Chapter 9
Spatial and Temporal Position Information Delivery to Mobile Terminals Using Audio Watermarking Techniques

Toshio Modegi
Dai Nippon Printing Co., Ltd., Japan

ABSTRACT
These authors are developing audio watermarking techniques that enable the extraction of embedded data by mobile phones. They applied acoustic interpolation of human auditory organs to embed data in full phone-line frequency ranges, where human auditory response is important for facilitating data extraction, using 3G mobile phones. They are interested in applying this technique to a mobile guide system for use in museums. In particular, they are considering applying audio watermarking techniques for synchronizing the stored contents of mobile terminals based on the spatial positions of the terminal and the temporal positions of playback contents in surrounding media. For this purpose, they are developing five linear spatial location identification codes that transfer to mobile terminals via two-channel stereo audio media that have embedded watermarks. They are also developing time codes that continuously transfer to mobile terminals via audio media. In this chapter, the authors initially describe their proposed audio watermarking algorithm and then present the main topic of novel audio watermarking applications for position information delivery to mobile terminals.

DOI: 10.4018/978-1-4666-2217-3.ch009
INTRODUCTION

In these days, “Ubiquitous Acoustic Spaces,” a term that has been defined by the author (Modegi, 2007), are growing to be popular. In these spaces, each sound source can emit certain link address information using audio signals, thereby allowing automatic access to related cyberspace using mobile terminals, such as mobile phones. For example, QR codes (quick response bar codes) shown during TV commercials can be used to indicate related shopping site URLs and these can be transmitted visually to allow site access by capturing with a mobile phone camera. However, it is not so simple to implement such site linking services in radio commercials.

A further example is a museum guide system, which the authors focus on in this chapter, where real object exhibits can be linked to their respective virtual museum sites via mobile terminals. This application requires that exhibit showcases send spatial position information to visitors via their mobile terminals. In museum video presentation areas, video equipment must send temporal position information to mobile terminals before the video contents are played back on the terminals. In order to implement these applications, the requisite technical topics are described in this chapter.

Proposed Basic Watermark Technology: Development of an Analogue-Robust Audio Watermark Embedding Technique, “G-Encoder Mark,” Enabling the Extraction of Embedded Data Using a 3G Mobile Phone

To implement this application, the authors aimed to develop a novel audio watermark technique to extract embedded data simply by directing a mobile phone at a loudspeaker that emits watermark embedded audio signals. They addressed the following problems.

Current 3G mobile phones and public phone networks cannot capture audio frequency components higher than 4 kHz. Recorded sound data are automatically compressed by the 3GPP specification format, so the sound quality can be highly degraded. Stego audio signals can also be distributed via analogue broadcasting, digital broadcasting, or IP network streaming, but frequency components higher than 4 kHz may be degraded by signal modulation or compression operations. Thus, data needs to be embedded in frequency components lower than 4 kHz when we implement audio watermark extraction functions for mobile phones. These frequency components are within the most sensitive auditory ranges, so we had to make extensive modifications to allow them to be detected by mobile phones. We also had to reduce additional auditory noises caused by these modifications during the playback of embedded signals.

The application of acoustic interpolation of human auditory organs (interpolation of missing signal components by the auditory stream segregation phenomenon) allowed us to embed data in full phone-line frequency ranges without adding any audible noises. Thus, the data signals were embedded in major auditory response ranges to facilitate data extraction using commonly available 3G mobile phones.

Proposed Main Functions: Provision for Extended Functions that Facilitate Spatial and Temporal Position Information Delivery of Embedded Audio Watermarks

The authors proposed an extended audio watermark method to extract new codes at the center of a stereo audio playback environment. Specifically, they developed five linear spatial location identification codes with embedded watermarks that can be transferred to mobile terminals via two-channel stereo audio media.