

효율적인 LSF 양자화기를 이용한 QCELP 성능개선

Performance Improvement of the QCELP using an Efficient LSF Coding

김 해 진*, 강 상 원*
(Haejin Kim*, Sangwon Kang*)

※본 연구는 한국과학재단 핵심전문연구(과제번호:951-0913-083-2)의 지원을 받아 수행된 것입니다.

요 약

본 논문에서는 LSF계수의 효율적인 양자화기인 improved PSVQ(IPSVQ) 방식을 제안하고, 이를 CDMA 디지털 이동통신용 음성부호화기인 8 kbps QCELP에 도입하여 QCELP의 성능을 개선하였다. 프레임당 40 비트의 DPCM 양자화방식 대신 27 비트에서 transparent한 성능을 제공하는 IPSVQ 방식을 이용하여 13 비트를 절약하고, 이렇게 절약된 비트들을 음질에 중요한 영향을 끼치는 계수들에 해당하는 코드북 및 피치이득 계수의 양자화에 추가로 할당하여 QCELP의 전체적인 성능을 개선하였다. 객관적인 평가 방법인 SNR 및 SEGSNR이 각각 0.9 dB와 0.4 dB 정도 개선되었으며, 주관적인 음질시험을 통해서 성능향상을 확인할 수 있었다.

ABSTRACT

In this paper, an efficient LSF quantizer, named improved PSVQ(IPSVQ), is proposed to apply in the 8 kbps QCELP speech coder. By using 27 bits IPSVQ instead of 40 bits DPCM quantizer per frame, we can save 13 bits/frame and allocate those bits to the codebook gain and the pitch gain parameters. Hence we improve the overall performance of the QCELP codec. The enhanced QCELP shows the performance improvement of 0.9 dB SNR and 0.4 dB SEGSNR. Informal listening tests also confirm the improvement in the speech quality.

I. 서 론

디지털 이동통신시스템에서 기존의 음성전송품질(toll quality)을 그대로 유지하면서 시스템의 용량을 증대시키기 위한 핵심기술로서 음성부호화기는 매우 중요하다. 현재 표준안으로 채택된 디지털 이동통신용 음성부호화기로는 유럽의 GSM 표준인 13 kbps RPE-LTP[1], 북미의 IS-54 TDMA 표준인 8 kbps VSELP[2] 및 IS-96 CDMA용 음성부호화기 표준인 QCELP[3]를 들 수 있다. 그리고 이러한 음성부호화기들과 같은 음질을 유지하면서 전송속도가 절반인 half-rate 음성부호화기와 조만간에 도래할 개인통신시스템에 정착하기 위해 고 음질의 음성부호화기에 대한 연구가 현재 진행중에 있다. 이러한 차세대 음성부호화기를 구현하기 위해서는 음성신호의 단구간(short term) 상관도를 나타내는 LPC 계수를 효율적으로 양자화하는 것이 매우 중요하다. LPC 계수는 양자화 후

LPC 합성필터의 안정도 검사가 어려우며 dynamic range가 커 양자화에 적합하지 않기 때문에 line spectral frequencies(LSF) 계수[4]로 변환하여 이를 양자화한다. 미래의 음성부호화기에 적합한 LSF 양자화기는 적어도 프레임당 30 비트 이하에서 transparent한 성능을 갖는 양자화기라야 경쟁력 있는 LSF 양자화기가 될 것이다.

LSF 계수의 양자화방식은 크게 스칼라 양자화(scalar quantization, SQ)방식과 벡터 양자화(vector quantization, VQ)방식이 있다. 스칼라 양자화방식은 이미 많은 연구가 수행되어 32~40 비트에서 평균 SD가 1 dB 이하인 양자화기들이 많이 발표되었다[5]. 그러나 30 비트 이하에서는 성능이 우수하지 못하므로 효율적인 저전송용 음성부호화기의 LSF 양자화 방식으로는 적합하지 못하다. 반면에 벡터 양자화방식은 LSF 계수들을 코드북으로 양자화하는 방식인데 성능은 우수하나 코드북을 저장하기 위한 메모리 용량이 커야하고, 입력벡터와 각 코드워드들과의 왜곡을 계산하는데 소요되는 계산량이 많다는 단점이 있다[6]. 이러한 단점을 보완하기 위해서 여러가지 변형된 VQ 방식들이 제시되었는데, 그중 대표적인 것이 벡터-스

*한양대학교 제어계측공학과
접수일자: 1996년 7월 25일

칼라 양자화방식[7]과 split VQ(SVQ)방식[8]이다. 벡터-스칼라 양자화방식은 LSF 계수 10개를 하나의 코드북으로 먼저 양자화한 후 양자화에러 신호를 스칼라 양자화기로 양자화하는 방식이다. SVQ는 1993년에 Paliwal과 Atal에 의해 제시되었는데 VQ의 복잡도와 메모리용량을 감소시키기 위해 LSF 계수 10개를 2개 부분으로 나누어 각각을 다른 VQ 양자화기로 양자화하는 방식으로서 24 비트에서 양호한 성능을 얻었다. 그리고 1994년에는 SVQ의 복잡도와 용량을 더욱 감소시키기 위해 partition을 4개로 나누었으며, 이로 인한 성능감소를 보완하기 위해 LSF 계수의 프레임간 높은 상관도를 이용해 SVQ에 예측기를 부가한 predictive split VQ(PSVQ)방식이 Marca에 의해 제안되었다[9]. 이 방식은 40 msec의 전송프레임을 갖는 half-rate 부호화기에는 적용 가능하나, 각 20 msec LSF 프레임당 전송속도가 일정하지 않음으로 현재의 디지털 셀룰라 프레임구조에는 사용할 수 없다. 본 논문에서는 성능 열화 없이 half-rate와 full-rate 모두에 적용가능한 PSVQ 방식을 제안하였다. 27 비트 IPSVQ(Improved PSVQ) 방식은 기존의 방법보다 프레임당 13 비트를 절약한 것으로 절약된 비트를 코드북 이득과 피치 이득에 더 추가하여 표준화된 CDMA 이동통신용 음성부호화기인 QCELP에 적용한 결과, 객관적인 평가 방법인 SNR 및 SEGSNR이 각각 0.9 dB와 0.4 dB 정도 향상되었으며, 주관적인 음질시험을 통해서 성능향상을 확인할 수 있었다. II장에서는 QCELP에 대하여, III장에서는 본 논문에서 제안된 IPSVQ 방식에 대하여, IV장에서는 QCELP의 성능개선에 대하여, 마지막으로 V장에서는 결론을 맺었다.

II. QCELP의 개요

Qualcomm사가 광대역 디지털 셀룰라 표준안으로 제시한 QCELP는 AT&T에서 제시된 8 kbps CELP에 음성의 activity를 이용하여 전송율을 8, 4, 2, 1 kbps로 변환시킨 것으로서 평균전송속도를 약 4 kbps로 줄이는 가변 전송율 음성부호화 방식이다. 이 방식은 양방향 통화 시스템에서 묵음시간이 전체 통화시간의 50%~60% 정도를 차지한다는 점을 이용하여 음성품질의 저하없이 낮은 전송률로 음성신호를 부호화할 수 있다는 장점이 있다[3]. 이는 CDMA 디지털 이동통신 시스템에서 채널의 용량을 늘려주는 중요한 역할을 한다.

그림 1은 QCELP 인코더의 구조를 나타낸 것이다. 자기상관 알고리즘을 이용한 LPC 해석이 음성신호 s(n)에 수행되어 LPC 계수가 계산된다.

LPC 계수는 효율적인 양자화와 순서성질에 의한 안정도 특성을 갖고 있는 LSF 계수로 변환되어 부호화된다. 이러한 LSF 계수는 음성의 포먼트 주파수와 포먼트 대역폭을 나타낸다.

그리고, LSF 양자화기의 성능평가를 위해 다음과 같이 표현되는 spectral distortion(SD)값[5]을 사용한다.

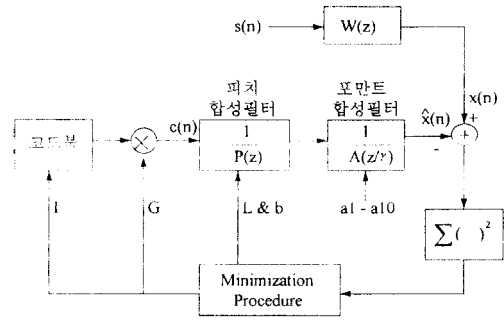


그림 1. QCELP 인코더의 기본구조도

$$SD = \left| \frac{1}{N_f} \sum_{n=1}^{N_f} \left(\frac{1}{\pi} \int_0^{\pi} (\log S_n(w) - \log \hat{S}_n(w))^2 dw \right) \right|^{1/2} \quad (1)$$

여기서, $S_n(w)$ 와 $\hat{S}_n(w)$ 는 각각 원음의 LPC power 스펙트럼과 양자화된 LPC power 스펙트럼을 나타내며, N_f 는 SD 계산에 사용된 전체 프레임 수를 표시한다. 설계된 양자화기는 transparent [8]한 성능을 보장해야 하는데, transparent한 성능이란 평균 SD가 1 dB 이하이며, 2~4 dB가 2%이내이고, 또한 4 dB 이상의 SD가 0%로서 사람의 귀로 LSF 계수의 양자화에러 영향을 거의 감지할 수 없는 수준이다. 각 LSF 계수는 그림 2에 보여진 차동(differential) 양자화기에 의해 부호화 되는데, bias가 제거된 각 LSF값에 이전 프레임에 근거한 예측값을 뺀 차이신호를 스칼라 양자화기로 양자화 한다. 이와같이 차동 부호화 방식은 LSF계수의 프레임간 상관관계만을 이용한 것으로 단순한 구조를 갖는 잇점이 있으나 transparent한 성능을 위해 프레임당 요구되는 비트가 많다.

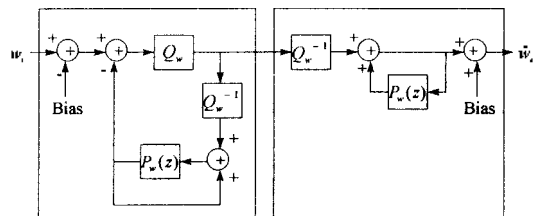


그림 2. QCELP에 사용된 LSF 양자화기

또한 QCELP 부호화기는 최적의 피치 및 코드북 파라미터를 찾기위해 분석/합성 기술을 이용한다. 즉 입력 음성신호와 합성 음성신호를 비교하여 그 오차를 최소화하는 피치 파라미터와 코드북 파라미터를 선택한 후 부호화한다.

III. 개선된 predictive SVQ

1. SVQ와 PSVQ의 양자화방식

SVQ방식[8]은 임의의 i번째 프레임의 p개 LSF계수, $(f_1^i, f_2^i, f_3^i, \dots, f_p^i)$ 를 M개의 subset으로 나누고 나누어

한 각 subset을 배디 양자화기로 양자화하는 방법이다. 일반적으로, M을 증가시킬수록 계산량과 복잡도는 줄어들지만 성능은 감소한다. 본 논문에서 고려한 SVQ는 실현 가능한 메모리용량과 계산량을 가지는 양자화기를 구현하기 위해 partition을 4로 나누었다. 각 subset의 크기는 실험에 의해 최적의 값이 설정되는데 기존의 연구결과[9]를 참조하여 10개 LSF계수를 3, 2, 2, 3의 subset들로 나누었다. SVQ는 partition을 증가시킬수록 성능이 감소하기 때문에 이러한 성능감소를 보완하기 위한 것이 predictive SVQ(PSVQ)[9]이다. PSVQ는 LSF 계수의 프레임간 높은 상관도를 이용해 두 프레임마다 한번씩 현재 프레임의 LSF값에서 이전프레임의 양자화된 LSF값을 뺀 에러신호를 양자화하는 방식이다. 즉, LSF 계수의 프레임간 상관도는 매우 높기 때문에 이를 이용한 에러신호의 dynamic range는 LSF 계수의 그것보다 훨씬 작아서 적은 비트로 양자화가 가능하다. 표 1은 training data로 사용된 음성신호의 임의의 i번째 프레임과 i-1번째 프레임간의 상관도를 나타내었다.

표 1. LSF 계수의 프레임간 상관도

j	1	2	3	4	5	6	7	8	9	10
ρ_j	0.82	0.87	0.87	0.91	0.91	0.88	0.87	0.85	0.84	0.82

예측 에러벡터값은 다음과 같이 표현될 수 있다.

$$\delta^{(i)} = \underline{f}^{(i)} - \hat{\underline{f}}^{(i-1)} \cdot \underline{a}^t \quad (2)$$

여기서, \underline{a} 는 예측계수의 벡터값이고, $\hat{\underline{f}}^{(i-1)}$ 는 $\underline{f}^{(i-1)}$ 의 양자화된 값이다. 프레임간 상관도의 이용은 DPCM 구조를 사용한 것이기 때문에 천이영역에서 slope 과부하와 한 프레임에서 발생한 에러가 다른 프레임에 영향을 미칠 우려가 있는데 이러한 것들을 방지하기 위해 예측을 두프레임마다 한번씩 하였다. PSVQ는 홀수번째 프레임은 LSF계수값을 4개의 subset으로 나눈다음 각각을 VQ 방식으로 양자화하고, 짝수번째 프레임은 해당된 LSF계수에서 이전프레임의 양자화된 LSF 계수값을 뺀 error 신호를 4개의 subset으로 나눈 다음 각각을 VQ방식으로 양자화하는 방법이다. 짝수번째 프레임은 error신호를 양자화하기 때문에 홀수번째 프레임보다 더 작은 비트를 할당해도 유사한 성능을 갖는다. 그러나 PSVQ는 half-rate의 전송프레임인 40 msec 동안은 전송속도가 일정하지만 LSF계수 해석구간인 20 msec동안에는 일정하지 못함으로 20 msec의 전송프레임을 이용하는 full-rate 음성부호화기에서는 사용할 수 없다. 따라서 본 논문에서는 QCELP의 성능을 향상시키기 위해, PSVQ를 성능열화 없이 full-rate 음성부호화기에서도 적용가능 하도록 개선한 improved PSVQ(IPSVQ) 양자화방법을 사용하였다.

2. IPSVQ 양자화방식

IPSVQ 양자화방식은 full-rate 음성부호화기에 적용가능하도록 매 LSF 전송프레임마다 LSF 계수를 직접 양자화하는 SVQ방식과 LSF 계수의 error신호를 양자화하는 differential SVQ방식이 동시에 사용된다. 그림 3은 IPSVQ의 기본구조를 나타낸 것이다. IPSVQ는 그림 3에서와 같이 홀수번째 프레임의 첫번째 및 두번째 subset은 LSF 계수를 직접 VQ방식으로 양자화하고, 나머지 2개의 subset은 현재프레임의 LSF계수값에서 이전프레임의 양자화된 LSF계수에 의한 예측값을 뺀 예측에러를 구한후 이를 VQ방식으로 양자화한다. 여기서 N 값은 각 subset의 dimension을 나타낸다. 그리고 짝수번째 프레임은 첫번째 및 두번째 subset은 예측에러값을 구한후 이를 VQ방식으로 양자화하며, 나머지 2개 subset은 LSF 값을 직접 VQ방식으로 양자화한다. 이렇게 함으로써 매 전송프레임마다 일정한 전송속도를 유지할 수 있다.

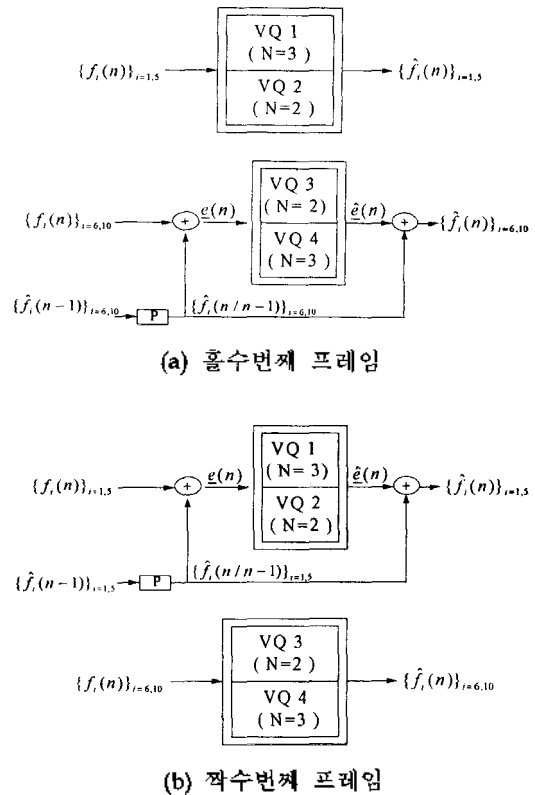


그림 3. IPSVQ(M=4)의 기본구조

이와같은 IPSVQ를 설계하기 위해서는 M(=4)개의 LSF subset들에 대한 코드북과 M개의 LSF 예측에러 subset 값들에 대한 코드북, 즉 총 2M개의 코드북이 설계되어야 한다. LSF subset들에 대한 코드북은 LBG 알고리즘에 의해 설계되는데 LSF계수값은 순서성질을 만족하기 때문에 설계된 코드북도 이를 만족해야 한다. 그러나 subset 벡터간에는 순서성질이 보장되지 않음으로 최적 코드워

드 탐색시 각 subset 벡터간 순서성질을 만족하는 코드워드만 탐색에 포함시켜야 한다. 즉 코드워드내의 LSF계수 값들은 순서성질을 만족하지만 인접 subset들간의 순서성질 만족여부를 알 수 없으므로, 현재 subset 양자화를 위한 코드벡터의 첫번째 요소값이 이전 subset의 양자화된 코드벡터의 마지막 요소값보다 큰 코드워드에 대해서만 탐색을 수행하면 프레임내 LSF값들간의 순서성질을 만족할 수 있다. 예측에러 신호의 코드북에 대해서도 순서성질을 검사하여야 한다. 예측에러값 자체로서는 순서성질의 만족여부를 알 수 없으므로 양자화된 예측에러값을 다시 LSF 계수로 복원한 후 순서성질을 검사하고 순서성질을 만족하는 예측에러 코드북에 대해서만 코드북 탐색을 하여야 한다. 그리고 partition이 3인 IPSVQ는 홀수번째 프레임의 첫번째 및 두번째 subset은 LSF 계수를 SVQ로 양자화하고, 나머지 1개 subset은 LSF 계수의 예측 에러신호를 구한후 이를 VQ 방식으로 양자화한다. 그리고 짝수번째 프레임은 첫번째 subset은 LSF 예측 에러값을 VQ 양자화하며, 나머지 2개 subset은 LSF 값을 직접 VQ 양자화한다. M=3인 IPSVQ는 계산량과 복잡도가 M=4인 IPSVQ보다 크지만 더 적은 비트로 우수한 성능을 가질 수 있다. 계산량 및 메모리용량 측면에서 27 비트 IPSVQ는 PSVQ와 같으며, 25 비트 IPSVQ(M=3)는 이들보다 메모리용량 및 계산량이 더 요구되나, 양자화 에러 측면에서 훨씬 유리하다.

3. Weighted distortion measure

VQ의 설계 및 탐색시 최적의 코드벡터를 찾기위해 어떠한 distortion measure를 사용할 것인가는 중요한 문제이다. 일반적으로 LBG 알고리즘[6]에서 사용되는 distortion measure는 Euclidean measure인데, 이것은 모두 LSF 계수에 똑같은 가중치를 부여한다. 그러나 비포먼트영역보다 포먼트영역의 LSF 값을 더 정확하게 양자화하는 것이 더 중요하다. 본 논문에서는 가중치를 부여하는 여러 가지 방법이 존재하나 간단하면서도 성능이 우수한 Farvardin이 제시한 가중치 부여방식을 사용하였다[5]. 이 방식의 기본적인 배경은 포먼트영역에 해당하는 LSF계수 값들이 매우 가까이 위치해 있다는데 있다. 그리고 가중치 값은 다음과 같이 표현된다.

$$w_j = (f_j^0 - f_{j-1}^0)^{-1} + (f_{j+1}^0 - f_j^0)^{-1} \quad (3)$$

여기서 $j=1, \dots, p$ 이며 경계치 값 f_0^0 와 f_{p+1}^0 는 각각 0과 0.5이다. 이러한 가중치는 매 프레임마다 구해져야하며 이를 이용한 distortion measure는 다음과 같이 표현된다.

$$d_s(f, \hat{f}) = \sum_{j=1}^p w_j (f_j - \hat{f}_j)^2 \quad (4)$$

4. IPSVQ 실험결과

제시된 IPSVQ 양자화기의 효율성이 시뮬레이션으로 평가되었다. LSF 양자화기의 코드북 설계를 위해 3.2 kHz로 대역제한 되어진 후 8 kHz로 샘플링된 88초 길이의 한국어 음성을 training 데이터로 사용하였으며, LSF 양자화기의 성능평가를 위해 설계시 이용되지 않은 9개의 한국어 단문이 사용되었다. 20 msec의 해석 프레임을 갖는 10차 LPC 분석이 20 msec의 Hamming window를 이용한 자기상관(auto correlation) 방식에 근거해서 이루어졌다.

성능평가를 위해 식(1)로 표현되는 spectral distortion (SD)값을 사용하였으며, VQ 코드북 설계 및 탐색시 식(4)로 표현되는 가중 distortion measure를 이용하였다. 표 2는 LSF 양자화기들의 시뮬레이션시 사용된 각 subset당 비트할당 값을 보였다. 앞서 언급한대로 PSVQ 양자화기의 경우는 짝수 및 홀수번째 프레임의 프레임당 비트수가 다르지만, IPSVQ 양자화기에서는 프레임당 비트수가 일정함을 알 수 있다. 표 3은 IPSVQ 방식의 SD 성능을 프레임당 사용된 비트에 따라 나타내었다. 4개의 subset

표 2. LSF 양자화기들의 성능분석시 비트할당

subsets	27 비트	27 비트 PSVQ		27 비트 IPSVQ		25 비트 IPSVQ	
	SVQ	홀수	짝수	홀수	짝수	홀수	짝수
(f_1, f_2, f_3)	8	8	6	8	7	8	10
(f_4, f_5)	6	7	5	7	5	7	7
(f_6, f_7)	6	7	6	5	7	10	7
(f_8, f_9, f_{10})	7	8	7	7	8	7	8
Total (비트/프레임)	27	30	24	27	27	25	25

표 3. IPSVQ 방식의 SD 성능

비트	SD (dB)	outliers (%)	
		2 ~ 4 dB	> 4 dB
27 (M=4)	0.93	2.31	0
26(M=4)	0.98	2.31	0
25(M=4)	1.11	3.54	0.07
25(M=3)	0.93	2.58	0

표 4. LSF 양자화방식들의 SD 성능비교

양자화방식	비트	SD (dB)	outliers (%)	
			2 ~ 4 dB	> 4 dB
SVQ (M=4)	28	1.03	2.40	0
PSVQ (M=4)	27	0.99	2.81	0
IPSVQ (M=4)	27	0.93	2.31	0
IPSVQ (M=3)	25	0.93	2.58	0
Differential SQ	40	0.79	0.66	0.083

을 이용하는 경우 프레임당 26 비트에서 transparent한 성능을 얻었으며, 3개의 subset인 경우는 프레임당 25 비트에서 transparent한 성능을 얻었다. 표 4는 SVQ 및 PSVQ 방식과 IPSVQ 방식과의 SD 성능비교를 보였다. 27 비트 IPSVQ 방식이 28 비트 SVQ 방식이나 27 비트 PSVQ 방식보다 성능이 우수함을 알 수 있다.

IV. QCELP의 성능개선

QCELP 부호화기의 음질에 주요한 영향을 끼치는 계수값들로 코드북 이득과 피치 이득을 들 수 있는데, 현재 이들 계수값들은 각각 3비트씩 양자화된다. 그런데 이보다 많은 비트를 할당하면 성능이 보다 향상될 수 있을 것이다. 표 5은 QCELP 구조에 코드북 이득만 비트할당을 늘렸을때의 성능향상을, 그리고 표 6는 피치이득만 비트할당을 늘렸을때의 성능향상을 나타낸것이다. 표에서와 같이 코드북 이득에 4 비트를 할당하면 SNR과 SEGSR이 각각 1.0 dB 및 0.4 dB정도 향상되며 피치 이득에 4 비트를 할당하면 SNR과 SEGSR이 각각 0.3 dB 및 0.1 dB 정도 향상됨을 볼 수 있다. 각 이득을 양자화하지 않았을 때의 성능도 최대한계치라는 의미로 포함하였다. 코드북 이득의 경우 3 비트 보다 4 비트를 할당한 경우 성능향상이 현저하지만, 4 비트할당과 양자화하지 않은 성능은 비슷하므로 4 비트 보다 많이 할당하더라도 성능증가가 크게 예상되지 않는다. 따라서 27 비트 IPSVQ를 QCELP에 도입시 기존의 LSF 양자화 방식과 비슷한 성능을 유지하면서 프레임당 13 비트를 절약할 수 있으며, 이렇게 절약된 비트를 실제 음질에 많은 영향을 끼치는 코드북 이득과 피치 이득에 매 부프레임마다 각각 1비트씩 더 추가함으로써 전체적인 성능개선을 이룰 수 있다. 피치 이득은 0~2를 16 레벨로 균등 양자화하였고, 코드북 이득은 dynamic range를 줄이기 위해 로그값을 취한 후 DPCM 방식으로 4 비트 양자화하였다.

표 5. 코드북 이득의 비트할당에 따른 성능

SNR/SegSNR in dB		
3 비트	4 비트	Unquantized
14.970 / 12.770	15.977 / 13.173	16.214 / 13.595

표 6. 피치 이득의 비트할당에 따른 성능

SNR/SegSNR in dB		
3 비트	4 비트	Unquantized
14.970 / 12.770	15.226 / 12.885	15.166 / 12.952

표 7에 기존의 QCELP 비트할당과 개선된 방식의 비트 할당이 나타나 있다. 표 8은 이러한 개념으로 IPSVQ를

QCELP에 부가했을때의 개선된 성능을 나타내었다. 객관적인 성능인 SNR 및 SEGSR이 각각 0.9 dB와 0.4 dB 정도 향상됨을 알 수 있다.

표 7. 기존의 방식과 개선된 방식의 비트할당(full-rate의 경우)

	QCELP의 비트할당				개선된 방식의 비트할당			
LPC frame	40				27			
Pitch subframe	10	10	10	10	11	11	11	11
Codebook subframe	10	10	10	10	10	10	10	10

표 8. IPSVQ를 도입한 QCELP의 성능

SNR/SegSNR in dB	
QCELP	개선된 QCELP
14.970 / 12.770	15.821 / 13.138

그림 4에 원래의 음성신호, QCELP 합성신호 및 개선된 QCELP 합성신호를 나타내었다. 개선된 QCELP는 기존의 방식에 비해 700~1300 샘플에 해당하는 영역에서 성능향상을 확인할 수 있다. 또한 주관적인 평가방법인 음질시험에 의해 음성의 명료도 및 자연성이 향상되었음을 확인할 수 있었다.

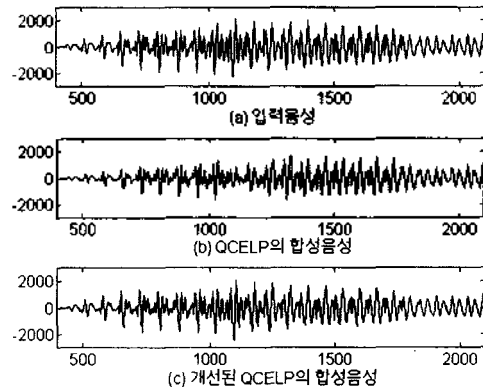


그림 4. 음성신호 '개인'의 합성신호 비교

V. 결 론

본 논문에서는 LSF계수의 효율적인 양자화기인 improved PSVQ(IPSVQ) 방식을 CDMA 디지털 이동통신용 음성부호화기인 8 kbps QCELP에 도입하여 QCELP의 성능을 개선하였다. 프레임당 27 비트에서 transparent한 성능을 제공하는 IPSVQ 방식을 이용하여 절약된 13 비트를 음질에 중요한 영향을 끼치는 계수들에 해당하는 코드북 및 피치이득 계수의 양자화에 추가로 할당하여

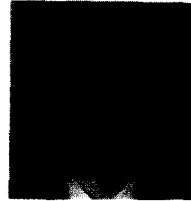
QCELP의 전체적인 성능을 개선하였다. 객관적인 성능인 SNR 및 SEGSNR이 각각 0.9 dB와 0.4 dB 정도 개선되었으며, 주관적인 유질시험을 통해서 성능향상을 확인할 수 있었다.

참 고 문 헌

1. P. Kroon, et al, "Regular-Pulse Excitation-A novel approach to effective and efficient multipulse coding of speech," IEEE Trans. ASSP, vol.ASSP-34, pp.1054-1063, Oct. 1986.
2. I. A. Gerson and M. A. Jasiuk, "Vector sum excited linear prediction(VSELP) speech coding at 8 kbps," Proc. IEEE ICASSP, vol.1, pp.461-464, April 1990.
3. TIA/EIA/IS-96 Interim Standard, Telecomm. Industry Association, April 1994.
4. F. K. Soong and B. H. Juang, "Line spectrum pair(LSP) and speech data compression," Proc. IEEE ICASSP, pp.1.10.1~1.10.4, Mar. 1984.
5. N. Sugamura and N.Farvardin, "Quantizer design in LSP speech analysis-synthesis," IEEE J. Select. Areas in Commun., vol.6, pp.432-440, Feb. 1988.
6. Y. Linde, A. Buzo and R. M. Gray, "An algorithm for vector quantizer design," IEEE Trans. Commun., vol.COM-28, pp.84-95, Jan. 1980.
7. J. Grass and P. Kabal, "Methods of improving vector-scalar quantization of LPC coefficients," Proc. IEEE ICASSP, pp. 657-660, Mar. 1984.
8. K. K. Paliwal and B. S. Atal, "Efficient Vector Quantization of LPC parameters at 24 Bits/frame," IEEE Trans. speech and audio proc. vol.1, pp.3-14, Jan. 1993.
9. J. B. Marca, "An LSF quantizer for the north-American half-rate speech coder," IEEE Trans. on vehicular tech. vol. 43, no.3, pp.413-419, August 1994.

▲김 해 진(Haejin Kim)

1971년 9월 28일생



1994년 2월: 창원대학교 제어계측공학과 졸업(학사)

1996년 2월: 한양대학교 제어계측공학과 졸업(석사)

현재: 고등기술원 정보통신연구실 연구원

▲강 상 원(Sangwon Kang)

현재: 한양대학교 제어계측공학과 조교수

14권 2E호 참조