

3D-DCT 계수를 적응적으로 이용한 비디오 워터마킹*

박 현,[†] 한 지석, 문 영식[‡]

한양대학교

Video Watermarking Scheme with Adaptive Embedding in 3D-DCT domain*

Hyun Park,[†] Ji-Seok Han, Young-Shik Moon,[‡]

Hanyang University

요 약

본 논문에서는 JND를 기반으로 한 3차원 지각 모델을 제시하고 비디오의 특성을 고려하여 3D-DCT 계수에 워터마크를 삽입하는 방법을 제안한다. 비디오는 인접한 프레임들간에 유사성이 있는 연속되는 프레임들로 구성된다. 만약 인접한 프레임들간에 유사성이 존재하는 영역, 즉 움직임이 없는 영역에 워터마크를 삽입한다면 워터마크는 인지되기 쉽다. 그러므로 워터마크의 투명성을 위하여 움직임이 있는 영역에 워터마크를 삽입하여야 하고 강인성을 위하여 삽입되는 워터마크의 양을 조절할 필요가 있다. 본 논문에서는 워터마크의 투명성과 강인성을 위하여 비디오의 특성을 고려하여 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. 본 논문에서는 실험을 통하여 비디오의 특성을 고려하지 않고 3D-DCT 계수를 이용하는 기존의 방법과 제안한 방법을 비교하였다. 그 결과 비록 PSNR은 유사하지만 JND를 기반으로 하였기 때문에 워터마크의 투명성이 보장되고 MPEG 압축 및 시간적 공격에 대한 강인성은 약 3~9% 정도의 성능향상이 있음을 확인하였다.

ABSTRACT

This paper introduces a 3D perceptual model based on JND(Just Noticeable Difference) and proposes a video watermarking scheme which is perceptual approach of adaptive embedding in 3D-DCT domain. Videos are composed of consecutive frames with many similar adjacent frames. If a watermark is embedded in the period of similar frames with little motion, it can be easily noticed by human eyes. Therefore, for the transparency the watermark should be embedded into some places where motions exist and for the robustness its magnitude needs to be adjusted properly. For the transparency and the robustness, watermark based on 3D perceptual model is utilized. That is, the sensitivities from the 3D-DCT quantization are derived based on 3D perceptual model, and the sensitivities of the regions having more local motion than global motion are adjusted. Then the watermark is embedded into visually significant coefficients in proportion to the strength of motion in 3D-DCT domain. Experimental results show that the proposed scheme improves the robustness to MPEG compression and temporal attacks by about 3~9%, compared to the existing 3D-DCT based method. In terms of PSNR, the proposed method is similar to the existing method, but JND guarantees the transparency of watermark.

Keywords : Watermarking, 3D-DCT, JND

접수일 : 2004년 10월 23일 ; 채택일 : 2005년 4월 21일

* 본 연구는 한국과학재단 목적기초연구 지원으로 수행되었음.

[†] 주저자 : hpark@cse.hanyang.ac.kr

[‡] 교신저자 : ysmoon@cse.hanyang.ac.kr

1. 서론

최근 몇 년 동안 디지털 멀티미디어의 발달과 더불어 멀티미디어(audio, image, video)에 대한 사람들의 관심이 증가하고 있다. 또한 인터넷과 같은 통신망의 발전에 따라 디지털 멀티미디어 콘텐츠(contents)가 쉽게 전송이 가능하다. 그러나 멀티미디어 표현물의 상업적 가치가 증가되는 상황에서 각종 멀티미디어 데이터들이 네트워크와 인터넷을 통해 쉽게 분배, 복제되고 있는 실정이다. 디지털 데이터의 불법 복제 및 변조 등에 대한 문제의 발생과 함께 이를 해결하기 위한 연구의 필요성이 발생하게 되었으며 이러한 문제를 해결하기 위해 최근에 디지털 데이터 내부에 정보를 은닉하여 멀티미디어 데이터의 저작권을 보호할 수 있는 연구가 활발히 진행되고 있다^(1,2). 대표적인 예로서 멀티미디어 데이터의 저작권 보호를 위해 디지털 데이터 내부에 개인의 정보를 은닉하여 확인하는 경우가 있는데 이러한 기술을 디지털 워터마킹(digital watermarking)이라고 한다.

워터마킹은 Steganography에서부터 출발한다. Steganography는 인증받지 못한 사용자는 접근할 수 없도록 하는 암호화(Cryptography)와는 달리 숨기고자 하는 대상 자체에 정보를 숨기는 방법이다⁽¹⁾. Steganography는 동서양을 통해서 오래 전부터 연구되어 왔다. 고대 워터마킹의 형태는 주로 중요한 편지나 문서 등에 은밀하게 마킹을 하거나 문서의 내용을 특수한 심벌로 변환하여 인증되지 않은 사람으로부터의 해석이나 변조 등을 막았다. 이러한 기술들은 시각적으로 확인할 수 있는 워터마킹 방법이라고 할 수 있다. 15세기경에는 중요한 그림이나 영상에 시각적으로 확인할 가능한 워터마크를 삽입하는 기술을 통하여 소유권을 인정하거나 변조를 막는 방법이 사용되었다. 고대의 문서와 그림에 대한 워터마킹 기술은 현대의 디지털 워터마킹에 유사하게 적용되고 있다. 예를 들면 소유권을 위해 문서 등에 사용되었던 마크(mark)는 디지털 워터마킹에서 워터마크(watermark)와 같고, 글자의 크기에 변형을 주거나 문서 전체에 걸쳐 마킹(marking)하는 방법은 현재 텍스트 파일 워터마킹에서 그대로 적용되고 있다⁽¹⁾. 고대의 워터마킹 기술은 기본적인 의도와 전체적인 과정에서 유사한 점이 있지만, 시각적으로 인식 가능한 워터마킹 방법이므로 현대의 디지털 정보에 그대로 적용하기에는 적합하지 못하다.

현대의 디지털 워터마킹은 인간의 시각과 청각이 완벽하게 디지털 데이터를 인식하지 못한다는 사실을 이용하게 된다. 즉, 디지털 데이터에 다른 정보를 삽입하여도 인간은 이를 인식하지 못한다. 이러한 특성을 이용하여 시각적으로 인식 불가능한 양의 정보를 인식 불가능한 위치에 삽입하게 된다^(2,3).

최근에는 DVD, 방송용 동영상 등의 디지털 멀티미디어 데이터가 발전하고 있고 이에 따라 비디오를 워터마킹하는 연구가 활발히 진행 중이다. 비디오 워터마킹에서는 정지영상의 워터마킹과는 다르게 많은 데이터를 갖고 있으며 시간적으로 유사한 프레임들의 연속이라는 점을 고려해야 한다. 비디오 워터마킹에서는 이웃한 프레임간의 관계에 의해서 정지영상 워터마킹에서는 나타나지 않는 현상이 있다. 예를 들면 움직임이 없는 영역에 워터마킹을 한다면 비디오를 재생중에 부자연스러운 재생이 일어날 수 있다. 움직임이 많은 곳은 인간의 시각에 덜 민감하다는 점을 고려한다면, HVS(Human Visual System)의 이론에 입각하여 움직임의 크기에 비례하게 워터마크를 삽입해야 한다.⁽⁸⁾

본 논문에서는 JND를 기반으로 한 3차원 지각 모델을 제시하고 전역 움직임에 비해 움직임이 큰 국부영역은 인간의 시각으로 인지하기 어렵다는 것을 이용하며 3D-DCT의 저주파 대역 계수의 크기를 적응적으로 이용하는 방법을 제안하였다. 3D-DCT의 저주파 대역 계수들은 공격에 강건하므로 움직임의 크기에 비례하게 삽입함으로써 워터마크의 강건성과 투명성을 높였고 실험을 통하여 비디오의 특징을 반영하지 않고 3D-DCT 계수를 사용하는 기존의 방법보다 성능이 우수함을 확인하였다⁽⁴⁾.

본 논문은 다음과 같이 구성된다. 2장에서는 본 논문이 제안하는 3D-DCT 계수를 적응적으로 이용하는 비디오 워터마킹 방법을 설명한다. 3장에서는 본 논문에서 제안하는 방법을 이용하여 실험한 결과를 보여주고, 마지막으로 4장에서 결론을 맺고, 앞으로의 향후과제에 대해서 논한다.

II. 제안하는 알고리즘

본 논문에서 제안된 비디오 워터마킹 알고리즘은 크게 워터마크 생성, 워터마크 삽입, 워터마크 검출 과정으로 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 큐브를 만들어 3D-DCT 계수를 생성하고 JND를 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 얻어지는 프레임들에 삽입과정에서 사용했던 3D-DCT를 적용하고 생성된 3D-DCT 계수와 워터마크의 상관관계를 이용하여 워터마크를 검출한다. 그림 1은 워터마킹 시스템의 흐름도를 나타낸다.



그림 1. 워터마킹 시스템 흐름도

2.1절에서는 제안하는 3차원 지각 모델을 설명하고 2.2절에서는 워터마크 생성방법, 2.3절에서는 워터마크 삽입방법, 2.4절에서는 워터마크 검출 방법에 대해서 설명한다.

2.1 3차원 지각 모델(Perceptual Model)

인간 시각 시스템(HVS)을 이해하고 영상처리에 이용하는 것은 수년간 연구 되어왔다. 그 중에서 2차원 시각 모델은 영상 압축에 사용되고 있는데 시각 모델을 이용한 통상적인 압축 방법은 영상의 특징을 고려하는 마스크를 양자화 테이블에 적용하는 것을 기반으로 한다^[5]. 이미지 의존적인 마스크는 JND 이론을 적용하여 만들어진다. 영상 압축에 사용되었던 시각 모델은 디지털 워터마킹으로 확장되어 사용되어지며 워터마크의 투명성과 강인성을 보장하기 위한 워터마크 삽입 강도의 상한선을 제공한다^[6]. 본 논문에서는 기존의 2차원 시각 모델을 확장한 3차원 지각 모델을 제안하고 인간 시각 시스템을 기반으로 하는 워터마킹 알고리즘을 제안한다.

2.1.1 2차원 시각 모델(Visual Model)

Watson은 이미지 의존적인 압축을 위한 DCT 양자화 테이블을 제안하였다^[5]. Watson은 양자화 상

수와 양자화 에러의 관계에서 민감도(frequency Sensitivity)를 정의하고 밝기 마스크(luminance masking)와 대조 마스크(contrast masking) 등을 통하여 이미지에 적응적인 JND 테이블을 만든다.

밝기 마스크를 만들기 위해 민감도가 필요한데 이는 JPEG 이미지의 압축 메카니즘을 이용한다^[5]. 이 방법은 이미지를 8x8 블록으로 나누어 블록마다 DCT를 수행한다. DCT된 8x8 블록의 계수는 64개의 기저 함수의 특성을 나타낸다. 64개의 각 계수는 이미지의 전체적인 특성을 나타내는 저주파 대역과 이미지의 상세한 부분을 나타내는 고주파 대역을 나타내는데 압축을 위해 고주파 대역의 대부분을 0으로 만드는 양자화 과정을 거친다. 양자화 과정은 식 1로 나타낼 수 있고, 양자화 과정에서 발생하는 에러는 식 2로 나타내어진다.

$$u_{ijk} = \text{Round}[c_{ijk} / q_{ij}] \tag{1}$$

$$e_{ijk} = c_{ijk} - u_{ijk}q_{ij} \tag{2}$$

c_{ijk} 는 DCT 계수, q_{ij} 는 양자화 상수, u_{ijk} 는 양자화된 DCT 계수, e_{ijk} 는 양자화 과정에서 발생하는 에러이다. 식 1과 식 2를 통해서 최대 양자화 에러는 식 3과 같이 나타낼 수 있음이 명백하다.

$$\max(e_{ijk}) = q_{ij} / 2 \tag{3}$$

따라서, 모든 에러는 $q_{ij} / 2$ 이하에서 인간의 시각에 invisible하며 이것은 민감도(sensitivity)를 정의할 만한 threshold value 값이 될 수 있다. 이는 식 4와 같이 나타낼 수 있다.

$$t_{ij} = q_{ij} / 2 \tag{4}$$

1.40	1.01	1.16	1.66	2.40	3.43	4.79	5.56
1.01	1.45	1.32	1.52	2.00	2.71	3.67	4.93
1.16	1.32	2.24	2.59	2.98	3.64	4.60	5.88
1.66	1.52	2.59	3.77	4.55	5.30	6.28	7.60
2.40	2.00	2.98	4.55	6.15	7.46	8.71	10.17
3.43	2.71	3.64	5.30	7.46	9.62	11.58	13.51
4.79	3.67	4.60	6.28	8.71	11.58	14.50	17.29
5.56	4.93	5.88	7.60	10.17	13.51	17.29	21.15

그림 2. 민감도(Sensitivity) 테이블

그림 2는 식 4를 이용하여 만든 8x8 블록에 해당하는 민감도 테이블이다⁽⁶⁾.

Watson은 전체 이미지의 밝기에 비해 국부적인 지역의 밝기가 더 밝은 경우 국부지역의 8x8 블록의 DCT 계수에 약간의 변화를 주어도 그 차이를 인지할 수 없다는 사실을 고려하여 민감도 테이블을 수정하였다. 이것을 밝기 마스크링이라고 한다. 또한 다른 주파수 에너지의 존재로 인한 어떤 주파수의 변화에 덜 민감한 정도를 이용하여 대조 마스크링을 한다. 식 5는 전체 이미지의 평균 밝기와 블록의 평균 밝기의 비율에 따라 민감도를 수정하는 식이다.

$$t_{ijk}^L = t_{ij}(C_{00k}/C_{00})^{a_r} \quad (5)$$

C_{00k} 는 k번째 블록의 DC값, C_{00} 는 전체이미지의 평균 DC값, t_{ijk}^L 는 밝기 마스크링을 적용한 민감도이다. a_r 는 상수이고 0.649가 제안되고 있다. Podilchuk의 방법은 Watson의 이미지 의존적인 시각 모델을 디지털 워터마킹으로 확장하였고 워터마크의 투명성과 강인성을 보장하기 위한 워터마크 삽입 강도의 상한선을 제공했다⁽⁷⁾. 본 논문에서는 Podilchuk이 이미지 의존적인 워터마킹을 위해 Watson의 시각 모델을 이용했듯이 비디오 의존적인 워터마킹을 위해 3차원 시각 모델을 제안한다.

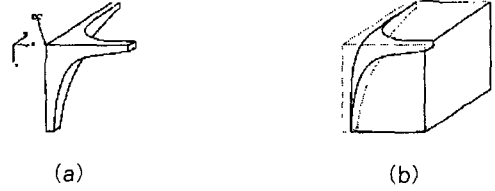
2.1.2 3차원 시각 모델(Perceptual Model)

본 논문에서는 2D 시각 모델을 3D 시각 모델로 확장하기 위해서 chan이 제안한 비디오 압축을 위한 3D-DCT 양자화 방법을 이용한다⁽⁹⁾. chan의 모델에 의하면 3D-DCT 계수 $g(u,v,w)$ 은 식 6과 같이 임계값 C에 의해서 시각적으로 중요한 계수와 시각적으로 중요하지 않은 계수로 나눈다. 쌍곡선의 특징으로 인해 임계값 C가 커지면 3차원 쌍곡선이 큐브의 원점에서 멀어지므로 시각적으로 중요한 계수의 개수가 많아진다. $g_i(u,v,w)$, $g_o(u,v,w)$ 는 각각 내부와 외부 영역을 나타낸다.

$$g(u,v,w) = \begin{cases} g_i(u,v,w), & \text{if } (u+1)(v+1)(w+1) \leq C, \\ g_o(u,v,w), & \text{if } (u+1)(v+1)(w+1) > C. \end{cases} \quad (6)$$

chan은 그림 3의 (a)와 (b)에 속한 계수들의 크기가 지수적으로 감소한다고 하였고 식 7의 지수함

수를 이용해서 그림 3의 (a), (b) 각 부분에 속한 계수의 양자화 상수를 구하였다.



(a) 시각적으로 중요한 계수 (b) 시각적으로 중요하지 않은 계수

그림 3. 3차원 쌍곡선에 의한 3D-DCT 계수의 분류

$$q(u,v,w) = \begin{cases} \text{if } g_i(u,v,w) \\ 255(1 - \frac{e^{-\beta_m(u+1)(v+1)(w+1)}}{e^{-\beta_m}}) + 1; \\ \text{if } g_o(u,v,w) \\ 255(1 - e^{-\beta_{om}(u+1)(v+1)(w+1)}); \end{cases} \quad (7)$$

본 논문에서는 chan의 방법에 의해 구하여진 양자화 상수를 이용하여 민감도(Sensitivity)를 구한다. Watson이 제안한 식 4에 의해서 3D-DCT 계수의 민감도는 식 8가 됨을 알 수 있다.

$$t(u,v,w) = q(u,v,w)/2 \quad (8)$$

비디오는 인접한 프레임들간에 유사성이 있는 연속되는 프레임들로 구성되어 있고 움직임이 존재하는 영역과 움직임이 없는 영역으로 나뉠 수 있다. Watson은 밝기 마스크링을 통해서 전역 밝기값에 비해 밝기값이 큰 국부지역의 DCT 계수의 변화는 인지하기 어렵다고 설명하였는데 본 논문에서는 이를 비디오에 적용하여 전역 움직임에 비해 움직임이 큰 국부지역의 3D-DCT 계수의 적당한 변화는 인지하기 어렵다고 보고 민감도를 수정하였다. 이는 상대적으로 움직임이 많은 블록에 적용적으로 워터마크를 삽입하겠다는 것으로써 본 논문에서는 이것을 움직임 마스크링(Motion masking)이라고 정의하고 Watson의 밝기 마스크링⁽⁵⁾과 유사하게 식을 유도한다. 국부지역의 움직임의 크기는 각 큐브의 움직임의 크기이고 이를 구하기 위해 그림 4의 A 영역을 참조한다. 그림 4의 A영역 즉, 시간축에 집중되어 있는 3D-DCT 계수들은 움직임을 나타내므로 A영역에 속하는 계수들을 이용하면 각 큐브의 움직임을 구할 수 있다.

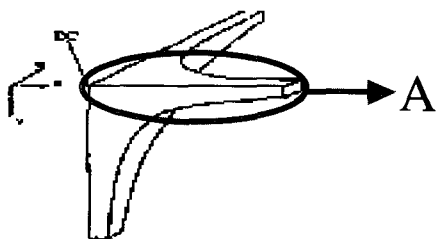


그림 4. 시간축 계수

국부 움직임의 크기는 식 9을 이용해서 A영역에 속한 계수의 크기의 총합으로 구한다. 전역 움직임의 크기는 식 10에 나타난 것 같이 국부 움직임의 크기의 평균을 구해 구한다. 식 11는 전역 움직임과 국부 움직임의 크기에 비례하게 민감도를 수정하는 움직임 마스크를 나타낸다.

$$M_k = \sum X_{uvw_k}, u < 3, v < 3, 0 < w < 8, g(u, v, w) \leq C \quad (9)$$

$$M_{global} = \frac{1}{m} \sum_{k=1}^m M_k, m \text{ 은 cube의 개수} \quad (10)$$

$$t_{(u,v,w)_k}^M = t_{(u,v,w)} (M_k / M_{global})^{\alpha_r} \quad (11)$$

X_{uvw} 는 DCT 계수, M_k 는 k번째 큐브의 움직임의 크기, M_{global} 은 전역 움직임의 크기이며 M_k 의 평균으로 구한다. α_r 는 상수이고 $t_{(u,v,w)_k}^M$ 는 움직임 마스크에 의해 수정된 민감도이다. 워터마킹에서 고려할 점은 워터마크의 강인성과 투명성이다. 본 논문에서는 워터마크의 투명성을 위해서 식 11와 같이 수정된 민감도를 이용해서 워터마크를 삽입하고 워터마크의 강인성을 위해서 저주파 대역에 해당하는 그림 3의 (a) 영역에 워터마크를 삽입한다.

2.2 워터마크의 생성

본 논문에서 사용할 워터마크는 삽입시 시각적으로 비가시성을 가지면서 이진 워터마크(binary watermark)보다 견고한 가우시안 분포 랜덤 벡터를 사용한다. 가우시안 분포를 가지는 랜덤 벡터를 생성하기 위해 식 12의 Box-Muller 변환을 사용하여 균등 분포(uniform distribution)를 갖는 랜덤

벡터 u_{i_1}, u_{i_2} 로부터 가우시안 분포를 갖는 R_i 를 생성하게 된다.

$$R_i = \sin(2\pi u_{i_1}) \sqrt{-2 \log u_{i_2}} \quad (12)$$

생성된 가우시안 랜덤 벡터 \mathbf{R} 은 식 2.13에 의해 변조된다. 변조된 랜덤 벡터 \mathbf{W} 는 민감도에 의해 크기 변화되어 3D-DCT 계수에 삽입되게 된다.

$$\mathbf{W} = \begin{cases} \mathbf{R} \times 1, & \text{if Watermarkbit} = 1 \\ \mathbf{R} \times (-1), & \text{if Watermarkbit} = 0 \end{cases} \quad (13)$$

2.3 워터마크 삽입

인간 시각 시스템을 고려하여 비디오를 워터마킹하기 위해서는 움직임이 있는 영역에 워터마크를 삽입하여야 한다. 비디오는 부피를 갖는 데이터의 집합이고 3차원 신호로 간주할 수 있다. 비디오의 3차원 신호를 3D-DCT하여 얻은 계수는 3차원의 큐브 안에 존재하게 되고 시공간적인 정보를 갖고 있으며 시간적 공격에 강하다⁽⁴⁾.

초기의 워터마킹 방법들은 시각적으로 인식하지 못하도록 하기 위해서 시각적으로 보다 덜 중요한 부분에 워터마크를 삽입하였다. 그러나 시각적으로 덜 중요한 부분에 삽입된 워터마크는 일반적인 압축 방법이나 영상처리에 의해 쉽게 제거될 수 있다. 따라서 강인성을 위해서는 시각적으로 중요한 부분에 워터마크를 삽입하여야 하고 워터마크 삽입과정은 시각적으로 중요한 계수를 선택하는 과정이 포함되어야 한다. 물론 이때 삽입되는 워터마크는 원영상의 화질저하를 가져오지 않는 범위내에서 삽입될 수 있어야 한다. 따라서 제안한 방법에서는 시각적으로 중요한 3D-DCT 계수를 선택하고 화질 저하를 막기 위해 JND를 기반으로 한 3차원 지각 모델을 이용하여 워터마크를 삽입하였다. 그림 5는 전체적인 삽입과정을 나타낸다.

워터마크 삽입의 첫 번째 과정에서는 부호화 작업을 하기 전의 비디오 프레임을 8프레임 기준으로 8x8x8 크기의 큐브로 나누고 각 큐브에 3D-DCT를 수행한다.

식 14는 일반적인 3D-DCT 식인데 k는 큐브의 인덱스이고 $f_k(x, y, z)$ 는 화소의 밝기 값 $F_k(u, v, w)$.

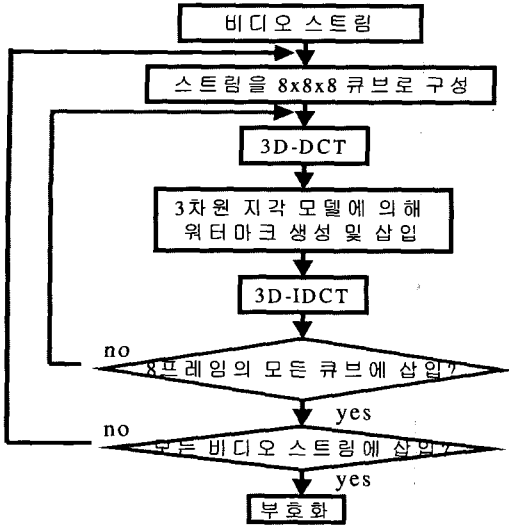


그림 5. 워터마크 삽입 흐름도

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

$$C(u) = \begin{cases} \sqrt{\frac{1}{N}}, & u = 0 \\ \sqrt{\frac{2}{N}}, & \text{otherwise} \end{cases} \quad (14)$$

는 3D-DCT의 계수를 나타낸다. 본 논문에서는 계산의 효율을 위해 식 14대신 3D-DCT의 분리성을 이용하여 x축, y축, z축으로 분리하여 구현하였다.

워터마크 삽입의 두 번째 과정은 3D-DCT 계수의 저주파 대역에 3차원 지각 모델을 적용하여 워터마크를 삽입하는 것이다. 3D-DCT 계수의 저주파 대역은 워터마크의 강인성을 보장하지만 화질저하를 초래할 수 있다. 제안하는 방법은 화질 저하를 방지하기 위해 3차원 지각모델에 근거하여 3D-DCT계수의 민감도를 구하고 움직임 마스킹을 통해 민감도를 수정하여 그림 3의 (a) 영역에 워터마크를 삽입한다.

그림 6을 보면 그림 3의 (a)영역에는 N개의 계수가 있고 비디오의 8프레임마다 M개의 큐브가 있으므로 M개의 원소를 갖는 N개의 그룹을 구할 수 있다. 제안하는 방법은 식 13에 의해 만들어진 가우시안 랜덤 벡터를 이용하여 그룹당 1 비트씩 N개의 그룹에 N개의 비트를 삽입한다. 삽입은 그림 3의 (a)

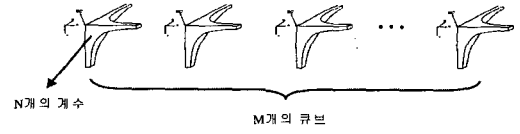


그림 6. 워터마크 삽입 영역

가 나타내는 전 영역에 삽입되는데, 워터마크의 capacity 효율을 높이기 위하여 식 15에서 보는 바와 같이 민감도 보다 큰 3D-DCT 계수는 민감도에 비례하게 삽입하고 그 외의 부분은 가중치를 두지 않고 삽입한다.

$$X_{n,k}^* = \begin{cases} X_{n,k} + \alpha t_{n,k}^M W_k, & \text{if } X_{n,k} > t_{n,k}^M \\ X_{n,k} + W_k, & \text{otherwise} \end{cases} \quad (15)$$

$t_{n,k}^M$ 는 움직임 마스킹으로 수정된 민감도, W_k 는 식 13에 의해 만들어진 워터마크, $X_{n,k}$ 는 그림 3의 (a)영역에 속하는 3D-DCT 계수이며 $X_{n,k}^*$ 는 워터마크가 삽입된 계수이다. α 는 비례상수으로써 초기값은 1이며, PSNR값에 따라 적응적으로 조정해준다. n은 그룹의 인덱스이며 k는 그룹에 속한 원소의 인덱스으로써 큐브의 개수와 같다. 마지막으로 워터마크가 삽입된 8x8x8 큐브를 3D-IDCT(Inverse DCT)한다. 3D-IDCT 식은 식 16과 같다.

$$f_k'(x, y, z) = \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} \sum_{w=0}^{N-1} C(u)C(v)C(w)F_k'(u, v, w) \cdot \cos\left[\frac{(2x+1)u\pi}{2N}\right] \cos\left[\frac{(2y+1)v\pi}{2N}\right] \cos\left[\frac{(2z+1)w\pi}{2N}\right] \quad (16)$$

$F_k'(u, v, w)$ 는 워터마크가 삽입된 DCT 계수이고 $f_k'(x, y, z)$ 는 3D-IDCT 결과로 역 변환된 워터마크가 삽입된 영상의 화소값이다. 첫 번째 과정부터 마지막 과정까지 모든 프레임에 대해서 수행하면서 워터마크를 삽입한다.

2.4 워터마크 검출

워터마크의 검출과정은 삽입과정과 유사하다. 그림 7은 워터마크 검출과정의 전체적인 흐름도이다.

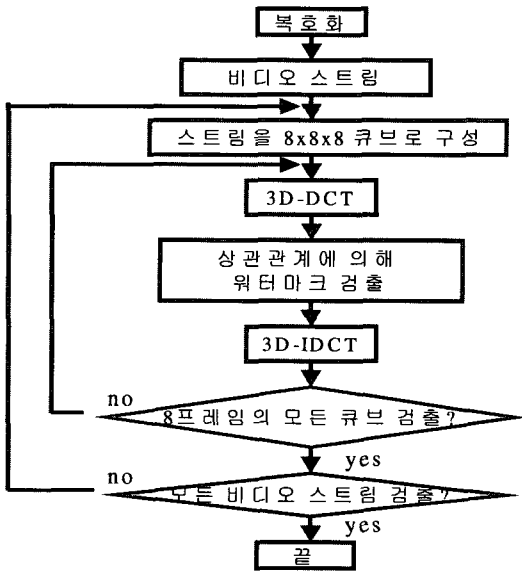


그림 7. 워터마크 검출 흐름도

먼저 압축된 동영상을 복호화하여 워터마크가 삽입된 프레임들을 얻는다. 연속되는 프레임들에 대해서 삽입과정과 동일한 방법으로 8x8x8 블록의 3D-DCT 계수를 얻는다. 식 17을 이용하여 워터마크된 계수 X_k^* 와 워터마크 W 의 정규화된 상관관계(normalized correlation) NC 를 구한다. 그리고 식 18를 통해 NC 의 부호를 검사하여 워터마크 비트를 검출한다.

$$NC = (X_n^*, W) = \sum_{k=0}^{M-1} \bar{X}_{k,n}^* \bar{W}_k \quad \left\{ \begin{array}{l} \bar{X}_{k,n}^* = \frac{X_{k,n}^*}{|X_n^*|} \\ \bar{W}_k = \frac{W_k}{|W|} \end{array} \right. \quad (17)$$

$$watermark\ bit = \begin{cases} 1, & \text{if } NC \geq 0 \\ 0, & \text{if } NC < 0 \end{cases} \quad (18)$$

III. 실험 결과

실험 결과는 비디오 워터마킹의 요구사항에 맞추어 Watermark capacity, Invisibility, Robust 측면에서 시행되었는데 Invisibility와 Robust의 실험 결과는 비디오의 특징을 고려하지 않고 3D-DCT의 계수를 사용하는 기존의 방법⁽⁴⁾과 비교하여 실험하였다. 기존의 방법⁽⁴⁾은 3D-DCT의 중간대역 계수

중에서 key값으로 결정된 계수의 시퀀스를 두 그룹으로 뽑아 계수치를 조정하여 삽입하고 역으로 검출하는 방식이었는데, JND과정이 없이 일괄적으로 삽입하여 비디오의 움직임 특성을 고려한 방법은 아니었다. 그림 8은 실험에 사용한 미스 아메리카 동영상이고 720x584 사이즈의 80프레임을 이용하였다.



그림 8. 미스 아메리카

워터마크의 검출을 평가하기 위해서는 사용된 워터마크 bits의 개수와 검출된 bits의 개수의 비율을 사용하였다.

$$검출율 = \frac{검출된\ watermark\ bits\ 의\ 개수}{삽입된\ watermark\ bits\ 의\ 개수} \times 100 \quad (19)$$

워터마크된 영상의 화질을 평가하기 위해 두 영상간의 차에 대한 측정값을 정량적인 수치로 표현하는 PSNR(Peak Signal to Noise Ratio)⁽⁶⁾을 이용하였다. 아래 식은 PSNR 값을 구하는 식이다.

$$PSNR = 10 \log_{10} \left[\frac{255^2}{MSE} \right] [dB] \quad (20)$$

$$MSE = \frac{1}{N^2} \sum_{x=1}^N \sum_{y=1}^N [f(x,y) - f'(x,y)]^2 \quad (21)$$

PSNR의 단위는 dB이며, $f(x,y)$ 는 원 영상이고, $f'(x,y)$ 는 워터마킹된 영상이다. 50dB은 원 영상을 나타내고 40dB이상이면 화질이 좋은 것으로

평가된다. 성능평가의 방법은 MPEG 압축과 시간적인 공격에 얼마나 강건한 결과를 나타내는가를 사용하였다. 시간적인 공격으로는 공모 공격 (Collusion attack), 프레임 제거 공격(Dropping attack) 등을 사용하였다.

3.1 Capacity Test

Data Capacity는 WMS(Watermarking minimum segment)를 기준으로 측정하였다^[10]. 64 bits/WMS 이상이 되면 유효한 것으로 제시되고 있으며, 비압축 영역에서 103bits/WMS의 Data Capacity 결과를 확인하였다.

3.2 Invisibility Test

MPEG 압축과 시간적인 공격없이 워터마크를 삽입한 후 바로 검출했을때의 검출률과 PSNR을 표 1에 나타내었다. PSNR이 40dB 이상일 경우 화질저하를 인간이 인지할 수 없는데 실험결과 40dB 이상이 나왔으므로 워터마크의 투명성 조건을 만족한다고 볼 수 있다^[11]. 또한 오검출률에 대한 실험결과를 표 2에 나타내었다. 랜덤 시퀀스를 삽입후 검출했을 경우와 워터마크를 삽입하지 않고 검출을 했을 경우 검출률이 약 50% 이하로 감소하는 것을 확인할 수 있었다. 이는 검출률이 약 50%이하로 떨어지면 삽입한 워터마크가 아닌 랜덤한 시퀀스가 삽입되었거나 워터마크가 없다는 것으로 생각 할 수 있다.

표 1. 공격 없을 때의 검출률과 PSNR

	검출 (%)	PSNR (dB)
제안하는 방법	100	44.83
기존의 방법	99.68	44.04

3.3 Robustness Test

3.3.1 압축 공격(Compression Attack)

MPEG 압축률을 5Mbps, 4Mbps, 3Mbps, 2Mbps, 1Mbps 등으로 구분하여 실험하였고 부호화 및 복호화 방법은 MPEG-2 를 사용하였다. 표 3에 실험결과를 나타내었다.

표 2. 각 8프레임당 오검출률과 평균 오검출률

	No attack	랜덤 시퀀스 삽입 후 검출	워터마크를 삽입하지 않고 검출
1~8 frame	100	51.42718	50.45631
9~16 frame	100	52.39806	43.66019
17~24 frame	100	54.33981	49.48544
25~32 frame	100	49.48544	48.51456
33~40 frame	100	49.48544	49.48544
41~48 frame	100	50.45631	47.54369
49~56 frame	100	53.36893	51.42718
57~64 frame	100	42.68932	39.77670
65~72 frame	100	50.45631	43.66019
73~80 frame	100	57.25243	49.48544
평균	100	51.13592	47.34951

표 3. MPEG 압축에 대한 검출률

	제안하는 방법	기존의 방법	성능개선(%)
5Mbps	99.99	98.80	1.20
4Mbps	99.99	98.64	1.37
3Mbps	99.95	98.48	1.50
2Mbps	97.55	96.67	0.91

3.3.2 공모 공격(Collusion Attack)

Collusion Attack은 워터마크를 공격하는 자가 프레임에서 워터마크를 제거하려는 것을 말하는데 영상처리나 기하학적 왜곡과 같은 공격과는 달리 인간이 인지할 수 없도록 프레임간 평균을 내거나 워터마크가 삽입되지 않은 프레임을 기존 프레임과 대체하는 등의 비교적 간단한 공격방법이다. Collusion 공격을 실험하기 위해서 본 논문에서는 연속된 프레임에서 한 개 또는 두 개의 프레임에 삽입된 워터마크를 임의로 제거하고 실험하였다. 표 4에 실험 결과를 나타내었다.

표 4. Collusion 공격에 대한 검출율

	제안하는 방법	기존의 방법	성능개선(%)
5Mbps +1프레임	99.63	94.34	5.61
5Mbps +2프레임	99.63	91.63	8.73

3.3.3 프레임 제거 공격(Dropping Attack)

Dropping은 워터마크를 공격하는 자가 연속된 프레임에서 프레임을 빼내는 것을 말한다. Dropping 공격을 실험하기 위해서 본 논문에서는 연속된 프레임에서 한 개의 프레임을 임의로 제거하고 실험하였다. 표 5에 실험 결과를 나타내었다.

표 5. dropping 공격에 대한 검출율

	제안하는 방법	기존의 방법	성능개선(%)
5Mbps +1프레임	98.12	94.98	3.31

실험 결과 MPEG 압축 및 시간축 공격에 대해서 3차원 지각 모델을 기반으로 계수의 크기에 비례하게 워터마크 정보를 삽입하는 방법이 계수의 크기를 고려하지 않은 기존의 방법보다 성능이 우수함을 알 수 있다.

IV. 결 론

본 논문에서는 JND를 기반으로 한 3차원 지각 모델(Perceptual Model)을 제시하고 비디오의 특성을 고려하여 3D-DCT 계수에 워터마크를 삽입하는 방법을 제안하였다.

본 논문에서는 워터마크의 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 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.

참 고 문 헌

- [1] M.D. Swanson, M. Kobayashi, A. Tewfik, "Multimedia Data-Embedding and Watermarking Technologies," *Proc. of IEEE*, vol. 86, no. 6, June 1988
- [2] M.D. Swanson, B. Zhu, and A. Tewfik, "Data Hiding for Video in Video and Other Applications," *IEEE Int. Conf. on Image Processing*, Oct. 1997
- [3] M.D. Swanson, B. Zhu, and A. Tewfik, "Robust Image Watermarking Using Perceptual Masking," *IEEE Int. Conf. on Image Processing*, Oct. 1996
- [4] 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
- [5] 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.
- [6] I. J. Cox, M. L. Miller and J. A. Bloom, "Digital Watermarking", *Morgan Kaufmann Publishers*, 2001
- [7] Podilchuk, C.I., Wenjun Zeng, "Image-adaptive watermarking using visual models," *Selected Areas in Communications, IEEE Journal vol.16*, pp. 525-539, May 1998
- [8] Raymond B. Wolfgang, Christine I. Podilchuk, Edward J. Delp, "Perceptual Watermarks for Digital Images and Video," *Proc. IEEE*, vol. 87 pp. 1108-1126, July. 1999.
- [9] 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.

- [10] Serdean, C.V. Ambroze, M.A. Tomlinson, M. Wade, J.G. "DWT-based high-capacity blind video watermarking, invariant to geometrical attacks," *Vision, Image and Signal Processing, IEE Proceedings*, vol. 150, pp.51-58, Feb. 2003
- [11] Gou-Min Wu, Yue-Ting Zhuang, Fei Wu, Yun-He Pan, "Self-Adaptive MPEG

Video Watermarking Based On Block Perceptual Features", *Proceedings of the Third International Conference on Machine Learning and Cybernetics*, pp.3871-3876, August 2004

- [12] 주낙근, 이재현, 김동서, "웨이블릿 계수 교환을 이용한 워터마킹 기법", *정보보호학회 논문지*, 13권 5호 pp.49-56, 2003
- [13] 최순영, 서영호, 유지상, 김대경, 김동욱, "DWT 기반 영상 압축기의 다해상도의 통계적 특성을 이용한 실시간 워터마킹 알고리즘", *정보보호학회 논문지*, 13권 6호 pp. 33-43, 2003

〈著者紹介〉



박 현 (Hyun Park)

1998년 2월: 원광대학교 전기공학과 졸업 (학사)
 2000년 8월: 한양대학교 전자계산학과 졸업 (석사)
 2001년 8월~현재: 한양대학교 컴퓨터공학과 박사과정
 2000년~2001년: (주)두루넷 기술연구소 소프트웨어팀
 <관심분야> 영상보안, 영상복원, 패턴인식 등



한 지 석 (Ji-Seok Han)

2002년 2월: 한양대학교 금속재료공학과 졸업 (학사)
 2004년 2월: 한양대학교 컴퓨터공학과 졸업 (석사)
 2004년~현재: LG전자 정보통신사업본부 단말연구소 연구원
 <관심분야> 영상보안, 영상복원, 패턴인식 등



문 영 식 (Young-Shik Moon)

1980년 2월: 서울대학교 공과대학 전자공학과 졸업 (학사)
 1982년 2월: 한국과학기술원 전기 및 전자공학과 졸업 (석사)
 1990년 3월: University of California at Irvine Dept. of Electrical and Computer Engineering. (박사)
 1982년~1985년: 한국전자통신연구소 연구원
 1989년~1990년: Inno Vision Medical 선임연구원
 1990년~1992년: 생산기술연구원 선임연구원
 1992년~현재: 한양대학교 컴퓨터공학과 정교수
 <관심분야> 영상보안, 영상처리, 컴퓨터비전, 패턴인식 등