

주 제

제로트리틀 이용한 스케일러블 오디오 부호화

창원대학교 석종원, 정영근, 한국전자통신연구원 장인선, 서정일, 장대영

차 례

I. 서론

II. 스케일러블 오디오 부호화 기술

III. EZW(Embedded Zerotree Wavelet) 오디오 부호화

IV. 결론

I. 서론

고품질의 서비스를 추구하는 사용자의 욕구를 충족시키기 위한 기술개발의 노력이 다양한 분야에서 끊임없이 시도되고 있다. 특히, 멀티미디어 기술의 진보에 따른 오디오 분야의 기술개발은 더욱 절실히 요구되고 있으나, 한정된 용량의 저장매체나 제한된 전송채널의 조건하에서의 고품질 오디오를 보장한다는 것은 매우 어렵다. 따라서, 효율적이고, 경제적인 방법을 이용하여 낮은 비트율로 고품질의 오디오를 제공하기 위한 신호 압축 기술이 필요하게 된다. 이 때문에 다양한 신호 압축 기술이 개발되어 사용되고 있으며, 다양한 압축 부호화 기술이 활발히 연구되고 있다.

오디오 기기에 있어서는 CD(Compact Disk), DAT(Digital Audio Tape)와 같은 디지털 오디오가 널리 사용되고 있었으나 방송, 통신 등과 같이 대역이 한정되어 있는 응용 분야에서 많은 데이터를 전송하

는 것은 커다란 문제였다. 15~20 kHz 정도의 광대역 오디오 신호를 압축하지 않고 PCM(Pulse Code Modulation) 방식으로 부호화 하려면 샘플당 16 비트 이상이 필요하며 HDTV의 경우 48KHz 샘플링 주파수에 5채널 오디오 신호를 표준으로 하고 있기 때문에 이를 전송하기 위해서는 3.84 Mbit/s 이상의 막대한 채널 용량이 필요하다. 따라서 전송과 저장에서의 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 80년대 후반부터 세계 각국의 여러 연구소에서는 CD 수준의 디지털 오디오 신호를 지각적인 음질을 떨어뜨리지 않고 압축하는 기술, 즉 지금까지의 아날로그 오디오 방송을 대체할 새로운 디지털 오디오 처리 기술을 개발하였다[1-7].

또한, 최근들어 MPEG 및 ITU-T 등에서는 전송 상황에 따라, 비트율을ダイナ믹하게 조절할 수 있는

계층적인 비트 스트림을 제공하는 bitrate scalable coding 기술 필요성 인식에 따라 이에 다른 표준화 작업을 준비 중에 있다[8]. MPEG에서 표준화가 진행 중인 Scalable Speech & Audio Coding 기술은 MPEG-4 GA에서 제공된 scalability가 32 kbps 이하의 비트율에서는 음질 열화가 심각하게 발생하는 단점과 음성에 대한 FGS(Fine Grain Scalability) 기능을 갖는 코덱이 없다는 단점을 보완하고, 단일구조의 scalable한 코덱을 개발하는 것을 목표로 한다. 즉, 네트워크를 통한 인터넷 방송 등에서는 오디오 스트림이 다양한 전송환경(cellular, WLAN, xDSL, FTTH, digital broadcasting, analogue modem, ISDN 등)에서 제공되게 되며 네트워크 상황에 따라 사용 대역폭이 가변적인 환경에서 제공되게 된다. 제공되는 오디오 콘텐츠의 QoS를 보장하기 위해서는 매우 낮은 비트율로부터 높은 비트율까지 단일 스트림 구조를 가지면서 스케일러블한 구조가 요구된다. 즉 높은 비트율에서는(>128kbps) Transparent audio quality를 제공하여야 하고, 낮은 비트율(<24kbps)에서도 오디오 신호를 효과적으로 제공하기 위한 기술이 요구된다. 이와 같은 가변 비트율 코덱 기술은 2개 이상의 같은 또는 서로 다른 부호화기를 연결하여 사용함으로써 scalability를 줄 수 있는 Large Step Scalability와 단일 코덱을 사용하여 scalability기능을 제공하는 FGS방식으로 나눌 수 있으며, MPEG-4 BSAC(Bit-Sliced Arithmetic Coding)이 FGS의 대표적인 예이다[7].

본 글에서는 단일 코덱을 사용하여 FGS 기능을 가지도록한 대표적인 방식인 BSAC 외에, zerotree를 이용하여 비트율을 가변적으로 조절할 수 있는 방식에 대하여 소개하고자 한다. zerotree는 원래 웨이블릿 영역에서 상위대역과 하위대역 사이의 각 계수가 가지는 특징을 이용하여 부호화하는 방식으로 영상 부호화에 많이 사용되어왔다. 최근 들어 이러한 방식

을 오디오에 적용하여 비트율을 가변적으로 조절하는 연구가 수행되어왔다.

II. 스케일러블 오디오 부호화 기술

스케일러블 부호와 방식이란 한정된 대역폭에서도 비트율을 가변적으로 조절하여 정보를 전달할 수 있는 부호화 방식이다. Bitrate scalability는 비트스트림의 일부분만을 가지고도 복호화가 가능하며 또한 의미 있는 신호를 만들어 낼 수 있는 방식을 말하는 것이다. 이러한 특징은 데이터 전송 시, 전송로가 전체 비트스트림을 전달하기 위해 필요한 대역폭의 제공을 보장하지 못하는 경우에 해결책으로 제시되어질 수 있다. MPEG-4에서는 이러한 scalability를 중요한 기능 중 하나로 정의하고 표준화상에서 정의하였고 large step scalable coding 방식과 fine grain scalable coding으로 구분하였다.

Large step scalability 방식은 하나의 base layer 코덱에 하나 이상의 같은 또는 서로 다른 코덱을 사용함으로써 scalability를 줄수있게 된다. 또한, 이방식은 매우 낮은 비트율(약 6kbps)을 지원하기 위해 core coder 를 사용할 수 있다. core coder 로는 음성 부호화방식인 MPEG-4 Narrow Band CELP 또는 TwinVQ를 사용한다.

반면 fine grain scalability는 MPEG-4 BSAC으로 대표되는 방식으로, 여러 코덱을 사용하여 scalability를 주는 방식이 아니라 단일 부호화기를 사용하여 scalability기능을 제공하는 방식이다. 이 기법은 양자화 후, 중요도에 따라 비트스트림을 구성하게 된다. 이 방식의 장점은 1kbps/ch 단위의 scalable layer를 제공한다는 것이다.

III. EZW(Embedded Zerotree Wavelet) 오디오 부호화

웨이블렛은 다해상도 신호 표현 및 분석 방법으로, 최근 들어 많은 분야에 응용이 되고 있으며, 특히 영상 및 오디오 신호 부호화에 많이 사용되고 있다. 한편 제로트리는 영상부호화에 처음 시도되었으며, 점진적 전송이 장점을 가지고 있다. 또한 웨이블렛 계수의 대역 간 상관관계를 이용하여 zerotree를 형성하고, 이를 중요성 지도(significance map)를 이용하여 부호화하는 기법이다.

1. 웨이블렛 변환

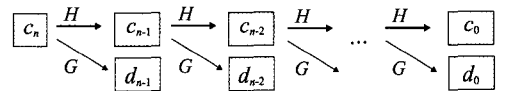
웨이블렛 변환은 1900년대 초, Haar에 의해서 처음 제안된 것으로 지금까지 신호 처리 분야에서 가장 널리 사용된 푸리에 변환(Fourier Transform)과 같은 직교 변환 방법으로 데이터나 함수등을 여러 다른 주파수 성분으로 나누어 각 스케일에 맞는 해상도를 가지고 나누어진 각 성분들을 해석하는데 사용한다. 웨이블렛 변환은 1980년대에 들어서면서부터 본격적인 연구가 진행되었으며 1990년 전까지 수학 분야에서 Morlet, Grossman등에 의해서 웨이블렛이 연구 되어 왔고 [9], 공학 분야에서는 Vetterli, Smith등에 의해서 부대역 부호화(Filter Banks, Subband Coding)가 연구되어 왔는데 [10], 1990년 후에 Mallat [11] 등에 의해서 이 두 분야가 하나의 이론이라는 것이 수학적으로 유도되어 웨이블렛으로 결합되었다. 웨이블렛 변환은 스케일과 시간이라는 두 변수로 표현되고 푸리에 변환과 같이 기저 함수들의 집합으로 표현되며 데이터와 함수들을 서로 다른 스케일 성분들로 분해하고 각 스케일에 해당하는 해상도와 연관된 각각의 성분들을 조사할 수 있고 공간에 대한 지역적 특성을 가지고 있어 전체 또는 지역적인 특

정도 분석 가능한 특성을 지닌다.

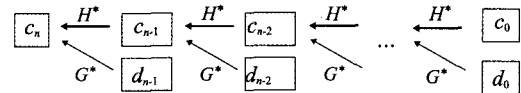
신호 처리에 있어서 사용하는 웨이블렛은 단순히 필터링한 것으로 생각할 수 있다. 즉, 고역 여파기와 저역 여파기를 지닌 두 개의 채널을 지나는 필터 뱅크(Filter Bank)를 통과하는 것으로 생각할 수 있다.

(그림 1)은 일반적인 1차원 신호의 웨이블렛 분해와 복원이 이루어지는 과정을 나타내고 있다. 여기서 G은 고역 여파기이고, H은 저역 여파기이다. 저역 여파기와 고역 여파기를 사용해서 두 개의 대역으로 분해된 신호를 downsampling하는 과정까지를 웨이블렛 분해라고 하며, 분해된 신호를 upsampling해서 다시 복원 필터를 사용하여 합치는 과정을 웨이블렛 복원이라고 한다.

한편, 웨이블렛 패킷은 웨이블렛 분해를 일반화한 것으로서, 좀더 풍부하게 신호해석을 할 수 있다. (그림 2)는 웨이블렛 패킷 분해방법을 나타낸 것으로서, 웨이블렛 분해가 저주파대역을 계속해서 분해하는데 비해, 웨이블렛 패킷은 저주파와 고주파 대역 모두를 분해하여 트리구조를 만드는 것을 나타내었다.

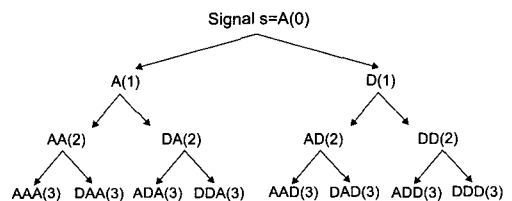


(a) 분해과정



(b) 복원과정

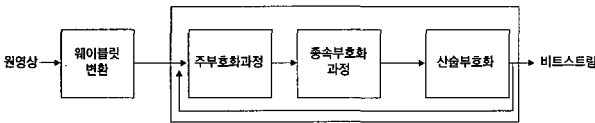
(그림 1) 웨이블렛 분해와 복원 과정



(그림 2) 웨이블렛 패킷 분해

2. EZW(Embedded Zerotree Wavelet)

EZW는 Shapiro[12]에 의해 제안된 것으로서, 웨이블릿 계수의 대역간 상관관계를 이용하여 제로트리(zero-tree)를 형성하고, 이를 중요성 지도(significance map)를 이용하여 부호화하는 기법이다. 중요성 지도는 지수 적으로 감소하는 문턱 치와의 비교를 통해 판단된 웨이블릿 계수의 중요성 여부를 나타낸 이진 비트평면(bit plane)이며, 중요성지도에서 비중요 계수로 판단된 계수들의 블록은 제로트리를 나타내는 심볼로 부호화된다. 입력계수는 네 개의 심볼(POS, NEG, IZ, ZTR)로 부호화되며, JPEG 알고리즘에서 사용되는 EOB(end-of-block)이 추가된다.



(그림 3) EZW 부호화 과정

EZW 알고리즘은 웨이블릿을 이용한 기존의 압축 방법에 비해 향상된 성능을 가지면서 점진적 전송에 유리하다. 현재, 이 알고리즘은 개선한 SPIHT(set partitioning in hierarchical tree) 알고리즘과, SFQ(space-frequency quantizer) 알고리즘 등에 이용되었다. 이 알고리즘들은 기존에 영상 부호화에 많이 이용되었으나, 최근에는 몇몇 논문에서 오디오 부호화에 응용하여 좋은 성능을 보였다[13-15].

EZW 부호화는 신호를 점진적으로 근사해 가는 방식이다. 즉, 부호화가 진행됨에 따라 신호 복원에 중요한 정보가 전송되어 신호가 조금씩 향상되는 방식이다. 이러한 부호화의 진행은 응용 분야에 따라 상이하지만 일반적으로 수신 측에서 요구하는 신호의

질을 만족시킬 때까지 계속된다.

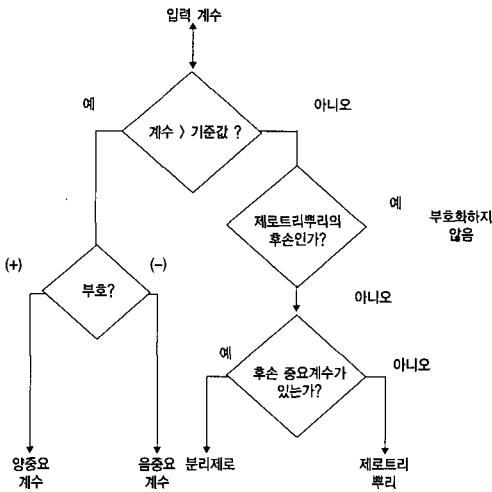
EZW 부호화는 (그림 3)에 보인 바와 같이 주부호화 과정(dominant pass), 종속 부호화 과정(subordinate pass)을 거치고 적응 산술 부호화(adaptive arithmetic coding)로 부호화된다. 이때 목표로 한 비트율에 이르지 못했을 경우 임계값을 낮추어 주부호화 과정부터 다시 시작한다.

일반적으로 영상신호 부호화에서는 주부호화 과정에서 웨이블릿 변환 계수 사이의 부모-자손 관계와 zerotree를 이용한다. 웨이블릿 변환 계수 사이에는 부모와 자손의 종속 관계가 있다.

각 웨이블릿 변환 계수는 하위 대역에서 같은 위치에 상응하는 계수와 부모-자손의 트리 구조를 이루고 있다. 일반적인 정지 영상에서는 부모 계수의 절대값의 크기가 특정 임계값보다 작다면 그의 자손 계수 또한 그 임계값보다 작을 확률이 높다. 이러한 사실을 만족하는 부모-자손의 트리를 zerotree라고 한다.

각 대역별로 웨이블릿 계수를 점진적으로 감소하는 임계값의 집합을 웨이블릿 변환 계수의 절대값과 비교하여 (그림 4)와 같은 방법으로 웨이블릿 변환 계수의 절대값이 임계값보다 크다면 그 계수는 중요 계수가 되고, 중요 계수는 계수의 부호가 양(+)이면 POS가 되고 음(-)이면 NEG가 된다. 중요 계수가 아닌 것은 후손 계수 중 하나라도 중요 계수가 있다면 IZ(Isolate Zero)가 되고 후손 계수가 모두 중요 계수가 아니라면 ZTR(ZeroTree Root)로 부호화된다.

종속 부호화 과정은 주부호화 과정이 끝난 후 중요 계수라고 판별된 계수에 대해서만 수행된다. 주부호화 과정에서 중요 계수라고 판단된 계수를 종속 부호화 리스트에 저장한다. 종속 부호화 리스트에 저장된 중요 계수는 주부호화 과정을 통해 이미 그 부호(sign)정보를 알고 있다. 종속 부호화 과정은 이미 알고 있는 부호 정보를 이용하여 원 계수의 값과 양자화



(그림 4) 계수의 중요도 판정 흐름도

된 값의 차값을 부호화 하는 과정으로 원 계수의 값의 크기가 복원된 계수의 크기보다 크면 1, 작으면 0을 부호화 한다.

1차 주부호화 과정과 종속 부호화를 거친 다음 산술 부호화를 거치게 되는데 목표로 했던 비트율에 이르면 부호화를 멈추지만 만약 그렇지 않다면 임계치의 값을 반씩 낮추어 가면서 주부호화 과정과 종속 부호화 과정을 반복해서 실행한다. 이러한 특징으로 수신 측에서 비트열을 전송받는 중 중간에 멈추어도 신호를 복원할 수 있게 된다.

EZW 부호화 방식은 중요 웨이블릿 변환 계수의 위치를 부호화하는데 매우 효율적이며 대상 신호의 웨이블릿 변환에 의해서 얻어진 에너지 집중 현상을 효과적으로 이용한 알고리즘으로 원하는 비트율까지 부호화 할 수 있을 뿐만 아니라 점진적인 전송을 가능하게 한다.

3. 심리음향모델

서브밴드 부호화 방식의 이점은 각 밴드의 양자화

잡음 수준을 신호의 특성에 따라, 그리고 청취자의 청각적 특성에 따라 다르게 할당할 수 있다는 데 있다. 심리 음향 모델을 이용하여 각 서브밴드에서 원음에 의해 마스킹되어 들을 수 없는 최대의 잡음 레벨을 결정할 수 있으며 이를 사용해서 각 서브밴드의 실제 양자화기를 조정하는 비트 할당을 함으로써 적은 비트를 가지고 주관적으로 양자화 잡음이 들리지 않도록 부호화할 수 있다.

MPEG에서는 두 가지의 심리음향 모델을 제공하는데 두 모델은 응용 분야에 따라 적절하게 선택되어 사용되어질 수 있다. 첫 번째 방법은 FFT 스펙트럼을 순음과 잡음 성분으로 나누어 각 성분에 의한 마스킹 임계값을 구한 후 절대 가청 한계(absolute threshold of hearing)를 고려하여 마스킹 임계값을 구한다. 계층 2에서는 일반적으로 심리음향 모델 1을 이용하여 마스킹 임계값을 구한다. 두 번째 방법은 FFT 스펙트럼을 청신경의 여기 모델인 스프레딩 함수와 콘볼루션하여 마스킹 임계값을 구하는 것으로 많은 계산량이 필요하지만 마스킹 특성을 보다 정확히 모델링할 수 있으므로 높은 압축율을 필요로 하는 응용분야에 이용한다. 이 절에서는 심리음향 모델 1에 비해 보다 좋은 성능을 보이는 모델 2를 중심으로 설명한다.

심리음향모델은 신호에너지(signal energy)에 의해 마스킹 되는 최대왜곡에너지(maximum distortion energy)를 계산하며, 이 에너지를 임계치(threshold)라고 한다.

심리음향 모델의 입력은 임계치 계산과정에 필요한 천이길이인 iblen, 지연(delay), 표본율(sampling rate)등이며, 출력은 SMR(Signal to Mask Ratio)과 임계치, 지연된 시간영역데이터(delayed time domain data), 창형태(window type), 예상 비트율 등이다.

[Step 1] 입력신호의 복소스펙트럼 (complex spectrum).

입력신호 s 를 Hann 윈도우를 곱하고, FFT(Fast Fourier Transform)을 수행하여 진폭응답 (magnitude response) $r(w)$ 와 위상응답(phase response) $f(w)$ 를 구한다.

[Step 2] 진폭응답과 위상응답의 예측.

현재 신호의 이전 두 프레임(frame)을 이용하여 진폭응답과 위상응답을 예측한다.

$$r_{pred}(w) = 2r_{t-1}(w) - r_{t-2}(w)$$

$$f_{pred}(w) = 2w_{t-1}(w) - w_{t-2}(w)$$

여기서, 현재 입력 프레임은 인덱스가 t 가 되고, 이전의 두 프레임의 인덱스는 각각, $t-1$ 과 $t-2$ 로 나타내었다.

[Step 3] 비 예측성 척도 (unpredictability measure) 계산.

$r(w)$ 와 $r(w)_{pred}$ 및 $f(w)$ 와 $f(w)_{pred}$ 를 이용하여 비예측성 척도 $c(w)$ 를 계산한다.

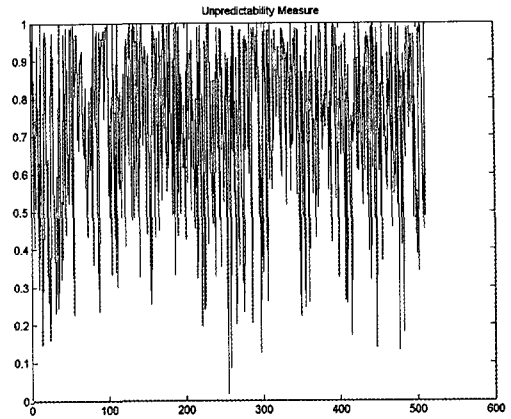
$$tmp1 = r(w)\cos(f(w) - r_{pred}(w)\cos(f_{pred}(w)))$$

$$tmp2 = r(w)\sin(f(w) - r_{pred}(w)\sin(f_{pred}(w)))$$

$$c(w) = \frac{\sqrt{tmp1^2 + tmp2^2}}{r(w) + |r_{pred}(w)|}$$

[Step 4] 임계치 계산 영역에서 에너지 및 가중 비 예측성 계산.

임계치 계산은 임계치 계산 영역에서 이루어지는데, 하나의 영역은 하나 이상의 스펙트럼 라인으로 이루어져 있고, 저주파에서 고주파로 갈수록 많은 스펙트럼 라인을 포함한다. MPEG 표준에서는 표본율에

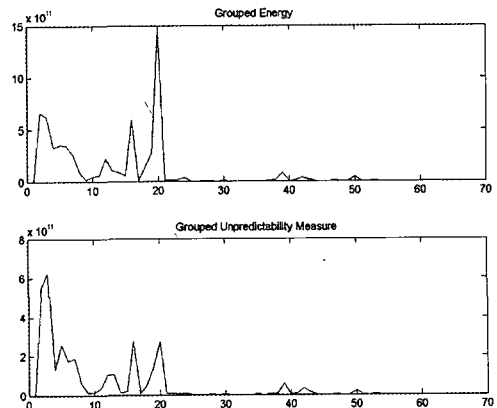


(그림 5) Unpredictability measure

따라 각각의 심리음향 매개변수표 (Psychoacoustic parameter table) 를 제공하고 있고, 이에 따라서 각 영역에 대하여 energy $e(b)$ 와 가중 비예측치 $c(b)$ 를 구한다.

$$e(b) = \sum_{w=w_{low}(b)}^{w_{high}(b)} r(w)^2$$

$$c(b) = \sum_{w=w_{low}(b)}^{w_{high}(b)} r(w)^2 c(w)$$



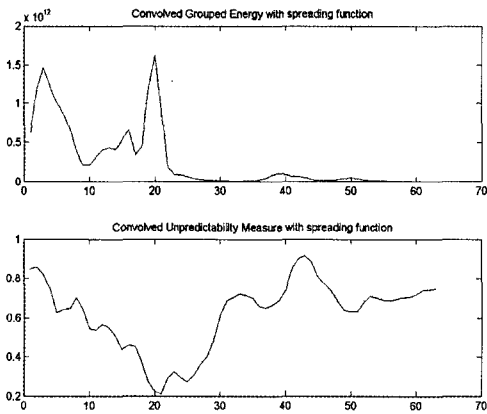
(그림 6) Grouped energy and unpredictability measure

[Step 5] 영역에너지 및 가중 비 예측치의 확산.

영역에너지 및 가중 비 예측치의 확산은 basilar 확산함수를 이용한다. 확산함수를 이용하여 영역에너지와 가중비 예측치를 콘볼루션 한다.

[Step 6] 순음지수 (tonality index) 계산.

정규화된 영역비 예측치는 순음지수로 바꿀 수 있다. 이때, 순음지수 0에서 1 사이의 범위에 있어야 한다. 순음지수는 1에 가까울수록 잡음(noise)에 가까운 성분이며, 0에 가까울수록 순음(tonal) 성분에 가까움을 나타낸다.



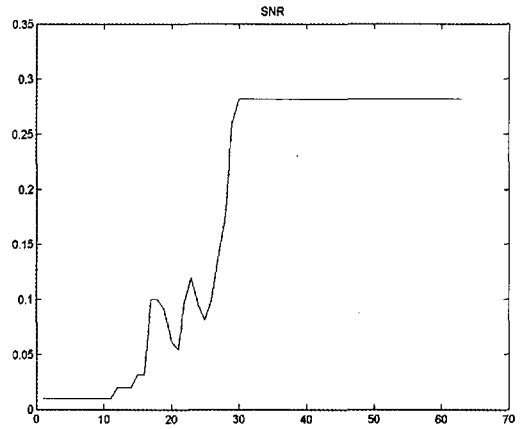
(그림 7) Convolved Grouped Energy & Unpredictability measure with spreading function

[Step 7] SNR, 파워비(Power ratio), 에너지 임계치 계산

각 영역에 대한 잡음마스킹순음(noise masking tone) NMT, 순음마스킹잡음(tone masking noise) TMN를 이용하여, 각 영역에서 요구되는 SNR을 구한다.

[Step 8] Pre-echo 조절과 quiet 임계치

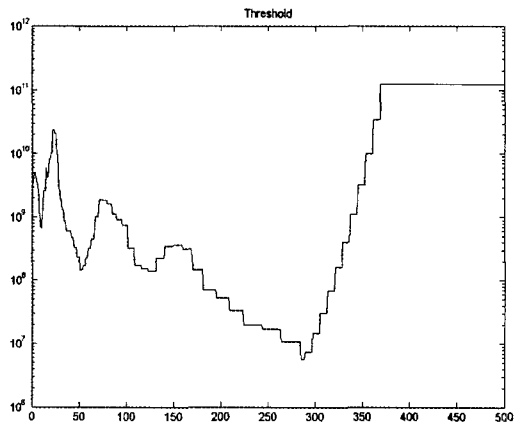
Pre-echo를 피하기 위해 quiet 임계치 및 이전블



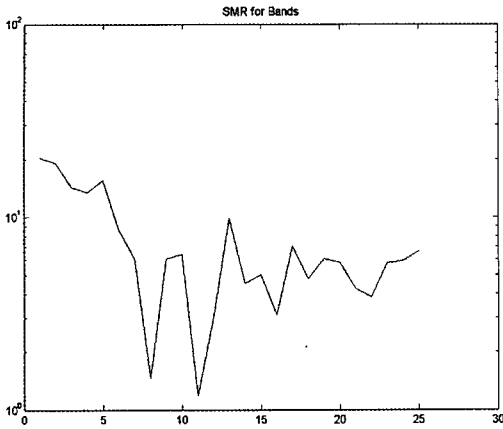
(그림 8) Masking index using TMN and NMT

력의 energy 임계치를 고려하여 pre-echo 조절치를 계산한다. Pre-echo는 변환블럭(transform block)의 끝부분 등에서 에너지가 적은 신호가 있을 경우, 역변환시 양자화에러가 복원된 신호의 전반에 걸쳐서 나타난 경우 발생한다.

[Step 9] SMR 계산



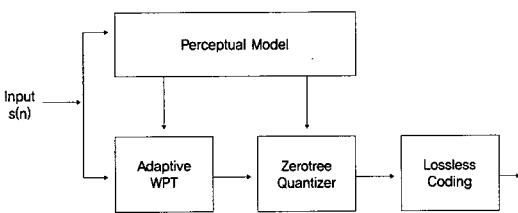
(그림 9) Masking Threshold



(그림 10) 각 서브밴드별 SMR

4. 부호화기 구조 및 양자화

(그림 11)은 전체적인 부호화 블록도를 나타내고 있다. 입력 오디오 샘플은 웨이블릿 필터뱅크 및 심리음향모델 블록으로 입력되고, 심리음향모델을 이용하여, 입력 신호의 마스킹 예상 비트율을 계산한다 [13].



(그림 11) 부호화기 구조

필터뱅크에 사용한 웨이블릿은 양정규직교(bi-orthogonal) 웨이블릿을 이용하였다. 이 필터뱅크를 이용하면 분해필터(analysis filter)와 합성필터(synthesis filter)가 다르기 때문에 부호화 및 복호화 시의 복잡도를 조절할 수 있고, 완전복원 및 anti-aliasing을 보장한다. 웨이블릿 필터뱅크는 입력신호

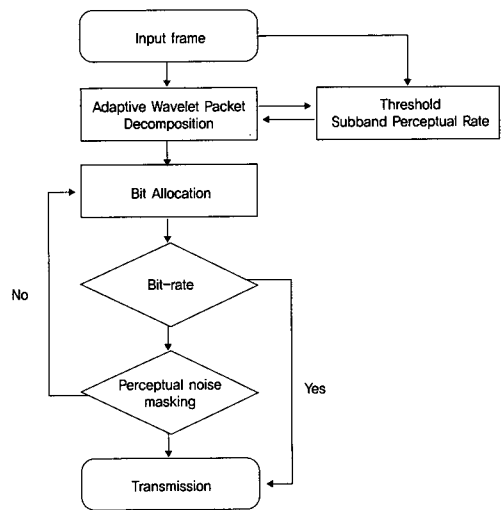
를 웨이블릿 패킷형태로 분해한 다음, 적응적으로 최적기저를 찾는다.

비트할당은 EZW 알고리즘을 목적하는 비트율에 도달할 때까지 반복한다. EZW 알고리즘은 맨처음 웨이블릿 계수를 저주파(coarse)에서 고주파(fine) 방향으로 tree 구조를 형성한다. 그 다음 계수를 POS, NEG, IZ, ZTR의 네가지 형태로 구분하고, 각각의 반복마다 임계치를 다시 계산한다.

POS는 positive를 의미하며, 계수의 크기가 임계치보다 크다는 것을 의미하고, NEG는 negative 이면서 계수의 크기가 임계치 보다 큰 것을 의미한다. ZTR은 계수와 하위 트리가 정해진 임계치 이하임을 나타내고, IZ은 이 계수는 임계치 이하지만 하위 tree 들이 정해진 임계치 이상임을 의미한다.

복호화기는 POS나 NEG의 처음 bit를 1로 복호하고, ZTR과 하위 트리는 0으로 복호한다. IZ 계수는 0으로 복호하고, IZ의 하위 트리는 POS나 NEG계수로 재해석한다.

매 반복마다, 임계치를 절반으로 줄인 후 다음 반복을 진행하게 된다. 또한, 반복할 때마다 지각에너지



(그림 12) 부호화기 흐름도

(perceptual energy)를 평가하여, 원하는 비트율에 도달하게 되면 알고리즘을 마치게 된다.

(그림 12)는 부호화기의 전체 흐름도를 나타내고 있다.

IV. 결 론

본 글에서는 웨이블릿에 기반 한 단일 코덱을 사용하여 비트율을 가변적으로 조정할 수 있는 제로트리를 이용한 오디오 부호화 기술에 대하여 소개하였다.

최근 들어 MPEG 및 ITU-T 등에서는 음성 및 오디오를 대상으로 하여 전송 상황에 따라, 비트율을 다이나믹하게 조절할 수 있는 계층적인 비트스트림을 제공하는 bit rate scalable coding 기술 필요성 인식에 따라 이에 다른 표준화 작업을 준비 중에 있다. 이러한 단일화된 코덱을 이용한 scalable 전송 기술은, 음성 및 오디오 스트림이 다양한 전송환경에서 제공되게 되며 네트워크 상황에 따라 사용 대역폭이 가변적인 환경에서 고품질의 다양한 서비스에 응용되어 유용하게 사용될 것으로 예상된다.

[참 고 문 헌]

- [1] D. Pan, "A Tutorial on MPEG/Audio Compression", *IEEE Multimedia*, pp.60-74, 1995.
- [2] S. Shilen, "Guide to MPEG-1 Audio Standard", *IEEE On Broadcasting*, Vol. 40, No. 4, December, 1994.
- [3] K.R.Rao and J.J.Hwang, *Techniques & Standards For ImageVideo & Audio Coding*, Prentice Hall, 1996.
- [4] D.Johnston, "Transform Coding of Audio Signal Using Preceptual Noise Criteria," *IEEE J. on Selected Areas in Comm.*, Vol.6, No.2, pp.314-323. Feb.1998
- [5] ISO/IEC 11818-7:1993 Information technology-Generic coding of moving picture and associated audio information Part 7: Advanced Audio Coding (AAC).
- [6] ISO/IEC 11172-3:1993 Information technology-Coding of moving pictures and associated audio for digital storage media at up to about 1.5 Mbit/s, Part3:Audio
- [7] ISO/IEC 144963-3 Information technology - Very Low Bitrate Audio-Visual Coding part3: Audio
- [8] ITU-T Temporary Document, TD 30 R1(WP 3/16), Report of Q.9/16 Meeting, Jan. 2004.
- [9] M.Vetterli, J.Kovacevic, "Wavelets and subband coding," Prentice Hall, 1995.
- [10] J.W.Woods, S D O'Neil, "Subband Coding of Images," *IEEE Trans. on Signal Processing*, vol. ASSP-34, No 5, Oct.1986, pp 1278-1288
- [11] S.Mallat, "A theory for multi-resolution signal decomposition: The wavelet representation," *IEEE Trans. on Pattern Anal. Machine Intell*, vol 11, No 7, July 1989.
- [12] J.Shapiro, "Image Coding Using Zerotrees of Wavelet Coefficients," *IEEE Trans. on Signal Processing*, vol.41, No.12 pp.3445-3462, Dec. 1993

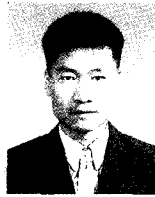
- [13] P.Srinivasan and L.H.Jamieson, "High quality audio compression using an adaptive wavelet packet decomposition and psychoacoustic modelling," *IEEE Trans. on Signal Processing*, vol 46, 1998.
- [14] D. Sinha and A.Tewfik., "Low bit rate transparent audio compression using adapted wavelets," *IEEE trans. on Signal Processing*, vol. 41, 1993.
- [15] D. Shinha and A.H.Tewfik, "Low Bit Rate Transparent Audio Compression Using Adapted Wavelet", *IEEE Trans. On Signal Processing*, Vol. 41, No.12, pp.3463–3479, Dec. 1993.
- [16] R.R.Coifman and M.V.Wickerhauser, "Entropy based algorithm for best basis selection," *IEEE Trans. on Information Theory*, vol. 38, 1992.
- [17] Mladen Victor Wickerhauser, *Adapted Wavelet Analysis from Theory to Software*, IEEE Press, 1994.



석종원

1993년 경북대학교 전자공학과 (공학사)
1995년 경북대학교 대학원 전자공학과 (공학석사)
1999년 경북대학교 대학원 전자공학과 (공학박사)
1999년 ~ 2004년 한국전자통신연구원 방송미디어 연구 그룹 선임연구원
2004년 ~ 현재 창원대학교 정보통신공학과 조교수

관심분야 : 멀티미디어 신호처리, 콘텐츠 보호관리 기술 등



장대영

1991년 부경대학교 전자공학과(공학사)
2000년 배재대학교 대학원 컴퓨터공학과 (공학 석사)

1991년 ~ 현재 한국전자통신연구원 전파방송연구 단방송미디어연구그룹 책임연구원
관심분야 : 오디오 신호처리, 음원분리, 객체기반

3차원 오디오 신호처리, MPEG, DMB 등



정영근

2006년 창원대학교 정보통신공학과 (공학사)
2006년 ~ 현재 창원대학교 정보통신공학과 석사 과정

관심분야 : 오디오 및 음성 신호처리, 콘텐츠 보호 관리 기술 등



장인선

2001년 충북대학교 전기전자공학부 (공학사)
2004년 포항공과대학교 대학원 컴퓨터공학과 (공학 석사)

2004년 ~ 현재 한국전자통신연구원 전파방송연구 단방송미디어연구그룹 연구원

관심분야 : 오디오 신호처리, 음원분리, 객체기반

3차원 오디오 신호처리, MPEG, DMB 등



서정일

1994년 경북대학교 전자공학과 (공학사)
1996년 경북대학교 대학원 전자공학과 (공학석사)
2005년 경북대학교 대학원 전자공학과 (공학박사)
1998년 ~ 1999년 LG 반도체 근무
1999년 ~ 2000년 현대 전자 근무
2000년 ~ 현재 한국전자통신연구원 전파방송연구 단방송미디어연구그룹 선임연구원

관심분야 : 음향신호처리, 음성신호처리, 객체기반 3차원 오디오 신호처리, MPEG, DMB 등