DISTRIBUTION INDEPENDENT BLIND WATERMARKING

S.M.E. Sahraeian[*]*, M.A. Akhaee, F. Marvasti*

Advanced Communications Research Institute (ACRI)
Department of Electrical Engineering, Sharif University of Technology
[*] Department of Electrical and Computer Engineering, Texas A&M University
msahraeian@tamu.edu, akhaee@ee.sharif.edu, marvasti@sharif.edu

ABSTRACT

In this paper, a new blind scaling based watermarking approach is presented. The host signal is assumed to be stationary Gaussian with first-order autoregressive model. Partitioning the host signal into two separate parts, the data is embedded in one part and the other is kept unchanged for blind parameter estimation. Driving the distribution of the decision variable we have suggested a maximum likelihood decoding algorithm which is independent of the host signal distribution and can be applied for any transform domains. The proposed algorithm is applied to both artificial Gaussian autoregressive signals as well as various test images. Experimental results confirm the independence of the decoder performance to the host signal distribution and its great robustness against common attacks.

Index Terms— Watermarking, Gaussian Ratio distribution, scaling based embedding, Maximum likelihood decoder.

1. INTRODUCTION

Effective digital watermarking requires matching of the watermark signal characteristics to those of the carrier signal. Various tradeoffs of watermarking have been thoroughly investigated in the last decade. An essential mechanism of attaining robust and/or high-capacity watermarking resides in selecting image features that can bear larger changes without much reflecting it perceptually and that are more immune to innocent or intentional attacks [1]. Two known ways to implement this principle are: multiplicative method [1] and scaling-based approach [2]. These approaches are often applied in some transform domain, which has the advantage of energy concentration in a few robust components of the host signal. Since correlation detection is suboptimal for both multiplicative and scaling-based watermarking, several alternative optimum and locally optimum decoders have been proposed [1], [3], and [4]. Recently, a heuristic multi-bit data hiding method based on the multiband wavelet transform and the empirical mode decomposition is introduced [5]. However, these approaches suffer from the gain attack which usually occurs in the transmission channel.

In this paper, we introduce a distribution independent blind data hiding scheme coupled with a maximum likelihood decoder. The host signal is assumed to be a first-order Gaussian autoregressive process. To instrument a blind decoder, the host signal is divided into two patches so that the unmodified portion is used for parameter estimation, while its twin portion bears scaling-based watermark signal. The detector uses a decision variable resulting from the sum of samples, which under Central Limit Theorem (CLT) converges to a Gaussian variate. Under this assumption we introduce a Distribution Independent Optimum Decoder (DIOD) that works under any kind of distribution model such as Gaussian and GGD and in any transform domain such as wavelet, contourlet, ridgelet, and FFT. The proposed decoder is shown to be highly robust against common attacks.

2. SIGNAL MODELING

In this section we introduce the model we consider for our watermarking algorithm. We assume that the base signal is Gaussian with first order Markov. The carrier model, v, results from the sum of samples of two correlated signals. This watermark carrier signal model is independent of the distribution of the host signal. This opens a wider vista in that the carrier can be created from any multimedia signal, be it speech, audio, image in spatial domain or transform domain such as FFT, DCT, DWT, contourlet, ridgelet, etc.

Let, $\mathbf{u}$ be a first order Markov sequence with mean μ_u, variance σ_u^2, and correlation coefficient ρ_u. The N samples $u_1, u_2, ..., u_N$ are split between two subsequences $\mathbf{a}$ and $\mathbf{b}$ consisting of the odd and even indexed terms, respectively: $a_i = u_{2i-1}$, $b_i = u_{2i}$, $i = 1, 2, ..., \frac{N}{2}$.

Now, let's define the cumulative sums of the subsequences: $x = \sum_{i=1}^{\frac{N}{2}} a_i = \sum_{i=1}^{\frac{N}{2}} u_{2i-1}$ and $y = \sum_{i=1}^{\frac{N}{2}} b_i = \sum_{i=1}^{\frac{N}{2}} u_{2i}$ and v as:

$$v = \frac{x}{y} = \frac{\sum_{i=1}^{\frac{N}{2}} a_i}{\sum_{i=1}^{\frac{N}{2}} b_i} = \frac{\sum_{i=1}^{\frac{N}{2}} u_{2i-1}}{\sum_{i=1}^{\frac{N}{2}} u_{2i}} \tag{1}$$

It is of great interest that the above definition makes our watermarking method highly robust against gain attack as the gain attack scales both x and y equally and do not affect v.

The Central Limit Theory (CLT) suggests that for a sufficiently large N, x and y are Gaussian random variables with mean, variance, and correlation coefficient as:

$$\begin{aligned} \mu = \mu_x = \mu_y &= \frac{N}{2}\mu_u, \\ \sigma_x^2 = \sigma_y^2 &= [\frac{N}{2} + \sum_{i=1}^{\frac{N}{2}-1} (N-2i)\rho_u^{2i}]\sigma_u^2, \\ \rho_{xy} &= \frac{\sum_{i=1}^{\frac{N}{2}} \sum_{j=1}^{\frac{N}{2}} \rho_u^{|2i-2j-1|}}{\frac{N}{2} + \sum_{i=1}^{\frac{N}{2}-1} (N-2i)\rho_u^{2i}} \end{aligned} \tag{2}$$

Thus, the numerator and denominator of v are two correlated normal random variables for large N.

Under small coefficients of variation (CoV), that is, $\frac{\sigma_x^2}{\mu_x^2} \ll 1$ and $\frac{\sigma_y^2}{\mu_y^2} \ll 1$, the distribution of $v = \frac{x}{y}$, which is the ratio of two

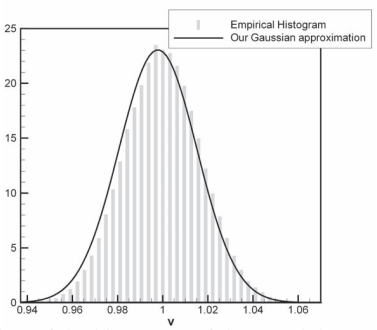


Fig. 1. Matching of the histogram of the v and the Gaussian approximation with mean and variance, respectively, in (5) and (7).

correlated random variables, for the case of non-zero means can be well approximated by a Gaussian distribution as:

$$f(v) = \frac{1}{\sqrt{2\pi}\sigma_v} e^{-\frac{1}{2\sigma_v^2}(v-\mu_v)^2} \tag{3}$$

This condition as verified in Section 5 is hold as an example case for the DWT wavelet approximation coefficients which have small coefficients of variation. The parameters μ_v and σ_v^2 of the approximated Gaussian distribution can be derived as follows.

Lets define $\Delta x = x - \mu_x$, $\Delta y = y - \mu_y$, and $\Delta v = v - \mu_v$. For a stationary process one has $\mu_x = \mu_y$ and $\sigma_x = \sigma_y$; furthermore small CoV implies that x and y signal excursions around their respective means are small, i.e., $\frac{\Delta x}{\mu_x} \ll 1$. Thus we have $\mu_v + \Delta v = \frac{\mu_x + \Delta x}{\mu_x + \Delta y}$, which can be approximated as:

$$\mu_v + \Delta v \simeq (1 + \frac{\Delta x}{\mu_x})(1 - \frac{\Delta y}{\mu_x}) \simeq (1 + \frac{\Delta x}{\mu_x} - \frac{\Delta y}{\mu_x} - \frac{\Delta x \Delta y}{\mu_x^2}) \tag{4}$$

Taking the expectation of both sides of (4) and considering the fact that As $E(\Delta x) = E(\Delta y) = E(\Delta v) = 0$, and $E(\Delta x \Delta y) = \rho_{xy}\sigma_x\sigma_y$, we have:

$$\mu_v = 1 - \frac{\rho_{xy}\sigma_x^2}{\mu_x^2} = 1 - \rho_{xy} A \tag{5}$$

where $A = \frac{\sigma_x^2}{\mu_x^2}$. Subtracting (4) from (5), we have:

$$\Delta v = \frac{\Delta x}{\mu_x} - \frac{\Delta y}{\mu_x} - \frac{\Delta x \Delta y}{\mu_x^2} + \frac{\rho_{xy}\sigma_x^2}{\mu_x^2} \tag{6}$$

Using zero skewness of Gaussian x and y variates, we have $E(\Delta x^2 \Delta y) = E(\Delta x \Delta y^2) = 0$ and considering the fourth moment of Gaussian we obtain $E(\Delta x^2 \Delta y^2) = \sigma_x^4(1 + 2\rho_{xy}^2)$. Finally, we can write the variance of the v variate:

$$\begin{aligned} \sigma_v^2 &= E(\Delta v^2) = 2\frac{\sigma_x^2}{\mu_x^2} + \frac{\sigma_x^4(1+\rho_{xy}^2)}{\mu_x^4} - 2\frac{\rho_{xy}\sigma_x^2}{\mu_x^2} \\ &= (2 - 2\rho_{xy})A + (1 + \rho_{xy}^2)A^2 \end{aligned} \tag{7}$$

Fig. 1 compares the histogram of v (bars) for a sample set of parameters $\mu_u = 1$, $\sigma_u = .2$, and $\rho_u = .7$ with the Gaussian approximation using (5) and (7). The good match between these two curves is an indication of the viability of this approximation.

3. PROPOSED BLIND WATERMARKING METHOD

3.1. Watermark embedding

The watermark is embedded by scaling the amplitude of the host signal in accordance with the message bit [2]. We assume a stationary Markov 1 sequence for the host signal $\mathbf{u}$. Given this cover sequence, we extract alternating subsample sequences $\mathbf{a}$ and $\mathbf{b}$ with odd and even indices, $a_i = u_{2i-1}$, $b_i = u_{2i}$. Then we embed data only in the $\mathbf{a}$ sequence based on the following scaling strategy:

$$a_i' = \begin{array}{ll} a_i \cdot \alpha & \text{For embedding 1} \\ a_i \cdot \frac{1}{\alpha} & \text{For embedding 0} \end{array} \tag{8}$$

where α is the strength factor and is larger than 1. Replacing $\mathbf{a}$ with $\mathbf{a}'$ in $\mathbf{u}$ yields the watermarked signal $\mathbf{u}'$. Therefore half of the samples are used for data hiding while the intact other half $\mathbf{b}$ lets as a reference.

3.2. Watermark decoding with Distribution Independent Optimum Decoder (DIOD)

Our ML-based decoder uses the cumulative sum statistics as in Section 2. The detector is then optimal for all distributions for which the sum of sequence terms converges rapidly enough to a Gaussian distribution.

The received sequence $\mathbf{u}'' = \mathbf{u}' + \mathbf{n}$ consists of the embedded subsequence $\mathbf{u}'$ contaminated by zero mean Additive White Gaussian Noise(AWGN) $\mathbf{n}$. Let's denote the subsequences with odd and even indices as $\mathbf{a}''$ and $\mathbf{b}''$, respectively, and their sums as $x'' = \sum_{i=1}^{\frac{N}{2}} a_i''$ and $y'' = \sum_{i=1}^{\frac{N}{2}} b_i''$. Under certain regularity conditions the x'' and y'' variables will behave as Gaussian random variables no matter what the distribution of $\mathbf{u}$.

The means and variances of these two random variables are calculated as $\mu_{x''|1} = \alpha\mu_x$, $\mu_{x''|0} = \alpha^{-1}\mu_x$, $\mu_{y''} = \mu_y = mu_x$, $\sigma_{x''|1}^2 = \alpha^2\sigma_x^2 + \sigma_N^2$, $\sigma_{x''|0}^2 = \alpha^{-2}\sigma_x^2 + \sigma_N^2$, and $\sigma_{y''}^2 = \sigma_x^2 + \sigma_N^2$, where the conditional variable is the polarity of the embedded bit. In these definitions σ_N^2 is the noise term in x'' and y''. This noise term can be shown to be $\sigma_N^2 = \frac{N}{2}\sigma_n^2$, where σ_n^2 is the variance of the AWGN channel noise $\mathbf{n}$. As given in [2], σ_n^2 can be easily estimate using the robust median estimator.

Moreover, The correlation coefficients between x'' and y'' under 1 and 0 embedding, represented by $\rho_{|1}'$ and $\rho_{|0}'$, can be computed in a straightforward way as in Section 2.

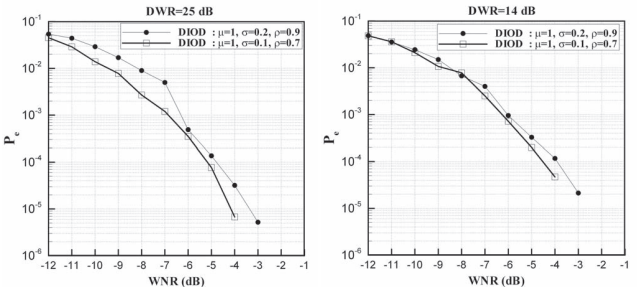
These are two correlated Gaussian random variables, x'' and y''. If we consider their ratio $v'' = \frac{x''}{y''}$, we revert to the model described in Section 2 and we can use the Gaussian distribution function suggested in(3). Accordingly, the ML decoder is given by $f(v''|1) \gtrless_0^1 f(v''|0)$. After some simplifications the ML decoder takes the form:

$$(\frac{1}{\sigma_{v|0}^2} - \frac{1}{\sigma_{v|1}^2})v^2 - 2(\frac{\mu_{v|0}}{\sigma_{v|0}^2} - \frac{\mu_{v|1}}{\sigma_{v|1}^2})v \gtrless_0^1 2\ln\frac{\sigma_{v|1}}{\sigma_{v|0}} + (\frac{\mu_{v|1}^2}{\sigma_{v|1}^2} - \frac{\mu_{v|0}^2}{\sigma_{v|0}^2}) \tag{9}$$

where $\mu_{v|1}$, $\mu_{v|0}$, $\sigma_{v|1}$, and $\sigma_{v|0}$ are the mean and variance of v for '1' or '0' embedding and can be computed in a similar way we compute μ_v and σ_v in Section 2.

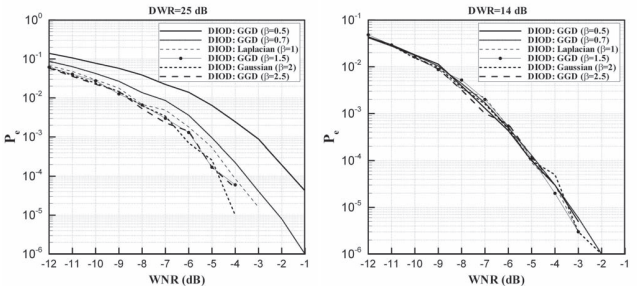
The host signal parameters μ_u, σ_u, and ρ_u are needed to detect the embedded bits. We can estimate these parameters from the reference signal b_i'', as:

$$\hat{\mu}_u = \mu_{b''}, \quad \hat{\sigma}_u = \sqrt{\max(\sigma_{b''}^2 - \sigma_n^2, 0)}, \quad \hat{\rho}_u = \sqrt{\rho_{b''}\frac{\hat{\sigma}_u^2 + \sigma_n^2}{\hat{\sigma}_u^2}}$$



(a) (b)

Fig. 2. Probability of error of different version of the DIOD for two test cases 1) $\mu_u = 1$, $\sigma_u = 0.2$, $\rho_u = 0.9$, 2) $\mu_u = 1$, $\sigma_u = 0.1$, $\rho_u = 0.7$. (a) DWR=25 dB, (b) DWR=14 dB.



(a) (b)

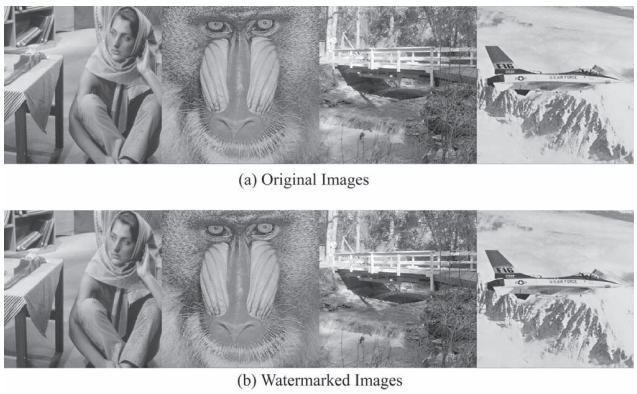
Fig. 3. Probability of error of the DIOD for six different i.i.d signals with the distributions of Gaussian, Laplacian, GGD with $\beta = \{0.5, 0.7, 1.5, 2.5\}$. (a) DWR=25 dB, (b) DWR=14 dB.

In conclusion, the proposed detector is optimal in the ML sense for any distribution of the u parent sequence that becomes Gaussian-like when summed over $N/2$ terms. We call this proposed decoding method as Distribution Independent Optimum Decoding (DIOD). Its distribution independence makes it suitable for implementation in different transform domains with any kind of distribution. Even it provides an optimum decoding for signals which has no specific distribution such as ridgelet coefficients.

It is noteworthy that here we considered odd indexed terms for data embedding and the even index terms for parameter estimation to find a closed form solution for the correlated host signal $\mathbf{u}$. For the case of independent and identically distributed (iid) signal, we can use any two random subsets of coefficients and only embed data in the coefficients of one subset and utilize the other one for parameter estimation. The indices of these subsets are produced by a random generator and the seed is sent to the decoder side through a secure channel. This way we can improve the security of the proposed technique. Moreover, for the case of the correlated signals, we can still implement these random subsets but we may have a a non-closed form solution.

4. PERFORMANCE EVALUATION

In this section, we investigate the performance of our watermarking method under AWGN attack of various strength. We fix Document-to-Watermark Ratio (DWR) in an experiment and plot the performance as a function of Watermark-to-Noise Ratio (WNR). We calculate 9 for two Gaussian first-order Markov sequences, namely, for the parameter sets $\mu_u = 1$, $\sigma_u = 0.2$, $\rho_u = 0.9$ and $\mu_u = 1$, $\sigma_u = 0.1$, $\rho_u = 0.7$. We test our DIOD decoder with DWRs of 25 dB and 14 dB. The results for WNRs in the range of $[-12, -1]$ are given in Fig. 2. We can see our decoder is highly robust even for very low WNRs.



(a) Original Images

(b) Watermarked Images

Fig. 4. (a)Original and (b) watermarked test images: *Barbara*, *Baboon*, *Bridge*, and *Plane*.

Next, we explore the performance of the DIOD decoder for different cover sequence distributions to substantiate the claim that our decoder is universal, that is, independent of the distribution. To verify this property, we test with distributions common in signal processing such as Gaussian, Laplacian, and Generalized Gaussian. The GGD distribution is a good model for the transform coefficients such as wavelet and contourlet. It is as $GG_{\sigma_x,\beta}(x) = C(\sigma_x, \beta)e^{-[\alpha(\sigma_x,\beta)|x-\mu|]^\beta}$, where $\alpha(\sigma_x, \beta) = \sigma_x^{-1}[\frac{\Gamma(\frac{3}{\beta})}{\Gamma(\frac{1}{\beta})}]^{\frac{1}{2}}$, $C(\sigma_x, \beta) = \frac{\beta\alpha(\sigma_x,\beta)}{2\Gamma(\frac{1}{\beta})}$, and σ_x is the standard deviation of x, β is the shape parameter, and $\Gamma(t)$ is the Gamma function.

We test the performance of our universal decoder for six different i.i.d GG with shape parameter varying over $\beta = \{0.5, 0.7, 1.5, 2.5\}$, which include the Gaussian and Laplacian cases. In all cases we set $\mu_u = 1$ and $\sigma_u = 0.2$. The results, given in Fig. 3 show that the proposed watermarking under with different DWRs and noise powers WNRs, has a high robustness against AWGN attack.

5. EXPERIMENTAL RESULTS

In order to evaluate the performance of the proposed technique on real condition, we have applied our watermarking method on several image signals and test theirs performance against various kinds of attacks. These results are obtained by averaging over 100 runs with 100 different pseudorandom binary sequences as the watermarking signal. We use Daubechies length-8 Symlet filters and embed the watermark information in the second approximation level using (8).

For this study, we use four natural images (*Barbara*, *Baboon*, *Bridge*, and *Plane*) of size 512×512. The original test images and their watermarked version using the proposed method with 16×16 block size and 64 bits message length are shown in Fig. 4. The strength factor α is fixed at 1.14 and it is assumed known to the decoder. The mean Peak-Signal-to-Noise-Ratio (PSNR) of the watermarked images are 41dB, 38dB, 35dB, and 36dB respectively.

First of all, we test the small coefficients of variation (CoV) assumption we made for the ratio signal in section 2. To this aim, the parameter $A = \frac{\sigma_u^2}{\mu_u^2}$ versus different approximation levels is plotted

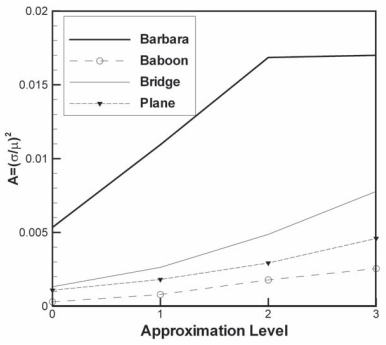


Fig. 5. Verifying the assumption of small $A = \frac{\sigma_u^2}{\mu_u^2}$ for different approximation levels.

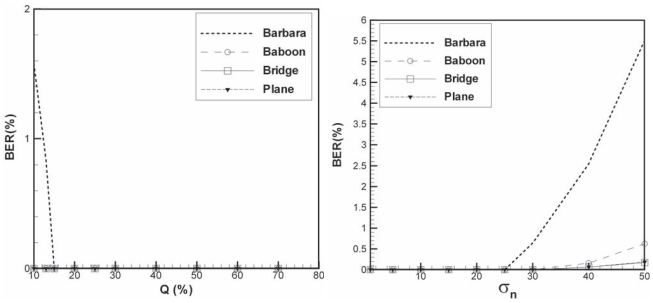


Fig. 6. Robustness of the proposed DIOD decoder for various test images under (a) JPEG compression attack, (b) AWGN attack.

in Fig. 5 for various test images. In this figure the zero approximation level represents the original image as host signal and further levels are first, second, and third approximation level of DWT decomposition. The validity of the assumed small A value is obvious in this figure.

Fig. 6 depicts the robustness results respectively against JPEG compression and AWGN attacks. The notable robustness of the DIOD method even against very low quality factors such as $Q = 10\%$ and intense noise attacks with $\sigma_n = 50$ verifies its outstanding quality for watermarking real signals.

Table 1 show the BER results of median filtering, and Gaussian low-pass filtering attacks for different test images with various window sizes. It can be seen that the proposed scheme is highly robust against these attacks.

Finally, we compare our watermarking algorithm with two of the recent blind watermarking techniques, Multiband Wavelets and Empirical Mode Decomposition (MWT-EMD) method [5], and Contourlet Scaling Based watermarking method (CSB) [6] for JPEG compression and AWGN attacks. The simulation results for *Baboon* test image are shown in Fig. 7. We see that the robustness of our method against JPEG and AWGN attacks are considerably better than MWT-EMD and CSB methods.

6. CONCLUSION

A distribution independent blind scaling-based data hiding approach suitable for multimedia signals has been introduced. The novelty of the algorithm is to create two alternating index sequences and embed into one of them while keeping the other unchanged. This enables us to obtain as sufficient statistics the ratio of summation of samples in two patches. Under reasonable assumption of Gaussian statistics as a result of CLT, a blind detector can be obtained. Furthermore, this detector is the optimum decoder independently of the host signal distribution and can be implemented in any transform domain with arbitrary distribution. Besides, it is highly robust against gain attack. Experimental results confirmed the distribution independence of the algorithm and its highly robustness under noise, JPEG attacks and filtering attacks.

Table 1. BER(%) Results of OD decoder under filtering attacks

Image	Median Filtering		Gaussian Filtering		
	3×3	5×5	3×3	5×5	7×7
Barbara	0.00	2.39	0.00	0.00	0.00
Baboon	0.00	5.03	0.00	0.00	0.00
Bridge	0.00	1.59	0.00	0.00	0.00
Plane	0.00	0.06	0.00	0.00	0.00

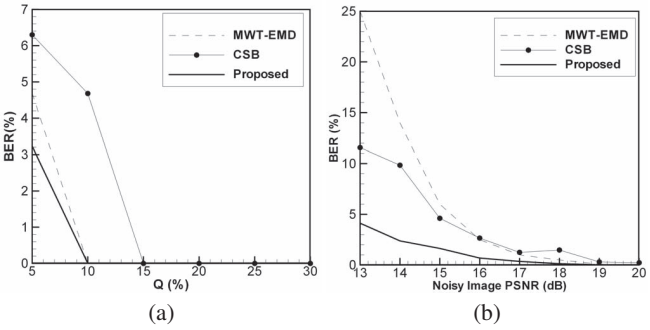


Fig. 7. Comparison between the proposed method and two other techniques: MWT-EMD method [5] and CSB [6] for *Baboon* test image under (a) JPEG compression, (b) AWGN attack.

7. REFERENCES

[1] M. Barni and F. Bartolini, *Watermarking Systems Engineering: Enabling Digital Assets Security and Other Applications*, CRC, 2004.

[2] M. A. Akhaee, S. M. E. Sahraeian, B. Sankur, and F. Marvasti, "Robust Scaling Based Image Watermarking Using Maximum Likelihood Decoder with Optimum Strength Factor," *IEEE Transaction on Multemedia*, accepted.

[3] Q. Cheng and T. S. Huang, "Robust optimum detection of transform domain multiplicative watermarks," *IEEE Trans. signal Process.*, vol. 51, no. 4, pp. 906–924, 2003.

[4] J. Wang, G. Liu, Y., Dai, and J. Sun, "Locally optimum detection for Barni's multiplicative watermarking in DWT domain," *Signal Processing*, vol. 88, pp. 117–130, 2008.

[5] N. Bi, Q. Sun, D. Huang, Z. Yang, and J. Huang, "Robust Image Watermarking Based on Multiband Wavelets and Empirical Mode Decomposition," *IEEE Transactions on Image Processing*, vol. 16, no. 8, pp. 1956–1966, 2007.

[6] M. A. Akhaee, S. M. E. Sahraeian, and F. Marvasti, "Contourlet Based Image Watermarking Using Optimum Detector in a Noisy Environment," *IEEE Transaction on Image Processing*, accepted.