2중 코드북 구조를 통한 CELP 음성부호화기의 성능 향상에 관한 연구

°김종우*, 김응곤**, 한승조*
* 조선대학교 전자정보통신공학부, ** 순천대학교 컴퓨터과학과

A Study on the performance improvement of the CELP coder by the structure of dual codebook

°Jong-Woo Kim*, Eung-Kon Kim**, Seung-Jo Han*
* School of Electronics and Information Communication Eng, Chosun Univ.
** Dept. of Computer Science, Sunchon National Univ.

요 약

본 논문에서는 CELP 부호화기의 계산량을 줄이면서도 고음질의 음성을 합성할 수 있는 코드북 구조를 제안한다. 제안한 코드북 구조는 불규칙 코드북과 희박 중첩형 코드북 두 개의 코드북의 합으로 여기신호를 표현한다. codebook Ⅰ에서 잔류신호와 오차가 적은 여기신호열을 구한 후, 이 여기신호열에 codebook Ⅱ의 여기신호열을 합하여 최적의 여기신호열을 구한다. 또한 이로 인한 전송 비트수의 증가를 막기위해 홀수 프레임에서는 두 개 코드북의 index를, 짝수 프레임에서는 codebook Ⅰ의 여기신호열은 그대로 사용하고 codebook Ⅱ에서만 검색하여 전송하는 방법을 사용하였다. 이러한 2중 코드북 구조는 두 개의 여기신호열의 합으로 표현되고 각각의 서로 다른 코드북 이득을 사용하기 때문에 정확한 이득을 표현할 수 있어 기존의 개선 알고리듬보다 더 나은 음질을 제공할 수 있다. 검색시간이 빠르고, 본 코드북 구조를 갖는 4.8kbps CELP형 부호화기를 설계하여 컴퓨터 모의 실험한 결과, 같은 전송률을 갖는 DoD CELP 부호화기보다 *seg*SNR가 0.53dB 더 높게 나타났다.

1. 서 론

디지털 기술의 발달과 함께 음성 부호화기의 성능은 꾸준히 발달되어 왔다. 이러한 음성 부호화법은 크게 파형 부호화법, 신호원 부호화법, 두 기법을 혼합한 혼성 부호화법 등으로 나뉜다. 이동 통신 시스템에서는 적은 데이터량으로도 높은 음질의 음성을 합성할 수 있는 혼성 부호화법을 사용하고 있다. 이러한 혼성 부호화법들 중에서도 CELP 음성 부호화기[1]는 16kbps이하의 전송률에서 가장 효율적인 압축 방식으로 이용되어 왔다. 음성부호화기는 최근 디지털 셀룰러나 PCS 같은 이동 통신시스템의 수요가 급증함에 따라 제한된 채널을 보다 효율적으로 사용하기 위한 연구가 계속되고 있다. 이에 음질의 저하 없이 전송률을 낮추기 위해 여러 가지 부호화기가 제안되었다[2-5]. 그 예로 ITU의 표준 음성 부호화기인 LD-CELP[2]와 CS-ACELP[3], VSELP[4], EVRC, QCELP등의 CELP방식의 부호화기가 개발되었다.

CELP 음성 부호화기는 Analysis-by-Synthesis 방법인, 입력 음성을 분석하여 필요한 파라미터를 추출하고 이를 이용, 코드북내에 저장된 코드워드와 합성하여 입력 음성과의 오차가 가장 적은 코드워드를 선택하는 방식에 기초를 둔다. 이러한 CELP 음성 부호화기는 여기신호(excitation signal)의 부호화과정인 최적의 코드워드를 찾는 코드북 구조에 따라 부호화기의 종류가 나뉘고 성능이 달라진다.

본 논문에서는 기존의 CELP 음성 부호화기의 코드북 구조를 개선함으로써 적은 비트로 더 나은 음질을 갖는 코드북 구조를 제안한다. 제안된 코드북 구조는 기존의 코드북 구조와는 달리 2중 코드북 구조를 적용함으로써 여기신호를 더욱 잘 표현할 수 있다.

2. CELP 음성 부호화기의 구조

CELP(Code Excited Linear Prediction) 음성 부호화기는 코드북내에 저장되어 있는 입력 여기신호열을 두 개의 시변 선형 회귀 필터를 통과시킴으로써 얻은 신호 중, 주어진 충실도 판정을 최적화 시키는 것을 선택하도록 구성되어 있다. 음성신호를 두 개의 시변 선형 회귀 필터로 필터링한 후 남아 있는 신호를 잔류신호(여기신호)라고 하는데 이 신호는 음성 신호보다 낮은 분산을 가지고 있어서 음성신호보다 더 쉽게 양자화할 수 있다. 그러나 낮은 전송률을 갖는 부호화기는 잔류신호를 위해서 할당된 비트수가 적으므로 직접 양자화가 힘들기 때문에 코드북을 이용하여 양자화 한다. 이 코드북의 구조에 따라 부호화기의 성능이 좌우된다. 잔류신호와 코드북내의 여기신호열과의 오차를 최소화시킨다고 해서 입력 음성신호와 합성신호의 오차가 최소화된다고 말할 수 없다. 그래서, CELP 음성 부호화기는 코드북내의 여기신호열과 입력 음성신호로부터 구한 파라미터를 이용하여 합성한 음성신호와 입력 음성신호를 비교하여 오차를 최소화하는 여기신호열로 잔류신호를 양자화하는 소위 합성에 의한 분석(Analysis-by-Synthesis)법을 사용한다. 그러

나 매번 음성을 합성해서 비교해야 하므로 매우 복잡한 구조를 가지며 많은 계산량을 필요로 한다. 또, 제한된 코드북구조 때문에 여기신호를 잘 표현하지 못하는 단점이 있다. 그림 1은 CELP 음성 부호화기의 기본 구조를 나타낸 것이다.

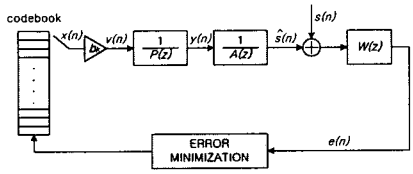


그림 1. CELP 음성 부호화기의 기본 구조

$P(z)$와 $A(z)$는 피치 필터와 스펙트럼 필터이고 이의 계수는 입력된 음성신호 $s(n)$을 이용하여 구한다. 피치 합성필터와 스펙트럼 합성필터는 다음과 같이 표현될 수 있다.

$$\frac{1}{P(z)} = \frac{1}{1 - bz^{-M}} \quad (1)$$

$$\frac{1}{A(z)} = \frac{1}{1 - a_1 z^{-1} - a_2 z^{-2} \cdots - a_m z^{-m}} \quad (2)$$

식(1)의 피치 필터 파라미터 b와 M은 개회로(Open Loop) 구조나 폐회로(Close Loop) 구조 중의 하나를 이용해서 얻어질 수 있는데, 폐회로 구조를 이용하여 구한 경우의 음질이 훨씬 우수함으로 일반적으로 폐회로 구조를 사용한다. 식(2)의 LPC 계수는 음성 압축시 양자화에 부적절하다고 알려져 왔다. 그래서 LPC 계수화 같은 스펙트럼에 관한 정보를 가지며, 안정성이 양자화시에도 보장되고 생동폭이 매우 좁은 특성을 가진 LSP 계수로 변환하여 사용한다.

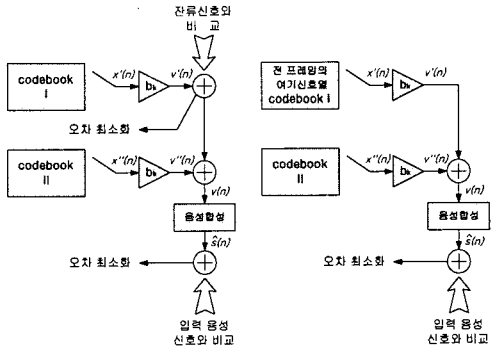
3. 2중 코드북 구조

본 논문에서 제안한 코드북 구조는 기존의 코드북 구조와는 달리 2중으로된 코드북으로 잔류신호와 합성음성신호를 2번 비교함으로써 여기신호를 잘 표현할 수 있다. 그림 2는 제안한 코드북 구조를 나타낸 것이다.

2중 코드북의 구조는 codebook Ⅰ은 불규칙코드북, codebook Ⅱ는 희박 중첩형 코드북으로 구성된다. codebook Ⅰ내의 여기신호열 $x'(n)$을 b_k를 통해 적절하게 이득을 조절한 $v'(n)$과 잔류신호 $\overline{v}(n)$과의 차이를 오차 가중 필터를 통해 오차가 가장 적은 여기신호열을 구한다. 그리고 codebook Ⅱ내의 여기신호열 $x''(n)$을 b_k를 통해 이득을 조절한 $v''(n)$과 codebook Ⅰ에서 구한 $v'(n)$을 더하여 합성 여기신호$v(n)$을 구한다. 그 후 기존의 CELP 음성부호화기와 마찬가지로 이 여기신호$v(n)$을 두 개의 시변 선형 필터를 통과시켜 음성을 합성하여 입력음성과 비교하여 오차를 최소화하는 여기신호열을 $x''(n)$을 구한다.

2중 코드북 구조는 codebook Ⅰ은 -1, 0, 1의 신호만을 갖는 크기가 N인 불규칙코드북으로 구성되어있고, codebook Ⅱ는 -1, 1과 80%이상의 0의 값을 갖고 인접 코드워드간에 상관성을 갖는 크기가 M인 희박 중첩형 코드북으로 구성되어 있다.

여기신호는 codebook Ⅰ과 codebook Ⅱ의 합으로 구하기 때문에 이론적으로 N×M의 크기의 불규칙 코드북을 가지는 CELP 부호화기의 성능을 낸다. 또한 코드북 검색시 N번 잔여신호와 비교하고 M번 음성을 합성하여 비교하므로, 코드북내의 모든 여기신호열으로 음성을 합성하여 비교하는 기존의 방법보다 코드워드 검색시간이 단축된다. 즉 희박 중첩형 코드북을 가지는 부호화기에 비해 검색 속도가 더 빠르고 더 나은 음질을 가지는 불규칙 코드북의 성능을 가진다는 장점이 있다.



(a)홀수 부프레임 (b)짝수 부프레임

그림 2. 2중 코드북 구조

2중 코드북 구조는 두 개 코드북 각각의 이득을 사용하기 때문에 정확한 이득을 표현할 수 있어 더 나은 음질을 합성할 수 있다. 그러나 이로 인한 코드북에 필요한 비트수가 증가하게 된다. 이러한 문제를 해결하기 위해서 다음과 같은 방법을 사용였다. 입력음성 분석할 때 한 프레임당 네 개의 부프레임으로 나뉘는데 첫째, 셋째 부프레임에는 두 개 모두의 코드북에서 검색 후, codebook Ⅰ, Ⅱ의 여기신호열 index와 각각의 이득 모두를 전송한다. 둘째, 넷째 부프레임에서는 앞 프레임에서 구한 codebook Ⅰ의 여기신호열은 그대로 이용하고 이득을 조절한 후 codebook Ⅱ에서만 검색하여, 두 개 코드북의 이득과 codebook Ⅱ의 여기신호열 index만을 전송한다. 음성신호는 각 부프레임간에는 비교적 유사한 성질을 가지고 있고, 음질은 codebook Ⅰ에서의 최적 여기신호열 선택 여부보다는 이득과 codebook Ⅱ에서의 여기신호열 선택에 따라 더욱 영향을 미친다. 그렇기 때문에 이와 같은 방법을 사용하여 음질의 저하를 최소화하면서 많은 비트수를 줄일 수 있다. 또한 짝수번째에는 codebook Ⅰ의 검색을 생략하기 때문에 검색속도도 향상된다. 표 1은 본 논문 실험에 사용된 4.8kbps CELP 음성부호와기의 비트수 할당의 예이다.

표 1. 4.8kbps CELP 음성부호화기의 비트수 할당 예

	frame															
	subframe				subframe				subframe				subframe			
codebook	Ⅰ		Ⅱ		Ⅰ		Ⅱ		Ⅰ		Ⅱ		Ⅰ		Ⅱ	
index gain	*i*	*g*	*i*	*g*	*i*	*g*	*i*	*g*	*i*	*g*	*i*	*g*	*i*	*g*	*i*	*g*
bit	5	4	6	4	·	4	6	4	5	4	6	4	·	4	6	4

4. 실험 및 성능 평가

제안된 방법의 성능 평가를 위하여 제안한 코드북 구조의 4.8kbps 음성 부호화기를 C 언어로 프로그래밍하여 컴퓨터 모의 실험을 하였다. 4.8kbps 음성 부호화기의 규격은 표 2와 같고 표 3과 같이 비트를 할당하였다.

표 2. 4.8kbps 음성 부호화기의 규격

항 목	값
Data Rate	4800 bit/sec
Sampling Rate	8kHz
Frame Size	240 samples(30ms)
Pitch Subframe Size	60 samples(7.5ms)
Codebook Subframe Size	60 samples(7.5ms)
LPC Filter Order	10
Pitch Predictor Order	1
Codebook Order	32×64 =2048

표 3. 4.8kbps의 음성 부호화기의 Bit 할당

Parameter	bit/7.5ms subframe	bit/30ms frame
LPC(LSP)	3,4,4,4,4,3,3,3,3,3	34
Pitch	7, 5, 7, 5	24
Pitch gain	5	20
Codebook	5+6, 6, 5+6, 6	34
Codebook gain	4+4	32
Total		144

한 프레임의 길이는 30ms(240samples)로 하였으며, 각 프레임의 부프레임은 7.5ms(60samples)씩 네 개로 구성하였다. LPC 분석을 위해 37.5ms(300samples)의 Hamming 창함수를 7.5ms(60samples)씩 중첩하여 사용하였다. 포먼트 필터는 자기상관관계 방법으로 10개의 LPC를 구하고 전송을 위해서 LSP(Line Spectrum Pair)로 변환하여 프레임간의 LSP 값의 차이를 34Bit로 스칼라 양자화를 사용하여 부호화 하였다. 피치 값은 첫 번째, 세 번째 부프레임에서는 부프레임 모든 구간에서의 피치 지연값을 7Bit로 부호화 하였고, 두 번째, 네 번째 부프레임에서는 바로 앞 부프레임과의 상대적 차를 5Bit로 부호화 하였다. 피치 이득은 각 부프레임당 5Bit로 코드북은 홀수 부프레임 11(5+6)Bit, 짝수 부프레임 6Bit 씩 전부 34Bit씩으로 부호화 하였으며, 코드북 이득은 각 부프레임당 8(4+4)Bit씩 32Bit로 부호화 하였다. 총 한 프레임당 144Bit로 부호화 하였으며, 4.8kbps의 전송률을 필요로 한다.

제안된 코드북 구조의 4.8kbps 음성 부호화기의 성능을 평가하기 위해 비교 기준 부호화기로 4.8kbps DoD CELP 음성 부호화기를 선정하였다. 성능평가에는 객관적 평가 방법인 *seg*SNR(*segmental* SNR)을 사용하였다. 평가 시료로는 20대 남녀화자 각각 2명이 다음 문장을 발음한 시료를 사용하였다. 표 4는 평가 결과를 나타낸다.

1. 영, 일, 이, 삼, 사, 오, 육, 칠, 팔, 구, 십
2. 여보세요. 안녕하십니까?
3. 지금은 전화를 받을 수 없으니, 다음에 다시 걸어 주십시오.

표 4. *seg*SNR 비교표

	4.8kbps DoD CELP	제안된 코드북 구조의 CELP
문장 1	12.67	13.15
문장 2	13.04	13..59
문장 3	12.31	12.87
평균	12.67	13.20

표 4에서 알 수 있듯이 제안된 코드북 구조를 가지는 음성 부호화기는 4.8kbps DoD CELP 음성 부호화기보다 *seg*SNR가 더 높게 나왔다. 제안된 코드북 구조를 이용한 음성부호화기의 *seg*SNR는 13.20dB로 4.8kbps DoD CELP 음성 부호화기의 12.67dB보다 0.53dB 더 높게 나왔다. 제안된 코드북 구조를 갖는 CELP 음성 부호화기는 동일한 조건에서 4.8kbps DoD CELP 음성 부호화기보다 더 우수한 성능을 보였다.

5. 결 론

본 논문에서는 2중 코드북 구조를 이용하여 코드북 검색 시간을 줄이면서도 더 나은 음질을 합성할 수 있는 코드북 구조를 제안하였다. 이 코드북 구조는 불규칙 코드북과 희박 중첩형 코드북 두 개로 구성되어 있어 두 개의 코드북의 합으로 음성의 여기신호를 표현함으로써 더 나은 음질을 합성할 수 있다. 또한 짝수 부프레임에서는 첫 번째 코드북의 여기신호열은 바로 앞 부프레임에서 구한 값을 그대로 사용하고 두 번째 코드북에서만 검색하여 전송함으로써 검색속도와 두 개의 코드북 이득 사용으로 인한 비트수의 증가를 줄였다.

2중 코드북 구조는 두 개의 코드북과 두 개의 코드북 이득을 사용하여 여기신호를 표현함으로써 다양한 특성을 가지는 음성신호를 잘 표현할 수 있었다. 이러한 코드북 구조를 가지는 음성 부호화기는 실시간 구현이 가능한 연산량을 가지고 성능을 분석한 결과 4.8kbps DoD CELP 음성 부호화기에 비해 성능이 우수함을 확인하였다.

참 고 문 헌

[1] M. R. Schroeder, B. S. Atal, "Code-Excited Linear Prediction(CELP) High-Quality Speech at Very Low Bit Rates," IEEE Proc. Int. Conf. Acoust. Speech and Signal Proc., pp. 937-940, 1985.

[2] "ITU Recommendation G.728 : Coding of speech at 16kbit/s using Low-Delay Code Excited Linear Prediction," 1992.

[3] "ITU Recommendation G.729 : Coding of speech at 8kbit/s using Conjugate_Structure Algebratic-Code-Excited Linear Predictive (CS-ACELP)," 1996.

[4] I. A. Gerson, M. A. Jasuik, "Vector-Sum Excited Linear Prediction(VSELP) Speech Coding at 8Kbps," IEEE Proc. Int. Conf. Acoust. Speech and Signal Proc., pp. 641-644, 1990.

[5] 최용수, 강홍구, 박상욱, 윤대희, "4kbps PSI-VSELP 음성 부호화 알고리듬" 한국음향학회지, vol. 15, No. 6, pp. 59-65, 1996.

[6] L. Torres and J. Huguet, "An omprovement on codebook search for vector quantization," IEEE Trans. Comm., COM-42, pp.208-210, 1994.

[7] 박호종, 권순영. "고성능 CELP 음성 압축기를 위한 새로운 코드북 구조" 한국음향학회지, Vol. 17, No. 2, pp. 43-49, 1998