SGN-14007
Introduction to Audio Processing

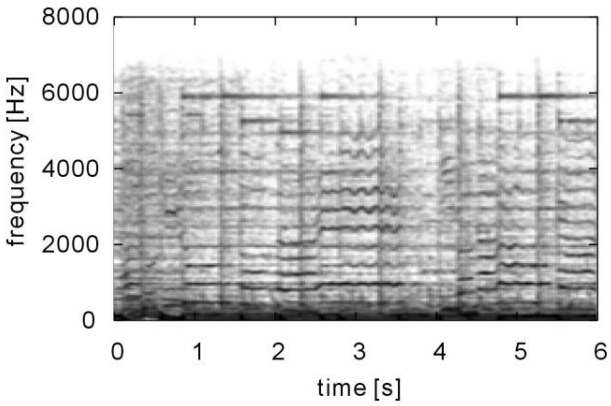
Project Work Introduction
Spring 2019

Practical arrangements

- Two topics (music signal processing and audio coding)
  - Implementing an audio signal processing algorithm
  - Done in two-person groups (doing the project alone is OK too)
  - Marked as Pass / Fail
  - Required for passing the course
  - Select one of the two given topics (or propose your own topic to us (we will need to approve it)
- Return the script and report by Friday 1 March 2019 (midnight) to Moodle (a project work page will appear)

Topic 1: Separation of drums from music signals

- Separation of harmonic and percussive components from music spectrogram (time-frequency domain)
  - Anisotropy: dependence on direction
  - Temporal continuity of harmonic sounds
  - Temporal localization of percussive sounds (continuity along frequency)



- Iterative search for such spectrograms that maximize the anisotropy (continuity over time vs. over frequency) [1]

- Together, the two spectrograms (harmonic / percussive) equal the original spectrogram: $H_{h,i} + P_{h,i} = W_{h,i}$

Ono, Nobutaka, et al. "Separation of a monaural audio signal into harmonic/percussive components by complementary diffusion on spectrogram." *Signal Processing Conference, 2008 16th European*. IEEE, 2008.

Topic 1: Separation of drums from music signals

- Hints to the implementation
  - Equations 24-30 of reference [1] describe the algorithm implementation, you do not need to care about the derivation of the iterative algorithm
  - Framewise processing (short-time Fourier transform) has been discussed at the lectures and at the exercises
  - Test material: drum and harmonic instruments (+ singing) separately, from which the total signal is obtained by mixing (drums + the rest)
    - Note that the algorithm works only for mono signals. If you want to process stereo signals, you may process both channels separately.
    - You may use audio material of your own or signals linked to the course web pages
  - Evaluation: signal-to-noise ratio
    - *s(t)* = original
    - *e(t)* = original minus separated

$$\mathrm{SNR} = 10 \log_{10} \left(\frac{\sum_t s(t)^2}{\sum_t e(t)^2} \right)$$

- Report
  - What kind of audio material is the algorithm limited to and why?
  - How should the separation quality be measured and assessed?

Topic 1: Instructions

1. Reading the scientific paper
   - N. Ono, K. Miyamoto, J. L. Roux, H. Kameoka and S. Sagayama, “Separation of a monaural audio signal into harmonic/percussive components by complementary diffusion on spectrogram,” in Proc. EUSIPCO, 2008
   - Equations 24-30 describe the algorithm. Feel free to skip the derivation of the algorithm and do not get scared by the math!

1. Implementing the algorithm described in the paper using Python

1. Evaluation and interpretation of the results (using some test data)

1. Writing a report about the results and your observations

Topic 2: Implementing a Psychoacoustic Model

- Exploit a uniform filterbank with critical sampling and perfect reconstruction along with a simplified psychoacoustic model to produce the masking threshold for each time-frequency point.
- Learn to create a pipeline for encoding, quantizing and decoding an audio signal

Topic 2: Instructions

1. Create an MDCT analysis-synthesis filterbank

1. Implement a psychoacoustic model, and compute masking thresholds from SPL levels of the signal

1. Quantize the sub-band time-domain signals

1. Implement a decoder and analyze the results

1. Write a report about the results and your observations
2. More detailed instructions can be found from Moodle

Returns

1. Python Script
   - Project_x.py
   - Return only a part of the test material to check the functionality of the algorithm (even one signal suffices) even though the evaluation results should be computed using some more data
2. Report
   - What problem is being solved in the project work?
   - How is the problem solved?
   - What assumptions were made?
   - Short description of the implementation. What stages does the algorithm consist of?
   - What is evaluated in your experiments, and how?
   - For project 2. include figures specified in the instruction pdf
   - Mention in the report how the work was divided in your group
   - Length <= 4 pages (depending on the figures)

General

- Questions regarding the project work in Moodle to course assistants
- We make an effort to read and answer the questions daily
- If the questions are related to Python code, the code has to be very well commented!
- We will try to arrange a support session for answering questions and advising related to the project work. More information about that on the course web page.
  - You may also ask questions at the end of the weekly exercise sessions