

QCELP에서 중첩된 코드북 검색의 개선

박 광 철[†] · 한 승 진^{††} · 이 정 현^{†††}

요 약

본 논문에서는 이동통신 시스템에서 사용할 수 있는 잡음에 강인하면서 음질이 개선된 QCELP 코드북 검색을 제안한다. 기존 QCELP는 고정 코드 북을 한 번 검색하지만, 본 논문에서는 두 번에서 다섯 번까지 검색을 하고 이를 실험해 본 결과 두 번 검색이 전송률에 따른 음질향상이 최적임을 알게 되었다. 따라서 본 논문에서는 두 번의 정밀 양자화를 통해 여기신호를 상세히 나타내으로써 음질을 향상시키는 개선된 QCELP 부호화기를 제안한다. 실험에서 잡음을 고려하지 않은 환경(강의실, 가정집, 거리, 연구실 등)에서 얻은 음성을 입력자료로 사용하고, 음질은 SNR과 segSNR을 이용하여 추정하였다. 실험 결과, 개선된 QCELP는 기존 QCELP 보다 SNR, segSNR에서 각각 38.35%, 65.51% 향상되었다.

키워드 : IMT-2000, CDMA, QCELP, 코드북, 부호기, 복호기, LPC

Improvement of Overlapped Codebook Search in QCELP

Kwang-Cheol Park[†] · Seung-Jin Han^{††} · Jung-Hyun Lee^{†††}

ABSTRACT

In this paper, we present the advanced QCELP codebook search improving the qualification of speech, which can make QCELP vocoder used in noise robust system. While conventional QCELP usually searches stochastic codebook once, we can find that two times search is the most suitable for improving the quality of speech after we did 2~5 times search. Consequently, the advanced QCELP vocoder represents excitation signal in detail using two times precise quantization and so improve the qualification of speech. In our experiment, we use the speeches collected from circumstance (such as lecture room, house, street, laboratory etc.) without regarding noise as input data and measure the speech qualification using SNR, segSNR. As the result of the experiment, we find that the advanced QCELP makes SNR and segSNR improved by 38.35% and 65.51% respectively compared with conventional QCELP.

Key word : IMT-2000, CDMA, QCELP, CODEBOOK, ENCODER, DECODER, LPC

1. 서 론

사람들은 일반적으로 전화 대화에서 50%이상의 휴지시간을 갖는다[1]. 이것을 효율적으로 이용하도록 제안한 알고리즘이 QCELP이다. 음성이 있다고 판단되면, 800MHz 대역에서의 QCELP는 8kbps 전송률로 분석/합성법을 사용하여 음성을 부호화한다. 그리고 배경 잡음만 있다고 판단되면 4kbps, 2kbps, 1kbps 전송률로 낮추어 가면서 분석/합성법을 사용하여 음성을 부호화한다[2]. 고정 코드북 검색은 부호화할 때 전송률이 8kbps인 경우 8번, 4kbps인 경우 4번, 2kbps인 경우 2번, 1kbps인 경우 1번씩 각 프레임마다 검색한다[3].

QCELP 음성 부호화기는 크게 STP(short-term prediction), LTP(long-term prediction), 고정 코드 북 검색의 과

정을 거쳐서 음성을 부호화한다. STP과정에서는 선형 예측 계수를 구하고, LPC계수를 LSP로 양자화 하여 통신 에러에 강인한 음성 부호기를 만든다. LTP 과정에서는 적응 코드 북을 이용해 피치 지연과 피치 이득을 구한다. 마지막으로 고정 코드북 검색과정에서는 코드북 인덱스와 이득을 구해 패킷화해서 전송한다. 복호화기에서는 이 과정을 역으로 수행해 합성 음성을 만든다.

QCELP의 다른 부호기와 구별되는 가장 큰 특징은 가변 전송률로 음성을 전송하는 것이다[4]. 이는 (그림 1)과 같이 음성 에너지의 값에 따라서 전송률을 8k, 4k, 2k, 1k로 전송을 하는 것이다. 기본적으로 음성에 배경 잡음 에너지가 있다고 가정하고, 각 전송률에 따른 문턱 값을 정한 후 음성 에너지 값에 따라 전송률을 정한다. 다음 프레임에서는 배경잡음 에너지를 증가시킴에 따라 각 전송률 문턱 값도 증가시킨다.

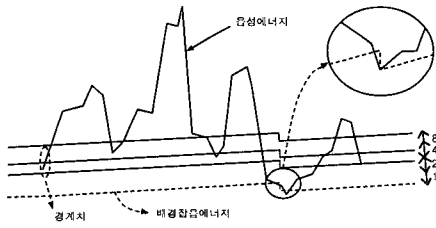
이렇게 일반적인 상황에서는 각 전송률에 따른 문턱 값이 계속 증가하지만 음성 에너지 값이 배경잡음 에너지 값

[†] 준 회 원 : LG전자 디지털미디어연구소 연구원

^{††} 정 회 원 : 인하대학교 대학원 전자계산공학과 박사수료

^{†††} 종신회원 : 인하대학교 전자전기컴퓨터공학부 교수

논문접수 : 2000년 5월 24일, 심사완료 : 2001년 1월 3일



(그림 1) QCELP의 전송률 결정 방법

보다 작아지게 되면 전송률은 최저 전송률을 선택하고, 배경잡음 에너지도 지금 측정된 음성 에너지 값으로 대체된다. 그러므로 배경 에너지 값은 작아지게 되고, 따라서 각 전송률 문턱 값도 같은 비율로 줄어들게 된다. 이렇게 유동적으로 문턱 값이 변하므로 주변의 잡음, 즉 일정한 에너지를 가진 소리들(예를 들어 달리는 자동차 소리, 주변의 음악 소리, 주변의 떠드는 소리 등)은 음성을 부호화하는데 적은 영향을 주게 된다. 반면에 갑작스런 소리(예를 들어 경적 소리, 폭발음 등)는 음성 부호화에 많은 영향을 준다. 따라서, 본 논문에서는 이러한 기존의 QCELP를 개선하여, 잡음에 강인한 개선된 QCELP 중첩 코드북 검색 방법을 제안한다.

2. 개선된 QCELP 부호기 및 복호기

IMT-2000시스템은 음성, 영상, 및 멀티 미디어를 전송할 수 있도록 전송 대역이 매우 넓어진다[5]. 그러므로 음성 부분에 대한 사용자의 요구사항은 음질에 관심을 가질 것이다. 기존 QCELP는 가변 전송률로 현재의 PCS인 협대역 CDMA 방식에 적합하다. 그러나 피치가 높은 여성음을 재생할 때 음질이 저하된다는 단점을 가지고 있다. 양질의 서비스를 제공하기 위해서는 이 점을 해결해야 한다.

음질을 향상시키는 방법으로는 인지적 가중 필터(perceptual weighting filter)와 postfilter를 사용하는 방법이 있다. 인지적 가중 필터는 인코딩을 할 때 LPC 분석에 의해 결정된 포먼트들에 해당하는 주파수 영역들의 영향을 줄임으로써 포먼트에서의 잡음을 감소시켜 음질을 향상시킨다. postfilter는 디코딩을 할 때 스펙트럼 계곡(영점)들을 약화시키고 포먼트들을 강조함으로써 음질을 향상시킨다. 비정상성(non-stationary)이 두드러지게 나타나는 구간에서는 20ms~30ms의 프레임에 대한 여기 신호를 부호화하기 위해 사용되는 비트를 동적으로 할당함으로써 음질을 향상시킬 수 있다.

본 논문에서 제안하는 개선된 QCELP는 음질을 향상시키기 위해서 여기 신호를 기존의 QCELP 방식보다 상세히 표현한다. 처음 코드북 검색을 마치고 두 번째 코드북 검색은 처음 검색한 후 발생한 잔차 신호보다 에러가 작은 잔차 신호를 양자화한다.

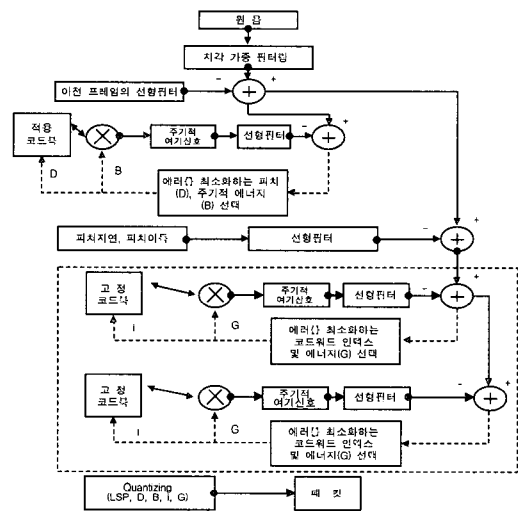
2.1 개선된 QCELP 부호기

지각 가중 필터링은 식 (2.1)을 이용하여 구하고, 선형 예측 계수(선형 예측 필터) LPC(Linear Predictive Coefficients)는 식 (2.2)를 이용하여 구한다.

$$W(z) = \frac{A(z)}{A(z/\gamma)} \tag{2.1}$$

$$= \frac{1 - \sum_{i=1}^p a_i z^{-i}}{1 - \sum_{i=1}^p a_i \gamma^i z^{-i}} \quad 0 \leq \gamma \leq 1$$

$$A(z) = 1 - \sum_{k=1}^b a_k z^{-k} \tag{2.2}$$



(그림 2) 개선된 QCELP 부호기

피치 지연(L또는 D)과 이득 (b)는 식 (2.3)을 이용하여 구한다.

$$P(z) = 1 - bz^{-L} \tag{2.3}$$

위의 식을 이용하여 역 필터로 여기 신호를 추출한 후 코드북 검색 시에 식 (2.4)를 이용하여 가장 작은 에러를 생성하는 코드북 인덱스와 이득을 찾는다.

$$E_c = \sum_{n=0}^{N_c-1} (x(n) - G\hat{y}_c(n))^2 \tag{2.4}$$

여기서 x(n)은 타겟 신호이고, $\hat{y}_c(n)$ 은 합성신호이다. 식 (2.4)를 최소화시키는 것은 식 (2.5)를 최소화하는 것과 같다[6].

고정 코드북 검색을 수행하고 여기 신호에서 잔차 신호를 뺀 잔여 신호를 다시 한 번 AbS(Analysis by Synthesis) 방법에 의해서 양자화를 한다.

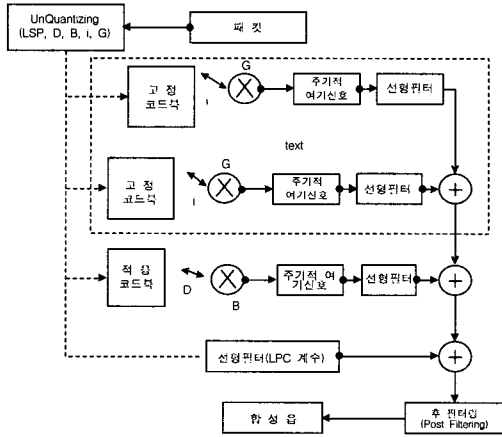
$$-2bE_{xxC} + b^2 E_{yyC}$$

$$E_{xxC} = \sum_{n=0}^{G_c-1} x(n)y_c(n)$$

$$E_{yyC} = \sum_{n=0}^{G_c-1} y_c(n)y_c(n) \tag{2.5}$$

2.2 개선된 QCELP 복호기

개선된 QCELP 복호기에서는 통신 선로상에서 패킷을 받아 역 양자화기를 통해 필요한 데이터를 구한 후 각 모듈에 전달한다.



(그림 3) 개선된 QCELP 복호기

각각의 모듈들은 이전의 입력신호와 자신이 만든 여기 신호를 더해 다음 단계의 여기 신호를 만든다. 여기 신호를 생성하는 모듈에서는 역 양자화기로부터 코드북 인덱스와 이득 값을 받고 주기적 여기 신호를 만든 후에 선형필터와 콘볼루션을 계산해서 합성 신호를 얻어낸다[7]. 그리고 피치를 생성하는 모듈에서는 피치 정보와 음색 정보를 가지고 있는 여기 신호를 생성한다. 그러므로 여기 신호와 성도여과기 계수(선형 예측 필터)를 콘볼루션 계산을 통해서 목적하는 원음을 생성한다.

2.3 개선된 코드북 검색 방법

코드북 검색에서 음성 합성은 식 (2.6)을 이용하여 합성한다.

$$\hat{s}_k(n) = \sum_{i=0}^n x_k(n-i)h_w(i) \quad (2.6)$$

코드북은 인접한 여기 신호들이 서로 두 샘플씩 쉬프트되어 중첩되어 있는 희박 중첩형 코드북(sparse overlapped codebook)을 사용한다[4]. 코드북의 구조 자체가 한 샘플씩 쉬프트되어 있으므로 연속되는 여기 벡터들을 순차적으로 합성시키고 오차 에너지를 구할 때, 이전 여기 벡터의 계산 결과를 이용하여 합성 필터링 및 오차의 계산을 간단하게 구할 수 있다. 현재의 여기 벡터의 계산은 식 (2.7)에 의해 이전 여기 벡터의 계산 결과를 이용하여 구한다. 이 과정을 식으로 나타내면 식 (2.7)과 같다.

$$\hat{s}_{k+1}(n) = x_{k+1}(0)h_w(n) + \hat{s}_k(n-1) \quad (2.7)$$

또한 두 가지 경우로 나타내면 식 (2.8)과 같이 구할 수

있다. 여기서 n은 코드북 벡터 크기를 나타내고, R은 쉬프트 크기를 나타낸다. 즉 중첩 코드북에서 옆으로 이동한 길이를 R을 이용해서 표현한다.

$$\hat{s}_{k+1}(n) = \begin{cases} \sum_{i=0}^n x_{k+1}(n-i)h_w(i) & 0 \leq n < R \\ \sum_{i=0}^{R-1} x_{k+1}(i)h_w(n-i) & R-1 \leq n < 2R \\ \hat{s}_k(n-R) & n \geq 2R \end{cases} \quad (2.8)$$

위의 식을 이용하여 타겟 신호와 비교할 합성 신호를 구한다. 처음 코드북 벡터를 비교할 때에는 콘볼루션을 이용하여 임펄스 응답을 구하지만, 두 번째 코드북 벡터를 비교할 때에는 R의 크기 만큼씩 선형 이동하여 계산하기 때문에 콘볼루션 과정을 거치지 않고 자기 회기 방법으로 임펄스 응답을 구할 수 있다. 또한 코드북 벡터가 0인 경우 계산을 하지 않고 단지 이동만으로 콘볼루션 계산을 수행할 수 있다. 그래서 클리핑 기법으로 코드북 벡터를 0으로 많이 만들어 계산 시간을 단축시킨다.

여리가 가장 작은 합성신호를 생성하는 인덱스와 이득을 검색하고, 각 파라미터들을 양자화한다.

첫 번째 코드북의 인덱스와 이득을 구한 후, 피치 검색하고 남은 여기 신호를 지금 구한 파라미터를 이용해 역 필터링 한다. 잔차 신호를 양자화하고 남은 신호만 걸러져 더 작은 에리 신호만 남게 된다. 이 에리를 수식으로 나타내면 식 (2.9)와 같다.

$$e(n) = s(n) - \sum_{j=-1}^1 \beta_j r(n-r-j) - \sum_{j=-1}^1 \alpha_j s(n-j) \quad (2.9)$$

여기서 e(n)은 다음 단계에서 쓰일 잔차 신호, 즉 타겟 신호를 말하며 e(n)을 타겟 신호로 하고, 다시 코드북 검색을 수행하여 최적의 코드북 벡터를 찾아 양자화를 수행한다. 첫 번째와 두 번째 코드북은 같은 것을 사용하기 때문에 추가적인 메모리 공간을 사용하지 않는다.

두 번째 코드북 검색에는 타겟 신호를 만들 때 곱셈 계산을 수행하는 콘볼루션 계산을 하는 합성을 하지 않고 타겟 신호를 만들 수 있기 때문에 계산 시간을 줄일 수 있다. 구현 프로그램에서는 첫 번째 코드북 인덱스를 찾을 때 콘볼루션 계산을 하기 때문에 코드북 인덱스를 찾은 다음에는 콘볼루션 계산을 하지 않고 잔차 신호에서 타겟 신호를 추출 할 수 있어 계산시간을 더 단축할 수 있다.

두 번째 코드북의 검색의 모델링을 수식으로 나타내면 식 (2.10)과 같다.

$$e(n) = s(n) - Gx_k(n) \otimes \alpha_s(n-j) - \sum_{j=-1}^1 \beta_j r(n-r-j) - \sum_{j=-1}^1 \alpha_j s(n-j) \quad (2.10)$$

여기서 G는 코드북 이득이고 k는 최적 코드북 인덱스이다. 그리고 ⊗연산은 콘볼루션을 나타내고, e(n)은 잔차 신호이다.

두번째 코드북 검색은 이전 신호의 여기 신호를 타겟 신호로 간주한다. 여기 신호는 음색을 나타내는 신호들의 집합이기 때문에 정밀한 양자화를 필요로 한다. 실험에서 두번째 코드북 인덱스와 이득을 추적해보면 처음 코드북 검색에서 찾은 이득보다 작은 값을 갖는다.

```

compute_cb ()
{
    get_impulse_response_pole(); //임펄스 응답 필터를 구한다.
    construct c(n); //코드북 벡터를 구성한다.
    complete c(n) in memory; //코드북을 메모리에 올린다.
    For(I = 0; I < n; I++) {
        err = -2 * G * Exy + G * G * Eyy;
        find_i_G_min_err(); //err이 최소인 i와 G를 찾는다.
        If err < min_err {
            Select_i_G(); //현재의 인덱스와 이득을 선택한다.
        }
    }
    Quantize_i_G (); //i와 G를 양자화한다.
    Make_residual(); // 잔차 신호를 만든다.
    Find_second(); //두번째 최적의 코드북 인덱스와 이득을 찾는다.
    Quantize_i_G (); // i와 G를 양자화한다.
}
    
```

(알고리즘 1) 코드북 검색 알고리즘

본 논문에서 이용하는 QCELP에서는 전송패킷의 코드북 인덱스에 7비트를 할당하고 이득에는 3비트를 할당한다[3]. 그러므로 두 번째 코드북 검색은 처음 코드북 검색과 같은 3비트를 할당받지만 더욱 정밀하게 양자화 될 수 있다. 이유는 양자화 해야 할 값의 범위가 처음보다 작아졌기 때문에 정밀한 양자화를 할 수 있기 때문이다. 두 번째 코드북 검색에도 코드북 인덱스에 7비트를 할당하고 이득에 3비트를 할당한다. 이 때 전송률을 계산하면 기존 QCELP는 최대 전송률이 8kbps이지만 개선된 QCELP는 12kbps가 된다. 그러나 평균 전송률을 계산해 보면 각각 5.5kbps와 8.0kbps가 된다.

3. 구현 및 실험

개선된 QCELP 부호기의 모의 실험은 펜티엄 II 450 MHz PC에서 C 언어로 구현된 인코더와 디코더에 의해 실시되었고, AD/DA 변환기로는 PC용 16비트 사운드 카드를 사용하였다. 실험에 필요한 음성은 실제 생활에서 적용될 수 있는 상황을 구현하기 위해 주변의 잡음을 거르지 않고 그대로 자료로 사용하였고, 다양한 장소(거리, 강의실, 연구실, 가정집)에서 마이크를 통해 16 비트, 16 kHz로 샘플링한 후 PCM방식으로 음성 파일을 저장하였다. 저장된 파일을 원래의 QCELP 음성 부호기와 개선된 QCELP 음성 부호기에 차례로 입력하여 두 부호기의 성능을 비교하였다.

본 논문에서 제시하는 알고리즘은 현재 상용화된 QC-ELP의 표준인 IS-95B를 기초로 하여 마이크로 소프트웨어의 Visual-C 5.0으로 구현하였다. 각 모듈은 각각 독립적으로 구현되어 있어 일부분 모듈을 향상시켜도 다른 모듈에 영향을 주지 않는다. 그리고 각 모듈은 실시간으로 동작한다는 장점이 있다. 그러므로 단말기에 C 또는 어셈블리어로 변환하기가 쉽다.

<표 1> 실험에 사용된 문장

문장번호	내 용
1	인하대학교 HCI 연구실입니다.
2	아이들이 활짝 웃는다.
3	너는 기분이 좋구나!
4	너는 오늘 시내에 가니?
5	내 말을 믿어라.
6	이 책은 재미있겠다.
7	그 애가 책을 읽느냐, 잡지를 보느냐?
8	이 편지를 읽어 주십시오.
9	내일 아침에 다시 나와 주십시오.
10	우리를 시험에 들게 하지 마옵시고, 다만 악에서 구원하십시오.
11	선생님께서는 따님이나 아드님을 뵈거나 두셨습니까?
12	하나, 둘, 셋, 넷, 다섯, 여섯, 일곱, 여덟, 아홉, 열.
13	일, 이, 삼, 사, 오, 육, 칠, 팔, 구, 십.
14	In the beginning was the Word.
15	And now abide faith, hope, love, these three ; but the greatest of these is love.

개선된 QCELP 알고리즘의 성능을 평가하기 위해 사용된 문장은 다음의 15 문장으로, 각각의 문장은 10대부터 50대까지 남녀 20명의 화자로 하여금 각각 3회씩 발성하도록 하여 음성 데이터 파일로 저장한 후 사용하였다.

기존의 QCELP 부호기에 대한 개선된 QCELP 부호기의 성능 향상을 알아보기 위해 본 논문에서는 SNR (Signal-to-Noise Ratio)과 segSNR(segmental SNR)을 각각 구하여 비교하였다[2]. 평가하기 위한 SNR과 segSNR은 각각 식 (3.1)과 식 (3.2)에 의해 구해진다.

$$\begin{aligned}
 SNR &= 10 \cdot \log_{10} \frac{\sum_n s^2(n)}{\sum_n e^2(n)} \quad (dB) \\
 &= 10 \cdot \log_{10} \left\{ \frac{\sum_n s^2(n)}{\sum_n [s(n) - \hat{s}(n)]^2} \right\} \quad (dB)
 \end{aligned}
 \tag{3.1}$$

$$segSNR = \frac{1}{M} \sum_{m=1}^M [SNR(m)] \quad (dB)
 \tag{3.2}$$

여기서 n은 프레임 크기이고, s(n)은 입력 음성 신호, $\hat{s}(n)$ 은 재생 합성된 음성 신호이다. 그리고 M은 세그먼트의 수를 나타낸다.

4. 결과 및 성능 평가

(그림 4), (그림 5), (그림 6)을 볼 때 입력데이터에 잡음

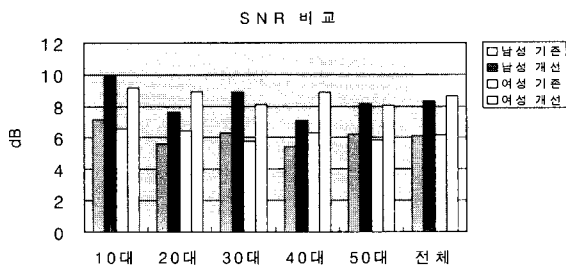
이 많지만 상대적으로 본 논문의 방법을 사용한 QCELP가 잡음에 강인하다는 것을 알 수 있다. 또한 남성 화자보다 여성 화자가 음질이 더 향상되었음을 알 수 있다.

이러한 이유는 잔차 에너지를 더욱 자세히 나타낼 수 있기 때문에 고주파 성분을 세밀하게 양자화 하므로 상대적으로 여성 음성을 더 자세히 양자화하기 때문이다. 결과적으로 남성 음성보다 여성 음성의 음질이 상대적으로 0.5dB 정도 더 나은 결과가 있었다.

SNR은 전체음성을 비교 대상으로 계산하기 때문에 문장에 대해서 비교할 때 유용하다.

(그림 4)를 보면 10대 음성이 상대적으로 음질이 더 좋은 결과를 볼 수 있다. 원인은 본 논문이 제안한 부호기는 비정상성 신호에 더 적응을 잘 하기 때문이며 10대의 음성은 불안정한 음성이기 때문에 비정상성을 나타내는 자료로 이용되어질 수 있다는 사실을 알게 되었다.

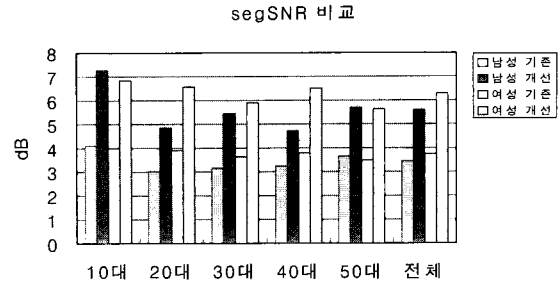
또한 남성들 중에서도 고음을 낼 수 있는 사람들은 상대적으로 기존 QCELP 보다 더 자세히 양자화 할 수 있어 음질이 더 향상되었다. 이것은 기존 QCELP가 무시한 잔차 에너지를 수정된 QCELP는 더 자세히 양자화해서 나타내기 때문에 상대적으로 음질의 향상이 두드러지게 나타난 것이다.



(그림 4) SNR 비교

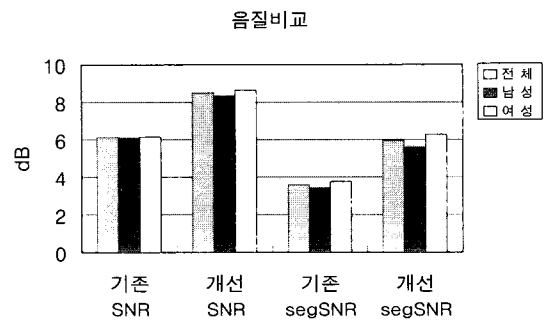
segSNR은 짧은 구간에서 음질을 측정할 수 있는 척도가 될 수 있다. 각 구간에서의 평균을 음질의 결과로 나타내 주기 때문에 합성된 파형이 얼마나 정확하게 표현되어 있는가를 나타내 준다.

(그림 5)는 상대적으로 기존 QCELP 보다 개선된 QCELP가 각 프레임마다 원음에 더 가깝다는 것을 나타낸다. 10대를 제외하고는 여성음성이 상대적으로 더 좋은 음질을 나타낸다. 이는 개선된 QCELP 부호기가 비정상성 음성을 충분히 분석을 하고 있다는 점을 나타낸다. 실제 합성된 음성을 들어보면 기존 QCELP 합성 음성보다 잡음이 적고 음성이 명확하게 들린다. 부가적인 효과로 코드북 검색을 많이 할수록 주위의 잡음이 사라지는 것을 실험 결과인 합성 음성을 들어서 확인 할 수 있었다. 그리고 원음에 잡음이 더 많을수록 음질의 향상률이 더 높아지는 것을 볼 수 있었고, 잡음도 음질에 영향을 준다는 사실을 알 수 있었다.



(그림 5) segSNR 비교

전체적으로 어느 정도 음질의 향상이 있었는가는 (그림 6)을 통해서 알 수 있다. 원음과의 비교 음질(SNR, segSNR)은 기존 QCELP에서는 각각 6.179282dB, 3.62999dB이고, 수정된 QCELP에서는 각각 8.557273dB, 6.019737dB이다. 결과적으로 수정된 QCELP는 각각 2.377991dB, 2.396738dB 정도 음질이 향상되었다.



(그림 6) 음질 비교

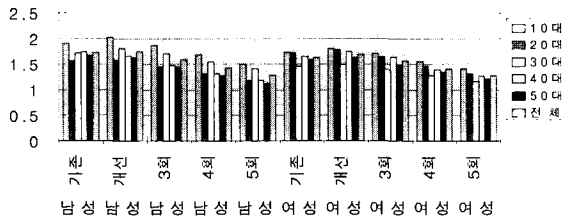
(그림 4), (그림 5)를 보면 전체적으로 음질이 낮은 것은 실험 대상 자료에 노이즈가 많은 상태의 음성이기 때문이다. 이것은 실제 휴대폰 사용장소가 일상 생활 공간에서 쓰이기 때문에 이를 고려한 실험을 하기 위해서 다양한 장소에서 녹음한 음성 자료를 사용하게 되었다. 본 논문의 실험을 통해서 실제 제한 조건을 수용할 수 있다는 장점이 있다. 잡음이 배제된 곳에서 녹음한 자료를 실험에 사용하게 되면 실제 생활과 거리가 있기 때문에 실제 활용 면에서 예상치 못한 상황이 되어 알고리즘을 적용하기 어려운 경우가 생길 수 있다. 그래서 잡음이 많은 경우 기존의 QCELP는 음질이 상당히 많이 떨어지고 있지만, 개선된 QCELP는 상대적으로 더 나은 음질을 나타내고 있다. 이 실험 결과를 통해서 개선된 QCELP는 잡음에 강인하다는 것을 실험을 통해서 증명할 수 있다.

또한 개선된 QCELP는 기존 QCELP보다 계산시간은 늘었지만, 현재는 MSM 칩의 성능이 향상되었고 사용자의 음질향상 요구로 서비스 업체와 제작자의 입장에서는 계산시간보다 음질향상을 중요시하기 때문에 계산량 증가는 무시할만한 수준이다.

(그림 7)은 코드북 검색을 1회에서 5회까지 여러 번 양자화를 해서 어느 정도 음질이 향상되는가를 나타낸 그래프이다. 코드북 검색 횟수가 많아질수록 음질도 또한 함께 향상됨을 보였다. 그러나 많은 코드북 검색은 작은 인덱스와 이득을 전송해야 하므로 많은 비트를 전송해야 하므로 전송률에 따른 음질의 향상 비율을 비교하는 실험을 하였다. (그림 7)에서 기준은 코드북 검색을 1회 하는 것이고, 개선은 코드북 검색을 2회하는 것이다.

이때 3회 이상의 코드북 검색을 하면 음질은 상당히 좋아지지만 그에 따른 전송률이 높아지기 때문에 넓은 대역폭을 가진 IMT-2000 시스템이라도 경제적으로 수용하기 어렵게 된다. 그러므로 위의 그림을 통해 2회의 코드북 검색이 음질과 전송률을 만족하는 최적의 방법이라는 것을 알 수 있다. 따라서, 본 논문에서는 2회 코드북 검색 방법을 IMT-2000 시스템의 음성 부호기로 제안한다.

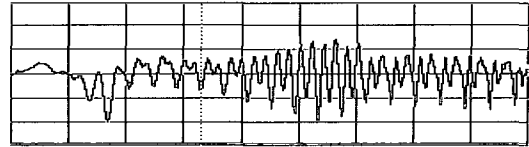
음 질 / 전송 률



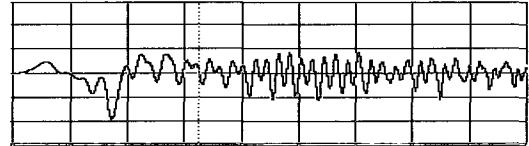
(그림 7) 전송률에 따른 음질 비교

(그림 8), (그림 9), (그림 10)은 문장 1번인 '인하대학교 HCI 연구실입니다.'의 일부 90ms 구간을 기존 QCELP부호기로 재생했을 때와 개선된 QCELP부호기로 재생했을 때의 재생 파형의 비교 결과를 나타낸다. 개선된 QCELP부호기가 더 원음에 가까운 파형을 보여주고 있다. 여기서 각 그리드(grid)는 10ms이다.

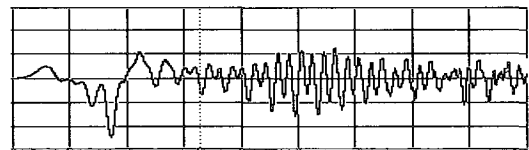
다음 <표 2>, <표 3>은 기존 QCELP 부호기와 개선된



(그림 8) 음성 파형



(그림 9) 기존 QCELP 부호기의 재생 파형



(그림 10) 개선 QCELP 부호기의 재생파형

QCELP 부호기의 결과 값을 비교해서 보여준다. 각각의 값들은 화자들에 대한 SNR, segSNR의 평균값을 나타낸다. <표 2>는 화자가 남성인 경우 각 문장에 대한 기존 QCELP 부호기와 개선된 QCELP 부호기의 음질을 비교하여 보여주며 SNR과 segSNR 각각 향상된 값을 볼 수 있다. <표 3>은 화자가 여성인 경우를 보여준다. 결과적으로 음질 향상률이 남성보다 높은 것으로 나타났다. 이러한 이유는 코드북 검색을 두 번 검색한 영향이기 때문이다.

<표 4>는 코드북 검색을 1회에서 5회까지 수행하였을 때 10대부터 50대까지 각각 SNR 값을 보여준다.

<표 5>는 코드북 검색을 1회에서 5회까지 수행하였을 때 10대부터 50대까지 각각 segSNR 값을 보여주며 segSNR이 SNR보다 낮은 값을 나타낸다. <표 6>은 기존 QCELP와 개선된 QCELP의 SNR과 segSNR 전체 평균을 보여주며 SNR과 segSNR은 각각 2.3dB 정도 향상함을 볼 수 있다.

<표 2> 남성 결과 (sSNR = segSNR)

	10대 남성				20대 남성				30대 남성				40대 남성				50대 남성			
	기 존		수 정		기 존		수 정		기 존		수 정		기 존		수 정		기 존		수 정	
	SNR	sSNR	SNR	sSNR	SNR	sSNR	SNR	sSNR	SNR	sSNR	SNR	sSNR	SNR	sSNR	SNR	sSNR	SNR	sSNR	SNR	sSNR
문장 1	7.187	4.586	9.963	8.17	5.831	3.367	8.135	5.549	6.637	3.711	9.337	6.693	5.875	3.535	7.911	5.663	6.55	3.506	8.449	5.583
문장 2	6.861	4.51	9.833	8.127	5.573	2.969	7.788	4.838	6.352	3.253	9.132	5.761	5.55	3.71	7.452	5.635	6.502	4.14	8.692	6.685
문장 3	6.591	3.755	9.355	6.41	5.314	3.061	7.41	4.785	5.954	3.064	8.322	4.973	5.569	3.323	7.748	4.931	6.658	3.794	8.432	5.838
문장 4	6.99	4.249	9.648	7.498	5.267	3.062	7.226	4.87	5.931	3.306	8.249	5.57	5.589	3.788	7.198	5.62	6.344	3.854	8.384	6.267
문장 5	7.154	3.905	10.29	7.511	5.397	2.781	7.482	4.439	6.569	3.413	9.361	6.143	4.468	2.567	5.92	3.804	5.578	3.285	7.649	5.105
문장 6	7.166	3.815	10.23	7.081	5.349	2.783	7.508	4.353	6.069	2.662	8.821	4.689	5.166	2.917	6.328	4.076	5.599	3.545	7.784	5.487
문장 7	7.283	4.256	10.1	7.769	5.457	3.185	7.631	5.312	5.896	3.134	8.626	5.601	6.6	4.026	8.178	5.989	5.589	3.41	7.535	5.379
문장 8	7.556	4.443	10.32	8.041	5.809	3.149	7.835	5.028	6.08	2.996	8.594	5.156	6.918	4.19	8.505	5.992	6.364	3.456	8.07	5.174
문장 9	7.044	4.195	9.786	7.545	5.529	3.148	7.675	5.191	6.085	3.095	8.806	5.502	4.651	2.907	6.384	4.407	5.673	3.207	7.879	5.228
문장 10	7.008	4.429	9.536	7.81	5.327	3.134	7.154	4.87	6.045	3.338	8.469	5.765	5.198	3.41	6.748	5.04	5.896	3.696	7.904	5.772
문장 11	6.942	4.371	9.959	7.922	5.497	3.324	7.647	5.386	6.052	3.242	8.853	5.903	5.445	3.57	7.296	5.187	6.078	4.054	8.327	6.427
문장 12	6.975	3.925	9.913	7.249	5.706	2.762	7.852	4.357	6.507	2.912	9.194	4.92	4.508	2.238	6.064	3.189	6.541	4.124	8.791	6.567
문장 13	7.261	3.915	9.754	6.521	5.645	2.845	7.701	4.413	6.551	3.115	9.02	5.085	5.08	2.35	6.424	3.209	5.997	3.68	7.969	5.675
문장 14	7.276	3.788	9.825	6.274	6.008	3.156	8.044	4.828	6.694	3.162	9.078	5.276	5.573	3.514	6.757	4.639	6.912	3.657	8.49	5.31
문장 15	7.133	3.193	9.798	5.296	5.944	2.749	8.103	4.298	6.65	2.824	9.321	4.721	4.877	2.352	6.453	3.345	6.347	3.23	8.341	4.832
평 균	7.095	4.089	9.887	7.282	5.577	3.032	7.68	4.835	6.271	3.148	8.879	5.45	5.404	3.226	7.024	4.714	6.175	3.643	8.18	5.689

<표 3> 여성 결과 (sSNR = segSNR)

	10대 여성				20대 여성				30대 여성				40대 여성				50대 여성			
	기 존		수 정		기 존		수 정		기 존		수 정		기 존		수 정		기 존		수 정	
	SNR	sSNR	SNR	sSNR	SNR	sSNR	SNR	sSNR	SNR	sSNR	SNR	sSNR	SNR	sSNR	SNR	sSNR	SNR	sSNR	SNR	sSNR
문장 1	6.648	4.197	9.293	7.048	6.816	4.557	9.428	7.985	5.17	3.233	7.282	5.19	6.153	3.955	8.783	6.87	5.612	3.681	7.945	6.014
문장 2	6.397	3.893	9.051	6.843	6.285	4.383	9.14	7.672	5.566	3.651	7.931	5.901	6.287	3.623	8.929	6.535	5.335	3.533	7.51	5.4
문장 3	5.979	3.751	8.424	6.175	6.3	3.902	8.655	6.3	5.211	3.665	7.431	5.317	6.136	3.835	9.002	6.763	6.454	3.892	8.731	6.02
문장 4	6.652	4.314	9.146	7.328	6.346	4.436	8.87	7.46	5.439	3.754	7.749	5.771	6.282	4.164	9.027	7.2	5.867	3.85	8.527	6.601
문장 5	6.187	3.63	8.827	6.415	6.219	3.734	8.859	6.603	5.718	3.908	8.396	6.502	5.732	3.448	8.456	6.088	5.071	2.902	7.577	5.008
문장 6	7.042	3.555	9.798	6.021	6.751	3.711	9.3	6.259	5.062	2.966	6.982	4.609	6.544	3.568	9.103	5.85	5.572	2.995	7.486	4.655
문장 7	6.791	4.258	9.362	7.146	6.73	4.273	9.485	7.57	5.205	3.412	7.7	5.434	6.386	3.771	9.164	6.866	5.695	3.47	7.759	5.434
문장 8	6.718	4.149	9.328	7.353	6.727	3.98	8.865	6.466	6.308	3.688	8.627	6.073	6.497	3.871	8.947	6.565	6.142	3.563	8.255	5.761
문장 9	6.399	3.788	9.085	6.728	6.365	3.477	9.012	5.926	5.628	3.159	8.038	5.269	6.383	3.734	9.081	6.381	5.464	3.398	7.822	5.543
문장10	6.568	4.046	9.24	7.139	6.248	3.95	8.62	6.715	6.244	3.879	8.76	6.611	6.014	3.716	8.554	6.464	5.817	3.273	8.295	5.629
문장11	6.246	4.403	8.903	7.428	6.182	4.013	8.927	6.943	5.486	3.982	7.957	6.463	6.106	3.984	8.792	6.712	5.883	3.779	8.401	6.356
문장12	6.359	3.938	9.096	7.013	6.169	3.252	8.832	5.758	5.98	3.732	8.672	6.599	5.927	3.403	8.677	6.064	5.508	3.232	8.013	5.55
문장13	6.809	4.009	9.236	6.671	6.564	3.53	8.747	5.573	6.08	3.528	8.328	5.971	7.009	4.2	9.421	7.125	6.18	3.669	8.251	5.985
문장14	6.875	4.053	9.445	6.593	6.056	3.574	8.165	5.517	6.718	3.857	9.284	6.378	6.229	3.913	8.551	6.194	6.433	3.874	8.32	5.972
문장15	6.678	3.778	9.371	6.449	6.682	3.643	8.942	5.78	6.581	3.869	9.076	6.217	6.369	3.457	8.949	5.884	5.945	2.947	8.19	4.64
평균	6.556	3.984	9.174	6.823	6.429	3.894	8.923	6.568	5.76	3.619	8.146	5.887	6.27	3.776	8.896	6.504	5.799	3.471	8.072	5.638

<표 4> SNR 비교

SNR	남 성					여 성				
	기 존	개 선	3회	4회	5회	기 존	개 선	3회	4회	5회
10대	7.095188	9.887117	11.2306	11.90252	12.29494	6.556491	9.173685	10.60966	11.41309	11.92407
20대	5.576879	7.679522	8.777955	9.464085	9.92345	6.429274	8.923099	10.19378	10.84611	11.27442
30대	6.27146	8.878803	10.24867	11.03196	11.54988	5.759769	8.14617	9.486806	10.23218	10.66701
40대	5.404423	7.024443	7.960901	8.518458	8.972689	6.27024	8.895578	10.27834	9.464085	9.92345
50대	6.175155	8.179681	9.12851	9.607412	9.995279	5.798512	8.072169	9.233539	9.961772	10.38019
전체	6.104621	8.329913	9.469328	10.10489	10.54725	6.162857	8.64214	9.960426	10.38345	10.83383

<표 5> segSNR 비교

segSNR	남 성					여 성				
	기 존	개 선	3회	4회	5회	기 존	개 선	3회	4회	5회
10대	4.089078	7.281573	9.639265	11.28281	12.42491	3.984119	6.823305	9.157148	10.94534	12.26792
20대	3.031572	4.834507	6.187649	7.202257	7.974372	3.894497	6.56841	8.550598	9.90606	10.89558
30대	3.14837	5.450452	7.41488	8.874345	10.05259	3.618934	5.887006	7.840962	9.418151	10.71253
40대	3.226351	4.713745	5.695777	6.399661	6.927985	3.77611	6.504045	8.712549	10.43294	11.69907
50대	3.642548	5.688631	7.028619	7.818777	8.459772	3.470598	5.638005	7.220929	8.449824	9.277015
전체	3.427584	5.593782	7.193238	8.315569	9.167925	3.748852	6.284154	8.296437	9.830463	10.97042

<표 7>은 코드북 검색을 1회에서 5회까지 검색하였을 때 10대부터 50대까지 전송률 평균을 보여준다. 검색 횟수가 많아질수록 전송률이 올라가는 것을 알 수 있다.

<표 8>은 각각의 전송률에 따른 음질 향상률을 보여준다. 이는 코드북 검색을 1회부터 5회까지 검색하고 10대부터 50대까지 전송률에 따른 음질 향상률을 나타낸다. 결과적으로 2회의 코드북 검색이 상대적으로 높은 값을 나타낸다.

<표 6> 전체적인 음질 비교

		전 체	남 성	여 성	
SNR	기존	6.133739	6.104621	6.162857	
SNR	개선	8.486027	8.329913	8.64214	
segSNR	기존	3.588218	3.427584	3.748852	
segSNR	개선	5.938968	5.593782	6.284154	
					음질향상%
SNR	음질향상	2.352288	2.225292	2.479283	38.34998
segSNR	음질향상	2.35075	2.166198	2.535303	65.51303

<표 7> 전송률 비교

전송률	남 성					여 성				
	기 존	개 선	3회	4회	5회	기 존	개 선	3회	4회	5회
10대	5.876736	8.526417	11.1761	13.82578	16.47546	6.109479	8.871244	11.63301	14.39478	17.15654
20대	5.493834	7.918079	10.34232	12.76657	15.19082	6.022412	8.742111	11.46181	14.18151	16.90121
30대	5.521606	7.969072	10.41654	12.864	15.31147	6.41806	9.415815	12.41357	15.41133	18.40908
40대	4.983919	7.121277	9.258635	11.39599	13.53335	6.109882	8.868084	11.62629	14.38449	17.14269
50대	5.906739	8.533452	11.16016	13.78688	16.41359	5.817969	8.434982	11.05199	13.66901	16.28602
전체	5.556567	8.013659	10.47075	12.92784	15.38494	6.09556	8.866447	11.63733	14.40822	17.17911

〈표 8〉 전송률에 따른 음질 비교

SNR+segSNR										
음질/전송률	남 성 기 준	남 성 개 선	남 성 3회	남 성 4회	남 성 5회	여 성 기 준	여 성 개 선	여 성 3회	여 성 4회	여 성 5회
10대	1.903142	2.013588	1.867366	1.676963	1.500404	1.725288	1.803241	1.6992	1.553232	1.410074
20대	1.56693	1.580438	1.447025	1.305468	1.1782	1.714225	1.772066	1.635377	1.463326	1.311741
30대	1.705995	1.798108	1.695722	1.547442	1.410868	1.461299	1.490384	1.395873	1.275058	1.161358
40대	1.731724	1.648326	1.475021	1.309067	1.174925	1.644279	1.736522	1.633444	1.383227	1.261326
50대	1.662119	1.62517	1.44775	1.263969	1.124376	1.593187	1.625395	1.488824	1.346959	1.206999
전체	1.715485	1.737495	1.591344	1.424867	1.281459	1.626054	1.683458	1.568818	1.402943	1.269231

5. 결 론

현재 이동통신에서 사용중인 기존 QCELP 는 가변 전송률로 현재의 PCS인 협대역 CDMA 방식에 적합하다. 그러나 피치가 높은 여성음성을 재생할 때 음질이 저하된다 단점을 가지고 있다. 본 논문에서는 양질의 서비스를 제공하기 위한 해결책으로 [8]에서 제안한 코드북 검색 방법이 IMT-2000의 음성 부호기로 적합하다는 것을 보였다. 실험 결과 여성음성에 대한 음질 향상률이 남성보다 0.5 dB 정도 개선됨을 보였고, 코드북 검색을 1회에서 5회까지 검색한 방법을 각각 비교해 본 결과 2회 검색이 가장 최적임을 알 수 있었다. 본 논문에서 제안한 시스템은 기존 시스템보다 SNR에서 38.35%, segSNR에서 65.51% 음질 향상이 있었고, 본 논문 방식에 의해 재생된 음질이 기존 방식에 의한 음질보다 잡음에 강인하다는 것을 실험결과를 통해 알 수 있었다.

향후 연구로는 본 논문의 코드북 검색과 피치 지연 및 피치 이득을 구하는 방법을 서로 연관시켜 분석해 계산시간을 줄일 수 있는 알고리즘과 코드북 검색 횟수가 많을수록 잡음이 감소되는 결과를 통해 기존보다 음질이 개선된 부호기를 연구하는 것이다.

참 고 문 헌

[1] 양진석, 개인휴대통신(PCS)을 위한 QCELP 부호기의 개선, 인하대학교 석사학위논문, 1997.
 [2] A. V. Oppenheim, R. W. Schaffer, *Discrete Time Signal Processing*, Prentice Hall, 1989.
 [3] B. S. Atal, "Predictive Coding of Speech at Low Bit Rates," IEEE Trans. on Communications, Vol. COM-30, No.4, pp.600-614, Apr. 1982.
 [4] W. Gardner, P. Jacobs, and C. Lee, "QCELP : A Variable Rate Speech Coder for CDMA digital Cellular," *Speech and Audio Coding for Wireless and Network Application(KAP)*, pp.85-92, 1993.
 [5] Hiragi K, Iwasaki S, Seki I, "Trends in the standardization of ITU-T SG11- Report on March 1999 general meeting," *NTT Review*, Vol.11 No.5, pp.83-86, Sept., 1, 1999.
 [6] H. K. Kim, "Adaptive Encoding of Fixed codebook in CELP Coders," *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing-Vol.1*, 1998.
 [7] A. M. Kondoz, *Digital Speech Coding for Low Bit Rate*

Communications Systems, John Wiley & Sons, 1994.

[8] 박광철의 3인, "QCELP에서 음질 향상을 위한 개선된 코드북 검색", 한국정보처리학회 추계 학술발표 논문집, 제6권 제2호, PR-23~PR-28, 1999.



박 광 철

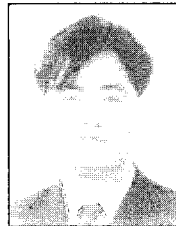
e-mail : teep@lge.co.kr

1995년 인하대학교 전자계산공학과 졸업 (공학사)

2000년 인하대학교 전자계산공학과 (공학석사)

2000년~현재 LG 전자 DM(연) Diznet팀

관심분야 : 음성코딩, 음성인식, 자연어처리, 리눅스



한 승 진

e-mail : softman@nlsun.inha.ac.kr

1989년 인하대학교 전자계산학과(이학사)

1992년 인하대학교 전자계산공학과(공학석사)

2001년 인하대학교 전자계산공학과 박사수료

1992년~1996년 대우통신 종합연구소

1996년~1998년 SK Telecom

2000년~현재 줘디오넷 부설연구소 책임연구원

1998년~현재 인하대학교 전전컴 강사

1998년~2000년 용인송담대학 인터넷 경영정보과 겸임교수

2000년~현재 인하공업대학 컴퓨터정보과 겸임교수

관심분야 : 음성신호처리, 컴퓨터 네트워크, Mobile IP, 멀티미디어 통신, IMT-2000



이 정 현

e-mail : jhlee@inha.ac.kr

1977년 인하대학교 전자공학과 졸업

1980년 인하대학교 대학원 전자공학과 (공학석사)

1988년 인하대학교 대학원 전자공학과 (공학박사)

1979년~1981년 한국전자기술연구소 시스템 연구원

1984년~1989년 경기대학교 전자계산학과 교수

1989년~현재 인하대학교 전자전기컴퓨터공학부 교수

관심분야 : 자연어처리, HCI, 정보검색, 음성인식, 음성합성, 컴퓨터구조