

양자화기 벡터 코드북을 이용한 HDTV 영상 적응 부호화

(Adaptive Coding Algorithm Using Quantizer Vector Codebook in HDTV)

金益煥*, 崔振秀**, 李光春*, 朴吉欽**, 河永浩**

(Ick Hwan Kim, Jin Soo Choi, Kwang Chun Lee, Kil Houm Park and Yeong Ho Ha)

要約

일반적인 영상 압축 방법은 연속 영상에 존재하는 시공간적인 중복성을 예측부호화(DPCM)와 변환 부호화 또는 이 두 방법을 결합한 예측 변환 부호화를 이용하여 제거하는데 기본을 두고 있다. 본 논문에서는 변위 프레임 차(displaced frame difference : DFD) 신호의 DCT 변환된 8×8 블럭 계수 각각을 인간 시각 특성과 버퍼의 상태에 의해 미리 추정된 전체 목표 왜곡값에 따라 네개의 양자화기 중 하나로 양자화하였다. 이때 블럭당 가능한 양자화기 선택 벡터(패턴)의 수는 4^6 가 된다. 이 벡터들을 전송하기 위해 부호화할 경우 너무 많은 부가 정보가 요구된다. 그러므로 본 논문에서는 벡터간의 가중된 절대차의 합(sum of weighted absolute difference : SWAD)을 이용하여 Y신호는 2048개, U, V 신호는 각각 512개로 양자화기 선택 벡터의 수를 제한하는 방법을 제안하였다. 제안한 방법으로 두가지의 양자화기 벡터 코드북을 구성하여 실험한 결과 주어진 전송률(17Mbps)내에서 객관적인 화질(PSNR)과 주관적인 화질이 양호함을 보였다.

Abstract

Video compression algorithms are based on removing spatial and/or temporal redundancy inherent in image sequences by predictive(DPCM) encoding, transform encoding, or a combination of predictive and transform encoding. In this paper, each 8×8 DCT coefficient of DFD(displaced frame difference) is adaptively quantized by one of the four quantizers depending on total distortion level, which is determined by characteristics of HVS(human visual system) and buffer status. Therefore, the number of possible quantizer selection vectors(patterns) is 4^6 . If this vectors are coded, too many bits are required. Thus, the quantizer selection vectors are limited to 2048 for Y and 512 for each U, V by the proposed method using SWAD(sum of weighted absolute difference) for discriminating vectors. The computer simulation results, using the codebook vectors which are made by the proposed method, show that the subjective and objective image quality (PSNR) are good with the limited bit allocation.(17Mbps)

*正會員, (株)金星社 TV 研究所
(TV Research Laboratory, Goldstar Corporation Limited)

**正會員, 慶北大學校 電子工學科
(Dept. of Elec. Eng., Kyungpook Nat'l Univ.)
接受日字 : 1993年 11月 5日

1. 서론

완전 디지털 HDTV 신호의 수평, 수직 해상도는 NTSC 신호의 약 2배가 되어, 이 신호의 전송이나 저장을 위해서는 디지털 NTSC 신호의 전송률보다 최소한 다섯배가 요구된다. HDTV 신호의 디지털 처리는 대략 1.2 Gbit/sec가 필요하다. 제한된 대역 폭내에서 HDTV 신호를 전송하기 위해서 보다 높은 압축 효과를 얻을 수 있는 방법에 관한 연구가 계속되고 있다.

본 논문에서는 완전 디지털 HDTV에서 사용하는 영상 압축 알고리즘에 관한 연구로써 계층적 블럭 정합 방법(hierarchical block matching algorithm)을 사용하여 영상에 존재하는 시간적인 중복성을 제거하였고, 움직임 벡터를 이용하여 구한 프레임간의 예측 에러(prediction error)를 2차원 DCT(discrete cosine transform)를 행함으로써 공간적인 중복성을 제거했으며, 인간 시각 특성을 이용하여 적응 양자화함으로써 진폭의 중복성을 제거하였다. 본 논문에서는 8 x 8 DCT 계수 각각을 주파수 민감도, 시공간적 마스크 효과와 대비 민감도를 이용한 인간 시각 특성과 버퍼의 상태에 따라 구한 양자화 허용 오차 추정치, 즉 전체 목표 왜곡값에 따라 양자화기를 달리 선택함으로써 적응 양자화하였다. 부호화를 간단히 하기 위해 양자화기의 종류를 네가지로 선택하였으며, Y 신호는 16 x 16(2H x 2V)의 대구획 블럭(superblock)을 구성하고, U 및 V 신호는 16 x 24(2H x 3V)의 대구획 블럭을 구성하였다. 대구획 블럭의 구성원인 각각 8 x 8 블럭의 같은 위치에 있는 계수들은 하나의 양자화기로 양자화한다. 따라서 대구획 블럭에 대한 양자화기 선택 패턴(벡터)은 464개가 된다. 이 선택 패턴들은 각 양자화된 계수의 부호화와 복호화에 사용되므로 수신단에 전송되어야 한다. 전송을 위해 이 선택 패턴들을 부호화할 경우 대구획 블럭당 128비트가 필요하게 되어 오히려 부가정보가 더 많아지게 된다.

따라서 본 논문에서는 양자화기 선택 패턴수를 Y 신호는 2048개(11비트), U, V 신호는 각각 512개(9비트)로 제한하는 방법을 제안하였다. 양자화기 종류를 나타내는 심볼로 구성된 각각의 양자화기 선택 벡터들로부터 제한된 수의 양자화기 선택 벡터를 가진 코드북을 구성하기 위해 벡터들간의 유사성의 측정자로서 가중된 절대차의 합을 이용하였다. 두 벡터간의 절대차의 합이 주어진 임계치보다 크면 다른 벡터로 간주하고, 임계치보다 작으면 같은 벡터로 간주한다. 이때 양자화기 선택 벡터의 수를 제한함으로써

발생하는 화질의 저하를 최대한 막기위해 두가지 사항을 고려하였다. 절대차의 합을 구함에 있어 레벨수가 많은 양자화기가 영상복원시에 많은 영향을 미치므로 레벨수의 차이가 많은 양자화기간의 절대차를 크게 하고 상대적으로 레벨수의 차가 적은 양자화기간의 절대차를 작게 하였다. 그리고 블럭내에서 저주파 위치의 계수를 양자화하는 양자화기는 고주파 위치의 계수를 양자화하는 양자화기보다 화질에 더 큰 영향(blocking effect)을 미치므로 양자화기의 상대적인 위치에 따라 가중치를 두가지 형태로 다르게 하는 방법을 제안하였다. 이는 양자화기간의 절대차를 선형적으로 저주파에서 높게 고주파에서는 낮게 가중하는 방법과 MTF(modulation transfer function)함수를 이용하여 비선형적으로 가중하는 방법을 말한다. 제안한 방식으로 구한 코드북을 이용하여 실험한 결과 할당된 비트(17Mbps)내에서 33 - 35dB 정도의 양호한 화질의 영상을 얻었다.

II. 영상 압축의 개요

본 논문에서 영상 신호를 압축하기 위해 사용하는 부호화기 루프는 그림 1과 같다. 부호화기 루프는 움직임 벡터, 인간 시각 특성과 버퍼 상태에 따라 추정된 양자화 허용 오차 추정치, 즉 전체 목표 왜곡 레벨(global target distortion level)과 제어 변수를 이용하여 움직임 보상된 차 신호를 DCT 변환한 후 적응 양자화한다.

움직임 보상기(motion compensator)는 버퍼에 저장된 이전 영상신호에 움직임을 적용하여 움직임 보상 신호를 구성한다. 움직임 보상신호와 원입력신호 간의 차를 구하여 원 영상에 비해 아주 작은 정보를 가지는 움직임 보상된 차 신호(displaced frame difference ; DFD)를 생성한다. 이 DFD 신호는 DCT 변환이 되고, 양자화기 벡터 선택기(quantizer

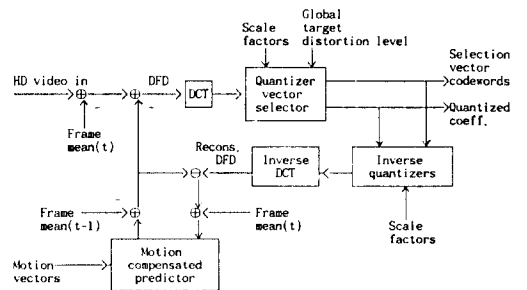


그림 1. 부호화기 루프 Fig. 1. Encoder loop.

vector selector)는 전체 목표 왜곡 레벨(값)에 따라 적절한 양자화기 벡터를 선택하여 그 벡터에 따라 각 계수들을 적응적으로 양자화한다. 여기서 양자화기 벡터는 주어진 계수 블록에 대해 적용될 수 있는 양자화기의 가능한 조합을 나타낸다. 이때 패턴의 수가 너무 많으므로 본 논문에서는 이 패턴의 수를 제한하는 방법을 제안하였다. 그리고 양자화기 벡터 대신에 그와 관계된 인덱스(index)를 전송함으로써 부호화 효율을 높였다.

Ⅲ. 전체 목표 왜곡값 추정

영상에 존재하는 중복성을 줄이기 위해 먼저 연속되는 두 프레임간의 시간적인 중복성은 계층적 블록 정합 알고리즘을 적용하여 움직임 추정^[1]으로 제거하였고, 공간적인 중복성을 줄이기 위해서는 추정된 움직임 벡터를 이용하여 현재 프레임을 추정하여 변위 프레임 차신호(DFD)를 구한 후 2차원 DCT 변환하였다.

양자화시에는 복원된 영상의 최종 관찰자가 인간이라는 점을 감안하여 양자화 오차를 인간 시각에 잘 띄지 않은 부분에 균일하게 분포시켜 같은 비트율이라도 인간이 느끼는 화질의 열화는 더욱 줄이도록 한다. 이를 위해 인간 시각 특성을 이용하여 8 x 8 블록의 각 주파수 대역에서 양자화시 허용될 수 있는 인식 문턱값을 구하고, 버퍼의 상태에 따라 발생하는 비트량을 제어하기 위해 인식 문턱값을 가중하는 목표 왜곡값을 추정하였다. 그리고 두 추정치를 이용하여 양자화시에 사용되는 전체 목표 왜곡값을 구하였다.

1. 인식 문턱값 추정

인간 시각 시스템의 인식 왜곡 특성을 고려하기 위해 영상에 독립적인 요소인 주파수 민감도(frequency sensitivity), 영상에 의존하는 성분인 시공

간적 마스크 효과(spatio-temporal masking effect) 및 대비 민감도와 같은 특성을 모델화하여 양자화기 선택에 이용되는 인식 문턱값을 추정한다. 그림 2는 인식 문턱값을 추정하는 인식 문턱값 연산기를 나타낸다.

주파수 민감도는 시각 시스템의 MTF(modulation transfer function)가 평평하지 않다는 특성을 이용한다. MTF의 곡선은 저주파보다 고주파에서 민감도가 떨어진다. 양자화 오차를 저주파보다 고주파에 많이 부여할 수 있으므로 고주파 대역에 대한 인식 문턱값은 저주파 대역의 인식 문턱값보다 높게 된다는 것을 뜻한다. 주파수 민감도를 구하기 위해 사용한 MTF곡선은 Nilli이 실험적으로 제안한 다음 식을 이용하였다.^[2]

$$H(\omega_j) = (0.2 + 0.45\omega_j) \exp(-0.18\omega_j) \quad (1)$$

여기서 ω_j 는 8 x 8 블록에 대한 공간주파수를 나타내며, 이는 HDTV 영상을 화면 높이의 3배 거리에서 시청한다고 가정하고 화면비는 16 : 9로 하였을 때 다음 식에 의해 구해진다.^[3] 여기서 N은 8이고, $i, j = 0, 1, \dots, N-1$ 이다.

$$\omega_j = \sqrt{\left(\frac{48i}{N-1}\right)^2 + \left(\frac{27j}{N-1}\right)^2} \quad (2)$$

인식 문턱값은 그 값이 클수록 인간 시각의 감지도가 떨어진다는 것을 고려하기 위해 MTF 함수의 역수를 취하여 값의 범위를 다음 식을 이용하여 변화시켰다.

$$F_s(i, j) = 0.5 \left(\frac{1}{MFT(\omega_j) + 0.3} \right) \quad (3)$$

시공간적 마스크 효과는 영상의 명암이 심하게 변하는 곳이나 그 부근에서 발생하는 양자화 오차는 인간 시각에 잘 띄지 않는 공간적 마스크 성질과 영상이 연속적으로 변하는 경우 화면내의 어느 고정된 위치에서 두 프레임간의 영상 값에 급격한 변화가 있을 때 HVS는 현재 프레임내의 정해진 위치에서의 고주파 상세함에는 민감도가 떨어진다는 시간적인 마스크 성질을 이용한다. 일반적으로 이러한 시공간 마스크의 측정자로서 주파수 영역에서의 AC 에너지 양^[4]을 사용하며, Rao는 주관적으로 개선된 영상을 얻기 위해 DCT 변환된 블록의 위쪽 부분을 수평, 왼쪽 부분을 수직, 나머지 부분은 대각선 성분의 구조로 나누어서 처리할 수 있다고 하였다.^[5] 인간 시각 시스

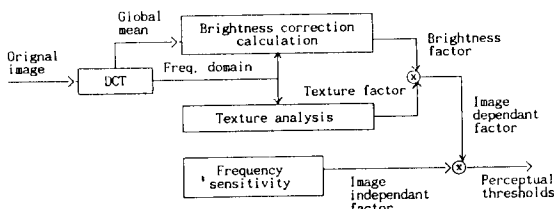


그림 2. 인식 문턱값 연산기
Fig. 2. Perceptual threshold calculator.

템(HVS)의 시공간적 매스킹 효과가 방향성에 따라 영향을 받으므로, 공간적 매스킹 효과를 고려하기 위해 먼저 식 (4) - (6)에 따라 수평, 수직, 대각선 성분의 추정치^[6] (T_{RH} , T_{RV} , T_{RD})를 구하였다.

$$T_{RH} = \sum_{\text{over } 8 \times 8 \text{ window}} MTF(0, j) \cdot \bar{I}_T^2(0, j) \quad (4)$$

$$T_{RV} = \sum_{\text{over } 8 \times 8 \text{ window}} MTF(i, 0) \cdot \bar{I}_T^2(i, 0) \quad (5)$$

$$T_{RD} = \sum_{\text{over } 8 \times 8 \text{ window}} MTF(i, j) \cdot \bar{I}_T^2(i, j) \quad (6)$$

where $i \neq 0$ and $j \neq 0$

여기서 $MTF(i, j)$ 는 인간 시각계의 변조 전달 함수이며, $\bar{I}_T(i, j)$ 는 평균이 제거된 영상의 DCT 변환된 계수들이다. 세가지 방향 성분에 대한 AC 에너지 양의 정도를 구하기 위해 식 (7) - (9)를 이용하여 각 방향성분의 AC 에너지(질감 요소 ; texture factor)를 구하였다.

$$T_H = T_{RH} + 0.50 \times T_{RD} \quad (7)$$

$$T_D = 0.25 \times T_{RH} + T_{RD} + 0.25 \times T_{RV} \quad (8)$$

$$T_V = 0.5 \times T_{RD} + T_{RV} \quad (9)$$

시간적인 매스킹을 고려하기 위해서는 이전 DCT 변환된 영상 정보와의 차를 구한 값을 이용하여, 식 (4) - (9)를 이용하여 세가지 방향성분의 질감요소를 구하였다. 그리고 시공간적 매스킹 효과를 동시에 고려하기 위해 식 (10)을 이용하여 세가지 성분의 질감요소를 추정하였다. 이 값은 방향성 질감요소의 전력(power)이므로 식 (11)을 이용하여 진폭(amplitude)으로 바꾸어 질감 보정인자(Texture Correction)를 구하였다.

$$T_{comb} = 0.5 \times T_{spatial} + 0.5 \times T_{temporal} \quad (10)$$

$$T_{correct} = 1 + \log(1 + 0.0001 \times T_{comb}) \quad (11)$$

대비 민감도는 인간 시각 특성은 균일한 자극에 대해 평균 밝기가 아주 어둡거나 밝을 때보다 중간 명암도일때 양자화 오차에 더 민감하다는 성질을 이용하였다. 웨버의 법칙^[7]에 근거하여 블록의 평균밝기에 따라 주파수 민감도와 시공간적 매스킹 효과를 고려하여 구한 추정값을 다음 식을 이용하여 구한 블록의 평균밝기에 대한 보정 인자(B_i)로 가중하였다. 여기서 I 는 블록의 평균밝기이다.

$$B_i = \left(\frac{1}{0.15 \cdot \log_{10}(I)} - 2.06 \right) \quad (12)$$

위 HVS 모델의 세가지 특성, 즉 영상에 독립적인 요소인 주파수 민감도, 영상에 의존하는 성분인 시공간적 매스킹 효과 및 대비 민감도를 고려하여 구한 각각의 가중치를 조합하여 최종 인식 문턱값을 생성한다. 이 최종 인식 문턱값은 인간 시각에 따른 8 x 8 블록의 각 주파수 대역에서 양자화시에 허용될 수 있는 양자화 오차를 추정하는데 이용하였다.

2. 목표 왜곡값 추정

계수들의 양자화시 인식 문턱값과 더불어 고려되어야 하는 것은 버퍼상태이다. 버퍼의 넘침(overflow)과 고갈(underflow)을 피하기 위해 버퍼의 상태가 기준 레벨보다 높으면 양자화시에 상대적으로 많은 양자화 오차를 주어 버퍼의 상태를 낮추고, 버퍼의 상태가 낮으면 양자화시에 작은 양자화 오차를 주어 버퍼의 상태를 높여 준다.^[8] 이를 위해 다음 식을 이용하여 버퍼의 상태에 따라 인식 문턱값을 가중하는 목표 왜곡값을 추정하였다.^[6]

$$D_t = \sigma_t \left(\frac{D_{t-1}}{\sigma_{t-1}} \right) e^{\frac{[(B_{t-1} - B_{ref})K_0 + (B_{t-1} - B_{t-2})]}{-b}} \quad (13)$$

여기서 B_{ref} 는 원하는 버퍼의 기준 레벨, B_t 는 t번째 프레임의 버퍼 상태, b 와 K_0 는 버퍼 제어 변수로 K_0 는 0과 1사이의 값을 가지며, D_t 는 t번째 프레임의 버퍼 충만도에 따른 목표 왜곡값이며, σ_t 는 t번째 프레임의 표준편차이다.

HVS 모델을 이용하여 추정한 인식 문턱값과 버퍼 충만도를 고려한 목표 왜곡값을 이용하여 대규모 블록당 8 x 8의 전체 목표 왜곡값을 구하고 이 값을 비트 할당 과정에서 사용한다. 그림 3은 전체 목표 왜곡값(레벨)을 구하기 위한 개략도를 나타낸다. 여기서 $I(t)$ 는 t번째 영상신호를 나타내고 $M(t)$ 는 t번째 영상의 평균값을 나타낸다. 이 두값을 입력으로 하여 perceptual processor는 최종 인식 문턱값을 생성하게 된다. 그리고 t번째 영상의 표준 편차값인 $\sigma(t)$ 와 버퍼충만도를 입력으로하는 rate processor가 목표 왜곡값을 계산한다. 그리고 최종 인식문턱값과 목표 왜곡값을 조합하여 전체 목표 왜곡값을 생성한다.

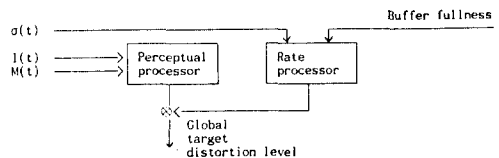


그림 3. 전체 목표 왜곡 레벨 연산기
Fig. 3. Global target distortion level calculator.

Ⅳ. 적응 양자화

1. 공간 변환

본 논문에서는 영상 신호의 공간적인 중복성을 제거하기 위해 변환 부호화 방법 중 한 방법인 DCT 변환을 사용하였다. 예측 에러(DFD)의 8 x 8 화소의 블록을 2차원 DCT^[5]를 사용하여 변환한다. 입력 신호 $f(i,j)$ (단, $i,j = 0,1, \dots, N-1$)의 2차원 DCT 변환 및 역변환은 각각 식 (14) 및 (15)와 같다.

$$F(u,v) = \frac{4c(u)c(v)}{N^2} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} f(i,j) \cos\left(\frac{(2i+1)u\pi}{2N}\right) \cos\left(\frac{(2j+1)v\pi}{2N}\right) \quad (14)$$

$$f(i,j) = \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} c(u)c(v)F(u,v) \cos\left(\frac{(2i+1)u\pi}{2N}\right) \cos\left(\frac{(2j+1)v\pi}{2N}\right) \quad (15)$$

여기서 $u, v = 0, 1, \dots, N-1$

$$c(w) = \begin{cases} \frac{1}{\sqrt{2}} & , w = 0 \\ 1 & , w = 1, 2, \dots, N-1 \end{cases}$$

2. DCT 계수의 적응 양자화

프레임당 발생하는 평균 비트수가 할당된 비트 예산내에 들도록 양자화기를 제어할 수 있어야 하며, 양자화 잡음을 가능한 영상 전체에 골고루 분포시키도록 양자화기를 조절할 수 있어야 한다. 프레임당 발생하는 비트수가 할당된 비트 예산내에 들면서 인간 시각 특성과 버퍼 상태를 고려하여 양자화 오차를 영상 전체에 골고루 분포시키기 위해서 다음과 같은 과정을 거친다.

우선 DCT 변환 회로는 8 x 8 주파수 계수 집합을 Y 신호는 2H x 2V 배열로, U, V 신호는 2H x 3V 배열로 무리를 짓는다. Y 신호의 경우 그 형태는 그림 4와 같으며 같은 주파수 대역을 가지는 계수 4개를 하나의 셀(cell)로 하여 64개의 셀로 이루어지는 대구획 벡터(superblock vector)를 구성한다.

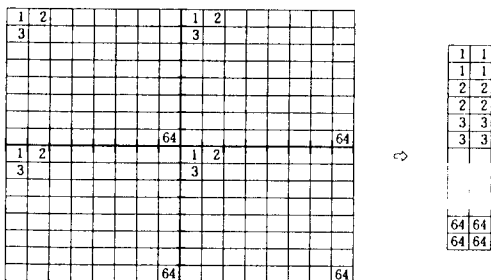


그림 4. 대구획 벡터
Fig. 4. Superblock vector.

여기서 각 셀에 존재하는 4개의 계수를 동일한 양자화기를 사용함으로써 양자화 과정을 단순화시킨다. 그리고 셀이 다른 경우에는 다른 종류의 양자화기를 사용할 수 있게 한다. 이는 DCT 변환 계수의 중요도가 주파수 대역에 따라 변한다는 특성을 고려한다고 볼 수 있다. 본 논문에서는 하드웨어의 복잡성을 줄이고, 부호화를 간략하게 하기 위해 네 종류의 비선형 양자화기로 제한시켰다.

64개의 셀을 하나의 대구획 벡터로 하여 각 셀당 전체 목표 왜곡 레벨에 따라 네가지 양자화기중에서 하나를 선택하면 대구획당 가능한 양자화기 선택 패턴의 가지수는 464개의 양자화기 선택 패턴이 된다. 각 양자화된 계수는 그 계수를 양자화한 양자화기에 따라 부호화하므로 수신단에서 복호화에 이 선택 패턴이 필요하게 된다. 이렇게 되면 양자화된 DCT 변환 계수 이외에 양자화를 하기 위해 사용된 양자화기 패턴도 전송해야 하므로 이를 코드로 표현하면 비트수는 대구획 벡터당 128비트가 되어 부가 정보가 너무 많게 된다. 따라서 대구획 벡터를 표현하기 위한 양자화기의 선택 패턴 수를 제한하는 것이 바람직하다. U, V 신호의 경우에는 8 x 8 DCT 변환 계수 집합을 2H x 3V 배열로 묶어 하나의 대구획 벡터로 정의하여 6개의 DCT 변환 계수가 하나의 셀을 이루게 된다.

$$q = |x - Q(x)| \quad (16)$$

Y 신호의 경우 제한된 수의 양자화기 선택 패턴을 가진 코드북에서 하나의 양자화기 벡터를 선택하는 방법은 다음과 같다. 양자화된 각 변환 계수의 양자화 오차(q)는 다음 식과 같이 정의한다.

$$e = |q - d| \quad (17)$$

여기서 $Q(x)$ 는 셀의 한 요소의 양자화된 값이다. 오차의 가지도를 주파수 대역에 따라 상대적으로 균등하게 하기 위해 인간 시각 특성과 버퍼 상태에 따라 구한 전체 목표 왜곡값과 양자화 오차가 동일하도록 하는 것이 바람직하나, 실체는 완전히 일치한다고 볼 수 없기 때문에 양자화기 선택에 따른 선택 왜곡 오차(e)가 발생한다. 이는 다음 식으로 정의된다.

여기서 d는 전체 목표 왜곡값이며, 각 셀에 대한 총 선택 오차는 2 x 2 셀 요소에 대해 식 (17)에 따른 각각의 선택 오차의 합이고, 대구획 벡터에 대한 총 선택 오차는 64개, 각 셀의 총 선택 오차의 합이다. 적절한 코드북 벡터를 선택하는 과정은 코드북내

의 모든 벡터를 차례로 검색하여 그 중 대구획 벡터의 총 선택 오차를 가장 작게 하는 벡터를 선택하는 것이다. 그림 5는 코드북을 이용하여 대구획 블록의 계수들을 양자화하는 과정을 나타내는 블록도이다.

양자화하려는 신호와 전체 목표 왜곡값과 계수 정규화 값이 선택 오차 신호 계산기에 입력되어 각 양자화기에 따라 양자화된 신호, 그에 따른 선택 오차와 비트 수를 계산하여 선택기로 보낸다. 선택기에서는 앞단에서 받은 정보를 코드북에 저장된 모든 양자화기 벡터를 차례로 검색하여 적절한 값을 조합기로 전송한다. 조합기는 서로 다른 코드북 벡터에 따라 양자화된 대구획 신호와 그것의 총 선택 오차, 그리고 발생하는 비트율을 차례로 문턱값 선택기로 보낸다. 그리고 문턱값 선택기는 그 중 최소 총 선택 오차를 발생하는 양자화기 벡터의 인덱스와 그 벡터에 따라 양자화된 신호를 출력하게 된다. 여기서 양자화는 앞에서 선택된 양자화기 벡터에 따라 각각의 8x8의 DCT 계수값을 한영상에 대해 구한 scale factor로 나누어 정규화한 후 양자화한다. 그리고 복호화시에는 다시 scale factor를 곱하여 역양자화한다. 따라서 scale factor 값을 전송하여야 하며 이를 전송하기 위한 부가정보도 또한 비트율에 포함하여야 한다.

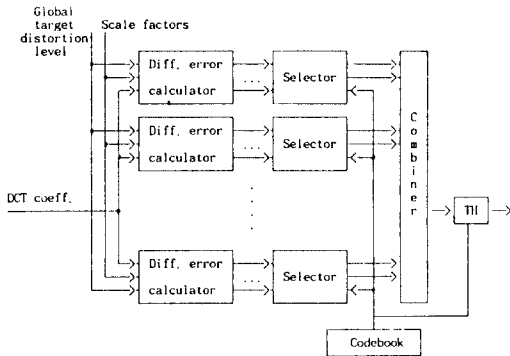


그림 5. 양자화기 벡터 선택기
Fig. 5. Quantizer vector selector.

V. 제한한 양자화기 벡터 코드북 생성 알고리즘

대구획 벡터당 발생 가능한 양자화기 선택 패턴은 4⁶개이며, 이를 부호화하여 전송할 경우 128비트가 추가로 소요된다. 따라서 본 논문에서는 대구획 벡터를 표현하기 위한 양자화기 선택 패턴수를 Y 프레임인 경우에 2048개로, U, V 신호는 각각 512개로 제한하는 방법을 제안하였다. 각각 부호화시 11비트, 9

비트가 필요하게 된다. 이때 코드북의 크기는 구현의 복잡도(complexity), 비트율의 압축, 화질측면을 고려하여 결정하였다. 이는 최적의 벡터를 결정하기 위해서는 코드북내를 모두 탐색하므로 codebook의 크기가 커질수록 구현의 복잡도는 더욱 커지기 때문이다.

실험에 사용된 양자화기는 DCT 변환된 계수들이 Laplacian 분포를 가진다는 특성^[9]과 양자화되는 계수들의 생동폭에 대한 조사를 근거로 하여 사역(dead zone)을 가지는 3 종류의 비선형 Laplacian Max 양자화기(35, 17, 9 레벨)와 계수값을 0으로 매핑하는 양자화기를 사용하였다. 양자화기 종류를 나타내는 심볼(3, 2, 1, 0)로 구성된 수 많은 양자화기 선택 벡터로부터 제한된 수의 양자화기 벡터를 만드는 가장 쉬운 방법은 첫번째 양자화기 벡터를 코드북에 저장하고 그 다음 양자화기 벡터부터는 코드북에 저장하고 있는 벡터와 하나라도 다를 경우에는 무조건 코드북에 저장하는 방법을 생각할 수 있다. 그러나 이 방법은 발생될 수 있는 양자화기 벡터 개수가 너무 많게 되므로 한정된 코드북에 다양한 양자화기 패턴을 저장할 수 없게 된다. 따라서 본 논문에서는 벡터들간의 유사성의 측정자로서 벡터 구성원들간의 절대차의 합을 이용하였다. 절대차 합이 주어진 임계값보다 크면 두 벡터를 다른 벡터로 구별하고 같은 경우는 같은 벡터로 간주한다.

양자화기 선택 벡터의 수를 화질의 큰 저하없이 제한시키기 위해 벡터들간의 절대차 합을 구함에 있어 두가지 사항을 고려하였다. 첫번째는 양자화기의 레벨수에 따라 다르게 가중을 하는 것으로, 양자화기를 나타내는 3, 2, 1, 0 코드로 구성된 벡터들간의 절대차를 구할 때 3과 2, 2와 1, 1과 0 사이의 절대차는 모두 1이나, 같은 1의 오차라도 양자화 레벨수가 많을수록 영상 복원시에 많은 영향을 미치므로 절대차를 크게 하기 위해, 본 논문에서는 각 양자화기의 양자화 레벨수에 해당하는 35, 17, 9, 0의 가중값을 3, 2, 1, 0의 심볼에 부여하여 절대차를 구하였다. 두번째는 양자화기의 주파수 대역에서 상대적인 위치에 따라 절대차에 가중을 다르게 하는 것으로, 저주파 위치의 계수를 양자화하는 양자화기는 고주파 위치의 양자화기보다 영상 복원시에 더 큰 영향을 미치므로 저주파 계수를 양자화하는 양자화기간의 차에 고주파 위치보다 상대적으로 큰 가중을 하였다. 가중함수로서 아래 식을 이용해 양자화기간의 절대차를 선형적으로 저주파에 높게 고주파에 낮게 가중을 하는 것과 식 (1)을 이용해서 비선형적으로 가중하는 것을 이용하였다.

$$W(x) = -0.014596x + 1 \quad (18)$$

두 벡터들간의 제안한 가중된 절대차 합(sum of weighted absolute difference ; SWAD)은 다음 식과 같다.

$$SWAD = \sum_{n=0}^{63} W_n |Q_c(n) - Q_t(n)| \quad (19)$$

여기서 W_n 은 주파수 대역의 상대적인 위치에 따른 가중치이고, $Q_c(n)$, $Q_t(n)$ 은 벡터들의 한 요소로서 양자화기의 종류에 따라 다르게 가중을 한 값이다.

이와 같이 주파수 대역의 상대적인 위치와 벡터의 한 요소에 대해 선택된 양자화기의 종류에 따라 가중치를 달리 부여하여 양자화기 벡터 코드북을 구성한다. 이는 일반적인 벡터양자화 코드북을 만들때 사용되는 입력벡터가 화소값 또는 화소값과 연관된 계수들로 이루지는 반면에 본 논문에서는 성질상 전혀 다른 양자화기를 나타내는 심볼들이 입력벡터로 사용되기 때문에 단순히 심볼들간의 절대차의 합만으로는 보다 다양한 양자화기 벡터 코드북을 구성할 수 없게 된다. 또한 양자화기 벡터 코드북내의 벡터와 선택될 수 있는 실제 벡터와의 선택 오차로 인해 유발되는 영상의 왜곡정도를 인간시각에 보다 덜 띄게 만들기 위해서 가중치를 부여하였다.

수많은 양자화기 선택 벡터로부터 제한된 수의 코드북 벡터를 구성하기 위한 전체적인 알고리즘은 그림 6과 같다. 먼저 초기의 양자화기 선택 벡터를 코드북의 구성벡터로 받고 아래 과정을 수행한다.

- (1) 임계치, 코드북의 첫번째 벡터와 첫번째 훈련 벡터를 입력으로 받는다.
- (2) 코드북 벡터와 후보 벡터 사이의 가중된 절대차의 합(SWAD)을 구한다.
- (3) SWAD가 임계치보다 작으면 후보벡터는 코드북 벡터와 유사한 것으로 간주하고 (4)를 수행하며, SWAD가 임계치보다 크면 현재 코드북 벡터와는 다르므로 코드북내의 다음 벡터와 (2)를 다시 수행한다. 이때 다음의 코드북 벡터가 없을 경우는 이 후보 벡터를 코드북에 저장하고 4)를 수행한다.
- (4) 다음 훈련 벡터를 후보 벡터로 입력하여 코드북 내의 첫번째 벡터와 (2)를 수행한다. 이때 더 이상의 후보벡터가 없고, 코드북의 벡터수가 원하는 갯수(Y:2048, U:512, V:512)일 경우는 여기서 멈추고, 코드북의 벡터수가 원하는 갯수가 아닐 경우는 다시 (1)부터 수행한다.

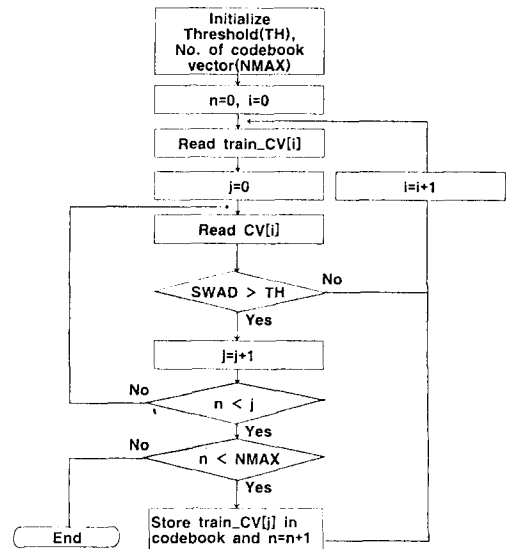


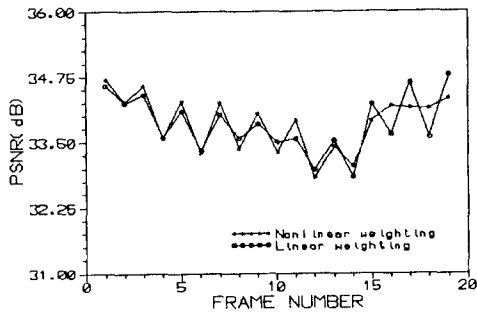
그림 6. 코드북 벡터 생성 순서도
Fig. 6. Flowchart for generating codebook vector.

IV. 실험 및 고찰

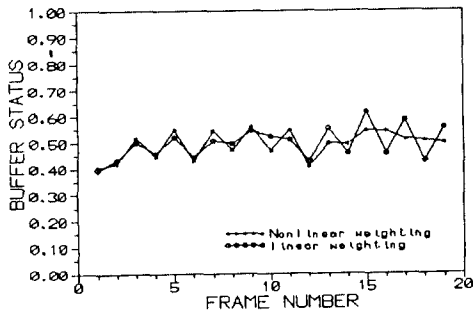
본 논문에서는 HDTV 신호인 1280 x 720 크기의 models 20 프레임, rider 20 프레임과 station 7 프레임의 Y, U, V 칼라 영상을 사용하여 실험을 하였다. 먼저 양자화기 선택 벡터의 훈련 데이터(training data)는 각 셀의 선택 오차를 최소화 하는 양자화기를 선택하여 models 영상과 rider 영상으로부터 구성했으며, 훈련 데이터로부터 제안한 방법으로 Y, U, V 각각 2048, 512, 512개의 코드북을 구성하였다. 본 논문에서는 코드북을 구성하기 위해 벡터간의 절대차 합을 구할 때 주파수 대역의 상대적인 위치에 따라 가중하는 두가지 함수, 비선형, 선형 함수 각각에 따라 코드북을 구성하고, 그 두 코드북에 대한 실험결과를 서로 비교하였다. 초기 코드북 벡터는 훈련 데이터 중 발생 확률이 높은 것을 임의로 10개 정도 선정하여 초기 코드북 벡터로 입력하였다. 실험에 앞서 초기상태로 초기 버퍼의 상태는 정상 상태라 할 수 있는 버퍼 크기의 1/2이 차있고, 이때 버퍼의 크기는 600 kbits 정도로 제한했으며, 첫번째 프레임은 원 영상 그대로 이미 전송되었다는 가정을 하였다. 전송률이 17 Mbps라는 전제하에 버퍼 제어를 하기위해 양자화된 계수와 코드북의 인덱스는 가변길이 부호화를 하였다. 본 알고리즘의 성능

분석은 객관적인 화질 평가 기준인 PSNR 값, 버퍼 상태와 주관적인 화질 평가에 의해 이루어졌다.

그림 7은 rider영상에 대해서 두가지 코드북을 이용하여 복원된 영상의 PSNR 값과 각 프레임당 버퍼 상태를 나타낸다. 그림 7의 (a)에서 전체 PSNR 값이 처음에는 떨어지다가 안정되는 경우를 볼수 있는데 이는 초기 가정중의 하나인 첫번째 프레임이 원영상 그대로 수신단에 전송되었다는 가정에 의해 생기는 현상이다. 안정된 후에 34 - 35dB 정도를 유지함을 볼 수 있다. 그림 7의 (b)에서는 rider 각 프레임의 버퍼 상태를 나타내는데 정상 상태인 버퍼 크기의 반 정도에서 버퍼 상태가 유동함을 볼 수 있다. 그림 8의 (a), (b)는 models 영상에 대한 실험 결과로서 models 영상이 rider 영상보다는 움직임이 작은 반면 영상의 상세함이 높기 때문에 초기 가정에 의한 PSNR의 떨어짐이 상당히 큰 것을 볼 수 있다. 역시 안정된 후에는 33 - 34 dB 정도를 유지한다. Models 각 프레임의 버퍼 상태는 유동의 폭은 커지만 버퍼의 넘침이나 고갈없이 기준레벨에서 유동함을 볼 수 있다. 부가 정보로서 대구획 벡터의 양자화기 선택 벡터의 인덱스를 가변 길이 부호화하였을 때 프레임당 15,000 - 30,000 비트 정도가 발생하였다.



(a)



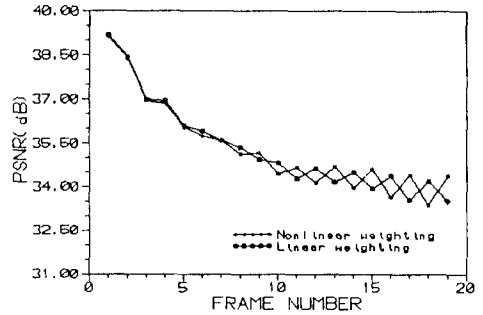
(b)

그림 7. Rider 복원 영상의 PSNR과 버퍼 상태

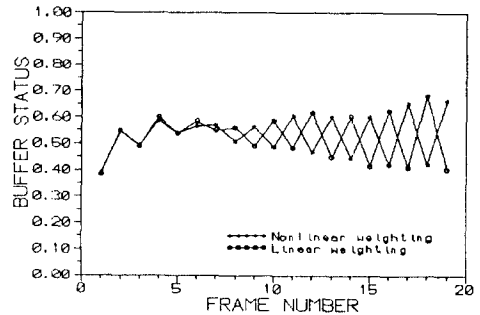
(a) PSNR (b) 버퍼 상태

Fig. 7. PSNR and buffer status of reconstructed "rider" image.

(a) PSNR, (b) Buffer status



(a)



(b)

그림 8. Models 복원 영상의 PSNR과 버퍼 상태

(a) PSNR (b) 버퍼 상태

Fig. 8. PSNR and buffer status of reconstructed "models" image.

(a) PSNR, (b) Buffer status

VII. 결론

본 논문에서는 공간적인 중복성을 제거하기 위해 계층적 블럭 정합 방법을 이용했으며, 공간적인 중복성은 DFD를 DCT 합으로써 제거하였다. 진폭의 중복성을 제거하기 위해 인간 시각 시스템의 특성을 분석하였다.

8 x 8 DCT 계수 각각을 적응적으로 양자화하기 위해 인간 시각 특성과 버퍼 상태에 따라 추정된 양자화 허용 오차 추정치, 즉 전체 목표 왜곡값을 이용하여 각 계수의 양자화기 선택 과정에서 이용하였다. 이때 부호화를 간단히 하기위해 양자화기 종류를 네

가지로 제한하고, 대구획 블록을 구성하였다. 대구획 블록당 양자화기 선택 패턴은 양자화된 계수들을 복호화할 때 이용하므로, 이 수가 너무 많으면 부가 정보가 증가한다. 본 논문에서는 이 수를 제한하는 방법을 제안하였다. 수많은 대구획 블록당 양자화기 벡터로부터 제한된 수의 코드북 벡터를 구성하기위해 벡터간의 유사성의 측정자로서 절대차 합을 이용했으며, 절대차 합을 구할 때 화질의 저하를 최대한 줄이기 위해 양자화기의 레벨수와 주파수 대역의 상대적인 위치에 따라 다르게 가중을 하여 절대차 합을 구하였다. 각 계수들은 인간 시각 특성과 버퍼의 상태에 따라 적응 양자화되고, 이때 대구획 블록당 양자화기 선택 패턴의 수를 제안한 방법으로 Y 신호는 2048개, U, V 신호는 각각 512개로 제한함으로써 화질의 큰 열화없이 부가 정보의 양을 감축시켰다.

실험 결과 객관적인 화질 평가 기준인 PSNR이 양호할 뿐만 아니라 주관적으로도 화질의 열화를 거의 관측할 수 없었다. 또한 버퍼의 제어 측면에서도 정상 상태 레벨에서 안정되었다. 향후, 제한된 수의 양자화기 선택 벡터를 이용하여 HDTV 신호를 적응 양자화하는 방법에 대해 더 많은 연구가 필요하다.

參 考 文 獻

- [1] Knauer et al, "HDTV encoder with forward estimation and constant rate motion vector," *U.S. Patent*, no.5,144, 423, Sep. 1, 1992.
- [2] N. B. Nill, "A visual model weighted cosine transform for image compression and quality assessment," *IEEE Transactions on Communications*, vol. com-33, pp.551-557, June, 1985.
- [3] D. L. McLaren and D. T. Nguyen, "Removal of subjective redundancy from DCT-coded images," *IEE Proceedings-I*, vol.138, no.5, pp.345-350, 1991.
- [4] A. N. Netravali and B. Prasada, "Adaptive quantization of picture signals using spatial masking," *Proceeding of the IEEE*, vol.65, no.4, April, 1977.
- [5] K. R. Rao and P. Yip, *Discrete Cosine Transform ; Algorithm, Advantages and Applications*, Academic Press.
- [6] Johnston et al, "High definition television coding arrangement with graceful degradation," *U.S. Patent*, no.5,128,756, Jul. 7, 1992.
- [7] Z. Xie and T. G. Stockham, "Toward the unification of three visual laws and two visual models in brightness perception," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 19, no.2, pp.379-387, March, 1989.
- [8] J. Zdepski, D. Raychauhuri, and K. Joseph, "Statistically based buffer control policies for constant rate transmission for compressed digital video," *IEEE Transactions on Communications*, vol.39, no.6, June, 1991.
- [9] R. C. Reininger and J. D. Gibson, "Distribution of the two-dimensional dct coefficients for image," *IEEE Transactions on Communications*, vol.

著 者 紹 介



金 益 煥((正會員)
1992年 2月 경북대학교 전자공학과(공학사). 1994年 2月 경북대학교 전자공학과(공학석사). 1994年 (주) 금성사 TV 연구소 연구원. 주관심 분야는 영상압축, HDTV 등임.



崔 振 秀(正會員)
1968年 10月 15日生. 1990年 2月 경북대학교 전자공학과(공학사). 1992年 2月 경북대학교 전자공학과(공학석사). 1994年 3月 경북대학교 전자공학과 박사과정 수료. 주관심 분야는 영상압축, HDTV, 영상처리 등임.



李 光 春(正會員)
1944年 5月 15日生. 1970年 2月 인하공과대학 전자공학과(공학사). 1970年 (주) 금성사 입사현재 同社 TV 연구소 연구위원, 한국센서학회 이사. 주관심 분야는 HDTV, 신호처리 등임.



朴 吉 欽(正會員)
1958年 11月 24日生. 1982年 2月 경북대학교 전자공학과(공학사). 1984年 2月 한국과학기술원 전기 및 전자공학과(공학석사). 1990年 2月 한국과학기술원 전기 및 전자공학과(공학박사). 1984年 4月 ~ 현재 경북대학교 전자공학과 조교수. 주관심 분야는 영상압축, 패턴인식, 영상데이터베이스 등임.

河 永 浩(正會員) 第 31卷 B編 第 5號 參照
현재 경북대학교 전자공학과 부교수