

압축 영역에서 시간축 확산 스펙트럼을 이용한 장면단위의 비디오 워터마킹

최 윤 희, 강 경 표, 최 태 선
 광주과학기술원 기전공학과
 Tel: 062-970-2419

Scene-Based Video Watermarking Using Temporal Spread Spectrum in Compressed Domain

Yun-Hee Choi, Kyung-Pyo Kang and Tae-Sun Choi
 Dept. of Mechatronics, Kwangju Institute of Science and Technology (K-JIST)
 E-mail: dals@sipl.kjist.ac.kr

요 약

This paper presents robust and efficient scene-based video watermarking method using visual rhythm (spatio-temporal slice) in compressed domain. Scene change can be detected easily using visual rhythm and video sequences are conveniently edited at the scene boundaries. Therefore, scene-based watermark embedding process is a natural choice. Temporal spread spectrum can be achieved by applying spread spectrum methods to visual rhythm. Additive Gaussian noise, low-pass filtering, median filtering and histogram equalization attacks are simulated for all frames. Frame sub-sampling is also simulated as a typical video attack. Simulation results show that proposed algorithm is robust and efficient in the presence of such kind of attacks.

1. 서 론

디지털 콘텐츠는 아날로그 형태의 콘텐츠와는 달리 반복적인 복제와 배포에도 불구하고 음질이나, 화질의 저하가 없는 것이 특징이다. 인터넷의 발달과 디지털 방송의 실용화로 디지털 미디어의 사용이 증가하면서 콘텐츠의 불법적인 복사와 저작권 보호에 대한 관심이 증가하게 되었다. 콘텐츠의 불법적인 복사는 제작자로 하여금 새로운 콘텐츠 창작의 의지를 저하시키고 장기적으로는 디지털 미디어 사업의 발전을 저해하는 요인이 된다. 이러한 문제에 대처하기 위해 디지털 미디어에 저작권이나 배포에 관련된 정보를 삽입하는 방법이 제안되었다. 디지털 워터마킹이란 저작권이나 배포 경로를 확인하기 위해 디지털 미디어에 사람이 감지할 수 없는 추가의 정보를 삽입하는 것을 말한다. 그간 비디오 워터마킹에 관련된 몇 가지 방법이 제시되었다[8]. M. K. Swanson은 Multiresolution 비디오 표현을 이용한 장면단위의 비디오 워터마킹 방법을 제시하였다[5]. 이 방법은 비디오가 장면단위로 분할되고 각각의 장면에서 시간축의 웨이블릿 변환을 적용하여 시간축 저주파 프레임과 고주파 프레임에 워터마크를 삽입한다. 이 방법은 워터마크를 장면단위로 삽입하기 위해 별도의 장면 전환 감지 알고리즘에 의해서 비디오가 장면단위로 나누어져 있다고 가정한다. 따라서 이 방법은 비디오를

장면단위로 나누기 위한 별도의 노력이 필요하다. C. T. Hsu는 압축된 비디오에서 DCT값에 기반한 워터마킹 방법을 제안하였다[6]. 이 방법은 압축된 비디오의 GOP 구조에 의존하기 때문에 GOP구조를 변화시키는 공격에 취약한 단점이 있다. 본 논문에서는 압축된 비디오에서 비주얼 리듬을 이용하여 장면을 분할하고 각 장면 단위로 워터마크를 삽입하는 방법을 제안한다. 비주얼 리듬은 연속되는 프레임에서 특정 DC값들을 취하여 구성한다. 따라서, 워터마크가 프레임 내의 모든 DC값에 삽입되는 것이 아니기 때문에 매우 효율적이며 화질의 열화를 최소화 시킨다. 장면 전환점은 비주얼 리듬으로부터 쉽게 구할 수 있으며 [1-3], 일반적으로 비디오는 장면단위로 편집되기 때문에 장면단위의 워터마킹 방법은 비디오 편집에 강인한 특성을 가진다. 비주얼 리듬은 비디오의 시간정보를 포함한 이차원 영상이므로 비주얼 리듬에 공간영역의 확산 스펙트럼 방법을 적용하면 시간축 확산 스펙트럼을 얻을 수 있다. 본 논문의 구성은 다음과 같다. 2장에서는 비주얼 리듬의 개념에 대해 간단히 설명하고, 3장에서는 제안하는 장면단위의 워터마킹 방법에 대해 설명한다. 4장에서는 실험결과와 분석으로 제안된 알고리즘의 성능을 살펴 볼 것이며, 5장에서는 결론을 제시한다.

2. 비주얼 리듬 (Visual rhythm)

비주얼 리듬은 3차원 비디오 콘텐츠에 대한 2차원 영상 표현으로, 연속된 비디오 프레임에서 특정 DC값들을 취하여 구성한다. $f(x,y,t)$ 를 픽셀 (x,y) 에서의 화소 값으로, $VR(z,t)$ 를 비디오 시퀀스 $V=\{f(x,y,t)\}$ 의 비주얼 리듬으로 정의하면, $VR(z,t)$ 은 다음과 같이 나타낼 수 있다.

$$VR(z,t) = f'(x(z), y(z), t) \quad (1)$$

여기서 $f' = T[f(x,y,t)]$ 이고, $x(z), y(z)$ 는 z 의 일차원 함수이다. 또한, T 는 시간 t 에서의 임의의 변환이다. $x(z)$ 와 $y(z)$ 는 각각의 프레임으로부터 비주얼 리듬의 한 열을 구성하기 위한 샘플링 함수이다. 따라서, 파라미터 $x(z)$ 와 $y(z)$ 에 따라서 다양한 종류의 비주얼 리듬을 얻을 수 있다. 예를 들어 일반적인 샘플링 함수를 다음과 같이 정의하면,

$$\begin{aligned} x(z) &= a \cos z + x_0 \\ y(z) &= b \sin z + y_0 \end{aligned} \quad (2)$$

샘플링 함수는 파라미터 a, b, x_0, y_0 의 값에 따라서 수평선, 수직선, 원, 타원 등 다양한 형태를 가질 수 있다. 그림 1(a)는 실제 비디오의 비주얼 리듬을 보여준다. 식 (2)에서 파라미터 $a=0, b=0, x_0=\text{const}, y_0=0$ 의 값을 사용하였다, 즉, 원래 프레임의 수직선으로 샘플링 하였다.

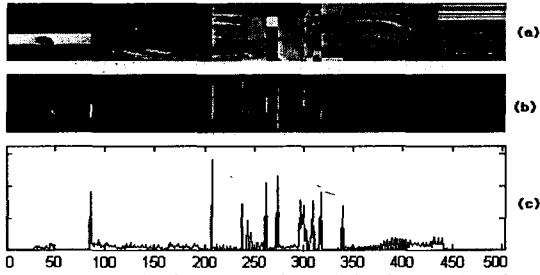


그림 1. 실제의 비디오로부터 구성한 비주얼 리듬영상

3. 비주얼 리듬을 이용한 장면 단위의 워터마킹

그림 3은 제안된 워터마킹 방법의 블록 다이어그램을 나타낸다. 블록 다이어그램의 윗부분은 워터마크 삽입 과정을, 아래 부분은 검출과정을 나타낸다. 워터마크를 삽입하기 위해 우선 주어진 비디오 V_0 로부터 비주얼 리듬을 구성한다. 다음으로 [1]에서 제안된 방법에 따라 장면전환점(Cut points)을 찾아낸다. 이 과정에서 전체 비주얼 리듬 영상이 몇 개의 장면단위의 서브영상으로 나뉘어 진다. 각각의 장면단위의 서브영상을 주파수 영역으로 변환하고 워터마크로서 의사 랜덤 잡음 시퀀스를 중간대역의 주파수 영역에 삽입한다. 워터마크가 삽입된 비주얼 리듬영상으로부터 다시 각각의 프레임을 복원하면 워터마크가 삽입된 비디오 시퀀스를 얻을 수 있다. 삽입된 워터마크는 삽입과정의 반대 과정을 거쳐 검출할 수 있다. 삽입과 마찬가지로 워터마크가 삽입되고 의도적인 또는 비의도적인 공격에 의해 변조된 비디오 시퀀스로부터 비주얼 리듬 영상을 구성하는 것이 검출의 첫번째 과정이다. 비주얼 리듬 영상을 주파수 영역으로 변환하고 삽입의 역과정을 거쳐 훼손된 의사 랜덤 잡음 시퀀스를 추출한다. 워터마크의 삽입을 확인하기 위하여 원래의 삽입된 워터마크와 추출된 워터마크 사이의 유사성을 측정한다. 이번 장의 나머지 부분에서는 각각의 과정에 대해서 자세히 설명한다.

3.1. 비주얼 리듬 영상의 구성

f 를 위한 변환 T 로서 여기서는 식 (3)에 나타난 바와 같이 8×8 블록 평균을 사용하였다.

$$f'(x, y, t) = \frac{1}{8} \sum_{i=0}^7 \sum_{j=0}^7 f(8x+i, 8y+j, t), \quad (3)$$

여기서 $x=0, 1, \dots, M/8, y=0, 1, \dots, N/8$ 이다.

MPEG-1, 2와 같은 압축된 비디오에서 하나의 프레임은 8×8 DCT 블록으로 이루어져 있다. 또한, 각각의 8×8 DCT 블록은 하나의 DC 계수와 63개의 AC 계수로 이

루어져 있다. 식 (3)의 $f(x, y, t)$ 는 8×8 DCT블록의 DC계수와 같은 값이다. DC계수는 각 8×8 블록 내의 픽셀들의 평균값의 8배와 같다. P와 B 프레임을 포함하여 빠르게 DC계수를 계산하는 알고리즘이 이미 많이 제안되었으므로 압축 비디오에서도 쉽게 비주얼 리듬을 구성할 수 있다.

3.2. 장면전환 탐지

장면의 경계점을 찾기 위해서 먼저 식 (4)와 (5)를 이용하여 $VR(z, t)$ 의 수직 에지 성분을 계산해야 한다. 식 (4)는 단순히 수직 성분의 에지를 찾는 수식이며, 식 (5)는 식 (4)에서 얻어진 에지를 수직 방향으로 프로젝션하는 것이다. 그림 1(b)는 그림 1(a)에 식 (4)를 적용한 결과이며, 그림 1(c)는 그림 1(b)에 식 (5)를 적용한 결과이다.

$$d(z, t) = |VR(z, t+1) - VR(z, t)| \quad (4)$$

$$d_p(t) = \sum_{z=0}^{n-1} d(z, t) \quad (5)$$

식 (5)의 결과와 [1]에 제시된 방법에 의해서 컷(cut), 와이프(wipe), 디졸브(dissolve) 등의 장면전환점을 찾아 낼 수 있다. 그러나 본 논문에서는 컷(cut)이 일어나는 장면전환만을 고려한다. 또한 워터마크의 삽입과 추출을 위해서는 아주 정확한 장면의 분할이 필요하지 않기 때문에, [1]에서 사용된 파라미터 K 의 값으로서 비교적 큰 값을 사용한다. 너무 작은 크기의 비주얼 리듬 영상에 워터마크를 삽입하면 워터마크 추출 성능이 떨어지기 때문에, 만일 장면 전환이 어떤 프레임 간격 d 보다 작은 프레임들 사이에서 일어나면 그 장면 전환점은 무시한다.

3.3. 워터마크 삽입

J. Cox 등은 정지영상에 확산 스펙트럼을 이용한 강인한 워터마킹 방법을 제안하였다. 이 방법은 전체 비주얼 리듬 영상에 바로 적용 가능하다. 그러나 장면단위로 워터마크를 삽입하면 비디오 편집에 강인하기 때문에 장면단위로 삽입하는 것이 바람직하다. 3.2절에 설명한 방법에 의해 전체 비주얼 리듬 영상은 각각의 장면을 이루는 N 개의 서브 영상으로 나눌 수 있다. 각각의 서브 영상은 다음과 같이 나타낼 수 있다.

$$VR_n = \{VR(z, t_s) | t_s \in S\}, \quad (6)$$

여기서 S 는 n 번째 장면에 속하는 모든 프레임들의 집합이며, $n=0, 1, \dots, N-1$ 이다. 이후 [4]에서 제시된 방법과 유사한 방법으로 각각의 VR_n 에 워터마크를 삽입한다.

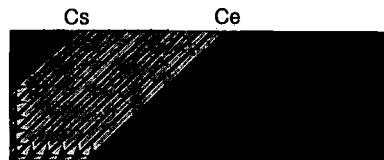


그림 2. 워터마크 삽입 대역

주파수 영역에서 워터마크의 삽입 대역은 [4]에서 제시된 대역과 다르다.

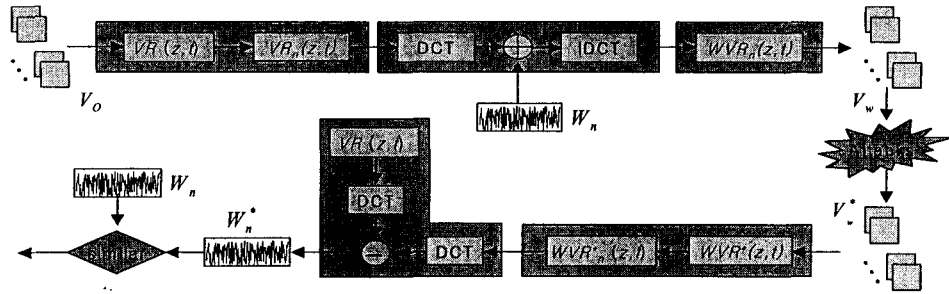


그림 3. 제안된 비디오 워터마킹 알고리즘의 블록 다이어그램

그림 2에 나타난 바와 같이 고정된 영역에 고정된 길이만큼 삽입하는 것이 아니라 응용에 따라 삽입 대역과 길이를 바꿀 수 있다. 여기서 C_s 는 삽입을 시작하는 열의 위치이며, C_e 는 삽입이 끝나는 열의 위치이다. 삽입 길이와 범위는 C_s 와 C_e 를 바꿈으로서 쉽게 변경할 수 있다. DCT 계수는 식 (7)과 같이 관계를 이용하여 변경한다.

$$v'_i = v_i(1 + \alpha x_i) \quad (7)$$

여기서 v_i 는 원래의 DCT계수이며, x_i 는 삽입될 의사 랜덤 잡음 시퀀스, α 는 워터마크의 강도를 조절하는 파라미터이며 v'_i 는 변경된 DCT계수의 값이다. 식 (7)은 DCT계수 값의 크기에 따라 삽입되는 워터마크의 크기를 조절하는 것이다, 즉, 큰 값의 계수에는 큰 값의 워터마크를 작은 값의 계수에는 작은 값의 워터마크를 삽입하는 것이다.

3.4. 워터마크 추출

본 논문에서는 각 서브 영상 VR_n 에 대해 워터마크의 삽입 유무를 결정하기 위해서 [4]에서 제시된 유사성 측정 함수 (similarity measure function)를 사용하였다.

$$sim(W_n, W_n^*) = \frac{W_n^* \cdot W_n}{\sqrt{W_n^* \cdot W_n^*}}, \quad (8)$$

여기서 W_n 과 W_n^* 은 각각 워터마크로서 삽입된 원래의 의사 랜덤 잡음 시퀀스와 공격에 의해 훼손된 의사 랜덤 잡음 시퀀스를 나타낸다. 만일 $sim(W_n, W_n^*)$ 가 어떤 문턱값 (Threshold) Th 보다 크다면, 우리는 W_n^* 이 원래의 워터마크 W_n 과 서로 일치한다고 판단할 수 있다.

4. 실험결과 및 분석

제안한 알고리즘을 검증하기 위해서, 본 논문에서는 352×288 크기의 300프레임의 table tennis 시퀀스를 사용하였다. 이 시퀀스는 131번째 프레임과 132번째 프레임 사이에 장면 전환이 있다.

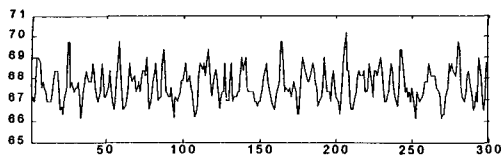


그림 4. 워터마크가 삽입된 비디오의 PSNR 분포
실험을 위해 파라미터 $\alpha = 0.2$, $C_s = 0.5 \times vrh$, $C_e = 1.5 \times vrh$ 를 사용하였다. 여기서 vrh 는 비주얼 리듬 영상의 높이를 나타낸다. 이 C_s 와 C_e 에 의해 주어진 삽입 영역은 DCT 주파수 영역에서 중간대역에 속한다.

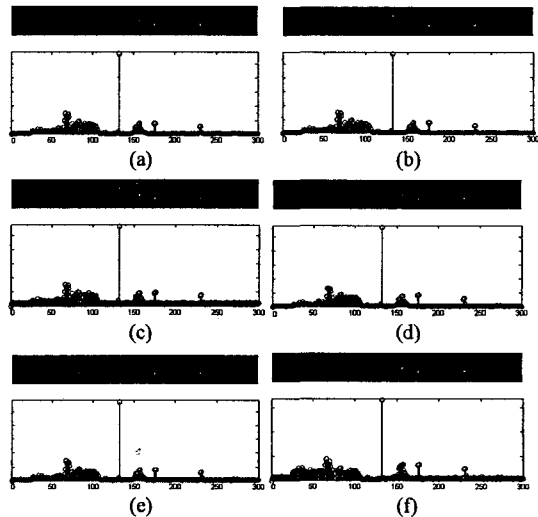


그림 5. 여러가지 공격에 대한 장면 전환 탐지 응답 (a) original, (b) watermark embedded, (c) Gaussian noise attacked, (d) lowpass filtering attacked, (e) median filtering attacked and (f) histogram equalization attacked sequences.

4.1. 워터마크된 비디오의 화질 평가

워터마크가 삽입된 비디오의 화질을 평가하기 위해서는 평가 기준을 정의하는 것이 필요하다. 본 논문에서는 모든 프레임의 PSNR의 평균값을 평가 기준으로 사용하였다. PSNR은 영상의 화질을 평가하기 위해서 널리 사용되며 다음과 같다.

$$PSNR[dB] = 10 \log_{10} \frac{255^2}{MSE}$$

여기서 MSE 는 Mean Square Error로서 평균 자승 오차를 나타낸다. 화소값의 범위는 0에서 255 사이에 존재한다고 가정한다. 그림 4는 워터마크가 삽입된 비디오의 PSNR 분포를 나타낸다. 300 프레임에 대한 평균 PSNR은 67.69였다. 이 값은 화질이 상당히 좋다는 것을

의미하는데 이는 제안된 방법이 프레임의 모든 열중에서 단지 8개의 열에만 워터마크를 삽입하기 때문이다.

4.2. 공격에 대한 장면전환 탐지

워터마크의 삽입이나 Gaussian noise, lowpass filtering, median filtering, histogram equalization과 같은 공격에 대한 장면전환 탐지의 응답을 실험하였다. 그림 5에 보인 바와 같이 모든 경우에 대해 강한 응답을 보였다. 그림 5(a)-(f)에 보인 바와 같이 모든 공격에 대해 최대 응답은 131번째에서 나타났다. 이것은 121과 132번째 프레임 사이에 장면전환이 나타난다는 것을 의미한다.

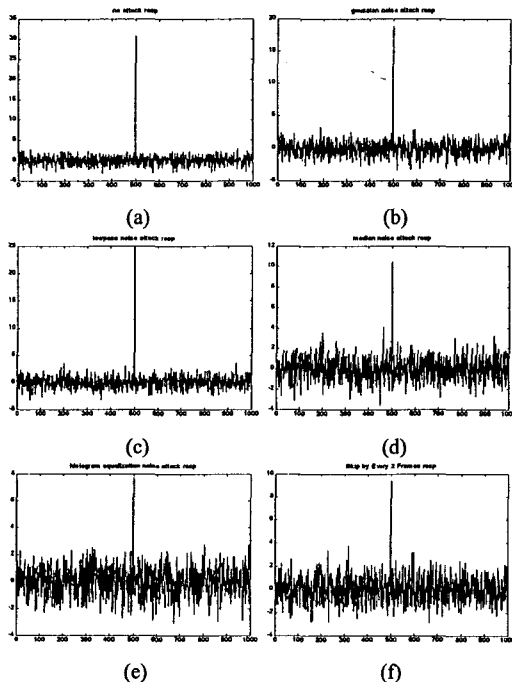


그림 6. 여러 가지 공격에 대한 워터마크 탐지 응답 (a) no attacked, (b) Gaussian noise attacked, (c) lowpass filtering attacked, (d) median filtering attacked, (e) histogram equalization attacked and (f) frame skipping attacked sequences.

4.3. 공격에 대한 워터마크 탐지

그림 6은 table tennis의 두 번째 장면(프레임 132 - 300)에 대한 Gaussian noise, lowpass filtering, median filtering, histogram equalization, frame skipping 등의 공격에 대한 응답을 나타낸다. 그림 6-(a)에 보여진 바와 같이 공격이 없는 경우 30.7293의 값을 나타내었다. 그림 6(b)는 noise variance 10의 additive Gaussian noise 공격에 대한 응답을 나타내며 18.7996을 보였다. 3x3 마스크를 사용하여 lowpass와 median filtering 공격에 대한 결과는 그림 6(c)와 6(d)에 보인 바와 같이 각각 24.8664과 10.4929를 나타내었다. 그림 6(e)는 histogram equalization 공격에 대한 응답을 나타내며 7.7082의 값을 나타내었는데 이는 제안된 알고리즘이 여전히 강인함을 보여준다. Frame skipping은 비디오 시퀀스에 대한 일반적인 공격이다. 그림 6(f)는 2프레임 간격의 스킵 공격에 대한

결과를 보여준다. 더 많은 프레임 스킵은 워터마크 탐지 응답의 감소를 가져올 것이다. 워터마크가 삽입된 장면의 크기, 즉, 하나의 장면을 이루는 프레임의 개수는 프레임 스킵 공격에 있어서 특히 중요한 요소이다. 장면의 크기가 크면 클수록 공격에 더욱 강한 응답을 보인다. 이 때문에 본 논문에서는 3.2절에서 설명한 바와 같이 너무 작은 간격 동안 발생한 장면전환은 무시하였다.

5. 결론

본 논문에서는 비주얼 리듬을 이용한 장면 단위의 워터마킹 방법이 제안되었다. 비주얼 리듬은 연속되는 프레임에서 어떤 변환 함수의 결과에 해당하는 DC 계수들만을 이용하여 구성하였다. 프레임 내의 모든 계수에 워터마크가 삽입되는 것이 아니기 때문에 제안된 방법은 매우 효율적이다. 제안된 방법을 검증하기 위해 additive Gaussian noise, low-pass filtering, median filtering, histogram equalization 등의 공격을 모든 프레임에 적용하고 워터마크 탐지 성능을 실험하였다. 또한 일반적인 비디오 시퀀스에 대한 공격으로서 프레임 스킵 공격에 대한 탐지 성능을 실험하였다. 실험결과, 제안된 방법이 이러한 종류의 공격에 강인하고 효율적이며 상당히 높은 화질을 유지함을 알 수 있었다.

참고문헌

- [1] M. G. Chung, H. Kim, and M. H. Song, "A Scene Boundary Detection Method," *Proceedings of IEEE International Conference on Image Processing*, pp. 933-936, 2000.
- [2] H. Kim, J. Lee, M. H. Song, "An Efficient Graphical Shot Verifier Incorporating Visual Rhythm," *IEEE International Conference on Multimedia Computing and Systems*, vol. 1, pp. 827-834, 1999.
- [3] C. W. Ngo, T. C. Pong, and R. T. Chin, "Video Partitioning by Temporal Slice Coherency," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 11, no. 8, pp. 941-953, August, 2001.
- [4] I. J. Cox, J. Kilian, F. T. Leighton and T. Shamon, "Secure Spread Spectrum Watermarking for Multimedia," *IEEE Trans. on Image Processing*, vol. 6, no. 12, pp. 1673-1687, December, 1997.
- [5] M. K. Swanson, B. Zhu and A. H. Tewfik, "Multiresolution Scene-Based Video Watermarking Using Perceptual Models," *IEEE Journals on Selected Areas in Communications*, vol. 16, no. 4, May 1998.
- [6] C. T. Hsu and J. L. Wu, "DCT-based Watermarking for Video," *IEEE Trans on Consumer Electronics*, vol. 44, no. 1, February, 1998.
- [7] M. D. Swanson, M. Kobayashi and A. H. Tewfik, "Multimedia Data-Embedding and Watermarking Technologies," *Proceedings of the IEEE*, vol. 86, no. 6, pp. 1064-1087, 1998.
- [8] F. Hartung and B. Girod, "Watermarking of Uncompressed and Compressed Video," *Signal Processing* 66, 283-301, 1998.