Diversity and Attack Characterization for Improved Robust Watermarking

Deepa Kundur, *Member, IEEE*, and Dimitrios Hatzinakos, *Senior Member, IEEE*

***Abstract*—We consider the use of novel communication tool sets to improve the performance of robust watermarking systems. In particular, we relate the effects of attacks on the watermark to signal interference in a fading environment and employ diversity transmission and channel estimation principles to improve performance. A nonstationary parallel binary symmetric channel (BSC) model of the watermark channel is introduced to more accurately characterize signal tampering and, hence, extract the watermark. Analysis of the system sheds light on novel strategies and domains to embed information to improve the performance of robust data hiding schemes. Simulation and testing verify our theoretical observations.**

***Index Terms*—Attack characterization, data communications, digital watermarking, robust data hiding, watermark channel, wavelet transform.**

I. Introduction

DIGITAL watermarking is the process by which a discrete data stream called a *watermark* is hidden within a *host* multimedia signal by imposing imperceptible changes on the signal. In many proposed techniques, this procedure entails the use of a secret key that must be used to successfully embed and extract the watermark. One major driving force for research in this area is the need for effective copyright protection scenarios for digital media. In such an application, a serial number or copy protection code is watermarked into the signal to protect to assign ownership or user rights. It is expected that an *attacker* will attempt to remove the watermark by intentionally modifying the *watermarked signal*. Thus, we must strive to embed the mark such that it is difficult to remove (without the use of the key) unless the marked signal is significantly distorted.

A popular analogy for watermarking is the process of data communications in which the goal is to effectively communicate the watermark information using information hiding techniques. Much of the work on robust digital watermarking is based on spread spectrum (SS) communication principles [1]–[9]. In SS watermarking, the embedded signal is generally a low-energy pseudo-randomly generated white noise sequence. It is *detected* by correlating the known watermark sequence with either the extracted watermark or the watermarked signal itself (if the host is not available for extraction). If the correlation factor is above a given threshold, then the watermark is detected. The antijamming properties of SS signaling makes it attractive for application in watermarking since a low energy, and hence imperceptible, watermark, that is robust to narrowband interference, can be embedded [8].

However, SS approaches have a number of limitations. SS signaling approaches are specifically vulnerable to the "near– far" problem [10]. For watermarking, this implies that if the energy of the watermark is reduced due to fading-like distortions on the watermark, any residual correlation between the host signal and watermark can result in unreliable detection [11]. In addition, they neither take into account spatial nonstationarity of the host image and attack interference nor readily incorporate adaptive techniques to estimate the statistical variations. Furthermore, the correlator receiver structures used for watermark detection are not effective in the presence of fading. Although SS systems in general try to exploit spreading to average the fading, the techniques are not designed to maximize performance.[1]

We hypothesize that many common multimedia signal distortions, including cropping, filtering, and perceptual coding, are not accurately modeled as narrowband interference. Instead, we believe that such signal modifications are *fading-like* on the watermark if embedded in an appropriate domain. The novel contribution of this work involves the application of communication diversity and channel estimation techniques, which are effective in fading environments to the problem of robust watermarking. Diversity is employed through watermark repetition and channel estimation through a *reference* watermark. Although it is well-known that repetition can improve the reliability of robust data hiding schemes, it is traditionally used to average out the effect of noise-like distortions. In this paper, we demonstrate that if properly engineered, repetition can often significantly improve performance and may be worth the apparent sacrifice in watermark bit rate. Specifically, if repetition is viewed as the application of communication diversity principles, we can demonstrate that proper selection of an appropriate watermark embedding domain with attack characterization can notably improve reliability.

It should be emphasized that the ideas presented in this work are meant to be employed within existing watermarking techniques and are not intended to replace well-established watermarking strategies such as modulation and SS watermarking.

A. Objectives of this Paper

The overall intent of this paper is to derive new insights into the digital watermarking problem in order to establish rules for effective algorithm design. This is accomplished, in general,

Manuscript received June 21, 2000; revised June 21, 2001. This work was supported by Communications and Information Technology Ontario (CITO). The associate editor coordinating the review of this paper and approving it for publication was Prof. Arnab K. Shaw.

The authors are with the Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON, Canada M5S 3G4 (e-mail: deepa@comm.toronto.edu; dimitris@comm.toronto.edu).

Publisher Item Identifier S 1053-587X(01)07780-7.

[1]SS is commonly used in wireless communications for its interference rejection capabilities of narrowband noise. It has no advantage in environments in which fading is prevalent. For such applications, path and antenna diversity are commonly used to overcome fading [10].

through the development of a framework for the analysis of a broad class of watermarking techniques.

Specifically, we will do the following.

A) We demonstrate that viewing watermark repetition as a communication diversity technique, we can identify appropriate domains to embed the watermark for improved robustness against fading-like attacks.

B) We use reference watermarking for attack characterization for optimal watermark estimation. Traditionally, a reference template has been used to undo geometric distortions; however, in this work, we employ it for the problem of watermark channel characterization.

C) We incorporate a new localized watermark channel model to practically implement our ideas. Analysis of this model casts light on appropriate domains in which to watermark. The importance of a watermark extraction stage that makes use of information concerning the attacker's actions in this setting is demonstrated.

We address watermark attacks that may be modeled as a noisy attenuating nonstationary communication channel. Geometric signal modifications such as scaling or affine transformations of images and video sequences do not fall within this class. Such distortions desynchronize the watermark signal, and hence, although the watermark signal exists within the host, it is not easily accessible. We assume in this work that the watermark is extracted/detected with ideal synchronization so that we may focus on the issue of reliable watermark retrieval in the presence of noise and scaling. We believe that geometric distortions, while tedious to address, are possible to overcome through appropriate signal computation and registration as discussed in [12]–[14].

The next section introduces the basic setting for the analytic work; the proposed assumptions and models are discussed. Section III highlights the main theoretical results. These observations and implications are verified through implementation of the novel principles into the robust reference watermarking algorithm reviewed in Section IV; results of testing are also summarized. Final remarks conclude the paper.

II. Models and Paradigms

In this section, we introduce the novel concepts proposed for our watermarking approach. They are treated in more abstract terms for pedagogical reasons and to motivate the theory of Section III. Section IV presents these same ideas in the context of a practical watermarking algorithm.

A. Novel Principles

1) Fading: Previous analytic work in the area of robust digital watermarking has assumed additive Gaussian watermark channels. That is, the effect (on the embedded watermark) of distortions on the overall watermarked signal is considered to be in the form of stationary additive Gaussian noise. Intuitively, however, it is clear that some degradations such as cropping or heavy linear filtering have the effect of completely destroying the watermark content in the associated components of the signal. For example, if the watermark is embedded in the spatial domain of an image, cropping the image to a quarter of its original area will annihilate those watermark signal components in the discarded region of the signal while leaving others intact. Similarly, if the watermark is placed in the discrete Fourier transform (DFT) components of the signal, then aggressive lowpass filtering will remove the existence of the watermark from high-frequency coefficients. This demonstrates how some very simple distortions have a varying (i.e., nonuniform) effect on the embedded watermark. That is, some watermark components are more severely distorted than others.

We assert in this paper that many watermark attacks are more appropriately modeled as fading like. Fading is a term used to describe the effect of a communication channel that attenuates the information-bearing signal amplitude in an unpredictable way. We do not assume a particular statistical or mathematical model for the degradations but believe that the traditional characteristics of a general fading processing are applicable. Specifically, these features include

- varying SNR with a possibility of an SNR of $-\infty$ representing a complete fade of the watermark signal which is unavoidable;
- unpredictability of SNR variations along the watermark channel prior to watermark transmission (i.e., embedding);
- independence of watermark signal attenuation in multimedia signal coefficients displaced significantly in space, frequency time or another component.

  The resulting degree of independence has an equivalent measure of coherence space, time, or bandwidth as in common SS communication environments.

It should be mentioned that these same characteristics are used to justify the use of a fading model for the classic partial-band jamming problem [15].

2) Diversity: One general way in which to improve reliability in an unknown, nonstationary environment susceptible to deep fades is to employ diversity. A communication channel can be broken into M independent subchannels, where the kth subchannel has associated capacity C_k for $k = 1, 2, \ldots, M$. Since, in a fading environment, some of these channels may have a capacity of zero at unpredictable times, to ensure reliable transmission, diversity principles are employed. Specifically, the same information is transmitted through each subchannel with the hope that at least one repetition will successfully be transmitted. For watermarking, we call this *coefficient diversity* because different coefficients within the host signal are modulated with the same information.

The sacrifice in employing diversity is the bandwidth expense since the same information is sent through M orthogonal resources. However, for many watermarking applications the payload is small. Furthermore, for video watermarking applications, there exists an abundance of data in which to embed the watermark information; therefore, watermark capacity is not an issue.

3) Channel Estimation and Postprocessing: In channel estimation, a training or reference sequence is employed to adjust the receiver filter to maximize detection reliability. Watermarking methods that do not attempt to depict the attacks fail to exploit the advantage of extraction after any signal modification and, hence, fundamentally operate in a nonoptimal manner. We evaluate and demonstrate the performance improvements of characterizing watermark attacks prior to extraction.

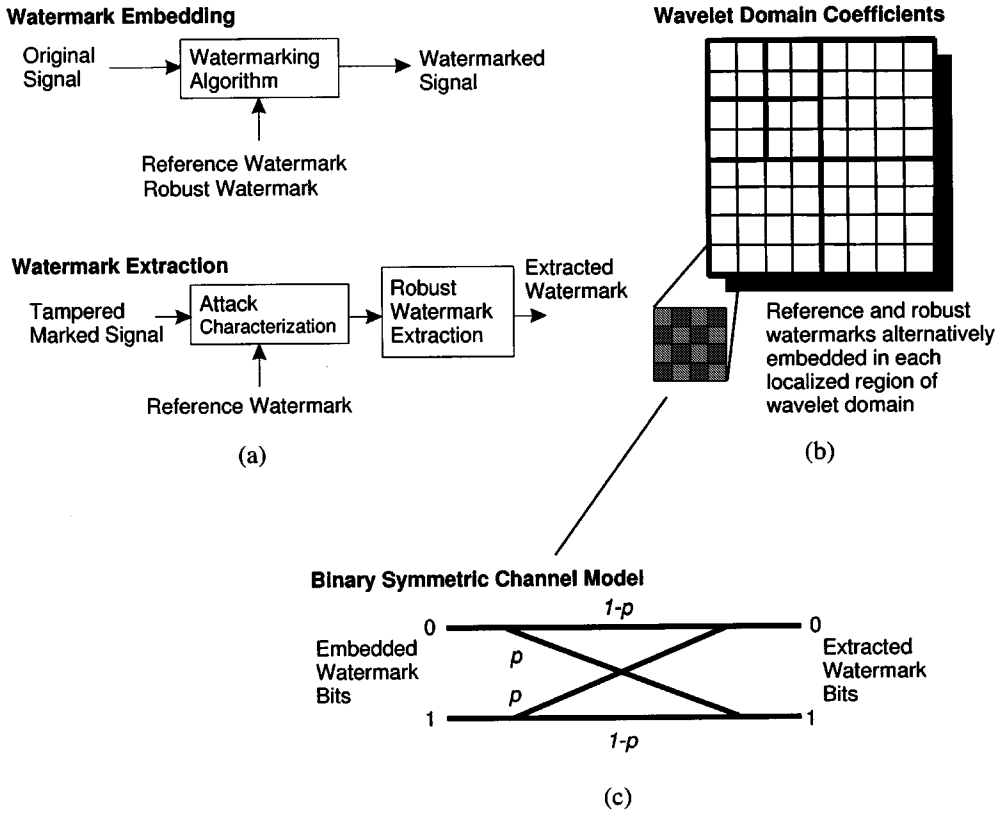

Fig. 1. Reference and robust watermarking employing coefficient diversity and attack characterization. (a) Overview of the watermark embedding and extracting scenarios. (b) We consider a 2-D host image; the watermark domain coefficients are divided into localized regions (outlined by the rectangular regions) so that reference and robust watermark bits are alternatively embedded in each region; the bold lines separate the different resolutions and wavelet detail coefficient classes (i.e., horizontal, vertical and diagonal). (c) The watermark in each region in modeled as passing through a binary symmetric channel (BSC) with a known probability of bit error estimated by the reference watermark.

Two questions naturally arise when incorporating coefficient diversity and channel estimation.

1) How do you combine the different extracted repetitions of the watermark to maximize the overall reliability of the system?
2) How do you define and characterize these "sub-channels" within the host signal to inherently promote robustness?

The subsequent sections address these issues. In Section II-B, we discuss integrating the various versions of the extracted watermark, and in Section III, we present analysis to identify reliable data embedding domains.

B. Context and Characterization

We limit the scope of our framework to the broad class of watermarking systems with the following basic characteristics.

1) The watermark w is binary and of length N_w bits. We denote the ith bit of w by $w(i)$ for $i = 1, 2, \ldots, N_w$.
2) The watermark information is repeatedly embedded $M \gg 1$ times within the host signal.
3) The embedding process occurs in the *watermark domain*. Specifically, an invertible continuous transformation $\mathbf{T}_w$ is applied to the host signal to produce coefficients in which the watermark bits are repeatedly inserted.
4) Each repetition of the watermark is embedded in a localized region of coefficients in the watermark domain. Fig. 1(a) and (b) elucidate the concept.
5) Each embedded watermark repetition is extracted separately. Thus, M different (and assumed independent) versions of the watermark $\hat{w}_1, \hat{w}_2, \ldots, \hat{w}_M$ are accessible from which to estimate the embedded watermark.
6) The watermarked signal may undergo distortions that affect the integrity of the embedded information. We assume that there exists a method of attack characterization such that each extracted watermark repetition has a known associated reliability. In our analysis, we make use of the probability of bit error measure. Specifically, $\hat{w}_k$ has an associated bit error likelihood of $\hat{p}_{Ek}$.

We incorporate diversity and channel estimation into our analysis framework through the use of watermark repetition and attack characterization, respectively. Each duplicate watermark is assumed to be separately extracted. Attack characterization is the process of measuring the reliability of each extracted watermark repetition. We do not specify how the characterization is performed here as this is an implementation issue discussed in Section IV, but we assume this reliability factor to be available.

Many proposed watermarking algorithms [6], [13], [16]–[20] (this is by no means an exhaustive list) are encompassed by this class of techniques or can be easily modified to fit this category. Although we restrict the watermark to be a bit sequence and the reliability measure to be the bit error rate, we believe the spirit of the results discussed in the paper holds for nonbinary watermarks with a different reliability measure such as the SNR.

1) Parallel BSC Watermark Channel Model: Given the characterization presented in the previous section, we can extract the

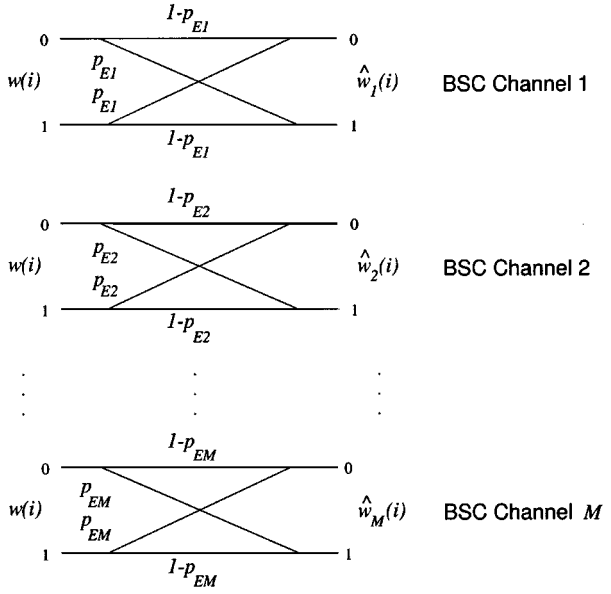

Fig. 2. Parallel binary symmetric channel model.

individual watermark repetitions to produce M estimates of the watermark $\hat{w}_1, \hat{w}_2, \ldots, \hat{w}_M$ with an associated probability of bit error p_{Ek}. Our framework is analogous to transmitting the watermark simultaneously along M independent BSCs, as shown in Fig. 2. The error probabilities $0 \leq p_{Ek} \leq 0.5$ are assumed to be known and independent of one another. If $p_{Ek} > 0.5$, then the output is complemented, and $1 - p_{Ek}$ is used as the probability of error parameter value. This type of localized characterization of the distortion in the watermark domain allows better modeling of nonstationary fading-like distortions. The basic stationary attack assumption precludes the benefits that diversity can provide and limits understanding into the advantages of using one watermarking domain over another.

There are important advantages to using the parallel BSC model. It is simple and the parameter p_{Ek} is easy to accurately estimate using an associated reference watermark we discuss in subsequent sections. In addition, a different parameter p_{Ek} for each w_k is incorporated, which provides a localized assessment of the attack in the watermark domain. Our attack characterization allows us to combine the repetitions based on a measure of their reliability to minimize the probability of watermark bit error.

2) Optimal Linear Watermark Estimation: To keep computational complexity low, we limit ourselves to linear watermark estimation. The overall extracted watermark $\hat{w}$ is computed as the weighted sum of the individual extracted repetitions. That is

$$\hat{w}(i) = \text{round}\left[\sum_{k=1}^{M} \alpha_k \hat{w}_k(i)\right] \tag{1}$$

where

round$[\cdot]$	integer round operator;
$\hat{w}(i)$ and $\hat{w}_k(i)$	ith watermark bits of $\hat{w}$ and $\hat{w}_k$, respectively;
α_k	associated scalar non-negative weight dependent on the estimated reliability factor $\hat{p}_{Ek}$ such that $\sum_{k=1}^{M} \alpha_k = 1$.

We show in [21] that the following assignment for α_k minimizes the bit error of $\hat{w}$ to produce an optimal linear watermark extraction:

$$\alpha_k = \frac{\log\left(\dfrac{1-\hat{p}_{Ek}}{\hat{p}_{Ek}}\right)}{\left(\displaystyle\sum_{j=1}^{M} \log\left(\frac{1-\hat{p}_{Ej}}{\hat{p}_{Ej}}\right)\right)}. \tag{2}$$

This linear estimation procedure is by no means the only alternative for combining the various extracted repetitions, but it is computationally simple, and it has been successfully implemented and tested [21]. Given the average watermark payload and order of magnitude of M in the range $10 \leq M \leq 100$ for still image watermarking, more sophisticated statistical estimation and combination principles may not be reliable. As pointed out by one anonymous reviewer of this paper, the formulation has the flavor of maximum likelihood estimation [22].

III. Analysis and Insights

In this section, we summarize our analysis to determine effective embedding strategies using our parallel BSC model of the watermark channel shown in Fig. 1.

A. Overview of the Math

Consider the bit $e_k(i)$ defined as

$$e_k(i) \triangleq w(i) \oplus \hat{w}_k(i) = \begin{cases} 1, & \text{if there is a bit error in } \hat{w}_k(i) \\ 0, & \text{otherwise.} \end{cases} \tag{3}$$

Similarly, we let

$$e(i) \triangleq w(i) \oplus \hat{w}(i) = \begin{cases} 1, & \text{if there is a bit error in } \hat{w}(i) \\ 0, & \text{otherwise.} \end{cases} \tag{4}$$

It easily follows from (1) that

$$e(i) = \text{round}\left[\sum_{k=1}^{M} \alpha_k e_k(i)\right] \tag{5}$$

which relates the bit errors of the individual extracted repetitions $e_k(i)$ to the bit error of the overall watermark estimate $e(i)$. Analysis of (5) is not straightforward due to the presence of the integer round operator. Alternatively, we consider the argument of this operator given by

$$\mathcal{E}(\mathbf{e}(i)) \triangleq \sum_{k=1}^{M} \alpha_k e_k(i) \tag{6}$$

where $\mathbf{e}(i) = [e_1(i)\, e_2(i) \cdots e_M(i)]^{\mathrm{T}}$. A bit error occurs in $\hat{w}(i)$ if $\mathcal{E}(\mathbf{e}(i)) > 0.5$. We can analyze the mean value of $\mathcal{E}(\mathbf{e}(i))$, $E\{\mathcal{E}(\mathbf{e}(i))\}$ to assess the reliability of the watermark channel. Although this is not a precise measure of the error rate of the system since a smaller $E\{\mathcal{E}(\mathbf{e}(i))\}$ does not necessarily guarantee a lower overall bit error rate, it does provide useful insight into the reliability of the watermarking procedure.

It is shown in Appendix A that

$$E\{\mathcal{E}(\mathbf{e}(i))\} \leq \frac{\overline{p}_E}{1-\overline{p}_E}\left[1 - \frac{D(q_a||q_b)}{\log\left(\frac{1-\overline{p}_E}{\overline{p}_E}\right)}\right] \tag{7}$$

where the average bit error rate is

$$\overline{p}_E = \frac{1}{M}\sum_{k=1}^{M} p_{Ek} \tag{8}$$

and $D(q_a||q_b)$ is the relative entropy given by [23]

$$D(q_a||q_b) = \sum_{k=1}^{M} q_a(k)\log\left(\frac{q_a(k)}{q_b k}\right) \tag{9}$$

$$q_a(k) = p_{Ek}/(M\overline{p}_E) \tag{10}$$

and

$$q_b(k) = (1-p_{Ek})/(M(1-\overline{p}_E)). \tag{11}$$

The sequences q_a and q_b are probability-like distributions; their elements are non-negative and sum to one. The bound of (7) is tight for small $\overline{p}_E$ and p_{Ek} close to a constant (i.e., $p_{Ek} \approx \overline{p}_E$ for all k). Specifically, the equality of (7) holds if and only if $p_{Ek} = 0$ for all k.

B. Implications

Observation of (7) reveals that the following possible tactics may be incorporated into a watermarking scheme to lower the value of $E\{\mathcal{E}(\mathbf{e}(i))\}$ and, hence, improve the robustness of the system in some way.

1) Reduce the value of the average bit error rate.

   Reducing the value of $\overline{p}_E$ decreases the term $\overline{p}_E/(1-\overline{p}_E)$ and increases the denominator term $\log((1-\overline{p}_E)/\overline{p}_E)$, which both serve to lower the overall bound.

   Many proposed watermarking methods attempt to gain performance by diminishing this average bit error rate. Signal processing strategies to imperceptibly embed a higher energy and, hence, more robust watermark are commonly employed.

   The deficiency of most watermarking methods is that they solely rely on embedding a stronger watermark using sophisticated human perceptual mathematical models for improved performance. Our next two theoretical observations shed light on a different strategy to increase robustness.

2) Embed the watermark such that the distributions q_a and q_b are dissimilar for a large class of distortions.

   For a fixed value of $\overline{p}_E$, we may reduce the performance bound by increasing the value of $D(q_a||q_b)$. The relative entropy is a measure of the distance between its two argument distributions [23]. Roughly, we can see that $D(q_a||q_b)$ is relatively large when $q_a(k) = p_{Ek}/(M\overline{p}_E)$ and its corresponding $q_b(k)\,(1-p_{Ek})/(M\overline{p}_E)$ are *dissimilar*.

   Assuming a fixed average probability of bit error, this requires that p_{Ek} vary in amplitude for different values of k, implying that we should embed the watermark in a domain for which the degree of distortion varies in each localized region containing a repetition on the watermark. As a result, the amplitude of p_{Ek} will be different for distinct values of k. This can be achieved by inserting the watermark in a domain that distributes the distortion more to certain coefficients, leaving others less affected.

3) For perfect watermark recovery, strive to localize the distortions on the watermarked signal.

   It is shown in Appendix B that the existence of $p_{Ek} = 0$ for at least one $k \in \{1, 2, \ldots, M\}$ implies that $\mathcal{E}(\mathbf{e}(i)) = 0$.

   Thus, if there exists a set of localized coefficients containing one complete repetition of the watermark that are unmodified by the distortion, then perfect watermark recovery is possible, as long as all the values of p_{Ek} are known. This translates to embedding the watermark in a domain that completely confines the distortion to a strict subset of the coefficients.

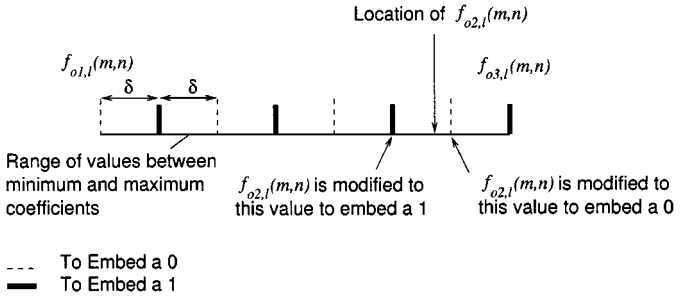

Fig. 3. Quantization procedure for watermarking. To embed the watermark, the median coefficient value $f_{o2,l}(m, n)$ is quantized to the nearest appropriate vertical line segment shown in bold (to embed a 1) or dashed (to embed a 0). The diagram provides an example case for $Q = 4$. The distance between each vertical line segment is δ, which is given by (14).

C. Discussion

Implications 2) and 3) relate the accuracy of the extracted watermark to the watermark domain in which the hidden data is embedded. By using diversity and attack characterization, it is possible to improve the effectiveness of the watermark to a broader class of distortions by inserting the mark in signal coefficients that localize these distortions. For example, to design a watermark robust against cropping, it would be wise to embed the mark in the spatial domain, which completely localizes the manipulation. Although a portion of the watermark is clipped out, the repetitions in the remaining signal are still accessible. Similarly, for robustness against filtering, the watermark should be embedded in the discrete Fourier domain that localizes the associated degradations.

It is also clear from simulation results (discussed in the next section) that a watermark embedded repeatedly in the spatial domain is significantly more robust (i.e., the watermark is perfectly recoverable unless the image is cropped beyond a certain threshold size) to cropping than if it were embedded in the frequency domain. Simulations demonstrate an analogous relationship with embedding in the Fourier domain and reliability against bandpass, highpass, or lowpass filtering. To make the watermark robust to both, a compromise would be to use the discrete wavelet domain for hiding the data.

TABLE I
PROPOSED ROBUST REFERENCE WATERMARK EMBEDDING METHOD. WE ASSUME THAT THE REFERENCE AND ROBUST WATERMARKS ARE OF THE SAME LENGTH (i.e., $N_w = N_r$) AND THAT THE COEFFICIENT SELECTION KEY SPECIFIES THAT ALL COEFFICIENTS ARE TO BE MARKED

1. Initialize user-defined variables:
   - Given: f the host image to be watermarked.
   - Given: $w(i)$, $i = 1, \ldots, N_w$, the robust watermark.
   - Select: $r(i)$, $i = 1, \ldots, N_r$, the reference watermark.
   - Select: $L \in \mathbb{Z}^+$, the maximum wavelet decomposition level.
   - Select: the particular DWT_t to apply on the host image.
   - Select: the quantization parameter Q.
2. Perform the Lth-level DWT_t on f to produce the detail images $f_{o,l}(m,n)$ for $o = 1,2,3$ and $l = 1,2,\ldots,L$, and an approximation image $f_{4,L}(m,n)$. That is,
$$\{f_{o,l}(m,n)\} := \mathrm{DWT}_t[f]. \tag{24}$$
3. Modify the wavelet coefficients to embed the watermarks:
```
for l = 1, 2, ..., L,
  for each localized spatial block R_s^l,
    i := 0
    for (m, n) ∈ R_s^l
      [f_o1,l(m,n), f_o2,l(m,n), f_o3,l(m,n)] := Sort(f_1,l(m,n), f_2,l(m,n), f_3,l(m,n))
      i := i + 1
      if i mod 2 = 1,
        f^w_o,l(m,n) := Quant(f_o1,l(m,n), f_o2,l(m,n), f_o3,l(m,n), Q, w(i))
      else,
        f^w_o,l(m,n) := Quant(f_o1,l(m,n), f_o2,l(m,n), f_o3,l(m,n), Q, r(i))
      end
    end
  end
end
```
4. Perform the Lth-level inverse DWT_t on the marked wavelet coefficients $f^w_{o,l}(m,n)$ to produce the marked image f_w. That is,
$$f_w := \mathrm{IDWT}_t\left[\{f^w_{o,l}(m,n)\}\right]. \tag{25}$$

The function $\mathrm{Sort}(a,b,c)$ sorts the input arguments in ascending order and the function $\mathrm{Quant}(\cdot)$ quantizes the corresponding input arguments to embed the watermark as discussed in Section 4.1.2.

To verify the insights derived in this work, we implement the proposed principles in a practical watermarking method discussed in the next section. Because of the importance of robustness against frequency and spatial domain distortions, our practical implementation embeds the watermark in the wavelet domain, which will be shown to localize these degradations. Not only does this domain allow more effective robust watermarking, but it is convenient for characterization of unlawful signal doctoring for telltale fragile watermarking, as discussed in [24]. Based on our analysis, we believe that the strengths of the scheme arise from the following factors.

- We embed the watermark in the wavelet domain, which localizes a diverse class of distortions such as spatial cropping and frequency domain filtering.
- The analysis assumes that the probabilities of bit errors p_{Ek} are known. Thus, to be robust to unknown distortions, estimation of the reliability of the watermark channels for each $\hat{w}_k$ must be employed. The proposed robust reference watermarking technique characterizes the locally varying distortions on the watermarked signal before extraction. In this way, the components of the watermark signal that is more accurate than the others can be identified.
- The proposed scheme is robust to a wider variety of distortions because diversity strategies are employed through the use of watermark repetition. Only one repetition of the watermark needs to be intact for reliable watermark recovery.

IV. IMPLEMENTATION AND TESTING

A. Robust Reference Watermarking

1) Role of the Reference Watermark: It has been shown in [25] and [26] that elementary characteristics of the signal distortions are easily estimated using a reference watermark. A reference watermark is one that is embedded into a signal for the purpose of detecting modifications. In fact, Tirkel *et al.* [12] have shown how reference marks can be used to help undo geometric distortions applied to images.

TABLE II
PROPOSED ROBUST REFERENCE WATERMARK EXTRACTION METHOD. WE ASSUME THAT THE REFERENCE AND ROBUST WATERMARKS ARE OF THE SAME LENGTH (i.e., $N_w = N_r$) AND THAT THE COEFFICIENT SELECTION KEY SPECIFIES THAT ALL COEFFICIENTS ARE TO BE MARKED

1. Initialize user-defined variables:
   - Given: $\hat{f}_w$ the host image to be watermarked.
   - Given: $r(i)$, $i = 1, \ldots, N_r$, the reference watermark.
   - Given: $L \in \mathbb{Z}^+$, the maximum wavelet decomposition level.
   - Given: the particular DWT_t to apply on the host image.
   - Given: the quantization parameter Q.
2. Perform the Lth-level DWT_t on $\hat{f}_w$ to produce the detail images $\hat{f}^w_{o,l}(m,n)$ for $o = 1,2,3$ and $l = 1,2,\ldots,L$, and an approximation image $\hat{f}^w_{4,L}(m,n)$. That is,
$$\{\hat{f}^w_{o,l}(m,n)\} := \text{DWT}_t\left[\hat{f}_w\right]. \tag{26}$$
3. Extract the robust and reference watermark repetitions:

```
k := 0
for l = 1,2,...,L,
    k := k + 1
    for each localized spatial block R_s^l,
        i := 0
        for (m,n) ∈ R_s^l
            [f_o1,l(m,n), f_o2,l(m,n), f_o3,l(m,n)] := Sort(f_1,l(m,n), f_2,l(m,n), f_3,l(m,n))
            i := i + 1
            if i mod 2 = 1,
                r̂_k((i − 1)/2) := Extract(f_o1,l(m,n), f_o2,l(m,n), f_o3,l(m,n), Q)
            else,
                ŵ_k(i/2) := Extract(f_o1,l(m,n), f_o2,l(m,n), f_o3,l(m,n), Q)
            end
        end
    end
end
M:=k
```

4. Calculate the BSC parameter estimates and compute the weights α_k:

```
for k = 1,2,...,M,
    p̂_Ek := (1/N_r) Σ_{i=1}^{N_w} r_k(i) ⊕ r̂_k(i)
    α_k := log((1 − p̂_Ek)/p̂_Ek) / (Σ_{j=1}^{M} log((1 − p̂_Ej)/p̂_Ej))
end
```

5. Combine the different robust watermark estimates

```
for i = 1,2,...,N_w,
    ŵ(i) := round[Σ_{k=1}^{M} α_k ŵ_k(i)]
end
```

where the function $\text{Sort}(a,b,c)$ sorts the input arguments in ascending order and the function $\text{Extract}(\cdot)$ extracts the watermark bit value from the DWT coefficient as discussed briefly in Section 4.1.2 and in more detail in [27].

We propose the scenario shown in Fig. 1(a) in which the host signal is embedded with both robust and reference watermarks. The reference watermark is assumed to be known at the extraction end as well. The two kinds of watermarks are placed in different coefficients of the host signal so that they do not interfere with one another. The tradeoff is that fewer repetitions of the robust watermark can be placed in the signal as a portion of the watermark "bit rate" is consumed by the presence of the reference watermark. Each embedded repetition of the robust watermark sequence, which we denote w_k, $k = 1, 2, \ldots M$ (where M is the total number of embedded copies), has an associated binary reference watermark sequence r_k, with the same statistical properties as w_k.[2] Fig. 1(b) demonstrates the embedding procedure where each w_k is placed in a localized region of the watermark domain. By localized region, we mean a relatively small closed compact region of regularly arranged coefficients. For example, this region could be a rectangular neighborhood of pixels. Fig. 1(b) shows how the wavelet domain can be divided into disjoint rectangular spatial neighborhood regions of coefficients at different resolutions. One robust and one reference watermark repetition are embedded into each of these regions. The bits of w_k are alternated with those of r_k such that an attack on the marked signal will reflect in the same way statistically on both w_k and r_k. Thus, if we let $\hat{w}_k$ and $\hat{r}_k$ be the extracted versions of w_k and r_k, respectively, after an attack, it is expected that the probability of bit error for $\hat{w}_k$ is equal to that for $\hat{r}_k$.

The probability of bit error of $\hat{w}_k$ is estimated as follows from r_k and $\hat{r}_k$

$$\hat{p}_{Ek} = \frac{1}{N_r}\sum_{k=1}^{N_r} r_k \oplus \hat{r}_k \tag{12}$$

where $\oplus$ is the exclusive OR operator.

[2]Note that since w_k is a repetition of the robust watermark, $w_k = w_j$ for all k and j. The reference watermarks $\{r_k\}$ do not necessarily have to be identical as long as their individual bit elements have the same statistical properties as that of the robust watermark. Both $\{w_k\}$ and $\{r_k\}$ are generated using a pseudo-random number emulating the same probability distribution.

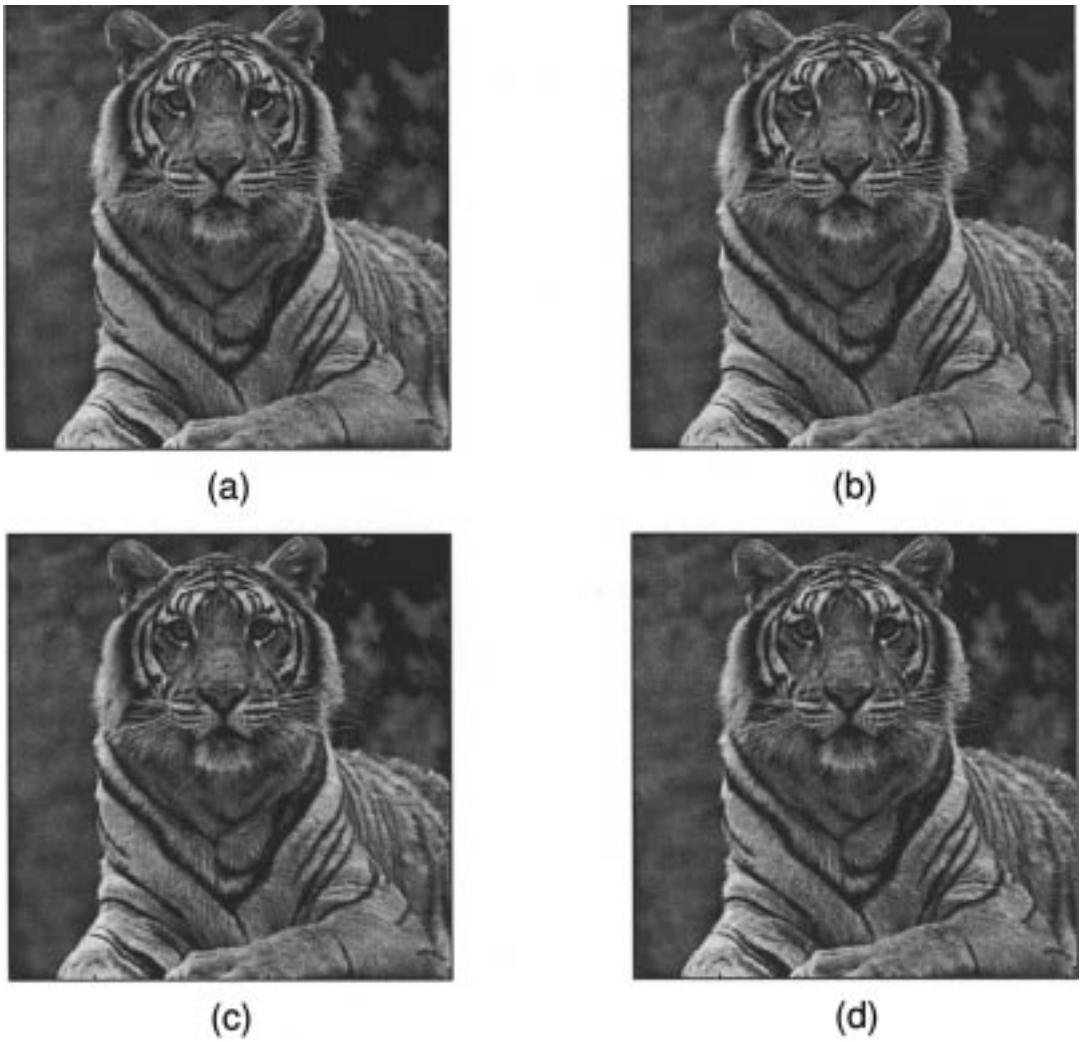

Fig. 4. Watermarked image results. (a) Host image. (b) Watermarked image using the Ohnishi and Matsui method. (c) Watermarked image using the Cox *et al.* method. (d) Watermarked image using the proposed robust reference watermarking approach.

The robust watermark is not used for estimation of the bit error as it may be unknown at the receiver. Use of the robust watermark, if it is known, will increase the probability of false positive detection that may not be appropriate for some applications [27].

It should be clarified that unlike traditional references used for identifying geometric distortions, our reference watermark r_k does not compromise the security of the technique. Its location and content is known only at the sender and the receiver. Because r_k and w_k are embedded in similar ways, removal of r_k is as difficult as w_k.

2) Data Embedding and Extraction: The technique employed to modify the host signal transform coefficients in order to embed the watermark can be arbitrary as long as it places the hidden data in a localized region. Because we test the algorithm on still images, we use a scheme which is bandwidth conservative to facilitate diversity; a technique proposed by the authors in [21], [28] which embeds information using quantization is put in practice. The method, which can also be applied for high capacity data hiding, is also well-documented in [27]. For reasons of space, we only summarize the approach and refer the reader to the relevant papers cited above for specifics.

The data embedding method is suited for the wavelet domain; we implement the method using the discrete wavelet transform (DWT). The Lth-level DWT of an image produces a sequence of $3L$ *detail* subimages corresponding to the horizontal, vertical, and diagonal details at each of the L resolution levels and a gross approximation at the coarsest resolution level. The different frequency orientations of the detail sub-images are represented with the variable o for which $o = 1, 2, 3$ corresponds to the horizontal, diagonal, and vertical details, respectively. The resolution level is denoted with the variable l, where a larger value of l corresponds to a coarser resolution level (i.e., larger scale). Thus, the detail coefficients for an Lth-level DWT of an image f are given by $f_{o,l}(m, n)$, where $o = 1, 2, 3, l = 1, 2, \ldots, L$, and (m, n) corresponds to the spatial location at the lth resolution. The gross approximation is represented with $f_{4,L}(m, n)$. Each subimage $f_{o,l}$ is comprised of pixels representing the coefficient values for various (m, n). The spatial locations are indexed from $(1, 1)$ to $(N_{fx}(l), N_{fy}(l))$ for an lth resolution subimage.

For the data-embedding algorithm, two basic parameters must be set: the quantization parameter Q and the coefficient selection key κ_c. The key κ_c is randomly generated and is used to select the exact locations in the wavelet domain in which to embed the watermark. If κ_c specifies that the mark is to be embedded at resolution l and location (m, n), then the following steps are performed.

1) The detail coefficients $f_{1,l}(m, n)$, $f_{2,l}(m, n)$ and $f_{3,l}(m, n)$ are sorted in ascending order. We denote these ordered coefficients by $f_{o1,l}(m, n)$, $f_{o2,l}(m, n)$, and $f_{o3,l}(m, n)$, where

$$f_{o1,l}(m, n) \leq f_{o2,l}(m, n) \leq f_{o3,l}(m, n) \tag{13}$$

such that $o1, o2, o3 \in \{1, 2, 3\}$ and $o1 \neq o2$, $o2 \neq o3$, and $o1 \neq o3$.

2) One watermark bit is embedded by modifying the median value of the detail coefficients at resolution l [i.e., $f_{o2,l}(m, n)$] at spatial location (m, n). To embed the watermark, we quantize $f_{o2,l}(m, n)$, as shown in Fig. 3. The range of values between $f_{o1,l}(m, n)$ and $f_{o3,l}(m, n)$ are divided into bins of width

$$\delta = \frac{f_{o3,l}(m,n) - f_{o1,l}(m,n)}{2Q - 1} \tag{14}$$

where Q is a key-specified quantization variable. To embed a watermark bit of value zero or one, $f_{o2,l}(m, n)$ is quantized to the nearest value specified in Fig. 3 by a dashed or bold line, respectively. The new watermarked coefficients are denoted $f^w_{o,l}(m, n)$.

To extract the embedded bit, the detail coefficients $\hat{f}^w_{o1,l}(m, n)$, $\hat{f}^w_{o2,l}(m, n)$, and $\hat{f}^w_{o3,l}(m, n)$ are, once again, sorted, and the value of the middle coefficient is used to assess the most probable bit embedded [27]. The watermark bit value is estimated from the relative position of $\hat{f}^w_{o2,l}(m, n)$. Using the same value of Q as for embedding, the watermark value is determined by finding the closest quantized value, which is specified by a dashed or solid vertical line in Fig. 3 to $\hat{f}^w_{o2,l}(m, n)$ and converting it to its associated binary number.

If this data embedding method is used naively (i.e., without attack characterization discussed in the overall robust reference watermarking method), the watermark is simply embedded several times; the most common extracted bit value is taken for the watermark estimate. If an equal number of ones and zeros is extracted, then a random guess is made to its value.

3) Algorithmic Overview: We present in Tables I and II the watermark embedding and extraction scenarios, respectively. Again, see [21], [27], and [28] for details.

B. Figures of Merit and Comparison

The focus of this paper has been on the incorporation of coefficient diversity and attack characterization to improve the reliability of robust watermarking systems. In order to demonstrate these advantages, we compare the performance of four watermark systems.

S1) The well-known spread spectrum technique by Cox *et al.*is presented in [6]. For compatibility with the proposed techniques, embedding is performed in the discrete wavelet domain. In addition, the original host image is not used for watermark detection. The method suggested by Piva *et al.* [29] is employed.

S2) The wavelet-based technique by Ohnishi and Matsui [17] hides information in the Haar wavelet domain. To embed a given watermark bit $w(i)$ into resolution l and spatial location (m, n), the difference between the minimum and maximum detail coefficients modulo 2 is forced to be equal to the watermark bit value.

S3) Our quantization-based data embedding method is discussed in Section IV-A2, in which the watermark is repeatedly embedded into the host signal but does not make use of the inherent diversity by characterizing the reliability of each extracted copy.

S4) The proposed robust reference watermarking method makes use of the data hiding technique of **S3)** but uses attack characterization to estimate the watermark optimally. Tables I and II summarize the algorithm.

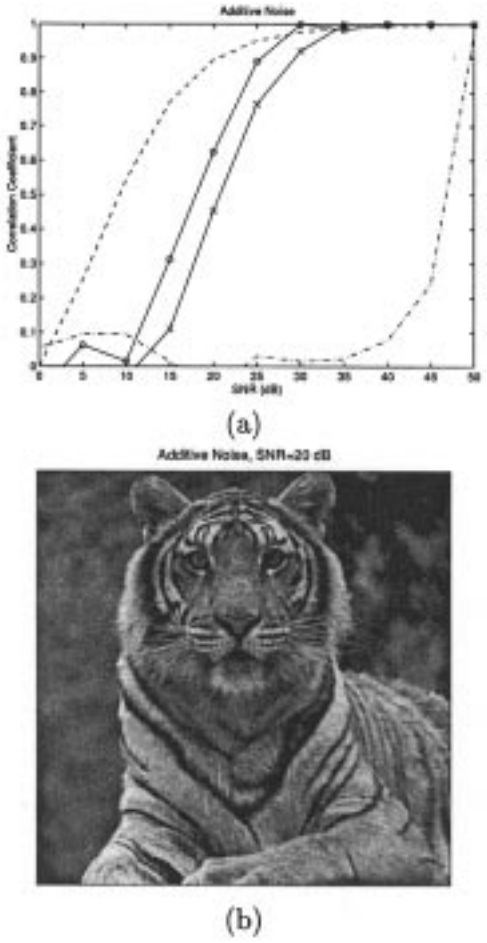

Fig. 5. Results for additive white Gaussian noise degradation. (a) Correlation coefficient versus SNR (the dashed and dash-dotted lines represent the results for the methods by Cox *et al.* and Ohnishi and Matsui, respectively; the solid lines with "x"s and "o"s correspond to watermark extraction by majority rule and weighted watermark estimation, respectively). (b) Degraded watermarked image at 20 dB SNR.

The correlation coefficient between the embedded and extracted watermarks given by

$$\rho(w, \hat{w}) = \frac{\sum_{i=1}^{N_w} w(i)\hat{w}(i)}{\sqrt{\sum_{i=1}^{N_w} w^2(i)} \sqrt{\sum_{i=1}^{N_w} \hat{w}^2(i)}} \tag{15}$$

where $w(i)$ and $\hat{w}(i)$ are the embedded and extracted watermark signals, respectively, and N_w is the length of the watermark that

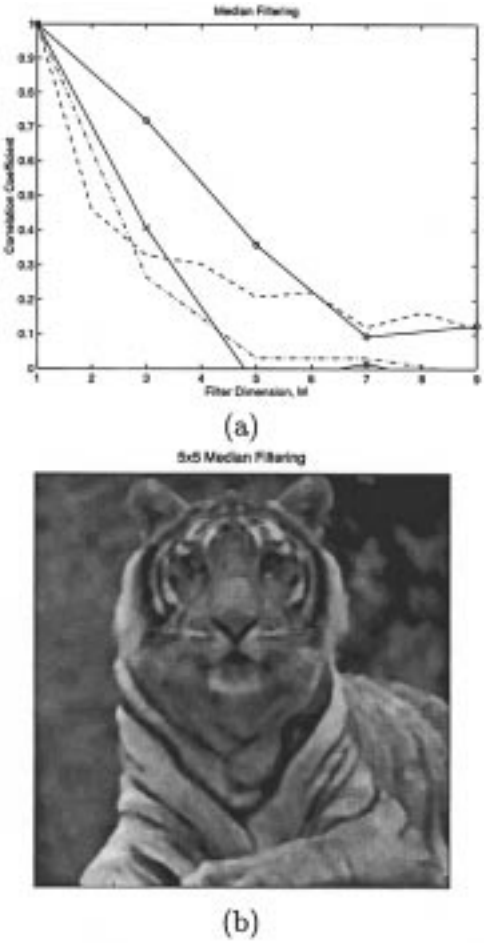

Fig. 6. Results for $F \times F$ median filtering degradation. (a) Correlation coefficient versus dimension of filter F (the dashed and dash-dotted lines represent the results for the methods by Cox *et al.* and Ohnishi and Matsui, respectively; the solid lines with "x"s and "o"s correspond to watermark extraction by majority rule and weighted watermark estimation, respectively). (b) Degraded watermarked image for 5×5 median filtering.

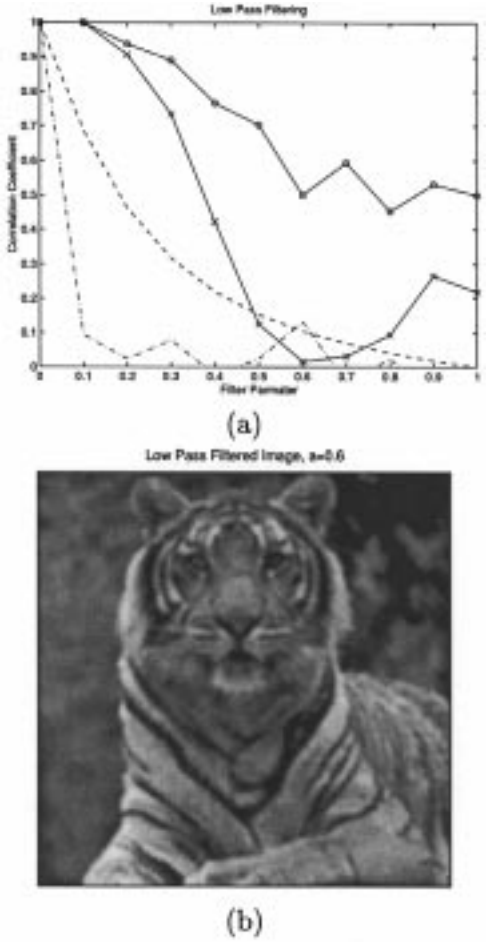

Fig. 7. Results for lowpass filtering degradation. (a) Correlation coefficient versus filter parameter μ_f (the dashed and dash-dotted lines represent the results for the methods by Cox *et al.*and Ohnishi and Matsui, respectively; the solid lines with "x"s and "o"s correspond to watermark extraction by majority rule and weighted watermark estimation, respectively). (b) Degraded watermarked image for $\mu_f = 0.6$.

is employed to assess the reliability of each system. A correlation threshold of 0.4, which corresponds to a false positive detection rate less than 10^{-10} for a 256 bit watermark, is used to detect the existence of the watermarks [27].

The difference in the relative performance of **S3)** and **S4)** provides an idea of the degree of effectiveness of the novel communication tool sets. A comparison with **S1)** and **S2)** gives an idea of the overall absolute performance improvement given by robust reference watermarking method.

The technique by Cox *et al.* [6] is one of the most influential algorithms proposed for robust watermarking. It is assessed to give an idea of the relative improvement that diversity provides over SS principles. The technique by Ohnishi and Matsui [17] is used for comparison because of its similarity to our proposed technique; both [17] and our data embedding approach place the watermark in a multiresolution domain and do so by imposing relative changes on the resulting detail coefficients without significant bandwidth consumption.

C. Simulation Results

In the implementation of our methods, we specifically make use of the Daubechies 10-point wavelet [30] for all simulations. Parameter values of $L = 4$ and $Q = 3$ were employed. The length of the reference watermark was set to the same value as the length of the robust watermark, which was 256 bits. In practice, any length of reference watermark can be used. The smaller the value of N_r, the poorer the estimate of p_{Ek}, but the more localized the assessment of the distortion. No user-specified parameters are required for the method in [17].The method in [6] was also implemented with the Daubechies 10-point wavelet, a randomly generated bipolar [i.e., $w(i) \in \{1, -1\}$] watermark for compatibility with our proposed techniques, and for scaling parameter $\alpha = 0.1$, as suggested in their paper.

We perform simulations on a 256×256 host image shown in Fig. 4(a). A 256-bit randomly generated equiprobable binary watermark was embedded. The reference watermark for our proposed technique was generated using the same distribution. The watermarked images using the method by Ohnishi *et al.* and the proposed reference watermarking algorithm are shown in Fig. 4(b)–(d), respectively. We next evaluate and compare the performance to various types of image distortions.

White Gaussian noise was added to the watermarked image to determine the robustness of the methods to stationary additive interference. Fig. 5 presents the results. Visible image

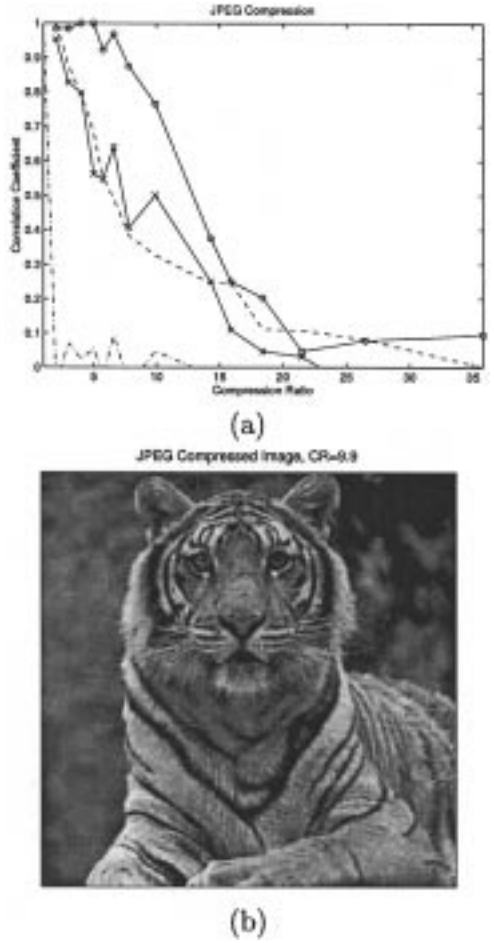

Fig. 8. Results for JPEG compression. (a) Correlation coefficient versus CR (the dashed and dash-dotted lines represent the results for the methods by Cox *et al.* and Ohnishi and Matsui, respectively; the solid lines with "x"s and "o"s correspond to watermark extraction by majority rule and weighted watermark estimation, respectively). (b) Degraded watermarked image for a CR of 9.9.

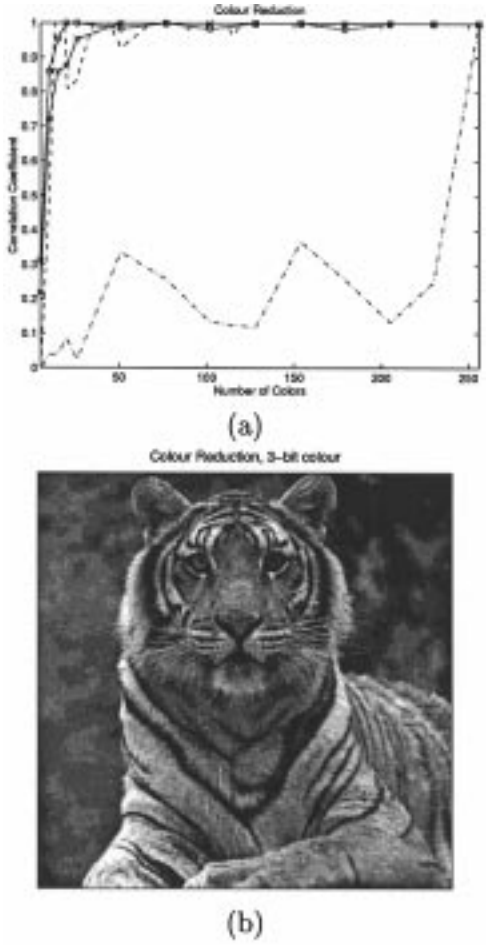

Fig. 9. Results for color reduction. (a) Correlation coefficient versus number of colors (the dashed and dash-dotted lines represent the results for the methods by Cox *et al.*and Ohnishi and Matsui, respectively; the solid lines with "x"s and "o"s correspond to watermark extraction by majority rule and weighted watermark estimation, respectively). (b) Degraded watermarked image for a color reduction to eight shades of gray (i.e., 3-bit image).

degradation was apparent around an SNR of 30 dB. The watermark, however, had a high correlation for even higher noise levels. Both **S3)** and **S4)** perform better than **S2)** for moderate to low SNRs. However, **S1)** demonstrates superior performance. Since the attack is not localized in space or frequency, it is not surprising that the performance of **S3)** and **S4)** is not better. Fig. 5(b) provides the watermarked image for a SNR of 20 dB, from which the watermark could still be reliably detected.

Similarly, an $F \times F$ median filter was applied on the watermarked image. **S3)** is comparable in performance to **S2)**, as shown in Fig. 6. However, **S4)** (robust reference watermarking) has significantly improved performance. **S1)** has varied performance.

The results for lowpass filtering with a radially symmetric blur of the form

$$h(m, n) = \mu_f^{\sqrt{m^2+n^2}} \Big/ K \tag{16}$$

where $K = \sum_{\forall (m,n)} h(m, n)$ are displayed in Fig. 7 for different values of μ_f. Use of diversity improves performance significantly. Similar results obtained for JPEG compression are presented in Fig. 8.

Excellent results for color reduction and cropping for our proposed approach are also evident. Fig. 9 shows that the proposed technique still detects the watermark with high precision even when the number of gray levels is reduced to 8; this corresponds to a 3-bit grayscale image. The use of reference watermarking offers little value in this situation since no significant performance is observed from **S3)** and **S4)**. The robustness stems from the nature of the data embedding process. Similarly, the results for image cropping are shown in Fig. 10. The localization of the watermark and attack characterization significantly improves the performance of the technique.

One of the most effective attacks against DCT-based watermarking has been proposed by Barnett and Pearson [31]. The distortion function known as the Laplacian removal (LR) attack operator makes two estimates (one positive denoted L_p and another negative denoted L_n) of the embedded watermark from the marked media using a series of Laplacian edge detections. The estimates $L_p - L_n$ are scaled with parameter α_{LR} and subtracted from the watermarked signal in order to attenuate the presence of the mark. Fig. 11 summarizes the results of this attack (as implemented in [31]) for varying values of α_{LR}.

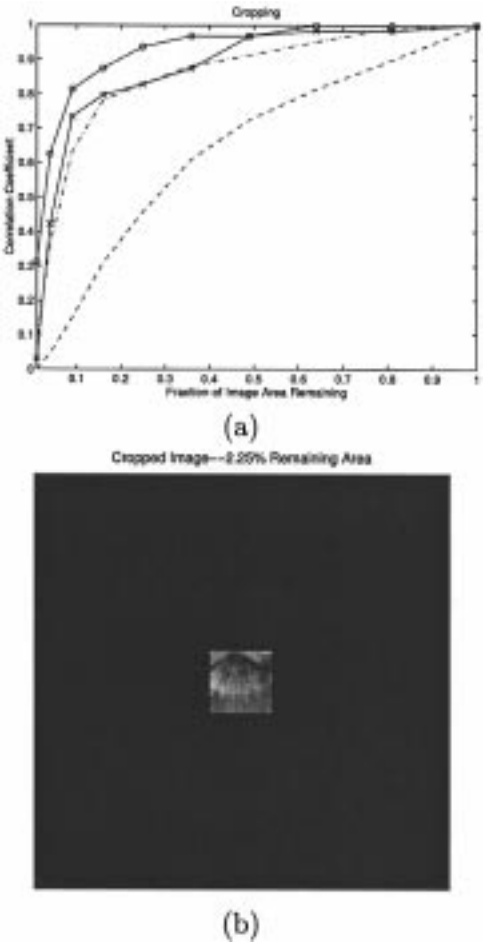

Fig. 10. Results for image cropping. (a) Correlation coefficient versus area of image remaining (the dashed and dash-dotted lines represent the results for the methods by Cox *et al.*and Ohnishi and Matsui, respectively; the solid lines with "x"s and "o"s correspond to watermark extraction by majority rule and weighted watermark estimation, respectively). (b) Cropped watermarked image for 2.25% area remaining.

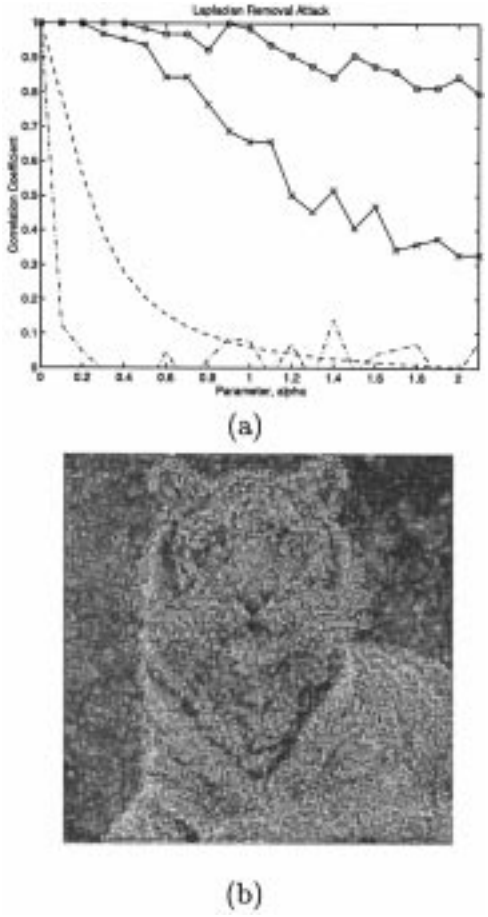

Fig. 11. Results for Laplacian attack operator: (a) Correlation coefficient versus attack parameter α_{LR} (the dashed and dash-dotted lines represent the results for the methods by Cox *et al.*and Ohnishi and Matsui, respectively; the solid lines with "x"s and "o"s correspond to watermark extraction by majority rule and weighted watermark estimation, respectively). (b) LR attacked image for $\alpha_{\mathrm{LR}} = 2.0$.

Fig. 11(b), in particular, demonstrates how the operator predominantly affects the high-frequency components of the watermarked image. The proposed technique **S4)** shows significant robustness to this attack. This can be explained since **S4)** embeds the watermark at all resolutions (including those near low frequencies) but, upon extraction, can isolate those components unaffected by the LR operator for reliable recovery. The success of **S3)** also demonstrates that quantization-based embedding strategies seem more robust to LR attacks than linear addition-based techniques, such as in **S1)**.

A more sophisticated version of the LR attack for DCT-based watermarking algorithms has been developed in [32]. Future work involves adapting this approach to DWT watermarking and assessing the robustness to the proposed scheme.

To summarize, it is evident from the results that use of diversity and channel estimation

- improves the absolute performance of a robust data embedding method for all degradations except additive white Gaussian noise;[3]
- significantly improves robustness against distortions that are localized in the watermark domain;
- has the least effect for degradations, such as additive noise and amplitude scaling, which are spread uniformly to the same degree to all coefficients in the watermark domain;
- broadens the class of distortions for which the watermark is resilient.

[3]For noise attacks, the proposed method is still detectable even when the fidelity of the attacked image is significantly reduced.

V. Final Remarks

In this paper, we apply the communication theory principles of diversity and channel estimation to improve the robustness of watermarking schemes. The use of coefficient diversity broadens the class of distortions for which the watermark is robust. In addition, we exploit the advantages of extracting the watermark *after* attacks by characterizing the distortions to more optimally process and detect the hidden information.

The nonstationary attack model introduced sheds light on the benefits of employing these novel tools and facilitates understanding into the advantages of using one watermark domain over another. As a result, we observe that techniques solely relying on embedding a stronger energy watermark may be supplemented with the proposed strategies in order to augment existing performance.

Our theoretical reflections are verified through the implementation of our ideas into the robust reference watermarking technique. Simulation results support our assertions and demonstrate the high degree of improvement achieved through application of our presented rationale.

APPENDIX A
ERROR BOUND DERIVATION

We prove (7) in this section, assuming $p_{Ek} \leq 0.5$ for $k = 1, 2, \ldots, M$. From the independence of e_k

$$E\{\mathcal{E}(\mathbf{e}(i))\} = \frac{\sum_{k=1}^{M} \log\left(\frac{1-p_{Ek}}{p_{Ek}}\right) E\{e_k(i)\}}{\sum_{k=1}^{M} \log\left(\frac{1-p_{Ek}}{p_{Ek}}\right)} \tag{17}$$

$$= \frac{\sum_{k=1}^{M} p_{Ek} \log\left(\frac{1-p_{Ek}}{p_{Ek}}\right)}{\sum_{k=1}^{M} \log\left(\frac{1-p_{Ek}}{p_{Ek}}\right)}. \tag{18}$$

Using the facts that $\log((1-p_{Ek})/p_{Ek}) \geq (1-p_{Ek})\log((1-p_{Ek})/p_{Ek})$ for $p_{Ek} \leq 0.5$

$$E\{\mathcal{E}(\mathbf{e}(i))\} \leq \frac{\sum_{k=1}^{M} p_{Ek} \log\left(\frac{1-p_{Ek}}{p_{Ek}}\right)}{\sum_{k=1}^{M} (1-p_{Ek}) \log\left(\frac{1-p_{Ek}}{p_{Ek}}\right)} \tag{19}$$

with equality if and only if $p_{Ek} = 0$ for all k. Using the log–sum inequality [23] to the denominator, we can show

$$E\{\mathcal{E}(\mathbf{e}(i))\} \leq \frac{\sum_{k=1}^{M} p_{Ek} \log\left(\frac{1-p_{Ek}}{p_{Ek}}\right)}{M(1-\overline{p}_E) \log\left(\frac{1-p_{Ek}}{p_{Ek}}\right)} \tag{20}$$

where $\overline{p}_E = (1/M)\sum_{k=1}^{M} p_{Ek}$ with equality if and only if $p_{Ek} = 0$ for all k.

The right-hand side of (20) can be expanded, rearranged, factored, and reduced to give (7).

APPENDIX B
ACCURATE WATERMARK RECOVERY

The watermark can be accurately recovered if $\mathcal{E}(\mathbf{e}) = 0$ since this implies that $e(i) = 0$ for all i. We prove the following.

If $p_{Ek} = 0$ for some $k \in \{1, 2, \ldots, M\}$, then $\mathcal{E}(\mathbf{e}(i)) = 0$.

Proof: Let $\mathcal{Z} = \{j | p_{Ej} = 0\}$. We assume in our analysis that $p_{Ej} = 0$ for some j (i.e., $\mathcal{Z} \neq \{\emptyset\}$). From (2) and (6)

$$\mathcal{E}(\mathbf{e}(i)) = \frac{\sum_{k=1}^{M} \log\left(\frac{1-p_{Ek}}{p_{Ek}}\right) e_k(i)}{\sum_{k=1}^{M} \log\left(\frac{1-p_{Ek}}{p_{Ek}}\right)} \tag{21}$$

$$= \frac{\sum_{k\notin\mathcal{Z}} \log\left(\frac{1-p_{Ek}}{p_{Ek}}\right) e_k(i) + \sum_{k\in\mathcal{Z}} \log\left(\frac{1-p_{Ek}}{p_{Ek}}\right) e_k(i)}{\sum_{k\notin\mathcal{Z}} \log\left(\frac{1-p_{Ek}}{p_{Ek}}\right) + \sum_{k\in\mathcal{Z}} \log\left(\frac{1-p_{Ek}}{p_{Ek}}\right)}. \tag{22}$$

It is easy to see that $\sum_{k\in\mathcal{Z}} \log((1-p_{Ek})/p_{Ek}) \gg \sum_{k\notin\mathcal{Z}} \log((1-p_{Ek})/p_{Ek})$ since the argument of $\log((1-p_{Ek})/p_{Ek})$ approaches infinity for $p_{Ek} \to 0$. Therefore

$$\mathcal{E}(\mathbf{e}) = \frac{\sum_{k\in\mathcal{Z}} \log\left(\frac{1-p_{Ek}}{p_{Ek}}\right) e_k}{\sum_{k\in\mathcal{Z}} \log\left(\frac{1-p_{Ek}}{p_{Ek}}\right)} = 0 \tag{23}$$

since $e_k(i) = 0$ for $k \in \mathcal{Z}$ because $w_k(i) = w(i)$ for $p_{Ek} = 0$.

REFERENCES

[1] R. G. van Schyndel, A. Z. Tirkel, and C. F. Osborne, "A digital watermark," in *Proc. IEEE Int. Conf. Image Process.*, vol. 2, 1994, pp. 86–90.

[2] A. Z. Tirkel, R. G. van Schyndel, and C. F. Osborne, "A two-dimensional digital watermark," in *Proc. Second Asian Conf. Comput. Vision*, vol. 2, Dec. 1995, pp. 504–508.

[3] A. Z. Tirkel, C. F. Osborne, and R. G. van Schyndel, "Image watermarking—A spread spectrum application," in *Proc. IEEE Int. Symp. Spread Spectrum Techn. Appl.*, vol. 2, 1996, pp. 785–789.

[4] R. G. van Schyndel, A. Z. Tirkel, I. D. Svalbe, T. E. Hall, and C. F. Osborne, "Algebraic construction of a new class of quasiorthogonal arrays for steganography," in *Proc. SPIE, Security Watermarking Multimedia Contents*, P. W. Wong and E. J. Delp, Eds., Jan. 1999, pp. 354–364.

[5] R. B. Wolfgang and E. J. Delp, "A watermark for digital images," in *Proc. IEEE Int. Conf. Image Process.*, vol. 3, 1996, pp. 219–222.

[6] I. J. Cox, J. Killian, T. Leighton, and T. Shamoon, "Secure spread spectrum watermarking for multimedia," *IEEE Trans. Image Processing*, vol. 6, pp. 1673–1687, Dec. 1997.

[7] X.-G. Xia, C. G. Boncelet, and G. R. Arce, "A multiresolution watermark for digital images," in *Proc. IEEE Int. Conf. Image Process.*, vol. 1, 1997, pp. 548–551.

[8] J. R. Smith and B. O. Comiskey, "Modulation and information hiding in images," in *Proc. First Int. Workshop Inform. Hiding*, ser. no. 1174 in Lecture Notes Comput. Sci., R. Anderson, Ed., May/June 1996, pp. 207–226.

[9] F. Hartung and B. Girod, "Digital watermarking of raw and compressed video," in *Proc. SPIE, Digital Compression Technol. Syst. Video Commun.*, vol. 2952, Oct. 1996, pp. 205–213.

[10] P. G. Flikkema, "Spread-spectrum techniques for wireless communications," *IEEE Signal Processing Mag.*, vol. 14, pp. 26–36, May 1997.

[11] B. Chen and G. W. Wornell, "Dither modulation: A new approach to digital watermarking and information embedding," in *Proc. SPIE, Security Watermarking Multimedia Contents*, vol. 3657, Jan. 1999.

[12] A. Z. Tirkel, C. F. Osborne, and T. E. Hall, "Image and watermark registration," *Signal Process.*, vol. 66, pp. 373–383, May 1998.

[13] J. J. K. ÓRuanaidh and T. Pun, "Rotation, scale and translation invariant digital image watermarking," in *Proc. IEEE Int. Conf. Image Process.*, vol. 1, 1997, pp. 536–539.

[14] M. Kutter, "Watermarking resisting to translation, rotation, and scaling," *Proc. SPIE*, 1998.

[15] M. K. Simon, J. K. Omura, R. A. Scholtz, and B. K. Levitt, *Spread Spectrum Communications Handbook*. Toronto, ON, Canada: McGraw-Hill, 1994.

[16] E. Koch and J. Zhao, "Toward robust and hidden image copyright labeling," in *Proc. Workshop Nonlinear Signal Image Process.*, I. Pitas, Ed., June 1995, pp. 452–455.

[17] J. Ohnishi and K. Matsui, "Embedding a seal into a picture under orthogonal wavelet transform," in *Proc. Int. Conf. Multimedia Comput. Syst.*, June 1996, pp. 514–521.

[18] G. W. Braudaway, "Protecting publicly-available images with an invisible image watermark," in *Proc. IEEE Int. Conf. Image Process.*, vol. 1, Oct. 1997, pp. 524–527.
[19] S. D. Servetto, C. I. Podilchuk, and K. Ramchandran, "Capacity issues in digital image watermarking," in *Proc. IEEE Int. Conf. Image Process.*, 1998.
[20] C. I. Podilchuk and W. Zeng, "Image-adaptive watermarking using visual models," *IEEE J. Select. Areas Commun.*, vol. 16, pp. 525–539, May 1998.
[21] D. Kundur and D. Hatzinakos, "Improved robust watermarking through attack characterization," *Opt. Express*, vol. 3, pp. 485–490, Dec. 7, 1998.
[22] S. Haykin, *Communication Systems*, third ed. Toronto, ON, Canada: Wiley, 1994.
[23] T. Cover and J. Thomas, *Elements of Information Theory*. Toronto, ON, Canada: Wiley, 1991.
[24] D. Kundur and D. Hatzinakos, "Digital watermarking for telltale tamper-proofing and authentication," in *Proc. IEEE—Special Issue on Identification and Protection of Multimedia Information*, vol. 87, July 1999, pp. 1167–1180.
[25] ——, "Toward a telltale watermarking technique for tamper-proofing," in *Proc. IEEE Int. Conf. Image Process.*, vol. 2, 1998, pp. 409–413.
[26] ——, "Semi-blind image restoration based on telltale watermarking," in *Proc. 32nd Asilomar Conf. Signals, Syst., Comput.*, vol. 2, 1998, pp. 933–937.
[27] D. Kundur, "Multiresolution digital watermarking: Algorithms and implications for multimedia signals," Ph.D. dissertation, Univ. Toronto, Toronto, ON, Canada, Aug. 1999.
[28] D. Kundur and D. Hatzinakos, "Digital watermarking using multiresolution wavelet decomposition," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, vol. 5, 1998, pp. 2969–2972.
[29] A. Piva, M. Barni, F. Bartolini, and V. Cappellini, "DCT-based watermark recovering without resorting to the uncorrupted original image," in *Proc. IEEE Int. Conf. Image Process.*, vol. 1, 1997, pp. 520–523.
[30] I. Daubechies, "Orthonormal bases of compactly supported wavelets," *Commun. Pure Appl. Math.*, vol. 41, pp. 909–996, Nov. 1988.
[31] R. Barnett and D. E. Pearson, "Attack operators for digitally watermarked images," *IEE Proc. Inst. Elect. Eng., Visual Image Signal Process.*, vol. 145, Aug. 1998.
[32] R. Barnett and D. E. Pearson, "Frequency mode lr attack operator for digitally watermarked images," *Electron. Lett.*, vol. 34, Sept. 1998.

Deepa Kundur (M'99) received the B.A.Sc., M.A.Sc., and Ph.D. degrees from Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON, Canada.

She is an Assistant Professor with the Edward S. Rogers, Sr. Department of Electrical and Computer Engineering, University of Toronto. She holds the title of Bell Canada Junior Chair-holder in Multimedia and is also an Associate of the Nortel Institute for Telecommunications. Her research interests span the areas of multimedia security, data hiding and covert communications, content-based multimedia processing, and nonlinear and adaptive communication algorithms.

Dr. Kundur is a Member of the IEEE Communications and Signal Processing Societies and the Professional Engineers of Ontario (PEO).

Dimitrios Hatzinakos (SM'98) received the Diploma degree from the University of Thessaloniki, Thessaloniki, Greece, in 1983, the M.A.Sc. degree from the University of Ottawa, Ottawa, ON, Canada, in 1986, and the Ph.D. degree from Northeastern University, Boston, MA, in 1990, all in electrical engineering.

In September 1990, he joined the Department of Electrical and Computer Engineering, University of Toronto, where now he holds the rank of Professor with tenure. He has also served as Chair of the Communications Group of the Department since July 1, 1999. His research interests are in the areas of digital communications and signal processing with applications to wireless communications, image processing, and multimedia. He has organized and taught many short courses on modern signal processing frameworks and applications devoted to continuing engineering education and given numerous seminars in the area of blind signal deconvolution. He is author/co-author of more than 100 papers in technical journals and conference proceedings, and he has contributed to six books in his areas of interest. His experience includes consulting through Electrical Engineering Consociates Ltd. and contracts with United Signals and Systems Inc., Burns and Fry Ltd., Pipetronix Ltd., Defense Research Establishment Ottawa (DREO), Vaytek Inc., Nortel Networks, and Vivosonic Inc.

Dr. Hatzinakos has been an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING since July 1998. He was also the Guest Editor for the Special Issue of *Signal Processing* on "Signal Processing Technologies for Short Burst Wireless Communications," which appeared in October 2000. He was a member of the IEEE Statistical Signal and Array Processing Technical Committee (SSAP) from 1992 until 1995 and Technical Program Co-Chair of the Fifth Workshop on Higher Order Statistics in July 1997. He is a member of EURASIP, the Professional Engineers of Ontario (PEO), and the Technical Chamber of Greece.