Copyright Protection in Virtual Communities through Digital Watermarking

Huayin Si
University of Warwick, UK

Chang-Tsun Li
University of Warwick, UK

INTRODUCTION

Although the development of multimedia processing techniques has facilitated the enrichment of information content, and the never-ending expansion of interconnected networks has constructed a solid infrastructure for information exchanges, meanwhile, the infrastructure and techniques have also smoothed the way for copyright piracy in virtual communities. As a result, the demand for intellectual property protection becomes apparent and exigent. In response to this challenge, digital watermarking has been proposed to serve this purpose.

The idea of digital watermarking is to embed a small amount of secret information—the watermark—into the host digital productions, such as image and audio, so that it can be extracted later for the purposes of copyright assertion, authentication and content integrity verification, and so forth. Unlike traditional watermarks printed on paper, which are visible to human eyes, digital watermarks are usually invisible and can only be detected with the aid of a specially designed detector. One characteristic distinguishing digital watermarking from cryptography, which separates the digital signature from the raw data/content, is that digital watermarking embeds the signature in the content to be protected. The superiority of this characteristic is that while cryptography provides no protection after the content is decrypted, digital watermarking provides "intimate" protection, because the digital signature/secret information has become an inseparable constituent part of the content itself after embedding. Because of the very characteristic, digital watermarking requires no secret channel for communicating the digital signature that cryptography does. So in the last decade, digital watermarking has attracted numerous attention from researchers and is regarded as a promising technique in the field of information security.

Various types of watermarking schemes have been developed for different applications. According to their natures, digital watermarking schemes could be classified into three categories: *fragile* watermarking, *semi-fragile* watermarking and *robust* watermarking. The schemes of the first two categories are developed for the purposes of multimedia authentication and content integrity verification, in which we expect the embedded watermark to be destroyed when attacks are mounted on its host media. More emphases of these schemes are placed on the capability of detecting and localizing forgeries and impersonations. The main difference between the two is that semi-fragile watermarking is tolerant to non-malicious operations, such as lossy compression within a certain compression ratio, while fragile watermarking is intolerant to any manipulations. Robust watermarking, on the other hand, is intended for the applications of copyright protection, wherein the watermarks should survive attacks aiming at weakening or erasing them provided the quality of the attacked content is still worth protecting. Therefore, the emphasis of robust watermarking schemes is placed on their survivability against attacks.

This article is intended to focus on robust watermarking schemes for the application of copyright protection. See Li and Yang (2003) and Lin and Chang (2001) for more details about fragile and semi-fragile schemes.

ROBUST WATERMARKING APPROACHES

Robust watermarking is applicable in the areas of copyright protection such as ownership identification/proof, copy control/copy prevention, fingerprinting/transaction tracking and so forth. Some common requirements for the robust watermarking schemes are:

- **Transparency:** The watermark should be invisible to human perception after embedded in the host media, so the impact on the perceptual quality is minimized.
- **Robustness:** Survivability against all kinds of malicious attacks and incidental manipulations, such as lossy compression and format trans-coding, should be maintained unless the manipulations have rendered the content useless in some sense.

- **Payload:** Payload (i.e., the embedding capacity) is important for applications such as "traitors" tracing. To trace the origin of pirated copies, unique secret information that identifies the recipient/buyer for each original copy has to be embedded when purchased. To avoid collusion of a number of buyers, such schemes should provide enough capacity to contain the information. Detailed treatment on collusion attack can be found in Trappe, Wu, Wang and Liu (2003).
- **Computing Complexity:** Complexity is expected to be low enough to enable online and real-time watermarking or detecting, especially for mobile devices without the aid of a computer.

These requirements are so conflicting that no watermarking scheme can provide a cure-all solution to fulfil all of them simultaneously. Researchers have to make trade-offs among the most important factors to suit the needs of the applications in question.

ROBUST WATERMARKING FOR IMAGE

From the perspective of an attacker, after attacking, the watermark should be removed from the image while the visual quality should not be significantly compromised. Therefore, for the attacker, the apparent way to attack the watermarked image is to modify the perceptually insignificant components greatly but to tamper the rest slightly, assuming the watermark is globally embedded. So to counter the attack, the designer may place the watermark in the perceptually significant components of the host images. However, just because the contents at the ideal positions to hide the watermark are so significant, even slight modification would become perceptible. Therefore, the designers have to make trade-offs between the robustness and transparency.

The basic idea of embedding watermark w in the original image C_o in order to create watermarked version C_w is essentially the addition of w to C_o using the conceptual model $C_w = w + C_o$. The robustness and transparency depends on the strength of watermark w. The characteristic of this model is the independency between the watermark and the host image. However, the image regions consisting of high-frequency components (i.e., regions with many details) can accommodate a higher degree of manipulations/distortions, while regions with low-frequency components (i.e., smoother regions without many details) tend to be less tolerant to manipulations. Without taking this phenomenon into account, the aforementioned conceptual model cannot be adopted in practice. Nevertheless, in order to make efficient use of network bandwidth and storage, lossy compression algorithms such as JPEG aiming at reducing redundancy and details by removing the high-frequency components, which are perceptually less significant, are widely adopted in multimedia applications. Therefore, embedding the watermark in the perceptually insignificant components is also not a feasible approach.

Aligning with these phenomenon and multimedia applications, Cox, Kilian, Leighton and Shamoon (1997) proposed the concept of Spread-Spectrum (SS) watermarking, which has inspired a great number of works in this field. Adopted from communication theory, the idea of SS watermarking is to treat the low-energy watermark as a narrow-band signal and spread it into multiple low-frequency components of the host image that is treated as a wide-band signal. By spreading the watermark throughout the significant spectrum, the watermark energy present in any single frequency is imperceptible, thus striking the balance between robustness and transparency elegantly. The principle of SS watermarking can be described as $C_w(i) = \alpha b_i w_i + C_o(i)$ where b_i is a pseudo-random wide-band noise sequence to spread the corresponding watermark bit w_i, and the modulated signal is mixed into the host signal with a weighting factor α used to balance the robustness and transparency of the watermark. However, high robustness of SS watermarking is gained at the expense of low payload (i.e., embedding capacity).

To further improve transparency, perceptual models based on the Human Visual System (HVS) have been proposed and incorporated in the watermark embedding process (Barni, Bartolini, & Piva, 2001). Feasible perceptual models facilitate adaptive watermark embedding in the components where HVS is less sensitive.

ROBUST WATERMARKING FOR VIDEO

Uncompressed video stream is composed of a sequence of still images that are called frames, so the classical approach is to use image watermarking schemes to mark every frame. However, such a frame-by-frame approach is at a high risk of attacks such as collusion (Doerr & Dugelay, 2004). Furthermore, uncompressed video takes huge storage. For commercial feasibility, video compression standards such as MPEGs have been developed based on the fact that adjacent frames are highly correlated. These standards break up frames and encode them into a block structure after removing the redundancy. The previous idea of utilizing the image watermarking directly is not applicable anymore. Moreover, the video industry