

# 모바일 방송용 비디오 및 오디오 압축 기술

박진수, 김지희, 이승준(엠큐브웍스(주))

## I. 서론

1990년대 이 후 급속히 발전한 디지털 신호처리, 저장매체 및 전송 기술로 인하여 멀티미디어 데이터를 접하는 일은 우리의 일상 생활이 되었다. 특히, 21세기 들면서 발달한 광대역 모바일 전송 기술, 저전력 컴퓨팅 기술은 모바일 환경에서도 멀티미디어 매체를 즐길 수 있게 하였으며, 최근에는 급기야 고정된 장소에서 큰 단말기를 통해서만 즐길 수 있었던 TV 방송 서비스를 모바일과 포터블의 영역에서도 가능하게 한 '모바일 방송'의 시대를 열게 하였다.

이러한 모바일 방송을 가능하게 한 대표적인 기술로는 모바일 컴퓨팅 기술, 멀티미디어 전송 기술, 멀티미디어 신호 압축 기술 등을 들 수 있으며, 본 논문에서는 이 중 비디오 및 오디오 등 멀티미디어 신호의 압축 기술에 대해 살펴본다. 현재 국내의 지상파 DMB(digital multimedia broadcasting)와 위성 DMB 서비스에는 비디오 압축 기술로서 H.264가 사용되고 있으며, 이 H.264는 DVB-H 등 유럽/북미에서의 모바일 방송 및

일본의 ISDB-T 모바일 방송 서비스를 위해서도 채택되었다. 오디오 압축 기술로는, 이른바 AAC+라고 불리고 있는 HE-AAC(high-efficiency AAC) 기술이 국내 위성 DMB, DVB-H, ISDB-T 서비스를 위해 채택되었으며, 압축률은 HE-AAC보다 다소 떨어지지만 에러강인성 및 Scalability 기능이 우수한 BSAC(bit sliced arithmetic coding) 기술이 국내 지상파 DMB 서비스를 위해 사용되고 있다.

먼저 II장에서는 MPEG-4 AVC(advanced video codec)이라고도 불리는 H.264 비디오 압축 기술에 대해 기술하며, 이 후 III장 및 IV장에서 각각 HE-AAC 및 BSAC 오디오 압축 기술에 대해 설명한 후, V장에서 향후 전망과 함께 본 논문을 정리한다.

## II. H. 264

### 1. H.264 개요

H.264동영상 압축 방식<sup>1)</sup>은 대표적인 표준화 단체인 ITU-T의 Video Coding Experts

Group과 ISO/IEC의 Moving Picture Experts Group이 연합하여 제정한 차세대 동영상 압축 방식으로, MPEG-4 Part 2등 기존의 동영상 압축 방식들에 비해 압축 효율이 현저히 개선되었으며, 다양한 네트워크 환경에 쉽게 부응할 수 있는 유연성을 갖추었다.

H.264는 동영상 데이터를 효율적으로 압축하는 계층인 VCL(Video Coding Layer)과 다양한 네트워크 환경에서 데이터를 전송 및 저장할 수 있도록 압축된 동영상 데이터의 형식을 변경시켜주는 NAL (Network Abstraction Layer)로 이루어져 있다.

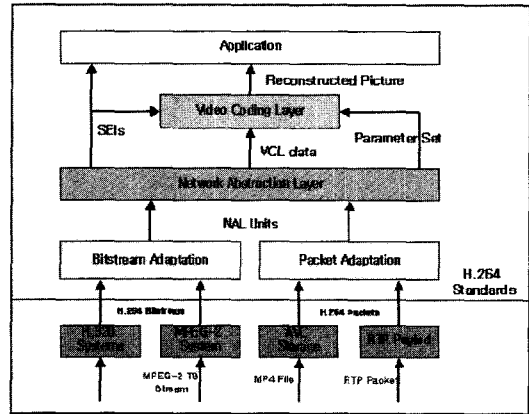
H.264는 기존의 동영상 압축 방식에 비해 높은 압축 성능과 네트워크 친화성의 장점을 가지는 반면, 부호기 및 복호기의 복잡도가 훨씬 증가하는 단점을 가지고 있다. 따라서, H.264가 실제 시장에서 널리 사용하기 위해서는, H.264의 원리 및 기술을 정확하게 이해하여 효율적인 부호기와 복호기를 구현하여 적용하는 기술이 매우 중요하다고 할 수 있다.

## 2. H.264 알고리즘 소개

### 2.1 NAL (Network Abstraction Layer)

다양한 디지털 통신망의 보급과 멀티미디어 콘텐츠 서비스의 확산은 서로 다른 특성을 갖는 네트워크 간의 데이터 교환을 빈번히 발생시켰으며, 이로 인해 서로 다른 특성을 갖는 네트워크 환경에서 전송되어야 하는 데이터는 네트워크들의 서로 다른 특성을 만족해야 한다. 이러한 요구사항을 만족시키기 위해 H.264에서는 NAL Unit이라 불리는 데이터 형식을 정의하여 네트워크 특성과 관계

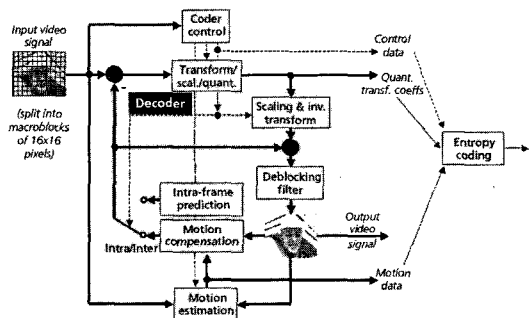
없이 데이터를 생성, 전송할 수 있도록 하였다. <그림 1>은 H.264 표준 복호기에서 생성된 데이터가 서로 다른 네트워크 환경에서 적응적으로 전송되는 것을 보여준다.



<그림 1> H.264 표준 복호기 구조

### 2.2 VCL (Video Coding Layer)

H.264 동영상 압축 방식의 기본 구조는 <그림 2>와 같이 기존 동영상 압축 방식을 토대로 형성되었으나, 각각의 세부 알고리즘의 변화로 기존 동영상 압축 방식보다 뛰어난 압축 성능을 보인다.



<그림 2> H.264 부호기의 블록도

#### 1) 화면간 예측 부호화

H.264의 움직임 추정 및 보상 방식은 기존

동영상 압축 방식의 블록 매칭 기법을 기반으로 아래와 같은 특징들을 추가하여 보다 정확하게 움직임을 보상하여 압축 효율을 개선 시켰다.

**\* 다양한 블록 단위를 사용한 움직임 보상방식 :**

H.264는 기존의 동영상 압축 방식의 움직임 보상방식에서 사용하는 블록들에 비해 <그림 3>과 같이 다양한 형태의 블록들을 사용하여 움직임 보상방식을 수행하므로 다양하고 복잡한 움직임 영역을 충분히 표현할 수 있게 되었다.

16x16 매크로 블록 (MacroBlock)	16 x 16 움직임 보상	16 x 8 움직임 보상	8 x 16 움직임 보상	8 x 8 움직임 보상
	0	0 1	0 1	0 1 2 3
8x8 단위 매크로 블록 (Sub MacroBlock)	8 x 8 움직임 보상	8 x 4 움직임 보상	4 x 8 움직임 보상	4 x 4 움직임 보상
	0	0 1	0 1	0 1 2 3

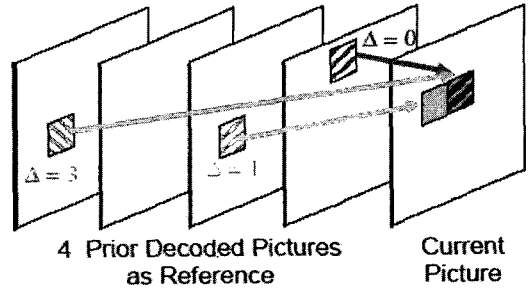
<그림 3> 다양한 블록단위의 움직임 보상방법

**\* 다중 참조 영상을 이용한 움직임 보상방식 :**

H.264는 다수의 재구성된 영상을 참조 영상으로 사용하여 움직임 벡터를 추정하고 보상함으로써 움직임을 좀 더 정확하게 표현할 수 있다. <그림 4>는 4개의 이전 영상을 참조 영상으로 사용한 움직임 보상의 예를 나타내며  $\Delta$ 는 현재 영상과 참조 영상 사이의 시간적인 간격을 나타낸다.

**\*1/4 화소 단위의 움직임 예측 :**

H.264는 좀더 정확한 움직임 벡터를 추정하기 위해 영상을 4배로 보간하여 움직임 예



<그림 4> 다중 참조 영상 움직임 보상방법

측 및 보상을 수행한다.

**2) 부정합이 없는 작은 블록크기 블록 변환**

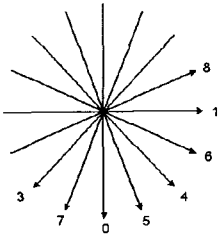
기존 동영상 압축 방식은 8x8 블록 단위의 변환 방식을 사용하였으나, H.264는 영상의 국부적인 특성에 더 적합하도록 4x4 블록단위로 변환을 수행한다. 이는 ringing artifact로 알려져있는 잡음을 제거하는데도 효과적이다.

**3) 공간영역의 화면 내 예측 부호화 기법**

기존 동영상 압축 방식의 대표적 방식인 MPEG-4 part 2는 휘도 및 색차에 대한 화면 내 예측을 변환된 영역에서 수행하는 반면, H.264는 공간 영역에서 수행한다. 이는 공간 영역의 방향성을 이용하여 다수의 예측 모드를 사용함으로써 압축 효율을 향상시킬 수 있다. <그림 5>은 4x4블록의 화면 내 예측 부호화 과정에서 사용하는 예측 모드를 나타낸다.

**4) 루프 필터**

부호기와 복호기 내부에 위치한 H.264의 루프 필터는 각 블록의 부호화 형태, 양자화 크기, 참조 영상 및 움직임 벡터 크기를 고려하여 필터링을 수행하므로 양자화 과정에서 발생하는 블록화 현상을 제거하고 화질 향상



예측 모드	예측 방향
0	Intra_4x4_Verical
1	Intra_4x4_Horizontal
2	Intra_4x4_DC
3	Intra_4x4_Diagonal_Down_Left
4	Intra_4x4_Diagonal_Down_Right
5	Intra_4x4_Verical_Right
6	Intra_4x4_Horizontal_Down
7	Intra_4x4_Verical_Left
8	Intra_4x4_Horizontal_Up

〈그림 5〉 화면 내 예측 부호화 모드

및 압축 효율을 향상시킨다.

### 5) 향상된 엔트로피 부호화

H.264의 엔트로피 부호화 방식은 크게 CAVLC(Context Adaptive Length Coding)와 CABAC(Context-based Adaptive Binary Arithmetic Coding)로 나누어지며, 선택적으로 사용할 수 있도록 하였다. 이들은 이미 부호화된 주변 블록들의 특성을 이용하여 부호화 과정을 수행하여 향상된 성능을 보인다.

### 6) 적응적 프레임/필드 부호화

H.264는 격자 주사 프레임 내의 이웃하는 두 행이 같은 패리티 필드 내의 이웃하는 두 행에 비해 낮은 상관관계를 갖는 경우에 효율적으로 부호화하기 위해 필드 부호화 방식을 선택적으로 사용할 수 있도록 하였다. H.264에서는 기존의 방식과는 달리 프레임/필드 부호화의 선택을 픽처 단계뿐 아니라 매크로 블록 단계에서도 선택할 수 있다.

## 3. H.264 응용분야

H.264는 높은 압축 성능과 우수한 네트워크 친화성으로 인해 국내 위성 및 지상파 DMB(Digital Mobile Broadcasting) 서비스와

유럽의 DVB-H(Digital Video Broadcasting-Handhelds) 서비스에서 동영상 압축 규격으로 채택되었다. 모바일 방송에 채택된 H.264의 구체적인 규격은 'Baseline profile'이라고 불리는 것으로, 이에는 위에서 언급한 CABAC, 적응적 프레임/필드 부호화 기법은 사용되지 않는다. H.264 Baseline profile은 모바일 방송에서만 아니라 모바일 환경에서의 메시징 및 화상전화 서비스를 위한 비디오 압축 방식으로도 채택되었다.

## III. MPEG-4 HE AAC

### 1. HE AAC 개요

MPEG-2 오디오의 등장 이후, 단순히 압축 효율을 높이기 위한 부호화의 관점을 넘어 이동 통신, 데이터베이스 접근 및 새로운 유형의 쌍방향 서비스 등 다양한 응용 분야에서의 적용을 목표로 새로운 기능의 확보에 중점을 둔 MPEG-4 오디오의 개발이 진행되었다. 결과적으로 MPEG-4 오디오는 2kbps의 낮은 비트 전송률 음성 부호화에서부터 채널 당 64kbps 이상의 고품질 오디오 부호화에 이르기까지 다양한 종류의 오디오 부호화 알고리즘을 포함하게 되었다.

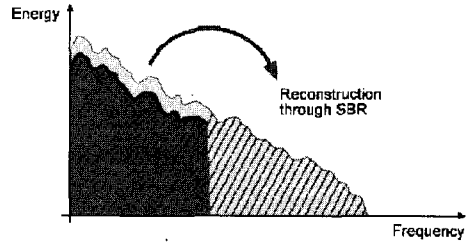
MPEG-4 HE AAC<sup>[2]</sup>는 MPEG-4 오디오의 다양한 부호화 알고리즘 가운데 일반적인 오디오 부호화를 담당하는 GA(General Audio Coding) 부호화 영역에 포함된 최신 부호화 기법의 하나로서, 채널 당 24kbps의 낮은 비트 전송률에서 CD 수준의 음질을 제공하는 저전송률 오디오 부호기법이다.

HE AAC의 높은 압축 효율은 이동통신 및

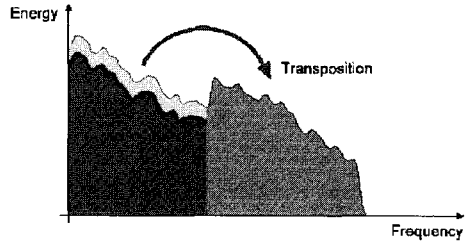
디지털 방송처럼 전송 대역이나 저장 용량이 제한되거나 매우 비싼 응용 분야에 적합하다고 할 수 있다. 또한 기존의 AAC와의 양방향 호환이 가능한 구조상의 유연성은 MPEG-4 HE AAC의 또 다른 장점이라고 할 수 있다.

## 2. SBR 알고리즘 소개

HE AAC는 MP3, AAC 등 기존의 지각적 오디오 부호기법의 한계를 극복하기 위해 대역폭 확장(Bandwidth extension) 기술의 하나인 SBR(Spectral Band Replication) 알고리즘을 AAC에 결합한 부호기이다. HE AAC의 주요 알고리즘인 SBR의 기본 원리는 오디오 신호의 고주파와 저주파 대역 사이에 높은 연관성이 존재한다는 사실을 이용하여, 저주파 대역의 정보를 이용해 고주파 대역 성분을 추정 및 복원하는 것이다. <그림 6>은 SBR을 이용해 고주파 대역을 복원하는 과정의 개념을 도식화한 것이다. 그림에 나타낸 것과 같이 SBR알고리즘은 저주파 스펙트럼 데이터를 고주파 대역으로 복사하는 전위(transposition) 과정과 소량의 추가 정보를 이용해 고주파 대역의 스펙트럼의 모양을 조정(adjustment)하는 과정으로 구성된다. HE AAC는 이러한 개념을 바탕으로 저주파 대역 성분은 AAC로 부호화하고, 부호화된 AAC 비트스트림에 고주파 대역 성분을 복원하기 위한 소량의 추가 정보를 포함하는 방식으로 설명될 수 있다. <그림 7>와 <그림 8>은 각각 HE AAC 부호기와 복호기의 구조를 나타낸 것이다.

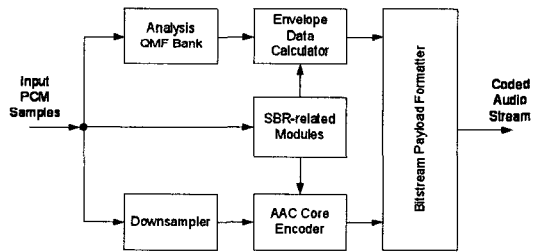


(a) 전위에 의한 고주파 생성

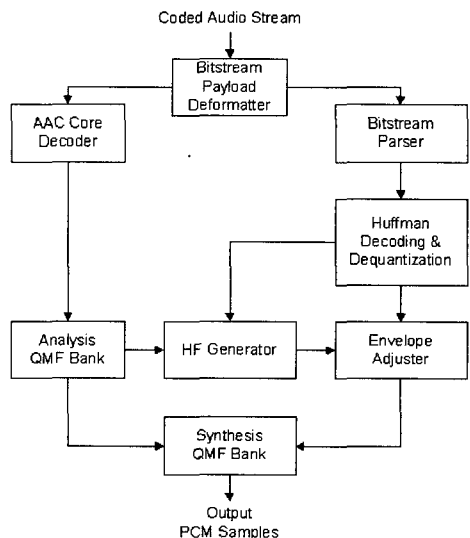


(b) 고주파의 포락선 조정

<그림 6> 고주파 대역의 복원 과정



<그림 7> MPEG-4 HE AAC 부호기의 구조



<그림 8> MPEG-4 HE AAC 복호기의 구조

HE AAC의 부호화는 다음과 같은 과정으로 수행된다. HE AAC 부호기에 입력된 오디오 신호는 먼저 down sampler를 거쳐 저주파 대역 성분으로 바뀐 후 AAC 부호기를 통해 부호화 된다. AAC 부호기로부터 출력된 비트스트림은 MPEG-2/4 AAC의 표준을 따르기 때문에 AAC 복호기를 이용해 복호화가 가능하다. 다만 이렇게 복호화된 오디오 신호는 down sampler를 통과한 후 부호화되었기 때문에 원본 오디오 신호와 비교해 표본화 주파수가 1/2로 줄어들게 된다. HE AAC 부호기의 입력 신호는 analysis QMF bank를 거쳐 스펙트럼 성분으로 변환된 후 envelope data calculator에 입력된다. 이 과정을 통해 스펙트럼의 고주파 영역의 특성을 표현하는 SBR 데이터가 추출된다. AAC 비트스트림에 비해 약 5-10%의 크기를 갖는 SBR 데이터는 MPEG-4의 표준에 따라 AAC 비트스트림에 포함된다.

부호화된 비트스트림이 HE AAC 복호기에 입력되어 복호화되는 과정은 아래와 같다. HE AAC 복호기에 입력된 비트스트림으로부터 AAC 데이터와 SBR 데이터가 분리된 후 AAC 데이터는 AAC 복호기에 입력된다. AAC 복호기를 통과한 저주파 대역 오디오 신호는 analysis QMF bank에 의해 주파수 성분으로 변환된다. SBR 데이터로부터 분석된 정보는 HF generator를 통해 저주파 대역 신호를 고주파 대역으로 복사하고 envelope adjuster에 의해 스펙트럼 모양을 조정하는데 사용된다. 각각 고주파 대역과 저주파 대역 성분에 해당하는 envelope adjuster와 analysis QMF bank의 출력은 synthesis QMF bank를 거쳐 최종적으로 원본 오디오

신호의 표본화 주파수와 같은 출력 오디오 신호를 생성한다.

### 3. HE AAC의 응용 분야

MPEG-4 HE AAC의 주요 이점은 뛰어난 오디오 음질을 제공하는 동시에 높은 부호화 효율을 갖는다는 것이다. 따라서, 플래시 메모리를 탑재한 휴대용 플레이어와 같은 파일 기반 저장 매체뿐만 아니라 디지털 방송 및 이동통신 네트워크와 같이 높은 압축 효율을 필요로 하는 응용 분야에 적합한 기술이라고 할 수 있다. 또한 전송 대역이나 저장 용량은 부호화될 오디오 채널의 수에 비례하기 때문에 높은 압축 효율은 다채널 환경에서도 유용하다고 할 수 있다. MPEG-4 HE AAC는 국내 위성 DMB와 유럽 DVB-H의 오디오 부호기 표준으로 채택되어 사용되고 있다.

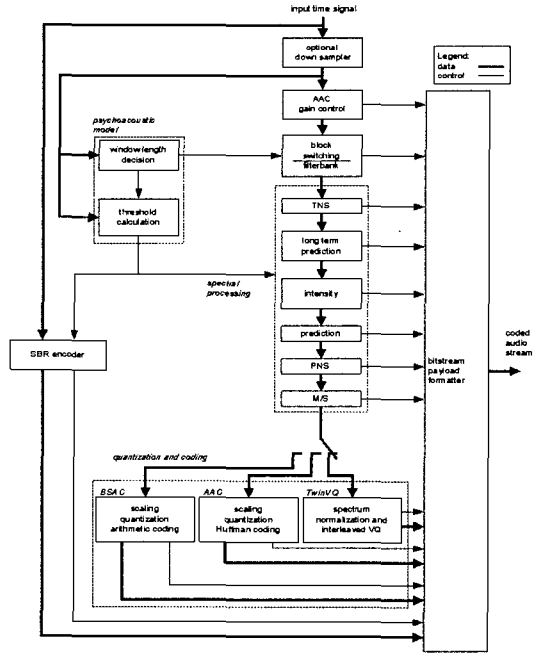
## IV. MPEG-4 BSAC

### 1. BSAC 알고리즘 소개

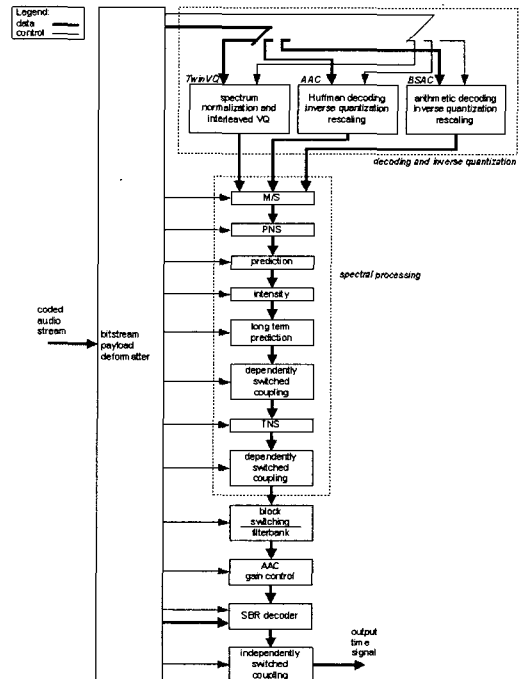
BSAC(Bit Sliced Arithmetic Coding)<sup>1)</sup>은 MPEG-4 GA에 포함된 오디오 부호기의 하나로써 AAC와 유사한 구조를 갖는다. <그림 9>는 MPEG-4 GA 부호기의 구조를 나타낸 것이다. MPEG-4 GA는 일반적인 지각적 오디오 부호화 알고리즘의 구조를 기반으로 한다. 즉, 시간 영역 오디오 신호와 스펙트럼 신호간의 변환을 위한 filterbank, 스펙트럼 신호의 양자화, 양자화된 스펙트럼 데이터에 대한 무손실 압축(noiseless coding) 등을 포함하고 있다.

AAC 및 BSAC의 부호화 과정은 무손실 압축을 제외하면 동일한 알고리즘이 적용된다. 입력된 오디오 신호는 filterbank를 거쳐 스펙트럼 성분으로 변환된다. 동시에 지각적 오디오 부호화의 가장 핵심적 과정인 psychoacoustic model을 통해 청각 기관이 소리를 인지하는 정도에 대한 물리적 한계를 표현하는데 사용되는 masking threshold가 계산된다. 이 masking threshold는 스펙트럼 성분의 양자화 과정에서 압축 효율을 높이는데 사용된다. 양자화 과정을 거치기 전에 스펙트럼 데이터는 TNS, PNS, M/S, I/S 등의 압축 도구들을 통해 더욱 압축된다. 최종적으로 양자화를 통과한 스펙트럼 성분은 무손실 압축을 거쳐 부호화된 비트스트림으로 변환된다. <그림 9>에 표시된 것처럼 BSAC과 AAC의 차이는 무손실 압축 방식의 관점에서 설명될 수 있는데, Huffman coding을 사용하는 AAC와 달리 BSAC에서는 arithmetic coding이 사용된다.

<그림 10>는 BSAC과 AAC의 복호화 과정을 나타낸 것이다. BSAC과 AAC의 복호화는 부호화 과정과 마찬가지로 무손실 압축 방식을 제외하면 동일한 과정을 거치게 된다. 입력된 비트스트림은 먼저 Huffman coding 또는 arithmetic coding 등의 무손실 압축 방식 및 역 양자화에 의해 스펙트럼 데이터로 변환된 후 TNS, PNS, M/S, IS 등의 압축 도구들을 통해 부호화 되기 전의 스펙트럼 데이터로 복원된다. 스펙트럼 데이터는 filterbank를 거쳐 시간 영역의 오디오 신호로 변환된다.



<그림 9> MPEG-4 GA 부호기의 구조



<그림 10> MPEG-4 GA 복호기의 구조

## 2. BSAC 알고리즘의 특징

BSAC과 AAC의 주요 차이점은 다음과 같이 요약할 수 있다. Noiseless coding의 수단으로서 AAC에서 사용되는 Huffman coding 대신 BSAC에서는 Arithmetic coding 방식이 적용된다. 또한, Fine grain scalability를 적용하여 채널 당 1 kbps 단위의 계층(layer)으로 부호화할 수 있다.

MPEG-4 BSAC의 대표적 특징 가운데 하나는 하나의 base layer와 여러 개의 enhancement layer를 사용하는 동시에 양자화된 스펙트럼 데이터에 bit-slicing 구조를 사용함으로써, 채널 당 1 kbps 단위의 fine grain scalability를 지원하는 것이다. <그림11>는 bit-slicing의 예를 나타낸 것이다. 무손실 압축을 적용할 스펙트럼 데이터가  $x[0]$ ,  $x[1]$ ,  $x[2]$ ,  $x[3]$ 일 때 이들에 대한 이진 표현을 그림과 같이 표시할 수 있다. Bit-slicing은 MSB부터 LSB까지 각각의 비트 값들을 화살표 방향으로 조합하는 것을 뜻한다. 이 bit-slice들은 MSB를 포함하는 base layer와 그 외의 bit-slice들로 이루어진 여러 개의 enhancement layer들로 나뉘어지고 각각의 layer들에 대해 binary arithmetic coding이 적용된다. 또한 Fine grain scalability를 MPEG-4 시스템을 사용하여 더욱 효율적으로 적용하기 위해 fine grain 오디오 데이터는 large-step layer 단위로 그룹화 될 수도 있다.

## 3. BSAC의 응용 분야

BSAC은 fine grain scalability를 적용함으로써 서비스 음질을 실시간으로 조절하는 것

	MSB				LSB	
$x[0] : 09$	0	1	0	0	1	
$x[1] : 00$	0	0	0	0	0	
$x[2] : 07$	0	0	1	1	1	
$x[3] : 11$	0	1	0	1	1	
	↓	↓	↓	↓	↓	

<그림 11> Bit-slicing의 예

을 가능하게 한다. 또한 이동 통신 네트워크나 DAB(Digital Audio Broadcasting)와 같이 전송 오류에 민감한 채널에서 오디오 신호의 음질을 개선할 수 있다는 것이 또 하나의 특징이라고 할 수 있다. MPEG-4 BSAC은 국내 지상파 DMB의 오디오 부호기 규격으로 채택되어 사용되고 있다.

## V. 결론

본 논문에서는 모바일 방송에서 사용되고 있는 비디오 및 오디오 신호의 압축 방법에 대해 소개하였다. H.264 및 HE-AAC, BSAC으로 대표되는 이 압축 기술들은 기존의 압축 방식에 비해 훨씬 더 나은 압축율을 제공하여 방송과 같이 제한된 대역폭 상황에서 더 나은 품질을 제공할 뿐 아니라 예러 강인성, scalability 등 우수한 기능성을 가지고 있어 모바일 방송 뿐 아니라 다양한 분야에서 채택되어 널리 사용될 것으로 기대된다. 실제로 H.264 비디오 압축 기술은 3GPP에서 MMS(mobile multimedia messaging) 및 화상전화 등 모바일 환경에서의 서비스 뿐 만 아니라, 차세대 HDTV나 DVD 등 고화질용 압축 포맷으로도 채택되어 그 응용분야가 매



우 넓을 것으로 예측된다.

### 참고 문헌

- [1] "Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG Document JVT-I050", 2003
- [2] "Draft ISO/IEC 14496-3 (Audio 3rd Edition), N6795", 2004.
- [3] "International Standard ISO/IEC 14496-3, Information technology - Coding of audio-visual objects - Part 3: Audio", 2001.

### 저자소개



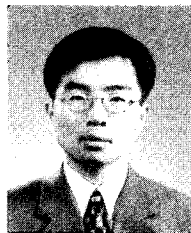
박진수

2003년 - 현재 엠큐브웍스(주) 전임연구원  
주관심 분야 오디오 신호 압축 및 전송, 모바일 멀티미디어, DMB



김지희

2003년 - 현재 엠큐브웍스(주) 전임연구원  
주관심 분야 비디오 신호 압축 및 전송, 모바일 멀티미디어, DMB



이승준

1997년 - 2000년 SK텔레콤 중앙연구원 선임연구원  
1997년 - 1998년 스탠퍼드대학 객원 연구원  
2001년 - 현재 엠큐브웍스(주) 연구소장  
주관심 분야 모바일 멀티미디어, DMB, 멀티미디어 신호 압축 및 전송