Digital Watermarks for Copyright Protection

NICHOLAS PAUL SHEPPARD'
REIHANEH SAFAVI-NAINI**
PHILIP OGBUNBONA***

Abstract

It is feared that the ease with which digital media can be copied will lead to a proliferation of copyright infringement. One proposed technical solution is digital watermarking, which embeds a hidden signal into host data that can be used in a variety of protocols that attempt to either prevent or deter copyright infringement. In this paper, we give a brief overview of digital watermarking and discuss some of the issues involved in providing effective digital watermarking systems for deterring copyright infringement.

1. Introduction

One of the great advantages of digital media over analogue media is that digital data can be reproduced easily and infinitely without any loss of fidelity. This boon, however, can also lead to a headache for copyright owners who wish to protect their works from unauthorised reproduction. According to high-profile copyright protection organisations such as SDMI and DVD-CCA, lack of protection from copyright violation is often given as a reason for the slow roll-out of digital media such as DVDs and digital television.

It is well-understood how encryption can protect digital data against access by unauthorised parties. Encryption scrambles the data according to a secret key in such a way that the usable data can only be recovered by a party knowing the key. This does nothing to prevent illicit copying or rebroadcasting by otherwise legitimate users, however, since they must know the secret key to be able to access the service.

---

* School of Information Technology and Computer Science, The University of Wollongong, NSW, 2522, Australia. E-mail: nps@uow.edu.au

** School of Information Technology and Computer Science, The University of Wollongong, NSW, 2522, Australia. E-mail: rei@uow.edu.au

*** Motorola Australia Research Centre, 12 Lord St, Botany, NSW, 2019, Australia. E-mail: pogunbon@arc.corp.mot.com

Digital Watermarks for Copyright Protection

Steganography is the science of hiding information in a host document in such a way that the hidden information can be recovered only by the intended recipient, and the host document remains useful for its conventional purpose. Digital watermarking is broadly used to describe the application of steganographic techniques to digital objects.

In this paper, we will focus upon digital watermarking for copyright protection. It has been proposed that hidden watermarks in digital objects can be used to control copying devices, prove copyright violation and trace unauthorised copies. Other applications for digital watermarking have been proposed in authentication, broadcast monitoring and captioning, which will not be described here.

We will give an introduction to the relevant technical aspects of digital watermarking, then describe proposed copyright protection mechanisms based on watermarking. We discuss the needs of watermarking systems that seek to deter copyright infringement, and discuss the limitations of watermarking technologies. Finally, we compare watermarking technologies to alternative mechanisms for copyright protection.

2. Digital Watermarking

A digital watermark is simply a signal to be embedded into a digital object which we will call the host. The watermark signal's strength is much lower than the strength of the host signal so that the host object retains its useful value after the watermark has been embedded.

Watermarks may be perceptible or imperceptible to a human observer. Most watermarking algorithms generate imperceptible watermarks that can only be detected by a machine. Imperceptible watermarks are less damaging to the host object, and are more secure against attackers wishing to destroy or manipulate the watermark. The watermarking approaches described in this paper invariably use imperceptible watermarks. Simple perceptible watermarks, however, have found use in the corners of network television footage, and a more sophisticated perceptible system is used by the Vatican Library.2

2.1 Example – LSB Embedding

A very simple imperceptible watermark, for example, can be embedded using LSB embedding, in which we replace the least significant (that is, rightmost) bit of each sample with a bit from some watermarking pattern. Figure 1 shows a 4x4 raster image with the least significant bit (of an 8-bit grey scale) highlighted.

1111111 1111110 1111101 1111100

Figure 1: A 4x4 host image.

We can think of our watermark as a binary bitmap pattern, that is, an image with only two colours which we will call '0' and '1' (which might be mapped to white and black for display purposes). The watermark pattern will have the same dimensions as the host image, and be embedded in the host image by replacing the least significant bit of each pixel with the corresponding bit in the watermark image. Figure 2 shows a 4x4 bitmap depicting the letter 'N' which we have embedded into the image of Figure 1 to form the watermarked image shown in Figure 3.

Figure 2: A 4x4 watermark bitmap.

A change in the least significant bit of a pixel is typically imperceptible to a human observer, but it is simple for a computer to detect the watermark by reading the values of the least significant bits. Figure 4 shows a host image and a version that has been watermarked using the bitmap shown in Figure 5.
Figure 4: (a) A host image and (b) its watermarked version.
2.2 Watermark Detection

In the context of copyright protection, we assume that we have a watermark detector (a computer programme or purpose-made machine) that, given a watermark pattern and an object, can test for the existence of the pattern in the object. Note that it is possible for a watermark pattern to exist in the host purely by chance -- it is easy to see that the pixel data of Figure 1 contains a "watermark" consisting of alternating 1s and 0s.

For this reason, the test is a statistical test. The detector returns a detection score that represents the detector's confidence that the pattern did not arise by chance. If the score is higher than a certain threshold, we assume that the watermark is present, otherwise we assume that it is absent.

Due to the statistical nature of the test, however, there remains a possibility that the detector will make a mistake. The detector may detect a watermark that was not inserted (a false positive), or fail to detect a watermark that was inserted (a false negative). Well-designed watermarking algorithms should minimise the number of false positives and false negatives, but there is a trade-off between the two and it is impossible to eliminate either in a non-trivial system.
2.3 Robustness

A watermark is said to be robust if its presence can still be detected after the host object has been manipulated in some way that does not damage the host beyond some acceptable level. The object may be manipulated in the course of innocent signal processing such as compression, or it may be manipulated maliciously by an attacker wishing to remove the watermark. For secure copyright protection, we require watermarks that are robust to any manipulation that results in an object that retains its commercial value.

The LSB embedding scheme is not robust. It is simple to embed another pattern into the least significant bits, erasing the original watermark without further damaging the image.

Many more sophisticated algorithms have been proposed that take advantage of the complexities of human perceptual systems to produce watermarks that cannot be so easily destroyed. Surveys of digital watermarking techniques are given by Swanson et al, Hartung and Kutter, and Langelaar et al.\(^3\)

3. Watermarking for Copyright Protection

3.1 Prevention

Various protocols have been proposed for building media players and recorders with built-in copy protection using watermarking.\(^4\) For example, a DVD writer could refuse to copy material in which it found a "no-copy" watermark.

For such systems to work, all consumer devices must have tamper-proof watermark detectors built into them. Even supposing that such devices were in use, a would-be pirate needs only to build his or her own device without a watermark detector, and such a task is not beyond the capabilities of professional pirates — or even of amateur hobbyists, as demonstrated by the DeCSS tool.\(^5\)

---


Copy prevention seems to be a very difficult technical, administrative and commercial challenge. Cracks of copy prevention systems are commonplace and many authors hold that copy prevention is futile\(^6\) or even potentially harmful.\(^7\) Nonetheless, devising copy prevention systems is a popular challenge with industry bodies such as SDMI, the DVD Forum and the 4C Entity. In this paper, however, we will focus upon mechanisms for deterrence.

3.2 Deterrence

3.2.1 Proof of Ownership

Suppose Alice creates an image and posts it on her web site. Bob takes a liking to the image, copies it and distributes it as his own. Since Bob's copy of the image is identical to Alice's, how can Alice prove that her copy is the original from which Bob made his copies?

If Alice had watermarked her image before she posted it on her web site, all of Bob's copies of the image would contain Alice's watermark (which is known only to Alice). Alice can then prove that Bob's copies descend from her original, since she can evidence the watermark but Bob cannot.

3.2.2 Fingerprinting

Suppose Alice sells a DVD containing the same film to each of Bob, Carol and Dave. Eve is found to have an unauthorised copy of the DVD, but cannot or will not reveal where she got it from. Since all of the copies of the DVD are identical, Alice has no way of knowing which of Bob, Carol or Dave made the copy.

Before manufacturing each DVD, Alice can generate an individual watermark for the DVD, called a fingerprint, and embed it into the DVD. By recording which customer has which fingerprint, Alice can trace the source of any unauthorised copies she finds by checking the fingerprint in the copy against her records.

4. Extensions to the Basic Watermarking Models

Unfortunately for would-be watermarkers, protecting one's intellectual property is not so simple as the basic models described above might suggest. A content creator cannot simply choose the (robust) watermarking scheme of his or her choice, embed some arbitrary watermark, and distribute the watermarked object safe in the knowledge that copyright infringers will be brought to justice. In

---


this section, we discuss implementation issues extending beyond the basic models described above.

4.1 Proof of Ownership

4.1.1 The Inversion Attack

Craver, et al.\(^8\) described a general procedure that defeated many, if not all, of the watermarking algorithms known at the time, known as the inversion attack. Using this procedure, it is possible for an imposter to create a counterfeit original object and watermark such that the imposter's watermark appears to be in the true original.

Any watermarking algorithm subject to the inversion attack, therefore, cannot be used to prove ownership since an invertible algorithm can be used to "prove" that anyone is the owner of an object.

Furthermore, since the test for the existence of a watermark pattern is statistical, it may be feasible for an imposter to repeatedly generate watermarks until he or she happens upon a watermark pattern with a detection score high enough to claim ownership.\(^9\) Hence, watermarkers should not be permitted to choose arbitrary watermark patterns -- either the pattern must be registered with a trusted authority, or patterns must be restricted to those that have some meaning to a human observer.

4.1.2 Detecting Infringement

The basic watermarking model does not consider the problem of detecting copyright infringement in the first place; it deals only with producing evidence against someone suspected of infringement. However, ownership watermarks can be used in an extended capacity by mechanisms that actively search for copyright violations.

A web spider is a piece of software that searches the Internet by examining World Wide Web pages and following the links on them to find more pages, and more links to follow. Such a piece of software could search the Internet for watermarked data and report suspect pages to the author, or to some responsible authority. Similar monitoring practices can, in principle, be used to police other media.

Attempting to monitor the entire Internet, as well as any other medium that could carry suspect data, seems a Herculean task and it

---

\(^8\) Craver, S., Memon, N., Yeo, B.L. and Yeung, M.M., 'Resolving rightful ownerships with invisible watermarking techniques', *IEEE Journal on Selected Areas in Communications*, 16, 1998, pp 573-586.

is not immediately clear that such a process is realistic. Nonetheless, even if spiders can find only a small fraction of copyright infringers, this may provide at least some discouragement to potential pirates.

4.2 Fingerprinting

4.2.1 Customers' Rights

Pfitzmann and Schunter\textsuperscript{10} point out that since the fingerprint is inserted by the seller alone, "unauthorised" copies can be produced either by an internal security breach in the seller, or by a malicious seller attempting to frame the buyer.

Qian and Nahrstedt\textsuperscript{11} note that this problem could be solved by requiring that a trusted third party be responsible for embedding, recording and verifying fingerprints. However, such an arrangement makes a lot of work for the third party. Furthermore, it is unclear where such a hard-working, trustworthy third party might be found.

Instead, a number of buyer-seller protocols have been proposed to prevent such a mishap.\textsuperscript{12} In these protocols, the buyer and seller exchange watermark information in such a way that the seller cannot reproduce the buyer's watermark.

These protocols require that the objects be manufactured on demand, which does not seem realistic for current distribution mechanisms based on physical media, but may become more realistic for any future network-based distribution.

4.2.2 Collusion and Framing

Suppose a party has possession of multiple copies of a fingerprinted object; for example, by simply buying several legitimate copies. Such a party can observe the differences between the different fingerprinted copies, and exploit these observations to create a new copy that does not contain any of the fingerprints contained in the original copies.

The party may try to remove the fingerprints altogether, or it may try to produce a copy containing a fingerprint belonging to an innocent


party. Furthermore, if the algorithm for watermarking is known or can be discovered, a pirate may attempt to embed the fingerprint of an innocent party into a single copy and thus frame the innocent party.

Clearly, if it is easy for pirates to frame legitimate buyers or otherwise construct copies containing spurious fingerprints, fingerprints are worthless as a method of copyright protection.

A collision-secure code is an encoding for fingerprints such that a fingerprint formed by combining fingerprints from the code contains sufficient information to identify at least one of the original fingerprints. Collusion-secure codes were introduced by Boneh and Shaw; other collusion-secures codes are proposed by Biehl and Meyer, Guth and Pfitzmann, and Dittmann et al. However, these codes grow impractically long very rapidly.

4.2.3 Privacy

The basic fingerprinting model assumes that the seller maintains a list of which buyers bought which products. Not all consumers may be pleased by such an arrangement.

Pfitzmann and Waidner propose an anonymous fingerprinting protocol that makes use of a registration centre (such as the buyer’s bank) to provide anonymity to the buyer. The seller can only identify a buyer if he or she obtains a copy of the product sold to the buyer, and co-operates with the registration centre. Other anonymous fingerprinting protocols are proposed by Pfitzmann and Sadeghi.

As for the solutions to the customers’ rights problems, anonymous fingerprinting requires a buyer-seller protocol and for products to be manufactured on demand.

---


5. Limitations of Watermarking

Watermarking technology cannot make the digital proof against copyright ownership disputes and piracy, even if we leave aside the technical and administrative problems that exist in current technologies. In this section, we discuss the limitations of watermarking technologies in providing copyright protection.

5.1 Imitations

Watermarking provides a mechanism for showing the existence of a copy made by an unintelligent mechanical process, such as photocopying, or duplication by a computer. We will refer to such unintelligent reproductions as direct copies.

We assume that watermarking leaves the semantic value of the data unchanged. Hence, watermarking does not, and cannot, provide any protection against more intelligent copying processes that extract the semantic meaning of a work and reproduce it independently. It is simple, for example, for musicians to copy a piece of music by playing it on their instruments. No watermark as we understand it could survive such a process, which we will refer to as imitation.

While imitation clearly requires more effort and typically greater skill than direct copying, it does not seem unreasonable to believe that potential copyright infringers might have ready access to the effort and skill required. Someone claiming to have written and performed a song is presumably a musician, for whom it is no trouble to play the music involved. Companies that might infringe on designs for buildings, machinery, etc. have ready access to draftsmen on their staff for whom it is no trouble to draw a diagram.

On the other hand, however, not all "imitation" is infringement even if direct copying might be. If Alice were to take a photograph of the Sydney Opera House from the top of the Harbour Bridge, for example, it is not an infringement for Bob to go to the top of the Harbour Bridge and take his own photograph, while it remains illegal for Bob to make unauthorised direct copies of Alice's photograph. Watermarking can distinguish between the two cases.

5.2 Before Publication

Current watermarking algorithms assume that the object to be watermarked is in its final, publishable state. However, ownership disputes can and do erupt at a much earlier stage in the creative process, particularly in cases where creation of the object involves more than one person.

It seems very ambitious, at least in the foreseeable future, to consider watermarking algorithms that can resolve issues of what amounts to a "creative contribution" by some participant in the creative process. Such algorithms would imply an algorithmic model of creativity that, so far as the present author is aware, does not exist.
A somewhat less ambitious task would be to consider an "incremental watermark" that existed in an object throughout the creative process. Such a watermark could, for example, protect an object against theft from the artists' studio and possibly provide some limited protection against disputes between co-workers. Such watermarks are a subject of further research.

5.3 Watermark This!

If a watermark is to meet imperceptibility requirements, it must be assumed that the host object contains some sort of "empty space" into which the hidden signal can reside, that is, redundant or insignificant portions of the object. Audio-visual data, on which most watermarking research has focused, contains much information that is imperceptible to a human listener or viewer and so a watermark can be hidden by manipulating this imperceptible data without significantly affecting the host data.

This, however, is not true of all data. Text, for example, contains little or no imperceptible information; even very minute changes to the text are readily apparent to a human reader.

Nonetheless, some determined researchers have developed watermarking algorithms for text17 that embed watermarks in text by making minute changes in line and word spacing. These watermarks are trivial to delete by re-watermarking the text and so seem of little or no realistic use for copyright protection. Other watermarks have been proposed that re-arrange the phrasing of the document18, but it is unclear what authors would make of a computer's version of their literary masterpiece.

Similar problems arise for simple designs such as company logos, flags and icons which do not contain sufficient host information in which to hide a watermark.

5.4 Fingerprint Transferral

Determining the owner of a fingerprint on an unauthorised copy of an object is not enough to conclude that the fingerprint's owner is guilty, since the copy may not have been made with the fingerprint owner's permission. For example, a buyer's legitimate copy may have


been stolen, then used by the thieves as the base for making pirate copies, or the legitimate copy may have been on-sold to a second owner who made the illegal copies.

Nonetheless, attaching the unauthorised copy to the identity of a buyer may be helpful in investigating suspects even if the legitimate buyer is not him- or herself the guilty party. If thieves have been apprehended for the theft of some fingerprinted digital recording, they can be further investigated concerning the origin of unauthorised copies.

Hence, fingerprinting may be useful as an investigative tool, but it does not seem to provide the conclusive proof of guilt that the watermarking literature tends to imply.

6. Watermark Administration

A number of observations in the foregoing sections imply that watermarking must be regulated for it to be an effective tool for copyright protection. Unlike prevention methods such as encryption and copy prevention, ownership watermarks and fingerprints must be able to satisfy an independent arbiter (who may not be a technical person) that they do, indeed, prove something.

An ownership watermark must be known (or reasonably believed) to be secure against inversion attacks, and watermark patterns must conform to some acceptable standards. Fingerprints must similarly be secure against framing and seller-side leaks.

So far as the present author is aware, no organisation yet exists to perform the necessary function of approving watermarking software and patterns for use. The European Union's TALISMAN Project\(^\text{19}\) appears to have been working in this direction, but, though the project proposed one watermarking algorithm, TALISMAN's Final Report ultimately states that

Watermarkers [i.e. vendors of watermarking software] have contracts with Copyright Owners. The Copyright Owner has or has not to go to a Copyright Authority to get certification, depending on upon the standardisation and regulations progresses.\(^\text{20}\)

In view of the observations in this section, until such "Copyright Authorities" exist, or existing organisations develop organs to deal


with these issues, it is unclear whether or not watermarking could be used to reliably resolve real-world copyright ownership disputes.

7. Alternatives to Watermarking

7.1 Copyright Registration

In order to ensure that one's work is recognised as being the original, the publication of the work can be made a matter of indisputable public record by registering the work with some trusted authority such as a collecting society. Explicit registration is often unnecessary since publication of a work is made a matter of public record by broadcast media. This process is referred to as timestamping by computer scientists, and (copyright) registration by everyone else.

Registration appears to have been an effective mechanism for protecting the ownership rights for non-digital media, and registration services continue to serve digital media.\(^{21}\) There is no apparent doubt in the public mind as to the origin of popular songs, films, books, etc., and even considerably more obscure works such as theme songs for television programmes are protected by the relevant collecting societies. Of course, disputes have arisen just as in any area of law, but the sky does not appear to be falling, even though digital media have been with us for many years now.

Registration overcomes some of the limitations of watermarking. It is much more flexible and adaptable than watermarking since anything that can be copied can be archived by a registration service. No amount of manipulation of a copy of an object can cause it to become de-registered in the way a watermark can be erased.

However, registration seems more expensive than watermarking, and authors may find formal registration more cumbersome than using a watermarker on their own computer. While the need to register watermarks may detract somewhat from the simplicity of using watermarks, watermark registration need only be a one-off process.

7.1.1 Imitations vs. Direct Copies

Unlike watermarking, registration makes no distinction between a direct copy and an imitation. This is useful where making an unauthorised imitation is illegal, as is typically the case for music, architecture, designs for machinery, and so on.

However, watermarking is able to distinguish between an (illegal) direct copy and a (legal) imitation. The Times reports a watermarking-like case, for example, in which the UK Automobile Association was

---

caught plagiarising Ordnance Survey maps. Ordnance Survey apparently inserted small "errors" into its maps that turned up in the Automobile Association's maps, showing that the latter had made a direct copy of the former's maps rather than an "imitation" of them. Without the watermark, the Automobiles Association's copied maps would (presumably) have been indistinguishable from maps drawn up by its own cartographers.

7.1.2 Making the Judgement

Watermarking provides a mechanism for easy, automated copy detection, whereas registration has typically required an expert human judge to compare the registered and suspect objects for their similarity or otherwise.

However, Relatable claims to have developed a product called TRM that achieves automated copy detection using registration. TRM is reportedly being considered as a solution for the recent copyright troubles of Napster. A central server is provided with a list of "fingerprints" (not the same "fingerprints" as used in watermarking literature, and in the rest of this paper) of copyrighted songs, and any song transferred through Napster's song-exchange service is compared with the fingerprint database to detect copyrighted transfers. We do not know how this product works, however, and nor are we aware of any public testing of its effectiveness.

8. Conclusion

We have described digital watermarking for copyright protection, and discussed some issues affecting development of effective watermarking systems for copyright protection.

Digital watermarking is still an immature technology. This is demonstrated in a technical sense by the experience of the Secure Digital Music Initiative in its hacking challenge, and by the success

---

of the StirMark benchmarking software in defeating watermarking algorithms.  

We have also outlined what we view as the administrative challenges remaining for copyright protection systems using watermarks. Companies such as DigiMarc, MediaSec and Verance are currently marketing watermarking solutions for a variety of applications including copyright protection, but the necessary infrastructure does not yet exist to support reliable copyright protection using watermarks. With no recognised independent mechanism for proving the efficacy of watermarking systems, copyright protection with watermarking is at best unproven, and in at least some cases, broken.

9. Acknowledgements

The authors would like to thank members of the Watermarking Mailing List for helpful discussions on this subject. This research is partially supported by the Motorola Australian Research Centre.

Bibliography

Advanced Communications Technologies and Services. AC019 TALISMAN.

J. Bale: "Twists in the plot cost AA map cheats £20m". The Times, page 1, 6 March 2001.


---


F. Mintzer, A. Cazes, F. Giordano, J. Lee, K. Magerlein, and F. Schiatterella: "Capturing and preparing images of Vatican Library


