

G.723.1 음성부호화기와 EVRC 음성부호화기의 상호 부호화 알고리즘

준희원 김 경 태*, 정 성 교*, 윤 성 완*, 박 영 철**, 윤 대 희*, 최 용 수***, 강 태 익***

An Efficient Transcoding Algorithm For G.723.1 and EVRC Speech Coders

Kyung-Tae Kim*, Sung-Kyo Jung*, Sung-Wan Yoon* *Associate Members*
Yong-Cheol Park**, Dae-Hee Youn*, Yong-Soo Choi***
Tae-Ik Kang*** *Regular Members*

요 약

서로 다른 음성 부호화기를 사용하는 유/무선 통신망의 연동에서 각 음성 패킷간 효율적인 변환 과정이 필요하다. 이러한 패킷 변환 과정을 위해서 과거에는 이중 부/복호화 방식을 이용하였다. 그러나, 두 음성 부호화기가 이중 부/복호화 방식으로 연동될 경우, 음질 저하 및 계산량 증가, 부가적인 전달 지연 등의 문제가 발생한다. 이 논문에서는 유/무선 통신 시스템에서 널리 사용되는 ITU-T G.723.1[1]과 TIA IS-127 EVRC(Enhanced-Variable-Rate-Codec)[2] 음성부호화기 간의 효과적인 연동을 위한 상호부호화 알고리즘을 제안하였다. 제안된 상호부호화 알고리즘은 크게 LSP(Line-Spectrum-Pairs) 변환, 개회로 피치 변환, 고속 적응코드북 검색, 고속 고정코드북 검색의 네 부분으로 나뉘어 진다. TMS320C62x DSP를 사용하여 구현해 본 결과, 제안된 상호부호화 알고리즘이 기존의 이중 부/복호화 과정에 비해 30% ~ 35% 정도 계산량을 개선하며, 적은 지연 시간으로 동등한 주/객관적 음질을 제공함을 확인하였다.

Key Words : tandem, transcoding; G.723; EVRC; speech coder

ABSTRACT

Interoperability is one the most important factors for a successful integration of the speech network. To accomplish communication between endpoints employing different speech coders, decoder and encoder of each endpoint coder should be placed in tandem. However, tandem coder often produces problems such as poor speech quality, high computational load, and additional transmission delay. In this paper, we propose an efficient transcoding algorithm that can provide interoperability to the networks employing ITU-T G.723.1[1] and TIA IS-127 EVRC[2] speech coders. The proposed transcoding algorithm is composed of four parts: LSP conversion, open-loop pitch conversion, fast adaptive codebook search, and fast fixed codebook search. Subjective and objective quality evaluation confirmed that the speech quality produced by the proposed transcoding algorithm was equivalent to, or better than the tandem coding, while it had shorter processing delay and less computational complexity, which is certified implementing on TMS320C62x.

* 연세대학교 전기전자공학부(kktae@mcs.p.yonsei.ac.kr), ** 연세대학교 정보기술학부(young@mcs.p.yonsei.ac.kr),

*** LG전자(cando@lge.com)

논문번호 : 020171-0418, 접수일자 : 2002년 4월 18일

※본 연구는 LG전자 지원으로 수행되었습니다.

I. 서론

현재 음성통신망은 사용자에게 다양한 품질의 서비스를 하기 위해서, 여러 음성 부호화기를 사용한다. 서로 다른 음성부호화기를 사용하는 두 음성 통신망의 통신 과정에서 음성부호화기의 연동이 발생한다. 즉, 다른 음성부호화기를 사용하는 두 endpoint 사이의 통신에서 한 음성부호화기의 비트열을 다른 음성부호화기의 비트열로 변환 시켜주는 작업이 필요하다. 이 역할을 위해 현재 두 음성부호화기를 이중 부/복호화 방식으로 연동하여 사용한다. 하지만, 이중 부/복호화 방식은 다음과 같이 여러 문제점을 발생 시킬 수 있다. (i) 음성 신호를 다른 음성부호화기를 사용해서 두번씩 부호화/복호화해야 하므로, 음질저하가 발생한다. (ii) 두 음성부호화기를 동시에 사용하므로 높은 계산량과 많은 메모리가 필요하다. (iii) LPC 분석이나 잔차신호 수정과정을 위해서 예견구간이 필요하므로 부가적인 지연이 발생한다. 서로 다른 부호화 과정을 갖는 두 음성부호화기의 연동에서 발생하는 문제를 해결하기 위해 효율적으로 송신 부호화기의 비트열을 수신 부호화기의 비트열로 변환시켜주는 과정이 필요하다. 이러한 변환과정을 상호부호화라 한다.

상호부호화에 관한 다양한 알고리즘이 연구되고 있다[3][4][5]. [4]에서 제안된 알고리즘들은 bitstream-mapping 방법을 적용하였다. 이 알고리즘들은 LSP 변환과 코드북 이득값 변환 과정에서 parameter-mapping 방법을 이용하여 적은 계산량으로 이중 부/복호화에 상응하는 음질을 제공한다. 이러한 bitstream-mapping 방법은 ITU-T G.729와 TIA IS-641과 같이 동일한 구조를 갖는 음성부호화기의 상호부호화에 적용이 가능하지만 서로 다른 프레임 크기와 부호화 과정을 갖는 부호화기 간의 상호부호화에는 적합하지 않다. 이러한 문제를 해결하기 위한 상호부호화 알고리즘에 관한 연구도 진행되어 왔다[3][5].

본 논문에서 VoIP(Voice-over-IP)에서 사용되는 G.723.1과 디지털 셀룰러 시스템에서 사용되고 있는 IS-127 EVRC간의 상호부호화 알고리즘을 제안했다. G.723.1와 EVRC 음성부호화기는 서로 다른 프레임 크기, 적응코드북과 고정코드북 구조를 가지므로, LSP 변환 과정 뿐만 아니라 여기신호 변환 과정이 필요하다. 제안된 상호부호화 알고리즘은 크게 LSP 변환, 개회로 피치 변환, 고속 적응코

드북 검색, 고속 고정코드북 검색으로 이루어진다. LSP 변환 과정에서 선형보간 방법을 이용하여 LSP 파라미터를 변환하므로 선형예측분석 과정의 예견구간으로 인한 부가적인 지연을 줄일 수 있다. 개회로 피치 변환 과정에서는 피치완만화를 통해서 수신된 적응코드북 피치를 개회로 피치로 변환시켜 적은 계산량으로 합성 신호의 음질을 개선시켰다. 여기 신호의 변환을 위해서는 적응 코드북과 고정 코드북 검색과정에 고속 알고리즘을 적용하여 연산량을 개선시켰다.

복잡도를 축정을 위해 제안된 상호부호화 알고리즘들을 고정 소수점 DSP를 사용하여 구현하였다. 이중 부/복호화 방식에 비해, 제안된 상호부호화 알고리즘들은 30~35%정도의 계산량 감소를 보였다. 성능평가를 위해 주/객관적 음질 평가를 수행한 결과, 제안된 알고리즘이 이중 부/복호화 과정과 동등한 음질을 제공함을 확인할 수 있었다.

II. G.723.1와 EVRC 비교

멀티미디어 통신용 음성 부호화 표준안으로 제정된 ITU-T G.723.1 음성부호화기[1]는 5.3 kbps와 6.3 kbps의 두개의 전송율을 제공하는 CELP(Code-Excited-Linear-Prediction) 유형의 부호화기이다. 현재 G.729와 함께 데이터 통신망을 통해 음성을 전송하고자 하는 VoIP 응용 분야에 널리 사용되고 있다. EVRC[2]는 1996년에 8 kbps 음성 부호화 표준안으로 채택되었다. 현재 우리나라 CDMA 디지털 셀룰러 시스템에서 8 kbps QCELP(Qualcomm-CELP)와 함께 서비스 중에 있으며 PCS에서 사용하고 있는 13 kbps QCELP에 버금가는 음질을 제공하는 우수한 성능의 음성 부호화기이다.

두 부호화기의 특징을 다음의 표 1로 간단히 정리하였다. 표 1에서 볼 수 있듯이 두 부호화기 모두 선형예측 기반의 분석-합성(Analysis-by-Synthesis) 구조를 가지고, 고정코드북 검색을 위해 ACELP (Algebraic-CELP) 방식을 사용하고 있다. 같은 차수의 LPC, LSP 양자화와 ACELP 기반의 고정코드북 구조는 두 부호화기를 효율적으로 연동할 수 있는 가능성을 보여준다. 그러나 두 부호화기는 프레임 구조와 적응코드북 구조에서 차이점을 보일 뿐만 아니라 고정코드북에서 검색하는 펄스 수가 서로 다르다. 이러한 부호화기 구조의 차이점 때문에 bitstream mapping 방식과 같은 상호

표 1. G.723.1와 EVRC의 특징 비교

	G.723.1	EVRC
Frame Size	30 ms	20ms
Lookahead	7.5 ms	10 ms
Bit-rate	5.3 / 6.3 kbps	8.55/4/0.8 kbps
Adaptive Codebook	5-tap integer pitch predictor	Fractional pitch predictor RCELP(Relaxation CELP)
Fixed codebook	MP-MLQ / ACELP	ACELP
Application	H.323 telephony Video telephony	CDMA DCS

부호화 알고리즘을 적용하는 것이 불가능하다. 본 논문에서 음성부호화기의 구조적 차이를 고려한 상호부호화 알고리즘을 제안한다. 다음 절에서 본 논문에서 제안하는 상호부호화 과정에 대해 자세히 설명한다.

III. G.723.1에서 EVRC로의 상호부호화 알고리즘

1. G.723.1에서 EVRC로의 상호부호화 알고리즘

제안된 G.723.1(5.3 kbps)에서 EVRC로의 상호부호화 알고리즘은 크게 LSP 변환과정, 개회로 피치 변환과정과 고속 고정코드북 검색과정으로 구성된다[5]. 그리고 두 부호화기의 프레임 길이를 고려하여 30ms 프레임길이를 갖는 G.723.1의 2개의 프레임이 20ms 프레임길이를 갖는 EVRC의 프레임 3개로 변환된다. G.723.1에서 EVRC의 상호부호화 과정은 그림 1과 같다.

상호 부호화 과정을 간략히 설명하면, 우선 G.723.1의 LSP를 프레임 크기를 고려하여 선형 보간하고 EVRC의 양자화 과정을 거쳐 복호화단으로 전송된다. 그리고 피치 완만화(pitch smoothing) 방법을 적용하여 개회로 피치값을 변환한다. 두 변환과정이 끝나면 G.723.1의 적응 코드북과 고정 코드북 성분으로 여기 신호를 만들고, 이 여기신호는

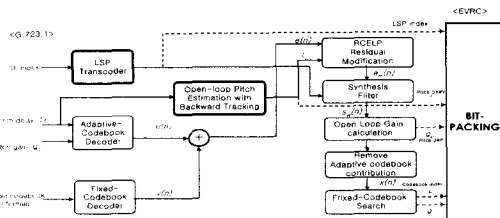


그림 1. G.723.1에서 EVRC로의 상호부호화 과정

개회로 피치 변환과정을 통해 얻어진 피치 컨투어(pitch contour)에 맞추기 위해 EVRC의 잔차신호 수정과정을 거쳐 수정된 음성신호를 만들게 된다. 수정된 음성신호에 대해 적응 코드북 검색을 수행하여 폐회로 피치와 이득값을 구하고, 제안된 고속 코드북 검색을 통해 고정코드북 파라미터를 추출한다.

1) 선형 보간을 이용한 LSP 변환 과정

제안된 LSP 변환과정에서는 G.723.1 프레임의 LSP 정보를 선형보간법[3]을 통해 EVRC의 LSP 값을 구한다. 선형 보간에 쓰이는 가중치는 (1)과 같다.

$$\begin{aligned}
 L_{B_i}(j) &= \alpha L_{A_i}(j) + (1 - \alpha)L_{A_1}(j) \\
 L_{B_2}(j) &= \beta L_{A_1}(j) + (1 - \beta)L_{A_2}(j) \\
 L_{B_3}(j) &= \gamma L_{A_1}(j) + (1 - \gamma)L_{A_2}(j)
 \end{aligned} \tag{1}$$

$$1 \leq j \leq 10$$

여기서 i 는 프레임 인덱스이며, L_{A_1} 와 L_{B_i} 는 각각 G.723.1과 EVRC의 번째 프레임의 LSP이다. 그리고 α, β, γ 는 보간상수이다.

이러한 직접 변환 방법은 LPC 계수를 구하기 위한 자기상관계수, Durbin Recursion과 LPC를 LSP로 변환하기 위한 과정이 필요하지 않으므로 이중 부/복호화 방식보다 계산량을 크게 줄일 수 있다. 그리고 이중 부/복호화 과정에서 얻어진 LPC 스펙트럼보다 원신호의 LPC 스펙트럼에 더 일치함을 확인했다. 또한, LPC 분석을 위한 예견구간이 필요하지 않으므로 부가적인 지연을 줄일 수 있다.

EVRC에서는 LPC분석에서 뿐만 아니라, 잔차신호 수정과정에서도 예견구간을 사용한다. 하지만 이 과정에서 실제로 EVRC의 예견구간 모두를 필요로 하지는 않는다. 잔차신호 수정과정에서 펄스를 30 샘플(3.74 ms)이상을 이동시키지 않는다는 특징을 이용해서, 제안된 방법에서는 잔차신호 수정과정에서 펄스의 위치를 찾기 위한 예견구간 검색 범위를 30 샘플로 제한시켰다. 검색범위를 30 샘플로 제한했을 경우에도 수정되기 전 방법으로 합성한 신호와 전혀 차이가 없음을 확인하였다.

수정된 방법은 잔차신호 수정과정에 필요한 10 ms의 예견구간을 모두 사용하지 않고 3.75 ms만 필요로 하므로, 결과적으로 제안된 상호부호화 알고리즘은 이중 부/복호화에서 일어나는 지연시간보다

6.25 ms 적은 지연시간을 제공할 수 있다.

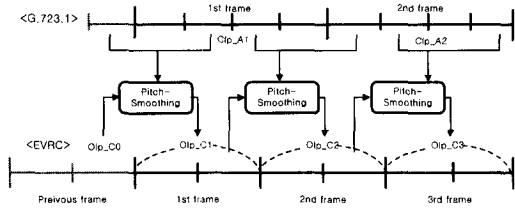


그림 2. G.723.1에서 EVRC으로의 피치 완만화를 이용한 개회로 피치 추정

2) 개회로 피치 변환

(1) 적응 코드북 피치를 이용한 개회로 피치 변환

G.723.1의 적응 코드북 피치정보와 복원된 여기 신호를 이용해서 개회로 피치를 추정한다. 개회로 피치추정은 시간영역에서 EVRC의 프레임에 가장 잘 대응되는 G.723.1의 부프레임의 피치값 주변에서만 이루어진다. 그러므로 개회로 피치 추정을 위해서 자기상관 계수 계산과 지역 최대값의 비교 과정에 소요되는 계산량을 줄일 수 있다. 그리고 여기 신호 영역에서 피치를 검색하므로 LPC 합성과정의 생략되어 이중 부/복호화 방법보다 훨씬 적은 연산량으로 개회로 피치를 결정할 수 있다. 피치 완만화를 적용한 경우가 적용하지 않은 경우보다 피치의 부정확성으로 인한 양자화 잡음을 줄일 수 있어 이중 부/복호화와 동등한 수준으로 주관적 음질을 개선할 수 있다.

(2) 피치 완만화를 이용한 개회로 피치 변환

만약, 현재 프레임에 대응되는 G.723.1의 부프레임의 피치값과 EVRC의 이전 부프레임의 피치가 서로 유사한 경우에 그림 2와 같이 피치 완만화 방법을 적용한다. 이 때 개회로 피치 추정의 후보가 되는 G.723.1의 피치값과 EVRC의 이전 부프레임의 피치값 각각에 대해 ±6 샘플 범위에서 식 (2)을 최대로 만드는 지연값을 결정한다.

$$R(k_i) = \sum_{n=0}^{79} s_w(n) s_w(n-k) \quad (2)$$

$$\left\{ \begin{array}{l} p_A - 3 \leq k_1 \leq p_A + 3 \\ p_B - 3 \leq k_2 \leq p_B + 3 \end{array} \right.$$

$i = 1, 2$

여기서 $s_w(n)$ 는 지각 가중된 음성신호, p_A 와 p_B

는 각각 G.723.1과 이전 부프레임의 적응코드북 피치값이며, k_1 과 k_2 는 해당 범위의 개회로 피치 지연값이다.

각 범위에서 R 를 최대로 하는 지연값, R 를 결정한 후, 기하평균 에너지로 정규화한다.

$$R'(m_i) = \frac{R(m_i)}{\sqrt{\sum_{j=0}^{159-m_i} \epsilon^2(j) \sum_{j=0}^{159-m_i} \epsilon^2(j+m_i)}} \quad (3)$$

$i = 1, 2$

여기서 m_1 과 m_2 는 각각 G.723.1과 이전 부프레임의 적응 코드북 피치 주변값 중에서 최대 상관계수를 가지는 지연값이다.

정규화된 각 범위에서의 최대 상관계수값은 아래의 과정과 같이 EVRC의 적응 코드북 피치에 가중치를 주어 비교하고, 개회로 피치값을 결정하게 된다. 아래의 판별식에 구체적인 EVRC의 개회로 피치값을 결정하는 과정을 보였다.

$$\text{if } \{ \{ (D_2 > D_1 + 6) \text{ and } (\beta_1 > 0.6\beta_2) \} \mid \{ (D_2 < D_1 - 6) \text{ and } (\beta_1 > 1.2\beta_2) \} \}$$

$$\text{then delay} = D_1$$

$$\text{else delay} = D_2$$

여기서 D_1 , β_1 는 이전 부프레임의 적응코드북 피치 주변에서 찾은 피치와 이득값이고, D_2 , β_2 는 G.723.1의 적응코드북 피치 주변에서 찾은 피치와 이득값이다.

3) EVRC의 고속 고정코드북 검색

EVRC의 고정 코드북 검색은 ACELP 구조를 기반으로 한다. 각 부프레임을 5 개의 트랙으로 나누어, 이중 세 트랙에서 2개의 펄스를 찾고 나머지 두 트랙에서 1개의 펄스를 찾는다. 기본적으로 depth-first tree search을 갖추고 있으며, 4가지 단일-펄스 검색 트랙 조합(codeword)에 대해서 검색한다. 제안된 상호부호화기에서는 계산량 감소를 위해 개회로 방식으로 4가지 codeword 중에서 한 가지만 선택하여 검색하는 고속 고정 코드북 검색 방식을 이용하였다. 이 고속검색을 위해서 먼저 각 트랙에 해당하는 음성신호의 에너지를 미리 계산한다. 그리고 에너지분포가 가장 적은 트랙 두개가

단일-필스 검색 트랙으로 설정된 codeword를 선택하여, 이 codeword에 대해서만 검색을 한다. 이 검색방법을 이용한 경우 큰 음질의 저하 없이 코드북 검색량을 1/4로 줄일 수 있다.

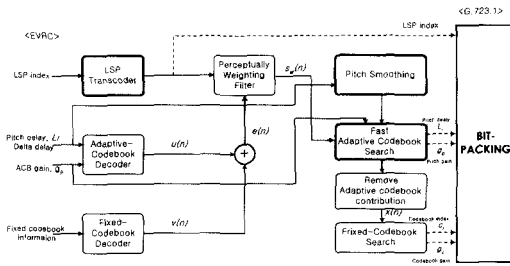


그림 3. EVRC에서 G.723.1로의 상호부호화 과정

2. EVRC에서 G.723.1로의 상호부호화 알고리즘

제한된 EVRC에서 G.723.1(5.3 kbps)로의 상호부호화 알고리즘은 반대 방향의 상호부호화의 경우와 비슷하게 LSP 변환과정, 개회로 피치 변환, 고속 적응 코드북 검색[6], 고속 고정 코드북 검색과정[6]으로 구성된다. 그리고 두 부호화기의 프레임 길이를 고려하여, 20ms 프레임길이를 갖는 EVRC의 세 프레임이 30ms 프레임길이를 갖는 G.723.1의 두 프레임으로 변환되며 그 과정은 그림 3과 같다. 반대방향의 상호부호화 과정과 같이 우선 EVRC의 LSP가 선형 보간을 통해 G.723.1의 LSP로 변환된다. 그리고 지각 가중 합성 필터를 구성하고 G.723.1의 개회로 피치 추정을 위한 목적 신호를 만든다. 개회로 피치 추정도 반대 방향의 상호부호화와 같이 두 음성 부호화기의 피치값을 이용한 피치 변환 방법을 적용하여 구현하였다. 개회로 피치를 결정한 후, 고속 적응코드북과 고속 고정코드북 검색을 통해 각각의 파라미터를 추출하게 되는데, 구체적인 과정을 3.2.1과 3.2.2에서 설명했다. 변환된 LSP와 적응 코드북과 고정 코드북 파라미터를 부호화하여 비트열을 생성하면 EVRC에서 G.723.1로의 상호부호화가 완료된다.

1) 고속 적응 코드북 검색

기존의 G.723.1 음성부호화기의 적응 코드북 검색 과정에서는 5차 피치예측기를 사용하여 지연과 이득을 동시에 검색한다. 이와 달리, 제안된 적응 코드북 변환과정은, [6]에 제시된 방법을 바탕으로, 먼저 지연을 1차 정수 피치예측기를 사용하여 검색하고, 구해진 지연값에서 5차 피치예측기의 이득값을 검색한다. 5차 피치예측기의 이득값 검색은 현재

프레임과 가장 인접한 EVRC의 부프레임의 적응 코드북 이득값에 따라 범위가 제한된다.

5.3kbps G.723.1의 적응코드북은 170 개의 5차 피치 예측기의 이득값 후보를 가진다. 제안된 고속 검색 방법에서는 미리 170 개 이득 후보값 각각에 대해 단일 피치예측기 이득값을 구해서 EVRC의 적응 코드북 이득값에 따라 분류해둔다. 이 때, 단일 피치예측기 이득값은 170 개의 5차 피치 예측기에 실제 음성신호를 통과 시켜 얻어낸 이득값이다. EVRC는 8 개(3 비트)의 적응코드북 이득값 후보를 가지므로, 170개의 이득값 후보는 8개의 범주로 나뉜다. 각 범주는 전체 170개의 후보값 중 크기는 34개, 작게는 7개만을 포함하고 있어, 검색범위를 크게 줄일 수 있다. EVRC의 적응 코드북 이득값을 이용한 G.723.1의 적응 코드북 이득값 검색 범위를 제한한 방법으로 음성신호를 부/복호화한 결과, 주관적 음질의 성능 저하가 없음을 확인하였다.

2) 고속 고정 코드북 검색

5.3 kbps G.723.1의 고정 코드북 검색을 위해 [6]에서 제안된 고정 코드북 고속 검색 방법을 사용했다. 즉, 기존의 nested-loop search 방법을 depth-first tree search으로 대체하였다. 이 알고리즘을 사용할 경우, 주관적 음질 저하 없이 검색 경우의 수를 1/16로 줄일 수가 있었다.

IV. 상호 부호화 알고리즘의 성능 평가

제안된 상호 부호화 알고리즘에 대해 객관적 음질평가와 주관적 음질평가 및 계산량 측정을 수행하여 이중 부/복호화 방법과 성능을 비교하였다. 객관적 음질평가의 척도로는 8 초 길이의 한국인 남녀 화자 문장에 대한 LPC-CD (LPC Cepstral-Distance)와 PSQM (Perceptual-Speech-Quality-Measure)을 사용하였고, 주관적 음질 평가를 위해 선호도 조사를 했다.

1. 객관적 음질 평가

기존의 이중 부/복호화 방법과 본 논문에서 제안한 상호부호화 알고리즘을 통해 얻어진 부/복호화된 음성 샘플의 LPC-CD와 PSQM을 측정하여 표 2에 정리하였다. 표 2에서 보는 바와 같이 여성화자와 남성화자 모두 LPC-CD와 PSQM 값이 이중 부/복호화의 경우보다 제안한 상호 부호화 방법이

더 작은 수치를 나타냈다. 이는 상호부호화기의 LPC 스펙트럼이 이중 부/복호화 과정의 LPC 스펙트럼보다 원신호에 가깝다는 결과와도 일치한다.

표 2. 객관적 음질평가(AtoB : G.723.1에서 EVRC로의 상호 부호화기, BtoA : EVRC에서 G.723.1로의 상호 부호화기)

	LPC-CD(dB)		PSQM	
	Female	Male	Female	Male
Tandem(AtoB)	5.51	4.46	3.31	3.24
Transcoding(AtoB)	5.14	3.56	3.08	2.74
Tandem(BtoA)	4.94	3.95	3.07	2.80
Transcoding(BtoA)	4.76	3.17	3.05	2.79

2. 주관적 음질 평가

주관적 음질평가로는 20명의 청취자에게 한국인 남/녀 화자 문장을 들려준 후, 우열을 가리는 선호도 조사를 수행하였다. 표 3은 이중 부/복호화 방법과 제안된 상호 부호화 방법의 선호도 평가 결과를 보여준다.

표 3. 주관적 선호도 평가 결과

Preference	G.723.1⇒EVRC		EVRC⇒G.723.1	
	Female	Male	Female	Male
Tandem	35 %	20 %	40 %	35 %
Transcoding	35 %	40 %	20 %	30 %
No Preference	30 %	40 %	40 %	35 %

주관적 평가의 결과, G.723.1에서 EVRC로의 상호부호화 알고리즘의 경우 이중 부/복호화 방식과 거의 비슷하거나 나은 음질을 제공하지만 EVRC에서 G.723.1로의 상호부호화의 경우 이중 부/복호화 방식보다 약간 음질이 저하되었다. 이 결과는 저전송률 부호화기에서 고전송률 부호화기로의 변환에서는 상호부호화기가, 반대 방향의 경우 이중 부/복호화 방식이 조금 더 우수한 성능을 보인다는 것을 보여준다.

3. 연산량 측정

이중 부/복호화 방법과 본 연구에서 제안한 상호 부호화 알고리즘의 계산량 비교를 위하여 TMS320-C62x의 simulator에서 소요되는 clock 수를 측정하였다. 이 실험은 단지 연산량을 비교하기 위한 목적이었으므로, 실험결과는 최적화되지 않은 것이다. 측정결과는 표 4에 나타내었다.

표 4. 상호부호화기와 Tandem의 연산량 비교

	G.723.1⇒EVRC		EVRC⇒G.723.1	
	Tandem	Transcoding	Tandem	Transcoding
LPC/LSP	930998.1	249303	306794.4	215881
Open-loop pitch estimation	263564.1	132238.3	213682.4	132302
Adaptive codebook search	620584.8	506014.2	826792	462816
Fixed codebook search	2104629.3	971666.1	624888.8	170008
Others	2954342.3	2954342.3	920985.3	920985.3
Total	6874118.6	4813563.9	2893142.9	1901992.3

제안된 상호부호화 알고리즘이 이중 부/복호화 방식에 비해, G.723.1에서 EVRC 방향의 경우 30.0%, EVRC에서 G.723.1 방향의 경우 34.3% 계산량이 감소되었음을 확인할 수 있다.

4. 지연 시간

같은 네트워크 조건에서 상호부호화기에 의한 지연과 이중 부/복호화 방식에 의한 지연시간의 차는 부호화 과정에서 발생하는 알고리즘 지연에 의해 결정된다. 제안된 상호부호화 알고리즘은 선형 예측 분석을 위한 예견구간이 필요하지 않기 때문에 예견 구간만큼의 지연시간의 이득을 볼 수 있다. EVRC에서 G.723.1로의 상호부호화 경우 G.723.1의 예견구간이 7.5 ms이므로 이중 부/복호화 방식에 비해 7.5 ms의 지연 시간을 단축시킬 수 있다. G.723.1에서 EVRC로의 상호부호화 경우 III 장에서 보였듯이 이중 부/복호화 방식에 비해 6.25 ms의 지연 시간을 단축할 수 있다.

V. 결론

본 논문에서는 음성 통신망에서의 이중 부/복호화로 발생하는 계산량 증가, 전달시간 지연 등의 문제를 극복할 수 있는 8.55kbps EVRC 음성부호화기와 5.3kbps G.723.1 음성부호화기 간의 상호부호화 알고리즘을 제안하였다. LPC-CD, PSQM 등의 객관적 음질평가와 주관적 선호도 음질평가 결과, 양방향 모두 제안된 상호부호화 알고리즘과 이중 부/복호화가 비슷한 음질을 제공할 수 있었다.

TMS320C62x로 시뮬레이션해 본 결과, 연산량은 이중 부/복호화 방법에 비해 약 30-35% 감소하였다. 이 알고리즘은 SMV(Selective-Mode-Vocoder)나 AMR(Adaptive-Multi-Rate)과 같은 CELP 유형의 음성부호화기 간의 상호부호화에 적용 가능하다.

참 고 문 헌

[1] ITU-T Rec. G.723.1 "Dual-rate Speech Coder For Multimedia Communications Transmitting at 5.3 and 6.3 kbit/s," 1996.

[2] TIA/EIA/IS-127 Enhanced variable rate codec, speech service option 3 for wideband spread spectrum digital systems, 1996.

[3] Sung-Wan Yoon, Sung-Kyo Jung, Young-Cheol Park, and Dae-Hee Youn, "An Efficient Transcoding Algorithm For G.723.1 And G.729A Speech Coders," in Proc. Eurospeech 2001, pp.2499-2502, 2001. Sep.

[4] Hong-Goo Kang, Hong Kook Kim, and Richard V. Cox "Improving transcoding capability of speech coders in clean and frame erased channel environments," Speech Coding, Proceedings of IEEE Workshop 2000, pp78-80, 2000

[5] Kyung Tae Kim, Sung Kyo Jung, Young Cheol Park, Yong Soo Choi, Dae Hee Youn, "An efficient transcoding algorithm for G.723.1 and EVRC speech coders," in Proc. IEEE VTS 54th Vehicular Technology Conference (VTC 2001), Vol. 3 ,pp. 1561-1564, Oct. 2001

[6] Sung-Kyo Jung, Young-Cheol Park, Sung-Wan Yoon, Kyoung-Tae Kim, and Dae-Hee Youn, "Efficient Implementation of ITU-T G.723.1 Speech Coder for Multichannel Voice Transmission and Storage," in Proc. of Eurospeech 2001, pp.2017-2020, 2001. Sep.

김 경 태(Kyung-Tae Kim)

준회원



2000년 2월 : 연세대학교 기계 전자공학부 졸업
 2002년 2월 : 연세대학교 전기 전자공학과 석사
 2002년 3월~현재 : 연세대학교 전기전자공학과 박사과정

<주관심분야> 음성신호처리

정 성 교(Sung-Kyo Jung)

한국통신학회논문지 28권 제3호 참고

윤 성 완(Sung-Wan Yoon)

한국통신학회논문지 28권 제3호 참고

박 영 철(Yong-Cheol Park)

한국통신학회논문지 28권 제3호 참고

윤 대 희(Dae-Hee Youn)

한국통신학회논문지 28권 제3호 참고

최 옹 수(Yong-Soo Choi)

한국통신학회논문지 28권 제3호 참고

강 태 익(Tae-Ik Kang)

한국통신학회논문지 28권 제3호 참고