Motivation definition

Virtual reality (VR) has dominated tech headlines in recent years with its ability to immerse its users in a virtual world. Gaming is the most well-known uses for VR. Moreover, virtual Reality has been adopted to education for teaching methods, to medical training and to military. Since the focus of the virtual reality research is graphics and hardware, the research on sound stayed in the background. However, for a truly immersive experience, Virtual Reality needs convincing sound to match(audio is essential), immersive graphics require equally immersive 3D audio that provides a natural listening experience. \cite{1} Otherwise, without high quality audio, users do not truly feel part of the virtual reality. \cite{8} Various studies have been carried out on this topic, which shows the benefits of spatial sound in VR.\cite{9}

A challenge for a realistic 3D audio in VR is complex virtual scenes with thousands of sounds sources \cite{2}. A basic concept of 3D audio can be simplified as the convolution of anechoic sound file and the matching HRTFs to its locations. For each source, the computational workload increases proportionally \cite{3}. To reduce computational complexity while maintaining the same auditory impression for the listener a a virtual scene, many different methods have been proposed. Some methods concern reducing the number of sound sources, some concern reducing the propagations paths.

Ambisonics is a method to record, modify, and recreate audio in 360 degree which reduces the number of propagation paths significantly for a complex virtual scene.

The advantage of Ambisonics over other surround technologies is that ambisonics introduces a format which is independent from the speaker configuration and the number of source signals. Unlike other sound techniques Ambisonics, encoding does not contain any specific original speaker information and does not require one specific loudspeaker setup to recreate the sound field. Another major benefit of this approach is that the recorded sound can be reproduced with a variable number of speakers at a variety of positions.

The number of the loudspeakers in the setup limits the number of the convolutions required for a binaural decoder to 2xN for a number of N loudspeakers. Increase in number of speakers doesn’t do any changes in number of convolutions. This works well for reducing the computational complexity for complex virtual scene. Therefore, Ambisonics has much potential to improve the virtual reality sound experience.

In this bachelor thesis, ambisonics method with different speaker layouts will be implemented and the performance will be tested and evaluated.