Spring 2008

Wavelets as a Tool for Musical Analysis

Daniel Guillot

Wavelets as a Tool for Musical Analysis

Daniel Guillot

Spring 2008

Contents

1 Introduction

A considerable amount of research has been performed to investigate the mathematical properties of music. Much attention has been given to using the Fourier Transform to analyze music, but using the Wavelet Transform to perform this analysis is a more novel concept. Much theory has been proposed in this field, but there are relatively few solid results readily available.

The wavelet transform has been used in the past to analyze various data signals, including images and audio. The wavelet transform can be used in imaging to denoise images or even to detect the presence of a specific object in an image. It is also capable of allowing slight modifications to an image in order to embed messages that can only be discovered by running the wavelet transform again. The same ideas hold for audio signals using different analyzing wavelets. Wavelets have been used to compress and denoise audio data and work has also been done in pitch detection, rhythm detection and octave detection[2].

I will first give an argument for why the Fourier transform is insufficient for musical analysis, based on its inherent inability to properly handle dynamic signals. I will then detail the basics of the wavelet transform and how it alleviates these problems by scaling and translations of a flexible mother function. Next, I will discuss some of the more significant results of the past 20 years in regard to the wavelet analysis of music. These results include using wavelets for denoising and compressing signals, for pitch and octave detection, and for timbre analysis and synthesis. I will then discuss my work with wavelets as a method of detecting the onset of notes in a piece of music. Finally, I will discuss the possible ramifications of using wavelets for musical analysis, with a focus on the ability to build new pieces of music with desired properties.

2 Why Wavelets and Not the Fourier Transform?

There has been a considerable amount of research using the Fourier transform for the analysis of musical signals. This method can be used to determine the strength of frequency components in musical signals, allowing for identification of the pitch. Work has also been done using the Fourier transform for instrument identification by analyzing its timber. Despite its wide use, however, the Fourier transform has limitations.

The standard Fourier transform is best equipped to handle the analysis of stationary signals. Musical signals, on the other hand, can be quite unpredictable and are more dynamic, having short, rapid changes[3]. Using the Fourier transform requires a trade off between accuracy and precision. For a more accurate reading, there is a large increase in the required precision, and therefore the processing time is greatly increased. Despite this, however, still some information is lost[2].

One method to help overcome this limitation is the windowed Fourier transform. In this, the Fourier transform is run on uniform windows of time. These

uniform windows, however, cause further limitations. Smaller windows give better information in regard to time but sacrifice information about the frequencies. Larger windows yield better information about the frequencies of the signal in exchange for less information about time. The windowed Fourier transform is most useful when the signal being analyzed is stationary within each window[3]. Again, however, music is rarely this stable.

Wavelets offer a better solution than the windowed Fourier transform, since they are translations and dilations of a chosen mother function, allowing for variable window sizes and locations as well as a function that is more suitable for the given signal. This effectively eliminates the limitations of the Fourier and windowed Fourier transforms. Wavelets are useful for dynamic signals such as speech and music[3].

Wavelet packets are even more flexible than wavelets and can provide information about the timber of an instrument, allowing for synthetic creation of tones similar to those of real instruments, as well as novel instruments. This is the most recent area of study in regard to wavelets and music.

3 Wavelets: A Quick Overview

As mentioned in the previous section, wavelets are based off of a mother function which is then dilated and translated, yielding windows of different sizes and locations in which the analysis in performed. In each window, the best fitting approximation based off of the mother function yields the wavelet coefficient for that window.

There is a rather natural comparison between wavelets and how music is written on a musical score. The sheet music corresponds to the analysis of a signal, containing its time-frequency information (Figure 1). Playing the music from the score is equivalent to the synthesis of a signal. A note is the smallest part of a piece of music, each note constituting part of a larger "dictionary" set. All of this is akin to other dictionaries in mathematics, including wavelets[1]. Specifically, one could think of a quarter note of middle C as a general mother function. We can then use this to generate an entire piece of music by translation, scaling and dilating this basic note. The base note can be moved to different locations in the piece of music, corresponding to different window locations. The duration of the note can be changed, which corresponds to different window sizes. Finally, the pitch of the note can be changed, moving it up or down the scale, which corresponds to the given wavelet coefficient for a given window.

The Fourier transform uses dilates of sine and cosine waves to obtain the synthesis equation:

$$f(x) = \int_{-\infty}^{\infty} F(s) e^{2\pi i s x} ds,$$

and accompanying analysis equation:

$$F(s) = \int_{-\infty}^{\infty} f(x) e^{-2\pi i s x} dx,$$

Figure 1: A few measures of a musical score as a time-frequency representation of the music itself.

if f is real valued. For functions that are periodic with period 2π, we have that

$$f(x) = a_0 + \sum_{k=1}^{\infty}(a_k \cos kx + b_k \sin kx),$$

where

$$a_0 = \frac{1}{2\pi}\int_0^{2\pi} f(x)dx$$
$$a_k = \frac{1}{\pi}\int_0^{2\pi} f(x)\cos kx dx$$
$$b_k = \frac{1}{\pi}\int_0^{2\pi} f(x)\sin kx dx.$$

With this representation, $f(x)$ is expressed as the sum of its Fourier series[4].

Wavelet transforms are based on a mother function, Φ, and its translations and dilations:

$$\Phi_{s,l}(x) = 2^{\frac{-s}{2}}\Phi(2^{-s}x - l),$$

where s corresponds to dilations and l to translations. These variables allow for the generation of various mother functions, Φ. Examples of wavelets include the basic Haar Wavelet, the Mexican Hat Wavelet, the Morlet Wavelet and the Meyer Wavelet (Figure 2).

The continuous wavelet transform of a function, $f(x)$ is given by

$$[W_\Phi f](s,l) = \frac{1}{\sqrt{|s|}}\int_{-\infty}^{\infty} f(x)\Phi(\frac{x-l}{s})dx,$$

with variable s and l as before. The wavelet series of a function is given by

$$f(x) = \sum_{s,l=-\infty}^{\infty} c_{s,l}\Phi_{s,l}(x),$$

where $c_{s,l}$ are the wavelet coefficients of the function, such that

$$c_{s,l} = [W_\Phi f](2^{-s}, l2^{-s}).$$

This wavelet transform and corresponding series are analogous to the Fourier transform and its corresponding series. The coefficients of the wavelet series

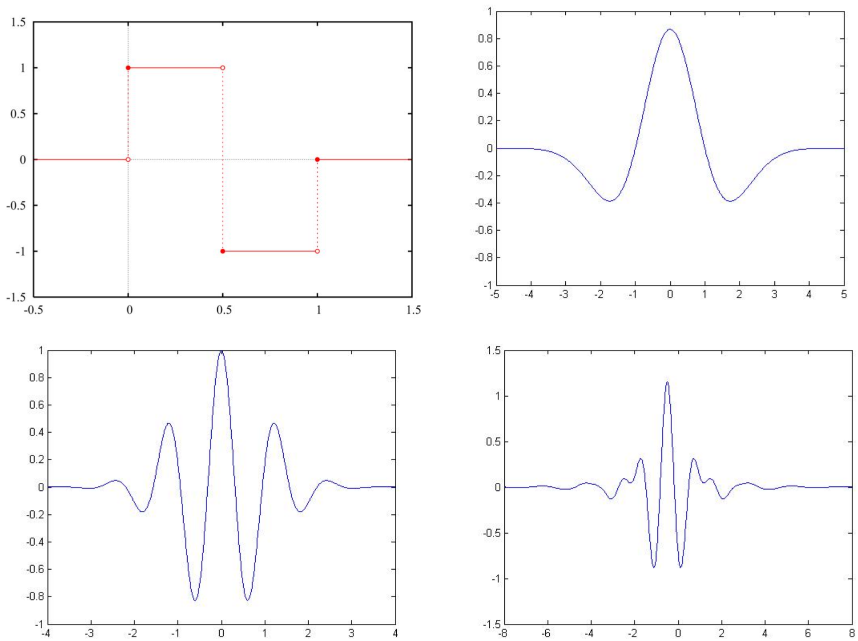

Figure 2: Example Wavelets: Haar, Mexican Hat, Morlet, and Meyer

allow for a compact representation of a function or signal, requiring only a finite number of coefficients to synthesize the original function with only small errors.

Wavelets form a subset of the wavelet packet transform, which are linear combinations of wavelets[4]. Wavelet packets allow for the determination of the Best Basis of a given signal, which may then be used for its analysis. Wavelet packets are well suited for signals that have both stationary and dynamic components, including music and fingerprints [3].

4 Recent Applications of Wavelets to Music

This section will examine some of the notable research that has been done with wavelets to analyze music over the past 20 years.

4.1 Compression

One of the most prominent uses of wavelets in the realm of music has been for data compression. In 1993, Sinha and Tewfik released research in which they used wavelets to create an audio synthesis and coding method. The standard of audio quality at the time was CD quality, which has a sampling rate of 44.1 kHz and a bit rate of 705 kilobytes per second. Previous methods had been employed to reduce the bit rate to 96 kilobytes per second while still maintaining near CD quality sound when synthesized. Sinha and Tewfik's method used optimal adaptive wavelet selection and wavelet quantization procedures with a dynamic dictionary. They found that their adaptive wavelet selection procedure could reduce the bit rate required to achieve near CD quality to at most 70 kilobits per second. If used in conjunction with the dynamic dictionary coding procedure, the bit rate could be reduced to 48-66 kilobits per second[5].

In 1996 Sablatash and Cooklev also published their research on using wavelet packets to compress high-quality audio signals. They suggested that, while the methods that Sinha and Tewfik developed were a great improvement, there was still room for improvement. Sablatash and Cooklev proposed that future research should focus on more sophisticated masking models, paying careful attention to the limitations of the human auditory system to filter out further unnecessary information to lower the bit rates even more [6].

4.2 General Applications

Another typical use of wavelets has been in the denoising of audio signals, providing a cleaner sound without sacrificing much of the sound quality. Wavelets make an excellent method of doing this because they are able to detect sharp changes in a signal. These sharp changes are most often noise related, and a filter can be applied to remove them.

Wavelets may also be used to change the length of a signal while maintaining pitch or to simply change the location of a pitch without altering anything else.

A related but slightly more complex idea is that wavelets may be used to preserve some aspects of music that are detectable, while changing others. One technique presented in [8] has been to use two different wavelets for analysis and synthesis.

Kussmaul also indicates that it is possible to analyze an image using wavelets and then use a different transform to generate music that corresponds to that image, allowing one to "hear" a painting. Alternatively, a composer could write a piece of music and transform it into a picture[8].

Coifman and others also propose that wavelets may have applications in the field of medicine. They suggest that a base signal of fifty healthy heartbeats could be used as a base model. Then, doctors could record the heartbeat of a patient and compare it to this base model, using wavelets, to determine if anything abnormal is occurring[11].

4.3 Timbre Analysis and Synthesis

Wavelets also have potential in not only the resynthesis of music, but also the synthesis of it. Provided a suitable wavelet packet, one could generate the data desired for a piece of music and then simply synthesize it. Using wavelets, Kronland-Martinet analyzed various instruments, developing a packet of wavelets corresponding to them. He also made some changes to these packets to create novel "instruments." He then showed that it is possible to generate novel music by using only some base data for the music and then synthesizing using one of the instrument packets[9].

4.4 Identifying an Unknown Piece

In 2004, Rein and Reisslein published a paper in which they used a wavelet dispersion measure coupled with a neural net to identify a given piece of music, regardless of who was performing the piece. After testing various wavelet families and scales and three different neural nets, they proposed a method which was capable of successfully identifying complex pieces of music that were not in the existing system's search database. Their method had a success rate of 78%. If a given piece of music were in the database, their method had a success rate of 100%. The authors suggest that this system of identifying music may be used to quickly and effectively search a database of music to identify an unknown piece of music via a web search engine for this purpose. Thus, a user with a sound bite of a piece of music whose title she/he does not know could go to the proposed search engine, input the sound clip, and the database would identify the piece and, if the piece itself exists in the database, the artist[12].

5 Personal Observations

5.1 Note Onset Detection

5.1.1 Homogenous Samples

Using the Haar wavelet, we can efficiently locate individual notes that are played in a sample, provided that the volume stays constant. The same method can be employed for chords, again assuming a constant volume, but with the further constraint that all chords must consist of the same number of notes.

This method relies upon a simple filter that preserves a given number of the wavelet coefficients, starting with the largest. If we assume that a measure will contain notes with durations no shorter than thirty-second notes, a number of coefficients equal to $4 \cdot 32m$, where m is the number of measures in the sample, is sufficient to detect the onset of each note. The 4 gives us a sufficient enough number to ensure that enough data around each note is preserved and to account for some volume fluctuation that is natural between notes of different frequencies. The 32 is a reasonable number, since most pieces of music will not contain notes of a shorter duration, much less an entire measure of them.

After applying this filter to a signal, we may run the reverse wavelet transform to see almost precisely where the notes begin. This can be seen in Figure 3.

5.1.2 Non-homogenous Samples

In samples where there are individual notes and/or chords of various amounts of notes, a slightly more sophisticated filter is required. In these cases, we may use a filter that can detect sharp decreases in the sine waves associated with the signal. Using such a filter, we may preserve only the onset of each note or chord without needing to worry about volume. After applying the reverse wavelet transform to this filtered sample, we may again see where each note is played, but some information about the note itself is lost.

5.2 Future Work

The above work is what I have been able to explore on my own until now. I have, however, considered some future options and basic implications.

The two filtering methods in the previous section may seem relatively trivial upon first glance. They do, however, have the ability to save time and processing in future analysis. These filters greatly reduce the amount of data needed to be stored and processed for analysis. When running more transforms to determine the pitch of a note by determining its frequency, for example, only a small period after each note or chord would need to be processed.

This does, however, lead to some data loss. Most importantly the duration of each note or chord is lost, and this is one of the integral parts of a piece of music.

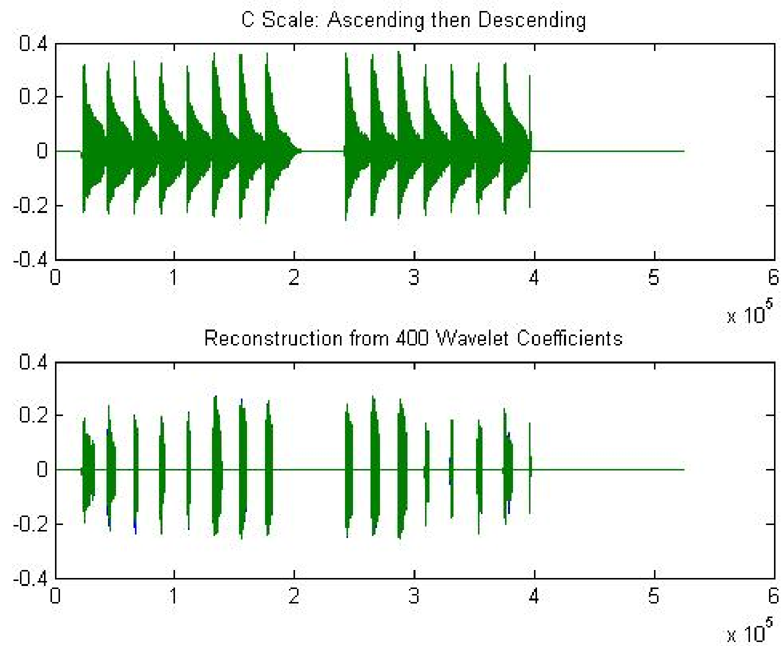

Figure 3: An ascending and descending C scale followed by the result after transforming and filtering out all but the largest 400 coefficients.

Also, the original motivation for this project was to try to apply wavelets to the synthesis of music to be used in music therapy. It has been noted that given a repeating structure, such as a strobe light, the brain will begin to emulate the frequency of the flash, provided it is within an acceptable range. Depression often stems from imbalances in the frequencies of the brain waves of the frontal lobes. Thus, music with some repetitive structure could be used to help renormalize these brain waves.

It may even be possible to use the Electroencephalogram (EEG) reading of a healthy brain to generate music in much the same way that Kussmaul suggested images be used to produce music. This healthy music, containing the properties of the healthy brain waves, could possibly renormalize, temporarily the brain waves of an unhealthy brain.

References

[1] K. Gröchenig. Foundations of Time-Frequency Analysis. Chapter 2. Boston: Birkhaäuser, 2001.

[2] S. Spiegelberg. "The Psychoacoustics of Musical Articulation." Eastman School of Music, January 2002.

[3] B. Hubbard. The World According to Wavelets: The Story of a Mathematical Technique in the Making. Wellesly: A.K. Peters Ltd., 1998.

[4] A. Graps. "An Introduction to Wavelets." *IEEE Computational Science and Engineering.* Vol. 2 No. 2 (1995).

[5] D. Sinha and A. Tewfik. "Low Bit Rate Transparent Audio Compression Using Adapted Wavelets." *IEEE Transactions on Signal Processing*, Vol. 41, No. 12 (1993) 3463-3479.

[6] M. Sablatash and T. Cooklev. "Compression of High-Quality Audio Signals, Including Recent Methods Using Wavelet Packets." *Digital Signal Processing*, 6. (1996) 96-107.

[7] W. Pielemeier, G. Wakefield, and M. Simoni. "Time-Frequency Analysis of Musical Signals." *Proceedings of the IEEE*, Vol. 84, No. 9 (1996) 1216-1230.

[8] C. Kussmaul. "Applications of Wavelets in Music: The Wavelet Function Library." Thesis, Dartmouth College. (1991).

[9] R. Kronland-Martinet. "The Wavelet Transform for Analysis, Synthesis, and Processing of Speech and Music Sounds." *Computer Music Journal*, vol 12, no 4 (1998) 11-20.

[10] R. Kronland-Martinet, J. Morlet, and A. Grossman. "Analysis of Sound Patterns Through Wavelet Transforms." *International Journal of Pattern Recognition and Artificial Intelligence*, Vol. 1 No. 2 (1987) 273-302.

[11] R. R. Coifman, Y. Meyer and M. V. Wickerhauser, "Wavelet analysis and signal processing", in: M. B. Ruskai et al., ed., Wavelets and Their Applications, Jones and Bartlett, Boston, 1992, pp. 153–178.

[12] S. Rein and M. Reisslein. "Identifying the Classical Music Composition of an Unknown Performance with Wavelet Dispersion Vector and Neural Nets (Extended Version)." *Proceedings of the IEEE Int. Conference on Acoustics, Speech, and Signal Processing.* (2004) 341-344.

[13] C. Cheng. "Wavelet Signal Processing of Digital and Audio with Applications in Electro-Acoustic Music." Thesis, Dartmouth College (1996).

[14] G. Tzanetakis, G. Essl, and P. Cook. "Audio Analysis using the Discrete Wavelet Transform." *In. Proc. WSES Int. Conf. Acoustics and Music: Theory and Applications (AMTA 2001).* Skiathos, Greece, 2001.

[15] D. Kammler. A First Course in Fourier Analysis, 2nd Edition. New York: Cambridge University Press, 2007.

[16] Y. Nievergelt. Wavelets Made Easy, 2nd Edition. Boston: Birkhäuser, 2001.