WEAK SIGNAL WATERMARK DETECTION THROUGH RAO-T HYPOTHESIS AND LIGHTWEIGHT DETECTION

Antonis Mairgiotis[1], Lisimachos P. Kondi[1], Yongyi Yang[2]

[1]Dept. of Computer Science and Engineering
University of Ioannina, P.O. Box 1186
45110, Ioannina, Greece

[2] Dept. of Electrical and Comp. Engineering
Illinois Institute of Technology
Chicago IL, 60616, USA

ABSTRACT

In this work, we investigate an asymptotically optimal blind zero-bit watermark detector in the wavelet domain. More specifically, assuming that the marginal distribution of detail coefficients is non-Gaussian, we model it with the Student's t probability density function. Furthermore, we assume that the embedding power of the hidden information is unknown, suggesting in this way a new test statistic based on the Rao hypothesis test. The proposed detector exhibits better performance in terms of detection sensitivity and robust properties compared with other known methods in the framework of non-Gaussian environment. Additionally, we investigate a fixed-parameterization approach towards a lightweight detection with regard of time complexity.

Index Terms— Watermarking, wavelet domain, Rao hypothesis, Student's t, lightweight detection.

1. INTRODUCTION

The need for copyright protection and the prevention of unauthorized infringement of intellectual property has highlighted the need for efficient solutions. Digital image watermarking has been proved a candidate for such a protection mechanism, proposing models that satisfy conflicting requirements like imperceptibility, robustness and capacity [1], [2]. Thus, a piece of additional hidden information, the watermark, resides in image data in an imperceptible manner. Working in a statistical framework, we detect its presence or absence based on the appropriate statistical test, resorting to known concepts from statistical decision theory [11]. Applying blind watermarking, we have no knowledge of the original image, and each proposed scheme should be robust against processing under legitimate use or even under adverse efforts known as attacks [1]-[3].

The problem of watermarking is a fairly mature field with a substantial understanding of the limits and the capabilities of the various methodologies that have been proposed all these years. But certain issues continue to be investigated, such as the even finer control of the parameters that determine the compromise between them [3], the better application of the image model [4]-[9], security issues [17] and the implementation of watermarking in the compressed domain [18].

Transform domain methods have been proved particularly useful because of their inherent robustness properties e.g. due to better HVS (Human Visual System) exploitation. Thus, various approaches towards optimal detection have been proposed. Many of them use appropriate probability distributions with the basic criterion of better fitting the empirical distribution of transform coefficients. Based on the DCT (Discrete Cosine Transform) or DWT (Discrete Wavelet Transform) domains, various blind watermark detectors have been proposed. Viewing the watermarking problem as the transmission of a weak signal through a noisy channel, we consider wavelet detail coefficients as noise.

It is well recognized that the Gaussian noise model is not an appropriate choice when trying to detect the hidden information [10]. Thereby, over the years researchers have proposed optimal solutions based on appropriate statistical models showing better results than the linear correlator detector [1]-[3], [11]. The most widely used models are Generalized Gaussian Density (GGD) and Cauchy distribution [4]-[9]. Alternatively, other models have been proposed depending on more parameters e.g. Gauss-Hermite expansion and Bessel-K-form distributions or different domain like Weibull distribution in DFT domain [23]. Recently, SNIG (Symmetric Normal Inverse Gaussian) has also been proposed [4]. All these works assume that the strength factor is known at the receiver's side.

Notice that, in essence we have very weak signals hidden in the noise through which they are transmitted. Rao detector is an asymptotically optimal detector which is independent of the knowledge of the embedding strength factor. The initial application of this particular test in the framework of non-Gaussian noise was discussed by Kay in [15]. In the work of Nikolaidis et. al. [8], for the first time in watermarking, a new watermark detector was proposed, which assumes no knowledge of the value of the embedding power. Authors suggested the usage of Rao hypothesis in additive watermarking through a GGD noise model. Then, based on the known Cauchy member of the SaS (Symmetric

Alpha Stable) family of distributions, Kwitt. et. al. [7] proposed a new detector based on the same hypothesis.They also proposed lightweight versions of their proposals and they investigated the computational and time complexity of known detectors using various methods of parameter estimation [9]. The investigation of the Rao-Cauchy detector in the NSCT (non-subsampled contourlet transform) domain has also been presented [14].

Recently, we considered the application of another statistical modeling the framework of the additive watermarking problem in the wavelet domain, the Student-t distribution [16]. The t-distribution has similar form to the normal distribution but longer tails. This model was found to provide good fit of the empirical data along with a new class of watermark detectors. As a consequence, we improved detection sensitivity along with robustness against attacks. In this work, we consider a new class of DWT based watermark detectors based on the Rao hypothesis test and the Student's t probability density function (pdf). By taking advantage of this prior's characteristics, we consider the improved performance of the proposed image model in the wavelet domain. In addition, we provide a fixed-parameterization implementation, towards a lightweight version of the GLRT (Generalized Likelihood Ratio) detector [16].

The structure of the article is as follows. In Section 2, we describe the proposed statistical model, whereas in Section 3 we describe the watermark detection problem in the wavelet domain. Based on the aforementioned definitions, in Section 4, we propose the Rao-hypothesis test using the t-prior model. In Section 5, we investigate the detection performance of the proposed Rao-test and in Section 6, we propose a light version of the basic GLRT detector. In Section 7, we provide experimental results. Finally, in Section 8, we present our conclusions.

2. STATISTICAL MODEL

In our effort to adapt to the behavior of wavelet coefficients and define new suitable detectors for our problem, we propose the use of Student's-t probability density function [16].

The definition of the t-distribution is given by:

$$St(x;\mu,\sigma^2,\nu)=\frac{\Gamma\left(\left(\nu+1\right)/2\right)}{\Gamma\left(\nu/2\right)\sqrt{\pi\nu\sigma^2}}\left(1+\frac{(x-\mu)^2}{\nu\sigma^2}\right)^{-\frac{\nu+1}{2}} \quad (1)$$

where μ is the mean and σ^2 the known scale parameter. Parameter ν defines the degrees of freedom where $\nu\in(0,\infty)$ and is responsible for the control of the tail length. This means that for small values of ν we have considerable weight in the tails. For larger values of ν $(\nu\to\infty)$, the distribution increasingly resembles the normal distribution giving less weight in the tails. The variance of the distribution, is $\sigma\nu/(\nu-2)$ for $\nu>2$ and infinite if $0<\nu<2$. The t-distribution can be thought of as a weighted sum of normal distributions with the same mean, but variance that depends inversely on the gamma distribution. The Student-t pdf may take the form of a marginalization of the joint distribution with respect to a hidden variable h meaning that:

$$St(x\mid\mu,\nu,\sigma)=\int N(\mu,\sigma\mid h)G(\nu/2,\nu/2) \quad (2)$$

where with N we denote the Normal distribution and G is the Gamma distribution defined by:

$$G(a,b)=\frac{b^a}{\Gamma(\alpha)}\exp\left(-bh\right)h^{a-1} \quad (3)$$

In order to estimate the aforementioned parameters, we resort to the known EM (Expectation Maximization) algorithm, where we can find the ML (Maximum Likelihood) solution [16]. Since the lengths of the tails are parameterized, the effect of this choice is the improved description of wavelet detail subband marginal coefficients. Based on this parameterization of the tail, we can fit the empirical data even better. Thus, we can uncover the small differences, which will help us detect the weak energy of watermark information.

3. WATERMARK DETECTION PROBLEM

Working in the wavelet domain, we consider N detail subband coefficients, without loss of generality. These coefficients have the role of host noise and are assumed to be independent identically distributed random samples from a probability density function (pdf). We can define them in a vector notation by $\mathbf{x}=[x_1,x_2,...,x_N]$. Following an additive embedding rule and given that the hidden information is given by $\mathbf{w}=[w_1,w_2,...,w_N]$, then the watermarked coefficients in some subband is given by the relation:

$$\mathbf{y}=\mathbf{x}+a\mathbf{w} \quad (4)$$

where a is the strength factor and $\mathbf{y}=[y_1,y_2,...,y_N]$ is the watermarked data vector. The presence or absence of the watermark enables us to define our problem as a binary hypothesis testing [11]:

$$\begin{aligned} &H_0: y_i=x_i & i=1,...,N \\ &H_1: y_i=x_i+aw_i & i=1,...,N \end{aligned} \quad (5)$$

where H_0 denotes the null hypothesis and H_1 denotes the alternative hypothesis. In case complete knowledge of the parameters of the distributions is not available, we can define the composite hypothesis testing [11].

4. RAO HYPOTHESIS

Based on statistical decision theory, we can apply three known test statistics: GLRT, Wald and Rao test [11]. The advantage of the Rao test against the alternatives lies in the hypothesis testing problem for non-Gaussian noise [15]. It also has the same asymptotic $(N\to\infty)$ detection performance as the GLRT and its main advantage stems from the fact that it is easier from a computational point of view. This happens because it is not necessary to perform parameter estimation under the alternative hypothesis. Thus, we could consider the problem at hand as:

$$\begin{aligned} H_0 &: a = 0,\ \theta_{s,H_0} = \nu,\ \sigma^2 \\ H_1 &: a \neq 0,\ \theta_{s,H_1} = \nu,\ \sigma^2 \end{aligned} \tag{6}$$

The unknown parameters of the noise (called nuisance parameters) θ_{s,H_0} and θ_{s,H_0} are not directly related to the problem of detection, but they have an impact in the form of the distribution in both cases under consideration. Thus, we resort to the Rao hypothesis test which requires only the ML estimates under the null hypothesis. In addition, we assume that the host noise follows the Student-t distribution.

Notice also, that this is an important degree of freedom, since it allows the embedding side to adapt the embedding strength to the signal at hand [2],[7], [9]. In what follows, we introduce a Rao hypothesis test conditioned on a Student's-t signal noise model. The Rao test decides H_1 if:

$$T_{RAO}(y) = \left.\frac{\partial \ln p(y;\Theta)}{\partial \alpha}\right|^{\mathrm{T}}_{\Theta=\hat{\Theta}} \left[\mathbf{I}^{-1}\left(\hat{\Theta}\right)\right]_{\alpha\alpha} \left.\frac{\partial \ln p(y;\theta)}{\partial \alpha}\right|_{\Theta=\hat{\Theta}} \overset{H_1}{>} thr \tag{7}$$

where $\mathbf{I}(\Theta)$ is the Fisher information matrix. Notice that $\hat{\Theta} = [0, \hat{\theta}_{s,\mathrm{H}_0}]$ denotes the ML estimates under H_0. The term

$$\left[\mathbf{I}\left(\hat{\Theta}\right)\right]^{-1}_{\alpha\alpha} = \left[\mathbf{I}_{\alpha\alpha}(\Theta) - \mathbf{I}_{\alpha\theta_s}(\Theta)\mathbf{I}^{-1}_{\theta_s\theta_s}(\Theta)\mathbf{I}_{\theta_s\alpha}(\Theta)\right]^{-1} \tag{8}$$

consists of the elements $\mathbf{I}_{\alpha\alpha}, \mathbf{I}_{\alpha\theta_s}, \mathbf{I}^{-1}_{\theta_s\theta_s}, \mathbf{I}_{\theta_s\alpha}$ which denote the partitions of the Fisher information matrix, which are defined in [7], [9]. After some calculus, we can shave the form of the proposed detector as [8], [11], [15]:

$$T_{RAO-T}(y) = \frac{\left(\sum_{i=1}^{N} w_i \frac{p'(y)}{p(y)}\right)^2}{\frac{1}{N}\sum_{i=1}^{N} w_i^2 \sum_{i=1}^{N}\left(\frac{p'(y)}{p(y)}\right)^2} > thr \tag{9}$$

where $p(.)$, $p(.)'$ denotes the pdf and its derivative of host noise. Applying this test using the t-prior we can define the Rao-t detector which is given by:

$$T_{RAO-T}(\mathbf{y}) = \frac{\left(\sum_{i=1}^{N} w_i \frac{(\nu+1)y_i}{\sigma^2\nu + y_i^2}\right)^2}{\sum_{i=1}^{N}\left(\frac{(\nu+1)y_i}{\sigma^2\nu + y_i^2}\right)^2} > thr \tag{10}$$

where we omitted the term with the watermark in the denominator since doesn't affect the results.

5. DETECTION PERFORMANCE

Assuming that we have large data records, it is known from Kay [11] that the proposed detector follows a chi-squared distribution under both detection hypotheses, meaning that:

$$\begin{aligned} H_0 &: T_{RAO-T} \sim \chi_1^2 \\ H_1 &: T_{RAO-T} \sim \chi_{1,\lambda}^2 \end{aligned} \tag{11}$$

where χ_1^2 denotes a Chi-Square distribution with one degree of freedom and $\chi_{1,\lambda}^2$ denotes a noncentral Chi-Square distribution with one degree of freedom and noncentrality parameter λ. We know that if a random variable follows a non-central $\chi_{1,\lambda}^2$ then it is equivalent to $N(\sqrt{\lambda},1)$

The noncentrality parameter λ is defined as:

$$\lambda = \alpha^2\left[\mathbf{I}_{\alpha\alpha}(0,\theta_s) - \mathbf{I}_{\alpha\theta_s}(0,\theta_s)\mathbf{I}^{-1}_{\theta_s\theta_s}(0,\theta_s)\mathbf{I}_{\theta_s\alpha}(0,\theta_s)\right] \tag{12}$$

Eventually we end up in the form,

$$\left[\mathbf{I}\left(\hat{\Theta}\right)\right]^{-1} = \left(\mathbf{I}_{\alpha\alpha}(\Theta)\right)^{-1} \tag{13}$$

since we know that due to the existence of a symmetric pdf, $\mathbf{I}_{\alpha\theta_s} = 0$ and Eq. (8) remains with only one term [9], [11], [15]. In Figs. 1, 2 we can observe the superiority of the proposed Rao-t based watermark detector compared with the known Rao-GGD [8] and Rao-Cauchy [7] detectors in the wavelet domain with (JPEG attack) and without attacks.

6. FIXED-PARAMETER LIGHTWEIGHT DETECTION

Motivated by Hernandez et al. [4] and Kwitt et al. [7] a critical question was raised, meaning whether the detection process actually benefits from an accurate estimate of the pdf's parameters or if it is possible to apply approximate or fixed parameter settings. Their findings point to the fact that we can find a fixed value setting. Thus, we can have high detection sensitivity over a whole set of natural images, with small deviations from the optimal performance.

Here, in the same spirit as [7], we investigate the detection performance of the proposed detector. Our target is to examine the detector's performance in accordance to runtime complexity with more or less accurate estimates of parameters or fixed parameter settings. In order to escape effort in the estimation of parameters e.g. based on ML estimation, the value of $c = 0.8$ has been proposed for DWT coefficients [4] using GGD prior. According to Kwitt et al., $c = 1$ is another proposed value that works efficiently. Based on the Cauchy distribution, we don't have any known reference to indicate which value may be a good approach except of the work of Kwitt et al. [7]. For an image independent value, $\gamma = 8$ seems to be a very good suggestion, while the value of $\gamma = 3$ seems to be impractical. With regard to the Bessel-K detector, we resort to the work of Bian and Liang [12] where, in order to alleviate the usage of estimated parameters for acceleration of the calculations, a value of $\delta = 0.85$ is suggested. Following a similar approach to [7], we derived the histogram of the values of the degrees of freedom parameter, using ML estimates. We created the corresponding histograms and we computed the mean values, describing the distribution of degrees of freedom values based on three different datasets that we can see in Table I. The values are related with the first and second level of wavelet transform for the referred sizes.

TABLE I. ESTIMATED VALUES OF DEGREES OF FREEDOM PARAMETER ν FOR THE HORIZONTAL DETAILS SUBBANDS

dataset (number of images)	size	1st level	2nd level
UCID (1000) [21]	256x256	1.18	1.25
Microsoft db (200) [20]	512x512	1.40	1.63
BOWS (10000) [22]	512x512	1.30	1.21

7. LIGHTWEIGHT DETECTION PERFORMANCE

Viewing watermark detection in the framework of lightweight detection, we control the performance of the detector when we predefine the parameter values (e.g., keeping them on a fixed value) and obtain a reduction in time complexity. Keeping some values fixed, we achieve a reduction (in a total of 1000 images) of about half of the total time of the proposed experiment. In Table II we observe that the proposed lightweight version of the proposed detection scheme verifies the results in [8]. More specifically, when we demonstrate the results in the table form, we utilize the area under the ROC (AUROC1) curve in the range of [0-0.1] describing the detector's performance at low false-alarm rates and we also make use of the total area under the ROC (AUROC2) describing the overall performance of the detector [12], [16]. For the quantification of the watermark's power in our experiments we used the known WDR (Watermark to Document ratio) definition [12], [16]:

$$WDR = 20\log_{10}\left(\frac{\|aw\|}{\|x_s\|}\right) \quad (14)$$

where the term "document" refers to the original host data [5] and x_s is the original host data in spatial domain. Thus, we can find fixed parameter settings for the proposed host t-based noise model, which result in competitive or superior detection performance compared with other similar lightweight versions of known GLRT detectors. Notice that, in this framework, we suggest $\nu = 1.25$ as a practical solution for a lightweight version of our proposal. The proposed value consists in a candidate value which its validity is verified from the extended experimental results of the proposed work in Table II.

8. CONCLUSIONS

In this work, we investigated the Rao hypothesis using the Student's t distribution and we verified its detection superiority compared with the known Rao-GGD and Rao-Cauchy based watermark detectors. In addition, we provided a fixed parameterization for Student's t distribution and we also verified its detection performance compared with other known detectors. By proposing a lightweight version we manage to alleviate the ML estimation of parameters without significantly sacrificing detection performance. Such a detection approach is a good match when the resources of the embedding side are constrained and thus allow only a simple embedding strategy [7]. Thus, depending on the application environment, one has the option to select the lightweight version due to the usefulness of the fast detector.

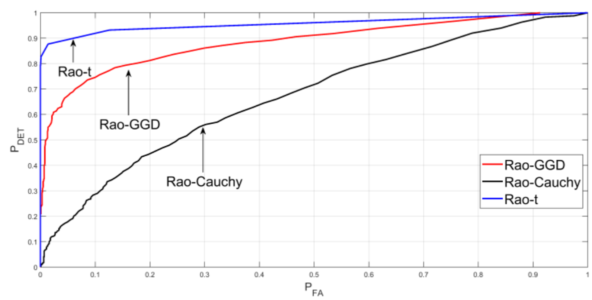

Fig. 1. Comparison of detectors using the Rao hypothesis (UCID images, 256x256), without attacks, WDR=-61dB.

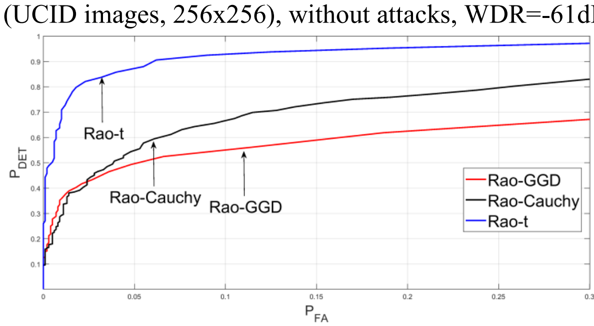

Fig. 2. Comparison of detectors using the Rao hypothesis (UCID, 256x256) under JPEG attack, WDR=-55dB.

TABLE II. AUROC (AUROC2, AUROC1) RESULTS FOR LIGHTWEIGHT VERSIONS OF GLRT BASED DETECTORS

WDR	*GGD (c=0.8) [4], [7]*	*Cauchy (γ=8) [7]*	*Bessel – K (δ=0.85) [12]*	*Student t (mean value: ν=1.25)*	*Student t (Cauchy member λ=1, ν=1)[16]*
-58	0.9855, 0.0905	0.9972, 0.0964	0.9039, 0.0939	0.9947, 0.0942	0.9868, 0.0907
-59	0.9690, 0.0862	0.9836, 0.0868	0.8836, 0.0819	0.9877, 0.0912	0.9722, 0.0846
-60	0.9475, 0.0817	0.9659, 0.0777	0.8391, 0.0690	0.9559, 0.0744	0.9478, 0.0728
-61	0.8644, 0.0514	0.9022, 0.0550	0.7748, 0.0515	0.8922, 0.0566	0.8761, 0.0494
-62	0.7079, 0.0229	0.8232, 0.0337	0.6333, 0.0234	0.7821, 0.0310	0.7399, 0.0248
-63	0.6337, 0.0156	0.6932, 0.0161	0.5764, 0.0161	0.6692, 0.0159	0.6212, 0.0123

9. 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, and R. L. Lagendijk, “Watermarking digital image and video data: A state-of-the-art overview,” *IEEE Signal Process. Mag.,* vol. 17, no. 5, pp. 20–46, Sep. 2000.

[4] J. Hernandez, M. Amado, and 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, pp. 55–68, Jan. 2000.

[5] A. Briassouli, P. Tsakalides, A. Stouraitis, “Hidden Messages in Heavy-Tails: DCT-Domain Watermark Detection Using Alpha-Stable Models,” *IEEE Trans. On Multimedia,* vol. 7, no. 4, pp. 700–715, Jan. 2005.

[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, pp. 273–284, Sep. 2001.

[7] R. Kwitt, P. Meerwald, A. Uhl, “Lightweight detection of Additive Watermarking in the DWT-Domain,” *IEEE Trans. On Image Processing*, vol.20, no.2, pp. 474-484.

[8] A. Nikolaidis, I. Pitas, “Asymptotically optimal detection for additive watermarking in the DCT and DWT domains,” *IEEE Trans.on Image Processing*, vol. 12, no.5, pp. 563-571, May 2003.

[9] R. Kwitt, P. Meerwald, A. Uhl, “A lightweight Rao-Cauchy detector for additive watermarking in the DWT-domain,” in Proc. of the *ACM Multimedia and Security workshop*, Oxford UK:ACM, pp.33-41, Sept.2008.

[10] S. A. Kassam, J. Bowman Thomas, *Signal detection in non-Gaussian noise*, Springer-Verlag, 1988.

[11] S. Kay, “Fundamentals of Statistical Signal Processing, Volume II: Detection Theory,” Prentice-Hall, 1998, vol. II.

[12] Y. Bian, S. Liang, “Locally Optimal Detection of Image Watermarks in the Wavelet Domain using Bessel K Form Distribution,” *IEEE Trans.on Image Processing*, vol. 22, no. 6, pp. 2372-2384, Feb 2013.

[13] S.M.M. Rahman, M.O.Ahmad, M.N.S. Swamy, “A New Statistical Detector for DWT-Based Additive Image Watermarking Using the Gauss–Hermite Expansion,” *IEEE Trans.on Image Processing*, vol. 18, no. 8, pp. 1782-1796, 2009.

[14] H.Sadreazami, M. Omair Ahmad, M. N. S. Swamy, “Multiplicative Watermark Decoder in Contourlet Domain Using the Normal Inverse Gaussian Distribution,” *IEEE Trans. On Multimedia*, vol18., no. 2, pp. 196-207, 2015.

[15] S. M. Kay, “Asymptotically optimal detection in incompletely characterized non-gaussian noise,” *IEEE Trans. on Acoustic, Speech and Signal Processing*, vol. 37, no.5, pp. 627-633, May 1989.

[16] A. Mairgiotis, Y. Yang, L. P. Kondi, “DWT-based additive image watermarking using the Student-t prior Information,” *IEEE Int. Workshop on Inf. Forensic and Security (WIFS),* pp. 1 – 6, Nov-Dec 2011.

[17] T. Bianchi, A. Piva, “Secure Watermarking for Multimedia Content Protection,” *IEEE Signal Processing Magazine*, pp. 87-96, March, 2013.

[18] A. Boho, G. Van Wallendael, A. Dooms, J. De Cock; G. Braeckman; P. Schelkens; B. Preneel; R. Van de Walle, “End-To-End Security for Video Distribution: The Combination of Encryption, Watermarking, and Video Adaptation,” *IEEE Signal Processing Magazine*, Vol. 30, No. 2, pp.97 – 107, Feb., 2013.

[19] H. Bi, Y. Liu, M. Wu, Y. Ge, “NSCT domain additive watermark detection using RAO hypothesis test and Cauchy distribution,” *Mathematical Problems in Engineering*, vol.2016.

[20] Microsoft Research Cambridge Object Recognition Image Database, http://research.microsoft.com/downloads.

[21] 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.

[22] P.Bas and T. Furon. Bows-2.http://bows2.ec-lille.fr/, December 2017.

[23] M. Barni, F. Bartolini, A. De Rosa, and A. Piva, “A new decoder for the optimum recovery of nonadditive watermarks,” *IEEE Trans.on Image Processing*, vol. 10, pp. 755–766, May 2001.