

디지털 이동통신을 위한 음성 부호기의 성능 분석

(A Performance Analysis of the Speech Coders for Digital Mobile Radio)

鄭 榮 謨*, 李 商 郁*

(Young Mo Chung and Sang Uk Lee)

要 約

최근 디지털 이동통신을 위하여 세계 각국에서 SBC-APCM, RPE-LPC, MPE-LTP, CELP 등의 음성 부호기가 제안되었다. 그러나 현재까지 이러한 여러 종류의 이동통신용의 음성 부호기들이 제안되고 있지만 그 성능에 대한 총체적인 비교, 평가 분석이 이루어지지 않고 있다. 따라서 본 논문에서는 디지털 이동통신에서의 음성 부호화 기법으로 유망한 SBC-APCM, RPE-LPC, MPE-LTP 그리고 CELP등에 대하여 랜덤 에러 및 Rayleigh 페이딩 채널 상에서의 burst 에러에 대하여 성능을 평가 분석하였다. 이들 부호기중 SBC-APCM, RPE-LPC, MPE-LTP에 대하여는 비트 sensitivity를 측정하여 이를 토대로 비트 선택성 에러교정 부호를 적용하였다. 그리고 CELP에서는 음질 개선의 한 방법으로 최적 양자기를 적용해본 결과 그 개선치는 별로 크지 않음을 확인할 수 있었다. 또한 CELP에서 채널 에러에 대한 강인성을 향상하기 위하여 Reed-Solomon 부호, Golay 부호, 길쌈 부호, 소거식 길쌈 부호등을 적용해본 결과 부호율 1/2의 길쌈 부호가 가장 좋다는 결론을 얻었다. 이상으로 부터 각 부호기의 성능 평가 결과 CELP, MPE-LTP, SBC-APCM 그리고 RPE-LPC등의 순으로 성능이 좋다는 결론을 얻을 수 있었다.

Abstract

Recently, four speech coding techniques, namely, SBC-APCM (sub-band coding adaptive PCM), RPE-LPC (regular pulse excitation linear predictive codec), MPE-LTP (multi-pulse excited long-term prediction) and CELP (code-excited linear prediction) are proposed for digital mobile radio applications. However, a performance comparison of these coders in the Rayleigh fading environment has not been made yet. In this paper, the performances of the four speech coders in the random bit error and burst error environment are investigated. For the channel coding of SBC-APCM, RPE-LPC and MPE-LTP, the sensitivity of output bit stream is measured and a bit selective forward error correction is provided according to the measured bit sensitivity. And for an attempt to improve the performance of CELP, an optimum quantizer is applied for transmitting scalar quantities in CELP. However, an improvement over the conventional approach is found to be negligible. For the channel coding of CELP, Reed-Solomon code, Golay code, convolutional code and punctured convolutional code are employed for comparison purpose. Convolutional code of rate 1/2 shows the best performance. Finally, from the simulation results, it is concluded that CELP is the best candidate for digital mobile radio and is followed by MPE-LTP, SBC-APCM and RPE-LPC.

*正會員, 서울대학교 制御計測工學科
(Dept. of Control and Instrumentation Eng., Seoul
Nat'l Univ.)

接受日字: 1989年 12月 23日

I. 서 론

현재의 이동통신 시스템에 음성뿐 아니라 데이터도 이동통신 가입자에 제공되도록 하여 궁극적으로

는 이동체에 ISDN 접속이 가능하게 되도록 현재 많은 연구가 진행중이다.^[1] 이동통신을 디지털화하기 위해서는 여러 문제들이 해결되어야 하는데, 그중 가장 중요한 것이 아날로그의 음성 신호를 디지털 신호로 변환하는 음성 부호화 기법의 선택이다. 이동통신에 사용되는 음성 부호화 기법은 사용되는 환경의 특성으로 인하여 다음과 같은 조건을 만족해야 한다. 첫째, 제한된 채널 대역폭으로 인하여 전송율이 16 kbps 이하가 되어야하고, 둘째 채널 에러에 강한 강인성을 가져야한다. 특히 이동통신 채널의 특성상 전송 에러는 burst 에러의 형태로 나타나므로 burst 에러에 강한 강인성을 가져야한다. 이러한 조건에서 toll quality의 음질을 달성하기 위하여 여러 음성 부호기들이 제안되었다. 이들중 가장 유망한 것들로서 SBC-APCM(sub-band coding adaptive PCM)^[2], RPE-LPC(regular pulse excitation linear predictive codec)^[3], MPE-LTP(multi-pulse excited long-term prediction)^[4], CELP(code-excited linear prediction)^[5] 등이 있다. SBC-APCM, RPE-LPC, MPE-LTP 등은 범유럽 공동의 디지털 이동통신망을 위하여 제안된 음성 부호기들이고 CELP는 최근 미국방성에서 제안된 음성 부호기이다.^[6] 이들 부호기들은 전송율 16 kbps 이하에서 지금까지 많이 사용되어온 CVSD(continuously variable slope delta modulation)와 비교하여 훌륭한 음질을 보이지만 전송 에러에 민감한 것이 단점이다. 따라서 최근에는 채널의 burst 에러에 대한 강인도를 증가시키기 위하여 에러 정정 부호를 음성 부호기에 적용시키는 것이 일반적인 추세이다. 그러나 지금까지 디지털 이동통신을 위한 여러 종류의 음성 부호기들이 제안되고 있지만 그 성능에 대한 총체적인 비교, 평가 분석이 이루어지지 않고 있다. [1]에서는 SBC-APCM, RPE-LPC 및 MPE-LTP에 대하여 랜덤 비트 에러 환경하에서 성능의 비교 분석이 이루어 졌으나 이동통신 채널 상에서의 성능 분석은 없었다. 또한 [6]에서는 CELP에 대한 성능 분석이 있었으나 다른 음성 부호기와 비교 분석이 없었다.

본 논문에서는 이들 네개의 음성 부호화 기법에 대하여 이동통신 채널에서 발생하는 burst 에러에 대한 성능을 컴퓨터 시뮬레이션을 통하여 비교 분석하여 이들 네 부호기에 대하여 객관적 평가자료를 제시하는 것을 목적으로 한다. 시뮬레이션 결과 CELP는 다른 세 음성 부호기에 비하여 전송율은 거의 반 정도이지만 이들과 거의 동등하거나 우수한 음질을 보임을 확인할 수 있었다. 본 논문에서는 전송율 7.2 kbps에서 CELP의 음질을 향상시키기 위한 목적으

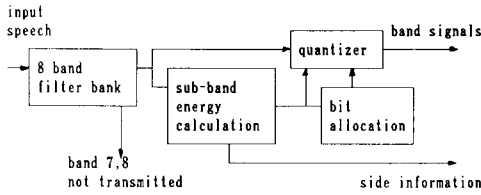
로 사용한 최적 양자기에 대한 실험 결과도 아울러 제시한다. 그리고 CELP에서는 채널 부호화를 할 수 있는 전송율의 여유가 많으므로 채널 에러에 강한 Golay 부호와 길쌈 부호등의 여러 에러 정정부호의 적용 결과도 제시한다. 본 논문에서 burst 에러는 소프트웨어 Rayleigh 페이딩 채널 시뮬레이터^[7]로 얻었으며, 최종적인 음질 평가는 단시간 SNR과 주관적인 청각 조사를 통하여 수행하였다.

본 논문의 구성은 다음과 같다. 제2장에서는 본 논문에서 시뮬레이션된 음성 부호기에 대한 고찰 및 CELP에서의 양자화에 대한 고찰을 한다. 제3장에서는 이들 부호기로 부터 측정된 음성 부호기의 비트 sensitivity 측정결과를 제시하며 이를 토대로 비트 선택적 에러 정정부호의 적용에 관하여 논한다. 아울러 제3장에서는 CELP에 적용된 여러 에러 정정부호에 관한 고찰을 한다. 다음 제4장에서는 시뮬레이션 결과를 제시하고, 마지막 제5장에서는 본 논문의 결론을 제시하였다.

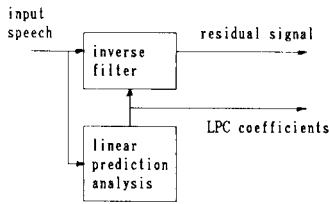
II. 각 음성 부호화 기법에 대한 고찰

1. SBC-APCM^[2]

일반적으로 음성 신호는 WSS(wide-sense stationary) 랜덤 프로세스라고 가정하지만 실질적으로는 음성음, 무성음에 따라 스펙트럼의 차이가 크다. 따라서 SBC에서 고정 비트 할당을 하는 것보다 동적 비트 할당을 하여 양자화 오차를 줄이는 것이 일반화되어 있다. 본 논문에서 사용한 SBC는 대역신호의 에너지 분포에 따라 동적 비트 할당을 하고 아울러 양자기도 신호의 크기에 따라 적응화를 하는 그림 1(a)와 같은 구조를 가진다. 여기에서 입력신호는 필터뱅크에 의해 각 0.5kHz의 대역폭을 가지는 8개의 대역신호로 분리된다. 각각의 필터뱅크는 32, 24, 16차의 QMF를 나무구조로 구성한 것으로 입력신호는 최초에 32차의 QMF에 의해 2개의 대역신호로 나뉘고 각 대역신호는 다시 24차의 QMF를 거친후 최종적으로 16차의 QMF에 의해 8개의 대역으로 나누어지게 된다. 3-4kHz 사이의 신호인 대역 7과 8은 음성통신의 대역인 3kHz를 벗어나므로 이 대역의 신호는 전송하지 않는다. 각각의 대역신호는 각 대역의 에너지 분포에 따라 비트할당이 되고 각 대역의 에너지 크기는 양자기를 순방향 적응(forward adaptation) 시키는데 이용된다. 따라서 각 대역의 에너지 정보를 수신측에 부정보(side information)의 형태로 보내주어야 한다. [8]에서는 부정보를 벡터 양자화를 하여 전송하나 계산량과 벡터 양자화에 필요한 구조의 복잡성을 고려할 때 [2]에서와 같이



(a) SBC-APCM



(b) Pulse-excited coders.

그림 1. 음성 부호기의 블록선도
Fig. 1. Blockdiagram for speech coders.

log PCM으로 양자화를 하는 것이 구조적인 측면에서 잇점이 크다. 실제로 부정정보의 전송율은 벡터 양자기에서 1kbps 정도 소요가 되고 log PCM에서 1.875 kbps 정도로 그리 큰 차이가 없다. 동적 비트할당에 의해 각 대역신호는 1-4 비트의 Max 양자기로 양자화되어 12 kbps의 전송율로 수신측에 전송된다. 본 논문에서는 부정정보를 log PCM 방식을 택하여 양자화하였고 이때의 비트 할당을 표 1에 제시하였다.

2. RPE-LPC^[3] MPE-LTP^[4]

RPE-LPC, MPE-LTP 그리고 CELP 등은 선형 예측 해석과 선형 예측 계수로 역필터링을 한다는 점에 있어 그림 1(b)와 같은 공통적인 구조를 가진다. 이러한 인간의 발성기관을 모델링한 음성 부호기에서 수신측에 전달되어야하는 정보는 그림 1(b)에서와 같이 필터의 계수와 잔여 신호인데, 이 중 잔여 신호의 전송 부담이 크다. 따라서 적은 전송율로 음질의 왜곡을 줄이는 잔여 신호의 전송법이 여러가지로 제안되었다. RPE-LPC는 잔여 신호를 N:1로 간축(decimate)하여 잔여신호의 1/N만을 전송하고, MPE-LTP에서는 음질의 왜곡을 최소화하는 몇개의 최적 펄스를 찾아 이들만을 전송하는 구조를 갖는다. 특히 MPE-LTP에서는 LTP(long-term predictor)가 있어 잔여신호내에서 핏치와 핏치 사이의 상관성을 제거하여 잔여신호내의 정보량을 좀 더 줄일 수 있게 하고, 아울러 합성음에서는 핏치와 핏치간의 상관성을 복원하여 유성음을 충실하게 재현하는 기능을 한다. 본 논문에서 시뮬레이션된 RPE-LPC는 범유럽 공동의 이동통신 시스템을 위하여 제안된 [3]의 구조를 그대로 사용하였다. 한편 [4]에서 제안된 MPE-LTP는 최적의 펄스를 [9]에서 제안한 준최적의 방법으로 찾는 구조를 가지고 있다. 이러한 방법에서는 필요한 갯수의 최적 펄스보다 적은 수의 펄스를 찾는 경우가 발생하는데 이때에는 부족한 펄스를 임의의 위치에 최소의 크기로 지정하는 방법을 사용한다. 그러나 최적 펄스의 양자화에는 8 레벨의 양자기가 사용되므로 전송한 경우 최소의 양자기 레벨이

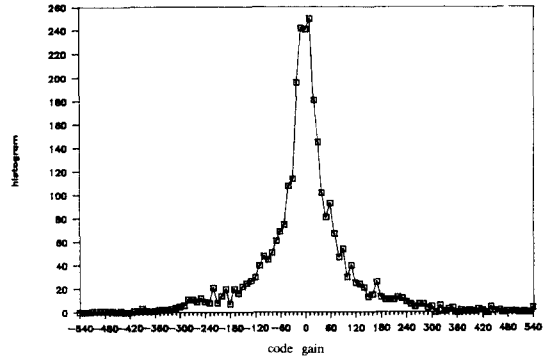
표 1. 비트 할당
Table 1. Bit allocation.

SBC-APCM		RPE-LPC		MPE-LTP		CELP	
frame length	16ms	frame length	19.5ms	frame length	20ms	frame length	20ms
6 energies	30	12 LAR's	52	8 PARCOR's	28	10 LSP's	36
band signals	192	52 RPE samples	208	4 LTP delays	28	4 LTP delays	28
		4 block maxima	20	4 LTP gains	8	4 LTP gains	12
		4 RPE grid positions	8	4 pulse positions	88	4 indices	40
				4 특성치	20	4 gains	28
				24 pulse samples	72		
				4 펄스 수	12		
sum	222	sum	288	sum	256	sum	144
bit rate	13.88kbps	bit rate	14.77kbps	bit rate	12.8kbps	bit rate	7.2 kbps
FEC	24	FEC	24	FEC	64	FEC	144
gross		gross		gross		gross	
bit rate	15.38kbps	bit rate	16.00kbps	bit rate	16.00kbps	bit rate	14.40kbps

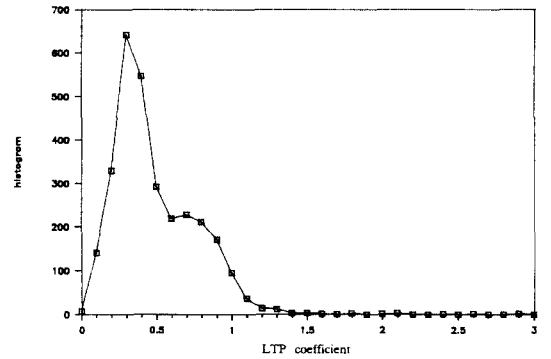
지정되더라도 양자화 오차는 1/8로서 상당히 크다. 따라서 최소 크기의 펄스를 임의의 위치에 지정하는 것보다, 펄스의 갯수를 지정해줌으로써 임의의 위치에서 펄스가 발생하는 것을 방지해 주는 것이 음질을 개선하는 한 방법이 된다. 실제 시뮬레이션 결과 이와같이 임의의 위치에 펄스가 발생되게 하는 것보다 최적 펄스의 갯수를 전송하는 방법이 1 dB 정도의 단시간 SNR의 이득이 있음을 확인할 수 있었다. 이때 추가적인 펄스 갯수를 나타내는 전송 비트의 부담은 펄스의 위치에 할당된 비트 수를 줄임으로서 해결하였다. 표 1에 RPE-LPC와 MPE-LTP의 비트 할당을 제시하였는데 MPE-LTP의 전송율은 [4]의 제안 전송율 보다 낮은 12.8 kbps이다.

3. CELP^[5]

CELP에서는 선형 예측 해석과 LTP를 가지는데 있어서 MPE-LTP와 동일한 구조를 가지나, 잔여 신호를 벡터 양자화하는 차이점이 있다. CELP에서의 벡터 양자화는 주어진 codebook내에서 청각의 감도에 따른 주관적인 에러 기준에 의해 이를 최소화하는 최적의 code를 찾아, 이 code의 지수(index)와 이때의 최적 이득(gain)을 찾는 것이 핵심 구조로 되어 있다. 본 논문에서 벡터의 크기는 40, codebook의 크기는 1024로 하였으며, codebook은 평균값이 0이고 분산이 1인 백색 가우시안 랜덤 숫자를 사용하여 구성하였다. 한편 CELP에서는 LPC 계수들을 LSP(line spectrum pairs)^[10]로 변화하여 양자화 한다. LSP는 기존의 반사 계수나 LAR(log area ratio)보다 분포영역이 좁아 양자화 오차를 줄일 수 있는 잇점이 있다. 표 1에 7.2 kbps CELP의 비트할당을 제시하였다. 이러한 구조의 CELP에서 양자화에 대한 고찰을하면 다음과 같다. CELP내에서 양자화를 하여야 하는 정보로는 LSP 주파수, code의 이득 그리고 LTP의 계수등이 있다. 시뮬레이션된 음성으로부터 추출된 약 3000개의 code의 이득과 LTP의 계수의 분포도를 그림 2에 제시하였다. 이러한 입력신호의 분포에 따라 양자화 이전의 신호와 양자화 이후의 신호에 대하여 제곱 평균 오차값을 최소화하는 기준에 따라 [11]에서 제안한 방법으로 최적 양자기를 구할 수 있다. 이러한 방법으로 구성한 최적양자기를 CELP에 적용시킨 결과를 표 2에 제시하였다. 표 2에서 보면 LSP 주파수, code의 이득과 LTP의 계수를 양자화하지 않았을때와 최적 양자기들을 사용하여 양자화하였을때의 차이는 약 0.36 dB 정도 나타나다. 한편 이러한 복잡한 구조의 최적 양자기 대신에 LSP 주파수의 양자화에는 [12]에서 제안한



(a) Code gain.



(b) LTP coefficient

그림 2. Code의 이득과 LTP 계수의 분포도
Fig. 2. Histograms of code gain and LTP coefficient.

표 2. 양자기에 따른 단시간 SNR
Table 2. Segmental SNR result with various quantizers.

no quantization	optimum quantization	proposed method
12.56 dB	12.20 dB	11.90 dB

방법과 같이 구조가 간단한 균일 양자기가 많이 사용되기도 한다. LSP 주파수들은 각기 분포의 범위가 다르지만 인접한 LSP 주파수의 차는 0에서 0.1 사이에 분포하는 특징을 가지므로 인접한 LSP 주파수의 차신호를 동일한 양자기로 양자화하는 DPCM 구조의 부호화를 사용한다. 그리고 양자기의 구조를 좀 더 간단히하기 위하여 code 이득의 양자화에도 구조가 간단한 log PCM 방식의 양자화를 적용시켰을때의 결과를 표 2의 세번째 항에 제시하였다. 이러

한 방식의 양자화와 최적 양자기와의 차이는 약 0.3 dB 정도로 아주 작음을 알 수 있다. 그리고 실제 청각 실험의 결과로도 음질의 차이는 거의 없었다. 그러므로 CELP에서 양자화로 인한 음질의 저하는 그리 크지 않다고 판단된다. 이상으로 부터 구조의 복잡성과 성능을 고려할 때 다음과 같은 양자기를 적용하는 것이 적당하다고 결론지을 수 있다. Code의 이득은 log PCM 방식을 사용하는 것이 바람직하고 LTP 계수의 양자기는 표 3에서 제시한 8레벨 최적 양자기를 적용하고 LSP 주파수의 양자화에는 4, 4, 4, 4, 4, 3, 3, 3, 3 비트가 할당된 DPCM이 적합하다.

표 3. LTP 계수의 최적 양자기
Table 3. Optimum quantizer for LTP coefficient.

index	output level	decision level
0	0.15402	0.216130
1	0.27824	0.330350
2	0.38246	0.448285
3	0.51411	0.604375
4	0.69464	0.796670
5	0.89870	1.077450
6	1.25620	2.031150
7	2.80610	

Ⅲ. 비트 sensitivity 및 채널 부호화

1. 비트 sensitivity 측정결과

앞 장에서의 음성 부호화 기법은 그림 3과 같은 프레임 구조를 갖도록 하였다. 이러한 프레임 구조 내에서 각각의 비트가 최종음질에 미치는 중요도를 측정하기 위하여 매 프레임마다 동일한 위치의 비트에 에러를 발생시켜 복원한 음성의 단시간 SNR을 측정한 것이 그림 4의 비트 sensitivity 측정 결과이다. 그림 4로부터 SBC-APCM에서는 부정정보량들과 저대역 신호들의 상위 비트들이 중요함을 알 수 있다. 부정정보 가운데서도 6개 대역의 에너지를 나타내는 비트들 중 상위 비트들의 비트 sensitivity가 높아 부정정보가 차지하는 0-30비트 사이에서 6개의 국부 최소점이 생김을 고찰할 수 있다. 그리고 RPE-LPC, MPE-LTP, CELP에서는 LAR(log area ratio), PARCOR (partial correlation) 계수, LSP 등의 선형예측 필터의 계수를 나타내는 비트들과 소블록내의 잔여 신호를 적응 양자화하는데 필요한 블록최대치가 매우 중요한 비트들임을 알 수 있다. 특히 블록 최대치 내에서는 부호(sign) 비트 보다 MSB(most significant bit)가 훨씬 중요하게 나타난다. 따라서 이러한 비트들의 중요도에 따라 비트들을 선택적으로 에러 교정 부호로 보호하는 비트 선택적 채널 부호화에 대하여 다음에서 논한다.

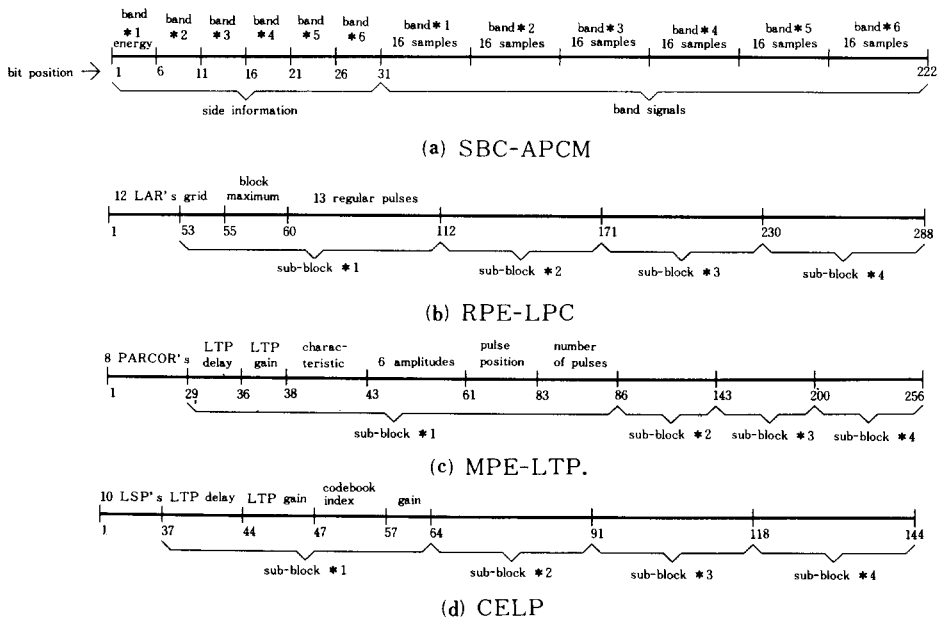


그림 3. 출력 프레임의 구조
Fig. 3. Structures of the output bit stream.

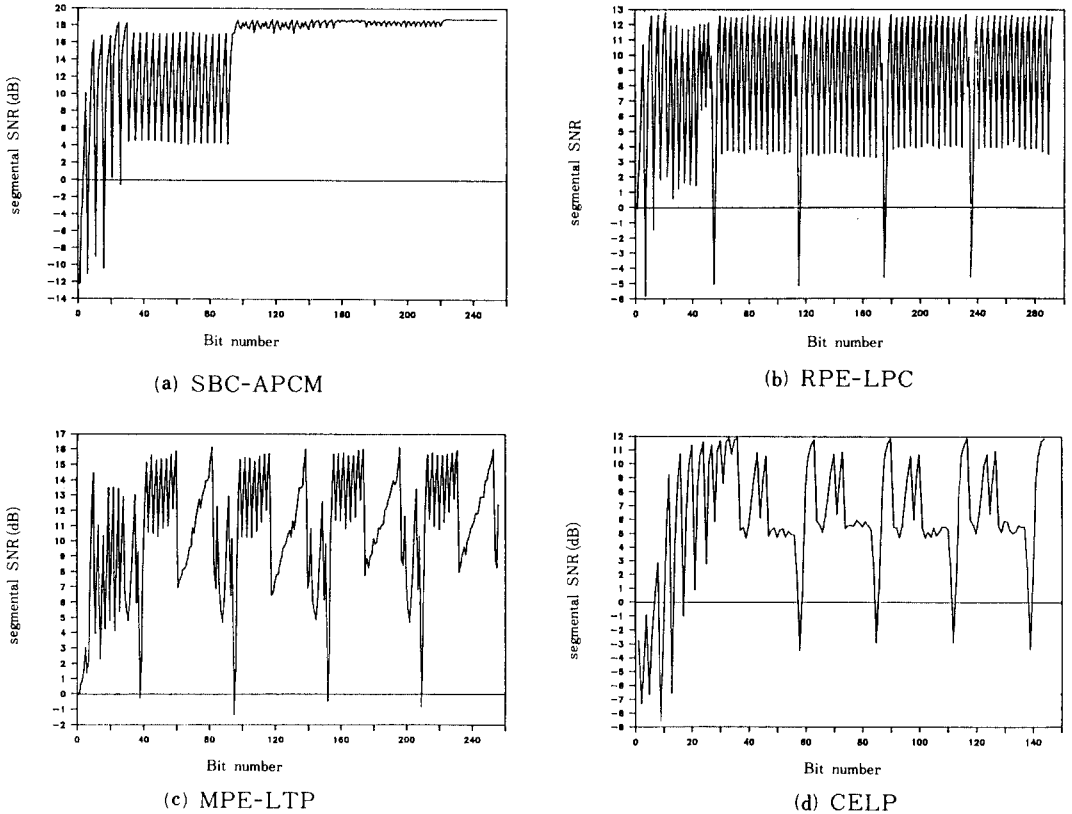


그림 4. 비트 sensitivity 측정결과
 Fig. 4. Bit sensitivity measure results.

2. 비트 선택적 채널 부호화

CELP를 제외한 세계의 음성 부호기들은 음성 신호의 전송율이 13-15 kbps 정도이므로 채널 부호화를 할 수 있는 여유는 1-3 kbps 정도이다. 따라서 비트 sensitivity 측정결과에 따라 상대적으로 중요한 비트들에 대하여 에러 교정 부호를 적용시켜 보호한다. 이동통신 채널에서의 비트 에러율은 최대 10^{-2} 정도로 예상을 하므로 에러 교정 부호의 적용목표는 비트 에러율 10^{-3} 에서는 음질이 에러가 없는 경우와 동일하도록 하고 비트 에러율 10^{-2} 에서는 음질의 왜곡이 나타나더라도 충분히 통화 내용을 이해할 수 있도록 하였다. 이러한 목적으로 블록 부호의 일종인 Reed-Solomon 부호와 확장(extended) Hamming 부호^[13]를 선택하여 각각의 음성 부호기에 적용하였다. SBC-APCM과 RPE-LPC에는 비트 선택적으로 Reed-Solomon 부호를 적용하였다. Reed-Solomon 부호는 디코딩 알고리즘이 복잡하다는 단점이 있지만 burst 에러 교정 능력이 있다는 장점이 있다.

여기에서는 $GE(2^4)$ 에서 정의되는 systematic (15, 9) Reed-Solomon 부호를 사용하였는데 이 부호는 3 심볼 에러 교정이 가능하다. MPE-LTP에는 1개의 비트 에러 교정과 2개의 비트 에러 검출이 가능한 systematic (16, 11) 확장 Hamming 부호와 (32, 26) 확장 Hamming 부호가 적용되었는데, 비트 sensitivity에 따라 8개의 (16, 11) 부호와 4개의 (32, 26) 부호가 사용되었다. Hamming 부호는 에러 교정 능력은 떨어지지만 디코딩이 간편하게 되므로 하드웨어의 실현이 용이하다는 장점이 있다. 최종적으로 채널 부호화를 하고난 출력 비트들에는 interleaving 기법을 적용하여 burst 에러에 대한 강인성을 높였다. 표1에 각 부호기에서 사용된 비트 선택적 에러 교정 부호의 비트 할당을 제시하였다.

3. CELP에서의 채널 부호화

전송율 7.2 kbps의 CELP에서는 채널 부호화를 적용할 수 있는 전송율의 여유가 많으므로 본 절에서는 Reed-Solomon 부호화 이것 이외에 강력한 에

러 정정 능력을 갖고 있는 Golay 부호화 길쌈 부호 (convolutional code)를 적용하여 이동 통신 채널에서 CELP의 성능을 어느정도 개선하는지를 비교 평가 하는데 그 목적을 두었다. (23, 12) Golay 부호는 블록 내에서 세계의 에러까지 정정이 가능한 부호로 여러개의 에러 정정이 가능한 유일한 완전 (perfect) 부호이다. (23, 12) Golay 부호는 $g(x) = 1 + x^2 + x^4 + x^6 + x^8 + x^{10} + x^{11}$ 과 같은 생성다항식으로 부호화되고 복호화는 Kasami 복호기^[13]로 용이하게 수행할 수 있다. (23, 12) Golay 부호를 적용하였을때 CELP의 전체 전송율은 13.8 kbps가 된다. 다음 길쌈 부호로는 부호율이 1/2이고 constraint length가 3인 것과 이의 출력 부호를 주기적으로 소거시켜서 구성한 부호율 2/3의 소거식 길쌈 부호 (punctured convolutional code)를 사용하였다. 부호율 1/2인 길쌈 부호의 생성다항식은 $g_1(x) = 1 + x^2$ 과 $g_2(x) = 1 + x + x^2$ 이고 이때의 최소 자유거리가 5가 된다. 이로부터 부호율 2/3인 소거식 길쌈 부호는 부호율 1/2인 길쌈 부호의 출력 부호 중 주기적으로 4개 마다 하나씩을 소거하여 얻게 되는데 이때의 최소 자유거리가 3이 된다. 소거식 길쌈 부호는 고부호율의 길쌈 부호를 필요로 할 때 사용되는데, 복호화 방법은 길쌈 부호에서와 동일하게 이루어 진다. 부호율 1/2인 길쌈 부호를 CELP에 적용하였을 때의 전송율은 14.4 kbps가 되고 부호율 2/3의 소거식 길쌈 부호를 적용하였을 때는 전송율 10.8 kbps가 된다. 표 4 에 여러 채널 부호기를 적용 하였을때 CELP의 전송율을 제시하였다.

표 4. 여러 채널 부호기에 대한 CELP의 전송율
Table 4. Bit rates for various channel coders.

RS	Golay	R=1/2	R=2/3
12.0 kbps	13.8 kbps	14.4 kbps	10.8 kbps

IV. 시뮬레이션결과 및 검토

앞에서 설명한 음성 부호기들은 랜덤 에러와 burst 에러에 대하여 에러교정 부호를 적용한 경우와 그렇지 않은 경우로 나뉘어 시뮬레이션되었다. 사용된 음성 데이터는 3-5초의 남자 음성 2개와 여자 음성 2 개로서 3.4 kHz로 대역 제한된 다음 8 kHz의 표본 화에 14 비트로 양자화된 것이다. 그림5에 Rayleigh 페이딩 채널에서의 시뮬레이션 블록 선도를 제시하였다. 음성 신호가 입력되면 음성 부호화를 한 다음 에러 교정 부호를 적용한다. Rayleigh 페이딩 채널에서는 에러가 burst 형태로 발생하므로 burst 에러에

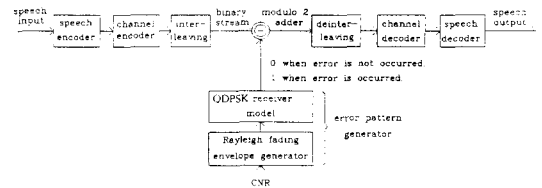


그림 5. 시뮬레이션 블록선도
Fig. 5. Simulation blockdiagram.

대한 강인성을 높이기 위하여 interleaving을 한다. Interleaving된 이진 출력열은 Rayleigh 페이딩 채널 시뮬레이터에 에러가 발생되는데, 이와 같이 에러가 발생한 이진 데이터 열은 수신측으로 전송되어 deinterleaver와 채널 복호기, 음성 복호기를 통하여 음성 신호로 복원된다. Rayleigh 페이딩 채널의 시뮬레이션은 [7]에서 제시한 방법과 같이 행하였으며 이때의 이동체의 속도는 55 mph, 반송파의 주파수는 450 MHz, 그리고 변복조 기법으로는 QDPSK (quadrature differential phase shift keying)로 하였다. 랜덤 비트 에러 시뮬레이션은 그림5에서 burst 에러 패턴 발생기 대신에 랜덤 에러 발생 함수를 사용하여 행하였다. 본 장에서는 먼저 이동 통신 채널상에서 CELP의 강인성을 높이기 위하여 적용된 여러채널 부호기에 대한 성능비교 결과를 제시한다. 이 결과로부터 가장 좋은 성능을 보이는 부호기를 선택한 다음 최종적으로 기존의 SBC-APCM, RPE-LPC MPE-LTP와 성능비교한 결과를 제시한다.

1. CELP에서의 채널 부호기 성능비교

랜덤 비트 에러는 비트 에러율 $10^{-3}, 5 \times 10^{-3}, 10^{-2}, 5 \times 10^{-2}, 10^{-1}$ 에 대하여 시뮬레이션 되었는데 각 비트 에러율에 대하여 최종적으로 복원된 음성을 원음과 비교하여 측정된 단시간 SNR의 결과를 그림6에 제시하였다. 그림6으로 부터 비트 에러율 10^{-2} 까지는 음질의 저하가 완만하게 일어나지만 그 이상에서는 음질의 저하가 급격하게 일어남을 관찰할 수 있다. 특히 비트 에러율 10^{-1} 에서 각 채널 부호기들은 모두 0 dB 보다 낮은 SNR을 보이므로 이 경우는 비교의 대상에서 제외하였다. 랜덤 비트 에러 발생시의 결과로는 10^{-2} 에서 최저 0 dB에서 최고 0.6 dB의 아주 작은 성능 저하를 보이므로 각 부호기들은 비트 에러율 10^{-2} 까지는 충분히 사용 가능하다고 판단 된다. 비트 에러율 5×10^{-2} 에서는 부호율 1/2의 길쌈 부호와 Golay 부호가 거의 동일한 성능을 보인다. 다음 이들의 부호기는 Rayleigh 페이딩 채널상에서 발

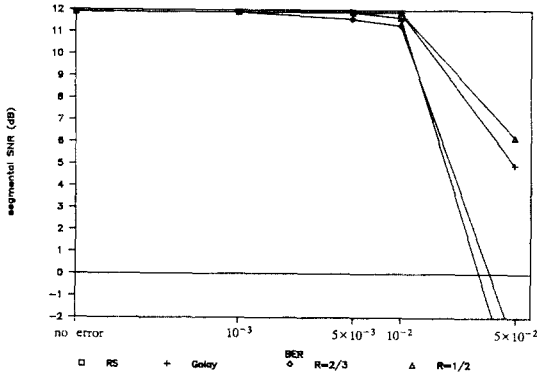


그림 6. 랜덤 비트 에러에서의 채널 부호기의 성능 비교
 Fig. 6. Performance comparison of channel coders in the random bit error environment.

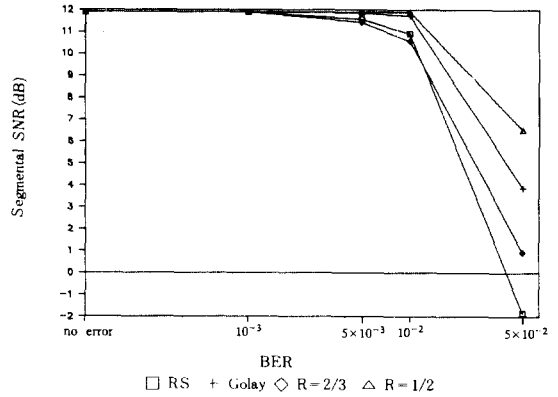


그림 7. Rayleigh 페이딩 채널에서의 채널 부호기의 성능 비교
 Fig. 7. Performance comparison of channel coders in the Rayleigh fading environment.

생하는 burst 에러에 대하여 시뮬레이션 되었는데 이때의 결과를 그림 7에 제시하였다. 그림 7의 결과는 그림 6의 결과와 거의 비슷한 추세를 보이지만 비트 에러율 10^{-2} 이하에서는 랜덤 비트 에러의 결과보다 단시간 SNR이 조금 낮음을 보인다. 그림 7로 부터 Rayleigh 페이딩 채널에서 모든 비트 에러율에 대하여 부호율 1/2의 길쌈 부호를 적용하였을때 SNR의 저하가 가장 적게 일어남을 관찰할 수 있다. 그리고 그 다음이 Golay 부호, Reed-Solomon 부호, 부호율 2/3의 소거식 길쌈 부호등의 순이다. 부호율 1/2의 길쌈 부호는 에러 발생율 10^{-2} 에서도 에러가 없을때와 비교하여 음질의 저하를 전혀 보이지 않을 만큼 Rayleigh 페이딩 채널에서 강력한 에러 정정 능력을 보인다. 에러 발생율 5×10^{-2} 에서 채널 부호화를 하지 않은 경우와 비교하면 길쌈 부호는 약 11.87 dB의 이득을 가져온 것을 알 수 있다. 채널 부호기를 포함한 음성 부호기가 가질 수 있는 전송율의 상한선을 16 kbps라 한다면 부호율 1/2의 길쌈 부호가 가장 강인한 에러 교정 부호라고 판단된다. 따라서 다음절의 각 부호기의 성능 비교에는 부호율 1/2의 채널 부호기를 적용한 CELP와의 비교결과를 제시한다.

2. 각 음성 부호기의 성능 비교

본 절에서의 성능 비교는 에러가 없는 경우와 비트 에러율 $10^{-3}, 5 \times 10^{-3}, 10^{-2}, 5 \times 10^{-2}$ 의 다섯가지 경우에 대하여 행하였다. 우선 에러가 없는 경우에 각 부호기의 음질을 비교하면 단시간 SNR적인 측면에서 SBC-APCM이 가장 우수한 성능을 보이고 그다음 MPE-LTP, RPE-LPC, CELP 등의 순이다. 실

제 청각 실험의 결과에도 이와 같은 추세를 보이고 있음을 알 수 있는데, SBC-APCM의 경우 대역 분할 필터들에 의한 약간의 음질의 왜곡이 감지되었다. 나머지 세개의 pulse excited 부호기들에서는 공통적으로 약간의 거칠은 음질을 보이는데 이는 여자의 음성에서 좀 더 심한 것을 확인할 수 있었다. 그러나 이상의 네가지 음성 부호기 모두 communication quality는 훨씬 상회하는 결과를 보였다.

다음 그림 8에 각 음성 부호기에 대하여 랜덤비트 에러에 대한 결과를 제시하였다. 비트 에러율이 10^{-2} 까지는 에러교정 부호의 적용 목적대로 단시간 SNR이 크게 저하되지 않으나 비트 에러율이 10^{-2} 보다 커지는 경우에는 단시간 SNR이 크게 저하된다. 그림 8의 결과로부터 SBC-APCM과 RPE-LPC의 비트 에러율 증가에 의한 단시간 SNR의 저하는 비슷한 비율로 나타남을 관찰할 수 있다. 이보다는 MPE-LTP에서의 성능 저하가 조금 완만함을 보이는데 이는 좀 더 많은 비트들이 에러 정정 부호에 의하여 보호되기 때문이다. 특히 CELP에서는 비트 에러율 10^{-2} 까지는 단시간 SNR의 저하가 전혀 없다.

다음 그림 9에는 burst 에러에 대한 단시간 SNR의 결과를 제시하였다. 그림 9의 burst 에러에 의한 결과는 랜덤 에러에서의 결과와 비교하여 보면 별로 큰 차이를 보이지는 않는다. 비트 에러율 10^{-2} 까지 각 부호기가 가지는 단시간 SNR의 곡선 순서는 랜덤 에러와 비교하였을때 변함이 없으나, 전체적으로 burst 에러의 경우 더 낮은 단시간 SNR을 보인다. 비트 에러율 10^{-2} 에서 burst 에러의 경우 랜덤 비트 에러의 경우보다 약 0.5-1 dB 정도 많은 왜곡을 보

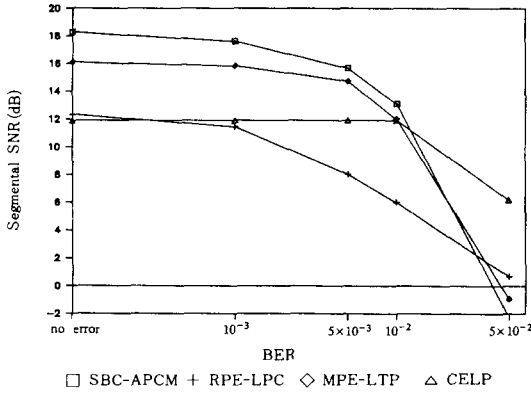


그림 8. 랜덤 비트 에러에서의 성능 비교
 Fig. 8. Performance comparison in the random bit error environment.

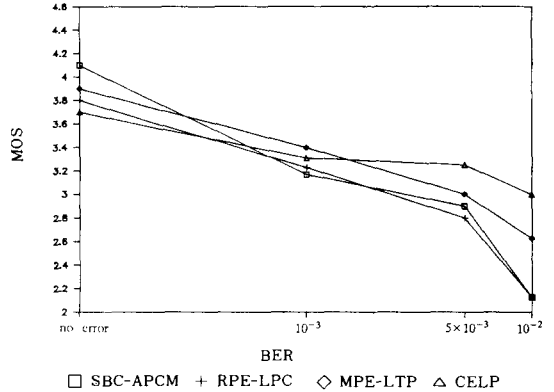


그림 10. Rayleigh 페이딩 채널에서의 주관적 청각실험 결과
 Fig. 10. Subjective listening test result in the Rayleigh fading environment.

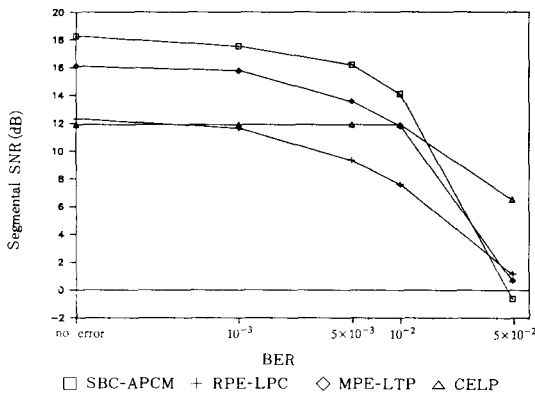


그림 9. Rayleigh 페이딩 채널에서의 성능 비교
 Fig. 9. Performance comparison in the Rayleigh fading environment.

이는데, 이것은 burst 에러의 경우 국부적으로 에러 정정 불능 상태가 발생하는 것으로 해석된다. 비트 에러율 5×10^{-2} 에서는 전체적으로 랜덤 에러에 의한 음질의 왜곡이 더 심한 것을 볼 수 있는데 이것은 랜덤 에러에 의하여 광범위하게 에러 정정 불능 상태가 발생하는 것에 기인한다. 한편 burst 에러에 대하여 시뮬레이션한 결과에 대해서는 주관적인 청각 조사가 실시되었다. 주관적인 청각 조사는 단시간 SNR이 실제의 음질을 비교 평가하는 데 있어서 인간의 청각을 잘 반영하지 못한다는 관점에서 발상된 것으로 현재 음성 부호기의 성능 비교에 많이 사용되고 있다. 주관적인 청각 조사방법에는 여러가지가 있으나 본 논문에서는 여러 명의 청취자가 5 단계로

판별하여 이들 값의 평균을 취하는 MOS (mean opinion score)^[14] 방식을 택하였다. 주관적인 청각 조사는 현재 이동 통신에서 고려되고 있는 비트 에러율인 비트 에러율 0에서 10^{-2} 까지의 경우로 제한하였으며 이때의 청취자는 12명, 사용된 음성 데이터는 64개로 하였다. 그림 10에 주관적인 청각 조사의 결과를 제시하였다. 주관적인 청각 조사의 결과는 단시간 SNR의 결과와는 상반된 양상을 보인다. 이중 가장 두드러진 것이 CELP와 MPE-LTP이다. MPE-LTP는 에러가 없는 경우에도 훌륭한 음질을 보이지만 비트 에러율이 커지는 경우에도 음질의 저하는 완만함을 보인다. CELP의 경우에는 모든 비트들이 에러 교정 부호로 보호됨으로써 에러가 많은 경우에도 음질의 손상이 거의 없음을 관찰할 수 있다. CELP의 경우 비트 에러율 10^{-2} 까지는 거의 음질의 왜곡이 나타나지 않고 비트 에러율 5×10^{-2} 에서도 음성의 이해도를 상실하지 않을 정도로 Rayleigh 페이딩 채널에서의 강인함을 보였다. 각 비트 에러율에 대한 MOS치를 평균하여 최종 평가를 표 5에 제시하였다. 채널 에러에 강인한 CELP가 가장 높은 결과를 보이고 그다음이 MPE-LTP이다. CELP는 비록 구조가 복잡하고 계산량이 많다는 단점이 있지만 이상의 결과로 부터 저전송율로 가장 훌륭한 성능을 가진다는 결론을 얻었다. CELP에서의 계산량을 줄이기 위한 효율적인 알고리즘이 Trancoso와 Atal^[17]에 의하여 제안되었는데, 제안된 알고리즘은 약 25 MOPS (millions of arithmetic operations per second)의 연산을 필요로 한다. 한편 SBC-APCM과 RPE-LPC는 약 1.5 MOPS 정도의 연산이, MPE-LTP

표 5. 평균 MOS 측정치
Table 5. Average MOS values.

SBC-APCM	RPE-LPC	MPE-LTP	CELP
3.075	2.992	3.233	3.315

에는 약 5 MOPS 정도의 연산이 필요하다고 알려져 있다.¹¹⁾ 이들과 비교하면 CELP는 상대적으로 많은 계산을 필요로 하지만 33 MOPS의 연산능력을 가지는 TMS 320C30^{11a)}과 같은 고성능 DSP(digital signal processor) chip의 발달로 조만간 실시간 하드웨어가 구현될 전망이다. 구조의 복잡성은 그렇게 문제가 되지 않는다고 판단된다. 한편 음성 부호기에서의 지연(delay)은 알고리즘에 의한 지연과 실제 프로세싱에 의한 지연 그리고 16 kbps의 직렬 전송에 의한 지연등에 의하여 결정된다. SBC-APCM과 RPE-LPC 그리고 MPE-LTP등의 부호기가 가지는 지연은 약 35ms-45ms 정도로 알려져 있다.¹¹⁾ CELP에서는 프레임의 길이가 20ms로 MPE-LTP와 동일하고, 전송율 또한 거의 동일하므로 CELP에서의 지연은 전송한 부호기에서의 지연과 크게 차이가 나지 않으리라 예상된다. 또한 CELP의 음질을 향상시키기 위하여 LTP 루프의 개선이라든지, 고주파 대역을 충실히 재생하는 기법등을 사용하여 음질을 향상시킨 사례가 있다.¹⁵⁾ 그러므로 이동 통신용의 음성 부호기로는 CELP가 가장 적합하다는 결론을 얻을 수 있다.

V. 결 론

본 논문에서는 디지털 이동통신에서의 음성부호화 기법으로 유망한 SBC-APCM, RPE-LPC, MPE-LTP 그리고 CELP등에 대하여 랜덤 에러 및 Rayleigh 페이딩 채널 상에서의 burst 에러에 대하여 성능을 평가 분석하였다. 또한 SBC-APCM, RPE-LPC, MPE-LTP에 대하여 비트 sensitivity를 측정하여 이를 토대로 비트 선택성 에러교정 부호를 적용한후 에러교정 부호를 적용하지 않은 경우와 비교하여 비트에러율 10^{-2} 에서 2-4 dB 정도의 이득이 있음을 보였다. 그리고 7.2 kbps CELP에서 음질 개선의 한 방법으로 최적양자기를 적용해본 결과 0.3 dB 정도의 SNR 향상이 있었다. 또한 CELP에서 채널 에러에 대한 강인성을 향상하기 위하여 Reed-Solomon 부호, Golay 부호, 길쌈 부호, 소거식 길쌈 부호등을 적용해본 결과 부호율 1/2의 길쌈 부호가 가장 좋다는 결론을 얻었다. 이상으로 부터 각 부호기의 성능평가 결과 burst 에러는 랜덤 에러의 결과와 비교하여 비

트 에러율 10^{-2} 에서 0.5-1 dB 정도의 단시간 SNR의 저하를 보이니, 모든 음성 부호기에 있어 거의 동일한 정도로 영향을 미치는 것을 보였다. 그리고 최종적인 주관적인 청각조사 결과 CELP, MPE-LTP, SBC-APCM 그리고 RPE-LPC 등의 순으로 성능이 좋다는 결론을 얻었다.

參 考 文 獻

- [1] J.E. Natvig, "Evaluation of six medium bit-rate coders for the Pan-European digital mobile radio system," *IEEE J. Select. Areas Commun.*, vol. SAC-6, pp.324-331, Feb. 1988.
- [2] J. Hagenauer, N. Seshadri and C. - E. Sundberg, "Variable-rate sub-band speech coding and matched channel coding for mobile radio channels," in *Proc. 38th IEEE Veh. Technol. Conf.*, 1988, pp.139-146.
- [3] P. Vary, R. Hofmann and K. Hellwig, "A regular pulse excited linear predictive codec," *Speech Commun.* vol. 7, pp. 209-215, 1988.
- [4] C. Galand, M. Rosso, P. Elie and E. Lancon, "MPE/LTP coder for mobile radio application," *Speech Commun.* vol. 7, pp. 167-178, 1988.
- [5] M.R. Schroeder and B.S. Atal, "Code-excited linear prediction (CELP): High quality speech at very low bit rates," in *Proc. ICASSP*, 1985, pp.937-940.
- [6] W.P. LeBlanc, S. Hanna and S.A. Mahmoud, "Performance of a low complexity CELP speech coder under mobile radio conditions," in *Proc. 39th IEEE Veh. Technol. Conf.*, 1989, pp.647-651.
- [7] S.U. Lee, Y.M. Chung, and J.M. Kim, "Performance comparison between the SBC and CVSD over the Rayleigh fading channel," in *Proc. 37th IEEE Veh. Technol. Conf.*, 1987, pp.512-518.
- [8] V. Lazzari, R. Montagna and D. Sereno, "Comparison of two speech codecs for DMR systems," *Speech Commun.*, vol. 7, pp. 193-207, 1988.
- [9] M. Berouti, H. Garten, P. Kabal and P. Mermelstein, "Efficient computation and encoding of the multiple excitation for LPC," in *Proc. ICASSP*, 1984, pp. 10.1.1-10.1.4.

- [10] P. Kroon and B.S. Atal, "Quantization procedures for the excitation in CELP coders," in *Proc. ICASSP*, 1985, pp. 1649-1652.
- [11] J.W. Modestino, N. Farvardin and M.A. Ogrinc, "Performance of block cosine image coding with adaptive quantization," *IEEE Trans. on Comm.*, vol. COM-33, pp. 210-217, Mar. 1985.
- [12] F.K. Soong and B. - H. Juang, "Line spectrum pairs (LSP) and speech data compression," in *Proc. ICASSP*, 1984, pp. 1.10.1-1.10.4.
- [13] S. Lin and D.J. Costello, Jr., *Error Control Coding: Fundamentals and Applications*, Englewood Cliffs: Prentice-Hall, 1983.
- [14] N.S. Jayant and P. Noll, *Digital Coding of Waveforms*. Englewood Cliffs: Prentice-Hall, 1984.
- [15] P. Kroon and B.S. Atal, "Strategies for improving the performance of CELP coders at low bit rates," in *Proc. ICASSP*, 1988, pp. 151-154.
- [16] J.P. Campbell, Jr., V.C. Welch and T.E. Tremain, "An expandable error-protected 4800 bps CELP coder (U.S. Federal standard 4800 bps voice coder)," in *Proc. ICASSP*, 1989, pp. 735-738.
- [17] I.M. Trancoso and B.S. Atal, "Efficient procedures for finding the optimum innovation in stochastic coders," in *Proc. ICASSP*, 1986, pp.2375-2378.
- [18] P. Papamichalis and R. Simar, Jr., "The TMS320C30 floating-point digital signal processor," *IEEE Micro*, vol. 8, pp. 13-29, Dec. 1988.

 著 者 紹 介

鄭 榮 謨 (正會員) 第25卷 第5號 參照
 현재 서울대학교 대학원
 제어계측공학과 박사과정

李 商 郁 (正會員) 第22卷 第1號 參照
 현재 서울대학교 공과대학
 제어계측공학과 부교수