

## Chapter 8

# A Psychoacoustic Model Based on the Discrete Wavelet Packet Transform

The perception of acoustic information by humans is based on the detailed temporal and spectral analysis provided by the auditory processing of the received signal. The incorporation of this process in psychoacoustical computational models has contributed significantly both in the development of highly efficient audio compression schemes as well as in effective audio watermarking methods. In this chapter, we present an approach based on the discrete wavelet packet transform, which closely mimics the multi-resolution properties of the human ear and also includes simultaneous and temporal auditory masking. Experimental results show that the proposed technique offers better masking capabilities and it reduces the signal-to-masking ratio when compared to related approaches, without introducing audible distortion. Those results have implications that are important both for audio compression by permitting further bit rate reduction and for watermarking by providing greater signal space for information hiding.

DOI: 10.4018/978-1-61520-925-5.ch008

## **8.1 INTRODUCTION**

Psychoacoustic modeling has made important contributions in the development of recent high quality audio compression methods (ISO/IEC 11172-3, 1993; Painter and Spanias, 2000; Pan, 1995) and has enabled the introduction of effective audio watermarking techniques (Swanson et al., 1998; Liu, 2004; Cox et al., 2002). In audio analysis and coding, it strives to reduce the signal information rate in lossy signal compression, while maintaining transparent quality. This is achieved by accounting for auditory masking effects, which make possible to keep quantization and processing noises inaudible. In speech and audio watermarking the inclusion of auditory masking has made possible the addition of information that is unrelated to the signal in a manner that keeps it imperceptible and can be effectively recovered during the identification process.

Most psychoacoustic models used in audio compression or watermarking, have so far utilized the short-time Fourier transform (STFT) to construct a time-varying spectral representation of the signal (Painter and Spanias, 2000; Cox, 1997; Bosi and Goldberg, 2003). A window sequence of fixed length is used to capture a signal section, resulting in a fixed spectral resolution. The STFT is applied on windowed sections of the signal thus providing an analysis profile at regular time instances. However, the STFT can provide only averaged frequency information of the signal and it lacks the flexibility of arbitrary time-frequency localization (Polikarg, online, 2006). Such a rigid analysis regime is in striking contrast with the unpredictably dynamic spectral-temporal profile of information-carrying audio signals. Instead, signal characteristics would be analyzed and represented more accurately by a more versatile description providing a time-frequency multi-resolution pertinent to the signal dynamics. The approaches included in the MPEG 1 standard and elsewhere allow the switching between two different analysis window sizes depending on the value of the signal entropy (ISO/IEC 11172-3, 1993), or the changes in the estimated signal variance (Lincoln, 1998). Greater flexibility, however, is needed. The wavelet transform presents an attractive alternative by providing frequency-dependent resolution, which can better match the hearing mechanism (Polikarg, online, 2006). Specifically, long windows analyze low frequency components and achieve high frequency resolution while progressively shorter windows analyze higher frequency components to achieve better time resolution. Wavelet analysis has found numerous signal processing applications including video and image compression (Abbate et al., 2002; Jaffard et al., 2001), perceptual audio coding (Veldhuis et al., 1998), high quality audio compression and psychoacoustic model approximation (Sinha and Tewfik, 1993).

Wavelet-based approaches have been previously proposed for perceptual audio coding. Sinha et al. (1993) used the masking model proposed in (Veldhuis et al.,

25 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: [www.igi-global.com/chapter/psychoacoustic-model-based-discrete-wavelet/56063](http://www.igi-global.com/chapter/psychoacoustic-model-based-discrete-wavelet/56063)

## Related Content

---

### MINTCar: A Tool Enabling Multiple Source Multiple Destination Network Tomography

Laurent Bobelin (2012). *Advancements in Distributed Computing and Internet Technologies: Trends and Issues* (pp. 86-111).

[www.irma-international.org/chapter/mintcar-tool-enabling-multiple-source/59679](http://www.irma-international.org/chapter/mintcar-tool-enabling-multiple-source/59679)

### Toward LessonCapture: A New Approach to Screencasting and Lecture Capture

Steve Garwood (2013). *Enhancing Instruction with Visual Media: Utilizing Video and Lecture Capture* (pp. 89-105).

[www.irma-international.org/chapter/toward-lessoncapture-new-approach-screencasting/75415](http://www.irma-international.org/chapter/toward-lessoncapture-new-approach-screencasting/75415)

### Bitrate Adaptation of Scalable Bitstreams in a UMTS Environment

Nicolas Tizonand Béatrice Pesquet-Popescu (2011). *Streaming Media Architectures, Techniques, and Applications: Recent Advances* (pp. 147-169).

[www.irma-international.org/chapter/bitrate-adaptation-scalable-bitstreams-umts/47518](http://www.irma-international.org/chapter/bitrate-adaptation-scalable-bitstreams-umts/47518)

### The Use of Eye Tracking as a Research and Instructional Tool in Multimedia Learning

Katharina Scheiterand Alexander Eitel (2018). *Digital Multimedia: Concepts, Methodologies, Tools, and Applications* (pp. 698-719).

[www.irma-international.org/chapter/the-use-of-eye-tracking-as-a-research-and-instructional-tool-in-multimedia-learning/189500](http://www.irma-international.org/chapter/the-use-of-eye-tracking-as-a-research-and-instructional-tool-in-multimedia-learning/189500)

### MPEG Standards and Watermarking Technologies

Jong-Nam Kimand Byung-Ha Ahn (2005). *Digital Watermarking for Digital Media* (pp. 182-214).

[www.irma-international.org/chapter/mpeg-standards-watermarking-technologies/8557](http://www.irma-international.org/chapter/mpeg-standards-watermarking-technologies/8557)