Introduction to Audio Analysis—A MATLAB Approach

T. Giannakopoulos and A. Pikrakis,
Elsevier—Academic Press, Waltham, MA (2014), 288 pp., hardbound, 79.96 USD, ISBN 978-0-0-8-099388-1.

This book provides a palatable introduction to the field of audio analysis, presenting concepts and examples using the MATLAB programming language as a guide. For readers with general audio, science or math backgrounds who would like to gain insight on the basic theory and practicalities of various audio analysis topics, this book is a good resource. To aid in the introduction of these topics, the book contains "the MATLAB Audio Analysis Library" with accompanying MATLAB files. These files contain useful codes for performing numerous audio analysis tasks, which are explained in the text. The book also contains exercises at the end of each chapter, which may be completed by the interested learner to help develop his or her audio analysis and programming skills. These exercises would also be useful in an instructional setting, providing good assignments for students in an audio analysis class (as an added bonus, they are even rated according to difficulty). Practical explanatory examples, which are helpful for interpreting and contextualizing the audio analysis theory presented, are included throughout the text.

The book is divided into three parts, with the first covering fundamental audio analysis topics and techniques that are directed to beginners in the field. The first chapter succinctly summarizes the book contents. Chapter 2 provides descriptions and MATLAB examples on the creation, playback, loading and storage of audio signals, which is necessary before analysis can begin. Signal transforms and filtering concepts are discussed in Chapter 3 including the introduction of the Fourier, Cosine and Wavelet Transforms. Aliasing is also considered. Chapter 4 introduces audio features and feature extraction techniques, which are essential for pattern recognition and machine learning tasks. Discussions of both time-domain and frequency-domain audio features are included.

The second part of the book delves into more advanced topics related to characterizing audio content. A reader who is already familiar with the basics may start with this part if so inclined. Chapter 5 describes how the features extracted (per Chapter 4) may be used to classify audio signals. This chapter concentrates on classification algorithms useful for classifying audio segments to predefined classes, including the Bayesian classifier, *k*-nearest-neighbor classifier, and support vector machines. Chapter 6 focuses on audio segmentation techniques, which are designed to separate an audio signal into its homogeneous content. Chapter 7 revisits the topic of classification techniques, discussing strategies that utilize the temporal evolution of audio signals. Topics include hidden Markov modeling, the Viterbi algorithm and the Baum–Welch algorithm.

The third part of the book provides a short summary of music information retrieval tasks, such as music identification, which is used by popular mobile applications to record a music signal and identify the song. The discussion of music information retrieval provided in Part III is only intended to provide a general insight into the topic. The text is not designed for the reader to gain an in-depth understanding of the music information retrieval tasks presented.

In short, this book is excellent starting point for those who would like to learn about audio analysis techniques and develop practical skills for using MATLAB to perform audio analysis tasks. While much of the information presented in Parts II and III would require additional reading to master the topics at hand, this book serves very well to provide fundamental theory and contextualized examples relevant to the audio analysis field.

Lauren M. Ronsse
Acoustics Program
Columbia College Chicago
Chicago, IL, USA
lronsse@colum.edu