A Wavelet-Based Video Watermarking Approach Robust to Re-encoding

Kil-Sang Yoo* ***Associate Member***, Won-Hyung Lee* ***Regular Member***

ABSTRACT

We present in this paper a method of digital watermarking for video data based on the discrete wavelet transform. In the proposed method, a watermark signal is inserted into the decompressed bitstream while detection is performed using the uncompressed video. This method allows detection if video has been manipulated or its format changed. We embed the watermark in the lowest frequency components of each frame in the un-coded video by using wavelet transform. The watermark can be extracted directly from the decoded video without access to the original video. Experimental results show that the proposed method gives the watermarked video of better quality and is robust against MPEG coding, down sampling and re-encoding to other type of video format such as MPEG4, H.264

Key Words : Wavelet, Watermarking, Encoding, Video, MPEG, DWT

Ⅰ. Introduction

The development of compression technology allows the wide-spread use of multimedia applications. Nowadays, digital images can be distributed via the World Wide Web to a large number of people in a cost-efficient way. The increasing importance of digital media, however, brings also new problems which are related with the duplication and manipulation of the multimedia content. There is a strong need for security services in order to keep the distribution of digital multimedia work both profitable for image owner and reliable for the customer.

In particular, digital video is usually compressed with MPEG-1[1] or MPEG-2[2] and will be stored and distributed in compressed format for video applications such as DVD (Digital Versatile Disc) and VOD (Video on Demand). Therefore, the watermark must be robust to MPEG coding.

Langelaar proposed a method to embed a label into the I frame in DCT domain[3]. Linnarz proposed a method to embed data into frame sequence with a predefined order for picture coding type[4]. These methods are weak to a change of the GOP structure of watermarked video on re-encoding MPEG2 to remove an embedded data.

Hartung proposed a method to embed a watermark in MPEG2 bit stream and detect it in the pixel domain[5]. This methods almost hide data into motion vector information in the MPEG-1 or MPEG-2 compressed domain. However, when once MPEG bitstream is re-encoded, the same motion vectors are not always selected. As a result, the embedded watermark can be easily erased by re-encoding video stream.

Barni proposed a scheme for the watermarking of MPEG-4 video objects[6]. The proposed algorithm works directly in the compressed domain thus reaching a high degree of flexibility and ease of use. They have proved that the proposed system presents some robustness against

※ 본 연구는 서울시 산학연협력사업으로 구축된 서울 미래형콘텐츠컨버전스 클러스터(SFCC) 및 교육부 BK21 사업의 지원으로 수행되었습니다.
* 중앙대학교 첨단영상대학원 영상공학과 영상콘텐츠보호 연구실(ksyoo@cau.ac.kr, whlee@cau.ac.kr)
논문번호 : KICS2007-07-302, 접수일자 : 2007년 7월 4일, 최종논문접수일자 : 2007년 12월 21일

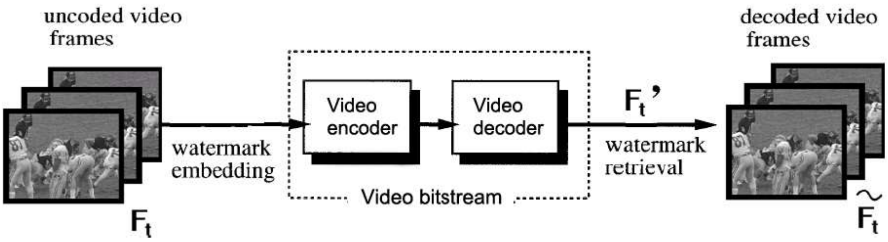

Fig. 1 The proposed video watermarking system in the uncoded and decoded domain

common manipulations such as re-coding at lower bit rates and frame dropping without re-encoding.

Profrock et al.[7] use Normed Centre of Gravity (NCG) to embed watermarking approach with robustness to H.264/AVC compression.

Furthermore, compressed domain watermarking introduces problems that a watermark embedding must be coupled tightly with a specific compression method. This coupling not only restricts the portability of the watermarking algorithm, but also imposes limitations set forth by the bitstream syntax and coding algorithm. The second problem is that of drift when the video is modified during watermark insertion. Drift occurs when (spatial or temporal) prediction is used and a predictor is modified without adjusting the residual to compensate for the new predictor. A compressed-domain watermarking technique must compensate for drift during watermark insertion to prevent drift from spreading and accumulating, leading to visible artifacts in the decoded video.

We consider in this paper the problem of watermarking for video data. Furthermore, we must modify the architecture of MPEG encoder and decoder in order to implement video watermarking and the generational copy control is hardly under consideration by the previous methods listed above.

In this paper, as an alternative to the conventional methods, we propose a new method of digital watermark for video data. As shown in Fig. 1, in this system the watermark is embedded into the uncoded video and is retrieved from the decoded video as shown Fig. 1. We demonstrates that the right algorithm can embed gaussian watermarks into the low frequency area while preserving good fidelity. Our embedding strategy is based on a DWT (discrete wavelet transform)[8]. An video data is transformed in the frequency domain applying DWT. As a consequence, the watermark is robust against MPEG-1, MPEG-2, MPEG-4, H.264 coding and re-encoding.

This paper is organized as follows. The watermark embedding and extracting approach is described in Section 2. In Section 3, the experimental results are shown. The conclusions of our study are stated in Section 4.

Ⅱ. The Proposed Watermarking Method

2.1 Color Representation

We embed the watermark into each luminance frame in uncompressed video using the DWT. In color representation, the perceptual attributes of color are brightness, hue and saturation. Brightness represents the perceived luminance. Hue of color refers to ‘‘redness”, ‘‘greenness” and so on. Saturation is that aspect of perception that varies most strongly as more and more white light is added to a monochromatic light. Fig. 2 shows a perceptual representation of the color space. Brightness (W*) varies along the vertical axis, hue varies along the circumference, and saturation (S) varies along the radial distance. Based on this representation, there are several coordinate systems that have come into existence for a variety of reason.

In NTSC format, image data consists of three components: luminance (Y), hue (I) and saturation (Q). The first component, luminance, represents

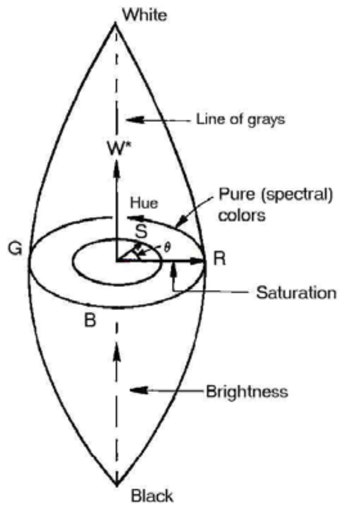

Fig. 2 Perceptual representation of the color space

grayscale information, while the last two components make up chrominance (color information). One of the main advantages of this format is that gray-scale information is separated form color data, so the same signal can be used for both color and black and white sets. Therefore, the watermark will be inserted in the luminance (Y) components of the color image in DWT domain.

Original frame can represent transform domain signal such as the DWT coefficients. First, an original frame in which the watermark is to be embedded is decomposed by a DWT. In our embedding procedure, watermarks are embedded into low frequency domain. The DWT converts a signal into low and high frequency area.

The watermark sequence which is a Gaussian sequence of owner number. The watermark is embedded according to PRNG (pseudo random number generator). That is to say, we created PRNG using seed key. And the watermark sequence permuted according to the created PRNG. In the watermark detection procedure, we perceive the watermark distribution location in the reverse method. Namely, we recreate PRNG by applying seed key used in the embedding procedure. We will describe the watermarking procedure in detail as the following.

2.2. Watermark embedding method

The first step is to extract the luminance component from the color image size. If the color image format is in RGB, then we need to convert it to YUV color space to get the luminance I_y, and then decomposed it through DWT in two levels. The wavelet transform is identical to a hierarchical sub band system, where the sub bands are logarithmically spaced in frequency. An image is first decomposed into four parts of high, middle, and low frequencies (i.e., LL1, HL1, LH1, HH1 sub-bands) by critically sub-sampling horizontal and vertical channels using sub-band filters as (1).

$$(W_\Psi)(a,b) = \frac{1}{|a|^{1/2}} \int_{-\infty}^{\infty} f(t)\Psi\left(\frac{t-b}{a}\right) dt \tag{1}$$

To obtain the next coarser scaled wavelet coefficients, the sub-band *LL1* is further decomposed and critically sub-sampled. We used the Haar wavelet transform that is a kind of discrete wavelet transform[8]. An video frame and its DWT decomposition are shown in Fig. 3.

Next, the subband HL2, LH2, HH2 is divided into blocks with 4x4 equal size. We embed the gaussian random sequence watermark in *LL2* that represent one watermark bit by modifying the wavelet coefficient with following equation:

We use a Gaussian sequence watermark *w*, where *w* sequence is chosen according to independent normal distributions with standard deviation . A secret key is chosen as the seed of predefined pseudo random number generator (or sometimes called as random permutation)[9]. The Gaussian sequence watermark is shuffled by using

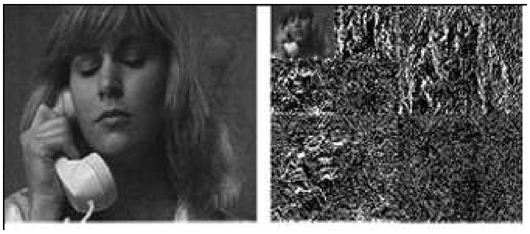

Fig. 3 Transformed Coefficients of DWT magnitude

PRNG. The *LL* band is modulated according to the formula (2) until watermark length times. A watermark embedding is done by using a watermark value w(i) ($1 \leq I \leq n$) to perturb $\tilde{x}_i$ as

$$m_i = |x_i - (x_{i-1} + x_{i+1})/2|,$$
$$\tilde{x}_i = (x_{i-1} + x_{i+1})/2 \cdot (1 + \alpha \cdot w_i), \quad if \quad m_i > Threshold$$
$$\tilde{x}_i = w_i, \quad Otherwise \tag{2}$$

After applying Eq. (2). The information data *w(n)* is embedded into the *LL*. Where $\tilde{x}_i$ is the modulated version of x_i, m_i is the check value for visibility, and α is a scaling factor which is the watermark strength and can be adjusted to achieve a reasonable compromise between the robustness of the watermark and its visibility. Large values of α lead to more robust schemes but the watermark becomes more visible because DWT coefficients are modified by a larger amount. After we insert the watermark, we do the DWT inverse transform to get the modified luminance component, and we compose the watermarked video frame with the modified luminance component, hue and saturation components.

2.3 Watermark Extraction Method

In blind watermarking techniques, where the un-watermarked original is not available at the decoder, the detector must synchronize, [10] with the spatial and temporal coordinates of the watermark signal for reliable detection. De-synchronization may occur as a result of a benign operation, such as changing the format to match a particular screen size (e.g., PanScan and letterbox) or as a result of a malicious attack to render the watermark undetectable. We classify such an extraction as blind watermarking algorithm.

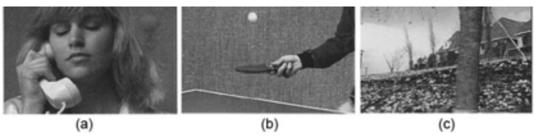

Fig. 4 Frames from “susie”, “table tennis”, and “°ower garden”

The original video data is not required in extracting watermarks.

To detect the watermark embedded in DWT domain, we use correlation. First, we extract the luminance component of watermarked color image and then decomposed it through DWT into two levels. The same decomposition used in the embedding is applied to both the original and embedded video. The watermark embedding locations are obtained by random permutation which is used in embedding procedure. We can extract the watermarks by equation (3) comparing the two values, $\tilde{x}_i$ and $(\overline{x_{i-1}}+\overline{x_{i+1}})/2$. Then the extracted watermarks are compared with the original watermarks generated by the *seed* key.

$$\tilde{w}_i = 0, \quad if \quad | \quad \tilde{x}_i - (\tilde{x}_{i-1} + \tilde{x}_{i+1})/2 \quad | \quad > Threshold$$
$$\tilde{w}_i = \frac{\tilde{x}_i - (\tilde{x}_{i-1} - \tilde{x}_{i+1})/2}{(\tilde{x}_{i-1} - \tilde{x}_{i+1})/2 \cdot \alpha}, \quad Otherwise \tag{3}$$

Ⅲ. Experimental results

We performed some numerical experiments with three video data (30 frames). The sample MPEG videos are obtained from David Sarnoff Research Center[11]. Three video sequences, “Susie”, “flower garden” and “table tennis” are used. The size of each luminance frame is 352 x 240 as shown in Fig. 4. Fig. 5. shows the relation between the number of the embedded watermark bits and PSNR. The PSNR of watermarked “Susie”, “flower garden” and “table tennis” video frames are larger than 51 dB.

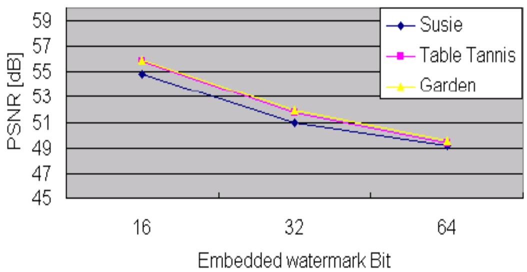

Fig. 5 The relation between the number of the embedded watermark bits and PSNR

3.1 On the Robustness against MPEG Coding

To compare the robustness of the proposed method against MPEG2 compression. Using the three watermarked video when the number of embedded watermark *w* = 1000, we performed numerical experiments with the MPEG codec source code where the GOP-length is 12, as shown in Fig. 6. Therefore, the total number of the data embedded into the I, P, B-frame (30 frames) and the I, P-frame (11 frames) are 30,000 and 11,000 bits, respectively. Fig. 7. shows the rate of the detection rate of the embedded data. It can be seen from Fig. 7. that the watermark can be almost extracted from the frames for "susie," while some bit errors occur in the case of "flower garden" and "table tennis" when the bitrate is less than 1.5 Mbit/s and 1 Mbit/s, respectively. However, the decoded data are degraded, too. The reason is as follows. The two test sequences "flower garden" and "table tennis" have much of the variation due to object motions and detailed patterns in comparison with the test sequence "susie." Hence, it is difficult to compress

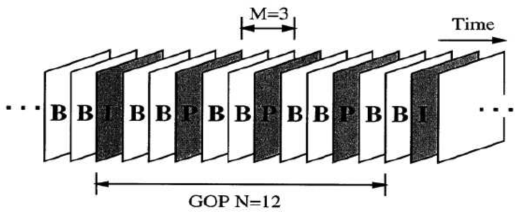

Fig. 6 An example of GOP including I, P, B-frames for [N,M]=[12,3]

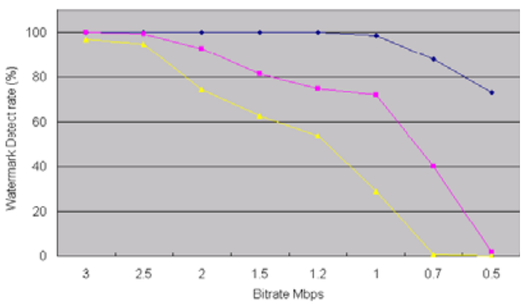

Fig. 7 Detect rate of the compressed video versus bitrate: "susie"(top), "table tennis", °ower garden"(bottom)

the two test sequences at low bitrate and "flower garden" and "table tennis" are degraded more than "susie" when compressed at the same bitrate. In the case of 1.5 Mbit/s and 1 Mbit/s, it is considered that this degradation caused the variation of watermarked wavelet coefficients and, as a result, the detect times per frame of the embedded watermark detected to 100 percent.

3.2 On the robustness against re-encoding to MPEG-4

As a new internet multimedia standard, MPEG4 provides the new mechanisms for video compression and distribution. We also investigate the robustness of the watermark to MPEG4 re-encoding. MPEG4 is fundamentally different from MPEG2 in that it is object based. We re-encode the same watermarked video data ten times over using Microsoft MPEG-4 video codec V3. As shown in Table. 1. the result, that is, detect rate per frame against MPEG-4 re-encoded video versus bitrate. The watermark can be clearly extracted from the each frames of the watermarked video stream.

Table 1. Detect rate per frame against Re-encoding to (MPEG-4)

(MPEG-4) Kbps	Susie	Table tennis	Flower Garden
984	100	43.6	32.4
768	100	38.9	14.1
512	80.4	32.3	8.7
384	37.3	7.9	1.2
128	3.2	0.5	0

3.3 On the robustness against re-encoding to H.264

Embedding the watermark in the uncompressed domain has the advantage, that the video data can be compressed with different standards and data rates. However, the embedded watermark has to be robust to the compression. At present in video data compression, the new H.264 standard, developed for a broad range of applications, provides the highest coding performance[12]. Compared with the MPEG2 standard, the H.264

Fig. 8 MPEG2 500bps(top), H.264 500bps(bottom)

Table 2. Detect rate per frame against Re-encoding to H.264

(H.264) Kbps	984	512	450	384	320	256	128
Detect rate(%)	99.4	91.3	82.5	35	28.6	14.4	1.06

standard achieves a three times lower data rate at the same video quality as shown Fig. 8.

By using a H.264 coder we compressed the "Susie" videos with different quality parameters [Single Pass Bit rate conversion]. As shown in Table. 2. the result, that is, detect rate per frame against H.264 compression with quality parameter between 128 and 984. The watermark can be extracted from the each frames of the watermarked video stream.

Ⅳ. Conclusion

Due to generalization of digital media and development of various multimedia data transmission devices, storage devices, and authoring tools, multimedia data may be easily copied and transformed via a network, thereby providing new services, but also generating a pirate copy of copyrighted contents of digital media. With the advent of digital video and digital broadcasting, copyright protection of video data has been one of the most important issues. In the proposed method, a watermark signal is inserted into the decompressed bitstream while detection is performed using the uncompressed video. This method allows detection if video has been manipulated or its format changed. We have proposed a method of digital watermarking for video data based on the DWT, and confirmed that the proposed watermarking method is able to extract the embedded data from each frame compressed with MPEG-1, MPEG-2, MPEG-4 and H.264 without using the original video data. As the watermark is embedded into the low-band of DWT. The embedded watermark is perceptually invisible and is robust.

The experimental results obtained by the computational simulation show that the presented method satisfies perceptual invisibility and it is sufficiently robust against MPEG coding, re-encoding, and different type of video transform attacks.

The proposed algorithm stands out for its simplicity, flexibility and low computational burden, thus being a suitable candidate for a number of novel and interesting applications, such as video fingerprinting for Set-Top Box.

Reference

〔1〕 ISO/IEC11172, Informationtechnology-Coding of moving pictures and associated audio for digital storage media at up to about 1.5Mbit/s, 1993.

〔2〕 ISO/IEC13818-2, "Information technology. Genericcoding of moving pictures and associated audio ,ITU-TRec.H.262," 1995.

〔3〕 Gerrit C. Langelaar, Reginald L, Lagendijk and Jan Biemond, "Real-time Labeling Methods for MPEG Compressed Video," *Proceeding of 18'th Symposium on Information Theory in the Benelux, Veldhoven*, The Netherlands, May 1997.

〔4〕 J.P.M.G. Linnartz μA¸ J.C. Talstra, "MPEG PTYMarks: Cheap Detection of embedded

Copyright Data in DVD-Video," *Proceeding of ESORICS'98 5'th European Symposium on Research In Computer Security*, Belgium, September 1998.

〔5〕 Frank Hartung and Bernd Girod, "Digital Watermarking of MPEG2-Coded Video in the Bit Stream Domain," *Proceeding of Int. Con. on Acoustics, Speech and Signal Processing,* vol. 4, pp. 2621-2624, April 1997.

〔6〕 Barni M., Bartolini F., Checcacci N., "Watermarking of MPEG-4 video objects," *IEEE Transactions on Multimedia*, vol 7, Issue 1, pp. 23-32, Feb. 2005

〔7〕 Dima Profrock, Mathias Schlauweg, Erika Muller., "A new uncompressed-domain video watermarking approach robust to H.264/AVC compression," 4th IASTED international conference on Signal processing, pattern recognition, and applications table of contents, pp. 99-104, 2006.

〔8〕 Albert Bogges Francis J. Narcowich., "A First Course in Wavelets with Fourier Analysis," *Prentice Hall*, Upper Saddle River, pp. 155-182, 2001.

〔9〕 Stefan Katzenbeisser, Fabien A. P, Petttcolas., "Information hiding techniques for steganography and digital watermarking," *Artech House*, Boston London, pp. 128-129, 2000.

〔10〕 E. Lin and E. Delp, "Temporal synchronization in video watermarking," *in Proc. SPIE Security and Watermarking of Multimedia Contents IV*, vol. 4675, San Jose, CA, pp. 478-490, Jan. 21.24 2002.

〔11〕 David Sarnoff Research Center, Inc., ftp://ftp.tek.com/tv/test/streams/Element\\/index.html

〔12〕 T. Wiegand, G. J. Sullivan, G. Bjontegaard, A. Luthra, "Overview of the H.264 / AVC Video Coding Standard," IEEE Transactions, Vol. 13, pp. 560-576, July 2003.

유 길 상 (Kil-Sang Yoo) 준회원

2003년 8월 중앙대학교 영상공학과 석사

2005년 9월 중앙대학교 영상공학과 박사수료

<관심분야> 디지털 영상처리, DRM, Watermarking

이 원 형 (Won-Hyung Lee) 정회원

현재 중앙대학교 첨단영상대학원 교수

현재 (사)한국컴퓨터게임학회 부회장

<관심분야> 디지털 영상 콘텐츠 보호, 컴퓨터게임