

동영상 부호화기에서의 적응적 양자화 매트릭스 생성

(An Adaptive Quantization Matrix Generation in Moving Image Encoder)

金正祐*, 李根泳*

(Jeongwoo Kim and Keunyoung Lee)

요 약

동영상 부호화에서 양자화는 부호 발생량 제어에 의해 손실을 발생시킨다. 부호화기에서 발생하는 대부분의 에러는 이 단계에서 발생하므로 잘 설계된 양자화기는 같은 비트율에서 주관적인 평가를 개선시킬 수 있다. 표준안에서 사용되는 양자화 매트릭스는 인간의 시각특성을 반영한 것이긴 하지만, 영상의 종류에 상관없이 일괄적으로 적용되므로 그 효율성은 저하된다. 그러므로 본 논문에서는 영상의 각 블록에 대하여 복잡도를 구하고, 이를 이용하여 개별적인 양자화 매트릭스를 작성하여 사용하였다. 이 방법을 정지영상과 동영상 표준안에 각각 적용하였으며, 그 결과 영상의 종류에 따라 약 0.05~0.82dB 정도의 PSNR 개선과 에지영역에서의 주관적인 평가의 개선을 얻을 수 있었다.

Abstract

Quantizer with bit rate control performs lossy compression in moving image coding. Efficient quantizer can improve subjective quality of reconstructed image at the same bit rate because most of errors are occurred at this stage. Although quantization matrix of current compression standards reflects human visual system, the performance of their quantization matrixes is not always optimal because they are not considering each image characteristics. We proposed a new adaptive optimal quantization matrix obtained by considering complexity of each image subblocks. It resulted in improvements of image quality about 0.05 ~ 0.82 dB and good subject quality at edge regions when we simulated the quantization method in various coding standards.

I. 서 론

우리는 매일 수많은 정보들이 쏟아져 나오는 정보의 홍수 속에서 살고 있다. 이들 정보는 문자, 음향, 영상, 일반 데이터등 여러 가지 형태를 하고 있다. 근래에는 이러한 다양하고 복합적인 정보들이 동일한 전달 미디어를 통하여 종합적으로 저장 및 전송되어지며, 이를 멀티미디어라 정의한다.^[1] 멀티미디어의 대부분을 차지하는 것은 영상과 음향이며, 이들 데이터의 양은 단

순한 방법에 의해 저장 및 전송되기에는 전달 미디어에 너무 큰 부담이 된다. 그러므로 멀티미디어가 얼마나 효율적이고 실현 가능한지는 데이터를 얼마만큼 압축할 수 있는가에 달려 있다고 해도 과언은 아니다.

고전적인 데이터 압축 방법은 크게 무손실 압축과 손실 압축으로 나누어진다. 무손실 압축은 데이터의 완전 복원이 가능한 반면 2:1 ~ 10:1 정도의 낮은 압축율을 가지며, 엔트로피 부호화, run-length 부호화, 예측 부호화 등이 여기에 속한다. 손실 압축은 압축된 데이터의 완전 복원은 불가능하지만 손실 정도에 따라서는 30:1 이상의 높은 압축율을 얻을 수 있다.^[2] 영상에 있어서 어느 정도의 손실은 사람의 시각특성상 잘 인식할 수 없다는 것을 이용한 압축방법이며, 대부

* 正會員, 成均館大學校 電子工學科

(Dept. of Electronics, Sungkyunkwan Univ.)

接受日字: 1997年10月20日, 수정완료일: 1998年2月3日

분의 손실은 양자화 과정에서 발생한다. 높은 압축율과 좋은 화질을 얻기 위해서는 이들 두 가지 압축방법을 적절히 결합하여 사용한다. 현재 세계 각국은 압축 방법에 관한 국제 표준화를 위해 노력하고 있다. 그 결과로서 정지영상에 대한 표준안인 JPEG^[3], 동영상에 관해서는 MPEG^[4,5]이 표준안으로 확정되었다. 이외에도 전화선을 이용한 데이터 전송을 위해 낮은 비트율을 목표로 하는 H.261^[6], H.263^[7] 등이 권고안으로 채택되었다.

MPEG과 같은 동영상 부호화의 표준안 내용을 크게 분류하면 프레임내 부호화 방법과 프레임간 부호화, 부호발생량 제어로 나눌 수 있다. 프레임내 부호화 방법은 DCT(Discrete Cosine Transform)를 이용한 변환 부호화 후 양자화기를 거치며, 이를 엔트로피 부호화를 이용하여 비트열이 생성된다. 이 과정에서 양자화기는 미리 약속된 양자화 매트릭스와 부호발생량 제어 방법에 의해서 생성된 계수값에 의해 양자화 과정을 수행한다. 부호화기에서 생성되는 대부분의 에러는 이 과정에서 발생한다. 그러므로, 양자화기가 영상의 특성을 얼마나 잘 반영하여 에러를 생성하는 지에 따라 복원되는 영상의 화질에 심각한 영향을 미친다. 양자화기에서 사용되는 매트릭스의 값은 충분한 데이터를 이용하여 실험을 거친 후 작성되었으며^[3], 비록 표준안에 포함되긴 하였으나 모든 영상에 적합하게 적용될 수는 없으므로 그 값에 대해서는 자유도를 인정하고 있다. 본 논문에서는 부호화 과정에서 입력된 영상의 특성을 분석하고, 인간의 시각 특성이 반영된 적응적인 양자화 매트릭스를 블록별로 능동적으로 생성하여 양자화 하는 방법을 제안하였다. 이를 통하여 주관적, 객관적 평가의 개선을 얻을 수 있었다.

2장에서는 기존 표준안에서 사용된 양자화 매트릭스를 분석하고, 3장에서는 이들 방법을 개선하기 위해 제안된 양자화 매트릭스 생성 과정에 관하여 설명하였다. 4장에서는 이 알고리즘을 이용하여 JPEG, moving-JPEG, MPEG과 같은 정지영상, 동영상 표준안에 적용한 후 결과를 비교한다. 마지막 5장에서는 본 논문에서 사용된 방법으로 얻은 결과에 관한 분석과 문제점 및 추후 연구방향에 관하여 서술하였다.

II. 표준안에서의 양자화 매트릭스

표준안에서의 변환 부호화 방법으로는 DCT가 사용된다. DCT는 최적 변환인 KLT에 가장 근접한 특성을 가지며 에너지 집중효과가 뛰어난 변환부호화 방법이다. 이로 인하여 주피수 영역으로 변환하고 난 후의

블록내 계수값들은 일정한 분포를 가진다. 계수값에 대하여 지그재그 스캐닝을 거쳐서 줄길이 부호화 등의 엔트로피 부호화를 사용하는 것 또한 이런 DCT의 특성을 잘 이용한 부호화 방법이다. 그림 1은 DCT를 기반으로 하는 일반적인 부호화기의 블록도이다.

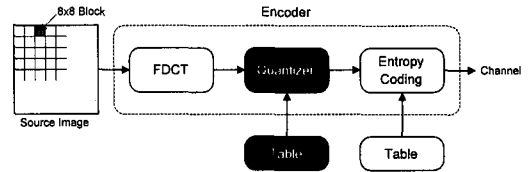


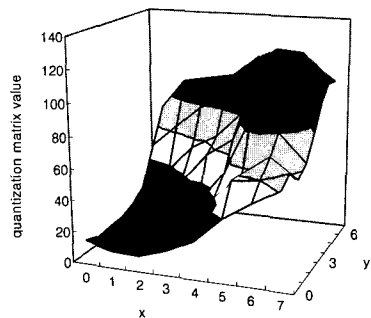
그림 1. DCT 기반 부호화기
Fig. 1. DCT-Based Encoder.

FDCT(Forward DCT)를 이용하여 변환된 블록내 (0,0) 좌표의 계수는 DC성분을 나타내며, 이 곳에서 멀어질수록 고주파 성분은 커진다. 사람의 시각은 저주파 성분의 변화에 민감하며, 반대로 고주파 성분의 변화에 둔감한 특성을 가진다. 그러므로 인간의 시각 특성을 충분히 이용한 양자화기의 경우, 변환계수 중 저주파 성분은 최대한 원래의 값을 유지하도록 하며 고주파성분 쪽으로 갈수록 계수값을 줄일 수 있도록 만들어져야 한다. 이렇게 구성된 양자화 매트릭스는 식 (1)과 같이 원래의 계수값에 적용된다.

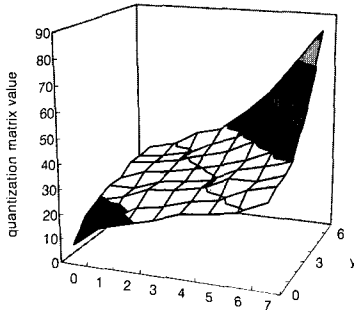
$$\hat{X}(x, y) = \text{round} \left(\frac{X(x, y)}{Q(x, y)} \right) \quad (1)$$

$$x, y = 0, \dots, 7$$

여기에서 $X(x, y)$ 는 변환부호화 후의 계수값, $Q(x, y)$ 는 양자화 매트릭스, $\hat{X}(x, y)$ 는 양자화 후의 계수값을 나타낸다. 그림 2에서는 이러한 특성을 이용한 표준안에서의 양자화 매트릭스 분포를 보여주고 있다.



(a) JPEG Luminance



(b) MPEG Intra

그림 2. 표준안 양자화 매트릭스 분포
Fig. 2. Quantization matrices for standards.

III. 제안한 양자화 매트릭스 생성 방법

그림 2에서의 양자화 매트릭스들은 충분한 양의 데이터에 의하여 일반화된 것이긴 하지만 영상의 종류와 프레임간, 혹은 프레임내의 관계에 상관없이 일률적으로 적용됨으로써 그 효율성이 저하된다. 그러므로 영상의 복잡도에 관한 평가척도를 이용하여 이에 적응적으로 대처할 수 있는 양자화 매트릭스는 화질에 있어서 주관적 평가의 개선을 가져올 수 있다. 이를 위하여 본 논문에서는 각 블록에 대한 원 영상의 표준편차를 구하고 이 값을 이용하여 그 블록에 적합한 양자화 테이블을 생성하여 적용하는 알고리즘을 제안한다. 양자화 매트릭스의 생성을 위하여 본 논문에서는 다음과 같은 수식 (2)를 제안하여 사용하였다.

$$q(x, y) = \frac{1}{1 + \sigma' e^{-\gamma(\sqrt{x^2+y^2} - center)}} \quad (2)$$

$$\sigma' = \frac{1}{10} \sqrt{\frac{n \sum x^2 - (\sum x)^2}{n(n-1)}} \quad (3)$$

단, center는 블록의 중심을 나타내며, γ 는 저주파와 고주파 경계에서의 기울기이며 실험적으로 0.5 ~ 1.2 정도의 값이 적당하다. σ' 은 표준편차를 1/10로 scale down한 값으로 블록 내 화소 값(x)과 화소 수(n)에 의해 계산되어진다. 이 값은 식 (2)에 대입되어 양자화 매트릭스를 구성하며, 이는 식 (1)에 의하여 영상에 적용된다. 사용된 σ' 값은 복호화기에서의 역 양자화를 위하여 부호화 후 전송된다.

1. σ' 에 의한 영향

그림 3은 σ' 에 의해 위의 식 (2)의 그래프가 어떻게 영향을 받는지 보여준다. 그래프에서의 변화를 확연하게 나타내기 위하여 1차원 식인 (2)'을 이용하였다.

$$q'(x) = \frac{1}{1 + \sigma' e^{-\gamma(x - center)}} \quad (2)'$$

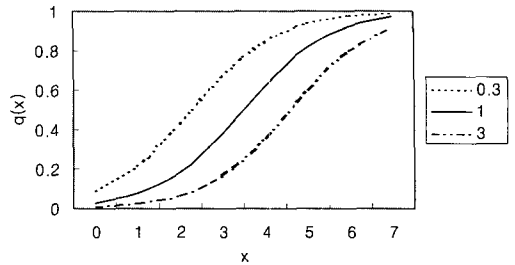


그림 3. σ' 의 변화에 따른 그래프
Fig. 3. Graph according to change of σ' .

복잡한 영상은 고주파 성분을 많이 포함하고 있으며, 이에 따라 계산된 σ' 의 값은 커진다. 이 값은 그림 3에서 보는 바와 같이 그래프를 우측으로 옮기는 역할을 하여 전체적으로 블록내의 DC값을 기준으로 넓은 영역에 걸쳐서 낮은 양자화값을 할당한다. 이로 인해 많은 고주파 성분이 살게되어 주관적인 평가의 개선을 가져오지만 부호화에 많은 비트를 할당하여 비트율은 증가한다. 반대로 낮은 σ' 은 그래프를 좌측으로 옮기는 역할을 하여 DC를 기준으로 매우 좁은 영역에 대해서만 부호화 하게 된다. 이 결과 고주파 성분의 대부분 값이 0이 되어 부호화 효율을 높이며 비트율을 감소시킬 수 있다.

2. γ 에 의한 영향

σ' 은 저주파와 고주파의 경계를 결정지으며, γ 는 이 경계에서의 기울기를 나타낸다. γ 의 값이 작을수록 그래프는 완만한 기울기를 가지며, 적당히 작은 값은 양자화 매트릭스 값의 경계 부분에서 발생할 수 있는 오차를 줄여준다. 그림 2에서 보는 바와 같이 JPEG 휘도 성분(Luminance)을 위한 양자화 매트릭스의 경우는 전체적으로 급격한 기울기와 중간 정도의 분산을 가지는 일반적인 형태로 고려된다. 이에 반해 MPEG Intra를 위한 양자화 매트릭스는 완만한 기울기와 높은 고주파 성분까지 포함하고 있으므로 인간의 시각 특성을 충분히 반영한 것으로 보기는 힘들다. 하지만, H.261에서 사용되었던 양자화 변수(mquant)를 이용한 선형 양자화기와 병행하고 있으므로 부호발생량 제어에 더 큰 비중을 두고 있음을 알 수 있다. γ 의 값이 0인 경우 식 (2)는 $1/(1+\sigma')$ 이 되며 이는 mquant를 이용한 선형양자화기와 유사하다. 그래프가 완만할수록 고주파 성분이 많이 포함되므로 PSNR은 증가할 수 있으나 그림 5에서 보는바와 같이 0.5 미만의 값에 대해서는 영상의 특성에 상관없는 PSNR 분

포를 가진다. 이는 인간의 시각특성을 가지는 양자화 매트릭스가 전혀 적용되지 않은 결과이다. 또한, 주관적인 평가에 있어서도 심한 블록화 현상이 발생하므로 본 논문에서는 실험적인 결과에 따라 0.5의 값으로 고정하였다.

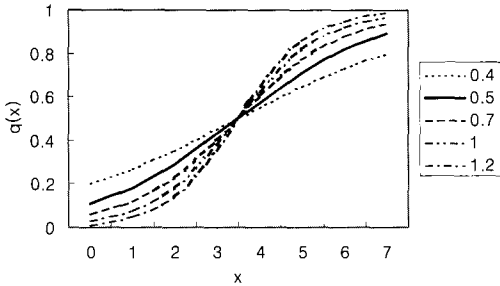


그림 4. γ 의 변화에 따른 그래프
Fig. 4. Graph according to change of γ .

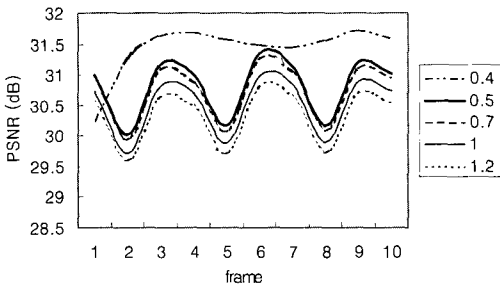


그림 5. γ 의 변화에 따른 PSNR (News 영상)
Fig. 5. PSNR distributions according to γ (News image).

3. 양자화 매트릭스의 스케일링

식 (4)와 그림 6은 양자화 매트릭스의 값을 결정하기 위하여 식 (2)에서 계산된 범위에 관한 스케일링 방법이다.

$$Q(x, y) = \frac{(upper - lower)}{(t_2 - t_1)} \times (q(x, y) - t_1) + lower \quad (4)$$

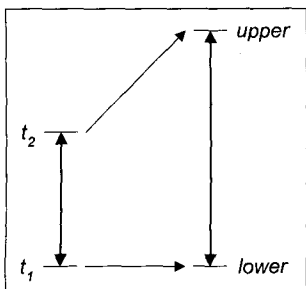
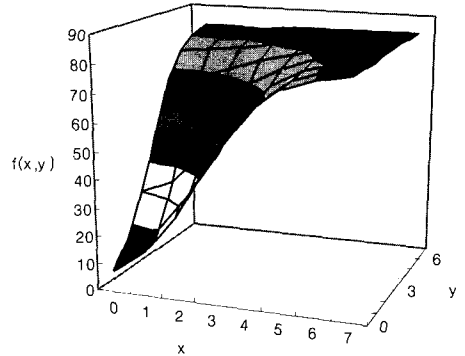
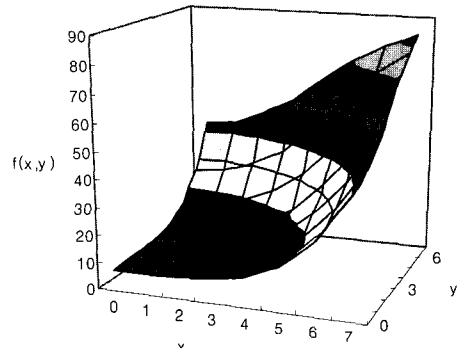


그림 6. 스케일링
Fig. 6. Scaling.

t_1 은 계산된 σ' 에 의한 최소값, t_2 는 최대값이며, upper와 lower는 각각 스케일링하고자 하는 최대, 최소값이다. JPEG Luminance의 경우 upper는 99, lower는 16이며 MPEG Intra의 upper는 83, lower는 8의 값을 가진다. 그림 7은 위 식에 의해 실제 계산된 예를 보여준다. 여기서 스케일링은 MPEG의 값을 이용하였으며, γ 는 0.5를 사용하였다.



(a) $\sigma' = 0.1$



(b) $\sigma' = 10.0$

그림 7. 생성된 양자화 매트릭스
Fig. 7. Generated quantization matrix.

IV. 실험결과

본 논문에서는 블록별 양자화 매트릭스를 생성하기 위하여 스케일링된 표준편차를 이용하므로 역 양자화를 위하여 이 값이 비트열에 포함되어야 한다. 이 값은 $0.00 < \sigma' < 10.00$ 의 범위로 제한하며 정수 변환 후 10비트로 부호화할 수 있다. 이로 인하여 각 영상에 대하여 약 0.16bpp 정도의 추가 부담이 생기며, 이를 프레임내에서 DPCM을 취한 후 엔트로피 부호화 하더라도 약 0.1bpp 정도의 부담은 피할 수 없다.

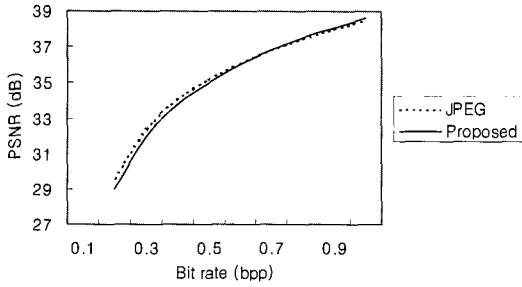


그림 8. Lenna 영상의 PSNR 분포
Fig. 8. PSNR distribution for Lenna image.

그림 8은 512×512 크기를 가지는 Lenna 영상에 관한 JPEG과의 비교 그래프이다. 그래프에서 보는 것처럼 0.8bpp이하에서는 0.1bpp 정도의 오버헤드로 인하여 JPEG보다 결과가 뒤떨어짐을 알 수 있다. 프레임내에서의 무손실 압축에는 한계가 있으므로 본 논문에서는 정지영상보다는 동영상에서의 프레임간 DPCM 과 허프만 부호화를 이용하여 오버헤드를 평균 0.02 ~ 0.04bpp 정도로 낮추었다. 이 때 사용된 허프만 부호 테이블은 suzie영상과 trevor영상(QCIF) 각 150 프레임씩, 총 300 프레임을 이용하여 구성하였다.

그림 9는 news영상(QCIF) 50 프레임, 그림 10은 akiyo영상(QCIF) 50 프레임에 대한 moving-JPEG 과의 비교 그래프이다.

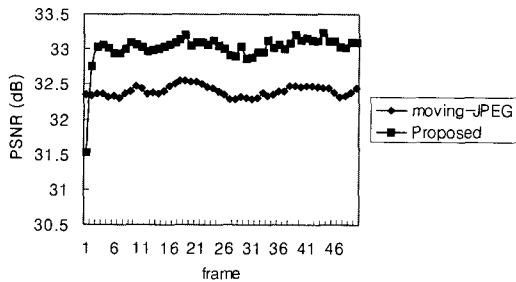


그림 9. News 영상에 대한 PSNR (0.9bpp)
Fig. 9. PSNR distribution for News image (0.9bpp).

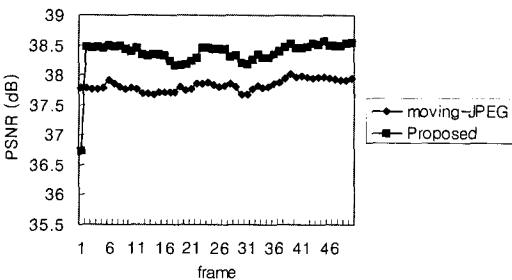


그림 10. Akiyo 영상에 대한 PSNR (0.9bpp)
Fig. 10. PSNR distribution for Akiyo image (0.9bpp).

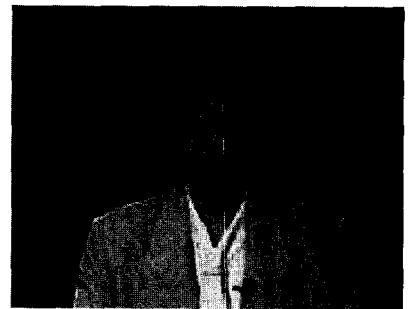
두 가지 그래프에서 첫 번째 프레임이 JPEG에 비해서 PSNR이 상당히 뒤떨어짐은 프레임간 DPCM을 위한 초기값과의 차이로 인한 결과로 생각된다. 그래프에서는 두 가지 영상에 관하여 전체적으로 평균 0.6dB 정도의 개선이 있었음을 알 수 있다.



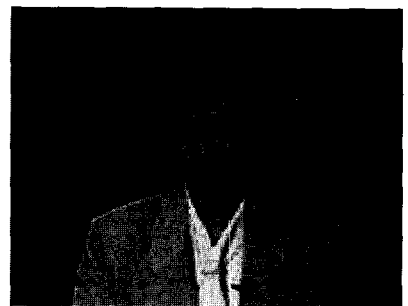
(a) news : moving-JPEG



(b) news : 제안된 방법



(c) akiyo : moving-JPEG



(d) akiyo : 제안된 방법

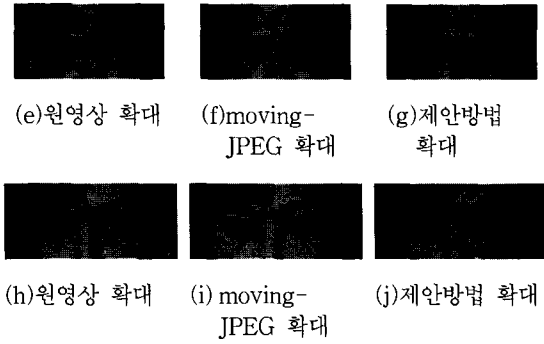


그림 11. 결과영상 (10번째 프레임)
Fig. 11. Result images (10th frame).

그림 11에서는 news 영상과 akiyo 영상의 10번째 프레임에 관한 결과 영상을 JPEG과 비교한 것이다. 제안된 방법이 고주파 성분이 많은 영역(에지 영역)에서 moving-JPEG에 비하여 주관적으로 훨씬 나은 결과임을 알 수 있다. 보다 효율적인 비교를 위하여 부분확대 하였다.

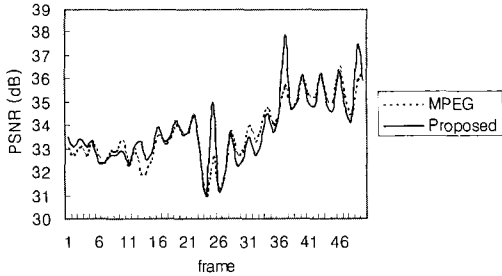


그림 12. Table tennis 영상에 대한 PSNR (1.5Mbps)
Fig. 12. PSNR distribution for Table tennis image (1.5Mbps).

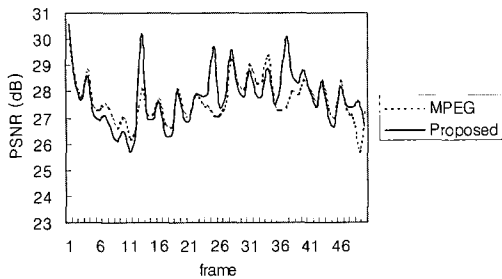


그림 13. Flower garden 영상에 대한 PSNR (1.5Mbps)
Fig. 13. PSNR distribution for Flower garden image (1.5Mbps).

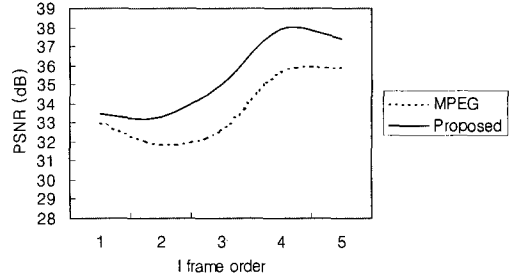


그림 14. Table tennis 영상의 I 프레임 PSNR (1.5Mbps)
Fig. 14. PSNR distribution for I frames in Table tennis image (1.5Mbps).

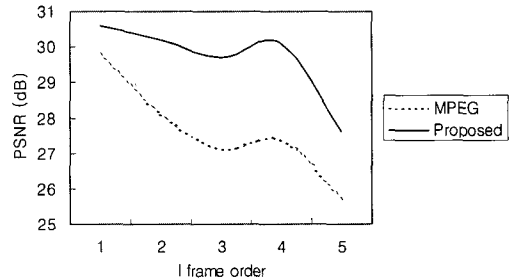
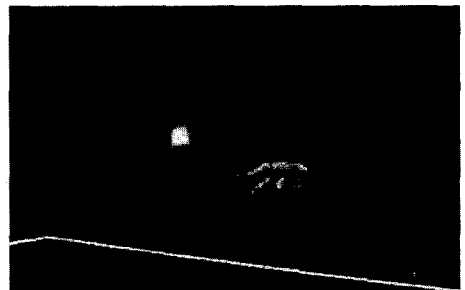
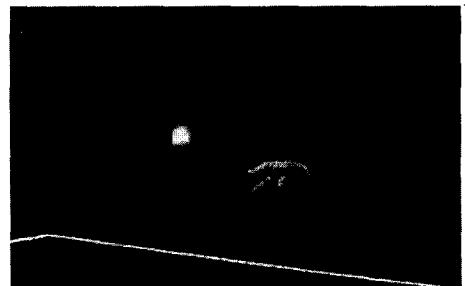


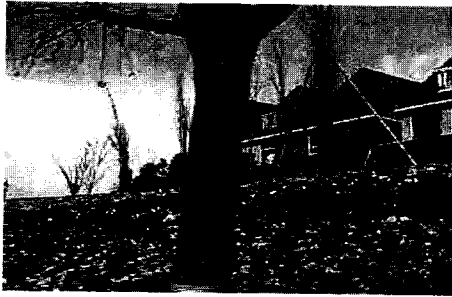
그림 15. Flower garden 영상의 I 프레임 PSNR (1.5Mbps)
Fig. 15. PSNR distribution for I frames in Flower garden image (1.5Mbps).



(a) Table tennis - MPEG



(b) Table tennis - 제안된 방법



(c) Flower garden - MPEG



(d) Flower garden - 제안된 방법

그림 16. 결과영상 (2번째 I 프레임)

Fig. 16. Result images (2nd I frame).

그림 12, 13은 Table tennis영상(SIF)과 Flower garden(SIF)영상에 대한 MPEG과의 PSNR 비교이다. $N = 12$, $M = 3$ 이며, 비트율은 1.5Mbps이다. moving-JPEG과는 달리 부호발생량 제어에 의하여 첫 번째 프레임에 대한 PSNR의 보상이 있었음을 알 수 있다. 전체적으로는 양자화기가 Intra 부호화에 속하므로 I 프레임에 대하여 우수함을 알 수 있으며, 이를 그림 14, 15에 나타내었다. P와 B프레임의 경우에는 양자화기의 추가 비트에 의한 부담이 I 프레임에서 얻은 PSNR을 소모하고 있는 것으로 생각된다. 표 1과 2는 실험에 사용된 영상들의 평균 PSNR을 나타내고 있다. moving-JPEG의 경우 약 0.5~0.6dB, MPEG의 경우에는 비트율과 영상에 따라 약 0.05~0.82dB 정도가 개선되었다.

표 1. 실험 영상의 평균 PSNR (moving-JPEG)

Table 1. Average PSNRs of test images (moving JPEG).

Method Sequences	moving-JPEG	제안된 방법
News	32.40 dB	33.01 dB
Akiyo	37.83 dB	38.36 dB

표 2. 실험 영상의 평균 PSNR (MPEG)
Table 2. Average PSNRs of test images (MPEG).

Method Sequences		MPEG	제안된 방법
Table tennis	512Kbps	29.64 dB	29.94 dB
	1Mbps	32.12 dB	32.27 dB
	1.5Mbps	33.74 dB	33.81 dB
Flower garden	512Kbps	23.32 dB	24.14 dB
	1Mbps	25.73 dB	25.78 dB
	1.5Mbps	27.64 dB	27.76 dB

V. 결론

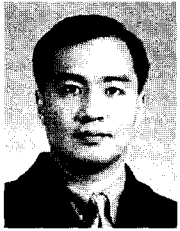
현재 사용중인 대부분의 동영상 압축 알고리즘은 DCT를 이용한 변화부호화 후 양자화를 거친다. 이때 사용되는 양자화 테이블은 비록 많은 실험을 거친 후 만들어진 것이긴 하지만 다양한 영상에 적응적으로 대처할 수 없다. 표준안에서는 단지 이에 관한 자유도만 인정하고 있으므로 본 논문에서는 영상의 복잡도에 따라 양자화 테이블을 능동적으로 바꿀 수 있는 구체적인 알고리즘을 제안하였다. 이를 정지영상에 적용한 결과 추가된 비트의 영향으로 낮은 비트율에서는 표준안보다 약간 뒤지는 면이 있으나 많이 사용되는 높은 비트율에서는 좋은 성능을 얻을 수 있었다. 또한, 오버헤드를 프레임간의 DPCM을 이용하여 더욱 줄일 수 있었던 동영상에서는 실험을 통하여 표1, 2에 보여지는 바와 같이 객관적인 평가척도(PSNR)의 개선을 입증할 수 있었으며, 주관적인 평가에 있어서도 고주파 부분에서의 화질 개선을 얻을 수 있었다. 영상의 복잡도 계산에는 부호발생량 제어시에 사용되는 값을 그대로 사용할 수 있으나 테이블 생성시에는 추가 계산량이 필요하므로 이에 관한 개선이 필요하다.

참고 문헌

- [1] 멀티미디어 통신연구회, *최신 MPEG 教科書* 아스키-출판국, 1994.
- [2] Rafael C. Gonzalez, Richard E. Woods, *Digital Image Processing: Addison-Wesley Publishing co.*, 1993.
- [3] W. B. Pennebaker and J. L. Mitchell, *JPEG: Still Image Compression Standard*. New York, NY: Van Nostrand Reinhold, 1993.
- [4] MPEG-1 ISO/IEC 11172, "Information

- Technology Coding of Moving Picture and Associated Audio for Digital Storage Media at up to about 1.5M bits/s," Aug 1993.
- [5] ISO/IEC/JTC1/SC2/WG11, MPEG 92/229, MPEG-2 Video, March 1992.
- [6] ITU-T Recommendation H.261, "Video codec for audiovisual services at $p \times 64$ kbits/s," Aug, 1990.
- [7] Draft ITU-T Recommendation H.263, "Video coding for low bit rate communication," May, 1996.
- [8] R. de Queriroz, C. K. Choi, Y. Huh, and K. R. Rao, "Wavelet Transforms in a JPEG-Like Image Coder," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 7, No. 2, pp. 419-424, April 1997.
- [9] Xudong Song, Timo Viero, Yrjö Neuvo, "Interframe DPCM with Robust Median-Based Predictors for Transmission of Image Sequences Over Noisy Channels," *IEEE Trans. Image Processing*, Vol. 5, No. 1, pp. 16-25, Jan. 1996.

 저 자 소 개



金 正 祐(正會員)

1967년 2월 12일생. 1991년 2월
성균관대학교 공과대학 전자공학과
(공학사). 1993년 2월 성균관대학
교 대학원 전자공학과(공학석사).
1996년 2월 성균관대학교 대학원
전자공학과(박사수료). 1996년 3월

~ 현재 성균관대학교 부설 과학기술연구소 연구전담
요원 근무. 주관심분야는 영상통신, 신호처리, 디지털
시스템설계

李 根 泳(正會員) 第 35卷 S編 第 1號 參照