

RPE검색을 이용한 CELP 보코더의 불규칙 코드북 검색

^o김대식, 하정미, 배명진,
송실대학교 정보통신공학과

(*)한기천, (o)김홍재, (o)유희영
(o)한국전자통신연구소 VLSI 구조 연구실

On a Reduction of Codebook Searching Time by using RPE Searching Technique
in the CELP Vocoder

^oD.S. KIM, J.M. HA, M.J. BAE,
Soongsil University

(*)G.C. HAN, (o)J.J. KIM, (o)H.Y. YOO
(o)Dept. of VLSI Structure, ETRI

* 본 연구는 전자통신연구소의 1995년도 수탁과제 연구지원비에 의해 이루어졌습니다. *

ABSTRACT

Code Excited Linear Prediction(CELP) speech coders exhibit good performance at data rates as low as 4800 bps. The major drawback to CELP type coders is their large computational requirements. In this paper, we propose a new codebook search method that preserves the quality of the CELP vocoder with reduced complexity. The basic idea is to restrict the searching range of the random codebook by using a searching technique of the regular pulse excitation(RPE). Applying the proposed method to the CELP vocoder, we can get approximately 48 % complexity reduction in the codebook search.

1. 서 론

음성신호를 메모리에 저장하거나 전송하기 위한 음성 부호화법에는 크게 과행부호화법, 신호원 부호화법, 혼성부호화법 등이 있다. 이들 중에서 혼성부호화법은 신호원부호화의 메모리 효율성과 고품질의 과행부호화법을 결합시킨 것이다. 혼성부호화법에서 포먼트 정보는 보통 선형 예측부호화법으로 부호화하게 되고, 그 나머지 잔여신호를 어떻게 부호화하느냐에 따라 RELP법, VELP법, MPLP법, CELP법 등[1]이 제안되어져 있다. 혼성부호화법들 중에서도 최근 이동통신용 부호화법으로 채택되고 있는 것은 CELP(code excited linear prediction)법이다[2].

그림 1-1은 CELP형 보코더의 한 블록도형을 나타낸다. 포먼트 합성필터는 10차의 LPC전극형 구조가 보통 적용된다.

LPC계수는 양자화시에 왜곡이 크기 때문에 LSP계수로 변환되어 부호화되고, 복호화시에 다시 LPC계수로 변환시킨다. LPC계수들은 보통 20 ms의 한 프레임마다 부호화하고, 인터플레이션을 통해 5 ms의 각 프레임마다 다르게 공급해준다. 여기서의 파라미터는 또한 5ms 부프레임마다 새롭게 바꾸어준다.

진행적인 CELP 보코더는 2개의 여기원을 사용한다. 첫 번째는 장기(피치)에측상태 또는 적응 코드북에서 취하고, 나머지 잔여 여기신호는 불규칙코드북에서 취한다. 저 전송율의 경우에는 코드북의 코드워드는 128개 정도이다. 이들 2개의 여기원은 그들에 해당하는 이득량이 곁해진 다음에 합해진다. 이것은 결합된 여기 시퀀스가 된다. 각 프레임의 여기출력은 다음 부프레임에서 활용할 수 있도록 적응코드북의 정기필터 상태를 바꾸는데 적용된다.

CELP 보코더의 부호화법에서는 포먼트 정보가 걸러지고 남은 잔여신호에 대해 벡터 양자화기법을 적용하여 부호화하기 때문에 잔여 신호성분 전송에 적용되는 데이터는 코드북의 인덱스가 된다. 따라서 전송율을 4.8 kbps이하로 낮출 수 있으며, 이러한 파라미터를 오류 정정용 부가코드와 함께 전송하면 전송감에도 강한 부호화법이 된다. 또한 이 부호화법은 합성에 의한 분석법을 적용하여 최적의 음질이 유지되도록 반복적으로 분석하기 때문에 주어진 저 전송율에서 음질이 우수하다. CELP 보코더는 때면 음성을 합성해서 비교해야 하므로 복잡한 구조를 갖는다. 특히 부호화시에 계산량을 많이 요구하는데, 피치필터의 계수를 구하는 과정과 불규칙 코드북에서 입력 여기신호를 검색하는 과정에서 대부분의 계산시간을 소모하게 된다[9-11].

주로 CELP 부호화법에서는 불규칙 코드북을 사용하여 음

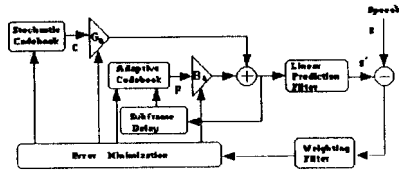


Fig. 1-1 A scheme of CELP speech coder
 그림 1-1. CELP 음성부호화기 구조

성신호의 잔여성분을 부호화하고 있다. 이러한 잔여성분의 부호화에 주로 적용되고 있는 검색법은 코드북 인덱스를 증가 시킴에 따라 원래의 음성신호에 가장 접근하는 합성파형을 찾는 교차 상관관계법을 사용하고 있다. 그러나 이러한 코드북 검색과정은 모든 코드북 인덱스구간에 대해 반복적으로 검색해야 하기 때문에 그 처리시간이 많이 소요되어 기존의 DSP칩으로도 실시간 처리가 어렵다. 또한 보코더의 처리과정이 복잡하면 이에 비례하여 전력소모가 증가하기 때문에 휴대형 송수화기의 전지 사용시간이 감소하게 된다[2][11].

따라서 본 논문에서는 CELP형 보코더의 코드북 검색시에 음질의 열하를 최소화하면서도 코드북 검색시간을 줄일 수 있는 새로운 코드북검색법을 제안하고자 한다.

2. CELP 보코더의 코드북검색

4.8 kbps 보코더에서 각 피치프레임은 두개의 코드북 부프레임으로 구성된다. 각 부프레임에 대해 음성부호화기는 코드북 인덱스 1과 코드북 이득 G를 결정해야만 한다. 이를 구하기 위해 LSP 인터플레이션은 부프레임마다 수행된다.

여기코드북은 단위변동 백색 가우시안 불규칙 시퀀스로 구성된 2^M 개의 코드백터로 이루어진다. $M = 7$ 을 갖는 코드북에는 128개의 내용이 들어있다. 코드북은 각 코드백터가 인접한 코드백터와 한 샘플만 다른 순환구조로 이루어져 있다. 따라서 순환구조의 코드북은 $L_c = 2^M \cdot L \cdot 1$ 길이(여기서 L은 포먼트필터의 차수)의 라인배열로 저장될 수 있다. 그렇지만 구현을 간단하게 하고 메모리 공간을 효율적으로 하기 위해 순환코드북은 $L_c = 2^M$ 개의 샘플길이로 적용된다.

코드북검색을 위해 평균 자승오차를 구하면 다음과 같다:

$$\begin{aligned} \text{MSE} &= \frac{1}{L_c} \sum_{n=0}^{L_c-1} (x(n) - x'(n))^2 \\ &= \frac{1}{L_c} \sum_{n=0}^{L_c-1} (x(n) - G y(n))^2 \end{aligned} \quad (2-1)$$

여기서 $x(n)$ 은 인자가중화된 음성신호이고, $y(n)$ 은 코드북이득 G=1일 때 코드북 인덱스 1번째의 코드백터로 인자가중화하여 합성한 음성으로 다음과 같다:

$$y(n) = h(n) * C(n), \quad 0 \leq n \leq L_c - 1 \quad (2-2)$$

여기서 $h(n)$ 은 현재의 부프레임에서 성도필터의 임펄스응답이다. 또한 코드북이 순환구조로 되었기 때문에 코드백터는 다음과 같다:

$$\begin{aligned} C_i(n) &= C(\text{mod}(n+1, 2^M)), \quad 0 \leq n \leq L_c - 1 \\ &= 0, \quad \text{그외} \end{aligned} \quad (2-3)$$

여기서 $C(n)$ 은 2^M 개의 내용을 갖는 클립된 가우시안 라인코드백이다.

평균자승오차를 최소화하는 것은 다음식을 최대화하는 것과 같다:

$$E_i = \frac{(E_{xy})^2}{E_{yy}} \quad (2-4)$$

여기서

$$E_{xy} = \sum_{n=0}^{L_c-1} x(n)y(n) \quad (2-5)$$

이고,

$$E_{yy} = \sum_{n=0}^{L_c-1} y(n)y(n) \quad (2-6)$$

이다. 주어진 코드워드 I에 대해 최적의 코드북이득 G는 다음과 같다:

$$G_I = \frac{E_{xy}}{E_{yy}} \quad (2-7)$$

이러한 검색과정은 코드워드 I가 허용되는 모든 범위에 대해 반복해야 하고, G_I 가 최대로 되는 인덱스 I와 코드북이득 G를 선택하게 된다.

코드북 검색과정에서 여기코드북의 순환성질을 이용하면 순환전일부선은 다음과 같이 계산과정을 줄일 수 있다:

$$\begin{aligned} y_i(n) &= h(0) C(\text{mod}(-1, 128)), & n=0 \\ &= y_{i-1}(n-1) * h(n) C(\text{mod}(-1, 128)), & 1 \leq n < N_b \\ &= y_{i-1}(n-1), & N_b \leq n < L_c \end{aligned} \quad (2-8)$$

여기서 L_c 는 부프레임의 길이이고, N_b 는 성도 임펄스응답이 두드러지게 지속되는 구간으로 20샘플 정도이다.

상기의 과정을 수행하여도 40샘플 길이의 부프레임에서 코드북을 검색하려면, 코드백터마다 컨벌루션계산에 20번의 곱셈과 덧셈이, 그리고 E_{xy} 를 계산하는데 40번의 곱셈과 덧셈

이 필요하게 된다. 따라서 부프레임에서 불규칙 코드북검색에 소요되는 연산량은 7680번 정도의 곱셈과 덧셈이 소요되는 많은 계산량이 된다.

3. 기존의 검색시간 단축법

코드북내의 가우시안 값들에 대해서 중심 잘라내기(center clipping)가 수행된다. 그 값은 원래 분산 1의 백색 가우시안 프로세스에서 얻어지고, 1.2 미만의 크기값은 영으로 만든다. 이렇게하면 약 75 % 정도가 영으로 채워지게 되어 임펄스들의 코드북을 생성하게된다. 완전한 가우시안 코드북과 중심-잘라내기를 수행한 코드북을 사용하여 보코디된 음성을 비교 하여보면 주목할만한 음질의 차이를 느끼지는 못하지만, 대신에 75 %값이 영으로 채워져 있게 되면 코드북검색에서 순환 컨벌루션의 실행에 소요되는 곱셈의 수는 1/4 정도로 줄이게 된다[6].

1번째 코드벡터로 합성된 신호 $y(n)$ 와 원래신호 $x(n)$ 사이의 교차상관관계 $E_{xy}(l)$ 는 다음과 같이 정의할 수 있다[2]:

$$\begin{aligned}
 E_{xy}(l) &= \sum_{n=0}^{L_c-1} x(n) y_1(n) \\
 &= \sum_{n=0}^{L_c-1} \left[x(n) \sum_{i=0}^n h(i) C(n-i) \right] \\
 &= \sum_{n=0}^{L_c-1} \left[C(n) \sum_{i=0}^{L_c-1-n} h(i) x(i+n) \right] \quad (3-1)
 \end{aligned}$$

여기서 다음과 같이 $z(n)$ 을 정의하고,

$$z(n) = \sum_{i=0}^{L_c-1-n} h(i) x(i+n), \quad 0 \leq n \leq L_c-1 \quad (3-2)$$

이를 대입하면 $E_{xy}(l)$ 는 다음과 같다:

$$E_{xy}(l) = \sum_{n=0}^{L_c-1} C(n) z(n) \quad (3-3)$$

코드북 검색시에 부프레임에 대해 $z(n)$ 을 한번만 계산하면 되기 때문에 800번의 곱셈-덧셈이 필요하다. 그리고 $E_{xy}(l)$ 의 계산에 영이 아닌 코드벡터 $C(n)$ 의 값에 대해서만 고려하면 코드벡터의 75 %가 영값인 경우에 1280번의 곱셈-덧셈이 소요된다. 따라서 40샘플길이의 부프레임인 경우에 $E_{xy}(l)$ 를 계산하는데는 약 2080번 정도의 곱셈-덧셈으로 단축할 수 있지만, 아직도 계산량이 많은 편이다.

4. RPE검색에 의한 코드북 검색법

13.2 kbps 보코더용 유럽 표준방식으로 알려진 정규펄스 여기방식(regular pulse excitation, RPE)은 불규칙 코드북 검색법 대신에 잔여신호를 정규펄스의 시퀀스로 고려하고 있다 [3]. 부프레임에서 원래음성에 대한 40개의 잔여펄스는 1/3로 테시메이션되어 13개의 펄스시퀀스가 APCM에 의해 3비트로 양자화되고, 1/3로 테시메이션할 때에 0-3까지 쉬프트시키면서 두드러진 에너지를 갖는 최적의 펄스위치를 찾게 된다. 이러한 정규펄스 여기법은 처리가 간단한 반면 부호화해야할 성분들이 많아져서, 8 kbps 이하에서는 잘 적용되지 않고 있다.

본 논문에서는 불규칙코드북의 저전송율과 RPE방식의 고속 검색성을 함께 활용한다. 즉, 불규칙 코드북의 구조는 그대로 사용하고, 코드북검색시에는 정규 펄스검색에서 사용하였던 잔여펄스 테시메이션법을 사용한다. 먼저 주어진 부프레임 구간의 원래음성에 대해 포만필터와 피치필터를 통과시켜 잔여신호 $e(n)$ 을 만든다. 잔여신호를 1/3의 정규펄스이기 (RPE) 간격이 되도록 테시메이션을 수행하여, 이 시퀀스 길이가 14, 13, 13이 되도록 3가지를 만든다. 잔여신호를 다시 시퀀스의 길이가 13이 되도록 다음과 같이 4가지로 나눈다:

$$e_m(i) = e(k+m+3*i), \quad i=0, \dots, 12 \quad m=0, \dots, 3 \quad (4-1)$$

여기서 m 은 테시메이션의 시점을 나타낸다. 4가지의 잔여신호에 대해 다음과 같이 에너지를 계산하고 최대값이 얻어지는 테시메이션 시점 m 을 선택하여 최적의 테시메이션된 잔여신호를 구하게 된다:

$$E(m) = \max_m \sum_{i=0}^{12} e_m^2(i), \quad m=0, \dots, 3 \quad (4-2)$$

이렇게 하여 찾은 잔여신호를 사용하여 합성한 음성신호를 기준으로 하여 불규칙 코드북에서 코드벡터를 3씩 건너뛰면

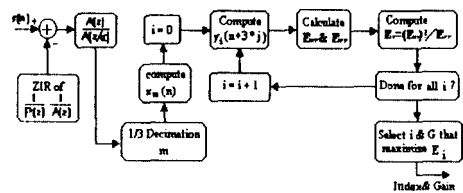


그림 4-1. 본 논문에서 제안한 코드북 검색과정
Fig. 4-1 The pitch search algorithm proposed in this paper.

서 합성한 신호와 인지가증화된 편차로 비교를 수행하게 된다. 이러한 과정은 코드백터의 증가를 3배한다는 점을 제외하고는 기존의 불규칙 코드북검색법과 동일하게 된다. 이상의 데시메이션 과정을 통해 코드북을 검색하면 개선과정이 기존의 감속임고리품에 비해 이론적으로 약 67 % 정도 감소하게 된다.

5. 실험 및 결과

컴퓨터 시뮬레이션에 이용한 장비는 IBM-PC /486 DX2 (50)시스템이며 여기에 음성신호를 입출력하기 위한 상용화된 16비트 AD/DA변환기를 인터페이스하여 8 kHz의 표본율로 데이터를 입력하였다. 각 시료에 대해 한 프레임의 길이를 160표본으로 하여 부프레임 길이를 40표본 단위로 처리하였다. 처리결과와 성능을 측정하기 위해 다음의 대표적인 문장을 20세의 여성, 25세의 남성 및 여성, 32세의 남성 화자가 각 5번씩 발성하여 시료로 사용하였다:

- 발성1: /인수네 꼬마는 천계소년을 좋아한다./
- 발성2: /예수님께서 천지창조의 교훈을 말씀하셨다./
- 발성3: /승실대 정보통신공학과 음성처리 연구실이다./
- 발성4: /창공을 헤쳐나가는 인간의 도전은 끝이 없다./
- 발성5: /공일이삼사오육칠팔구/

코드북검색은 그림 4-1과 같이 일반적인 CELP 보코더의 불규칙 코드북검색 과정을 C-언어로 구현하여 수행하였다. 코드북검색의 성능을 비교하기 위해 컴퓨터 시뮬레이션 과정을 크게 두 가지로 나누어 수행하였다. 우선 코드북검색 구간(코드북 인덱스 0에서 127까지)에 대해 인덱스 1를 1씩 증가하여 검색하는 순차(full) 코드북검색법을 수행하였다. 두 번째 과정은 본 논문에서 제안한 방법으로써, 주어진 부프레임에 대해 잔여신호를 1/3로 데시메이션하여 이를 잔여 여기신호로 합성한 신호를 기준으로 코드북검색을 3씩 건너뛰는 검색법을 적용하였다.

두 처리과정의 코드북검색 시간차를 구하기 위해 상기 발성들에 대해 1초 단위의 평균 검색시간을 구하여 표 5-1과 표 5-2에 나타내었다. 기존의 순차 코드북검색법은 평균 9.32초가 소요되었고 제안한 방법으로 평균 4.82초가 소요되어 약 48.3 %의 시간절약이 이루어 졌다. 여기서 시간 측정 치는 컴퓨터의 기종에 따라 다르기 때문에 상대적인 시간 단축을 평가에 고려하였다. 한편 제안한 검색법은 무잡음 상태에서 순차 코드북검색에 비해 객관적인 코드북검색 이득이 11.27 dB에서 10.75 dB로 낮아져서 -0.52 dB 정도로 열화되었으나, 주관적으로는 음질의 차이를 느끼지 못하였다.

Table 5-1. Performance evaluation of prediction gain

Method \ SNR(dB)	Clean	20	0
Full search	11.27	10.13	5.61
Proposed search	10.75	9.35	4.47
Degradation	-0.52	0.78	-1.14

Table 5-2. Searching time of random codebook

Method \ Time	Average Searching Time
Full search	9.32 sec
Proposed search	4.82 sec

6. 결 론

CELP 부호화기는 음성신호를 합성하여 입력 음성신호와 비교하는 합성에 의한 분기법을 사용함으로써 음질이 매우 우수하다. 그렇지만 방대한 계산량 때문에 기존의 DSP칩으로도 실시간 구현에 어려움이 있다. CELP 보코더에서 불규칙 코드북검색 시간은 부호화시간의 50 % 정도를 차지하고 있다. 따라서 우리는 CELP 보코더의 불규칙 코드북 검색시간을 개선하는 코드북검색법 하나를 새로이 제안하였다.

본 논문에서 코드북은 지층음에서도 유질이 우수한 불규칙 코드북을 적용하고, 검색법은 현재 13.2 kb/s 유희표준 보코더법으로 채택하고 있는 정규펄스여기(RPE)법을 적용하였다. 주어진 부프레임에서 포먼트필터와 미치필터를 통기한 원래음성의 잔여신호에 대해 정규펄스여기법으로 1/3로 데시메이션하여 여기시킨 합성신호를 기준신호로 사용하였다. 따라서 코드북 검색구조는 1/3정도로 간단하게 된다.

실제음성에 대해 제안한 코드북검색을 수행하였을 때, 순차 코드북검색에 비해 객관적인 예측이득이 평균 -0.52 dB로 저하되었으나, 검색시간은 순차 코드북검색법에 비해 약 48 % 정도가 단축되었다.

참 고 문 헌

- [1] L.R. Rabiner and R.W. Schafer, *Digital Processing of Speech Signal*, Prentice-Hall, 1978.
- [2] A.M. Kondoz, *Digital Speech*, John Wiley & Sons, 1994.
- [3] E.F. Deprettere and P. Kroon, "Regular Excitation Reduction for Effective and Efficient LP-Coding of Speech", *IEEE, Proc. Int. Conf. on Acoustics, Speech and Signal Processing*, pp.965-968, 1985.
- [4] G. Davison and A. Gersho, "Complexity Reduction Methods

- for Vector Excitation Coding", Proc. Int. Conf. on Acoustics, Speech and Signal Processing, 1986.
- [5] A.L. Guyader, D. Massaloux, and J. P. Petit, "Robust and Fast Code-Excited Linear Predictive Coding of Speech Signals", Proc. Int. Conf. on Acoustics, Speech and Signal Processing, 1989.
- [6] J.P. Campbell, Jr., V.C. Welch, and T.E. Tremaino, "An Expandable Error-protected 4800 bps CELP Coder(U.S.Federal Standard 4800 bps Voice Coder)", Proc. Int. Conf. on Acoustics, Speech and Signal Processing, 1989.
- [7] W.B. Kleijn et al., "fast Methods for the CELP Speech Coding Algorithm", IEEE Trans., Acoustics, Speech and Signal Processing, Vol.38, No.8, pp.1330-1341, Aug., 1990.
- [8] 강 진덕, 배 성근, 배 명진, 김 형래, 한 기천, 변 경진, 유 하영, "Positive ACF를 이용한 CELP 코덱의 피처검색구간 단축에 관한 연구", Proc. of KITE Summer Conference '94, pp291-294, VOL.17. NO.1, 1994.
- [9] D.S. KIM, M.J. BAE, J.J. KIM, K.J. BYUN, K.C. HAN, H.Y. YOO, "On a Reduction of Pitch Searching Time by Preliminary Pitch in the CELP Vocoder," J. Acoust. Soc. Korea, Vol.13, No.2E, pp.51-57, Jul. 1994.
- [10] D.S. KIM, M.J. BAE, J.J. KIM, K.J. BYUN, K.C. HAN, H.Y. YOO, "On a Reduction of Pitch Searching Time by Preprocessing in the CELP Vocoder", J., Acoust., Soc., Korea, Vol.13, No.3, pp.33-40, Jun.,1994.
- [11] H.Y. Yoo, J.J. Kim, K.J. Byun, K.C. Han, D.K. Kim, J.S. Kim, H.B. Lee, and M.J. Bae, "Efficient DSP Design for Vocoder Application," IEEE Electron Devices Society, Proceeding of IEEE 1995 CICC, pp.189-192, May 1-5, 1994.