Applications of the Wavelet Transform
at the Level of Pitch Contour

Clifton Kussmaul
Dartmouth College, Dept of Math and Computer Science
Hanover, NH 03755, USA

email: kussmaul@dartmouth.edu

ABSTRACT

This paper describes an investigation of the use of wavelet transforms to analyze and synthesize pitch contours. The wavelet transform analyzes a signal in terms of dilations and translations of a given basis function, and thereby provides better time-frequency resolution than most methods based on Fourier analysis. The current investigation examines possibilities at the level of note events, such as used in MIDI, as opposed to the level of samples in a digitized sound. The author has developed a general-purpose library of functions to manipulate signals at either the sample or note-event level, and is using this library to explore applications in analysis and synthesis.

INTRODUCTION

Wavelet analysis is a recent technique in signal processing which improves on the time and frequency resolution tradeoff found in most analysis methods. Several researchers (e.g. Richard Kronland-Martinet, 1988) have examined applications of wavelets in electro-acoustic music at the timbral level, in which samples in a digitized sound are manipulated. This paper extends these ideas to the level of note events, such as are used in the MIDI protocol. Representing musical information at higher levels of abstraction, such as the note event level rather than the sample level, makes it possible to transform and manipulate this information in new and different ways.

This presentation provides an overview of wavelet theory, a description of possible applications of pitch contour manipulations, a description of the current implementation, and examples of transformations which can be performed using wavelet transforms.

THE WAVELET TRANSFORM

The wavelet transform is a method for improving the analysis of signals in both time and frequency. Originally developed for seismic analysis, wavelets have found applications in a variety of fields. Kronland-Martinet [1988] provides an introduction to wavelets for the computer musician; a more technical introduction is given by Strang [1989], among others. What follows is a brief overview; see one of these other sources for more specific information.

ICMC 483
In the *time domain*, signals are represented as magnitudes sampled at given times. This is the most intuitive representation, since this is how we experience signals in the real world. Signals can also be represented in the *frequency domain*, as magnitudes of sinusoids at given frequencies, with the implicit assumption that the signal is periodic. This is the basis for additive synthesis; sine waves of different frequencies are added together to produce more complex sounds. A variety of representations have been developed which attempt to combine the best features of both time and frequency representations; most involve transforming overlapping segments of the signal in time. A common example is the *spectrograph*, used in speech processing, which displays frequency content as a function of time.

In the wavelet domain, a signal is represented by the weighted sum of a family of *basis functions* which are *dilations* and *translations* of an initial wavelet, rather than representing a signal as a sequence of samples in time or as a sum of infinite sinusoids. Dilation is the stretching or compressing of the wavelet, while translation is simply relocating the wavelet in time. The advantage of using wavelets approach is that the small dilations can be used to capture local or transient detail, while large dilations can capture broader features of the signal. A particular wavelet can be chosen to satisfy a variety of criteria, such as smoothness, orthogonality to its dilations and translations, or its frequency content. Strang (1989) derives several different wavelets by way of example. The rectangular *Haar wavelet* is shown below with a sample dilation and translation.

The wavelet transform of a signal is determined by finding the *detail* in the signal at each *level of resolution*; that is, for each value of the dilation parameter values are found at all appropriate translations. In essence, this is done by convolving the input signal with the appropriately dilated and translated wavelet function. The result is a set of wavelet coefficients for the various values of the dilation and translation variables. Mallat (1989) describes an efficient recursive algorithm for computing the wavelet transform.

**PITCH CONTOUR MANIPULATION**

In most musical traditions, the fundamental unit of sound is the note; melodies and harmonies are described and perceived in terms of pitches, durations, and amplitudes of notes, with less emphasis placed on the internal structure of the sound. Traditional music theory consists of organizing and abstracting these events to even higher levels. These generalizations are less true of electro-acoustic music; the timbres used are often a central, creative part of the piece, with melody, harmony, and rhythm often playing less central roles. Powerful tools have been developed to digitally synthesize and transform sounds.
In view of the tendency to perceive and organize music at higher levels of abstraction, it seems logical to extend the tools used for sample-level manipulations to the level of note events, or pitch contour. It should be possible to take the techniques used to transform digitized sounds and apply them to melodic lines or harmonic progressions, with a variety of potential applications.

As an analysis tool, a wavelet analysis of pitch contour could make it possible to look for regularities at many different levels in a composition. Formal music analysis often involves describing structures, such as key relationships, at several different levels, such as within a phrase, movement, and piece. A wavelet analysis could also serve as a front end to a system for automated musical analysis, such as described by Tenney [1980]. Just as the Fourier transform of a sampled sound can provide insight into the way the sound is produced, a frequency or wavelet-based analysis of the pitch contour of a composition might reveal regularities not immediately obvious to the listener or analysis system. In addition, since the shape of the wavelet can be chosen to facilitate detection of certain features in the signal, a wavelet analysis could be optimized for specific tasks.

As a synthesis tool, wavelets could assist a composer in creating variations on an initial line or progression, by altering some subset of the wavelet coefficients. These manipulations are described in more detail later. Spyridis and Roumeliotis [1983] developed a more ambitious version of this in the Fourier domain; they analyzed a set of songs in a common style, and used the Fourier transform of the pitch contour to build a model of conditional transition probabilities, which were then used to generate new songs. Incorporating wavelet theory into their approach would be a logical extension. Other computer-aided composition systems which use "listening lists of motivic patterns" (e.g. Cope, 1987) might also benefit from using a more abstract representation of pitch contours.

OVERVIEW OF IMPLEMENTATION

The current implementation consists of a library of general-purpose functions for manipulating signals. Signals can be read and written from a variety of standard formats, including MIDI files and NeXT, Sun, Macintosh, and Sound Designer II sound files. A variety of functions exist to manipulate, transform, and compare signals. A user can access these functions through an interpreter or through a family of simple commands in a UNIX environment; eventually more sophisticated graphical interfaces may be developed. The intent is to make the system as general-purpose and flexible as possible, so that different sorts of signals can be analyzed for various purposes on almost any UNIX-based machine.

EXAMPLES

The wavelet transformations described by Krontz and Martinet [1988] can be extended to the level of pitch contour quite easily. The analyzing wavelet can be changed in the reconstruction, allowing aspects of the form of the piece to be preserved while changing the character of the individual motives. By changing the values of the wavelet coefficients, it would be possible to decrease the average interval size locally while increasing the overall range of the piece, or vice versa. Finally, changing the set of translations and dilations used in the reconstruction provides another way of making broad changes while preserving other characteristics.
CONCLUSIONS

This paper has described some of the ways in which the wavelet transform can be applied to pitch contour information rather than sampled sound. The intent has been to provide an overview and suggestions for future work, rather than to restate information presented elsewhere.

REFERENCES


