P05 – Multimedia analysis

1. Software requirements

Install the “librosa” Python library for basic audio/music processing:

- On a command shell, type: **`pip install librosa`**
- On Windows: If an error message asks you to download and install the Microsoft Visual C++ compiler for Python 2.7, follow the link in the message and so

2. Speaker identification

2.1. Make it run

Open the Python script **`speaker-id.py`** and make it run. If you installed Anaconda and **`librosa`** correctly and the train/test folders lies in the same directory as the code, it should work without any adaptations.

(The script performs speaker recognition on a tiny corpus of 5 speakers by extracting MFCC features and building a GMM model for each training file as discussed in the lecture; then performing maximum likelihood classification of each test utterance.)

2.2. Make yourself comfortable with the code

Work through the code and get a feeling for its workings. Do you recognize that most real work is delegated to the respective libraries, while the script merely contains “glue code”? This is the reason Python is so well respected in the data analytics community.

2.3. Perform a speaker recognition experiment

The script is a starter for a classical speaker identification experiment after Reynolds et al. (as discussed in the lecture). Complete the code and perform the experiment. How would you e.g. compute recall and precision per speaker / for the whole data set? Can you implement that into the script?

2.4. Extend it

Can you enrich this experiment with additional speakers from other audio sources? Can you for example also identify the sound of different bird voices (theoretical/practical → see e.g. http://www.imageclef.org/lifeclef/2017/bird)? Can you use it for musical genre recognition, or Shazam-like audio fingerprinting? What additional methods would you need? Where could you learn about them? Is that still AI – or are methods with roots in AI merely one amongst several components?

Build a prototype for at least one of the suggestions above (birds, music, own idea).