

VoLTE를 위한 4G 음성압축기술 EVS

정경훈 3GPP TSG SA WG4 부의장
삼성전자 DMC 연구소 수석연구원
주기현 삼성전자 DMC 연구소 책임연구원
성호상 삼성전자 DMC 연구소 수석연구원
오은미 삼성전자 DMC 연구소 마스터

1. 머리말

2010년대에 이르러 4세대에 진입한 이동통신 시스템은 내부적으로는 2, 3, 4세대 기술이 혼재되어 구성되며 각각의 기술들이 다른 속도와 주기로 진화하고 있다. 패킷교환 기술이 전면적으로 사용된 4세대에 이르러 무선 액세스, 코어 네트워크가 동시에 교체되었고 IP Multimedia Subsystem(IMS)와 같은 신기술이 도입되었으나 음성압축기술은 3세대 초반에 개발된 기술이 그대로 사용되어 통화품질과 망 용량을 제약하게 되었다. TTA저널을 통해 4세대 음성압축기술로 LTE와 동시에 개발이 시작되어 2014년말 표준화 완료를 앞두고 있는 Enhanced Voice Service(EVS) 코덱의 개발배경, 설계목표, 그리고 목표에 맞추어 개발된 코덱의 구조와 성능평가 결과를 VoLTE 시스템과의 연동 측면에서 정리하고자 한다.

2. 세대별 음성통화 기술의 발전현황 및 EVS 개발의 기술적 배경

1990년대 초, 중반에 상용화된 GSM, IS-95의 경우처럼 이동통신 디지털화의 초창기에서는 무선 액세스, 코어 네트워크, 음성압축기술(코덱) 등 이동통신 시스템을 구성하는 주요 기술들이 동시에 개발되었다. 제품 교체주기가 짧은 단말기의 경우 신기술의 도입이 빠르게 이루어지는 반면 막대한 투자가 요구되는 유, 무선 네트워크 분야의 기술은 점진적인 개량이 이루어지는 경향이 있었는데 각각 다른 속도로 진화하기 시작한 2세대 이동통신 기술들이 2000년대 초반 W-CDMA, cdma2000 등 3세대로 접어들었을 때 무선 액세스 기술의 경우 현격한 변화(W-CDMA) 또는 점진적 개량(cdma2000)이 이루어졌으며 코어넷 기술의 경우, 양 시스템 모두 소폭의 개량이 이루어졌다. 한편 음성압축기술은 2세대 초반 제안되었던 Full Rate(FR), Qualcomm

Code-Excited Linear Prediction(QCELP) 기술이 통화품질과 네트워크 용량 확보에 약점을 드러내어 2세대 후반에 각각 Adaptive Multi-Rate(AMR), Enhanced Full Rate(EVRC) 기술로 교체된 후 계속 사용되었다.

3세대를 위한 음성압축기술로 AMR-Wideband(AMR-WB), EVRC-B 등이 개발되었으나 기존 기술을 대체하지 못하였고 Long Term Evolution(LTE), Evolved Packet Core(EPC) 등 기존 기술과 크게 다른 무선 액세스와 코어 네트워크가 출현한 4세대가 시작하였는데 그 동안 이동통신 기술의 양대 산맥으로 GSM, W-CDMA 기술과 10여 년간 경쟁해 온 CDMA 기반 3GPP2 기술들이 이 시점부터 시장에서 퇴출되기 시작하였으며 패킷 네트워크를 선점하기 위해 제안된 WiMAX 기술이 안착하지 못하여 4세대 이동통신 시장은 결국 LTE로 단일화되었다. 새로운 이동통신 기술의 상용화를 음성통화 개통으로 증명하던 2, 3세대와 달리 LTE는 데이터 통신에 제한되어 서비스가 시작된 후 상당한 시간이 경과되어 통화 기술로 Voice over LTE(VoLTE) 기술이 상용화되었는데 기존 회선교환 기술이나 인터넷 기반 VoIP 기술과 달리 유무선 네트워크의 자원관리에 IMS를 적용하고 기존 서킷 네트워크와의 핸드오버에 복잡한 절차를 사용하는 등 향상된 품질이 아니라 안정된 회선교환 기술의 통화품질 수준을 지원하는 것만으로도 기술적으로 벽찬 부분이 적지 않았기 때문이다. 2012년 한국에서 세계 최초로 상용화되어 사용범위가 전국에 도달한 VoLTE는 4세대 기술인 LTE, EPC, IMS에 3세대 초반 기술인 AMR-WB로 구성되어 네트워크의 수준이 높은 한국시장에서도 이미 Skype, Face Time 등 Over the Top(OTT) 서비스의 최신 음성압축기술을

체험한 고객들에게 품질로 크게 인정받지 못하였으며 패킷교환망의 전송환경으로 인해 23.85kbps라는 기존 대비 두 배 가까운 전송속도를 소모하여 망 용량관리 측면에서도 크게 인정받지 못하면서 LTE의 통화기술 지원 가능성 여부를 확인하는데 만족하여야 하였다.

2, 3세대 이동통신망의 경우 유선통화망과 접속하여 통화를 처리하므로 통화상대를 고려하지 않고 네트워크와 단말의 변경 혹은 교체가 가능하였다. 따라서 이동통신망을 개통하는 즉시 타 이동통신 사업자 혹은 해외 사업자 단말과의 통화가 지원되었으며 단말의 모델이 해당 주파수를 지원하는 한 Universal Integrated Circuit Card(UICC) 교체를 통해 통신사업자를 손쉽게 바꿀 수 있었다. 그러나 VoLTE의 경우 단말 간에 Session Initiation Protocol(SIP), Session Description Protocol(SDP)와 IMS를 통해 직접 메시지를 교환하는데 이들 인터넷 프로토콜 메시지의 형태가 기존 통신기술에서 사용되는 명확한 ASN.1 포맷과 달리 규격의 이해나 구현방법에 따라 해석의 차이가 발생할 수 있는 텍스트로 구성되어 있으므로 두 단말의 교섭절차가 유사한 논리에 기반하지 않을 경우 통화성립이 지체되거나 의도하지 않은 조건으로 통화가 교섭되는 문제가 있으며 이러한 제약으로 인해 현재 상용화된 VoLTE 서비스는 개별 사업자의 망 내에서만 가능하며 로밍도 가능하지 않은 실정인데 모든 통신사업자가 이미 전국망 수준의 VoLTE 서비스를 제공하고 있는 한국에서는 이러한 문제점을 해결하는 '사업자 간 UICC 이동성 제공을 위한 VoLTE 단말 규격'이 TTA에서 2년 간의 작업 끝에 제정되어 정부 고시로 7월 1일부터 발매되는 모든 단말에 적용될 예정이다.

<표 1> EVS 설계기준 (Design Constraints)

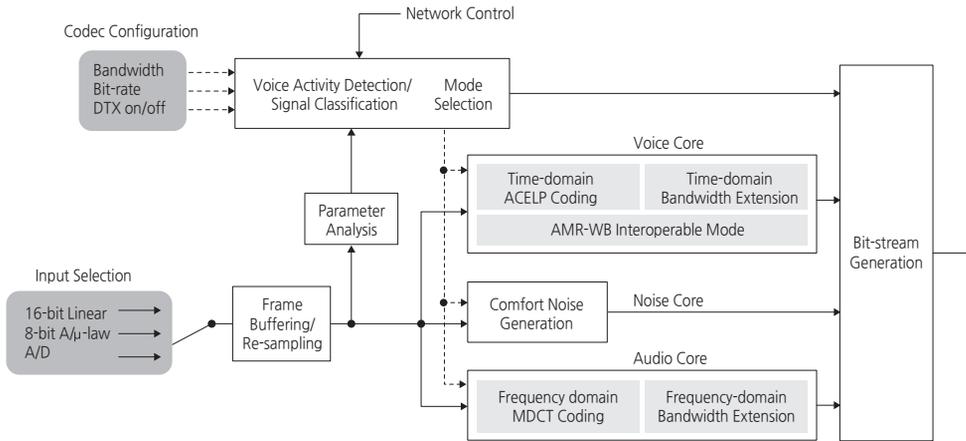
Parameter	Design constraints
Sampling frequency	8, 16, 32, and 48kHz
Audio bandwidth	NB, WB, SWB, and optionally FB
Audio channels	Mono and optionally stereo
Bit rates	5.9kbps VBR, 7.2, 8.0, 9.6, 13.2, 16.4, 24.4, 32, 48, 64, 96, 128kbps, and all AMR-WB modes
Algorithmic delay	Less than or equal to 32ms
Frame length	20ms
Complexity	Less than or equal to 88 wMOPS
Memory size	RAM≤100, ROM≤100 kWords
Supplementary features	Jitter buffer management, packet loss concealment, rate switching, DTX

3. EVS 개발목표

3GPP에서 2007년부터 필요성과 기술적 타당성 검토를 시작한 EVS는 <표 1>과 같은 설계목표를 정하고 2011년부터 본격적인 기술개발을 시작했다[1]. AMR, AMR-WB에서 사용되었던 Constant Bit-Rate(CBR) 압축기술에 기반하고 있으나 기존 기술들과 비교하여 가장 큰 차이점은 입력신호의 대역폭이 다양하며 통화 중에도 변경될 수 있다는 점이다. 특히 32kHz, 48kHz의 sampling rate는 음성 이외의 음악이나 배경잡음이 혼재될 경우 압축성능이 급격히 악화되었던 기존 코덱의 제한을 극복하기 위해 지원하기로 결정되었으며 Opus와 같이 인터넷의 무료 압축기술들이 이미 이러한 넓은 대역폭을 지원하는 점도 설계기준 설정에 영향을 끼쳤다. 또한, 7.2kbps에서 128kbps 까지의 넓은 bit-rate 범위를 지원하며 기존 VoLTE 단말들을 배려한 AMR-WB 호환모드도 포함되었다.

한편 EVRC 등의 CDMA 네트워크에 주로 사용되던 Variable Bit-Rate(VBR) 압축기술도 성능목표는 강제하지 않는 평균 5.9kbps 모드로 포함되어 있다. LTE와 같은 패킷 교환망의 전송

조건은 기존 회선 교환망과 달리 전송시간이 일정하지 않을 수 있고(지터) 패킷들이 전송 도중 대량으로 손실될 수 있는데 이는 기본적으로 망 운용이 품질관리 보다 전송효율을 극대화하기 위한 스케줄링에 기인하기 때문이다. EVS는 패킷 교환망의 이러한 문제점들을 효율적으로 극복하기 위해 Jitter Buffer Management(JBM), Packet Loss Concealment(PLC) 등의 기능을 디코더에 내장하여 단말들의 성능을 상향 평준화 하고자 하였다. 2001년 AMR-WB가 마지막으로 표준화된 이후 13년 만에 새로운 코덱이 표준화되었고 그동안의 반도체 기술의 발전을 감안하여 연산능력과 메모리 요구사항은 다소 상향되었다. 이에 따라 다양한 입력 신호에 대해 더 우수한 성능을 보장하며 다양한 기능이 구현되었다. 실제 동작모드는 VBR 2(NB, WB), NB 4, WB 11, AMR-WB 호환모드 9, SWB 8, FB 7등 총 41가지로 구성되어 1992년 GSM 상용화 이후 ETSI, 3GPP에서 개발되어 온 모든 음성 코덱의 동작모드 수의 합보다 많다.



[그림 1] EVS Encoder 구조

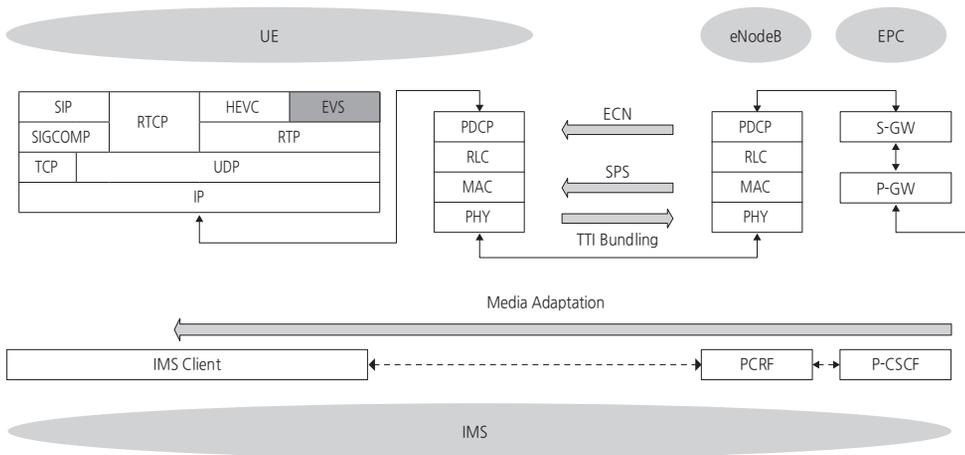
4. EVS 개발결과

설계기준 확립 후 2011년부터 본격적인 기술개발이 시작된 EVS는 2013년 상반기에 제안된 기술 중 상위 5건을 선정하는 Qualification Phase를 진행하여 삼성전자, NTT, Fraunhofer, Qualcomm, Huawei의 5사의 제안기술이 이 단계를 통과하였고 이들 5사의 기술들을 바탕으로 표준화가 진행되어 성능평가 후 2014년 말에 개발 완료될 예정이다.

4.1 EVS Encoder 구조 분석

[그림 1]은 <표 1>의 요구사항을 만족하도록 개발된 EVS 코덱의 엔코더 구조이다. 기존 음성압축 기술들과의 가장 큰 차이점은 입력신호의 압축이 음성 코어(Voice Core)와 오디오 코어(Audio Core)에서 이루어지며 입력신호를 분석하여 적절한 코어를 동적으로 할당하는 데 있다. 음성 코어는 기존 기술들에서 사용되어 오던 음성압축에 특화된 Code-Excited Linear Prediction(CELP) 방식으로 입력신호를 압축하며 광대역(WB, Wideband), 초광대역(SWB, Super Wideband), 전체대역(FB,

Fullband) 모드에서는 고주파 대역 신호를 저주파 대역 주파수 파형을 이용하여 압축하는 Bandwidth Extension(BWE) 기법으로 처리한다. 오디오 코어는 MP3, AAC 등 음악 파일의 압축과 저장에 사용되는 Modified Discrete Cosign Transform(MDCT) 방식을 사용하여 넓은 대역의 입력신호를 처리하게 되며 음성 코어처럼 BWE 기법을 적용하여 고주파 대역의 신호를 효율적으로 압축한다. 각 음성 프레임의 내용을 분석하여 적절한 코어와 압축방법을 적응적으로 선택하여 최적의 통화품질과 네트워크 용량을 실현하는 구조를 갖고 있다. 이러한 병렬코어 구조는 MPEG에서 2009년 개발된 Unified Speech and Audio Coding(USAC) 코덱의 구조와 유사하나 USAC가 특정 유무선 통신 네트워크를 고려하여 설계되지 않아 코덱의 algorithm 지연이나 복잡도 측면에서 실시간 통화용으로 사용하기 어려운 반면 EVS는 HSPA, LTE 등 3GPP 패킷교환 네트워크에 최적화되어 개발되었다. EVS 각 모드의 비트 수가 Transport Block Size(TBS) 등 모뎀의 파라미터에 맞추어져 있으며 디코더는 JBM, PLC 등 이러한 네트워크에서 빈번히 발생하는 망 부하



[그림 2] Voice over LTE (VoLTE) 시스템 구조

혹은 패킷 손실로 인한 품질손상을 최소화하는 구조를 갖고 있다. 이에 더해 EVS는 기존의 VoLTE의 기본 codec인 AMR-WB와 비트스트림간의 호환성을 유지하는 AMR-WB 호환(IO, Interoperable) mode를 지원하여, 기존 사용자와도 원활한 서비스를 지원할 수 있다.

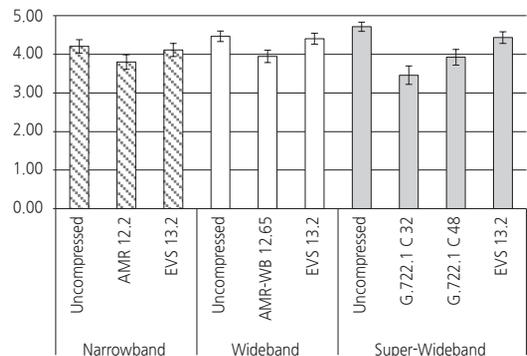
4.2 VoLTE 연동

[그림 2]에서 VoLTE 시스템은 단말(UE)과 기지국(eNodeB), 코어 네트워크(EPC), 그리고 IMS로 구성된다. 음성, 오디오 코어는 입력신호에 의해서만 조절되는 것이 아니며 네트워크에 의해서도 조절되는데 [그림 2]에서 설명된 VoLTE 시스템의 구조도에서 상대 단말과 음성 패킷이 전송되는 구간에 있는 기지국과 라우터가 패킷의 IP 헤더에 임박한 망 부하를 표시하는 Explicit Congestion Notification (ECN) 또는 단말이 RTP 헤더나 RTCP 신호를 통해 직접 비트 레이트 조절을 요청하는 미디어 적응(Media Adaptation) 기능이 사용될 수 있다. 전송 상황이 극히 악화될 경우 통화품질을 희생하고 음성정보를 중복해서 전송하는 Channel-aware 모드,

Partial-redundancy 모드 등도 제공된다.

4.3 성능평가 결과

[그림 3]은 Qualification Phase를 통과한 EVS 코덱의 성능을 전문 평가업체(Dynastat)에서 기존기술들과 Mean Opinion Score(MOS) 기준으로 비교한 결과이다[2]. 본 평가결과에서는 0.1 이상의 차이가 있을 경우 품질의 차이가 통계적으로 의미가 있으며, [그림 3]은 협대역, 광대역, 초광대역 3가지 조건에서의 상대평가 결과를 나타낸다. EVS는 협대역 13.2kbps 모드에서 MOS=4.11로 AMR(3.80) 대비



[그림 3] EVS 성능평가 결과

큰 폭의 성능향상을 가능하게 하며 디지털화 후 압축하지 않은 상태의 음성신호 품질(4.21)에 근사한 수준이다. 광대역 13.2kbps 모드에서도 EVS는 MOS=4.41로 AMR-WB 12.65kbps(3.95) 대비 크게 향상된 품질을 제공하며 비 압축상태의 품질(MOS=4.47)에도 근접한다. EVS WB 13.2는 현재 국내 VoLTE 서비스에서 사용되는 AMR-WB 23.85kbps 수준의 통화품질을 절반 수준으로 제공할 수 있으므로 각 사업자의 LTE 주파수 대역 중 전국망, 해외단말의 로밍에 사용되어야 하는 가장 중요한 대역의 용량을 크게 증가시킬 수 있을 것으로 예상된다.

EVS가 가장 강점을 가질 것으로 예상하는 초광대역에서 EVS는 13.2kbps(MOS=4.44)으로 기존 초광대역 코덱인 G.722.1 C 32 kbps(MOS=3.46), 48 kbps 을 모두 능가한다. 본 결과는 오류가 없는 채널에서 평가하였으며 EVS에 내장된 JBM와 PLC로 인해 오류가 있는 전송조건에서 이러한 차이는 그대로 유지되거나 더욱 커진다[2]. EVS를 구성하는 각 모듈의 최적화가 완료되면 EVS의 성능은 이 수준을 더욱 상회할 것으로 예상된다.

5. 맺음말

지금까지 4세대 이동통신을 위한 음성압축기술 EVS의 표준화 배경과 설계목표, 그리고 성능평가 결과를 요약하였다. 음성압축기술은 무선 액세스 기술보다 긴 교체주기를 가지며 그만큼 목표 설정과 개발 기간이 길게 소요되는 경향이 있다. EVS는 2008년부터 부분적으로 상용화가 시작된 4세대 이동통신 시스템의 마지막 핵심기술로 2, 3세대 음성압축 기술과 근본적으로 다른 설계철학과 목표에 기반하여 개발되어 압도적인 통화 품질과 망 용량을 가능하게 하여 VoLTE 확산을 위한 기폭제 역할을 수행함은 물론 다양한 입력신호를 융통성 있게 소화하여 새로운 서비스와 단말의 출현을 가능하게 할 것이다. 

[참고문헌]

- [1] 3GPP S4-110710 EVS Permanent Document #4 (EVS-4): EVS design constraints 1.0 (2011)
- [2] 3GPP S4-130292 EVS Qualification Test Global Analysis Report #2, Complete Results for the Subjective Tests and Results for Terms of Reference Tests (2013)

정보통신 용어해설

입체 영상 Stereoscopy [방송]

독립적인 두 개의 카메라 렌즈를 통하여 왼쪽 영상과 오른쪽 영상으로 획득된 스테레오스코픽 영상(입체 영상).

*유사어: 3D 영상

