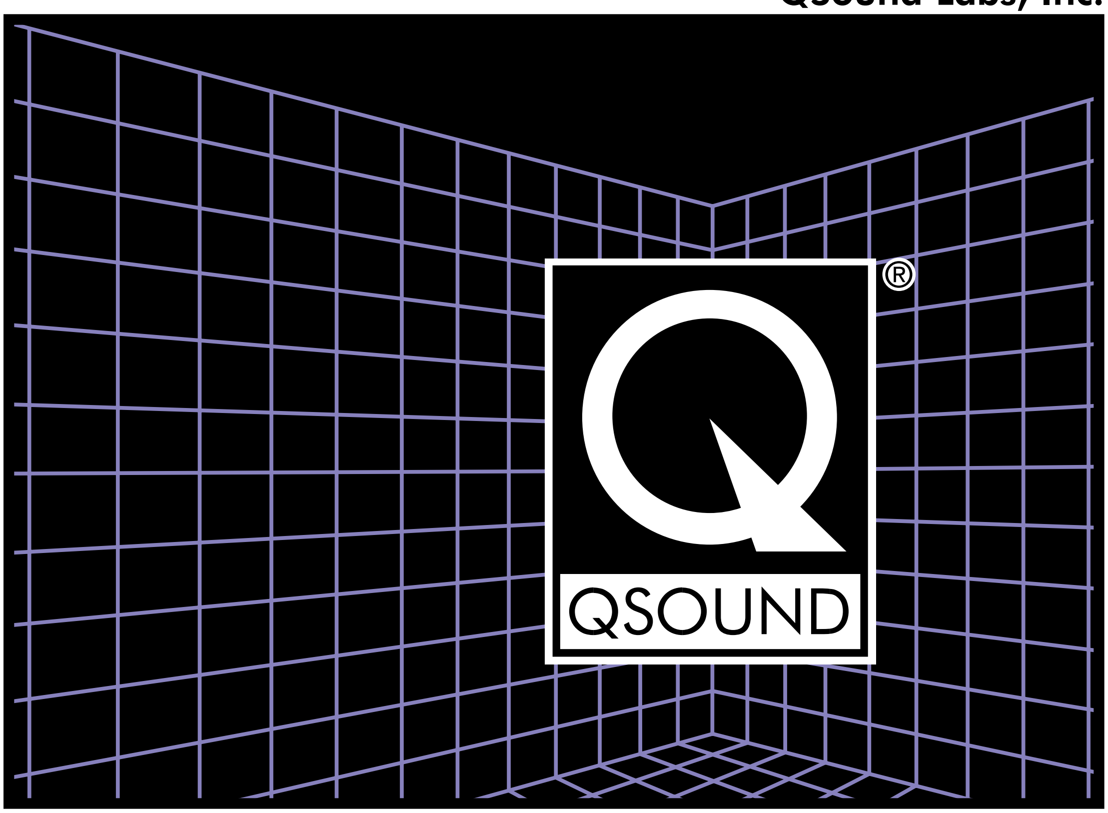
QSound® Q3D™
Positional 3D Audio

QSound Labs, Inc.



Rev 4.4 11 / 2008

Introduction

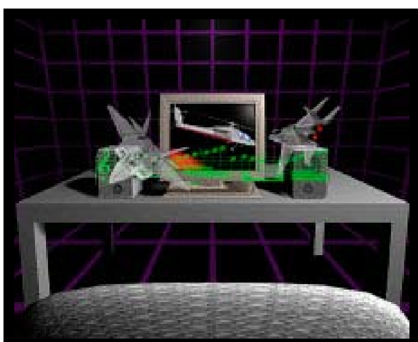
Abstract

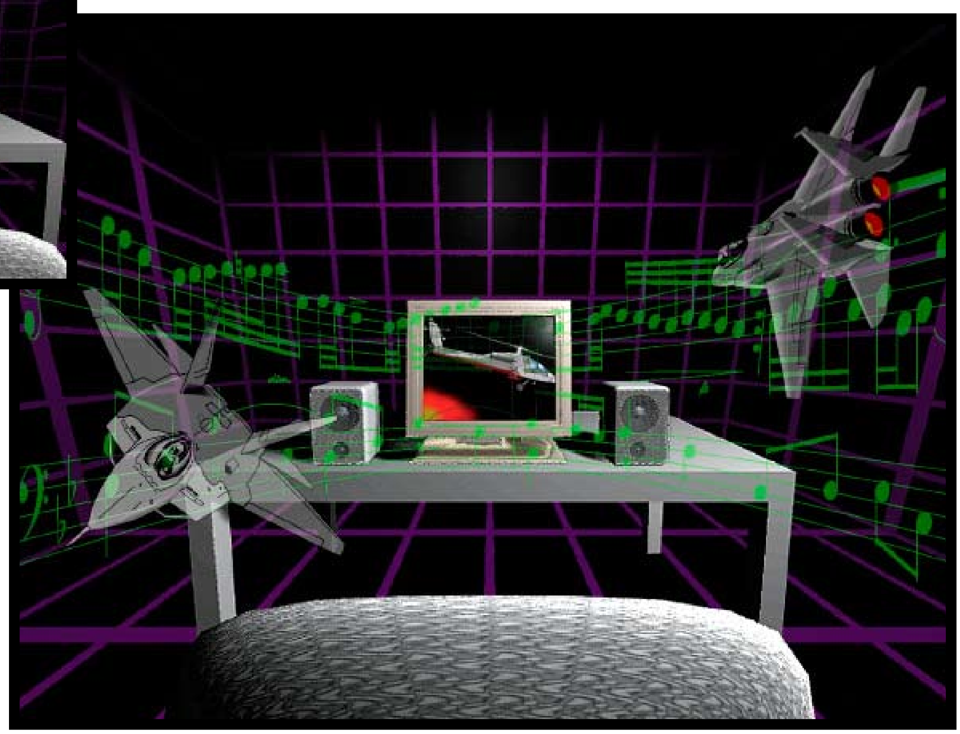
This document provides an overview of the Q3D three-dimensional positional audio process offered by QSound Labs, Inc.

Q3D dramatically extends the ability of standard sound generation and delivery systems to provide an immersive, life-like audio experience. In particular, Q3D enables real-time placement of multiple sound sources under software control, making it an ideal technology for virtual reality and video games.

Intended Audience

The document is suitable for all general reference purposes. It is, however, directed primarily to producers and manufacturers of software and hardware audio products.





Figures 1 & 2: Stereo vs QSound

A conceptual representation of the difference between plain stereo and QSound 3D stereo audio, showing how normal stereo imaging, (above) may be converted into an immersive soundfield, heightening the listener's impression of realism (right).

Development of Q3D

First let us consider a classic approach to 3D audio, which leads directly to a valid solution for the headphone output target. Then we will consider the unusual approach QSound Labs took to the provision of positional 3D audio for speakers

Traditional Binaural Hearing Analysis

Analysis of directional hearing in human beings can be undertaken by a number of methods. One approach is to use an artificial human head with microphone elements in place of the inner ears.

By placing real sound sources around the head, it is possible to measure, as a function of sound position relative to the listener, such characteristics as:

- Interaural Time Difference • This is the difference in the arrival times of a particular sound element at the two ears.
- Interaural Intensity Difference • This is the difference in amplitude measured at the two ears for a given sound and location.
- Head-Related Transfer Function • This is defined as the frequency-dependent effect on amplitude and delay primarily resulting from the interaction of sound waves with the outer ear (pinnae) and the shape of the head, and to a lesser extent the upper body.

Using such analytic methods it is possible to derive a filtering function that can be applied to an arbitrary monophonic sound signal, creating two output signals incorporating 'cues' mimicing the stimulus occurring at the ears if the sound was generated at a specific location in actual space.

This approach, properly termed binaural synthesis, and often loosely referred to simply as HRTF, is perfectly valid for headphone-targeted 3D audio, and indeed QSound Labs employs binaural synthesis and simplified derivatives thereof for headphone-targeted 3D.

However, binaural synthesis cannot be applied directly to speakers. The appropriateness of the signals is destroyed by crosstalk (leakage of left signal into right ear and vice versa). Further, unless the room is a totally sound-absorbing (anechoic) laboratory test environment, the signals will be further modified by the room characteristics.

Crosstalk Cancellation for Loudspeakers

Researchers such as Atal and Schroeder have formulated mathematical models that attempt to remove crosstalk. Although these models are correct in principle they have significant drawbacks in practical application. Cancelling crosstalk is computationally intensive and extremely sensitive to speaker/listener geometry. It tends to "color" the resultant signals audibly, making it unsuitable for professional applications such as music production.

It is for these reasons that QSound Labs took a fundamentally different approach to the development of algorithms for 3D audio effects over loudspeakers.

The QSound Approach to Speaker 3D

QSound inventors Danny Lowe and John Lees' original research was based on phenomena experienced in the recording studio as a result of experimentation with multiple microphones. Eventually they were able to achieve the impression of sound apparently originating from thin air, far beyond the normal limitations of the stereo sound stage.

After initial analysis of audio tape recordings, a series of experiments was conducted using a Hewlett Packard signal generator and an Eventide Clockworks Precision Delay. Though the latter was hardly a laboratory-grade instrument, the two men were able to demonstrate the possibilities well enough to raise interest and research capital.

- Armed with the results of this crucial early work, Lowe and Lees assembled a team of scientists and audio professionals.
- Software tools for sound analysis and modeling, test signal generation and statistical data analysis were custom developed in house, exploiting the newly available power of purpose-designed DSP (digital signal processor) IC's and desktop computers.
- QSound entered into an intense research phase, modeling sound wave propagation from loudspeakers, extrapolating promising theories for further exploration.
- Human listeners were presented with carefully designed test signals which had been passed through the initial process algorithms and delivered through stereo speakers in several controlled environments.
- Listener responses were captured using a custom data entry system consisting of multiple hand-held electronic keypads and a central data collector.
- As the system operators adjusted various characteristics of the test signals and process algorithms, analysis of the data was performed using a combination of software and skilled human judgement.
- As trends emerged, the processes were refined, retuned, retested and further refined in an iterative procedure taking place over many months. A key element in this task was the participation of team members having a background in the professional audio recording industry.
- In all, over 550,000 listening tests were performed, using a variety of subjects in order to produce averaged results leading to algorithms effective for virtually any listener.

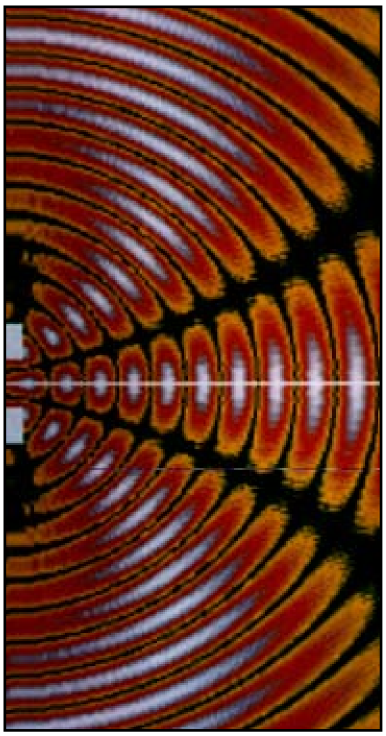


Figure 3
QSound modeling software generated this simulation of sound wave propogation from a stereo speaker pair. Circa 1989.

In short, rather than take a legitimate binaural model, and imperfectly and inefficiently adapt it to speakers through additional crosstalk cancellation, the QSound team concentrated directly on the target environment: the human percept of sound location as presented through stereo speakers.

The result was arguably the world's most natural-sounding, effective, and *inherently efficient* positional 3D audio algorithms for stereo speakers. In practical consumer applications such as desktop computers and mobile devices, the efficiency advantage over the combination of binaural synthesis with crosstalk cancellation should not be underestimated, as it has a direct impact on system compute and memory requirements, and therefore, product cost.

This initial algorithm development became the basis of the first commercial applications developed by QSound, the core of the QSystem hardware processor that was applied to the production of professional music releases and movie soundtracks, and became the foundation for what was to become the most complete suite of 3D algorithms available.

The Q3D 3D Mix Engine

Positional 3D Audio Fundamentals

Positional 3D audio processors are designed to accept multiple arbitrary monophonic audio streams and apply 3D localization filters, resulting in a combined stereo output incorporating positional cues in the form of time, phase and amplitude differentials between output channels. The result is that the listener perceives the individual sounds as originating from specific locations in space. (This principle can be extended to the multi-channel case, e.g. surround sound cards and speaker systems for personal computers.)

Conceptually, though not always literally, the basic signal path for each stream—mono in, stereo out—is repeated for the number of streams in play, and the output of individual stream processes is summed to produce the net output.

As the introductory material has indicated, significantly different algorithms are employed for speaker and headphone targets.

By adjusting the filtering parameters appropriately, the perceived location of a given sound stream—virtual sound source—may be controlled. In the majority of practical applications for positional audio, it is desirable to be able to make such adjustments in real time without audible artifacts. This makes possible the control of individual source positions in accordance with real-time user input, thus enabling highly realistic audio for virtual reality and 3D gaming applications.

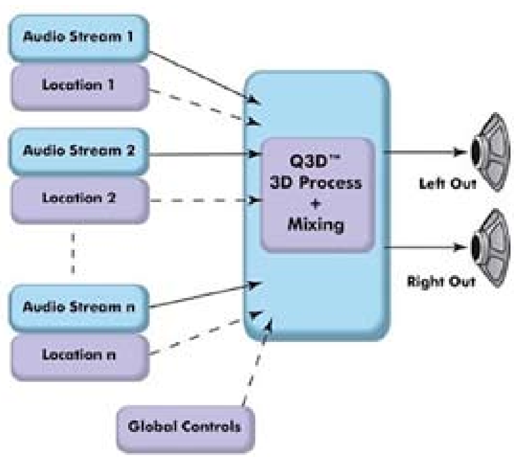


Figure 4: Multi-channel QSound Positional 3D Audio Mixer

Figure 4 is a highly simplified representation of a multi-channel positional mixer, emphasizing the independence of multiple input streams and control directives.

Positional 3D audio, like conventional "panning" of individual sound sources (i.e. left to right positioning by controlling relative signal level in each channel), is fundamentally part of the mixing process and is to be sharply contrasted with stereo enhancement / expansion algorithms (e.g. QSound Labs QXpander™) that act on already-mixed stereo streams, producing stereo output with increased apparent width of the sound stage.

Mixing and Playback Mechanics

Figure 4 shows a block at the heart of the Q3D process labeled "QSound Process + Mixing." The caption accurately describes this as a simple 3D mixer.

In practical applications, merely being able to impart a 3D position on an audio stream is a small component of a useful real-time audio engine. It must also play back and manage multiple audio streams in every other respect.

Under program control, sampled digital audio is loaded into memory buffers representing individual sound sources, then played, repeated, attenuated and otherwise manipulated on a per-stream (per virtual sound source) basis. The audio playback engine resembles a wavetable music synthesizer in many respects. The task of managing audio buffers is significant in itself, given that typical games may require many simultaneous audio sources. Keeping compute overhead to a minimum (i.e. achieving high efficiency) requires sophisticated and highly optimized real-time processing routines.

In addition to summing the outputs from all streams downstream of 3D positioning, up-stream sub-mixes of individual signals may also need to be produced to send to the reverberation module, and the output of the reverberation module added back to the master output.

Examples of Secondary Parametric Control

As with 3D position, secondary parameters must be independently applicable to multiple individual input streams.

Real-time control of the volume (relative signal level) of individual sounds is a basic necessity of the mixing function, for numerous reasons apart from the distance-related attenuation that is built into the positional process.

Perhaps the most obvious requirement is simply to control when a sound plays and when it is silent.

Less obvious is the need to control the rate of sample playback, and by so doing, affect the relative pitch of the sample.

For example, when a sound-generating object is in motion with respect to the listener, the perceived pitch of the sound is affected; this is the well-known 'Doppler shift' phenomenon. Pitch appears to increase as the object approaches the listener, and to drop as it moves away. To reproduce the sound of moving objects realistically, it is therefore necessary to control the rate of individual sound sample playback.

There are other reasons to control the rate of sample playback. For example, in video games, many sounds such as those of automobile engines are efficiently produced using a short sample which is continuously looped, i.e. played over and over again seamlessly. Controlling playback rate imparts the impression of such an engine running at different speeds.

Note that high-quality time-variant control of parameters such as volume and pitch, with low compute bandwidth and without significant audible artifacts, requires sophisticated interpolation algorithms and well optimized coding.

Simulating the Acoustic Environment

When a source produces sound in the real world, the physical characteristics of a given environment, whether mountain valley, concert hall or living room, determine the densities, strengths, delay and decay times associated with sound reflections that make their way back to the listener via a multitude of indirect paths. For example, in a small room, reflections of the original sound come so quickly as to blend together almost immediately into a wash of noise called reverberation, whereas in larger spaces they can take long enough to arrive that they will be distinguished as discrete echoes. Every real-world physical space, whether an open field, a stadium or a closet, has a unique reverberant characteristic.

Echo and reverberation effects are therefore a secondary but important aspect of realistic sound design. Q3D provides processing to create the impression of realistic virtual acoustic spaces within which individual sound sources can have an acoustic context. If the audio application using the positional engine is sufficiently sophisticated, it can select from a menu of preset Q3D reverberant acoustic environments to match, for example, various physical environments in a 3D game.

Selected History of Q3D

1990 to 2004

QSystem and QSystem II hardware 3D audio processors used to produce the first 3D-encoded music projects in the history of recording. The original QSystem provided ten independent static positional channels, while QSystem II provided four static and four dynamically-positionable channels. Sound position could be controlled in real time using a custom intelligent joystick controller, and the mix engineer's moves recorded by QSystem II for later playback.

An early form of real-time positional Q3D is adapted to DOS gaming in conjunction with normal mixing techniques.

QMixer™ and QMixer32 become the first fully positional, self-contained 3D software engines for multimedia software production, and amongst the first software applications to use the Intel® MMX® instruction set. QMixer thus predates Windows®95, DirectSound®, and the DirectSound3D API introduced by Microsoft years later to provide a standard interface for 3D audio applications on the Windows platform.

QSys/TDM™ plug-in for Digidesign Pro Tools™ and QTools/AX™ for DirectX-compatible editors bring positional audio capabilities to popular audio editing software.

QSoft3D™, a 3D mix engine for integration with sound card drivers, enables OEM's to add full 3D audio capabilities to existing stereo platforms, at overall CPU loads similar or better than hardware-implemented positional 3D.

QVE™, a modular WDM 3D audio mix engine with additional optional effects and audio control components, in conjunction with audio stream interceptor technology from QSound, enables a broad range of PC implementations from standalone self-contained software applications to the provision of advanced feature sets for sound cards and USB audio devices.

microQ™ brings Q3D to portable platforms, enabling hand-held 3D gaming.

For Further Information, Contact:
QSound Labs, Inc.

Head Office: 102, 2816-11 St. NE, Calgary, Alberta, Canada T2E 7S7
Telephone: (403) 291-2492 Facsimile: (403) 250-1521

Email oem.support@qsound.com
Web: http://www.qsound.com

While every effort is made to ensure the accuracy of the information presented in this document, QSound Labs cannot be held responsible for the dynamic nature of the marketplace, which may result in changes to product specifications and availability. Please note the revision date below and confirm the currency of all data with QSound Labs, Inc.

Rev 4.4 11 / 2008