ALGORITHM FOR IDENTIFYING THE NOTES*

xizheng ma

This work is produced by The Connexions Project and licensed under the
Creative Commons Attribution License †

Algorithm For Identifying the Notes

Step 1: Convert the music segment into frequency domain

Because our recorded music files are chopped into fixed finite length segments, all we have to do is apply DFT to the sample. Considering speed of computation, we decided to use the method of computing with the DFT matrix .

We know that DFT samples are defined by

,

Where

Thus can be expressed in matrix form as

Where

By using the matrix method to compute the DFT, we can calculate the DFT of many pieces of same length and same sample rate music segments at the same time using the same matrix. This will increase the overall speed of our program and enable tune identification for lengthy music that consists of thousands of segments to be analyzed.

Step 2: Find the peaks in the frequency domain representation of the segmented music. Identify the peaks whose frequencies are near integer multiples of each other and group them into the same category.

We know that notes played by instruments have harmonics which the frequencies are integer multiples of the base frequency. Often the note that's being played is the base frequency or the first odd harmonic of the base frequency. Thus, peaks which frequencies are near integer multiples of each other indicate that they are mostly likely to be the harmonics of a base frequency and the note that's being played lies within it.

To identify the harmonics we used findpeaks function in Matlab and grouped them in a matrix which the i-jth entry is the frequency of the jth peak over the frequency of the ith peak. We then look for near integer entries in the matrix to determine peaks whose frequencies have integer multiples relationship and categorize them together. The peak with the lowest frequency in each category is the base band while the other peaks in the same category are its harmonics.

Step3: Use the power of peaks to determine the most significant categories.

Because there may be noise and other random vibrations when an instrument is played, there are categories which do not contain the note that's played by the instrument. On the other hand, when multiple notes are played at the same time (aka a chord), there may be multiple categories which contain notes that are played. To correctly analyze the notes that are being played, we pick out the category which has the most power and categories whose powers are larger than one-tenth of that. The way we calculated the power of each category is by summing up the power(frequency*peak^2) of all the harmonics in that category. After

*Version 1.1: Dec 19, 2011 3:19 pm -0600

†http://creativecommons.org/licenses/by/3.0/

this process, ambient noise and vibrations should be excluded and the remaining categories should contain the notes that are being played.

Step 4: Determine the notes that lie in each category.

Mostly, the base frequency of each category is the note that's played by the instrument. However, we've learned that the note played by some instruments is either the base frequency or the first odd harmonic of the base frequency. To determine which note is being played, we compared the sum of the power of odd harmonics versus the sum of the power of the even harmonics in each category. When the power of even harmonics is larger than a certain times of the power of the odd harmonics, we say that the note is the base frequency; otherwise, we say that the note is the first odd harmonic of the base frequency.