

## **Description for the general public**

Spatial audio processing plays an important role in entertainment and audio communication applications. Spatial hearing is no longer reserved for immersive audio over headphones or multiple loudspeakers but it slowly enters the communication domain, in which better understanding of individual speakers in group or conference calls is offered when spatially reproducing voices of distant speakers. For such emerging advanced multimedia presentation, spatial processing of the sound scene recorded using devices equipped with several microphones are required.

Over the recent years we have seen a rapid development of various audio signal analysis methods which allow for a more efficient and effective automation in many areas, in particular in recognition of speech or voice, estimation of spatial location and acoustic source identification. These techniques have been developed as processing a single audio channel recording the source signal of interest from close proximity. On the other hand, in real life scenarios, the recorded microphone signals inevitably contain other sounds which are not of interest to the given application.

In this project we aim to explore the possible methods of integration of the spatial audio processing and speech processing, which have evolved over the years independently from one another. Of particular interest in this project is the spatial audio with methods that are based on the analysis of the recorded sound scene, which could benefit from audio-based machine learning which performs the classification of sound sources and speaker identification. The goal of the project is to develop methods, in which these areas will not be logically separated, but when combined, they could be used jointly to offer better performed in both spatial audio and signal analysis such as source recognition.

At early project stages we will investigate the ways to integrate the processing known in the aforementioned domains and will seek the methods which offer notable benefits. For instance, we will aim to determine which spatial acoustic parameters could be used as auxiliary features in source classification and speaker recognition methods. We will also focus on incorporating knowledge gained from source signal analysis using machine learning techniques into the spatial audio framework. An expected result of integrating the machine learning techniques with spatial sound analysis could for instance be a single algorithm that identifies the source with higher accuracy than it would be the case with all modules used separately. The outcome of the project would therefore be not only the development in each research area but also insights into creating synergies between them.