

G.723.1 음성부호화기와 G.729A 음성부호화기의 상호 부호화 알고리즘

정희원 윤성완*, 정성교*, 박영철**, 최용수***, 윤대회*

An Efficient Transcoding Algorithm Between G.723.1 And G.729A Speech Coders

Sung-Wan Yoon*, Sung-Kyo Jung*, Young-Cheol Park**, Young-Soo Choi***, Dae-Hee Youn*

Regular Members

요 약

유/무선 통신 시스템에서는 통신망마다 서로 다른 음성 부호화기를 사용하므로 음성신호는 두 번의 부/복호화 과정을 거치게 되어 음질저하, 연산량 증가, 그리고 전달 지연 증가 등의 문제가 발생된다. 본 논문에서는 위의 문제점들을 개선하기 위하여 유/무선 음성 통신에 사용되는 음성 부호화기의 상호 부호화 알고리즘을 제안하고, 5.3 kbps G.723.1의 패킷과 8 kbit/s G.729A의 패킷을 서로 변환하는 방법을 제안한다. 여러 가지 음성 신호에 대해 객관적 음질 평가와 주관적 선호도 평가를 수행한 결과, 제안된 상호 부호화 알고리즘이 이중 부/복호화보다 짧은 전달 지연 시간과 26-37% 적은 연산량으로 동등한 음질의 음성신호를 복호화함을 확인하였다.

ABSTRACT

Digital communication systems tend to support multiple speech coders. To accomplish communication between endpoints employing different speech coders, decoder and encoder of each endpoint coder should be placed in tandem. Tandem coding produces problems such as poor speech quality, high computational load, and additional transmission delay. In this paper, we propose an efficient transcoding algorithm for a legitimate communication between 5.3 kbps G.723.1 and 8 kbps G.729A coders. The efficient transcoding algorithm reduces about 26-37% of the complexity in the encoding part. Subjective preference tests as well as objective quality evaluation confirmed that the proposed transcoding algorithm could produce equivalent speech quality to the tandem coding with the shorter processing delay and less computational complexity.

I. 서 론

최근 들어 무선통신망과 유선네트워크망 사이의 음성 데이터를 저장, 전송하려는 시도가 많이 이루어지고 있다. 이때 각각의 통신망에서 서로 다른 음성부호화기를 사용할 경우, 음성부호화기는 두 번의 부/복호화 과정을 거치게 되는데, 이를 tandem

이라고 하며, 음질저하, 연산량 증가, 지연시간 증가 등의 문제점들을 갖게 된다.

이와 같은 문제점들의 근본적 원인은 음성신호가 음성부호화기를 두번 통과하기 때문이다. 이중 부/복호화로 인해 발생하는 문제점들을 해결하기 위해서는 한쪽 통신망의 압축된 비트열을 PCM 신호로 복호화한 후, 다른 쪽 통신망에서 다시 한번 부/복

* 연세대학교 전기전자공학과 (yocello@assp.yonsei.ac.kr),

** 연세대학교 신호처리연구센터,

*** (주)LG전자

논문번호 : K01021-0110, 접수일자 : 2000년 1월 10일

※ 본 연구는 (주)LG전자가 지원한 유/무선 통신용 음성 부호화기의 상호부호화 알고리즘 연구의 일부로서 수행되었습니다.

호화 과정을 거치는 것이 아니라, 전송되어온 비트열을 그대로 한번의 복호화 과정을 거치게 하는 것이다. 이러한 방법을 상호부호화(transcoding)라고 한다.

본 논문에서는 유/무선 통신망에서 많이 사용되고 있는 5.3 kbit/s G.723.1^[1]과 8 kbit/s G.729A^[2]사이의 상호 부호화 알고리즘을 제안하였다. 제안된 상호 부호화 알고리즘은 LSP 변환과정, 개회로 피치 변환과정, 그리고 고속 적응/고정 코드북 검색과정으로 구성된다. 두 부호화기의 프레임 길이를 고려하여 30ms 프레임을 갖는 G.723.1의 프레임 한 개와 10ms 프레임을 갖는 G.729A의 세 개의 프레임이 서로 변환된다.

여러 가지 음성 신호로 모의 실험한 결과, 본 논문에서 제안한 상호 부호화 알고리즘이 두 번의 부/복호화 과정을 거친 경우보다 짧은 전달 지연 시간과 적은 연산량으로 동등한 음질의 음성 신호를 복호화하는 것을 확인하였다. 본 논문의 구성은 다음과 같다. 2장에서는 G.723.1 음성부호화기와 G.729A 음성부호화기의 알고리즘을 설명하고, 3장에서는 G.723.1과 G.729A로의 상호 부호화 과정을 설명한다. 4장에서는 주관관적 평가를 통하여 제안된 알고리즘과 기존의 이중 부/복호화의 성능을 비교하고, 끝으로 5장에서 결론을 맺는다.

II. G.723.1과 G.729A의 알고리즘 비교

본 논문에서 사용한 G.723.1과 G.729A의 부호화 과정 알고리즘이 모듈별로 표 1에 비교 정리되어 있다.

표 1. G.723.1과 G.729A의 알고리즘 비교

	ITU-T G.723.1	ITU-T G.729A
Frame	30ms/frame, 4 subframe	10ms/frame, 2 subframe
LPC&LSP	10th LPC&LSP	10th LPC&LSP
적응코드북	5-tap predictor, integer pitch	1-tap predictor, 1/3 fractional pitch
고정코드북	ACELP for 5.3 kbps	ACELP

III. G.723.1 음성부호화기와 G.729A 음성 부호화기의 상호 부호화 알고리즘

1. G.723.1에서 G.729A로의 상호 부호화 제안된 G.723.1(5.3 kbps)에서 G.729A로의 상호 부호화 알고리즘은 크게 LSP 변환과정, 개회로 피치 변환과정으로 구성된다^[3]. 그리고 두 부호화기의 프레임 길이를 고려하여 30ms 프레임길이를 갖는 G.723.1의 1개의 프레임이 10ms 프레임길이를 갖는 G.729A의 프레임 3개로 변환된다. G.723.1에서 G.729A의 상호 부호화 과정은 그림 1과 같다.

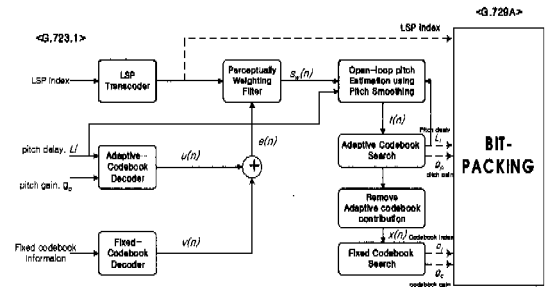


그림 1. G.723.1에서 G.729A로의 상호 부호화

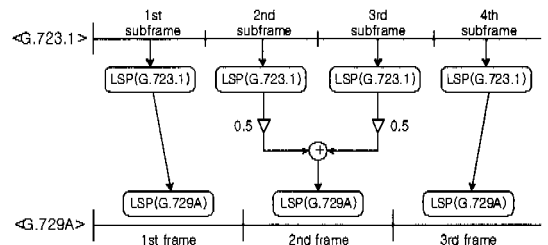


그림 2. 선형 보간을 이용한 LSP 변환 과정

가. 선형 보간을 이용한 LSP 변환 과정

제안된 LSP 변환과정에선 G.723.1 프레임의 LSP 정보를 그림 2와 같이 선형보간법을 통해 G.729A의 3개의 프레임에 해당하는 LSP값을 구한다. G.723.1의 첫 번째와 네 번째 서브프레임의 LSP는 각각 G.729A의 첫 번째 프레임과 세 번째 프레임의 LSP 정보로 쓰인다. 그리고 G.723.1의 두 번째, 세 번째 서브프레임의 LSP는 선형 보간을 이용해 G.729A의 두 번째 프레임 LSP 정보로 변환된다.

이로 인해 LPC 계수를 구하기 위한 과정이 필요 없게 되어 이중 부/복호화를 하는 경우에 비해 연산량이 크게 줄어든다. 그리고 [4]와 같이 지연시간을 계산한 결과, G.729A의 LPC 계산에 필요한 5ms의 예견구간으로 인한 지연이 생기지 않기 때문에 이중 부/복호화에서 일어나는 지연시간 증가의 문제를

해결하였다. 또한 이중 부/복호화는 과정에서 서로 다른 양자화 방법으로 인해 LSP가 크게 왜곡되는데 비해, 제안된 방법에선 선형보간법을 이용해 변환된 LSP값과 G.729A의 부호화기에서 계산된 LSP값이 거의 일치했다. 그림 3에는 유성을 부분의 LPC 스펙트럼이 나타나 있다. 그림에서 보는 바와 같이 본 논문에서 제안한 LSP 변환 과정의 스펙트럼이 G.723.1과 G.729A로 이중 부/복호화한 경우보다 음질에 많은 영향을 미치는 저주파 부근과 포먼트 부근에서 원래의 G.729A의 스펙트럼과 더 유사함을 알 수 있다.

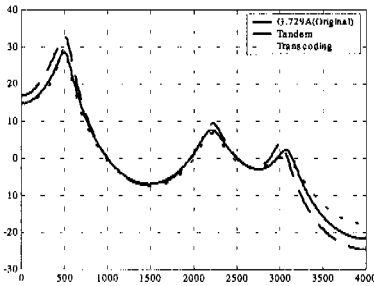


그림 3. 이중 부/복호화와 상호부호화의 LPC 스펙트럼 비교

나. 개회로 피치 변환

상호 부호화하려는 해당 프레임에서 G.723.1의 적응 코드북 피치정보를 이용해 지각 가중된 신호 도메인에서 개회로 피치를 추정한다. 즉, 현재 프레임에 대응되는 G.723.1의 서브프레임의 적응 코드북 피치값과 G.729A의 이전 서브프레임의 적응 코드북 피치가 서로 10 샘플 미만으로 차이가 날 때에는 G.723.1의 적응 코드북 피치값을 G.729A의 개회로 피치로 결정한다.

만약, 현재 프레임에 대응되는 G.723.1의 서브프레임의 적응 코드북 피치값과 G.729A의 이전 프레임의 2번째 서브프레임의 적응 코드북 피치가 서로 10 샘플 미만으로 차이가 날 때에는 그림 4와 같이 피치 완만화 방법을 적용한다. 이때에는 개회로 피치 추정의 후보가 되는 G.723.1의 적응 코드북 피치값과 G.729A의 이전 서브프레임의 적응 코드북 피치값을 중심으로 ±3샘플 주변에서 (1)의 $R(k_i)$ 를 최대로 만드는 구간 최대 지연값(local maximum delay)을 결정한다.

$$R(k_i) = \sum_{n=0}^{79} s_w(n) \cdot s_w(n-k_i) \cdot \begin{cases} P_A - 3 \leq k_1 \leq P_A + 3 \\ P_B - 3 \leq k_2 \leq P_B + 3 \end{cases} \quad (1)$$

여기서 $S_w(n)$ 은 지각 가중된 음성신호, P_A 와 P_B 는 각각 G.723.1과 G.729A의 적응 코드북 피치값이며, k_1 과 k_2 는 해당 범위의 개회로 피치 지연값이다.

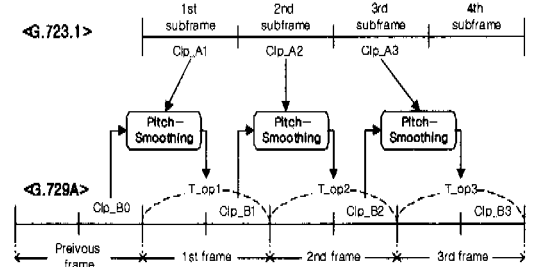


그림 4. 피치 완만화를 이용한 개회로 피치 추정

각 범위에서 (1)의 $R(k_i)$ 를 최대로 만드는 지역 최대 지연값(local maximum delay)을 결정한 후, 이를 아래와 같이 지역 최대 지연값에서의 에너지로 정규화한다.

$$R'(t_i) = \frac{R(t_i)}{\sqrt{\sum_n s_w^2(n-t_i)}}, \quad i=1,2 \quad (2)$$

여기서 t_i 는 각 범위에서의 지역 최대 지연값이다. 즉, t_1 과 t_2 는 각각 G.723.1과 G.729A의 적응 코드북 범위에서의 지역 최대 지연값이다.

정규화된 각 범위에서의 local maximum값은 G.729A의 적응 코드북 피치값에 가중치를 주어 비교하고, 개회로 피치값을 결정하게 된다. 즉, G.729A의 local maximum값이 G.723.1의 local maximum값보다 3/4배 이상 크게 되면 G.729A의 local maximum값을 개회로 피치로 결정하고, 그 반대의 경우 G.723.1의 local maximum값을 개회로 피치로 정한다.

피치 완만화를 이용한 개회로 피치 추정 방법도 좁은 범위 내에서 자기상관 계수를 비교하기 때문에 이중 부/복호화의 경우보다 훨씬 적은 연산량으로 개회로 피치를 결정할 수 있게 된다.

2. G.729A에서 G.723.1로의 상호 부호화

제안된 G.729A에서 G.723.1(5.3 kbps)로의 상호 부호화 알고리즘은 크게 LSP 변환과정, 개회로 피치 변환, 고속 적응 코드북 검색, 고속 고정 코드북 검색과정으로 구성된다⁵⁾. 그리고 10ms 프레임길이 를 갖는 G.729A의 세 프레임이 30ms 프레임길이

를 갖는 G.723.1의 한 프레임으로 변환되며 그 과정은 G.723.1의 비트열이 G.729A의 비트열로 변환되는 과정과 유사하다.

G.729A에서 G.723.1로의 상호부호화에선 LSP와 개회로 피치 변환과정은 반대방향의 상호부호화 과정과 유사하므로, 이 절에선 적용 코드북 고속검색에 대해 자세히 기술하도록 하겠다.

가. 고속 적응 코드북 검색

본 논문에서 사용한 고속 적응 코드북 검색은 [6]에서 사용한 방법에 기반하고 있다. 5차 피치 예측기의 이득과 지연을 동시에 검색하지 않고 지연값, L_{cp} , 을 먼저 결정한 다음 이득값을 검색하는 순서로 적응 코드북 검색이 이루어진다.

지연 L_{cp} 가 결정되면 5차 피치 예측기의 이득값을 검색하게 되는데 기존의 G.723.1(5.3 kbps)의 적응 코드북 이득 검색 과정에서는 170개의 인덱스를 가진 이득값표 전체를 검색한다. 본 논문에서 제안한 적응 코드북 변환 과정에선 현재 서브프레임과 가장 인접한 G.729A의 서브프레임의 적응 코드북 이득값에 g_b 에 따라 5차 피치 예측기의 이득값 검색 범위를 제한한다. 즉, G.729A의 적응 코드북 이득값 의 값의 분포에 따라 170개의 후보값 중 미리 선택된 85개의 값만 검색한다.

나. 고속 고정 코드북 검색

5.3 kbps G.723.1의 고정 코드북 탐색은 ACELP 구조를 기반으로 이루어지며, 각 부프레임을 4개의 트랙으로 나누고 각 트랙마다 1개의 펄스를 검색하는 nested-loop search를 사용한다. 이로 인해 펄스 위치를 검색할 때 $8 \times 8 \times 8 \times 8 = 4096$ 의 경우의 수에 대해 분석-합성 방법으로 위치를 검색하며, 실제적인 구현에서는 마지막 펄스의 검색을 제한함으로써 연산량을 줄인다. 본 논문에서는 G.729A에서 사용한 depth-first tree search에 기반한 고속 고정 코드북 탐색방법을 적용하였다. Depth-first tree search 과정에선 검색하고자 하는 4개의 펄스를 처음 2개와 나머지 2개를 쌍으로 나누어 각각 동시에 검색하게 되므로 펄스 위치의 경우의 수는 $2 \times \{(8 \times 8) + (8 \times 8)\} = 256$ 으로 줄어 들게 된다.

IV. 상호 부호화 알고리즘의 성능 평가

1. 객관적 음질평가

객관적 음질평가의 척도로 사용한 LPC-CD는 비

교하고자 하는 두 음성신호의 cepstral 계수 (Cepstral Coefficient)의 왜곡을 나타낸 척도이며, PSQM은 300-3400Hz의 전화 대역폭상에서 음성부호화기의 객관적 음질 평가방법으로 주관적 평가와 높은 상관도를 갖고 있는 ITU-T의 권고안이다^[7].

기존의 이중 부/복호화 방법과 본 논문에서 제안한 상호부호화 알고리즘을 사용해 음성신호를 부/복호화 했을 때의 LPC-CD와 PSQM 결과를 표 2에 정리하였다.

표 2. 객관적 음질 평가결과

	LPC-CD(dB)		PSQM	
	Female	Male	Female	Male
Tandem(AtoB)	3.98	3.90	2.45	2.44
Transcoding(AtoB)	3.66	3.54	2.22	2.17
Tandem(BtoA)	4.17	3.65	2.47	2.43
Transcoding(BtoA)	3.86	3.22	2.46	2.27

*AtoB : from G.723.1 to G.729A

*BtoA : from G.729A to G.723.1

표 2에서 보는 바와 같이 여성화자와 남성화자 모두 LPC-CD와 PSQM 값이 이중 부/복호화의 경우보다 본 논문에서 제안한 상호 부호화 방법이 더 작은 수치를 나타냈다. 이는 그림 3의 LPC 스펙트럼 비교와 일치하는 결과로 볼 수 있다. 실험 결과, G.723.1에서 G.729A로의 연결의 경우 이중 부/복호화와 상호부호화 각각의 LPC-CD 차이는 여성화자의 경우 0.32 dB, 남성화자의 경우 0.36 dB로 나타났다. G.729A에서 G.723.1로의 연결의 경우 LPC-CD의 차이는 여성의 경우 0.31 dB, 남성의 경우 0.40 dB로 나타났다. PSQM 결과 역시 양방향 모두 상호 부호화 방법을 적용한 경우가 이중 부/복호화 방법을 적용한 경우보다 0.01에서 0.27까지 낮은 수치를 나타냈다. 객관적 평가 결과 이중 부/복호화보다 본 논문에서 제안한 상호 부호화 방법이 더 좋은 성능을 보였다. 이는 최종 복호화단의 합성필터를 구성하는 LPC 계수를 계산하는 과정에서 이중 부/복호화 방법보다 상호 부호화 방법으로 구한 LPC 계수가 단일 부/복호화의 LPC 계수에 더 유사하기 때문에 오차로 인한 음질의 저하를 최소화할 수 있고, 객관적 평가의 척도의 수치도 이중 부/복호화보다 좋은 성능을 나타낸 것으로 생각할 수 있다.

2. 주관적 음질평가

주관적 음질평가로는 30명의 청취자에게 한국인 남녀 화자 문장을 들려준 후, 우열을 가리는 선호도 조사를 수행하였다. 표 3은 이중 부/복호화와 상호 부호화 방법의 선호도 평가 결과를 보여준다.

표 3. 주관적 선호도 평가결과

	G.723.1→G.729A		G.792A→G.723.1	
	Female	Male	Female	Male
Tandem	30%	20%	26.7%	30%
Transcoding	36.7%	33.3%	13.3%	40%
No Preference	33.3%	46.7%	60%	30%

G.723.1에서 G.729A로의 연결에선 여성화자의 경우, 이중 부/복호화와 상호부호화를 선호하는 비율이 각각 30%와 36.7% 였고, 음질의 우열을 가리지 못하는 비율은 33.3%였다. 남성 화자의 경우엔, 이중 부/복호화를 선호하는 비율은 20%에 그치는데 반해, 상호 부호화를 선호하는 비율은 33.3%로 나타났다. G.729A에서 G.723.1로의 연결에선 여성 화자의 경우, 이중 부/복호화를 선호하는 비율은 26.7%로 나타났고 나머지 73.3%는 상호부호화를 선호하거나 음질의 우열을 가리지 못했다. 남성의 경우엔 30%는 이중 부/복호화를 선호하였고, 나머지 70%는 상호부호화를 선호하거나 음질의 우열을 가리지 못하였다. 주관적 평가의 결과로 볼 때, G.729A에서 G.723.1로의 연결에선 상호부호화 알고리즘이 이중 부/복호화 방법에 비해 음질 저하가 거의 일어나지 않는다고 할 수 있다.

3. 연산량 측정

이중 부/복호화 방법과 본 논문에서 제안한 상호 부호화 알고리즘의 연산량 비교를 위하여 TI사의 TMS320C62x DSP chip을 사용하여 simulator에서 소요되는 clock을 측정하였다. G.723.1에서 G.729A로의 연결에선 LPC 분석과 개회로 피치 추정 과정에 상호부호화 알고리즘이 적용되었고, 그 반대 방향의 경우는 LPC 분석, 개회로 피치 추정, 적응 코드북 검색 과정, 그리고 고정 코드북 검색 과정에 적용되었다. 표 4에는 각 모듈별로 기존의 방법과 제안한 방법이 음성신호를 부호화하는데 걸리는 clock을 MIPS로 환산하여 정리되어있다. 본 논문의

DSP 구현과정은 단순히 연산량의 비교를 위한 것이었기 때문에 일반적으로 DSP 구현의 연산량 감소를 위한 최적화 과정은 생략되었다. 따라서 표 4에 정리되어 있는 연산량 수치들은 DSP 구현의 관점에서 볼 때 최적화된 수치가 아님을 미리 명시해 둔다. 그렇지만 이중 부/복호화와 상호 부호화의 연산량은 전체 부호화 과정에서 상호 부호화 알고리즘이 적용된 부분에서만 차이가 나고, 나머지 부분은 동일하기 때문에 표 4의 수치들은 연산량 비교의 측면에서 객관적인 결과라 할 수 있다.

표 4. Tandem과 Transcoding의 연산량 비교

연산량 (MIPS)	G.723.1→G.729A		G.792A→G.723.1	
	Tandem	Transcoding	Tandem	Transcoding
LPC&LSP	6.41	2.36	6.93	5.55
Olp	0.94	0.21	1.54	1.19
Acbk	2.45	2.45	10.14	6.34
Fcbk	4.30	4.30	10.50	2.16
Others	4.04	4.04	8.05	8.05
Total	18.15	13.37	37.17	23.28

G.723.1에서 G.729A로의 연결에선 이중 부/복호화 방식과 상호 부호화 방법을 사용하여 G.723.1의 비트열에서 G.729A의 압축된 비트열을 생성하는 데에는 18.15 MIPS, 13.37 MIPS가 소요됐다. 따라서 본 논문에서 제안한 상호 부호화 방법이 기존의 방법보다 약 26% 적은 연산량으로 동등한 음질을 얻어냄을 확인하였다. G.723.1에서 G.729A로의 연결에선 이중 부/복호화 방식과 상호 부호화 방법을 사용하여 G.729A의 비트열에서 G.723.1의 압축된 비트열을 생성하는 데에는 각각 37.17 MIPS, 23.28 MIPS가 소요됐다. 전체 부호화 과정에 걸리는 시간을 보면 상호 부호화 방법이 기존의 방법보다 약 37% 적은 연산량으로 동등한 음질을 얻어냄을 확인하였다.

V. 결론

본 논문에서는 유/무선 통신망에서의 음성신호의 이중 부/복호화로 야기되는 연산량 증가, 전달시간 지연 등의 문제를 최소화할 수 있는 상호 부호화 알고리즘을 제안하였다. 제안된 알고리즘은 8 kbps G.729A로 부호화된 패킷과 5.3 kbps G.723.1의 음

성패킷을 상호 부호화함으로써 이중 부/복호화를 할 경우보다 적은 연산량과 짧은 전달 시간 지연으로 동등한 음질을 얻을 수 있었다.

상호 부호화 알고리즘의 성능 평가 결과, LPC-CD, PSQM 등의 객관적 음질평가에선 양방향 모두 이중 부/복호화보다 성능이 개선되었음을 확인하였다. 주관적 선호도 음질평가에선 이중 부/복호화와 비교해 음질의 저하가 없음을 확인하였다. 연산량은 이중 부/복호화 방법에 비해 26-37% 감소하였다.

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] ITU-T Rec. G.729 Annex A "Reduced Complexity 8 kbit/s CS-ACELP Speech Codec," 1996.

[3] 윤성완, 정성교, 박영철, 윤대회, "5.3 kbps G.723.1에서 8 kbps G.729A로의 상호 부호화 알고리즘," 한국통신학회 추계종합학술발표회 논문집 하권, pp 1331-1334, 2000.

[4] H.G Kang, H.K Kim, and Cox. R.V, "Improving transcoding capability of speech coders in clean and frame erased channel environment," Proc. IEEE Speech Coding Workshop, pp 78-80, 2000.

[5] 윤성완, 정성교, 박영철, 윤대회, "8 kbps G.729A에서 5.3 kbps G.723.1로의 상호 부호화 알고리즘," 신호처리 합동 학술대회 논문집 제13권 제1호, pp 823-826, 2000.

[6] 정성교, 박영철, 윤성완, 차일환, 윤대회, "효율적인 다채널 구현을 위한 ITU-T G.723.1 음성부호화기 고속 알고리즘 제안", 한국음향학회 학술발표대회 논문집 제19권 제1호, pp 67-70, 2000.

[7] ITU-T Rec. P.861 "Objective Quality Measurement Of Telephoneband(300-3400Hz) Speech Codecs," 1996.

윤 성 완(Sung-Wan Yoon)

1999년 2월 : 연세대학교 전자공학과 졸업

2001년 2월 : 연세대학교 전기전자공학과 석사

2001년 3월~현재 : 연세대학교 전기전자공학과 박사과정

<주관심 분야> 음성신호처리, 실시간 구현, VoIP

정 성 교(Sung-Kyo Jung)

한국통신학회 논문지 제26권 1B호 참고

박 영 철(Young-Cheol Park)

한국통신학회 논문지 제24권 9B호

최 용 수(Yong-Soo Choi)

한국통신학회 논문지 제26권 1B호 참고

윤 대 회(Dae-hee Youn)

한국통신학회 논문지 제26권 제 1B호 참조