

계층적 피라미드 구조와 DCT 영역에서의 분류 벡터 양자기를 이용한 점진적 영상 전송

(Progressive Image Transmission Using Hierarchical Pyramid Structure and Classified Vector Quantizer in DCT Domain)

朴 涉 亨*, 李 商 郁*

(Seop Hyeong Park and Sang Uk Lee)

要 約

본 논문에서는 계층적 피라미드 구조와 DCT(discrete cosine transform) 영역에서의 벡터 양자기를 이용한 비손실 점진적 영상 전송 기법을 제안하였다. 계층적으로 구성된 피라미드 신호에 변환 부호화를 적용하여 공간 영역에 존재하는 잉여(redundancy)를 줄이고, 분류 벡터 양자기를 이용하여 DCT 계수를 부호화하였다. 변환 블록의 AC 분산을 사용한 간단한 분류기를 사용하였으며, training 영상으로부터 mirror set를 구성하여 견실한 codebook을 설계할 수 있었다. 영상 전송에서는 영상 피라미드의 상층부터 하층의 순서로, 그리고 각 층의 피라미드 신호에서는 AC 분산값이 큰 부류부터 작은 부류의 순으로 전송하여 점진적 영상 전송이 가능하도록 하였다. 제안한 부호화 기법은 0.3bpp 이하의 전송율에서도 좋은 성능을 보였으며, 0.409bpp에서는 아주 우수한 성능을 나타내는 것을 알 수 있었다. 제안한 부호화 기법은 영상 데이터 감축 성능도 우수할 뿐만 아니라, 비손실 점진적 영상 전송도 가능한 기법이다.

Abstract

In this paper, we propose a lossless progressive image transmission scheme using hierarchical pyramid structure and classified vector quantizer in DCT domain. By adopting DCT to the hierarchical pyramid signals, we can reduce the spatial redundancy. Moreover, the DCT coefficients can be encoded efficiently by using classified vector quantizer in DCT domain. The classifier is simply based on the variance of a subblock. Also, the mirror set of training set of images can improve the robustness of codebooks. Progressive image transmission can be achieved through following processes: from top to bottom level of planes in a pyramid, and from high to low AC variance class in a plane. Some simulation results with real images show that the proposed coding scheme yields a good performance at below 0.3 bpp and an excellent result at 0.409 bpp. The proposed coding scheme is well suited for lossless progressive image transmission as well as image data compression.

*正會員, 서울대학교 制御計測工學科
(Dept. of Control and Instrumentation Eng., Seoul
Nat'l Univ.)
接受日字: 1989年 1月 21日

I. 서 론

점진적 영상 전송(progressive image transmission)
기법은 전화선 등과 같은 협대역 채널(low-bandwidth)

을 이용하여 사용자와의 대화식 영상 전송을 하고자 할 때 많이 사용될 수 있는 기법이다. 예를 들어서 2400bps 채널을 통하여 $256 \times 256 \times 8\text{bits}$ 의 크기를 갖는 영상을 raster scan의 형태로 전송하기 위해서는 약 3분 40초 정도의 긴 시간이 필요하다. 그러나 전송되는 영상 데이터의 순서를 raster scan과 다른 형태로 바꾸면, 데이터가 전송됨에 따라서 처음에는 영상의 개략적인 형태를 복원하고 점차적으로 화질이 향상되도록 할 수 있다. 이러한 점진적 영상 전송 기법에서는 처음에는 가능하면 적은 양의 데이터를 이용하여 영상의 개략적인 형태를 표현하며, 사용자의 요청에 의하여 추가 정보를 전송하여 점진적으로 향상된 화질을 지닌 영상을 복원할 수 있도록 한다. 그리고 최종적으로 원 영상의 완벽한 복원이 가능해야 한다.

지금까지 점진적 영상 전송을 위한 영상 부호화 기법에 대한 다각적인 연구들이 수행되었으며,¹⁻⁷ 가장 널리 사용되고 있는 기법으로는 계층적 부호화^{1,2,4} 변환 부호화⁵ 그리고 벡터 양자화⁷ 등이 있다. 1983년에 Burt와 Adelson에 의해서 제안된 라플라시안 피라미드 부호화 기법⁴은 피라미드 구조를 이용한 계층적 부호화의 일종으로 점진적 영상 전송에 적합한 구조를 지니고 있으나, 양자화 오차가 누적되기 때문에 비손실(lossless) 전송이 불가능한 단점이 있다.⁷ 이러한 단점을 극복하기 위하여 전단의 라플라시안 피라미드 신호의 양자화 오차를 다음단의 라플라시안 피라미드 신호를 구할 때 보상할 수 있는 제환(feedback) 기법을 이용할 수 있다.^{8,9}

그림 1은 비손실 라플라시안 피라미드 부호화 시스템의 블록도이다. 그림 1에서 입력 영상으로부터 구성되는 계층적인 피라미드 형태의 영상들을 가우시안 평면이라 하고, 이웃하는 가우시안 평면 사이의

실제로 전송되는 차영상을 라플라시안 평면이라 하며, 이들 평면들로 구성되는 영상 피라미드들을 각각 가우시안 피라미드, 그리고 라플라시안 피라미드라 부른다. 라플라시안 피라미드 부호화 시스템은 피라미드 구성부와 라플라시안 신호의 부호부로 나눌 수 있으며, 피라미드의 구성과정은 2차원 신호의 간축-보간(decimation-interpolation) 과정으로 설명할 수 있다.¹⁰ 따라서 라플라시안 피라미드 부호화 기법과 같이 계층적 피라미드 구조를 이용한 점진적 영상 전송 기법에서는 다음과 같은 두가지 문제를 생각할 수 있다. 첫째는 피라미드 영상을 구성하는 것으로서 이 과정에서는 간축, 보간 필터의 성능이 매우 중요하다. 두번째는 전송할 신호의 부호화로서 여기에서는 전송할 신호의 특성에 적합한 부호화 기법의 선택이 중요한 문제로 부각된다.

영상 피라미드 구성부에서는 Burt의 필터를 사용하는 것 보다는 간축, 보간 과정을 모두 고려하여 설계된 최적 half-band FIR 필터를 사용하면 좋은 성능을 얻을 수 있다.¹⁰ 다음으로 라플라시안 피라미드 신호의 부호화를 살펴보면, 공간상에서 부호화를 시도하는 기법 중에는 엔트로피 부호화를 이용한 스칼라 양자기를 이용하거나,⁴ 고감축율을 얻기 위하여 벡터 양자기를 이용한 방법 등이 있고,^{11,12} 변환 영역에서의 부호화 기법으로는 DCT (discrete cosine transform) 부호화와 스칼라 양자기를 결합한 방법이 있다.⁶ 변환 영역에서 부호화를 수행하면 공간 영역의 중복성(redundancy)을 제거할 수 있기 때문에 비교적 높은 감축율을 얻을 수 있다. 따라서 변환 부호화를 이용하면 라플라시안 피라미드 신호에 남아 있는 중복성을 더욱 줄일 수 있어서 효과적이다. 한편, 잘 알려진 바와 같이 DCT 계수를 부호화하기 위해서는 스칼라 양자기를 사용하는 것보다 벡터 양자기를 사용하는 것이 유리하다. 변환 영역에서의 벡터 양자화 기법은 변환 부호화의 장점과 벡터 양자화의 장점을 함께 이용할 수 있는 효율적인 부호화 기법으로 알려져 있다.¹³

본 논문에서는 비손실 전송이 가능한 점진적 영상 전송 기법을 제안하였다. 제안한 부호화 기법은 계층적 피라미드 구조를 사용한 것으로 이는 기본적으로 라플라시안 피라미드 부호화 시스템의 구조를 이용한 것으로 생각할 수 있다. 영상 피라미드를 구성하기 위한 필터로는 [10]에서 제안한 half-band 필터를 그대로 사용하였고, 라플라시안 피라미드 신호의 효율적인 부호화를 위해서 변환 부호화 기법중에서 가장 널리 사용되는 DCT를 사용하였으며, 변환 계

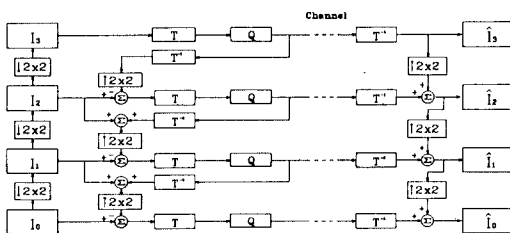


그림 1. Lossless 라플라시안 피라미드 부호화 시스템의 블록선도

Fig. 1. Block diagram of lossless Laplacian pyramid coding system.

수의 부호화에는 분류 벡터 양자기를 이용하였다. 제안한 부호화 방법은 0.4bpp 이하의 저전송율에서 33dB 이상의 영상을 복원할 수 있었으며, 여타의 블럭 부호화에서 발생하는 블럭현상(blocking effect)이 현저히 줄어든 것을 확인할 수 있었다.

II. DCT 영역에서의 분류벡터 양자기를 이용한 라플라시안 피라미드 신호의 점진적 부호화

본 논문에서는 DCT 영역에서의 부호화 기법을 고려하였다. 먼저 부호화에 앞서 라플라시안 피라미드 신호의 특성을 살펴보면 다음과 같다. 라플라시안 피라미드 신호에는 원 영상의 가장자리(edge) 성분이 주로 나타난다. 그 이유는 라플라시안 피라미드를 원 영상의 대역통과(bandpassed) 성분들의 집합으로 생각할 수 있기 때문이다. 그리고 라플라시안 피라미드 신호는 일반적인 영상 신호와 달리 공간 평면에서의 자기 상관 계수가 0.3~0.7 정도로 원 영상의 0.85~0.95와 비교해 볼 때 공간에서의 잉여(spatial redundancy)가 상당히 감소했으나 공간에서의 잉여가 여전히 존재하는 것을 알 수 있다. 따라서 변환 부호화를 이용하면 공간 영역에 존재하는 잉여를 더욱 줄일 수 있고, 다양한 형태의 가장자리들이 존재하는 공간 영역에서 부호화를 수행하는 것에 비해서 효율적인 부호화를 수행할 수 있다.

변환 부호화 기법에는 여러가지의 직교 변환(orthogonal transformation)이 사용될 수 있으나, 고속 알고리즘으로 구현할 수 있는 변환 기법 중의 하나인 DCT가 현재 가장 널리 사용되고 있다. 변환 계수의 부호화에는 스칼라 양자기(SQ: scalar quantizer), 벡터 양자기(VQ: vector quantizer) 등이 사용될 수 있으나 벡터 양자기가 갖는 여러가지 장점 때문에 최근에는 변환 영역에서의 벡터 양자화 기법, 특히 DCT와 분류 벡터 양자기(classified VQ: CVQ)를 결합한 DCT-CVQ 영상 부호화에 관한 연구가 최근에 활발히 이루어지고 있다.^{14,15,16,17}

한편, 벡터 양자기의 문제점으로 가장자리 손상(edge degradation), 전송율과 벡터 차수가 증가함에 따라서 지수적으로 증가하는 계산량, 그리고 codebook 설계시 사용된 training set의 특성과 부호화될 영상의 특성이 다를 때 발생하는 불일치(mismatch) 현상 등이 있다. 이러한 벡터 양자기의 문제점을 해결하기 위한 시도 중의 하나가 CVQ(classified VQ)이다.¹⁸ 즉, 다양한 특성을 가진 입력 신호를 특정한 기준에 의해서 분류하고, 각 부류별로 별도의 code-book을 구성하면 codebook에 적응성을 부여하는 결

과가 되어 어느 정도 현실한 codebook을 구성할 수 있게되며 아울러 벡터 양자기의 복잡성을 줄일 수 있다. 그러나 DCT-CVQ에서 부블럭의 크기가 증가할수록(일반적으로 8×8 또는 16×16) 변환된 블럭 전체를 벡터 양자화하는 것은 계산상의 부담때문에 불가능하다. 따라서 변환된 블럭을 적당한 차수의 벡터들로 분할(partition)하는 product code 구조를 사용한 벡터 양자기를 각 부류마다 사용하여 계산량을 줄일 수 있다.^{14,16}

DCT-CVQ 영상 부호화 시스템의 부호화 부의 블럭선도를 그림 2에 도시하였다. 양자화 과정을 다음과 같이 설명할 수 있다. 먼저 영상을 N×N 크기의 블럭으로 나눈 2차원의 벡터 소오스를 $\{x(n) | n=0, 1, 2, \dots\}$ 라고 하면, n번째 영상 블럭 x(n)에 N×N 2차원 DCT를 취하여 DCT 영역의 2차원 벡터 X(n)으로 변환한다. 다음에 분류화 알고리즘을 적용하여 부블럭에 부류를 할당하고, 각 부류에 따른 DCT 블럭의 분할 형태에 따라서 m_i개의 벡터 $\{X_{i1}(n) | i=1, 2, \dots, m_i\}$ 들로 분할한다. 분할된 벡터들은 각 부류에 대한 sub-codebook 집합에서 해당 벡터의 codebook을 사용하여 양자화하고 해당 codebook의 인덱스(index)를 전송한다.

따라서 DCT-CVQ를 구현하기 위해서는 분류기 설계(classifier design), 분할 방식(partition scheme), 비트 할당(bit allocation) 등과 같은 사항들을 고려해야 한다.

1. 분류기 설계

부류를 결정하는 분류기로는 가장자리 분류기(edge classifier)¹⁸와 AC 분산(variance)분류기¹⁹ 등이 대표적으로 사용된다. 가장자리 분류기(edge classifier)에 근거한 CVQ는 인간 시각의 정신 심리학

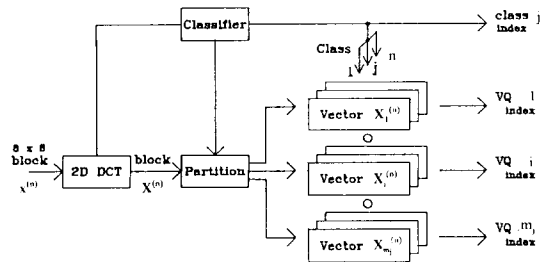


그림 2. DCT-CVQ의 부호화부의 블럭선도
Fig. 2. Block diagram of the encoding part of DCT-CVQ.

(psychophysics)에 바탕을 둔 효과적인 기법으로 알려져 있다. 분류화된 각 부류들에 대해서 각 부류의 발생 확률과 평균 왜곡 예상치에 따른 최적의 codebook을 구성할 수 있어서 가장자리 손상 방지와 계산량 감소를 얻을 수 있다. 가장자리 방향에 따른 분류기를 사용할 경우에는 가장자리의 특성이 단순할수록 효과적이다. [18]에서는 공간 영역에서 부블럭의 크기가 4×4 와 5×5 인 벡터에서 직선 형태의 가장자리만을 대상으로 각각 31 부류와 43부류로 분류하여 어느 정도의 효과를 보았으나, 부블럭의 크기가 이보다 커질 경우에는 좋은 결과를 얻을 수 없었다. 더우기 8×8 크기의 부블럭으로 나누어 블럭 부호화를 수행할 경우를 고려해볼 때, 발생할 수 있는 가장자리의 종류가 너무 다양하므로 공간 영역에서의 가장자리 분류기를 적용하면 좋은 결과를 기대할 수 없다.

그리고 가장자리 특성에 의한 분류기를 효과적으로 설계하기 위해서는 부류를 세분하는 것이 유리하다. 반면에 부류가 세분화될수록 분류기에 의한 효과는 증대되지만, 실제적으로 codebook을 설계할 때 필요한 다양한 가장자리 특성을 지닌 training set의 확보가 어려울 뿐만 아니라, 부호화 과정에서 부류 인덱스를 전송할 때 사용되는 부가 정보량(side information)이 증가하기 때문에 이에 대한 조정이 필요하다. 따라서 라플라시안 피라미드 신호에 있어서는 가장자리 분류기보다 AC 분산 분류기 등과 같은 단순한 분류 알고리즘을 적용하는 것이 유리하다고 판단하였으며, 이같은 사실은 실험적으로도 알 수 있었다. AC 분산에 의한 분류기법은 각 부블럭이 차지하는 AC 분산의 크기를 기준으로 부류를 결정하는 단순한 방법을 사용한다. 여기에서 AC 분산이 큰 블럭은 high activity 블럭을 의미한다. 라플라시안 피라미드 신호에서 high activity block은 대부분 원 영상의 가장자리가 포함된 블럭들이다. 그 이유는 라플라시안 피라미드 신호들은 준대역통과(quasi-bandpassed) 신호이기 때문에 가장자리가 없는 저주파 성분이 강한 블럭은 주로 상층의 라플라시안 신호만으로 표현이 가능하고, 가장자리 성분이 강한, 즉 고주파 성분이 강한 부분은 여러 층의 라플라시안 피라미드 평면에 걸쳐서 그 성분들이 나타나기 때문이다. 실제로 최하층의 라플라시안 피라미드 신호를 살펴보면 주로 원영상의 가장자리에 해당하는 화소들만이 큰 값을 가지는 것을 볼 수 있다. 따라서 AC 분산의 크기에 의해서 부류를 결정하면 가장자리 성분이 강한 블럭부터 차례로 순서를 정하는 것과 같은 효과를 볼 수 있다.

본 논문에서는 부류의 수를 실험적으로 8로 정하였으며 블럭의 AC 분산의 분포로부터 각 부류의 중심값(centroid)을 구하여 이를 각 부류의 대표값으로 사용하였다. 이 방법은 [19]에서 4 부류를 사용한 것의 확장으로 생각할 수 있다. [19]에서는 입력 영상마다 적응적으로 4개의 부류를 할당하였으나, 벡터 양자화에서는 정해진 부류별로 codebook을 별도로 구성해야 하므로 고정된 분류기준에 의해서 부류를 할당하는 것이 현실적으로 유리하다.

2. 벡터의 분할 방식(partition scheme)

벡터의 분할방식은 각 부류마다 동일한 방식을 채택하였다. 그림 3에는 본 논문에서 사용한 분할 방식을 도시하였다. 제시한 분할방식은 zig-zag 방식^[15]으로 DCT 영역에서 구역(zone)을 형성하고, 구역밖의 적은 에너지를 갖는 높은 sequency 계수들을 생략(truncate)하는 구역 부호화(zonal coding) 개념에서 유래한 것이다. 이 방식은 같은 sequency대의 계수들에 남아있는 상관성을 이용할 수 있으며, 큰 분산치일수록 낮은 차수로 벡터가 구성되므로 전송을 증가에 의한 왜곡치 감소에 유리하다. 또한 전송을 증가와 벡터 차수 증가에 지수적으로 비례하는 벡터 양자화의 계산상의 복잡성을 줄일 수 있는 장점이 있다.

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 블럭에서의 벡터의 분할 형태
Fig. 3. Partition scheme of the DCT block.

3. 비트 할당

비트 할당에서는 다음의 두가지를 고려하여야 한다. 즉, 전체적인 전송율이 주어지면 먼저 각 층(plane)의 라플라시안 피라미드 신호에 할당해야 할 전송율, 즉 층간(inter-plane) 전송율을 정해야 한다. 각 층에 할당할 전송율이 결정되면 층내(intra-plane)의 각 부류별로 전송율을 할당해야 한다. 실험에

의하면 [4]와 같은 시스템에서는 전체적인 결과가 각 층 간의 비트 할당에 따라 민감한 차이를 나타내지만, 양자화 오차를 보상할 수 있는 그림 1과 같은 시스템에서는 층 간의 비트 할당의 영향이 거의 없는 안정된 결과를 얻을 수 있다. 따라서 본 논문에서 사용한 것과 같이 양자화 오차를 보상할 수 있는 시스템에서는 층 내의 비트 할당만 고려하면 된다. 그리고 각 층에 배정하는 전송율의 영향이 매우 적기 때문에 점진적 영상 전송을 위한 전송율 배정에 있어서도 많은 유연성을 가질 수 있는 장점이 있다.

각 층에 대한 비트가 할당되면 부류에 대한 전송율의 할당문제와 각 부류 내의 벡터들에 대한 전송율 조정 문제가 발생한다. DCT 블럭에서 발생하는 왜곡치를 최소화 시키는 전송율 즉 비트 할당을 위해서 사용되는 것이 벡터 소오스에 대한 asymptotic DRF이다. 이러한 예측식을 training set로부터 코드북을 설계하는 LBG 알고리즘^[20]의 특성에 결부시킴으로써 다음과 같은 최소화 문제 형태로 비트 할당 문제를 표현할 수 있다. 먼저 입력 DCT 블럭의 부류를 j ($j=1,2,\dots,n$)라 하고, 분할된 벡터들을 v_i , ($i=1,2,\dots,m_j$)라 하자. 그러면 비트 할당 문제는 다음과 같은 최소화 문제 형태로 귀결된다.^[21]

$$\begin{aligned} \text{Minimize } D &= \sum_{j=1}^n P_j \left(\sum_{i=1}^{m_j} D_{ji} \right) \\ \text{subject to } B &= \sum_{j=1}^n P_j \left(\sum_{i=1}^{m_j} b_{ji} \right) \end{aligned} \quad (1)$$

여기에서 b_{ji} 는 양의 정수이어야 하며, P_j 는 부류 j 의 발생 확률이고, D_{ji} 는 부류 j 의 m_j 개 벡터들 중에서 i 번째 벡터에서 발생될 평균 왜곡치를, b_{ji} 는 부류 j 의 i 번째 벡터에 할당되는 비트를 각각 의미한다. 그리고 D 와 B 는 각각 평균 왜곡치와 부호기 전체의 평균 전송율에 해당된다. LBG 알고리즘에서는 codebook을 설계하기 위하여 training 영상을 사용하기 때문에 $\{P_j\}$ 와 $\{D_{ji}\}$ 는 training 영상들로부터 쉽게 구할 수 있다. 그리고 (1)식에 주어진 최소화 문제는 Lagrangian multiplier를 이용하는 non-linear programming 기법을 이용하여 풀 수 있다.^[21]

4. codebook 설계

라플라시안 피라미드 부호화 기법은 원 영상을 부호화하는 대신에 원 영상으로부터 피라미드 구조의 라플라시안 피라미드 신호를 구성하고 이들을 부호화하는 방식이다. 따라서 라플라시안 피라미드의 각 층의 통계적 특성이 다르다면, 각 층마다 별도로 codebook을 구성하여야 한다. 그러나 실제로 부호화

시스템을 구성할 때 각 층의 라플라시안 피라미드 신호에 대한 codebook을 별도로 구성하기 위해서는 많은 양의 계산과 메모리를 필요로 하기 때문에, 모든 층에 함께 사용할 수 있는 codebook을 구성하는 것이 바람직하다.

최상층의 신호를 제외한 각 층의 라플라시안 피라미드 신호들은 원 영상에 비해 공간 평면에서의 상관도가 줄어든 비슷한 특성을 보이고, DCT 영역에서는 원 영상에 비해서 저 sequency 계수들의 분산이 상대적으로 감소된 특성을 공통적으로 나타낸다. 그 이유는 전단에서 부호화된 신호들은 주로 저 sequency 계수들이고, 양자화 오차가 바로 다음단의 라플라시안 신호의 저 sequency 계수 부분에 나타나기 때문이다. 그리고 이러한 양자화 오차들은 앞단의 라플라시안 신호의 양자화에 배정된 전송율에 의해서 분산의 크기가 결정되기 때문에 전송율을 조정하여 모든 층의 라플라시안 신호들의 DCT 영역에서의 특성을 어느정도 유사하게 조절할 수 있다. 이런 방법을 사용하여 본 논문에서는 codebook set을 하나만 구성하여 모든 층의 라플라시안 신호를 부호화 하였다.

Codebook 작성에는 LBG 알고리즘^[20]을 사용하였다. 잘 알려진대로 벡터 양자화에서는 주어진 training set로부터 codebook을 구성하기 때문에 다양한 특성을 가진 training set의 확보가 중요하다. 본 논문에서 사용한 바와 같이 단순히 AC 분산의 크기로 부류를 구분하는 경우에는 가장자리의 방향은 전혀 고려할 수 없기 때문에 각 부류별로 가능한 모든 방향의 가장자리 성분을 가진 training set를 확보해야 한다. 그리고 공간 영역에서의 가장자리의 방향은 DCT 영역에서의 계수의 분포와 밀접한 관계가 있으므로^[22] 이러한 training set를 확보한다는 것은 DCT 영역의 벡터를 이용하여 codebook을 작성하기 위해서도 중요한 것이다. 그러나 이런 training set를 확보한다는 것은 현실적으로 대단히 어렵다. 따라서 다음과 같이 인위적으로 codebook 설계에 필요한 training set를 얻었다. 그림 4에서 블럭(a)가 주어지면 (b)에서 (d)까지의 블럭들은 블럭 (a)를 각각 90°, 180°, 270° 회전시킴으로써 얻을 수 있다. 블럭 (e)는 블럭 (a)를 x축에 대하여 대칭이동시켜 구할 수 있으며, (f)에서 (h)까지의 블럭들은 블럭 (e)를 각각 90°, 180°, 270° 회전시킴으로써 얻을 수 있다. 이들을 통칭하여 mirror set라 하겠다. Training set로 사용되는 입력 영상으로부터 mirror set를 구하고, 이를 이용하여 codebook을 구성하면 공간 영역에서는 다양한 방향의 가장자리 성분을 균등하게 확보할

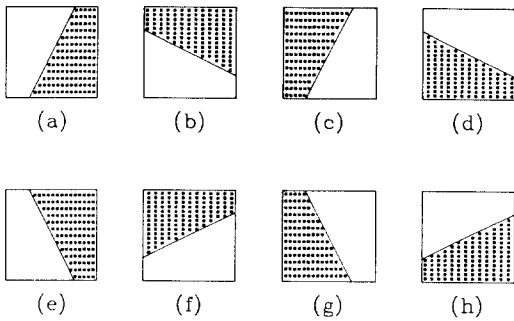


그림 4. mirror set에서의 가장자리 방향
Fig. 4. Edge directions in a mirror set.

수 있으므로 결과적으로 강인한(robust) codebook을 구성할 수 있는 장점이 있다.

5. 비손실 점진적 영상 전송

그림 5에는 본 논문에서 제안한 부호화 기법의 블록도를 도시하였다. 그림 5에서 T는 이산 여현 변환을 의미하고 T⁻¹는 이산 여현 변환의 역변환을 의미한다. 그리고 SQ와 CVQ는 각각 스칼라 양자기와 분류 벡터 양자기를 의미한다. 점진적 영상 전송은 다음과 같은 과정을 통하여 구현될 수 있다. 라플라시안 피라미드 신호는 위층부터 아래층의 순

서(top-bottom)로 전송한다. 그리고 각층의 라플라시안 신호는 AC 분산 분류기에 의해서 결정된 부류별로 전송하며, 전송 순서는 AC 분산이 큰 부류부터 작은 부류 순으로 차례로 전송한다. 최하층을 제외한 모든 층의 양자화 오차는 다음 단의 라플라시안 피라미드 신호를 구하는 과정에서 보상되지만, 최하층의 신호를 부호화한 후에는 최종적인 잔여(residual) 오차가 존재하며, 비손실 영상 전송을 구현하기 위해서는 이 잔여 오차를 전송해야 한다.

비손실 부호화 기법에서 전체의 전송률 R_t는 다음과 같이 정의된다.

$$R_t = R_p + R_e \tag{2}$$

여기에서 R_p는 progressive bit rate를 의미하며, R_e는 식(3)과 같이 표시되는 잔여 오차 신호의 1차 엔트로피를 의미한다.

$$R_e = -\sum_i P_e(i) \log_2 P_e(i) \tag{3}$$

(3)식에서 i는 잔여 오차 신호의 회도값을, 그리고 P_e(i)는 i의 발생 확률을 의미하며, P_e(i)는 잔여 오차 신호의 히스토그램으로 부터 직접 구할 수 있다.

점진적 영상 전송 기법에서는 누적 정보 전송율이 증가할수록 R_e는 감소하여 0으로 수렴하지만, R_p가 증가하기 때문에 R_t를 최소화하기 위한 잔여 오차 신호의 부호화도 중요하다. 이 오차 신호에 대해서는 제안한 CVQ를 계속적으로 적용한 다단

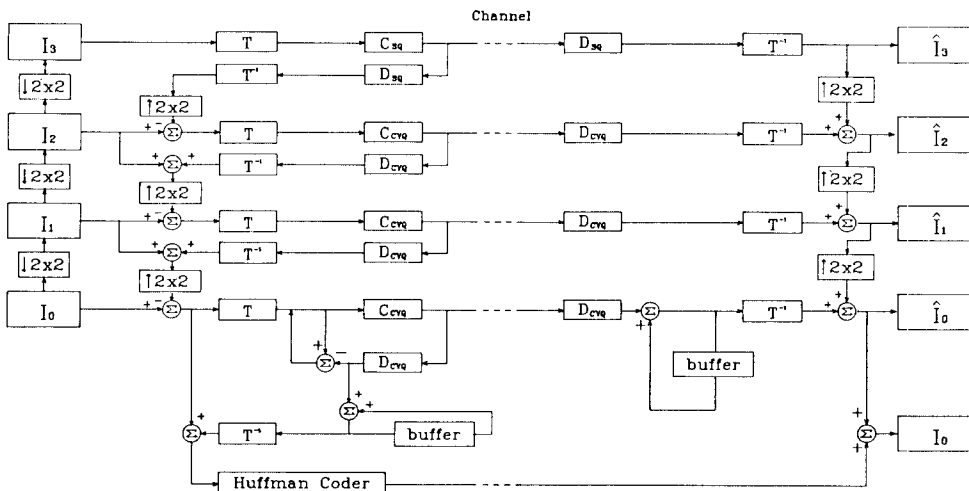


그림 5. 제안한 LPC-DCT-CVQ의 블록선도
Fig. 5. Block diagram of proposed LPC-DCT-CVQ.

(multi-stage) 부호화 등을 수행하여 비손실 영상 전송을 구현할 수도 있으나,⁷⁾ 실제적으로는 엔트로피 부호기를 적용하는 것이 바람직하다.

III. 실험 결과 및 검토

본 논문에서 제안한 부호화 기법에 대한 컴퓨터 시뮬레이션을 수행하였다. 부호화 시뮬레이션에서는 라플라시안 피라미드 신호를 두가지 방법으로 부호화하고 이를 비교하였다. 첫번째 방법은 제안한 부호화 기법으로 이를 LPC-DCT-CVQ(Laplacian pyramid coding using DCT-CVQ)라 하겠다. 두번째 방법은 [9]에서와 같이 각 층의 라플라시안 피라미드 신호의 부호화에 Chen과 Smith의 적응 DCT 부호화 기법¹⁰⁾을 적용한 것으로 이를 LPC-ADCT(Laplacian pyramid coding using adaptive DCT)라 하겠다.

실험 결과는 객관적 비교 척도인 (4)식의 PSNR과 주관적인 화질 평가를 병행하였다.

$$PSNR=10\log \{255^2/M^{-1} \sum_{i=0}^{M-1} (x_i-y_i)^2\} \text{ dB} \quad (4)$$

(4)에서 M은 부호화하는 영상의 화소 수이고, x_i 는 원 영상의 화소의 휘도를, 그리고 y_i 는 복원된 영상의 화소의 휘도를 각각 의미한다.

DCT-CVQ를 구성하기 전에 다음의 두가지 문제를 고려하였다. 첫째는 입력 영상으로부터 라플라시안 피라미드를 구성할 때 층의 갯수를 결정하는 문제이다. 주어진 영상의 휘도 분포가 통계적으로 독립이라고 가정하면, 입력영상을 정확히 부호화하기 위한 최소한의 전송율은 휘도 분포의 1차 엔트로피에 의해서 결정된다.¹²⁾ 이러한 가정하에서 비손실 전송을 고려할 때 식(5)가 성립하는 범위에서 피라미드의 층수를 계속 증가시키는 것이 바람직하다.

$$H(G_i) \geq 1/4 H(G_{i+1}) + H(L_i), \quad i=0, 1, \dots, N-1, \quad (5)$$

여기에서 $H(\cdot)$ 는 각 평면의 1차 엔트로피를 의미한다. 식(5)가 성립하면 G_i 대신 G_{i+1} 과 L_i 를 부호화함으로써 부호화에 필요한 최소한의 전송율을 줄일 수 있다. 표 1에는 각 층의 가우시안 평면과 라플라시안 평면의 1차 엔트로피를 비교하였다. 표 1을 살펴보면 층수가 증가함에 따라서 $H(G_i)$ 와 $1/4 H(G_{i+1}) + H(L_i)$ 의 차이가 감소하다가, 3층에서는 $H(G_3)$ 가 $1/4 H(G_4) + H(L_3)$ 보다 큰 것을 볼 수 있다. 이것은 더이상 층을 분리하면 전송해야 할 신호의 엔트로피, 즉 필요한 최소한의 전송율

표 1. 층에 따른 가우시안 피라미드와 라플라시안 피라미드 신호의 엔트로피 변화

Table 1. Dependence of first-order entropy of Gaussian and Laplacian pyramids on the level of planes.

$H(G_0)$	(256×256)	7.440438
$H(G_1)$	(128×128)	7.431952
$H(G_2)$	(64×64)	7.387117
$H(G_3)$	(32×32)	7.263348
$H(L_0)$	(512×512)	4.011750
$H(L_1)$	(256×256)	4.292453
$H(L_2)$	(128×128)	5.008270
$H(L_3)$	(64×64)	5.624935
$1/4 H(G_1) + H(L_0)$		5.871860
$1/4 H(G_2) + H(L_1)$		6.150441
$1/4 H(G_3) + H(L_2)$		6.855049
$1/4 H(G_4) + H(L_3)$		7.464077

image : LENA (512×512×8 bits)
Entropy=7.445072(bpp)

이 증가한다는 것을 의미한다. 따라서 본 논문에서는 이에 근거하여 하나의 피라미드를 4층으로 구성하였다. 이때 0층, 1층, 2층, 3층의 영상의 크기는 각각 512×512, 256×256, 128×128, 64×64이다

둘째는 codebook을 설계하는 문제이다. 본 논문에서는 최상층의 라플라시안 평면(L_3)만을 DCT와 스칼라 양자기를 이용하여 부호화 하였고, 그 이외의 평면들은 DCT-CVQ를 이용하여 부호화하였다. 그 이유는 윗층으로 올라갈수록 더 많은 bit가 할당되므로,¹¹⁾ L_3 에 VQ를 사용하면 복잡도가 증가하게 되어 쉽게 구현할 수 없기 때문이다. 그리고 L_3 는 원 영상의 1/64로 크기가 감소하므로 L_3 에 할당되는 전송율은 전체 전송율에 커다란 영향을 미치지 못하기 때문에 스칼라 양자기를 사용해도 무방하다.

II.4 절에서 설명한 것과 같이 본 논문에서는 하나의 codebook set를 구성하여 최상층을 제외한 모든 층의 라플라시안 피라미드 신호를 부호화 하였다, training set로 사용한 영상들로 부터 라플라시안 피라미드를 구성하고, 각 층별로 별도의 training set를 구성하였다. 이들로 부터 각각 설계된 codebook을 사용하여 부호화 시뮬레이션을 수행한 결과, 1층의 라플라시안 신호들만 사용하여 설계된 codebook을 사용한 경우에 가장 우수한 결과를 얻을 수 있었다. codebook을 설계하기 위해서 사용한 training 영상들은 512×512의 크기를 갖는 영상들이며, 모두 8장을 사용하였는데 부호화 실험에 사용한 "LENA"는 training 영상에 포함되지 않았다. 표 2에는 LBG

알고리즘에 의해서 구한 각 부류의 AC 분산값의 중심값을 보였다. 표 3에는 codebook 설계에 사용된 각 부류별, 그리고 각 벡터에 대한 비트 할당표를 보였다. 표 3에서 첫번째 벡터는 차수가 1인, DCT 계수 중에서 DC 성분을 나타내며 각 숫자는 그림 3의 분할 형태에서 표시한 숫자를 나타낸다. 그리고 부류 1과 8은 각각 분산이 가장 작은 부류와 가장 큰 부류를 의미한다. 비트가 할당된 형태를 살펴보면 DC에 1, 2bit가 할당되었고 DC 부근의 저sequency 계수들에 비교적 균등하게 할당된 것을 알 수 있는데, 이것은 라플라시안 피라미드 신호의 DCT 영역에서의 특성이 원 영상의 특성과는 많이 다르다는 것을 의미한다.

그림 6에서는 실험에 사용한 원 영상과 점진적으로 복원된 영상을 보였다. 그림 6에서, (a)는 실험에 사용된 512×512 크기의 "LENA" 원 영상이고, (b)는 L₃만을 부호화하여 복원된 영상이고, (c)와 (d)는 (a)에 L₂와 L₁이 각각 추가되어 보원된 영상이다.



(a) (b)



(c) (d)



(e)

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

- (a) 원영상
- (b) 0.031bpp (PSNR = 24.24dB)
- (c) 0.081bpp (PSNR = 26.31dB)
- (d) 0.170bpp (PSNR = 30.13dB)
- (e) 0.409bpp (PSNR = 33.78dB)

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

- (a) original image.
- (b) 0.031bpp (PSNR = 24.24dB).
- (c) 0.081bpp (PSNR = 26.31dB).
- (d) 0.170bpp (PSNR = 30.13dB).
- (e) 0.409bpp (PSNR = 33.78dB).

즉, (d)는 최하층의 라플라시안 피라미드 신호가 제외된 중간형태의 영상으로 0.17bpp의 저전송율에서도 비교적 뚜렷한 천체적인 모습을 볼 수 있다. (e)는 최하층의 라플라시안 피라미드 신호가 전송되어

표 2. 각 부류의 분산의 대표값

Table 2. Representative variance of each class.

class	variance for centroid
1	79.50
2	260.51
3	475.54
4	766.82
5	1165.41
6	1695.41
7	2486.97
8	3946.54

표 3. codebook 설계에 사용된 비트 할당표

Table 3. Bit allocation table for codebook design.

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

화질이 향상된 영상이다. (e)에서 볼 수 있듯이, 0.409bpp에서는 원 영상에 매우 가까운 화질을 지닌 영상이 복원되는 것을 알 수 있다.

표 4에서는 LPC-DCT-CVQ 기법에 의해 단계적으로 복원된 영상과 동일한 전송율 하에서 LPC-ADCT 기법에 의해 복원된 영상의 PSNR 값을 비교하였다. 표 4에서 알 수 있듯이 LPC-ADCT 기법도 비교적 좋은 결과를 나타낸다. 이것은 라플라시안 피라미드 신호의 부호화에는 DCT 기법이 유리하다는 것을 의미한다. 그리고 제안한 LPC-DCT-CVQ 기법은 LPC-ADCT에 비해서 약 0.8-1.3dB 이상 우수한 성능을 나타내는 것을 알 수 있다.

표 5에는 단계적 누적 전송율에 따른 전체 정보 전송율의 변화와 복원된 영상의 PSNR 값을 제시하였다. 점진적으로 정보 전송율이 증가함에 따라서 전체 정보 전송율은 급격하게 감소하는 추세를 보이다가 11번째 단계에서 최소값(4.793bpp)을 나타내고 이후에는 다시 증가하는 것을 볼 수 있다. 12번째 단계 이후는 잔여 오차 신호를 전송하는 부분으로 분산 값이 큰 6,7,8 부류가 나타나지 않는 것을 볼 수 있다. 이것은 잔여 오차 신호와 그 이전의 라플라시안 영상 신호들과의 분산 분포가 다르다는 것을 의미한다. 따라서 잔여 오차 신호의 특성에 맞추어 codebook을 구성하는 것이 바람직하지만, 이것은 이미 만들어진 codebook 중에서 분산이 작은 부류를 세분화하는 것과 같은 의미이므로 기존의 codebook을 이용하는 것과 커다란 차이가 없다. 따라서 이 이후로는 DCT-CVQ를 적용하는 것보다는 엔트로피 부호기를 사용하는 것이 유리하다. 그리고 11번째 단계의 전체 정보 전송율의 최소값은 실험에 사용된 원 영상의 1차 엔트로피인 7.445에 비하여 훨씬 적은 것을 알 수 있다. 이것은 정보 압축과 함께 비손실 영상 전송이 가능하다는 것을 의

표 4. LPC-DCT-CVQ와 LPC-ADCT의 결과 비교

Table 4. Comparison of LPC-DCT-CVQ and LPC-ADCT.

bit rates	LPC-DCT-CVQ PSNR (dB)	LPC-ADCT PSNR (dB)
0.170	30.13	29.42
0.268	31.21	30.82
0.333	32.52	31.46
0.386	33.45	32.20
0.409	33.78	32.46

표 5. 전체 정보 전송율과 복원 영상의 PSNR
Table 5. Total bit rates and PSNR of reconstructed images.

transmission			B _p (bpp)	B _e (bpp)	B _t (bpp)	PSNR (dB)
stage	plane	class				
1	L ₃		0.031	5.803	5.834	24.24
2	L ₂	all	0.081	5.429	5.510	26.31
3	L ₁	all	0.170	4.888	5.058	30.13
4	L ₀	8	0.235	4.853	5.088	30.52
5		7	0.268	4.777	5.045	31.21
6		6	0.303	4.689	4.992	31.94
7		5	0.333	4.611	4.944	32.52
8		4	0.363	4.537	4.900	33.00
9		3	0.381	4.457	4.843	33.45
10		2	0.401	4.403	4.804	33.71
11		1	0.409	4.384	4.793	33.78
12	L ₀	5	0.460	4.380	4.840	33.82
13		4	0.483	3.362	4.845	33.97
14		3	0.511	4.337	4.848	34.12
15		2	0.531	4.311	4.842	34.24
16		1	0.543	4.299	4.843	34.29

미한다.

제안한 부호화 방법은 0.3bpp 정도에서도 우수한 성능을 나타내며, 0.409bpp에서는 대단히 좋은 성능을 나타내는 것을 알 수 있다. 또한 복원된 영상은 저 전송율에서도 시각적으로 좋은 화질을 보여주었고, 0.409bpp에서는 만족할만한 화질을 나타냈다. 특히 원 영상에 직접 변환 부호화를 적용했을 때 발생하는 블럭 현상이 현저하게 감소한 것을 확인할 수 있었다.

IV. 결 론

본 논문에서는 계층적 피라미드 구조와 DCT 영역에서의 벡터 양자기를 이용한 점진적 영상 전송 기법을 제안하였다. 부호화 부에서는 DCT 영역에서의 분류 벡터 양자기를 사용하였다. 분류기로는 블럭의 AC 분산의 크기에 의한 간단한 분류기를 사용하였으며, codebook의 설계 과정에서는 입력 영상의 mirror set를 사용하여 codebook의 견실성을 향상시켰다. 그리고 영상 전송은 영상 피라미드의 상층부터 하층의 순서로, 그리고 각 층의 피라미드 신호에서는 AC 분산이 큰 부류부터 작은 부류의 순서로 전송하여 점진적 전송이 가능하도록 하였다.

실제 영상을 사용한 실험에 의하면 제안한 부호화 기법은 0.3bpp이하에서도 좋은 성능을 보였으며,

0.409bpp에서는 복원 영상의 PSNR이 33.78dB로 아주 우수한 성능을 나타내는 것을 알 수 있었다. 제안한 방법에 의하여 복원된 영상은 객관적 평가 척도인 PSNR 값도 높게 나왔을 뿐 아니라 주관적 평가 척도인 화질 감상에서도 우수한 결과를 보였다. 특히 저전송율에서도 여타의 블럭 현상이 현저히 줄어든 것을 확인할 수 있었다. 그리고 전송율이 점진적으로 증가할수록 전체 전송율이 급격하게 감소하여 원 영상의 1차 엔트로피보다 작은 값을 나타내었다. 따라서 제안한 부호화 기법은 영상 데이터 감축 성능도 우수하고, 비손실 점진적 영상 전송도 가능한 기법임을 알 수 있다.

參 考 文 獻

- [1] S.L. Tanimoto, "Image transmission with gross information first," *Computer, Graphics, Image Processing*, vol. 9, pp. 72-76, Jan. 1979.
- [2] K. Sloan and S.L. Tanimoto, "Progressive refinement of raster images," *IEEE Trans. on Computers*, vol. C-28, pp. 871-874, Nov. 1979.
- [3] K. Knowlton, "Progressive transmission of grey scale and binary pictures by simple, efficient, and lossless encoding schemes," *Proc. IEEE*, vol. 68, pp. 885-896, July 1980.
- [4] P.J. Burt and E.H. Adelson, "The Laplacian pyramid as a compact image code," *IEEE Trans. on Communications*, vol. COM-31, no. 4, pp. 532-540, Apr. 1983.
- [5] K.N. Ngan, "Image display techniques using the cosine transform," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-32, pp. 173-177, Feb. 1984.
- [6] K.H. Tzou, "Progressive image transmission: a review and comparison of techniques," *Optical engineering*, vol. 26, no. 7, pp. 581-589, July 1987.
- [7] L. Wang, M. Goldberg, "Progressive image transmission by transform coefficient residual error quantization," *IEEE Trans. Commun.*, vol. COM-36, pp. 75-87, no. 1, Jan. 1988.
- [8] ISO TC 97/SC2/WG8 Document N640, Jan. 1988.
- [9] 박섭형, 이상욱, "변환 부호화 기법을 이용한 라플라시안 피라미드 부호화," 한국통신학회 영상통신 연구 발표회 논문집, pp.44-48, 1988. 10.
- [10] 박섭형, 이상욱, "영상 피라미드를 위한 최적 half-band FIR 필터," 전자공학회 논문지, 제 25권 제 7 호, pp.112-121. 1988. 7.
- [11] C.H. Chen, "Laplacian pyramid image data compression," *IEEE Proc. ICASSP-87* vol. 2, pp. 737-739, Apr. 1987.
- [12] A. Tran, et a., "An efficient pyramid image coding system," *IEEE Proc. ICASSP-87* vol. 2, pp. 744-747, Apr. 1987.
- [13] N.M. Nasrabadi and R.A. King, "Image coding using vector quantization: A review," *IEEE Trans. Commun.*, vol. COM-36, pp. 957-971, Aug. 1988.
- [14] Y. Du, "On optimization of band partition for vector quantization of large blocks," *Signal Processing Conf. EUSIPCO-88*, pp. 1517-1520, 1988.
- [15] M. Breeuwer, "Transform coding of images using directionally adaptive vector quantization," *IEEE ICASSP-88*, pp. 788-791, 1988.
- [16] C. Fong and M. Baraniecki, "Adaptive transform coding with multi-band vector quantization for video signals," *IEEE ICASSP-88*, pp. 1184-1187, 1988.
- [17] S.C. Kwatra, "Vector quantization in image coding," *Proceedings of international workshop on image coding*, Seoul, Korea, Aug. 1987.
- [18] B. Ramamurthi and A. Gersho, "Classified vector quantization of images," *IEEE Trans. Commun.*, vol. COM-34, pp. 1105-1115, Nov. 1986.
- [19] 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.
- [20] 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.
- [21] A. Segall, "Bit allocation and encoding for vector sources," *IEEE Trans. Inform. Theory*, vol. IT-22, no. 2, Mar. 1976.
- [22] D.S. Kim and S.U. Lee, "Image vector quantizer based on a classification in the DCT domain," *TENCON-87*, pp. 413-417, Aug. 1987.
- [23] R.G. Gallager, *Information theory and reliable communication*, John Wiley & Sons, Inc., 1968. *

著 者 紹 介

朴 涉 亨 (正會員) 第25卷 第7號 參照
현재 서울대학교 대학원
제어계측공학과 박사과정
재학중.

李 商 郁 (正會員) 第25卷 第12號 參照
현재 서울대학교 제어계측
공학과 부교수.