The 21st Electrical Engineering Conference Nov. 12-13 King Mongkut's Univ. of Technology, Thonburi

A New Criteria for Quality Estimation of Watermarking Technique

Masahiro Iwahashi*, Jarree Chaicharn** and Pongsak Mahachoklertwattana**

* Faculty of Engineering, Nagaoka University of Technology, (JICA expert to JICA Thammasat University Project) Nagaoka, Niigata, 940-2188, JAPAN, tel:+81-258-47-9546

** Department of Electrical Engineering, Faculty of Engineering, Thammasat University, Khlong-Luang, Pathum-Thani 12121, THAILAND, tel:02-516-4551~3

Abstract

Digital watermarking has been proposed as a solution to copyright protection and authentication of multimedia data in network environment. In this paper, we present a new criteria for quality estimation of watermarking technique which embeds a signature into a digital signal in frequency domain and extracts the signature without the original image. Experimental results verify our theoretical analysis and the criteria, which can be used for optimum design of more advanced watermarking methods.

1. Introduction

Security of digital multimedia contents, such as copyright protection or authentication, is an important issue in digital communications. Recently, watermarking has been attracting many researchers' attention as a solution to this problem, since it makes possible to identify the author, owner, or authorized consumer of the data by means of permanently embedding a signature into an image, video, or audio data [1-7].

Existing watermarking methods are categorized into two groups; the time domain (space domain) method [4] and the frequency domain method [5-7]. The latter method uses a frequency decomposition technique, e.g. subband, wavelet, discrete cosine transform (DCT), etc., and it is reported to have better performance than the time domain method [5-7]. However, up to now, it is concluded as a result of numerical experiments, without any theoretical verification.

This paper focuses on the watermarking based on the frequency decomposition, in which a signature sequence is embedded into an image, and can be extracted without using the original image. We address a theoretical analysis on ; (a) *quality* of embedded signal and (b) *error rate* of reading the signature. We also make clear how to estimate watermarking parameters under a given quality specification in noisy environment. The analysis in this report can be used for optimum design of more advanced watermarking methods.

2. Basic Procedure of Watermarking

2.1 Watermarking

As stated previously, watermarking is a technique which secretly embeds a signature or ID number into digital data [1-3]. Difference between original and embedded should be perceptually invisible, however, only one who knows a key can decode the signature by means of digital signal processing on the embedded signal. In addition to that, reading the signature should be robust against damages, such as noise, lossy compression, etc., on the embedded signal. This new technique is becoming the key to secure digital communications.

2.2 Embedding and reading the signature

In this paper, we deal with the watermarking procedure described in [10]. It embeds an N-bit signature $s(n)$ ($n=0,1,...,N-1$) with a parameter I_s (embedding intensity) to attain robustness. Letting $x_o(n)$ and $x_e(n)$ be original signal and embedded signal that contain the signature information, the signature embedding rule is described as follows:

1. If $s(n)$ is zero, and $int[x_o(n)/I_s]$ is an odd number, then assign $(int[x_o(n)/I_s]+1)\cdot I_s$ to $x_e(n)$.
2. If $s(n)$ is one, and $int[x_o(n)/I_s]$ is an even number, then assign $(int[x_o(n)/I_s]+1)\cdot I_s$ to $x_e(n)$.
3. Otherwise, assign $int[x_o(n)/I_s]\cdot I_s$ to $x_e(n)$.

where $int[\cdot]$ means integer part of the value. From the rule, we can infer that the larger the I_s is, the more robust the watermarking becomes, but the lower embedded signal's quality is. Therefore we should choose the smallest I_s so that the watermarking have sufficient robustness against unexpected damages.

Figure 1 illustrates the relationship between $x_o(n)$ and $x_e(n)$. We can see that $x_e(n)$ becomes a multiple of I_s. Namely,

$$x_e(n) = L(n)I_s \tag{1}$$

where $L(n)$ is an integer. To judge whether embedding signature bit is zero or one, the formula:

$$s(n) = L(n) \textit{ modulo } 2 \tag{2}$$

is used to decode the signature.

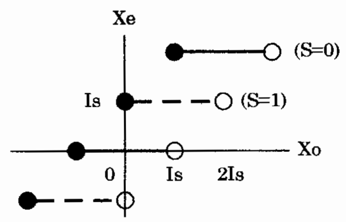


Fig.1 $x_o(n)$ is quantized according to $s(n)$.

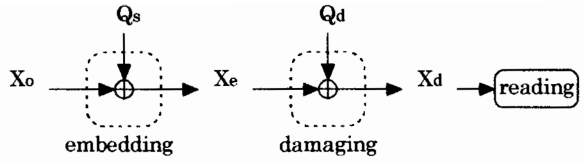


Fig.2 Watermarking model.

2.3 Damaging model

In this paper, we study on a simple damaging model in figure 2 where

$x_e(n) = x_o(n) + q_s(n)$ (3)

$x_d(n) = x_e(n) + q_d(n)$ (4)

to clarify our discussion. We study on the case of adding uniformly distributed $q_d(n)$ between $\pm I_d$ (damaging intensity = I_d) whose probability density function (PDF) is given by

$$P_{qd}(x) = \begin{cases} 1/(2I_d) & , \quad |x| \le I_d \\ 0 & , \quad |x| > I_d \end{cases} \qquad (5)$$

as damage on $x_e(n)$.

Since $x_e(n)$ will be changed due to $q_d(n)$, we should use rounding R[·] on $x_d(n)$ before reading the signature,

$x_d'(n) = R[x_d(n) / I_s]I_s$ (6)

Reading is successfully done for $I_d<I_s/2$ as we can see from figure 1. Otherwise reading error occurs.

3. Theoretical Analysis

3.1 Quality of the signal

To estimate quality of $x_d(n)$, we use mean square error (MSE) defined by the following formula:

$\varepsilon = M[x_d(n)-x_o(n)]$ (7)

where

$$M[x(n)] = \frac{1}{N}\sum_{n=0}^{N-1} x^2(n) \qquad (8)$$

In case of figure 2, equation (7) becomes

$$\varepsilon = \int_{\infty}^{\infty} P_{qs}(x)\,x^2 dx + \int_{\infty}^{\infty} P_{qd}(x)\,x^2 dx \qquad (9)$$

from equations (3) and (4), where $q_s(n)$ has no correlation with $q_d(n)$.

3.2 Reading performance

As described in **2.3**, there is no error in reading the signature when $I_d< I_s/2$. However error occurs in case of $I_s/2< I_d< 3I_s/2$. Notice that there is no error again for $3I_s/2< I_d< 5I_s/2$. Namely error appears periodically as indicated in figure 3(a) for uniformly distributed $q_d(n)$. In the figure, the shadowed areas represent the regions of error occurrence.

This can be generally expanded to arbitrary damaging $q_d^*(n)$ whose PDF is $P_{qd}(x)$ as in figure 3(b). Then, we define error rate of reading the signature by

$$\eta = \int_{\infty}^{\infty} P_{qd}(x) R(x) x dx \qquad (10)$$

where

$$R(x) = \begin{cases} 1 & , \quad \frac{(4i+1)I_s}{2} < x < \frac{(4i+3)I_s}{2} \\ 0 & , \quad \textit{else above} \end{cases} \qquad (11)$$

and (i=0,1,2,···) to estimate reading performance of the watermarking in noisy environment.

4. Watermarking in Frequency Domain

4.1 Frequency decomposition

The frequency decomposition method such as wavelet and DCT has been utilized for the watermarking as in figure 4 [6,7]. Analysis filters $H_b[\cdot]$ (b=0,1,···, B-1) followed by down sampler ↓B[·] decomposes $x_o(n)$ into B kinds of band signals $y_{oi}(m)$, (i=0,1,···, B-1, m=0,1,···, N/B-1)

$y_{oi}(m) = \downarrow B[H_i[x_o(n)]]$ (12)

We assume that an N/B bit signature s(m) is embedded into only k-th band signal $y_{ok}(m)$, while other bands have no embedding, namely,

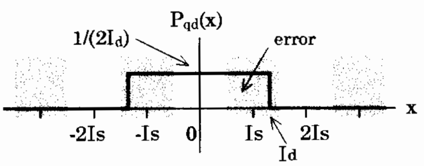
$y_{ek}(m) = y_{ok}(m)+q_s(m)$ (13)

$y_{ej}(m) = y_{oj}(m)$ (14)

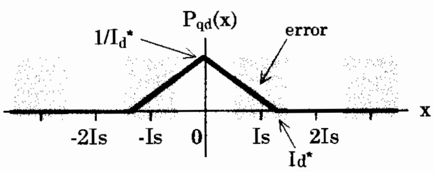
(j=0,1,···,k-1, k+1,···,B-1)

Up sampler ↑B[·] and synthesis filters $G_b[\cdot]$ (b=0,1,···, B-1) are used to get the embedded signal

$$x_e(n) = \sum_{i=0}^{B-1} G_i\left[\uparrow B\left[y_{ei}(m)\right]\right] \qquad (15)$$



(a) PDF of $q_d(n)$. [original]



(b) PDF of $q_d^*(m)$. [filtered by $(1+z^{-1})/2$]

Fig.3 Probability density function of the damage.

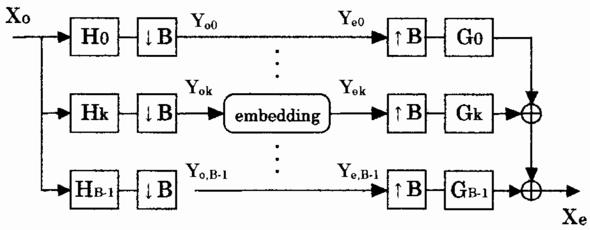


Fig.4 Embedding in frequency domain.

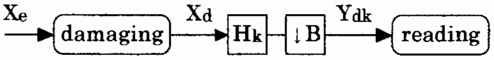


Fig.5 Damaging and reading model.

4.2 Reading and damaging

In this paper, we study on case that embedded signal is damaged in time (space) domain and the signature is decoded in frequency domain as illustrated in figure 5.

$$x_d(n) = x_e(n)+q_d(n) \quad (16)$$

$$y_{dk}(m) = \downarrow B[H_k[x_d(n)]] \quad (17)$$

The filters H_b and G_b are basis functions of DCT or another orthogonal transform [8], or analysis and synthesis filters of wavelet or subband filter bank [9]. Figure 3 satisfies the perfect reconstruction (PR) condition [9], namely

$$x_e(n) = x_o(n) \quad (18)$$

in case of no embedding.

4.3 Quality of the signal

Because of the PR condition, difference between $x_d(n)$ and $x_o(n)$ becomes

$$x_d(n)-x_o(n) = G_k[\uparrow B[q_s(m)]]+q_d(n) \quad (19)$$

Substituting equation (19) into equation (7), we get

$$\varepsilon = M[G_k[\uparrow B[q_s(m)]]+q_d(n)] \quad (20)$$

Because of no correlation between $q_s(m)$ and $q_d(n)$ and Parseval's theorem [9], equation (20) becomes

$$\varepsilon = |G_k|^2 \cdot M[q_s(m)]/B+M[q_d(n)] = \frac{|G_k|^2}{B}\int_{-\infty}^{\infty} P_{qs}(x)\,x^2\,dx + \int_{-\infty}^{\infty} P_{qd}(x)\,x^2\,dx \quad (21)$$

where

$$|G_k|^2 = \sum_i g_k^2(i) \quad (22)$$

and $g_k(i)$ are filter coefficients of G_k. The signature is embedded into only k-th band of B bands. Therefore the first term is divided by B.

4.4 Reading performance

From equations (13)-(17) and the PR condition, embedded and damaged signal becomes

$$y_{dk}(m) = y_{ek}(m)+q_d^*(m) \quad (23)$$

where

$$q_d^*(m) = \downarrow B[H_k[q_d(n)]] \quad (24)$$

Reading error rate η of the frequency method is calculated by substituting $P_{qd}(x)$ of $q_d^*(m)$ into equation (10). Total performance is evaluated with both of η and ε as described in 5..

5. Simulation Results

5.1 Time domain method

Time domain method is equivalent to figure 3 where

$$B=1, \quad H_0(z)=G_0(z)=1 \quad (25)$$

In this case, values of $q_s(n)$ and $q_d(n)$ are distributed uniformly. Therefore equation (9) and (10) become

$$\varepsilon = \frac{1}{3}\left(I_s^2 + I_d^2\right) \quad (26)$$

and

$$\eta = 0 \quad \text{for} \quad 0<\xi^{-1}<1/2$$

$$\eta = \begin{cases} k\xi/2 & , \quad k=2,4,6,\cdots \\ 1-k\xi/2 & , \quad k=1,3,5,\cdots \end{cases} \quad (27)$$

where

$$(2k-1)/2 < \xi^{-1} < (2k+1)/2, \quad \xi = I_s/I_d$$

respectively.

5.2 Frequency domain method

We deal with a basic two channel (B=2) wavelet method given by

$$\begin{pmatrix} H_0(z) & G_0(z) \\ H_1(z) & G_1(z) \end{pmatrix} = \begin{pmatrix} p(1+z^{-1}) & q(1+z) \\ p(1-z^{-1}) & q(1-z) \end{pmatrix} \quad (28)$$

where

$$pq=1/B \quad (29)$$

as an example. In this case, using (α, β),

$$(\alpha, \beta)=(2q^2, 2p) \quad (30)$$

equation (9) becomes

$$\varepsilon = \frac{1}{3}\left(\frac{\alpha}{B} I_s^2 + I_d^2\right) \quad (31)$$

and PDF of $q_d^*(m)$ is given by

$$P_{qd^*}(x) = \begin{cases} \frac{1}{I_d^*}\left(1-\frac{|x|}{I_d^*}\right) & , \quad 0<|x|<I_d^* \\ 0 & , \quad else \ above \end{cases} \quad (32)$$

where $I_d^* = \beta I_d$ as in figure 3(b). Therefore equation (10) becomes

$$\eta = 0 \quad \text{for} \quad 0<\xi^{-1}<1/2$$

$$\eta = \begin{cases} 2\xi \sum_{j=1}^{k/2} Q(4j-2) & , \quad k=2,4,6,\cdots \\ 2\xi \sum_{j=1}^{(k-1)/2} Q(4j-2) + Q^2(2k-1) & , \quad k=1,3,5,\cdots \end{cases} \quad (33)$$

where

$$(2k-1)/2 < \xi^{-1} < (2k+1)/2$$

$$\xi = I_s/(\beta I_d), \quad Q(i)=1-\xi i/2$$

5.3 Watermarking gain

From equations (27) and (33), η becomes zero for $2I_d^*<I_s$ where $I_d^*=\beta I_d$ ($\beta=1$ for time domain). However, when $2I_d^*/3<I_s<2I_d^*$,

$$\begin{cases} \eta = 1 - \dfrac{I_s}{2 I_d} & (time\ /\ space) \\ \eta = \left(1 - \dfrac{I_s}{2 I_d^*}\right)^2 & (frequency) \end{cases} \tag{34}$$

Figure 6 indicates that the result of the experiment gave good agreement with the value of our theoretical estimation. In this experiment, we used 2048 point auto regressive (AR) model as x(n) and random binary data as s(n). Then, we can determine embedding intensity I_s for a given watermarking specification (I_d, η), namely,

$$\begin{cases} I_{s\_t} = 2 I_d (1 - \eta) \\ I_{s\_f} = 2\beta I_d (1 - \sqrt{\eta}) \end{cases} \tag{35}$$

where _t and _f indicate time domain and frequency domain respectively. From (26), (31), (35), quality of embedded and damaged signal for above embedding intensity is estimated by

$$\begin{cases} \varepsilon_{\_t} = \dfrac{I_d^2}{3}\left\{1 + 4(1-\eta)^2\right\} \\ \varepsilon_{\_f} = \dfrac{I_d^2}{3}\left\{1 + 4\left(1 - \dfrac{4\alpha\beta^2}{B}\sqrt{\eta}\right)^2\right\} \end{cases} \tag{36}$$

In this report, we define *"watermarking gain"*

$$G_{WT} = 10\log_{10}\frac{\varepsilon_{\_t}}{\varepsilon_{\_f}} \quad [dB] \tag{37}$$

as a criteria for optimum design of a new watermarking method. Substituting equations (29), (30) and (36) into (37), G_{WT} becomes

$$G_{WT} = 10\log_{10}\frac{1 + 4(1-\eta)^2}{1 + 4(1-\sqrt{\eta})^2} \quad [dB] \tag{38}$$

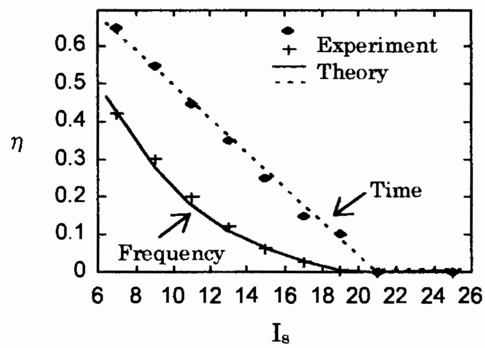


Fig.6 Comparison of the two methods.
I_d=10, (p,q)=(0.5,1.0)

We can see, from equation (38), that: 1) the frequency domain method is advantageous over time domain method by 0.65[dB] and 1.31[dB] when reading error rate is 5% and 1% respectively. 2) The performance is independent of input signal and filter coefficients (p,q).

6. Conclusion Remarks

In this report, we have analyzed signal quality and reading error rate of the watermarking technique. As a result of our analysis, it became possible to theoretically compare performance of the two methods, the time domain method and the frequency domain method, and to determine embedding intensity for a given specification of reading error rate. We have also defined the *"watermarking gain"* as a new criteria for optimum design of some new watermarking methods.

References

[1] I.J.Cox, J.Kilian, F.T.Leighton, T.Shamoon, "Secure Spread Spectrum Watermarking for Multimedia," IEEE Trans. Image Processing, vol.6, no.12, pp.1673-1687, Dec.1997.

[2] E.Koch, J.Rindfrey, J.Zhao, "Copyright Protection for Multimedia Data," Proc. International Conf. Digital Media and Electronic Publishing, pp.6-8, Dec. 1994.

[3] R.Schyndel, A.Trikel, C.Osborne, "A Digital Watermarking," Proc. IEEE International Conf. Image Processing, 1994.

[4] K.Tanaka, Y.Nakamura, K.Matsui ,"Embedding Secret Information into a dithered Multi-level Images," Proc. IEEE Military Communications Conf., pp.216-220, 1990.

[5] X.C.Xia, C.G.Boncelet and G.R.Arce, "A Multiresolution Watermark for Digital Images," Proc. IEEE International Conf. Image Processing, no.436, pp.548-551, Oct. 1997.

[6] B.Tao, B.Dickinson ,"Adaptive Watermarking in the DCT Domain," Proc. IEEE International Conf. Acoustic, Speech, Signal Processing, pp.2985-2988, 1997.

[7] A.Piva, M.Barni, F.Bartolini, V.Cappellini, "DCT-based Watermark Recovering without Resoring to the Uncorrupted Original Image," Proc. IEEE International Conf. Image Processing, no.1117, pp.520-523, Oct. 1997.

[8] N.S.Jayant, P.Noll, "*Digital coding of waveforms,*" Englewood Cliffs, Prentice-Hall, 1984.

[9] A.N.Akansu, R.A.Haddad, "*Multiresolution signal decomposition,*" New York: Academic, 1992.

[10] J.Onishi, K.Matsui, "A Method of Watermarking with Multiresolution Analysis and PN Sequence," Trans. IEICE DII, vol.J80-DII, no.11, pp.3020-3028, Nov.1997.