

주 제

임베디드 가변 비트율 광대역 음성 코덱 기술 및 표준화 동향

충북대학교 정규혁, 서상원, 이인성

차례

I. 서론

II. 임베디드 형태의 가변 비트율 광대역 음성 코덱 기술

III. 임베디드 가변 비트율 음성 코덱 표준화(G.729 EV)

V. 결론

I. 서론

현재 국내에는 언제 어디서나 좀더 나은 품질의 이동 통신 서비스를 다양하게 제공 받고자 하는 사용자의 요구가 증가하고 있다. 이동통신 시스템은 소비자의 요구에 부응하기 위해 별개의 형태로 제공되어 왔던 전화, 데이터 통신, 그리고 방송의 3가지 서비스를 IP로 통합·융합시키는 방향으로 발전해가고 있다. 최근 광대역 통신망(BcN : Broadband convergence Network)을 기본 인프라하는 기술과 서비스의 개발이 활발히 진행되어 왔고, 이미 상용화의 단계에 진입하고 있다. 세계적으로도 초고속 인터넷 망 보급의 증가로 인해 인터넷 망을 통한 멀티미디어 서비스 분야의 기술개발이 활발히 연구 될 전망이다[1].

음성 통신 서비스는 IP 주소를 기반으로 종단간의

채널 설정을 통해 정보를 압축하고 패킷화하여 전달하는 멀티미디어 서비스 중 가장 기본적으로 제공되어야 하는 서비스이고, 음성 압축 기술은 서비스 관련 핵심 기술의 주요한 부분을 차지하고 있다[2][3]. 이와 같은 음성 통신에 관련된 인터넷 응용 기술을 VoIP(Voice over IP)라 한다.

현재 VoIP 기술은 완전히 해결되거나 성숙 단계에 있지 않으며, 이와 관련된 표준화 논쟁은 제조업체들의 경쟁적 위치 강화와 맞물려 매우 복잡한 양상으로 전개되고 있다[4]. H.323은 현재 가장 폭 넓게 사용되고 있는 신호계층간의 프로토콜이며 이를 기반으로 하는 시스템을 위한 QoS 제어 기술 관련 표준화 작업이 ITU-T SG16/WP2 Q.24[4]에서 추진되고 있으며, 멀티미디어 신호 압축을 위한 관련 표준화가 Q.6(영상관련), Q.9/10(음성관련)에서 진행되고 있다. 특히 음성 압축 기술과 관련하여 Q.9과 Q.10에서는 가변비트율 음성 코덱에 대한 표준화

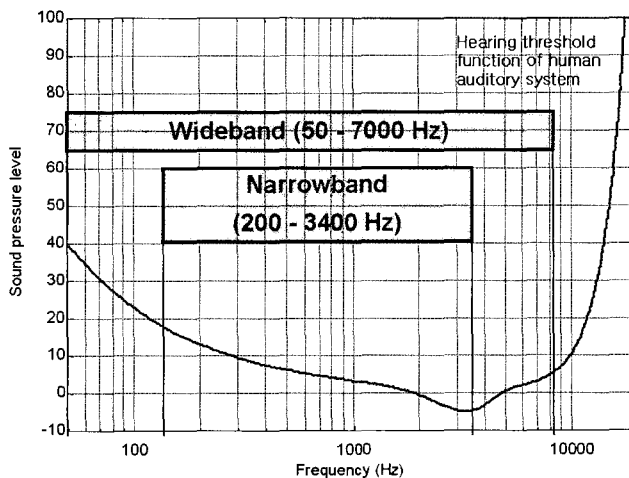
를 진행하고 있다[4] [5] [6]. Q.10에서는 VoIP에서 널리 사용되고 있는 협대역 코덱인 G.729의 확장 버전(G.729 EV/annex J)에 대한 표준화가 진행중이고 현재(2006년 4월) 관련 권고안이 작성중에 있다[7]. Q.9에서는 기존 표준화 코덱과의 상호 연동성을 배제한 코덱을 2008년에 완성할 목표로 표준화하고 있다. Q.9에서는 Q.10의 G.729 EV와의 차별화를 위해 다음의 몇가지 기능을 추가하였다. 손실 환경에서 우수한 보상기능과 음질 향상을 위한 wideband core의 도입, 그리고 상위 계층에 오디오 정보처리가 가능한 super wideband 계층에 대한 요구 사항도 정의하고 있다[8].

본 논문에서는 표준화된 협대역 코덱을 기반으로 광대역 코덱으로 확장시키는 임베디드 광대역 코덱에 대해 기술한다. 2장에서는 기존의 확장 기술들의 개요와 국내외 관련기술 개발 현황을 소개하고, 3장에서는 SG16 Q10에서 마무리 단계에 있는 G.729 EV(Embedded Variable bit rate)에 대해 소개하고자 한다.

II. 임베디드 형태의 가변 비트율 광대역 음성 코덱 기술

1. 임베디드 가변비트율 음성 코덱의 개요

이동통신 및 인터넷 통신의 가장 큰 특징은 넓은 대역폭을 사용한 통신이 가능하게 됨으로써 음성에서 동영상까지 다양한 멀티미디어 데이터를 전송할 수 있다는 것이다. 특히 음성 통신에 있어서 기존의 300-3400Hz의 협대역 음성 신호 뿐만 아니라 50-7000Hz의 광대역 음성 통신도 가능하게 되었다. 50-300Hz의 낮은 주파수의 포함은 자연성과 현재성, 그리고 편안함을 증가시키며 3400-7000Hz까지의 높은 주파수의 포함은 더 정확한 마찰음의 구분으로 인해 명료도를 증가시킴으로써 화자 인식의 용이성을 가져다 준다[9]. (그림 1)은 광대역 및 협대역 신호의 주파수 대역과 인간의 청각인지 시스템을 나타낸다. 광대역 음성 신호는 인간의



(그림 1) 광대역 음성의 주파수 대역폭 및 인간 청각인지 시스템 [10]

청각 인지에서 중요한 3400-7000Hz까지의 대역을 포함하고 있어 더욱 자연스러운 음질을 유지할 수 있다[10].

IP 기반의 패킷 스위칭 방식에서 채널용량이 가변적이기 때문에 고정된 크기의 음성 패킷을 전송하는 것은 비효율적이다. 이를 위해 채널의 전송 상태에 따라 음성 패킷의 양을 조절하여 주어진 채널환경에서 최적의 통신이 가능하도록 하는 광대역 코덱, 즉 AMR-WB(Adaptive Multimedia Rate Wide-band)가 3GPP에서 표준화되었고, 3GPP2에서는 CDMA2000용 광대역 코덱인 VRM-WB코덱을 표준화하였다. 이같은 계층형 단일 코덱은 채널환경에 따라 여러 종류의 비트율로 다양한 음질을 제공하고 있지만 각 레이어들이 독립적인 비트스트림을 가지게 되고 인코더와 디코더의 동작도 각 레이어마다 독립적으로 작동한다.

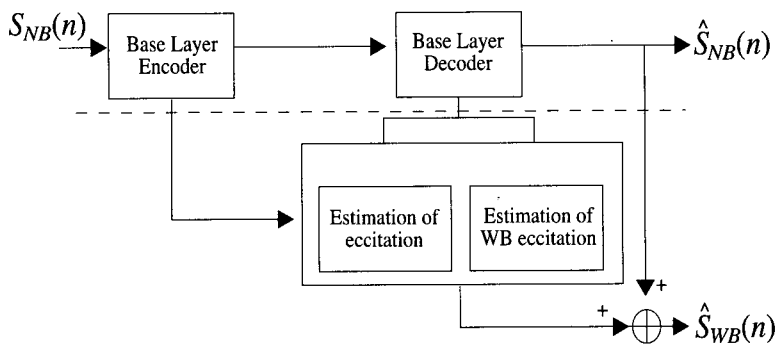
임베디드 가변비트율 음성코덱은 인터넷 전화(VoIP)같은 기존의 서비스 가입자들을 수용하면서 광대역 음성 서비스를 제공하기 위해서 현 시스템에 응용되고 있는 표준화 코덱들과의 상호 연동성을 제공할 수 있는 코덱으로 정의한다. 임베디드 가변비트율 음성코덱의 특징은 기본 계층에 표준 협대역 코덱을 포함시키고 기본 계층이 처리하지 못하는 부분(양

자화 오차나 고대역 신호)들에 대해 추가 비트를 계층별로 할당하여 품질을 향상할 수 있다는 것이다. 따라서 명료도 높은 음질을 제공할 수 있고, 기존의 시스템의 많은 변경없이 사용할 수 있다. 현재의 코덱 기술과는 달리 송수신 쌍방의 단말기의 통신 능력이 다른 경우에도 통신이 가능하다는 장점을 가지고 있다. 이런 특성 때문에 넓은 대역폭을 이용한 통신 서비스가 상용화되기 전부터 많은 연구가 진행되어 왔다[11] - [32].

2. 임베디드 가변비트율 음성 코덱의 기술 개발 현황

전화신호 대역폭에서 50-7000Hz의 넓어진 대역폭으로 증가시킴으로써 음성의 명료도와 자연성을 증가시킬 수 있다. 최근에 50-7000Hz 대역의 신호로 음성 통신을 하는 광대역 음성 서비스에 관심이 모아지고 있다.

전화선 대역의 통신에서 광대역 서비스로 관심이 변화함에 따라 부호화된 협대역 신호로 광대역 신호를 추정하는 연구가 관심을 받아왔다. 이를 대역폭 확장 기술(Bandwidth Extension)이라 하고 음성 뿐 아니라 오디오 코덱에서도 SBR(Spectral



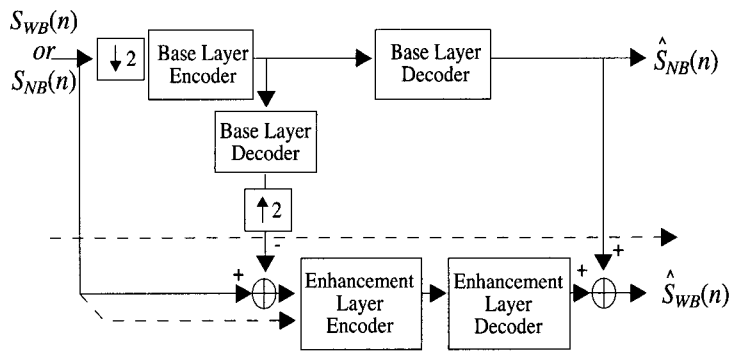
(그림 2) 대역 확장 기술의 기본 블록도

Bandwidth Replica)과 같은 기술들이 사용되고 있다[11]. 협대역 신호와 광대역 신호가 어느 정도로 상관관계가 있다는 가정 하에 출발된 이 연구는 송신측에서의 특별한 변경없이 수신측 단말기상에서의 변경만으로 기존 시스템과 상호 연동이 가능하다는 장점을 갖는다. (그림 2)는 광대역 신호 복원을 위한 대역 확장 기술의 기본 블록도를 나타낸 것이다. 코딩된 협대역 비트스트림에 대한 정보로 스펙트럼 포락선, 에너지, 그리고 여기신호등을 추정한다[12]. 스펙트럼 포락선의 추정은 VQ[13], GMM[14], HMM[15], 그리고 정현파를 이용하는 방법[16]이 있으며 포락선 정보를 협대역 신호의 포락선정보(LPC)를 이용할 경우 여기신호의 추정방법에는 주기적인 임펄스 입력이나 spectral folding[17],

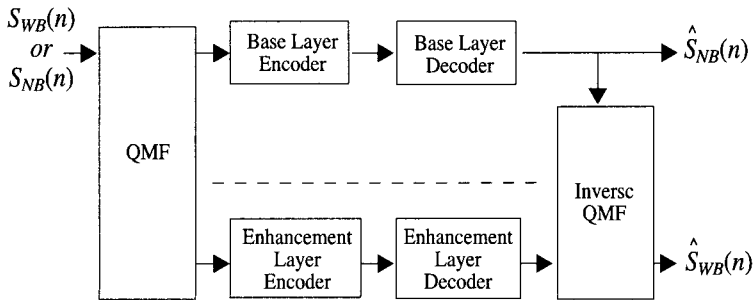
BP-MGN방법[18] 등이 사용되고 있다. 하지만 대역폭 확장 기술은 다양한 음질 서비스의 한계와 광대역 코덱의 음질에 비해 낮은 명료도를 갖는다.

통신 대역폭의 증가로 협대역 신호 외에 추가적인 정보 전송이 가능해 Multi-layer coding기법에 관한 연구들이 표준화 진행과 맞물려 다양하게 진행되었다. Multi-layer coding 기법은 협대역 코덱이 사용된 기본계층(Base layer)과 가변 비트율에 따른 다층의 향상계층(enhancement layer) 코더가 결합되어 구현된다. 현재까지 발표된 연구 결과들은 향상계층의 목적신호가 되는 신호에 따라 크게 SNR scalable 구조[19] - [22]와 대역 분할 구조로 나뉜다[23] - [28].

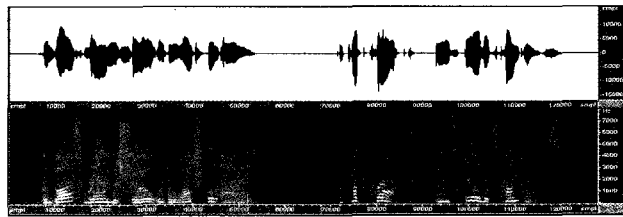
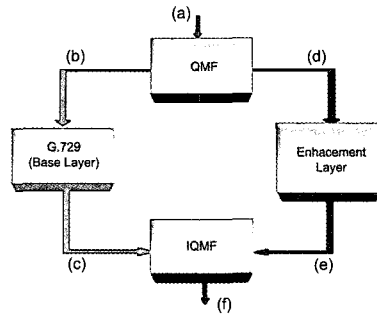
(그림 3)과 (그림 4)는 기본 구조에 대한 블록도



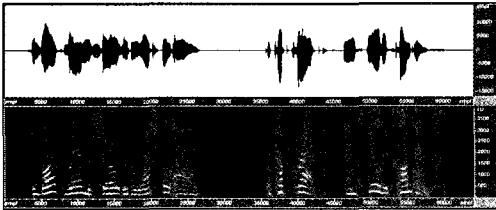
(그림 3) SNR scalable 방식의 가변비트율 코덱의 기본 블록도



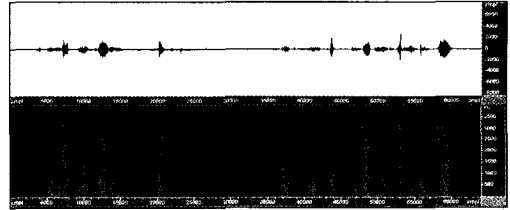
(그림 4) 대역 분할 방식의 가변비트율 코덱의 기본 블록도



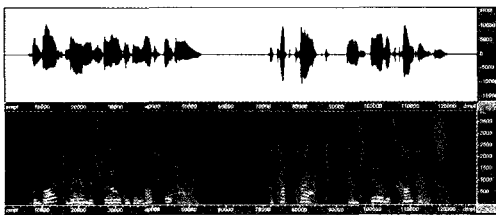
(a) 광대역 입력 음성



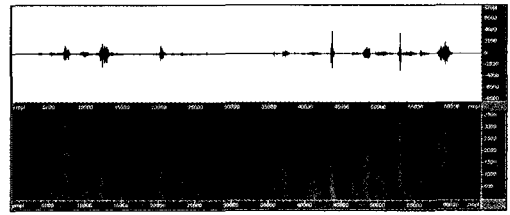
(b) 저대역 입력 음성 (QMF 출력)



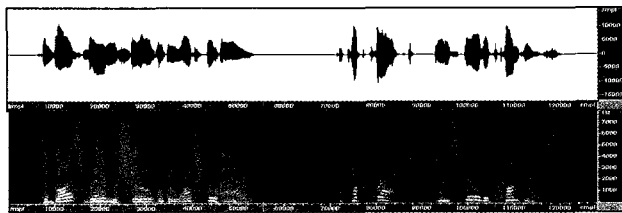
(d) 고대역 입력 음성 (QMF 출력)



(c) 저대역 합성 음성 (G.729)



(e) 고대역 합성 음성 (하모닉-CELP 방법)



(f) 광대역 음성의 최종 출력

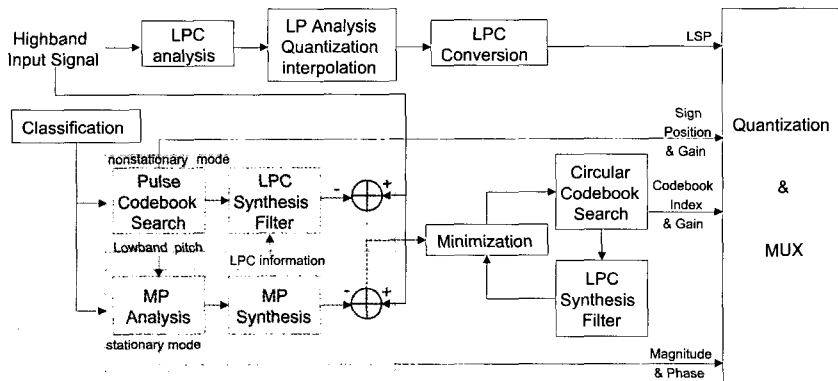
(그림 5) 대역 분할 구조 코덱의 파형들

를 보여준다. (그림 3)에서는 SNR scalable 방식의 가변비트율 코덱의 기본 블록도를 나타내고 있다. 입력신호가 협대역 또는 광대역 신호 모두에서 동작하며 광대역 신호가 입력될 경우 다운 샘플링된 신호가 기본계층에 입력된다. 향상계층을 위한 입력은 기본계층의 양자화 오차나 현재의 코딩 레이어의 양자화 오차를 양자화한다[19] - [22].

(그림 4)에서는 대역 분할 방식 가변비트율 코덱의 기본 블록도를 나타낸다. 광대역 신호의 입력일 경우 (그림 5)와 같이 저대역 (0-4000Hz) 신호와 고대역(4000-8000Hz) 신호로 나누어 지고, 각각 협대역 표준화 코덱과 제안되는 고대역 코덱 방법으로 부호화하게 된다. 이런 구조를 가진 코덱은 sub-band 코덱이라 정의한다.

일반적으로 고대역 신호의 정보는 저대역 신호 정보보다 포먼트 특성과 주기적인 특성이 미약하지만 최종 광대역 음질에 있어서 명료도에서 차이를 보인다. 이에 따라 향상계층, 즉 고대역 신호의 부호화 방법에 대한 Transform 방법[26] - [28], CELP 방법[23] - [25], 정현파 모델 방법[29] 등이 연구되었다. 특히 국내에서 Q.10에 기고서를 제출한 삼성

과 ETRI를 중심으로 몇몇의 대학에서 이와 관련된 많은 연구가 진행되었다[21] [28] - [32]. 많은 코덱에 사용되는 CELP방법은 고대역 신호 부호화에 있어 대역 분할과 합성시 생기는 알리아싱으로 인하여 음질의 왜곡이 발생한다[23]. 이때문에 파라미터 기반의 방식이나 Transform 방법을 대부분 사용한다. 하지만 Transform 방법은 비트율이 충분한 경우 효율적이지만 가변비트율 방식에서 향상계층에 필요한 비트율이 낮아질 경우 주파수 왜곡으로 인한 음질 저하가 발생한다[33]. 이와 같은 문제 해결을 위한 대표적인 접근방법인 DCT 계수의 효율적인 양자화 방법이 다양하게 연구되었다[28] - [32] [34]. 예를 들어 DCT 계수를 하나의 신호로 간주하고 피크점의 위치와 크기를 양자화하는 방법[32]이나 고대역신호를 다수의 밴드로 분할하여 각각의 특성에 적합하게 양자화하는 방법[30] 등이 있다. 이와 다른 새로운 접근방법으로 저전송률 협대역 코덱에서 성능이 입증된 HVXC(Harmonic Vector eXcitation Coder)와 비슷한 하모닉-CELP방법을 이용한 연구도 발표되었다[29]. 고대역에 존재하는 하모닉 성분을 효율적으로 부호화한다면 음질 향상을 이룰 수



(그림 6) 하모닉-CELP방법을 이용한 고대역 향상계층의 블록도[29]

있다는 가정하에 분석-합성 구조의 매칭 퍼슈잇 (Matching Pursuit)과 CELP모형을 다단계 구조를 이용해 고대역 신호를 부호화한다. (그림 6)은 하모닉-CELP방법을 이용한 고대역 향상계층의 블록도를 나타내고, <표 1>은 A/B 테스트 결과를 나타낸다. Test 1에서 하모닉-CELP방법을 이용한 15.8 kbps 광대역 코덱은 G.722(48kbps)보다 여자 음성에 대하여 30.0% 높은 선호도를 보였으며, 남자 음성에서 대하여 15.0%의 높은 선호도를 보였다. 또한 전체적으로 성별에 상관없이 62.3%라는 높은 선호도를 보였다. Test 2에서 하모닉-CELP방법을 이용한 광대역 코덱은 G.722.2(12.85kbps)보다 여자 음성에 대하여 11.2% 높은 선호도를 보였지만, 남자 음성에 대하여는 8.8% 낮은 선호도를 보였고, 전체적으로 1.2% 높은 선호도를 보였다.

(표 1) 하모닉-CELP방법을 이용한 광대역 음성 코덱의 A/B 테스트 결과[29]

Coder	Test 1		Test 2	
	G.722 (48kbps)	하모닉-CELP 방법 (15.8kbps)	G.722.2 (12.85kbps)	하모닉-CELP 방법 (15.8kbps)
Female	35.0%	65.0%	44.4%	55.6%
Male	42.5%	57.5%	54.4%	45.6%
Total	37.7%	62.3%	49.4%	50.6%

2004년에는 ETRI와 삼성종합기술원이 공동으로 음성 코덱 분야에서 국내 최초로 핵심 원천 기술 개발 및 품질검증을 완료한 차세대 광대역 음성 코덱 기술이 TTA 표준으로 확정되었다.

“계층형 다중 비트율 광대역 음성 코덱”이라 불리는 이 코덱은 대역 분할 방식을 기반으로 기본 계층을 기존의 인터넷 전화와 호환될 수 있도록 G.729, G.723.1, G.711을 사용하고 고대역 신호에 대해서 부가적으로 품질 향상 계층 및 광대역 계층을 가진 가변 비트율 기반 방식이다.

대역 분할 방식을 기반으로 저대역에서의 향상계

층이 존재하고, 고대역 신호에서는 MLT 기반 Transform 코딩 방식을 통해 분석되는 기본 계층이 있다[34].

III. 임베디드 가변 비트율 음성 코덱 표준화(G.729 EV)[7]

1. G.729 EV코덱의 개요

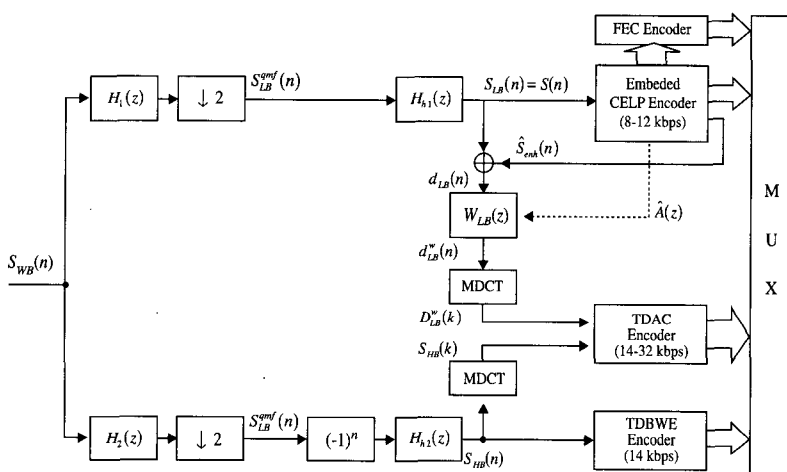
임베디드 가변비트율 음성 코덱 표준화의 시작은 1999년 Qualcomm사의 제안으로 Q.9에서 시작되었다. 응용분야와 주요 기술 규격들을 정의 하는 과정에서 기존의 음성 코덱과의 상호연동성 제공여부에 대해 많은 논의가 있었고, Q.10에서 “Software tools for signal processing standardization activities and maintenance and extension of existing voice coding standards” 라는 주제 하에 VoIP에서 널리 사용되고 있는 H.323 아래의 음성 코덱들의 수정이나 향상에 대한 검토 및 논의가 이루어졌다. 즉, G.722.1 extension, G.729 EV, G.729B enhancement, G.711/G.728/G.727의 revision 들이 있으며 이들 중 France Telecom사에서 제안된 G.729 EV (Embedded Variable Bit rate)의 기고서에 대한 표준화가 현재(2006년) 마무리 단계에 있다.

G.729 EV 코덱은 인터넷 전화(VoIP) 서비스 가입자들을 수용하면서 광대역 VoIP 서비스를 제공하기 위해서 인터넷 전화 (VoIP) 서비스에 널리 사용되는 G.729와의 상호 연동성을 제공하는 형태의 가변 비트율 음성 코덱이다. G.729EV 코덱은 8-12kbps에서 기존의 협대역 통신망과의 호환성을 유지하면서 음질 향상이 가능하고, 14-32kbps범위에서는 2kbps단위로 자유롭게 비트율을 생성하며

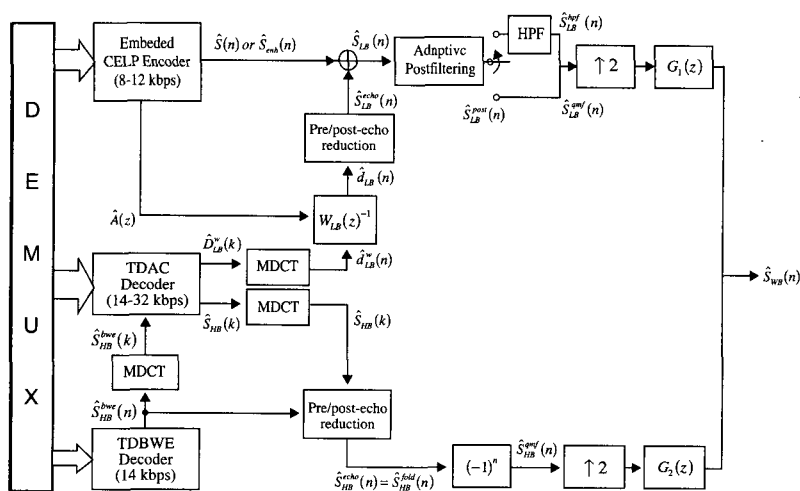
수신측에서는 송신측에서 부호화한 신호의 비트율이 다르더라도 복호화할 수 있다. 이러한 특성은 유무선 융합통신환경에서 허락되어지는 대역폭이 변화하는 경우 통신 서비스가 끊어지지 않고 지속적으로 이루어질 수 있도록 한다.

2. G.729 EV 코덱의 알고리즘

G.729 EV 코덱은 기본계층인 G.729 코덱과 저대역 신호의 양자화 오차를 양자화하는 SNR scalable 구조, 대역폭 확장 기술, 그리고 저대역과



(그림 7) G.729 EV 코덱의 인코더 블록도[7]



(그림 8) G.729 EV 코덱의 디코더 블록도[7]

고대역 신호로 나누어 분석하는 대역 분할 방식이 혼재되어 구현되어 있다. <표 2>는 G.729 EV 코덱의 특성을 정리하였고, (그림 7)과 (그림 8)은 G.729 EV 코덱의 인코더와 디코더 블록도를 나타낸다.

기본적인 동작은 20ms단위로 동작하고 슈퍼프레임(super frame)이라고 정의한다. 결과적으로 G.729 EV코덱의 한 프레임단위에서 10ms로 동작하는 G.729는 두 번을 동작하게 된다. 320샘플단위로 입력되는 광대역 신호는 64tap을 가지는 QMF($H_1(z)$, $H_2(z)$, 그리고 다운 샘플러)를 통과하여저대역 신호와 고대역 신호로 분리된다. 저대역 신호는 50Hz cut-off 주파수를 가지는 고대역 통과 필터를 통해 Layer1(G.729)의 목적신호를 만든다. Layer 1의 LP 잔여신호의 양자화 오차는 합성 필터를 통과시켜 Layer 2의 목적신호를 만든다. Layer 2에서는 tri-pattern 구조의 펄스를 G.729의 고정코드북과 같은 위치에서 검색한다. 즉, 한번의 검색 펄스는 3개의 위치를 가지고 중심의 펄스 코드북은 G.729 고정북과 같은 위치에서 ± 1 의 크기를 가지고 ± 1 의 위치에서 부가적인 펄스는 신호의 특성에 맞게 적응적인 크기를 갖는다.

<표 2> G.729 EV 코덱의 특성 정리

샘플링율	8kHz / 16kHz			
	기본 계층	향상계층		
전송률(kbps)	8	12	14	16-32
모드	Layer1	Layer2	Layer3	Layer4 -Layer12
방식	G.729A	ACELP	TD-BWE	TDAC
		NB	WB	
지연(ms)	25		48.9375	
프레임 크기(ms)	10		20	
음질	G.722 56 kbps 이상			
복잡도	37 WMOPS 이하			
VAD/CNG	있음			

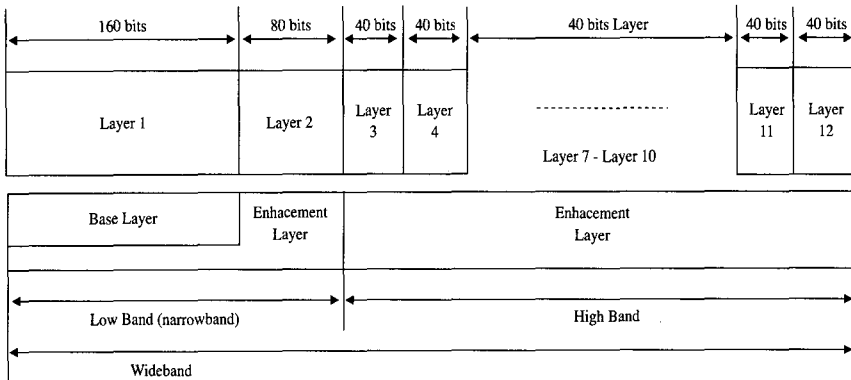
QMF의 출력인 고대역 신호는 스펙트럼 반전을 위해 (-1)ⁿ로 스펙트럼 폴딩(spectral folding)을 한 후 MDCT 변환을 한다. 계수들은 TDAC 인코더와 TD-BWE코더로 입력되어 부호화 된다.

G.729 EV 디코더는 수신된 비트율과 레이어의 수에 따라 다음과 같이 동작한다.

- 8kbps 전송률 (Layer 1) : 기본계층 (G.729) 디코더로 복호화하게 되고, 최종 출력은 고대역 신호를 0으로 채운 후 최종 음성신호를 출력한다.
- 12kbps 전송률 (Layer 1, 2) : 기본계층 디코더와 향상계층 디코더를 통해 복호화 되고, 최종 출력은 Layer 1과 같다.
- 14kbps 전송률 (Layer 1, 2, 3) : 12kbps에서의 복호화된 음성과 TD-BWE 디코더의 출력이 더해져서 최종 음성이 출력된다. TD-BWE의 디코더의 출력은 스펙트럴 폴딩되고, 3000Hz cut-off 주파수를 가지는 고대역 통과 필터를 거치지 않는다.
- 14kbps 전송률 이상 (Layer 1, 2, 3, 4+) : 이전 레이어의 출력(CELP 디코더와 TD-BWE 디코더)과 TDAC 디코더에서 복호화된 신호가 합쳐져 최종 광대역 음성이 출력된다. 각 계층들은 2kbps 단위의 scalability를 유지한다.

(그림 9)는 비트스트림 포맷을 나타낸다. 예를 들어 24kbps의 전송률 일 경우 슈퍼 프레임당 480비트가 전송된다; Layer 1(160bits) + Layer 2(80bits) + Layer 3 (40bits) + Layer 4-8 (200bits).

코덱의 지연은 <표 3>과 같이 48.9375ms이다. 슈퍼 프레임 길이 (20ms), MDCT Look ahead (20ms), G.729 LPC Look ahead (5ms), 그리



(그림 9) G.729 EV 코덱의 비트스트림 포맷

고 QMF 필터 (3.9375ms)를 가진다. 협대역(NB) 모드나 저지연 모드(LOW_DELAY)에서 코덱의 지연은 25ms로 감소된다.

3. G. 729 EV코덱의 음질 테스트

2005년 8월 ITU-T SG 16 회의에서 ETRI, 프랑스테레콤, 지멘스, 마쓰시다, 마인드스피드, 보이스에이지 4개 기관의 코덱이 품질 시험을 인증받았으며 프랑스테레콤의 주도하에 기술통합을 추진하였다. G.729 EV 코덱의 최종 테스트가 06년 3월에 ToR로 합의된 모든 조건을 만족하면서 품질 인증을 받았으며 (그림 10)은 주관적 음질 평가 결과를 나타내었다. (그림 10)은 G. 729 EV 코덱의 음질이 기존의 표준 코덱보다 나은 음질이거나 동등한 음질임을 나타낸다. 또한 12kbps 전송률에서 G.711의 음질과 동등한 결과를 보였고 32kbps에서 현재까지의 표준화된 광대역 코덱들보다 음성 뿐 아니라 음악에서도 좋은 음질을 보였다[35].

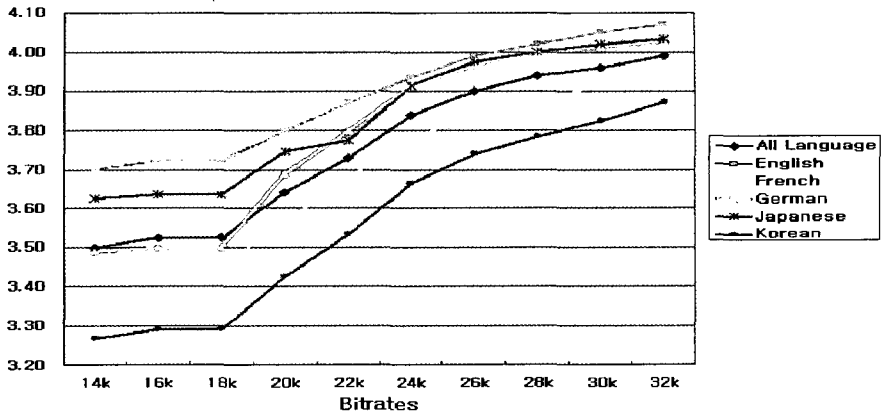
4. G. 729 EV 코덱의 응용

유선 및 무선 음성 통신 뿐만 아니라, 저장장치 등의 다양한 응용 분야에서 기존 협대역 음성 코덱과의 호환성을 지원하는 동시에 품질의 음성 서비제 제공하는 효과를 기대할 수 있다. 특히 인터넷 전화 서비스의 가장 많은 불만 사항은 무엇보다 에코(echo) 문제를 포함한 통화품질에 대한 불만이기 때문에 G.729 EV 코덱의 첫 번째 응용대상이 될 것이다. 또한 광대역 음성신호 뿐 아니라 음악신호에 대해서도 현재까지의 표준화된 광대역 코덱보다 우수한 품질을 제공하기 때문에 인터넷에서의 게임, 비디오, 화상회의, IP-TV 등 다양한 멀티미디어 서비스에의 응용도 가능할 것이다.

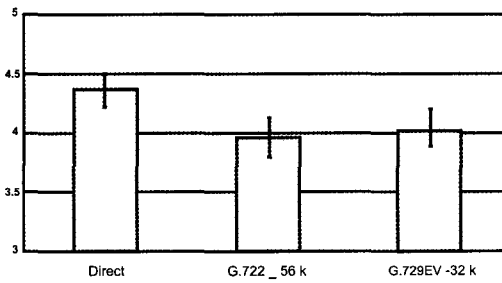
유럽 및 북미의 NGN 인터넷 전화(VoIP) 서비스를 위한 국제표준으로 널리 사용될 것으로 전망되며 국내에서도 관련 국제표준 권리 확보에 성공함으로써 Wibro, BcN 및 홈 네트워크, 유무선 인터넷 전화 서비스에 폭넓게 사용될 것으로 전망된다.

V. 결 론

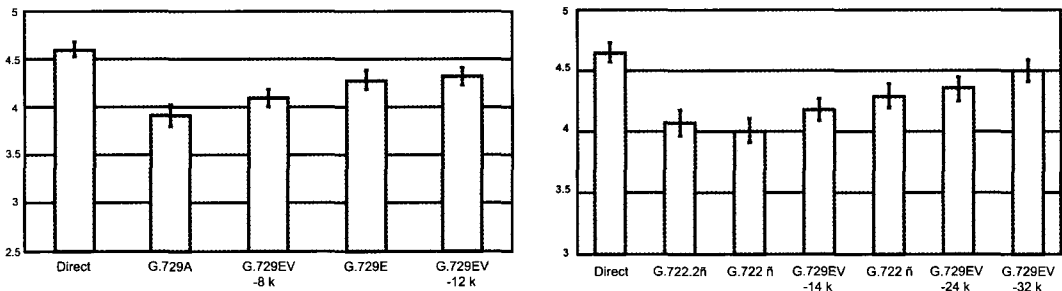
현재 통신과 방송용 단말 장치에서는 수많은 코덱



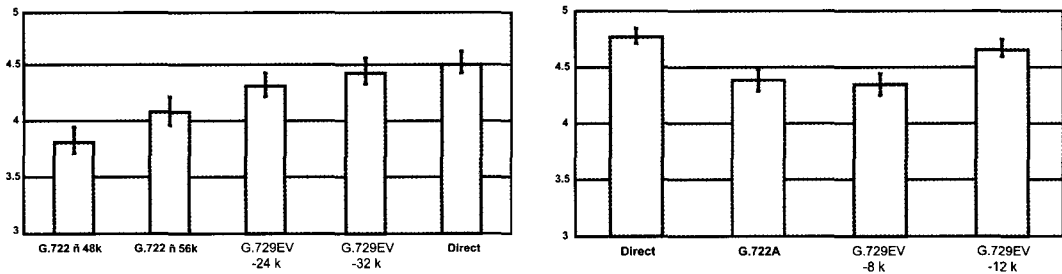
(가) 각 전송률과 음성 DB에 따른 주관적 음질 평가 결과



(나) 32kbps에서의 음악 신호에 대한 주관적 평가 결과



(다) 기존 표준화 코덱과의 주관적 음질 평가 결과 (clean speech / French)



(라) 기존 표준화 코덱과의 주관적 음질 평가 결과 (15dB 배경 잡음 / German)

(그림 10) G. 729EV 코덱의 주관적 음질 평가 결과[35]

들이 사용되고 있고, 응용목적에 맞게 매우 다양한 방식이 사용되고 있다. 이러한 코덱의 다양성은 IP기반의 통신망을 통해 통신하려는 경우 유선과 무선이나 고정과 이동간 융합형 서비스를 위해서 트랜스 코딩을 필요로 한다. 그러나 트랜스 코딩은 품질의 저하를 야기시키고, 단말기나 망장비 개발을 위해 많은 비용이 필요하게 된다.

이런 이유로 음성 신호의 대역폭을 응용 목적에 따라 분할 가변화시킴으로써 다양한 통신망에서도 상호간의 연동성을 제공하는 통일된 코덱이 필요하게 되었고, 최근 ITU-T SG16에서는 압축 기술 분야 중 Q.10의 G.729EV 코덱과 Q.9의 VBR-EV 코덱에 대한 표준화가 진행중에 있다.

본 논문에서는 G.729EV와 이에 관련한 기술현황들을 살펴보았다. 표준화 마무리 단계에 있는 G.729 EV 코덱은 유럽과 북미의 인터넷 전화 서비스를 위한 국제 표준으로 사용될 것으로 전망되며 국내에서도 표준화에 참여에 관한 국제표준 권리를 확보한 상태여서 현재 국가적인 차원에서 추진중인 IT-839전략에서와 같은 분야에서도 다양하게 사용될 것이다.

[참 고 문 헌]

- [1] 정보통신부, Conference Proceedings, VoIP Grand Conference, 2005년 7월
- [2] 이미숙, 김도영, “음성코덱의 특성과 협대역 음성코덱의 표준 기술 분석”, 주간기술동향 18권 6호, 2003. 12
- [3] 김태규, 김도영, 김영선, “VoIP 계층형 광대역 고품질 음성 코덱 협상 처리 기술 분석”, 전자통신동향분석, 2004. 11
- [4] ITU-T WP3/SG16 Q.9, “Report of the Q9/16 meeting”, 2005
- [5] ITU-T WP3/SG16 Q.10, “Report of the Q10/16 meeting”, 2005
- [6] ITU-T WP3/SG16, “Report of the third meeting of Study Group 16(Geneva, 3-13 April 2006)”, Jun. 2005
- [7] ITU-T WP3/SG16 Q.10, “Detailed description of the G.729EV candidate coder”, Apr. 2006
- [8] ITU-T WP3/SG16 Q.9, “Discussion on the remaining issues in Question 9/16”, Apr. 2006
- [9] VoiceAge TM, “Wideband Speech coding standards and applications”, sep. 2001
- [10] Tero Piirainen, “Details in Voice Coding: Advanced Speech Coding for VoIP”, Dec. 2000
- [11] A Laaksonen, “Bandwidth extension in high-quality audio coding”, master thesis Helsinki University of Technology, 2005
- [12] Kallio, Laura, “Artificial Bandwidth Expansion of Narrowband Speech in Mobile Communication Systems”, master thesis Helsinki University of Technology, 2003
- [13] Kallio, Laura, “Artificial Bandwidth Expansion of Narrowband Speech in Mobile Communication Systems”, Helsinki University of Technology, 2003
- [14] K. Y. Park and H. S. Kim, “Narrowband to wideband conversion of speech using

- GMM based transformation”, in Proc. of ICASSP, vol. 3, pp. 1843–1864, 2000
- [15] P. Jax and P. Vary, “Wideband extension of telephone speech using a hidden markov model”, in Proc. of IEEE Workshop on Speech Coding, 2000
- [16] J. Epps, Wideband extension of narrowband speech for enhancement and coding, Ph.D. thesis, University of New South Wales, Sep, 2000
- [17] J. Makhoul and M. Berouti, “High-frequency regeneration in speech coding systems”, in Proc. of ICASSP '79, Volume: 4, Apr 1979 pp. 428 – 431
- [18] Y. Qian and P. Kabal, “Dual-mode wideband speech recovery from narrowband speech”, in Proc. of EUROSPEEH, Geneva, Swiss pp. 1433–1437, Sep. 2003
- [19] Ramprashad, SA., “A Two stage hybrid embedded speech/audio coding structure”, in Proc. of ICASSP '98, pp. 337–340, May. 1998
- [20] Koishida, K., Cuperman, V., Gersho, A., “A 16-kbit/s bandwidth scalable audio coder based on the G.729 standard”, in Proc. of ICASSP '00, pp. 1149–1152, May. 2000
- [21] Kyung Tae Kim, Sung Kyo Jung, Young Cheol Park, and Dae Hee Youn, “A new bandwidth scalable wideband speech/audio coder”, in Proc. of ICASSP '02, pp. 657–660, May. 2002
- [22] Hui Dong, Gibson, J. D., “Structures for SNR scalable speech coding”, IEEE Trans. audio, speech, and Language processing, vol. 13, pp. 545–557, 2005
- [23] Kataoka, A., Kurihara, S., Sasaki, S., Hayashi, S.: A 16kbps/sec wideband speech codec scalable with G.729, in Proc. of EUROSPEECH, pp. 192–195, 1989
- [24] Javovo, R., Montagna, R., Perpsino, F., Sereno, D., “Some experiment of 7 kHz audio coding at 16 kbits/sec”, in Proc. of ICASSP 97', pp. 1491–1494, 1997
- [25] Sung-Kyo Jung, Kyung-Tae Kini, Hong-Goo Kang, “A bit-rate/bandwidth scalable speech coder based on ITU-T G.723.1 standard”, in Proc. of ICASSP 04', pp. 285–288, 2004
- [26] Ramprashad, SA., “Embedded coding using a mixed speech and audio coding paradigm”, International Journal of speech technology, vol. 2, pp. 359–372, 1999
- [27] Kovesi, B., Massaloux, D., Sollaud, A., “A scalable speech and audio coding scheme with continuous bitrate flexibility”, in Proc. of ICASSP 04', pp. 273–276, 2004
- [28] Ji-Hyuk You, Chul-Man Park, Jung-Il Lee, Chang-Beom Ahn, Seoung-Jun Oh, Hochong Park, “Magnitude-Sign Split Quantization for Bandwidth Scalable Wideband Speech Codec”, 6th

- Pacific-Rim Conference on Multimedia. Lecture Notes in Computer Science, pp. 489--499, 2005
- [29] 정규혁, 안영욱, 김종학, 이인성, “광대역 음성 부호화기를 위한 매칭퍼슈잇 알고리즘과 CELP방법을 이용한 고대역 부호화 방법”, 한국음향학회지 제 23권 제 1호, pp. 21-29, 2006
- [30] 송재중, 박호중, 김 무영, 김도석, 김정수, “광대역 신호 압축기를 위한 주파수 대역 특성에 선택적인 양자화 방법”, 음향학회지 제 20권 제 7호, pp. 76-82, 2001
- [31] 이우석, 박호중, 손창용, 이영범, “대역폭 계층 구조의 광대역 음성 부호화기 개발”, 음향학회지 제 23권 제 6호, pp. 481-487, 2004
- [32] 오연선, 신재현, 이인성, “MLT 여기신호를 이용한 광대역 음성 부호화기 설계”, 음향학회지 제 24권 제 5호, pp. 230-237, 2005
- [33] Lajos Hanzo, F. Clare, A. Somerville and Jason P. Woodard, Voice Compression and Communications (John Wiley & Sons Ltd., New York, 2001), pp. 531-564
- [33] B. Kovesi, D. Massaloux and A. Sollaud, “A scalable speech and audio coding scheme with continuous bitrate flexibility”, in Proc. of ICASSP 04', vol. 1, pp. 273-276, 2004
- [34] 정보통신단체 표준 TTAS. KO-01.0045, “계층형 다중 비트율 광대역 음성 코덱”, 2004. 6월
- [35] Christophe Beaugeant, “G.729 EV : An 8-32kbps scalable wideband speech and audio coder bitstream interoperable with G.729”, 2006 IMTC Forum ITU-T Workshop, 2006



정규혁

2004년 충북대학교 전기전자공학 (공학사)
2006년 충북대학교 전파공학과 (공학석사)
2006년 ~ 현재 충북대학교 전파공학과 박사과정
관심분야 : 음성/오디오 부호화, 이동통신, 디지털
신호처리



서상원

2004년 충북대학교 전기전자공학과 학사
2004년 ~ 현재 충북대학교 전파공학과 석사과정
관심분야 : 음성/오디오 부호화, 통신신호처리



이인성

1983년 연세대학교 전자공학과 (공학사)
1985년 연세대학교 전자공학과 (공학석사)
1992년 Texas A&M University 전기공학과
(공학박사)
1993년 ~ 1995년 한국전자 통신연구원 이동통신
기술연구단 선임연구원
1995년 ~ 현재 충북대학교 전기전자공학부 정교수
관심분야 : 음성/영상 신호 압축, 이동통신, 적응필터