A Low Complexity Spatial Localization System

4567 (J-3)

Robert C. Maher
EuPhonics, Inc.
Boulder, CO 80301-6115, USA

**Presented at
the 103rd Convention
1997 September 26–29
New York**

AN AUDIO ENGINEERING SOCIETY PREPRINT

A Low Complexity Spatial Localization System

Robert C. Maher

EuPhonics, Inc.
4840 Pearl East Circle, Suite 201E
Boulder, CO 80301-6115 USA
`http://www.euphonics.com`

Abstract

A low complexity procedure intended to approximate the directional cues of natural listening situations is described. The procedure accepts a monophonic (single channel) input signal and a set of parameters which specify the orientation and motion of a simulated sound source relative to the listener. The resulting stereo (two channel) output signal is appropriate for use with headphones or a cross talk compensated loudspeaker system. The approach is suited for use in personal computer systems with limited computational resources.

1. Introduction

Spatial localization is the method by which a listener perceives direction and distance of a sound source in three dimensions. A listener is able to localize a sound source primarily by interpretation of the differences in the received sound field at the two ears. Therefore, it is of practical interest to simulate electronically the acoustical effects that would occur for a sound source located at an arbitrary direction and distance from the intended listener.

1.1 Scope

The spatial localization procedure described in this paper is intended to approximate the directional cues of natural listening situations, but with very low computational complexity compared with existing solutions. At present, azimuth (relative angle of the source in the horizontal plane) and distance cues are the primary emphasis, while elevation information is supported with scaleable resolution in order to minimize computation. In situations where more than the minimum computational resources are available the algorithm is extendable to provide enhanced perceptual performance.

Among the important features of the algorithm is a procedure to derive the correct interchannel amplitude, frequency, and phase effects that would occur in the natural environment for a sound source moving with a particular trajectory and velocity relative to the listener. Although the correct effects can be obtained directly from measurements of the head-related transfer function (HRTF) from every point in space to each of the listener's ears, the resulting plethora of measurements is inconvenient or impractical for use in a real time system due to the large number of calculations required and the large amount of storage needed for all of the coefficients. Moreover, the problem of switching smoothly and imperceptibly from one set of HRTF coefficients to another in the case of moving sound sources is a serious impediment to this "brute-force" method.

In a case where the audio stream is localized using HRTF convolution the computation requirements can be very high. Assuming a 44.1kHz sample rate and a pair of 5 msec HRTF impulse responses, each audio sample requires about 440 multiply-accumulate operations to create a pair of left and right output samples. This corresponds to nearly 20 million multiply operations per second for one stream, or 160 million multiplies per second for 8 streams. Including processing for distance, Doppler, reverb, etc., quickly swells the computation requirements still further. When one realizes that typical hardware accelerators provide perhaps 50 million multiplies per second that must be shared simultaneously between music synthesis, sample rate conversion, audio effects, cross talk compensation, and other *concurrent* audio requirements, it becomes clear that the raw computation needs of direct HRTF processing exceed the available resources by an order of magnitude.

Therefore, in the algorithm described in this paper a *parametric* method is employed. The parameters provided to the localization algorithm describe explicitly the required directional changes (magnitude, phase, frequency shift, etc.) for the signals arriving at the listener's ears. Furthermore, the parameters are easily interpolated so that simulation of arbitrary movements from one source direction to another can be performed with much lower computational requirements than interpolation of the HRTF impulse responses themselves.

The localization algorithm currently uses one global on/off signal for loudspeaker cross talk compensation and approximately ten parameter sets per localized voice (eight fundamental parameter sets and potentially two additional parameters to support reverberation mixing and Doppler effect).

1.2 Background

Directional audio systems for simulating sound source localization are becoming increasingly commonplace and well known. Similarly, the principal mechanisms for sound source localization by human listeners have been studied systematically since the early 1930's. The essential aspects of source localization consist of the following features, or *cues*:

- **Interaural time difference (ITD)** -- the difference in arrival times of a sound at the two ears of the listener, primarily due to the path length difference (if any) between the sound source and each of the ears.

- **Interaural intensity difference (IID)** -- the difference in sound intensity level at the two ears of the listener, primarily due to the shadowing effect of the listener's head.

- **Head diffraction** -- the wave behavior of sound propagating toward the listener involves diffraction effects in which the wavefront bends around the listener's head, causing various frequency-dependent interference effects.

- **Effects of pinnae** -- the external ear flap (pinna) of each ear produces high frequency diffraction and interference effects that depend upon both azimuth and elevation of the sound source.

The combined effects of ITD, IID, head diffraction, and the pinnae is contained implicitly in the head related transfer functions for each ear at each combination of azimuth and elevation angles. The details of the HRTF vary from person to person, so creating a suitable localization impression for a wide range of listeners requires considerable ingenuity.

Other familiar cues that are present in normal listening surroundings include:

- Discrete reflections from nearby surfaces
- Reverberation
- Doppler and other time-variant effects due to relative motion between source and listener
- Listener experience with the characteristics of common sound sources

A large number of studio techniques have been developed in order to provide listeners with the impression of spatially distributed sound sources [1]. Most of the techniques involve provision for short time delays, gain panning, echo generation, artificial reverberation, and so forth, which simulate, via loudspeakers or headphones, the cues that would occur in a natural listening environment with real sound sources.

An interesting example of sound source spatial simulation for loudspeaker presentation was reported by John Chowning in the early 1970's [2]. Chowning's method included provision for the simulation of directional intensity, Doppler frequency shift, and direct-to-reverberant ratio as a function of distance. Chowning chose not to use ITD or other head-related cues since the intended application was four-channel loudspeaker playback in which the precise location of the listener(s) was not known.

Additional work has been performed in the area of *binaural* recording. Binaural methods involve recording a pair of signals that represent as closely as possible the acoustical signals that would be present at the ears of a real listener. This goal is often accomplished in practice by placing small microphones at the ear positions of a mannequin head, with the mannequin then located at the desired "listening" position. Thus, the naturally occurring time delays, diffraction effects, etc., are generated acoustically during the recording process. During playback the recorded signals are delivered individually to the listener's ears, typically via isolated headphones. In this way any directional information present in the original recording environment is available for the listener during playback.

The basic principles of binaural (e.g., a dummy head recording with headphone playback) and stereophonic (e.g., a multimicrophone recording with loudspeaker playback) reproduction have been practiced for many years. The fundamental concepts were summarized by Snow in 1953 [3]. Snow stated that:

> It has been aptly said that the binaural system transports the listener to the original scene, whereas the stereophonic system transports the sound source to the listener's room.

It is also of interest to be able to produce a binaural listening experience with loudspeakers rather than headphones. The major difference between headphone and loudspeaker playback is leakage, or *cross talk,* from the left speaker into the right ear and from the right speaker into the left ear when loudspeakers are used, while little if any leakage occurs for headphones. Eliminating or reducing loudspeaker cross talk is referred to as *cross talk compensation*, or CTC.

The initial work on CTC was reported by Schroeder and Atal of Bell Laboratories in the early 1960's [4]. In their original paper Schroeder and Atal describe the cross talk problem of loudspeaker reproduction and propose a solution involving preprocessing of the left and right loudspeaker signals. Other important work in CTC has been described by Damaske [5], and Cooper and Bauck [6, 7].

A refinement of the binaural recording method is to simulate the head-related effects by convolving the desired source signal with a pair of measured (or estimated) head-related transfer functions. Laboratory demonstrations during the 1980's showed this process to be a useful method of simulating directional information [8]. More recently, the term *auralization* has been used to describe binaural simulation systems in which the head-related effects are combined with the desired effects of the simulated listening environment (reflections, reverberation, multiple sound sources, etc.). A description of the features and goals of auralization systems is available in the paper by Kleiner, et al. [9].

In summary, the extensive prior art in stereo spatial sound localization simulation involves the creation of signal processing models which generate acoustical cues that are reminiscent of the effects familiar to human listeners in natural surroundings.

1.3 Outline of Paper

The remaining sections of this paper are organized as follows. First, an overview of the proposed spatial localization procedure is given, followed by a summary of each signal processing block. Next, the major considerations for a low-complexity implementation of the algorithm are discussed. Finally, a concluding section with suggestions for further work is provided.

2. Description

The proposed spatial localization procedure provides a set of audible modifications which produce the impression that a sound source is located at a particular azimuth, elevation, and distance relative to the listener. In a typical implementation the input signal to the process is a single channel (monophonic) recording of the desired sound source. The output from the localization process is a two channel (stereophonic) pair of signals that are presented to the listener from a pair of conventional hi-fi loudspeakers or optionally via stereo headphones. If loudspeakers are used the process includes a cross talk canceling network to reduce signal leakage from the left loudspeaker into the right ear and from the right loudspeaker into the left ear.

The conceptual description of the proposed low-computation auditory spatial localization procedure is depicted in Figure 1. For each *voice* (monophonic audio signal) to be localized a sample rate converter is provided. The sample rate converter performs frequency shifting in order to simulate the Doppler effect due to the radial velocity of a moving source relative to the listener. Next, the frequency shifted signal is passed to a parametric equalizer filter (EQ) stage. The EQ block produces two (left and right) filtered output signals, L_A and R_A, which have been spectrally adjusted to correspond to the head-related effects for the desired source direction and distance. The filtered signals are then passed through a variable differential delay which simulates the correct interaural time difference (ITD) for the desired source direction. The delayed signals (L_D and R_D) are passed through a signal splitter which sums a portion of the left and right signals to FRONT left and right accumulators and BACK left and right accumulators. The FRONT and BACK accumulators are shared by all the voices to be localized by the system. The BACK accumulator signals are passed through a decorrelation stage to simulate the diffuse quality of sound sources when located behind the listener's head. Finally, the left and right processed signals are mixed with artificially reverberated versions of the source signals to simulate the desired acoustical properties of the listening environment. If headphones are to be used for playback the composite left and right signals are passed through optional headphone EQ filters and on to a conventional headphone amplifier. If loudspeaker playback is desired, the composite left and right signals are passed through a cross talk compensation (CTC) system before being sent to a conventional stereo power amplifier.

Each of the above mentioned processes are described in detail next.

2.1 Spatial Localization Modules

Among the important features of the proposed system is the use of a modular framework. The processing modules are intended to be self-contained hardware or segments of DSP code that may already be available "off the shelf". It is also expected that the modules can be re-used by several algorithms sharing the same platform (e.g., a music synthesizer or effects processor).

2.1.1 Rate Converter

The rate converter module performs frequency shifting on the input voice. The frequency shift is proportional to the simulated radial velocity of the source relative to the listener (Doppler effect). The fractional sample rate factor by which the frequency changes is given by the expression $1-\frac{\vec{v}_{radial}}{c}$, where $\vec{v}_{radial}$ is the radial velocity (positive away from the listener, negative toward the listener), and c is the speed of sound (~343 m/sec in air at room temperature). The rate converter function is accomplished using a fractional phase accumulator to which the sample rate factor is added for each sample. The resulting phase index is the location of the next output sample in the input data stream. If the phase accumulator contains a noninteger value, the output sample is generated by interpolating the input data stream. This process is analogous to a wavetable synthesizer with fractional addressing.

2.1.2 Equalizer and Gain

As mentioned above, the acoustical signal from a sound source arrives at the listener's ears modified by the acoustical effects of the listener's head, body, ear pinnae, and so forth, comprising the HRTF. In the proposed system the HRTF frequency responses are approximated using a low order parametric filter. The control parameters of the filter (cutoff frequencies, low and high frequency gains, resonances, etc.) are derived once in advance from actual HRTF measurements using an iterative procedure which minimizes the discrepancy between the actual HRTF and the low order approximation for each desired azimuth and elevation. This low order modeling process is helpful in situations where the available computational resources are limited.

In a minimal formulation of this procedure the HRTF approximation filter for each ear could be a first order shelving equalizer of the Regalia and Mitra type [10]. This structure is shown in Figure 2. In this minimal example the equalization and gain process is controlled by three parameters for each channel (left and right): G, K, and b. The G parameter controls the overall gain. K controls the gain at high frequencies, while b selects the bandwidth of the low frequency portion of the response. Setting $K = 1$ causes the equalizer to have a constant (flat) gain for all frequencies, while $K>1$ and $K<1$ correspond to high frequency boost and cut, respectively. Thus, in this simple case the HRTFs for each combination of azimuth and elevation are simulated by choosing the G, K, and b parameters that cause the response of the equalizer filter to match the general shape of the HRTF frequency response magnitude. Note that the parameters can also be used to model the frequency response changes that occur due to alterations in source-to-receiver distance.

2.1.3 Left and Right Delays

Simulation of the interaural time difference between sound arrivals at the left and right ear can be accomplished with a pair of interpolated delay lines. The maximum interaural delay of approximately 700 microseconds occurs for azimuths of 90° and 270°. This corresponds to less than 32 samples at a 44.1kHz sample rate. Note that the delay needs to be applied to the "far" ear signal channel only.

In a minimal implementation, the interaural delay can be calculated using the Woodworth empirical formula [11]:

$$\text{DELAY} = 257\text{microseconds}*(\Theta + \sin(\Theta)) \qquad (1)$$

where DELAY is for the far ear relative to the near ear, and Θ is the azimuth angle in radians. If the required delay is not an integer number of samples the delay line can be interpolated to estimate the value of the signal between the explicit sample points.

2.1.4 Decorrelation and Front and Back Factors

Creating a convincing "behind the head" source location impression is difficult, particularly if a pair of loudspeakers are used for playback. In the proposed system the front and back hemispheres are processed differently to overcome the ambiguity often found in localization systems. Specifically, sound sources to be simulated behind the head are passed through a decorrelation system which provides a diffuse sound quality that is typical of sources that are out of the listener's field of view. The decorrelation can be accomplished with a cascade of all-pass filters, a simple reverberation scheme, or some other method [12].

In order to avoid an audible discontinuity when a sound source passes from front to back or from back to front, the proportion of the signal sent through the decorrelation process is gradually increased or decreased smoothly as a function of azimuth and elevation.

Note that in order to reduce the computational complexity of the localization procedure the decorrelation system is shared by all the voices to be localized. That is, the output of each localized voice (shaded portion of Figure 1) consists of a left and right "front" portion ($L_{F,n}$, $R_{F,n}$) and a left and right "back" portion ($L_{B,n}$, $R_{B,n}$), which are then additively combined with the corresponding outputs of the other voices.

2.1.5 Reverberation

A listener gains a strong impression of the acoustical surroundings by means of echoes and reverberation. These cues provide information about the size, geometry, and materials making up the listening space. In the proposed localization procedure the important reverberation cues are provided by a shared reverberation system which receives the unprocessed monophonic signals from each voice. To a first approximation reverberant effects are non-directional, so sharing a single reverberation system for all voices gives a suitable result with minimum computation.

The direct-to-reverberant energy ratio gives the user a strong impression of distance: sources close to the listener have direct (non-reverberant) energy greater than the reverberant energy, while sources located at greater distance are dominated by the reverberant sound field. This effect can be simulated by controlling the gain factor (G) in each voice, since this factor controls the direct sound without altering the reverberation level.

2.1.6 Output Equalization

Once the entire simulated sound field is created as two output channels (L_T and R_T), the user may select to listen via headphones or loudspeakers. In either case it may be desirable to compensate for the frequency response characteristics of the playback system or for the personal preferences of the listener. If headphones are used, the output equalization may be chosen as the inverse of the headphone response so the signals arriving at the listener's ears are spectrally uncolored [13, 14]. If loudspeakers are used, the effects of signal leakage (cross talk) from each loudspeaker to the opposite-side ear of the listener must be reduced or eliminated, as described next.

2.1.7 Cross Talk Compensation

The primary problem with loudspeaker reproduction of directional audio effects is cross talk between the loudspeakers and the listener's ears. Delivering only the left audio channel to the left ear and only the right audio channel to the right ear either requires the use of headphones or the inclusion of a cross talk compensation (CTC) system prior to the loudspeakers.

The main principle of CTC is to generate signals in the audio stream that will acoustically cancel the cross talk components at the position of the listener's ears. Because the cancellation occurs in the acoustic field itself, the effectiveness of the CTC system depends largely on the degree to which the transfer functions from each speaker to each ear of the listener can be determined.

The mathematical development of the seminal Schroeder-Atal CTC system is as follows [15]. Referring to the configuration of Figure 3, the left and right channel unprocessed spectral domain signals, $L(\omega)$ and $R(\omega)$, are sent through the CTC network to generate $L_p(\omega)$ and $R_p(\omega)$, the left and right loudspeaker spectral domain signals. Denoting the transfer function from a speaker to the same-side ear as $S(\omega)$ and to the opposite ear (cross talk) as $A(\omega)$, the total acoustic spectral domain signal at each ear is given by

$$\begin{aligned} L_E(\omega) &= S(\omega)\cdot L_P(\omega) + A(\omega)\cdot R_P(\omega) \\ R_E(\omega) &= S(\omega)\cdot R_P(\omega) + A(\omega)\cdot L_P(\omega) \end{aligned} . \tag{2}$$

Note that the transfer functions $A(\omega)$ and $S(\omega)$ are essentially the HRTFs corresponding to the particular azimuth, elevation, and distance of the loudspeakers relative to the listener's ear locations. These transfer functions take into account the diffraction of the sound around the listener's head and body, as well as any spectral properties of the loudspeakers.

The desired result is to have $L_E = L$ and $R_E = R$ so that the listener receives the L and R channels without cross talk. This can be accomplished by properly generating L_P and R_P from L and R using the CTC network. Substituting $L_E = L$ and $R_E = R$ into the previous equation, dropping the explicit ω notation for simplicity, and collecting terms gives

$$\boxed{\begin{aligned} L_P &= \left(\frac{S}{S^2 - A^2}\right)\cdot L - \left(\frac{A}{S^2 - A^2}\right)\cdot R = \left(\frac{1}{S}\right)\cdot\left(\frac{1}{1-(A/S)^2}\right)\cdot\left(L - \frac{A}{S}R\right) \\ R_P &= \left(\frac{S}{S^2 - A^2}\right)\cdot R - \left(\frac{A}{S^2 - A^2}\right)\cdot L = \left(\frac{1}{S}\right)\cdot\left(\frac{1}{1-(A/S)^2}\right)\cdot\left(R - \frac{A}{S}L\right) \end{aligned}} . \tag{3}$$

The resulting CTC network is shown in Figure 4. Thus, if the ipsi- and contra-lateral HRTFs (approximately S and A) are known, at least theoretically the required filters in Figure 4 can be computed numerically.

In order to realize the network depicted in Figure 4 it is necessary that the denominator factor $1-(A/S)^2$ be a causal function so that the recursive filter $1/(1-(A/S)^2)$ can be implemented. This requirement is met if A/S is causal, which is at least possible since A and S are causal and the delay of A is presumably always greater than the delay of S. Schroeder and other authors state that A/S is typically causal, but it is necessary either to verify this in practice, or to make suitable approximations [6]. Other practical problems with determining A/S are that the z-transform representation of S may contain zeros near the unit circle which cause large peaks in the inverse filter, or S may contain zeros outside the unit circle which cause a two-sided (non-causal) inverse function.

In practice it is necessary to reduce the computational complexity of the CTC system shown in Figure 4 by simplifying the filter blocks. Appropriate simplifications include the use of low-order approximations, simple delay lines, and so forth.

3. Discussion

3.1 Implementation Considerations

The localization process involves several steps for each active voice. At the decorrelation point in Figure 1, the left and right streams from all active voices are mixed into four (or six) accumulators: left front, right front, left back, and right back (and left reverb and right reverb, if available). The left and right *back* streams are decorrelated and summed with the corresponding left and right *front* and *reverb* streams, creating left and right processed signals. The processed signals are sent to the left and right headphones of the user, or passed through a cross talk compensation network if a pair of loudspeakers are to be used for playback.

In a typical multimedia personal computer the localization process is implemented using a hardware accelerator of some type. Simultaneous processing of 4 to 8 localized audio streams is a common requirement in the context of computer games. The control information for the localization process is calculated by the game program and passed to the accelerator for processing. In addition to the localization process the hardware accelerator is simultaneously called upon to perform music synthesis, audio effects, and other audio-related processes. Thus, it is important to view the computational cost of the localization process in terms of available accelerator resources for concurrent processing.

3.2 Evaluation of the Parametric Approach

A minimal implementation of the parametric localization approach described in this paper requires a total of approximately 10 million instructions per second (MIPS) to localize 8 streams at a 44.1kHz sample rate (including shared decorrelation and loudspeaker CTC processing, but not including the reverberation computation). This level of computation has been found to be suitable for typical hardware accelerators supporting concurrent wavetable music synthesis for personal computer multimedia use.

A minimal implementation provides surprisingly good azimuth and distance results for sound effects common in computer games (gun shots, explosions, engine noises, etc.). Elevation cues are objectively less dependable with such a low level of computation, but listeners report the subjective impression of strong elevation effects when presented with sounds for which they are predisposed to assume elevation (jet flyovers, rocket launches, etc.). Increasing realism is obtained if additional MIPS are available.

3.3 Importance of scalability

A significant advantage of the proposed parametric approach in the context of concurrent audio processing (music synthesis, effects, and localization simultaneously) is its ability to scale smoothly to a lower level of computation if the resources of the hardware accelerator are needed for higher priority tasks. This is accomplished by assigning a perceptual priority to each of the processing blocks so that features of less perceptual importance can be deleted while the most important cues are retained. For example, a simulated sound source moving rapidly past the listener from left to right provides gain, delay, filter, and Doppler cues in the left and right output channels. If the available computational resources are limited, retaining the gain and delay elements but bypassing the filter and Doppler elements provides a largely acceptable localization impression with substantially reduced computation.

4. Conclusion

In this paper a parametric spatial localization system has been described that is suitable for use in multimedia personal computer systems. The system accepts a monophonic input signal to be localized, and a set of parameters which are used to control the process. This parametric approach incorporates the sharing of resources when multiple signals are processed simultaneously, and supports efficient algorithm scalability to allow the required computation to be increased or decreased according to system loading.

Additional work is needed to develop automated tools for the determination of the required localization parameters based on a set of HRTF measurements. Furthermore, it would be desirable to provide a mechanism for "fitting" the parameters to a particular playback system and listener. Development of these and other features is currently underway.

5. References

[1] Eargle, J. M., *Handbook of Recording Engineering,* New York: Van Nostrand Reinhold, 1986.

[2] Chowning, J., "The Simulation of Moving Sound Sources," *J. Audio Eng. Soc.,* vol. 19, no. 1, pp. 2-6, 1971.

[3] Snow, W. B., "Basic Principles of Stereophonic Sound," *J. SMPTE,* vol. 61, pp. 567-589, 1953.

[4] Schroeder, M. R., and Atal, B. S., "Computer Simulation of Sound Transmission in Rooms," *1963 IEEE International Conv. Record,* Part 7, pp. 150-155, March 25-28, 1963.

[5] Damaske, P., "Head-Related Two-Channel Stereophony with Loudspeaker Reproduction," *J. Acoust. Soc. Am.,* vol. 50, pp. 1109-1115, 1971.

[6] Cooper, D. H., and Bauck, J. L., "Prospects for Transaural Recording," *J. Audio Eng. Soc.,* vol. 37, pp. 3-19, 1989.

[7] Bauck, J. L., and Cooper, D. H., "Generalized Transaural Stereo and Applications," *J. Audio Eng. Soc.,* vol. 44, pp. 683-705, 1996.

[8] Wenzel, E. M., Foster, S. H., and Wightman, F. L., "Realtime Digital Synthesis of Localized Auditory Cues over Headphones," *Proc. ICASSP,* 1989.

[9] Kleiner, M., Dalenbäck, B., and Svensson, P., "Auralization--An Overview," *J. Audio Eng. Soc.,* vol. 41, pp. 861-875, 1993.

[10] Regalia, P. A., and Mitra, S. K., "Tunable Digital Frequency Response Equalization Filters," *IEEE Trans. ASSP,* vol. ASSP-35, pp. 118-120, 1987.

[11] Woodworth, R. S., *Experimental Psychology,* New York: Holt, 1938.

[12] Kendall, G. S., "The Decorrelation of Audio Signals and Its Impact on Spatial Imagery," *Computer Music J.,* vol. 19:4, pp. 71-87, 1995.

[13] Bauer, B. B., "Stereophonic Earphones and Binaural Loudspeakers," *J. Audio Eng. Soc.,* vol. 9, pp. 148-151, 1961.

[14] Begault, D. R., *3-D Sound for Virtual Reality and Multimedia,* Cambridge, MA: Academic Press Professional, 1994.

[15] Schroeder, M. R., "Computer Models for Concert Hall Acoustics," *Am. J. Physics,* vol. 41, pp. 461-470, 1973.

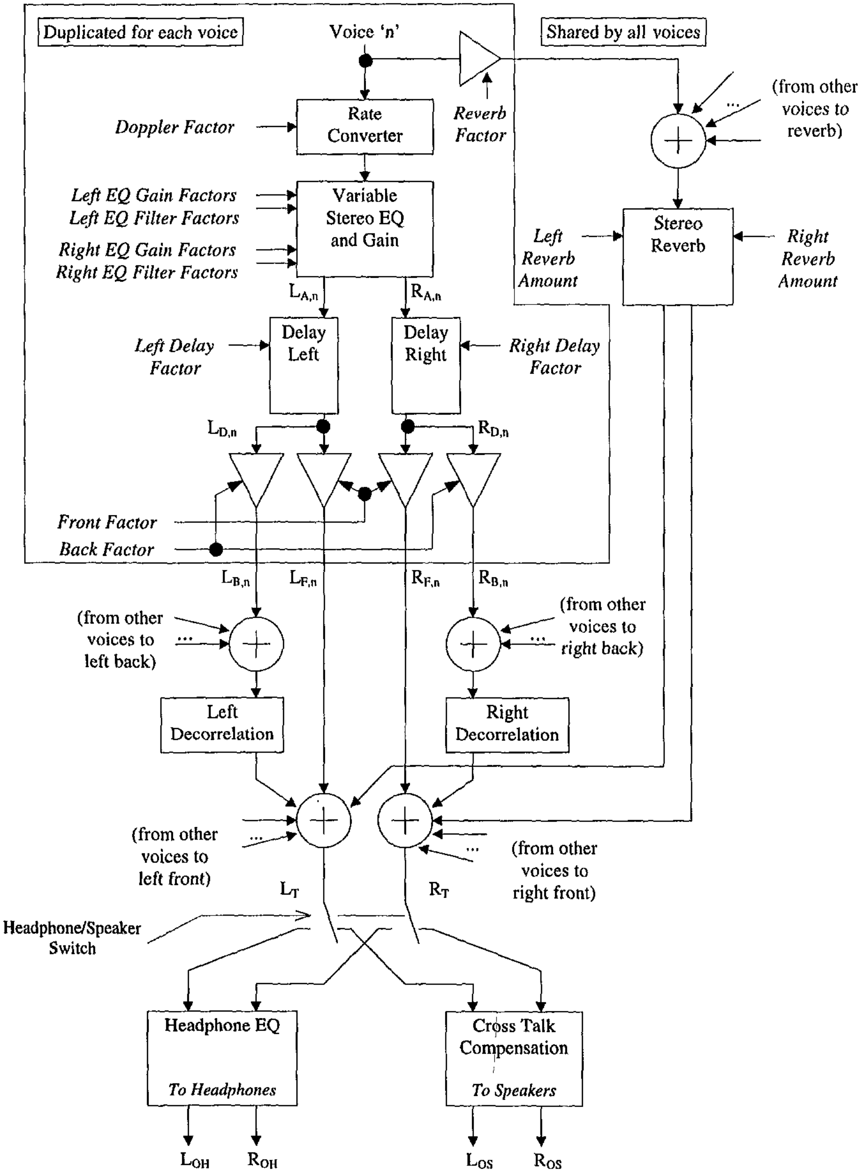

Figure 1: Proposed low-computation spatial localization framework

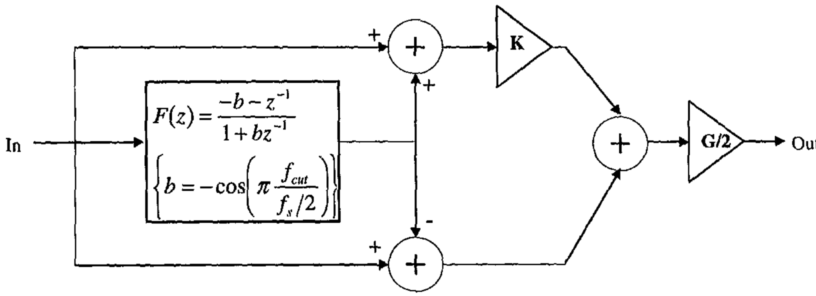

Figure 2: Example of a minimal spectral shaping implementation

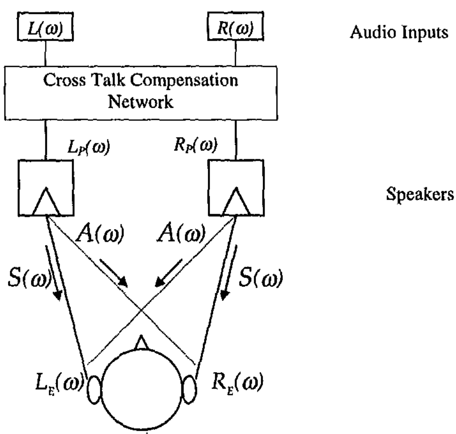

Figure 3: Schroeder-Atal cross talk compensation (CTC) concept

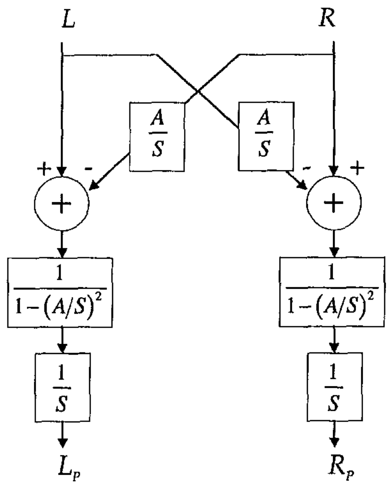

Figure 4: CTC block diagram