

# 선택적 비트 플레인 부호화를 이용한 오디오 주파수 계수의 무손실 부호화 기술

## Lossless Coding of Audio Spectral Coefficients Using Selective Bit-Plane Coding

유 승 관\*, 박 호 종\*, 오 승 준\*, 안 창 범\*, 심 동 규\*, 백 승 권\*\*, 강 경 옥\*\*  
(Seung-Kwan Yoo\*, Hochong Park\*, Seung-Jun Oh\*, Chang-Beom Ahn\*,  
Donggyu Sim\*, Seung-Kwon Beak\*\*, Kyoung-Ok Kang\*\*)

\*광운대학교 VIA 멀티미디어 센터, \*\*ETRI 방송미디어 연구그룹  
(접수일자: 2007년 12월 6일; 채택일자: 2008년 1월 10일)

본 논문에서는 오디오 부호화기의 양자화 된 주파수 계수 인덱스를 무손실 부호화 하는 새로운 방법을 제안한다. 기존의 무손실 부호화기는 계수의 통계적 특성을 이용하여 허프만 부호화 하고 있으나 간단한 구조로 인하여 최고의 성능을 제공하지 못한다. 본 논문에서는 비트 플레인 변환과 런LENGTH 부호화 기술을 사용하여 성능이 향상된 새로운 주파수 계수의 무손실 부호화 기술을 제안한다. 제안하는 방법은 주파수 계수의 양자화 인덱스를 비트 플레인 변환을 거쳐 상관도가 높은 1차원 비트열로 변환하고 이를 런LENGTH 부호화 하고 최종적으로 런LENGTH 심볼을 허프만 부호화 한다. 또한 주파수 대역을 크게 3개의 구역으로 나누고 각 구역에서 비트 플레인 부호화 방법을 선택적으로 사용하도록 하여 성능을 추가로 향상시킨다. 제안한 무손실 부호화 성능을 엔트로피에 의한 이론적 비트 수로 측정하면 기존의 AAC 무손실 부호화에 비하여 최대 6% 성능 향상을 제공한다.

**핵심용어:** 오디오 부호화기, 무손실 부호화, 비트 플레인 부호화, 허프만 부호화

**투고분야:** 뉴미디어 분야 (13.3)

In this paper, new lossless coding method of spectral coefficients for audio codec is proposed. Conventional lossless coder uses Huffman coding utilizing the statistical characteristics of spectral coefficients, but does not provide the high coding efficiency due to its simple structure. To solve this limitation, new lossless coding scheme with better performance is proposed that consists of bit-plane transform and run-length coding. In the proposed scheme, the spectral coefficients are first transformed by bit-plane into 1-D bit-stream with better correlative properties, which is then coded into run-length and is finally Huffman coded. In addition, the coding performance is further increased by applying the proposed bit-plane coding selectively to each group, after the entire frequency is divided into 3 groups. The performance of proposed coding scheme is measured in terms of theoretical number of bits based on the entropy, and shows at most 6% enhancement compared to that of conventional lossless coder used in AAC audio codec.

**Key words:** Audio Codec, Lossless Coding, Bit Plane Coding, Huffman Coding

**ASK subject classification:** New Media (13.3)

## I. 서론

오디오 신호를 부호화 하는 대표적 구조는 변환 부호화 (transform coding) 방식이며, 입력 신호를 주파수

영역으로 변환하고 주파수 계수를 심리음향 모델을 기반으로 양자화 하고, 양자화 계수를 무손실 부호화 (lossless coding) 한다 [1-6]. 무손실 부호화는 엔트로피 (entropy) 부호화 원리를 기반으로 양자화 인덱스에 해당하는 입력 심볼 (symbol)의 발생 확률을 고려하여 평균적으로 최소의 비트로 부호화 하는 것을 목표로 한다. 가장 널리 사용되는 방법은 심볼 또는 심볼 벡터마

책임저자: 박 호 종 (hcpark@kw.ac.kr)  
139-701 서울 노원구 월계동 447-1 광운대학교 전자공학과  
(전화: 02-940-5104; 팩스: 02-913-9057)

다 부호화 비트 패턴을 미리 할당하는 허프만 (Huffman) 부호화와, 각 심볼을 연속적으로 처리하면서 하나의 비트 패턴을 출력하는 Arithmetic 부호화 방식이 있으며, 기존의 비디오, 영상, 오디오 표준 부호화기에 널리 사용되고 있다 [7].

무손실 부호화의 궁극적 목표는 적은 수의 비트를 사용하여 해당 심볼을 부호화 하는 것이다. 그러나 가장 대표적인 오디오 부호화기인 MPEG-4 AAC (advanced audio coding)에서 사용하는 방법은 주어진 심볼을 그대로 부호화 하는 매우 기본적인 형태를 사용하여 오디오 신호의 통계적 특성 및 변화 형태를 최대한 활용하지 못하고, 그에 따라 최고의 무손실 부호화 성능을 제공하지 못한다 [1, 2]. 따라서 추가적인 처리를 통하여 무손실 부호화의 성능을 향상시킬 수 있는 가능성이 있으며, 본 논문에서는 새로운 무손실 부호화 방법을 통하여 AAC의 무손실 부호화 방법에 비하여 우수한 성능의 부호화를 제공하도록 한다. 동일한 주파수 계수에 대하여 보다 적은 비트를 사용하는 무손실 부호화를 제공하면 전체 부호화 비트 수의 감소를 얻고 오디오 부호화기의 성능을 향상시키게 된다.

제안하는 부호화 방법의 기본 구조는 최초 입력 심볼에 복원 가능한 변환을 적용하여 부호화에 보다 유리한 통계적 특성을 가지는 “새로운 심볼” 을 생성하고 이 심볼을 무손실 부호화하는 것이다. 구체적으로 본 논문에서는 비트 플레인 (bit plane) 변환과 런LENGTH (run length) 부호화를 통하여 새로운 심볼에 해당하는 런LENGTH를 구하고, 이를 최종적으로 무손실 부호화 한다. 또한, 각 주파수 대역별로 비트 플레인 변환에 의한 부호화 성능을 분석하고, 이를 기반으로 선택적 부호화를 통하여 전체 무손실 부호화 성능을 향상시키도록 한다. 제안하는 부호화 방식의 성능을 기존의 표준 방법과 비교하기 위하여 본 논문에서는 AAC 부호화기의 동작을 가정하고, AAC에서 정한 양자화 규격에 따라 주파수 계수를 양자화 하여 얻은 양자화 인덱스를 입력 심볼로 정의하고 제안하는 무손실 부호화기에 적용하여 성능을 측정하고 기존 방법의 성능과 비교한다.

본 논문의 구성은 다음과 같다. 2장에서 제안하는 비트 플레인 부호화 방법에 대하여 설명하고, 3장에서는 제안하는 방법을 선택적으로 사용하는 방법을 설명하고, 마지막으로 4장에서는 제안하는 부호화의 성능을 AAC 허프만 부호화의 성능과 비교하고, 향후 성능 개선을 위한 추가 연구 내용을 설명한다.

## II. 비트 플레인 부호화

### 2.1. 개요

그림 1은 일반적인 무손실 부호화 과정을 보여준다. 기존의 방법은 그림 1 (a)와 같이 입력 심볼을 직접 허프만 부호화 하며, 심볼을 다차원 벡터로 구성하여 부호화 하거나 여러 개의 허프만 테이블을 제공하여 선택적 부호화를 할 수 있다. 반면, 본 논문에서 제안하는 부호화는 그림 1 (b)와 같이 입력 심볼을 직접 허프만 부호화 하는 것이 아니라 복원 가능한 변환 과정을 거쳐서 입력 심볼들의 전체적인 상관도가 변하는 새로운 심볼로 변환한 후에 허프만 부호화 하는 구조이다. 이렇게 함으로써 직접 부호화 하는 것과 다른 통계적 특성을 활용하게 되며, 만일 심볼 변환을 통하여 부호화에 유리한 통계적 특성이 나타난다면 부호화 성능은 향상될 수 있다.

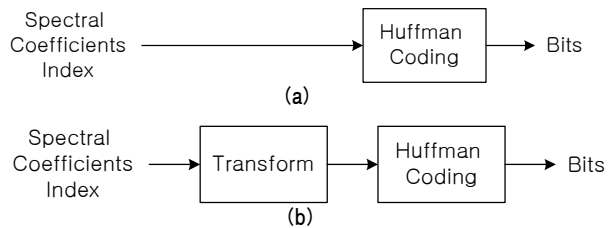


그림 1. 주파수 계수의 무손실 부호화 과정. (a) 기존 방법. (b) 제안하는 방법  
Fig. 1. Lossless coding of spectral coefficients (a) Conventional method (b) Proposed method.

본 논문에서는 그림 1 (b)의 변환 과정에서 비트 플레인 변환을 사용한다. 단, 본 논문에서 제안하는 비트 플레인 변환 및 부호화는 일반적으로 오디오 또는 영상 부호화에서 많이 사용되는 비트 단위의 비트 플레인 방식과는 구조가 다르다. 기존의 비트 플레인 부호화는 그림 2 (b)와 같이 각 심볼을 MSB-LSB의 비트로 표현하고 MSB부터 시작하여 LSB까지 각각 비트 플레인을 생성하며, MSB 비트 플레인 부터 차례대로 부호화 하여 전달하면 비트 중요도에 따라 정보가 전달되므로 Scalable 기능이 제공된다. 이 방법을 사용하는 가장 대표적인 오디오 부호화기가 MPEG-4 BSAC(bit-sliced arithmetic coding)이다 [2].

반면, 본 논문에서 제안 하는 비트 플레인 부호화는 각 심볼의 크기를 레벨 (level)별로 분해하여 비트 플레인을 정의한다. 구현 방법의 예는 그림 2 (c)에 주어진다. 입력 심볼을 레벨에 따라 각각 비트 플레인으로 변환하고 각 비트 플레인의 비트를 레벨을 증가시키면서 연결하여 1차원 비트 스트림을 만들고, 비트 정보를 최

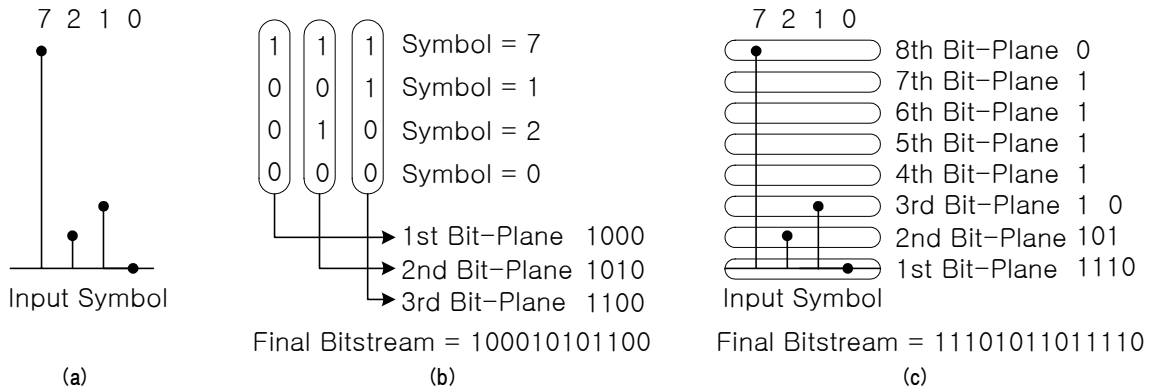


그림 2. 비트 플레인 부호화의 예 (a) 입력 심볼 (b) 기존의 비트 플레인 부호화 (c) 제안하는 비트 플레인 부호화  
 Fig. 2. Example of bit-plane coding (a) Input symbols (b) Conventional bit-plane coding (c) Proposed bit-plane coding.

종적으로 부호화 한다. 이때, 필요 없는 비트 정보는 제외하고 비트 스트림을 생성한다.

제안하는 비트 플레인 부호화 과정을 구체적으로 설명하면 다음과 같다. 현재 프레임에 해당하는  $N$ 개의 주파수 계수 양자화 인덱스에 절대값 취한 값을  $x[n](0 \leq n < N)$ 라 하고, 식(1)에 따라  $x[n]$ 을 영 레벨 ( $k=0$ )부터 시작하여 각 레벨  $k$ 에 대한 비트 플레인 값  $Q_k[n]$ 을 구한다.

$$Q_k[n] = \begin{cases} 1, & \text{if } x[n] > k \\ 0, & \text{if } x[n] = k \\ null, & \text{else} \end{cases} \quad (1)$$

예로, 그림 2에 대한  $Q_k[n]$ 을 구하면 다음과 같다.

$$\begin{aligned} Q_7[n] &= \{0, \quad , \quad , \quad \}, \\ Q_3[n] &= Q_4[n] = Q_5[n] = Q_6[n] = \{1, \quad , \quad , \quad \}, \\ Q_2[n] &= \{1, \quad , 0, \quad \}, \quad Q_1[n] = \{1, 0, 1, \quad \}, \\ Q_0[n] &= \{1, 1, 1, 0\} \end{aligned}$$

이때, null 값은 제외하고 표시 하였다.  $Q_k[n]$ 에서 null 값을 가지는  $n$  위치는 복원 과정에서 스택 (stack) 성질에 의하여 추가 정보 없이 정해질 수 있으므로 부호화에서 제외된다. 또한 각  $Q_k[n]$ 의 레벨값  $k$ 는 별도로 전송할 필요가 없고  $k=0$  부터 시작하여  $k$ 를 증가시키면서  $Q_k[n]$ 의 비트를 1차원으로 나열하여 새로운 비트 열  $B[i]$ 를 정의하여 전송하면 필요한 모든 정보가 전달된다. 만일 특정 레벨에서 비트가 모두 0이면 그 이상의 레벨은 필요가 없으므로 자동으로 비트 생성이 종료되며, 따라서 그림 2에 대한  $B[i]$  값은  $B[i] = \{1,1,1,0,1,0,1,1,0,1,1,1,1,0\}$ 이 된다.

복원기에서  $B[i]$ 로부터  $x[n]$ 을 복원하는 과정은 다음과 같다.  $k=0$  부터 시작하여  $k$ 를 증가시키면서 진행하며, 만일 특정  $n = n_0$ 에서  $Q_p[n_0]=0$  이면  $r > p$  인 모든 레벨  $r$ 에 대하여  $Q_r[n_0] = null$ 이다. 따라서 식 (2)에 따라  $B[i]$ 로부터  $Q_k[n]$ 을 복원한다.

$$Q_k[n] = \begin{cases} B[i++], & \text{if not null} \\ 0, & \text{else} \end{cases} \quad (2)$$

다음, 식 (3)에 따라  $Q_k[n]$ 로부터  $x[n]$ 을 복원한다.

$$x[n] = \sum_{k=0} Q_k[n] \quad (3)$$

이와 같이 정의된 비트열  $B[i]$ 를 전송하면 계수 인덱스 절대값  $x[n]$ 의 복원이 가능하다. 특히, 이 과정에서  $B[i]$ 에는 일반적으로 동일한 비트 값이 연속적으로 발생하게 되며 연속적으로 발생하는 동일한 비트열을 런 (run)이라 하고 “1”의 런과 “0”의 런이 반복적으로 발생한다. 이 경우 각 런의 길이 (length)만 전달하면 정보의 손실 없이 복원이 가능하다. 예로 그림 2의 심볼에 대하여 다음과 같은  $B[i]$ 의 런 길이 정보만 전송하면 된다.

$$\text{런 길이 정보} = \{3, 1, 1, 1, 2, 1, 4, 1\}$$

최종적으로 런 길이 정보를 허프만 부호화 하여 최종 비트 열을 생성하면 무손실 부호화 과정이 종료된다. 이와 같이 최초 심볼을 그대로 부호화 하지 않고 각 레벨 별로 비트 플레인을 구하고 1차원 비트 열로 나열한 후 런랭스를 최종 심볼로 출력함으로써 최초 심볼과 다른

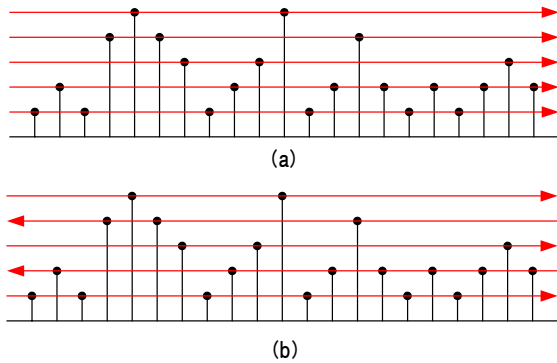


그림 3. 비트 플레인 부호화의 스캔 방향 (a) 일반적인 비트 스캔 (b) 연속적 비트 스캔  
 Fig. 3. Scan directions for bit-plane coding. (a) Conventional bit scan. (b) Continuous bit scan.

통계적 특성의 새로운 심볼을 얻게 되고, 만일 새로운 심볼의 통계적 특성이 부호화에 유리하게 나타나게 되면 기존 방법에 비하여 부호화 성능을 향상시킬 수 있다.

비트 플레인 부호화에서 모든 심볼은 양수가 되어야 스택 성질을 활용할 수 있으므로 주파수 계수 양자화 인덱스는 모두 절대값을 취한 다음 비트 플레인 변환을 한다. 따라서 사인 (sign) 정보는 독립적으로 부호화 하며, 각 Non-Zero 인덱스의 사인을 1 비트 부호화 한다.

비트 플레인 변환에서 레벨별로 만든 비트 플레인을 1 차원으로 연결할 때 일반적으로 그림 3 (a)와 같이 모든 레벨에서 동일한 방향으로 비트 스캔을 한다. 그러나 심볼의 연속적 연결 구조를 고려하면 그림 3 (b)와 같이 심볼이 연속적으로 연결되는 구조로 비트 스캔 하는 것이 부호화에 좀 더 효율적이다. 실험을 통하여 연속 비트 스캔이 약간 우수한 부호화 성능을 제공하는 것을 확인 하였고, 본 논문에서는 연속 비트 스캔을 기준으로 비트 플레인 부호화를 진행한다.

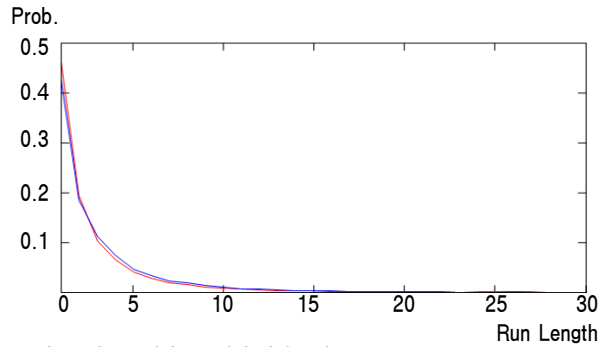


그림 4. 비트 플레인 부호화의 런레스 확률 분포  
 Fig. 4. Histogram of run-length of bit-plane coding.

### 2.2. 비트 플레인 부호화의 이론적 성능

비트 플레인 부호화를 통해 런레스 열이 생성 되면 런레스의 통계적 특성을 이용하여 허프만 부호화를 하게 된다. 그림 4는 매우 상이한 성질을 가지는 두 개의 오디오 신호 Input#1과 Input#2에 대하여 각 런레스의 확률 분포를 동시에 나타낸 것이다. 그림에서 보듯이 두 신호의 주파수 특성 차이에도 불구하고 런레스의 통계적 특성에는 큰 차이가 없으며, 모두 런레스 = 1의 발생 확률이 0.5에 근접하여 매우 높고, 런 길이가 길어지면 발생 확률이 점점 감소하는 것을 알 수 있고, 이것은 허프만 부호화하기에 매우 유리한 성질이다.

최종 런레스 부호화 과정은 위와 같은 확률적 특성을 기반으로 허프만 테이블을 만들어서 비트 스트림을 구하게 되며, 먼저 엔트로피를 기반으로 이론적인 비트 수를 구하여 성능을 분석해 보기로 한다. 엔트로피에 따른 이론적인 비트수를 구하기 위하여 발생하는 각 런레스 값  $i$ 에 대하여 발생 확률  $p_i$ 를 구하고,  $\sum_i p_i \log_2 \left( \frac{1}{p_i} \right)$ 으로부터 평균 엔트로피를 구한다. 다음, 평균 엔트로피와 발

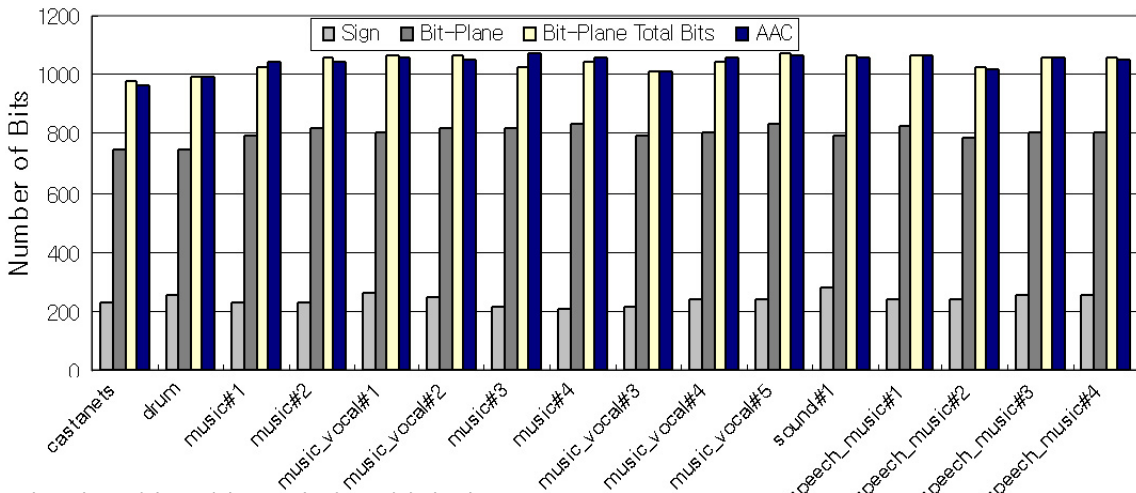


그림 5. 비트 플레인 부호화와 AAC 허프만 부호화의 비트 수  
 Fig. 5. The number of bits for bit-plane coding and AAC Huffman coding.

표 1. 엔트로피 기반으로 계산한 비트 플레인 부호화의 비트 수(프레임 평균)  
 Table 1. Number of bits for bit-plane coding based on entropy(frame average).

	# of Sign Bits(A)	Entropy (B)	# of Runs(C)	# of Bit-Plane Bits(D=B×C)	# of Total Bits(E=A+D)	# of AAC Huffman Coding Bits(F)	Bit Ratio(E/F)
Input#1	207.6	2.989	276.9	827.7	1035.3	1079.9	95.9%
Input#2	228.8	2.861	299.9	858.1	1086.9	1051.6	103.4%

생하는 총 런 수를 곱하여 최종 비트 플레인 비트 수를 구하고 마지막으로 사인 비트를 더하여 최종 부호화 비트 수를 구한다. Input#1과 Input#2에 대하여 이와 같이 구한 이상적인 비트 수를 프레임 평균으로 구하면 표 1과 같다. 또한, 동일한 주파수 계수에 대하여 AAC 허프만 부호화에 필요한 비트 수를 구하였다.

표 1로부터 Input#1에서는 비트 플레인 부호화의 성능이 AAC 허프만 부호화보다 우수하고, Input#2에서는 AAC 허프만 부호화 보다 성능이 약간 저하되는 것을 알 수 있다. 그림 5는 다양한 특성의 입력 신호에 대하여 동일한 실험을 진행하여 얻은 결과이다. 그림에서 보면 각 입력 신호의 특성에 따라 부호화 성능에 많이 차이가 나타나고, Input#1은 비트 플레인 부호화의 성능이 매우 우수한 경우이고, Input#2는 AAC 허프만 부호화의 성능이 매우 우수한 대표적인 경우이다.

### III. 선택적 비트 플레인 부호화

#### 3.1. 개요

오디오 신호의 주파수 특성은 오디오 신호의 종류에 따라 많은 차이를 보이며, 오디오 신호 내에서도 프레임 별로 매우 변화가 심하게 나타난다. 또한 일반적인 오디오 신호의 주파수 계수는 주파수 대역에 따라 특성의 변화가 나타난다. 무손실 부호화의 성능을 향상시키기 위하여 심볼의 특성에 적합한 통계 모델을 사용하여야 하고, 주어진 심볼과 통계 모델이 맞지 않으면 부호화 성능이 급격히 저하된다. 따라서 상이한 성질을 가지고 있는 심볼들이 혼합하여 나타나는 경우에 부호화 성능을 향상시키기 위하여 선택적 부호화 개념이 필요하다. 예로, 기존의 AAC 허프만 부호화 방법은 다수의 통계 모델에 따라 다수의 허프만 테이블을 만들어 최적의 테이블을 선택하여 사용한다.

그러나 이 방법은 통계 모델만 선택할 뿐이고 부호화 방법에는 차별을 하지 않아 아직 선택적 부호화 범위에 한계가 있다. 본 논문에서 제안하는 선택적 부호화 방법은 서로 다른 부호화 방법 중에서 각 심볼 그룹에 가장

적합한 부호화 방법을 선택하여 선택의 범위를 넓히고 성능 향상을 얻도록 한다. 즉, 주어진 심볼에 대하여 부호화 방법을 선택하고, 선택된 부호화 방법에서 다시 적절한 통계 모델을 선택하여 부호화 하면 단순히 통계 모델만을 선택하는 방법에 비하여 부호화 성능을 향상시킬 수 있을 것이다.

두 가지의 부호화 방법 중에서 선택하는 경우, 두 부호화 방법이 근본적인 차이를 가지고 있고 그에 따라 한 가지 방법에서는 성능이 저하되지만 다른 방법으로 부호화 하면 성능이 향상되는 효과를 얻어야 한다. 예로, 허프만 부호화와 Arithmetic 부호화 중에서 선택을 하면 두 방법이 매우 유사하므로 실질적인 선택의 효과는 크지 않다. 따라서 본 논문에서는 기존의 허프만 부호화 방법과 제안하는 비트 플레인 부호화를 두 부호화 방법으로 제공하고, 이 중에서 최종 부호화 방법을 선택하도록 한다. 아래 3가지 방법의 선택적 부호화의 성능을 분석하였고, 가장 성능이 뛰어난 섹션 그룹 기반의 선택적 부호화를 사용한다.

#### 3.2. 프레임 기반 선택적 부호화

오디오 신호의 특성이 프레임 단위로 변하므로 프레임마다 기존의 허프만 부호화 방식과 비트 플레인 부호화 방식 중 더 좋은 성능을 가지는 것을 선택하는 방법을 사용하면 성능 향상을 얻을 수 있다. 표 2는 프레임 별로 부호화 방법을 선택할 경우의 평균 성능을 보여주며, Input#1의 경우 전체 프레임의 약 81.3%에 대하여 비트 플레인 부호화 방식의 성능이 더 우수하며, 프레임 기반 선택 부호화를 통하여 4.6%의 비트 감소를 얻을 수 있다.

표 2. 프레임별 부호화 방법 선택에 의한 성능  
 Table 2. Coding performance of frame-based selective coding.

	Input #1	Input #2
# of AAC Huffman Coding Bits(A)	1079.9	1051.6
Frame Selection Ratio	Huffman Selected	18.7%
	Bit-Plane Selected	81.3%
# of Frame-Based Selective Coding Bits(B)	1030.0	1045.7
Bit Ratio(B/A)	95.4%	99.4%

### 3.3. 스케일 팩터 밴드 기반 선택적 부호화

오디오 신호의 특성이 프레임 단위로 변하므로 프레임 단위의 선택이 효과를 보이지만, 프레임의 성능을 더 세분화 하여 분석하면 프레임 내에서도 대역별로 부호화 방법에 따라 성능의 변화가 심하다. 그림 6은 오디오 신호의 한 프레임에서 기존의 AAC 허프만 부호화와 비트 플레인 부호화에 필요한 비트 수를 AAC 스케일 팩터 (scalefactor) 밴드 단위로 나타낸 것이며, 이로부터 스케일팩터 밴드 단위로 선택적 부호화를 하면 성능 향상이 생길 것이라는 것을 알 수 있다.

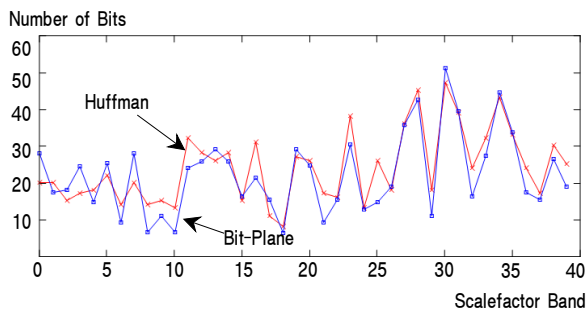


그림 6. 각 스케일 팩터 밴드별 부호화 비트수  
Fig. 6. The number of coding bits for each scalefactor band.

그러나 모든 스케일 팩터 밴드 단위로 선택적 부호화를 하게 되면 오버헤드 (overhead) 비트의 부담이 커지게 된다. 더욱이, 비트 플레인 부호화는 각 심볼의 값을 그대로 부호화 하는 것이 아니라 심볼의 진행 방향으로 비트 스캔 하여 비트 플레인을 구하므로 심볼의 진행 특성이 중요한 정보가 되고, 이와 같은 차이점이 부호화 선택에서 중요한 요인으로 작용한다. 만일, 비트 플레인 부호화를 작은 밴드 단위로 적용하면 심볼의 진행 흐름을 충분히 반영하기 어렵고 이에 따라 심볼들의 상관도가 전체적으로 낮아지게 되어 비트 플레인의 효율이 낮아지게 된다. 따라서 비트 플레인 부호화가 효력을 가지려면 스케일팩터 단위 보다 좀 더 큰 단위로 부호화하여 비트 플레인 부호화의 이점을 가질 수 있도록 하여야 한다.

### 3.4. 섹션 그룹 기반 선택적 부호화

따라서, 본 논문에서는 스케일팩터 밴드 단위로 선택하지 않고 섹션 (section) 그룹 기반으로 선택적 부호화한다. 먼저, 입력 심볼에 대하여 AAC 허프만 부호화 할 경우에 사용하는 섹션 정보를 얻은 후, 그림 7과 같이 한 프레임의 전체 6개 섹션을 3개의 그룹으로 분리하고 앞의  $p$  개 섹션과 뒤의  $q$  개 섹션은 AAC 허프만 부호화를

표 3. 제안하는 섹션 그룹 기반의 선택적 부호화 모드

Table 3. Proposed selective coding modes based on section groups.

Mode Index	Sections with AAC Huffman Coding	Mode Index	Sections with AAC Huffman Coding
0	None	8	0, 1, N-1, N-2
1	N-1	9	0, 1, 2
2	N-1, N-2	10	0, 1, 2, N-1
3	0	11	0, 1, 2, N-1, N-2
4	0, N-1	12	0, 1, 2, 3
5	0, N-1, N-2	13	0, 1, 2, 3, N-1
6	0, 1	14	0, 1, 2, 3, N-1, N-2
7	0, 1, N-1 (All for Mode8 Case)	15	All

#0	#1	#2	#3	#4	#5
Huffman Coding		Bit-Plane Coding			Huffman Coding

그림 7. 섹션 그룹 기반의 부호화 선택 구조

Fig. 7. Structure of coder selection based on section groups.

적용하고, 가운데 나머지 섹션들은 비트 플레인 부호화를 적용하여 최적의 성능을 가지는  $p$ 와  $q$  조합을 검색한다. 이 때  $0 \leq p \leq P$ ,  $0 \leq q \leq Q$  로 제한하면 최대  $(P+1) \times (Q+1)$  가지의 조합 중에서 최적의 한 조합을 선택하게 된다. 이와 같이 하면 비트 플레인 부호화가 적용되는 가운데 섹션 그룹들의 범위가 비트 플레인 부호화의 특징을 충분히 살릴 수 있을 만큼의 크기가 된다. 또한, 스케일팩터 밴드를 섹션 단위로 결합하므로 AAC 허프만 부호화에서의 허프만 테이블 선택도 자동적으로 최적화 된다.

본 논문에서는 표 3과 같이 8가지 조합 (Mode8)과 16가지 조합 (Mode16)을 가지는 섹션 그룹 기반의 선택적 부호화 모드를 제안한다. 이 때, Mode8에서 Mode Index = 7, Mode16에서 Mode Index = 15는 전체 섹션을 AAC 허프만 부호화 하는 경우를 의미한다. 예로, Mode Index = 8은 프레임에서 첫 섹션과 두 번째 섹션, 그리고 제일 마지막 2개 섹션이 AAC 허프만 부호화 되고 나머지 섹션들은 비트 플레인 부호화 된다. 이 방식을 구현하기 위해서는 Mode Index 정보 전송을 위하여 3 또는 4 비트의 추가 비트가 필요하지만 전체 부호화 비트 수는 감소하게 되어 우수한 부호화 성능을 낼 수 있다.

## IV. 성능 평가

본 논문에서 제안하는 섹션 그룹 기반의 선택적 비트 플레인 부호화 방법의 성능을 AAC 허프만 부호화의 성

표 4. 섹션 그룹 기반 선택적 부호화의 성능

Table 4. Performance of section group-based selective coding.

		Input #1		Input #2	
		Mode8	Mode16	Mode8	Mode16
# of AAC Huffman Coding Bits(A)		1079.9		1051.6	
# of Proposed Coding Bits	Ideal Number of Bits(B)	1019.1	1015.1	1036.0	1032.4
	Ideal Bit Ratio(B/A)	94.4%	94.0%	98.5%	98.2%
	Real Number of Bits(C)	1045.8	1046.1	1049.4	1048.3
	Real Bit Ratio(C/A)	96.8%	96.9%	99.8%	99.7%

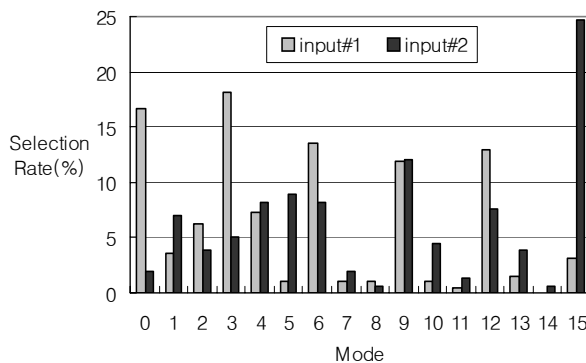


그림 8. Mode16에서 각 모드가 선택된 비율

Fig. 8. Mode histogram for Mode16.

능과 비교하여 정리하면 표 4와 같다. 대표적으로 입력 신호 Input#1과 Input#2에 대한 것만 정리하였으며, 엔트로피 기반으로 계산한 이론적 비트 수와 기존의 AAC 허프만 부호화의 비트 수를 비교하면 제안한 방법의 성능이 우수한 것을 알 수 있다. 예로, Input#1에 대하여 Mode16을 적용하면 6%의 비트 감소를 얻을 수 있다.

그림 8은 Input#1과 Input#2에 대하여 Mode16에서 각 모드가 선택되는 비율을 보여준다. 입력의 성질에 따라 선택되는 모드 분포에 매우 큰 차이를 가지고, 이를 통하여 모드 선택을 통하여 매우 효과적으로 최적의 선택적 부호화가 진행되는 것을 알 수 있다.

이상의 비트 수는 엔트로피를 기반으로 이론적으로 구한 성능이며, 구체적인 런LENGTH 허프만 부호화를 실시할 때의 실제 비트 수를 구하여 보았다. 실제 필요한 비트 수를 구하기 위해 각 런LENGTH의 통계를 구하고 이를 기반으로 허프만 테이블을 설계하였고, 표 5에 각 런LENGTH에 적용되는 허프만 코드의 비트 길이가 정리되어 있다. 런LENGTH의 최대값이 정해져 있지 않기 때문에 최대 런LENGTH를 63으로 제한하였고, 64 이상은 AAC에서 사용하는 ESCAPE 동작을 사용하여 처리 하였다.

표 5. 런LENGTH 부호화를 위한 허프만 테이블

Table 5. Huffman table used for run-length coding.

Symbol (Run Length)	1	2	3	4	5	6	7	8	9	10	...	60	61	62	63	ESC
Bit Length	1	3	3	4	5	5	5	6	6	7	...	14	13	13	15	9

이와 같이 설계된 허프만 테이블을 이용하여 Input#1과 Input#2의 실제 비트수를 구하면 표 4와 같이 기존의 AAC 허프만 부호화와 비교하여 Bit Ratio가 각각 96.8%와 99.7%가 된다. 실제로 구현한 값과 이론적인 값 사이에 어느 정도의 차이가 있음을 알 수 있으나, 이는 런LENGTH 허프만 부호화를 하나의 허프만 테이블만 사용하도록 매우 간단하게 구현하였기 때문이며, 이를 개선시키기 위하여 모드별 및 주파수 대역별로 독립적인 테이블을 선택적으로 사용하거나 다차원으로 허프만 부호화 하는 방법을 사용할 수 있다. 이 문제는 일반적인 허프만 부호화의 최적화에 대한 것이므로 본 논문에서는 다루지 않는다. 즉, 본 논문에서는 비트 플레인 부호화를 사용하는 선택적 부호화 방법을 통하여 기존의 허프만 부호화에 비하여 성능 향상을 얻을 수 있는 것을 이론적으로 확인하였으며, 구체적인 런LENGTH에 대한 최적 허프만 부호화는 추가의 연구 과제로 남겨 둔다.

## V. 결론

본 논문에서는 기존의 오디오 부호화기에서 사용하는 주파수 계수의 허프만 부호화기에 비하여 성능이 향상된 새로운 무손실 부호화기를 제안하였다. 오디오 신호의 주파수 계수의 특성은 주파수 대역마다 변하므로 각 대역에 적합한 부호화 방법을 사용하여야 부호화 성능의 향상을 얻을 수 있고 이를 위하여 선택적 부호화 방법을 사용할 수 있다. 특히 부호화 방법 선택에서 특성이 매우 다른 두 가지의 부호화 방법 중에서 선택을 하여야 선택의 범위가 넓어지고 선택에 의한 성능 향상이 증가된다. 이를 위하여 본 논문에서는 비트 플레인을 통하여 입력 심볼을 새로운 심볼로 변환하여 부호화 하는 새로운 무손실 부호화 방법을 제안하고, 이를 주파수 섹션 그룹 기반으로 선택적으로 사용하여 전체 주파수 계수의 무손실 부호화의 성능을 향상시켰다. 엔트로피를 기반으로 이상적인 비트 수를 계산하면 기존의 AAC 허프만 부호화에 비하여 최대 6%의 성능 향상을 얻는 것을 확인하였다. 실제 부호화 성능을 가장 간단한 허프만 부호화를

가정하고 구하면 이론적 성능에 비하여 약간 저하되지만, 대역별 또는 모드별 허프만 테이블을 선택하는 방법과 다차원 부호화 등을 통하여 허프만 부호화의 최적화를 진행하면 이론적 성능에 근접할 것이며, 현재 허프만 부호화 최적화에 대한 연구를 진행 중에 있다.

### 감사의 글

본 연구는 정보통신부 및 정보통신연구진흥원의 IT신성장동력 핵심기술개발사업 [2007-S005-01, AV코덱 고도화를 통한 리치미디어 방송기술 개발]과 2007년 광운대학교 교내 연구비 지원으로 이루어졌습니다.

### 참고 문헌

1. ISO/IEC 13818-7, "Generic Coding of Moving Pictures and Associated Audio Information - Part 7 : Advanced Audio Coding," 1997.
2. ISO/IEC 14496-3, "Coding of Audio-Visual Objects - Part 3 : Audio, Subpart 4 : General Audio Coding - AAC, TwinVQ, BSAC," 2001.
3. 3GPP TS 26.290, "Extended AMR Wideband Codec: Transcoding Function," 2002.
4. 3GPP TS 26.403, "Enhanced aacPlus General Audio Codec : Encoder Specification AAC Part," 2004.
5. John Watkinson, *The MPEG Handbook MPEG-1, MPEG-2, MPEG-4*, (Second Edition, Focal Press, 2004)
6. Marina Bosi and Richard E. Goldberg, *Introduction to Digital Audio Coding and Standards*, (Springer, 2002)
7. Khalid Saywood, *Lossless Compression Handbook*, (Academic Press, 2003)

### 저자 약력

• 유 승 관 (Seung-Kwan Yoo)

2006년 2월 : 광운대학교 전자공학과(공학사)  
 2008년 2월 : 광운대학교 전자공학과(공학석사)  
 2008년 1월~현재 : (주)엠텍비전 연구원  
 ※ 주관심 분야 : 음성/오디오 신호처리, 통신 신호처리

• 박 호 중 (Hochong Park)

1986년 2월 : 서울대학교 전자공학과(공학사)  
 1987년 12월 : Univ. of Wisconsin-Madison, 전자공학과(M.S.)  
 1993년 5월 : Univ. of Wisconsin-Madison, 전자공학과(Ph.D.)  
 1993년 9월~1997년 8월 : 삼성전자 선임연구원  
 1997년 9월~현재 : 광운대학교 전자공학과 교수  
 ※ 주관심 분야 : 음성/오디오 신호처리, 통신 신호처리, 영상 신호처리

• 오 승 준 (Seoung-Jun Oh)

1980년 2월 : 서울대학교 전자공학과(공학사)  
 1982년 2월 : 서울대학교 전자공학과(공학석사)  
 1988년 5월 : 미국 Syracuse University 전기/컴퓨터공학과(공학박사)  
 1982년 3월~1992년 8월 : 한국전자통신연구원 멀티미디어연구실 실장  
 1992년 9월~현재 : 광운대학교 전자공학과 교수  
 ※ 주관심분야 : 비디오 데이터 처리, 비디오코덱시스템, 비디오 및 영상압축

• 안 창 범 (Chang-Beom Ahn)

1981년 2월 : 서울대학교 전자공학과(공학사)  
 1983년 8월 : 한국과학기술원 전기 및 전자공학과(M.S.)  
 1986년 8월 : 한국과학기술원 전기 및 전자공학과(Ph.D.)  
 1986년 9월~1991년 4월 : University of California, Irvine 연구조교수  
 1991년 4월~1992년 8월 : 생산기술연구원 전자정보센터 부교수  
 1992년 9월~현재 : 광운대학교 전자공학과 교수  
 ※ 주관심 분야 : 생체 신호처리, 의학영상시스템, 영상 신호처리

• 심 동 규 (Donggyu Sim)

1993년 2월 : 서강대학교 전자공학과(공학사)  
 1995년 2월 : 서강대학교 전자공학과(공학석사)  
 1999년 2월 : 서강대학교 전자공학과(공학박사)  
 1999년 3월~2000년 9월 : 현대전자 기반기술 연구소 선임연구원  
 2000년 10월~2002년 3월 : 바로비전 뉴미디어 연구소 선임연구원  
 2002년 4월~2005년 2월 : Univ. of Washington Senior Research Engineer  
 2005년 3월~현재 : 광운대학교 컴퓨터공학과 조교수  
 ※ 주관심분야 : 신호처리, 영상처리, 컴퓨터비전, 멀티미디어

• 백 승 권 (Seung-Kwon Beak)

1999년 2월 : 한국항공대학교 항공전자공학과(공학사)  
 2001년 2월 : 한국정보통신대학교 대학원 전자공학과(공학석사)  
 2005년 8월 : 한국정보통신대학교 대학원 전자공학과(공학박사)  
 2005년 8월~현재 : 한국전자통신연구원 방송미디어연구그룹 선임연구원  
 ※ 주관심분야 : 오디오 신호처리, 음원분리, MPEG, DMB

• 강 경 옥 (Kyoung-Ok Kang)

1985년 2월 : 부산대학교 물리학과(학사)  
 1988년 2월 : 부산대학교 대학원 물리학과(석사)  
 2004년 2월 : 한국항공대학교 대학원 항공전자공학과(박사)  
 1991년 2월~현재 : 한국전자통신연구원 방송미디어연구그룹 3D미디어연구팀장, 책임연구원  
 ※ 주관심분야 : 오디오 신호처리, MPEG-7 및 TV-Anytime 기반 맞춤형 방송