Rao based watermark detector through two-level hierarchical prior in transform domain

Antonis Mairgiotis
Dept. of Computer Science and Engineering, Tech. Educ. Inst. of Thessaly
Larissa, Greece
mairgiotis@gmail.com

Christos Koliopanos
Technological Educational Institute of Epirus
Arta, Greece
koliopanos@yahoo.com

Lisimachos P. Kondi
Dept. of Computer Science and Engineering, University of Ioannina,
Ioannina, Greece
lkon@cse.uoi.gr

ABSTRACT

In this work, we propose a transform based blind zero-bit watermark detector, designed based on a hierarchical, two-level image prior. This model is applied through Rao hypothesis test, where we detect the hidden information with unknown amplitude. The proposed method is suitable for wavelet domain watermark detection without the need to estimate the appropriate parameters under the alternative hypothesis. The proposed system shows efficient robustness against a variety of known attacks. The experimental results show that the proposed detector has a good performance with or without attacks in terms of ROC (receiver operating characteristics) curves compared with other known state-of-the-art statistical detectors.

KEYWORDS

Watermark detection, Rao detector, two-level hierarchical prior, wavelet domain.

1 INTRODUCTION

In recent years, the evolution of multimedia and internet technologies allows people to create, edit, store and deliver the digital content of media like images, sound and video. Watermarking comes as an alternative technology, where the protection of intellectual property is embodied in the actual content of the digital medium and is maintained even after its transmission. Watermarking can find applications such as copyright protection, ownership identification or copy protection [1]-[3].

PCI 2017, September 28–30, 2017, Larissa, Greece

ACM ISBN 978-1-4503-5355-7/17/09…$15.00
https://doi.org/10.1145/3139367.3139444

In a watermarking system, we are interested in additional properties such as the robustness against malicious or unintentional attacks and of the amount of information that can be incorporated in our data reliably and in an imperceptible manner. We refer to robust watermarking when the watermark is still detectable after various removal procedures e.g. accidental or malicious attacks. Notice that, even though someone knows the hidden information, this is not sufficient to remove the hidden message without knowledge of various parameters and secret keys. The concept of security means that unauthorized parties aren't capable of recovering the watermark even if they know the embedding or the detection technique that is applied. In our work, we develop a blind watermarking methodology, meaning that we don't have any access to the original host data [1], [2].

Among the most popular techniques based on a transform domain are those using the DWT (Discrete Wavelet Transform) or the DCT (Discrete Cosine Transform) domain. Considering the DWT domain, the good properties of localization, the multi-resolution properties, the HVS (Human Visual System) exploitation and the low computational complexity are powerful features for selecting the wavelet domain. In the case of the DCT domain, the wide usage and applications e.g. in coding formats, have become diachronic choices even in the case of watermarking [5]-[9]. Recently, the Contourlet Transform (CT) has become a powerful alternative in terms of embedding the secret information [12].

Spread Spectrum (SS) is a standard technique where the power spectrum of the hidden data is spread over all frequencies, thus making its detection more difficult. This technique has proven to be efficient, robust and cryptographically secure [1], [4]. Thus, it is difficult for an attacker to identify the information of interest, a fact that increases the security of the system. Typically, watermarks are embedded over an important part of host features using an additive or multiplicative technique [4]-[11].

A watermarking scheme can be seen as a communication task consisting of two basic components: i) the embedding stage and ii) the detection stage. During embedding we can follow an additive or a multiplicative rule of watermark casting depending on whether we want the embedding to exploit image variations according to the amplitude of the selected transform coefficients.

Usually, the optimal watermark detector is based on the statistical modeling of the distribution of transform's coefficients and the appropriate test statistic.

In what follows, in Section II we are going to refer to the problem of watermark detection and in Section III, to the Rao hypothesis test and the proposed detector. We will then look at the results related to comparison with other known state of the art Rao based watermark detectors.

2 PROPOSED DETECTOR

2.1 Watermark detection problem

In this work, following the additive rule, the detection problem can be defined as a binary hypothesis problem [15]:

$$\begin{aligned} \mathbf{H_1}&: \mathbf{Y}[i] = \mathbf{X}[i] + \gamma \mathbf{W}[i] \\ \mathbf{H_0}&: \mathbf{Y}[i] = \mathbf{X}[i] \end{aligned} \tag{1}$$

where $\mathbf{X}[i]$ is the known host image sequence and $\mathbf{W}[i]$ is the watermark signal sequence of interest. Notice that $i = 1, \dots, N$ for N coefficients of our interest. The parameter γ is the known strength parameter which defines the level of watermark's embedding power. Watermark detection, in the view of binary decision, is the procedure of trying to answer if a host image contains or not the watermark information based on some kind of test statistic [1]-[3], [15].

Usually, the optimal watermark detector is based on the statistical modeling of the distribution of transform's coefficients and the appropriate test statistic. Based on the observations of Y[n] and by using a specific statistical detector, we can derive a test statistic e.g Likelihood Ratio Test (LRT) [15]. The LRT is defined as:

$$\mathrm{LRT}(\mathbf{Y}) \underset{\mathrm{H_0}}{\overset{\mathrm{H_1}}{\gtrless}} \mathrm{T} \tag{2}$$

where the threshold T can be calculated based on Neyman-Pearson criteria defining a level of the probability of false alarm, or based on Bayesian criteria resorting to a cost function [15].

2.1 Rao hypothesis test and proposed detector

In many cases, due to blind watermarking we can't have any access to the watermark's strength knowledge from the receiver's side. As a consequence, in this work, we will propose the application in the transform domain, of the Rao hypothesis test based on a hierarchical image model that was applied in the spatial domain [13]. The proposed detector is based on previous work [13], but with the distinct difference that in this work we investigate its application in the transform domain. Using the Rao hypothesis we manage to alleviate parameter's estimation where, in addition, we can derive an optimal detector equivalent to that of a GLRT (Generalized Likelihood Ratio Test). That means that the imaging model of our interest concerns the coefficients of the wavelet domain. For efficient blind watermark detection, we propose to use a hierarchical distribution to model the detail subband wavelet coefficients. The proposed pdf assumes that the coefficients X(i, j), in every location (i, j) of our interest follow a Gaussian pdf, given by:

$$\mathbf{X}(i,j) \sim \mathrm{N}\left(0, v_k^{-1}(i,j)\right) \tag{3}$$

where $v_k^{-1}(i,j)$ is the variance parameter. Based on the assumption that the coefficients are independent and that every subband in a specific level of the DWT, has N coefficients in total, we can write the joint pdf as:

$$p(\mathbf{X}_k(i); \boldsymbol{\nu}_k) \propto \prod_{k=1}^{3} \prod_{i=1}^{N} v_k^{-1}(i) \exp\left(-\frac{1}{2} v_k^{-1}(i) \mathbf{X}(i)^2\right) \tag{4}$$

where k denotes the corresponding sub-band. For notational convenience we used a single index –i –. Notice that we used the three detailed subbands of the same decomposition level.

Since the proposed prior has the flexibility to be adapted to the local information, meaning in a coefficient base, we define a hyper-prior on them in order to avoid the problem of over-fitting. Thus, a Gamma pdf is applied where:

$$p(\boldsymbol{\nu}_k(i); m, l) \propto \exp(-m(l-2)\boldsymbol{\nu}_k(i)), \; k=1, 2, 3 \tag{5}$$

where m, l are the parameters of Gamma distribution.

In addition, in the watermarking problem the watermark signal has very low power compared with the power of the host signal. Based on the binary hypothesis problem defined in Eq. (1) we can derive the likelihood ratio test in Eq. (2). Given that the original wavelet coefficients can be considered as a realization of i.i.d. random variables following the proposed distribution, the watermark detection problem can be formulated as the detection of a deterministic signal of unknown amplitude in the proposed distribution noise, meaning the host image's coefficients. This is actually a composite hypothesis testing problem, which can be formulated as a two-sided parameter test. The null hypothesis (H_0) and the alternative hypothesis of this test is given by:

$$\begin{aligned} \mathbf{H_0}&: \gamma = 0 \\ \mathbf{H_1}&: \gamma \neq 0 \end{aligned} \tag{6}$$

For this test there isn't any UMP (Uniform Most Powerful) test. But, since for the problem at hand we have large data records, we can define an asymptotical test equivalent to the GLRT known as Rao test [15]. In this case, we need only to compute the Maximum Likelihood (ML) under the null hypothesis.

Thus, we can define the Rao hypothesis test for the hierarchical model in the transform domain as [13], [15]:

$$\mathrm{T}_{\mathrm{Hier-Rao}}\left(\mathbf{Y}; \mathbf{v}_{/\mathrm{H_o}}\right) = \frac{\left[\sum_{k=1}^{3} \sum_{i=1}^{N} \mathbf{w_k}(i) \mathbf{v_k}(i) \mathbf{Y_k}(i)\right]^2}{\frac{1}{2N}\left[\sum_{k=1}^{3} \sum_{i=1}^{N} \mathbf{w_k}(i)^2\right] \sum_{k=1}^{3} \sum_{i=1}^{N} \left(\mathbf{v_k}(i) \mathbf{Y_k}(i)\right)^2} \tag{7}$$

We chose to use the Rao test for our observations, for two main reasons: except the practical value of the proposed detector, it is only the watermark shape and not the knowledge of the strength parameter γ (invariant with respect to the strength), that is necessary for the Rao detector. In order to estimate the parameters l, m of the Gamma hyper-prior we followed similar approach as in [13], where for the more critical parameter l, the selection was based on an empirical approach in such a way that the histogram of the normalized wavelet coefficients would best fit a standard Gaussian pdf.

3 RESULTS AND DISCUSSION

3.1 *Comparison with other similar state-of-the art detectors in transform domain*

In order to test the performance of the proposed watermark detector based on Rao hypothesis, we have to compare the

$$\mathrm{WDR} = 10\log\left(\frac{\sigma_W^2}{\sigma_x^2}\right)\mathrm{dB} \quad (8)$$

proposed hierarchical based Rao test with the known Rao-GGD [7] and Rao-Cauchy tests [9], [10]. The methods were chosen since they were the most known Rao-test based methods for the transform domain additive watermarking problem. In addition, the usefulness of the proposed detector is validated through a methodology that doesn't take into account the watermark's power or any other additional complexity due to the embedding stage (e.g. multiplicative rule). In Table 1, we can see the proposed test statistics, for the pdfs of comparison:

Table 1: Test statistics of comparison

Generalized Gaussian Distribution definition [7], [8]:

$p(\mathbf{x}|\alpha, c) = \frac{c}{2\alpha\Gamma(1/c)} e^{-\left|\frac{x}{\alpha}\right|^c}$,
c>0: shape parameter, α>0: scale parameter

RAO - GGD test statistic:

$$l(\mathbf{y}) = \left(\sum_{i=1}^{N} \mathrm{sgn}(\mathbf{y}(i))\mathbf{w}(i)|\mathbf{y}(i)|^c\right)^2$$

Cauchy Distribution definition [9], [10]:

$p_x(x) = \frac{1}{\pi}\frac{\gamma}{\gamma^2+(x-\delta)^2}$,
γ: data dispersion, δ: location parameter

RAO- Cauchy test statistic:

$$l(\mathbf{y}) = \left[\sum_{i=1}^{N} \frac{\mathbf{y}(i)w(i)}{\gamma^2 + \mathbf{x}(i)^2}\right]^2 \frac{8\gamma^2}{N}$$

The evaluation of the performance of the proposed detector is based on ROC curves. The implementation of the detectors and all the experiments were conducted in MATLAB environment [16]. Thus, using Monte-Carlo simulation we follow two kinds of experiments. At first, we run the "random watermark" experiment, where we derive 1000 watermarks using SS methodology and we embed them in known images. Subsequently, for statistical significance reasons, we run the "random image" experiment based on a dataset of images [17].
The DWT domain is the domain of our choice, trying to exploit the inherent exploitation of the HVS and other useful features of the transform. More specifically, we make use of the second level of wavelet transform, trying to combine the good detection performance along with the robustness properties of the detector in use. Regarding the choice of DWT wavelet filter, the decomposition level and the embedding subband, we employ the Daubechies-8 2D separable filters [14].
In this study, in order to have various cases, the images of our interest are the four known images as shown in Figure 1 ("Lena", "Boat", "Bridge" and "Barbara") of size 512x512. Thus, based on the mean PSNR (Peak-Signal-to-Noise-Ratio) as it is defined in Eq. 10 we can have a visual appearance of watermarked images for the WDR (Watermark- to-Document-Ratio) quantification method of watermark energy embedding defined in Eqs. 9.
Thus, we define the "Watermark to Document Ratio" as in [9], [10], [13]:

Figure 1: Four known original images (a) "Lena", (b) "Boat", (c) "Bridge", (d) "Barbara"

with

$$\sigma_w^2 = \frac{1}{N}\sum_{i=1}^{N} W[i]^2,\ \ \sigma_x^2 = \frac{1}{N}\sum_{i=1}^{N} X[i]^2 \quad (9)$$

the corresponding watermark and host signal powers respectively. With N is denoted the number of data samples and the term "document" refers to the original host data. The known PSNR metric is defined as:

$$\mathrm{PSNR} = 10log_{10}\left(\frac{\max(\mathbf{X})^2}{\|\mathbf{Y}-\mathbf{X}\|^2}\right)\mathrm{dB} \quad (10)$$

3.2 *Detection performance without attacks– random watermark experiment*

The performance comparison of the proposed detection schemes is based on the four known images. Our experiments in Fig. 2 show that the proposed detector exhibits the same or better performance when it works without some kind of attack. Only in the case of image "Boat", performance seems slightly reduced. For practical purposes, this is not a problem due to the low level of watermark energy that we study. With regard to the other

images the performance of the proposed detector sensitivity is the same or even much better.

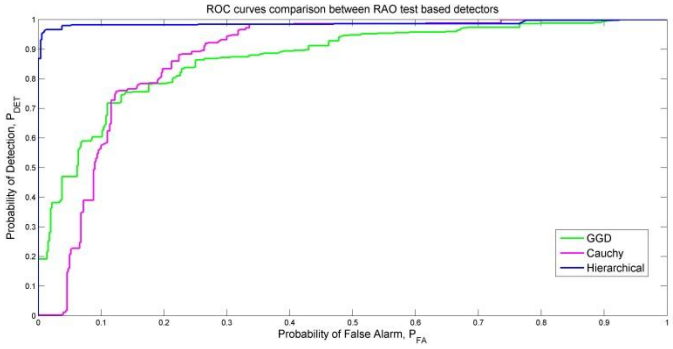
3.3 *Detection performance under attacks*

To extract reliable results, our detector needs to be tested for its performance under various attacks, so judging for its robustness property. Here, we test the performance of the proposed technique against common watermarking attacks such as: JPEG compression, AWGN (Additive White Gaussian Noise) and median filtering.

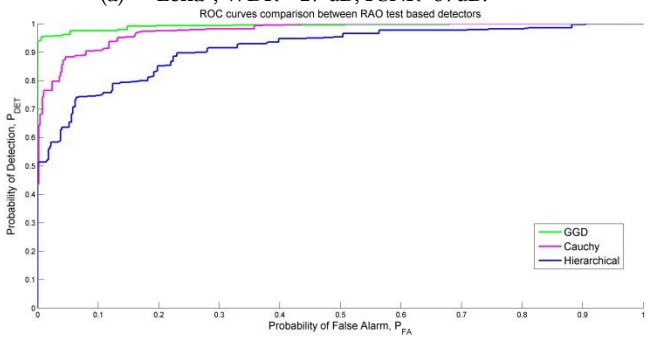
JPEG attack: in the first attack, we investigate the effect of JPEG compression using a known quality factor. As demonstrated in Fig. 3, the proposed method is highly robust against JPEG with different quality factors up to 50%. In all cases, the resistance of the proposed watermarking scheme is superior compared with the other cases.

Median filtering attack: using median filtering with window size of 5x5, we test the robustness of the proposed detector under this particular attack. As we can observe in Fig. 4, only in the first case of image Lena the proposed detector is slightly inferior compared with the corresponding Cauchy based attack. In all the other images, the robustness of the proposed detector is clearly superior compared with the other two detectors.

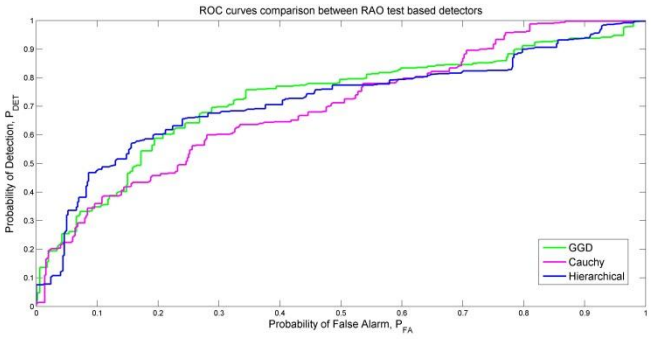
AWGN attack: in the second attack, the proposed technique is tested against AWGN attack. As we can observe in Fig. 5, the method has high resistance against this specific attack appearing at same levels with other known statistical methods.



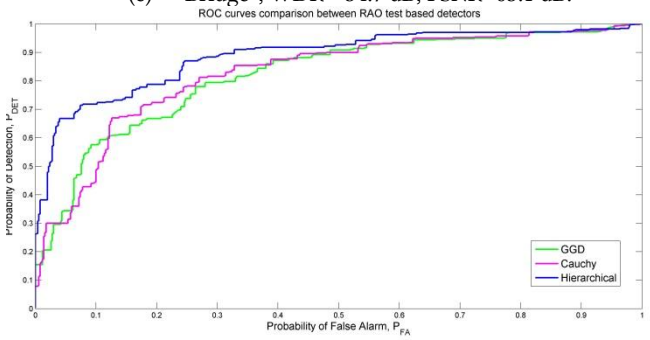
(a) "Lena", WDR=-27 dB, PSNR=57dB.



(b) "Boat", WDR=-31.9dB, PSNR=60 dB.

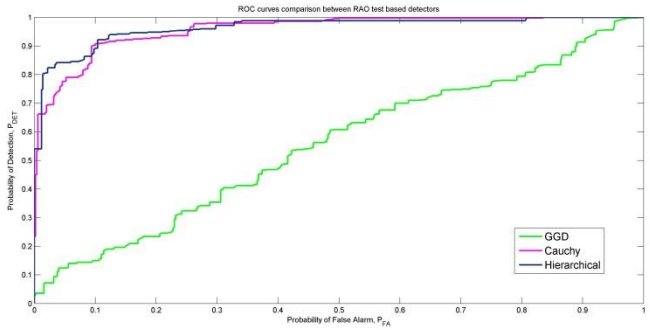


(c) "Bridge", WDR=-34.7 dB, PSNR=63.1 dB.

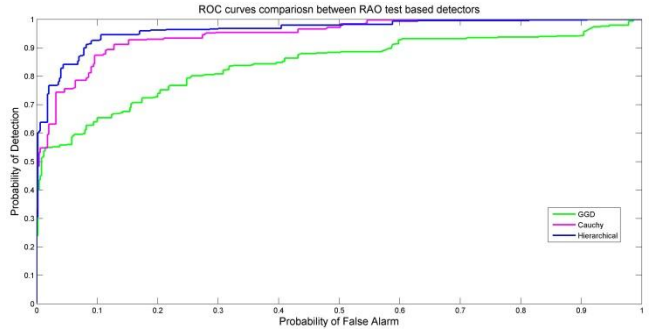


(d) "Barbara", WDR=-32.6 dB, PSNR=62.9 dB.

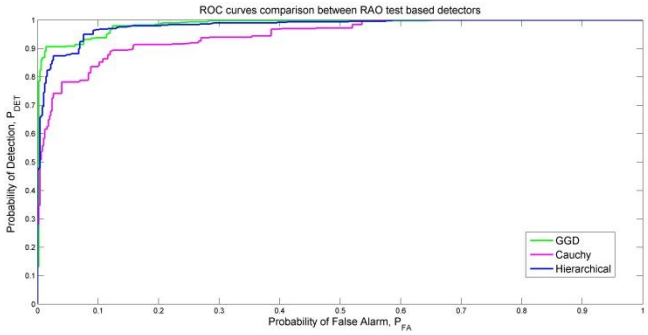
Figure 2 Detection performance comparison of the four known images (a) "Lena", (b) "Boat", (c) "Bridge", (d) "Barbara", without any kind of attacks



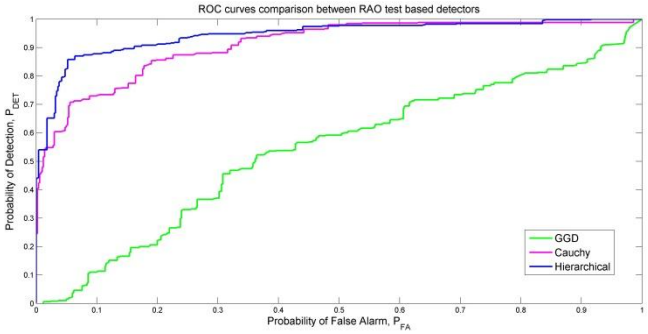
(a)"Lena", WDR=-27dB, PSNR=57dB.



(b)"Boat", WDR=-29.78dB, PSNR=56.4dB

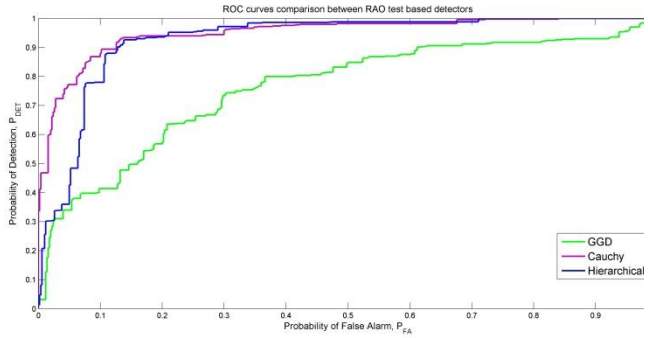


(c) "Bridge", WDR=-32.6dB, PSNR=57.8dB.

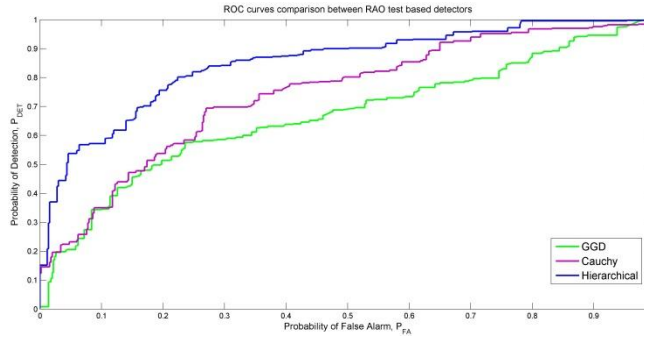


(d) "Barbara", WDR=-30.6 dB, PSNR=58dB.

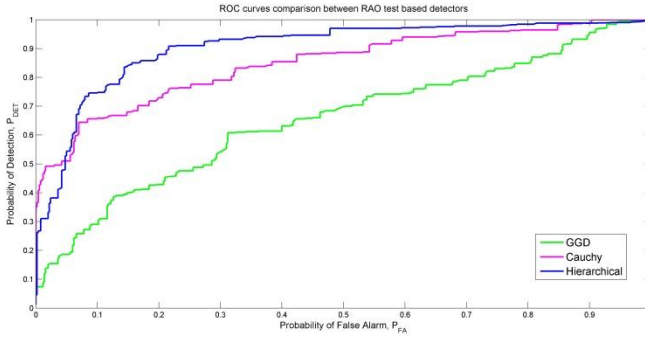
Figure 3. Detection performance comparison of the four known images (a) "Lena", (b) "Boat", (c) "Bridge", (d) "Barbara", under JPEG attack, with quality factor of 50%.



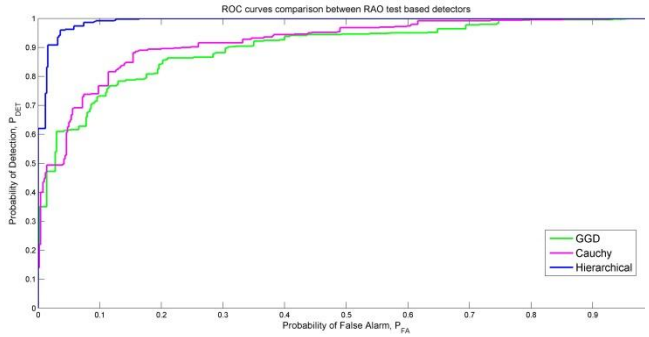
(a)"Lena", WDR=-29.4, PSNR=62 dB.



(a)"Boat", WDR=-29.7, PSNR=56.4 dB.

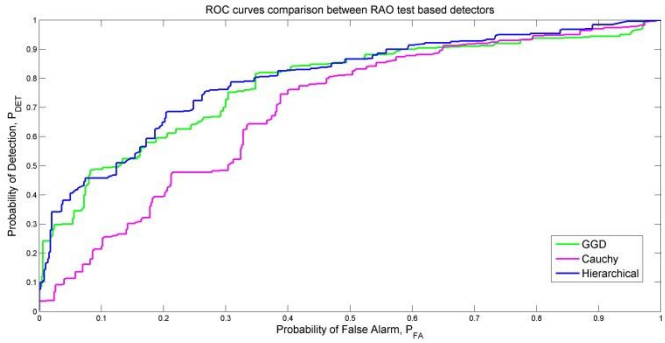


(c) "Bridge", WDR=-34.7, PSNR=63.1 dB.

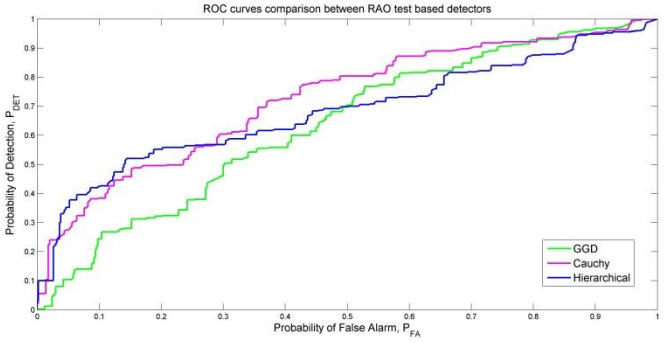


(d) "Barbara", WDR=-32.6, PSNR=62.9 dB.

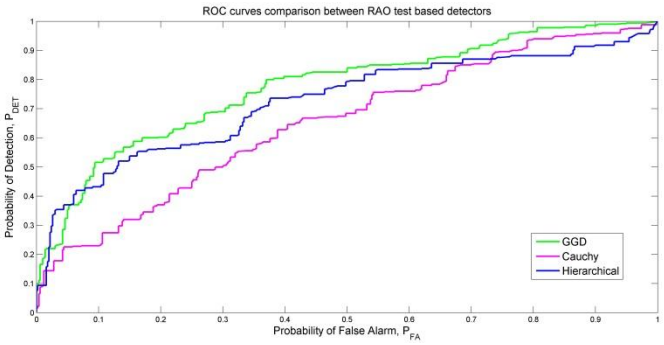
Figure 4. Detection performance comparison of the four known images (a) "Lena", (b) "Boat", (c) "Bridge", (d) "Barbara", under median filtering with window size equal to 5x5.



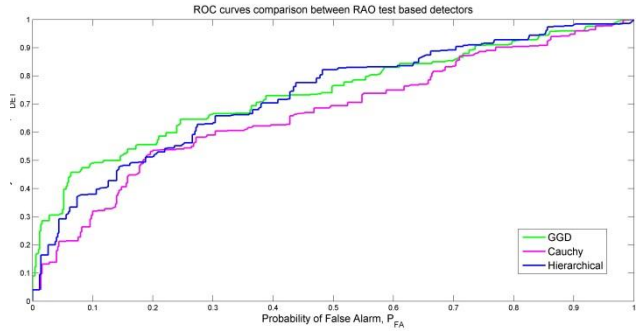
(a)"Lena", WDR=-28.4 dB, PSNR=52 dB.



(a)"Boat", WDR=-25 db, PSNR=51.7 dB.



(c) "Bridge", WDR=-27.64 db, 52.2 dB.



(d) "Barbara", WDR=-25.4 db, PSNR=52.2 dB.

Figure 5.Detection performance comparison of the four known images (a) "Lena", (b) "Boat", (c) "Bridge", (d) "Barbara", under AWGN.

3.4 *Experiments in a dataset of images*

In order to enhance the reliability of our results and verify the statistical significance of our proposal, we test the performance of the proposed detector using many more images by resorting to a dataset of images like the UCID image dataset [17]. Notice, that we produced the grey-scale versions of images at the size of 256x256 pixels in order to test our detector under more difficult circumstances. As we can observe from this kind of the experiment, the proposed detector has better detection sensitivity than the Rao-Cauchy detector, but has slightly inferior behavior than the Rao-GGD detector. Generally, we can say that the proposed detector has competitive performance compared to state-of-the-art detectors.

4 CONCLUSIONS

We investigated the validity of our proposal in the transform based additive watermarking problem and we further proposed a new watermark detector in wavelet domain. Motivated by the Rao hypothesis test, we employed a two-level hierarchical prior trying to model the host signal's coefficients. As our experimental results indicated, the proposed detector has competitive performance than other known state-of-the-art detectors using the same hypothesis. The robustness of the proposed detectors, using the aforementioned prior, against known attacks like JPEG compression, additive white Gaussian noise and median filtering has been studied.

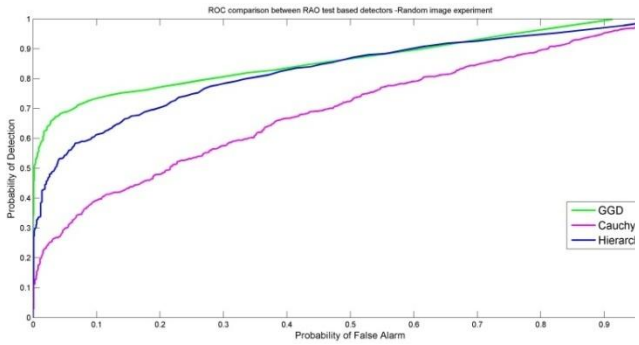


Figure 5. Detection performance comparison using a dataset of images [17]

REFERENCES

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

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

[3] G. C. Langelaar, I. Setyawan, R. L. Langendijk, "Watermarking digital image and video data", IEEE Signal Processing Magazine, vol. 17,no. 5, p.20 – 46, September 2000

[4] Cox IJ, Kilian J, Leighton FT, Shamoon T. "Secure spread spectrum watermarking for multimedia". IEEE Trans Image Process 1997;6(12):1673–87.

[5] 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

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

[7] A Nikolaidis, I. Pitas, "Asymptotically optimal detection for additive watermarking in the DCT and DWT domains", IEEE Trans. Image Process., vol. 12, no. 5, pp. 563-571, May. 2003.

[8] 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

[9] R. Kwitt, P. Meerwald, A. Uhl. "A lightweight Rao-Cauchy detector for additive watermarking in the DWT-Domain", Proc. of the 10th ACM workshop on Multim.and Sec, 33-42, Oxford, 2008

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

[11] Q. Cheng, T. S. Huang, "Robust optimum detection of transform domain multiplicative watermarks", IEEE Trans. On Image Processing, vol. 51, no. 4, p. 906 - 924, 2003

[12] Hamidreza Sadreazami; M. Omair Ahmad; M. N. S. Swamy, "A Study of Multiplicative Watermark Detection in the Contourlet Domain Using Alpha-Stable Distributions", IEEE Transactions on Image Processing, 23(10), pp. 4348 - 4360, 2014

[13] A. Mairgiotis, N. P. Galatsanos, Y. Yang, "New Additive Watermark Detectors Based On A Hierarchical Spatially Adaptive Image Model", IEEE Transactions on Information Forensics and Security 3(1): 29-37, 2008

[14] I. Daubechies, "Orthonormal bases of compactly supported wavelets," Commun. Pure Appl. Math., vol. 41, pp. 909–996, Nov. 1988.

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

[16] http://www.mathworks.com

[17] G. Schaefer and M. Stich, "UCID—an uncompressed colour image database", in Proceedings of SPIE, Storage and Retrieval Methods and Applications for Multimedia, vol. 5307, San Jose, CA, USA:SPIE, Jan. 2004, pp.472-480

,