

멀티채널 오디오(MPEG-2 AAC) 인코더 및 디코더의 실시간 구현

Real-time Implementation of Encoder and Decoder for Multi-channel Audio(MPEG-2 AAC)

홍진우, 김진웅, 박재홍, 양재우

Jin-Woo Hong*, Jin-Woong*, Jae-Hong Park*, and Jae-Woo Yang**

*ETRI-무선방송기술연구소 방송기술연구부

**ETRI-교환전송기술연구소 통신단말연구부

Jwhong@etri.re.kr

요약

본 논문에서는 ISO/IEC MPEG-2 AAC 표준을 기반으로 한 멀티채널 오디오 인코더 및 디코더(MASIC) 시스템의 실시간 구현 기술에 대해서 기술한다. MPEG-2 AAC 기술은 멀티채널 오디오 부호화 방식의 국제 표준으로써, 지금까지 개발된 멀티채널 오디오 부호화 방식중 최신의 기술이며, 압축율과 오디오 품질이 가장 우수한 것으로 알려져 있다. MASIC 시스템은 인코딩 및 디코딩 기술의 실시간 처리를 위하여 범용 DSP 인 TMS320C6701 을 사용하였고, 멀티채널 오디오의 고속 입력과 출력을 위한 디지털 인터페이스를 가지고 있으며, 개인용 컴퓨터와의 인터페이스를 위한 PCI 기술이 적용되어 다양한 입출력 모드를 지원하는 특징을 갖는다.

1. 서론

방송, 화상전화, 영상회의, 대화형 멀티미디어 서비스(정보, 교육, 오락등)와 같은 시각통신은 기본적으로 오디오를 필요로 하고, 미래의 문자, 그래픽, 팩스, 정지영상, 전자메일 등도 음성이나 오디오를 하나의 수단으로 제공하게 될 것이다. 또한, 전송 미디어로서 뿐만 아니라 패키지 미디어로서 오디오가 갖는 장점도 매우 다양하다. 이러한 상황을 인지하여 보면 오디오 신호를 압축하기 위한 연구는 매우 중요하고, 필요한 기술이 된다.

따라서, MPEG(moving Picture Experts Group) 오디오 그룹에서는 1992 년에 스테레오 채널 오디오 부호화 방식으로 MPEG-1 오디오 국제 표준(표준번호 ISO/IEC 11172-3)을 제정하였으며, 1994 년 멀티채널 오디오 부호화 방식으로 MPEG-2 오디오 BC(Backward Compatibility)(표준번호 : ISO/IEC 13818-3)를 제정하였다. 또한, MPEG-4 오디오 국제표준(표준번호 ISO/IEC 14496-3, 버전 1)을 1999 년 초에 제정하였으며, 현재 MPEG-4 오디오 버전 2 와 MPEG-7 오디오 국제표준의 제정을 위하여 노력하고 있다.

그러나, MPEG-2 오디오 BC 는 MPEG-1 오디오와의 호환성을 유지하여야 한다는 제한 때문에 압축율과 오디오 품질면에서 다른 멀티채널 오디오 부호화 방식(예를 들면, Dolby Lab.의 AC-3 와 AT&T 의 MPAC 등)보다 열악한 결과를 나타나게 되었다. 따라서, MPEG 오디오 그룹에서는 MPEG-1 오디오와 호환성을 유지하지 않는(NBC : Non-Backward Compatibility) 멀티채널 오디오 부

호화 방식인 MPEG-2 AAC(Advanced Audio Coding) 국제표준을 1997 년에 제정하였다.

MPEG-2 AAC 부호화 방식은 지금까지 개발된 멀티채널 오디오 부호화 방식중 최선 기술(state of the art)로써, 압축률 및 품질면에서 가장 우수한 것으로 인정받고 있으며, 제한된 대역폭을 갖는 분야로부터 고품질을 요구하는 응용분야까지 적용할 수 있는 특징을 가지고 있다. 또한, 관련 기술들이 멀티미디어용 오디오 부호화 기술인 MPEG-4 오디오의 일부 기술로 활용되고 있으며, 48KHz 샘플링 주파수를 갖는 5 채널 오디오 신호에 대해 320kbps의 데이터 율(즉, 채널당 64kbps)에서 ITU-R 에서 규정한 구분할 수 없는(indistinguishable) 오디오 품질을 갖는다.

MPEG-2 AAC 부호화 알고리즘은 필수 기술(required tool)과 선택 기술(Optional tool)로 구분되는 핵심 기술(Core tool)의 조합으로 구성되어 있으며, 복잡한 디지털 신호 처리 및 알고리즘의 반복 처리에 의해 계산량이 매우 많아 실시간으로 구현하기 위해서는 다양한 고려사항이 요구되고 있는 상황이다.

본 논문에서는 이러한 요구사항을 고려하여 실시간으로 동작되는 MPEG-2 AAC 인코더 및 디코더 시스템(이하 MASIC : Multimedia Audio and Speech Integrated Codec 시스템이라 함)을 개발하였으며, 이와 관련된 하드웨어 및 소프트웨어 기술에 대해 설명한다.

2. MASIC 시스템의 기능 및 규격

MASIC 시스템은 부동 소수점을 갖는 범용 디지털 신호처리기인 TI사의 TMS320C6701 DSP 를 이용하여 구성하였으며, 개인용 컴퓨터(PC) 환경에서 오디오 저장 도구로 활용할 수 있도록 PCI 인터페이스를 적용하였다. MASIC 시스템의 기능 및 외부 시스템과의 인터페이스를 그림 1에 도시하였다.

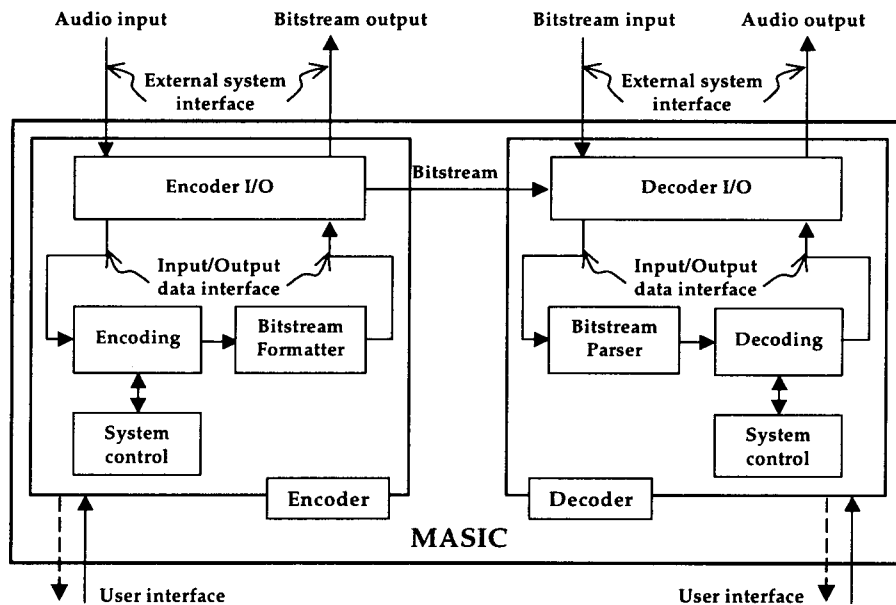


그림 1. MASIC 시스템의 기능 및 외부 인터페이스 구성도

MASIC 시스템은 크게 입력장치(Input Box, MASIC-I), 출력장치(Output Box, MASIC-O), 인코더 보드(MASIC-ENC), 그리고 디코더 보드(MASIC-DEC) 등 4 개의 모듈로 구성되어 있으며, 각 모듈의

기능 및 규격은 다음과 같다.

2.1 입력장치(MASIC-I)

- 멀티채널 오디오 신호의 입력
- 최대 8 채널의 아날로그 오디오 입력 인터페이스
- AES/EBU 형식을 갖는 최대 8 채널(스테레오 4 개 채널)의 디지털 오디오 입력 인터페이스
- 24 비트 양자화를 갖는 32KHz, 44.1KHz, 48KHz 의 샘플링 주파수 사용
- ADI(Alesis Digital Interface) 형식을 갖는 다중화된 멀티채널 오디오 데이터의 광출력
- 입력 오디오 신호의 감시 기능

2.2 출력장치(MASIC-O)

- 멀티채널 오디오 신호의 출력
- 최대 8 채널의 아날로그 오디오 출력 인터페이스
- S/PDIF 형식을 갖는 최대 8 채널(스테레오 4 개 채널)의 디지털 오디오 출력 인터페이스
- 24 비트까지의 양자화로 구성된 디지털 신호의 아날로그 변환(D/A 변환)
- ADI(Alesis Digital Interface) 형식을 갖는 다중화된 멀티채널 오디오 데이터의 광입력
- 출력 오디오 신호의 감시 기능

2.3 인코더 보드(MASIC-ENC)

- MPEG-2 AAC LC(Low Complexity) profile 의 인코딩
- 2 개의 TMS320C6701 DSP 사용
- 최대 5.1 채널까지의 실시간 인코딩 처리
- 모노, 스테레오, 3/0, 2/2, 3/2 채널의 구성(LFE 는 선택사항)
- 48KHz 에서 96kbps(스테레오), 256kbps(4 채널), 320kbps(5.1 채널)의 비트스트림 생성
- PC 와 MASIC-ENC 간의 데이터 전송을 위한 PCI 인터페이스
- MASIC-I 와의 ADI 형식의 인터페이스 제공
- 데이터 입력 및 출력에 의한 4 가지 동작 모드의 제공
- AES-3 형식을 갖는 인코딩 비트스트림의 출력

2.4 디코더 보드(MASIC-DEC)

- MPEG-2 AAC main profile 의 디코딩 및 최대 8 채널까지의 실시간 디코딩 처리
- 1 개의 TMS320C6701 DSP 사용
- 모노, 스테레오, 3/0, 2/2, 5+1, 7+1 채널의 구성(LFE 는 선택사항)
- PC 와 MASIC-DEC 간의 데이터 전송을 위한 PCI 인터페이스
- MASIC-O 와의 ADI 형식의 인터페이스 제공
- 데이터 입력 및 출력에 의한 4 가지 동작 모드의 제공
- AES-3 형식을 갖는 인코딩 비트스트림의 입력

3. 하드웨어 구성

MASIC 시스템에서 가장 중요한 보드인 DSP 보드의 전체적인 구성도는 그림 2와 같으며, 각 기능부의 기능은 다음과 같다.

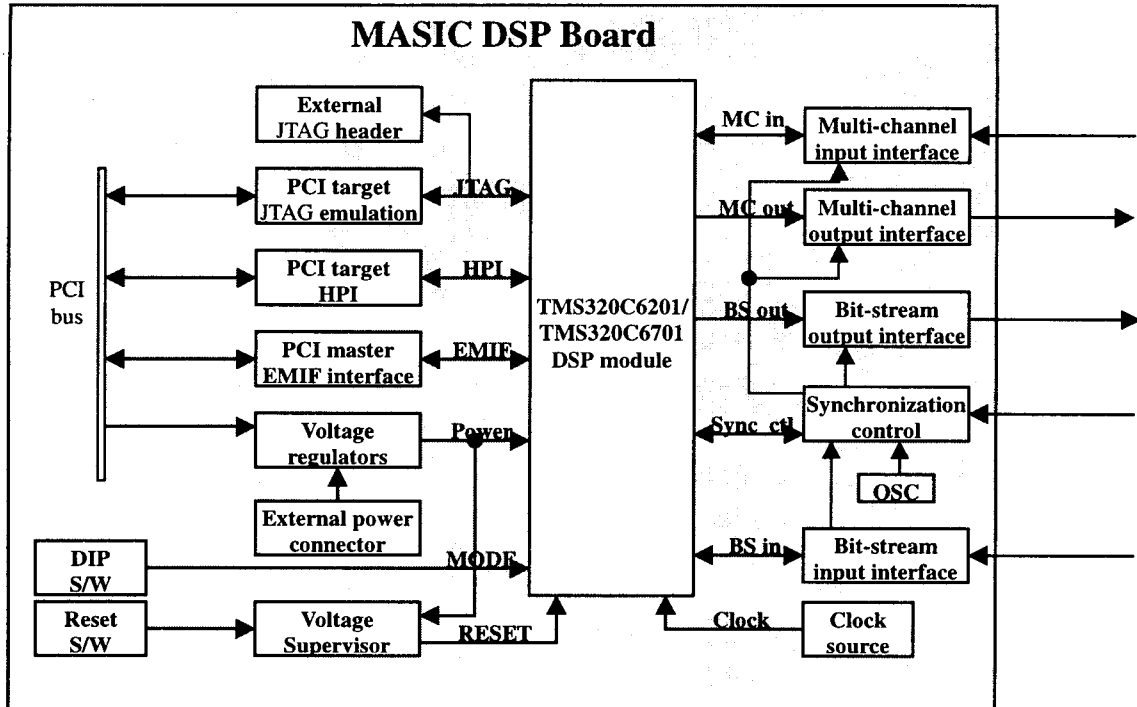


그림 2. DSP 보드의 구성도

3.1 입력/출력 기능부

입력 기능부는 그림 3에 보여준 것처럼 PC의 하드 디스크 또는 외부 시스템으로부터 아날로그 또는 디지털 형식의 오디오 데이터를 받고, 이 데이터를 입력 데이터 형식에 따라 인코딩 기능부로 전달하며, 외부 시스템과 인코딩 기능부로부터 디코딩 기능부로 비트 스트림 데이터를 직접 전달할 수 있다. 반대로 출력 기능부는 디코딩 기능부에서 디코딩된 디지털 데이터를 받고, PC의 하드 디스크 또는 외부 시스템에 아날로그 또는 디지털 형식으로 오디오 데이터를 전달하며, 인코딩 기능부로부터 외부 시스템으로 인코딩된 비트스트림을 직접 전달할 수 있다. 입력/출력 기능부는 오디오 신호 입출력 및 비트스트림 입출력 등의 서브 기능부로 구분된다.

3.2 인코딩 기능부

인코딩 기능부는 신호 입력, 압축처리, 데이터 출력, 인코딩 제어, 전달 미디어 인터페이스 등의 서브 기능부로 구성되며, 입력/출력 기능부의 오디오 입력 기능부로부터 디지털 오디오를 받아 MPEG-2 AAC의 LC profile에 따라 인코딩을 수행한 후 비트스트림을 외부 시스템에 전달한다.

3.3 디코딩 기능부

디코딩 기능부는 전달 미디어 인터페이스, 데이터 입력, 복원처리, 신호 출력, 디코딩 제어 등의 서브 기능부로 구성되며, 외부로부터 인코딩된 비트스트림을 받아 MPEG-2 AAC main profile 에 따라 디코딩을 수행한 후 오디오 출력 기능부로 디지털 데이터를 전달한다.

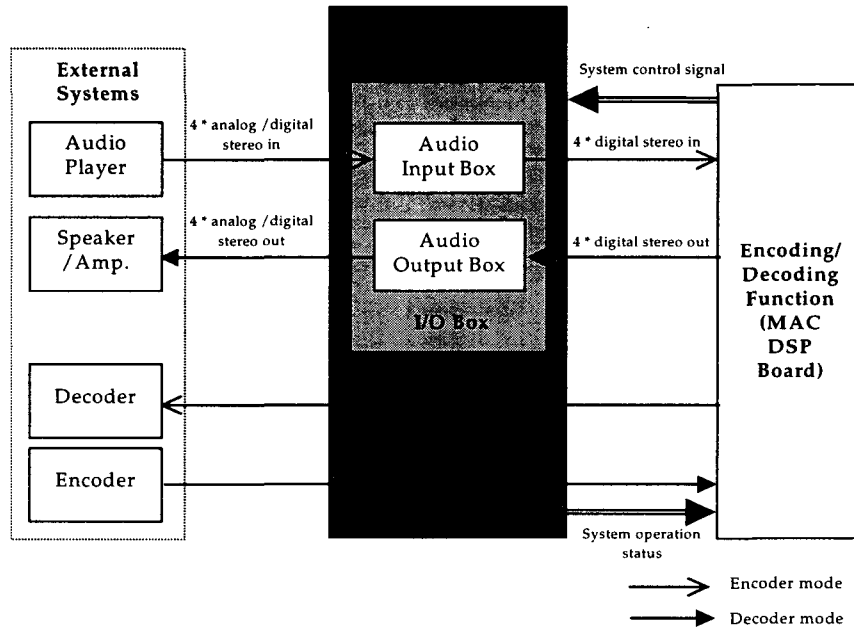


그림 3. MASIC 시스템의 각 기능부별 인터페이스

4. 실시간 소프트웨어 구성

현재의 시스템에서 사용하고 있는 TMS320C6701 DSP의 버전이 샘플 버전이고, 완벽한 제품이 아직 판매되고 있지 않기 때문에 실시간으로 소프트웨어를 구성하기 위해서는 추가적인 기법들이 적용되어야 한다. 현재 알려진 TMS320C6701 DSP의 문제점은 외부 메모리를 사용하는 경우 동기화된 메모리를 사용할 지라도 0 wait-states로 처리할 수 없고, 사용하는 메모리 종류에 따라 추가적인 wait-states가 필요하다는 것과 LDDW 명령어를 사용할 수 없다는 것이다. 그러나, 내부 메모리를 이용하는 경우 이러한 문제점이 해결되기 때문에 가능한 한 내부 메모리를 이용하여 데이터를 처리하여야 하지만 내부 메모리의 크기가 작아서 이를 처리하기 위한 노력이 필요하다.

4.1 인코더의 구현

다른 DSP를 이용한 실험과 C언어로 구성된 프로그램의 시뮬레이션으로부터 MPEG-2 AAC LC의 경우 채널당 약 80 MIPS의 수행능력이 필요하다는 결과를 도출하여 TMS320C6701 DSP의 경우 5.1 채널을 실시간으로 처리하기 위해서는 적어도 2개의 DSP가 필요하다는 사실을 알게 되었다. 첫번째 실행 단계에서 완전한 AAC 알고리즘이 요구되는 메모리의 양과 필요한 성능을 추측하기 위하여 하나의 DSP에서 컴파일되었고, 최적화가 전혀 되지 않은 상태에서 전체 코드양과 MIPS를 측정하였다. 2번째 단계에서 모든 루틴에 대한 몇가지 최적화 레벨을 가지고 프로그램을 컴파일하였다.

마지막 단계에서 인코더의 모듈들이 2개의 DSP에 각각 분리되어 나뉘어 졌는데, 첫번째

DSP 에는 심리음향 모델링부, 필터뱅크부, TNS(Temporal Noise Shaping)부, Intensity/Coupling 부, M/S 부가, 두번째 DSP 에는 나머지의 AAC 알고리즘 모듈이 각각 수행되도록 하였다. 편의상 첫번째 DSP 의 알고리즘을 심리음향 블록, 두번째 DSP 의 알고리즘을 루프(loop) 블록이라고 한다. 두 블록사이의 데이터 전송을 위하여 트랜스포트 데이터 프로토콜을 정의하였고, 추가로 트랜스포트 구조에 멀티채널 오디오를 위한 인터페이스를 추가하였다.

표 1 은 최적화의 3 가지 단계에 대한 결과를 나타낸다. 사용하고자 하는 하드웨어인 TMS320C6701 을 160MHz 에서 동작시킬 때 48kHz 샘플링에서 실시간으로 처리하기 위해서는 34 백만 실행 사이클이내가 되어야만 한다.

표 1. TI 컴파일러 Ver. 2.2 를 사용한 AAC 스테레오 인코더(C source)의 실행시간

Trial	Encoder Million of Cycles/Sec. Worst case		Notes
	PSYCH	LOOP	
Un-optimised	>130	450	Original C-Code
-O0	31	34	Single precision + CMDCT assembler
-O1	28	23	Single precision + CMDCT assembler
-O2	23	29	Single precision + CMDCT assembler

AAC 스테레오에 대한 DSP 의 전체 메모리는 프로그램과 데이터(SDRAM)를 위해 1.67 Mbyte, 테이블 및 버퍼(SBSRAM)를 위해 21,3 Kbyte, 그리고 HEAP 메모리를 위해(SDRAM) 288 KByte 이었고, 전체적으로는 1,97 Mbyte ~ 500 kwords 가 필요하였다. 다음 수행 단계에서 메모리 overlap 기법이 적용되었다. TMS320C6701 DSP 코어가 0 wait-states 를 가지고 외부 메모리를 액세스하지 못한다는 사실을 고려하여 모든 외부 메모리의 데이터들을 내부 메모리로 DMA 를 통하여 복사하였다. 이러한 과정은 내부 메모리가 제한된 크기를 갖고 있기 때문에 매우 조심스럽게 구현되어야만 한다. 만일 데이터가 내부 메모리에 위치한다면 LDDW 명령이 빠른 메모리 액세스를 위해서 사용될 수 있다.

이 단계에서 표준 C 벡터 라이브러리들은 그러한 LDDW 명령어를 사용하여 최적화된 어셈블리 언어로 변경하였다. 다음으로 DSP 사이클을 줄이기 위하여 알고리즘의 수정을 가하였고, 필터뱅크, 블록 스위칭, 주기성분 계산과 콘볼루션과 같은 심리음향 모델이 DSP 어셈블리 언어로 대체되었다. 이러한 과정에서 요구되는 정밀도가 최소한으로 수정되도록 주의깊은 검사가 이루어졌고, 특히, 나누기, 대수함수, 제곱 함수 등과 같은 C 레벨의 벡터 계산은 아주 최적화가 잘된 DSP 에 어셈블리 언어로 재 작성되었다.

또한, 양자화와 부호화를 포함하는 루프 블록중 모든 스펙트럼 라인을 처리할 때마다 내부 루프를 호출하는 양자화기 코어와 허프만 코드 비트 카운터 등에 대해서 몇가지 알고리즘을 수정하였고, 간략화 함으로써 최적화를 수행하였다. 이 결과 표 2 와 같이 2 개의 TMS320C6701 DSP 의 거의 100%를 사용하는 멀티채널 MPEG-2 AAA 인코더의 실시간 처리가 가능하였고, 각각의 DSP 는 약 2Mbyte 의 SDRAM 을 필요로 하는 것으로 나타났다. 본 논문에서는 160MHz 클럭을 사용하

는 경우 샘플 DSP 칩에서 클럭 여유가 거의 없는 인터페이스 구조로 인하여 SBSRAM 에서 가끔 데이터 에러가 발생하였기 때문에 이 메모리를 사용하지 않고, 전적으로 half clock 으로 동작하는 SDRAM 을 사용하였다.

표 2. 최종적인 멀티채널 AAC 인코더의 DSP 실행시간

Trial	Encoder Million of Cycles/Sec. Worst case		Notes
	PSYCH	LOOP	
-O3	3.3	3.4	Realtime break condition for Loops

4.2 디코더의 구현

디코더의 경우 인코더와는 달리 160MHz로 동작하는 1개의 DSP를 이용하여 8채널까지 AAC 비트스트림을 디코딩하도록 구현하였다. 디코더의 구현도 인코더와 마찬가지로 C언어로 구성된 소스 프로그램을 이용하여 실시간 동작을 위한 최적화를 수행하였다. 디코더를 구현할 당시에 TMS320C6701 DSP가 생산되지 않았기 때문에 시뮬레이터를 이용하여 프로그램을 구현하였다. 그러나, 시뮬레이터가 DSP의 외부 메모리 액세스를 위한 메모리 뱅크 충돌과 wait-states 등을 고려할 수 없기 때문에 목표 시스템에서의 처리와는 실행시간에 있어서 많은 차이를 나타나게 되었다. 특히, 48kHz의 샘플링 주파수를 갖는 디코더의 경우 실시간으로 처리되기 위해서는 한 프레임당 3.424백만 사이클보다 작아야 한다. 이것은 스테레오 오디오 한 프레임을 디코딩하기 위하여 856,000 사이클보다 작아야 한다는 것을 의미한다.

따라서, C언어로 구성된 프로그램의 실행시간을 분석하였고, 이중 가장 실행시간에 중요한 요인으로 작용하는 모듈을 최적화하기로 하였다. 실험 결과 실시간의 제약보다 약 3배의 사이클이 필요하고, 디코더의 실시간 최적화를 위하여 허프만 코딩 모듈과 필터뱅크 모듈이 가장 많은 실행 사이클을 소비하는 것으로 분석되었다. 그래서, IMDCT의 필터뱅크 모듈은 어셈블리 언어로 재구성되었고, TI의 linear assembler가 일부의 기능에서는 최적으로 가능한 코드를 만들지 못하기 때문에 일일이 손으로 코드를 분석하면서 최적화를 수행하였다. 또한, 허프만 코드를 디코딩하기 위하여 빈도가 높은 코드표가 빨리 찾아지도록 모든 코드표가 재구성하였으며, 이와 관련된 코드도 재구성되었다. 이결과 표 3과 같은 결과를 얻었다.

표 3. 시뮬레이터를 이용한 디코딩 처리 결과(스테레오 채널)

Function	Cycles per frame	
Huffman Coding	329,600	76%
Filter Bank	29,400	7%
Other Tools	75,000	17%
TOTAL	434,000	100%

이 결과를 이용하여 실제의 하드웨어 시스템인 TMS320C6701 DSP 보드에서 실행해본 결과 프레임당 약 6백만 사이클이 소요되었고, 이것은 앞에서 언급한 것과 같은 DSP의 문제점에 기인한 것으로 분석되었다. 따라서, 또 다른 최적화가 다음과 같이 이루어졌다.

- DSP 의 내부 프로그램 캐쉬(16KWords)를 스위칭함 : 1,800,000 cycles/frame
- 전체 시스템 스택을 DSP 의 내부 데이터 메모리로 배치 : 1,200,000 cycles/frame
- 중요한 코어 함수들을 최적화된 어셈블리 언어로 재작성 : 1,050,000cycles/frame
- 메모리 overlap 기법의 적용 및 DMA 를 통한 다양한 외부 메모리 버퍼의 내부 메모리 버퍼로 swap : 678,000 cycles/frame

이 결과는 앞에서 분석한 8 채널의 실시간 조건에 부합되는 것이지만 이러한 노력은 TMS320C6701 DSP 의 외부 메모리 인터페이스 기능이 미약하기 때문에 이루어진 것이다.

5. 결 론

범용 DSP 인 TMS320C6701 을 이용하여 멀티채널 오디오 부호화 방식인 MPEG-2 AAC 인코더 및 디코더 시스템(MASIC)을 실시간으로 구현한 결과를 기술하였다. MASIC 시스템은 입력장치(Input Box, MASIC-I), 출력장치(Output Box, MASIC-O), 인코더 보드(MASIC-ENC), 그리고 디코더 보드(MASIC-DEC) 등 4 개의 모듈로 구성되어 있으며, 인코더의 경우 160MHz 로 동작하는 2 개의 TMS320C6701 DSP 를 이용하여 LC profile 을 5 채널까지, 디코더의 경우 1 개의 TMS320C6701 DSP 를 이용하여 8 채널까지 실시간으로 처리할 수 있다.

MASIC 시스템은 멀티채널 오디오의 저작 tool, 디지털 오디오 방송(DAB), 디지털 텔레비전 방송, AOD(Audio on Demand) 또는 MOD(Music on Demand), 그리고 MP3 플레이어의 차세대 버전인 AAC 플레이어 등에 응용될 수 있다. 이 시스템은 현재 세계에서 처음으로 MPEG-2 AAC 국제 표준을 실시간으로 구현한 시스템으로써, 1999 년 5 월초에 독일 뮌헨에서 개최된 106 차 AES(Audio Engineering Society) Convention 에 전시하여 많은 관심과 호평을 받았으며, 서울에서 개최되는 KOBA 99 에도 전시될 예정이다.

감사의 글

본 논문은 정보통신부의 국제공동연구지원사업인 “차세대 멀티미디어용 고품질 오디오 처리 기술 개발”과제의 일환으로 독일 Fraunhofer 연구소와 공동으로 수행한 연구결과로써 정보통신부 담당자와 양 기관의 관련 연구원들의 노력에 감사를 드립니다.

참고문헌

- [1] G. Theile, "The New Sound Format 3/2-Stereo," The 94th AES Convention, Berlin, preprint 3550, Mar. 1993.
- [2] ISO/IEC, "Information technology – Generic coding of moving pictures and associated audio information, Part 7: Advanced audio coding", 1997.
- [3] ITU-R Recommendation BS. 1116, "Methods for the subjective assessment of small impairments in audio systems including multichannel sound systems," Geneva, 1994.
- [4] Texas Instrument, TMS320C62x/C67x CPU and Instruction Set Reference Guide, 1998.
- [5] ISO/IEC JTC1/SC29/WG11 MPEG. International Standard IS 13818-7 Information Technology - Generic Coding of Moving Pictures and Associated Audio, Part 7: Advanced Audio Coding, 1997.