

DCT-VQ 알고리즘과 가변 블록 선택에 의한 비트율 개선

Improvement of Bit Rate by Selection of Variable Blocks and DCT-VQ Algorithm

조 태 봉*, 조 재 황*

(Tae Bong Cho*, Che Hwang Cho*)

요 약

본 논문에서는 가변블록에 대해 DCT와 VQ를 적용함으로써 복원시 영상의 품질을 원하는 수준으로 유지하면서 비트율을 감소시키는 방법을 제안한다. 각 블록에 대해 임계값에 따라 블록의 크기를 세분화 시켜 최종적인 블록에 대해 DCT와 VQ를 적용한다. 제안된 방법에서 12개의 영상에 대한 PSNR과 비트율을 비교했을 때 기존의 방법보다 비트율이 크게 감소함을 보여주고 있다.

ABSTRACT

In this paper, we propose a new method to reduce a bit rate with preserving the desired quality of the reconstructed image in applying the DCT and VQ coding based on variable blocking. The subblocking is resulted from comparison with given threshold, or the pixels of the block are replaced with the mean of the block. The final block of subblocking over the threshold value is compressed by DCT-VQ. The proposed method is lower bit rate than conventional method when it is compared with that for 12 images.

I. 서 론

영상 데이터 압축의 목적은 영상의 품질을 원하는 수준으로 유지하면서 영상을 나타내는 데이터량을 최소화하는 것이다. 압축의 근거인 영상의 중복성(redundancy)을 제거함으로써 영상의 품질을 일정수준이 되도록 유지하면서 상당한 데이터량을 감소시킬 수 있다. 공간영역에서 행하는 압축은 계산량이 적고 소프트웨어적 구현이 용이하나 큰 압축을 기대할 수 없는 반면, 주파수 영역에서의 압축은 공간영역에서 보다 계산량이 많지만 비트율(bit rate)을 크게 낮출 수 있는 장점이 있다[1].

주파수 영역에서의 대표적인 압축방법인 이산어현변환(DCT: Discrete Cosine Transform)을 적용하는 경우, 일부 몇 개의 변환 계수만이 변환 영역에서 큰 값을 가지고 대부분의 나머지 계수들은 무시할 수 있을 만큼 작은 값을 갖게 되는 에너지 집중화(energy compaction)를 이용하여 집중된 몇 개의 변환계수를 부호화 함으로써 큰 압축을 기대할 수 있다[2]-[4]. 이러한 주파수 영역 변환 부호화의 대표적인 방법으로 KLT(Karhunen-Loeve Transform), DFT(Discrete Fourier Transform), DCT, DST(Discrete Sinc Transform) 등이 있다[2]. KLT는 변환 방식에서 최적 변환이지만

각 영상마다 공분산을 구해야 하고, 다른 변환과 달리 고속 알고리즘이 없어 하드웨어나 소프트웨어적으로 구현하기가 어려운 단점이 있다. DCT는 에너지 집중현상 측면에서 KLT와 가장 유사한 성능을 가지고 있으며 고속 알고리즘이 있어 구현이 용이하여 변환을 이용한 데이터 압축방법으로 많이 이용되고 있다[5].

또한 벡터 양사화(VQ: Vector Quantization)는 데이터 원이 통계적으로 독립된 심벌로 구성되더라도 스칼라 대신 벡터로 조합된 신호를 부호화 함으로써 높은 압축률을 얻을 수 있다는 Shannon의 왜곡률 이론(Rate-Distortion)에 근거를 두고 있다[6]. VQ는 부호화와 복호화로 구성되는데, 먼저 부호화는 영상을 몇 개의 성분을 갖는 학습 벡터로 구성하고, 구성된 학습벡터를 이용하여 반복적인 근접화를 통해 코드북을 생성해주는 k-means 알고리즘 등을 이용하여 이미 정해진 개수의 코드벡터로 구성된 코드북을 생성하여 부호화기를 구성한다. 입력벡터는 부호화기의 코드북으로부터 일련의 특성벡터에 대한 가장 근사한 코드벡터를 찾아 그 코드벡터에 부여된 이진 표현 값의 색인을 전송하고, 복호화는 수신된 색인의 이진 표현 값에 해당하는 코드북의 코드벡터 값으로 치환함으로써 영상을 복원한다[7]-[9].

본 논문에서는 블록의 크기를 일정하게 고정하지 않고, 하위 블록의 완성도에 따라 블록의 크기를 가변시켜 DCT-VQ의 알고리즘을 적용한 방법을 제안한다.

* 동산대학교 전기전자공학과
접수일자: 1998년 9월 17일

II. 이론적 배경-DCT와 VQ

대역 압축의 분야에서 일반적으로 사용되고 있는 DCT는 직교변환으로 알려져 있으며, DCT에 의해 변환된 영상은 원영상을 주파수 대역별로 분해할 때 낮은 주파수 성분의 에너지는 크지만, 높은 주파수 성분의 에너지는 작아지는 특징을 이용하여 대역을 압축한다. DCT의 수행 과정은 영상을 $N_s \times N_s$ 의 블록으로 나누고, 이 블록 내의 화소에 대해 DCT의 연산을 행한다. DCT의 변환 식은 다음과 같다[5].

$$y_{kl} = \frac{4c(k)c(l)}{N_s^2} \sum_{i=0}^{N_s-1} \sum_{j=0}^{N_s-1} x_{ij} \cos\left(\frac{(2i+1)k\pi}{2N_s}\right) \times \cos\left(\frac{(2j+1)l\pi}{2N_s}\right) \quad (1)$$

$$c(k), c(l) = \begin{cases} \frac{1}{\sqrt{2}} & \text{if } k, l = 0 \\ 1 & \text{otherwise} \end{cases}$$

여기서 $k, l = 0, 1, \dots, N_s$ 이다.

수신측에서 적용될 IDCT(Inverse Discrete Cosine Transform) 식은 다음과 같다.

$$x_{ij} = \sum_{k=0}^{N_s-1} \sum_{l=0}^{N_s-1} y_{kl} \frac{4c(k)c(l)}{N_s^2} \cos\left(\frac{(2i+1)k\pi}{2N_s}\right) \times \cos\left(\frac{(2j+1)l\pi}{2N_s}\right) \quad (2)$$

여기서 $i, j = 0, 1, \dots, N_s$ 이다.

DCT변환 알고리즘의 성능을 좌우하는 것은 영상의 변환 계수중 전송하기 위한 계수의 선택과 계수를 어느 정도 정확한 값으로 전송하는가에 달려있다[5]. DCT변환 후 계수의 선택 방법에는 영역 부호화와 임계치 부호화가 있다. 영역 부호화는 변환 계수에서 정해놓은 영역에 속한 계수들을 전송하고 그 외 계수들은 전송하지 않는 방법이며 임계치 부호화는 한 블록의 변환 계수에서 그 값이 정해진 임계값보다 큰 계수들만 전송하는 방식으로 임계값을 정하기 위한 다양한 방법들이 제시되었다[10].

VQ는 크게 두 가지의 과정으로 구성된다. VQ를 구성하는 부호화기와 복호화기에는 학습 알고리즘으로 구한 동일한 코드북이 존재한다. 부호화기는 i 번째 입력벡터 $X_i = (x_{i0}, x_{i1}, \dots, x_{i(k-1)}) \in R^k$ 에 대해 코드북으로부터 거리 오차가 가장 적은 코드벡터 $Y_j = (y_{j0}, y_{j1}, \dots, y_{j(k-1)}) \in R^k$ 에 대한 색인 j 를 송신한다. 입력벡터 X_i 와 이에 해당하는 코드벡터 Y_j 사이의 거리오차 $d(X_i, Y_j)$ 는 일반적으로 유클리디안 거리(Euclidean distance)를 사용하며 다음과 같다.

$$d(X_i, Y_j) = \sum_{l=0}^{k-1} (x_{il} - y_{jl})^2 \quad (3)$$

복호화기에서는 수신된 색인 j 에 대응되는 코드벡터를 치환하여 복원한다. VQ에서 압축이 가능한 이유는 원래의 데이터 값을 전송하는 것이 아니고 일련의 조합으로 이루어진 코드벡터를 나타내는 이진 값으로 표현된 색인만을 전송하기 때문이다. VQ에 의한 압축에서 중요한 것은 압축하고자 하는 영상의 특징을 가장 적절히 표현하는 특징벡터의 결정과 코드북 생성을 위한 학습 알고리즘의 선택이다.

III. 실험 및 결과

제안된 방법을 모의 실험 하기 위해 256 그레이 레벨(grey level)을 갖는 512×512 화소의 영상 12개를 사용하며 그 영상의 종류는 표 1과 같다.

표 1. 실험 영상

Table 1. Test Images.

No.	Images
1	Airplane
2	Boat
3	Couple
4	Crowd
5	Goldhill
6	Lake
7	Lena
8	Man
9	Milkdrop
10	Peppers
11	Woman
12	Zelda

제안된 방법은 먼저 영상을 16×16 블록으로 분할하여 해당 블록에 대해 평균값을 구한 후 평균값으로 대표된 복원될 영상에 대해 PSNR(Peak Signal to Noise Ratio)를 구하여 40dB 이상이 될 때 평균값으로 대체하고, 미만일 경우에는 8×8 로 재분할 과정을 거친다. 8×8 블록에서 평균값을 구한 후 평균값으로 대체된 복원될 영상의 PSNR이 35dB 이상일 경우 평균값으로 대체하고, 그렇지 않을 경우 4×4 로 재분할하여 평균값으로 대체된 결과의 PSNR이 35dB 이상일 경우 평균값으로 대체하고, 35dB 미만일 경우 DCT-VQ를 행하여 시각적 잡음을 유발하지 않고, 비트율을 감소시킨다. 제안된 방법에서 해당 블록의 임계값은 실험을 통하여 구해진 결과로써 16×16 에 대해 블록의 임계값을 40dB와 35dB를 다르게 취할 경우 복원된 영상의 PSNR의 값은 거의 동일하게 나오지만 화소간의 상관성이 적은 영상에 대해서는 블록킹 효과가 두드러지게 나타남으로써 최상위 블록에 대해서는 40dB를 기준으로 하

였다. 일반적으로 40dB이상의 블록은 구성화소의 변화가 거의 없으므로 블록내의 화소 평균값으로 대체해도 복원 화질에 큰 영향을 미치지 않는 것으로 확인되었다. 이러한 가변 블록은 영상에서 활성도의 분포에 따라 세밀한 부분은 변환 블록의 크기를 작게 하여 블록의 크기물 가변 함으로써 원하는 영상 품질을 유지하면서 비트율을 감소시킬 수 있다. 제안된 방법의 순서도는 그림 1에 주어진다.

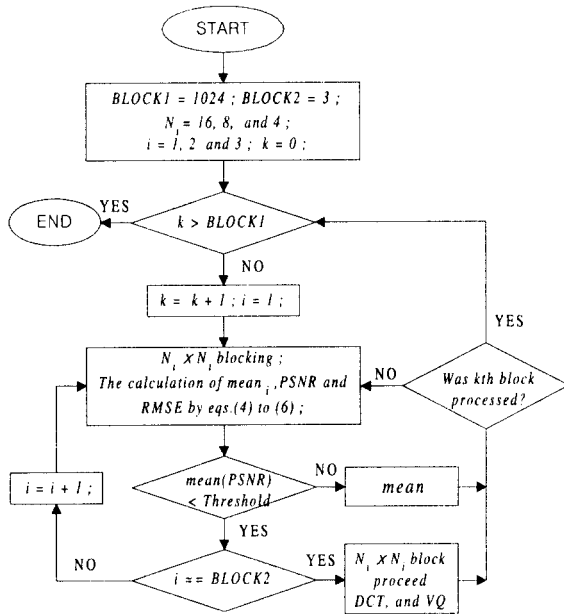


그림 1. 제안된 알고리즘 순서도
Fig. 1. Flowchart of the proposed algorithm.

그림 1에서 mean, PSNR, RMSE(Root Mean Square Error)에 대한 수식은 다음과 같다.

$$mean_i = \frac{1}{N_i \times N_i} \sum_{n=1}^{N_i \times N_i} x_n \quad (4)$$

$$PSNR = 20 \log_{10} \left(\frac{255}{RMSE} \right) \quad (5)$$

$$RMSE = \sqrt{\frac{1}{N^2} \sum_{i=1}^N \sum_{j=1}^N (f_{ij} - g_{ij})^2} \quad (6)$$

블록내의 평균값을 구하는 식은 식 (4)로 주어지고, x_n 는 블록내의 n 번째 화소값, $N_i \times N_i$ 는 i 번째 블록의 화소의 개수이며, f_{ij} 와 g_{ij} 는 i 번째 열과 j 번째 행의 원 영상과 복원된 영상의 화소값을 각각 나타낸다. PSNR과 RMSE에 대해서는 하위블록에 적용할 경우에는 $N = 16, 8, 4$ 이며, 전체 영상에 이용될 경우 $N = 512$ 이다. 그림 2에서는 12개의 영상에 대해 DCT를 가한 후 영역 부호화의 방법

을 적용하여 $4 \times 4, 8 \times 8, 16 \times 16$ 의 각각 해당 블록에 대해 DC항과 AC항 5개의 계수를 선택했을 때 PSNR의 변화를 나타낸다.

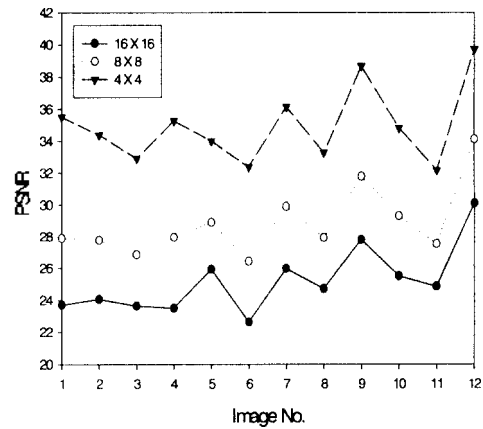


그림 2. DCT에 의한 PSNR
Fig. 2. PSNR by DCT.

그림에서 블록의 크기가 클수록 PSNR은 낮아지는 반면 비트율은 감소한다.

그림 3은 12개의 영상에 대해 제안된 가변 블록 기반 DCT-VQ를 적용한 경우의 PSNR을 나타내고 있다.

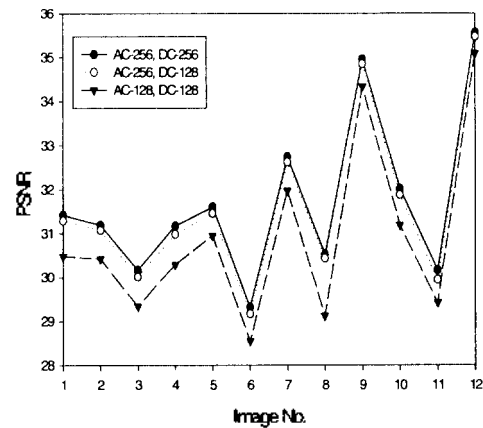


그림 3. 제안된 방법에 의한 PSNR
Fig. 3. PSNR by the proposed method.

그림 2와 그림 3으로부터 제안된 방법은 4×4 블록에 DCT를 적용할 때 보다 PSNR이 낮으나, 8×8 블록에 대해서는 높게 나타난다.

그림 3에서 나타나듯이 DC항에 대한 코드북의 크기에 따른 PSNR의 변화가 거의 없다. 따라서 DC항의 코드북 크기를 작은 값을 택함으로써 비트율을 감소할 수 있다. 그림 4는 12개의 영상에 대한 제안된 방법을 적용했을 경우 비트율의 분포를 나타낸다.

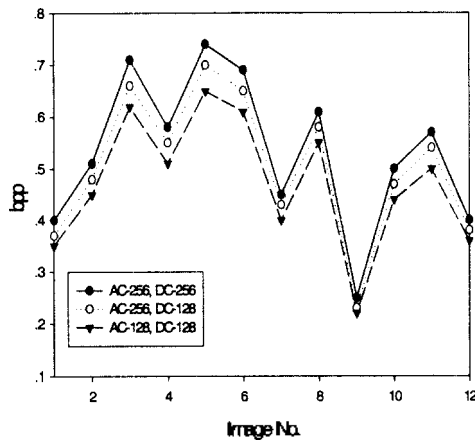


그림 4. 제안된 방법에 의한 비트율
Fig. 4. Bit Rate by the proposed method.

기존 DCT-VQ에서 블록의 크기를 4×4로 하고, DC항과 AC항의 코드북의 크기를 각각 256으로 했을 경우 비트율은 1bpp가 된다. 제안된 방법에 대한 비트율 계산식은 다음과 같다.

$$bpp = \frac{\sum_{i=1}^4 BIT_i \times num_i}{512^2} \quad (6)$$

여기에서 BIT_1 과 num_1 은 16×16블록의 평균에 대한 PSNR이 임계값보다 작아 그 블록의 평균으로 대표될 때 정보 비트와 블록의 크기를 구분할 수 있는 구분 비트에 대한 비트율(0.03390 = (8bits+2bits)/(16×16) pixels)과 그 개수를 의미한다. BIT_2 와 num_2 는 8×8블록에 대한 비트율(0.1526bpp)과 개수이고, BIT_3 와 num_3 는 4×4블록에 대한 비트율(0.625bpp)과 개수, BIT_4 는 4×4블록에 대한 DC항의 코드북을 8비트, AC항의 코드북을 8비트로 구성할 때 비트율(1.125bpp)이다. 하지만 DC항의 코드북을 7비트, AC항의 코드북이 8비트일 때 비트율(1.0625bpp)이 되고, num_4 는 블록의 개수를 의미한다. 그림 2, 3, 4를 통해 제안된 방법은 4×4 블록보다는 PSNR은 낮아지나 비트율이 크게 개선되고, 8×8 블록보다는 PSNR이 높다. 따라서 블록의 크기를 가변적으로 선택한 방법이 고정된 블록기반 변환 부호화 방법을 이용하는 경우보다 응용에 따라 다양하게 적용될 수 있다.

IV. 결 론

제안된 방법은 임계값에 의해 블록을 가변적으로 재분할하여 화소간의 변화가 적은 블록에 대해 평균값을 전송하고, 화소간의 변화가 많은 부분에 대해서는 블록의 크기를 작게하여 DCT와 VQ를 적용함으로써 시각적으로 잡음을 유발하지 않으면서 비트율을 감소할 수 있다. 본 논

문에서 알 수 있듯이 화소간의 변화가 많은 영상에 대해서는 비트율이 크게 줄어들지는 않지만 화소간의 변화가 적은 영상에 대해서는 비트율이 크게 감소됨을 알 수 있다. 따라서 비트율을 감소하기 위해 블록에 대해 최적의 임계값을 찾는 방법과, DC항과 AC항의 각 계수의 코드북에 대해 최적의 크기를 찾고, 블록을 재분할하는 과정을 세분화하여 비트율을 감소시킬 수 있는 방법이 앞으로의 연구 과제이다.

참 고 문 헌

1. K. R. Rao and J. J. Hwang, *Techniques & Standards For Image Video & Audio Coding*, Prentice Hall, pp. 43-83, 1996.
2. V. Bhaskaran and K. Konstantinides, *Image and Video Compression Standards*, KAP, pp. 53-85, 1995.
3. R. C. Reinninger and J. D. Gibson, "Distributions of Two-Dimensional DCT Coefficient for Images," *IEEE, Trans. Comm.*, Vol. COM-31, no. 6, pp. 835-839, Jun. 1983.
4. W. H. Chen and C. H. Smith, "Adaptive Coding of Monochrome and Color Images," *IEEE, Trans. Comm.*, Vol. COM-25, pp. 1285-1292, Nov. 1977.
5. R. J. Clarke, *Digital Compression of Still Images and Video*, Academic Press, pp. 53-86, 1995.
6. T. Murakami, K. Asai and E. Yamazaki, "Vector Quantiser of Video Signals," *Electronics Letter*, Vol. 7, pp. 1005-1006, 1982.
7. Y. Linde, A. Buzo and R. M. Gray, "An algorithm for Vector Quantizer design," *IEEE Trans. Comm.*, Vol. COM-28, pp. 84-95, Jan. 1980.
8. R. M. Gray, "Vector Quantization," *IEEE ASSP Mag.*, pp. 4-29, Apr. 1984.
9. A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*, KAP, pp. 345-400, 1992.
10. M. Rabbani and P. W. Jones, *Digital Image Compression Techniques*, SPIE Press, pp. 102-128, 1991.

▲ 조 태 봉 (Tae-Bong Cho)



1996년 2월: 동신대학교 전자공학과 졸업(공학사)
1996년 8월 ~ 현재: 동신대학교 전기 전자공학과(석사과정)
*주관심분야: 동영상압축, 영상처리, 패턴인식

▲ 조 제 황 (Che-Hwang Cho)

한국음향학회지 제16권 1호 참조