

논문 2004-41SP-6-28

# Hybrid 무손실 오디오 부호화기의 설계

## (A Design of Hybrid Lossless Audio Coder)

박 세 형\*, 신 재 호\*

(Se Hyoung Park and Jaeho Shin)

### 요 약

본 논문에서는 정수 웨이블릿 변환(Integer Wavelet Transform)과 선형 예측(Linear Prediction)을 이용한 Hybrid 무손실 오디오 부호화 알고리즘을 제안하였다. 제안된 부호화기는 입력된 신호를 적당한 프레임으로 분할 후, 정수 웨이블릿 변환과 정수 예측기를 사용하여 입력된 오디오 신호내의 상관관계(Correlation)를 제거하였다. 그리고 신호내의 상관성이 제거된 신호를 엔트로피 부호화(Entropy Coding) 하였다. 엔트로피 부호화기로는 정수 입력원의 특성에 따라 최적의 압축 효율을 갖는 옵션을 선택적으로 처리하는 Adaptive Golomb-Rice 부호화기법을 이용하였다. 제안된 부호화 방식은 모든 연산을 정수 연산으로 가능케 하므로 기존의 실수를 사용하는 연산과 비교하여 연산 속도의 개선과 시스템의 복잡성을 낮추어 고속처리 및 저전력화가 가능하다. 또한 각 프레임은 독립적으로 부호화되고, 부호화된 데이터는 프레임헤더와 바이트 단위로 정렬이 되도록 데이터 포맷을 설계하여, 압축 부호화된 데이터의 이동, 찾기, 편집이 편리하도록 하였다.

### Abstract

This paper proposes a novel algorithm for hybrid lossless audio coding, which employs an integer wavelet transform and a linear prediction model. The proposed algorithm divides the input signal into frames of a proper length, decorrelates the framed data using the integer wavelet transform and linear prediction and finally entropy-codes the frame data. In particular, the adaptive Golomb-Rice coding method used for the entropy coding selects an optimal option which gives the best compression efficiency. Since the proposed algorithm uses integer operations, it significantly improves the computation speed in comparison with an algorithm using real or floating-point operations. When the coding algorithm is implemented in hardware, the system complexity as well as the power consumption is remarkably reduced. Finally, because each frame is independently coded and is byte-aligned with respect to the frame header, it is convenient to move, search, and edit the coded, compressed data.

**Keywords :** Lossless audio coding, Integer wavelet transform, Golomb-Rice code

### I. 서 론

디지털화 된 멀티미디어 신호는 방대한 데이터량을 가지고 있으므로, 이를 효율적으로 저장·전송하기 위해 압축부호화 하게 된다. 일반적인 멀티미디어 신호의 압축 부호화기는 인간의 시청각적 특성을 이용하여 압축하게 되는데, 대표적인 방법으로 MPEG, AC-3 등을 들 수 있다. MPEG 오디오의 경우 인간의 청각 특성을 모델링한 심리음향모델을 적용하여 오디오 신호를 압축하며, 압축된 데이터는 원 신호에 비해 1/6에서 최대

1/12정도의 크기로 부호화 된다. 복원된 오디오 신호의 음질은 CD 음원의 수준과 동등한 결과를 나타내 음질의 저하를 크게 느끼지 못하게 된다.

그런데 MPEG, AC-3와 같은 오디오 부호화기의 음질은 CD 원음과 같은 수준이나, 실제로 복원된 오디오 신호는 인간의 귀에 들리지 않는 잡음이 들어가 있어 원래의 신호를 완벽히 복원할 수는 없다. 음향합성용 음원, 고음질 DVD 오디오, 임체음향신호, 초음파 등의 디지털 오디오 신호는 MPEG과 같은 손실 압축방식을 사용하게 되면 원래의 신호의 특징을 잃게 되거나 또는 귀에 들리지 않는 잡음이 들어가 용도에 맞지 않게 된다. 따라서 원 오디오 신호의 특징을 보존할 수 있는 무손실 압축부호화 방법이 필요하다. 또한 Zip, Rar와 같

\* 정희원, 동국대학교 전자공학과  
(Department of Electronics, Dongguk University)  
접수일자: 2004년6월29일, 수정완료일: 2004년7월9일

은 일반문서용 데이터 압축방식은 무손실 압축부호화 기법이지만, 오디오 신호에 적용할 경우 거의 압축이 되지 않는다. 이에 본 논문에서는 오디오 신호에 적용하기 위한 무손실 오디오 부호화기를 제안한다.

기존의 무손실 오디오 부호화는 신호의 중복성 제거 방법에 따라 크게 2가지로 구분할 수 있는데, 첫째 음성 신호처리에 많이 사용되어진 예측모델(Prediction Model)을 이용하는 방법과 DCT, DWT와 같은 변환(Transform)을 이용하는 손실부호화모델(Lossy Coding Model)방법이 있다<sup>[1][2][8][10][11][12]</sup>. 예측모델은 음성신호의 압축부호화에서 많이 사용되어졌는데, 대표적으로 DPCM(Differential PCM), ADPCM등을 들 수 있고, 오디오 신호의 경우는 T. Robinson이 제안한 Shorten<sup>[1]</sup>의 방법을 들 수 있다. 손실부호화모델은 M. Purat이 제안한 LTAC( Lossless Transform Audio Coding)<sup>[2]</sup>을 들 수 있다.

일반적으로 무손실 부호화방법은 손실 부호화방법에 비해 압축률이 많이 떨어진다. 따라서 무손실 부호화방법은 원신호의 무손실 보존 이외의 다른 장점으로 이를 보완할 필요가 있다. 예로 무손실 부호화기는 간단한 구조로 연산량이 적어야 하며 부호화된 데이터로부터 직접 이동(move), 찾기(search), 편집(edit) 등이 쉬워야 한다. 또한 다양한 오디오 신호의 포맷을 지원할 수 있어야 하며, 데이터의 필요에 따라 쉽게 손실-무손실 형태의 오디오 신호를 제공할 수 있는 형태이어야 한다.

본 논문에서는 디지털 오디오의 무손실 부호화를 위해 손실부호화모델과 예측모델을 이용한 Hybrid 무손실 부호화 알고리즘을 제안하였다. 리프팅을 이용한 정수 웨이블릿 변환을 사용하는 부호화기에서는 LTAC과 같이 역 변환하는 과정을 수행하지 않도록 하였고, 정수 웨이블릿 변환된 신호를 간단한 정수계수를 갖는 예측을 수행하여 상관제거 효율을 높였다. 이는 제안된 알고리즘의 모든 연산을 정수연산으로 가능하게 한다. 또한 신호내의 중복성이 제거된 신호는 효율적인 부호화를 위해 CCSDS(Consultative Committee for Space Data System)에서 제안한 Adaptive Golomb Rice 부호화 알고리즘을 이용하였다.

본 논문의 구성은 II장에서 무손실 부호화기 시스템의 기본 구조인 프레임, 중복성 제거를 위한 Intra-channel 상관제거 방법을 기술하고, III장에서는 새로운 부호화기의 알고리즘을 제안한다. 그리고 IV장에서 시뮬레이션 및 시뮬레이션 결과를 검토하고, 마지막으로 결론을 맺는다.

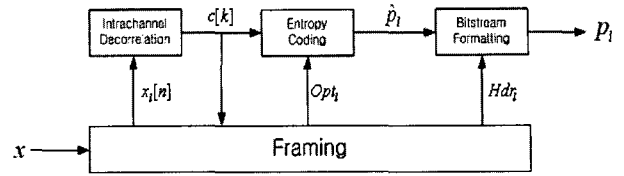


그림 1. 무손실 오디오 부호화 시스템 블록도  
Fig. 1. Block diagram of lossless audio coding system.

## II. 무손실 오디오 부호화 시스템

### 2.1 프레임(Framing)

일반적인 무손실 오디오 부호화기는 그림 1과 같다. 시간영역의 입력신호  $x$ 는 프레임 블록으로 들어가  $N$ 개의 샘플을 갖는 프레임으로 나누어 진다. 이 때 프레임의 길이는 부호화기의 메모리와 시스템의 지연을 고려하여 적절한 값이 선택되어지는데, 일반적으로  $N$ 은 512에서 4096의 범위에서 선택되어 진다.

프레임 블록은 간단한 기능을 수행하지만, 디지털 오디오 응용에서 필요한 기능을 갖는다. 프레임에서는 디지털 오디오 신호를 동일한 시간 구간의 프레임으로 나누어 주는데, 이 때 프레임 구간이 너무 짧거나 너무 길면, 부호화 또는 전송할 때 시스템의 오버헤드가 발생하게 되므로, 적절한 길이로 입력데이터를 나누어 주는 기능을 갖게 된다. 또한 프레임은 데이터의 편집성을 증가시키기 위한 각종 정보를 제공하게 된다. 입력신호의 샘플링주파수, 샘플의 해상도, 오디오 채널 수, 프레임내의 샘플 수, 압축된 프레임의 길이 등 원신호의 정보와 압축된 데이터의 정보를 제공한다.

프레임된 신호  $x_i[n]$ 은 전후의 다른 프레임에 독립적으로 부호화를 하게 되므로, 메모리를 절약할 수 있고, 오디오 데이터의 편집성을 제공할 수 있게 된다. 또한, 압축된 신호에 랜덤엑세스(Random Access)하기 위한 기본적인 단위는 프레임이 된다.

### 2.2 상관제거

프레임된 신호  $x_i[n]$ 은 중복성 제거를 위해 상관제거 후 다른 프레임에 독립적으로 부호화된다.  $x_i[n]$ 은 상관제거 블록의 입력으로 들어가 프레임내의 상관제거된 출력  $c_i[k]$ 를 구하게 되는데, 상관제거 블록은 무손실 부호화 시스템에 따라 일반적으로 선형 예측기(Predictor)를 사용하거나 변환기(Transformer)를 사용하게 된다. 상호상관제거기의 출력  $c_i[k]$ 는 신호간의 중복성이 제거되었으므로 엔트로피 부호화기를 통해 부호

화되고, 프레임에서 제공되는 헤더정보와 함께 비트 스트림  $p_i$ 로 포맷화되어 저장 또는 전송되어 진다.

Shorten에서는 예측방법을 이용하는 2가지 다른 방식을 제안하였는데, Durbin's 알고리즘을 이용하여 선형예측 필터의 계수를 계산하고 부호화하는 방식과 고정된 정수계수를 가지는 예측필터를 통해 예측기의 출력에 정수가 되도록 하여 정수연산만으로 구현되는 방식을 제안하였다. Shorten에서 제안된 방식 중 Durbin's 알고리즘을 이용하여 선형예측 필터의 계수를 결정하는 방식은 FIR 또는 IIR타입의 예측필터와 양자화기를 같이 이용하여 예측하는 형태로 많이 연구되어 지고 있다 [8][9][10][11]. 이러한 방식은 프레임된 신호  $x_i[n]$ 에 대해 Durbin's 알고리즘을 통해 필터계수를 매 프레임마다 계산하고, 예측된 신호와 필터계수를 부호화하여 전송하게 된다. 이는 매 프레임마다 필터계수를 계산하여야 하고, 필터계수의 유한길이문제로 완벽 복원이 어려운 단점이 있다. 또한 예측을 이용한 방식은 협대역 신호인 음성신호에는 적당하지만 일반적인 광대역 신호인 음향신호에는 좋은 성능을 발휘하지 못한다.

다음 상관제거 방법으로 변환기를 이용하는 경우로 대표적인 방식은 변환부호화를 들 수 있다. 변환부호화는 멀티미디어 신호를 압축하기 위해 DCT, DWT를 사용하여, 시간(공간)영역의 신호를 다른 영역으로 변환시킨 후 부호화하는 방법이다. 그러나 유한비트의 입력(오디오의 경우 8비트 또는 16비트)신호 또는 정수입력 신호를 변환한 결과는 무한비트의 실수값이 나와 부호화 과정에서 양자화에 의해 에러가 발생하게 되어 결과적으로 손실부호화(Lossy coding)가 된다.

M. Purat의 제안된 LTAC은 손실부호화(또는 변환부호화) 모델을 이용한 무손실 부호화 방식이다. 손실부호화모델을 이용한 무손실 부호화기는 그림 2와 같은 구조로 이루어져 있다. 일반적인 손실부호화 기법의 부호화 과정에서 원래의 신호  $x_i[n]$ 과 부호화된 신호  $\hat{c}[k]$ 는 식 (1)과 같이 표현할 수 있다.

$$\hat{c}[n] = Q[T\{x_i[n]\}] \tag{1}$$

이 때,  $T(\cdot)$ 는 변환(Transform)을  $Q[\cdot]$ 는 양자화를 말한다.

손실 부호화기에서 변환된 신호는 양자화 되고, 양자화 된 신호는 엔트로피 부호화되어 전송되어 진다. 그러나 비선형 연산인 양자화시 발생하는 에러 때문에 복호화기에서 원 신호의 완벽 복원이 불가능하므로 부호

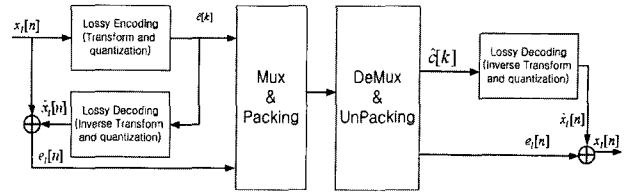


그림 2. 손실부호화모델을 이용한 무손실 부호화 시스템의 블록도

Fig 2. Block diagram of lossless coding system using lossy coding model.

화기에서는 부호화된 신호의  $\hat{c}[k]$ 를 역과정을 통해 원 신호와의 에러신호  $e_i[n]$ 를 구하여 부호화된 신호  $\hat{c}[k]$ 와 다중화하여 전송하게 된다. 식 (2), (3)은 에러신호  $e_i[n]$ 과 원 신호의 복원과정을 나타낸다.

$$e_i[n] = x_i[n] - Q\{T^{-1}\{\hat{c}[k]\}\} \tag{2}$$

$$x_i[n] = Q\{T^{-1}\{\hat{c}[k]\}\} + e_i[n] \tag{3}$$

LTAC에서는 DCT를 이용하여 신호를 변환한 후, 이를 스케일링 하고, 양자화를 하게 된다. 이때 입력  $x_i[n]$ 이 정수값을 가져도, DCT 계수  $c[k]$ 는 실수가 된다. 스케일 팩터와 양자화 스텝사이즈를 1로 하여 양자화를 하게 되면, 이 결과는 간단히 정수변환을 위한 소수버림(Truncation)이 되고, 양자화된  $\hat{c}[k]$ 를 얻게 된다. 이와 같이 정수화된 신호를 엔트로피 부호화를 하여 전송하게 되면, 복호화기에서는 부호화된 신호를 역 스케일링과 역 DCT를 수행한 후, 다시 양자화 한 결과  $\hat{x}[n]$ 을 얻게 되는데, 이는 손실부호화의 일종인 변환부호화와 동일한 결과를 갖게 된다. 따라서 무손실 복원을 위해서 부호화기에서는 양자화시 발생된 에러값을 재 전송해 주어야 한다. 그러므로 LTAC의 부호화기에서는 복호화 과정과 같이 양자화 된 신호  $\hat{c}[k]$ 를 이용하여 근사화된 신호  $\hat{x}[n]$ 와 에러값  $e_i[n]$ 을 구하여 전송하므로 원 신호를 완벽 복원하게 된다.

### III. 제안된 Hybrid 무손실 부호화기 및 데이터 포맷

#### 3.1. 제안된 알고리즘

제안한 무손실 부호화기는 프레임된 신호 사이의 상관을 제거하기 위해 정수 웨이블릿 변환을 사용하였다. 입력된 신호의 중복성이 제거된 신호는 Adaptive Rice 부호화를 위해 세그먼트별로 다시 예측을 통해 다시 한번 상관을 제거한 후 부호화하였다. 그림 3은 제안된 부호화 알고리즘의 블록도이다.

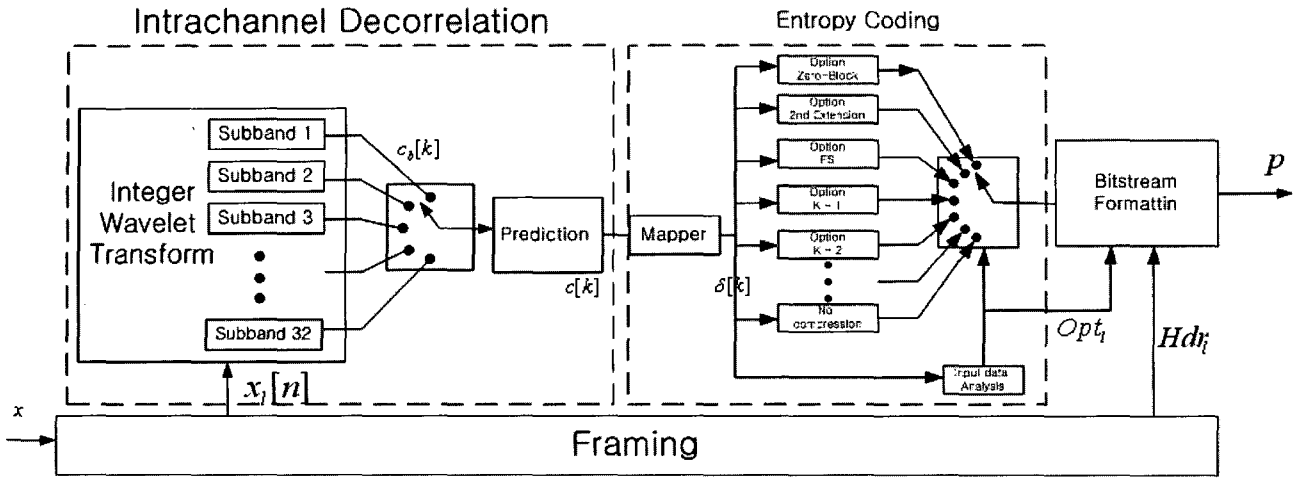


그림 3. 제안된 Hybrid 무손실 부호화기  
 Fig. 3. Proposed Hybrid lossless Coder.

입력신호  $x$ 는 프레임링 블록을 거쳐 길이  $N$ 을 갖는 프레임링된 신호  $x_l[n]$ 로 나뉘어 진다. 이때 프레임의 길이  $N$ 은 512에서 4096의 길이로, 샘플링 주파수가 44.1kHz인 경우 약 12ms에서 93ms의 구간길이를 갖게 되는데 이는 압축, 복호시 처리 및 지연시간과 메모리의 사용을 적절히 유지할 수 있게 한다. 프레임링된 신호  $x_l[n]$ 은 식 (4)와 같다,

$$x_l[n] = x[lN + n] \quad (4)$$

이 때  $n = 0, 1, 2, \dots, N-1$  이고  $l$ 은 프레임 번호로  $l = 0, 1, 2, \dots$  이다.

프레임별로 분할된  $x_l[n]$ 은 정수 웨이블릿 변환되는데, 웨이블릿 변환은  $L$ 레벨 패킷변환 후  $2^L$ 대역으로 분할하여  $c_b[k]$ 를 구하게 된다. 이때  $b$ 는 0에서  $2^L - 1$ 까지의 대역을 가리킨다. 그러나 일반적인 음향 신호가 아닌 음성의 경우 샘플링 주파수가 낮아 대역을 많이 나누는 것은 비효율적이기 때문에 웨이블릿 패킷변환 레벨은 0에서 최대 5레벨까지 입력신호에 따라 다르게 선택할 수 있다. 여기서 0레벨은 정수 웨이블릿 변환을 수행하지 않고 간단한 예측 후 Adaptive Rice 부호화를 한 경우이다. Dyadic 5레벨 정수 웨이블릿 패킷변환을 하게 되면 32개의 대역은 16( $N=512$ )에서 128( $N=4096$ )개의 샘플로 구성되고, 각각 독립적인 세그먼트 블록별로 예측을 하게 된다. 예측된 세그먼트 블록은 Adaptive Rice 부호화기의 입력으로 들어가 부호화된다. 마지막으로 엔트로피 부호화된 신호는 프레임 헤더와 전체 오디오 신호의 헤더정보 등과 함께 비트스트림으로 출력되어 진다.

웨이블릿 변환은 인접한 프레임 간의 오버랩 없이 신호 변환이 가능하여 프레임마다 독립적으로 부호화하는 방식에 적합하고, 또한 웨이블릿 패킷변환에 의해 전체 대역을 인위적으로 분할하지 않아도 트리구조의 32 서브밴드로의 분리가 가능하다. 정수 웨이블릿 변환은 리프팅을 이용하여 정수 입력에 대해 근사화된 웨이블릿 변환 정수계수 신호를 얻을 수 있고 또한 역변환이 가능하다. 본 논문에서는 프레임링된 신호  $x_l[n]$ 사이의 상관제거를 위해 정수 웨이블릿 변환을 이용하였다 [3][4][6].

리프팅을 이용한 정수 웨이블릿 변환은 예측(Prediction)과 갱신(Updating)의 과정을 이용하여 웨이블릿 변환을 구현하게 된다. 리프팅의 장점은 빠른 속도와 적은 메모리의 사용, 비선형연산인 정수 대 정수변환의 변환과 그 역과정도 변환과정과 동일하게 구현된다는 것이다.

리프팅을 이용한 정수웨이블릿 변환과정은 분해단계에서 짝수번째 신호  $x_e$ 와 홀수 번째 신호  $x_o$ 로 분리되고, 이는 예측과 갱신의 단계를 거쳐 식 (5), (6)과 같이 하이패스된 신호  $d$ 와 로우패스된 신호  $s$ 를 구하게 된다.

$$d = x_e - \lfloor P(x_o) \rfloor \quad (5)$$

$$s = x_o + \lfloor U(d) \rfloor \quad (6)$$

리프팅을 통해 정수 웨이블릿 변환된 신호  $d$ 는 인접 신호의 차 즉  $x_e$ 를 이용하여  $x_o$ 를 예측한 결과이고,  $s$ 는 인접신호의 평균을 구하는 과정으로  $d$ 를 이용하여  $x_e$ 의 고주파 성분을 제거해 주는 과정으로 볼 수 있다.

그리고 적절한 웨이블릿 필터의 선택은 입력 신호에 대해 예측과 갱신 과정을 잘 수행할 수 있어 신호의 상관 관계를 잘 제거할 수 있다. 그러나 입력 신호에 따라 항상 다른 웨이블릿 필터를 찾는 것은 많은 연산량을 필요로 하게 되며, 사용된 모웨이블릿을 매번 전송해야 하는 부담이 생기게 된다. 본 논문에서는 무손실 부호화기의 구조가 간단하고 적은 연산량으로도 적절한 상관 제거가 가능토록 하기 위해 리프팅을 이용한 정수 웨이블릿 변환 후 변환된 신호에 다시 간단한 정수계수를 갖는 선형 예측을 이용하여 입력신호에 독립적인 웨이블릿 필터 사용의 단점을 보완하였다. 식 (7)과 (8)은 제안된 알고리즘에서 사용된 Biorthogonal (9-7)필터와, Interpolation(4,4)필터를 이용한 정수 웨이블릿 변환의 구현식이다.<sup>[3][4][6]</sup>

$$\begin{aligned}
 d'[k] &= x_o[k] + \alpha(x_e[k] + x_e[k+1]) + 1/2 \\
 s'[k] &= x_e[k] + \beta(d'[k] + d'[k-1]) + 1/2 \\
 d[k] &= d'[k] + \gamma(s'[k] + s'[k+1]) + 1/2 \\
 s[k] &= s'[k] + \delta(d[k] + d[k-1]) + 1/2 \\
 \alpha &= -203/128 \\
 \beta &= -217/4096 \\
 \gamma &= 113/128 \\
 \delta &= 1817/4096
 \end{aligned}
 \tag{7}$$

$$\begin{aligned}
 d[k] &= x_o[k] - \frac{9}{16}(x_e[k] + x_e[k+1]) \\
 &\quad - \frac{1}{16}(x_e[k-1] + x_e[k+2]) + 1/2 \\
 s[k] &= x_e[k] + \frac{9}{32}(d[k-1] + d[k]) \\
 &\quad - \frac{1}{32}(d[k-2] + d[k+1]) + 1/2
 \end{aligned}
 \tag{8}$$

정수 웨이블릿을 통해 상관제거된 입력  $c_b[k]$ 는 대역별로 예측기를 통해 다시 한번 상관제거를 하게 된다. 이때 Durbin's 알고리즘을 이용하거나 실수계수를 갖는 예측기는 많은 연산량과 복잡한 구조를 가지게 되기 때문에 1차 지연을 갖는 예측기를 이용하였다. 예측 계수  $P[k]$ 는 식 (9)와 같다.

$$P[k] = \{1, -1\}
 \tag{9}$$

상관제거된 신호  $c[k]$ 를 엔트로피 부호화하기 위해 본 논문에서는 CCSDS에서 제안한 Adaptive Golomb Rice 부호화 알고리즘을 이용하였다. Rice 부호화 알고리즘은 입력되는 블록별로 독립적으로 부호화되고, 복

호하기 위해 추가적인 정보가 필요 없어 부호화된 데이터로부터 직접 편집이 가능하고, 적은 메모리의 사용과 구조가 간단한 특징을 갖고 있다.

일반적인 Rice 부호는 입력정수  $c$ 를 unary부호와 이진부호로 나누어 부호화 하게 된다. 입력 정수  $c$ 를  $m=2^l$ 로 나눈 후 몫은 unary부호로 나머지 값은  $l$ 비트의 이진부호로 부호화한다. Rice 부호에서  $l$ 은

$$l = \lceil \log_2(E(|c|)) \rceil
 \tag{10}$$

과 같이 구해진다.

입력신호  $c$ 는  $n$ 비트인  $J$ 개의 샘플로 나뉘어져 각각의 세그먼트별로 전처리기로 전달되어진다. 이때  $J$ 는 CCSDS에서는 8에서 64정도의 값을 갖을 때 최적의 성능을 나타낸다. 입력 세그먼트  $c[k]$ 는 정수값을 갖기 때문에 Rice 부호화를 하기 위해서는 입력 세그먼트량의 정수로 바꾸는 맵핑 과정이 필요하다. 그리고 예측기를 거친 값을 일반적인 부호비트의 할당만으로 입력신호의 맵핑은 비효율적이기 때문에 식 (11)과 같이 맵핑한다.

$$\delta[k] = \begin{cases} 2c[k] & 0 \leq c[k] \\ 2|c[k]| - 1 & c[k] < 0 \end{cases}
 \tag{11}$$

이와 같이 맵핑된 입력신호  $\delta[k]$ 는 Adaptive Rice 부호화기에서 최적의 부호화 효율을 갖는 옵션을 선택하기 위해 각각의 옵션에 대해 부호화에 할당되는 비트 수를 계산한다. Adaptive Rice 부호화기는 4가지의 옵션을 갖는데, Zero-block 옵션, 2차 확장옵션(2nd Extension), FS(Fundamental Sequence)옵션, 그리고 Split 옵션이다<sup>[13]</sup>.

### 3.2 데이터 포맷의 설계

본 논문에서는 제안된 부호화기를 통해 부호화된 신호를 유용하게 사용하기 위한 무손실 부호화 데이터포맷을 설계하였다. 그림 4는 제안된 부호화기의 압축 부호화된 정보를 비트스트림으로 포맷화하기 위한 데이터 포맷이다. 제안된 데이터 포맷은 압축된 오디오 전체 데이터의 기본 특징을 정의한 데이터 헤더와 각각의 프레임별 정보를 가지는 프레임 헤더로 구성되어져 있다.

전체 오디오 샘플수를 나타내는 TSN은 4바이트의 길이를 가지는데 최대 4GB길이의 데이터까지 표현 할 수 있다. CDI는 2바이트의 길이로 입력 신호의 정보와 압축방법에 대한 정보를 가지고 있게 되는데, CN은 3비트 길이로 부호화된 데이터의 채널 수를 정의하고 최

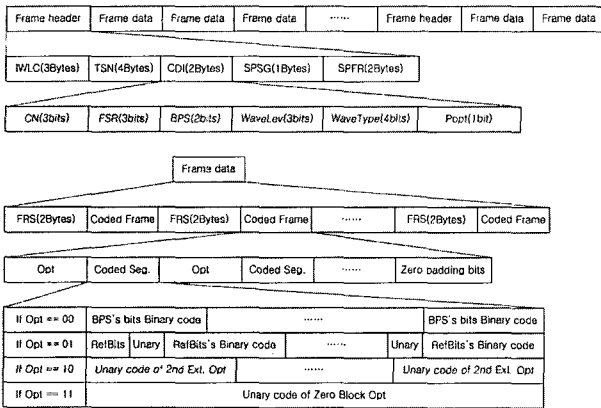


그림 4. 제안된 부호화기를 위해 설계된 데이터 포맷  
 Fig. 4. A design of data format for proposed coding algorithm.

대 6채널까지 설정이 가능하다. *FSR*은 3비트의 길이로 오디오 데이터의 샘플링 주파수를 나타내며 *BPS*는 2비트의 길이로 샘플의 해상도를 나타낸다. 그리고, 3비트 길이의 *WaveLev*는 웨이블릿 레벨을 나타내는데, 0-7까지의 레벨 값을 가지게 된다. *WaveType*은 모 웨이블릿의 종류, *Popt*는 예측기 옵션으로 *Popt* = 0 이면, 예측기를 사용하지 않는 경우이다. 그리고, *SPSSG*는 세그먼트 블럭당 샘플수, *SPFR*은 프레임당 샘플수를 의미한다. *CDI*는 오디오 데이터 뿐 아니라 입력신호의 종류에 따라 다양한 형태로 사용자가 정의하여 사용하므로 오디오 뿐만 아니라 초음파, 센서데이터, 영상 등에도 적용할 수 있다.

프레임 헤더는 각각 프레임의 정보를 갖게 되는데 부호화된 프레임 포맷은 데이터 편집성을 증가시키기 위해 약간의 오버헤드를 감수하고 바이트 단위로 정렬하기 위해 Zero padding을 하였다. 또한 프레임의 오류를 검출하기 위해 CRC 코드를 사용하여 프레임에 대한 오류검출코드 추가할 수 있다. 프레임 데이터는 먼저 매 프레임의 길이를 나타내는 2바이트의 *FRS*로 시작한다. *FRS*는 압축된 프레임의 길이를 바이트 단위로 표현하였다. 다음은 *Opt*로 프레임 내의 각각의 세그먼트 블럭별 Adaptive Rice 부호화의 옵션값으로 *Opt*는 2비트의 길이로 4가지의 옵션을 나타낸다.

각 프레임은 독립적으로 부호화되어 연속된 앞뒤의 프레임에 관계없이 완벽복원이 가능하다. 그리고 부호화된 데이터는 동일한 프레임구조가 연속될 경우 프레임 헤더에 프레임의 정보를 가지게 된다. 또한 프레임 데이터가 바이트 단위로 정렬이 되어 있고 매 프레임 데이터의 앞에 프레임 길이정보가 있기 때문에 이동과

찾기가 쉽고, 데이터의 내용을 편집하여 프레임의 정보가 변할 경우 변한 프레임에 대한 새로운 프레임헤더의 삽입으로 전체 압축 부호화된 데이터의 수정 없이 편집이 가능토록 하였다. 따라서 제안된 포맷은 약간의 오버헤드를 갖게 되지만, 압축 부호화된 데이터의 효율적인 사용이 가능토록 하였다.

#### IV. 시뮬레이션 결과 및 검토

제안된 무손실 부호화기의 성능을 실험하기 위해 본 논문에서 설계한 무손실 데이터 포맷을 이용하여 다양한 음원에 대한 실험을 실행하였다. 일반적으로 무손실 부호화기는 입력신호에 따라 매우 다른 압축률을 갖게 된다. 또한 상호상관계거기의 성능이 입력신호의 특성에 따라 민감하게 반응을 하게 된다. 예로 Shorten과 같은 예측기법을 이용한 방식은 음성, 고주파 성분이 적은 음원 등에는 상당히 좋은 결과를 갖는 반면에 록 음악과 같이 비트가 강하고 고주파 성분이 큰 음원에 대해서는 좋지 않은 결과를 나타낸다. 반면 LTAC은 대부분의 음원에 좋은 성능을 나타내지만, Shorten에 비해 음성과 같은 신호에 대해서는 좋지 않은 결과를 나타내게 된다.

따라서 제안한 무손실 부호화기의 성능을 실험하기 위해서 일반 오디오 CD(Sony Music. Max9 album)에서 17개의 모든 트랙의 처음부터 20초간의 데이터를 추출하여 사용하였다. 사용된 음원을 분석해 보면 음성, 도입부의 악기 솔로연주, 비정재적인(non-stationary) 타악기 연주, 컴퓨터 음악 등 다양한 형태의 음원을 얻을 수 있었다. 그리고 추출된 데이터는 44.1kHz의 샘플링 주파수를 갖고, 16비트의 샘플해상도, 그리고 모노 채널을 갖는다.

또한 제안된 무손실 부호화기는 신호에 독립적인 모 웨이블릿을 이용하여 정수 웨이블릿 변환을 하므로 웨이블릿 변환에 사용되는 모 웨이블릿에 따라 다른 성능을 발휘하게 된다. 따라서 본 연구에서는 대표적인 Biorthogonal (9-7)필터와, Interpolation(4,4)필터를 이용하여 시뮬레이션을 하였다. 또한 하드웨어의 부담을 덜기위한 간단한 1차 예측만을 이용한 부호화과정도 시뮬레이션을 하였다. 본 논문에서는 사용한 필터에 따라 IWL1형, IWL2형 그리고, IWL3형으로 구분하였는데 각각의 특징은 표1과 같다.

본 논문에서는 위의 3가지 형태의 파라미터를 사용한 부호화기를 가지고 표본 시디로부터 추출한 데이터

표 1. IWL1형, IWL2형과 IWL3형의 파라미터  
Table 1. Parameters of IWL1, IWL2, and IWL3 types.

Type	Wavelet type	Frame Size	SegSize	Multi Level
IWL1	Biorthogonal(9,7)	1024	32	4
IWL2	Interpolation(4,4)	2048	64	5
IWL3	No operation	1024	128	0

를 사용하여 부호화를 실시하였다. 제안된 무손실 부호화기는 역변환이 가능한 정수 웨이블릿 변환과 정수계수의 예측기를 사용하므로 부호화된 데이터의 완벽복원이 가능하였다. 각각의 결과는 대표적인 예측기법인 Shorten, 손실부호화 모델을 이용한 LTAC과 비교하여 표2와 그림 5에 나타내었다. 부호화 비교를 위해 압축률은 다음과 같이 계산하였다.

$$\text{Ratio (\%)} = 100 \frac{\text{Compressed file Size in byte}}{\text{Original file Size in byte}} \quad (12)$$

제안된 방식은 평균 68~69%로 압축이 됨을 확인할 수 있었는데, 이는 Shorten에 비해서는 같거나 더 좋은 결과를 얻을 수 있었으나 손실 부호화 모델인 LTAC에 비해서는 약 2%의 압축률이 떨어짐을 확인할 수 있었다.

그러나 제안된 방식은 [3]에 의하면 일반적인 DWT를 사용할 경우 보다 bior(9,7)의 경우 64%의 Interpolation ( $N, \tilde{N}$ ) 필터인 경우 100%의 연산량을 줄일 수 있게 된다. 따라서 정수 웨이블릿을 사용하게 되면 웨이블릿 필터(Symmetric and Even인 경우)의 길이가  $h$ , 프레임 사이즈  $N$ , 웨이블릿 레벨  $L$ 인 경우 최대  $hLN/2 + L$ 의 곱과  $hLN$  개의 합이 필요하게 된다.

[9]에 의하면 DCT의 경우  $(1/2)N \log 2N$ 의 곱과 약  $(3/2)N \log 2N - N + 1$ 의 합 연산이 필요하게 된다. 그러나 DCT의 경우 부동 소숫점(Floating point)의 실수 연산인데 반하여 정수 웨이블릿 변환은 고정소숫점(Fixed point)연산으로 계산이 되거나, 필터의 계수가  $x/2^n$ 의 형태를 가지고 있어 알고리즘에 따라 정수 곱과 쉬프트연산으로도 구현이 가능하므로 DCT에 비하여 하드웨어적인 비용과 복잡성을 줄일 수 있다.

그리고, LTAC의 경우 손실 부호화 모델을 이용하므로 정수입력 신호에 대하여 변환된 결과가 실수값이므로 양자화를 걸쳐야 하고, 다시 한번 역변환을 통해 에러값을 계산해야 하는 부담이 생기게 되어, 부호화기에서는 DCT와 IDCT를 동시 수행하여야 하나, 제안된 알

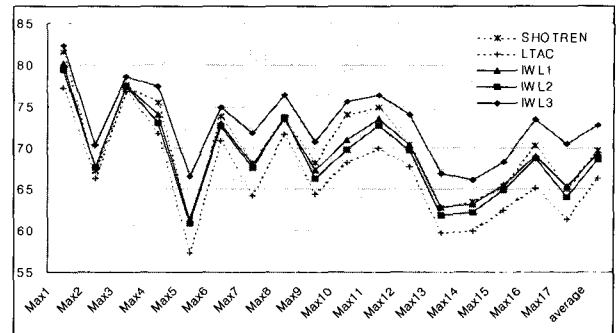


그림 5. Shorten, LTAC과 제안된 알고리즘의 압축률[%]  
Fig. 5. Compression ratio[%] of Shorten, LTAC and proposed algorithms.

표 2. Shorten, LTAC과 제안된 알고리즘의 압축률[%]  
Table 2. Compression ratio[%] of Shorten, LTAC and proposed algorithms.

Track NO.	Shorten	LTAC	IWL1	IWL2	IWL3
Max1	81.53	77.21	80.22	79.43	82.36
Max2	67.17	66.28	67.58	67.65	70.32
Max3	77.59	76.85	77.49	77.40	78.65
Max4	75.48	71.74	73.96	73.00	77.41
Max5	60.89	57.31	61.34	60.89	66.58
Max6	73.78	70.84	72.90	72.56	74.82
Max7	68.19	64.29	67.94	67.62	71.84
Max8	73.42	71.65	73.50	73.66	76.32
Max9	68.14	64.44	67.34	66.32	70.74
Max10	74.02	68.15	70.98	69.80	75.52
Max11	74.89	69.98	73.48	72.70	76.33
Max12	70.34	67.71	70.36	69.61	73.99
Max13	62.82	59.73	62.90	61.90	66.93
Max14	63.46	59.92	63.20	62.18	66.09
Max15	65.35	62.49	65.54	64.91	68.26
Max16	70.21	65.18	68.95	68.65	73.39
Max17	65.06	61.43	65.30	64.09	70.47
average	69.70	66.31	69.22	68.58	72.67

고리즘의 경우 정수 역 웨이블릿 변환과정이 필요 없으므로 부호화시 LTAC보다 많은 계산량이 감소된다. 따라서 제안된 알고리즘은 실수 연산 없이 정수 연산만으로도 구현이 가능하며 메모리의 소모가 적어 하드웨어의 비용을 낮출 수 있다. 따라서 최근 고속이동통신에서 요구되는 고속처리와 저전력화가 가능하다.

또한 IWL3와 같이 웨이블릿 변환없이 1차예측기와 Adaptive 라이스 부호화를 이용한 방식으로 부호화를 실행해 평균 약 72%로 압축이 됨을 볼 수 있었다. IWL3의 경우 매우 적은 연산량과 모든 정수연산만으로 가능하게 되고, CCSDS의 부호화기와 같은 기능을 할 수 있다. 따라서 오디오 뿐 아니라 영상등 다른 정수 데이터를 효과적으로 부호화할 수 있게 된다.

### V. 결 론

본 논문에서는 정수 웨이블릿과 예측기법을 이용한 무손실 부호화기와 이를 위한 데이터 포맷을 제안하였다. 제안된 알고리즘은 기존의 알고리즘과 거의 대등한 성능을 갖지만 실수연산이 필요치 않게 되어 시스템의 복잡성을 낮추고 하드웨어의 비용을 줄일 수 있어 고속 이동통신에서 요구되는 저전력화가 가능하다. 부호화기에서 LTAC과 같이 역변환이 필요치 않으며, 부동소수점 연산을 정수연산만으로도 계산이 가능하므로 고속처리가 가능하다. 그리고 응용되는 분야에 따라 웨이블릿 패킷의 고대역 성분을 제거하여 손실 복원형태의 데이터를 제공할 수 있다.

또한 제안된 알고리즘과 데이터 포맷은 데이터의 이동, 복사, 편집, 그리고 랜덤 액세스가 프레임단위로 가능하게 되어 부호화된 데이터의 유용성을 높일 수 있었다. 그리고 부호화기의 성능에 따라 다양한 옵션을 이용한 부호화가 가능하게 되어, 웨이블릿 변환없이 간단한 예측만을 이용하여 부호화하거나, 오디오 신호 이외의 초음파, 센서 데이터, 영상과 같은 정수 데이터의 효과적인 부호화가 가능하다.

### 참 고 문 헌

[1] T. Robinson, "SHORTEN: Simple lossless and near-lossless waveform compression", Cambridge Univ.Eng. Dept., Cambridge, UK, Tech.Rep.156, 1994.  
 [2] M.Purat, T.Liebchen and P.Noll, "Lossless transform coding of audio signals," Proc. 101nd AES Conv., Munich, Germany, 1997, preprint 4414.  
 [3] I. Daubechies and W. Sweldens, "Factoring wavelet transforms into lifting step," J. Fourier Anal. Appl., vol. 4, no.3, pp.247-269, 1998.

[4] G. Strang and T. Nguyen, Wavelet and Filter Bank. Wellesley, Cambridge, 1996.  
 [5] M. Vetterli and J. Kovacevic, Wavelets and Sub-band Coding. Prentice Hall, Englewood Cliffs, NJ, 1995.  
 [6] W. Sweldens. "The Lifting scheme: A custom-design construction of biorthogonal wavelet," Appl. Comput. Harmon. Anal., vol. 3, no. 2, pp.186-200, 1996.  
 [7] N. S Jayant and P. Noll. Digital coding of waveforms. Printice Hall, Engewood Cliffs, NJ, 1984  
 [8] M. Hans, R.W. Schafer, "Lossless Compression of Digital Audio," IEEE Signal Preccessing Magazine, July, 2001.  
 [9] Y. Zeng, G. Bi, and A.R. Leyman, "New Polynomial Transform Algorithm for Multidimensional DCT", IEEE Trans. on Signal Processing, vol. 48. no. 10. Oct. 2000.  
 [10] P. Craven, M.Law, and J.stuart, "Lossless compression using IIR prediction filters," Proc. 102ndAES Conv., Munich, Germany, 1997, preprint 4415.  
 [11] A. Bruekers, A.Oomen, and R.van der Vlenuten, "Lossless coding for DVD audio, " Proc. 101st AES Conv., Los Angeles, Ca, Nov.1996, preprint 4358.  
 [12] C. D. Giurcaneanu, I. Tabus, and J. Astola. "Adaptive context based sequential prediction for lossless audio compression," In Proc. Eusipco-98 IX European Signal Processing Conference, volume 4, pp. 2349-2352, Rhodes, Greece, Sept. 1998.  
 [13] S.Y. Park, S.H. Park, and J.H. Shin,"A lossless and lossy audio coding method in wavelet domain," ICEIC Proc., June , 2002.  
 [14] K. Sayood, Introduction to Data Compression, Morgan Kaufmann,2000.  
 [15] Album "Max9". Sony Music. April. 2002.

### 저 자 소 개



**박 세 형**(정회원)  
 1997년 동국대학교 전자공학과 학사 졸업  
 1999년 동국대학교 전자공학과 석사 졸업.  
 2001년 동국대학교 전자공학과 박사 수료.

<주관심분야: 신호처리, 객관적심리음향평가, 오디오 부호화 >



**신 재 호**(정회원)  
 1979년 서울대학교 전자공학과 학사 졸업.  
 1982년 서울대학교 전자공학과 석사 졸업.  
 1987년 서울대학교 전자공학과 박사 졸업.

1988년 현재 동국대학교 전자공학과 교수  
 <주관심분야: 디지털 신호처리, 워터마크, 오디오 부호화 >