Statistical Invisibility for Collusion-Resistant Digital Video Watermarking

Karen Su, *Student Member, IEEE*, Deepa Kundur, *Senior Member, IEEE*, and Dimitrios Hatzinakos, *Senior Member, IEEE*

***Abstract*—In this paper, we present a theoretical framework for the linear collusion analysis of watermarked digital video sequences, and derive a new theorem equating a definition of statistical invisibility, collusion-resistance, and two practical watermark design rules. The proposed framework is simple and intuitive; the basic processing unit is the video frame and we consider second-order statistical descriptions of their temporal inter-relationships. Within this analytical setup, we define the linear frame collusion attack, the analytic notion of a statistically invisible video watermark, and show that the latter is an effective counterattack against the former. Finally, to show how the theoretical results detailed in this paper can easily be applied to the construction of collusion-resistant video watermarks, we encapsulate the analysis into two practical video watermark design rules that play a key role in the subsequent development of a novel collusion-resistant video watermarking algorithm discussed in a companion paper.**

***Index Terms*—Linear collusion, robust imperceptible digital video watermarking, statistical invisibility, watermarking attack.**

I. Introduction

THE PURPOSE of a digital watermark is to embed auxiliary information into a host digital signal by imposing imperceptible signal changes. The technology has potential for covert communications, fingerprinting, signal tagging, and media copyright control among other applications. Current work has led to a number of practical algorithms and a first wave of commercial products using watermarks to provide added-value to content. Since the first papers on digital watermarking (focusing on images) appeared in the early 1990s [2]–[4], the publication rate has approximately doubled every year.

In this work, we concentrate on the challenge of *digital video watermarking*. Much of the academic and industrial interest in digital video watermarking has centered on the design of a copyright protection system for MPEG-2 coded video distributed on Digital Versatile Disk (DVDs) [5]. A video watermarking system had been designed by the Galaxy Group to complement the existing content scrambling system (CSS) that is part of the DVD standard; the technology is now called WaterCast and is being applied in the automatic monitoring of digital video broadcasts. The growing appeal of video watermarking for more general applications is evidenced by the number of proposals for digital TV transmission [6], satellite broadcast monitoring [7], video on demand distribution [8], and authenticating video surveillance for use as legal evidence [9]. The flexibility of watermarking concepts for use with new data types has also been demonstrated through preliminary work with MPEG-4 video objects and parameters [10], [11].

Although in their raw form video streams are sequences of image frames, the complexity and flexibility of the solution space for the watermarking of video is significantly greater than that for images due to the presence of the time dimension. Video watermarking is distinct from image watermarking, in part, because there is more data available to both the attacker as well as to the watermarker. This additional volume allows the payload to be more redundantly and reliably embedded possibly by exploiting more sophisticated temporal masking characteristics of human perception. In the same vein, the attacker has the liberty to make greater use of correlations in the signal volume to devise more effective watermark estimation or removal attacks. One important class of such attacks which is the focus of this paper is known as *multiple frame linear collusion*.

Collusion occurs when collections of video frames are analyzed or combined with the ultimate goal of producing a mark-free copy of the original. The frames may form a temporally continuous subsequence, or come from greatly varied parts of the video. The key idea is the exploitation of temporal redundancy, either of the host video or the watermark, to estimate the redundant component. In the case of multiple frame linear collusion, distinct watermarked video frames are scaled and added to form a resulting frame that provides an estimate of either the host or the watermark. If the scaling is the same for each frame, then the overall attack consists of average a set of watermarked video frames. Intuitively, this operation has the effect of amplifying the component of the watermark or host that is the same from frame to frame and attenuating the part that differs. To date, the collusion attack has not been well studied, most probably because of the research focus on watermarking still images. However, its growing importance is evidenced by the publication of recent papers in the area [12]–[14].

The objectives of this paper are two-fold:

1) To provide a theoretical framework for the analysis of collusion-resistant video watermarking algorithms. We intend our framework to be general and compatible with a broad class of existing video watermarking schemes.

Manuscript received October 24, 2002; revised April 30, 2003. This work was supported by the Natural Sciences and Engineering Research Council (NSERC) and by the Communications and Information Technology Ontario (CITO). The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Anna Hac.

K. Su was with the University of Toronto, Toronto, ON, Canada. She is now with the Laboratory for Communication Engineering, University of Cambridge, Cambridge, U.K.

D. Kundur was with the University of Toronto, Toronto, ON, Canada. She is now with the Electrical Engineering Department, Texas A&M University, College Station, TX 77843-3128 USA (e-mail: deepa@ee.tamu.edu).

D. Hatzinakos is with the Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON, Canada.

Digital Object Identifier 10.1109/TMM.2004.840617

2) To devise collusion-resistant video watermark design strategies inspired by our theoretical observations. A bridge between theory and practical applicability is of interest in this work.

The collusion problem was first addressed by Swanson *et al.* [15]. However, the proposed solution involves a two-level hierarchy of transforms and is considered to be highly complex. One of the contributions in this work is to provide a framework in which lower complexity collusion-resistant video watermarks can be devised. Recent work by Trappe *et al.* [14] has focused on collusion-resistant digital fingerprinting that can identify colludes; their work makes use of effective anti-collusion codes for CDMA type watermarking using theory of combinatorial designs. The work presented in this paper formulates the problem using statistical notions rather than coding theory and thus involves a framework that is complementary to that proposed in [14].

We begin in Section II by presenting a mathematical framework for the analysis of collusion attacks. Then, in Section III, we define multiple frame linear collusion, and the notion of *statistical invisibility* for watermarking which we prove theoretically to be an effective counterattack to collusion. We close the paper with a discussion of the theoretical results and their implications to existing video watermarking algorithms in Section IV, followed by concluding remarks in Section V.

II. Nomenclature

A robust watermarking system is comprised of two basic components: the *embedder* that inserts the watermark and, the *detector* that detects or extracts the mark. The inputs to the embedder are the host video sequence, a key K, and a binary data message vector V_i. The key is a sequence of bits encapsulating all secret parameters and components of the watermarking system, i.e., block sizes for block-based algorithms or seeds for random number generators. The payload capacity (in bits) of the system is determined by the length of the binary vector V. Fig. 1(a) summarizes the set-up.

The original or host video sequence is denoted $U_k(m,n)$, where k is a time or frame number index set, m and n are row and column indices, respectively. The domain of a variable x is denoted $\mathrm{D}(x)$; for instance, $\mathrm{D}(k) = \{1,\ldots,K_U\}$, $\mathrm{D}(m) = \{1,\ldots,M_U\}$, and $\mathrm{D}(n) = \{1,\ldots,N_U\}$, i.e., the video sequence consists of K_U frames, each of size $M_U \times N_U$ pixels.

The embedded watermark signal $W_k(m,n)$ is defined over the same domain as the host $U_k(m,n)$. The embedder produces a watermarked video $X_k(m,n)$ sequence obtained by linear combination of the watermark with the host data

$$X_k(m,n) = U_k(m,n) + \alpha_k(m,n) \cdot W_k(m,n) \qquad (1)$$

where $\alpha_k(m,n)$ represents a general scaling factor (i.e., local or global). No matter how the watermark is actually embedded, all data hiding procedures can be expressed in this form by defining the watermark as the difference between the watermarked and host signals (and setting $\alpha_k(m,n)$ to 1 or some other appropriate function based on the details of the embedding algorithm). In general, the signal $\alpha_k(m,n)W_k(m,n)$ is dependent on the message V_i, the key K, and the host $U_k(m,n)$ for which the specific relation is determined by the particular embedding algorithm. For the remainder of this paper, we focus on $W_k(m,n)$ and $\alpha_k(m,n)$ to characterize the affects of the watermark embedding process on the host and only implicitly consider their possible dependence on V_i, K, and $U_k(m,n)$.

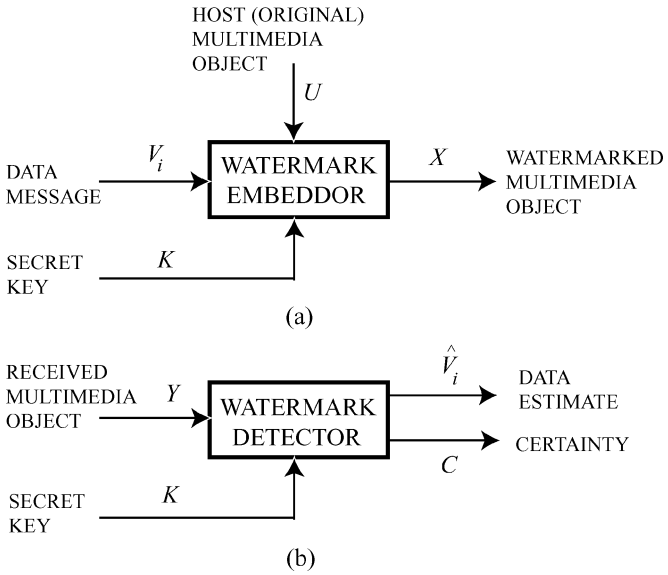

Fig. 1. Overview of the basic robust video watermarking scenario. (a) Watermark embedding entails the modification of the host signal U using a secret key K to embed a payload V_i to produce a watermarked video sequence X and (b) watermark detection involves extraction of an estimate of the payload $\hat{V}_i$ from a possibly attacked and watermarked video sequence Y using the secret key K (also known at the receiver). A measure of the certainty of the payload estimate C may also be available.

The blind watermark detection process in our formulation is displayed in Fig. 1(b). The inputs to the detector are the secret key K (also used at the embedder) and a possibly degraded watermarked signal $Y_k(m,n)$. The watermark detector produces two outputs, C and $\hat{V}_i$, indicating the certainty of the watermark's presence and the extracted data message respectively. C is a value between 0 (watermark absent) and 1 (watermark present) indicating the degree to which the watermark is detected. $\hat{V}_i$ is a binary vector of bits that must be compared to the embedded message V_i to measure the bit error rate of the system.

Overall, as a general practical principle, robust imperceptible video watermarks are designed such that

- the watermark is present and detectable in every frame. A common measure used in practice involves the correlation coefficient between the watermarked frame and the scaled watermark $\rho(X_k, \alpha_k W_k)$ such that

$$\rho(X_k, \alpha_k W_k) > 1 - \tilde{\gamma}. \qquad (2)$$

For robustness to standard image processing on each frame, we ideally would like $0 < \tilde{\gamma} < 0.5$. That is, the watermark detection measure in the ideal case of no tampering is between 0 and 0.5. The authors believe that this broad assumption range encompasses the necessary constraints to guarantee robustness in the presence of incidental distortions with a reasonably small false positive detection rate. For the situation in which α_k is a constant

for all k, this restriction is only applied to the correlation between the watermarked frame and the watermark $\rho(X_k, W_k)$ which is commonly found in the academic literature.

- the watermark is imperceptible in every frame. The power of the watermark $\mathbf{E}[(X_k - U_k)^2]$ is often restricted by existing methods as follows:

$$\mathbf{E}[(X_k - U_k)^2] < \xi. \tag{3}$$

Note that in practice for imperceptibility we would like ξ to be reasonably small. For instance it has been found experimentally that for a 8-bit grayscale image, a typical value for ξ for each pixel can range between 2 and 10.

These restrictions to every frame are implicitly assumed for the practical success of the watermark scheme in the remainder of the paper.

III. Multiple Frame Linear Collusion

A. Linear Collusion

In this paper, we say that Type I collusion arises when large numbers of "visually dissimilar" video frames are marked via linear combination with a fixed watermark pattern. For instance, in [16], it is shown that statistical processing can be used to recover a good approximation of a nonframe dependent watermark pattern. This is commonly found in many existing video watermarks [11], [7], [17].[1] Type II collusion arises when large numbers of "visually similar" frames are marked via linear combination with independent watermark patterns. An example of such an attack is frame averaging. This case is relevant, for instance, to video watermarks that use different two-dimensional pseudo-noise (PN) sequences to mark each frame [18]. We next formulate our framework and present some propositions useful for video watermark design resistant to linear collusion.

Definition 1: Given two random variables A and B with finite means and variances, we say that A *is an* ϵ-optimal Mean Square Error (MSE) estimate of B if and only if

$$\mathbf{E}[(\hat{A} - B)^2] < \epsilon \tag{4}$$

where $\mathbf{E}$ is the expectation operator, $\hat{A} = \sqrt{\mathsf{var}(B)/\mathsf{var}(A)}(A - \mathbf{E}A) + \mathbf{E}B$, and $\mathsf{var}(\cdot)$ is the variance of the argument random variable.

In Definition 1, we normalize random variable A to reflect the same mean and variance as B; thus, we also require $0 < var(A), var(B), |\mathbf{E}A|, |\mathbf{E}B| < \infty$ as stated. The next definition specifies the class of attacks considered in this work.

Definition 2: Given a set of watermarked video frames $X_k = U_k + \alpha_k W_k$, $k = 1, \ldots, n$, linear collusion is the process of forming a linear combination of the frames

$$\begin{aligned}\overline{X} &= \sum_{k=1}^{n} \beta_k X_k \\ &= \sum_{k=1}^{n} \beta_k U_k + \sum_{k=1}^{n} \beta_k \alpha_k W_k\end{aligned} \tag{5}$$

such that $\overline{X}$ provides an ϵ -optimal MSE estimate of

a) the watermark component $\overline{W} = \sum_{k=1}^{n} \beta_k \alpha_k W_k$, where we require $\rho(W_k, W_l) > 1 - \delta$ for $k, l = 1, 2, \ldots, n$ and some $0 < \delta \ll 1$,
   or
b) the host component $\overline{U} = \sum_{k=1}^{n} \beta_k U_k$, where we require $\rho(U_k, U_l) > 1 - \delta$ for $k, l = 1, 2, \ldots, n$ and some $0 < \delta \ll 1$,
   for which $\rho(A, B) = \mathsf{cov}(A, B)/\sqrt{\mathsf{var}(A)\mathsf{var}(B)}$ is the statistical correlation coefficient between random variables A and B (where $\mathsf{cov}(\cdot, \cdot)$ denotes covariance), and $\tilde{\epsilon} = \epsilon/(2(\mu_2 - \mu^2))$ (where we let $\mathbf{E}B = \mu$, $\mathbf{E}B^2 = \mu_2$).

In the case of a), we refer to the attack as Type I collusion; in the case of b), as Type II collusion.

In the subsequent analysis of this paper we let $\overline{X} = \sum_{k=1}^{n} \beta_k X_k$, $\overline{U} = \sum_{k=1}^{n} \beta_k U_k$, and $\overline{W} = \sum_{k=1}^{n} \beta_k \alpha_k W_k$ exclusively. The conditions in Definition 2 on $\rho(W_k, W_l)$ and $\rho(U_k, U_l)$ are required for the watermark (Type I) or host (Type II) to be "well-defined" notions; only when this occurs do we consider $\overline{X}$ to be a valuable information-bearing estimate. The values of ϵ, δ are considered to be "small" (i.e., $0 < \epsilon, \delta \ll 1$) to provide a reasonable estimate of the watermark or host. In addition, our definition of collusion involves the addition of *spatially synchronized* (i.e., unshifted, registered) raw format frames. Spatial correlations are not accounted for in the attack formulation. Thus, for notational simplicity we drop reference to the pixel values (m, n) and, subsequently, formulate our problem using random variables instead of random fields.

Definition 2 encapsulates frame averaging attacks by setting $\beta_k = 1/n$, as well as more sophisticated linear temporal filters by allowing β_k to take on arbitrary values. It is also general enough to encompass combining arbitrary sets of frames that are not necessarily in a temporally continuous sequence relative to the video.

The remainder of this section presents two useful propositions.

Proposition 1: Given two random variables A and B with finite means and variances, A is an ϵ-optimal MSE estimate of B if and only if $\rho(A, B) > 1 - \tilde{\epsilon}$. That is,

$$\begin{aligned}&\mathbf{E}[(\hat{A} - B)^2] < \epsilon \\ &\qquad\Longleftrightarrow \\ &\rho(A, B) > 1 - \tilde{\epsilon}\end{aligned} \tag{6}$$

where as discussed before, $\rho(A, B) = \mathsf{cov}(A, B)/\sqrt{\mathsf{var}(A)\mathsf{var}(B)}$ is the statistical correlation coefficient between random variables A and B and $\tilde{\epsilon} = \epsilon/(2\mathsf{var}(B))$.

Proof: See Appendix A.

Through Proposition 1 we can see the relationship between the statistical correlation and the MSE estimate; the smaller the value of ϵ, the greater the lower bound on correlation. We next present a necessary condition for each type of linear collusion.

Proposition 2: Assuming that the scaled watermarks $\alpha_k W_k$ are independent of the host frames U_k , then a necessary condition for each of the two forms of linear collusion described in Definition 2 is given by

$$\rho(\overline{X}, \overline{U}) < \sqrt{\tilde{\epsilon}(2 - \tilde{\epsilon})} \text{ (Type I)} \tag{7}$$

$$\rho(\overline{X}, \overline{U}) > 1 - \tilde{\epsilon} \text{ (Type II)} \tag{8}$$

[1]Note: The watermarks do not have to be embedded in the spatial domain, the analysis presented here is relevant to watermark and host signals considered in their embedding domain.

where $\overline{X} = \sum_{k=1}^{n} \beta_k X_k$, and $\overline{U} = \sum_{k=1}^{n} \beta_k U_k$, i.e.,

$$\rho(\overline{X}, \overline{U}) > \sqrt{\tilde{\epsilon}(2-\tilde{\epsilon})} \Longrightarrow \text{Type I linear collusion is not possible, and} \quad (9)$$

$$\rho(\overline{X}, \overline{U}) < 1 - \tilde{\epsilon} \Longrightarrow \text{Type II linear collusion is not possible.} \quad (10)$$

Proof: See Appendix B.

Proposition 2 only gives a necessary condition for Type I linear collusion since we also require the watermark to be a well-defined quantity that we mathematically express as $\rho(W_k, W_l) > 1 - \delta, \forall k, l = \{1, 2, \ldots, n\}$. For instance, the condition of (7) is not sufficient if the watermark pattern is independent in each frame, because Type I collusion can be evaded; the linear combination of the watermarked frames will not enhance the watermark. More specifically, if independent watermark patterns are used to mark each of the video frames, and if the necessary condition $\rho(\overline{X}, \overline{U}) < \sqrt{\tilde{\epsilon}(2-\tilde{\epsilon})}$ is met, then

$$\rho(\overline{X}, \overline{W}) > 1 - \tilde{\epsilon}. \quad (11)$$

However, no information about the watermarks is revealed according to Lemma 1 which gives

$$\rho^2(\overline{X}, \overline{W}) > (1-\tilde{\epsilon})^2 = \sum_{k=1}^{n} \rho^2(\overline{X}, \beta_k \alpha_k W_k) \quad (12)$$

where all of the watermark terms on the right hand side of (12) are unknown uncorrelated random variables, and, thus, the watermark is not a well-defined and estimated notion.

When considering Type II collusion, we observe that if the same watermark $W_k = W, \forall k = \{1, 2, \ldots, n\}$ is used to mark each of the video frames that obey $U_k = U, \forall k = \{1, 2, \ldots, n\}$, and even if the necessary condition $\rho(\overline{X}, \overline{U}) > 1 - \tilde{\epsilon}$ is met, the colluded host will still include a scaled version of the watermark pattern. If we consider the case for which $\beta_k = 1/n$:

$$\overline{X} = \overline{U} + \overline{W} = U + W \sum_{k=1}^{n} \frac{\alpha_k}{n}. \quad (13)$$

Therefore the collusion attempt can fail to sufficiently separate the host from the watermark and a mark-free copy cannot be obtained. As in the case of Type I collusion, Proposition 2 gives only a necessary condition for Type II collusion. The condition is not sufficient since proper design of the watermark can, as shown, provide protection against these attacks.

Through the subsequent analysis, it will become evident that it is possible to achieve protection from Type I and II collusion by imposing certain design criteria on the watermark. These criteria are formalized mathematically in the next section as *statistical invisibility*.

B. Statistical Invisibility

Definition 3: Given a sequence of host video frames U_k, $k = 1, \ldots, n$ and watermarked video frames $X_k = U_k + \alpha_k W_k$, we say that the video watermark W_k *is* statistically invisible if and only if the correlation coefficient between any two host frames a and b is equal to that between the two corresponding watermarked frames, i.e., $\rho(U_a, U_b) = \rho(X_a, X_b) \forall a, b \in \{1, \ldots, n\}$.

We refer to this property as statistical invisibility since an attacker analyzing the video sequence in a frame-by-frame manner does not observe any statistical difference between the host and watermarked sequences. For the remainder of the analysis the following assumptions are made about the statistics of the watermark, host, and scaling factors.

- A1) The video frames share a common finite mean and variance (average power), i.e., $\mathbf{E}U_k = \mu_U$ and $var(U_k) = \sigma_U^2$.
- A2) The watermarks W_k are zero-mean, i.e., $\mathbf{E}W_k = 0$ and share a common nonzero finite variance $\sigma_W^2 > 0$. Consequently, $\mathbf{E}X_k = \mathbf{E}U_k$.
- A3) The scaling factors α_k share a finite second moment $\mathbf{E}\alpha^2$.
- A4) The watermarks W_k, scaling factors α_k and host U_k are independent of one another.

The following proposition gives an alternate representation of the statistical invisibility criterion which is useful in developing a watermark design principle.

Proposition 3: Under assumptions (A1)–(A4)

$$\rho(X_a, X_b) = \rho(U_a, U_b) \forall a, b \in \{1, 2, \ldots, n\} \text{ (statistical invisibility)} \quad (14)$$

$$\Longleftrightarrow$$

$$\rho(U_a, U_b) = \frac{\mathbf{E}\alpha_a \alpha_b}{\mathbf{E}\alpha^2} \rho(W_a, W_b) \forall a, b \in \{1, 2, \ldots, n\}. \quad (15)$$

Proof: See Appendix C.

Another interpretation is provided in Proposition 4.

Proposition 4: Under assumptions (A1)–(A4)

$$\rho(U_a, U_b) = \frac{\mathbf{E}\alpha_a \alpha_b}{\mathbf{E}\alpha^2} \rho(W_a, W_b) \forall a, b \in \{1, 2, \ldots, n\}$$

$$\Longleftrightarrow$$

$$\rho(\overline{X}, \overline{U}) = \rho(X_a, U_a) \forall a \in \{1, 2, \ldots, n\}. \quad (16)$$

Proof: See Appendix D.

Next, we present a sufficient condition for robustness to linear collusion; this expression will enable us to show a direct relationship between statistical invisibility and collusion resistance.

Proposition 5: Under assumptions (A2) and (A4)

$$\rho(\overline{X}, \overline{U}) = \rho(X_a, U_a) \forall a \in \{1, 2, \ldots, n\}$$

$$\Longrightarrow$$

$$\sqrt{\tilde{\epsilon}(2-\tilde{\epsilon})} < \rho(\overline{X}, \overline{U}) < 1 - \tilde{\epsilon} \quad (17)$$

which provides sufficient conditions for the resistance to Type I and Type II linear collusion attacks.

Proof: See Appendix E.

Therefore, $\rho(\overline{X},\overline{U}) = \rho(X_a, U_a)$ ensures the inability of an attacker to obtain an ϵ-optimal MSE estimate of the watermark (Type I) or host (Type II) for $\tilde{\epsilon} \ll 1$[2], and is sufficient for robustness to linear collusion.

Having assembled all of the necessary ingredients, we state our theorem on the relationship between statistical invisibility and multiple frame collusion.

Theorem 1: Under assumptions (A1)–(A4), the following statements are equivalent:

1) $\rho(X_a, X_b) = \rho(U_a, U_b) \forall a, b \in \{1, 2, \ldots, n\}$,
2) $\rho(U_a, U_b) = (\mathbf{E}\alpha_a\alpha_b/\mathbf{E}\alpha^2)\rho(W_a, W_b) \forall a, b \in \{1, 2, \ldots, n\}$, and
3) $\rho(\overline{X},\overline{U}) = \rho(X_a, U_a) \forall a \in \{1, 2, \ldots, n\}$.

Property (1) describes the statistical invisibility condition; property (2) defines a host-dependent watermark design criterion; and property (3) ensures that a watermark satisfying these criteria exhibits statistical resistance to Type I and Type II linear collusion attacks.

Proof: Proposition 3 shows the equivalence between properties (1) and (2), and Proposition 4 shows the equivalence between properties (2) and (3). Proposition 5 ensures that in practice all three properties are sufficient for robustness to linear collusion. ∎

C. Video Watermark Design Principles

In this section we make use of our analysis results to determine two video watermark design principles robust to linear collusion. Proposition 5 states that under assumptions (A2) and (A4), $\rho(X_a, U_a) = \rho(X_b, U_b)$ for all $a, b \in \{1, 2, \ldots, n\}$ guarantees collusion resistance in practice. Furthermore, we assert:

Proposition 6: Under assumptions (A2) and (A4)

$$\rho(X_a, U_a) = \rho(X_b, U_b) \tag{18}$$

$$\Longleftrightarrow$$

$$\frac{\mathbf{E}\alpha_b^2}{\mathbf{E}\alpha_a^2} = \frac{\mathsf{var}(U_b)}{\mathsf{var}(U_a)}. \tag{19}$$

Proof: See Appendix F.

In other words, if the energy of the basic watermark signals W_k embedded into each frame is kept constant over the video sequence, modulating the per-frame embedding strengths $\mathbf{E}\alpha_k^2$ proportionally to the variances of the host frames ensures that the condition in (18) is met. The idea of watermark strength adaptation according to some function of the image variance, both at global and local scales, is a popular rule of thumb used by many image watermarks. As it turns out, it also plays a role in maintaining statistical invisibility and thus becomes our first video watermark design principle:

The second moment of the watermark scaling factors should be adapted proportionally to the variance of the host video frames, i.e.,

$$\frac{\mathbf{E}\alpha_b^2}{\mathbf{E}\alpha_a^2} = \frac{var(U_b)}{var(U_a)}.$$

[2]The reader should note that for $\bar{\epsilon} \leq 0.29$, it is guaranteed that $\sqrt{\bar{\epsilon}(2-\bar{\epsilon})} < 1 - \bar{\epsilon}$.

Under assumptions (A1)–(A4), property (2) of Theorem 1 guarantees robustness to linear collusion in practice, and we can identify a second watermark design principle:

The correlation of the watermarks embedded into each pair of video frames should be matched to the correlation of the host frames themselves, i.e.,

$$\rho(U_a, U_b) = \frac{\mathbf{E}\alpha_a\alpha_b}{\mathbf{E}\alpha^2}\rho(W_a, W_b) \forall a, b \in \{1, 2, \ldots, n\}.$$

A necessary and sufficient condition for the statistical invisibility of a video watermark is given by property (2). This implies that in order for the watermarks embedded into two video frames to be statistically invisible, their correlation must differ from that of the host frames only by a scaling factor. In the trivial case, where a constant strength watermark is embedded into each frame, i.e., $\alpha_a = \alpha_b = \overline{\alpha}$, we require that $\rho(W_a, W_b) = \rho(U_a, U_b)$. This implies that highly correlated video frames should be watermarked with highly correlated watermark patterns, and vice versa. Our second design principle is a more precise mathematical statement of the hypothesis originally proposed by et. al. Swanson *et al.* in [15] that visually similar regions of video sequences should be marked with consistent watermarks.

IV. Discussion

Although the theoretical results presented here are strictly correct only when the specified assumptions hold, we believe that they are still indicative of the behavior to be expected under more general circumstances. Assumption (A1), which states that the video frames share a common finite mean and variance, may seem at first like a restrictive condition. It is expected to hold approximately over scenes with similar lighting arrangements. But more importantly, it can be shown that if it does not hold:

$$\left|\frac{\rho^2(U_a, U_b)}{\rho^2(X_a, X_b)} - 1\right| = \left|\frac{\mathbf{E}\alpha^2\sigma_W^2(var(U_a) + var(U_b) - 2\sqrt{var(U_a)var(U_b)})}{var(U_a)var(U_b) + 2\mathbf{E}\alpha^2\sigma_W^2\sqrt{var(U_a)var(U_b)} + (\mathbf{E}\alpha^2\sigma_W^2)^2}\right|. \tag{20}$$

Since the image power is clearly greater than that of the watermark, the dominant term on the right hand side of (20) is the product $var(U_a)var(U_b)$ in the denominator. The expression then tends to 0 even when $|var(U_a) - var(U_b)|$ is quite large, as long as the images have reasonably large variances.[3] Thus, $\rho(U_a, U_b) \approx U(X_a, X_b)$ and the statistical invisibility criterion is approximately met and collusion resistance is loosely supported for a broader range of cases than is immediately evident.

Assumption (A2) holds for all existing spread spectrum watermarks. Assumption (A3), which considers the second

[3]Typical values might be $var(U) = 2500 \pm 500$ for an 8-bit grayscale image. For instance, the variances of some commonly used test images are: barbara: 2982, goldhill: 2423, and boat: 2179.

moment of the scaling factors is more difficult to interpret. If a global scaling factor is used, it is clear that the assumption is valid. On the other hand, if the factors are locally derived from an image property like the Noise Visibility Function (NVF) [19], the analysis without this assumption becomes more complicated. Assumption (A4) is reasonable when the watermark pattern has a large spatial spread, and also applies for a global scaling factor.

Accepting these assumptions as valid, we can consider what Theorem 1 tells us about the collusion-resistant properties of some existing video watermarks. First we look at the algorithms in which the same noise-like watermark pattern is embedded into each frame of a video sequence, i.e., where $W_a = W_b \forall\, a,b \in \{1,2,\ldots,n\}$ [7]. By definition, since the same pattern is used to generate the watermark for each frame, $\rho(W_a, W_b) = 1$ regardless of the correlations of the underlying video frames. At the other end of the spectrum are the video watermarking strategies in which different noise-like patterns are embedded into each video frame [18]. One typical scenario is to generate W_a and W_b as two statistically independent patterns. In this case, $\rho(W_a, W_b) = 0$.

In [12], the authors propose to construct a video watermark based on their powerful Discrete Fourier Transform (DFT) domain image watermark [20]. To address the collusion problem, the watermarking key is changed every L frames. Thus the resulting watermark frames have pairwise correlations $\rho(W_a, W_b)$ equal to 1 when $\lceil a/L \rceil = \lceil b/L \rceil$, and 0 otherwise. From a statistical invisibility perspective, this mark may offer better performance than the two extreme cases considered in the last paragraph. In particular, temporally distant watermark frames are uncorrelated and we expect that the underlying video frames will likewise be negligibly correlated. Temporally adjacent watermark frames are also likely to be identical and the underlying video frames are expected to have a high correlation. However, the collusion-resistance of the mark depends strongly on the selection of L. Observe that when $L = \infty$ we get the same result as in the first of the extreme cases above (i.e., $\rho(W_a, W_b) = 1$), and when $L = 1$ we get the second, (i.e., $\rho(W_a, W_b) = 0$).

Finally, a more sophisticated video watermark based on the Temporal Wavelet Transform (TWT) has been proposed [15]. The watermark frame correlations range over [0,1], but since the watermark is constructed in a three-dimensional space, it is difficult to study the two-dimensional frame-by-frame correlations analytically. The watermark is designed to be temporally layered by embedding the mark separately into static and dynamic components of the video, therefore, we expect it to exhibit good statistical invisibility. However, it employs two levels of transforms, as well as two layers of visual masking, making it one of the most computationally complex approaches.

We hope that the practical design strategies highlighted in this work will serve to inspire more feasible, but robust watermark constructions. The sequel to this paper attempts at the development of such an algorithm. Future directions for this work include developing methods to achieve matched host-watermark correlation while analyzing the effect of nondimension-preserving transformations, such as scaling, on the statistical correlation coefficients.

V. Final Remarks

In Part I of this two-part paper, we developed a simple yet effective framework for the statistical analysis of linear collusion resistance in video watermarking. An important component that has been added to the body of watermarking research is the derivation of equations describing properties that a watermark should possess in order to resist this class of attacks. Through this theoretical work we defined a desirable watermark property called *statistical invisibility* in which the presence of a collusion-resistant watermark is not revealed using linear statistical tools. In other words, given a set of second-order statistics describing the relationships between the frames of a video sequence, we have shown that collusion resistance is achieved if these statistics are not affected by the watermarking procedure. We propose a new theorem and two video watermark design principles summarizing this result. Finally, we considered how a number of existing video watermarking techniques measure up in terms of statistical invisibility, and found that there is indeed some room for improvement. The second installment of this work [1] proposes a novel video watermarking approach that attempts to make more effective use of our two watermark design principles. Evaluation and simulation of this algorithm allows the ideas presented in this paper to be assessed under more practical conditions.

Appendix

A. Proof of Proposition 1

Proof: Let $\mathbf{E}B = \mu$, $\mathbf{E}B^2 = \mu_2$, and

$$\hat{A} = \sqrt{\frac{var(B)}{var(A)}}(A - \mathbf{E}A) + \mathbf{E}B = \sqrt{\frac{\mu_2 - \mu^2}{var(A)}}(A - \mathbf{E}A) + \mu$$

so that $\hat{A}$ is a normalized version of A with the same mean and variance as B. If A is an ϵ-optimal MSE of B, then from Definition 1

$$\begin{aligned}
\epsilon &> \mathbf{E}[(\hat{A} - B)^2] && (21)\\
&= 2\mu_2 - 2\mathbf{E}\hat{A}B && (22)\\
&\Longleftrightarrow \\
\mathbf{E}\hat{A}B &> \mu_2 - \frac{\epsilon}{2} && (23)\\
&\Longleftrightarrow \\
cov(\hat{A}, B) &= \mathbf{E}\hat{A}B - \mu^2 \\
&> \mu_2 - \mu^2 - \frac{\epsilon}{2} \\
&= var(B) - \frac{\epsilon}{2} && (24)\\
&\Longleftrightarrow \\
\rho(\hat{A}, B) &> 1 - \tilde{\epsilon} && (25)
\end{aligned}$$

where in the final step both sides of the inequality were divided by $var(B) > 0$. From the properties of the correlation coefficient, $\rho(\hat{A}, B) = \rho(A, B)$. Therefore, $\rho(A, B) > 1 - \tilde{\epsilon}$, where $\tilde{\epsilon} = \epsilon/2(\mu_2 - \mu^2)$. ■

B. Proof of Proposition 2

From [21], we have the following Lemma:

Lemma 1: Given a set of independent random variables $Z_1, Z_2, \ldots Z_n$ with finite means and variances, and their sum $Z = \sum_{i=1}^{n} Z_i$, the squared correlation coefficient of the composite random variable with itself is equal to the sum of the squared coefficients of its correlations with each of the independent components, i.e.,

$$\begin{aligned}\rho^2(Z,Z) &= \rho^2\left(Z, \sum_{i=1}^{n} Z_i\right) \\ &= \sum_{i=1}^{n} \rho^2(Z, Z_i) \\ &= 1. \end{aligned} \tag{26}$$

Proof (Type I): By Definition 2, $\overline{X}$ is an ϵ-optimal MSE estimate of the watermark component $\overline{W}$. Therefore from Proposition 1

$$\rho(\overline{X}, \overline{W}) > 1 - \tilde{\epsilon}. \tag{27}$$

Since $\overline{X} = \overline{U} + \overline{W}$, where $\overline{U}$ and $\overline{W}$ are independent, we apply Lemma 1 to give

$$\rho^2(\overline{X}, \overline{X}) = 1 = \rho^2(\overline{X}, \overline{U}) + \rho^2(\overline{X}, \overline{W}). \tag{28}$$

Equations (27) and (28), imply

$$\rho(\overline{X}, \overline{U}) < \sqrt{\tilde{\epsilon}(2 - \tilde{\epsilon})}$$

where $0 \leq \sqrt{\tilde{\epsilon}(2-\tilde{\epsilon})} \leq 1$ is a monotonically increasing function of $\tilde{\epsilon}$ (and also of ϵ). ■

Proof (Type II): Similarly by Definition 2, $\overline{X}$ is an ϵ-optimal MSE estimate of the host component $\overline{U}$, and from Proposition 1 we have the necessary condition:

$$\rho(\overline{X}, \overline{U}) > 1 - \tilde{\epsilon}.$$

■

C. Proof of Proposition 3

Proof: Given the assumptions, the following hold:

$$\begin{aligned}\mathbf{E}X_k &= \mathbf{E}(U_k + \alpha_k W_k) \\ &= \mathbf{E}U_k + \mathbf{E}\alpha_k W_k \\ &= \mathbf{E}U_k + \mathbf{E}\alpha_k \mathbf{E}W_k \\ &= \mathbf{E}U_k \\ &= \mu_U. \end{aligned} \tag{29}$$

$$\begin{aligned}\mathbf{E}X_k^2 &= \mathbf{E}(U_k + \alpha_k W_k)(U_k + \alpha_k W_k) \\ &= \mathbf{E}U_k^2 + 2\mathbf{E}\alpha_k \mathbf{E}W_k \mathbf{E}U_k + \mathbf{E}\alpha_k^2 \mathbf{E}W_k^2 \\ &= \sigma_U^2 + \mu_U^2 + \mathbf{E}\alpha^2 \cdot \sigma_W^2. \end{aligned} \tag{30}$$

$$\begin{aligned}\mathbf{E}X_a X_b &= \mathbf{E}(U_a + \alpha_a W_a)(U_b + \alpha_b W_b) \\ &= \mathbf{E}U_a U_b + \mathbf{E}\alpha_a \mathbf{E}W_a \mathbf{E}U_b \\ &\quad + \mathbf{E}\alpha_b \mathbf{E}W_b \mathbf{E}U_a + \mathbf{E}\alpha_a \alpha_b \mathbf{E}W_a W_b \\ &= \mathbf{E}U_a U_b + \mathbf{E}\alpha_a \alpha_b \mathbf{E}W_a W_b. \end{aligned} \tag{31}$$

Therefore, $\forall\, a, b \in \{1, \ldots, n\}$,

$$\begin{aligned}\rho(X_a, X_b) &= \frac{cov(X_a, X_b)}{\sqrt{var(X_a) var(X_b)}} \\ &= \frac{\mathbf{E}X_a X_b - \mathbf{E}X_a \mathbf{E}X_b}{\sqrt{(\mathbf{E}X_a^2 - \mathbf{E}^2 X_a)(\mathbf{E}X_b^2 - \mathbf{E}^2 X_b)}} \\ &= \frac{\mathbf{E}U_a U_b + \mathbf{E}\alpha_a \alpha_b \mathbf{E}W_a W_b - \mathbf{E}U_a \mathbf{E}U_b}{\sigma_U^2 + \mu_U^2 + \mathbf{E}\alpha^2 \cdot \sigma_W^2 - \mu_U^2} \\ &= \frac{cov(U_a, U_b) + \mathbf{E}\alpha_a \alpha_b cov(W_a, W_b)}{\sigma_U^2 + \mathbf{E}\alpha^2 \cdot \sigma_W^2} \\ &= \frac{\rho(U_a, U_b)\sigma_U^2 + \mathbf{E}\alpha_a \alpha_b \cdot \rho(W_a, W_b)\sigma_W^2}{\sigma_U^2 + \mathbf{E}\alpha^2 \cdot \sigma_W^2} \end{aligned} \tag{32}$$

where $cov(\cdot,\cdot)$, $var(\cdot)$, and $\mathbf{E}^2(\cdot)$ denote the covariance, variance and squared expectation of the argument random variables. For statistical invisibility in Definition 3 we require that $\rho(X_a, X_b) = \rho(U_a, U_b)$. Therefore, equating the right hand side of (32) to $\rho(U_a, U_b)$, we conclude that under assumptions (A1)–(A4),

$$\begin{aligned}&\rho(X_a, X_b) = \rho(U_a, U_b) \forall a, b \in \{1, 2, \ldots, n\} \\ &\qquad \text{(statistical invisibility)} \\ &\qquad\qquad \Longleftrightarrow \\ &\rho(U_a, U_b) = \frac{\mathbf{E}\alpha_a \alpha_b}{\mathbf{E}\alpha^2} \rho(W_a, W_b) \forall a, b \in \{1, 2, \ldots, n\}. \end{aligned}$$

■

D. Proof of Proposition 4

Proof: Consider the simplest case where $n = 2$:

$$\begin{aligned}\rho(\overline{X}, \overline{U}) &= \rho(\beta_1 X_1 + \beta_2 X_2, \beta_1 U_1 + \beta_2 U_2) \\ &= \sqrt{\frac{var(\beta_1 U_1 + \beta_2 U_2)}{var(\beta_1 U_1 + \beta_2 U_2) + var(\beta_1 \alpha_1 W_1 + \beta_2 \alpha_2 W_2)}}. \end{aligned} \tag{33}$$

Similarly

$$\rho(X_a, U_a) = \sqrt{\frac{var(U_a)}{var(U_a) + \mathbf{E}\alpha_a^2 \cdot \sigma_W^2}}. \tag{34}$$

Therefore

$$\rho(\overline{X}, \overline{U}) = \rho(X_a, U_a) \forall\, a \in \{1, 2\} \tag{35}$$

$$\Longleftrightarrow$$

$$\frac{var(\beta_1 U_1 + \beta_2 U_2)}{var(U_a)} = \frac{var(\beta_1 \alpha_1 W_1 + \beta_2 \alpha_2 W_2)}{\mathbf{E}\alpha^2 \cdot \sigma_W^2}. \tag{36}$$

The left-hand side of (36) can be reduced by applying the assumption that the video frames share a common variance:

$$\begin{aligned}\text{L.H.S. of (36)} &= \frac{\beta_1^2 \sigma_U^2 + 2\beta_1 \beta_2 cov(U_1 U_2) + \beta_2^2 \sigma_U^2}{\sigma_U^2} \\ &= \beta_1^2 + \beta_2^2 + 2\beta_1 \beta_2 \cdot \rho(U_1, U_2). \end{aligned} \tag{37}$$

The right-hand side can also be reduced by employing the assumption that the scaling factors share a common second moment:

$$\text{R.H.S. of (36)} = \frac{\beta_1^2\mathbf{E}\alpha^2\cdot\sigma_W^2+2\beta_1\beta_2\mathbf{E}\alpha_1\alpha_2\cdot cov(W_1,W_2)+\beta_2^2\mathbf{E}\alpha^2\cdot\sigma_W^2}{\mathbf{E}\alpha^2\cdot\sigma_W^2}$$
$$= \beta_1^2+\beta_2^2+\frac{2\beta_1\beta_2\mathbf{E}\alpha_1\alpha_2}{\mathbf{E}\alpha^2}\rho(W_1,W_2). \quad (38)$$

The desired conclusion can then be obtained by combining (37) and (38), thus establishing the relationship between statistical invisibility and robustness to linear collusion (for the case of $n=2$):

$$\rho(\overline{X},\overline{U}) = \rho(X_a,U_a)\forall\, a\in\{1,2\} \quad (39)$$
$$\Longleftrightarrow$$
$$\rho(U_1,U_2) = \frac{\mathbf{E}\alpha_1\alpha_2}{\mathbf{E}\alpha^2}\rho(W_1,W_2). \quad (40)$$

In the case of $n>2$, it can be shown that since

$$\rho(\overline{X},\overline{U}) = \rho(X_a,U_a)\forall\, a\in\{i,j\} \quad (41)$$
$$\Longleftrightarrow$$
$$\rho(U_i,U_j) = \frac{\mathbf{E}\alpha_i\alpha_j}{\mathbf{E}\alpha^2}\rho(W_i,W_j) \quad (42)$$

holds pairwise $\forall\{i,j\}\subseteq\{1,2,\ldots,n\}$, then the more general form, presented in this proposition, also holds. ■

E. Proof of Proposition 5

Proof: In this work we have made the implicit assumptions for each video frame stated in (2) and (3) presented here once again for frame a:

$$\rho(X_a,\alpha_aW_a) > 1-\tilde{\gamma} \text{ (robustness to individual frame distortions)} \quad (43)$$
$$\mathbf{E}[(X_a-U_a)^2] < \xi \text{ (imperceptibility to a human observer).} \quad (44)$$

Therefore, for $0<\tilde{\gamma}<0.5$ and $0<\tilde{\epsilon}\ll 1$, from the robustness restriction

$$\rho(X_a,\alpha_aW_a) > 1-\tilde{\gamma} \quad (45)$$
$$\Longleftrightarrow$$
$$\rho(X_a,U_a) < \sqrt{\tilde{\gamma}(2-\tilde{\gamma})} \quad (46)$$
$$\Longrightarrow$$
$$\rho(X_a,U_a) < 1-\tilde{\epsilon} \quad (47)$$

where $1-\tilde{\epsilon} > \sqrt{\tilde{\gamma}(2-\tilde{\gamma})}$ for the assumed ranges of $\tilde{\gamma}$ and $\tilde{\epsilon}$.

Similarly, for $0<\xi\ll 1$ and $0<\tilde{\epsilon}\ll 1$, from the imperceptibility restriction

$$\mathbf{E}[(X_a-U_a)^2] < \xi \quad (48)$$
$$\Longrightarrow$$
$$\mathbf{E}[(\hat{X}_a-U_a)^2] < \xi \quad (49)$$
$$\Longleftrightarrow$$
$$\rho(X_a,U_a) > 1-\tilde{\xi} \quad (50)$$
$$\Longleftrightarrow$$
$$\rho(X_a,U_a) > \sqrt{\tilde{\epsilon}(2-\tilde{\epsilon})} \quad (51)$$

where $\tilde{\xi}=\xi/(2(var(U_a))$ and we make use of the properties that $\mathbf{E}[(\hat{X}_a-U_a)^2]\le\mathbf{E}[(X_a-U_a)^2]$ since $\hat{X}_a$ is just a scaled version of X_a such that it minimizes the variance of $\hat{X}_a-U_a$, and $1-\tilde{\xi}>\sqrt{\tilde{\epsilon}(2-\tilde{\epsilon})}$ for the ranges of ξ and $\tilde{\epsilon}$ considered.

Thus, $\sqrt{\tilde{\epsilon}(2-\tilde{\epsilon})} < \rho(X_a,U_a) < 1-\tilde{\epsilon}$ implicitly in our formulation. Therefore, in practice

$$\rho(\overline{X},\overline{U}) = \rho(X_a,U_a)\ \forall a\in\{1,2,\ldots,n\}$$
$$\Longrightarrow$$
$$\sqrt{\tilde{\epsilon}(2-\tilde{\epsilon})} < \rho(\overline{X},\overline{U}) < 1-\tilde{\epsilon}.$$

From Proposition 2 the statements above are both sufficient conditions for linear collusion. ■

F. Proof of Proposition 6

Proof: From (34), the expression for the correlation coefficient between a host video frame and its corresponding watermarked copy is

$$\rho(X_a,U_a) = \sqrt{\frac{var(U_a)}{var(U_a)+\mathbf{E}\alpha_a^2\cdot\sigma_W^2}}.$$

Therefore the following equalities hold $\forall a,b\in\{1,2,\ldots,n\}$:

$$\rho(U_a,X_a) = \rho(U_b,X_b)$$
$$\Longleftrightarrow$$
$$\sqrt{\frac{var(U_a)}{var(U_a)+\mathbf{E}\alpha_a^2\cdot\sigma_W^2}} = \sqrt{\frac{var(U_b)}{var(U_b)+\mathbf{E}\alpha_b^2\cdot\sigma_W^2}} \quad (52)$$
$$\Longleftrightarrow$$
$$var(U_a)\cdot\mathbf{E}\alpha_b^2\cdot\sigma_W^2 = var(U_b)\cdot\mathbf{E}\alpha_a^2\cdot\sigma_W^2$$
$$\Longleftrightarrow$$
$$\frac{\mathbf{E}\alpha_b^2}{\mathbf{E}\alpha_a^2} = \frac{var(U_b)}{var(U_a)}. \quad (53)$$

■

REFERENCES

[1] K. Su, D. Kundur, and D. Hatzinakos, "A novel statistically invisible digital video watermark for collusion resistance," *IEEE Trans. Multimedia*, vol. 7, no. 1, pp. 61–75, Feb. 2005.

[2] K. Tanaka, Y. Nakamura, and K. Matsui, "Embedding secret information into a dithered multi-level image," in *Proc. IEEE Military Communications Conf. MILICOM '90*, vol. 1, Sep. 1990, pp. 216–220.

[3] A. Z. Tirkel, G. A. Rankin, R. G. van Schyndel, W. J. Ho, N. Mee, and C. F. Osborne, "Electronic water mark," in *Proc. Int. Conf. Digital Image Computing, Techniques and Applications—DICTA '93*, Dec. 1993, pp. 666–672.

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

[5] J. A. Bloom, I. J. Cox, T. Kalker, J.-P. M. G. Linnartz, M. L. Miller, and C. B. S. Traw, "Copy protection for DVD video," *Proc. IEEE*, vol. 87, pp. 1267–1276, Jul. 1999.

[6] V. Darmstaedter, J.-F. Delaigle, D. Nicholson, and B. Macq, "A block based watermarking technique for MPEG2 signals: Optimization and validation on real digital TV distribution links," in *Proc. Eur. Conf. on Multimedia Applications, Services and Techniques*, 1998, pp. 190–206.
[7] T. Kalker, G. Depovere, J. Haitsma, and M. Maes, "A video watermarking system for broadcast monitoring," in *Proc. SPIE*, vol. 3657, Jan. 1999, pp. 103–112.
[8] C. Griwodz, O. Merkel, J. Dittmann, and R. Steinmetz, "Protecting VoD the easier way," *ACM Multimedia*, pp. 21–28, 1998.
[9] R. Barnett, "Digital watermarking: Applications, techniques, and challenges," *Electron. Commun. Eng. J.*, vol. 11, no. 4, pp. 173–183, Aug. 1999.
[10] J. Dittmann, F. Nack, A. Steinmetz, and R. Steinmetz, "Interactive watermarking environments," in *Proc. IEEE Int. Conf. Multimedia Computing and Systems*, 1998, pp. 286–294.
[11] F. Hartung, P. Eisert, and B. Girod, "Digital watermarking of MPEG-4 facial animation parameters," *Comput. Graph.*, vol. 22, no. 4, pp. 425–435, Jul.-Aug. 1998.
[12] F. Deguillaume, G. Csurka, and T. Pun, "Countermeasures for unintentional and intentional video watermarking attacks," in *Proc. SPIE*, vol. 3971, Jan. 2000, pp. 346–357.
[13] K. Su, D. Kundur, and D. Hatzinakos, "A content-dependent spatially localized video watermark for resistance to collusion and interpolation attacks," in *Proc. IEEE Int. Conf. Image Processing*, vol. 1, 2001, pp. 818–821.
[14] W. Trappe, M. Wu, and K. J. R. Liu, "Collusion-resistant fingerprinting for multimedia," in *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing*, vol. 4, 2002, pp. 3309–3312.
[15] M. D. Swanson, B. Zhu, and A. T. Tewfik, "Multiresolution scene-based video watermarking using perceptual models," *IEEE J. Select. Areas Commun.*, vol. 16, no. 4, pp. 540–550, May 1998.
[16] M. Holliman, N. Memon, and M. M. Yeung, "On the need for image dependent keys for watermarking," in *Proc. IEEE Symp. Content Security and Data Hiding in Digital Media*, Newark, NJ, 1999.
[17] I. J. Cox, J. Kilian, F. T. Leighton, and T. Shamoon, "Secure spread spectrum watermarking for multimedia," *IEEE Trans. Image Process.*, vol. 6, pp. 1673–1687, Dec. 1997.
[18] B. G. Mobasseri, "Exploring CDMA for watermarking of digital video," *Proc. SPIE*, vol. 3657, pp. 96–102, Jan. 1999.
[19] S. Voloshynovskiy, A. Herrigel, N. B. , and T. Pun, "A stochastic approach to content adaptive digital image watermarking," *Lecture Notes Computer Science*, vol. 1768, pp. 212–236, Sep. 2000.
[20] S. Pereira and T. Pun, "Robust template matching for affine resistant image watermarks," *IEEE Trans. Image Process.*, vol. 9, no. 6, pp. 1123–1129, Jun. 2000.
[21] G. Grimmett and D. Stirzaker, *Probability and Random Processes*. Oxford, U.K.: Oxford Univ. Press, 1992.

Karen Su (S'98) received the B.A.Sc. degree in electrical engineering, honors athematics option with distinction, from the University of British Columbia, Vancouver, BC, Canada, in 1999, and the M.A.Sc. degree in electrical engineering from the University of Toronto, Toronto, ON, Canada, in 2001. She is currently pursuing the Ph.D. degree in engineering at the University of Cambridge, Laboratory for Communication Engineering, Cambridge, U.K. Her research interests are in the areas of coding theory, wireless communications, digital watermarking, and digital signal processing.

Deepa Kundur (S'93–M'99–SM'03) was born in Toronto, ON, Canada. She received the B.A.Sc., M.A.Sc., and Ph.D. degrees, all in electrical and computer engineering, in 1993, 1995, and 1999, respectively, from the University of Toronto.

In January 2003, she joined the Electrical Engineering Department at Texas A&M University, College Station, where she is a member of the Wireless Communications Laboratory and an Assistant Professor. From September 1999 to December 2002, she was an Assistant Professor at the Edward S. Rogers Sr. Department of Electrical and Computer Engineering, University of Toronto where she was Bell Canada Junior Chair-holder in Multimedia. Her research interests include multimedia and network security for digital rights management, video cryptography, data hiding and steganography, covert communications, and nonlinear and adaptive information processing algorithms.

Dr. Kundur has been on numerous technical program committees and has given tutorials at ICME 2003 and Globecom 2003 in the area of digital rights management. She is a Guest Editor for the PROCEEDINGS OF THE IEEE Special Issue on Enabling Technologies for Digital Rights Management. She was the recipient of the 2002 Gordon Slemon Teaching of Design Award and the 2002 Best Electrical Engineering Professor Award (Spring) presented by the ECE Club at the University of Toronto.

Dimitrios Hatzinakos (M'90–SM'98) received the Diploma degree from the University of 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, Toronto, ON, where now he holds the rank of Professor with tenure. He also serves 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 120 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 served as an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING from 1998 through 2002 and was Guest Editor for the special issue on Signal Processing Technologies for Short Burst Wireless Communications for Elsevier's *Signal Processing*, in October 2000. He was a member of the IEEE Statistical Signal and Array Processing Technical Committee (SSAP) from 1992 through 1995 and Technical Program co-Chair of the 5th Workshop on Higher-Order Statistics in July 1997. He was co-organizer and Technical program co-Chair of the IEEE Toronto Centenial Workshop on Wireless Communications, held in Toronto in October 2003. He is a member of EURASIP, the Professional Engineers of Ontario (PEO), and the Technical Chamber of Greece.