

論文95-32B-12-8

# 낮은 비트율에서 정지 영상 코딩을 위한 영역 적응 CELP 부호화기

(A region-adaptive CELP image coder for still images at low bit rates)

朴容徹\*, 車日煥\*\*, 尹大熙\*\*

(Yong-Cheol Park, Il-Whan Cha, and Dae-Hee Youn)

## 요 약

본 논문에서는 0.5 bpp 이하의 낮은 비트율에서 효율적인 정지 영상 부호화 방법으로 영역 적응 CELP 코딩 방법을 제안하였다. 제안된 방식은 영상을 minimum spanning tree를 이용하여 통계적 특성이 유사한 영역으로 분할한 후 각 영역에 대해 이차원 선형 예측 모델을 적용하여 예측 계수를 구한다. 부호화는  $8 \times 8$  크기의 블록 단위로 수행하며 블록안에 여러 영역이 존재하는 경우 각 영역에 해당되는 예측 계수를 사용하여 영상을 합성한다. 컴퓨터 시뮬레이션 결과, 제안된 방식은 기존의 블록 적응 CELP 방식과 비교하여 경계 부분에서 특히 우수한 영상을 재생시킬 수 있었다. 또한 JPEG DCT 부호화 방법과의 성능 비교에서도 0.5 bpp 이하의 비트율에서 JPEG DCT 부호화 방법이 블록 왜곡 및 에지 계단화 현상을 나타내는 반면 제안된 CELP 방식에서는 이들 현상이 줄어든 시각적으로 향상된 영상을 재생시킬 수 있었다.

## Abstract

In this paper we propose a region-adaptive CELP image coder for still images at low bit rates below 0.5 bpp. The proposed method partitions the image into stochastically similar regions by the minimum spanning tree method and finds prediction coefficients for each region using a 2-dimensional linear prediction model. Coding is carried out on  $8 \times 8$  blocks and when there are several regions included in a block, an image is synthesized using the prediction coefficients of each region. Computer simulation results show that the proposed method allows improved synthesized image over conventional block-adaptive CELP methods, especially at edges. In addition, performance comparison with the JPEG DCT method shows that while the JPEG method shows block distortion and staircase effects (ragged edges) at bit rates below 0.5 bpp, the proposed CELP method shows improved synthesized images with such effects reduced.

## I. 서 론

\* 正會員, LG 電子 映像 미디어 研究所

(LG Electronics Inc. Image Media Lab.)

\*\* 正會員, 延世大學校 電子工學科 音響 音聲 및 信號 處理 研究室

(Yonsei University Dept. of Electronic Engineering)

接受日字: 1995年8月22日, 수정완료일: 1995年12月4日

기존의 rate-distortion 이론에 근거한 부호화 방법의 압축률이 한계에 다달음에 따라 인간의 시각 특성을 부호화 과정에 이용하므로써 높은 압축율을 얻고자 하는 제 2세대 부호화 방법<sup>[1]</sup>이 제안되었다. 제 2세대 부호화 방법중의 하나인 영역 기반 부호화 방법은 인간의 시각이 영역의 경계와 그 내부를 독립적으로

처리하며 영역 내부의 질감보다는 경계에 더욱 민감하다는 특성을 고려한 부호화 방법으로서 영상을 통계적 특성이 유사한 영역으로 분할한 후 영역 내부는 이차원 polynomial 함수로 모델링하고 영역 경계 정보는 contour 코딩을 통해 부호화한다. 이 방법은 기존의 블럭 적응 부호화 방식들이 모자이크성 영상을 재생하는 40 - 50:1의 높은 압축율에서 영역 경계 및 그 내부의 평균을 보존할 수 있어 시각적으로 보다 의미있는 영상을 재생시킬 수 있다. 그러나 20 - 30:1의 압축율에서는 영역 내부를 평균값이외의 고차 다항식으로 근사화하여도 그 성능 향상이 뚜렷하지 못하다<sup>12)</sup>. 또한 이 방법은 영역 단위로 부호화 과정을 수행함에 따라 블럭 경계 현상을 보이지 않는 반면 영역 경계가 두드러지게 되어 자연스러운 영상을 재생하기 어렵다. 더욱이 영역 분할 과정을 통해 얻은 contour가 실제 영역의 경계와 일치하지 않을 경우 시각적으로 민감하게 느껴지는 false contour가 발생하게 된다.

본 논문에서는 낮은 비트율 (0.5 bpp 이하)에서의 효율적인 정지 영상 부호화 방법으로서 영역 단위로 각기 다른 선형 예측 모델을 이용하여 영상을 합성하는 영역 적응(region-adaptive) CELP (Code Excited Linear Prediction) 부호화기를 제안하였다. CELP 부호화기는 분석-합성 (analysis-synthesis) 부호화 방법으로 선형 예측 모델(Linear Prediction Model)로부터 구한 합성 필터에 코드북에서 선택한 여기 신호를 통과시켜 신호를 합성한다. 따라서 이 방법은 신호를 직접 부호화하는 대신 모델 파라미터와 여기 신호 인덱스만을 부호화하여 높은 압축율을 얻을 수 있으며 음성 신호의 경우 8 Kbps이하의 낮은 비트율에서 대표적인 부호화 방법으로 활발히 연구되어지고 있다<sup>13)</sup>. 영상 신호를 대상으로 제안된 기존의 CELP 방식들은 영상 신호의 통계적 특성이 nonstationary한 것을 보상해주기 위해 일정 크기의 블럭 단위로 선형 예측 모델을 새로이 계산하는 블럭 적응(block-adaptive) 부호화 방식이다<sup>14-7)</sup>. 이 방식들은 통계적 특성이 일정한 블럭들을 낮은 비트율에서 효율적으로 부호화하는 반면 한 블럭 안에 통계적 특성이 상이한 여러 영역이 존재하는 경우 시각적으로 민감하게 느껴지는 경계 부분에서 왜곡이 두드러지게 되는 단점을 지닌다. 이와 비교하여 본 논문에서 제안한 방식은 영상을 통계적 특성이 유사한 영역으로 분할하고 각 영역에 대한 선형 예측 모델을 구한다. 부호

화는 블럭 단위로 수행하며 한 블럭 안에 여러 영역이 존재하는 경우 각 영역에 해당되는 선형 예측 계수를 이용하여 영상을 합성한다. 따라서 한 블럭을 동일한 선형 예측 모델을 사용하는 경우와 비교하여 경계 부분을 효율적으로 합성할 수 있으며 블럭 안에 포함되는 여러 영역을 하나의 여기 신호를 사용하여 동시에 재생하므로써 각 영역을 독립적으로 부호화할때 발생하는 영역 경계 현상 및 false contour를 줄일 수 있다. 또한 이웃하는 블럭의 재생 화소값이 현 블럭의 영상을 재생시키기 위한 합성 필터의 초기 조건으로 작용하므로써 기존 DCT (Discrete Cosine Transform) 부호화기<sup>19)</sup>와 같은 블럭 적응 방식에서 보이는 블럭 경계 현상도 줄어 들게 된다. 이와함께 제안된 방식에서는 각 블럭들을 영역의 경계를 포함하는 에지 블럭과 영역 내부에 존재하는 평탄 블럭으로, 평탄 블럭은 다시 그 블럭의 에너지에 따라 몇 개의 클래스로 분류하여 각 클래스마다 독립적인 여기 신호 코드북을 설계하므로써 부호화기의 성능을 향상시킬 수 있다.

본 논문의 구성은 다음과 같다. II장에서는 영상 신호를 선형 예측 모델인 자기 회환 모델 (autoregressive model)로 모델링하고 임의의 모양을 갖는 영역 단위로 모델 파라미터를 구하는 방식에 대해 살펴 보았다. III장에서는 영역 적응 CELP 부호화기에 대한 설명으로서 MST (Minimum Spanning Tree)를 이용하여 영상을 분할하는 방법<sup>18)</sup>과 이때 얻은 영역 경계 정보를 이용하여 블럭 단위로 부호화를 수행하는 방법에 대해 설명하였다. IV장에서는 제안된 부호화기의 성능을 컴퓨터 시뮬레이션을 통해 기존의 블럭 적응 CELP 부호화기 및 JPEG (Joint Photographic Expert Group)에서 제안한 DCT 부호화기<sup>19)</sup>와 비교하였으며 V장에서 결론을 맺었다.

## II. 영역 적응 이차원 선형 예측

### 1. 영상 모델

이차원 영상 신호  $x(m,n)$ 은 다음과 같은 autoregressive (AR) 모델로 모델링 할 수 있다<sup>10)</sup>.

$$x(m,n) = \sum_k \sum_l a(k,l)x(m-k,n-l) + a_0 + e(m,n) \quad (1)$$

$$(k,l) \in W$$

위에서  $a(k,l)$ 은 선형 예측 계수를,  $e(m,n)$ 은 여기 신호 (excitation signal) 또는 예측 오차 신호

(prediction error signal)를,  $a_0$ 는 영상 신호의 평균값이 영 이상인 것을 보상해주는 bias 값이다. 또한  $k, l$ 의 범위를 나타내는  $W$ 는 예측창(prediction mask)으로서 현 화소를 예측하는데 사용되는 주위 화소의 위치를 나타내며 그 형태에 따라 causal, semicausal, noncausal 모델로 분류된다<sup>[11]</sup>. 본 논문에서는 그림 1과 같은 quarter plane causal 모델을 선택하였으며 이때 예측기의 차수  $P$ 는 현 화소 위치  $(0,0)$ 을 제외한 예측창의 크기인  $Q^2 - 1$ 이 된다.

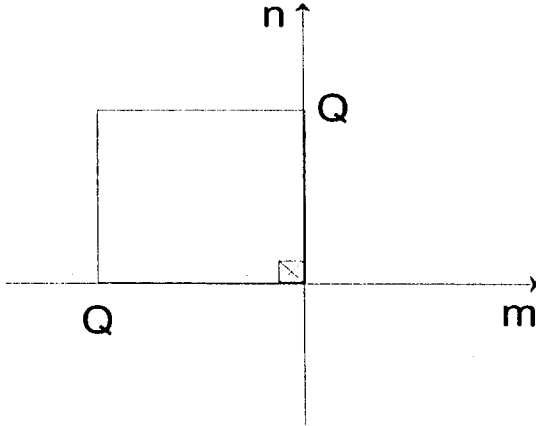


그림 1. Quarter-plane 예측창

Fig. 1. Quarter-plane prediction mask.

기존의 블록 적응 모델에서는 영상의 통계적 특성이 비교적 짧은 구간에서는 stationary하다는 가정하에 그림 2(a)와 같이 일정 크기의 블록 단위로 모델 파라미터를 계산한다. 그러나 이 모델은 음성 신호에서의 일차원 적응 모델을 단순히 이차원 블록으로 확장한 것으로서 영상 고유의 이차원적 특성을 반영하지 못한다. 따라서 본 논문에서는 그림 2(b)에서와 같이 영상을 그 특성이 유사한 영역으로 분할한 후 각 영역을 단위로 모델 파라미터를 구하는 영역 적응 모델<sup>[10]</sup>로 모델링하였다.

## 2. 모델 파라미터 추정

입의 형태를 갖는 영역  $R_i$ 에서의 최적 선형 예측 계수는 아래와 같이 정의되는 예측 오차 에너지  $E$ 를 최소화하도록 구해진다.

$$E = \sum_m \sum_n e^2(m, n) \quad (2a)$$

$$= \sum_m \sum_n [x(m, n) - \sum_k \sum_l a(k, l)x(m-k, n-l) - a_0]^2 \quad (2b)$$

$E$ 를 최소화하는 예측 계수를 구하기 위해  $\partial E / \partial a$

$(k, l) = 0$ 과  $\partial E / \partial a_0 = 0$ 을 풀면 다음과 같은 정규식을 얻는다.

$$\sum_k \sum_l a(k, l) \phi(k, l; i, j) + a_0 s(i, j) = \phi(0, 0; i, j) \quad (3a)$$

$$\sum_k \sum_l a(k, l) s(k, l; i, j) + a_0 N_s = s(0, 0) \quad (3b)$$

위에서  $N_s$ 는  $e(m, n)$ 이 구해지는 범위의 화소수이며  $\phi(\cdot)$ 와  $s(\cdot)$ 는 각각 다음과 같이 정의된다.

$$\phi(k, l; i, j) = \sum_m \sum_n x(m-k, n-l)x(m-i, n-j) \quad (4a)$$

$$s(k, l) = \sum_m \sum_n x(m-k, n-l) \quad (k, l) \in W \quad (4b)$$

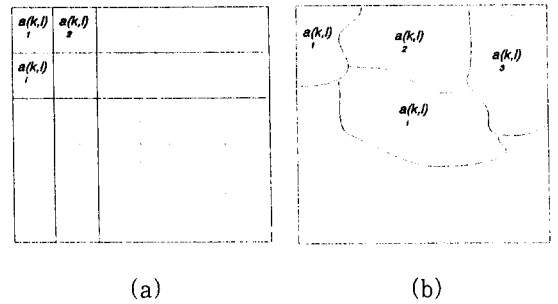


그림 2. 선형 예측 적응 모델

(a) 블록 적응 모델 (b) 영역 적응 모델

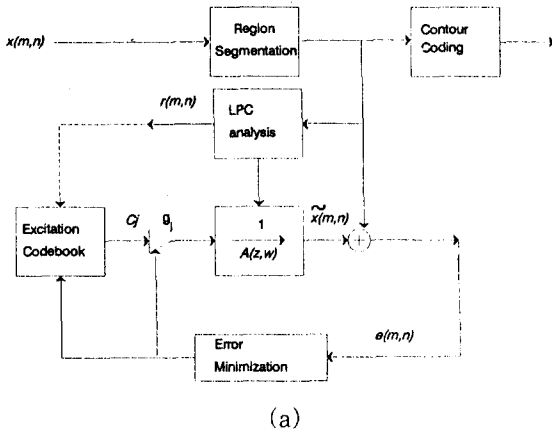
Fig. 2. adaptive LPC models.

(a) block-adaptive model (b) region-adaptive model

선형 예측 모델 파라미터인  $a(k, l)$ 과  $a_0$ 를 구하는 방법은 (2)와 (4)에서  $(m, n)$ 의 범위에 따라 자기 상관법(autocorrelation method)과 공분산법(covariance method)으로 구분된다<sup>[11]</sup>. 자기 상관법의 경우, 영역  $R_i$ 를 벗어난 화소값을 영으로 가정한 후  $e(m, n)$ 의 값이 영이 아닌 모든 위치를  $(m, n)$ 의 범위로 하며 공분산법의 경우에는 인접 영역의 화소값에 대해 특별한 가정을 하지않은 상태에서  $(m, n)$ 의 범위를 영역  $R_i$ 로 제한한다. 이차원 선형 예측 모델에서는 예측창이 유한할 경우 일차원 모델과는 달리 자기 상관법을 사용하여도 (1)에서와 같이  $a(k, l)$ 과  $a_0$ 를 계수로 사용하는 IIR (Infinite Impulse Response) 합성 필터의 안전성을 보장해주지 못한다<sup>[12]</sup>. 따라서 본 논문에서는 예측 오차 에너지 측면에서 유리한 공분산법을 사용하여 모델 파라미터를 구하였으며 [10]에서와 같은 일차원 좌표 변환을 통해 (3)과 (4)를 행렬식으로 나타낸 후 Cholesky decomposition 방

법을 사용하여 해를 구하였다.

한 영역  $R_i$ 에 의해 결정된다.



$$x(m, n) = \sum_k \sum_l a_i(k, l) x(m-k, n-l) + a_{0i} + e(m, n) \quad (k, l) \in W \quad (5)$$

이와같이 제안된 방식에서는 한 블록을 합성할때 각 화소가 속한 영역에 따라 서로 다른 모델 파라미터를 사용하므로 부가 정보로서 영역 경계 정보를 contour coding을 통해 수신단에 전송한다. 제안된 부호화기의 각 부분에 대한 보다 자세한 설명은 다음과 같다.

2. MST를 이용한 영상 분할과 영역 경계 부호화

$N \times N$  크기의 영상은 각 화소를 노드 (node)로 규정하고 4-방향으로 이웃하는 노드를 연결하는 브랜치 (branch)를 정의하므로써  $N_2$ 개의 노드  $V$ 와  $2N(N-1)$ 개의 브랜치  $E$ 로 구성되는 그래프  $G = (V, E)$ 로 표현할 수 있다. 각각의 브랜치에는 노드간의 특성을 나타내는 가중치를 부가하는데 본 논문에서는 화소 값 차의 절대값을 이용하였다. MST는 그래프  $G = (V, E)$ 의 모든 노드로 구성되는 트리 (tree) 중에서 브랜치에 할당된 가중치의 합이 최소가 되는 트리를 말하며 Kruskal의 알고리즘<sup>[8]</sup>을 사용하여 효율적으로 구할 수 있다. MST를 이용한 기본적인 영상 분할 방법은 MST를 구성하는 브랜치를 그 가중치 값이 큰 순서대로 영역의 수 만큼 자른 후 자기 독립된 트리에 포함되는 노드들로서 하나의 영역을 구성한다. 그런데 이 방법에서는 영상 분할 과정시 이웃하는 4-방향 노드간의 국부적인 특성만이 반영되므로 의미있는 분할 영상을 얻기 힘들다. 따라서 이 방법을 수정한 Recursive MST (RMST), Minimax MST (MMST)가 제안되었으며<sup>[8]</sup> 본 논문에서는 이 중 시각적으로 보다 의미있는 영역 분할 결과를 얻을 수 있는 RMST 방법을 사용하여 영상을 분할하였다. 또한 한 영역을 구성하는 노드 수가 일정한 값 이상일 경우에만 브랜치를 자를 수 있도록 알고리즘을 수정하므로써 일반적으로 제약 조건없이 분할하였을때 발생하는 아주 작은 영역들을 효과적으로 제거할 수 있도록 하여 영역 경계를 부호화하는데 필요한 비트수를 줄일 수 있도록 하였다. 그림 4와 같은  $512 \times 512$  크기의 256 밝기를 갖는 Lena 영상을 MST를 이용하여 40개의 영역으로 분할한 결과를 그림 5에 나타내었다. 그림 5는 화소를 중심으로 그것을 둘러싼 4개의 링크(link)를 정의할때 경계 화소들 사이의 링크를 표시한 그림이다.

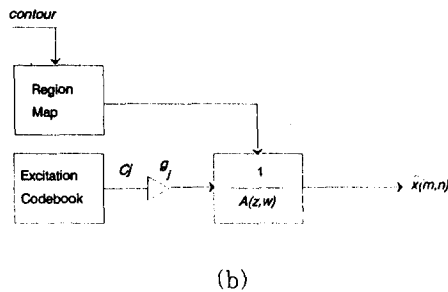


그림 3. 영역 적응 CELP 부호화기  
(a) 인코더 (b) 디코더  
Fig. 3. Block diagram of region-adaptive CELP coder.  
(a) Encoder (b) Decoder

III. 영역 적응 CELP 부호화기

1. 부호화기 기본 구조

영상 분할기, 영역 경계 부호화기, 이차원 블록 CELP 부호화기로 구성되는 영역 적응 CELP 부호화기의 블록 구성도를 그림 3에 나타내었다. 제안된 부호화기에서 영상 분할기는 입력 영상 신호를 MST를 이용하여 통계적 특성이 유사한 몇 개의 영역으로 분할하며 LPC 분석단에서 각 영역을 단위로 모델 파라미터를 구한다. 부호화는  $L \times L$  크기의 블록 단위로 수행하는데 여기 신호를 코드북에서 선택한 후 합성 필터  $H(z_1, z_2)$ 를 통과시켜 영상 신호를 재생한다. 이때 합성 필터의 임펄스 응답은 현 화소가 속한 영역에 의해 결정된다. 즉 아래식에서와 같이 현 화소  $x(m, n)$ 을 입력 여기 신호  $e(m, n)$ 과 주위 화소  $x(m-k, n-l)$ 를 이용하여 합성하는 경우에 모델 파라미터  $a_i(k, l)$ 과 bias 값  $a_0$ , 는 그림 2-b에서와 같이 현 화소가 속

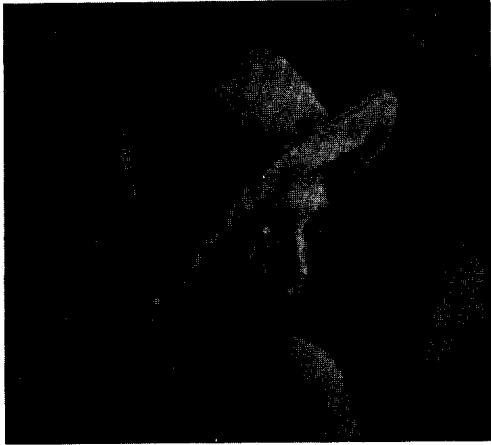


그림 4. Lena 원 영상 (512×512 pixels, 8 bit/pixel)

Fig. 4. Original Lena image (512×512 pixels, 8 bit/pixel)



그림 5. MST를 이용한 영역 분할(영역 수 = 40)

Fig. 5. Segmentation using MST (number of region = 40)

영역 경계를 부호화하는 일반적인 방법으로는 4-방향 chain 코드나 chain-difference 코드를 이용하는 방법<sup>[13,14]</sup>이 있다. chain 코딩 방법은 chain의 시작 좌표, 경계를 구성하는 링크의 방향 정보, 그리고 경계의 끝을 나타내는 end 코드를 하나의 package로하여 영역 경계를 나타낸다. 이 방법은 링크의 방향을 나타내는 정보가 상당히 중복된다. chain difference 코드에서는 이를 해결하기 위해 절대 방향을 기준으로 링크의 상대적인 방향 변화를 부호화한다. 그러나 이들

방법 모두는 경계의 분기점을 단순히 경계의 끝으로 인식하고 분기점마다 새로히 경계의 시작 위치를 부호화하므로 경계의 시작 위치를 코딩하는데 비트를 너무 많이 소모하는 단점을 지닌다. 따라서 본 논문에서는 이와같은 단점을 해결하고 4-방향 영역 경계의 특성을 살려 효율적으로 영역의 경계를 나타낸 Biggar가 제안한 방법<sup>[21]</sup>으로 영역 경계를 부호화하였다.

### 3. 최적 입력 벡터 선택 방법

$L \times L$  크기 블록 내의 영상 신호를  $X = \{x(m, n); m, n = 0, 1, \dots, L-1\}$ 로, 같은 크기의 재생 영상 신호  $\hat{x}(m, n)$  및 입력 벡터  $c(m, n)$ , 에러 신호  $e(m, n)$ 를 각각  $X, C, E$ 로 나타내고 다음과 같이  $f$ 번째의 항만이 1의 값을 갖고 모두 영인 벡터를 정의하면,

$$e_f = [ \underbrace{0001 \dots 000}_f ] \quad f = 0, 1, \dots, L-1 \quad (6)$$

이차원  $X, C, E$  신호를 열 방향을 우선으로하여  $L^2$  차원의 일차원 벡터로 나타낸  $x, \hat{x}, c, e$ 는 다음과 같이 표현할 수 있다.

$$x = [ e_0 X \dots e_f X \dots e_{L-1} X ] \quad (7a)$$

$$\hat{x} = [ e_0 \hat{X} \dots e_f \hat{X} \dots e_{L-1} \hat{X} ] \quad (7b)$$

$$c = [ e_0 C \dots e_f C \dots e_{L-1} C ] \quad (7c)$$

$$e = [ e_0 E \dots e_f E \dots e_{L-1} E ] \quad (7d)$$

재생 신호  $\hat{x}$ 를 합성 필터  $H(z_1, z_2)$ 의 임펄스 응답과 입력 벡터  $c$ 와의 행렬식으로 나타내기 위해서 다음과 같은 식들을 정의한다.

이차원 IIR 합성 필터의  $z$ -변환식은 아래와 같다.

$$H(z_1, z_2) = \frac{1}{A(z_1, z_2)} = \sum_{i,j=0}^{\infty} h_{i,j} z_1^{-i} z_2^{-j} \quad (8)$$

윗식에서 합성 필터 계수만을 모아 다음에서와 같이 이차원 행렬  $F$ 로 나타낸다.

$$F = [ h_{i,j}; i, j = 0, 1, 2, \dots ] \quad (9)$$

무한 크기의 이차원 행렬  $F$ 를 분석 구간의 크기  $L \times L$ 로 제한한 것을  $H_T$ 라 정의할때

$$H_T = [ h_{i,j}; i, j = 0, 1, 2, \dots, L-1 ] \quad (10a)$$

$$= \begin{bmatrix} h_{00} & h_{01} & \dots & h_{0,L-1} \\ h_{10} & h_{11} & \dots & h_{1,L-1} \\ \vdots & \vdots & \dots & \vdots \\ h_{L-1,0} & h_{L-1,1} & \dots & h_{L-1,L-1} \end{bmatrix} \quad (10b)$$

$H_T$ 를 이차원 각 방향으로  $p$ 와  $q$  샘플만큼 지연시킨 결과는 다음과 같다.

$$z_1^p z_2^q H_T = [ z_1^p z_2^q h_{i,j} ; i, j = 0, 1, 2, \dots, L-1 ] \quad (11a)$$

$$= [ h_{i-p, j-q} ; i, j = 0, 1, \dots, L-1 ] \quad (11b)$$

$$h_{k,l} = 0 \text{ if } k < 0 \text{ or } l < 0$$

$$= \begin{bmatrix} 0 & \dots & 0 & \dots & \dots & 0 \\ 0 & \dots & h_{00} & h_{01} & \dots & h_{0, L-q} \\ 0 & \dots & h_{10} & h_{11} & \dots & h_{1, L-q} \\ \vdots & \dots & \vdots & \vdots & \dots & \vdots \\ 0 & \dots & h_{L-p, 0} & h_{L-p, 1} & \dots & h_{L-p, L-q} \end{bmatrix} \quad (11c)$$

(11)의  $L \times L$  크기의 행렬은 (6)의 벡터  $e_i$ 를 이용하여  $L^2$  차원의 벡터  $\mu_{pq}$ 로 나타낼수 있으며

$$\mu_{pq} = [ e_{00} z_1^p z_2^q H_T \dots e_{i1} z_1^p z_2^q H_T \dots e_{L-1, L-1} z_1^p z_2^q H_T ] \quad (12)$$

다음과 같이  $\mu_{00}$ 를 첫번째 열로,  $\mu_{L-1, L-1}$ 를 마지막 열로 하는  $L^2 \times L^2$  크기의 행렬  $H$ 를 정의한다.

$$H = [ \mu_{p,q} ; p, q = 0, 1, 2, \dots, L-1 ] \quad (13)$$

현 블럭의 최적 입력 벡터와 그 gain을 구할때 과거 블럭의 영향을 없애주기 위하여 원 신호에서 합성 필터의 초기 메모리에 의해 얻어지는 zero 입력에 대한 응답을 빼준 신호를 기준 신호  $\hat{x}$ 로 하여 재생 신호와의 오차 신호를 구한다. 그리고 입력 벡터에 의한 재생 신호를 구할때 합성 필터의 메모리를 초기화 시킨다. 따라서 현 블럭내에서 (8)과 같이 주어지는 합성 필터는 (10)의 FIR (Finite Impulse Response) 필터로 나타낼 수 있으며 재생 신호  $\hat{x}_j$ 와 입력 벡터  $e_j$ 와의 관계는 (13)을 이용하여 다음과 같은 행렬식으로 표현할 수 있다. 아래에서  $g_j$ 는 입력 벡터  $e_j$ 에 대응되는 gain이다.

$$\hat{x}_j = g_j H e_j \quad (14)$$

기준 신호  $\hat{x}$ 와 재생 신호  $\hat{x}_j$ 와의 차신호 에너지를 구하면 다음과 같다.

$$E_j = \| e_j \|^2 = \| \hat{x} - \hat{x}_j \|^2 = \| \hat{x} - g_j H e_j \|^2 \quad (15)$$

윗식에서  $\| \cdot \|$ 는 Euclidean vector norm으로서 다음에서와 같이 벡터의 크기를 나타내며

$$\| x \| = (|x_1|^2 + \dots + |x_n|^2)^{1/2} = (x^T \cdot x)^{1/2} \quad (16)$$

(15)를 이용하여  $\partial E_j / \partial g_j = 0$ 의 해를 구하므로써 입

력 벡터  $e_j$ 의 최적 gain을 다음과 같이 결정할 수 있다.

$$g_{j, opt} = \frac{\hat{x}_j^T H e_j}{\| H e_j \|^2} \quad (17)$$

최적 입력 벡터는 (17)과 (15)로부터 구해지는 아래와 같은 예측 오차 에너지  $E_j$ 를 최소화하도록 코드북에서 선택한다.

$$E_j = \| \hat{x} \|^2 - \frac{\hat{x}_j^T H e_j}{\| H e_j \|^2} \quad (18)$$

코드북은 예측 오차 신호와 유사한 확률 밀도 함수를 갖도록 설계되는데 음성 신호에서는 Gaussian random noise를 주로 이용하는 반면<sup>[15]</sup> 영상 신호에서는 직접 예측 오차 신호로부터 LBG 알고리즘<sup>[6]</sup>을 사용하여 설계한다.

#### IV. 실험 및 결과 고찰

영역 적응 CELP 부호화기와 이와의 성능 비교를 위한 블럭 적응 CELP 부호화기의 실험 조건을 표 1에 나타내었다. 여기 신호의 코드북은  $512 \times 512$  크기 및  $256 \times 256$  크기와 256 밝기 단계를 갖는 7 장의 영상을 대상으로 예측 오차 신호를 구한 후 LBG 알고리즘을 사용하여 설계하였다. 이때 영역 적응 CELP 부호화기에서는  $8 \times 8$  크기의 각 블럭들을 영역 경계를 포함하는 에지 블럭과 영역 내에 존재하는 평탄 블럭으로, 평탄 블럭은 다시 예측 오차 신호의 분산값에 따라 2 개의 클래스로 분류하여 (문턱치 = 20) 각각 독립적인 코드북을 설계하였다. III장에서 설명된 바와 같이 영역 적응 CELP 부호화기에서는 블럭 단위로 영상 신호를 합성할때 합성 필터의 임펄스 응답 및 bias 값이 현 화소의 영역 인덱스에 따라 변하게 된다. 따라서 각 화소의 영역 인덱스를 부가 정보로 수신단에 전송해 주어야 하는데 이는 영역 경계 정보를 부호화하여 전송하므로써 해결할 수 있다. 본 논문에서는 영역 경계를 나타내는 그림 5와 같은 line diagram을 4 방향 chain 코드로 나타낸 후 Biggar가 제안한 방법<sup>[2]</sup>을 사용하여 부호화하였다. 그림 5의 Lena 영상 line diagram의 경우 링크 수가 14793으로 18732 비트를 사용하여 부호화할 수 있었으며 다른 실험 영상에 있어서도 화소 당 평균 0.06 - 0.07 bpp (bit / pixel)로 영역 경계 정보를 표현할 수 있었다.

표 1. CELP 부호화기의 실험 조건  
Table 1. Experiment conditon of CELP coder.

부호화 방법 파라미터	블럭 적응 CELP	영역 적응 CELP
예측기 차수 P	3	
모델 파라미터 추정	공분산법	
분석 구간	(32 × 32) 블럭	영역 단위
합성 구간	(8 × 8) 블럭	
코드북 크기	1024	1024
코드북 수	1	3

블럭 적응 모델과 영역 적응 모델에서의 비트율은 다음과 같이 구할 수 있다.

블럭 적응 모델에서의 비트율 :

$$\begin{aligned} \text{비트율(bpp)} = & (P \times \text{BIT\_LPC} + \text{BIT\_a}_0) / \\ & (M \times M) \\ & + (\text{BIT\_Code} + \text{BIT\_g}) / (I \times I) \end{aligned} \quad (19)$$

영역 적응 모델에서의 비트율 :

$$\begin{aligned} \text{비트율(bpp)} = & (P \times \text{BIT\_LPC} + \text{BIT\_a}_0) \times \\ & \text{영역 수} / (N \times N) \\ & + (\text{BIT\_g}_s \times \text{평탄 블럭 수} + \text{BIT\_g}_e \times \text{에지 블럭 수}) / (N \times N) \\ & + \text{BIT\_Code} / (I \times I) \\ & + \text{BIT\_CLASS} \times \text{평탄블럭수} / (N \times N) \\ & + \text{BIT\_Edge} / (N \times N) \end{aligned} \quad (20)$$

(19)과 (20)에서 BIT\_LPC, BIT\_a<sub>0</sub>, BIT\_g, BIT\_g<sub>s</sub>, BIT\_g<sub>e</sub>, BIT\_CLASS, BIT\_Edge는 각각 선형 예측 계수, 바이어스, 이득치(g=블럭 적응, g<sub>s</sub>=영역 적응 평탄 블럭, g<sub>e</sub>=영역 적응 에지 블럭), 블럭 클래스 및 영역 경계를 나타내기 위해 할당된 비트수를 말하며 N×N은 입력 영상의 크기를, M×M은 합성 구간의 크기, I×I는 블럭 적응 모델에서 예측 계수를 구하는 분석 구간의 크기를 나타낸다. 본 실험에서는 그 크기를 각각 N = 512, M = 32, I = 8로 하였으며 표 1과 같은 실험 조건에서 코드북 크기를

1024로하고 표 2와 같은 비트 할당으로 512×512 크기의 lena 영상을 부호화했을때 전체 비트율은 0.32 bpp가 된다.

표 2. 선형 예측 모델 파라미터 비트 할당  
Table 2. Bit allocation of LPC parameters.

부호화 방법 파라미터	블럭 적응 CELP (비트수)	영역 적응 CELP (비트수)
a(k,l)	10	10
a <sub>0</sub>	8	10
g	8	에지 블럭 6 평탄 블럭 5

Lena 영상을 대상으로 모델 파라미터에 대한 비트 할당을 조절하면서 구한 여러 비트율에서 각 방법의 PSNR을 그림 6에 나타내었다. 이 결과를 보면 JPEG DCT 방법, 영역 적응 CELP, 블럭 적응 CELP 순으로 부호화 성능을 보이고 있다. 그러나 대부분 영상 압축의 응용에 있어 재생 영상의 최종 평가는 인간에 의해 이루어지며 따라서 PSNR보다 주관적인 시각 판단이 더욱 중요하다. 0.32 bpp에서 부호화된 결과를 그림 7, 8, 9에, 그리고 각 방법의 차이가 두드러진 부분을 확대한 그림을 그림 10, 11, 12에 각각 나타내었다. 그림 7과 8, 10과 11을 비교해보면 영역 적응 부호화기가 전반적으로 경계 부분을 잘 재생하고 있음을 알 수 있다.

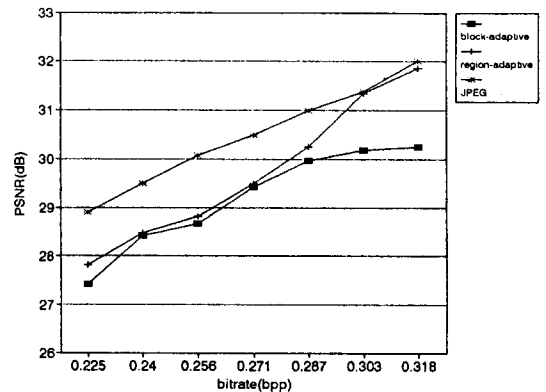


그림 6. 재생된 lena 영상의 PSNRs  
Fig. 6. PSNRs of decoded Lena image at various bit rates.

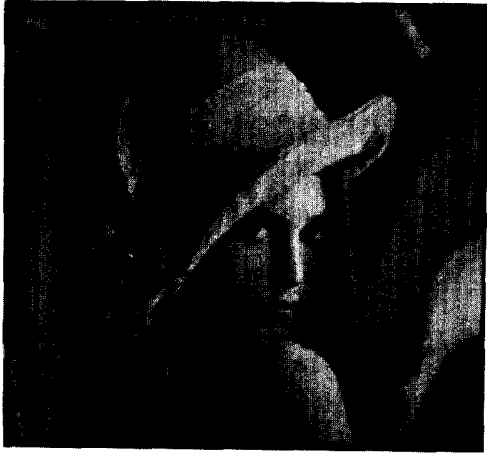


그림 7. 블럭 적응 CELP 방법으로 재생된 영상 (0.32 bpp)  
 Fig. 7. Decoded image from block-adaptive CELP coder (0.32 bpp).

장 우수한 영상을 재생시킴을 알 수 있다.

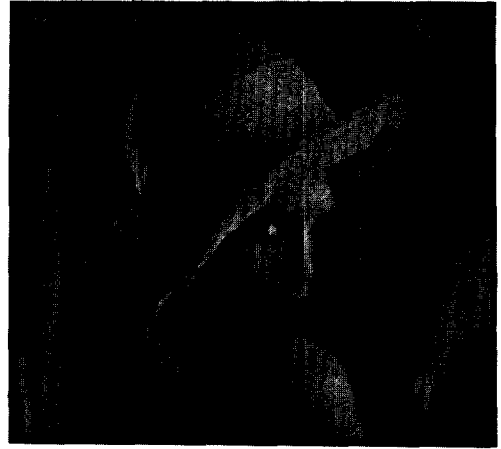


그림 9. JPEG DCT 방법으로 재생된 영상 (0.32 bpp)  
 Fig. 9. Decoded image from JPEG DCT coder (0.32 bpp).

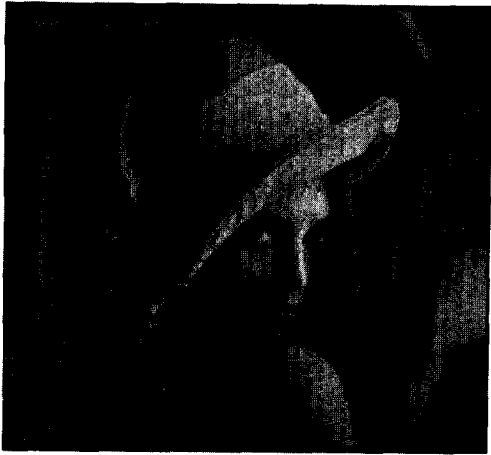


그림 8. 영역 적응 CELP 방법으로 재생된 영상 (0.32 bpp)  
 Fig. 8. Decoded image from region-adaptive CELP coder (0.32 bpp).



그림 10. 블럭 적응 CELP 방법으로 재생된 확대 영상 (0.32 bpp)  
 Fig. 10. Enlarged image decoded from block-adaptive CELP coder (0.32 bpp).

특히 블럭 적응 방법은 lena 영상의 모자 끝 부분에서와 같이 블럭내에 평균값 차이가 심한 영역이 존재할 경우 수평 또는 수직 방향으로 선 모양의 에러를 발생시킴을 관찰할 수 있다. JPEG 부호화기는 그림 9와 12에서 보듯이 평탄 부분에서 심한 블럭 현상을 보이며 깃털 부분과 같이 고주파수 성분이 많이 존재하는 영역에서는 CELP 방식보다 뚜렷한 영상을 재생시키는 반면 에지의 계단화 현상을 보인다. 따라서 전체적인 화질 평가에 있어 제안된 방식이 시각적으로 가

### V. 결 론

본 논문에서는 낮은 비트율에서의 효율적인 정지 영상 부호화 방법으로서 영역 적응 CELP 코딩 방법을 제안하고 기존의 블럭 적응 CELP 방식 및 JPEG 방법과 성능 비교를 수행하였다. 컴퓨터 시뮬레이션 결과, 제안된 방식은 기존의 블럭 적응 CELP 방식과 비





그림 11. 영역 적응 CELP 방법으로 재생된 확대 영상 (0.32 bpp)

Fig. 11. Enlarged image decoded from region-adaptive CELP coder. (0.32 bpp)



그림 12. JPEG DCT 방법으로 재생된 확대 영상 (0.32 bpp)

Fig. 12. Enlarged image decoded from JPEG DCT coder. (0.32 bpp)

교해 경계 부분에서 우수한 영상을 재생시킬 수 있었다. 또한 0.5 bpp 이하의 비트율에서 JPEG 방법이 블럭 왜곡 및 에지의 계단화 현상을 나타내는 반면 제안된 방식에서는 이들 현상이 줄어든 시각적으로 향상된 재생 영상을 얻을 수 있었다.

본 연구와 관련하여 앞으로는 VSELP<sup>[16]</sup>와 같은 구조를 갖는 코드북을 설계하여 시스템의 복잡도를 줄이는 방법 및 YUV 컬러 영상 부호화시 Y 영상으로부

터 구한 영역 경계 정보를 U,V 신호에 적용하므로써 영역 경계 정보에 소요되는 비트수를 상대적으로 줄이는 방법, 그리고 시각 특성을 고려한 여러 가중 필터를 설계하여 코드북 선택시에 사용하므로써 부호화기의 성능을 향상시키는 방법에 대한 연구가 요망된다.

### 참 고 문 헌

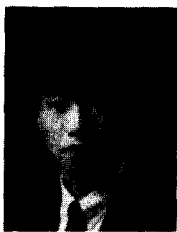
- [1] M. Kunt, M. Benard and R. Leonardi, "Recent results in high-compression image coding," *IEEE Trans. Circuit and Systems*, vol. 34, pp. 1306-1336, November 1987.
- [2] M. J. Biggar, O. J. Morris and A. G. Contantinides, "Segmented-image coding : performance comparison with the discrete cosine transform," *IEE Proc. Pt.F., Commun. Radar and Signal Process.*, vol. 135, no. 2, pp. 121-131, April 1988.
- [3] B. S. Atal, V. Cuperman and A. Gersho, Edited, *Advances in speech coding*, Kluwer Academic Publishers, 1991.
- [4] C. Horne and E. F. Deprettere, "Multi-pulse and regular-pulse LP coding of images," *Proc. 1987 IEEE Int. Conf. Acoust., Speech, Signal Processing*, pp. 1083-1086.
- [5] V. J. Mathews, R. W. Waite and T.D. Tran, "Image compression using vector quantization of linear (one-step) prediction errors," *Proc. 1987 IEEE Int. Conf. Acoust., Speech, Signal Processing*, pp. 733-736.
- [6] K. K. Paliwal, "Low bit-rate image coding using 2-D linear prediction and 2-D stochastic excitation," *Proc. IEEE 1988 Int. Conf. Acoust., Speech, Signal Processing*, pp. 800-803.
- [7] L. Torres and E. Arias, "Stochastic vector quantization of images," *Proc. 1992 IEEE Int. Conf. Acoust., Speech, Signal Processing*, III, 385-388.
- [8] O.J. Morris, M.de J. Lee and A.G.

- Constantinides, "Graph theory for image analysis : an approach based on the shortest spanning tree." *IEE Proc. Pt. F., Commun. Radar and Signal Process.*, vol. 133, pp. 146-152, April 1986
- [9] G. K. Wallace, "The JPEG still picture compression standard." *Commun., ACM.* vol. 34, pp. 30-44, April 1991.
- [10] P. Maragos, R. M. Mersereau and R. W. Schafer, "Two-dimensional linear predictive analysis of arbitrary-shaped regions." *Proc. 1983 IEEE Int. Conf. Acoust., Speech, Signal Processing.* V, 237-240.
- [11] A. K. Jain, "Advances in mathematical models for image processing." *Proc. IEEE*, vol. 69, pp. 502-528, May 1981.
- [12] T. L. Marzetta, "Two-dimensional linear prediction : autocorrelation arrays, minimum-phase prediction error filters, and reflection coefficients arrays." *IEEE Trans. Acoust., Speech, Signal Process.* vol. 28, no. 6, pp. 725-733, Dec. 1980.
- [13] H. Freeman, "Computer processing of line drawings." *Comput. Surveys*, vol. 6, pp. 57-97, 1974.
- [14] M. Eden, and M. Kocher, "On the performance of a contour coding algorithm in the context of image coding, Part 1 : Contour segment coding." *Signal Processing*, vol. 8, pp. 381-386, July 1985.
- [15] M. R. Schroeder and B. S. Atal, "Code-Excited Linear Prediction(CELP) : High quality speech at very low bit rates." *Proc. 1985 IEEE Int. Conf. Acoust., Speech, Signal Processing.* pp. 937-940.
- [16] I. Gerson and M. Jasiuk, "Vector Sum Excited Linear Prediction(VSELP)." *IEEE workshop on Speech Coding for Telecommunications*, pp. 66-68, September 1989.

---

 저 자 소 개
 

---



朴 容 徹(正會員)

1963년 4월 30일생. 1987년 2월 연세대학교 전자공학과 졸업(공학사). 1989년 2월 연세대학교 대학원 전자공학과 졸업(공학석사). 1994년 2월 연세대학교 대학원 전자공학과 졸업(공학박사). 1994년 3월 ~ 10

월 한국 전자 통신 연구소, 1994년 10월 ~ 현재 LG 전자 영상 미디어 연구소 선임 연구원. 주관심 분야는 A/V 신호처리, 영상 데이터 압축

尹 大 熙(正會員) 第 31卷 B編 第 10號 參照

현재 연세대학교 전자공학과 교수

車 日 煥(正會員) 第 31卷 B編 第 9號 參照

현재 연세대학교 전자공학과 교수