

DWT 기반 H.263에 적합한 적응 양자화기에 관한 연구

(A Study on the Adaptive Quantizer Appropriate to DWT-based H.263)

沈優星*, 李堉植*, 鄭相淵*, 許道根*

(Woo-Sung Sim, Yong-Sik Lee, Sang-Yun Jung, and Do-Geun Huh)

요 약

동영상 표준안인 H.263의 성능평가를 위해 사용되고 있는 TMN5는 변환부호화방법으로 DCT를 기본으로 사용하고 있어 블록킹 현상의 문제점을 갖고 있다. 따라서 본 논문에서는 TMN5의 구조를 그대로 유지하면서 DCT를 DWT로 대체하고 이 구조에 적용하는 양자화기를 설계하였다. 실험의 결과로 평균 PSNR 측면에서 휘도 신호는 평균 0.35dB 정도의 손실을 보였으나 색차 신호는 평균 2dB 정도 개선되었다. 주관적 화질평가에서도 블록킹 현상이 감소되므로 TMN5에 비하여 우수한 화질의 영상을 얻을 수 있다. 또한 같은 전송속도에서 더 많은 프레임은 전송함으로 화면 드롭 현상을 줄일 수 있다.

Abstract

TMN5 has been used to evaluate the performance of H.263, the ITU-T/LBC standard of moving picture coding. As DCT is used as transform coding methods in TMN5, blocking effect exists seriously in reconstruction picture at low bit rate.

In this paper, DWT is substituted for DCT and adaptive quantizer is designed while the structure of TMN5 is maintained to be excellent in the transplantation of conventional system.

It has shown that average PSNR is reduced by 0.35dB in luminance but increased by 2dB in chrominance comparing with that of TMN5. As block effect is reduced in reconstruction picture, the subjective picture quality is better than that of TMN5. Picture drop is, also, reduced because much more frames are transmitted at same bit rate.

I. 서 론

정보화 사회로의 변화가 급속하게 전개되고 각종 신호처리를 위한 통신 기술이 빠르게 발전함에 따라 문자, 음성, 정지영상뿐만 아니라 동영상까지를 포함하는 다양한 멀티미디어 통신 서비스가 국내외적으로 상용화 되어가고 멀티미디어 관련 기술의 상업적 이용이

커다란 관심을 불러일으키고 있다. 이러한 시대적 흐름에 맞춰 64Kbps 이하의 전송속도를 갖는 초저속 동영상 부호화에 관한 연구 결과로 ITU-T/LBC 그룹에서 H.263이라는 국제 표준이 제정되었다^{[1][2]}.

공중전화망이나 이동통신망과 같은 초저속망에서 동영상 전송을 목적으로 하는 H.263은 일반적인 동영상에 존재하는 3가지 중복성, 즉 신호 성분들간에 존재하는 중복성, 프레임간에 존재하는 시간적 중복성, 그리고 화면 내에 존재하는 공간적 중복성을 효과적으로 제거함으로써 고압축을 수행하는 하이브리드 방식의 알고리즘이다^{[1][2][3]}. 신호 성분간에 존재하는 중복성은 휘도 신호와 색차 신호의 비율을 2:1로 정의함으로써 제거되

* 正會員, 圓光大學校 電子工學科

(Department of Electronic Engineering Wonkwang University)

接受日字:1999年5月17日, 수정완료일:1999年10月21日

고, 화면간에 존재하는 시간적 중복성은 움직임 추정/보상(Motion Estimation/Compensation : ME/MC) 기법에 의해 제거되며, 화면 내에 존재하는 공간적 중복성은 변환부호화와 양자화를 통해서 제거하고 있다.

H.263 권고안은 기본적인 비트열의 구조만을 정의하였을 뿐 ME/MC, 변환부호화 및 양자화기의 세부적이며 구체적인 알고리즘을 정의하고 있지 않다^[5]. H.263 권고안의 이러한 특징은 부호화 알고리즘에 어느 정도의 융통성을 제공함으로써 H.263 권고안을 만족하는 다양한 부호화기의 설계를 가능하도록 하고 있다. 이러한 관점에서 H.263의 성능평가를 위해서 사용된 TMN5(Test Model Near-term 5)와 여러 국제표준화 기관에서 변환 부호화 방법으로 DCT를 채택하고 있다. 하지만 DCT는 높은 비트율의 환경에서는 좋은 성능을 나타내지만 고압축 또는 초저속 동영상 부호화에선 심한 블록킹 현상으로 인한 급격한 화질 열화를 가져오는 단점이 있다.

본 논문에서는 이러한 DCT 변환부호화의 문제점을 개선하고자 TMN5 구조를 갖는 DWT 기반의 H.263과 이에 적합한 적응 양자화기를 제안한다. 제안한 DWT 기반의 H.263은 DCT를 DWT로 대체하여 블록단위의 3레벨 DWT 변환을 수행하고 DWT 변환계수의 특성과 인간 시각적 특성을 고려한 새로운 양자화기를 설계한다.

본 논문의 구성은 아래와 같다. 제 II장에서는 제안한 DWT 기반 H.263의 기본 구조와 DWT 변환부호화에 대해서 설명하고, 제 III장에서는 DWT 기반 H.263에 적합한 적응 양자화기를 제안한다. 제 IV장에서는 제안한 DWT 기반 H.263과 TMN5의 모의 실험을 통한 성능을 평가하고 제 V장에서 본 논문의 결론을 맺는다.

II. DWT 기반의 H.263

1. DWT 기반 H.263의 전체 구성

최근 H.263의 성능을 향상시키고자 DWT 기반 H.263 부호화기가 많이 제안되고 있으나 이들은 블록 기반 부호화로 인한 블록킹 현상을 줄이고자 입력된 영상 전체를 DWT하여 부호화 한다^{[1][12]}. 이와 같은 시스템은 기존의 DCT와 같은 블록 단위의 변화부호화를 채택하고 있는 시스템과의 호환성이 문제가 된다.

따라서 본 논문에서 제안한 DWT 기반 H.263은 이미 개발된 기존의 DCT 기반 시스템과의 이식성을 고려하여 부호화 과정 및 비트열 구조를 TMN5의 방식 그대로 유지[1][4][5] 하면서 DWT 변환 계수 특성과 인간 시각적 특성을 고려한 양자화기를 제안한다. 제안한 DWT 기반 H.263의 부호화기는 그림 1과 같다.

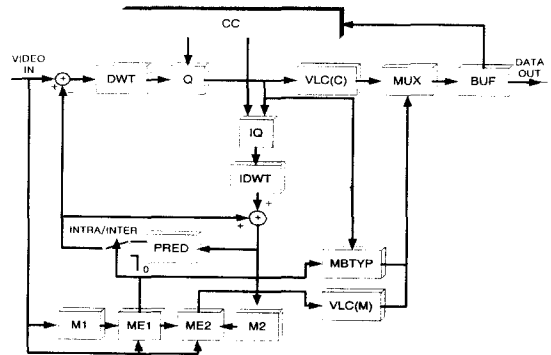


그림 1. DWT 기반의 H.263 부호화기
Fig. 1. DWT-based H.263 coder.

DWT 기반 H.263의 부호화 과정은 먼저 입력되는 현재 프레임과 M1에 기억되어 있는 이전 프레임의 원 영상을 MB(Macro Block : 16×16)단위로 정합시켜 정수단위의 움직임 벡터(Motion Vector : MV)를 ME1에서 추정한다. 두 영상을 정합시키는 정합 기준은 SAD(Sum Absolute Difference), MAD(Mean Absolute Difference), MSD(Mean Square Difference)등이 있으나 연산량이 적고 하드웨어 구현이 용이한 SAD를 채택한다. MB의 크기를 N×N, 현재 프레임의 MB를 S_k(x, y)라고 하고, 현재 프레임의 MB를 기준으로 -N≤m<N, -N≤n<N 인 탐색 영역 내에서 이전 프레임의 MB를 S_{k-1}(x+m, y+n)라 할 때 각 (m, n)에 대한 SAD는 식 (1)과 같이 정의된다.

$$SAD(m, n) = \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} |S_k(x, y) - S_{k-1}(x+m, y+n)| \quad (1)$$

탐색 영역 내의 m, n에 대한 최소 SAD값이 어떤 임계값 이상이면 현재 프레임의 MB를 INTRA 모드로 분류하고 임계값 미만이면 움직임 벡터 (m, n)을 갖는 INTER 모드로 분류한다. 현재 프레임의 MB가 INTRA 모드이면 ME/MC없이 MB를 직접 DWT하고

INTER 모드이면 M2에 저장되어 있는 재생된 이전 프레임의 기준으로 ME2에서 반 화소 단위로 MV를 추정한다. 반 화소 단위의 MV에 대응하는 재생된 이전 프레임의 MB를 PRED에서 보상하고 이 보상된 MB와 현재 프레임 MB의 차이를 DWT한다. DWT계수를 본문에서 제시한 양자화기로 양자화하여 가변장 부호화(VLC(c))한다. 또한 현재 프레임 MB의 모드 형태(MBTYP)와 반화소 단위 MV를 가변장 부호화(VLC(M))하여 이들 가변장 부호들을 다중화 하여 비트열을 발생시킨다. 재생된 영상을 프레임 메모리 M2에 저장하기 위해 역 양자화(Inverse Quantization : IQ)와 역 DWT(Inverse DWT : IDWT) 블록이 필요하다.

2. DWT 변환 부호화

웨이블릿 변환을 이용하여 영상을 다해상도로 분해한 후 영상을 압축하는 방법은 기존의 DCT를 이용한 방법보다 높은 압축율에서 블록킹 현상이 일어나지 않아 인간 시각 특성(Human Visual System : HVS)에 보다 적합한 압축 방법으로 평가되면서 차세대 압축 알고리즘으로 주목받고 있다. 또한 여러 국제 표준기관에서 이의 성능을 향상시키기 위한 많은 연구가 이루어지고 있다^{[6][7][10]}.

웨이블릿 변환은 Fourier 변환과 같이 직교특성을 만족하는 기저함수들의 집합으로 신호를 분해하는 방법 중 하나이다. Fourier 변환은 주파수영역에서만 신호 해석이 가능한 반면 웨이블릿 변환은 시간영역과 주파수영역에서 신호를 해석할 수 있는 변환 방법이다^{[8][9][13]}. 모 웨이블릿(mother wavelet) $\psi(t)$ 를 확장, 이동시켜 식 (2)와 같이 웨이블릿 함수 $\psi_{a,b}(t)$ 를 얻을 수 있다.

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right) \tag{2}$$

여기서 a 는 주파수영역에서의 신호 확장을 나타내는 스케일링변수이고 b 는 시간영역에서의 이동을 나타내는 이동변수이다. 이때 $\psi_{a,b}(t)$ 는 $t=b$ 주변에 집중 분포하게 된다. 임의의 연속된 신호 $f(t)$ 에 대한 연속 웨이블릿 변환(Continues Wavelet Transform : CWT)은 식 (3)으로 정의된다.

$$CWT[f_{a,b}(t)] = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} f(t) \psi\left(\frac{t-b}{a}\right) dt \tag{3}$$

여기서 스케일링변수 a 를 변화시켜 웨이블릿의 위치를 시간영역과 주파수영역에서 변화시킬 수 있으며 이동변수 b 를 변화시켜 시간영역의 중심점을 변화시킬 수 있다. CWT를 이산 웨이블릿 변환(Discrete Wavelet Transform : DWT)으로 표현하기 위해 $j, k \in Z$ (Z 는 정수공간)에 대하여 스케일링 변수 $a = a_0^j$, 이동변수 $b = ka_0^j b_0$ ($a_0 > 1, b_0 \neq 0$)로 한다. 특히 $a_0 = 2, b_0 = 1$ 일 때 기저 함수들은 정규직교 기저가 되고 식 (4)과 같이 표현된다.

$$\psi_{j,k}(t) = 2^{-j/2} \psi(2^{-j}t - k) \tag{4}$$

식 (4)를 이용하여 DWT를 정의하면 식 (5)와 같다.

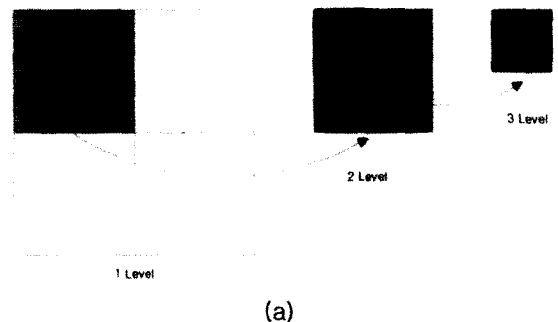
$$DWT[f_{j,k}(t)] = \int_{-\infty}^{\infty} f(t) \psi_{j,k}(t) dt \tag{5}$$

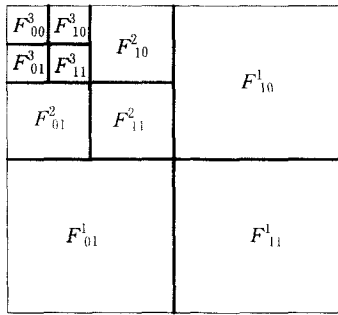
본 논문에서는 ME/MC의 기본단위가 되는 크기 $N \times N$ 인 MB를 다시 크기 $M \times M$ 블록으로 나누어 각 블록에 DWT를 수행한다. 차수 M 인 저역 필터 $h(n)$ 과 고역 필터 $g(n)$ 를 이용해 데이터 블록 $f(i, j)$ 를 수평방향으로 필터링 하여 다운 샘플링 하고, 이 결과를 수직방향으로 다시 필터링 하여 다운 샘플링 하는 식 (6)과 (7)을 이용해 데이터 블록을 DWT한다^[8].

$$f_L(i, j) = (h * f)(2k) = \sum_{n=0}^{M-1} h(n) f(2k - n) \tag{6}$$

$$f_H(i, j) = (g * f)(2k) = \sum_{n=0}^{M-1} g(n) f(2k - n) \tag{7}$$

식 (6)과 식 (7)을 이용해 DWT를 하면 하나의 최저 주파수대역과 3개의 고주파대역으로 나누어지며 더 세분화된 분해를 위해서 최저 주파수대역에 대해 DWT를 반복 수행한다. 이런 방법으로 블록을 3레벨 DWT





(b)

그림 2. 3레벨 DWT : (a) 3레벨 DWT 과정 (b) 3레벨 DWT의 부대역
 Fig. 2. 3 level DWT : (a) Process of 3 level DWT (b) The subband of 3 level DWT.

하면 10개의 부대역으로 분해할 수 있다. 그림 2(a)는 $M \times M = 8 \times 8$ 블록에 대한 3레벨 DWT 과정을 나타내고 그림 2(b)는 3레벨 DWT의 부대역을 나타내는데 F_{rs}^n 은 n 레벨의 rs 부대역임을 의미한다.

III. DWT 기반 H.263에 적합한 양자화기 설계

DCT 기반의 TMN5 부호화기는 INTRADC 계수를 위한 1개의 양자화기와 Non-INTRADC 계수를 위한 31개의 양자화기를 가지고 있다^{[1][5]}. INTRADC 계수는 양자화 간격이 8인 균일 양자화기로 양자화 되고 Non-INTRADC 계수는 양자화 간격이 2에서 64까지의 짝수 값을 갖는 31개의 균일 양자화기로 양자화 된다. TMN5에서는 MB 계층의 양자화 간격을 결정하는 QP (Quantization Parameter)를 변화시켜 버퍼에 유입되는 비트열을 제어한다. 따라서 QP 는 부호화된 비트열의 발생률에 따라 1에서 31까지의 값을 갖고 양자화 간격은 $2QP$ 가 된다. 그리고 INTRA 프레임을 양자화하기 위한 초기 QP 값은 16으로 한다. 변환계수 $C(i, j)$ 를 양자화 계수 $QC(i, j)$ 로 양자화할 때 INTRA 블록의 DC 계수는 식 (8)을, INTRA 블록의 AC계수는 식 (9)를 이용하여 양자화 한다. 또한 INTER 블록의 DC계수와 AC계수는 식 (10)에 의해 양자화 된다. 양자화된 계수를 지그재그로 주사하면서 가변장으로 부호화하여 전송한다.

$$QC(0, 0) = C(0, 0) // 8 \quad \text{for INTRA DC} \quad (8)$$

$$QC(i, j) = \text{sign}[C(i, j)] \times [|C(i, j)| / 2QP]$$

$$\text{for INTRA AC} \quad (9)$$

$$QC(i, j) = \text{sign}[C(i, j)] \times [|C(i, j)| - QP/2] / 2QP$$

$$\text{for INTER DC \& AC} \quad (10)$$

DWT 변환에서 고주파 대역보다 저주파 대역에 에너지가 집중되는 특성은 DCT와 유사하지만 같은 대역에서의 에너지 분포는 한 부분으로 집중하지 않고 원영상의 에지 성분에 따라 고루 분포한다. 이러한 DWT 계수특성 때문에 TMN5에서 이용했던 양자화기를 사용하면 심한 영상의 열화를 가져온다. 정지영상의 압축시 DCT를 DWT로 대체하고 DWT 변환계수를 양자화하기 위한 방법^[6]이 제안되었지만 이 방법을 영상전환 및 영상회의 등과 같은 동영상 응용에 적용할 경우 버퍼 상태의 고려 없이 동일한 양자화기를 사용하므로 많은 비트열을 발생시켜 프레임율이 현저하게 저하된다. 따라서 본 논문에서는 이러한 DWT 계수 특성과 인간 시각 특성을 고려하여 동영상을 부호화하기 위한 새로운 양자화기 행렬 H 의 각 원소 h_{ij} 를 식(11)과 같이 정의한다.

$$h_{ij} = \frac{A}{g_{ij}} \quad (11)$$

$$g_{ij} = g \left(\frac{\sqrt{i^2 + j^2}}{M} \right) \quad (12)$$

$$g(x) = (a_0x + a_1x + a_2x^2 + a_3x^3) e^{(a_4x + a_5x^2)} \quad (13)$$

여기서, $a_0 = 0.79406835$ $a_1 = -1.6394898$
 $a_2 = 0.61415342$ $a_3 = 0.46967190$
 $a_4 = 4.2767498$ $a_5 = -4.8919091$

식 (13)의 $g(x)$ 는 인간 시각적 특성을 반영하기 위한 JPEG 양자화기 Q_{JPEG} 의 원소를 일반화한 함수로 a_n 은 인간 시각적 특징을 잘 반영하는 계수이다^[6]. 식(12)에서 DWT 블록 크기가 $M \times M$ 일 때 $0 \leq i \leq M-1$, $0 \leq j \leq M-1$ 인 정수 i, j 에 대하여 $x = \sqrt{i^2 + j^2} / M$ 을 식 (13)에 대입한 값을 g_{ij} 라 한다. 각각의 g_{ij} 를 식 (11)에 대입하여 새로운 양자화기 행렬 H 의 i 행 j 열 원소 h_{ij} 를 구한다. 식 (11)의 A 는 양자화기의 이득으로 TMN5의 양자화 계수 QP 에 대응된다. $A = 12$, $M = 8$ 일 때 새로운 양자화기 행렬 H 는 그림 3(a)의 JPEG 양자화기 행렬 Q_{JPEG} 과 가장 유사하다. 이때의 H 를 H_{JPEG} 이라 하며 H_{JPEG} 의 각 원소 $(h_{ij})_{JPEG}$ 은

식 (14)를 이용해 구할 수 있으며 그 결과는 그림 3(b)와 같다.

$$(h_{ij})_{JPEG} = \frac{12}{g_{ij}} \quad (14)$$

DWT 계수는 부대역에 따라 고르게 분포하므로 각 부대역에 속하는 DWT 계수는 동일 양자화 간격으로 균일 양자화 해야 한다. 따라서 양자화기 행렬 H 를 DWT의 부대역으로 나누고 각 부대역에 대응하는 H 의 원소를 평균하여 그 부대역의 양자화 간격으로 한다. 그러므로 그림 2에서 F_{rs}^n 부대역의 양자화 간격 h_{rs}^n 를 식 (15)와 같이 정의한다.

$$h_{rs}^n = \frac{2^{n+1}}{M} \sum_{i=\frac{rM}{2}}^{(r+1)M-1} \sum_{j=\frac{sM}{2}}^{(s+1)M-1} h_{ij} \quad (15)$$

식 (15)에서 r 과 s 는 $r=s=0$ 를 제외한 0 또는 1의 값을 가진다. 또한 레벨 n 은 1에서 $\log_2 M$ 까지의 값을 가지며 $n = \log_2 M$ 이면 식(15)를 식(16)과 같이 정의한다.

$$h_{rs}^n = \sqrt{2} h_{ij} \quad (16)$$

식 (15)와 식 (16)에서 정의된 F_{rs}^n 부대역의 양자화 간격 h_{rs}^n 을 이용하여 DWT 기반 H.263의 양자화기 행렬 $H_{H.263}$ 을 얻는다.

$$H_{H.263} = [h_{rs}^n] \quad (17)$$

$A=12, M=8$ 일 때 식(15), 식(16), 식(17)을 이용하여 구한 DWT기반 H.263의 양자화기 행렬 $H_{H.263}$ 은 그림 3(c)와 같다.

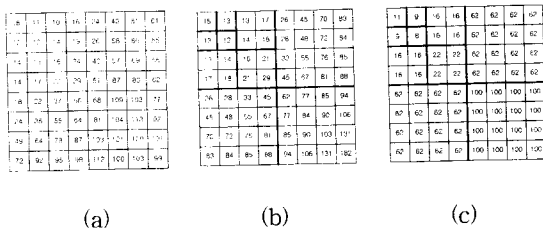


그림 3. 양자화 행렬 :
 (a) Q_{JPEG} (b) H_{JPEG} (c) $H_{H.263}$
 Fig. 3. Quantization matrix :
 (a) Q_{JPEG} (b) H_{JPEG} (c) $H_{H.263}$

TMN5는 전송속도에 맞게 양자화 파라메타 QP 를 변화시켜 버퍼를 제어하며 QP 는 1~31까지의 값을 갖는다. 따라서 양자화 이득 A 는 QP 에 대응되므로 A 도 1~31까지의 값을 갖는다. 이렇게 설계된 양자화기 행렬 $H_{H.263}$ 를 이용하여 식 (18)과 같이 모든 DWT 계수 (INTRADC, Non-INTRADC) $C(i, j)$ 를 양자화 한다.

$$QC(i, j) = \text{sign}[C(i, j)] \times [|C(i, j)| / H_{H.263}(i, j)] \quad (18)$$

초기 INTRA 프레임의 양자화를 위한 A 의 초기값은 12로 한다. $H_{H.263}$ 양자화기에 의해 양자화된 계수 $QC(i, j)$ 는 DCT와 분포 특성을 달리하므로 대역별 주사순서는 그림 4(a)와 같이 10개의 대역을 지그재그 주사하지만 각 대역 내에서의 계수별 주사순서는 그림 4(b)와 같이 주사하여 전송할 때 0의 값을 갖는 양자화 계수가 더욱 연속하게되어 RUN, LEVEL, LAST값만을 이용하는 가변장 부호화시 더욱 높은 압축율을 가져올 수 있다^[6].

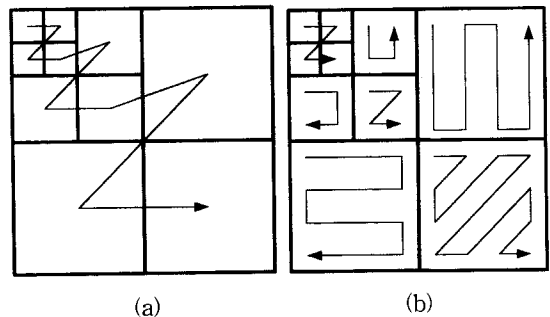


그림 4. DWT 계수의 주사순서 :
 (a) 대역별 주사순서 (b) 계수별 주사순서
 Fig. 4. The scanning method of DWT coefficient :
 (a) Each subband of scanning method
 (b) Each coefficient of scanning method

IV. 실험 결과 및 고찰

본 논문에서 제안한 DWT 기반 H.263의 성능을 분석하기 위해서 ITU-T에서 H.263의 평가를 목적으로 사용한 3가지 QCIF(144×176)영상을 이용한다. 사용된 영상은 주로 영상전환이나 영상회의 시스템에 적합한 영상으로 한 명의 화자가 화면의 중심에 위치하는 Susie, Carphone, Foreman 영상의 각 100 프레임을 가지고

실험한다. 이들 영상은 움직임의 크기가 조금씩 차이가 있는데 Susie 영상은 다른 영상에 비하여 비교적 움직임이 적고 Foreman 영상은 다른 영상에 비해 다소 많은 움직임을 갖는다. 객관적인 평가를 위해서 움직임 추정은 ± 15 의 탐색 영역에서 모든 후보 블록을 순차적으로 비교하여 가장 유사한 블록을 찾는 전역 탐색 기법을 사용하며 최적의 정합된 블록을 찾는 정합 기준으로는 SAD를 사용한다. 또한 재생된 영상의 객관적인 화질을 평가하기 위해서 원 영상과의 PSNR을 비교한다.

표 1과 표 2는 33Kbps와 24Kbps 환경에서 TMN5와 제안한 방식의 성능을 보인다. 평균 PSNR은 휘도신호(Y)와 색차신호(Cb, Cr)에 대하여 한 장의 INTRA 프레임을 포함한 전송된 프레임 각각의 PSNR을 평균한 것이다. 휘도 신호는 제안된 방식이 각 영상별로 약 0.35dB정도 다소 낮은 PSNR을 나타냈으나 색차 신호는 1~2.7dB 정도의 개선을 보였고 움직임이 많은 Carphone 영상과 Foreman 영상에서 한 프레임을 더 전송할 수 있어 TMN5보다 높은 프레임율을 나타낸다.

표 1. 33Kbps에서의 성능

Table 1. Performance on 33Kbps.

33Kbps		DCT-based H.263			DWT-based H.263		
QCIF Sequence		Susie	Carphone	Foreman	Susie	Carphone	Foreman
Average PSNR	Y	35.0668	31.7385	29.6223	34.6968	31.2353	29.3945
	Cb	41.6132	37.0662	35.4043	44.4207	38.0456	37.9587
	Cr	41.3371	37.1374	36.6531	44.0362	38.5708	38.6403
Coded Frame Number		18	14	10	18	15	11
Frame Rate (Frame/sec)		5.4	4.2	3	5.4	4.5	3.3

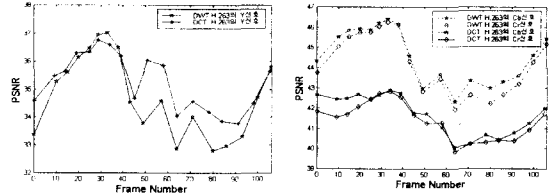
표 2. 24Kbps에서의 성능

Table 2. Performance on 24Kbps.

24Kbps		DCT-based H.263			DWT-based H.263		
QCIF Sequence		Susie	Carphone	Foreman	Susie	Carphone	Foreman
Average PSNR	Y	34.1479	31.1025	29.2531	33.8574	30.7046	28.9892
	Cb	40.9197	36.9042	35.3079	43.7742	37.6333	36.3891
	Cr	40.6146	36.9071	36.6436	43.4349	38.0570	37.6093
Coded Frame Number		14	11	7	14	12	8
Frame Rate (Frame/sec)		4.2	3.3	2.1	4.2	3.6	2.4

Susie 영상에 대한 33Kbps 환경에서 전송된 프레임 별 휘도 신호의 PSNR 변화를 그림 5(a)에 나타내고 색차 신호의 PSNR 변화를 그림 5(b)에 나타내었다. 휘

도 신호의 경우 움직임이 비교적 적은 40 프레임 이전에서는 TMN5보다도 높은 PSNR을 보이지만 움직임이 많은 40 프레임 이후에서는 제안된 방법의 PSNR이 감소한다. 색차 신호의 경우 움직임의 정도에 관계없이 제안된 방법이 항상 높은 PSNR을 보인다.

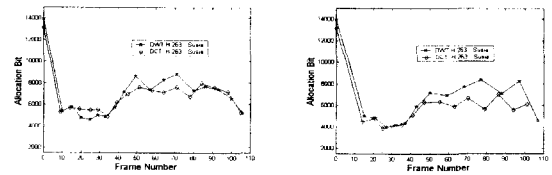


(a) (b)

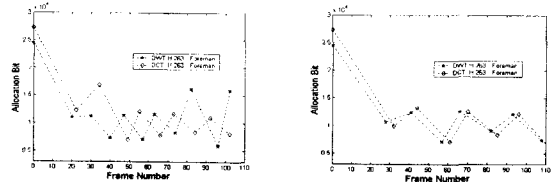
그림 5. Susie 영상에 대한 33Kbps에서 각 프레임에 따른 PSNR : (a) Y 신호에 대한 PSNR (b) Cb, Cr 신호에 대한 PSNR

Fig. 5. PSNR resulted in each frame on 33Kbps of Susie : (a) PSNR on Y signal (b) PSNR on Cb & Cr signals.

그림 6은 33Kbps와 24Kbps 환경에서 Susie 영상과 Foreman 영상의 각 프레임에 할당되는 비트 수를 나타내고 있다. 제안된 방법이 평균적으로 할당되는 비트수가 적어 높은 압축율을 보여 결과적으로 영상의 전송 프레임율을 높인다.



(a) (b)



(c) (d)

그림 6. Susie 와 Foreman영상의 각 프레임에 따른 비트 할당 : (a) 33Kbps에서 비트 할당(Susie) (b) 24Kbps에서 비트 할당(Susie) (c) 33Kbps에서 비트 할당(Foreman) (d) 24Kbps에서 비트 할당(Foreman)

Fig. 6. Bit allocation resulted in each frame of Susie and Foreman : (a) Bit allocation of 33Kbps(Susie) (b) Bit allocation of 24Kbps (Susie) (c) 33Kbps에서 비트 할당(Foreman) (d) 24Kbps에서 비트 할당(Foreman).

공간 중복성만을 이용하여 부호화하는 INTRA 프레임은 전송속도에 무관하게 항상 일정 양의 비트를 발생시키므로 이는 전체 부호화 성능을 좌우할 수 있다. 표 3은 INTRA 프레임, 즉 각 영상의 1번 프레임을 전송 처리한 결과를 나타낸다. 할당된 비트를 TMN5와 비교하면 Susie 영상에서는 제안된 방법이 약 800비트가 더 할당되었지만 Carphone 영상과 Foreman 영상의 경우는 약 4000비트가 적게 할당되어 높은 압축율을 보인다.

표 3. Intra Frame에 대한 성능
Table 3. Performance on intra frame.

Intra Frame	DCT-based H.263			DWT-based H.263			
	Susie	Carphone	Foreman	Susie	Carphone	Foreman	
Allocation Bit	13,147	24,744	27,302	13,935	20,901	24,439	
PSNR	Y	34.5011	32.8839	32.0742	34.3977	32.2647	31.9672
	Cb	42.6482	37.8357	36.7307	44.3294	37.4109	37.4645
	Cr	41.8293	38.3593	38.9311	43.7741	38.5562	38.8437

그림 7은 INTRA 프레임의 화질을 보인 것으로 DWT 기반 H.263의 화질이 DCT 기반 H.263보다 에지 영역에서 우수함을 보인다. 그림 8은 전송 속도가 33Kbps 일 때 Foreman 영상 22번과 63번의 INTER 프레임 화질을 보인 것으로 DWT 기반 H.263이 DCT 기반 H.263보다 블록킹 현상이 줄어들음을 알 수 있다. 64Kbps, 33Kbps, 24Kbps 환경에서 제공된 영상이외의 여러 영상에 대하여 실험한 결과도 DWT 기반 H.263의 영상이 INTRA 프레임과 INTER 프레임 모두에서 블록킹 현상이 어느 정도 제거되어 전체적으로 TMN5 보다 우수한 화질을 보인다.

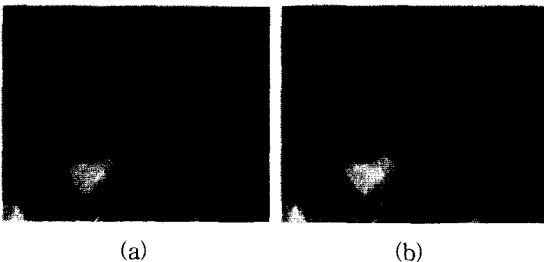


그림 7. INTRA 프레임의 화질 :
(a) DCT-based H.263 (b) DWT-based H.263
Fig. 7. Subjective quality of image in Intra frame :
(a) DCT-based H.263 (b) DWT-based H.263.

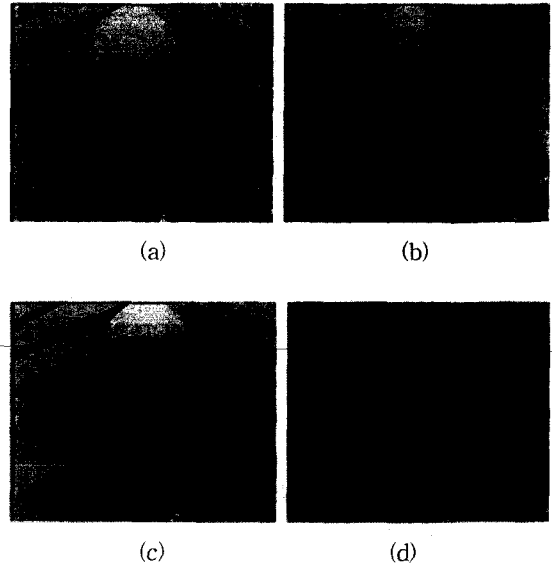


그림 8. 33Kbps에서 INTER 프레임의 주관적인 화질 비교 :
(a) DCT-based H.263(22번 프레임)
(b) DWT-based H.263(22번 프레임)
(c) DCT-based H.263(63번 프레임)
(d) DWT-based H.263(63번 프레임)
Fig. 8. Subjective quality of image in Inter frame on 33Kbps :
(a) DCT-based H.263(22nd frame)
(b) DWT-based H.263(22nd frame)
(c) DCT-based H.263(63rd frame)
(d) DWT-based H.263(63rd frame)

V. 결론

본 논문에서 제안한 DWT기반 H.263은 TMN5 구조를 그대로 유지하면서 DCT를 DWT로 대체하고 이 구조에 알맞은 적응 양자화기를 설계하였다. DWT 기반 H.263은 TMN5의 구조를 그대로 유지함으로써 기존 시스템과의 이식성이 우수하며 DWT는 DCT에 비해 압축율이 높아 같은 전송 속도에서 보다 많은 프레임을 전송할 수 있기 때문에 화면 드롭 현상이 감소된다. 뿐만 아니라 객관적 화질 평가의 기준이 되는 평균 PSNR 측면에서 휘도 신호는 평균 0.35dB정도의 손실을 보였으나 색차 신호에서는 평균 2dB정도가 개선된다. 주관적 화질평가에서도 TMN5에서 문제시되었던 블록킹 현상을 줄일 수 있어 보다 질 좋은 영상을 얻을 수 있다.

참 고 문 헌

- [1] ITU-T Recommendation H.263, "Video Coding for Low Bit Rate Communication," Mar. 1996.
- [2] ITU-T Recommendation H.263 Version 2, "Video Coding for Low Bit Rate Communication," Jan. 1998.
- [3] ITU-T Recommendation H.261, "Video Codec for Audiovisual Services at P×64 Kbits," Mar. 1993.
- [4] ITU-T SG 15, Working Party 15/1 Expert's Group on Very Low Bitrate Visual Telephony, "Video Codec Test Model TMN3," July 1994.
- [5] ITU-T SG 15, Working Party 15/1 Expert's Group on Very Low Bitrate Visual Telephony, "Video Codec Test Model TMN5," Jan. 1995.
- [6] R. de Queiroz, C. Choi, Y. Huh, J. Hwang and K. R. Rao, "Wavelet Transforms in a JPEG-like Image Coder," *Proc. of SPIE Conf. Visual Commun. and Image Processing*, vol. 2308, pp. 1662~1673, Sep. 1994.
- [7] Ya-Qin Zhang, "Very Low Bitrate Video Coding Standards," *Proc. SPIE VCIP'95*, vol. 2501, pp. 1016~1023, May 1995.
- [8] Marc Antonini, Michel Barlaud, Pierre Mathieu and Ingrid Daubechies, "Image Coding Using Wavelet Transform," *IEEE Trans. on Image Processing*, vol. 1, no. 2, pp. 205~220, April 1992.
- [9] Stephane G. Mallat, "A Theory for Multiresolution Signal Decomposition : The Wavelet Representation," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 11, no. 7, pp. 674~693, July 1989.
- [10] Detlev Marpe and Hans. L. Cycon, "Very Low Bit-Rate Video Coding Using Wavelet-Based Techniques," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 9, no. 1, pp. 85~94, Feb. 1999.
- [11] Mutsumi Ohta and Satoshi Nogaki, "Hybrid Picture Coding with Wavelet Transform and Overlapped Motion-Compensated Interframe Prediction Coding," *IEEE Trans. on signal processing*, vol. 41, no. 12, pp. 3416~3424, December 1993.
- [12] Detlev Marpe and Hans L. Cycon, "Very Low Bit-Rate Video Coding Using Wavelet-Based Techniques," *IEEE Trans. on circuits and systems for video technology*, vol. 9, no. 1, pp. 85~94, February 1999.
- [13] C. Sidney Burrus, Ramesh A. Gopinath and Haitao Guo, "Introduction to Wavelets and Wavelet Transforms," *Prentice-Hall International Inc*, 1998.

저 자 소 개



沈優星(正會員)

1994년 2월 원광대학교 전자공학과 졸업. 1996년 2월 원광대학교 전자공학과 석사 졸업. 1996년 3월~현재 원광대학교 전자공학과 박사과정. <관심분야> 영상처리, 영상통신, 통신시스템



李埴植(正會員)

1998년 2월 원광대학교 전자공학과 졸업. 1998년 3월~현재 원광대학교 전자공학과 석사과정. <관심분야> 영상처리, 영상통신, 통신시스템



鄭相淵(正會員)

1998년 2월 원광대학교 전자공학과 졸업. 1998년 3월~현재 원광대학교 전자공학과 석사과정. <관심분야> 영상처리, 영상통신, 통신시스템

許道根(正會員)

1975년 2월 울산대학교 전자공학과 졸업. 1980년 2월 경희대학교 전자공학과 석사. 1990년 3월 경희대학교 전자공학과 박사. 1980년~현재 원광대학교 전자공학과 교수. <관심분야> 영상처리, 영상통신, 통신시스템