

MPEG-4 스케일러블 오디오 부호화 방식에 관한 관찰

A survey of MPEG-4 scalable audio coding

김연배, 박성희, 서양석
삼성종합기술원 신호처리 LAB

요약

지금까지 알려진 오디오 부호화 방식중 실제로 스케일러블 부호화 방식은 없었다. MPEG-4 에서는 스케일러블 부호화라는 새로운 기능을 중요한 요구사항으로 받아들여 개발하였다. 스케일러블 부호화(scalable coding) 기능이란 주어진 자원의 일부분을 가지고 의미 있는 정보를 재생할 수 있게 하는 부호화방식이다. 본 논문에서는 ISO MPEG-4 에서 추진하고 있는 스케일러블 부호화방식인 large step scalable coding 방식과 fine grain scalable coding 방식에 대해 알아보고 그 응용분야에 대해 살펴본다.

1 서론

스케일러블 부호화방식이란(scalable coding) 주어진 자원의 일부분을 가지고 의미 있는 정보를 재생할 수 있게 하는 부호화방식이다. 스케일러블 부호화 방식에는 bitrate scalability, bandwidth scalability, complexity scalability 등이 있다. 이중 가장 의미 있는 방식중의 하나가 bitrate scalability 이다. Bitrate scalability 는 비트스트림의 일부분만을 가지고도 복호화가 가능하며 또한 의미 있는 신호를 만들어 낼 수 있는 방식을 말하는 것이다. 이러한 특징은 데이터 전송시, 전송로가 전체 비트스트림을 전달하기 위해 필요한 대역폭의 제공을 보장하지 못하는 경우에 해결책으로 제시되어질 수 있다.

그동안 비디오분야에서의 scalability 는 종종 거론이 되었으나, 오디오분야에서는 거의 거론되지 않았었다. 오디오분야에서는 다루었던 scalability 는 MPEG-1 에서 사용한 bandwidth scalability 와 MPEG-2 BC 방식이다. MPEG-1 의 bandwidth scalability 는 표준화에서의 정의가 아니라 각 응용제품에서 구현한 것으로서, 합성필터의 대역중 낮은 주파수 대역에 대한 데이터만을 사용함으로써 대역 제한된 신호를 만들어내는 것이었다. MPEG-2 BC 방식은 그 의미대로 MPEG-2 의 다채널에 대한 비트스트림중 스테레오부분에 대해서 MPEG-1 audio decoder 가 재생할 수 있게 하는 형태로 제공된다.

MPEG-4 에서는 이러한 scalability 를 중요한 기능 중 하나로 정의하고 표준화상에서 정의하고자 하였다. 여러 가지 형태의 scalability 에 대한 요구사항이 있지만, 여기서는 가장 의미 있는 방식인 bitrate scalability 를 그 기능으로 선정하였다. 이방식으로, 부호화기에서는 높은 비트율로 비트스트림을 작성하고, 전송시 통신선로의 대역폭에 따라 원 비트스트림의 일부만이 수신단에 전송된다.

수신단 측면에서는 전송된 비트스트림을 이용하여 복호화기의 복잡도 또는 사용자의 음질 요구사항에 따라 자유롭게 음악을 재생할 수 있다.

본 논문에서는 MPEG-4 에 기술된 scalability 중 General Audio 에 대해 알아보고 scalable coding 의 응용에 대해 살펴본다.

2. Scalable Audio Coding

2.1 개요

현재 MPEG-4 audio 부분에는 크게 2 가지의 scalable audio coding 방식이 있다.

- Large Step Scalability : scalable AAC
- Fine Grain Scalability : BSAC

Large Step Scalability 는 각 단계의 간격이 큰 것을 의미하고 Fine Grain Scalability 는 간격이 세밀한 방식이다. Large Step Scalability 에서는 각 단계의 간격을 약 8kbps/ch 이상으로 하였고 Fine Grain Scalability 에서는 1kbps/ch 로 하였다.

2.2 Large Step Scalability

이방식은 2 개 이상의 같은 또는 서로 다른 부호화기를 연결하여 사용함으로써 scalable coder 를 형성한다. 또한, 이방식은 매우 낮은 비트율(약 6kbps)을 지원하기 위해 core coder 를 사용할 수 있다. core coder 로는 음성 부호화방식인 MPEG-4 Narrow Band CELP 또는 TwinVQ 를 사용한다. scalable AAC 은 다음과 같은 구성을 지원한다.

- 1) AAC layer only
- 2) Narrow-band CELP base layer + AAC
- 3) TwinVQ base layer + AAC

AAC 또는 TwinVQ 의 coding layer 는 mono 또는 stereo/joint stereo 일 수 있다.

1) AAC layer only

이방식은 각 scalable 단계마다 AAC 방식을 이용하여 전체적인 scalable coder 를 구성하는 것이다. 예를 들면, 3 단계의 scalability - 24kbps, 32kbps, 40kbps- 를 구현한다고 하자. 먼저, 원신호를 24kbps AAC coding 을 한다. 다음 단계에서는 원신호(original input signal)와 첫 단계에서 처리한 결과와의 차에 대해 8kbps AAC coding 을 한다. 마지막 단계에서는 원신호와 두번째 단계에서 처리한 결과와의 차를 8kbps AAC coding 하여, 각 단계에서 만들어진 비트스트림을 결합함으로써 scalability 를 구현하는 방식이다. 그림 1 은 2 단계 scalable AAC 의 블록도이다.

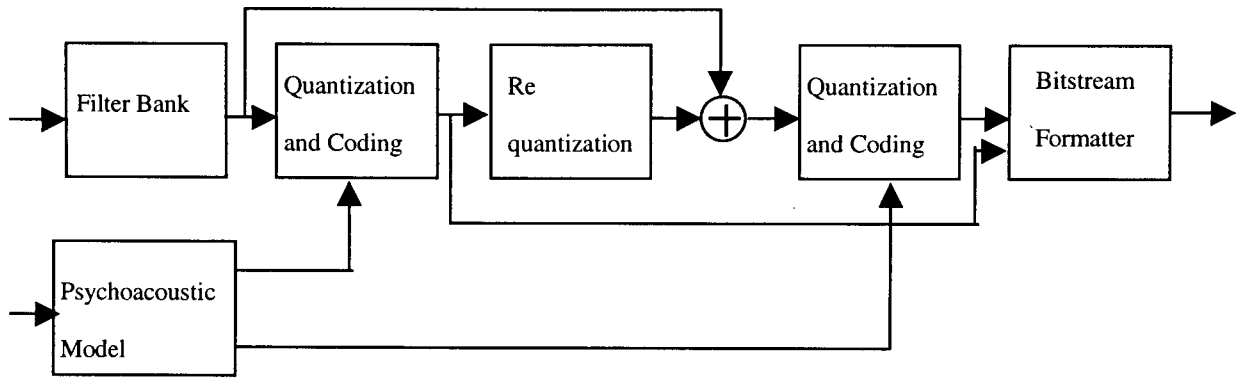


그림 1.2 단계 scalable AAC 부호화기의 Block diagram

2) Narrow-Band CELP base layer + AAC

이방식은 core coder 로써 MPEG-4 NB CELP 방식을 사용하고 그 오차신호를 AAC 또는 TwinVQ 방식을 이용하여 구현하는 것이다. 이방식은 먼저 입력신호를 MPEG-4 NB-CELP 방식으로 부호화 한 후, 부호화된 신호를 다시 NB-CELP 방식으로 복호화한다. 다음 단계의 부호화기의 입력은 원 입력신호와 복호화된 신호의 차를 이용한다.

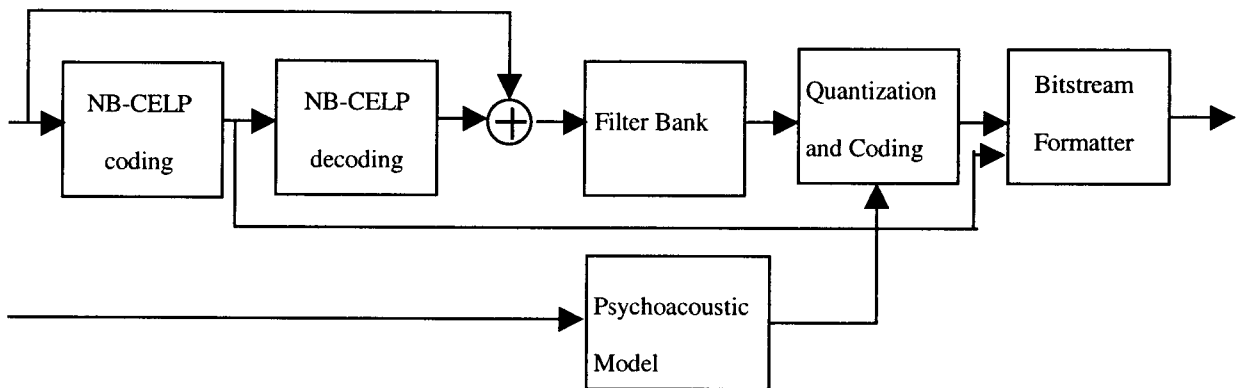


그림 2.2 단계 NB-CELP + AAC 부호화기의 Block diagram

3) TwinVQ base layer + AAC

이방식은 core coder 로써 TwinVQ 방식을 쓰고 그 오차신호를 AAC 또는 TwinVQ 방식을 이용하여 구현하는 것이다. 이방식은 NB CELP+AAC 방식과는 다르게, TwinVQ 와 AAC/TwinVQ 방식이 주파수 영역의 데이터를 다루기 때문에 단지 주파수 영역에서의 신호차이를 다음 단계에서 부호화한다.

표 1 은 scalable AAC 의 동작 모드에 대해 각 layer 의 수와 각 coder 별 채널구성을 요약한 것이다.

Mode	# of Layer	core coder present	# of channel for core coder	# of channel for Layer 1	# of channel for Layer 2	# of channel for Layer 3
0	1	O	1			
1	1	O	2			
2	1	X	0	1		
3	1	X	0	2		
4	2	O	1	1		
5	2	O	1	2		
6	2	O	2	2		
7	3	O	1	1	1	
8	3	O	1	1	2	
9	3	O	1	2	2	
10	3	O	2	2	2	
11	3	X	0	1	1	
12	3	X	0	1	2	
13	3	X	0	2	2	
14	4	X	0	1	1	1

표 1. scalable AAC 의 동작 모드

2.3 Fine Grain Scalability

이 기능을 제공하는 tool 이름이 BSAC(Bit-Sliced Arithmetic Coding)이다. scalable AAC 처럼 여러 codec 을 사용하여 scalability 를 주는 방식이 아니라 단일 부호화기를 사용하여 scalability 기능을 제공하는 방식이다. 기본적인 codec 의 구조는 AAC 방식을 사용하며 scalability 를 주기위한 bitstream 작성은 BSAC 기술을 사용하여 구성한다. 이방식에서 제공하는 scalability 는 1kbps/ch 단위의 scalable layer 를 제공한다. BSAC 방식은 fine grain scalability 를 제공하는 기술이지만 coding efficiency 측면에서는 top layer 에서 기존의 MPEG2-AAC 와 거의 동일하다.

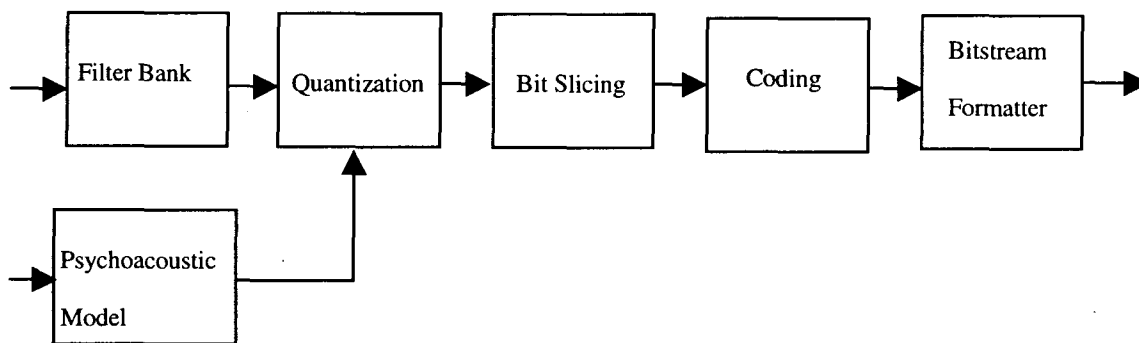


그림 3. BSAC 부호화기의 Block diagram

그림 4 는 BSAC bitstream 구조에 대한 개념도를 나타낸다.

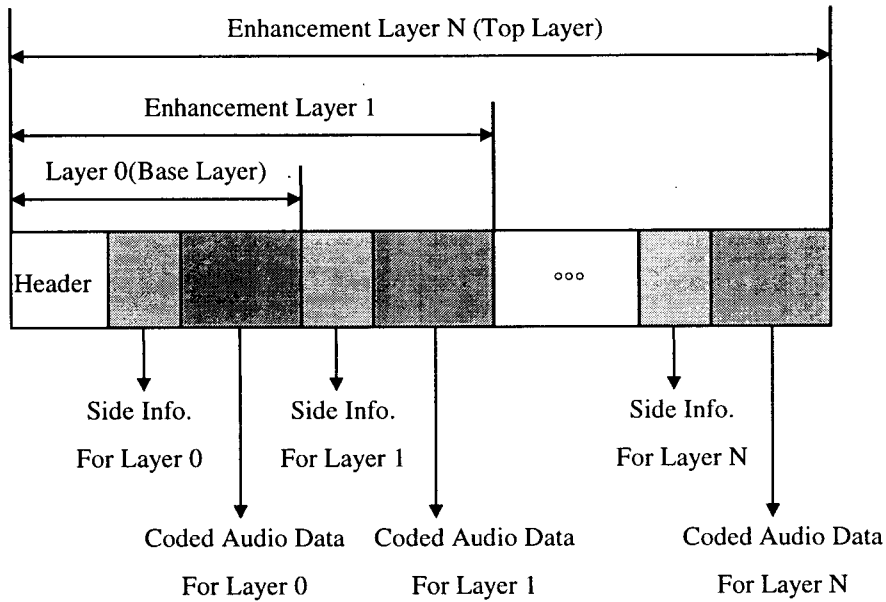


그림 4. BSAC bitstream 구조의 개념도

3. 스케일러블 부호화의 응용분야

그림 5는 오디오 서버를 통한 AOD(Audio on Demand)의 간단한 구성을 나타낸다. 이때 오디오 서버의 총 전송 가능 용량은 1024kbps 라고 하고, 10 명의 사용자(오디오 사용자 1 ~ 10)가 그림과 같은 서비스를 받고 있다고 가정하자. 11 번째 사용자가 128kbps 의 서비스를 요청 하였을 경우 각 부호화방식에 따른 처리를 살펴보자.

1) non scalable coding 방식을 사용하는 경우

- ㄱ) 현재 가용한 1024kbps 의 용량을 10 명의 사용자가 모두 사용하고 있으므로, 11 번째 사용자의 요구는 무시되어 질 수 있다.
- ㄴ) 11 번째 사용자에게 서비스를 하려면 사용자 5 부터 사용자 10 사이의 두 사용자를 선택하여 그 서비스를 128kbps 에서 종료하고 64kbps 로 낮추어 새로운 환경에 대한 재접속을 설정하고 11 번째 사용자에게 128kbps 의 서비스를 제공한다.

2) scalable coding 방식을 사용하는 경우

scalable coding 방식의 경우 기존 사용자에게 대한 재접속 과정이 없이 처리된다. 이것은 2 경우로 나누어 생각해 볼 수 있다.

ㄱ) large step scalable 방식의 경우

만일 large step scalable coding 방식이 64+32+32 로 되어 있다고 가정하면, 사용자 5 부터 사용자 10 중 3 명의 사용자의 비트율을 128 에서 96 으로 감소시키고 11 번째 사용자에게 96kbps 의 서비스를 제공한다.

ㄴ) fine grain scalable 방식의 경우 사용자 1 과 사용자 2 는 64kbps 에서 56kbps 로 감소시키고, 사용

자 5 부터 사용자 10 까지는 128kbps 에서 112kbps 로 감소시킴으로써 사용자 11 에게 112kbps 의 서비스를 할 수 있다.

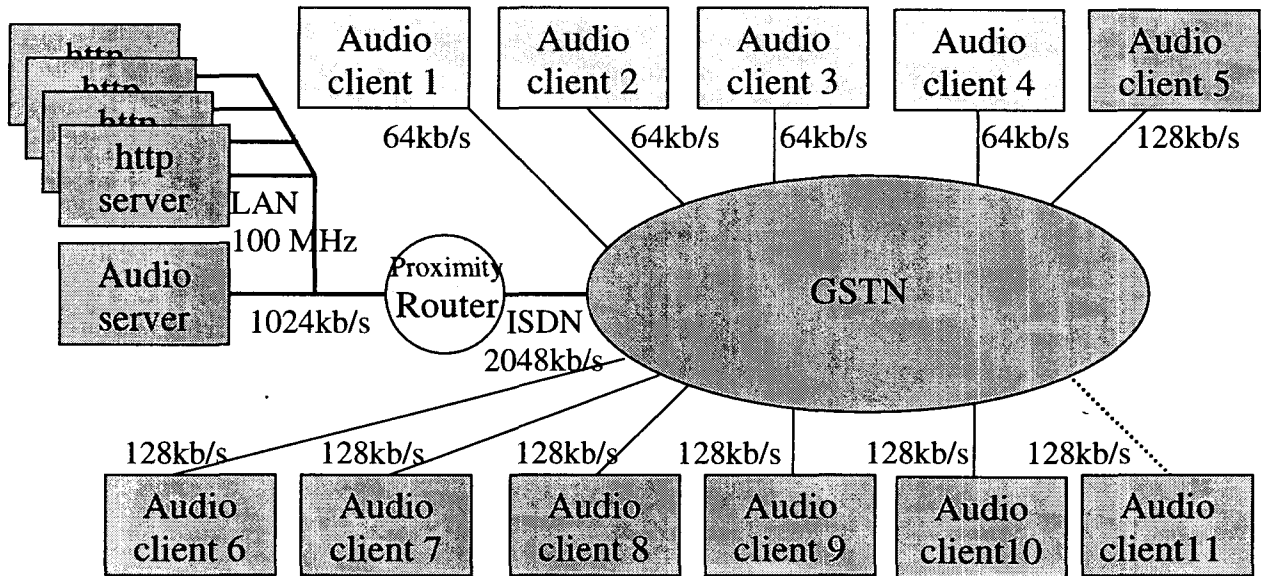


그림 5. 오디오서버를 통한 AOD 서비스 구성의 예

4. 결 론

지금까지 MPEG-4 오디오의 scalable coding 에 관한 기술을 살펴보았다. scalable coding 방식은 기존에 효율적인 압축만을 지향하는 부호화 방식(MPEG-1, MPEG-2) 과는 다르게 많은 응용분야를 가지고 있다. 기존의 방식에서는 주로 사용하는 응용분야가 저장용 매체를 이용하는 것이었다. 즉, 컴퓨터의 파일이나 CD-ROM, DVD 등을 이용하여 안정적인 데이터 전송을 가정하는 방식인 것이다. 그러나, 앞으로는 이러한 저장매체뿐 아니라 인터넷과 같은 통신을 통한 음악서비스가 본격화되어 진다. 통신선로상의 용량이 무한대라면 기존의 방식도 별 무리 없이 적용될 수 있으나, 한정된 자원을 공유해야 하는 현재의 상황에서는 scalable coding 방식이 가장 적합한 방식이라 할 수 있다. 즉, 기존의 방식으로는 서비스가 불가능한 경우라도 scalable coding 을 이용하여 서비스가 가능해지므로 더 많은 사용자에게 서비스를 제공할 수 있다. 또한, 이방식을 사용하면 사용자에게 제공되는 서비스의 품질에 대한 차별화도 가능하다. 예를 들면, 사용자가 지불하는 요금에 따른 차등적인 서비스가 가능하다.

참고문헌

1. ISO/IEC JTC1/SC29/WG11 MPEG, International Standard ISO/IEC IS 11172-3 "Information technology - Coding of movinf pictures and associated audio for digital storage media up to about 1.5Mbps, Part 3:Audio"
2. ISO/IEC JTC1/SC29/WG11 MPEG, International Standard ISO/IEC IS 13818-3 "Generic Coding of Moving pictures and Associated Audio, Part 3:Audio"
3. ISO/IEC JTC1/SC29/WG11 MPEG, Draft International Standard ISO/IEC IS 14496-3 "Information technology -Generic Coding of Audiovisual Objects, Part 3:Audio"