

음질향상 알고리즘을 내장한 MPEG-1 오디오 디코더의 Embedded OS 플랫폼에의 구현

홍성민[†], 박규식^{**}

요 약

본 논문에서는 음질향상 알고리즘을 내장한 MPEG-1 오디오 디코더를 Embedded OS(Linux)기반 플랫폼에 구현하였다. 네트워크 대역폭과 저장 공간의 제한점을 고려한 MP3, AAC, OGG 등과 같은 손실 오디오 압축 코덱들은 공통적으로 고주파 대역의 정보가 손실되는데, 이러한 고주파 신호성분 손실은 결국 표준 CD음질을 가지는 오디오 신호보다 제한된 저주파 대역만을 재생할 수 있게 된다. 본 논문은 손실된 고주파 대역의 신호성분과 하모닉 성분을 효과적으로 추정 및 복원할 수 있는 음질 향상 알고리즘을 MPEG-1 오디오 디코더 내부에 각 계층 I, II, III 특성별로 최적화하여 Embedded Linux 플랫폼으로 구현하였다. 기존의 MPEG-1 오디오 디코더와의 비교 청취 실험을 통해 본 논문에서 구현된 시스템이 신호 스펙트럼 및 음질 면에서 향상되었음을 확인할 수 있었다.

An Implementation of Sound Enhanced MPEG-1 Audio Decoder on Embedded OS Platform

Sung-Min Hong[†], Kyu-Sik Park^{**}

ABSTRACT

In this paper, we implement a sound-enhanced MPEG-1 audio decoder on embedded OS Platform. Low bit rate lossy audio codecs such as MP3, OGG, and AAC for mitigating the problems in storage space and network bandwidth suffer a major common problem such as a loss of high frequency fidelity of audio signal. This high frequency loss will reproduce only a band-limited low-frequency part of audio in the standard CD-quality audio. In order to overcome this problem, we embedded a sound enhancement algorithm into the MPEG-1 audio decoder and then the algorithms optimized according to the characteristic of the MPEG-1 audio layer I, II, III were implemented on an embedded OS platform. From the experimental results with spectrum analysis and listening test, we confirm the superiority of the proposed system compared to the standard MPEG-1 audio decoder.

Key words: MP3 Audio(MP3 오디오), Sound Enhanced(음질향상), Embedded OS(임베디드OS), MPEG-1 Decoder(MPEG-1 디코더)

1. 서 론

오늘날 휴대용 멀티미디어 시스템이나 모바일 기

기에서 MP3 재생 기능은 필수적인 기본사양으로 자리 잡아가고 있다. 유·무선으로 광범위하게 확대되어가는 네트워크를 통한 최근의 인터넷 패러다임은

※ 교신저자(Corresponding Author): 박규식, 주소: 경기도 용인시 수지구 죽전동 산44-1 단국대학교(448-701)
전화: 031)8005-3230, FAX: 031)8005-3228,
E-mail: kspark@dankook.ac.kr
접수일: 2007년 2월 20일, 완료일: 2007년 7월 6일

[†] 준회원, 단국대학교 정보컴퓨터학부
(E-mail: kilikan@dankook.ac.kr)

^{**} 단국대학교 정보컴퓨터학부

※ 본 연구는 2005학년도 단국대학교 대학연구비의 지원으로 연구되었음.

거대한 음악 산업에 획기적인 구조 변화를 가져오게 하였으며, 인터넷 사용자들은 보다 쉬운 방법과 적은 비용으로 오디오를 다운로드 받고 공유할 수 있게 되었다. 그러나 표준 CD 음질의 오디오 신호를 인터넷 상에 유통하거나 또는 모바일 기기에서 처리하기 위해서는 44.1 kHz 샘플링 비율의 큰 용량을 가지는 디지털 데이터를 처리할 수 있어야 한다. 이러한 문제는 일반적으로 네트워크 대역폭과 저장 공간의 축소를 위해 오디오 신호의 샘플링 비율을 낮추거나 MP3, OGG, AAC와 같은 오디오 데이터 압축 기법을 이용하여 해결할 수 있지만, 공통적으로 고주파수 (high frequency) 대역의 정보가 손실된다는 문제가 발생한다. 예를 들어, MP3(MPEG-1 오디오 계층 III)로 압축된 오디오 신호는 일반적으로 128 kbps 비트율 기준으로 자체 프로토콜에 의하여 16 kHz까지의 대역-제한된 오디오 신호만을 재생할 수 있는데 이는 낮은 샘플링 비율(low sampling rate)의 오디오 신호로 해석될 수 있다. 낮은 샘플링 비율의 오디오 신호는 표준 CD 음질의 오디오 신호보다 제한된 저주파 대역만을 재생할 수 있게 된다. 오디오 신호의 저주파 성분은 높은 에너지를 형성하고 있으므로 고주파 성분보다 중요하게 처리된다. 하지만 고주파 성분은 오디오 신호에 포함된 위치정보와 명료도, 재생 환경 등의 다양한 정보를 제공하고 있다. 만약 MP3로 디코딩된 오디오 신호에서 손실된 고주파 대역의 정보를 효과적으로 추정 및 복원할 수 있다면 표준 CD 음질의 오디오 신호에 존재하는 풍부한 오디오 정보를 제공할 수 있을 것이다.

기존의 압축코덱에 의해 손실된 오디오 신호의 손실 대역 추정 및 복원 연구는 인코더(encoder) 정보를 이용하는지 여부에 따라 크게 Non-blind와 Blind 방법 2가지로 구분할 수 있다. Non-blind 방법은 인코더를 이용해 압축 전 원 오디오 신호의 고주파 정보를 추출, 저장, 전송하고 전송된 정보를 이용해 디코더(decoder)에서 고주파 영역의 오디오 정보를 복원하는 방법이다. 이러한 방법으로는 독일의 Coding Technology사에서 개발한 SBR(Spectral Band Replication)[1]과 핀란드 VLSI 사의 PLUS-V[2]라는 알고리즘이 대표적으로 사용되고 있다. SBR 알고리즘은 2002년 MPEG-4 AAC 오디오 표준(HE AAC)으로 채택되었으며, 또한 기존 MP3 알고리즘에 탑재해 MP3Pro[3]라 불리는 오디오 코덱으로 개

발되기도 하였다. 한편, 2001년 제안된 PLUS-V는 인코더에서 원 오디오 신호의 고주파 대역에 해당하는 하모닉 정보를 추출해 전송하고, 디코더에서는 이를 이용하여 고주파 대역을 추정하는 방법을 사용하고 있다. 이러한 Non-blind 방식은 공통적으로 인코더를 사용해 원 오디오 신호의 고주파 정보를 추출, 전송하고 디코더에서 고주파 신호를 복원함으로써 비교적 정확한 고주파 신호 복원이 가능하나, 고주파 정보 추출을 위한 별도의 인코더와 고주파 정보 전송을 위한 별도의 추가 비트, 그리고 연산량이 높다는 단점이 있다.

한편 고주파 대역 보상 알고리즘으로 별도의 인코더 없이도 전송된 압축 신호만을 이용해 고주파 신호 성분을 추정할 수 있는 대표적인 Blind 방식으로는 HFR(High Frequency Regeneration)[4] 방식이 있다. 2003년 Liu, et al. 이 제안한 HFR 알고리즘은 압축된 오디오 신호의 저주파 성분을 지수 함수로 모델링하여 고주파 성분을 추정하는 방법이다. 이러한 Blind 방식의 알고리즘은 별도의 인코더를 필요로 하지 않고 오직 디코딩된 오디오 신호만을 이용해 고주파를 복원할 수 있다는 점과 연산량이 Non-blind 방식에 비해 비교적 낮다는 장점을 가지고 있다. 하지만 재생 오디오 음질이 Non-blind 방식에 비하여 떨어지고, 실시간 시스템으로 구축하기에는 여전히 높은 연산량을 가지고 있다.

본 연구에서는 기존 Blind 방식 알고리즘의 문제점인 높은 연산량을 개선한 MPEG-1 오디오 음질향상 알고리즘[5]을 MPEG-1 오디오 디코더 자체에 내장해 향상된 음질의 MPEG-1 오디오 디코더를 리눅스 기반의 Embedded 플랫폼에서 구현하였다.

본 논문은 다음과 같이 구성된다. 먼저 II장에서는 본 논문에서 제안한 음질향상 알고리즘과 이 알고리즘이 내장된 MPEG-1 계층 I, II 와 계층 III 디코더 시스템에 대한 각각의 구현 방법을 소개한다. III장에서는 II장의 각 계층별 제안 시스템을 구현한 Embedded Linux 플랫폼과 본 연구의 실험결과를 논하고, 마지막으로 IV장에서는 결론으로 끝을 맺는다.

2. MPEG-1 오디오 디코더 계층 I, II, III별 음질향상 알고리즘

본 연구에서 제안한 시스템은 오디오 인코딩 과정

에서 손실된 고주파수 신호 성분을 복원하기 위한 음질 향상 알고리즘을 MPEG-1의 각 계층 I, II, III별 특성에 맞추어 디코더에 내장하여 구현한 것으로, 다음과 같은 특징을 갖는다.

첫째, 시스템 복잡도를 낮추면서 범용적으로 적용 가능한 Blind 방식의 고주파 복원 알고리즘을 채택하여 MPEG-1 오디오 인코더에서 부가적인 정보 추출 없이도 동작 가능하도록 디코더 내부에 적용하였다.

둘째, MPEG-1 디코더에 각 계층 I, II, III별 최적화된 음질향상 알고리즘을 내장함으로써, 각 계층별 음질 향상 알고리즘에 소모되는 연산량을 최소화하였다. MPEG-1 오디오 계층 I, II에서는 서브밴드(Subband) 대역별 신호의 에너지를 계산하여 손실된 고주파 성분의 에너지 포락선을 추정하였으며, 복원된 고주파 성분에 에너지 포락선 스케일 팩터(scalefactor)를 적용하였다. 반면, 계층 III에서는 576개 스펙트럼 데이터 정보로부터 각 대역별 에너지를 계산하였으며 이를 이용하여 손실된 고주파 대역의 에너지 포락선을 추정하였다. 계층 I, II와는 다르게 고주파 대역 성분 보상시 정현파 모델링[6]을 이용해 스펙트럼의 국부 최대값(local maximum)을 추정해서 신호의 하모닉 성분을 보상하였고, 최종적으로는 대역별 이득(gain)을 조절하여 각 오디오 유형별로 음색을 조절할 수 있게 하였다.

본 연구에서 적용된 MPEG 오디오 음질향상 알고리즘은 기존의 Liu, et al[4]가 제안한 HFR 알고리즘을 개선 발전시킨 것으로 다음과 같은 차이점을 갖는다.

첫째, 대역별 에너지 계산을 이용한 유동 차단 주파수 계산함으로써 MPEG-1 인코딩 데이터율에 관련 없이 유동적으로 고주파 성분을 복원할 수 있는 장점이 있다.

둘째, 위의 유동 차단 주파수 계산 방법을 이용해 실제 차단-주파수와 복원되는 고주파 신호 간에 발생할 수 있는 스펙트럼 갭(spectral gap)을 극복할 수 있는 장점이 있다.

셋째, 고주파 신호 복원시 정확한 하모닉(harmonic) 주파수 성분을 보상해 재생 음에서 발생할 수 있는 반향 왜곡 음을 완화시킬 수 있다.

또한 제안 알고리즘은 기존의 알고리즘에 비해 연산량을 약 2~3배 정도로 감소시켜 실시간 구현에 적합한 복잡도를 가지고 있다. 다음 절들에서는 본

논문에서 제안한 음질향상 알고리즘과 각 MPEG-1 계층 I, II, III 별로 최적화된 알고리즘 구조에 대해 상세히 설명하였다.

2.1 고주파 음질 향상 알고리즘

본 연구에서 적용된 알고리즘은 오디오 신호의 대역별 에너지로부터 지수 함수 포락선(Envelop)을 추정하여 고주파 스펙트럼 및 피치 하모닉 성분(계층 III에만 사용)을 복원한다. 지수 함수를 이용한 에너지 포락선 추정은 최소 자승(least square) 접근법을 사용하였다.

2.1.1 에너지 계산

본 시스템에 적용된 알고리즘은 기본적으로 32개의 대역별 에너지를 활용하는데, 이 경우 한 대역별 주파수 폭은 약 689 Hz로 설정된다. 여기서 대역별 신호 성분이 나타나는 약 23번째 밴드(16 kHz)까지 일반적인 MPEG-1 오디오의 신호가 존재하는 영역이며 인코딩 설정사항에 따라 16 kHz 이상의 대역에도 에너지가 존재하는 경우도 있다.

2.1.2 대역별 에너지를 이용한 유동 차단 주파수 계산

Liu[4]의 HFR 알고리즘은 입력 신호의 특성에 상관없이 차단(cut-off) 주파수를 16 kHz로 가정하여 16 kHz~17 kHz까지 1 kHz의 고주파 신호만을 복원하였다. 그러나 MPEG-1 오디오 코덱으로 압축/복원된 결과 신호는 항상 16 kHz 차단 주파수만을 갖지 않는 경우도 있어, 만약 실제 입력 신호의 차단 주파수가 16 kHz 이하라면 HFR로 복원된 결과 신호는 스펙트럼 상에 깊은 갭(spectrum gap)을 초래하는 현상이 나타난다. 본 연구에서는 이러한 문제점을 해결하기 위하여 각 밴드별 에너지를 구하여 차단 주파수를 유동적으로 찾아 차단 주파수 부분에서의 스펙트럼 갭 현상을 효과적으로 제어하였다. 그림 1에 보는 바와 같이 16 kHz 부근의 대역별 에너지 비율을 조사하여 얻어진 임계값을 활용해 차단 주파수를 찾아 고주파 복원 지점으로 정하였다.

2.1.3 지수 함수 에너지 포락선 계산

일반적으로 오디오 신호의 각 밴드별 평균 에너지 포락선(energy envelope)을 검출해보면 지수 함수적으로 감소하는 현상을 발견할 수 있다. 이러한 스펙

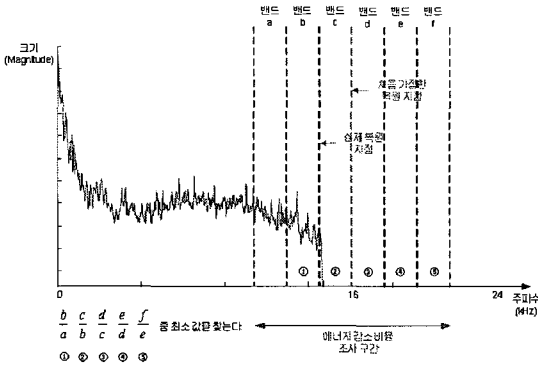


그림 1. 대역별 에너지를 이용한 유동 차단 주파수 결정

트럼 포락선은 원 오디오 신호의 음색과 관련된 부분으로 음질 왜곡을 방지하고 부드러운 음질 재생을 위해서는 필수적이다. 본 연구에서는 그림 1에서 구한 차단 주파수 이후의 밴드 에너지 포락선을 최소화승(least square) 접근법을 이용하여 추정하였다. 최소화승 접근법은 에너지 포락선의 근사 함수를 $y = be^{ax}$ 같이 지수 함수 형태로 가정하고 지수 함수를 선형 함수로 변형한 후 최소 자승 접근법에 의해 상수 A, B 를 다음과 같이 계산할 수 있다[4].

$$A = \frac{m(\sum_{i=1}^m x_i y_i) - (\sum_{i=1}^m x_i)(\sum_{i=1}^m y_i)}{m(\sum_{i=1}^m x_i^2) - (\sum_{i=1}^m x_i)^2}$$

$$B = \frac{(\sum_{i=1}^m x_i^2)(\sum_{i=1}^m y_i) - (\sum_{i=1}^m x_i y_i)(\sum_{i=1}^m x_i)}{m(\sum_{i=1}^m x_i^2) - (\sum_{i=1}^m x_i)^2}$$

여기서 x_i 와 y_i 는 각 대역별 인덱스와 에너지를 나타내며, 지수 함수 계수 a, b 는 $a = A, b = e^B$ 변환을 통해 구할 수 있으며 고주파 성분 복원시 스케일 팩터(scalefactor)로 사용된다.

2.1.4 고주파 피치 하모닉 추정(계층 III에만 적용)

고주파 대역을 보상할 때 하모닉 성분을 고려하지 않고 단순히 저주파 대역에서 구한 에너지 포락선 스케일 팩터만을 곱하여 고주파 대역 성분을 복원해 준다면 잘못된 하모닉 성분의 추가로 인해 반향음(reverberation)이 생성될 수 있기 때문에 정확한 하모닉 성분의 추정 및 보상이 요구된다. 본 연구에서는 정현파 모델링 기법[5]을 이용하여 스펙트럼에서

국부 최대값(local maximum)을 계산한 후 각 피크 성분들 간의 주파수 간격을 히스토그램으로 표현하고, 히스토그램의 크기가 가장 큰 주파수를 각 프레임에서의 하모닉으로 선정하였다.

2.1.5 고주파 성분 복원 및 피치 하모닉 보상

고주파 정보 복원은 2.1.3 절에서 구해진 에너지 포락선 함수를 이용하여 고주파 대역 구간에 저주파 대역의 스펙트럼을 복사하게 되며 이때 복원되는 대역별 에너지의 크기는 에너지 포락선 스케일 팩터에 따르게 된다. 고주파 복원된 신호는 최종적으로 하모닉 추정치를 보상하게 되는데, 만약 추정하고자 하는 대역에 하모닉이 포함되어 있지 않다면 각 주파수 성분들에 동일한 에너지 포락선 스케일 팩터를 곱해 주고, 하모닉이 포함되어 있을 경우는 하모닉을 제외한 성분들의 에너지 합을 전체 추정 에너지의 90%가 되도록 스케일 팩터를 곱해준 다음 하모닉 주파수에는 잔여 에너지를 할당해 주었다. 다음의 그림 2는 각각 추정된 에너지 포락선과 이에 따른 하모닉 보상과 고주파 대역 복원 과정을 설명하고 있다.

2.2 MPEG-1 오디오 계층 I, II, III 디코더별 음질향상 알고리즘 구현

2.2.1 MPEG-1 오디오 계층 I, II 디코더

다음의 그림 3과 4는 MPEG-1 오디오 계층 I, II 디코더에 구현한 제안 음질향상 시스템을 보여주고 있다.

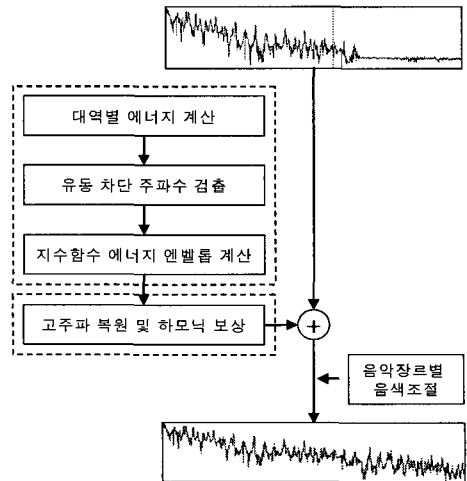


그림 2. 디코더 내부에서의 고주파 대역 복원 과정

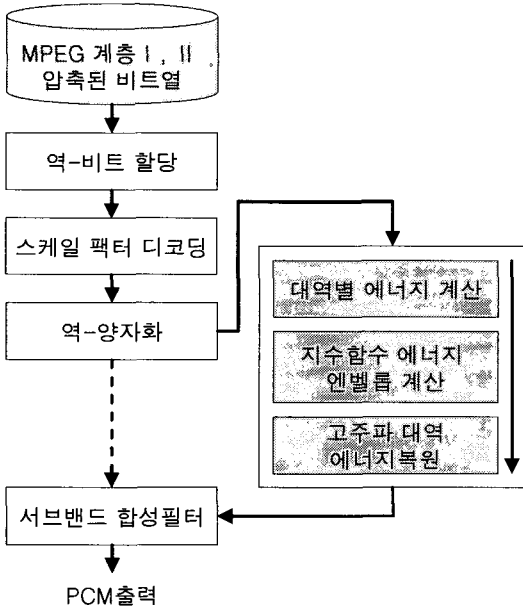


그림 3. MPEG 오디오 계층 I, II 디코더에 내장된 음질 향상 알고리즘

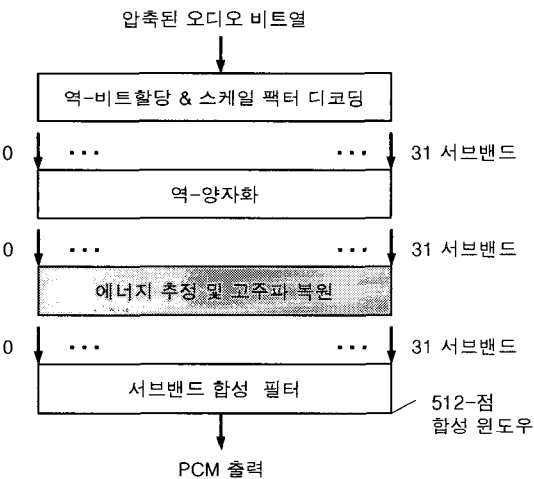


그림 4. MPEG-1 오디오 계층 I, II 디코더에 구현된 음질 향상 알고리즘

MPEG-1 오디오 계층 I, II 는 32개 서브밴드 대역을 갖는 서브밴드 분석-합성 필터를 기반으로 동작되기 때문에 본 연구에서 제안된 음질향상 알고리즘은 디코더의 서브밴드 합성필터 단계 바로 전에 적용하게 된다. 손실된 고주파 대역에서의 스펙트럼 에너지 포락선은 디코더 내의 역-양자화(de-quantization) 과정 후 각 대역별 신호의 에너지 평균값을 이

용해 추정하였고 이때 차단 주파수 지점도 같이 결정된다. 손실된 고주파 성분은 추정된 에너지 포락선의 스케일 팩터 값에 근거해 저주파 성분을 고주파 대역에 복사해 복원하였으며 최종적으로는 서브밴드 합성필터를 통과시켜 음질 향상된 오디오 신호를 재생한다. MPEG-1 계층 I, II 비트열(bitstream)은 오직 32개 서브밴드 대역별 샘플 값을 포함하고 있기 때문에 스펙트럼 영역의 하모닉 추정을 위해서는 별도의 추가적인 연산과정이 필요하다. 이러한 이유로 본 논문에서는 MPEG-1 계층 I, II에 2.1.4절의 고주파 영역 하모닉 보상을 적용하지 않고 있으며 다음 절에 설명하는 계층 III에만 이를 적용하였다.

2.2.2 MPEG-1 오디오 계층 III 디코더

MPEG-1 계층 III는 계층 I, II에 비해 그림 5과 같이 MDCT(Modified DCT) 과정과 허프만 코딩(Huffman coding) 그리고 과도신호(transient signal) 처리 과정이 추가된 시스템이다. 계층 III에서는 MDCT와 허프만 코딩으로 인해 연산량은 상당히 증가 되었으나 기존의 계층 I, II 보다 향상된 음질을 갖는다.

계층 III는 계층 I, II와 달리 심리음향 모델 2(Psychoacoustic Model 2)가 적용되고 정상신호(stationary signal)와 과도신호의 시간-주파수 해상도를 동적으로 조절하기 위해 MDCT(Modified

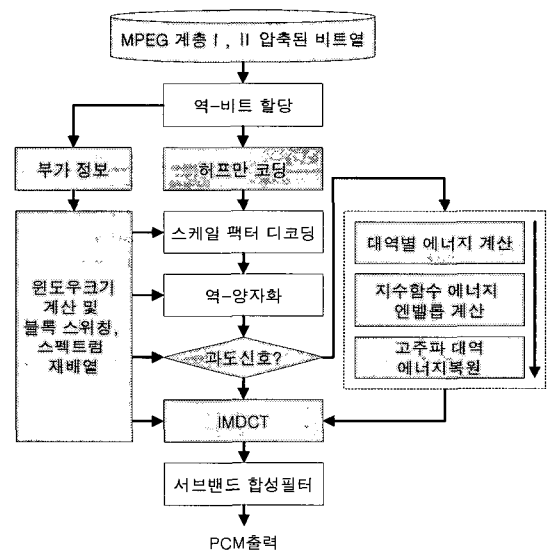


그림 5. MPEG 오디오 계층 III 디코더에 내장된 음질 향상 알고리즘

DCT) 블록 변환을 위한 윈도우 길이를 가변적으로 처리할 수 있다. 여기서 과도신호는 신호가 급격히 변하는 부분을 말하는 것으로 이 경우 정상신호에 사용되는 길이가 긴 윈도우(주파수 해상도가 높음)를 사용하게 되면 소위 프리-에코(pre-echo)라는 인공잡음이 생성되며 이를 방지하기 위해서는 긴 윈도우를 짧은 윈도우(시간 해상도가 높음)로 바꿀 필요가 있다. MPEG-1 오디오 계층 III는 이러한 동적 시간-주파수 해상도 구현을 위해 긴(long) 윈도우, 짧은(short) 윈도우, 이들 윈도우 전이를 위한 스타트(start)/스톱(stop) 윈도우를 지원하고 있다. 본 연구에서는 긴 윈도우를 사용하는 정상신호 경우에만 제안된 음질향상 알고리즘을 적용하였다.

그림 6은 긴 윈도우에서 음질향상 알고리즘을 적용한 경우를 보여준다. MPEG-1 오디오 계층 III에서 역-양자화 과정을 거치게 되면 576개 MDCT 스펙트럼 값을 갖게 된다. 이 때 고주파 복원 대상인 16 kHz ~ 22.05 kHz 주파수 대역은 MDCT의 432번째부터 576까지의 스펙트럼 계수 영역에 해당된다. 576개 MDCT 스펙트럼 계수값은 18개 샘플 블록별로 32개 서브밴드 필터뱅크별 데이터를 표현하기 때문에 블록별로 에너지를 계산하고 이에 따른 에너지 포락선을 구하게 된다. MDCT 블록별 에너지 포락선을 구한 다음에는 대역별 에너지 비율을 이용해 차단 주파수를 찾아 고주파 대역의 손실된 스펙트럼 정보를

복원하게 된다. 이때 계층 III에는 음향반향 왜곡을 방지하기 위해 고주파 대역의 하모닉 성분을 추정하고 이를 복원 과정에 적용하게 된다. 이전에 설명한 바와 같이 보상된 고주파 대역에 하모닉 성분이 포함되어 있지 않다면 각 주파수 성분들에 동일한 에너지 포락선 스케일 팩터를 곱해주고, 하모닉이 포함되어 있을 경우는 하모닉을 제외한 성분들의 에너지 합을 전체 추정 에너지의 90%가 되도록 스케일 팩터를 적용하였다.

3. 시스템 환경 및 실험 결과

3.1 시스템 환경

본 연구에서는 음질향상 알고리즘이 내장된 MPEG-1 오디오 디코더를 Intel PXA255기반의 Huins사 ARM 보드에 Kernel 2.4.19 기반 Redhat OS 플랫폼에 구현하였다. 호스트 PC는 Cygwin과 Toolchain을 설치하여, 컴파일 환경을 조성하였다. 다음의 그림 7은 본 연구에서 구현한 ARM 플랫폼의 구성을 나타낸다. 또한, 제안된 알고리즘은 Xing Technology의 Xing MPEG 오디오 라이브러리를 사용하고 있는 MPEG-1 오디오 디코더 Ximp v0.1.15를 수정해 구현 하였다.

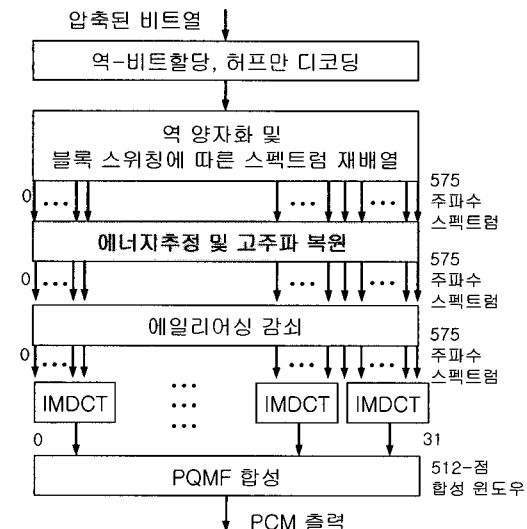


그림 6. MPEG-1 오디오 계층 III 디코더에 구현된 음질 향상 알고리즘

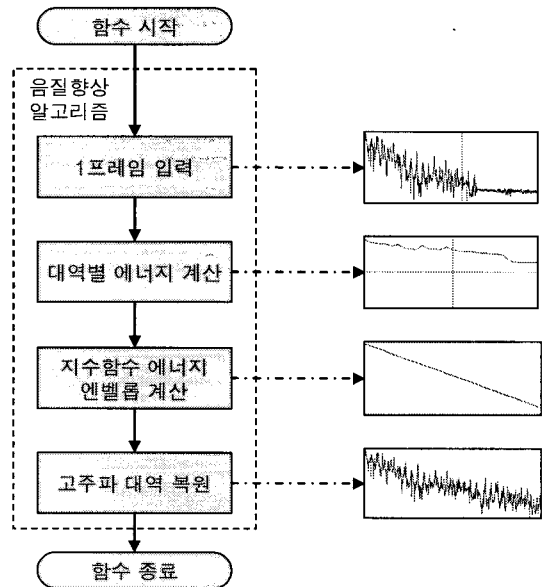


그림 7. Embedded OS상에 구현된 MPEG-1 오디오 음질 향상 시스템

3.2 실험 결과

본 연구에서 제안된 시스템의 성능 확인을 위하여 Rock, HipHop, Jazz, Classic 등의 각 장르별 다양한 오디오 신호를 MPEG-1 오디오의 각 계층 I, II, III로 압축해 그림 8의 제안된 디코더 시스템에 입력으로 사용하여 실험을 수행하였다. 모든 장르의 오디오 신호에 대해 제안된 디코더가 표준 MPEG-1 오디오로 디코딩된 신호보다 개선된 오디오 음질을 제공함을 확인할 수 있었으며 그림 8에서 보는 바와 같이 표준 MPEG-1 오디오의 자체 알고리즘으로 인해 손실된 고주파 영역을 충실히 재생해 내는 것을 확인하였다.

다음의 그림 9은 각각 Rock, HipHop, Classic, Jazz 음악에 대한 코딩 전 오디오 신호와 표준 MPEG-1 오디오 계층 III(계층 I, II도 유사한 결과를 보이기 때문에 생략함)로 디코딩된 결과 신호, 본 논문에서 제안한 시스템의 결과 신호 그리고 Liu의 HFR (High Frequency Regeneration)[4] 알고리즘의 결과 신호를 하나의 그래프에 나타내어 각 알고리즘의 성능을 비교 한 것이다.

그림 9에서 보듯이 본 논문에서 제안한 알고리즘이 Liu의 알고리즘보다 더 충실하게 고 주파수 대역의 신호를 추정 및 복원하고 있는 것을 알 수 있다(이

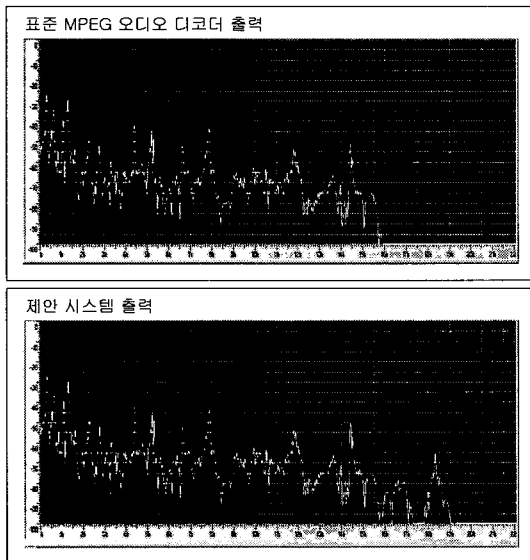
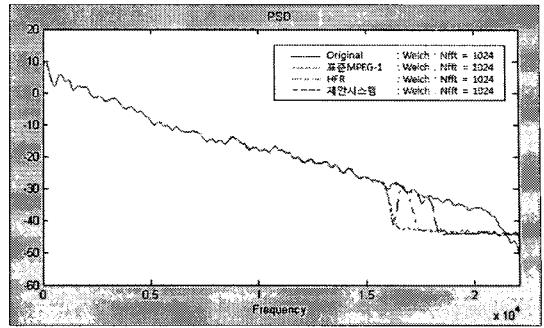
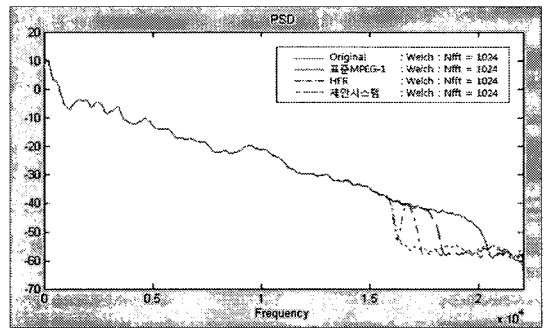


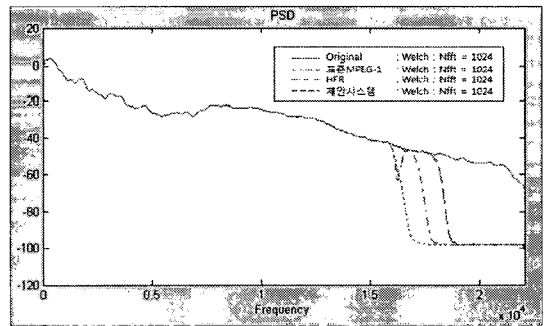
그림 8. 표준 MPEG-1으로 디코딩된 오디오신호와 제안 시스템으로 고주파성분이 복원된 오디오 신호 스펙트럼 비교



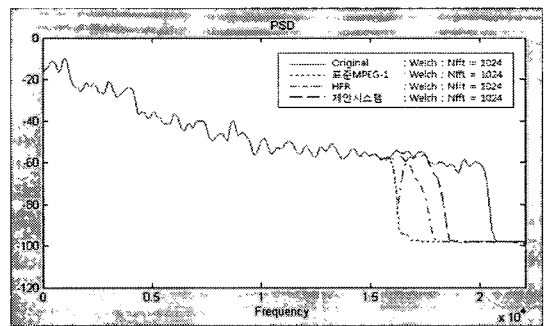
(a) Rock



(b) HipHop



(c) Jazz



(d) Classic

그림 9. 원 오디오신호, 표준 MPEG-1 오디오 계층 III로 디코딩된 결과 신호, 제안 시스템의 결과 신호, Liu의 HFR(4) 알고리즘의 결과 신호 비교

실험에서는 18 kHz까지 복원하고 있으나 최대 22 kHz까지 가능함). 또 하나 주목할 점은 디코딩된 신호의 차단 주파수인 16 kHz 근처에서 보듯이 기존의 Liu의 알고리즘은 추정 복원된 신호의 차단 주파수 근처에서 스펙트럼 갭을 보이는 반면 제안된 알고리즘은 신호의 대역별 에너지를 이용하여 차단 주파수를 유동적으로 찾기 때문에 스펙트럼 갭 현상을 효과적으로 제어할 수 있다는 점이다.

다음의 표 1은 한개 프레임 당 ARM 플랫폼에서 처리해야할 데이터 연산량을 비교한 것이다. 2절에서 설명한 바와 같이 MPEG-1 오디오 디코더 내부에 알고리즘이 구현되기 때문에 별도의 FFT 연산과정 등이 필요 없어 비교적 낮은 연산량으로도 구현이 가능하며, 표에서 보듯이 제안 알고리즘이 논문[4]의 방식에 비해 약 2~3 배 정도의 연산량을 절감할 수 있는 것을 알 수 있다.

표 1. ARM 플랫폼에서의 1 프레임당 연산량 비교

항 목	제안 알고리즘	논문[4]
CPU	PXA255	PXA255
샘플링 비율	44.1khz	44.1khz
샘플당 비트수	16 bits	16 bits
알고리즘에 의한 지연시간	1.04ms	1.92ms
뎁스(MIPS)	0.5	0.92
메모리(KB)	18	75

4. 결 론

본 연구에서는 MPEG-1 오디오 신호에서 손실된 고주파 신호를 효율적으로 추정 및 복원할 수 있는 음질향상 알고리즘을 내장한 MPEG-1 오디오 디코더를 실시간으로 ARM Embedded OS 플랫폼에 구현하였다. 오늘날의 다양한 휴대용 멀티미디어 시스템이나 모바일 기기에서 MPEG-1 오디오 계층 3 디코더 기능은 기본적인 사양으로 적용되고 있다. 본 연구에서 제안된 시스템은 표준 MPEG-1 오디오 디코더에 내장되어 실시간으로 구현될 수 있는 시스템으로 상용화된 시스템에 별도의 부가장치 없이도 바로 적용할 수 있다는 장점이 있다. 본 연구팀은 향후 본

제안 시스템의 상용화된 모바일 기기 탑재를 위하여 다양한 음질 실험을 진행하고 있다.

참 고 문 헌

- [1] Martin Dietz, Lars Liljeryd, Kristofer Kjörling, and Oliver Kunz, "Spectral Band Replication, a Novel Approach in Audio Coding," *112th AES Convention*, Germany, April 2002.
- [2] "PLUS-V Specification : Public Document," VLSI Solution, Finland, Dec 17, 2001.
- [3] Thomas Ziegler, Andreas Ehret, Per Ekstrand, and Manfred Lutzky, "Enhancing mp3 with SBR: Features and Capabilities of the new mp3PRO," *112th AES Convention*, May 10-13, 2002.
- [4] Chi-Min Liu, Wen-Chieh Lee, and Han-Wen Hsu, "High frequency reconstruction for band-limited audio signals," *Proc. of 6th Conf. on Digital audio Effects*, London, UK, Sep. 8-11, 2003.
- [5] Sang-heon Oh, Won-Jung Yoon, Youn-ho Cho, Kyu-Sik Park, and Ki-Man Kim, "A New Spectral Enhancement Algorithm In MP3 audio," *IEEE Transactions on Consumer Electronics*, Vol. 52, No. 1, pp. 196-199, 2006.
- [6] R.J. McAulay and T.F. Quatieri, "Speech Analysis-Synthesis Based on a Sinusoidal Representation," *IEEE Transactions on ASSP*, Vol. 34, No. 4, pp. 744-754, 1986.



홍 성 민

2000년 7월 상명대학교 졸업
 2004년 7월 대한민국 해군 정보통신 병과 중위 계대
 2007년 3월 단국대학교 정보컴퓨터과학과 졸업(이학석사)



박 규 식

1986년 (미) Polytechnic University 졸업

1988년 (미) Polytechnic University 졸업(공학석사)

1994년 (미) Polytechnic University 졸업(공학박사)

2001년~현재 단국대학교 정보컴

퓨터학부 부교수

관심분야 : 음성 및 오디오 신호처리, DSP 구현 및 설계, 감성인식, 멀티미디어 정보처리