Music Analysis

Josiah Boning

TJHSST Senior Research Project
Computer Systems Lab, 2007-2008

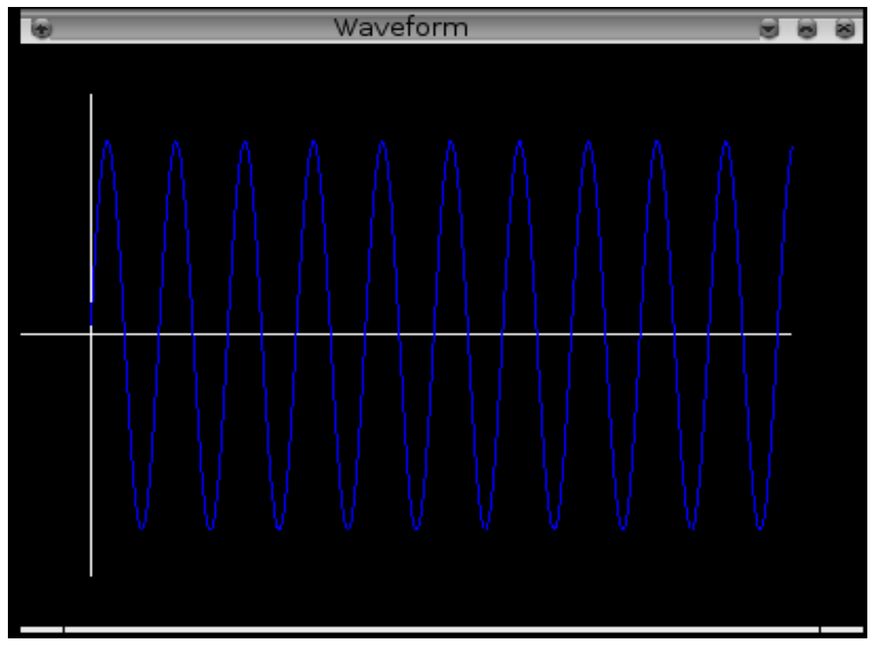
Abstract

Although music is one of the most universal aspects of human culture, it is very difficult to define. Most definitions of music have been dependent on attributes such as rhythm, melody, and harmony, which are extremely subjective, so the ability to identify music has been limited to humans. This project aims to better define “music” by applying machine learning techniques to music analysis and recognition, allowing computers to autonomously identify whether a given audio sample is musical in nature.

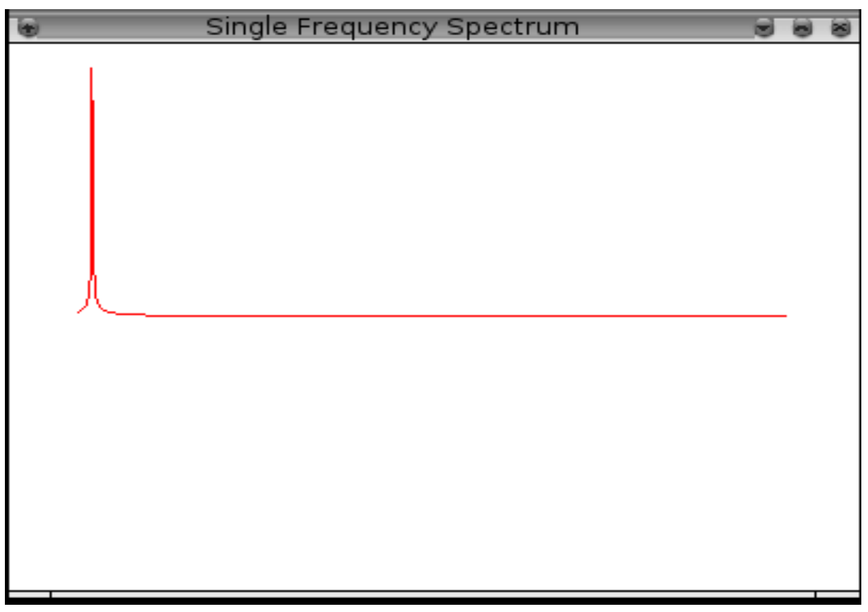
Background

Computers have already been used to perform analysis of music. Research has shown that different genres of music can be distinguished by fractal dimension and that machine learning techniques could successfully identify musical genres[2][1]. Other research has attempted to deconstruct music in terms of rhythmic and melodic patterns, and even looked at writing software to generate music conforming to such patterns[3]. Howeve, each instrument has a different sound quality, and composers write music with these timbral differences in mind. Simply analyzing the notes on sheet music precludes the use of these differences in the analysis. Audio recordings, in contrast, allow analysis of exactly what the composer intended his audience to hear.

Waveform



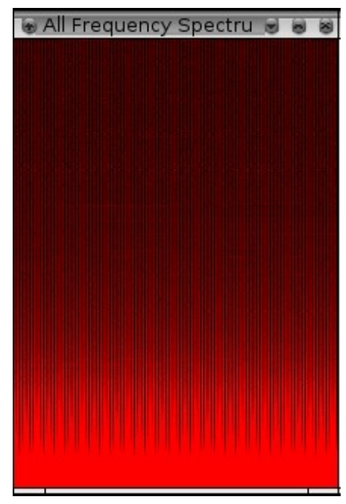
Single Fourier Transform



Fourier Transforms

A Fourier transform decomposes a wave into its constituent frequencies. A sine wave (left) has a Fourier transform with a spike at a single frequency (below left). The transforms can also be performed over several time windows (below).

Multiple Fourier Transforms



Fractal Dimension

The equations below are calculated by numeric methods to yield the fractal dimension of the audio waveform. Since different genres of music are distinguishable by their fractal dimensions, it is reasonable to suspect that music itself might be distinguishable by its fractal dimesion[2].

$$\lim_{\tau \to 0} \left(2 - \frac{\log\left(\frac{1}{b-a} \int_a^b \left| \max_{|x-t|<\tau}(f(t)) - \min_{|x-t|<\tau}(f(t)) \right| dx \right)}{\log \tau} \right)$$

$$\lim_{\tau \to 0} \left(2 - \frac{\log\left(\frac{1}{b-a} \int_{x=a}^{x=b} \left[\frac{1}{\tau^2} \int_{t_1=0}^{\tau} \int_{t_2=0}^{\tau} \left| f(x+t_1) - f(x-t_2)^{\alpha} \right| dt_1\, dt_2 \right]^{1/\alpha} dx \right)}{\log \tau} \right)$$

Works Cited

[1] Basili, Roberto, Alfredo Serafini, and Armando Stellato. 2004. ”Classification of Musical Genre: A Machine Learning Approach.” Presented at the 5th International Conference on Music Information Retrieval.

[2] Bigerelle, M., and A. Iost. 2000. ”Fractal Dimension and Classification of Music.” Chaos, Solitons & Fractals. 11(14):2179-92.

[3] Leach, Jeremy, and John Fitch. 1995. ”Nature, Music, and Algorithmic Composition.” Computer Music Journal. 19(2):22-23.