임베디드 시스템 기반 MPEG-4 BSAC 오디오 최적화 구현

(Impelementation of Optimized MPEG-4 BSAC Audio based on the embedded system)

황진용*,박종순*,오화용**,김병일**,장태규***
Jin-Yong Hwang, Jong-Soon. Park, Hwa-Yong Oh, Byoung-Ii Kim, Tae-Gyu Chang

Abstract - 본 논문에서는 MPEG-4 Version2 Audio 표준에 근거하여 낮은 연산부담을 갖는 독자적인 엘고리즘을 적용한 MPEG-4 BSAC Audio 디코더를 개발하였다. 개발된 BSAC 디코더는 32bit RISC 구조를 갖는 Intel Xscale Processor 기반 시스템에 최적화하여 구현 및 평가를 수행하였다. 수행속도 증가 및 연산 정밀도 향상을 위해 각 기능 블록별 기능 및 구현 원리 연구와 32 bit 연산 구조를 파악하여, 이를 고정소수점 연산 구조로 구현함으로써 성능을 향상시켰다. 유한비트에 따른 오차 영향을 최소화하기 위해 데이터의 표현 범위에 대한 연구를 통해 근사화 오차를 최소화 하여 연산 정밀도를 향상 시키고자 하였다. 비선형 양자화기 및 filter bank 등 상대적으로 높은 연산 부담을 갖는 기능 블록은 Table look-up, 보간법, 지수연산 제거, pre/post scrambling 기법 등을 적용하여 최적화 하였다. 최종적으로 개발된 BSAC 디고더는 32 bit 연산 구조의 X-scale 프로세서를 탑재한 Development Board와 WindowsCE OS로 구성된 타겟 system에 이식하여 performance 평가하였으며, 높은 연산 정밀도 및 빠른 수행속도를 확인할 수 있었다. 주관적인 청각 평가에서도 MPEG-4 reference 디코더와의 음원의 차이가 거의 없음을 확인하였다.

Key Words : MPEG-4 BSAC, 고정소수점, Audio decoder, fine grain scalability

1. 장 서론

디지털 신호처리 기술의 발전과 하드웨어 성능의 비약적인 향상에 따라, 다양한 형태의 오디오 멀티미디어 서비스가 생성 제공되고 있으며, 유선 환경기반의 서비스에 국한되지 않고 DMB, 개인휴대단말과 같은 무선 환경 서비스로 영역이 점차 확대되고 있다. 향상된 오디오 압축기법 및 전송을 위한 인프라의 급속한 발전이 이루어 지는 중에도 효율적인 오디오 코더의 개발의 필요성이 계속 요구되고 있다. 이는 제한된 대역폭을 갖는 네트워크를 통해 고품질의 오디오서비를 사용하고자 하는 요구가 급속도로 늘어나기 때문이다.

MPEG-4 Version2의 BSAC(Bit Slice Arithmetic Coding) 오디오 표준은 제한된 밴드폭 및 정보손실이 발생하기 쉬운 채널 환경에서 오차의 강인성과 scalability를 제공하며, 적은 비트율에서 고품질의 오디오 재생이 가능하다.[1] BSAC 오디오 표준은 고품질의 오디오 재생을 위해 Transform 기법과 심리음향 모델을 적용한 양자화기를 통해 효율적인 압축을 실현한다. 하지만 압축된 Data를 재생하기 위해서는 디코더에 높은 연산 부담을 요구한다. 이는 모바일 환경 등의 적용에서는 중대한 문제점이 된다.

본 논문에서는 MPEG-4 버젼2 audio의 기능에 충실하면서 DMB 및 개인휴대단말과 같은 모바일 단말에서의 응용을 위해 32-bit X-scale Processor 기반의 H/W를 타겟으로 복잡도 및 연산부담을 최소화하는 독자적인 MPEG-4 BSAC Audio 디코더를 개발하였다. 수행속도 증가를 위해 32 bit 연산 구조를 파악하고, 이를 고정소수점 연산 구조 방식으로 구현함으로써 성능을 향상시킬 수 있었다. 유한비트 연구를 통해 유한비트 근사화 오차를 최소화 할 수 있는 수치 표현 범위를 설정하고 이를 코드에 적용하였다. 비선형 양자화기에서는 Table Look-up 방식으로 높은 연산 부담을 갖는 지수연산을 대체하였고, 또한 보간법을 이용 소요메모리를 최적화 할 수 있었다. 개발된 BSAC 디코더는 Sophia社의 Sandgate II Board와 WindowsCE OS로 구성된 system에 적용 평가하였다. 원곡의 36% 수준의 빠른 디코딩과 음질의 손실이 적은 고음질의 BSAC 디코더를 확보할 수 있었다.

저자 소개
*中央大學校 電子電氣工學部 碩士課程
**中央大學校 電子電氣工學部 博士課程
***中央大學校 電子電氣工學部 敎授

2. 장 MPEG-4 BSAC 디코더 구조

MPEG-4 Audio 버전2(ISO/IEC 14496-3) 표준을 따르는 BSAC 오디오 디코더는 기존 AAC 디코더의 구성에서 noiseless coding 부분을 제외한 다른 기능블록은 그대로 사용한다. noiseless coding 기능블록은 fine grain scalability (1Kbps) 기능을 효과적으로 부여하기 위해 Huffman coding 대신에 bit sliced arithemetic coding으로 대체되었다.[1][3]

MPEG-4 BSAC decoder의 구조는 그림 1과 같이 BSAC decoding, De-quantization, Spectral decoding, Fiter bank의 네 부분의 중요한 기능 블록으로 나누어진다. 새로 도입된 noiseless coding 방식인 bit sliced arithmetic coding은 기존의 Huffman coding이 각각의 스펙트럼을 심볼 단위로 코드워드를 할당하는 데 반하여 arithmetic coding은 심볼들의 연속적인 시퀀스에 따라서 코드워드가 가변 길이로 할당되는 방식이다. 이는 channel 환경에 따라 bit rate을 가변적으로 부여하는데 효과적이다.

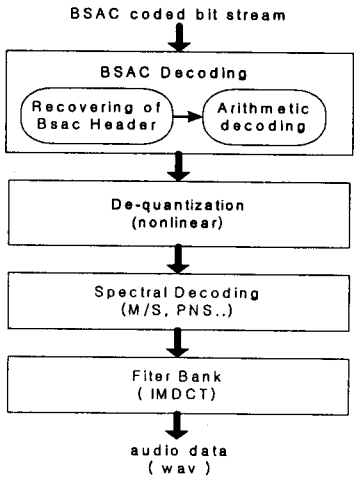

그림 1. 개발한 MPEG-4 BSAC 디코더 구조

역양자화 과정에서는 스펙트럼이 갖는 값의 범위에 따라 비트효율을 높일 수 있는[4] 비선형 양자화기를 사용한다. 압축 효율을 높이기 위해 spectral decoding에서는 mid/side stereo, PNS(Perceptual Noiseless Substitution)가 사용된다. filter bank에서는 IMDCT 알고리즘을 기반으로 주파수 계수로 표현된 data를 시간 영역의 오디오 신호로 변환 시킨다. 이는 시간영역에서 50% 중첩된 신호를 변환하도록 하는 기법으로 TDAC(Time Domain aliasing cancellation)에 기반한 변환 방식이다.[2] 또한 프레임간 양자화 스텝의 크기가 서로 다른 프레임을 중첩하여 신호를 생성하기 때문에 블록간 불연속 현상을 억제하는 효과를 준다.

비선형양자화기는 상대적으로 적은 비트로 audio data를 표현하는데 효과적이지만, 지수연산으로 인해 디코더에 높은 연산 부담을 요구한다. IMDCT 기능 블록은 sine파의 곱셈 및 덧셈 연산의 반복 loop로 구성되어 있어 디코딩 시간이 길어지는 문제점이 있다.

3. 장 32-bit 연산기반 BSAC 디코더 최적화 기법

본 논문에서는 32-bit 연산을 하는 X-scale 프로세서 기반의 SandgateII Board에 적용을 목적으로 오디오 디코더의 C-code를 최적화 하였다. 아래에는 본 논문에서 제안한 최적화 기법을 기술하였다.

먼저 수행속도 최적화를 위해 고정 소수점 연산구조로 수행될 수 있도록 연산 구조를 변환하고 유한비트 근사에 따른 오차영향 분석 및 최적 알고리즘 개발을 수행하였다.

비선형 양자화를 수행하는 역양자화 기능 블록은 지수 연산을 Table Look-up 방식으로 대체하여 연산 부담을 낮추고[4], 이 과정에서 발생할 수 있는 근사화 오차를 최소화 할 수 있도록 하였다. 비선형 양자화 식은 식 (1)와 같다. 3/4승 지수 연산은 디코더의 성능을 저하 시키는 요소중 하나이므로 1+(1/3) 형태로 변환하여 부호부 연산은 곱의 형태로 변환하고 1/3에 대해서만 Table look-up 방식으로 대체하여 디코더의 Throuput을 향상시켰다. 이에 따라 테이블의 수치 범위도 $(2^{13})^{4/3} = 17.33$에서 $(2^{13})^{1/3} = 4.333$으로 낮아져서 같은 비트폭에 대해서 더 많은 유효자리를 확보와 함께 연산 정밀도를 향상할 수 있었다.

$$x_invquant = sign(x_quant) \cdot |x_quant|^{\frac{4}{3}} \quad (1)$$

IMDCT 연산 과정은 sin 함수의 곱셈과 덧셈의 반복 수행으로 인해 디코딩 시간에 절대적인 부분을 차지한다. 본 연구에 사용한 IMDCT 알고리즘은 전/후처리 과정(pre/post-scrambling)과 FFT 연산에 의하여 시간 영역의 aliasing 현상의 제거와 함께 빠른 수행시간을 가지는 연산법을 적용하였다.[5] 반복 연산을 수행시 유한비트 표현에 따른 누적오차가 발생하는데, 미리 계산된 Table 값을 참조하는 방식을 사용하면 근사화 오차의 누적을 막을 수 있을 뿐만 아니라 연산 부담도 낮출 수 있다.[2][5]

4. 장 X-scale 프로세서 기반 BSAC 디코더 구현

본 논문에서는 Intel의 Xscale 프로세서인 PXA270 MPU을 탑재한 Shopia社의 SandgateII Board를 타겟 system으로 사용하였다. Board에 탑재된 OS는 Embeded VC++로 쉽게 Application 개발이 가능한 WindowsCE를 선택하였다. Host PC에서 embeded visual C++4.0 프로그램을 사용하여 MPEG-4 BSAC 디코더를 개발하였다. 프로그램 디버깅은 Microsoft社에 WidowsCE용 SDK를 사용하였다. SDK는 타겟 board와 같은 emulation 환경을 제공하여 손쉬운 디버깅 및 simulation 환경을 제공한다. 아래 그림 2은 위와 같은 개발 환경 구성도를 나타내고 있다. Host PC에서 BSAC 디코더를 개발하고 SDK를 사용하여 디버깅을 한 후 최종 결과물인 BSAC 디코더 파일을 타겟 board에 이식하여 수행속도, 메모리 사용량, 음질 테스트 등을 수행하였다.

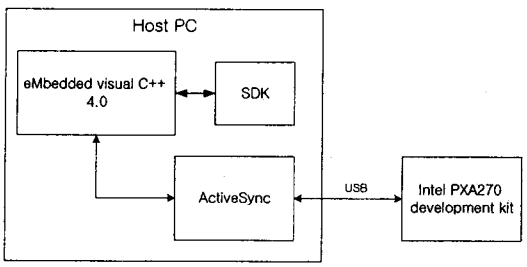

그림 2. 본 논문에서 사용한 개발환경 구성도

5. 장 실험 및 결과

본 논문에서 개발된 BSAC 디코더는 그림 3와 같이 32 bit floating point 연산을 하는 reference 디코더와 비교하여 평가를 수행하였다. BSAC 디코더의 입력 스트림은 MPEG reference 인코더로부터 얻은 flexmux format으로 mp4 확장자를 갖는다. 44.1KHz의 16bit 정밀도를 갖는 테스트 스트림으로 평가를 진행 하였다.

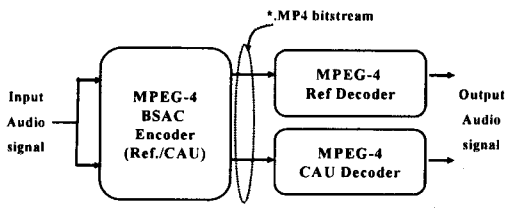

그림 3. BSAC 디코더 동작 시험 환경

아래 그림 4와 그림 5은 0에서 15MHz까지 5초동안 sweep 신호에 대한 디코딩 결과를 나타내고 있으며 sweeping이 잘 되는 것을 확인할 수 있다.

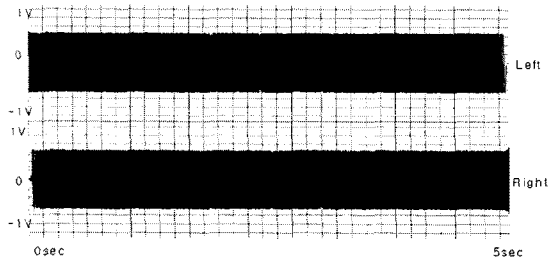

그림 4. 개발한 BSAC 디코더로 출력된 sweep신호의 좌우 채널 시간영역 파형의 예

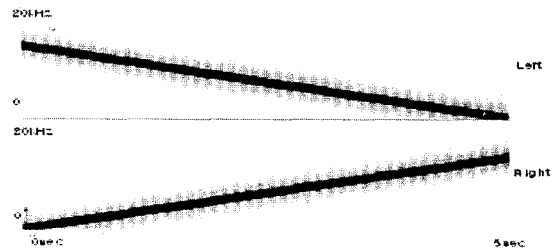

그림 5. 개발한 BSAC 디코더로 출력된 sweep신호의 좌우 채널 주파수 파형의 예

디코더의 성능을 평가하기 위해 MPEG-4의 reference 디코더와 비교 평가 하였다. 아래 표 1과 같이 개발한 디코더의 디코딩 시간은 원곡의 36%로 우수한 디코딩 성능을 나타내었으며,. floating point로 구현한 reference 디코더보다 약 6배 빠른 성능을 나타내었다.

표 1. 개발 BSAC 디코더와 MPEG-4 reference 디코더와 디코딩 시간 비교

곡	재생	BSAC 디코더		reference 디코더	
	시간	처리시간	비율	처리시간	비율
5BWV1011.mp4	3:23	1:16	0.37	7:45	2.29
Don't kill me tonight.mp4	3:34	1:24	0.39	7:16	2.03
Anywhere.mp4	4:01	1:34	0.39	7:43	1.92
Lindsay Everything UR.mp4	3:44	1:28	0.39	6:31	1.75
Shut up.mp4	3:46	1:29	0.43	6:27	1.71
Schon lacht.mpr	4:26	0:48	0.18	7:36	1.71
Tonight the heartaches.mp4	3:26	1:19	0.38	7:43	2.24
평균	3.29	0.86	0.36	6.71	1.95

본 연구에서는 개발된 BSAC 디코더의 음질 성능을 평가하기 위해 ISO/IEC 14496-3 Conformance Test 문서에 정한 기준을 이용하여 reference 디코더에 대해 PSNR을 측정하여 음질 평가를 실시 하였다. 측정식은 식 4와 같고 PSNR 결과 표 2와 같이 평균 77dB로 연산 정밀도가 높음을 알 수 있다. 주관적인 청각 테스트에서도 음원의 차이를 느낄 수 없는 수준임을 확인하였다.

표 2. 개발한 BSAC 디코더의 PSNR 평가

곡	PSNR	재생시간
5BWV1011.mp4	79.2	3:23
Don't kill me tonight.mp4	73.8	3:34
Anywhere.mp4	71.5	4:01
Lindsay Everything UR.mp4	70.0	3:44
Shut up.mp4	84.1	3:46
Schon lacht.mpr	79.4	4:26
Tonight the heartaches.mp4	78.5	3:26
평균	76.64	3.29

6. 장 결론

본 논문에서는 MPEG-4 Version2 Audio 표준을 바탕으로 32bit 연산구조를 갖는 X-scale 프로세서 Target system에 최적화 하기 위해 고정소수점 연산 구조로의 변환 및 이에 따른 근사화 오차를 최소화 하기 위한 유한비트 연구를 수행 적용하였다. 개발된 BSAC 디코더는 Target system에 이식 평가한 결과 실시간 디코딩이 가능하였으며, 디코딩 수행시간 도reference 디코더에 약 18% 수준으로 현격히 개선되었음을 확일 할수 있었다. PSNR 측정결과도 76dB로 연산정밀도가 높음을 알 수 있었고, 주관적인 청각 평가에서도 reference 디코더와 대등한 음질 성능을 갖는 것을 확인 하였다.

참 고 문 헌

[1] Heiko Purnhangen. MPEG-4 Audio(Final Comittee Draft 14496-3 AMD1, document, Juli 1999)

[2] Ku, D.S.; Kim, P.J.; Jeong, L.S.; Yang, D.H.; Hwang, J.Y.; Kim, J.B,"A design on the vector-processor of MDCT/IMDCT algorithm for digital audio",SICE 2003 Annual Conference,Volume 1,pp 1132-1135, 2003

[3] Fang-Chu Chen; Te-Ming Chiu, "Scalefactor based bit shift FGS audio coding", AINA 2005, Volume 2, pp 28-30 March 2005

[4] Tsung-Han Tsai, "A high quality re-quantization /quantization method for MP3 and MPEG-4 AAC audio coding", IEEE,. International Symposium on Volume 3, pp.26-29 May 2002

[5] Po-sheng Wu; Yin-Tsung Hwan,"Efficient IMDCT core designs for audio signal processin", SIPS 2003. IEEE Workshop, pp. 275-280, Aug. 2003