

HRTF(Head Related Transfer Function) for 3D Audio

Mehmet Guney

Sumit Ahuja

Overview

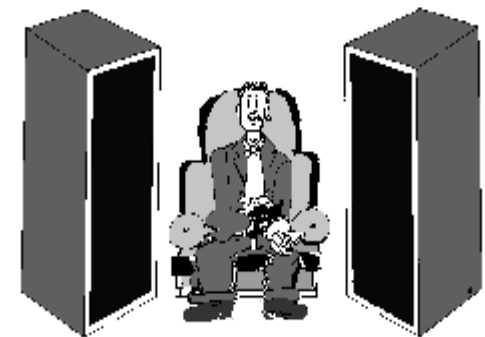
- What is 3D Audio ?
- Head related transfer function (HRTF).
- DSP in 3D Audio.
- Companies/Products.

What is 3D Audio?

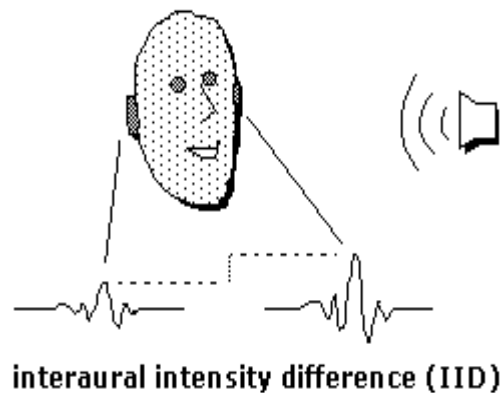
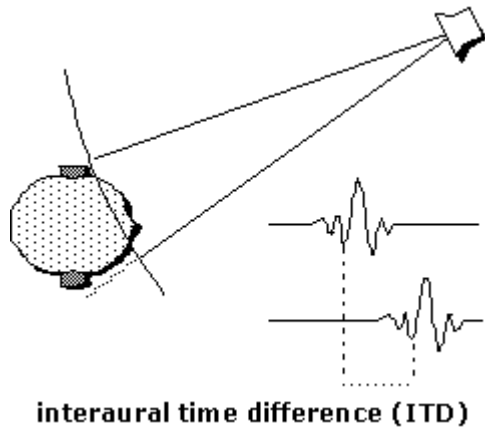


- headphone-based systems (binaural system), using the appropriate HRTF's (Head Related Transfer Function) to process the sound and to get the exact duplication of what the ear would hear in a natural situation

- loudspeaker-based systems (stereophonic system), allowing the creation of spatial auditory images with the help of a set of loudspeakers around the listener.



Head related transfer function (HRTF)



■ Interaural time delay (ITD)

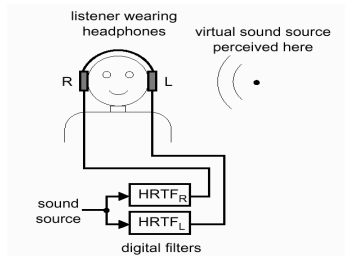
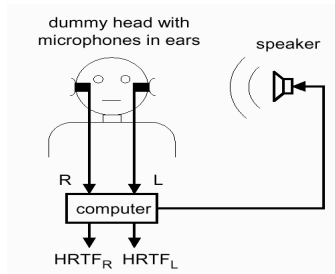
sounds coming from different angles will take different time to reach listener's ears

■ Interaural intensity difference (IID)

Human head absorbs different amount of sound signals coming from different angles.

■ The Pinna Diffractive Effect and the Frequency Vibration of Aural Canal to Sound Wave.

Head related transfer function (HRTF)



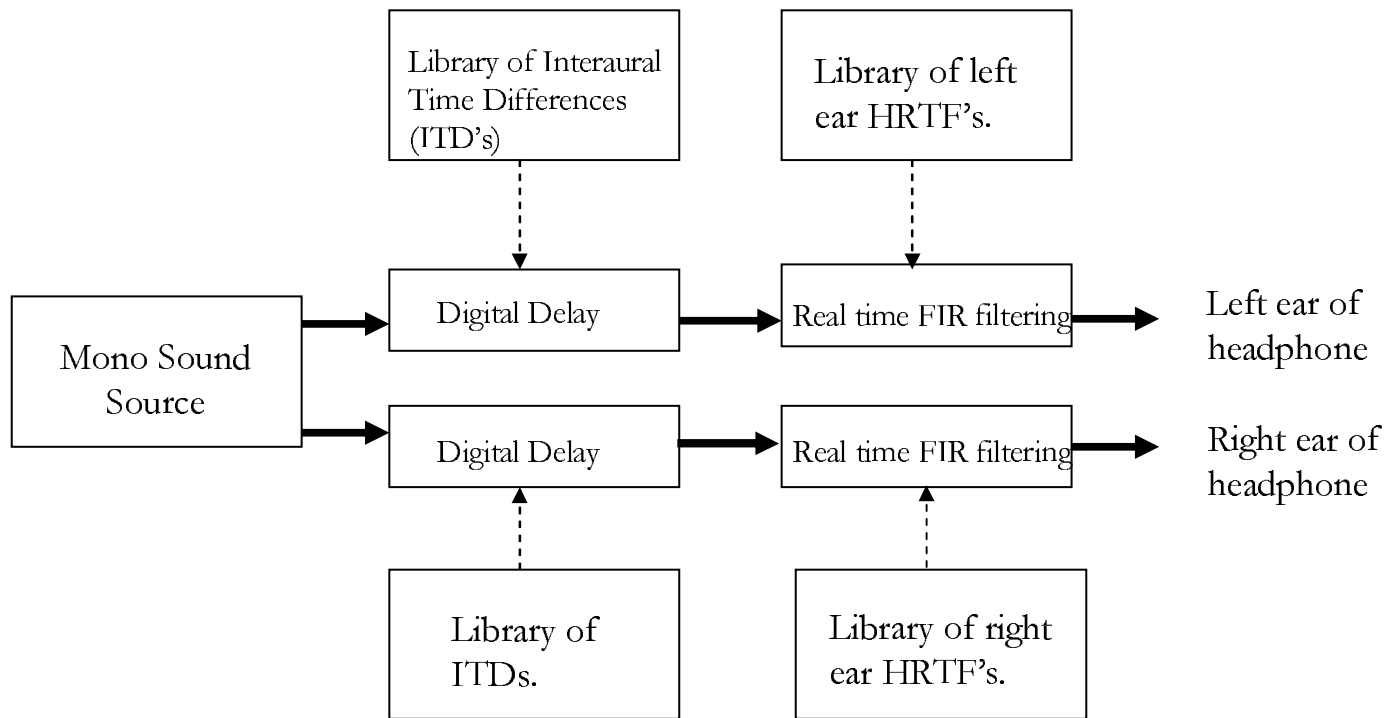
Courtesy: MIT MEDIA LAB.

■ **HRTF** is the transfer function that defines the sound of a point in the space.

■ ***Binaural Recording***: It is repeated for many location of the sound source relative to the listener head for the entire audio space.

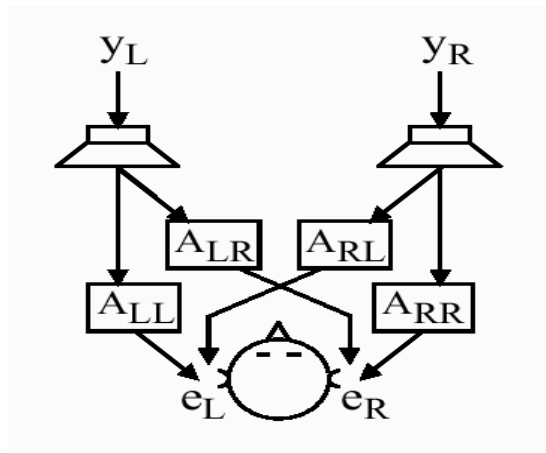
■ ***Binaural Synthesis*** : Obtained data base can be used as a specification for a pair of digital audio filters to reproduce the sound localization cues for each ear

Head Related Transfer Function (HRTF)



The Delays and FIR Filters can be updated in real time to synthesize moving sound source.

Crosstalk Effects



- Is it possible that left ear hear voice coming from left source and right ear from right source?
- The effect is known as crosstalk.
- Signals can be generated for the person sitting exactly in between to reduce the undesired signal because of the other source.
- **Condition:** The listener should sit in Sweet spot.

DSP in 3d Audio

$$y = H x$$

$$\mathbf{y} = \begin{bmatrix} Y_L \\ Y_R \end{bmatrix} \quad \mathbf{H} = \begin{bmatrix} H_{LL} & H_{RL} \\ H_{LR} & H_{RR} \end{bmatrix} \quad \mathbf{x} = \begin{bmatrix} X_L \\ X_R \end{bmatrix}$$

Where

- \mathbf{X} is the column vector representing the signal input to the filter.
- \mathbf{H} contains the filter values
- \mathbf{Y} is the column vector representing the signal at the eardrums of the listener

Time Domain Processing

- To go in frequency domain we use this equation of DFT.

$$\underbrace{\begin{bmatrix} X_1 \\ \dots \\ X_N \end{bmatrix}}_{\text{Freq. Domain}} = F \cdot \underbrace{\begin{bmatrix} x_1 \\ \dots \\ x_N \end{bmatrix}}_{\text{Time Domain}}$$

- F is the $N \times N$ DFT Matrix. Its computational cost is $O(N^2)$.

Frequency Domain Processing

$$\begin{bmatrix} a & d & c & b \\ b & a & d & c \\ c & b & a & d \\ d & c & b & a \end{bmatrix}$$

$$\begin{bmatrix} a & d & c & b \\ b & a & d & c \\ c & b & a & d \\ d & c & b & a \end{bmatrix} = F^{-1} \begin{bmatrix} A & 0 & 0 & 0 \\ 0 & B & 0 & 0 \\ 0 & 0 & C & 0 \\ 0 & 0 & 0 & D \end{bmatrix} F, \quad \text{with} \quad \begin{bmatrix} A \\ B \\ C \\ D \end{bmatrix} = F \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}$$

- Starting from HRTF matrix we can obtain circulant matrix.
- It can be converted to Diagonal Matrix by applying following method

$$\mathbf{r} = [\mathbf{0} | \mathbf{I}] \mathbf{F}^{-1} \mathbf{D} \mathbf{F} \mathbf{s}$$

- Computational Cost with FFT $N(\log N)$.
- So for longer sequences frequency domain filtering is cheaper.

Companies/Applications

- Dolby Digital (Home theater).
- THX from Lucasarts.
- Wavearts : develops Wavesurround for WinAmp and other 3 D products.
- Qsound Labs 3D audio for Internet,PC/Multimedia,consumer electronics and healthcare markets.
- Creative Labs (PC sound Cards).
- AM3D.