EUROGRAPHICS Workshop on Multimedia (2004)
N. Correia, J. Jorge, T. Chambel, Z. Pan (Editors)

Optimum Detection of Multiplicative Watermarks for Digital Images in the DWT Domain

Zhongwei Sun*, Rui Xue and Dengguo Feng

State Key Laboratory of Information Security, Institute of Software , Chinese Academy of Sciences

Beijing, People's Republic of China

Abstract

Watermark detection plays a crucial role in digital watermarking. It has traditionally been tackled using correla-tion-based techniques. However, the correlation-based detection is not the optimum choice when the host media doesn't follow a gaussian distribution or the watermark is not embedded in the host media in an additive way. A discrete wavelet transform (DWT) domain multiplicative watermark detection algorithm for digital images is propo-sed in this paper, which exploits the imperceptibility constraint of watermarking. By formulating the watermark detection as weak signal detection in non-gaussian noise, the proposed algorithm is derived according to statistical inference theory. With the wavelet coefficients modeled by generalized gaussian distribution (GGD), the optimum decision threshold for the detector is obtained by applying Neyman-pearson criteria. The superiority of the novel detector in performance is confirmed through Monte Carlo simulations.

Keywords: Digital watermarking, Multiplicative embedding, Discrete wavelet transform, Generalized gaussian dis-tribution, Weak signal detection.

1. Introduction

Digital processing and transmission of multimedia information permit convenient access to multimedia data, but the very same properties also lead to the problems regarding copyright protection. Digital watermarking has been proposed as a mean to protect an owner's right or tracing pirate [CMB02]. Watermarking is the process of embedding hidden copyright information into the original multimedia data, and the embedding is performed in such a way that the watermark is imperceptible under normal observation conditions. Driven by the need of the protection of intellectual property rights, a great deal of research has been carried out in this field.

Generally speaking, a watermark can be embedded into an image either in the spatial domain or in the transformed domain. Although the requirements of a watermarking system are related to the application, the robust detection of the embedded watermark is needed for all applications. Digital watermarking is similar to the communication task, in which the original media plays the role of channel noise. Based on the assumption of additive gaussian white noise (AGWN) channel, most researchers use some kind of correlation-based detector to detect the watermarks. According to the signal detection theory [Poo94], the correlation-based detector is not optimum when the host signal doesn't

* Corresponding author. E-mail: sunzwcn@yahoo.com.cn

© The Eurographics Association 2004.

delivered by EUROGRAPHICS DIGITAL LIBRARY www.eg.org diglib.eg.org

follows a gaussian distribution or the watermark is not embedded in the host media in an additive way. Aiming at overcoming the problem of the unreasonable assumption of AGWN channel, Vidal and Sayrol [VS98] assume that, after a whitening process before watermark detection, the host data samples in the spatial domain follow a Cauchy distribution. As transform domain watermarking schemes are robust to many signal processing and compression attacks, they have attracted a lot of attention recently. In the discrete cosine transform (DCT) domain, Hernendez *et al.* [HAF00] propose to use the generalize Gaussian distribution (GGD) to model the alternative current coefficients for additive watermark detector. Barni *et al.* [BBDP01] construct a multiplicative watermark detector in the discrete Fourier transform (DFT) domain. However, both detection schemes have assumed that the embedding strength is known to the detector, and this assumption is impractical for blind watermark detection, which does not need the original images in the detection phase. Cheng and Huang present a multiplicative watermark detection scheme based on local optimum detection (LOD) in [CH03]. The LOD is actually a one-sided hypothesis test, and they have implicitly made the assumption that the embedding strength factor is positive. As a matter of fact, the embedding strength factor can be positive or negative, but the watermark detector does not have the knowledge about this for blind watermark detection, so their method is effective only when the embedding strength factor is positive.

The imperceptibility is one of the basic requirements for invisible watermarks. It has made the watermark detection a weak signal detection problem. Based on the theory of statistical inference and weak signal detection in non-gaussian noise, this paper presents a discrete wavelet transform (DWT) domain multiplicative watermark detection algorithm for digital images. The rest of this paper is organized as follows. The multiplicative watermark detection algorithm is introduced in section 2. The implementation of the proposed detector is discussed in section 3. The experimental results are presented in section 4. The conclusion is given in section 5.

2. Multiplicative watermark detection

The watermark is embedded into the detailed wavelet coefficients of the host image in a multiplicative way:

$$y_i = (1+\alpha w_i)x_i, \quad i = 1,\cdots,n \tag{1}$$

where $X = \{x_1, x_2, \cdots x_n\}$ is a sequence of the original DWT coefficients, $W = \{w_1, w_2, \cdots w_n\}$ is a sequence of watermark signals, α is the embedding strength factor, $Y = \{y_1, y_2 \cdots y_n\}$ is a sequence of watermarked data. The watermark W is a zero-mean sequence, whose components are independent and uniformly distributed in [-1,1]. W is independent of X. Due to the imperceptibility requirement, $|\alpha| < 1$. Taking the inverse transform, we get the watermarked image.

From the viewpoint of statistical theory, the watermark detection process can be treated as a statistical inference process. We define two hypotheses: the null hypothesis is that a given watermark is not present, and the alternative hypothesis is that it is present. Consequently, the watermark detection problem can be mathematically formulated as following binary hypothesis testing problem[Tre68]:

$$\begin{aligned} H_0 &: y_i = x_i \quad , & i = 1,\cdots,n \\ H_1 &: y_i = x_i(1+\alpha w_i), & i = 1,\cdots,n \end{aligned} \tag{2}$$

As the original image is not available at the detection stage, it must be modeled by noise, while the watermark is the desired signal. The goal of the designed detector is to decide whether or not there is a watermark in the received image, based on the statistical properties of the given data. Let the DWT coefficients have a pdf of $p_x(x_i)$. Under H_0, the pdf of y_i is

$$p_{y_i}(y_i; H_0) = p_x(y_i) \tag{3}$$

Under H_1, as $|\alpha m_i| < 1$, the pdf of y_i is

$$\begin{aligned} p_{y_i}(y_i; \alpha, H_0) &= \frac{1}{|1+\alpha m_i|} p_x\left(\frac{y_i}{1+\alpha m_i}\right) \\ &= \frac{1}{1+\alpha m_i} p_x\left(\frac{y_i}{1+\alpha m_i}\right) \end{aligned} \tag{4}$$

It has been shown in image coding applications that the high-frequency DWT coefficients can be reasonably well approximated by zero-mean GGD, and its probability density function (pdf) is given by the expression [CYV00, ABMD92]

© The Eurographics Association 2004.

$$p_x(x) = Ae^{-|\beta x|^c} \quad (5)$$

where $\beta = \frac{1}{\sigma}\left(\frac{\Gamma(3/c)}{\Gamma(1/c)}\right)^{1/2}$, $A = \frac{\beta c}{2\Gamma(1/c)}$, $\Gamma(\cdot)$ is a Gamma function, σ is the standard deviation of the distribution, c is the shape parameter. Following the same assumption as in [CH03], where the wavelet coefficients are independently and identically distribution (i.i.d.), we can see that the joint probability density of the observed sequence Y is simply the product of the probability densities of the single observation y_i, i.e.,

$$p_Y(Y;H_0) = \prod_{i=1}^{n} Ae^{-|\beta y_i|^c} \quad (6)$$

$$p_Y(Y;\alpha,H_1) = \prod_{i=1}^{n} \frac{A}{1+\alpha m_i} e^{-\left|\frac{\beta y_i}{1+\alpha m_i}\right|^c} \quad (7)$$

The log likelihood ratio between $p_Y(Y;\alpha,H_1)$ and $p_Y(Y;H_1)$ is:

$$\begin{aligned} l(Y) &= \ln \frac{p_Y(Y;\alpha,H_1)}{p_Y(Y;H_0)} \\ &= \ln p_Y(Y;\alpha,H_1) - \ln p_Y(Y;H_0) \end{aligned} \quad (8)$$

It can be seen from the binary hypothesis testing problem defined by (2) that the null hypothesis H_0 and the alternative H_1 are equivalent to $\alpha = 0$ and $\alpha \neq 0$ respectively. From the imperceptibility constraint, the embedding strength factor is assumed to be small. Using a first-order Taylor expansion of $\ln p_Y(Y;\alpha,H_1)$ about $\alpha = 0$, we have

$$l(y) = \left.\frac{\partial \ln p_Y(Y;\alpha,H_1)}{\partial \alpha}\right|_{\alpha=0} \cdot \alpha \quad (9)$$

According to the definition of likelihood ratio test [Poo94], $l(Y)$ is compared with a decision threshold λ_0 to decide if a given image contains a watermark. For blind watermark detection, α is unknown to the detector. Since α can take positive or negative values, the hypothesis testing problem formulated by (2) is a two-side test, the detector decides H_1 if $\left.\frac{\partial \ln p_Y(Y;\alpha,H_1)}{\partial \alpha}\right|_{\alpha=0} > \frac{\lambda_0}{\alpha}$ for $\alpha > 0$, and the detector decides H_0 if $\left.\frac{\partial \ln p_Y(Y;\alpha,H_1)}{\partial \alpha}\right|_{\alpha=0} > \frac{\lambda_0}{\alpha}$ for $\alpha < 0$, the choice of which depends on the unknown parameter α, so the detector is unrealizable in the sense of likelihood ratio test or local optimum detection (LOD). For the testing statistic $\left.\frac{\partial \ln p_Y(Y;\alpha,H_1)}{\partial \alpha}\right|_{\alpha=0}$, we define a decision rule for the aforementioned testing problem as follows:

$$\left|\left.\frac{\partial \ln p_Y(Y;\alpha,H_1)}{\partial \alpha}\right|_{\alpha=0}\right| \underset{H_0}{\overset{H_1}{\gtrless}} \lambda \quad (10)$$

where $\lambda = \lambda_0 / |\alpha|$. By using (7), $\left.\frac{\partial \ln p_Y(Y;\alpha,H_1)}{\partial \alpha}\right|_{\alpha=0}$ can be expressed as

$$\left.\frac{\partial \ln p_Y(Y;\alpha,H_1)}{\partial \alpha}\right|_{\alpha=0} = \sum_{i=1}^{n} c|\beta y_i|^c w_i - \sum_{i=1}^{n} w_i \quad (11)$$

Since W is a zero-mean sequence, $\sum_{i=1}^{n} w_i = 0$, the decision criteria defined in (10) is equivalent to

$$\left|\sum_{i=1}^{n} c|\beta y_i|^c w_i\right| \underset{H_0}{\overset{H_1}{\gtrless}} \lambda \quad (12)$$

3. Actual implementation of the detector

3.1 pdf estimation of DWT coefficients

The discrete wavelet transform for two-dimensional images can be implemented by Mallat algorithm [Mal89]. When a two-dimensional DWT is applied to an original image, one low frequency subband and three high frequency subbands are achieved. This process is continued with the low frequency subband, till the desired levels are achieved. An example of such decomposition with 3 levels is shown in figure 1, where the arrows indicate the order of decreasing importance.

© The Eurographics Association 2004.

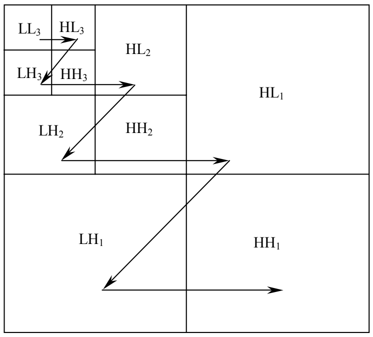


Figure 1: *Wavelet decomposition of the image*

The distribution of the DWT coefficients is the basis for the development of detector that will be proposed. It can be seen in (5) that the GGD of the high frequency DWT coefficients is completely determined by the parameters σ and c. Following the approach proposed in [BS99], where the parameters are solved by minimizing relative entropy, we fit the histogram of the high-frequency subband coefficients using GGD to estimate the parameters σ and c. The relative entropy between the assumed distribution and the empirical distribution is calculated as

$$\Delta H(\sigma_i, c_c) = -\sum_{k=1}^{n} h_k \log \frac{F_x(c_n)}{h_k} \qquad (13)$$

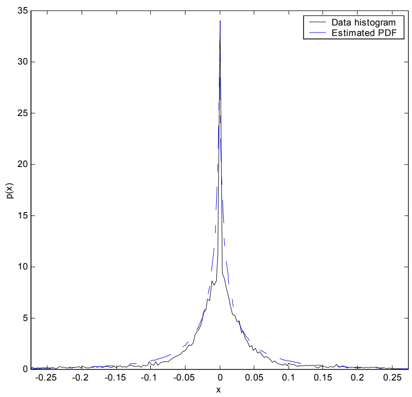


Figure 2: *Histograms of the subband and the pdfs using GGD for Lena with* $c_i = 0.4413$, $\sigma_i = 0.0913$

The smaller the $\Delta H(\sigma_i, c_i)$, the better the fit is. Figure 2 shows the histograms of the coefficients in subband HL_2 of Lena decomposed through DWT using db7/9 filter, together with plots of fitted density function of GGD. It can be seen that the empirical pdfs from histograms are in good agreement with the pdfs using the GGD.

3.2 Threshold selection

It is the decision threshold that determines the performance of the detector. Let $T(Y) = \left.\frac{\partial \ln p_Y(Y;\alpha, H_1)}{\partial \alpha}\right|_{\alpha=0}$. When H_0 holds $y_i = x_i$. The sufficient statistic $T(Y)$ becomes

$$T(Y) = \sum_{i=1}^{n} c\left|\beta x_i\right|^c w_i \qquad (14)$$

As $x_i s$ are independent and follow a stationary GGD and $w_i s$ are i.i.d, $c\left|\beta x_i\right|^c w_i$ are i.i.d random variable. By invoking the central limit theorem, the statistic $T(Y)$ is an asymptotically gaussian random variable when n is large enough. The mean of $T(Y)$ is

$$\begin{aligned}\mu = E(T(Y); H_0) &= E(\sum_{i=1}^{n} c\left|\beta x_i\right|^c w_i) \\ &= c\left|\beta\right|^c \sum_{i=1}^{n} (E(\left|x_i\right|^c) E(w_i)) = 0\end{aligned} \qquad (15)$$

And the variance of $T(Y)$ is

$$\begin{aligned}\sigma^2 = \mathrm{var}(T(Y); H_0) &= \mathrm{var}(\sum_{i=1}^{n} c\left|\beta x_i\right|^c w_i) \\ &= c\left|\beta\right|^c \sum_{i=1}^{n} (E(\left|x_i\right|^{2c}) E(w_i^2))\end{aligned} \qquad (16)$$

Since w_i s are independent and uniformly distributed in [-1,1], we have

$$E(w_i^2) = 1/3 \qquad (17)$$

The term $E(\left|x_i\right|^{2c})$ is evaluated by

$$E(\left|x_i\right|^{2c}) = \int_{-\infty}^{\infty} \left|x_i\right|^{2c} p(x_i) dx_i = \frac{2A\Gamma(2+1/c)}{c\beta^{2c+1}} \qquad (18)$$

The resulting variance is

$$\begin{aligned}&\sigma^2 = \mathrm{var}(\sum_{i=1}^{N} c\left|\beta x_i\right|^c w_i) = \\ &c\left|\beta\right|^c \sum_{i=1}^{N} (E(\left|x_i\right|^{2c}) E(w_i^2)) = \frac{2nA\Gamma(2+1/c)}{3\beta^{c+1}}\end{aligned} \qquad (19)$$

It can be shown that under H_0, the asymptotic pdf of $T(Y)$ doesn't depend on the unknown parameter α. Define $Q(Y)=\int_x^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2\sigma^2}}dt$. By applying Neyman-pearson criteria, namely, the decision threshold λ is chosen in such a way that detection probability P_D is maximized while the false alarm probability P_{FA} is constrained to an acceptable value, λ can be computed through the following relation:

$$
\begin{aligned}
P_{FA} &= P_{T_{Norm}}(|T_{Norm}(Y)| > \lambda; H_0) \\
&= P_{T_{Norm}}(T_{Norm}(Y) > \lambda; H_0) + P_{T_{Norm}}(T_{Norm}(Y) < -\lambda; H_0) \\
&= 2Q(\lambda) \qquad (20)
\end{aligned}
$$

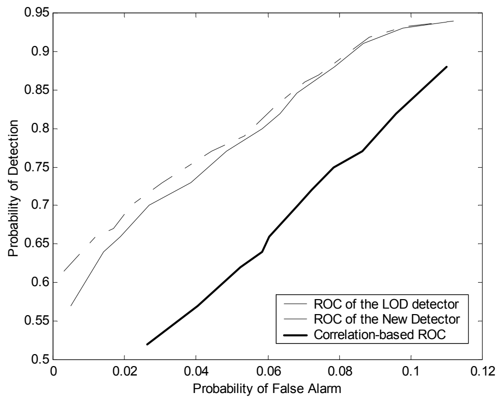
or equvalently

$$\lambda = \sigma Q^{-1}(P_{FA}/2) \qquad (21)$$

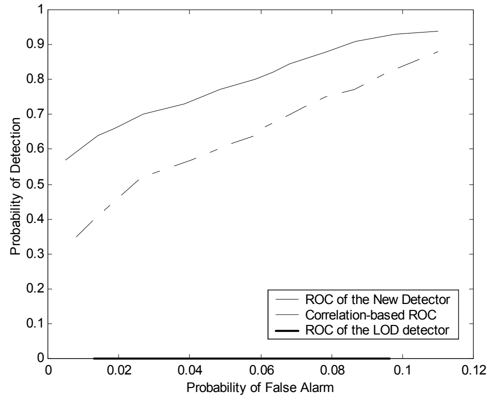
where $Q^{-1}(x)$ is inverse of $Q(x)$.

4. Experimental Results

Experiments are conducted to measure the performance of the proposed watermark detector. In particular, the Lena image of size 512*512 is used as the host image. It is decomposed through DWT into three levels using db7/9 filter. The watermark is then inserted by modifying significant wavelet coefficients. The significant wavelet coefficients are those with large magnitude. We generate 5000 watermark sequences and use Monte Carlo simulations to measure the performance of the proposed watermark detector. The detection performance is measured by receiver operation characteristic (ROC), which plots the variation of the detection probability against the probability of false alarm.



(a) $\alpha = 0.08$



(b) $\alpha = -0.08$

Figure 3: *Experimental and theoretical ROC curves*

At every Monte Carlo experiment, the test statistic $T(Y)$ is calculated according to equation (12). If the absolute value of $T(Y)$ is above the threshold λ under H_1, the watermark is detected. While if the absolute value of $T(Y)$ is above the threshold λ under H_0, a false alarm occurs. In this manner, the experimental ROCs are derived.

For comparison purpose, similar experiments are also conducted for correlation-based detector and LOD detector. Following the approaches of [CH03] and [CKLS97], we design correlation-based detector and LOD detector respectively, without explicitly exploit the visual masking. The test statistic $T_{Cor}(Y)$ for corre-

© The Eurographics Association 2004.

lation-based detector and the test statistic $T_{LOD}(Y)$ for LOD detector are defined as follows:

$$T_{Cor}(Y) = \left| \sum_{i=1}^{N} y_i w_i \right| \quad (22)$$

$$T_{LOD}(Y) = \frac{1}{N} \sum_{i=1}^{N} c \left| \beta_i y_i \right|^c w_i \quad (23)$$

Experimental results are shown in figure 3. Figure 3(a) shows that for the positive embedding strength, the performance of the proposed detector and the LOD detector are close to each other, though the LOD detector performs slightly better than the new detector. And the performance of both detectors outperforms the correlation-based detector. But for the negative embedding strength, as shown in figure 3(b), the probability of detection for LOD detector is close to zero, while the proposed detector still outperforms the correlation-based detector. It demonstrates that the proposed detector has a robust detection performance.

5. Conclusion and future work

This paper deals with the issues of watermark detection for digital images in the DWT domain. As the original image is not available during watermark detection, it is modeled by noise. Therefore, the watermark detection can be seen as a statistical decision problem. By formulating the watermark detection as weak signal detection in non-gaussian noise, a multiplicative watermark detection scheme is derived according to statistical inference theory. Two key problems in the construction of the corresponding watermark detector, namely, the probability density function estimation of the DWT coefficients and the decision threshold selection for the proposed detector, are solved. Experimental results demonstrate that the proposed detector presents a superior performance over the correlation-based detector and the LOD detector.

Acknowledgements

This work is supported in part by the Chinese National 973 Project (G1999035802) and National Science Foundation of China (60025205 and 60373048). We also thank the anonymous EGMM 2004 reviewers for the suggentions to improve the paper.

References

[ABMD92] Antonini M., Barlaud M., Mathieu P. Daubechies A.: Image coding using wavelet transform. *IEEE Trans. Image Processing*, 1992, 1(2): 205-220.

[BBDP01] Barni M., Bartolini F., De Rose A., Piva M.: A new decoder for the optimum recovery of non-additive watermarks. *IEEE Trans. on Image Processing,* 2001, 10(5): 755-765.

[BS99] Buccigrossi R. W , Simoncelli P.: Image compression via joint statistical characterization in the wavelet domain. *IEEE Trans Image Processing*, 1999, 8(12):1688-1701.

[CH03] Cheng Q., Huang T.S.: Robust optimum detection of transform domain multiplicative watermarks. *IEEE Trans. Signal processing*, 2003, 51(4): 906-924.

[CKLS97] Cox I.J., Killian T., Leighton T., Shamoon T.: Secure spread spectrum watermarking for multimedia. *IEEE Trans. On Image processing*,1997, 6(12):1673-1687.

[CYV00] Chang S.G., Yu B., Vetteri M.: Adaptive wavelet thresholding for image denoising and compression. *IEEE Trans. Image Processing*, 2000, 9(9): 1532-1546.

[CMB02] Cox I.J., Miller L., Bloom A.: *Digital watermarking*. Morgan Kaufmann Publishers, 2002

[HAF00] Hernández J. R., Amado M., and Fernando Pérez-González: DCT-domain watermarking techniques for still images: Detector performance analysis and a new structure. *IEEE Trans. Image Processing* 2000, 9(1): 55-68.

[Mal89] Mallat, S.: A theory of multi-resolution signal decomposition: The wavelet representation. *IEEE trans. on pattern analysis and machine intelligence*, 1989, 11(7): 674-693.

[Poo94] Poor H.V.: *An introduction to signal and estimation*. New York: Springer-Verlag, 1994.

© The Eurographics Association 2004.

[Tre68] Trees H.L.V.: *Detection, estimation and modulation theory*. New York: Wiley, 1968

[VS98] Vidal J., Sayrol E. Optimum watermark detection and embedding in digital images. *Proc. IEEE workshop on multimedia signal processing*, California, 1998.

© The Eurographics Association 2004.