

# 3D-DCT 계수를 적응적으로 이용한 비디오 워터마킹

한지석 신현빈<sup>o</sup> 문영식

한양대학교 컴퓨터공학과

{jshan, hbshin<sup>o</sup>, ysmoon}@cse.hanyang.ac.kr

## Video WaterMarking Scheme with Adaptive Embedding in 3D-DCT domain

Ji Seok Han Hyun Bin Shin<sup>o</sup> Young Shik Moon

Dept. of Computer Science Engineering, Hanyang University

### 요약

비디오는 인접한 프레임들간에 유사성이 있는 연속된 프레임들로 구성된다. 만약 인접한 프레임들간에 유사성이 존재하는 영역, 즉 움직임이 없는 영역에 워터마크를 삽입한다면 워터마크는 인지되기 쉽다. 본 논문에서는 워터마크의 투명성과 강인성을 위하여 이와 같은 비디오의 특성을 고려한 3D-DCT 계수를 이용한다. 즉, 3D-DCT 압축을 위한 양자화 상수에서 민감도를 유도하고 전역적인 움직임에 비해 지역적인 움직임이 큰 영역의 민감도를 조절한 후 움직임의 크기에 비례해서 시각적으로 중요한 계수에 워터마크를 삽입한다. 실험을 통하여 비디오의 특성을 고려하지 않고 3D-DCT 계수를 이용하는 기존의 방법과 비교해서 PSNR은 유사하지만 JND를 기반으로 하였기 때문에 워터마크의 투명성을 보장하였고 MPEG 압축 및 시간적 공격에 대한 강인성은 약 5% 정도의 성능향상이 있음을 확인하였다.

### 1. 서론

최근 몇 년 동안 디지털 멀티미디어의 발달과 더불어 멀티미디어(audio, image, video)에 대한 사람들의 관심이 증가하고 있다. 그에 따라 상업적 가치가 증가되는 상황에서 각종 멀티미디어 데이터들이 네트워크와 인터넷을 통해 쉽게 분배, 복제되고 있는 실정이며, 디지털 데이터 내부에 정보를 은닉하여 멀티미디어 데이터의 저작권을 보호할 수 있는 연구가 활발히 진행되고 있는데, 대표적인 예로서 멀티미디어 데이터의 저작권 보호를 위해 디지털 데이터 내부에 개인의 정보를 은닉하여 확인할 수 있는 기술을 디지털 워터마킹(digital watermarking)이라고 한다. 이전의 워터마킹 기술과 달리 현대의 디지털 워터마킹은 인간의 시각과 청각이 완벽하게 디지털 데이터를 인식하지 못한다는 사실을 이용하게 된다. 본 논문에서는 JND를 기반으로 한 3차원 지각 모델을 제시하고 전역 움직임에 비해 움직임이 큰 구부영역은 인간의 시각으로 인지하기 어렵다는 것을 이용하여 3D-DCT의 저주파 대역 계수의 크기를 적응적으로 이용하는 방법을 제안하였다.

### 2. 제안된 워터마킹 방법

본 논문에서 제안된 비디오 워터마킹 알고리즘은 크게 워터마크 생성, 워터마크 삽입, 워터마크 검출과정으로 나누어진다. 먼저 워터마크를 생성하기 위하여 가우시안 분포를 갖는 랜덤 벡터를 생성한다. 가우시안 분포를 갖는 랜덤 벡터를 생성하기 위해서는 Box-Muller 변환을 사용한다. 워터마크 삽입과정에서는 8프레임 단위로 8x8x8 블록을 만들어 3D-DCT 계수를 생성하고 JND(Just Noticeable Difference)를 기반으로 한 3차원 지각 모델을 이용하여 저주파 대역에 워터마크를 삽입한다. 3차원 지각 모델을 이용하여 워터마크의 투명성을 보장했고 워터마크의 강인성을 위하여 공격에 강한 저주파 대역에 삽입하였다. 검출과정은 복호화 과정 후에 일어나는 프레임들에 삽입 과정에서 사용했던 3D-DCT를 적용하고 생성된 3D-DCT 계수와 워터마크의 상관관계를 이용하여 워터마크를 검출한다. 그림 1은 워터마킹 시스템의 흐름도를 나타낸다.

#### 2.1 3차원 지각 모델(Perceptual Model)

인간 시각 시스템(Human Visual System)을 이해하고 영상 처리에 응용하는 것은 수년간 연구 되어왔다. 그 중에서 2차원 시각 모델(Visual Model)은 영상 압축에 사용되고 있는데 시각

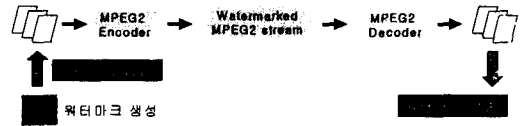


그림 1. 워터마킹 시스템 흐름도

모델을 이용한 통상적인 압축 방법은 영상의 특징을 고려하는 JND(Just Noticeable Difference) 이론을 기반으로 한 마스크를 양자화 테이블에 적용하는 것을 기반으로 한다[1]. 이러한 시각 모델은 워터마킹으로 확장되어지는데, 본 논문에서는 기존의 2차원 시각 모델을 확장한 3차원 지각 모델(Perceptual Model)을 제안하고 인간 시각 시스템을 기반으로 하는 워터마킹 알고리즘을 제안한다. 2D 시각 모델을 3D 지각 모델로 확장하기 위해서 chan이 제안한 비디오 압축을 위한 3D-DCT 양자화 방법을 이용한다[2]. chan의 모델에 의하면 식 1과 식 2에 의하여 3차원 큐브를 그림 2와 같이 시각적으로 중요한 계수와 시각적으로 중요하지 않은 계수로 나눈다.

$$S(u, v, w) = (u+1)(v+1)(w+1) \leq C \quad (\text{식 1})$$

$$S(u, v, w) = (u+1)(v+1)(w+1) > C \quad (\text{식 2})$$



(a) 시각적으로 중요한 계수 (b) 시각적으로 중요하지 않은 계수

그림 2. 3차원 쌍곡선에 의한 3D-DCT 계수의 분류

chan은 그림 2의 (a)와 (b)에 속한 계수들의 크기가 지수적으로 감소한다고 하였고 식 3의 지수합수를 이용해서 그림 2의 (a), (b) 각 부분에 속한 계수의 양자화 상수를 구하였다

$$q(u, v, w) = \begin{cases} \frac{e^{-\beta_{in} S(u, v, w)}}{255(1 - \frac{e^{-\beta_{in} S(u, v, w)}}{e^{-\beta_{in} S(u, v, w)}}) + 1} & S(u, v, w) \leq C \\ \frac{e^{-\beta_{out} S(u, v, w)}}{255(1 - e^{-\beta_{out} S(u, v, w)})} & S(u, v, w) > C \end{cases} \quad (\text{식 3})$$

chan의 방법에 의해 구하여진 양자화 상수를 이용하여 민감도(Sensitivity)를 구한다. Watson이 제안한 방법에 의해서 3D-DCT 계수의 민감도는 식 4가 됨을 알 수 있다.

$$t(u,v,w) = q(u,v,w) / 2 \quad (식 4)$$

본 논문에서는 이것을 움직임 마스크(Motion masking)이라고 정의하고 Watson의 밝기 마스크(luminance masking)과 유사하게 식을 유도한다. 국부지역의 움직임의 크기는 각 큐브의 움직임의 크기이고 이를 구하기 위해 그림 3의 A 영역을 이용하면 각 큐브의 움직임을 구할 수 있다.

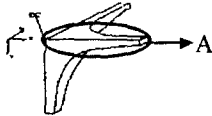


그림 3. 시간축 계수

국부 움직임의 크기는 식 5를 이용해서 A영역에 속한 계수의 크기의 총합으로 구한다. 전역 움직임의 크기는 식 6에 나타낸 것과 같이 국부 움직임에 대한 평균 크기를 구한다. 식 7은 전역 움직임과 국부 움직임의 크기에 비례하게 민감도를 수정하는 움직임 마스크를 나타낸다.  $X_{uvw}$ 는 DCT 계수,  $M_k$ 는 k번째 큐브의 움직임의 크기,  $M_{global}$ 은 전역 움직임의 크기이며  $M_k$ 의 평균으로 구한다.  $a_T$ 는 상수이고  $t_{(u,v,w),k}^M$ 는 움직임 마스크에 의해 수정된 민감도이다.

$$M_k = \sum X_{uvw}, \quad u < 3, v < 3, 0 < w < 8, S(u,v,w) \leq C \quad (식 5)$$

$$M_{global} = \frac{1}{m} \sum_{k=1}^m M_k, \quad m \text{은 cube 의 개수} \quad (식 6)$$

$$t_{(u,v,w),k}^M = t_{(u,v,w)} (M_k / M_{global})^{a_T} \quad (식 7)$$

위터마킹에서 고려할 점은 위터마크의 강인성과 투명성이다. 본 논문에서는 위터마크의 투명성을 위해서 식 7과 같이 수정된 민감도를 이용해서 위터마크를 삽입하고 위터마크의 강인성을 위해서 저주파 대역에 해당하는 그림 2의 (a) 영역에 위터마크를 삽입한다.

### 2.2 위터마크의 생성

본 논문에서 사용할 위터마크는 삽입시 시간적으로 비가시성을 가지면서 이전 위터마크(binary watermark)보다 견고한 가우시안 분포 랜덤 벡터를 사용한다. 가우시안 분포를 가지는 랜덤 벡터를 생성하기 위해 식 8의 Box-Muller 변환을 사용하여 균등 분포(uniform distribution)를 갖는 랜덤 벡터  $u_1, u_2$ 로부터 가우시안 분포를 갖는  $R_i$ 를 생성하게 된다.

$$R_i = \sin(2\pi u_1) \sqrt{-2 \log u_2} \quad (식 8)$$

생성된 가우시안 랜덤 벡터 R은 식 9에 의해 변조된다. 변조된 랜덤 벡터 W는 민감도에 의해 크기 변화되어 3D-DCT 계수에 삽입되게 된다.

$$W = \begin{cases} R \times 1, & \text{if Watermark bit} = 1 \\ R \times -1, & \text{if Watermark bit} = 0 \end{cases} \quad (식 9)$$

### 2.3 위터마크 삽입

비디오는 부피를 갖는 데이터의 집합이고 3차원 신호로 간주할 수 있다. 비디오의 3차원 신호를 3D-DCT하여 얻은 계수는 3차원의 큐브안에 존재하게 되고 시공간적인 정보를 갖고 있으

며 시간적 공격에 강하다[3]. 이때 삽입되는 위터마크는 원영상의 화질저하를 가져오지 않는 범위내에서 삽입될 수 있어야 한다. 따라서 제안한 방법에서는 시간적으로 중요한 3D-DCT 계수를 선택하고 화질 저하를 막기 위해 JND를 기반으로 한 3차원 지각 모델을 이용하여 위터마크를 삽입한다. 그림 4는 전체적인 삽입과정을 나타낸다.

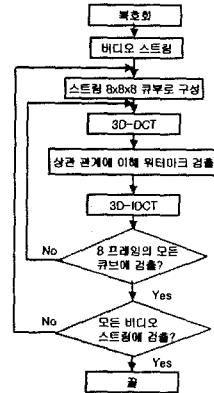


그림 4. 전체적인 삽입과정

위터마크 삽입의 첫 번째 과정에서는 부호화 작업을 하기 전의 비디오 프레임을 8프레임 기준으로 8x8x8 크기의 큐브로 나누고 각 큐브에 3D-DCT를 수행한다. 식 10은 일반적인 3D-DCT를 계산하는 식으로써 k는 큐브의 인덱스이고  $f_k(x,y,z)$ 는 화소의 밝기 값,  $F_k(u,v,w)$ 는 3D-DCT의 계수를 나타낸다.

$$F_k(u,v,w) = C(u)C(v)C(w) \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} \sum_{z=0}^{N-1} f_k(x,y,z) \cos \left[ \frac{(2x+1)u\pi}{2N} \right] \cos \left[ \frac{(2y+1)v\pi}{2N} \right] \cos \left[ \frac{(2z+1)w\pi}{2N} \right] \quad (식 10)$$

$$C(u) = \begin{cases} \sqrt{1/N}, & u = 0 \\ \sqrt{2/N}, & \text{otherwise} \end{cases}$$

본 논문에서는 계산의 효율을 위해 식 10 대신 3D-DCT의 분리성을 이용하여 x축, y축, z축으로 분리하여 구현하였다. 위터마크 삽입의 두 번째 과정은 3D-DCT 계수의 저주파 대역에 3차원 지각 모델을 적용하여 위터마크를 삽입하는 것이다. 3D-DCT 계수의 저주파 대역은 위터마크의 강인성을 보장하지만 화질저하를 초래할 수 있다. 제안하는 방법은 화질 저하를 방지하기 위해 3차원 지각모델에 근거하여 3D-DCT 계수의 민감도를 구하고 움직임 마스크를 통해 민감도를 수정하여 그림 2의 (a) 영역에 위터마크를 삽입한다.

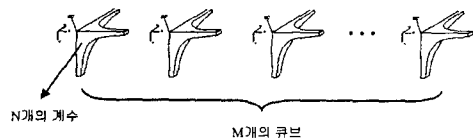


그림 5. 위터마크 삽입 영역

그림 5를 보면 그림 2의 (a)영역에는 N개의 계수가 있고 비디오의 8프레임마다 M개의 큐브가 있으므로 M개의 원소를 갖는 N개의 그룹을 구할 수 있다. 제안하는 방법은 식 9에 의해 만들어진 가우시안 랜덤 벡터를 이용하여 그룹당 1비트씩 N개의 그룹에 N개의 비트를 삽입한다. 식 11에서 보는 바와 같이 민감도 보다 큰 3D-DCT 계수는 민감도에 비례하게 위터마

크를 삽입하게 된다.  $S_{n,k}$ 는 움직임 마스크로 수정된 민감도,  $W_k$ 는 식 9에 의해 만들어진 위터마크,  $X_{n,k}$ 는 그림 3의 A영역에 속하는 3D-DCT 계수이며  $X'_{n,k}$ 는 위터마크가 삽입된 계수이다.  $\alpha$ 는 비례상수이고  $n$ 은 그룹의 인덱스이며  $k$ 는 그룹에 속한 원소의 인덱스로서 큐브의 개수와 같다.

$$X'_{n,k} = \begin{cases} X_{n,k} + \alpha S_{n,k} W_k, & \text{if } X_{n,k} > S_{n,k} \\ X_{n,k} + W_k, & \text{otherwise} \end{cases} \quad (\text{식 11})$$

마지막으로 위터마크가 삽입된 8x8x8 큐브를 3D-IDCT(Inverse DCT)한다. 3D-IDCT 식은 식 12과 같다. 이때  $F'_k(u, v, w)$ 는 위터마크가 삽입된 DCT 계수이고  $f'_k(x, y, z)$ 는 3D-IDCT 결과로 역 변환된 위터마크가 삽입된 영상의 화소값이다.

$$f'_k(x, y, z) = \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} \sum_{w=0}^{N-1} C(u)C(v)C(w)F'_k(u, v, w) \cos\left[\frac{(2x+1)u\pi}{2N}\right] \cos\left[\frac{(2y+1)v\pi}{2N}\right] \cos\left[\frac{(2z+1)w\pi}{2N}\right] \quad (\text{식 12})$$

첫 번째 과정부터 마지막 과정까지 모든 프레임에 대해서 수행하면서 위터마크를 삽입한다.

2.4 위터마크 검출

위터마크의 검출과정은 삽입과정과 유사하다. 그림 6은 위터마크 검출과정의 전체적인 흐름도이다. 먼저 압축된 동영상을 복호화하여 위터마크가 삽입된 프레임들을 얻는다. 연속되는 프레임들에 대해서 삽입과정과 동일한 방법으로 8x8x8 블록의 3D-DCT 계수를 얻는다. 식 13을 이용하여 위터마크된 계수  $X'$ 와 위터마크  $W$ 의 정규화된 상관관계(normalized correlation)  $NC$ 를 구한다. 그리고 식 14를 통해  $NC$ 의 부호를 검사하여 위터마크 비트를 검출한다.

$$NC = (X_n, W) = \sum_{k=0}^{M-1} \bar{X}'_{k,n} \bar{W}_k \quad \begin{cases} \bar{X}'_{k,n} = \frac{X'_{k,n}}{|X_n|} \\ \bar{W}_k = \frac{W_k}{|W|} \end{cases} \quad (\text{식 13})$$

$$\text{watermark bit} = \begin{cases} 1, & \text{if } NC \geq 0 \\ 0, & \text{if } NC < 0 \end{cases} \quad (\text{식 14})$$

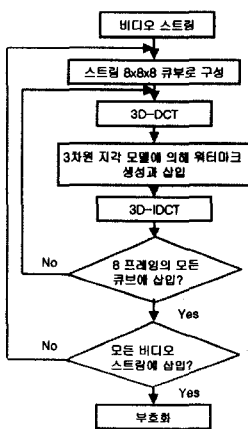


그림 6. 전체적인 검출과정

3. 실험결과

실험은 새로 제안한 방법으로 미스 아메리카를 통하여 이루어졌고, 실험을 통하여 비디오의 특성을 고려하지 않고 3D-DCT 계수를 이용하는 기존의 방법과 비교해서 PSNR은 유사하지만 JND를 기반으로 하였기 때문에 위터마크의 투명성을 보장하였고 MPEG 압축 및 시간적 공격에 대한 강인성은 약 5% 정도의 성능향상이 있음을 확인하였다. MPEG 압축과 시간적인 공격없이 위터마크를 삽입한 후 바로 검출했을때의 검출률과 PSNR(Peak Signal to Noise Ratio)을 그림 7에 나타내었다. 그림 8은 여러 가지 경우에서의 검출률이다.

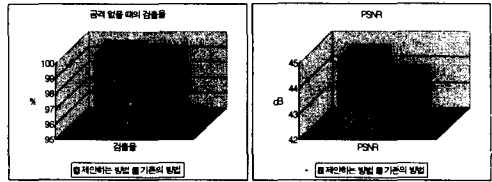


그림 7.

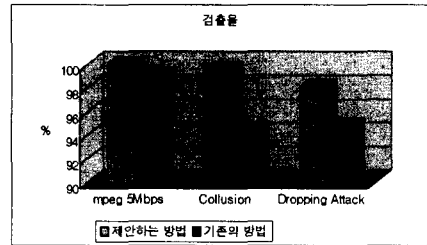


그림 8.

4. 결론 및 향후과제

본 논문에서는 JND(Just Noticeable Difference)를 기반으로 한 3차원 지각 모델(Perceptual Model)을 제시하고 비디오의 특성을 고려하여 3D-DCT 계수에 위터마크를 삽입하는 방법을 제안하였다.

본 논문에서는 위터마크의 투명성과 강인성을 위하여 비디오의 특성을 고려하여 3D-DCT 계수를 이용하였다. 3D-DCT 압축을 위한 양자화 상수에서 민감도를 유도하고 전역적인 움직임에 비해 지역적인 움직임이 큰 영역의 민감도를 조절한 후 움직임의 크기에 비례해서 시간적으로 중요한 계수에 위터마크를 삽입하였고, 실험을 통하여 성능향상을 확인할 수가 있었다.

향후 연구 과제로는 미스 아메리카보다 움직임이 많은 다른 동영상에 대한 실험이 필요하다.

참고문헌

- [1] Andrew B. Watson, "DCT quantization matrices visually optimized for individual images," *Human Vision, Visual Processing, and Digital Display IV, Proc. SPIE* 1913, pp. 202-216, 1993.
- [2] Chan R.K.W., Lee M.C., "3D-DCT quantization as a compression technique for video sequences," *Virtual Systems and Multimedia, 1997. VSMM '97. Proceedings., International Conference*, pp.188-196, Sept. 1997.
- [3] J. H. Lim, D. J. Kim, H. T. Kim, C. S. Won "Digital Video Watermarking Using 3D-DCT and Intra-Cubic Correlation," *Proc. of SPIE In Security and Watermarking of Multimedia Contents III*, vol.4314, pp.64-72, 2001