Abstract - Sounds that we hear come from all directions and distances. Individual sounds can be distinguished by pitch, tone, loudness, and by their location in space. The spatial location of a sound is what gives the sound its three-dimensional aspect.

Many researches have taken place in the field of 3-D sound effects and have been gradually applied to multimedia devices specifically in theaters and gaming. To achieve these effects many applications adopt Head-Related Transfer Functions (HRTFs), well known for excellent sound localization effect. The HRTF is a linear function that is based on the sound source's position and takes into account many of the cues humans use, to localize sounds.

In the present work, we plan to investigate application of these on a DSP processor.

I- INTRODUCTION
Humans use auditory localization cues to help locate the position in space of a sound source. There are eight sources of localization cues:

1) Interaural Time Difference (ITD)
2) Interaural Level Difference (ILD)
3) Pinna response
4) Head shadow
5) Head motion
6) Shoulder echo
7) Early echo response / Reverberation
8) Vision

Following are the primary cues which play an important role in 3-D sound localization

1) Interaural time difference

Interaural time difference describes the time delay between sounds arriving at the left and right ears. This is a primary localization cue for interpreting the lateral position of a sound source.

2) Interaural level difference

Interaural level difference describes the intensity difference between sounds arriving at the left and right ears. The sound wave reaching the far ear will be attenuated compared to the near ear, creating an interaural level difference between the ears (ILD).

3) Head shadow

Head shadow is a term describing a sound having to go through or around the head in order to reach an ear. The head can account for a significant attenuation (reduced amplitude) of overall intensity as well as provides a filtering effect. The filtering effects of head shadowing cause one to have perception problems with linear distance and direction of a sound source.

4) Pinna Effect

Pinna response describes the effect that the external ear, or pinna, has on sound. Higher frequencies are filtered by the
pinna in such a way as to affect the perceived lateral position, or azimuth, and elevation of a sound source. The response of the pinna "filter" is highly dependent on the overall direction of the sound source.

II - DUPLEX THEORY FOR SOUND SOURCE LOCALIZATION.

According to the "duplex theory", the primary localization cues for interpreting the lateral position of a sound source are Interaural Time Differences (ITD) and Interaural Level Differences (ILD).

Humans generally perceive a sound source to be originating closer to the ear where it arrives earliest and with the greatest amplitude. If we vary the sound source location by azimuth (lateral position as related to directly in front of the listener), the ITDs and ILDs would vary as well.

If the head were not in the way, it would be a trivial matter to calculate the additional distance the sound would have to travel in order to reach the far ear, and then calculate the appropriate ITD and ILD. The problem is significantly complicated by the presence of the head however, as there may not even be direct paths from the source to the far ear. Instead we have to account for the waves reaching the ears after propagating along the surface of an irregular object.

III - HEAD RELATED TRANSFER FUNCTIONS

When sound waves are propagated from a vibrating source to a listener, the pressure waveform is altered by diffraction caused by the torso, shoulders, head and pinnae. In engineering terms, these propagation effects can be captured by two transfer functions, HL and HR that specify the relation between the sound pressure of the source and the sound pressures at the left and right ear drums of the listener.

These Head Related Transfer Functions (HRTFs) are acoustic filters that vary both with frequency and with azimuth, elevation and range to the source. If a monaural sound signal representing the source is passed through these filters and heard through Headphones, the listener will hear a sound that seems to come from a particular location in space. Appropriate variation of the filter characteristics will cause the sound to appear to come from any desired spatial location.

Modeling the Head Related Transfer Functions:

Several attempts have been made to model HRTFs, both to understand their behavior and to simplify the binaural synthesis process. This system identification task has been complicated by four major problems:

- The difficulty of approximating the effects of wave propagation and diffraction by simple, low-order parameterized filters.
- The complicated joint dependence of the HRTFs on azimuth, elevation and range.
- The lack of a quantitative criterion for measuring the accuracy of an approximation.
- The great person-to-person variability of HRTFs.

To avoid these drawbacks the practically recorded impulse responses are used for generating the 3D effect.
IV - IMPLEMENTATION

For the implementation we need the impulse responses of what our left and right ear sense. These are difficult to generate manually and are more convenient to be calculated experimentally. Thus for our implementation also we used the impulse responses of our left and right ears experimentally calculated at various horizontal angles and vertical angles. These experiments have been done by MIT and the impulse responses were available in anonymous domain for academic utility.

**Study of impulse responses**

**General characteristics:** The general nature of the impulse response is shown in figure below:

![General nature of Impulse Response](image)

As can be seen above the dummy ear has sensed the response after certain duration. This is due the fact that the sound source was kept about 0.4 meters away from the source. For this reason the sound was detected by the dummy ears after $0.4/330 = 1.2$ ms which corresponds to around 50th sample in the response. It is also observed that the sound is heard again and again with decreasing amplitude in time at a duration of about 10 samples corresponding to 0.22ms. This is due to the face that many sound waves reach the ear after reflection from internal parts of ear, outer ear shape, head, shoulders etc.

**Effect of change in azimuth angle:** The following figures demonstrate the impulse responses of left and right ears at 0 degree elevation but azimuth being 0 degree in one case and 45 degree in another.

![Impulse response perceived by left and right ear at 0 degree elevation and 45 degree azimuth.](image)
impulse responses perceived by the left and right ear. As expected, the right ear hears the sound earlier than the left ear as the source is now near the right ear. Also the nature of the reverberations perceived by left and the right ear are different.

**Effect of change in elevation and azimuth angle**

The following impulse responses are the response of the left at and right ear when the source is kept at 50 degrees of elevation and 48 degrees azimuth angle.

As expected the sound is received early by the right ear than the left. But there are also considerable changes in the nature of the reflected sound. As the elevation changes the sound reflected by outer ear, head and shoulders get reflected in very complex manner and cannot be in generally explained.

**Impulse response perceived by left and right ear at 50 degree elevation and 48 degree azimuth.**

**Frequency (DFT) response perceived by left and right ear at 50 degree elevation and 48 degree azimuth.**
Sound waves that come from different directions in space are differently scattered by the listener’s outer ears, head, shoulders, and upper torso. The scattering leads to an acoustical filtering of the signals appearing at left and right ears. The filtering can be described by a complex response function—the anatomical transfer function (ATF), also known as the head-related transfer function (HRTF). Because of the ATF, waves that come from behind tend to be boosted in the 1000 Hz frequency region, whereas waves that come from the forward direction are boosted near 3000 Hz. The most dramatic effects occur above 4000 Hz: In this region, the wavelength is less than 10 cm and details of the head, especially the outer ears, or pinnae, become significant scatterers. Above 6000 Hz, the ATF for different individuals becomes strikingly individualistic, but there are a few features that are found rather generally. In most cases, there is a valley-and-peak structure that tends to move to higher frequencies as the elevation of the source increases from below to above the head. The peak near 7000 Hz is thought to be a particularly prominent cue for a source overhead. The direction-dependent filtering by the anatomy, used by listeners to resolve front–back confusion and to determine elevation, is also a necessary component of externalization.

V - CONCLUSION
The modeling of HRTFs is a challenging problem in system identification. Alternate ways of generating 3D effect is through the convolution of impulse responses with the input (monaural) sample file. A simple ellipsoidal model of the head that includes downward and backward displacements of the ears can provide accurate values for the ITD for an HRTF model, and can be used to generate the impulse responses.

References:

[1] Gestur Bjorn Christianson

[2] Sun Inc., World Wide Web,

