

Edge 정보에 근거한 피라미드 영상 부호화

김해성^o 김남철 심영석
 경북대학교 전자공학과

Pyramidal Image Coding using Edge Information

Hae Sung Kim, Nam Chul Kim, Young Serk Shim
 Dept. of Electronics, Kyungpook National Univ.

Abstract

Some modification of Laplacian pyramidal coding have been done and tested for a test image. From our studies, it appears that 3-D predictor which exploits inter- and intra-plane redundancy simultaneously somewhat improves the image quality. We also have tried to reduce the bit rate by only sending zeroth plane image values that correspond to estimated edge points which can be obtained from the first plane. For both cases, the feedback of quantization error in the previous plane has been in the reconstruction of each plane. Subjective and SNR tests show the better performance of the studied methods over the conventional one.

I. 서론

지금까지 주로 연구되어온 영상부호화 기법들은 공간영역에서의 DPCM, 변환영역에서 DCT 등을 이용한 transform coding, DPCM 과 transform coding의 장점을 이용한 hybrid coding 등이 있다. 이러한 방법들은 정보이론과 부호화이론에 근거하여 영상 데이터사이의 redundancy를 제거한 후 영상에 적절한 codeword를 할당하는데 주력하는 부호화 방법이다.[1]

2세대 영상부호화 기법으로 분류되는 피라미드 부호화(pyramidal coding)는 보내고자 영상과 이 영상을 low-pass filtering 한 영상과의 차이를 전송하는 방식으로, 영상을 계층적으로 구성한후 각 평면 별로 부호화하여, 점진적으로 전송하는 방법이다.[2] 이 기법은 0.75 bpp 정도의 비트율에서 좋은 성능을 보이지만 그 이하로 비트율을 낮추기는 매우 어렵다.

본 논문에서는, 보내고자 하는 영상과 이전 평면으로부터 재생된 영상과의 차이를 전송하며, bottom plane에서의 예측오차를 줄이기 위하여 3-D predictor 를 이용하고, 비트율을 품질의 저하없이 0.5 bpp 정도로 낮추기 위하여, 수신된 128x128 영상으로 부터 bottom plane 에서의 edge 의 위치를 판단하여, edge부분만 전송하는 방법등을 제안하였다.

II. 5-Plane Laplacian Pyramidal coding(LPC)

(1). Gaussian plane 및 Laplacian plane 구성

원래의 영상 $G_L(i, j)$ 에서 sampling rate 를 낮추면서 low-pass filtering 을 동시에 행하여 $G_{L+1}(i, j)$ 을 구성한다. 그 구성식은 다음과 같다.

$$G_{L+1}(i, j) = \sum_{m=-1}^1 \sum_{n=-1}^1 w(m, n) G_L(2i+m, 2j+n)$$

여기서, $w(m, n) = \hat{w}(m) \hat{w}(n)$

$$\hat{w}(0)=a$$

$$\hat{w}(-1)=\hat{w}(1)=1/4$$

$$\hat{w}(-2)=\hat{w}(2)=1/4-a/2$$

a : weighting factor

l : plane level

다음, 오차신호 $L_{\mathbf{L}}(i,j)$ 를 구성하며 그 구성식은 다음과 같다.

$$L_{\mathbf{L}}(i,j)=G_{\mathbf{L}}(i,j)-4 \sum_{m=-2}^2 \sum_{n=-2}^2 w(m,n)G_{\mathbf{L}}((i-m)/2,(j-n)/2)$$

이때, 합계 계산시 $(i-m)/2,(j-n)/2$ 가 정수가 되는 위치만 고려된다.[4] 따라서, 위와 같은 과정을 반복하여 Gaussian plane과 Laplacian plane을 원하는 수준의 평면까지 구성할 수 있다.

(2) Quantization & Reconstruction

Entropy coding 시는 uniform quantizer 가 좋은 성능을 가지므로, uniform quantizer를 사용하였다. sampling rate 가 낮은 평면은 quantizer level수를 많이 하고 step size도 작게 하여 양자화 하며, sampling rate가 높은 평면은 영상에서 변화가 심한 부분이므로 level수를 작게 하면서 step size를 크게 하여 부호화한다.[5] quantizer 구성식은 다음과 같다.

$$C_{\mathbf{L}}(i,j)=mn, \quad \text{if } (m-1/2)n < L_{\mathbf{L}}(i,j) < (m+1/2)n$$

m : 양자화기의 level 수
n : step size

가장 높은 평면에서는 $G_4=L_4$ 이며 수신단에서는 다음과 같이 영상을 재구성한다.

$$R_{\mathbf{L}} = C_{\mathbf{L}} + \text{Interpolation of } \{R_{\mathbf{L}+1}\}$$

여기서, $R_{\mathbf{L}}$: reconstructed signal
 $C_{\mathbf{L}}$: quantized signal

III. 5-Plane Modified Pyramidal Coding(MPC)

(1) MPC with 3-D Predictor

LPC 방식에서는 송,수신단이 각기 따로 구성되어 있으므로, 각 평면에서 왜곡이 누적될수 있으므로, 제안한 방법은 각 평면의 왜곡이 다음 평면에서 보상될수있는 Laplacian plane 을구성 하였으며 그 구성식은 다음과 같다,

$$L_{\mathbf{L}}(i,j)=G_{\mathbf{L}}(i,j)-4 \sum_{m=-2}^2 \sum_{n=-2}^2 w(m,n)R_{\mathbf{L}}((i-m)/2,(j-n)/2)$$

실험결과, 최종단에서의 오차신호들이 어느정도 상관성을 갖고있어, 오차신호의 분산을 줄이기 위하여 다음과 같이 3차원 예측을 하였다.

$$\hat{L}_0(i,j)=(\hat{L}_0(i,j-1)+\hat{L}_0(i-1,j))/3 + \hat{L}_1(i+1,j)/3 \times \text{DIV}$$

여기서, DIV 는 각 평면에서의 오차의 분산의 비율 만큼 고려해 준 것이다.

(2) MPC with edge detector

128x128 의 재구성된 영상을 살펴보면 영상의 평탄한 지역및 적은 변화를 갖는 edge부분은 거의 구성된 상태이므로, 최종단에서는 큰 변화를 갖는 edge부분만 보완하면 되는 것을 알 수있다. 따라서,overhead 없이 edge만 전송하기 위하여 일반적으로 사용되는 3x3 Sobel operator[3]를 사용하여 128x128 양자화된 오차신호에서 edge의 위치를 찾았다.

x_1	x_2	x_3
x_4	VAL	x_5
x_6	x_7	x_8

$$\text{VAL}=(H^2+V^2)^{\frac{1}{2}}$$

$$H=(x_1+2x_4+x_6)-(x_3+2x_5+x_8)$$

$$V=(x_1+2x_2+x_3)-(x_6+2x_7+x_8)$$

VAL 이 threshold 보다 크면 edge 로 판단하게 된다. 이렇게, edge 부분만 추출하여

부호화함으로써 비트율을 0.5 bpp 정도로 낮출수있다.

IV. 시뮬레이션 결과 및 고찰

기존의 피라미드 부호화 방법은 plane-8 까지 영상을 구성하지만, plane-5 이상은 영상 전체에 미치는 영향이 적으므로, plane 을 제한하여 구성을 간단히한다음, plane-4 의 16X16 영상은 8 bit PCM 으로 부호화 하였다. Gaussian Plane 을 구성할 때의 weighting factor a 는 GIRL 영상의 경우 0.5 일때 기존의 피라미드 부호화에서 가장 좋은 품질을 나타내었다.

제안한 방법들에 대하여 각 plane 에서의 bit 할당을 거의 같게하여 비트율(여기서는 엔트로피)에 대한 SNR 을 조사하였다. 여기서 SNR 은 다음과 같이 정의 된다.

$$SNR[dB]=10\log_{10} \frac{255^2}{M^2 \sum_{i=1}^M \sum_{j=1}^M [G_o(\lambda_i, j) - R_o(\lambda_i, j)]^2}$$

표 1 을 보면 기존의 방법에 비교하여 약 1-1.2 [dB]의 성능 개선이 있으며, 제안한 방법으로 비트율을 0.5 bpp 정도로 낮추어도 기존의 방법과 비슷한 성능을 보였다.

그림1-4 는 기존의 방법과 제안한 방법들에 의하여 부호화한 영상으로, 그림3은 그림2 와 비교할 때 눈부근 등을 포함하여 전체적으로 선명한 것을 볼수있다. 그림4 는 bottom plane에서 찾은 약 6000 pel 의 edge에 대하여 3 level 로 양자화하여 단지 0.09 bpp 만 할당하여 부호화한 영상으로 그림2 와 비슷한 품질을 보인다.

V. 결 론

기존의 피라미드 부호화는 0.5 bpp 정도로 전송률을 낮출때 배경과 edge 부분에 양자화 잡음이 그대로 나타나 시각상의 품질이 매우나쁘다.

표 1. Performace comparison of various Pyramidal Coding methods

Method	a	bite_rate[bpp]	SNR [dB]
A	0.5	0.762	31.82
A	0.6	0.750	31.20
B	0.6	0.750	32.42
C	0.5	0.756	32.72
D	0.5	0.516	31.43
D	0.6	0.500	31.58

[Note]

A.LPC B.MPC C.MPC with 3-D predictor

D.MPC with edge operator

제안한 방법에서는 Laplacian plane 을 각 plane 에서의 왜곡을 보상하여 주는 방식으로 구성하여 영상의 품질을 향상시켰다. 또한 0.5 bpp 정도로 비트율을 낮추기 위하여 최종단에서 Sobel operator 를 이용하여 edge 부분만 전송함으로써 시각적으로 구분할 수 없을 정도로 영상의 품질을 유지할 수 있었다.

참고문헌

- [1] M.KUNT et al., "Second Generation Image Coding Techniques", Proc. IEEE, vol.73, no.4, pp549-574, Apr 1985.
- [2] P.J.BURT and E.H.ADELSON, "The Laplacian Pyramid as a Compact Image Code", IEEE Trans. Commun. vol.COM-31, no.4, Apr 1983.
- [3] W.K.PRATT. "Digital Image Processing", John Wiley & Sons, 1978.
- [4] P.J.BURT, "Fast filter Transformation for Image Processing", Comput. Graphics, Image Processing, vol.16, pp.20-51, 1981.
- [5] J.MAX, "Quantizer for Minimum Distortion", IRE Trans. Inform. Theory, pp.16-21, Mar 1960.



그림 1. Original GIRL
(256x256 8 bit pixels)



그림 2. LPC
(bit rate 0.762 bpp)



그림 3. MPC with 3-D predictor
(bit rate 0.756 bpp)



그림 4. MPC with edge detector
(bit rate 0.516 bpp)