

범용 DSP를 이용한 MPEG-2 오디오 부호화기의 성능 개선

오현오*, 김성윤*, 윤대희*, 차일환*, 이준용**

*연세 대학교 전자공학과 음향, 음성 및 신호처리 연구실

**KBS 기술 연구소

An Enhancement of the MPEG-2 Audio Encoder Using General DSPs

Hyen-O Oh*, Sung-Youn Kim*, Dae-Hee Youn*, Il-Whan Cha*, Jun-Yong Lee**

*Yonsei University, ASSP Lab.

**KBS Technical Research Institute, System Division

Abstract

The ISO(International Standard Organization) has standardized MPEG-2 audio. The MPEG-2 audio compression algorithm is based upon subband analysis and exploits the human auditory characteristics to achieve a low bit rate with minimum perceptual loss of audio signal quality.

This thesis presents an enhanced MPEG-2 audio encoder using multiple TMS320C30 general purpose DSP's. The developed system is made up of five slave boards and one master board. Each slave board performs subband analysis and psychoacoustic parameter calculation for one channel, and the master board manages bit allocation, quantization, and bit-stream formatting for all channels. Parallel processing and pipelining techniques are used in hardware structure and fast algorithms are applied in each subroutine to implement a real-time process. The implemented system supports multichannel up to 5.1 and various bitrates.

I. 서론

기존의 아날로그 디스크(LP)를 대체하여 탄생한 컴팩트 디스크(CD)를 필두로 오디오 신호의 디지털화는 한층 가속화되고 있는 추세이다. 아날로그 오디오에 비하여 디지털 오디오가 갖는 장점은 대역폭과 동적 영역(Dynamic Range)이 넓고 복사할 때 음질 손상이 없다는 점이다. 그러나 15~20 KHz 정도의 광대역 오디오 신호를 압축하지 않고 PCM(Pulse Code Modulation)방식으로 부호화하면 샘플당 16비트 이상이 필요하며 HDTV(High Definition Television)의 경우 48KHz 표본화 주파수에 5채널 오디오 신호를 표준으로 하고 있기 때문에 이를 전송하기 위해서는 3.84 Mbit/s 이상의 막대한 채널 용량이 필요하다. 전송과 저장에서의 이러한 문제점을 줄이기 위해서는 신호의 압축이 불가피하며 압축 부호화한 오디오 신호는 복호화한 후의 주관적 음질이 기존 음질과 거의 동일하도록 유지되는 효과적인 압축 기법이 사용되어야만 한다.

이에 대한 대응책으로 국제 표준화 기구(ISO :

[†] 이 논문은 한국 방송 공사의 연구비 지원에 의한 것임.

International Standard Organization) 산하의 동영상 전문가 그룹(MPEG:Moving Picture Experts Group)에서는 동영상과 더불어 CD 수준의 디지털 오디오를 1.5 Mbit/s로 압축할 수 있는 MPEG-1 표준안 ISO/IEC 11172-3을 1992년에 확정하였다[1]. 그러나 MPEG-1 표준안은 스테레오 신호만을 기준으로 표준화 하였기 때문에 HDTV와 같은 방송 매체에 적용할 경우 다채널, 음성 다중 등의 부가 서비스에 대한 지원이 어려우며, 5채널 이상의 오디오를 처리하기 위해서는 여전히 높은 전송률을 필요로 한다. MPEG-2(ISO/IEC 13818-3)는 위와 같은 조건을 만족하면서 6 ~ 40 Mbit/s의 전송률로 동영상과 다채널 디지털 오디오를 압축할 수 있는 방식으로 94년 11월에 국제 표준안이 결정되었으며 97년 2월에 불분명했던 내용을 명확히 정의한 표준안을 새롭게 발표하였다[2][3].

MPEG-2 오디오 표준안의 기본 특성은 다음과 같다. 먼저 좌, 우 채널 외에 중앙 및 좌, 우 서라운드, 저음 보강(Low Frequency Enhancement) 채널을 추가한 5.1채널을 지원하여 현장감, 음상 정위 등의 효과를 가져오며 최대 7개 언어의 음성 다중

방송을 지원할 수 있다. 또한 다채널 오디오 신호의 중복성을 제거할 수 있는 다채널 부호화 방식이 제안되었으며 음성과 같이 대역이 작은 신호를 위한 새로운 표본화 주파수인 24KHz, 22.05KHz, 16KHz를 지원하게 되었다. 비트율은 32Kbps에서 최대 1130Kbps를 지원하여 다양한 응용 분야에 맞게 선택할 수 있으며 MPEG-1과의 호환이 가능하도록 표준화하였다.

MPEG에서 사용된 오디오 압축 알고리즘은 서브밴드 부호화와 심리 음향 모델을 기본으로 하는데, 서브밴드 부호화를 통해서 신호의 통계적 중복성을 제거하고 심리 음향 모델을 통해서 귀의 청각 특성, 특히 마스킹 현상을 이용하여 신호의 지각적인 중복성을 제거한다. 마스킹은 음압이 다른 두 음이 존재할 때 음압이 큰 음에 의해 작은 음이 들리지 않는 현상으로서 달팽이관의 기저막에 있는 청신경의 반응에 의해서 일어나며, 주로 주파수 영역에서 설명되어 진다[4]. 또한 MPEG-2에서는 채널 간의 중복성을 제거하여 압축율을 보다 향상시킬 수 있는 여러 가지 다채널 복합 부호화를 지원한다.

본 논문에서는 TI(Texas Instrument)사의 범용 디지털 프로세서인 TMS320C30 칩을 이용하여 구현되었던 MPEG-2 계층 2(Layer II) 오디오 부호화 실시간 병렬처리 시스템의 성능을 개선하였다[5][6]. 제작된 시스템은 각 1 채널의 오디오 신호를 처리하기 위한 Slave 보드 5개와 채널 매트릭싱(Matrixing), 비트 할당, 양자화, 비트열 포매팅, 그리고 PC와의 인터페이스를 위한 Master 보드 1개로 이루어져 있으며 각각의 보드에는 TMS320C30 칩이 두 개씩 장착되어 있다.

구현된 하드웨어내에서 높은 비트율로 다채널 부호화가 가능하도록 몇가지 고속 알고리즘을 적용하였으며 메모리 배치와 반복 루프의 최적화를 통해 서브루틴별 수행시간을 개선하였고 각 CPU사이의 적절한 업무분담과 동기화를 통해 코딩 시간 지연(Coding Delay)을 최소화하였다.

II. 개선된 시스템 하드웨어

구현된 실시간 부호화 시스템의 전체적인 구성도를 그림 1에 나타내었다. 전체 시스템은 5개의 Slave 보드와 1개의 Master 보드, 그리고 이것을 서로 연결해 주는 Back plane 보드로 구성되어 있다. 각 Slave 보드에서는 한 채널 데이터를 입력 받아 서브밴드 분석과 심리 음향 모델링을 수행하고 Master 보드에서는 각 Slave 보드에서 처리된 결과

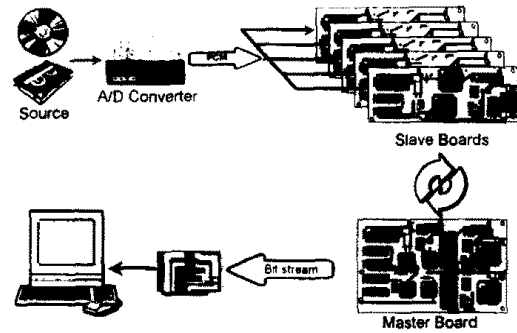


그림 1. MPEG-2 부호화기의 전체 구성도

를 모아서 최종적인 비트열을 만들어 PC로 전송한다.

각 채널별 서브밴드 분석과 심리 음향 모델을 담당하는 Slave 보드의 하드웨어 구성은 다음과 같다.

- 두개의 TMS320C30 CPU
- CPU간의 정보 전달을 위한 듀얼 포트 RAM(2K word)
- 프로그램과 데이터를 저장하기 위한 로컬 RAM (32K word * 2)
- stand-alone 동작을 위한 프로그램이 담겨 있는 로컬 ROM (32K word * 2)
- Master 보드로의 데이터 전송을 위한 듀얼 포트 RAM (2K word)
- 메모리 어드레스를 디코딩하여 웨이트 수를 변경해주는 회로

Slave 보드는 각 채널별로 들어온 1채널 입력 샘플에 대해서만 처리를 하기 때문에 다채널 확장시 별도의 성능 개선이 필요하지 않다. 다만 stand-alone 동작을 하기위해 로컬 ROM과 로컬 RAM을 병행하여 사용하게 되는데, 이에 따른 문제점을 하드웨어적으로 해결하기 위해 메모리 어드레스 디코딩 과정에 레디(Ready) 신호를 발생시키는 회로를 추가하였다. 어드레스 디코딩 소자로는 GAL22V10을 사용하였다.

다채널 확장을 위해 개선된 Master 보드의 구성은 그림 2와 같다. 최대 5채널에 대해 비트 할당, 양자화, 그리고 비트열 포매팅을 수행하는 Master 보드는 두 개의 TMS320C30, 96K ROM과 64K RAM, 데이터 전송을 위한 2K 듀얼 포트 RAM, 리셋 회로와 클럭 회로, 어드레스 디코딩 회로로 이루어져 있고 Slave 보드와의 데이터 교환을 위하여 Back Plane 커넥터를 통하여 각 Slave 보드의 듀얼 포트 RAM과 연결되어 있다. 다채널 처리시 확대되는 Master 보드의 역할을 대비하여 한 개의 CPU가 추가되었고 사용자 요구사항을 PC로부터 받아들일

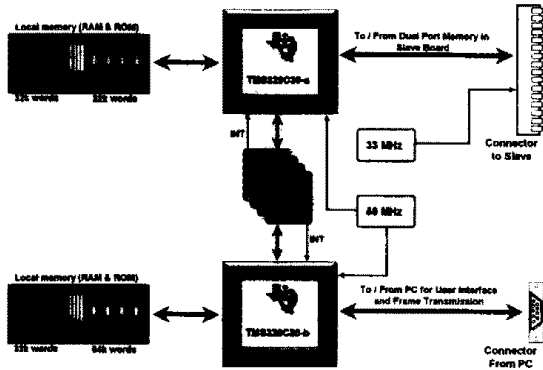


그림 2. 개선된 Master 보드의 구성도

수 있는 로직이 추가 되었다. 또, stand-alone 시스템을 구축하기 위하여 ROM과 RAM의 크기를 늘렸다. 그리고 Master 보드에 장착될 CPU는 최대 50MHz CPU에서부터 Slave와 같은 기존의 33MHz CPU까지를 지원할 수 있도록 Dual Clocking 회로를 추가하였다.

입력된 16비트 PCM 샘플은 Slave 보드의 TMS320C30에 내장된 직렬 포트를 이용하여 메모리에 저장되고 부호화 된다. 최종 처리가 끝난 비트열은 TMS320C30에 내장된 타이머와 직렬 포트(Serial Port)를 이용하여 Master에서 PC로 전송되고 PC에서는 직렬-병렬 변환기(Serial-to-Parallel Converter)와 DMA(Direct Memory Access)를 사용하여 전송된 비트열을 하드 디스크에 저장한다. 저장한 데이터는 다채널 D/A로 전송되어 사용자가 원하는 채널 구성대로 재생하여 음질을 평가할 수 있고, 비디오 비트열과 통합하기 위한 시스템으로도 전송할 수 있다.

완성된 MPEG-2 오디오 부호화기는 stand-alone 동작을 위해 로컬 ROM에 필요한 모든 소프트웨어를 저장해두고 리셋시에 프로그램을 로컬 RAM으로 옮겨와 부호화과정을 수행하게 된다.

III. 소프트웨어 성능 개선

MPEG-2 실시간 부호화기의 소프트웨어를 구성하는데 있어서 중요한 점은 구현된 하드웨어의 성능하에서 최상의 음질과 높은 비트율, 그리고 다양한 선택사항을 지원할 수 있도록 하는 것이다. 이를 위해서 각 처리단계에 실시간 구현에 적합한 고속 알고리즘을 적용하고 각각의 서브루틴을 최적화시킨 후 이들의 수행시간을 바탕으로 프로세서간에 적절한 업무 분담을 하는 작업이 필요하다.

MPEG-2 부호화 알고리즘은 그림 3에 나타난 것

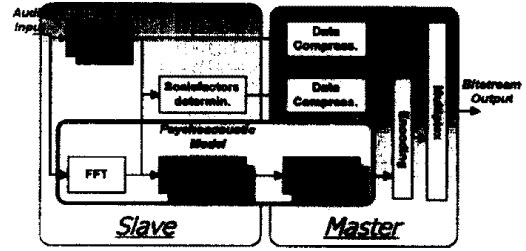


그림 3. MPEG-2 부호화 과정의 할당

과 같이 Slave 보드에서 수행되는 작업과 Master 보드에서 수행되는 작업으로 크게 나눌 수 있다. 각 서브루틴이 채널별로 독립적으로 처리될 수 있는 성격인지, 아니면 모든 채널의 정보를 함께 지녀야 처리 가능한지의 여부에 따라 작업 할당이 이루어 지는데, 전자의 경우 Slave 보드에 할당하고 후자는 Master 보드에서 처리해야만 한다. Slave에서 처리 가능한 작업을 Master로 가져오게 되면 Master는 최대 5채널 데이터를 순차적으로 처리해야 하기 때문에 5배 이상 수행 시간이 연장되고 작업량에 부담을 갖게 된다. 따라서 Slave에서는 가능한 모든 작업을 처리해서 전송에 필요한 최소한의 결과만을 Master로 넘겨주도록 해야한다.

그림 4는 5채널 처리를 기준으로 각 CPU별로 할당된 작업과 개선된 CPU사이의 인터럽트를 통한 동기화 과정을 나타낸 흐름도이다. 입력 샘플 버퍼에서부터 Slave-b까지는 1채널에 대해 보여주고 있고, Master의 처리과정은 5채널을 모두 처리하고 있음을 나타낸다. Slave-b까지의 처리는 5개의 Slave 보드에서 각각의 입력 샘플에 대해 같은 방법으로 병렬적으로 처리된다. 그림에서 화살표는 각 서브루틴의 수행이 끝나고 다음 과정에 필요한 데이터를 듀얼포트 RAM을 통해 다른 CPU에 넘겨주면서 인터럽트를 발생시키는 것을 나타낸다. 한 프레임에 대한 처리를 하는 동안에도 새로 들어오는 입력 샘플을 받기 위해 입력 버퍼는 2개를 사용하여 스위

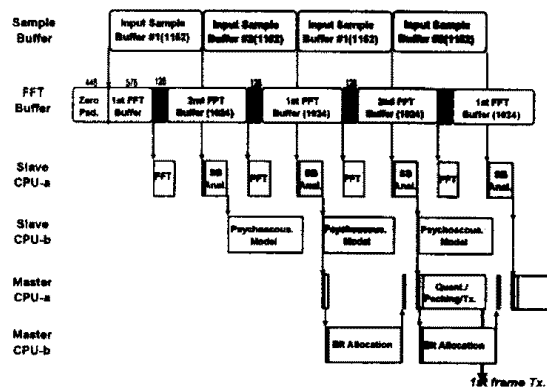


그림 4. 각 CPU의 작업 할당과 동기화

칭시키고 있다. 각 CPU의 서브루틴 처리가 1152개의 다음 프레임 샘플이 모두 들어오기 전에 끝나지 못하면 실시간 처리 실패가 발생한다.

수행시간 개선을 위해 한 프레임 처리를 위한 36회의 32x64 분석행렬 연산이 많은 수행시간을 요구하는 서브밴드 분석과정의 경우 행렬의 커널이 DCT(Discrete Cosine Transform)의 커널과 유사한 점을 이용하여 입력 벡터를 변환하여 고속 DCT로 치환할 수 있는 알고리즘을 적용하였다[7]. 고속 DCT로는 범용 DSP칩에 적용할 때 가장 우수한 성능을 보이는 Lee's 고속 DCT 알고리즘을 사용하였다[8]. 또한 Slave-b에서 전담하고 있는 심리 음향 모델의 경우 로그 연산 과정 및 반복 루틴의 최적화를 통해 수행시간을 개선하였다. Master 보드에서 수행되는 과정중 채널수가 증가함에 따라 가장 많은 수행시간이 필요한 비트 할당의 경우 5채널 640Kbps를 할당하기 위해서는 62.5msec나 걸려서 구현된 하드웨어상에서 실시간 처리가 사실상 불가능하다. 본 논문에서는 비트율에 따른 음원별 비트 할당 인덱스의 통계적 특성을 고려하여 사전에 비트를 할당하는 방식으로 수행시간을 줄이는 알고리즘을 사용하였다. 사전 할당되는 양에 따라 수행시간의 조절이 가능하므로 실시간 처리가 가능한 선에서 비트율에 따라 음질 저하가 생기지 않는 최적의 값을 찾아 적용하였다.

IV. 실험 결과

구현된 MPEG-2 실시간 부호화기의 실험은 다음과 같이 이루어졌다. 5채널의 아날로그 음원을 제공하는 THX 프로세서를 A/D 변환기(Analog-to-Digital Converter)에 연결한 다음 A/D의 출력을 각 Slave 보드의 직렬 포트에 연결하여 부호화한 후 생성된 최종 비트열은 PC로 전송하여 하드 디스크에 저장하고, PC상에서 표준 복호화 프로그램으로 오프라인(off-line) 복호화하여 음질을 평가하였다.

표 1. Slave 보드 처리 루틴의 수행시간

CPU	서브 루틴	수행 시간	
		(msec)	load
a	서브밴드 분석	6.22	25.9%
	스케일 팩터 코딩	1.81	7.5%
	FFT & 파워 스펙트럼	3.65	15.2%
	CPU - a 합계	11.68	48.7%
b	심리 음향 모델	18.28	76.2%
Slave 전체 수행 시간		29.96	124.9%

표 2. Master 보드 처리 루틴의 수행시간

CPU	서브 루틴	2채널, 192Kbps		5채널, 640Kbps	
		msec	load	msec	load
a	양자화	1.33	5.5%	3.42	14.3%
	비트열 포맷	2.38	9.9%	6.21	25.9%
	비트열 전송	2.31	9.6%	7.52	31.3%
	CPU-a 합계	6.02	25.0%	17.15	71.5%
b	비트 할당	7.03	29.3%	17.52	73.0%
Master 전체 수행시간		13.05	54.3%	34.67	144.5%

소프트웨어의 성능 개선 및 프로그램 최적화를 통해 향상된 Slave 보드 처리 루틴들의 수행시간을 표 1에 나타내었다. 측정된 수행시간은 시스템 클럭 주파수가 33MHz일 때를 기준으로 하였다. 이때, 'load'는 표본화 주파수가 48KHz일 때, 한 프레임(1152샘플) 구간의 길이인 24msec에 대한 상대적인 비율이다. 수행시간이 일정치 않은 서브 루틴들에 대해서는 가장 길 때의 시간으로 표시했다. Lee's 고속 DCT를 이용하여 분석 행렬을 대체한 서브 밴드 분석과정은 기존 수행시간 10.86msec의 57.3%인 6.22msec만에 처리되어 가장 개선된 수행 성능을 보였다. 이때, Slave-a에서 실행되는 서브루틴들은 처리되는 동안에 계속해서 입력 샘플을 받기 위한 인터럽트 루틴이 수행되기 때문에 실제 수행 시간은 측정된 것보다 빠르다고 말할 수 있다. Slave 보드에서 처리되는 루틴의 전체 'load'의 합이 약 125%이므로 시스템 클럭주파수가 조금 높아지면 한 개의 CPU로 처리할 수 있게 될 것이다.

Master에서 수행되는 루틴은 채널수와 비트율에 따라 처리해야하는 데이터 양이 달라지므로 수행시간도 각기 다르다. 채널 증가에 따라 가장 많은 수행시간이 걸리는 비트 할당 과정의 경우 사전할당을 통해 실시간 처리가 가능하도록 수행시간을 조절할 수 있다. 2채널까지의 처리는 비트율에 상관없이 사전할당을 사용하지 않고도 실시간 처리가 가능하여 기존의 MPEG 비트 할당 알고리즘을 사용하였고, 3채널 이상부터는 비트율에 따라 적절히 사전할당을 하는 방법을 택하였다. 2채널 192Kbps와 5채널 640Kbps인 경우에 대해 Master에서 수행되는 루틴들의 최적화된 수행시간은 표 2와 같다.

그림 4에서 첫 번째 샘플이 입력되어 최종 비트열이 전송될 때까지의 시간인 부호화 시간 지연(coding delay)은 5채널 640Kbps일 때 94.16msec이다.

그림 5는 44.1KHz 표본화 주파수에서 같은 입력

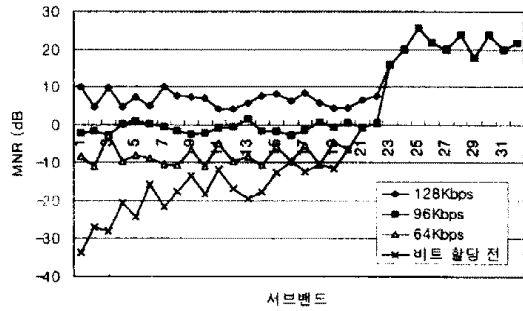


그림 5. 비트율에 따른 최종 MNR 곡선

샘플에 대해 구한 SMR 곡선을 바탕으로 채널당 64, 96, 128 Kbps로 비트 할당을 했을 때 최종적인 MNR 곡선을 나타내고 있다. 비트율이 상승함에 따라, 비트 할당된 대역에서 대략 6-10dB 정도의 MNR 상승이 이뤄지고 있음을 볼 수 있다. 128Kbps의 비트율에서 전 대역의 MNR이 5-10dB 이상에 분포하므로 이론적으로 모든 서브밴드의 양자화 잡음이 신호에 의해 마스킹되고 있다고 말할 수 있다. 실제 음질을 비교하였을 때, 채널당 128Kbps의 비트율인 5채널 640Kbps에서 원음과 같은 복원음을 들을 수 있었다. 비트 할당에 따른 최종 MNR 곡선을 비교함으로써 부호화된 비트열의 비트율에 따른 음질 차이를 설명할 수 있다.

V. 결론

본 논문에서는 범용 DSP 프로세서를 이용하여 구현된 MPEG-2 계층 2 오디오 부호화기의 성능을 개선하였다. 5개의 Slave 보드와 1개의 Master 보드로 구성된 부호화기는 최대 5.1 채널을 지원한다. 부호화된 비트열은 PC와의 인터페이스를 통해 하드 디스크에 저장할 수 있어 부호화기의 성능을 쉽게 평가할 수 있도록 하였고 실시간 복호화기에 직접 연결도 가능하도록 설계하였다. 구현된 부호화기는 MPEG-2 계층 2의 48KHz 표본화율에서 5.1채널, 640Kbps까지 실시간 처리되었으며, 따라서 그 이하의 모든 비트율과 표본화율, 채널 조합이 가능하다.

소프트웨어의 고속화 및 최적화는 2개의 CPU로 처리하던 Slave 보드 수행 루틴을 50MHz로 Up-clocking시 한 개의 CPU만으로 처리가 가능하게 하였다. 또한 4개의 CPU가 있어야 실시간 처리가 가능한 Master 보드의 다채널 처리도 고속 알고리즘을 적용하여 2개의 CPU만으로 처리할 수 있게 하였다. 사용되는 CPU의 숫자가 줄면 제작 비용상의 이득뿐만 아니라 HDTV 방송등의 응용시에 중요한

문제인 부호화 시간 지연(coding delay)도 줄일 수 있다. 최적화된 프로그램은 범용 DSP 프로세서를 이용하여 개발된 것이므로 다른 프로세서나 ASIC 등의 응용에 활용할 수 있다.

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.703 "Generic Coding of Moving Pictures and Associated Audio - CD 13818-3(Part3. MPEG-Audio)", Mar., 1994.
- [3] ISO/IEC JTC1/SC29/WG11 No.1519 "Generic Coding of Moving Pictures and Audio - CD 13818-3(Part3. MPEG-Audio)", 2nd Edition, Feb., 1997.
- [4] E. Zwicker, *Psychoacoustics*. Springer-Verlag, New York, 1982.
- [5] Texas Instrument, *TMS320C30 User's Guide*, 1992.
- [6] 김성윤, "범용 DSP를 이용한 MPEG-2 부호화기의 실시간 구현", 연세대학교 석사 학위 논문, Dec., 1995.
- [7] K. Konstantinides, "Fast Subband Filtering in MPEG AUDIO Coding.", *IEEE Signal Processing Letters*, vol. 1, no. 2, pp. 26-28, Feb., 1994.
- [8] 오현오, *et al.*, "범용 DSP칩에서의 구현을 위한 고속 DCT 알고리즘의 성능 비교 분석", *제10회 신호처리합동학술대회 논문집*, vol. 10, no. 1, pp. 375-378, Sep., 1997.