

대역폭 계층 구조의 광대역 음성 부호화기 개발

Bandwidth Scalable Wideband Speech Codec

이 우 석*, 박 호 중*, 손 창 용**, 이 영 범**

(Woo-Suk Lee*, Hochong Park*, Chang-Yong Son**, Yong-Beom Lee**)

*광운대학교 전자공학과, **삼성종합기술원 HCI 연구실

(접수일자: 2004년 4월 23일; 채택일자: 2004년 7월 21일)

본 논문에서는 대역폭 계층 구조의 광대역 음성 부호화기 구조와 이를 위한 상위 대역 부호화기를 제안한다. 상위 대역 부호화기는 상위 대역 신호를 다수의 주파수 밴드로 분할하고, 각 밴드 신호를 DCT 변환하여 양자화 한다. DCT 계수는 크기와 부호로 분리하여 독립적으로 처리하며, 각각의 특성에 적합한 양자화 방법을 개발하여 성능을 향상시킨다. 또한, 하위 대역의 부호화된 이득 파라미터를 상위 대역의 양자화에 이용함으로써 부호화 성능을 향상시키는 방법을 제안한다. G.729E로 하위 대역을 부호화하고 제안한 부호화기로 상위 대역을 처리하는 대역폭 계층 구조의 광대역 음성 부호화기를 개발하였고, 주관적 음질을 측정하여 24kbps G.722.1보다 우수한 음질을 가지는 것을 확인하였다.

핵심용어: 광대역 음성 부호화기, 계층 구조의 부호화기, 대역폭 계층 구조, DCT

투고분야: 음성처리 분야 (2.2)

In this paper, the structure of bandwidth scalable wideband speech codec and its high-band codec are proposed. In the high-band codec, the signal is divided into frequency bands, and each band is quantized in DCT domain. The DCT coefficients are splitted into magnitude and sign, and each is quantized independently by a specialized method based on its characteristics. In addition, the quantized gain parameter in the low-band codec is utilized in the high-band codec for an enhanced performance. The bandwidth scalable wideband speech codec using G.729E for low-band and the proposed codec for high-band is developed, and it is confirmed that the proposed codec has better subjective performance than 24kbps G.722.1.

Keywords: Wideband speech codec, Scalable codec, Bandwidth scalability, DCT

ASK subject classification: Speech signal processing (2.2)

I. 서론

최근에 음성 통신의 응용 분야가 다양해짐에 따라 고품질의 음성 통신에 대한 필요성이 크게 부각되고 있다. 이에 따라 16kHz로 샘플링 되고 7kHz 대역폭을 가지는 광대역 음성 신호의 전달이 필요하고 이를 위한 새로운 광대역 음성 부호화기의 개발이 요구된다[1]. 또한, 디지털 통신은 아날로그 통신에서 주로 사용하던 회선 스위칭 방식이 아니라 데이터와의 통합을 위하여 패킷 (packet) 스위칭 방식을 사용하며, 이 방식에서의 채널 용량은 시간에 따라 가변적이므로 고정된 크기의 음성 패킷을 계

속 전달하는 것은 매우 비효율적인 통신이 되고, 패킷 손실이 발생하여 음질이 저하되는 문제점을 가질 수 있다. 이를 해결하기 위하여 패킷을 여러 개의 계층으로 구성하여 채널의 전송 용량에 따라 전송 계층의 양을 조절하여 주어진 채널 환경에서 최적의 통신이 가능하도록 하며, 이 기능을 제공하는 계층 구조의 부호화기 (scalable codec) 개발이 필요하다. 계층 구조의 음성 부호화기에서 전달되는 계층의 양이 감소하면 약간의 음질 저하는 발생하지만, 주어진 채널 전송률에 최적화하여 데이터를 전송하므로 계층 구조를 갖지 않는 부호화기에서 발생하는 완전한 패킷 손실을 방지할 수 있어 통신의 품질을 향상시킬 수 있다.

계층 구조의 음성 부호화기에서 가장 일반적인 구조는 대역폭에 따라 계층을 할당하는 것으로서, 입력 음성 신

책임저자: 박 호 중 (hcpark@mail.kw.ac.kr)
139-701 서울시 노원구 월계동 447-1
광운대학교 전자공학과
(전화: 02-940-5104; 팩스: 02-913-9057)

호를 대역 별로 부호화하여 계층적인 패킷을 만들고, 채널의 가변적인 전송률에 따라 패킷을 선택적으로 전달하는 구조를 가진다. 예로 7khz 대역폭 통신을 기본으로 하고 채널 전송률이 감소하면 협대역에 해당하는 패킷만을 전송하여 협대역 통신을 할 수 있도록 한다.

대역폭 계층 구조의 광대역 음성 부호화기에 대한 많은 연구 결과가 있다[2, 3, 4]. 일반적으로 하위 대역과 상위 대역을 독립적으로 처리하여, 협대역 통신에서의 호환성을 유지하기 위하여 하위 대역은 기존의 표준 협대역 음성 부호화기를 사용한다. 상위 대역 신호는 보통 Transform 부호화 구조와 파라미터 기반의 방법을 사용하며, 상위 대역 신호의 특성을 고려하여 신호를 부호화한다. 본 논문에서는 새로운 대역폭 계층 구조의 음성 부호화기의 전체 구조를 결정하고, 상위 대역 신호의 부호화를 위한 새로운 구조와 양자화 방법을 연구하고 각 블록에 대한 설계를 하여 최종 대역폭 계층 구조의 광대역 음성 부호화기를 개발한다. 이때, 하위 대역은 호환성을 위하여 G.729와 G.723.1 부호화기를 기준으로 설계한다. 또한, 개발한 대역폭 계층 구조의 광대역 음성 부호화기의 성능을 주관적 평가를 통하여 기존의 광대역 음성 부호화기와 비교한다.

II. 제안하는 음성 부호화기 구조

본 논문에서 제안하는 대역폭 계층 구조의 광대역 음

표 1. 제안하는 음성 압축기의 규격

Table 1. Specification of the proposed speech codec.

Sampling Rate	16 khz
Bit Rate(High-Band Only)	8 kbps
Frame Size	30 msec
Low Pass Filtering × 2	3 msec
Look-Ahead	5 msec

성 부호화기의 전체 구조는 그림 1과 같다. 광대역 입력 신호는 저역 통과 필터링과 Down Sampling을 통하여 협대역 신호로 변환되고, 이를 G.729 또는 G.723.1 표준 협대역 음성 부호화기로 부호화한다. 상위 대역 부호화기는 광대역 입력 신호로부터 상위 대역에 해당하는 정보를 추출하여 제안하는 음성 부호화 방법에 의하여 부호화한다. 이 때, 상위 대역 부호화기는 하위 대역의 부호화된 정보를 활용하여 부호화 성능을 높이도록 한다.

제안하는 음성 부호화기의 규격은 표 1과 같다. 하위 대역 부호화기로 G.729 뿐만 아니라 G.723.1을 사용하기 위하여 프레임 크기를 30msec로 하였고, 만일 G.729가 사용되면 3개의 10msec 프레임을 한 번에 모아 처리하도록 한다. Look-Ahead는 G.729의 5msec를 초과하지 않게 설계 하였다.

III. 제안하는 상위 대역 부호화기

3.1. 전체 구조

제안하는 부호화기에서 하위 대역 신호는 표준 협대역 음성 부호화기로 처리하므로 하위 대역 신호의 부호화

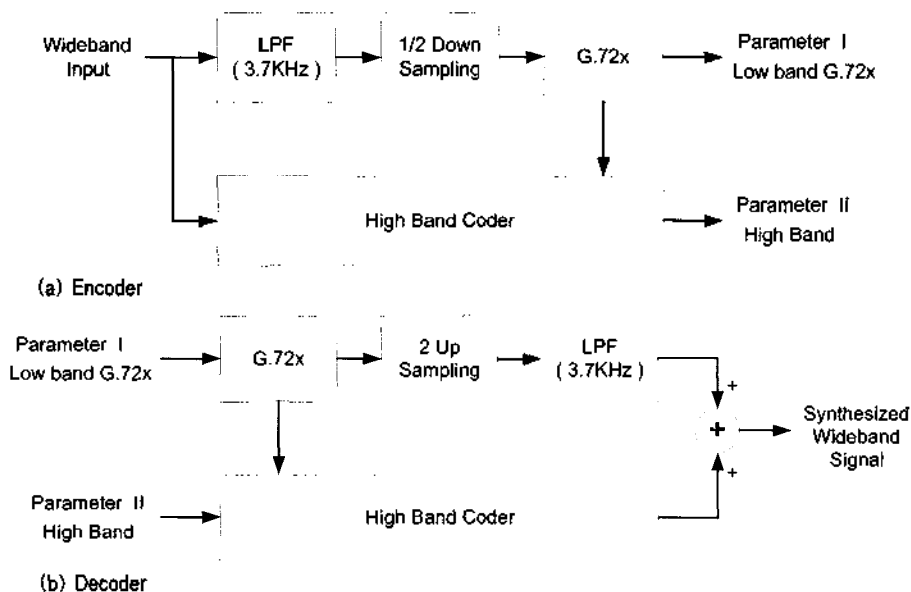


그림 1. 제안하는 대역폭 계층 구조의 광대역 음성 부호화기 구조

Fig 1. Structure of the proposed bandwidth scalable wideband speech codec

과정에서 새롭게 추가되는 부분은 없으며, 단지 하위 대역의 양자화 된 정보를 상위 대역에서 활용하는 구조를 가진다. 그러나 상위 대역에서는 표준 부호화 방법이 없으므로 본 논문은 상위 대역 신호의 부호화를 위한 새로운 방법의 개발에 초점을 맞춘다. 이 때, 상위 대역 패킷은 항상 전달되는 것이 아니라 채널 상황에 따라 전달 여부가 결정되므로 상위 대역 부호화기는 반드시 프레임 단위의 독립적인 동작만으로 구성되고 과거의 정보를 활용하는 구조는 피하도록 설계한다. 즉 프레임 사이의 예측 동작이 제한되고 프레임 내에서 모든 연산이 이루어져야 한다.

상위 대역 신호는 그림 2와 같이 청각 필터를 이용하여 주파수 대역으로 분할하고 주파수 영역에서 양자화한다. 먼저, 광대역 입력 신호에 청각 필터를 적용하여 임계 대역을 고려한 4개의 대역 통과 신호로 분해한다. 청각 필터로 일반적인 Gammatone 필터를 이용하고, 필터의 충격 응답은 다음의 식으로 표현된다[5, 6].

$$g(t) = at^{N-1} e^{-2\pi b(f_c t)} \cos(2\pi f_c t + \phi)$$

여기서 f_c 는 중심 주파수, $B(f_c)$ 는 대역폭, a , b , ϕ 는 필터의 세부 특성을 결정하는 파라미터이다. 본 논문에서 사용하는 중심 주파수와 대역폭은 표 2에 정리되어 있다. 청각 필터의 구조는 필터링에 의한 시간 지연을 최소화하고 합성 단계에서 프레임 사이의 순환 구조를 막기 위하여, 분해 필터는 IIR 구조를 사용하고 합성 필터는 FIR 구조를 이용한다. 필터의 차수는 분해 및 합성 필터링 과정으로 인한 시간 지연이 Look-Ahead인 5msec 이내가 되도록 정하였다.

다음, 각각의 대역 통과 된 신호를 DCT (discrete cosine transform) 하여 주파수 영역으로 변환 한다. 즉, 30msec 크기의 프레임에 대하여 각 대역별 Gammatone 출력 신호는 480 샘플이며, 이 신호를 각

표 2. 각 대역의 DCT 계수 인덱스 : DCT 크기 = 480
Table 2. DCT index of each frequency band : DCT size=480

	Center Freq. (Hz)	Band-width (Hz)	DCT Coefficient		
			First Index	Last Index	# of Coeff.
Band 0	4000	457	220	263	44
Band 1	4800	543	264	317	54
Band 2	5800	651	318	383	66
Band 3	7000	780	384	425	42
Sum					206

각 480-point DCT 하여 주파수 영역으로 변환한다. 이 때, DCT 입력 신호가 대역 통과 신호이므로 해당 대역의 DCT 계수만 양자화 하여 전달하도록 하며, 각 대역별 해당 DCT 계수 인덱스는 표 2에 정리되어 있다. 따라서 총 206개의 DCT 계수에 대한 양자화기를 설계하면 상위 대역 부호화가 완성된다. 각 대역 통과 신호의 DCT 과정에서 윈도우를 적용하지 않고 직접 DCT 하며, 이로 인하여 발생하는 경계 효과에 의한 왜곡은 합성 단계에서 청각 필터를 거치면서 시간적인 중첩 현상이 발생하므로 사라지게 된다.

3.2. DCT 계수의 크기와 부호 분리

앞에서 구한 DCT 계수를 효율적으로 양자화 하기 위하여 DCT 계수 사이의 상관관계를 이용하여야 한다. 표 3은 인접한 DCT 계수 사이의 상관관계와 인접한 DCT 계수 크기 (magnitude) 사이의 상관관계를 구한 것이다. 입력으로 NTT 한국어 DB에서 목음을 제외한 10분 분량의 음성을 만들어 사용하였다. 표 3에서 볼 수 있듯이, DCT 계수들 사이의 상관관계는 거의 없지만, 그 계수의 크기만을 취하면 비교적 높은 상관관계가 존재하게 된다. 따라서 DCT 계수를 크기와 부호 성분으로 분리하여 양자화 함으로써 그 성능을 높이는 방법을 제안하며, 그림 3은 각 대역의 DCT 계수를 크기와 부호로 분리하여 양자화 하는 전체적인 구조를 보여준다.

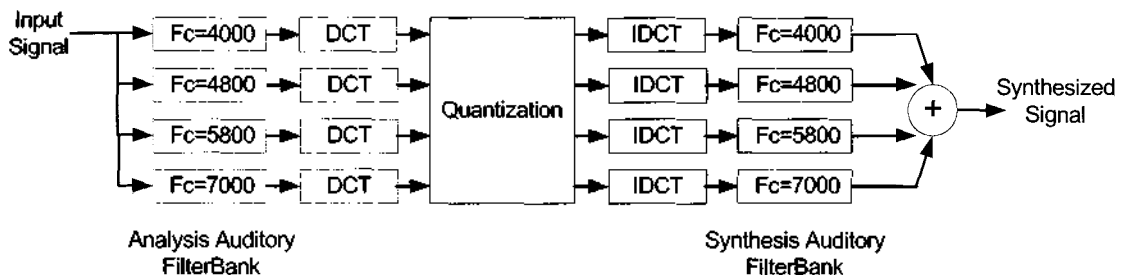


그림 2. 제안하는 상위 대역 부호화기 구조
Fig 2. Structure of the proposed high-band codec

표 3. DCT 계수 및 그 크기 사이의 상관관계
Table 3. Correlation coefficient of DCT coefficients and DCT coefficient magnitudes

	Correlation between index k and k+1	
	DCT coefficient	DCT coeff. magnitude
Band 0	0.0047	0.5901
Band 1	0.0018	0.6162
Band 2	0.0056	0.6492
Band 3	0.0001	0.6469

3.3. DCT 계수 크기의 양자화

DCT 계수 크기는 높은 상관관계를 가지므로, 이를 하나의 신호로 간주하고 다시 한번 DCT 하여 Energy Compaction 효과를 얻도록 한다. 이와 같이 두 번째 단계에서 구한 DCT 계수를 "두 번째 DCT 계수"라 하고, 대역 n의 두 번째 DCT 계수를 $dct[n][k]$ 로 표시한다. 예를, 대역 0에 해당하는 44개의 DCT 계수 크기를 44차 DCT 하여 44개의 새로운 두 번째 DCT 계수를 구하며, 그 결과 대역 0의 주파수 영역에 균일하게 분포하던 에너지가 두 번째 DCT 영역에서는 낮은 차수 쪽으로 집중된다.

표 4는 두 번째 DCT에 의한 에너지 집중 효과를 실험적으로 확인한 결과이다. 상관도가 높은 신호를 DCT 함으로써 두 번째 DCT에서는 낮은 차수에 에너지가 모이게 되고, 표 4에서와 같이 모든 대역에서 $dct[n][0] \sim$

표 4. 두 번째 단계의 DCT에 의한 에너지 분포
Table 4. Energy distribution by the 2nd stage DCT

	2nd DCT Size	$dct[n][0] \sim dct[n][10]$	$dct[n][0] \sim dct[n][20]$	$dct[n][0] \sim dct[n][30]$
Band 0	44	82.23 %	88.58 %	93.83 %
Band 1	54	81.68 %	87.02 %	91.29 %
Band 2	66	82.91 %	87.29 %	90.83 %
Band 3	42	84.96 %	90.56 %	95.32 %

$dct[n][10]$ 에 전체 에너지의 80% 이상이 집중하게 되고, $dct[n][0] \sim dct[n][30]$ 에 90% 이상의 에너지가 존재한다. 수치적인 결과 이외에, 각 대역에서 $dct[n][0] \sim dct[n][30]$ 만으로 양자화 없이 신호를 복원하여 청취하면 원음과 동일한 음질을 유지하는 것을 확인할 수 있었다.

이와 같은 통계 결과와 주관적인 평가를 바탕으로 각 대역의 $dct[n][k]$ 중에서 상위 차수 정보를 버리고 $dct[n][0] \sim dct[n][30]$ 만을 양자화 하는 양자화기를 설계하였다. DC 값, 즉 $dct[n][0]$ 은 다른 계수 값에 비해 매우 크고 각 대역의 전체 에너지를 결정하는 매우 중요한 정보이므로 다른 계수들과 함께 벡터 양자화 하지 않고 독립적으로 예측 스칼라 양자화 한다. 표 5는 각 대역의 DC 값들 사이의 상관관계를 구한 결과이다. 하위 대역을 처리하는 G.729 또는 G.723.1의 고정 코드북 이득은 하위 대역의 전체 에너지와 밀접한 관계를 가지므로

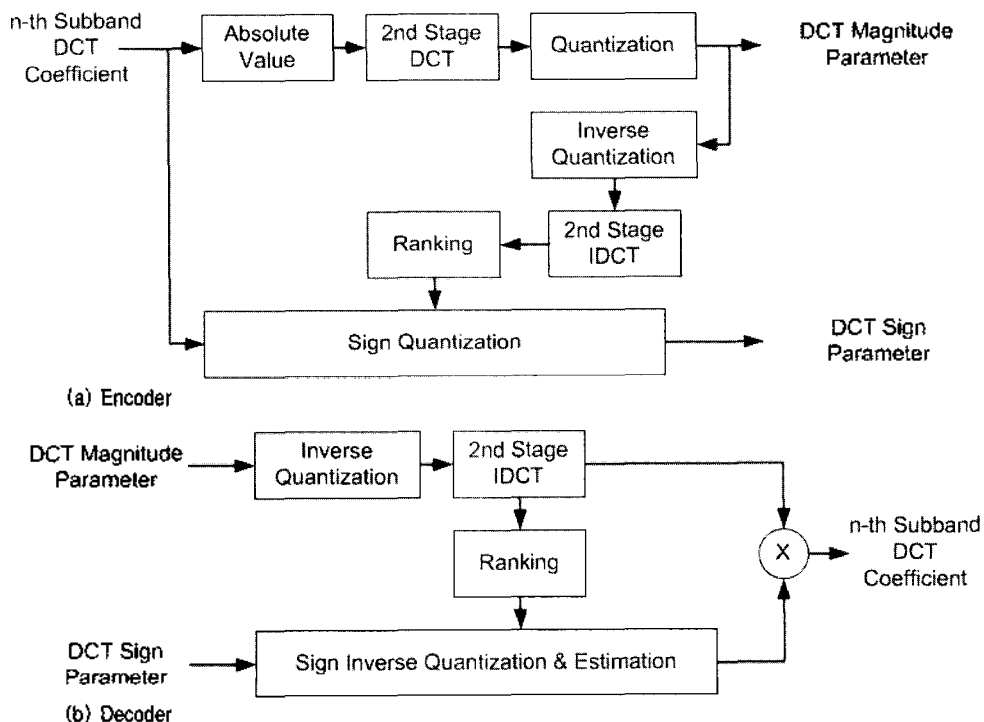


그림 3. DCT 계수의 양자화
Fig 3. Quantization of DCT coefficients

표 5. 파라미터 사이의 상관관계

Table 5. Correlation of parameters

Parameters	Correlation Coefficient	Parameters	Correlation Coefficient
CB Gain : DC(0)	0.8239	DC(0) : RMS(0)	0.9912
DC(0) : DC(1)	0.8311	DC(1) : RMS(1)	0.9911
DC(1) : DC(2)	0.8506	DC(2) : RMS(2)	0.9869
DC(2) : DC(3)	0.8327	DC(3) : RMS(3)	0.9909

30msec 동안의 평균 이득은 대역 0의 DC 값과 상당히 높은 상관도를 가지게 된다. 따라서 대역 0의 DC 값은 하위 대역에서 기본적으로 전달되는 고정 코드북 이득을 이용하여 예측 양자화 하도록 한다. 또한 각 대역의 DC 값들 사이의 상관관계도 매우 높으므로 한 대역의 DC 값으로 그 다음 대역의 DC 값을 예측하여 오차를 양자화 하는 예측 스칼라 양자화를 하며, 1차 AR 모델로 예측하고 분포를 고려하여 μ -Law Companding 하여 양자화 한다.

DC를 제외한 나머지 계수들 $\text{dct}[n][1] \sim \text{dct}[n][30]$ 를 벡터 양자화 하기 위하여 계수들의 정규화와 이를 위한 RMS 양자화가 필요하다. 각 대역 내에서 DC와 RMS의 상관도가 표 5에서 같이 상당히 높으므로 각 대역에서 DC로부터 RMS를 예측하여 양자화 하는 예측 스칼라 양자화를 하고, 1차 AR 모델을 사용하며 분포를 고려하여 μ -Law Companding을 이용하였다.

30개의 정규화된 계수들은 3개의 10차 부벡터 (sub-vector)로 분할하여 각 부벡터를 9 비트로 벡터 양자화를 한다. 총 4개의 대역이 있고 각각 3개의 부벡터, $\text{dct}[n][1] \sim \text{dct}[n][10]$, $\text{dct}[n][11] \sim \text{dct}[n][20]$, $\text{dct}[n][21] \sim \text{dct}[n][30]$ 를 가지므로 총 $4 \times 3 \times 9 = 108$ 비트가 필요 하지만, 각 대역의 중요도를 구하여 양자화에 차등을 주어 신호의 왜곡 없이 비트를 줄이도록 한다. 즉, 가장 중요한 두 대역에서는 3개의 부벡터를 모두 양자화 하여 전달하고, 세 번째 중요 대역에서는 2개의 부벡터 $\text{dct}[n][1] \sim \text{dct}[n][10]$ 과 $\text{dct}[n][11] \sim \text{dct}[n][20]$ 만을 전달하고, 가장 중요도가 낮은 대역에서는 1개의 부벡터 $\text{dct}[n][1] \sim \text{dct}[n][10]$ 만 양자화 하여 전달한다. 결국, 총 9개의 부벡터를 양자화 하여 전달하고, 각 부벡터는 9 비트로 벡터 양자화 하므로 총 $9 \times 9 = 81$ 비트를 사용한다. 대역의 중요도는 각 대역의 에너지를 이용하여 간단하게 구하고, 중요도를 전달하기 위하여 4 비트를 할당한다.

3.4. DCT 계수 부호의 양자화

DCT 계수의 부호는 거의 랜덤 신호의 형태를 가지며

로 각 DCT 계수 부호를 1 비트로 양자화 할 경우 총 206 비트가 필요하게 된다. 그러나 IDCT를 정현파의 선형 조합 관점으로 보면 정현파의 부호는 정현파의 위상에 해당하고 주파수 영역에서의 위상은 청각적 음질에 큰 영향을 주지 않는다. 특히 크기가 매우 작은 DCT 계수의 부호는 IDCT 과정에서 거의 영향을 미치지 못하므로 부호를 전달하지 않고 임의로 정하여 IDCT 하여도 큰 왜곡이 발생하지 않는다. 이와 같은 현상은 크기가 작은 DCT 계수의 부호를 임의로 정하여 합성하여도 거의 음질의 차이를 느끼지 못하는 것을 통하여 검증하였다.

따라서 본 논문에서는 DCT 계수의 부호를 중요도에 따라서 선택적으로 양자화 하는 방법을 제안한다. 즉, 30msec 프레임의 총 206개의 DCT 계수를 크기 순으로 정렬하여 크기가 큰 115개의 DCT 계수의 부호는 1 비트로 양자화 하여 전달하고, 나머지 크기가 작은 DCT 계수의 부호는 전달하지 않고 복원기에서 예측하여 사용한다. 이렇게 함으로써 매우 적은 비트로 부호 정보를 전달할 수 있다. 이 때, 복원기에서도 동일한 크기 순서를 알기 위하여 그림 3과 같이 양자화 된 DCT 계수의 크기를 기준으로 정렬을 한다.

복원기에서 부호를 전달받지 못한 DCT 계수의 부호는 해당 주파수에서 이전 프레임과의 연결성을 유지하여 왜곡을 최소화하기 위하여 이전 프레임으로부터 부호를 예측하여 사용한다. 즉, 인덱스 k 의 DCT 계수 부호는 이전 프레임의 동일한 인덱스 k 에 해당하는 정현파의 마지막 값이 현재 프레임의 새로운 정현파의 시작과 최대한 부드럽게 연결될 수 있도록 그 부호를 정한다. 이 때, 부호가 전달되지 않는 DCT 계수 인덱스는 계속 변하므로 특정 주파수 성분에 오류가 집중되는 현상은 발생하지 않는다. 만일 패킷 손실에 의하여 이전 프레임에 대한 정보가 없을 경우에는 전달되지 못한 부호는 임의로 정하여 사용하며, 이로 인한 왜곡은 거의 발생하지 않는 것을 청취 평가를 통하여 확인하였다.

3.5. 비트 할당

제안하는 상위 대역 부호화기의 비트 할당은 여러 성

표 6. 제안하는 상위 대역 부호화기의 비트 할당
Table 6. Bit allocation of the proposed high-band coder

		Band 0	Band 1	Band 2	Band 3	Sum
DCT Magnitude	DC	6	6	6	6	24
	RMS	4	4	4	4	16
	VQ	9 × 9				81
	Classification	4				4
DCT Sign		115				115
Total Sum						240

표 7. 제안하는 음성 부호화기의 주관적 성능 평가
Table 7. Subjective performance of the proposed codec

	Proposed	No Preference	G.722.1
Male	70.0%	27.5%	2.5%
Female	55.0%	40.0%	5.0%

능 분석과 청취 평가를 통하여 최적화 하였고, 표 6에 그 결과가 정리되어 있다.

IV. 성능 분석

본 논문에서 제안한 대역폭 계층 구조의 광대역 음성 부호화기의 성능을 주관적 선호도 선택 방법으로 측정하였다. 음질 비교를 위하여 하위 대역 신호를 11.8kbps G.729 Annex E 음성 부호화기로 부호화하고 상위 대역은 제안한 부호화기를 사용하여 최종 대역폭 계층 구조의 음성 부호화기를 완성하였다. 음성 신호는 한국 여성과 남성 각각 5명이 발성한 총 20개 문장을 사용하였고, 주관적 음질은 G.722.1 24kbps 표준 광대역 음성 부호화기를 기준으로 평가하였다[7]. 표 7에 정리되어 있듯이 본 논문에서 제안하는 19.8kbps 대역폭 계층 구조의 광대역 음성 부호화기가 24kbps G.722.1보다 우수한 성능을 가지는 것을 확인하였다.

V. 결론

본 논문에서는 대역폭 계층 구조의 광대역 음성 부호화기의 구조를 제안하고 상위 대역 부호화를 위한 새로운 형태의 부호화기를 개발하였다. 입력 신호를 임계 대역 단위로 분해하여 주파수 영역으로 변환하고, DCT 계

수의 크기와 부호를 분리하여 양자화 함으로써 그 성능을 높이는 방법을 제시 하였다. DCT 계수 크기의 상관관계를 활용하기 위하여 두 번째 단계의 DCT를 실시하여 에너지 집중 효과를 통하여 효율적인 양자화를 하였고, DCT 계수 부호는 계수의 중요도에 따라 선택적으로 양자화 하며, 이를 통하여 많은 비트를 절약할 수 있었다. 제안한 대역폭 계층 구조의 음성 부호화기는 24kbps G.722.1 부호화기보다 우수한 성능을 가지는 것을 주관적 성능 평가를 통하여 확인하였다.

본 논문에서 제안한 부호화기는 아직 최적화되지 못한 상태이다. 대역 사이의 차등 양자화를 위하여 보다 지능적인 비트 할당 방법을 개발하는 것이 필요하고, DC와 RMS의 양자화에도 대역별 차등을 주는 것이 필요하다. 또한, 마스킹 효과를 기반으로 대역별 중요도를 보다 정확하게 계산하는 것이 필요하다. 이와 같은 추가 최적화 과정을 통하여 추가적인 성능 향상이 가능할 것으로 기대된다.

감사의 글

본 연구는 삼성종합기술원의 연구비 지원으로 이루어졌습니다.

참고 문헌

- 3GPP TS 26.190 AMR wideband speech codec: Transcoding functions, 2001.
- K. Koishida, V. Cuperman, and A. Gersho, "A 16-kbit/s bandwidth scalable audio coder band on the G.729 standard," Proc. IEEE ICASSP, pp.1149-1152, 2002.
- A. McCree, "A 14kbps wideband speech coder with a parametric highband model," Proc. IEEE ICASSP, 1153-1156, 2000.
- Kyung Tae Kim, Sung Kyo Jung, Young Cheol Park, and Dae Hee Youn, "A new bandwidth scalable wideband speech/audio coder", Proc. IEEE ICASSP, pp.657-660, 2002.
- E. Ambikairajah, J. Epps, and L. Lin, "Wideband speech and audio coding using Gammatone filter bank", Proc. IEEE ICASSP, 773-776, 2001.
- G. Kubin, and W. B. Kleijn, "On speech coding in a perceptual domain", Proc. ICASSP, 205-208, 1999.
- ITU G.722.1, Coding at 24 and 32 kbit/s for hands-free operation in systems with low frame loss, 1999.

저자 이력

● 이 우 석 (Woo-Suk Lee)

2002년 2월: 광운대학교 전자공학과(공학사)

2004년 2월: 광운대학교 전자공학과(공학석사)

2004년 1월- 현재 : 삼성전자 연구원

* 주관심 분야: 음성 신호처리, 통신 시스템 및 단말기

● 박 호 중 (Hochong Park)

1986년 2월: 서울대학교 전자공학과(공학사)

1987년 12월: Univ. of Wisconsin-Madison, 전자공학과(M.S.)

1993년 5월: Univ. of Wisconsin-Madison, 전자공학과(Ph.D.)

1993년 9월- 1997년 8월 : 삼성전자 선임연구원

1997년 9월- 현재 : 광운대학교 전자공학과 부교수

* 주관심 분야 : 음성/오디오 신호처리, 통신 신호처리, 영상 신호처리

● 손 창 용 (Chang-Yong Son)

1999년 2월: 한양대학교 제어계측공학과(공학사)

2001년 2월: 한양대학교 제어계측공학과(공학석사)

2001년 2월-현재 : 삼성종합기술원 연구원

* 주관심분야: 음성/오디오 신호처리, 임베디드 시스템

● 이 영 범 (Yong-Beom Lee)

1987년 2월: 서강대학교 전자공학과 (공학사)

1993년 2월: Michigan State University, 전기공학과 (M.S)

2000년 10월: Michigan State University, 전기공학과 (Ph.D)

2001년 1월~: 삼성종합기술원 전문 연구원

* 주관심 분야: 음성인식, 음성압축, HCI, Interaction 기술