Automatic Musical Genre Classification Of Audio Signals

- **Features extraction process** : “The means and standard deviations of these features are calculated over a “texture” window of 1 second consisting of 40 “analysis” windows of 20 milliseconds (512 samples at 22050 sampling rate). The feature calculation is based on the Short Time Fourier Transform (STFT). that can be efficiently calculated using the Fast Fourier Transform (FFT) algorithm”

- **Features extracted :**
  [Timbral features] Spectral Centroid, Spectral Roll-off, Spectral Flux, Zero Crossing, Low Energy
  [Rhythm] Beat Histogram (steps to achieve this histogram : Full Wave Rectification, Low Pass Filtering, Downsampling by a 16 factor, Normalization, Autocorrelation), from which are extracted the periods (in bpm) of the highest 3 peaks and their amplitudes (relative to the sum of all amplitudes on the histogram). More info on the Beat Histogram in summary of “Comparative study on content-based…”

- **Genres :** 6 musical genres (one of which, classical, is subdivided into another 4 subgenres), and 3 speech genres

- **Dataset :** collected from radio, CDs and the web. One of the biggest issues is that pieces from an album were all labeled with the same genre, which is not really realistic (ex : in Sting’s album, there are Jazz songs as well as Rock songs)

- **Evaluation :** through 10-fold cross validation, with 4 different tests : 1) differentiating music from speech **2) classifying between the 6 music genres** 3) classifying between the 3 speech genre 4) classifying between the 4 classic music subgenres

  2) is what interests us the most : 62% accuracy using Gaussian classifier (similar to Gaussian Mixture Model but with only 1 Gaussian per class)

- Conclusion : need to expand the genre hierarchy, both in width and **DEPTH** (which has actually not been done by the subsequent researchers)