A Color Image Watermarking Method for Embedding Audio signal

Sang Jin Kim* and Chung Hwa Kim**
* Dept. of Computer & Information, Chosun-College of Science & Technology, Gwangju, 290, Korea
Tel : +82-62-230-8828 Fax : +82-62-230-8801 E-mail: sjkim@mail.chosun-c.ac.kr
** Dept. of Electronics, Chosun University, Gwangju, 375, Korea
Tel : +82-62-230-7068 Fax : +82-62-232-3369 E-mail: jhdkim@mail.chosun.ac.kr

Abstract: The rapid development of digital media and communication network urgently brings about the need of data certification technology to protect IPR (Intellectual property right). This paper proposed a new watermarking method for embedding contents owner's audio signal in order to protect color image IPR. Since this method evolves the existing static model and embeds audio signal of big data, it has the advantage of restoring signal transformed due to attacks. Three basic stages of watermarking include: 1)Encode analogue 1D owner's audio signal using PCM and create new 3D audio watermark; 2)Interleave 3D audio watermark by linear bit expansion and 3)Transform Y signal of color image into wavelet and embed interleaved audio watermark in the low frequency band on the transform domain. The results demonstrated that the audio signal embedding in color image proposed in this paper enhanced robustness against lossy JPEG compression, standard image compression and image cropping and rotation which remove a part of image.

1. INTRODUCTION

The rapid growth of digital media and communication ork brings about the need for IRP protection technology for digital multimedia. Digital watermarking may solve this kind of problem by directly embedding ownership data in multimedia contents.

Image watermarking technology includes the method of embedding watermark in a spatial domain and transform domain. Spatial domain method corrects LSB of image pixel in general, but is not robust to such attack as low pass filtering. For this reason, most methods prefer to correct pixel data on the transform domain using Fourier transform, DCT, or wavelet transform. However, these methods have also weak robustness to such image processing as cropping, rotation, and lossy JPEG compression. Reference [7] is color image watermarking method that embeds watermark in B signal where visual sensitivity is least in RGB mode. But it has also weak robustness to general image processing. Since watermarking methods in service at home and abroad currently provide only several bits of data, they have been used for ownership claim rather than for aggressive opposition to digital contents hacking.

This paper proposes a digital media IRP protection algorithm to embed owner's audio signal on the DWT domain of Y signal of color image. As the audio watermark used for this paper uses owner's audio of big data, ownership claim can be made in auditory. Moreover when the attack removing a part of image is made, loss data can be easily restored through applying interleaving. In addition, to make security strong, non-blind method is selected in extracting audio signal.

2. AUDIO WATERMARKING TRANSFORM AND INTERLEAVING

2.1. Payload of audio signal

Owner's audio signal data to be embedded in image contents is decided according to original image size, wavelet transform level, and marking space in the transform domain. Accordingly, the audio watermark data proposed in this paper is calculated by

$$\text{Watermark data} = \left(\frac{X_x \times X_y}{4^{l+1}}\right) \times 8\,[\text{bits}] \qquad (1)$$

where X_x and X_y indicate the size of x and y axis of original image, l indicates the number of wavelet transform level of image, and $1/4^{l+1}$ indicates marking space size equivalent to low frequency band in level l.

For example, if the size of original image is 512x512 and audio signal is embedded in the 1-level wavelet transform domain, data to be embedded is found as $(512 \times 512/4^2) \times 8 = 131{,}072$[bits].

2.2. Transform of audio signal into audio watermark

This section will explain the method for transforming owner's analogue 1D audio signal into audio watermark about image contents obtained through an audio input device. The audio used in this paper is analogue sound of Korean pronunciation: "*Bon Yeongsangmul E Daehan Soyugown Eun Kim Sang Jin Ege Itseumnida*(The ownership of this image contents is held by Kim Sang Jin in English)". The wave of audio signal is shown in Fig. 1.

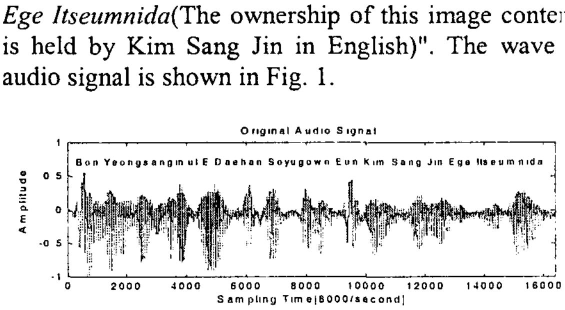
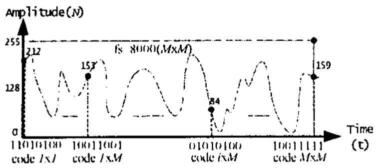
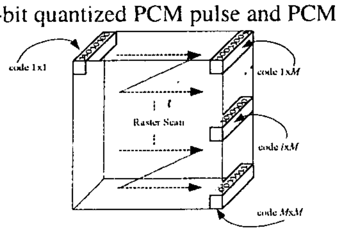


Fig. 1. Wave of analog audio signal " Bon Yeongsangmul E Daehan Soyugown Eun Kim Sang Jin Ege Itseumnida'.

To transform 1D analogue audio signal into 3D digital audio watermark, the audio signal is placed in new 3D space using PCM, as shown in Fig. 2, via such transform as sampling, quantization, and coding.



(a) 8-bit quantized PCM pulse and PCM code



(b) 3D audio watermark

Fig. 2. Transform sequence for audio watermark.

The proposed 3D audio watermark consists of the following transform stages.

Stage 1. Specify 1D analogue audio signal data which can be embedded using Equation (1).

Stage 2. Take sample signal of analogue audio signal as 8kHz sampling frequency : PAM pulse.

Stage 3. Quantize with 8-bit level per sample : PCM pulse.

Stage 4. Encode by 8 bits per the quantized sample signal : PCM code.

Stage 5. Arrange bit strings obtained through coding in new 3D space using Raster Scan and obtain audio watermark : Digital audio watermark.

For original image of 512x512 in size, when the low frequency band transformed into 1-level wavelet is set as marking space, the audio watermark according to the above procedure is shown in Fig. 2.

2.3. Interleaving of audio watermark

Interleaving is a communication term and the process of rearranging by specific mixed rules after repeating data bit strings to be sent. Although bit error occurs in data bit due to external noise or unauthorized external attack, data rearranged with specific rules is arranged to original data bit strings through the inverse process of interleaving (de-interleaving) and a repeat pattern is examined to restore loss bit.

Watermarking system also causes bit error in embedded watermark data due to external noise or illegal transform of work. If interleaving is applied to watermarking system, error correction efficiency can be enhanced even though burst error occurs by dispersing error by de-interleaving, by changing into independent error, and by correcting error.

Equation (2) is linear bit-expansion expression applied in Chae and Manjunath's[8] watermarking method.

$$P_N(x,y) = \left(2^N - 1\right)\left(\frac{P(x,y) - P_{\min}}{P_{\max} - P_{\min}}\right) \qquad (2)$$

where $P_N(x,y)$ is the result of expanding the pixel of n-bit watermark $P(x,y)$ to linear N-bit and $P_{\max}$ and $P_{\min}$ are the maximum and minimum values of pixel . If $P_{\max}$ of the pixel of audio watermark is $2n$-1 and is $P_{\min}$ zero, Equation (3) can be induced from Equation(2). The result of linear bit-expanded bit strings repeats bit string with a regular pattern. In other words, if n-bit pixel is expanded by linear $N(nK)$, n bit is repeated K times. Then R is defined as n-bit repeat factor. Using this, the proposed method applies the interleaving of audio watermark as follows.

$$\begin{aligned} P_N(x,y) &= \frac{\left(2^N - 1\right)}{\left(2^n - 1\right)} P(x,y) = \frac{\left(2^{nK} - 1\right)}{\left(2^n - 1\right)} P(x,y), \\ &= \left[1 + 2^n + 2^{2n} + 2^{3n} + \cdots + 2^{(K-1)n}\right] \cdot P(x,y) \\ &= \left[\sum_{i=1}^{K} 2^{(i-1)n}\right] \cdot P(x,y) \\ &= R \cdot P(x,y), \; for \; N = nK, \; K = 1, 2, 3 \cdots \end{aligned} \qquad (3)$$

First, to obtain four repeated bit strings(K=4), bit-expand ($P_i \mid i = 1,2,3,4$) 8-bit(n=8) audio watermark(P) per pixel by linear 32 bits(nK=32). In 32-bit expanded bit string, $P_1 = P_2 = P_3 = P_4$ is established from MSB to LSB and four 8-bit strings are relocated to new space respectively. Then the 32-bit repeat factor R becomes decimal number "65793" and 32-bit binary number "00000001...00000001". Using this principle, audio watermark(128x128) is interleaved as shown in Fig. 3.

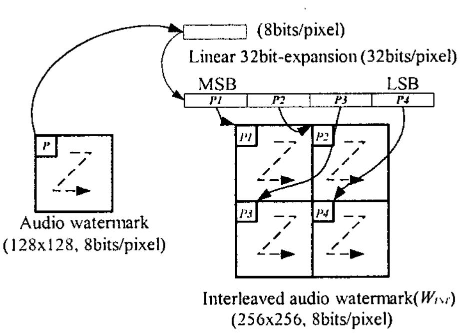


Fig. 3. Interleaving block diagram using linear 32-bit expansion.

Therefore, interleaved audio watermark becomes four times bigger than original watermark in size(256x256) and has the same size as marking space of low frequency band transformed into 1-level wavelet from original image of 512x512.

3. AUDIO WATERMARK EMBEDDING AND EXTRACTING

3.1. Marking space of color image

To apply the proposed watermarking system to gray level image as well as color image, Y signal as luminance component of color image is used for watermarking. RGB color image mode is transformed into YIQ mode as NTSC transmission standard by coordinates transform of Equation (4).

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.275 & -0.321 \\ 0.212 & -0.523 & 0.311 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \qquad (4)$$

where Y represents the luminance of color and I and Q represent the hue and saturation attributes of color image. The marking space of this method used low frequency band on the Y signal wavelet transform domain in order to gain non-visual processing and robustness of audio watermark to be embedded.

3.2. Embedding of interleaved audio watermark

To embed audio watermark in marking space, Equation (5) is used.

$$Z_{LL}(x,y) = Y_{LLS}(x,y) + alpha \times W_{INT}(x,y) \qquad (5)$$

where $Y_{LLS}(x,y)$ is the coefficient of LL band that is low frequency band on the wavelet transform domain of original image and $W_{INT}(x,y)$ is interleaved audio watermark. Transform the result obtained by Equation (5) by inverse wavelet, gain embedded Y signal, and finally transform into RGB mode, which is the inverse process of Equation (4) using I and Q signal. And watermarked color image can be obtained.

3.3. De-interleaving for extracting audio watermark

In the method proposed in this paper, the extraction of embedded watermark requires original image. Fig. 4 shows the procedure of extracting watermark, but the order of extraction is the inverse of watermark embedding. First process for watermark extraction is to obtain low frequency band coefficient by conducting wavelet transform, in both watermark-embedded image and original image. Second is to obtain P_i from four locations relocated in the process of interleaving by Equation (6).

$$W_{INT}(x,y) = \left(\frac{Z_{LL}(x,y) - Y_{LL}(x,y)}{alpha} \right) \qquad (6)$$

If audio watermark data is transformed due to external attack, go through the following de-interleaving process in order to obtain the optimum audio watermark pixel through restoring.

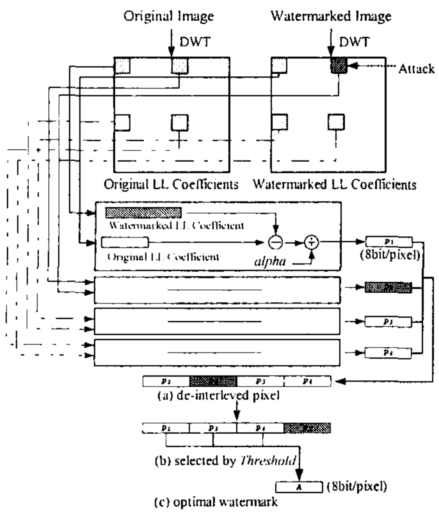


Fig. 4. Watermark extraction using de-interleaving.

First, find average deviation D_i of P_i. Average deviation refers to variance by calculating and averaging the absolute value to deviation of each measurement from arithmetic mean. By using mean m, the following is given.

$$D_i = | A_i - m | \qquad (7)$$

Then calculate standard deviation σ using Equation (7). Standard deviation refers to variance by calculating the square root to squares mean and is given by

$$\sigma = \sqrt{\frac{1}{4}\sum_{i=1}^{4} D_i^{\,2}} \qquad (8)$$

In case of normal distribution, 68.27% of standard deviation cases are included between $m-\sigma$ and $m+\sigma$, 95.45%, between $m-2\sigma$ and $m+2\sigma$, and 99.73%, between $m-3\sigma$ and $m+3\sigma$.

In the proposed method, to decide transform of audio watermark pixel data P_i due to attack, threshold value was set as 68.27% between $m-\sigma$ and $m+\sigma$. If the average deviation of four extracted pixel data D_i is less than threshold, it is considered that there is no loss due to external attack. Then the pixel value with minimum mean deviation(min{ D_i }) is selected as the optimum watermark data. On the other hand, if there is bigger deviation than threshold, it is considered that there is loss due to external attack. Then by finding mean deviation D_k again using the rest pixel value($P_k \mid k \le i$) except pixel at that time, select the pixel value of minimum average deviation(min{ D_k }) as the optimum audio watermark pixel value.

If extracted optimum audio watermark is decoded through the inverse of Fig. 1, contents owner's 1D analogue audio signal will be obtained.

4. SIMULATION RESULTS OF TEST IMAGE

This section will verify the robustness of watermarking proposed in this paper through simulation. For similarity between original image and audio watermark-embedded image, and between original audio signal and extracted audio signal, PSNR Equation (9) is used.

$$PSNR = 10\log_{10}\frac{255^2}{MSE(X,Y)} \qquad (9)$$

where X indicates original a signal and Y a comparative signal.

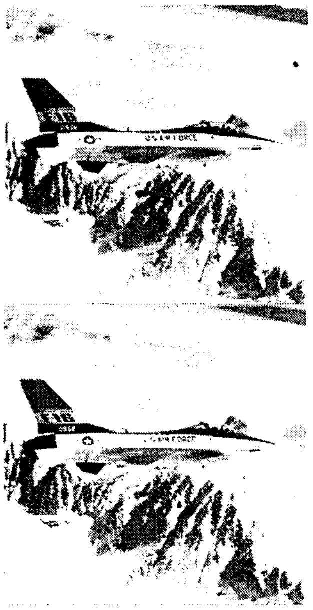
To evaluate the excellence of the proposed method, as shown in Table 1, a variety of test images are tested against lossy JPEG compression, image cropping, and image rotation.

Table 1 Audio signal extracted by proposed method, unit[dB]

Original images	Lenna	House	Airplane	Peppers
Watermarked images	32.8006	34.8865	38.4797	33.5604
JPEG qt.90%	21.7466	24.0738	23.7683	22.1230
JPEG qt.80%	18.5078	20.0995	20.1175	19.4218
JPEG qt.70%	16.6779	17.8840	18.1938	17.4035
JPEG qt.60%	15.4743	16.3439	16.5789	15.7742
Cropping 10%	37.3252	36.6316	36.9936	37.4998
Cropping 30%	31.5719	32.0883	32.4872	32.9642
Cropping 50%	29.8389	30.0284	30.0743	30.4155
Cropping 70%	28.6722	28.8247	28.6046	28.8005
Rotation 10D	28.5979	29.4994	29.6468	30.1909
Rotation 20D	21.0878	22.4389	24.6131	24.4214
Rotation 30D	18.6178	19.9192	22.0149	22.5051
Rotation 40D	18.3306	19.3358	21.5413	22.2497

Fig. 5a is "Airplane" test image of 512x512 in size.

Fig. 5b is the result of embedding interleaved audio watermark in the test image of Fig. 5a when weight *alpha* is 0.1. It suggests that 39.30dB quality is maintained and there is little visual image distortion, as compared with original image.

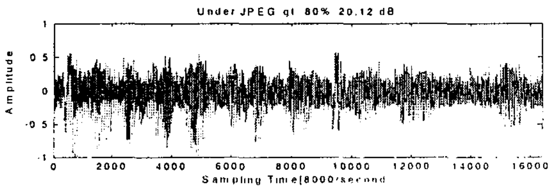


(a) "Airplane" test image (b) Watermarked image (*alpha* 0.1, *PSNR 39.30*dB).
Fig. 5. Test image and embedded image.

Fig. 6 to 8 are the simulation results of "Airplane" test image. Fig. 6a is distorted image of 32.76dB due to JPEG quality 80% attack and Fig. 6b is the audio signal of 20.12dB extracted from Fig. 6a. Fig. 7a is distorted image of 4.18dB where a part of image is removed due to cropping 70% and Fig. 7b is the audio signal of 28.60dB extracted from Fig. 7a. Fig. 8a is distorted mage of 11.60dB where a part of image is cropped and pixel value is transformed due to image rotation 20 degrees and Fig. 8b is the audio signal of 24.61dB extracted from Fig. 8a.

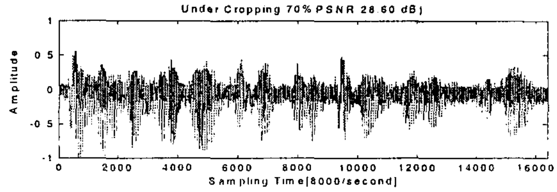
It is suggested that the audio signal extracted from various attacks maintains robustness against attack of removing a part of image.

(a) Image attacked against JPEG qt. 80% (*PSNR* 32.7603dB)



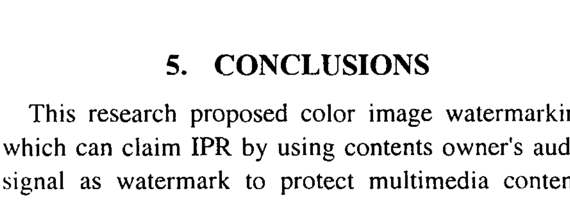
(b) Audio signal extracted from JPEG qt. 80%
Fig. 6. JPEG compression attack and extracted audio signal.

(b) Image attacked against cropping 70% (*PSNR* 4.1883dB)



(b) Audio signal extracted from cropping 70%
Fig. 7. Image cropping attack and extracted audio signal.

(c) Image attacked against rotation 20 degrees (*PSNR* 11.6015dB)



(b) Audio signal extracted from rotation 20 degrees
Fig. 8. Image rotation attack and extracted audio signal.

5. CONCLUSIONS

This research proposed color image watermarking which can claim IPR by using contents owner's audio signal as watermark to protect multimedia contents ownership. To enhance robustness, interleaving was applied to audio watermark using linear bit expansion in low frequency band on the wavelet domain. The proposed method can restore loss data by de-interleaving even though a part of audio watermark data bit string is lost due to external attack. Accordingly, this method has the advantage of applying to general gray scale image because of use of Y and of maintaining imperceptibility and robustness by embedding watermark in DWT domain. Furthermore, although watermark data is removed or there is other transform caused by image improvement, they can be restored by de-interleaver since interleaving is used.

In future, audio perception evaluation will be applied in order to claim ownership more reasonably by using audio perception algorithm in extracted audio signal.

References

[1] W. Bender, D. Gruhl, and N. Morimoto, Techniques for Data Hiding, Proc. SPIE, Storage and Retrieval for Image and Video Database III, vol. 2420, San Jose, pp.164-173, 1995.

[2] R. G. van Schyndel, A. Z. Tirkel and C. F. Osborne, A Digital Watermark, Proc. IEEE Int. Conf. Image Processing, vol. 2, Austin, pp.86-90, 1994.

[3] J. J. K. O Ruanaidh, W. J. Dowling, and F. M. Boland, Phase Watermarking of Digital Images, Proc. IEEE Int. Conf. Image Processing, vol. 3, Switzerland, pp. 239-242., 1996.

[4] A. Piva, M. Barni, F. Bartolini, and B. Cappellini, DCT-based Watermark Recovering Without Resorting to the Uncorrupted Original Image, Proc. IEEE Int. Conf. Image Processing, Santa Barbara, CA, pp. 520-523, 1997.

[5] M. Barni, F. Bartolini, "A DWT-based Technique for Spatio-frequency Masking of Digital Signatures, Proc. SPIE/IS&T Int. Conf. Security Watermarking Multimedia Contents, vol. 3657, pp.31-39, 1999.

[6] D. Kundur and D. Hatzinakos, Digital Watermarking using Multiresolution Wavelet Decomposition, Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing, vol. 5, pp. 2969-2972, 1998.

[7] M.Kutter, F Jordan, and F. Bossen, Digital Signature of Color Images Using Amplitude Modulation, in Proc. SPIE Electronic Imaging97, Storage and Retrieval for Image and Video Databases V, pp. 518-526, 1997.

[8] Chae JJ, and Manjunath BS, "A Robust Embedded Data from Wavelet Coefficients", Proceedings of the SPIE EI'98, vol. 3312, pp. 308- 317, 1998.