CHAOTIC WATERMARK SEQUENCES FOR CORRELATION-BASED SCHEMES

Anastasios Tefas *Nikos Nikolaidis* *Ioannis Pitas*

Department of Informatics, Aristotle University of Thessaloniki
Box 451, Thessaloniki 54124, GREECE
{tefas,nikolaid,pitas}@zeus.csd.auth.gr

ABSTRACT

In this paper, an overview of watermarking schemes based on chaotic generators and correlation detection is presented. Statistical properties of watermark sequences generated by piecewise-linear Markov maps are exploited for both additive and multiplicative watermark embedding. Correlation/spectral properties of such sequences are easily controllable, a fact that reflects on the watermarking system performance. A family of chaotic maps, namely the skew tent map family, is used in temporal and transform-domain watermarking schemes. The chaotic watermarking framework is applied successfully to audio signals, demonstrating its superiority with respect to both robustness and inaudibility.

1. INTRODUCTION

The design of robust techniques for copyright protection and content verification of multimedia data became an urgent necessity in the last years. This demand has been lately addressed by the emergence of a variety of watermarking methods. Such methods target towards hiding in the original data an imperceptible and undetectable signal which conveys information about the host medium (owner or authorized user, transaction or product ID, etc). For a review of existing schemes and a detailed discussion on the main requirements of a watermarking scheme, the interested reader may consult [1].

So far, many approaches have attempted to statistically analyze the performance of watermarking schemes in terms of detection reliability by addressing the problem in a communication framework (see for example [2, 3, 4]). In these papers, the statistical properties of watermarking schemes based on pseudorandom watermark signals and correlation detectors were derived, among others. In [3], the authors investigate the performance of white and lowpass-filtered pseudorandom watermarks concluding that the former are ideal when no distortions are inflicted on the image, whereas the latter provide additional robustness against lowpass distortions.

Watermarking techniques based on chaotic systems appeared in the literature already in the first years of watermarking research [5]. An overview of early chaotic watermarking techniques can be found in [6]. In this paper, an overview of watermarking schemes based on chaotic generators and correlation detection is presented. Statistical properties of watermark sequences generated by piecewise-linear Markov maps are exploited for both additive and multiplicative watermark embedding. The major advantage of chaotic sequences is their easily controllable spectral/correlation properties, a fact that makes them a good alternative to the widely used pseudorandom signals [7, 8]. Chaotic watermarks can be either embedded in the temporal/spatial domain or in a transform domain where their corrlation/spectral properties can be exploited more efficiently for obtaining robust watermarking schemes. Such a transform-domain audio watermarking technique is presented in this paper. The scheme involves multiplicative embedding of high-frequency chaotic watermarks in the low frequencies of the Discrete Fourier Transform (DFT). The corresponding watermarking scheme guarantees robustness against lowpass attacks, along with enhancement of the detector reliability. The complete theoretical justification and statistical analysis of correlation-based additive and multiplicative schemes employing chaotic watermarks can be found in [8, 9].

This work has been supported by the European IST Project - ECRYPT

2. WATERMARKING SYSTEM MODEL

Within a watermarking system, the watermark generation functional block aims at constructing a sequence $\boldsymbol{w}$, $w[i] \in \mathscr{R}$, of N samples using an appropriate generation function G, $\boldsymbol{w} = G(K,N)$, where K denotes the watermark key. Watermark embedding aims at inserting the watermark signal $\boldsymbol{w}$ in the host signal $\boldsymbol{f}$ in a way that ensures imperceptibility and robustness under intentional or unintentional attacks. For the model under study, either additive watermark embedding $\boldsymbol{f}_w = \boldsymbol{f} + p\boldsymbol{w}$, or multiplicative embedding $f_w(n) = f(n) + p\, w(n)\, |f(n)|$ is considered, where $\boldsymbol{f}_w$ is the watermarked signal and p is a constant that controls the watermark embedding power, which will be called hereafter watermark embedding factor. Obviously, p is closely related to the watermark perceptibility. In the following, we will describe the basic formulation of correlation-based detection for additive watermarks. The corresponding equations for multiplicatively embedded watermarks are straightforward to derive [9] and only the peculiarities that arise in this case will be presented.

Watermark detection can be formulated as a binary hypothesis test, the two hypotheses being the following:

- H_0: The test signal $\boldsymbol{f}_t$ contains the watermark $\boldsymbol{w}_d$, i.e., $\boldsymbol{f}_t = \boldsymbol{f}_o + p\boldsymbol{w}_d$, $\boldsymbol{f}_o$ being the host signal.
- H_1: The test signal $\boldsymbol{f}_t$ does not contain the watermark $\boldsymbol{w}_d$.

The two events mentioned above can be summarized in the following formula:

$$\boldsymbol{f}_t = \boldsymbol{f}_o + p\boldsymbol{w}_e \tag{1}$$

where the watermark $\boldsymbol{w}_d$ is indeed embedded in the signal if $p \neq 0$ and $\boldsymbol{w}_e = \boldsymbol{w}_d$ (event H_0), and it is not embedded in the signal if $p = 0$ (no watermark is present, denoted hereafter as event H_{1a}) or $\boldsymbol{w}_e \neq \boldsymbol{w}_d$ (wrong watermark presence, denoted hereafter as event H_{1b}).

The correlation between the signal under investigation $\boldsymbol{f}_t$ and the watermark $\boldsymbol{w}_d$ is given by:

$$c = \frac{1}{N}\sum_{n=0}^{N-1} f_t[n]w_d[n] = \frac{1}{N}\sum_{n=0}^{N-1} (f_o[n]w_d[n] + pw_e[n]w_d[n]) \tag{2}$$

In order to decide on the valid hypothesis, c is compared against a suitably selected threshold T. For a given threshold, the system performance can be measured in terms of the probability of false alarm $P_{fa}(T)$, (i.e., the probability to detect a watermark in a signal that is not watermarked or is watermarked with a different watermark) and the probability of false rejection $P_{fr}(T)$ (i.e., the probability to erroneously neglect the watermark existence in the signal). The plot of P_{fa} versus P_{fr} is called the *receiver operating characteristics* (ROC) curve of the corresponding watermarking system. This curve conveys all the necessary system performance information.

For the watermark sequences that will be studied in this paper, i.e., the sequences generated by piecewise linear Markov maps, the correlation output c is normally distributed (see Section 3). Thus, its distribution can be fully determined in terms of its mean $\mu_{c|H_0}$,

$\mu_{c|H_1}$, and variance $\sigma^2_{c|H_0}$, $\sigma^2_{c|H_1}$, which can be derived in a straightforward manner:

$$\mu_c = E[c] = \frac{1}{N}\sum_{n=0}^{N-1} E[f_o[n]]E[w_d[n]] + \frac{1}{N}\sum_{n=0}^{N-1} pE[w_e[n]w_d[n]] \quad (3)$$

$$\begin{aligned}\sigma_c^2 = \; & E[c^2] - E[c]^2 = \frac{1}{N^2}\left[\sum_{n=0}^{N-1}\Big(E[f_o^2[n]]E[w_d^2[n]] + \right. \\ & p^2E[w_d^2[n]w_e^2[n]] + 2pE[f_o[n]]E[w_e[n]w_d^2[n]]\Big) + \\ & \sum_{n=0}^{N-1}\sum_{m=0,m\neq n}^{N-1}(E[f_o[n]f_o[m]]E[w_d[n]w_d[m]] + \\ & pE[f_o[n]]E[w_d[n]w_e[m]w_d[m]] + \\ & pE[f_o[m]]E[w_e[n]w_d[m]w_d[n]] + \\ & \left. p^2E[w_e[n]w_e[m]w_d[n]w_d[m]]\Big)\right] - \mu_c^2 \end{aligned} \quad (4)$$

By examining (3), (4), one can easily conclude that several higher order moments (statistics) need to be evaluated if μ_c, σ_c^2 are to be computed. To proceed in such an evaluation, an assumption about the statistical properties of the host signal has to be adopted. In our case, the host signal will be assumed to be wide-sense stationary, obeying a first order exponential autocorrelation function model [4]:

$$R_{f_o}[k] = \mu_{f_o}^2 + \sigma_{f_o}^2\beta^k, \quad k \geq 0, \;\; |\beta| \leq 1 \quad (5)$$

where β is the parameter of the autocorrelation function and $\sigma_{f_o}^2$ is the host signal variance.

In the preceding analysis we studied the system model for additive watermark embedding in the temporal domain. In the case where watermark embedding takes place in a transform domain (e.g., DFT) the watermark is usually embedded in specific coefficients and the embedding rule is multiplicative instead of additive. As an example, a multiplicative audio watermarking system will be described [9]. Multiplicative embedding is employed in this case for exploiting masking properties of the human auditory system (HAS). Moreover, watermark embedding to a specific frequency band provides increased performance and inaudibility [9]. We consider x and X to be the source signal and its DFT coefficients, correspondingly. Watermark embedding is performed by modifying the magnitude $F = |X|$ of the DFT coefficients of a specific band. The corresponding watermark sequence can be described by the following formula:

$$W(n) = \begin{cases} W_o(i), & \text{if } aN_s \leq n \leq bN_s, 0 \leq i < N-1 \\ W_o'(i), & \text{if } (1-b)N_s \leq n \leq (1-a)N_s, \\ & 0 \leq i < N-1, \\ 0, & \text{otherwise} \end{cases} \quad (6)$$

where $n = 0, 1, \ldots, N_s - 1$ and coefficients a, b $(0 < a < b \leq 0.5)$ control the frequency terms that will be modified. The watermark signal W_o that is used for the construction of W consists of N samples, where $N = \lceil (b-a)N_s \rceil$, and it is generated as described previously. W_o affects a specific low frequency subband of the host signal, around coefficient 0, according to a multiplicative superposition rule:

$$F'(n) = F(n) + pW(n)F(n) \quad (7)$$

where F' is the watermarked audio signal and p is a constant that controls the watermark embedding power. Due to the symmetry of the DFT magnitude, a reflected version of the signal $W_o'(i) = W_o(N-i-1)$ is also embedded in the low frequency components around coefficient $N_s - 1$. Correlation detection is also utilized in this case to examine whether a test audio signal F_t, described by equation (7), contains a watermark W_t or not:

$$c = \frac{1}{N}\sum_{n=0}^{N-1}(F(n)W_t(n) + pW(n)F(n)W_t(n)) \quad (8)$$

In order to reach a decision about the signal being watermarked or not, c is compared against a suitably selected threshold T.

3. EMPLOYING CHAOTIC SEQUENCES IN WATERMARKING SCHEMES

Sequences generated by chaotic maps constitute an efficient alternative to pseudorandom watermarking sequences. A chaotic discrete-time signal $x[n]$ can be generated by a chaotic system with a single state variable by applying the recursion:

$$x[n] = f(x[n-1]) = f^n(x[0]) = \underbrace{f(f(\ldots(f}_{n \text{ times}}(x[0]))\ldots)) \quad (9)$$

where $f(\cdot)$ is a nonlinear transformation that maps scalars to scalars and $x[0]$ is the system initial condition. The notation $f^n(x[0])$ is used to denote the n-th application of the map.

Let $p_n(\cdot)$ denote the probability density function of the n-th iterate $x[n]$. A linear operator P_f can be defined such that:

$$p_n(\cdot) = P_f\{p_{n-1}(\cdot)\} = P_f^n\{p_0(\cdot)\} \quad (10)$$

This operator, which is referred to as the Frobenius-Perron (FP) operator [10], describes the time evolution of the density $p_n(\cdot)$ for a particular map. Although, in general, the densities at distinct iterates n will differ, there can be certain choices of $p_0(\cdot)$ such that the densities of subsequent iterates do not change, i.e.,

$$p(\cdot) = P_f^n\{p(\cdot)\}, \quad \forall n \quad (11)$$

Such a density $p(\cdot)$, is referred to as the *invariant density* of the map $f(\cdot)$, and constitutes a fixed point of the FP operator. The invariant density plays an important role in the computation of time-averaged statistics of time series from nonlinear dynamics.

A rich class of 1-D chaotic systems that are particularly amenable to analysis are the eventually expanding, piecewise-linear Markov maps. The statistics of Markov maps can be determined in closed form. For a detailed definition of the matrices (i.e., FP matrix and basis correlation matrix) and vectors involved in statistics calculations, one may consult [11], where, a strategy for computing these statistics, was developed. By using the FP matrix, the higher order correlation statistics of Markov maps can be derived. To do so, the FP matrix and the basis correlation matrix must be expanded in a sufficient dimension [11]. For example, for calculating the autocorrelation function of a chaotic sequence, the FP matrix $\boldsymbol{P}_1$ and the corresponding basis correlation matrix are needed. According to (3) and (4) the highest order correlation statistic required for evaluating the mean value and the variance of the detector in the temporal-domain watermarking system is of third order and the corresponding FP matrix that need to be evaluated is $\boldsymbol{P}_3$.

From the preceding discussion one can conclude that a chaotic sequence $\boldsymbol{x}$ is fully described by the map $f(\cdot)$ and the initial condition $x[0]$. By imposing certain constraints on the map or the initial condition, sequences of infinite period can be obtained. Thus, if we consider two finite sequences $\boldsymbol{x}, \boldsymbol{y}$ generated by the iterative application of the same map on two distinct initial conditions $x[0], y[0]$, respectively, that belong to the same chaotic orbit, there will be an integer $k > 0$ such that:

$$x[0] = f^k(y[0]) \quad \text{or} \quad y[0] = f^k(x[0]) \quad (12)$$

The corresponding samples $x[n], y[n]$ are associated through the following expression for a suitably selected $k > 0$ (sequence shift):

$$y[n] = f^n(y[0]) = f^n(f^k(x[0])) = x[n+k] \;\text{ or }\; x[n] = y[n+k] \quad (13)$$

Having described how a chaotic sequence $\boldsymbol{x}$ can be generated in the interval $[0, 1]$, the corresponding chaotic watermark sequence is given by:

$$\boldsymbol{w} = \boldsymbol{x} - d\boldsymbol{1} \quad (14)$$

where d is a constant that controls the range of the watermark sequence, and $\mathbf{1}$ is the unit vector. By substituting (14) in (3) and (4) and considering that $w_d[n] = w_e[n+k]$, according to (13), it is straightforward to derive the mean value and the variance of the correlation c. The constant value d is usually chosen to be the mean value of the chaotic sequence $\boldsymbol{x}$ in order to have a DC free watermark which, according to [4], results in better system performance. Moreover, by subtracting the test signal mean value prior to detection, we can decrease the variance of the correlation, thus obtaining better system performance.

Although samples of Markov chaotic watermarks are correlated for small $k > 0$, since they posses exponential autocorrelation function and $\boldsymbol{w}_d$ is a shifted version of $\boldsymbol{w}_e$, the Central Limit Theorem for random variables with small dependency [12] may be used in order to establish that the correlation c in eq. (2) attains a Gaussian distribution, even in the case of wrong watermark presence (assuming that N is sufficiently large).

4. THE SKEW TENT MAP

In this section, analysis techniques presented so far are being exemplified using the *skew tent* map which is a piecewise linear Markov map. The skew tent map can be expressed as:

$$\mathscr{T} : [0,1] \rightarrow [0,1] \tag{15}$$

$$\mathscr{T}(x) = \begin{cases} \frac{1}{\alpha}\, x & , \ 0 \le x \le \alpha \\ \frac{1}{\alpha - 1}\, x + \frac{1}{1-\alpha} & , \ \alpha < x \le 1 \end{cases} \ , \ \alpha \in (0,1)$$

A trajectory $t[k]$ of the dynamical system is obtained by iterating this map i.e.,

$$t[k] = \mathscr{T}(t[k-1]) = \mathscr{T}^k(t[0]) \tag{16}$$

The invariant density of the skew tent map is uniform. Following the methodology described in [11], the statistical properties of sequences produced using the skew tent map can be derived. The analytical expressions for the first, second and third order correlation statistics required for evaluating the performance of watermarking schemes based on the skew tent map can be evaluated using the Frobenius-Perron operator approach [11]. The power spectral density of the skew tent map sequences can be shown to be:

$$S_t(\omega) = \frac{1 - e_2^2}{12(1 + e_2^2 - 2e_2 \cos\omega)} \tag{17}$$

where $e_2 = 2\alpha - 1$ is an eigenvalue of the corresponding Frobenius-Perron matrix. Thus, by varying the parameter α either highpass ($\alpha < 0.5$), or lowpass ($\alpha > 0.5$) sequences can be produced. For $\alpha = 0.5$ the symmetric tent map is obtained. Sequences generated by the symmetric tent map posses white spectrum, since the autocorrelation function becomes the Dirac delta function. The control over the spectral properties is very useful in watermarking applications, since the spectral characteristics of the watermark sequence are directly related to watermark robustness against common types of attacks, such as filtering and compression.

5. EXPERIMENTAL RESULTS AND DISCUSSION

Various experiments were conducted to demonstrate the efficiency of the chaotic watermarking sequences when employed in an audio watermarking scheme either additively or multiplicatively. For this purpose, music audio signals sampled at 44.1 KHz with 16 bits per sample, were utilized. All sets of experiments were performed by employing chaotic watermark signals generated by the skew tent map, using a total number of 10000 keys. The system detection performance was measured in terms of the ROC curves (plots of P_{fa} versus P_{fr}) under the worst case assumption for P_{fa} evaluation corresponding to the signal being watermarked by a watermark, different than the one used in detection (event H_{1b}). A watermark embedding factor p that resulted in watermarked signals with SNR=30dB has been used in all cases of the additive embedding in the temporal domain. In the experiments with multiplicative embedding a panel of listeners was asked to listen all watermarked sequences for choosing watermarks that are just below the audibility threshold, ensuring a fair comparison.

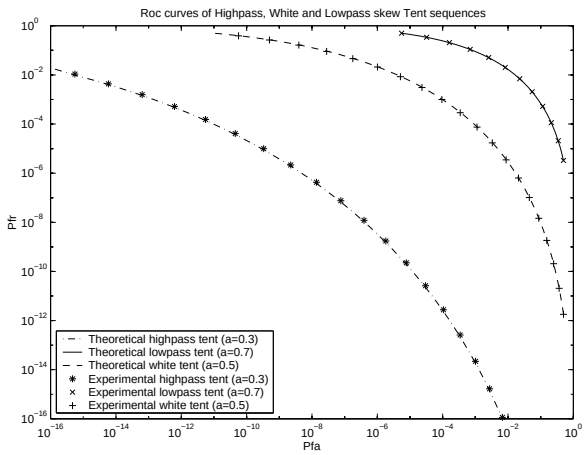

Figure 1: Receiver Operating Characteristics for watermarking schemes based on highpass, lowpass and white skew tent chaotic watermarks.

The ROC curves for lowpass ($\alpha = 0.7$), white ($\alpha = 0.5$) and highpass ($\alpha = 0.3$) skew tent chaotic watermarks embedded additively in the temporal domain were theoretically and experimentally evaluated. The superior performance of the highpass tent chaotic watermarks can be easily observed in Figure 1. The performance of the watermarking system is considerably inferior for white tent watermarks whereas the worst performance is observed when lowpass watermarks are used. However, it is obvious that in case of lowpass attacks, such as filtering or compression, the lowpass watermark will be more robust. In order to take advantage of the superior correlation properties of highpass watermarks even in the case of lowpass attacks one can perform embedding in another domain and not in the spatial one. Moreover, if a highpass watermark is embedded in the low frequencies of the DFT domain, the watermark becomes robust to lowpass attacks while retaining its correlation properties. To achieve this, a multiplicative embedding scheme similar to the one described in Section 2 can be utilized. A large number of experiments were devoted to investigate the robustness of this watermarking scheme against lowpass attacks. In order to compare the performance of the resulting audio watermarking scheme against the performance of alternative techniques, experiments were conducted for two competitive watermark embedding schemes. A correlation detector (applied in the appropriate domain) was used in all three schemes.

The first alternative embedding scheme involved white pseudorandom watermark sequences ($w(i) \in \{-1, 1\}$) multiplicatively embedded in the same low frequency subband ($a = 0.01$, $b = 0.11$) of the DFT domain, producing watermarked signals with SNR=23 db. The second scheme was based on the time-domain audio watermarking technique presented in [13]: a bipolar white pseudorandom watermark $w(n)$ ($w(n) \in \{-1, 1\}$) was modulated according to the amplitude of the original audio samples $m(n)$ using a multiplicative law:

$$w'(n) = p \mid m(n) \mid w(n) \tag{18}$$

where p denotes the embedding strength. In the next stage $w'(n)$ was shaped using a lowpass Hamming filter with cut-off frequency of 2205 Hz, in order to improve imperceptibility and robustness to lowpass attacks. The resulting filtered watermark signal $w''(n)$ was embedded in the time domain of the original signal $m(n)$:

$$m_w(n) = m(n) + w''(n) \tag{19}$$

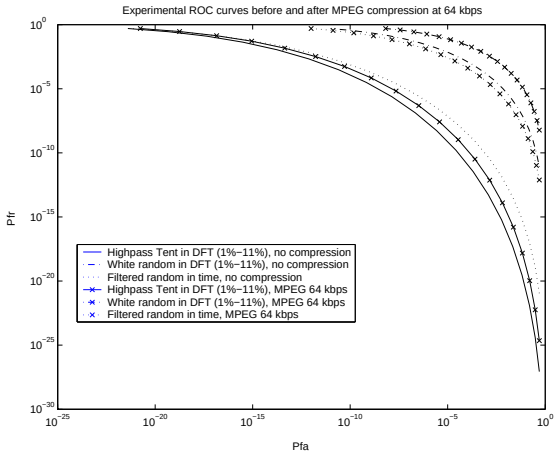

Figure 2: ROC curves for the three watermarking schemes (tent and white watermarks in the DFT domain and prefiltered watermarks in the time-domain) after MPEG compression at 64 kbps.

thus, producing the watermarked signal $m_w(n)$ (SNR=22 db). Watermarks generated using the previously described procedure will be called hereafter "time-domain pseudorandom watermarks".

The superior performance of highpass tent watermarks embedded over the low DFT frequencies, against the alternative techniques described above in the case of MPEG-I layer III encoding at 64 kbps is illustrated in Figure 2. Further experiments have shown that the proposed watermarking scheme outperforms the other two schemes when mean and median filtering, with window of length 3 and 5, as well as other attacks, such as subsampling and cropping of the audio signal, are applied prior to detection [9].

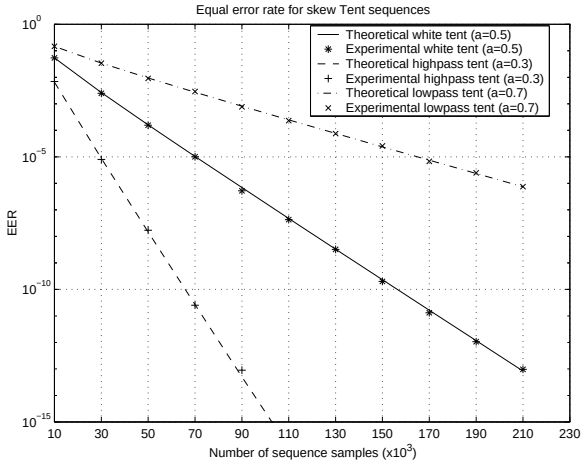

Figure 3: Equal error rate of watermarking schemes based on skew tent maps versus the number of watermarked data samples.

Another important aspect that can be studied by exploiting the theoretical analysis presented in the previous Sections, is the minimum number of watermarked data samples required for a watermarking scheme based on correlation detection in order to achieve a certain performance. This number can be estimated by setting the desired P_{fa} and P_{fr} values and using (3), (4). The Equal Error Rate (EER) i.e., the operating state where $P_{fr} = P_{fa}$ versus the number of watermarked data samples is plotted in Figure 3 for two systems based on tent chaotic watermarks and on additive embedding in the temporal domain. It can be observed that the number of samples required, for a reliable watermarking scheme (e.g. EER $\approx 10^{-12}$), is 80000 for a highpass spectrum watermark and this number increases to 190000 samples for a white watermark. For a lowpass tent watermark the minimum number is much larger.

6. CONCLUSIONS

In this paper, a review on chaotic watermarks generated by Markov maps and their watermarking related statistical properties is presented. Highpass chaotic watermarks prove to perform better than white ones whereas lowpass watermarks have the worst performance when no distortion is inflicted on the watermarked signal. Chaotic watermarks attaining high-frequency spectrum were embedded in the lowest frequency subband of the DFT domain, obeying a multiplicative rule. The statistical properties of the correlation detector were also studied. The proposed technique guaranteed enhancement of the system detection reliability, imperceptibility and great robustness to various attacks.

REFERENCES

[1] "Identification & protection of multimedia information," *Special issue on Proceedings of the IEEE*, vol. 87, no. 7, July 1999.

[2] J.R. Hernandez and F. Perez-Gonzalez, "Statistical analysis of watermarking schemes for copyright protection of images," *Proceedings of the IEEE*, vol. 87, no. 7, pp. 1142–1166, July 1999.

[3] T. Kalker, J-P. Linnartz, and G. Depovere, "On the reability of detecting electronic watermarks in digital images," in *Proc. of EUSIPCO'98*, Rodos, Greece, September 1998.

[4] J.-P. Linnartz, T. Kalker, and G. Depovere, "Modeling the false alarm and missed detection rate for electronic watermarks," in *Proc. of 2nd Information Hiding Workshop*, Oregon, USA, April 1998, pp. 329–343.

[5] G. Voyatzis and I. Pitas, "Chaotic mixing of digital images and applications to wateramrking," in *Proc. of ECMAST'96*, Louvain-la-Neuve, Belgium, 28-30 May 1996, pp. 687–694.

[6] N. Nikolaidis, S. Tsekeridou, A. Nikolaidis, A. Tefas, V. Solachidis, and I. Pitas, "Applications of chaotic signal processing techniques to multimedia watermarking," in *Proceedings of the IEEE workshop on Nonlinear Dynamics in Electronic Systems*, Catania Italy, May 18-20 2000, pp. 1–7.

[7] S. Tsekeridou, V. Solachidis, N. Nikolaidis, A. Nikolaidis, A. Tefas, and I. Pitas, "Statistical analysis of a watermarking system based on bernoulli chaotic sequences," *Elsevier Signal Processing, Sp. Issue on Information Theoretic Issues in Digital Watermarking*, vol. 81, no. 6, pp. 1273–1293, 2001.

[8] A. Tefas, A. Nikolaidis, N. Nikolaidis, V. Solachidis, S. Tsekeridou, and I. Pitas, "Performance analysis of correlation-based watermarking schemes employing markov chaotic sequences," *IEEE Trans. on Signal Processing*, vol. 51, pp. 1979 –1994, July 2003.

[9] A. Giannoula, A. Tefas, N. Nikolaidis, and I. Pitas, "Improving the detection reliability of correlation-based watermarking schemes," in *2003 IEEE International Conference on Multimedia and Expo (ICME 2003)*, Baltimore, USA, 2003, pp. 6–9.

[10] A. Lasota and M.C. Mackey, *Probabilistic Properties of Deterministic Systems*, Cambridge Univ. Press, 1985.

[11] S.H. Isabelle and G.W. Wornell, "Statistical analysis and spectral estimation techniques for one-dimensional chaotic signals," *IEEE Trans. on Signal Processing*, vol. 45, no. 6, pp. 1495–1506, June 1997.

[12] Patrick Billingsley, *Probability and Measure*, Wiley, 1995.

[13] P. Bassia, I. Pitas, and N. Nikolaidis, "Robust audio watermarking in the time domain," *IEEE Transactions on Multimedia*, vol. 3, no. 2, pp. 232–241, June 2001.