# Chapter V
# Damageless Watermark Extraction Using Nonlinear Feature Extraction Scheme Trained on Frequency Domain

**Kensuke Naoe**
*Keio University, Japan*

**Yoshiyasu Takefuji**
*Keio University, Japan*

## ABSTRACT

*In this chapter, we propose a new information hiding and extracting method without embedding any information into the target content by using a nonlinear feature extraction scheme trained on frequency domain. The proposed method can detect hidden bit patterns from the content by processing the coefficients of the selected feature subblocks to the trained neural network. The coefficients are taken from the frequency domain of the decomposed target content by frequency transform. The bit patterns are retrieved from the network only with the proper extraction keys provided. The extraction keys, in the proposed method, are the coordinates of the selected feature subblocks and the neural network weights generated by the supervised learning of the neural network. The supervised learning uses the coefficients of the selected feature subblocks as the set of input values, and the hidden bit patterns are used as the teacher signal values of the neural network, which is the watermark signal in the proposed method. With our proposed method, we are able to introduce a watermark scheme with no damage to the target content.*

## INTRODUCTION

In this chapter, we present a new model of digital watermark that does not embed any data into the content, but is able to extract meaningful data from the content. This is done by processing the coefficients of the selected feature subblocks to the trained neural network. This model trains a neural

network to assign predefined secret data, and use the neural network weight and the coordinates of the selected feature subblocks as a key to extract the predefined code. This means that it would not damage the content at all. The proposed method is an improvement from the paper of our research project, which was published before (Ando & Takefuji, 2003).

In Section 2, we discuss the background surrounding digital watermarks, frequency transformation, and neural networks. We demonstrate the characteristics, and discuss what techniques are useful for implementing digital watermarks. In Section 3, we propose the method of damageless watermark embedding and extraction for still image. In Section 4, we provide experiment results for testing its robustness and fragileness. In Section 5 and 6, we conclude with a discussion of the proposed method, and indicate some future works of the proposed method.

## BACKGROUND

In this section, we discuss the background surrounding digital watermarks, and we go deeply to the backgrounds and researches in frequency transformation and neural networks, which consist of important modules for the proposed method.

### General Background Surrounding Digital Watermarks

Recently, with the rapid progress of information technologies and the emergence of the Internet, digital multimedia contents are easily distributed on the network. This circumstance helped to open digital contents to the public without difficulty, even for ordinary computer users, but also helped illegal distribution and copying of contents. Due to the characteristics of digital contents, digital contents are easy to make an exact copy and to alter the content itself. This became a main concern for authors, publishers, and legitimate owners of the contents. Therefore, digital watermark became a key technology for protecting the copyrights. Digital watermark protects unauthorized change of the contents and assures legal uses for its copyright.

There are several ways to protect digital content. One example is to encrypt the content and to share the decryption key between the author and the observer. But this method prevents other observers without a decryption key from accessing the content. This feature avoids a free distribution and circulation of the content through the network, which, most of the time, is not desirable to the author of the content. Digital watermark only embeds data to the content, and this feature does not avoid the distribution of the content.

Watermarking techniques are one technique of information hiding techniques (Katzenbeisser & Petitcolas, 2000). The research in information hiding has a history (Kahn, 1996), namely, the researches in digital watermark and steganography have been active. Both are very similar, but the applications are different (Reither & Rubin, 1998). Digital watermark can be classified in several ways. The first classification is by the perceptibility of the watermark signal to humans. A perceptible watermark has various usages but because it limits the utility of the content, most of the research in this area has focused on imperceptible watermarking techniques. Digital watermark is often embedded imperceptibly to human receptors to avoid contaminating the content. For imperceptible images, the human visual system (HVS) model is often used. There are many still image watermark researches that make use of HVS model (Delaigle, De Vleeschouwer, & Macq, 1998; Kim, Byeong, & Choi, 2002; Reither & Rubin, 1998; Swanson, Zhu, & Tewfil, 1996; Westen, Lagendijk, & Biemond, 1996). For imperceptible audio, a psychoacoustic model is often used. The basic idea of the psychoacoustic model is that human ears are insensitive to the change in phase and amplitude, but very sensitive to the change in the time domain. Also, humans have

## Related Content