

# OAK DSP Core를 이용한 MPEG 계층 II 오디오 복호화기 구현

김수현, 김진호, 이창원, 김현중, 차형태  
승실대학교 전자공학과

## Implementation of MPEG Layer II Audio Decoder on OAK DSP Core

Soo-hyun Kim, Jin-ho Kim, Chang-won Lee, Hun-joong Kim, Hyung-tai Cha  
Dept. of Electronic Engineering, Soongsil Univ.  
E-mail : shkim@mmslab.soongsil.ac.kr

### 요 약

본 논문에서는 MPEG-1 계층 II와 MPEG-2 계층 II LSF 오디오 복호기를 OAK DSP Core를 이용하여 실시간 응용이 가능하도록 구현하였다. Ungrouping시 이용되는 테이블을 효율적으로 사용하였으며 합성필터부의 RAM과 ROM의 크기 그리고 각 부분의 연산에 필요한 연산량을 최적화하기 위하여 알고리즘을 효율적으로 적용하였고 불필요한 연산 부분을 제거하거나 최적화 하였다.

[1]은 1992년에 MPEG 회의에서 압축과 신장 그리고 전송을 위한 표준 방식이 결정되었다. 5.1채널과 최대 256 kbps이하의 오디오 전송률을 가진 MPEG-2 규격 [2]은 낮은 Sampling 주파수를 사용하며 1994년에 MPEG 회의에서 결정되었다[3][4].

본 논문에서 사용되는 OAK DSP Core는 16-bit 고정 소수점 연산, 16-bit 데이터와 프로그램 버스와 36-bit의 barrel shifter를 지니며, 통신과 전자 응용 분야에 좋은 성능을 지니도록 설계되어 있으며, 40 MIPS의 계산 능력을 갖고 있다.

### I. 서 론

현대 기술의 급속한 발전은 대용량의 디지털 정보의 처리와 전송을 가능하게 하였고 이에 따라 급속하게 이전의 아날로그 기술을 대체해 나가고 있다. MPEG (Moving Picture Export Group) 오디오, 비디오 규격은 디지털 방송에 적합하도록 만들어 졌으며 DBS, HDTV, DAB, Multimedia등에 적용이 되고 있다.

최대 1.5 Mbps이하의 비디오 전송률과 2 채널, 최대 384 Kbps이하의 오디오 전송률을 가진 MPEG-1 규격

### II. MPEG 오디오 복호기의 개요

MPEG 오디오 부호기에서 만들어진 비트열은 그림 1에서와 같이 크게 4가지 부분으로 나눌 수 있으며 다음과 같다. 첫째, Header부분은 동기신호, 비트열의 압축률, 압축 방법, Sampling 주파수, Stereo 또는 Mono 방송의 여부 등에 대한 기본적인 정보를 가지고 있으며 Error에 가장 민감한 부분이다. 둘째, CRC(Cyclic Redundancy Code)는 비트열의 주요 부분에 Error의 발생 여부를 검사하는 부분이며 동기 신호를 제외한

Header부분에서부터 Scfsi(Scalefactor selection information)까지 주어진 16bit 방정식에 의하여 계산된다. 셋째로 Audio Data 부분은 실제 부호화된 오디오 신호에 대한 정보를 가지고 있으며 Bit 할당량(Bit Alloc.), 크기정보 선택(scfsi), 크기정보(scfs) 그리고 Sample들로 이루어져 있다. 이 정보들을 바탕으로 역양자화와 합성 필터를 거쳐 PCM 신호로 복호화된다. 마지막으로 Ancillary부는 부가 정보를 전송하고자 할 때 사용되는 부분이다.

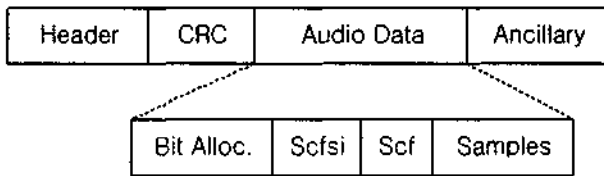


그림 1. MPEG 계층 II 비트열의 구조

이 비트열이 복호기에 입력되고 동기신호가 검출되면 Parser부에서 각 부분별로 정보들을 분류하여 저장한다. 압축된 Sample들까지 복호화가 끝나면 역양자화기에서 역양자화되고 크기정보를 이용하여 원래의 크기로 복원된다. 이 신호들은 합성필터부에서 합성되고 PCM 신호로 재생되어 출력하게 된다.

### III. OAK DSP Core를 이용한 MPEG 오디오 구현

MPEG 오디오 비트열에는 각 Frame의 시작을 표시하는 동기 신호가 들어있으며 12개의 '1'로 구성되어 있다. 동기 신호는 Sampling 주파수가 고정되었을 때 각 Frame간의 간격이 일정하다는 사실을 이용하여 검출될 수 있다. 이때 계층 II에 사용되는 한 Frame의 길이는 식(1)과 같이 계산할 수 있다.

$$\text{Bytes} = 144 \times \text{Bitrate} / \text{Sampling\_Frequency} \quad (1)$$

Parser부에서는 동기 검출된 비트열로부터 Header 정보, Bit allocation, Scale factor 정보와 Sample Data 들을 분석하고 저장한다. Header 정보는 비트열의 특성에 관한 자세한 기록, 즉, Sampling 주파수, Bitrate,

Stereo mode등의 정보를 가지고 있다. Header 정보에 들어 있는 ID 비트에 의해 MPEG-1('1') 비트열과 MPEG-2('0') 비트열이 구분된다. MPEG-2 계층 II LSF(Low Sampling Frequency) mode는 MPEG-1 계층 II의 알고리즘에 MPEG-2 계층 II의 Bit allocation 정보만을 추가함으로써 다른 프로그램 수정 없이 구현할 수 있다.

scalefactor는 각 Subband 당 3개의 scalefactor가 존재한다. 인접한 scalefactor가 비슷한 값을 가지는 경우에 대해 비트를 줄이기 위한 방법으로 표 1과 같이 2비트의 scfsi을 사용한다. 따라서 scfsi를 이용하여 scalefactor들을 복호화할 수 있다.

scfsi	scalefactor
00	scalefactor 0
	scalefactor 1
	scalefactor 2
01	scalefactor 0,1 동일
	scalefactor 2
10	scalefactor 0,1,2 동일
11	scalefactor 0
	scalefactor 1,2 동일

표 1. scfsi 정보에 대한 scalefactor

Sample Data는 Bit의 양에 따라 Grouping을 해서 보내는 경우가 있는데 이것은 다음과 같이 Ungrouping할 수 있다. Ungrouping 정보를 테이블화하여 사용하였으며 3비트로 표현되는 3개의 테이블들을 한 개의 테이블로 사용하였다.

```

for i=0,1,2
    sample[i] = c%nlevels;
    c = c/nlevels;
end

```

여기에서 c는 Grouping된 sample이며, nlevels는 Grouping된 3개 Sample의 Bit수이다. Syncword가 연속된 '1'로 구성되어 있으므로 이와 혼동을 피하기 위해 Sample Data의 MSB가 반전되어 있으므로 RND 명

령을 사용하여 1 cycle에 보상하도록 구현하였다.

역양자화부에서는 복호된 Sample Data에 역양자화 계수를 더하고 곱한 후에 Scalefactor를 곱하여 구해진다.

$$\begin{aligned} S' &= C \times (\text{sample} + D); \\ S &= S' \times \text{Scalefactor}; \end{aligned}$$

C와 D는 역양자화 계수이며 Scalefactor는 Parser부에서 해석된 크기 조정 계수이다.

Subband 합성은 다음의 식 (2)에 의해 처리된다.

$$Y[i] = \cos\left[(16+i)(2k+1)\frac{\pi}{64}\right] \times S_k \quad (2)$$

$i = 0, 1, 2, \dots, 63$   
 $k = 0, 1, 2, \dots, 31$

식(2)의 방정식을 이용하여 32개의 입력  $S_k$ 로부터 64개의  $Y[i]$ 를 구하여 사용해야 하므로 Stereo mode를 기준으로 한 Frame당 2304개의 Sample을 처리해야 하므로 오디오 복호화기 실시간 구현에서 합성필터 블록이 가장 많은 연산량을 차지하였다. 이러한 연산량을 줄이기 위해 cosine 함수의 대칭성을 이용하여 32개의  $Y[i]$ 만으로 64개의  $Y[i]$ 를 구하도록 하였다. 즉,

$$\begin{aligned} Y[i] &= -Y[32-i] & i = 0, 1, 2, \dots, 15 \\ Y[i+32] &= Y[64-i] & i = 1, 2, 3, \dots, 15 \\ Y[16] &= 0 \\ Y[48] &= -\sum S_k & k = 0, 1, 2, \dots, 31 \end{aligned} \quad (3)$$

이렇게 구하여진  $Y[i]$ 는 64개씩 Shifting되는  $V_i$ 로 대입된다.

$$\begin{aligned} V[i] &= V[i-64] & i = 1023, \dots, 65, 64 \\ V[i] &= Y[i] & i = 63, \dots, 2, 1, 0 \end{aligned} \quad (4)$$

$V_i$ 는 그림 2에서와 같이 1024개의 Word로 이루어져 있으나 Cosine 함수의 대칭성을 이용하여 512개의 Word만 저장하고 이로부터 1024개의 계수들을 구할 수 있다.  $U_i$ 는  $V_i$ 를 8개의 블록으로 만들어 1블록의 128개에서 앞 뒤 32개씩을 취하여 만들어진다. 이  $U_i$ 에

Window 계수  $D_i$ 를 곱하여 512개의  $W_i$ 를 만들고,  $W_i$ 를 32개의 계수를 가진 16개의 블록으로 만들어 각 블록을 더하여 32개의  $S_i$ 를 만든다. 또 32개의 PCM값인  $S_i$ 를 구할 때  $U_i$ 와  $W_i$ 와 같은 RAM을 거치지 않고 직접 계산함으로써 불필요한 RAM 부분을 줄일 수 있다.  $V_i$ 부터  $S_i$ 까지는 한 경로를 사용하여 계산함으로써 설계의 효율성을 기할 수 있고, 불필요한 출력부분의 시간 지연을 최소화 할 수 있다.

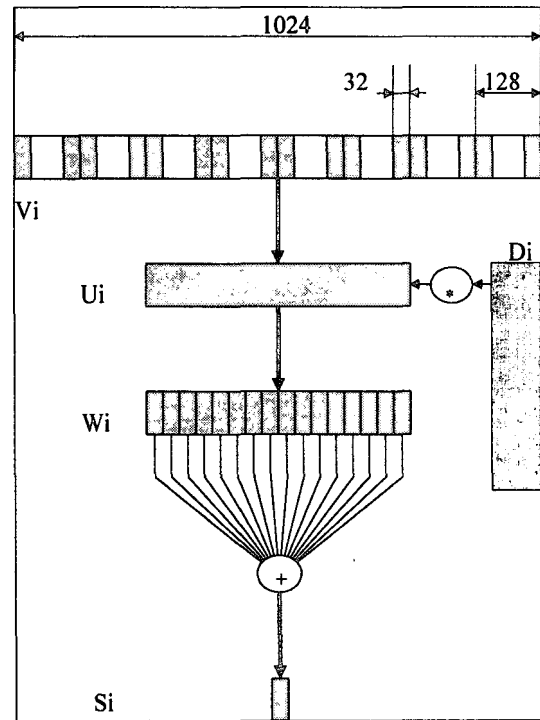


그림 2. MPEG Subband 합성구조

$$\begin{aligned} S_0 &= \sum_{k=0}^7 [V_{64k} D_{64k} - V_{64k+32} D_{64k+32}] \\ S_i &= \sum_{k=0}^7 [V_{64k+i} D_{64k+i} + V_{64k+i+47} D_{64k+i+32}] & i = 1, 2, \dots, 15 \\ S_{16} &= \sum_{k=0}^7 [V_{64k+63} D_{64k+48}] \\ S_i &= \sum_{k=0}^7 [V_{64k+79-i} D_{64k+32+i} - V_{64k+32-i} D_{64k+i}] & i = 17, 18, \dots, 31 \end{aligned} \quad (5)$$

이때  $V_i$ 와  $S_i$ 의 관계는 주어진 식 (5)로부터 직접 구

할 수 있다. MPEG 계층 II는 Subband 당 32개의 Sample 중 최대 30만의 샘플을 사용하여 처리되므로 32개의 Sample 모두를 처리할 필요가 없으며 연산량 및 ROM의 양을 줄일 수 있다.

### III. 성능 평가 및 실험 결과

본 논문에서는 MPEG 계층 II 복호기에 대한 검증으로서 어셈블러 프로그램의 성능을 OAK DSP core의 시뮬레이션을 통해 측정된 결과 약 13 MIPS의 연산량을 필요로 하였다. Word 단위를 16비트로 했을 때 이 알고리즘이 갖는 프로그램 ROM의 크기는 약 2.5 K Word이며, 약 4 K Word의 데이터 ROM과 약 2 K Word의 데이터 RAM으로 구성되어 있다.

독일의 Institut fuer Rundfunktechnik(IRT)에서 제공하는 표준안 pattern 테스트 벡터 7개를 입력으로 실험한 결과 최대 4비트의 오차를 가짐을 확인하였으며 그림 3은 이 중 192 kbit/s의 bitrate, 44.1 kHz의 Sampling 주파수를 가진 테스트 벡터의 결과를 나타낸다. 그림 3의 상위는 복호화된 원래의 값을, 중간은 시뮬레이션 결과를, 하위는 오차를 나타낸다.

### IV. 결론

본 논문에서는 OAK DSP Core를 이용하여 MPEG 계층 II 오디오 복호기를 구현하였다.

Ungrouping시 사용되는 3개의 테이블들을 1개의 테이블로 효율적으로 사용하였으며, cosine 함수의 대칭성과 하나의 경로 사용으로 합성필터부의 알고리즘을 효율적으로 적용하고, 불필요한 연산 부분을 제거함으로써 RAM과 ROM의 크기 그리고 각 부분의 연산에 필요한 연산량을 최적화하였다.

실험 결과에서와 같이 사람이 거의 인지하지 못하는 4비트 정도의 오차를 가짐으로써 MPEG 계층 II는 16 비트 system으로의 구현이 적절함을 알 수 있었다.

### 참고문헌

[1] ISO/IEC 11172-3, "Coding of Moving Pictures and Associated Audio for Digital Storage Media

at up to 1.5 Mbit/s Audio Part," *International Standard*, Nov., 1992.

[2] ISO/IEC 13818-3, "Generic Coding of Moving Pictures and Associated Audio-Video Part," *International Standard*, Nov., 1994.

[3] Mark, K. and Karlheinz, B., *Applications of Digital Signal Processing to Audio and Acoustics*, Kluwer Academic Pub., 1998.

[4] 차형태, "MPEG 오디오 합성 필터의 구현을 위한 연구," 제 13회 음성통신 및 신호처리 워크샵 논문집, pp. 79-84, 8월 16일, 1996.

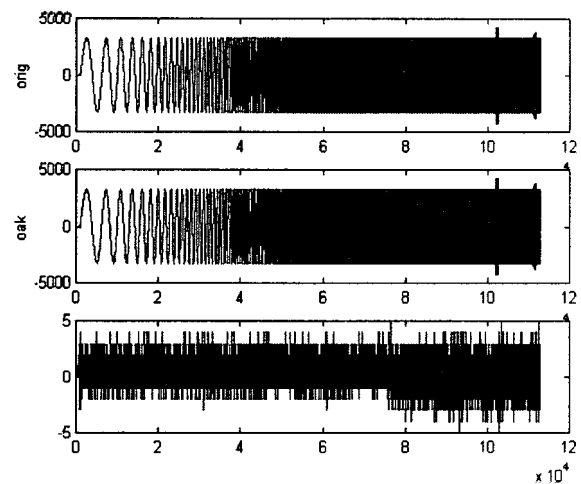


그림 3. 실험 결과 비교