VIRTUAL SURROUND SOUND IMPLEMENTATION USING DECCORRELATION FILTERS AND HRTF

Jakes Bejoy

Center for computer research in music and acoustics

Stanford University

ABSTRACT

Carma space project is focused on exploring the various spatial techniques available today and this paper is an exploration into virtual surround sound technology making use of decorrelation filters to provide an impact on the spatial imagery. Virtual Surround Sound system makes it possible to reproduce multi channel recording over an ordinary pair of stereo headphones. In the field of spatial hearing, signal decorrelation is known to have a dramatic impact on the perception of sound imagery.Decorrelating the sound input signal gives a feeling of diffuseness and also brings about a great amount externalization in a headphone reproduction.

1. INTRODUCTION

Many may wonder how it was possible to cram five speakers and a sub-woofer into a set of two-speaker headphones, while still enjoying the same immersive experience generated by multi speaker playback system. This is possible through the virtual surround sound technology. This may sound a bit contradictory (multi channel play back through headphones) but considering that we only have two ears to capture the auditory information irrespective of the different sound sources gives a good justification. Decorrelation is a process by which a single source signal is transformed in such a way that they sound the same but their waveforms are altered.Decorrelation occurs in sound synthesis when there are slight differences between the sound synthesized for output channels. In this case a 5.1 virtual field is generated using KEMAR HRTF's. Only the left surround and the right surround signals are decorrelated so as to produce a diffused field (similar to the late field of reverberant concert hall)

2. WHY DECORRELATION

The degree to which sounds are decorrelated has proven to be a significant predictor of spatial effects. Even in natural environment decorrelation occurs as a by-product of acoustic or electronic processes that often change the sound of the source. Even in recording studios, vocals are sometimes recorded twice on separate tracks so that the micro variations in recording produces decorrelation. Decorrelation has atleast five effects on the perception of spatial imagery [1],

1.Coloring is eliminated
2.They can produce diffuse sound fields.
3.Produce enough externalization in headphone playback.
4.Image shift can be eliminated.
5.Precedence effect can be defeated.

3. IMPLEMENTING DECORRELATED SIGNALS

There are several ways of introducing decorrelation.The easiest way to conceptualize the creation of decorrelation signals is through convolution.To produce a pair of output signal with a specified correlation measure, an input signal can be convolved with of two similar signals that are correlated with each other by specified amount.

3.1.Definition of correlation measure

The correlation measure of two signals y1(t) and y2(t),can be determined by cross correlation function Δt.

$$\Omega(\Delta t) = \lim_{t\to\infty} 1/2T \int_{-T}^{+T} y_1(t)\, y_2(t + \Delta t)dt$$

The user can specify the correlation measure for each pair of output signals in a range from +1 through -1.Lets consider the three basic conditions of correlation measure, which could be applied to the exemplar signals. They are

3.1.1. The two signals are identical

If the two signals have a correlation measure of 1,it assumes that the signals are identical and the peak value of the correlation measure with greatest absolute value is +1..

3.1.2 The two signals are out of phase

If the two signals have a correlation measure of -1,it assumes the signals are identical but out of phase and the peak value of the correlation function with the greatest absolute value is -1.

3.13 The two signals are dissimilar

If the two signals have a correlation measure of 0,it assumes that the signals are dissimilar with the peak value of the correlation function with the greatest absolute value being 0.

Optimum condition for producing decorrelation is subjective but it's always good to iterate between the correlation measure values from -1 to 1 and then choose the most appropriate value for spatializing.

3.2 Building Decorrelated Filters.

In most practical applications of decorrelation the input signal is monophonic and correlating the input signal by specified correlation measure produces the pair of put signals. The computation of the coefficient of the FIR filter could be obtained by the frequency domain specification via the IFFT.The magnitude is set to unity and the phase is constructed from the random number generation of sequences A and B whose correlation measure could be varied from -1 to +1.

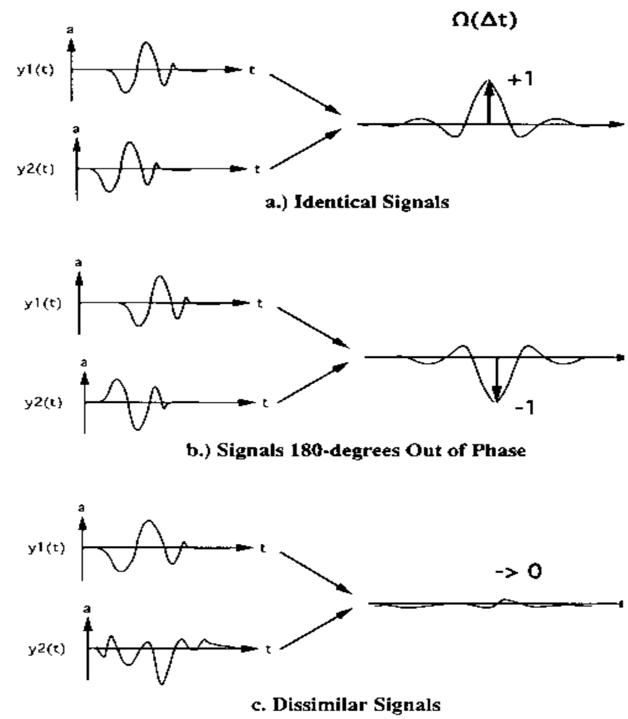

Figure 1.Signals with three different correlation measure categories.

4. ALTERNATE METHOD

4.1.Problem with the previous method

In the previous method the selection of the phase is random and this can give way to unexpected results in the decorrelated output as the power distribution could be affected leading to extreme coloring of the out put signal.

4.2 Lauridsen Decorrelator

An alternate method is using the Lauridsen decorrelator [2].This decorrelator could be viewed as two FIR comb filters(*hL and hR*) with

two taps each for surround left and surround right(since we are decorrelating only the rear channels).The impulse response of these filters are illustrated in Figure 2.A time delay of $\partial = 10$ ms(440 samples) is used between the taps, which is determined experimentally.
The choice of the time delay ∂ is a subtle compromise between the amount of widening and the sound diffuseness. The greater the ∂ is the more diffuse the sound field is going to be, and at some point it will lead to confusion.

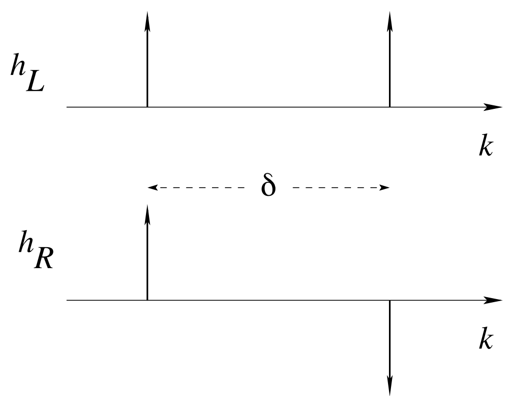

Figure 2. Impulse response of left and right Lauridsen decorrelation filters.

The delay $\partial = 10$ms (440 samples is experiment chosen to produce stereo sounds for the application concerned.

5. HRTF MEASUREMENT

The ability of humans to use spatial cues to estimate the spatial location of a target is of great practical and research importance.Recently, advances in computational power and acoustic measurement techniques have made it possible to empirically measure, analyze, and synthesize the spectral cues which influence spatial hearing.
In this experiment an extensive set of HRTF measurements of KEMAR [3] dummy head head microphone was considered. The measurements consist of the left and right ear impulse responses from a Realistic Optimus Pro 7 loudspeaker mounted 1.4 meters from the KEMAR. Maximum length (ML) pseudo-random binary sequences were used to obtain the impulse responses at a sampling rate of 44.1 kHz. A total of 710 different positions were sampled at elevations from -40 degrees to +90 degrees. Also measured were the impulse response of the speaker in free field and several headphones placed on the KEMAR.

5.1 HRTF selection for Virtual Surround Field

In order to create a 5.1 surround virtual field the impulse responses at the left ear was taken for five locations in the virtual surround field. The centre speaker having zero elevation and an azimuth of 0.The front left and right speakers where chosen to have zero elevation and an azimuth 30˚ and 330˚ respectively. Here it has to be noted that in order to get a faithful reproduction of the virtual left speaker it is necessary to combine the input signal with impulse responses at 330˚ and 30˚ so that we capture their reflections. This applies to the right speaker also. For the left and right surround speakers it is necessary to use the impulse responses at zero elevation and azimuth 110˚ and 250˚

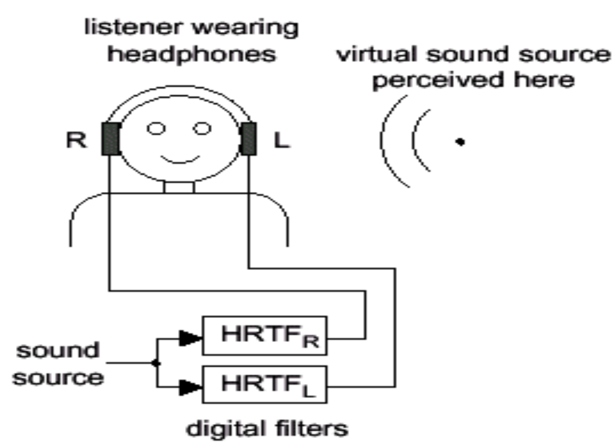

Figure 3.Virtual Surround sound synthesis using HRTF's

6. REPRODUCTION OF MOVING SOURCES

In order to produce moving sources a specific size of the sample is taken at each instant of time and windowed. Now this chunk of the sample is convolved with the HRTF's arranged in zero elevation with the lowest resolution being 5 degrees. Here apart from the direct path a first order reflection from the ground at an elevation of 20 degrees is also considered. The nature of the first order reflection is based on the material of the reflection surface and a shelving filter was designed to model the response. An attenuation is also added considering the air absorption.

7. HRTF INTERPOLATION

For a realistic reproduction of moving sources with no discontinuities it is important to have a very low resolution of the HRTF's in any elevation or azimuth. The lowest resolution of the KEMAR HRTF is 5 degrees. This produce an audible artifact like a click while directly switching between the HRTF's.To avoid this problem and ensure that enough HRTF's are available for a smooth movement of the sound source , adjacent HRTF's are interpolated[6].

8. CONCLUSION

A virtual surround sound system is implemented with the ability to incorporate moving sources. Decorrelation filters that are incorpated in the algorithm adds spatial elements to the audio field. The next step in is to be able enable the listener to choose the listening space that is desired for the audio playback. Acoustic environmental modeling is a technique that could be used for creating a virtual listening space.In this method any reverberant acoustic space could be modeled by computing the impulse response and modeling the transfer function using various shelving and parametric filters. In this way, the listener could choose to listen to the audio from a set of predefined acoustic spaces and also could change the parameters of the filters to model his own listening environment. Another application is to produce music in a binaural set up and to study the most efficient way to compose music for a virtual field by considering the psycho acoustic factors.

9.REFERENCES

[1] Gary S Kendall, E. "The Decorrelation of Audio signals and its impact on spatial Imagery*" published in the Computer Music journal Vol 19,No.4,(Winter 1995),pp 71-87*

[2] Irwan,R and Aarta M Ronald. "Two to Five Channel Sound Processing" *published in the Computer Music journal of the Audio Eng. Society Vol.50,No.11,2002 November*.

[3] Bill Gardner and Keith Martin "HRTF Measurement of a KEMAR Dummy head Micro phone"*MIT Lab May,1994*.

[4] G.Bloch,G.Assayag,O.Warusfel,J.P Jullien "Spatializer:from room acoustics to virtual acoustics",*Proc.Int. Music Conference,1992*

[5] William G Gardner, "3d Acousstic and Environmental Modelling"*Wave Arts Inc March 15 1999.*

[6] Fakheredine Keyrouz and Klaus dipoled "A new HRTF interpolation approach for fast synthesis of dynamic interactive appilcation" *August 15 1999.*