VSELP 부호화의 연산량 감소법에 관한 연구

Complexity Reduction Methods for Vector Sum Excited Linear Predictive Coding

Complexity Reduction Methods for Vector Sum Excited Linear Predictive Coding

Advisor : Yung Hwan Oh

by

Sung-Joo Kim

Department of Computer Science

Korea Advanced Institute of Science and Technology

A thesis submitted to the faculty of the Korea Advanced Institute of Science and Technology in partial fulfillment of the requirement for the degree of Master of Science in the Department of Computer Science.

Daejon, Korea
1993. 12. 16.

Approved by

Major Advisor

VSELP 부호화의 연산량 감소법에 관한 연구

김 승 주

위 논문은 한국과학기술원 석사학위논문으로
학위논문 심사위원회에서 심사 통과하였음.

1993 년 12 월 16 일

심사위원장 오 영 환 (인)

심사위원 맹 승 렬 (인)

심사위원 이 광 형 (인)

MCS 김 승 주, *Sung-Joo Kim, Complexity Reduction Methods for Vector Sum*
923084 *Excited Linear Predictive Coding,* VSELP 부호화의 연산량 감소법에 관한 연
구, *Department of Computer Science, 1994, 56p, Advisor : Yung-Hwan Oh.*

Abstract

The Vector Sum Excited Linear Predictive(VSELP) coding has the potential for producing high quality synthetic speech at bit rates as low as 4.8kbps. Unfortunately, the computational complexity of VSELP coding is prohibitive for real-time applications.

In this paper, three techniques to reduce the computational complexity of the VSELP coding are proposed. Firstly, the computational load to calculate the weighted synthesis filter responses to every basis vectors is significantly reduced by using overlapped sparse basis vectors. Secondly, the proposed preprocessing step reduces the search space efficiently before applying every linear combinations of the basis vectors to the VSELP codebook search procedure. It determines whether the combination coefficient of each basis vector can be fixed using heuristics so that the number of combinations decreases. Finally, to determine the pitch period only pre-selected candidates are applied to the pitch period search procedure. The candidates are selected by comparing them with the ideal excitation sequence which is produced by inverse filtering.

Then, all the proposed techniques for complexity reduction are combined to develop a VSELP coder with reduced complexity. It has been shown that the proposed coder retains good quality of synthetic speech and it is more than twice as fast as the original coder.

사랑하는 부모님께 이 논문을 드립니다.

목 차

Abstract

I. 서론 1

II. 음성 부호화기 4

2.1 음성 부호화의 개요 4

2.2 음성 부호화 방법 5

2.3 벡터 여기 부호화(Vector Excitation Coding) 7

2.3.1 벡터 여기 부호화의 구성 7

2.3.2 합성부 9

2.3.3 분석부 12

2.4 VSELP 부호화(Vector Sum Excited Linear Predictive Coding) 16

2.4.1 여기 코드북의 구조 16

2.4.2 여기 코드북의 탐색 17

2.4.3 VSELP 부호화의 문제점 20

III. VSELP 부호화의 연산량 감소법 21

3.1 개선된 VSELP 부호화의 특징 21

3.2 여기 신호 탐색에서의 연산량 감소법 22

3.2.1 중첩 희소 기저벡터 22

3.2.2 기저 벡터 부호 결정의 전처리 25

3.2.3 적응 코드북 탐색의 전처리 31

3.3 제안된 VSELP 부호화기 34

IV. 실험 및 평가 37

4.1 실험 환경 37

4.1.1 실험 자료 38

4.1.2 평가 방법 38

4.2 실험 결과 40
4.2.1 기존 VSELP 부호화기의 성능 평가 40
4.2.2 기저 벡터의 중첩도와 희소도에 따른 음질 변화 42
4.2.3 전처리 기저 벡터 선택 방법의 비교 44
4.2.4 적응 코드북 탐색 전처리의 유효성 49
4.2.5 제안된 VSELP 부호화기의 성능 평가 51
V. 결론 및 검토 55
참고문헌

I. 서론

음성 통신은 보편적이고 자연스러운 정보 교환의 수단으로, 현대 문명의 필수적인 요소이다. 또한, 통신 장비가 저렴해지고 개인화가 이루어짐에 따라 음성 통신에 대한 수요는 급격히 증가하는 추세이다. 한편, 간편한 정보 전달 수단으로 음성을 사용하는 새로운 응용 분야가 늘어나고 있다. 일례로 개인용 컴퓨터의 보급과 함께 소규모의 자동응답시스템이나 학습시스템에 음성을 이용하려는 시도가 이루어지고 있다.

음성을 디지탈(digital) 전송 선로를 통해 전송하거나 기억 장소에 저장하기 위해서는 먼저 음성 부호화가 수행되어야 한다. 이때 전송 선로나 기억 장소를 효과적으로 이용하려면 적은 정보량의 음성 부호화 방법이 요구된다.

일반적으로 음성 부호화 방법은 파형 부호화(waveform coding) 방법과 음원 부호화(source coding) 방법으로 구분된다. 음성 신호를 단순히 물리적으로 관측되는 파형으로 규정하는 파형 부호화 방법은 음질이 좋은 반면, 많은 정보량을 사용하는 단점이 있다. 음원 부호화 방법은 인간의 음성 생성 모델(speech production model)에 근거하여 음성의 조음(articulatory) 제어 파라미터와 음원(source) 생성 제어 파라미터를 분석한 후 각각을 부호화하는 방법이다. 음원 부호화 방법은 음성 생성 과정상의 제어 파라미터만을 이용하여 음성을 효과적으로 부호화하는 장점이 있으나, 기반이 되는 음성 생성 모델 자체의 단순성으로 인해 좋은 음질을 얻지 못하는 단점이 있다. 최근에는 파형 부호화와 음원 부호화를 조합한 복합형 부호화(hybrid coding) 방법이 개발되었다. 복합형 부호화는 두 방법의 장점을 반영하여 적은 정보량으로 좋은 음질을 유지한다[3].

대표적인 복합형 부호화 방법인 벡터 여기 부호화(vector excitation coding)는 음성을 8kbps이하의 정보량으로 부호화하며, 양질의 재생음을 제공하는 우수한 음성 부호화 방법이다. 벡터 여기 부호화에서는 부호화된 여기 신호(excitation signal)가 시간에 따라 변하는 음성 합성 필터(speech synthesis filter)를 통과

할 때 생성되는 신호로 음성을 표현한다[2]. 이때 여기 신호는 미리 정해둔 코드북내의 코드벡터중 합성 음성의 지각가중오류(perceptually weighted error)를 최소화하는 코드벡터로 벡터 양자화(vector quantization)된다. 그리고 음성 합성 필터의 계수는 시간적으로 변하는 조음 특성을 반영하며, 일반적으로 음원 부호화의 선형 예측 분석법(linear predictive analysis)을 이용하여 추정된다. 현재 벡터 여기 부호화 방법은 다양한 변형이 제안되고 있다. 특히 CELP 부호화(Code Excited Linear Predictive Coding)와 VSELP 부호화(Vector Sum Excited Linear Predictive Coding)는 각각 무선통신을 위한 CDMA (Code Division Multiple Access) 및 TDMA(Time Division Multiple Access) 표준안의 음성 부호화 방법으로 채택되어 사용되고 있다[5][6][7]. 국내에서도 최근 효율적인 음성 부호화의 중요성이 인식되면서 벡터 여기 부호화에 대한 연구가 진행중이다[9]-[13].

벡터 여기 부호화의 일종인 VSELP 부호화는 여기 신호로 미리 정한 기저 벡터의 선형 조합을 이용하는 특징이 있다. VSELP 부호화는 우수한 음질과 낮은 전송률을 갖는 장점이 있으나, 여기 신호를 결정하는데 막대한 연산량이 소모되는 단점이 있다. 북미 이동통신에 사용될 VSELP 부호화기는 실시간 동작을 위해 현재 40MIPS의 연산 능력을 갖는 전용 프로세서를 이용하여 구현되고 있고[4], 상당한 제작 비용이 요구된다. 따라서 VSELP 부호화기의 제작 비용을 낮추고 응용 범위를 넓히기 위해서는 VSELP 부호화의 연산량 감소법이 개발되어야 한다.

본 논문에서는 VSELP 부호화 방법의 여기 신호 탐색 과정에 필요한 연산량을 줄이는 세가지 방법을 제시한다. 첫째, 여기 신호 탐색시 중첩 희소 기저 벡터(overlapped sparse basis vectors)를 사용하여 기저 벡터의 필터 응답을 구하는 계산량을 감소시킨다. 둘째, 일부 기저 벡터의 선형 조합 계수를 경험적 지식(heuristics)을 바탕으로 미리 결정하는 전처리 과정을 수행하므로써 여기 신호의 탐색 범위를 줄인다. 셋째, VSELP 부호화의 피치 주기 탐색 과정에는 역필터링을

이용한 피치 주기 후보 검출을 선행하여 탐색 범위를 줄인다. 다음으로 이상의 세가지 연산량 감소 방법을 조합한 VSELP 부호화기를 설계 및 구현하고, 기존의 탐색 방법을 이용한 VSELP 부호화기 및 타 부호화기와 성능을 비교한다.

본 논문의 구성은 다음과 같다. II 장에서는 음성 부호화 전반에 대하여 설명하고, 본 연구의 기반이 되는 벡터 여기 부호화 및 VSELP 부호화기에 대해 기술한다. III 장에서는 VSELP 부호화의 여기 신호 탐색 과정의 연산량을 줄이기 위한 세가지 방법을 제시하고, 이를 이용하여 개선된 VSELP 부호화기를 구현한다. IV 장에서는 다양한 조건에서의 연산량 감소 방법에 대한 성능 평가 결과를 기술하고, 개선된 VSELP 부호화기를 기존 VSELP 부호화기 및 타 부호화기와 비교한다. 마지막으로 V 장에서는 결론 및 향후 연구 방향에 대하여 기술한다.

II. 음성 부호화기

음성은 디지탈 매체를 통해 전송 및 저장되거나, 디지탈 장비를 이용하여 분석되기에 앞서 부호화되어야 한다. 본 장에서는 음성 부호화 전반을 간략히 설명하고, 다양한 음성 부호화 방법중 성능이 우수한 백터 여기 부호화에 대해 기술한다. 그리고, 백터 여기 부호화의 연산량을 감소시키기 위해 설계된 VSELP 부호화기에 대해 기술하고 문제점을 살펴본다.

2.1 음성 부호화의 개요

음성 부호화(Speech Coding)란 음성 신호내의 필요한 정보를 손실없이 디지탈화(digitalize)하는 과정이다[1]. 음성 부호화는 크게 두가지 목적을 가지고 있다. 첫번째는 아날로그 신호인 음성을 컴퓨터와 같은 디지탈 장비를 이용하여 분석, 인식, 합성등의 처리를 수행하기 위함이고, 다른 하나는 디지탈 전송로나 기억장소내에 전송, 저장하기 위함이다. 전자의 경우 단순히 음성을 디지탈화하는 이외에 특징 파라미터를 분석 추출해내는 단계까지를 넓은 의미의 음성 부호화로 볼 수 있다.

음성 부호화의 가장 기초적인 방법으로는 음성 파형을 일정한 주기로 표본화(sampling)한 후 각 표본값을 선형적으로 양자화시키는 PCM(Pulse Code Modulation)이 있다. 전화망을 위한 PCM방식의 음성 부호화 규격은 64kbps이다[13]. 최근에는 16kbps이하의 적은 정보량을 사용하는 음성 부호화 방법들이 개발되어 응용되고 있다.

음성 부호화에 있어서, 정보량 감소의 요인은 음성 정보가 가지고 있는 중복성(redundancy) 때문이다. 원리적으로 음성은 음성 생성 기관의 비교적 느리고 제한된 움직임의 결과이므로 일반적인 음파와 동일한 방법으로 표본화된 음성 정보에는 많은 중복성이 내포된다. 또한 인간의 청각 기관도 제한된 범위의 인지 능력을 가지

므로 인간의 청각 능력으로 구분되지 못하는 음성 정보는 중복성으로 간주될 수 있다. 구체적으로 음성 부호화에 이용되는 음성 정보의 중복성에는 파형상의 표본값 크기의 불균등 분포 특성, 인접한 표본간 혹은 피치(pitch) 주기만큼 떨어진 표본간의 연관성 등이 있다. 또 음성을 주파수 성분별로 분석할 경우 대역별 에너지의 불균등 분포 특성도 음성 정보 감소에 이용될 수 있다(8).

음성을 전송하거나 저장하는데 사용되는 음성 부호화 방법은 매우 다양하다. 음성 부호화 방법의 성능 비교에는 정보량, 재생음의 음질, 처리 속도의 세가지 요소가 고려된다. 이중 정보량은 단위 시간의 음성 부호화에 필요한 비트수인 bps(bits per second)로 측정된다. 음성 통신에서 전송선의 대역폭과 음성 저장시 기억 용량은 사용되는 음성 부호화 방법의 정보량과 밀접한 관계가 있다. 다음으로 음질은 그 측정 방법이 매우 다양하며, 객관적인 음질 평가 방법으로 원음과 재생음 오류간의 에너지 비율에 해당하는 SNR(signal to noise ratio)을 많이 사용한다. 마지막으로 처리 속도는 단위 시간의 음성 부호화에 필요한 연산량으로 평가될 수 있다. 처리 속도는 사용 환경에 따라 음성 부호화의 제한 요건으로 작용한다. 일례로 음성 통신을 위한 음성 부호화 알고리즘은 부호화 및 복호화시 실시간 처리를 제한 조건으로 하며, 일정 수준 이하의 전송률을 갖고, 음질을 향상시키는 방향으로 설계된다.

2.2 음성 부호화 방법

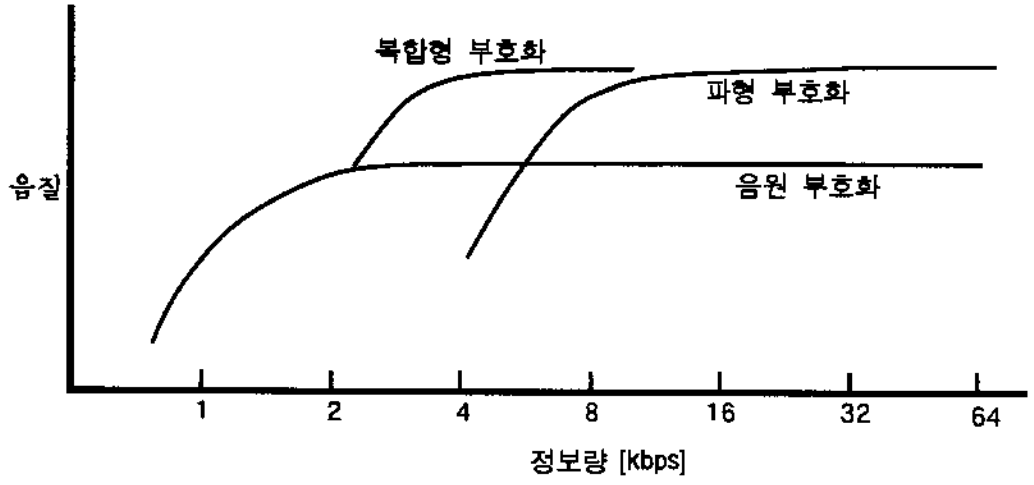
음성 부호화 방법은 인간의 음성 생성 과정을 고려하는가의 여부에 따라 파형 부호화와 음원 부호화로 구분된다(1). 음성 생성 과정을 고려하지 않고 단지 음성 파형의 통계적 특성만을 이용하는 파형 부호화는 음성을 단순한 1차원 시계열(time series) 정보로 규정한다. 파형 부호화의 목적은 관측된 1차원 시계열을 한정된 정보량으로 기술하면서 오류를 최소화하는데 있다. 파형 부호화는 충분한 정보량을 사용할 경우 음성 파형을 원음 그대로 재현하여 음질이 뛰어난 장점을 가지나, 적은 정보

량으로 부호화하는 경우에 음질이 급격히 저하되는 문제가 있다.

반면 음원 부호화는 인간의 음성 생성 모델을 전제로 하는 음성 부호화 방법이다. 음성은 음원, 즉 성도내의 공기의 흐름이 조음기관(articulatory)에 의해 방해를 받거나, 공명(resonance)을 일으킬 때 생성된다. 음성 생성 모델은 음원과 조음기관의 공명 특성을 각각 여기 신호(excitation signal)와 음성 합성 필터(speech synthesis filter)로 모델링한다. 음원 부호화는 음성 생성 모델을 기반으로 특정 음성을 생성하기 위한 음성 합성 필터의 계수를 추출하고 그 값을 부호화한다. 또 음원에 해당하는 여기 신호를 일정 주기의 펄스(pulse train)나 백색 잡음 신호(white random noise signal)로 대체한다. 음원 부호화 방법은 음성을 몇개의 파라미터들의 시계열로 부호화하므로 정보량이 적은 장점이 있으나, 음성 생성 모델의 단순성과 여기 신호 부호화의 어려움으로 양질의 재생음을 제공하지 못하는 단점이 있다.

최근에는 파형 부호화 및 음원 부호화의 장점을 조합하여 적은 정보량으로도 좋은 음질을 유지할 수 있는 복합형 부호화(hybrid coding) 방법이 개발되었다. 복합형 부호화는 음원 부호화의 음성 합성 필터를 사용하여 조음기관의 공명 특성에 해당하는 음성의 스펙트럼 포락(spectral envelope)을 재현하고 파형 부호화 방법을 이용하여 여기 신호를 생성한다. 복합형 부호화 방법에는 다음의 4가지 잇점이 있다[8].

1. 음원 정보를 주기 펄스나 잡음 신호로 모델링하는 대신 파형 부호화하므로 양질의 재생음을 얻을 수 있다.
2. 음성의 조음기관 공명 특성이 제거된 여기 신호는 파형 부호화 방법으로도 적은 정보량으로 부호화가 가능하다.
3. 조음기관 공명 특성은 음원 부호화에 의해 효과적으로 부호화된다.
4. 음원 정보의 파형 부호화로 여기 신호의 유/무성 결정 및 피치 주기 결정의 오류를 회피할 수 있다.



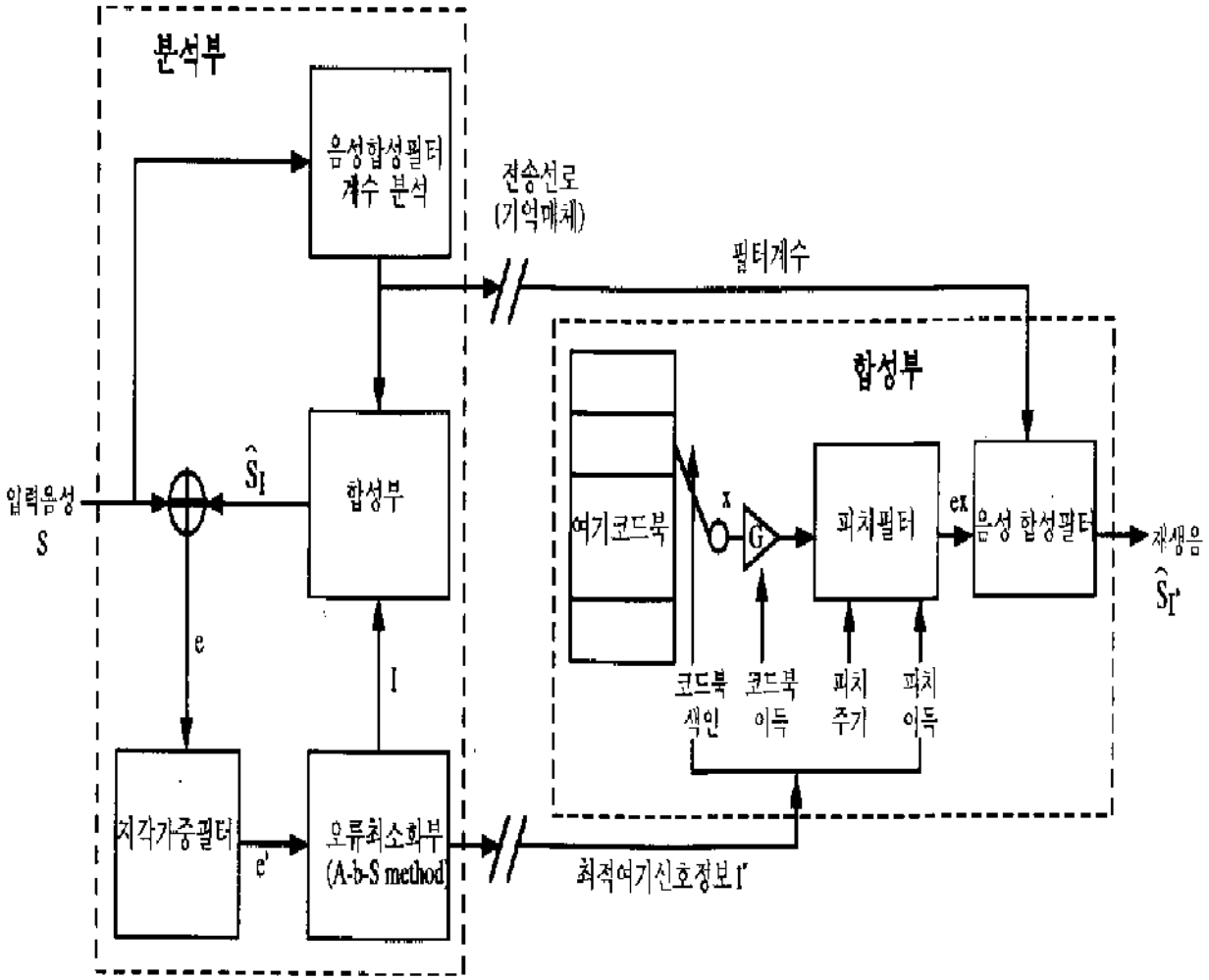
〔그림 2.1〕 음성 부호화 방법의 음질 및 정보량

2.3 벡터 여기 부호화 (Vector Excitation Coding)

복합형 부호화 방법중 일정 시간 단위의 여기 신호를 벡터 양자화하는 방법을 포괄적으로 벡터 여기 부호화라 한다[2]. 본 절에서는 벡터 여기 부호화의 일반적인 구성과 각 부분의 기능 및 동작을 기술한다.

2.3.1 벡터 여기 부호화의 구성

벡터 여기 부호화는 입력 음성을 부호화하는 분석부와 부호화된 음성을 재생하는 합성부로 구성된다. 벡터 여기 부호화의 분석부는 입력 음성으로부터 합성에 필요한 합성 필터 계수 및 여기 신호 정보를 추출한다. 분석부에서 부호화된 합성 필터 계수 및 여기 신호 정보는 전송선로나 기억 매체를 통해 합성부에 전달된다. 합성부는 여기 신호 정보를 이용하여 여기 신호를 생성하고 음성 합성 필터를 여기시켜 음성을 재생한다. 그림 2.2은 일반적인 벡터 여기 부호화의 구성도이다.



[그림 2.2] 벡터 여기 부호화기의 구성

2.3.2 합성부

벡터 여기 부호화의 합성부는 분석부에서 부호화된 필터 계수와 여기 신호 정보를 이용하여 음원 부호화의 방법으로 음성을 재생한다. 합성부의 동작 순서는 다음과 같다. 먼저 매 합성 단위마다 부호화된 여기 코드북 색인에 해당하는 코드벡터 x 를 구한다. x 에 이득(gain) G를 곱하고 피치 주기를 재현하는 피치 필터(pitch filter)를 통과시켜 합성을 위한 여기 신호 ex 를 생성한다. 여기 신호 ex 가 다시 음성의 조음 특성을 재현하는 음성 합성 필터를 거치면 최종 합성음 $\hat{S}$가 생성된다.

1) 여기 코드북(excitation codebook)

여기 코드북은 벡터 여기 부호화의 분석부에서 여기 신호를 벡터 양자화하기 위해 사용되며, 합성부에서는 음성 재생을 위해 분석부와 동일한 코드북이 사용된다. 여기 코드북의 형태는 탐색시 연산량과 밀접한 관계를 가지며, 합성음의 음질에도 크게 영향을 미쳐 벡터 여기 부호화 방법의 주요 연구 분야가 되었다.

여기 코드벡터는 입력 음성에서 성도 특성 및 피치 주기가 제거된 잔차 신호(residual signal)를 표현하기 위해 사용되므로 매 샘플이 독립적이고 가우스 분포(Gaussian distribution)를 따르는 부작위(random) 코드벡터가 이용된다. 그러나 코드북 탐색시 계산량의 감소를 위해 실제 구현되는 벡터 여기 부호화기에서는 중첩 구조의 코드북(overlapped codebook)을 이용하거나, 중앙-절단(center clipping)을 통해 코드벡터 내에 0이 많이 포함된 희소 코드벡터(sparse codevector)를 사용한다. VSELP 부호화에서는 기저 벡터(basis vector)의 선형 조합에 해당되는 특수한 여기 코드북을 사용하여 연산량을 감소시키고, 전송 오류에 강한 장점을 갖는다.

2) 피치 필터

여기 신호의 피치 주기성을 재현하기 위해 여기 코드북에서 선택된 코드벡터는 피치 필터를 통과하게 된다. 피치 필터는 1차 선형 예측 필터로 특성 함수(characteristic function)는 식 2.1과 같다.

$$H_p(z) = \frac{1}{1 - \beta z^{-L}} \qquad (2.1)$$

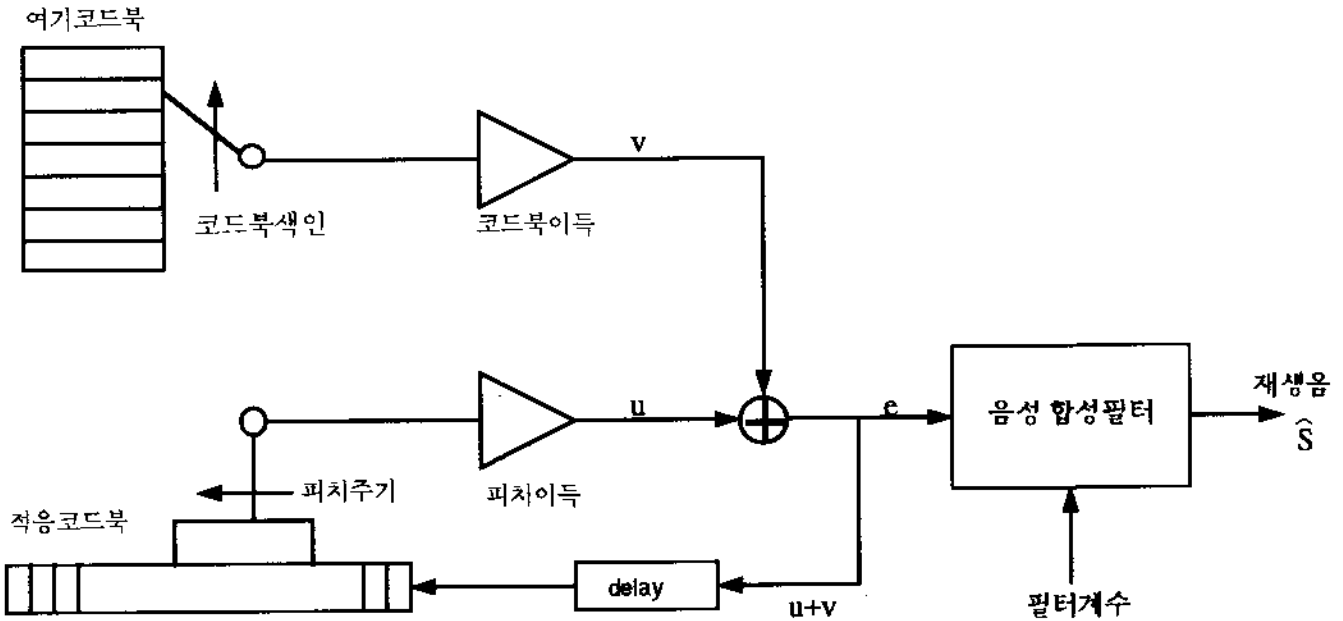
단, β 는 피치 필터 계수, L은 피치 주기

피치 주기와 피치 필터 계수의 피치 정보 분석은 기존의 음원 부호화의 피치 분석법을 사용하거나 지각가중오류를 이용하여 탐색에 의해 결정할 수 있다[8][26]. 전자의 경우, 분석이 간편하고 연산량이 적은 장점이 있으나 분석 방법의 정확도가 최종 합성음의 음질에 크게 영향을 미치는 단점이 있다. 따라서 일반적인 벡터 여기 부호화에서는 미리 정한 피치 정보 세트중 최소 오류를 갖는 피치 정보를 합성에 의한 분석법으로 선택한다.

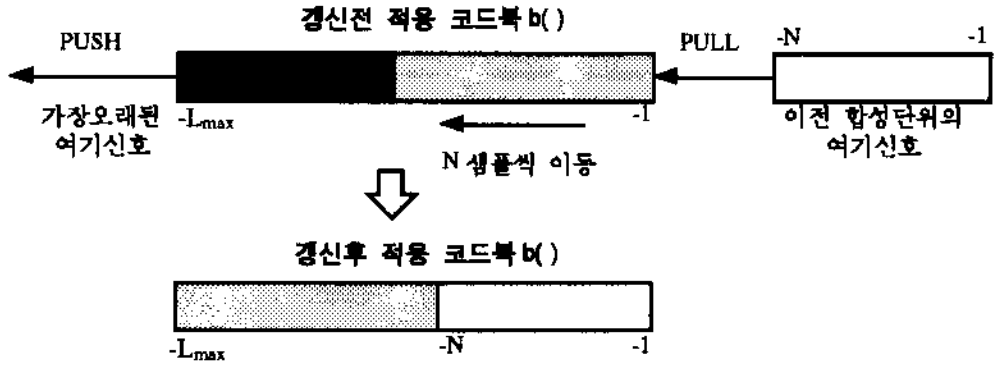
피치 정보를 합성에 의한 분석법으로 결정할 때, 피치 필터의 영입력응답(zero-input response)과 영상태응답(zero-state response)을 분리하면, 합성부는 그림 2.3과 같이 적응 코드북(adaptive codebook)을 사용하여 재구성될 수 있다. 적응 코드북은 이전 합성 단위에서 사용된 여기 신호를 저장하며, 피치 주기를 색인으로 피치 필터의 영입력응답을 재현한다. 적응 코드북은 최대 피치 주기 L_{max} 개의 과거 여기 신호 샘플을 저장하는 1차원 행렬로 실현된다. 식 2.2는 피치 주기 L인 경우의 적응 코드벡터 b_L를 구하는 방법이다.

$$b_L(n) = \begin{cases} b(n-L) & \text{if } n-L < 0 \\ b_L(n-L) & \text{if } n-L \geq 0 \end{cases} \qquad (2.2)$$

단, b(-i)는 적응 코드북의 i번째 샘플값이다.



[그림 2.3] 적응 코드북을 사용한 벡터 여기 부호화기의 합성부



〔그림 2.4〕 적응 코드북의 갱신

그림 2.4는 적응 코드북의 코드북 갱신 과정이다. 적응 코드북은 선입선출의 큐(queue)와 같이 동작하며 항상 최근 L_{max}개의 과거 여기 신호 샘플이 저장된다.

적응 코드북을 도입하면 피치 정보 탐색은 여기 코드북 탐색과 유사하게 진행되며, 일반적으로 탐색 범위를 줄이기 위해 피치 정보 탐색과 여기 코드북 탐색 과정은 순차적으로 수행된다.

3) 음성 합성 필터

피치 필터를 통과한 최종 여기 신호는 조음 특성을 반영하는 음성 합성 필터를 여기시켜 합성음을 생성한다. 음성 합성 필터는 음성 분석 분야에서 다양한 분석 방법이 연구되고 있다. 현재 대부분의 벡터 여기 부호화에서는 분석의 용이성과 분석 결과의 양자화 특성을 고려하여 선형 예측 분석을 이용한 all-pole필터가 사용되고 있다.

2.3.3 분석부

벡터 여기 부호화의 분석부는 입력 음성 신호와 합성음간의 지각가중오류가 최소

화되도록 제어 파라미터들을 결정한다. 분석부내의 오류 최소화 부분은 합성에 의한 분석 기법을 이용하여 미리 정해진 여기 신호 정보 세트중에서 가장 원음에 가까운 합성음을 재생하는 정보를 선택한다.

지각가중필터(perceptual weighting filter)는 청각 기관의 masking effect를 이용한 것으로, 여기 신호 탐색시에 최소화 목표인 오류파(error signal)를 인간의 청각 특성을 고려하여 실제 귀에서 인지되는 오류파가 되도록 변화시킨다. 지각가중필터의 특성 함수는 식 2.3과 같으며, 이때 γ 는 bandwidth expansion factor로 0에 가까울수록 오류파의 주파수 특성은 원음의 주파수 특성과 유사하게 된다.

$$W(z) = \frac{A(z)}{A(z/\gamma)} \qquad (2.3)$$

단, $\gamma = 0.8$이고, $1/A(z)$는 음성 합성 필터의 특성 함수이다.

1) 음성 합성 필터 계수의 분석 및 부호화

합성부에 포함된 음성 합성 필터의 계수를 분석하는 방법은 사용되는 필터의 종류에 따라 결정되며, 일반적으로 선형 예측 분석을 이용하여 선형 예측 계수를 구한 후, 이를 양자화 특성이 우수한 PARCOR(PARtial CORrelation) 계수나 LSP (Line Spectrum Pair)로 변환하여 부호화한다.

선형 예측 분석이란 현재의 음성 샘플이 과거 몇개 샘플의 선형 조합으로 예측된다는 가정에서 출발한 음성 분석 방법이다. 현재 음성의 샘플값을 x(n) 이라 하고 과거 p 개의 샘플값을 x(n-1), x(n-2), ..., x(n-p) 라고 하면 선형 예측에 의한 예측오차 e(n)은 식 2.4로 정의된다.

$$e(n) = |x(n) - \hat{x}(n)| = \left| x(n) - \sum_{i=1}^{p} \alpha_i x(n-i) \right| \qquad (2.4)$$

단, α_k는 선형 예측 계수

선형 예측 분석에서는 식 2.4의 e(n)에 대한 평균자승오류(mean square error)를 최소화하는 선형 예측 계수 α_k를 구하고, 추정된 α_k는 음성의 조음 특성을 반영하게 된다.

선형 예측 분석은 가정 자체의 제약으로 정확한 음성 스펙트럼 포락을 재현하지 못하며, 이를 해결하기 위한 방법으로 pole-zero model의 켑스트럼 분석(cepstral analysis)을 이용한 Homomorphic vocoder가 제안된 바 있다[16][17]. 그러나 이 방법은 분석에 많은 연산량이 요구되고, 켑스트럼 계수의 양자화에도 어려움이 많아 널리 사용되지 않는다.

2) 여기 신호 정보의 탐색

벡터 여기 부호화의 여기 신호 생성시에는 피치 주기 및 이득, 여기 코드북 색인 및 이득에 대한 정보가 필요하다. 여기 신호 정보의 탐색시 적응 코드북의 도입으로, 피치 정보는 여기 코드북 탐색과 분리되어 순차적으로 탐색될 수 있다. 물론, 순차적인 탐색 과정을 사용하면 준최적의 여기 신호가 선택되는 문제가 있으나, 음질에는 큰 영향을 주지 않으며, 탐색 범위를 효과적으로 줄일 수 있는 장점이 있다.

적응 코드북과 여기 코드북의 색인 및 이득은 다음의 탐색 과정을 순서대로 적용하여 결정한다. s, $\hat{s}$, e는 모두 길이 L의 벡터로 각각 여기 신호 탐색 구간내의 원음, 합성음 및 지각가중오류신호(perceptually weighted error signal)이다. v는 현재 탐색중인 코드북의 코드벡터를 이용한 여기 신호이고, u는 이전 탐색단계에서 선택된 최적 여기 신호이다. 즉, u는 적응 코드북 탐색인 경우에는 영벡터가 되고, 여기 코드북 탐색중인 경우에는 적응 코드북에서 이미 선택된 최적 적응 코드벡터에 이득을 곱한 벡터가 된다. 고려중인 여기 벡터 $v^{(i)}$는 해당 코드벡터 $x^{(i)}$와 이득 g_i의 곱으로 나타낸다.

$$v^{(i)} = g_i x^{(i)} \qquad (2.5)$$

다음의 F 와 W 는 모두 L x L 크기의 행렬로 j 번째 열은 각각 j 번째 필터 입력으로 임펄스가 입력될 때에 해당하는 음성 합성 필터와 지각가중필터의 영상태 임펄스응답(zero-state impulse response)이다. 따라서 i 번째 여기 벡터를 여기 신호로 이용한 합성음 $\hat{s}^{(i)}$는 식 2.6과 같이 표시할 수 있다.

$$\hat{s}^{(i)} = \hat{s}^{(0)} + F(u + v^{(i)}), \qquad 1 \le i \le N \tag{2.6}$$

단, $\hat{s}^{(0)}$는 음성 합성 필터의 영입력응답이다.

i 번째 코드벡터를 여기 신호로 이용할 경우의 지각가중오류 신호 $e^{(i)}$는 탐색 과정의 목표오류파 $e^{(0)}$와 코드벡터 $x^{(i)}$의 음성 합성 필터 및 지각가중필터를 통과한 결과 $y^{(i)}$를 정의하면 식 2.7로 정리된다.

$$e^{(i)} = W(s - \hat{s}^{(i)}) = e^{(0)} - WFv^{(i)} = e^{(0)} - g_i\, y^{(i)} \tag{2.7}$$

단, $e^{(0)} = W(s - \hat{s}^{(0)}) - WFu$ 이고, $y^{(i)} = WFx^{(i)}$

코드북 탐색 과정은 지각가중오류신호 $e^{(i)}$의 에너지 E_i를 최소화하는 코드벡터 색인 i 를 찾는 과정이다. 이때, E_i 의 g_i에 대한 편미분식을 0로 하여 정리하면 코드북 탐색 과정은 식 2.8의 $match_i$ 를 최대로 하는 색인을 찾는 과정과 동일하다.

$$match_i = g_i(2e^{(0)T} y^{(i)} - g_i y^{(i)T} y^{(i)}) = \frac{\left(e^{(0)T} y^{(i)}\right)^2}{y^{(i)T} y^{(i)}} \tag{2.8}$$

결국 적응 코드북 및 여기 코드북의 탐색은 각 코드북의 모든 코드벡터에 대하여 $match_i$ 를 계산하고 $match_i$ 의 최대값을 갖는 코드북 색인 i 와 해당되는 최적 이득 g_i 를 부호화하는 과정으로 실현된다. 하나의 코드벡터에 대한 적합도 $match_i$ 를 계산하는데 필요한 곱셈-덧셈 연산(multiply-add operation)의 횟수 U 는 표 2.1과 같다.

〔표 2.1〕 코드북 탐색 과정의 코드벡터별 연산량

항목	연산량
$WFx^{(i)}$	N(N+1)/2
$<e^{(0)}, y^{(i)}>$	N
$<y^{(i)}, y^{(i)}>$	N
$match_i$	2
U	(N+4)(N+1)/2

2.4 VSELP 부호화 (Vector Sum Excited Linear Predictive Coding)

벡터 여기 부호화기의 여기 신호 탐색의 연산량 감소를 위해 VSELP 부호화기가 제안되었다〔4〕. VSELP 부호화기는 여기 코드북으로 기저 벡터의 선형 조합을 사용하는 특징이 있다. 본 절에서는 VSELP 부호화의 코드북 구조와 탐색 알고리즘을 기술하고 소요 연산량을 계산한다.

2.4.1 여기 코드북의 구조

벡터 여기 부호화의 코드북 탐색과정에서 모든 코드벡터의 적합도 $match_i$ 를 계산하는데는 막대한 연산량이 필요하다. 코드북 탐색의 연산량을 줄이기 위해 많은 선행 연구가 있었고, VSELP 부호화도 이러한 맥락에서 개발된 벡터 여기 부호화의 변형이다.

VSELP 부호화기에서 2^M 개의 여기 코드벡터 u_i 은 식 2.9와 같이 M 개 기저벡터의 선형조합으로 생성된다.

$$u_i(n) = \sum_{m=1}^{M} \theta_{im} v_m(n), \qquad 0 \le i \le 2^M - 1 \qquad (2.9)$$

단, $\theta_{im} = \begin{cases} +1 & \text{if (bit m of codeword i) = 1} \\ -1 & \text{if (bit m of codeword i) = 0} \end{cases}$

식 2.9와 같이 코드벡터가 기저벡터의 선형조합으로 생성되므로, 전송 매체의 오류로 코드벡터 색인의 비트 하나가 변하더라도 해당되는 기저 벡터의 부호만 변하여 전체 코드벡터의 형태는 유지되는 전송오류에 강한 특성을 갖는다[7].

코드벡터의 색인 j 가 i 에 대한 1의 보수(1's complementary)라고 하면 $u_i = -u_j$ 의 관계가 성립된다. 이때 u_i 를 코드벡터 u_j 의 보수(complementary)라고 한다. 결과적으로 여기 코드북내의 2^M 개의 코드벡터는 2^{M-1} 개의 서로 다른 코드벡터와 그의 보수들로 구성된다. 한편 코드벡터 u_i 와 그 보수의 여기 코드북 탐색시 적합도는 동일하고, 이들의 최적 이득값은 부호만 바뀌는 성질이 있다. 따라서 VSELP 부호화의 여기 코드북 탐색은 처음 2^{M-1}개의 코드벡터에 대해서만 수행되며, 코드벡터 색인의 중복성은 코드벡터 이득을 항상 양의 값으로 고정시키는데 사용된다. 즉, 2^{M-1} 개의 여기벡터에 대한 탐색 결과 선택된 최적 코드벡터의 이득이 음인 경우에는 보수벡터의 색인을 부호화하여 이득을 양의 값으로 바꾼다. 이것은 여기 코드벡터 이득의 부호비트를 여기 코드북 색인에 포함시키는 결과가 된다.

2.4.2 여기 코드북의 탐색

VSELP 부호화의 최적 여기신호는 식 2.10을 최소화하는 벡터 b'_L 과 f_i 에 의해 결정된다. 이때, 이득 β 과 γ 는 벡터 b'_L 과 f_i 가 선택된 후에 식 2.10을 최소화하도록 동시에 결정된다.

$$E' = \sum_{n=0}^{N-1} \left(p(n) - \beta b'_L(n) - \gamma f_i(n)\right)^2 \qquad (2.10)$$

식 2.10에서 b'_L과 f_i는 각각 L 번째 적응 코드벡터 b_L과 i 번째 여기 코드벡터 u_i에 대한 가중합성필터(weighted synthesis filter) H(z)의 영상태응답이다. 목표 오류파 p는 앞서 식 2.7의 $e^{(0)}$ 에서 v가 영벡터인 경우와 같고, H(z)는 식 2.11과 같이 음성 합성 필터 특성과 지각가중필터 특성의 곱으로 결정된다.

$$H(z) = W(z)F(z) = \frac{A(z)}{A(z/\gamma)} \frac{1}{A(z)} = \frac{1}{A(z/\gamma)} \qquad (2.11)$$

실제 탐색 과정은 적응 코드북을 탐색한 후에 여기 코드북을 탐색하는 순서로 진행된다. 먼저 적응 코드북 탐색에서는 식 2.12를 최소화시키는 b_L^*을 구한다. 이 과정에는 기존 벡터 여기 부호화의 적합도 $match_i$를 이용한 코드북 탐색법이 적용된다.

$$E'_L = \sum_{n=0}^{N-1} \left(p(n) - \beta' b'_L(n)\right)^2 \qquad (2.12)$$

다음으로 여기 코드북 탐색은 식 2.13을 최소화하는 f'^*_i를 구한다.

$$E'_i = \sum_{n=0}^{N-1} \left(p(n) - \gamma' f'_i(n)\right)^2 \qquad (2.13)$$

식 2.13에서 f'_i는 코드벡터 u_i의 가중합성필터응답 f_i를 적응 코드북 탐색에서 결정된 b_L^*에 대해 직교화한 결과로 식 2.14와 같다. VSELP 부호화에서는 f_i의 직교화를 통하여 이미 결정된 b_L^*에 대한 최적 여기 코드벡터를 선택한다.

$$f'_i(n) = \sum_{m=1}^{M} \theta_{im} q'_m(n) \qquad (2.14)$$

식 2.14에서 q'_m은 각 기저벡터 v_m에 대한 H(z)의 영상태응답 q_m을 b_L^* 과 직교화한 결과이다. 식 2.14와 같이 VSELP 부호화기에서는 2^M개의 여기 코드벡터에 대한 가중합성필터응답 f_i의 직교화를 M 개의 기저 벡터의 직교화만으로 수행할 수 있다.

식 2.13의 최소화는 γ 에 대한 편미분을 0으로 하여 정리하면 식 2.15의 최대화로 바꾸어 표현된다.

$$match_i' = \frac{(C_i)^2}{G_i} \tag{2.15}$$

단,

$$\begin{cases} C_i = \sum_{n=0}^{N-1} f'_i(n)p(n) \\ G_i = \sum_{n=0}^{N-1} (f'_i(n))^2 \end{cases}$$

이때 C_i와 G_i는 각각 R_m과 D_{mj}를 이용하여 식 2.16과 2.17로 나타낼 수 있다.

$$C_i = \frac{1}{2}\sum_{m=1}^{M} \theta_{im} R_m \tag{2.16}$$

$$\left(R_m = 2\sum_{n=0}^{N-1} q'_m(n)\, p(n), \quad 1 \le m \le M \right)$$

$$G_i = \frac{1}{2}\sum_{j=2}^{M}\sum_{m=1}^{j-1} \theta_{im}\theta_{ij} D_{mj} + \frac{1}{4}\sum_{j=1}^{M} D_{jj} \tag{2.17}$$

$$\left(D_{mj} = 4\sum_{n=0}^{N-1} q'_m(n)\, q'_j(n), \quad 1 \le m \le j \le M \right)$$

한편, 코드북 탐색시에 코드벡터 색인 u 가 바로 이전 색인 i 와 v 번째 비트의 값만 다르다고 할 때, C_u, G_u은 식 2.16, 2.17을 이용하여 식 2.18, 2.19 같이 재귀적으로 정리된다.

$$C_u = C_i + \theta_{uv} R_v \tag{2.18}$$

$$G_u = G_i + \sum_{j=1}^{v-1} \theta_{uj}\theta_{uv} D_{jv} + \sum_{j=v+1}^{M} \theta_{uj}\theta_{uv} D_{vj} \tag{2.19}$$

$(\theta_{uv} = -\theta_{iv}\ ,\ \theta_{um} = \theta_{im}\ ,\ \text{if } m \neq v)$

여기 코드북 탐색시 코드벡터 색인으로 binary gray code를 사용하면 연속된 코드벡터의 색인 i 와 u는 하나의 비트만 다르게 되고, 식 2.18, 2.19와 같이 C_u, G_u는 이전 코드벡터의 C_i와 G_i를 이용하여 재귀적으로 계산될 수 있다.

2.4.3 VSELP 부호화의 문제점

벡터 여기 부호화 방법의 공통적인 문제는 합성에 의한 분석법으로 여기 신호 정보를 선택하기 위하여, 모든 여기 신호에 대한 합성음을 생성하고 원음에 대한 지각 가중오류의 에너지를 계산하는데 많은 연산량을 필요로 하는 점이다.

VSELP 부호화에서는 기저 벡터의 선형 조합을 이용한 여기 코드북으로 탐색에 필요한 연산량이 상당히 감소된다. 기저 벡터의 갯수 M 이 11이고 코드벡터의 길이 N 이 60인 VSELP 여기 코드북에 대해 벡터 여기 부호화의 코드북 탐색 과정을 적용하면 $2^{M-1}U$ (≒ 1,998,848)번의 곱셈-덧셈 연산이 필요하다. 그러나, VSELP 여기 코드북의 특성을 이용한 2.4.2절의 탐색 알고리즘을 적용하면 곱셈-덧셈 연산의 횟수가 53,174번으로 줄어 약 97%의 연산량 감소 효과를 얻는다. 또한 VSELP 여기 코드북 탐색 알고리즘은 직교화를 통해 최적 여기 코드벡터를 찾는 장점이 있다.

그러나 VSELP 부호화의 여기 코드북 탐색에도 초당 약 7백만번의 곱셈-덧셈 연산이 소요되고 그외에 적응 코드북 탐색 과정과 음성 합성 필터 계수 분석에 사용되는 연산량을 종합하면 실시간 구현을 위해서는 초당 약 2천만번 정도의 연산 속도를 갖는 전용 프로세서가 요구된다. 실제로 북미 이동통신의 표준안으로 제안된 8kbps VSELP 부호화기는 40MIPS의 연산 능력을 갖는 DSP칩(digital signal processing chip)을 이용하여 구현되고 있다(4). 이러한 VSELP 부호화의 복잡도는 부호화기 제작에 소요되는 비용을 높여 응용에 제한이 된다. 따라서 VSELP 부호화기의 제작 비용을 낮추고 응용 범위를 확장시키기 위해 VSELP 부호화의 연산량을 감소시키는 연구가 필요하다.

III. VSELP 부호화의 연산량 감소법

VSELP 부호화는 근본적으로 합성에 의한 분석법을 사용하므로 여기 신호 탐색에 많은 연산량을 필요로 한다. 본 장에서는 VSELP 부호화기의 여기 신호 탐색 과정에 필요한 연산량을 감소시키는 개선된 VSELP 부호화 방법을 제안한다.

3.1 개선된 VSELP 부호화의 특징

개선된 VSELP 부호화는 기존 VSELP 부호화의 구성을 바탕으로 하고, 여기 신호 탐색 과정에는 연산량을 감소시키는 다음의 세가지 방법이 적용된다.

1. 중첩 회소 기저 벡터를 이용한 여기 코드북을 사용한다. 따라서 여기 코드북 탐색 과정에 필요한 기저 벡터에 대한 가중합성필터의 영상태응답 q_m을 빠르게 계산할 수 있다.
2. 경험적인 지식을 바탕으로 미리 결정 가능한 기저 벡터 부호를 전처리 과정에서 선별하여 고정시킨다. 여기 코드북 탐색은 나머지 기저 벡터의 선형 조합에 대해서만 적용되므로 match$_i$의 계산 횟수가 줄어든다.
3. 적응 코드북 탐색에서는 역필터링(inverse filtering) 기법을 이용하여 이상적인 여기 신호(ideal excitation signal)를 구하고 각 적응 코드벡터와의 유사도를 측정하여 피치 주기의 후보를 검출한다. 적응 코드북의 탐색은 선발된 후보 피치 주기에 대해서만 적용되므로 탐색 범위가 줄어든다.

위의 연산량 감소 방법은 모두 내부 파라미터값에 따라 연산량 감소율이 가변적이고 음질이 변화될 수 있는 특징이 있다. 중첩 회소 기저 벡터의 경우, 중첩 정도와 회소도에 따라 연산량 감소율이 변화한다. 기저 벡터 부호 결정의 전처리 방법도 대상 기저 벡터의 수에 따라 다른 연산량 감소율을 보인다. 마지막으로 적응 코드북 탐

색의 전처리 방법에서는 선발할 후보 수가 연산량 감소율을 결정한다. 첫번째 방법의 경우에는 중첩 정도와 희소도가 큰 기저 벡터의 사용이 음질의 저하를 초래할 수 있다. 두번째와 세번째 방법은 전처리에 의한 탐색 범위 감소법으로 적용 정도에 따라서 준최적 코드벡터가 선택될 수 있다.

3.2 여기 신호 탐색에서의 연산량 감소법

본 절에서는 개선된 VSELP 부호화에 사용한 세가지 연산량 감소법을 기술한다. 중첩 희소 기저 벡터의 사용과 기저 벡터 부호 결정의 전처리 과정은 여기 코드북 탐색시 계산량을 감소시키며, 피치 주기 후보 검출의 전처리 과정은 적응 코드북 탐색 범위를 줄인다.

3.2.1 중첩 희소 기저벡터

VSELP 부호화기에서 여기 코드북 탐색시 기저 벡터의 가중합성필터응답 q_m의 계산 과정은 전체 여기 코드북 탐색 연산량의 37.8%를 차지하여 연산량 감소의 주요 대상이 된다. 본 논문에서는 기저 벡터의 가중합성필터응답 계산에 필요한 연산량 감소를 위해 중첩 희소 기저 벡터의 사용을 제안한다.

1) 필터 응답 계산의 개선법

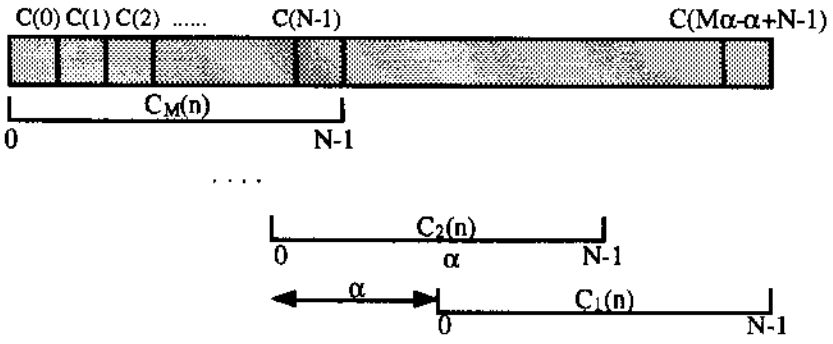
벡터 여기 부호화에서 코드벡터의 필터응답을 구하는 convolution과정은 벡터 길이가 N 일때 N(N+1)/2번의 곱셈-덧셈 연산을 필요로 하며, 이를 개선하기 위한 다양한 방법이 연구되었다[13][16]. 중첩되고 희소한 코드북을 이용하는 방법[22]은 코드벡터 convolution연산의 계산량을 효과적으로 감소시키며, 코드북 저장을 위한 기억 장소도 줄이는 우수한 방법이다.

중첩 구조의 코드북이란 연속되는 코드벡터가 α 개의 샘플씩 이동되어 중첩된 코드북을 말한다. 이때, α를 중첩도로 정의하며 코드벡터의 중첩 정도를 나타낸다. 길이 N인 M개의 코드벡터에 대한 중첩 구조 코드북은 그림 3.1과 같이 $(M-1)\alpha+N$ 개의 샘플로 이루어진 1차원 행렬로 실현된다. 중첩구조 코드북의 i번째 코드벡터 $C_i(n)$는 식 3.1과 같이 i-1 번째 코드벡터 $C_{i-1}(n)$과 α 개의 새로운 샘플들로 이루어진다.

$$C_i(n) = \begin{cases} C(M\alpha - i\alpha + n) & \text{if } \alpha > n \geq 0 \\ C_{i-1}(n-\alpha) & \text{if } N > n \geq \alpha \end{cases}, \quad M \geq i > 1 \qquad (3.1)$$

중첩구조 코드북을 사용하면 식 3.1의 코드벡터 C_i의 필터 응답 C_i'는 식 3.2와 같이 재귀적으로 정리할 수 있다. 이때 h는 필터의 임펄스응답(impulse response)이다.

$$\begin{aligned} C_i'(n) = C_i(n) * h(n) &= \sum_{k=0}^{n} C_i(k)h(n-k) \quad , M \geq i > 1, N > n \geq \alpha \\ &= \sum_{k=0}^{\alpha-1} C_i(k)h(n-k) + \sum_{k=\alpha}^{n} C_{i-1}(k-\alpha)h(n-k) \\ &= \sum_{k=0}^{\alpha-1} C_i(k)h(n-k) + C_{i-1}'(n-\alpha) \end{aligned} \qquad (3.2)$$



〔그림 3.1〕 중첩 구조 코드북

따라서 M 개의 코드벡터의 필터응답 C_i'는 식 3.3으로 계산되며, 전체 계산량은 $N(N+1)/2 + (M-1)\{\alpha(\alpha+1)/2 + \alpha(N-\alpha)\}$ 가 된다.

$$\begin{cases} C_1'(n) = \sum_{k=0}^{n} C_1(k)h(n-k) \quad , N > n \geq 0 \\ C_i'(n) = \begin{cases} \sum_{k=0}^{n} C_i(k)h(n-k) & , \alpha > n \geq 0 \\ \sum_{k=0}^{\alpha-1} C_i(k)h(n-k) + C_{i-1}'(n-\alpha) & , N > n \geq \alpha \end{cases} \\ \quad , M \geq i > 1 \end{cases} \tag{3.3}$$

희소 코드북이란 가우스 분포를 갖는 코드북 내부의 샘플값들을 중앙-절단(center-clipping)에 의해 0이 많이 포함되도록 수정한 것이다. 선행 연구의 결과 실험적으로 코드벡터내에 샘플의 90%가 0인 희소 코드북을 사용하여도 음질이 저하되지 않는 것으로 밝혀졌다[20]. 중첩구조 코드북의 식 3.3에서 i 번째 코드벡터의 α 개의 새로운 샘플중 $\sigma\alpha$ 개가 0인 경우, 전체 계산량의 σ 배만큼의 연산량을 추가로 감소시킬 수 있다. 이때, σ 를 코드북의 희소도(sparsity)라고 하며 코드북내의 각 샘플값이 0일 확률을 나타낸다.

2) 중첩 희소 기저 벡터의 이용

VSELP 부호화기에서는 코드벡터가 M 개의 기저 벡터의 선형 조합으로 결정된다. 이때 기저 벡터는 백색 잡음 신호를 이용하며, 결과적으로 선형 조합된 2^M 개의 코드벡터도 백색 잡음 신호가 된다. 제안된 중첩 희소 기저 벡터는 M 개의 기저 벡터를 중첩도 α 에 희소도 σ 의 중첩 희소 코드북으로 구현한다[11].

중첩 희소 기저 벡터의 가중합성필터 응답 q_m 은 식 3.3과 유사한 방법으로 계산된다. 따라서 M 이 11일때 중첩도 6에 희소도 0.9인 중첩 희소 기저 벡터를 사용하면 q_m 계산을 위한 곱셈-덧셈 연산은 528회로 기존 기저 벡터를 위한 연산량 20,130

율 97.4% 감소시킨다.

중첩 희소 기저 벡터를 이용한 VSELP 여기 코드북의 코드벡터는 희소 구조 여기 코드북의 여기 코드벡터와 유사한 형태를 보여주며, 실험적으로 크게 음질을 저하시키지 않음이 밝혀졌다[11].

3.2.2 기저 벡터 부호 결정의 전처리

VSELP 부호화의 여기 코드북 탐색 과정은 M 개 기저 벡터의 선형 조합 계수 θ_{i^*m}의 부호 결정 과정으로 볼 수 있다. 기존의 방법은 2^{M-1} 개의 모든 조합의 여기 코드벡터에 대한 적합도 $match_i'$를 계산하고 최대값을 갖는 코드벡터의 색인 i^*로 θ_{i^*m}를 결정한다. VSELP 부호화에서 모든 코드벡터의 적합도 계산 과정은 $match_i'$를 2^{M-1} 회 계산하여 전체 여기 코드북 탐색 과정의 44.2%에 해당하는 많은 연산량이 소요된다.

본 논문에서는 전처리 과정을 통해 각 기저 벡터의 선형 조합 계수를 음질의 저하 없이 미리 결정할 수 있는지 판별하여 고정시키고, 결정되지 않은 나머지 기저 벡터의 선형 조합 계수를 후처리에서 기존의 방법으로 결정하는 새로운 여기 코드북 탐색 알고리즘을 제안한다. 본 논문에서는 기저 벡터 부호를 기저 벡터의 선형 조합 계수 θ_{i^*m}와 동일한 의미로 사용한다.

1) 전처리 과정의 기저 벡터 부호 결정 방법

VSELP 부호화기의 여기 코드북 탐색 과정은 식 3.4의 E_i'를 최소화한다. 이때, p, q_m' 는 모두 N차원 벡터이다.

$$E_i' = \| p - \gamma' \sum_{m=1}^{M} \theta_m q_m' \|^2 \qquad (3.4)$$

식 3.4의 γ' 에 대한 편미분을 0으로 하여 정리하면 E_i' 의 최소화는 식 3.5의

match$_i$'의 최대화로 변환된다.

$$match_i' = \frac{<p, \sum_{m=1}^{M} \theta_{im} q_m'>^2}{\| \sum_{m=1}^{M} \theta_{im} q_m' \|^2} = \|p\|^2 \cos^2 \omega_{pf'} = \frac{(C_i')^2}{G_i'} \qquad (3.5)$$

$\|p\|^2$는 여기 코드북 탐색 과정에서 상수로 고정되므로, match$_i$'의 최대화는 M 개의 벡터 q_m'의 선형 조합 결과와 벡터 p간의 편각 $\omega_{pf'}$를 최소화한다. 따라서 M 개의 선형 조합 계수 $\theta_{i'm}$는 벡터 q_m'의 부호를 + 혹은 -로 선형 조합한 결과가 벡터 p 와 가장 유사한 방향을 갖도록 결정된다.

한편 R_m은 두 벡터 q_m'과 p의 내적으로 식 3.6으로 나타낼 수 있고, R_m의 부호는 q_m' 과 p 사이의 편각 $\omega_{pq'}$에 의해 결정된다.

$$R_m = 2<q_m', p> = 2\|q_m'\|\|p\|\cos\omega_{pq'} \qquad (3.6)$$

M 개의 벡터 q_m'을 더하거나 빼서 목표 벡터 p 와 같은 방향의 벡터를 만들 때, p와 q_m'사이의 편각이 $\pi/2$ 보다 작으면 더하고, $\pi/2$ 보다 크면 빼는 직관적인 방법을 생각할 수 있다. 실제 R_m의 부호와 기존의 여기 코드북 탐색에 의해 결정된 기저 벡터의 선형 조합 계수 $\theta_{i'm}$와의 일치 정도를 실험을 통해 조사한 결과, 약 54초의 음성 자료를 부호화하는 과정에서 79,112번의 R_m과 $\theta_{i'm}$의 부호 비교 결과중 85.9%에 해당하는 67,971번의 비교에서 R_m과 $\theta_{i'm}$의 부호가 일치되는 밀접한 관계를 나타내었다.

그러나 항상 $\theta_{i'm}$를 R_m의 부호로 결정하여 여기 코드북 탐색을 생략하는 경우에는 재생음의 SNRseg (segmental signal to noise ratio)가 기존의 탐색 방법에 비해 약 1dB까지 감소하는 음질 저하를 나타내었다. 따라서 R_m의 부호를 이용한 $\theta_{i'm}$의 부호 결정은 적절한 조건하에서만 수행될 필요성이 있다. 본 논문에서는 기저 벡터 부호 결정 전처리 대상을 선별하기 위하여 경험적 지식을 사용한 4가지 방법을 제시한다.

2) 기저 벡터 부호 결정 전처리 대상의 선별법

전처리 과정의 대상을 결정하기 위해 앞선 실험에서 R_m과 θ_{i^*m}의 부호가 일치하지 않는 11,141번의 경우와 부호가 일치하는 67,971번의 경우에 대해 각각 q'_m의 p 방향 성분의 제곱 SP와 p, q'_m사이의 편각의 cosine 제곱 SC를 식 3.7, 3.8과 같이 계산하여 통계 처리하였다.

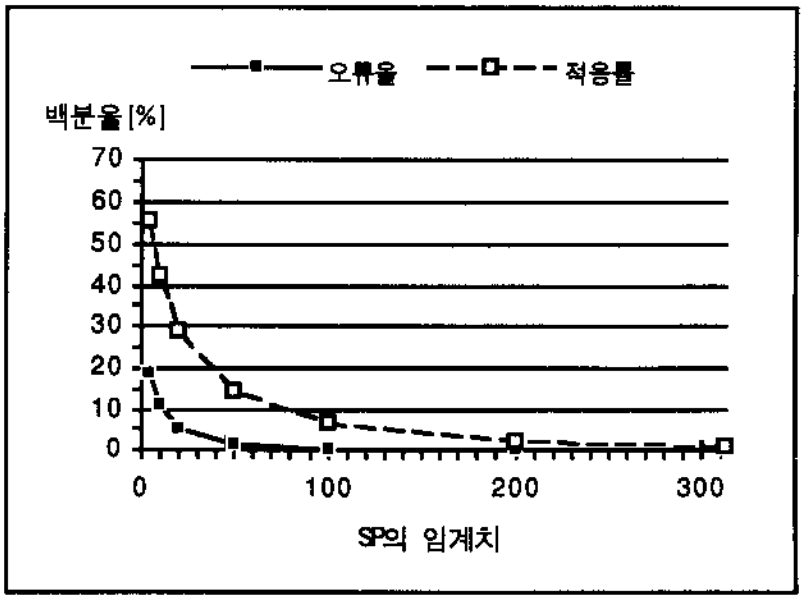
$$SP = 4\|q'_m\|^2\cos^2\omega_{pq'} = \frac{R_m^2}{\|p\|^2} \qquad (3.7)$$

$$SC = \cos^2\omega_{pq'} = \frac{SP}{D_{mm}} \ , \ D_{mm} = 4\langle q'_m, q'_m\rangle \qquad (3.8)$$

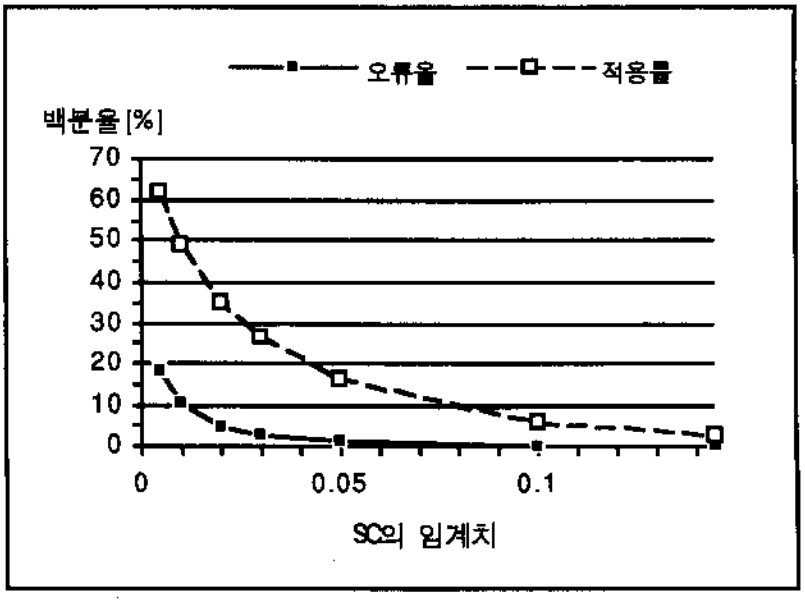
SP와 SC는 모두 q'_m과 p 의 상관성을 측정한 값으로 SP는 방향과 크기를 함께 고려하고, SC는 방향만을 고려한 척도이다. 전처리 과정의 범위 결정을 위하여 SP와 SC를 사용한 이유는 SP 또는 SC가 클수록 θ_{i^*m}이 R_m의 부호와 일치될 가능성이 크다는 실험적 결과에 기반을 둔다. 따라서 기저 벡터 부호 결정 전처리의 대상 선택 방법으로 SP 또는 SC값이 미리 정한 임계치보다 큰 기저 벡터를 선택하는 2가지 방법을 사용할 수 있다.

그림 3.2의 실선은 SP가 임계치보다 크고 R_m과 θ_{i^*m}의 부호가 일치하지 않는 경우의 백분율이고 점선은 SP가 임계치보다 크고 R_m과 θ_{i^*m}의 부호가 일치하는 경우의 백분율이다. 그림 3.3는 SC에 대한 동일한 통계 처리의 결과이다.

그림 3.2, 3.3에서 실선은 R_m과 θ_{i^*m}의 부호가 일치하지 않는 경우에 SP 또는 SC값이 임계치보다 클 가능성으로 전처리에 의한 오류율을 나타낸다. 한편 점선은 R_m과 θ_{i^*m}의 부호가 일치하는 경우에 SP 또는 SC값이 임계치보다 클 가능성으로 여기 코드북 탐색시 전체 기저 벡터중 평균 전처리 적용 대상의 비율이 된다.



〔그림 3.2〕 SP의 임계치를 이용한 전처리 대상 선택



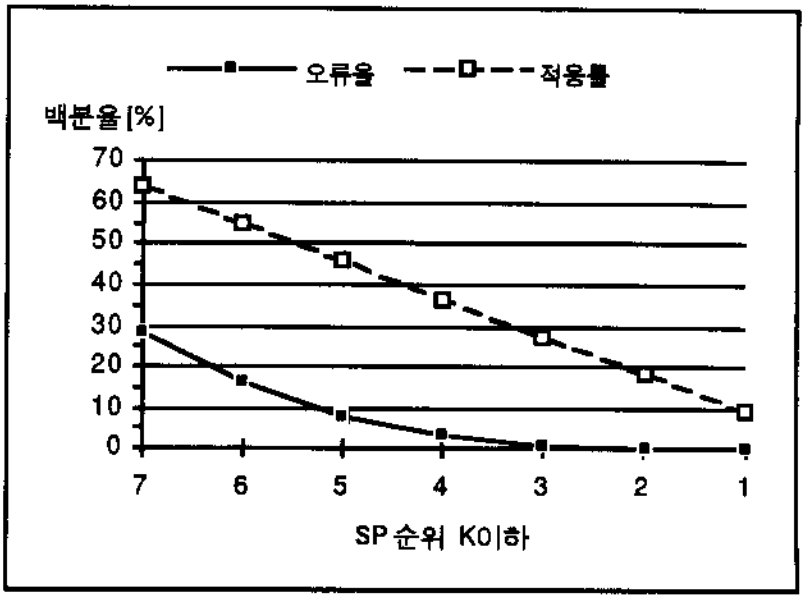
〔그림 3.3〕 SC의 임계치를 이용한 전처리 대상 선택

최적의 임계값은 전처리에 의한 오류를 최소화하면서, 여기 코드북 탐색시 평균 전처리 적용 대상의 수를 최대화한다. 그러나, SP와 SC의 임계치를 상수로 고정할 경우 전처리에 의한 오류를 5%까지 허용해도 평균 전처리 적용 대상의 수가 SP 사용시 2.8개, SC 사용시 3.4개로 적어 연산량이 크게 감소되지 않는다.

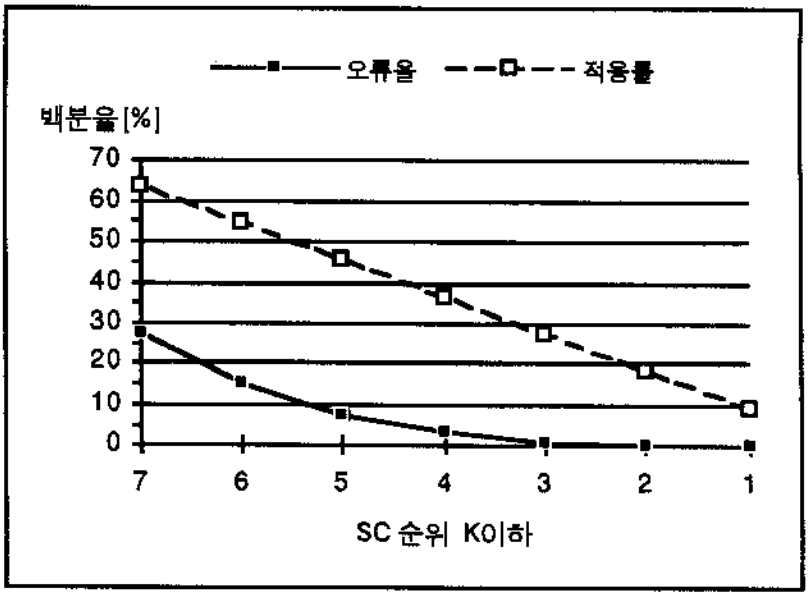
전처리 적용 대상의 수를 늘이기 위해서는 SP와 SC의 임계치를 매 여기 코드북 탐색시마다 가변적으로 결정하는 방법이 필요하다. 본 논문에서는 임계치를 가변적으로 결정하는 2가지 방법으로 SP와 SC 값이 큰 순서로 K개를 전처리 과정의 대상으로 선택하는 K-선택법(K-selection)을 제안한다.

SP와 SC 값의 K-선택법으로 전처리 과정의 대상을 선택하는 방법에서 적절한 K를 정하기 위하여 앞 실험에서의 R_m과 $\theta_i{}^*{}_m$의 부호가 일치하지 않는 11,141회의 경우에 대한 SP 및 SC값의 순위를 구하여 통계 처리하였다. 그림 3.4과 3.5의 실선은 각각 총 7,192회의 부호 비교중 R_m과 $\theta_i{}^*{}_m$의 부호가 일치하지 않으면서 SP 혹은 SC의 순위가 K이하인 경우의 백분율이다. 따라서 K-선택법을 이용한 전처리 적용 대상의 선택시 오류율과 같다. 그림 3.4와 3.5의 점선은 K값의 전체 기저 벡터수에 대한 비율로써 전처리 대상의 적용률이다.

그림 3.4와 3.5에서 SP 또는 SC의 K-선택법을 이용한 전처리 대상의 수를 4개로 정하여도 오류율은 5%이하로 유지되었다. 이상의 통계 자료로부터 SP 또는 SC가 큰 순서로 K개의 기저 벡터를 전처리 대상으로 결정하는 방법이 상수 임계치를 사용하는 방법에 비해 전처리에 의한 오류가 적고, 평균 전처리 적용 대상의 수도 많은 우수한 방법임을 알았다. 한편 전처리 대상 설정을 위한 내부 변수로는 SP, SC 모두 비슷한 결과를 주었으나 유사도 측정시 방향과 크기를 모두 고려하는 SP를 사용할 경우에 약간의 오류 감소가 있었다.



〔그림 3.4〕 SP의 K-선택법을 이용한 전처리 대상 선택



〔그림 3.5〕 SC의 K-선택법을 이용한 전처리 대상 선택

3.2.3 적응 코드북 탐색의 전처리

VSELP 부호화에서 적응 코드북 탐색 과정은 여기 코드북 탐색 과정외에 가장 많은 연산량을 필요로 하는 과정이다. 적응 코드북 탐색 과정의 연산량을 줄이기 위한 방법에는 다음의 세가지 방법이 있다.

1. 적응 코드북의 중첩성을 이용하여 연속되는 적응 코드벡터의 가중합성필터 응답을 재귀적으로 계산하는 방법
2. 피치 주기의 느린 변화 특성을 이용하여 델타 탐색(delta-search)하는 방법
3. 전처리 과정에서 피치 주기의 예상 후보를 선정하여 탐색 범위를 축소하는 방법

위의 세가지 방법은 서로 조합되어 함께 사용될 수 있으나, 일반적으로 전처리 과정에서 선정된 피치 주기 후보들은 연속하지 않으므로 전처리 과정과 재귀적 계산법은 동시에 적용되지 못한다.

본 논문에서는 적응 코드북 탐색의 전처리를 통한 계산량 감소 방법중 역필터링을 이용한 피치 주기 후보 선택법에 대해 기술한다.

1) 역필터링을 이용한 후보 탐색

VSELP 부호화의 적응 코드북 탐색은 행렬식 3.9를 최소화하는 L^* 을 구한다.

$$E_L = \| p - \beta' b'_L \|^2 = \| p - H(\beta' b_L) \|^2 \qquad (3.9)$$

단, H 는 가중합성필터의 임펄스응답이다.

한편 가중합성필터 H(z)를 통과하여 적응 코드북 탐색의 목표 오류파 p를 생성해 낼 수 있는 이상적인 여기 신호(ideal excitation signal) x를 가정하고, 적응 코드북의 출력 $\beta' b_L$과 x와의 차를 d_L이라 한다. 이때 x는 p를 가중합성필터 H(z)의 역필터링(inverse filtering)한 결과이다[21].

$$d_L = x - \beta' b_L \quad , p = Hx \tag{3.10}$$

위에 식 3.9는 d_L를 이용하면 식 3.11과 같이 정리된다.

$$E_L' = \| p - H(x - d_L) \|^2 = \| Hd_L \|^2 \tag{3.11}$$

식 3.11에서 적응 코드북 탐색 과정은 d_L에 대한 가중합성필터 H(z)의 응답 Hd_L의 에너지를 최소화하는 과정으로 해석된다. 경험적으로 입력 신호 d_L의 에너지가 적을수록 필터 응답 Hd_L의 에너지가 적어질 가능성이 크므로 적응 코드북 탐색의 전처리 과정에서는 d_L의 에너지를 기준으로 K개의 피치 후보를 선택한다.

d_L의 에너지는 식 3.12와 같고, β'에 대한 편미분을 0으로 하여 정리하면 식 3.12를 최소화하는 과정은 식 3.13을 최대화하는 과정으로 전환할 수 있다.

$$\| d_L \|^2 = \| x - \beta' b_L \|^2 \tag{3.12}$$

$$\frac{(x^T b_L)^2}{\| b_L \|^2} = \| x \|^2 \cos^2 \omega_{x b_L} \tag{3.13}$$

결과적으로 적응 코드북 탐색의 전처리 과정에서 피치 주기 후보 선택의 기준은 적응 코드벡터 b_L과 이상적인 여기 신호 x 간의 편각 $\omega_{x b_L}$의 크기가 된다. 즉, 피치 주기 후보는 $\omega_{x b_L}$가 작은 순서대로 선택된다.

2) 전처리 과정을 통한 연산량 감소

행렬식 3.10에서 가중합성필터 H(z) 의 임펄스응답을 나타내는 행렬 H는 N 차원 삼각행렬(triangular matrix)이다. 따라서 이상적인 여기 신호 x는 역치환법(back substitution method)에 의해 N(N+1)/2 번의 곱셈-덧셈 연산을 사용하여 구할 수 있다. 표 3.1에 피치 주기 후보 선택의 전처리 과정에서 추가로 사용된 연산량을 나타내었다.

전처리 과정에서 K개의 피치 주기 후보가 결정되면 각 후보를 식 3.9에 적용하여 최적 피치 주기를 선택한다. 적응 코드북의 크기를 N_L이라고 할때, 일반적인 벡터 여기 부호화의 코드북 탐색 과정을 이용하여 적응 코드북을 탐색할 경우에 필요한 연산량은 N_LU로 결정된다. 그러나 전처리 과정에서 K개의 피치 후보를 미리 선택하면 탐색 연산량은 KU로 줄어든다. 코드 벡터의 길이 N 이 60이고 적응 코드북의 크기를 128이라 할때, 전처리 과정에서 12개의 피치 후보를 미리 선별하면 연산량은 총 40,870회로 기존의 방법에 비해 약 84%의 감소 효과를 얻는다. 또한 전처리 과정을 이용한 방법은 재귀적 방법을 이용한 적응 코드북 탐색법보다도 약 2배의 속도 향상을 얻을 수 있다.

〔표 3.1〕 적응 코드북 탐색의 전처리에 필요한 연산량

항목	연산량
x	$N(N+1)/2$
$x^T b_L$	N_LN
$\|\| b_L \|\|^2$	N_LN
$\frac{(x^T b_L)^2}{\|\| b_L \|\|^2}$	$2N_L$
총계	$(4N_L+N)(N+1)/2$

(단, N_L는 적응 코드북의 크기)

3.3 제안된 VSELP 부호화기

본 절에서는 3.2절의 여기 신호 탐색에서의 연산량 감소법을 이용하여 개선된 VSELP 부호화기를 제안한다. 표 3.2에는 제안된 VSELP 부호화기의 파라미터별 비트 할당량을 보인다. 제안된 부호화기는 8kHz 표본화하여 16비트로 선형 양자화된 입력 신호를 4.8kbps의 정보량으로 부호화하여 약 27 : 1로 정보량을 감소시킨다.

제안된 VSELP 부호화기의 각부 동작은 기본적으로 DoD 4.8kbps CELP 부호화기의 방법을 기반으로 한다[22]. 단기 예측 필터의 계수는 매 30ms 분석 프레임(frame)마다 10차 선형 예측 계수를 구한 후, LSP 파라미터로 변환하여 양자화한다. 적응 코드북 및 여기 코드북 탐색은 프레임내에 4개의 7.5ms 서브프레임(subframe)마다 탐색되어 부호화된다. 피치 이득 β 및 여기 코드벡터 이득 γ 는 피치 주기 및 여기 코드벡터 색인이 결정된 이후에 지각가중오류를 최소화하도록 동시에 결정된다. 여기 코드벡터 이득 γ는 VSELP 여기 코드북의 중복성을 이용하여 항상 양의 값을 갖도록 조정된다. 최종적으로 β 및 γ는 LBG(Linda-Buzo-Gray) 알고리즘으로 학습시킨 코드북을 이용하여 비선형 양자화된다. β 와 γ를 위한 코드북은 각각 5비트, 4비트의 색인을 사용한다.

(표 3.2) 개선된 VSELP 부호화기의 특징

	Linear Predictor	적응코드북	여기코드북
Update	30 ms	7.5 ms	7.5 ms
parameter	10 LSP	256 codeword	2048 codeword
Bits per Frame	34 (3,4,4,4,4, 3,3,3,3,3)	index: 8,6,8,6 gain: 5x4	index: 11x4 gain: 4x4

(프레임당 2비트는 사용하지 않음)

적응 코드북은 DoD CELP 부호화기에서 제안된 147샘플의 1차원 행렬을 사용하며, 과거 147샘플의 여기 신호를 저장한다. DoD CELP 부호화기에서는 16에서 143까지의 128개 정수 피치 주기와 128개의 소수 피치 주기를 사용하여 음질을 향상시키고 있다. 제안된 VSELP 부호화에서도 동일한 소수 피치 주기를 이용하며, 탐색 과정은 DoD CELP 부호화기의 소수 피치 주기 탐색법을 사용한다[22]. 한편 연산량의 감소를 위해 피치 주기는 홀수번째 프레임에서는 256개의 색인을 모두 탐색하고 짝수번째에서는 이전 프레임의 피치 주기를 중심으로 연속되는 64개의 피치 주기만을 탐색하는 델타 탐색법을 이용한다. 또한 소수 피치 주기 탐색은 1차 정수 피치 주기 탐색 결과로 얻은 최적 정수 피치 주기 주변만을 탐색하는 계층적 탐색(hierarchical search)으로 수행된다. 정수 피치 주기 탐색시에 제안한 역필터링을 이용한 후보 선택 방법을 적용하고, 선택된 후보중 최적 피치 주기를 일반적인 코드북 탐색 방법의 적합도 계산법을 이용하여 결정한다.

여기 코드북 탐색 방법은 근본적으로 기존의 VSELP 여기 코드북 탐색 방법을 따른다. 먼저 적응 코드북 탐색 결과 선택된 최적 적응 코드벡터의 가중합성필터응답 b_L를 계산한다. 기저 벡터의 가중합성필터 응답 계산과정은 중첩 희소 기저 벡터를 사용할 때의 재귀적 계산법을 도입하여 연산량을 감소시킨다. 다음으로 직교화 과정을 수행하고 R_m과 D_m을 계산한다. 적합도 계산과정 이전에 전처리에 의한 기저 벡터 부호 결정을 수행한다. 마지막으로 부호 결정이 안 된 나머지 기저 벡터들의 조합에 대해 적합도를 계산하여 최적 여기 코드벡터를 결정한다. 여기 신호 탐색 과정이 끝나면 적응 코드벡터의 이득 β과 여기 코드벡터 이득 γ 을 동시에 계산한다.

이상에서 제안된 VSLEP 부호화기의 여기 신호 탐색 방법은 다음의 4가지 내부 파라미터를 갖는다.

1. 중첩도 : 기저 벡터의 중첩정도로 인접 기저 벡터와 몇 샘플 이동하여 중첩되었는가로 측정된다.

2. 희소도 : 기저 벡터내의 각 샘플이 0일 확률로 측정된다.

3. 전처리 기저 벡터 결정 방법 : 기저 벡터 부호 결정의 전처리 과정에서 대상 기저 벡터를 선택하는 방법으로 SP 혹은 SC에 대한 상수 임계치를 사용하거나 K-선택법을 사용할 수 있다.

4. 피치 주기 후보 수 : 적응 코드북 탐색 전처리에서 선택할 피치 주기 후보의 수

본 논문에서는 음질의 저하를 최소화하면서 연산량의 감소를 최대화하는 적절한 내부 파라미터값을 실험적으로 유도하도록 한다.

IV. 실험 및 평가

본 장에서는 3장에서 제안된 세가지 연산량 감소 방법의 성능을 평가하고 제안된 각각의 방법에 필요한 내부 파라미터값을 실험적으로 결정하였다. 그리고 제안된 세가지 연산량 감소 방법을 적용한 VSELP 부호화기를 구현하고, 그 성능을 기존 VSELP 부호화기 및 DoD CELP 부호화기와 비교하였다.

4.1 실험 환경

제안된 연산량 감소 방법의 성능을 평가하기 위해서 각 방법을 적용한 VSELP 부호화기를 구현하였다. 구현된 부호화기는 미리 부호화하여 파일 형태로 저장된 음성 신호를 입력으로 사용하고 부호화기를 거쳐 재생된 음성 신호를 출력 파일에 저장한다. 부호화기 구현을 위해서 SUN sparc-II workstation이 사용되었다.

VSELP 부호화의 기저 벡터는 학습을 통해 최적화가 가능하다[4]. 그러나 본 실험에서는 실험의 용이성을 위해 모든 실험에 사용한 기저 벡터는 백색 잡음 신호를 이용하여 만들었고, 학습 과정은 고려하지 않았다.

3.3절에서 VSELP 부호화 과정중 피치 이득과 여기 코드북 이득의 부호화시에 비선형 양자화를 위한 코드북을 LBG 알고리즘으로 학습시키도록 제안한 바 있다. 이득 양자화를 위한 코드북은 새로운 기저 벡터를 사용할 때마다 학습되어야 하는 번거로움이 있다. 본 연구에서는 연산량 감소에 따른 음질의 변화 측정이 목적이므로 이득의 부호화 과정은 모든 실험에서 생략하였다. 다만 제안된 VSELP 부호화기와 타 부호화기의 성능 비교시에는 이득 부호화 방법을 적용하였다.

4.1.1 실험 자료

음성의 다양한 음운 환경 및 화자 특성을 반영하여 표 4.1과 같이 약 54초분의 음성을 실험 자료로 수집하고 모든 실험에서 입력 신호로 사용하였다. 자료 음성은 IBM-PC 환경의 음성 분석용 패키지인 CSL(computerized speech lab.)을 이용하여 표본주파수(sampling frequency) 8kHz에 16비트 양자화하는 PCM 방식으로 부호화하였다.

〔표 4.1〕 음성 자료

자료명	언어	화자성별	길이
문장 1	한국어	남	9.21 sec
문장 2	영 어	남, 여	29.20 sec
문장 3	한국어	여	8.19 sec
문장 4	한국어	남	2.55 sec
문장 5	한국어	남	4.80 sec

4.1.2 평가 방법

음성 부호화의 성능 평가에는 음질, 정보량, 처리 속도의 세가지 요소가 고려된다. 본 논문에서 제안된 VSELP 부호화의 연산량 감소 방법은 부호화의 정보량에는 전혀 변동을 주지 않고, 내부 파라미터값에 따라서 음질의 저하를 유발할 수 있는 특징이 있다. 따라서 제안된 여기 신호 탐색 방법의 평가 실험은 해당 내부 파라미터값을 변화시키면서 연산량 감소율을 계산하고 음질의 변화를 측정하여 비교하는 방식으로 진행되었다. 이러한 평가 실험의 비교 기준으로 연산량 감소율과 음질 저하 정도

를 고려한 식 4.1의 성능 평가 점수를 사용하였다.

SCORE = Q x (속도 향상) - (음질 저하) (4.1)

단, Q = (최대 허용 음질 저하) / (최대 속도 향상)

식 4.1에서 속도 향상은 제안된 방법의 연산량 감소 항목에 대해 기존 VSELP 부호화기 연산량과 제안된 방법 사용시의 연산량을 구해 그 비율로 측정하였다.

음질 측정 방법으로는 계산이 용이하고 주관적 음질 평가와 유사한 결과를 주는 SNRseg를 사용하였다. SNRseg는 SNR(signal to noise ratio)을 일정 시간 단위로 구하여 평균을 취하는 방법으로, SNR은 식 4.2과 같이 원음과 잡음의 에너지 비율로 계산되며, 이를 때 30ms마다 계산한 후 평균하였다.

$$SNR = 10\log_{10}\left(\frac{\sum_{t} S(t)^2}{\sum_{t} |S(t) - \hat{S}(t)|^2}\right) \text{ [dB]} \qquad (4.2)$$

단, S(t) 는 원음, $\hat{S}(t)$는 재생음

한편 최대 허용 음질 저하는 주관적인 판단에 의해 기존 VSELP 부호화기의 SNRseg값의 약 6%에 해당하는 0.5dB까지의 감소를 허용하였다.

식 4.1은 속도 향상에 대한 가중치 Q 를 사용하여 특정 내부 파라미터값으로 제안된 연산량 감소 방법의 최대 속도 향상을 얻어도 음질이 허용치 이하로 떨어지면 평가 점수는 음수값이 되도록 고안되었다. 한편, 실험에서 제안된 연산량 감소 방법 사용시 음질이 개선되는 경우에는 음질 저하 항목을 0으로 간주한다. 연산량 감소 방법의 성능 평가에 제안된 평가 점수를 이용함으로써 최대 허용 음질 저하의 한계를 항상 유지하면서 음질과 속도 증가 사항을 동시에 고려한 성능을 용이하게 평가할 수 있었다.

4.2 실험 결과

4.2.1 기존 VSELP 부호화기의 성능 평가

본 논문에서는 성능 비교 실험의 기준 시스템으로 기존의 여기 신호 탐색 방법을 이용한 VSELP 부호화기를 구현하고 성능을 평가하였다.

구현된 VSELP 부호화기의 적응 코드북 탐색시, 1차 정수 피치 주기 탐색에서는 전처리 과정을 거치지 않고 모든 피치 주기의 적합도를 계산하므로, 적응 코드북 자체의 중첩성을 이용한 재귀적 계산법을 적용할 수 있다. 즉, 피치 주기 L이 적응 코드벡터의 길이 60보다 클 때는 연속되는 적응 코드벡터가 1샘플씩 이동되어 중첩된 코드북의 코드벡터로 간주된다. 따라서 적응 코드벡터의 가중합성필터 응답을 계산할 때 중첩 코드북을 위한 재귀적 계산법을 적용하여 연산량을 감소시킨다.

표 4.2과 4.3에 기존 여기 신호 탐색법을 이용한 VSELP 부호화기의 적응 코드북 탐색 과정 및 여기 코드북 탐색 과정의 단계와 소요 연산량을 기술하였다. 또 연산량 감소 방법의 음질 평가 기준을 위해 기존 VSELP 부호화기를 이용하여 표 4.1의 음성 자료를 부호화한 결과를 표 4.4에 나타내었다. 이때 이득의 양자화는 수행하지 않았다. 기존 VSELP 부호화기의 평균 SNRseg은 8.387 dB였다.

〔표 4.2〕 기존 VSELP 부호화기의 적응 코드북 탐색의 연산량

항 목		곱셈-덧셈 연산량	예
1차 정수 피치 주기 탐색	(60≥L≥16)	45U	87,840 (77.9%)
	(143≥L>60)	83(3N+2)	15,106 (13.4%)
2차 소수 피치 주기 탐색		5U	9,760 (8.7%)
전체계산량			112,706 (100%)

〔표 4.3〕 기존 VSELP 부호화기의 여기 코드북 탐색의 연산량

항 목	곱셈-덧셈 연산량	예
$b'_L(n) = b_L(n) * h(n)$	N(N+1)/2	1830 (3.4%)
$q_m(n) = v_m(n) * h(n)$	MN(N+1)/2	20130 (37.9%)
$q'_m(n)$	2MN+ N	1380 (2.6%)
R_m	M(N+1)	671 (1.3%)
D_{mj}	M/2(M+1)(N+1)	4026 (7.6%)
$match_0$	M(M+3)/2 + 2	79 (0.1%)
$match_i$	$(2M+1)(2^{M-1}-1)$	23529 (44.2%)
β, γ	2N+M+2NM+12	1463 (2.8%)
전체계산량		53175 (100%)

〔표 4.4〕 기존 VSELP 부호화기의 음질 평가

음성자료	길 이	SNRseg
문장 1	9.21 sec	7.60 [dB]
문장 2	29.20 sec	8.26 [dB]
문장 3	8.19 sec	10.33 [dB]
문장 4	2.55 sec	6.38 [dB]
문장 5	4.80 sec	8.42 [dB]
전 체	53.98 sec	8.387 [dB]

4.2.2 기저 벡터의 중첩도와 희소도에 따른 음질 변화

기저 벡터의 가중합성필터 응답 계산시 연산량 감소를 위하여 제안한 중첩 희소 기저 벡터의 사용에 따른 음질의 변화와 연산량 감소율을 측정하였다. 이하 α 샘플씩 이동하여 중첩되고 희소도가 σ 인 중첩 희소 기저 벡터를 $v(\alpha, \sigma)$로 표기한다.

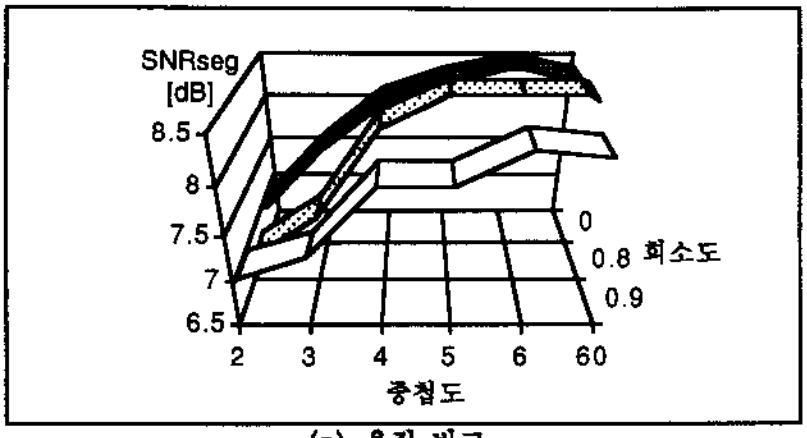
기저 벡터 $v(\alpha, \sigma)$를 사용하는 VSELP 부호화기의 속도 향상 정도는 식 4.3으로 결정되며, 본 실험에서는 최대 속도 향상을 $v(2, 0.9)$를 사용할 때의 속도 향상 66.7로 정한다. 따라서 중첩 희소 기저 벡터 이용법의 성능 평가 점수를 위한 계수 Q는 0.0075 (= 0.5 / 66.7)가 된다.

$$(\text{속도 향상}) = Oq / (1-\sigma)[N(N+1)/2 + (M-1)\{\alpha(\alpha+1)/2 + \alpha(N-\alpha)\}] \quad (4.3)$$

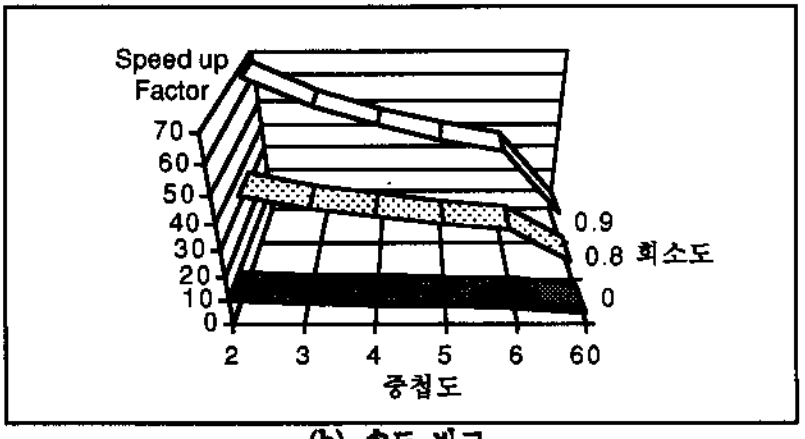
단, Oq는 기존 VSELP 부호화기의 q_m 계산 연산량 (= 20,130)

그림 4.1은 중첩시 이동시키는 샘플의 수를 2에서 6까지 변화시키고 희소도를 0.9, 0.8 및 0 으로 변화시킨 중첩 희소 기저 벡터를 이용하는 VSELP 부호화기의 음질, 속도 향상 및 성능 평가 점수이다.

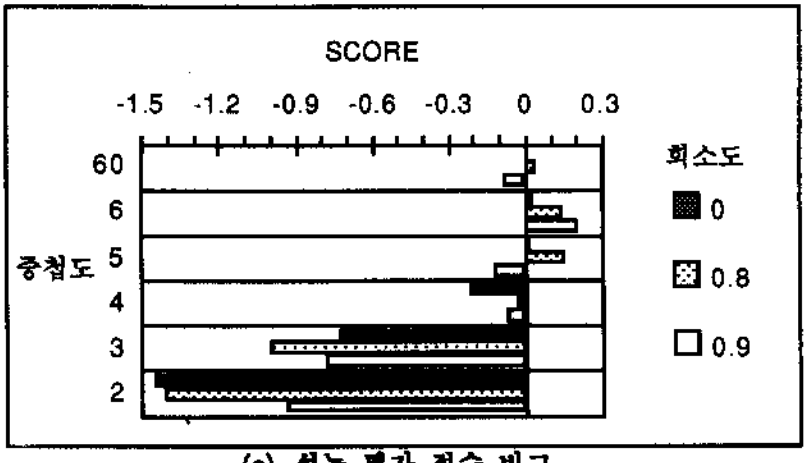
그림 4.1c에서 기저 벡터 $v(6, 0.9)$를 사용할 경우에 가장 높은 성능 평가 점수를 나타냈다. 이때 SNRseg가 8.305dB로 기존 VSELP 부호화기에 비해 0.08dB 감소하였고, q_m 계산 과정에서 38.1배의 속도 향상이 있었다.



(a) 음질 비교



(b) 속도 비교



(c) 성능 평가 점수 비교

〔그림 4.1〕 중첩 회소 기저 벡터 사용법의 성능 평가

4.2.3 전처리 기저 벡터 선택 방법의 비교

기저 벡터 부호 결정의 전처리 과정은 임의의 기저 벡터가 SP 및 SC에 대한 일정 조건을 만족시킬 때 선형 조합 계수 $\theta_{i,m}$를 R_m의 부호로 결정한다. 전처리 과정에서 K개의 기저 벡터의 부호가 미리 결정되면 $match_i'$의 계산시 속도 향상은 식 4.4와 같다. 최대 속도 향상은 K가 7일 경우인 68.2로 하고, 이때 성능 평가 점수의 계수 Q는 0.0073이 된다.

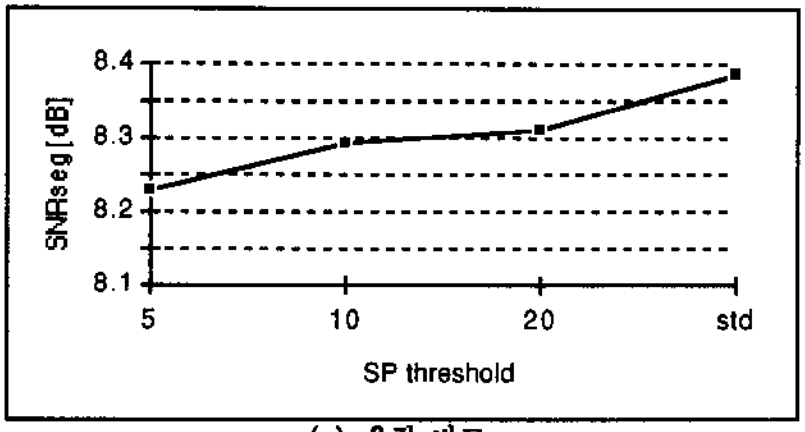
$$(\text{속도 향상}) = Om / (2^{M-k}-1)(2M+1) \qquad (4.4)$$

단, Om 는 기존 VSELP 부호화기의 $match_i'$ 계산 연산량 (= 23,529)

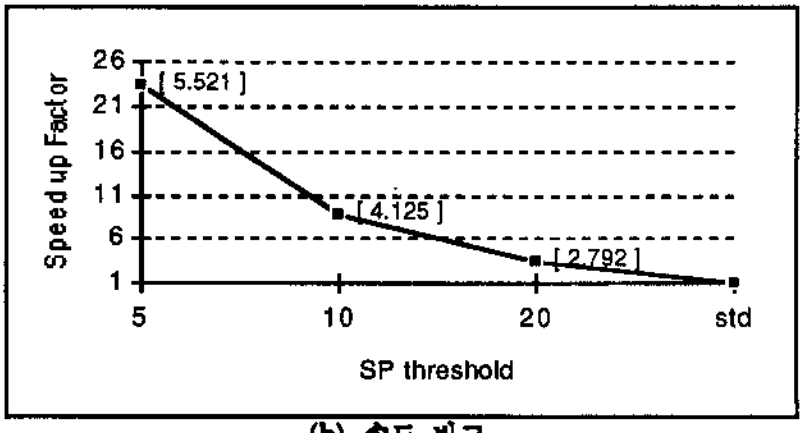
기저 벡터 부호 결정의 전처리 과정에 이용되는 각 기저 벡터의 SP 및 SC는 R_m 및 D_{mm}을 이용하여 계산되고, 추가로 $\| p \|^2$의 계산과 K개의 후보 선택 과정에 연산량이 소요되나 속도 향상 계산식에서 무시될 수 있는 정도이다.

그림 4.2와 4.3에 전처리 대상을 각각 SP 및 SC의 임계치를 이용하여 결정할 경우에 실험 결과를 나타내었다. 그림 4.2b와 4.3b에서 괄호안의 수는 평균적으로 전처리 과정에서 부호 결정된 기저 벡터의 수이다. 그림 4.4와 4.5는 각각 SP와 SC를 이용하여 그 값이 큰 K개의 기저 벡터를 전처리 대상으로 결정한 경우의 실험 결과이다.

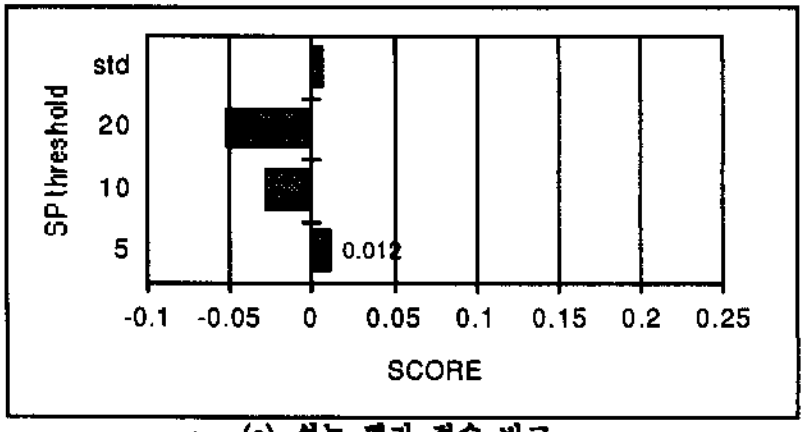
그림 4.2에서 4.5까지의 4가지 전처리 기저 벡터 결정 방법을 모두 고려해 보았을 때 SP값이 큰 6개의 기저 벡터를 전처리 대상으로 선택하는 방법이 가장 높은 성능 평가 점수를 나타냈다. 이때 음질의 저하는 기존 VSELP 부호화기에 비해 SNRseg가 0.17dB 감소되는 정도였고, $match_i'$ 계산 과정의 속도는 33배 향상되었다.



(a) 음질 비교

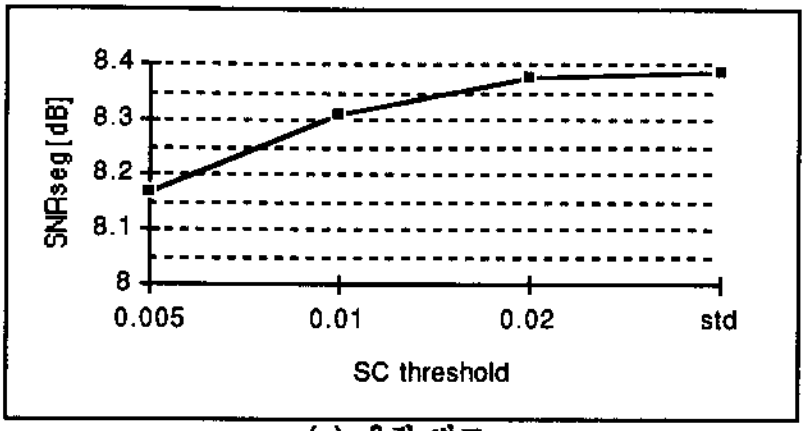


(b) 속도 비교

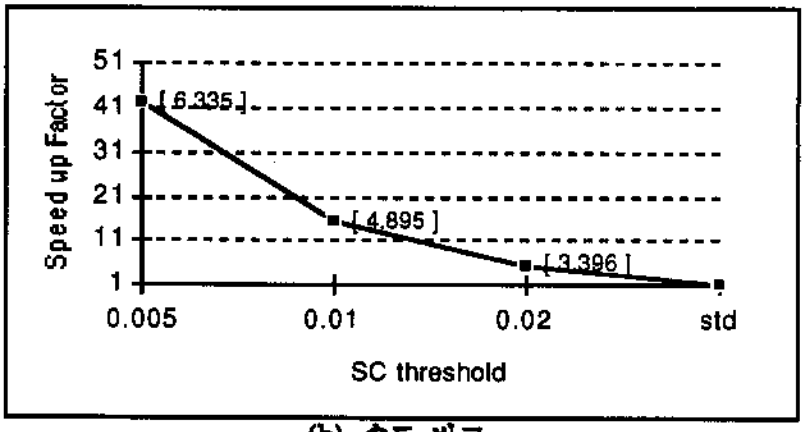


(c) 성능 평가 점수 비교

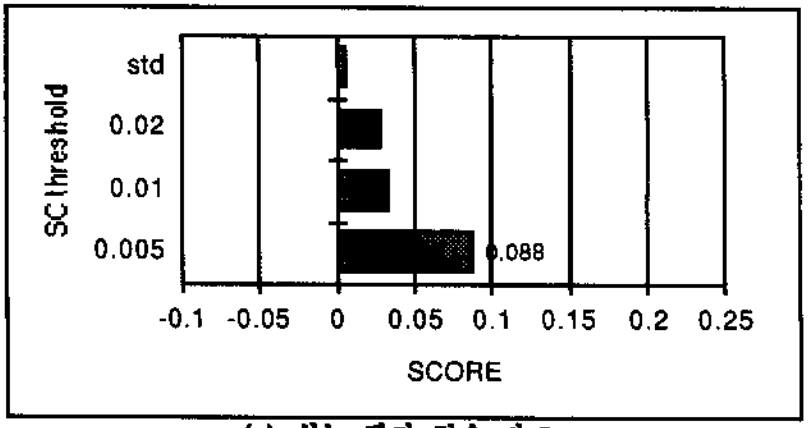
〔그림 4.2〕 SP의 임계치를 이용한 기저 벡터 부호 결정의 전처리



(a) 음질 비교

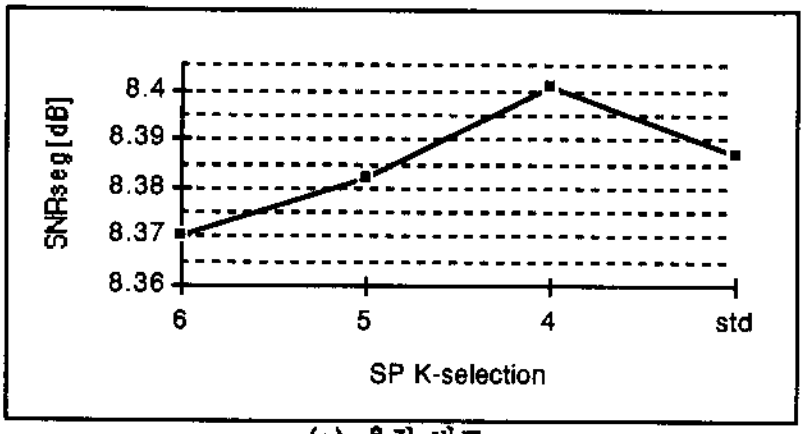


(b) 속도 비교

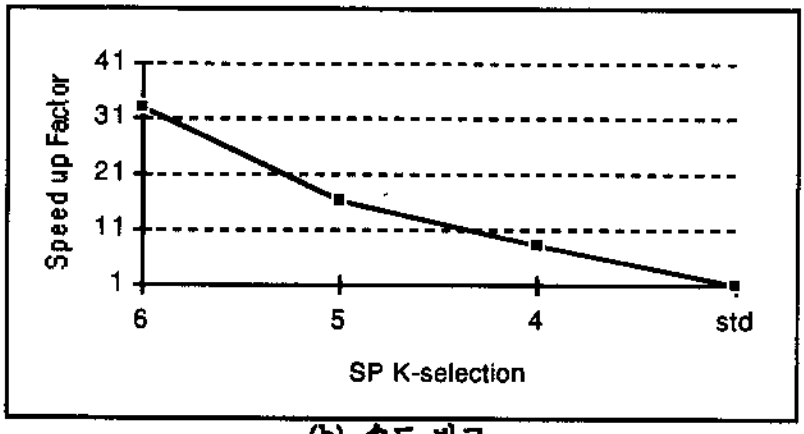


(c) 성능 평가 점수 비교

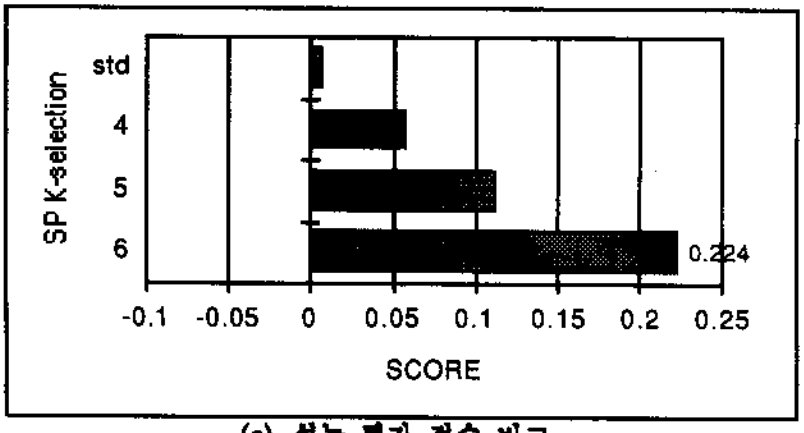
〔그림 4.3〕 SC의 임계치를 이용한 기저 벡터 부호 결정의 전처리



(a) 음질 비교

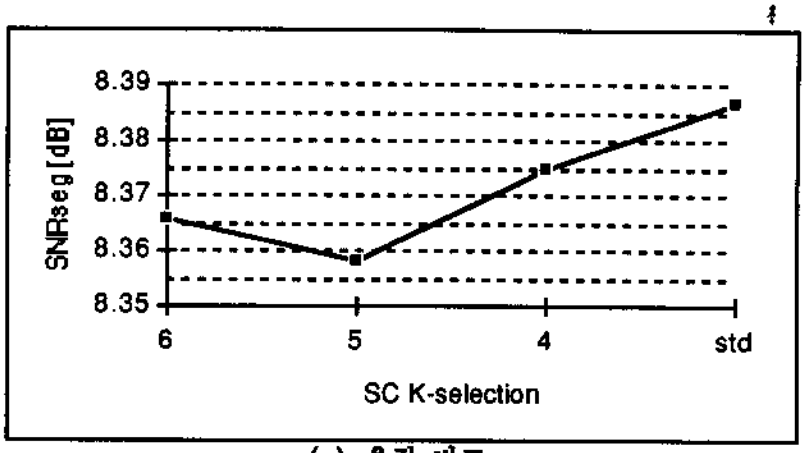


(b) 속도 비교

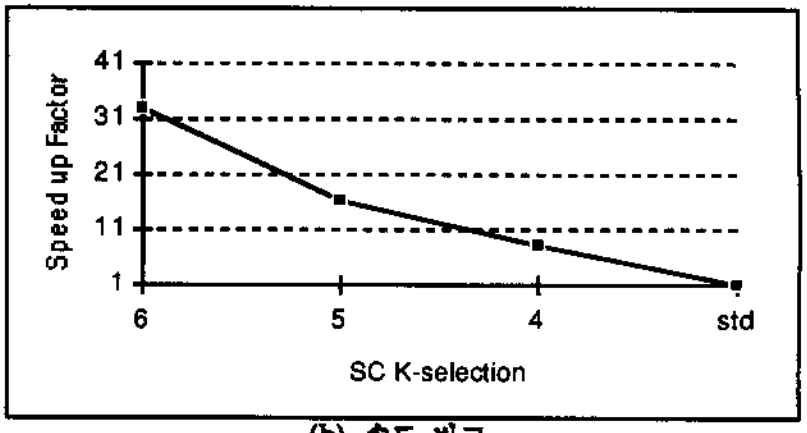


(c) 성능 평가 점수 비교

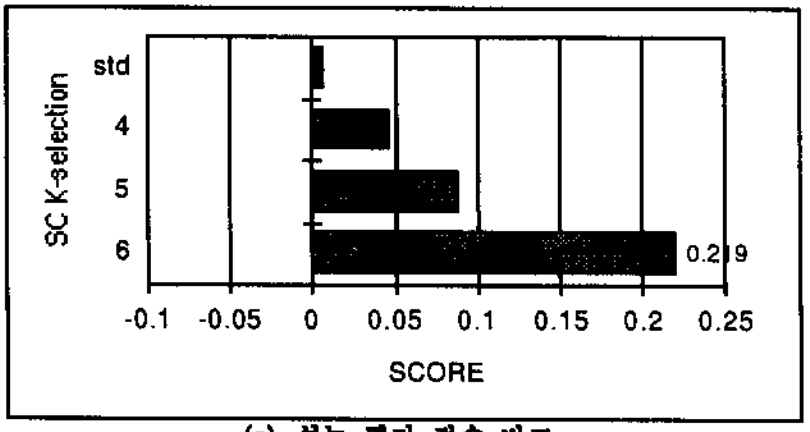
[그림 4.4] SP의 K-선택법을 이용한 기저 벡터 부호 결정의 전처리



(a) 음질 비교



(b) 속도 비교



(c) 성능 평가 점수 비교

〔그림 4.5〕 SC의 K-선택법을 이용한 기저 벡터 부호 결정의 전처리

4.2.4 적응 코드북 탐색 전처리의 유효성

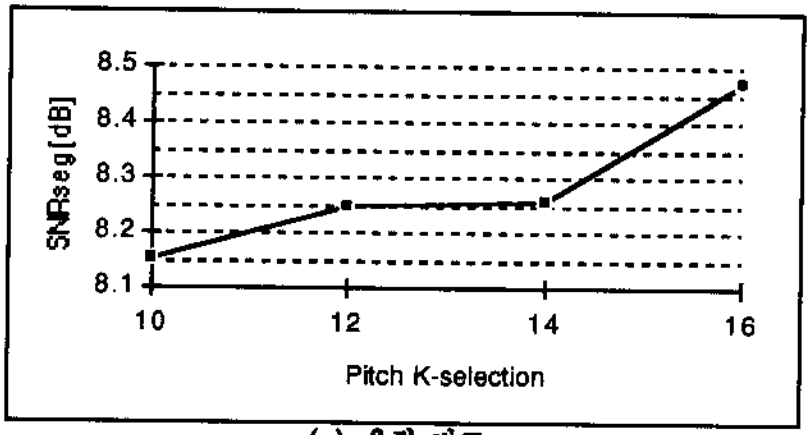
역필터링을 이용하여 구한 이상적인 잔차파 x 와의 편각을 기준으로 피치 주기의 후보를 선택하는 전처리 과정은 결정된 후보의 수에 따라 속도 및 음질에 영향을 준다.

전처리 과정에서 K개의 피치 주기 후보가 선택될 때 표준 VSELP 부호화기의 1차 정수 피치 주기 탐색 과정의 속도 향상은 식 4.5와 같다. 최대 속도 향상은 후보 수를 10으로 할 경우의 2.78로 정하였고, Q값은 0.18이다.

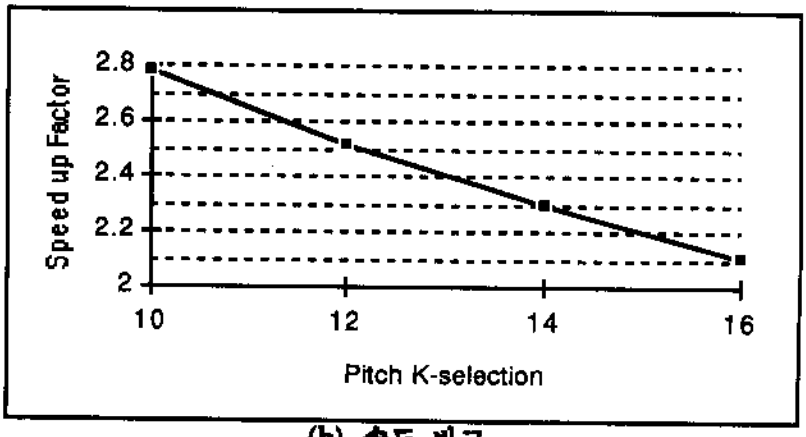
$$(\text{속도 향상}) = Op / [(4N_L+N)(N+1)/2 + KU] \qquad (4.5)$$

단, Op 는 기존의 정수 피치 주기 탐색의 연산량 (=102,946)

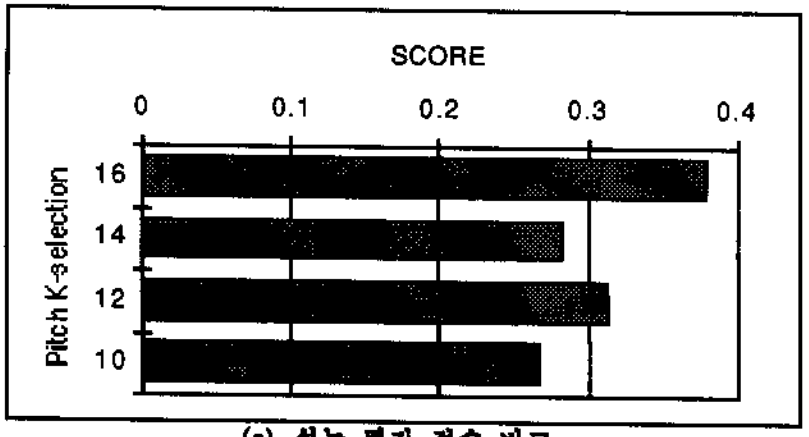
그림 4.6은 전처리 과정의 후보 수를 10에서 16까지 2씩 증가시킬 때 음질, 속도 향상 및 성능 평가 점수를 나타낸다. 그림 4.6c에서 적응 코드북 탐색의 전처리 과정의 후보 수를 16개로 정할 경우 성능 평가 점수가 가장 높았다. 이때 SNRseg는 8.471dB였고, 1차 정수 피치 주기 탐색의 속도를 2.11배 향상시켰다.



(a) 음질 비교



(b) 속도 비교



(c) 성능 평가 점수 비교

[그림 4.6] 적응 코드북 탐색 전처리 과정의 성능 평가

4.2.5 제안된 VSELP 부호화기의 성능 평가

본 논문에서 제안한 세가지 연산량 감축 방법의 실험 결과, 각각의 방법에 대해 가장 높은 성능 평가 점수를 받은 내부 파라미터값을 이용하여 VSELP 부호화기를 구현하였다. 구현된 VSELP 부호화기에서 사용한 내부 파라미터값은 다음과 같다.

1. 중첩도 : 6샘플씩 어긋나게 중첩시킨 기저 벡터 이용
2. 희소도 : 전체 샘플의 90%가 0인 기저 벡터 사용
3. 전처리 기저 벡터 결정 방법 : SP 값이 큰 순으로 6개의 기저 벡터를 기저 벡터 부호 결정 전처리의 대상으로 결정
4. 피치 주기 후보 수 : 적응 코드북 탐색의 전처리 과정에서 16개의 피치 주기 후보를 선택함

표 4.5와 4.6에는 제안된 VSELP 부호화기의 적응 코드북 탐색 과정 및 여기 코드북 탐색 과정의 단계와 소요 연산량을 기술하였다. 제안된 VSELP 부호화기는 기존 방식에 비해 적응 코드북 탐색 과정에서 1.9배, 여기 코드북 탐색 과정에서 4.9배의 속도 향상을 보인다.

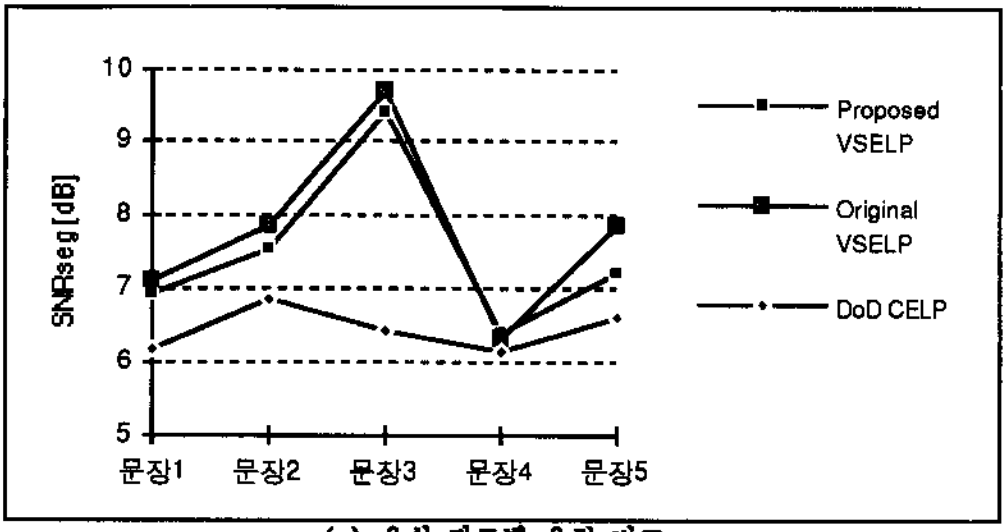
〔표 4.5〕 제안된 VSELP 부호화기의 적응 코드북 탐색의 연산량

항 목		곱셈-덧셈 연산량	예
1차 정수 피치 주기 탐색	전처리	$(4N_t+N)(N+1)/2$	17,446 (29.9%)
	후처리	KU (K=16)	31,232 (53.4%)
2차 소수 피치 주기 탐색		5U (변동 없음)	9,760 (16.7%)
전체계산량			58,438 (100%)

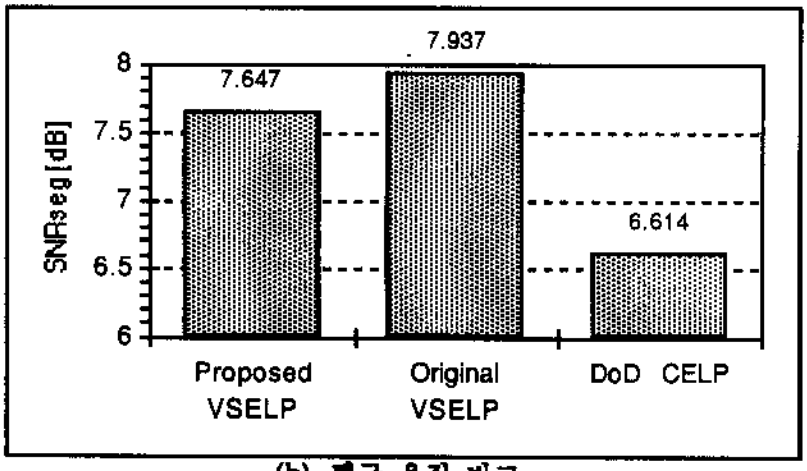
〔표 4.6〕 제안된 VSELP 부호화기의 여기 코드북 탐색의 연산량

항 목	곱셈-덧셈 연산량	예
$b'_L(n) = b_L(n) * h(n)$	N(N+1)/2	1830 (17.0%)
$q_m(n) = v_m(n) * h(n)$ ($\sigma = 0.9, \alpha = 6$)	$(1-\sigma)[(N+1)N/2 + (M-1)\{\alpha(\alpha+1)/2 + \alpha(N-\alpha)\}]$	528 (4.9%)
$q'_m(n)$	2MN+ N	1380 (12.8%)
R_m	M(N+1)	671 (6.2%)
D_{mj}	M/2(M+1)(N+1)	4026 (37.4%)
$\|\| p \|\|^2$	N	60 (0.6%)
SP	2M	22 (0.2%)
K selection	KM (비교 연산)	
$match_0$	M(M+3)/2 + 2	79 (0.7%)
$match_i$	$(2M+1)(2^{M-k}-1)$	713 (6.6%)
β, γ	2N+M+2NM+12	1463 (13.6%)
전체계산량		10772 (100%)

그림 4.7은 제안된 VSELP 부호화기의 성능을 기존 VSELP 부호화기 및 DoD CELP 부호화기와 비교한 결과이다. DoD CELP는 현재 미국 연방 무선 통신을 위한 음성 부호화의 표준안으로 채택되어 사용되고 있다. 실험 결과 제안된 VSELP 부호화기는 기존 VSELP 부호화기에 비해 SNRseg가 0.29dB가 감소하는 음질 저하를 보였으나, 청취 실험 결과 거의 구분되지 않는 정도였다. 벡터 여기 부호화 방법의 하나인 DoD CELP 부호화기에 비해서는 SNRseg가 1.3dB가량 큰 우수한 음질을 나타냈다. 한편 제안된 VSELP 부호화기의 처리 속도는 기존 VSELP 부호화기에 비해 평균 2배 가량 향상되었다.



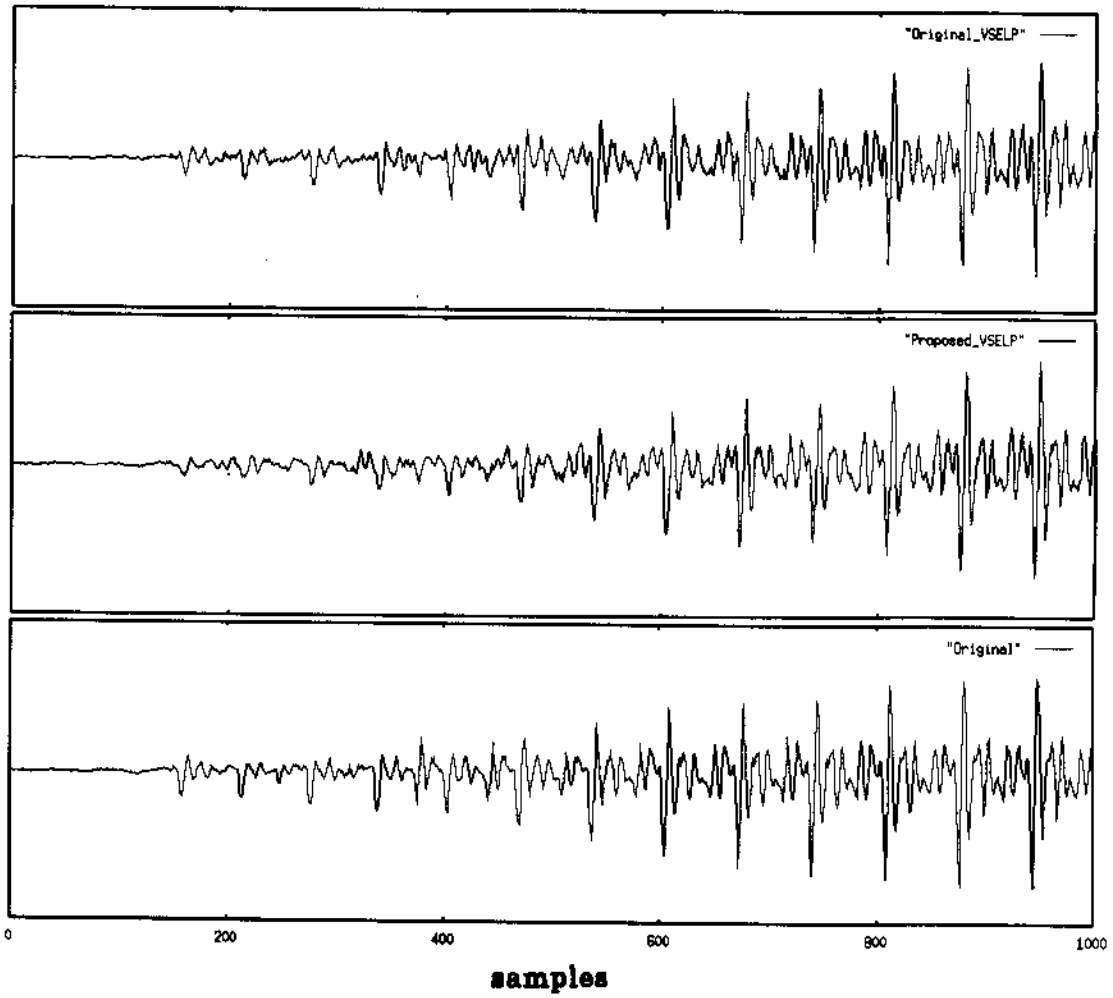
(a) 음성 자료별 음질 비교



(b) 평균 음질 비교

〔그림 4.7〕 제안된 VSELP 부호화기의 음질 평가

그림 4.8에는 남성 화자가 발성한 모음 '아'의 전반부를 기존 VSELP 부호화기 및 제안된 VSELP 부호화기로 부호화한 결과를 나타낸다. 그림 4.8에서 두 부호화 방법이 뚜렷한 차이없이 원음의 파형을 재현하고 있음을 확인할 수 있다.



〔그림 4.8〕 VSELP 부호화기 재생음의 파형 비교

V. 결론 및 검토

본 논문에서는 다양한 음성 부호화 방법중 정보량과 음질면에서 우수한 성능을 보이는 VSELP 부호화의 처리 속도를 향상시키기 위한 연구에 대하여 기술하였다. VSELP 부호화의 대부분의 연산량은 입력 신호와 유사한 합성 신호를 생성하기 위한 적응 코드북 탐색 및 여기 코드북 탐색 과정에서 사용된다. 본 논문에서는 VSELP 부호화의 적응 코드북 탐색 및 여기 코드북 탐색 과정에 필요한 연산량을 감소시키기 위하여 세가지 방법을 제시하고 이를 VSELP 부호화기 구현에 적용하였다. 성능 평가 실험을 통해 제안된 방법들이 기존의 탐색 방법에 비해 음질을 크게 저하시키지 않으면서 효과적으로 연산량을 감소시켜 처리 속도가 향상됨을 확인하였다.

VSELP 부호화의 여기 코드북 탐색 과정에서는 기저 벡터의 가중합성필터 응답을 구하는 과정과 각 여기 코드벡터에 대한 적합도 계산 과정에 많은 연산량이 소요된다. 본 논문에서는 먼저 중첩되고 희소한 기저 벡터를 이용하는 여기 코드북을 제안하여 기저 벡터의 필터 응답 계산 과정의 연산량을 감소시켰다. 또한, 적절한 중첩도 및 희소도 결정을 위해 다양한 기저 벡터를 이용한 부호화 실험을 수행하였다. 실험 결과 6샘플씩 이동되어 중첩되고 기저 벡터내에 90%의 샘플이 0으로 채워진 중첩 희소 기저 벡터를 사용할 때, 기존의 필터 응답 계산과정의 연산량을 97%까지 감소시켰고, 이때 음질은 기존의 방법에 비해 SNRseg이 0.08dB 감소하였으나 청취 실험에서는 거의 구별할 수 없었다.

두번째로 기저 벡터에 대한 선형 조합 계수를 결정하는 전체 여기 코드벡터의 적합도 계산 과정에서는, 전처리 과정을 이용하여 음질 저하를 초래하지 않는 기저 벡터 부호를 미리 결정하는 방법을 제안하였다. 결과적으로 전처리 과정 이후에 여기 코드북 탐색 범위가 줄어들어 연산량이 감소되었다. 실험에서는 기저 벡터 부호 결정 전처리의 대상을 선별하기 위해 제안된 4가지 조건을 이용한 부호화를 수행하였다.

실험 결과 각 기저 벡터에 대한 가중합성필터응답의 목표오류파 방향 성분의 크기를 이용한 K-선택법이 가장 좋은 결과를 나타내었고, K를 6으로 할 때 음질의 저하는 SNRseg가 0.17dB 감소하는 정도였다.

마지막으로 적응 코드북 탐색 과정에 역필터링을 이용한 후보 선택 과정을 전처리 과정으로 도입하여 연산량을 감소시켰다. 후보의 수를 변화시키면서 부호화 실험을 수행한 결과, 전처리 과정에서 16개의 후보를 미리 선별하면 음질의 저하없이 1차 정수 피치 주기 탐색 과정의 연산량을 52.7% 감소시킬 수 있었다.

본 논문에서는 제시한 세가지 연산량 감소 방법을 이용하여 VSELP 부호화기를 구현하고, 이를 기존 VSELP 부호화 및 타 부호화 방법과 비교하였다. 제안된 VSELP 부호화기는 적응 코드북 및 여기 코드북 탐색 과정에서 기존의 방법에 비해 각각 48.2% 및 79.7%의 연산량 감소 효과를 보였다. 그리고 음질면에서는 기존의 VSELP 부호화기에 비해 약간의 음질 저하가 있었으나 청취 실험에서는 거의 구별되지 않았고, DoD 4.8kbps CELP 부호화기에 비해서는 우수한 성능을 보였다.

본 논문에서는 연산량을 감소시키는 세가지 방법을 평가하기 위해 음질과 속도 향상 정도를 동시에 고려한 평가 점수를 설정하였다. 그러나 본 평가 방법에 대해서는 앞으로 음질 측정법이나 평가 요소별 가중치등에 개선의 여지가 있다. VSELP 부호화기를 구현할 때, 피치 이득과 여기 코드벡터 이득은 서로 독립적으로 양자화하였으나, 이를 동시에 벡터 양자화할 경우 음질에 개선이 있을 것으로 기대된다[4]. 또한, 본 연구에서는 제안된 중첩 회소 기저 벡터의 작성을 위해 난수 발생기를 이용하였으나, 학습을 통해 최적 기저 벡터를 생성하는 방법이 연구되어야 할 것으로 판단된다.

참고 문헌

[1] S. Saito, K. Nakata, Fundamentals of Speech Signal Processing , ACADEMIC PRESS, 1985

[2] G. Davidson, M. Yong, A. Gersho, "Real-time vector excitation coding of speech at 4800 bps", Int. Conf. Acoust. Speech. Signal Processing, 1987, pp.2189~2192

[3] Proposed EIA/TIA Interim Standard, "Wideband Spread Spectrum Digital Cellular System Dual-Mode Mobile Station - Base Station Compatibility Standard", April 21, 1992

[4] I. A. Gerson and M. A. Jasiuk, "Vector Sum Excited Linear Prediction (VSELP) Speech Coding at 8kbps", Int. Conf. Acoust. Speech. Signal Processing, 1990, pp.461~464

[5] G. Roy and P. Kabal, "Wideband CELP Speech Coding at 16kbits/sec", Int. Conf. Acoust. Speech. Signal Processing, 1991, pp.17~20

[6] J. H. Chen, M. J. Melchner, R. V. Cox and D. O. Bowker, "Real-Time Implementation and Performance of a 16kbps Low-Delay CELP Speech Coder", Int. Conf. Acoust. Speech. Signal Processing, 1990

[7] Motorola Inc. , "Vector Sum Excited Linear Prediction (VSELP) 13000 Bit Per Second Voice Coding Algorithm Including Error Control for Digital Cellular", Technical Description, Oct 18, 1989

[8] S. Furui, Digital Speech Processing, Synthesis and Recognition , Dekker, 1991

[9] 오 영 환, "무선 통신을 위한 음성압축 알고리즘의 개발", 삼성정보통신연구소 위탁과제 최종보고서, 1993

[10] 김 승 주, 오 영 환, "프레임 보간을 이용한 가변율 CELP코딩 ", 제10회 음성통신 및 신호처리 워크샵 논문집, pp.190~194, 1993

[11] 박승종, 김승주, 오영환, "중첩 희소 기저 벡터를 이용한 VSELP 부호화에 관한 연구", 충청지부 정보과학회 추계학술발표논문집, pp.87~91, 1993

[12] 성기철, 문상재, "VSELP 부호화기의 음질 및 주기탐색 개선에 관한 연구", 신호처리합동학술대회 논문집 제6권 1호, pp.49~53, 1993

[13] 이진익, "부호여기 선형예측 부호화기의 연산량 감축 및 음질 향상에 관한 연구", 한국과학기술원 전자공학과 박사학위논문, 1989

[14] 배성근, 백금란, 배명진, "CELP형 보코더에서 피치 검색오류의 개선", 제10회 음성통신 및 신호처리 워크샵 논문집, pp.185~189, 1993

[15] I. Trancoso, B. Atal, "Efficient Procedures for Finding the Optimum Innovation in Stochastic Coders", Int. Conf. Acoust. Speech. Signal Processing, 1986, pp.2375~2378

[16] J. Chung, R. Schafer, "Vector Excitation Homomorphic Vocoder", in Advances in Speech Coding, Kluwer Academic Publishers, pp.235~244

[17] S. Hoseini, M. Nakhai, M. Tabiani, "Real-time Implementation of the Homomorphic vocoder", Int. Conf. Signal Processing Applications & Technology, pp.1391~1399, 1993

[18] F. Jean, C. Kuo, H. Wang, "Spectral Sensitivity Weighted Transform Coding for LSP Parameters", EUROSPEECH '93, pp.1089~1092

[19] T. Svendsen, "Efficient Quantization of Speech Spectral Information", EUROSPEECH '93, pp.1143`1149

[20] G. Davison, A. Gersho, "Complexity reduction Methods for Vector Excitation Coding", Int. Conf. Acoust. Speech. Signal Processing, 1986, pp.3055~3058

[21] M.E. Ahmed, "Fast Methods for Code Search in CELP", IEEE trans. speech and audio processing, Vol. 1, NO. 3, 1993, pp.315~325

[22] Federal Standard 1016, "Analog to Digital Conversion of Radio Voice by 4.8kbps CELP". Feb. 14, 1991

[23] V. Ramamoorthy and N. Jayant, "Enhancement of ADPCM Speech by Adaptive Postfiltering", AT&T Bell Lab tech. report Vol. 63, No. 8, Oct. 1984

[24] K. M. Ponting, S. M. Peeling, "The use of variable frame rate analysis in the speech recognition", Computer Speech and Language, pp.169~179, Apr. 1991

[25] B. H. Juang, D. Y. Wong, A. H. Gray. Jr., "Distortion performance of vector quantization for LPC voice coding", IEEE Trans. ASSP, pp.294~304, Apr. 1982

[26] B. S. Atal, Advances in Speech Coding, Kluwer Academic Publishers, 1991

[27] S. Furui, "On the role of spectral transition for speech perception", J. Acoust. Soc. Am, Oct., pp.1016~1025, 1986

감사의 글

본 논문이 완성되기까지 따뜻한 격려와 지도를 아끼지 않으신 오영환 교수님께 진심으로 감사를 드립니다. 그리고 바쁘신 중에 심사를 위해 수고하여 주신 맹승렬 교수님과 이광형 교수님께도 깊은 감사를 드립니다.

연구에 많은 조언과 독려를 해주신 삼성전자의 김홍국 부장님과 음성에 대한 부족한 이해를 자상한 설명으로 보충해 주신 전자통신연구소의 지민제님께도 깊은 감사를 드립니다.

처음부터 끝까지 세심하게 논문 지도를 해준 실험실 선배 하진형, 환진형, 연준형에게 감사드리고, 언제나 관심있게 지켜봐주신 명현형, 호승형, 상문형과 졸업하신 영선형, 정석형에게도 고마운 마음을 전합니다. 실험실의 크고 작은 일들을 함께한 동료 성진과 후배 승종, 상호에게 고맙다는 말을 전하고 싶고, 그외에 보람차고 즐거운 과학원 생활이 되도록 함께해준 모든 분들께 감사를 드립니다.

멀리서 격려와 사랑을 보내주신 부모님께 가장 큰 감사를 드립니다. 부모님의 사랑은 어려운 시기에 항상 큰 힘이 되었읍니다. 형의 역할도 제대로 하지 못한 저를 잘 따라주고, 자신의 일에 최선을 다하는 동생 승욱, 승언에게 미안함과 고마운 마음을 함께 전합니다.

이 력 서

성 명 : 김 승 주

생년월일 : 1970년 4월 8일

출생지 : 서울특별시

본 적 : 제주도 남제주군 성산읍 수산리 552번지

학 력

1988. 3 ~ 1992. 2 : 한국과학기술대학 전산학과 졸업 (B.S.)

1992. 3 ~ 1994. 2 : 한국과학기술원 전산학과 졸업 (M.S.)