound source to be always more than 1 m away, or thereabouts.

9 MacroFX implementation

Sensaura MacroFX provides a multi-zone system for 3D-sound synthesis and features five distinct proximal regions as shown below. (Six in all, including far-distance.)



In the present context, the expression near-field is defined to mean that volume of space around the listener's head within a distance of about 1 m from the centre of the head. For practical reasons, as described above, it is also useful to define a closeness limit and a distance of 0.2 m has been chosen here. These limits have both been chosen purely for descriptive purposes, based respectively on the typical HRTF measurement distance and the closest simulation distance one might wish to create in a game. It is also important to note that the listener hearing the sound *only* in a single ear represents the ultimate closeness, as would be the case if he or she was wearing a single earphone. This, too, can be simulated. The limiting case of near-field effects is to have the sound appear inside the listener's head, as occurs with conventional headphone sound images. This is the fifth of the five modes of Sensaura MacroFX operation, as listed below.

Zone 0: Far-Distance mode – constant intensity.

Zone 1: Distant mode – obeys 1/R^n law.

Zone 2: Near-Field FX algorithm – creates proportional effects close to the head.

Zone 3: In the ear (whisper mode): left ear.

Zone 4: In the ear (whisper mode): right ear.

Zone 5: Inside the head.

10 Effects

Sensaura MacroFX can be used to create a variety of special 3D-sound effects for interactive games, including the following.

1. Whispering in the ear.
2. Wind noise (especially when skiing, running, cycling, falling etc.).
3. Earphone and headphone simulation (for air-traffic, personal communications).
4. Ultra-close fly-by (for bullets, missiles, birds, bats, insects and so on).

Sensaura MacroFX can also be used to enhance conventional simple effects, such as underwater simulation.

For further information please contact:

Email: dev@sensaura.com
WWW: www.sensaura.com
Tel: +44 20 8848 6636

Sensaura Ltd
Dawley Road, Hayes
Middlesex UB3 1HH
United Kingdom

A Scipher Company

DEVPC/004/0001/1