Published in the Int. Conf. on Information Technology'2000, Las Vegas, 2000

Watermarking applications and their properties

Ingemar J. Cox, Matt L. Miller and Jeffrey A. Bloom
NEC Research Institute
4 Independence Way
Princeton, NJ 08540
ingemar,mlm,bloom@research.nj.nec.com

Abstract

We describe a number of applications of digital watermarking and the examine the common properties of robustness, tamper resistance, fidelity, computational cost and false positive rate. We observe that these properties vary greatly depending on the application. Consequently, we conclude that evaluation of a watermarking algorithm is difficult without first indicating the context in which it is to be applied.

1. Introduction

Watermarking methods are often evaluated based on the common properties of robustness, tamper resistance, and fidelity. However, examination of these properties without careful consideration of the application can often be misleading. A watermark designed to serve security needs of the CIA must meet different requirements than one intended for annotating home video. Thus, it is inappropriate to evaluate these two watermarks according to the same standards.

In this paper, we examine how the requirements for watermarking can vary with application. Section 2 briefly describes eight existing and proposed applications of watermarks. This is followed, in Section 3, with a discussion of several properties of watermarking systems, and how their definition and importance depend on the application.

2 Applications

This section describes seven applications of watermarking: broadcast monitoring, owner identification, proof of ownership, authentication, transactional watermarks, copy control and covert communication.

2.1 Broadcast monitoring

In 1997, a scandal broke out in Japan regarding television advertising. At least two stations had been routinely overbooking air time. Advertisers were paying for thousands of commercials that were never aired [17]. The practice had remained largely undetected for over twenty years, in part because there were no systems in place to monitor the actual broadcast of advertisements.

There are several types of organizations and individuals interested in broadcast monitoring. Advertisers, of course, want to ensure that they receive the air time purchased from broadcasting firms. Musicians and actors want to ensure that they receive accurate royalty payments for broadcasts of their performances.[1] And copyright owners want to ensure that their property is not illegally rebroadcast by pirate stations.

We can use watermarks for broadcast monitoring by putting a unique watermark in each video or sound clip prior to broadcast. Automated monitoring stations can then receive broadcasts and look for these watermarks, identifying when and where each clip appears. Commercial systems have been deployed for a number of years and the basic concepts have a long history [16, 3, 20, 12, 4].[2]

2.2 Owner identification

Although a copyright notice is no longer necessary to guarantee copy rights, it is still recommended. The form of the copyright notice is usually "©*date, owner*". On books and photographs, the copyright is placed in plane sight. In movies, it is appended to the end of the credits. And on prerecorded music, it is placed on the packaging.

One disadvantage of such text copyright notices is that they can often be removed from the protected material.

[1] A recent spot-check by the Screen Actor's Guild found an average of $1000 in underpaid royalties per hour of US television programming [2].

[2] The earliest reference we have found [16] is assigned to the Muzak Corporation, famous for providing "elevator music", and may be the source of the many rumors that Muzak contained subliminal messages.

Packaging can be lost, movies can have the credits cut off, and images can be spatially cropped. A digital watermark can be used to provide complementary copyright marking functionality because it becomes an integral part of the content, i.e. the copyright information is embedded in the music to supplement the text notice printed on the packaging.

The Digimarc corporation has marketed a watermarking system designed for this application. Their watermark embedder and detector are bundled with Adobe's popular image processing program, Photoshop. When the detector finds a watermark, it contacts a central database to identify the watermark's owner (who must pay a fee to keep the information in the database).

2.3 Proof of ownership

Multimedia owners may want to use watermarks not just to *identify* copyright ownership, but to actually *prove* ownership. To illustrate the problem, let's quickly introduce some characters who are well known in the watermarking literature. Suppose Alice creates an image and puts it on her website, with a copyright notice "©Alice 2000". Bob then steals the image, uses an image processing program to replace the copyright notice with "©Bob 2000", and then claims to own the copyright himself. How can the dispute resolved?

Traditionally, Alice could register the image with the Copyright Office by sending a copy to them. The Copyright Office archives the image, together with information about the rightful owner. When the dispute between Alice and Bob comes up, Alice contacts the Copyright Office to obtain proof that she is the rightful owner. If Alice did not register the image, then she should at least be able to show the film negative. However, with the rapid acceptance of digital photography, there might never have been a negative.

In theory, it is possible for Alice to use a watermark embedded in the image to prove that she owns it. However, this is not a trivial problem, as Craver *et al* [11] have noted.

2.4 Authentication

As both still and video cameras increasingly embrace digital technology, the ability for undetectable tampering also increases. The content of digital photographs can easily be altered in such a way that it is very difficult to detect what has been changed. In this case there is not even an original negative to examine. There are many applications where the veracity of an image is crucial, especially in legal cases and medical imaging.

Authentication is a well studied problem in cryptography [23]. Friedman [13, 14] first discussed its application to create a "trustworthy camera" by computing a cryptographic signature that is associated with an image. If even one bit of one pixel of the image is modified, it will no longer match the signature, so any tampering can be detected. However, this signature is metadata that must be transmitted along with the photograph, perhaps in a header field of a particular image format. If the image is subsequently copied to another file format that does not contain this header field, the signature will be lost, and the image can no longer be authenticated.

A preferable solution is to embed the signature directly into the image using watermarking. This eliminates the problem of ensuring that the signature stays with the image. It also opens up the possibility that we can learn more about what tampering has occurred, since any changes made to the image will also be made to the watermark. Thus, there are several systems that can indicate the rough location of changes that have been made to the image. There are also systems designed to allow certain changes, such as JPEG compression [18, 19], and only disallow more substantial changes, such as removing an individual from a crime scene.

2.5 Transactional watermarks (Fingerprinting)

Monitoring and owner identification applications place the same watermark in all copies of the same content. However, electronic distribution of content allows each copy distributed to be customized for each recipient. This capability allows a unique watermark to be embedded in each individual copy. Transactional watermarks, also called fingerprints, allow a content owner or content distributor to identify the source of an illegal copy. This is potentially valuable both as a deterrent to illegal use and as a technological aid to investigation.

One possible application of transactional watermarks is in the distribution of movie dailies. During the course of making a movie, the result of each day's photography is often distributed to a number of people involved in its production. These dailies are highly confidential, yet occasionally, a daily is leaked to the press. When this happens, studios quickly try to identify the source of the leak. Clearly, if each copy of the daily contains a unique transactional watermark that identifies the recipient, then identification of the source of the leak is much easier.

Another application of transactional watermarks was deployed by the DiVX corporation. DiVX marketed a modified version of DVD. One of the security measures implemented in DiVX hardware was a transactional watermark that could be used to identify a player used for piracy. If illegal copies of a DiVX movie turned up on the black market, DiVX could use the watermark to track them to the source.

2.6 Copy Control

Transactional watermarks as well as watermarks for monitoring, identification, and proof of ownership do not *prevent* illegal copying. Rather, they serve as powerful deterrents and investigative tools. However, it is also possible for recording and playback devices to react to embedded signals. In this way, a recording device might inhibit recording of a signal if it detects a watermark that indicates recording is prohibited. Of course, for such a system to work, all manufactured recorders must include watermark detection circuitry. Such systems are currently being developed for DVD video [6] and for digital music distribution [5]. Interestingly, the use of watermarks in video to control equipment dates back to at least 1989 [7] and in audio to perhaps 1953 [24].

2.7 Covert communication

One of the earliest applications of watermarking, or more precisely, data hiding, is as a method of sending secret messages. The application has been formulated by Simmons [22] as "the prisoner's problem", in which we imagine two prisoners in separate cells trying to pass messages back and forth. Their problem is that they cannot pass these messages directly, but rather, must rely on the prison warden to act as a messenger. The warden is willing to carry innocuous messages between them, but will punish them if he finds that, for example, their messages relate to a plan for escape. The solution is to disguise the escape-plan messages by hiding them in innocuous messages. There are several commercially available programs designed for this application, including StegoTools [1].

3 Properties

There are a number of papers that have discussed the characteristics of watermarks [9, 21, 15, 25]. Some of the properties discussed are robustness, tamper resistance, fidelity, computational cost, and false positive rate. In practice, it is probably impossible to design a watermarking system that excels at all of these. Thus, it is necessary to make tradeoffs between them, and those tradeoffs must be chosen with careful analysis of the application. In addition, the application can affect the very definition of a property.

In the following subsections, we look at each of the five properties listed above, and discuss how its importance and definition varies with application.

3.1 Robustness

A watermark is said to be robust if it survives common signal processing operations such as digital-to-analog-to-digital conversions and lossy compression. More recently, there has been an increased concern that video and still image watermarks also be robust to geometric transformations.

Robustness is often thought of as a single-dimensional value, but this is incorrect. A watermark that is robust against one process may be very fragile against another. In many applications, robustness to all possible processing is excessive and unnecessary.

Usually, a watermark must survive common signal processing only between the time of embedding and the time of detection. For example, in television and radio broadcast monitoring, the watermark need only survive the transmission process. For television, this means lossy compression, analog transmission, and some small amount of horizontal and vertical translation. It need not survive rotation, scaling, high-pass filtering, or any of a wide variety of distortions that do not occur during broadcast.

In some cases, robustness may be completely irrelevant, or even undesirable. Watermarks used for covert communication need not be robust at all, if the cover media will be transmitted digitally without compression. A watermark for simple authentication, which just indicates whether the media has been altered, should be fragile.

On the other hand, when the signal processing between embedding and detection is unpredictable, the watermark may need to be robust to every conceivable distortion. This is the case for owner identification, proof of ownership, fingerprinting, and copy control. It is also true for any application in which hackers might want to remove the watermark.

3.2 Tamper resistance

Tamper resistance refers to a watermarking system's resistance to hostile attacks. There are several types of tamper resistance. Depending on the application, certain types of attacks are more important than others. In fact, there are several applications in which the watermark has no hostile enemies, and tamper resistance is irrelevant. Some basic types of attack are

- *Active* attacks. Here the hacker tries to remove the watermark or make it undetectable. This type of attack is critical for many applications, including owner identification, proof of ownership, fingerprinting, and copy control, in which the purpose of the mark is defeated when it cannot be detected [10]. However, it is not a serious problem for authentication or covert communication.

- *Passive* attacks. In this case, the hacker is not trying to remove the watermark, but is simply trying to determine whether a mark is present, i.e. is trying to identify a covert communication. Most of the scenarios above are not concerned with this type of attack. In fact, we

might even advertise the presence of the mark so that it can serve as a deterrent. But for covert communication, our primary interest is to prevent the watermark from being observed.

- *Collusion* attacks. These are a special case of active attacks, in which the hacker uses several copies of one piece of media, each with a different watermark, to construct a copy with no watermark [8]. Resistance to collusion attacks can be critical in a fingerprinting application, which entails putting a different mark in each copy of a piece of media. However, the number of copies that we can expect the hacker to obtain varies greatly from application to application. For example, in the DiVX application, a hacker can buy any number of DiVX players, and play one movie on all of them to obtain any number of differently-watermarked copies. On the other hand, in the film-studio dailies application, each employee can only obtain one copy of the watermarked material. A collusion attack would require that several employees conspire to steal the material, which is an unlikely prospect.

- *Forgery* attacks. Here, the hacker tries to embed a valid watermark, rather than remove one. These are our main security concern in authentication applications, since, if hackers can embed valid authentication marks, they can cause the watermark detector to accept bogus or modified media. In addition, as pointed out by Craver *et al* [11], this type of attack is a serious concern in proof of ownership.

3.3 Fidelity

A watermark is said to have high fidelity if the degredation it causes is very difficult for a viewer to perceive. However, it only needs to be imperceptible at the time that the media is viewed. If we can be certain that the media will be seriously degraded before it is viewed, we can rely on that degradation to help mask the watermark. Such a case occurs when we watermark video that will be transmitted over NTSC, or audio that will be transmitted over AM radio. The quality of these broadcast technologies is so low that our initial fidelity need not be very good. Conversely, in HDTV and DVD video and audio, the signals are very high quality, and require much higher fidelity watermarks (though, of course, the quality of the content remains the same - a bad movie is a bad movie whether on VHS or DVD).

In some applications, we can accept mildly perceptible watermarks in exchange for higher robustness or lower cost. For example, Hollywood dailies are not finished products. They are usually the results of poor transfers from film to video. Their only purpose is to show those involved in a film production the raw material that has been shot so far. A small visible distortion caused by a watermark will not diminish their value.

3.4 Computational cost

Different applications require the embedders and detectors to work at different speeds. In broadcast monitoring, both embedders and detectors must work in (at least) real time. The embedders must not slow down the media production schedule, and the detectors must keep up with real-time broadcasts. On the other hand, a detector for proof of ownership will be valuable even if it takes days to find a watermark. Such a detector will only be used during ownership disputes, which are rare, and its conclusion about whether the watermark is present is important enough that the user will be willing to wait.

Furthermore, different applications require different numbers of embedders and detectors. Broadcast monitoring typically requires a few embedders and perhaps several hundred detectors at different geographic locations. Copy control applications may need only a handful of embedders but millions of detectors. Conversely, in the fingerprinting application implemented by DiVX, in which each player embeds a distinct watermark, there would be millions of embedders and only a handful of detectors. In general, the more numerous a device needs to be for a given application, the less it must cost.

The wide variation in dollar cost and in speed requirements means that there is a wide variation in the required computational efficiency of watermark embedders and detectors.

3.5 False positive rate

A false positive is a detection of a watermark in a piece of media that does not actually contain that watermark. When we talk of the false positive rate, we refer to the number of false positives we expect to occur in a given number of runs of the detector. Equivalently, we can discuss the probability that a false positive will occur in any given detector run. There are two subtly different ways to define this probability, that are often confused in the literature. They differ in whether the watermark or the media is considered to be the random variable.

In the first definition, the probability of a false positive is the probability that, given a fixed piece of media and a randomly-selected watermark, the detector will report that the watermark is in the media. The watermarks are drawn from a distribution that is defined by the design of a watermark generation system. Typically, watermarks are generated by either a bit-encoding algorithm or by a Gaussian, independent random number generator. In many cases, probability of false positives, according to this first definition,

is actually independent of the piece of media, and depends *only* on the method of watermark generation.

In the second definition, the probability of a false positive is the probability that, given a fixed *watermark* and a randomly-selected *piece of media*, the detector will detect the watermark in the media. The media is chosen from the distribution of natural media, which is defined by either nature or Hollywood, depending on the application. This distribution is very different from that defined by the watermark generation system, and thus probabilities based on this definition can be quite different from those based on the first definition.

In most applications, we are more interested in the second definition of false positive probability than in the first. However, in a few cases, the first definition is also important, such as in the case of fingerprinting, where the detection of a random watermark in a given image might lead to a false accusation of theft.

The probability of false positives that is required depends on the application. In the case of proof of ownership, the detector is used so rarely that a probability of 10^{-6} should suffice to make false positives unheard of. On the other hand, in the copy control application, millions of watermark detectors are constantly being run on millions of pieces of media all over the world. If one piece of unwatermarked media consistently generates false positives, it could cause serious trouble. For this reason, the false positive rate should be infinitesimal. For example, the general consensus is that watermark detectors for DVD video should have a false positive rate of 1 in 10^{12} frames [6].

4 Summary and conclusion

Watermarking is a technology that can serve a wide variety of applications, each of which may have very different requirements. Each application dictates a different tradeoff between the properties of robustness, tamper resistance, fidelity, and false positive rate. Moreover, for many of these properties, their very definitions can be dependent on application.

We conclude that a single set of standards should not be applied to all proposed watermarking systems. Instead, a separate set of standards should be applied to each system according to the application for which it is intended. One size does not fit all.

References

[1] http://www.informatik.tu-muenchen.de/ stowasse/security.html.

[2] http://www.sag.org/pressreleases/prla990826_spotchecks.html, 1999.

[3] C. R. Abbey and H. H. Pursel. Data channel monitor. *United States Patent*, (3,415,947), 1968.

[4] D. E. H. amd C. M. Solar. Automatic monitor for programs broadcast. *United States Patent*, (4,025,851), 1977.

[5] A. E. Bell. The dynamic digital disk. *IEEE Spectrum*, 36(10):28–35, 1999.

[6] J. A. Bloom, I. J. Cox, T. Kalker, J.-P. Linnartz, M. L. Miller, and B. Traw. Copy protection for DVD video. *Proceedings of the IEEE*, 87(7):1267–1276, 1999.

[7] R. S. Broughton and W. C. Laumeister. Interactive video method and apparatus. *United States Patent*, (4,807,031), 1989.

[8] I. Cox, J. Kilian, F. T. Leighton, and T. Shamoon. Secure spread spectrum watermarking for multimedia. *IEEE Trans. on Image Processing*, 6(12):1673–1687, 1997.

[9] I. Cox and M. L. Miller. A review of watermarking and the importance of perceptual modeling. In *Proceedings of SPIE, Human Vision & Electronic Imaging II*, volume 3016, pages 92–99, 1997.

[10] I. J. Cox and J.-P. Linnartz. Some general methods for tampering with watermarks. *IEEE Trans. on Selected Areas of Communications*, 16(4):587–593, 1998.

[11] S. Craver, N. Memon, B.-L. Yeo, and M. Yeung. Resolving rightful ownerships with invisible watermarking techniques: Limitations, attacks and implications. *IEEE Trans. on Selected Areas of Communications*, 16(4):573–586, 1998.

[12] M. G. Crosby. Communication including submerged identification signal. *United States Patent*, (3,845,391), 1974.

[13] G. L. Friedman. The trustworthy camera: restoring credibility to the photographic image. *IEEE Trans. On Consumer Electronics*, 39(4):905–910, 1993.

[14] G. L. Friedman. Digital camera with apparatus for authentication of images produced from an image file. *U.S. Patent*, (5,499,294), 1996.

[15] F. Hartung and M. Kutter. Multimedia watermarking techniques. *Proceedings of the IEEE*, 87(7):1079–1107, 1999.

[16] E. F. Hembrooke. Identification of sound and like signals. *United States Patent*, (3,004,104), 1961.

[17] D. Kilburn. Dirty linen, dark secrets. *Adweek*, 1997.

[18] C.-Y. Lin and S.-F. Chang. A robust image authentication algorithm surviving jpeg compression. In *SPIE Storage and Retireval of Image/Video Databases*, 1998.

[19] C.-Y. Lin and S.-F. Chang. Issues and solutions for authenticating mpeg video. In *Proc. IS&T/SPIE Syposium on Electronic Imaging: Science and Technology (EI'99) - SPIE Security amd Watermarking of Multimedia Contents*, 1999.

[20] T. Ohsawa and M. Karita. Automatic telecasting or radio broadcasting monitoring system. *United States Patent*, (3,760,275), 1973.

[21] F. A. P. Petitcolas, R. Anderson, and M. G. Kuhn. Information hiding - a survey. *Proceedings of the IEEE*, 87(7):1062–1077, 1999.

[22] G. J. Simmons. The prisoners' problem and the subliminal channel. In *Proc. CRYPTO'83*, pages 51–67. Plenum Press, 1984.

[23] D. R. Stinson. *Cryptography: Theory and Practice*. CRC Press, 1995.

[24] W. M. Tomberlin, L. G. MacKenzie, and P. K. Bennett. System for transmitting and receiving coded entertainment programs. *United States Patent*, (2,630,525), 1953.

[25] R. B. Wolfgang, C. I. Podilchuk, and E. J. Delp. Perceptual watermarks for digital images and video. *Proc. of the IEEE*, 87(7):1108–1126, 1999.