Perceptual Watermarks for Digital Images and Video

RAYMOND B. WOLFGANG, STUDENT MEMBER, IEEE,
CHRISTINE I. PODILCHUK, MEMBER, IEEE, AND EDWARD J. DELP, FELLOW, IEEE

Invited Paper

The growth of new imaging technologies has created a need for techniques that can be used for copyright protection of digital images and video. One approach for copyright protection is to introduce an invisible signal, known as a digital watermark, into an image or video sequence. In this paper, we describe digital watermarking techniques, known as perceptually based watermarks, that are designed to exploit aspects of the human visual system in order to provide a transparent (invisible), yet robust watermark. In the most general sense, any watermarking technique that attempts to incorporate an invisible mark into an image is perceptually based. However, in order to provide transparency and robustness to attack, two conflicting requirements from a signal processing perspective, more sophisticated use of perceptual information in the watermarking process is required. We will describe watermarking techniques ranging from simple schemes which incorporate common-sense rules in using perceptual information in the watermarking process, to more elaborate schemes which adapt to local image characteristics based on more formal perceptual models. This review is not meant to be exhaustive; its aim is to provide the reader with an understanding of how the techniques have been evolving as the requirements and applications become better defined.

Keywords*—*Digital watermarking, multimedia security, perceptual models.

I. INTRODUCTION

In the past several years there has been an explosive growth in digital imaging technology and applications. Digital images and video are now widely distributed on the Internet and via CD-ROM. One problem with a digital image is that an unlimited number of copies of an "original" can be easily distributed and/or forged. This presents problems if the image is copyrighted. The protection and enforcement of intellectual property rights has become an important issue in the "digital world." Recently, Congress passed the Digital Millennium Copyright Act (DMCA) which addresses copyright issues for digital content. This act makes it illegal to attempt to circumvent any technological measure, including digital watermarking, that effectively protects an owner's intellectual property rights of digital content. This move by Congress should act as a catalyst to further advance the current state-of-the-art technology as well as help define new applications for watermarking.

Many approaches are available for protecting digital images and video; traditional methods include encryption, authentication, and time stamping. In this paper we present algorithms for image authentication and forgery prevention known as *digital watermarking*. A digital watermark is a signal that is embedded in a digital image or video sequence that allows one to establish ownership, identify a buyer, or provide some additional information about the digital content.

This paper focuses on invisible watermarks that are designed to exploit perceptual information in the watermarking process known as *perceptual watermarks*. We will first describe the problem, purpose, and requirements of digital watermarking. An overview of visual models developed for image processing applications is then presented. The review of watermarking techniques begins with some examples of image-independent perceptual watermarks, that is, watermarking techniques which are based on the modulation transfer function (MTF) of the human eye only, not the particular characteristics of the individual images or video frames. Next, we cover the class of perceptual techniques known as *image-adaptive watermarks*, that is, watermarks which depend not only on the frequency response of the human eye, but properties of the image itself. In this way, one can maximize the watermark strength (robustness) while satisfying the transparency requirement. The image-adaptive watermarking schemes are the main emphasis of this paper. We conclude with some examples of video

Manuscript received April 4, 1998; revised December 13, 1998. This work was supported in part by a grant from the AT&T Foundation.

R. B. Wolfgang and E. J. Delp are with the Video and Image Processing Laboratory (VIPER), School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN 47907-1285 USA (e-mail: wolfgang@ecn.purdue.edu; ace@ecn.purdue.edu).

C. I. Podilchuk is with Bell Laboratories, Lucent Technologies, Murray Hill, NJ 07974 USA (e-mail: chrisp@bell-labs.com).

Publisher Item Identifier S 0018-9219(99)04947-6.

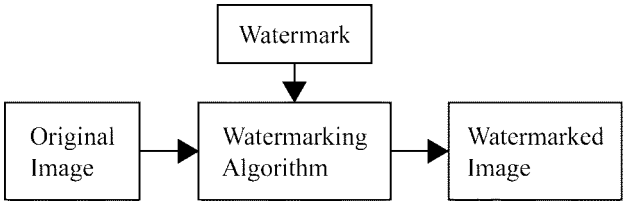


Fig. 1. Block diagram of a watermarking algorithm.

watermarking and their unique requirements. Although watermarking for other media, such as audio, is outside the scope of this paper, the general notion of using perceptual models to determine the watermarking strategy also applies.

This paper describes image watermarking from a technical perspective. It is important to note that any technique that allows a user to assert their ownership of digital content must also be placed in the context of intellectual property right law. In the final analysis, how "well" a watermarking technique works depends on how effectively the technique protects an owner's intellectual property rights in a court of law [1]. Watermarking is a special case of data hiding or steganography. Steganography is the practice of encoding secret information in a communications channel in a manner such that the very existence of the information is concealed [2], [3]. Typically, in steganography, the secret information contains all the value, and the communications channel or vessel used to hide it is not of value in itself. In digital watermarking, however, the secret message (watermark) is of little or no value on its own, and the vessel or communications channel (digital image or video) is of value. The effective coupling of the secret message to the image makes the watermarking technique effective [4], [5].

Overviews of digital watermarking are presented in [6]–[8] as well as other papers in this PROCEEDINGS.

A. What is the Watermarking Problem?

Digital watermarking aids owners in asserting their intellectual property rights on the works of art or digital content they create. These rights are particularly difficult to enforce with digital images, since it is easy to copy and distribute perfect copies of an original image. Fig. 1 shows that the watermarking process consists of a marking algorithm that inserts the watermark into an image. The watermark may be inserted in the spatial domain or spatial frequency domain. As part of the watermarking process, a testing or verification algorithm must be defined that tests an image to see if a particular watermark is contained in the image. It may be also desirable for the testing procedure to determine if the image has been altered and to supply localization information as to where the image was altered. It is our feeling that to assert ownership that is consistent with current intellectual property right law, the watermarking technique must support the use of third-party cryptographic-based digital time stamping [9] that is embedded in the image through the watermarking process. In [10], two general scenarios were identified where copyright enforcement is needed.

1) *Invisible Watermarking for Content and/or Author Authentication*: An example of this scenario is the digital recording of images from newsworthy or historical events. The images should be watermarked upon capture so that the news services can be sure that an image has not been altered. The unrestricted distribution of copies of the images is much less a concern here than verifying an image's origin and content. This is a very critical issue in the protection of historical images and images used in courts of law as evidence [11].
2) *Invisible Watermarking for Detecting Unauthorized Copies of Images*: An example of this scenario is proving ownership of an image. The mere presence of the owner's mark in a suspect image can prove the theft of that image by the unauthorized possessor. This type of watermark can be used to label the customer who is a potential illegal distributor. The mark would identify the original purchaser whose copy has been illegally distributed.

Requirements and properties differ for effective digital watermarks in the above scenarios. For authentication, it is important that even slight changes to the image be detected and localized. Embedding a false mark must be practically impossible and the watermark should be sensitive to "information altering or destroying" image manipulations. Ideally, for these types of applications, the watermarks should be insensitive to some common, innocent transformations, such as compression, but should be sensitive to image transformations that alter the information, such as replacing a portion of the image. The challenge here from a signal processing perspective is to provide a watermark that can distinguish between information altering and simple signal altering transformations. These types of watermarks are known as *semi-fragile watermarks*. *Fragile watermarks*, on the other hand, are not robust to any lossy transformations of the image.

When an image is marked using a fragile watermark, an attacker does not want changes that they make to the image to alter the mark; it is desired by the attacker to have an altered image "pass" as authentic. This is not the case in the second scenario. Here, an attacker wants to remove the watermark at a minimal loss in image quality. In this way, the true owner cannot verify the presence of the watermark in the image, thus greatly reducing any ownership claim. The watermark, therefore, must remain in the image after many types of attacks. These attacks include compression, filtering, printing and rescanning, and geometric attacks such as cropping, resampling, and rotation. These attacks can include sophisticated methods for removing the watermark [12]. Furthermore, users must not be able to attack the watermark through collusion by comparing multiple copies of the image, each marked with a different watermark. Watermarks that are resistant to these attacks are known as *robust watermarks*.

Techniques that require the original image to verify the mark are known as *private watermarks*, while approaches

that do not require the original image for verification are known as *public watermarks*. Private watermarks are typically used for applications requiring a robust watermark, such as identifying the buyers to prevent illegal duplication and distribution. Public watermarks are necessary for some of the fragile watermarking applications, such as authentication and tamper detection, when the original is not available for watermark detection. It is also possible to describe watermarking techniques that embed visible information in an image. These visible watermarks are typically verified by inspection [10]. Visible watermarking techniques are outside the scope of this paper.

In this paper, we will describe transparent watermarking schemes for image and video data that are designed to use perceptual information based on human visual system models. There are three principles that characterize perceptually based watermarks for robust applications [5].

1) *Transparency:* The watermark is not visible in the image under typical or specified viewing conditions.
2) *Robustness to Attacks:* The watermark is detected with low probability of error after the image has undergone linear or nonlinear operations such as those mentioned above.
3) *Capacity:* The ability of the watermarking scheme to be able to verify and distinguish between different watermarks with a low probability of error as the number of differently watermarked versions of an image increases.

Additional tools are available to protect images from unauthorized viewing and modification. These include encryption techniques and digital time stamping [9], [13]. Encryption disguises the content of an image and makes it "unreadable" to most applications. Only users who possess the decryption "key" can convert the encrypted image data back to its original form. Asymmetric encryption (i.e., public key cryptography) can be used for the authentication of a digital image. Authentication does not protect the content of the image but proves who created the image. If an image is altered, even by one bit, authentication techniques will indicate that the image has been changed. We feel that the use of "pure" encryption has limited use in digital images and video because this technique results in unviewable content. Digital watermarking compliments encryption technology by providing a way to protect digital content in its original viewable form. Time stamping determines the image owner and the time at which an image is created. One trivial way to prove ownership of an image is to process the earliest time stamp of the image. We believe that time stamping is absolutely critical to the success of any multimedia security system. We feel that all image watermarking techniques must use some form of third-party cryptographic time stamping to be successful. In many cases the time stamp can be used as part of the watermark generation. Time stamps also thwart the "rewatermarking attack" described in [14].

B. Why Use Visual Models?

The simplest form of a perceptually based watermarking scheme relies on incorporating the watermark into the perceptually insignificant parts of an image in order to guarantee transparency. An example of such a scheme is to make sure that the watermark signal amplitude is quite low by inserting information only into the low-order bits in an image. However, watermarks which are embedded in this way are easily removed or altered through mild filtering or random bit flipping of the lower bits without "perceptually" affecting the image quality. If the amplitude of such watermarking schemes is increased to make the scheme more robust, the mark may become visible. For this reason, techniques have been examined that use more sophisticated ways of incorporating perceptual knowledge into watermarking schemes in order to provide robustness as well as transparency.

Several effective frequency domain techniques have been introduced which provide a framework to incorporate some type of frequency weighing in order to take advantage of the frequency sensitivity of the eye. These techniques range from simple frequency selection based on common-sense rules to frequency weighing based on models of the human visual system. These techniques may be image independent or dependent but usually do not take advantage of local image characteristics. Typically, they are based only on viewing conditions. Image-adaptive techniques not only take advantage of general frequency sensitivity but also rely on adapting to local image properties in order to provide maximum performance in terms of robustness while maintaining the transparency constraint.

We begin by introducing visual models that have been developed in the context of still image compression. Such models are ideally suited for the watermarking problem since thresholds developed to determine the visibility of compression artifacts can be used to determine the perceptual upper bound on the watermark signal strength. This section is followed by a general review of watermarking techniques for still images. The review starts with the simplest techniques, motivated by the transparency requirement, and continues with more elaborate frequency-domain techniques motivated by robustness as well as transparency. We then describe image-adaptive techniques, which, depending on the accuracy of the visual model, should provide the optimum performance in terms of robustness and transparency. We then describe how some of the perceptually based still image watermarking techniques have been extended to video watermarking.

II. Visual Models—Overview

There has been much work over the years on trying to understand the human visual system as well as using this knowledge for image and video applications [15]. In particular, a very useful application for perceptual models is in the area of source coding or compression [16]. While traditional coding techniques take advantage of signal statistics to remove redundancy, it is ultimately the viewer

who decides how well the compressed version represents the original image or how objectionable the compression artifacts are. Perceptual models allow one to take advantage of characteristics of the human visual system in order to remove irrelevancy as well as redundancy in designing optimal compression algorithms.

Here we describe three properties of the human visual system that have been studied in the context of image coding: frequency sensitivity; luminance sensitivity; and contrast masking. Most of the early work on perceptually based image coding has utilized the frequency sensitivity of the human visual system as described by the modulation transfer function (MTF). The MTF describes the human eye's sensitivity to sine wave gratings at various frequencies [17]. From such a model, given that the minimum viewing distance is fixed, it is possible to determine a static just noticeable difference (JND) threshold for each frequency band. The JND threshold is such that changes in the frequency content in the image in the particular frequency band below the threshold are not noticeable [18]. These thresholds can be used for both quantization and bit allocation in image compression [19]. Frequency sensitivity provides a basic visual model that depends only on viewing conditions and is independent of image content.

A further refinement can be achieved by extending the visual model to include luminance sensitivity. Luminance sensitivity is a way to measure the effect of the detectability threshold of noise on a constant background [17]. This phenomenon depends on the average luminance value of the background as well as the luminance level of the noise. For the human visual system, this is a nonlinear function. Since luminance sensitivity takes advantage of local luminance levels, it is important that the frequency framework chosen for the compression technique allows for some local spatial control.

Frequency sensitivity and luminance sensitivity are good starting points in utilizing properties of the human visual system for image compression. However, frequency sensitivity depends only on viewing conditions; also, luminance sensitivity is a conservative estimate of visual masking, which does not model masking properties due to high-frequency details or texture. Ideally we would like a more dynamic model that allows for finer control of the masking process. The addition of contrast masking allows for even more dynamic control of the JND threshold levels. Contrast masking refers to the detectability of one signal in the presence of another signal; the effect is strongest when both signals are of the same spatial frequency, orientation, and location [20], [21]. The most effective visual models should take into account frequency sensitivity, local luminance sensitivity, and contrast masking.

The choice of the frequency decomposition of the image used in an encoder can affect not only the performance of the compression system, but also how effectively visual masking can be utilized. For compression, the frequency transformation should decorrelate the data and compactly represent the original signal energy in a minimum number of nonzero coefficients. In order to utilize visual properties effectively, the frequency decomposition should be selected to allow control of the spatial frequency location of the quantization distortion. Ideally, the addition of quantization distortion to one frequency coefficient should not show up in coefficients that are not adjacent to the one that was perturbed. The frequency decomposition should also mimic the human visual system's structure in order to gain the most in terms of masking. For a human, this structure is a set of filters with frequency spacing of approximately 1.5 octaves and an angular width of 40° [20].

The discrete cosine transform (DCT) and other uniform filter banks [22] satisfy the criterion of controlling the frequency location of the quantization distortion but do not provide a good model of the human visual system's structure. This presents a difficulty in creating masking models since there is a mismatch between the underlying structure of the model and the structure of the transform. However, it is worth studying how visual models can be utilized in a DCT framework, since it is the current building block for still image and video compression standards. The cortex transform [20] was introduced as a way to produce a transform that corresponds to the known structure of the human eye; a corresponding coder based on this transform is described in [19].

JPEG is the current international standard for color still image compression [23]. The most basic version of this coder, referred to as the baseline sequential codec, consists of decomposing the original image into nonoverlapping blocks, obtaining a DCT of each block, quantizing the transform coefficients, and entropy coding the quantized coefficients. It is very desirable to see how we can take advantage of visual models within this framework, although it has already been mentioned that block-based DCT's are not ideal in terms of mimicking the human visual system's structure.

The quantizer step size that is used in JPEG for each DCT coefficient is given by a quantization table which is specified as an input to the encoder. Since JPEG allows the user to specify a quantization table for each image, it should be possible to derive a "perceptually optimal" table. In fact, in [24] a set of formulas for determining the perceptually optimal quantization tables for both luminance and chroma components, given the image size, monitor white point, and viewing distance, is presented. These formulas were derived by running a large set of subjective experiments to determine the detectability of each DCT basis function.

From these experiments, a perceptual model for predicting the detection thresholds based only on the viewing conditions and global properties of the visual system was derived. This model takes into account frequency sensitivity in determining the optimum quantization table but does not take into account the image-dependent components of luminance sensitivity and contrast masking. This has been addressed in [25], where this approach has been extended to determine an image-dependent quantization table that incorporates not only the global conditions, but also accounts for local luminance and contrast masking. An iterative approach is presented to determine an image-dependent quantization table that provides a specified level of visual

distortion. Since JPEG allows for only one quantization table for all the image blocks, it is difficult to take full advantage of the local properties as given by the model in [25].

The work in [26] introduces additional local quantizer control by using the previous model to drive a prequantizer which zeros out all DCT coefficients below the locally derived JND threshold. Such an approach is compliant with the JPEG bit-stream specification while allowing some amount of local control as given by the visual model. The perceptual coder with the optimized quantization table and adaptive prequantizer is shown to yield significant improvement over the standard implementation of JPEG. Other perceptually based algorithms that have been proposed for image compression include [27], which is based on a subband decomposition and quantization step sizes derived from frequency and luminance sensitivity and contrast masking. An overview of using visual models for signal compression is presented in [16].

A different image-compression model has been developed in [28]. This model uses both frequency sensitivity and spatial masking based on edge content. The spatial masking is a modified version of the spatial masking model presented in [29]. The main principle of the spatial masking is that edges in an image are able to mask signals of much greater amplitude than regions of near-constant intensity. For a given image, a tolerable-error level (TEL) may be formed for each pixel. This quantity is similar in concept to the JND. The TEL provides the magnitude that a pixel can change without the change becoming visible. The development of the TEL is further described in [28]. A third model, which is based on contrast masking is described in [30] and [31].

The JPEG bit-stream specification limits the amount of perceptual fine tuning that can be incorporated into the encoder [23]. This is also true for other coding schemes, where the overhead information needed for the visual model is prohibitively large and hence may reduce the effects gained by using the model. However, for the watermarking application we are not limited by the amount of bits needed to transmit the perceptual information. Some of the models introduced in this section for image compression will be utilized later in developing very effective watermarking schemes.

III. Perceptual Watermarking for Still Images

A. Motivation

There are two basic modalities for image watermark embedding: spatial-domain techniques (spatial watermarks) and spatial frequency-domain techniques (spectral watermarks). This section first describes several spatial watermarking algorithms that rely on some form of perceptual knowledge. Many of the spatial watermarking techniques provide simple and effective schemes for embedding an invisible watermark into an image but are not robust against common attacks. Another way to mark an image is to transform it into the frequency domain using a DCT, wavelet, or other type of transform. The mark is incorporated directly into the transform coefficients of the image. These types of algorithms commonly use frequency sensitivity of the human visual system to ensure that the watermark is invisible. Many of these techniques, however, are not image adaptive. The algorithms described in Sections IV and V fall into the category of image-adaptive watermarks. They use formal visual models to determine how to embed the watermark by taking advantage of image characteristics to provide perceptual transparency as well as robustness to attack.

B. Perceptual Watermarking in the Spatial Domain

The watermarking scheme described in [32] is based on spread-spectrum communications. A linear feedback shift register with n stages can be used to form pseudorandom binary sequences with periods as large as $2^n - 1$. M sequences achieve this maximum period and have very desirable autocorrelation and randomness properties [9]. Two types of sequences may be formed from an m sequence: *unipolar* sequences $\{0, 1\}$ and *bipolar* sequences $\{-1, 1\}$. An extended m sequence has properties similar to an m sequence [33] but has length 2^n. One creates an extended m sequence by appending a zero to the end of the longest run of zeros in an m sequence. Let X be a 512×512 grayscale image, and w a bipolar extended m sequence of length 512. The watermark W consists of 512 circularly shifted copies of w with random phase, one in each row of X. W is then added to X to form the watermarked image Y

$$Y = X + W. \tag{1}$$

To verify whether an image Z has the watermark in it, a row-by-row spatial cross correlation between Z and W is obtained. Let z be a row of Z. The spatial crosscorrelation function between z and w is

$$R_{zw}(\alpha) = \sum_j [z(j) - E[z]]w(j - \alpha) \tag{2}$$

where $E[z]$ is the mean of row z and α is the offset. R_{zw} is examined for the presence of peaks indicating high correlation. If the peaks for each row are below a threshold, Z does not contain the watermark. Note that this technique is a public technique in that the original image is not needed to verify the presence of the watermark in Z.

This watermarking technique is robust to small amounts of distortion introduced in Y (i.e., the peak of R_{zw} is still statistically significant). It can also accommodate multiple watermarks. However, the mark is not robust to random bit flipping in the lower two bit planes, or more sophisticated image processing attacks.

The spatial watermarking scheme proposed in [34] and [35] is known as the *Variable-W Two-Dimensional Watermark (VW2D)*. The m sequence is reshaped into two-dimensional watermark blocks, which are added and verified on a block-by-block basis. VW2D can detect local alterations in an image on a block-wise basis [36].

Many other spatial domain watermarks exist. One algorithm partitions the pixels of the original image into

two sets, A and B, using a random partition S [37]. The luminance of pixels in set A is increased by an integer k, where k is small enough to maintain the imperceptibility of the alterations. Verification is performed with S and k; the original image is not required. A similar method (known as the patchwork method) chooses a set of pairs of pixels and increases the difference between the value of each pixel in the pair [38]. A second method presented in [38], known as texture block coding, inserts a textured patch into an area of the image with the same texture. This is a good example of using common-sense rules to determine where signal alterations will be least noticeable, which results in good visual quality of the marked image. This scheme, however, requires that the image contain relatively large areas of texture; the technique is also vulnerable to low-pass filtering. In other words, transparency is met at the expense of robustness.

Some transparent spatial watermarks exploit the fact that humans detect changes in different colors unequally. One such algorithm embeds sinusoidal patterns exclusively in the yellow–blue color component of an image [39]. Another performs amplitude modulation only in the blue component [40]. Several invisible spatial watermarks have been developed which exploit color dependency of an image [41], [42].

Another method selectively filters a subset of pixels in an image (known as signature pixels) to embed a watermark [43]. The locations of signature pixels are determined by the watermark. To embed the watermark, a pixel is replaced with a filtered version of itself. Verification is performed by first filtering the image, then identifying the prefiltered pixels.

A watermarking algorithm has been developed specifically for halftone images [44]. This algorithm exploits the fact that different halftone patterns will produce a perceptually similar gray field in an image. The halftone pattern of the image is modified to incorporate a watermark. The watermark is verified by inspection as follows: a transparent sheet (known as the screen) with a particular halftone pattern is overlaid on a printed version of the watermarked image. Upon sliding the sheet into alignment with the printed image, a visible watermark appears. This process is equivalent to optical correlation of the sheet with the image. The mark is invisible in the printed image. Another perceptual spatial algorithm [45] segments the image into blocks, much like VW2D. The amplitude of the mark is adapted to the corresponding image block using local statistics.

C. Perceptual Watermarking in the Frequency Domain

1) A DCT Domain Watermark: An early frequency domain method for digital watermarking of images proposed in [46] is also based on the concept of spread-spectrum communications. The technique is motivated by both perceptual transparency and robustness. The results show that the types of image distortions to which this technique is robust include cropping, very low data-rate JPEG compression, printing, and rescanning, as well as collusion

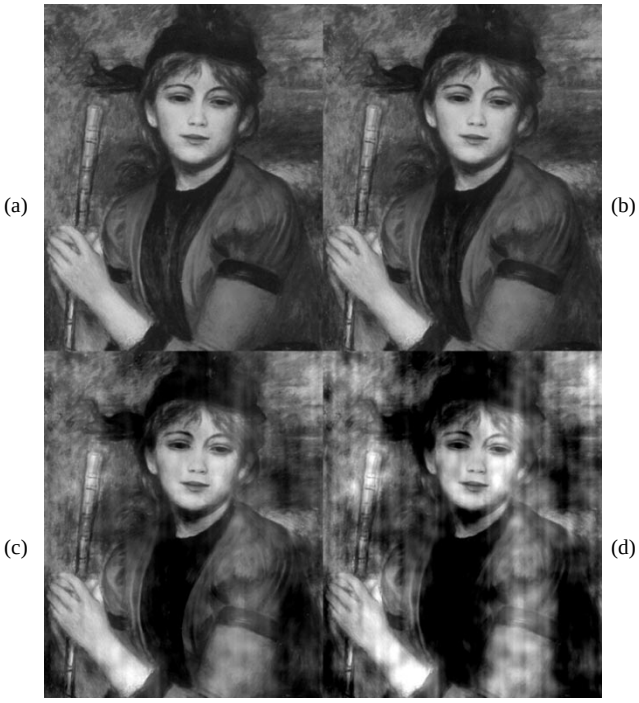


Fig. 2. Example of DCT-spread-spectrum technique in [46]: (a) original; (b) $a = 0.1$; (c) $a = 0.5$; and (d) $a = 1.0$.

with several independently watermarked images. One of the significant contributions in this work is the realization that the watermark should be inserted in the perceptually significant portion of the image in order to be robust.

The watermark W is a sequence of normally distributed, zero-mean unit-variance random numbers. A DCT is performed on the entire image and W is inserted in a predetermined range of low frequency coefficients. Let X be the original image, Y be the watermarked image, and X_D and Y_D be the DCT coefficients of X and Y, respectively. Let $X_D(u, v)$ and $Y_D(u, v)$ be the (u, v)th DCT coefficient in X_D and Y_D, respectively. $W(u, v)$ is the (u, v)th element in the watermark sequence; a is a scale factor which prevents unreasonable values for $Y_D(u, v)$. The marking algorithm is expressed as

$$Y_D(u, v) = X_D(u, v)(1 + aW(u, v)). \tag{3}$$

Alternate marking methods are described in [46]. Taking the inverse transform of Y_D to form Y completes the marking procedure. Fig. 2(a) shows an original image as well as watermarked versions [Fig. 2(b)–(d)] of the original with $a = 0.1, 0.5$, and 1.0, respectively. The authors propose an empirically determined value of 0.1 for a and choose to insert the watermark in the 1000 lowest frequency non-DC DCT coefficients. The first step of the verification procedure is to extract a "copy" of the watermark from a possibly altered image Z. Z_D is the DCT of Z. W^* is the extracted version of the watermark

$$W^*(u, v) = \frac{1}{a}\left[\frac{Z_D(u, v)}{X_D(u, v)} - 1\right]. \tag{4}$$

A similarity measure between W^* and W is determined as follows:

$$\gamma(W, W^*) = \frac{W \cdot W^*}{\sqrt{W^* \cdot W^*}}. \quad (5)$$

If an image has not been watermarked with W, then γ can be modeled as a zero mean random variable. If W^* differs only slightly from W (i.e., W is indeed present in Z, although slightly altered), then $E[\gamma] \gg 0$. A hypothesis test on γ determines if W is present in the image. This technique accommodates multiple watermarks, and withstands a much wider range of attacks than the transparency-only based spatial techniques. This technique requires the original image for verification.

This algorithm is one of the earliest attempts at providing image adaptability in the watermark embedding scheme. This is due to the fact that the watermark strength depends on the intensity of the DCT coefficients of the original image. In this way, the watermark signal can be quite strong in the DCT coefficients with large intensity values and is attenuated in the areas with small DCT coefficients. Inserting the watermark into the perceptually significant components and adapting the watermark strength by the strength of the DCT component provides a watermark that is quite robust and transparent. However, because the DCT is obtained on the entire image rather than the usual block-based approach commonly found in image and video compression schemes, the transform does not allow for local spatial control of the watermark insertion process. In other words, the addition of a watermark value to one DCT coefficient affects the entire image. This scheme may benefit from a perceptual model that determines the optimal weights for the DCT coefficients, but the framework needs to be modified in order to obtain finer control of watermark adaptability to image characteristics and the human visual system. A wavelet-based version of the technique is described in [47].

Another global method also modulates the DCT coefficients but uses a one-dimensional bipolar binary sequence for W [48]. The marking procedure consists of sorting the DCT coefficients of the image according to their absolute magnitude. A percentage of total energy, P, is defined and the largest n coefficients are identified relative to P. The watermark is then added to the n largest AC coefficients. A larger P increases the number of coefficients that are marked but increases the probability that the watermark will be perceptible. This technique also requires the original image for verification.

The method in [49] is a slight modification of previous work [46], [48], where the authors allow the user to determine the scaling factor and coefficients to be marked. The user-defined scaling factor and watermark length will greatly influence the effectiveness of this scheme both in terms of transparency and robustness.

2) Other Transform-Based Approaches: An early DCT-based technique is presented in [50] and [51]. The image is segmented into 8×8 nonoverlapping blocks and the DCT of each block is obtained similar to JPEG. A random subset of the blocks is chosen and a triplet of midrange frequency coefficients is slightly altered to encode a binary sequence. This seems to be a reasonable approach for adding some sort of perceptual criterion. Watermarks inserted into the high frequencies are most vulnerable to attack, whereas the low-frequency components are perceptually significant and very sensitive to alterations; such alterations may make the watermark visible. This scheme provides reasonable results on average, although a more image-dependent scheme could provide better quality and robustness. The original image is not required for verification.

The basic ideas introduced in [50] are further extended in [52] by introducing the watermark embedding in the quantization process of the midfrequency coefficients. The result is two schemes that also do not require the original image for watermark verification. The first scheme embeds a linear constraint among the selected DCT coefficients; the second defines a circular detection region in the DCT domain similar in concept to vector quantization [53]. An improvement to [50] is presented in [54] and [55]. The original image is segmented into 8×8 blocks. Image blocks that contain either sharp edges, or have little texture are not marked.

The method in [56] embeds the watermark in the phase information of the discrete Fourier transform (DFT) of an image; here the authors argue that the phase information should be watermarked since this is the perceptually significant part of the original image data. A second DFT-based technique uses properties of the DFT to create a watermark resistant to geometric attacks (rescaling, translation, and rotation) [57].

All of the techniques described above do not explicitly incorporate visual models into the marking algorithm even though many of the techniques are image dependent. In Section IV we describe how visual models can enhance the performance of the watermarking technique relative to the three principles described in Section I.

IV. Image-Adaptive Watermarking

A. Perceptual Watermarking Based on Image-Adaptability

We begin by reviewing some of the requirements that are necessary to provide a useful and effective robust watermarking scheme. We will briefly discuss the requirements as an introduction to the use of formal perceptual models in meeting these requirements.

In Section I, we discussed the three principles for watermarking: transparency; robustness; and capacity. Transparency refers to the perceptual quality of the data being protected. The digital watermark should also be robust to alterations of the image. Ideally, the amount of signal distortion necessary to remove the watermark should degrade the image quality to the point of becoming commercially valueless. Capacity may also be a critical feature for applications where the watermark identifies the buyer or end user. Capacity refers to the ability to detect individual watermarks with a low probability of error as the number of differently marked versions of the image increases. It does not refer to placing multiple watermarks in an individual image. The

watermarking technique should provide a way to insert the maximum number of distinguishable watermarks. To best meet these three requirements, we would like the watermark to adapt to the local image characteristics as well as viewing conditions in order to provide the strongest watermark (most robust) while satisfying the transparency constraint.

The above principles of transparency, robustness, and capacity introduce a challenging problem from the signal processing perspective. The most straightforward way to introduce a transparent watermark results in a watermark that is very vulnerable to attack. For example, placing a watermark in the least significant bits or in the high-frequency components results in very good image quality but can be destroyed with simple quantization or low-pass filtering. In other words, from a signal processing viewpoint, the requirements of transparency and robustness conflict with each other. A similar conflict exists from a perceptually based viewpoint. In order to provide transparency, it makes sense to consider perceptually insignificant components for watermark insertion. However, to provide robustness it makes sense to insert a watermark in the perceptually significant components.

B. Algorithm Description

The two techniques described here, the image-adaptive DCT (IA-DCT) approach as well as the image-adaptive wavelet (IA-W) approach [4], [5], [58] have been motivated by the excellent results presented in [46]. The goal is to investigate if a more image-adaptive scheme has any advantages over the results that have already been achieved with the scheme introduced in [46], which is based on a simple heuristic weighing of the frequency components and no local adaptability.

The frequency decomposition for the image-adaptive DCT algorithm is based on an 8×8 DCT. Unlike the decomposition in [46], the block-based approach provides local control that allows for incorporating local visual masking effects. A benefit of such a scheme (and other block-based DCT schemes) is that if the images are stored as compressed JPEG bit streams, the watermarks can be inserted directly to the partially decompressed bit stream [55], [59].

It is important to note that for the watermarking problem, all the local information extracted from the visual models can be utilized in determining the watermark embedding algorithm. The local information is stored in the JND. In the applications addressed here, the original image is available for verification and the JND thresholds can be obtained directly from this image. Actually, the JND thresholds can be estimated from the watermarked image fairly accurately; this means that this technique could be used in applications where the original image is not available for watermark verification.

The image-adaptive DCT-based approach uses the visual model described in Section II for image compression in a JPEG framework [25]. Recall that the JND thresholds derived from the visual model consist of an image-independent part based on frequency sensitivity and an image dependent part based on luminance sensitivity and contrast masking. They have been derived in the context of image compression to determine the maximum amount of quantization noise that can be tolerated at every frequency location without affecting the visual quality of the image (under the specific viewing conditions used in the model). In the context of image watermarking, the JND thresholds can be used to determine the maximum amount of watermark signal that can be tolerated at every location without affecting the visual quality of the image. The watermark W consists of a sequence of Gaussian random numbers with zero mean and unit variance, as presented in [46].

We present two watermarking schemes: the IA-DCT and IA-W scheme. The embedding process for both the IA-DCT and IA-W schemes can be described in general as

$$Y(u, v) = \begin{cases} X(u, v) + J(u, v)W(u, v), & |X(u, v)| > J(u, v) \\ X(u, v), & \text{otherwise} \end{cases} \tag{6}$$

where $X(u, v)$ are the frequency coefficients of the original image X, $Y(u, v)$ are the watermarked image coefficients, $W(u, v)$ is the watermark sequence, and $J(u, v)$ is the JND calculated for each frequency coefficient. A block diagram of the general image-adaptive perceptual watermarking scheme is illustrated in Fig. 3.

The watermark encoder for the IA-DCT scheme is described in (7) at the bottom of the page, where $X_D(u, v, b)$ is the unmarked DCT coefficient at position (u, v) of block b, $Y_D(u, v, b)$ is the watermarked DCT coefficient, $W(u, v, b)$ is the watermark sequence, and $J(u, v, b)$ is the JND threshold. For some applications we have *a priori* knowledge about some of the image transformations that will be applied to the watermarked image and it would be beneficial to take advantage of this knowledge in the watermarking process. For instance, if we know that the image will be low-pass filtered, it is best to avoid placing the watermark sequence in the high-frequency components. In general, however, watermark insertion is not limited only to perceptually significant parts of the image. The goal in this approach is to take advantage of the full capacity of the image in order to place a maximum strength watermark sequence that will be very difficult to alter or remove.

Note that since the watermark is generated from a normal distribution, watermark insertion will occasionally result in values that exceed the JND. Informal studies show that exceeding the JND occasionally does not result in any

$$Y_D(u, v, b) = \begin{cases} X_D(u, v, b) + J(u, v, b)W(u, v, b), & |X_D(u, v, b)| > J(u, v, b) \\ X_D(u, v, b), & \text{otherwise} \end{cases} \tag{7}$$

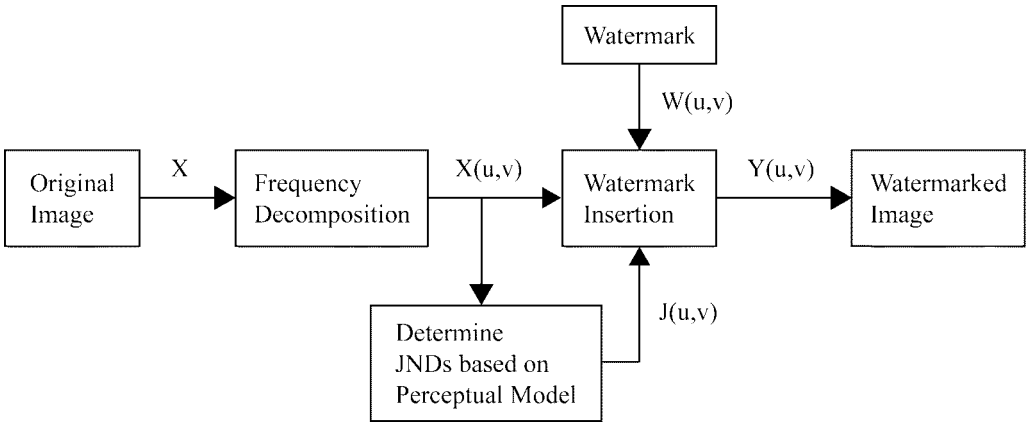


Fig. 3. General block diagram of the IA watermarking schemes.

visibly objectionable results. This might signify that there are other masking effects that could be incorporated into the visual models. Currently, the watermark is only inserted into the luminance component of the image.

For the IA-W scheme [4], [5], [58], frequency sensitivity thresholds are determined for a hierarchical decomposition using 9-7 biorthogonal filters [60]. Due to the hierarchical decomposition, this approach has the advantage of consisting of watermark components that have varying spatial support. This provides the benefits of both a spatially local watermark and a spatially global watermark. The watermark component with local spatial support is suited for local visual masking effects and is robust to attacks such as cropping. The watermark component with global spatial support is robust to operations such as low-pass filtering. Due to the hierarchical nature of such an approach, this scheme is more robust to certain types of distortions than the DCT-based framework [4], [5], [58].

The wavelet framework consists of a four-level decomposition. The visual model used here is much simpler than the one used in the DCT-based scheme. A wavelet-based JND $J_{l,f}$ is determined for each frequency band based on typical viewing conditions. Here l denotes the resolution level where $l = 1, 2, 3, 4$ and f denotes the frequency orientation where $f = 1, 2, 3$. The details of the experiments and resulting weights can be found in [60].

Adding image dependent components as in the DCT-based approach could further refine this model. However, even this simple visual model yields very good results and the hierarchical framework provides a robust watermark as well as finer control of watermark insertion than can be obtained using a block-based scheme. Results comparing the wavelet-based scheme to the DCT-based scheme are described in [5] and [58].

The watermark encoder for IA-W is described by (8) at the bottom of the page, where $X_{l,f}(u, v)$ is the unmarked wavelet coefficient at position (u, v) in resolution level l and frequency orientation f, $Y_{l,f}(u, v)$ is the watermarked wavelet coefficient, $W_{l,f}(u, v)$ is the watermark sequence, and $J_{l,f}$ corresponds to the JND at level l and frequency orientation f for the 9-7 biorthogonal filters. As for IA-DCT, the watermark is inserted only in the luminance component of the image.

Watermark verification for the approach in [46] as well as the IA-DCT and IA-W schemes are based on classical detection theory. The test image is subtracted from the original image and the correlation between the normalized signal difference and a specific watermark sequence is determined. The correlation is compared to a threshold to determine whether the test image contains the watermark in question. The correlation detection scheme for the IA-DCT scheme can be expressed as

$$W_s^*(u, v, b) = X_D(u, v, b) - Z_D(u, v, b) \tag{9}$$

$$W^*(u, v, b) = \frac{W_s^*(u, v, b)}{J(u, v, b)} \tag{10}$$

$$\rho = \frac{W \cdot W^*}{\sqrt{E_W \cdot E_{W^*}}} \tag{11}$$

where $W \cdot W^*$ denotes the dot product, $W_s^*(u, v, b)$ denotes the possible watermark in the test image scaled by the JND thresholds $J(u, v, b)$, $W^*(u, v, b)$ denotes the watermark in the test image Z, ρ is the normalized correlation coefficient between W and W^*, and $E_W = W \cdot W$. If W is similar to W^*, the correlation coefficient approaches one. If W and W^* are independent, ρ is normally distributed with zero mean. Therefore, the probability of ρ exceeding a certain threshold can be directly obtained from the normal distribution. Comparing the correlation coefficient to a threshold T completes the verification process. This threshold can be modified according to the tradeoff between the desired probability of detection, P_D, and the probability of false identification (false alarm), P_F. Hence

$$\begin{aligned} &\rho > T \quad && \text{watermark } W \text{ detected} \\ &\rho \le T \quad && \text{watermark } W \text{ not detected.} \end{aligned} \tag{12}$$

$$Y_{l,f}(u, v) = \begin{cases} X_{l,f}(u, v) + J_{l,f} W_{l,f}(u, v), & |X_{l,f}(u, v)| > J_{l,f} \\ X_{l,f}(u, v), & \text{otherwise} \end{cases} \tag{8}$$

Any prior knowledge about attacks should be incorporated either at the embedding or verification steps. For instance, if it is known that the image is to be low-pass filtered in some way, the high-frequency components should be avoided for watermark embedding. At verification, the potential watermark sequence should be "whitened" before the correlation step in order to achieve better detection results. The work presented in [46] offers several techniques to estimate the degradations given the test image and the original image.

The verification for the wavelet scheme is also based on correlation. What is different in IA-W is that the correlation is obtained separately for each subband (l, f) as follows:

$$W^*_{l,f,s}(u, v) = X_{l,f}(u, v) - Z_{l,f}(u, v) \tag{13}$$

$$W^*_{l,f}(u, v) = \frac{W^*_{l,f,s}(u, v)}{J_{l,f}} \tag{14}$$

$$\rho(l, f) = \frac{W_{l,f} \cdot W^*_{l,f}}{\sqrt{E_{W_{l,f}} \cdot E_{W^*_{l,f}}}}, \qquad l = 1, 2, 3, 4 \text{ and } f = 1, 2, 3. \tag{15}$$

The average for each resolution level l is obtained as

$$\rho(l) = \frac{1}{N_f} \sum_{f=1}^{N_f} \rho(l, f), \qquad l = 1, 2, 3, 4 \tag{16}$$

where N_f is the number of frequency orientations. In this case $N_f = 3$. Evaluating the correlations separately at each resolution can be used to our advantage. For instance, cropping the image will impact the watermark values in the lower layers more than in the higher layers. This is due to the fact that the bands in higher layers (and the corresponding watermark sequence) correspond to a smaller spatial support. Likewise, any type of low-pass filtering operation will affect the higher layer watermark coefficients more than the lower layer coefficients. In this case the layers with low correlation values would be discarded from the computation of the average. Similarly, the average correlation over a certain frequency orientation is obtained as

$$\rho(f) = \frac{1}{N_l} \sum_{l=1}^{N_l} \rho(l, f), \qquad f = 1, 2, 3 \tag{17}$$

where N_l is the number of levels; in this case $N_l = 4$. By evaluating the correlations separately for each frequency orientation, one can take advantage of any strong structure that is associated with the original image where the watermark sequence would be much stronger than in other parts of the image. The final test statistic is the maximum average correlation value over all the possible resolution levels as well as frequency orientations

$$\rho = \max_{l,f} \{\rho(l), \rho(f)\}. \tag{18}$$

Figs. 4 and 5 show images marked with IA-DCT and IA-W. The corresponding watermark is displayed directly underneath the image. The watermarks in certain images benefit greatly from the image-adaptive nature of the marking procedure. The images in Fig. 4 are fairly uniform (in texture, brightness, etc.) and cannot take full advantage

Fig. 4. (a) IA-DCT watermarked images. (b) IA-DCT watermarks corresponding to (a). (c) IA-W watermarked images. (d) IA-W watermarks corresponding to (c).

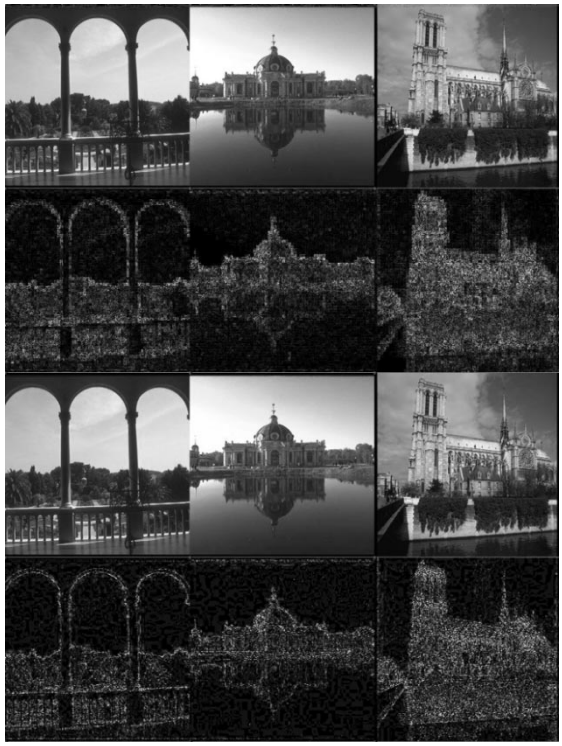


Fig. 5. (a) IA-DCT watermarked images. (b) IA-DCT watermarks corresponding to (a). (c) IA-W watermarked images. (d) IA-W watermarks corresponding to (c).

(a) (b)

(c) (d)

Fig. 6. Uncompressed and JPEG compressed versions of the marked Garden image: (a) uncompressed image; (b) bpp = 0.6, (c) bpp = 0.4; and (d) bpp = 0.18.

of the image-adaptive nature of the watermark. Fig. 5 illustrates sample images that contain highly nonuniform properties where image adaptability is particularly useful. The adaptability is demonstrated by the varying strength of the watermark signal in smooth areas of the image and highly textured areas of the image. In Fig. 4 the watermarks are much less distinct. Note that if a nonimage-adaptive marking method were used for the images of Fig. 5, the maximum amplitude of the mark would have to be much lower than that of the IA algorithms in order to keep the mark imperceptible in the regions of little texture.

C. Robustness to Attacks

Here we discuss the experimental results on the algorithms' robustness to attack. More detailed robustness experiments and comparisons for the IA-DCT and IA-W approaches described here can be found in [4], [5], and [58]. A detailed set of experiments for the spread spectrum approach can be found in [46]. It has been noted earlier that the requirements of transparency and robustness to attack conflict with each other.

1) JPEG Compression: For many applications, it is very likely that a watermarked image may be compressed for more efficient storage or transmission. It is therefore important to examine whether the watermarking schemes can survive JPEG compression. The lowest data rate achievable for an image without producing any visible distortions is highly dependent on the original image, and can range typically from 2 bits per pixel (bpp) for a highly complex scene to 0.2 bpp for a simple scene. An average rate of approximately 1 bpp produces reasonable quality for a typical image.

Table 1
Robustness of IA-DCT to JPEG Compression of Garden Image

Bits per pixel	1.18	0.8	0.6	0.4	0.27	0.18
ρ	0.98	0.61	0.47	0.29	0.17	0.08

Fig. 6 illustrates a watermarked image that has been JPEG compressed to different data rates. The uncompressed watermarked image appears in Fig. 6(a). Fig. 6(b) is the watermarked image compressed from an original 8 to 0.6 bpp. Fig. 6(c) is compressed to 0.4 bpp, and Fig. 6(d) to 0.18 bpp. Table 1 shows the watermark detection values for the IA-DCT scheme at different data rates. The detection results are very good for JPEG compression for all but the poorest quality images. Note that it would be sufficient to be able to detect the watermark only for the data rates that produce acceptable image quality. For instance, the compressed image at 0.18 bpp has little commercial value

Fig. 7. (a) Watermarked Russia image and (b)–(d) three scaled versions.

due to the poor image quality as shown in Fig. 6. A more detailed study of robustness to JPEG compression can be found in [4], [5], and [58], where compression results are shown for a variety of images at different data rates comparing the spread spectrum approach with the IA-DCT and IA-W approaches. All three schemes are shown to be effective against JPEG compression with the IA-W approach yielding the best overall performance.

2) Rescaling: Another common image transformation that may occur to a watermarked image is rescaling. In this case, we examine whether the watermark can survive low-pass filtering followed by subsampling and interpolation. Fig. 7(a) illustrates the undisturbed IA-DCT watermarked image with a decimated version, followed by interpolation shown in Fig. 7(b)–(d). Fig. 7(b) has been decimated by two in each direction, Fig. 7(c) has been decimated by four in each direction, and Fig. 7(d) has been decimated by eight in each direction. In addition, all of the images were previously converted from RGB space to YCbCr space. The watermark detection values for the three decimated examples are given in Table 2. Although the IA-DCT watermark survives the subsampling introduced here, the authors in [5] show that subpixel decimation or interpolation will result in poor watermark detection results for the IA-DCT scheme due to the registration problem for a coherent detection scheme.

3) Cropping: So far, we have examined attacks where in some sense we are throwing away frequency components of the watermarked image signal. In one case we discard frequency information through compression, where the high-frequency components are discarded in order to provide an efficient representation; in the second case we use low-pass filtering in order to avoid aliasing in the rescaling case. It is also interesting to examine the dual case, where we are discarding spatial (instead of spectral) information through cropping, which is a common editing operation. Fig. 8(a) illustrates the undisturbed watermarked image, followed by this image cropped to 1/4 of the original size [Fig. 8(b)], 1/16 of the original size [Fig. 8(c)], and 1/64 of the original size [Fig. 8(d)].

Table 2
Robustness of IA-DCT to Subsampling and Interpolation of Russia Image

Scaling factor:	Unscaled	2	4	8
ρ	1.0	0.9	0.75	0.66

We assume that the original, unmarked image is available for the verification process and we are able to register the cropped watermarked image with the original image. Each cropping experiment results in perfect correlation for the IA-DCT scheme due to the local spatial control that this framework allows. This means that since we can register the original image with the cropped watermarked

Fig. 8. Cropped versions of a watermarked image.

images, we only calculate the correlation value over the cropped area. If the frequency framework were based on a global transform, the verification would be affected since any amount of cropping will affect every watermark value in the image by convolving the watermarked coefficient with a sinc function determined by the window size of the cropping operation. Cropping still affects the capacity of the image because the amount of cropping that is considered acceptable will determine the maximum length watermark that can be recovered from the image. For more details on these experiments as well as others, please refer to [4], [5], and [58]. Additional work presented in [61] examines matching the watermarking and compression frameworks and the effects on watermark robustness.

D. Capacity

Capacity refers to the ability to distinguish multiple watermarks with low probability of error as the number of differently watermarked versions of an image increases. In the IA watermarking procedures, a frequency coefficient is not watermarked if that coefficient is below its JND. The more locations within an image we can mark, the more information bits we can hide and the more unique watermarks we can distinguish with low probability of error. The tradeoff we encounter is capacity versus robustness. The more information bits we try to pack into an image, the less robust the information bits become. For example, if the watermark sequence is 512 elements long and 2048 coefficients in an image are above the JND, then four information bits can be embedded in the image. If the watermark sequence is 1024 bits long, only 2 bits can be inserted. The longer length makes these 2 bits more robust to image processing attacks. Longer sequences provide more robustness at the expense of capacity since the number of bits that we can insert into an image decreases as the length increases.

Table 3
Length of IA-DCT watermark for images in Figs. 4 and 5

Figure	Left	Center	Right
Figure 4	22804	28504	15359
Figure 5	15160	14745	27278

As an example, in Table 3 we show the lengths of the IA-DCT watermarks for the images in Figs. 4 and 5. Note how the length of the sequence varies for each image. This is in contrast to some of the other techniques, such as [46], where the authors suggest using a watermark sequence of length 1000 for all images. The lengths of the watermark sequences for the images in Figs. 4 and 5 are not surprising. The images with large smooth areas [such as Fig. 4 (right) and Fig. 5 (center)] carry shorter sequences than the images with large areas of texture [such as Fig. 4 (center) and Fig. 5 (right)]. This example shows that for the IA schemes, capacity is highly dependent on the

particular image characteristics. Other factors that affect the capacity include the type and amount of degradation that the watermark should survive. An example is cropping, which directly affects the watermark capacity simply by reducing the size of the image available for testing. This issue, as well as the collusion attack, is discussed in further detail in [62].

E. Testing IA-DCT Without the Original Image

A possible revision to IA-DCT as applied to JPEG images [59] is proposed in [63] and avoids the use of the original unmarked image in the verification procedure. In this technique, it is assumed that the original image has already been JPEG compressed. The marking procedure is similar to IA-DCT, except a different subset of DCT coefficients are marked. This subset is called the feature vector, denoted $\{X_F\}$. If a DCT coefficient, $X_D(u, v)$, is larger than 1/2 of its corresponding quantization table value, $Q(u, v)$, it is included in $\{X_F\}$

$$X_D(u, v) \in \{X_F\}, \quad X_D(u, v) > \frac{Q(u, v)}{2}. \tag{19}$$

The watermark W is a sequence of Gaussian distributed random numbers with zero-mean and unit variance that is added to the elements of $\{X_F\}$

$$Y_D(u, v) = \begin{cases} X_D(u, v) + W(u, v), & X_D(u, v) \in \{X_F\} \\ X_D(u, v), & X_D(u, v) \notin \{X_F\}. \end{cases} \tag{20}$$

To verify the presence of W in a test image Z, the feature vector $\{Z_F\}$ is obtained. A correlation measure c is found between $\{Z_F\}$ and W. Let

$$\mu = \frac{\sum_{i=1}^{n} Z_D(u,v)W(u,v)}{n}, Z_D(u,v) \in \{Z_F\} \tag{21}$$

$$\sigma^2 = \frac{\sum_{i=1}^{n} [Z_D(u,v)W(u,v) - \mu]^2}{n-1}, Z_D(u,v) \in \{Z_F\} \tag{22}$$

$$c = \frac{\mu\sqrt{n}}{\sigma}. \tag{23}$$

W is assumed to be uncorrelated with $\{Z_F\}$. A threshold test is performed on c to determine if the W under test is present in Z.

V. Other Image-Adaptive Techniques

A subband watermarking algorithm is described in [30]. An original image, X, is first passed through a Gabor filter bank and the energy E_k in each band k is obtained. Next, an array of random numbers the same size as X is generated, and then low-pass filtered to form G. G is modulated at each center frequency of the filter bank f_k to form G_k. The energy of G_k is D_k. The watermark can then be formed

$$W = \sum_k \alpha_k G_k \tag{24}$$

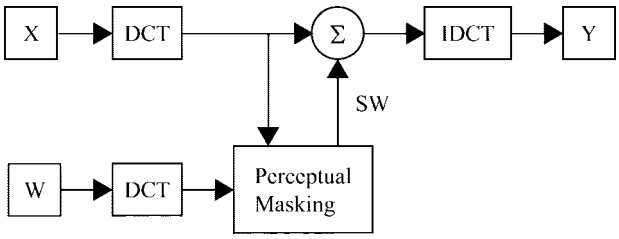


Fig. 9. Block diagram of method in [64].

$$\alpha_k = \begin{cases} 1, & E_k > D_k \\ 0, & E_k \le D_k \end{cases} \tag{25}$$

$$Y = X + W. \tag{26}$$

The verification procedure first passes Y through the filter bank, and attempts to extract G from each subband that had G embedded in it originally. If a sufficient number of subbands are determined to contain the watermark, the image is authentic. This technique is robust to JPEG compression and tolerates common attacks such as low-pass filtering. As in the IA-W technique, certain subbands may be excluded from consideration if *a priori* knowledge of certain attacks is available. The authors have also used their visual model in another watermarking technique described in [31].

The image watermarking technique in [64] first obtains the DCT of an original image X using 8 × 8 blocks. Masking thresholds M are defined for each block based on the DCT coefficients. The watermark for an individual block, $W(b)$, is a reshaped m sequence of size 8 × 8. Different $W(b)$ are used for each block in X. The DCT of $W(b)$ is obtained to form $W_D(b)$, and a scale factor $S(b)$ for the block is derived that is designed to maximize the amplitude of $W(b)$, while keeping each element in $W(b)$ below the corresponding value in M

$$Y_D(u, v, b) = X_D(u, v, b) + S(b) \cdot W_D(u, v, b). \tag{27}$$

The process is repeated for each block. To ensure that the addition of W is imperceptible, spatial domain correction, if necessary, may be employed on the marked block Y as follows. A tolerable error level (TEL) value is obtained for each pixel in X. If

$$|Y(i, j) - X(i, j)| > \text{TEL}(X(i, j)) \tag{28}$$

for any pixel (i, j) in Y, $S(b)$ for the block is reduced and the corresponding block $X(b)$ is remarked with the lower scale factor. Fig. 9 shows the block diagram for this method, where S is the entire array of masking blocks $S(b)$, and W is the array of watermark blocks.

The watermark verification procedure is similar to the IA-W method. A hypothesis test is performed on the normalized cross-correlation coefficient between the extracted watermark W^* and W. If the result is above a threshold, T, the image is authentic. As in previous techniques, T is determined according to the desired probability of detection and probability of false alarm. An original image and a

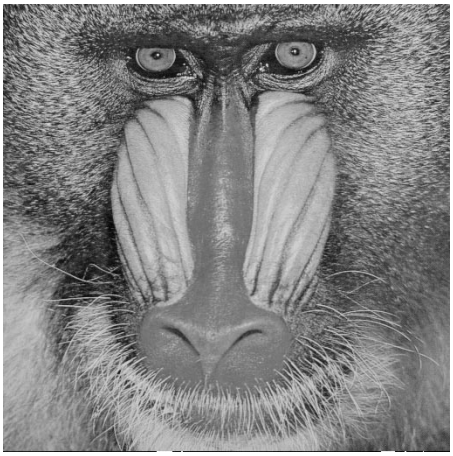


Fig. 10. Original baboon image.

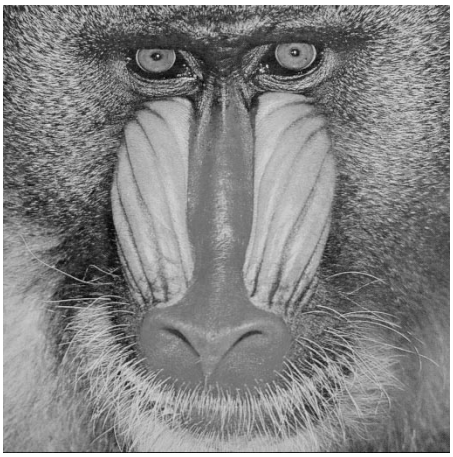


Fig. 11. Watermarked baboon image.

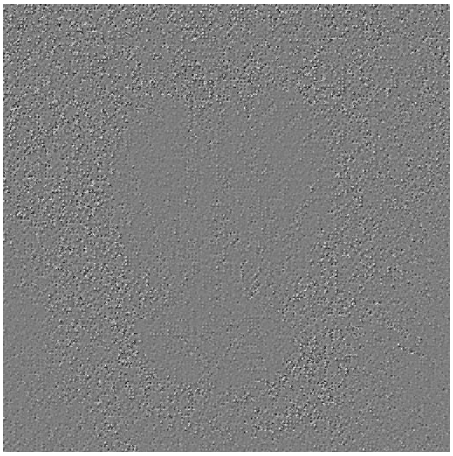


Fig. 12. Watermark for baboon image.

watermarked image are in Figs. 10 and 11; the watermark generated with this technique is shown in Fig. 12.[1]

A wavelet-based algorithm that also scales the watermark on a block basis is described in [65]. The watermark W for this technique is much smaller than the original image; copies of W are tiled throughout the subbands of a wavelet decomposition of X. This protects against cropping, but may be susceptible to collusion attacks. W is a $2m \times 2n$ array of random numbers, where m and n are positive integers. The original image X (or excerpt of X to be marked) should be of size $m2^R \times n2^R$, where R is also a positive integer. This allows for a wavelet decomposition of X up to R levels. A K level decomposition of X is obtained. This forms $3K+1$ wavelet subband images. The lowest resolution subband, of size $m2^{R-K} \times n2^{R-K}$, is not marked. The remaining $3K$ subbands are segmented into $m \times n$ blocks. The subband at the fth frequency orientation and lth resolution level is denoted as $X_{l,f}$; block b in $X_{l,f}$ is denoted $X_{l,f}(b), l = 1, 2, \ldots, K$ and $f = 1,2,3$. A one-level wavelet decomposition of W is then obtained. This produces one $m \times n$ baseband image w_0 and three subband images, denoted w_1, w_2, and w_3, each corresponding to a different frequency orientation.

Each block $X_{l,f}(b)$ is marked with a scaled version of w_f, for $f = 1,2,3$

$$Y_{l,f}(b) = X_{l,f}(b) + \varphi_{l,f}\sqrt{S_{l,f}(b)} \cdot w_f. \tag{29}$$

The scale factor $S_{l,f}(b)$ is defined as the saliency of the block and is a measure of how large the amplitude of the embedded mark can be made while still remaining invisible. $S_{l,f}(b)$ is based on the definition in [66]

$$S_{l,f}(b) = \sum_{u,v} C(u,v)|F_{l,f}(u,v,b)|^2. \tag{30}$$

C is the contrast sensitivity matrix described in [66], and u and v are spatial frequency variables. $F_{l,f}(u,v,b)$ is the Fourier transform of the block $X_{l,f}(b)$

$$\varphi_{l,f} = \frac{\alpha E[X]}{\max_b \sqrt{S_{l,f}(b)}}. \tag{31}$$

α is a user defined constant adjusted to ensure invisibility of the watermark; $\alpha = 0.1$ was used in [65].

To verify an image Z with this technique, the same wavelet decomposition used for X is obtained for Z. The different versions of w_f are extracted from all subbands $Z_{l,f}$. For each frequency orientation f, the versions of w_f are averaged to form w_f^*. W^* is then constructed from the original w_0, and the extracted w_1^*, w_2^*, and w_3^*. A threshold test is then performed on the normalized cross correlation between W and W^* to determine if W is present in Z.

VI. Video Watermarks

There has been extensive research in developing spatial masking models for still images and the use of these models in compression. Developing more sophisticated models, which include temporal masking for the encoding of video sequences, remains an open research problem. Some of this is due to the limitations of current video formats. For instance, in order to take advantage of temporal masking effects, we would need to sample the video at a much higher

[1] These images have been provided by M. Swanson, B. Zhu, and A. Tewfik of the University of Minnesota. Used with permission of the authors.

Fig. 13. Video frames watermarked with IA-DCT interpolation.

rate than the current 30 frames/s. For this reason, much of the work on using visual models for video compression consists of simple rules, such as taking advantage of masking effects at scene changes.

Video watermarking poses some unique requirements not applicable to still image watermarking because of the additional attacks to which that video is subject, such as frame shuffling and interframe collusion, as well as unique applications such as DVD. Another important issue that appears for video watermarking, in particular for watermarking of compressed bit streams, is the constraint on the total data rate of the watermarked sequence and the detection complexity. For example, it is important not to watermark each frame of a video as an independent image. If a different watermark is used for each frame, an attacker can compare frames that change very little within a scene. Simple collusion between these frames can remove a large part of the watermark. Also, the computational effort needed to test each frame in real time could be impractical for many applications. Using the same watermark for each frame also poses problems, since an attacker could then collude with frames from completely different scenes. One method that achieves a tradeoff between marking every frame independently and using the same watermark for each frame is described below. This method, an extension to IA-DCT, marks each I frame in an MPEG sequence and allows the motion vectors to carry the mark over to subsequent frames.

A. Extension of the IA-DCT Technique to Video

A straightforward extension of the IA-DCT technique discussed here for still images has been extended to video [67]. The JND's that are based on spatial masking properties do not, however, apply to temporal masking; this means that watermarking each individual frame based on the spatially derived JND's will result in visible temporal distortion. A possible way to help reduce this effect within the MPEG framework is to take advantage of the motion vectors, which are available in the MPEG compressed bit stream to propagate the watermark from frame to frame. Although this produces visually better results than individual still image watermarking, block artifacts remained in the video sequence. The best visual quality was obtained by using the IA-DCT watermarking technique at every I frame and applying a simple linear interpolation of the watermarks to the frames between two consecutive I frames. The overhead necessary to encode the displaced frame difference (DFD), which now consists of the difference between watermarks as well as the displaced signal difference from the original video sequence is negligible, typically adding only 1–3% in additional data rate. In the watermark encoding process the interpolation technique adds an additional delay consisting of the difference between I frames (typically 15 frames). Several frames of a video sequence that have been encoded using the IA-DCT interpolation method are shown in Fig. 13; Fig. 13(a) represents the watermarked frames, and Fig 13(b) displays the watermark for each individual frame. Ideally, a more formal visual model that takes into account temporal as well as spatial masking should be used in the watermarking process.

B. Watermarking of MPEG-2

In [68], a technique to embed a watermark into an MPEG-2 sequence is presented. The authors propose a watermarking scheme that is similar in spirit to the approach of [46]. For each 8×8 DCT block of the video, a corresponding block is chosen from the watermark signal and transformed using the DCT. The two transformed blocks are added to form the watermarked DCT block. These coefficients are then quantized and Huffman encoded.

The goal is to produce a watermarked video sequence that does not exceed the data rate of the original sequence. The watermarked coefficients are transmitted only if the rate for that coefficient does not increase. The authors claim that typically, 10–20% of the DCT coefficients are altered. Due to the motion-compensated structure of the coder, the authors also address the problem of drift due to the incorrect compensation based on the watermarked blocks. The resulting scheme is of low complexity and allows a watermark signal of several bytes/s.

C. Scene-Adaptive Video Watermarking

Another method watermarks a video sequence on a scene-by-scene basis [69]. A video sequence is first segmented

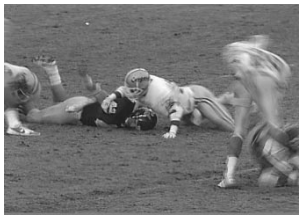


Fig. 14. Original football frame.

Fig. 15. Watermarked football frame.

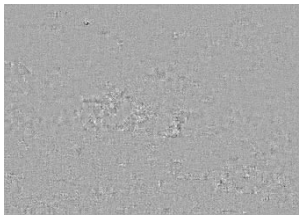


Fig. 16. Watermark for football frame.

into scenes, and a temporal wavelet transform of each scene is obtained. The elements of the video that change little temporally (for example, background) are located in the low time-resolution bands of the transform. The parts with fast motion would be decomposed into higher resolution bands. This effectively isolates the two types of motion in the scene. A different watermark is then placed in each resolution band. This results in the same watermark being embedded in the slow-changing parts of a scene to avoid collusion. The watermarks added to the video with large amounts of motion are different and change frequently throughout the scene. Different watermarks are placed in each scene (except perhaps for scenes with near-identical backgrounds). An original video frame, marked frame, and the corresponding watermark are shown in Figs. 14–16.[2]

VII. Watermarking Standards

A. Standards

The digital versatile disk (DVD) [70] is the latest technology that has been developed to deliver data to the consumer. One problem that had delayed the development of the DVD standard is the protection of copyrighted content [71]. Several media companies initially refused to provide DVD material until the copy-protection problem had been addressed. One way to secure the content on a DVD is to link a watermark verification process to the proper functioning of the DVD player [72]. For instance, the player's output port would be enabled only upon verification of the watermark.

There are several ongoing efforts in standardizing copy protection technology. Most of these involve the use of watermarking and/or encryption techniques or other mechanisms including analog approaches for making images or video either not viewable or recordable. The Data Hiding Subgroup (DHSG) of the Copy Protection Technical Working Group (CPTWG) has issued several calls for proposals in the area of data hiding and watermarking [73]. Another group that is very interested in the copy protection of images and video is the Copyright Issues Group of the Digital Audiovisual Council (DAVIC) [74]. The study of intellectual property rights is also part of the developing MPEG-7 standard [75].

[2] Different frames from the same sequences have originally appeared in [69]. Used with permission of the authors.

VIII. Conclusion

We have described recent developments in the digital watermarking of images and video where the watermarking schemes are designed to exploit properties of the human visual system in order to provide extremely high-quality original content (transparent watermarks) while providing an effective means of protecting intellectual property rights (robustness to attacks). More work is needed to describe the performance criteria of these techniques with respect to robustness and capacity as the applications become better defined. In the case of video watermarking, advances in understanding the psychovisual aspects of spatiotemporal masking could make a significant impact on designing more effective watermarking schemes.

The protection of intellectual property rights is perhaps one of the last major barriers to the "digital world." There is hope.

Acknowledgment

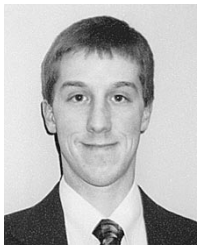
The authors would like to thank Prof. A. Tewfik, B. Zhu, M. Swanson, M. Kobayashi, and B. Chau for providing an advanced copy of [7], as well as example images from their work.

References

[1] E. J. Delp, "Watermarking: Who cares and does it work?" presented at the ACM Multimedia and Security Workshop, Bristol, U.K., Sept. 1998.
[2] R. Anderson and F. Petitcolas, "On the limits of steganography," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 463–473, May 1998.
[3] E. T. Lin and E. J. Delp, "A review of data hiding in digital images," in *Proc. Image Processing, Image Quality, Image Capture and Systems Conf. (PICS)*, Savannah, GA, Apr. 25–28, 1999, pp. 274–278.
[4] C. I. Podilchuk and W. Zeng, "Digital image watermarking using visual models," in *Proc. SPIE Int. Conf Human Vision and Electronic Imaging II*, vol. 3016, San Jose, CA, Feb. 10–13, 1997, pp. 100–111.
[5] ——, "Image-adaptive watermarking using visual models," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 525–539, May 1998.
[6] R. B. Wolfgang and E. J. Delp, "Overview of image security techniques with applications in multimedia systems," in *Proc.*

SPIE Int. Conf. Multimedia Networks: Security, Displays, Terminals, and Gateways, vol. 3228, Dallas, TX, Nov. 4/5, 1997, pp. 297–308.
[7] M. D. Swanson, M. Kobayashi, and A. H. Tewfik, "Multimedia data embedding and watermarking technologies," *Proc. IEEE*, vol. 86, pp. 1064–1087, June 1998.
[8] I. J. Cox and M. L. Miller, "A review of watermarking and the importance of perceptual modeling," in *Proc. SPIE Int. Conf. Human Vision and Electronic Imaging—II*, San Jose, CA, Feb. 10–13, 1997, pp. 92–99.
[9] B. Schneier, *Applied Cryptography*, 2nd ed. New York: Wiley, 1996.
[10] F. Mintzer, G. W. Braudaway, and M. Yeung, "Effective and ineffective digital watermarks," in *Proc. IEEE Int. Conf. Image Processing*, vol. 3, Santa Barbara, CA, Oct. 26–29, 1997, pp. 9–12.
[11] W. J. Mitchell, *The Reconfigured Eye: Visual Truth in the Post-Photographic Era*. Cambridge, MA: MIT Press, 1992.
[12] F. Petitcolas and R. Anderson, "Weaknesses of copyright marking systems," in *Proc. ACM Multimedia and Security Workshop (ACM Multimedia'98)*, Bristol, UK, Sept. 1998, pp. 55–62. [Online]. Available WWW: http://www.cl.cam.ac.uk/~fapp2/watermarking/stirmark.
[13] D. Stinson, *Cryptography, Theory and Practice*. Boca Raton, FL: CRC Press, 1995.
[14] S. Craver, N. Memon, B.-L. Yeo, and M. Yeung, "Can invisible watermarks resolve rightful ownerships?" in *Proc. SPIE Int. Conf. Storage and Retrieval for Image and Video Databases—V*, vol. 3022, San Jose, CA, Feb. 13–14, 1997, pp. 310–321.
[15] T. G. Stockham, "Image processing in the context of a visual model," *Proc. IEEE*, vol. 60, pp. 828–841, July 1972.
[16] N. S. Jayant, J. D. Johnston, and R. J. Safranek, "Signal compression based on models of human perception," *Proc. IEEE*, vol. 81, pp. 1385–1422, Oct. 1993.
[17] T. N. Cornsweet, *Visual Perception*. New York: Academic, 1970.
[18] C. R. Carlson and R. Cohen, "A simple psychophysical model for predicting the visibility of displayed information," in *Proc. Society for Information Display*, vol. 21, 1980, pp. 229–245.
[19] A. B. Watson, "Efficiency of an image code based on human vision," *J. Opt. Soc. Amer. A*, vol. 4, no. 12, pp. 2401–2417, Dec. 1987.
[20] ——, "The cortex transform: Rapid computation of simulated neural images," *Comput. Vision, Graphics, Image Processing*, vol. 39, no. 3, pp. 311–327, Sept. 1987.
[21] G. E. Legge and J. M. Foley, "Contrast masking in human vision," *J. Opt. Soc. Amer.*, vol. 70, no. 12, pp. 1458–1471, Dec. 1980.
[22] M. Vetterli and J. Kovacevic, *Wavelets and Subband Coding*. Englewood Cliffs, NJ: Prentice-Hall, 1995.
[23] W. B. Pennebaker and J. L. Mitchell, *JPEG: Still Image Data Compression Standard*. New York: Van Nostrand Reinhold, 1993.
[24] H. A. Peterson, A. J. Ahumada, Jr., and A. B. Watson, "Improved detection model for DCT coefficient quantization," in *Proc. SPIE Int. Conf. Human Vision, Visual Processing, and Digital Display—IV*, vol. 1913, Feb. 1993, pp. 191–201.
[25] A. B. Watson, "DCT quantization matrices visually optimized for individual images," in *Proc. SPIE Int. Conf. Human Vision, Visual Processing and Digital Display—IV*, vol. 1913, Feb. 1993, pp. 202–216.
[26] R. J. Safranek, "JPEG compliant encoder utilizing perceptually based quantization," in *Proc. SPIE Int. Conf. Human Vision, Visual Processing, and Digital Display—V*, vol. 2179, San Jose, CA, May 1994, pp. 117–126.
[27] R. J. Safranek and J. D. Johnston, "A perceptually tuned sub-band image coder with image-dependent quantization and post-quantization data compression," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, Glasgow, U.K., 1989, pp. 1945–1948.
[28] B. Zhu and A. H. Tewfik, "Low bit rate near-transparent image coding," in *Proc. SPIE Int. Conf. Wavelet Applications for Dual Use*, vol. 2491, Orlando, FL, 1995, pp. 173–184.
[29] B. Girod, "The information theoretical significance of spatial and temporal masking in video signals," in *Proc. SPIE Int. Conf. Human Vision, Visual Processing, and Digital Display*, vol. 1077, 1989, pp. 178–187.
[30] J.-F. Delaigle, C. De Vleeschouwer, and B. Macq, "A psychovisual approach for digital picture watermarking," *J. Electronic Imaging*, vol. 7, no. 3, pp. 628–640, July 1998.
[31] ——, "Watermarking algorithm based on a human visual model," *Signal Processing*, vol. 66, no. 3, pp. 319–335, May 1998.
[32] R. G. van Schyndel, A. Z. Tirkel, N. Mee, and C. F. Osborne, "A digital watermark," in *Proc. IEEE Int. Conf. Image Processing*, vol. 2, Austin, TX, Nov. 1994, pp. 86–90.
[33] U.-C. G. Fiebig and M. Schnell, "Correlation properties of extended m-sequences," *Electron. Lett.*, vol. 29, no. 20, pp. 1753–1755, Sept. 30, 1993.
[34] R. B. Wolfgang and E. J. Delp, "A watermark for digital images," in *Proc. IEEE Int. Conf. Image Processing*, vol. 3, Lausanne, Switzerland, Sept. 16–19, 1996, pp. 219–222.
[35] ——, "Techniques for watermarking digital imagery: Further studies," in *Proc. Int. Conf. Imaging Science, Systems, and Technology*, vol. 1, Las Vegas, NV, June 30–July 3, 1997, pp. 279–287.
[36] ——, "Fragile watermarking using the VW2D watermark," in *Proc. SPIE Int. Conf. Security and Watermarking of Multimedia Contents*, vol. 3657, San Jose, CA, Jan. 25–27, 1999.
[37] N. Nikolaidis and I. Pitas, "Robust image watermarking in the spatial domain," *Signal Processing*, vol. 66, no. 3, pp. 385–403, May 1998.
[38] W. Bender, D. Gruhl, N. Morimoto, and A. Lu, "Techniques for data hiding," *IBM Syst. J.*, vol. 35, nos. 3/4, pp. 313–336, 1996.
[39] D. J. Fleet and D. J. Heeger, "Embedding invisible information in color images," in *Proc. IEEE Int. Conf. Image Processing*, vol. 1, Santa Barbara, CA, Oct. 26–29, 1997, pp. 532–535.
[40] M. Kutter, F. D. Jordan, and F. Bossen, "Digital signature of color images using amplitude modulation," in *Proc. SPIE Int. Conf. Storage and Retrieval for Image and Video Databases—V*, vol. 3022, San Jose, CA, Feb. 13–14, 1997, pp. 518–526.
[41] G. W. Braudaway, "Protecting publicly-available images with an invisible image watermark," in *Proc. IEEE Int. Conf. Image Processing*, vol. 1, Santa Barbara, CA, Oct. 26–29, 1997, pp. 524–527.
[42] M. Yeung and F. Mintzer, "An invisible watermarking technique for image verification," in *Proc. IEEE Int. Conf. Image Processing*, vol. 2, Santa Barbara, CA, Oct. 26–29, 1997, pp. 680–683.
[43] P. Kuosmanen, J. Astola, K. Davídsson, and K. Halonen, "Digital watermarking through filtering," in *Proc. IEEE Nonlinear Signal and Image Processing Workshop, CD ROM*, Mackinac Island, MI, Sept. 8–10, 1997.
[44] K. T. Knox and S. Wang, "Digital watermarks using stochastic screens—A halftoning watermark," in *Proc. SPIE Int. Conf. Storage and Retrieval for Image and Video Databases—V*, vol. 3022, San Jose, CA, Feb. 13–14, 1997, pp. 310–316.
[45] G. C. Langelaar, J. van der Lubbe, and R. L. Lagendijk, "Robust labeling methods for copy protection of images," in *Proc. SPIE Int. Conf. Storage and Retrieval for Image and Video Databases—V*, vol. 3022, San Jose, CA, Feb. 13–14, 1997, pp. 298–309.
[46] I. Cox, J. Kilian, F. T. Leighton, and T. Shamoon, "Secure spread spectrum watermarking for multimedia," *IEEE Trans. Image Processing*, vol. 6, pp. 1673–1687, Dec. 1997.
[47] X.-G. Xia, C. G. Boncelet, and G. R. Arce, "A multiresolution watermark for digital images," in *Proc. IEEE Int. Conf. Image Processing*, Santa Barbara, CA, Oct. 26–29, 1997, vol. 3, pp. 548–551.
[48] F. M. Boland, J. J. K. Ó. Ruanaidh, and C. Dautzenberg, "Watermarking digital images for copyright protection," in *Proc. Int. Conf. Image Processing and its Applications*, Edinburgh, U.K., July 1995, pp. 321–326.
[49] M. Barni, F. Bartolini, V. Cappellini, and A. Piva, "A DCT-domain system for robust image watermarking," *Signal Processing*, vol. 66, no. 3, pp. 357–372, May 1998.
[50] E. Koch and J. Zhao, "Toward robust and hidden image copyright labeling," in *Proc. IEEE Workshop Nonlinear Signal and Image Processing*, Neos Marmaras, Halkidiki, Greece, June 20–22, 1995, pp. 452–455.
[51] S. Burgett, E. Koch, and J. Zhao, "Copyright labelling of digitized image data," *IEEE Commun. Mag.*, vol. 36, pp. 94–100, Mar. 1998.

[52] A. G. Bors and I. Pitas, "Image watermarking using DCT domain constraints," in *Proc. IEEE Int. Conf. Image Processing*, vol. 3, Lausanne, Switzerland, Sept. 16–19, 1996, pp. 231–234.
[53] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*. Boston, MA: Kluwer, 1992.
[54] M. Holliman, N. Memon, B.-L. Yeo, and M. Yeung, "Adaptive public watermarking of DCT-based compressed images," in *Proc. SPIE Int. Conf. Storage and Retrieval for Image and Video Databases—VI*, vol. 3312, San Jose, CA, Jan. 28–30, 1998, pp. 284–295.
[55] D. Benham, N. Memon, B.-L. Yeo, and M. Yeung, "Fast watermarking of DCT-based compressed images," in *Proc. Int. Conf. Imaging Science, Systems, and Technology*, vol. 1, Las Vegas, NV, June 30–July 3, 1997, pp. 243–252.
[56] J. J. K. O. Ruanaidh, W. J. Dowling, and F. M. Boland, "Phase watermarking of digital images," in *Proc. IEEE Int. Conf. Image Processing*, vol. 3, Lausanne, Switzerland, Sept. 16–19, 1996, pp. 239–242.
[57] J. J. K. Ó. Ruanaidh and T. Pun, "Rotation, scale and translation invariant spread spectrum digital image watermarking," *Signal Processing*, vol. 66, no. 3, pp. 303–317, May 1998.
[58] C. I. Podilchuk and W. Zeng, "Perceptual watermarking of still images," in *Proc. IEEE Workshop Multimedia Signal Processing, CD ROM*, Princeton, NJ, June 1997.
[59] ——, "Watermarking of the JPEG bitstream," in *Proc. Int. Conf. Imaging Science, Systems, and Technology*, Las Vegas, NV, June 30–July 3, 1997, pp. 253–260.
[60] A. B. Watson, G. Y. Yang, J. A. Solomon, and J. Villasenor, "Visibility of wavelet quantization noise," *IEEE Trans. Image Processing*, vol. 6, pp. 1164–1175, Aug. 1997.
[61] R. B. Wolfgang, C. I. Podilchuk, and E. J. Delp, "The effect of matching watermark and compression transforms in compressed color images," in *Proc. IEEE Int. Conf. Image Processing*, vol. 1, Oct. 4–7, 1998, pp. 440–444.
[62] S. Servetto, C. Podilchuk, and K. Ramchandran, "Capacity issues in digital image watermarking," in *Proc. IEEE Int. Conf. Image Processing*, Chicago, IL, Oct. 4–7, 1998, vol. 1, pp. 445–449.
[63] W. Zeng and B. Liu, "On resolving rightful ownerships of digital images by invisible watermarks," in *Proc. IEEE Int. Conf. Image Processing*, vol. 1, Santa Barbara, CA, Oct. 26–29, 1997, pp. 552–555.
[64] M. D. Swanson, B. Zhu, and A. H. Tewfik, "Transparent robust image watermarking," in *Proc. 1996 Int. Conf. Image Processing*, vol. 3, Lausanne, Switzerland, Sept. 16–19, 1996, pp. 211–214.
[65] D. Kundur and D. Hatzinakos, "A robust digital image watermarking method using wavelet-based fusion," in *Proc. IEEE Int. Conf. Image Processing*, vol. 1, Santa Barbara, CA, Oct. 26–29, 1997, pp. 544–547.
[66] T. A. Wilson, S. K. Rogers, and L. R. Myers, "Perceptual based hyperspectral image fusion using multiresolution analysis," *Opt. Eng.*, vol. 34, no. 11, pp. 3154–3164, Nov. 1995.
[67] S. Servetto and C. Podilchuk, Lucent Technologies Bell Labs, Murray Hill, NJ, Lucent Technologies Internal Tech. Memo, 1998.
[68] F. Hartung and B. Girod, "Watermarking of uncompressed and compressed video," *Signal Processing*, vol. 66, no. 3, pp. 283–301, May 1998.
[69] M. D. Swanson, B. Zhu, and A. H. Tewfik, "Multiresolution scene-based video watermarking using perceptual models," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 540–550, May 1998.
[70] J. Taylor, *DVD Demystified: The Guidebook for DVD-Video and DVD-Rom*. New York: McGraw Hill, 1997.
[71] S. Rupley, "What's holding up DVD?," *PC Mag.*, vol. 15, no. 20, p. 34, Nov. 19, 1996.
[72] I. J. Cox and J.-P. Linnartz, "Public watermarks and resistance to tampering," in *Proc. IEEE Int. Conf. Image Processing*, Santa Barbara, CA ,Oct. 26–29, 1997.
[73] Data Hiding Subgroup of the Copy Protection Technical Working Group. (1997, July 1). Call for proposals. [Online]. Available WWW: http://www.dvcc.com/dhsg.
[74] Digital Audio Visual Council (DAVIC). Call for proposals. [Online]. Available WWW: http://www.davic.org.
[75] "MPEG-7: Context and objectives (version 6—San Jose)," International Standards Org., San Jose, CA, Doc. ISO/IEC JTC1/SC29/WG11 N2082, Feb. 1998.



Raymond B. Wolfgang (Student Member, IEEE) was born in West Chester, PA. He received the B.S.E.E. degree (with distinction) and the M.S. degree in mechanical engineering from the Pennsylvania State University, University Park, in 1993. He is currently pursuing the Ph.D. degree in electrical and computer engineering at Purdue University, West Lafayette, IN.

In the summer of 1991, he worked at the August Rüggeberg Company, Marienheide, Germany. His research interests include image and video processing, communications, and multimedia security.

Mr. Wolfgang graduated as a University Scholar from the Pennsylvania State University. He is a member of Tau Beta Pi and Eta Kappa Nu.

Christine I. Podilchuk (Member, IEEE) received the B.S., M.S., and Ph.D. degrees in electrical engineering from Rutgers University, New Brunswick, NJ, in 1984, 1986, and 1988, respectively.

Since then, she has been working in the research division of Bell Labs, Lucent Technologies, Murray Hill, NJ (formerly part of AT&T), in the Visual Communications Research Department. Her research interests are in the general area of image processing and computer vision. She is particularly interested in video compression and transmission, image detection and recognition, image restoration, and digital watermarking of multimedia.

Dr. Podilchuk served as an Associate Editor of IEEE TRANSACTIONS ON IMAGE PROCESSING from 1996 to 1998. She is currently a member of the IEEE Image and Multidimensional Signal Processing Technical Committee.

Edward J. Delp (Fellow, IEEE) was born in Cincinnati, OH. He received the B.S.E.E. (cum laude) and M.S. degrees from the University of Cincinnati, OH, and the Ph.D. degree from Purdue University, West Lafayette, IN.

From 1980 to 1984, he was with the Department of Electrical and Computer Engineering at the University of Michigan, Ann Arbor. Since August 1984, he has been with the School of Electrical and Computer Engineering at Purdue University, West Lafayette, IN, where he is a Professor of electrical engineering. During the summer of 1998, he was a Visiting Professor at the Tampere International Center for Signal Processing at the Tampere University of Technology in Finland. He has consulted for various companies and government agencies in the areas of signal and image processing, robot vision, pattern recognition, and secure communications. His research interests include image and video compression, image processing, multimedia security, medical imaging, parallel processing, multimedia systems, nonlinear filtering, communication, and information theory.

Dr. Delp is a member of Tau Beta Pi, Eta Kappa Nu, Phi Kappa Phi, Sigma Xi, ACM, and the Pattern Recognition Society. He is a Fellow of the SPIE and a Fellow of the Society for Imaging Science and Technology (IS&T). In 1997, he was elected Chair of the Image and Multidimensional Signal Processing (IMDSP) Technical Committee of the IEEE Signal Processing Society. From 1994 to 1998, he was Vice-President for Publications of IS&T. He was Cochair of the 1999 SPIE/IS&T Conference on Security and Watermarking of Multimedia Contents, the General Cochair of the 1997 Visual Communications and Image Processing Conference (VCIP), the Program Chair of the IEEE Signal Processing Society's 9th IMDSP Workshop, and the General Cochairman of the 1993 SPIE/IS&T Symposium on Electronic Imaging. From 1984 to 1991, he was a member of the editorial board of the *International Journal of Cardiac Imaging*. In the past, he has been an Associate Editor of IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE and IEEE TRANSACTIONS ON IMAGE PROCESSING. Since 1992, he has been a member of the editorial board of the journal *Pattern Recognition*. Since 1994, he has been an Associate Editor of the *Journal of Electronic Imaging*. In 1990, he received the Honeywell Award and in 1992 the D. D. Ewing Award for excellence in teaching.