

H.263의 화질 개선을 위한 적응 양자화기 설계

The Design of Adaptive Quantizer to Improve Image Quality of the H.263

신 경 철*, 이 광 형**
(Shin Kyung Cheol*, Lee Kwang Hyung**)

요 약

H.263은 영상전화, 영상회의 등의 서비스를 64Kbps 이하 전송로에서 가능하게 하는 ITU-T의 국제 표준이다. 이 권고안에서는 움직임 추정/보상, 변환 부호화, 양자화방법을 기본으로 사용하고 있다. H.263의 성능평가에 사용된 TMN5는 변환부호화 방법으로 DCT를 사용하고, DCT 변환계수를 양자화하기 위한 양자화기가 제시되고 있다. 본 논문에서는 TMN5의 구조를 그대로 유지하면서 인간 시각 특성을 고려하여 DCT 계수를 효과적으로 양자화할 수 있는 적응 양자화기를 제안한다. 제안된 DCT 기반 H.263의 양자화기는 같은 전송 속도에서 TMN5보다 더 많은 프레임 전송 처리함으로써 화면 드롭현상을 줄일 수 있었다. 또한 객관적 화질 평가를 위한 평균 PSNR에서 TMN5보다 휘도 신호는 -0.3~+0.7dB의 차이를 보이고 색차 신호에서는 1.5dB 정도의 개선을 나타냈다. 결과적으로 주관적 화질평가에서는 TMN5에 비하여 더욱 선명한 화질의 영상을 얻을 수 있었다.

ABSTRACT

H.263 is an international standard of ITU-T that can make the service such as video phone, video conference in the transmission line less than 64Kbps. This recommendation draft has used motion estimation/compensation, transform coding and quantizing methods. TMN5 used for the performance estimation of H.263 has fundamentally used DCT in transform coding method and presented quantizer for quantizing the DCT transform coefficient. This paper is presenting adaptive quantizer effectively able to quantize DCT coefficient considering the human visual sensitivity while the structure of TMN5 is maintaining. As quantizer that proposed DCT-based H.263 could make transmit more frame than TMN5 in a same transfer speed, it could lower the frame drop effect. And the luminance signal appeared the difference of -0.3~+0.7dB in the average PSNR for the estimation of objective image quality and the chrominance signal appeared the improvement in about 1.5dB in comparison with TMN5. As a result it can attain the better image quality compared to TMN5 in the estimation of subjective image quality.

I. 서 론

문자, 도형, 음성, 정지영상뿐만 아니라 동영상까지 다양한 표현형식을 갖는 정보들은 디지털 신호처리, 디지털 통신, 반도체, 컴퓨터 등의 기술발전에 힘입어 주로 영상과 음성을 중심으로 결합된 멀티미디어의 형태로 저장 또는 전송되어진다. 이러한 정보의 표현이나 전송을 위한 국제 표준 규격들이 권고되고 있다. 이 중에서 64Kbps 이하의 전송속도를 갖는 초저속 동영상 부호화에 관한 표준안으로 ITU-T/LBC 그룹에서 H.263이라는 국제 표준안이 제정되었다[1][2].

공중전화망이나 이동통신망과 같은 초저속망에서 동영상 전송을 목적으로 하는 H.263은 일반적인 동영상 압축기법에

쓰이는 4가지 중복성을 효과적으로 제거함으로써 고압축을 수행하는 하이브리드 방식의 알고리즘이다[1][2][3]. 신호 성분 간에 존재하는 중복성은 휘도 신호와 색차 신호의 비율을 2:1로 정의함으로써 제거되고, 통계적 발생 확률에 존재하는 통계적 중복성은 데이터의 출현 확률 분포를 활용하여 평균 비트수를 줄이는 것으로 가변장 부호화에 의해 제거되고, 화면간에 존재하는 시간적 중복성은 움직임 추정/보상(ME/MC : Motion Estimation/Compensation) 기법에 의해 제거되며, 화면내에 존재하는 공간적 중복성은 변환 부호화와 양자화를 통해서 제거하고 있다[4][5]. H.263은 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[5]. H.263 권고안의 이러한 특징은 부호화 알고리즘에 어느 정도의 융통성을 제공함으로써 H.263 권고안을 만족하는 다양한 부호화기의 설계를

* 경남정보대학 전자정보과

** 숭실대학교 정보통신전자공학부

접수일자: 1999년 5월 26일

가능하도록 하고 있다. 여기서는 TMN5의 변환부호화로 DCT를 채택하여 입력 영상을 부호화한다. DCT는 다른 변환 방법과는 달리 고속 알고리즘과 순환 구조를 가지고 있어서 비교적 처리가 간단하고 다중 차원으로서의 확장을 용이하게 하는 장점으로 인해 많은 국제 표준에서 이를 기본 변환 부호화로 삼고 있다[8][9][11]. 본 논문에서는 DCT 변환부호화의 성능을 향상시키고자 DCT 변환계수의 특성과 인간 시각 특성(Human Visual Sensitivity : HVS)을 고려한 적응 양자화를 제안한다. 제안된 양자화 방법은 TMN5보다 높은 프레임율을 나타내며 객관적, 주관적 화질 또한 개선되었다. 본 논문의 구성은 II장에서는 DCT 기반 H.263의 기본 구조와 DCT 변환부호화에 대해서 설명하며, III장에서는 DCT 기반 H.263에 적합한 적응 양자화를 제안한다. IV장에서는 제안한 DCT 기반 H.263과 TMN5의 모의 실험을 통한 성능을 분석하며, V장에서 본 논문의 결론을 맺는다.

II. DCT 기반 H.263 : TMN5

2.1. TMN5의 전체 구성

TMN5는 H.263 권고안을 만족하는 영상전화 또는 영상의 시스템으로 ME/MC의 모든 과정은 MB(Macro Block : 16×16)단위로 이루어지고 변환부호화는 8×8 블록에 대하여 DCT를 수행한다. TMN5의 부호화기를 그림 1에 나타낸다[1][5].

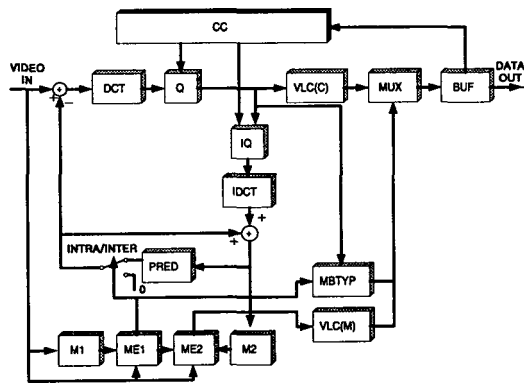


그림 1. DCT 기반 H.263 : TMN5
Fig. 1. DCT-based H.263 : TMN5.

DCT를 기반으로 하고 있는 TMN5의 부호화 과정을 살펴보면 먼저 현재 프레임의 입력 영상이 ME1(Motion Estimator 1) 블록에서 프레임 메모리 M1에 저장되어 있는 이전 프레임의 원 영상과 MB 단위로 블록 정합방식을 이용하여 정수 단위의 움직임 벡터(Motion Vector : MV)를 추정한다. 이 때 이전 영상과의 차이 정도를 식 (1)의 정합기준 SAD(Sum Absolute Difference)값에 따라 각각의 MB마다 INTRA/INTER 모드를 결정한다[5][6].

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

만약 현재 MB이 INTRA 모드이면 ME/MC없이 바로 DCT하고, INTER 모드이면 프레임 메모리 M2에 저장되어 있는 이전에 재생된 영상을 기준으로 ME2 블록에서 반화소 단위의 정확한 MV를 추정한 후 PRED에서 MV에 해당하는 블록을 가져와서 현재 블록과의 차신호를 구하여 이를 DCT한다. DCT 계수들은 양자화하여 가변장 부호화(VLC(C))를 하고 MB의 형태(MBTYP)와 MV도 가변장 부호화(VLC(M))를 하여 비트열을 발생시킨다. 재생된 영상을 프레임 메모리 M2에 저장해 놓기 위해서는 역양자화(Inverse Quantization : IQ)와 역DCT(Inverse DCT : IDCT) 블록이 필요하다.

2.2 DCT 변환 부호화

변환 부호화에서 1차원 또는 다차원의 입력 신호는 공간이나 시간영역으로부터 변환영역으로 매핑된다. 여기서 매핑은 유일하고 가역적인 직교특성을 만족한다. 이 경우 에너지는 변환영역에서도 보존되며 원 신호는 역변환에 의해 완전히 복원된다. 또한 1차원적인 변환을 연속으로 처리함으로써 다차원 변환을 수행하게 된다. 변환 부호화는 직교특성을 만족하기 때문에 역변환 과정이 순방향 변환 과정과 동일과정으로 이루어지며 모든 특성을 그대로 계승한다. 변환부호화는 여러 응용분야 가운데 특히 영상 또는 음성과 같은 신호의 압축에 크게 기여한다[6][8][9].

DCT는 1976년 Jain에 의해 입증되어 영상과 음성을 부호화하는 많은 국제표준화 기관(ISO, IEC, ITU-T, JPEG, MPEG 등)에서 채택되어 널리 이용되고 있다. DCT 수행과정은 입력된 영상 데이터를 $N \times M$ 의 블록으로 나누고 블록 내의 각 픽셀들에 대하여 식 (2)와 같이 DCT 연산을 행하며 그 역변환 과정은 식 (3)과 같다[6][8][9].

$$X_{u,v}^c = \frac{4}{NM} c_u c_v \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} x_{n,m} \cos\left[-\frac{(2n+1)u\pi}{2N}\right] \cos\left[-\frac{(2m+1)v\pi}{2M}\right] \quad (2)$$

여기서, $u=0,1,\dots,N-1, v=0,1,\dots,M-1$

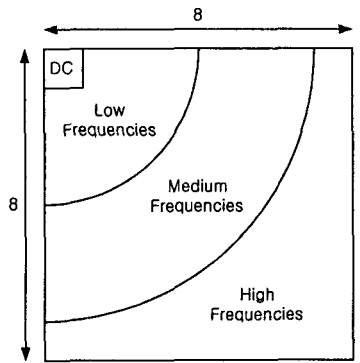
$$c_l = \begin{cases} 1/\sqrt{2} & l=0 \\ 1 & l \neq 0 \end{cases}$$

$$x_{n,m} = \sum_{u=0}^{N-1} \sum_{v=0}^{M-1} c_u c_v X_{u,v}^c \cos\left[\frac{(2n+1)u\pi}{2N}\right] \cos\left[\frac{(2m+1)v\pi}{2M}\right] \quad (3)$$

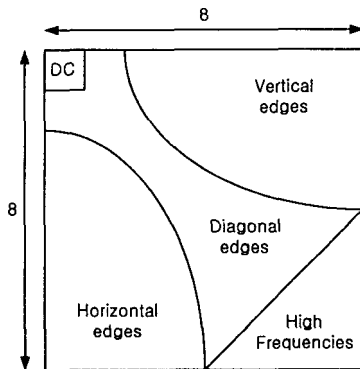
여기서, $n=0,1,\dots,N-1, m=0,1,\dots,M-1$

식 (2)에서 알 수 있듯이 DCT는 원 영상에 기저함수 cosine을 곱하여 변환한다. 최저 주파수를 나타내는 $u=0, v=0$ 인 경우를 DC 계수라 하고 나머지 계수를

AC 계수라 하며 u 와 v 가 클수록 고주파 성분임을 나타낸다. DC 계수는 블록 내 픽셀값의 평균값을 의미하고, AC 계수는 블록 내에 존재하는 에지 성분을 나타낸다. 그림 2는 8×8 블록에 대한 DCT 변환영역의 주파수 분포와 각 방향에 따른 에지 성분의 분포를 보이고 있다[11].



(a) Frequency distribution



(b) Edge distribution

그림 2. DCT 영역의 주파수 및 에지 분포
Fig. 2. Frequency and edge distribution of DCT domain.

식 (2)에 의하여 영상을 DCT 변환하면 변환계수들은 저주파 영역에서는 큰 값을 갖지만 고주파 영역에서는 작은 값을 갖는다. 따라서 영상의 대부분의 에너지가 DC 성분을 중심으로 하는 저주파 영역으로 모이는 DCT 계수 특성으로 인해 저주파 영역의 변환계수를 양자화하여 전송하고 고주파 영역의 변환계수는 양자화하는 과정에서 버려지는 방법으로 영상을 압축하게 된다. DCT는 다른 변환방법과는 달리 실수 부분만을 이용하여 변환하므로 빠른 속도로 변환처리가 가능하고 다중 차원으로의 확장을 용이하게 한다. DCT는 높은 비트 올에서는 좋은 성능을 나타내지만 낮은 비트 올의 환경에선 심한 블록화 현상이 나타나 화질 열화를 가져오는데 이는 H.263과 같이 64Kbps이하의 대역폭을 갖는 초저속 영상통신에서는 커다란 문제점이 된다. 따라서 DCT 변환계수 특성과 인간 시각 특성을 고려한 양자화기가 절대적으로 요구되어진다.

III. 제안한 DCT 기반 H.263에 적합한 적응 양자화기

DCT 기반의 TMN5 부호기 및 복호기에서는 INTRADC 계수값을 위한 1개의 양자화기와 Non-INTRADC 계수값을 위한 31개의 양자화기를 가지고 있다[1][5]. INTRADC 계수값은 양자화 간격 8을 갖는 균일 양자화기로 양자화되고 Non-INTRADC 계수값은 양자화 간격 2에서 64까지의 짝수 값을 갖는 31개의 양자화기로 균일 양자화한다. TMN5에서는 MB계층의 양자화 간격을 결정하는 QP(Quantization Parameter)를 변화시켜 버퍼에 유입되는 비트열을 제어한다. 따라서 QP는 부호화된 비트열의 발생률에 따라 1에서 31까지의 값을 갖고 양자화 간격은 $2QP$ 가 된다. 그리고 INTRA 프레임을 양자화하기 위한 초기 QP값은 16으로 한다. 변환계수 $C(i, j)$ 를 양자화 계수 $QC(i, j)$ 로 양자화하는데 있어서 INTRADC는 식 (4), Non-INTRADC 중 INTRA 블록의 AC는 식 (5), INTER 블록의 DC와 AC의 경우는 식 (6)에 의해 양자화 된다[5].

$$QC(0, 0) = C(0, 0) / 8 \tag{4}$$

$$QC(i, j) = \text{sign}[C(i, j)] \times [|C(i, j)| / 2QP] \tag{5}$$

$$QC(i, j) = \text{sign}[C(i, j)] \times [(|C(i, j)| - QP/2) / 2QP] \tag{6}$$

TMN5에서는 DC 계수의 중요성은 충분히 고려가 되었으나 모든 AC 계수는 동일한 양자화기를 사용하여 양자화하기 때문에 영상의 에지 성분에 대한 손실을 많이 가져온다. 따라서 본 논문에서는 인간 시각 특성을 만족하기 위해 그림 2와 같은 DCT 변환계수의 분포를 고려하여 선명한 화질의 영상을 제공할 수 있는 양자화기를 설계한다.

정지영상의 압축시 변환부호화로 DWT(Discrete Wavelet Transform)를 적용하고 이에 따른 인간 시각 특성을 만족하는 양자화기를 설계하려는 연구[10][12]가 시도되었지만 이를 영상전환 및 영상회의 등과 같은 동영상 응용에 적용할 경우 DCT를 사용하고 있는 기존 시스템과의 호환성이 문제가 되며 또한 버퍼 상태의 고려없이 동일한 양자화기를 사용하므로 많은 비트열을 발생시켜 프레임율이 현저하게 저하된다. 따라서 본 논문에서는 정지영상의 양자화 설계과정으로부터 H.263 동영상 부호화기 TMN5의 DCT 계수특성을 고려한 새로운 양자화기 행렬 $H_{H.263}$ 를 제안하고 버퍼에 입력되는 비트열에 따라 버퍼 상태를 고려하여 전체 부호화기의 성능을 좌우할 수 있는 새로운 양자화 이득 A 를 제안한다. 제안한 H.263 동영상 부호화기의 양자화기 행렬 $H_{H.263}$ 의 각 원소 $(h_{ij})_{H.263}$ 를 식(7)과 같이 정의한다.

$$(h_{ij})_{H.263} = \frac{A}{g_{ij}} \tag{7}$$

$$g_{ij} = g\left(\frac{\sqrt{i^2 + j^2}}{M}\right) \quad (8)$$

$$g(x) = (a_0x + a_1x^2 + a_2x^3 + a_3x^4)e^{(a_4x + a_5x^2)} \quad (9)$$

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

식 (7)과 (8)에서 DCT 블록 크기가 $M \times M$ 이면 i, j 는 0에서 $M-1$ 까지의 값을 갖는다. $0 \leq i \leq M-1, 0 \leq j \leq M-1$ 에 대해 식 (8)에 의해 $x = \sqrt{i^2 + j^2}/M$ 을 구하여 식 (9)에 대입해서 g_{ij} 를 구한다. 각각의 g_{ij} 를 식 (7)에 대입하여 새로운 양자화기 행렬 $H_{H.263}$ 를 설계한다. 이 때 A 는 양자화기의 이득으로 특히 A 가 12의 값을 가질 때 Q_{JPEG} 의 각 원소 $(q_{ij})_{JPEG}$ 과 가장 유사한 양자화기 행렬 H_{JPEG} 의 각 원소 $(h_{ij})_{JPEG}$ 을 얻을 수 있다.

$$(q_{ij})_{JPEG} \cong (h_{ij})_{JPEG} = \frac{12}{g_{ij}} \quad (10)$$

식 (9)에서 a_n 은 $\epsilon = \|Q_{JPEG} - H_{JPEG}\|$, 즉 JPEG에서 인간 시각 특성을 잘 반영하는 것으로 제시된 양자화기 Q_{JPEG} 과 H_{JPEG} 의 오차를 최소화하기 위해 찾아진 값들이다[7][10]. 그림 3의 (a)는 JPEG에서 인간 시각 특성을 가장 잘 만족하는 것으로 제시된 Q_{JPEG} 을 나타내고, (b)는 $A=12$ 인 $H_{H.263}$ 즉 H_{JPEG} 을 나타내며, (c)는 $A=15$ 일 때의 $H_{H.263}$ 을 나타낸다.

16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	68	109	103	77
24	35	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

(a) Q_{JPEG}

15	13	13	17	26	45	70	83
13	12	14	18	28	48	72	84
13	14	16	21	33	55	76	85
17	18	21	29	45	67	81	88
26	28	33	45	62	77	85	94
45	48	55	67	77	84	90	106
70	72	76	81	85	90	103	131
83	84	85	86	94	106	131	182

(b) $H_{H.263} = H_{JPEG} (A=12)$

19	16	16	21	32	56	88	104
16	16	17	22	34	60	90	105
16	17	20	26	42	69	95	107
21	22	26	37	56	83	101	110
32	34	42	56	78	97	106	118
56	60	69	83	97	105	113	133
88	90	95	101	106	113	129	164
104	105	107	110	118	133	164	227

(c) $H_{H.263} (A=15)$

그림 3. 양자화 테이블
Fig. 3. Quantization table.

양자화 이득 A 의 값은 TMNS에서 부호화시 발생하는 비트의 과다할당으로 인한 버퍼의 오버플로우를 방지하기 위해서 QP를 결정했던 방법과 동일한 방법으로 결정하고 1에서 31까지의 값을 가진다. TMNS에서는 식 (5), (6)과 같이 AC 계수에 대해선 DCT 계수 특성의 고려없이 QP값에 의해서 동일한 가중치(weight)를 갖는 양자화 행렬로 양자화 한다. 하지만 제한한 양자화 행렬은 양자화 이득 A 값만을 전송하여 AC 계수의 중요도에 따라 각기 다른 가중치를 갖는 양자화기 행렬 $H_{H.263}$ 을 구성하여 DCT 계수를 양자화 함으로 화질의 개선을 가져올 수 있다. 초기 INTRA 프레임의 양자화를 위한 A 의 초기값은 12로 한다. 이렇게 만들어진 양자화 행렬 $H_{H.263}$ 으로 DCT 계수 $C(i, j)$ 를 식 (11)과 같이 양자화 한다.

INTRADC or Non_INTRADC :

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

$H_{H.263}$ 에 의해 양자화된 계수들은 그림 4와 같이 지그재그 주사한 후 가변장 부호화를 이용하여 전송하게 된다[6].

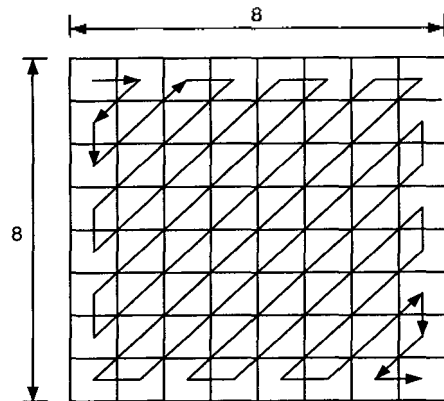


그림 4. DCT 계수의 지그재그 주사
Fig. 4. Zig-Zag scanning of DCT coefficients.

IV. 실험 결과 및 고찰

본 논문에서 제안한 DCT 기반 H.263의 적응 양자화기의 성능을 분석하기 위해서 ITU-T에서 H.263의 평가를 목적으로 사용한 3가지 QCIF(144×176, 4:2:0)영상을 이용한다[5]. 사용된 영상은 주로 영상전환나 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 ±7의 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 1. 33Kbps에서의 성능 비교

Table 1. Performance comparison of 33Kbps.

33Kbps	TMN5			제안한 양자화기			
	Susie	Carphone	Foreman	Susie	Carphone	Foreman	
QCIF Sequence							
Average PSNR	Y	35.0668	31.7585	29.6223	35.7275	31.4816	29.2991
	Cb	41.6132	37.0662	35.4043	43.9245	38.9653	36.5308
	Cr	41.3371	37.1374	36.6531	43.6368	38.9208	37.9597
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 비교

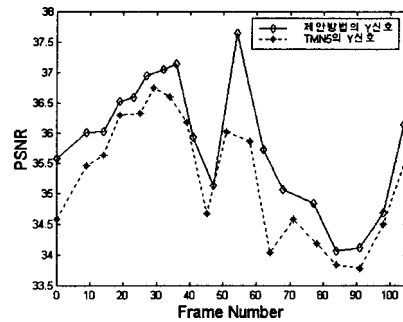
Table 1. Performance comparison of 24Kbps.

24Kbps	TMN5			제안한 양자화기			
	Susie	Carphone	Foreman	Susie	Carphone	Foreman	
QCIF Sequence							
Average PSNR	Y	34.1479	31.1025	29.2831	35.0750	30.9800	28.9383
	Cb	40.9197	36.9042	35.3079	43.2828	38.4181	36.3784
	Cr	40.6146	36.9071	36.6436	43.0598	38.4121	37.9043
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	

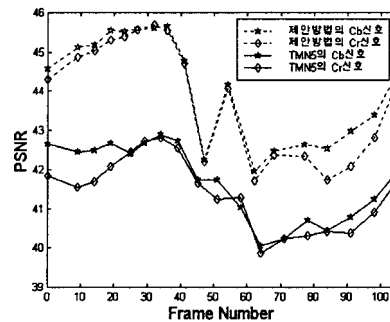
표 1과 표 2는 33Kbps와 24Kbps 환경에서 TMN5와 제안한 방식을 비교한 것이다. 표 1에서 평균 PSNR은 휘도 신호(Y)와 색차 신호(Cb, Cr)에 대하여 한 장의 INTRA 프레임을 포함한 전송된 프레임 각각의 PSNR을 평균한 것이다. 휘도 신호는 제안된 방식이 Susie 영상의 경우 평균 0.7dB 이상의 개선을 나타내고 나머지 영상에서 0.2~0.3dB 정도의 손실을 가져왔으며 색차 신호는 평균 1.5dB 정도의 개선을 보였고 움직임이 많은 Carphone 영

상과 Foreman 영상에서 한 프레임씩을 더 전송 처리함으로써 TMN5보다 높은 프레임율을 나타낸다.

그림 5와 그림 6은 각각 Susie와 Foreman 영상에 대한 33Kbps 환경에서 전송된 프레임별 휘도 신호와 색차 신호의 PSNR 변화를 나타낸다. 제안한 양자화기를 사용할 때 Susie 영상은 휘도 신호와 색차 신호 모두 TMN5보다 높은 PSNR을 보이지만 Foreman 영상은 휘도 신호에서는 약간의 손실을 보이는 반면 색차 신호의 경우 역시 높은 PSNR을 보인다.

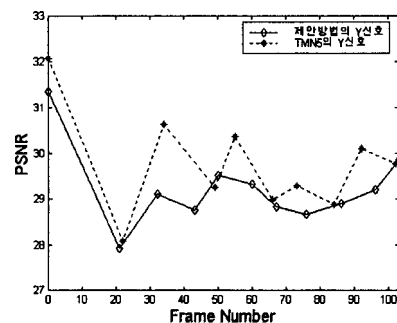


(a) Y 신호에 대한 PSNR

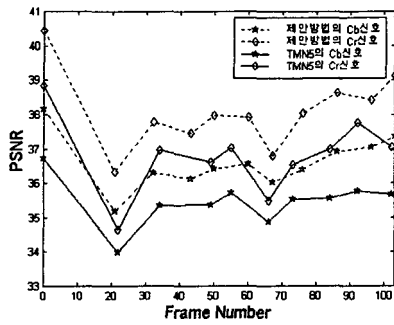


(b) Cb, Cr 신호에 대한 PSNR

그림 5. Susie 영상에 대한 33Kbps에서 각 프레임에 대한 PSNR
Fig. 5. PSNR for each frame of 33Kbps(Susie).



(a) Y 신호에 대한 PSNR

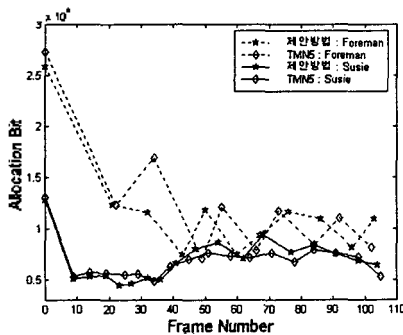


(b) Cb, Cr 신호에 대한 PSNR

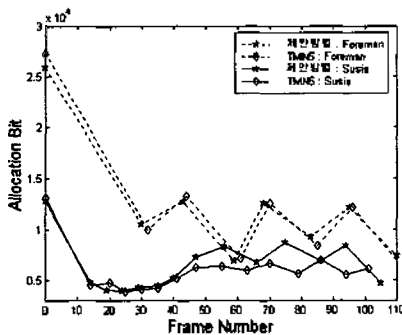
그림 6. Foreman 영상에 대한 33Kbps에서 각 프레임에 대한 PSNR

Fig. 5. PSNR for each frame of 33Kbps(Foreman).

그림 7은 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.



(a) 33Kbps



(b) 24Kbps

그림 7. Susie와 Foreman 영상의 각 프레임에 대한 비트 할당
Fig. 7. Bit allocation for each frame(Susie and Foreman).

그림 8은 INTRA 프레임의 주관적인 화질의 차이를 보이기 위한 것이며 그림 9는 전송 대역폭이 33Kbps일 때 Carphone 영상과 Foreman 영상의 INTER 프레임에 대한 주관적인 화질의 차이를 보인다.



(a) TMNS

(b) 제안한 양자화기

그림 8. INTRA 프레임의 주관적인 화질 비교

Fig. 8. Subjective image comparison of INTRA frame.



(a) TMNS(20th frame)

(b) 제안한 양자화기(20th frame)



(c) TMNS(63th frame)

(d) 제안한 양자화기(63th frame)

그림 9. 33Kbps에서 INTER 프레임의 주관적인 화질 비교

Fig. 9. Subjective image comparison of INTER frame(33Kbps).

주관적인 화질 평가를 위해서 56Kbps, 40Kbps, 33Kbps, 24Kbps에서 실험한 결과 제안된 양자화기를 사용한 영상이 INTRA 프레임과 INTER 프레임 모두에서 블록화 현상이 어느 정도 제거되고 영상의 에지 성분이 향상되어 전체적으로 TMNS보다 우수한 화질의 영상을 제공한다.

V. 결론

본 논문에서는 H.263의 시험모델 TMNS에서 사용하는 변환부호화 방법인 DCT를 유지하면서 DCT 변환계수의

분포 특성 및 인간 시각 특성을 고려한 적응 양자화기를 제안하였다. 제안된 H.263은 양자화기를 제외한 모든 시스템의 전체적인 구조는 TMN5에 기준 하였기 때문에 기존 시스템에의 이식성이 우수하고 같은 전송 속도 환경에서 움직임이 많은 영상의 프레임율을 TMN5보다 높임으로 화면 드롭 현상을 줄일 수 있었다. 또한 객관적 화질 평가를 위한 평균 PSNR 측면에서 휘도 신호는 평균 -0.3~+0.7dB 정도의 차이를 보이고 색차 신호에서는 평균 1.5dB 정도의 개선을 가져옴으로 결과적으로 더 좋은 화질을 가져왔을 뿐만 아니라 주관적 화질 평가에서도 TMN5에서 문제시 되어왔던 블록화 현상이 어느 정도 제거된 선명한 영상을 얻을 수 있었다.

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 Bit rate Visual Telephony, "Video Codec Test Model TMN3," July 1994.
5. ITU-T SG 15, Working Party 15/1 Expert's Group on Very Low Bit rate Visual Telephony, "Video Codec Test Model TMN5," Jan. 1995.
6. K. R. Rao, J. J. Hwang, "Techniques and Standards of Image, Video, and Audio Coding," Prentice Hall PTR, 1996.
7. ISO/IEC JTC1 10918-1] ITU-T Rec. T.81, "Information Technology-Digital Compression and Coding of Continuous-tone Still Images : Requirement and Guidelines," 1994.
8. S. C. Chan and K. L. Ho, "A New Two-Dimensional Fast Cosine Transform Algorithm," IEEE Trans. Signal Process., vol. 39, pp. 481~485, Feb. 1991.
9. N. I. Cho and S. U. Lee, "Fast Algorithm and Implementation of 2D Discrete Cosine Transform," IEEE Trans. Circuits Systems, vol. 38, pp. 297~305, Mar. 1991
10. 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.
11. J. Y. Nam and K. R. Rao, "Image Coding Using a Classified DCT/VQ based on Two Channel Conjugate Vector Quantization," IEEE Trans. Circuits and System for Video Technology, vol. 1, pp. 325~336, Dec. 1991.
12. R. de Queiroz, C. K. Choi, Young Huh, and K. R. Rao, "Wavelet Transforms in a JPEG-Like Image Coder," IEEE Trans. Circuits and Systems for Video Technology, vol. 7, pp. 419~424, Apr. 1997.

▲신 경 철(Kyung Cheol Shin)



1987년 2월 : 원광대학교 전자공학과 (학사)
 1990년 2월 : 숭실대학교 대학원 전자공학과(석사)
 1995년 2월 : 숭실대학교 대학원 전자공학과 박사과정 수료
 1994년~현재 : 경남정보대학 전자정보과 조교수

※주관심분야 : 멀티미디어, 영상신호처리, 영상압축

▲이 광 형(Kwang Hyung Lee)



1968년 2월 : 서울대학교 전자공학과 (학사)
 1972년 2월 : 서울대학교 전자공학과 (석사)
 1987년 2월 : 중앙대학교 전자공학과 (박사)
 1973년~1982년 : 일본 동경대학 전자공학과 연구원

1993년~1995년 : 대한전자공학회 인공지능, 신경망, 퍼지 시스템 연구회 위원장.

1994년~현재 : 국제 전기전자학회(IEEE Korea Council) 학술이사

1994년~1995년 : 대한 퍼지 및 지능 시스템 학회 국제 담당이사

1994년~현재 : 대한전자공학회 평의원 이사

1982년~현재 : 숭실대학교 정보통신전자공학부 교수

※주관심분야 : 지능 신호처리, 퍼지신경망 음성인식 및 합성, 멀티미디어 통신처리, ATM 망