Recognition of Musical Tones

Digital Signal Processing, 2021/2022

Introduction

A melody is a sequence of tones produced by a musical instrument, each tone being characterized by a pitch and by a duration. Pitch is an auditory sensation and it is, in general, a subjective attribute of sound. However, if the musical instrument produces quasi-periodic sounds, then pitch corresponds to the fundamental frequency of vibration and can be measured in an objective way. Duration can be defined in a simpler way since it stands for the time-interval in which the quasi-periodic sound is produced.

In Western music, pitch cannot take arbitrary values since it is quantized in a logarithmic scale. The set of admissible pitch values can be defined as follows. First, the frequency range is divided in octaves starting at a reference frequency A (typically A=440 Hz). Then, each octave is divided in 12 intervals equally spaced in a logarithmic scale. These intervals are denoted as A, A#, B, C, C#, D, D#, E, F, F#, G, G#. The difference between consecutive notes is denoted as a semitone and the frequency of a musical note is obtained by multiplying the frequency of the previous note by $\sqrt[12]{2}$. The symbol # increases the pitch of the tone by a semitone.

Tones can also be represented by an integer number i, and the frequency associated is given by

$$f_i = 2^{\frac{i}{12}} f_0$$

where f_i stands for the pitch of the $i-th$ tone and f_0 stands for the reference frequency.

Music can be graphically represented using a shaft representation with 5 horizontal lines (see Figure 1). This is a time-frequency representation. Each note is represented by a symbol that includes a circle. The vertical position of the circle in the shaft defines the pitch. The shape of the symbol characterizes the duration.

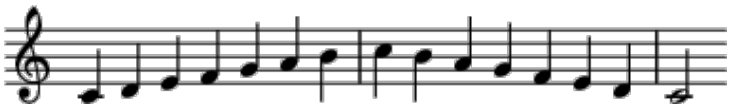

Figure 1 - C Major scale (from Wikipedia)

The objective of the project is to recognize the musical tones from a melody produced by a musical instrument. The work is split into three parts and should be described in a report.

Part 1 - Basic Music Analysis

Perform the following steps:

1. Visualize the signal stored in the file greensleeves.wav. The melody was produced by a keyboard and sampled at 44 KHz.
2. Observe the shape of the signal associated to the first 9 tones.
3. Discuss if the signal can be down sampled before pitch recognition.
4. Visualize the Fourier spectrum and the auto-correlation for each note.
5. Compute the time-frequency representation (spectrogram) of the signal.

Part 2 - Pitch Recognition in Segmented Sounds

1. Manually segment each note *e.g.,* by observing the signal in the time domain and its Fourier spectrum.
2. Build an algorithm to automatically recognize the pitch in each segment.
3. Evaluate the performance of the recognition algorithm.

Part 3 - Pitch Recognition and Sound Segmentation

1. Build an algorithm to simultaneously segment the tones and recognize the pitch of each tone.
2. Evaluate the performance of the tone segmentation and recognition algorithm for the segment corresponding to the first 9 notes and then for the complete signal. Suggestion: synthesize a signal using only the fundamental frequency for each note.

Report

Each group should write a report describing the algorithms that were developed and the achieved results.

The report should be written in Portuguese if both students are Portuguese. Otherwise, the report should be written in English.

The maximum number of pages is 10 pages and the report should be written with a font size of 12 pt.