Binaural Demonstration of a Schizophrenic Episode

Irfan Adil Navaz (Student ID: 301074887 Exam no: Y3928473)

CONTENTS

I	**Introduction**			1
II	**Background Theory**			1
	II-A	**Binaural Audio**		1
	II-B	**Schizophrenia**		1
	II-C	**Impulse Response**		1
	II-D	**Head Related Impulse Responses**		2
	II-E	**Convolution**		2
	II-F	**Digital Audio Filters**		2
		II-F1	**Low-pass Filter (LPF)**	2
		II-F2	**High-pass Filter (HPF)**	2
		II-F3	**Band-pass Filter (BPF)**	2
		II-F4	**Band-stop Filter (BSF) or Notch Filter**	2
		II-F5	**Shelving Filter**	2
		II-F6	**Peaking Filter**	2
		II-F7	**All-pass Filter**	2
	II-G	**DAFX**		3
		II-G1	**Flanger**	3
III	**Evaluation**			3
References				3

I. INTRODUCTION

The objective or aim of this project is to successfully create a binaural soundscape representing an episode of schizophrenia, which is a mental illness that can cause a person to hallucinate and undergo delusions, both visually and auditory. By using various voice lines and other 'paranoia' inducing sounds, we can create an almost realistic recreation of a schizophrenic episode

In this report, we aim to compile, read, process and convolve audio sources in such a way, we create a soundscape that surrounds and immerses the listener in an unusual schizophrenic episode. In Matlab, we use various audio processing tools and HRIR to successfully create this soundscape. In this project, we are using convolution and HRIR spatial processing to create a sort of 360 degree environment. We can use a variety of audio processing tools that includes but is not limited to digital audio filters, digital audio effects and so on. Audio sources involved in this project were initially to be recorded, but instead sourced from online resources, as to obtain high quality audio that can serve our purposes of this project.

More applications that use 3D sound to produce an immersive auditory experience are becoming commonplace as spatial audio technologies advance in both recording and reproduction. With the advent of extensive spatial audio tools for sound capture and 3D sound rendering over headphones and speaker arrays, soundscape design can now take advantage of these resources. .

II. BACKGROUND THEORY

A. Binaural Audio

Binaural audio simulates the inter aural time and level variations that take place between the ears to produce a realistic, spatial sound experience. This technology enhances the realism of auditory experiences by providing an accurate sense of direction, distance, and depth for sound sources in a virtual environment. It is widely used in virtual reality, gaming, and entertainment.[1] When it comes to localizing sounds in relation to their heads, humans excel at it. Using the ways that sound is altered on its way from the source to the ears—particularly the ways that sound is altered differently in each ear—the human brain is able to localize sounds.[2]

Binaural cues, such as the inter aural time difference (ITD) and inter aural level difference (ILD), are a result of these variations in sound between the ears.

Numerous factors, most notably the shape of the head, pinnae, and upper torso, influence these cues. The distance that sound travels to each ear, attenuation from occlusion (the head shadow effect), and reflections (from the upper torso and the pinnae) can all have an impact on these. Furthermore, the brain can deduce source direction from the pinna's creation of direction-dependent spectral cues.[3]

B. Schizophrenia

A severe mental illness characterised by abnormal reality interpretation. A combination of hallucinations, delusions, and severely abnormal thinking and behaviour that can be debilitating can be symptoms of schizophrenia. These symptoms can interfere with day-to-day functioning. Schizophrenia patients need lifelong care.

Over time, there may be phases when symptoms get worse and times when they get better. Certain symptoms might not go away. Schizophrenia symptoms in men usually appear in their early to mid-20s. Symptoms in women usually start in their late 20s. Schizophrenia is not often diagnosed in children and is not frequently diagnosed in adults over 45.

C. Impulse Response

Playing a sound or an impulse in a space elicits an impulse response. This impulse can sound like a sine sweep, which is a sine tone that pitches up through the audible frequency spectrum, or it can sound like a starter pistol, clapboard/slate, balloon popping, etc., which are short, percussive sounds. Because of the space's distinct acoustics, this impulse creates a momentary image of the ambience that is specific to it.

The audio that is produced is recorded using microphones. We can observe how the acoustics of the space or object influence the timbre of the resultant sound by keeping the initial impulse in mind. Ideally, the recording would be edited to remove the initial impulse, leaving only the space's acoustic response.

D. *Head Related Impulse Responses*

HRIRs, or head-related impulse responses, enable the generation of virtual acoustic environments. Given that the human auditory system is capable of extremely accurate sound localization under ideal circumstances, having an HRIR database with a high spatial resolution is helpful in producing realistic-sounding scenes.[4] HRTF is an essential element. It describes how sound waves enter the head and ears and are filtered. HRTFs offer comprehensive details regarding the modifications made to a sound's frequency content during its journey from the source to the ears. Accurate spatial audio reproduction is made possible by the mathematical representation of this information. An essential component of an HRIR database is spatial resolution, or the density of directions at which the HRIRs are recorded. This is because it influences the accuracy of incident direction selection.[5]

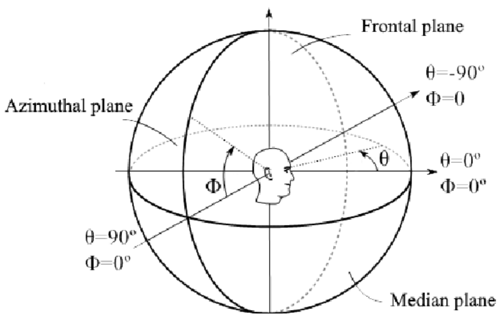

Fig. 1. Parameters in HRIR

There are still obstacles in the way of obtaining a genuinely accurate and personalized HRIR, despite progress. There are constant challenges due to things like head movement, the need for high-fidelity measurements, and the requirement for real-time processing. Subsequent investigations endeavor to tackle these concerns and enhance spatial audio technologies even more.

One of the more advanced techniques frequently employed in audio production is convolution. Its accurate transfer of space and object timbres to other signals is beneficial for both standard processing applications and sound design. Convolution can be a great tool for producers of all stripes, offering a vast array of realistic and otherworldly sonic possibilities.

E. *Convolution*

Convolution is a method of cross-synthesis, combining two audio sources in such a manner that, in the frequency domain, those frequencies they have in common will be emphasized proportionately, and those they do not share will be minimized. Cross-synthesis is the process of imparting some or all of the properties of one signal to another.[6]

Another aspect of the convolution process is how those frequencies linger, become smeared, and eventually disappear in the time domain. The sources may both be digital audio files of finite length, or one finite file and the other real-time input (potentially infinite). Convolution is now widely used to create reverb of higher quality, filter audio, and impart specific qualities of one sound file to another.

The convolution process does not involve oscillating wave forms. Instead, we work with two audio sources, an input signal and an impulse response. The sound that will be impacted is called the input signal, and the impulse response contains the sonic characteristics of the area or thing that we will add to the input signal.

Since any sound can be used as an impulse response or an input signal, convolution really has no limits. This is a great tool for sound design because it allows you to record and intentionally select timbres for any desired effect.

F. *Digital Audio Filters*

Digital audio filters are essential components in the field of signal processing, particularly in audio processing and music production. They manipulate the frequency content of digital audio signals, allowing for various effects such as equalization, noise reduction, and modulation.

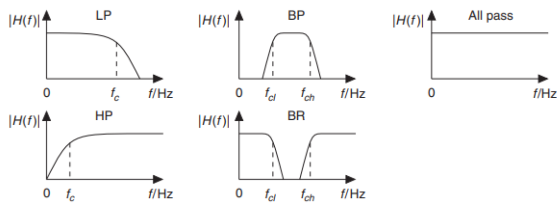

Fig. 2. Filter Classification

*1) **Low-pass Filter (LPF):*** Allows frequencies below a certain cutoff point to pass through while attenuating higher frequencies.

*2) **High-pass Filter (HPF):*** Allows frequencies above a certain cutoff point to pass through while attenuating lower frequencies.

*3) **Band-pass Filter (BPF):*** Allows a specific range of frequencies to pass through, attenuating frequencies outside that range.

*4) **Band-stop Filter (BSF) or Notch Filter:*** Attenuates a specific range of frequencies while allowing frequencies outside that range to pass through.

*5) **Shelving Filter:*** Boosts or cuts frequencies above or below a specified cutoff point.

*6) **Peaking Filter:*** Boosts or cuts frequencies around a specified center frequency.

*7) **All-pass Filter:*** Does not affect the amplitude of the signal but changes the phase relationship of different frequency components.

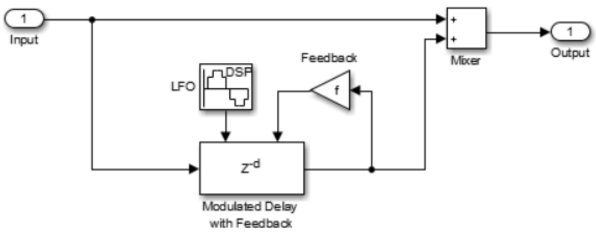

Fig. 3. Block Diagram for Flanger

G. *DAFX*

The term "digital audio effects" (DAFX) describes how audio signals are altered and modified through the use of digital processing methods. By applying these effects, audio recordings can have their sound altered in different ways, improving or changing the auditory experience.

Control over variables like depth, rate, feedback, wet/dry mix, and more is frequently offered by DAFX. Automation increases expressiveness by enabling these parameters to change over time.

*1) **Flanger**:* The use of magnetic tape reels gave rise to the term "flanging". In the early days of audio production, the effect was created by engineers manipulating the rims or flanges of two tape machines. In order to accomplish this, one tape machine had to be manually sped up or slowed down, delaying the signals.

When two turntables or tape machines are playing the same recording and you alternately brake them by lightly pressing a finger on each flange, this is known as "flanging." The outputs of both playback devices are summated to produce the audible effect. A chorus, phaser, and flanger are closely related devices that all modify the phase relationships between multiple signals. However, each effect has its unique characteristics.

With so many creative options for modifying the timbre and spatial qualities of sound, the flanger is still a highly expressive and adaptable tool in the audio engineer's toolbox. The flanger is still a standard in many different musical genres, whether it is employed delicately to add depth or very loudly for a striking effect.

A straightforward mathematical formula that describes the delay modulation can be used to implement the flanger effect. Applying a time-varying delay to the original signal is one popular technique. The following formula yields a basic flanger effect:

$$y(t) = x(t) + x(t - D(t)) \tag{1}$$

The modulation of the delay (D(t)) over time holds the secret to the flanger effect. Generally, this modulation function is implemented as a low-frequency oscillator (LFO). The delay is adjusted using a periodic waveform produced by the LFO. Waveforms like sine, triangle, and sawtooth waves are frequently used for modulation.

The delay is applied in accordance with the discrete time intervals at which the input signal is sampled. The time-varying delay effect is produced by updating the modulation function at each time step.

It's crucial to remember that more complex flanger implementations could call for extra settings and factors like stereo processing, feedback, and filtering. Furthermore, the way the flanger effect is implemented can vary depending on the kind of filter that is used and the desired results. Digital signal processing methods are frequently incorporated into advanced flanger designs for effective and adaptable real-time applications.

III. Evaluation

We have successfully created a binaural soundscape that depicts a schizophrenic episode, using audio processing and spacialization tools in matlab.

An incredibly immersive and engrossing project has been produced by employing four audio sources to create a binaural soundscape and sophisticated audio processing and spatialization techniques. An unparalleled degree of audio realism has been made possible by the painstaking integration of binaural recording technology, providing listeners with a three-dimensional auditory experience. The four audio sources were carefully chosen and worked with to create a rich, dynamic soundscape that highlights the project's versatility. The use of audio processing has improved the quality even more, guaranteeing a smooth transition between the various components of the composition. With the deft use of spacialization techniques, the audience is submerged in a sonic world that goes beyond conventional stereo recordings, gaining a sense of depth and directionality.

References

[1]J. Thiemann and S. van de Par, "or aided and unaided ears," EURASIP Journal on Advances in Signal Processing, vol. 2019, no. 1, Feb. 2019, doi: 10.1186/s13634-019-0604-x.

[2] T. Wendt, "A Computationally-Efficient and Perceptually-Plausible Algorithm for Binaural Room Impulse Response Simulation," Dec. 04, 2014. https://doi.org/10.17743/jaes.2014.0042

[3]J. E. Summers, "Auralization: Fundamentals of Acoustics, Modelling, Simulation, Algorithms, and Acoustic Virtual Reality," Journal of the Acoustical Society of America, Jun. 01, 2008. https://doi.org/10.1121/1.2908264

[4]C. Roads, S. T. Pope, A. Piccialli, and G. De Poli, Musical Signal Processing. Routledge, 2013. [Online].

[5]J. Y. Hong, J. He, B. Lam, R. Gupta, and W. Gan, "Spatial Audio for Soundscape Design: Recording and Reproduction," Applied sciences, Jun. 16, 2017. https://doi.org/10.3390/app70606A multiple model high-resolution head-related impulse response database f27

[6]"The Basics of Convolution in Audio Production." https://www.izotope.com/en/learn/the-basics-of-convolution-in-audio-production.html