

H.263 성능 개선을 위한 DWT 양자화기 설계

이 용 식, 정 상 연, 심 우 성, 허 도 근

원광대학교 전자공학과

전화 : (0653) 850-6742 / 팩스 : (0653) 857-3999

The Design of DWT Quantizer to Improve the H.263

Yong Sik Lee, Sang Yean Jung, Woo Sung Sim, Do Geun Huh

Dept. of Electronic Eng. Wonkwang University

E-mail : forlover@electro.wonkwang.ac.kr

Abstract

TMN5 used to test the performance of H.263 is a DCT-based H.263 which DCT is used as transform coding, but DWT is used instead of DCT and adaptive quantizer of being appropriate to DWT is designed in this paper while the structure of TMN5 is maintained. DWT-based H.263, which is proposed in this paper, is excellent in transplation of the existing system and can improve frame rate by transmitting more Frame. Also, the average PSNR, objective quality of image, has been lost at degree of average 0.35dB in luminance signals, but has been improved over a average of 2dB in chrominance signals. Subjective quality of image has been improved as blocking effect, which has seriously occurred in DCT based H.263, is diminished. DWT-based H.263 can, therefore, advance video quality of image comparing with the existing ways.

I. 서론

64Kbps 이하의 초저속망에서 동영상 전송을 목적으로 하는 H.263은 일반적인 동영상 압축기법에 쓰이는 3가지 중복성, 즉 신호 성분들간에 존재하는 중복성, 화면간에 존재하는 시간적 중복성, 그리고 화면내에 존재하는 공간적 중복성을 효과적으로 제거함으로써 고압축을 수행하는 하이브리드 방식의 알고리즘이다[1][2][4]. 신호 성분간에 존재하는 중복성은 휘도 신호와 색차 신호의 비

율을 2:1로 정의함으로써 제거되고, 화면간에 존재하는 시간적 중복성은 움직임 추정/보상(Motion Estimation/Compensation : ME/MC) 기법에 의해 제거되며, 화면내에 존재하는 공간적 중복성은 변환부호화와 양자화를 통해서 제거하고 있다.

H.263 권고안은 H.263의 성능평가를 위해서 쓰였던 TMN5(Test Model Near-term 5)와는 달리 기본적인 비트열의 구조만을 정의하였을 뿐 ME/MC, 변환부호화 및 양자화기의 세부적이며 구체적인 알고리즘을 정의하고 있지 않다[1][2]. H.263 권고안의 이러한 특징은 부호화 알고리즘에 어느 정도의 융통성을 제공함으로써 H.263 권고안을 만족하는 다양한 부호화기의 설계를 가능하도록 하고 있다.

본 논문에서는 낮은 비트율 또는 초저속 동영상 부호화에서 블록화 현상을 가져오는 DCT의 단점을 개선하고자 DWT 기반의 H.263과 이에 적합한 적응 양자화기를 제안한다. 기존의 TMN5가 변환부호화에서 DCT를 사용하지만 제안한 DWT 기반의 H.263은 DCT를 DWT로 대체하여 블록단위의 3레벨 DWT 변환을 수행하고 DWT 변환계수의 특성을 이용한 새로운 양자화기를 설계한다.

II. DWT 변환 부호화를 갖는 H.263

II-1. DWT 변환 부호화를 갖는 H.263의 구조

제안한 DWT 기반 H.263의 부호화기는 이미 개발된 기존 시스템과의 이식성을 높이기 위하여 부호화 과정 및 비트열 구조를 TMN5의 방식 그대로 유지하고 있다 [1][2]. 단, DWT 변환 계수 특성을 고려한 양자화기를 새롭게 제안한다. 제안한 DWT 기반 H.263의 부호화기를 그림 1에 보였다.

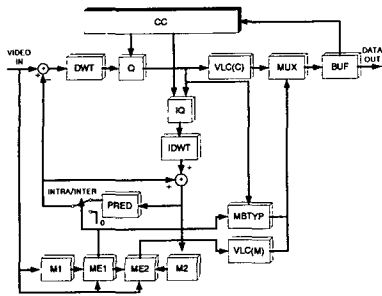
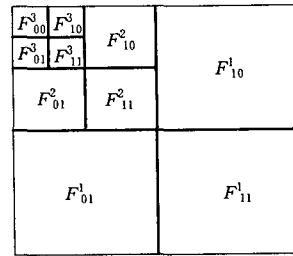


그림 1. DWT 기반의 H.263 부호화기



(b)

그림 2. 3레벨 DWT : (a) 3레벨 DWT 과정
(b) 3레벨 DWT의 부대역

II-2. DWT 변환 부호화

웨이블릿 변환을 이용하여 영상을 다해상도로 분해한 후 영상을 압축하는 것은 기존의 DCT를 이용한 방법보다 높은 압축율에서 블록화 현상이 일어나지 않아 인간 시각 특성(Human Visual Sensitivity : HVS)에 보다 적합한 압축 방법으로 평가되면서 차세대 압축 알고리즘으로 주목받고 있다.

DWT는 식 (1)과 같이 정의 할 수 있으며 본 논문에서는 ME/MC의 기본 단위가 되는 MB를 다시 크기 $M \times M$ 블록으로 나누어 블록에 대하여 DWT를 수행한다.

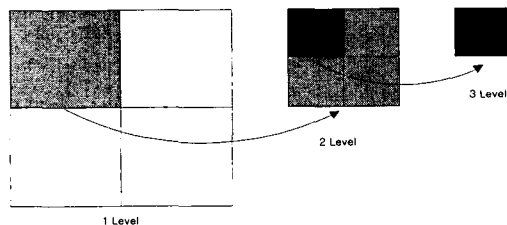
$$DWT[f_{i,k}(t)] = \int_{-\infty}^{\infty} f(t) \psi_{i,k}(t) dt \quad (1)$$

차수 N 인 저역 필터 $h(n)$ 과 고역 필터 $g(n)$ 를 이용해 데이터 블록 $f(i, j)$ 를 수평방향으로 필터링하고 이 결과를 수직방향으로 다시 필터링하면 입력된 데이터 블록의 DWT를 식 (2)와 (3)과 같이 얻을 수 있다[5].

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

$$f_H(i, j) = (g * f)(2k) = \sum_{n=0}^{N-1} g(n) f(2k-n) \quad (3)$$

윗 식을 이용해 DWT를 하면 하나의 최저 주파수대역과 3개의 고주파대역으로 나누어지며 더 세분화된 분해를 위해서 최저 주파수대역에 대해 DWT를 반복 수행한다. 이런 방법으로 8×8 블록을 3레벨 DWT 하면 10개의 부대역으로 분해할 수 있다. 그림 2(a)는 8×8 블록에 대한 3레벨 DWT 과정을 나타내고 그림 2(b)는 3레벨 DWT의 부대역을 나타내는데 F^r_{rs} 은 r 레벨의 (r, s) 부대역임을 의미한다.



(a)

III. 제안한 양자화기 설계

DWT 변환에서 고주파 대역보다 저주파 대역에 에너지가 집중되는 특성은 DCT와 유사하지만 같은 대역에서의 에너지 분포는 한 부분으로 집중하지 않고 원 영상의 에지 성분 에 따라 고주 분포한다. 이러한 DWT 계수특성 때문에 TMN5에서 이용했던 양자화기를 사용하면 심한 영상의 열화를 가져온다. 정지영상의 압축시 DCT를 DWT로 대체하고 DWT 변환계수를 양자화하기 위한 방법 [3]이 제안되었지만 이 방법을 영상전화 및 영상회의 등과 같은 동영상 응용에 적용할 경우 버퍼 상태의 고려없이 동일한 양자화기를 사용하므로 많은 비트열을 발생시켜 프레임율이 현저하게 저하된다. 따라서 본 논문에서는 동영상 부호화에 있어서 DWT 계수특성을 고려한 새로운 양자화기 행렬 H 를 얻기 위해 H 의 각 원소 h_{ij} 를 식(4)와 같이 정의한다.

$$h_{ij} = \frac{A}{g_{ij}} \quad (4)$$

$$g_{ij} = g \left(\frac{\sqrt{i^2 + j^2}}{M} \right) \quad (5)$$

$$g(x) = (a_0 x + a_1 x + a_2 x^2 + a_3 x^3) e^{(a_4 x + a_5 x^2)} \quad (6)$$

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 (4)와 식(5)에서 DWT 블록 크기가 $M \times M$ 이면 i, j 는 0에서 $M-1$ 까지의 값을 갖는다. 식 (6)에서 a_n 은 $\epsilon = \|Q_{JPEG} - H\|$, 즉 JPEG에서 인간 시각 특성을 잘 반영하는 것으로 제시된 양자화기 Q_{JPEG} 과 행렬 H 의 오차를 최소화하기 위해 찾아진 값들이다[3]. $0 \leq i \leq M-1, 0 \leq j \leq M-1$ 에 대해 $x = \sqrt{i^2 + j^2} / M$ 을 구하여 식 (6)에 대입해서 g_{ij} 를 구한다. 각각의 g_{ij} 를 식 (4)에 대입하여 새로운 양자화기 행렬 H 를 설계한다. 이 때 A 는 양자화기의 이득으로 특히 A 가 12의 값을 가질 때 Q_{JPEG} 과 가장 유사한 양자화기 행렬 H_{JPEG} 의 각 원소 $(h_{ij})_{JPEG}$ 을 얻을 수 있다.

$$(h_{ij})_{JPEG} = \frac{12}{g_{ij}} \quad (7)$$

DWT 계수는 부대역에 따라 고르게 분포하므로 각 부대역에 속하는 DWT 계수는 동일 양자화 간격으로 균일 양자화 해야 한다. 따라서 양자화 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 (8)과 같이 정의한다.

$$h_{rs}^n = \frac{2^{n+1}}{M} \sum_{i=-\frac{rM}{2}}^{\frac{(r+1)M}{2}-1} \sum_{j=-\frac{sM}{2}}^{\frac{(s+1)M}{2}-1} h_{ij} \quad (8)$$

식 (8)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(8)를 식(9)와 같이 정의한다.

$$h_{rs}^n = \sqrt{2} h_{ij} \quad (9)$$

식 (8)와 식 (9)에서 정의된 양자화 간격 $H_n(r, s)$ 로 H.263의 양자화기 행렬 $H_{H.263}$ 을 얻는다.

$$H_{H.263} = [h_{rs}^n] \quad (10)$$

$M=8$ 일 때 식(8)(9)(10)을 이용한 각 대역별 양자화기 행렬 $H_{H.263}$ 은 그림 3(c)와 같다.

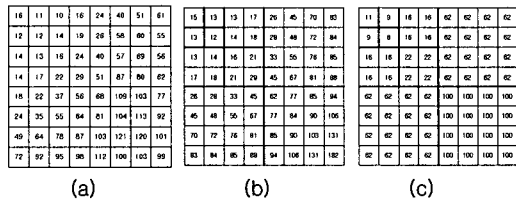


그림 3. 양자화 행렬 : (a) Q_{JPEG} (b) H_{JPEG} (c) $H_{H.263}$

양자화 이득 A 의 값은 TMN5에서 부호화시 발생하는 비트의 과다할당으로 인한 버퍼의 오버플로우를 방지하기 위해서 QP를 결정했던 방법과 동일한 방법으로 결정하고 1에서 31까지의 값을 가진다[1][2]. 이렇게 만들어진 양자화기 $H_{H.263}$ 으로 모든 DWT 계수(INTRADC, Non-INTRADC) $C(i, j)$ 를 동일하게 식 (11)과 같이 양자화 한다.

$$QC(i, j) = sign[C(i, j)] \times [|C(i, j)| / H_{H.263}(i, j)] \quad (11)$$

초기 INTRA 프레임의 양자화를 위한 A 의 초기값은 12로 한다. $H_{H.263}$ 양자화기에 의해 양자화된 계수값은 그림 4(a)와 같은 순서로 10개의 대역을 주사하고 각 대역 안에서 계수별 주사순서는 그림 4(b)와 같이 주사하여 전송하게 된다.

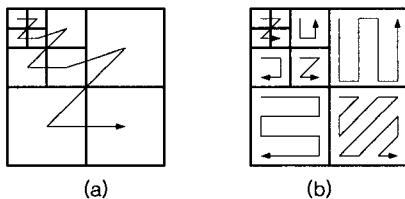


그림 4. DWT 계수의 주사순서 : (a) 대역별 주사순서

(b) 계수별 주사순서

III. 실험 결과 및 고찰

본 논문에서 제안한 DWT 기반 H.263의 성능을 분석하기 위해서 ITU-T에서 H.263의 평가를 목적으로 사용한 Susie, Carphone, Foreman의 QCIF(144×176)영상 100 프레임을 이용한다. Susie 영상은 다른 영상에 비하여 비교적 움직임이 적고 Foreman 영상은 다른 영상에 비해 다소 많은 움직임을 갖는다. 객관적인 평가를 위해서 움직임 추정은 ±15의 탐색 영역을 갖는 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 방식을 비교하고 있다. 표 1에서 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 프레임씩을 더 전송 처리함으로써 TMN5보다 높은 프레임율을 나타낸다.

표 1. 33Kbps에서의 성능

33Kbps		DCT-based H.263			DWT-based H.263		
QCIF Sequence	Susie	Car phone	Fore man	Susie	Car phone	Fore man	
Average PSNR	Y	35.06	31.75	29.62	34.68	31.29	29.39
	Cb	41.61	37.06	35.40	44.42	38.04	37.95
	Cr	41.33	37.13	36.65	44.03	38.57	38.64
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

24Kbps		DCT-based H.263			DWT-based H.263		
QCIF Sequence	Susie	Car phone	Fore man	Susie	Car phone	Fore man	
Average PSNR	Y	34.14	31.10	29.28	33.86	30.70	28.98
	Cb	40.91	36.90	35.30	43.77	37.65	36.38
	Cr	40.61	36.90	36.64	43.43	38.06	37.60
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 프레임별 휘도 신호의 PSNR 변화를 그림 5(a)에 나타내고 색차 신호의 PSNR 변화를 그림 5(b)에 나타내었다. 휘도 신호의 경우 움직임이 비교적 적은 40 프레임 이전에서는 TMN5보다도 높은 PSNR을 보이지만 움직임이 많은 40 프레임 이후에서는 제안된 방법의 PSNR이 감소한다. 색

차 신호의 경우 움직임의 정도에 관계없이 제안된 방법이 항상 높은 PSNR을 보인다.

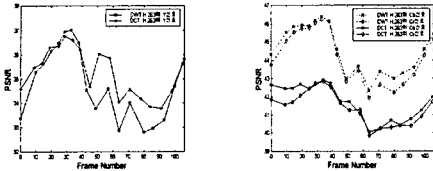


그림 5. Susie 영상에 대한 33Kbps에서 각 프레임에 따른 PSNR : (a) Y 신호에 대한 PSNR
(b) Cb, Cr 신호에 대한 PSNR

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

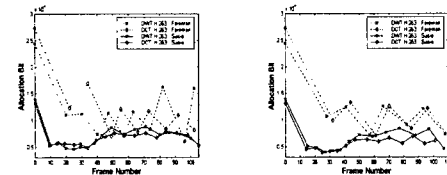


그림 6. Susie와 Foreman 영상의 각 프레임에 따른 비트 할당 : (a) 33Kbps에서 비트 할당
(b) 24Kbps에서 비트 할당

그림 7과 8은 각각 INTRA 프레임과 INTER 프레임의 화질을 주관적으로 평가하기 위한 것으로 그림 8은 전송 속도가 33Kbps일 때 Foreman 영상의 22번과 63번 프레임의 화질을 보인다.



그림 7. INTRA 프레임 : (a) DCT-based H.263
(b) DWT-based H.263

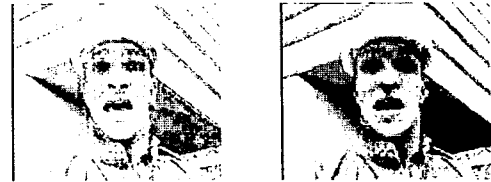


그림 8. 33Kbps에서 INTER 프레임의 주관적인 화질 비교 :
(a) DCT-based H.263(63번 프레임)
(b) DWT-based H.263(63번 프레임)

주관적인 화질 평가를 위해서 64Kbps, 33Kbps, 24Kbps에서 실험한 결과 제안된 DWT 기반 H.263의 영상이 INTRA/INTER 프레임 모두에서 블록화 현상이 어느 정도 제거되어 전체적으로 TMN5보다 우수한 화질의 영상을 제공한다.

IV. 결론

본 논문에서는 H.263의 시험모델인 TMN5에서 사용하는 변환부호화 방법인 DCT를 DWT로 대체하고 이 구조에 알맞은 적응 양자화기를 제안하였다. 제안된 DWT 기반의 H.263은 전체적인 구조는 TMN5에 기준 하였기 때문에 기존 시스템에의 이식성이 우수하고 같은 전송속도 환경에서 움직임이 많은 영상의 프레임율을 TMN5보다 높임으로 화면 드롭 현상을 줄일 수 있었다. 또한 객관적 화질 평가를 위한 평균 PSNR 측면에서 휘도 신호는 평균 0.35dB정도의 손실을 보았으나 색차 신호에서는 평균 2dB정도의 개선을 가져옴으로 결과적으로 더 좋은 화질을 가져왔을 뿐만 아니라 주관적 화질평가에서도 TMN5에서 문제시 되어왔던 블록화 현상을 줄임으로 우수한 영상을 얻을 수 있었다.

참고 문헌

- [1] ITU-T Recommendation H.263, "Video Coding for Low Bit Rate Communication," Mar. 1996.
- [2] ITU-T SG 15, Working Party 15/1 Expert's Group on Very Low Bitrate Visual Telephony, "Video Codec Test Model TMN5," Jan. 1995.
- [3] 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.
- [4] Ya-Qin Zhang, "Very Low Bitrate Video Coding Standards," Proc. SPIE VCIP'95, vol. 2501, pp. 1016~1023, May 1995.
- [5] Marc Antonini, Michel Barlaud, Pierre Mathieu and Ingrid Daubechies, "Image Coding Using Wavelet Transform," IEEE Trans. on Image Processing, vol. 1, no. 2, April 1992.