

저전송율 보코더의 성능개선에 관한 연구

On an Improving Performance of Low Bit-Rate Speech Coder

박 영 호*, 홍 성 훈**, 배 명 진**

(Young Ho Park*, Seong Hoon Hong**, Myung Jin Bac**)

요 약

본 논문에서는 잔차신호를 모델링하기 위해 사용되는 동적희박대수코드북에 대해 분석하고 성능이 향상된 새로운 대수코드북 구조 및 검색과정을 제안하였다. 제안된 알고리즘은 대수 코드북의 단점인 계산량의 증가 없이 개선시켰다. 먼저 기존에 단순히 부호비트만을 검색하는 것에 대해 다양한 펄스 진폭의 선택을 가능하게 하였다. 그리고 동일 트랙상에서 두 펄스를 선택하게 하였으며 추가 계산량이 필요 없는 무성음에서 유성음으로의 천이 구간 검출기를 이용하여 LSF 보간 시 발생하는 천이구간에서의 LP지연을 최소화하였다. 제안된 알고리즘을 이용한 5.6kbps 음성부호화기는 전화선상의 음질을 시료로 하여 주관적 음질면에서 6.3kbps MP-MLQ와 동등하였으며 MNRU Q = 15dB에서는 MP-MLQ에 비해 약간의 음질 열화가 발생하였다.

ABSTRACT

This paper analyzes the dynamic sparse algebraic codebook used to model a residual signal and proposes a new algebraic codebook structure as well as a searching process with improved performance. The proposed algorithm improves the disadvantage of algebraic codebook without increased computation. First, this paper makes it possible to select various pulse amplitudes differently from the conventional method which looks up the sign bit simply. In addition, two pulses are made to be selected on the same track and the delay of LP is made to be minimized at the transient period occurred when LSF interpolation is performed using a transient detector from unvoiced to voiced without requiring additional computation. For speech quality on the telephone line 5.6kbps speech coder using the proposed algorithm was equivalent to the 6.3kbps MP-MLQ in the viewpoint of subjective speech quality. However, speech degradation was caused a little compared to the MP-MLQ where MNRU Q = 15dB.

I. 서 론

최근에 디지털 이동통신 및 개인통신망의 발달로 증가하는 이동통신 가입자를 보다 많이 수용하고 서비스의 질을 높이기 위해 음성부호화기에 대한 많은 연구가 진행되고 있다. 이 중 가입자의 수용율을 증가시키는데 있어서 효과적인 방법 중의 하나는 부호화기의 전송율을 낮추는 방법이다. 이론적으로 디지털 셀룰라 망에서 보코더의 전송율이 1/2로 낮아질 경우 기존 대비 약 2배의 가입자를 수용할 수 있다고 알려져 있다. 그런 목적으로 현재 최고 600bps대까지 보코더의 전송율이 낮아지고 있다. 그러나 전송율 감소와 그에 따른 음질의 저하는 가입자로 하여금 통신 서비스 질에 대한 불만을 불러일으킨다. 따라서 전송율과 음질이라는 두 지표를 적정수준으로 맞추는데 있어 이동통신사업자의 어려움이 예상된다.

현재까지 발표된 음성부호화기 중 가장 많은 연구가 이루어지고 있는 방식이 CELP(Code Excited Linear Prediction) 방식이다. 이 방식은 4.8kbps내외의 전송율에서 양호한 음질을 얻을 수 있으며 ITU-T, TTA/EIA 등 여러 국제 표준화 기구를 통해 다양한 응용분야에서 표준화가 이루어지고 있다. 특히 국내에서는 PCS의 출현으로 CELP 부호화기에 대해 많은 관심이 집중되고 있다[1].

현재 국내에서 디지털 셀룰라 서비스에 사용되고 있는 보코더로는 Qualcomm사의 IS-96 QCELP, IS-127 EVRC [7]가 있다. 이들 두 부호화기의 구조적인 차이로는 IS-127이 부호화단에 모토롤라 소유권의 노이즈 억압 방식을 채용하고 있으며 Relaxation CELP의 피치 contour, 대수 CELP의 동적 희박 대수 코드북(dynamic sparse algebraic codebook)을 적용했다는 차이가 있다. 또한 유선망을 통한 화상회의를 목적으로 ITU-T에서 1996년에 표준화한 G.723.1 Dual-rate speech coder[3] 역시 대수 코드북(5.3kbps)을 사용하며 인터넷 폰, 화상회의에 응용하기 위해 국내에서 많은 연구가 이루어지고 있다[9][15][16][17]. 잔

* 서울이동통신 중앙연구소

** 숭실대학교 정보통신공학과

접수일자: 1998년 9월 21일

차 신호를 모델링하는 데 기존의 통계적 방식대신 대수적 방식을 사용하는 이들 부호화기의 간차코드북은 음성 코딩알고리즘의 성능지표중의 하나인 메모리요구량을 현저히 줄일 수 있으며 검색속도 역시 통계적 방식에 비해 향상시킬 수 있다. 더구나 DFS(Depth-First Search)검색방식은 기존의 Fast Search Nested Loop 검색법에 비해 약간의 음점저하(G.729A의 경우 SNR:약 0.2dB감소)로 검색시간을 현저히 향상시킬 수 있는 방식이다. 그러나 대수적 코드북의 단점으로는 동일 트랙상에서 단지 하나의 펄스만이 선택 가능하다는 점, 각 펄스위치에서 진폭 검색시에 복잡도를 줄이기위해 ± 1 만의 제한된 부호비트로 코드어의 진폭 다양성을 제한한다는 점이 있다.

따라서 본 논문에서는 이 같은 문제점을 효과적으로 극복할 수 있는 방식을 제안하고 그 성능을 모의실험을 통해 검증해 보고자 한다.

본 논문의 구성은 다음과 같다. II장에서는 대수적 코드북 검색 법 중 널리 사용되는 두가지 방식의 검색법에 대해 설명하고 III장에서는 LSF보간과 제안하는 동적 최적 대수 코드북의 구조와 기존 검색에 비해 개선된 방식에 대해 논한다. IV장에서는 개선된 방식에 대해 음질평가 및 결과를 고찰하고 V장에서 결론을 맺는다.

II. 일반적인 대수 코드북 검색법

Adoul et. al.에 의해 제안된 대수 검색법은 통계적 방식의 잔차신호 모델링법보다 적은 메모리 사용량, 검색시간 단축으로 널리 사용되고 있다. 특히 G.729A, G.723.1, EVRC에 적용되었으며 DFS의 대수적 코드북이 ITU-T에 제출된 여러 4kbps 보코더 [5][6] 및 6.3kbps CS-ACELP, 11.8kbps CS-ACELP, IS-641등에 적용되고 있다. 본 장에서는 기존의 부호화기에 적용되는 대수적 코드북의 검색법 중 Fast Search Nested-Loop법과 DFS법에 대해 논한다.

2.1 일반적인 대수 코드북 구조 및 검색법

고정 코드북은 가중된 입력음성과 가중화된 합성음 사이의 MSE(Mean Squared Error)를 최소로 하는 펄스위치를 선택한다. 코드어를 결정하는 변수가 펄스의 위치와 이득이므로 이득을 일정하게 한 후 각 펄스위치에서 오차를 구하여 위치를 결정한 후 이득을 구한다. 식(1)에 이에 대한 수식적인 표현을 나타내었다. 여기서 S는 가중된 입력음성, S'는 가중된 합성음성, g는 이득이다.

$$\begin{aligned} \min_k \|\mathcal{O}\|^2 &= \min_k \|S - S'\|^2 = \min_k \|S - [P - gA_k H^T]\|^2 \\ &= \min_k \|X - gA_k H^T\|^2 \end{aligned} \quad (1)$$

식(1)을 다시 쓰면

$$\min_k \|\mathcal{O}\|^2 = \min_k \left\{ \|X\|^2 - \frac{(X(A_k H^T)^T)^2}{\|A_k H^T\|^2} \right\} \quad (2)$$

이다.

여기서

$$H = \begin{bmatrix} h(0) & 0 & 0 & \dots & 0 \\ h(1) & h(0) & 0 & \dots & 0 \\ h(2) & h(1) & h(0) & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ h(L-1) & h(L-2) & h(L-3) & \dots & h(0) \end{bmatrix} \quad (3)$$

이고 H는 L×L 하삼각 Toeplitz행렬이고 h(n)의 임펄스 응답으로 형성된 행렬이다. A_k는 인덱스 k의 코드벡터, X는 목적벡터이다. 식(2)의 오른쪽 두번째 항이 최대가 될 때 식(1)을 만족하므로

$$\begin{aligned} \max_k \frac{(X(A_k H^T)^T)^2}{\|A_k H^T\|^2} &= \max_k \frac{((X H^T) A_k^T)^2}{\alpha_k^2} \\ &= \max_k \frac{(D A_k^T)^2}{\alpha_k^2} \end{aligned} \quad (4)$$

과 같이 쓸 수 있다.

여기서 D는 역필터링(backward filtering)된 목적벡터이다. 식(1)을 만족하기 위해서는 최적의 펄스위치와 최적 펄스진폭이 필요한데 두 인자를 발견하기 위해서는 기본 판별식(criterion)이 필요하다. N개의 영이 아닌 펄스를 가지는 코드벡터의 집합 중 최적의 코드벡터를 결정하는 기본적인 판별식은 다음과 같이 나타낸다.

$$k = \max_k^{-1} [\Theta_k[N]] \quad (5)$$

여기서

$$\Theta_k[N] = \left[\frac{(D A_k^T)^2}{\alpha_k^2} \right] \text{ 이다.}$$

그리고 판별식의 분모는

$$\alpha_k^2 = \sum_{n=0}^N S_n^2 \mathcal{O}(P_n, P_n) + 2 \sum_{i=1}^N \sum_{j=i+1}^N S_n S_m \mathcal{O}(P_i, P_j) \quad (6)$$

분자는

$$D A_k^T = \sum D_n S_n \text{ 이며}$$

여기서 $\mathcal{O}(\cdot)$ 는 식(7)로 표현되는 h(n)의 자기상관식이다.

$$\mathcal{O}(i, j) = \sum_{k=0}^L h(k-i+1)h(k-j+1) \quad (7)$$

S_n S_m는 미리 설정된 펄스의 부호이다.

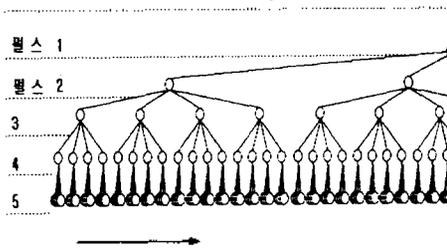


그림 1. 일반적인 완전검색법

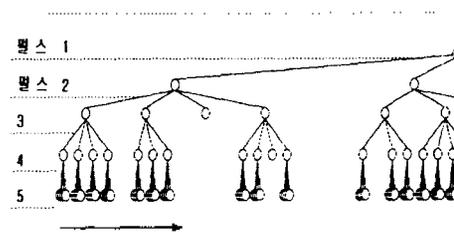


그림 2. Fast Search Nested-Loop 검색법

2.2 Fast Search Nested-Loop 검색법(3)(7)(8)(13)

식(1)을 만족하는 최적의 펄스위치를 검색하기 전에 먼저 코드북을 구성하는 테이블을 결정해야 한다. 특히 대수 코드북은 트랙내의 구성샘플위치에 따라 음질차이가 발생하므로 일반적으로 각 트랙마다 이전 트랙과의 한 샘플 위치씩 차이를 주는 구조를 가지고 있다. 그러나 부프레임 길이가 길거나 트랙당 전송할 펄스의 갯수가 많을 경우 G.723과 같이 적수위치/홀수위치 테이블을 따로 만들어 선택된 비트와 Grid 한비트를 함께 전송하는 경우도 있다. 멀티펄스방식에서 펄스를 선택할 위치와 진폭을 제한한 형식의 대수코드북은 ISPP(Interleaved Single Pulse Permutation)를 이용하여 구성되며 ISPP(L, E)으로 표현한다. 여기서 L 는 부프레임의 길이, E 는 트랙의 수이다. 예로 ISPP(40,5)는 5개의 트랙에 각 샘플이 8개씩 할당되어 $5 \times 3 = 15$ 의 위치비트와 5개의 부호비트로 전체 20비트가 코더어를 나타내는데 필요하다. 표 1은 ISPP(40,5)의 특별한 경우이다.

표 1. 8kbps CS-ACELP 코드북[4]

펄스	부호	위치
I_0	± 1	0,5,10,15,20,25,30,35
I_1	± 1	1,6,11,16,21,26,31,36
I_2	± 1	2,7,12,17,22,27,32,37
I_3	± 1	3, 8, 13, 18, 23, 28, 33, 38 4, 9, 14, 19, 24, 29, 34, 39

ISPP를 이용하여 구성된 코드북을 검색하기 위해서는 식(5)의 판별식을 이용하여야 하며 분자보다 분모를 계산하는데 많은 계산량이 요구되어진다. 따라서 검색법의 초점은 분모의 계산량을 줄이는데 있다. Salami는 많은 계산

량을 요구하는 식(6)을 계산하는 방식을 사용하였다. 일명 Focused 검색이라고 하는 이 검색법은 마지막 펄스 검색 전 N-1번째까지 펄스 검색을 행한 후 미리 결정된 문턱값을 넘는 경우에만 마지막 N번째 펄스를 검색하는 검색법이다.

그림 1은 일반적인 대수 코드북의 완전검색을 나타내고 있다. 모든 하부 트리에 대해 검색을 하고 있으며 선택될 확률이 적은 트리구조까지 검색을 행하고 있다. 반면 Focused 검색을 하는 그림 2는 두 개의 문턱값을 설정하여 불필요한 하부트리를 검색하는 단점을 개선하였다.

2.3 Depth First 검색법(2)

Adoul에 의해 제안된 Depth First 검색법은 Fast Search Nested-Loop법에 비해 약간의 음질을 희생하며(G.729A의 경우 SNR : 0.2dB감소) 계산량을 감소시킨 방식이다. N_p 개의 여기펄스는 N_m 개의 펄스로 구성된 M 부분집합으로 나누어진다. 검색은 부분집합 #1에서부터 시작되며 트리 구조로 검색을 실시하므로 다음 부분 집합이 검색되며 검색은 위치트랙안에 할당된 펄스순서를 변경하면서 반복적으로 검색된다. 먼저 펄스는 두 펄스쌍의 부분집합안에서 검색된다. 트랙 T_{k0} 와 $T_{(k0+1)modL}$ 트랙의 개수로 구성된 쌍 (i_0, i_1)에 대해 최적의 펄스위치를 검색한다. 그다음에 $T_{(k0+2)modL}$ 트랙의 개수와 $T_{(k0+3)modL}$ 트랙의 개수안의 펄스들의 가능한 위치결합 (i_2, i_3)에 대해 검색한다. 같은 과정을 (i_4, i_5), (i_6, i_7)... 등에 대해 수행한다. 각 레벨에서 검색 판별식은 그 레벨에서 이용 가능한 펄스에 대해서만 수행한다. 트랙에 할당된 펄스의 수를 변경하여 N_k 번의 반복을 수행한다. 그림 3은 검색에 대한 구조를 나타내고 있다. Focused 검색법과는 다르게 처음 부분집합부터 후보펄스를 선택하기 때문에 다른 검색이 가능하나 잘못된 패스 선택은 음질의 저하를 가져올 수 있다. 계산량적인 면에서 부프레임 길이가 40, 한 트랙안에 펄스위치의 수가 8, 트랙이 5개인 경우 3번의 트랙안에 펄스 위치를 변경하여 수행 할 경우 $3 \times 4 \times 8 \times 8 = 768$ 번을 수행한다. 전체 검색에 비해 2.34%의 검색만을 필요로 하는 계산량이다.

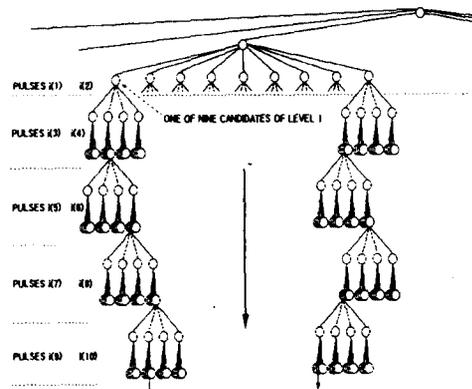


그림 3. Depth First Search 검색법

표 2. 6.4kbps CS-ACELP 트랙 테이블

펄스	부호	위치
i_0	± 1	1, 3, 6, 8, 11, 13, 16, 18, 21, 23, 26, 28, 31, 33, 36, 38
i_1	± 1	0, 1, 2, 4, 5, 6, 7, 9, 10, 11, 12, 14, 15, 16, 17, 19, 20, 21, 22, 24, 25, 26, 27, 29, 30, 31, 32, 34, 35, 36, 37, 39

III. 개선된 알고리즘

3.1 유/무성음 분리에 의한 LSF보간

CELP형의 보코더는 20~30ms의 프레임 구간에 대해 LSF계수를 구한 후 인접 프레임간의 불연속에 의한 음질 저하를 최소화하기 위해 부프레임 단위로 보간을 행한다. 그러나 음의 상태에 따른 구분없이 일률적인 인접 프레임간 보간으로 무성음에서 유성음으로의 천이구간에서 원치 않는 LP값의 지연이 발생한다. 그 결과 재생음에서 유성음 시작시 왜곡발생의 원인이 된다. 따라서 무성음에서 유성음으로의 천이구간을 효율적으로 검출하는 알고리즘이 필요하다. 본 논문에서는 Durbin 알고리즘 동작시 무가적으로 발생하는 LAR(Log Area Ratio)계수값을 가지고 별도의 추가 계산량을 필요치 않는 간단한 무성음에서 유성음으로의 천이시점을 검출하였다. LAR(1), LAR(2)의 문턱값을 결정하기 위해 영교차율, 에너지 분포를 가지고 실험적으로 구간 분리하였다. 그 다음 그 결과를 기준값으로 사용하여 두개의 LAR 변수의 문턱값을 결정하였다.

구하여진 무성음에서 유성음으로의 천이 문턱값을 사용하여 천이시에 결정논리에 따른 프레임간 보간으로 프레임간 보간을 하였다.

3.2 개선된 코드북 검색법

1장에서 설명한 바와 같이 대수 코드북은 멀티펄스구조에 대해 트랙상에서의 일장간격의 펄스위치와 제한된 진폭으로 복잡도를 감소시킨 방식이다. 기존의 대수코드북의 구조를 보면 다음과 같은 문제점을 발견할 수 있다. 첫째, 멀티펄스에 대해 가능성이 있는 일부 펄스위치만을 검색하는 부분적인 하이퍼큐브방식을 사용한다. 둘째, 각 트랙에서 단지 하나의 펄스만을 검색하므로 동일 트랙상에 다른 위치에 있는 두 펄스가 최적의 값을 갖더라도 준최적만을 이루는 펄스를 검색한다. 셋째, 발견된 모든 펄스가 일정한 진폭을 가지고 있다고 가정하고 ± 1 이라는 규준화된 진폭만을 검색한다. 따라서 본 논문에서는 이를 개선하기 위하여 표 3과 같은 코드북을 구성하였다. 검색 방법은 DFS검색을 사용한다. 첫번째와 두번째 문제를 해결하기 위하여 트랙 1, 2에서는 한 펄스를 각 트랙에서 검색하여 발견하며 트랙3에서는 두개의 펄스를 검색한다. 이렇게 하는 이유는 위에 설명한 바와 같이 트랙3에서 다른 위치에 있는 펄스가 최적으로 판단되었을 경우 검색 불가능한 것을 방지하기 위함이다. 또한 동일 위치에 두 펄스가 최적이라 판단되었을 경우 진폭부호는 두 펄스의 합인 진폭 ± 2 를 전송한다.

부호비트의 측면에서 한 트랙에서 두개의 펄스를 발견

할 경우 부호 1비트는 잉여성분으로 볼 수 있다[2]. 즉, 펄스a, b가 서로 다른 부호를 가지고 있다 할지라도 절대값으로 보아 두 펄스의 위치는 교환이 가능하다. 만약 두 펄스의 부호가 같다면($S_1 = S_2$) 낮은 위치의 펄스는 P_a 에, 높은 위치의 펄스는 P_b 에 둔다. 그리고 두 부호가 같지 않다면 낮은 위치의 펄스는 P_b 에, 높은 위치의 펄스는 P_a 에 둔다. 따라서 두 펄스의 위치와 S_1 의 부호만 전송하면 수신단에서 두 펄스의 위치 및 부호를 복원할 수 있다. 그림 4에 이에 대한 흐름도를 나타내었다.

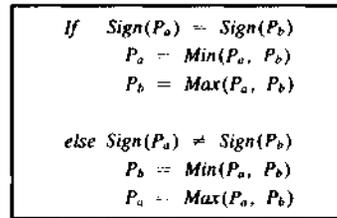


그림 4. 잉여비트 제거 알고리즘

세번째 문제를 해결하기 위하여 다음과 같은 방식을 제안하였다. 기존 대수적 코드북의 단점은 펄스진폭비트를 +1의 부호성분만 나타내어 다양한 펄스 진폭 변화를 나타낼 수 없었다. 또한 각 펄스마다 N의 진폭변화를 설정할 때 트랙이 L개인 경우 \sum 펄스위치 + (N×L)만큼의 비트가 필요하다. 즉, 코드북 크기의 급격한 증가로 검색 시 많은 계산량이 필요로 된다.

이를 개선하기 위해서 먼저 부프레임내에서의 펄스진폭을 미리 설정한다. 식(8)로 진폭을 결정하기 위한 목적진폭신호를 만든다. 식(8)은 펄스설정을 위해 목적벡터와 피치성분을 제거한 원음성 그리고 LPC성분을 제거한 원음성과의 적절한 혼합비율로 구성되어있다. 여기서 $\theta_1, \theta_2, \theta_3$ 는 0과 1사이의 값이어야 하며 본 논문에서는 $\theta_1 = \theta_2 = 0.5, \theta_3 = 0.1$ 을 사용하였다. 그 결과 추출된 펄스 진폭은 부프레임 구간내에서 식(9)을 이용하여 정규화 한다.

$$Y = \theta_1 \frac{D}{\|D\|} + \theta_2 \frac{P}{\|P\|} + \theta_3 \frac{S}{\|S\|} \tag{8}$$

$$\Delta = Q \left(\frac{V_p}{\max_k |V_x|} \right) \tag{9}$$

여기서 $\max |V_x|$ 는 영이 아닌 진폭펄스의 최대치를 나타내는 정규화 인자이고 $Q(\cdot)$ 는 양자화 함수이다.

본 논문에서 사용하는 진폭의 경우 $\pm 1/\pm 0.5$ 크기의 진폭을 사용하므로 트랙2의 경우 선택된 진폭에서 부호 비트만을 사용한다. 트랙 1, 2에서는 Δ 함수가 미리 설정된 문턱값에 따라 $\pm 1/\pm 0.5$ 의 진폭을 결정한다. 검색과정은 기존의 Depth-First Search법과 동일하다. 단, 행렬 $\Phi(i,j)$ 를 계산하는데 제한된 알고리즘으로 미리 계산되어진 진폭을 식(10)과 같이 결합시킨다. 이렇게 함으로써 진폭이 q개의 경우의 수를 가지더라도 진폭을 검색하기 위한 계산량의 증가는 발생하지 않는다.

$$\Phi'(i, j) = S_i S_j \Phi(i, j)$$

(10)

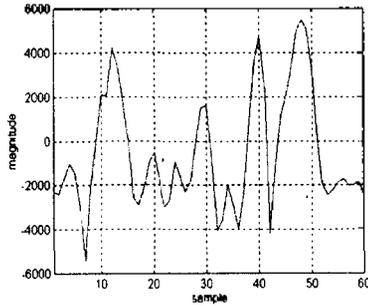


그림 5. 목적벡터 파형 (부 프레임 : 60샘플)

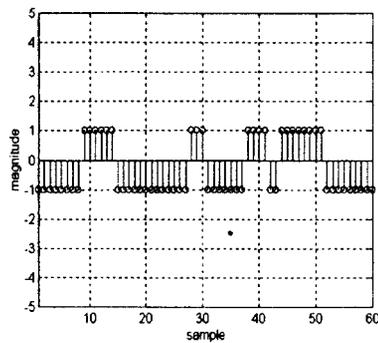


그림 6. 기존방식의 목적벡터 진폭추출

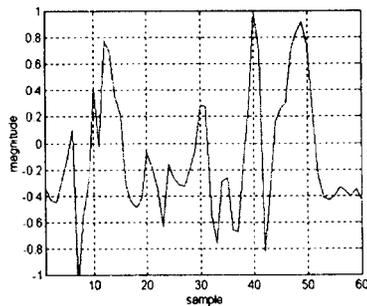


그림 7. 제안한 방식을 이용한 진폭추출 목적벡터

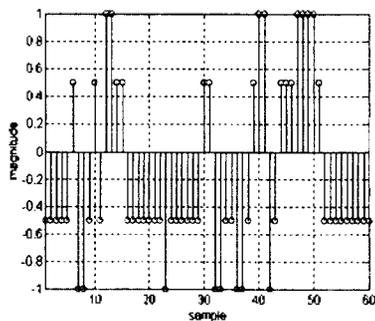


그림 8. 제안한 방식을 이용해 추출한 진폭파형

표 3. 본 논문의 대수적 코드북 구조

트래	부호	위치
T ₀	±1, ±0.5	0, 8, 16, 24, 32, 40, 48, 56
T ₁	±1, ±0.5	2, 10, 18, 26, 34, 42, 50, 58
T ₂	±1	4, 12, 20, 28, 36, 44, 52, (60) 6, 14, 22, 30, 38, 46, 54, (62)

IV. 실험 및 결과

본 논문에서 개선한 알고리즘을 평가하기 위하여 컴퓨터 시뮬레이션을 실행하였다. 이용된 장비로는 IBM Pentium (166MHz)과 음성신호를 입출력 하기 위한 16비트AD/DA 변환기를 인터페이스로 하여 8kHz로 샘플링 하였다.

이상의 방법에 대한 성능을 평가하기 위해 아래와 같은 대표적인 문장을 20대 중, 후반의 남녀화자가 각각 5번씩 발성한 음성을 시료로 사용하였다. 시료는 두드러진 피크를 가지지 않고 노이즈가 30dBA를 가진 방에서 녹음하였다.

1. 인수네 꼬마는 천재소년을 좋아한다.
2. 예수님께서 천지창조의 교훈을 말씀하셨다.
3. 송설대 정보통신과 음성통신 연구실이다.
4. 창공을 헤쳐 나가는 인간의 도전은 끝이 없다.
5. 공일이삼사오육칠팔구

원음성은 일반인을 이용하여 채취하였으며 그 이유는 훈련된 발화자보다는 일반 전화 사용자의 음성을 정확히 반영한다고 볼 수 있기 때문이다. 녹음된 원 시료는 전화기 특성을 나타내기 위하여 전화기 특성을 모델링한 변형된 IRS(Modified-Intermediate Reference System) 가중 필터를 통과시켰다. IRS필터를 통과한 음성신호는 P.56 음성 레벨 측정 도구를 이용하여 -26 ± 1 dBov로 정규화 하였다. 그리고 주관적 음질평가를 하기 위해 MOS(Mean Opinion Score)를 사용하였으며 MOS의 신뢰성을 위해 기준 신호의 사용타입으로 P.81을 준수한 MNRU Ver.2.0를 사용하였다. MNRU의 Q값은 5dB간격으로 입력 음성에 첨가하였다. 본 논문에서는 진폭파형의 양자화 스텝을 네 단계($\pm 1, \pm 0.5$)로 하였으며 식(8)의 결정인자는 $\theta_1 = \theta_2 = 0.5, \theta_3 = 0.1$ 으로 실험적인 값을 사용하였다. 그림 5~8은 제안된 알고리즘의 진폭추출에 대한 기존 방식과의 비교파형이다. 기존 방식에 의한 두 레벨(± 1) 진폭파형을 나타내는 그림 6보다 극소의 추가 계산량으로 네 레벨($\pm 1, \pm 0.5$)에 대해 그림 8과 같이 진폭파형을 추출 하였다. 표 4는 주관적 음질평가를 나타낸다. 비교 대상으로 유사한 전송율을 가진 6.3kbps MP-MLQ와 비교하였다. 수행결과 주관적 평가에서 MP-MLQ와 차이가 없었으나 MNRU Q = 15dB에서 약간의 음질의 저하를 느낄 수 있었다. 그림 9~12는 시료 1의 /인수네 꼬마는 천재소년을 좋아한다/에 대한 MP-MLQ와 제안된 알고리즘을 이용한 부호화기의 스펙트로그램이다. 각각 MNRU Q = 15, 30dB에 대해 나타내었다.

표 4. 주관적 음질평가결과(MNRU: dB)

코덱	MNRU	15	20	25	30
	MP-MLQ		2.88	3.28	3.53
제안된 방식을 이용한 5.6kbps CELP		2.84	3.31	3.54	3.76

표 5. 5.6kbps ACELP의 바트할당

파라미터	부프레임 0/2	부프레임 1/3
LPC코드북	24	
적응코드북	7	2
잔차코드북	20	20
이득코드북	12	12
합계	170	

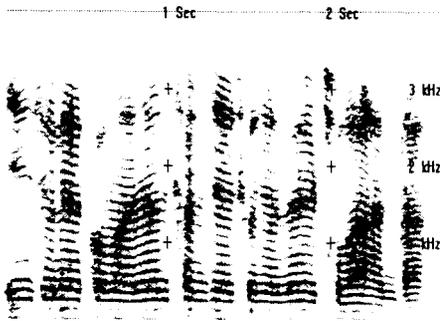


그림 9. 제안된 방식의 부호화기 합성음 (Q = 30 dB) /인수네 꼬마는 천재소년을 좋아한다/

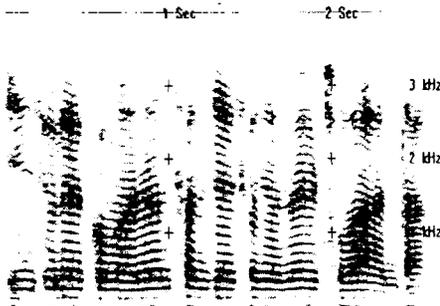


그림 10. MP-MLP 부호화기의 합성음 (Q = 30 dB)

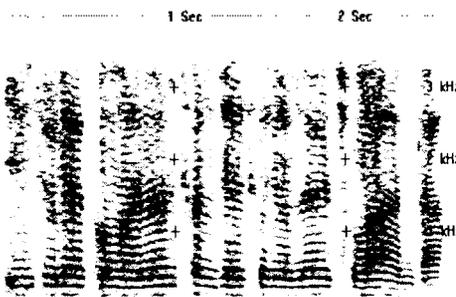


그림 11. 제안된 방식의 부호화기 합성음 (Q = 10 dB) /인수네 꼬마는 천재소년을 좋아한다/

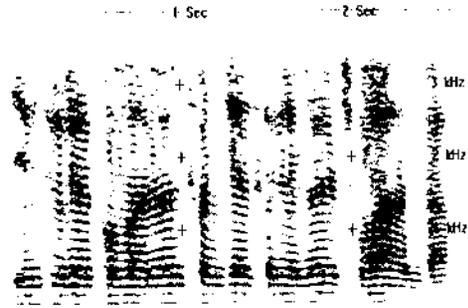


그림 12. MP-MLP 부호화기의 합성음 (Q = 15 dB)

V. 결 론

본 논문에서는 대수적 코드북을 이용한 음성부호화기의 성능을 개선시켰다. 로그영역비를 이용한 무성음에서 유성음으로의 천이검출과 제안된 대수 코드북을 이용하여 펄스진폭의 다양성을 개선하였으며 위치선택에 있어서의 제한을 비트증가없이 완화 시켰다. 구체적으로 먼저 무성음에서 유성음으로의 천이시 발생하는 위치 않는 LP지언을 제거하기위해 무성음에서 유성음의 천이구간에 대해 결정논리에 의한 프레임간 보간을 하였다. 대수 코드북에 시는 코드북의 구조를 개선하여 2개의 펄스를 한 트랙에서 검출하여 대수코드북이 가지는 단점을 보완하였다. 진폭면에서는 펄스진폭의 다양성을 주기위해 네가지 경우에 대한 부호비트를 설정하였다. 결과적으로 코드북의 크기가 증가되었으나 진폭의 사전 설정에 의한 검색법의 단순화로 계산량의 증가없이 성능을 향상시켰다. 제안된 알고리즘에 대한 음질검사를 실시한 결과 주관적 음질에서 MP-MLQ와 차이가 없었으며 MNRU Q = 15dB에서 제안된 알고리즘을 이용한 부호화기의 약간의 음질저하를 느낄 수 있었다.

참 고 문 헌

1. A. M. Kondoz, *Digital Speech - Coding for Low bit rate communication system*, 1994.
2. ITU-T DRAFT, *ITU-T Draft description of Annex D to Recommendation G.729 - 6.4kbps CS-ACELP Speech coding algorithm*, May, 1998.
3. ITU-T Recommendation G.723.1, March, 1996.
4. ITU-T Recommendation G.729, July, 1996.
5. Matsushita Electronics Inc.co., *High Level description of proposed 4kbps Speech coder and Qualification Test Results*, ITU-T SG16 Delayed-133, Feb, 1998.
6. Voxware, Inc., *High Level description of Voxwares 4kbps Candidate CODEC*, ITU-T SG16 Delayed-93, Feb, 1998.
7. TTA/EIA/IS-127, *Enhanced Variable Rate Codec, Speech Service Option 3 for Wideband Spread Spectrum Digital Systems*, 1997.
8. 박영호, 외 4명, "피치동기를 이용한 대수 CELP의 성능개선", 제13회 음성통신 및 신호처리 워크샵, 1996, 8, pp. 388-392.

9. Y.H. Park, J.C. Yang, S.M. Sohn and M.J. Bae, "On a Time Reduction of Pitch Searching by the Regular Pulse Search Technique in the CELP Vocoder," IEEE Comm. Society, Proc. of MILCOM'97, Vol.1, Nov., 1997.
10. C.Laflamme, J-P.Adoul, H.Y.Su, and S.Morissette, "On Reducing Computational Complexity of codebook Search in CELP Coder Through the Use of Algebraic Codes", Proc. ICASSP, pp.177-181, 1990.
11. R. Salami et al., "A Toll Quality 8kbps speech Codec for the Personal Communications System(PCS)", IEEE Trans. Vehic. Tech., vol 43, no 3, Aug. 1994, pp. 806-816.
12. V. K. Varma, "Testing Speech Coders for usage in Wireless Communication system." Proc. IEEE Workshop on speech coding for Telecommunications, Saint-Adele, Quebec, Canada, Oct, pp.13-15, 1993.
13. C.Laflamme, J-P Adoul et. Al., "16KBPS WIDEBAND SPEECH CODING TECHNIQUE BASED ON ALGEBRAIC CELP", Proc. of ICASSP, pp.13-16, 1991.
14. M. R. Schroeder and B. S. Atal, "Code-excited linear prediction : high quality speech at very low bit rates", Proc. IEEE ICASSP-85, pp.937-940, 1985.
15. S.H. Hong, M.J. Bae, K.J. Byun and J.J. Cha, "On a Pitch Search Technique with Correlation Characteristic of Quantization Error in CELP Type Vocoder," The Sixth Western Pacific Regional Acoustics Conference, Proceeding of WESTPRAC-VI'97, Vol.1, pp.106-111, Nov., 1997.
16. M.J. BAE, W.C. LEE and S.B. IM, "On a New Speech Coding Technique using the Nonuniform Sampling and the Separated Bandwidth of Speech Signals," Intrenational Symposium on Signal Processing and its Applications, Proceedings of ISSPA'96, Vol.1, pp.13-16, August 26-28, 1996.
17. M.J. BAE and W.R. JO, "On a Fast Pitch Search of CELP Type Vocoder using Decimation Technique," IEEE Region-10, Proceeding of TENCON'96, Vol.1, pp.204-208, Nov. 27-29, 1996.

▲박 영 호 (Young Ho Park) 1970년 12월 26일생



1995년 2월 : 숭실대학교 전자공학과 (학사)
 1997년 2월 : 숭실대학교 전자공학과 (공학석사)
 1996년 11월 ~ 현재 : 서울이동통신 중앙연구소 주임연구원 재직 중

※주관심분야: 음성압축, 실시간 DSP설계

▲홍 성 훈 (Seong Hoon Hong) 1975년 2월 13일생



1997년 2월 : 숭실대학교 정보통신공학과(학사)
 1997년 3월 ~ 현재 : 숭실대학교 정보통신공학과 석사과정 재학중

※주관심분야: 음성부호화, 음성 합성

▲배 명 진 (Myung Jin Bae)

현재 : 숭실대학교 정보통신공학과 교수
 (제14권 1E호 참조)