Lab 5: Music Classification

1. Introduction

The goal of this lab is to build a music classifier. We will use standard classification algorithms and the audio features we have studied in the course.

2. Steps

- Taxonomy definition: we will decide on a specific problem related to classification (genre, mood, singer, artist, voice vs instrumental, male vs female, country, production style, etc)
- Annotated dataset: we will collectively build a music collection for the selected taxonomy.
- Feature extraction: we will extract descriptors related to timbre, rhythm melody, harmony and spatial location using existing toolboxes and frameworks.
- Automatic classification: we will use standard classification algorithms implemented under the WEKA software.

3. Description

Step 1: **Taxonomy definition**

Step 2: **Building a dataset**

Task: You have to organize the collection of music files by applying the highest quality checking standards. You have to:

0) contribute with files from your own collection according to the taxonomy
1) check for appropriateness of files with respect to each of the labels
2) check for duplicates, maybe using different encoding parameters. Even though 128kbps will be ok, it's better to keep the highest quality ones. Use stereo mp3 files.
3) ensure the right metadata is included for all the files: song title, artist, album title. Generate a .cvs file for each file including metadata in the following order. Example:

It's so hard, John Lennon, Imagine

4) the structure of the files will be one folder per label.

Step 3: **Feature extraction.**

Task: Extract any feature that you can from the final collection:

1) Install the assigned toolbox: MIRToolbox, Essentia, MARSYAS, Sonic Annotator (Sonic Visualizer), jAudio.
2) Analyze the collection.

emilia.gomez@upf.edu

3) Generate a cvs file with the extracted descriptors, where the first line will indicate the name of the descriptors and each line will correspond to a computed descriptor. Example:

filename, descriptor1, descriptor2,… ,descriptorN, toolbox, 'genre',

1.wav, 0, 0.5, 25, MIRToolbox, 'pop'

Expected result: cvs file with computed descriptors

Step 4: **Automatic classification**.

Task: Build a classifier in Weka for music classification.

1) Install Weka. Perform automatic classification with the features you have extracted.

2) Merge your feature values with the ones provided by other people. Perform automatic classification.

3) Evaluate the final accuracy.

Expected result: automatic music classifier.