

MPEG 심리음향 모델-II 알고리즘의 ASIC 구현을 위한 고정 소수점 연산 최적화

정회원 이근섭*, 박영철**, 윤대희*

Fixed-point Processing Optimization of MPEG Psychoacoustic Model-II Algorithm for ASIC Implementation

Keun-Sup Lee*, Young-Cheol Park**, Dae Hee Youn* *Regular Members*

요 약

본 논문에서는 MPEG 오디오 Layer-III 부호화 과정 중 심리음향 모델 과정을 ASIC으로 구현하기 위하여 고정 소수점 연산기에 적합하도록 최적화를 수행하였다. 최적화 과정은 크게 부호화기의 음질을 고려하여 프로세서의 데이터 워드 길이를 결정하는 과정과 자주 사용되는 초월 함수를 고정 소수점 연산을 통해 구현하는 것으로 구성된다. 데이터 워드 길이를 결정하기 위하여 심리음향 모델 과정의 고정 소수점 연산 오차와 이 오차가 비트 할당 과정에 영향을 미칠 확률 사이의 관계를 통계적 모델로 정의하였다. 여기서 정의된 모델을 사용하여 고정 소수점 연산 오차에 의한 영향이 1% 이내가 되도록 24비트의 데이터 워드를 선택하였다. 최적화된 고정 소수점 심리음향 모델을 사용한 MP3 부호화기의 음질은 부동 소수점 부호화기에 비해 ITU-R의 음질평가 점수를 기준으로 평균 -0.2 이내의 구분하기 힘든 수준의 음질 저하를 보였다

Key Words : MP3 encoder, Fixed-point optimization, Psychoacoustic model, Data Word length, Statistical model,

ABSTRACT

The psychoacoustic model in MPEG audio layer-III (MP3) encoder is optimized for the fixed-point processing. The optimization process consists of determining the data word length of arithmetic unit and the algorithm for transcendental functions that are often used in the psychoacoustic model. In order to determine the data word length, we defined a statistical model expressing the relation between the fixed-point operation errors of the psychoacoustic model and the probability of alteration of the allocated bits due to these errors. Based on the simulations using this model, we chose a 24-bit data path and constructed a 24-bit fixed-point MP3 encoder. Sound quality tests using the constructed fixed-point encoder showed a mean degradation of -0.2 on ITU-R 5-point audio impairment scale.

1. 서론

디지털 오디오 기기에서 MPEG 오디오 비트열을 얻기 위해서는 실시간으로 오디오 알고리즘의 부호화 과정을 수행해야 한다. 디지털 오디오 알고리즘

의 부호화 과정은 크게 T/F 변환(Time-to-Frequency Mapping), 심리음향 모델, 비트 할당 과정, 마지막으로 비트열 생성 과정으로 이루어지며, 대부분의 연산량이 비트할당 과정과 심리음향 모델에 집중되어 있다^[1]. MPEG 오디오 부호화 과정들

* 연세대학교 전기전자공학과 미디어 통신 신호처리 연구실(taraji@cyclon.yonsei.ac.kr, dhyoun@yonsei.ac.kr),

** 연세대학교 정보기술학부(young00@dragon.yonsei.ac.kr)

논문번호 : KICS2004-07-107, 접수일자 : 2004년 7월 20일

중 심리음향 모델은 주파수 영역으로 변환된 입력 신호를 이용하여 마스킹 곡선을 구하는 과정이다. 만약 데이터 워드 길이가 충분히 길지 않은 고정 소수점 연산 프로세서를 이용하여 부호화 과정을 구현할 경우, 부동 소수점 연산에 비해 마스킹 임계치 값의 오차가 크게 발생할 수 있다. 이 경우 심리음향 모델의 오차가 발생할 수 있으며 청각적으로 음질이 저하될 수 있다. 또한 심리음향 모델에서 사용되는 사인, 코사인 함수나 로그연산 등과 같은 초월함수 연산은 고정 소수점 프로세서의 연산기로는 구현이 용이하지 않다^[2]. 따라서 초월함수를 고정 소수점 연산기를 이용하여 소프트웨어로 구현할 경우, 일반적으로 적지않은 연산량이 필요하며 연산 오차가 발생할 가능성도 크다. 이러한 초월함수 연산 오차 역시 심리음향 모델의 정확성에 영향을 줄 수 있으며, 부호화기의 음질의 저하를 막기위해 심각하게 고려해야 할 문제이다^[2].

기존의 실시간 MPEG 오디오 부호화기는 주로 PC 기반, DSP 기반 그리고 RISC 프로세서 기반으로 구현되었으며, 주어진 구현 환경에 최적화하여 구현하는 방법에 대한 연구가 이루어져 왔다. PC 기반의 구현은 주로 Pentium 프로세서에 적합한 알고리즘 선택과 SIMD (Single Instruction stream-Multiple Data stream) 명령어를 사용한 연산량 최적화가 주로 연구되어 왔으며^{[3][4]}, DSP 기반 부호화기 구현은 부동 소수점 연산기나 고정 소수점 연산기를 사용하여 DSP가 가진 자원을 최대한 활용하는 방향으로 구현되어 왔다^[5]. 또한 RISC 프로세서는 적은 하드웨어 자원으로 인한 저전력, 저비용 구현이 장점이며, 32비트의 넓은 데이터 워드 길이로 높은 정확도를 갖는 부호화기를 구현할 수 있다^[6]. 그러나 부호화 알고리즘에 적합한 구조에 대한 연구나 프로세서 선택 기준에 관한 연구, 그리고 그에 맞는 최적화에 대한 연구는 미미했다.

본 논문은 MP3 부호화기 과정 중 복잡도가 높은 심리음향 모델 과정을 ASIC으로 구현하기 위하여 최적의 구조를 선택하고 고정 소수점 연산기에 적합한 연산 최적화를 수행하였다. 하드웨어의 크기와 구현 비용에 가장 큰 영향을 미치는 데이터 워드 길이를 결정하기 위하여, 데이터 워드 길이와 부호화기의 음질 사이의 관계를 통계적으로 모델링하였다. 심리음향 모델에서 발생한 연산 오차는 비트 할당 과정에 영향을 끼쳐 음질 열화를 일으키게 된다. 본 논문은 심리음향 모델의 연산 오차와 비트 할당 과정과의 관계를 분석하고, 허용 가능한 최대 고정 소수점 연산

오차를 이 오차로 인해 비트 할당 과정이 영향을 받을 확률 값에 대한 식으로 나타내었다. 이 관계를 통해 심리음향 모델의 연산 오차가 비트 할당 과정에 미치는 영향을 수치화할 수 있게 되었으며, 이 결과로부터 심리음향 모델을 고정 소수점 연산으로 구현하는데 필요한 기준을 마련하였다.

모든 최적화 과정은 부호화기의 음질에 영향을 미치는 연산 정확도를 높이고 이와 함께 연산량과 하드웨어 자원이 최소화되도록 하였다. 연산 정확도와 연산량과의 반비례 관계를 고려하여 최적화된 데이터 워드 길이를 결정하였으며, 심리음향 모델에서 자주 사용되는 초월함수를 구현하기에 적절한 크기의 테이블 크기를 결정하였다.

본 논문에서 사용한 프로세서 구조는 1개의 MAC (Multiplication and Accumulation) 연산기를 갖는 전형적인 DSP 구조를 가정하였으며 본 연구 결과는 ASIC 구현 뿐만 아니라 일반 범용 프로세서에 대한 선택 기준으로도 활용될 수 있다.

II. 프로세서의 데이터 워드 길이 결정

MP3 부호화 과정에서 고정 소수점 연산을 수행하는 부분은 PCM 오디오 입력으로부터 심리음향 모델의 정보를 바탕으로 비트 할당을 하는 과정까지이다. 이들 부호화 과정들 중 특히 비트 할당 과정은 오디오 부호화 과정에서 발생하는 양자화 오차와 심리음향 모델에서 얻은 마스킹 임계치가 비교되는 과정이며, 동시에 각각에서 발생한 고정 소수점 연산 오차의 영향을 함께 받는 과정이므로 고정 소수점 연산 오차 최적화 과정에서 가장 중요한 부분이다. 또한 이 과정에서의 연산 오차는 양자화 잡음의 제어를 부정확하게 만들고 나아가 부호화기의 음질을 떨어뜨릴 수 있으므로, 부호화기의 성능을 유지하기 위해서는 이 과정을 기준으로 고정 소수점 연산을 최적화할 필요가 있다.

비트 할당 과정 중 양자화 오차와 마스킹 임계치를 비교하여 잡음 제어를 하는 알고리즘은 다음과 같다^[7].

$$\begin{aligned}
 & \text{if } \sum_{i \in sb} \{ABS(xr(i)) - ix(i)\}^4 \times 2^{(global_gain - scalefactor(sb) - 210)/4} > Xmin(sb), \\
 & \text{then } scalefactor(sb) = scalefactor(sb) + 1 \\
 & \text{where, } xr(i) : i \text{ th MDCT coefficient, } ix(i) : i \text{ th quantized coefficient,} \\
 & \quad Xmin : \text{Masking Threshold, } ABS() : \text{absolute value function,} \\
 & \quad sb : \text{scalefactor band index}
 \end{aligned}
 \tag{1}$$

스케일 팩터 밴드 단위로 MDCT 계수의 양자화 오차를 구하여 마스킹 임계치와 비교하고, 양자화 오차가 더 클 경우에 대해 스케일 팩터를 증가시켜 양자화 오차가 귀에 들리지 않도록 양자화 잡음을 제어한다. 그러나 만약 비교 대상이 되는 두 항의 양쪽에 고정 소수점 연산에 의한 오차가 더해지게 되면 비교 결과가 달라지게 되고, 이러한 과정이 반복되면 결국 스케일 팩터는 전혀 다른 값을 갖게 된다. 결과적으로 위의 과정을 거쳐 잡음 제어를 성공적으로 수행했다 해도 양자화 잡음이 귀에 들릴 수가 있다. 따라서 비트 할당 결과가 부동 소수점 연산 결과와 최대한 같게 만들기 위해서는 충분한 데이터 워드 길이를 확보하여 고정 소수점 연산에 의한 오차를 최소한으로 줄여야 한다.

허용 가능한 고정 소수점 연산 오차의 양을 측정하기 위하여, 먼저 식 (1)의 조건문을 다음과 같은 연산 오차 모형으로 가정하였다.

$$\begin{aligned} & \text{if } \sum_{i \in sb} E_Q^2 > Xmin(sb) + E_{fix} \text{ then ...} \\ \Leftrightarrow & \text{if } K \cdot E_Q^2 > Xmin(sb) + E_{fix} \text{ then ...} \end{aligned} \quad (2)$$

왼쪽 항은 식 (1)에서의 양자화 오차 에너지를 나타내며, 오른쪽 항은 마스킹 임계치를 나타낸다. E_Q 는 양자화오차, K 는 sb 번째 스케일 팩터 밴드의 폭이며, $Xmin(sb)$ 는 마스킹 임계치, E_{fix} 는 마스킹 임계치를 구할 때 발생한 고정 소수점 연산 오차이다. 심리음향 모델의 연산 오차만을 고려하기 위해 양자화 오차 에너지를 계산할 때 발생하는 고정 소수점 연산 오차는 고려하지 않았다.

이때 식 (1)의 과정에서 고정 소수점 연산 오차로 인해 조건문의 판단이 바뀌지 않기 위해서는 $K \cdot E_Q^2$ 값과 $Xmin(sb)$ 의 차이가 오차 범위인 E_{fix} 값 범위 이내에 있지 않아야 한다. 만약 $K \cdot E_Q^2 > Xmin(sb)$ 인 경우를 가정할 때, 고정 소수점 연산 오차가 더해져서 $K \cdot E_Q^2 < Xmin(sb) + E_{fix}$ 로 부등호의 방향이 바뀌는 경우가 발생하면 스케일 팩터 값이 부동 소수점 연산 결과와 다른 값을 갖게 되고 양자화 잡음 제어가 부동 소수점 연산에서의 의도한 것과는 다르게 얻어지게 된다. 따라서 이러한 경우가 발생하지 않도록 E_{fix} 값이 충분히 작아야 한다. 적절한 고정 소수점 연산 오차의 양을 얻기 위해 본 논문에서는 다음과 같은 통계적 모형을 사용하였다.

$$Prob \left\{ \left| \sum E_Q^2 - Xmin \right| \leq |E_{fix}| \right\} \leq \alpha \quad (3)$$

위 식은 마스킹 임계치와의 차이가 E_{fix} 값보다 작은 확률이 임의의 값 보다 작은 것을 나타낸다. 만약 랜덤 변수 E_Q 가 각각 독립이고 균일하게 분포되어있으며 K 의 폭을 갖는 스케일 팩터 밴드 내의 랜덤 변수들의 합을 P_Q 라고 가정하면, central-limit 정리에 의해 다음과 같은 식을 얻을 수 있다^{[8][9]}.

$$\sqrt{K} \cdot \left(\frac{1}{K} P_Q - E[P_Q] \right) \xrightarrow{K \rightarrow \infty} N(0, \sigma_{P_Q}), \quad N: \text{normal distribution} \quad (4)$$

위의 식을 통해서 임의의 M 에 대해 아래와 같은 식을 얻을 수 있다.

$$Prob \left\{ \sqrt{K} \cdot \left(\frac{1}{K} P_Q - E[P_Q] \right) \leq M \right\} \xrightarrow{K \rightarrow \infty} \frac{1}{2} \left[1 + \operatorname{erf} \left(\frac{M}{\sqrt{2} \cdot \sigma_{P_Q}} \right) \right] \quad (5)$$

P_Q 와 $Xmin$ 의 차를 Q 라고 정의하고, K 값이 충분히 크면 다음과 같은 식을 얻을 수 있다.

$$Prob \left\{ \sqrt{K} \cdot \left(\frac{1}{K} Q - E[Q] \right) \leq M \right\} \cong \frac{1}{2} \left[1 + \operatorname{erf} \left(\frac{M}{\sqrt{2} \cdot \sigma_Q} \right) \right] \quad (6)$$

$Q = P_Q - Xmin$

$$Prob \left\{ Q \leq |E_{fix}| \right\} \cong \frac{1}{2} \left[1 + \operatorname{erf} \left(\frac{|E_{fix}| - K \cdot E[Q]}{\sqrt{2K} \cdot \sigma_Q} \right) \right] \leq \alpha \quad (7)$$

$$\begin{aligned} Prob \left\{ |Q| \leq |E_{fix}| \right\} &= Prob \left\{ Q \leq |E_{fix}| \right\} - Prob \left\{ Q < -|E_{fix}| \right\} \\ &\cong \frac{1}{2} \left[\operatorname{erf} \left(\frac{|E_{fix}| - K \cdot E[Q]}{\sqrt{2K} \cdot \sigma_Q} \right) - \operatorname{erf} \left(\frac{-|E_{fix}| - K \cdot E[Q]}{\sqrt{2K} \cdot \sigma_Q} \right) \right] \leq \alpha \end{aligned} \quad (8)$$

그런데 위의 식 (8)에서 부등호의 왼쪽 항은 $\operatorname{erf}()$ 함수의 특성상 식 (9)와 같은 성질을 발견할 수 있으므로, 다음 식 (10)과 같이 $|E_{fix}|$ 의 최대값에 관한 식을 얻을 수 있다.

$$\begin{aligned} & \operatorname{erf} \left(\frac{|E_{fix}| - K \cdot E[Q]}{\sqrt{2K} \cdot \sigma_Q} \right) - \operatorname{erf} \left(\frac{-|E_{fix}| - K \cdot E[Q]}{\sqrt{2K} \cdot \sigma_Q} \right) \\ & \geq 2 \cdot \left\{ \operatorname{erf} \left(\frac{|E_{fix}| + K \cdot E[Q]}{\sqrt{2K} \cdot \sigma_Q} \right) - \operatorname{erf} \left(\frac{K \cdot E[Q]}{\sqrt{2K} \cdot \sigma_Q} \right) \right\} \end{aligned} \quad (9)$$

$$|E_{fix}| \leq \sqrt{2K\sigma_Q} \cdot \operatorname{erfinv} \left(\alpha + \operatorname{erf} \left(\frac{K \cdot |E[Q]|}{\sqrt{2K\sigma_Q}} \right) \right) + K \cdot |E[Q]| \quad (10)$$

위의 식 (10)은 고정 소수점 연산 오차가 양자화 오차와 마스킹 임계치의 비교에 영향을 미칠 확률이 α 이하가 될 때의 고정 소수점 연산 오차의 최대값을 나타낸다. 밴드의 길이인 K 값이 충분히 크다는 가정이 있으므로 (식 (5),(6)), K 값이 큰 높은 주파수 쪽으로 갈수록 위의 모델은 좀더 정확한 값을 갖게 된다.

이 값을 사용하여 부호화에 적합한 데이터 워드 길이를 결정하기 위하여 실제의 오디오 신호에 대해 실험을 하였다. 다양한 장르의 음악 10곡에 대해서 실험을 수행하였으며, 각 밴드마다 양자화 오차가 많이 발생하는 채널 당 48 kbps 부호화에 대해 고정 소수점 연산 오차를 측정하였다. 심리음향 모델에서 발생하는 연산 오차의 영향을 나타내기 위해 심리음향 모델의 고정 소수점 연산 오차만을 사용하였다. 먼저 각 워드길이별 고정 소수점 연산 오차를 측정된 후 그 최대값을 식 (8)에 대입하여 연산 오차 비율 값의 최대값을 얻었다. 그림 1은 이렇게 얻은 연산 오차 비율 값을 나타낸 것이다.

값이 0.01라는 것은 비트 할당 과정이 고정 소수점 연산 오차의 영향을 받을 확률이 1% 이내라는 의미이다. 실험 결과로부터 고정 소수점 연산 오차의 영향을 받을 확률이 1% 내외 정도(가로 실선)가 되기 위해서는 대략 23 비트 이상의 데이터 워드 길이를 사용해야 한다는 결과를 얻을 수 있었다. 그림 2는 24비트에 대한 각 스케일 팩터 밴드별 최대 오차(실선)를 나타낸 것으로서, 식 (10)을 사용하여 값이 0.007일 때의 최대허용 오차(점선)를 함께 나타낸 것이다. 높은 주파수로 갈수록 실선과 점선의 차이가 줄어들고 있으므로 K 값이 큰 높은 주파수에서 통계 모델의 정확성이 높다는 것을 보여주고 있다.

이 통계 모델을 검증하기 위하여 실제로 부동 소수점 연산 결과와 고정 소수점 연산 결과를 비교하여 고정 소수점 연산 오차의 영향을 받는 경우는 전체의 몇 퍼센트인지를 측정하였다. 만약 부동 소수점 연산과 고정 소수점 연산을 각각 사용하여 식 (1)의 비교를 수행할 때 부동호의 방향이 서로 다르면 고정 소수점 연산 오차의 영향을 받는 것으로 기준을 정하였다. 오디오 부호화 실험에 많이 사용되는 Suzanne Vega의 "Tom's Diner"라는 곡의 샘플을 사용하여 각 워드 길이별 고정 소수점 연산에 대해 측정된 실험치는 표 1과 같다. 표의 결과는 통계적 모델이 실제 실험의 결과와 대체로 잘 맞다는 것을 보여주고 있다.

플을 사용하여 각 워드 길이별 고정 소수점 연산에 대해 측정된 실험치는 표 1과 같다. 표의 결과는 통계적 모델이 실제 실험의 결과와 대체로 잘 맞다는 것을 보여주고 있다.

III. 고정 소수점 연산을 통한 초월 함수의 구현

로그, 사인 함수 등의 사칙 연산으로 구할 수 없는 초월 함수들을 고정 소수점 연산으로 계산하는

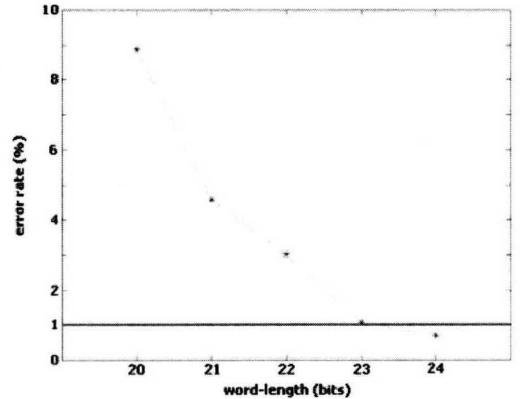


그림 1. 데이터 워드 길이 별 고정 소수점 연산 오차에 대한 비트할당 연산 오차 비율

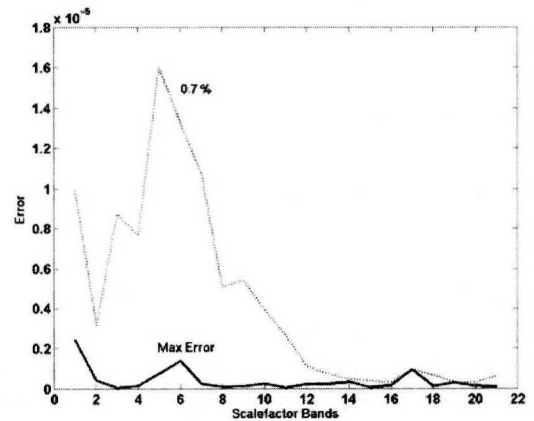


그림 2. 24 비트 워드 길이에 대한 통계 모델 (점선, $\alpha = 0.007$)과 실측 최대 오차 값(실선)의 비교

표 1. 워드 길이에 따른 고정 소수점 연산 오차 비율

bits	24	23	22	21	20
α	0.006	0.011	0.031	0.046	0.089
Rate	0.006	0.014	0.032	0.056	0.096

데에는 일반적으로 테이블을 사용한 방법^[10]과 테일

러 급수와 같은 몇 차의 다항식으로 근사화 하는 두가지 알고리즘이 많이 사용된다. 급수를 사용한 방법은 차수를 높일 경우 매우 높은 정확도를 얻을 수 있지만, 차수가 높아짐에 따라 연산량도 함께 급격히 증가하는 단점이 있다. 반면 테이블을 사용하여 근사화 하는 방법은 연산량을 적지만, 정확도를 높이기 위해 많은 테이블을 필요로 한다. MP3 부호화기에서 초월 함수가 많이 사용되는 과정은 연산량이 많은 심리음향 모델이므로 연산의 정확성과

면, 빈번하게 사용되는 초월 함수의 연산량이 늘어나는 것은 실시간 구현을 어렵게 하므로, 본 논문에서는 테이블을 사용한 근사법을 선택하였다.

그림 4는 테이블 크기에 따른 로그 함수 근사화의 신호 대 잡음비(SNR)를 나타낸 것이다. 알고리즘의 특성상 테이블 크기는 2의 거듭 제곱이어야 한다^[10]. 그림과 같이 테이블 크기가 증가함에 따라 신호 대 잡음비가 증가하지만, 그 증가율은 점점 줄어들고 있다. 초월 함수에서 로그 함수 외에도 지수

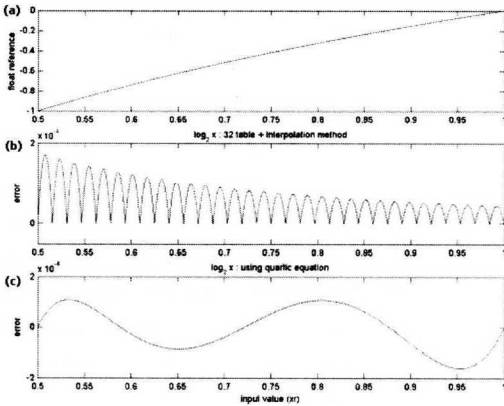


그림 3. 32개의 테이블을 사용한 알고리즘과 4차 함수를 사용한 알고리즘의 연산 오차 비교; (a) 참값 (b) 32개의 테이블을 사용한 방법의 오차 (c) 4차 급수를 사용한 방법의 오차

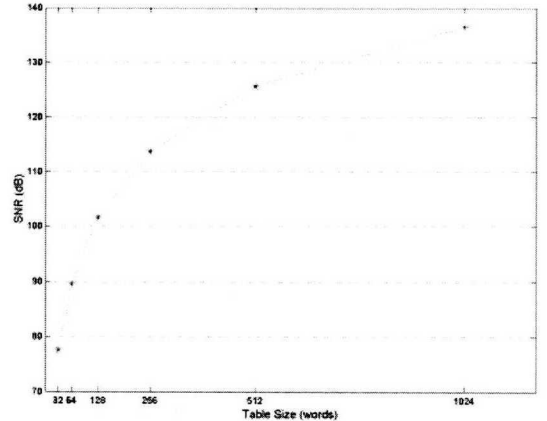


그림 4. 테이블 크기에 따른 로그 함수 근사화의 신호 대 잡음비

함께 연산량도 함께 고려되어야 한다.

그림 3은 위의 두 알고리즘을 사용하여 로그 함수를 근사화한 결과를 비교한 그림이다. 테이블을 사용한 방법은 미리 계산된 로그 테이블과 함께 선형 보간법을 사용하여 근사화했으며^[10], 급수를 사용한 방법은 로그 함수를 4차 함수로 근사화한 것이다. 그림 3(a)는 0.5~1.0사이의 x값에 대한 $\log_2(x)$ 의 참값을 나타낸 것이고, 3(b)와 3(c)는 각각 32개의 테이블을 사용한 방법과 4차의 급수를 이용한 방법에 대한 오차 값을 나타낸 것이다. 실험 결과 그림 3(b)와 같이 32개의 테이블을 사용할 때 4차 함수 알고리즘과 유사한 오차 특성을 얻을 수 있었다. 이 결과보다 더 높은 정확도를 얻기 위해서는 급수를 사용한 방법이 차수를 늘리기 위해 연산량을 더 높여야 하는데 비해서 테이블 알고리즘은 연산량과 관계없이 테이블 크기만을 늘려주면 된다. 테이블은 하드웨어 상에서 비교적 부담이 적은 ROM을 사용하기 때문에 테이블이 조금 더 늘어나는 것은 부호화기 구현에 큰 부담이 되지 않는 반

함수도 사용되므로, 테이블의 크기와 정확성을 함께 고려하여 신호 대 잡음비의 증가가 둔화되기 시작하는 256 크기의 테이블을 선택하였다

IV. 음질 평가 결과

본 논문은 앞서 설명한 통계 모델에 의한 결과로부터 고정 소수점 연산 오차가 비트 할당 과정에 미치는 영향의 적정선을 1%로 정하고, 24비트의 데이터 워드 길이를 갖는 DSP 구조를 선택하였다. 24비트의 데이터 워드 길이를 사용한 고정 소수점 연산 최적화를 통하여 부호화된 MP3 비트열의 음질을 평가하기 위하여 ITU-R 권고안 BS.1378에 정의된 ODG (Object Difference Grade)라는 객관적 지표를 사용하였다^[11]. 이 방법은 원음과의 음질 차이를 0부터 4 사이의 ODG 값으로 나타내는 것이다.

본 실험에서는 고정 소수점 연산을 사용하여 부호화한 후 부동 소수점 연산으로 복호화한 음원의 ODG 값과 부동 소수점 연산을 사용하여 부호화한

후 부호화 한 음원의 ODG 값과의 차이를 평가의 지표로 정하였다. 두 값의 차이가 0 이면 청각적으로 고정 소수점 연산이 음질 열화가 없는 경우이며, 값이 작아질수록 음질이 낮아지는 것을 의미한다. 객관적 음질 평가를 위해 사용된 음원은 5~20초 정도의 짧은 음악 샘플들이며, 장르와 사용된 악기의 특징 등을 고려하여 선택된 23개의 스테레오 샘플들을 사용하였다. 각 음원은 채널 당 64 kbps, 즉 128 kbps의 스테레오 비트열로 부호화되었다.

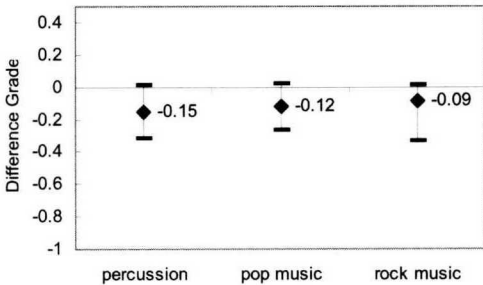


그림 5. 24 비트 고정 소수점 연산의 음질 평가 결과

그림 5는 음질 평가 결과를 나타낸 것이다. 실험에 사용된 샘플들을 그림과 같이 크게 타악기, 팝 음악, 락 음악으로 나누어 음질의 분포를 나타내었다. 실험 결과 사용된 샘플들의 특성이나 장르 등에 크게 영향을 받지 않고 있으며, 부동 소수점 연산 결과와 고정 소수점 연산 결과 사이의 음질 차이는 최대 -0.5를, 평균적인 음질 저하는 -0.2를 넘지 않는 것으로 나타났다. 따라서 고정 소수점 연산에 의한 음질 열화가 구분하기 힘든 수준임을 알 수 있다.

V. 결론

본 논문에서는 MPEG 오디오 Layer-III 부호화 과정 중 심리음향 모델 과정을 하드웨어 모듈로 ASIC 구현을 하기에 앞서, 고정 소수점 연산기로 구현하기에 적합하도록 최적화를 수행하였다. 심리음향 모델 과정은 오디오 부호화 과정 중 양자화의 기준이 되는 마스크 임계치를 구하는 과정으로서, 많은 연산량과 높은 정확도를 요구하는 과정이다.

최적화 과정을 통해 심리음향 모델에서 자주 사용되는 초월함수를 구현하기에 적절한 크기의 테이블 크기를 결정하였으며, 고정 소수점 연산의 데이

터 워드 길이가 비트 할당 과정에 미치는 영향을 통계적으로 모델링 하였다. 여기서 구한 통계적 모델로부터 24비트의 데이터 워드 길이를 결정하였으며, 24비트를 사용할 경우 고정 소수점 연산 오차가 비트 할당에 영향을 미칠 확률이 1% 이하가 됨을 확인하였다. 이 통계적 모델은 ASIC 구현 뿐만 아니라 범용 DSP를 사용하여 부호화기를 구현할 때 사용할 DSP를 선택하는 객관적인 기준으로 활용될 수 있다. 또한 최적화 과정이 수행된 고정 소수점 심리음향 모델을 사용한 MP3 부호화기의 음질은 부동 소수점 부호화기에 비해 ITU-R의 음질평가 점수를 기준으로 평균 -0.2 이내의 구분하기 힘든 수준의 음질 저하를 보였다.

참고 문헌

- [1] Hyen-O Oh, Joon-Seok Kim, Chang-Jun Song, Young-Cheol Park and Dae-Hee Youn, "Low Power MPEG/AUDIO Encoders Using Simplified Psychoacoustic Model and Fast Bit Allocation," *IEEE Trans. On Consumer Electronics*, pp 613-621, 2001.
- [2] Nikil Jayant, et al., "Signal Compression Based on Models of Human Perceptual Noise Criteria," *IEEE Journal of Selected Areas on Communication*, 6(2), Feb. 1988
- [3] Yuichiro Takamizawa, Tochiyuki Nomura, and Masao Ikekawa, "High-Quality and Processor-Efficient Implementation of an MPEG-2 AAC Encoder," *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing*, Vol.2, pp.985-988, May 2001
- [4] Antonio Servetti, Alessandro Rinotti, Juan Carlos De Marin, "Fast Implementation of the MPEG-4 AAC Main and Low Complexity Decoder," *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing*, Vol.5, pp.249-252, 2004
- [5] Dong-Yan Huang, Xuesong Gong, Daqing Zhou, Toshio Miki, and Sanae Hotani, "Implementation of the MPEG-4 Advanced Audio Coding Encoder on ADSP-21060 SHARC," *Proc. IEEE International Symposium on Circuits and Systems*, Vol.3, May 1999

- [6] Tadashi Sakamoto, Maiko Taruki, and Tomohiro Hase, "A Fast MPEG-Audio Layer-III Algorithm for a 32-Bit MCU," *IEEE Trans. On Consumer Electronics*, Vol.45, No.3, Aug 1999.
- [7] ISO/IEC 11172-3, *Coding of Moving Pictures and Associated Audio for Digital Storage Media at up to about 1.5 Mbit/s (Part 3: MPEG-Audio)*
- [8] Peyton Z. Peebles, Jr., *Probability, Random Variables, and Random Signal Principles*, McGraw-Hill, 1993
- [9] Oliver Derrien, Maurice Charbit, Pierre Duhamel, "Statistical Model for the Quantization Noise in the MPEG Advanced Audio Coder Application to the Bit Allocation Algorithm," *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing*, Vol. 2, pp. 1849-1852, 2002
- [10] Keun-Sup Lee, Hyen-Oh Oh, Young-Cheol Park, Dae Hee Youn, "High Quality MPEG-Audio Layer III Algorithm for a 16-bit DSP," *Proc. IEEE International Symposium on Circuit and Systems*, Vol. II, pp.205-208, May 2001
- [11] ITU Radiocommunication Study Group 6, *Draft Revision to Recommendation ITU-R BS.1387 - Method for objective measurements of perceived audio quality*

이 근 설(Keun-Sup Lee)

정회원

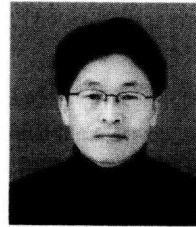


1997년 2월 : 연세대학교 전자공학과 졸업
 1999년 8월 : 연세대학교 전기컴퓨터공학과 석사
 1999년 9월~현재 : 연세대학교 전기전자공학과 박사과정

<관심분야> 디지털 신호처리, 오디오 신호처리,

박 영 철(Young-Cheol Park)

정회원

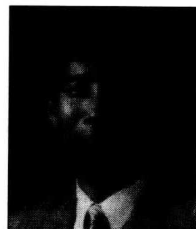


1986년 2월 : 연세대학교 전자공학과 졸업
 1988년 2월 : 연세대학교 전자공학과 석사
 1993년 2월 : 연세대학교 전자공학과 박사
 현재 : 연세대학교 정보기술학부 교수

<관심분야> 디지털 신호처리, 오디오 신호처리, 음성 신호처리, 적응 신호처리

윤 대 희(Dae Hee Youn)

정회원



1977년 2월 : 연세대학교 전자공학과 졸업
 1979년 8월 : Kansas State Univ. 석사
 1982년 8월 : Kansas State Univ. 박사
 현재 : 연세대학교 전기전자공학과 교수

<관심분야> 디지털 신호처리, 적응 신호처리, 음성 신호처리, 오디오 신호처리