VoIP 계층형 광대역 고품질 음성 코덱 협상 처리 기술 분석

A Call Processing Method for the VoIP Wideband High Quality Speech Codec

강태규(T.G. Kang)	VoIP서비스기술팀 책임연구원
김도영(D.Y. Kim)	VoIP서비스기술팀 책임연구원, 팀장
김영선(Y.S. Kim)	통합서비스연구그룹 책임연구원, 그룹장

유선 네트워크, 무선 이동통신 네트워크, 인터넷 등을 통합하는 유무선 통합 네트워크(BcN)에서는 VoIP 기술을 사용하게 될 것이다. TTA 표준으로 2004년 7월에 제정된 VoIP 계층형 광대역 고품질 음성 코덱은 핵심계층에 G.711, G.723.1, G.729를 사용하므로 10종의 PT를 설정하여 코덱을 협상한다. 이로 인하여 자기자신의 코덱 이외에도 G.711, G.723.1, G.729 등과 상호 호환이 되는 장점을 갖는다. 본 고는 신규로 제정된 VoIP 계층형 광대역 고품질 음성 코덱을 네트워크에서 사용할 수 있도록 호 처리에 대한 표준화를 추진하여야 하는데 이를 위한 표준 기술을 설명하고, 코덱과 호처리 관계 및 표준화 기술을 근거로 한 코덱 협상 처리 기술을 설명한다. 코덱 협상 처리 기술로서 PSTN/MSC 연동 코덱 협상 방안과 All IP 코덱 협상 방안으로 구분하였다. All IP 코덱 협상 방안에서는 발신, 착신, MGC, 착신서버에서 호환성을 위한 호 처리 기능을 제공한다. 본 고의 호 처리 기술을 적용하면, VoIP 계층형 광대역 고품질 음성 코덱은 기존 네트워크 장치 기능을 수정하지 않고 사용할 수 있다.

I. 서 론

인터넷 전화는 유선 PSTN(Public Switched Telephone Network) 및 무선 이동통신들과 SIP/SDP (Session Iniation Protocol/Session Description Protocol) 호 설정 절차를 이용하여 연동한다[1]-[3]. 2003년 2월 현재 PSTN 가입자 수 2,322만 명, 이동통신 PCS(Personal Communications Service) 가입자 수 3,241만 명, 인터넷 사용자 수 2,627만 명이다. 네트워크는 유선, 무선, 인터넷 가입자 수가 동등하게 구성되었고, 이들 네트워크들을 통합한 시너지 효과를 위하여 BcN(Broadband convergence Network) 기술을 개발하고 있다. VoIP (Voice over Internet Protocol) 기술은 유선, 무선, 인터넷들과의 연동하는 단계에서 통합하는 단계로 발전시킬 수 있는 BcN의 핵심 기술이 될 것이다.

한국전자통신연구원 VoIP 서비스 기술팀에서는 VoIP의 문제점이었던 인터넷 음성 전화 음질을 개선하기 위하여 VoIP 계층형 광대역 고품질 음성 코덱(Wideband Codec for Internet Telephony Interoperable High-Quality Vocoder for VoIP: WIT)을 개발하였다. 개발된 VoIP 계층형 광대역 고품질 음성 코덱은 TTA 표준으로 상정하여 2004년 7월에 표준으로 제정되었다[4]. 표준으로 제정된 VoIP 계층형 광대역 고품질 음성 코덱은 3.4Hz에서 7kHz 대역까지 고대역부와 G.711, G.723.1, G.729의 저대역부(핵심계층 및 개선계층)로 구성되었다. 이러한 특성에 따라 협대역 음성 코덱에 비하여 음질이 향상되었으며, 기존의 협대역 코덱들과 호환이 가능한 장점을 갖는다.

유무선 통합 네트워크(BcN)는 다수의 네트워크가 존재하는 상황에서 다수의 코덱이 사용될 것이다. 다수의 코덱을 사용하는 환경에서는 호 설정 과

정에서 코덱과 협상을 수행하여야 한다. VoIP 계층형 광대역 고품질 음성 코덱은 G.711, G.723.1, G.729를 저대역부 코덱으로 사용하기 때문에 코덱 협상과정에 대한 표준 작업을 추진하여야 한다. 본 고에서는 IETF 코덱 협상하는 표준 규격을 바탕으로 TTA 표준 코덱인 VoIP 계층형 광대역 고품질 음성 코덱 협상 호 처리 방안을 제시하였다.

본 고에서는 네트워크에서 음성 코덱에 대한 현황 분석으로써 네크워크 음성코덱, 코덱 사용현황, 표준화 동향, 네트워크와 코덱간의 관계를 제 II장에서 설명하였다. 제 III장에서는 IETF VoIP 코덱 호 처리 표준 분석으로 SIP/SDP 코덱 협상, RTP 포맷, 페이로드 타입 및 포맷 표준, 코덱과 호처리 관계를 설명하였다. 제 IV장에서는 VoIP 계층형 광대역 고품질 음성 코덱 호 처리 절차 특성과 호환성에 따른 코덱 협상 방안을 설명하였다.

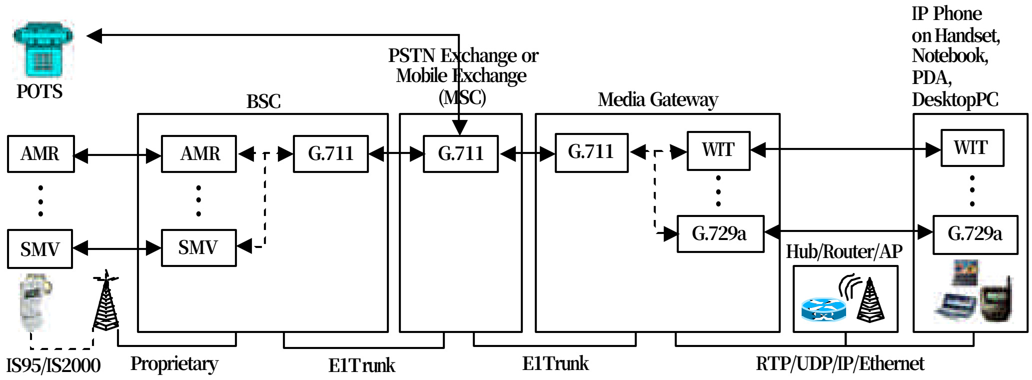
II. 네트워크 음성 코덱 현황

인터넷 음성 전화에서 사용하는 코덱은 기존의 PSTN 및 이동통신 네트워크 사용자들과 통화하여야 하기 때문에 코덱이 상호 호환되어야 한다. 본 장에서는 네트워크 음성 코덱, VoIP 코덱, 코덱의 사용 현황, 코덱 표준화 동향을 설명한다.

1. 네트워크별 사용 음성 코덱

PSTN 및 이동통신은 네트워크에서 제공하는 코덱 중심으로 사용한 반면, 인터넷은 단말에서 코덱을 선택할 수 있기 때문에 사용자 중심으로 사용한다. 즉, 인터넷에서는 단일 코덱 사용을 강요할 수 없다. PSTN 음성 코덱 G.711 사용자이면서, 동시에 ADSL/VDSL 서비스 가입으로 인터넷 전화용 코덱 사용자일 수 있다. 이동통신 네트워크에서도 이동통신 코덱 및 교환기 코덱 기반에서 VoIP 기술에 의한 인터넷 전화용 코덱을 사용하게 될 것이다. 데이터 서비스를 중심으로 하는 휴대인터넷에서도 음성 서비스를 제공하기 위하여 인터넷 전화용 코덱을 사용하게 될 것이다. 어느 한순간에 하나의 통일된 코덱을 사용하는 것은 불가능하다. 기존 PSTN 코덱, 이동통신 코덱, 이동통신 교환기 코덱 등이 존재하는 상황에서 기존 VoIP 코덱과 새롭게 표준화되어 사용하게 될 인터넷 전화용 코덱이 공존하게 될 것이다. 또한, 기존 PSTN 코덱, 이동통신 코덱, 이동통신 교환기 코덱 등을 경유하지 않고 인터넷 전화 코덱간에 통신하는 비율이 증가할 것이다.

유무선 통합 네트워크(BcN)는 PSTN, 이동통신 네트워크, VoIP 인터넷이 동시에 존재한다[5]-[7]. (그림 1)은 네트워크에서 사용하는 음성 코덱을 설명한다. PSTN과 이동통신 교환기(MSC)는 G.711 코덱을 사용한다. 인터넷에서는 코덱으로 G.723.1, G.729, WIT 등 다양하게 사용할 수 있다. PSTN의 POTS는 코덱이 없는 더미 단말이고, 이동통신 단말기는 제조사별로 상호 상이한 코덱을 채택하여 사용하고 있다. 다양한 코덱을 사용하더라도 상호 통화가 가능하도록 MSC에서는 G.711을 사용하였다. VoIP 인터넷 음성 전화용 코덱은 하드폰 또는 소프



(그림 1) 네트워크 음성 코덱

트폰의 제조사 별로 코덱을 달리하여 사용한다. 코덱 사용에 대한 제한이나 규약이 없으며, 타 제품과 연동하기 위하여는 호 처리 과정에서 코덱 정보를 상호 이해할 수 있도록 코덱 및 코덱 페이로드 포맷을 IETF에 표준으로 등록하여야 한다.

IP 폰은 IP 소프트폰과 IP 하드폰으로 구성할 수 있으며, IP 소프트폰은 노트북, 데스크톱 PC, PDA 등에 설치하여 사용할 수 있다. IP 폰과 미디어 게이트 간에는 IP(RFC791)/UDP(RFC768)/RTP(RFC 1889/RFC3550)에 음성 미디어 데이터를 전달한다.

본 고의 주된 관심사는 WIT 코덱을 활용할 수 있도록 네트워크에서 호 처리과정에서의 코덱 협상하는 방법에 있다.

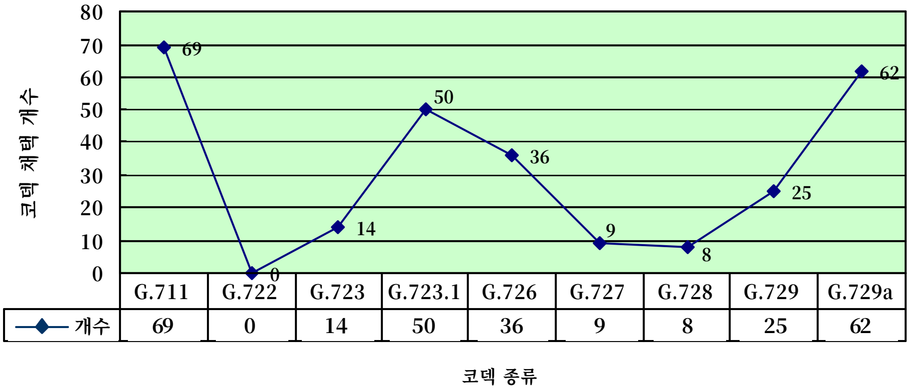
2. VoIP 코덱 사용 현황

PSTN에서는 코덱의 종류가 ITU-T G.711 하나였으나, VoIP와 이동통신의 출현으로 코덱의 종류가 증가하였다. VoIP 게이트웨이에서의 각 코덱에 대한 구현 개수는 (그림 2)와 같다[8]. 모든 게이트웨이에서 G.711을 선택하여 개발하였다. VoIP의 상징 코덱인 723.1은 50개, G.729a는 62개를 나타내고 있다. G.722를 제공하는 게이트웨이는 전혀 없는 0으로 나타나고 있다.

(그림 2)는 VoIP에서는 G.711, G.723.1, G.729가 일반적으로 사용되는 것을 나타내고 있다. G.729a와 G.729는 상호 호환된다. TTA에서 2004년 7월에 제정된 VoIP 계층형 광대역 고품질 음성 코덱은 사용 빈도가 높은 G.711, G.723.1, G.729와 상호 호환이 가능하다. ITU-T, 3GPP, 3GPP2에서의 표준화된 VoIP 계층형 광대역 고품질 음성 코덱은 자기 자신과 동일한 코덱하고만 통화가 가능한 반면, TTA VoIP 계층형 광대역 고품질 음성 코덱은 자기 자신뿐만 아니라, G.711, G.723.1, G.729 코덱들과 상호 호환 가능한 것이 큰 특징이다. TTA 광대역 코덱을 새롭게 채택하여 시스템 또는 단말에 적용하더라도, G.711, G.723.1, G.729 코덱을 이미 사용하고 있는 PSTN, 이동통신 네트워크, 인터넷 전화 등의 상대 시스템 및 상대 단말에게 TTA 광대역 코덱을 사용하라고 강요할 필요성이 없는 장점을 갖는다. 인터넷 전화는 전세계 어느 누구나와도 통화가 가능하여야 하기 때문에 모든 상대편에게 자기 자신이 사용하고 있는 코덱 사용을 강요할 수 없다. TTA 광대역 코덱은 호환성이 있기 때문에 상대 시스템이나 제 3의 시스템에게 트랜스 코덱 기능을 별도로 요구하지 않고 TTA VoIP 계층형 광대역 고품질 음성 코덱 자신이 상대 시스템의 능력에 맞추어 코덱을 제공할 수 있다. 그러나, ITU-T, 3GPP, 3GPP2에서의 표준화된 광대역 음성 코덱은 G.711, G.723.1, G.729 코덱들과 호환성이 제공되지 않기 때문에 상대 시스템 또는 제 3의 시스템에서 트랜스 코덱 기능을 제공하여야 한다.

3. 코덱 표준화 동향

VoIP에서는 G.729, G.723.1 등을 사용할 수 있



(그림 2) 코덱별 구현 개수

<표 1> 유선 및 무선의 코덱 표준화

구분	네트워크	코덱	표준화 단체	특성
유선	PSTN	G.711	ITU-T	64kbps
	인터넷	G.729a G.723.1 WIT	ITU-T ITU-T TTA	8kbps 5.3/6.3kbps G.723.1/G, 729a 호환
무선	GSM 비동기 3GPP	(GSM-AMR-NR) AMR-WB/G.722.2	ITU-T/3GPP/ETSI	4.75~12.2kbps/8mode 6.6~23.85kbps/9mode 50Hz~7kHz
	동기 3GPP/2	(QCELP/EVRC) SMV	TIA	9.6~14.4kbps/4mode

WIT: Wideband codec for Internet Telephony　EVRC: Enhanced Variable Rate Codec
ACELP: Algebraic Code Excited Linear Prediction　SMV: Selectable Mode Vocoder

다. PSTN은 ITU-T G.711 코덱을 사용하여 전송하여야 할 데이터가 64kbps인데 반하여 VoIP 코덱에서는 8kbps이면 가능하기 때문에, 전송 자원을 8배나 더 많이 효율적으로 사용할 수 있다. 이와 같이 기존 회선 방식의 ITU-T G.711 코덱 고정 방식보다는 코덱 선택에 대한 확장성을 제공하기 때문에 보다 효율적인 코덱을 개발할 동기가 생기고, 선택할 수 있는 기회가 발생하게 되었다.

유선 및 무선, 인터넷에서 사용하는 대표적인 코덱은 <표 1>과 같다. 유선은 PSTN과 인터넷으로 구분하였으며, 무선은 3GPP와 3GPP2로 구분하였다. 유선에 대한 코덱은 모두 ITU에서 표준화를 제정하였으며, 무선에 대한 코덱은 북미와 유럽의 지역적인 구분으로 인하여 표준 단체가 구분되었다.

유선에서 사용했던 G.711, G.729, G.723.1 등과 같은 코덱은 300~3400Hz 범위에서 음성신호를 코딩하였다. 유선 및 무선에서 음성 통화 품질을 향상시키기 위하여 50~7000Hz 범위의 음성 신호를 코딩하는 것을 표준으로 채택하는 추세이다. 비동기식 3GPP에서는 광대역 코딩 표준을 AMR-WB로 2001년에 채택하였고, ITU-T는 G.722.2를 2002년 초에 표준으로 채택하였다. 이 두 코덱은 동일한 것으로서 상호 다른 표준 기관에서 동일하게 표준으로 채택하였다. WIT는 TTA 표준의 근간이 되는 코덱으로 2004년 7월에 TTA 표준으로 제정되었다[9],[10].

4. 인터넷 음성 코덱 전달 요구사항

회선 교환 방식 특성과는 달리 인터넷 패킷의 특성은 분실, 중복, 순서 바뀜, 분할 전송, 다중 전송이 가능하다. 분할 전송은 하나의 패킷 손실로 인한 연속된 패킷 손실을 초래하므로 가능한한 하지 않는 것이 좋다. 코덱과 인터넷 패킷간의 관계에 대한 요구사항은 다음과 같다.

- 패킷의 경계는 코덱 프레임 경계와 동시에 존재하여야 한다. 그러므로, 하나의 패킷에는 하나 이상의 완전한 코덱 프레임들로 구성하여야 한다.
- 하나의 코덱 데이터의 최소 단위는 MTU(Message Transfer Unit) 이상의 크기가 되지 않는 한 하나의 패킷 내에 코덱 데이터들을 엇갈리게 패킷화해서는 안된다.
- 만약 코덱 프레임 크기가 MTU보다 크다면, 페이로드 포맷은 IP fragmentation에 의존하지 말아야 하고 코덱 자신의 메커니즘을 사용하여야 한다. 이 메커니즘은 fragment를 디코딩하는 코덱 특정 정보를 갖고 있어야 한다[11].
- 코덱은 정해진 표준화 RTP 포맷에 따라 고정된 길이로 전달된다. 전달된 RTP 패킷 길이로 코덱을 유추할 수 있으나, 동일한 패킷 길이를 갖는 코덱이 많은 경우에 하나의 코덱을 유추할 수 없다.
- 코덱 데이터의 전송도중 분실에 대하여 복구하기 위해 재전송하지 않는다.
- 코덱 데이터의 전송 도중 순서 바뀜은 RTP (Real-time Protocol) 데이터를 보고 탐지할 수 있고, 수신측에서 순서를 맞추어 디코딩하여야 한다.

Ⅲ. VoIP 코덱 협상 표준 분석

VoIP 코덱을 사용하기 위하여는 발신과 착신간의 코덱에 대한 협상 절차가 선행되어야 한다. 본 장에서는 SIP/SDP에 의한 호 처리 코덱 협상 절차, RTP, 코덱 타입 표준, 코덱 페이로드 포맷 표준, 코덱과 호 처리 절차 관계 등을 설명한다.

1. SIP/SDP 코덱 협상

IP 영역에서의 호 처리 절차는 SIP/SDP에 의해 이루어진다. SIP/SDP에 의한 호 설정과정에서 상호 통신가능한 코덱 정보를 주고 받으면서 코덱이 선정된다. M=audio 49230 RTP/AVP 0이라면, 음성 코덱 중에 0번을 사용한다는 것을 의미한다. 음성 코덱 0번은 RFC3551에 의거 PCMu 코덱을 의미한다(<표 2> 참조)[12]-[14]. PCMu 코덱은 RTP 코덱(codecs)에 실리게 되고, 코덱 0번은 PT(Payload Type)에 실려 착신측에 전달된다. PT에 따라 정해진 코덱 페이로드 포맷에 따라 착신측에 전달된다. 코덱별 페이로드 포맷은 IETF AVT 워킹 그룹에서 표준화하고 있다[15]-[20].

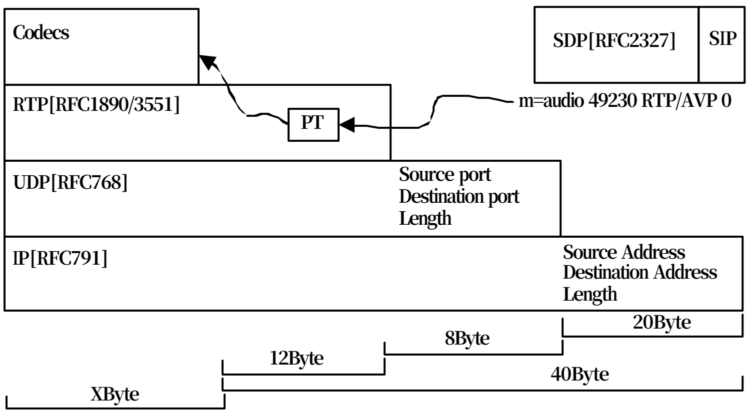
<표 2> IETF Payload Type 번호

PT No	PT Name	PT No	PT Name
0	PCMU	20	unassigned
1	reserved	21	unassigned
2	reserved	22	unassigned
3	GSM	23	unassigned
4	G723	24	unassigned
5	DVI4	25	CelB
6	DVI4	26	JPEG
7	LPC	27	unassigned
8	PCMA	28	nv
9	G722	29	unassigned
10	L16	30	unassigned
11	L16	31	H261
12	QCELP	32	MPV
13	CN	33	MP2T
14	MPA	34	H263
15	G728	35~71	unassigned
16	DVI4	72~76	reserved
17	DVI4	77~95	unassigned
18	G729	96~127	dynamic
19	reserved		

발신측에서 착신측으로 전달하는 인터넷 음성 패킷은 IP(Internet Protocol), UDP(User Datagram Protocol), RTP(A Transport Protocol for Real-Time Applications), codecs 내용으로 (그림 3)과 같이 패킷화하여 전달된다. UDP, RTP의 헤더는 20, 8, 12바이트로 고정되어 전달될 수 있으나, 코덱의 특성에 따라 RTP 페이로드가 상이하므로 전달되는 패킷의 전체 길이는 다양하게 나타날 수 있다.

2. RTP 포맷

RTP는 RFC1889/RFC3550에 의거 V(version:2),

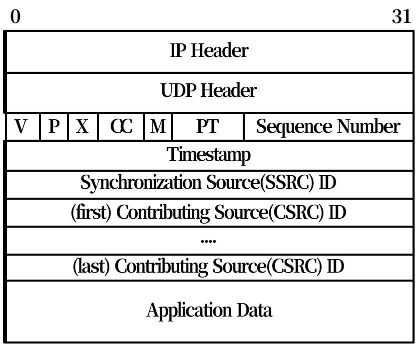


(그림 3) 코덱 결정 정보 관계

P(padding:1), X(extension:1), CC(CSRC count:4), M(marker:1), PT(payload type:7), SN(sequence number:16), Timestamp(32), SSRC(synchronization source:32) ID, CSRC(contributing sources) ID로 (그림 4)와 같이 구성된다. 코덱 정보는 PT로 RTP 페이로드에 실리는 코덱 번호를 나타낸다. SIP/SDP의 호 설정 과정에서 코덱을 결정하기 때문에 시스템에서 어떤 코덱을 사용할 것인지 알 수 있지만, RTP PT를 보고서도 어떤 코덱인지 알 수 있다.

PT는 7비트이기 때문에 128개로 한정하여 표시할 수 있다. 코덱에 대한 PT를 IETF에서 표준으로 제정하고 있다. 수신측은 이해할 수 없는 PT가 오면 무시한다. PT는 하나의 코덱만 표시할 수 있기 때문에 하나 이상의 코덱 데이터는 하나의 패킷으로 전달할 수 없다. 하나 이상의 코덱 데이터를 상대에게 전달하고자 할 경우에는 상호 다른 IP 주소와 포트를 사용하여 전달하여야 한다.

RTP sequence number와 timestamp에 의해 UDP/IP의 특성에 의해 패킷 분실(packet loss) 또는 패킷 순서(misorder)가 바뀌는 것을 탐지할 수 있다. RTP를 갖는 ADU(Application Data Units)의 특징은 수신측에서 해석할 수 있는 충분한 정보를 갖고 있다. 예를 들면, 이전 음성 코덱 데이터가 있는 ADU를 손실하였다손 치더라도, 현재 받은 ADU로 해석할 수 있다. 이러한 특성은 인터넷 특성에 따른 것이다. RTP 정보에 의해 패킷이 분실된 것을 발견하면, 코덱에 정보를 제공하여 코덱 내의 분실된 패킷을 복구하는 기능을 트리거하도록 한다[21].



(그림 4) RTP 포맷

3. 코덱 페이로드 타입

127개의 정해진 코덱 정의 공간을 IETF에서 표준화하고 있다. RTP(RFC1890/RFC3551)에 정의되지 않은 코덱은 수신측에서 무시하여야 한다. MIME(RFC2045/2048/3555)에 따라 인터넷에 미디어 데이터를 전달할 수 있다.

PT 번호 95까지는 고정된 번호이고, 96부터 127까지의 32개의 dynamic payload type으로 사용할 수 있다. 새로운 응용이 발생하면, 96에서 127 범위 내에서 사용하여야 한다.

4. 코덱 페이로드 포맷 표준

VoIP에서 음성 코덱 데이터를 RTP에 포함하여 전달하는 포맷(RTP 페이로드)을 IETF Transport Area Audio/Video Transport(avt) 워킹 그룹에서 정의한다. VoIP 음성 코덱별로 정해진 표준 포맷으로 RTP에 실려 전달되어야 수신측에서 전달받은 음성 코덱 데이터를 정해진 표준으로 읽어들여 자연 음성으로 처리한다.

AVT(Audio/Video Transport) 워킹그룹은 UDP/IP 상위에 오디오와 비디오를 실시간으로 전달하는 프로토콜을 정의하는 것을 목표로 한다. 이 프로토콜을 RTP(Real-time Transport Protocol: RFC1889/3550)이라 하고, RTP에 대한 프로파일(RFC1890/3551)과 페이로드 포맷을 정의한다. 새로운 코덱 자체를 검토 대상으로 하지 않고, 새로운 코덱에 대한 포맷을 수용하는 방향으로 추진하고 있다. 인터넷 드래프트의 페이로드 포맷은 MP3, H.264, iLBC, MIDI, VMR-WB(Variable-Rate Multimode Wideband) Audio Codec 등이 있다. RFC 페이로드 코덱은 JPEG(RFC2035/2435), H.261(RFC2032), MPEG1/MPEG2(RFC2038/2250), H.263(RFC2190), H.263(RFC2429), DTMF(RFC 2833), MPEG4(RFC3016), G722.1(RFC3119), AMR/AMR-WB(RFC3267), Comport Noise(RFC3389), EVRC(Enhanced Variable Rate Codecs)/SMV(Selectable Mode Vocoders: RFC 3558)로 제정되어 있다.

5. 코덱과 호 처리 절차 관계

호 설정 과정중에 발신측의 코덱과 착신측의 코덱이 협상하는 과정에서 결정된다. 결정된 코덱은 송신측과 수신측에서 (그림 5)에서와 같이 selector에 의해 encoder와 decoder가 선택된다. Encoder에 의해 산출된 미디어 데이터는 코덱(payload type format)에 따라 packetizing된다. Packetizing 결과는 RTP/UDP/IP 패킷으로 착신측에 전달된다.

코덱 데이터 수신측에서는 인터넷을 통하여 RTP/UDP/IP에 의해 수신되면, 이들 데이터를 depacketizer에 의해 payload type(코덱)을 알 수 있고, 이에 해당되는 decoder에 의해 음성으로 변환된다.

코덱 처리 관련 호 처리에서의 역할과 시스템 자체에서의 역할을 정립하여야 한다. 코덱 처리를 위한 호 처리 역할은 송신측 코덱 능력과 수신측 코덱 능력을 문의, 일치, 확인하는 것이며, 코덱 처리 관련 시스템 자체에서의 역할은 코덱 데이터의 패킷화, 송수신, 패킷 순서 정렬 후 decoder에 전달, encoding/decoding 등이다. 모든 단말이 동일한 코덱을 갖거나, 수신측에서 수신된 모든 코덱을 수신할 수 있는 능력이 보장된다면, RTP PT를 이용하여 코덱을 식별할 수 있기 때문에 호 처리 과정에서 코덱에 대한 협상 과정은 필요 없다. 그러나, 인터넷은 코덱에 대한 제한이 없기 때문에 다양한 코덱 데이터가 송수신될 수 있고, 수신된 모든 코덱을 갖는 것은 불가능하다. 호 설정 과정에서 코덱 협상 절차가 없다면, 수신측에서 송신측의 코덱이 없는 경우에도 송신측과 수신측이 상호 연결이 될 수 있다. 수신측에서 RTP PT 정보를 이용하여 코덱을 인지할 수는 있지만, selector가 코덱을 선정할 수 없기 때문에 음성 호 통화는 불가능하게 된다. 이러한 문제점을 해소하기 위하여 호 설정과정에서 반드시 코덱 능력 협상이 이루어져야 한다.

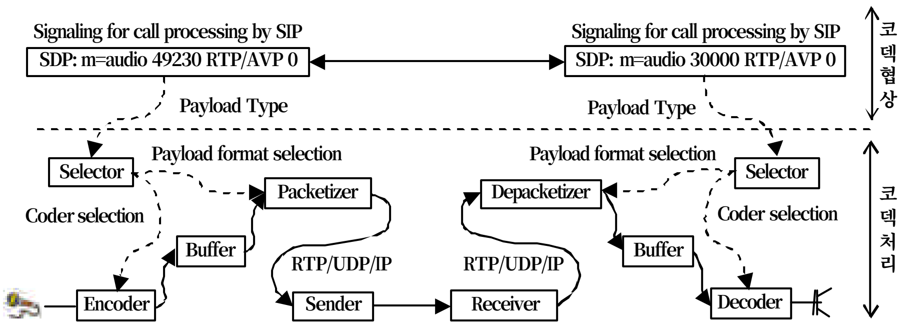
호 통화중에 코덱을 변경하는 경우에는 회의 통화 시에 새로운 화자가 추가될 시에 발생하며, 이때 코덱 처리 능력에 대한 협상 과정을 진행하여야 한다. 코덱의 특성상 여러 모드로 구성되어 표준화된 경우에는 수신측에서 해당 코드의 모든 모드를 갖고 있다는 가정 아래 호 통화 도중에 코덱 능력 협상 과정없이 모드 변경은 가능하다. 모드에 대한 정보는 RTP 페이로드 내에 정의된 정보를 이용한다.

Ⅳ. 광대역 고품질 음성코덱 협상

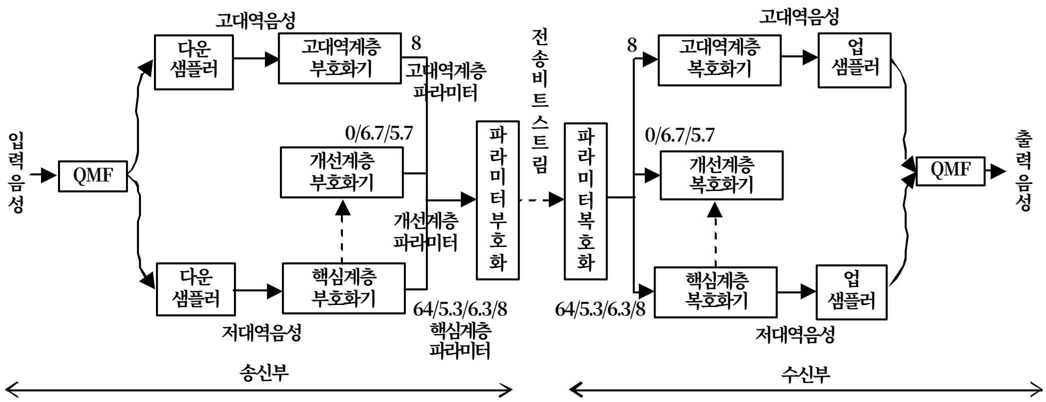
통화 서비스를 위하여는 호 설정이 이루어지고 난 후에 미디어 데이터 송수신에 의한 통화가 이루어진다. 코덱 변환 기능을 위하여는 호 처리 기술이 선행되어야 한다. 코덱 변환 호 처리 기술이 없으면, 코덱 협상하는 과정에서 통화가능한 코덱이 없다는 오류 메시지와 함께 호 설정이 안되기 때문이다.

1. VoIP 계층형 광대역 고품질 음성 코덱의 특성

VoIP 계층형 광대역 고품질 음성 코덱(WIT)은 협대역 및 광대역 음성 코덱을 이용하여 음성을 전달하는 패킷 전화에 사용할 수 있고, (그림 6)과 같이 고대역계층 부호화기, 개선계층 부호화기, 핵심계층 부호화기로 구성한다.



(그림 5) 코덱과 호 처리 절차 관계



(그림 6) WIT 부호화기

<표 3> WIT 호환 코덱

	A node	B node	전송률(kbps)	ms/frame octet	default ms/frame	응용분야
1	WIT	WIT	64/6.3/5.3/8+ 0/6.7/5.7/4+8			지사간 VoIP, VPN, elearning
2	WIT	G.711A	64		20	PSTN
3	WIT	G.711u	64		20	PSTN
4	WIT	G.723.1 5.3	5.3	20	30	VoIP 호환
5	WIT	G.723.1 6.3	6.3	24	30	VoIP 호환
6	WIT	G.729	8	10	20	VoIP 호환

저대역 음성은 G.711, G.723.1 및 G.729 중 하나의 핵심계층 부호화기를 통해 핵심계층 파라미터로 부호화된다. 핵심계층 부호화기로 G.723.1과 G.729를 사용하는 경우에 개선계층 부호화기를 통해 개선계층 파라미터를 구한다. 고대역 음성은 고대역계층 부호화기를 통해 고대역계층 파라미터로 변환된다. 각 계층 부호화기에서 부호화된 각 파라미터들은 파라미터 부호화 블록을 통해 비트스트림으로 변환되어 전송된다.

먼저 수신된 비트스트림으로부터 각 계층에 해당하는 파라미터의 인덱스를 추출한다. 추출된 파라미터 인덱스는 30ms의 음성 프레임에 해당하는 파라미터를 얻기 위해 복호화된다. 이 파라미터들은 핵심계층 파라미터, 개선계층 파라미터와 고대역계층 파라미터 등이다. 핵심계층 파라미터와 개선계층 파라미터는 핵심계층 복호화기와 개선계층 복호화기를 통해 저대역 음성으로 복호화되고, 고대역계층 파라미터는 고대역계층 복호화기를 통해 고대역에 해당하는 음성으로 복호화된다. 복호화된 저대역 음성과 고대역 음성은 업-샘플러와 QMF(Quadrature Mirror Filter)를 통해서 16kHz의 샘플링 주파수를 갖는 광대역 음성으로 복원된다.

2. VoIP 계층형 광대역 고품질 음성 코덱의 호환성

VoIP 계층형 광대역 고품질 음성 코덱 WIT는 <표 3>과 같이 6종류의 코덱과 상호 호환성을 갖고 음성 통화할 수 있는 장점을 갖는다.

WIT에서 호환가능한 코덱 WIT, G.711A, G.711u, G.723.1 5.3, G.723.1 6.3, G.729는 각각 별개의 코덱으로 식별 가능하므로, 호 설정 코덱 협상과정에서 개별적 코덱으로 협상하여야 한다. 저대역부 개선계층에 대한 협상방법은 추후 연구대상으로 한다. WIT, G.711A, G.711u, G.723.1 5.3, G.723.1 6.3, G.729로 offer한 경우에 answer 측에서 WIT가 아닌 G.711A, G.711u, G.723.1 5.3, G.723.1 6.3, G.729로 answer 하는 경우에는 answer 순서의 의미에서 협상 완료하여 미디어 통신한다. WIT-WIT로 협상하는 경우에는 WIT 저대역 핵심계층 G.711A, G.711u, G.723.1 5.3, G.723.1 6.3, G.729 중 어느 것을 선택하여 협상하여야 한다.

WIT는 WIT, G.711A, G.711u, G.723.1 5.3, G.723.1 6.3, G.729 코덱과 개별적으로 협상하는 반면, VBR-WB 코덱인 경우에는 mode-set=0,1,2로 호 설정과정에서 협상된 후에 호 처리 도중에 RTP 내에 CMR(Codec Mode Request)에 의해 모드 변환을 수행한다[21].

3. PSTN/MSC 연동 코덱 협상

WIT 코덱을 갖는 VoIP는 기존 PSTN 교환기와 연동하거나 이동통신 교환기 MSC와 연동할 수 있다. PSTN 교환기와 이동통신 교환기 MSC는 G.711 코덱을 사용하므로 WIT 코덱을 사용하는 게이트웨이 입장에서는 동일한 구성노드로 인식한다. PSTN/MSC 연동 호 처리 시나리오는 다음과 같이 구분할 수 있다.

- IP(WIT)-(G.711)PSTN/MSC
- IP(WIT)-(G.711)PSTN/MSC(G.711)-(WIT) IP
- PSTN/MSC(G.711)-(WIT) IP(WIT)-(G.711) PSTN/MSC

PSTN/MSC 연동 호 처리 시나리오에서는 offer/answer 코덱 협상 호 처리방식과 TFO(Transcoding Free Operation) 방식을 이용할 수 있다. IP(WIT)(G.711)PSTN/MSC 또는 PSTN/MSC(G.711)-(WIT) IP(WIT)-(G.711) PSTN/MSC는 offer/answer 코덱 협상 호 처리 방식을 사용하여야 한다. 기존 망에서 G.711 코덱으로 협상하므로 기존 네트워크 코덱으로 통신하기 위하여 G.711 코덱으로 부호화한다. IP(WIT)-(G.711)PSTN/MSC는 웹투폰 또는 폰투웹의 형태로 VoIP 서비스에서 가장 일반적인 시나리오이다. PSTN/MSC(G.711)-(WIT) IP(WIT)-(G.711) PSTN/MSC는 국제구간 및 시외구간의 전화 비용을 절감하기 위하여 통신사업자 네트워크 형태와 기업체의 지사간 구성된 네트워크 형태이다.

IP(WIT)-(G.711)PSTN/MSC(G.711)-(WIT) IP에서는 3GPP(www.3gpp.org) TFO 방식[22]과 같이 PSTN 또는 이동통신(mobile) 구간 내의 통화에서 탄뎀 통화를 회피하는 기능이다. TFO는 G.711 변환 전달하는 과정에서 발생하는 음질 저하 부분을 제거하고 WIT의 음질을 그대로 전달할 수 있는 장점이 있으나, 기존 네트워크 표준을 수정하여야 하기 때문에 실현하기 위하여는 추후 연구가 필요하다.

4. All IP 코덱 협상

All IP 호 처리 절차는 PSTN이나 이동통신 교환기 MSC 처럼 G.711 하나의 코덱을 사용하는 것이 아니라 발신 및 착신의 단말이 상호 상이한 코덱을 사용하기 때문에 코덱에 대한 호 처리 절차로 IETF의 offer/answer 방식을 사용한다. Offer/answer 코덱 협상 호 처리 기능은 발신측에서 수신 받기를 원하는 코덱 정보를 착신측에게 전송하여 제시(offer)하면, 착신측은 발신측에서 제시한 코덱 중에서 수신받고자 하는 코덱을 SDP를 이용하여 응답(answer)한다. Offer/answer 방식을 이용하여 발신과 착신간의 코덱을 협상한 후에 선택된 코덱으로 통화하도록 한다. Offer/answer 코덱 협상 호 처리에 대한 표준은 IETF(www.ietf.org.) MMUSIC 워킹 그룹에서 2002년 6월에 "An Offer/Answer Model with the SDP" RFC 3264로 제정되었다.

WIT는 WIT, G.711A, G.711u, G.723.1 5.3, G.723.1 6.3, G.729 상호 호환이 가능하므로, WIT 코덱을 갖고 있는 위치에 따라 다음과 같이 호 처리 기능을 구분할 수 있다.

- 발신측 WIT 코덱 호 처리 기능
- 착신측 WIT 코덱 호 처리 기능
- MGC WIT 코덱 호 처리 기능
- 착신측 WIT 호출 코덱 서버 기능

발신측 WIT 코덱 호 처리 기능은 발신측에서 offer SDP로 전송 시에 WIT, G.711A, G.711u, G.723.1 5.3, G.723.1 6.3, G.729 코덱 정보를 착

신측에게 제시함으로써 착신측에서 제공 가능한 코덱을 선택하도록 한다. 착신측 WIT 코덱 호 처리 기능은, 발신측에서 보내온 offer에 의한 답으로 착신측에서 WIT, G.711A, G.711u, G.723.1 5.3, G.723.1 6.3, G.729 코덱 정보를 발신측에서 제시한 코덱에 응답하여 통화할 수 있도록 한다. MGC WIT 코덱 호 처리 기능은 발신측과 착신측 중간에서 WIT 코덱이 탑재한 경우에 발신측 코덱과 착신측에게 WIT, G.711A, G.711u, G.723.1 5.3, G.723.1 6.3, G.729 offer/answer SDP를 전송하여 코덱을 선택할 수 있도록 한다. 착신측 호출 WIT 코덱 서버 기능은 코덱 변환 엔진을 갖는 전용 서버를 구축하고, 착신측이 발신측에게 제시한 WIT 코덱을 수신할 수 있는 코덱이 없는 경우에 착신측에서 WIT 코덱 서버에 WIT 코덱 변환 기능을 호출한다.

V. 결 론

인터넷에서 기존에 제공하던 PSTN과 이동통신의 음성 전화가 가능해짐에 따라 인터넷 전화에 대한 인터넷 음성 코덱 기술이 중요한 핵심 기술이 되었다. 본 고에서는 TTA에서 VoIP를 위한 광대역 고품질 음성 코덱이 2004년 7월에 제정됨에 따라 제정된 코덱 특성, 네트워크와 코덱간의 관계, 호 설정 과정에서의 코덱 협상에 대한 표준화 동향을 설명하였다.

TTA VoIP 계층형 광대역 고품질 음성 코덱은 G.711, G.723.1, G.729 코덱과도 상호 호환성을 갖는다. G.711, G.723.1, G.729 코덱과의 상호 호환성에 따라 통신하고자 하는 상대 시스템이 VoIP 계층형 광대역 고품질 음성 코덱을 갖고 있지 않는 경우에는 G.711, G.723.1, G.729 중의 하나를 선택하여 통화할 수 있다. 호환성을 제공하지 않는 경우에는 상대 시스템 또는 제3의 시스템에 트랜스코덱 기능을 사용하여야 한다.

VoIP 계층형 광대역 고품질 음성 코덱을 호 설정 과정에서의 PSTN/MSC 연동 호 처리와 ALL IP 호 처리로 구분하여 코덱 협상 처리 방안을 제시하였다. PSTN/MSC 연동 호 처리는 IP(WIT)-(G.711) PSTN/MSC, IP(WIT)-(G.711) PSTN/MSC (G.711)-(WIT) IP, PSTN/MSC(G.711)-(WIT) IP(WIT)-(G.711) PSTN/MSC로 구분한다. Offer/answer 방식은 PT를 정의하여 사용하여야 한다. ALL IP인 경우에는 모든 노드에 광대역 음성 코덱이 있을 수도 있고, 발신측, 착신측, MGC, 착신측 WIT 호출 코덱 서버에만 있을 수도 있다.

VoIP 계층형 광대역 고품질 음성 코덱의 저대역 핵심계층을 고려한 코덱 협상을 수행한다. WIT와 G.711A, G.711u, G.723.1 5.3, G.723.1 6.3, G.729 저대역 개선 계층을 갖고 있는 상대인 경우에 대한 호 협상도 해야 한다. WIT는 고대역부는 동일하나, 저대역부 G.711A, G.711u, G.723.1 5.3, G.723.1 6.3, G.729를 선택적으로 사용할 수 있다. WIT의 저대역부 G.711A, G.711u, G.723.1 5.3, G.723.1 6.3, G.729에 대한 협상도 해야 한다. 협상하고자 하는 코덱에 대한 정보는 SDP로 전달하고 RTP PT에 정의하기 위하여는 IETF PT 번호를 부여받거나, 동적으로 정의하여 사용하는 97부터 127 중에 선택하여 정의한 후에 사용하여야 한다. PT는 4비트로 정의되어 있어서 총 127개만 사용 가능하고, 동적으로 사용 가능한 개수는 30개이다. WIT-저대역 핵심계층 5개, 저대역 개선계층 5개가 PT로 지정되어 사용되어야 한다.

VoIP 계층형 광대역 고품질 음성 코덱은 저대역 핵심계층에 G.711, G.723.1, G.729를 사용하므로 10종의 PT를 설정하여 코덱 협상함으로써 광대역 고품질 음성 코덱, G.711, G.723.1, G.729 등과 상호 호환이 되는 장점을 갖는다. 코덱 협상 처리 방법이 정착이 되면, VoIP 계층형 광대역 고품질 음성 코덱은 기존 네트워크 장치 기능을 수정하지 않고 사용할 수 있다. VoIP 계층형 광대역 고품질 음성 코덱을 네트워크에서 원활하게 사용하기 위하여 PT 표준 정의 및 코덱 협상 과정을 추진하여야 한다.

참 고 문 헌

[1] G. Camarillo, E. Burger, H. Schulzrinne, and A. van Wijk, "Transcoding Services Invocation in the Ses-

sion Initiation Protocol," draft-camarillo-sip-deaf-02.ps, Feb. 17, 2003.
[2] J. Rosenberg, H. Schulzrinne, G. Camarillo, A.R. Johnston, J. Peterson, R. Sparks, M. Hand-ley, and E. Schooler, "SIP: Session Initiation Protocol," RFC 3261, Internet Engineering Task Force, June 2002.
[3] M. Handley and V. Jacobson, "SDP: Session Description Protocol," RFC 2327, Internet Engineering Task Force, Apr. 1998.
[4] TTA, "계층형 다중 비트율 광대역 음성 코덱," TTAS.KO-01.0045, 2004.
[5] 강태규, "차세대 인터넷 서비스 고도화를 위한 IETF의 인터넷 지능망 표준화 동향 분석," 주간기술동향 통권 953호, 2000. 7. 4., pp.1-17.
[6] 강태규, "유무선 통합 네트워크에서의 VoIP를 위한 공통 논리 기능 구조 분석," 전자통신동향분석, 2002. 10., pp.47-54.
[7] 강태규, "유무선 통합 VoIP 미디어게이트웨이 기술 동향," 주간기술동향, 통권 1101호, 2003. 6. 25., pp.1-11.
[8] DAI, "Voice Over Packet(IP and ATM)," The DAI Perspective a Global Telelcom Technology Series Edition 43, 2001. 10.
[9] 이미숙, "음성코덱의 특성과 협대역 음성 코덱 표준기술 분석," 주간기술동향 통권 1129, 2004. 1. 20., pp.1-14.
[10] 이미숙, "광대역 코덱의 기술 및 표준화 동향," *TTA Journal*, 2004. 3., pp.65-71.
[11] M. Handley and C Perkins, "Guidelines for Writers of RTP Payload Format Specifications," IETF RFC 2736, Dec. 1999.
[12] N. Charlton, M. Gasson, G. Gybels, M. Spanner, and A. van Wijk, "User Requirements for the Session Initiation Protocol(SIP) in Support of Deaf, Hard of Hearing and Speech-impaired Individuals," RFC 3351, Internet Engineering Task Force, Aug. 2002.
[13] S. Floyd and L. Daigle, "IAB Architectural and Policy Considerations for Open Pluggable Edge Services," RFC 3238, Internet Engineering Task Force, Jan. 2002.
[14] J. Rosenberg and H. Schulzrinne, "An Offer/Answer Model with Session Description Protocol(SDP)," RFC 3264, Internet Engineering Task Force, June 2002.
[15] J. Rosenberg, J. Peterson, H. Schulzrinne, and G. Camarillo, "Best Current Practices for Third Party Call Control in the Session Initiation Protocol," internet draft, Internet Engineering Task, Force, June 2002, Work in progress.
[16] J. Rosenberg, "A Framework for Conferencing with the Session Initiation Protocol," internet, draft, Internet Engineering Task Force, Nov. 2002, Work in progress.
[17] R. Sparks, "The SIP Refer Method," internet draft, Internet Engineering Task Force, Dec. 2002, Work in progress.
[18] B. Biggs, R. Dean, and R. Mahy, "The Session Initiation Protocol(SIP)," internet draft, Internet Engineering Task Force, May 2002, Work in progress.
[19] G. Camarillo, H. Schulzrinne, and E. Burger, "The Source and Sink Attributes for the Session, Description Protocol," internet draft, Internet Engineering Task Force, Sep. 2002, Work in progress.
[20] G. Camarillo, J. Holler, G. Eriksson, and H. Schulzrinne, "Grouping of m Lines in SDP," internet draft, Internet Engineering Task Force, Feb. 2002, Work in progress.
[21] Sassan Ahmadi, "Real-Time Transport Protocol (RTP) Payload and File Storage Format for the Variable-Rate Multimode Wideband(VMR-WB) Audio Codec, IETF draft-ietf-avt-rtp-vmr-wb-00.txt, May 28, 2004.
[22] 3GPP TS 22.053 V5.0.0(2002-06), "Tandem Free Operation(TFO); Service Description; Stage 1(Release 5)," 2002. 6.