METADATA

Title: Audio, speech and music processing and analysis

Other Titles: -

Language: Greek

Authors: Koutras. A., Associate Professor, UOP, Alexandraki, Ch., Assistant Professor, HMU, Zarouchas, Th., Dr., CTI Diophantus, Zervas, P., Assistant Professor, UOP, Hatziantoniou, P., Laboratory and Teaching staff, UPATRAS

ISBN: 978-618-228-115-4

Subject: NATURAL SCIENCES AND AGRICULTURAL SCIENCES, ENGINEERING AND TECHNOLOGY, HUMANITIES AND ARTS, LAW AND SOCIAL SCIENCES

Keywords: Acoustics / Audio Signal Processing / Time Frequency Analysis / Speech Processing / Speech Synthesis

Bibliographic Reference: Koutras, A., Alexandraki, C., Zarouchas, T., Zervas, P., & Chatziantoniou, P. (2023). Audio, speech and music processing and analysis [Undergraduate textbook]. Kallipos, Open Academic Editions. http://dx.doi.org/10.57713/kallipos-347

Abstract

This textbook is divided into four main sections. The first section, titled "Audio" provides an introduction to audio signals and digital sound. It presents the main functions of the human auditory system and the fundamental concepts of psychoacoustics and spatial acoustics. The behavior of sound in open and closed spaces is described, along with basic concepts from the theory of signals and systems for the analysis and description of room acoustics. Additionally, it presents the most significant descriptors used in the analysis, processing, and recognition of audio signals. Furthermore, it discusses some major applications of Digital Sound and Acoustics Technology. In the second section, titled "Speech" the basic concepts and theories of digital speech processing are presented. Its aim is to design and implement basic classifiers for speech recognition, speaker identification, and other applications. Moreover, it provides an overview of the basic stages of text-to-speech conversion, along with an analysis of the technologies and algorithms that are proposed in the literature. The third section, titled "Music", provides an introduction to musical signals and music information. It covers the structure of music information and the techniques for encoding musical signals. Furthermore, it introduces a relatively new research area known as Music Information Retrieval (MIR) and explores its main applications. Lastly, it presents a variety of algorithmic problems in the field of MIR. In the fourth and final part of the textbook, titled "Lab Exercises", eleven indicative lab exercises/projects are presented. These exercises offer readers the opportunity to test and apply the concepts that were discussed in all previous chapters.

The Project is funded by the National Development Programme 2021-2025 of the Ministry of Education and Religious Affairs and implemented by the Special Account for Research Funds of the National Technical University of Athens and the Hellenic Academic Libraries Link.