OPTIMAL DETECTION FOR MULTIPLICATIVE WATERMARKS EMBEDDED IN DFT DOMAIN

V. Solachidis and I.Pitas

Department of Informatics
University of Thessaloniki
Thessaloniki 54124, Greece
Tel,Fax: +30-2310996304
e-mail: {vasilis/pitas}@zeus.csd.auth.gr

ABSTRACT

This paper deals with the statistical analysis of the behavior of a blind copyright protection watermarking system based on pseudorandom signals embedded in the magnitude of the Fourier transform of the host data. The host data that the watermark is embedded in is one-dimensional and non-white, following a specific model. The analysis performed involves theoretical evaluation of the statistics of the Fourier coefficients and an optimum detector design for multiplicative embedding. It is proved that the optimum detector is not the widely used correlator. Finally, experimental results are presented in order to show the proposed detector's efficiency versus that of the correlator detector.

1. INTRODUCTION

The risk of illegal copying, reproduction and distribution of copyrighted material is becoming more threatening with the all-digital evolving solutions adopted by content providers, system designers and users, thus it made the watermark protection an essential requirement for multimedia distribution.

The watermark is a signal that contains information about the copyright owner and it is embedded permanently in the multimedia data.

The watermarks must be robust to distortions such as those caused by image processing algorithms. Image processing does not modify only the image but may also modify the watermark as well. Thus, the watermark may become undetectable after intentional or unintentional image processing attacks. should not decrease the image quality. A general watermarking framework for copyright protection has been presented in [1, 2] and describes all these issues in detail.

Correlation based detection of watermarked signals is involved in the majority of the watermarking techniques in the literature. However, correlator detector is optimal and minimizes the error probability only in cases when the domain of the signal follows a Gaussian distribution. There are papers in the literature that propose detectors, different than correlator, in the cases that the host data does not follow Gaussian distribution [3]-[4].

In [3] the embedding domain is DCT. The DCT domain distribution is modelled as a generalized Gaussian distribution. Then, the maximum likelihood (ML) criterion is used in order to derive the optimal detector structure.

In [4] the watermark is embedded in the magnitude of the DFT domain. In this case also the authors assume that the Fourier magnitude distribution does not follow the generalized Gaussian distribution. They propose the Weibull one due to the fact that it is a non-zero distribution only on the positive axes and it represents a big distribution family.

In the present paper, the watermark is also assumed to be embedded in the magnitude of the DFT domain. Moreover, we assume that the signal is not white and it follows a specific model. Then, the DFT magnitude distribution is analytically calculated and it is different than the Weibull [4]. Finally, we construct the optimal detector according to the Neyman-Pearson criterion.

The paper is organized as follows. In section 2 the watermarking system model is presented. In the next section the signal model is presented and the distribution of DFT magnitude coefficients is shown. Then, in section 4, the construction of the optimal detector is depicted. In sections 5 and 6, the experimental results, the conclusions and the future work is presented.

2. WATERMARKING SYSTEM MODEL

Let $s(i)$ be the samples of a host signal $\mathbf{S}$ with length N. Let also $S(i)$ be the Fourier transform of $s(i)$ and $M(i)$, $P(i)$ the magnitude of the Fourier transform ($M(i) = |S(i)|$) and its phase, respectively. Suppose that $S_R(k)$ and $S_I(k)$ denote the real and the imaginary part of $S(k)$ respectively. As mentioned in the introduction, the watermark embedding is performed in the Fourier domain and particularly in its magnitude. Thus, from the magnitude of the Fourier transform $\mathbf{M}$, we get the watermarked magnitude. Thus, let us assume that $\mathbf{M}'$ be the watermarked magnitude generated by the watermark embedding function f.

$$\mathbf{M}' = f(\mathbf{M}, \mathbf{W}, p). \tag{1}$$

In the previous formula, $\mathbf{W}$ denotes the samples of the watermark sequence $\mathbf{W}$. This sequence is produced by a random generator. We assume that $W(i)$ consists solely of 1 and -1 and that it is uniformly distributed. Thus, the mean of the watermark sequence $\mathbf{W}$equals zero and in case that f has linear form, it is ensured that the mean of the watermarked magnitude remains unaltered. This property increases both the watermarked signal imperceptibility and its robustness. The parameter p that is employed in equation 1

This work has been supported by the European Project IST-1999-10987 CERTIMARK

is a real number that determines the watermark strength. Increase in the value of p results in more robust watermarks but less imperceptible.

The embedding function is multiplicative where the watermarked magnitude results from the equation below:

$$\mathbf{M}' = \mathbf{M} + \mathbf{MW}p = \mathbf{M}(1 + \mathbf{W}p). \quad (2)$$

In order to compute the final watermarked signal $\mathbf{s}'$ (in spatial domain) the inverse Fourier transform is applied on the watermarked magnitude $\mathbf{M}'$ and the initial phase $\mathbf{P}$.

$$\mathbf{s}' = IFFT(\mathbf{M}', \mathbf{P}). \quad (3)$$

Given a possibly watermarked signal $\mathbf{y}$, watermark detector aims at clarify whether $\mathbf{y}$ hosts a certain watermark $\mathbf{W}$. Watermark detection can be expressed as a hypothesis test where two hypotheses are possible:

- H_0: signal $\mathbf{y}$ does not host watermark $\mathbf{W}$
- H_1: signal $\mathbf{y}$ hosts watermark $\mathbf{W}$

It should be noted that hypothesis H_0 can occur either in case that signal $\mathbf{y}$ is not watermarked (hypothesis H_{0a}) or in case that the signal $\mathbf{y}$ is watermarked by another watermark $\mathbf{W}'$, where $\mathbf{W} \neq \mathbf{W}'$ (hypothesis H_{0b}). The events $\mathbf{W}$,$\mathbf{W}'$ are mutually excluded and their intersection gives the hypothesis H_0.

The performance of the watermarking method depends mainly on the selection of the watermark detector d. Correlator is the most commonly used watermark detector. It has been used in many watermarking methods concerning not only spatial domain watermarking but also watermarking in frequency domain. According to this detector, the test statistic arises out of correlation between the watermark and the possibly watermarked media (sound, image).

$$d = \frac{1}{N} \sum_{i=1}^{N} S_i W_i \quad (4)$$

In order to decide on a valid hypothesis, the detector d is compared against a suitably selected threshold T. The evaluation of the watermark method can be measured by the false alarm (P_{fa}) and the false rejection (P_{fr}) probabilities. False alarm probability is the type I error that is the probability to reject hypothesis H_0 even though it is true, where in our case it is the probability to detect a watermark $\mathbf{W}$ in a signal that is not watermarked by the watermark $\mathbf{W}$. Correspondingly, false rejection is the type II error that is the probability not to detect a watermark $\mathbf{W}$ in a signal that is watermarked by the watermark $\mathbf{W}$ (accept H_0 even it is false).

In most of watermarking methods, hypothesis H_0 is accepted when the detector output is greater than a threshold T. Thus, false alarm and false rejection probabilities can be expressed as:

$$P_{fa} = prob\{d > T/H_0\}, P_{fr} = prob\{d < T/H_1\}.$$

The calculation of the above probabilities can be performed if the detector distribution for both hypotheses is known. Thus, assuming that the $f_0(x)$, $f_1(x)$ are the probability density functions for the hypotheses H_0 and H_1 respectively, the error probabilities are given by:

$$P_{fa} = \int_{T}^{\infty} f_1(x)dx,$$

$$P_{fr} = \int_{\infty}^{T} f_0(x)dx.$$

According to the above equations, P_{fa} and P_{fr} depend on the threshold T. A possible change of T increases the one probability and decreases the other. Thus, apart from the detector, an appropriate threshold should be selected. In many cases the detector is expressed as an sum or a product of almost independent terms that obey the same distribution. Thus, according to the central limit theorem the detector (the detector logarithm in case of product) obey Gaussian distribution. Thus, in that case, the error probabilities can be written as

3. SIGNAL MODEL AND DISTRIBUTION OF DFT MAGNITUDE COEFFICIENTS

A basic step for the optimal detector construction is the computation of the distribution of the domain in which the watermark is embedded. Thus, in this section, the distribution of the DFT magnitude coefficients of a signal will be computed. This procedure will be performed for signals that are modeled as in [5]. In that model, the signal is assumed to be an ergodic and wide-sense stationary stochastic process. The signal is modeled as:

$$E(s(i)) = \mu_s \ \forall i = 0, ..., N-1 \quad (5)$$

$$E(s(i), s(i+D)) = F_{s,s}(D), \ \forall i = 0, ..., N-1 \quad (6)$$

A first order separable autocorrelation function model will be assumed:

$$F_{s,s}(D) = \mu_s^2 + \sigma_s^2 a^{|D|} \quad (7)$$

where σ^2 is the signal variance:

$$\sigma_s^2 = E(s^2) - E(s)^2 \quad (8)$$

and a is a real-valued constant typically equal to $a = 0.9, ...0, 99$. From equations (5) and (7) it is shown that the correlation depends only on the absolute difference of the signal indexes. DFT transform is given by the following equation:

$$S(k) = \sum_{i=0}^{N-1} s(i) e^{-\frac{j2\pi ik}{N}} = \quad (9)$$

$$\sum_{i=0}^{N-1} s(i) \cos\left(\frac{-2\pi ik}{N}\right) + js(i) \sin\left(\frac{-2\pi ik}{N}\right)$$

It should be noted that if a tends to zero, the autocorrelation approaches a Dirac distribution. Although the signal is correlated for small D, the Central Limit Theorem for random variables with small dependency [6] may be used in order to establish that the DFT transform (eq.9) of the signal attains a Gaussian distribution. Thus, it can be proved that the mean of $S(k)$ is given by:

$$\begin{aligned} \mu_{S(k)} &= E[S(k)] = E\left[\sum_{i=0}^{N-1} s(i) e^{-\frac{j2\pi ik}{N}}\right] \quad (10) \\ &= \begin{cases} 0, & \mathbf{k} \neq \mathbf{0} \\ E[s(n)]N, & \mathbf{k} = \mathbf{0} \end{cases} \end{aligned}$$

where $E(\cdot)$ denotes the expected value. The variance of $S(k)$ will be computed separately for its real $S_R(\mathbf{k})$ and imaginary $S_I(\mathbf{k})$ part according to the following formula:

$$\sigma^2_{S_R(k)} = E[S_R(k)^2] - E[S_R(k)]^2 =$$

$$= \sum_{i=0}^{N-1}\sum_{l=0}^{N-1}\cos\left(\frac{-2\pi ik}{N}\right)\cos\left(\frac{-2\pi lk}{N}\right)E[s(i)s(l)] - \mu_{S_R(k)}^2 \quad (11)$$

By substituting equation (5) in the equation (11) we result in

$$\sigma_{S_R(k)}^2 = \sum_{i=0}^{N-1}\sum_{l=0}^{N-1}\cos\left(\frac{-2\pi ik}{N}\right)\cos\left(\frac{-2\pi lk}{N}\right)\left(m^2+\sigma^2 a^{|j-m|}\right) - \mu_{S_R(k)}^2 \quad (12)$$

The final results for the variances of $S_R(k)$ and $S_I(k)$ are given by the equations 14 and 15 respectively.

The next step is to calculate the distribution of the Fourier magnitude $|\mathbf{S}(\mathbf{k})|$. By observing the equation 10 we conclude that all the terms except the DC term, have zero mean. If variances of $S_R(k)$ and $S_I(k)$ were equal then we could conclude that the distribution of $|\mathbf{S}(\mathbf{k})| = \sqrt{S_R(k)^2 + S_I(k)^2}$ is the Rayleigh distribution [7]:

$$|S(k)| \sim f_s(s) = \frac{s}{\sigma^2}\exp\left(-\frac{s^2}{2\sigma^2}\right),\ x > 0.$$

However, the variances of real and imaginary part of $\mathbf{X}(\mathbf{k})$ are equal only in the case of signals that can be modeled as independent identically distributed (i.i.d) random variables ($a = 0$). Thus, for any other case we have to compute the probability density function of a signal

$$z = \sqrt{x^2+y^2}$$

where $x \sim N(0,\sigma_1^2)$, $y \sim N(0,\sigma_2^2)$ and $\sigma_1 \neq \sigma_2$. It is proved that the pdf of the random variable z is given by the equation below:

$$f_z(z) = \frac{z}{\sigma_1\sigma_2}\exp\left(-\frac{\sigma_1^2+\sigma_2^2}{4\sigma_1^2\sigma_2^2}z^2\right) I_0\left(0, \frac{\sigma_2^2-\sigma_1^2}{4\sigma_1^2\sigma_2^2}z^2\right) \quad (13)$$

where I_0 denotes the modified Bessel function and σ_1, σ_2 are the standard deviations of the real and imaginary part of $X(k)$.

4. OPTIMAL DETECTOR

In the next section the optimal watermark detector will be evaluated by using the likelihood ratio test (LTR). The above procedure will be performed for both additive and multiplicative embedded watermarks. According to the Neyman-Pearson theorem, in order to maximize the probability of detection (P_D) for a given $P_{fa} = e$, decide on H_1 if

$$L(\mathbf{M}') = \frac{p(\mathbf{M}'; H_1)}{p(\mathbf{M}'; H_0)} > T \quad (17)$$

where the threshold T can be found from

$$P_{fa} = \int_{\mathbf{M}': L(\mathbf{M}')>T} p(\mathbf{M}'; H_0)d\mathbf{M}' = e \quad (18)$$

The test of eq.(17) is called Likelihood ratio test (LTR). In the sequel the probability density function of the watermarked signal will be computed for both the watermarked with a known and an unknown watermark. Due to the fact that the watermark consists of ± 1 and its mean value is zero then its pdf is:

$$f_w(w) = \begin{cases} 0.5 & ,w=1 \\ 0.5 & ,w=-1 \\ 0 & ,otherwise \end{cases} \quad (19)$$

According to the embedding formula 2, it can be easily proved that the pdf of the watermarked signal is equal to:

$$f_{M'}(x) = \frac{1}{2}\left[\frac{1}{1+p}\ f_M\left(\frac{x}{1+p}\right) + \frac{1}{1-p}\ f_M\left(\frac{x}{1-p}\right)\right] \quad (20)$$

By substituting f_M with the probability density function of distribution in equation (13) in the above equation 20, we result in:

$$f_{M'}(x) = \frac{x}{4\sigma_1\sigma_2}\cdot \quad (21)$$
$$\left[\frac{1}{(1+p)^2}\exp\left(-\frac{\sigma_1^2+\sigma_2^2}{4\sigma_1^2\sigma_2^2}\frac{x^2}{(1+p)^2}\right) I_0\left(0, \frac{\sigma_2^2-\sigma_1^2}{4\sigma_1^2\sigma_2^2}\frac{x^2}{(1+p)^2}\right) + \frac{1}{(1-p)^2}\exp\left(-\frac{\sigma_1^2+\sigma_2^2}{4\sigma_1^2\sigma_2^2}\frac{x^2}{(1-p)^2}\right) I_0\left(0, \frac{\sigma_2^2-\sigma_1^2}{4\sigma_1^2\sigma_2^2}\frac{x^2}{(1-p)^2}\right)\right]$$

Thus, using the above equations, both probabilities $p(M'(k), H_0)$, $p(M'(k), H_1)$ can be calculated. In the case of hypothesis H_0, for each $M'(k)$, the probability of $W(k) = 1$ equals the probability of $W(k) = -1$, $(P(W(k) = 1) = P(W(k) = -1) = 1/2)$. Therefore, the above equation (21) give the probabilities for hypothesis H_0. Hence,

$$p(M'(k); H_0) = \frac{M'(k)}{4\sigma_1\sigma_2}\cdot \quad (22)$$
$$\left[\frac{1}{(1+p)^2}\exp\left(-\frac{\sigma_1^2+\sigma_2^2}{4\sigma_1^2\sigma_2^2}\frac{M'(k)^2}{(1+p)^2}\right) I_0\left(0, \frac{\sigma_2^2-\sigma_1^2}{4\sigma_1^2\sigma_2^2}\frac{M'(k)^2}{(1+p)^2}\right) + \frac{1}{(1-p)^2}\exp\left(-\frac{\sigma_1^2+\sigma_2^2}{4\sigma_1^2\sigma_2^2}\frac{M'(k)^2}{(1-p)^2}\right) I_0\left(0, \frac{\sigma_2^2-\sigma_1^2}{4\sigma_1^2\sigma_2^2}\frac{M'(k)^2}{(1-p)^2}\right)\right]$$

In the case of hypothesis H_1, the signal is watermarked by the known watermark W. Thus, the probability is:

$$p(M'(k); H_1) = \frac{M'(k)}{2\sigma_1\sigma_2(1+W(k)p)^2}$$
$$\exp\left(-\frac{\sigma_1^2+\sigma_2^2}{4\sigma_1^2\sigma_2^2}\frac{M'(k)^2}{(1+W(k)p)^2}\right) I_0\left(0, \frac{\sigma_2^2-\sigma_1^2}{4\sigma_1^2\sigma_2^2}\frac{M'(k)^2}{(1+W(k)p)^2}\right). \quad (23)$$

Assuming independence between the samples of S, we conclude that:

$$p(\mathbf{M}'; H_j) = \prod_{i=0}^{N-1} p(M'(k); H_j) \ \ ,j = 0, 1 \quad (24)$$

Using equations 22 and 23, equation 17 is given by the equation 16

5. EXPERIMENTAL RESULTS

In this section, experiments are performed in order to verify the superiority of the proposed detector against the correlator. The experiments are performed on one dimensional digital signals.

In order to construct signals with the desired autocorrelation properties (7) we filter a random white uniformly distributed signal S of zero mean value with an IIR filter

$$H(z) = \frac{1-a}{1-az^{-1}}.$$

$$\sigma^2_{S_R(k)} = -\frac{1}{2}\sigma_s^2 \frac{-2a\cos\left(2\frac{\pi k}{N}\right)\left(2a^N(1+a^2)+a^2(N-2)-N-2\right)-N+a^4N-6a^2+6a^2a^N+2a^2\cos\left(4\frac{\pi k}{N}\right)\left(a^N-1\right)}{2a^2\cos\left(4\frac{\pi k}{N}\right)+4a^2-4a\cos\left(2\frac{\pi k}{N}\right)(1+a^2)+1+a^4} \tag{14}$$

$$\sigma^2_{X_I(k)} = -\frac{1}{2}\sigma_s^2 \frac{\left(-2a^2\cos\left(4\frac{\pi k}{N}\right)\left(a^N-1\right)-2aN\cos\left(2\frac{\pi k}{N}\right)(a^2-1)+N(a^4-1)+2a^2(a^N-1)\right)}{2a^2\cos\left(4\frac{\pi k}{N}\right)+4a^2-4a\cos\left(2\frac{\pi k}{N}\right)(1+a^2)+1+a^4} \tag{15}$$

$$L(M') = \prod_{k=1}^{N-1} \frac{\frac{2}{(1+W(k)p)^2} I_0\left(0, \frac{\sigma_2^2-\sigma_1^2}{4\sigma_1^2\sigma_2^2}\frac{M'(k)^2}{(1+W(k)p)^2}\right)}{\frac{1}{(1+p)^2}\exp\left(-\frac{\sigma_1^2+\sigma_2^2}{4\sigma_1^2\sigma_2^2}\frac{2p(W(k)-1)M'(k)^2}{(1+W(k)p)^2(1+p)^2}\right) I_0\left(0, \frac{\sigma_2^2-\sigma_1^2}{4\sigma_1^2\sigma_2^2}\frac{M'(k)^2}{(1+p)^2}\right) + \frac{1}{(1-p)^2}\exp\left(-\frac{\sigma_1^2+\sigma_2^2}{4\sigma_1^2\sigma_2^2}\frac{2p(W(k)+1)M'(k)^2}{(1+W(k)p)^2(1-p)^2}\right) I_0\left(0, \frac{\sigma_2^2-\sigma_1^2}{4\sigma_1^2\sigma_2^2}\frac{M'(k)^2}{(1-p)^2}\right)} \tag{16}$$

This filtering creates a signal having an autocorrelation function that has the form:

$$R_{SS}(k) = \frac{1-b}{1+b}\sigma_s^2 a^k$$

that is identical to eq.(7) for $\mu_s^2 = 0$. The variance of the filtered signal equals to $(1-a)/(1+a)\sigma_s^2$.

In the next step the embedding procedure is performed according to the formula (2). Then, the watermarked signal that arises is detected using correlator (eq.4) and the proposed detector (eq.16). In order to estimate both false alarm and false rejection probabilities both correct and erroneous keys have been used during detection.

The above procedure is executed for many different keys. Due to the Central Limit Theorem for products [8], the distribution of $L(x)$ is log-normal. Consequently, the distribution of $\log(L(x))$ is normal, where $\log(x)$ is the natural logarithm of x. In the following the proposed detector will be the $\log(L(x))$ instead of $L(x)$. Let $dr(x)$ and $de(x)$ be the distributions of the detector outputs for detecting correct and erroneous watermarks respectively. Due to the fact that both distributions obey the normal distribution, calculating the empirical mean and standard deviation by approximating the empirical pdf with a normal one, Receiver Operator Characteristic (ROC) curves for both detector outputs are constructed. ROC curves is the measure that will be used for comparing the detector qualities.

The above procedure is performed for several values of parameter a. In figure 1 the performance of the proposed detector against the correlator is shown. The value of the parameter a is 0.95. In this figure three ROC curves are shown. The first one represents the performance of correlator. The second and the third results for the proposed detector by using the unknown parameter a in the detection procedure or by estimating the parameter a respectively.

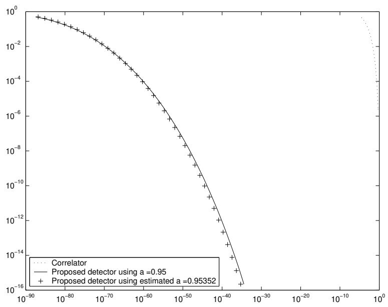

Figure 1: ROC curves of correlator , the proposed detector by using the known parameter a and the proposed detector after estimating the parameter a). $a = 0.95$

6. CONCLUSIONS AND FUTURE WORK

This paper deals with the statistical analysis of the behavior of a blind copyright protection watermarking system based on 1-D pseudorandom signals embedded in the magnitude of the Fourier transform of the data and the designing of an optimum detector. A multiplicative embedding method is examined and experimental results are performed in order to show the detector's efficiency against correlator. Expansion in higher data dimensions (2-D,3-D) are under consideration for future work.

REFERENCES

[1] G. Voyatzis and I. Pitas, "Protecting digital-image copyrights: A framework," *IEEE Computer Graphics and Applications*, vol. 19, no. 1, pp. 18–24, January/February 1999.

[2] G. Voyatzis and I. Pitas, "The use of watermark in the rotection of digital multimedia products," *IEEE Proceedings Special issue on Identification and protection of multimedia information*, vol. 87, no. 7, pp. 1197–1207, July 1999.

[3] Juan R. Hernandez, Martin Amado, and Fernando Perez-Gonzalez, "Dct-domain watermarking techniques for still images: Detector performance analysis and a new structure," *IEEE Transactions on Image Processing*, vol. 9, no. 1, pp. 55–68, January 2000.

[4] M.Barni, F.Bartolini, A.DeRosa, and A.Piva, "A new decoder for the optimum recovery of nonadditive watermarks," *IEEE Transactions on Image Processing*, vol. 10, no. 5, pp. 755–766, May 2001.

[5] G. Depovere, T. Kalker, and J.-P. Linnartz, "Improved watermark detection reliability using filtering before correlation," in *Proc. of ICIP'98*, 4-7 October 1998, vol. I, pp. 430–434.

[6] Patrick Billingsley, *Probability and measure*, Wiley, 1995.

[7] S.M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*, Prentice Hall, NJ, USA, 1998.

[8] Athanasios Papoulis, *Probability, random variables, and stochastic proccesses*, McGraw-hill, New York, 1991.