

특집논문-05-10-3-01

유무선망에서 사용되는 디지털 음성 부호화 기술 동향

윤 병 식^{a)b)*}, 최 송 인^{b)}, 강 상 원^{a)}

Digital Speech Coding Technologies for Wire and Wireless Communication

Byungsik Yoon^{a)b)*}, Songin Choi^{b)}, Sangwon Kang^{a)}

요 약

디지털 음성 부호화기는 디지털 통신 시스템의 음성 압축 수단으로 이용되어 왔으며 심한 채널에러와 한정된 주파수 자원과 같은 디지털 무선 통신 시스템 환경에 적합하게 더욱 발전해왔다. 또한 디지털 통신 기술의 비약적인 발전에 따라 사용자는 더욱 높은 수준에 해당하는 음성 서비스를 요구하게 되고 그로 인하여 단순히 의사 전달에 해당하는 음성 서비스에서 높은 수준의 멀티미디어 콘텐츠를 수용할 수 있는 음성 및 오디오 부호화기의 기술개발로 급격히 전환 되고 있다. 본 논문에서는 유선 및 무선 디지털망에서 사용되어지는 음성 / 오디오 부호화기의 기술에 대하여 살펴보고 이와 관련된 표준화 활동 및 기술동향에 대하여 알아본다. 또한 향후 유무선 디지털망의 발전에 따른 디지털 음성 / 오디오 부호화기술의 발전 방향에 대하여 언급한다.

Abstract

Throughout the history of digital communication, the digital speech coder is used as speech compression tool. Nowadays, the speech coder has been rapidly developed in the area of mobile communication system to overcome severe channel error and limitation of radio frequency resources. Due to the development of high performance communication system, high quality of speech coder is needed. This kind of speech coder can be used not only in communication services but also in digital multimedia services. In this paper, we describe the technologies of digital speech coder which are used in wire and wireless communication. We also present a summary of recent speech coding standards for narrowband and wideband applications. Finally we introduce the technical trends of next generation speech coder.

Keywords : Digital Speech coding, Vocoder, Codec, Speech coder

I. 서 론

최근 들어 새로운 디지털 무선 통신 시스템의 출현과 유

선 통신 시스템의 비약적인 발전과 더불어 높은 수준의 디지털 음성 부호화기의 개발이 진행되어 왔다. 초창기의 디지털 음성부호화기는 단순히 아날로그 음성의 파형을 샘플링하여 양자화하는 PCM(Pulse Coded Modulation) 방식을^[1] 채택하였으나 근래의 디지털 음성부호화기는 인간의 발성 원리와 청각 특성등을 분석하여 최적의 매개 변수를 찾아 이를 부호화하는 기법으로 발전되고 있다^[2]. 또한 기

a) 한양대학교 전자컴퓨터공학부
Hanyang University, Electronics & Computer Engineering Department
b) 한국전자통신연구원 이동통신연구단
ETRI, Mobile Telecommunication Research Division

존의 음성 부호화기는 대화 위주의 의사 전달이 가능한 협대역 (0.2 ~ 3.4 kHz) 음성 부호화기들이 사용되어 왔으나, 최근의 음성 부호화기는 낮은 수준의 오디오 서비스와 높은 수준의 음성 서비스가 가능한 광대역 (0.05 ~ 7 kHz) 음성 부호화기들이 등장했을 뿐 아니라 기존의 방송 영역에서나 가능한 다양한 멀티미디어 콘텐츠들을 수용할 수 있는 오디오 부호화 영역의 주파수 대역까지 (15 ~ 20 kHz) 처리할 수 있는 부호화기들이 디지털 통신 시스템에 적용되고 있는 실정이다.

디지털 통신 및 방송 시스템이 각기 발전함에 따라 두 시스템간의 구분이 모호해질 뿐 아니라 두가지 시스템이 하나로 융합되는 방향으로 기술 진화가 이루어지고 있다. 따라서 디지털 음성 부호화기는 화상통신, 서킷 통신망, 패킷통신망, 멀티미디어 전송, 오디오 방송 등의 기술적 발전과 높은 품질의 음성 서비스, 자연친화적 음성 서비스 등의 사용자의 요구사항 변화에 맞추어 발전하고 있다.

음성 / 오디오 부호화기 개발에서 중요한 요소는 전송속도, 다양한 환경 하에서의 음성 품질, 음성 부호화 지연시간, 복잡도로 나누어질 수 있다. 전송속도 관점에서 음성부호화기를 본다면 협대역 음성 부호화기는 평균 전송속도 4 kbit/s에서 Toll 품질 (Toll quality) 수준의 음성부호화기들이 개발되었으며, 광대역음성부호화기는 24 kbit/s 급에서 개발이 되었고 최근의 오디오 부호화 기술은 약 48 kbit/s에서 CD 수준의 오디오 품질을 나타내는 부호화기들이 개발을 완료하였다. 음성 품질 관점에서 본다면 다양한 무선 채널 환경 변화와 배경 잡음하에 음성 품질의 현격한 성능 저하를 보이지 않아야 하며 이는 실제 응용분야에 중요한 요소로 작용한다. 부호화 지연 시간은 디지털 음성 부호화를 위하여 입력 데이터를 수집 및 처리하는데 요구되는 시간을 의미한다. 일반적으로 낮은 전송 속도의 음성 부호화기는 높은 지연 시간을 가지며 이러한 지연 시간은 디지털 통신 시스템의 지연 시간과 맞물려 라인 에코와 같은 형태로 전체적인 음성 품질을 저하 시킨다. 복잡도는 음성 부호화기를 실제 DSP(Digital Signal Processor) 등으로 구현할 때 DSP가 요구하는 계산량과 메모리 크기를 의미하며 다양한 기능을 하나의 단말로 융합하려는 추세에 따라 적은 용량의 복잡도를 가지는 음성 부호화기가 바람직하다^[3].

본 논문에서는 다양한 네트워크에서 사용중인 디지털 음성부호화 기술의 동향, 표준화 동향, 향후 발전 방향에 대하여 언급한다. 먼저 현재까지 발전해온 디지털 음성 부호화 기술에 대하여 언급하고 각 기술별 특성과 응용분야에 대하여 설명한다. 특히 ITU-T, 3GPP, 3GPP2 등과 같은 표준화 단체에서 제정된 음성부호화 기술들에 대하여 알아본다. 그리고 마지막으로 현재 개발중인 디지털 음성부호화기의 기술동향과 향후 발전 방향에 대하여 언급한다.

II. 음성 부호화기 표준화

통신 시스템이 발전함에 따라 디지털 음성 부호화기도 더불어 발전해왔다. 이러한 발전은 각종 통신 표준화 단체의 표준 통신 시스템의 일부로 채택되어 왔는데, 특히 근래 들어 디지털 이동통신 시스템의 비약적인 발전에 따라 각 이동 통신 시스템에 적합한 디지털 음성 부호화기가 개발되어왔고 진화되어 왔다.

이와 관련한 대표적인 표준화 단체는 UN산하 각국 통신 표준화를 담당하는 ITU(International Telecommunication Union)가 있으며, 유럽지역의 각종 통신 표준화를 담당하는 ETSI(European Telecommunications Standards Institute) 및 3GPP(3rd Generation Partnership Project)가 있고, 북미지역의 각종 통신 표준화 업무를 수행하는 TIA (Telecommunication Industry Association) 및 3세대 이동통신 시스템 표준화 기구인 3GPP2(3rd Generation Partnership Project 2)가 있다. 또한 그밖의 관련 표준화 기구로 ISO/IEC(International Organization for Standardization / International Electrotechnical Commission), IETF(Internet Engineering Task Force) 등이 있다. 본 단락에서는 표준화 단체별 디지털 음성 부호화기의 종류와 특성에 대하여 알아보고 실제 응용 분야에 대하여 살펴본다.

1. ITU-T 표준화 방식

ITU-T는 1972년 64 kbit/s 전송 속도의 G.711 음성 부호화기를 표준화하였다^[4]. G.711은 8kHz의 표본 주파수로

표 1. ITU-T 표준화 음성 부호화기 특징
Table 1. The Properties of ITU-T Speech coder

음성부호화기	표준화 연도	전송속도	코딩알고리즘	알고리즘 지연	응용분야
G.711	1972	64 kbit/s	PCM	-	유선통신
G.726	1990	16, 24, 32, 40 kbit/s	ADPCM	-	유선통신
G.728	1992	16 kbit/s	LD-CELP	0.625 msec	유선통신
G.723.1	1996	5.6, 6.3 kbit/s	ACELP, MP-MLQ	37.5 msec	VoIP
G.729	1996	8 kbit/s	CS-ACELP	15 msec	VoIP
G.722 (WB)	1988	64 kbit/s	SB-ADPCM	-	유선통신
G.722.1 (WB)	1999	24, 32 kbit/s	ATC	40 msec	VoIP
G.722.2 (WB)	2003	6.6 ~ 23.85 kbit/s	AMR-WB	20 msec	IMT-2000 VoIP

음성을 샘플링하고, 한 샘플당 8bit를 할당하여 PCM 코딩을 수행한다. 이 음성 부호화기는 다시 2가지 방식으로 나누어지는데, 미국과 일본 지역은 μ -law PCM 방식을 사용하고 유럽은 A-law PCM 방식을 사용한다. 이 두가지 방식은 상용 전자 교환기간 음성 통화에서 가장 기본적으로 사용하는 부호화 방식으로 간단하면서 음성 왜곡이 적지만 전송속도가 높은 단점을 가지고 있다.

그후, ADPCM(Adaptive Differential PCM) 방식을 이용한 G.726이 표준화되었다^[5]. 이 음성부호화기는 전송속도에 따라 16, 24, 32, 40 kbit/s로 나누어진다. 1992년에는 16 kbit/s 전송속도의 G.728 LD-CELP(Low Delay - Code Excited Linear Prediction) 음성부호화기를 표준화하였고^[6], 1996년에는 두개의 전송 속도를 갖는 G.723.1 음성부호화기를 표준화하였는데^[7], Algebraic 코드북의 여기신호구조의 5.6 kbit/s 와 MP-MLQ(Multi - Pulse Maximum Likelihood Quantization) 구조의 6.3 kbit/s 전송 속도를 갖는다.

또한 같은해에 8 kbit/s 전송 속도를 가지는 CS-ACELP (Conjugated Structure - Algebraic CELP) 음성 부호화기를 표준화 하였다[8]. 이 방식은 15 msec 의 짧은 알고리즘 지연 시간을 가지고 있으며 최초로 8 kbit/s 급의 전송속도에서 Toll 음질을 보여준 음성부호화기이다. G.729 음성부호화기는 사용되는 응용분야에 적합하게 여러 가지 부록들이 추가로 표준화되었다. 예를들어 계산량이 절반으로 줄어든 G.729A 모델이 있고 음성의 목음 상태를 파악하여 전송속도를 줄이는 G.729B 모델 등이 있다. G.723.1과 G.729 음성 부호화기는 특히 IP 텔레포니 시스템으로 널리

사용되고 있다. 앞서 언급한 음성 부호화기는 협대역 음성 부호화기의 표준화 상황이다.

최근 통신 시스템의 전송 주파수 대역이 넓어짐에 따라 사용자들은 보다 높은 품질의 음성 부호화기 성능을 요구하게 되었다. 광대역 음성부호화기들은 이러한 요구에 부흥하기 위하여 ITU-T에서 표준화로 제정되었다. 1988년, 음성의 주파수 대역폭이 7kHz에 해당하고 전송속도가 64 kbit/s 되는 SB-ADPCM(Sub Band - ADPCM) 음성부호화기가 G.722로 표준화가 되었다^[9]. 1999년에는 24 kbit/s, 32 kbit/s 두개의 전송속도를 가지고 MLT(Modulated Lapped Transform) 방식을 사용하는 G.722.1 음성부호화기를 표준화하였으며 2002년에는 3GPP에서 표준화가 완료된 AMR-WB(Adaptive Multi Rate - Wideband) 음성부호화기의 일부 모드를 G.722.2로 표준화를 하였다^[10]. 광대역 음성부호화기는 상호 통화시 음성의 자연성 (Naturalness) 및 인지성을 (Intelligibility) 높여 자연스러운 대화가 가능할 뿐아니라 낮은 수준의 오디오레벨의 응용분야까지 영역을 확장 시킬 수 있어 향후 통신 시스템의 음성 통화에 높은 이용이 기대된다.

2. 유럽 표준화 방식

1988년 ETSI가 창립된 이래 유럽지역의 통신 표준화는 크게 발전해 왔다. 아울러 2세대 디지털 이동통신 시스템의 개발로 디지털 음성 부호화기가 표준화 단체로부터 주목을 받기 시작하였고 GSM(Global System for Mobile) 이동통신 시스템에 적합한 이동통신용 부호화기의 표준을 위하여

표 2. 유럽 및 북미 표준화 음성 부호화기 특징

Table 2. The properties of European and North America speech coder

음성부호화기	표준화 연도	전송속도	코딩알고리즘	알고리즘 지연	응용분야
GSM FR 코덱	1988	13 kbit/s	RPE-LTP	20 msec	GSM
GSM HR 코덱	1995	3.45 kbit/s	PSI-CELP	24.4 msec	GSM
EFR GSM 코덱	1996	12.2 kbit/s	ACELP	20 msec	GSM, TDMA
AMR	1999	4.75 ~ 12.2 kbit/s	ACELP	25 msec	GSM, WCDMA
AMR-WB	2001	6.6 ~ 23.85 kbit/s	ACELP	25 msec	GSM, WCDMA
IS-54	1989	8 kbit/s	VSELP	20 msec	TDMA
IS-96	1994	8 kbit/s (VBR)	QCELP	24 msec	CDMA
IS-733	1995	13 kbit/s (VBR)	QCELP	24 msec	CDMA
IS-127	1996	8 kbit/s (VBR)	RCELP, ACELP	30 msec	CDMA
SMV	2000	8.5 kbit/s (VBR)	eX-CELP	25 msec	CDMA 2000
VMR-WB	2004	12.65 kbit/s (VBR)	ACELP	35 msec	CDMA 2000

ETSI 산하 SMG(Special Mobile Group) 분과에서 표준화를 담당하였다.

1988년 GSM에서는 RPE-LTP(Regular Pulse Excitation Long Term Prediction)을 GSM 06.10으로 표준화 하였다. 13 kbit/s의 전송속도를 가지며 GSM의 full rate 채널에서 동작이 가능하다. 또한 GSM half rate 채널에 동작이 가능하도록 5.6 kbit/s 전송속도를 가지는 VSELP(Vector Sum Excited Linear Predictio) 형태의 음성부호화기도 GSM 06.20 으로 표준화 되었다^[11]. 그러나 Full rate 및 Half rate 채널에서 음성 부호화기 성능은 기존의 유선 전화 성능보다 크게 떨어져 사용자의 불만이 쌓이게 되었고 마침내 ETSI는 1996년 12.2 kbit/s 전송속도를 가지고 ACELP 알고리즘을 바탕으로 구성된 EFR-GSM (Enhanced Full Rate - GSM) 음성부호화기를 표준화 하였다^[12]. EFR-GSM은 강력한 에러 은닉 기술과 전송효율을 높일 수 있는 VAD(Voice Activity Detection) / DTX (Discontinues Transmission) 등이 장착되어이동통신 시스템에서 유선 전화 서비스 정도의 품질을 제공한다.

3GPP 표준화 단체에서 3세대 이동통신 시스템 개발이 되면서 이 시스템에 적합한 음성 부호화기가 개발 되었다. 3GPP 산하 SA(System Aspect) WG4(Working Group 4)에서는 1999년 유럽형 3세대 이동통신 시스템인 WCDMA(Wideband Code Division Multiple Access)에서 사용될 협대역 음성호화기로 AMR 음성 부호화기를 표준화 하였다^[13]. AMR 음성부호화기는 최소 4.75 kbit/s 의 전송속

도를 갖는 모드부터 시작하여 최고 12.2 kbit/s의 전송속도를 갖는 모드까지 총 8개 모드중 하나로 음성을 부호화한다. 각 부호화 모드는 무선 채널의 상태나 네트워크의 상태에 따라 변경된다. 예를 들어 무선 채널 환경이 나빠거나 혹은 많은 통화 접속이 이루어지는 환경에서는 낮은 전송속도로 부호화가 이루어지고 반대의 경우는 높은 전송속도로 음성을 부호화 하여 전반적인 음성 품질을 높이고 이동 통신 시스템 효율을 높일 수 있는 장점이 있다.

2001년 3GPP에서는 WCDMA Release 4 규격의 일환으로 AMR-WB를 표준화 하였다. 이 음성 부호화기는 그림 1과 같이 주파수 대역폭이 높아 이동통신 시스템에서 획기적인 통화 품질을 개선할 수 있을 뿐 아니라 음성 이외의

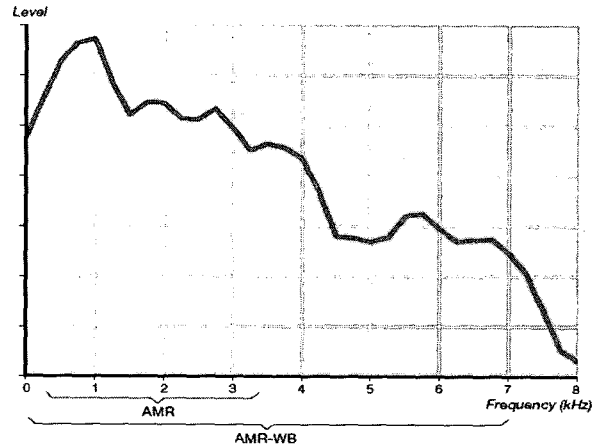


그림 1. NB 및 WB 음성부호화기 주파수 특성

Fig1. The property of bandwidth frequency between NB and WB

다양한 멀티미디어 서비스를 제공할 수 있다. 이 음성 부호화기는 앞서 언급한 ITU-T에서 IP 텔레포니 응용 분야로 표준화가 되었기 때문에 향후 이동통신 시스템과 IP 텔레포니 시스템간의 호환성도 보장된다.

3. 북미 표준화 방식

북미 통신 산업의 표준화를 담당하는 TIA에서는 1989년 북미 TDMA (Time Division Multiple Access) 이동통신 시스템의 음성부호화기로 VSELP(Vector Sum Excited Linear Prediction)을 표준화하였다^[14]. 이 음성부호화기는 8 kbit/s 전송 속도로 동작하며 여기 신호는 128개 벡터들의 합으로 표시하는 특징이 있다. 기술 발전이 이루어짐에 따라 1998년 VSELP 음성부호화기를 대체하는 8kbit/s 전송속도의 ACELP 기반의 IS-641 음성부호화기를 표준화 하였다. 1994년 TIA는 북미 CDMA 이동통신 시스템 개발을 위하여 QCELP(Qualcomm Code Excited Linear Prediction)라는 음성부호화기를 IS-96으로 표준화 하였다^[15].

QCELP는 CDMA 특성을 살리기 입력 음성에 따라 전송 속도를 변화하는 특성을 가진다. 목음이 50% 이상 차지하는 양방향 통화 특성에 따라 목음 구간의 전송속도를 줄임으로써 무선 주파수 자원을 더욱 효율적으로 이용할 수 있다. 그러나 낮은 음성 품질로 인한 문제가 제기되어 개선된 음성 품질을 보여주는 13kbit/s 전송속도의 QCELP 음성 부호화기가 1995년 개발되어 IS-733으로 표준화 되었다^[16]. 또한 8 kbit/s의 전송 속도로 유선 전화 품질의 서비스가 가

능한 EVRC 음성부호화기가 CDMA 이동통신에서 IS-127로 채택되었는데 EVRC 음성 부호화기는 여기신호 발생을 위하여 RCELP(Relaxed Code Excited Linear Prediction) 방식과 ACELP 방식을 혼합하여 채택하였다^[17].

북미 3세대 이동통신 시스템용 음성 부호화기를 개발하기 위하여 3GPP2 TSG-C(Technical Study Group - C) WG1.1(Working Group 1.1)에서 SMV(Selectable Mode Vocoder)를 2000년 표준화 하였다^[18]. 주파수 채널 환경과 네트워크 환경에 따라 CDMA 이동통신 Rate-set I에서 프리미엄 모드, 표준 모드, 이코노미 모드로 나누어지게 되며 SMV 음성 부호화기는 각 모드에 맞게 최적의 음성 부호화 전송속도로 부호화하게 된다. 2004년에는 북미 3세대 이동통신 시스템에 채용할 VMR-WB(Variable-rate Multimode Wideband) 광대역 음성부호화기가 표준화 되었다^[19]. 이 음성부호화기는 CDMA 이동통신 Rate-set II에서 광대역 음성부호화가 가능하도록 설계 되었다. 또한 유럽의 3세대 이동통신에서 채택한 AMR-WB와 상호 호환이 가능하다.

4. 기타 표준화 단체

일본의 경우, PDC(Personal Digital Cellular) 시스템에 적용한 PDC-FR 음성부호화기는 1990년에는 6.7 kbit/s 전송속도의 VSELP를 사용하였으나 1999년부터 ACELP 계열의 PDC-EFR 음성부호화기로 바꾸어 사용하고 있으며, PDC-HR 음성부호화기로는 3.45 kbit/s PSI-CELP(Pitch

표 3. 기타 표준화 음성부호화기 특징
Table 3. The properties of other standardized speech coder

음성부호화기	표준화 연도	전송속도	코딩알고리즘	알고리즘 지연	응용분야
PDC FR 코덱	1990	6.7 kbit/s	VSELP	20 msec	PDC
PDC HR 코덱	1995	5.6 kbit/s	PSI-CELP	40 msec	PDC
PDC EFR 코덱	1999	6.7 kbit/s	ACELP	20 msec	PDC, GSM
IMBE	1993	6.4 kbit/s	MBE	78.75	INMARSAT
HXVC	1999	2.0, 4.0 kbit/s	Harmonic coder	56 msec	Broadcast
LD WB -CELP	1990	3.85 ~ 23.8 kbit/s	CELP	45 msec	Broadcast
MIL-STD-3005	1996	1.2, 2.4 kbit/s	MELPe	240 msec	Military
RFC-3951	2004	13.3, 15.2 kbit/s	iLBC	30 msec	VoIP

Synchronous Innovation CELP)를 사용하고 있다^[20].

국제 해상 위성 (INMARSAT)은 위성을 통한 양방향 통화 서비스를 제공하는데, Inmarsat-B 시스템으로는 16kbit/s APC-MLQ(Adaptive Predictive coder-MLQ), Inmarsat-Aero(Skyphone) 시스템에서는 9.6kbit/s MP-LPC (Multi-pulse-LPC), Inmarsat-M 시스템에서는 4.15 kbit/s IMBE (Improved Multi Band Excitation) 음성부호화기를 사용하고 최근 성능이 향상된 AMBE(Advanced MBE)를 사용하고 있다^[21].

미국 국방성 DOD(Department of Defense) 는 군사용 목적의 표준 음성부호화기로 전송 속도는 낮지만 음성의 명료성이 높은 2.4 kbit/s LPC-10e 음성부호화기(FS-1015)를 1984년에 표준화 하였고 1990년 4.8 kbit/s 전송속도의 DOD CELP(FS-1016)를 표준화하였다. 이후 2.4 kbit/s MELPe(Enhanced Mixed Excitation Linear Prediction) 를 표준화 하였다^[22]. ISO/IEC의 MPEG(Moving Picture Experts Group) 오디오 분과에서는 1999년 6 ~ 64 kbit/s 해당하는 다양한 오디오 및 음성 부호화기에 대한 표준화를 완료하였는데 주요 응용분야는 디지털 방송용 오디오 전송을 목표로 개발되었다. 따라서 양방향 통화 서비스를 위한 기존의 음성부호화기와 비교하여 몇가지 차이가 존재한다. 일단 다양한 음성 및 오디오 부호화기를 구성하여 네트워크의 특성에 맞게 전송율이 가변적으로 변화하는 전송 속도 변환이 가능하여야 하며, 아울러 대역폭 변환도 가능

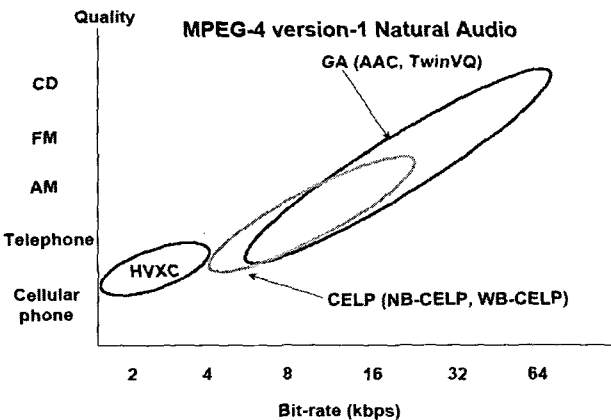


그림 2. MPEG 음성부호화기 전송속도별 성능
Fig 2. The performance of MPEG speech coders

하여야 한다. 그림 2는 각 음성 및 오디오 부호화기의 전송 속도별 품질을 나타낸다. MPEG-4 오디오 분과에서는 이러한 특성을 만족하는 3개의 음성부호화기를 개발하였는데. 유성음의 하모닉성분을 여기신호로 이용하는 HVXC (Harmonic vector excitation coding)과 협대역 음성부호화기인 NB-CELP(Narrowband - CELP)와 광대역 음성부호화기인 WB-CELP(Wideband - CELP)를 표준화 하였다^[23].

2004년 IETF(Internet Engineer Taskforce Team) AVT (Audio Video Transport) WG에서 iLBC (Internet Lowbit Rate Codec)을 표준화 하였다^[24]. iLBC는 기존의 CELP 음성부호화기가 이전 프레임의 메모리들을 사용함으로써 생기는 에러의 영향을 최소화하기 위하여 이전 프레임의 메모리를 사용하지 않고 현재 프레임에 대해서만 부호화함으로써 인터넷상에서 빈번히 일어난패킷손실이나 지연에 매우 강인하게 설계되어 향후 VoIP등의 인터넷 멀티미디어 응용에 사용되어질 예정이다. 그림 3은 패킷 손실에 대한 iLBC의 성능을 보여준다.

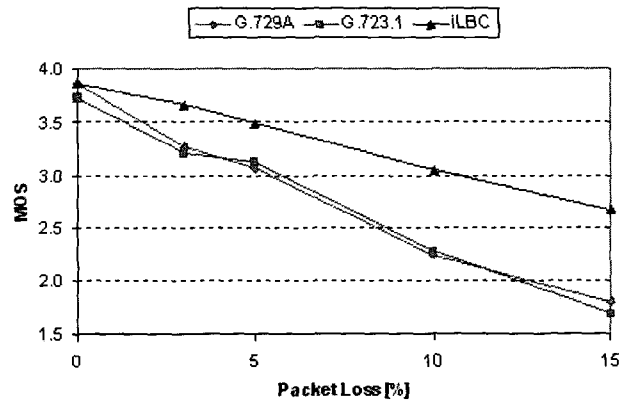


그림 3. iLBC 음성부호화기의 패킷손실 성능
Fig. 3. The performance of iLBC in packet loss

Ⅲ. 음성부호화기 발전 방향

통신 기술의 발전과 더불어 최근의 음성부호화기들의 개발 방향은 유무선 통신 시스템에서 고품질의 멀티미디어

응용 서비스 혹은 고품질의 양방향 통화 서비스가 가능한 음성부호화기들이 주목받고 있다. 또한 무선채널 상황, 네트워크 자원 등의 네트워크의 특성에 적응하는 음성부호화기들의 개발이 이루어지고 있다.

현재 ITU-T SG16(Study Group 16) Q.9(Question Group 9)에서는 패킷통신망에서 적응적으로 전송 속도 변화가 가능하고 서로 다른 네트워크에서 호환이 가능한 VBR(Variable Bit Rate) 음성부호화기의 표준화를 진행 중에 있다.

특히 QCELP나 SMV와 같이 음성의 묵음에 따른 전송 속도를 변화하여 가변 평균 전송속도를 정하는 MSC-VBR(Multi-Rate Source Controlled - VBR) 음성 부호화기와, 네트워크의 상태나 자원에 따라서 전송속도를 변화하는 EV(Embedded VBR) 음성부호화기가 각각 표준화 중에 있다. 음성부호화기들에 대한 광대역 음성품질 보장을 포함한 요구사항에 대한 정의가 완료되었고 특히 G.729를 코어 코덱으로 사용하고 비트율 및 대역이 가변적으로 변화하는 G.729EV에 대한 표준화가 진행중에 있으며 이 표준화에 참여한 업체들로부터 제안된 음성부호화기의 성능을 검증하고 선택하기 위한 작업들이 2005년 현재 수행중이다. 3GPP에서는 SA WG4에서 PSS(Packet Switched Streaming), MMS(Multimedia Messaging Service), MBMS(Multimedia Broadcast Multicast Service)용 음성 및 오디오 부호화기에 대한 표준화를 수행하였다^{[25][26]}.

최근 두개의 음성/오디오 부호화기에 대한 표준화가 완료되었는데, AMR-WB+ 음성부호화기는 ACELP와 TCX(Transform Coded Excitation) 기술을 이용하여 6-48 kbit/s의 전송속도와 20 kHz의 주파수 대역을 갖는 음성 혹은 오디오에 대한 부호화를 수행할 수 있으며 기존의 AMR-WB와 상호 호환이 가능하다. Enhanced AAC+(Advanced Audio Codec plus)는 SBR(Spectral Band Replication) 기술과 BCC(Binaural Cue Coding) 방식을 이용하여 최고 48 kHz로 샘플링된 오디오 신호를 48 kbit/s 전송 속도로 부호화가 가능하다.

OMA(Open Mobile Alliance) BCAST(Broadcast) STI(Standard Transcoding Interface) 분과에서는 다양한 네트워크에서 멀티미디어 데이터들의 호환을 위하여 표준

화된 멀티미디어 변환 인터페이스 및 변환 엔진에 대한 연구가 진행 중에 있다. 현재 서로 다른 이 기종망 간에 영상, 오디오, 음성의 효율적인 전송 및 변환을 위한 표준화 작업이 진행 중에 있다.

IV. 결 론

본 논문에서는 현재까지 표준화된 음성 및 오디오 부호화기에 대하여 그 특징과 응용분야에 대하여 설명하였다. 또한 ITU-T, 유럽, 북미 그리고 다른 기타 단체에서 표준화된 음성부호화기들에 대하여 살펴보았다. 각 음성부호화기는 그 성능과 전송 속도 그리고 통신 시스템의 특성에 맞게 최적으로 설계 되었으며 성공적으로 사용되어 왔다. 그러나 앞으로 도래할 차세대 VoIP, 멀티미디어 스트리밍, 3G 및 4G 시스템은 좋은 품질의 음성 및 오디오서비스는 물론이고 영상과 데이터를 포함한 다양한 형태의 서비스를 제공할 것이므로 유선망을 이용하는 많은 가입자들이 IMT-2000과 같은 무선망을 이용하게 될 것이다.

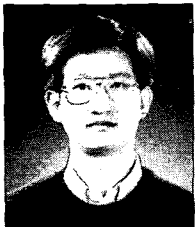
현재 IMT-2000 서비스를 위해 할당된 주파수는 송수신 합쳐 230MHz로서 많은 대역이 할당되어 있으나, IMT-2000 시스템을 사용하게 될 가입자 수가 매우 많을 것이고 제공될 서비스도 상당히 다양할 것이다. 그리고 패킷 네트워크를 통한 전화 서비스로서 현재 각광을 받고 있는 VoIP 시스템의 다음 버전인 차세대 VoIP 시스템의 가장 중요한 특징 중 하나는 고품질 음성 및 오디오이므로, face-to-face 음질에 해당하는 음성 서비스와 CD음질 오디오의 제공은 IMT-2000 시스템 및 차세대 VoIP에서 필연적이다.

특히 VoIP 응용에서는 scalability 기능을 가지면서 자연성과 명료도가 우수한 음성 및 오디오 부호화기가 요구된다. 그러나 국내에서 수행중인 네트워크의 트래픽 상태와 단말의 처리속도에 따라 최적의 음질을 제공하는 embedded 개념의 음성 및 오디오 부호화 기술에 대한 연구는 미미한 실정이므로, 관련 핵심 기술 도입시 외국 방식 및 핵심부품을 그대로 사용하게 되어 많은 특허료와 개발비용이 지불될 것으로 생각된다.

참고 문헌

- [1] J. F. Boag : "The end of the first pulse-code modulation era in the UK" The Post Office Electrical Engineers Journal, 71, Part 1, pp 2-4, April 1978
- [2] B.S. Atal and M. R. Schoeder, "Stochastic coding of speech at very low bit rates," in Proc. Int. Conf. Comm., 1984, pp. 1610-1613
- [3] W.B. Kleijn and K.K. Paliwal, Speech Coding and synthesis, Elsevier, 1995.
- [4] ITU-T Rec. G.711, "Pulse code modulation (PCM) of voice frequencies," Nov. 1988.
- [5] ITU-T Rec. G.726, "40, 32, 24, 16kbit/s adaptive differential pulse code modulation (ADPCM)," Dec. 1990.
- [6] ITU-T Rec. G.728, "Coding of speech at 16kbit/s using low-delay code-excited linear prediction," 1992.
- [7] ITU-T Rec. G.723.1, "Dual-rate speech coder for multimedia communications transmitting at 5.3 and 6.3 kbit/s," May 1996.
- [8] ITU-T Rec. G.729, "Coding of speech at 8 kbit/s using conjugate-structure algebraic code-excited linear prediction (CS-ACELP)," May 1996.
- [9] ITU-T Rec. G.722, "7 kHz audio-coding with 64 kbit/s," Nov. 1988.
- [10] ITU-T Rec. G.722.2 "Wideband coding of speech at around 16 kbit/s using Adaptive Multi-rate Wideband (AMR-WB)," Jan. 2002.
- [11] Rec. GSM 06.20, "European digital cellular telecommunications system; Half-rate speech; Part 1: Half rate speech processing functions," Dec. 1995.
- [12] Rec. GSM 06.60, "Digital cellular telecommunications system(Phase 2+) (GSM); Enhanced Full Rate(EFR) speech transcoding," July 1999.
- [13] 3GPP Rec. TS 26.090 "AMR speech coder; Transcoding functions," Jun. 2002.
- [14] TIA/EIA IS-54, "Cellular system dual-mode mobile station-base station compatibility standard," May, 1990.
- [15] TIA/EIA IS-96 "Speech Service Option Standard for Wideband Spread Spectrum Digital Cellular System," Apr. 1994.
- [16] TIA/EIA/IS-733, "High rate service option 17 for wideband spread spectrum communication system," Mar. 1997.
- [17] TIA/EIA/IS-127, "Enhanced variable rate codec, speech service option 3 for wideband spread spectrum digital systems," Jan. 1997.
- [18] TIA/EIA/IS-893, "Selectable mode vocoder service option for wideband spread spectrum communication system," Jun. 2005.
- [19] TIA-1016. "CDMA2000 wideband codec algorithm description," Oct. 2004.
- [20] K. Mano et al. "Design of a pitch synchronous innovation CELP coder for mobile communications," IEEE J. Select. Areas Commun., vol. 13, no. 1, pp. 31-41, 1995.
- [21] S. Dimolitsas et al. "Evaluation of Voice Codec Performance for the Inmarsat Mini-M System," Proceedings, 10th Int. Digital Satellite, Brighton, England, May, 1995.
- [22] L. Supplee, et al. "MELP: The new federal standard at 2400 bps," ICASSP-97, Munich, Germany.
- [23] ISO/IEC 14496-3, Information technology coding of audio-visual objects, Part 3: Audio, Subpart 2: Speech Coding HV C, Nov. 1998.
- [24] S. V. Andersen et al. "iLBC - a linear predictive coder with robustness to packet losses," Speech Coding, IEEE Workshop Proceedings, pp. 23-25, Oct. 2002.
- [25] 3GPP Rec. TS. 26.290, "Audio codec processing functions; Extended Adaptive Multi-Rate - Wideband (AMR-WB+) codec; Transcoding functions," 2004.
- [26] 3GPP Rec. TS. 26.401, "General audio codec audio processing functions; Enhanced aacPlus general audio codec; General description," 2004.

저자 소개



윤 병 식

- 1990년 2월 : 경북대학교 공과대학 전자공학과 졸업 (학사)
- 1990년 2월 : 경북대학교 공과대학 전자공학과 졸업 (학사)
- 1992년 1월~현재 : 한국전자통신연구원 이동통신연구단 선임연구원
- 2004년 3월~현재 : 한양대학교 전자컴퓨터 공학부 박사과정
- 주관심분야 : 음성 코딩, 멀티미디어 신호처리, 이동통신 시스템

저 자 소 개



최 송 인

- 1982년 2월 : 광운대학교 응용전자공학과 졸업 (학사)
- 1987년 2월 : 광운대학교 전자계산기공학과 졸업 (석사)
- 1982년 7월~현재 : 한국전자통신연구원 이동통신연구단 책임연구원 (현 융합단말연구팀장)
- 주관심 분야 : 음성 신호처리, 영상 신호처리, 이동통신 시스템



강 상 원

- 1980년 : 한양대학교 전자공학과 졸업 (학사)
- 1982년 : 서울대학교 전자공학과 졸업 (석사)
- 1990년 : Texas A&M대 전기공학과 졸업 (박사)
- 1982년~1994년 : 한국전자통신연구소 근무
- 1994년~현재 : 한양대학교 전자컴퓨터공학부 교수
- 주관심분야 : 음성코딩, 신호처리 및 디지털통신