Asymptotically optimal detection for
additive watermarking in the DCT and
DWT domains

Athanasios Nikolaidis    Ioannis Pitas *
Department of Informatics
Aristotle University of Thessaloniki
Box 451, Thessaloniki 540 06, GREECE Tel,Fax: +3031-996304
E-mail: pitas@zeus.csd.auth.gr

* Corresponding Author
Abstract

Most of the watermarking schemes that have been proposed until now employ a correlation detector (matched filter). The current paper proposes a new detector scheme that can be applied in the case of additive watermarking in the DCT (Discrete Cosine Transform) or DWT (Discrete Wavelet Transform) domain. Certain properties of the probability density function of the coefficients in these domains are exploited. Thus, an asymptotically optimal detector is constructed based on well known results of the detection theory. Experimental results prove the superiority of the proposed detector over the correlation detector.

EDICS number: 5-AUTH
I. INTRODUCTION

Image watermarking as a tool for copyright protection has attracted a lot of attention in the last few years [1]-[5]. The main trend is to superimpose the watermark on the image in either an additive or a multiplicative way either in the spatial or in transform domains [6]-[9]. In either case, the detection stage involves a simple correlator for the majority of the reported techniques. This is done under the general implicit assumption that the image pixels or coefficients on which the watermark is embedded can be modeled as white, i.i.d. Gaussian noise, and that the watermark is a deterministic signal, independent of the image and synchronized with it. If this is indeed the case, the correlator is the optimal detector in the sense that, under a constant probability of false alarm, the probability of false rejection is minimized [10].

However, two major errors may occur, while using unconditionally the correlator as a watermark detector: a) the detector structure of the correlator is employed without previously resorting to a binary hypothesis model (watermark exists/does not exist) and applying a proper hypothesis test, b) the data distributions are neither i.i.d. nor can they be adequately well described by the Gaussian pdf (probability density function) as is frequently the case. In fact, there is no way to describe any image as a set of data samples that comply with a certain pdf.

An attempt to use the spatial domain pdf in watermarking is proposed in [11]. A whitening stage is introduced before detection. The data samples are assumed to follow a Cauchy distribution afterwards. An adequately good estimation of the data pdf can only be made in some transform domains (e.g. DFT, DCT, DWT). In this case, the pdf of certain transform coefficients cannot be modeled as Gaussian ones. Attempts to model the pdf of transform coefficients correctly and to render the detector structure compatible with the presumed watermark binary hypothesis testing have only been done lately. In [12], a Weibull distribution is assumed for the magnitude of the DFT domain coefficients, and an optimum decoder is derived, based on the Neyman-Pearson criterion, while assuming a multiplicative embedding rule, i.e. when the watermark power that is embedded on each coefficient is proportional to the magnitude.
of the respective coefficient. Apart from this, a similar rationale is adopted in [13], where the embedding domain is that of a band of DCT coefficients excepting the DC term. A generalized Gaussian distribution, as well as an additive embedding rule are assumed in this case. The ML (maximum likelihood) criterion is employed in order to derive a detector structure that is optimal under the assumption that the watermark power is known [13].

The present paper proposes a detector structure that displays superior performance compared to other detectors, including the correlator, under certain assumptions. Watermarking in the DCT and DWT domains is considered, and the pdf assumption for the coefficients in these domains is the same as in [13]. In order to derive the optimal detector structure, a Rao test that is equivalent with a GLRT (generalized likelihood ratio test) is employed. The resulting detector is asymptotically optimal, meaning that it is optimal under the assumption of a large data record. The derivation of the optimal detector, given the additive embedding model and assuming knowledge of the watermark except for the watermark power, is presented in Section II. In Section III, a description of the algorithms involved in estimating the generalized Gaussian distribution parameters is presented. Sections IV and V present the specifics for watermark embedding in the DCT and DWT domains, respectively. In Section VI, experimental results confirm the superiority of the proposed detector compared to the correlator, even in the presence of compression attacks. Finally, certain conclusions about the performance of the proposed detector are drawn in Section VII.

II. OPTIMAL DETECTOR STRUCTURE

As explained in the introduction, the embedding model is that of an additive watermark inserted in a transform domain (e.g. DCT, DWT) and modulated by a factor corresponding to the watermark power, which is unknown during detection. Thus, the hypothesis test under consideration throughout this paper will be of the form:

\[ H_1 : Y[k] = X[k] + \alpha W[k] \]
\[ \mathcal{H}_0 : Y[k] = X[k] \]  

where \( X \) is the original transformed image that follows a certain pdf model, \( W \) is the watermark that is embedded in the original transformed image, and \( \alpha \) is an amplitude parameter that corresponds to a watermark power. The statistics of \( X \) and \( \alpha \) are generally unknown during watermark detection.

It is generally agreed that the performance of a watermarking scheme relies heavily on the design of the detector. Most of the methods that have been proposed until now employ a correlator-detector that demonstrates an acceptable performance in many applications. This detector is easy to implement and is based on a simple correlation of the possibly watermarked (or even attacked) image and the watermark that is probably embedded on the image:

\[ D(Y) = \sum_k Y[k] W[k] \]  

where \( k \) are the indices of the transform coefficients where the watermark is supposed to be embedded, \( Y \) is the watermarked (and possibly attacked) image, and \( W \) is the two-dimensional watermark. The probability distribution of the correlator-detector is assumed to be normal under both hypotheses \( \mathcal{H}_1 \) (the watermark under test exists) and \( \mathcal{H}_0 \) (no watermark exists). Accordingly, the probabilities of false alarm (\( P_{FA} \)) and false rejection (\( P_{FR} \)) are defined as:

\[ P_{FA} = Pr \{ D > \gamma'; \mathcal{H}_0 \} = Q \left( \frac{\gamma'}{\text{std}(D; \mathcal{H}_0)} \right) \]
\[ P_{FR} = Pr \{ D < \gamma'; \mathcal{H}_1 \} = 1 - Pr \{ D > \gamma'; \mathcal{H}_1 \} = 1 - Q \left( \frac{\gamma' - E(D; \mathcal{H}_1)}{\text{std}(D; \mathcal{H}_1)} \right) \]

where \( \gamma' \) is the decision threshold, \( E(D; \mathcal{H}_1) \) is the mean value of the detector under hypothesis \( \mathcal{H}_1 \), \( \text{std}(D; \mathcal{H}_0) \) and \( \text{std}(D; \mathcal{H}_1) \) are the standard deviations of the detector under hypotheses \( \mathcal{H}_0 \) and \( \mathcal{H}_1 \), respectively, and:
\[ Q(x) = \int_x^{\infty} \frac{1}{\sqrt{2\pi}} \exp \left( -\frac{1}{2}t^2 \right) \, dt \]  

If \( W[k] \in \{-1, 1\} \), then \( E(D; H_1) = \alpha N \) and \( \text{std}(D; H_0) = \text{std}(D; H_1) = \sigma \sqrt{N} \), where \( N \) is the number of data samples and \( \sigma^2 \) is the variance of the original transformed image. The detector performance is measured by using the receiver operating characteristics curve (ROC), which displays \( P_{FA}, P_{FR} \) for various detection thresholds \( \gamma' \). According to the detection theory, this detector is expected to be optimal, that is, to present the lowest possible equal error rate (EER), in the case that the model noise (in our case, the original image) is i.i.d. WGN (White Gaussian Noise), meaning that it has a white spectrum and follows the normal distribution:

\[ p(X[k]) = \frac{1}{\sqrt{2\pi}\sigma^2} \exp \left[ -\frac{1}{2\sigma^2}(X[k] - \mu)^2 \right] \]  

where \( \mu \) is the mean value and \( \sigma^2 \) is the variance of the distribution. Under this assumption, and assuming that the watermark is deterministic, the correlator-detector proves to perform better than any other detector. However, in practice, most transformed images cannot be considered as white Gaussian noise. On the contrary, different images may have completely different spectral content and the gray levels or the transform coefficients of the image may not comply with the assumption of a Gaussian distribution. This leads to the conclusion that, in reality, most of the existing watermark detectors perform only suboptimally. Thus, a more sophisticated detector structure in conjunction with a better, image-dependent approximation of the pdf of the embedding domain coefficients is necessary in order to achieve superior detector performance.

A step towards the design of optimal detection has been done in [13], where a detector is designed based on the ML decision rule and on the fact that the DCT transform coefficients, in which the watermark is embedded, follow a generalized Gaussian distribution with zero mean value:
\[ p(X[k]) = \frac{c_1(\beta)}{\sqrt{\sigma^2}} \exp \left( -c_2(\beta) \left| \frac{X[k]}{\sqrt{\sigma^2}} \right|^\beta \right) \]  

(6)

where:

\[ c_1(\beta) = \frac{\Gamma \left( \frac{3}{2} (1 + \beta) \right)}{(1 + \beta) \Gamma \left( \frac{1}{2} (1 + \beta) \right)} \]  

(7)

\[ c_2(\beta) = \frac{\left[ \Gamma \left( \frac{3}{2} (1 + \beta) \right) \right]^{\frac{1}{\beta + \gamma}}}{\left[ \Gamma \left( \frac{1}{2} (1 + \beta) \right) \right]^{\frac{1}{\beta + \gamma}}} \]  

(8)

and \( \beta > -1 \) (called the shape parameter). \( \sigma^2 \) is the variance and \( \Gamma(x) \) is the Gamma function:

\[ \Gamma(x) = \int_0^\infty t^{x-1} \exp(-t)dt \]  

(9)

The generalized Gaussian pdf reduces to the Gaussian for \( \beta = 0 \) and to the Laplacian for \( \beta = 1 \), while it tends to the uniform pdf for \( \beta \to -1 \). A generalized Gaussian pdf for \( \beta = 1.5 \) and \( \sigma = 1 \) is shown in Figure 1. It has been proven that both DCT and DWT domain coefficients follow this pdf, has already been investigated in coding applications [14]-[15]. The same assumption will be used subsequently.

![Generalized Gaussian pdf for \( \beta = 1.5 \) and \( \sigma = 1 \).](image)

Fig. 1. Generalized Gaussian pdf for \( \beta = 1.5 \) and \( \sigma = 1 \).

Kay [16] has proven that, for such a problem, the Rao test has asymptotically optimal performance that is equivalent to that of a generalized likelihood ratio test (GLRT). This means
that, if the noise pdf is symmetric, then the performance of this detector is equal to that of a 
clairvoyant GLRT (one that is designed with a priori knowledge of the noise parameters). The 
so-called Rao detector can be written as:

\[ D_R(Y) = \frac{\left( \sum W[k] \frac{p'(Y[k])}{p(Y[k])} \right)^2}{\frac{1}{N} \sum W^2[k] \sum \left[ \frac{p'(Y[k])}{p(Y[k])} \right]^2} \]  (10)

where \( N \) is the number of samples that have been watermarked, \( p \) is the pdf and \( p' \) is the 
derivative of the pdf of the image under hypothesis \( H_0 \). Since the detector 10 is asymptotically 
optimal, practically it is indeed optimal for large images.

Let us now consider the case of a known binary watermark (\( W[k] \in \{-1, 1\} \)). Let us also assume 
that the possibly watermarked and transformed image approximately follows a generalized 
Gaussian distribution with the same parameters as that of the original transformed image. This 
assumption can easily be proven, since the watermarked transformed image is a linear function 
of the original transformed image according to hypothesis \( H_1 \), as can be seen in Equation (1):

\[ p_Y(Y[k]) = p_X(Y[k] - \alpha W[k]) \]  (11)

Thus, the watermarked coefficients also comply with a generalized Gaussian distribution, with 
a slight shift in the mean value, due to watermark power \( \alpha \). Thus, our GWGN assumption 
for the watermarked coefficients is valid, and parameter estimation can be performed using the 
watermarked image instead of the original one, which is not available in the detection stage anyway.

Formula (10) can be rewritten as:

\[ D_R(Y) = \frac{\left( \sum sgn(Y[k]) W[k] Y[k] \left| \frac{1 - \beta}{1 + \beta} \right| \right)^2}{\sum |Y[k]| \left| \frac{1 - \beta}{1 + \beta} \right|} \]  (12)

where \( sgn(x) \) is the signum function and \( \beta \) is the shape parameter of the generalized Gaussian 
pdf. It has been shown ([18]-[19]) that, for large data records, this detector follows a chi-squared
distribution under both detection hypotheses. More specifically, the distribution of the detector under $\mathcal{H}_0$ is $\chi^2$, that is a chi-squared distribution with one degree of freedom, whereas, under $\mathcal{H}_1$, it is $\chi^2(1, \lambda)$, that is a non-central chi-squared distribution with one degree of freedom and non-centrality parameter $\lambda$:

\[
D_R(Y) \sim \chi^2 \quad \text{under } \mathcal{H}_0
\]
\[
D_R(Y) \sim \chi^2(1, \lambda) \quad \text{under } \mathcal{H}_1
\]  

(13)

The performance of this detector is the same as that of the generalized log-likelihood ratio test ([16]). 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:

\[
P_{FA} = 2Q(\sqrt{\gamma'})
\]
\[
P_{FR} = 1 - Q(\sqrt{\gamma' - \sqrt{\lambda}}) - Q(\sqrt{\gamma' + \sqrt{\lambda}})
\]  

(14)

where $\gamma'$ is the detection threshold. The watermark detection performance depends on the so-called non-centrality parameter $\lambda$ defined as:

\[
\lambda = N\alpha^2 i(\alpha) = N\alpha^2 \int_{-\infty}^{\infty} \frac{(dp/w)^2}{p(w)} dw
\]  

(15)

where $i(\alpha)$ is the Fisher information, also called the intrinsic accuracy, of a DC level $\alpha$ in generalized Gaussian noise with pdf $p(w)$, based on a single sample. In general, no closed form of the Fisher information exists. However, for the case of the generalized Gaussian pdf for a certain range of values of the parameter $\beta (-1 < \beta < 3)$, the intrinsic accuracy is given by
\begin{equation}
\hat{\beta}(\alpha) = \frac{4\Gamma\left(\frac{3-\beta}{2}\right)\Gamma\left(\frac{2}{\beta}(1+\beta)\right)}{(1+\beta)^{3/2}\sigma^2 \Gamma^2\left(\frac{1}{2}(1+\beta)\right)}
\end{equation}

where the notation \( \beta \) is used instead of \( \beta_Y \) for notation similarity. Of course, this is a biased estimate, since it considers all watermarked coefficients as i.i.d. random variables. However, it is the best approximation we can have in order to assess the experimental performance of the new detector.

It is interesting to see that \( P_{FR} \) decreases monotonically with \( \lambda \). This means that, for increased watermark power \( \alpha \) or for larger sample number \( N \), the probability of false alarm decreases, as expected. Apart from that, the detector reaches its worst performance when the generalized Gaussian distribution reduces to a Gaussian one, since the intrinsic accuracy \( \hat{\beta}(\alpha) \) reaches its lowest value: \( \hat{\beta}(\alpha) = 1/\sigma^2 \). The dependency of the non-centrality parameter \( \lambda \) on the shape parameter \( \beta \) can be seen in Figure 2, for \( \sigma = 0.007 \), \( \alpha = 0.255 \) and \( N = 90112 \). \( \lambda \) becomes minimal (and \( P_{FR} \) maximal) for \( \beta = 0 \), i.e. when the generalized Gaussian pdf reduces to the normal pdf. In this case, the correlator (2) would perform optimally. However, this is almost never the case in the subband domain watermark detection discussed in this paper.

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{dependence.png}
\caption{Dependence of the non-centrality parameter \( \lambda \) on the shape parameter \( \beta \) for \( \sigma = 0.007 \), \( \alpha = 0.255 \) and \( N = 90112 \).
}
\end{figure}

Figure 3a shows an example of the pdf of the detector output for 1000 watermarks of power
Fig. 3. Simulated distributions for detection of a binary watermark in generalized Gaussian noise. (a) Rao detector for $\beta = 2.7037$. (b) Correlator. (c) Rao detector for $\beta = 0$.

$\alpha = 0.001$ inserted in a set of data samples of length $N = 20000$ that follow a generalized Gaussian distribution with $\beta = 2.7037$ and $\sigma = 0.05$. The left hand distribution (almost a delta function at 0) refers to detection under hypothesis $H_0$ and the right hand one to detection under hypothesis $H_1$. Figure 3b shows the corresponding distributions for the classic correlator of Equation (2). We can notice that, in the case of the new detector, the $P_{FA}$ and $P_{FR}$ are practically zero for a wide range of threshold values, while, in the case of the correlator, the distributions are hardly distinguishable and produce large $P_{FA}$ and $P_{FR}$. The corresponding ROC curves are displayed in Figure 4a. The dotted line corresponds to the proposed detector and the solid line to the correlator (hardly visible close to the right hand border of Figure 4a). Clearly, Rao detector has much better performance than the correlator.

In the case $\beta = 0$, the performance of the proposed detector degrades heavily, as can be seen in Figure 3c, where the right hand distribution can no longer be distinguished from the left hand one. As the ROC curves in Figure 4b show, the correlator now performs slightly better than the proposed detector.

III. Generalized Gaussian parameter estimation

As derived in Equation (12), the detector structure depends, apart from the watermarked image and the watermark under test, on the value of the $\beta$ parameter of the generalized Gaussian
Fig. 4. ROC curves for correlator and Rao detector: (a) Generalized Gaussian distribution ($\beta = 2.7037$).

(b) Gaussian distribution ($\beta = 0$).

approximation of the transform coefficients. In order to assess the experimental performance of the new detector, we should also have an estimate of $\sigma_Y$, so that the Fisher information in Equation (16) can be calculated.

A. Estimation of variance $\sigma_Y^2$

The variance of the generalized Gaussian distribution that fits best to the data can of course be estimated using the sample variance:

$$\hat{\sigma}_Y^2 = \frac{1}{N} \sum_{i=1}^{N} (Y_i - \hat{\mu}_Y)^2 \tag{17}$$

where:

$$\hat{\mu}_Y = \frac{1}{N} \sum_{i=1}^{N} Y_i \tag{18}$$

is the sample mean, $N$ is the number of image data samples and $Y_i$ are the data samples of the watermarked image under test. The estimates are quite satisfactory, since the number of data samples are of the order of $10^4$ for image domains larger than $100 \times 100$. This holds even for second-level DWT subbands of images of size $512 \times 512$ (that is, for images of size $128 \times 128$).
B. Estimation of shape parameter $\beta_Y$

As far as the shape parameter $\beta_Y$ is concerned, we wish to estimate it in a computationally efficient way. For this reason, the technique proposed in [20] is adopted. It is based on exploiting the relation between the variance $\sigma_Y^2$, the mean of the absolute values $E[|Y|]$ and the shape parameter $\beta_Y$:

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

(19)

also called the generalized Gaussian ratio function. Based on this relation, the following steps can be taken to estimate the shape parameter $\beta_Y$:

1. We compute the estimate of the variance $\hat{\sigma}_Y^2$ by using Equations (17) and (18).
2. In addition, we compute the estimate of the mean of the absolute values using $\hat{E}[|Y|] = (1/N) \sum_{i=1}^{N} |Y_i - \hat{\mu}_Y|$.
3. We calculate the ratio $\rho = \hat{\sigma}_Y^2 / \hat{E}^2[|Y|]$.
4. We solve the equation $\hat{\beta}_Y = r^{-1}(\rho)$, e.g. by using a lookup table.

Therefore, the distribution of the watermarked data samples can be fully approximated by a generalized Gaussian distribution with parameters $\hat{\sigma}_Y^2$ and $\hat{\beta}_Y$.

IV. WATERMARK EMBEDDING IN THE DCT DOMAIN

An investigation of the possible exploitation of the (non-DC) DCT coefficient distribution has been previously done in [13]. However, the ML (maximum likelihood) decision rule that was employed produced a different detector structure from the one proposed in this paper, which was optimal under the assumption that the watermark power $\alpha$ is known. Our approach, which assumes that the watermark power is unknown to the detector, is much more realistic. The resulting detector structure is optimal for the asymptotic case, which is satisfied in our experiments, since the number of samples is adequately large.

The dataset on which we embed the watermark is the set of the non-DC coefficients resulting from a $8 \times 8$-block DCT transformation ([21], p.156):
\[
C(u,v) = 4 \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} X(i,j) \left[ \frac{\pi}{2N} u \cos(2i + 1) \right] \left[ \frac{\pi}{2N} v \cos(2j + 1) \right] \quad u,v = 0, 1, 2, \ldots, N - 1
\]

where \(X(i,j)\) is the intensity image, \(C(u,v)\) is the DCT coefficients matrix, and \(N = 8\). It is proven in [22] that a generalized Gaussian model is more suitable for all DCT coefficients (but the DC), because it provides the lowest MSE (mean-squared error) during quantization in coding and compression applications. Although the univariate distribution of each coefficient is proven to be in accordance with the GWGN model, we can consider that the same distribution holds for all the DCT coefficients involved in watermarking, although this is not quite correct in practice. However, this is an assumption that is made in many applications for coding in subband domains. Notably, constant generalized Gaussian distribution parameters are usually considered for all images ([22]). The assumption that all coefficients follow the same pdf is crucial, in order to obtain a reasonable estimate of this distribution, based on a sufficient sample number. Thus, when estimating the parameters of the generalized Gaussian distribution, we can take all coefficient samples into account. An example is given in Figure 5, where the watermarked DCT coefficients of the Baboon image are approximated by a generalized Gaussian pdf, with shape parameter \(\beta = 1.76\) and variance \(\sigma^2 = 0.0171\).

In coding applications, the DCT coefficient data set is usually traversed from top left to bottom right in a zig-zag order (the DC term is located in the upper left corner of the block). We can choose to use a diagonal band of low-to-middle frequency coefficients for watermark embedding in order to achieve robustness to lowpass filtering and compression. For example, we can choose those coefficients \(C(u,v)\) that satisfy \(3 \leq u + v \leq 6\) for a \(8 \times 8\) DCT block. This coefficients block is displayed in Figure 6.

Another aspect we could consider is that of visual masking. For this purpose, base luminance thresholds \(T(i,j)\) should be defined for each DCT coefficient, because each coefficient corresponds to a frequency to which the human visual system has a different level of sensitivity
Fig. 5. Watermarked data samples histogram of DCT domain of Baboon and its generalized Gaussian pdf model fitting.

\[
\log T(i, j) = \log \left( \frac{T_{\text{min}}}{r + (1 - r) \cos^2 \left( \arcsin \left( \frac{2F_{i,0}F_{0,j}}{F_{i,0} + F_{0,j}} \right) \right)} \right) + K \left\{ \log \left( \frac{\sqrt{F_{i,0}^2 + F_{0,j}^2} }{F_{\text{min}}^2} \right) \right\}^2 \tag{21}
\]

where \( F_{i,0} \) and \( F_{0,j} \) are the vertical and horizontal spatial frequencies (cycles/degree) of the DCT-basis functions, \( T_{\text{min}} \) is the minimum value of \( T(i, j) \), \( F_{\text{min}} \) is its corresponding spatial
frequency, $r$ is the obliqueness effect parameter and $K$ is the parabola steepness parameter.

The introduction of a visual masking step in our scheme does not affect our model, since hypothesis $\mathcal{H}_1$ would be:

$$\mathcal{H}_1 : Y[k] = X[k] + \alpha T[k] W[k]$$  \hspace{1cm} (22)

However, $T[k]$ is chosen to be independent of the image data and fixed for a certain DCT block size. Thus, it can be considered as an integral part of the watermark generation, which is also known during detection.

V. WATERMARK EMBEDDING IN THE DWT DOMAIN

As pointed out in [22] as well as in the classic paper of Mallat on wavelets [24], subband image data can be fairly well represented by generalized Gaussian distributions. This holds for all frequency bands but the lowest one. An example is given in Figure 7, where the watermarked coefficients of the HL1 subband of the Baboon image are approximated by a generalized Gaussian pdf with shape parameter $\beta = 1.79$ and variance $\sigma^2 = 0.0137$.

![Data histogram and Estimated PDF](image)

**Fig. 7.** Watermarked data samples histogram of HL1 subband of Baboon and its generalized Gaussian pdf model fitting.

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, as well as
a smoothed image are produced in each stage. This can be repeatedly performed up to the desired resolution level. This is illustrated in Figure 8 for a 2-level DWT. Therefore, the 2-D DWT can mathematically be expressed by:

\[
\begin{align*}
  f^J_{LL}(u, v) &= \sum_{i_1=0}^{K-1} \sum_{i_2=0}^{K-1} g(i_1) g(i_2) f^{J-1}_{LL}(2u - i_1)(2v - i_2) \\
  f^J_{LH}(u, v) &= \sum_{i_1=0}^{K-1} \sum_{i_2=0}^{K-1} g(i_1) h(i_2) f^{J-1}_{LL}(2u - i_1)(2v - i_2) \\
  f^J_{HL}(u, v) &= \sum_{i_1=0}^{K-1} \sum_{i_2=0}^{K-1} h(i_1) g(i_2) f^{J-1}_{LL}(2u - i_1)(2v - i_2) \\
  f^J_{HH}(u, v) &= \sum_{i_1=0}^{K-1} \sum_{i_2=0}^{K-1} h(i_1) h(i_2) f^{J-1}_{LL}(2u - i_1)(2v - i_2)
\end{align*}
\]  

(23)

where \( J \) is the level of the 2-D DWT, \( K \) is the filter length, \( g(n) \) and \( h(n) \) are the impulse responses of the lowpass and highpass filters, respectively, and \( f^0_{LL}(u, v) = f(u, v) \) is the original image. The inverse procedure is followed in order to reconstruct the image. The filters \( g(n) \) and \( h(n) \) are usually orthonormal, meaning that the filters are the same both for the analysis and the reconstruction stage. That is, if \( G(z) \) and \( H(z) \) are the transfer functions of the analysis
filters and $\tilde{G}(z)$ and $\tilde{H}(z)$ are the reconstruction filters, then $G(z) = \tilde{G}(z)$ and $H(z) = \tilde{H}(z)$. They are also necessarily biorthogonal in the sense that analysis and reconstruction stage are symmetric, that is: $G(z) = \tilde{H}(z)$ and $H(z) = \tilde{G}(z)$. In our experiments we employed the Haar filter for which $g(n) = h(n) = 1/\sqrt{2}$, $n \in \{0, 1\}$.

In the case of DWT, more sophisticated methods of visual masking can be incorporated than in the case of DCT. In [25], several masking tools that are integral part of the JPEG 2000 standard are discussed, that demonstrate a significant improvement compared to what was included in the original DCT-based JPEG standard. However, these tools usually imply a modulation of the wavelet coefficients that may substantially affect the detector performance. Thus, we shall not consider them in our experiments.

VI. Experimental Results

In order to confirm the superiority of the new detector compared to the classic correlator, we conducted a number of experiments on many images. We report the tests on 8 different real images of size $512 \times 512$ having varying image content for several SNR values (equivalently, watermark power values). These images are “Baboon”, “Lena”, “Peppers”, “Boat”, “Aerial”, “Couple”, “Elaine” and “Bridge”. 100 different binary pseudo-random watermarks were embedded in the coefficients of either the DCT or the DWT domain, and were afterwards detected by the proposed detector, the correlator and the detector proposed in [13], having always in mind that the watermark power is not known at the detection stage. We should note here that the detector proposed in [13] is defined as:

$$I(Y) = \sum_{k} \left( |Y[k]|^{2/\beta} - |Y[k] - W[k]|^{2/\beta} \right)$$

(24)

and that the pdf parameters are considered to be the same for all coefficients. Its output is proven to follow a Gaussian pdf under both hypotheses, as in the case of the correlator detector. The performance of all detectors was evaluated under no attack as well as under
JPEG compression (in the case of DCT embedding) and SPIHT compression (in the case of
DWT embedding).

Results for DCT domain embedding are shown in Figures 9 and 10. The original and water-
marked versions of the Bridge image of size 512×512 are shown in Figures 9a and 9b respectively,
for SNR≈43dB. Because of the high SNR value, no visual artifacts can be noticed. Figure 9c
shows the ROCs for all considered detectors, when no attack (i.e. intentional or unintentional
image processing operation) is imposed. The superiority of the proposed detector is obvious,
whereas the other two methods exhibit much inferior performance. A curve that is derived
theoretically based on the estimated values of the pdf parameters, is delineated in the same
figure with x. The respective experimental curve demonstrates lower performance, due to the
imperfect fitting of the estimated pdf to the experimental data. After a moderate JPEG com-
pression attack of quality factor 95% the performance of all detectors degrades but the new
detector still demonstrates a much better performance, as can be seen in Figure 9d.

An interesting counterexample can be observed in Figure 10. In this case, the watermarked
DCT coefficients are better approximated by a Gaussian distribution, since the estimated value
of the parameter β is close to zero. This explains the fact that the correlator performs slightly
better than the proposed detector, as is depicted in Figure 10c. The theoretical performance
is slightly better than that of the correlator and, thus, is not visible in the Figure. From our
experiments on several images it has been established that such cases are not very likely to occur
in practice. Namely, only one image in our study behaved this way. The JPEG compression
attack affects the performance of the detectors unevenly, since the correlator presents the best
performance whereas the performance of the other two degrades much more.

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 smoothed image that cor-
responds to the LL2 subband. If we would like to exploit the smoothed image (corresponding
to the LL2 subband) as well, we could perform an 8 × 8-block DCT to it and watermark the
properly chosen coefficients, as described in the previous section. Figures 11a and 11b show the original and watermarked Boat image of size $512 \times 512$. We cannot perceive any visual difference between the two, since the watermark power is rather low (SNR\approx 37db). Figure 11c shows the ROC curves resulting from the watermark detection in subband HH1 of the Boat image, for 100 different binary watermarks embedded additively in the subbands of the 2-level DWT transformation (except LL2). The new detector displays, again, better performance, especially when
Fig. 10. (a) Original Elaine image. (b) Elaine watermarked in the DCT domain (SNR≥44dB). (c) ROCs without attack. (d) ROCs after JPEG compression (95%).

compared to the log-likelihood-based detector. It is also superior to the correlator detector. We can notice that the experimental performance is, again, worse than the theoretical one but it is substantially superior to that of the other detectors. The same observations also hold after a moderate SPIHT compression (3bpp), as is shown in Figure 11d.

Similar conclusions can be drawn from Figure 12. The approximation of the watermarked dataset by the generalized Gaussian is obviously superior to the Gaussian assumption made by the correlator. The log-likelihood based detector displays a much inferior performance in this
Fig. 11. (a) Original Boat image. (b) Boat watermarked in the DWT domain (SNR≈42dB). (c) ROCs without attack. (d) ROCs after SPIHT compression (3bpp).

case, as its ROC curve is located in the top right corner of the diagram. The performance of the detector derived from the log-likelihood function also degrades a lot, due to lack of an estimate for the watermark power in the detection stage.

VII. CONCLUSIONS

A new watermark detector scheme has been proposed in this paper. It is derived from well grounded theoretical results of statistical detection theory. A detector that is equivalent to a
Fig. 12. (a) Original Bridge image. (b) Bridge watermarked in the DWT domain (SNR~32dB). (c) ROCs without attack. (d) ROCs after SPIHT compression (3bpp).

cleirvoyant GLRT is introduced, that is asymptotically optimal in GWGN pdf models. Asymptotic performance is ensured in the case of watermarking since the datasets are adequately large. A generalized Gaussian distribution is assumed both in the DCT and DWT watermark embedding domains. The pdf parameters can be estimated efficiently from the watermarked data instead of the original ones, without distorting the hypothesis model. The detector is proven to perform optimally in practice when no attacks are imposed, but also under moderate compression levels, considering the watermark power employed in our experiments. For wa-
termarked images of lower SNR level, the detector would still be superior to the others under harsher compression. Results are compared to those of the correlator-detector and the detector derived based on the log-likelihood ratio function, when no knowledge of the watermark power is available. The proposed detector has poor performance only in the rarely encountered case when the dataset is better described by a Gaussian distribution.

REFERENCES


List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Generalized Gaussian pdf for $\beta = 1.5$ and $\sigma = 1$</td>
<td>7</td>
</tr>
<tr>
<td>2</td>
<td>Dependence of the non-centrality parameter $\lambda$ on the shape parameter $\beta$ for $\sigma = 0.007$, $\alpha = 0.255$ and $N = 90112$.</td>
<td>10</td>
</tr>
<tr>
<td>3</td>
<td>Simulated distributions for detection of a binary watermark in generalized Gaussian noise. (a) Rao detector for $\beta = 2.7037$. (b) Correlator. (c) Rao detector for $\beta = 0$.</td>
<td>11</td>
</tr>
<tr>
<td>4</td>
<td>ROC curves for correlator and Rao detector: (a) Generalized Gaussian distribution ($\beta = 2.7037$). (b) Gaussian distribution ($\beta = 0$).</td>
<td>12</td>
</tr>
<tr>
<td>5</td>
<td>Watermarked data samples histogram of DCT domain of Baboon and its generalized Gaussian pdf model fitting.</td>
<td>15</td>
</tr>
<tr>
<td>6</td>
<td>$8 \times 8$ DCT block and selected coefficients for watermark embedding.</td>
<td>15</td>
</tr>
<tr>
<td>7</td>
<td>Watermarked data samples histogram of HL1 subband of Baboon and its generalized Gaussian pdf model fitting.</td>
<td>16</td>
</tr>
<tr>
<td>8</td>
<td>2-level two-dimensional DWT with all subbands displayed.</td>
<td>17</td>
</tr>
<tr>
<td>9</td>
<td>(a) Original Bridge image. (b) Bridge watermarked in the DCT domain (SNR$\approx$43dB). (c) ROCs without attack. (d) ROCs after JPEG compression (95%).</td>
<td>20</td>
</tr>
<tr>
<td>10</td>
<td>(a) Original Elaine image. (b) Elaine watermarked in the DCT domain (SNR$\approx$44dB). (c) ROCs without attack. (d) ROCs after JPEG compression (95%).</td>
<td>21</td>
</tr>
<tr>
<td>11</td>
<td>(a) Original Boat image. (b) Boat watermarked in the DWT domain (SNR$\approx$42dB). (c) ROCs without attack. (d) ROCs after SPIHT compression (3bpp).</td>
<td>22</td>
</tr>
<tr>
<td>12</td>
<td>(a) Original Bridge image. (b) Bridge watermarked in the DWT domain (SNR$\approx$32dB). (c) ROCs without attack. (d) ROCs after SPIHT compression (3bpp).</td>
<td>23</td>
</tr>
</tbody>
</table>