

유한 상태 분류 벡터 양자기를 이용한 라플라시안 피라미드 부호화 기법

(Laplacian Pyramid Coding Technique using a
Finite State-Classified Vector Quantizer)

朴 涉 亨* 李 商 郁*

(Seop Hyeong Park and Sang Uk Lee)

要 約

본 논문에서는 라플라시안 피라미드 구조와 DCT 영역에서의 계층적 유한 상태 분류 벡터 양자기 (FS-DCT-CVQ)를 이용한 영상 부호화 기법을 제안하였다. 먼저 라플라시안 피라미드 구조에서 DCT-CVQ를 이용하는 경우의 고정 전송율에서의 최적 비트 할당 문제를 기술하였고, 최적 비트 할당이 이루어진 점근적 경우에서 각 라플라시안 피라미드 평면을 스칼라 양자화하는 경우에 대한 부호화 이득을 구하였다. 다음으로, 라플라시안 피라미드 구조를 이용하여 전체 최소 평균 자승 오차의 측면에서 부호화 이득을 얻을 수 있다는 것을 실험적으로 보였다. 마지막으로, 라플라시안 피라미드의 계층적 상관성을 이용하는 FS-DCT-CVQ를 제안하였다. 제안한 부호화 기법은 전송율 0.192bpp, 0.385bpp에서 각각 30.33dB 32.45dB의 높은 PSNR을 보였다.

Abstract

In this paper, we propose an image coding scheme which combines the Laplacian pyramid structure and a hierarchical finite state classified vector quantizer in the DCT domain, namely FS-DCT-CVQ. First, an optimal bit allocation problem for fixed rates DCT-CVQ on the Laplacian pyramid structure is described. In an asymptotic case, with an optimal bit allocation, a coding gain over scalar quantization of each Laplacian plane is derived. Second, it is experimentally shown that the Laplacian pyramid structure provides a considerable coding gain in the sense of total MMSE (minimum mean squared error). Finally, we propose an FS-DCT-CVQ which exploits the hierarchical correlation between the Laplacian planes. Simulation results on real images show that the proposed coding scheme can reconstruct an image with 30.33 dB at 0.192 bpp, 32.45 dB at 0.385 bpp, respectively.

I. 서 론

라플라시안 피라미드 부호화 기법⁽¹⁾은 점진적 영상 전송에 적합한 구조를 지니고 있으며, 입력 신호를 특정한 주파수 대역별로 분리한 후에 부호화한다

*正會員, 서울大學校 制御計測工學科
(Department of Control and Instrumentation Eng.,
Seoul National University)
接受日字: 1989年 9月 19日

는 측면에서 subband 부호화 기법^[2,3]과 유사하다. 이러한 구조에서는 입력 신호를 주파수 대역별로 분리하는데 사용되는 필터^[4,5] 분리 대역 수의 결정, 각 대역의 특성에 적합한 부호화 기법의 선택 등이 중요한 문제이며, 특히 주어진 전송율 하에서 최소 왜곡을 얻기 위해서는 왜곡율 함수(distortion-rate function)에 근거한 최적비트 할당^[6]이 필요하다.

본 논문에서는 라플라시안 피라미드 구조와 DCT 영역에서의 계층적 유한 상태(finite state) 분류벡터 양자기를 이용한 영상 부호화 기법을 제안하였다. 먼저 라플라시안 피라미드 신호를 부호화하기 위한 방법으로 DCT-CVQ 기법^[7]을 사용하는 경우의 최소 왜곡을 갖는 최적 비트 할당식을 유도하였고, 스칼라 양자기에 대한 DCT-CVQ의 점근적(asymptotic) 부호화 이득(coding gain)을 구하였다. 또한 라플라시안 피라미드 구조가 갖는 장점을 보이기 위하여, 해석적으로 간단한 스칼라 양자기에 대하여 라플라시안 피라미드 구조가 갖는 MMSE(minimum mean squared error) 측면에서의 이득을 이론적으로 구하였으며, 실제 영상을 이용하여 1.0bpp 미만의 저 전송율에서의 부호화 이득을 실험적으로 구하였다. 그 결과 라플라시안 피라미드 구조는 0.5bpp이하의 저 전송율에서 큰 이득을 갖는 것을 알 수 있었다. 한편 라플라시안 피라미드는 원 영상의 준대역통과(quasi-band-passed) 신호들의 집합으로 볼 수 있으므로 원 영상의 edge 부분만이 나타나는 라플라시안 신호의 인접한 평면 사이에는 층간의 상관성이 존재한다. 특히 인접한 라플라시안 평면의 동일한 위치에 있는 두 블럭의 AC분산값은 서로 밀접한 관계를 갖는 것을 알 수 있었으며, 이는 부호화하고자 하는 블럭의 AC분산값이 상층의 같은 위치에 있는 블럭의 AC분산값에 큰 영향을 받는다는 것을 의미한다. 본 논문에서는 이러한 계층적 상관성을 이용하여 DCT영역에서의 FSCVQ(finite state classified vector quantizer)를 구성하였다. 이를 본 논문에서는 FS-DCT-CVQ(finite state discrete cosine transform classified vector quantizer)라 부르겠다.

II. 라플라시안 피라미드 구조에서의 최소 왜곡을 갖는 최적 비트 할당

본 절에서는 라플라시안 피라미드 구조에 DCT-CVQ를 도입하는 경우에 있어서의 최적 비트 할당식을 유도하고, 스칼라 양자기에 대한 부호화 이득에 관해서 서술하기로 한다. 라플라시안 피라미드 구조에서는 입력 영상 G_0 가 주어지면, L 층의 가우시안

피라미드 $\{G_\ell, \ell=0, 1, \dots, L-1\}$ 와 라플라시안 피라미드 $\{L_\ell, \ell=0, 1, \dots, L-1\}$ 를 구성한다. 이때 입력영상의 크기가 $N \times N$ 이라하면, ℓ 층의 영상의 크기는 $4^{-\ell} N \times N$ 이다 [6]에서는, 스칼라 양자기를 사용하여 라플라시안 피라미드 부호화에서의 최적 비트 할당식을 다음과 같이 유도하였다. 각 라플라시안 평면 L_ℓ 에 할당된 전송율을 R_ℓ 라 하고, 전체 전송율을 R 이라 하면, 다음의 식이 성립한다.

$$R = \sum_{\ell=0}^{L-1} \frac{1}{4^\ell} R_\ell. \quad (1)$$

그리고, L_ℓ 에서의 평균자승오차를 d_ℓ , 최종적으로 복원된 영상의 평균자승오차를 D 라하면, 다음이 성립한다.

$$D = \sum_{\ell=0}^{L-1} d_\ell. \quad (2)$$

또한, L_ℓ 의 분산을 σ_ℓ^2 , 원 영상의 분산을 σ^2 이라 하면 다음 식이 성립한다.

$$\sigma^2 = \sum_{\ell=0}^{L-1} \sigma_\ell^2. \quad (3)$$

(2), (3)식은 이상적인 저대역 통과 필터를 사용한 경우에 성립하고, 필터의 천이 대역(transition band)이 넓어질수록 부정확해진다. 스칼라 양자기를 사용했을 때 왜곡값은 다음과 같이 주어진다.^[8]

$$d_\ell = \epsilon^2 \sigma_\ell^2 2^{-2} R_\ell. \quad (4)$$

여기에서 ϵ^2 은 양자화 계수를 나타낸다. 그러면, 전송율이 주어진 경우의 비트 할당 문제는 다음과 같이 표현된다.

$$\min \sum_{\ell=0}^{L-1} d_\ell, \quad (5)$$

subject to (1).

잘 알려진 Lagrange multiplier 기법을 이용하면, 다음과 같은 해를 얻을 수 있다.^[6]

$$R_\ell = \frac{3R}{4(1-\alpha)} - \frac{1}{3} + \frac{\alpha L}{(1-\alpha)} + \ell + \frac{1}{2} \log_2(\sigma_\ell^2) - \frac{1}{2} \log_2 \left\{ \prod_{k=0}^{L-1} (\sigma_k^2) 4^{-k} \right\}^{3/(4(1-\alpha))}, \quad (6)$$

$$D(R) = \epsilon^2 \left\{ \prod_{\ell=0}^{L-1} (\sigma_\ell^2) 4^{-\ell} \right\}^{3/(4(1-\alpha))} \frac{4}{3} (1-\alpha) \cdot 2^{-2(3R/(4(1-\alpha)) + (1-\alpha)/(1-\alpha) - 1/3)}, \quad (7)$$

여기에서 $\alpha = (1/4)^L$.

라플라시안 피라미드 신호를 DCT-CVQ를 이용하여 부호화하는 경우의 비트 할당은 다음과 같이 구할 수 있다. k차원 벡터 양자기의 DM(distortion measure)로서 (8)과 같이 정의되는 평균자승오차(mean squared error : MSE)인 ℓ_2 norm을 사용한다고 가정한다.

$$d(x, y) = \frac{1}{k} \sum_{i=1}^k (x_i - y_i)^2 \tag{8}$$

이때, 전송율 R이 충분히 큰 경우의 최소 왜곡 경계치는 다음과 같이 주어진다.⁹⁾

$$D_L^k(R) = N^{-2/k} \frac{k}{(k+2)\pi} \{ (k/2) \Gamma(k/2) \}^{2/k} \parallel P(x) \parallel_{k/(k+2)} \tag{9}$$

여기에서 k는 벡터 x의 차수이고 p(x)는 x의 확률 밀도 함수를 의미하며 p(x)의 ℓ_s norm은 다음과 같이 정의된다.

$$\parallel P(x) \parallel_s = \left[\int p(x)^s dx \right]^{1/s} \tag{10}$$

일반적으로 memoryless 독립 동일 분산(iid:independent identically distributed) source에 대하여 $\parallel p(x) \parallel_{k/(k+2)}$ 는 (11)과 같이 표시되는 벡터 분산 $\bar{\sigma}^2$ 에 비례하는 형태로 나타난다.⁹⁾

$$\bar{\sigma}^2 = \left(\prod_{i=1}^k \sigma_i^2 \right)^{1/k} \tag{11}$$

이때, 편의상 다음과 같이 정의한다.

$$\parallel p(x) \parallel_{k/(k+2)} = P_k \bar{\sigma}^2 \tag{12}$$

$$C_k = \frac{1}{(k+2)\pi} \{ (k/2) \Gamma(k/2) \}^{2/k} \tag{13}$$

그리고 $r=R/k$ 라 하면 (9)식은 다음과 같이 표시할 수 있다.

$$D_L^k(R) = k C_k P_k 2^{-2r} \bar{\sigma}^2 \tag{14}$$

DCT-CVQ에서 부류의 수를 n이라 하고 L_i 에서 j 부류에 속하는 $M \times M$ 크기의 DCT블럭에서 분할된 벡터의 수를 m_{ij} , 그리고 j번째 부류의 i번째 벡터의 차수를 k_{ij} 라고 하면,

$$\sum_{i=1}^{m_{ij}} k_{ij} = M^2 \text{ for all } \ell, j \tag{15}$$

그리고 ℓ 번째 층의 왜곡값을 D_ℓ 이라 하면 전체 왜곡값은 다음과 같이 표시된다.

$$D^* = \sum_{i=0}^{L-1} D_i \tag{16-a}$$

$$D_i = \frac{1}{M^2} \sum_{j=1}^n Pr_{i,j} \sum_{i=1}^{m_{ij}} k_{ij} C_{k_{ij}} P_{k_{ij}} 2^{-2r_{i,j}} \bar{\sigma}^2_{i,j} \tag{16-b}$$

여기에서, $Pr_{i,j}$ 는 L_i 에서의 j번째 부류에 속한 벡터의 발생 확률을 의미하며, $r_{i,j}$ 와 $\bar{\sigma}^2_{i,j}$ 는 각각 ℓ 번째 층의 j번째 부류의 i번째 벡터의 평균 전송율과 벡터 분산을 나타낸다. 그리고 전체 평균 전송율은 다음과 같이 주어진다.

$$R = \frac{1}{M^2} \sum_{i=0}^{L-1} \frac{1}{4^i} \sum_{j=1}^n Pr_{i,j} \sum_{i=1}^{m_{ij}} k_{ij} r_{i,j} \tag{17}$$

이때 전송율이 주어질 경우의 최적 비트 할당 문제를 다음과 같이 표현할 수 있다.

$$\min \sum_{i=1}^{L-1} \frac{1}{M^2} \sum_{j=1}^n Pr_{i,j} \sum_{i=1}^{m_{ij}} k_{ij} C_{k_{ij}} P_{k_{ij}} 2^{-2r_{i,j}} \bar{\sigma}^2_{i,j} \tag{18}$$

subject to (17).

역시 Lagrange multiplier 기법을 이용하여 위 문제의 해를 구하기 위하여 확장함수를 다음과 같이 정의한다.

$$f(r_{i,j}, \lambda) = \sum_{i=1}^{L-1} \frac{1}{M^2} \sum_{j=1}^n Pr_{i,j} \sum_{i=1}^{m_{ij}} k_{ij} C_{k_{ij}} P_{k_{ij}} 2^{-2r_{i,j}} \bar{\sigma}^2_{i,j} + \lambda \left\{ \sum_{i=0}^{L-1} \frac{1}{4^i} \frac{1}{M^2} \sum_{j=1}^n Pr_{i,j} \sum_{i=1}^{m_{ij}} (k_{ij} r_{i,j}) - R \right\} \tag{19}$$

f를 $r_{i,j}$ 에 대하여 편미분을 한 것을 0으로 두고, 이를 풀면 다음의 식을 얻을 수 있다.

$$r_{i,j} = \frac{3R}{4(1-\alpha)} - \frac{1}{3} + \frac{\alpha L}{(1-\alpha)} + \ell + \frac{1}{2} \log_2 (C_{k_{ij}} P_{k_{ij}} \bar{\sigma}^2_{i,j}) - \frac{1}{2} \log_2 \left\{ \prod_{p=0}^{L-1} \prod_{q=1}^n \prod_{s=1}^{m_{pq}} (C_{k_{pqs}} P_{k_{pqs}} \bar{\sigma}^2_{pqs})^{k_{pqs}/M^2} \right\}^{3/4(1-\alpha)} \tag{20}$$

$$D(R) = \left\{ \prod_{p=0}^{L-1} \prod_{q=1}^n \prod_{s=1}^{m_{pq}} (C_{k_{pqs}} P_{k_{pqs}} \bar{\sigma}^2_{pqs})^{k_{pqs}/M^2} \right\}^{3/4(1-\alpha)} \frac{4}{3} (1-\alpha)^{-2} 2^{-2\{3R/4(1-\alpha) + L\alpha/(1-\alpha) - 1/3\}} \tag{21}$$

여기에서 $\alpha = (1/4)^L$

(7)식의 왜곡을 함수를 $D_{LP-sq}(R)$ 이라 하고, (21)식의 왜곡을 함수를 $D_{LP-DCT-CVQ}(R)$ 라 하면, 라플라시안 피라미드 구조에서 스칼라 양자기에 대한 DCT-CVQ의 이득, 즉 $G_{sq/DCT-CVQ}$ 를 다음과 같이 구할 수 있다.

$$G_{sq/DCT-CVQ} = D_{LP-sq}(R) / D_{LP-DCT-CVQ}(R) - (\log_2 n) / M^2 \left\{ \varepsilon^2 \cdot \left[\frac{\sum_{p=0}^{L-1} (\sigma_p^2)^{1-p}}{\sum_{p=0}^{L-1} \sum_{q=1}^n \prod_{s=1}^{m_{pq}} (C_{k_{pqs}} P_{k_{pqs}} \bar{\sigma}_{pqs}^2)^{k_{pqs} / M^2 4^p}} \right]^{2^{L-k-1-\alpha}} \cdot 2^{-N \log_2 n / M^2} \right\} \quad (22)$$

여기에서 $\alpha = (1/4)^L$ 이고, $(\log_2 n) / M^2$ 은 부류 인덱스를 전송하는데 소요되는 비트 수를 의미한다. (22)식은 전송율 R이 충분히 큰 경우에만 적용할 수 있는데, 이때 이득은 전송율과는 무관하며 입력 신호의 스펙트럼 분포와 각 피라미드 평면의 DCT계수의 분산 분포, 그리고 피라미드 층수 L에 따라서 달라지는 것을 알 수 있다. 또한 DCT계수를 특정한 분포로 가정하더라도^[10] 분류기의 특성에 따라서 $\bar{\sigma}_{pqs}^2$ 가 달라지므로 (22)에서 구한 이득을 직접 계산하는 것이 쉽지 않다. 한편, $n=1, L=1, k=1$ 인 경우를 가정하면, (22)에서 구한 이득은 PCM에 대한 변환 부호화 기법의 부호화 이득^[8]과 일치하는 것을 알 수 있다.

III. 라플라시안 피라미드 구조의 부호화 이득 및 최적 피라미드 층수의 결정

본 절에서는 원 영상을 직접 부호화하는 경우에 대한 피라미드 구조를 이용하여 부호화하는 경우의 부호화 이득 및 주어진 전송율 하에서 최소 왜곡을 갖는 피라미드 층수의 결정에 대하여 논하기로 한다. 본 논문에서는 입력 신호는 1차 Markov신호의 가분리 2차원 모델^[8]로 가정한다. 그리고 원 영상과 모든 피라미드 신호는 스칼라 양자기를 사용하여 부호화한다고 가정한다.

1차원 Markov신호의 전력 스펙트럼 밀도(psd: power spectral density) 함수는 다음과 같이 주어진다.^[8]

$$S_{xx}(\omega) = \frac{1-\rho^2}{1+\rho^2-2\rho \cos \omega} \sigma_x^2, \quad (23)$$

여기에서 ρ 는 자기상관계수를 나타낸다. 이때 1차원

입력 신호에 대하여 L층의 라플라시안 신호를 구성하고 (24)를 이용한다.^[11]

$$\int_{\alpha}^{\beta} \frac{1-\rho^2}{1+\rho^2-2\rho \cos \omega} d\omega = 2 \tan^{-1} \left\{ \left(\frac{1+\rho}{1-\rho} \right) \tan(\beta/2) \right\} - 2 \tan^{-1} \left\{ \left(\frac{1+\rho}{1-\rho} \right) \tan(\alpha/2) \right\}. \quad (24)$$

그러면 l층의 분산 $\sigma_{x_i}^2$ 은 다음과 같이 구할 수 있다

$$\sigma_{x_i}^2 = 2 \int_{\alpha_{i+1}}^{\alpha_i} S_{xx}(\omega) d\omega, \quad (25)$$

$$\alpha_i = \begin{cases} (1/2)^i \pi, & \ell = 0, 1, \dots, L-1. \\ 0, & \ell = L. \end{cases} \quad (26)$$

2차원 가분리 모델에 대해서는 psd가 (23)식에 주어진 1차원 psd의 곱으로 주어지므로 이와 유사한 방법으로 각 피라미드 평면의 분산을 구할 수 있다. 이때 피라미드를 L층으로 분리하여 얻을 수 있는 부호화 이득은 다음과 같다.

$$G_L = 10 \log_{10} (D_{sq} / D_{LPC-sq}) \text{ (dB)}. \quad (27)$$

(27)에서 D_{sq} 는 (4)에서 $k=1$ 인 경우이고, D_{LPC-sq} 는 (7)에서 구한 값을 의미한다. 그림 1은 입력 신호를 $\rho=0.95$ 인 1차 Markov 신호의 2차원 가분리 모델로 가정한 경우에 전송율 변화에 따른 이득을 나타낸 것이다. 그림 1에서 알 수 있듯이 4bpp 이상에서는 2층 피라미드의 이득이 가장 크고, 그 이하에서는 3층 피라미드의 이득이 가장 크다. 이러한 이론적인 이득은 전송율이 충분히 크다는 가정하에서 유도된 것이므로 1.0bpp 미만의 저 전송율에서는 정확한 추세를 알 수 없다. 따라서 1.0bpp 미만의 전송율에서는 실제 영상에 대한 실험값을 제시하기로 한다. 실험에 사용된 영상은 512×512 크기를 갖는 LENA 영상이고, 부호화 기법으로는 전송율을 변화시키기 용이하고 저 전송율 부호화가 가능해야 하는데, 본 실험에서는 Chen과 Smith가 제안한 ADCT부호화 기법^[12]을 사용하였다. 그림 2는 이러한 결과를 1.0bpp 미만의 전송율에서 복원된 영상의 PSNR 값을 보인 것이다. 이 결과를 살펴보면 1.0bpp 미만의 저 전송율에서는 동일한 부호화 기법을 적용하더라도 라플라시안 피라미드 구조를 이용한 경우가 원영상을 직접 부호화하는 경우에 비하여 우수한 성능을 나타내며, 0.4bpp 이하에서는 약 1-4 dB 정도의 이득이 있다는 것을 알 수 있다. 이상의 결과로부터 라플라시

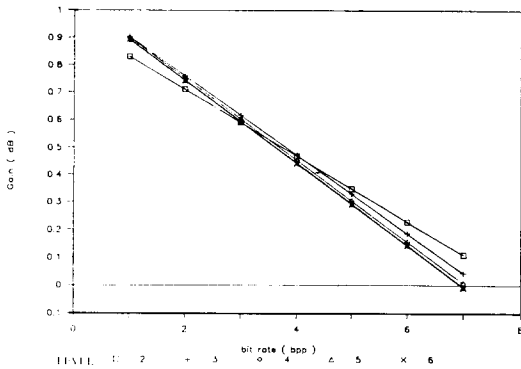


그림 1. 피라미드 층수와 전송율에 따른 이론적 부호화 이득 ($\rho=0.95$)
 Fig. 1. Theoretical coding gain for pyramid levels and bit rates ($\rho=0.95$).

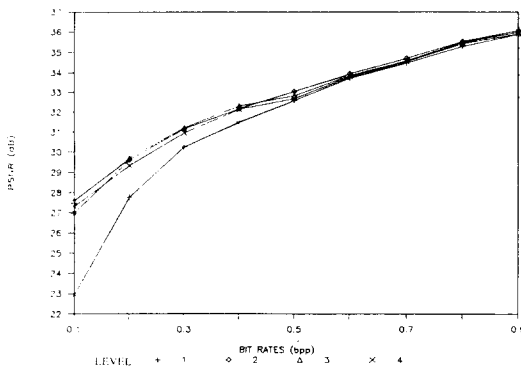


그림 2. ADCT 기법에 의한 부호화 결과
 Fig. 2. Coding results obtained by ADCT.

안 피라미드 구조를 사용하면 부호화 이득을 얻을 수 있다는 사실을 알 수 있다. 그리고 ADCT를 이용하는 경우는 3층의 피라미드를 구성할 때 가장 높은 이득을 얻을 수 있다. 그러나 (7), (18)에서 알 수 있듯이 라플라시안 피라미드 부호화 기법에서의 비트 할당 및 최소 왜곡값은 입력 영상의 스펙트럼 분포와 필터의 성능, 사용된 부호화 기법의 성능과 밀접한 관련이 있기 때문에 피라미드의 최적 층수는 이러한 요인에 따라서 변할 수 있다.

IV. 계층적 상관성을 이용한 FSVQ (finite state vector quantizer)

FSVQ는 memory VQ의 일종으로 인접한 영역에

있는 벡터 사이의 상관성을 이용하여 더 높은 감축율을 얻고자하는 기법이다.^{13, 14, 15} 본 절에서는 라플라시안 피라미드 구조에서 인접한 평면사이의 계층적 상관성을 이용한 FSVQ에 대하여 설명하기로 한다. 라플라시안 신호들은 준대역통과(quasi-bandpassed) 신호이기 때문에 피라미드 평면 사이에 상관성이 존재한다. 원 영상에서 가장자리 성분이 강한 신호는 여러 층의 평면에 걸쳐서 그 성분들이 나타나며 영상을 작은 크기의 부분력으로 세분하면 이러한 부분에 해당되는 부분력들은 AC분산이 큰 high activity 블록에 속한다. 반면에 가장자리 성분이 약하거나 없는 신호들은 주로 상층의 라플라시안 신호들만으로 표현이 가능하며 이에 해당되는 부분력들의 AC분산값은 작게 나타난다. 따라서 이웃한 피라미드 평면에서 동일한 위치에 있는 부분력들의 AC분산값은 높은 상관 관계를 갖는다. L_1 의 8×8 크기의 블록과 같은 위치에 있는 L_{i+1} 에서의 4×4 의 크기를 갖는 블록을 8×8 블록의 '대응 블록'이라고 하겠다. 그림 4는 L_1 의 각 블록들의 AC분산값과 L_2 의 대응 블록들의 AC분산값의 관계를 보여주는 산포도이다. 그림에서 가로 축은 L_2 의 4×4 블록의 AC분산값을, 세로 축은 L_1 의 8×8 블록의 AC분산값을 각각 나타낸다. 그림에서 알 수 있듯이 두 블록의 AC분산값은 어느 정도 선형적인 상관 관계를 가지고 있다. 즉, L_2 에 있는 블록의 AC분산값을 안다면, L_1 에 있는 블록의 AC분산이 가질 수 있는 값의 범위를 예측할 수 있으며 이러한 성질을 이용하여 FSVQ를 구성할 수 있다. 특히 라플라시안 신호에 대해서는 AC분산값에 의한 분류기가 효과적이므로⁽⁷⁾ AC분산값에 의해서 상태를 결정하면 분류 범위를 좁힐 수 있어서 효과적이다.

1	2	3	4	5	6	7	8
2	3	4	5	6	7	8	9
3	4	5	6	7	8	9	10
4	5	6	7	8	9	10	11
5	6	7	8	9	10	11	12
6	7	8	9	10	11	12	13
7	8	9	10	11	12	13	14
8	9	10	11	12	13	14	15

그림 3. DCT 블록에서의 벡터의 분할 형태
 Fig. 3. Partition scheme in the DCT block.

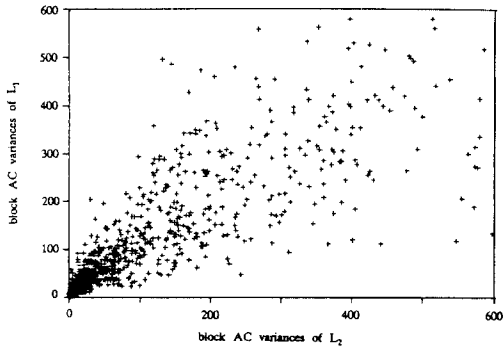


그림 4. L_2 와 L_1 의 블록 AC 분산 값의 산포도
 Fig. 4. Scatter diagram of block AC variances in L_2 and L_1 .

한편, 라플라시안 피라미드 신호는 원 영상에 비해서 상관성이 크게 줄어든 신호이지만 DCT등의 변환을 이용하면 더욱 높은 압축율을 얻을 수 있다.^[7] 본 논문에서는 라플라시안 피라미드 신호를 부호화하기 위하여 8×8 크기의 블록에 DCT를 취한 후 AC 분산의 크기에 따라서 8부류로 분류하고, 이를 작은 크기의 벡터로 분할한 후에 부호화하는 DCT-CVQ 방식을 채택하였다.^[7] 분할 방식은 그림 3 과 같은 가장 널리 사용되는 분할 방식을 선택하였으며, 모든 평면에 동일한 분할 방식을 사용하였다. 이때, 한 블록 내에 여러개의 벡터가 존재하기 때문에 이 모든 벡터의 상태를 정의하여 FSVQ를 구성하는 것이 쉽

지 않으므로 한 블록의 상태를 정의하고 그 블록 내의 모든 벡터는 같은 상태를 갖는다고 가정하였다. L_{i+1} 에서의 4×4 블록의 AC 분산값에 따라서 실험적으로 5 부류로 구분하고, 이 부류 인덱스를 L_i 에서의 8×8 블록의 '상태'로 정의하였다. 이때, 차기 상태 함수(next state function)는 대응 블록의 AC 분산값의 크기에 의해서 쉽게 결정된다. 부류를 나누는 방법은[11]에서와 같이 주어진 training 집합에서 각 부류에 동일한 수의 블록이 할당되도록 문턱값을 정하였다. 각 블록의 상태가 결정되면, 동일한 상태를 갖는 블록들로 각 상태의 training 집합을 구성하고 이들로부터 각 상태의 codebook을 설계하였다. 즉, 제안한 FS-DCT-CVQ는 p 개의 상태와 각 상태별로 q 개의 부류가 존재하여 전체적으로는 $p \times q$ 개의 부류가 있는 CVQ로 간주할 수 있다. 이때, p 개의 상태는 이미 알고 있는 정보로부터 예측할 수 있으므로, 부류 인덱스 전송에 필요한 비트 수는 $\log_2 q$ 이다. 따라서 본 논문에서 제안하는 FS-DCT-CVQ의 비트 할당은 $p \times q$ 개의 부류가 있는 CVQ의 비트 할당으로 대체하였으며, II 절에서 유도한 식을 이용하여 비트 할당을 수행하였다.

V. 실험 결과 및 검토

그림 5 에서 본 논문에서 제안한 FS-DCT-CVQ를 이용한 라플라시안 피라미드 부호화 기법의 블록도를 도시하였다. 그림 5 에서 T는 이산 여현 변환을 의미하고 T'는 이산 여현 변환의 역변환을 의미한다. 그리고 SQ는 스칼라 양자기를 의미하고, CVQ

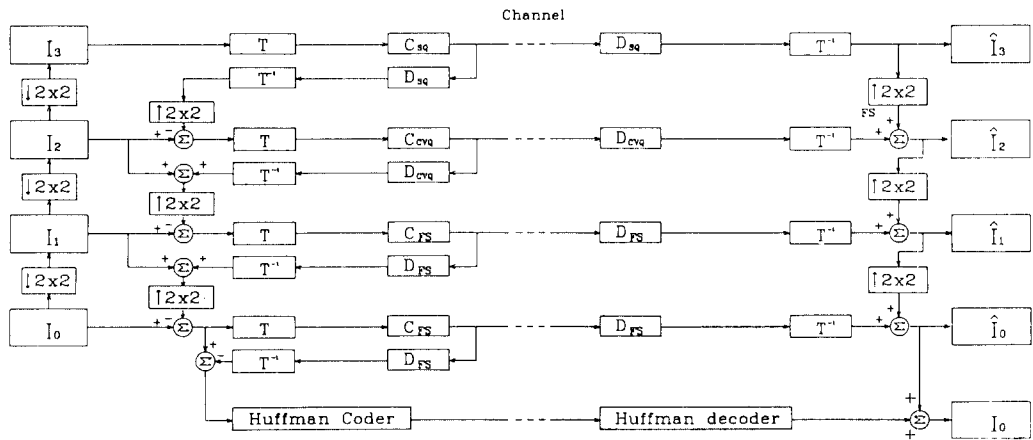


그림 5. DCT-CVQ와 FS-DCT-CVQ를 이용한 부호화 시스템의 블록선도
 Fig. 5. A block diagram of encoding scheme using DCT-CVQ and FS-DCT-CVQ.

는 [7]에서 제안한 AC 분산 분류기를 이용한 분류 벡터 양자기, FS는 본 논문에서 제안한 FS-DCT-CVQ를 의미한다. 그리고 C와 D는 각 부호화 기법의 부호기와 복호기를 나타낸다. 부호화 실험에 사용된 영상은 512×512크기를 갖는 LENA영상과 BABOON 영상으로 LENA는 training 집합에 포함되지 않았으며 BABOON은 training 집합에 포함되었다. 피라미드 층수는 4로 하였으며, 피라미드 구성을 위한 필터로는 [5]에서 제안한 half-band FIR 필터를 사용하였다. III 절에서는 ADCT를 사용하여 부호화할 경우에 3층이 최적인 것을 보였는데, 여기에서는 L₂에 DCT-CVQ를 사용했기 때문에 층수가 증가되더라도 ADCT를 사용한 경우에 비해서 적은 왜곡값을 얻을 수 있기 때문에 그 효과를 상쇄할 수 있고, 한 층의 라플라시안 평면에 FS-DCT-CVQ를 적용하기 위해서는 전단의 라플라시안 평면이 필요하기 때문에 4층의 피라미드를 구성하였다. L₃는 ADCT부호화 기법으로 부호화하였고, L₂는 DCT-CVQ 기법으로, L₁과 L₀는 본 논문에서 제안한 FS-DCT-CVQ 기법으로 부호화하였다. II 절에서 유도한 비트 할당식에서 k=1인 경우가 스칼라 양자기를 의미하므로, 각 피라미드 평면에 여러가지 부호화 기법을 혼용한 경우에도 본 논문에서 유도한 비트 할당식을 적용할 수 있다. 일반 영상의 DCT계수는 DC는 균일 분포, AC는 Laplacian 분포를 갖는다고 알려져 있으나,^[6] 라플라시안 신호의 평균은 0이므로 DCT 계수의 DC도 0을 중심으로 한 Laplacian 분포와 유사하게 나타난다. 따라서 본 논문에서는 모든 DCT계수가 Laplacian 분포를 갖는다고 가정하고 비트 할당식을 적용하였다. 그러나 그림 5에서 알 수 있듯이 본 논문에서 사용한 부호화 기법에서는 전단의 라플라시안 신호의 양자화 오차가 다음단의 피라미드 신호를 구하는 과정에서 보상되기 때문에 비트 할당 과정에서도 이를 고려해야 한다. 이에 관한 자세한 알고리즘은 부록에

제시하였다. 표 1, 2에는 FS-DCT-CVQ의 각 상태와 부류 별 발생 확률 분포 및 비트 할당표를 제시하였다. 표를 살펴보면 각 상태별로 각 부류의 분포가 다른 것을 알 수 있으며, 이것은 각 상태의 특성에 맞는 별도의 codebook을 구성하여 효과적인 부호기를 구성할 수 있다는 것을 의미한다.

표 2. FS-DCT-CVQ에서의 비트 할당표.

Table 2. Bit allocation table in FS-DCT-CVQ.

상태	부류	벡터														
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0
	1	2	1	1	0	0	0	0	0	0	0	0	0	0	0	0
	2	2	1	2	2	0	0	0	0	0	0	0	0	0	0	0
	3	2	2	0	2	0	0	0	0	0	0	0	0	0	0	0
	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0
	1	1	2	1	0	0	0	0	0	0	0	0	0	0	0	0
	2	2	2	2	2	0	0	0	0	0	0	0	0	0	0	0
	3	2	3	3	4	2	0	0	0	0	0	0	0	0	0	0
	4	2	3	4	4	3	2	0	0	0	0	0	0	0	0	0
	5	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
	6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	7	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0
2	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0
	1	1	3	2	1	0	0	0	0	0	0	0	0	0	0	0
	2	2	3	3	3	1	0	0	0	0	0	0	0	0	0	0
	3	2	3	3	4	3	2	0	0	0	0	0	0	0	0	0
	4	2	3	4	5	4	4	2	1	0	0	0	0	0	0	0
	5	1	3	3	4	6	7	4	3	3	0	0	0	0	0	0
	6	1	3	2	2	1	4	5	0	2	0	0	0	0	0	0
	7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	1	1	2	2	1	0	0	0	0	0	0	0	0	0	0	0
	2	2	3	3	3	1	0	0	0	0	0	0	0	0	0	0
	3	2	3	4	4	3	1	0	0	0	0	0	0	0	0	0
	4	2	3	4	5	5	4	2	0	0	0	0	0	0	0	0
	5	2	4	4	5	6	6	4	3	2	0	0	0	0	0	0
	6	2	3	4	6	7	8	8	7	4	3	1	0	0	0	0
	7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	2	1	3	3	3	2	0	0	0	0	0	0	0	0	0	0
	3	2	4	4	4	3	1	0	0	0	0	0	0	0	0	0
	4	2	4	5	5	5	3	1	0	0	0	0	0	0	0	0
	5	2	4	5	6	6	6	4	1	0	0	0	0	0	0	0
	6	2	5	6	7	8	8	7	5	3	1	0	0	0	0	0
	7	2	5	6	9	10	10	10	10	8	6	4	2	1	1	0

표 1. 각 상태와 부류 별 확률 분포표(단위: %)

Table 1. Probability distribution of blocks per state and class (unit: %).

상태	부류	확률							
		0	1	2	3	4	5	6	7
0		7.007	2.856	0.256	0.049	0.000	0.000	0.000	0.000
1		1.931	6.128	2.856	0.439	0.110	0.024	0.012	0.012
2		0.085	2.661	5.774	2.954	0.916	0.269	0.061	0.000
3		0.000	0.244	2.905	5.298	2.710	1.172	0.366	0.110
4		0.000	0.049	0.940	4.224	8.948	12.125	13.013	13.672

Codebook 작성에는 LBG 알고리즘^[16]을 사용하였다. DCT-CVQ와 FS-DCT-CVQ의 codebook을 설계하기 위하여 512×512의 크기를 갖는 8장의 영상으로부터 라플라시안 피라미드 신호들을 구성하고, 각 부류별로 가능한 모든 방향의 가장자리 성분을 가진 training set를 확보하기 위하여 [7]에서 제안한 mirror set를 이용하였다.

부호화 결과는 다음과 같다. 먼저 고 주파수 성분이 매우 강한 BABOON 영상에 대해서는 0.295bpp에서 22.45dB의 PSNR 값을 보였으며, LENA 영상에 대해서는 0.192bpp에서는 30.33 dB, 0.385bpp에서는 32.45dB의 PSNR 값을 나타냈다. 그러나 피라미드 층수를 5로 하여 LENA영상을 부호화한 경우에는 0.195 bpp에서 29.89 dB, 0.392 bpp에서 32.24 dB로 4층의 경우에 비하여 약간의 성능저하가 있었다. 그림 6에 LENA 원 영상과 제안한 부호화 기법을 이용하여 점진적으로 복원된 영상을 보였다. 복원된 영상은 저 전송율에서도 시각적으로 좋은 화질을 보여 주었고, 특히 원 영상에 직접 변환 부호화를 적용했을 때 발생하는 블럭 현상이 현저하게 감소한 것을 확인할 수 있었다.

VI. 결 론

본 논문에서는 라플라시안 피라미드 신호를 부호화하기 위한 방법으로 DCT 영역에서의 분류 벡터 양자기(DCT-CVQ)와 계층간의 AC분산 상관성을 이용한 유한 상태 VQ 기법을 결합한 FS-DCT-CVQ 부호화 기법을 제안하였다. 먼저 라플라시안 피라미드 구조에서의 최적 비트 할당 및 이 구조의 최소 왜곡 측면에서의 부호화 이득을 분석하였다. DCT-CVQ 기법을 적용하여 부호화할 때의 최소 왜곡을 갖는 최적 비트 할당 식을 유도하였으며, 스칼라 양자기에 대한 DCT-CVQ의 점진적 부호화 이득을 구하였다. 그리고 이론적으로 추측하기 힘든 1.0 bpp 미만의 저 전송율에서 라플라시안 피라미드 구조가 갖는 부호화 이득을 실험적으로 구하였다. 그 결과 ADCT를 이용하여 부호화하는 경우 라플라시안 피라미드 구조는 0.5 bpp 미만의 저전송율에서 약 1-4dB의 이득을 갖는 것을 알 수 있었다.

실제 영상을 사용한 실험에 의하면 제안한 부호화 기법은 0.192bpp에서도 30.33 dB의 좋은 성능을 보였으며, 0.385 bpp에서는 복원 영상의 PSNR이 32.45 dB로 아주 우수한 성능을 나타내는 것을 알 수 있었다. 제안한 방법에 의하여 복원된 영상은 객관적 평가 척도인 PSNR 값도 높게 나왔을 뿐 아니라

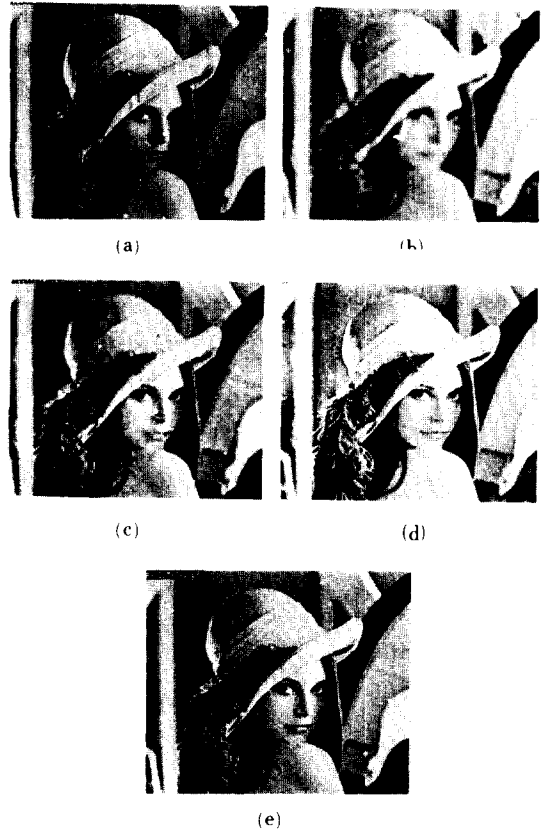


그림 6. 원 영상과 점진적으로 부호화된 영상 (outside training set)

- (a) 원 영상
- (b) 0.050bpp (PSNR = 24.80 dB)
- (c) 0.088bpp (PSNR = 26.91 dB)
- (d) 0.192bpp (PSNR = 30.33 dB)
- (e) 0.385bpp (PSNR = 32.45dB)

Fig. 6. Original image and progressively reconstructed images(outside training set).

- (a) original image.
- (b) 0.050bpp (PSNR = 24.80dB).
- (c) 0.088bpp (PSNR = 26.91dB).
- (d) 0.192bpp (PSNR = 30.33dB).
- (e) 0.385bpp (PSNR = 32.45dB).

주관적 평가 척도인 화질 감상에서도 우수한 결과를 보였다.

【부 록】

전단에서의 양자화 오차를 다음단의 라플라시안 평면을 구할 때 보상하는 경우^[7]의 비트 할당 문제는 closed form으로 표현할 수 없으나, 다음과 같은 알고리즘을 통하여 구할 수 있다.

step I : Given pyramid level L and bit rate R ,
construct all Laplacian planes, $\{L_\ell, \ell = 0, 1, \dots, L-1\}$.

step II : Calculate $\bar{\sigma}_{\ell i}^2$, $i = 1, 2, \dots, m_{\ell j}$, $j = 1, 2, \dots, n$, $\ell = 0, 1, \dots, L-1$.

step III : Let $\ell = L-1$.

step IV : Calculate $r_{\ell i}$, $i = 1, 2, \dots, m_{\ell j}$, $j = 1, 2, \dots, n$.

step V : Encode L_ℓ .

step VI : If $L=1$, go to step VII.
else
 $R \leftarrow R - 4^{-L} R_1$,
 $L \leftarrow L - 1$,
 $\ell = L - 1$,
 Reconstruct L_ℓ ,
 Calculate $\bar{\sigma}_{\ell i}^2$, $i = 1, 2, \dots, m_{\ell j}$,
 $j = 1, 2, \dots, n$,
 Go to step IV.

step VII : Stop.

參 考 文 獻

- [1] P.J. Burt and E.H. Adelson, "The Laplacian pyramid as a compact image code," *IEEE Trans. on Commun.*, vol. COM-31, no. 4, pp. 532-540, April 1983.
- [2] J.W. Woods and S.D. O'Neil, "Subband coding of images," *IEEE Trans. on Acoust., Speech and Signal Processing*, vol. ASSP-34, pp. 1278-1288, Oct. 1986.
- [3] M. Vetterli, "Multi-dimensional sub-band coding: some theory and algorithms," *Signal Processing*, vol. 6, pp. 97-112, 1984.
- [4] P.J. Burt, "Fast filter transforms for image processing," (*Computer, Graphics, Image Processing*, vol. 16, pp. 20-51, 1981).
- [5] 박섭형, 이상욱, "영상 피라미드를 위한 최적 half-band FIR 필터", 전자공학회 논문지, 제 25권 제 7호, pp. 112~121, 1988. 7.
- [6] J.L. Salinas and R.L. Baker, "Laplacian pyramid encoding: optimum rate and distortion allocations," *Proc. ICASSP-89*, pp. 1957-1960, 1989.
- [7] 박섭형, 이상욱, "계층적 피라미드 구조와 DCT영역에서의 분류벡터 양자기를 이용한 점진적 영상 전송 기법", 전자공학회 논문지, 제 26권 제 8호, pp. 100~110, 1989. 8.
- [8] N.S. Jayant and P. Noll, *Digital coding of waveforms*, Englewood Cliffs, NJ: Prentice-Hall, 1984.
- [9] Y. Yamada, S. Tazaki and R.M. Gray, "Asymptotic performance of block quantizers with difference distortion measures," *IEEE Trans. Inform. Theory*, vol. IT-26, pp. 6-14, Jan. 1980.
- [10] R.C. Reininger and J.D. Gibson, "Distributions of the two-dimensional DCT coefficients for images," *IEEE Trans. Commun.*, vol. COM-31, no. 6, pp. 835-839, June 1983.
- [11] A.P. Prudnikov, Yu. A. Brychkov, and O.I. Marichev, *Integrals and series*, vol. 1 *Elementary functions*, Gordon and Breach science publishers, 1986.
- [12] W. Chen and C.H. Smith, "Adaptive coding of monochrome and color images," *IEEE Trans. Commun.*, vol. COM-25, no. 11, pp. 1285-1292, Nov. 1977.
- [13] J. Foster, R.M. Gray and M.O. Dunham, "Finite state vector quantization for waveform coding," *IEEE Trans. Inform. Theory*, vol. IT-31, pp. 348-355, May 1985.
- [14] R. Aravind and A. Gersho, "Low-rate image coding with finite-state vector quantization," *Proc. ICASSP-86*, pp. 137-140, April 1986.
- [15] T. Kim, "New finite state vector quantizers for images," *Proc. ICASSP-88*, pp. 1180-1183, April 1988.
- [16] Y. Linde, A. Buzo, and R.M. Gray, "An algorithm for vector quantizer design," *IEEE Trans. Commun.*, vol. COM-28, pp. 84-95, Jan. 1980.

著 者 紹 介

朴 涉 亨 (正會員) 第25卷 第7號 參照
현재 서울대학교 대학원 제
어계측공학과 박사과정 재학
중. 주관심분야는 디지털 신
호처리 등임.

李 商 郁 (正會員) 第25卷 第7號 參照
현재 서울대학교 제어계측공
학과 부교수