적응 임계값을 사용한 워터마킹 시스템
(A Watermarking System using Adaptive Thresholds)

오 상 헌 † 박 성 욱 †† 김 병 준 †††
(Sang-Heun Oh) (Sung-Wook Park) (Byung-Jun Kim)

요 약 본 논문에서는 이산 웨이블릿 변환에 기반한 워터마크 시스템을 제안 하였다. 제안한 시스템의 특징으로는 워터마크 삽입 시 화질과 워터마크의 정보량과의 트레이드-오프, 그리고 화질과 워터마크의 강인성의 트레이드-오프를 적절하게 조절 할 수 있다. 또한, 워터마크를 추출 시 워터마크 삽입 영상의 공격에 따라 임계값이 재설정되고 훼손된 정도를 측정하여 추출된 비트 중에서 가장 영향을 받지 않은 비트를 선택한다. 마지막으로 워터마크 검출을 위한 새로운 응답 방법을 제안하였다. 실험결과 제안된 워터마킹 시스템은 여러 종류의 신호처리와 의도적인 공격에도 높은 강인성을 보였다.

키워드 : 워터마크, 임계값 재설정, 훼손 측정

Abstract In this paper, a discrete wavelet transform (DWT)-based watermarking system is proposed. The main feature of proposed system is that the embedding system uses adaptive thresholds to control the trade-off between the quality of the watermarked image and the capacity of the watermark, and the trade-off between the quality and robustness of the watermarked image. Also, the extracting system rebuilds threshold according to various attacks and decides a watermark bit from the least distorted coefficient after measuring the distortion of coefficient. Finally, a new measure to detect the uniqueness of watermark is proposed. The experimental result shows that the proposed watermarking system is robust against conventional signal processing and intentional attacks.

Key words : watermark, rebuild threshold, measure distortion

1. Introduction

With the recent growth of networked multimedia systems, various techniques are needed to prevent illegal copying, forgery and distribution of digital audio, images and video. One way to improve one's claim of ownership over an image, for instance, is to place a low-level signal directly into the image data. This signal, known as a digital watermark, uniquely identifies the owner and can be easily extracted from the image. It is also desirable to determine where and how much the multimedia file has been changed from the original.

† 정 회 원 : 삼성전자 DM연구소 연구원
sam8080@samsung.co.kr
†† 비 회 원 : 삼성전자 DM연구소 연구원
sungwook park@samsung.com
††† 정 회 원 : 삼성전자 IT Center 연구원
bjkim@samsung.co.kr
논문접수 : 2002년 1월 28일
심사완료 : 2002년 11월 4일

Current techniques described in the literature for the watermarking of images can be grouped into two classes: spatial domain techniques [1][2] which embed the data by directly modifying the pixel values of the original image and frequency domain methods which embed the data by modulating the frequency domain coefficients. Frequency domain techniques can be further divided by the transforms they are using. Discrete Cosine Transform (DCT)-based embedding techniques [3][4] and wavelet-based embedding techniques [5][6][7][8][9][10] are mostly researched.

Some algorithms [7][8] for watermarking a wavelet transformed image focus on the optimized embedding strength. In [7], weighing factor is decided by pixel-wise masking. A mask measures the orientation of sub-band, level of sub-band, local brightness, and texture activity in the neighbors of the pixel. In spite of optimized scaling factor, [7] [8] are not robust

against low-pass filtering and compression because it does not embed watermark in the coarsest sub band. In [9], watermarks are embedded by mapping coefficient to one of two fixed thresholds according to watermark bit, but [9] is not robust to conventional signal processing since fixed thresholds can not extract accurate watermark from distorted coefficients. [10] embeds a watermark by modifying a coefficient to a value that is adaptive to maximum and minimum value in the non-overlapped running window. In [10], the trade off between quality of image and the robustness of watermark is tunable but does not have a parameter to control the trade off between the capacity of watermark and the image quality.

We propose a wavelet based watermarking method that is robust to conventional signal distortions. The advantages of our wavelet-based technique lie in the management of trade off and robustness. One is between robustness and quality of watermarked image and another is between quality and capacity of watermark during watermark embedding. The other is adjusting threshold for watermark detection according to attack. In addition, a new measure to detect the response from the watermark and smart selection system from repeatedly extracted watermark bits are proposed.

2. Proposed Watermarking System

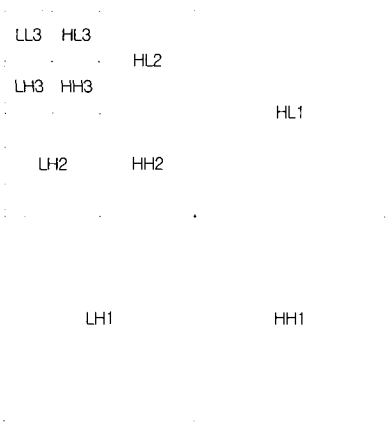

Figure 1 DWT pyramid decomposition of an image

In this section, we propose a watermark system using a discrete wavelet transform (DWT). The embedding method is described in section 2-1, and the method for extracting embedded watermarks is described in section 2-2. The method for deciding one bit among three bits is explained in section 2-3, and the method for detecting watermark only without extracting is also described in section 2-4.

2.1 Embedding

We describe the watermark embedding method using the classified coefficients after wavelet decomposition as shown in Fig.1. The embedding system inserts watermarks by modifying the classified wavelet coefficients in the sub bands, LH3, HL3 and HH3 including the lowest frequency sub-band LL3. The details are described below.

First, the two thresholds T1 and T2 are defined by α, $\Delta\alpha$ and Cmax. Cmax is the absolute maximum value of the sub band, α is the ratio of a threshold to Cmax, and $\Delta\alpha$ is the distance between T2/Cmax and T1/Cmax. Second, coefficients are classified with two thresholds. Finally, classified coefficients are mapped to either of the thresholds according to watermark bit. A watermark is embedded three times for the case of distortion by attacks.

Setting thresholds (j LH, HH, HL, LL) (sub-band level, l 1,2,3)

T2(j, l) αCmax(j, l)

T1(j, l)=(α $\Delta\alpha$)×Cmax(j, l)

Classifying Coefficients

Find the coefficients Ci (i 0,1,..,M) which satisfys T1(j, l)<|Ci|<T2(j, l).

Watermark embedding by modifying classified Ci

Set T (T2(j, l) T1(j, l))/β, β >1 and

If W(k) 1 and |Ci|>T1(j, l)+ΔT, then Ci sign (Ci)×T2(j)

If W(k) 0 and |Ci|<T2(j, l) ΔT, then Ci sign(Ci) ×T1(j)

Fig.2 shows the concept of embedding. The distance of the two thresholds can control the trade off between the quality of the image and the

robustness of the watermark. If the distance of two thresholds, T1 and T2, is getting larger, the quality of the image is being degraded while robustness to attacks increases. $\Delta\alpha$ defines the distance of two thresholds. Also ΔT controls the trade-off between the quality of the watermarked image and the capacity for embedding watermark bits. Mapping coefficients to thresholds degrades the quality of the image. The worst case is that a coefficient which is close to one threshold, is changed to the other threshold. ΔT can relieve this situation. As described in this paper, the size of ΔT is inversely proportional to β. This means that the number of coefficients for embedding is getting less as β is getting less, but the quality of the watermarked image is getting better. The proposed embedding system saves $\Delta\alpha$, the embedded position and watermark for detection.

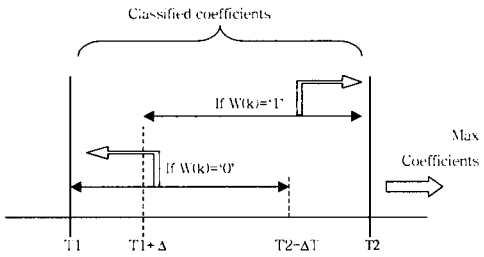

Figure 2 The concept of classification and embedding

2.2 Extracting

The watermark extractor recovers the embedded watermarks using three parameters $\Delta\alpha$, corresponding embedded position and the original watermark W(k). If common image processing or intentional attacks distort watermarked coefficients, it will be hardly possible to extract correct watermarks with the same thresholds that has been used for embedding. Therefore, the extractor needs to rebuild T2(j,l) and T1(j,l) dynamically against various attacks. Thresholds are rebuilt without α and Cmax of the original image. T1 is rebuilt by averaging coefficients in which the watermark "0"s are embedded and T2 by averaging coefficients in which the watermark "1"s are embedded. When all embedded watermarks are "0" or "1", one of the two thresholds can not be estimated. In those cases, the missed thresholds can be estimated by using $\Delta\alpha$. T2 is estimated with $(\Delta\alpha \times Cmax)-T1$ and T1 is estimated with $T2-(\Delta\alpha \times Cmax)$, respectively. Cmax is measured from the watermarked image.

Read watermarked coefficients

Using the embedded position, read corresponding wavelet coefficients, $\hat{C}_i$ (i=1,2..M)

Rebuild Thresholds

Read W(k) and set i=k, $0 \le x < M$, $0 \le y \le M$

If W(i)="1", then assign $|\hat{C}i|$ to set Cx,

If W(i)="0", then assign $|\hat{C}i|$ to set Cy.

T2=E[Cx]

T1=E[Cy]

Extract watermark

If $|Ci| < (T1(\alpha,j)+T2(j,l))/2$, then $\hat{W}(i)$="0".

If $|Ci| > (T1(\alpha,j)+T2(j,l))/2$, then $\hat{W}(i)$="1".

(j=LH, HH, HL, LL) (sub band level, l=1,2,3)

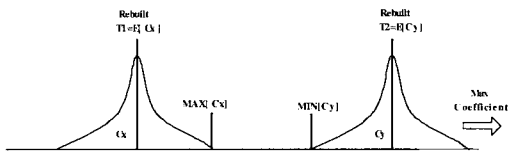

Figure 3 The concept of proposed extracting and decision method

2.3 Deciding watermark bits

As watermark is repeatedly embedded three times, the detecting system needs to select one bit among the three sets of extracted watermark bits after extracting all bits of the watermark. When a set of watermark bits among the three sets of watermark bits is selected by major voting [9], two distorted bits can neglect one correct bit. To solve this problem the proposed system analyzes coefficients to give each coefficient the degree of distortion and select a bit from the least distorted coefficient. As watermarked coefficients are more distorted, coefficients are more dispersed from the initial threshold. Therefore the degree of distortion

can be measured by the variance of coefficients where a watermark is embedded with the same threshold. In the proposed system, we estimate the degree of variation using the distance from the minimum value of *Cy* to the maximum value of *Cy* since a large variance of *Cy* and *Cy* makes a small value of distance. *Cx* and *Cy* are sets of coefficients which are supposed to be embedded with the same thresholds. Fig. 3 shows the concept of extracting and decision.

Measure distortion

Dist(i, iter)= MIN[C*x*(iter)]-MAX[C*y*(iter)],

where Dist(i, iter) is the distortion of coefficient C(i+Niter).

$0 \le i < N$, N is the number of watermark bit in one watermark set and, $0 \le iter \le 2$, is the number of watermark set.

Find the least distorted coefficient

index(i)=MAX[Dist(i, iter)], where MAX[] returns iter value whose Dist(i, iter) has maximun value, $0 \le iter \le 2$.

Select one bit from the coefficient C(i+Nindex(i)).

W'(i)=func{C(i+Nindex(i))}, where func is a function for extracting watermark in section 2.2.

2.4 Detecting the watermark response

For the case when the detector only need to prove the uniqueness of the watermark among other watermarks instead of extracting watermark data, we suggest measuring the difference between rebuilt T2 and rebuilt T1 as the detecting response instead of using many other measures. When the incorrect watermark is used to rebuild thresholds, the average of *Cx* and the average of *Cy* tend to be close each other. Therefore the difference of two rebuilt thresholds tend to be zero mean. On the contrary, the difference of two rebuilt thresholds with correct watermark will be larger than the difference with incorrect watermark. From this fact, we chose the difference of thresholds for our measure to detect the uniqueness of the watermark among watermarks. The advantage of the proposed measure is that the measurement is available without extracting watermark, like correlation or Cox's similarity measure [3]. As shown in equation (4), when a correct watermark is used watermark response will be close to 1.

$$D = \sum_{j\ LL3, LH3, HL3, HH3} (T2')(j) - (T1')(j)) \Big/ \sum_{j\ LL3, LH3, HL3, HH3} ((\Delta\alpha(j) \times Cmax(j)),$$

where T's are rebuilt thresholds (4)

3. Experimental Results

In order to evaluate the proposed digital watermark method, we applied 8bits/pixel and 256256 pixel images, "Camman", "Lenna" and "Baboon"(which are shown in Fig. 4). We set $\Delta\alpha$ small in the sub-band LL since human visual system is very sensitive to low frequency. We also set $\Delta\alpha$=0.02 for LL, $\Delta\alpha$=0.2 for other area, and β is set to 3. We randomly generate 80 bits as watermark.

(a) Camman (b) Lenna (c) Baboon

Figure 4 Original Images

3.1 The trade-off of watermarking

To shows β(=(T2 T1)/T) controlling the trade-off between the capacitance and the quality of watermarked image, the same watermark is embedded to test images two times with varying β. One watermark is embedded in LL3 and the other is in other sub-bands. $\Delta\alpha$ is set to 0.02 for LL3 and 0.1 for other sub-bands. In Fig. 5 horizontal axis is and left vertical axis is available number of coefficients in LL3, HL3, LH3 and HH3. Right vertical axis shows PSNR of watermarked image. When β is decreased, the quality (circle markers) is increased and the capacitance (square markers) is decreased.

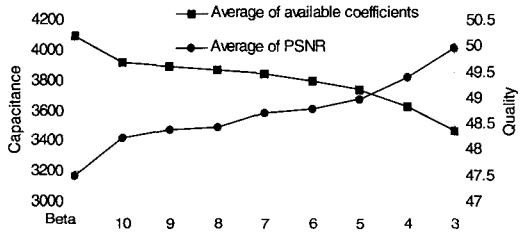

Figure 5 The Capacitance and the quality with various β

Original images are watermarked with varying β and two different patterns. One pattern is that watermark is embedded one time in the LL sub-band and two times in the other sub-bands. The other pattern is that two times in LL sub-band and one time in other sub-bands. $\Delta\alpha$ is set 0.02 for LL sub-band and 0.2 for other sub-bands. Table 1 shows relationship between the quality of the watermarked image with β and $\Delta\alpha$. Experimental results show that the quality of watermarked images increases as β decreases and more embedding in LL sub-band, which is defined by less $\Delta\alpha$ than other sub-bands, increases the quality of the watermarked image.

Table 1 Qualities of watermarked images according to β and $\Delta\alpha$

Sub-band	Embed once in LL sub-band, twice in other sub-bands			Embed twice in LL sub-band, once watermark in other sub-bands	
	Lenna (dB)	Baboon (dB)	Camman (dB)	Lenna (dB)	Baboon (dB)
b=1.5	46.07	49.86	45.26	47.44	50.71
b=3	45.66	49.7	44.4	47.86	50.25
b=5	45.47	48.8	44.11	46.5	49.27
b=?	43.7	46.86	43.94	45.06	47.12

3.2 The robustness against common image processing

We compare our watermarking algorithm with another wavelet-based technique developed by Inoue[9]. Both of algorithms are using thresholds to embed in the wavelet domain and the same thresholds are equally used for both of algorithms in this test. Also the algorithms to select one bit among three bits are compared too. One is voting method and the other is the proposed measure distortion method, M.D. Watermarks are embedded two times in LL sub-band and one times in other

Table 2 Robustness against common image processing with and without the proposed method

Common Processing	Gaussian filter	Sharpening	Median filter (2 2)	Median filter (3 3)	Scaling (128 128)	Scaling (512 512)
Voting (% of Error)	0	7.5	10	5	10	6.25
M.D (% of Error)	0	2.5	1.2	0	1.2	0
Inoue[9] (% of Error)	7.9	25.4	20.8	14.6	1.7	13.3
Quality of processed image (dB)	25.19	18.41	20.77	23.61	23.29	29.08

JPEG Compress ratio	60	50	40	30	20	10
Voting (# of Error)	0	0	0	0	1.2	2.5
M.D (# of Error)	0	0	0	0	0	0
Inoue[9] (% of Error)	2.1	2.1	2.5	2.1	3.3	18.3
Quality of decompressed image (dB)	31.26	30.43	29.57	28.55	27.25	25.33

sub-band of images.

The experimental result is shown in Table 2.

For the comparison, we test the robustness against common image processing in "Stirmark Attack"[11] with the watermarked image "Camman". We choose following attacks from stirmark attack.

- Sharpening : Edges of image are emphasized in this attack.
- Gaussian filter : An image is filtered by Gaussian filter.
- Median filter : An image is filtered by 22 and 33 median filter.
- JPEG : An image is compressed by JPEG compression.
- Image scaling (from a 256256 image to 512512 and 128128) is tested, too.

Table 2 shows that the proposed watermark method is robust against common image processing especially against JPEG compression. The proposed measure distortion method reduces more errors than voting method and show more robust than Inoue's algorithm[9].

3.3 The response of the watermark

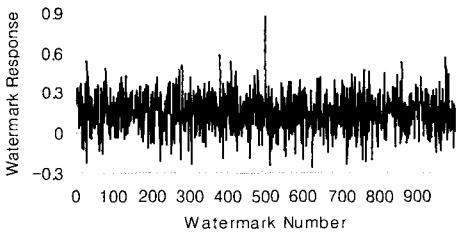

Figure 6 The response of detector to 1000 randomly generated watermarks

Watermark response detection, which is proposed in section 2.4, is tested. We first randomly generate 1000 watermarks and replace the correct watermark with the 500^{th} watermark. As shown in Fig.6, we can clearly see one unique watermark at 500^{th} watermark.

3.4 Robustness against intentional attack

To experiment robustness against intentional attacks, warping attack is done to the eye part of "Lenna" image, Fig.7 (a), because eye part includes many kinds of frequency components, noise is added to "Camman", Fig.7-(b), because most part of this image are composed of low frequency components. Watermarks are embedded two times in LL sub-band and three times in other sub-band of both images.

(a) Warped "Lenna"(23.30dB)

(b) Noised "Camman" (19.05dB)

Figure 7 Experimented intentional attack (a), (b)

Through watermark extracting, 1 error bit is occurred from warping attack but 4 error bits from noising. Although there are many error bits at noise attack, detecting is still available. When detecting response is tested with randomly generated 1000 watermarks and replace the correct watermark with the 500^{th} watermark. Though noising attack degrades image quality to 19.05dB (Fig.7 (b)). The watermark detector still indicates 500^{th} watermark is correct one,

which response the highest, as shown in Fig.8.

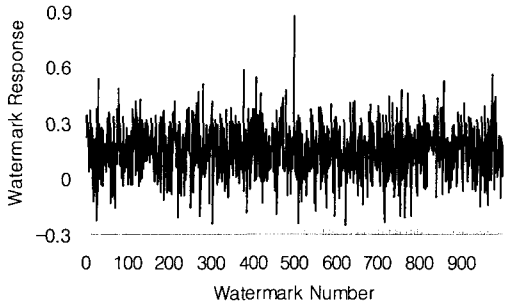

Figure 8 The detector response to 1000 different watermarks

4. Conclusion

In this paper, we have introduced a new watermarking system using adaptive thresholds based on the discrete wavelet transform (DWT). Watermark is embedded in the coefficients after the discrete wavelet transform. There are two features with embedding system. First, the trade-off between the quality and the robustness can be controlled by . Second, the trade-off between the quality and the capacitance is controlled by . These two factor set embedding thresholds according to an original image.

The watermark extracting is performed without the original image. The feature of extracting system is that thresholds are dynamically adjusted for distorted coefficients.

The other feature of the proposed system is that a bit is selected from the least distorted coefficient to overcome the problem of voting system. Finally we suggest new detecting measure for proving the uniqueness of watermark.

We have also implemented numerical examples for several kinds of attacks, such as scaling, compression with JPEG, cropping, Gaussian filtering, sharpening warping and adding noise. It is found that wavelet-based watermark approach we proposed in this paper is robust to those attacks.

References

[1] Mitchell D. Swanson, Mei Kobayashi, Ahmed H. Tewfik, "Multimedia Data_Embedding and watermarking Technologies," Proc. IEEE Trans. Image processing, Vol 86,pp1064-1087, June 1998.

[2] G. Voyatzis, I. Pitas, "Embedding Robust Watermarks by Chaotic Mixing," Proc. DSP'97, Satorini, Greece, Vol.2, pp.1121-1124, June 1997.

[3] I. J. Cox, J. Kilan, F. T. Leighton and T. Shannon, "Secure spread spectrum watermarking for multimedia," IEEE Trans. Image Processing. Vol. 6. no. 12. pp. 1673-1687, 1997.

[4] Jiwu Huang, Y.Q. Shi, Yi Shi, "Embedding image watermarks in dc components," IEEE Trans. Circuits and Systems for Video Technology, Vol. 10, pp. 604-608, Sept. 2000.

[5] Xiang-Gen Xia, Charles G. Boncelet, Gonzalo R, Arce, "Wavelet transform based watermark for digital images," Proc. OCIS'98, Vol.3, pp.497-511, November 1998.

[6] Houng-Jyh Mike Wang, Po-Chi Su, C. C. Jay Kuo, "Wavelet-based digital image watermarking," Proc. OCIS'98, Vol.3, pp.491-496, November 1998.

[7] M. Barni, F. Bartolini, A. Piva, "Improved Wavelet-Based Watermarking Through Pixel-Wise Masking," IEEE Tras. Image Processing, Vol. 10, no. 5. pp783-791, MAY 2001.

[8] D. Taskovski, S. Bogdanova, M. Bogdanov, "A low resolution content based watermarking of images in wavelet domain," Proc. ISPA'2001, Proceedings of the 2nd International Symposium, Pula, Croatia , pp. 604-608, June 2001

[9] H. Inoue, A. Miyajaki, A. Yamamoto, T. Katsura, "A Digital Watermark Technique Based on the wavelet Transform and Its Robustness on Image Compression and Transformation," Proc. IEICE. Trans. Funadamentals, Vol.E82-A, pp.2-10, January 1999.

[10] Liehua Xie, G.R. Arce, "A class of authentication digital watermarks for secure multimedia communication," IEEE Trans. Image Processing, Vol. 10, pp. 1754-1764 , Nov. 2001.

[11] http://www.cl.cam.ac.uk/~app2/watermarking/stirmark

오 상 헌

1999년 성균관대학교 전자공학과 졸업(공학사). 2001년 본대학원 전기전자컴퓨터공학부졸업(공학석사). 2001년~현재 삼성전자 DM연구소연구원. 관심분야는 워터마크, 정보보호 시스템

박 성 욱

1993년 연세대학교 전자공학과 졸업(공학사). 1995년 연세대학교 본대학원 전자공학과 졸업(공학석사). 1998년 연세대학교 본대학원 전자공학과 졸업(공학박사) 1998년~2000년 삼성전자 중앙연구소 전임연구원. 2000년~현재 삼성전자 DM연구소 책임연구원. 관심분야는 멀티미디어 신호처리, 오디오 신호처리

김 병 준

1992년 경북대학교 전자공학과 졸업(공학사). 1992~현재 (주)삼성전자 재직. 관심분야는 저작권 보호 시스템 연구, 암호 시스템, 워터마크