

# Multimedia Watermarking Techniques

FRANK HARTUNG, STUDENT MEMBER, IEEE, AND MARTIN KUTTER

## *Invited Paper*

*Multimedia watermarking technology has evolved very quickly during the last few years. A digital watermark is information that is imperceptibly and robustly embedded in the host data such that it cannot be removed. A watermark typically contains information about the origin, status, or recipient of the host data. In this tutorial paper, the requirements and applications for watermarking are reviewed. Applications include copyright protection, data monitoring, and data tracking. The basic concepts of watermarking systems are outlined and illustrated with proposed watermarking methods for images, video, audio, text documents, and other media. Robustness and security aspects are discussed in detail. Finally, a few remarks are made about the state of the art and possible future developments in watermarking technology.*

**Keywords**— Audio, image, multimedia, review, video, watermarking.

## I. INTRODUCTION

Multimedia production and distribution, as we see it today, is all digital, from the authoring tools of content providers to the receivers. The advantages of digital processing and distribution, like noise-free transmission, software instead of hardware processing, and improved reconfigurability of systems, are all well known and obvious. Not so obvious are the disadvantages of digital media distribution. For example, from the viewpoint of media producers and content providers, the possibility for unlimited copying of digital data without loss of fidelity is undesirable because it may cause considerable financial loss. Digital copy protection or copy prevention mechanisms are only of limited value because access to cleartext versions of protected data must at least be granted to paying recipients which can then produce and distribute illegal copies. Technical attempts to prevent copying have in reality always been circumvented.

One remaining method for the protection of intellectual property rights (IPR) is the embedding of digital watermarks into multimedia data. The watermark is a digital code

unremovably, robustly, and imperceptibly embedded in the host data and typically contains information about origin, status, and/or destination of the data. Although not directly used for copy protection, it can at least help identifying source and destination of multimedia data and, as a “last line of defense,” enable appropriate follow-up actions in case of suspected copyright violations.

While copyright protection is the most prominent application of watermarking techniques, others exist, including data authentication by means of fragile watermarks which are impaired or destroyed by manipulations, embedded transmission of value added services within multimedia data, and embedded data labeling for other purposes than copyright protection, such as data monitoring and tracking. An example for a data-monitoring system is the automatic registration and monitoring of broadcasted radio programs such that royalties are automatically paid to the IPR owners of the broadcast data.

The development of watermarking methods involves several design tradeoffs. Watermarks should be robust against standard data manipulations, including digital-to-analog conversion and digital format conversion. Security is a special concern, and watermarks should resist even attempted attacks by knowledgeable individuals. On the other hand, watermarks should be imperceptible and convey as much information as possible. In general, watermark embedding and retrieval should have low complexity because for various applications, real-time watermarking is desirable. All of these (partly contradicting) requirements and the resulting design constraints will be discussed in more detail throughout the paper.

The paper is organized as follows. Section II gives an introductory explanation of the terms used, as well as a few remarks about the historical aspects of watermarking. In Section III, common design requirements and principles are explained that apply to all watermarking techniques, independent of the actual application. Sections IV–VII review various watermarking techniques that have been proposed for formatted text data, images, video, and audio, respectively. Watermarking of other media, including three dimensional (3-D) data and 3-D animation parameters, is discussed in Section VIII. Section IX gives detailed insight

Manuscript received October 20, 1997; revised March 26, 1998.

F. Hartung was with the Telecommunications Laboratory, University of Erlangen–Nuremberg, 91058 Erlangen, Germany. He is now with Ericsson Eurolab, Research Department, 52134 Herzogenrath, Germany.

M. Kutter is with Signal Processing Laboratory, Swiss Federal Institute of Technology, 1015 Lausanne, Switzerland.

Publisher Item Identifier S 0018-9219(99)05174-9.

into security issues, namely attacks against watermarks, and shows the relations between watermarking and cryptology. In Section X, we extrapolate the recent development of watermarking technology and watermarking applications and try to forecast future trends. Section XI summarizes and concludes this paper on multimedia watermarking techniques.

## II. STEGANOGRAPHY AND WATERMARKING—HISTORY AND TERMINOLOGY

### A. History

The idea to communicate secretly is as old as communication itself. First stories, which can be interpreted as early records of covert communication, appear in the old Greek literature, for example, in Homer's *Iliad*, or in tales by Herodotus. The word "steganography," which is still in use today, derives from the Greek language and means covert communication. Kobayashi [67] and Petitcolas *et al.* [99] have investigated the history of covert communication in great detail, including the broad use of techniques for secret and covert communication before and during the two World Wars, and steganographic methods for analog signals. Although the historical background is very interesting, we do not cover it here in detail. Please refer to [67] and [99] for an in-depth investigation of historic aspects.

Paper watermarks appeared in the art of handmade papermaking nearly 700 years ago. The oldest watermarked paper found in archives dates back to 1292 and has its origin in Fabriano, Italy, which is considered the birthplace of watermarks. At the end of the thirteenth century, about 40 paper mills were sharing the paper marked in Fabriano and producing paper with different format, quality, and price. They produced raw, coarse paper which was smoothed and postprocessed by artisans and sold by merchants. Competition not only among the paper mills but also among the artisans and merchants was very high, and it was difficult to keep track of paper provenance and thus format and quality identification. The introduction of watermarks helped avoiding any possibility of confusion. After their invention, watermarks quickly spread over Italy and then over Europe, and although originally used to indicate the paper brand or paper mill, they later served as indication for paper format, quality, and strength and were also used to date and authenticate paper. A nice example illustrating the legal power of watermarks is a case in 1887 in France called "Des Decorations" [41]. The watermarks of two letters, presented as pieces of evidence, proved that the letters had been predated and resulted in considerable sensation and, in the end, in the resignation of President Grévy. For more information on paper watermarks, watermark history, and related legal issues, please refer to [144], an extensive listing of over 500 references.

The analogy between paper watermarks, steganography, and digital watermarking is obvious, and in fact, paper watermarks in money bills or stamps [135] actually inspired the first use of the term watermarking in the context of digital data.

The idea of digital image watermarking arose independently in 1990 [131], [132] and around 1993 [20], [136]. Tirkel *et al.* [136] coined the word "water mark" which became "watermark" later on. It took a few more years until 1995/1996 before watermarking received remarkable attention. Since then, digital watermarking has gained a lot of attention and has evolved very quickly, and while there are a lot of topics open for further research, practical working methods and systems have been developed. In this paper, we introduce the concepts and illustrate them with some of the work that has been published. While attempting to be as complete as possible, we can still only give a rough overview.

### B. Terminology

Today, we are of course concerned with digital communication. As in classical analog communication, also in digital communication there is interest for methods that allow the transmission of information hidden or embedded in other data. While such techniques often share similar principles and basic ideas, there are also important distinguishing features, mainly in terms of robustness against attacks. Several names have been coined for such techniques. However, the terms are often confused, and therefore it is necessary to clarify the differences.

*Steganography* stands for techniques in general that allow secret communication, usually by embedding or hiding the secret information in other, unsuspected data. Steganographic methods generally do rely on the assumption that the existence of the covert communication is unknown to third parties and are mainly used in secret point-to-point communication between trusting parties. As a result, steganographic methods are in general not robust, i.e., the hidden information cannot be recovered after data manipulation.

*Watermarking*, as opposed to steganography, has the additional notion of robustness against attacks. Even if the existence of the hidden information is known it is difficult—ideally impossible—for an attacker to destroy the embedded watermark, even if the algorithmic principle of the watermarking method is public. In cryptography, this is known as *Kerckhoffs law*: a cryptosystem should be secure, even if an attacker knows the cryptographic principles and methods used but does not have the appropriate key [117]. A practical implication of the robustness requirement is that watermarking methods can typically embed much less information into host data than steganographic methods. Steganography and watermarking are thus more complementary than competitive approaches. In the remainder of this paper, we focus on watermarking methods and not on steganographic methods in general. For an overview of steganographic methods the reader is referred to [67], [99], and [124].

*Data hiding* and *data embedding* are used in varying contexts, but they do typically denote either steganography or applications "between" steganography and watermarking, which means applications where the existence of the embedded data are publicly known, but there is no need

to protect it. This is typically the case for the embedded transmission of auxiliary information or services [125] that are publicly available and do not relate to copyright protection or conditional access functionalities.

*Fingerprinting* and *labeling* are terms that denote special applications of watermarking. They relate to copyright protection applications where information about originator and recipient of digital data is embedded as watermarks. The individual watermarks, which are unique codes out of a series of codes, are called “fingerprints” or “labels.”

*Bit-stream watermarking* is sometimes used for data hiding or watermarking of compressed data, for example, compressed video.

The term *embedded signatures* has been used instead of “watermarking” in early publications. Because it potentially leads to confusion with cryptographic digital signatures [117], it is usually not used anymore. Cryptographic signatures serve for authentication purposes. They are used to detect alterations of the signed data and to authenticate the sender. Watermarks, however, are only in special applications used for authentication and are usually designed to *resist* alterations and modifications.

*Visible watermarks*, as the name says, are visual patterns, like logos, which are inserted into or overlaid on images (or video), very similar to visible paper watermarks. However, the name is confusing since visible watermarks are not watermarks in the sense of this paper. Visible watermarks are mainly applied to images, for example, to visibly mark preview images available in image databases or on the World Wide Web in order to prevent people from commercial use of such images. A visible watermarking method devised by Braudaway *et al.* [16] combines the watermark image with the original image by modifying the brightness of the original image as a function of the watermark and a secret key. The secret key determines pseudorandom scaling values used for the brightness modification in order to make it difficult for attackers to remove the visible mark.

### III. DIGITAL WATERMARKING

#### A. Requirements

The basic requirements in watermarking apply to all media and are very intuitive.

- 1) A watermark shall convey as much information as possible, which means the watermark data rate should be high.
- 2) A watermark should in general be secret and should only be accessible by authorized parties. This requirement is referred to as security of the watermark and is usually achieved by the use of cryptographic keys.
- 3) A watermark should stay in the host data regardless of whatever happens to the host data, including all possible signal processing that may occur, and including all hostile attacks that unauthorized parties may attempt. This requirement is referred to as robustness of the watermark. It is a key requirement for copyright protection or conditional access applications, but less important for applications where the watermarks

are not required to be cryptographically secure, for example, for applications where watermarks convey public information.

- 4) A watermark should, though being unremovable, be imperceptible.

Depending on the media to be watermarked and the application, this basic set of requirements may be supplemented by additional requirements.

- 1) Watermark recovery may or may not be allowed to use the original, unwatermarked host data.
- 2) Depending on the application, watermark embedding may be required in real time, e.g., for video fingerprinting. Real-time embedding again may, for complexity reasons, require compressed-domain embedding methods.
- 3) Depending on the application, the watermark may be required to be able to convey arbitrary information. For other applications, only a few predefined watermarks may have to be embedded, and for the decoder it may be sufficient to check for the presence of one of the predefined watermarks (hypothesis testing).

In the following, a few of the mentioned requirements and the resulting design issues are highlighted in more detail.

1) *Watermark Security and Keys*: If security, i.e., secrecy of the embedded information, is required, one or several secret and cryptographically secure keys have to be used for the embedding and extraction process. For example, in many schemes, pseudorandom signals are embedded as watermarks. In this case, the description and the seed of the pseudorandom number generator may be used as key. There are two levels of secrecy. In the first level, an unauthorized user can neither read or decode an embedded watermark nor can he detect if a given set of data contains a watermark. The second level permits unauthorized users to detect if data are watermarked, however, the embedded information cannot be read without having the secret key. Such schemes can, for example, embed two watermarks, one with a public key and the other with a secret key. Alternatively, a scheme has been proposed which combines one or several public keys with a private key and embeds one combined public/private watermark, rather than several watermarks [48]. When designing an overall copyright protection system, issues like secret key generation, distribution, and management (possibly by trusted third parties), as well as other system integration aspects have to be considered.

2) *Robustness*: In the design of any watermarking scheme, watermark robustness is typically one of the main issues, since robustness against data distortions introduced through standard data processing and attacks is a major requirement. Standard data processing includes all data manipulation and modification that the data might undergo in the usual distribution chain, such as data editing, printing, enhancement, and format conversion. “Attack” denotes data manipulation with the purpose of impairing, destroying, or removing the embedded watermarks. Section IX-B below revisits attacks and gives remedies that help to make watermarks attack resistant.

Although it is possible to design robust watermarking techniques, it should be noted that a watermark is only robust as long as it is not public, which means as long as it cannot be read by everyone. If watermark detector principle and key are public, and even if only a “black-box” watermark detector is public, the watermark is vulnerable to attacks [28], [64]. Hence, public watermarks, as sometimes proposed in the literature, are not robust unless every receiver uses a different key. This however is difficult in practice and gives rise to collusion attacks.

3) *Imperceptibility*: One of the main requirements for watermarking is the perceptual transparency. The data embedding process should not introduce any perceptible artifacts into the host data. On the other hand, for high robustness, it is desirable that the watermark amplitude is as high as possible. Thus, the design of a watermarking method always involves a tradeoff between imperceptibility and robustness. It would be optimal to embed a watermark just below the threshold of perception. However, this threshold is difficult to determine for real-world image, video and audio signals. Several measures to determine objectively perceived distortion and the threshold of perception have been proposed for the mentioned media [75]. However, most of them are still not perfect enough to replace human viewers or listeners who judge the visual or audio fidelity through blind tests. Thus, in the design of watermarking systems, it is usually necessary to do some testing with volunteers. The second problem occurs in combination with post watermarking processing, which might result in an amplification of the embedded watermark and make it perceptible. An example is zooming of watermarked images, which often makes the embedded watermarks visible, or contrast enhancement, which may amplify highly frequent watermark patterns that are otherwise invisible.

4) *Watermark Recovery With or Without the Original Data*: Watermark recovery is usually more robust if the original, unwatermarked data are available. Further, availability of the original data set in the recovery process allows the detection and inversion of distortions which change the data geometry. This helps, for example, if a watermarked image has been rotated by an attacker. However, access to the original data is not possible in all cases, for example, in applications such as data monitoring or tracking. For other applications, like video watermarking, it may be impractical to use the original data because of the large data volume, even if it is available. It is, however, possible to design watermarking techniques that do not need the original for watermark extraction. Most watermarking techniques perform some kind of modulation in which the original data set is considered a distortion. If this distortion is known or can be modeled in the recovery process, explicitly designed techniques allow its suppression without knowledge of the original. In fact, most recent methods do not require the original for watermark recovery. In some publications, such techniques are called “blind” watermarking techniques [2], [1].

5) *Watermark Extraction or Verification of Presence for a Given Watermark*: In the literature, two different types of watermarking systems can be found: systems that embed

a specific information or pattern and check the existence of the (known) information later on in the watermark recovery—usually using some sort of hypothesis testing—and systems that embed arbitrary information into the host data.

The first type, verification of the presence of a known watermark, is sufficient for most copyright-protection applications.

The second type, embedding of arbitrary information, is, for example, useful for image tracking on the Internet with intelligent agents where it might not only be of interest to discover images, but also to classify them. In such cases, the embedded watermark can serve as an image identification number. Another example where arbitrary information has to be embedded are applications for video distribution where, e.g., the serial number of the receiver has to be embedded.

Although most presented methods or systems are designed for either watermark extraction or verification of presence for a given watermark, it should be noted that in fact both approaches are inherently equivalent. A scheme that allows watermark verification can be considered as a 1-bit watermark recovery scheme, which can easily be extended to any number of bits by embedding several consecutive “1-bit watermarks.” The inverse is also true: a watermark recovery scheme can be considered as a watermark verification scheme assuming the embedded information is known.

## B. Basic Watermarking Principles

The basic idea in watermarking is to add a watermark signal to the host data to be watermarked such that the watermark signal is unobtrusive and secure in the signal mixture but can partly or fully be recovered from the signal mixture later on if the correct cryptographically secure key needed for recovery is used.

To ensure imperceptibility of the modification caused by watermark embedding, a perceptibility criterion of some sort is used. This can be implicit or explicit, host data adaptive or fixed, but it is necessary. As a consequence of the required imperceptibility, the individual samples (e.g., pixels or transform coefficients) that are used for watermark embedding can only be modified by an amount relatively small to their average amplitude.

To ensure robustness despite the small allowed changes, the watermark information is usually redundantly distributed over many samples (e.g., pixels) of the host data, thus providing a “holographic” robustness, which means that the watermark can usually be recovered from a small fraction of the watermarked data, but the recovery is more robust if more of the watermarked data are available for recovery.

As said before, watermark systems do in general use one or more cryptographically secure keys to ensure security against manipulation and erasure of the watermark.

There are three main issues in the design of a watermarking system.

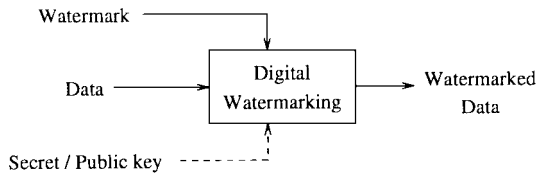


Fig. 1. Generic digital watermarking scheme.

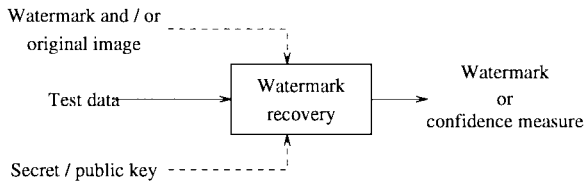


Fig. 2. Generic watermark recovery scheme.

- 1) Design of the watermark signal  $\mathbf{W}$  to be added to the host signal. Typically, the watermark signal depends on a key  $\mathcal{K}$  and watermark information  $\mathbf{I}$

$$\mathbf{W} = f_0(\mathbf{I}, \mathcal{K}). \quad (1)$$

Possibly, it may also depend on the host data  $\mathbf{X}$  into which it is embedded

$$\mathbf{W} = f_0(\mathbf{I}, \mathcal{K}, \mathbf{X}). \quad (2)$$

- 2) Design of the embedding method itself that incorporates the watermark signal  $\mathbf{W}$  into the host data  $\mathbf{X}$  yielding watermarked data  $\mathbf{Y}$

$$\mathbf{Y} = f_1(\mathbf{X}, \mathbf{W}). \quad (3)$$

- 3) Design of the corresponding extraction method that recovers the watermark information from the signal mixture using the key and with help of the original

$$\hat{\mathbf{I}} = g(\mathbf{X}, \mathbf{Y}, \mathcal{K}) \quad (4)$$

or without the original

$$\hat{\mathbf{I}} = g(\mathbf{Y}, \mathcal{K}). \quad (5)$$

The first two issues, watermark signal design and watermark signal embedding, are often regarded as one, specifically for methods where the embedded watermark is host signal adaptive.

Figs. 1 and 2 illustrate the concept. Fig. 1 shows the generic watermarking scheme for the embedding process. The input to the scheme is the watermark, the host data, and an optional public or secret key. The host data may, depending on the application, be uncompressed or compressed, however, most proposed methods work on uncompressed data. The watermark can be of any nature, such as a number, text, or an image. The secret or public key is used to enforce security. If the watermark is not to be read by unauthorized parties, a key can be used to protect the watermark. In combination with a secret or a public key, the watermarking techniques are usually referred to as secret and public watermarking techniques, respectively. The output of the watermarking scheme are the modified, i.e., watermarked,

data. The generic watermark recovery process is depicted in Fig. 2. Inputs to the scheme are the watermarked data, the secret or public key, and, depending on the method, the original data and the original watermark. The output of the watermark recovery process is either the recovered watermark or some kind of confidence measure indicating how likely it is for the given watermark at the input to be present in the data under inspection.

Many proposed watermarking schemes use ideas borrowed from spread-spectrum radio communications [25], [43], [101]. They embed a watermark by adding a pseudonoise (PN) signal with low amplitude to the host data. This specific PN signal can later on be detected using a correlation receiver or matched filter. If the parameters like amplitude and the number of samples of the added PN signal are chosen appropriately, the probabilities of false-positive or false-negative detections are very low. The PN signal has the function of a secret key. The scheme can be extended if the PN signal is either added or subtracted from the host signal. In this case, the correlation receiver will calculate either a high-positive or high-negative correlation in the detection. Thus, 1 bit of information can be conveyed. If several such watermarks are embedded consecutively, arbitrary information can be conveyed.

#### IV. TEXT DOCUMENT WATERMARKING

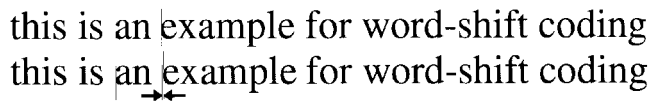
Methods for embedding information into text documents have been used for a long time by secret services.

For text watermarking, we have to distinguish between methods that hide information in the semantics, which means in the meaning and ordering of the words, and methods that hide information in the format, which means in the layout and the appearance.

The first class designs a text around the message to be hidden. In that sense, the information is not really embedded in existing information, but rather covered by misleading information. This class of techniques is outside the scope of this paper and will not be considered here. In the following, we concentrate on the latter type of information-embedding methods which use an existing text document into which data are embedded.

Formatted text is probably the medium where watermarking methods can be defeated most easily. If the watermark is in the format, then it can obviously be removed by “retyping” the whole text using a new character font and a new format where “retyping” can be either manual or automated using optical character recognition (OCR). OCR systems are still not perfect for many applications today and often need human supervision. Thus, removal of watermarks either yields bad results (single characters are wrong, due to OCR) or is expensive. The goal is to make watermark removal more expensive than obtaining the right to copy from the copyright owner. If this goal is achieved, text watermarking makes sense, though it can be defeated [14].

Text watermarking has applications wherever copyrighted electronic documents are distributed. Important examples are virtual digital libraries where users may download



this is an example for word-shift coding  
this is an example for word-shift coding

Fig. 3. Example for word-shift coding.

copies of documents, for example, books, but are not allowed to further distribute them or to store them longer than for a certain predefined period. In this type of application, a requested document is watermarked with a requester specific watermark before releasing it for download. If later on illegal copies are discovered, the embedded watermark can be used to determine the source.

Brassil *et al.* [14], [15], [84], [85], [91] have extensively worked on text watermarking. They propose three different methods for information embedding into text documents: line shift coding; word-shift coding; and feature coding. In line-shift coding, single lines of the document are shifted upwards or downwards by very small amounts. The information to be hidden is encoded in the way the lines are shifted. Similarly, words are shifted horizontally in order to modify the spaces between consecutive words in word-shift coding. An example for word-shift coding is shown in Fig. 3. Both methods are applicable to the format file of a document or to the bitmap of a page image. While line-shift coding can rely on the assumption that lines are uniformly spaced, and thus does not necessarily need the original for watermark extraction, the original is required for extraction in word-shift coding, since the spaces between words are usually variable. The third method, feature coding, slightly modifies features such as the length of the end lines in characters like *b, d, h*, etc. Among the three presented methods, line-shift coding is the most robust in the presence of noise but also most easily defeated. The authors again argue that although the described methods can theoretically be defeated, it requires interactive human intervention and is expensive in practice. The presented methods are robust enough to resist printing, consecutive photocopying up to ten generations, and rescanning [85].

## V. IMAGE WATERMARKING

Most watermarking research and publications are focused on images. The reason might be that there is a large demand for image watermarking products due to the fact that there are so many images available at no cost on the World Wide Web which need to be protected.

Meanwhile, the number of image watermarking publications is too large to give a complete survey over all proposed techniques. However, most techniques share common principles. Thus, we try to point out the common ideas first, before we explain some selected methods in more detail to illustrate how the principles are applied in practice.

The watermark signal is typically a pseudorandom signal with low amplitude, compared to the image amplitude, and usually with spatial distribution of one information (i.e., watermark) bit over many pixels. A lot of watermarking methods are in fact very similar and differ only in parts or

single aspects of the three topics: signal design; embedding; and recovery.

The information that is embedded is usually not important for the watermarking itself. However, there are methods that are designed to embed and extract one out of a codebook of codes, and thus cannot accommodate arbitrary information [27], [72]. Other proposed schemes modulate the codes available in the codebook with arbitrary information bits and can thus accommodate arbitrary messages. Although some authors distinguish strictly between the two types, they are in fact conceptually very close.

The watermark signal is often designed as a white [136], [139] or colored pseudorandom signal with, e.g., Gaussian [27], uniform, or bipolar [33], [72], [76], [93], [136], probability density function (pdf). In order to avoid visibility of the embedded watermark, an implicit or explicit spatial [7], [66], [126], [146] or spectral [66], [105], [106], [126], [130], [146] shaping is often applied with the goal to attenuate the watermark in areas of the image where it would otherwise become visible. The resulting watermark signal is sometimes sparse and leaves image pixels unchanged [33], [74], but mostly it is dense and alters all pixels of the image to be watermarked. The watermark signal is often designed in the spatial domain, but sometimes also in a transform domain like the full-image discrete cosine transform (DCT) domain [27] or block-wise DCT domain [69].

The signal embedding is done by addition [78], [93], [139] or signal-adaptive (i.e., scaled) addition [2], mostly to the luminance channel alone, but sometimes also to color channels, or only to color channels [73]. The addition can take place in the spatial domain, or in transform domains such as the discrete Fourier transform (DFT) domain [113], the full-image DCT domain [3], [27], [105], the block-wise DCT domain [7], [47], [69], [78], [106], [151], the wavelet domain [71], [72], [143], the fractal domain [34], [96], [109], the Hadamard domain [59], [111], the Fourier–Mellin domain [114], [115], or the Radon domain [150]. It is often claimed that embedding in the transform (mostly DCT or wavelet) domain is advantageous in terms of visibility and security [3]. However, while some authors argue that the watermarks should be embedded into low frequencies [27], [114], other argue that they should rather be embedded into the medium [3], [36], [56] or high frequencies. In fact, it has been shown [122], [123] that for maximum robustness watermarks should be embedded signal adaptively into the same spectral components that the host data already populate. For images and video, these are typically the low frequencies.

As said before, watermark signal generation and watermark embedding are often treated jointly. For some proposed methods, they cannot be regarded separately, especially if the watermark is signal adaptive [3], [22], [23], [78], [148].

The watermark recovery is usually done by some sort of correlation method, like a correlation receiver or a matched filter. Since the watermark signal is often designed without knowledge of the host signal, crosstalk between watermark signal and host data is a common problem in

watermarking. In order to suppress the crosstalk, many proposed schemes require the original, unwatermarked data in order to subtract it before watermark extraction. Other proposed methods apply a prefilter [38], [73], [82], [139] instead of subtracting the original. Yet other methods do not suppress the crosstalk [105]. Some researchers propose to use more sophisticated detectors than just simple correlation detectors, e.g., maximum *a-posteriori* (MAP) detectors [3]. Like for embedding, several domains have been proposed for watermark extraction, often corresponding to the domain that is proposed for embedding or for signal design. There are fewer publications where watermark embedding and extraction are proposed in different domains.

Before we look at some specific watermarking techniques in the different domains, we give a brief chronological overview of early watermarking methods.

The year 1993 can be considered the beginning of the digital image watermarking era, although other publications from the early 1990's, such as Tanaka *et al.* [131], [132], already introduced the idea of tagging images to secretly hide information and assure ownership rights. Caronni [20], [21] describes an overall system to track unauthorized image distribution. He proposes to mark images using spatial signal modulation and calls the process tagging. A tag is a square of size  $N * N$ . In a first step, all possible locations in an image where a tag could possibly be placed are identified by calculating the local region variance of size  $N * N$  in the image and comparing it to empirically identified upper and lower limits. Only locations with minimal variance are used for tagging. A tag is a square with a constant value proportional to the maximum image brightness within the square and decaying outside the border. A selected image area is tagged by adding or subtracting the tag and a random, zero mean, noise pattern. Both the tag location and the noise sequence are key dependent. One selected tag location hides 1 bit and is only tagged if the bit to embed is set to one. To recover an embedded bit, the difference between the original and the tagged image is computed. Then the mean of a supposedly tagged location is compared to the neighboring mean to determine the bit value. In addition to the marking process, Caronni also suggests to use the correlation coefficient between the original and the tagged image as a measure for the image degradation due to the tagging process. A correlation coefficient of one indicates that the two images are identical, whereas for distorted images the value decreases toward zero.

In the same year, approaches and ideas for digital image watermarking were proposed by Tirkel *et al.* [136] in their 1993 publication entitled *Electronic Water Mark*. In this early publication on digital watermarking, the authors already recognized the importance of digital watermarking and proposed possible applications for image tagging, copyright enforcement, counterfeit protection, and controlled access to image data. Two methods were proposed for grayscale images. In the first approach, the watermark in form of an  $m$ -sequence-derived PN code is embedded in the least significant bit (LSB) plane of the image data. To

**Table 1**  
Sample Cipher Key Table

$\Delta_i$	...	-4	-3	-2	-1	0	1	2	3	4	...
$c_i$	...	0	0	1	1	0	1	0	0	1	...

gain full access to the LSB plane without introducing much distortion, the image is first compressed to 7 bits through adaptive histogram manipulation. This method is actually an extension to simple LSB coding schemes in which the LSB's are replaced by the coding information. The watermark decoding is straightforward since the LSB plane carries the watermark without any distortion. In the second approach, the watermark, again in form of an  $m$ -sequence-derived code, is added to the LSB plane. The decoding process makes use of the unique and optimal autocorrelation function of  $m$ -sequences [86]. A modified version of the paper was published in 1994 [139] titled *A Digital Watermark*, and being the first publication explicitly mentioning, and hence defining, the term digital watermarking. In 1995 [137], the idea of using  $m$ -sequences and LSB addition was extended and improved by the authors through the use of two-dimensional (2-D)  $m$ -sequences which resulted in more robust watermarks.

About the same time Matsui and Tanaka [90] published a paper called "Video Steganography: How to Secretly Embed a Signature in a Picture," in which several watermarking techniques were proposed for image watermarking. Their first method is based on a predictive coding scheme for gray scale images. Predictive coding schemes exploit the correlation between adjacent pixels by coding the prediction error instead of coding the individual gray scale values. A digital image is scanned in a predefined order traversing the pixels  $\{x_i\}; i \in N$ . The set of pixels is then coded using a predictive coding scheme by keeping the first value  $x_1$  and replacing subsequent values  $x_i$  by the difference  $e_i$  between adjacent pixels

$$e_i = x_i - x_{i-1}. \quad (6)$$

To embed a watermark in form of a binary string, Matsui and Tanaka introduce a cipher key table which assigns a corresponding bit  $c_i$  to all possible differences  $\Delta_i$ . An example of such a table is given in Table 1. The correspondence between bit values and the differences is kept secret. To embed a bit  $b$ , select a pixel  $x_i$  with its corresponding difference  $e_i$ . Check in the cipher table if the bit value  $c_i$  corresponding to  $\Delta_i = e_i$  has the same value as bit  $b$ . If this is the case, proceed to the next bit, otherwise select the closest value to  $e_i$  in the cipher table that has the appropriate bit value. The watermark can be recovered by looking up the bit in the coding table. The second method modifies the ordered dithering scheme for binary pictures. A dithering scheme consists of comparing the monotone level of pixels within a pixel block with a position-dependent threshold and turning "on" those pixels with a value above the threshold. The location dependent thresholds are given in a square matrix of size  $N * N$  called dither matrix with entries  $d_{pq}^{(n)}$ , where  $n$  denotes an ordering number between zero and  $N^2 - 1$  and  $p$  and  $q$  the

6	7	8	9
5	0	1	10
4	3	2	11
15	14	13	12

Fig. 4. Sample dither matrix: dot-concentrated type.

corresponding matrix line and column, respectively. Fig. 4 shows a sample dithering matrix. Given the dither matrix, the corresponding thresholds  $T$  are defined as

$$T = \left( d_{pq}^{(n)} + \frac{1}{2} \right) \times \frac{R}{N^2} \quad (7)$$

where  $R$  defines the dynamic brightness range of the image. To dither an image, it is first divided into adjacent blocks of the same size as the dither matrix. Then all values in each block are compared to the corresponding threshold value and modified accordingly. Now let the set of threshold pairs be defined as

$$S_k = \{(x_i, x_j)_k | x_i - x_j = k; i, j = 0, 1, \dots, N; i \neq j\} \quad (8)$$

where  $x_{i,j}$  denote thresholds in the dither matrix. Further, let  $(y_i, y_j)_k$  be the output signal of  $x_i, x_j$  and assuming the values of  $(0, 0)_k, (0, 1)_k, (1, 0)_k$ , and  $(1, 1)_k$ . Only the two pairs  $(0, 1)$  and  $(1, 0)$  are considered for data embedding.

To embed a bit  $b$ , an output pair  $(y_i, y_j)_k$  is selected, and  $y_i$  is compared with the bit value  $b$ . If the values are equal, the pair is left unchanged, otherwise  $y_i$  and  $y_j$  are swapped. In order to decode an embedded signature, the above described procedure is inverted. Again, the pairs  $(0, 0)_k$  and  $(1, 1)_k$  are disregarded. The third scheme is proposed to watermark facsimile documents. Facsimile documents are scanned with a horizontal resolution of about 8.23 pixels/mm and then compressed using run length encoding (RLE) followed by modified Huffman coding (MH). The embedding process modifies the run lengths between two subsequent, changing pels. If a one is to be embedded, the run length is forced to be even, whereas for a zero the run length is forced to be odd. For valid embedding, the original run length has to be larger than one. Decoding an embedded bit is achieved by looking at the decoded run length. Their last method is based on the modification of DCT coefficients in a progressive transmission scheme. The watermark bits are embedded by modifying the rounding rule for the quantized coefficients such that the resulting coefficients are odd or even, depending on the watermark bits.

It was soon recognized that digital watermarking and digital modulation, and especially direct sequence spread-spectrum modulation [40], [102], [119], [140], share similar concepts, and it was proposed to consider digital watermarking as communication in non-Gaussian noise. First theoretical approaches were proposed by Smith [120].

A more in depth analysis of 2-D multipulse amplitude modulation was given by Hernandez *et al.* [53].

Since the above-mentioned first publications, the interest and research activities on watermarking have largely increased. Multimedia content providers and distributors are especially interested in working solutions. In the following, we present some of the more recent work and start the overview with methods working in the spatial domain.

Bender *et al.* [6] propose two methods for data hiding. In the first method, called "Patchwork," randomly selected pairs of pixels  $(a_i, b_i)$  are used to hide 1 bit by increasing the  $a_i$ 's by one and decreasing the  $b_i$ 's by one. Provided that the image satisfies some statistical properties, the expected value of the sum of the differences between the  $a_i$ 's and  $b_i$ 's of  $N$  pixel pairs is given by  $2N$

$$\sum_N a_i - b_i = \begin{cases} 2N, & \text{for watermarked pairs} \\ 0, & \text{for nonwatermarked pairs.} \end{cases} \quad (9)$$

In the second approach, called "Texture Block Coding," the watermark is embedded by copying one image texture block to another area in the image with a similar texture. To recover the watermark, the autocorrelation function has to be computed. A remarkable feature of this technique is the high robustness to any kind of distortion, since both image areas are distorted in a similar way, which means that the watermark recovery by autocorrelation still works.

Pitas and Kaskalis propose signature casting on digital images [93], [103], [104], which is based on the same basic idea as the patchwork algorithm proposed by Bender *et al.* [6]. The watermark  $S = \{s_{m,n}\}$  consists of a binary pattern of the same size as the original image and where the number of "ones" is equal to the number of "zeros." The original image  $I$ , with luminance values  $x_{m,n}$  at location  $m$  and  $n$ , is divided into two sets  $A$  and  $B$  of equal size in the following way:

$$\begin{aligned} A &= \{x_{mn} \in I, s_{mn} = 1\} \\ B &= \{x_{mn} \in I, s_{mn} = 0\}. \end{aligned} \quad (10)$$

The watermark is superimposed by changing the elements of the subset  $A$  by a positive integer factor  $k$ , e.g.,  $A' = \{x_{mn} + k, x_{mn} \in A\}$ . The watermarked image is then given by the union of  $A'$  and  $B$ . To verify the presence of a watermark, hypothesis testing [97] is applied. The test statistic  $q$  is defined as the normalized difference between the mean  $\bar{a}'$  of set  $A'$  and the mean  $\bar{b}$  of set  $B$

$$q = \frac{\bar{b} - \bar{a}'}{\sigma_{A'}^2 + \sigma_B^2} \quad (11)$$

where  $\sigma_{A'}^2$  and  $\sigma_B^2$  defines the sample variance of set  $A'$  and  $B$ , respectively. The test statistic is then compared with a threshold to determine if there is a watermark. The method is immune to subsampling followed by up-sampling and resists to JPEG compression with a compression factor of 1:4.

An improved version of this idea has been proposed Langelaar *et al.* [78], [82]. The image is tiled into square blocks with a size being a multiple of eight. A single bit is embedded by iteratively modifying a pseudorandomly



selected block. Each selected block has a pseudorandom pattern  $P$ , with equal number of “1” and “0” assigned to it. To embed a bit with a value of “1,” the scaled pattern  $k \times P$ , where  $k$  is a predefined scaling factor defining the initial minimal watermark strength, is added to the block. For a bit with a value of “0,” the scaled pattern is subtracted from the block. Let  $I_0$  be the mean of all pixel values within the block for which the corresponding pattern value is zero, and  $I_1$  the mean of the remaining pixels. Further, let  $D_{\text{high}} = I_1 - I_0$  be the difference between the two means, and  $D_{\text{low}} = \hat{I}_1 - \hat{I}_0$  be the difference between the means after JPEG compression of the block with a predefined quality factor  $Q$ . If a “0” is to be embedded, the pattern  $P$  is iteratively subtracted from the block until both differences,  $D_{\text{high}}$  and  $D_{\text{low}}$  are below zero or the maximum number of iterations has been reached. If a “1” is to be embedded the pattern is iteratively added to the block until both differences,  $D_{\text{high}}$  and  $D_{\text{low}}$ , are above a predefined threshold  $T$  or the maximum number of iterations has been reached. An embedded bit can be extracted by again computing the difference  $D_{\text{high}}$  between the two means  $I_1$  and  $I_0$ . The sign of this difference is then used to determine the embedded bit value. Tests with the parameters set to block size  $32 \times 32$ , threshold  $T = 1$ , initial scaling factor  $k = 4$  and maximum number of iterations six, indicate that the method features decent robustness toward JPEG compression with a bit error rate of about 5% for 85% JPEG quality and 20% for 60% JPEG quality. In a second method the authors propose watermarking in the DCT domain by setting DCT-coefficients below a selected scan line to zero.

To increase the performance of the block base spatial watermarking methods, Bruyndonckx *et al.* [17] suggest the used of pixel classification. Pixels within pseudorandomly selected blocks are classified into zones (1 and 2) of homogeneous luminance values. The classification is based on three types of contrast between zones: hard contrast; progressive contrast; and noise contrast. Each zone is then further subdivided into two categories  $A$  and  $B$  based on a grid defined by the coder. Each pixel is thus assigned to one of four zone/category combinations, e.g.,  $1/A, 1/B, 2/A$ , and  $2/B$ . A bit  $b$  is embedded by modifying the zone/category means to satisfy the following constraints:

$$\begin{aligned} \text{if } b = 0: \quad & m_{1B}^* - m_{1A}^* = S \\ & m_{2B}^* - m_{2A}^* = S \\ \text{if } b = 1: \quad & m_{1A}^* - m_{1B}^* = S \\ & m_{2A}^* - m_{2B}^* = S \end{aligned} \quad (12)$$

where  $m_{1A}^*, m_{1B}^*, m_{2A}^*$ , and  $m_{2B}^*$  are the modified zone/category mean values and  $S$  the watermark embedding strength. The modification of the mean values is done by applying equal luminance variations for all pixels belonging to the same zone. To increase robustness the authors suggest to perform redundant bit embedding and use error-correcting codes. Good robustness to JPEG compression is reported.

In order to increase the performance of spread-spectrum watermarking in the spatial domain Kutter *et al.* [73], [74] propose a method which exclusively works with the blue image component, in the RGB color space, in order to maximize the watermark strength while keeping visual artifacts minimal. Further, they propose to preprocess the image prior to watermark decoding in order to predict the embedded watermark. This concept improves the robustness significantly and is applicable to any spread-spectrum watermarking in the spatial domain. The method embeds a watermark in form of a binary number through amplitude modulation in the spatial domain. A single bit  $b$  is embedded at a pseudorandomly selected location  $(i, j)$  by either adding or subtracting, depending on the bit, a value which is proportional to the luminance at the same location

$$B_{i,j} \leftarrow B_{i,j} + \alpha(-1)^b L_{i,j} \quad (13)$$

where  $B_{i,j}$  describes the blue value at location  $(i, j)$ ,  $L_{i,j}$ , the luminance at the same location, and  $\alpha$ , the embedding strength. To recover an embedded bit, an estimate of the original, nonwatermarked, value is computed using linear combination of neighboring pixels in a cross shape

$$\hat{B}_{i,j} = \frac{1}{4c} \left( \sum_{k=-c}^c B_{i+k,j} + \sum_{k=-c}^c B_{i,j+k} - 2B_{j,k} \right) \quad (14)$$

where  $c$  defines the size of the cross-shaped neighborhood. The bit value is determined by looking at the sign of the difference  $\delta_{i,j}$  between the pixel under inspection and the estimated original. In order to increase robustness, each signature bit is embedded several times, and to extract the embedded bit the sign of the sum of all differences  $\delta_{i,j}$  is used. Fig. 5 illustrates an image composition example. The two watermarked images on the top are used to generate the new composite image on the bottom. Given the appropriate keys, both original watermarks can be recovered. Extensions to this method allow increased robustness and even watermark recovery after geometrical attacks [76] and printing–scanning.

Maq *et al.* [37], [87] introduce watermarking adapted to the human visual system (HVS) using masking and modulation. In their scheme, the watermark in form of a spatially limited binary pattern is low-pass filtered, frequency modulated, masked, and then added to the host image. A secret key is used to determine the modulation frequencies and the watermark embedding location. The masking process uses an extension of the masking phenomena for monochromatic signals, also called gratings. To further adapt the watermark to the image, a shaping mask, based on morphological homogenized areas of high frequencies, is used. Watermark recovery is performed by demodulation followed by a correlation function.

In a very different approach, Voyatzis and Pitas watermark images by inserting logo like patterns using torus automorphisms [141], [142]. A 2-D torus automorphism can be considered as a spatial transformation of planar regions which belong to a square 2-D area. It is defined in the subset

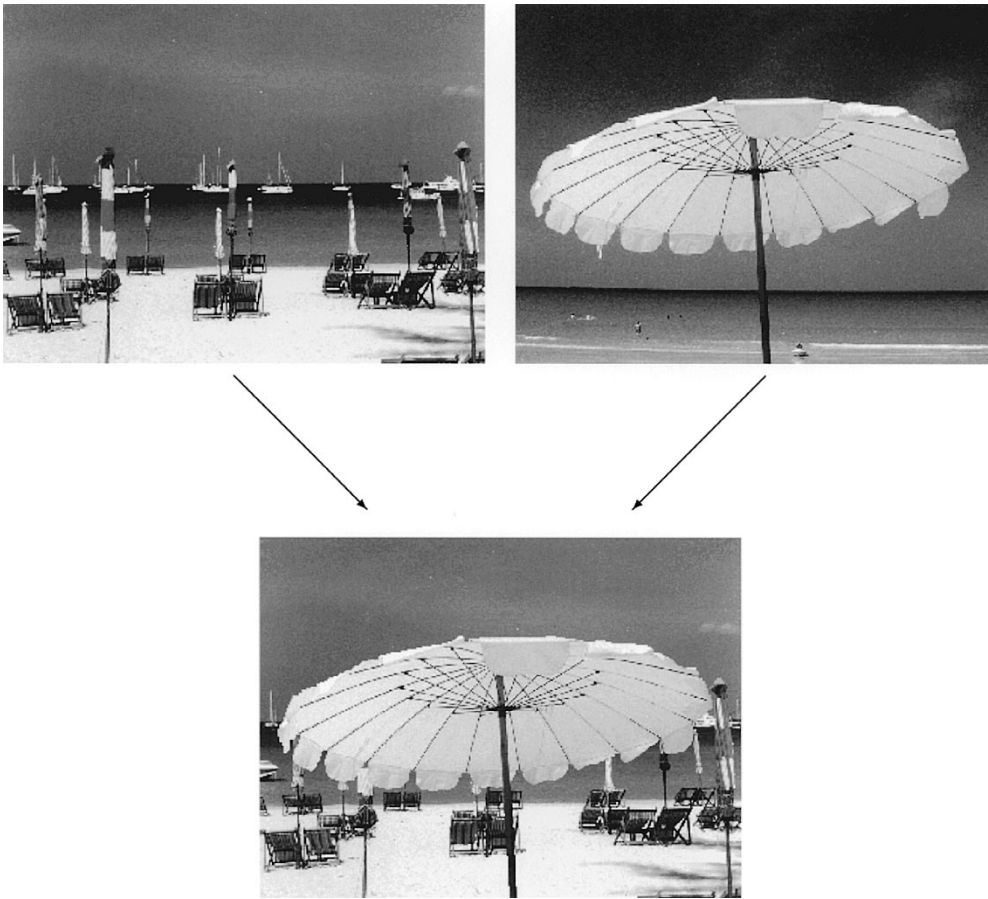


Fig. 5. Image composition. The umbrella of the “umbrella” image is pasted onto the “beach” image. The watermarks from both images can be recovered from the composed image.

$U = [0, 1) \times [0, 1) \subset R^2$  by

$$\mathbf{r}' = \mathbf{A}\mathbf{r}, \quad \begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \pmod{1}. \quad (15)$$

Iterated actions of  $\mathbf{A}$  on a point  $\mathbf{r}_0$  form a dynamical system which can be expressed like a map

$$\mathbf{r}_{n+1} = \mathbf{A}^n \mathbf{r}_0 \pmod{1} \quad \text{or} \quad \mathbf{r}_{n+1} = \mathbf{A}\mathbf{r}_n \quad (16)$$

An example for a well-known automorphism in dynamics is the “cat map,” defined as

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \pmod{1}. \quad (17)$$

The set of points  $\{\mathbf{r}_0, \mathbf{r}_1, \mathbf{r}_2, \dots\}$  is called an orbit of the system. Roughly speaking, such a system mixes the points in a chaotic way. Under certain circumstances, the automorphisms may have periodic orbits, which means that after  $T$  iterations the current point is equal to the initial point, e.g.,  $\mathbf{A}^T \mathbf{r}_0 = \mathbf{r}_0$ . Fig. 6 shows an example of an automorphism using the cat map.

To sign an image, a watermark in the form of a square binary image, with a size smaller than the original image, is first mixed using the automorphism  $\mathbf{A}_N$ . The resulting, mixed watermark is then overlaid on a selected block in the original image using an embedding function such as LSB modification. Watermark recovery is performed

by first extracting the mixed watermark from the signed image followed by reconstructing the watermark using the automorphism  $\mathbf{A}_{N-T}$ , where  $T$  is the automorphism period for the given system. Using more sophisticated overlaying methods will increase the robustness of the method.

Raymond and Wolfgang [147], [148] propose a watermarking technique to verify image authenticity based on an approach similar to the  $m$ -sequence approach suggested by Schyndel *et al.* for the one-dimensional case [139] and Tirkel *et al.* for the two dimensional case [137]. A random sequence generated by using linear feedback shift registers is mapped from  $\{0,1\}$  to  $\{1, -1\}$ , arranged into a suitable block and added to the image. To locate where an image has been forged, the algorithm overlays the watermarked image block with the watermark block, computes the inner product, and compares the result to the ideal value. Let the cross-correlation function  $R_{XY}(\alpha, b\beta)$  of two blocks  $X$  and  $Y$  be defined as

$$R_{XY}(\alpha, b\beta) = \sum_i \sum_j X(i, j) Y(i - \alpha, j - \beta) \quad (18)$$

then the test statistic  $\delta$  for a block, given the original image block  $X$ , the watermark block  $W$ , the watermarked image block  $Y$ , and the probably forged image block  $Z$ , is defined as

$$\delta = R_{YW}(0, 0) - R_{ZW}(0, 0). \quad (19)$$



**Fig. 6.** Example automorphism with the “cat map.” (a) is the original image. (b)–(d) show the automorphism of (a) after one, two, and ten iterations, respectively.

If the watermark is unchanged,  $\delta = 0$ . When  $\delta$  is greater than a defined tolerance, the block fails the watermark test. The method detects any kind of image filtering, and the authors claim that an improved version can even accommodate JPEG compression.

Watermark embedding not based on spread-spectrum modulation but quantization has been proposed by Chen and Wornell [24]. Their method is called quantized index modulation (QIM) and is based on a set of  $N$ -dimensional quantizers. The quantizers satisfy a distortion constraint and are designed such that the reconstruction values from one quantizer are “far away” from the reconstruction points of every other quantizer. The message to be transmitted is used as in index for quantizer selection. The selected quantizer is then used to embed the information by quantizing the image data in either the spatial or DCT domain. In the decoding process, a distance metric is evaluated for all quantizers and the index of the quantizer with the smallest distance identifies the embedded information. The authors show that the performance of the resulting watermarking scheme is superior to standard spread-spectrum modulation without watermark weighting.

Besides spatial domain watermarking related to modulation it was proposed by Maes *et al.* [89] to modify geometric features of the image. The method is based on a dense line pattern, generated pseudorandomly and representing the watermark. A set of salient points in the

image is then computed, for example, based on an edge detection filter. The detected points are then warped such that a significantly large number of points are within the vicinity of lines. In the detection process, the method verifies if a significantly large number of points are within the vicinity of lines.

Related to spatial domain watermarking schemes are methods based on fractal image compression. The idea to use this approach has first been proposed by [109]. In fractal image compression the image is coded using the principles of iterated function systems and self similarity [116]. The original image is divided into square blocks  $R_k$  called range blocks. Further, let  $F$  be a set of mapping functions  $f_k$ , which are a composed of a geometric transformation  $g_k$  and a massic transformation  $m_k$ . The mapping functions work on domain blocks  $D_k$ , which are larger than range blocks. The geometric transformation consists of moving the domain block  $D_k$  to the location of the range block  $R_k$  and reducing the size of the domain block to the size of the range block. The massic transformation adjusts the intensity and orientation of pixels in the domain block after geometric transformation. Massic transformations include rotation by  $90$ ,  $180$ , and  $-90^\circ$ , reflection at midhorizontal and cross-diagonal axis, as well as identity mapping. To compress an image for all range blocks  $R_k$ , the best combination of domain block  $D_k$  and mapping function  $f_k$  has to be found such that the difference between the range block  $R_k$  and

the mapped domain block  $f_k(D_k)$  is minimal. That means that the encoding includes a spatial search over all possible domain blocks. Decoding is accomplished by iterating over the coded mapping functions using any initial image. To embed a bit into this scheme a range block is pseudorandomly selected. The corresponding search space  $S_k$  for the range blocks is then split up into two subsearch spaces  $S_k^1$  and  $S_k^2$  of equal size. Each subspace is assigned to a bit value, and the current range block is encoded by searching only in the subspace corresponding to the bit value of the current bit. To recover an embedded bit, the image is again compressed, however this time using the full domain block search space. Then for a marked range block the location of the corresponding domain block reveals the embedded bit value. The algorithm was tested against JPEG compression and showed good robustness down to a compression quality of about 50%. A drawback of this technique is the slow speed due to the fractal compression scheme.

A very similar approach has been proposed by Davern and Scott [34]. The only difference is that they do not encode the entire image, but only a user-defined range region based on a user-defined domain region. Given the two regions, the watermark encoding is equivalent to the system proposed by Puate and Jordan in that the domain region is divided into two parts and, depending on the bit value, one or the other region is used for encoding a range block. This idea of watermarking using spatial domain fractal image coding has been extended to DCT blocks by Bas *et al.* [4].

Efficient watermarking in the DCT domain was first introduced by Koch *et al.* [18], [68], [69]. As in the JPEG compression scheme, the image is first divided into square blocks of size  $8 \times 8$  for which the DCT is computed. From a pseudorandomly selected block, a pair of midfrequency coefficients is selected from 12 predetermined pairs. To embed a bit, the coefficients are then modified such that the difference between them is either positive or negative, depending on the bit value. In order to accommodate lossy JPEG compression, the quantization matrix is taken into account when altering the DCT coefficients. This method shows good robustness to JPEG compression down to a quality factor of 50%.

Bors and Pitas [12], [13] suggest a method that modifies DCT coefficients satisfying a block site selection constraint. The image is first divided into blocks of size  $8 \times 8$ . Certain blocks are then selected according to a Gaussian network classifier decision. The middle range frequency DCT coefficients are then modified, using either a linear DCT constraint or a circular DCT detection region, to convey the watermark information. In the first approach, the linear constraint is defined as

$$Y = FQ \quad (20)$$

where  $F$  is the modified DCT coefficient vector and  $Q$  the weighting vector provided by the watermark. The constraint is imposed by changing the DCT coefficients based on a least-squares criterion. The second algorithm defines circular regions around the selected DCT frequency

coefficients. The selected frequencies are then quantized according to

$$\|F - Q_k\|^2 = \min_{i=1}^H \|F - Q_i\|^2 \quad \text{then} \quad F = Q_k \quad (21)$$

where  $Q_i, 1 = 1, \dots, H$  is the set of coefficient vectors provided by the watermark. In the watermark recovery process, the algorithm first verifies the DCT coefficient constraint for all blocks followed by the location constraint. The algorithm can accommodate JPEG compression for a compression ratio of 13:1 and 18:1 using the linear DCT constraint or the circular DCT detection region, respectively.

Swanson *et al.* [129], [130] suggest a DCT domain watermarking technique, based on frequency masking of DCT blocks, which is similar to methods proposed by Smith and Comiskey [120]. The input image is split up into square blocks for which the DCT is computed. For each DCT block, a frequency mask is computed based on the knowledge that a masking grating raises the visual threshold for signal gratings around the masking frequency. The resulting perceptual mask is scaled and multiplied by the DCT of a maximal length PN sequence. This watermark is then added to the corresponding DCT block followed by spatial masking to verify that the watermark is invisible and to control the scaling factor. Watermark detection requires the original image as well as the original watermark and is accomplished by hypothesis testing. The authors report good watermark robustness for JPEG compression, colored noise, and cropping.

Tao and Dickinson [133] introduce an adaptive DCT-domain watermarking technique based on a regional perceptual classifier with assigned sensitivity indexes. The watermark is embedded in  $N$  AC DCT coefficients. The coefficients are selected as to have the smallest quantization step sizes according to the default JPEG compression table. The selected coefficients  $x_i$  are modified as follows:

$$\hat{x}_i = x_i + \max \left[ x_i \alpha_m, \text{sign}(x_i) \frac{D_i}{\kappa} \right] \quad (22)$$

where  $\alpha_m$  defines the noise sensitivity index for the current block,  $D_i$  the quantization step for  $X_i$ , and  $\kappa$  satisfies  $5 \leq \kappa \leq 6$ . It should be noted that the watermark signal is not generated randomly. Various approaches exist to determine the noise sensitivity by efficiently exploiting the masking effects of the HVS. The authors propose a regional classification algorithm which classifies the block in one of six perceptual classes. The classification algorithm exploits luminance masking, edge masking, and texture masking effects of the HVS. Namely the perceptual block classes from one to six are defined as: edge; uniform; low sensitivity; moderately busy; busy; and very busy, in descending order of noise sensitivity. Each perceptual class has a noise-sensitivity index assigned to it. Watermark recovery requires the original image as well as the watermark and is based on hypothesis testing. Experiments show that the method resists JPEG compression down to a quality

of 5% and can accommodate random noise with a peak signal-to-noise ratio (PSNR) of 22.1 dB.

Podilchuk [107], [108] introduces perceptual watermarking using the just noticeable difference (JND) to determine an image-dependent watermark modulation mask. The watermark modulation onto selected coefficients in either the DCT or wavelet transform domain can be described as

$$I_{u,v}^* = \begin{cases} I_{u,v} + \text{JND}_{u,v} \times w_{u,v}, & \text{if } I_{u,v} > \text{JND}_{u,v} \\ I_{u,v}, & \text{otherwise} \end{cases} \quad (23)$$

where  $I_{u,v}$  are the transform coefficients of the original image,  $w_{u,v}$  are the watermark values, and  $\text{JND}_{u,v}$  is the computed JND based on visual models. For DCT coefficients, the author suggest using a perceptual model defined by Watson based on utilizing frequency and brightness sensitivity as well as local contrast masking. This model provides image-dependent masking thresholds for each  $8 \times 8$  DCT block. Watermark detection is based on the correlation between the difference of the original image and the image under inspection and the watermark sequence. The maximum correlation is compared to a threshold to determine whether an image contains the watermark in question. Experiments showed that the watermark scheme is extremely robust to JPEG compression, cropping, scaling, additive noise, gamma correction, and printing-xeroxing-scanning. For attacks involving a geometrical transformation, the inverse operation has to be applied to the image before the watermark-detection process.

Piva *et al.* describe another DCT-based method which exploits the masking characteristics of the HVS [105]. The watermark consists of a pseudorandom sequence of  $M$  real numbers with normal distribution  $\mathbf{X} = \{x_1, x_2, \dots, x_M\}$ . The coefficients of the  $N \times N$  DCT of the original image  $I$  are reordered into a vector using a zig-zag scan. From this vector,  $M$  coefficients, starting at position  $L + 1$ , are selected to generate the vector  $\mathbf{T} = \{t_1, t_s, \dots, t_M\}$ . The watermark  $\mathbf{X}$  is embedded into  $\mathbf{T}$  according to

$$t'_i = t_i + \alpha |t_i| x_i \quad (24)$$

where  $\alpha$  determines the watermark strength. The modified coefficients replace the nonmodified coefficients before the watermarked image  $I'$  is reconstructed. In order to enhance the robustness visual masking is applied as follows:

$$y''_{ij} = y_{ij}(1 - \beta_{ij}) + \beta_{ij} y'_{ij} = y_{ij} + \beta_{ij}(y'_{ij} - y_{ij}) \quad (25)$$

where  $\beta_{ij}$  is a weighting factor taking into account the characteristics of the HVS. A simple way of choosing  $\beta_{ij}$  is the normalized sample variance at pixel  $y_{ij}$  defined as the ratio between the sample variance for a square block with center at  $y_{ij}$  and the maximum of all block variances. As in most schemes, watermark detection is performed by comparing the correlation  $z$  between the watermark and the possibly corrupted signed DCT coefficients  $\mathbf{T}^*$  with a threshold  $S_z$ . The correlation  $z$  is defined as

$$z = \frac{\mathbf{X} \cdot \mathbf{T}^*}{M} = \frac{1}{M} \sum_{i=1}^M x_i t_i^* \quad (26)$$

The threshold  $S_z$  is adaptive and given as

$$S_z = \frac{\alpha}{3M} \sum_{i=1}^M |t_i^*|. \quad (27)$$

Experimental results demonstrate that the watermark is robust to several image processing techniques (for example, JPEG compression, median filtering, and multiple watermarking) and geometrical distortions (after applying the inverse geometric transformation).

Frequency-domain watermarking was first introduced by Boland *et al.* [8] and Cox *et al.* [27], who independently developed perceptually adaptive methods based on modulation. Cox *et al.* draw parallels between their technology and spread-spectrum communication since the watermark is spread over a set of visually important frequency components. The watermark consists of a sequence of numbers  $\mathbf{x} = x_1, \dots, x_n$  with a given statistical distribution, such as a normal distribution  $N(0, 1)$  with zero mean and a variance of one. The watermark is inserted into the image  $V$  to produce the watermarked image  $V'$ . Three techniques are proposed for watermark insertion

$$v'_i = v_i + \alpha x_i \quad (28)$$

$$v'_i = v_i(1 + \alpha x_i) \quad (29)$$

$$v'_i = v_i e^{\alpha x_i} \quad (30)$$

where  $\alpha$  determines the watermark strength and the  $v_i$ 's are perceptually significant spectral components. Equation (28) is only suitable if the values  $v_i$  do not vary too much. Equations (29) and (30) give similar results for small values of  $\alpha x_i$ , and for positive  $v_i$ 's (30) may even be viewed as an application of (28) where the logarithms of the original values are used. In most cases (29) is used. The scheme can be generalized by introducing multiple scaling parameters  $\alpha_i$  as to adapt to the different spectral components and thus reduce visual artifacts. To verify the presence of the watermark, the similarity between the recovered watermark  $\mathbf{X}^*$ , given by the difference between the original image  $V$  and the possibly tampered image  $V^*$ , and the original watermark  $\mathbf{X}$  is measured. The similarity measure is given by the normalized correlation coefficient

$$\text{sim}(\mathbf{X}, \mathbf{X}^*) = \frac{\mathbf{X}^* \cdot \mathbf{X}}{\sqrt{\mathbf{X}^* \cdot \mathbf{X}^*}} \quad (31)$$

Robustness tests showed that the method resists JPEG compression (at a quality factor of 5% and no smoothing), dithering, fax transmission, printing-photocopying-scanning, multiple watermarking, and collusion attacks. For the experiments, the watermark was of length 1000 with  $N(0, 1)$  [where  $N(\mu, \sigma)$  represents a normal distribution with mean  $\mu$  and variance  $\sigma$ ],  $\alpha$  was set to 0.1, and the watermark was embedded into the 1000 strongest DCT coefficients using (29). Boland *et al.* propose a similar technique based on a hybrid between amplitude modulation and frequency shift keying and suggest the use of different transform domains such as DCT, wavelet transform, Walsh-Hadamard transform, and the fast Fourier transform (FFT).

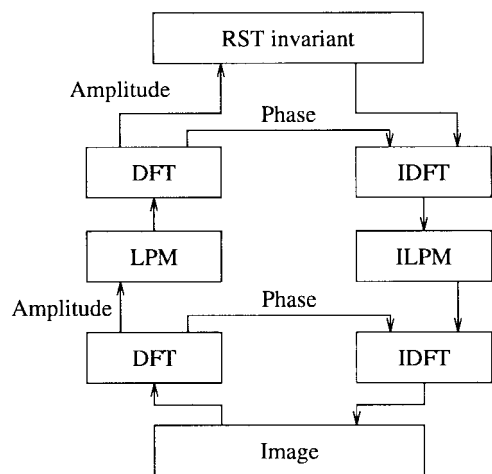


Fig. 7. RST invariant watermarking scheme.

Ruanaidh *et al.* propose watermarking by modification of the phase in the frequency domain [112], [113]. To embed a bit the phase of a selected coefficient  $F(k_1, k_2)$  of an  $N_1$  by  $N_2$ , DFT is modified by adding a small  $\delta$

$$\angle F(k_1, k_2) \leftarrow \angle F(k_1, k_2) + \delta \quad (32)$$

in order for the watermarked image to be real, the phase must satisfy negative symmetry, which leads to the additional modification

$$\angle F(N_1 - k_1, N_2 - k_2) \leftarrow \angle F(N_1 - k_1, N_2 - k_2) + \delta. \quad (33)$$

Coefficients are only modified if their relative power is above a given threshold. If the original image is available, the watermark can easily be recovered by comparing the phase. In case the original is not available, Ruanaidh suggests prequantizing the original phase prior to modifying it. Then deviations between the quantized states could be used to convey the data.

In another publication, Ruanaidh *et al.* explicitly design a watermarking technique invariant to translation, rotation, and scaling [114]. The method is a hybrid between DFT and log-polar mapping. The process is depicted in Fig. 7. In a first step, the DFT of the image is computed. One of the DFT properties is that a shift in the spatial domain results in a phase shift in the frequency domain. Keeping only the amplitude for further processing makes the image translation invariant. In the second step, rotation and scale invariance is achieved by mapping the amplitude from the Cartesian grid to a log-polar grid. Consider a point  $(x, y) \in \mathbb{R}$ , then the mapping is defined as

$$\begin{aligned} x &= e^\mu \cos \theta \\ y &= e^\mu \sin \theta \end{aligned} \quad (34)$$

where  $\mu \in \mathbb{R}$  and  $0 < \theta < 2\pi$ . One can easily see that this is a one-to-one mapping and that rotation and scaling in the Cartesian grid are converted to a translation of the  $\mu$  and  $\theta$  coordinates, respectively. Computing again the DFT of the log-polar map and keeping only the amplitude results in a rotation and translation invariant. Taking the

Fourier transform of a log-polar map is equivalent to computing the Fourier–Mellin transform. Hence combining the two steps results in a rotation, scale, and translation (RST) transformation invariant. The watermark takes the form of a two dimensional spread-spectrum signal in the RST transformation invariant domain. In a test, a 104-bit watermark was embedded into an image. The watermarked image was then rotated by  $143^\circ$  and scaled by 75% along each axis. The embedded watermark was recovered from this image. Further, the method resists JPEG compression down to a quality factor of 75% and cropping to 50% of the original image size. This approach, which is actually the first one which was especially designed as to resist to geometrical attacks, has interesting aspects and ideas and might trigger a new way of approaching the design of future watermarking techniques. A variation of this idea based on the Radon transform has been proposed by Wu *et al.* [150].

Embedding the watermark using a multiresolution decomposition has first been proposed by Boland *et al.* [8]. As for schemes working in other transformation domains, the watermark is usually given by a pseudorandom 2-D pattern. Both the image and watermark are decomposed using a 2-D wavelet transform, and in each subband of the image a weighted version of the watermark is added. Watermark decoding is, as usual, based on a normalized correlation between an estimate of the embedded watermark and the watermark itself. Various wavelet based schemes have been proposed [58], [71], [151], [152]. The difference between the schemes usually lies in the way the watermark is weighted in order to decrease visual artifact.

In this section we have presented several different watermarking methods. It can be recognized that most watermarking methods are based on the same basic principle: small, pseudorandom changes are applied to selected coefficients in the spatial or transform domain. These changes are later on identified by correlation or correlation-like similarity measures. Usually, the number of modified coefficients is much larger than the number of information bits to be encoded. This can be considered as redundant embedding and leads to implicit robustness. As we have seen, the watermark embedding domain may have a substantial influence on the watermark robustness. Spatial domain watermarking schemes are in general less robust toward noise like attacks, for example, due to lossy JPEG compression. However, a big advantage is the fact that the watermark may easily be recovered if the image has been cropped or translated. This is less obvious if the frequency domain is used. Cropping in the spatial domain results in a substantially large distortion in the frequency domain, which usually destroys the embedded watermark. The same is true for the full-frame DCT domain. If DCT blocks are watermarked, it is important to know the block position for successful watermark decoding. The wavelet domain has very similar drawbacks because the wavelet transform is neither shift nor rotation invariant. Most proposed methods watermark in the spatial domain. This is probably due to the simplicity and efficiency of such methods. The number of publications on DCT-based methods is also large.

## VI. VIDEO WATERMARKING

Video sequences consist of a series of consecutive and equally time-spaced still images. Thus, the general problem of watermarking seems very similar for images and video sequences, and the idea that image watermarking techniques are directly applicable to video sequences is obvious. This is partly true, and there are a lot of publications on image watermarking which conclude with the remark that the proposed approach is also applicable to video. Since image watermarking has been covered in great detail in Section V, we do not repeat it here, even if some of them carry the word video in the title [26]. However, there are also some important differences between images and video which suggest specific approaches for video.

One important difference is the available signal space. For images, the signal space is very limited. This motivates many researchers to employ implicit or explicit models of the HVS, in order to reach the threshold of visibility and to embed a watermark as robust as possible without sacrificing image quality. Examples have been cited in Section V. For video, the available signal space, i.e., the number of pixels, is much larger. On the other hand, video watermarking often imposes real-time or near-real-time constraints on the watermarking system. As a consequence, it is less important, and for many applications even prohibitively complex, to use watermarking methods based on explicit models of the HVS. Complexity in general is a much more important issue for video watermarking applications than it is for image watermarking applications.

For individual watermarking, i.e., fingerprinting, of video sequences (for example, embedding of a receiver ID), this problem is even more severe because video sequences are usually stored in compressed format. Uncompressed storage and on-the-fly compression, or decompression, watermarking, and recompression, are usually not feasible for this kind of application, unlike for images. Thus, such applications may require compressed-domain watermarking, as presented in [47], [49], and [80] and discussed below.

Another point to consider is that the structure of video as a sequence of still images gives rise to particular attacks, for example, frame averaging, frame dropping, and frame swapping [47], [126]. At frame rates of 25–30 Hz, as they are used in television, this would possibly not be perceived by the casual viewer. A good watermarking scheme, however, should be able to resist to this kind of attack, for example, by distributing watermark information over several consecutive frames. On the other hand, it might be desirable to retrieve the full watermark information from a short part of the sequence. It depends on the application of which of those two competing requirements is realized (or both, e.g., by embedding a multiscale watermark with more than one temporal scale [126] or progressive watermark transmission [33]).

While a lot of research has been published on image watermarking, there are fewer publications that deal with video watermarking. However, the interest in such techniques is high, for example, the emerging digital versatile

disk (DVD) standard which will contain a copy protection system employing watermarking.<sup>1</sup> The goal is to mark all copyrighted video material such that DVD standard compliant players or recorders will refuse to play back or record pirated material.

In the following, some watermarking methods exploiting uncompressed or compressed video properties are discussed. Some other methods that have been proposed but are in fact image watermarking techniques applied to image sequences with or without subsequent compression are not discussed here.

Hartung and Girod [47]–[49] have concentrated on watermarking of compressed video for fingerprinting applications. They employ a straightforward spread-spectrum approach and embed an additive watermark into the video. The watermark is generated using a PN signal with the same dimensions as the video signal that is modulated with the information bits to be conveyed. Each information bit is redundantly embedded into many pixels. For each compressed video frame, the corresponding watermark signal frame is DCT transformed on an  $8 \times 8$  block-by-block basis, and the resulting DCT coefficients are added to the DCT coefficients of the video as encoded in the video bitstream. This is done for *I*, *P*, and *B* frames. A rate control is realized by individually comparing the number of bits for each encoded watermarked DCT coefficient versus the corresponding encoded unwatermarked coefficient. Due to variable length coding, the watermarked coefficient may or may not need more bits for encoding than the unwatermarked one. If more bits are required, and the bit rate of the video sequence may not be increased, the coefficient is not used for embedding. Due to the inherent redundancy in the watermark, the watermark information can still be conveyed as long as enough coefficients can be embedded. Visible artifacts, as they could be produced due to the iterative structure of hybrid video coding, are avoided by applying a drift compensation scheme. The added drift compensation signal is the difference of the motion compensated predictions from the unwatermarked and the watermarked sequence. Fig. 8 shows a basic block diagram of the method. The bit stream has to be parsed and the watermark has to be transformed with the DCT. However, the method does not require full decompression and recompression. The complexity of the scheme is in the same order of magnitude as decompression, and the embedded watermarks pertain in the video after decompression. The scheme is compatible with all DCT-based hybrid compression schemes, for example, MPEG-2, MPEG-4, and ITU-T H.263. MPEG-4 has tools for compression of arbitrarily shaped objects. For nonrectangular border blocks of such objects, the shape-adaptive DCT (SA-DCT) [118] is used instead of the DCT. The watermarking scheme is also applicable to such border blocks, only that the DCT of the watermark has to be replaced by the SA-DCT. The watermark is recovered from the decompressed video by correlation using the same PN sequence that was used

<sup>1</sup> As of April 1999, two competing proposals from two different industry consortia are under evaluation.

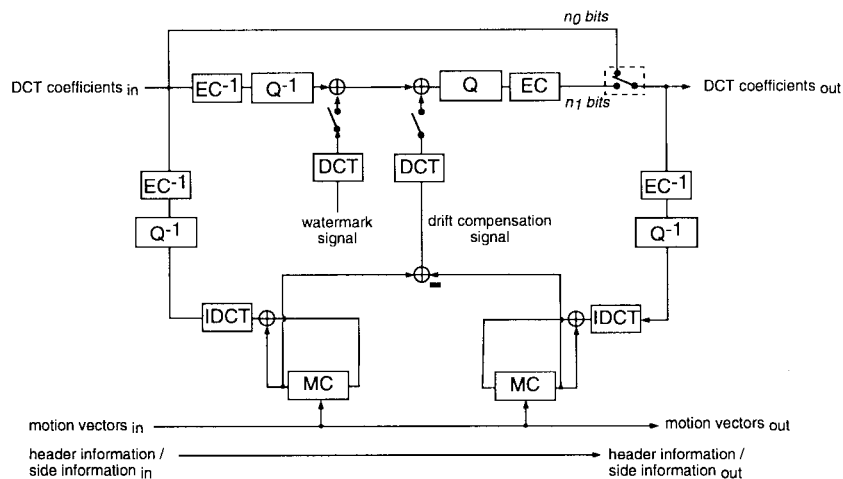


Fig. 8. Block diagram of watermark embedding into DCT coefficients of compressed video.

for generation of the embedded watermark signal. Typical watermark data rates are up to 50 bits/s, depending on the robustness requirements. The watermarks are robust against standard signal processing and with a modified watermark detector, as proposed in [50] also, to a certain extent, against geometrical distortions like shift, zoom, and rotation.

Jordan *et al.* [62] have proposed a method for the watermarking of compressed video that embeds information in the motion vectors of motion-compensated prediction schemes. Motion vectors pointing to flat areas are slightly modified in a pseudorandom way. Because the blocks pointed to by the original and the modified vectors are very similar (there is not much detail), this does not introduce any visible artifacts. The embedded information can be retrieved directly from the motion vectors, as long as the video is in compressed format. After decompression, the watermark can still be retrieved by first recompressing the video. This works because during recompression the watermarked motion vectors will be found with a probability high enough to statistically recover the watermark. The complexity of the method is negligible.

Hsu and Wu present a watermarking method [56], [57] for compressed video which is an extension of their method for images [55] and which modifies middle-frequency DCT coefficients in relation to spatially (for I-frames) or temporally (for P- and B-blocks) neighboring blocks. The coefficients are forced to assume a smaller or larger value than the corresponding neighboring coefficients, depending on the watermark sample to be embedded into the specific coefficient. The watermark signal is a visual pattern, like a logo, consisting of binary pixels. Prior to embedding, the watermark signal is spatially scrambled such that it can be recovered from a cropped version of the video. A drawback of the scheme is that for watermark extraction the watermarked video, the unwatermarked video, and the watermark have to be known.

In [80], Langelaar *et al.* propose two different information embedding schemes for compressed video. According to the different robustness and the definitions that we made in Section II, we call one of the methods a data-hiding method

and the other a watermarking method. The data-hiding method adds the label directly in the MPEG-1 or MPEG-2 bit stream by replacing variable length codes (VLC'S) of DCT coefficients. In MPEG (and other hybrid coding schemes), the quantized DCT coefficients are encoded using run/level encoding and subsequent variable length coding. In the MPEG-2 code tables there exist pairs of codes which represent the same run and levels that deviate only by one from each other. One of the codes is then assigned a "1," the other one a "0." The idea is to find VLC's in the bit stream for which such a "similar" code exists, and to eventually replace one by the other such that the bit to be embedded is coded in the choice of the VLC. In principle, this could be done for intra- and intercoded blocks, but the authors alter only intracoded blocks. Still, they can embed up to 8 kbits/s into TV resolution video. The authors do admit, however, that the label can be removed easily by decompression and recompression without seriously affecting the video quality. The watermarking method is more complex, but also more robust. It is based on discarding parts of the compressed video bitstream. For each information bit to be embedded, a set of  $n$   $8 \times 8$ -blocks is pseudorandomly taken from the video frame and, also pseudorandomly, divided into two subsets of equal size.  $n$  typically varies between 16 and 64. For each of the two subsets, the energy of the high-frequency DCT coefficients is measured. In order to embed the bit, the energy of the high-frequency coefficients in one or the other subset is reduced by removing high-frequency coefficients. The principle is illustrated in Fig. 9. For ease of understanding, consecutive blocks are used, rather than blocks randomly taken from the image. The information bit can be extracted by selecting the same set of blocks, dividing it into the same subsets, and comparing the energy of the high-frequency coefficients in each of the two subsets. Thus, the selection of blocks is the secret key involved. The method requires only partial decoding and no re-encoding. For TV resolution, up to 400 bits/s can be embedded. However, the robustness is limited. Re-encoding increases the error rate of the embedded bits much, and the method does not resist re-encoding using another group-



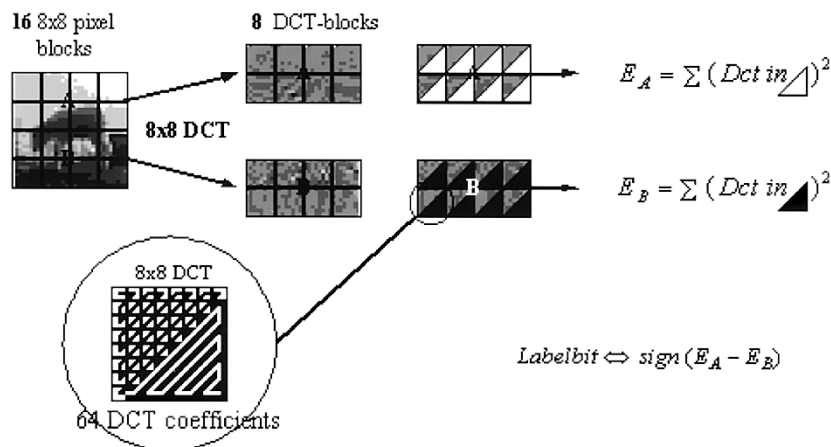


Fig. 9. Principle of DCT watermarking by comparison of the energy in the high-frequency coefficients. (Courtesy of G. Langelaar.)

of-picture (GOP) structure, since the DCT coefficients of a block are different depending on whether the frame is encoded as I, P, or B frame (however, in this case it is possible to extract the watermark by decoding and re-encoding the sequence with the same GOP structure that it had during watermarking [77]). Since DCT coefficients of the video are removed, care must be taken to adjust the parameters properly [79] in order to avoid visible blurring.

Swanson *et al.* [126], [127] propose a multiscale watermarking method working on uncompressed video which has some interesting properties. In a first step, the video sequence to be watermarked is segmented into scenes. Each scene is handled as an entity in the following. A temporal wavelet transform is then applied to each video scene, yielding temporal low-pass and high-pass frames. The watermark to be embedded is not an arbitrary message, but rather a unique code identifying the IPR owner and taken from a predefined codebook. In the design of the watermark, an explicit model of the HVS is employed in order to exploit spatial and temporal masking. Also, the watermark is designed with a signal-dependent key and thus avoids deadlock problems, as addressed in [30]. The watermark is embedded into each of the temporal components of the temporal wavelet transform, and the watermarked coefficients are then inversely transformed to get the watermarked video. Thus, the watermark has some components that change over time, while others do not or only slowly change over time, since they are embedded in the coefficients representing low temporal frequencies. This allows robustness against attacks like frame averaging, frame dropping, and the detection of the watermark from a frame of the scene without knowledge of its actual index. This is a property that the other video watermarking methods mentioned here do not automatically have. (Other video watermarking schemes could, however, achieve that with appropriate design of the watermark that they embed.) The watermark detection is done by hypothesis testing (the watermark is there or the watermark is not there). Experimental results show the robustness of the scheme against additive noise, MPEG video compression, and even



Fig. 10. Example for the structure of I, P, and B frames in a GOP.

frame drop. A disadvantage of the scheme is that it has a very high complexity, since it involves a forward and a backward wavelet transform, and an explicit model of the HVS including a blockwise DCT.

Linnartz *et al.* [83] propose to embed information encoded in the GOP structure of the MPEG-2 compressed video. In MPEG-2, video frames can be encoded in three different ways: as intracoded I frames coded JPEG like and without reference to other frames; as P frames predicted from previous frames; or as B frames bidirectionally predicted from previous and following frames. I frames are needed as random access points. Usually, there is a maximum distance between two successive I frames in order to allow random access with a maximum delay. The frame type is signaled in the frame header and can be switched randomly from frame to frame. The set of frames from one I frame (including the I frame) to the next (excluding the next) is referred to as GOP (see Fig. 10). Possible GOP structures are for example “IPPP,” “IBBPBBPBBPBB,” “IBBBBBBBB,” or “IPBPBBB,” and in fact there are  $2^{N-1}$  possible GOP structures for GOP’s of  $N$  frames. A popular GOP size is, for example,  $N = 12$ , thus allowing as many as 2048 different variations. However, most available video codecs use a fixed GOP size and structure, and never use most of the admissible GOP structures. The idea for data embedding is to purposely use those (irregular) GOP structures, that are very unlikely, to embed information. Linnartz *et al.* propose a scheme where they embed 6 bits of information per GOP, which means very few bytes per second. The method can only be employed during compression, not after compression where the GOP structure is already fixed. Also, information embedded as such is not resistant to decompression. Thus, decompression and recompression would already remove this information completely. Another disadvantage might

be that this type of watermark contradicts efforts to improve coding efficiency using rate-distortion optimized rate control [145], because such rate-distortion optimized video codecs are not restricted to a predefined GOP structure. A plus of the method is certainly that its complexity is negligible.

Darmstaedter *et al.* [33] propose to embed a spatial-domain low-pass spread-spectrum watermark into  $8 \times 8$  pixel blocks of video sequences. The blocks are first classified according to their activity. Blocks with low activity are not watermarked. A low-pass pseudorandom pattern is then added to each selected block. In principle, each block (64 pixels) conveys one bit watermark information, but the bits are redundantly repeated over several blocks and several frames. Also, the authors apply an error correcting code. After watermark embedding, the sequence is compressed using MPEG-2 compression. Watermark extraction is done in the spatial domain after decompression using a correlation concept with thresholding. In order to achieve error-free watermark retrieval for compression down to a video bit rate of 6 Mbit/s, the authors embed one bit of watermark information into a total of 162 000 pixels.<sup>2</sup> The authors have verified the method, including real transmission over digital satellite links, and optimized the embedding parameters manually. Depending on block mean and block variance, the individual pixels (PCM encoded with 8 bit) are modified by up to  $\pm 6$ .

Dittmann *et al.* [39] apply two previously proposed still image watermarking methods [44], [69] to video. The video is decompressed prior to watermarking and recompressed after watermarking. The authors are not precise about video formats, encoding parameters, or other details, but they admit that after recompression, and using an error correcting BCH (31, 6, 15) code, residual bit error rates of 1–5% for the watermark information bits remain. Already with slight attacks like format conversion from MPEG-2 to Quicktime, the bit error rates increase significantly. Thus, at least the parameters of the scheme are obviously not chosen adequately.

Deguillaume *et al.* [36] propose to embed a spread-spectrum watermark into 3-D blocks of video by employing a 3-D DFT and adding to the transform coefficients. The watermark is composed of the real watermark and an auxiliary pattern, called template, that is easy to detect even under geometric attacks and that can be used to undo such attacks to enable retrieval of the real watermark. The blocks that are processed consist of typically 16 or 32 frames. Since the template is embedded into the 3-D log-log-log map of the DFT, it is not affected by zoom and shift [115]. Results are reported for an 88-bit watermark embedded into 3-D blocks of 32 CIF resolution ( $352 \times 288$  pixels) frames each (giving a watermark data rate of 1 bit per 36 864 pixels). The reported bit error rates are 0% after high-quality compression (bit rate 4.75 Mbits/s for CIF 25 Hz [35]), but without attack, and they go up to around 20% in the presence of aspect-ratio changes and frame-rate

<sup>2</sup>64 bits are embedded into 25 frames of ITU-R 601 resolution video ( $720 \times 576$  pixels).

changes, even though the changes are recognized with help of the template and compensated. Thus, it seems that the parameters of the scheme should be chosen such that the watermark is embedded more robustly than presented in the simulations.

Busch *et al.* [19] apply a still-image watermarking method working on DCT blocks [69] to video sequences. The watermarks are embedded into the luminance component of uncompressed video and retrieved after decompression. In order to improve the invisibility of the watermarks, especially at edges, blocks are selected for watermarking depending on the block activity. For watermarking and watermark retrieval of a 64-bit watermark into each frame of ITU-R 601 video (that means into 5280 pixels/bit) and subsequent MPEG-2 compression at 4–6 Mbit/s, bit error rates between  $\approx 0$  and 50% are reported, depending on the sequence. For critical sequences, the authors propose to introduce additional temporal redundancy by embedding the watermark into several consecutive frames and averaging in the retrieval. For individual difficult sequences, averaging over 50 frames (corresponding to the embedding of one watermark bit into 264 000 pixels) still yields bit error rates of a few percent, and the authors propose averaging over an even higher number of frames for synthetic video.

Kalker *et al.* [65] have developed a video watermarking method for video broadcast monitoring applications which they call JAWS (just another watermarking system). For the sake of low complexity, both watermark embedding and detection are performed in the spatial domain, which means prior to compression and after decompression, respectively. The embedded watermark consists of watermark patterns of size  $128 \times 128$  drawn from a white random process with Gaussian distribution that are repeated (tiled) to fill the whole video frame. In order to avoid visible artifacts, the watermark is, on a pixel-by-pixel basis, scaled with a scaling factor which is derived from an activity measure. The activity measure is computed using a Laplacian high-pass filter. The same watermark is embedded into several consecutive video frames. For watermark detection, a correlation detector is used after applying a spatial prefilter that reduces cross talk between video signal and watermark. Since the watermark must be detected even in the presence of spatial shifts, a search over all possible shifts is performed. Since the watermark signal is generated by tiling of a smaller watermark pattern, only  $128 \times 128$  positions have to be searched, according to the size of the watermark pattern. In order to reduce complexity, the search and correlation is done in the FFT domain. Further, only the phase information of the FFT is used in the correlation. This method of detection has been previously proposed for pattern recognition and is referred to as symmetrical phase only filtering (SPOMF). In order to embed arbitrary watermark information, the watermark signal is designed using several different basic watermark patterns. The information is encoded in the choice of the basic patterns and their relative positions. The watermark can convey up to about 35–50 bits/s, but for applications

that require less watermark information per second the watermark data rate is reduced for increased robustness [63]. The method is claimed to be robust against MPEG-2 compression down to 2 Mbits/s, format conversion, scaling, and addition of noise.

Summarizing the above mentioned watermarking methods for video, a few general observations can be made.

- 1) The proposed methods span a wide complexity range from very low complexity to considerable complexity including, e.g., wavelet transforms and models of the HVS. In general however, the more complex methods seem to embed the watermarks with higher robustness.
- 2) Most methods operate on uncompressed video; only a few methods can embed watermarks directly into compressed video. For watermarking of compressed video watermarks can be embedded in the DCT coefficients [47], [49], [80], in the motion vectors [62], or in side information like the GOP structure [83].
- 3) The reported watermark data rates are between a few hundred bits per second and a few bits per second for television resolution video. It seems that if robustness is a real concern realistic data watermark data rates are not higher than a few bits per second to a few dozen bits per second. However, this is sufficient for most applications, including DVD.

## VII. AUDIO WATERMARKING

Compared to images and video, audio signals are represented by much less samples per time interval. This alone indicates that the amount of information that can be embedded robustly and inaudibly is much lower than for visual media. An additional problem in audio watermarking is that the human audible system (HAS) is much more sensitive than the HVS, and that inaudibility is much more difficult to achieve than invisibility for images.

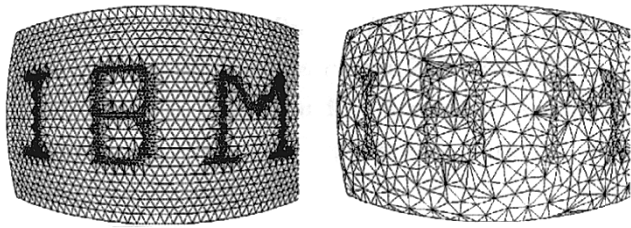
Boney *et al.* [11] propose a spread-spectrum approach for audio watermarking. They use a PN sequence that is filtered in several stages in order to exploit long-term and short-term masking effects of the HAS. In order to exploit long-term masking, a masking threshold for each overlapping block of 512 samples is calculated and approximated using a tenth-order all-pole filter which is then applied on the PN sequence. Short-term masking is additionally exploited by weighting the filtered PN sequence with the relative time-varying energy of the signal in order to attenuate the watermark signal where the audio signal energy is low. Additionally, the watermark is low-pass filtered by using a full audio compression/decompression scheme as low pass, in order to guarantee that it survives audio compression. A high-pass component of the watermark is also embedded which improves watermark detection from uncompressed audio pieces but is expected to be removed by compression. The authors denote the two spectral components of the watermark by “low-frequency watermark” and “coding error watermark.” The watermark

can be extracted by hypothesis testing using the original and the PN sequence and by employing a correlation method. Experimental results show the robustness of the scheme to MPEG-1 layer III audio coding, to coarse PCM quantization using word lengths down to 6 bits/sample instead of 16 bits/sample as for the original, and additive noise.

Bassia and Pitas [5] apply a very straightforward time-domain spread-spectrum watermarking method to audio signals. They report robustness against audio compression, filtering and resampling.

Tilki and Beex [134] have developed a system for an interactive television application where they embed information into the audio component of a television signal. The embedded information is detected from the acoustic signal emitted from the television receiver. Though the system is designed for analog transmission, the principle could similarly be applied to digital signals. The information to be embedded is partitioned in blocks of 35 bits. Each information bit is modulated using a sinusoidal carrier of a specific frequency and low amplitude and added to the audio signal. The simplified principle is that if the sinusoidal carrier for a specific bit is present in the signal, the bit is “1,” otherwise it is “0.” The frequencies of the sinusoidal carriers are above 2.4 kHz, thus at frequencies where the HAS is less sensitive, no explicit model of the HAS is employed. In order to reduce interference from the audio signal itself, the audio signal is attenuated at frequencies above 2.4 kHz. Thus, the principle involves a fidelity loss of the host signal which seems acceptable for the envisaged application. In order to increase the robustness, the information bits are protected by a cyclic redundancy code (CRC) and bit repetition. In order to compensate frequency shift of the whole signal, for example, after analog recording and playback with inaccurate speed, a frequency locking mechanism is applied using five special sinusoidal carriers of known frequency. Thus, the scheme is robust against room noise and video tape recording.

Bender *et al.* [6] propose several techniques for watermarking which are applicable to audio. They call the techniques spread-spectrum coding, echo coding, and phase coding. Direct sequence spread-spectrum coding is performing biphasic shift keying on a carrier wave by using an encoded binary string and pseudorandom noise. The code introduces perceptible noise into the original sound signal, but by using adaptive coding and redundant coding the perceptible noise can be reduced. Echo coding is a method which employs multiple decaying echos to place a peak in the cepstrum at a known location. The result is that moderate amounts of data can be hidden in a form that is fairly robust versus “analog” transmission. Phase coding is a method that employs the phase information as a data space. For the encoding, a Fourier transform is applied and the phase values of each frequency component are lined up as a matrix; binary information can be embedded into this matrix by modifying the phase component. Since the human HAS is not very sensitive to the distortion to the phase of the sound, it can be used to encode data without introducing much audible distortion to the original sound.



**Fig. 11.** Embedding of visible watermarks into 3-D meshes by local variation of the mesh density. (Figure taken with kind permission from [94].)

## VIII. WATERMARKING OF OTHER MULTIMEDIA DATA

Most watermarking research, publications, and products are dedicated to images. Less has been published on video, audio, and formatted text watermarking, and even less on watermarking of other media. However, the underlying basic ideas are certainly applicable to almost all kinds of digital data.

Ohbuchi *et al.* [94], [95] have proposed methods for embedding visible and invisible watermarks into 3-D polygonal models. Such models comprise primitives like points, lines, polygons, and polyhedrons, which are attributed by their geometry and their topology. Ohbuchi *et al.* propose to modify geometry or topology for watermarking. In detail, they propose two different methods for embedding of invisible watermarks for models consisting of triangular meshes. The first method pseudorandomly selects sets of four adjacent triangles and embeds information by displacing the vertices of the four triangles in a specific way by up to 1% of the shortest edge of the rectangular bounding box of the entire 3-D model. The authors claim that the modifications are imperceptible and that the method is resistant to cropping if the watermark information is repeated several times over the 3-D model and to local deformation. The second method pseudorandomly selects tetrahedron from the mesh and embeds information in the volume ratio of consecutive tetrahedron by modification of vertices. This method is robust against cropping and local deformation. A third method embeds visible watermarks into meshes by local variation of the mesh density, as shown in Fig. 11.

The emerging video compression standard MPEG-4 features additional functionalities, besides common video compression, such as model-based animation of 3-D head models using so-called facial animation parameters (FAP's). These are parameters like "rotate head," "open mouth," or "raise right corner-lip." The head model used at the receiver is either a predefined generic head and face model or a particular model that can be transmitted using so-called facial definition parameters (FDP's). The tool for face animation allows the compression of head-and-shoulder scenes, for example, in video telephony applications, with bit rates below 1000 bits/s. In [46], Hartung *et al.* propose a spread-spectrum method for watermarking of MPEG-4 FAP's. The watermarks are additively embedded into the animation parameters. Smoothing of the spread-spectrum watermark by low-pass filtering and an adaptive amplitude

attenuation prevents visible distortions of the animated head models. The watermarks can be retrieved by correlation from the watermarked parameters, but also from video sequences showing 3-D head models animated with the watermarked parameters, even after modifications such as block-based compression. Fig. 12 shows examples of video frames from a sequence rendered from a 3-D head model and animation parameters. In this case, the parameters first have to be estimated from the sequence. An interesting point is that the watermark is not contained in the waveform representation of the depicted object (the pixels), but in the semantics (the way the head and face move).

## IX. WATERMARK APPLICATIONS, SECURITY, ROBUSTNESS, AND CRYPTOANALYSIS

### A. Applications

We have already seen in Section III that the requirements and the design constraints for watermarking technologies strongly depend on the final application. For obvious reasons there is no "universal" watermarking method. Although watermarking methods have to be robust in general, different levels of required robustness can be identified depending on the specific application-driven requirements.

In authentication applications, the watermarks have to resist only to certain attacks. Among all possible watermarking applications, authentication watermarks require the lowest level of robustness. The purpose of such watermarks is to authenticate the data content. For example, data can be watermarked such that the watermark can accommodate lossy compression, but they are destroyed as soon as the data are manipulated in a different way.

Applications such as data monitoring and tracking require a higher level of robustness. The main purpose is to detect or identify stored or transmitted data. Examples are automatic monitoring of radio broadcast for billing purposes or identification of images on the World Wide Web with the help of web crawlers. For such applications, the watermarks have to be easily extractable and must be reasonably robust, for example, against standard data processing like format conversion and compression.

In fingerprinting applications, watermarks are embedded that identify the recipient of each individual distributed copy. The purpose is to have a means to trace back pirated copies to the recipient who pirated it. Fingerprinting applications require a very high level of robustness against data processing and malicious attacks.

Watermarking for copyright protection is used to resolve rightful ownership and requires the highest level of robustness. However, robustness alone is not sufficient for such applications. For example, if different watermarks are embedded in the same data, it must still be possible to identify the first, authoritative, watermark. Hence, additional design requirements besides mere robustness apply, as discussed below.

In the following, we go into more details on how to resist malicious attacks and elaborate on design constraints for copyright protection applications of watermarking.



**Fig. 12.** Example frame from a video sequence rendered from (a) a 3-D head model and watermarked animation parameters and (b) a similar frame after subsequent MPEG-2 video compression at 600 kbit/s.

### B. Watermark Robustness

Robustness against attacks is a major watermarking requirement. Absolute robustness against all possible attacks and their combinations may be impossible to achieve. Thus, the practical requirement is that a successful attack must impair the host data to the point of significantly reducing its commercial value before the watermark is impaired so much that it cannot be recovered. In fact, with appropriate design, fairly high robustness can be achieved, but it should be pointed out that robustness always has to be traded against watermark data rate and imperceptibility, and the optimum tradeoff depends on the application.

1) *Classification of Attacks:* Following the classification in [50], four different types of attacks can be identified.

- 1) “Simple attacks” (other possible names include “waveform attacks” and “noise attacks”) are conceptually simple attacks that attempt to impair the embedded watermark by manipulations of the whole watermarked data (host data plus watermark) without an attempt to identify and isolate the watermark. Examples include linear and general nonlinear filtering, waveform-based compression (JPEG, MPEG), addition of noise, addition of an offset, cropping, quantization in the pixel domain, conversion to analog, and gamma correction.
- 2) “Detection-disabling attacks” (other possible names include “synchronization attacks”) are attacks that attempt to break the correlation and to make the recovery of the watermark impossible or infeasible for a watermark detector, mostly by geometric distortion like zooming, shift in spatial or temporal (for video) direction, rotation, shear, cropping, pixel permutations, subsampling, removal or insertion of pixels or pixel clusters, or any other geometric transformation of the data.
- 3) “Ambiguity attacks” (other possible names include “deadlock attacks,” “inversion attacks,” “fake-watermark attacks,” and “fake-original attacks”) are attacks that attempt to confuse by producing fake original data or fake watermarked data [54]. An example is an inversion attack [30]–[32] that

attempts to discredit the authority of the watermark by embedding one or several additional watermarks such that it is unclear which was the first, authoritative watermark.

- 4) “Removal attacks” are attacks that attempt to analyze the watermarked data, estimate the watermark or the host data, separate the watermarked data into host data and watermark, and discard only the watermark. Examples are collusion attacks [121], denoising, certain nonlinear filter operations [81], or compression attacks using synthetic modeling of the image (e.g., using texture models or 3-D models). Also included in this group are attacks that are tailored to a specific watermarking scheme and combat it by exploiting conceptual cryptographic weaknesses of the scheme that make it vulnerable to a specific attack.

It should be noted that the transitions between the groups are sometimes fuzzy and that some attacks do not clearly belong to one group. Collusion attacks could be argued to be a group of its own, since they require, unlike the other attacks, more than one differently watermarked copy of the data. However, since they attempt to reconstruct the unwatermarked original host data, and thus remove the watermark(s), the classification as a “removal attack” holds.

In the following, remedies are given that make watermarks more robust against malicious attacks.

2) *Remedies Against Simple, Waveform-Based Attacks:* As already mentioned, noise-like distortions, for example, due to lossy compression, result in a distorted watermark signal in the watermark recovery or verification process. There are two main remedies against such attacks: increasing the embedding strength or applying redundant embedding. Increasing the embedding strength is straightforward and efficient in many cases, especially if appropriate masking according to the properties of human perception is used to determine the maximum allowable embedding strength. Redundant embedding can be performed in many ways. In the spatial domain it might consist of embedding a watermark many times and then taking a majority vote in the recovery process. A more efficient technique could include the use of error-correcting codes [52], possibly

even with soft-decision decoding [51]. Both increasing the watermark strength and introducing redundancy either increase the watermark visibility/audibility or decrease the watermark data rate. Further, as pointed out before, it should be noted that there is a tradeoff between watermark robustness on one hand and watermark imperceptibility and watermark data rate on the other hand.

3) *Geometrical Distortions and Remedies:* Watermarks are typically most vulnerable to geometrical distortions. The reason is that, for most proposed watermarking methods, the watermark detector has to know the exact position of the embedded watermark. Geometrical distortions tend to destroy the synchronization such that watermark embedding and watermark detection are misaligned and do not fit anymore.

Simple geometric attacks include affine transforms, clipping, and cropping. Remedies against such attacks are difficult if the watermarking algorithm has not explicitly been designed to withstand such attacks [114]. For this “simple” geometrical attacks, the challenge consists of finding the original watermark reference within the host data. For watermarking schemes which require the original image to recover the watermark this may not be a real problem, since the geometrical distortion can be estimated from the two images and inverted. If the watermarking scheme does not have the original data available for the watermark recovery, many schemes still allow the reference recovery by using a full search over all possible manipulations using some kind of correlation criteria between the image and the watermark modulation sequence. If the geometrical distortion consists of simple cropping, translation, or rotation, this process is feasible. However, if the attack consists of any affine transform this becomes very intensive computationally. Another way to resist geometrical attacks is based on embedding a watermark reference within the host data. Gruhl and Bender [45] propose embedding invisible crosses into the image by modifying the LSB image plane. Later detection of the crosses allows exact determination of the undergone attack and thus its reversal. If resistance to cropping has also to be assured, the row and column information can be encoded in addition to the crosses. One simple way of doing so would, for example, consist of changing the horizontal and vertical spacing between crosses depending on the location within the image. Although fully functioning, this system is not very robust since the reference can very easily be removed or destroyed. Another example is the embedding of sinusoidal patterns in the color channel using a visibility metric to ensure invisibility, as proposed by Fleet and Heeger [42]. An extension of the method of Gruhl and Bender has been proposed by Kutter [76] in which a spatial watermark pattern is embedded four times into the host image by using predetermined horizontal and vertical shifts. In the recovery process an autocorrelation function of an estimated watermark pattern can be computed to determine the affine distortion. Applying the inverse transform then allows full recovery of the watermark. A more sophisticated geometrical attack is based on jittering [70], [100], [138].

Jittering cuts the data set in small chunks, than randomly removes or duplicates small pieces and then sticks the small chunks back together. If done in a smart way, this alteration introduces only little or even no perceptible artifacts. This attack has proven to be very efficient in removing watermarks for many algorithms. Remedies exist against this attack, depending on the algorithm. For example, the method proposed by Kutter *et al.* [74] resists jittering if the image under inspection is low-pass filtered before the watermark extraction process. For other methods this remedy might work as well.

4) *Watermark Removal Attacks and Remedies:* Collusion attacks are attacks that use several copies of the same host data with different embedded watermarks. Several types of collusion attacks have been examined by Cox *et al.* [27] and Stone [121]. In the following, a watermark observation refers to a watermarked data representation in any domain, e.g., spatial or frequency domain. The first attack is called statistical averaging, in which a new data set is created by taking the average of all available watermark observations. A second attack creates a new data set by taking the average of the minimum and maximum of all watermark observations. The third approach is based on introducing negative correlation as follows:

$$\hat{w} = \begin{cases} w_{\max}, & \text{if } w_{\text{median}} \leq \frac{(w_{\max} - w_{\min})}{2} \\ w_{\min}, & \text{otherwise} \end{cases} \quad (35)$$

where  $w_{\text{median}}$ ,  $w_{\min}$ , and  $w_{\max}$  are the median, minimum, and maximum of the all watermark observations. Stone shows that for the image watermarking scheme proposed by Cox *et al.* [27] and a watermark with uniform distribution, at least four watermark observations are required for a successful attack. In general, all these statistical attacks can successfully destroy embedded watermarks even if only a few watermarked data sets are available. Another collusion attack interleaves the different watermarked copies of the same data [121]. Small parts of different watermarked data sets are taken and reassembled in a new data set. A remedy against collusion attacks is to limit the available number of watermarked copies. Alternatively, it has been proposed to use collusion-secure codes to design watermarks [9], [10]. The drawback is that the code lengths increase exponentially with the number of codes.

If the watermark detector device is available, the Oracle attack, first proposed by Perrig [98] and further developed by Cox and Linnartz [28], [29], can be used to destroy the embedded watermark. Such a scenario is, for example, possible in copy control systems for digital media, such as the DVD. The watermark detector can be used to experimentally deduce its behavior and then destroy the watermark. Although commonly believed that this approach involves an extremely high complexity, the authors illustrate that this is not true and claim the complexity to be of order  $O(N)$ , where  $N$  is the number of data samples, for most watermarking system. If the watermark inserter is available, another attack is based on predistorting the original data set. The difference between the watermarked data set and

original data set is used to predistort the original data set through subtraction. The newly watermarked predistorted data set is then very unlikely to contain the watermark. One remedy against a predistortion attack is based on encryption using a random session key. Given a binary watermark  $W$  to be embedded into a set of data, it is first encrypted using a random encryption key  $k$  resulting in  $W_k$ . The key is then appended to the encrypted watermark to give the new watermark  $\hat{W}_K$ , which is then embedded into the host data set. The watermark detector can recover the embedded watermark and decrypt it. The predistortion attack fails because the watermark inserter is not deterministic anymore due to the fact that the embedded watermark changes each time.

A histogram-based attack called *Twin Peaks* for fixed depth bimodal watermarks has been proposed by Maes [88]. To illustrate the concept of the attack, let us consider an image histogram with a peak at the intensity level  $P$ . Further, let us assume that the image was watermarked with a uniformly distributed watermark with a bimodal amplitude of  $\pm d$ . In this case, the watermarking process maps 50% of the values from  $P$  to  $P + d$ , and the other 50% from  $P$  to  $P - d$ . The peak in the original histogram at intensity  $P$  is therefore replaced by two peaks at intensities  $P - d$  and  $P + d$  (hence the name *Twin Peaks*), both having half the height of the original peak. By looking at the histogram of a watermarked image, it is possible to determine the embedded watermark by detecting close by peaks with similar amplitude. The original value may then be estimated and substituted into the watermarked image in order to destroy the embedded watermark. Based on this idea, the author show how to successfully destroy embedded watermarks. The performance of the attack may be improved when a prediction of the embedded watermark is used instead of the watermarked image. The prediction is computed by filtering the image with a high-pass filter which can be seen as taking the difference between a pixel value and the local mean computed in a squared wind of size  $3 \times 3$ .

### C. Remedies Against Watermark Ambiguities

As mentioned at the beginning of this section, to resolve rightful ownership, it must be possible to determine the authoritative watermark in case several watermarks are present in a data set.

1) *Timestamps*: To determine who first signed a set of data, timestamps (provided by trusted third parties) should be used [117], [149]. Let  $X$  be the data to be time stamped and  $H$  the corresponding hash value. The owner sends an official request  $R_n = (H_n, I_n)$ , where  $I_n$  is the owners identification string, to an official third party time stamping service (TSS). The TSS produces a timestamp  $TS_n$

$$TS_n = S_k(n, I_n, H_n, T_n, I_{n-1}, H_{n-1}, T_{n-1}, L_n) \quad (36)$$

where  $n$  is the request number,  $T_n$  the time of the request, and  $S_k$  indicates that the message is signed with the public key of TSS.  $L_n$  is known as the linking string defined as

$$L_n = H(I_{n-1}, H_{n-1}, T_{n-1}, L_{n-1}) \quad (37)$$

and is used to avoid that the timestamp requester and the TSS collude to produce any timestamp they want. The TSS then waits for the next request and returns the new identification  $I_{n+1}$  of the originator. If someone challenges a timestamp  $TS_n$ , the owner can prove that is was stamped after and before those by  $I_{n-1}$  and  $I_{n+1}$ , respectively. If their documents are also called in question they can get in touch with  $I_{n-2}$  and  $I_{n+2}$ , and so on.

Because digital time stamping involves a trusted third party, the question might arise why to use watermarking in combination with timestamping since this is very similar to traditional copyright registration and protection of copyright laws.

2) *Noninvertible Watermarks*: Until the publications of Craver *et al.* [30]–[32] it was believed that with the help of the original, nonwatermarked data set one can easily prove rightful ownership. Craver *et al.* showed that having the original is not sufficient and introduced the expression of invertible watermarking schemes. Given an original data set  $I_o$  to be watermarked with  $W_1$

$$\hat{I}_o = I_o \oplus W_1 \quad (38)$$

where  $\hat{I}_o$  is the watermarked original and the operator  $\oplus$  represents watermark insertion. Craver *et al.* showed that certain watermarking methods are invertible and allow reverse engineering to produce a counterfeit original

$$I_c = \hat{I}_o \ominus W_2 \quad (39)$$

where  $I_c$  it the counterfeit original and  $\ominus$  the inversion process. Let further assume that  $D$  is a watermark decoder function with a binary output of “0” and “1” for watermark absent and watermark present, respectively. This scenario creates an ownership deadlock because the rightful owner can show that his watermark is presents in the signed data and counterfeit original

$$\begin{aligned} D(I_o, \hat{I}_o, W_1) &= 1 \\ D(I_o, I_c, W_1) &= 1. \end{aligned} \quad (40)$$

However, the attacker can also show that his watermark  $W_2$  is present in the watermarked original, as well as in the original

$$\begin{aligned} D(I_c, \hat{I}_o, W_2) &= 1 \\ D(I_c, I_o, W_2) &= 1. \end{aligned} \quad (41)$$

Hence it is not possible to resolve rightful ownership since all claims from both parties are legally speaking equivalent. Some watermarking techniques are inherently invertible and the question is how to make them noninvertible or how to avoid this problem. Meanwhile, several methods have been devised to construct noninvertible watermarks [92], [110], [128]. The general idea in most methods is to make watermarks noninvertible by making them signal dependent, for example, by using one-way hash functions. In this case, it is computationally infeasible for an attacker to create a counterfeit original because it depends on



Fig. 13. Demonstration of the StirMark 2.2 attack.

the watermark, which in turn depends on the counterfeit original which is not yet existing.

It should also be noted that in applications where the owner of the data is undisputed, like, for example, in labeling applications where a serial number is embedded into different copies of distributed data, the above concerns do not apply.

#### D. Robustness Test Utilities and Watermark-Removal Software

Similar to conditional access and copy-prevention mechanisms, the existence of watermarking technology and its potential possibilities have stimulated individuals to come up with attempts to defeat watermarking. Examples are publicly available tools to test the robustness of image watermarking techniques. Unzign [138] is a utility that works for images in JPEG format. In version 1.1, Unzign introduces pixel jittering in combination with a slight image translation. For many proposed watermarking techniques, the embedded watermarks are efficiently destroyed. However, besides removing the watermark, Unzign version 1.1 introduces severe artifacts. An improved version 1.2 has been released. Although the artifacts were decreased, its watermark destruction capability decreased as well.

StirMark [70], [100] is a simple generic tool to test the robustness of image watermarking techniques. It simulates resampling to emulate a printing–scanning procedure and applies minor geometric distortions (stretching, shearing, shifting, and rotation) followed by resampling and bilinear or Nyquist interpolation. In addition, small and smoothly distributed errors are introduced into all sample values. Applying StirMark only once introduces a practically unnoticeable quality loss in the image. The author claims that his tool removes all current watermarks. Fig. 13 demonstrates the affect of the StirMark attack on a test image containing a grid and a natural image, and its StirMark 2.2 attacked version. From visual inspection, it can be confirmed that the effect of the attack is not visually annoying in the image, and is only evident in the grid. However, this attack is quite successful if the watermarking method does not account for it [50].

## X. THE FUTURE OF DIGITAL WATERMARKING

The interest in watermarking technology is high, both from academia and industry. The interest from academia is reflected in the number of publications on watermarking and in the fact that conferences on watermarking and data hiding are being held. The interest from industry is evident in the number of companies in the field that have been founded within the past few years.

Besides research activities in universities and industry, several international research projects funded by the European Community have the goal to develop practical watermarking techniques. TALISMAN [61] (ACTS project AC019, “Tracing Authors’ rights by labeling image services and monitoring access network”) aims to provide European Union service providers with a standard copyright mechanism to protect digital products against large scale commercial piracy and illegal copying. The expected output of TALISMAN is a system for protecting video sequences through labeling and watermarking. OCTALIS [60] (ACTS project P119, “Offer of Content through Trusted Access Links”) is the follow-up project of TALISMAN and OKAPI with the main goal of integrating a global approach to equitable conditional access and efficient copyright protection and to demonstrate its validity on large scale trials on the Internet and European Broadcasting Union (EBU) network.

International standardization consortia are also interested in watermarking techniques. The emerging video compression standard MPEG-4 (ISO/IEC 14 496), for example, provides a framework that allows the easy integration with encryption and watermarking. The DVD industry standard will contain copy control and copy protection mechanisms that use watermarking to signal the copy status of multimedia data, like “copy once” or “do not copy” flags.

Despite the many efforts that are underway to develop and establish watermarking technology, watermarking is still not a fully mature and understood technology, and a lot of questions are not answered yet. Also, the theoretical fundamentals are still weak, and most systems are designed heuristically.

Another drawback is that fair comparisons between watermarking systems are difficult [75]. As long as methods and system implementations are not evaluated in a con-



sistent manner using sophisticated benchmarking methods, the danger exists that weak and vulnerable systems and *de facto* standards are produced that result in spectacular failures and discredit the entire concept.

Thus, the expectations into watermarking should be realistic. It should always be kept in mind that every watermarking system involves a tradeoff between robustness, watermark data rate (payload), and imperceptibility. The invisible 10 000-bit-per-image watermark that resists all attacks whatsoever is an illusion (realistic numbers are approximately two orders of magnitude lower). Even when designed under realistic expectations, watermarks offer robustness against nonexperts but may still be vulnerable to attacks by experts.

Although proof of ownership was the initial thrust for the technology, it seems that there is a long way to go before watermarking will be accepted as a proof in court, and it is likely enough that this may never happen. In copyright-related applications, watermarking must be combined with other mechanisms like encryption to offer reliable protection.

Still, there exist enough applications where watermarking can provide working and successful solutions. Specifically for audio and video it seems that watermarking technology will become widely deployed. The DVD industry standard, as an example, will use watermarking for the copy protection system. Similarly, there exist plans to use watermarking for copy protection for Internet audio distribution. Broadcast monitoring using watermarking is another application that will probably widely be deployed for both audio and video.

Whether the development of watermarking technology will become a success story or not is an interesting yet unclear question. Watermarking technology will evolve, but attacks on watermarks as well. Careful overall system design under realistic expectations is crucial for successful applications.

## XI. CONCLUSIONS

In this overview paper, we reviewed the most important aspects, design requirements, system issues, and techniques for digital watermarking. The historical roots of digital watermarking derive mainly from steganography, the art of data hiding. Although digital watermarking and steganography are in some sense similar, the main difference lies in the notion of robustness for digital watermarks. Watermark robustness is one of the major design issues, besides imperceptibility. We have shown that the various digital watermarking applications, such as data tracking, data monitoring, and copyright protection, result in corresponding design issues and algorithm requirements. Some schemes require the original data set in order to recover an embedded watermark and others do not. Further, in some publications methods are proposed that allow full watermark extraction, whereas in other publications techniques are presented which only allow verification if a given watermark is present in the data under investigation. We have emphasized that these two approaches are inherently equivalent in that

a watermark-extraction scheme can be transformed into a watermark-verification scheme and vice versa. Although often associated to still images, video, and audio, digital watermarking is also applicable to other digital data such as text, 3-D meshes, or face animation parameters. We have elaborated on numerous watermarking techniques for still images, video, audio, text, and other multimedia data. It has been pointed out that a majority of techniques are inherently similar and based on modulation with a PN signal, often in combination with masking, for the embedding process and some kind of hypothesis testing using correlation in the watermark recovery process. Designing watermarking methods does not only have to consider robustness against standard data processing, but also robustness against malicious attacks. Several classes of attacks have been outlined, and remedies were given to make watermarks attack resistant. As a general statement, it can be said that watermarks should be sufficiently overdesigned and contain enough redundancy to ensure resilience against attacks. For copyright enforcement, additional aspects have to be considered. One problem is to prove who first watermarked data if several watermarks are present in the data. Solutions to this problem might consist of digital time stamping or watermark registration. Further, it has been shown that robustness is not sufficient to resolve rightful ownership, even if the original data are available. In addition, the used watermarking method needs to be noninvertible. Several techniques have been proposed to render invertible methods noninvertible, including hashing and time stamping. Although working systems are already available, research in digital watermarking has to continue. There is a huge demand from content providers and IPR owners. The market is currently far from being saturated and many more companies are expected to be founded in the near future. The question whether digital watermarks will be used as legal proof in court is not yet decided and difficult to answer. There are, however, other applications, like multimedia copy protection systems and data broadcast monitoring, where we will see watermarking in operation.

## ACKNOWLEDGMENT

The authors would like to thank Dr. I. Cox, Prof. E. Delp, Dr. A. Herrigel, Dr. T. Kalker, Prof. M. Kobayashi, D. Kundur, S. Moskowicz, Prof. I. Pitas, Prof. T. Pun, and Dr. J. Zhao for sharing their views on the future of watermarking technology. Significant parts of Section X are a summary of their contributions. The authors would further like to thank Dr. J. K. Su and the anonymous reviewers for their suggestions which helped to improve the quality of the paper. The second author thanks Prof. Ebrahimi, Swiss Federal Institute of Technology, Lausanne, for introducing him to the presented topic and is grateful for the technical discussions, insights, and hints.

## REFERENCES

- [1] R. J. Anderson and F. Petitcolas, "On the limits of steganography," *IEEE J. Select. Areas Commun. (Special Issue on Copyright and Privacy Protection)*, vol. 16, pp. 474–481, May 1998.

- [2] M. Barni, F. Bartolini, V. Cappellini, and A. Piva, "A DCT-domain system for robust image watermarking," *Signal Processing (Special Issue on Watermarking)*, vol. 66, no. 3, pp. 357–372, May 1998.
- [3] M. Barni, F. Bartolini, V. Cappellini, A. Piva, and F. Rigacci, "A M.A.P. identification criterion for DCT-based watermarking," in *Proc. Europ. Signal Processing Conf. (EUSIPCO '98)*, Rhodes, Greece, Sept. 1998.
- [4] P. Bas and J.-M. Chassery, "Using fractal code to watermark images," in *Proc. Int. Conf. Image Processing (ICIP)*, vol. 1, Chicago, IL, 1998.
- [5] P. Bassia and I. Pitas, "Robust audio watermarking in the time domain," in *Proc. European Signal Processing Conf. (EUSIPCO 98)*, Rhodes, Greece, Sept. 1998.
- [6] W. Bender, D. Gruhl, and N. Morimoto, "Techniques for data hiding," in *Proc. SPIE*, vol. 2420, San Jose, CA, Feb. 1995, p. 40.
- [7] D. Benham, N. Memon, B.-L. Yeo, and M. Yeung, "Fast watermarking of DCT-based compressed images," in *Proc. Int. Conf. Image Science, Systems, and Technology (CISST '97)*, Las Vegas, NV, June 1997, pp. 243–253.
- [8] F. M. Boland, J. J. K. Ó Ruanaidh, and W. J. Dowling, "Watermarking digital images for copyright protection," in *Proc. Int. Conf. Image Processing and Its Applications*, vol. 410, Edinburgh, U.K., July 1995.
- [9] D. Boneh and J. Shaw, "Collusion-secure fingerprinting for digital data," in *Advances in Cryptology—Proc. CRYPTO '95 (Lecture Notes in Computer Science)*, vol. 963, Don Copper-smith, Ed. Berlin, Germany: Springer, 1995, pp. 452–465.
- [10] —, "Collusion-secure fingerprinting for digital data," *IEEE Trans. Inform. Theory*, vol. 44, pp. 1897–1905, Sept. 1998.
- [11] L. Boney, A. H. Tewfik, and K. H. Hamdy, "Digital watermarks for audio signals," in *Proc. EUSIPCO 1996*, Trieste, Italy, Sept. 1996.
- [12] A. Bors and I. Pitas, "Embedding parametric digital signatures in images," in *EUSIPCO-96*, Trieste, Italy, Sept. 1996.
- [13] —, "Image watermarking using DCT domain constraints," in *Proc. Int. Conf. Image Processing (ICIP)*, Lausanne, Switzerland, Sept. 1996.
- [14] J. Brassil, S. Low, N. Maxemchuk, and L. O'Gorman, "Electronic marking and identification techniques to discourage document copying," *IEEE J. Select. Areas Commun.*, vol. 13, pp. 1495–1504, Oct. 1995.
- [15] —, "Hiding information in document images," in *Proc. 29th Annu. Conf. Information Sciences and Systems (CISS 95)*, Johns Hopkins Univ., Baltimore, MD, Mar. 1995, pp. 482–489.
- [16] G. W. Braudaway, K. A. Magerlein, and F. C. Mintzer, "Color correct digital watermarking of images," U.S. Patent 5530 759, June 1996.
- [17] O. Bruyndonckx, J. J. Quisquater, and B. Macq, "Spatial method for copyright labeling of digital images," in *Proc. IEEE Workshop Nonlinear Signal and Image Processing*, Halkidiki, Greece, June 1995.
- [18] S. Burgett, E. Koch, and J. Zhao, "A novel method for copyright labeling digitized image data," Fraunhofer Inst. Comput. Graphics, Darmstadt, Germany, Tech. Rep., Sept. 1994.
- [19] C. Busch, W. Funk, and S. Wolthusen, "Digital watermarking: From concepts to real-time video applications," *IEEE Comput. Graphics Applicat.*, pp. 25–35, Jan. 1999.
- [20] G. Caronni, "Ermitteln unauthorisierter Verteiler von maschinenlesbaren Daten," ETH, Zürich, Switzerland, Tech. Rep., Aug. 1993.
- [21] —, "Assuring ownership rights for digital images," in *Proc. VIS 95, Session "Reliable IT Systems"*, Vieweg, Germany, 1995.
- [22] B. Chen and G. W. Wornell, "Digital watermarking and information embedding using dither modulation," in *Proc. IEEE Workshop Multimedia Signal Processing*, Los Angeles, CA, Dec. 1998.
- [23] —, "Dither modulation: A new approach to digital watermarking and information embedding," in *IS&T/SPIE's 11th Annu. Symp., Electronic Imaging '99: Security and Watermarking of Multimedia Contents*, vol. 3657, San Jose, CA, Jan. 1999.
- [24] —, "An information-theoretic approach to the design of robust digital watermarking systems," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing 1999 (ICASSP '99)*, Phoenix, AZ, Mar. 1999.
- [25] G. Cooper and C. McGillem, *Modern Communications and Spread Spectrum*. New York: McGraw-Hill, 1986.
- [26] I. Cox, J. Kilian, T. Leighton, and T. Shamon, "Secure spread spectrum watermarking for images, audio and video," in *Proc. IEEE Int. Conf. Image Processing (ICIP 96)*, Lausanne, Switzerland, Sept. 1996.
- [27] —, "Secure spread spectrum watermarking for images, audio, and video," NEC Res. Inst., Princeton, NJ, Tech. Rep. 95-10, 1995.
- [28] I. J. Cox and J.-P. Linnartz, "Some general methods for tampering with watermarks," *IEEE J. Select. Areas Commun. (Special Issue on Copyright and Privacy Protection)*, vol. 16, pp. 587–593, May 1998.
- [29] I. J. Cox, J.-P. Linnartz, and T. Shamon, "Public watermarks and resistance to tampering," in *Proc. Int. Conf. Image Processing (ICIP)*, 1997.
- [30] S. Craver, N. Memon, B.-L. Yeo, and M. Yeung, "Can invisible watermarks resolve rightful ownerships?," IBM, IBM Res. Rep. RC 20509, July 1996.
- [31] —, "On the invertibility of invisible watermarking techniques," in *Proc. IEEE Int. Conf. Image Processing 1997 (ICIP '97)*, vol. 1, Santa Barbara, CA, Oct. 1997, pp. 540–543.
- [32] —, "Resolving rightful ownerships with invisible watermarking techniques: Limitations, attacks, and implications," *IEEE J. Select. Areas Commun. (Special Issue on Copyright and Privacy Protection)*, vol. 16, pp. 573–586, May 1998.
- [33] V. Darmstaedter, J.-F. Delaigle, D. Nicholson, and B. Macq, "A block based watermarking technique for MPEG-2 signals: Optimization and validation on real digital TV distribution links," in *Proc. European Conf. Multimedia Applications, Services, and Techniques—ECMAST '98*, Berlin, Germany, May 1998.
- [34] P. Davern and M. Scott, "Fractal based image steganography," in *Lecture Notes in Computer Science: Information Hiding*, vol. 1174. Berlin, Germany: Springer, 1996, pp. 279–294.
- [35] F. Deguillaume. (1999, Jan.). Video watermarking—MPEG-2 video samples used for 3D-DFT video watermarking tests. [Online]. Available WWW: <http://cuiwww.unige.ch/deguilla/WM/wm.html>.
- [36] F. Deguillaume, G. Csurka, J. Ó Ruanaidh, and T. Pun, "Robust 3D DFT video watermarking," in *IS&T/SPIE's 11th Annu. Symp., Electronic Imaging '99: Security and Watermarking of Multimedia Contents*, vol. 3657, San Jose, CA, Jan. 1999.
- [37] J. F. Delaigle, D. De Vleeschouwer, and B. Macq, "Low cost perceptible digital picture watermarking method," in *Proc. ECMAST'97*, Milan, Italy, May 1997, pp. 153–167.
- [38] G. Depovere, T. Kalker, and J.-P. Linnartz, "Improved watermark detection reliability using filtering before correlation," in *Proc. IEEE Int. Conf. Image Processing 1998 (ICIP 98)*, Chicago, IL, Oct. 1998.
- [39] J. Dittmann, M. Stabenau, and R. Steinmetz, "Robust MPEG video watermarking technologies," in *Proc. ACM Multimedia '98*, Bristol, U.K., Sept. 1998.
- [40] R. C. Dixon, *Spread Spectrum Systems with Commercial Applications*. New York: Wiley, 1994.
- [41] O. Emery, "Des filigranes du papier." *A.T.I.P. Bull.: Bull. de l'Association Technique de l'Industrie Papetière*, vol. 12, no. 6, pp. 185–188, 1958.
- [42] D. Fleet and D. Heeger, "Embedding invisible information in color images," in *Proc. IEEE Int. Conf. Image Processing 1997 (ICIP '97)*, Santa Barbara, CA, vol. 1, Oct. 1997, pp. 532–535.
- [43] P. G. Flickema, "Spread-spectrum techniques for wireless communications," *IEEE Signal Processing Mag.*, vol. 14, pp. 26–36, May 1997.
- [44] J. Fridrich, "Methods for data hiding," State Univ. New York, Binghamton, Tech. Rep., 1997.
- [45] D. Gruhl and W. Bender. (1995). Affine invariance. [Online]. Available WWW: <http://nif.www.media.mit.edu/DataHiding/affine/affine.html>.
- [46] F. Hartung, P. Eisert, and B. Girod, "Digital watermarking of MPEG-4 facial animation parameters," *Comput. Graphics*, vol. 22, no. 3, pp. 425–435, 1998.
- [47] F. Hartung and B. Girod, "Digital watermarking of raw and compressed video," in *Proc. SPIE Digital Compression Technologies and Systems for Video Commun.*, vol. 2952, Oct. 1996, pp. 205–213.
- [48] —, "Fast public-key watermarking of compressed video," in *Proc. IEEE Int. Conf. on Image Processing 1997 (ICIP '97)*, vol. 1, Santa Barbara, CA, Oct. 1997, pp. 528–531.

- [49] ———, “Digital watermarking of uncompressed and compressed video,” *Signal Processing (Special Issue on Copyright Protection and Access Control for Multimedia Services)*, vol. 66, no. 3, pp. 283–301, 1998.
- [50] F. Hartung, J. K. Su, and B. Girod, “Spread spectrum watermarking: Malicious attacks and counterattacks,” in *Proc. SPIE Security and Watermarking of Multimedia Contents 99*, San Jose, CA, Jan. 1999.
- [51] F. Hartung, “Digital watermarking and fingerprinting of uncompressed and compressed video,” Ph.D. dissertation, Telecommun. Lab., Univ. Erlangen–Nuremberg, Erlangen, Germany, 1999.
- [52] J. R. Hernández, F. Pérez-González, and J. M. Rodríguez, “The impact of channel coding on the performance of spatial watermarking for copyright protection,” in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing 1998 (ICASSP 98)*, Seattle, WA, vol. 5, May 1998, pp. 2973–2976.
- [53] J. R. Hernández, F. Pérez-González, J. M. Rodríguez, and G. Nieto, “Performance analysis of a 2-d multipulse amplitude modulation scheme for data hiding and watermarking still images,” *IEEE J. Select. Areas Commun.*, vol. 16, pp. 510–524, 1998.
- [54] M. Holliman and N. Memon, “Counterfeiting attacks on linear watermarking schemes,” in *Proc. IEEE Multimedia Systems '98, Workshop Security Issues in Multimedia Systems*, Austin, TX, June 1998.
- [55] C.-T. Hsu and J.-L. Wu, “Hidden signatures in images,” in *Proc. IEEE Int. Conf. Image Processing (ICIP 96)*, Lausanne, Switzerland, Sept. 1996, pp. 223–226.
- [56] ———, “Digital watermarking for video,” in *Proc. of DSP'97*, Santorini, Greece, July 1997.
- [57] C.-T. Hsu, “Digital watermarking for images and videos,” Ph.D. dissertation, Commun. Multimedia Lab., National Taiwan Univ., 1997.
- [58] H. Inoue, A. Miyazaki, A. Yamamoto, and T. Katsura, “A digital watermark based on the wavelet transform and its robustness on image compression,” in *Proc. Int. Conf. Image Processing (ICIP)*, Chicago, IL, 1998.
- [59] A. Johnson and M. Biggar, “Digital watermarking of video/image content for copyright protection and monitoring,” ISO, ISO Doc. ISO/IEC JTC1/SC29/WG11 MPEG97/M2228, July 1997.
- [60] P. Jones. Octalis. [Online]. Available WWW: <http://www.cordis.lu/esprit/src/octalis.htm>.
- [61] ———. Talisman. [Online]. Available WWW: <http://www.cordis.lu/esprit/src/talisman.htm>.
- [62] F. Jordan, M. Kutter, and T. Ebrahimi, “Proposal of a watermarking technique for hiding/retrieving data in compressed and decompressed video,” ISO/IEC Doc. JTC1/SC29/WG11 MPEG97/M2281, July 1997.
- [63] T. Kalker, private communication.
- [64] ———, “Watermark estimation through detector observations,” in *Proc. IEEE Benelux Signal Processing Symposium '98*, Leuven, Belgium, Mar. 1998.
- [65] T. Kalker, G. Depovere, J. Haitsma, and M. Maes, “A video watermarking system for broadcast monitoring,” in *Proc. SPIE IS&T/SPIE's 11th Annu. Symp., Electronic Imaging '99: Security and Watermarking of Multimedia Contents*, vol. 3657, Jan. 1999.
- [66] M. S. Kankanhalli, Rajmohan, and K. R. Ramakrishnan, “Content-based watermarking of images,” in *Proc. ACM Multimedia '98*, Bristol, U.K., Sept. 1998.
- [67] M. Kobayashi, “Digital watermarking: Historical roots,” IBM Research, Tokyo Res. Lab., Tech. Rep., Apr. 1997.
- [68] E. Koch, J. Rindfrey, and J. Zhao, “Copyright protection for multimedia data,” *Digital Media and Electronic Publishing*, 1996.
- [69] E. Koch and J. Zhao, “Toward robust and hidden image copyright labeling,” in *Proc. Workshop Nonlinear Signal and Image Processing*, Marmaros, Greece, June 1995.
- [70] M. Kuhn. (1997, Nov.). StirMark. [Online]. Available WWW: <http://www.cl.cam.ac.uk/mgk25/stirmark/>.
- [71] D. Kundur and D. Hatzinakos, “A robust digital image watermarking method using wavelet-based fusion,” in *Proc. IEEE Int. Conf. Image Processing 1997 (ICIP 97)*, vol. 1, Santa Barbara, CA, Oct. 1997, pp. 544–547.
- [72] ———, “Digital watermarking using multiresolution wavelet decomposition,” in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing 1998 (ICASSP 98)*, vol. 5, Seattle, WA, May 1998, pp. 2969–2972.
- [73] M. Kutter, F. Jordan, and F. Bossen, “Digital signature of color images using amplitude modulation,” in *Proc. Electronic Imaging 1997 (EI 97)*, San Jose, CA, Feb. 1997.
- [74] ———, “Digital signature of color images using amplitude modulation,” *J. Electron. Imaging*, vol. 7, no. 2, pp. 326–332, Apr. 1998.
- [75] M. Kutter and F. Petitcolas, “A fair benchmark for image watermarking systems,” in *Proc. SPIE IS&T/SPIE's 11th Annu. Symp., Electronic Imaging '99: Security and Watermarking of Multimedia Contents*, vol. 3657, Jan. 1999.
- [76] M. Kutter, “Watermarking resisting to translation, rotation, and scaling,” in *Proc. SPIE Int. Symp. on Voice, Video, and Data Communication*, Nov. 1998.
- [77] G. Langelaar, private communication.
- [78] C. Langelaar, J. C. A. van der Lubbe, and R. L. Lagendijk, “Robust labeling methods for copy protection of images,” in *Proc. Electronic Imaging*, San Jose, CA, Feb. 1997, vol. 3022, pp. 298–309. [Online]. Available WWW: <http://www-it.et.tudelft.nl/~gerhard/home.html>.
- [79] G. Langelaar, R. Lagendijk, and J. Biemond, “Watermarking by DCT coefficient removal: Statistical approach to optimal parameter settings,” in *Proc. SPIE IS&T/SPIE's 11th Annu. Symp., Electronic Imaging '99: Security and Watermarking of Multimedia Contents*, vol. 3657, Jan. 1999.
- [80] G. C. Langelaar, R. L. Lagendijk, and J. Biemond, “Real-time labeling methods for MPEG compressed video,” in *Proc. 18th Symp. Information Theory in the Benelux*, Veldhoven, The Netherlands, May 1997.
- [81] ———, “Removing spatial spread spectrum watermarks by nonlinear filtering,” in *Proc. Europ. Signal Processing Conf. (EU-SIPCO '98)*, Rhodes, Greece, Sept. 1998.
- [82] G. C. Langelaar, J. C. A. van der Lubbe, and J. Biemond, “Copy protection for multimedia data based on labeling techniques,” in *Proc. 7th Symp. Information Theory in the Benelux*, Enschede, The Netherlands, May 1996. [Online]. Available WWW: <http://www-it.et.tudelft.nl/~gerhard/home.html>.
- [83] J.-P. Linnartz. (1998). MPEG PTY marking. [Online]. Available WWW: <http://diva.eecs.berkeley.edu/linnartz/pty.html>.
- [84] S. Low and N. Maxemchuk, “Performance comparison of two text marking methods,” *IEEE J. Select. Areas Commun. (Special Issue on Copyright and Privacy Protection)*, vol. 16, pp. 561–572, May 1998.
- [85] S. Low, N. Maxemchuk, J. Brassil, and L. O’Gorman, “Document marking and identification using both line and word shifting,” in *Proc. Infocom '95*, Boston, MA, Apr. 1995.
- [86] H. D. Lüke, *Korrelations-signale* (in German). Berlin, Germany: Springer, 1992.
- [87] B. Macq, J.-F. Delaigle, and C. De Vleeschouwer, “Digital watermarking,” *SPIE Proc. 2659: Optical Security and Counterfeit Deterrence Techniques*, Mar. 1996, pp. 99–110.
- [88] M. Maes, “Twin peaks: The histogram attack on fixed depth image watermarks” in *Lecture Notes in Computer Science*, vol. 1525. Berlin, Germany; Springer, 1998, pp. 290–305.
- [89] M. J. J. B. Maes and C. W. A. M. Overveld, “Digital watermarking by geometric warping,” in *Proc. Int. Conf. Image Processing (ICIP)*, vol. 1, Chicago, IL, 1998.
- [90] K. Matsui and K. Tanaka, “Video-steganography,” in *Proc. IMA Intellectual Property Project*, vol. 1, Jan. 1994, pp. 187–206.
- [91] N. F. Maxemchuk and S. Low, “Marking text documents,” in *Proc. IEEE Int. Conf. Image Processing 1997 (ICIP '97)*, vol. 3, Santa Barbara, CA, Oct. 1997, pp. 13–16.
- [92] G. Nicchiotti and E. Ottaviani, “Non-invertible statistical wavelet watermarking,” in *Proc. Europ. Signal Processing Conf. (EUSIPCO '98)*, Rhodes, Greece, Sept. 1998.
- [93] N. Nikolaidis and I. Pitas, “Copyright protection of images using robust digital signatures,” in *Proc. ICASSP '96*, Atlanta, GA, May 1996.
- [94] R. Ohbuchi, H. Masuda, and M. Aono, “Embedding data in three-dimensional polygonal models,” in *Proc. ACM Multimedia '97*, Seattle, WA, Nov. 1997.
- [95] ———, “Watermarking three-dimensional polygonal models through geometric and topological modifications,” *IEEE J. Select. Areas Commun. (Special Issue on Copyright and Privacy Protection)*, vol. 16, pp. 551–560, May 1998.

- [96] J.-M. Chassery, P. Bas, and F. Davoine, "Self-similarity based image watermarking," in *Proc. Europ. Signal Processing Conf. (EUSIPCO '98)*, Rhodes, Greece, Sept. 1998.
- [97] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*. New York: McGraw-Hill, 1991.
- [98] A. Perrig, "A copyright protection environment for digital images," Diploma dissertation, École Polytechnique Fédérale de Lausanne, Switzerland, Feb. 1997.
- [99] F. A. P. Petitcolas, R. J. Anderson, and M. G. Kuhn, "Information hiding—A survey," this issue, pp. 1062–1078.
- [100] —, "Attacks on copyright marking systems," in *Lecture Notes in Computer Science: Information Hiding*. Berlin, Germany: Springer, 1998, pp. 218–238.
- [101] R. Pickholtz, D. Schilling, and L. Milstein, "Theory of spread spectrum communications—A tutorial," *IEEE Trans. Commun.*, vol. COM-30, pp. 855–884, May 1982.
- [102] R. L. Pickholz, D. L. Schilling, and L. B. Milstein, "Theory of spread-spectrum communications—A tutorial," *IEEE Trans. Commun.*, vol. 30, pp. 855–884, May 1982.
- [103] I. Pitas, "A method for signature casting on digital images," in *Proc. Int. Conf. Image Processing (ICIP)*, Lausanne, Switzerland, Sept. 1996.
- [104] I. Pitas and T. H. Kaskalis, "Applying signatures on digital images," in *Proc. IEEE Workshop Nonlinear Image and Signal Processing*, Neos Marmaros, Greece, June 1995, pp. 460–463.
- [105] A. Piva, M. Barni, E. Bartoloni, and V. Cappellini, "DCT-based watermarking recovering without resorting to the uncorrupted original image," in *Proc. IEEE Int. Conf. Image Processing (ICIP)*, vol. 1, Santa Barbara, CA, 1997, p. 520.
- [106] C. Podilchuk and W. Zeng, "Watermarking of the JPEG bit-stream," in *Proc. Int. Conf. Imaging Science, Systems, and Applications (CISST '97)*, Las Vegas, NV, June 1997, pp. 253–260.
- [107] C. I. Podilchuk, "Digital image watermarking using visual models," in *Proc. Electronic Imaging*, vol. 3016, San Jose, CA, Feb. 1996.
- [108] C. I. Podilchuk and W. Zeng, "Perceptual watermarking of still images," in *Proc. of Workshop Multimedia Signal Processing*, Princeton, NJ, June 1997.
- [109] J. Puate and F. Jordan, "Using fractal compression scheme to embed a digital signature into an image," in *Proc. SPIE Photonics East'96 Symp.*, Boston, MA, Nov. 1996.
- [110] L. Qiao and K. Nahrstedt, "Watermarking schemes and protocols for protecting rightful ownership and customer's rights," *J. Visual Commun. Image Representation*, vol. 9, no. 3, pp. 194–210, Sept. 1998.
- [111] M. Ramkumar and A. Akansu, "On the choice of transforms for data hiding in compressed video," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing 1999 (ICASSP '99)*, Phoenix, AZ, Mar. 1999.
- [112] J. J. K. Ó Ruanaidh, F. M. Boland, and O. Sinnén, "Watermarking digital images for copyright protection," in *Proc. Electronic Imaging and the Visual Arts*, Florence, Italy, Feb. 1996.
- [113] J. J. K. Ó Ruanaidh, W. J. Dowling, and F. M. Boland, "Phase watermarking of digital images," in *Proc. Int. Conf. Image Processing (ICIP)*, vol. 3, Sept. 1996, pp. 239–242.
- [114] J. J. K. Ó Ruanaidh and T. Pun, "Rotation, scale and translation invariant digital image watermarking," in *Proc. IEEE Int. Conf. Image Processing 1997 (ICIP 97)*, Santa Barbara, CA, vol. 1, Oct. 1997, pp. 536–539.
- [115] —, "Rotation, scale, and translation invariant spread spectrum digital image watermarking," *Signal Processing (Special Issue on Watermarking)*, vol. 66, no. 3, pp. 303–318, May 1998.
- [116] K. Sayood, *Introduction to Data Compression*. New York: Morgan Kaufmann, 1996, ch. 13.
- [117] B. Schneider, *Applied Cryptography*. New York: Wiley, 1996, ch. 4.
- [118] T. Sikora, "Low complexity shape-adaptive {DCT} for coding of arbitrarily shaped image segments," *Image Commun. (Special Issue on Coding Techniques for Very Low Bit-Rate Video)*, vol. 7, nos. 4–6, Nov. 1995.
- [119] M. K. Simon, J. K. Omura, R. A. Scholtz, and B. K. Levit, *Spread Spectrum Communications Handbook*. New York: McGraw-Hill, 1994.
- [120] J. R. Smith and B. O. Comiskey, "Modulation and information hiding in images," in *Lecture Notes in Computer Science: Information Hiding*, vol. 1174. Berlin, Germany: Springer, 1996, pp. 207–226.
- [121] H. S. Stone, "Analysis of attacks on image watermarks with randomized coefficients," NEC Res. Inst., Princeton, NJ, Tech. Rep., May 1996.
- [122] J. K. Su and B. Girod, "On the imperceptibility and robustness of digital fingerprints," submitted for publication.
- [123] —, "Power spectrum condition for L2-efficient watermarking," submitted for publication.
- [124] M. D. Swanson, M. Kobayashi, and A. H. Tewfik, "Multimedia data embedding and watermarking technologies," *Proc. IEEE (Special Issue on Multimedia Signal Processing)*, vol. 86, pp. 1064–1087, June 1998.
- [125] M. Swanson, B. Zhu, and A. Tewfik, "Data hiding for video-in-video," in *Proc. IEEE Int. Conf. Image Processing 1997 (ICIP '97)*, vol. 2, Santa Barbara, CA, Oct. 1997, pp. 676–679.
- [126] M. Swanson, B. Zhu, and A. Tewfik, "Multiresolution video watermarking using perceptual models and scene segmentation," in *Proc. IEEE Int. Conf. Image Processing 1997 (ICIP '97)*, vol. 2, Santa Barbara, CA, Oct. 1997, pp. 558–561.
- [127] M. Swanson, B. Zhu, and A. H. Tewfik, "Multiresolution scene-based video watermarking using perceptual models," *IEEE J. Select. Areas Commun. (Special Issue on Copyright and Privacy Protection)*, vol. 16, pp. 540–550, May 1998.
- [128] M. Swanson, B. Zhu, A. H. Tewfik, and L. Boney, "Robust audio watermarking using perceptual coding," *Signal Processing (Special Issue on Watermarking)*, vol. 66, no. 3, pp. 337–356, May 1998.
- [129] M. D. Swanson, B. Zhu, and A. H. Tewfik, "Robust data hiding for images," in *Proc. IEEE Digital Signal Processing Workshop*, Loen, Norway, Sept. 1996, pp. 37–40.
- [130] —, "Transparent robust image watermarking," in *Proc. of Int. Conf. Image Processing (ICIP)*, Lausanne, Switzerland, Sept. 1996.
- [131] K. Tanaka, Y. Nakamura, and K. Matsui, "Embedding secret information into a dithered multilevel image," in *Proc. 1990 IEEE Military Commun. Conf.*, Sept. 1990, pp. 216–220.
- [132] —, "Embedding the attribute information into a dithered image," *Syst. Comput. Japan*, vol. 21, no. 7, 1990.
- [133] B. Tao and B. Dickinson, "Adaptive watermarking in the DCT domain," in *Proc. Int. Conf. Image Processing (ICIP)*, Lausanne, Switzerland, Sept. 1996.
- [134] J. F. Tilki and A. A. Beex, "Encoding a hidden digital signature onto an audio signal using psychoacoustic masking," in *Proc. 7th Int. Conf. Digital Signal Processing Applications & Technology*, Boston, MA, Oct. 1996, pp. 476–480.
- [135] A. Tirkel, private communication.
- [136] A. Tirkel, G. Rankin, R. van Schyndel, W. Ho, N. Mee, and C. Osborne, "Electronic water mark," in *Proc. DICTA 1993*, Dec. 1993, pp. 666–672.
- [137] A. Tirkel, R. van Schyndel, and C. Osborne, "A two-dimensional watermark," in *Proc. DICTA 1993* (1997, July). UnZign watermark removal software. [Online]. Available WWW: <http://altern.org/watermark/>.
- [138] R. G. van Schyndel, A. Z. Tirkel, and C. F. Osborne, "A digital watermark," in *Proc. Int. Conf. Image Processing (ICIP)*, vol. 2, 1994, pp. 86–89.
- [139] A. J. Viterbi, *CDMA: Principles of Spread Spectrum Communication*. Reading, MA: Addison-Wesley, 1995.
- [140] G. Voyatzis and I. Pitas, "Applications of toral automorphisms in image watermarking," in *Proc. Int. Conf. Image Processing (ICIP)*, vol. 3, Lausanne, Switzerland, Sept. 1996, pp. 237–240.
- [141] —, "Chaotic mixing of digital images and applications to watermarking," in *Proc. Europ. Conf. Multimedia Applications, Services, and Techniques (ECMAST)*, Louvain-la-Neuve, Belgium, May 1996.
- [142] H. Wang and C. C. J. Kuo, "An integrated progressive image coding and watermark system," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing 1998 (ICASSP 98)*, vol. 6, Seattle, WA, May 1998, pp. 3721–3723.
- [143] J. Weiner and K. Mirkes, *Watermarking* (no. 257 in Bibliographic Series). Appleton, WI: Inst. Paper Chemistry, 1972.
- [144] T. Wiegand, M. Lightstone, D. Mukherjee, T. G. Campbell, and S. K. Mitra, "Rate-distortion optimized mode selection for very low bit rate video coding and the emerging H.263 standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 6, pp. 182–190, Apr. 1996.
- [145] R. B. Wolfgang, C. I. Podilchuk, and E. J. Delp, "Perceptual watermarks for digital images and video," this issue, pp. 1108–1126.

- [147] R. B. Wolfgang and E. J. Delp, "A watermark for digital images," in *Proc. Int. Conf. Image Processing (ICIP)*, Lausanne, Switzerland, Sept. 1996, pp. 219–222.
- [148] —, "A watermarking technique for digital imagery: Further studies," in *Proc. Imaging Science, Systems, and Technology*, Las Vegas, NV, June–July 1997, pp. 279–287.
- [149] —, "Overview of image security techniques with applications in multimedia systems," in *Proc. SPIE Int. Conf. Voice, Video, and Data Commun.*, Dallas, TX, Nov. 1997.
- [150] M. Wu, M. L. Miller, J. A. Bloom, and I. J. Cox, "A rotation, scale, and translation resilient public watermark," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing 1999 (ICASSP '99)*, Phoenix, AZ, 1999.
- [151] X. Xia, C. Boncelet, and G. Arce, "A multiresolution watermark for digital images," in *Proc. IEEE Int. Conf. Image Processing 1997 (ICIP '97)*, vol. 1, Santa Barbara, CA, Oct. 1997, pp. 548–551.
- [152] W. Zhu, Z. Xiong, and Y.-Q. Zhang, "Multiresolution watermarking for images and video: a unified approach," in *Proc. Int. Conf. on Image Processing (ICIP)*, Chicago, IL, 1998.



**Martin Kutter** received the B.Sc. degree from the Technikum Winterthur Ingenieurschule, Switzerland, in 1989 and the M.Sc. degree in electrical engineering from the University of Rhode Island, Kingston, in 1996. He is currently pursuing the Ph.D. degree at the Signal Processing Laboratory, Swiss Federal Institute of Technology, Lausanne, Switzerland.

From 1992 to 1994, he was working in the R&D department of a company in the medical industry. His research interests include digital watermarking, cryptography, data compression, and image morphing.



**Frank Hartung** (Student Member, IEEE) received the M.Sc. degree in electrical engineering from the Technical University of Aachen, Germany. He was a Ph.D. student at the Telecommunication Lab of the University of Erlangen–Nuremberg, Erlangen, Germany, where he worked on video watermarking and video compression.

Since the spring of 1999, he has been with the Research Department of Ericsson Eurolab, Herzogenrath, Germany, working on multimedia. His research interests include digital watermarking of video and other multimedia data, video compression and transmission, multimedia systems and technology, and telecommunications technology.