

VSELP 부호화기의 음질 및 주기탐색 개선에 관한 연구

正會員 成 基 鐵* 正會員 文 相 在*

A Study on Improving Voice Quality and Pitch Searching of the VSELP Coder

Ki-Cheol Sung*, Sang-Jae Moon* *Regular Members*

要 約

본 논문에서는 기존의 VSELP 음성 부호화기의 성능을 개선하는 방법들을 제안하였다. 혼합 방법에 의한 주기추출법을 제안하여 기존 VSELP 음성 부호화기의 피치 주기 탐색시간을 줄였으며, 장기 필터(long term filter)의 출력 여기신호에서의 양자화 오차로 인해 발생할 수 있는 주기추출 오류를 개선시켰다. 또한 주기성을 향상시키고 잡음을 줄이기 위해 본 논문에서 제안한 주기성 향상 필터를 부호기의 여기신호에 적용하였으며, 합성된 음성신호에서 음의 명료도를 높이기 위해 적응 일차 필터를 적용하였다. 상기한 방식과 필터를 적용한 결과 주기탐색시간이 기존에 대해 26%로 줄었으며, 최종합성음의 MOS(mean opinion score) 측정에서 기존 합성음은 3.19를 나타낸데 비해 개선된 합성음의 4.04를 나타내었다.

Abstract

This paper presents methods for improving the performance of the VSELP speech coder. The hybrid method is employed for pitch period searching. Pitch searching time is reduced and pitch detection error, caused by quantization error of excitation signal of encoder in VSELP coder, is reduced by this method. This paper also adopts a pitch period enhancement filter and an adaptive first order filter. In this result, pitch period searching time is reduced to 26%, and MOS of reconstructed speech signal is increased by 3.19 to 4.04.

I. 서 론

*慶北大學校 電子工學科
Dept. of Electronics, Kyungpook Nat'l Univ.
論文番號 : 93257
接受日字 : 1994年 2月 28日

지금까지 개발된 음성 부호화 방식으로는 크게 파형 부호화(waveform coding) 방식과 신호원 부호화(source coding) 방식 그리고 이 둘의 결합 형태인 혼합 부호화(hybrid coding) 방식이 많이 연구되고 있

다. 대표적인 혼합 부호화 방식으로는 CELP(code excited linear predictive)부호화 방식⁽¹⁾이 있으며, 현재 이 CELP부호화 방식은 여러가지의 변형된 형태의 알고리즘으로 연구되고 있다^(2, 3). 그 중에서도 VSELP(vector-sum excited linear predictive)부호화 방식⁽⁴⁻⁶⁾은 북미의 이동통신용 음성 부호화 방식의 표준안⁽⁷⁾으로 채택되었다.

음성신호는 유성음 부분에서 대체로 2.5ms에서 18.25ms의 주기성을 가진다. 주기추출 방법으로는 크게 개루프 방법(open loop method)과 폐루프 방법(closed loop method)이 있으며, 최근에는 이들의 결합 형태인 혼합 방법(hybrid method)이 연구되고 있다⁽⁸⁾. 개루프 방법은 입력음성신호나 오차신호(residual signal)에서 직접 주기를 구한다. 폐루프 방법은 입력음성과 장기 필터 상태(long term filter state)를 이용하여 구한다. VSELP 부호화 방식에서는 폐루프 방법을 이용한다. 폐루프 방법은 합성에 의한 해석(analysis-by-synthesis) 과정에서 모든 주기에 대한 탐색을 필요로 하므로 주기탐색시 많은 시간이 소요된다. 또한 장기 필터의 출력은 가중 합성 필터에서 합성되어 원음과 비교되며, 이때 합성음에는 양자화 오차가 존재하게 된다. 음성신호의 유성음에서 이러한 음성신호사이의 주기는 유사성을 가진다. 그러나 이 양자화 오차는 현재 음성신호의 주기로써 이웃한 음성신호주기와 유사한 주기가 아닌 바람직하지 않은 이웃한 음성신호주기의 2배 주기나 3배 주기를 최적주기로 선택하는 것과 같은 주기추출 오류를 심화시킨다.

VSELP 부호화기의 부호화의 합성음에는 전송 계수(parameter)의 양자화 오차나 주기추출 오류로 인한 잡음 성분이 나타난다. 음성 합성에서 비슷한 주기성을 띠는 인접한 여러개의 부프레임(subframe)에서 가중치를 두어 구한 합성음에서는 음성신호 성분은 보다 주기성이 향상되고 주기성이 없는 잡음 성분은 상대적으로 줄어든다^(12, 13). 이와 같은 효과를 이용하여 VSELP 음성 부호화기의 합성음에 나타나는 잡음 영향을 감소시킬 수 있다. 그러나 이러한 효과는 이용한 합성음에서는 주기성의 평활화(smoothing)로 인해 음의 병료도가 떨어지는 현상이 나타나므로 이것을 적절하게 보상 해 주는 필터가 필요하다.

본 논문에서는 상기된 개념을 바탕으로 주기탐색시 필요한 계산량을 줄여 처리 시간을 단축시키면서 바람직하지 않은 2배 주기나 3배 주기와 같은 주기추출 오류를 줄일 수 있는 새로운 혼합 방법에 의한 주

기추출 알고리즘을 제안한다. 또한 유사한 주기성을 가지는 인접한 여러개의 부프레임을 사용하여 주기성을 향상시키면서 잡음을 줄이는 주기성 향상 필터를 적용한다. 그리고 음의 병료도가 떨어지는 현상을 보상하기 위해 적응 일차 필터를 도입하여 전체적인 음질을 개선시킨다.

II. VSELP 음성 부호화기

이동통신용 디지털 음성 부호화기 표준안으로 북미에서 채택된 VSELP는 CELP 음성 부호화기의 변형으로 8kbps의 비교적 낮은 전송률에서 양호한 음질을 가진다. 그리고 CELP 부호화기가 지닌 코드북 탐색에 많은 계산량을 요구하는 단점을 보완한 독특한 구조의 코드북을 사용하고 있다. VSELP 음성 부호화기는 40개 7조의 기저 벡터(basis vector)를 포함하는 여기 코드북(excitation codebook)을 이용하므로 채널 에러에 강한 특성을 갖는다. 음성 처리는 프레임(frame)단위 혹은 프레임을 4개로 나눈 부프레임 단위로 행해진다.

1. 부호기

그림 1은 8kbps VSELP 부호화기의 부호기를 나타낸다. 부호기는 여기신호를 위한 두개의 코드북과 주기를 위한 하나의 적응 코드북을 이용한다. 먼저 입력된 음성신호를 Chebyshev type II 고역 필터에 통과시킨다. LPC 계수는 FLAT(fixed point covariance lattice) 알고리즘을 적용하여 프레임단위로 계산되며, 매 부프레임에 대해 인접 프레임의 LPC 계수와 보간된 값이 사용된다. 프레임 에너지도 LPC 계수와 마찬가지로 보간된 값이 사용된다. 각각 가중 필터(perceptual weighting filter) $W(z)$ 는 여기신호의 양자화 오차에 대한 청각적 특성 개선을 위해 오차신호 이전에 사용된다. 그리고 각각 가중 필터를 거친 입력음성에서 가중 합성 필터 $H(z)$ 의 영입력 응답이 계산된다. 가중 합성 필터는 각각 가중 필터와 일반적인 LPC 합성 필터⁽¹⁴⁾가 결합된 형태이다.

코드북 탐색은 순차적으로 이루어지는데 먼저 주기를 위한 적응 코드북이 탐색된 후 VSELP 여기 코드북 L, H 가 차례로 탐색된다. 코드북 탐색은 독립적으로 이루지며 나중에 탐색되는 코드북은 먼저 탐색된 코드북에서 선택된 코드벡터와 직교화(orthogonal)된다. 코드북 탐색시 최적 코드벡터는 가중 합성 필터를 거친 여기신호와 가중 합성 필터의 영입력

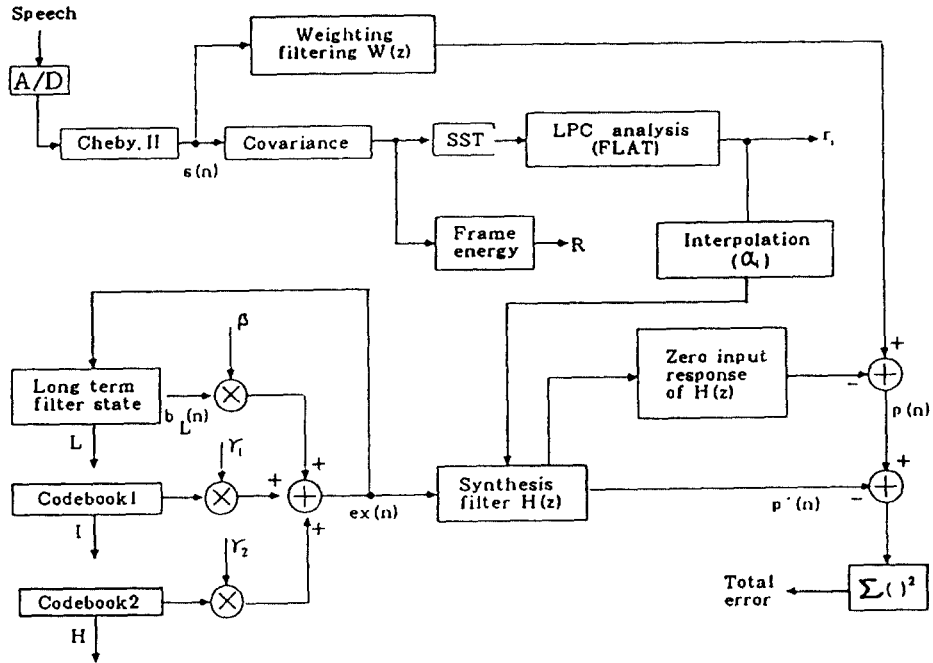


그림 1. VSELP 부호기 선도
Fig 1. VSELP encoder block diagram.

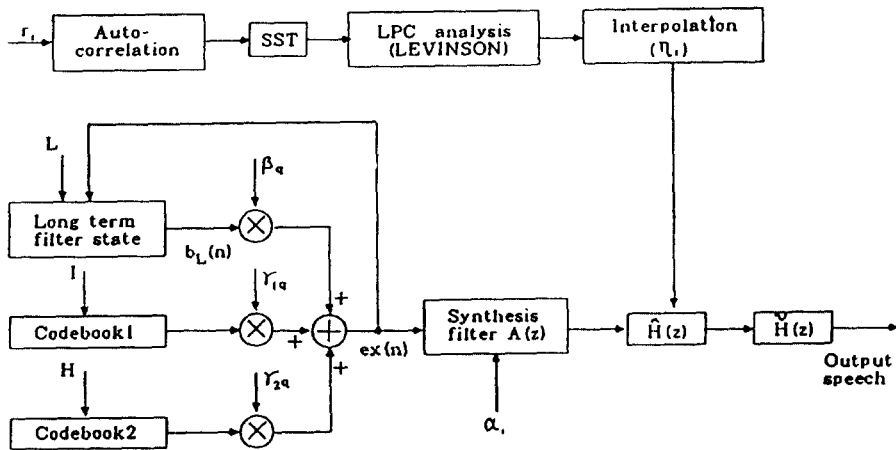


그림 2. VSELP 복호기 선도
Fig 2. VSELP decoder block diagram.

응답이 제거된 음성신호 사이의 오차자승합(total squared error)을 최소로 하는 값이 선택된다. 각 코드북을 이용하여 구한 여기신호 $ex(n)$ 은 다음 부프레임 처리를 위해 장기 필터 상태의 갱신(update)에 사용된다.

2. 복호기

부호기에서 부호화되어 전송된 계수들은 복호기에 서 복호화되어 음성합성에 이용된다. 그림 2⁽⁷⁾는 복호기의 선도이다. 복호기의 음성 합성 필터는 부호기의 가중 합성 필터와 다른 LPC 합성 필터를 사용하며, 음질 개선을 위해 적응 스펙트럼 후단 필터(adaptive spectral postfilter)를 사용하고 있다. 이 후단 필터는 ADPCM 음성부호 화기에서 합성음의 청각적 특성을 개선시키기 위해 제안되었다⁽¹⁰⁾. 또한 후단 필터를 거친 합성음의 명료도(brightness)를 개선하기 위한 일차 필터가 사용된다. 아래의 (1)식은 일차 필터에 대한 식이다. 여기서 τ 는 모든 음성신호에 대해 고정된 값을 가진

$$\bar{H}(z) = 1 - \tau z^{-1} \quad (1)$$

다. 본 논문에서 제안하는 적응 일차 필터에서는 τ 가 가변적이다.

III. 주기탐색

1. 기존의 페루프 방법에 의한 주기추출

VSELP에서 사용하는 페루프 방법에 의한 주기추출에서는 주기를 부프레임마다 구하며 장기 예측 필터의 응답은 (2)식으로 주어진다. 여기서 $\lfloor x \rfloor$ 는 내림 함수(floor function)를 의미하고, L 은 프레임의 첫 번째 L 샘플에 대한 지연을 나타낸다. β 는 필터의 계수이며, N 은 부프레임 길이다.

$$B_n(z) = \frac{1}{1 - \beta z^{-L} \sum_{l=0}^{n+L} z^{-l \cdot L}}, \quad 0 \leq n \leq N-1 \quad (2)$$

적응코드북은 20에서 146의 지연 각각에 대해서 탐색되고, (3)식과 같이 가중 합성 필터의 영입력 응답이 제거된 입력음성 $p(n)$ 과 가중 합성 필터에서 합성된 장기 필터

$$E = \sum_{n=0}^{N-1} \{p(n) - \beta b'_L(n)\}^2 \quad (3)$$

출력 $b'_L(n)$ 사이의 오차 자승합을 최소화하는 값이 최적 주기로 선택된다. 여기서 β 는 각 지연에 대한 적응코드북의 이득이다. (3)식을 β 에 대해 미분을 취한 후 전개하면 아래와 같이 (4)식을 최대로 하는 값이 주기 L 로 선택된다. 주기 L 에 대해서 C_L

$$\frac{(C_L)^2}{G_L} \quad (4)$$

$$\text{, 여기서 에너지: } G_L = \sum_{n=0}^{N-1} b_L'^2(n)$$

$$\text{상호상관관계: } G_L = \sum_{n=0}^{N-1} p(n)b'_L(n)$$

은 양의 값만이 고려되고 음의 값이 되는 경우에는 주기성이 없는 것으로 간주한다. 주기탐색시간을 줄이기 위해 $b'_L(n)$ 은 (5)식으로 나타내고 $z_L(n)$ 은 과거 값을 이용하여 순환적으로 계산된다. (6)식에서 $r(n)$ 은 장기필터의 상태를, $h(n)$ 은 가중 합성 필터의 임펄스 응답을 나타낸다.

$$\begin{aligned} L &\leftarrow \mathcal{O}_L^n \\ b'_L(n) &= \sum_{i=0}^n z_L(n-iL) \end{aligned} \quad (5)$$

$$z_L(n) = \sum_{i=0}^{\min(n, L-1)} r(i-L)h(n-i) \quad (6)$$

모든 $h(n)$ 에 대한 $z_L(n)$ 의 계산은 많은 계산량이 필요하므로 이 계산량을 줄이기 위해 (7)식과 같이 부분 임펄스 응답(truncated impulse response) $h'(n)$ 이 사용된다. 여기서 N_T 는 부분 임펄스 응답의 길이를 나타내며 21이 사용된다.

$$h'(n) = \begin{cases} h(n) & \text{for } 0 \leq n \leq N_T-1 \\ 0 & \text{for } N_T \leq n \end{cases} \quad (7)$$

위의 (5), (6), (7)식을 이용하여 (4)식의 에너지 G_L 은 식 (8), (9)와 같이 된다.

1. $L < N$ 일때

$$G_L = \sum_{n=0}^{N-1} b_L'^2(n) \quad (8)$$

2. $L \geq N$ 일때

$$G_j = E_{j-1} + \sum_{n=0}^{N-1} b_j^2(n) \quad (9)$$

, 여기서 $E_j = E_{j-1} + z_{j-1}^2(N-1) - z_j^2(N-1)$

이와 같이 페루프 방법에 의한 주기추출은 입력음성과 가중 합성 필터를 지친 장기 필터 출력 $b_j(n)$ 과의 오차자승합을 최소화하는 주기 L 을 구하는 것이다. 이때 $b_j(n)$ 은 장기 필터 출력 $b_l(n)$ 이 가중 합성 필터에서 합성된 신호이고, $b_l(n)$ 이 양자화된 신호인 $ex(n)$ 이 생성된 것이므로 $b_j(n)$ 에는 양자화 오차가 존재하게 된다. 따라서 이 오차가 존재하는 신호 $b_j(n)$ 대신에 오차가 존재하지 않는 신호를 이용하여 주기를 추출한다면 좀 더 정확한 주기를 추출할 수 있다. 기존의 FSELP에서는 주기 탐색시의 계산량을 줄이기 위해 (6), (7), (8), (9)과 같은 식을 이용한다. 그러나 근본적으로 모든 예측가능한 주기, 즉 20에서 146의 지연 각각에 대해 장기 필터 상태가 (5)식과 같이 계산되어야 하므로, 즉 가중 합성 필터에서 합성이 이루어져야 하므로 많은 계산량이 요구된다. 따라서 주기탐색시의 모든 예측가능한 주기에 대해서 해야 하는 계산량을 줄이고, 선택적인 조수의 가능주기에 대해서는 기존의 방법같이 계산한다면 전체적인 계산량을 줄일 수 있다.

2. 제안한 혼합 방법에 의한 주기추출

혼합 방법에서의 주기탐색은 개루프단에서의 주기탐색과 페루프단에서의 주기탐색으로 구성된다. 먼저 개루프단에서 입력음성을 이용하여 후보(candidate) 주기들을 구하고, 이 후보 주기만을 이용하여 페루프단에서 최적 주기를 구한다.

개루프단에서는 가능한 모든 주기, 즉 20에서 146의 지연에 대해 (10)식과 같이 오차자승합 E 를 계산한다. C 가 음의 값을 가지는 경우는 주기성이 없으므로 간주

$$E = \sum_{n=0}^{N-1} \{p(n) - p(n-k)\}^2 \quad , 20 \leq k \leq 146$$

$$= \sum_{n=0}^{N-1} p^2(n) + \sum_{n=0}^{N-1} p^2(n-k) - 2 \sum_{n=0}^{N-1} p(n)p(n-k) \quad (10)$$

$$= G_j + G_k - 2C$$

, 여기서 에너지: $G_j = \sum_{n=0}^{N-1} p^2(n)$

에너지: $G_k = \sum_{n=0}^{N-1} p^2(n-k)$

상호상관관계: $C = \sum_{n=0}^{N-1} p(n)p(n-k)$

한다. 여기서 $p(n)$ 은 가중 합성 필터의 입력역 응답이 재기된 입력음성이며 $p(n-k)$ 는 입력음성의 과거 값이다. k 는 지연(lag), 즉 주기를 나타낸다.

G_k 의 계산량은 이전 값의 생성을 통하여 줄일 수 있다. G_k 에 대한 식을 다시 (11), (12), (13)식으로 표현된다.

$$G_k = \sum_{n=0}^{N-1} p^2(n-k) = \sum_{n=0}^{N-1} d_k^2(n) \quad (11)$$

, 여기서 $d_k(n) = p(n-k)$, $0 \leq n \leq N-1$

$$G_k = \sum_{n=0}^{N-1} d_k^2(n) \quad , k=20 \quad (12)$$

$$G_k = G_{k-1} - d_{ps}^2(N-1) + d_{ps}^2(0) \quad , k \neq 20 \quad (13)$$

(13)식에서 $d_{ps}(N-1)$ 은 이전 주기에 해당되는 에너지의 부분이면서 현재 주기에 해당하는 에너지에 속하지 않는 부분을 나타내며 $d_{ps}(0)$ 는 현재 주기에 해당되는 에너지의 부분이면서 이전 주기에 해당하는 에너지에 속하지 않는 부분을 나타낸다. G_k 은 이전 주기에 대한 에너지를 말한다. 위의 과정에 의해 후보 주기는 오차자승합이 작은 순서대로 여러 개가 선택된다.

페루프단에서는 개루프단에서 구한 몇개의 후보주기만을 이용하여 기존의 FSELP 페루프 방법과 같이 (4)식을 최소화하는 방법에서 최적주기를 구한다.

이와 같은 혼합 방법에 의한 주기추출에서, 먼저 개루프단에서 후보주기들을 구하기 위해 (10)식을 이용하는데, 입력음성과 비교되는 신호로 입력음성의 과거 값 $p(n-k)$ 을 생성하여 그대로 사용한다. 이때 $p(n-k)$ 에는 양자화 오차가 없다. 그러나 (3)식의 기존의 방식에서는 입력음성과 비교되는 신호로 입력음성의 과거 여기신호 $cx(n)$ 을 생성하여 가중 합성 필터에서 합성된 신호 $b_l(n)$ 을 사용한다. 이 $b_l(n)$ 에는 양자화 오차가 존재한다. 따라서 개루프단에서 구한 후보주기들중 최적후보주기는 기존의 FSELP 페루프 방법으로 구한 최적주기보다 좀 더 정확한 값을 가진다. 그러나 실제 음성 합성에 이용되는 것은 각 코덱부의 출력값을 더한 여기신호가 음성 합성에 이용되기 때문에 혼합 방법의 개루프단에서 구한 최

적후보주기를 바로 입력음성의 최적주기로 이용하는 것보다는 몇개의 후보주기를 구한후 이 후보주기에 한해서 다시 페루프단에서 (3)식을 최대화하는 최적주기를 선택하는 것이 보다 좋은 결과를 나타낸다.

음성신호의 유성음에서는 이웃한 신호들사이에서의 주기는 급격한 변화없이 유사성을 가진다. (3)식과 같은 기존의 방식을 이용할때 $b_j(n)$ 의 양자화 오차로 인해 주기추출에서 오차를 가지게 되는데, 이 오차가 이웃 신호들 사이의 주기에서 2배 주기나 3배 주기같은 급격한 변화를 일으킨다. 그러나 제한한 혼합 방법에 의한 주기추출은 개루프단에서 오차없는 신호를 이용해 후보주기를 구하고, 이 후보주기에 한정해서 페루프단에서 최적주기를 구하게 되므로 양자화 오차로 인해 주기가 급격하게 변하는 현상이 줄어든다. 따라서 혼합 방법에 의한 주기추출은 주기의 평활 효과를 가진다.

혼합 방법의 개루프단에서는 오차사승합을 구하기 위해 입력신호와 비교되는 신호로 입력음성의 과거 값을 갱신하여 사용하므로 기존 VSELP의 (5), (6)식과 같은 합성이 불필요하며, 에너지 G_k 가 (11), (12), (13)식에 의해 $L=20$ 을 제외하고 나머지 탐색주기 21에서 146에 대해 검색에 있어서 두번의 연산만이 필요하다. 페루프단에서는 (4)식의 계산이 후보주기에 한정되어 이루어지며, 최적주기는 이 한정된 후보주기를 이용하여 구해진다. 따라서 혼합방법에서는 주기탐색시간이 단축된다. 즉, 계산량이 줄어든다.

IV. 주기성 향상 필터와 적응 일차 필터

유사한 주기성을 가지는 이웃한 여러개의 부프레임의 각신호에 가중치를 두어 더해진 음성신호에서 음성성분은 보다 주기성이 향상되고 잡음성분은 상대적으로 줄어든다^(12,13). 그러나 유성음과 무성음의 교차점이나 유성음과 묵음(silent)의 교차점에서는 오히려 특성이 나빠질 수도 있다. 본 논문에서는 가중해서 더해야하는 이웃신호와의 거리를 가까이 유지하여 유성음과 무성음의 교차점이나 유성음과 묵음의 교차점에서 특성이 나빠지는 현상을 줄이면서, 음성의 주기성을 향상시키는 주기성 향상 필터를 제안한다. 이것은 아래와 같은 형태를 가짐으로서 가능하다. 먼저 분모항에 의해 주기성

$$H_p(z) = \frac{a_0 + \sum_{l=1}^n a_l z^{-l}}{b_0 + \sum_{l=1}^m b_l z^{-l}} \quad (14)$$

을 향상시키고, 이 향상된 주기성을 가진 신호를 가지고 분자항에서 다시 주기성을 향상시킨다. 이와같은 방법은 짧은 필터 길이를 가지고서도, 주 가중해서 더하는 이웃 신호사이의 거리를 가까이 유지하면서도 실제보다 긴 필터의 효과를 가진다. 여기서 L_n , L_d 는 각각 주기를 나타내는데, 예로 L_n 은 현재 부프레임의 주기를 L_d 는 한 부프레임 이전의 주기를 나타낸다. 식에서 n 과 m 은 각각 분모분자의 필터 차수를 나타내며 임의적으로 확장 가능하다. 그러나 합성음의 잡음성분의 정도에 따라 적절하게 선택할 필요가 있다. 필터 계수 a_l , b_l 는 인접 부프레임과의 가중 정도를 나타내는 고정된 값이며 음질에 많은 영향을 준다. 따라서 적절한 필터 계수를 찾는 것이 중요하다. 필터의 계수는 아래식의 조건을 만족하는 범위 내에서 값을 임의적으로 가변시켜가며,

$$a_i > a_{i+1}, \quad b_j > b_{j+1}, \quad a_i \geq b_i \quad (15)$$

합성음의 청취를 통해 계수의 만족여부를 결정한다. 본 논문에서는 해밍 윈도우(hamming window)상에서 값을 변화시키면서 구했다.

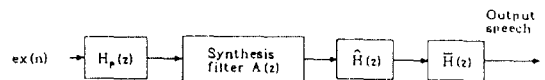


그림 3. 제안된 VSELP 부호화기의 주기성 향상 필터와 적응 일차 필터

Fig. 3. Proposed pitch period enhancement filter and adaptive first order filter in the VSELP decoder.

주기성 향상 필터 $H_p(z)$ 는 그림 3과 같이 합성 필터 이전에 적용되며, 세계의 코드북의 출력값을 결합한 여기신호 $ex(n)$ 을 바탕으로 하여 여러개의 부프레임으로부터 주기성이 향상된 새로운 여기신호를 만든다. 부호화기의 복호기⁽⁷⁾에서 세계의 코드북으로부터 결합된 여기신호 $ex(n)$ 은 합성필터에서 합성되고, 이 값을 적응 후단 필터를 거쳐 최종 합성음을

출력한다. 필터의 출력값은 단위이득을 보장하기 위해 크기조정이 이루어지며 이것을 위한 식은 (16)식과 같다. 여기서 $ex(n)$ 은 세개의 코드북에서 만들어진 여기신호로 필터의 입력이며, $ex_p(n)$ 은 주기성 향상 필터를 통과한 여기신호

$$ex_{ps}(n) = scale \cdot ex_p(n) \quad 0 \leq n \leq N-1 \quad (16)$$

$$, 여기서 scale = \sqrt{\frac{\sum_{n=0}^{N-1} ex^2(n)}{\sum_{n=0}^{N-1} ex_p^2(n)}}$$

를 나타낸다.

주기성 향상 필터를 적용하여 합성된 음성신호는 음의 명료도가 떨어지는 현상을 나타낸다. 기존의 VSELP 부호화기에서는 음의 명료도를 높이기위해 (1)식과 같은 일차 필터를 사용하고 있다. 아래의 (17)식에 이것을 다시 나타내었다. 음성신호의 주파수

$$\bar{H}(z) = 1 - \tau z^{-1} \quad (17)$$

특성은 매 프레임에 대해 어느정도 가변적이며, 일차 필터 적용시 매 프레임마다 가변되는 음성신호의 특성을 반영할 필요가 있다. 기존의 (17)식의 일차 필터는 모든 음성 신호에 대해 고정된 특성을 가지는 프리엠프시스 필터라고 볼 수 있으며, 따라서 매 프레임의 신호특성을 반영하여 가변적인 필터특성을 가지는 적응 일차 필터는 모든 음성 신호에 대해 고정된 특성을 가진다. 즉 일정한 τ 를 가지는 기존의 일차 필터에 비해 좋은 특성을 가진다. 아래식은 가변적 필터특성을 가지는 프리엠프시스 필터⁽¹¹⁾를 나타낸다.

$$\bar{H}(z) = 1 - az^{-1} \quad (18)$$

여기서 a 는 매 부프레임마다 가변되는 값이며, 최적 a 값은 복호기의 부프레임에 대한 합성음의 자기상관 함수 $R(0)$ 와 $R(1)$ 의 비를 이용하여 아래식과 같이 구해진다⁽¹¹⁾.

$$a = \frac{R(1)}{R(0)} \quad (19)$$

(19)식의 필터는 음성신호의 가변적 특성을 반영하기는 하나 실험결과 a 의 값이 0.75~0.97의 값의 범위를 가짐으로 인해 음질이 나빠지는 결과를 나타냈다.

기존의 VSELP는 (17)식에서 τ 를 0.4로 사용하고 있다. 실제 기존의 VSELP 부호화기에서 실험결과 $\tau = 0.4 \pm 0.1$ 근처에서 최상의 음질을 나타냄을 확인하였다. 본 논문에서는 이 점을 개선하여 음의 명료도를 효과적으로 개선시킬 수 있는 아래식과 같은 적응 일차 필터

$$\bar{H}(z) = 1 - \psi az^{-1} \quad (20)$$

를 적용하였다. 이 필터는 (19)식의 프리엠프시스 필터의 계수 a 에 가중치 ψ 를 두어서 $\psi a = 0.5 \pm 0.1$ 의 주위값을 가지도록 한것이다. 여기서 0.4 대신에 0.5를 사용한것은 주기성 향상 필터 적용시 음의 명료도가 조금 떨어지는 현상을 보상하기 위한 것이다. 가중치 ψ 는 모든 음성신호에 대해 고정된 값이며, $\psi a = 0.5 \pm 0.1$ 을 만족하는 범위내에서 임의적으로 가변시키면서 청취실험을 통해 구하였다. 이 필터는 기존의 프리엠프시스 필터⁽¹¹⁾의 특성을 그대로 가지며, 기존의 VSELP 부호화기의 일차 필터의 고정된 특성을 보완한다.

V. 실험 및 결과

본 논문에서 제안한 혼합 방법에 의한 주기추출과 주기성 향상 필터의 성능평가를 위해 기존의 VSELP 음성 부호화기에 적용하였다. 실험장비는 HP워크스테이션과 12비트 AD-DA 변환 시스템을 사용하였으며, C 언어로 구현하여 실험하였다.

1. 혼합 방법에 의한 주기추출

먼저 개루프단의 입력음성과 입력음성의 과거 값을 이용하여 오차자승합을 최소로 하는 값을 찾는다. 후보주기는 오차자승합이 작은 순서대로 5개가 선택된다. 이 후보주기들은 페루프단에서 여기신호에 대한 최적 주기탐색에 이용된다. 이것은 기존의 페루프 주기추출 방법을 이용한다.

본 논문에서 적용한 혼합 방법에 의한 주기추출은 기존의 VSELP의 주기탐색시 걸리는 시간을 줄여 빠른 탐색이 가능케 하였다. 표 1은 12비트로 양자화된 4초간 음성 데이터를 사용하여 실험한 기존방법과 제안된 방법의 결과를 나타내었다. 여기서 백분율은 기존의 방법에 대한 제안된 방법의 처리시간 비율이며, 실험 음성 데이터는 남녀 각각 10명을 선택하였다. 실험 결과 제안된 방법은 기존방법에 대해 평균적으

표 1. 음성 데이터 처리 시간

Table 1. Processing time for speech data.

	conventional method	proposed method	%
male	27.0	7.1	26.3
female	26.4	6.9	26.1

로 26%의 시간이 소요되었다.

본 논문에서 제안한 주기추출 방법은 여기신호의 양자화 오차로 인한 바람직하지 않은 2배 주기나 3배 주기를 최적주기로 선택하는 검색오류를 개선시켰다. 그림 4는 “제 전화번호는 232-6868입니다”라는 남자음성에 대하여 기존의 페루프 방법에 의한 추출 주기 그리고 본 논문에서 적용한 혼합 방법에 의한 추출주기를 나타내었다. 실제 한국인 남자음성의 경우 유성음부분에서 대개 35~65의 주기성을 가지며, 이웃한 신호사이에는 주기의 유사성을 가진다. 또한 무성음이나 묵음 부분에 근접한 부분에서는 주기의

급격한 변화를 나타낸다. 그림에서 (b)의 기존의 페루프 방법에 의해 구해진 주기에서는 주기의 연속된 부분과 급격한 변화부분의 구분없이 전체적으로 급격한 변화를 보인다. 즉, 2배 주기나 3배 주기같은 주기추출 오류가 많이 나타나고 있다. 그림 (c)의 혼합 방법에서는 급격하게 변화하는 부분과 연속성을 가지는 부분이 구별되고 있다. 2배 주기나 3배 주기가 기존의 페루프 방법보다 적게 나타남을 알 수 있다. 본 논문에서 적용한 이와같은 혼합방법은 2배 주기나 3배 주기와 같은 주기추출 오류를 줄임으로써 음질에 있어서 기존 VSELP의 페루프에 의한 주기탐색보다 좋은 결과를 나타내었다.

2. 주기성 향상 필터와 적응 일차 필터

본 논문에서 적용한 주기성 향상 필터에서 n 과 m 을 2로 하여 현재 부프레임의 주기와 이전 부프레임의 주기만을 고려하였다. 필터 계수는 $a_0=1.0, a_1=0.55, a_2=0.25, b_0=1.0, b_1=0.42, b_2=0.15$ 를 사용

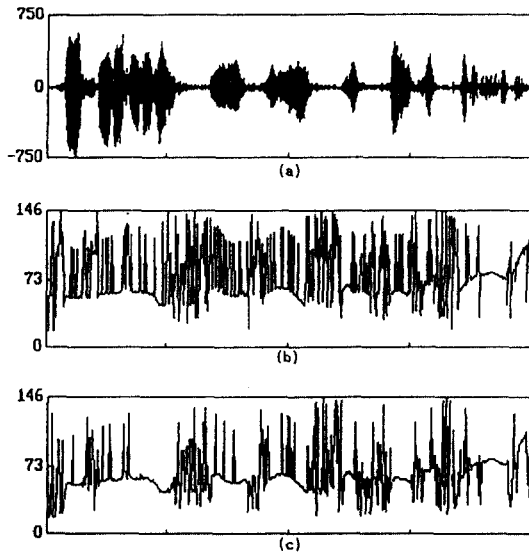


그림 4. 주기추출 방법에 따른 주기추출 결과의 비교. (a) 원음. (b) 기존 방법에 의해 결정된 주기 그리고 (c) 제안된 방법에 의해 결정된 주기

Fig 4. Comparison of pitch detection methods. (a) Original speech signal, (b) pitch periods determined by the conventional VSELP method, and (c) pitch periods by the proposed method.

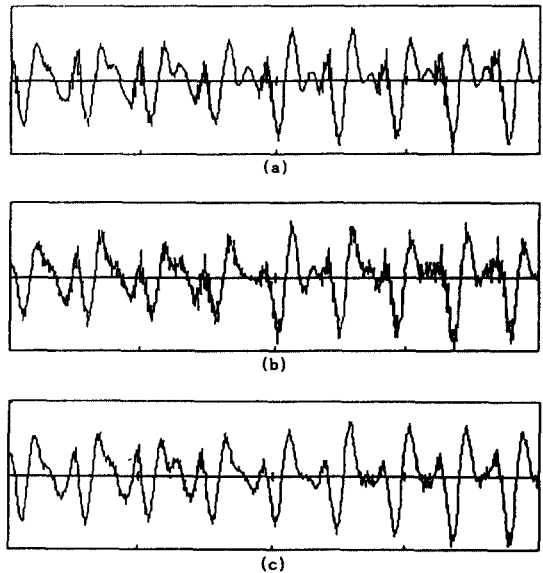


그림 5. (a) 원음과 (b) 기존의 VSELP 부호화기에 의한 합성음 그리고 (c) 향상된 주기성을 가지는 합성음의 파형

Fig 5. Waveforms of (a) original speech signal, (b) reconstructed speech signal by the conventional VSELP coder, and (c) reconstructed speech signal with pitch period enhancement.

하였다. 일차 필터에서 ψ 는 0.75를 사용하였다. 그림 5는 원음과 기존의 VSELP 음성 부호화 방식에 의해 합성된 음성신호, 그리고 주기성 향상 필터가 적용된 합성음성신호를 나타낸다. 그림 5에서 그림(b)의 기존의 VSELP 부호화기에 의해 합성된 합성음에서는 잡음성분이 많이 나타나고 있지만, 그림(c)의 제안된 필터를 이용한 합성음에서는 전체적으로 잡음성분이 줄어들었음을 알 수 있다.

주기성 향상 필터 적용시 정확한 주기를 이용하는 것이 필수적인데 잘못된 주기를 이용한 주기성 향상 필터의 적용은 오히려 음질을 악화시킬 수 있다. 실험결과 단순히 혼합방법에 의한 주기추출방법을 적용하는 것 보다는 혼합방법에 의해 정확한 주기를 구한 후 이 주기를 주기성 향상 필터에 적용할 때에 음질에서의 개선이 뚜렷했다. 표2는 음성 품질 검사법의 하나인 MOS^[11]를 이용한 합성음의 개선 정도를 측정할 것이다.

표 2. MOS 측정
Table 2. MOS measurement

	MOS scale					average MOS score
	5	4	3	2	1	
conventional method(%)	7.02	33.33	35.09	21.05	3.51	3.19
proposed method(%)	26.32	54.39	15.79	0.50	0.00	4.04

실험은 4초간 남녀음성 데이터 20개에 대해 19명의 20대 남녀를 대상으로 실시하였다. 표에서 보면 기존의 VSELP 부호화기의 합성음에 대해서는 MOS = 3.19를 나타냈으며, 개선된 합성음에서는 MOS = 4.04를 나타내었다. 또한 기존의 방법에서는 MOS등급 3의 비율이 가장 높게 나타나는 반면 제안된 방법에서는 4에서 가장 높은 비율을 나타내고 있다. 이 결과에서 보듯이 전체적으로 음질이 개선되었음을 알 수 있다.

VI. 결 론

본 논문에서는 주기추출 시간을 줄이면서 기존의 VSELP부호화 방식에서 장기 예측 필터 출력 여기신호의 양자화 오차로 인해 발생할 수 있는 주기추출 오류를 개선시키는 혼합 방법에 의한 주기추출 알고리즘을 제안하였다. 혼합 방법은 먼저 개루프단에서

입력음성과 양자화 오차가 없는 입력음성의 과거 값을 이용하여 오차자승합을 최소화하는 방법에서 여러개의 후보주기를 구하고, 개루프단에서는 이 후보주기를 이용하여 최적 주기를 선택한다. 또한 VSELP 음성 부호화기의 복호기에서 합성음의 주기성을 향상 시키면서 잡음을 줄이는 주기성 향상 필터와 합성음에 잡음의 필요도를 직설하게 강조시킬 수 있는 적응 일차 필터를 적용하였다. 주기성 향상 필터는 비슷한 주기성을 띠는 인접한 여러 개의 부프레임에 가중치를 부여 현재 음성신호에 반영한다. 적응 일차 필터는 매 프레임의 음성특성의 변화를 따라가며 음의 명확도를 높인다.

실험 결과 제안된 혼합 방법은 주기탐색시간을 기존의 개루프 방법에 비해 26%로 줄였으며, 2배 주거나 3배 주기를 최적 주기로 선택하는 것과 같은 주기추출 오류를 줄였다. 또한 주기성 향상 필터와 적응 일차 필터를 적용한 VSELP 음성 부호화기의 합성음에서 합성음의 과잉과 정위실험을 통하여 보다 높은 명확도를 가지며 잡음 성분이 줄어든 개선된 음질을 얻어내었다. 여러 명의 화자들 통한 MOS평가에서 기존의 합성음이 3.19의 점수를 나타내며 비해 개선된 합성음은 4.04를 나타냈다.

參 考 文 獻

1. M. R. Schroeder and B. S. Atal, "Code excited linear prediction(CELP) : High quality speech at very low bit rates," Proc. ICASSA, vol. 1, no.25.1.1, pp.937-940, Mar. 1985.
2. J. - H. Chen, Richard V. Cox, Y. -C. Lin, N. J. Yant, and M. J. Melchner, "A low delay CELP coder for the CCITT 16 kb/s speech coding standard," IEEE J. selected areas in communications, vol. 10, no. 5, June 1992.
3. Proposed Federal Standard 1016, Telecommunications: analog to digital conversion of radio voice by 4,800 bit/second code excited linear prediction(CELP), Pre Third Draft, July, 1990.
4. Motorola Inc., Vector sum excited linear prediction (VSELP) 13000 bit per second voice coding algorithm including error control for digital cellular, technical description, Oct., 1989.
5. I. A. Gerson and M. A. Jasiuk, "Vector sum excited linear prediction (VSELP) speech coding

- at 8kbps," Proc. ICASSP, pp. 461-464, Apr. 1990.
6. M. H. Sunwoo and S. Park, "Real-time implementation of the VSELP on a 16-Bit DSP Chip," IEEE Trans. on Consumer Electronics, vol. 37, no 4, Nov. 1991.
 7. EIA/TIA-IS-54-A, 2.1.3.3.2 Speech Coding.
 8. I. A. Gerson and M. A. Jasiuk, "Techniques for improving the performance of CELP-type speech coders," IEEE J. on Selected Areas in Communications, vol.10, no. 5, June 1992.
 9. B. S. Atal and M. R. Schroeder, "Predictive coding of speech signals and subjective error criteria," IEEE Trans on Acoustics, Speech, and Signal Processing, vol. ASSP-27, no. 3, June 1979.
 10. N. S. Jayant and V. Ramoorthy, "Adaptive postfiltering of 16 kb/s ADPCM speech," Proc. ICASSP, pp.829-832, Apr. 1986.
 11. R. Viswanathan and J. makhoul, "Quantization properties of Transmission Parameters in linear predictive system," IEEE Trans. on Acoustics, Speech, and Signal Processing, vol. ASSP-23, no.3, June 1975.
 12. R. H. Frazier, S. Samsam, L. D. Braidia, and A. V. Oppenheim, "Enhancement of speech by adaptive filtering," Proc. ICASSP, pp. 251-253, Apr. 1976.
 13. J. S. Lim, A. V. Oppenheim, and L. D. Braidia, "Evaluation of and Adaptive comb filtering method for enhancing speech degraded by white noise addition," IEEE Transaction on Acoustics, Speech and Signal Processing, vol. ASSP-26, no. Aug. 1978.
 14. S. Furui and M. M. Sondi, Advances in speech signal processing, Marcel Dekker, Inc., 1992.

文 相 在 (Sang-Jae Moon)

정희원

1948년생

1972년 2월 : 서울대학교 공과대학공업교육과 전자전공
(공학사)

1974년 2월 : 서울대학교 대학원 전자공학과(공학석사)

1985년 6월 : 미국 UCLA 통신공학 전공(공학박사)

1985년 7월 ~ 1986년 7월 : 미국 UCLA Postdoctor

1985년 7월 ~ 1986년 7월 : 미국 OMNET주식회사 Consultant

1974년 12월 ~ 현재 : 경북대학교 공과대학 전자공학과 교수

※주관심분야 : 디지털 통신시스템, 부호기술, 통신정보보호, 음성통신 등



成 基 鐵 (Ki-Cheol Sung) 정희원

1966년 1월 22일생

1992년 2월 : 경북대학교 전자공학과(학사)

1994년 2월 : 경북대학교 전자공학과(석사)

1994년 1월 ~ 현재 : 삼성사 영상MEDIA연구소 연구원

※주관심분야 : 음성통신, 음성 및 영상처리 분야