EECS 225D Lecture 27

Music Synthesis

Ketan Patel, kpatel@cs.berkeley.edu

1.0 Computers in Music

Prof. Gold separated the involvement of computers in music into 3 classes:

1. Direct Synthesis
2. Analysis-Based Synthesis
3. Musique Concrete

2.0 Direct Synthesis

Direct synthesis is the use of functional and physical models to create the source material for the music (i.e., composing).

Another direct synthesis application is the reproduction of the sounds created by acoustic instruments. Acoustic reproduction often lends insight into the nature of the particular instrument. In addition, it is often easier to create unique and original sounds by starting from a reproduction of a particular instrument.

3.0 Analysis-Based Synthesis

The nature of analysis-based synthesis is to start with an analysis of existing sounds or music, transforming the sound representation into a representation that extracts particular features. Then by modifying the results of the analysis and resynthesizing the sound, we can create a new sound.

3.1 Types of Analysis

What types of analysis are used? We can often apply many of the same techniques used in speech. Prof. Gold listed the following as possibilities for analysis:

1. LPC
2. Channel Vocoding
3. Homomorphic Vocoding
4. Sine-Transform Coding (same concept used in telharmonium)

At this point, Prof. Gold wanted to show that speech analysis techniques could be applied to music with various degrees of success depending on the bandwidth used. He played a tape of music that had been sent through a vocoder at different rates to illustrate the difference in quality given bandwidth.

3.2 Applications of Analysis-Based Synthesis

Possible applications of analysis-based synthesis include:

1. Enhancement
2. Bandwidth Compression
3. Transformations

3.2.1 Enhancement

One possible application is music enhancement. For example, if we are given a scratchy record, can we remove the scratches?

The concept is to analyze the sound of the record in a parametric way. If we are able to separate the parts of the various parameters that correspond to the noise we want to remove, we can modify the parameters and resynthesis the sound without the noise.

As an example, Prof. Gold stepped through the stages for noise removal followed by an existing system that worked for speech or single source music. The stages he presented were:

1. Original Source: the original audio waveform.
2. Spectrum Analysis: the waveform in frequency domain.
3. Log Spectrum Analysis: the log of step 2.
4. Cepstral: application of this filter separates the noise from the music.
5. Synthesis: resynthesis of the noise-free music.

Prof. Gold then showed a block diagram of the actual system that was built.

All of the analysis methods listed in the previous section except for the sine-transform basically performs a deconvolution that separates excitation and impulse response. Once separated, each can be dealt with separately (i.e. modified, transformed, etc.) and recombined to synthesis the new sound.

3.2.2 Transformation

Various transformations that can be applied to music include speeding up, slowing down, changing pitch, and frequency scaling.

The idea behind frequency scaling was to compress the spectral information of a sound into a smaller and lower range of frequencies in order to present to hearing impaired ears a similar spectral envelope shape in frequencies that could be heard better.

Prof. Gold played a tape of speech with various transformations applied. First the original was heard, followed by the sound sped up, slowed down, pitch changed (i.e. sex transform), and finally with frequency scaling.

The concept behind transformations is to split the sound into small frames of 20ms or so. Then run analysis on this frame of data to create a set of parameters that could be used to reproduce the frame of data. Now if want to produce sound that is twice as fast, we create frames from the parameters that are only 10 ms long instead of 20 ms. Similarly, we can create sounds that are twice as slow. For things like pitch changing, we simply tweak the parameter that corresponds to the attribute we are trying to change for each frame.

It should be noted that sine transform coding can also be used to apply transformations. STC finds the frequency and amplitudes of the major peaks of the spectrum. Moving these peaks allows for pitch changes. Since STC is a frame based system, the output can be speeded up or slowed down.

4.0 Musique Concrete

The concept of musique concrete is to digitize music and using the same notes, recombine the sounds to create new melodies.

Prof. Gold's example was a Louis Armstrong cadenza that included all the notes of the scale. In addition, this particular recording of this particular cadenza has unique attributes such as attack and decay times, harmonic relationships, and variations of the spectrum within each note. These attributes can't be easily reproduced, but the recording can be digitized. Once digitized we can recombine the notes to create new music preserving the attributes of the original notes but now in an altogether different melody.