

논문 2006-43SP-3-7

# EVRC와 G.729A 간의 상호부호화

( A Transcoding Algorithm between EVRC and G.729A )

권 구 락\*, 고 성 제\*\*

( Goo-Rak Kwon and Sung-Jea Ko )

## 요 약

본 논문에서는 EVRC와 G.729A 음성부호화기를 위한 상호부호화 알고리즘을 제안한다. 다른 음성 표준을 사용하는 통신망 간에 음성신호를 연동시키는 가장 간단한 방법은 이중 부/복호화 (tandem coding) 방법이다. 이 방법은 두 번의 부/복호화 과정을 거치기 때문에 많은 계산량이 요구되며 아울러 음성 지연이 발생하게 된다. 이러한 문제점들을 개선하기 위하여 상호부호화를 사용한다. 상호부호화는 LSP (Line Spectral Pair) 변환과 피치 지연 변환 그리고 지연 시간 단축 알고리즘을 통하여 수행한다. 제안된 알고리즘은 18~22%의 적은 계산량과 5~10 ms의 짧은 지연으로 상호 부/복호화에 상응하는 음성 품질을 제공함을 실험을 통해 확인할 수 있다.

## Abstract

This paper presents an effective algorithm for transcoding between the Enhanced Variable Rate Codec(EVRC) and G.729A. The simplest way to communicate between heterogeneous speech networks is the cascade connection of two different codecs, called tandem coding. However, tandem coding not only produces high computational loads, but also makes long delay. These problems can be solved by using the transcoding algorithm. The proposed algorithm consists of LSP (Line Spectral Pair) conversion, pitch delay conversion and algorithm for reduction of delay. Experimental results show the proposed algorithm produces lower computational complexity, shorter algorithm delay, and similar speech quality when compared with the tandem algorithm.

**Keywords :** 후처리기법, SCS, QCS, 웨이브렛 변환

## I. 서 론

유선 전화망은 펄스부호 변조 (Pulse Code Modulation, PCM) 방식의 ITU G.711<sup>[1]</sup> 음성 부호화기를 사용하는 데에 반해 최근 활발히 보급되고 있는 Voice over Internet Protocol (VoIP) 망에서는 Code Excited Linear Prediction (CELP) 방식의 ITU G.723<sup>[2]</sup> 과 ITU G.729<sup>[3][4]</sup> 음성부호화기가 널리 사용되고 있다. 그리고 이동통신 환경에서도 CELP 방식인 Adaptive Multi Rate (AMR)와 Enhanced Variable Rate Codec (EVRC)<sup>[5]</sup> 음성부호화기 등이 표준으로 널리 사용되고 있다.

멀티미디어 통신과 네트워크 환경이 급속히 발전해 감에 따라 서로 다른 네트워크 간 통신의 필요성이 강조되고 있다. 각각의 통신망에서 사용되는 음성신호를 연동시키려면 두 번의 부/복호화 과정을 거치는 tandem coding이 필요하다. Tandem 방식에서는 먼저 원천 음성부호화기(Source Coder)로부터 전송된 비트열 (Bit-stream)을 복호화(Decoding)하여 PCM 형식의 음성으로 합성한 후, 이를 다시 목적 음성부호화기 (Target Coder) 형식으로 부호화(Encoding)하는 과정으로 상호부호화를 수행한다. 이 때 통화품질의 열화와 함께 많은 계산량이 요구되며 전달 시간이 지연되는 문제가 발생한다. 이러한 문제점들을 해결하기 위하여 두 번의 부/복호화 과정을 한 번의 상호부호화 과정으로 바꾸어 계산량과 전달 시간을 줄이면서도 통화품질을 유지하는 기술이 중요한 문제로 대두되고 있다.<sup>[5]-[7]</sup>

\* 정회원, \*\* 평생회원, 고려대학교 전기전자전공공학부 (School of Electrical Engineering, Korea University).  
접수일자: 2005년9월13일, 수정완료일: 2006년4월5일

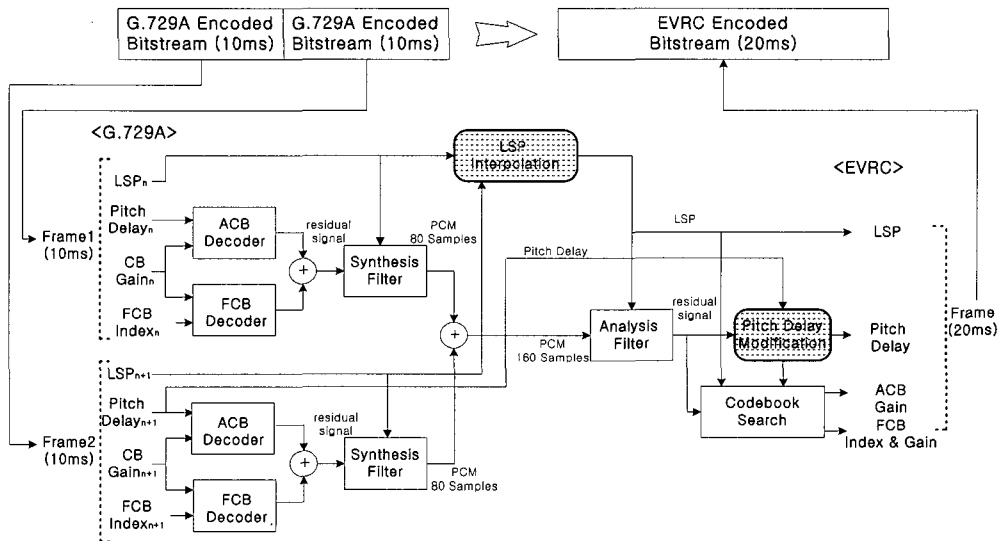


그림 1. G.729A에서 EVRC로의 상호부호화 블록 다이어그램  
 Fig. 1. Block diagram of transcoder from G.729A to EVRC.

다음은 G.729 CS-ACELP (Conjugate Structure - Algebraic Code Excited Linear Prediction)<sup>[8]</sup>와 EVRC (Enhanced Variable Rate Codec) 음성부호화기에 관해서 비교한다.

1.1. G.729A와 EVRC의 비교

G.729A와 EVRC의 특징을 비교하여 표 1에 정리하였다. 두 음성 부호화기의 차이점은 프레임의 크기와 부 프레임의 수, 피치 지연의 업데이트 주기 그리고 부호화 알고리즘 등이 있다.

본 논문에서는 CS-ACELP를 기반으로 하는 G.729A와 RCELP를 기반으로 하는 EVRC 간의 상호부호화

알고리즘을 제안하였다. 이 방법은 두 번의 부/복호화 과정을 거치기 때문에 많은 계산량이 요구되며 아울러 음성 지연이 발생하게 된다. 이러한 문제점들을 개선하기 위하여 상호부호화를 사용한다. 상호부호화는 LSP (Line Spectral Pair) 변환과 피치 지연 변환 그리고 지연 시간 단축 알고리즘을 통하여 수행한다.

본 논문은 다음과 같이 구성되어 있다. II장에서는 G.729A에서 EVRC로의 상호부호화를, III장에서는 EVRC에서 G.729A로의 상호부호화를 설명한다. IV장에서는 양방향 상호부호화에 관한 성능 평가를 수행하고 끝으로 V장에서는 이 논문의 결론을 맺는다.

II. G.729A에서 EVRC로의 상호부호화

2.1 전체적인 구조

한 프레임이 10 ms로 이루어지는 G.729A의 비트열을 20 ms의 프레임으로 이루어진 EVRC 비트열로 변환하기 위하여 G.729A의 두개의 프레임을 EVRC 하나의 프레임으로 변환하였다. 제안된 상호부호화 알고리즘의 전체적인 구조는 그림 1과 같다.

본 논문에서 제안된 G.729A에서 EVRC로의 상호부호화 알고리즘은 LSP 변환과 피치 지연 변환 과정으로 이루어져 있다. G.729A의 복호화 과정을 거쳐 얻어진 LSP와, 폐회로 피치 지연을 EVRC 부호화기에 전송함으로써 LPC 분석 과정과 피치 지연 추정 과정에서의 계산량을 감소시키고 지연 시간을 줄였다.

표 1. G.729A와 EVRC의 특징 비교  
 Table 1. Comparison of G.729A and EVRC.

	G.729A	EVRC
Input PCM	16 bit, 8kHz PCM	16 bit, 8kHz PCM
Frame	10ms (80 samples) 2 subframe	20ms (160 samples) 3 subframe
Look-ahead	5ms	10ms
LPC&LSP	10 <sup>th</sup> LPC&LSP	10 <sup>th</sup> LPC&LSP
Pitch Delay	Update every subframe (5ms)	Update every frame (20ms)
Coding Algorithm	CS-ACELP	RCELP

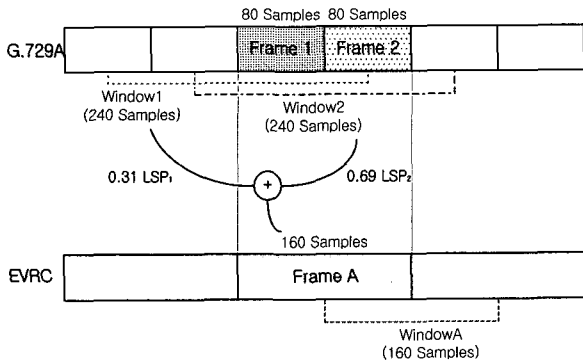


그림 2. G.729A에서 EVRC로의 LSP 변환  
Fig. 2. LSP transcoding from G.729A to EVRC.

## 2.2 LSP 변환

G.729A와 EVRC는 선형예측 부호화의 방법을 사용하고 있다. 각 프레임마다 LPC 분석을 수행하게 되는데 그림 2를 보면 G.729A와 EVRC는 LPC 분석을 위한 샘플 구간의 위치를 다르게 적용하고 있다.

G.729A는 LPC 분석을 위하여 현재 프레임에 해당하는 80 샘플과 이전 프레임의 120 샘플 그리고 다음 프레임의 40 샘플, 총 240 샘플을 고려한다. 반면 EVRC는 그림 2에서처럼 현재 프레임의 절반에 해당하는 80 샘플과 다음 프레임의 80 샘플, 총 160 샘플을 고려하여 LPC 분석을 수행한다. LPC 분석 후 얻어진 10 개의 계수들은 안정하고 양자화에 효과적인 LSP로 변환된다.

제한된 LSP 변환은 G.729A의 두 프레임에서 생성된 LSP를 EVRC 한 프레임의 LSP로 선형보간하는 과정으로 이루어진다. 선형보간 과정은 다음과 같다. 그림 2와 같이 G.729A의 프레임 1과 프레임 2를 EVRC의 프레임 A로 변환 하는데 LPC 분석 구간을 살펴보면 프레임 2의 구간이 프레임 1에 비하여 프레임 A에 유사함을 알 수 있다. 따라서 프레임 2에 가중치를 두어 식 (1)과 같이 선형보간 하였다.

$$LSP_{EVRC_n} = 0.31LSP_{G.729_{n-1}} + 0.69LSP_{G.729_n} \quad (1)$$

식 (1)의 계수들은 프레임이 인접할수록 유사성이 커지는 LSP의 특징에 의하여 실험을 통해 결정되었다. 제한된 알고리즘을 통해 선형 예측 분석과정에서의 계산량을 단축할 수 있었다.

## 2.3 피치 지연 변환

EVRC와 G.729A는 피치 지연을 구하는 방법이 서로 다르다. G.729A는 10 ms 프레임 단위로 개회로 분석을

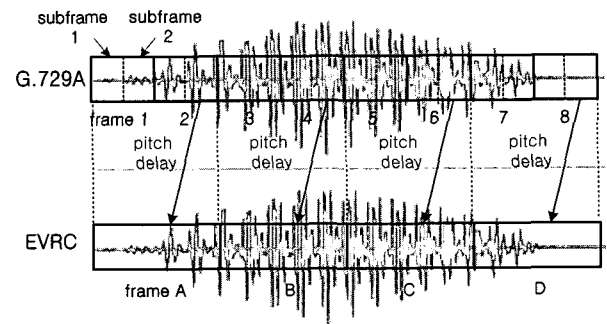


그림 3. 피치 지연 변환 알고리즘  
Fig. 3. Conversion algorithm of pitch delay.

거친 후에 5 ms 서브프레임 단위로 폐회로 분석을 수행하여 한 프레임당 2 개의 피치 지연을 전송하는 반면 EVRC는 RCELP 알고리즘을 기반으로 폐회로 분석 없이 개회로 피치 탐색만으로 결정된 피치 지연을 20 ms 한 프레임 단위로 전송한다. 따라서 5 ms로 이루어진 G.729A 서브프레임의 4 개 피치 지연을 20 ms EVRC 한 프레임의 피치 지연으로 변환하는 과정이 필요하다. 또한 EVRC는 20~120 샘플 구간에서 피치 지연을 구하는 반면 G.729A에서는 20~143 샘플 구간에서 피치 지연을 구한다. 따라서 단순히 G.729A의 피치 지연 값을 선형보간하여 사용할 경우 결과 음성 품질이 떨어지게 된다.

제한된 피치 지연 변환 과정은 다음과 같다. 그림 3과 같이 G.729A의 두 번째 프레임의 두 번째 서브프레임이 갖는 피치 지연 값을 EVRC 프레임의 피치 지연 값으로 대체한다. 이때 G.729A에서 가져온 피치 지연이 120 보다 큰 경우에는 이전 프레임의 피치지연을 취한다.

피치 지연 보정 과정에서는 G.729A에서 얻어온 피치 지연의 서브프레임  $\pm 3$  구간 내에서 자기 상관함수를 최대로 만드는 피치 지연과 EVRC 이전 프레임에서의 피치 지연을 비교하여 자기 상관함수를 최대로 하는 피치 지연을 택하였다.

## 2.4 지연 시간 단축 알고리즘

EVRC 부호화 과정에서는 LPC 분석을 위하여 그림 4 (a)와 같이 이전 프레임 80 샘플, 현재 프레임 160 샘플, 이후 프레임 80 샘플의 총 320 샘플의 음성 버퍼를 사용하게 된다. 그러나 상호부호화에서는 LPC 분석 과정을 생략하여 그림 4(b)와 같이 240 샘플의 음성 버퍼만을 사용함으로써 10 ms의 지연시간을 줄였다.

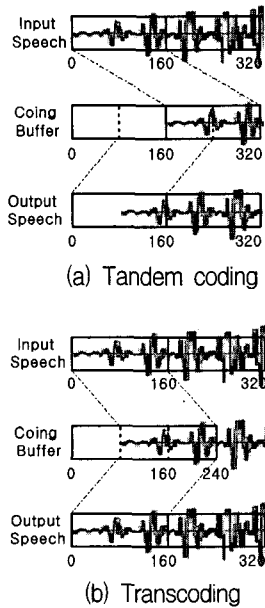


그림 4. 상호부호화를 이용한 지연시간 단축  
Fig. 4. Delay-time reduction using transcoding.

III. EVRC에서 G.729A로의 상호부호화

3.1 전체적인 구조

한 프레임이 20 ms로 이루어지는 EVRC의 비트열을 10 ms의 프레임으로 이루어진 두 개의 G.729A의 비트열로 변환하였다. 제안된 상호부호화 알고리즘의 전체적인 구조는 그림 5와 같다.

EVRC에서 G.729A의 상호부호화 알고리즘도 반대의 경우와 마찬가지로 LSP 변환과 피치 지연 변환 과정으로 이루어져 있다. EVRC 복호화 과정을 거쳐 얻어진 LSP와 피치 지연을 G.729A부호화기에 전송함으로써 LPC 분석 과정과 피치 지연 추정 과정에서의 계산량을

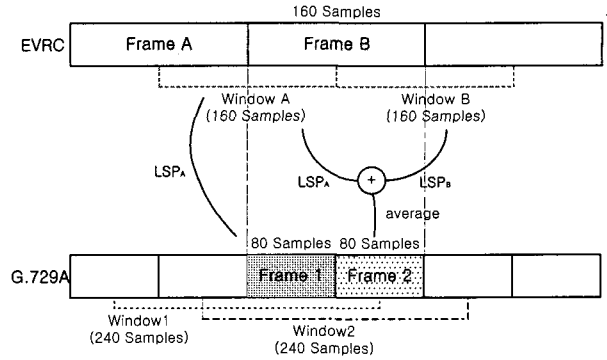


그림 6. EVRC에서 G.729A로의 LSP 변환  
Fig. 6. LSP transcoding from EVRC to G.729A.

감소시키고 지연 시간을 줄였다.

3.2 LSP 변환

EVRC에서 G.729A로의 LSP 변환은 EVRC와 G.729A의 선형예측분석 구간에 따라 LSP 직접전달과 선형보간으로 이루어진다.

EVRC와 G.729A의 LPC 분석 구간은 그림 6과 같으며 LSP 변환 식은 식 (2)와 같다.

$$LSP_{G.729, a} = LSP_{EVRC, a}$$

$$LSP_{G.729, b} = 0.5LSP_{EVRC, a} + 0.5LSP_{EVRC, b} \tag{2}$$

3.3 피치 지연 변환

G.729A에서 피치 지연을 구하는 알고리즘은 개회로 피치 분석을 통해 대략의 피치 지연을 얻은 후에 폐회로 피치 분석을 거쳐 정확한 피치 지연을 얻어내는 과정을 거친다. EVRC에서 G.729A로의 상호부호화에서는 EVRC의 피치 지연을 G.729A의 개회로 피치 지연으로

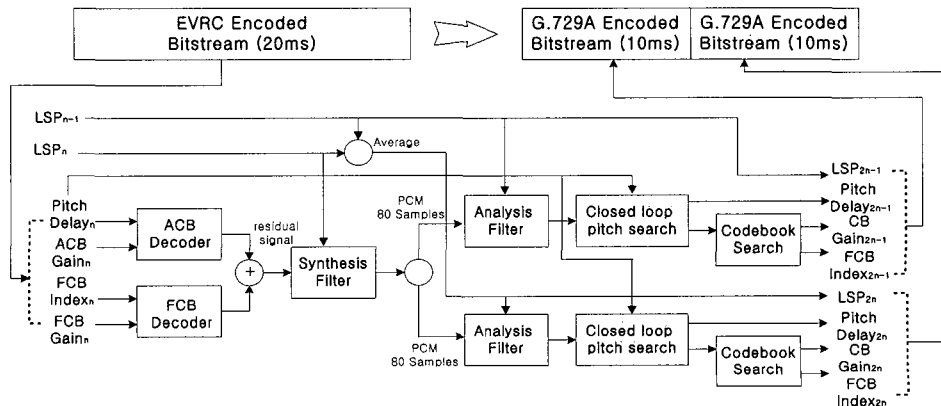


그림 5. EVRC에서 G.729A로의 상호부호화 블록 다이어그램  
Fig. 5. Block diagram of transcoder from EVRC to G.729A.

대체시킴으로서 연산량을 줄일 수 있었다.

G.729A에서 EVRC로의 상호부호화에서는 대체된 피치 지연을 보정해주는 과정을 거쳤으나 EVRC에서 G.729A로의 상호부호화에서는 추가적인 보정 과정 없이 사용하였다. EVRC의 피치 지연 범위는 (20~120샘플) G.729A의 피치 지연 범위에 (20~143샘플) 포함될 뿐 아니라 G.729A 부호화 과정에서 폐회로 피치 분석을 거쳐 정확한 피치 지연을 얻어 내기 때문이다.

### 3.4 지연 시간 단축 알고리즘

G.729A에서 EVRC로의 상호부호화와 마찬가지로 EVRC에서 G.729A로의 상호부호화시에도 G.729A의 부호화 과정의 LPC 분석을 위한 음성 버퍼를 그림 7과 같이 240 샘플에서 200 샘플로 단축시킴으로서 음성 지연을 줄였다.

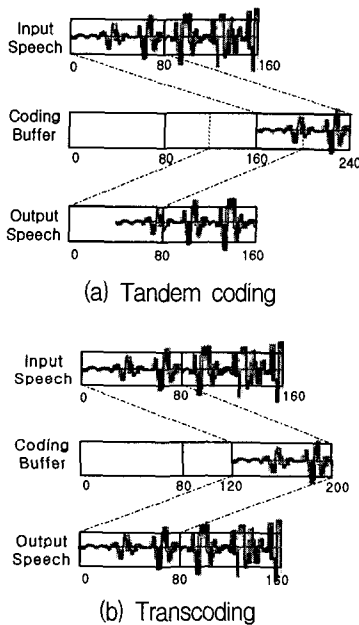


그림 7. 상호부호화를 이용한 지연시간 단축  
Fig. 7. Delay-time reduction using transcoding.

## IV. 성능 평가 및 분석

상호부호화와 이중 부/복호화 간의 성능 비교를 위하여 계산량, 지연시간, 주/객관적 음질 평가를 실시하였다. 실험용 음성 데이터는 남녀 각각 3명으로부터 5 초 분량의 16 비트 8 kHz로 샘플링된 한국어 음성을 사용하였다. 음성 파일은 8kHz로 표본화 되었고 13 비트 선형 PCM 형식이며 전송 과정의 오류는 없다고 가정한다. EVRC 및 G.729A 그리고 상호부호화 알고리즘은

표 2. 계산량 비교 (WMOPS)

Table 2. Comparison of computational complexity.

	G.729A→EVRC		EVRC→G.729A	
	Male	Female	Male	Female
Tandem Coding	21.177	21.358	19.649	19.637
Transcoding	17.279	17.400	15.343	15.368
Reduction Rate	18%	19%	22%	22%

고정 소수점 기반의 C 언어로 구현을 하였다.

### 4.1 계산량

상호부호화와 이중 부/복호화 알고리즘의 계산량 비교를 위하여 WMOPS (Weighted Million Operations Per Second)를 측정하였다. LPC 분석과정과 피치지연 분석과정을 단축시킴으로써 표 2에서와 같이 18~22%의 계산량 감소를 확인할 수 있다.

### 4.2 지연 시간

음성 통신 시스템에서 총 지연 시간은 알고리즘 지연, 처리 그리고 전송 지연을 합한 지연을 말한다. 그러나 전송 지연은 네트워크 상황에 따라 달라지고 처리 지연 또한 앞에서 계산량 비교를 통하여 비교를 하였으므로 본 절에서는 알고리즘의 지연시간만을 비교하였다. EVRC의 경우 LPC 분석을 위한 320 개의 샘플 중 80 샘플을 이후의 프레임에서 취하므로 한 프레임의 길이인 20 ms외에 10 ms의 알고리즘 지연이 발생한다. G.729A 또한 LPC 분석을 위한 240 개의 샘플 중에 40 샘플을 이후 프레임에서 취하므로 한 프레임의 길이 10 ms외에 5 ms의 알고리즘 지연이 발생한다. 이에 비해 상호부호화에서는 음성 버퍼 조절을 통해 이후 프레임에서의 샘플을 사용하지 않음으로써 표 3과 같이 지연 시간을 단축하였다.

표 3. 지연 시간 비교

Table 3. Comparison of delay-time.

	G.729A→EVRC	EVRC→G.729A
Tandem Coding	20ms + 15ms	20ms + 15ms
Transcoding	20ms + 5ms	20ms + 10ms

### 4.3 객관적 음질 평가

객관적 음질 평가를 위해 ITU P.862 PESQ (Perceptual Evaluation of Speech Quality)<sup>[9]</sup>를 사용하

표 4. 객관적 음질 평가 비교  
Table 4. Comparison of PESQ score.

	G.729A→EVRC		EVRC→G.729A	
	Male	Female	Male	Female
Tandem Coding	3.507	3.378	3.587	3.365
Transcoding	3.294	3.100	3.394	3.334

표 5. LSP변환만을 이용한 G.729A에서 EVRC로 상호 부호화  
Table 5. Transcoding (G.729A → EVRC) using LSP conversion.

	G.729A→EVRC	
	Male (PESQ / WMOPS)	Female (PESQ / WMOPS)
Tandem Coding	3.507 / 21.177	3.462 / 21.358
Transcoding	3.522 / 19.052	3.429 / 19.299

표 6. 주관적 음질 평가 비교  
Table 6. Comparison of subjective speech quality.

	G.729A→EVRC		EVRC→G.729A	
	Male	Female	Male	Female
Tandem Coding	40%	45%	35%	25%
Transcoding	30%	25%	30%	30%
No Preference	30%	40%	35%	45%

였다. 제안된 상호부호화 알고리즘과 이중 부/복호화를 거친 음성 샘플의 음질 평가 결과를 표 4에 정리하였다. 이중 부/복호화에 비해 상호부호화에서 약간의 음질 저하가 생기는 것을 알 수 있었다.

음질 저하의 원인은 개회로 피치 분석을 거쳐 개회로 피치 지연을 찾는 G.729A와 달리 한번의 피치지연만을 거치는 EVRC의 알고리즘의 영향이 크다. 또한 G.729A의 피치지연 범위가 EVRC 보다 크기 때문에 보정해주는 과정에서도 오차가 발생한다. G.729A에서 EVRC로의 상호부호화시 음질만을 고려할 경우 음성 버퍼 조절을 통한 지연시간 단축과 피치 지연과정을 생략하고 LSP만을 전달하는 것이 가장 뛰어난 음질을 얻을 수 있었으며 이때의 결과는 표 5와 같다.

4.4 주관적 음질 평가

상호부호화 알고리즘에 의해 생성된 음성신호의 주관적 음질 평가를 위해 성인 남녀 각각 20명에게 선호

도 평가를 실시하였다. 각각 상호부호화와 이중 부/복호화 샘플을 무작위적 순서로 들려주고 어느 음성 샘플의 품질이 더 뛰어난지 또는 두 음성 품질을 구분할 수 없는지 (No Preference)를 평가하였다. 표 6은 주관적 음질 평가의 결과를 정리한 것이다. G.729A에서 EVRC로의 상호부호화 경우 약간의 선호도 저하가 발생하였지만 전반적으로 이중 부/복호화와 거의 동등한 음질을 제공함을 알 수 있었다.

V. 결 론

본 논문에서는 기존의 이중 부/복호화 방법이 갖는 문제점을 개선하기 위하여 LSP 변환과 피치 지연 변환 알고리즘을 적용하여 상호부호화를 수행하였다. LSP 변환을 통해 LPC 분석 과정을 생략하여 실행 시간을 단축시킬 수 있도록 하였다. 또한 LPC 분석 과정이 생략되는 만큼 LPC 분석을 위한 음성 버퍼의 크기를 줄여 알고리즘 지연 시간을 줄일 수 있었다. 피치 지연 변환 알고리즘을 통해 한쪽 음성 부호화기에서 추정된 피치 주기를 다른 음성 부호화기에서 약간의 보정을 거쳐 이용하도록 함으로써 계산량을 줄이도록 하였다.

본 논문을 통해 구현된 G.729A와 EVRC간의 상호부호화기는 EVRC를 사용하는 디지털 셀룰라망과 G.729A를 사용하는 PDA, VoIP 간의 통신에서 적은 계산량과 짧은 지연으로 상호 부/복호화에 상응하는 음성 품질을 제공하여 실시간 통신에 보다 효과적으로 이용될 수 있다.

참 고 문 헌

- [1] ITU-T Rec. G.711, Pulse code modulation of voice frequencies, 1988.
- [2] ITU-T Rec. G.723.1, Dual-rate Speech Codec for Multimedia Communications Transmitting at 5.3 and 6.3 kbit/s, 1996.
- [3] TIA/EIA/IS-127 Enhanced variable rate codec, speech service option 3 for wideband spread spectrum digital systems, 1996.
- [4] ITU-T Rec. G.729 Annex A, "Reduced Complexity 8kbit/s CS-ACELP Speech Codec," 1996.
- [5] K. -T. Kim, S. -K. Jung, Y. -C. Park, Y. -S. Choi, and D. -H. Youn, "An efficient transcoding algorithm for G.723.1 and EVRC speech coders," IEEE 54th Vehicular Technology Conference, vol. 3, pp. 1561-1564, 2001.

- [6] 윤성완, 정성교, 박영철, 윤대희, "8kbps G.729A에서 5.3 kbps G.723.1로의 상호 부호화 알고리즘," 신호처리 합동 학술대회 논문집 제 13권 제 1호, pp. 832-826, 2000년10월
- [7] 광영진, 정지민, 권구락, 임정석, 황인호, 이경훈, 고성제 "G.729A에서 EVRC로의 상호부호화," 대한전자공학회 하계학술대회 논문집, 제 26 권, 제 1호, 2248-2251쪽, 2003년 7월
- [8] ITU-T Rec. G.729, "Coding of Speech at 8kbit/s CS-ACELP Speech Coder," 1996.
- [9] ITU-T Rec. P.862, Perceptual evaluation of speech quality (PESQ), an objective method for end-to-end speech quality assessment of narrowband telephone networks and speech codecs, 2000.

---

 저 자 소 개
 

---



권 구 락(정회원)  
 1997년 경일대학교  
 전자공학과 학사 졸업  
 1999년 성균관대학교 전기전자 및  
 컴퓨터공학부 석사 졸업  
 2002년~현재 고려대학교  
 전기전자전파공학부 박사  
 과정 재학

<주관심분야 : 영상 신호 처리, 멀티미디어 압축  
 및 콘텐츠 저작권 보호>



고 성 제(평생회원)  
 1980년 고려대학교  
 전자공학과 학사 졸업  
 1986년 State Univ. of New York  
 at Buffalo 전기 및 컴퓨  
 터공학과 석사 졸업  
 1988년 State Univ. of New York  
 at Buffalo 전기 및 컴퓨터  
 공학과 박사 졸업

1981년~1983년 대우통신 중앙연구소 연구원

1988년~1992년 Univ. of Michigan, Dearborn  
 전기 및 컴퓨터공학과 조교수

1996년 11월 IEEE APCCAS best paper award

1997년 12월 대한전자공학회 해동논문상 수상

1999년 11월 한국통신학회 LG 학술상

1997년 IEEE Senior member, 2000년 IEE Fellow

2004년 고려대학교 훌륭한 공대 교수상  
 (학술부문)

1992년~현재 고려대학교

전기전자전파공학부 정교수

<주관심분야 : 영상 처리 및 압축, 멀티미디어  
 통신>