

MPEG-4 AAC를 위한 고속 양자화기

황진용, 김병일, 장태규
 중앙대학교 신호 및 통신 시스템 연구실

A new quantization method for the realization of the low complexity MPEG-4 AAC encoder

Jin-Yong Hwang, Byoung-Il Kim and Tae-Gyu Chang
 Chung-ang University, Signals and communications LAB.

Abstract - 본 논문에서는 MPEG-4 AAC 표준에 근거하여 개인휴대단말과 같은 모바일 단말의 용용을 위해 연산 복잡도를 낮춘 새로운 양자화기 구조를 제안하였다. 기존 MPEG-4 AAC 인코더의 양자화기는 rate control 과 distortion control을 동시에 수행하는 순환 이중루프 구조를 갖으며, 높은 연산부담으로 인해 실시간 인코딩 및 시스템의 소비전력을 증가시키는 중요한 요인으로 작용한다. 본 논문에서는 이러한 제약조건을 극복하기 위해 허용 잡음 레벨과 대역별 양자화 잡음을 정합하는 잡음 포락선 정합기법을 제안하였고, 이를 응용하여 낮은 소비전력 실현이 가능한 단일루프 구조의 단순화된 양자화기 구조를 제안하였다. 또한 제안한 양자화기를 적용한 독자적인 구조의 MPEG-4 AAC 인코더를 구현하여 검증 및 성능평가를 수행하였다.

1. 서 론

디지털 신호처리 기술의 발전과 하드웨어 성능의 비약적인 향상에 따라, 다양한 형태의 오디오 멀티미디어 서비스가 생성 제공되고 있으며, PC 및 DSP 기반의 Storage 형태 서비스에 국한되지 않고 PDA, 스마트폰, DMB, 컴팩트 디지털 오디오 장치와 같은 개인휴대단말로 영역이 점차 확대되고 있다. 현재 개인휴대단말 분야의 용용에서는 MPEG-1 Layer-III(MP3)가 시장의 대부분을 차지하고 있다. 하지만 스마트폰과 DMB와 같은 무선환경 기반의 스트리밍 서비스에서는 제한된 대역폭내에 유연한 코딩 기술이 필요하다. MPEG-4 AAC는 기존 MP3에 비해 낮은 bitrate에서의 코딩효율이 뛰어나고 제한된 밴드폭내에서 유연한 코딩이 가능하다.[1] 그러나 MPEG-4 AAC 압축 알고리즘의 개인휴대단말 용용에서는 높은 연산부담으로 인해 실시간 인코딩 및 전력소비 문제가 여전히 해결 되어야 할 과제로 남아있다.

MPEG-4 AAC 인코더는 여러 가지 기능 블록으로 구성되어 있다. 이중 양자화기는 전체 인코더 연산 복잡도의 60% (64Kbps, LC profile) 이상을 차지한다.[1] 표준 MPEG-4 AAC 인코더의 양자화기는 목표 bitrate를 만족시키기 위해 rate control을 하는 Inner loop와 대역별 양자화 잡음에 의한 왜곡을 허용잡음수준(Masking threshold) 이하로 분포시키기 위한 Outer loop의 순환 이중루프 구조로 되어 있다. 순환 이중루프 구조로 인해 발생하는 연산 복잡도를 낮추기 위해서는 새로운 양자화기 구조가 필요하다.

본 논문에서는 기존 MPEG-4 AAC의 양자화기를 대체하기 위해 잡음 포락선 정합 알고리즘 및 이에 기초한 낮은 연산 부담을 갖는 새로운 양자화기를 제안하였다. 잡음 포락선 정합 알고리즘은 대역별 양자화 잡음을 허용잡음레벨 포락선과 일치하도록 정합하는 기법이다. 제안한 알고리즘을 적용한 양자화기는 대역별 양자화 잡음에 의한 distortion control loop를 제거함으로써 기존 양자화기의 연산 부담을 줄일 수 있다. 최종적으로 잡음 포락선 정합 기법에 기초한 양자화기를 적용한 MPEG-4 AAC 인코더를 개발하였으며, 기존 Reference 코더에 비해 50% 이상의 성능향상을 얻을 수 있었다.

2. 본 론

2.1 MPEG-4 AAC 양자화기 구조 및 원리

MPEG-4 AAC 인코더에 적용되는 양자화는 식 1과 같다.[2]

$$X_{quant} = Round(|mdct_coeff|^{\frac{3}{4}} \times Gain)$$

$$= Round(|mdct_coeff|^{\frac{3}{4}} \times 2^{-\frac{3}{16}(common_sf - sf[sb])})$$

(1)

where, sb: the number of scalerfactor band

식에서 보듯이 MPEG-4 AAC 인코더에 적용되는 양자화는 주파수 계수들의 비선형 compression을 수행하는 $(X(k))^{\frac{3}{4}}$ 부분과 양자화 잡음과 coding에 사용되는 비트수를 scalerfactor band별로 조절하기 위한 gain factor로 구성된다.[3] gain factor는 coding에 사용되는 bit수를 조절하기 위한 common scalerfactor(common_sf)와 band별 왜곡을 허용잡음수준 이내로 조절하기 위한 scalerfactor(sf[sb])에 의해 결정이 된다. MPEG-4 AAC에서는 이 두가지 제약조건을 만족시키기 위해 순환 이중루프 구조에 의해 common scalerfactor와 scalerfactor band별 scalerfactor값을 결정한다. 본 논문에서 제안한 잡음 포락선 정합 알고리즘은 band별 양자화 잡음이 허용잡음

레벨 포락선과 일치하는 gain값과 scalerfactor값을 수식에 의해 직접 산출할 수 있게 한다.

2.2 잡음 포락선 정합 알고리즘

양자화 잡음 power는 식 2와 같이 양자화 스텝크기에 의한 확률적 기대치로 얻을 수 있다.[4]

$$q_e^2 = \frac{\Delta^2}{12}$$

(2)

식 1에서 scalerfactor band내의 k번째 주파수 계수의 양자화 스텝 크기는 식 3과 같이 compression에 의해 발생하는 양자화 스텝크기와 해당 scalerfactor band의 gain을 곱함으로써 얻을 수 있다.

$$\Delta_T(k) = \Delta_{3/4}(k) \times Gain$$

(3)

$$where, Gain = 2^{\frac{3}{16}(common_sf - sf[sb])} = 2^{\frac{3}{16}\gamma(sb)}$$

, k : index of mdct coefficient

$\Delta_T(k)$ 에 의해 발생하는 양자화 잡음 power $Q^2(k)$ 는 선형 양자화로 해석할 수 있고 식 4와 같다.

$$Q^2(k) = \frac{\Delta_T^2(k)}{12} = \frac{(\Delta_{3/4}(k) \times Gain)^2}{12}$$

(4)

따라서 scalerfactor band별 양자화 잡음 power $P_Q(s)$ 식 5와 같이 나타낼 수 있다.

$$P_Q(sb) = \sum_{k=\min(sb)}^{\max(sb)} Q^2(k) = \sum_{k=\min(s)}^{\max(s)} \frac{\Delta_T^2(k)}{12}$$

$$= \frac{(\Delta_{3/4}(k) \times Gain)^2}{12}$$

scalerfactor band별 허용잡음레벨을 $T(sb)$ 라 할때 양자화 잡음의 왜곡조건을 만족시키기 위해서는 $P_Q(sb) \leq T(sb)$ 이어야 하며, 이 조건을 만족시키는 대역별 gain 값은 식 6과 같이 얻을 수 있다.

$$\sum_{k=\min(sb)}^{\max(sb)} \frac{\Delta_{3/4}^2(k) Gain^2}{12} \leq T(sb)$$

$$\Rightarrow Gain^2 \leq \frac{12 \times T(s)}{\sum_{k=\min(s)}^{\max(s)} \Delta_{3/4}^2}$$

(6)

이므로 최종적으로 대역별 scalerfactor $\gamma(s)$ 는 식 7의 조건을 만족시킨다.

$$\gamma(s) \leq \frac{8}{3} \log_2 \left(\frac{12 \times T(s)}{\sum_{k=\min(s)}^{\max(s)} \Delta_{3/4}^2} \right)$$

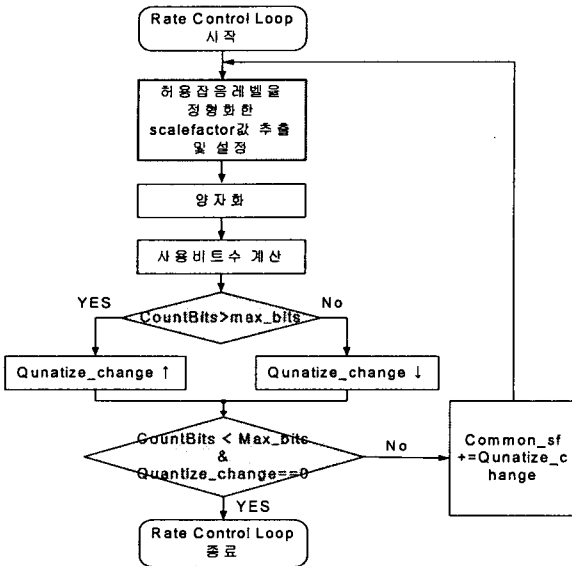
(7)

식 7에 의해 산출된 scalerfactor값의 최대값을 적용한 대역별 양자화 잡음 power $P_Q(sb)$ 는 허용잡음레벨의 포락선을 정합하게 된다.

2.3 잡음 포락선 정합 알고리즘 적용 양자화기 구조 및 원리

식 7은 MPEG-4 AAC에 적용되는 비선형 양자화식으로부터 scalerfactor band별 허용잡음 포락선을 정합할 수 있도록 양자화 잡음을 분포시키는 scalerfactor값을 직접 산출할 수 있음을 수식적으로 증명하였다. 하지만 산출된 scalerfactor값은 인코딩 과정에서 요구하는 bitrate를 만족시키지 못한다. bitrate를 맞추기 위해 대역별 scalerfactor값을 일괄적으로 상향 또는 하향 조절을 하여도 대역별로는 여전히 허용잡음레벨을 정형화 하고 있기 때문에 대역별 왜곡을 조정해주는 추가적인 노력이 불필요하게 된다. 그림 1은 본 논문에서 제안한 잡음 포락선 정합 알고리즘을 적용한 단일루프 구조의 양자화기를 나타내고 있다. 제안한 양자화기는 MPEG-4 AAC의 rate control loop와 비슷한 구조를 가진다. 다만 대역별 허용잡음레벨 포락선을

정합하도록 양자화 잡음을 분포시키는 scalefactor값을 산출하는 블록이 추가 되었다.



〈그림 1〉 잡음 포락선 정합 기법 적용 양자화기 구조

2.4 구현 및 성능 평가

본 논문에서 제안한 잡음 포락선 정합 알고리즘 및 이를 적용한 양자화기의 성능평가를 위해 ISO/IEC MPEG-4 AAC 표준에 근거하여 양자화기를 재구성한 MPEG-4 AAC 인코더를 개발하였다. 고성능 오디오 압축 구조의 성능 평가를 위하여 44.1KHz의 샘플율을 갖는 CD 오디오 신호를 추출하고 MPEG-4 AAC Reference 인코더와 성능을 비교 평가하였다.

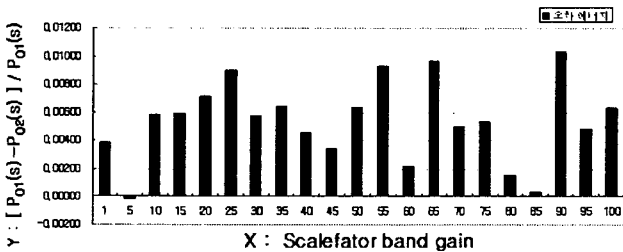
제안한 잡음 포락선 정합 기법에 대한 이론적 검증은 위해 1차적으로 Matlab simulation을 통해 측정값과의 오차를 비교 하였다. 측정값 $P_{Q1}(sb)$ 은 식 8과 같이 입력샘플 $X(K)$ 와 양자화 및 역양자화를 과정을 거쳐 estimation된 $\hat{X}(K)$ 와의 차에 의해 발생하는 양자화 잡음 Power이다.

$$P_{Q1}(sb) = \sum_{k=\min(sb)}^{\max(sb)} [\hat{X}(K) - X(K)]^2 \quad (8)$$

제안한 잡음 포락선 정합 기법에 의한 양자화 잡음 Power $P_{Q2}(sb)$ 은 식 5를 적용하였다. 오차에 대한 환산은 식 9와 같이 측정값에 대하여 잡음 포락선 정합 기법에 의해 산출된 양자화 잡음의 오차로 나타내었다.

$$\text{error} = \frac{P_{Q1}(sb) - P_{Q2}(sb)}{P_{Q1}(sb)} \quad (9)$$

그림 2는 10000개의 입력 샘플 $X(K)$ 를 $0 \leq X(K) \leq 2^{24}$ 범위 내에서 random하게 발생시키고 gain factor를 1에서 100까지 증가 시키면서 식 9와 같은 잡음 포락선 정합에 의해 산출된 양자화 잡음 오차에 대한 simulation 결과를 보여 주고 있다. simulation 결과 식 5에 의해 산출된 양자화 잡음 power는 측정된 실제 양자화 잡음 power와 차이가 없음을 보여 주고 있으며 잡음 포락선 정합 알고리즘이 이론적으로 이상 없음을 반증하고 있다.



〈그림 2〉 잡음 포락선 정합에 의한 양자화 잡음 산출 오차

잡음 포락선 정합 기법을 적용한 개발 인코더의 수행속도 평가를 위해 표 1과 같이 reference 인코더와 비교 하였다. 개발된 인코더는 reference 인코더에 비해 128 Kbps에서 55%, 64 Kbps 77% 정도 수행속도 개선이 되었음을 확인할 수 있다. 연산 복잡도 및 프로세싱에 따른 소비전력도 수행속도와 비례하여 개선효과를 가져올 것으로 판단된다. 또한 낮은 bitrate에서 개선효과가 더욱 크게 나타났다. 이는 reference 인코더의 경우 낮은 bitrate에서는 허용잡음수준 이내로 양자화 잡음을 분포 시키기 어렵기 때문에 더 많은 횟수의 순환 이중루프 반복 수행을 요구하게 된다.

〈표 1〉 encoding 수행시간 비교(Reference and proposed encoder)

* Pentium-4 3.0 GHz									
Song (44.1KHz)	Play time	128 Kbps				64 Kbps			
		Reference encoder		Proposed encoder		Reference encoder		Proposed encoder	
		Time	ratio	Time	ratio	Time	ratio	Time	ratio
Blue christmas	4:06	2:47	0.68	1:57	0.22	3:57	0.96	1:47	0.21
I'll take care of you	3:40	3:12	0.87	1:44	0.47	6:19	1.72	1:33	0.43
Kiss me	3:39	2:37	0.72	1:42	0.47	6:22	1.74	1:38	0.45
Suite_No1_BW1007	5:25	4:46	0.82	1:12	0.22	5:08	0.95	1:10	0.22
The voice	4:23	3:59	0.90	2:11	0.50	8:23	1.91	2:01	0.46
Mean	4:14	3:24	0.80	1:32	0.36	6:02	1.42	1:27	0.34

[ratio = encoding time / play time]

3. 결 론

본 논문에서는 MPEG-4 AAC 인코더의 실시간 인코딩 및 저전력 실현을 목적으로 잡음 포락선 정합 알고리즘 및 이를 적용한 새로운 구조의 양자화기를 제안하였다. 또한 실제 응용을 목적으로 제안된 양자화기를 적용한 MPEG-4 AAC 인코더를 개발하였다. 제안된 잡음 포락선 정합 알고리즘은 Matlab simulation을 통해 허용잡음레벨의 포락선과 일치하는 scalefactor band별 양자화 잡음을 산출할 수 있음을 확인하였고 이론적 검증을 통해 유도 과정에 오류가 없음을 확인하였다. 또한 제안한 알고리즘 및 양자화기를 적용한 개발 인코더의 연산 복잡도 및 소비 전력 평가를 위해 reference 인코더와의 수행속도를 비교 하였다. 개발된 MPEG-4 AAC 인코더는 reference 인코더 대비 128 Kbps에서 55%, 64 Kbps에서 77% 개선된 성능을 나타내었으며 소비전력 또한 processing 시간에 비례하므로 이와 비슷한 수준의 개선효과를 기대할 수 있다.

[참 고 문 헌]

- [1] S.W. Huang, T.H. Tsa, L.G. Chen, "A low complexity design of psycho-acoustic model for MPEG-2/4 advanced audio coding", Consumer Electronics, IEEE Transactions on Volume 50, Issue 4, Nov. 2004 Page(s):1209 - 1217
- [2] ISO/IEC JTC1/SC29/WG11 "ISO/IEC 14496-3, General Audio (GA) Coding: AAC/TwinVQ," ISO/IEC, May 1998
- [3] Derrien, O.; Charbit, M.; Duhamel, P.; "Statistical model for the quantization noise in the MPEG advanced audio coder. Application to the bit allocation algorithm" Acoustics, Speech, and Signal Processing, 2002. Proceedings. (ICASSP '02). IEEE International Conference on Volume 2, 2002 Page(s):1849 - 1852
- [4] K. Saywood, Introduction to Data Compression, 2nd ed. Morgan Kaufmann, 2000.