

무선 음성 오디오 통합코덱 기술

한국전자통신연구원 김도영, 이미숙, 정해원 광주과학기술원 김홍국

차 례

- I. 서론
- II. 가변비트율 음성코덱 기술 및 표준화 동향
- III. 가변비트율 오디오 코덱 기술 및 표준화 동향
- IV. 가변대역 음성 오디오 통합코덱 기술 및 표준화 동향
- V. 결 론

I. 서론

사람이 서로 현장에서 얼굴을 마주보고 하는 대화나 회의, 연주홀과 같은 공간에서의 현장감 있는 오디오 정보를 시간과 공간의 제약을 받지 않고 그대로 재생하고자 하는 인간의 노력은 전화, 라디오, 그리고 TV 서비스의 성공에 후속하여 데이터 통신망으로 출발하였던 인터넷에서 이를 통합하여 실현하고자 고품질의 미디어 획득과 재생, 부호화 및 복호화, 그리고 통신기술 분야에서 급속히 이루어지고 있다.

특히 BcN(Broadband convergence Network)과 같은 차세대 통신망을 구축하여 지금까지 서로 다른 통신망에서 별개의 서비스로 제공되어 왔던 전화와 데이터 통신, 방송의 3가지 서비스(TPS; Triple Play Service)를 인터넷 프로토콜로 통합하여 제공하기 위한 연구개발은 활발히 진행되어 국내에서도 이미 상용화 진입 준비단계에 있다. 이러한 경향은

〈표 1〉에서와 같이, 세계적으로 초고속 인터넷 포트수의 증가에 따라 인터넷전화 서비스 분야[16]뿐 아니라 향후 오디오를 비롯한 IP TV와 멀티미디어 분야 등에서도 빠르게 확산될 것임을 전망할 수 있게 한다.

〈표 1〉 세계의 초고속인터넷 보급과 인터넷전화(VoIP) 사용자 비율[17]

구 분	2004년	2008년 전망	신장률
초고속인터넷 보급	1억 3500만 포트	2억 6000만 포트	192%
인터넷전화(VoIP) 사용자 비율(%)	2.7%	25.7%	952%

이와 같은 IP 멀티미디어 서비스 제공을 위해 음성과 오디오, 비디오 정보를 응용목적에 맞도록 적절히 부호화하고 복호화하기 위한 코덱 기술은 서비스 품질을 결정하는 매우 주요한 핵심기술이다[2]. 특히 음성과 오디오 코덱은 미디어 신호특성이 매우 유사함에도 불구하고 최근까지도 그 응용서비스가 각기 서

로 달라 기술적으로나 표준체계 측면에서 별개로 취급되어 왔다. 표준화 주관기관도 서로 달라서 음성 코덱은 ITU-T SG16을 중심으로 주로 전화나 인터넷 전화(VoIP), 음성 컨퍼런싱과 같은 G 계열의 대화형(Conversational) 서비스용 표준코덱을, 오디오 코덱은 ISO/MPEG을 중심으로 방송, 다운로드 및 미디어 저장과 같은 MPEG 계열의 방송형(Broadcasting) 서비스 표준코덱을 개발하였다. 그러나 IP TV, 대화형 홈쇼핑 등과 같은 통신과 방송의 융합 서비스, 인터넷 공동게임, 멀티미디어 통신 및 메시징 서비스와 같은 대화형 음성 오디오 통합서비스가 인터넷에서 확산됨에 따라 음성과 오디오 정보를 1 개의 단일코덱으로 개발하기 위한 기술과 국제표준이 급년부터 본격적으로 추진되고 있다. 특히 ITU-T SG16에서는 무선 통신망에서 매우 유효하며 유무선 간 상호연동이 용이한 임베디드형 가변대역 음성 오디오 멀티코덱 표준[1]을, MPEG에서는 scalable 음성 오디오 코덱 표준개발을 진행하고 있다.

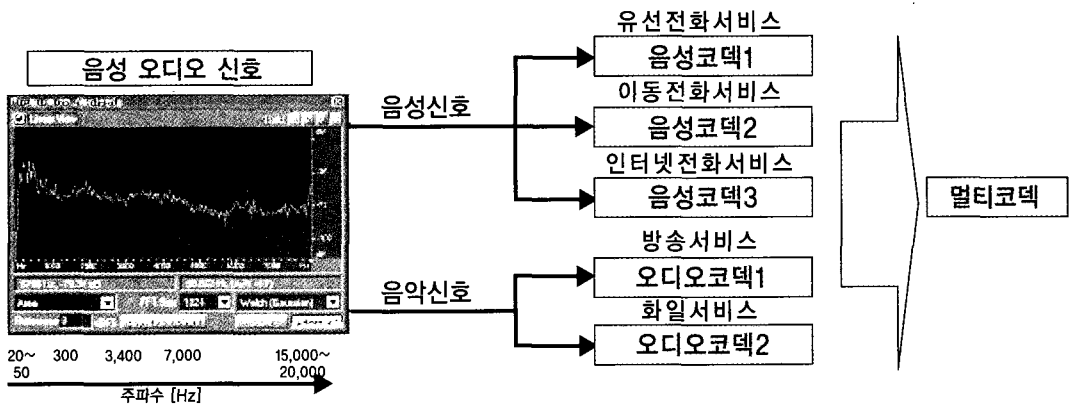
즉, 음성 오디오 멀티코덱은 미디어 신호의 대역폭을 응용목적에 따라 세부 대역별로 분할 가변시킴으로서 다양한 통신망에서 고도의 통신 연동성을 제공

하는 통일된 부호화 및 복호화 기술을 말한다. 즉, 현재의 개별 코덱기술을 개개의 비타민에 비교한다면 멀티코덱 기술은 종합비타민에 비유할 수 있을 것이다. (그림 1)에 기존의 코덱에 비교한 멀티코덱의 개념을 나타내었다.

이에 본 논문에서는 휴대인터넷, 무선랜, 이동통신망과 같은 공중 무선망에서 단일 코덱을 사용하여 음성과 오디오 정보를 함께 부호화할 수 있으며 무선 환경에서 가용 대역폭의 변화에 대해 매우 적응이 용이하고, 장치들의 통신 및 신호처리 능력에 대한 scalability의 장점을 가진 가변비트율 음성 및 오디오 코덱, 그리고 음성 오디오 통합 코덱 기술 및 관련 표준화 동향에 대해 기술한다.

II. 가변비트율 음성코덱 기술 및 표준화 동향

가변비트율 음성 코덱은 협대역(300~3400Hz)과 광대역(50~7000Hz) 음성 신호를 모두 처리할 수 있으며, 임베디드 형태로 구성된 비트스트림을 통해 여



(그림 1) 가변대역 멀티코덱의 개념

러 종류의 비트율을 지원함으로써 다양한 품질의 음성 서비스를 제공할 수 있는 코덱으로 정의한다. 즉, 가변비트율 음성 코덱은 비트스트림이 임베디드 형태로 구성되어 있기 때문에 인코딩된 비트스트림으로부터 얼마만큼의 비트스트림을 추출해서 디코딩하느냐에 따라 재생된 음성의 대역폭과 품질이 달라진다.

기존의 음성 코덱은 고정된 대역폭과 비트율을 제공한다. 따라서 어떤 대역폭을 갖는 신호를 처리하는가에 따라 협대역 코덱과 광대역 코덱으로 분류된다. 3GPP 표준인 AMR-NB(Adaptive MultiRate-NarrowBand)와 AMR-WB(Adaptive MultiRate-WideBand)는 여러 종류의 비트율로 다양한 음질을 제공하고 있지만, 비트스트림이 임베디드 형태가 아니고 각 비트율에 따라 독립된 비트스트림을 가지기 때문에 비트율이 달라지면 인코더의 동작도 달라진다[2].

ITU-T의 WP3/SG16의 Q9와 Q10에서는 가변비트율 음성코덱에 대한 표준화를 진행하고 있다. Q10에서는 프레임 사이즈가 10ms로 적고 이동망과의 연동이 용이하여 인터넷 전화(VoIP) 서비스에서 널리 사용되고 있는 협대역 코덱인 G.729와 비트스트림 레벨의 상호연동성을 제공하는 가변비트율 광대역

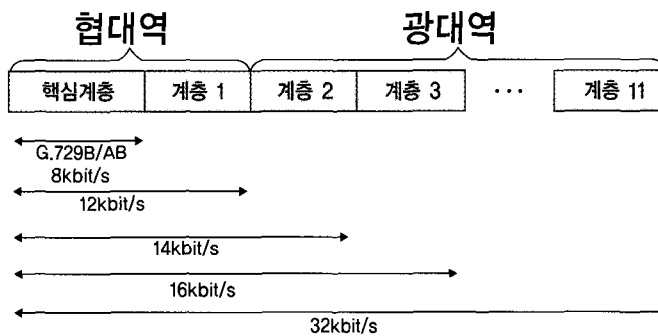
(표 2) G.729EV의 주요 기술규격

코덱 명칭	G.729EV (G.729 Annex J)
표준화 시기	2006. 4 (예정)
전송률(kbit/s)	8, 12, 14~32 가변비트율 (11 모드)
방식	Hybrid
지연(ms)	60 이하
프레임 크기(ms)	20
음질	> G.722 56kbit/s
복잡도	35 WMOPS 이하
VAD/CNG	있음
주 응용분야	유무선 인터넷 전화 및 IP 멀티미디어 통신

(주) VAD는 발생과 목음을 검출, 처리하는 기능, CNG는 Comfort Noise Generation

음성 코덱에 대한 표준화를 2006년 4월을 목표로 진행하고 있으며, Q9에서는 기존 코덱과의 상호연동성보다는 패킷손실 보상기능이 우수한 가변비트율 코덱표준을 2008년을 목표로 표준화 하고 있다. 또한 상위 계층에 오디오 정보처리가 가능한 Super Wideband 계층에 대한 요구사항도 정의하고 있다.

가변비트율 음성코덱의 표준화는 1999년 Qualcomm의 제안으로 Q9에서 시작되었다. 기술적 요구사항과 응용분야를 정의하는 과정에서 기존의 음성코덱과의 상호연동성 제공여부에 대해 많은 논의가 있었다. 특히, 프랑스테레콤에서는 기존 인터



(그림 2) G.729EV 코덱의 비트스트림 구조

넷전화(VoIP) 서비스 가입자들을 수용하면서 광대역 VoIP 서비스를 제공하기 위해서 인터넷전화(VoIP) 서비스에 널리 사용되는 G.729와의 상호연동성을 제공하는 형태의 가변 비트율 음성코덱(G729EV; Embedded Variable Bit Rate)의 표준화를 제안하였으며 ITU-T에서는 <표 2>에 그 주요 규격을 정리한 바와 같이 2006년 4월까지 가변 비트율 음성코덱의 표준화를 완료할 예정이다.

G729EV 코덱은 12개의 계층으로 이루어졌으며 핵심계층은 8 kbit/s G.729와 비트스트림 레벨의 상호연동성을 제공한다. 즉, G.729EV로 인코딩한 비트스트림 중 핵심계층의 비트스트림을 G.729로 디코딩할 수 있다. 계층 1은 12 kbit/s로 동작하며 G.729보다 향상된 협대역 품질을 제공한다. 계층 2(14 kbit/s)부터 계층 11(32 kbit/s)까지는 2 kbit/s 단위로 디코딩 가능하며 광대역 신호를 재생한다. 이러한 계층 구조 때문에 계층수를 늘려갈수록 향상된 음질을 제공하며, 광대역 VoIP 서비스에 G.729를 사용하는 기존 사용자들을 그대로 수용할 수 있다는 장점을 가지고 있다.

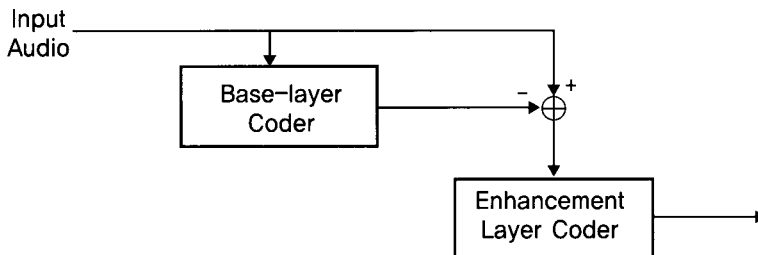
또한 무선환경에서와 같이 휴대형이며 저전력 단말기에 사용할 수 있도록 기존 G.729 코덱에 비해 메모리와 계산량을 최소화하려는 노력을 하고 있다. 이러한 저복잡도, 높은 대역폭 적응성, 기존 인터넷전화 서비스와 별도의 미디어 게이트웨이를 사용하지

않고도 상호연동이 가능한 점, 유선전화에서 사용하는 G.711 코덱보다 우수한 품질 때문에 향후 Wibro와 같은 휴대인터넷, 무선랜, 이동통신망, BcN 등에서 고품질 인터넷전화 서비스에 널리 사용될 것으로 전망되고 있다.

III. 가변비트율 오디오 코덱 기술 및 표준화 동향

가변비트율 오디오 코덱 기술은 가변비트율 음성코덱과 같이 전달된 비트스트림의 크기에 따라 가변적으로 오디오 재생 품질을 얻을 수 있다. 즉, 전송 채널의 상황이 열악할 경우, 적은 비트율로 오디오 신호를 표현함으로써 채널 오류에 의한 오디오 품질의 열화보다는 우수한 오디오 품질을 얻을 수 있다. 한편, 음성코덱과의 차이점은 가변비트율 오디오코덱은 다채널 오디오 신호에 대해 전송 채널의 상황에 따라 다른 수의 오디오 채널을 보낸다는 데 있다[10]. 예를 들면, 스테레오 오디오에 대해, 전송채널이 열악할 경우에는 왼쪽과 오른쪽 중 한 채널만을 보내기도 한다.

가변비트율 오디오 코덱 기술은 크게 multi-layer coding과 bitplane coding으로 나뉘며, 두가지 방식이 혼재되어 구현되기도 한다. 가변오디오 코덱들

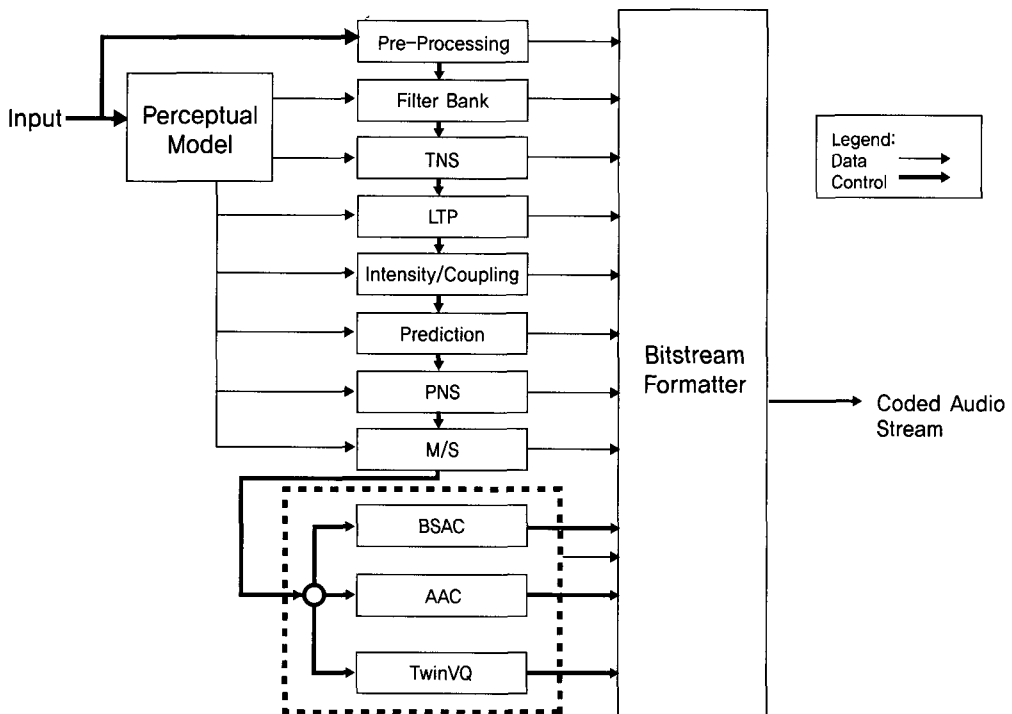


(그림 3) Multi-layer coder 의 개념도

은 계산량과 가변되는 비트의 정도 (granularity)로 비교된다 [3]. Multi-layer coding 기법에서는 오디오 코덱은 기본적인 base-layer coder와 가변비트율에 따른 다층의 enhancement layer coder의 결합으로 구현된다. 현재의 encoding layer는 그 전 layer의 양자화오차를 양자화한다. (그림 3)은 enhancement layer가 한 개인 경우의 multi-layer coder를 보여 준다. 따라서 granularity는 각 layer에서 부가되는 비트율이 되며, 복잡도는 layer의 수에 비례하게 된다. 대표적인 예로는 MPEG-4 AAC (Advanced Audio Coder)로 enhancement layer coder와 base-layer coder는 동일한 방식으로 동작

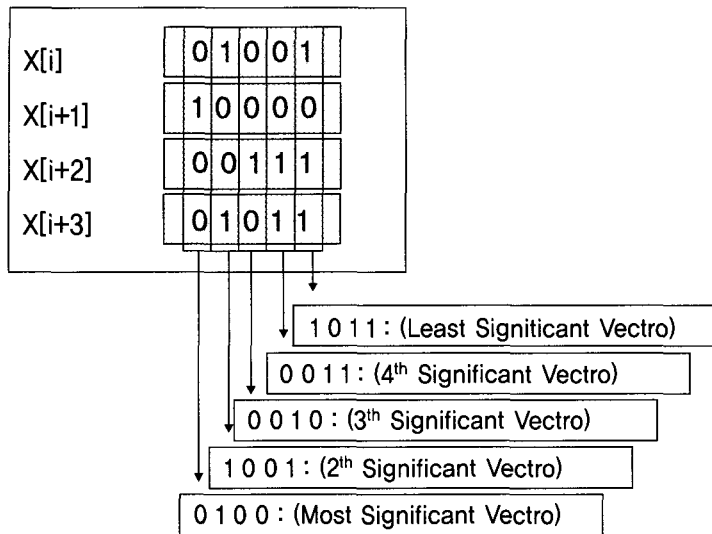
한다[6][7].

반면 bitplane coding 기법은 양자화한 후, 중요도에 따라 비트스트림을 구성한다. 따라서 multi-layer coding에 비해 적은 계산량으로 가변비트율 오디오 코덱을 구현할 수 있다는 장점이 있는 반면, 각 bit의 중요도를 나타내는 bitplane의 설계가 알고리즘의 성능 및 전송률의 overhead와도 관련이 있다. Bitplane 설계방법은 매 분석 프레임마다 가변적으로 bit의 중요도를 판단하는 방법, 프레임에 관계없이 중요도를 미리 정해놓은 fixed bitplane ordering 방법. 전자의 방법은 매 프레임마다 비트의 순서를 부가 정보로 전송해야하는 overhead가 있



(그림 4) MPEG-4 AAC/BSAC/Twin VQ 부호화기 블록도

* TNS (Temporal Noise Shaping), LTP (Long Term Prediction), PNS (Perceptual Noise Substitution), M/C (Multi Channel), TwinVQ (Transform-Domain Weighted Interleave VQ)



(그림 5) BSAC의 기본 개념

는 방식이며, 일반적으로는 dynamic bitplane ordering으로 구현된다. 후자의 경우의 대표적인 예는 MPEG-4의 BSAC (Bit Sliced Arithmetic Coding) coder이다[8].

BSAC은 MPEG-4 Version 2에서 small step scalability를 가지는 서비스에 사용하기 위한 목적으로 표준화 되었다. Large step scalability는 여러 코더를 필요로 하고 8kbit/s 이상의 scalable step size를 가지는데 비해, BSAC은 그 자체만으로 1kbit/s의 step size를 가지는 장점이 있다. BSAC은 (그림 4)에서 보는 것과 같이 MPEG-4 AAC의 Quantization and Coding 부분을 대체하여 사용된다. 또한 large step scalability의 경우 시간영역에서 scalable 코딩 방법인 것과 대조적으로, BSAC의 경우 spectral coefficient를 처리하는 주파수 영역의 scalable 코딩 방법이다.

(그림 5)는 BSAC의 기본적인 개념을 보여 준다[8]. 기존 MPEG-4와 다르게 BSAC에서는 양자화된 스펙트럼 데이터 (spectral data)와 크기정보 (scale

factor)에 대해 코딩을 할 때, AAC에서 허프만 코딩 (Huffman coding)을 사용하는 것과 달리 산술 코딩 (arithmetic coding)을 사용한다. 일반적으로 MSB (Most Significant Bit)의 정보가 LSB (Least Significant Bit)의 정보보다 음질에 많은 영향을 미치기 때문에 환경에 따라 LSB를 전송하지 않더라도 음질은 큰 손상을 받지 않는다. 따라서 BSAC에서는 (그림 5)에서 보는 것과 같이 LSB로 이루어진 최저 중요 벡터(Least Significant Vector)를 지우거나 전송하지 않음을 통해 작은 단위 가변성 (small step scalability)를 제공한다. 64kbit/s로 라디오, 96kbit/s로 CD급의 음질을 제공하는 MPEG-4 BSAC은 채널당 16kbit/s에서 64kbit/s의 가변범위를 가지게 된다. 현재 국내 지상파 DMB 오디오 표준으로 채택되었다. 최근 MPEG에서는 entropy coding 기법을 사용한 MPEG-4 Audio Lossless Coding (ALC)[9]을 표준화를 진행하려 왔으며, 특히 그 안에서 MPEG AAC 기반의 scalable lossless audio coding을 중점으로 표준화를 추진하고 있다.

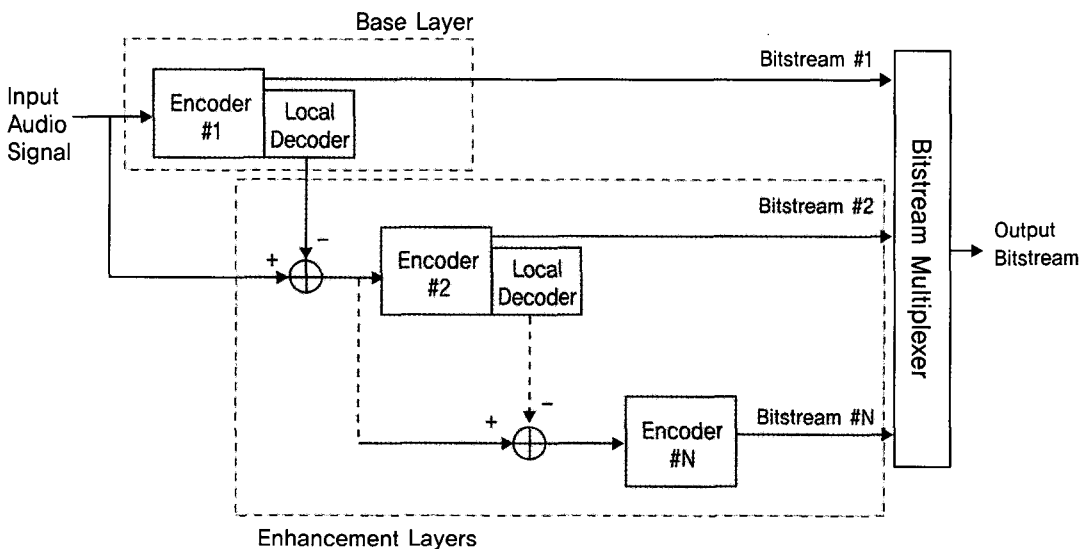
IV. 가변대역 음성 오디오 통합코덱 기술 및 표준화 동향

지금까지 가변비트율을 갖는 음성 코덱과 오디오 코덱 알고리즘에 대해 각각 제 2 절과 제 3 절에서 기술하였다. 최근 들어 네트워크 기술의 발달로 보다 넓은 대역폭의 전송이 가능해지면서 협대역 음성부호화에서 광대역 음성부호화, 그리고 오디오 부호화로의 중요성이 대두되고 있다. 특히, 입력 신호 대역과 네트워크 상태에 따라 서비스 품질을 관리하기 위하여, 가변비트율 및 가변대역 오디오 부호화기의 요구가 증대되어 오고 있다.

일반적으로 기존의 서비스 장치와의 상호연동성을 고려하여 가변대역 오디오 부호기는 기존의 부호화기를 포함하는 임베디드 형태를 갖게 된다. 3400 Hz 정도의 대역폭으로 정의되는 기존의 음성 대역(협대역) 부호화기와의 7.2 kHz까지의 대역폭을 갖는 광대역 음성부호화기와의 임베디드 형태의 부호화기가

개발되어 왔다. 특히, ITU-T 표준 G.722[12]는 quadrature mirror filter를 사용하여 저주파 대역(저대역, 기존의 협대역)과 고주파 대역(고대역)으로 나누고, 고대역에 대해서는 16kbit/s의 ADPCM을 적용하고, 저대역에 대해서는 48, 40, 32kbit/s 중의 하나의 ADPCM을 적용하여 음성을 전송한다. 즉, 저대역 음성에 대해 48kbit/s로 압축을 하지만, 내부 파라미터는 오직 16kbit/s에 해당하는 압축 비트스트림으로 변경한다. 그리고 60 level adaptive quantizer는 48, 40, 32kbit/s의 모든 비트율에서 임베디드 형태로 공유하도록 설계되었다.

최근 들어 ITU-T SG 16의 Question 9에서는 “Variable bit rate coding of speech signal”이라는 연구과제로 임베디드 보코더 (EV: embedded vocoder)에 대한 표준화가 진행 중이다[13]. 이 표준화에 관한 terms of reference (ToR)[13]에 따르면, 개발될 보코더는 협대역 6.4kbit/s를 그 base-layer coder로 하여, 압축률의 증가와 대역폭의 증가가 점



(그림 6) Embedded hierarchical coding의 원리

진적으로 일어나야 한다. 즉, EV는 가변 비트율 기능(variable bit-rate), 가변대역폭의 기능을 가지면서, 음질의 점진적 향상을 보여야 한다. 또한 ITU-T SG 16의 Question 10에서는 “G.729 based embedded variable Bit-Rate (G.729EV) extension to the ITU-T G.729 speech codec”라는 주제 하에 기존의 8 kbit/s G.729 음성부호화기에 협대역에서 광대역의 가변대역폭, 2 kbps 단위의 granularity를 제공하는 G.729EV 부호화기에 대한 표준을 진행하고 있다.

MPEG에서는 음성과 오디오의 scalability를 갖게 하기 위해 MPEG-4 version 1에서 scalability를 제공하는 표준[4]을 정하였다. Large step scalable audio coding 이라고 불리는 이 표준은 계층적 임베디드 코딩(hierarchical embedded coding)을 이용하여 만들어졌다. 이 방식은 여러 개의 코더를 통해 독립적인 비트스트림은 생성한 후, 이 비트스트림을 하나로 묶어서 큰 비트스트림으로 만든다.

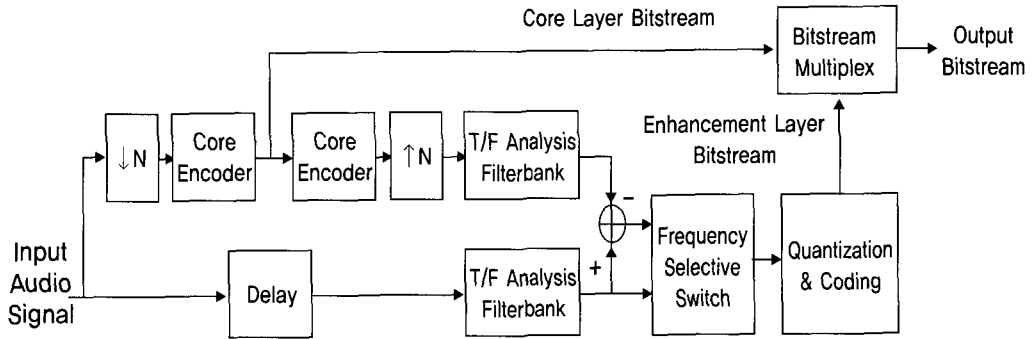
계층적 임베디드 코딩의 원리를 보면(그림 6)에서와 같이, 먼저 Base layer coder라고 부르는 첫 번째 코더는 입력된 오디오 신호 중에 코더에서 처리 가능한 대역의 신호를 인코딩한다. 첫 번째 코더에서 인코딩된 비트스트림은 복원과정을 통해 오디오 신호로 변환되고, 이것을 원본 오디오 신호를 비교하여 첫 번째 코더에서 처리하지 못하는 부분의 데이터를 구한다. 나머지 신호(residual signal)는 Enhancement layer coder의 하나에 해당하는 두 번째 코더의 입력으로 들어가게 된다. 위와 동일한 방법으로 반복적인 작업을 통해서 scalability를 얻게 된다. 하지만 BSAC 방식의 경우 scalable step size가 채널당 1kbit/s 인데 비해, large step scalability 방식을 통해 얻은 scalable step size는 채널당 8kbit/s 이상으로 큰 값을 가진다. 최근에 가변대역 음성 오디오 통합코덱의 하나로써 base layer coder로는 협대역 표준 음성부호화기인 8kbit/s G.729를 사용하고

enhancement layer로는 MPEG-1 Layer 3 (MP3)를 사용하여 전체 비트율을 64kbit/s인 가변대역 음성 오디오 부호화기를 설계하였다 [15]. Segmental SNR과 청취실험을 통해, 입력신호가 주로 음성인 경우는 제안한 방식이 64kbit/s의 MP3에 비해 좋은 성능을 보였으며, 입력신호가 음악인 경우는 그 성능이 비슷함을 확인하였다.

MPEG-4 오디오에서는 2~4 개의 코더를 사용하여 large step scalability를 구현한다. 서로 다른 코더를 사용하여, 가변비트를 뿐만 아니라 음성 및 오디오 대역까지 확장이 가능한 가변대역의 성질을 가진다. 일반적인 MPEG-4 large step scalable 오디오 인코더의 구조는(그림 7)과 같다. MPEG-4 large step scalable 오디오 인코더는 그림에서 보는 것과 같이 non-T/F base layer 코더와 T/F 코딩 방식이 사용된 여러 개의 enhancement layers로 구성되어 있다. non-T/F base layer 코더는 enhancement layer 코더보다 낮은 표본화율에서 동작하며, 코어(Core) 코더라고 부른다. (그림 7)에서의 동작 순서를 설명하면 다음과 같다.

(과정 1) 입력 오디오 신호는 downsampling된 후 core 인코더에 의해 처리된다. 생성된 비트스트림은 비트스트림 다중화 블럭(bitstream multiplexer)에 전해지고, 동시에 core 디코더에 의해 복원된 후 upsampling 과정을 통해 T/F analysis filterbank를 통과하게 된다. 이 때 사용되는 analysis filterbank로는 MDCT가 사용된다.

(과정 2) 입력 오디오 신호는 상기의 과정 1)에서 upsampling까지 걸리는 지연시간만큼 지연을 해준 뒤, 과정 1)과 동일한 analysis filterbank를 통해 주파수 영역



(그림 7) MPEG-4 large step scalable audio encoder의 구조

의 데이터로 변환된다.

(과정 3) 과정 1)과 2)의 차이 (residual 신호)를 구한 후 차이 신호를 Frequency Selective Switch (FSS) 장치를 통해 양자화할 데이터를 구한다.

(과정 4) Spectral coefficient를 양자화하여 비트스트림을 만든다.

(과정 5) 과정 1)에서 만들어진 비트스트림과 과정 4)에서 만들어진 비트스트림을 합친다.

입력 신호의 특성에 따라 base layer 코더가 AAC[7]나 TwinVQ[14]와 같은 T/F 코더가 사용될 수 있다. 이 경우 위의 코어 코더 사용방식과 달리 샘플링 레이트 변환 과정이 빠지게 되고, 오디오 입력 신호를 처리하는데 더 좋은 음질을 얻을 수 있다. 주로 사용되는 코더로는 협대역 CELP, TwinVQ, AAC가 있다. 협대역 CELP 코더는 base layer에서만 사용되고, TwinVQ는 base layer와 enhancement layer에 사용되는데, enhance-ment layer에 사용하려면 base layer에도 TwinVQ인 경우만 가능하다. AAC 코더는 base layer, enhancement layer 모두 사용된다. 실질적으로, 코딩하려는 오디오의 특성에 따라 각 코더의 조합이 다른 성능을 나타내기 때문

에, 오디오의 특성에 맞는 코더의 조합을 사용하는 것이 바람직하다. 예를 들면, 음성신호를 압축하기 위해서는 base layer에서 CELP와 enhancement layer에 AAC 코더를 사용하는 것이 좋은 성능을 보이고, 일반적인 오디오 및 음악신호에서는 base layer에 TwinVQ, enhance-ment layer에 AAC 코덱을 사용하는 것이 좋은 성능을 보인다.

V. 결론

본 논문에서는 음성과 오디오 정보를 통합하여 하나의 코덱을 이용하여 부호화하고 복호화하기 위한 통합코덱과 이 코덱에 가변비트율 특성을 부가하여 대역폭의 변화가 심한 무선통신망에서 효과적으로 활용할 수 있으며 유선통신망에서의 장치나 서비스와의 상호연동성이 우수한 가변대역 음성 오디오 통합코덱 기술개발 및 표준화 동향에 대해 기술하였다.

이러한 가변대역 특성을 가진 음성 오디오 멀티코덱 기술은 IEEE802.11e와 같은 무선랜에서의 서비스 클래스에 따른 QoS 분리처리 기술, 서비스 대역폭 할당을 매우 동적으로 수행할 수 있는 Wibro의 가변대역 특성과 우수한 조화성을 보유하고 있어 향후

Wibro, 무선랜 등에서 인터넷전화(VoIP) 서비스, 오디오 컨퍼런싱, 원격교육, 실시간 인터넷 게임, 대화형 IP TV와 같은 서비스에서 고품질의 미디어 서비스를 제공하는데 널리 사용될 것으로 전망된다.

[참고문헌]

- [1] ITU-T SG16, "Report of SG16 Meeting," 2005.8
- [2] 이미숙, 김도영, "음성코덱의 특성과 협대역 음성 코덱의 표준 기술분석," 주간기술동향 18권 6호, 2003.12
- [3] A. Aggarwal and K. Rose, "Approaches to improve quantization performance over the scalable advanced audio coder," *112nd AES Convention*, Munich, Germany, Reprint 5557, May 2002.
- [4] ISO/IEC JTC1/SC29, "Information technology - very low bitrate audio-visual coding," ISO/IEC IS 14496 (Part 3, Audio), 1998.
- [5] B. Grill, "A bit rate scalable perceptual coder for MPEG-4 audio," *103rd AES Convention*, New York, NY, Reprint 4620, Sept. 1997.
- [6] ISO/IEC JTC1/SC29, "Information technology - generic coding of moving pictures and associated audio," ISO/IEC IS 13818-3 (Part 3, Audio), 1995.
- [7] M. Bosi, et. al., "ISO/IEC MPEG-2 Advanced Audio Coding," *J. Audio Engineering Soc.*, vol. 45, pp. 789-814, Oct. 1997.
- [8] S. Park, Y. Kim, and Y. Seo, "Multi-layer bit-sliced bit-rate scalable audio coding," *103rd AES Convention*, New York, NY, Reprint 4520, Sept. 1997.
- [9] T. Liebchen, "An introduction to MPEG-4 audio lossless coding," in *Proc. ICASSP*, Montreal, Canada, pp. 1012-1015, May 2004.
- [10] C. Faller, R. Haimi-Cohen, P. Kroon, and J. Rothweiler, "Perceptually-based joint-program audio coding," *113rd AES Convention*, Los Angeles, CA, Reprint 5684, Oct. 2002.
- [11] R. Geiger, J. Herre, J. Koller, and K. Brandenburg, "IntMDCT- A link between perceptual and lossless audio coding," in *Proc. ICASSP*, Orlando, FL, pp. 1813-1816, Apr. 2002.
- [12] ITU-T Recommendation G.722, "7 kHz audio-coding within 64 kbit/s," Nov. 1988.
- [13] ITU-T Temporary Document, TD 30 R1(WP 3/16), Report of Q.9/16 Meeting, Jan. 2004.
- [14] A. Jin, et. al, "Scalable audio coder based on quantizer units of MDCT coefficients," in *Proc. ICASSP*, Phoenix, AZ, pp. 897-900, Mar. 1999.
- [15] 이길호, 김홍국, "고품질 오디오 서비스를 위한 스케일러블 오디오 부호화 기법," 제18회 신호처리합동학술대회 논문집, p. 9, 2005년 10월.
- [16] 김도영, 김영선, "인터넷 전화 기술현황 및 전망," 한국통신학회지 제21권 4호, pp. 170~179, 2004년 4월.
- [17] 정보통신부, Conference Proceedings, VoIP Grand Conference, 2005년 7월.



김도영

1985년 성균관대학교 전자공학과 공학사
1987년 성균관대학교 대학원 전자공학과 공학석사
1987년 ~ 현재 VoIP 기술팀장, 광대역통합망연구단 BcN 핵심기술연구그룹 멀티미디어통신연구팀장(책임연구원)

2000년 ~ 현재 VoIP 포럼 H.323 기술분과위원장, VoIP 네트워크 기술분과위원장
2001년 ~ 현재 TTA VoIP PG 부의장, 광대역 음성코덱 WG 의장
관심분야 : 고품질 유무선 인터넷전화, BcN 전화, 고속 실시간 데이터 처리, 멀티미디어 서비스 게이트웨이 등



정해원

1980년 한국항공대학교 항공통신정보공학과 공학사
1982년 한국항공대학교원 항공전자공학과 공학석사
1982년 ~ 현재 한국전자통신연구원 광대역통합망연구단 BcN 핵심기술연구그룹장(책임연구원)
1999년 한국항공대학교원 항공통신정보공학과 공

학박사
2005년 ~ 현재 한국통신학회 교환 및 라우팅연구회 부의장, 대한전자공학회 이사
관심분야 : IPv6, 라우터 및 이더넷 스위치, 무선랜, 홈네트워크링



김홍국

1988년 서울대학교 공과대학 제어계측공학과 공학사
1990년 한국과학기술원 전기 및 전자공학과 공학석사
1990년 ~ 1998년 삼성종합기술원 선임연구원
1994년 한국과학기술원 전기 및 전자공학과 공

학박사
1998년 ~ 2003년 AT&T Labs-Research, Senior Member Technical Staff
2003년 8월 ~ 현재 광주과학기술원 정보통신공학과 부교수
관심분야 : 광대역 음성부호화, 저전송률 오디오부호화, 분산음성인식, robust 음성인식, 대용량 음성인식시스템



이미숙

1991년 호서대학교 전자공학과 공학사
1993년 호서대학교 대학원 전자공학과 공학석사
2001년 한국과학기술원 전기전자공학과 공학박사
2002년 ~ 현재 한국전자통신연구원 BcN 핵심기술연구그룹 멀티미디어통신연구팀 선임연구원
관심분야 : 음성 및 오디오 코덱, 인터넷전화 등