Signal and Communication Lab
DII University of Brescia - ITALY

Digital Audio Processing

Lab.5 - PCP and Multirate Filter banks - May 18, 2014

Alessio Degani, [alessio.degani@ing.unibs.it]
http://www.ing.unibs.it/alessio.degani

Ref. Book: Information Retrieval for Music and Motion, Chap. 3, *Meinard Müller*, Springer

1 Pitch Class Profile (Chroma features)

The Chroma features (Pitch Class Profile) is an audio descriptor that is suitable for chord/key estimation or music similarity measures.
The PCP can be calculated with two simple steps:

- Subband decomposition of music signal. Each band represents a semitone in a equal-tempered chromatic scale.
- Computing the short-time power for each subband (short-time=sliding window).
- Sum the coefficient of the same pitch class.

Open the wave file *lab5.wav* and compute the PCP.

1.1 Filter bank - Subband decomposition

For the subband decomposition of the musical signal, we use a bank of tuned constant-Q band-pass filter. The center frequency of each filter is set to the frequency of the musical notes with a bandwidth of half semitone.
We use 88 band-pass filter, from note $A0$ to $C8$ corresponding to the keys of a standard piano. Using the **MIDI** notation, we take the pitches in the range $p = [21 \dots 108]$. The central frequency of the pass-band filter can be calculated as follow:

$$f(p) = f_{ref} \cdot 2^{\left(\frac{p-69}{12}\right)} \tag{1}$$

In this case we assume $f_{ref} = 440$ Hz. Hence the filter bank is constant-Q, we use $Q = 25$ for all filters. That means that for a filter with center frequency f_i, the corner frequencies are calculated as $f_{left} = f_i - \frac{f_i}{2Q}$ and $f_{right} = f_i + \frac{f_i}{2Q}$. For speed reason, we use a multirate filter bank. For the high frequencies filters ($p \in [93 : 108]$) we use the musical signal at his sampling frequency, for the mid frequencies ($p \in [57 : 92]$) we use the musical signal downsampled by a factor of 5 and for the low frequencies ($p \in [21 : 56]$) we use the musical signal downsampled by a factor of 25. For example, if we have a 22050 Hz sampled audio file, we have 3 different sample rate version of the same signal at frequencies 22050 Hz, 4410 Hz and 882 Hz.

1.1.1 Downsampling

The operation of downsampling by a factor D is made by two steps:

- Anti-alias low pass filtering at normalized frequency $\omega_{AA} = \frac{1}{2D}$
- Signal decimation of factor D

Use the Matlab command *filter* for filtering and design a $N = 1000$ samples FIR anti-alias filter (LP) using the command *fir1*.

> **NOTE: 1. The normalized frequency $\omega = [0 : 1]$ in matlab correspond to the range of analog frequency $f = [0 : \frac{f_s}{2}]$**
> **2. FIR filter has only the feed-forward polynomial $B(z)$, so $A(z) = 1$!**

1.1.2 Multirate filter bank

Design all the 88 filters as elliptic 8^{th} order band-pass with $Q = 25$, passband ripple $PB_r = 1$ dB and stopband attenuation $SB_a = 50$ dB. For elliptic filter design in Matlab, see the command *ellip*. For filtering use the **zero-phase** forward-backward filter command *filtfilt* in order to obtain no phase distortion.

> **NOTE: Pay attention to the normalized frequency ω that change with different sample rate!**

1.2 Decimation and Energy summation - Short-Time Power

In order to get the PCP representation you have to obtain the same length output for each sample rate. In order to obtain this you have to decimate the signals according to their sample rate. You have to window the signal and sum the square value of the samples within a sliding window of length w_l and hop w_h depending on sample rate:

- for 22050 Hz: $w_l = 501$, $w_h = 250$
- for 4410 Hz: $w_l = 101$, $w_h = 50$
- for 882 Hz: $w_l = 21$, $w_h = 10$

For each windowed samples $w[n]$ compute the squared sum $a \cdot \sum_n w[n]$ where a is a energy correction constant $a = 1$ for 22050 Hz sampled signals, $a = 5$ and $a = 25$ for the 4410 Hz and 882 Hz sampled signals.

1.3 Pitch Class Profile

In order to obtain the 12-dimensional PCP representation, you have to sum all of the energy coefficient of the same pitch class. For example, the PCP coefficient for the pitch class C is calculated by summing the energy of the coefficient from $C1$ to $C8$ ($C0$ is not considered in our implementation since our filter bank starts from $A0$). Normalize each PCP vector in order to have the maximum value equal to 1.