

Audio and Music Processing

Spezielle Kapitel aus Informatik: Audio and Music Processing
344.023
KV, 2std., SS 2013

Exercise Track

The Framework

- the framework takes 5 parameters:
 - -i WAVFILENAME
 - -d OUTPUTDIR
 - -o ONSETGROUNDTRUTHFILE (optional)
 - -t TEMPOGROUNDTRUTHFILE (optional)
 - -b BEATGROUNDTRUTHFILE (optional)
- the output (in the output directory) consists of 6 files:
 - WAVFILENAME.onsets (containing estimated onset times)
 - WAVFILENAME.bpms (containing the estimated tempo)
 - WAVFILENAME.beats (containing the estimated tempo)
 - and .eval files for each task (only if you specified the groundtruth)
- you can load the onset and the beat data into the Sonic Visualizer!
 - first load the .wav file
 - then the estimated onset times (File → Import Annotation Layer)

Preliminaries

- I prepared a little Framework in Java
 - you will soon be able to download the framework via KUSSS
 - the framework provides simple ways to
 - read audio files (.wav only)
 - compute the STFT (giving you the energy, the phase and the unwrapped phase of each bin for each frame)
 - evaluate the output of your system (given the ground truth)
 - you are allowed to change EVERYTHING in this framework
 - EXCEPT the handling of the input parameters, the evaluation function and the output routine for the results
 - please document the changes you made in the description of your work (see later)
- You will also receive a set of test data incl. ground truth
- And: Do not hesitate to contact me if you have questions, problems with the framework, ...!

Goals of the Exercise

- Implement at least 2 onset detection methods
 - i.e., implement at least 2 detection functions
 - you are allowed to use the same pre-/post processing/peak picking method for both functions
- Implement at least 1 tempo extraction method
- Implement at least 1 beat detection method
- You have to decide on a final combination of onset detection / tempo extraction / beat detection methods
 - but you may also come up with ideas of how to combine multiple onset detection / tempo extraction methods into one system
 - your tempo extraction may also be based on other features than the detected onsets
 - this final system will be used for the evaluation

Test Data

- You are provided with a data set
 - 18 excerpts (info: files train12 and train17 are missing)
 - length 30 seconds
 - includes various genres
 - in general: you can assume nearly (!) constant tempo over these 30 seconds
 - still it might improve the results if you allow for slight tempo variations...
- For each excerpt you are also provided with 3 text files
 - NAME.onsets is a list of all onset times in the excerpt
 - NAME.bpms contains the tempo of the excerpt
 - in fact this is manually compiled data
 - humans are rarely in agreement of the tempo of a piece of music
 - these files include 2 tempi and a weighting
 - e.g. 120 60 0.8 means that 80% of the people say that the tempo is 120 bpm and 20% of the people think that the tempo is 60 bpm
 - for our evaluation we only use the tempo which received the majority of votes
 - NAME.beats is a list of all beat locations in the excerpt
 - The second column contains beat labels (metrical level) and can be ignored

Submission

- Deadline: 23.6.2013 (submission via email)
- You have to submit
 - your program
 - source code and binary (in one .jar file)
 - I have to be able to run your program with “java -jar JARNAME.jar”
 - a description of your program and the experiments you did
 - 2-3 pages
 - a description of which methods you chose, how you modified/combined them and how you came up with your final system
 - especially: describe the experiments you did, things you tried, ideas you had, ...
 - it is more important for me to see that you experimented with different methods, tried many different things, tried to understand the difficulties, ... than the performance of your final system
 - curiosity and creativity will be rewarded!

Evaluation

- The Framework includes an evaluation procedure
- Onset detection:
 - every onset estimate which is within +/- 50 ms of an actual onset is counted as a true positive (O_{CD})
 - for each actual onset only 1 onset estimate is allowed; others are counted as FP (O_{FP})
 - every onset estimate outside of this window is counted as a FP (O_{FP})
 - every onset for which no onset estimate is inside this window is counted as a FN (O_{FN})
 - Important measure here: F-measure
- Tempo Estimation
 - if the tempo estimate is within +/- 4% within the actual tempo this is counted as correct
 - there is also a second measure: the tempo estimate is compared to double/half/third/3 times the actual tempo (using the same window)
- Beat Detection
 - As for onset detection but with a window of +/- 70 ms

Discussion of your Submission on the 26.6.2013

- Attendance is compulsory!
- I will talk about
 - the different approaches you implemented
 - problems and pleasant surprises with/about your implementations
 - the evaluation of your systems on an independent test set
- And we will discuss your experiences with
 - the exercise track
 - and the lecture in general