

웨버의 법칙을 고려한 DCT 영상 부호화에 관한 연구

正會員 李 恩 國* 正會員 김 장 복**

A Study on the DCT Image Coding Considering Weber's law

Eun Gug Lee*, Chang Bock Kim** *Regular Members*

要 約

본 논문에서 인간의 시각특성을 고려한 적응적 DCT 영상 부호화 방법이 제안되었다. 인간의 시각은 어두운 부분에서는 잡음에 민감하고, 밝은 부분에서는 잡음에 상대적으로 덜 민감하다. 이러한 특성은 Weber의 법칙으로 설명된다. 즉, Weber의 법칙은 인간의 감지가능차(j.n.d)가 영상의 밝기에 비례함을 말한다. 그러므로 이러한 특성이 영상처리에 응용될 때 밝은 영역에서 잡음을 줄이는 것 보다는 어두운 영역의 잡음을 줄이는 것이 중요하다. 따라서 본 논문에서는 DCT 계수의 AC 계수를 그 블록의 평균적 밝기를 의미하는 DC 계수로 정규화함으로써 상대적으로 어두운 부분의 AC 계수를 세밀하게, 밝은 부분의 AC 계수를 거칠게 양자화하였다. 주관적 평가 결과 낮은 비트율에서, 제안된 부호화 방법이 Weber의 법칙을 고려하지 않은 방법보다 향상됨을 알 수 있었다.

ABSTRACT

In this paper, a DCT image coding algorithm using the human visual property is proposed. Human visual is relatively sensitive to noise in the darker region, insensitive to noise in the brighter region. This property was proved by Weber's law through psycovisual experiment. Weber's law states that the just noticeable difference(j.n.d.) is propotional to intensity. Therefore, the implication of this observation for image processing is that reducing noise in the darker region is more important than reducing noise in the brighter region. In this proposed coding scheme AC coefficients in the darker region are more finely quantized than those in the brighter region. Results showed that, at low bit rate, the subjective quality of reconstructed images by proposed coding scheme is improved than that of coding scheme without considering human visual property.

I. 서 론

*大宇電子 映像研究所
**弘益大學校 電子工學科
論文番號 : 93-67

영상 정보의 통신을 실현하기 위해서는 방대한 양의 디지털 영상 정보를 저장하거나 대역폭이 한정된 채널을 통해 전송해야 하므로 영상 데이터의 압축이

질실히 요구된다.^(1,9)

영상에 적용되는 압축기법은 영상내에 중복되는 정보가 많기 때문에 이러한 중복성(redundancy)을 제거함으로써 데이터를 압축하는데 일반적으로 영상 부호화에 중복성을 제거하는 방법으로는 예측 부호화(Predictive Coding), 변환 부호화(Transform Coding), 혼합 부호화(Hybrid Coding)방식 등으로 대별할 수 있다. 예측 부호화 방식은 현재 부호화할 화소(Pel: picture element)의 실제값과 이미 부호화된 화소들로부터 얻을 수 있는 예측값과의 차이를 전송하는 방식으로 DM, DPCM, ADPCM 등이 있으며, 화소당 1~2 비트까지 압축이 가능하고 하드웨어 구현이 비교적 간단하여 실시간 처리가 가능하지만 전송 오차가 누적된다는 단점이 있다. 변환 부호화 방식은 영상을 다른 영역으로 변환할 경우 그 변환 계수가 일부 특정 영역에 집중됨을 이용하여 불필요한 변환 계수 성분을 제거함으로써 압축 효과를 가져온다. 이러한 변환중에는 WHT(Walsh-Hadamard Transform), KLT(Karhunen-Loeve Transform), DCT(Discrete Cosine Transform) 등이 있으며, 화소당 1비트 이하의 낮은 비트율로 압축이 가능하고 전송시의 잡음에 의한 오차가 분산되는 특성을 가지고 있으나 하드웨어 구현이 복잡하여 실시간 처리가 어려운 것이 단점이다. 변환 부호화 방법중 KLT는 에너지 집중도와 계수들 간의 상관성 제거에 있어서는 최적변환이지만 커널(kernel)이 영상 의존적이며 효율적인 고속 알고리즘이 없으므로 그 성능이 KLT와 유사하고 고속 알고리즘이 존재해 실시간 처리가 가능한 DCT가 데이터 압축, 필터링 등에 많이 사용된다.⁽²⁾ 혼합 부호화 방식은 예측 부호화 방식과 변환 부호화 방식을 적당히 혼합하여 각각의 단점들을 보완한 방식으로 DCT/DPCM, WHT/DPCM 등이 있다.

이 외에도 여러 개의 샘플을 하나의 벡터로 취급하여 양자화한 벡터 양자화(Vector Quantization)방법과 실시간 처리가 가능하고 하드웨어 구현이 비교적 용이한 BTC(Block Truncation Coding)방법 등이 있으며, 최근에는 인간 시각 체계에 입각한 고압축을 위한 제 2세대 부호화 기법등이 활발히 연구되고 있다.^(8,13)

본 논문에서는 기본적으로는 변환 부호화 방식의 하나인 DCT를 이용하고 인간의 시각에 좀 더 적합한 영상 부호화에 대하여 연구하였다. 이를 위하여 영상의 변환 부호화에서 최적이라고 알려져 있는 8×8 또

는 16×16 크기의 블럭에 대해 변환을 수행하였다.⁽⁵⁾ 그리고 인간의 시각특성을 반영하기 위하여 각 블럭의 AC 계수를 DC 계수로 정규화하였으며, 양자화 방법은 Chen과 Smith가 제안한 방식⁽³⁾을 이용하여 각 블럭을 변화도에 따라 네개의 부류로 나누어 비트 맵과 양자화를 그 블럭에 맞게 적용적으로 실행하였다.

II. 영상신호와 시각현상

인간의 시각은 빛의 세기에 대하여 비선형적 반응 특성을 나타내며, 또한 공간 주파수의 모든 성분에 동일한 반응을 보이는 것이 아니라 특정 대역에서 보다 더 민감하게 반응한다는 특성을 이용하면 불필요한 성분은 제거하고 반대로 필요한 성분은 보존함으로써 보다 효율적인 부호화를 할 수가 있다.

흑백 영상의 세기(intensity)는 식(1)과 같이 표현된다.⁽⁶⁾

$$I = k \int_0^{\infty} C(\lambda) S_{BW}(\lambda) d\lambda \quad (1)$$

여기서 $C(\lambda)$ 는 파장이 λ 이고 특정 평면을 통과하는 파의 에너지 분포이며, $S_{BW}(\lambda)$ 는 입력 센서의 스펙트럼 특성이고 k 는 비례상수이다. I 는 흑백영상의 밝기를 나타내게 된다. 식(1)은 단위 면적 당 전력이므로 이 값은 항상 유한한 양수가 된다.

색의 밝기 차이에 대하여 인간의 눈의 반응이 비선형임은 이미 알려져 있다. 같은 형태의 두 물체를 밝기나 빛의 강도의 차이만으로 인간이 얼마나 잘 구별해 낼 수 있는가하는 판정 방법의 하나로 감지가능차(j.n.d.: just noticeable difference)라는 방법이 있다. 감지가능차는 여러가지 방법에 의하여 정의되기도 하고 측정될 수도 있다. 심리적인 실험을 통한 한 가지 방법으로 다음의 그림 1과 같은 시각 자극체를 가정하자. 내부영역은 균일한 밝기의 단색이고 밝기는 I 나 $I + \Delta I$ 로 임의로 바꿀 수 있다. 외부영역도 단색의 균일한 밝기로 내부의 밝기가 I 일 때는 $I + \Delta I$ 로 혹은 내부의 밝기가 $I + \Delta I$ 일 때는 I 로 선택할 수 있게 되어 있다. 관찰자에게 이 밝기의 변화과정에서 안과 밖 중 어느 쪽이 더 밝은가를 질문한다. ΔI 가 클 경우에는 대부분의 응답자는 $I + \Delta I$ 가 더 밝다고 바른 대답을 한다. 그러나 ΔI 의 크기가 줄어 들수록 바른 선택의 비율은 반 정도로 줄어 든다. ΔI 의 크기를 줄일수록 응답자의 바른 응답의 비율이 줄어드는데 응답자의 약 75%가 바른 대답을 하였을 때의 ΔI

를 I에서의 감지가능차로 정의할 수 있다.⁽⁶⁾

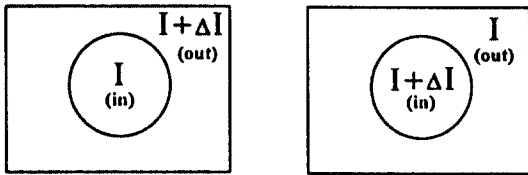


그림 1. 밝기의 판별의 실험에 사용되는 2개의 시각 자극체

Fig. 1. Two stimuli used in an intensity discrimination experiment

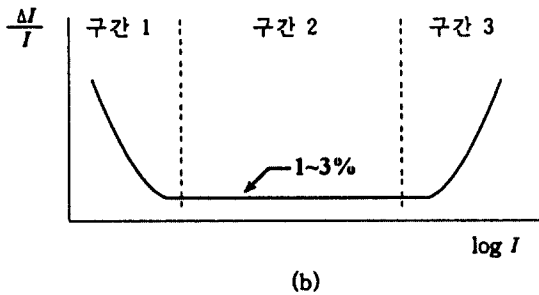
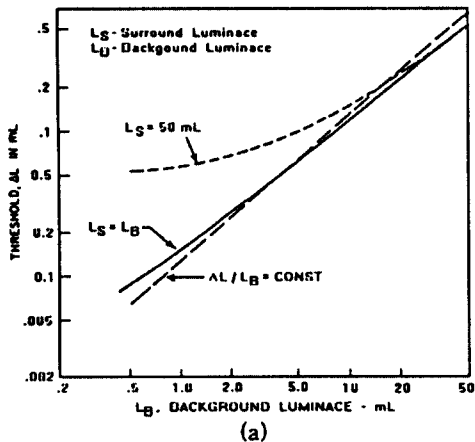


그림 2. Weber의 법칙
Fig. 2. Weber's law

그 결과가 그림 2에 나와 있다. 그림 2에서는 I의 넓은 범위에서 다음 관계식이 성립한다.

$$\frac{\Delta I}{I} \approx constant, \quad 1 \sim 3\% \quad (2)$$

식(2)의 관계를 Weber의 법칙이라 한다.⁽¹⁰⁾ Weber의 법칙에 따르면 ΔI 는 I에 비례한다. Weber의 법칙은 시각뿐만 아니라 인간의 다른 감각 기관(예를 들면, 청각, 후각, 미각, 촉각 등)에도 모두 적용되는 법칙으로 알려져 있다. 식(2)에서 만일 ΔI 가 작은 값이라면,

$$\frac{dI}{I} = d(\log I) \approx constant \quad (3)$$

위의 식(3)에서 살펴 볼 수 있는 바와 같이 I의 넓은 범위에 걸쳐서, 감지 가능차의 $\Delta I / I$ 는 일정하다.

이 밝기의 판별에 대한 실험 결과는 I가 증가할수록 ΔI 도 따라서 증가한다는 것을 보여준다. 이러한 현상들은 영상 처리등에 이용될 수 있는데, 같은 크기의 잡음이 영상에 더하여지면 그 영향은 명암의 대비가 큰 영역보다는 작은 영역에서 더 두드러지게 나타난다. 또한 밝은 영역보다는 어두운 영역에서 잡음은 눈에 띄기 쉬우며, 밝고 균일한 영역에서의 잡음이 어둡고 명암 대비가 큰 영역에서 보다 눈에 더 잘 띄게 된다. 이러한 현상은 Weber의 법칙등으로 설명될 수 있으며, 명암 대비가 큰 대부분에서는 신호의 크기가 크며 잡음을 더 많이 마스크하는 공간 마스크 현상으로 설명된다.⁽¹⁴⁾

이러한 문제를 보완하기 위하여 다음의 가정을 세우고, 이에 따른 양자화 방법을 모색해 본다.

가정 1. 인간의 시각체계는 밝기 성분 보다는 색의 변화에 덜 민감하다. 즉 밝기 정보만을 갖고 있는 흑백 영상만으로도 충분히 시각 정보를 얻을 수가 있다. 그리고 색의 밝기에 따라 변화에 따른 인간의 시각적 감지 능력이 다르므로 이러한 특성을 이용하여 영상을 부호화하는 것은 타당하다.⁽¹⁰⁾ 또한 DCT 영상 부호화 방법에서 영상의 질은 비트할당으로 인한 양자화 오차가 지배적이다. 이러한 양자화로 인해 생기는 잡음에 대하여 인간의 시각적인 특성을 이용하는 것은 중요하다.

가정 2. 인간의 시각이 일상적으로 접하는 빛의 밝기가 그림 2의 구간 2에 존재한다고 볼 수 있으며⁽⁹⁾, 이 구간에서 Weber의 법칙 $dI/I = const.$ 는 일리가 있다. 따라서, 각 블록의 DCT 계수들을 분석해 볼때 DC 계수는 그 블록의 평균적 밝기를 의미하며, AC 계

수는 그 블록의 밝기 변화에 대한 정도를 나타낸다. 따라서 DC 성분을 I, AC 성분을 ΔI 로 가정하는 것은 타당하다.

$$\frac{\Delta I}{I} = \frac{F[u,v]}{F[0,0]}, \quad \text{단, } u,v \neq 0 \quad (4)$$

가정 3. 그림 2에서 구간 1의 경우, 실제로 정상적인 화상에서는 매우 어두운 화상으로 나타나고, 시각정보도 거의 없다고 가정한다. 물론 데이터 자체를 가지고 취급하는 경우에는 밝기를 강조한다든지 하면 이 부분의 정보도 나타내야 한다. 그러나 본 논문에서 취급하는 화상데이터의 압축에는 이 경우를 제외한다. 이 부분을 전체 밝기의 10% 정도로 한다. 즉 본 논문에서 취급하는 영상의 화소의 밝기는 256 단계이므로 20~30 단계까지를 이 구간으로 한다.

가정 1,2,3을 종합하면 256 단계의 영상에서, 각 블록의 DCT 계수들은 양자화 잡음에 대하여 눈에 더 잘 띄는 어두운 부분에서는 세밀하게(finely) 양자화함으로써 오차를 줄이고, 상대적으로 잡음에 대하여 인간 시각이 둔감하게 반응하는 밝은 부분에서는 거칠게(coarsely) 양자화함으로써 시각적으로 보다 나은 결과의 영상을 얻게 된다.

본 논문에서 사용하는 2차원 DCT는 단순히 1차원 DCT에서 2차원으로 확장함으로써 얻을 수 있다. 2차원 DCT는 두 번의 1차원 DCT로 구해질 수 있는데 1차원 DCT의 고속 알고리즘은 Chen등에 의해서 제시되었다.

따라서 $N \times N$ 2차원 DCT의 정식은 다음과 같다.

$$F[u,v] = \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} 4f(i,j) \cos \frac{(2i+1)u\pi}{2N} \cos \frac{(2j+1)v\pi}{2N}, \quad 0 \leq u,v \leq N-1 \quad (5a)$$

$$f(i,j) = \frac{1}{N^2} \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} w(u)w(v)F(u,v) \cos \frac{(2i+1)u\pi}{2N} \cos \frac{(2j+1)v\pi}{2N}, \quad 0 \leq i,j \leq N-1 \quad (5b)$$

이러한 2차원 DCT를 적용하는 과정은 1차원 DCT와 그 역변환을 먼저 각 행에 대해서 적용하고 그 결과를 이용 각 열에 대해 적용하여 식(5)와 같이 2차원 DCT를 구현하게 된다.⁽⁷⁾

III. 부호화 방법

인간의 시각 특성 중 빛의 세기에 대한 비선형적 반응 특성, 즉 Weber의 법칙은 변환 부호화에 효과적으로 응용될 수 있다. 즉, 변환된 각 계수에 대하여 Weber의 법칙을 적용한 후, 부호화하면 시각에 좀더 적합한 영상을 얻을 수가 있다.⁽⁶⁾

한 블록에서 변환된 계수들은 DC성분과 AC성분으로 나눌 수 있는데 DC성분은 그 블록의 평균을 의미하며, AC성분은 그 블록의 변화도를 의미한다. 따라서 DC성분을 조사함으로써 그 블록의 평균적 밝기를 알 수가 있다. 식(7)과 같이 각 블록의 AC 계수들을 DC 계수로 정규화시키면 그림 3에서 보여주는 것과 같이 서로 다른 블록의 어느 한 AC 성분의 크기가 같을 경우, 그 블록의 평균적 밝기가 밝은 부분보다 어두운 부분, 즉 DC 계수의 값이 큰 부분보다 작은 부분의 AC 계수가 상대적으로 세밀하게 양자화된다. 따라서 인간의 시각이 어두운 부분에서 잡음에 대하여 더욱 민감하므로 이러한 시각 특성이 고려된 양자화를 할 수가 있다.

$$AC \text{ 계수} = \frac{AC \text{ 계수}}{DC \text{ 계수}}, \quad \text{단, } DC \text{ 계수} \geq 20 \quad (6)$$

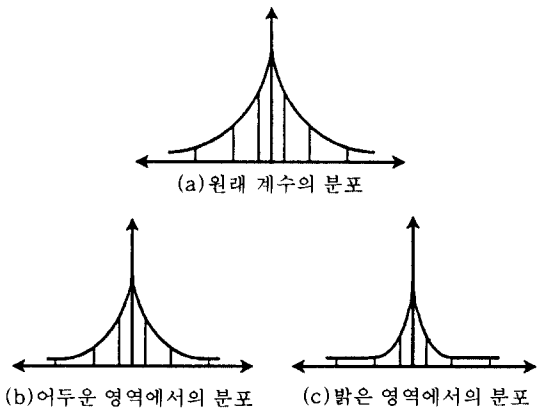


그림 3. 시각특성을 고려한 양자화 특성(8 level의 경우)
Fig. 3. Quantization characteristics considering visual property

시각특성에 맞게 정규화된 DCT 계수들은 비트할당에 의하여 서로 다른 비트수 $[l_{(m,n)}]$ 로 양자화된다. 따라서 1 비트에서 8 비트까지의 서로 다른 양자화기가 8개 필요하다. 이러한 양자화기를 사용할 때 입력 신호는 먼저 단위분산(unit variance)을 갖도록 정규화된다.⁽¹⁰⁾ 즉,

$$\bar{F}_i(u,v) = \frac{F_i(u,v)}{\sigma(u,v)}, \quad (u,v) \neq (0,0) \quad (7)$$

이러한 계수들은 평균이 0(zero mean)이고 분산이 1(unit variance)인 입력을 위해 설계된 $l_{(m,n)}$ -비트의 양자화기로 양자화된다. 비트가 할당되지 않은 계수들은 전혀 처리되지 않으며 수신단에서 비트할당표를 참고하여 처리되지 않은 계수들은 0으로, 즉 그 계수의 평균 값으로 대신한다.

양자화를 위한 비트 할당 방법은 비율왜곡이론(rate distortion theory)을 이용해 비트 할당표를 구성하였으며, Chen과 Smith가 제안한 적응적 부호화 방식⁽³⁾을 도입하여 각 블럭을 변화도에 따라 4개로 분류하여 각각 비트 할당표를 달리하였다.

다음은 비율왜곡이론에 의한 비트할당 방법이다.⁽¹²⁾ n개의 계수중 i번째에 b_i 비트를 할당했을 때 오차 및 평균자승오차는 다음과 같다.

$$E_i = KR^{-p} = K2^{-pb_i} \quad (8)$$

$$E_{av} = \frac{1}{N} \sum_{i=1}^N \sigma_i^2(u,v) K2^{-2b_i} \quad (9)$$

여기서 $R=2^b$, $p=2$, K는 상수, N은 한 블럭의 계수의 갯수이며, σ_i^2 는 i번째의 분산이다.

식(9)에서 E_{av} 를 최소로 하기위하여 총 할당 비트를 B이라 하면

$$\sum_{i=1}^N b_i = B \quad (10)$$

이고, 다음과 같이 상수 D로 바꾸면

$$\sigma_i^2(u,v) K2^{-2b_i} = D \quad (11)$$

이므로 각 분산 계수에 대한 비트 할당식은 다음과 같다.

$$b_i = \frac{1}{2} (\log_2 \sigma_i^2(u,v) - \log_2 D) \quad (12)$$

식(12)를 식(10)에 대입하면

$$\begin{aligned} 2B &= \sum_{i=1}^N (\log_2 \sigma_i^2(u,v) - \log_2 D) \\ &= \sum_{i=1}^N \log_2 \sigma_i^2(u,v) - N \log_2 D \end{aligned} \quad (13)$$

이므로 식(12)와 (13)으로 부터 다음과 같이 정리할 수 있다.

$$\begin{aligned} b_i &= \frac{1}{2} [\log_2 \sigma_i^2(u,v) - \frac{1}{N} \{ \sum_{i=1}^N \log_2 \sigma_i^2(u,v) - 2B \}] \\ &= \frac{B}{N} + \frac{1}{2} [\log_2 \sigma_i^2(u,v) - \frac{1}{N} \sum_{i=1}^N \log_2 \sigma_i^2(u,v)] \end{aligned} \quad (14)$$

따라서 결과적으로 i번째 계수에 할당될 비트수는 식(15)과 같다.

$$b_i = \frac{B}{N} + \frac{1}{2} [\log_2 \sigma_i^2 - \frac{1}{N} \log_2 (\prod_{i=1}^N \sigma_i^2)] \quad (15)$$

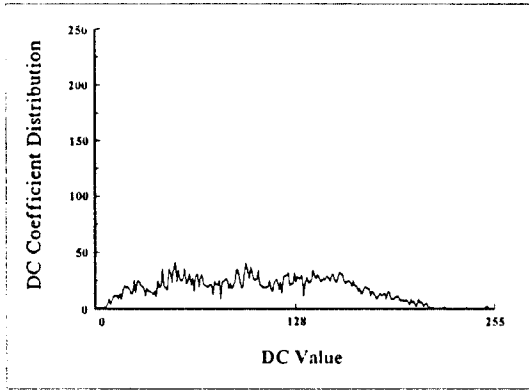
한편, 총 평균자승오차(MSE)는 할당된 비트에 대한 양자화 오차와 할당되지 않은 계수로 인한 절단오차(truncation error)로 구성된다.

본 논문에서는 Lloyd-Max 양자화기를 사용하였으며 양자화를 위해 사용한 DCT 계수들의 확률분포로는 DC 계수에는 Uniform 밀도함수를 사용하였고, 나머지 AC 계수는 Laplacian 밀도함수를 사용하였다. 식(16)은 본 논문에서 사용한 AC 계수에 대한 확률밀도함수이다.⁽¹²⁾

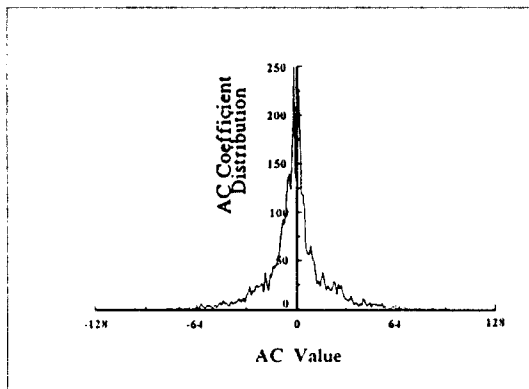
$$P_{(u,v)}(x) = \frac{1}{\sigma^2(u,v)} \exp\left[-\frac{2}{\sigma^2(u,v)} |x|\right] \quad (16)$$

그림 4는 영상의 DCT 계수들의 분포를 보여준다. 그림 4(a)는 DC계수에 대한 확률분포이며 (b)는 AC 계수에 대한 확률분포를 나타낸다. 그림에서 알 수 있듯이 DC 계수와 AC 계수의 확률 분포가 수 많은 샘플을 대상으로 확률분포를 구할때에 각각 Gaussian, Laplacian 밀도함수의 형태에 접근함을 알 수가 있다. 따라서 본 논문에서는 DC 계수에는 경계면의 휘도 차이를 피하기 위해 8 비트를 할당하여 Uniform

밀도함수를, 나머지 AC 계수는 Laplacian 밀도함수를 적용할 수가 있다.¹¹⁰⁾



(a) DC 계수의 분포

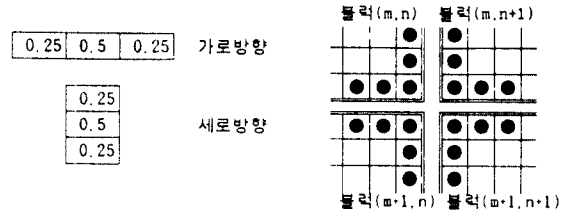


(b) AC 계수의 분포

그림 4. 실험 영상의 DCT 계수 분포
Fig. 4. DCT coefficient distribution of test images

DCT를 사용한 영상 부호화 방식에서는 영상전체를 한번에 변환하지 않고 블록으로 나누어 처리를 행한다. 따라서 고압축률로 영상을 부호화할 경우, 블록간의 경계부분에 불연속적인 경계가 발생하는 블럭화 현상을 줄이기 위한 후처리로, 복원영상의 블럭의 경계부분에 대하여 LPF(Low Pass Filter)를 취하는 방안을 사용하였다. 전체 영상에 LPF를 취하면 영상이 전체적으로 흐려지는 현상이 나타나지만, 블럭의 경계부분에만 LPF를 취할 경우 그러한 현상을 크게 눈에 띄지 않는다. 그림 5의 (a)는 블럭화 현상 보정을 위한 LPF 마스크를 나타낸다.⁴⁾ (b)는 후처리를 하지않은 결과, (c)는 후처리를 행한 결과의

그림이다.



(a) 경계면 LPF 마스크



(b) 후처리를 하지않았을 경우



(c) 후처리를 하였을 경우

그림 5. 블럭화 현상 보정을 위한 후처리
Fig. 5. Post processing for adjusting the blocking effect

IV. 실험 및 결과 고찰

본 논문에서는 Chen 등이 제안한 부호화 방식³⁾에 인간의 시각 특성을 도입하여 실험을 하였다.

그림 6과 그림 7은 제안 부호화 방식의 전체 계통도이다.

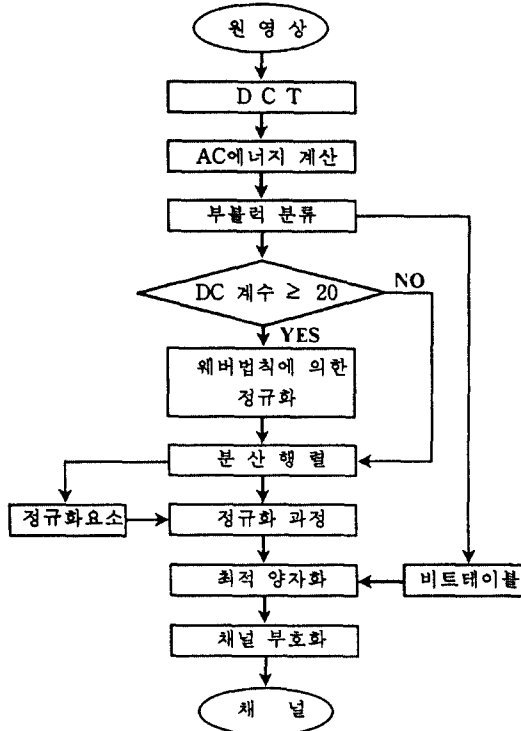


그림 6. 부호화 과정
Fig. 6. Encoding process

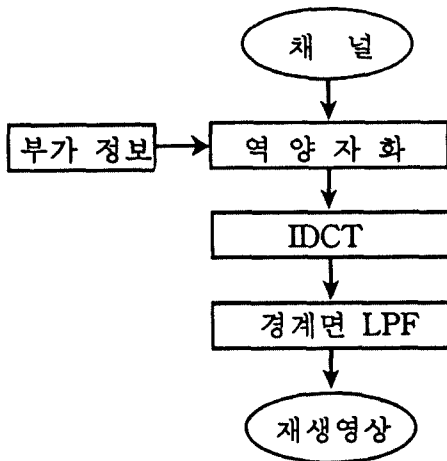


그림 7. 복호화 과정
Fig. 7. Decoding process

부가 정보는 압축된 영상 정보를 복호화하기 위한 정보값으로 본 논문에서 필요한 부가 정보는 활동도 분류표, 정규화 요소, 네개의 비트할당표이다.

이러한 부가 정보는 식(17)과 같다.

$$B_T = B_{CM} + B_{NF} + B_{BT} \quad (17)$$

단, B_T : 총 부가 정보량

B_{CM} : 활동도 표(Classification Map)

B_{NF} : 정규화 요소(Normalization Factor)

B_{BT} : 비트 할당표(Bit Allocation Map)

512×512 크기의 영상을 전송하는데 필요한 부가 정보는 다음과 같다.

$$B_{CM} = \frac{\text{블럭 수} \times 2 \text{ 비트}}{512 \times 512}$$

$$B_{NF} = \frac{4 \times N \times N \times 8 \text{ 비트}}{512 \times 512}$$

$$B_{BT} = \frac{4 \times N \times N \times 8 \text{ 비트}}{512 \times 512}$$

정규화 요소의 값은 좀 더 정확한 값을 보존하기 위하여 각 계수에 대한 분산행렬로부터 정규화된 값을 보존한다. 각 부가 정보는 부블럭이 16×16의 경우 $B_{CM} = 0.0078 \text{ bits/pel}$, $B_{NF} = 0.0312 \text{ bits/pel}$, $B_{BT} = 0.0312 \text{ bits/pel}$ 로 총 부가 정보는 $B_T = 0.0702 \text{ bits/pel}$ 이 된다. 그러나 송수신단에서 압축률을 고정시킬 경우 비트할당표에 대한 정보는 부가정보로 취급하지 않으므로 그 때의 총 부가정보는 $B_T = 0.0390 \text{ bits/pel}$ 이 된다.

실험에 사용한 영상은 512×512 해상도의 4개의 영상(LENA, PEPPER, MAN, CROWD)을 사용하였으며, 그림 8은 LENA 영상, 그림 9는 PEPPER 영상이며 각각 0.5 bpp의 비율로 압축된 영상을 재생한 데이터이다. (a)는 원영상의 사진이고, (b)는 단순히 DCT만을 이용한 것이며, (c)는 Chen방식으로 부호화한 것이고, (d)는 후처리를 하지않은 제안한 방식의 사진이다.

영상의 평가방법에는 객관적 방법과 주관적 방법이 있다. 수치적으로 품질을 평가하는 객관적 방법에는 식(18)의 PSNR(Peak to peak signal to noise ratio)을 사용하였다. 표 1과 그림 10은 각 방식에 대한 객관적 성능 평가 결과를 나타낸다. PSNR을 사용하는 객관적 평가는 신호의 에너지에 대한 충실도

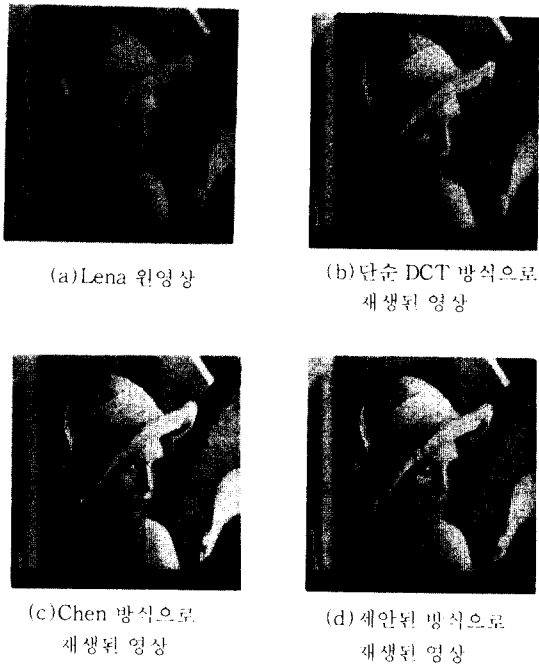


그림 8. 각 방식으로 부호화된 LENA 영상 (0.25bpp, 16×16)

Fig. 8. LENA images coded by each method (0.25bpp, 16×16)

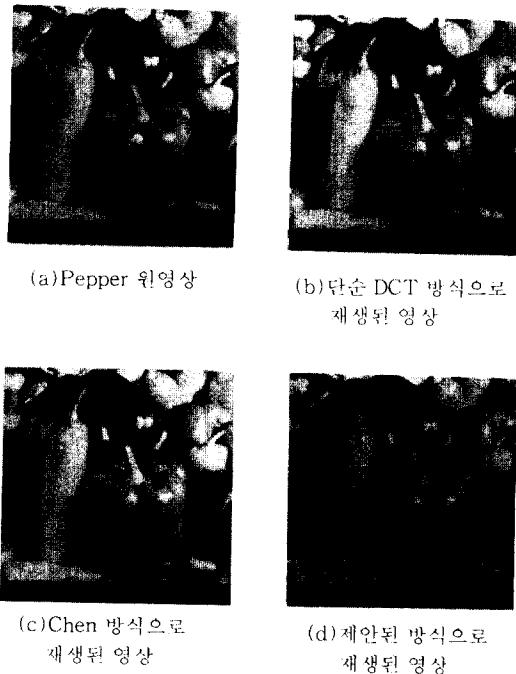


그림 9. 각 방식으로 부호화된 PEPPER 영상 (0.25bpp, 16×16)

Fig. 9. PEPPER images coded by each method (0.25bpp, 16×16)

는 효과적으로 반영하지만 사람의 시각 특성은 정확하게 반영하지 못한다. 따라서 시각적으로 더 우수한 영상이 PSNR은 더 낮은 경우가 발생할 수 있으며, 표 1로부터 제안한 방식이 PSNR면에서는 기존의 방식보다 낮은 결과를 보임을 알 수 있다.

$$PSNR = 10 \log_{10} \frac{255^2}{MSE} \text{ (dB)} \quad (18)$$

$$\text{단, } MSE = \frac{1}{N^2} \sum_{i=1}^{M+1} \sum_{j=1}^{N+1} [f(i,j) - f'(i,j)]^2$$

표 1. 객관적 평가 결과

Table 1 Results of objective estimation

(a) LENA 영상 (8×8)

방법 bpp	단순 DCT	Chen	Proposed	Proposed with LPF	PSNR
0.25	27.20	29.19	29.07	28.95	dB
0.50	29.83	32.92	32.44	32.32	dB
1.00	33.18	35.92	35.30	35.21	dB

(b) LENA 영상 (16×16)

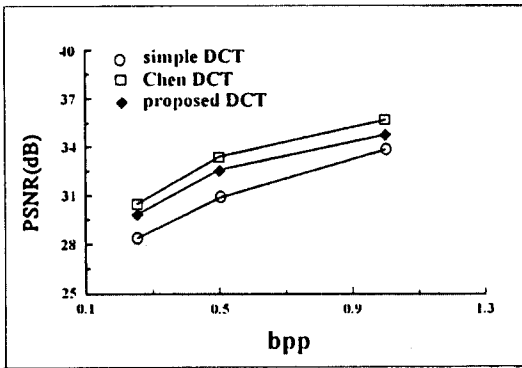
방법 bpp	단순 DCT	Chen	Proposed	Proposed with LPF	PSNR
0.25	28.32	30.42	30.12	29.79	dB
0.50	30.83	33.38	32.76	32.56	dB
1.00	33.84	35.69	34.89	34.74	dB

(c) PEPPER 영상 (8×8)

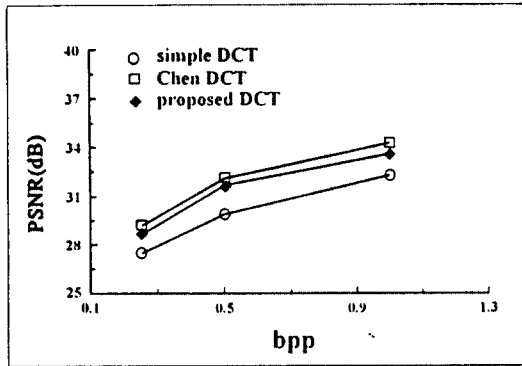
방법 bpp	단순 DCT	Chen	Proposed	Proposed with LPF	PSNR
0.25	26.67	28.72	28.22	27.91	dB
0.50	29.04	31.91	31.25	31.03	dB
1.00	31.96	34.54	33.73	33.64	dB

(d) PEPPER 영상 (16×16)

방법 bpp	단순 DCT	Chen	Proposed	Proposed with LPF	PSNR
0.25	27.44	29.17	28.91	28.64	dB
0.50	29.89	32.12	31.83	31.69	dB
1.00	32.31	34.28	33.69	33.60	dB



(a) LENA 영상 (16×16)



(b) PEPPER 영상 (16×16)

그림 10. 객관적 평가 결과

Fig. 10. Results of objective estimation

표 3. 주관적 평가 결과

Table 3 Results of subjective estimation

(a) LENA 영상 (8×8)

방식 bpp	단순 DCT	Chen	Proposed	Proposed with LPF	PSNR
0.25	2.2	3.2	3.4	3.5	등급
0.50	3.8	4.7	4.9	4.9	등급
1.00	5.2	6.4	6.4	6.4	등급

(b) LENA 영상 (16×16)

방식 bpp	단순 DCT	Chen	Proposed	Proposed with LPF	PSNR
0.25	2.0	3.2	3.7	3.9	등급
0.50	4.0	4.8	5.1	5.2	등급
1.00	5.0	6.5	6.6	6.6	등급

(c) PEPPER 영상 (8×8)

방식 bpp	단순 DCT	Chen	Proposed	Proposed with LPF	PSNR
0.25	1.8	2.7	3.0	3.1	등급
0.50	4.2	4.9	5.2	5.3	등급
1.00	5.3	6.4	6.6	6.6	등급

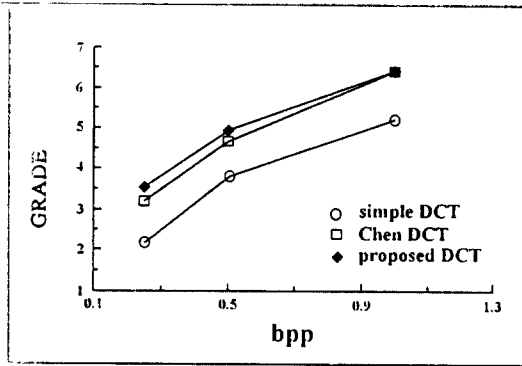
(d) PEPPER 영상 (16×16)

방식 bpp	단순 DCT	Chen	Proposed	Proposed with LPF	PSNR
0.25	1.7	2.9	3.6	3.9	등급
0.50	4.4	5.2	5.6	5.7	등급
1.00	5.1	6.3	6.4	6.4	등급

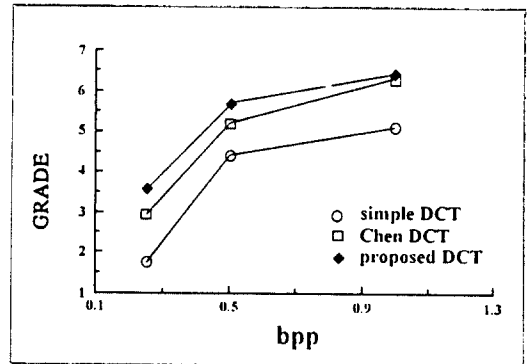
표 2. 주관적 평가 기준

Table 2 Criteria of subjective estimation

평가 내용	등급
식별되지 않음 (Not noticeable)	7
거의 식별되지 않음 (Just noticeable)	6
명확히 식별되지만 훼손정도가 가벼움 (Definitely noticeable but only slight impairment)	5
훼손됐지만 거부감은 들지 않음 (Impairment not objectionable)	4
다소 거부감이 생김 (Somewhat objectionable)	3
명확히 거부감이 생김 (Definitely objectionable)	2
지나치게 거부감이 생김 (Extremely objectionable)	1



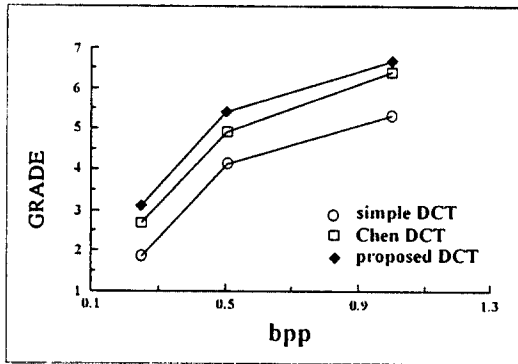
(a) LENA 영상



(b) PEPPER 영상

그림 12. 주관적 평가 결과 (16×16)

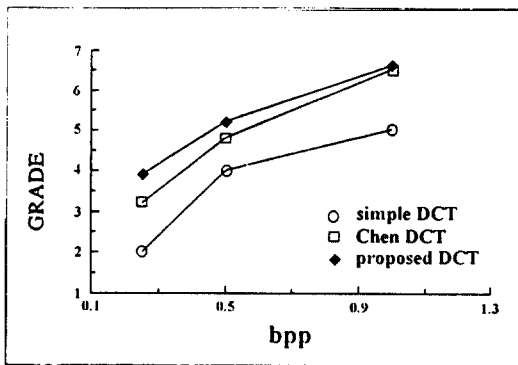
Fig. 12. Results of subjective estimation (16×16)



(b) PEPPER 영상

그림 11. 주관적 평가 결과 (8×8)

Fig. 11. Results of subjective estimation (8×8)



(a) LENA 영상

주관적 평가는 표 2에 나타난 기준으로 평가하였으며(4)(10), 10명의 대학원 학생들에게 원래의 영상과 처리된 영상을 같이 화면에 보여주어 두영상을 비교하도록 하였다. 실험에 사용한 컴퓨터는 1024×768 해상도를 가진 IBM PC 486 DX /33이다. 주관적 평가결과 표 3과 그림 11, 12에서 알 수 있듯이 고압축 영상에서 제안한 방식이 좀 더 우수한 성능을 나타냄을 확인하였다.

표 3은 모든 평가자들의 평가결과를 종합한 후 평균을 취한 결과이다.

본 논문에서 제안한 방법으로 처리된 영상을 비교 분석한 결과, 수치적 평가 방법인 PSNR은 Chen 방식보다 다소 떨어지나 이것은 인간 시각에는 덜 민감하나 수치적으로는 많은 영향을 미치는 블러의 대안 양자화가 거칠게 수행되었기 때문이다. 따라서 그 블러에서 발생하는 오차가 전체 신호대 잡음비를 낮추는 원인이 된다. 또한, 복호화 과정에서 블러간의 경계면에 LPF 처리를 하였기 때문에 시각적으로는 부드럽게 보이지만 LPF 처리된 데이터는 원래의 데이터보다 거리가 덜 멀어질 수 있다.

표 1에 나타난 바와 같이 제안한 방식은 추시적인 면에서는 Chen 방식 보다 그 성능이 떨어지나 영상 평가의 궁극적인 대상은 사람이므로 주관적인 평가치가 더 가치가 있다. 표 3에서 주관적 평가 결과, 제안한 방식이 낮은 비트율에서 더욱 우수한 성능을 나타내는데 그 이유는 제안한 방식이 잡음에 대한 시각의 민감한 정도를 이용해 압축을 하였기 때문에 잡음이 두드러지지 않는 비트율에서는 큰 효과가 나타나

지 않지만 잡음이 많이 나타나는 고압축의 비트율에서는 제안한 방식의 효과가 크게 나타남을 알 수가 있다.

V. 결 론

본 논문에서는 인간 시각체계의 특성과 영상압축에서 우수한 성능을 나타내는 변화부호화 방식인 DCT와 결합하여 주관적 화질이 우수한 부호화 방법을 연구하였다. 영상을 저장, 전송할 경우 영상 데이터의 방대한 양으로 인하여, 훼손이 따르는 영상압축이 불가피하다. 이때 영상의 손실 부호화를 수행함에 있어 인간 시각체계를 이용하면 고압축률에서 시각적으로 보다 우수한 화질을 나타내는 영상 부호화를 구현할 수가 있다.

인간의 시각은 Weber의 범칙이 설명하듯 밝기에 따라 그 차이를 감지하는 능력이 다르다. 즉, 밝은 영역에서 보다는 어두운 영역에서의 잡음이 더 눈에 띄기 쉬우며 명암 대비가 큰 곳에서 보다는 균일한 영역에서 잡음이 눈에 더 잘 띄게 된다. 이러한 시각특성을 이용하여 8×8 또는 16×16 의 DCT 변환 계수들을 새롭게 양자화하였다. 따라서 같은 비트율에서 기존의 방식과 비교했을 때 영상의 시각적 품질을 향상시킬 수 있었다.

이상과 같은 부호화 방법으로 컴퓨터 모의 실험한 결과 기존의 Chen과 Smith방식에 비하여 1.0(bpp) 이하의 고압축율에서 시각적으로 더 우수한 품질의 영상을 얻을 수 있었다.

차후 연구 과제로는 빛의 세기, 영상의 공간 세밀성, 공간 주파수, 색신호에 대한 시각의 반응 등을 정확하게 수학적으로 모델링하는 연구가 필요하며 이러한 특성들을 종합하여 영상 부호화에 최대한 적합하게 적용하는 연구가 필요하다.

참 고 문 헌

1. A. K. Jain, "Image Data Compression : A Review," *IEEE Proc.*, Vol.69, pp.349-389, 1981.
2. P. A. Wintz, "Transform Picture Coding," *IEEE Proc.*, Vol.60, pp.809-820, 1972.
3. Wen-Hsuiung Chen and C. Harrison Smith, "Adaptive Coding of Monochrome and Color Images," *IEEE Trans. Comm.*, Vol.COM-25, pp.

- 1285-1292, 1977.
4. 남승진, 최성남, 전중남, 박규태, "시각특성과 DCT를 이용한 영상부호화에 관한 연구," *한국통신학회논문지*, Vol.17, No.4, pp.323-335, April 1992.
5. K. R. Rao and P. Yip, *Discrete Cosine Transform*, Academic Press, 1990.
6. Jae S. Lim, *Two Dimensional Signal and Image Processing*, Prentice-Hall, 1990.
7. Byeong Gi Lee, "A New Algorithm to Compute the Discrete Cosine Transform," *IEEE Trans. Acoust., Speech, Signal Processing*, Vol.ASSP-32, pp.1243-1245, Dec. 1984.
8. Majid Rabbani and Paul W. Jones, *Digital Image Compression Techniques*, SPIE Optical Engineering Press, 1991.
9. Arun N. Netravali and John O. Limb, "Picture Coding : A Review," *IEEE Proc.*, Vol.68, pp.366-406, March 1980.
10. A. K. Jain, *Fundamentals of Digital Image Processing*, Prentice-Hall, 1989.
11. Murat Kunt, Michel Benard and Riccardo Leonardi, "Recent Results in High-Compression Image Coding," *IEEE Tran. Circuits and Systems*, Vol.CAS-34, pp.1306-1336 Nov. 1987.
12. N. S. Jayant and Peter Noll, *Digital Coding of Waveforms: Principles and Applications to Speech and Video*, Prentice-Hall, 1984.
13. Allen Gersho and Robert M. Gray, *Vector Quantization and Signal Compression*, KAP, 1992.
14. Arun N. Netravali and Birendra Prasada, "Adaptive Quantization of Picture Signal Using Spatial Masking," *IEEE Proc.*, Vol.65, No.4, April 1977.



李 恩 國(Eun Gug Lee) 正會員
1969年 4月 22日生
1991年 2月:弘益大學校 電子工學
科 卒業(工學士)
1993年 2月:弘益大學校 大學院 電
子工學科 卒業(工學碩
士)
1993年 1月~現在:大宇電子 映像
研究所 研究員

※研究關心分野:信號處理 및 映像壓縮



김 장 복(Chang Bock Kim)終身會員
1951年 4月 22日生
1974年 2月:延世大學校 電子工學
科 卒業(工學士)
1976年 2月:延世大學校 電子工學
科 卒業(工學碩士)
1987年 8月~1988年:獨逸 Braun-
schweig 공대 Post-
Doc.

1979年 3月~現在:弘益大學校 電子工學科 教授

※研究關心分野:映像壓縮 및 通信網