
Machine Listening for Music and Sound Analysis

Dr.-Ing. Jakob Abeßer

Fraunhofer IDMT

Jakob.abesser@idmt.fraunhofer.de

Machine Listening

Combine **signal processing** and
machine learning to extract
information from sound & music

Overview

■ Lecture Structure

■ Fundamentals

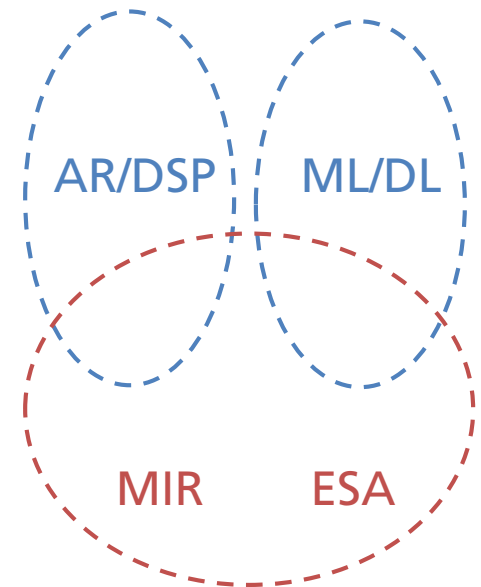
- L1 - Audio Representations & DSP
- L2 - Machine Learning & Deep Learning

■ Applications

- L3 & L4 - Music Information Retrieval
- L5 & L6 - Environmental Sound Analysis

■ Additional Content

- Insights into projects & current research @ Fraunhofer IDMT
- Open student topics



Own

Overview

■ Seminar Structure

- S1 – Introduction to Python programming
- S2 – Basics: Audio processing, machine learning, and deep learning
- S3 – Music classification
- S4 – Sound classification

■ Notes

- Programming in IPython notebooks / Google Colaboratory
- Additional course material (audio samples, libraries)

Course Website

<https://www.machinelisting.de>

Further Resources: Books

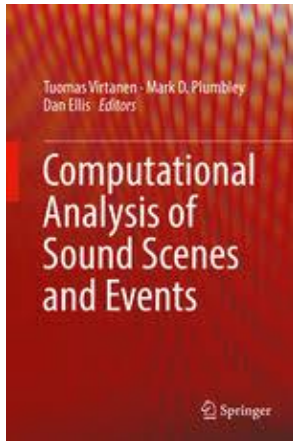


Fig. 1

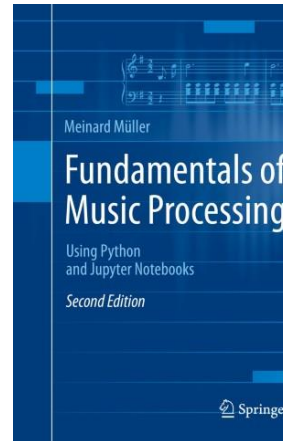


Fig. 2

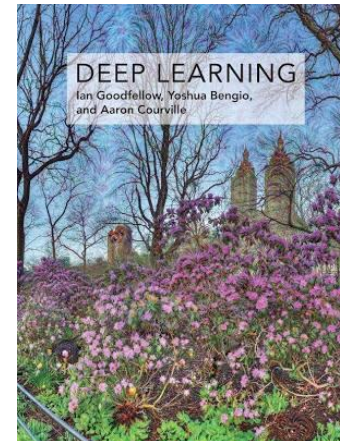


Fig. 3

- Virtanen, T., Plumbley, Mark D., and Ellis, Dan: Computational Analysis of Sound Scenes and Events, Springer, 2018.
- Müller, M.: Fundamentals of Music Processing – Using Python and Jupyter Notebooks, Springer, 2021.
- Goodfellow, I., Bengio, Y., and Courvill, A.: Deep Learning, The MIT Press, 2016.

Further Resources: Webpages

- Machine Learning / Deep Learning
 - <https://www.deeplearningbook.org/>
 - <http://www.coursera.org> (online courses)
 - <http://www.udemy.com> (online courses)
 - <https://machinelearningmastery.com/deep-learning-books/>

Further Resources: Webpages

■ Music Information Retrieval

- <https://www.audiolabs-erlangen.de/FMP> (iPython notebooks)
- <https://musicinformationretrieval.com> (iPython notebooks)
- <https://audiolabs-erlangen.de/PCP> (Preparation Course Python Notebooks)
- <https://github.com/meinardmueller/libfmp> (Python package for music processing)
- <https://github.com/meinardmueller/synctoolbox> (Music synchronization)
- <https://github.com/meinardmueller/libtsm> (Time-scale modifications & pitch shifting)

Further Resources: Webpages

- Environmental Sound Recognition
 - <http://dcase.community/> (DCASE challenges & workshop)

Further Resources: Programming Libraries

- General

 - numpy, scipy, scikit-learn, matplotlib, pandas

- Machine Learning / Deep Learning

 - scikit-learn, tensorflow, keras, (pytorch)

- Audio & Music Processing (Python)

 - pysox, soundfile (audio I/O & manipulation)

 - librosa, madmon, libfmp, synctoolbox, libtsm (audio & music processing)

 - Music21, MeloSpyLib (symbolic music processing)

 - (MIR Toolbox – Matlab)

Acknowledgments

- Meinard Müller (International Audio Laboratories)
- Sebastian Stober (Otto-von-Guericke-University Magdeburg)
- Patrick Aichroth (Fraunhofer IDMT)

- Michael Taenzer (Fraunhofer IDMT)
- Christian Kehling (TU Ilmenau / Fraunhofer IDMT)

- Christian Dittmar (Fraunhofer IIS)
- Estefanía Cano Ceron (AudioSourceRE)
- Christof Weiß (International Audio Laboratories)
- Daniel Gärtner (MusicDNA)

Images

Fig. 1: <https://media.springernature.com/w306/springer-static/cover-hires/book/978-3-319-63450-0>

Fig. 2: <https://media.springernature.com/w306/springer-static/cover-hires/book/978-3-030-69808-9>

Fig. 3: <https://mitpress.mit.edu/books/deep-learning>