

## 가변 임계값을 갖는 JPEG 영상의 양자화 방법

윤 석 현\*

### JPEG Image Quantization Method with Variable Thresholds

Seok-Hyun Yoon\*

#### 요 약

본 논문에서는 JPEG 영상의 오차(PSNR)의 개선을 위해 반올림 양자화와 내림 양자화의 장점을 조합하여 가변 임계값을 갖는 양자화 방법(VTQM: Variable Threshold Quantization Method)을 제안하고, 실험에 의해 VTQM의 비트율 대비 PSNR을 표준 JPEG 방법의 경우와 비교하였다. VTQM에서는 양자화 방법에 따라 영상의 압축률에 영향을 미치는 영역에서는 압축한 결과값의 길이가 줄어들 수 있도록 임계값을 조정하여 양자화하고, 그 외의 경우에는 원시 영상의 값을 유지하는 방향(반올림)으로 양자화 한다. 그 결과 최대한 화질을 유지하면서 비트율을 감소시킬 수 있다. 몇 개의 샘플 영상에 대한 실험 결과, VTQM은 기존의 반올림과 내림 양자화 방법보다 같은 비트율에서 PSNR이 각각 0.2 ~ 0.3dB, 0.4 ~ 0.5dB 향상되는 것을 보여주었다.

#### Abstract

In this paper, VTQM(Variable Threshold Quantization Method) is proposed, which enhances PSNR by composing the advantages of round-off and truncation quantization methods. From vast experiments, PSNR's of JPEG images with the method are compared with that of standard JPEG methods. At the boundary which affects the compression ratio, VTQM quantizes with a new threshold to reduce the length of compressed data; otherwise, VTQM employs the round-off quantization method to preserve the original image information. As a result, VTQM reduces bit rate while preserving the image quality. The experimental results show that VTQM enhances PSNR by 0.2 ~ 0.3dB, 0.4 ~ 0.5dB with the same bit rate than those of the traditional truncation and round-off quantization methods, respectively.

---

\* 청강문화산업대학 컴퓨터소프트웨어과 교수

## I. 서론

영상 데이터를 압축하는 목적은 전달하고자 하는 영상 정보를 손상시키지 않는 범위 내에서 가능한 한 적은 데이터로 영상을 표현하는데 있다. JPEG은 정지영상 압축의 표준이다(1~7). JPEG 표준은 압축방법에 따라 크게 DCT(Discrete Cosine Transform)를 이용한 손실 압축 방법과 DPCM(Differential Pulse Code Modulation)을 이용한 비손실 압축으로 크게 나뉘어지며 JPEG의 모드는 연속 DCT 모드(Sequential DCT-based mode), 점진 DCT 모드(Progressive DCT-based mode), 연속 비손실 모드(Sequential lossless mode)와 계층 모드(Hierarchical mode) 등으로 나뉘어진다. JPEG 베이스라인은 DCT를 근간으로 하는 JPEG 압축 방법에 사용되어야 되는 공통의 특성을 의미한다.

본 논문에서는 JPEG 베이스라인 과정 중에서 양자화 단계에 대하여 가변 임계 양자화 방법(VTQM: Variable Threshold Quantization Method)을 제안한다. 기존의 양자화 방법으로는 내림 양자화와 반올림 양자화가 대표적이며 표준안에서는 반올림 양자화를 사용할 것을 권고하고 있다. 같은 양자화 테이블을 사용할 경우 내림 양자화는 반올림 양자화에 비하여 압축률이 증가(비트율이 감소)하지만 상대적으로 화질이 나빠지는 단점이 있다. 본 논문에서 제안하는 VTQM은 반올림 양자화와 내림 양자화의 장점을 조합하여, 양자화로 인해 압축률에 영향을 미치는 경우와 기타의 경우로 나누어 각각 다른 임계값을 이용하여 원영상의 정보를 유지하면서 압축률을 높여준다.

기존의 양자화 방법에 관한 연구들은 주로 영상에 최적인 양자화 테이블을 구하는 연구(8, 9)등이 있으며 (8)의 경우 영상에 따라 같은 비트율(bpp: bit per pixel)에서 PSNR이 0.8 ~ 1.7dB정도 개선되는 것으로 발표되고 있으나 이 경우 양자화 테이블을 구하기 위하여 기존의 JPEG 부호화기보다 수백, 수천배 혹은 그 이상의 연산 시간이 필요하다. 이러한 연구들은 양자화 테이블을 구하기 위하여 많은 추가의 연산이 필요한 반면, 본 논문에서 제안한 VTQM의 경우 연산량이 표준 JPEG에서 거의 추가되지 않으면서도 우수한 성능을 보여주고 있다.

## II. 가변 임계값을 가지는 양자화 방법

### 2.1 JPEG 베이스라인 표준안

JPEG 표준안은 정지 영상 데이터의 압축 규약으로서 영상 압축 방법에 대한 상세한 기술보다는 기본적인 구성과 기능, 정밀도 및 오차에 대한 권고 사항만이 명시되어 있고, 실제 구현에 관련된 부분에는 상당한 융통성을 부여하고 있다. 따라서 제작자는 기술적 요구, 효율 및 응용 등에 따라 조금씩 다른 JPEG 시스템을 구현할 수 있다(1~3, 10). JPEG 인코더는 원 영상을 입력받아 압축을 행하는 기능을, 디코더는 압축된 화일을 읽어서 복원 영상을 만들어 주는 기능을 수행하며 이들은 서로 대칭적인 구조를 가지고 있다. <그림 1>은 JPEG 인코더의 처리과정을 나타낸 것이다. 우선 입력 영상은 8×8 크기의 블록으로 나뉘어진다. 그런 다음 각각의 8×8 영상 블록은 DCT(Discrete Cosine Transform)를 통하여 64개의 계수(Coefficients)가 계산되고 양자화 테이블에 의해 양자화된다. 그런 다음 DC 계수는 DPCM(Differential Pulse Code Modulation)에 의해 부호화되며 AC 계수는 런랭스 부호화된다. JPEG의 경우 런랭스 부호화방법은 DCT 계수를 지그재그 스캔하고 이를 허프만 부호화(Huffman coding), 수학적 부호화(Arithmetic coding) 등의 가변 길이 부호 엔트로피 부호화(Variable Length Code Entropy Encoding) 단계를 거쳐 압축된다(8, 10~12).

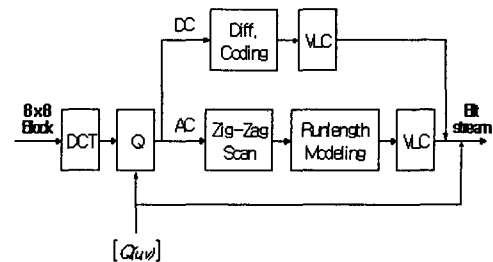


그림 1. JPEG 베이스라인 부호화기

본 절에서는 VTQM을 설명하기 위하여 양자화 방법과 허프만 부호화만 설명하기로 한다. FDCT의 결과값들은 실수로 계산되며, 식 (1)과 같이 양자화 과정을 거쳐 정수로 표현된다(9).

$$F^Q(u, v) = \lfloor \frac{F(u, v)}{Q_F \cdot Q(u, v)} + \theta \rfloor, \quad 0 \leq \theta \leq 1 \dots (1)$$

우선 각각의 FDCT 계수  $F(u, v)$ 들은 대응하는 양자화 테이블의 계수  $Q(u, v)$ 로 나뉘어지며, 비트율을 조정하기 위하여  $Q_F$ (Quality factor)로 다시 나뉘어지게 된다. 그러므로  $Q_F$ 가 커지면 비트율은 작아지나 상대적으로 화질이 떨어지게 된다(13). 양자화 테이블은 JPEG에서 권고한 테이블을 참조하거나 원영상에 가장 알맞은 테이블을 구하여 사용할 수 있다(8~10). 그런 다음 그 결과를 정수로 변환한다. 이때 임계값  $\theta$ 에 따라 반올림 양자화( $\theta = 0.5$ )와 내림 양자화( $\theta = 0$ )로 구분되며 JPEG에서는 일반적으로 반올림을 권고하고 있다.

FDCT 과정을 거쳐 양자화된 영상은 부호화 과정을 거친다. 부호화 과정은 크게 그룹화와 허프만 부호화의 두 단계로 나뉜다. JPEG에서는 양자화된 영상의 DC 성분과 AC 성분을 나누어서 그룹화 한다. DC계수를 그룹화하기 위해서는 바로 전  $8 \times 8$  블록의 DC 성분과의 차이를 DPCM을 써서 비손실 압축한다. 한편 AC계수의 그룹화는 DC계수의 그룹화와 비슷하지만, 지그재그 스캔을 하면서 런랭스 부호화하는 것이 다르다. 이때, 다수의 양자화된 AC 계수들은 0이 되기 때문에 런랭스 부호화할 경우 데이터가 압축된다.

*symbol-1 symbol-2*

Code : (런-랭스, 그룹번호) (AMPLITUDE) ... (2)

식 (2)에서 AC 코드를 보여 주고 있다. 여기서 런랭스는 지그재그 스캔 과정에서 몇 개의 0 값 뒤에 그 AC 성분이 오는지를 나타내며 맨 뒤에 EOB를 넣어서  $8 \times 8$  블록의 마지막임을 표시한다. 이 결과는 최종적으로 DC 성분을 위한 허프만 테이블과 AC 성분을 위한 허프만 테이블을 이용하여 압축되어 부호화되며 이 때 허프만 테이블은 JPEG 표준안에서 주어진 테이블을 이용하거나 영상에 따라 최적화된 테이블을 이용하게 된다(3.11).

## 2.2 가변 임계값을 가지는 양자화 방법

본 절에서는 JPEG 영상의 비트율 대비 PSNR을 개선

하기 위하여 가변 임계값을 가지는 양자화 방법(VTQM)을 제안한다. 영상 압축에 있어서 중요한 성능 평가 요소는 비트율(압축률)과 화질이다. 일반적으로 이 두 가지 요소는 상충 관계에 있어서 비트율을 작게 하면 화질이 나빠지고, 화질을 개선하면 비트율이 커지게 된다. 이런 관점에서 JPEG 표준에서 일반적으로 많이 쓰이는 반올림 양자화와 내림 양자화의 경우를 비교해 보자. 앞 절의 결과에서 보는 바와 같이 일반적으로 같은  $Q_F$ 에서 내림 양자화의 그룹화 결과는 반올림 양자화의 결과에 비하여 작은 정수로 표시되며 결과적으로 그룹번호가 작게 될 가능성이 커지게 된다. 한편 JPEG 영상의 허프만 부호화과정을 살펴보면 일반적으로 그룹번호가 작을수록 허프만 부호화된 결과도 적은 비트수로 표시되며(JPEG 허프만 테이블 참조) *symbol-2(AMPLITUDE)*를 표현하는데 필요한 비트

수도 작게 된다. 극단적으로  $\frac{F(u, v)}{Q_F \cdot Q(u, v)} = 0.6$ 의 경우 내림 양자화를 하면 1개의 전체 심볼이 없어지고 다음 심볼의 런랭스만 증가하게 된다. 그러므로 같은  $Q_F$ 에서는 내림 양자화를 할 경우 반올림 양자화에 비하여 작은 비트수로 영상을 표현할 수 있게 된다. 그러나 내림 양자화는 반올림 양자화에 비하여 양자화·역양자화 과정에서 많은 오차를 가지게 되므로 영상의 화질은 떨어지게 된다.

본 논문에서 제안하는 VTQM은 그룹번호가 낮을수록 허프만 부호화 과정에서 필요한 비트수가 작아진다는 점에 착안하여, 양자화방법에 따라 압축률에 영향을 주는 부분과 그렇지 않은 부분으로 나누어서 각기 다른 임계값을 적용한다. 예를 들어 식 (1)에서  $\frac{F(u, v)}{Q_F \cdot Q(u, v)} = 1.6$ 의 경우를 가정하자. 이 값은 양자화 방법( $\theta$ 의 값)에 따라 1 또는 2로 표현될 것이다. 그런데 1과 2는 서로 다른 그룹에 속해 있으므로 허프만 부호화 할 경우 다른 비트 수로 표현될 것이며 *symbol-2*를 표시하는데 필요한 비트 수도 달라진다. 한편  $\frac{F(u, v)}{Q_F \cdot Q(u, v)}$ 가 2.6인 경우에는 양자화 되면 2 혹은 3이 되므로 같은 그룹에 속하게 되어 허프만 부호화해도 같은 비트수로 표현되므로 JPEG 압축 파일의 크기에는 영향을 미치지 않게 된다. 본 논문에서는 다음과 같은 두 개의 성질을 관찰하여 양자화 방법에 따라 압축률에 영향을 주는  $\frac{F(u, v)}{Q_F \cdot Q(u, v)}$ 의 값의 범위를 그룹경계영역으로 정의한다.

**관찰 1** 그룹번호와 그룹내의 AMPLITUDE의 갯수와의 관계

양자화 테이블에서 그룹 n에는 2<sup>n</sup>개의 AMPLITUDE들이 존재하며 따라서 그룹 내에서 AMPLITUDE를 표현하기 위하여 n-bit의 symbol-2가 필요하다.

**관찰 2**

AMPLITUDE N은 그룹 [ ln|N| + 1 ] 에 속한다.

**정의 1** 그룹경계영역, 경계영역합수

양자화 방법에 따라 다른 그룹으로 되는  $\frac{F(u,v)}{Q_F \cdot Q(u,v)}$ 의 값의 범위를 그룹경계영역이라고 하며 이를 판별하기 위한 부울(Boolean) 함수 Bound(X)를 식으로 표시하면 다음과 같다.

$$\text{Bound}(X) = \begin{cases} 1, & \text{if } 2n-1 \leq X < 2n, \text{ 단 } n \text{은 } 0 \text{ 이상의 정수} \\ 0, & \text{기타의 경우} \dots\dots\dots (3) \end{cases}$$

$\frac{F(u,v)}{Q_F \cdot Q(u,v)}$ 의 값이 그룹경계영역 내의 값일 경우  $(\text{Bound}(\frac{F(u,v)}{Q_F \cdot Q(u,v)})=1)$ 에는 양자화 방법(식(1)의  $\theta$  값)에 따라 양자화 되면 2n-1 혹은 2n로 된다. 그러면 관찰 2에 의해서 각각의 경우 n 혹은 n+1 그룹에 속하게 된다.

기존의 양자화 방법에서는  $\frac{F(u,v)}{Q_F \cdot Q(u,v)}$ 가 그룹 경계 영역의 값이 아닌 경우에도 똑같은  $\theta$ 가 적용된다. 예를 들어 반올림 양자화에서  $\theta$ 는 0.5로 내림 양자화에서  $\theta$ 는 0으로 고정된 상수이다. 그러나 VTQM에서는 PSNR의 향상을 위하여 이 값을 상수로 두지 않고  $\frac{F(u,v)}{Q_F \cdot Q(u,v)}$ 의 값에 의해 결정되는 변수로 한다. 즉 Bound( $\frac{F(u,v)}{Q_F \cdot Q(u,v)}$ )의 값에 따라  $\theta$ 의 값을 다르게 정하는 것이다. 앞서 언급한대로 VTQM에서는 그룹경계영역이 아닌 경우에는 화질을 좋게 하기(오차를 줄이기) 위하여 반

올림 양자화와 같이  $\theta$ 는 0.5로 정하고 그룹경계영역 내에서는  $\theta$ 를 다르게 정한다. 본 논문에서 제안하는 VTQM은 다음과 같이 정의된다.

**정의 2** VTQM, 양자화 상수 :  $\theta V$

$$F^Q(u,v) = \left\lfloor \frac{F(u,v)}{Q_F \cdot Q(u,v)} + 0.5 - \theta V \times \text{Bound}\left(\frac{F(u,v)}{Q_F \cdot Q(u,v)}\right) \right\rfloor \dots\dots\dots (4)$$

예를 들어  $\frac{F(u,v)}{Q_F \cdot Q(u,v)} = 1.74$ , 양자화 상수  $\theta V = 0.3$ 인 경우를 살펴보자. 정의 1에 의해 이 값은 그룹 1과 그룹 2를 나누는 그룹경계영역 내에 존재하고  $(\text{Bound}(\frac{F(u,v)}{Q_F \cdot Q(u,v)})=1)$ , 따라서  $F^Q(u,v)$  값은 1로 양자화 되어 그룹 1이 된다.  $\theta V = 0$ 인 경우 VTQM은 반올림양자화와 동일한 결과가 나오며  $F^Q(u,v)$  값은 2로 양자화 되어 그룹 2가 된다. 따라서 반올림 양자화 방법은 VTQM의 특수한 경우로 생각할 수 있다.

### III. 실험 방법 및 결과

본 절에서는 제안된 VTQM의 성능을 실험을 통하여 측정한다. 이를 위하여 우선 VTQM에서 최적의 양자화 상수  $\theta V$ 를 구하기 위하여  $Q_F$ 와  $\theta V$ 를 바꾸어 가면서 실험하였다. 그리고 주어진  $\theta V$ 를 이용한 VTQM과 다른 양자화 방법들 간의 성능을 비교하였다. 본 논문에서는 Independent JPEG Group에서 제공한 JPEG Src Ver. 5(14)를 수정하여 사용하였으며, 양자화 테이블과 허프만 테이블은 모두 JPEG 표준안에서 제시된 것을 사용하였다. 실험 영상으로는 각각 512x512 크기의 Lenna 영상, Zelda 영상, Babbon 영상, Wool 영상, Olea 영상 등을 사용했다. 이들 영상은 여러 영상처리 분야의 성능 측정 및 평가에서 가장 널리 사용되는 표준 영상들이다.



그림 2. Baboon 영상

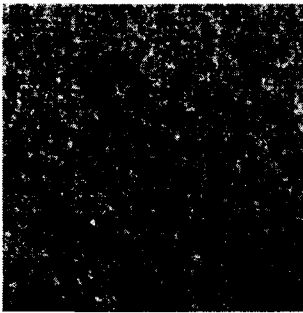


그림 3. Wool 영상

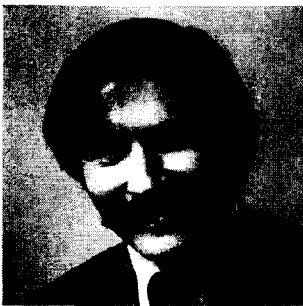


그림 4. Oleh 영상

일반적으로 영상 압축 알고리즘의 성능 및 압축된 영상의 화질을 평가하기 위한 척도로는 압축률(Compression Ratio), 비트율(bpp) 대비하여 SNR, PSNR, RMSE (Root Mean Square Error) 등이 사용되어 왔다[15]. 이 중에서 본 논문에서는 영상의 비트율과 PSNR을 성능 측정의 척도로 한다. PSNR의 정의는 아래의 식과 같다. 식에서  $N_2$ 은 영상의 전체 화소수이고  $f(i, j)$ 는 원영상의 화소값이며  $f'(i, j)$ 는 복원된 영상의 화소값이다.

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

$$MSE = \frac{1}{N^2} \sum_{i=1}^{N-1} \sum_{j=1}^{N-1} |f(i, j) - f'(i, j)|^2 \dots\dots (5)$$

우선, VTQM에서 최적의 양자화 상수  $\theta_V$ 를 구하기 위하여 QF와  $\theta_V$ 를 바꾸어 가면서 실험하였다. <그림 5>는 Lenna 영상에 대한 결과이다. 실험 결과 양자화 상수  $\theta_V$ 가 0.15일 때 가장 PSNR이 우수하였으며 다른 영상의 경우에도 비슷한 결과를 얻었다. 양자화 상수  $\theta_V$ 가 커지면 더 낮은 비트율을 얻게 되나, 오차도 커지게 되어 비트율 대비 PSNR은 작아지는 결과를 보였다.

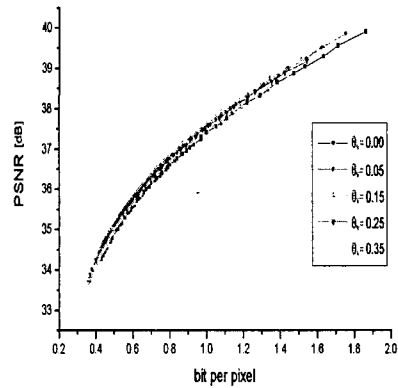


그림 5. 양자화 상수  $\theta_V$ 에 따른 비트율과 PSNR(Lenna 영상)

앞에서 측정된 최적의  $\theta_V$  값 0.15를 이용하여 VTQM과 다른 양자화 방법들 간의 성능을 비교하였다. 그림 6과 표 1은 양자화 방법에 따른 비트율과 화질과의 관계에 대한 실험 결과를 나타낸 것이다. Lenna 영상, Zelda 영상과 Olea 영상의 경우에는 비트율을 0.5~2.0까지 실험하였고 Baboon 영상과 Wool 영상의 경우에는 1.0~2.0까지 실험하였다. 이것은 Baboon 영상이 상대적으로 압축이 작게 되기 때문이다.

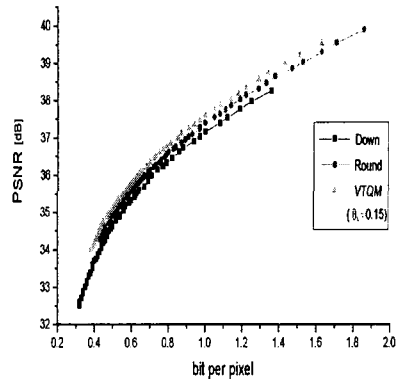


그림 6. 양자화 방법에 따른 비트율과 PSNR(Lenna 영상)

실험 결과 Lenna 영상의 경우 VTQM은 반올림 양자화에 비하여 비트율(bpp)에 따라 PSNR이 약 0.2 ~ 0.3dB 개선되었으며 내림 양자화를 적용한 경우보다는 약 0.4 ~ 0.5dB 개선되는 결과를 보여주고 있다. 다른 영상들의 경우에도 비슷한 결과를 보여주고 있다. 이것은 기존의 양자화 테이블을 변환을 이용한 양자화 방법 연구에서 같은 비트율(bpp)에서 PSNR이 0.8 ~ 1.7dB정도 개선되는 것에 비하여 상대적으로 성능 향상이 작다. 그러나 기존의 연구에서는 양자화 테이블을 구하기 위해 많은 추가 연산이 필요한 것과 달리 본 논문에서 제안한 VTQM은 연산량이 표준 JPEG에서 거의 추가되지 않으면서도 우수한 성능을 보여주고 있다.

#### IV. 결론

본 논문에서 제안한 VTQM은 JPEG표준안에서 권고한 양자화 테이블과 허프만 테이블을 이용하면서, DCT 계수를 정수부분의 값에 따라 각각 다른 임계값을 적용하므로서 화질을 반올림 양자화와 비슷하게 유지하면서 압축률을 증가(비트율을 감소)시킬 수 있다.

본 논문에서 제안한 VTQM과 반올림 양자화를 적용한 표준 JPEG을 비교하여 실험한 결과 흑백의 Lenna 영상의 경우, 비트율(bpp)에 따라 PSNR이 약 0.2 ~ 0.3dB 개선되었으며 내림 양자화를 적용한 경우보다는 약 0.4 ~ 0.5dB 개선되는 결과를 보여주고 있다. 또한 다른 영상들도 비슷한 결과를 보여주고 있다.

기존의 양자화 방법에 관한 연구들은 주로 영상에 최적인 양자화 테이블을 구하는 연구(8, 9)등이 있으며 [8]의 경우 영상에 따라 같은 비트율(bpp)에서 PSNR이 0.8 ~ 1.7dB정도 개선되는 것으로 발표되고 있다. 이러한 연구들은 양자화 테이블을 구하기 위하여 많은 추가의 연산이 필요한 반면, 본 논문에서 제안한 VTQM의 경우 연산량이 표준 JPEG에서 거의 추가되지 않으면서도 우수한 성능을 보여주고 있다.

#### 참고문헌

- [1] G. K. Wallace, "The JPEG still picture compression standard," Comm of the ACM, vol. 34, pp. 31-44, April 1991.
- [2] G. K. Wallace, "The JPEG still picture compression standard," IEEE Trans. on Consumer Electronics, vol. 38, pp. 18-34, Feb. 1992.
- [3] ISO/IEC JTC1 10918-1 } ITU-T Rec. T.81, Information technology-Digital compression and coding of continuous-tone still images : Requirements and guidelines, 1994.
- [4] ISO/IEC 10918-2 } ITU-T Rec. T.83, Information technology - Digital compression and coding of continuous-tone still images: Compliance testing, 1995.
- [5] ISO/IEC 10918-3 } ITU-T Rec. T.84, Information technology - Digital compression and coding of continuous - tone still images : Extensions, 1996.
- [6] R. C. Gonzalez and R. E. Woods, Digital Image Processing, Addison - wesley, 1992.
- [7] W. Pennebaker, "JPEG technical specification, revision 8," Working Document No. JTC1/SC2 /WG10/JPEG-8-R8, Aug. 1990.
- [8] S. W. Wu and A. Gersho, "Rate - constrained picture adaptive quantization for JPEG baseline coders," ICASSP '93, Minneapolis, MN, pp. V.389 - 392, April 1993.
- [9] J. J. Hwang and K. H. Yang, "Optimal variable quantization for JPEG extensions," IEEE Intl. Conf. on Consumer Electronics, Chicago, IL, pp.266-267, June 1995
- [10] W. B. Pennebaker and J. L. Mitchell, JPEG Still Image Data Compression Standard, Van Nostrand Reinhold, New York, 1993.

- [11] G. A. Baxes, Digital Image Processing, Wiley - Interscience, 1994.
- [12] 최우영, 박래홍, "정지영상 압축기술," 한국통신학회지, 제10권, 제4호, pp.4-16, 1993년 4월.
- [13] 김성훈, 오승준, "DCT 부호화 영상의 블록화 현상 제거 알고리즘," 정보과학회논문지(A), 제24권, 제9호, pp.888-896, 1997년 9월.
- [14] <http://www.ijg.org/>
- [15] 허봉식, 김민환, "블록 DCT기반의 시각 적응적 이미지 압축에 관한 연구," 정보과학회논문지(A), 제22권, 제10호, pp.1405-1415, 1995년 10월.

## 저자소개

제7권 제3호 (2002. 9) pp.148 참조