OPTIMAL WATERMARK DETECTORS FOR TRANSFORM DOMAIN WATERMARKING

Vassilios Solachidis *Athanasios Nikolaidis* *Ioannis Pitas*

Department of Informatics
Aristotle University of Thessaloniki
Box 451, Thessaloniki 540 06, GREECE
E-mail: `pitas@zeus.csd.auth.gr`

ABSTRACT

This paper deals with the construction of optimal watermark detectors for watermarks embedded in a transform domain of a signal. Discrete Cosine (DCT), Discrete Wavelet (DWT) and Discrete Fourier transforms (DFT) are included in the analysis. Experimental results are presented in order to show the proposed detectors' efficiency as opposed to that of the correlator detector.

1. INTRODUCTION

Digital services development created new requirements for multimedia security and copyright protection techniques. The risk of illegal copying, reproduction and distribution of copyrighted material is becoming more threatening with the all-digital solutions adopted by content providers, system designers and users, thus rendering watermark-based protection an essential requirement for multimedia distribution.

Robustness is one of the most important issues of a watermarking scheme. It is obvious that the detector performance significantly affects the method's robustness. Therefore, the detector has to be designed appropriately in order to increase watermark robustness and to decrease false alarms. Correlation based detection of watermarked signals is employed in the majority of watermarking techniques in the literature. However, the correlator-detector is optimal and minimizes error probability only when the signal samples are white Gaussian noise. There are papers in the literature that propose detectors, different than the correlator, aiming to deal with cases in which the host data do not follow a Gaussian distribution [6]-[7].

In [6] the embedding domain is the DCT. The DCT coefficient 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 [7] the watermark is embedded in the magnitude of the DFT domain. In this paper the authors assume, that the Fourier magnitude follows the Weibull distribution, due to the fact that it is a non-zero distribution only along the positive axis and represents a large distribution family.

In the present paper, we overview two nonlinear detectors that have been recently proposed for watermarking in transform domains [1, 8, 12]. Section 2 describes the detector developed for DCT and DWT domain watermarking, while Section 3 is concerned with the DFT domain detector. Experimental results for both detectors are provided in Section 4 and, finally, certain conclusions about the detector performance are drawn in Section 5.

2. OPTIMAL DETECTION IN THE DCT AND DWT DOMAINS

2.1. Detector structure

In the case of watermark embedding in the DCT and DWT domains ([1]) we have adopted the additive watermarking model, which assumes that the watermark is modulated by a factor corresponding to the watermark power.

The need to introduce a watermark detector other than the well-known correlator detector arises from the fact that, in practice, most image transform coefficients cannot be considered as white Gaussian noise. This leads to the conclusion that watermarking algorithms that are based on a correlator-detector perform only suboptimally. In order to tackle with this problem, we have developed a more sophisticated detector structure in conjunction with a better, image-dependent approximation of the pdf of the embedding domain coefficients.

We assume that all the DCT or DWT transform coefficients $X[\boldsymbol{k}]$, in which the watermark is embedded, follow a generalized Gaussian distribution with zero mean value and the same parameters β and σ^2 [2]:

$$p(X[\boldsymbol{k}]) = \frac{c_1(\beta)}{\sqrt{\sigma^2}} \exp\left(-c_2(\beta)\left|\frac{X[\boldsymbol{k}]}{\sqrt{\sigma^2}}\right|^{\frac{2}{1+\beta}}\right) \tag{1}$$

where $c_1(\beta)$ and $c_2(\beta)$ are constant expressions β is called the *shape* parameter ($\beta > -1$) and σ^2 is the variance. This assumption has been proven to hold both for DCT and DWT domain coefficients, and has already been employed in coding applications [3].

It has also been proven in [4] that, for such a problem, a Rao test or, equivalently, a generalized likelihood ratio test (GLRT), has asymptotically optimal performance (i.e. it is indeed optimal for large images). This test can be written as:

$$D_R(Y) = \frac{\left[\sum W[\boldsymbol{k}]\frac{p'(Y[\boldsymbol{k}])}{p(Y[\boldsymbol{k}])}\right]^2}{\frac{1}{N}\sum W^2[\boldsymbol{k}]\sum\left[\frac{p'(Y[\boldsymbol{k}])}{p(Y[\boldsymbol{k}])}\right]^2} \tag{2}$$

where $Y[\boldsymbol{k}]$ are the watermarked coefficients, $W[\boldsymbol{k}]$ are the watermark values, N is the number of coefficients that have been watermarked, p is the pdf and p' is the derivative of the pdf of the image transform coefficients under hypothesis H_0 (watermark not present).

Let us now consider the case of a known binary watermark ($W[\boldsymbol{k}] \in \{-1, 1\}$). Let us also assume that the possibly watermarked transform coefficients approximately follow a generalized

Gaussian distribution with the same parameters as the original image transform coefficients. This can easily be proven to hold for small values of watermark power. Thus, parameter estimation can be performed using the watermarked image transform coefficients instead of the original, which are not available in the detection stage.

Formula (2) can be rewritten as:

$$D_R(Y) = \frac{\left(\sum sgn(Y[\boldsymbol{k}])W[\boldsymbol{k}]|Y[\boldsymbol{k}]|^{\frac{1-\beta}{1+\beta}}\right)^2}{\sum |Y[\boldsymbol{k}]|^{\frac{2(1-\beta)}{1+\beta}}} \quad (3)$$

where $sgn(x)$ is the signum function and β is the shape parameter of the generalized Gaussian pdf. It has been shown that, for large data records, this detector follows a chi-squared distribution under both detection hypotheses. Since a random variable that follows a $\chi'^2(1, \lambda)$ distribution is equivalent to the square of a normal random variable with mean value $\sqrt{\lambda}$ and variance equal to 1, it follows that the probabilities of false alarm and false rejection (P_{FA} and P_{FR}) which characterize the detector performance are given by:

$$\begin{aligned} P_{FA} &= 2Q(\sqrt{\gamma'}) \\ P_{FR} &= 1 - Q(\sqrt{\gamma'} - \sqrt{\lambda}) - Q(\sqrt{\gamma'} + \sqrt{\lambda}) \end{aligned} \quad (4)$$

where γ' is the detection threshold. The watermark detection performance depends on the so-called *non-centrality parameter* λ. For a certain range of values of the parameter β ($-1 < \beta < 3$), parameter λ can be defined as:

$$\lambda = N\alpha^2 \frac{4\Gamma(\frac{3-\beta}{2})\Gamma(\frac{3}{2}(1+\beta))}{(1+\beta)^2\sigma^2\Gamma^2(\frac{1}{2}(1+\beta))} \quad (5)$$

where α is the watermark power and $\Gamma(x)$ is the Gamma function. It is interesting to see that P_{FR} decreases monotonically with λ. This means that, for increased watermark power α or for larger sample number N, the probability of false alarm decreases, as expected.

2.2. Generalized Gaussian parameter estimation

As derived in formula (3), the detector formula involves the value of the parameter β of the generalized Gaussian approximation of the transform coefficients. We wish to estimate this parameter in a computationally efficient way. For this reason, the technique proposed in [5] is adopted. This is based on exploiting the relation between the variance σ_Y^2, the mean of the absolute values $E\left[|Y|\right]$ and the shape parameter β_Y:

$$r(\beta_Y) = \frac{\sigma_Y^2}{E^2\left[|Y|\right]} = \frac{\Gamma\left(\frac{1}{2}(1+\beta_Y)\right)\cdot\Gamma\left(\frac{3}{2}(1+\beta_Y)\right)}{\Gamma^2(1+\beta_Y)}, \quad (6)$$

also called the *generalized Gaussian ratio function*. After computing the estimate of the variance $\hat{\sigma}_Y^2$ using the sample variance, and the estimate of the mean of the absolute values using $\hat{E}\left[|Y|\right] = (1/N)\sum_{i=1}^{N}|Y_i - \hat{\mu}_Y|$, we calculate the ratio $\rho = \hat{\sigma}_Y^2/\hat{E}^2\left[|Y|\right]$ and finally solve the equation (6) for β_Y using e.g. a lookup table.

2.3. Embedding in DCT and DWT domains

The dataset on which we embed the watermark in the case of DCT domain watermarking is the set of the non-DC coefficients resulting from a 8 × 8-block DCT transformation. We estimate the parameters of the generalized Gaussian distribution based on all coefficient samples. We can choose to use a band of low-to-middle frequency coefficients for watermark embedding in order to achieve robustness to lowpass filtering and compression.

As far as DWT domain watermarking is concerned, all frequency bands but the lowest one can be fairly well represented by generalized Gaussian distributions. The 2-D DWT decomposes an image into space-frequency subbands by applying lowpass and corresponding highpass filters to the original image at each dimension and subsequently downsampling the result by a factor of 2. In this way, the so-called *detail* images are produced, as well as a smoothed image. This can be repeatedly performed up to the desired resolution level.

3. OPTIMAL DETECTOR IN THE DFT DOMAIN

3.1. Watermarking system model

In the DFT case [8] the proposed detector can be applied on a host multidimensional signal of any dimension D. Let $\mathbf{s} = s(i_1, i_2, ..., i_D) = s(\mathbf{i})$ be the original signal, where D is the signal dimension. Let also $S(\mathbf{k})$ be the Fourier transform of $s(\mathbf{i})$ and $M(\mathbf{k})$, $P(\mathbf{k})$ be the magnitude of the Fourier transform ($M(\mathbf{k}) = |S(\mathbf{k})|$) and its phase, respectively. The embedding procedure can be either additive or multiplicative.

3.2. Probability density function of the magnitude of DFT coefficients

In order to extract an optimal detector, the probability density function (pdf) of the magnitude of the DFT coefficients has to be calculated. We assume that the samples of the host signal can be modelled as independent identically distributed (i.i.d.) random variables obeying a distribution function with the following properties:

$$E(s(\mathbf{i})) = E(s_r(\mathbf{i})) + i\ E(s_i(\mathbf{i})) = \mu_{s_r} + i\ \mu_{s_i},\ \forall \mathbf{i} \in \mathbf{N} \quad (7)$$

$$E(s(\mathbf{i})s(\mathbf{i}+\mathbf{d})) = 0,\ \forall \mathbf{i}, \mathbf{d} \in \mathbf{N}, \mathbf{d} \neq \mathbf{0} \quad (8)$$

$$Var(s_r(\mathbf{i})) = \sigma_r^2, \quad Var(s_i(\mathbf{i})) = \sigma_i^2, \quad \forall\ \mathbf{i} \quad (9)$$

where $E(\cdot)$ is the expectation and Var$(\cdot)$ is the variance operator.

In the sequel, the distribution of each coefficient $S(\mathbf{k})$ will be calculated. Assuming that some of the $N_1, N_2, ..., N_D$ (at least one) are greater than 30, we can conclude, according to the Central Limit Theorem, that the distribution of each DFT coefficient is Gaussian [9]. Thus, the statistics of the DFT coefficients can be easily calculated. The mean of all the coefficients except the DC term equals zero. The variances of the real and the imaginary part for each coefficient are equal for almost all $\mathbf{k}$. Thus, both the real and the imaginary part of the Fourier coefficients follow Gaussian distributions: $S_R(\mathbf{k}) \sim N(0, \sigma_R^2)$ and $S_I(\mathbf{k}) \sim N(0, \sigma_I^2)$. Therefore, since $\sigma_R(\mathbf{k}) = \sigma_I(\mathbf{k}) \stackrel{\triangle}{=} \sigma^2$, the pdf $f_M(x)$ of $M(\mathbf{k}) = |S(\mathbf{k})| = \sqrt{S_R^2(\mathbf{k}) + S_I^2(\mathbf{k})}$ follows the Rayleigh distribution [10]:

$$|S(\mathbf{k})| \sim f_M(x) = \frac{x}{\sigma^2}\exp\left(-\frac{x^2}{2\sigma^2}\right),\ x > 0. \quad (10)$$

3.3. Likelihood ratio test (LRT)

Since the pdf of the magnitude of the DFT coefficients $M(k)$ is known, an optimal detector can be computed using the likelihood ratio test (LRT). According to the Neyman-Pearson theorem [10] and in order for the probability of correct detection to be maximized, the following quantity has to be computed:

$$L(\mathbf{M}') = \frac{p(\mathbf{M}'; H_1)}{p(\mathbf{M}'; H_0)} > T, \tag{11}$$

where $\mathbf{M}'$ is the vector of all the watermarked samples $M'(\mathbf{k})$, H_0, H_1 are the detection hypotheses (H_0: the data are not watermarked or are watermarked by another watermark W', H_1: the data are watermarked by the watermark W).

In the sequel, both pdfs $p(\mathbf{M}', H_0), p(\mathbf{M}', H_1)$ will be computed in order to find the optimal detector. For the hypothesis H_0 we shall adopt the case that the data is watermarked by another (unknown) watermark W' because this is the worst of the two subcases since, in this case, the signal variance is greater. Since the watermark consists of equiprobable values ± 1, it can be easily proven that, in the case of an additive watermark, the pdf $f_{M'_{add}}(x)$ of the watermarked signal is equal to [11]:

$$f_{M'_{add}}(x) = \frac{1}{2}[f'_M(x-\alpha) + f'_M(x+\alpha)], \ \ x > 0 \tag{12}$$

Thus, using formula (12), both pdfs can be calculated. In the case of hypothesis H_0: $P(W(\mathbf{k}) = 1) = P(W(\mathbf{k}) = -1) = 1/2$. Hence for the additive case, the probability that $M'(\mathbf{k})$ is watermarked by another unknown watermark is given by:

$$\begin{aligned} p_{add}(M'(\mathbf{k}); H_0) = \frac{1}{2}\Bigg[& \frac{M'(\mathbf{k}) - \alpha}{\sigma^2} \exp\left(-\frac{(M'(\mathbf{k}) - \alpha)^2}{2\sigma^2}\right) \\ & + \frac{M'(\mathbf{k}) + \alpha}{\sigma^2} \exp\left(-\frac{(M'(\mathbf{k}) + \alpha)^2}{2\sigma^2}\right)\Bigg] \end{aligned} \tag{13}$$

In the case of hypothesis H_1, the signal $M'(\mathbf{k})$ is watermarked by the *known* watermark W. Thus, the probability for the additive case is:

$$p_{add}(M'(\mathbf{k}); H_1) = \tag{14}$$

$$\frac{M'(\mathbf{k}) + W(\mathbf{k})\alpha}{\sigma^2} \exp\left(-\frac{(M'(\mathbf{k}) + W(\mathbf{k})\alpha)^2}{2\sigma^2}\right) \tag{15}$$

We assume that only a subset of the DFT coefficients $M(\mathbf{k}),\ \mathbf{k} \in \mathbf{A}\ \mathbf{A} = \{\mathbf{k} \in \mathbf{N} : \mathbf{W}(k) \neq 0\}$ is watermarked, for reasons to be explained in section 4.2. Assuming independence between the samples of $M'(\mathbf{k})$, we conclude that:

$$p(\mathbf{M}'; H_j) = \prod_{\mathbf{k} \in \mathbf{A}} p(M'(\mathbf{k}); H_j) \ \ , j = 0, 1 \tag{16}$$

Using formulae (13) and (14), the optimal detector (11) takes the form:

$$\prod_{\mathbf{k} \in \mathbf{A}} \frac{2(M'(\mathbf{k}) + W(\mathbf{k})\alpha)}{\begin{pmatrix} (M'(\mathbf{k}) - \alpha) \exp\left[\frac{M'(\mathbf{k})\alpha(1+W(\mathbf{k}))}{\sigma^2}\right] + \\ (M'(\mathbf{k}) + \alpha) \exp\left[\frac{M'(\mathbf{k})\alpha(W(\mathbf{k})-1)}{\sigma^2}\right] \end{pmatrix}} \geq T \tag{17}$$

The above formulae can be easily calculated also for the multiplicative embedding case. If the aforementioned independence assumption (16) is not valid, the derived detectors are suboptimal.

The estimator of the Rayleigh parameter σ is given by:

$$\hat{\sigma} = \sqrt{\frac{\sum_{\mathbf{k} \in A} M(\mathbf{k})^2}{2|\mathbf{A}|}}. \tag{18}$$

where $|\mathbf{A}|$ denotes the cardinality of set $\mathbf{A}$. Unfortunately, $M(\mathbf{k})$ is unknown during blind watermark detection. Therefore, $M'(\mathbf{k})$ is used instead of $M(\mathbf{k})$ in (18).

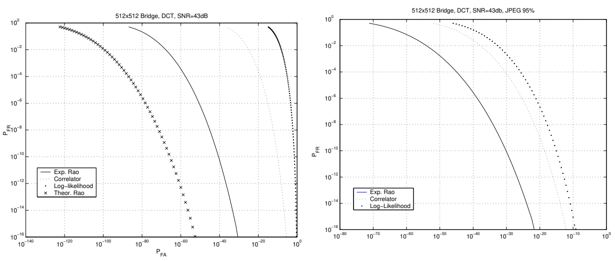
The optimal detector analysis has been performed also for colored signal models [12].

4. EXPERIMENTAL RESULTS

4.1. DCT and DWT domains

In order to confirm the superiority of the new detector compared to the classic correlator, and to the correlator proposed in [2], we conducted a number of experiments on several real images and for several SNR values. Experiments were performed by embedding different binary pseudo-random watermarks in the coefficients of either the DCT or the DWT domain, and afterwards detecting them by the proposed detector, the correlator and the detector proposed in [2], having always in mind that the watermark power is not known at the detection stage. For this reason, a constant value of 1 is implied for watermark power in the case of the detector in [2]. The performance of all detectors was evaluated on undistorted images as well as JPEG compressed images (in the case of DCT embedding) and SPIHT compressed images (in the case of DWT embedding).

An example of a result for DCT domain embedding is shown in Figure 1. The original image is the BRIDGE image of size 512×512. The watermarked one has SNR≃43dB. Figure 1(a) shows the ROC curves (plots of P_{FA} versus P_{FR} for several detection thresholds) for all considered detectors, when no attack is imposed. The superiority of the proposed detector is obvious, whereas the other two methods exhibit similar performance. A curve that is derived theoretically, based on the estimated values of the pdf parameters, is delineated in the same Figure with crosses. After a moderate JPEG compression attack of quality factor 95%, the performance of all detectors degrades. However, the new detector still demonstrates a better performance, as it can be seen in Figure 1(b).

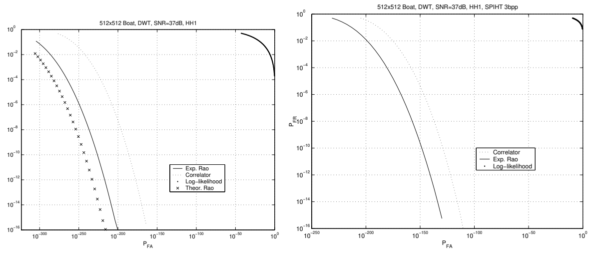


(a) (b)

Figure 1: (a) ROCs without attack for watermarked Bridge image (SNR≃43dB); (b) ROCs after JPEG compression (95%).

In the case of DWT domain embedding, the watermark is embedded in all coefficients that correspond to subbands of a 2-level transformation except for the LL2 subband. Figure 2(a) shows the

ROC curves resulting from the watermark detection in subband HH1 of the BOAT image for SNR$\simeq$37dB, for additive embedding in the subbands of the 2-level DWT transformation. The new detector displays, again, better performance, especially when compared to the log-likelihood-based detector. It is also superior to the correlator. The same observations also hold after a moderate SPIHT compression (3bpp), as is shown in Figure 2(b).



(a) (b)

Figure 2: (a) ROCs without attack for watermarked Boat image (SNR$\simeq$37dB); (b) ROCs after SPIHT compression (3bpp).

4.2. DFT domain

Simulation experiments have been performed in order to verify the superiority of the proposed detector (17) against the classic correlator. The experiments have been performed on many digital images. The gray scale 512×512 LENNA image was used as a host image in this presentation. The watermark has been embedded in the middle frequencies for to the following reasons: a) modifications in the low frequencies of the Fourier transform will cause visible changes in the spatial domain, b) compression affects mostly the high frequencies of the Fourier transform. Considering that the zero frequency term (DC term) is in the center of the transform domain, the watermark is embedded in a ring that covers the middle frequencies.

The watermark embedding procedure has been performed either additively or multiplicatively. Then, the detector output has been calculated using the proposed optimal detector and the classical correlator. In order to calculate the false alarm and false rejection probabilities experimentally, 100 correct and erroneous keys have been used during detection. The watermarked image was altered by intentional attacks for the examination of the watermark robustness against attacks.

The watermark power used in the experiments is chosen to be $\alpha = 1000$ in case of additive and $\alpha = 0.1$ in case of multiplicative embedding. The watermark is invisible and the watermarked image PSNR is 46 dB (for both embedding types). As shown in Figure 3, the performance of the proposed detector, measured by the ROC curves, is much better than that of the correlator. It has also been proven by experiments that the proposed algorithm is very robust to JPEG compression and median filtering compared to the correlator detector.

5. CONCLUSIONS

This paper deals with the design of optimal detectors for additive or multiplicative watermarking in the DCT, DWT and DFT domains. Experiments have been performed in order to display the superiority of the proposed detectors against the classical correlator detector.

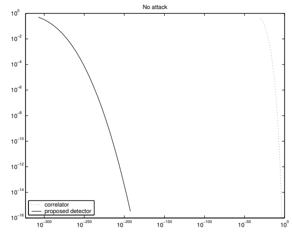


Figure 3: ROC curves for both detectors (multiplicative embedding)

6. REFERENCES

[1] A. Nikolaidis and I. Pitas, "Asymptotically Optimal Detection for Additive Watermarking in the DCT and DWT Domains", in *IEEE Trans. on Image Processing*, **12**(5), pp. 563-571.

[2] J. R. Hernández, M. Amado and F. Pérez-González, "DCT-Domain Watermarking Techniques for Still Images: Detector Performance Analysis and a New Structure", in *IEEE Trans. on Image Processing*, **9**(1), pp. 55-68.

[3] R. C. Reininger and J. D. Gibson, "Distributions of the two-dimensional DCT coefficients for images", in *IEEE Trans. on Communications*, **31**(6), pp. 835-839, 1983.

[4] S. M. Kay, "Asymptotically Optimal Detection in Incompletely Characterized Non-Gaussian Noise", in *IEEE Trans. on Acoustics, Speech and Signal Processing*, **37**(5), pp. 627-633, 1989.

[5] K. Sharifi and A. Leon-Garcia, "Estimation of Shape Parameter for Generalized Gaussian Distributions in Subband Decompositions of Video", in *IEEE Trans. on Circuits and Systems for Video Technology*, **5**(1), pp. 52-56.

[6] 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.

[7] 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.

[8] V.Solachidis and I.Pitas, "Non-linear detector of watermarked signal embedded in fourier domain," in *EURASIP Workshop on,Nonlinear Signal and Image Processing (NSIP 2003)*, Grado - Gorizia, Italy, June 8-11 2003.

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

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

[11] A. Papoulis, *Probability & Statistics*, Prentice Hall, 1991.

[12] V.Solachidis and I.Pitas, *Optimal detection for multiplicative watermarks embedded in DFT domain*, IEEE International Conference on Image Processing (ICIP 2003), Barcelona, Spain, 14-17 September 2003