International Journal of Innovative Computing, Information and Control
Volume **6**, Number **3(B)**, March **2010**

ICIC International ©2010 ISSN 1349-4198
pp. 1209–1220

A DIGITAL WATERMARK FOR STEREO AUDIO SIGNALS USING VARIABLE INTER-CHANNEL DELAY IN HIGH-FREQUENCY BANDS AND ITS EVALUATION

Kazuhiro Kondo and Kiyoshi Nakagawa

Graduate School of Science and Engineering
Yamagata University
4-3-16 Jonan, Yonezawa, Yamagata 992-8510, Japan
{ kkondo; nakagawa }@yz.yamagata-u.ac.jp

Received December 2008; revised May 2009

Abstract. *We propose a watermarking algorithm for stereo audio signals which embeds data using delay values of the high-frequency channel signals. Since the stereo image perception of the human auditory system is known to be relatively insensitive to phase in the high-frequency region, we replace the two high-frequency channel signals with one middle channel signal, and embed data as the delay between the channels. Blind detection of embedded data is possible using the correlation between high-frequency channels at the detector. Embedded data rate was 7 to 30 bps depending on the host audio signal. Embedded data was detected with little or no errors for added noise at 20 dB SNR and above. MP3 and AAC coders, as well as sample rate conversion (both down-sampling and up-sampling) did not affect the data. Low-pass filtering and high-pass filtering also did not affect the data when cut-off frequencies were set to leave reasonable amount of frequency ranges intact. The embedded audio quality was shown to be source-dependent, with quality equivalent to MP3 coded audio with some sources, but significantly lower for other sources.*

Keywords: Audio signal, Watermark, Stereo, Inter-channel delay, Blind detection

1. **Introduction.** With the recent advances in multimedia signal processing technology, it has become relatively easy to edit, copy, store and transmit digital multimedia content. This situation calls for intellectual ownership protection, and detection/protection against unauthorized tampering. Digital watermarking, or information hiding, is a strong candidate for such protection. Digital contents can be "marked" with copyright information using watermarks, for instance.

Although the volume of research in information hiding (watermarking) has been in images and video [1, 2, 3, 4], we recently see new proposals for information hiding in speech and audio [5]. Most of these take advantage of the human auditory system (HAS) in order to hide information into the host speech or audio signal without causing significant perceptual disturbances. Temporal masking and frequency masking [6] are two of the properties of HAS that are utilized most to hide information in audio signals without noticeable degradations. These are also used in MPEG audio coding standards [7, 8, 9] to compress the audio signals without noticeable coding noise.

Phase coding has been regarded as an effective method to embed data with minimum impact on the perceived quality [1, 10]. Phase of the original audio is replaced with reference phase according to the data. Since the HAS is known to be relatively insensitive to small phase distortions [11, 12, 13], this method can potentially embed data without affecting the host signal quality significantly. In this paper, we use this property to code