

소 특 집

광대역 음성부호화 알고리즘 및 그 응용

이인성, 신재현

충북대학교 전기전자공학부

I. 서 론

차세대 이동 통신 및 인터넷 통신 시스템의 가장 큰 특징은 넓은 대역폭을 사용하며, 음성에서 동영상까지 다양한 멀티미디어 데이터를 전송할 수 있다는 것이다. 기존의 음성 위주의 통신 방법과 달리 데이터와 영상 전송이 가능하게 되며, 음성통신에 있어서도 기존의 200-3400 Hz의 협대역 음성 신호 뿐만 아니라 50-7000 Hz의 광대역 음성 통신도 가능하게 된다. 200-3400 Hz의 전화신호 대역폭에서 50-7000 Hz의 넓어진 대역폭으로 증가시킴으로 음성의 명료도와 자연성을 증가시킬 수 있으며, 바로 앞에서 대화하는 느낌을 얻을 수 있다. 협대역 음성과 비교하면 50-200 Hz의 낮은 주파수의 포함은 자연성과 현실성, 편안함을 증가시키며 3400 Hz로부터 7000 Hz까지의 높은 주파수의 포함은 더 정확한 마찰음의 구분으로 인해 더 높은 명료도를 증가시킴으로 화자 인식의 용이성을 가져다 준다.^[6]

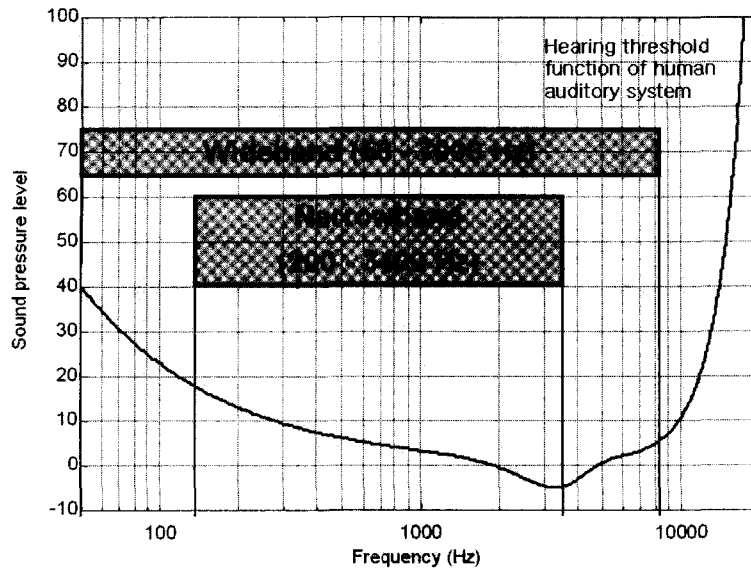
광대역 음성신호는 디지털 전송을 위해 16 kHz 주기로 샘플링된다. 각각의 샘플이 16 bit 정수로 표현한다면 실제 비트률은 256 kbit/s가 된다. 이러한 음성 데이터를 낮은 전송률에서 전송하기 위해서는 효율적인 음성 부호화 알고리즘이 필요하게 되며, 최근 ITU-T, 3GPP/ESTI에서는 광대역 음성부호화 표준화 작업을 수행하여 16 kbit/s대의 전송률 광대역 음성 부호화 표준 방식을 채택하였다. 3GPP/ESTI와 ITU-T에 의해 채택된 방식은 협대역 음성부호화 방식에서 많이 사용하는 Algebraic Code Excited Linear

Prediction(ACELP) 구조의 적응 다전송률 광대역 부호화(AMR-WB: Adaptive Multi-Rate WideBand) 방식이다. 6.6 kbit/s-23.85 kbit/s까지의 전송률에서 동작되는 AMR-WB 방식의 선정은 최초의 유선과 무선에 같은 방식의 선정이라는 상당한 중요한 의미를 갖게 된다. 유무선상에서 동일한 방식을 선정함으로써 유무선 통신 시스템간의 트랜스코딩에 대한 필요성을 제거하고, 통신 시스템 분야의 광범위한 응용 및 서비스의 구현을 쉽게 가능하게 할 것이다.^[5]

본 논문은 II장에서 ITU-T에서 선정한 광대역 음성부호화 방식인 G.722^[1], G.722.1^[2], G.722.2^[3] 광대역 음성 부호화 알고리즘의 각 방식별 특징 및 구조를 설명하고, 장단점 비교 및 그 응용 분야를 언급하였으며, 특히 가장 최근에 채택된 AMR-WB 표준안을 보다 자세히 설명하였다. III장에서는 광대역 음성부호화 알고리즘의 응용 분야 및 그 사용 가능성을 언급하였으며, IV장에서 결론을 맺는다.

II. 광대역 음성부호화 알고리즘

협대역 음성신호(NB: NarrowBand, 30 Hz-3400 Hz)에서는 '스' 발음 등의 치찰음에서 중요한 성분인 높은 주파수를 표현할 수 없는 문제점을 갖고 있는데 반해, 광대역 음성신호(WB: Wide-Band, 50-7000 Hz)에서는 자연스러움, 음성의 명확성을 증가시킬 수 있다. 광대역 음성 신호를 사용함으로써 현재 셀룰러 시스템이나 PSTN에



〈그림 1〉 광대역 음성의 주파수 대역폭 및 인간 청각인지 시스템^[7]

서 사용하는 협대역 음성신호 서비스들보다 우수한 음질 서비스가 제공 가능하며 높은 음질의 오디오 서비스의 제공이 더욱 용이하게 된다. 〈그림 1〉은 광대역 및 협대역 신호의 주파수 대역폭 및 인간의 청각 인지 시스템을 나타낸다. 광대역 음성 신호는 인간의 청각 인지에서 중요한 3400-7000 Hz까지의 대역을 포함하고 있어 더욱 자연스러운 음질을 유지할 수 있다^[7].

음성부호화 표준 작업은 ITU-T 중심으로 협대역 음성 신호에서 주로 수행되어졌으며, 64 kbit/s μ -law, A-law PCM 방식(G.711), 32 kbit/s ADPCM(G.726), 16 kbit/s LD-CELP(G.728), 8 kbit/s CS-ACELP(G.729), 6.3 kbit/s 및 5.3 kbit/s의 G.723.1 방식이 광범위한 분야에서 사용되고 있다. 음성 부호화 연구의 대부분이 협대역 음성신호에만 집중되어 있었지만 더 넓은 대역폭까지 포함하는 입력 음성으로 확대함으로써 음성 품질을 개선을 이룰 수 있었다. 1988년에 최초의 광대역 음성 부호화 표준인 G.722를 ITU-T에서 결정되었으며, G.722는 구현이 간단하고 48, 56, 64 kbit/s의 전송률에서 좋은 성능을 나타낸다. 1995년에 ITU-T에서 광대역 부호화기에 대한 새로운 표준화 작업이 시작되었

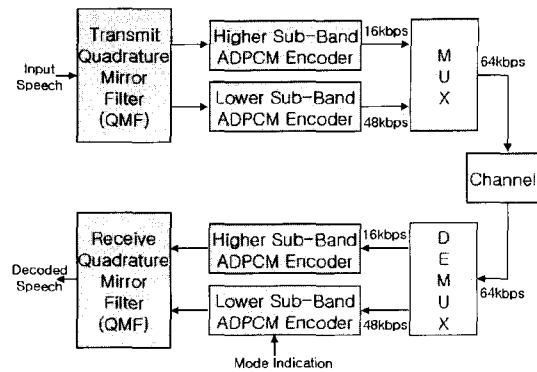
고, 1999년에 24.32 kbit/s 전송률의 G.722.1을 결정하였다. 더욱 최근에는 새로운 광대역 음성 부호화 표준 작업이 16 kbit/s 주변의 전송률에서 ITU-T와 ETSI/3GPP에서 수행되었다. AMR-WB라 불리는 새로운 부호화 방식은 GSM과 3세대 무선시스템에서 광대역 부호화기로 3GPP에 의해 2000년에 결정되었으며, 그 후 AMR-WB 방식은 ITU-T 표준화에 참여하여 2001년에 6.6 kbit/s-23.85 kbit/s의 전송률에서 광대역 부호화기 표준으로 선정되었다.^[6] 본 장에서는 이러한 광대역 음성부호화 표준 방식들의 특징 및 구조를 중심으로 설명하며 3가지 방식의 주요 특징을 비교할 예정이다. 특히 AMR-WB의 주요 구조는 더욱 자세히 설명할 예정이다.

1. ITU-T G.722 방식^[1]

G.722 방식은 1988년 권고된 16 kHz의 샘플링률을 갖는 광대역 음성 신호 적용에 대한 64 kbit/s ITU-T 표준 방식이다. 본래 G.726 협대역 표준과 같은 기술인 ADPCM(Adaptive Differential Pulse Code Modulation) 방식을 사용하며 광대역 음성신호를 두개의 부밴드 신호로 나누어 각각 ADPCM 부호화 한다. 부호

화기를 위해 음성신호는 16kHz로 샘플링된 후, 같은 대역폭을 갖는 두 개의 부밴드로 분해된다. 부밴드의 분해는 유한 임펄스 응답을 가지는 QMF (Quadrature Mirror Filters)를 사용하여 수행된다. G.722 표준안에서 사용하는 QMF는 24개의 계수를 갖는다. 이러한 구성은 3ms의 전체 지연시간을 가지며, 전체 주파수 범위에서 1dB보다 적은 총손실을 가지게 된다. 복호화기에서 양자화된 부프레임 신호들은 부호화기에서 사용된 같은 QMF 필터로 다시 합성해낸다. QMF 필터를 사용함으로써 겹치는 부분에서의 왜곡을 제거할 수 있다. 복원된 부밴드 신호들은 합성신호를 만들기 위하여 서로 합해진다. 최적의 비트 할당률로 낮은 주파수 밴드에 6bits/sample를 높은 주파수 밴드에 2bits/sample를 할당한다. 팩스나 모뎀에서 데이터와 음성은 같은 채널에 전송하기 위하여 낮은 주파수 대역에서는 다양한 6, 5, 4bits/sample의 값을 사용할 수 있으며, 높은 주파수 대역에서는 2bits/sample로 고정한다. 낮은 주파수 대역에서의 1bit/sample씩의 감소는 데이터 전송을 위해 8kbit/s를 할당할 수 있다. 따라서 부호화기는 64kbit/s 오디오 전송, 56kbit/s 오디오+8kbit/s 데이터 전송, 48kbit/s 오디오+16kbit/s 데이터 전송 세 가지 모드로 동작할 수 있다.

기본 부호화기의 구조는 다음과 같다. 전체 대역폭의 입력 신호 $x(n)$ 은 높은 주파수 대역성분 $X_h(n)$ 과 낮은 대역성분 $X_L(n)$ 의 두 개의 부대역폭 신호로 분리되어진다. 0-8000 Hz의 주파수 대역폭은 0-4000 Hz와 4000-8000 Hz 두 대역으로 나뉜다. 0-4000 Hz의 낮은 주파수 대역은 높은 주파수 대역보다 매우 높은 비율의 신호 에너지를 가지고 있으며, 4000-8000 Hz의 높은 주파수 대역보다 주관적으로 더 중요한 대역이 된다. 낮은 주파수 대역은 MODE 1의 8000 sample×6bits/sample=48kbps에서 6bits/sample ADPCM 부호화 방법을 사용하여 부호화된다. 적은 중요성을 갖는 4-8 kHz 대역은 16 kbits/s에서 2bits/sample을 사용하여 부호화된다. 부호화된 신호들은 디지털 채널을 통해 전



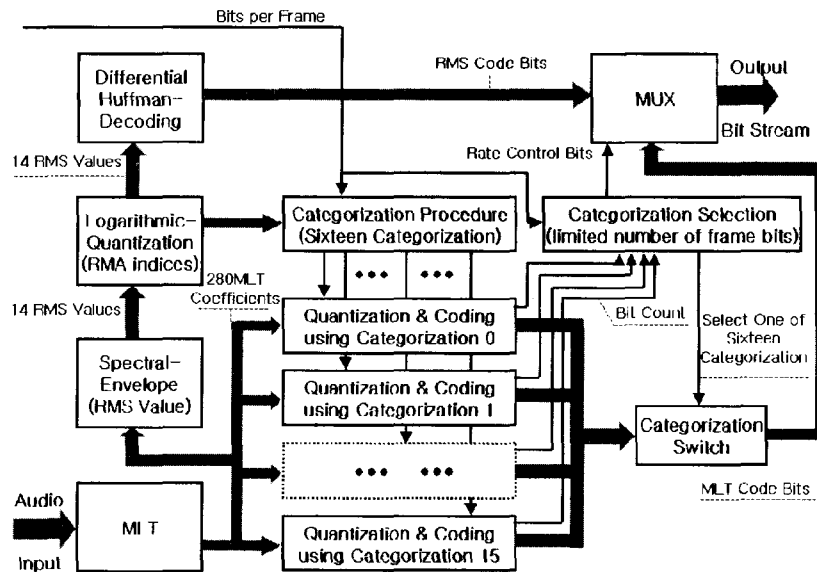
〈그림 2〉 G.722 광대역 음성부호화기의 구조도

송되기 위해 다중화된다^[5]. 수신단에서는 송신단의 역동작을 수행하며 최종적으로 QMF 합성필터를 사용하여 음성신호를 합성해낸다. 전체적 G.722의 구조는 〈그림 2〉에 나타나 있다.

2. ITU-T G.722.1 방식^[2]

1999년 표준화된 G.722.1 광대역 음성부호화 방식은 MLT (Modulated Lapped Transform)를 기반으로 지각 동기 심리음향 양자화 모델과 잔차 주파수 영역의 계수들을 부호화하기 위해 Huffman 부호화 방법을 사용한다. 16 kHz 주기로 샘플링된 320개의 오디오 샘플 (20ms)은 한 프레임을 구성하며 16 kbits/s, 24 kbits/s, 32 kbits/s 전송률에서 프레임당 320, 480, 640 비트를 출력한다.

〈그림 3〉에서의 부호화기의 구성도를 보면, 오디오 입력신호는 블록킹 효과에서 효율적인 MLT를 사용하여 시간 영역에서 주파수영역으로 투영된다. 현재 프레임의 320개 샘플들은 MLT에서 320개의 이전 블록과 함께 주어진 하나의 블록으로 구성된다. MLT를 통해 320개 주파수 영역 샘플들로 출력된다. 따라서 주파수 분해능은 8000 Hz/320=25 Hz를 갖게 된다. 7kHz 대역을 취하기 위해 280 주파수 신호 요소만을 취하며 부호화된다. MLT 계수들은 먼저 각각은 500 Hz의 주파수 범위로 표현되고 280/12=20개의 계수들을 14개의 같은 대역폭 범위로 그룹지어진다. 각각의 대역에서 RMS (Root Mean Square)



〈그림 3〉 G.722.1의 부호화기 구성도

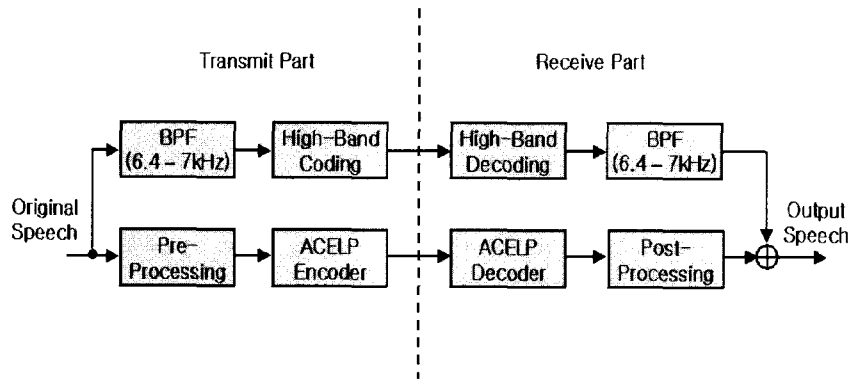
값은 스펙트럼 포폭선의 평가를 위해 계산되어지며, 각 대역의 RMS 값들의 도움으로 MLT 값은 대수적 영역으로 바뀌어지고, MLT 계수들은 지각 모델을 통하여 다른 단계별 크기를 사용하여 양자화 된다. 양자화 작업은 양자화와 부호화 파라미터의 어떤 집합이 각 영역에 할당되어진 범주로 여겨지는 초기 분류 값을 계산함으로써 수행된다. 총 16개로 분류되며 각 분류에 따라 비트 할당률이 계산되어진다. 최고의 비트 할당률이 정해진 후, 연관된 카테고리들의 파라미터로부터 MLT 계수들은 양자화되고, 최종적으로 허프만 부호화된다^[5].

3. AMR-WB(G.722.2)^[3]

AMR-WB 음성 부호화기는 AMR-NB와 ITU-T G.729, G.723.1에 사용되어진 ACELP (Algebraic Code Excitation Linear Prediction) 방식을 사용한다. AMR-WB 음성 부호화기는 전송률에 따라 23.85, 23.05, 19.85, 18.25, 15.85, 14.25, 12.65, 8.85, 6.6 kbits/s의 9개 모드로 동작된다. 또한 부호화기는 GSM의 비연속적인 전송 동작과 다른 시스템에서 저 전송률에 따른 소스 종속적인 배경 잡음에 사용

하기 위해 배경잡음모드를 포함하고 있다. GSM에서는 1.75 kbit/s의 전송률을 사용한다. 12.65 kbit/s 모드와 그 이상의 모드에서는 높은 음질의 광대역 음성을 제공한다. 최하위 두 모드 8.85 kbits/s와 6.6 kbit/s는 간단한 라디오 채널이나 네트워크 과부하 등에 일시적으로 사용하기 위해서 제안되었다.

AMR-WB 음성부호화기는 16 kHz 샘플링 주기에서 동작되며, 20 ms 프레임 단위로 부호화된다. 〈그림 4〉와 같이 50-6400 Hz와 6400-7000 Hz의 두 개의 주파수 대역을 나누어지며 계산량의 감소와 주파수 범위 내에서 중요한 정보들에 비트 할당을 맞추기 위해 각각 부호화된다. 낮은 주파수대역은 ACELP 알고리즘을 사용하여 부호화한다. 선형예측(LP) 분석은 매 20 ms 프레임당 한번씩 수행되며 16차 LPC 계수를 사용한다. 고정 및 적응 여기 코드북은 매 5 ms 마다 최적의 코드 값으로 검색된다. 모든 과정은 12.8 kHz 샘플링 주기에서 수행이 된다. 높은 주파수 대역은 낮은 대역의 파라미터와 랜덤 여기신호를 사용하여 복호화기에서 재구성된다. 고대역의 이득은 저대역의 음성 정보에 적절하게 조정된다. 고대역의 스펙트럼은 저대역에서



〈그림 4〉 AMR-WB의 전체적 구성도

의 LP 필터로부터 생성한 LP 필터를 사용하여 재구성된다.

AMR-WB 부호화기는 VAD(Voice Activity Detector), GSM의 DTX(Discontinuous Transmission)기능, 3G의 SCR(Source Controlled Rate)기능과 같은 고정 전송률 음성 및 채널 부호화 모드를 포함한다. AMR-WB는 유력한 무선 환경에서 음질에 최적화되도록 음성 채널 부호화 사이의 전송률이 능동적으로 변할 수 있는 장점이 있다. 또한 현존하는 협대역 표준안들보다 음질이 월등한 반면, 다전송률이 고 적응적이기 때문에 전송 에러에 매우 강인하다. AMR-WB 방식은 GSM full-rate 채널, GSM EDGE Radio Access Network(GERAN) 8-Phase Shift Keying(8-PSK) Circuit Switched 채널, UTRAN(3G Universal Terrestrial Radio Access Network) channel, VoIP(Voice over Internet Protocol)을 위한 패킷 어플리케이션 등 다양한 통신망에 사용 가능하도록 설계되어졌다.

1) AMR-WB의 부호화기

AMR-WB 음성 코덱은 비트율이 6.6, 8.58, 12.65, 14.25, 15.85, 18.25, 19.85, 23.05, 23.85 kbits/s의 9개의 음성부호화기로 구성되어 있다. 각 부호화기는 CELP(Code-Excited Linear Predictive) 구조에 기반을 두고 있다. 16차 LP(Linear Prediction) 단구간 합성 필

터는 다음과 같이 사용한다.

$$H(z) = \frac{1}{\hat{A}(z)} = \frac{1}{1 - \sum_{i=1}^m \hat{a}_i z^{-i}}$$

$\hat{a}_i, i=1, \dots, m$: 양자화된 LP 계수,
 $m=16$: 예측 계수의 차수

장구간 또는 피치 합성필터는 다음과 같이 주어지는데, 적응 코드북 검색 방법을 사용하여 구현된다.

$$\frac{1}{B(z)} = \frac{1}{1 - g_p z^{-T}}$$

T 는 피치 지연시간,
 g_p 는 피치 이득

CELP 음성 합성모델은 단구간 LP 합성필터의 입력에서 여기신호가 적응 및 고정 코드북의 두 개의 여기벡터를 합함으로써 구성된다. 음성은 단구간 합성필터로부터 이러한 코드북들로부터 적절히 선택된 두 개의 벡터를 공급함으로써 합성된다. 하나의 코드북에서 최적의 여기 시퀀스는 원본음성과 합성된 음성 간의 오차가 인지가중적 감쇄추정에 관련하여 최소화되는 분석-합성 탐색과정을 사용하여 선택되어진다.

음성 부호화기는 12.8 kHz로 샘플링된 신호의 LPC, LTP, 그리고 고정 코드북 파라미터들을 분석하도록 설계되어 있고, 20 ms의 음성 프레임

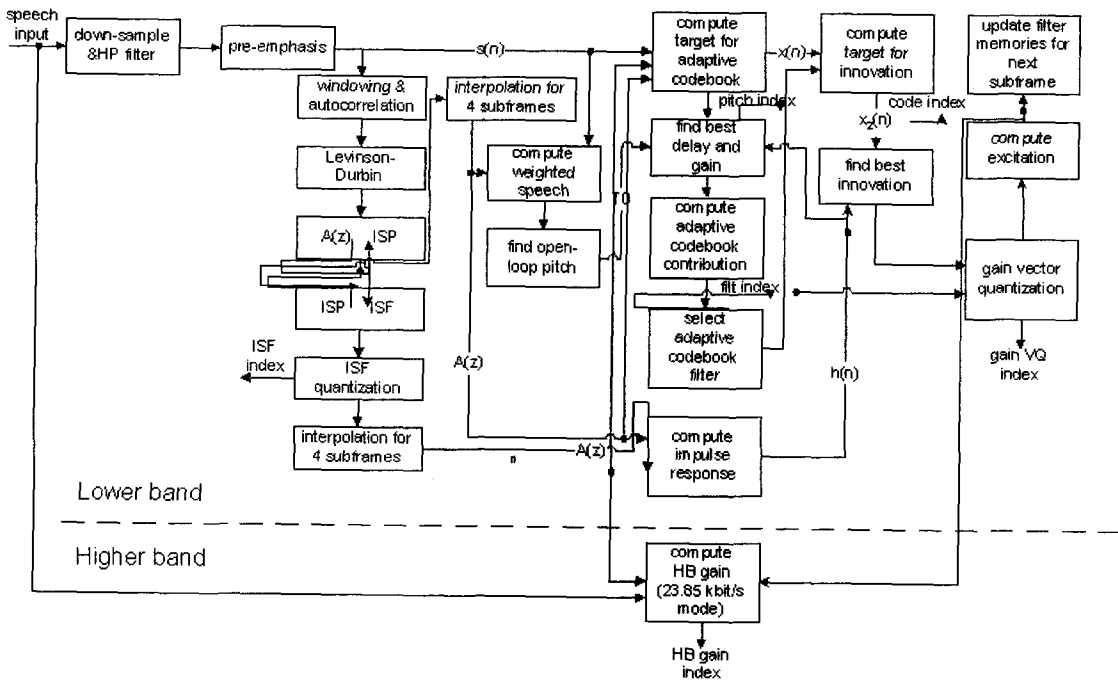
마다 동작한다. 각각의 프레임에서 음성신호는 CELP 모델의 파라미터들인 LP 필터 계수, 적응 및 고정 코드북의 인덱스와 이득을 추출하기 위하여 분석되어지고, 이러한 파라미터들에 추가로 고대역 이득 인덱스는 23.85 kbit/s 모드에서만 계산되어진다. 이러한 파라미터들은 부호화되어 전송되어지는데, 복호기에서 이러한 파라미터들은 다시 복원한다. 음성은 LP 합성 필터를 통과한 재구성된 여기 신호를 필터링함으로써 합성되어진다.

부호화기의 신호의 흐름은 <그림 5>와 같다. 샘플링 주파수를 12.8 kHz로 낮춘 후에 고대역 필터링과 전처리 강조 필터링을 수행하며, LP 분석은 프레임당 한번 수행된다. LP 파라미터들은 ISP(Immittance Spectrum Pairs)로 변환되고, S-MSVQ(Split-MultiStage Vector Quantization) 방식으로 벡터 양자화한다. 음성 프레임은 각각 5 ms(12.8 kHz 샘플링 비율의 64 샘플)의 4개의 부프레임으로 나눈다. 적응 및 고정 코드북 파라미터들은 매 부프레임마다 전송

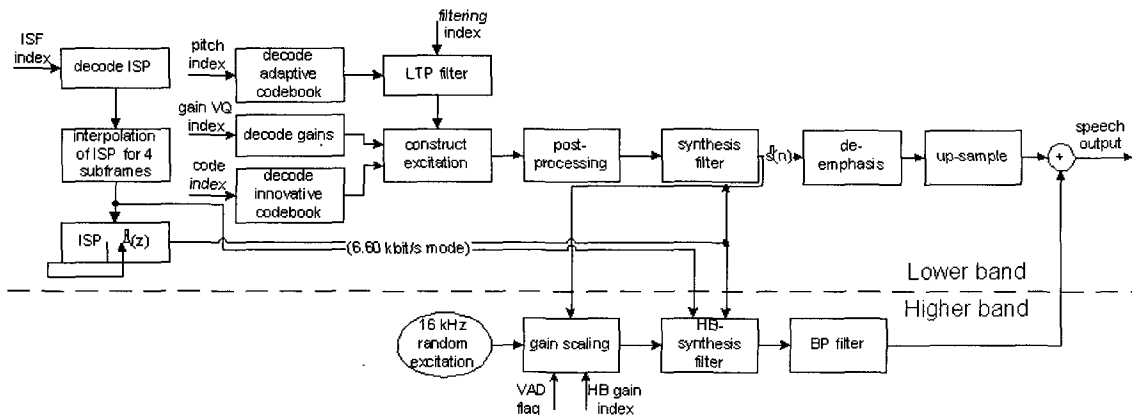
된다. 양자화되고 양자화되기 전의 LP 파라미터 또는 이것들의 서로 삽입된 변형들은 부프레임에 의존하여 사용되어진다.

다음 동작은 4개의 각 부프레임에서 반복 수행된다.

- 목표신호 $x(n)$ 은 LP 잔여신호와 여기신호 사이의 오차를 필터링하여 경신된 필터들의 초기 상태를 가진 가중된 합성 필터 $W(z)H(z)$ 를 통한 LP 잔여신호를 필터링함으로써 계산되어진다. 이것은 가중된 음성 신호로부터 가중된 합성필터의 Zero input 응답을 빼는 ZIR 방법과 같다.
- 가중된 합성필터의 임펄스 응답 $h(n)$ 을 계산한다.
- 페루프 피치합성은 목표신호 $x(n)$ 이 임펄스 응답 $h(n)$ 을 사용한 개방루프 피치 지연 주변을 찾음으로써 피치 지연과 이득을 찾기 위해 수행된다. 모드와 피치 지연값에 의존하는 샘플 분석의 1/4 또는 1/2의 분수 값의 피치값이 사용된다. 분수값의 피치 탐색과정에서 보간 필



<그림 5> AMR-WB(G.722.2)의 부호화기 구성도^[3]



〈그림 6〉 AMR-WB(G.722.2)의 복호화기의 구성도^[9]

터는 지역통과 주파수응답을 갖는다.

- 목표신호 $x(n)$ 은 여과된 적응 코드벡터의 적응 코드북 기여값을 제거함으로써 경신되고 새로운 목표신호 $x_2(n)$ 은 최적의 값을 찾기 위해 고정 대수 코드북 탐색에서 사용된다.
- 적응 및 고정 코드북의 이득은 고정 코드북 이득에서 적용되는 MA(Moving average) 예측을 갖는 6-7 비트로 벡터 양자화된다.

마지막으로 필터의 메모리는 결정된 여기신호를 사용하여 다음 부프레임에서 목표신호를 찾으므로써 경신된다.

2) AMR-WB의 복호화기

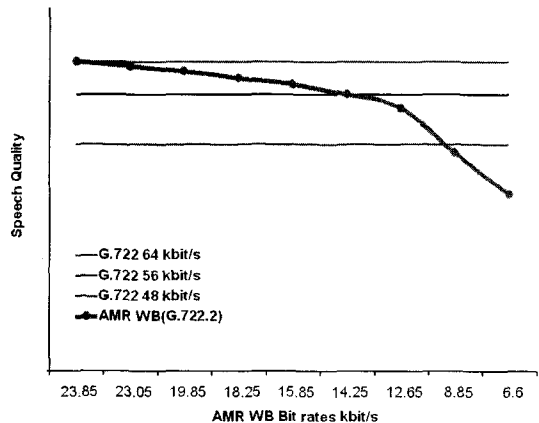
복호화기에서의 신호의 흐름도는 〈그림 6〉과 같다. 복호화기에서는 전송된 인덱스들은 수신된 비트스트림으로부터 추출된다. 이러한 인덱스들은 각각의 전송된 프레임에서 코더의 파라미터들을 얻기 위하여 복호화된다. 파라미터들에는 ISP 벡터, 4개의 분할된 피치 지연, 4개의 LTP 필터링된 파라미터, 4개의 혁신적인 코드벡터, 4개의 벡터 양자화된 피치와 혁신적 이득의 사이 있다. 23.85 kbit/s 모드에서는 또한 고대역 이득 인덱스가 복호화된다. ISP 벡터는 LP 필터 계수로 변환되며, 각 부프레임에서 LP 필터들을 얻기 위해 보관된다.

그 다음 각 64샘플의 부프레임에서 여기신호는

기대되는 이득들에 의하여 크기가 변한 적응 및 고정 코드벡터가 합해짐으로써 구해지고, 12.8 kHz의 음성은 LP 합성 필터를 통과한 여기신호를 필터링함으로써 재구성되며, 재구성된 음성은 역강조되어진다. 마지막으로 재구성된 음성은 12.8kHz에서 16kHz로 업샘플링되고, 16kHz로 샘플링된 6kHz로부터 7kHz까지의 주파수 대역을 갖는 고대역 음성이 더해진다.

4. ITU-T 광대역 음성부호화 알고리즘 비교

ITU-T 광대역 음성부호화 알고리즘들의 주요 특징 및 차이점은 〈표 1〉에 나타나 있다. 또한 AMR-WB(G.722.2)의 음질을 G.722와 비교한 결과는 〈그림 7〉에 나타나 있다. G.722.1



〈그림 7〉 AMR-WB와 G.722와의 음질 비교^[7]

〈표 1〉 ITU-T 광대역 음성부호화 알고리즘 비교^[6]

표준 방식	G.722	G.722.1	AMR-WB(G.722.2)
표준화 결정시기	1988	1999	2000, 2001
전송률(kbit/s)	48, 56, 64(embedded)	24, 32	23.85, 23.05, 19.85, 18.25, 15.85, 14.25, 12.65, 8.85, 6.6
부호화 방식	Sub-Band ADPCM	Transform Coding	Algebraic Code Excited Linear Prediction(ACELP)
전송지연(ms)	1.625	40	25
프레임 크기(ms)	0.125	20	20
Lookahead(ms)	1.5	20	5
음 질	우수한 음질 (64 kbit/s)	조건에 따른 음질 변화 심함 음악에서 좋은 음질 보임 핸드프리에 적합 저패킷 손실을 환경에 적합 24 kbit/s ≥ G.722@56 kbit/s 32 kbit/s ≥ G.722@64 kbit/s	12.65 kbit/s 이상에서 좋은 음질 음악에서 좋은 음질 보이지 않음 15.85 kbit/s ≥ G.722@56 kbit/s 23.05 kbit/s ≥ G.722@64 kbit/s
계산량	10 MIPS	<15 MIPS	38 WMOPS
RAM 크기	1 kwords	2 kwords	5.3 kwords
고정소수점 사양	Bit-exact	Bit-exact C	Bit-exact C
부동소수점 사양	준비안됨	Annex B	준비됨
VAD/DTX/CNG	준비안됨	준비안됨	가능함
주된 응용분야	ISDN+화상회의	VoPN+G.722와 동일	3G Wireless+G.722.1과 동일

방식은 패킷 손실이 심한 환경에서는 음질 저하가 비교적 심하게 나타나며 AMR-WB는 음악에서 좋은 음질을 보이지 않는 약점을 가지고 있다. AMR-WB는 비교적 많은 계산량을 요구하고 있으며 14.25 kbit/s 이상의 전송률에서 AMR-WB는 56 kbit/s G.722 방식과 동일한 음질을 보인다. 전송지연을 비교한다면 G.722.1 방식은 40 ms 알고리즘 지연을 보이며, AMR-WB는 25 ms 알고리즘 지연, G.722는 1.625 ms의 매우 적은 알고리즘 지연을 갖고 있다.

III. 광대역 음성부호화 알고리즘의 응용

광대역 음성 부호화 알고리즘의 응용 분야는

기존의 협대역 음성 부호화 알고리즘이 사용하고 모든 분야에서 사용 가능하며 더욱 자연스럽게 명료한 음질로 인해 디지털 방송 등에도 사용 가능하리라 사료된다. 또한 최근에 많은 관심을 보이고 있는 IP기반 패킷 네트워크에 적용된다면 기존의 인터넷 전화 보다 더욱 음질이 개선된 서비스가 가능하리라 본다. 다음은 광대역 음성부호화 알고리즘이 사용 가능한 분야를 간략히 설명한다. ^{[6][6]}

광대역 전화통신 :

AMR-WB는 기존의 PSTN 전화보다 음질이 향상되고, 자연스러움을 제공함.

오디오 비디오 원격회의 :

현재 G.722 방식을 사용함. AMR-WB 사용 시 G.722에 비교하여 1/2 이상의 전송률 감소 가능함.

3세대 이동 통신시스템 :

높은 음질의 오디오와 음성 전송이 가능한 멀티미디어 서비스는 3세대 이동통신의 주요 목적 중에 하나임. 3세대 이동통신의 음성 전화통신 응용에서 전통적인 PSTN망보다 뛰어난 음질을 전송하기 위한 무선 서비스 가능함. 최근 AMR-WB 3GPP의 표준 방식으로 채택함.

광대역 Packet 네트워크와 ISDN에서의 고품질 전화

DSL, Packet Cable, ATM, Frame Relay, 광대역 패킷 네트워크에서 실시간 IP-based 프로토콜과 결합하여 고품질 음성 전화 가능함.

패킷 네트워크에 대한 광대역 전화 통신

협대역 음성 VoIP 시스템을 고품질 음성 통신으로 변경 가능함. QoS에 영향을 미치는 지연시간, 지터, 패킷손실, 음성압축에 의한 음질저하 등의 개선이 요구됨.

인터넷 응용

인터넷 방송, 음성 채팅, 멀티미디어 실시간 강의, 언어 교육시스템 등 인터넷 활용에 사용 가능함.

디지털 전파 방송

광대역 음성은 디지털 AM 라디오 방송과 뉴스에 대한 오디오, 비디오 방송에 사용되며, TV 프로그램 및 폐쇄회로 강좌에 사용될 수 있음.

IV. 결 론

차세대 이동통신 시스템에서 제공되어야 할 서비스의 종류와 질은 기존의 고정 통신망과 구별을 두지 않고 제공되어야 하며 궁극적으로는 초고속 정보망과도 연결되어 동일한 통신 서비스를 제공하여야 한다. 예를 들자면 이동통신 시스템에서도 유선통신 시스템과 같이 무선 화상 전화 서비스를 비롯하여 다양한 멀티미디어 서비스가 제공되어야 한다. 이러한 필요성에 의해 음성

대역폭을 50 Hz-7000 Hz로 확대시킨 광대역 음성부호화 알고리즘에 대한 관심이 증가되고 있다. 본 논문에서는 ITU-T 광대역 음성부호화 방식인 G.722, G.722.1, G.722.2(AMR-WB) 부호화 방식을 간략하게 소개하고 각 방식의 특징을 비교하였다. 특히 가장 최근에 표준화된 AMR-WB 방식은 16 kbps 전송률에서 좋은 음질을 갖고 있어 차세대 이동통신 및 인터넷에 응용 가능하다. AMR-WB 방식은 협대역 음성부호화 방식과 같은 ACELP 구조를 갖고 있으며 전송 채널의 상태에 적응적으로 전송률을 변할 수 있는 구조로 구성되어 유무선 채널에 사용 가능하며 그 사용범위가 크게 증가되리라 예측되어진다. 이밖에 광대역 음성부호화기는 화상회의 시스템, 고품질 디지털 AM 라디오 방송, 고품질 전화 등에 사용 가능하며 응용분야는 더욱 확대되리라 예상되어진다.

참 고 문 헌

- [1] ITU-T Recommendation G.722, "7 kHz Audio-Coding within 64 kbit/s," 1988.
- [2] ITU-T Recommendation G.722.1, "Coding at 24 and 32 kbit/s for Hands-free Operation in Systems with Low Frame Loss," 1999.
- [3] ITU-T Recommendation G.722.2, "Wideband Coding of Speech at around 16 kbit/s Using Adaptive Multi-rate Wideband(AMR-WB)," 2002.
- [4] 3GPP TR 26.976 "AMR-WB Speech Codec Performance Characterization," 3GPP Technical Specification.
- [5] Lajos Hanzo, F. Clare A. Somerville, Jason P. Woodard, "Voice Compression and Communications," IEEE Press Series on Digital & Mobile Communication, pp.341-384, 2001.
- [6] VoiceAge™, "Wideband Speech Cod-

ing Standards and Applications,” September. 2001.

- [7] Tero Piirainen, “Details in Voice Coding : Advanced Speech Coding for VoIP,” December. 2000.

저자 소개



李寅誠

1983년 2월 연세대학교 전자공학과 학사, 1985년 2월 연세대학교 전자공학과 석사, 1992년 12월 Texas A&M University 전기공학과 박사, 1986년 5월~1987년 7월 : 한국통신 연구개발단 전임연구원, 1993년 2월~1995년 9월 : 한국전자통신연구원 이동통신기술연구단 선임연구원, 1995년~현재 : 충북대 전기전자공학부 부교수, <주관심 분야 : 음성 및 영상신호 압축, 이동통신, 적응필터>



申載賢

1997년 2월 충북대학교 전파공학과 학사, 2001년 3월~현재 : 충북대학교 전파공학과 석사과정, <주관심 분야 : 음성코딩, LSP 양자화, 적응필터>