WAVELET TRANSFORM BASED METHODS IN DIGITAL AUDIO PROCESSING

Robert GALAMBOS
Advisor: Laszlo SUJBERT

I. Introduction

Nowadays the computational capacity of computers and DSP-s increased radically letting the audio processing move to the digital domain, where new possibilities are available. Beside the well known tools, and algorithms for analyzing and synthesizing digital signals there are special solutions and algorithms adapted to the features of the human perception.

II. Time and frequency resolution

The digital signal can be processed in time domain or transformed – e.g. with the Fourier transform – into the time-frequency plane. The time-frequency resolution is limited by the uncertainty principle [1], that means that the resolution of the time and the frequency can not be arbitrary large at the same time.

A. Fourier transform

The basic element of the continuous-time Fourier transform is an infinitely long complex periodic wave. The transformation of this wave is a so called Dirac impulse. We can not talk about time resolution, because the time domain signal is infinitely long.

By sampling this continuous signal we get to the discrete-time Fourier transform. Still there is no time resolution because in time domain the sampled signal has infinite number of samples. To be able to localize the signal we need to multiply it with a window function, and "cut out a piece" in the time domain.

This is the so called short-time Fourier transform. The multiplication of the signal and the window function in time domain, is convolution in the frequency domain, so the signal itself is modified by the convolution with the window functions Fourier transform [2]. The short-time Fourier transform has the same window function for each frequency location, resulting an equidistant grid on the time-frequency plane. Because the human ear recognizes the frequencies in a logarithmic manner, the Fourier transform might not produce as good result as a transformation would do that follows this logarithmic manner.

B. Wavelet transform

The wavelet transform can be seen as a modification of the short-time Fourier transform. The window function multiplied signal is replaced by the so called mother wavelet, a special function with scale and shift parameters. This is the source of the daughter wavelets. They are derived from the mother wavelet by shifting and scaling. The mother wavelet must vanish at zero frequencies, so it has a band-pass like characteristic. So the average of the mother wavelet must be zero, it has to be a "wave". The daughter wavelets will have another central frequency and bandwidth in the frequency domain, so the time-frequency plane grid structure will not be equidistant.

From the multi-rate signal processing point of view the wavelet transform can be implemented as filter bank based calculation, where lower frequencies are represented more accurate in the frequency parameter, and higher frequencies are represented more accurate in the time parameter.

Following the filter bank approach, and because the wavelets must vanish at zero frequency it is not possible to cover the whole frequency range with the wavelets. There will be always a blind area under the lowest band. There is a so called scaling function that covers the near zero frequency range.

Because the wavelet filter bank is built up in this way, it is possible to use sub-band coding [3] to calculate the wavelet transform of a signal. This can be done in two ways. We can use parallel bandpass filters, or an iterative filter bank structure. The latter has the advantage that the calculation complexity is less.

III. Practical occurrences of Wavelet transform

There are several situations where wavelet transform can be used instead of the usual methods (e.g. Fourier transform). With the advantage of fitting better to the human ears perception the wavelet transform is fairly a better choice for audio applications.

Decomposing the signal into wavelets makes it possible to filter the noise on each decomposition to reduce the overall noise of the signal. Each decomposition has a threshold level, and if the amplitude falls under this level, the gain of the given part will be muted or dumped [4].

Watermarks are special signs in the audio (or image) to store information, which can be used to identify the author of the signal, or to store copyrights. Wavelet transform can be used for giving another representation of the signal producing more robust watermarks [5].

The wavelet decomposition can be used for data compression. By decomposing the signal to wavelets, it is possible to simplify the dimensions needed to describe the signal [6].

Wavelet transform can be used in special audio effect algorithms, to produce a higher quality output. For example there is a big advantage of using wavelet transform in a phase vocoder algorithm instead of using the usual short-time Fourier transform. In this case the fine details of the signal remain after the processing, and less distortion is applied.

As the short-time Fourier transform the wavelet transform can be used in audio identification algorithms, by creating a fingerprint from the wavelet coefficients. This fingerprint must be robust against noises and other distortions on the audio signal. It is used for example for automatic music detection.

For modelling musical instruments it could be a good choice to use wavelet based analysis and/or synthesis and so allow the model to use less computational capacity.

IV. Conclusions

The wavelet transform is a transform that comes closer to the human perception because of its logarithmic manner. According to this it is possible to achieve better results with this transformation than with the usual Fourier transform. The difference is in the resolution of the time-frequency plane. Where the short-time Fourier transform has equidistant grids the wavelet transform has more option. It is possible to achieve a frequency resolution on the needed level in high and low frequency ranges. Beneath these features in some cases the wavelet transform can be calculated with a small amount of computation power by using the so called fast wavelet transform.

References

[1] C. K. Chui, Ed., *An Introduction to Wavelets*, Academic Press, 1992.

[2] G. Kaiser, Ed., *A Friendly Guide to Wavelets*, Birkhauser, 1994.

[3] S. Mallat, Ed., *A Wavelet Tour of Signal Processing, Second Edition*, Academic Press, 1999.

[4] E. B. Guoshen Yu, Stephane Mallat, "Audio denoising by time-frequency block thresholding," *IEEE Transaction on Signal Processing*, 56(5):1830–1839, 2008.

[5] A. M. Ali Al-Haj, "Digital audio watermarking based on the discrete wavelets transform and singular value decomposition," *European Journal of Scientific Research*, 39(1):6–21, 2010.

[6] M. Zanartu, "Audio compression using wavelet techniques," Tech. Rep., Purdue University, 2005.