Digital Watermarking and Steganography

Kuanchin Chen
Western Michigan University, USA

INTRODUCTION

Sharing, disseminating, and presenting data in digital format is not just a fad, but it is becoming part of our life. Without careful planning, digitized resources could easily be misused, especially those that are shared across the Internet. Examples of such misuse include use without the owner's permission, and modification of a digitized resource to fake ownership. One way to prevent such behaviors is to employ some form of copyright protection technique, such as digital watermarks.

Digital watermarks refer to the data embedded into a digital source (e.g., images, text, audio, or video recording). They are similar to watermarks in printed materials as a message inserted into the host media typically becomes an integral part of the media. Apart from traditional watermarks in printed forms, digital watermarks may also be invisible, may be in the forms other than graphics, and may be digitally removed.

INFORMATION HIDING, STEGANOGRAPHY, AND WATERMARKING

To many people, information hiding, steganography, and watermarking refer to the same set of techniques to hide some form of data. This is true in part because these terms are closely related to each other, and sometimes they are used interchangeably.

Information hiding is a general term that involves message embedding in some host media (Cox, Miller, & Bloom, 2002). The purpose of information hiding is to make the information imperceptible or to keep the existence of the information secret. Steganography means "covered writing," a term derived from the Greek literature. Its purpose is to conceal the very existence of a message. Digital watermarking, however, embeds information into the host documents, but the embedded information may be visible (e.g., a company logo), or invisible (in which case, it is similar to steganography).

Steganography and digital watermarking differ in several ways. First, the watermarked messages are related to the host documents (Cox et al., 2002). An example is the ownership information inserted into an image. Second, digital watermarks do not always have to be hidden. See Taylor, Foster, and Pelly (2003) for the applications of visible watermarks. However, visible watermarks are typically not considered steganography by definition (Johnson & Jajodia, 1998). Third, watermarking requires additional "robustness" in its algorithms. Robustness refers to the ability of a watermarking algorithm to resist from removal or manipulation attempts (Acken, 1998; Craver, Perrig, & Petitcolas, 2000). This characteristic deters attackers by forcing them to spend an unreasonable amount of computation time and/or by inflicting an unreasonable amount of damage to the watermarked documents in the attempts of watermark extraction. Figure 1 shows that there are considerable overlaps in the meaning and even the application of the three terms. Many of the algorithms in use today are, in fact, shared among information hiding, steganography, and digital watermarking. The difference relies largely on "the intent of use" (Johnson & Jajodia, 1998). Therefore, discussions in the rest of the article on watermarking also apply to steganography and information hiding, unless specifically mentioned otherwise.

To be consistent with the existing literature, a few common terms are described below. *Cover work* refers to the host document (text, image, multimedia, or other media content) that will be used to embed data. The data to be embedded is called the *watermark* and may be in the form of text, graphic, or other digital format. The result of this embedding is called a *stego-object*.

CHARACTERISTICS OF EFFECTIVE WATERMARKING ALGORITHMS

Watermarking algorithms are not created equal. Some will not survive from simple image processing operations, while others are robust enough to deter attackers from some forms of modifications. Effective and robust

Figure 1. Information hiding, steganography, and digital watermarking

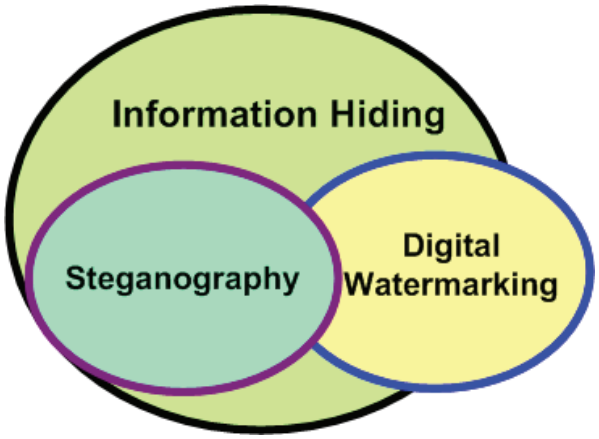

image watermarking algorithms should meet the following requirements:

- **Modification tolerance:** They must survive common document modifications and transformations (Berghel, 1997; Zheng, Liu, Zhao, & Saddik, 2007);
- **Ease of authorized removal:** They must be detectable and easily removable by authorized users (Berghel, 1997); and
- **Difficulty for unauthorized modifications:** They also must be difficult enough to discourage unauthorized modifications.

In addition to the above requirements for image watermarking algorithms, Mintzer, Braudaway, and Bell (1998) suggest the following for watermarking digital motion pictures:

- **Invisibility:** The presence of the watermark should not degrade the quality of motion pictures;
- **Unchanged compressibility:** The watermark should not affect the compressibility of the media content; and
- **Low cost:** Watermark algorithms may be implemented in the hardware as long as they only add insignificant cost and complexity to the hardware manufacturers.

The main focus of these requirements concerns the capabilities of watermarking algorithms to survive various attacks or full/partial changes to the stego-object. However, the fundamental requirement for most algorithms is unobtrusiveness. Unless the goal of using an algorithm is to render the host medium unusable or partially unavailable, many watermarking algorithms will not produce something perceptibly different from the cover work. However, theoretically speaking, stego-objects are hardly the same as the cover work when something is embedded into the cover work.

When it comes to watermarking text documents, most of the above requirements apply. A text watermarking algorithm should not produce something that is easily detectable or render the resulting stego-object illegible. Different from many image or multimedia watermarking techniques which produce imperceptible watermarks, text watermarking techniques may render a visible difference if the cover work and stego-object are compared side by side.

DIGITAL WATERMARKS IN USE

Authentication of the host document is one important use of digital watermarks. In this scenario, a watermark is inserted into the cover work resulting in a stego-object. Stripping off the watermark should yield the original cover work. Common uses of authentication watermarks include verification of object content (Mintzer, Braudaway, & Bell, 1998), and copyright protection (Acken, 1998). The general concept of watermarking works in the following way.

$\boldsymbol{W} + \boldsymbol{M} \rightarrow \boldsymbol{S}$, where $\boldsymbol{W}$ is the cover work, $\boldsymbol{M}$ is the watermark, and $\boldsymbol{S}$ is the stego-object. The + operator embeds the watermark $\boldsymbol{M}$ into the cover work $\boldsymbol{W}$, producing the stego-object $\boldsymbol{S}$. (1.1)

The properties of watermarks used for authentication imply the following:

$$\boldsymbol{S} - \boldsymbol{M'} = \boldsymbol{W'}, \boldsymbol{W'} \cong \boldsymbol{W} \text{ and } \boldsymbol{M'} \cong \boldsymbol{M}, \quad (1.2)$$

where $\boldsymbol{S}$ is the stego-object, $\boldsymbol{M'}$ is the watermark to be stripped off from $\boldsymbol{S}$, $\boldsymbol{W'}$ is the object with the $\boldsymbol{M'}$ stripped off. Theoretically, $\boldsymbol{W'}$ cannot be the same as $\boldsymbol{W}$ since most watermarking algorithms permanently change $\boldsymbol{W}$. However, invisible or imperceptible watermarks typically render an object that is perceived the same as the cover work in human eyes or ears. For this reason, the $\boldsymbol{W'}$ and $\boldsymbol{W}$ should be "perceived" as identical or similar. As (1.2) concerns watermarking for authentication, the main requirement is that the decoded watermark $\boldsymbol{M'}$