

# 대역폭 계층 구조의 광대역 음성 부호화기를 위한 상위 대역 부호화기 연구

## High-Band Codec for Bandwidth Scalable Wideband Speech Codec

김 영 보\*, 정 병 학\*, 손 창 용\*\*, 성 호 상\*\*, 박 호 종\*

(Youngvo Kim, Byounghak Jeong, Chang-Yong Son\*, Ho-Sang Sung\*, Hochong Park)

\*광운대학교 전자공학과, \*\*삼성종합기술원 Computing Lab

(접수일자: 2005년 5월 30일; 수정일자: 2005년 7월 20일; 채택일자: 2005년 8월 9일)

본 논문에서는 대역폭 계층 구조의 광대역 음성 부호화기를 위한 8kbps 상위 대역 부호화기를 제안한다. 광대역 입력 음성 신호는 하위 대역 신호와 상위 대역 신호로 분리되고, 하위 대역 신호는 표준 협대역 음성 부호화기로 처리하고, 상위 대역 신호는 제안하는 부호화기로 처리하는 구조를 가진다. 제안한 상위 대역 부호화기는 입력 신호를 부프레임 단위로 MLT 변환하고 MLT 계수를 크기와 부호로 분리하여 각각의 특성에 적합하도록 양자화 한다. MLT 계수 크기는 다수의 시간과 주파수 성분을 함께 가지는 밴드로 분할하고, 각 밴드 신호는 2차원 DCT 변환하여 양자화 하며, 하위 대역의 에너지 정보를 이용하여 양자화 성능을 향상시킨다. MLT 계수 부호는 각각의 중요도를 판정하여 일부 부호만 선택적으로 양자화 하는 방법을 사용한다. 제안한 상위 대역 부호화기를 포함하는 19.8kbps 광대역 부호화기의 객관적 성능과 주관적 성능을 측정하였으며, 32kbps G.722.1보다 우수한 성능을 가지는 것을 확인하였다.

**핵심어:** 광대역 음성 부호화기, 대역폭 계층 구조, MLT, 2차원 DCT

**투고분야:** 음성처리 분야 (2.2)

In this paper, the high-band codec for bandwidth scalable wideband speech codec is proposed. The wideband input speech signal is separated into low-band signal and high-band signal, and the low-band signal is encoded by the standard narrow-band speech codec and the high-band signal is encoded by the proposed codec. In the high-band codec, the signal is transformed into frequency domain by MLT on a subframe basis, and MLT coefficients are splitted into magnitude and sign for quantization. The magnitudes of MLT coefficients are arranged into several time-frequency bands and each band is quantized in 2D-DCT domain, where the low-band information is utilized for better performance. The sign of MLT coefficient is quantized based on a priority selection process with the weighting measurement. The objective and subjective performance of wideband speech codec including the proposed high-band codec is measured, and it is confirmed that the proposed codec has better performance than 32kbps G.722.1.

**Keywords:** wideband speech codec, bandwidth scalability, MLT, 2D-DCT

**ASK subject classification:** Speech Signal Processing (2.2)

### I. 서론

현재 디지털 통신에서 사용되는 대부분의 음성 부호화기는 8kHz로 샘플링 되어 4kHz의 대역폭을 가지는 협대역 음성 신호를 제공하고 있다. 그러나 통신 기술이

발달하고 통신 서비스가 다양해짐에 따라 많은 영역에서 고품질의 음성 서비스가 요구되고 있고, 또한 통화 품질이 통신의 품질 평가에서 매우 중요한 기준으로 활용되고 있다. 이러한 조건을 충족시키기 위해 16kHz로 샘플링 되고, 7kHz 대역폭을 가지는 광대역 음성 신호의 전달이 필요하게 되었다. 또한, 새로운 광대역 음성 통신에서는 기존의 협대역 시스템과의 호환성을 유지하기 위하여 하위 대역과 상위 대역 신호를 계층적으로 압축하

여 전달하는 계층적 구조의 음성 부호화기가 필요하다. 이러한 요구에 따라 대역폭 계층 구조의 광대역 음성 부호화기에 대한 많은 연구가 진행되고 있으며, 협대역 통신에 대한 호환성을 유지하기 위하여 표준 협대역 음성 부호화기를 이용하여 4kHz 이하의 하위 대역 신호를 압축하고, 4kHz 이상의 상위 대역 신호는 보통 Transform 구조와 파라미터 기반의 방법을 사용하며 상위 대역 신호의 특성을 고려하여 신호를 부호화 한다[1-4].

본 논문에서는 광대역 입력 신호를 4kHz를 기준으로 상위 대역 신호와 하위 대역 신호를 분리하여 하위 대역 신호는 표준 협대역 음성 부호화기인 G.729, G.729E, G.723.1 부호화기를 기준으로 부호화 하는 구조를 가정 하며[5-6], 본 논문의 핵심 연구 내용은 4~7kHz의 상위 대역 신호의 부호화를 위하여 8kbps 전송률을 가지는 부호화기를 연구하여 최종적으로 대역폭 계층 구조의 광대역 음성 부호화기를 개발하는 것이다. 특히, 개발하는 상위 대역 부호화기는 계층적 구조의 부호화기의 제약점에 의하여 과거의 정보를 사용하지 않고 프레임 단위로 독립적인 동작이 가능하도록 하며, [4]에 발표되었던 구조를 기반으로 문제점을 분석하여 새로운 변환 구조와 양자화 방법을 새로 개발하여 향상된 성능의 상위 대역 부호화기를 개발하였다.

본 논문의 구성은 다음과 같다. 2장에서 제안하는 상위 대역 부호화기의 동작을 상세하게 설명하고 3장에서 개발한 음성 부호화기의 성능을 측정하고 4장에서 결론 및 앞으로의 연구 방향을 설명한다.

## II. 제안하는 상위 대역 부호화기

### 2.1. 전체 구조

본 논문에서 제안하는 대역폭 계층 구조의 광대역 음성 부호화기의 전체 구조는 그림 1과 같다. 광대역 입력 신호는 Quadrature Mirror Filter (QMF)를 통하여 하위 대역과 상위 대역 신호로 분리되고, 하위 대역 신호는 G.729 또는 G.723.1 표준 협대역 음성 부호화기로 부호화하며, 상위 대역 신호는 제안하는 음성 부호화 방법에 의하여 부호화한다. 이 때, 상위 대역 부호화기는 하위 대역의 부호화된 에너지 정보 (예로, 고정 코드북 이득)와 상위 대역의 상관관계를 활용하여 부호화 성능을 높이도록 한다. 상위 대역 부호화기는 입력 광대역

표 1. 제안하는 상위 대역 음성 부호화기의 규격

Table 1. Specification of proposed high-band speech codec.

Bit Rate	8 kbps
Frame Size	30 msec
Subframe Size	5 msec
Look-Ahead	5 msec

음성 신호로부터 분리된 4~7kHz 대역의 신호를 처리하도록 설계하며, 표 1과 같은 규격을 가진다. 상위 대역 부호화기의 전송률은 8kbps이고, 표준 8kbps 협대역 음성 부호화기를 사용할 경우 총 16kbps 전송률을 가지며, 하위 대역에 개선 계층에 4kbps가 추가될 경우 전체 전송률은 20kbps이다. 상위 대역의 프레임 크기는 ITU G.729와 G.723.1을 하위 대역 부호화에 모두 사용할 수 있도록 하기 위하여 30ms로 정하였으며, Look-Ahead는 G.729의 5ms를 초과하지 않도록 설계하였다[5].

제안하는 상위 대역 부호화기는 Transform Coding 구조에 기반을 둔다. 입력 신호를 주파수 영역으로 변환하여 주파수 계수를 구하고, 주파수 계수에 대한 일련의 처리를 거친 후 청각적 특성을 고려하여 양자화 한다.

그림 2는 제안한 상위 대역 부호화기의 전체 구조를 보여준다. 상위 대역 부호화의 첫 단계로서 입력 신호에 대하여 5msec 크기의 부프레임 단위로 Modulated Lapped Transform (MLT)를 실시한다. 효율적인 부호화기의 설계를 위하여 MLT 계수를 직접 양자화하지 않고 계수 값들에 대해 크기 (Magnitude) 정보와 부호 (Sign) 정보를 분리하여 각각의 특성을 활용하여 독립적으로 양자화 한다[4]. 또한 시간적으로 진행되는 MLT 계수 크기 값들 사이의 시간 축 상관관계도 활용하기 위하여 부프레임 단위의 주파수 계수를 시간 축으로 배열하여 2차원적인 구조를 구성하고 추가적인 2차원 Discrete Cosine Transform (DCT)을 통하여 시간 축과

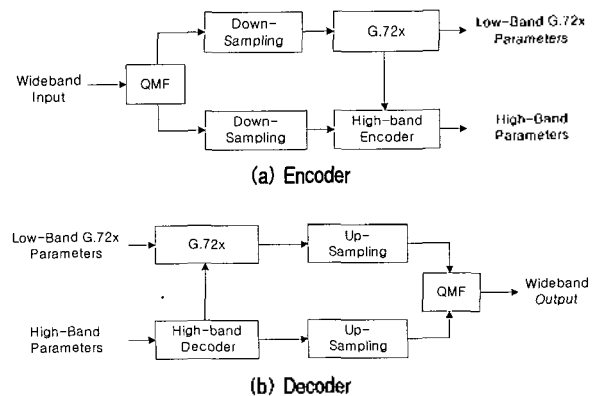


그림 1. 제안하는 대역폭 계층 구조의 광대역 음성 부호화기의 전체 구조

Fig. 1. Overall structure of proposed bandwidth scalable wideband speech codec.

주파수 축에 대하여 동시에 상관관계를 활용하도록 한다. 이 방법은 천이 구간에서와 같이 시간적으로 진행되는 각각의 주파수 성질에 매우 상이한 특성을 가질 경우 효율이 저하될 수 있지만 음질에 더 큰 영향을 미치는 정상 구간에서의 성능 향상을 위하여 활용한다. 2차원 DCT 계수는 DC 성분과 나머지 성분으로 분리하여 각각 독립적으로 양자화 하며, 대역의 중요도에 따라 차별화된 양자화 방법과 비트 할당을 사용한다. MLT 계수의 부호는 [4]와 동일하게 양자화된 MLT 계수 크기의 정보를 활용하여 중요한 일부 계수의 부호만 선택하여 전달하고 나머지 부호는 복원기에서 임의로 정하여 사용한다.

이와 같이 짧은 부프레임 단위로 MLT를 실시하고 이를 다시 2차원 구조로 DCT 하는 방법을 통하여 시간과 주파수 영역에서의 상관관계를 동시에 활용하도록 하였으며, [4]와 차별화 되는 새로운 구조이다. 또한, 새롭게 제안하는 2-D DCT 계수의 양자화를 위하여 2차원 데이터의 에너지 집중 성질에 따라 1차원 재배치를 포함하는 새로운 양자화 방법을 사용한다.

**2.2 MLT 변환**

Transform 부호화기에서 Transform의 크기와 방식은 주파수 계수의 성질을 규정하고 양자화에 큰 영향을 미치는 중요한 요소이다. 본 논문에서는 프레임 경계에서 양자화로 인한 왜곡을 줄이는데 효과적이고 Aliasing에 대해 우수한 성질을 가지는 것으로 알려진 MLT를 사용하였다[6]. 즉, 그림 3과 같이 주어진 30msec 프레임의 입

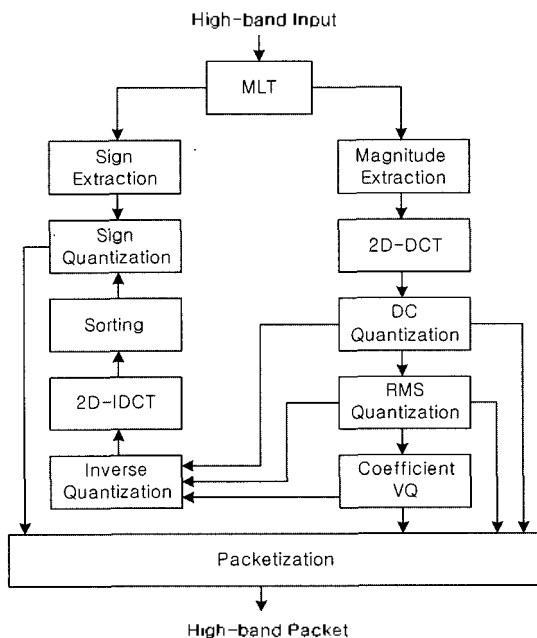


그림 2. 상위 대역 Encoder의 세부 구조  
Fig. 2. Detailed structure of high-band encoder.

력 신호를 6개의 5msec 부프레임으로 분리하고, 각 부프레임에 대하여 5msec Look-Ahead를 포함하여 총 10msec 크기의 MLT를 계산하며, 이 때 이웃한 MLT 윈도우사이에는 50%의 중첩을 가진다. 따라서 프레임 전체에 대하여 5msec의 Look-Ahead를 가지게 된다.

그림 4는 각 부프레임 단위로 진행되는 MLT으로부터 구한 주파수 계수를 2차원으로 배열하는 과정을 도식화한 것이다. 각 부프레임에 대하여 MLT는 80 샘플(10msec)에 적용되고 부프레임마다 40개의 MLT 계수가 생성된다. 이 중에서 4.0~7.0kHz에 해당하는 30개의 MLT 계수만을 취하게 되며, 그림 5에서 가로 형태로 나열되어 있다. 각 부프레임별 30개의 MLT 계수는 다시 5개의 600Hz 단위의 주파수 대역으로 분리하여 처리한다.

이와 같은 동일한 동작이 부프레임마다 반복 적용되며 각 부프레임별로 구하여진 30개씩의 MLT 계수를 시간 축으로 나열하여 시간과 주파수 축에서 2차원으로 표현하면 그림 4와 같이 6×30크기의 행렬로 표시되고, 한 프레임에 대한 600Hz 대역은 6×6 크기의 부행렬로 표현된다. 즉, 프레임에 대한 총 180개의 MLT 계수는 6×6 크기의 행렬 데이터를 가지는 5개의 대역으로 구성된다.

**2.3 MLT 계수의 크기와 부호 분리**

MLT 계수의 효율적인 양자화를 위하여 계수들 사이의 상관관계를 충분히 활용하여야 한다. 그러나 MLT 계수들 사이의 시간과 주파수에 대한 상관관계를 측정해보면 평균적으로 0.01 정도로 매우 작다. 이를 해결하기 위하여 본 논문에서는 계수의 크기 (Magnitude)와 부호 (Sign) 성분을 분리하는 방법을 사용하며, MLT 계수 크기만을 취하여 크기에 대한 상관관계를 측정해보면 주파수 영역과 시간 영역에서 평균적으로 0.57 정도의 상관관계를 가지는 것을 확인할 수 있다. 각 부프레임별로 주파수 영역에서 인접한 MLT 계수 사이의 상관관계와 인접한 MLT 계수 크기 사이의 상과 상관관계를 측정할 결과는 표 2에 정리되어 있고, 5개의 대역에 대해 시간

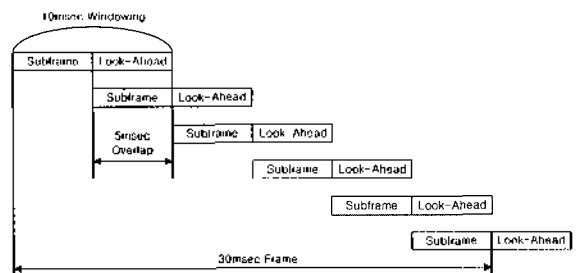


그림 3. 부프레임 단위의 MLT 변환 구조  
Fig. 3. Structure of subframe-based MLT.

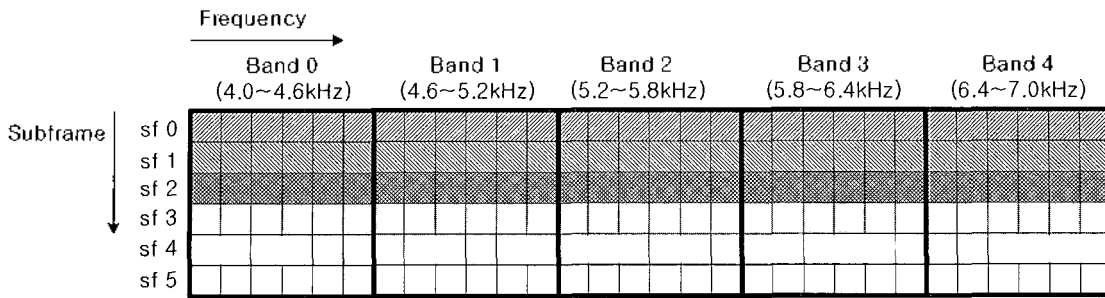


그림 4. 부프레임 단위의 MLT 구조와 MLT 계수들의 2차원 배열 구조  
 Fig. 4. Subframe-based MLT structure and two-dimensional structure of MLT coefficients.

영역에서 구한 상관관계 계수는 표 3에 있다.

위에서 구한 통계적 분석 결과에 의해 MLT 계수는 부호를 분리하여 크기의 성분으로 상관관계를 높일 수 있고, 이러한 구조를 이용하여 양자화기의 성능을 높일 수 있다. 또한, 주파수 영역과 시간 영역에서 매우 비슷한 상관관계 계수 값을 가지는 점을 활용하기 위하여 그림 4와 같은 시간과 주파수에 대한 2차원 계수 배열을 통하여 시간과 주파수에서 동시에 에너지를 압축하는 2차원 DCT를 사용하며, 1차원적인 처리에 비하여 에너지 집중(compaction)에서 큰 성능 향상을 얻을 수 있다.

### 2.4 MLT 계수 크기의 양자화

#### (a) 전체 구조

MLT 계수 크기는 시간과 주파수 측면의 상관관계를 이용하기 위해 2차원 DCT로 처리하여 에너지 집중 효과를 향상시키며, 2차원 DCT의 효과를 검증하기 위하여 2

차원 DCT 계수의 에너지 분포를 분석하였다. 2차원 DCT 계수의 에너지 분포를 1차원적으로 분석하기 위하여 각 대역별로 6×6의 2차원 DCT 계수를 각 Element의 평균 에너지 순으로 순위를 구하여 에너지 순으로 1차원으로 나열하고, 이를 k[0] ~ k[35]로 표시하며, k[0] ~ k[35] 순서에 따라 에너지의 집중 정보를 구하면 표 4와 같다. DC 값에 해당하는 k[0]에 75% 이상의 에너지가 집중되고, 36개의 계수 중에서 첫 10개 정도의 계수에 90% 이상의 에너지가 분포한다. 이를 통하여 2차원 DCT로부터 시간-주파수 영역에서 에너지의 집중 효과를 얻을 수 있고, 이를 활용하여 효율적인 양자화를 할 수 있게 된다. 2차원 DCT 계수를 에너지 분포에 따라 1차원으로 재배열 하는 기준은 매 프레임마다 변하는 것이 최적이지만 본 논문에서는 각 대역별로 각 계수의 평균 에너지를 구하여 모든 프레임에 대하여 고정된 방식으로 1차원 배열을 한다.

2차원 DCT에 의한 에너지 집중의 효과를 청각적으로 검증하기 위하여 다음과 같은 실험을 하였다. 180개의 2차원 DCT 계수를 모두 사용하지 않고 일부분의 계수만 사용하여 신호를 합성하도록 하였으며, 각 대역별로 대역 에너지를 구하고 큰 에너지 대역 순서에 따라 전달하는 2차원 DCT 계수의 수를 36개, 27개, 19개, 11개로 한정하여 총 129개의 계수만 사용하여도 청각적으로 차이가 없는 결과를 얻을 수 있었다.

MLT 계수 부호의 중요도를 정하기 위하여 양자화된 MLT 계수 크기를 구하여야 하며, 이를 위하여 내부적으로

표 2. 주파수 영역에서 인접한 MLT 계수의 상관관계와 MLT 계수 크기의 상관관계

Table 2. Average correlation coefficients between adjacent MLT coefficients and adjacent magnitudes of MLT coefficients in frequency domain.

	MLT coefficient	Magnitude of MLT coefficient
Subframe 0	0.0049	0.5647
Subframe 1	0.0125	0.5744
Subframe 2	0.0076	0.5645
Subframe 3	0.0001	0.5714
Subframe 4	0.0138	0.5745
Subframe 5	0.0019	0.5796

표 3. 시간 영역에서 인접한 MLT 계수의 상관관계와 MLT 계수 크기의 상관관계  
 Table 3. Average correlation coefficients between adjacent MLT coefficients and adjacent magnitudes of MLT coefficients in time domain.

	MLT coefficient	Magnitude of MLT coefficient
Band 0	0.0047	0.5843
Band 1	0.0016	0.5894
Band 2	0.0034	0.5739
Band 3	0.0001	0.5745
Band 4	0.0025	0.5736

표 4. 2차원 DCT 계수의 에너지 집중 효과  
 Table 4. Energy Compaction of 2D-DCT.

	DC k[0]	k[0]-k[5]	k[0]-k[11]	k[0]-k[19]	k[0]-k[27]	k[0]-k[35]
Band 0	77.6 %	87.46 %	92.39 %	95.13 %	97.59 %	100 %
Band 1	76.7 %	86.57 %	91.71 %	94.61 %	97.32 %	100 %
Band 2	76.6 %	87.02 %	91.74 %	94.63 %	97.38 %	100 %
Band 3	77.5 %	87.38 %	91.56 %	94.51 %	97.28 %	100 %
Band 4	76.7 %	87.13 %	91.53 %	94.45 %	97.25 %	100 %

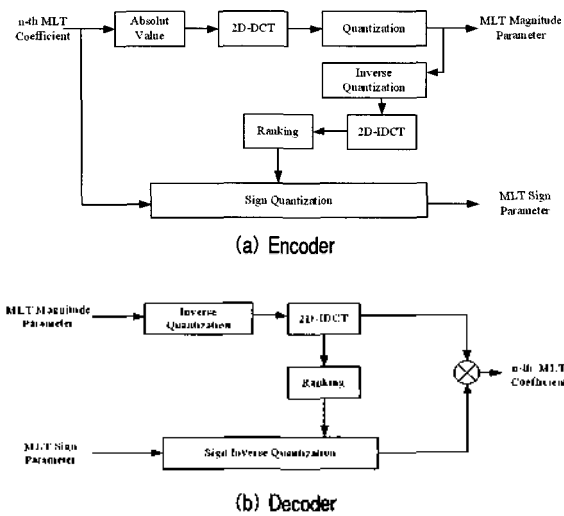


그림 5. MLT 계수의 양자화 구조  
 Fig. 5. The quantization structure of MLT coefficients.

로 MLT 계수 크기의 역양자화기와 크기의 순서를 구하는 모듈을 포함한다. 이상과 같이 MLT 계수의 통계적 특성을 분석하고, 양자화에 적합하도록 추가적인 변환을 적용하여 최종 MLT 계수의 양자화 방법을 개발하였고, 이를 정리하면 그림 5와 같다.

(b) DC 양자화

2차원 DCT 계수의 DC 값은 다른 DCT 계수들에 비하여 매우 크고, 각 대역의 전체 에너지를 결정하는데 매우 중요하기 때문에 다른 값들과 함께 벡터 양자화 하지 않고, 각 대역의 DC 값 모두 스칼라 양자화를 한다. 한편, CELP 구조를 가지는 G.729 또는 G.723.1의 고정 코드북 이득은 음성의 전체 에너지와 밀접한 관계를 가지고, 표 5와 같이 첫 번째 대역 (Band 0)의 DC 값과 상당히 높은 상관도를 가지게 된다. 따라서 하위 대역의 추가적인 에너지 계산을 하지 않고 기본적으로 전달되는 고정 코드북 이득을 양자화에 이용을 할 수 있다. 또한 각 대역의 DC 값들 사이의 상관도 역시 매우 높다는 것을 알 수 있으며, 이 성질을 바탕으로 예측 스칼라 양자화를 실시한다.

(c) 계수 양자화

DC를 제외한 나머지 계수들을 벡터 양자화 하기 위해

서 계수들의 정규화가 필요하고, 이를 위하여 RMS의 양자화가 필요하다. 표 5에서 볼 수 있듯이 RMS 값은 각 대역 내의 DC 값과 높은 상관도를 가지므로 로그 영역에서 예측 양자화를 한다.

정규화된 DCT 계수는 벡터 양자화 하며, 비트의 사용 효율을 높이기 위하여 각 대역의 RMS 크기 순서에 따라서 다른 VQ 구조와 비트 수를 할당한다. 본 논문에서는 RMS가 큰 두 개의 대역에서 대하여 35개 (5-6-8-8-8 Split)의 계수를 모두 양자화하고, 3번째 대역은 27개 (5-6-8-8 Split)의 계수, 4번째 대역은 19개 (5-6-8 split)의 계수, 마지막 대역은 11개(5-6 Split)의 계수를 양자화 하여 전달한다. 이를 다른 관점으로 보면, 모든 대역에 대하여 35차 벡터를 5-6-8-8-8차로 Split 하고, 각 대역별로 전달할 부벡터 (subvector)와 전달하지 않을 부벡터를 선택적으로 정하는 개념이다. 예로, 4번째 대역에서는 첫 3개의 부벡터는 전달하고 나머지 2개의 부벡터는 전달하지 않는다. 또한, Split 과정에서 상위 부벡터의 차수를 8로 통일하여 동일한 VQ 코드북을 사용하고 이로부터 필요한 코드북 크기를 줄이도록 하였다. VQ를 위한 비트 할당은 청각적 중요도의 차이를 반영하기 위하여 대역별 순위에 따라 동일한 크기의 부벡터에 대하여 서로 다른 비트를 할당하며, 대역 순위에 따라 총  $36 + 28 + 21 + 11 + 7 = 103\text{bits}$ 를 할당한다.

2.5 MLT 계수 부호의 양자화

MLT 계수의 부호는 불규칙하고 상관관계가 매우 낮은 특성을 가지므로 180개의 모든 부호를 적은 비트로 양자화가 매우 어렵다. 즉 하나의 MLT 계수 부호 양자화를 위해 1 bit가 필요하지만 부호 양자화에 180 비트를 사용할 수는 없다. MLT를 정현파들의 선형 합 관점에서 보면 부호는 정현파의 위상으로 볼 수 있으며, 실험을 통하여 크기가 작은 MLT에 해당하는 부호를 전달하지 않고 임의로 정하여도 이에 의한 청각적 왜곡이 크지 않다는 것을 확인하였다. 따라서 모든 부호를 전달하지 않고 중요도에 따라 일부 부호만 선택하여 전달하는 방식을 제안한다.

본 논문에서 제안하는 부호화기에서는 MLT 계수 부호 양자화를 위해서 92 비트를 할당하였으며, 그림 5와 같이 복호화된 MLT 계수 크기 정보를 가지고 크기의 순서대로 순번을 정하여 92개의 MLT 계수에 대해서만 부호를 1비트로 양자화 하여 전달한다. 이와 같은 부호 선택

표 5. DC와 RMS 값에 대한 상관관계  
 Table 5. Correlation of DC and RMS parameters.

Parameters	Correlation Coefficient	Parameters	Correlation Coefficient
CB gain : DC[0]	0.87	DC[0] : RMS[0]	0.99
DC[0] : DC[1]	0.90	DC[1] : RMS[1]	0.98
DC[1] : DC[2]	0.83	DC[2] : RMS[2]	0.97
DC[2] : DC[3]	0.87	DC[3] : RMS[3]	0.99
DC[3] : DC[4]	0.87	DC[4] : RMS[4]	0.99

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

		Band 0	Band 1	Band 2	Band 3	Band 4	Total
MLT Magnitude	DC	5	5	5	5	5	25
	RMS	4	4	4	4	4	20
	Vector Quantization	36 + 28 + 21 + 11 + 7 = 103bits					103
MLT Sign	Sign Quantization	92					92
Total							240

방법을 통하여 순서에 대한 추가적인 정보의 전달 없이 상위 대역 복원기가 선택된 부호를 알 수 있게 된다.

부호를 전달 받지 못하는 88개의 계수들에 대하여 (i) 임의로 선택하는 방법; (ii) 주파수 영역에서 동일한 위치의 계수 값들에 대하여 이전 프레임의 정현파와 현재 프레임의 정현파가 최대한 부드럽게 연결되도록 해당 주파수 계수의 부호를 정하는 방법; (iii) 부호가 할당된 계수들 사이에서의 평균값의 부호로 예측하는 방법 등을 사용하였으나 뚜렷한 성능의 차이를 보여주지 못하여 랜덤 함수를 이용하여 부호를 정하도록 하였다.

2.6 비트 할당

제안하는 상위 대역 부호화기에서 각 파라미터에 할당된 비트는 표 6에 정리되어 있다.

III. 성능 분석

제안한 상위 대역 부호화기의 성능 평가를 위하여 하위 대역에 G.729E를 사용하고 상위 대역에 제안한 부호화기를 사용하는 19.8kbps 광대역 부호화기를 제작하여 이용하였다[5]. 객관적 성능 평가를 위하여 WPESQ (Wideband Perceptual Evaluation of Speech Quality) 를 사용하였으며, 표 7과 같이 32kbps G.722.1 보다 우수한 성능을 확인할 수 있다[7].

제안하는 부호화기의 주관적 품질을 측정하기 위한 간단한 A/B 선호도 실험을 진행하였다. 한글 남성 10명과 여성 10명이 각 2분장씩 발성한 40개의 실험 데이터를 이용하였고, 20-30대 남녀 대학생 각각 10명씩 총 20명의 청취 실험자를 대상으로 실시한 32kbps G.722.1에 대한 선호도 결과가 표 8에 정리되어 있으며, 이를 통하여 제안한 계층 구조의 음성 부호화기가 청각적으로 우수한 성능을 가지는 것을 확인할 수 있다.

표 7. 제안하는 음성 부호화기의 객관적 성능 평가  
Table 7. Objective performance of proposed speech codec.

	Bitrate	WPESQ
G.729E + Proposed High-Band Codec	19.8 kbps	3.333
ITU-T G.722.1	32 kbps	3.296

표 8. 제안하는 음성 부호화기의 주관적 성능 평가  
Table 8. Subjective performance of proposed speech codec.

	Proposed Codec Preferred	G.722.1 Preferred	No Preference
Male Speech	70%	10%	20%
Female Speech	80%	5%	15%

IV. 결론

본 논문에서는 계층적 음성 부호화기를 위한 새로운 상위 대역 부호화기를 제안하였다. 제안된 상위 대역 부호화기는 시간과 주파수 영역에서의 상관관계를 적극적으로 이용하기 위하여 2차원 구조의 데이터와 2차원 DCT를 사용하였으며, 양자화 효율을 증가시키기 위하여 계수의 크기와 부호를 분리하여 독립적으로 양자화 하였다. 또한, 부호 양자화는 계수 크기에 따라 선택적으로 실시하여 적은 비트로 부호를 양자화 하도록 하였다. 제안한 음성 부호화기의 객관적 성능과 주관적 성능을 측정 하였으며, 제안한 상위 대역 부호화기는 낮은 전송률을 가지면서도 기존의 광대역 음성 부호화기에 비하여 우수한 성능을 가지는 것을 확인하였다.

본 논문에서 제안한 구조에 대하여 잡음 환경 및 프레임 오류 환경에서의 성능을 분석하고 이 환경에서의 성능 개선을 위한 개선 방법에 대한 연구가 진행 중이다. 또한, 전체 동작에서 아직 성능 저하를 유발하는 부분이 존재하며 이 부분의 추가 개선을 통하여 음성 압축기의 성능을 높이도록 할 예정이다. 제안한 부호화기의 큰 특징 중에 하나인 2차원 DCT 과정에서 입력 신호의 특성을 고려하지 않아 각 프레임에 최적의 동작을 수행하지 못하고 있다. 즉, 부프레임 사이의 상관도가 높은 프레임과 그렇지 않은 프레임에 대하여 2차원 DCT 계수의 분포는 매우 다를 것이며, 이에 따라 2차원 DCT 구조를 입력 특성에 따라 6×6 DCT, 2번의 3×6 DCT, 또는 6번의 1×6 DCT 등에서 선택하는 구조에 대하여 연구를 진행하고 있다. 또한, 부프레임 단위로 구한 2차원 DCT 계수를 크기 순서에 따라 1차원으로 배열하는 과정에서 고정된 방식보다는 각 프레임의 특성에 따라 가변적이 방식이 필요하다.

## 감사의 글

본 논문은 삼성전자의 연구비 지원과 2004년도 연구년 수행에 의한 연구 결과입니다.

## 참고 문헌

1. A. McCree, "A 14kbps wideband speech coder with a parametric highband model," Proc. IEEE ICASSP, 1153-1156, 2000.
2. K. Koishida, V. Cuperman and A. Gersho, "A 16-kbit/s bandwidth scalable audio coder band on the G.729 standard," Proc. IEEE ICASSP, 1149-1152, 2002.
3. Kyung Tae Kim, Sung Kyo Jung, Young Cheol Park, and Dae Hee Youn, "A new bandwidth scalable wideband speech/audio coder", Proc. IEEE ICASSP, 657-660, 2002.
4. 이우석, 박호종, 손창용, 이영범, "대역폭 계층 구조의 광대역 음성 부호화기 개발," 한국음향학회지, 23 (6), 400-406, 2004. 8.
5. ITU-T Rec. G.729, "Coding of speech at 8kbit/s using conjugate-structure algebraic-code-excited linear-prediction(CS-ACELP)," 1996.
6. ITU-T Rec. G.722.1, "Coding at 24 and 32 kbit/s for hands-free operation in systems with low frame loss," 1999.
7. ITU-T Recommendation P.862, "Perceptual evaluation of speech quality," Version 1.2-2, August 2002.

---

## 저자 약력

---

### • 김 영 보 (Youngvo Kim)

2003년 2월: 광운대학교 전자공학과(공학사)  
 2005년 2월: 광운대학교 전자공학과(공학석사)  
 2005년 1월~현재: LG전자 연구원  
 ※ 주관심 분야: 음성/오디오 신호처리, 이동통신 시스템

### • 정 병 학 (Byounghak Jeong)

2003년 2월: 광운대학교 전자공학과(공학사)  
 2005년 2월: 광운대학교 전자공학과(공학석사)  
 2005년 1월~현재: LG전자 연구원  
 ※ 주관심 분야: 음성/오디오 신호처리, 이동통신 시스템

### • 손 창 용 (Chang-Yong Son)

1999년 2월: 한양대학교 제어계측공학과(공학사)  
 2001년 2월: 한양대학교 제어계측공학과(공학석사)  
 2001년 2월~현재: 삼성종합기술원 연구원  
 ※ 주관심 분야: 음성/오디오 신호처리, 임베디드 시스템

### • 성 호 상 (Ho-Sang Sung)

1996년 2월: 한양대학교 제어계측공학과(공학사)  
 1999년 2월: 한양대학교 제어계측공학과(공학석사)  
 1999년 6월~2003년 8월: 한국전자통신연구원 연구원  
 2003년 9월~현재: 삼성종합기술원 연구원

※ 주관심 분야: 음성/오디오 신호처리, 임베디드 시스템

### • 박 호 종 (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, 영상신호처리, 이동통신 시스템