

MPEG-2 AAC 오디오 인코더 및 디코더 구현

홍진우, 장대영, 김진웅
ETRI-무선방송기술연구소 방송기술연구부

Implementation of Audio Encoder and Decoder Using MPEG-2 AAC

J. W. Hong, D. Y. Jang, and J. W. Kim
ETRI-Radio Broadcasting Lab., Broadcasting Technology Department
jwhong@etri.re.kr

요약

본 논문에서는 멀티채널 오디오 부호화 방식인 MPEG-2 AAC(Advanced Audio Coding) 국제 표준을 수용한 AAC 인코더 및 디코더의 실시간 구현에 대해 기술한다. 범용 DSP 인 TMS320C6701 DSP를 이용한 하드웨어 플랫폼과 이 플랫폼에서 실시간으로 동작되는 인코더와 디코더 소프트웨어를 설계, 개발(MASIC 시스템)하였다.

구현한 MASIC 시스템은 오디오 입력 장치, 출력 장치, 인코더 보드, 그리고 디코더 보드로 구성되어 있으며, 개인용 컴퓨터의 PCI 슬롯을 이용하여 인코더의 경우 최대 6 채널의 오디오를, 디코더의 경우 8 채널의 오디오를 실시간 동작으로 처리할 수 있다. 인코더 및 디코더의 실시간 처리를 위한 소프트웨어 최적화 기술 및 인코더와 디코더의 연동시험에 대해서도 기술하며, 개인용 컴퓨터에서 실시간으로 수행되는 스테레오 AAC 디코더 소프트웨어의 개발 결과를 기술한다.

1. 서론

ISO/IEC JTC1 산하 MPEG(moving Picture Experts Group) 오디오 그룹에서는 1992년에 스테레오 채널 오디오 부호화 방식으로 MPEG-1 오디오 국제 표준(표준번호 ISO/IEC 11172-3)을 제정하였으며, 1994년 멀티채널 오디오 부호화 방식으로 MPEG-2 오디오 BC(Backward Compatibility)(표준번호 : ISO/IEC 13818-3)를 제정하였다.

그러나, MPEG-2 오디오 BC는 MPEG-1 오디오와의 호환성을 유지하여야 한다는 제한 때문에 압축율과 오디오 품질면에서 다른 멀티채널 오디오 부호화

방식보다 열악한 결과를 나타나게 되었다. 따라서, MPEG 오디오 그룹에서는 MPEG-1 오디오와 호환성을 유지하지 않는(NBC : Non-Backward Compatibility) 멀티채널 오디오 부호화 방식인 MPEG-2 AAC(Advanced Audio Coding) 국제표준을 1997년에 제정하였다.

MPEG-2 AAC 부호화 방식은 지금까지 개발된 멀티채널 오디오 부호화 방식중 최신 기술(state of the art)로써, 압출율 및 품질면에서 가장 우수한 것으로 인정받고 있으며, 제한된 대역폭을 갖는 분야로부터 고품질을 요구하는 응용분야까지 적용할 수 있는 특징을 가지고 있다. 또한, 관련 기술들이 멀티미디어 용 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-R에서 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 설명하고, 개인용 컴퓨터에서 실시간으로 수행되는 스테레오 AAC 디코더 소프트웨어의 개발 결과를 기술한다.

2. MASIC 시스템의 기능 및 규격

MASIC 시스템은 부동 소수점을 갖는 범용 디지털 신호처리기인 TI사의 TMS320C6701 DSP를 이용하여 구성하였으며, 개인용 컴퓨터(PC) 환경에서 오디오 저작 도구로 활용할 수 있도록 PCI 인터페이스를 적용하였다.

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

2.1 입력장치(MASIC-I)

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

2.2 출력장치(MASIC-O)

- 멀티채널 오디오 신호의 출력 및 신호 감지 기능
- 최대 8 채널의 아날로그 오디오 출력 인터페이스
- S/PDIF 형식을 갖는 최대 8 채널(스테레오 4개 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

2.3 인코더 보드(MASIC-ENC)

- MPEG-2 AAC LC(Low Complexity) profile의 인코딩
- 2개의 TMS320C6701 DSP 사용
- 최대 5.1 채널까지의 실시간 인코딩 처리
- 모노, 스테레오, 3/0, 2/2, 3/2 채널의 구성(LFE는

선택사항)

- 48KHz에서 96kbps(스테레오), 256kbps(4 채널), 320kbps(5.1 채널)의 비트스트림 생성
- PC와 MASIC-ENC 간의 데이터 전송을 위한 PCI 인터페이스
- MASIC-I와의 ADI 형식의 인터페이스 제공
- 데이터 입력 및 출력에 의한 4가지 동작 모드의 제공
- AES-3 형식을 갖는 인코딩 비트스트림의 출력

2.4 디코더 보드(MASIC-DEC)

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

3. 하드웨어 구성

MASIC 시스템의 전체적인 구성도는 그림 1과 같으며, 각 기능부의 기능은 다음과 같다.

3.1 입력/출력 기능부

입력 기능부는 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 기능부

로 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 profile으로 디코딩을 수행한 후 오디오 출력 기능부로 디지털 데이터를 전달한다.

그림 2는 인코딩 기능부와 디코딩 기능부에 중심이 되는 DSP 보드의 내부 구성도를 나타낸다.

4. 실시간 소프트웨어 구성

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 5.1채널을 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 DSP에는 심리음향 모델링부, 필터뱅크부, TNS(Temporal Noise Shaping)부, Intensity/Coupling부, M/S 부가, 두번째 DSP에는 나머지의 AAC 알고리즘 모듈이 각각 수행되도록 하였다. 편의상 첫번째 DSP의 알고리즘을 심

리음향 블록, 두번째 DSP의 알고리즘을 루프(loop)블록이라고 한다. 표 1은 최적화의 3가지 단계에 대한 결과를 나타낸다. 사용하고자 하는 하드웨어인 TMS320C6701을 160MHz에서 동작시킬 때 48kHz 샘플링에서 실시간으로 처리하기 위해서는 3.4백만 실행 사이클이내에 되어야만 한다.

AAC 스테레오에 대한 DSP의 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 kwords가 필요하였다. 다음 수행 단계에서 메모리 overlap 기법이 적용되었다. TMS320C6701 DSP 코어가 0 wait-states를 가지고 외부 메모리를 액세스하지 못한다는 사실을 고려하여 모든 외부 메모리의 데이터들을 내부 메모리로 DMA를 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.

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, 간략화함으로써 최적화를 수행하였다. 이 결과 표 2와 같이 2개의 TMS320C6701 DSP의 거의 100%를 사용하는 멀티채널 MPEG-2 AAA 인코더의 실시간 처리가 가능하였고, 각각의 DSP는 약 2Mbyte의 SDRAM을 필요로 하는 것으로 나타났다. 본 논문에서는 160MHz 클럭을 사용하는 경우 샘플 DSP 칩에서 클럭 여유가 거의 없는 인터페이스 구조로 인하여 SBSRAM에서 가끔 데이터 에러가 발생하였기 때문에 이 메모리를 사용하지 않고, 전적으로 half clock으로 동작하는 SDRAM을 사용하였다.

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 프로그

램을 이용하여 실시간 동작을 위한 최적화를 수행하였다. 디코더를 구현할 당시에 TMS320C6701 DSP가 생산되지 않았기 때문에 시뮬레이터를 이용하여 프로그램을 구현하였다.

따라서, C 언어로 구성된 프로그램의 실행시간을 분석하였고, 이중 가장 실행시간에 중요한 요인으로 작용하는 모듈을 최적화하기로 하였다. 실험 결과 실시간의 제약보다 약 3 배의 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. 이 결과 표 3 과 같은 결과를 얻었다.

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

- DSP의 내부 프로그램 캐쉬(16KWords)를 스루칭함 : 1,800,000 cycles/frame
- 전체 시스템 스택을 DSP의 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 DMA를 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 DSP의 외부 메모리 인터페이스 기능이 미약하기 때문에 이루어진 것이다.

4.3 PC 용 AAC 디코더 플레이어의 구현

본 논문에서는 개발된 디코더를 이용하여 범용 PC(MMX 기능을 가진 펜티엄 PC, 윈도우 OS)에서 실시간으로 수행되는 AAC 디코더 플레이어 소프트웨어를 그림 3 과 같이 개발하였다. 프로그래밍 언어로는 C/C++ 와 어셈블리 언어를 사용하였으며, AAC main 프로파일과 LC 프로파일을 디코딩하며, 모노 채널 및 스테레오 채널 오디오를 재생할 수 있다.

CPU 점유율은 main 프로파일의 경우 Pentium MMX 266MHz 에서 20% 이내로, LC 프로파일의 경우 15% 이내로 처리가 가능하고, GUI 인터페이스, Play, Stop, Pause 버튼, Progressive bar, Play List 지원, Repeat 지원, Information display 등의 기능이 있다.

5. 결론

범용 DSP인 TMS320C6701을 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 160MHz로 동작하는 2 개의 TMS320C6701 DSP를 이용하여 LC profile을 5 채널까지, 디코더의 경우 1 개의 TMS320C6701 DSP를 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.

감사의 글

본 논문은 정보통신부의 국제공동연구지원사업인 "차세대 멀티미디어용 고품질 오디오 처리 기술 개발"과제의 일환으로 독일 Fraunhofer 연구소와 공동으로 수행한 연구결과로써 정보통신부 담당자와 양기관의 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] ITU-R Recommendation BS. 1116, "Methods for the subjective assessment of small impairments in audio systems including multichannel sound systems," Geneva, 1994.
- [3] 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.

- [4] M. Bosi, K. Brandenburg, S. Quackenbush, L. Fielder, etc., "ISO/IEC MPEG-2 Advanced Audio Coding," Journal of AES, Vol. 45, No. 10, Oct. 1997.

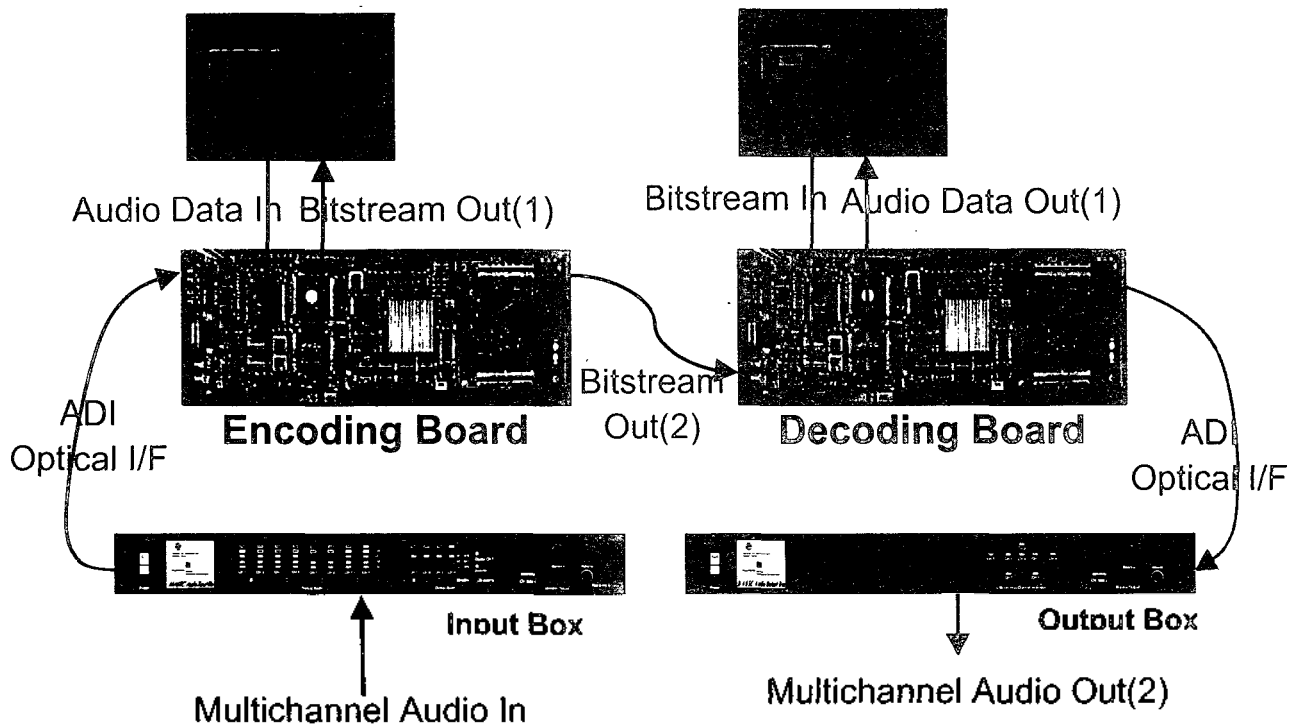


그림 1. MPEG-2 AAC 인코더 및 디코더(MASIC) 시스템의 전체 구성도

표 1. AAC 스테레오 인코더의 실행시간

처리과정	PSYCH (Million Cycles/s)	LOOP (Million Cycles/s)	내용
비최적화	>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

표 2. 최종적인 멀티채널 인코더의 DSP 실행시간

처리과정	PSYCH (Million Cycles/S)	LOOP (Million Cycles/S)	내용
-O3	3.3	3.4	Loop 에 대한 실시간 break 조건 사용

표 3. 시뮬레이터를 이용한 디코딩 처리 결과(2CH)

함수	Million Cycles / frame	
Huffman Coding	329,600	76%
Filter Bank	29,400	7%
Other Tools	75,000	17%
합계	434,000	100%

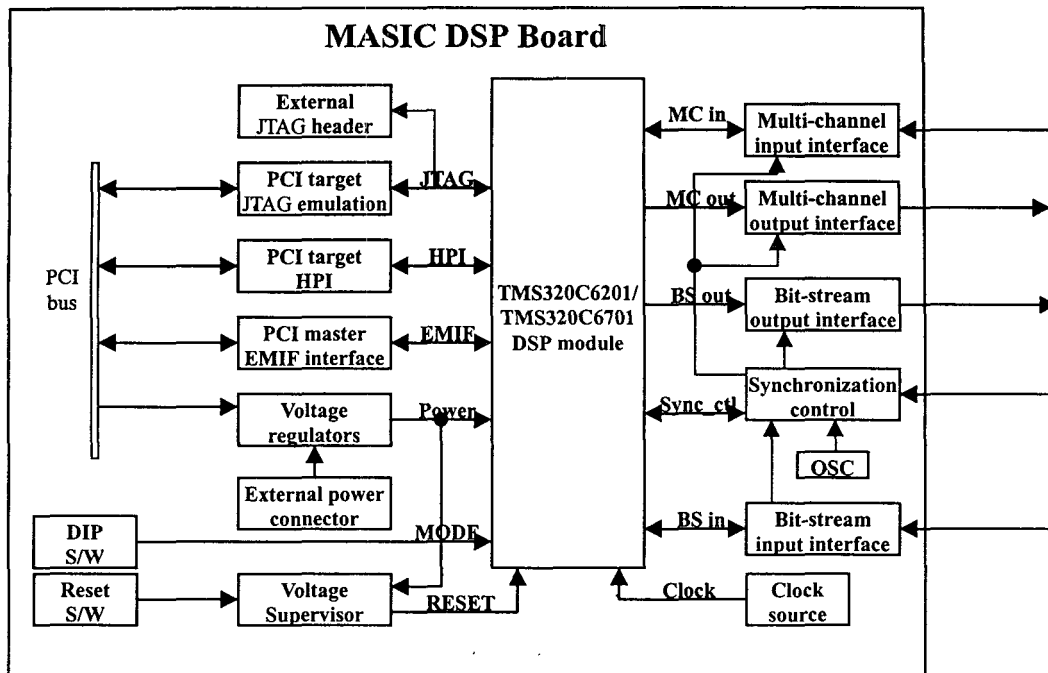


그림 2. 인코더 및 디코더용 DSP 보드의 내부 구성도

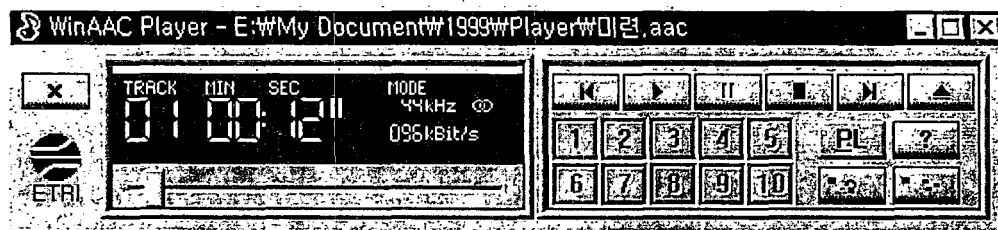


그림 3. AAC 디코더 플레이어의 윈도우 표현