



WAVELET IN AUDIO PROCESSING FOR REAL-TIME SYSTEM

KALE R.H. AND MOHOD N.P.

M.E. Computer Science and Engg. dept., Sipna C.O.E.T., Sant Gadge Baba University, Amravati, MS, India.

*Corresponding Author: Email- rohini.kale@gmail.com, nikita.mohod@gmail.com

Received: February 21, 2012; Accepted: March 15, 2012

Abstract- Digital processing of audio on personal computers is becoming more and more common. Increasing hardware performance and decreasing price broadens possibilities and quality. Even today's standard PC's are capable of processing CD-quality audio data in real time, making it affordable even for amateurs and small studios to work in the digital domain. Real time audio processing allows modified audio to be heard while it is processed. Although needing much CPU power, it significantly improves professional digital audio: only when the effect of a changed parameter or setting (e.g. volume of an audio track) can be heard instantly, the desired parameter combination can be found in an acceptable time scale[4]. Real time filters also improve non-destructive audio editing possibilities and can reduce the needed disk space for filtered sections. This thesis will evaluate the wavelet theory for the use in real time digital audio processing[6]. Wavelets provide a new way of gathering frequency information from musical signals. Contrary to the traditionally employed technique for doing that based on Fourier transforms - the STFT - time information is not lost in a portion of analyzed audio data. This property (along with others, which are discussed later in this thesis) promises that wavelets provide efficient and suitable algorithms for real-time digital audio processing.

Keywords- wavelet, audio processing, discrete wavelet transform, similar value decomposition., orthogonal bi-orthogonal.

Citation: Kale R.H. and Mohod N.P. (2012) Wavelet in Audio Processing for Real-Time System. Journal of Pattern Intelligence, ISSN: 2230-9330 & E-ISSN: 2230-9349, Volume 2, Issue 1, pp.-15-18.

Copyright: Copyright©2012 Kale R.H. and Mohod N.P. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Introduction

The ever-increasing illegal manipulation of genuine audio products has been a dilemma for the music industry. This situation calls for immediate, yet effective, solutions to avoid further financial losses and intellectual property violations. Audio watermarking has been proposed as a possible solution, since this technology embeds copyright information into audio files as a proof of their ownership. In this paper, we propose an effective, robust, and an inaudible audio watermarking algorithm. The effectiveness of the algorithm has been brought by virtue of applying a cascade of two powerful mathematical transforms; the discrete wavelets transform (DWT) and the singular value decomposition (SVD)[5]. Experimental results will be presented in this paper to demonstrate the effectiveness of the proposed algorithm. Recent unauthorized copying and distribution of digital audio has been greatly facilitat-

ed by the availability of powerful personal computers, low-cost and reliable storage devices, broadband communication networks, and many audio recording and editing software. This alarming situation, has created a need for the protection and enforcement of intellectual property rights for digital media, to prevent its illegal copying and reproduction. Such an urgent need is particularly relevant to the music industry, which is seeking for reliable solutions to problems associated with copyright protection of music files. Data protection techniques, such as encryption, are insufficient for protecting the music industry's intellectual properties. Digital watermarking technology, on the other hand, is now attracting attention as a new method of protecting against unauthorized copying of digital multimedia files that includes image, audio and video components.

Choosing a Wavelet for processing Musical Signals

Audio signals come in many different flavors. Classical music has different characteristics than speech, and both are again different to pop music. This thesis focuses on musical audio signals, without distinction of the characteristics. Speech signals could be regarded as a subset, so most assumptions for musical signals remain valid for speech signals. The idea is to develop algorithms that work well for many kinds of audio signals in real time.

Requirements Quality

The quality of wavelet decomposition especially depends on the ability of approximating the signal with wavelets. When the applied wavelet does not resemble the shape of the analyzed signal, the wavelet coefficients will not extract the main "features" of the signal- resulting in many non-zero wavelet coefficients to approximate the signal. Thus, the better the analysis, the fewer significant wavelet coefficients result- they can be described as "concentrated" coefficients Musical signals are always some kind of smooth wave, significantly smoother than pictures Pictures may have sharp edges, fine lines and high contrast. Short filters corresponding to non-smooth wavelets like Daubechies 2 have proven to approximate well pictures. Musical signals, however, lead to the requirement of a sufficiently smooth wavelet, or in other words, a high regularity is preferred. The size of the transition band of low pass and high pass filters is an important factor, too. Larger transition bands (i.e. low steepness), cause high overlapping of low pass and high pass bands. So the output bands of the filter bank are not separated well, and aliasing effects are enforced when the coefficients are change. Especially in applications where the wavelet coefficients are related directly to frequency (i.e. in pitch shifting), highly separated low pass and high pass frequency response is important. Recursive wavelet filters have been designed which greatly decrease the transition band, however they need a special implementation and could not be researched further for this thesis. Furthermore, linear phase response is crucial for high quality audio filters. When the Furthermore, linear phase response is crucial for high quality audio filters. When the filters do not have at least an approximate linear phase, certain frequencies are delayed in the wavelet domain. The inverse transform undoes this phase distortion. However, when the wavelet coefficients are changed, unwanted modifications may occur to the frequencies, which are "out of phase". Linear phase response can be achieved by using symmetric filters. Last, but not least, different wavelets have different temporal localization. Wavelets with short compact support can localize an event's time better than others. So, for exact temporal analysis, a short wavelet is required, the faster decay the better. This conflicts with the ability of separation of the frequency bands and smoothness- there, longer filters provide better results [6].

Real-time Aspects

Especially in a real time environment, wavelet transforms lead to the requirement of a sufficiently fast algorithm so that the processor is able to compute the forward and inverse wavelet transform faster than the resulting chunk is played. In the example of chunks of 23ms duration, any processing of the chunk may not take more than 23ms- otherwise the flow of chunks will have breaks. The faster the processing has been completed, the better, as the re-

maining processor time can be used for additional processing on the audio signal, operating system tasks, etc. Additionally, some headroom is required, so that the real-time environment operates stable at any time, also when high peaks of processor usage occur. This headroom needs to be especially large for operating systems with preemptive multitasking, as the system may interrupt the chunk processing at any time for other tasks. As the length of the filters directly affects computation time of analysis and resynthesis, shorter filters are preferred. However, in general, more vanishing moments and smaller transition bands lead to longer filters. As this is preferred for audio filters, a reasonable compromise of filter length has to be found. Computers normally process integer numbers faster than floating point numbers. It would be an idea to use one of the integer-based wavelet transforms, e.g. following the procedure described. However, integer sample values are not very well suited for high quality sound processing, resulting in round off errors, unsmooth waves, causing alias effects. High-quality audio processing systems can be assumed to work with signals in a floating point format, so using an integer transform would be of little benefit, while reducing overall quality. All modern PCs are equipped with fast floating point processors, so the performance impact is not very important. Also, the increasing popularity of other programs using floating-point calculations extensively (i.e. 3D games) plays a role for processor manufacturers to develop high performance floating point processing units.

Common Wavelets and their Properties

Some selected wavelets and their properties are presented in this section. In general, constructing a wavelet is not a very difficult task. However, highly sophisticated mathematics is involved when wavelets with special "good" properties are wished. All wavelets presented here were designed with such specialties, so their construction has not been trivial at all (maybe except Haar).

Haar Wavelet

The Haar wavelet is a special one. It has only 2 filter coefficients, so a long transition band is guaranteed. The wavelet function is a square wave; smooth audio signals cannot be approximated well. It is the only wavelet that is at once symmetric and orthogonal. Regarding computation speed, it is perfect for real-time processing. However, the quality is not sufficient: any modification of wavelet coefficients results in strong aliasing.

Daubechies Wavelets

The compactly supported and orthogonal wavelets created by Ingrid Daubechies in the late 1980's gained much attention. They were one of the first to make discrete wavelet analysis practicable. She constructed them by designing orthogonal filters with maximum flatness of the frequency response at 0 and one half the sampling rate (*maxflat filters*). So the restriction for design was the highest number of vanishing moments for a given support width. For a given number of vanishing moments p , the filters have $2p$ coefficients. The minimum support constraint leads to maximum temporal resolution. The resulting filters and wavelets are called Daubechies p or just D_p . For the special case of $p=1$, the resulting wavelet is Haar. Most Daubechies wavelets are not symmetric- in contrary, some are very asymmetric. For small $p>1$, they are not smooth but still continuous. With increasing p , the wavelet function

becomes smoother. For example, the D2 wavelet has singularities at the points $p/2n$ (p and n integer) where it is left-differentiable but not right-differentiable. Due to the flatness, the filters do not separate the frequency bands very well. The steepness of the filter's frequency response is proportional to the square root of $2p$.

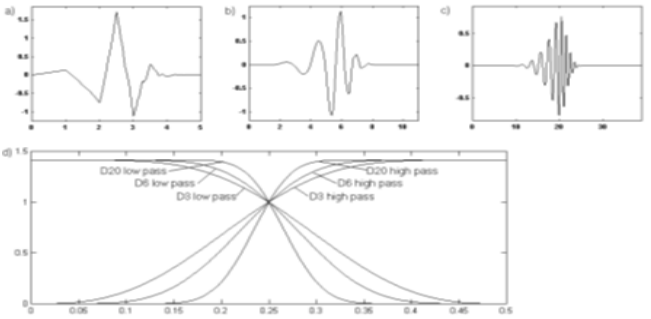


Fig. 1-

Daubechies wavelet family Most Daubechies wavelets are not symmetric- in contrary, some are very asymmetric. For small $p > 1$, they are not smooth but still continuous. With increasing p , the wavelet function becomes smoother. For example, the D2 wavelet has singularities at the points $p/2n$ (p and n integer) where it is left-differentiable but not right-differentiable. Due to the flatness, the filters do not separate the frequency bands very well. The steepness of the filter's frequency response is proportional to the square root of $2p$. Fig. shows a) D3 wavelet, b) D6 wavelet, c) D20 wavelet. In d), the respective filter responses are plotted. It can easily be seen that the higher the order p , the steeper the transition curve.

Other Orthogonal Wavelets

Daubechies constructed a series of other orthogonal wavelets: "symlets" have similar good features like the Daubechies family (compact support, p vanishing moments) but they were designed with the requirement to optimize symmetry and linear phase. Still, as it is impossible for orthogonal wavelets, they are not perfectly symmetric. Another family of wavelets (also constructed by I. Daubechies) are the so-called Coiflets. She constructed them on request of R. Coifman³⁶, who needed wavelets similar to the Daubechies family, but with an additional constraint on the scaling function: not only the wavelet function, but also the scaling function has to have p vanishing moments. This has the advantage that the approximation coefficients can be approximated by the signal samples themselves. However, the support, and therefore the length of the filters, is longer (length of filter $6p$ instead of $2p$ ³⁷), so this additional property costs efficiency.

A special wavelet family is the one of Meyer wavelets. The wavelet and scaling function are constructed in the frequency domain with an auxiliary function. Their support is infinite, but still the functions have a fast decay. They are infinitely differentiable; furthermore they are symmetric and orthogonal, but have no vanishing moments. FIR Filters cannot be constructed, so a filter bank implementation is not possible.

Crude Wavelets

In [6], wavelets which lack many interesting properties are called

"crude": the *Morlet wavelet* and the *mexican hat*³⁸ both have an explicit expression for y , but a scaling function cannot be constructed. They have neither compact support, nor vanishing moments, and are not orthogonal. Due to these limitations, filters cannot be calculated, and only the forward CWT is possible. They are useful for mathematical demonstrations, as the wavelet function exists as a formula.

Biorthogonal Wavelets

There exist a number of well-studied biorthogonal wavelets. The major advantage of biorthogonal wavelets is the possibility to create symmetric transforms: both wavelet and scaling function are symmetric. This requires an odd length of both analysis filters. Biorthogonal wavelet functions and scaling functions are different for analysis and resynthesis, so for a filter bank transform, 2 analysis filters and 2 different resynthesis filters need to be used. Common practice for biorthogonal transforms is to indicate the analysis wavelet and scaling function with $y\sim$ and $f\sim$, respectively. It is apparent that the filters may have different properties for analysis and resynthesis. Consequently, useful properties for analysis are designed into the analysis filters (e.g. vanishing moments) while the resynthesis filters may be designed in respect to useful properties for reconstruction (e.g. regularity). *Battle and Lemarié* introduced biorthogonal wavelets based on polynomial splines. For splines of degree m , the resulting wavelet function has $m+1$ vanishing moments. Unlike Daubechies wavelets, they are not compactly supported; finite filters can only be approximated by cutting off at the edges. However, the wavelet function has exponential decay, so reasonably truncating the filters does not introduce much error. Polynomial spline wavelet functions can be specified explicitly in the frequency domain, and since they are polynomial splines, they are $m-1$ times continuously differentiable, resulting in quite smooth wavelets. For odd m , these wavelets are symmetric. An orthogonalization scheme allows making the Battle- Lemarié family of filters orthogonal. In short, spline wavelets provide maximum regularity with symmetry and minimum support. Other biorthogonal wavelets are Binlets, also based on splines. They are symmetric, have short support and the coefficients are binary: all coefficients of a filter are integers divided by the same power of 2. This allows efficient implementation on computers- division by a power of 2 is "natural" for computers.

Decision

The parameterization possibilities of the wavelet transform provide a high degree of flexibility on its properties and performance. By fixing the wavelet and its parameters for the transform used in the example applications, the flexibility of the wavelet transform would be lost. It would degrade its potential; therefore no definitive choice shall be made. For example, when high separation of the frequency bands is needed, long filters with high demands on processing power are needed. By providing the length of the filters as a parameter to the user, the quality can be adjusted with respect to the performance of the computer. However, some decisions can be taken- mostly by exclusion. The demand for linear phase leads to symmetric biorthogonal wavelets. A high degree of regularity and frequency band separation is preferred. On the other hand, temporal resolution is not a major concern- steep filters are more important. Biorthogonal spline wavelets provide all

these properties. Studies of wavelet transforms for audio or audio like signals agree on this. Consequently, symmetric spline-based wavelets or Battle-Lemarié wavelets will be used for the example applications. Other wavelets are included for comparison purposes.

Conclusion

The paper contains a description of the algorithm which allows us to perform the wavelet transform in real time. The algorithm works on the basis of calculating the optimal extension (overlap) of signal segments, and subsequent performance of the modified transform.

In the future it would be convenient to improve the computational effectivity by reducing redundant computations at the borders of the segments, as it follows from the Algorithm 3.5. Also, it should not be very difficult to generalize the SegWT method to include biorthogonal wavelets and more general types of decimation because the parameters of SegWT can be chosen in a fairly general way. Another important part of the future work is the derivation of an efficient counterpart to the introduced method- the segmented inverse transform. In fact, we made first experience, in which it turned out, above all, that the time lag in the consecutive forward inverse processing will be, unfortunately, always nonzero.

Acknowledgement

I express my sincere gratitude to Resp. Dr. A.D. Gawande Head of the Department, Computer Science & Engineering & Resp. Dr. S.A. Ladhake for providing their valuable guidance and necessary facilities needed for the successful completion of this seminar throughout. I am also obliged to our principal, Resp. Dr. S.A. Ladhake who has been a constant source of inspiration throughout.

Lastly, but not least, I thank all my friends and well-wishers who were a constant source of inspiration.

References

- [1] Darlington D., Daudet L. and Sandler M. (2002) *The 5th Int. Conf. on Digital Audio Effects (DAFX)*.
- [2] Strang G. and Nguyen T. (1996) *Wavelets and Filter Banks*.
- [3] Rajmic P. (2004) *PhD Thesis, Brno University of Technology*.
- [4] Dutilleul P. (1989) *Time-Frequency Methods and Phase Space, Inverse Problems and Theoretical Imaging*, 298-304.
- [5] Nason G.P. and Silverman B.W. (1995) *Wavelets and Statistics*, 103 of *Lecture Notes in Statistics*, 281-300.
- [6] Matteo F. and Johnson Steven G. (2000) <http://www.fftw.org/>.
- [7] Tony F. <http://wwwusers.cs.york.ac.uk>.
- [8] Ghael S., Sayeed A. and Baraniuk R. <http://www.dsp.rice.edu>.