Correlation Detection of Asymmetric Watermark

Jin S. Seo and Chang D. Yoo

Korea Advanced Institute of Science and Technology, Department of EECS,
373-1 Kusong-dong, Yusong-gu, Daejeon 305-701, Korea
pobi@eeinfo.kaist.ac.kr, cdyoo@ee.kaist.ac.kr

Abstract. This paper proposes a novel method to detect Furon's asymmetric watermark by using a correlation detector that is mathematically tractable and simple. The performance of the proposed method is tested under various conditions. The experimental results matched the theoretical results well, showing that the correlation detector can indeed be used for the detection of asymmetric watermark. The proposed detector is applied to both single and multiple bit embedded watermark. Bit error rate (BER), obtained from the experiment, was compared to the one obtained from the theory.

1 Introduction

With the advent of Internet, there has been an explosive growth in the use of digital media. Since digital media is easily reproduced and manipulated, anyone is potentially capable of incurring considerable financial loss to the media producers and content providers. In this respect, digital watermarking is essential.

Most of the existing watermarking methods use symmetric key, that is to say the same key or pattern is used in the embedding and detection. Thus the secrecy of the key is shared by the embedder and detector. In situations where the detector must be available to the public, the secrecy can be divulged by tampering the detector. Based on public-key crypotography, T. Furon addresses this problem with asymmetric watermarking. In his work [1], the presence of filtered watermark that is considered as an output of a filtered random process is detected using only the knowledge of the magnitude of frequency response of the filter.

In this paper, a simple and mathematically tractable detection method is proposed for the detection of asymmetric watermark. It is based on the theory of detecting a known signal in noisy channel. The known signal is the power spectrum of the embedded watermark, and the noise is the estimation error. The estimation error of the power spectrum is assumed to be additive, uncorrelated and Gaussian noise. By using periodogram averaging in the power spectrum estimation, the assumptions are satisfied. The advantages of periodogram averaging over periodogram used in [1] are the reduction of the variance of the power specturm estimate and computational load in estimating the power spectrum. The optimum threshold of the correlator output is set using the Neyman-Pearson

H.-Y. Shum, M. Liao, and S.-F. Chang (Eds.): PCM 2001, LNCS 2195, pp. 638–645, 2001.
© Springer-Verlag Berlin Heidelberg 2001

lemma that maximizes the probability of correct detection for a given false alarm probability.

Embedding one bit of information is not sufficient for the real application, such as DVD copy protection [2], thus we modified T. Furon's method to increase information rate. By using PN sequences in filter shaping, multiple bits of information can be conveyed. BER, obtained from the experiment, followed the one obtained from the theory favorably. The reliability of the correlation detector is verified by the experiment.

This paper is organized as follows. Section 2 explains the asymmetric watermark embedding. Section 3 describes the proposed detection method. Section 4 describes the multiple bit embedding and detection. Section 5 shows the experimental results.

2 Correlation detector

For asymmetric watermark embedding, filtered watermark pattern is embedded after interleaving. As in [3], the symmetric method given in [4] is translated into asymmetric method. The filtered watermark pattern is embedded into the interleaved DFT magnitude coefficients of the image as shown in Fig. 1. Details of embedding are in [1] and [3].

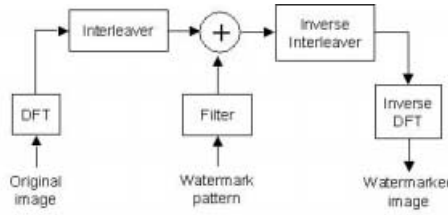

Fig. 1. Asymmetric watermark embedding in DFT domain

In order to detect the embedded watermark, the following binary hypothesis test was used in the interleaved domain as in [3]. The DFT magnitude coefficients of the received image are denoted by r_u. The interleaved signal of r_u is denoted by $\tilde{r}_u$. The Fourier transform of the filter used in the embedding is denoted by $H(f)$.

- H_0 : r_u is not watermarked, so $\tilde{r}_u$ is a white noise. The power spectrum of $\tilde{r}_u$ is as follows:

$$g_0(f) = \mu_{\tilde{r}_u}^2 \delta(f) + \sigma_{\tilde{r}_u}^2. \tag{1}$$

- H_1 : r_u is watermarked, so $\tilde{r}_u$ is a colored noise. The power spectrum of $\tilde{r}_u$ is as follows:

$$g_1(f) = \mu_{\tilde{r}_u}^2 \delta(f) + \sigma_{\tilde{r}_u}^2 + \gamma^2 \mu_{\tilde{r}_u}^2 (|H(f)|^2 - 1). \tag{2}$$

Under this hypothesis the power spectrum is shaped by $|H(f)|^2$.

The detection is formulated as detecting a known signal in noisy channel. The known signal is the power spectrum of the embedded watermark, and the noise is the estimation error. The power spectrum estimation error is assumed to be additive. First, we observe three properties of periodogram for the frequencies $\{f_k = k/N, 0 \le k \le N/2\}$ where N is the length of r_u [5].

P1) The mean of the periodogram is given by

$$\text{Mean}\{I_N(f_k)\} = P_{\tilde{r}_u}(f_k) + \mathcal{O}(N^{-1}) \tag{3}$$

where $I_N(f) = \frac{1}{N}|\sum_{k=0}^{N-1} \tilde{r}_u(k)e^{j2\pi kf}|^2$ which is the periodogram of $\tilde{r}_u$ and $P_{\tilde{r}_u}(f)$ is the true power spectrum of $\tilde{r}_u$.

P2) The variance of the periodogram is given by

$$\text{var}\{I_N(f_k)\} = \begin{cases} 2P^2_{\tilde{r}_u}(f_k) + \mathcal{O}(N^{-1}) & k = 0 \text{ or } \frac{N}{2} \\ P^2_{\tilde{r}_u}(f_k) + \mathcal{O}(N^{-1}) & \text{otherwise.} \end{cases} \tag{4}$$

P3) The covariance of the periodogram at different frequencies is given by

$$\text{cov}\{I_N(f_k), I_N(f_l)\} = \mathcal{O}(N^{-1}) \quad k \neq l\ . \tag{5}$$

From P1, the mean of the estimation error at the frequencies f_k can be regarded as zero. From P2, the variance of the estimation error σ_e^2 is given by equation (7) for $k \neq 0$ and $N/2$:

$$\begin{aligned} \sigma_e &= \sigma^2_{\tilde{r}_u} + \gamma^2\mu^2_{\tilde{r}_u}(|H(f_k)|^2 - 1) \\ &\approx \sigma^2_{\tilde{r}_u} \quad \text{when } \gamma\mu_{\tilde{r}_u} \ll 1 \end{aligned} \tag{6}$$

when the watermark exists. In the case that the watermark does not exist, σ_e is also given by $\sigma^2_{\tilde{r}_u}$. From P3, the estimation error for the frequencies f_k is uncorrelated.

Second, the probability of distribution of the estimation error should be considered. Through the periodogram averaging, the distribution of the estimation error becomes Gaussian. Periodogram averaging consists of three steps. First, the N-length sequence $\tilde{r}_u$ is subdivided into K nonoverlapping segments, where each segment has length M. This results in the K data segments. For each segment, the periodogram is computed for the frequencies $\{f_k = k/M, 0 \le k \le M/2\}$. By averaging K data segments, the power spectrum estimate is obtained. The most important advantage of periodogram averaging is the distribution of the estimation error can be regarded as normal by the central limit theorem. Due to the averaging of K independent periodograms, the distribution of the estimation error approaches normal distribution. In practice, the normal approximation is good even for the small value of K. The second advantage of periodogram averaging is the reduced variance of the estimation error by the factor of K. Thus the variance of the estimation error is $\sigma_e^2 = \sigma^4_{\tilde{r}_u}/K$. But periodogram averaging

reduces the resolution by the factor of K. There is a trade-off between the variance and the resolution of the power spectrum estimate. The required resolution is determined by the frequency response of the filter. The third advantage of periodogram averaging is the reduction of computational load by using K small size DFT (M point) instead of one large size DFT (N point).

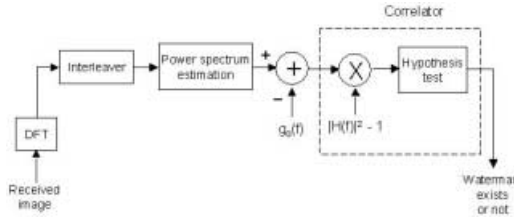

Fig. 2. Proposed correlation detector of asymmetric watermark

From the above, the estimation error can be regarded as additive, uncorrelated and Gaussian, then the detection problem can be transformed into the following hypothesis test shown in Fig. 2. A test function t, that can be modeled as some nominal value ρs plus noise with variance σ_e^2, is obtained by subtracting g_0 from the power spectrum estimation as follows:

$$t(f_k) = I_N^K(f_k) - g_0(f_k) = \rho s(f_k) + \sigma_e n(f_k) \tag{7}$$

where I_N^K is the averaged periodogram obtained from K data segments, $\{f_k = k/M, 1 \leq k \leq M/2 - 1\}$, $s(f_k) = |H(f_k)|^2 - 1$ and $n(f_k)$ is distributed as multivariate normal distribution $N(0, I)$. It needs to determine whether $H_0 : \rho = 0$ or $H_1 : \rho > 0$ from $t(f_k)$ that is distributed as $N(\rho s, \sigma_e^2 I)$. From the Fisher-Neyman factorization theorem, the sufficient statistic for the parameter ρ is

$$m = \frac{< s, t >}{\sigma_e (< s, s >)^{1/2}}. \tag{8}$$

The statistic m is distributed as $N(\frac{\rho\sqrt{E_s}}{\sigma_e}, 1)$ where $< s, t >= \sum_{k=1}^{M/2-1} s(f_k)t(f_k)$ and $E_s =< s, s >$ [6]. In communication theory, this hypothesis test based on m is known as correlation detector.

Let m_0 be the detection threshold in determining whether $\rho = 0$ ($m < m_0$) or $\rho > 0$ ($m \geq m_0$). By the Neyman-Pearson lemma, the false alarm probability P_{FA} is given by

$$P_{FA} = \int_{m_0}^{\infty} (2\pi)^{-1/2} e^{-x^2/2} dx = \frac{1}{2} erfc\Big(\frac{m_0}{\sqrt{2}}\Big). \tag{9}$$

For a certain value of P_{FA}, the threshold m_0 can be set by solving the above equation. The detection reliability is determined by the threshold m_0. The cor-

rect detection probability P_D is given by

$$P_D = \int_{m_0}^{\infty} (2\pi)^{-1/2} e^{-\left(x-\frac{\rho\sqrt{E_s}}{\sigma_e}\right)^2/2} dx$$
$$= \frac{1}{2} erfc\Big((m_0 - \frac{\rho\sqrt{E_s}}{\sigma_e})/\sqrt{2}\Big). \quad (10)$$

In the Furon's method [3], the probability density function of the estimated power spectrum (periodogram) is assumed Laplacian, while in the proposed method that involves periodogram averaging in the power spectrum estimation, Gaussian is assumed. The Gaussian assumption makes the detection problem more mathematically tractable.

3 Multiple bit embedding and detection

To increase information rate from a single bit to L bits, the frequency bins are divided into L bands; $(l-1)\frac{0.5}{L} < f < l\frac{0.5}{L}$ where $l = 1, 2, \cdots, L$. Each of the information bit is modulated by a PN sequence. In the l-th band, $|H(f)|$ is shaped as follows:

$$|H(f)|^2 = 1 + (-1)^{mes(l)} R_l(f) \quad (11)$$

where $R_l(f)$ is generated by a PN sequence with zero sum and $mes(l)$ is a information bit as shown in Fig 3. For the length of each PN sequence D, periodogram averaging factor K should satisfy the following inequality

$$K \leq \frac{1}{C}\Big(\frac{N}{2LD}\Big) \quad (12)$$

to meet the resolution requirement (experimental choice of C was 3).

In the detection, the power spectrum is estimated from K periodogram averaging and the sufficient statistic m_l of the l-th band is defined as follows:

$$t_l(f_k) = \rho R_l(f_k) + \sigma_e n(f_k) \quad (13)$$

where $(l-1)\frac{0.5}{L} < f_k < l\frac{0.5}{L}$.

$$m_l = \frac{\sum_k R_l(f_k) t_l(f_k)}{\sigma_e (\sum_k R_l^2(f_k))^{1/2}} \quad \text{for } l = 1, 2, \ldots, L. \quad (14)$$

For each band, the detection is performed by testing the three hypotheses:

- H_0 : The watermark does not exist. ($\rho = 0$)
- H_1 : The embedded information is 0. ($\rho > 0$)
- H_2 : The embedded information is 1. ($\rho < 0$)

in the model m_l is distributed as $N(\frac{\rho\sqrt{E_{R_l}}}{\sigma_e}, 1)$ where $E_{R_l} = \sum_k R_l^2(f_k)$.

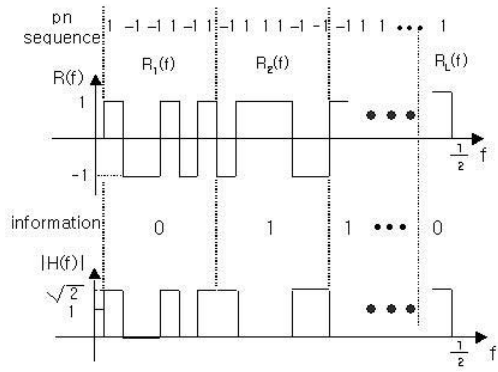

Fig. 3. Multiple bit embedding

4 Experimental results

In order to validate the proposed scheme, we tested our scheme on 512 by 512 "Lena" image. The embedding is the same as in [3] and [4] except that HVS (human visual system) was not used ($\gamma = 0.22$).

To show the validity of the proposed detection method, the outputs of the correlation detector, obtained from ten different interleavers and 100-tap FIR filters, were averaged. Fig. 4 shows the behavior of the statistic m versus the averaging factor K. On the whole, the correlation output was well-matched to the theoretical expectation. Even in the case of $K = 1$ (periodogram), the experimental correlator output followed the theoretical expectation. This result shows that the correlation detector can indeed be used for asymmetric watermarking. As K increases, the correlation output slowly decreases due to the reduced resolution. It was observed that the variance of m was a little higher than 1. This is attributed to the imperfection of the interleaver.

To compare the proposed method with that of Furon's [3], the performance of each method in detecting watermark from a JPEG compressed image is evaluated. Fig. 5 shows the averaged detection value normalized to the value obtained for uncompressed image. In the Furon's method, the normalized output varies roughly between -1 and 1, while in the proposed method the normalized output varies roughly between 0 and 1. Thus, we can conclude that the performances of two methods are similar. Robustness tests against other attacks, such as noise addition, enhancement filtering and malicious attacks, are required for more accurate comparison.

To investigate the possibility of embedding and detecting multiple bits, we calculated the bit error rate (BER) from the experiments using 20 different interleavers and PN sequences for a given false alarm probability. Fig. 6 shows the BER $(1 - P_D)$ versus the embedded bits L. BER, obtained from the experiment, was a little higher than that obtained from the theory. This is attributed to possible loss of information due to the quantization (256 levels) of watermarked

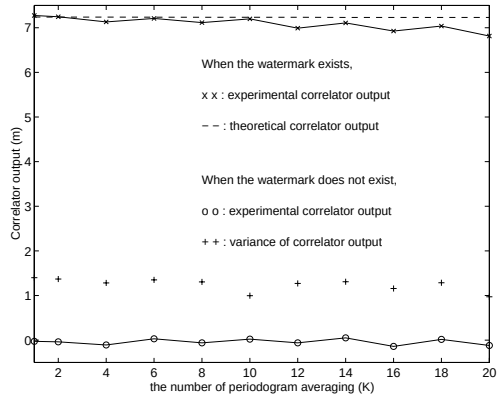

Fig. 4. Correlator output Vs. averaging factor

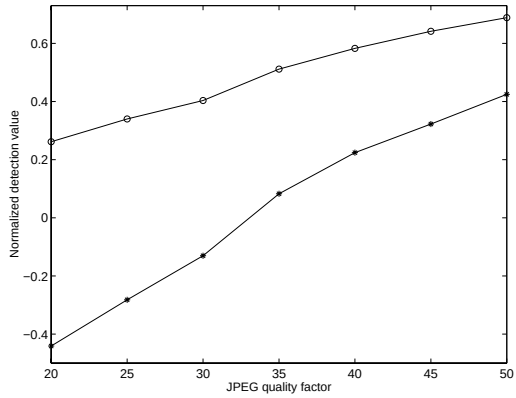

Fig. 5. Normalized detection value Vs. JPEG quality factor: ∘ proposed detector, ∗ Furon's detector [3]

image after watermark embedding and the imperfection of the interleaver. In the power spectrum estimation, periodogram averaging ($K = 8$) was used. As the embedded bits L increases, the number of DFT points allocated for each bit is decreased; thus the BER increases. Only several bits of information can be embedded without severe error rate for the data length $N = 16200$. The possible information rate is proportional to the data length N. To increase the information rate further, N must be increased.

5 Conclusion

A novel and simple detector for an asymmetric watermarking is proposed and tested. With certatin assumptions, a correlation detector is used in detecting

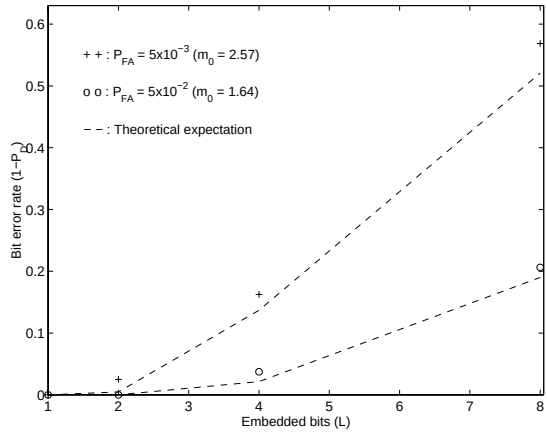

Fig. 6. Bit error rate Vs. Embedded bits

asymmetric watermark. It is based on the theory of detecting a known signal in noisy channel. The correlation detector output was well-matched to the theoretical expectation, showing that the correlation detector can indeed be used for the detection of asymmetric watermark. The proposed detector is applied to both single and multiple bit embedded watermark. Multiple bits of information are embedded by using PN sequence in filter shaping. From the experiment, several bits of information can be embedded without severe error rate. Testing the robustness of the proposed method and applying the detector to spatial domain additive watermarking remain as further works.

References

1. Teddy Furon and Pierre Duhamel, "An asymmetric public detection watermarking technique," in *Proc. of the 3rd Int. Work. on Infomation Hiding*, Dresden, Sept 1999.
2. J. A. Bloom, I. J. Cox, T. Kalker, J.-P. G. Linnartz, M. L. Miller, and C. B. S. Traw, "Copy protection for DVD video," *Proceedings of IEEE*, vol. 87, no. 7, pp. 1267–1276, July 1999.
3. Teddy Furon and Pierre Duhamel, "Robustness of asymmetric watermarking technique," in *Proc. IEEE Int. Conf. Image Processing*, Vancouver, Canada, Sept 2000.
4. A. De Rosa, M. Barni, V. Cappellini, and A. Piva, "Optimum decoding of non-additive full frame DFT watermarks," in *Proc. of the 3rd Int. Work. on Infomation Hiding*, Dresden, Sept 1999.
5. Boaz Porat, *Digital Processing of Random Signals: Theory & Methods*, Prentice Hall, 1994.
6. Louis L. Scharf, *Statistical Signal Processing: Detection, Estimation, and Time Series Analysis*, Addison-Wesley, 1991.