Chapter VI

Audio and Speech Watermarking and Quality Evaluation

Ronghui Tu, University of Ottawa, Canada

Jiying Zhao, University of Ottawa, Canada

Abstract

In this chapter, we will introduce digital watermarking systems and quality evaluation algorithms for audio and speech signals. Originally designed for copyright protection, digital watermarking is a technique that embeds some data to the host signal to show the copyright information. The embedded data, or the watermark, cannot be eliminated from the host signal with normal use or even after some attacks. In the first part of this chapter, we introduce a generic framework of digital watermarking systems. Several properties and application of the watermarking systems are then described. Focused on audio and speech signals, we divide the watermarking algorithms into robust and fragile ones. Different concerns and techniques are described for each category. In the second part of this chapter, a novel watermarking algorithm for audio and speech quality evaluation is introduced together with some conventional quality evaluation standards.
Introduction

Digital watermarking is a technique to embed a collection of bits into a signal. We refer to this collection of bits as a watermark. A well-designed watermarking algorithm will keep the watermark imperceptible to users. In other words, the resulting signal remains almost the same quality as the original signal. Watermarks can be embedded into audio, image, video, and other formats of digital data. There are different concerns to different formats of data in watermarking system design. In this chapter, we narrow our scope to audio and speech watermarking systems, or simply audio watermarking systems.

There are several ways to categorize audio watermarking systems. They can be classified as temporal and spectral watermarking. Temporal watermarking algorithms embed watermarks into audio signals in their temporal domain. Spectral watermarking algorithms embed watermarks in certain transform domain, such as Fourier transform domain, wavelet domain, or cepstrum domain.

Watermarking algorithms can also be categorized as robust and fragile ones. Watermarks in robust algorithms cannot be removed by common signal processing operations. On the contrary, a fragile watermark will be changed if the host audio is modified. In this chapter, we will follow this classification.

In recent years, digital watermarking algorithms boomed rapidly, especially in the image watermarking field. Compared to image, the research on audio watermarking is not as mature. The most important reason is the difference between the human visual system (HVS) and human auditory system (HAS). In general, HAS is more sensitive to distortions than the visual system. Therefore, it is challenging to implement imperceptible audio watermarks.

The rest of this chapter is organized as follows. In the first section, a generic digital watermarking framework is presented. A variety of properties and applications of watermarking algorithms are described. In the second section we provide a classification of attacks and a benchmark for audio watermarking systems. Several techniques for robust audio watermarking algorithms are described in the third section. Example algorithms for each technique are introduced. The fourth section is designated for fragile audio watermarking. Since there are not many algorithms for fragile audio watermarking, we combine semi-fragile watermarking in this section. In the last section, we introduce traditional audio and speech quality evaluation methods and a watermark-based evaluation algorithm.

Introduction to Digital Watermarking Technology

With the rapid growth of Internet, personal computers, and a variety of new digital devices, the digital format of media becomes more and more popular. The variety of software makes it convenient for consumers to create, manipulate, and store the digital multimedia data. Internet and wireless network provide a channel to transmit and to exchange the multimedia information. In the recent decades, these new formats of data have brought many changes to our life. However, they also pose the danger of illegal copy, redistribution, and various
Related Content

Maximizing ANLP Evaluation: Harmonizing Flawed Input
Adam Renner, Philip M. McCarthy, Chutima Boonthum-Denecke and Danielle S. McNamara (2012). *Applied Natural Language Processing: Identification, Investigation and Resolution* (pp. 438-456).
[www.igi-global.com/chapter/maximizing-anlp-evaluation/61064?camid=4v1a](www.igi-global.com/chapter/maximizing-anlp-evaluation/61064?camid=4v1a)

Multimodal Emotion Recognition
[www.igi-global.com/chapter/multimodal-emotion-recognition/45495?camid=4v1a](www.igi-global.com/chapter/multimodal-emotion-recognition/45495?camid=4v1a)

Mining and Visualizing the Narration Tree of Hadiths (Prophetic Traditions)
Aqil Azmi and Nawaf AlBadia (2012). *Cross-Disciplinary Advances in Applied Natural Language Processing: Issues and Approaches* (pp. 239-257).
[www.igi-global.com/chapter/mining-visualizing-narration-tree-hadiths/64591?camid=4v1a](www.igi-global.com/chapter/mining-visualizing-narration-tree-hadiths/64591?camid=4v1a)

Data Mining
Martin Atzmueller (2012). *Applied Natural Language Processing: Identification, Investigation and Resolution* (pp. 75-94).
[www.igi-global.com/chapter/data-mining/61043?camid=4v1a](www.igi-global.com/chapter/data-mining/61043?camid=4v1a)