

# 콜레스키 분해와 골롬-라이스 부호화를 이용한 무손실 오디오 부호화기 설계

정전대<sup>†</sup>, 신재호<sup>\*\*</sup>

## 요 약

무손실 오디오 부호화기에 있어서 선형예측기 및 이에 적합한 엔트로피 부호화기의 설계가 가장 중요한 부분이다. 본 논문에서는 공분산 방법에 콜레스키 분해를 이용하여 선형예측기의 계수를 계산하였고, 그 결과를 다항 예측기와 비교하여 예측 에러가 최소화되는 선형예측기를 선택하도록 하였다. 엔트로피 부호화기는 골롬-라이스 부호를 사용하였고, 골롬-라이스 부호화기의 매개변수를 계산하기 위해 블록기반 매개변수 예측 방법과 LOCO-I, RLGR의 순차 적용 방법을 적용하였다. 실험 결과 블록기반 매개변수 예측 방법과 제안 방식의 예측기를 이용하면 자기상관 방법과 레빈슨-더빈을 사용하는 FLAC 무손실 부호화기보다 2.2879%~0.3413% 압축률이 향상되는 결과를 나타내었고, 제안 방식의 예측기와 LOCO-I 순차 적용 방법을 이용한 경우는 2.2381%~0.0214% 압축률이 향상되는 결과를 나타내었다. 그러나 제안 방식의 예측기와 RLGR 순차 적용 방법을 이용한 경우는 특정 신호에서만 압축률이 향상되었다.

## Design of a Lossless Audio Coding Using Cholesky Decomposition and Golomb-Rice Coding

Cheon-Dae Cheong<sup>†</sup>, Jae-Ho Shin<sup>\*\*</sup>

## ABSTRACT

Design of a linear predictor and matching of an entropy coder is the art of lossless audio coding. In this paper, we use the covariance method and the Cholesky decomposition for calculating linear prediction coefficients instead of the autocorrelation method and the Levinson-Durbin recursion. These results are compared to the polynomial predictor. Both of them, the predictor which has small prediction error is selected. For the entropy coding, we use the Golomb-Rice coder using the block-based parameter estimation method and the sequential adaptation method with LOCO-I and RLGR. The proposed predictor and the block-based parameter estimation have 2.2879%~0.3413% improved compression ratios compared to FLAC lossless audio coder which use the autocorrelation method and the Levinson-Durbin recursion. The proposed predictor and the LOCO-I adaptation method could improved by 2.2879%~0.3413%. But the proposed predictor and the RLGR adaptation method got better results with specific signals.

**Key words:** Lossless audio coding(무손실 오디오 부호화), Cholesky decomposition(콜레스키 분해), Levinson-Durbin recursion(레빈슨-더빈 반복), Golomb-Rice coder(골롬-라이스 부호화), LOCO-I(로코-아이), Sequential adaptation(순차 적용)

※ 교신저자(Corresponding Author): 신재호, 주소: 서울시 중구 필동3가 26번지(100-715), 전화: 02)2260-3336, FAX: 02)2277-1274, E-mail: jhshin@dongguk.edu  
접수일: 2008년 6월 27일, 완료일: 2008년 9월 16일

<sup>†</sup> 정희원, 동국대학교 전자공학과 박사과정  
(E-mail: jdc1004@dongguk.edu)

<sup>\*\*</sup> 정희원, 동국대학교 전자공학과 교수

## 1. 서 론

무손실 오디오 부호화 기술은 맨 처음 오디오 제작 분야나 오디오 파일들에 의해 주로 이용되었으나, 손실 오디오 부호화기에 의해 생성된 음원의 음질에 만족하지 못하는 사용자들을 중심으로 무손실 음원 및 무손실 오디오 부호화에 대한 관심이 늘어나고 있다.

무손실 부호화란 원본 데이터를 압축 혹은 부호화한 상태에서 원래의 정보를 완벽히 복원하는 것을 말하고, 일부의 정보를 잃어버린 상태로 복원하는 경우 손실 부호화라 한다. 현재 인터넷에서 많이 사용되는 무손실 오디오 부호화기는 FLAC [1], Monkey's audio [2], WavPack [3] 등이며, MS 사의 무손실 오디오 부호화기인 무손실 WMA, APPLE 사의 무손실 오디오 부호화기인 ALAC 등 주요 회사에서도 무손실 오디오 부호화기를 제작하였으며, MPEG-4 ALS(Audio Lossless)[4] 국제 표준이 진행 중이다.

대부분의 무손실 오디오 부호화기는 선형 예측 부호화기와 엔트로피 부호화기로 이루어져 있다. 어떠한 선형 예측 부호화기와 엔트로피 부호화기를 사용하는가에 따라 그 성능이 결정된다.

선형 예측 부호화 방법에는 크게 자기상관 방법과 공분산 방법으로 나눌 수 있다. 선형 예측 부호화 방법 중 자기상관 방법은 레빈슨-더빈의 고속 알고리즘을 이용할 수 있고, 합성필터의 출력이 항상 안정적이라는 장점이 있다. 반면에, 자기상관 방법은 입력신호에 윈도우를 사용해야 함으로 경계효과 등으로 인하여 필터 출력의 정확성이 떨어지고, 입력신호가 임펄스응답인 경우에 필터 출력의 정확성이 더욱 떨어지는 단점이 있다[5].

본 논문에서 무손실 오디오 부호화기를 설계함에 있어서 공분산 방법에 콜레스키 분해(Cholesky decomposition)를 이용한 선형예측기와 SHORTEN [6]의 다항 예측기를 이용하였고, 두 예측기 중에 예측 에러가 최소화되는 선형예측기의 계수와 선형예측기의 종류를 부호화하도록 하였다. 콜레스키 분해를 이용한 공분산 방법은 합성필터가 안정적이지 않을 수 있는 단점이 있지만 자기상관 방법에 비하여 정확한 예측 결과를 얻을 수 있어서 예측 에러의 에너지를 최소화 할 수 있는 장점이 있다[7]. 또한, 출력 필터의 안정성 문제로 인하여 예측 에러의 에너지가

커지는 경우 다항 예측기가 선택되어짐으로 전체 부호화기의 안정성에는 문제가 없도록 설계하였다.

무손실 오디오 부호화의 엔트로피 부호화기는 콜롬-라이스 부호화 방법이 가장 많이 사용된다. 콜롬-라이스 부호화 방법은 크게 블록기반 매개변수 예측 방법[8]과 순차 적응 방법[9,10]으로 나눌 수 있다. 콜롬-라이스 부호화에서 블록기반 매개변수 예측 방법은 입력 프레임의 평균이나 통계적 특성을 이용하여 부호화할 프레임에 사용할 매개변수를 계산한다 [8]. 순차 적응 방법에는 LOCO-I [9], RLGR [10], 프레임 평균 이용 방법 등이 있다. 블록기반 매개변수 예측 방법을 이용하면 모든 프레임의 매개변수 정보를 저장해야하는 단점이 있는 반면에 복호화 시에 단일 매개변수를 이용함으로써 복호화 속도가 빠른 장점이 있다. 순차 적응 방법은 매개변수에 대한 정보를 저장할 필요가 없어서 부호화 시 저장 공간을 절약할 수 있는 장점이 있는 반면에 부호화 및 복호화 시에 매 신호 별로 매개변수를 계산해야 함으로 속도가 떨어지는 단점이 있다. 본 논문에서는 블록기반 매개변수 예측 방법과 LOCO-I, RLGR 순차 적응 방법을 비교 검토하였다.

인터넷의 속도향상 및 저장기술의 발전과 더불어 무손실 부호화 기술은 다양한 분야에 적용되고 있다. 정지영상 분야에서는 JPEG-LS [9] 등의 국제표준이 진행 중이며, 의료분야에서도 의료영상이나 ECG (electrocardiogram) 데이터를 저장하는데 무손실 부호화 기술을 적용하고 있다. 이와 같이 고음질 및 고화질을 특징으로 하는 미디어 기술은 점점 더 무손실 부호화 기술을 이용하는 방향으로 바뀌고 있다. 인터넷의 전송대역폭이 점점 더 늘어나고 있지만, 많은 양의 데이터를 효율적으로 전송하거나 저장하기 위해서는 무손실 부호화 기술에 대한 지속적인 연구가 필요하다.

본 논문의 2장에서는 무손실 오디오 부호화에 대하여 개괄적으로 설명한다. 3장에서는 무손실 오디오 부호화기의 핵심이라고 할 수 있는 선형 예측 부호화 및 콜레스키 분해를 이용한 공분산 방법에 대하여 설명한다. 4장에서는 콜롬-라이스 엔트로피 부호화 및 블록기반 매개변수 예측 방법과 순차 적응 방법에 대하여 설명한다. 5장에서는 본 논문에서 제안하는 콜레스키 분해를 이용한 공분산 방법 및 다항 예측기를 이용한 예측기에 대하여 설명하고, 블록기

반 대개변수 예측 방법과 LOCO-I, RLGR 순차 적응 방법을 적용한 경우에 대하여 실험결과를 설명하고, 6장에서 결론을 맺는다.

## 2. 무손실 오디오 부호화

일반적으로 무손실 오디오 부호화기는 그림 1과 같이 입력신호를 작은 블록으로 나누는 프레임 단계, 프레임 내의 상관관계를 제거하는 상관관계 제거 단계, 그리고 엔트로피 부호화 단계로 이루어졌다. 많은 오디오 데이터를 한꺼번에 처리하기 힘들므로 일정한 블록 크기로 입력 오디오 신호를 나누게 된다. 이렇게 나눈 블록을 프레임 혹은 세그먼트라고 하며, 일반적으로 프레임 크기가 커질수록 압축률이 향상되지만 선형 예측 부호화나 변환 부호화 시의 성능과 상쇄관계를 가짐으로 적당한 크기를 가지도록 설계해야 한다.

오디오 신호를 시공간에 나타내면 시간에 대하여 연속적인 파형으로 나타난다. 인접한 샘플들 사이의 값들이 유사하여 상관관계가 있는 것처럼 보인다. 그러나 여기에 LZW(Lempel-Ziv-Welch)와 같이 모든 샘플들이 독립적이라고 가정하는 만능 엔트로피 부호화기를 적용하게 되면 실제 엔트로피보다 더 높은 엔트로피를 가지기도 한다. 만약, 신호가 진정으로 랜덤하다면, 어떠한 장기상관(long-term correlation) 특성이나 저주파 대역에 에너지가 집중되는 현상 등은 나타나지 않게 된다. 그러나 오디오 신호는 인접 프레임 간에 유사한 파형을 나타내는 장기상관 특성을 가지고 있고, 이러한 장기상관 특성에 의한 중복성 등을 제거하기 위하여 모든 오디오 부호화기는 엔트로피 부호화 이전에 상관관계 제거 단계를 거치게 된다.

상관관계 제거 단계에 주로 사용하는 방법은 선형 예측 부호화와 변환 부호화 방법이 주로 사용된다. 무손실 오디오 부호화기 설계에 있어서 상관관계 제거 단계가 가장 핵심 분야이며, 사용하는 방법에 따

라 예측기반 오디오 부호화 및 변환기반 오디오 부호화로 나눌 수 있다.

현재 대부분의 무손실 오디오 부호화기는 변환기반 오디오 부호화 방법 보다는 예측기반 오디오 부호화 방법이 선호되고 있다. Liebchen [4]의 경우 맨 처음에는 DCT 기반의 변환 오디오 부호화 방법을 사용하는 MPEG-4 ALS를 제안하였다가, 나중에는 선형 예측 부호화 방법을 사용하는 MPEG-4 ALS를 제안하였다. Malvar [10]는 LOT(Lapped Orthogonal Transform)와 RLGR 엔트로피 부호화 방법을 이용하는 변환기반 오디오 부호화 방법을 제안하였고, 현재 MS사의 무손실 WMA에 사용되고 있다.

선형 예측 부호화는 오디오 신호처리나 음성처리에서 주로 사용하는 기법으로써, 선형 예측 모델의 정보를 이용하여 오디오 신호나 음성 신호를 압축된 형태로 표현하는 기법이다. 선형 예측 모델은 현재의 샘플 값을 이전 샘플 값들의 선형조합을 이용해 예측한다. 원 신호 값과 선형 예측 모델에 의해 생성된 예측 신호 값과의 차이가 에러 값이 되며, 오디오 부호화기는 이 에러 값을 양자화 및 엔트로피 부호화하여 전송한다. 원 신호 값 대신에 에러 값만 전송됨으로 원 신호의 크기가 많이 줄어드는 효과가 있다. 또한, 이 에러 값은 원 신호에 비해 작은 편차를 가지고, 원 신호에 비해 평평한 스펙트럼을 가지므로 양자화 및 엔트로피 부호화하기에 적합한 형태가 된다[5].

엔트로피 부호화란 정보이론에서 나오는 무손실 데이터 압축 기법으로써, 오디오, 영상, 문서 데이터 압축 등의 다양한 분야에 이용되고 있다. 신호는 어떠한 정보량을 가지고 있는데, 나타나는 빈도가 낮으면 정보량이 커지고, 나타나는 빈도가 높으면 정보량이 작아진다. 즉, 나타나는 빈도가 높은 신호는 작은 정보량으로 표현할 수 있다. 엔트로피 부호화에서 가장 많이 사용되는 기법은 런-LENGTH(run-length) 부호화, 허프만(Huffman) 부호화, 산술 부호화 등이 있다.

데이터의 통계적 특성이 기하분포를 가지는 경우 곱셈 부호화[8] 등의 방법이 많이 사용된다. 이 외에

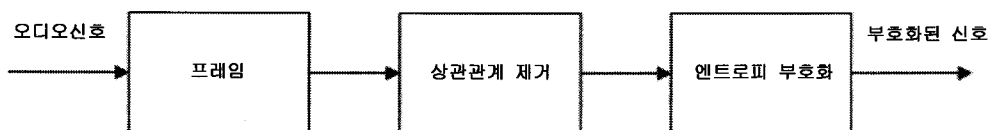


그림 1. 무손실 오디오 부호화기

Monkey's audio에서는 레인지(range) 부호화기를 사용하고[2], MPEG-4 ALS에서는 BGMC(Block Gilbert Moore Coding)와 콜롬-라이스 부호화 중에 선택하도록 하고 있다[4].

### 3. 선형예측 부호화

#### 3.1 선형 예측 모델

선형 예측 모델 혹은 자기회귀(AR) 모델에서 현재 샘플  $x(n)$ 은 식 (1)과 같이  $M$ 개의 이전 신호 샘플들의 선형 조합으로 표현할 수 있다.

$$\hat{x}(n) = \sum_{i=1}^M a_i x(n-i) \tag{1}$$

계수  $a_i$ 가 주어진다면, 식 (1)은  $M$ 차의 FIR(finite impulse response) 시스템으로 간주할 수 있다. 선형 예측에서는 이 계수 값을 예측하는 방법이 가장 중요한 부분이며 크게 자기상관 방법과 공분산 방법으로 나눈다.

#### 3.2 자기상관 방법

자기상관 방법은 식 (2)와 같이 무한 구간에서 예측 에러를 최소화 하는 방법이다.

$$E = \sum_{-\infty}^{\infty} |x(n) - \hat{x}|^2 \tag{2}$$

식 (2)의 예측 에러  $E$ 를 최소화하기 위해 편미분을 재료로 두면, 식 (3)의 위너-호프(Wiener-Hopf) 방정식을 얻을 수 있다.

$$\sum_{k=1}^M r(k-i)a_k = r(i), \text{ for } i=1, \dots, M$$

$$\text{where, } r(i) = E\{x(n)x(n-i)\} = \sum_{n=0}^{N-1+i} x(n)x(n-i) \tag{3}$$

식 (3)은 자기회귀 모델에서 자기상관 함수에 대한 율-워커(Yule-Walker) 방정식과 동일하다. 식 (3)에서  $r(i)$ 는 입력신호의  $i$ 번째 자기상관 함수이고,  $r(0)$ 는 입력 신호의 편차가 된다. 식 (3)의 해법에는 가우시안 소거법, QR 분해, 콜레스키 분해 등이 있다. 어떠한 방법을 이용하던지  $O(N^3)$ 의 곱셈이 필요하지만, 토프리트(Toeplitz) 행렬인 경우 레빈슨-더빈 알고리즘을 사용할 수 있고, 이 경우  $O(N^2)$ 의 복잡도로 낮출

수 있다. 자기상관 방법을 이용할 경우 반사 계수의 범위가 항상 1이하여서 합성필터가 안정적인 수 있다는 장점이 있다. 반면에 식(2)를 보면 입력신호는 유한하지만 무한 구간에서 예측 에러를 계산함으로 부가적으로 윈도우를 사용해야 되며, 이로 인한 경계효과가 필연적으로 발생한다. 그 결과 필터 출력의 정확성이 떨어지게 되며, 임펄스와 같은 신호에 대해서는 필터 출력의 정확성이 더욱 떨어지게 된다.

#### 3.3 공분산 방법

공분산 방법은 식 (4)와 같이 유한 구간에서 예측 에러를 최소화하는 방법을 사용하다.

$$E = \sum_{n=0}^{M-1} |x(n) - \hat{x}(n)|^2 \tag{4}$$

식 (4)를 최소화하는 율-워커 방정식은 식 (5)와 같이 구할 수 있다.

$$\sum_{k=1}^M c(k,i)a_k = c(i,0) \tag{5}$$

$$\text{where, } c(k,i) = \sum_{n=0}^{M-1} x(n-i)x(n-k)$$

식 (5)의 해는 가우시안 소거법, 콜레스키 분해 등을 이용해 구할 수 있다. 식 (5)를 행렬형태로 표시하면 식 (6)과 같다.

$$Ca = c \tag{6}$$

식 (6)에서 행렬  $C$ 는 공분산 함수  $c(i,j)$ 를 원소로 가지는 행렬이고, 행렬  $a$ 와  $c$ 는 각각  $a_j$ 와  $c(i,0)$ 를 열벡터로 가지는 행렬이다. 콜레스키 분해를 하기 위해서 행렬  $C$ 는 식 (7)을 만족하는 양의 대칭 행렬이어야 한다.

$$C = VDV^T \tag{7}$$

행렬  $V$ 와 행렬  $D$ 의 원소는 식 (8)을 이용해서 구한다.

$$V_{ij}D_j = c(i,j) - \sum_{k=1}^{j-1} V_{ik}D_k V_{jk}, \tag{8}$$

$$1 \leq j \leq i-1$$

$$D_i = c(i,i) - \sum_{k=1}^{i-1} V_{ik}^2 d_k, d_1 = c(1,1), i \geq 2$$

식 (6)의 해는 식 (9)를 이용하여 계산할 수 있고, 식 (10)과 같이 선형예측기의 계수  $a_i$ 를 구할 수 있다.

$$Y = DV^T a, \quad Y_1 = c_i$$

$$Y_i = c_i - \sum_{j=1}^{i-1} V_{ij} Y_j, \quad 2 \leq i \leq M \quad (9)$$

$$a_i = \frac{Y_i}{d_i} - \sum_{j=i+1}^M V_{ji} a_j, \quad 1 \leq i \leq M-1 \quad (10)$$

공분산 방법에서 행렬 C는 양의 대칭행렬이기는 하지만 일반적으로 토플리츠 행렬이 아니기 때문에 레빈슨-더빈 알고리즘을 이용할 수 없고, 가우시안 소거법이나 콜레스키 분해를 이용한다[5]. [11]에서는 콜레스키 분해를 이용한 공분산 방법에서 행렬 C가 토플리츠 행렬일 경우 자기상관 방법의 반사 계수와 동일한 역할을 수행하지만, 행렬 C가 토플리츠 행렬이 아닌 경우 이를 자기상관 방법의 반사 계수와 동일 시 하는 것의 문제에 대하여 지적하였다.

공분산 방법은 합성필터의 안정성을 보장할 수 없지만, 자기상관 방법에 비하여 필터 출력의 정확성을 보장할 수 있어서, 예측 에러의 에너지를 줄일 수 있다 [7,11]. 콜레스키 분해를 자기상관 방법의 토플리츠 행렬에 적용할 수도 있지만, 일반적으로 고속 알고리즘을 이용할 수 있는 레빈슨-더빈 방법을 선호한다[7]. [12]에서는 공분산 방법과 콜레스키 분해를 MPEG-4 ALS에 적용하였고, 필터가 불안정할 경우 자기상관 방법과 레빈슨-더빈 알고리즘을 사용하도록 하여 기존의 MPEG-4 ALS보다 0.1%의 성능을 개선하였다.

### 3.4 다항 예측기

SHORTEN [6]에서는 선형예측기외에 식 (11)의 정수 계수만을 사용하는 간단한 다항 예측기를 이용한다.

$$\begin{aligned} \hat{x}_0(n) &= 0 \\ \hat{x}_1(n) &= x(n-1) \\ \hat{x}_2(n) &= 2x(n-1) - x(n-2) \\ \hat{x}_3(n) &= 3x(n-2) - 3x(n-1) \\ &\quad + x(n-3) \end{aligned} \quad (11)$$

이 예측기의 예측 에러  $e_i(n) = x(n) - \hat{x}_i(n)$  는 식 (12)와 같은 반복 방법으로 구할 수 있다.

$$\begin{aligned} e_0(n) &= x(n) \\ e_1(n) &= e_0(n) - e_0(n-1) \\ &= x(n) - x(n-1) \\ e_2(n) &= e_1(n) - e_1(n-1) \\ &= x(n) - 2x(n-1) + x(n-2) \\ e_3(n) &= e_2(n) - e_2(n-1) \\ &= x(n) - 3x(n-1) \\ &\quad + 3x(n-2) - x(n-3) \end{aligned} \quad (12)$$

식 (12)를 이용하여 한 프레임에서 절대 평균 합을 계산하고, 이를 최소화 하는 예측기를 선택한다. 선택된 예측기에 대한 정보는 부호화 시에 포함된다. 무음 프레임의 경우  $e_0(n)$ 나  $e_1(n)$  만으로 프레임을 부호화할 수 있어서 엔트로피 부호화 시 압축률을 많이 줄일 수 있는 장점이 있다.

SHORTEN의 다항 예측기는 MPEG4-SA (structured audio)와 AudioPaK, FLAC 등에도 사용되었다.

## 4. 엔트로피 부호화

### 4.1 골롬-라이스 부호

골롬 부호는 허프만 부호와 마찬가지로 가변길이 부호이지만, 허프만 부호가 데이터 자체에 기반을 두는데 비해, 골롬 부호는 확률분포에 기반을 둔 부호화 기법이다. 골롬 부호는 확률분포가 지수적으로 감소하는 확률모델에서 최적 부호화 방법으로 알려져 있다[8].

골롬 부호는 맨 처음 양의 정수  $m$ 이 단일 매개변수로 주어진다. 양의 정수  $m$ 은 식 (13)에서 부호화할 정수  $n$ 을 몫과 나머지로 표현할 때 제수의 역할을 한다.

$$\begin{aligned} n &= qm + r = \left\lfloor \frac{n}{m} \right\rfloor m + r \\ q &= \left\lfloor \frac{n}{m} \right\rfloor \\ r &= n - qm \\ c &= \lceil \log_2 m \rceil \end{aligned} \quad (13)$$

부호화할 정수  $n$ 은 식 (13)을 이용하여,  $q$ ,  $r$ ,  $c$ 의 세 값을 구하고,  $q$ 는 일진법(0의 시퀀스)으로 부호화하고,  $r$ 은 맨 처음  $r$ 의  $c-1$  비트는  $2^{c-m}$  값을 양의 정수로 이진 부호화하고, 나머지  $c$  비트들은 이진 부호화한다. 0의 시퀀스 다음에는 정지비트 1을 두어서 나머지 부분과 구분한다.

골롬-라이스 부호는 골롬 부호화에서 매개변수  $m$ 이 2의 멱승( $m=2^k$ )인 경우에 해당한다. 이 경우, 골롬 부호에서 나머지  $r$ 의  $c-1$  비트 부분은  $2^{c-m}=0$ 이 됨으로 부호화하지 않아도 된다. 그리고 식 (13)의 나누기와 모듈로 연산은 쉬프트와 마스크 연산만으로 계산이 가능하기 때문에 고속 알고리즘이나 하드웨어 구현 시 용이한 장점이 있다. FLAC, ALAC 등

의 무손실 오디오 부호화기는 콜롬-라이스 엔트로피 부호화 방법이 사용되었다.

4.2 블록기반 매개변수 예측 방법

매개변수  $k$ 를 이용하여 양의 정수  $s$ 을 콜롬-라이스 부호화할 경우  $\lfloor s/2^k \rfloor + 1$  비트의 일진법 부분과  $k$  비트의 나머지 부분이 필요하므로 전체 길이는 식 (14)와 같이 구할 수 있다.

$$l(s, k) = 1 + k + \left\lfloor \frac{s}{2^k} \right\rfloor \quad (14)$$

기하분포를 가지는 소스의 단일 매개변수가  $\rho$ 인 경우, 식 (15)를 이용하여 최적 매개변수  $k$ 를 구할 수 있다.

$$k(\rho) = \max \left\{ 0, \left\lceil \log_2 \frac{\log(\phi)}{\log(\rho^{-1})} \right\rceil \right\}, \quad (15)$$

where,  $\phi = \frac{\sqrt{5} + 1}{2}$

소스의 평균이  $\bar{s}$ 인 경우, 식 (16)을 이용하여 최적 매개변수  $k$ 를 구할 수 있다[8].

$$k(\bar{s}) = \max \left\{ 0, \left\lceil \log_2(\bar{s}) - 0.05 + \frac{0.6}{\bar{s}} \right\rceil \right\} \quad (16)$$

오디오 부호화 시에 프레임 단위로 처리함으로써 소스의 평균  $\bar{s}$ 를 계산하기가 쉽다. 소스의 발생확률  $\rho$ 를 알아야 하는 식 (15)보다 소스의 평균을 이용할 수 있는 식 (16)이 최적 매개변수  $k$ 를 계산하는데 더 편리하다.

4.3 순차 적응 방법

블록기반 매개변수 예측 방법은 부호화 시 매개변수의 정보를 포함시켜야 하고, 블록 별로 평균이나 소스의 발생확률을 계산해야 하는 단점이 있다. 반면에 순차 적응 방법을 이용하면 이전에 부호화한 신호의 정보를 이용하여 현재 부호화할 신호에 사용할 매개변수 값을 결정한다. 이 방법을 이용하면 매개변수에 대한 정보를 저장할 필요 없이 부호화 및 복호화가 가능하다.

JPEG-LS에서는 Wienberger의 LOCO-I (Low COMplexity LOSSless COMpression for Images) 방법을 이용하여 매개변수를 결정한다[9]. 현재 샘플  $x_{t+1}$ 를 부호화하기 위해 식 (17)과 같이  $S_t$ 와  $N_t$ 를

정의한다.  $S_t + N_t$  값은 지금까지 샘플들에 대한 절대 합이 되며,  $N_t$ 는 지금까지 음의 샘플의 개수가 된다.

$$S_t = \sum_{i=1}^t (|x_i| - u(x_i)), \quad N_t = \sum_{i=1}^t u(x_i) \quad (17)$$

$S'(t)$ 를 식 (18)과 같이 정의한다.  $S'(t) \leq 2t$  인 경우,  $S(t)$ ,  $N(t)$ ,  $t - N(t)$  의 세 값을 비교하여,  $S(t)$ 가 가장 큰 경우  $k=1$ 을 매개변수로 하여 식 (19)의  $M(x_i)$  값을 콜롬-라이스 부호화한다.  $t - N(t)$ 가 가장 큰 경우  $k=0$ 을 매개변수로 하여 식 (19)의  $M(x_i)$  값을 콜롬-라이스 부호화하고,  $N(t)$ 가 가장 큰 경우  $k=1$ 을 매개변수로 하여 식 (20)의  $M(x_i)$  값을 콜롬-라이스 부호화한다.  $S'(t) > 2t$  인 경우 식 (21)을 만족하는  $k$ 를 구한 다음,  $k+1$ 을 매개변수로 하여 콜롬-라이스 부호화 한다.

$$S'(t) = S(t) + \frac{t}{2} - \frac{t}{8} \quad (18)$$

$$M(x_i) = 2|x_i| - u(x_i), \quad u(x_i) = 1 \text{ if } x_i < 0. \quad (19)$$

$$M(x_i) = M(-x_i - 1) \quad (20)$$

$$t2^k \leq S'(t) < t2^{k+1} \quad (21)$$

식 (18)은 양방향 기하분포를 가지는 소스에서 과거의 입력을 이용하여 최적 매개변수  $k$ 를 계산하기 위한 최대가능 예측기(maximum likelihood estimator)이다.

[10]에서는 런-랭스 부호화 및 콜롬-라이스 부호화를 이용한 RLGR 부호기에 대하여 제안하였다. RLGR 부호화기에서는 긴 부호를 출력한 경우 매개변수  $k$ 를 상향 조절하고, 짧은 부호를 출력하는 경우 매개변수  $k$ 를 하향 조절하는 순차 적응 방식을 이용하여 매개변수  $k$ 를 계산한다. 현재 값  $x$ 를 매개변수  $k$ 를 이용하여 부호한 경우, 식 (22)를 이용하여 매개변수  $p$ 를 계산한 다음, 매개변수  $p$ 값에 따라서 식 (23)과 같이  $k_r$  매개변수 값을 조절한다. 그 다음 샘플을 부호화할 경우 식 (24)에 의해 갱신된 매개변수  $k$  값을 이용하여 부호화한다.

$$k_r = kL, \quad p = \left\lfloor \frac{x}{2^{k_r}} \right\rfloor = x \gg k_r \quad (22)$$

$$\begin{aligned} \text{if } p < 0, k_r &= k_r - 2 \\ \text{if } p = 0, k_r &= k_r \\ \text{if } p > 0, k_r &= k_r + p + 1 \end{aligned} \quad (23)$$

$$k = \frac{k_r}{L} \quad (24)$$

[10]에서는 상수  $L$ 의 값은 4를 사용하도록 권고하고 있다.  $L$ 값은 스케일된 매개변수  $k_r$ 에 따라서 적응 속도를 조절하는 매개변수의 역할을 한다.

LOCO-I 순차 적응 방법은 JPEG-LS 표준에 채택되었고, RLGR 순차 적응 방법은 MS 사의 무손실 WMA 등에 사용되고 있다.

### 5. 제안 방식

본 논문에서는 선형예측기의 예측 계수를 계산하기 위해 공분산 방법에 콜레스키 분해를 이용하였고, 이를 다항 예측기와 비교하여 예측 에러를 최소화하는 예측기를 선택하여 예측 계수와 예측기 종류를 저장하도록 하였다. 예측 에러를 엔트로피 부호화하는데 콜롬-라이스 부호화 방법을 이용하였다. 콜롬-라이스 부호화 시 식 (16)을 이용한 블록기반 매개변수 예측 방법과 LOCO-I 순차 적응 방법 및 RLGR 순차 적응 방법을 이용하였다. 그림 2에 제안 방식의 블록 다이어그램을 나타내었다.

다항 예측기가 선택되는 경우 다항 예측기의 종류 (0~3)가 부가정보로 저장되고, 선형예측기가 선택

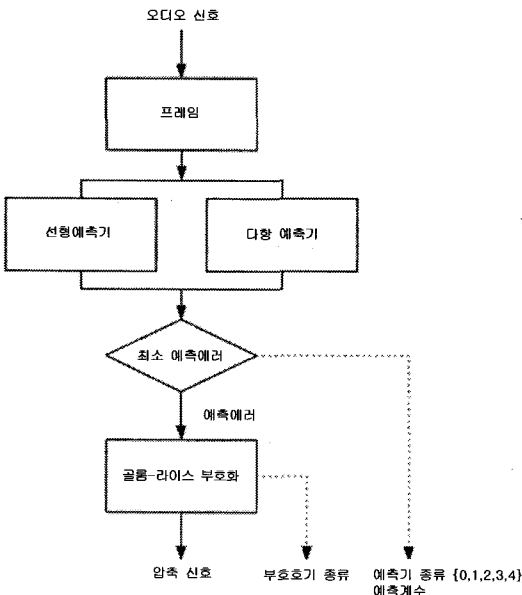


그림 2. 제안 알고리즘의 블록 다이어그램

표 1. SQAM 데이터

번호	신호종류	시간	파일크기 (바이트)
1	Electronic tune	0:35	6169508
2	Violoncello	0:37	5303964
3	Trumpet	0:23	3146652
4	Horn	0:31	4576268
5	Glockenspiel	0:35	4568104
6	Harpsichord	0:23	2891112
7	Soprano	0:28	4171220
8	Bass	0:30	4385644
9	Quartet	0:28	4070604
10	Female speech	0:23	3384860
11	Male speech	0:22	3167160

되는 경우 예측기 종류는 4가 되도록 하였다. 무음 프레임인 경우 다항 예측기 0번이 선택되고, 이 경우 예측 계수를 저장하지 않는다.

실험에 사용한 데이터는 표 1의 EBU SQAM (Sound Quality Assessment Material) 데이터를 이용하였다.

#### 5.1 선형예측기 필터의 차수

선형예측기에서 필터의 차수를 결정하는 것은 매우 중요한 문제이다. 에러 신호의 편차를 작게 하기 위해 예측 차수를 증가시키면 디코더와 인코더의 복잡도가 높아지고, 처리해야할 데이터 및 매개변수의 개수도 증가한다. 반면에 차수가 높아지면 신호의 변화가 심하더라도 예측을 잘 할 수 있는 장점이 있다.

필터의 차수  $M$ 을 결정하는 문제는 식 (25)와 같이 에러 신호의 평균에 대한 편차를 최소화하는 것이다.

$$\sigma_e^2(M) = \frac{1}{N-M-1} \sum_{k=M}^{N-1} e^2(n) \quad (25)$$

최적 필터 차수는 통계모델의 모델 선택을 이용하여 결정하는데, 그 기준을 정보기준이라고 한다. 정보기준에는 식 (26)의 AIC(Akaike Information Criteria), 식 (27)의 PHI(Pukkila Hannan-Quinn Information Criteria), 식 (28)의 MDL(minimum description length) 등의 방법이 있다[13].

$$AIC(M) = N \ln \sigma_e^2(M) + 2M \quad (26)$$

$$\Phi(M) = N \ln \sigma_e^2(M) + 2M \ln(\ln N) \quad (27)$$

$$MDL(M) = N \ln \sigma_e^2(M) + M \ln N \quad (28)$$

그림 3은 샘플 오디오 신호(violin.wav, 44.1k 샘플링)의 한 프레임(1024개)에 대하여 필터 차수의 변화에 따른 정보기준을 나타내었다. AIC나 PHI의 경우 최적 필터 차수가 31~32에서 최소화가 되는데, MDL의 경우는 3에서 최적 필터 차수를 나타내었다.

표 2를 적용한 실험에서도 필터 차수를 0~32개로 늘려가면서 테스트 해 보았는데, 실제로 32개 이상에서 가장 높은 압축률을 나타내었으나, 필터 차수가 12개 이상에서는 1 차수 증가 할 때 마다 파일당 수백 바이트 정도의 절감 효과 밖에는 얻을 수 없었다. 32

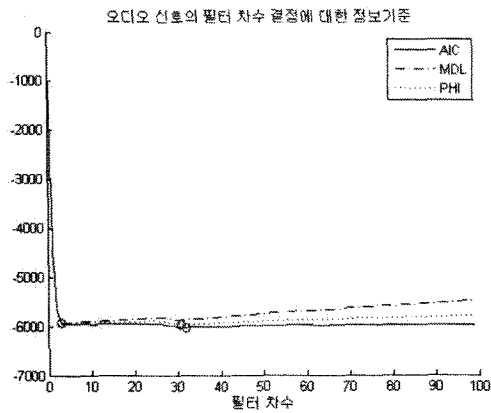


그림 3. 오디오 신호의 필터 차수 결정에 대한 정보기준

표 2. 자기상관 방법에 레빈슨-더빈을 이용하는 경우의 실험 결과

번호	블록기반 예측방법 (압축률%)	LOCO-I (압축률%)	RLGR (압축률%)
1	16.5037	16.8163	17.2584
2	39.1655	39.2419	41.4886
3	31.7345	31.7826	33.9150
4	28.2593	28.2890	30.6118
5	36.2271	36.2919	38.6747
6	47.7642	47.8289	49.9174
7	41.7353	41.8330	43.8998
8	40.9592	41.1214	43.3247
9	46.6459	46.6803	49.0536
10	34.9552	35.3070	36.4440
11	36.7871	36.9930	38.5747

개 이상에서는 오히려 압축률이 더 줄어드는 경우도 있었다. 제안 방식에서는 공분산 방법에 클래스키 분해를 이용한 예측 계수 계산 시 필터 차수를 32로 하였다.

### 5.2 제안 방식과 자기상관 방법 비교

제안 방식에서는 선형예측기와 다항 예측기 중에 예측 에러를 최소화하는 예측기를 선택하도록 하였다. 실험결과 다항 예측기가 선택되는 경우는 전체 프레임 중에 1% 미만이었고, 심지어는 맨 앞의 두 프레임 정도만 다항 예측기가 선택되고 나머지는 전부 선형예측기가 선택되어서 다항 예측기를 도입한 효과가 미비한 신호도 있었으나, 무음 프레임이 많은 신호에서는 다항 예측기로 인한 압축효과가 커졌다.

그림 4~그림 6과 표 2~표 3은 블록기반 매개변수 예측 방법, LOCO-I 순차 적용 방법, RLGR 순차

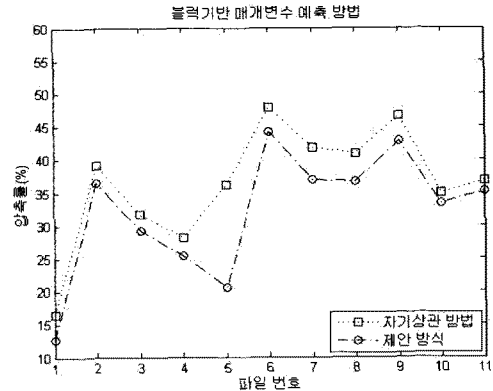


그림 4. 블록기반 매개변수 예측 방법

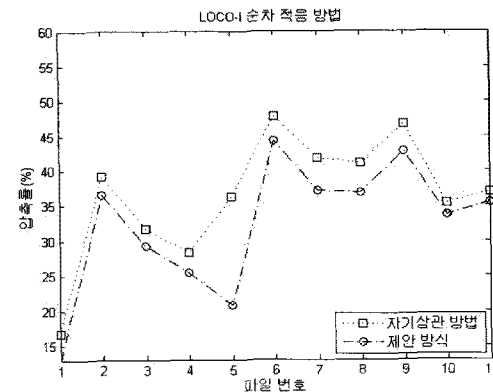


그림 5. LOCO-I 순차 적용 방법



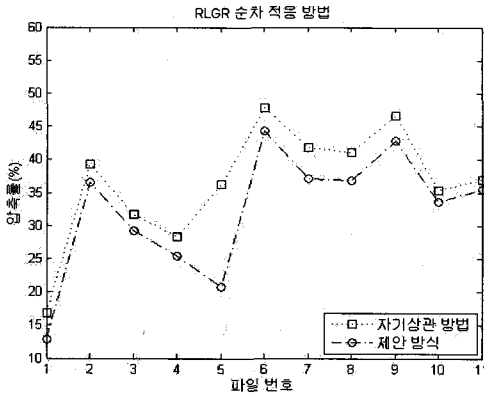


그림 6. RLGR 순차 적응 방법

표 3. 제안 방식의 실험결과

번호	블록기반 예측방법 (압축률%)	LOCO-I (압축률%)	RLGR (압축률%)
1	12.6920	12.8450	12.9247
2	36.4832	36.5312	39.1437
3	29.2116	29.2619	31.5317
4	25.3940	25.4438	27.8887
5	20.5385	20.6986	21.5725
6	44.1923	44.2625	47.8051
7	37.0095	37.0694	39.5613
8	36.6726	36.8128	39.4231
9	42.7892	42.8212	45.5866
10	33.3604	33.6804	35.0796
11	35.2480	35.4457	37.2883

적용 방법에 대하여 자기상관 방법에 레빈슨-더빈을 이용하는 경우와 제안 방식의 압축률을 비교하였다.

제안 방식을 사용한 경우 자기상관 방법보다 블록기반 매개변수 예측 방법의 경우 15.6886%~1.5391% 압축률이 향상되었고, LOCO-I 순차 적응 방법의 경우 15.5934%~1.5473%, RLGR 순차 적응 방법의 경우 17.1022%~1.2863% 압축률이 향상되었다. 세가지 방법 모두 5번 신호에서 제안 방식이 가장 높은 압축률을 나타내었고, 11번 신호에서 제일 적은 압축률을 나타내었다. 유독 높은 결과가 나타난 5번 신호를 제외하면 제안 방식이 일반적으로 3~4% 정도 자기상관 방법보다 높은 압축률을 나타내었다.

제안 방식이 자기상관 방법에 비해 15% 이상의 높은 압축률을 나타내었지만, 이는 SQAM 데이터의 특성에 기인한 결과로 보인다. 락 음악이나 클래식

교향곡과 같이 변화가 많은 오디오 신호의 경우 자기상관 방법이 제안 방식보다 좋은 결과를 나타내는 경우도 있었지만, 클래식 소품이나 SQAM 데이터의 경우 제안 방식이 더 좋은 성능을 나타내었다.

### 5.3 제안 방식과 다른 무손실 오디오 부호화기 비교

표 4는 무손실 압축 포맷으로 가장 많이 사용하는 FLAC 1.2.1 및 Monkey's audio 3.99, Wavpack 4.11 무손실 오디오 부호화기를 표 1의 데이터에 적용한 결과를 나타내었다. FLAC과 Monkey's audio, Wavpack 무손실 부호화기는 최대 압축률을 얻을 수 있도록 각각 '-8', 'High Compression', '-hh' 옵션을 적용하였다. FLAC에 사용되는 예측기는 자기상관 방법에 레빈슨-더빈을 사용하는 선형예측기와 다항예측기 중에 선택하도록 되어 있고, 엔트로피 부호화기는 한 프레임을 다중 영역으로 나누고, 다중 영역에서 블록기반 매개변수 예측 방법을 이용하여 곱률-라이스 부호화한다.

표 3 및 표 4에서 알 수 있듯이 제안 방식에 블록기반 매개변수 예측 방법을 사용한 경우 FLAC 1.2.1 알고리즘 보다 2.2879%~0.3413% 압축률이 향상되었다. LOCO-I 순차 적응 방법을 적용하면 FLAC 1.2.1 보다 2.2381%~0.0214% 압축률이 향상되었다. 반면에 RLGR 순차 적응 방법을 적용하면 최대 1.0976% 좋아지는 경우가 있으나 최대 2.7453% 까지 압축률이 저하되는 경우도 있었다. RLGR 순차 적응 방법을 사용한 경우 1번, 5번 신호 외에 나머지 신호

표 4. 여러 가지 무손실 오디오 부호화기에 대한 실험결과

번호	FLAC (압축률%)	WavPack (압축률%)	Monkey's audio (압축률%)
1	13.7585	17.6797	14.7472
2	37.1065	36.9443	33.6371
3	29.9270	29.7422	26.8000
4	27.6819	28.3045	24.4194
5	22.6701	29.7815	22.6566
6	45.0598	44.4283	39.6163
7	38.8551	37.6450	34.8356
8	37.9958	37.5250	35.4267
9	44.0272	42.9633	41.3951
10	33.7018	32.5720	31.7792
11	35.7154	33.3693	32.6285

는 FLAC 1.21에 비해 더 나쁜 결과가 나왔다.

제안 방식을 표 4의 FLAC 1.21, Monkey's audio 3.99, Wavpack 4.11 무손실 오디오 부호화기와 압축률을 비교해보면, Monkey's audio 무손실 부호화기가 전반적으로 가장 좋은 성능을 나타내었다. 제안 방식은 1번과 5번 신호에서만 가장 좋은 성능을 나타내었다. 그러나 FLAC 1.21과 Wavpack 4.11 보다는 전반적으로 좋은 성능을 나타내었다. Wavpack 4.11의 경우 10번과 11번 신호에서만 제안 방식보다 좋은 성능을 나타내었다.

## 6. 결 론

실험 결과 공분산 방법에 콜레스키 분해를 이용하는 제안 방식이 기존의 자기상관 방법과 레빈슨-더빈을 이용한 방식보다 17.1022%~1.5391% 압축률이 향상되었다. 이는 공분산 방법에 콜레스키 분해를 사용하는 제안 방식이 자기상관 방법에 레빈슨-더빈을 사용하는 방식에 비하여 좀 더 정확한 필터 출력을 보장해서 예측에러가 줄어들었기 때문으로 사료된다.

곱셈-라이스 부호화의 매개변수를 계산하기 위해, 블록기반 매개변수 예측 방법과 LOCO-I 및 RLGR 순차 적용 방법을 사용하였으나 큰 차이점은 없었다.

곱셈-라이스 부호화기와 MPEG-4 ALS의 BGMC와 같은 엔트로피 부호화기를 동시에 사용하는 방법이나, [13]과 같이 자기상관 방법과 공분산 방법을 부호화할 신호의 특성에 따라 전환하는 방법 등을 사용하면 좀 더 좋은 성능의 무손실 오디오 부호화기를 설계할 수 있으리라 사료된다.

## 참 고 문 헌

- [1] <http://flac.sourceforge.net/>
- [2] <http://www.monkeysaudio.com/>
- [3] <http://www.wavpack.com/>
- [4] T. Liebchen, T. Moriya, N. Harada, Y. Kamamoto, and Y. Reznik, "The MPEG-4 Audio Lossless Coding (ALS) Standard - Technology and Applications," *119th AES Convension*, pp. 1-14, 2005.
- [5] J. Makhoul, "Linear prediction: A tutorial review," *Proceedings of The IEEE*, Vol.63, pp. 561-580, 1975.
- [6] T. Robinson, "Shorten: Simple lossless and near-lossless waveform compression," Tech. Rep. CUED/F-INFENG/TR.156, Cambridge University, UK, 1994.
- [7] B.W. Dickinson, "Reflection Coefficient Estimation Using Cholesky Decomposition," *IEEE Trans. on ASSP.*, Vol.27, No.2 pp. 146-149, 1979
- [8] A. Said, "On the Determination of Optimal Parameterized Prefix Codes for Adaptive Entropy Coding," *Hewlett-Packard Lab. Report HPL-2006-74*, Palo Alto, CA., 2006.
- [9] M. J. Weinberger, G. Seroussi, and G. Sapiro, "The LOCO-I Lossless Image Compression Algorithm: Principles and Standardization into JPEG-LS," *IEEE Trans. on Image Processing*, Vol.9, pp. 1309-1324, 2000.
- [10] H. S. Malvar, "Adaptive Run-length/ Golomb-Rice encoding of Quantized Generalized Gaussian Sources with Unknown Statistics," *Proc. Data Compression Conference*, pp. 23-32, 2006.
- [11] J. D. Gibson, "On Reflection Coefficients and the Cholesky Decomposition," *IEEE Trans. on ASSP.*, pp. 93-96, 1977.
- [12] Y. Kamamoto, T. Moriya, N. Harada, and C. Kos, "Enhancement of MPEG-4 ALS Lossless Audio Coding," *NTT Technical Review*, Vol.5, No.12, 2007.
- [13] H. Akaike, "A New Look at the Statistical Model Identification," *IEEE Trans. on Automatic Control*, Vol.19, pp. 716-723, 1974.



정 전 대

1995년 동국대학교 전자공학과  
학사  
1997년 동국대학교 전자공학과  
공학석사  
1997년~현재 동국대학교 전자공  
학과 박사과정  
2002년~2008년 (주)제론블루

2008년~현재 (주)포티스

관심분야 : 신호처리, 음향신호처리, 오디오 부호화



신 재 호

1979년 서울대학교 전자공학과  
학사  
1982년 서울대학교 전자공학과  
공학석사  
1987년 서울대학교 전자공학과  
공학박사  
1983년~1988년 명지대학교 조교수

1988년~현재 동국 대학교 전자공학과 교수

관심분야 : 신호처리, 음향신호처리, 음향위터마킹