ECEN 4/5532: Digital Signal Processing Lab

Lecture Notes: Lab 2

Instructor: Prof. Farhad Pourkamali-Anaraki

University of Colorado at Boulder

Spring 2016

College of Engineering & Applied Science

UNIVERSITY OF COLORADO BOULDER

Lab 2 Overview

- Digital signal processing methods for **content-based** music information retrieval.

- We will explore **higher** level musical features:
  1. Rhythm
  2. Tonality and Chroma

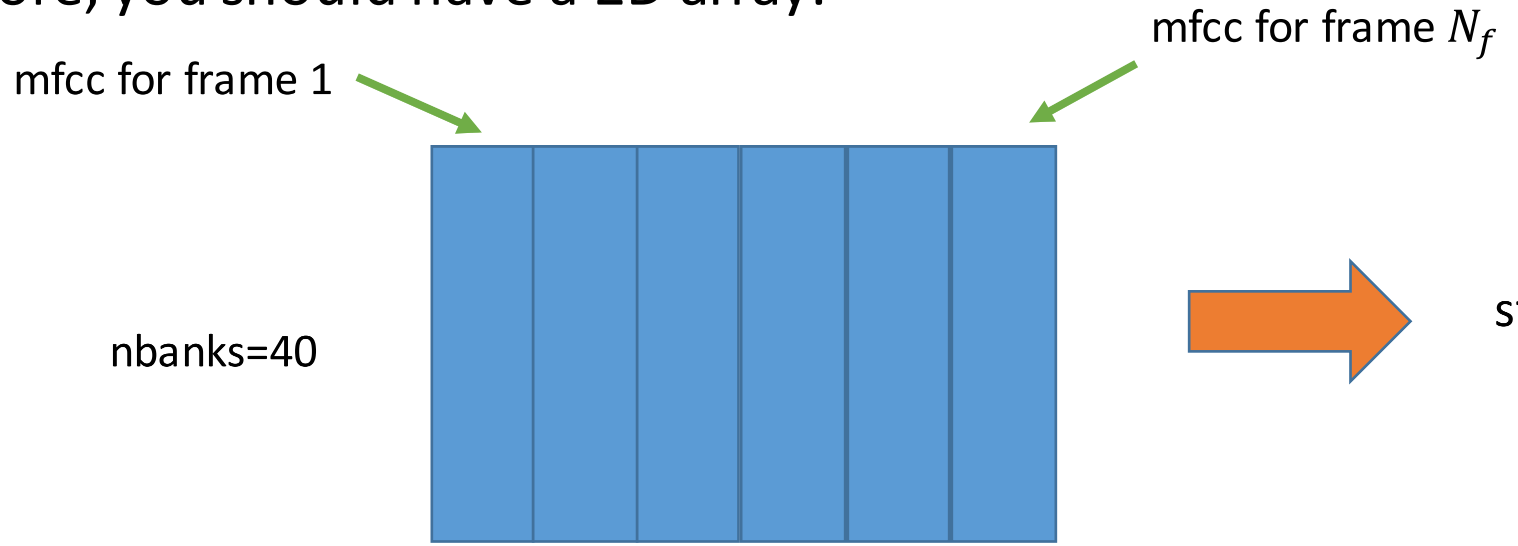
- In lab 3, we are going to **classify** tracks according to the **features** discussed:

Spectral decomposition

- The mfcc measured in **decibels** (dB) is defined by:

$$10\log_{10}(\text{mfcc})$$

- Therefore, you should have a 2D array:



store this matrix of size
$nbanks \times N_f$

Background

- Given two vectors of attributes or features:

$$\cos(\theta) = \frac{\langle a, b \rangle}{\|a\| \|b\|} = \frac{\sum_{i=1}^{n} a_i b_i}{\sqrt{\sum_{i=1}^{n} a_i^2} \sqrt{\sum_{i=1}^{n} b_i^2}}$$

measure of similarity

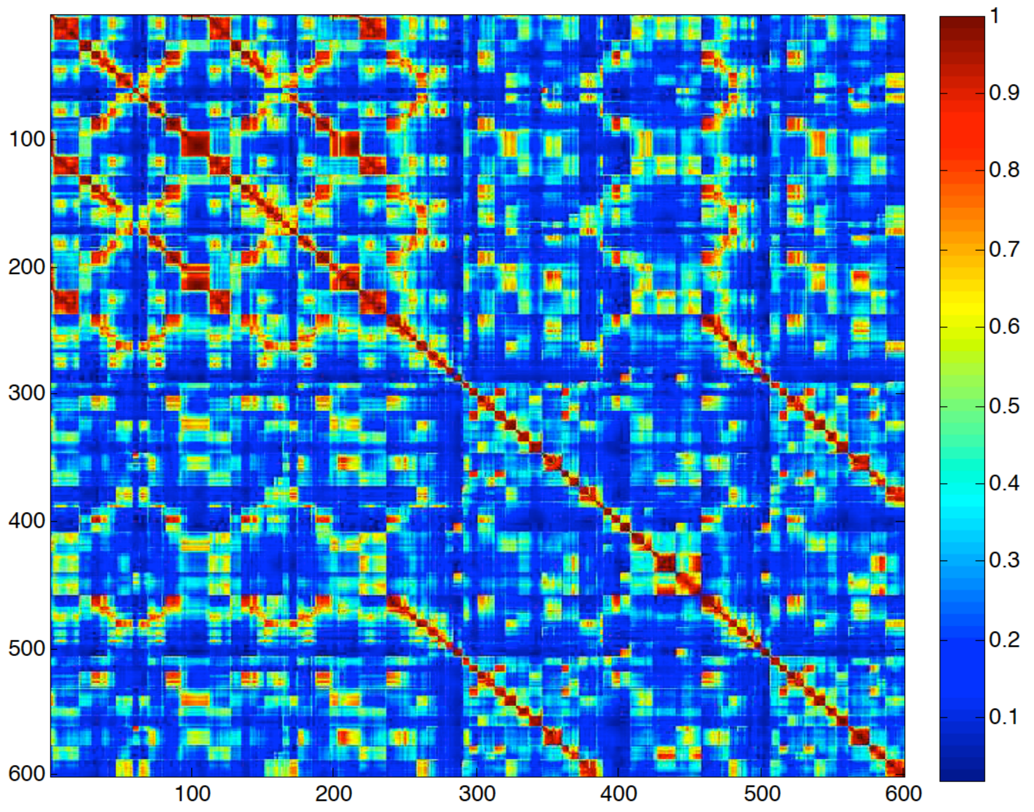
- The resulting similarity ranges from −1 meaning exactly opposite, to 1 meaning exactly the same, with 0 indicating orthogonality, and in-between values indicating intermediate similarity or dissimilarity.

Similarity Matrix

- $S(i,j)$ measures the **cosine** of the angle between the spectral signatures of frames i and j:

$$\mathrm{S}(\mathrm{i},\mathrm{j}) = \frac{\langle \mathtt{mfcc}(:,\mathrm{i}),\ \mathtt{mfcc}(:,\mathrm{j})\rangle}{\|\mathtt{mfcc}(:,\mathrm{i})\|\|\ \mathtt{mfcc}(:,\mathrm{j})\|} = \sum_{\mathrm{k}=1}^{\mathrm{nbanks}} \frac{\mathtt{mfcc}(\mathrm{k},\mathrm{i})\ \mathtt{mfcc}(\mathrm{k},\mathrm{j})}{\|\mathtt{mfcc}(:,\mathrm{i})\|\|\ \mathtt{mfcc}(:,\mathrm{j})\|}.$$

Example of a Similarity Matrix



Rhythm

- The presence of **repetitive patterns** in the temporal structure of music.

- We will compute a vector B such that $B(l)$ quantifies the presence of similar spectral patterns for **frames** that are l frames apart.

- The lag associated with the **largest** entry in the array B is a good candidate for the period in the rhythmic structure.

A First Estimate of the Rhythm

$$B(l) = \frac{1}{N_f - l} \sum_{n=1}^{N_f - l} S(n, n + l), \quad l = 0, \ldots, N_f - 1$$

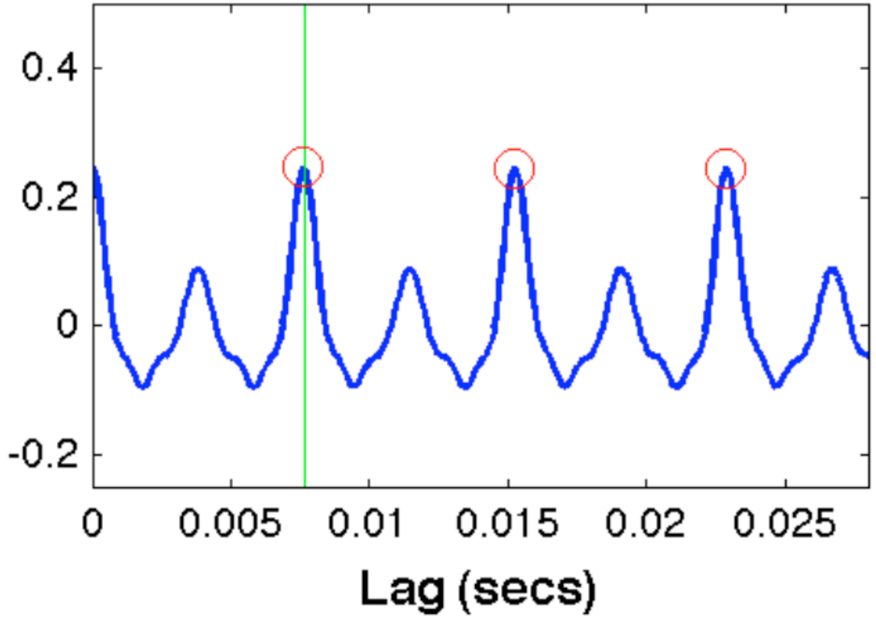
entries on the l-th
upper diagonal

A Better Estimate of the Rhythm

- Basic Idea:
  - A lag l will be a good candidate for a rhythmic period, if there are many i and j such that if $S(i, j)$ is large then $S(i, j + l)$ is also large.

$$AR(l) = \frac{1}{N_f(N_f - l)} \sum_{i=1}^{N_f} \sum_{j=1}^{N_f - l} S(i, j)S(i, j + l), \quad l = 0, \ldots, N_f - 1.$$

A Better Estimate of the Rhythm



$$l \times \frac{K}{f_s}$$

Figure 2: Rhythm index $AR(l)$as a function of the lag l

Rhythmic Variations Over Time

- Consider short time **windows** formed by 20 frames:

$$\mathrm{AR}(l, m) = \frac{1}{20(20-l)} \sum_{i=1}^{20} \sum_{j=1}^{20-l} S(i + m*20, j + m*20) S(i + m*20, j + m*20 + l),$$

$$\text{for } l = 0, \ldots, 19, \quad \text{and } m = 0, \ldots, N_f/20 - 1.$$

Norm and Inner Product in MATLAB

```
>> v1 = [1;2;3]

v1 =

     1
     2
     3

>> v2 = [3;4;5]

v2 =

     3
     4
     5

>> v1'*v2

ans =

    26

>> dot(v1,v2)

ans =

    26
```

Norm and Inner Product in MATLAB

```
>> v1 = [1;2;3]

v1 =

     1
     2
     3

>> norm(v1)

ans =

    3.7417

>> sqrt(v1'*v1)

ans =

    3.7417
```