Transform based additive data hiding based on a hierarchical prior

ANTONIS MAIRGIOTIS[1,2], GEORGE STYLIOS[2,3]
1. School of Medicine, Department of Hygiene and Epidemiology
University of Ioannina, School of Medicine,
University Campus, Ioannina 45110, Greece
mairgiot@cs.uoi.gr, mairgiotis@gmail.com
GREECE
2. Department of Applications of Information Technology in Management and Economics
Multimedia Lab
Technological Educational Institute of Ionian Islands
3100, Lefkada
GREECE
3. Department of Computer Engineering and Informatics,
26500, Patras
GREECE
[1,2]mairgiot@cs.uoi.gr, mairgiotis@gmail, [2,3]gstylios@yahoo.gr, gstylios@teiion.gr

Abstract: - This work aims at additive watermark detection using a hierarchical prior with two levels in the DWT (Discrete Wavelet Transform) domain. Using this prior we construct a new test statistic which exhibits comparable performance with regard to other state of the art methods. The experimental results on known watermark images demonstrates the high detection sensitivity of the proposed prior in the transform domain along with its improved robust properties.

Key-Words: - watermark detection, watermarking, wavelet domain, hierarchical prior

1 Introduction

Data hiding methods are drawing attention last decade, as one of the useful methods of protecting copyright and security of digital multimedia contents. Digital image watermarking is a data hiding technique where by embedding a piece of information in a host image we ensure the copyright protection, integrity checking, protection of intellectual property etc [1], [5]. The embedding procedure take place either to spatial domain [3] or in transform domain [5]-[9]. The most applied transforms are DCT (Discrete Cosine Transform) and DWT transforms whereas the exploitation of their properties provide us with watermark detectors that have high detection performance along with good robust properties [6]-[10].

Usually, the secret information is embedded in perceptually significant spectral components of the image in the transform domain. More specific this signal is a spread spectrum watermark signal where its insertion in specific image coefficients helps to avoid the degradation of image quality as this perceived by HVS (Human Visual System). The simpler detector that we could apply is Linear Correlator (LC) [4], which is an optimal detector if the host image's coefficients follow Gaussian statistics. However, is common knowledge that wavelet transform's coefficients obey non-Gaussian statistics, following distributions with more heavy tails [7], [8]. In watermarking literature, many researchers have shown that, using subband representation of natural images, the histograms of the corresponding coefficients have heavier tails and are more sharply peaked at zero.

As a consequence the LC detector has a suboptimal behavior, a fact that has led us to search for alternative distributions for the problem at hand.

In watermarking literature Generalized Gaussian Density (GGD) and Cauchy distribution as a member of SaS (Symmetric alpha Stable) family have been the most applied distributions in image watermarking problem [6]-[8].

Inspired from image recovery problems [3] a spatially adaptive prior applied in image watermarking problems. The present paper, based on the same prior, proposes a detector structure that has been proved successful in additive watermarking problem [3] in spatial domain. More specific, we investigate the suitability of the hierarchical prior in the problem of additive watermarking problem in transform (DWT) domain.

The rest of the paper is organized as follows. In Section 2, we set our problem as a binary hypothesis problem. Then in subsections 2.1 and 2.2 the

watermark embedding and detection procedures are explained. In Section 3, we define the hierarchical prior of this work and in Section 4 we present the experimental results.

2 Problem Formulation

2.1 Additive Watermarking Problem as a binary hypothesis problem

In order to derive a binary hypothesis test we assume that DWT coefficients are i.i.d (independent and identically distributed) random variables drawn from some underlying probability density function (pdf).

Suppose we have N coefficients of an image in wavelet domain. Then, defining the host signal's coefficients in a vector form as $\mathbf{x} = [x[1], x[2], ..., x[N]]$ and the watermark signal as $\mathbf{w} = [w[1], w[2], ..., w[N]]$ the additive watermark embedding rule can be denoted as:

$$\mathbf{y} = \mathbf{x} + \gamma \mathbf{w} \tag{1}$$

where γ reflects the trade-off between image fidelity and robustness of the watermark.

Following the embedding scheme as in the previous rule, we can define the additive watermarking problem, as a binary hypothesis problem:

$$\begin{aligned} H_0 &: \mathbf{y} = \mathbf{x} \\ H_1 &: \mathbf{y} = \mathbf{x} + \gamma \mathbf{w} \end{aligned}, \ \gamma > 0 . \tag{2}$$

where H_0 denotes the null hypothesis (we don't have any watermark or we have different watermark) and H_1 denotes the alternative hypothesis (we have the watermarked signal). The actual test is performed without knowledge of the original, unwatermarked image, thus we have a blind watermark detection.

Definining likelihood ratio as $\Lambda(\mathbf{y})$, and after the application of logarithm the log-likelihood ratio becomes:

$$\Lambda(\tilde{\mathbf{y}}';\tilde{\mathbf{a}}) = \log\left\{\frac{p(\tilde{\mathbf{y}}';\tilde{\mathbf{a}},H_1)}{p(\tilde{\mathbf{y}}';\tilde{\mathbf{a}},H_0)}\right\} \underset{H_0}{\overset{H_1}{\gtrless}} 0 \tag{3}$$

2.1 Watermark Embedding

In the embedding process, we apply the DWT transform to our image and in consequence we make use of the second level sub-bands of wavelet transform. Using the spread-spectrum watermarking system the secret information that we embed is a spread spectrum watermark [1], [5]. More specific, is a pseudorandom sequence (PRS), where its security is based on a key existence. This key initializes the pseudorandom generator's seed.

It is noticeable, that the anti-jamming properties of these systems increase the robustness of the watermark, while the pseudorandom modulation of the hidden information signal increases the achieved security [7].

After watermark's embedding process, we apply the inverse DWT transform in order to take the marked image. In all of the experiments, we quantify these watermarked images using 8-bpp accuracy in spatial domain before watermark detection.

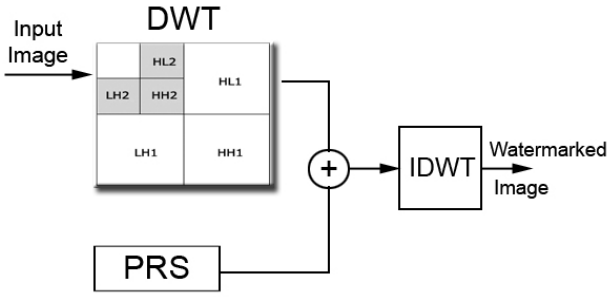

Fig.1 Block diagram of the embedding process in wavelet domain

2.2 Watermark Detection

In our case, we are interested for the verification of the existence of the watermark signal i.e the watermark detection. The watermark detection procedure is taking place directly in wavelet domain on the second level detail band's coefficients of wavelet transform. In Fig. 2 we see a block diagram of the detection process in wavelet domain.

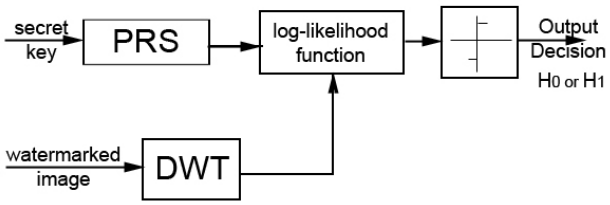

Fig.2 Block diagram of the detection process in wavelet domain

3 Problem Solution

3.1 Proposed prior

We assume that wavelet coefficients obey a Gaussian pdf, given by

$$x[i] \sim N\left(0, a^{-1}[i]\right) \tag{4}$$

where a^{-1} is the variance parameter at coefficient location i. Assuming independence between wavelet coefficients, we can write the joint pdf between the sub-bands as:

$$p(\mathbf{x};\tilde{\mathbf{a}}) \propto \prod_{k=1}^{K}\prod_{i=1}^{N}\left[a_k^{1/2}[i]\exp\left(-\frac{1}{2}a_k[i]\mathbf{x}_k[i]^2 \right)\right] \quad (5)$$

where

$\tilde{\mathbf{a}} = \left[\mathbf{a}_1^{\,T}, \mathbf{a}_2^{\,T}, \mathbf{a}_3^{\,T} \right]^T$, $\mathbf{a}_k = \left[a_k[1], a_k[2], ..., a_k[N] \right]^T$, denotes the corresponding variance parameters, K is the number of sub-bands we use and indexes $k = 1, 2, 3$ are the horizontal, vertical and diagonal detail subbands of wavelet transform. The total number of coefficients in every band is equal to N.

The pdf in (5) allows the flexibility that the variance parameters can vary between every coefficient. This is desirable for modeling the non-stationary properties of the image (e.g., edges) as these described from image's coefficients in wavelet domain. Unfortunately, we have as many variance parameters $a_k(i)$ as the number of transform's coefficients. Thus, to avoid the problem of over-fitting, we model these parameters as random variables, and define a hyper-prior on them. Our choice of hyper-prior is Gamma pdf, which is of the form:

$$p(a_k[i];m,l) \propto a_k^{\frac{l-2}{2}}[i]\exp\left\{ -m(l-2)a_k[i] \right\}, \quad k = 1,2 \quad (6)$$

where m, l are the parameters of Gamma distribution. The choice of such a distribution is justified by the fact that Gaussian and Gamma families are conjugated.

For this definition of Gamma pdf we have that the expected value is $E\left[a_k(i) \right] = l\left(2m(l-2) \right)^{-1}$ and the variance is: $Var\left[a_k(i) \right] = l\left(2m^2(l-2)^2 \right)^{-1}$.

In order to accelerate the speed of computations, we propose to "marginalize" the unknown variance parameters, leading to a Bayesian detector, given by [3], [4]:

$$\mathrm{B}(\tilde{\mathbf{y}}';m,l) = \log\left\{ \frac{\int p(\tilde{\mathbf{y}}' \mid \tilde{\mathbf{a}}, H_1)\, p(\tilde{\mathbf{a}};m,l)\, d\tilde{\mathbf{a}}}{\int p(\tilde{\mathbf{y}}' \mid \tilde{\mathbf{a}}, H_0)\, p(\tilde{\mathbf{a}};m,l)\, d\tilde{\mathbf{a}}} \right\} \underset{H_0}{\overset{H_1}{\gtrless}} 0 \quad (7)$$

After some algebra and resorting to Gamma integrals, we can show that test statistic is given by:

$$T_B(\tilde{\mathbf{y}}';m,l) = \sum_{k=1}^{3}\sum_{i=1}^{N}\log\left(\frac{(\mathbf{y}_k'(i) - \mathbf{w}_k''(i))^2 + 2m(l-2)}{(\mathbf{y}_k'(i))^2 + 2m(l-2)} \right) \underset{H_0}{\overset{H_1}{\gtrless}} T \quad (8)$$

where $\mathbf{w}_k''(i) = \gamma \mathbf{w}_k$.

4 Experimental Results

In this section we present experimental results that demonstrate the performance of the derived class of detectors. We conducted two kinds of experiments, where in the first one we compare the detection performance of our detector with the known GGD based watermark detector and in the second one we see the detector's performance under intentional or unintentional attacks, e.g JPEG compression. The quantification of the detection performance is based on Receiver Operating Characteristic (ROC) curves. Two known images in image processing community, Lena and Bridge, are used with size of 512x512.

Fig.3 Test images of "Lena" and "Bridge".

A set of 100 different randomly generated 1-bit spread-spectrum watermarks were used for each of the two images at a specified WDR. For each watermark, we evaluated the test statistic twice, once with the watermark and once with the unwatermarked data. Then, the histograms of test statistic for the two cases are then computed based on which the ROC curve is generated using a moving threshold. We usually call this procedure "random watermarks", since for fixed images we used random watermark images to obtain ROCs for fixed images.

In order to quantify the strength of the watermark relative to its host signal, we use the so called watermark to document ratio (WDR) [10], which is defined as

$$WDR = 10\log_{10}\left(\frac{\sigma_w^2}{\sigma_x^2} \right), \quad (9)$$

where $\sigma_w^2 = \frac{1}{N}\sum_k w^2[k]$ and $\sigma_x^2 = \frac{1}{N}\sum_k x^2[k]$ which are the powers of the watermark and host signal, respectively.

The most known detectors that are based on wavelet transform for the additive watermarking problem, are GGD based detectors. In this work as in [3] we apply an "adaptive" wavelet GGD model, meaning that we have a different GGD model for each wavelet band. Details of these wavelet based detectors are provided in the Appendix of [3]. The wavelet filters that have been used in this work are the Daubechies-8 2D separable filters.

4.1 Performance of detectors without attacks

In Fig. 4 and Fig.5 we show the performance results achieved by two DWT domain detectors for image Lena and Bridge. The first one, is based on a GGD model in the DWT domain (GGD-based detector) and the second one is the detector proposed, which is based on the hierarchical model and the Bayesian methodology of parameter treatment (Bayesian-based detector).

In Fig. 4 for very weak watermarks we can see that the proposed detector is superior compared with the state of the art GGD based detector. In Fig. 5 we can see that the two detectors have almost the same detection performance.

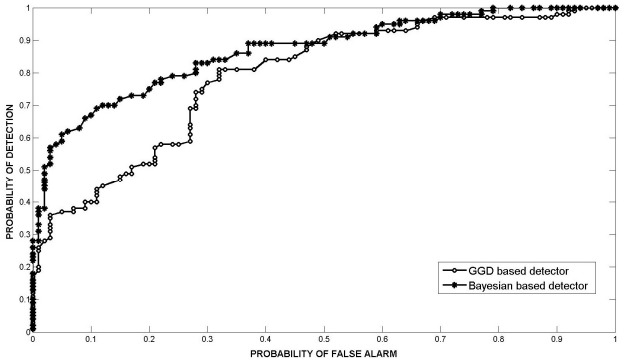

Fig. 4 ROC curves for detection performance comparison between GGD based and Bayesian based (proposed prior) wavelet detectors – Image Lena (WDR=-49dB)

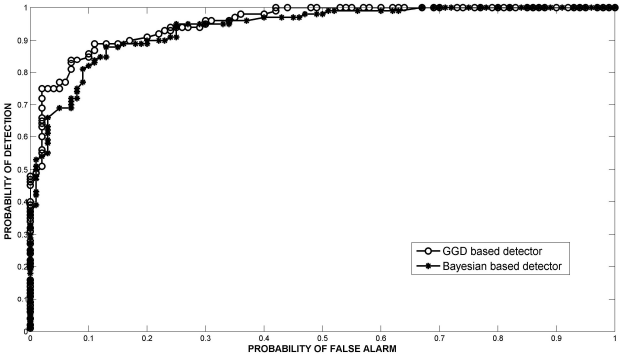

Fig. 5 ROC curves for detection performance comparison between GGD based and Bayesian based (proposed prior) wavelet detectors – Image Bridge (WDR=-47dB)

4.2 Performance of detectors under intentional/unintentional attacks

It is common practice when we examine the performance of a new watermark detector, to test its behavior under intentional or unintentional attacks (e.g JPEG compression) or some kind of filtering (e.g wiener filtering). In this work, we present the performance of our proposal compared with the known GGD based detector in wavelet domain. In order to do that we apply two kind of experiments. In the first one we have the known JPEG compression format and in the second one we apply a Wiener filtering followed by an additive white Gaussian noise addition. In Figures 6-9, we can observe the performance of the aforementioned detectors under attacks applied on the test images of this work.

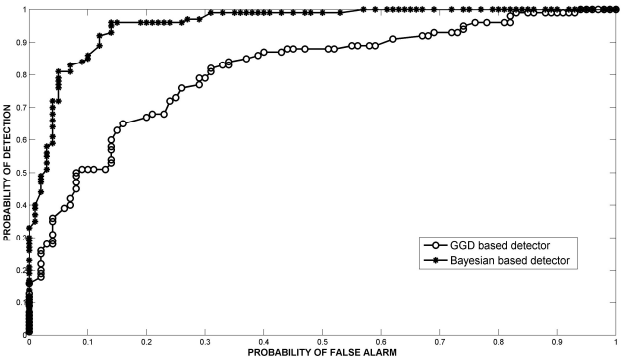

Fig. 6 ROC curves for detection performance comparison between GGD based and Bayesian based (proposed prior) wavelet detectors under JPEG attack – Image Lena (WDR=-47dB)

ISBN: 978-960-474-297-4

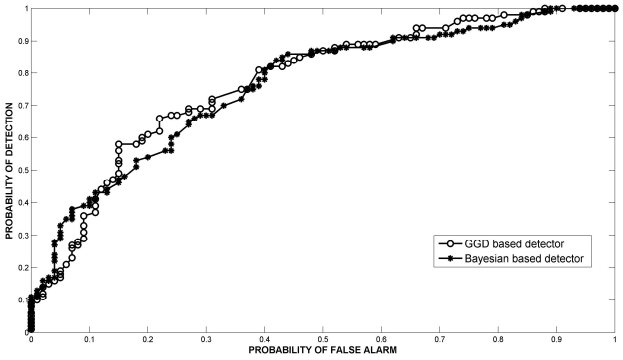

Fig. 7 ROC curves for detection performance comparison between GGD based and Bayesian based (proposed prior) wavelet detectors under JPEG attack – Image Bridge (WDR=-47.1dB)

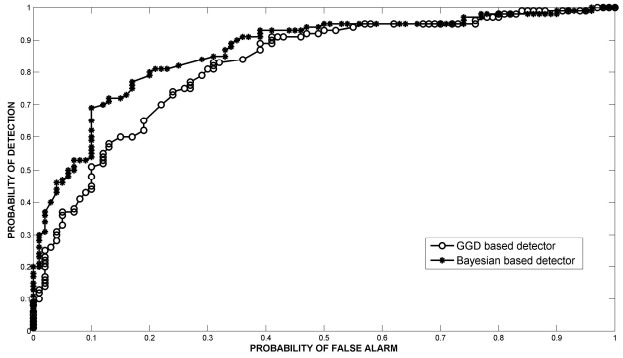

Fig. 8 ROC curves for detection performance comparison between GGD based and Bayesian based (proposed prior) wavelet detectors – Image Lena (WDR=-42dB) after Wiener filtering plus awgn attack.

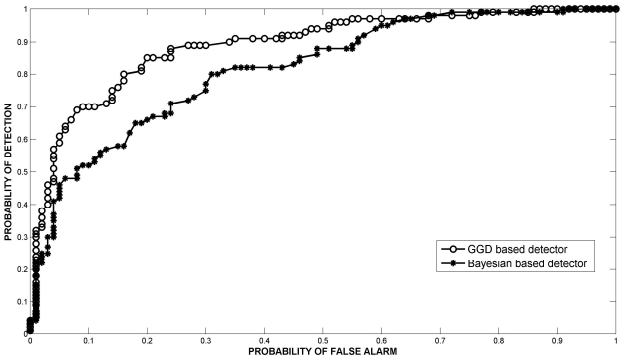

Fig. 9 ROC curves for detection performance comparison between GGD based and Bayesian based (proposed prior) wavelet detectors – Image Bridge (WDR=-45dB) after Wiener filtering plus awgn attack.

5 Conclusion

In this work we developed a transform based data hiding approach using a hierarchical, two level model for the problem of image additive watermarking. Based on the proposed prior we derived a new class of watermark detectors validating the fact that this model is also applicable to transform domain and it is a suitable solution for the problem of additive watermarking.

References:

[1] I. Cox, M. Miller and J. Bloom, J. Fridrich, T. Kalker, *Digital Watermarking and Steganography*, 2nd edition, Morgan Kaufman, 2008

[2] G. Chantas, N. P. Galatsanos and A. Likas, "Bayesian restoration using a new nonstationary edge preserving image prior", IEEE Trans. On Image Processing, Vol. 15, No. 10m pp. 2987-2997, October 2006

[3] A. K. Mairgiotis, N. P. Galatsanos and Y. Yang, "New additive watermark detectors based on a hierarchical spatially adaptive image model", IEEE Trans. Information and Security, vol. 3, no. 1, pp. 29-37, 2008

[4] S. M. Kay, *Fundamentals of Statistical Signal Processing: detection Theory*, vol. 2, Prentice Hall, 1998

[5] M. Barni, F. Bartolini, *Watermarking Systems Engineering, Enabling Digital Assets Security and Other*, Marcel Dekker, 2004

[6] Q. Cheng and T. S. Huang, "An additive approach to transform domain information hiding and optimum detection structure", IEEE Trans. On Multimedia, vol. 3, no.3, Sept. 2001

[7] A. Briassouli, P. Tsakalides, A. Stouraitis, "Hidden Messages in Heavy-Tails: DCT-Domain Watermark Detection Using Alpha-Stable Models", IEEE Trans. On Multimedia, 7(4):700-712, Aug. 2005

[8] R. Kwitt, P. Meerwald, A. Uhl, "Lighweight detection of Additive Watermarking in the DWT-Domain", vol. 20, no.2, p. 474-484, Feb. 2011

[9] J. R. Hernandez, M. Amado, F. perez-Gonzalez, "DCT-Domain Watermarking Techniques for Still Images: Detector Performance Analysis and a New Structure", IEEE Trans. on Image Processing, vol. 9, no. 1, Jan. 2000

[10] J. J. Eggers, B. Girod, "Quantization effects on digital watermarks", Signal Processing, vol. 81, no. 3, 2001

ISBN: 978-960-474-297-4