

## DCT 변환계수의 인덱스의존 (줄길이, 레벨) 부호화

## (Frequency Dependent (Runlength, Level) Coding of DCT Coefficients)

李鍾和\*, 姜東郁\*\*\*, 姜守遠\*\*, 張俊鎬\*\*, 李忠雄\*\*\*

(Jong-Hwa Lee, Dong-Wook Kang, Su-Won Kang, Joon-Ho Chang and Choong Woong Lee)

## 要約

본 논문에서는 확률분포함수의 구조적인 차이를 이용하여, DCT계수를 효율적으로 부호화하는 새로운 접근에 대하여 논의하였다. 신호원 모델에 따라 신호를 분리함으로써 신호의 엔트로피를 감소시킬 수 있음을 해석적으로 증명하였으며, (줄길이, 레벨) 심볼들을 그 심볼의 인덱스에 따라 적응적으로 부호화함으로써 평균부호길이가 감소함을 실험을 통하여 보였다.

## Abstract

In this paper, we have addressed a new approach to encode the DCT coefficients efficiently by exploiting the structural differences among the probability distribution functions of them. We have proved analytically that if signals are separated according to their source models, the entropy of the separated signals decreases below that of the unseparated signals. And simulation results show that the average code length decrease when each (runlength, level) symbol is encoded adaptively according to the index associated with the symbol.

## 1. 서론

디지털 영상을 효과적으로 전송하거나 저장하기 위해서는 신호에 내포되어 있는 중복성을 제거하여 데이터율을 감축하는 (대역압축) 과정이 필수적이다.

대역압축기법은 크게 예측부호화와 변환부호화로 구분할 수 있는데, 고압축을 위해서는 이 두 부호화 기법을 결합한 하이브리드 방식이 많이 이용되고 있다.<sup>1,2</sup> 움직임보상 변환부호화가 가장 대표적인 하이브리드 방식인데, 시간방향의 신호의 중복성은 예측의 일종인 움직임보상을 통하여 제거하고, 2차원 공간에서의 신호의 중복성은 2차원 변환부호화를 통하여 제거하는 구조로 되어있으며, 2차원 변환을 위한 변환기로서는 DCT 커널(kernel)을 이용하는 것이 보편화되어있다.<sup>3,4</sup>

이러한 방식의 부호화시스템은 움직임예측/보상.

\* 正會員, 韓國放送公社 技術研究所

(Tech. Research Institutue, KBS)

\*\* 準會員, \*\*\* 正會員, 서울大學校 電子工學科

(Dept. of Elec. Eng., Seoul Nat'l Univ.)

接受日字 : 1993年 11月 19日

DCT, 양자화기 및 엔트로피부호화기로 구성된다. 부호화시스템의 전체성능을 향상시키기 위해서, 양자화기의 경우 시각적응적 양자화기법, 엔트로피부호화의 경우는 DCT계수의 통계적인 특성을 이용하는 기법이 연구되고 있다. 시각적응적 양자화란 신호가 공간적으로 자주 또는 급히 변하는 경우, 시각적으로 일종의 해상도 차단효과(masking effect)가 있음을 이용하여 가능한한 양자화의 폭을 넓혀 양자화된 계수의 엔트로피를 최소화하고자 양자화 기법을 말하며, 엔트로피부호화는 부호화하고자 하는 심볼의 엔트로피에 접근하는 평균부호길이로써 심볼을 부호화하는 과정을 말한다.<sup>15)</sup>

움직임보상 변환부호화시, 양자화된 DCT 계수들은 대부분 '0'이므로 이 현상을 이용한 적절한 부호화 방법이 필요하다. 이를 위해 일반적으로 (줄길이, 레벨) 부호화를 사용하는데, 이는 2차원 DCT블럭을 특정 순서로 스캐닝한 후, 연속하는 '0'의 줄길기와 이어지는 '0'이 아닌 계수를 부호화하는 방법이다. 이때 부호화할 신호의 수는 '0'이 아닌 계수의 갯수에 의해 결정되기 때문에, 엔트로피부호화의 효율은 각 신호의 엔트로피를 얼마나 떨어뜨리도록 신호를 표현하는가에만 좌우된다. 일반적으로 스캐닝은 저주파수 성분으로부터 고주파수 성분의 차례로 하게 되는데, 지그재그 스캐닝의 경우 (줄길이, 레벨) 심볼의 엔트로피가 다른 스캐닝방법에 비하여 낮다는 것이 알려져 있다.

한편, DCT부호화는 구획부호화의 일종으로서, 인접 구획간의 DCT계수 사이에는 상당히 큰 상관성을 갖고 있으므로, 고도의 부호화효율을 얻기 위해서 여러개의 구획을 동시에 부호화하는 기법을 이용할 수 있다.<sup>16,17)</sup> 그러나, 이 기법은 여러개의 구획을 동시에 부호화함으로써 구획간의 상관성을 충분히 제거할 수는 있지만 부호화하기 위한 구조가 복잡하다는 단점이 있다.

본 논문에서는 각 구획을 독립적으로 부호화하면서, 부호화의 효율을 향상시키기 위한 방법을 제시한다. 이 방법은 DCT계수의 각 (줄길이, 레벨) 심볼마다 그 발생위치와 관련된 정보인 인덱스를 할당하고, (줄길이, 레벨) 특성이 그 인덱스에 따라 현격하게 다르다는 성질을 이용한다. 이를 위해 발생된 신호들을 인덱스에 따라 가상적인 몇개의 신호원의 출력으로서 모델링하고, 이 신호원 모델에 따라 각각의 신호원에 정합된 허프만 부호화를 행하는 것이다. 제 II장에서는 신호를 분리하여 부호화함으로써 얻을 수 있는 엔트로피 감소효과에 대하여 분석하며, 제 III장에서는 DCT 계수의 (줄길이, 레벨) 심볼을 그 심볼이 존재하는 인덱스에 따라 가상적인 신호원으로 분리하는 방법으로서 (줄길이, 레벨) 심볼의 clustering기법을 제시하

고, 분리된 심볼의 분포특성의 변화를 보이며, 부호기와 복호기 구성을 보인다. 제 IV장에서는 신호원 모델에 따라 신호를 분리함으로써 발생한 엔트로피의 감소와 각 신호원에 정합된 부호책을 사용하여 허프만 부호화하여 얻은 비트율 감축 결과를 보인다.

## II. 신호 분리에 따른 엔트로피 감소효과

### 1. 신호원 분리 모델

메시각 이산적 랜덤변수 X와 Y를 발생시키는 두개의 상호 독립적인 신호원이 있다고 가정하자. 또한, X와 Y의 알파벳과 확률밀도함수는 각각  $\Omega_X$ ,  $\Omega_Y$  및  $\{p_i\}$ ,  $\{q_i\}$ 라고 가정하자. 이때 n번째 출력에서 랜덤변수 X와 Y의 관찰값을 각각  $x(n)$  및  $y(n)$ 로 나타낸다. 여기서 메시각 두 관찰값중의 하나를 확률  $\alpha$ 와  $1-\alpha$ 로 선택하면, 이 선택에 의하여 새로이 얻어지는 계열  $z(n)$ 은 다음과 같은 알파벳과 확률밀도함수를 갖는 새로운 집합적 신호원의 출력으로 고려할 수 있다:

알파벳 :

$$\Omega_Z = \Omega_X \cup \Omega_Y \tag{1}$$

확률밀도함수 :

$$\Pr(Z = z_i) = s_i = \begin{cases} \alpha p_i, & z_i \in \Omega_X - \Omega_Y \\ \alpha p_i + (1-\alpha)q_i, & z_i \in \Omega_X \cap \Omega_Y \\ (1-q_i), & z_i \in \Omega_Y - \Omega_X \\ 0, & \text{otherwise} \end{cases} \tag{2}$$

이 경우,  $z(n)$ 을 발생시키는 신호원을 두개의 성분 신호원으로 구성된 복합신호원이라고 일컫으며, 랜덤변수 Z를 복합기호, X와 Y를 성분기호라고 한다.

(2-성분 복합신호원 모델)

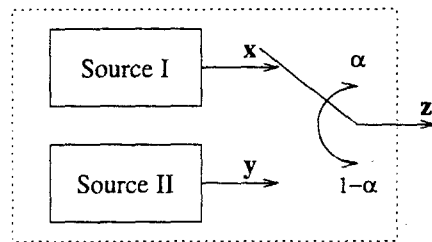


그림 1. 2-성분 복합신호원 모델

Fig. 1. Two-component composite source model.

다음에 이러한 2-성분 복합신호원 모델에 있어서 의미있는 두가지의 정보량을 정의한다. 이 두 정보량

은 앞으로 본 논문에서 다루려고 하는 엔트로피 부호기의 도달가능한 최소평균부호길이의와 매우 밀접하게 관련되어 있다.

정의 1: (성분정보량과 복합정보량) 그림 1과 같은 2-성분 복합신호원을 고려하자. X와 Y 각성분기호의 엔트로피를 그 선택확률  $\alpha$ 와  $1-\alpha$ 로 가중하여 합한 평균 엔트로피를 이 2-성분 복합신호원의 성분정보량이라고 하고, 복합기호 Z의 엔트로피를 복합정보량이라고 한다. 한편 이 정보량들을 각각  $H_n(Z)$ 과  $H_s(Z)$ 로 나타낸다:

$$H_n(z) = \alpha H(X) + (1 - \alpha)H(Y) \tag{3}$$

$$H_s(Z) = H(Z) \tag{4}$$

단,  $H(X)$ ,  $H(Y)$ 와  $H(Z)$ 는 각각 기호 X, Y와 Z의 엔트로피를 나타낸다.

정의 2: (신호원분리이득) 그림 1의 2-성분 복합신호원에서, 복합정보량에서 성분정보량을 뺀 나머지가 이 복합신호원의 신호원분리이득이라고 하고, 이를  $SG(Z)$ 라고 나타낸다:

$$SG(Z) = H_s(Z) - H_n(Z) \text{ (bits)} \tag{5}$$

이제 2-성분 복합신호원 모델에서 신호원분리이득을 해석적으로 분석해 보자. 먼저 복합정보량은 복합기호의 확률밀도함수로부터 다음과 같이 얻을 수 있다.

$$H_s(Z) = \sum_{z \in \Omega_Z} s_z \cdot \log \frac{1}{s_z} \tag{6}$$

식 (2)에서 주어진 확률밀도함수를 적용하여 정리한 후 적당한 수식 전개를 거치면 다음과 같은 결과식을 얻을 수 있다. (부록 A 참조)

$$\begin{aligned} H_s(Z) &= \alpha H(X) + (1 - \alpha)H(Y) \\ &+ \alpha \log \frac{1}{\alpha} \Pr\{z_i \in \Omega_x - \Omega_y | X\} \\ &+ (1 - \alpha) \log \frac{1}{1 - \alpha} \Pr\{z_i \in \Omega_y - \Omega_x | Y\} \\ &+ \alpha \left[ \sum_{z_i \in \Omega_x \cap \Omega_y} p_i \cdot \log \frac{1}{s_i} - \sum_{z_i \in \Omega_x \cap \Omega_y} p_i \cdot \log \frac{1}{p_i} \right] \\ &+ (1 + \alpha) \left[ \sum_{z_i \in \Omega_x \cap \Omega_y} q_i \cdot \log \frac{1}{s_i} - \sum_{z_i \in \Omega_x \cap \Omega_y} q_i \cdot \log \frac{1}{q_i} \right] \end{aligned} \tag{7}$$

따라서 신호원분리이득은

$$\begin{aligned} SG(Z) &= \alpha \log \frac{1}{\alpha} \Pr\{z_i \in \Omega_x - \Omega_y | X\} \\ &+ (1 - \alpha) \log \frac{1}{1 - \alpha} \Pr\{z_i \in \Omega_y - \Omega_x | Y\} \\ &+ \alpha \left[ \sum_{z_i \in \Omega_x \cap \Omega_y} p_i \cdot \log \frac{1}{s_i} - \sum_{z_i \in \Omega_x \cap \Omega_y} p_i \cdot \log \frac{1}{p_i} \right] \\ &+ (1 - \alpha) \left[ \sum_{z_i \in \Omega_x \cap \Omega_y} q_i \cdot \log \frac{1}{s_i} - \sum_{z_i \in \Omega_x \cap \Omega_y} q_i \cdot \log \frac{1}{q_i} \right] \end{aligned} \tag{8}$$

이 된다.

식 (8)로부터 알 수 있는 바와 같이 신호원분리이득은 기호 X와 Y에 대한 선택의 불확실성과 통계적인 특성이 다른 두 기호의 결합으로 인하여 발생하는 기호의 통계적인 확률분포에 있어서의 증가된 불확실성에 의해서 발생한다. 이 두가지 요인에 의한 효과를 확인하기 위하여 다음절에서 두가지 전형적인 경우에 있어서 신호원분리이득을 해석한다.

2. 서로 분리된 알파벳을 갖는 성분신호원을 갖는 복합신호원

Lemma 1: 2-성분 복합신호원에서 두 성분기호들이 서로 분리된 알파벳을 갖는 경우, 신호원분리이득은 0보다 크거나 같고, 최대 1 비트이다.

(증명) (8)식에  $\Omega_X \cap \Omega_Y = \emptyset$ 을 대입하면,

$$\begin{aligned} SG(Z) &= \alpha \log \frac{1}{\alpha} + (1 - \alpha) \log \frac{1}{1 - \alpha} \\ &= H(\alpha) \geq 0 \end{aligned} \tag{9}$$

따라서 신호원분리이득은 항상 0보다 크거나 같고, 그 값은 선택 확률  $\alpha$ 만의 함수이다. 정보이론에 의하면  $\alpha = \frac{1}{2}$  일 때 최대값 1을 갖는다. (증명끝)

3. 동일한 알파벳을 갖는 성분신호원을 갖는 복합신호원

Lemma 2: 2-성분 복합신호원에서 두 성분기호들이 서로 동일한 알파벳을 갖는 경우, 신호원분리이득은 0보다 크거나 같다.

(증명) 두 성분신호원의 알파벳이 동일하므로,

$$\Omega_Z = \Omega_X = \Omega_Y = \Omega_X \cup \Omega_Y = \Omega_X \cap \Omega_Y \tag{10}$$

이다.

(8)식과 (11)식으로부터, 신호원분리이득은

$$\begin{aligned} SG(Z) &= \alpha \left[ \sum_{z_i \in \Omega_z} p_i \log \frac{1}{s_i} - \sum_{z_i \in \Omega_z} p_i \cdot \log \frac{1}{p_i} \right] \\ &+ (1 - \alpha) \left[ \sum_{z_i \in \Omega_z} q_i \log \frac{1}{s_i} - \sum_{z_i \in \Omega_z} q_i \log \frac{1}{q_i} \right] \end{aligned} \tag{11}$$

이 된다. 이 경우 대수함수의 convexity 성질로부터

$$\sum p_i \cdot \log \frac{1}{s_i} - \sum p_i \log \frac{1}{p_i} \geq -\log(\sum S_i) = 0 \quad (12)$$

$$\sum q_i \cdot \log \frac{1}{s_i} - \sum q_i \cdot \log \frac{1}{q_i} \geq -\log(\sum S_i) = 0 \quad (13)$$

을 만족하므로, 신호원분리이득은 다음과 같은 식을 만족한다.

$$SG(Z) \geq -\log(\sum S_i) = 0 \quad (14)$$

단, 여기서 등식은  $s_i = p_i = q_i, \forall z_i \in Q_z$ 의 필요충분조건으로 성립한다. (증명끝)

이상의 두 성질로부터 우리는 다음과 같은 성질을 유추할 수 있다: 임의의 신호를 분포특성이 다른 두 신호원으로 분리하여 부호화하는 경우, 분리하지 않고 부호화하는 경우에 비하여 평균부호길이가 줄어든다. 특히 두 분포특성이 현격히 다르면 다를수록 평균부호길이가 더욱 짧아진다. 이러한 유추가 가능하기 위하여는, 평균부호길이가 신호의 엔트로피의 감소에 비례하여 감소한다는 가정이 필요한데, 본 논문에서 이용하고 있는 허프만 부호화는 이 가정을 충분히 만족하고 있다. 이는 제 IV장의 실험결과로부터 확인할 수 있을 것이다.

### III. 인덱스의존 (줄길이, 레벨) 부호화

#### 1. (줄길이, 레벨) 심볼의 인덱스

일반적인 DCT부호화에서 영상은 8 x 8의 중첩되지 않은 작은 블록들로 나뉘게 되고, 이 블록들은 DCT 변환되어 8 x 8의 DCT계수 블록으로 표현되며, 양자화과정을 거치게 된다. 이때 DCT계수를 양자화하면 일반적으로 대부분의 계수가 '0'이 되므로 이른바 (줄길이, 레벨) 부호화기법을 이용하게 된다. 이를 위해 양자화된 변환계수는 저주파성분부터 고주파성분까지에 이르는 지그재그 스캐닝을 통해 1차원으로 나열되며, '0'인 계수의 연속횟수를 줄길이라고 하고 뒤따르는 '0'이 아닌계수의 값을 레벨이라 할 때, 양자화된 변환계수는 (줄길이, 레벨) 심볼로 나타낼 수 있게 된다. 그리고 블록내에서 더이상 '0'이 아닌 계수가 없을 때 블록종료(EOB) 심볼을 달아 주게 되며, 이러한 심볼들은 미리 정하여진 부호책내에서 가변길이부호화하게 된다.

이제 (줄길이, 레벨) 심볼의 인덱스를 정의한다. 먼저 DCT블록에 있어 지그재그 스캐닝을 통해 그림 2와 같이 고유한 번호 ZSO(zigzag scanning order)를 할당한다. DCT 블록의 계수들을 이 ZSO

에 따라 1차원으로 배열할 때, 각 (줄길이, 레벨) 심볼의 스캐닝 시작점이 되는 ZSO, 즉 이전 '0'이 아닌 계수의 바로 다음 위치를 그 심볼의 인덱스라고 한다. 예를 들어 이전에 '0'이 아닌 계수가 ZSO n에 있었다면 현재 심볼의 인덱스는 (n+1)이 된다.

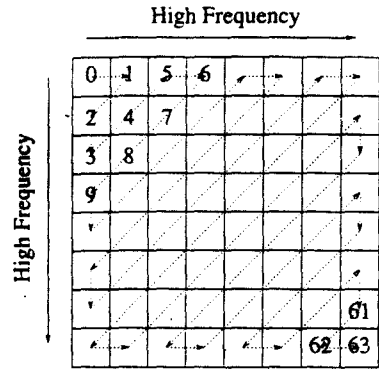


그림 2. DCT변환계수의 일반적인 지그재그 스캐닝  
Fig. 2. The conventional zigzag scanning of DCT coefficients.

(줄길이, 레벨)심볼의 인덱스를 MPEG이나 H.261등 일반적인 하이브리드 codec에서의 프레임내와 프레임간의 두가지 모드에 대해 설명하면 다음과 같다. 즉 프레임내 부호화의 경우, DC계수는 AC계수에 비해 매우 다른 통계적 특성을 갖고 있기 때문에 각각 분리하여 처리하며, 따라서 최초 스캐닝위치는 ZSO 1이 되고, 인덱스 범위는 1부터 63까지가 된다. 그리고 프레임간 부호화의 경우, DC계수는 AC계수와 비슷한 특성을 갖기 때문에 다른 AC와 같이 가변길이부호화하므로 최초 스캐닝위치는 ZSO 0이 되어, 인덱스 범위는 0부터 63까지가 된다.

#### 2. 인덱스에 따른 (줄길이, 레벨)심볼의 분포특성

양자화된 변환계수의 (줄길이, 레벨)을 입력심볼로 하는 허프만 부호화는 매우 우수한 특성을 나타내지만, 변환계수의 구조적 중복성을 더 잘 이용하여 그 성능을 더욱 향상시킬 수 있다. 즉, 부호화할 (줄길이, 레벨) 심볼의 분포특성이 그 심볼이 존재하는 인덱스에 따라 매우 다르다는 성질을 이용하여, 심볼의 발생위치에 따라 서로 다른 부호책을 사용하여 심볼을 부호화함으로써 부호화효율을 향상시킬 수 있다. 허프만 부호의 평균부호길이가 입력 심볼의 분포에 의해 결정된다는 점을 상기하면 이 점을 용이하게 이해할 수 있다.

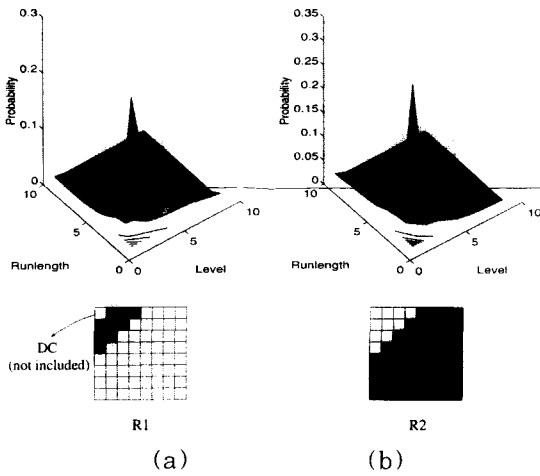


그림 3. 프레임내 DCT변환계수의(줄길이, 레벨) 심볼 분포특성. (a)저주파수 인덱스에서의 심볼 분포, (b)고주파수 인덱스에서의 심볼 분포

Fig. 3. The typical (runlength, level) distribution of DCT coefficients in the intraframe mode. (a)at low frequency index. (b)at high frequency index.

정지영상(intraframe mode)과 움직임보상잔류신호(interframe mode)에 대한 DCT계수의 특성은 현격히 다르다. 정지영상의 경우, 저주파수를 나타내는 인덱스에서의 (줄길이, 레벨) 심볼의 분포 특성은 짧은 줄길기와 넓은 범위의 레벨 분포로 나타나고, 고주파수를 나타내는 인덱스에서의 심볼은 비교적 긴 줄길기와 극히 좁은 범위의 레벨 분포를 갖는다.

한편, 움직임보상잔류신호를 변환한 경우에는, 저주파수를 나타내는 인덱스에서의 (줄길이, 레벨) 심볼의 분포특성은 비교적 다양한 줄길기와 좁은 범위의 레벨 분포를 갖고, 고주파수를 나타내는 인덱스에서의 심볼은 상당히 한정된 줄길기와 좁은 레벨 분포를 갖는다.

본 논문에서는 “Mobile and Calendar” 및 “Flower Garden”을 포함한 10가지 표준동영상을 사용하여 0.4bpp로 MPEG에 따른 부호화를 실시하므로써 프레임내와 프레임간 부호화시에 대해 각각 (줄길이, 레벨) 심볼의 분포특성을 구하였다. 그림 3과 그림 4는 각각 정지영상과 움직임보상잔류신호에 대한 DCT계수의 저주파수 인덱스와 고주파수 인덱스에서 (줄길이, 레벨) 심볼의 분포특성을 나타낸다.

3. 인덱스의존 (줄길이,레벨) 부호화

1) 인덱스의 Clustering

전술한 바와 같이 프레임내 부호화시에는 63개, 프레임간 부호화시에는 64개의 인덱스가 존재하므로 각각에 있어 최대 63, 64개로 (줄길이, 레벨) 심볼을 분리할 수 있다. 그러나 이 최대수치는 많은 부호책과 구현상의 복잡함으로 실제 적용키 어려우며, 따라서 주어진 분리갯수상에서 엔트로피감소가 최대가 되도록 하는 동일한 부호책을 이용하도록 인덱스를 clustering할 필요가 있다.

지그재그 스캔된 DCT계수 블록 R을 N영역으로 분리하면 식(15)와 같다.

$$R_k = \begin{cases} d_k \leq ZSO \leq d_{k+1} - 1 & k = 1, \dots, N - 1 \\ d_k \leq ZSO \leq d_{k+1} & k = N \end{cases} \quad (15)$$

여기서 프레임내 부호화시는  $d_1 = 1, d_{N+1} = 63$ , 프레임간 부호화시는  $d_1 = 0, d_{N+1} = 63$ 이다. 위의 정의로부터

$$\cup_i R_k = R \quad (16)$$

$$R_k \cap R_l = \emptyset, \quad k \neq l \quad \text{일때}$$

이다.

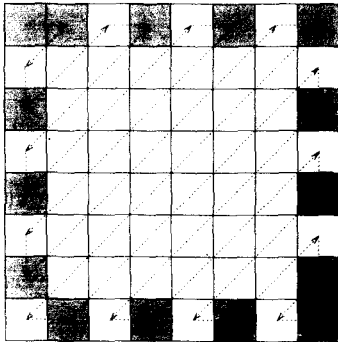
한편  $d_k$ 가  $d_1$ 부터  $d_{N+1}$ 까지의 모든 값을 가질 수 있지만, 그림 5와 같이 경계점(boundary points)사이의 DCT변환계수가 비슷한 특성을 가지며, 지그재그

그림 4. 프레임간 DCT변환계수의 (줄길이, 레벨) 심볼 분포특성. (a) 저주파수 인덱스에서의 심볼 분포, (b) 고주파수 인덱스에서의 심볼 분포

Fig. 4. The typical (runlength, level) distribution of DCT coefficients in the interframe mode. (a) at low frequency index. (b) at high frequency index.

스캔라인의 중간에서 끊어 분리할 필요가 없다면, 그림 5와 같이 표시된 경계점의 수로 영역을 분리할 수 있다. 따라서 그림 4에 따르면 최대 분리 갯수는 15가 된다.

여기에서 줄길이  $r$ 과 레벨  $l$  및 인덱스  $i$ 가 될 심볼의 확률을  $P_i(r,l)$ 라 할 때, 인덱스가 영역  $R_k$ 에 있는 (줄길이, 레벨) 심볼과 관련한 불규칙변수  $X_{R_k}$ 의 확률을 다음과 식(17)과 같이 나타낼 수 있다.




 Possible Boundary Points

그림 5. DCT변환계수의 분리시 가능한 경계점  
Fig. 5. Possible boundary points for separating of DCT coefficients in a block.

$$\text{Prob}\{X_{R_k} = (r,l)\} = \frac{\sum_{i \in R_k} pi(r,l)}{\alpha_{R_k}} \quad (17)$$

여기서  $\alpha_{R_k}$ 는 심볼이 영역  $R_k$ 에 있는 인덱스를 가질 확률이며, 식(18)과 같다.

$$\alpha_{R_k} = \sum_{i \in R_k} \sum_r \sum_l pi(r,l) \quad (18)$$

그리고 각 불규칙변수  $X_{R_k}$ 의 엔트로피를  $H(X_{R_k})$ 라 할 때, 분리 갯수  $N$ 에 대한 전체 엔트로피  $H_N(d_1, \dots, d_{N+1})$ 가 최소가 되도록 경계점  $d_k$ 를 구해야 한다.

$$H_N(d_1, \dots, d_{N+1}) = \sum_{k=1}^N \alpha_{R_k} H(X_{R_k}) \quad (19)$$

이제  $N$ 개로 분리하여 얻게 될 최소 엔트로피를  $H_N^*$ 이라 하면, 분리하지 않았을 때의 엔트로피  $H_1$ 은  $H_N^*$  ( $N > 1$ ) 보다 항상 크거나 같게 된다.  $d_1, \dots, d_{N+1}$ 의 경우 수가 많지 않으므로 모든 조합들에 대해 엔트로피를

구한 후 최소가 되는  $d_1, \dots, d_{N+1}$ 을 찾아 낼 수 있다.

### 2) 부호화 과정

변환계수를 표현하는 심볼의 분포특성이 인덱스에 따라 서로 다른 모양을 하고 있으므로, 각각의 특성에 정합된 부호책을 이용함으로써, 부호화의 효율을 높일 수 있다. 즉, 정합된 각각의 부호가 우수한 경우 각 부호의 평균부호길이는 각 부호의 입력 심볼의 엔트로피에 접근하게 되는데, 제 II장에서 살펴본 정리 1과 성질 1, 2로부터 알 수 있는 바와 같이, 신호원을 분리하여 부호화하는 경우 평균부호길이가 분리하지 않는 경우에 비하여 감소하게 된다.

그러나 신호원을 분리하기 위하여 오버헤드 비트가 추가된다면, 이 오버헤드 비트가 신호원 분리를 통하여 얻는 엔트로피의 감소 즉 평균부호길이의 감소를 상쇄하기 때문에 종합적인 이득이 없어진다. 따라서 오버헤드 정보없이 신호원 자체의 구조적인 중복성을 이용하여 신호원을 분리할 필요가 있다.

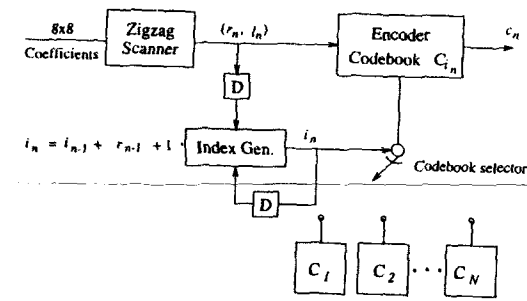
프레임내 변환된 변환계수와 프레임간 변환된 변환계수의 분포특성의 현격한 차이를 이용하기 위하여 원영상과 움직임보상잔류신호를 적응적으로 선택하여 변환하는 적응적 프레임내/프레임간 변환부호화는 MPEG등 대부분의 움직임보상부호화 시스템에서 이미 이용하고 있다.<sup>[4]</sup> 이 경우에는 부호화할 심볼의 분리를 위하여 새로운 오버헤드가 필요하지 않다. 즉 움직임 보상을 위하여 전송하는 움직임 정보를 이용하여 심볼의 분리가 가능하다.

본 연구에서는 프레임내/프레임간의 차이 뿐만 아니라, 신호의 저주파수 인덱스와 고주파수 인덱스에서의 (줄길이, 레벨) 심볼의 분포특성의 차이를 이용하기 위한 방법을 제안한다. 오버헤드를 추가하지 않고 주어진 심볼  $z_n$ 을 저주파수성분 또는 고주파수성분으로 분리하는 방법은, 바로 전에 부호화된 심볼의 주파수가 얼마인가에 따라서 그 심볼의 가상적인 신호원을 판정함으로써 이루어진다. 즉 각 구획내에서 심볼이 순차적으로 복호됨에 따라 심볼이 나타내는 인덱스를 알 수 있으므로, 직전에 복호된 심볼의 인덱스에 따라 현재 부호화할 심볼의 신호원을 판정하여, 그 판정된 신호원에 정합된 부호책을 이용하여 부호화한다.

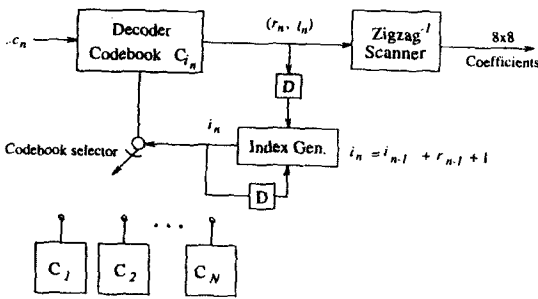
각각의 신호원에 정합된 부호책은 3.2절에서와 같이 여러 영상들을 실제 부호화하여 얻은 심볼의 통계적인 특성으로부터 구하였다. 즉 인덱스 clustering을 통하여 최적의 영역으로 분할한 후, 식 (17)의 확률 분포로부터 허프만부호책을 구성하였다.

### 3) 부호화기와 복호화기

제안된 방법의 부호화기와 복호화기의 기능구획도를 그림 6에 보인다.



(a)



(b)

$(r_n, l_n)$   $(r, l)$  symbol  
 $i_n$  index of  $(r_n, l_n)$   
 $c_n$  codeword assigned to  $(r_n, l_n)$   
 D delay element

그림 6. 제안된 기법의 부호화기와 복호화기.  
 (a) 부호화기, (b) 복호화기

Fig. 6. The encoder and decoder of the proposed scheme. (a) encoder, (b) decoder.

부호화기에서는 지그재그 스캔된 8x8 블록에서의 심볼  $(r_n, l_n)$ 을 발생시키고, 그 심볼의 인덱스  $i_n$ 은 그전 심볼의 인덱스  $i_{n-1}$ 과 줄길이  $r_{n-1}$ 을 이용하여 계산한다. 그리고 구해진 인덱스로 해당되는 부호책을 선택하게 되고 부호화된 심볼  $c_n$ 이 출력된다. 이전에 부호화된 심볼의 인덱스와 줄길이를 새로운  $(r, l)$  심볼을 부호화하기 전에 알 수 있으므로 제안된 기법에서는 부호책선택을 위해 복호기로 보내야 할 추가정보 비트가 필요하다.

복호화기는 부호화기의 역과정이며, 현재 입력된 심볼  $c_n$ 의 인덱스는 이전 심볼의 인덱스와 줄길이를 부터 계산되고, 이에 정합된 부호책을 선택하여 심볼

$(r_n, l_n)$ 을 얻게 된다. 그리고 이 심볼은 다음 심볼의 인덱스 계산과 부호책을 선택하는데 다시 사용된다.

한편 부호화기와 복호화기에서 필요한 부호책은 신호원의 분리 갯수만큼 필요하며, 따라서 분리 갯수의 증가에 따라 부호책의 수도 증가하게 된다. 그런데 부호책은 충분한 훈련열(training sequence)에 대해 미리 구할 수 있으며, 약속된 부호책이 부호화기와 복호화기양측의 일정한 메모리에 부호책 갯수만큼 저장되어 있는 것으로 충분하다.

IV. 실험 결과

본 논문에서는 MPEG에서 이용하는 표준동영상인 "Mobile and Calendar" 및 "Flower Garden" 등을 포함한 10 종류의 영상에 대해 (줄길이, 레벨) 심볼의 확률분포를 구하였고, 실험영상으로는 확률분포 발생에 포함시키지 않은 "Football"을 사용하였다. 즉 각각 10 프레임으로 구성된 10 종류의 704x480 휘도영상에 대해, 평균부호율 0.4 bpp로 움직임보상 DCT변환부호화를 실시하여 (줄길이, 레벨) 심볼의 확률분포를 구하였다. 그리고 프레임내/프레임간 변환의 선택은 매 구획마다 변환하고자 하는 신호의 전력의 작은 쪽을 선택하였다. 본 논문의 제 II장에서 살펴본 신호원 분리에 의한 신호의 엔트로피 감소효과를 확인하기 위하여, (줄길이, 레벨) 심볼을 가상의 신호원으로 분리하여 가면서 엔트로피의 감소 경향을 비교하였다. 또한 (줄길이, 레벨) 심볼을 부호화하는 방법으로 각 신호원마다 64개의 부호어를 갖는 변형된 2차원 허프만 부호화를 사용하는 경우의 평균부호길이를 비교하여 부호화효율의 개선도를 계산하였다.

한편 각 부호책은 64개의 부호어로 고정하여 구하였으므로 분리된 신호원에 대해 동일한 크기의 부호책을 갖게 되며, 따라서 신호원을 분리하지 않은 방법에 대해, 신호원 분리를 갯수만큼 부호책을 갖게 된다. 이 경우 한 부호책의 크기는 1 Kbits 미만이므로 본 논문에서 제시하는 바와 같이 하드웨어의 복잡도와 성능개선 효율을 감안하여, 프레임내 부호화시 14개, 프레임간 부호화시 15개를 적용할 때, 전체 부호책수는 29개이며, 따라서 전체 부호책의 크기는 30 Kbits내외가 된다. 이는 제안된 기법의 성능개선 효과와 추가정보 bit의 전송이 필요없음을 고려하면, 하드웨어 구현상 매우 단순한 문제이다.

그림 7은 프레임내 변환 모드에서 인덱스의존 (줄길이, 레벨) 부호화를 적용한 경우, 분리된 가상 신호원의 수에 따른 심볼의 엔트로피와 평균부호길이를

보여준다. 분리된 신호원의 수는 DC 성분을 제외한 최대 63개이다.

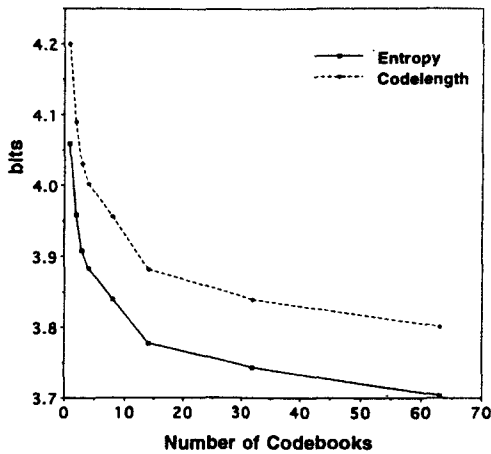


그림 7. 분리 신호원의 수에 따른 신호의 엔트로피와 평균부호길이(프레임내 변환)

Fig. 7. Entropy and average codelength according to the number of virtual sources. (intraframe mode)

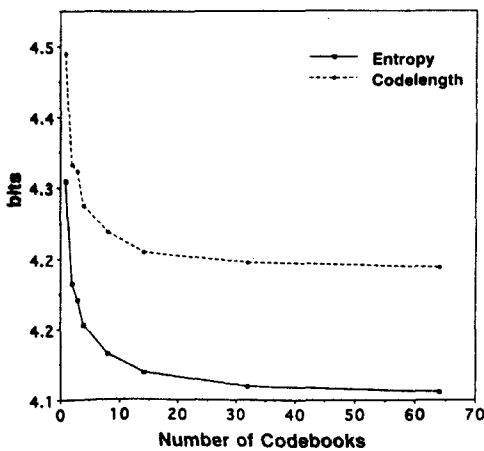


그림 8. 분리 신호원의 수에 따른 신호의 엔트로피와 평균부호길이(프레임간 변환)

Fig. 8. Entropy and average codelength according to the number of virtual sources. (interframe mode)

그림 8은 프레임간 변환 모드에서의 인덱스의존 (줄길이, 레벨) 부호화를 적용한 경우의 결과를 보여준다. 프레임간 모드에서는 DC를 AC와 동일하게 양자화하

므로, 최대 64개의 신호원으로의 분리가 가능하다.

그림 7과 그림 8로부터 알 수 있는 바와 같이 분리하는 신호원의 수를 늘림에 따라 엔트로피와 평균부호길이가 감소한다. 즉 신호원의 수를 14(15)개로 늘릴때까지는 엔트로피와 평균부호길이의 감소효과가 있으나, 그 이상으로 늘리는 경우 감소효과가 뚜렷하지 않다는 것을 알 수 있다. 따라서 프레임내 부호화시 14개, 프레임간 부호화시 15개로 신호원을 분리하는 것이 성능개선과 복잡도를 고려할 때, 적절한 최대수치임을 알 수 있다. 표 1에 신호원의 수가 2, 3, 4, 8, 14(15)개일때의 각 변환 모드에서의 엔트로피와 평균부호길이의 감소율을 나타내었다.

표 1. 분리되는 신호원 수에 따른 신호의 엔트로피 및 평균부호길이의 감소율

Table 1. Ratios of reduction of entropy and average codelength of symbols according to the number of source partitions.

신호원의 수	엔트로피 감소율 (%)		평균부호길이 감소율 (%)	
	프레임내	프레임간	프레임내	프레임간
2	2.49	3.33	2.62	3.46
3	3.75	3.86	4.01	3.66
4	4.36	4.70	4.69	4.73
8	5.43	5.63	5.80	5.52
14 또는 15	6.96	6.23	7.59	6.14

### V. 결론

본 논문에서는 변환부호화 시스템에서 많이 사용하고 있는 (줄길이, 레벨) 2차원 허프만 부호화의 성능을 향상시키기 위한, 신호원 분리의 개념과 신호원 분리에 의한 부호화율의 감축에 대하여 논하였다. 신호원을 분리함으로써, 분리하지 않는 경우에 비하여, 평균 엔트로피가 감소함을 해석적으로 증명하였으며, 이와 아울러 신호원 분리 기법을 엔트로피 부호화에 적용하는 경우 평균부호길이가 감소함을 여러가지 MPEG 동영상의 부호화를 통하여 실증하였다. 심볼을 분리하기 위한 가상 신호원의 수를 늘릴수록 부호화효율은 더욱 증가하지만, 가상 신호원 수의 증가에 따라 필요한 허프만 부호책의 수가 증가하고 이에 따라 시스템의 복잡도가 증가하므로, 분리할 신호원의 갯수에 대한 최적의 선택에 대한 연구가 필요하다. 또한 구획간의 상관성을 이용하기 위하여 여러개의 구획을 동시에 부호화하는 경우에 본 논문에서 제시



한 신호원 분리의 기법을 도입함으로써 부호화효율이 더욱 개선될 것으로 기대된다.

부 록

A. (7)식의 유도:

복합기호 Z에 대한 알파벳과 확률밀도함수를 나타내는 (1)식과 (2)식을 다시 써보면,

$$\Omega_z = \Omega_x \cup \Omega_y \tag{A-1}$$

$$\Pr(Z = z_i) = s_i = \begin{cases} \alpha p_i, & z_i \in \Omega_x - \Omega_y \\ \alpha p_i + (1 - \alpha)q_i, & z_i \in \Omega_x \cap \Omega_y \\ (1 - q_i), & z_i \in \Omega_y - \Omega_x \\ 0, & \text{otherwise} \end{cases} \tag{A-2}$$

인 관계를 만족한다.

이제, 이 확률분포로부터 불규칙변수 Z의 엔트로피를 구해보면,

$$\begin{aligned} H(Z) &= \sum_{i \in \Omega_z} s_i \cdot \log \frac{1}{s_i} \\ &= \left( \sum_{z_i \in \Omega_x - \Omega_y} + \sum_{z_i \in \Omega_x \cap \Omega_y} + \sum_{z_i \in \Omega_y - \Omega_x} \right) s_i \cdot \log \frac{1}{s_i} \end{aligned} \tag{A-3}$$

식(A-1)과 (A-2)를 대입하여 정리하면

$$\begin{aligned} H(Z) &= \sum_{z_i \in \Omega_x - \Omega_y} \alpha p_i \cdot \log \frac{1}{\alpha p_i} \\ &+ \sum_{z_i \in \Omega_x \cap \Omega_y} [\alpha p_i + (1 - \alpha)q_i] \cdot \log \frac{1}{s_i} \\ &+ \sum_{z_i \in \Omega_y - \Omega_x} (1 - \alpha) \cdot \log \frac{1}{(1 - \alpha) q_i} \end{aligned} \tag{A-4}$$

이 된다. 여기서 (A-4)식의 우변 제 1항은 다음과 같이 표현된다:

$$\begin{aligned} \text{제 1항} &= \sum_{z_i \in \Omega_x - \Omega_y} \alpha p_i \cdot \left[ \log \frac{1}{p_i} + \log \frac{1}{\alpha} \right] \\ &= (\alpha \log \frac{1}{\alpha}) \sum_{z_i \in \Omega_x - \Omega_y} p_i + \alpha \sum_{z_i \in \Omega_x - \Omega_y} p_i \cdot \log \frac{1}{p_i} \\ &= (\alpha \log \frac{1}{\alpha}) \Pr\{z_i \in \Omega_x - \Omega_y | X\} \\ &+ \alpha \left( \sum_{z_i \in \Omega_x} - \sum_{z_i \in \Omega_x \cap \Omega_y} \right) (p_i \cdot \log \frac{1}{p_i}) \\ &= (\alpha \log \frac{1}{\alpha}) \Pr\{z_i \in \Omega_x - \Omega_y | X\} \\ &+ \alpha H(X) - \alpha \sum_{z_i \in \Omega_x \cap \Omega_y} p_i \cdot \log \frac{1}{p_i} \end{aligned} \tag{A-5}$$

이 된다. 같은 방법으로, 식 (A-4)의 제 3항을 정리하면,

$$\begin{aligned} \text{제 3항} &= \left[ (1 - \alpha) \log \frac{1}{1 - \alpha} \right] \Pr\{z_i \in \Omega_y - \Omega_x | Y\} \\ &+ (1 - \alpha) H(Y) - (1 - \alpha) \sum_{z_i \in \Omega_x \cap \Omega_y} q_i \cdot \log \frac{1}{q_i} \end{aligned} \tag{A-6}$$

이 된다.

식 (A-5)와 (A-6)의 결과를 식 (A-4)에 대입하여 정리하면,

$$\begin{aligned} H(Z) &= \alpha H(X) + (1 - \alpha) H(Y) \\ &+ \alpha \log \frac{1}{\alpha} \Pr\{z_i \in \Omega_x - \Omega_y | X\} \\ &+ (1 - \alpha) \log \frac{1}{1 - \alpha} \Pr\{z_i \in \Omega_y - \Omega_x | Y\} \\ &+ \alpha \left[ \sum_{z_i \in \Omega_x \cap \Omega_y} p_i \cdot \log \frac{1}{s_i} - \sum_{z_i \in \Omega_x \cap \Omega_y} p_i \cdot \log \frac{1}{p_i} \right] \\ &+ (1 - \alpha) \left[ \sum_{z_i \in \Omega_x \cap \Omega_y} q_i \cdot \log \frac{1}{s_i} - \sum_{z_i \in \Omega_x \cap \Omega_y} q_i \cdot \log \frac{1}{q_i} \right] \end{aligned} \tag{A-7}$$

이 된다. (증명 끝)

參 考 文 獻

- [1] Jae S. LIM, *Two-dimensional Signal and Image Processing*, Prentice-Hall, pp.612-617, 1990.
- [2] M. Kunt, et al., "Recent results in high-compression image coding," *IEEE Trans. on CAS*, vol. CAS-34, no. 11, pp. 1306-1336, Nov. 1987.
- [3] W.B.Pennebaker and J.L.Mitchell, *JPEG Still Image Data Compression Standard*, Van Nostrand Reinhold, pp. 169-173, 1993.
- [4] Motion Picture Expert Group, "Coding of moving pictures and associated audio for digital storage media at up to about 1.5Mbit/s," *ISO-IEC/JTC1 SC29, Committee Draft 11172*, Nov. 1991.
- [5] N. Garcia, et al., "Optimization of HDTV codec algorithms," *Telettra Review* 45, pp. 67-82, 1990.
- [6] B. Macq, "A universal entropy coder for transform or hybrid coding," *PCS90*.

Boston, 1990.

[7] S. G. Lee, et al., "Entropy reduction by interleaving DCT coefficients of MCP error images," *International Workshop on HDTV '92*, vol. 1, pp.

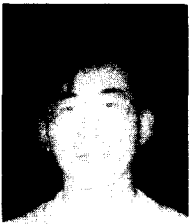
47.1-47.8, 1992.

[8] R.J.McEliece, *The Theory of Information and Coding*. Addison-Wesley Pub. Co. Inc., pp. 275-279, 1977.

著 者 紹 介

李 鍾 和(正會員) 第 27卷 第 2號 參照  
 현재 한국방송공사 재직중

姜 東 郁(正會員) 第 28卷 第 1號 參照  
 현재 서울대학교 전자공학과 박사  
 과정 재학중



姜 守 遠(準會員)  
 1971年 5月 17日 生. 1993年 서울  
 공대 전자공학과 졸업. 1994年 현  
 재 동 대학원 석사과정 재학중. 주  
 관심 분야는 영상신호처리 등임.



張 俊 鎬(準會員)  
 1968年 5月 10日 生. 1990年 서울  
 공대 전자공학과 졸업. 1992年 동  
 대학원 석사과정 졸업. 공학석사  
 학위 취득. 1994年 현재 동 대학  
 원 박사과정 재학중. 주관심 분야  
 는 영상신호처리 및 컴퓨터비전  
 등임.

李 忠 雄(正會員) 第 28卷 第 1號 參照  
 현재 서울대학교 전자공학과 교수.  
 서울대학교부설 뉴미디어통신 공동  
 연구소 소장