

MPEG-2 오디오 부호화기 설계

정남훈, 박성욱, 김준석, 유선국†, 윤대희, 차일환

연세대학교 전자공학과, †연세대학교 의용공학과

Design of An MPEG-2 Audio Encoder Chip

N.H. Jeong, S.W. Park, J.S. Kim, S.K. Yoo†, D.H. Youn, I.W. Cha

Dept. of Electronic Eng. Yonsei University, †Dept. of Medical Eng. Yonsei University

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

요약

본 논문에서는 VLSI 기술에 바탕을 둔 top-down 접근 방식에 의하여 MPEG-2 오디오 부호화 알고리즘을 구현하였다. MPEG-2 오디오 부호화기의 알고리즘은 많은 연산량을 갖고 이질적인 특성을 갖는 알고리즘들이 복합적으로 존재한다. 그러므로, 부호화기를 효과적으로 구현하기 위해서는 알고리즘 수준에서 구조적 수준에 이르기까지 많은 고찰이 이루어져야 한다. 본 논문에서는 우선 전체 부호화 알고리즘을 분석하여 이들을 다시 작업이라고 정의된 작은 부-알고리즘으로 나누었다. 다음으로, 분할된 작업들은 시간과 공간을 최대한 활용할 수 있도록 적절한 작업 순서를 부여하고, 좀더 큰 모듈들로 모으는 클러스터링을 수행하였다. 마지막으로 이러한 분석 결과를 바탕으로, 실시간으로 동작하는 5.1 채널 MPEG-2 오디오 부호화기를 설계하였다. 설계된 시스템은 두 개의 하드웨어 블록과 한 개의 ASIC형 DSP 프로세서를 갖는 이질적인 다중 프로세서의 형태를 갖는다. 설계된 오디오 부호화기는 0.6 μ m 표준 셀 기술을 이용하여 단일 칩으로 제작되었으며, PC에 탑재 가능한 시험 기판을 제작하여 동작을 검증하였다.

1. 서론

MPEG 오디오 부호화 알고리즘은 화상 회의, DCC, DVCR, CD-I, DAB, HD-TV 오디오 등의 여러 응용 분야를 갖고 있다. 이러한 여러 응용 분야들은 크게 단방향 서비스와 양방향 서비스로 나누어질 수 있다. 화상 회의와 같은 양방향 서비스의 경우, 모든 사용자들은 부호화기와 복호화기를 동시에 보유하고 있어야 한다. 그러나 부호화/복호화기의 필요성에도 불구하고, 부호화 알고리즘의 복잡성 및 많은 계산량으로 인하여 부호화기의 구현에 대한 연구는 복호화기에 비해 상대적으로 적게 이루어져 왔다.

MPEG 오디오 부호화 알고리즘은 많은 계산량을 갖는 변환 과정과 데이터의 특성에 의해 결정되는 비트 할당 및 양자화 과정으로 이루어져 있다. 이러한 이질적이고 계산량이

많은 알고리즘을 일반적인 프로세서로 구현하려면, 필연적으로 많은 수의 프로세서를 사용하는 큰 시스템을 설계하여야 한다[3]. 그러나, 일반적인 프로세서들은 특정한 응용분야에 항상 최선의 해결책이 되도록 설계되어 있지 않기 때문에 이러한 설계방식은 하드웨어의 이용도 측면에서 그리 바람직하지 못하다. ASIC 설계 기술을 이용함으로써 이러한 비효율적인 문제를 극복하고 비용 측면에서 훨씬 효율적인 오디오 부호화기를 설계할 수 있다.

2장에서는 고정 소수점 시뮬레이션을 통한 MPEG-2 오디오 부호화기의 분석에 대하여 기술한다. 이러한 분석을 바탕으로, 효과적인 시스템의 분할 및 시간 할당을 수행하였다. 3장에서는 실시간 제한 조건을 만족시키기 위한 과정에 대하여 설명한다. 최종적으로 설계된 MPEG-2 오디오 부호화기의 구성은 4장에 기술하였다. 마지막으로 5장에서 동작 검증 및 합성 결과로 끝을 내었다.

2. 알고리즘 분석

2.1 MPEG-2 오디오 부호화 알고리즘

그림 1은 MPEG-2 다채널 오디오 부호화 계층 II의 부호화 과정이다[1][2].

시스템의 입력으로 들어간 시간 영역 신호는 분석 필터(AF)와 매트릭싱(matrixing) 모듈을 통과하여 서브밴드 영역의 신호로 변환된다. 크기인자 부호화(scf coding) 모듈은 각 서브밴드의 크기인자를 찾아서 크기인자 선택정보와 함께 부호화한다. 결정된 크기인자는 다채널 부호화 과정중에 동적 전송채널 할당(dyn TC allocation)의 정보로 사용된다. 동적 비트 할당(bit-allocation) 과정에서는 심리 음향 모델(PAM)에 의해 구해진 SMR(Signal to Mask Ratio)을 이용하여 양자화기를 결정한다. 마지막으로, 서브밴드 샘플을 양자화한 후 최종 비트열을 생성함으로써 부호화 과정을 마치게 된다.

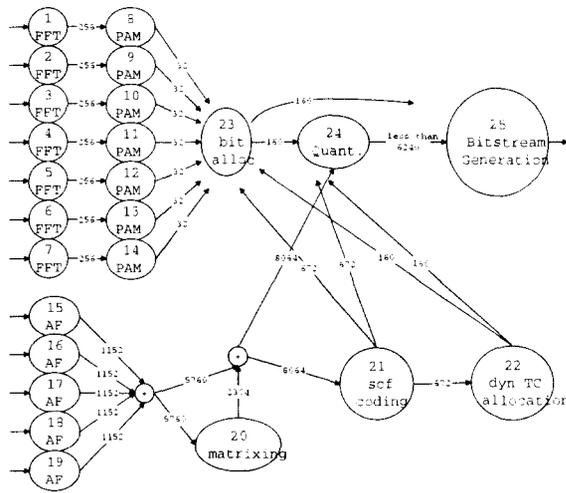


그림 1: MPEG-2 Audio Layer II Encoding Procedure

2.2 행위수준 분할 및 스케줄링

표 1은 고정소수점 시뮬레이션을 통해 분석한 각 세부 과정에서 소요 사이클수와 연산량의 분포를 나타낸다.

표 1: Number of operations for MPEG-2 multichannel audio layer II encoding algorithm

task	# of op. (unit: 1000 cyc)	ratio (%)
5ch AF	600	26.1
7ch FFT (PAM step1)	360	15.7
7ch PAM step2 ~ step8	840	36.5
matrixing	40	1.7
scf coding	65	2.8
dynamic channel allocation	5	0.2
bit-allocation	180	7.8
quantization	130	5.7
bitstream generation	80	3.5
total	2,300	100.0

실시간 제한 조건 하에서 효율적인 부호화기를 설계하기 위해서는 먼저 작업 수준의 분할이 이루어져야 한다[4][5]. 분할의 최종 목표는 주어진 시간과 공간의 제약 조건 하에서 최소의 면적과 최대의 성능을 얻는 것이다. 먼저, 전체 시스템을 그림 1과 표 1과 같이 분할하였다. 시스템의 크기를 줄이기 위하여, 작업간의 유사성을 부호화기의 분할을 위한 주요 척도로 사용하였다. 유사한 동작들을 갖는 작업 모듈들은

하나의 큰 모듈로 모아지고, 이렇게 만들어진 모듈은 동일한 datapath를 사용한다. 작업간의 전송비용은 시스템 분할의 두번째 척도가 된다. 전송비용을 줄이기 위하여, 높은 전송비용을 갖는 작업 모듈들은 하나의 모듈로 모아지게 된다. 표 2는 시스템 분할의 결과를 요약한 것이다.

작업간의 분할을 수행한 후, 각 작업의 특성을 고려하여 적절하게 작업의 흐름을 조정함으로써 각 모듈간의 전송 비용을 효과적으로 줄일 수 있다. 그림 2는 전송비용을 감소시킨 부호화 시스템의 작업 수행 흐름도이다.

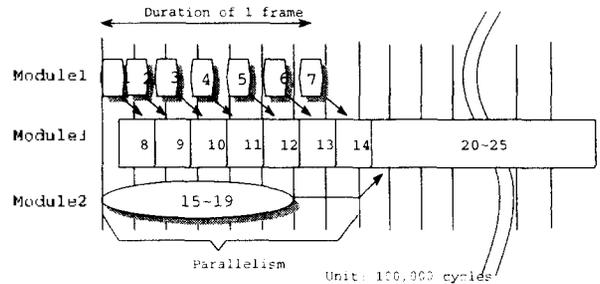


그림 2: Task Schedule of the Encoder

3. 실시간 동작을 위한 고려

제해진 시간 안에 부호화기가 동작하기 위해서는, 각 모듈들이 적절한 시간 안에 할당된 작업을 수행할 수 있도록 조정되어야 한다.

3.1 분석 필터뱅크와 FFT

표 1에 의하여, 모듈-2(분석 필터뱅크)는 할당된 작업을 수행하는 데 약 600,000 사이클을 소요한다. 표본화 주파수가 48KHz일 때 사용 가능한 총 사이클 수가 640,000 사이클임을 고려하면, 부호화기는 사용 가능한 대부분의 사이클을 분석 필터링 과정에 사용하게 된다. 이러한 이유때문에, 수학적 방법이나 구조적인 방법을 이용하여 분석 필터링 과정에 요구되는 사이클 수를 줄여야 한다. 수리적인 접근 방식에서 보면, Konstantinos[6]가 제안한 알고리즘을 이용함으로써 연산량을 줄일 수 있다. 구조적인 접근 방식에서 보

표 2: The result of clustering

# of iter.	clusters (with more than 2 elements)
1	(21,24)
2	(20,21,24)
3	(20,21,24,25)
4	(15,16,~20,21,24,25), (22,23)
5	(15,16,~20,21,22,23,24,25)
6	(1,2,~13,14), (15,16,~21,25)

면, 필터뱅크 연산의 정형적인 특성을 이용하여 FSM(Finite State Machine)에 의해 구동되는 하드웨어를 이용함으로써 효과적으로 구현할 수 있다.

모듈-1(FFT)을 설계하는데 있어서도, 분석 필터뱅크와 유사한 방식을 적용시킬 수 있다. 본 논문에서는 Cooley-Tukey의 FFT 알고리즘을 FSM과 한 개의 곱셈기, 두 개의 덧셈기를 이용하여 구현하였다.

3.2 심리 음향 모델(PAM)

표 1로부터, PAM (step 2~8) 과정의 수행 시간은 전체 사용 가능한 640,000 사이클보다 훨씬 많은 840,000 사이클임을 알 수 있다. 알고리즘 분석 결과, 개별/전체 마스크 임계값을 구하는 과정이 연산의 대부분을 차지하고 있다. 본 논문에서는 로그와 지수 연산을 수행할 수 있는 별도의 유닛을 설계하고, 마스크 곡선의 구간 선형성(piecewise linear)[1]을 이용하여 연산에 소요되는 시간을 효과적으로 줄였다.

3.3 기타 연산 유닛의 설계

크기 인자의 계산이나 비트 할당 과정, 그리고 비트 패킹 과정은 많은 비교 판단 및 반복 연산을 수행하는 과정으로 이루어져 있다. 이러한 작업들은 보통 한 번 수행하는데 3~6 사이클 정도를 사용하며, 매우 빈번하게 사용되기 때문에, 전체 연산량의 측면에서 많은 부분을 차지하게 된다. 본 논문에서는 전체 연산량을 줄이기 위해서 이들을 위한 별도의 연산 유닛(가속기)을 설계하였다.

4. 시스템의 구현

오디오 부호화기는 하드웨어로 설계한 FFT 모듈과 AF 모듈, 그리고 주문형 프로세서로 설계된 프로세서 모듈과 부가 논리 회로로 구성된다.

4.1 주문형 프로세서 구현

설계된 DSP 프로세서는 기본적으로 프로그램 시퀀서와 데이터 번지 발생기, 그리고 datapath의 세 가지 요소들로 구성되어 있고, 효과적인 동작을 위해서 프로그램 메모리와 데이터 메모리는 서로 나누어져 있다.

프로그램 시퀀서의 주된 기능은 명령어를 패치하고 해석하는 것이다. 프로그램 시퀀서는 파이프라인에 필요한 적절한 제어신호를 발생시키고, 부피와 함수 호출을 부가 사이클 없이 수행할 수 있도록 지원한다. 데이터 주소 발생기는 메모리에 있는 피연산자를 접근할 때 사용할 주소를 효과적으로 발생시킨다. 효과적인 주소 발생을 위하여 별도의 덧셈기 및 8개의 레지스터를 갖고 있고, DSP에 사용되는 기본적인 주소 발생을 지원한다.

데이터 경로 처리기는 레지스터 파일과 연산기들로 구성되어 있다. 연산기는 ALU와 MAC 유닛을 포함하여 배럴 쉬프터(barrel shifter), 로그와 지수 연산 유닛 기타 오디오 부호화를 위해 특성화된 연산 유닛을 갖고 있다.

4.2 MPEG-2 Audio Encoder System

부호화기를 구성하는 각 모듈들은 알고리즘상의 데이터

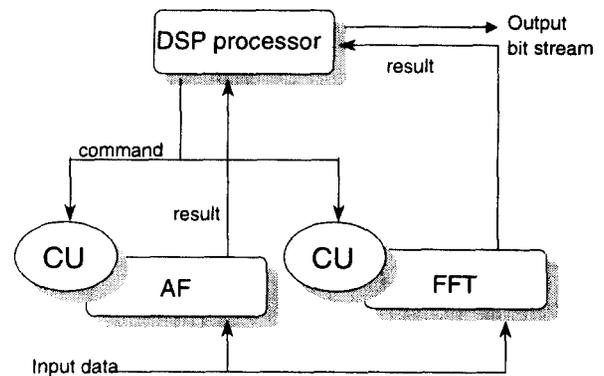


그림 3: Configuration of the Audio Encoder (The DSP processor (module-3) works as a bus-arbitrator and task scheduler. AF and FFT takes commands from the DSP processor and return the produced results.)

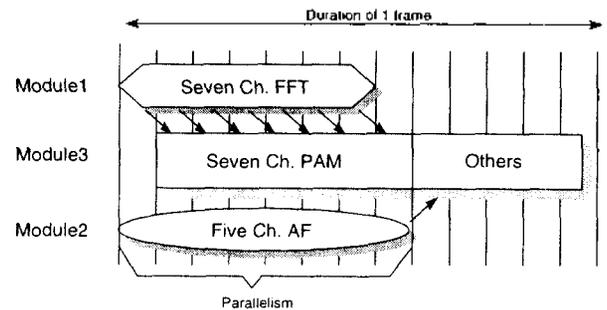


그림 4: Refined Schedule of the Audio Encoder (All modules are accelerated respectively. The pipelinings between module-1 and -3, and module-2 and -3 are refined to only have negligible overheads)

동시성을 극대화하도록 배치되어 있고, 그에 맞는 시간 할당을 하였다. 부호화기의 배치 및 시간 할당은 각각 그림 3과 그림 4에 나타내었다. 각 모듈간의 효과적인 정보 전송을 위하여, 오디오 부호화기는 두 개의 분리된 메모리 블록으로 구성된 공유 메모리 구조를 갖고 있다. 또한 그림 4에서 보는 바와 같이 FFT 과정과 FFT 과정을 제외한 심리 음향 모델링 과정의 연산 시간이 비슷하게 함으로써 파이프라인을 효과적으로 수행할 수 있도록 하였다. 시스템의 최종적인 시간 정보는 표 3에 나타내었다. 설계된 시스템은 파이프라인과 병렬처리 기법을 동시에 사용함으로써 약 568,000 사이클만에 모든 작업을 마칠 수 있고 좀 더 높은 비트율을 지원할 수 있는 여유도 갖게 되었다.

5. 결과

표 3: Timing analysis of finally designed audio encoder

task	module number	# of op.(1000cyc)	
		개별	동시성
1ch FFT	1	45	358 =315+43
1ch PAM	3	43	
7ch FFT	1	315	
5ch AF	2	350	
matrixing	3	30	210
scf coding	3	40	
bit-allocation	3	50	
quantization	3	60	
bitstream generation	3	30	
total			568

설계된 MPEG-2 오디오 부호화기는 0.6 μ CMOS M3 기술에 의하여 구현되었다. 부호화기의 동작은 시스템의 설계 단계별로 고 수준에서 저 수준에 이르기까지 모두 검증되었다.

행위 수준의 검증은 C 언어를 이용한 컴퓨터 시뮬레이션과 VHDL을 이용한 SYNOPSIS 상에서의 결과를 비교하여 수행하였다. 그리고 클럭에 기반하여 부호화 과정의 시간 정보를 추출하였다. 행위 수준의 검증이 끝난 후, VHDL을 이용하여 RT(Register-Transfer) 수준이나 더 낮은 수준으로 기술한 후 합성하였다. 합성된 각 모듈의 동작은 동작 수준의 결과와 비교하여 검증하였다. 합성 후의 시간 정보 또한 행동 수준의 시간 정보와 같은 결과를 나타내었다.

표 4: Maximum time delay paths

signal name	delay[ns]
DSP (from data register to memory input port)	27
DSP (from pointer register to internal memory address port)	5.84

표 1에 나타난 바와 같이, 최악 상황에서 시간 지연은 27 ns이 발생하였다. 27 MHz의 시스템 클럭을 사용할 경우 설계된 시스템은 모든 연산들을 37 ns 안에 수행할 수 있어야 한다. 그러므로, 최악 상황에서도 설계된 시스템은 사용 가능한 시간의 약 73%만을 사용할 수 있다.

제작된 부호화기는 최종적으로 시험 기판에 탑재 하여 실시간 검증을 마쳤다. 구현된 오디오 부호화기는 68,311 개이

트에 총 373x389의 크기를 갖는다. 그림 5는 제작된 칩의 내부를 보여준다.

MPEG-2 AUDIO ENCODER

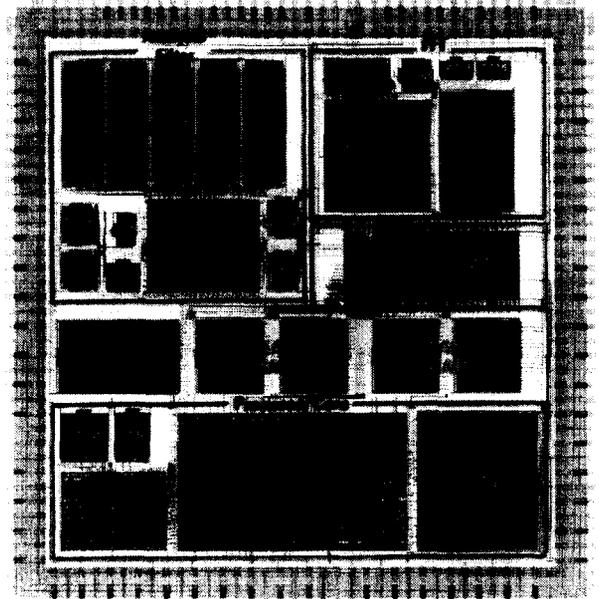


그림 5: The Micro-photograph of the Audio Encoder

참고 서적

- [1] ISO/IEC JTC1/SC29 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. AUDIO)"
- [2] ISO/IEC JTC1/SC29 No.703. "Generic Coding of Moving Pictures and Associated Audio - CD 13818-3 (Part3. MPEG-AUDIO)," Mar. 1994.
- [3] S. Y. Kim, et al., "An Real-Time Implementation of the MPEG-2 Audio Encoder," *IEEE Trans. Consumer Elect.*, vol. 43, no.3, pp. 593-597, Aug. 1997.
- [4] Frank Valud, "A Survey of Behavioral-Level Partitioning Systems," Tech. Rep. #91-71, Dept. of Inf. and CS, UC Irvine, Oct. 1991.
- [5] R. Camposano, and R. Brayton, "Partitioning Before Logic Synthesis," *Proc. Int. Conf. Computer-Aided Design*, 1987.
- [6] Konstantinos Konstantinides, "Fast Subband Filtering in MPEG Audio Coding," *IEEE Signal Processing Letters*, vol. 1, no. 2, pp. 26-28, Feb. 1994.