Hal Berghel

Watermarking Cyberspace

The use of watermarks is almost as old as paper manufacturing. Ancients poured their half-stuff slurry of fiber and water onto mesh molds to collect the fiber, then dispersed the slurry within deckle frames to add shape and uniformity, and finally applied great pressure to expel the water and cohere the fiber. This process hasn't changed too much in 2,000 years, even with the benefit of automation. One by-product of this process is the watermark—the technique of impressing into the paper a form, image, or text derived from the negative in the mold, as the paper fibers are squeezed and dried.

Paper watermarks have been in wide use since the late middle ages. Their earliest use seems to have been to record the manufacturer's trademark on the product so that the authenticity could be clearly established without degrading the aesthetics and utility of the stock. In more recent times, watermarks have been used to certify the composition of the paper, including the nature of the fibers used. Today, most developed countries also watermark their paper currencies and postage stamps to make forgery more difficult.
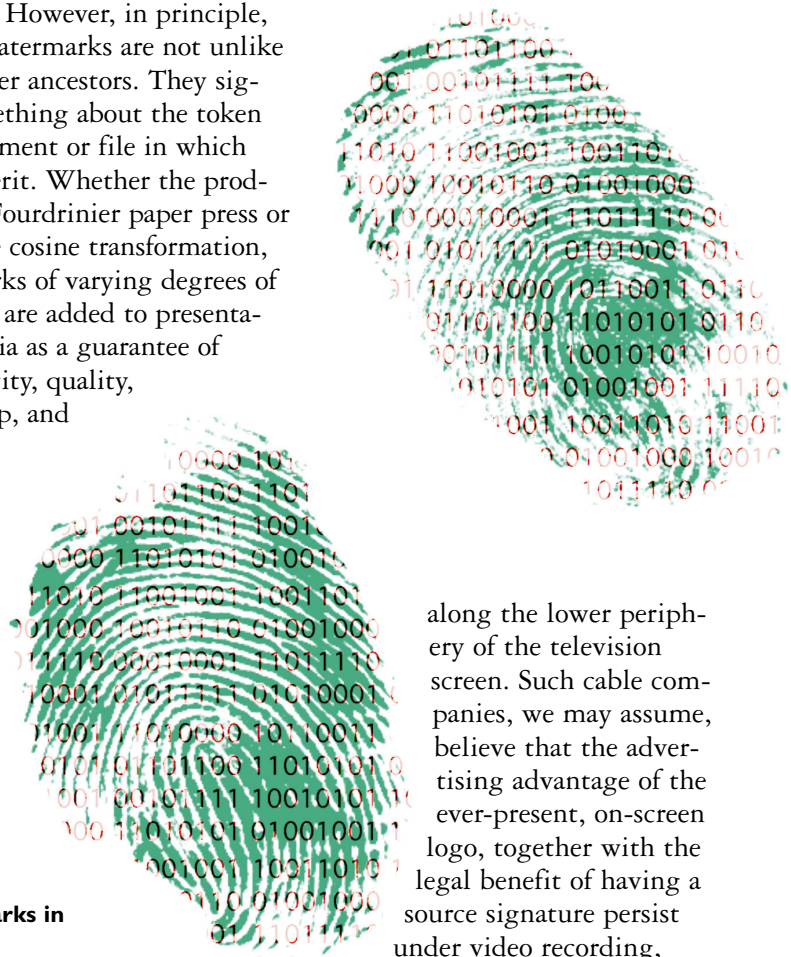
The digitization of our world has expanded our concept of watermarking to include immaterial, digital impressions for use in authenticating ownership claims and protecting proprietary interests. However, in principle, digital watermarks are not unlike their paper ancestors. They signify something about the token of a document or file in which they inherit. Whether the product of a Fourdrinier paper press or a discrete cosine transformation, watermarks of varying degrees of visibility are added to presentation media as a guarantee of authenticity, quality, ownership, and source.

Watermarks in Context

A digital watermark is a digital signal or pattern inserted into a digital document (text, graphics, multimedia presentations). As such, it is a form of electronic watermark much like the corporate logos used by the cable television industry to identify the source of the program, typically along the lower periphery of the television screen. Such cable companies, we may assume, believe that the advertising advantage of the ever-present, on-screen logo, together with the legal benefit of having a source signature persist under video recording, more than offset the aggregate user annoyance and distraction.

Digital watermarks extend these advantages to digital documents. A signal or pattern may be digitally imposed on a docu-

TERRY MIURA

VISIBLE WATERMARKS ARE A MORE OVERT

means of discouraging theft and unauthorized use both by

reducing the commercial value of a document and making it

obvious to the criminally inclined that the document's

ownership has been definitively established.

ment prior to sale or distribution. The persistence of the watermark under transmission, and some common forms of transformation, contribute to our ability to authenticate copies. This, in turn, should enable us to protect our ownership rights of digital information, even in the undisciplined, anarchistic world of the Internet (see Figure 1).

Before going into detail about what digital watermarking is, we'll first explain what it is not. Digital watermarking is not encryption, which also involves file transformation. It is a common practice nowadays to encrypt digital documents so that they become unviewable without a decryption key. Unlike encryption, however, digital watermarking leaves the original image or file basically intact and recognizable.

Further, decrypted documents are free of any residual effects of encryption, whereas visible digital watermarks are designed to be persistent in viewing, printing, or subsequent retransmission or dissemination.

Digital watermarking differs from digital fingerprinting, which produces a "meta" file that describes the contents of the source file. Cyclic redundancy checking and checksum algo-

Figure 1: Digitized copy of artwork from a 16th century Aztec manuscript. Note the circular digital watermark is most visible against a light background. Faint watermarks tend to hide in the intense, foreground imagery.

rithms are both simple uses of file fingerprinting for error detection applications. A more advanced use of fingerprinting is found in RSA Data Security's use of message digests for authentication purposes. Digests are the result of applying a hashing algorithm (for example, MD5, SHA) to a document or file to produce an identifying bit string (fingerprint). If the receiver's hash algorithm produces the same message digest for the file as the sender's, the file is authentic. Of course, this assumes that sender and receiver use the same software, hence the same hash algorithm.

Fingerprints may also serve

as digital signatures. If the message digest just discussed were further encrypted, converted to plain text, and attached to the original file or message in transit, the plaintext version of the message digest (fingerprint) would also serve as a digital signature for the original file. While both fingerprints and signatures accompany unaltered source documents, signatures, like their penned counterparts, are embedded in the document itself even if in encrypted form.

Watermarks in Use

Authentication is just one use of digital watermarking. Both symmetric and asymmetric hashing algorithms can be used to embed a unique digital imprint on a document or file. If the removal of an imprint yields the original document (which is to say that the stripped watermark is identical to the embedded watermark), then the copy is authentic. Once again, this assumes the stripping algorithm is available to the end user. Such authentication techniques are usually associated with some sort of encryption for the distribution of keys, programs, and so forth, which are related to the watermarked documents.

In addition, watermarks are also used as a check for non-repudiable duplication and transmission. In this case, the owner, creator, or sender imprints a watermark unique for each receiver. The watermark holds under subsequent re-transmission, so the authorized source of unauthorized copies may be easily identified after extraction. A collateral benefit is

that the intended recipient of a document token could always be identified.

However, these applications really only apply to the class of invisible watermarks. Visible watermarks (as in Figure 1) contribute to document and transmission security in different ways. To illustrate, visible watermarks are a more overt means of discouraging theft and unauthorized use both by reducing the
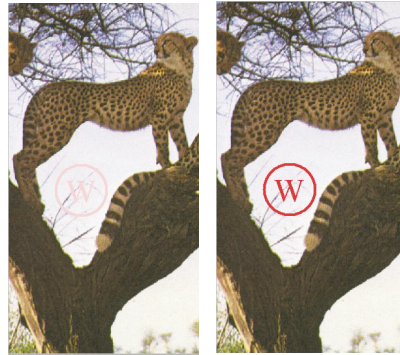
Figure 2. Two watermarked images identical but for the intensity of the image

commercial value of a document and making it obvious to the criminally inclined that the document's ownership has been definitively established. Invisible watermarks only have this effect if the digital thief is aware of the technology and the possibility that watermarks may be present on a document of interest.

There are several characteristics of effective watermarks. For one, they must be difficult or impossible to remove. For another, they must survive common document modifications and transformations such as cropping and compressing image files. They must also, in principle at

least, be easily detectable and removable by authorized users with such privileges (law enforcement agencies). Invisible watermarks should also be imperceptible, while visible watermarks should be perceptible enough to discourage theft but not perceptible enough to decrease the utility or appreciation of the document.

Watermarking Practice
Watermarking techniques tend to divide into two categories—text and image—according to the type of document to be watermarked. In the case of imagery, several different methods enable watermarking in the spatial domain from simply flipping low-order bits of selected pixels to superimposing watermark symbols over an area of a graphic. Spatial domain watermarking is illustrated in Figure 2. This demonstrates the degree of visibility of the watermark. Considerable latitude is available, in terms of placement, size, and intensity, to blend the watermark into a graphic.

Another spatial watermarking technique uses color separation. In this way, the watermark appears in only one of the color bands. This renders the watermark visually subtle so it is difficult to detect during regular viewing. However, the watermark appears immediately when the colors are separated for printing. This renders the document useless to the printer unless the watermark can be removed from the color band. This approach is used commercially for journalists to inspect digital pictures from a photo-stockhouse before buying

unwatermarked versions.

An alternative to spatial watermarking is frequency domain watermarking. In this case, transforms like the Fast Fourier Transform (FFT) alter the pixel values of the image for chosen frequencies. Since high frequencies will be lost by compression or scaling, the watermark signal is applied to lower frequencies, or applied adap-

technique. However, there is more of a trade-off here between invisibility and decodability, since the watermark is in effect applied indiscriminately across the spatial image.

Watermarking can be applied to text images as well. Three proposed methods are: text line coding, word-space coding, and character encoding. For text line coding, the text lines of a docu-

alteration is practically imperceptible. The effectiveness of such watermarking is confirmed in Figure 3b. Even with the affected lines set apart in red, it is still difficult to determine the lines are elevated.

For word-shift coding, the spacing between words in a line of justified text is altered. The plain text in Figure 4a has three words shifted right one pixel.

Figure 4b highlights the affected words.

The remaining text watermarking technique involves character coding. This involves minor alterations to the shapes of characters such as clipping a serif imperceptibly, or extending a descender. An advantage of these methods over those applied to picture images is that, by combining two or three of these to one document, two documents with different watermarks cannot be spatially registered to extract the watermark. Of course, the watermark can be defeated by retyping the text.
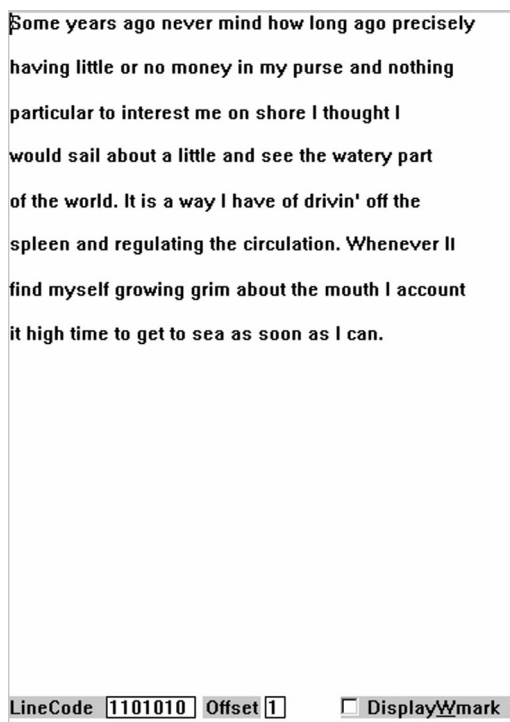

Figure 3a. Text with lines 1, 2, 4 and 6 elevated from normal position by one pixel
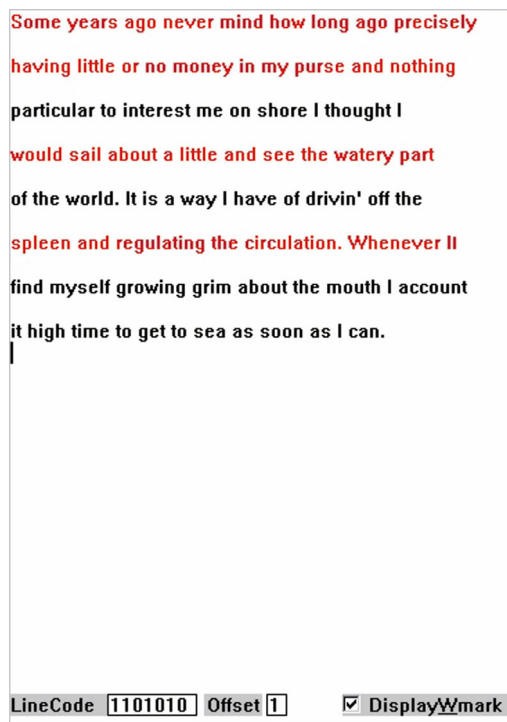

Figure 3b. Elevated lines highlighted

tively to frequencies containing important information of the original picture (feature-based schemes). Since watermarks applied to the frequency domain will be dispersed over the entirety of the spatial image upon inverse transformation, this method is not as susceptible to defeat by cropping as the spatial

ment page are shifted imperceptibly up or down. For a 40-line text page, for instance, this yields 2^{40} possible code words.

Figure 3a illustrates text line coding as it would appear to the casual reader. According to the line code box, the first, second, fourth and sixth lines are elevated by one pixel, although the

Limitations of Digital Watermarking

As of this writing, a counterfeiting scheme has been demon-

Some years ago never mind how long ago precisely having little or no money in my purse and nothing particular to interest me on shore I thought I would sail about a little and see the watery part of the world. It is a way I have of drivin' off the spleen and regulating the circulation. Whenever I find myself growing grim about the mouth I account it high time to get to sea as soon as I can.
I

Figure 4a. Text with three words offset by one pixel

Some years ago never mind how long ago precisely having little or no money in my purse and nothing particular *to* interest me on shore I thought I would sail about a little and see the watery part of *the* world. It is a way I have of drivin' off the spleen and regulating the circulation. Whenever I find *myself* growing grim about the mouth I account it high time to get to sea as soon as I can.
I

Figure 4b: Text with offset words highlighted

proceeds in two stages. First, the watermarked signature is "removed" from the watermarked copy. The watermark signature is the "difference" between the original (white) and the water-marked copy of the original (blue). Next, the extracted signature (blue) is compared against the original signature (gold). Identity signifies authenticity of the copy.

The counterfeiting scheme (see Figure 5b) works by first creating a counterfeit water-marked copy (violet) from the genuine watermarked copy (blue) by effectively inverting the genuine watermark. This inversion produces a counterfeit signature (violet) as well.

The trick is the original image and bona fide signature stand in the same relationship to the watermarked image as the counterfeit image and counterfeit signature (see Figure 5c). Thus, the technique of establishing legitimate ownership by recovering the signature watermark by comparing a watermarked image with the original image breaks down. While it may be demonstrated that at least one recipient has a counterfeit watermarked copy, it cannot be determined who it is.

This research suggests not all watermarking techniques will be useful in resolving ownership disputes in courts of law. There will likely be noncommercial applications, or those with limited vulnerability to theft, where "good enough watermarking" will suffice. More sensitive applications may require noninvertable or nonextracting watermarking techniques.

strated for a class of invertible, feature-based, frequency domain, invisible watermarking algorithms. This counterfeiting scheme could be used to subvert ownership claims because the recovery of the digital signature from a watermarked image requires a comparison with an original. We may illustrate the point simply with graphics.

Standard watermarking (see Figure 5a) involves the creation of a watermarked image by encoding a signature into an original image. Authentication

IN THE CASE OF IMAGERY, SEVERAL DIFFERENT METHODS

enable watermarking in the spatial domain from simply

flipping low-order bits of selected pixels to superimposing

watermark symbols over an area of a graphic.

The Future of Watermarking

The enormous popularity of the World Wide Web in the early 1990s demonstrated the commer-cial potential of offering multime-dia resources through the digital networks. Since commercial inter-ests seek to use the digital net-works to offer digital media for profit, they have a strong interest in protecting their ownership rights. Digital watermarking has been proposed as one way to protect such interests. Though much research remains before watermarking systems become robust and widely available, there is much promise they will contribute signif-icantly to the protec-tion of proprietary interests of electronic media. Collateral technology also will be necessary to auto-mate the process of authentication, nonre-pudiable transmission, and validation.

Pointers:

Some of this material was adapted from H. Berghel and L. O'Gorman, "Protect-ing Ownership Rights through Digi-tal Watermarking," *IEEE Com-puter 29*, 7 (1996), 101–103.

A good overview of how coun-terfeiters could attack water-marked images based on the correlation of the differences between samples is reported in H. Stone, "Analysis of Attacks on Image Watermarks with Ran-domized Coefficients," NEC Research Institute Technical Report, May 17, 1996. The counterfeiting scheme described here will appear in S. Craver, N. Memon, B. Yeo, and M. Yeung, "Resolving Rightful Ownerships with Invisible Watermarking Techniques: Limitations, Attacks and Implications," *IEEE J. on Selected Areas of Communications*, Dec. 1997.

The latter two articles also con-tain useful references to the water-marking literature. Information on IBM's digital library projects are to be found at www.ibm.com/ IBM/ibmgives/ diglib.htm.

The images in Figures 2 and 3 were taken from our digital watermarking demonstration pro-gram available as freeware for noncommercial use through the author's ftp site (see bio). **C**

HAL BERGHEL (www.acm.org/~hlb) is a professor of computer science at the Univer-sity of Arkansas and a frequent contributor to the literature on cyberspace.

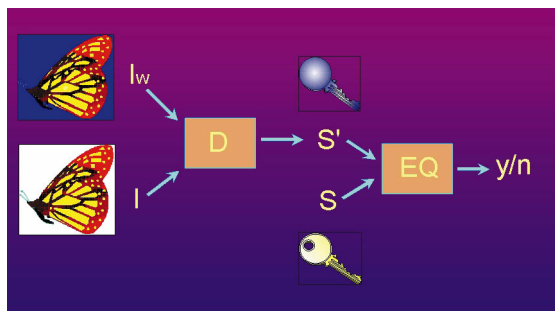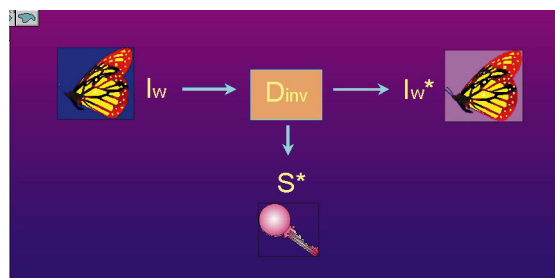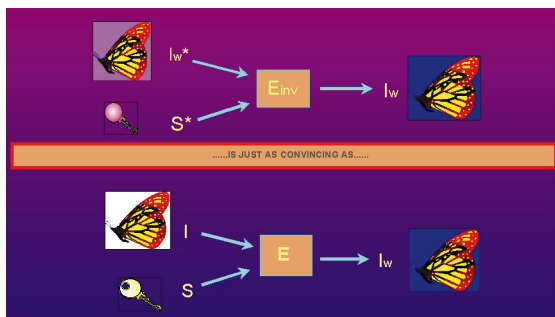**Figure 5a.** Basic watermarking technique

Figure 5b. Watermark "inversion" for counterfeiting

Figure 5c. Counterfeit logic