

객관적 음질 평가를 위한 PEAQ의 성능 평가 및 분석

박세형 류승완 박정렬 신재호
 동국대학교 전자공학과

Analysis and Evaluation of PEAQ : Objective Method for Perceived Audio Quality Measurement

Se-Hyoung Park, Seung-Wan Ryu, Jeong-Yeol Park, Jaeho Shin
 Dept. of Electronic Eng. Dongguk Univ.

Abstract - 디지털방송, DAB 등과 같은 디지털 오디오 방송 서비스를 위한 디지털 시스템을 설계하기 위해서는 오디오 음질을 평가하기 위한 방법이 필수적이다. 기존의 방식은 인간의 귀를 이용한 주관적 방식을 이용함으로써 많은 시간과 비용을 들이게 되며, 음질평가를 하는 사람의 주관적 의견에 많이 좌우하게 된다. 그러나 최근 ITU-R에서는 오디오 음질의 객관적 평가를 위한 BS.1387(PEAQ)를 제안함으로써 많은 시간과 비용을 절감하고 신뢰할 수 있는 결과를 얻게 되었다. PEAQ는 인간의 귀에서의 신호의 처리과정과 인식과정을 심리음향모델과 인식모델로 분리하여 구성함으로써 주관적 평가의 SDG(Subjective Difference Grade)에 대응하는 ODG(Objective Difference Grade)를 구하게 된다. 본 논문에서는 이러한 PEAQ의 심리음향 모델과 인식 모델을 원리와 과정을 평가·분석하였다.

및 전송 등 미디어 서비스에 많은 변화를 주고 있다. 그러나, 영상과는 달리 오디오 신호의 음질 평가 방식에는 적절한 객관적인 평가 방식이 제안되지 않아 인간의 귀에 의지하는 주관적인 평가방식을 이용하고 있다. 오디오 신호에 대한 주관적인 평가방식은 ITU-R에서 제안한 BS.1116을 이용하고 있으나 이 방법은 많은 비용과 시간의 제약, 그리고 사용방법에 제한적인 요소가 많아 이를 보완할 방법이 필요하였다.

PEAQ(Perceptual Evaluation of Audio Quality : ITU-R Recommendation BS.1387)는 기존의 주관적 음질평가 방법을 보완하기 위한 방법으로 제안되어졌다. 특히, 오디오 신호에 대한 SNR과 왜곡(distortion)의 측정과 같은 기존의 객관적 평가방식은 지각부호화를 이용하는 시스템에서의 적용이 불합리하기 때문이다. 또한, MPEG 오디오와 같이 표준적인 부호화기의 설계 시, 표준에서 요구되는 사항에 대한 최적의 결과를 도출하기 위해 구현과정에서의 변형은 음질에 영향을 미치게 되고 이러한 영향을 평가할 수 있는 방법이 필요하다. 그리고 디지털 방송, 온라인을 이용한 AOD등과 같이 연속적인 음질의 모니터링이 필요한 경우 주관적인 평가 방법은 전혀 사용할 수 없게 된다. 이와 같이 새로운 디지털 미디어 서비스의 환경변화는 기존의 주관적인

1. 서론

디지털 기술의 급속한 발전에 따라 기존의 미디어의 저장 및 전송 등 모든 처리과정은 점차 디지털화 되어 가고 있다. 최근 디지털 오디오 신호에 대한 압축기술 중 특히 심리 음향 모델에 근거를 둔 지각부호화 방식은 CD수준의 높은 음질을 보장하며 저장

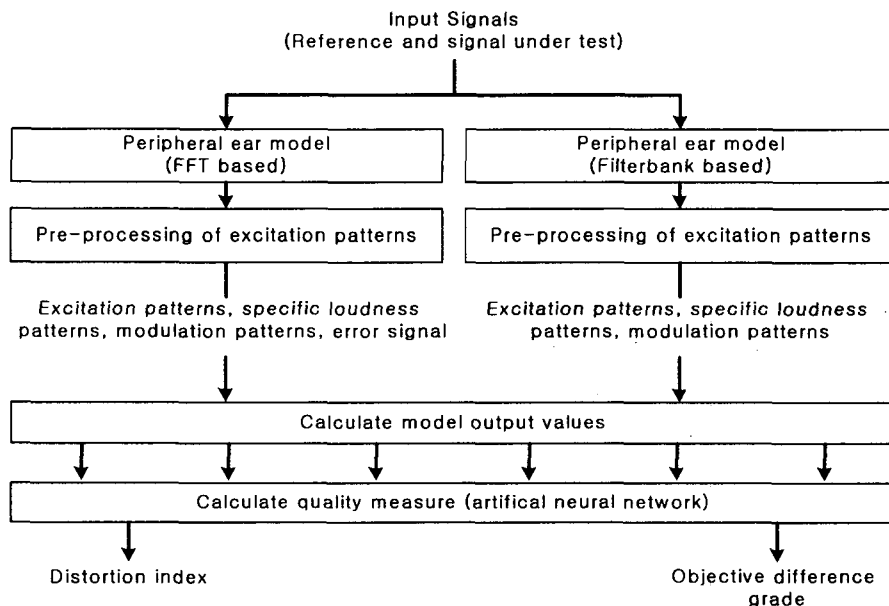


그림 1. PEAQ의 전체블록(BS.1387)

평가방법을 보완하며 다양한 분야에서의 음질 평가를 위한 방법이 필요하게 되었다.

본 논문에서는 PEAQ를 분석·평가함으로서 인간의 청각 시스템을 더욱 잘 나타낼 수 있는 새로운 지각 모델을 설계하거나, 지속적으로 발전되는 음향압 측정방식에 대응한 새로운 평가방법을 설계하는 등 발전하는 디지털 오디오 기술의 분석을 목적으로 하였다. 2장에서는 전체적인 PEAQ의 과정을 심리음향모델과 인식모델로 나누어 분석하고, 결론에서는 PEAQ에서 분석·평가 과정에서 추가적으로 연구가 필요한 부분에 대해 알아보고 결론을 맺었다. 본 논문에서의 식의 표기에서 k 는 주파수 인덱스로 n 은 시간영역에서의 샘플 블럭의 인덱스를 나타낸다.

2. 심리음향 모델과 인식모델

PEAQ는 크게 심리음향모델(Psychoacoustic)부분과 이 출력의 결과를 처리하는 인식모델(Cognitive Model) 두 가지로 나눌 수 있다. 심리음향모델은 인간 귀의 구조를 디지털 필터 등을 이용하여 구현하는 것이고, 인식모델은 심리음향모델의 출력을 입력으로 하여 귀의 세포와 뇌에서 처리과정과 인간의 인식과정을 구현하는 것으로 볼 수 있다. 그림1은 PEAQ의 전체적이 처리과정을 보여준다.

그림1에서와 같이 PEAQ는 두개의 심리음향 모델을 가지고 있는데 FFT기반의 심리음향모델과 필터뱅크 기반의 심리음향 모델로 나누어진다. 여기서 각각의 모델은 용도에 따라 Basic버전과 Advanced버전으로 나뉘어 용도에 따라 두 버전이 달리 사용되어 지고 있다. 표 1은 적용되는 용도에 따라 사용가능한 버전을 보여준다.

표 1. 버전별 응용분야

Application	Brief Description	Version
표준 코덱 구현 평가	오디오 코덱의 구현 시 발생하는 영향평가.	Basic/Advanced
서비스 라인 음질평가	방송등 미디어 서비스라인 점에서의 모니터링	Basic
온라인 모니터링	연속적인 서비스에서 오디오 전송을 모니터링	Basic
장치·연결	일련의 장치 혹은 회로의 세부적인 분석	Advanced
코덱 확인	특별한 코덱의 성능확인	Advanced
코덱 개발	새로운 코덱의 개발과정에서의 평가	Basic/Advanced
네트워크 planning	주어진 제약하에서 전송 네트워크의 비용과 이행을 최적화하기 위한처리	Basic/Advanced
주관적 평가의 보완	listening test의 보완수단	Basic/Advanced

전반적인 처리과정은 레퍼런스 신호(Reference signal)와 평가하고자 하는 신호(Signal under test)¹⁾는 모델의 입력으로 들어가 심리음향적 처리과정을 통해 음질의 인식과정에 필요한 파라미터를 생성하게 되고 이 파라미터는 인식모델의 입력으로 MOV (Model Output Variable)을 구하여 인식모델의 인공 신경망을 통해 객관적 차이인 ODG(Objective Difference Grade)를 구하게 된다.

2.1 FFT기반 심리음향모델

심리음향모델은 인간의 청각기관을 필터 등 전기적인 요소로 분리 재구성하여 이를 구현하는데, 이는 크게 세 부분으로 나뉘어 진다.

첫째, 외이와 중이로서 외이는 컷바퀴를 통해 소리를 모으는 과정과 함께 소리의 음원을 인식할 수 정보를 생성하는 역할을 수행하게 되고, 중이는 외이를 통해 들어온 신호를 인간의 귀에 적당한 주파수 대역에 대해 증폭과 감쇠를 해주는 필터 역할을 하게 된다. 이 과정은 주로 디지털 필터를 이용하여 구현할 수 있다.

둘째, 달팽이관(Cochlea)과 BM(Basilar Membrane)으로 구성된 부분은 심리음향처리과정의 가장 중요한 부분 중 하나로서 일종의 주파수 스펙트럼 분석기로 생각할 수 있다. 그러나 모든 주파수 성분을 알아내는 정밀한 스펙트럼과는 달리 인간의 청각기관은 특별한 주파수 간격으로 신호를 분해하는 스펙트럼과 같다. 특히 위와 같은 주파수 간격을 임계대역(Critical Band)이라고 한다. 임계대역은 심리음향 모델에서 상당히 중요한 역할을 하게 되는데 E.Zwicker에 의해 다음과 같이 정의 되었다.²⁾

$$z/Bark = 7 \cdot \operatorname{arsinh}\left(\frac{f/HZ}{650}\right) \quad (1)$$

임계대역은 80~18,000Hz의 주파수 대역을 25개의 대역으로 분할하고 각 대역에 Bark라는 단위를 사용한다.

(1)과 같이 임계대역은 비선형적이기 때문에 많은 불편함이 따른다. 그러나 대부분의 심리음향 모델은 임계대역(또는 Bark스케일)으로 가청주파수 대역을 분할한다. 이외에 사용되는 주파수의 대역분할방식은 1/3 옥타브대역, ERB(Equivalent Rectangular Bandwidth)등을 사용하기도 한다. 이 과정은 크게 두 가지 방식으로 구현될 수 있는데 첫째, DFT(FFT), DCT, DWT등과 같이 트랜스폼을 이용한 후 그룹화에 의한 방법과 필터뱅크를 이용하는 방법이 있다.

마지막 부분은 내이의 외이세포(Outer ear Cell)와 내이세포(Inner ear Cell)로 구성되어 있는데 외이세포는 신호의 달팽이관을 통해 들어온 BM(Basilar Membrane)의 움직임을 통해 귀의 소리의 세기에 따라 귀의 손상을 막기 위한 세기의 조절 기능과 함께

1) Signal under test : 평가받는 신호, 이후 테스트신호라 함.

신호를 내이세포에 전달하고 내이세포는 이 신호를 바탕으로 신경계에서 이용 가능한 신호로 변환하여 준다. 이러한 부분은 신호의 세기에 따라 적응(Adaptation)과 정류(Rectification)등을 통하여 이루어지게 된다.

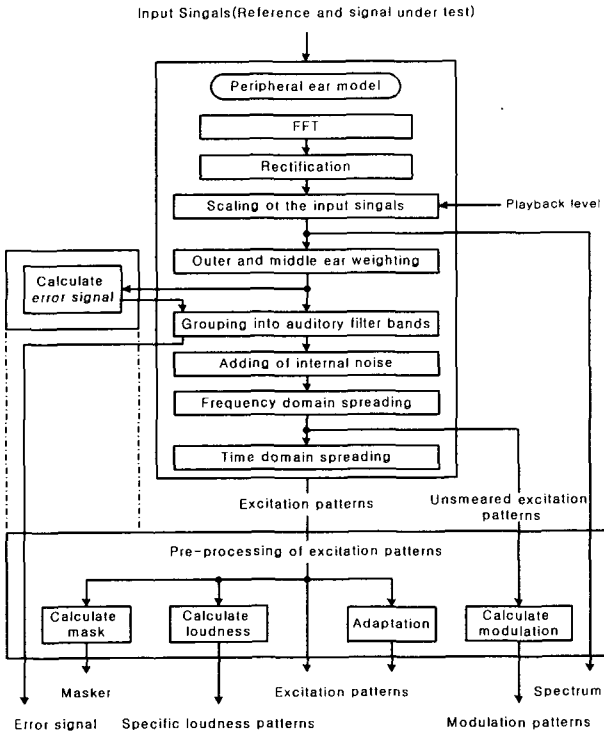


그림 2. FFT기반 심리음향모델의 처리과정(BS.1387)

이와 같이 인간의 청각기관은 그림2나 6와 같이 트랜스폼이나 필터뱅크를 이용하는 심리음향모델로 모델링 되어 질 수 있고, 이는 디지털 신호처리 기술을 이용하여 구현할 수 있게 된다.

그림 2는 FFT기반의 심리음향모델의 과정을 보여 준다. 레퍼런스 신호와 테스트신호는 48kHz의 샘플링 주파수를 가지고 A/D변환되어 심리음향모델의 입력으로 주어진다. 매 입력은 2048개 샘플씩 FFT의 입력으로 주어지게 된다. 이 때, 사용되는 윈도우는 Hanning 윈도우가 사용되어지고, 오버랩사이즈는 1024이다.

FFT후 입력되는 신호는 입력신호의 SPL(Sound Pressure Level)에 따라 스케일링 되는데 이는 다음의 식과 같다.

$$fac = \frac{10^{\frac{Lp}{20}}}{Norm} \quad (2)$$

$$F[k, n] = fac \cdot F_f[k, n] \quad (3)$$

여기서 $F_f[k, n]$ 은 입력신호의 FFT계수를 의미한다. 스케일링된 FFT계수는 외이와 중이를 모델링한 필터에 의해 필터링 되는데, 이는 PEAQ에서는

가중치 함수를 이용하며 이는 다음의 식과 그림과 같은 주파수 응답을 갖게 된다.

$$W[k]/dB = -0.6 \cdot 3.64 \cdot \left(\frac{f[k]}{kHz}\right)^{-0.8} + 6.5 \cdot e^{-0.6 \cdot \left(\frac{f[k]}{kHz} - 3.3\right)^2} - 10^{-3} \cdot \left(\frac{f[k]}{kHz}\right)^{3.6} \quad (4)$$

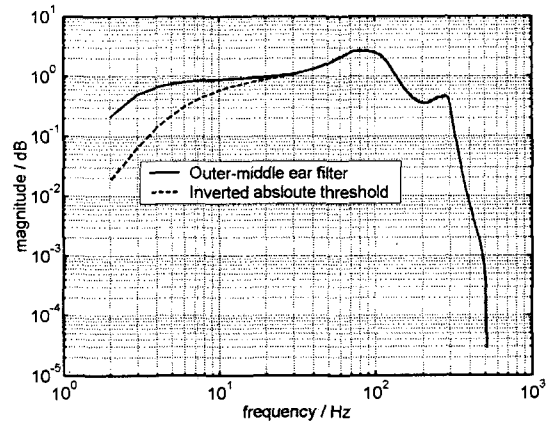


그림 3. 가중치 함수의 주파수 응답

PEAQ는 이 과정을 FFT계수의 가중치로서 연산을 수행하였다. 그러나 이 과정은 DC 제거필터와 함께 FFT를 수행하기 전에 디지털 필터로서 구현될 수 있다.

PEAQ에서는 양귀(Binaural)특성을 고려하기 위해 각각의 입력신호의 채널별로 처리를 하였다. 그러나 이러한 평가는 각각의 채널이 독립된 채널과 같은 평가를 낼 뿐, 실제 음질의 평가에서 사용자의 환경 즉 라우드 스피커를 이용하거나 또는 헤드폰을 이용할 경우에 따른 음질의 차이를 고려하기 위해서는 HRTF(Head Related Transfer Function)을 이용할 수 있다. 그리고 AC-3, AAC등의 2채널 이상의 음질을 테스트하기 위한 음상정위, 공간감, 확산감등의 요소에 대한 평가요소와 과정이 없다.(특히 PEAQ에서는 Binaural에 대한 영향에 대한 음질평가 항목이 미미하다.)

FFT기반의 음향모델은 FFT의 특성에 따라 시간과 주파수의 해상도에 trade-off를 가지게 된다. 이는 빠른 연산속도를 가지고 있지만 시간-주파수의 분해능의 제약을 갖게 한다. 그리고 주파수의 해상도는 DC부터 $F_s/2$ 까지의 주파수 대역을 선형적으로 분할하게 됨으로 인간이 인지하는 임계대역과 다른 주파수 분해 능력을 갖게 된다. 따라서 가중치된 FFT계수는 임계대역으로 그룹화 되어져야 한다. 이과정은 FFT와 같이 고속연산기능을 가지고 있지만 인간의 청각기관을 모델링하기 위한 과정에서의 단점이 된다. 이러한 과정은 웨이브렛을 이용한 DWPT(Discrete Wavelet Packet Transform), FAFT(Fast Approximation Fourier Transform)과 같은 방식을 이용하는 것도 좋은 방법이 될 것이다.

PEAQ의 FFT기반의 심리음향모델은 2가지 버전을 제안하는데, Basic버전에서는 각각의 임계대역이 0.25 Bark의 대역폭으로 총 109개의 대역으로 나누어 그룹화하고, Advanced 버전에서는 0.5Bark로 55개의 대역으로 그룹화하게 된다. 여기서 임계대역으로 그룹화되기 전에 레퍼런스 신호와 테스트신호의 차를 계산하여 에러신호 즉 노이즈 패턴 MOV(Model Output Valuable)를 구하게 된다.

다음 그룹화 된 신호는 실제 인간의 귀에서 처리과정에서 발생하는 손실을 보상하는데 이때 추가되는 내부노이즈는 주파수에 의존적이고 다음과 같은 값을 가지게 된다.

$$P_{Thres}[k] = 10^{0.1456 \cdot \left(\frac{f_c[k]}{kHz}\right)^{-0.8}} \quad (5)$$

여기서 $f_c[k]$ 는 각각의 그룹화 된 임계대역의 중심 주파수를 말한다. 그리고 그룹화 된 FFT계수 $P_e[k, n]$ 은 다음 식을 통해 피치패턴(Pitch Pattern)

$$P_p[k, n] = P_e[k, n] + P_{Thres}[k] \quad (6)$$

$P_p[k, n]$ 을 구하게 된다.

지금까지의 처리 결과인 피치 패턴은 이제 심리음향 모델의 주된 요소인 마스크 효과를 적용하게 된다. 마스크 효과는 주파수 마스크와 시간 마스크의 두 가지로 분리하여 볼 수 있다. 먼저 주파수 마스크는 다음의 여기패턴에서 보는 바와 같이 특정한 주파수 성분의 세기에 의해 주변의 신호가 마스크 되는 현상을 보여주는데 마스크 되는 과정은 주파수 성분의 크기에 따라 다르게 되고, 또한 높은 주파수 측과 낮은 주파수 측으로의 마스크 곡선이 다르게 된다. 이러한 결과는 다음의 그림4에 나타나 있다.

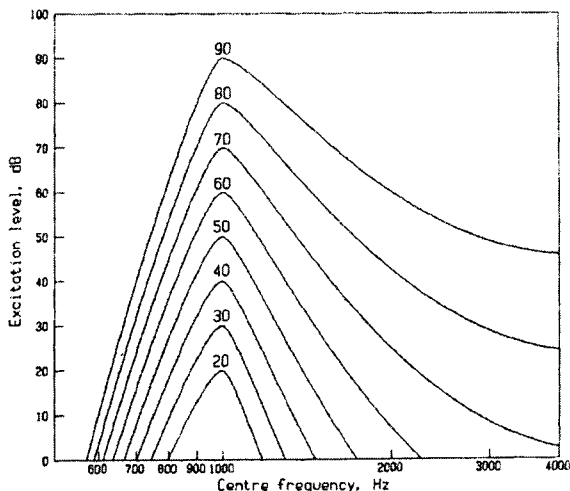


그림 4 신호의 크기에 따른 여기패턴(Terhardt). [2]

2) $f_c[k]$ 의 대역은 참고문헌 [1]을 참조바람.

주파수 마스크는 대부분의 지각부호화 오디오 코덱에서 주요한 압축기법으로 데이터 량을 줄이는데 가장 많은 영향을 미친다. 주파수 마스크의 여기패턴은 Unsmearred 여기패턴 이라 하고, 이는 시간영역에서의 마스크 효과가 적용되지 않은 상태에서의 신호의 패턴을 말하게 된다. 일반적으로 주파수 마스크는 스프레딩(Spreading)이라 하는데 이는 마스크 이외의 주파수 성분이 여기패턴에 의해 평균화되어지는 것이다. 즉 이 값은 입력신호의 라우드니스 값(주파수 영역)의 평균값과 평균값의 차를 구하기 위해 이용된다. 이 과정은 시간 영역에서의 마스크 즉, 시간영역에서 신호의 smoothing효과를 배제한 상태에서의 값을 의미한다.

다음은 주파수 마스크효과를 적용한 후 시간영역에서의 마스크 즉 smoothing효과를 적용하는 과정이다. 시간영역에서의 spreading은 신호의 envelope를 smoothing하게 하는 효과를 나타내는데, 이는 다음의 그림에서 보는 바와 같이 시간적으로 큰 소리에 의해 시간상 앞뒤(-50~200ms)의 소리가 마스크 되는 현상을 모델링 하는 과정을 의미하고, 다음의 그림은 시간 마스크에 대해 설명한다.

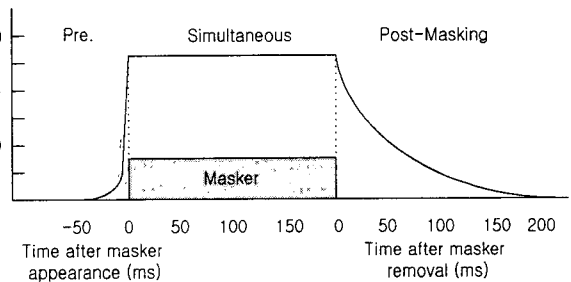


그림 5. 시간영역의 마스크

PEAQ에서의 FFT기반의 심리음향모델의 단점은 Non-stationary 신호와 같이 매우 급격히 변화하는 신호의 경우 FFT기반의 블록 샘플길이가 2048이기 때문에 변화가 빠른 신호에 대한 적절한 대응을 하지 못하게 된다. 따라서 입력 블록의 stationary/Non-stationary 특성을 구하기 위한 Transient 검출기법등의 추가가 필요하다고 생각된다.

마지막으로 시간영역에서의 마스크과정을 통과한 신호는 여기패턴(Excitation Pattern)은 인식모델의 입력으로 가기 전 전 처리과정(Pre-Processing)을 거쳐 에러신호, 마스크, 라우드니스 패턴, 변조패턴, 주파수 스펙트럼 등의 변수로서 인식과정으로 넘겨지게 된다.

2.2 필터뱅크기반 심리음향모델

필터뱅크 기반의 심리음향 모델은 신호의 스펙트럼을 구하기 위한 과정이 FFT가 아닌 필터뱅크를 이용한 점이 구별된다. PEAQ에서는 비선형적인 필터뱅크를 이용하여 40개의 대역으로 입력신호를 필터링하여 사용하는데 이때 필터는 다음과 같은 In-

phase 필터와 Quadrature 필터를 사용한다.

$$h_I[k, n] = h_p[k, n] \cos(2\pi f_c[k] \cdot (n - N_k/2)/F_s)$$

$$h_Q[k, n] = h_p[k, n] \sin(2\pi f_c[k] \cdot (n - N_k/2)/F_s)$$

$$h_p[k, n] = (4/N_k) \cdot \sin^2(\pi n/N_k) \quad (7)$$

이때 N_k 는 임펄스 응답의 길이로 최대 1456에서 최소 52개 이다. 필터뱅크를 통과한 값들은 1/32로 서브 샘플링되어 매 192 샘플이 하나의 블록을 구성하여 FFT기반의 심리음향모델과 같은 처리과정을 거치게 된다. 이러한 필터뱅크의 구현은 좋은 시간해상도와 주파수 해상도를 가지며 필요에 따라 필터뱅크를 다시 구현함으로써 많은 유연성을 제공할 수 있게 된다. 다음의 그림 5는 필터뱅크 기반의 심리음향모델의 과정을 보여준다.

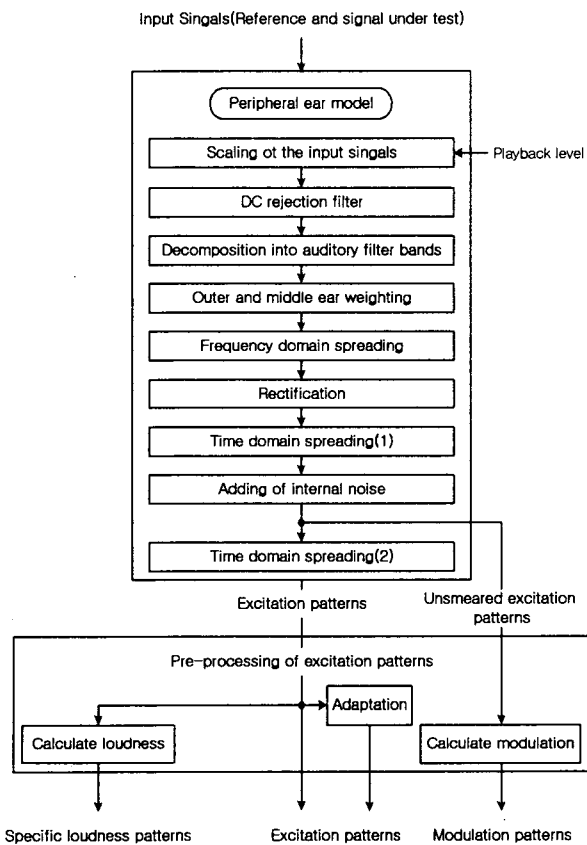


그림 6 필터뱅크기반 심리음향모델(BS.1387)

PEAQ에서는 Non-linear filter bank를 이용하는데 이 과정 후 FFT기반의 경우와 마찬가지로 주파수·시간 마스크 처리과정을 거치게 된다. 그러나 [4]에서 제안한 방법과 같이 비대칭 Gamma Chirp 필터뱅크를 이용함으로써 필터뱅크, 주파수, 시간 마스크의 과정을 하나의 필터뱅크과정으로 처리할 수 있게 되며,

이러한 Gamma Chirp 필터뱅크의 경우 낮은 차수의 IIR필터로 구현됨으로 많은 연산량을 줄일 수 있는 대안이 될 수 있을 것이다.

2.3. 인식모델

PEAQ의 인식모델은 인간의 뇌와 같이 심리음향 모델의 결과를 가지고 인간이 인식하는 것과 같은 주요한 요인들의 값을 구하게 된다. 이 과정은 MOV를 계산하는 과정과 MOV를 이용하여 주관적인 평가방법인 SDG와 유사한 ODG(Objective Difference Grade)를 구하는 인공신경망을 회로로 되어 있다. 먼저 인식모델에서 사용하는 MOV 표2와 3과 같이 각각 7가지와 5가지로 구분되어진다.

표 2. Advanced 버전의 MOV

MOVs	목적
RmsNoiseLoudAsymA	왜곡에서의 loudness
RmsModDiffA	변조에서의 변화 (roughness에 관련하여)
AvgLinDistA	선형 왜곡(주파수 응답 등)
Segmental NMRB	Noise-to-Mask Ratio
EHSB	에러의 harmonic 구조

표 3 Basic버전의 MOV

MOVs	목적
WinModDiff1 _B AvgModDiff1 _B AvgModDiff2 _B	변조도의 차(related to Roughness)
RmsNoiseLoud _B	왜곡성분의 라우드니스
BandwidthRef _B BandwidthTest _B	선형왜곡성분(주파수 응답 등)
RelDistFrame _B	귀에 들리는 왜곡된 주파수
Total NMR _B	Noise-to-Mask Ratio
MFPD _B ADB _B	검출확률
EHS _B	에러의 harmonic 구조

계산된 표 2,3의 MOV들은 인공지능신경망을 통하여 ODG로 예측되는데 이때 사용되는 신경망은 1층의 hidden layer를 가지고 있고, 입력I(Basic : 11, Adv. : 5)와 노드J(Basic : 3, Adv. : 5)를 가지는 신경망을 구성하고 그 식은 다음과 같다.

$$sig(x) \cong \frac{1}{1 + e^{-x}}$$

(8)

$$DI = w_y[J] + \sum_{j=0}^{J-1} (w_y[j] \cdot sig(w_x[I, j] + \sum_{i=0}^{I-1} w_x[i, j] \cdot \frac{x[i] - a_{min}[i]}{a_{max}[i] - a_{min}[i]})))$$

(9)

$$ODG = b_{min} + (b_{max} - b_{min}) \cdot sig(DI)$$

(10)

표 4. b_{min}, b_{max} 값

	b_{min}	b_{max}
ODG	-3.98	0.22

인식과정은 MOV를 계산하는 과정과 인공신경망을 이용하는 ODG를 구하는 과정으로 구성되어 지고, 이를 이용하여 객관적인 음질 평가를 하게 되어진다. 그러나 MOV에서 일반적인 음질평가에서 고려되어지는 라우드니스(Loudness), 러프니스(Roughness), 왜곡(Distortion), 하모닉 디스토션(Harmonic Distortion)등은 있으나 샤프니스(Sharpness), 플럭테이션(Fluctuation), 리듬(Rhythm)에 대한 평가항목은 제외되었다. 또한 최근 [3]에서 제안되는 심리음향 요소인 Cubic distortion tone등의 고려가 필요하겠다.

3. 결 론

객관적 음질 평가 방법의 연구는 평가방법의 연구와 더불어 기존은 오디오 코더의 성능을 향상 시킬 수 있는 진보된 심리음향 모델을 제시하게 될 것이다. 또한 디지털 미디어 서비스의 품질을 유지 관리하는데 많은 도움이 될 것이라 생각한다. PEAQ는 주관적 음질평가 방법과 함께 보완하여 디지털 오디오의 음질평가를 쉽고 편리하게 해 줄 수 있다.

ODG를 이용하여 SDG와의 유효한 처리과정을 Phase 1, 2, 3으로 분리하여 평가하는데 이 부분에 대한 분석은 본 논문에 포함하지 않았다.

지금까지 본 논문에서 살펴본 바와 같이 PEAQ에서의 다음과 같은 보완 및 개선점이 있음을 확인할 수 있었다.

1. Playback Scaling의 개선.
2. HRTF와 같이 외이 효과나 서라운드 시스템의 음상 효과의 변위 측정.
3. 새로운 트랜스폼의 적용.
4. 라우드스피커 사용시의 측정과 헤드폰 등 사용자의 환경에 따른 평가방법.
5. Transient Detection 기법 적용.
6. Gamma Chirp 필터의 이용한 필터뱅크의 개선.
7. Sharpness, CDT등 새로운 심리음향 요소의 추가.

[참 고 문 헌]

- [1] ITU-R Rec. BS.1387, "Method for Objective Measurements of Perceived Audio Quality (PEAQ)," International Telecommunication Union, Geneva, Switzerland, 1998.
- [2] E. Zwicker and H. Fastl, *Psychoacoustics: Facts and Models*, 2nd edition, Springer-Verlag, Berlin Heidelberg, 1999.
- [3] David J M Robinson & Malcolm J Hawksford, "Time-Domain Auditory Model for The Assessment of High-Quality Coded Audio," 107th convention of AES, preprint 5017, Sep. 1999.
- [4] Toshio Irino, Masashi Unoki, "An Analysis/Synthesis Auditory Filterbank Based on an IIR Implementation of the Gammachirp" *Journal of the Acoustic society*. vol 20, pp.397-406, Nov. 1999.