

낮은 연산 부담을 갖는 MPEG-4 AAC 인코더 개발에 관한 연구

김병일, 김동환, 장태규, *장홍엽

중앙대학교 전자전기공학부, *㈜삼성전자
전화 : 02-820-5318 / FAX : 02-812-1293

Development of an MPEG-4 AAC encoder of low implementation complexity

Byoung-Il Kim, Dong-Hwan Kim, Tae-Gyu Chang and Heung-Yub Chang

School of Electrical and Electronics Engineering, Chung-Ang University
E-mail : tgchang@cau.ac.kr

Abstract

This paper presents a new structure of MPEG-4 AAC encoder. The proposed encoder directly shapes quantization noise distribution according to the energy distribution curve and thereafter performs adjustment of the offset level of the noise distribution to meet the given bit rate. The direct noise shaping and the bit rate matching scheme of the proposed encoder algorithm significantly alleviate the problem of conventional encoder's processing burden related with the employment of the precise psychoacoustic model and iteration intensive quantizer. The encoder algorithm is implemented on ARM processor with fixed-point arithmetic operations. The audio quality of the implemented system is observed comparable to those of commercially available encoders, while the complexity of the implementation is drastically reduced in comparison to the conventional encoder systems.

I. 서론

제한된 채널 밴드 폭을 통해 멀티미디어 데이터의 효율적인 전송을 위해 다양한 오디오 코더가 개발되어 왔으며, 향상된 정보 전송을 위한 인프라의 급속한

발전이 이루어지는 중에도 효율적인 압축 코더의 개발의 필요성이 계속 요구되었다. 이는 제한된 대역폭을 갖는 네트워크를 통해 고품질 멀티미디어 정보를 사용하고자 하는 요구가 급속도로 늘어나기 때문이다[1][2].

고품질 오디오 인코더에서는 Transform 코딩 기법과 심리음향 모델이 적용된 양자화기를 채택하고 있다.[3][4] 정밀한 심리음향 모델과 정교한 양자화기 알고리즘은 오디오의 고품질 압축 재생을 실현시키나 이의 구현에 따른 높은 계산 부담이 요구된다. 특히, 모바일 환경 등의 적용에서 높은 계산 부담은 중대한 문제점이 된다.

본 논문에서는 모바일 단말에서의 응용을 위해 복잡도를 줄인 새로운 MPEG-4 AAC(Advanced Audio Coding) 인코더를 개발하였다. 제안된 알고리즘에서는 정교한 심리음향 모델과 양자화기 블록 대신에 direct noise shaping 과 bit rate matching 구조로 계산 부담을 줄일 수 있다.

제안된 알고리즘에서는 DCT(Discrete Cosine Transform)과정으로부터 대역별 에너지 분포를 얻어내고 이를 통해 직접 양자화 잡음 분포를 조절하고 나서 주어진 bit rate 에 맞는 잡음 분포의 임계 레벨의 조절을 실행한다. 또한 LC(Low Complexity) profile 을 대상으로 하는 MPEG-4 AAC 인코더를 개발하고 이의

고정 소수점 연산 구조를 도출하여 ARM 프로세서 상에 구현하였다. 구현된 시스템의 오디오 성능은 기존의 인코더와 비교해서 구현의 복잡도는 확실히 줄인데 반해서, 음질 성능은 상용 인코더들과 대등함을 확인하였다.

II. MPEG-4 AAC 인코더의 구조

MPEG-4 AAC 인코더의 일반적 구조는 그림 1 과 같이 filterbank, spectral processing, 심리음향 모델, quantizaion/noiseless coding 의 네 부분의 중요한 기능 파트로 나뉜다.

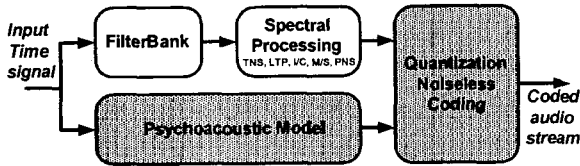


그림 1. MPEG-4 AAC 인코더

먼저 filterbank 에서는 DCT 를 기반으로 시간 영역의 오디오 신호를 주파수 계수로 변환된다. 압축 효율을 높이기 위해 Temporal Noise Shaping, Long Term Prediction, Intensity stereo, Mid/Side stereo, Perceptual Noise Substitution 등과 같은 다양한 형태의 spectral data processing 이 행해진다. 심리음향 모델은 인간의 청각특성을 이용하기 위한 임계 마스킹 레벨을 산출하는 것이다. 계산된 masking thresholds 는 양자화기에서 주어진 비트율에서 최고의 오디오 성능을 얻기 위해 사용된다. 일반적으로 청각 특성을 반영한 Spreading function 은 FFT 계수를 사용하여 masking threshold 를 산출한다. 양자화기는 이중 루프 구조를 이용하는데, 이들 중 inner 루프는 목표 비트율을 조절하며, outer 루프는 양자화 잡음을 마스킹 레벨보다 작도록 조절하여 대역별 잡음의 분포를 제어한다. 무손실 압축을 제공하는 Huffman 코딩은 inner 루프 안에서 실행된다.

많은 반복 수행이 요구되는 이중루프 구조의 양자화기는 인코딩 과정에서 큰 처리부담을 가지고 있는 블록들 중 하나이다. 게다가 기존의 양자화

알고리즘의 무한 루프수행에 따른 강제 종료 조건으로 인해 임의의 대역에서 양자화 잡음이 과도하게 분포될 가능성이 있으며, 특히 낮은 비트율에서는 더욱 이러한 가능성이 증가하여 오디오 성능의 심각한 열화를 야기할 수 있다. 기존의 인코더에서 DCT 와 DFT 모듈을 사용하는 것은 프로세서의 계산 부담을 증가시키는 주요 원인이었다. 다음 장에서는 위에서 말한 계산 부담과 음질 열화를 발생 시키는 문제점들을 해결하기 위한 새로운 인코더 구조를 기술하겠다.

III. 제안된 오디오 인코더 알고리즘

본 논문에서는 계산 부담이 향상된 심리음향 모델 및 낮은 비트율 대상이나 모바일 단말 응용에 적합한 양자화기를 적용한 새로운 오디오 인코더 알고리즘을 제안하였다. 제안된 양자화 알고리즘은 그림 2 에 나타난 것과 같이 direct noise shaping 과정과 level matching 과정의 연속적인 동작으로 구성된다.

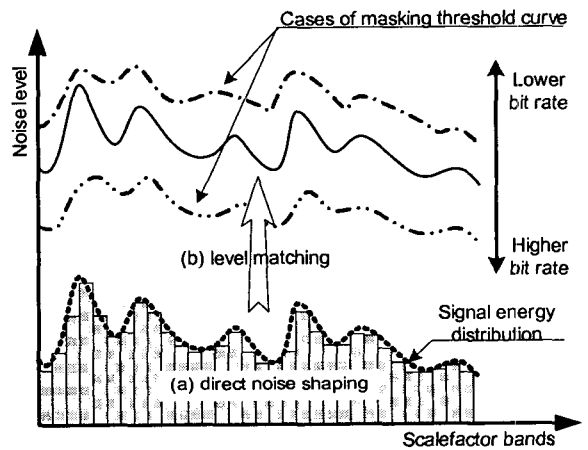


그림 2. direct noise shaping 과 level matching 에 기반한 제안된 양자화기 알고리즘

Direct noise shaping 은 DCT 로부터 산출된 데이터 블록내의 신호 에너지 분포를 통하여 scalefactor 대역별로 상대적 이득을 조절하는 과정이다. Level matching 은 사용된 비트율을 목표 비트율 아래로 떨어질 때까지 양자화 잡음 레벨을 증가시키기 위해

common gain 을 조절하는 과정을 나타낸다. 이러한 과정에서 최상의 품질을 보장하는 오디오 압축은 목표 비트율 내에서 최대의 비트를 사용할 때에 이루어진다. 이것은 양자화 동안에 잡음의 분포 형태를 먼저 이루고, 목표 비트율을 맞추는 동안 이 형태를 유지하는 과정으로부터 오는 특성이다. 또한 제안된 알고리즘은 전 대역에 걸친 양자화 잡음이 마스킹 레벨보다 작게 분포시킬 수는 없더라도 이보다 초과된 양자화 잡음의 분포를 고르게 분포시키는 이점을 제공하여 특정대역에 양자화 잡음의 편중을 통한 심각한 음질 열화를 줄일 수 있다.

IV. 구현 및 결과

본 논문에서는 개발된 오디오 인코더의 성능을 MP3 와 MPEG-2 AAC 를 포함한 기존의 인코더와 비교하였다. 제안한 인코더는 32 bit 고정 소수점 연산구조를 사용한 ARM 프로세서에 구현되었다. 128kbps 와 32kbps 에 대해 제안한 인코더와 MPEG 표준 인코더에서 측정한 양자화기 이중루프의 반복횟수를 그림 3 과 4 에 나타내었다. 그림과 같이 기존의 양자화기는 평균 반복횟수가 제안한 양자화기보다 세 배 더 많이 필요로 하는 것을 쉽게 관찰할 수 있다.

특히, 그림 4 와 같이 저비트율에서 기존 인코더의 양자화기 반복 횟수의 분포가 거의 일정하게 나타나는 것은 양자화 잡음이 마스킹 임계 레벨 아래로 조절하지 못할 때 발생하는 비정상 종료 조건에 의한 현상이다. MPEG reference 인코더와는 대조적으로 제안된 알고리즘은 평균 5 회의 반복 수행을 하는 것을 확인할 수 있으며 이를 통하여 제안한 알고리즘의 효율성을 확인하였다.

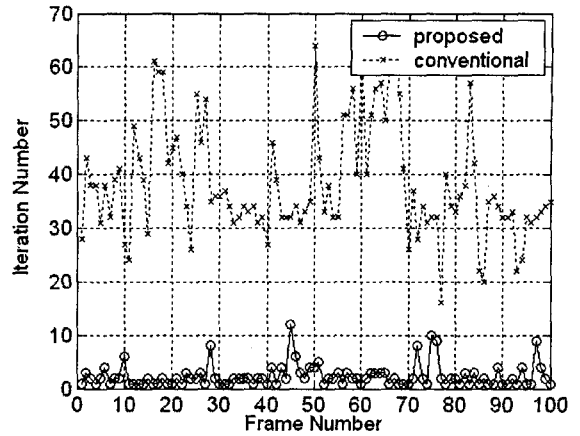


그림 3. 양자화 루프의 반복횟수 측정 (128Kbps)

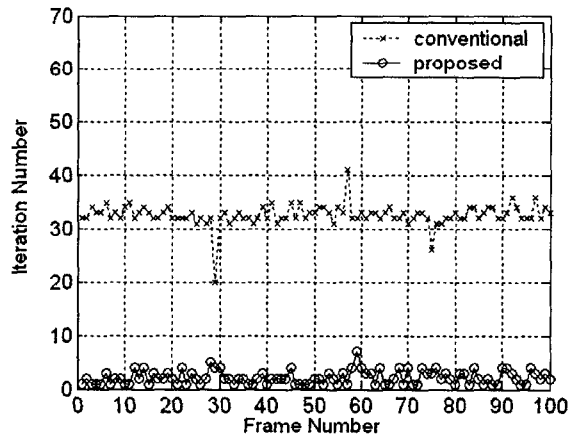


그림 4. 양자화 루프의 반복 횟수 측정 (32Kbps)

MPEG-4 AAC 인코더를 통해 오디오 정보를 인코딩하기 위한 총 수행 시간을 표 1 에 나타내었다. 제안한 알고리즘에서는 기존의 방식에 비해 128 kbps 에서는 단지 10%, 32kbps 에서는 15%의 연산량을 필요로 하였다.

표 1. 인코딩 수행시간 시간[ms]

	Proposed		Conventional	
	128Kbps	32Kbps	128Kbps	32Kbps
1	4949.177	1904.408	47725.316	13865.439
2	4623.372	2045.028	46147.410	13800.656
3	4606.716	1993.972	45649.009	13744.582
4	4653.765	1946.889	45879.296	13765.424
5	4507.633	1657.288	45477.217	13671.095
Avg	4,668.13	1,909.51	46,175.64	13,769.43

제안한 인코더는 32-bit 고정 소수점 연산 구조로 ARM 프로세서에서 개발되고 구현되었다. 다양한 비트율에서 측정한 MIPS 를 표 2 에 나타내었다.

표 2. ARM 프로세서의 MIPS 측정

Cycle Block	32Kbps		128Kbps	
	MIPS	Core Cycle	MIPS	Core Cycle
File I/O	0.864	1.132	2.479	3.249
FilterBank	4.854	6.135	14.056	17.789
TNS	7.859	10.232	12.612	16.398
MS	0.512	0.687	1.469	1.934
Quantization/Huffman	9.094	11.917	32.765	42.588
Bitstream Formatter	0.614	1.016	2.231	3.686
Total	23.80	31.12	65.61	85.64

V. 결론

본 논문에서는 간소화된 심리음향 모델과 양자화 알고리즘을 적용해서 개발한 MPEG-4 AAC 인코더를 제안하였다. 개발된 인코더는 특히 낮은 비트율에서 고음질을 유지하면서 MPEG 에 의해 기존에 제안됐던 인코더에 비해 10~15%의 계산 부담을 현격히 감소한 것을 확인하였다. 개발된 인코더 알고리즘은 32-bit 고정 소수점 연산 구조를 사용한 ARM 프로세서에서 구현되었다. 또한 오디오 성능의 주관적인 테스트는 MP3 와 MPEG-2 AAC 의 성능과 비교하여 수행되었고, 상용 코더와 대등한 음질성능을 갖는 것을 확인하였다.

References

[1] K. Sayood, *Introduction to Data Compression*, Morgan Kaufmann, 1996.

[2] A. Puri, T. ChenK, *Multimedia Systems, Standards, and Networks*, Marcel Dekker, 2000.

[3] E. Zwicker, H. Fastl, *Psychoacoustics*, Springer, 1999.

[4] K. C. Pohlmann, *Principles of Digital Audio*, McGraw Hill, 2000.

[5] ISO/IEC 14496-3 Coding of audio-visual objects Part 3: Audio, 1998.

[6] J. Princen and A. Bradley, "Analysis/Synthesis Filter Bank Design Based on Time Domain Aliasing Cancellation," *IEEE Trans. on ASSP*, Vol. 32, pp. 1153-1161, Oct. 1986.

[7] T. Painter and A. Spanias, "Perceptual Coding of Digital Audio," *IEEE Pro.* Vol. 88, no. 4, pp. 451-513, April 2000.