An Improved Detection Technique for Spread Spectrum Audio Watermarking with a Spectral Envelope Filter

Sarah Jung, Jongwon Seok, and Jinwoo Hong

ABSTRACT— We propose an improved algorithm for detecting audio watermarks based on a spread spectrum in the spectral domain. Since the energy of a watermark is much smaller than that of the cover audio data, pre-processing to reduce the effect of the cover data is needed to reliably extract watermarks. We introduce a spectral envelope filter as a pre-process that enhances detecting performance by filtering out the intrinsic spectral character of cover data. The proposed watermarking structure can be easily included in the compression system and can extract watermarks from partially decompressed spectral data. Our experimental results demonstrate that with a bit error rate of around 10 dB against general attacks, the proposed detecting scheme works better than detectors without the spectral filter.

I. INTRODUCTION

Digital watermarking technology is now drawing attention as a new method of protecting unauthorized copying of digital content [1]-[5]. A digital watermark is an imperceptible signal added to digital multimedia data (namely, audio, video, or image), which should remain even after several signal processes or potential attacks. Among various schemes, spread spectrum watermarking is expected as a promising technology due to its robustness.

Recent efforts have focused on combining watermarking with perceptual coding [6]. Since almost all distribution of multimedia content is carried out in the form of compressed streams, integrating the structure of the two processes, compression and watermarking, is desirable for achieving timing efficiency in future encoding systems. Combining is especially beneficial in spread spectrum watermarking, because the psychoacoustic process of the compression system can be shared with the watermarking system, which requires a heavy computational burden.

Until now, most studies on spread spectrum audio watermarking have dealt with detecting procedures performed in the time domain. In such systems, a whole decompression process is needed to extract a watermark from a compressed audio stream, which adds to the time delay. The time delay makes the extraction scheme inefficient, especially when only watermark detection is required without playing the data. The system proposed in [7] deals with watermark detection in the spectral domain. The spectral transform used in the system is not only complicated in itself, but the pre-process done prior to the correlation also needs a discrete cosine transform (DCT) and an inverse DCT (IDCT), which increases the amount of required calculations significantly.

In this letter, we propose an enhanced detection technique for audio watermarking using an envelope filter in the spectral domain. The filter can reduce the effect of cover data considerably with little computational increase. In the proposed system, the watermarked audio data is first filtered to extract the spectral envelope, and then the filtered data is correlated with a pseudo noise (PN) sequence in a spectral form. The proposed watermarking system makes it easy to combine the compression system and can detect a watermark from partially

Manuscript received July 26, 2002; revised Nov. 6, 2002.
This work is supported by the Ministry of Information and Communication (MIC) of Korea.
Sarah Jung (phone: +82 42 860 1558, e-mail: sjung@etri.re.kr), Jongwon Seok (e-mail: jwseok@etri.re.kr), and Jinwoo Hong (e-mail: jwhong@etri.re.kr) are with Broadcasting Content Protection Research Team, ETRI, Daejeon, Korea.

decompressed spectral data by improving detection performance in the spectral domain.

II. PROPOSED WATERMARKING SCHEME

Figure 1 diagrams the embedding scheme. The structure is based on the generalized spread spectrum watermarking method. The PN sequence is generated in the time domain, but inserted in the spectral domain. This scheme is adaptable for a compression system since the PN sequence is added to the spectral coefficients of input audio data after a psychoacoustic procedure. The spectral transform can be a fast Fourier transform, DCT, modified DCT (MDCT), or such. The output form of the watermarked audio data can be a pulse coded modulation signal or a compressed stream. For example, the spectral transform is an MDCT and the post-embedding process consists of quantization and Huffman encoding if the watermark embedding function is included in an MPEG-2 advanced audio coding (AAC) encoding system.

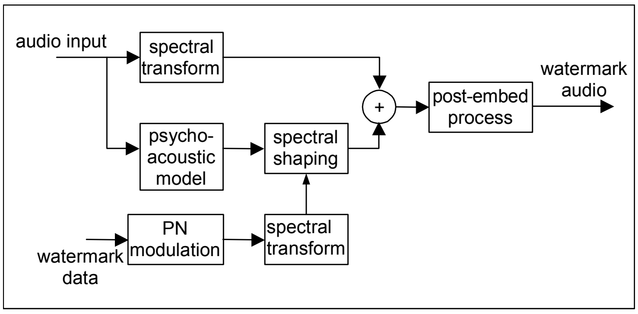

Fig. 1. Diagram of the watermark embedding scheme.

Figure 2 diagrams the proposed detecting scheme. In the detector of the spread spectrum watermark, the watermark is determined from the correlation value between the input watermarked data and the template PN sequence. Thus, the cover data plays as a noise source and degrades the detecting performance significantly due to its stronger energy. This effect is more serious in detecting audio watermarks than image watermarks because since audio watermarks have more sensitive auditory qualities, the energy of audio watermarks is smaller than that of image watermarks. In general, the original cover data is not available for the watermark detector. Therefore a pre-process which filters out the effects of cover data is required to extract an audio watermark efficiently without accessing the original data. We propose a new filter in the spectral domain which reduces the noise variance of the correlation value in the detector by removing the spectral envelope of the input watermarked audio spectrum prior to

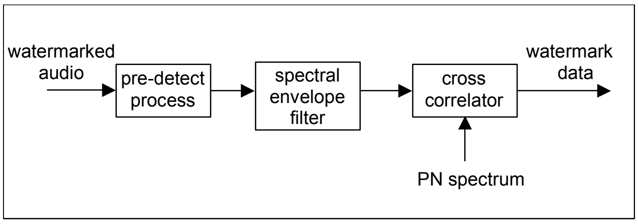

Fig. 2. Diagram of the watermark detecting scheme.

correlation. The spectral envelope of the watermarked audio data contains mostly information on spectral features of the cover audio data and has little relevance for the features of the PN spectrum [8]. Thus, the proposed spectral envelope filter (SEF) can filter out the effect of cover data. By doing this, the detection performance in the spectral domain can be improved and the noise variance reduced. The procedure of the SEF can be summarized as follows:

- Obtain a spectral envelope vector $\overline{Y}$ by low pass filtering (LPF) input watermarked audio spectrum vector $\overline{X}$ in dB magnitude; $\overline{Y} = LPF(\log(\overline{X}))$.
- Remove $\overline{Y}$ from $\log(\overline{X})$.
- Convert the filtered data into linearly scaled data; $\overline{X}_{ef} = \exp(\log \overline{X} - \overline{Y})$.

We can prevent information loss that is due to the large difference of scale in low and high band data by using dB magnitude values. The effects of the SEF are similar to *high quefrency liftering in the cepstral domain* [8]. The proposed SEF is processed in the spectral domain. Thus, it is useful to detect watermarks from compressed data which are generated mostly using a spectral transform. Figures 3 and 4 illustrate the effect of the SEF on detecting results. Figure 3 shows the

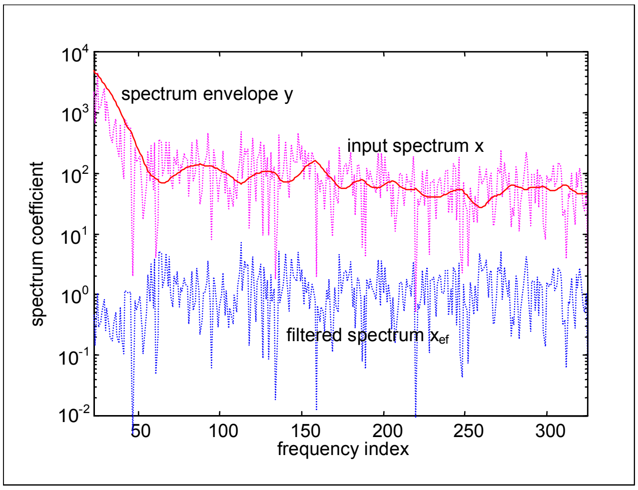

Fig. 3. Variance reduction through spectral envelope filtering (SEF); watermarked input audio spectrum *x*, spectral envelope *y*, and x_{ef} after SEF.

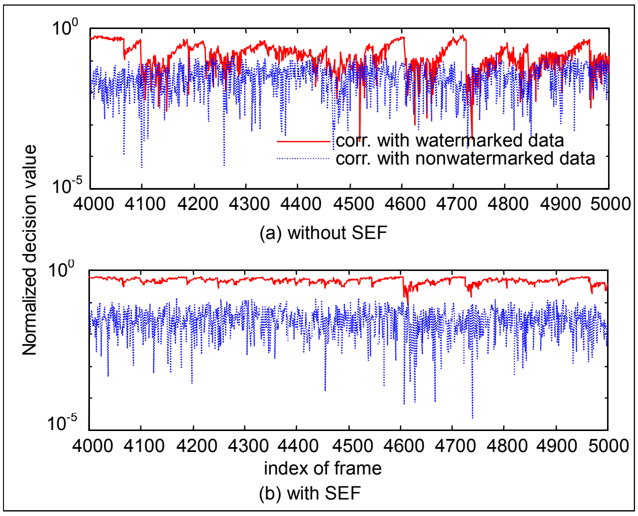

Fig. 4. Normalized correlation values for watermarked and non-watermarked data: (a) without SEF, (b) with SEF.

spectrum variance and Fig. 4 shows the normalized correlation values in the detector. As Fig. 3 shows, the spectrum variance of the data is largely reduced after the SEF is used. By doing this, the noise variance is lowered and the detecting performance is enhanced. In Fig. 4, the normalized correlation values with and without the SEF are presented. To assess the performance of the SEF, the correlation values of non-watermarked data are also plotted. Figure 4 reveals an obvious difference between correlation values in plot (b) with the SEF. In plot (a) without the SEF, the difference is not clear and even overlapped for several frames.

In the proposed detecting structure, the pre-detect process is the inverse of a post-embedding process. For example, the pre-detect process consists of Huffman decoding and inverse quantization if the detecting function is included in an MPEG-2 AAC decoding system.

III. EXPERIMENTAL RESULTS

We prepared six audio pieces for our experiments. Each audio piece had a 30 s duration and 16 bits of resolution with a sample frequency of 44.1 kHz. We embedded and detected information in the DCT domain at a rate of 86 bps. The robustness of the proposed detector was tested against several attacks: amplitude compression, band-pass filtering, echo addition, equalization, and MPEG-2 AAC compression with 64 kbps/mono. We also tested the detector without the SEF to compare it with the proposed detector. Table 1 shows the results for the test described above. We see that the bit error rate (BER) of the proposed detector is better than that of the detector without the SEF by approximately 10 dB.

Table 1. Detection results for the attacks on the BER.

Type of attack	Without SEF	With SEF
No attack	7.624E-2	5.870E-3
Amplitude compression	7.611E-2	6.321E-3
Band-pass filtering	1.302E-1	1.671E-2
Echo addition	1.091E-1	1.161E-2
Equalization	1.177E-1	6.644E-3
AAC 64 kbps	8.697E-2	9.614E-3

IV. CONCLUSION

We have presented an improved detection technique for spread spectrum watermarks based on envelope filtering in the spectral domain. The proposed system accomplished reliable detection by reducing the noise variance in the correlation without the original audio data. Since the detection is executed in the spectral domain, the technique can extract watermarks from partially decompressed spectral data. Experimental results showed that the proposed detecting scheme enhanced the BER performance by around 10 dB against common signal processing attacks, such as amplitude compression, filtering, and data compression.

REFERENCES

[1] S. Winkler, E.D. Gelasca, and T. Ebrahimi, "Perceptual Quality Assessment for Video Watermarking," *Proc. of ITCC*, 2002, pp. 90-94

[2] I.J. Cox, M.L. Miller, and J.A. Bloom, *Digital Watermarking*, Morgan Kaufmann, San Francisco, 1999.

[3] Jongwon Seok, Jinwoo Hong, and Jinwoong Kim, "A Novel Audio Watermarking Algorithm for Copyright Protection of Digital Audio," *ETRI J.*, vol. 24, no. 3, June 2002, pp. 181-189.

[4] P. Bassia, I. Pitas, and N. Nikolaidis, "Robust Audio Watermarking in the Time Domain," *IEEE Trans. on Multimedia*, vol. 3, no. 2, June 2001, pp. 232-241.

[5] A. Shiozaki, J. Tanimoto, M. Iwata, "A Digital Image Watermarking Scheme Withstanding Malicious Attacks," *IEICE Trans. Fundamentals*, vol. E-83-A, no. 10, Oct. 2000, pp. 2015-2022.

[6] F. Siebenhaar, C. Neubauer, and J. Herre, "Combined Compression/Watermarking for Audio Signals," *Proc. AES* 110th, 2001, preprint 5344.

[7] D. Kirovski and H. Malvar, "Robust Spread Spectrum Audio Watermarking," *Proc. ICASSP*, 2001, pp. 1345-1348.

[8] A.V. Oppenheim, *Applications of Digital Signal Processing*, Prentice Hall, New Jersey, 1978.