

Audio Processing and Indexing

E.M. Bakker

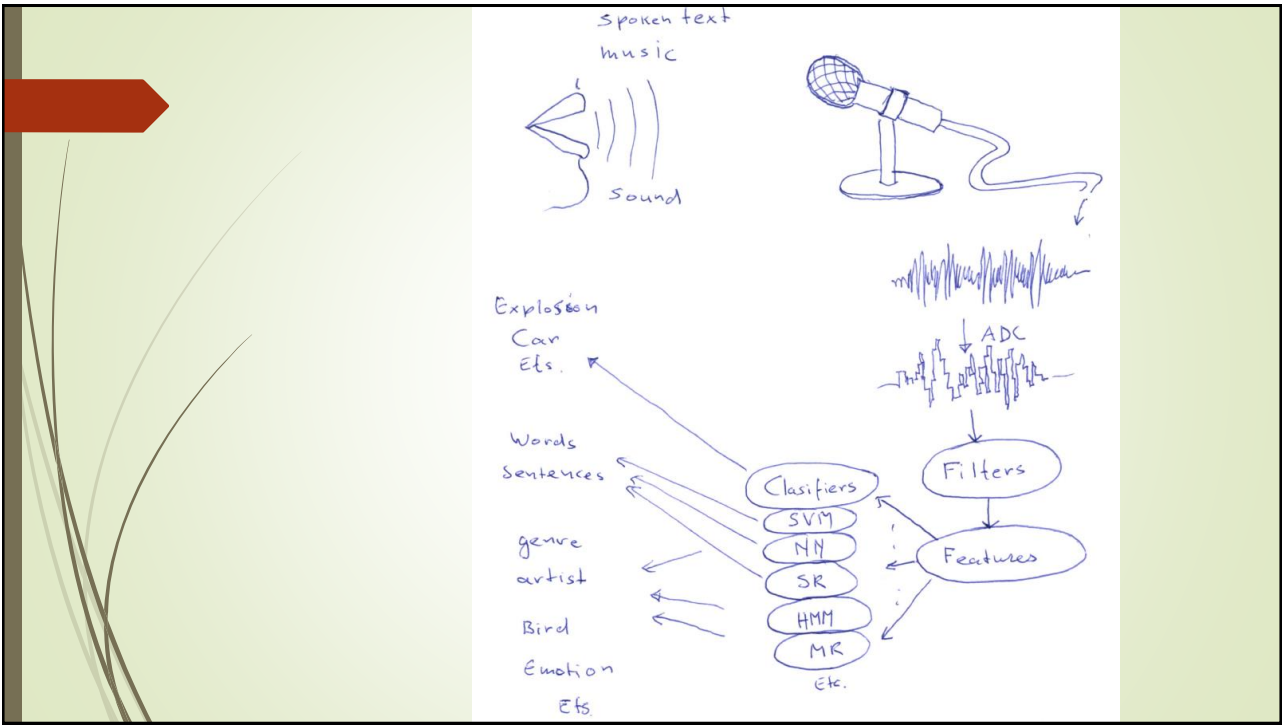
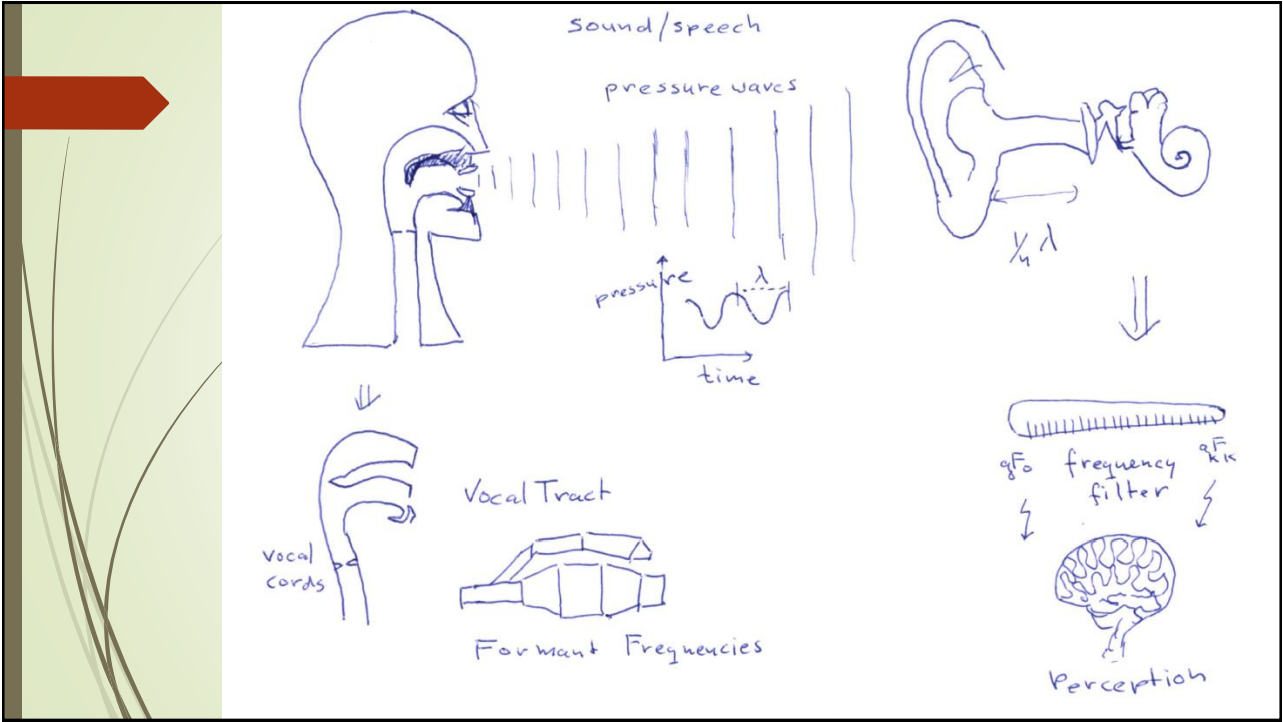
Overview

- 5-9 Organization and Introduction
- 12-9 Audio Production and Processing + Vocal Tract Workshop@Home
- 17-9 ADC and an Algebraic Introduction to FT
- 24-9 FFT + FFT Workshop@Home
- 3-10 No Class: Leidens Ontzet
- 10-10 Project Proposals (presentations by students) + Student Paper Selection
- 17-10 Audio Features and Data Sets + Audio Features Workshop@Home
- 24-10 Machine Learning + Machine Learning Workshop@Home
- 31-10 Student Paper Presentations I.
- 7-11 Student Paper Presentations II.
- 14-11 Student Paper Presentations III.
- 21-11 Student Paper Presentations IV.
- 28-11 TBA
- 5-12 **Final Project Presentations Demo**
- 19-12 Project Deliverables:
 - Final Technical Project
 - Scientific/technical paper (4-8 pages), code
 - Web Site (or github)

Online Project Progress Meetings

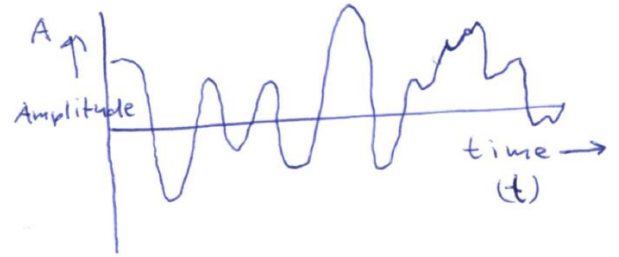
Grading (6 ECTS):
 Presentations (20%) and Project (40%) (60% of grade).
 Class discussions,
 attendance, and workshops (40% of grade).

See: <http://liacs.leidenuniv.nl/~bakkerem2/api/>



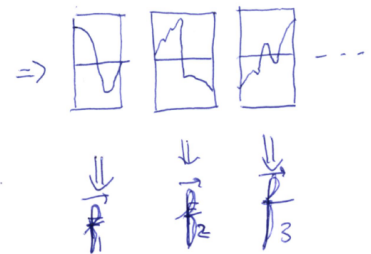
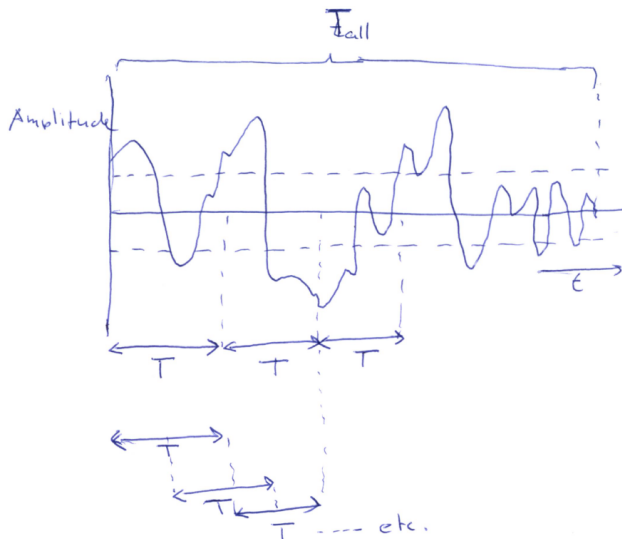
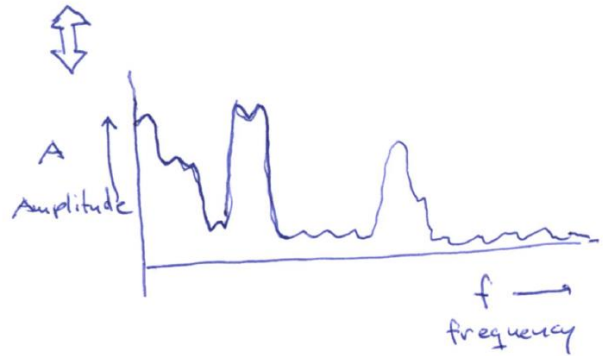
Digitizing Sound

Time Domain

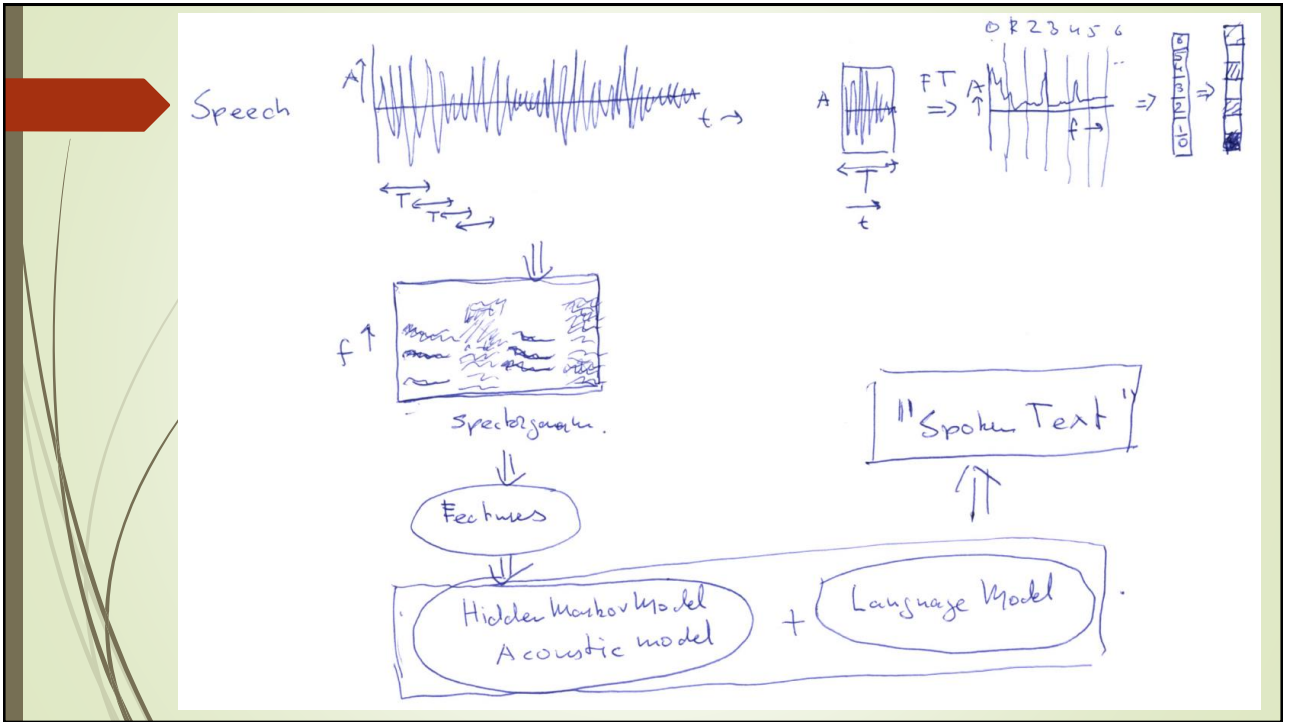


Fourier
Transform

Frequency
Domain



a series of feature vectors.




Applications

- ASR
- TTS
- Music Recognition
- Genre Recognition
- Sound Recognition
- Bird Recognition
- Sound location
- Language Recognition
- Timing
- Speaker Recognition
- Voice Cloning
- Singing
- Emotion Recognition
- Chord Recognition
- Music Transcription
- Cover Song Detection
- Drum Transcription
- Neural Audio Synthesis
- User Interfaces for Music Discovery
- Style Transfer
- Etc. Etc.



<https://dblp.org/db/conf/index.html>

- ▶ ISMIR <https://dblp.org/db/conf/ismir/index.html>
 - ▶ Proceedings: <https://www.ismir.net/conferences/>
- ▶ Interspeech <https://dblp.org/search?q=interspeech>
 - ▶ Proceedings: <https://www.isca-speech.org/archive/>
- ▶ Eurasip <https://dblp.org/db/journals/ejasmp/index.html>



Presentation Examples

- ▶ Chunyan Ji, et al., A review of **infant cry analysis** and classification. **EURASIP J. Audio Speech Music. Process.** **2021**.
- ▶ S. Kahl et al., BirdNET: A deep learning solution for **avian diversity monitoring**. Ecological Informatics, 2021.
- ▶ Y. Xiang et al., A speech enhancement algorithm based on **a non-negative HMM** and Kullback-Leibler divergence. EURASIP J. Audio Speech Music. Process., 2022.
- ▶ H. Ankishan, **Blood pressure prediction** from speech recordings. Biomedical Signal Processing and Control, 2020.
- ▶ S. Hantke et al., **WHAT IS MY DOG TRYING TO TELL ME? THE AUTOMATIC RECOGNITION OF THE CONTEXT AND PERCEIVED EMOTION OF DOG BARKS.** **Xxxx.**
- ▶ J. Lee, The Emotion is Not One-Hot Encoding: Learning with Grayscale Label for **Emotion Recognition** in Conversation. **Interspeech 2022.**
- ▶ J. Gao et al., Black-box **adversarial attacks** through speech distortion for speech **emotion recognition**. EURASIP J. Audio Speech Music. Process., 2022.
- ▶ K. Choi et al., LISTEN, READ AND IDENTIFY: **MULTIMODAL SINGING LANGUAGE IDENTIFICATION OF MUSIC.** **ISMIR2021.**
- ▶ Etc.

Previous Projects I

- Second Voice Generation
- Robustness of Musical Genre Identification
- Improved Mobile Song Recognition
- An iOS App using Bliss for Improved Communication through Text-To-Speech
- ScoreAid
- Emotion Recognition
- Instrument Detection
- Musical Instrument Recognizer (Annotation)
- **Audio Feature Extraction with Deep Belief Networks**
- Audio Morphing
- Audio Indexing the 1.000.000 song data set
- Chord Recognition
- Audio Phantom Materialization
- Harmonic Model Based Audio Transformations
- Content-Based Music Similarity, Visualization and Automatic Play-List Generation.
- Indexing and Predicting Bands from Unknown Songs
- Interpolation between Different Instruments
- Modular Synthesizer
- Hit Predictor
- **Pitch Perfector**
- Inter-Voice Morphing

Previous Projects II

- A Steglbiza implementation using traditional digital signal processing techniques: Steganography in music through tempo modulation
- Transfer Learning limited edition sounds (VGG16)
- **Tempo Extraction From Electroencephalography Using a Single EEG Channel**
- Monophonic Music Generation with LSTM Recurrent Neural Networks
- Blind speech signal separation from stereo sound input
- Gaussian Process Audio Model For **Audio Textures Modelling And Synthesis**

Previous Projects III

- A Deep Learning Approach to Instrument Detection and Chord Estimation via Frequency-based Feature Extraction
- **Open-Air Acoustic Delay-line Memory using a Micro-controller**
- Midioke
- Towards Real Time Audio Mosaicing
- Artist Recognition with Convolutional Recurrent Neural Networks
- **Automatically Identifying and Fixing Single Channel Audio Defects in Stereo Audio**
- **Information transmission by multilevel pitched audio**
- Event Detection using Wavelet Packets.
- Dynamic Play-list Generation.
- Automatic Song Remixer.
- **Isolating Voices in Sacred Harp Singing.**
- Voicetrack: A geolocation- and voice authentication-based password recovery mechanism.
- Song year prediction based on dynamic range values.
- Quotes to Characters.
- Arduino Peripheral Sound Imitation System.

API 2022 Fall Projects (Selection)

- **Automatic Sheet Music Generation**
 - <https://api-2022-final.vercel.app/>
- **Investigating Wav2Vec 2.0 language adaptability**
 - <https://github.com/fjatelnicki/audio-project>
- **A Tremolo Plugin**
 - <https://github.com/fRedelaar/API-project2022-tremolo>
- **Giving Agents a Voice**
 - https://github.com/jumorisse/Giving_Agents_a_Voice-Emergence_of_Language_in_Auditory_Multi-Agent_Games
- **Emotion within Speech**
 - <https://github.com/RylinnM/Speech-Emotion-Recognition>
- **Snake Jazz**
 - <https://github.com/MTVDN/API-Project>
- **Voice Cloning and Translation**
 - <https://drive.google.com/drive/folders/11CP1Y3v-FleoNunYTWRa6YQsvyDG2a0x?usp=sharing>

API 2022 Spring Projects (Selection)

- **Music Analyzer - Information Extraction and Visualization**
 - <https://github.com/arvindeva/API-final-project>
 - <https://api-final-project.vercel.app/>
- Exploring the Limits of DDSP
 - https://github.com/skylerf1/API_Final/
- Generative Adversarial Networks for Audio Generation: A Comparative Study
 - <https://github.com/ernestvmo/API-project>
- Mapping LPC predictor coefficients to reflection coefficients of an N-tube 2-Dimensional waveguide filter in real-time
 - <https://tinyurl.com/3hebyhx8>
- Music Genre Classification and Noise Separation
 - https://github.com/BingxuanW/api_project/tree/master
- Audio Processing and Indexing Musical Instrument Identification
 - <https://github.com/thomasmaliappis/api>
- **MUSICNN FOR DMX LIGHTING SOFTWARE AUTOMATION**
 - <https://github.com/poldemort/Musicnn-for-DMX>

Overview

- 5-9 Organization and Introduction
- 12-9 Audio Production and Processing + Vocal Tract Workshop@Home
- 17-9 ADC and an Algebraic Introduction to FT
- 24-9 FFT + FFT Workshop@Home
- 3-10 No Class: Leidens Ontzet
- 10-10 Project Proposals (presentations by students) + Student Paper Selection
- 17-10 Audio Features and Data Sets + Audio Features Workshop@Home
- 24-10 Machine Learning + Machine Learning Workshop@Home
- 31-10 Student Paper Presentations I.
- 7-11 Student Paper Presentations II.
- 14-11 Student Paper Presentations III.
- 21-11 Student Paper Presentations IV.
- 28-11 TBA
- 5-12 **Final Project Presentations Demo**
- 19-12 Project Deliverables:
 - Final Technical Project
 - Scientific/technical paper (4-8 pages), code
 - Web Site (or github)

Online Project Progress Meetings

Grading (6 ECTS):
 Presentations (20%) and Project (40%) (60% of grade).
 Class discussions,
 attendance, and workshops (40% of grade).

See: <http://liacs.leidenuniv.nl/~bakkerem2/api/>