

# 휴대용 기기를 위한 고음질 MPEG-1/2 계층-III 복호화기 실시간 구현

황태훈\*, 오현오\*, 이규하\*, 이근섭\*, 박영철\*\*, 윤대희\*

연세대학교 전기·컴퓨터 공학과\*, 연세대학교 신호처리연구센터\*\*

## A Real-Time Implementation of a High-Quality MPEG-1/2 Layer-III Decoder for Portable Devices

Tae-Hoon Hwang\*, Hyen-O Oh\*, Kyu-Ha Lee\*, Keun-Sup Lee\*, Young-Cheol Park\*\*, Dae-Hee Youn\*

Dept. of Electrical & Computer Eng., Yonsei University

Center for Signal Processing Research, Yonsei University

E-mail: hth@cyclon.yonsei.ac.kr

### 요 약

본 논문에서는 최근 휴대용 오디오 기기 등에서 활발하게 사용되고 있는 MP3 (MPEG-1,2 계층-III) 오디오 복호화 알고리즘을 실시간 구현하였다. 휴대용 기기에 적합한 저전력 설계를 위하여 16비트 고정 소수점 범용 DSP인 모토로라 DSP56654를 이용하였고, 연산량을 줄이기 위한 작업을 수행하였다. 또한 음질 열화를 최소화 하고 CD 수준의 고음질을 얻기 위해서 각 복호화 과정에 대한 최적의 고정소수점 연산을 연구하였다.

구현된 복호화기는 약 40MIPS 정도의 연산량으로 90dB이상의 SNR을 갖는 최종 PCM 샘플을 생성한다.

### 1. 서 론

최근 인터넷이 보편화되면서 일반인들에게도 널리 알려진 MP3 포맷은 MPEG-1/2 LSF(Low Sampling Freq.)의 계층 III을 말하는 오디오 압축 방식으로 일반적으로 2채널을 지원하고 있다.[1][2]

특히 휴대형 MP3 재생기가 개발되면서 MP3 포맷은

새로운 디지털 오디오 매체로서 주목을 받게 되었고, 이에 따라 MP3 복호화 칩은 많은 시장 수요를 창출할 것으로 기대되고 있다.

본 논문에서는 저전력 범용 DSP인 Motorola DSP56654를 이용하여 MP3 복호화 알고리즘을 실시간 구현한다. DSP56654는 Motorola DSP566xx 계열의 16비트 고정소수점 연산 DSP와 32-bit Micro-controller인 MCORE가 결합되어 있는 복합 프로세서이다. 이것은 PDA나 핸드폰과 같은 휴대용 응용에 적합하도록 저전력으로 개발된 칩이다[3]. 16비트 고정 소수점 연산 DSP는 MP3와 같은 고음질 오디오 신호를 처리하기엔 내부 연산 오차가 커서 부적합하지만, 저전력 설계가 가능하여 휴대형 기기에 장착하기에 유리하다.

본 논문에서는 16비트 고정 소수점 연산 DSP를 사용하면서도 음질 열화가 생기지 않는 CD 수준의 고음질 MP3 복호화기의 구현을 위해 서브루틴별로 다른 해상도를 적용한 음질 최적화 방법을 제안하며, 이에 따른 연산량 증가를 최소화하기 위해 몇 가지 고속 알고리즘을 적용하여 구현한다.

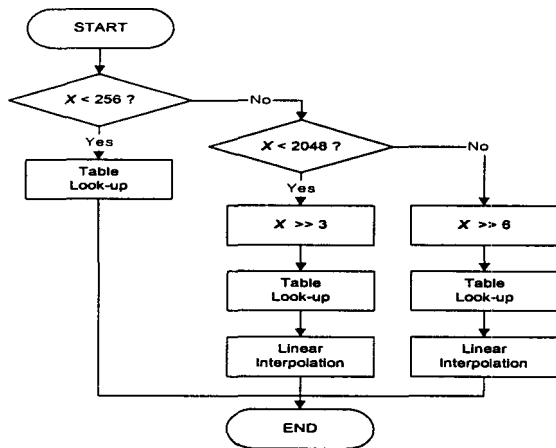


그림 1. 테이블 찾기 알고리즘

## 2. 저전력 설계를 위한 연산량 감소

저전력의 실시간 MP3 복호화기를 구현하기 위해서는 복호화 과정 중 대부분의 산술 연산을 차지하고 있는 역양자화 과정과 합성필터 뱅크 과정에 대한 연산량을 줄이기 위한 작업이 수행되어야 한다.[4]

### 2.1 역 양자화 과정

역 양자화 과정은 식 (1)과 같이 절대값을 취하는 과정과 정수 값에 4/3제곱을 하는 과정으로 나누어진다. 식 (1)에서  $X$ 는 양자화된 샘플값이고,  $Y$ 는 역 양자화된 값인 스펙트럼의 계수이다. 또한  $Sign(\ )$ 는 부호를 구하는 함수이다.

$$Y = Sign(X) \times |X|^{\frac{4}{3}} \dots\dots\dots(1)$$

$$\begin{aligned} |X|^{\frac{4}{3}} &= |X| \times |X|^{\frac{1}{3}} \\ &= |X| \times |(X/8) \times 8|^{\frac{1}{3}} \\ &= |X| \times |X/8|^{\frac{1}{3}} \times 8^{\frac{1}{3}} \\ &= |X| \times |X'|^{\frac{1}{3}} \times 2 \dots\dots\dots(2) \\ &= |X| \times |X'/8|^{\frac{1}{3}} \times 2 \\ &= |X| \times (|X'|^{\frac{1}{3}} \times 2) \times 2 \end{aligned}$$

식 (1)의 4/3 제곱 연산은 구현하는데 많은 연산량을 필요로 하기 때문에 미리 계산결과를 테이블로 만들어 놓고, 입력 값에 따라 결과 값을 찾는 table-look-up 과정을 이용하여 구할 수 있다.

그러나 테이블의 크기는 입력값의 범위에 따라 달라

지는데 입력값이 0~8191 사이의 정수이므로 테이블의 크기는 8192개가 필요하다. 그러나 역양자화 연산에서 8K word 이상의 테이블 메모리를 사용하기에는 구현상에 무리가 따르기 때문에 테이블 크기를 줄일 수 있는 방법을 생각해야 한다.

식 (2)와 같은 방법을 이용하여 모의 실험을 수행한 후 테이블의 개수를 256개로 만들었다. 그림 1은 역양자화 과정의 입력값에 따른 4/3 제곱값을 테이블을 이용하여 찾는 알고리즘이다. 입력값이  $x < 256$  인 경우에 대해서는 테이블에서 직접 찾아 사용하고,  $256 \leq x < 2048$  인 경우는 입력값을 3비트 쉬프트하고,  $2048 \leq x < 8192$  인 경우는 입력값을 6비트 쉬프트하고 테이블에서 찾은 후 선형보간법을 수행하게 된다.

선형 보간법은 쉬프트 연산에서 잃게 되는 나머지 부분에 대해 수행되므로 table-look-up 방식에 따른 오차를 줄일 수 있다. 또한 역양자화 과정의 대부분의 입력값이 256보다 작은 특성을 갖기 때문에 오차의 값이 크지 않다.

### 2.2 고속 합성 필터 뱅크

MP3의 서브밴드 합성 필터 뱅크 과정은 기존의 MPEG 오디오 계층-II에서 사용하는 방법과 동일하므로 32개의 PCM 샘플을 얻기 위해 64\*32의 MAC 연산을 수행하는 대신 32-point DCT(Discret Cosine Transform)를 이용하는 고속 필터 뱅크 알고리즘으로 대체할 수 있다[5]. 이 때 DCT의 경우 많은 고속 알고리즘이 존재하며 구현하고자 하는 시스템에서 가장 효율적인 성능을 내는 DCT 방법을 수행하면 수행시간을 보다 개선할 수 있다. 본 구현에서는 프로그램의 길이를 최소화 하면서도 DSP의 MAC연산을 최대한 활용할 수 있는 일반적인 32-point DCT 방법을 채택하였다.

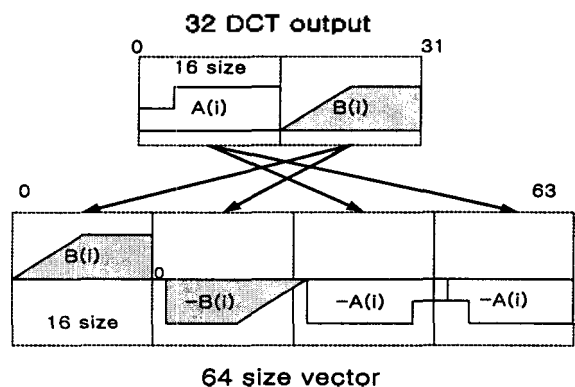


그림 2. 기존의 매트릭싱 결과와 고속알고리즘의 결과의 유사성

고속 알고리즘의 매트릭싱 과정에 의해 생성된 32개의 DCT출력은, 그림 2와 같은 과정을 통하여 64개의 벡터로 재구성된다. 64개의 샘플벡터들은 1024크기의 윈도우 연산버퍼(V버퍼)에 저장된다. 이와 같은 방식으로 64개 샘플 벡터를 구성하면 실제 정보량보다 2배 많은 데이터를 저장해야한다. 64크기의 샘플 벡터 구성 과정을 거치지 않고 32포인트 DCT 출력 데이터로 윈도우 연산버퍼를 구성하면, 윈도우 연산버퍼의 전체 크기를 512개로 줄일 수 있다.

그러나 이러한 방법은 윈도우 연산버퍼의 어드레싱이 복잡해지는 단점이 있다. 본 논문에서는 어드레싱의 문제점을 DSP에서 제공하는 병렬수행이 가능한 명령어와 모듈로-어드레싱 기법을 효율적으로 사용하여 해결하였다.

### 3. 음질향상을 위한 최적화 과정

고정 소수점 연산은 부동 소수점 연산에 비해 산술 연산이 비교적 간단하여 전력 소모가 적고 속도도 빠르다. 따라서 저전력 실시간 복호화기를 위해서는 고정 소수점 연산만으로 모든 과정을 구현하는 것이 가장 유리하다. 하지만 표현할 수 있는 수의 범위와 해상도가 제한되어 있다. 이런 이유로 고정 소수점 연산의 결과로 나오는 오차는 복호화 과정 중 계속 누적되어 복호화된 오디오 신호를 왜곡시키는 문제를 일으키게 된다. 복호화된 PCM 샘플이 부동 소수점 연산과 비교하여 만족할만한 수준의 오차범위(최하위 1비트 이하) 내에 들어오도록 하기 위해서는 서브루틴들을 선별하여 32비트 이중 해상도 연산을 적용해야 한다.

이를 위해 본 논문에서는 MP3 복호화 과정 중 산술 연산이 요구되는 서브루틴들에 대해 5가지 방법으로 나누어 고정 소수점 연산 모의실험을 수행하였다. 5가지 모의실험에 사용된 각 서브루틴별 해상도는 그림 3과 같다. 예를 들어 방법 IV는 역양자화 과정부터 합성필터뱅크 과정의 매트릭싱 과정까지를 32 비트로 처리하고 그 이후의 과정에 대해 16비트로 연산하는 것을 의미한다.

구분	스케일 팩터 / 허프만	역 양자화	IMDCT	합성필터뱅크
방법 I	16bits			
방법 II	16bits	32bits	16bits	
방법 III	16bits	32bits		16bits
방법 IV	16bits	32bits		16bits
방법 V	16bits	32bits		

그림 3. 음질 향상을 위한 최적화 과정

각 연산 방법에 의해 복호화된 PCM 신호의 음질을 객관적으로 평가하기 위해 부동 소수점 연산 복호화기로 복호화된 샘플과의 차이로부터 SNR을 측정하였다.

비교를 위한 부동 소수점 복호화기는 Fraunhofer에서 제공하는 L3DEC.exe를 사용하였고, 평가에 사용된 비트열은 SQAM(Sound Quality Assessment Material) CD 내의 오디오 샘플들을 Fraunhofer의 L3ENC.exe를 통해 부호화 한 것으로 사용하였다[6][7].

본 논문에서 오차를 평가하기 위한 SNR은 식 (3)과 같이 정의한다.

$$SNR = \{20 \times \log_{10}(2^{16} - 1)\} - \{20 \times \log_{10} 2 \times RMS_{pcm}\}$$

$$RMS_{pcm} = \sqrt{\sum_n [S_{test}(n) - S_{ref}(n)]^2} \dots \dots \dots (3)$$

식 (3)에서  $S_{test}$ 은 고정소수점으로 복호화된 PCM 샘플이고,  $S_{ref}$ 은 부동 소수점으로 복호화된 PCM 샘플 신호이다. 식 (3)에 의하면 고정소수점 연산을 통해 복호화된 PCM 신호가 부동 소수점 연산 복호화기의 결과와 비교하여 평균 최하위 1비트 이하의 오차가 있을 때 90dB 이상의 SNR을 얻을 수 있다. 각 방법들에 따른 최종 SNR 결과는 제 4장에서 정리하였다.

### 4. 실험 및 결과

음질 향상을 위한 최적화 과정 방법 5가지 중 방법 I, 방법 IV, 방법 V에 대한 SNR 결과를 표 1에 정리하였다. 표 1에서 볼 수 있듯이 방법 V만이 실험에 사용된 모든 테스트 비트열에 대해 90dB 이상의 SNR 결과를 갖는다. 따라서 본 논문에서는 방법 V를 사용하여 실시간 구현한다.

표 2는 구현된 복호화기의 각 과정별 MIPS 단위의 수행시간을 나타낸다. ADS보드상에서 실제 수행시간 측정은 DAC 인터럽트를 이용한 방법을 사용했다[8]. 방법 V의 연산량은 대략 40 MIPS 정도인데, 방법 I일 때보다 전체적으로 약 58.1% 정도의 증가량을 보인다. 특히 합성필터뱅크 과정이 전체 연산량의 약 40%를 차함을 볼 수 있다.

표 3은 어셈블리 프로그램에 사용된 메모리의 크기로써 방법 V일 때 데이터 메모리, 프로그램 메모리, 테이블 메모리의 크기가 각각 4K, 9.7K, 5.7K word 정도이다. 방법 V의 메모리 양은 이중 해상도를 사용하는 루틴들에서 방법 I에 비해 데이터 메모리 사용이 두배 증가하여 전체적으로 15.5% 증가하였다.

표 1. 5가지 방법에 대한 SNR 결과 (단위: dB)

SQAM	설명	방법 I	방법 IV	방법 V
frer07_1	Electronictune	74.8004	89.5919	91.9694
horn23_2	Horn	58.5831	89.8692	91.9196
gspl35_1	Glockenspiel	60.1806	90.0852	92.0550
gspl35_2	Glockenspiel	53.1447	90.0277	91.9946
harp40_1	Harpsichord	45.7475	90.0938	92.0636
bass47_1	Bass	37.0026	89.9642	91.9694

표 2. 서브 루틴별 연산량 측정 결과 (단위: MIPS)

구분	방법 I	방법 V	증가율(%)
스케일팩터/허프만	6.69	6.69	0
역양자화	5.23	6.84	30.8
IMDCT	5.38	8.80	63.6
합성필터뱅크	7.02	15.78	124.8
기타	1.01	1.79	77.2
합계	25.23	39.90	58.1

표 3. 사용된 메모리의 양 (단위: word)

메모리	방법 I	방법 V	증가율(%)
프로그램	3.5K	4K	14.3
데이터	7.8K	9.7K	24.4
테이블	5.5K	5.7K	3.6
합계	16.8K	19.4K	15.5

실시간 복호화기의 성능평가를 위해, SQAM CD에 있는 곡들을 Fraunhofer에서 제공하는 부호화기로 부호화 하여 테스트 비트열로 이용하였다. 구현된 실시간 MP3 복호화기로 복호화 한 PCM 샘플들과 Fraunhofer에서 제공하는 복호화기에 의한 PCM 값과의 비교 결과인 최종 SNR은 방법 V에 대해서 모든 테스트 비트열에 90dB 이상을 보이는 것을 확인하였다.

## 5. 결론

본 연구에서는 16비트 고정 소수점 범용 DSP인 DSP56654를 이용하여 MP3 복호화기를 구현하였다. 저전력을 위해 많은 산술 연산을 수행하는 역양자화 과정과 합성필터뱅크 과정의 연산량을 감소시켰다. 고정 소수점 연산으로 인한 음질 저하를 줄이고, CD수준의 고음질을 얻기 위해 5가지 방법에 의한 고정 소수점 시뮬레이션을 통하여 각 서브루틴별 해상도를 결정하였다.

이러한 과정을 통해 약 40MIPS의 연산량으로 90dB SNR의 고음질을 제공하는 실시간 MP3 복호화기를 구

현 할 수 있었으며 구현된 MP3 복호화기는 저전력 DSP를 사용하므로 가정용 오디오 기기, 특히 휴대형 오디오 기기에 사용하기 적합하다.

## 6. 참고 문헌

- [1] ISO/IEC JTC1/SC29/WG11 No.71 "Coding of Moving Pictures and Associated Audio for Digital Storage Media at up to about 1.5 Mbit/s - CD 11172-3(Part3. MPEG-Audio)"
- [2] ISO/IEC JTC1/SC29/WG11 No.1519 "Generic Coding of Moving Pictures and Audio - CD 13818-3(Part3. MPEG-Audio)", 2nd Edition, Feb., 1994.
- [3] Motorola, DSP56654 Baseband Digital Signal Processor User's Manual, 1999
- [4] 김진원, 정남훈, 김준석, 이근섭, 이충용, "MPEG-1 계층 III 오디오 복호화기의 VLSI 설계," 제 12회 신호처리합동학술대회, pp. 847-850, 1999
- [5] K. Konstantinides, "Fast Subband Filtering in MPEG AUDIO Coding.", IEEE Signal Processing Letters, vol. 1, no. 2, pp. 26-28, Feb., 1994.
- [6] <http://www.iis.fhg.de/amm>
- [7] <http://www.tnt.uni-hannover.de/project/mpeg/audio/sqam>
- [8] 김성운, 오현오, 윤대회, "범용 DSP를 이용한 MPEG-2 오디오 부호화기" ,Telecommunication Review 8권, 3호,1998년