

EVRC의 고속 구현 알고리즘

Fast Implementation Algorithms for EVRC

정성교*, 최용수**, 김남건***, 윤대희*

(Sung-Kyo Jung*, Yong-Soo Choi**, Nam-Gun Kim***, Dae-Hee Youn*)

*연세대학교 전기·컴퓨터공학과 음향·음성 및 신호처리 연구실,

LG 정보통신 정보시스템 연구소, *SK Telecom 중앙연구원 IMT-2000 개발그룹

(접수일자: 2000년 4월 4일; 수정일자: 2000년 11월 7일; 채택일자: 2000년 12월 26일)

EVRC (Enhanced Variable Rate Codec)는 북미 및 우리 나라 CDMA 디지털 셀룰러 시스템에 채택되었으며 8kbps의 전송률에서 우수한 성능을 갖는 부호화기이다. 본 논문에서는 복잡한 알고리즘으로 인해 많은 계산량을 갖는 EVRC 부호화기를 성능 저하 없이 고속으로 구현할 수 있는 알고리즘을 제시한다. 제안된 고속 알고리즘에서는 효율적인 피치 검색과 고정 코드북 탐색 과정이 구현되는데, 고정 코드북 탐색 과정에서는 펄스 위치 조합의 수를 제한하는 방법과 줄여진 임펄스 응답을 사용하여 연산량을 기존의 방법의 70% 정도로 감소시킨다. 주관적인 음질 평가를 통해 제안된 고속 EVRC 알고리즘이 기존의 방법에 비해 적은 계산량에 구현되지만 음질의 저하는 초래하지 않는다는 것을 확인하였다.

핵심용어: EVRC, 부호화기, 고속 알고리즘

투고분야: 음성처리 분야 (2.2)

EVRC (Enhanced Variable Rate Codec) has been adopted as a standard coder for the CDMA digital cellular system in North America and Korea, and known to provide good call quality at 8kbps. In this paper, fast implementation algorithms for EVRC encoder are proposed. The proposed algorithms are based on both efficient pitch detection scheme and fast fixed codebook search algorithm. In the codebook search, computational complexity is reduced down to 70% of the original EVRC by limiting the number of pulse position combination and by using a truncated impulse response. The proposed algorithms enable us to implement the EVRC with much smaller computational works. Also, informal subjective tests confirmed that the difference in the speech quality between the original EVRC and the proposed method was indistinguishable.

Key words: EVRC, Coder, Fast algorithm

Subject classification: Speech signal processing (2.2)

1. 서 론

현재 북미와 우리 나라의 CDMA DCS (Digital Cellular System)에서 통화 품질의 개선을 위해 IS-127 EVRC[1]를 채택하여 서비스 중에 있다. 또한 EVRC는 GSM AMR (Adaptive Multi-Rate) 부호화기와 함께 차세대 이동통신인 IMT-2000의 음성 부호화 표준안으로 알려져 있다. 기존의 CELP (Code-Excited Linear Prediction) 부호화기와는 달리 EVRC는 일반화된 분석-합성 (generalized analysis-by-synthesis)[2] 구조를 사용하는 RCELP (Relaxation CELP) 알고리즘[3]에 기반한다. RCELP는 선형 보간으로 얻은

피치 컨트롤에 적합하도록 신호를 수정한 후 적응 코드북과 고정 코드북을 사용하여 여기 신호를 부호화한다. 이러한 방법은 피치에 할당되는 비트를 줄일 수 있지만 부호화 과정에서 많은 계산량이 요구된다는 문제가 있다.

본 논문에서는 EVRC의 효과적인 실시간 구현을 위해 피치 검색 과정과 가장 많은 연산량을 요구하는 고정 코드북 탐색 과정에서의 계산량을 줄인 고속 알고리즘을 제안한다. 우선 피치 검색 과정에서는 간축된 도메인에서 지연 검색 과정을 간단히 함으로써 피치 추정 과정에서 소요되는 연산량을 개선하였다. 고정 코드북 검색 과정에서는 분석-합성 방법에 의해 펄스의 위치를 정하므로 검색하는 펄스 위치의 조합에 제한을 두고 임펄스 응답의 길이를 줄여 연산량을 개선하였다.

본 논문의 구성은 다음과 같다. 2장에서는 EVRC의 부호화 과정에 대해 살펴보고 3장에서는 EVRC의 피치 검색

책임저자: 정성교 (skjung@lethe.yonsei.ac.kr)
120-749 서울시 서대문구 신촌동 134번지
연세대학교 전기·컴퓨터공학과 음향·음성 및 신호처리 연구실
(전화: 02-2123-2863; 팩스: 02-2123-4584)

과정과 고정 코드북 탐색 과정에서의 고속 알고리즘을 제안한다. 그리고 4장에서는 제안된 고속 알고리즘에 대해 성능 평가를 수행한다. 끝으로 5장에서 결론을 맺는다.

II. EVRC 음성 부호화기

EVRC는 우리 나라와 북미 CDMA 디지털 셀룰러 시스템에서 채택되어 서비스 중에 있으며 부호화 알고리즘으로 일반화된 분석-합성 방법에 기반한 RCELP 알고리즘을 사용한다. 그리고 제한된 채널을 효과적으로 사용하기 위해 세 가지의 전송률을 가변적으로 사용하는데 이는 입력 신호에 따라 부호화가 내부에 있는 RDA (Rate Determination Algorithm)에서 결정된다. EVRC에서는 현재 프레임에서 개회로 검색으로 찾은 피치와 이전 프레임에서 찾은 피치를 선형 보간을 하여 얻은 단순화된 피치 컨트롤러를 적용 코드북 여기를 위한 지연 컨트롤러로 사용한다. 프레임 단위로 정수 피치를 전송하므로 서브프레임 단위로 실수 피치를 전송하는 기존의 CELP 방법에 비해 피치에 할당되는 비트를 줄일 수 있으며, 고정 코드북의 여기 신호와 채널 손상 방지를 위한 보호 비트 등에 비트 할당을 늘려 복호화된 신호의 음질을 향상시킬 수 있는 점이 있지만 부호화 과정에서 많은 계산량을 요구한다.

부호화 과정은 다음과 같다. 고역 필터링과 잡음 억제의 전처리 과정을 통과한 입력 음성에 대해 20ms의 프레임 단위로 선형예측 분석을 수행한다. 구해진 선형예측 계수는 양자화 과정과 전송 과정에서 효과적인 LSP (Line Spectrum Pairs) 파라미터로 변환한다. 그리고 전처리된 음성 신호를 현재와 이전 프레임에서 얻은 LSP 파라미터를 보간하여 얻은 역필터에 통과시켜 잔차 신호를 얻은 후 개회로 검색으로 지연을 추정하고 개회로 장구간 예측 이득을 구한다. 그리고 앞의 과정에서 구한 17개의 자기 상관 계수와 개회로 장구간 예측 이득을 이용하여 입력 신호의 특성에 맞도록 세 가지 전송률 (8, 4, 1kbps) 중에서 하나를 선택한다. 음성은 최대 전송률과 중간 전송률로 부호화되고 배경 잡음은 최소 전송률로 부호화된다.

적용 코드북의 지연 컨트롤러는 현재 프레임과 이전 프레임의 피치를 선형 보간하여 구한 선형 피치 컨트롤러를 사용하고 고정 코드북 검색에 사용되는 목적 신호를 만들기 위해 이전 과정에서 수정된 잔차 신호와 현재 프레임의 지연 컨트롤러를 가지고 잔차 신호 수정을 위한 목적 신호를 만들고 이 신호에 최적으로 정합되도록 원래의 잔차 신호를 수정한다.

수정된 잔차 신호를 합성 필터에 통과시켜 수정된 음성 신호를 구하고 가장 필터를 사용하여 수정된 음성 신호를 가중한다. 수정된 가중 음성 신호에서 적용 코드북의 기여를 제거하여 고정 코드북 탐색을 위한 목적 신호를 생성한다. 고정 코드북은 ACELP (Algebraic CELP) 알고리즘에 기반하며 코드북 탐색은 서브프레임 단위로 행한다. 코드북은 최대 전송률인 경우에는 ± 1 의 크기를 갖는 8개의 펄스를, 중간 전송률인 경우에는 3개의 펄스를 포함한다. 다음으로 본 논문에서 연산량을 개선한

피치 지연 추정과 고정 코드북 탐색 과정에 대하여 설명한다.

2.1. 피치 지연 추정

피치 추정은 전처리된 신호에 대하여 행하지 않고 포먼트 성분이 제거된 선형예측 잔차 신호에 대하여 프레임 단위로 행하여진다. 예측 잔차 신호는 전처리된 음성 신호를 각 구간의 보간된 LPC를 필터 계수로 갖는 선형 예측 필터에 통과시켜 얻는다. 피치 지연이란 선형 예측 잔차 신호의 자기상관 함수를 최대화 하는 지연을 뜻한다. EVRC에서는 피치 지연을 추정하기 위해 두 개의 추정 윈도우에 대해 각각 독립적으로 행하는데, 첫 번째 윈도우는 현재 프레임의 전구간을, 두 번째 윈도우는 현재 프레임의 반과 다음 프레임의 반을 포함한다.

효율적인 지연 검색을 위해 우선 4:1로 간축된 잔차 신호에 대해 피치를 구한다. 이 과정에서 과거 피치와의 연속성을 가지도록 역방향 추정 (backward tracking)을 수행한다. 간축된 도메인에서 결정된 지연을 이용하여 최적 지연 추정을 위한 검색 구간을 정하고 결정된 구간 내에서 잔차 신호의 자기상관 함수가 최대값을 갖는 최적 지연과 장구간 예측 이득을 구한다. 간축된 도메인에서의 지연을 추정할 때와 같은 방법으로 과거 지연과의 연속성을 위해 역방향 추정을 수행한다. 두 개의 추정 윈도우에 대해 각각 독립적으로 지연과 장구간 예측 이득을 구한 후 순방향 추정으로 두 지연과 이득을 결합하여 최종적인 지연과 이득을 얻는다.

2.2. 고정 코드북 탐색

우선 고정 코드북 탐색을 위한 목적 신호 $x_m(n)$ 을 다음과 같이 구한다.

$$x_m(n) = \hat{s}_m(n) - g_p \lambda(n), \quad 0 \leq n < L \quad (1)$$

여기서 수정된 음성 신호 $\hat{s}_m(n)$ 과 필터링된 적용 코드북의 여기 신호 $\lambda(n)$ 은 각각 수정된 잔차 신호와 적용 코드북의 여기 신호를 가중 합성 필터에 통과시켜 구한다. 그리고 g_p 는 양자화된 폐회로 이득이다.

고정 코드북은 저장과 계산량에 장점을 가지고 있는 대수 코드북 (algebraic codebook)이며, 그 구조는 ISPP (Interleaved Single-Pulse Permutation) 디자인에 기반한다. 코드북 탐색은 원래 신호와 합성 신호 사이의 가중된 평균 자승 오차를 최소화하는 코드북 인자와 코드북 이득을 찾는 과정으로 서브프레임 단위로 행한다.

2.3. 최대 전송률인 경우의 대수 코드북의 구조

최대 전송률인 경우의 고정 코드북은 35비트 대수 코드북을 사용한다. 이 코드북에서는 모든 코드북 벡터는 크기가 ± 1 인 펄스를 최대 8개를 포함하고 있으며 그 길이는 55이다. 한 서브프레임은 각각 11개의 펄스 위치를 갖는 5개의 트랙으로 나뉘어지며 5개의 트랙에 대해 8

개의 펄스를 탐색하므로 2개의 펄스를 갖는 이중-펄스 트랙 (double-pulse track)과 하나의 펄스를 포함하는 단일-펄스 트랙 (single-pulse track)이 존재한다. 단일-펄스 트랙의 쌍은 4가지 중에서 하나가 선택되며 2비트의 코드 워드를 사용하여 부호화하고 이를 수신단에 전송한다. 그리고 단일-펄스 트랙에 대해서는 9비트를 사용하고 이중-펄스 트랙에 대해서는 각각 8비트를 사용하여 두 펄스의 위치와 부호를 부호화한다. 따라서, 대수 코드북의 부호화에는 총 35비트가 필요하다.

가중된 입력 신호와 가중된 합성 신호 사이의 평균 자승 오차를 최소화시키도록 대수 코드북을 검색하는 코드북 탐색 과정은 아래 식 (2)를 최대화시키는 코드 벡터 c_k 를 찾는 과정이다.

$$T_k = \frac{C_k}{E_k} = \frac{(d' c_k)^2}{c_k' \Phi c_k} \quad (2)$$

여기서 $d = H' x_w$ 는 식 (1)에서 주어진 목적 신호 x_w 를 가중 합성 필터 H 에 통과시켜 얻은 역방향 필터링된 신호이며 $\Phi = H'H$ 는 가중 합성 필터의 임펄스 응답의 상관행렬이다. 벡터 d 와 행렬 Φ 는 코드북을 탐색하기 전에 미리 계산하며 그 연산량은 서브프레임 길이의 제곱에 비례한다.

EVRC에서는 최적의 코드북 벡터를 결정하는 코드북 탐색을 간단하게 하기 위해 두 가지 방법을 사용한다. 우선 적절한 참조 신호로 사용하여 그 신호의 부호 정보를 이용하여 펄스의 부호를 미리 정한다. 사용하는 참조 신호는 잔차 도메인에서의 목적 신호와 역방향 필터링된 신호와의 가중된 합으로 이루어진 신호이다. 펄스 부호가 정해지면, 다음으로 최적 펄스 위치를 찾기 위해 많은 계산량을 요구하는 nested-loop 탐색 방법 대신 depth-first tree 탐색 방법을 반복적으로 사용하여 효과적으로 펄스 위치를 찾는다. 최대 전송률인 경우에 우선 8개의 펄스를 4쌍의 펄스로 분류하고 펄스의 위치는 한 번에 1쌍씩 순차적으로 결정한다. 이러한 과정을 4가지 경우의 단일-펄스 트랙과 이중-펄스 트랙의 조합에 대해 수행한다.

III. 제안된 고속 알고리즘

본 논문에서는 최대 전송률인 경우의 EVRC 부호화 과정을 구성하는 개회로 피치 지연 검색 과정과 고정 코드북 탐색 과정에서 적용될 수 있는 고속 알고리즘을 제안한다. 피치 검색 과정에서는 유성음 구간에서의 검색 과정을 간단히 함으로써 계산량을 감소시켰다. 고정 코드북 탐색 과정에서는 검색하는 펄스 위치를 제한하고 줄여진 임펄스 응답을 사용하였다. 제안된 알고리즘에 대해 객관적인 음질 평가 방법을 사용하여 이에 대한 성능 평가를 수행한다. 객관적인 음질 평가 방법에는 SNR과 segSNR 그리고 심리음향을 고려한 BSD를 사용하였다. EVRC에

서 일반화된 분석-합성 방법을 사용하므로 원 신호 대신에 수정된 음성 신호를 객관적 평가를 위한 참조 신호로 사용하였다.

3.1. 효율적인 피치 지연 검색 과정에서의 계산량 감소

본 논문에서는 간속된 도메인에서의 지연 검색 과정에서 소요되는 계산량을 개선시키기 위한 방법을 제안하고자 한다. EVRC에서는 잔차 신호의 자기상관 함수를 이용하는 개회로 지연 검색을 수행한다. 고속 검색을 위해 우선 간속된 잔차 신호 영역에서 지연을 구하고 이 값을 이용하여 최적 지연을 검색하기 위한 구간을 정한다. 일반적으로 피치가 존재하는 구간은 [20,120]이므로 4:1로 간속된 도메인에서 검색은 [5,30]에서 식 (3)과 같이 주어지는 자기상관함수를 최대화시키는 지연 d 를 찾는다.

$$r(d) = \sum_{k=0}^{40-d} \varepsilon_d(k) \varepsilon_d(k+d) ; \quad d \in [5,30] \quad (3)$$

여기서 $\{\varepsilon_d(n)\}$ 은 간속된 도메인의 잔차신호이다.

제안된 방법에서는 식 (4)와 같이 간속된 도메인에서의 검색 구간을 여섯 구간으로 나누고 구간별 지연을 순차적으로 검색한다. 각 구간에서 식 (4)의 자기상관함수를 최대화시키는 지연 d_i 가 결정되면 식 (5)와 같이 간속된 잔차 신호의 상관함수 $r_{\max}(d_i)$ 가 잔차 신호의 에너지 $r(0)$ 의 일정비를 넘으면 검색을 중단하고 해당하는 지연을 간속된 지연으로 결정한다.

$$r(d_i) = \sum_{k=0}^{40-d_i} \varepsilon_d(k) \varepsilon_d(k+d_i) ; \quad d_i \in R_i, \quad i = 1, 2, \dots, 6 \quad (4)$$

$$\frac{r_{\max}(d_i)}{r(0)} \geq THR_i ; \quad d_i \in R_i, \quad i = 1, 2, \dots, 6 \quad (5)$$

여기서 R_i 는 여섯 검색 구간인 [5,8], [9,12], [13,16], [17,20], [21,24], [25,30]을 각각 나타내며 THR_i 는 각 검색구간에서 지연을 결정하기 위해 실험적으로 구한 문턱값이다.

제안한 방법을 피치 검색 과정에 실제 적용한 결과, 각각 3명의 남녀 음성에 대해 모두 기존의 방법으로 구한 지연과 일치함을 확인하였다. 컴퓨터 모의 실험을 통해 각 방법에 대해 피치 검색 과정에서 소요되는 시간을 측정한 결과를 살펴보면, 기존의 방법에 비해 제안된 방법이 피치 검색 과정에서 평균적으로 약 21%의 계산량을 개선하였음을 확인하였다.

3.2. 고정 코드북 탐색 과정의 계산량 감소

고정 코드북 탐색 과정은 CELP 부호화기에서 가장 많은 계산량을 필요로 하는 과정이다. 실시간 구현을 위해서는 우선적으로 이 과정에서의 고속화 알고리즘 연구가 필요하다. 본 논문에서는 고정 코드북 탐색 과정에서의

계산량을 개선시키는 알고리즘을 제안한다.

3.2.1. 펄스 위치 제한에 의한 펄스 위치 조합의 감소

EVRC에서는 효과적인 탐색을 위해 depth-first tree 탐색 방법을 반복하여 펄스의 위치를 결정한다. 8개의 펄스 위치를 찾기 위해 1쌍씩 묶어 4쌍의 펄스로 분류하고 한 번에 한 쌍씩 순차적으로 결정하는데 이 과정을 4가지의 코드워드에 대해 반복한다. 이 경우 가능한 펄스 위치 조합은 1936이다. 실제 구현에서는 이보다 적은 위치 조합에 대해서 탐색을 한다. 펄스 부호의 사전 조절을 위해 사용하는 참조 신호의 절대값의 크기를 이용하여 홀수번째 펄스 위치에 대하여 11개의 펄스 위치에서 6개로 줄였다. 이 경우의 가능한 펄스 위치 조합은 1056이다.

본 논문에서 제안된 방법에서는 단일-펄스 트랙과 이중-펄스 트랙의 네 가지 조합을 나타내는 코드워드를 줄이거나 펄스 위치에 제한을 두어 가능한 펄스 위치 조합을 감소시켰다. 제안된 방법에 대한 가능한 펄스의 위치 조합과 그에 따른 SNR을 표 1에 정리하였다.

표 1. 펄스 위치 제한에 따른 위치 조합과 객관적 음질평가 결과
Table 1. Possible pulse-position combinations and objective measure results by pulse position restriction.

FCB Search Method 1	Possible Position Combination	SNR (dB)
Original Method	{(6x11)x4}x4	1056 17.268
M 1-1	{(6x11)x4}x3	792 17.268
M 1-2	{(5x11)x2+(5x5)x2}x4	640 17.130
M 1-3	{(3x11)x4}x4	528 17.140

'M 1-1'은 네 가지 코드워드 중에서 세 번째 코드워드를 제외하여 25%의 위치 조합을 줄인다. 약 40%의 위치 조합을 감소시킨 'M 1-2'는 세 번째와 네 번째의 펄스 쌍에 비해 첫 번째와 두 번째의 펄스 쌍은 많은 펄스 위치 조합에 대해서 탐색하였다. 'M 1-3'은 기존의 방법에 서처럼 참조 신호의 크기 정보를 이용하여 펄스 위치에 제한을 두어 50%의 위치 조합을 감소시킨다. 홀수 번째 펄스의 가능한 위치를 11개에서 3개로 줄이면 기존의 방법보다 조합의 수가 반으로 준다. 위치 조합 감소에 대한 객관적 평가 결과를 보여준다. 객관적 평가 결과, 위치 조합 조합을 반으로 줄인 경우에 대해서 성능의 저하가 거의 없음을 확인하였다.

3.2.2. 임펄스 응답의 길이 감소

고정 코드북 탐색은 분석-합성법에 기반하여 원 신호와 재생 신호 사이의 차이를 최소화하는 코드 벡터와 이득을 구하는 과정이다. 코드 벡터와 가중 합성 필터의 임펄스 응답과의 필터링으로 합성 신호를 생성한다. 그림 1의 (a)는 기존의 EVRC 부호화기를 보여주고 있다.

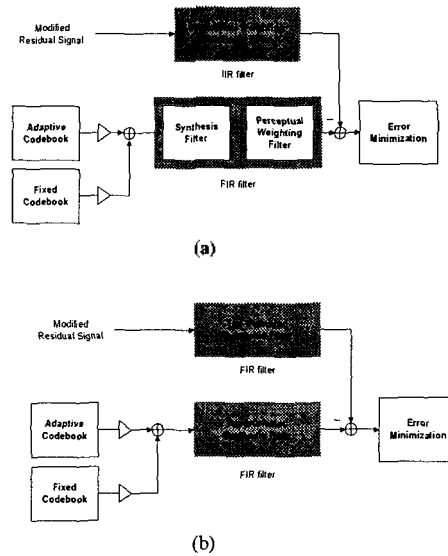


그림 1. 기존의 EVRC 부호화기와 제안된 부호화기 비교

- (a) 기존의 부호화기
- (b) FIR 가중 합성 필터를 사용하는 부호화기

Fig. 1. Comparison of EVRC encoder and proposed encoder, (a) Traditional EVRC encoder, (b) Proposed encoder using FIR weighted synthesis filter.

고정 코드북 탐색을 위한 목적 신호를 만들 때는 식 (6)과 같이 주어지는 IIR 가중 합성 필터를 사용하지만 코드북 탐색 과정에서는 이 필터의 임펄스 응답으로 이루어지는 FIR 가중 합성 필터를 사용한다.

$$H_1(z) = \left(\frac{1}{\hat{A}_q(z)} \right) W(z) = \frac{1}{\hat{A}_q(z)} \left(\frac{\hat{A}(z/\gamma_1)}{\hat{A}(z/\gamma_2)} \right) \quad (6)$$

여기서 $W(z)$ 는 가중필터이며 $\hat{A}(z)$ 와 $\hat{A}_q(z)$ 는 각각 서브프레임에서 선형 보간하여 얻은 LPC 계수와 이를 양자화하여 얻은 LPC 계수를 필터 계수로 갖는 예측 필터이다. γ_1 과 γ_2 는 각각 0.9와 0.5이며 가중 필터의 주파수 응답을 결정한다. 서브프레임 길이와 같은 가중 합성 필터를 사용하므로 임펄스 응답의 길이를 줄인다면 코드북 탐색 과정의 계산량을 줄일 수 있다. 그러나 임펄스 응답의 길이를 줄이면 줄여진 뒷부분의 손실로 인해 왜곡이 발생하여 음질의 저하를 초래한다. 이와 달리 G.729A[5]에서처럼 양자화된 LPC 계수에 기반하는 가중 필터를 사용하면 이러한 문제를 해결할 수 있다. 본 논문에서는 가중 필터를 다음과 같이 설정하며 γ_3 을 0.85로 고정하였다.

$$H_2(z) = \frac{1}{\hat{A}_q(z)} \left(\frac{\hat{A}_q(z)}{\hat{A}_q(z/\gamma_3)} \right) = \frac{1}{\hat{A}_q(z/\gamma_3)} \quad (7)$$

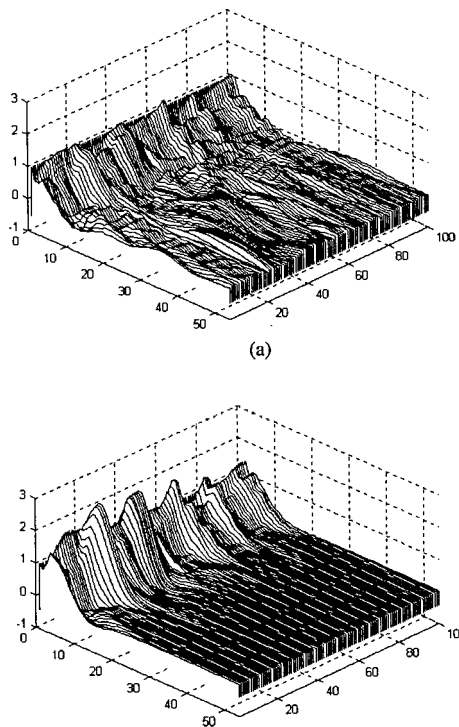


그림 2. 가중 합성 필터의 임펄스 응답 비교

(a) $\gamma_1 = 0.9$ $\gamma_2 = 0.5$ (b) $\gamma_3 = 0.85$

Fig. 2. Comparison of impulse response of weighted synthesis filter.

(a) $\gamma_1 = 0.9$, $\gamma_2 = 0.5$, (b) $\gamma_3 = 0.85$.

그림 2의 (a)와 (b)는 각각 식 (6)과 식 (7)의 두 가중 합성 필터의 임펄스 응답을 보여주고 있다. 그림 2의 (a)를 보면 임펄스 응답의 뒷부분이 길게 늘어짐을 알 수 있다. 그러므로 뒷부분의 영향을 무시하고 응답의 길이를 줄인다면 이로 인한 왜곡이 발생한다. 그러나 그림 2의 (b)와 같은 가중 합성 필터의 임펄스 응답이 20탭 이내에 존재하므로 20탭 정도의 임펄스 응답을 사용하여도 기존의 가중 합성 필터를 사용한 경우에 비해 주관적 음질은 저하되지 않음을 실험을 통해 확인하였다.

줄여진 임펄스 응답을 사용하면 임펄스 응답의 상관행렬이 계산과 목적 신호를 계산할 위한 필터링 과정 그리고 필터 상태의 갱신에서의 계산량을 줄일 수 있다. 식 (7)과 같은 합성 필터의 임펄스 응답을 사용하면 평탄한 스펙트럼을 갖는 음성 신호에 대하여 다소 음질의 저하를 초래한다[5]는 문제점이 있지만 그러한 음성 신호의 발생 빈도는 매우 낮으므로 실제 응용에 있어서 그 차이를 느낄 수 없다. 그러나 임펄스 응답의 길이를 심하게 줄이면 성능 저하가 발생한다. 이는 목적 신호를 구할 때와 코드북을 탐색할 때 필터링 과정이 서로 다르기 때문이다. 그림 1의 (a)에서와 같이 코드북 탐색의 목적 신호는 IIR 필터링을 통해 생성하지만 코드북을 탐색할 때에는 FIR 필터링 과정을 수행한다. 그러므로 심하게 줄여진 임펄스 응답을 효과적으로 사용하기 위해서는 목적 신호

를 생성하는 과정에서도 FIR 필터를 사용하는 것이 적절하다. 본 논문에서는 [6]에서 제안된 방법을 적용하여 줄여진 임펄스 응답으로 야기되는 왜곡을 최소화하였다. 임펄스 응답의 길이를 유한한 탭으로 줄인다면 주파수 특성은 보다 평탄해지고 임펄스 응답의 길이가 서브프레임의 길이보다 작게 했을 때는 필터의 길이가 작을수록 강한 인지 가중 (perceptual weighting)을 야기한다.

그림 1의 (b)와 같이 목적 신호와 합성 신호를 생성하는데 동일한 FIR 가중 합성 필터를 사용하므로 줄여진 임펄스 응답에 의한 왜곡을 최소화할 수 있다. 또한 기존의 방법과 달리 IIR 합성 필터 대신에 FIR 합성 필터를 사용하기 때문에 필요한 필터 메모리의 갱신 과정과 hangover 연산 과정이 불필요하게 된다. 또한 줄여진 임펄스 응답을 사용하면 인지 도메인의 목적 신호와 임펄스 응답의 상호상관인 역방향 필터링 신호와 임펄스 응답의 상관행렬을 구할 때 계산량을 줄일 수 있다. 표 2는 사용된 임펄스 응답의 길이에 따른 객관적 음질 평가 결과를 보여준다. 검색 과정에서 강한 가중으로 인해 SNR 측면에서 성능 저하가 발생하였다. 그러나 이러한 결과는 항상 주관적 음질과 일치하는 것은 아니다.

표 2. 제안된 방법의 임펄스 응답의 길이와 객관적 음질평가 결과

Table 2. Taps of impulse response and results of objective measure by using FIR weighted synthesis filter.

FCB Search Method 2	Tap number of Impulse response	SNR (dB)
Original Method	55	17.268
M 2-1	10	10.820
M 2-2	7	10.675
M 2-3	5	10.560

IV. 주관적 성능 평가

주관적인 음질 평가를 위해 기존의 방법에 의해 복원된 신호와 본 논문에서 제안된 방법에 의해 복원된 신호를 듣고 음질의 우열을 가리는 선호도 평가 (preference test)를 수행하였다. 두 신호를 임의의 순서대로 두 번씩 들려준 다음 둘 중에서 음질이 나은 하나를 선택하게 하였다. 두 음성 신호에 대하여 음질의 차이를 느끼지 못한 경우에 대한 선택도 포함하였다. 주/객관적인 음질 평가를 위한 음성 데이터는 여성 화자와 남성 화자에 대해 각각 2문장을 사용하였다. 주관적 평가는 훈련된 10명의 평가자에 의하여 헤드셋을 사용하여 이루어졌다.

본 논문에서 제안된 고속 알고리즘은 프레임 단위로 수행되는 피치 검색 과정에서 3.1절에서 제안한 방법을 적용하고 고정 코드북 탐색 과정에서 위치 조합을 기존의 방법에 비해 반으로 줄이고 7탭의 FIR 가중 합성 필터를 사용한다. 그리고 제안된 고속 알고리즘에서 성능

평가를 수행한다. 표 3은 제안된 고속 알고리즘에 대한 주관적 음질 평가 결과를 보여준다. 고정 코드북 여기 신호의 양자화 과정에서 심한 가중 함수를 사용하여 객관적 음질 평가에서는 성능 저하를 보였지만 주관적 음질 평가에서는 기존의 방법에 상응하는 결과를 얻었다. 제안된 방법에 대한 주관적 평가 결과를 보면, 평가자의 25%가 기존의 방법의 음질을 선호하였고 20%가 제안된 방법의 음질을 선호하였다. 나머지 55%는 기존의 방법과 제안된 방법의 음질의 차이를 구분하지 못하였다.

표 3. 주관적 음질 평가 결과

Table 3. Results of subjective quality measure.

Perferences	Percentage (%)
EVRC	25.0
Fast EVRC	20.0
No preference	55.0

표 4는 컴퓨터 모의 실험을 통해 부호화 과정에서 소요되는 계산량을 측정한 결과이다. 각 방법에 대해 C 프로그램 상에서 모의 실험을 한 결과, 기존의 방법에 비해 제안된 방법이 부호화 과정에서 약 30%의 계산량을 개선하였다. 실시간 구현을 하였을 때에도 이와 비슷한 결과를 얻을 것으로 예측된다.

표 4. 개선된 계산량 비교

Table 4. Complexity comparison of EVRC encoder and proposed encoder.

Encoding Algorithm	Complexity (%)
Original Algorithm	100.00
Proposed Algorithm	68.73

V. 결론

본 논문에서는 현재 우리 나라의 CDMA 디지털 셀룰러폰용 음성 부호화기인 8kbps QCELP (Qualcomm CELP)의 음질을 개선시키기 위해서 새롭게 채택된 EVRC 부호화기의 알고리즘 분석을 바탕으로 많은 계산량을 요구하는 부호화 과정에서의 효율적인 고속 알고리즘을 제안하였다. 고속 알고리즘은 분석-합성 방법을 사용하는 CELP 음성 부호화기에서 가장 많은 계산량을 요구하는 고정 코드북 탐색 과정에 초점을 두었다.

EVRC는 많은 계산량을 요구하는 부호화 알고리즘이므로 실시간 구현을 위해서는 고속 알고리즘에 대한 연구가 필요하다. 부호화 과정에서 가장 많은 계산량을 차지하는 부분은 고정 코드북 탐색 과정으로 전체 계산량의 약 48% 정도를 차지한다. 먼저 펄스 위치의 제한을 두어 가능한 펄스 위치의 조합을 줄였다. 기존의 방법에 비해

펄스의 위치를 반으로 줄였을 때 주/객관적 음질 평가를 통해 음질의 저하가 없음을 확인하였다. 또 심하게 줄인 임펄스 응답을 사용하는 ETIR 방법을 적용하였다. 필터링 과정에서의 계산량 감소를 위해 가중 합성 필터도 수정하였다. 이 방법 또한 음질의 저하가 없음을 주관적 평가를 통해 확인하였다. 그리고 피치 추정 단계에서의 연산량도 개선시켰다. 제안된 고속 알고리즘을 사용하면 기존의 방법과 비교하여 부호화 과정의 평균 계산량을 감소시킬 수 있으며 실제로 CDMA 환경에서의 EVRC 단말기에 적용한다면 기존의 방법에 비해 부호화 과정에서의 전력 소모를 줄일 수 있다.

EVRC 음성 부호화기는 부호화 과정 중에서 고정 코드북 탐색 과정 다음으로 많은 계산량을 요구하는 잔차 신호의 수정 과정에서의 계산량 개선을 위한 연구를 진행하고 있다. 그리고 제안된 고속 알고리즘을 고정 소수점 DSP 프로세서를 사용하여 실시간으로 구현하고자 한다.

감사의 글

본 논문은 현대전자산업 주식회사가 지원한 CDMA 핵심 기술 개발을 위한 공동 연구의 일부로서 수행된 것입니다.

참고 문헌

1. QUALCOMM Inc., Proposed TIA/EIA/PN-3292 Standard - Enhanced Variable Rate Codec, Speech Service Option 3 for Wideband Spread Spectrum Digital Systems, Official Ballot Version, April 19, 1996.
2. W. B. Kleijn, *Analysis-by-Synthesis Speech Coding Based on Relaxed Waveform-Matching Constraints*, Ph. D. dissertation, Delft University of Technology, 1991.
3. W. B. Kleijn, P. Kroon, and D. Nahumi, "The RCELP Speech-Coding Algorithm," *European Transactions on Telecommunications*, vol. 5, no. 5, pp. 573-582, Sept/Oct. 1994.
4. Shihua Wang, et al, "An Objective Measures for Predicting Subjective Quality of Speech," *IEEE J. Select. Areas Commun.*, vol. 10, no. 5, pp. 819-829, June 1992.
5. Redwan Salami, Claude Laflamme, Bruno Bessette, and Jean-Pierre Adoul, "ITU-T G.729 Annex A : Reduced Complexity 8 kb/s CS-ACELP Codec for Digital Simultaneous Voice and Data," *IEEE Communication Magazine*, pp. 56-63, Sept. 1997.
6. H. Ohmura, and K. Mano, "A Low-Complexity 4.6 kbit/s Speech Coder Based on Extremely Truncated Impulse Response," *Proc. of IEEE Workshop on Speech Coding*, pp. 63-64, 1997.

▲ 정 성 교 (Sung-Kyo Jung)



1997년 2월 : 연세대학교 전자공학과 (공학사)
1999년 2월 : 연세대학교 대학원 전자공학과 (공학석사)
1999년 3월 ~ 현재 : 연세대학교 대학원 전기전자공학과 박사과정

※ 주관심분야 : 음성부호화, 실시간 구현, VoIP

▲ 김 남 건 (Nam-Gun Kim)



1998년 2월 : 연세대학교 전자공학과 (공학사)
2000년 2월 : 연세대학교 대학원 전자공학과 (공학석사)
2000년 3월 ~ 현재 : SK텔레콤 전임연구원

※ 주관심분야 : 음성부호화, 실시간 구현, 네트워크 프로토콜

▲ 최 응 수 (Yong-Soo Choi)

1993년 2월 : 연세대학교 전자공학과 졸업
1995년 2월 : 연세대학교 전자공학과 석사
1999년 8월 : 연세대학교 전기컴퓨터공학과 박사
1999년 9월 ~ 현재 : LG전자 선임연구원

※ 주관심분야 : 음성부호화, 음성신호처리, 실시간 구현

▲ 윤 대 희 (Dae-Hee Youn)

한국 음향학회지 제18권 제7호 참조