

9th International Symposium on ElectroAcoustic Technologies

2024-11-02/03

Place:

Shenzhen, China

Title:

Auditory Testing of the New 3GPP Speech and Audio Codec

Authors:

Magnus Schäfer

Jan Reimes

Abstract:

In recent years, there have been a lot of scientific advancements in the field of immersive audio technology and the topic has its first consumer applications as well. While these first applications mostly focus on playback or streaming scenarios, it can be expected that communication use cases will follow soon. Thus, a suitable transmission system for immersive audio communication will be necessary.

Accordingly, the 3rd generation partnership project (3GPP) as the corresponding standardization body initiated a work item to develop a new low-delay speech and audio codec for services with stereo or immersive audio communication. This codec is called Immersive Voice and Audio Services (IVAS). It will support different stereo and immersive audio formats at bit rates ranging from 13.2 to 512 kbps and it is designed as an extension to the latest 3GPP codec, Enhanced Voice Services (EVS). Within the standardization process, a joint codec candidate was developed by a large collaborative group of companies. This candidate must demonstrate its capabilities and performance in auditory experiments.

This presentation provides an insight into designing and conducting the listening tests for these experiments and analyzing their results from the point of view of one of the listening labs and the global analysis lab that were responsible for this stage of the standardization process. The listening tests were designed for different qualifications of test subjects (naïve and experienced), they utilized different playback systems (headphones and multi-channel loudspeaker setups) and used two distinct rating paradigms (degradation category rating and MUSHRA: multiple stimulus with hidden reference and anchor). All tests assessed the performance of the IVAS codec candidate in comparison to the EVS codec. The IVAS codec candidate performed better than EVS by either providing better quality at comparable bit rates or using less data to achieve the same quality.