

이동통신에 적용되는 음성통화 기술 및 코덱

글. (사)한국정보통신기술사회 방송통신전파분과위원회

개요

이동통신은 이동 중에도 음성 및 데이터를 원활하게 이용하려는 요구에 맞추어 지속해서 발전해 왔습니다. 최근 영상 서비스와 SNS 이용이 급증하면서 데이터 사용량이 크게 증가하고 있으며, 많은 이용자들이 데이터 전송 기술의 발전(고속, 저지연, 저손실)은 쉽게 체감하는 반면, 상대적으로 이용 시간이 짧은 음성통화 기술의 발전은 크게 인식하지 못하는 경우가 많습니다. 그러나 이동통신 세대가 변화함에 따라 음성통화 기술 또한 지속해서 발전해 왔습니다. 본 글에서는 2세대 이후 이동통신의 주요 세대별 음성통신 기술과 아날로그 음성을 디지털로 변환하는 음성코덱 기술에 대해 살펴보려고 합니다.

이동통신의 특징과 기술의 발전

전송로를 기준으로 통신을 구분하면 무선통신과 유선통신으로 나눌 수 있으며, 이동통신의 가장 큰 특징은 무선통신이면서도 이동성을 고려한 서비스라는 점입니다.

Microwave 통신과 같은 고정 무선통신의 경우, 유선에 비해 구조가 다소 복잡하지만, 전송로의 변화가 날씨나 장애물과 같은 비교적 단순한 요인에 의해 결정된다고 볼 수 있습니다. 반면, 이동통신은 단말이 이동하는 환경을 전제로 해야 하므로 도플러 효과나 페이딩과 같은 급격한 채널 변화에 대응하는 기술이 필수적이며, 배터리 사용 시간, 단말의 크기, 안테나 성능 등의 이동형 단말 제약 사항을 고려해야 하는 어려움이 있습니다. 따라서 이동통신 기술은 성능 향상과 함께 이러한 제약을 극복하기 위한 다양한 기술이 지속해서 발전해 왔습니다.

이동통신의 성능 발전은 이동성 개선과 수용량 및 속도 개선의 두 가지 측면에서 이루어졌습니다. 이동성 개선을 위해 디지털 시스템이 도입되었으며, 디지털 신호처리 기술을 적용함으로써 이동 중 발생하는 채널 변화를 빠르게 보정할 수 있도록 하였습니다. 또한, 수용량 및 속도 개선을 위해 이용 주파수 대역폭을 확대하고, QAM(Quadrature Amplitude Modulation)과 같은 변조 방식을 도입하였으며, Multiple Access 기술을 활용하여 동시 접속 가능한 단말 수를 증가시켰습니다. 이와 함께 MIMO(Multiple Input Multiple Output) 기술을 적용하여 공간 다이버시티 이득을 확보하고, 음성 압축률을 높이는 기법을 채택하면서 지속적인 성능 향상을 이루어 왔습니다.

또한 이동통신의 제약 사항을 극복하기 위한 기술도 발전해 왔습니다. 무선 환경에 따라 데이터 복구 확률을 높이는 AMC(Adaptive Modulation and Coding) 기술이 적용되었으며, Diversity 기법을 통해 선택적인 무선 성능 향상을 이루었습니다. 또한, 단말의 배터리 사용 시간을 절감하기 위해 Sleep Mode 기

능이 도입되었고, Hybrid ARQ(Hybrid Automatic Repeat Request) 기법을 활용하여 데이터 복구 속도를 높이는 등 전송 신뢰성을 강화하는 다양한 기술이 지속적으로 발전하고 있습니다.

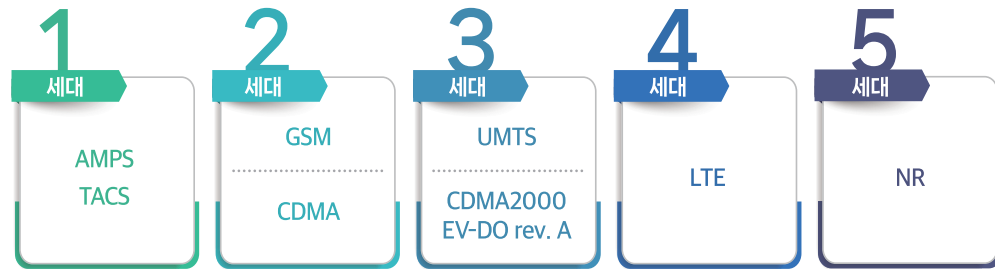


그림 1. 이동통신 세대별 구분

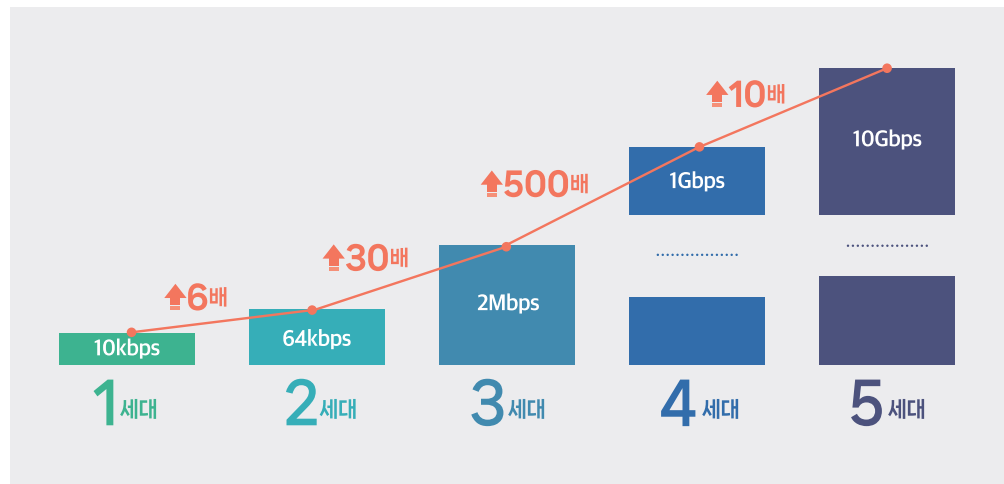


그림 2. 이동통신 세대별 속도 / 출처 : 중앙전파관리소 자료 가공

음성통화 기술

아날로그 신호인 소리를 디지털로 변환하는 이유 중 하나는 디지털 처리 기법을 활용하면 채널 변화로 인해 발생하는 오류에 효과적으로 대응할 수 있기 때문입니다. 디지털 변환 방식 중 가장 기본적인 PCM(Pulse Code Modulation)은 상대적으로 많은 대역폭이 필요하기에 이동통신 환경에서는 적합하지 않습니다. 이동통신에서는 방송과 달리 음악과 같은 다양한 음원보다는 사람의 목소리라는 단순한 음원의 특성을 분석하고 이를 최적화하여 전송되는 정보의 양을 줄이는 데 초점을 맞추고 있습니다. 이를 위해 도입된 기술이 바로 보코더(Vocoder = Voice + Encoder)입니다. 보코더는 음성의 특징적인 파라미터를 추출하고 이를 압축하는 것을 목적으로 합니다.

예를 들어, 일반적인 대화에서는 발화자와 청자가 번갈아 가며 대화를 진행하기 때문에, 연속적으로 말을 할 때는 짧은 주기로 음성 정보를 전송하지만, 상대방의 목소리를 듣고 있을 때는 전송할 소리가 없으므로 'Silent Packet'을 긴 주기로 전송하거나, 샘플링 주파수의 제약을 적용하고, 음성에서 나타나는 특정 주파수 변화의 특성만을 추출하는 등의 방식으로 효율적인 처리가 이루어집니다.

이동통신에서는 이러한 보코더 기술과 더불어, 세대별로 전송 가능한 데이터 속도, 기지국 및 채널 카드의 수용 용량 등을 고려하여 다양한 코덱을 적용함으로써 제한된 자원을 보다 효율적으로 활용할 수 있도록 지속해서 기술을 발전시키고 있습니다.

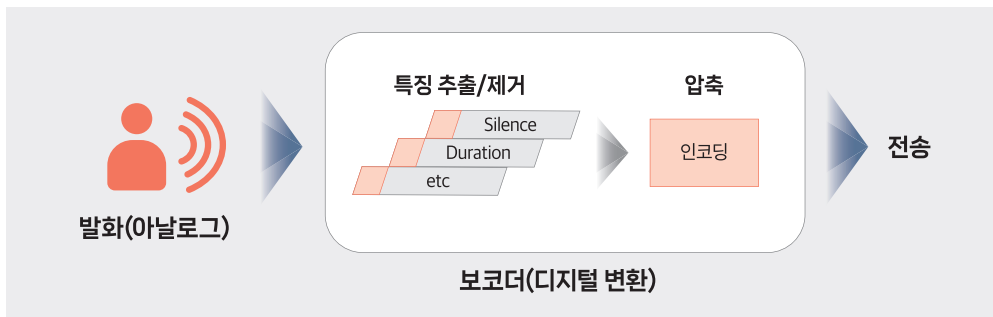


그림 3. 보코더 개요도

이동통신 세대별 음성통화 방식

1) 2세대(CDMA, GSM) 음성통신

2세대 통신 방식에는 국내에서 사용하는 동기식 방식인 CDMA(Code Division Multiple Access)와 비동기식 방식인 GSM(Global System for Mobile Communications)이 있습니다. 2세대 이동통신에서는 통화 전송로를 점유하여 사용하는 서킷 서비스 방식을 채택하였습니다. 서킷 서비스 방식의 주요 특징으로는, 통신 연결 경로가 사전에 결정되어 전송되므로 안정적인 서비스 제공이 가능하며, 일정한 품질의 통신이 유지될 수 있다는 장점이 있습니다. 그러나 회선을 점유하는 방식이기 때문에 전체 네트워크의 효율성이 저하될 수 있다는 단점도 존재합니다. 초기 2세대 이동통신 서비스는 음성통화 전용으로 설계되었으나, 기술이 발전하면서 GPRS(General Packet Radio Service)나 IS-95B와 같은 기술을 통해 음성과 데이터 서비스를 동시에 제공할 수 있게 되었습니다.

2세대 음성통신에 사용된 코덱을 살펴보면, CDMA의 경우 초기에는 QCELP(Qualcomm Code Excited Linear Prediction)를 사용하였으며, 이후 음성 품질 향상을 위해 EVRC(Enhanced Variable Rate Codec)가 도입되었습니다. 두 코덱 모두 8kbps의 전송 속도를 가지며, 특히 EVRC는 잡음이 심한 환경에서 음성 신호가 포함된 주파수 영역을 강조하고, 그 외의 대역은 억제하는 소프트웨어 기법을 적용하여 음성 품질을 개선하는 특징을 가지고 있습니다. 한편, GSM에서는 AMR(Adaptive Multi-Rate) 방식을 채택하였습니다. AMR 코덱은 3세대 이동통신에서도 활용되는 방식이므로 이에 대한 자세한 설명은 3세대 음성통신에서 다루도록 하겠습니다.

2) 3세대(UMTS : Universal Mobile Telecommunication Service) 음성통신

3세대 이동통신부터는 음성과 데이터를 동시에 제공할 수 있도록 시스템이 구성됩니다. 음성은 서킷 서비스 방식으로 제공되며, 데이터는 패킷 서비스 방식으로 제공됩니다. (2세대 통신도 기술 발전에 따라 2.5세대인 GPRS와 같은 기술에서는 패킷 서비스가 동시에 제공되기도 합니다.) 서킷 서비스와 패킷 서비스 방식을 함께 사용하게 되면서, 서킷 서비스를 제공하는 네트워크 파트와 패킷 서비스를 제공하는 네트워크 파트가 분리되어 망 구성이 복잡해졌고, 서킷 서비스와 패킷 서비스의 도메인을 별도로 관리해야 하는 단점이 있었습니다. 그러나 2세대와 달리, 3세대에서는 빠른 데이터 전송이 가능해졌고, 이를 통해 영상 통화와 같은 멀티미디어 서비스가 가능해졌습니다.

3세대 UMTS의 음성코덱은 GSM에서 사용하던 AMR을 그대로 사용하였습니다. AMR 코덱은 4.75kbps

에서 12.2kbps 사이의 8가지 전송률을 제공하며, 음성패킷 내의 CMR(Codec Mode Request)을 통해 비트레이트를 동적으로 변경할 수 있습니다. AMR은 200Hz에서 3,400Hz의 대역을 처리할 수 있으며, 샘플링 주파수는 8kHz를 사용합니다. AMR 코덱은 음성이 없는 구간에서는 Comfort Noise를 발생시키고, 패킷 발생 주기를 늘려 전송량을 줄여 효율성을 높이는 방식으로 동작합니다.

AMR 코덱은 처리할 수 있는 대역폭에 따라 AMR-NB(Narrow Band)와 AMR-WB(Wide Band)로 구분됩니다. 엄밀히 말하면, 기존 GSM에서 사용되던 AMR 코덱은 AMR-NB이며, 우리나라의 3세대 이동통신에서도 주로 AMR-NB가 사용되었지만, 일부 단말기에는 AMR-WB가 탑재되어 사용할 수 있었습니다.

3) 4세대(LTE : Long Term Evolution) 음성통신

4세대 이동통신은 UMTS의 복잡한 구조를 단순화한 특징이 있습니다. 이를 위해 네트워크 측면에서는 서킷 서비스 파트를 제외하고, 패킷 서비스만을 제공하는 All IP망을 구성하였으며, 네트워크와 단말 간에는 복잡한 상태 관리 대신 접속(Connected)과 휴지(Idle) 두 가지 상태만을 정의하여 효율성을 높였습니다.

이와 같이 서킷 서비스 구조를 제외하면서 음성 서비스를 제공하기 위해 다른 네트워크를 사용해야 했습니다. 초기 LTE 네트워크에서는 3G망을 활용하여 음성 서비스를 제공하는 방식을 사용했으며, 이를 CSFB(Circuit Switched FallBack)라고 합니다. CSFB 방식에서는 음성통화 착/발신이 발생하면 3G 네트워크로 전환된 후, 3G 시그널링을 통해 음성통화 접속 절차가 진행됩니다. 통화가 끝나면 3G 네트워크에서 전달하는 정보 메시지를 통해 해당 셀이 LTE 영역으로 돌아갈 수 있는지 확인하고, LTE로 되돌아가는 과정을 거칩니다. 이 방식은 네트워크 투자비를 절감할 수 있는 장점이 있지만, 호처리의 복잡성 및 LTE와 UMTS 코어망 간의 정보 교환, 그리고 호 연결 지연 등의 단점이 존재합니다.

이를 보완하기 위해 IMS(Internet Multimedia Subsystem)를 도입하고, LTE에서 음성을 제공하는 VoLTE(Voice over LTE) 서비스가 시작되었습니다. VoLTE는 QoS(Quality of Service)를 확보하기 어려운 VoIP(Voice over IP)와 달리, LTE QoS 관리 방식을 통해 음성 품질을 보장합니다. 음성에 필요한 최소 전송 속도를 GBR(Guaranteed Bit Rate)로 보장하고, 음성통화 연결 시 SIP(Session Initiation Protocol)를 이용하여 음성패킷의 QoS를 보장하기 위해 전용 전송로를 단말과 네트워크 간에 구성하여 해당 전송로의 QoS를 최대로 높이는 방식으로 음성 품질을 보장합니다.

LTE는 기존 서비스 대비 넓은 대역폭과 높은 변조방식을 적용할 수 있기 때문에 초기 VoLTE는 AMR-NB의 확장인 AMR-WB 코덱을 사용했습니다. AMR-WB는 AMR-NB보다 확장된 16kHz의 샘플링 레이트와 20~7,000Hz의 대역폭을 사용하기 때문에 음질이 보강된 음성통화를 제공할 수 있으며, AMR-NB와 같은 가변 비트레이트(6.6~23.85kbps)를 제공하여 채널 환경에 따라 변경 가능하기 때문에 이동통신에 적합한 코덱으로 자리잡았습니다.

초기 VoLTE에서는 AMR-WB 코덱을 사용했지만, 기술 발전에 따라 고품질 음성코덱의 필요성이 커졌고, EVS(Enhanced Voice Service) 코덱이 추가로 적용되었습니다. 기존 음성코덱은 제한적인 샘플링 주파수로 인해 손실이 발생하는데, EVS는 다양한 샘플링 주파수와 비트레이트를 도입하여 더욱 향상된 음성 품질을 제공합니다. EVS 코덱은 8kHz에서 48kHz까지 샘플링 주파수를 지원하며, 오디오 대역폭도 NB(Narrow Band), WB(Wide Band), SWB(Super Wide Band) 외에도 옵션으로 FB(Full Band)도 사용할 수 있습니다. 또한, 비트레이트는 AMR-WB보다 더 넓은 5.9~128kbps를 선택할 수 있으며, 통화 중에는 코덱 변경이 AMR과 유사한 방식으로 이루어집니다.

4세대 초반 이후, 국내 사업자들은 자사 간 통화에 EVS 코덱을 적용했지만, AMR 코덱도 함께 사용하여 호처리 시 코덱 협상이 이루어집니다. 코덱 협상은 Offer/Answer 모델을 사용하여 발신 단말이 가능한 모든 코덱 리스트를 전송하면, 착신 단말은 이를 기반으로 지원하는 코덱과 우선순위에 따라 응답하는 방식으로 통화가 이루어집니다. 호처리 과정에서 EVS와 AMR 코덱은 SDP(Session Description Protocol) 메시지에 함께 전송되어, EVS 코덱을 지원하지 않는 단말과도 통화가 이루어질 수 있습니다. 또한, G.711 등의 레거시 코덱과의 통화는 네트워크 내부에서 트랜스코딩을 통해 AMR로 변환되어 진행되기도 합니다.

4) 5세대(NR : New Radio) 음성통신

5세대 이동통신(5G)의 가장 큰 특징은 4세대와 비교하여 훨씬 넓은 주파수 대역폭을 사용하고, 확장된 MIMO 기술을 통해 전송 성능을 크게 향상시킨다는 점입니다. 또한, 클라우드 및 가상화를 활용하여 유연한 코어망 구조를 제공하고, 다양한 기술을 통해 QoS와 보안 효율성을 확장하는 데 중점을 두고 있습니다. NR(New Radio) 역시 LTE와 마찬가지로 All IP 구조를 가지고 있어 IMS를 통한 음성통신을 제공하지만, 네트워크의 구조에 따라 통화 방식에 차이가 발생할 수 있습니다.

5G 초기에는 규격이 완벽하게 확립되지 않아 액세스망은 LTE와 5G를 혼용하고, 코어는 LTE(EPC)를 사용하는 NSA(Non-Stand Alone) 방식으로 구축되었습니다. 이후에는 5G 액세스와 5G 코어(5GC)를 사용하는 SA(Standalone) 방식으로 전환할 수 있게 되었습니다.

NSA 방식은 이동성 관리와 같은 제어 작업이 EPC를 통해 이루어지며, LTE와 5G를 동시에 사용할 수 있어 용량 측면에서는 장점이 있을 수 있습니다. 하지만, 5G의 핵심 기술인 초저지연통신(URLLC, Ultra-Reliability and Low Latency Communication)과 같은 유즈 케이스를 제공하기 어려운 단점도 있습니다. NSA에서는 LTE 네트워크를 사용하기 때문에 음성통화는 LTE와 동일하게 VoLTE(Voice over LTE) 방식을 사용합니다.

SA 방식에서는 두 가지 음성 서비스 제공 방법이 존재합니다. 첫 번째 방법은 초기에 LTE와 비슷하게 LTE 네트워크로 전환 후 VoLTE를 사용하는 방식으로, 이를 EPS-FB(Evolved Packet System FallBack)라고 부릅니다. 이 방식은 LTE와 5G를 오가며 음성통화를 진행하기 때문에, CSFB 방식처럼 호처리 및 가입자 정보 관리에 어려움이 있을 수 있습니다. 두 번째 방법은 LTE와의 연동 없이 5G 액세스와 코어를 통해 음성통화를 직

접 처리하는 방식으로, 이를 VoNR(Voice over NR)이라고 합니다. VoNR은 5G 규격에서 정의하는 QoS를 제공받아 음성 품질이 더 신뢰성이 높고 향상된 통화를 제공할 수 있습니다.

5G 시대에 진입했지만, 음성코덱은 여전히 LTE에서 사용되던 EVS와 AMR 코덱을 계속 사용하고 있습니다.

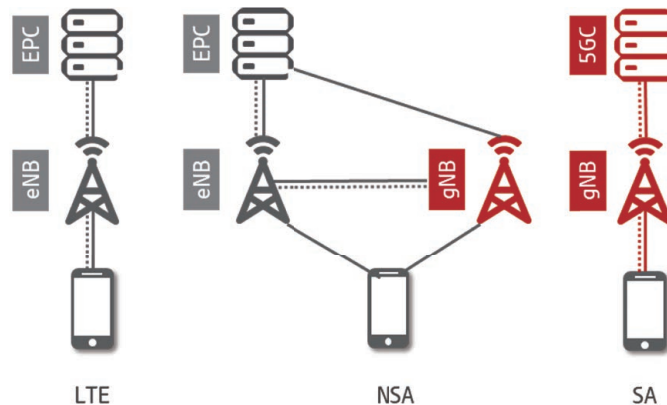


그림 4. LTE와 5G SA/NSA

향후 발전 방향

6G 네트워크의 발전은 주로 패킷망 기반으로 진행될 것이며, 저궤도 위성을 활용한 광역 커버리지 구성과 더 빠른 전송 속도를 지원하는 기술이 적용될 것으로 예상됩니다. 이러한 발전 방향을 반영하여 음성 통화의 코덱도 두 가지 주요 방향으로 발전할 것으로 보입니다. 하나는 전송 효율성을 극대화하는 방법이고, 다른 하나는 사용자의 체감 품질을 향상시키는 방법입니다.

전송 효율성 측면에서는 위성을 이용한 음성통신에서 적은 용량의 데이터 전송이 중요합니다. 수백 킬로미터 이상 떨어진 위성으로 데이터를 송신하려면 단말기의 출력이 높아져야 하지만, 현재와 같은 스마트폰 폼팩터에서는 안테나와 배터리의 제한이 존재합니다. 이를 해결하기 위해 시맨틱 코덱과 AI 기반 압축률 향상 방법 등이 연구되고 있으며, 이러한 기술을 통해 기존 코덱보다 적은 데이터로 동일한 품질의 음성을 전달할 수 있다는 결과들이 발표되고 있습니다.

체감 품질 측면에서는 스마트폰 카메라의 화질 발전과 VR 기기의 대중화에 따라 화상 통화와 실감 통화 서비스가 중요해지고 있습니다. 이에 따라 이러한 서비스를 지원하기 위한 새로운 코덱이 적용될 것으로 예상됩니다. 현재의 음성코덱은 주로 말(Speech) 중심의 압축 방식에 집중하고 있지만, 미래에는 주변 소리까지 함께 전달하여 현장감을 높일 수 있는 음성 전송 방식이 채택될 것으로 보입니다. 이는 사용자에게 더욱 몰입감 있는 통화 경험을 제공할 것입니다.

맺음말

지금까지 이동통신에서 사용되는 음성통화 방식과 코덱에 대해 살펴보았습니다. 음성통신은 신뢰성을 높이기 위한 기술을 요구하며, 세대별로 이를 달성하기 위해 QoS 기술, 디지털 프로세싱 적용 등 다양한 기술이 반영되었습니다. 음성코덱은 높은 압축율을 제공함으로써 전송되는 정보의 양을 줄이면서도 체감되는 통화 품질을 높이는 방향으로 발전해 왔습니다.

다가올 미래에는 실감통신과 같은 고효율의 코덱을 활용한 커뮤니케이션이 더욱 중요해질 것입니다. 이를 위해 높은 전송률과 낮은 에러율을 구현할 수 있는 이동통신 기술과 다채널 오디오와 같은 체감적인 요소가 반영된 코덱의 개발이 필요합니다. 또한, 단말의 프로세싱 파워가 증가하면서 방송에서 사용되는 고급 코덱을 이동통신용 단말에서도 사용할 수 있는 가능성도 기대할 수 있습니다. 이와 같은 기술 발전은 음성통화뿐만 아니라 멀티미디어 커뮤니케이션에서도 중요한 역할을 할 것입니다. 

