

# 하모닉 코더를 위한 직선과 이산코사인변환 (DCT)을 이용한 하모닉 크기값 (Magnitude) 양자화 기법

## Method of Harmonic Magnitude Quantization for Harmonic Coder Using the Straight Line and DCT (Discrete Cosine Transform)

최 지 욱\*, 정 규 혁\*, 이 인 성\*  
(Ji-Wook Choi\*, Gyu-Hyeok Jeong\*, In-Sung Lee\*)

\*충북대학교 전파공학과  
(접수일자: 2008년 4월 15일; 채택일자: 2008년 5월 20일)

본 논문은 음성신호의 주파수를 반으로 나누어 양자화 계수 추출에 직선과 DCT (Discrete Cosine Transform)를 혼합하여 사용한 양자화 기법을 제안한다. 하모닉 부호화기에서의 하모닉은 가변이므로 저주파에서의 하모닉만을 오버샘플링하여 차원을 고정한다. 고정된 차원의 스펙트럼 포락선을 직선으로 표현한 후 직선이 불연속인 주파수에서의 샘플값만을 전송한다. 이 논문에서는 저주파에서의 하모닉은 복원을 위해 고정된 직선의 계수를 사용하였고 고주파에서의 하모닉은 가변 DCT를 적용하여 직선과 DCT가 결합된 형태의 양자화기법을 제안한다. 제안된 양자화 기법의 성능평가는 스펙트럴의 크기값에 대한 Spectral Distortion (SD)값을 사용하였다. 실험결과 HVXC보다 제안된 양자화 기법의 SD가 0.3dB 정도 개선되었다.

**핵심용어:** 이산 코사인 변환, 직선, 오버샘플링, HVXC

**투고분야:** 음성처리 분야 (2.1, 2.2)

This paper presents a method of quantization to extract quantization parameters using the straight-line and DCT (Discrete Cosine Transform) for two splitted frequency bands. As the number of harmonic is variable frame to frame, harmonics in low frequency band is oversampled to fix the dimension and straight-lines present a spectral envelope, then the discontinuous points of straight-lines in low frequency is sent to quantizer. Thus, extraction of quantization parameters using the straight-line provides a fixed dimension. Harmonics in high frequency use variable DCT to obtain quantization parameters and this paper proposes a method of quantization combining the straight-line with DCT. The measurement of proposed method of quantization uses spectral distortion (SD) for spectral magnitudes. As a result, The proposed method of quantization improved 0.3dB in term of SD better than HVXC.

**Keywords:** DCT, Straight-line, over-sampling, HVXC

**ASK subject classification:** Speech Signal Processing (2.1, 2.2)

### I. 서론

오늘날 이동통신 시스템의 발전과 수요의 증가는 통신 시스템의 다양한 수용능력과 더 나은 서비스를 요구한다. 이러한 수요의 증가는 낮은 전송률과 음질향상의 문제점을 부각시킨다. 전송률이 4-16 kbit/s 주위의 Vector eXcitation Coding (VXC)로 잘 알려진 Code Excited Linear Prediction (CELP)는 디지털 모바일 폰과 전화에서 많이 사용되어진다 [1]. 그러나 CELP는 전송률 4 kbit/s 아래

에서 그 성능이 저하된다. 반면에 기본주파수 (Fundamental frequency), 크기 (Magnitude), 위상 (Phase)을 갖는 정현파 성분의 선형합으로 정의하는 정현파 모델은 낮은 비트 전송률로 음성신호를 부호화하는 효율적인 기술로 알려져 왔다. 좋은 음질을 유지하면서 더 낮은 전송률을 추진하기 위해 Mixed Excited Linear Prediction (MELP), Waveform Interpolation (WI),와 Harmonic Vector eXcitation coding (HVXC)와 같이 Linear prediction Coding (LPC)와 정현파 코딩을 효율적으로 결합시키는 연구가 진행되었다 [10][11][12]. HVXC는 LPC잔여신호의 파라미터 표현을 활용하고 전송률 4 kbit/s 아래에서 좋은 음질을 얻을 수 있다 [5][1].

OGG Vorbis 코덱은 동일한 비트율에서 MP3 (MPEG Audio Layer-3)보다 음질이 좋고, AAC와는 음질이 비슷하거나 그 이상의 음질을 보장한다 [9]. 통신 시스템의 수용능력이 좋아짐에 따라 고비트율의 음성신호의 음질 개선에 오디오 신호처리 방식을 적용하는것이 음질 개선에 좋은 결과가 예상되었다. 따라서 4 kbit/s 이상에서 제안된 코더의 음질을 개선하기 위해 오디오 코덱인 OGG Vorbis에서 스펙트럴 커브를 직선으로 표현하는 방법을 음성신호 양자화에 부분적으로 적용하게 되었다. 이와 같은 방법은 비트율이 증가할수록 더 나은 성능을 기대할 수 있으며, 직선이 적용되지 않은 나머지부분은 DCT를 사용함으로써 코드북 효율을 증가시킨다.

HVXC는 양자화 계수 추출에서 고정된 44개를 사용하므로 차원고정에서 하모닉에 대한 정보손실이 예상된다. 이점을 보완하기 위해 주파수를 반으로 나누어 저주파와 고주파로 양자화 계수 추출방법을 다르게 적용하였다. 저주파는 22개에서 64개로 차원을 고정하고 스펙트럼 포락선을 직선으로 표현하여 사용하였고 고주파는 DCT를 적용하였다.

이 논문에서는 유성음의 크기 (magnitude)만을 양자화 하였고, 성능평가는 스펙트럴 크기값에 대한 Spectral Distortion (SD) 을 사용하였고 제안된 방법이 더 낮은 SD를 나타내었다.

## II. 하모닉 코더

각각의 하모닉 요소의 주파수가 기본주파수의 정수배 일때 사인파형 또는 하모닉 요소들의 연속에 의해 신호가 생성된다면 이 신호는 하모닉이라 한다.

실제로 음성과 음악의 종류를 포함하는 많은 신호들은 하모닉 모델을 따른다. 정현파 모델은 기본주파수 (Fundamental Frequency), 크기 (Magnitude), 위상 (Phase)의 세가지 하모닉 성분의 사인파의 합으로 표현할 수 있다 [2]. 정현파 모델에 기초하여 여기신호 표현은 다음과 같다.

$$s(n) = \sum_{m=1}^L A_m \cos(w_m n + \psi_m) \quad (1)$$

여기서  $A_m$  과  $\psi_m$  은 주파수가  $w_m$  인 정현파 성분에 대한 크기와 위상을 나타내며,  $L$  은 정현파형의 개수를 나타낸다. 하모닉 성분만으로도 유성음 구간의 여기신호는 대부분의 음성 신호를 표현할 수 있으므로, 스펙트럴 기본 모델을 이용하여 근사화가 가능하다. 그러므로 식 (2)

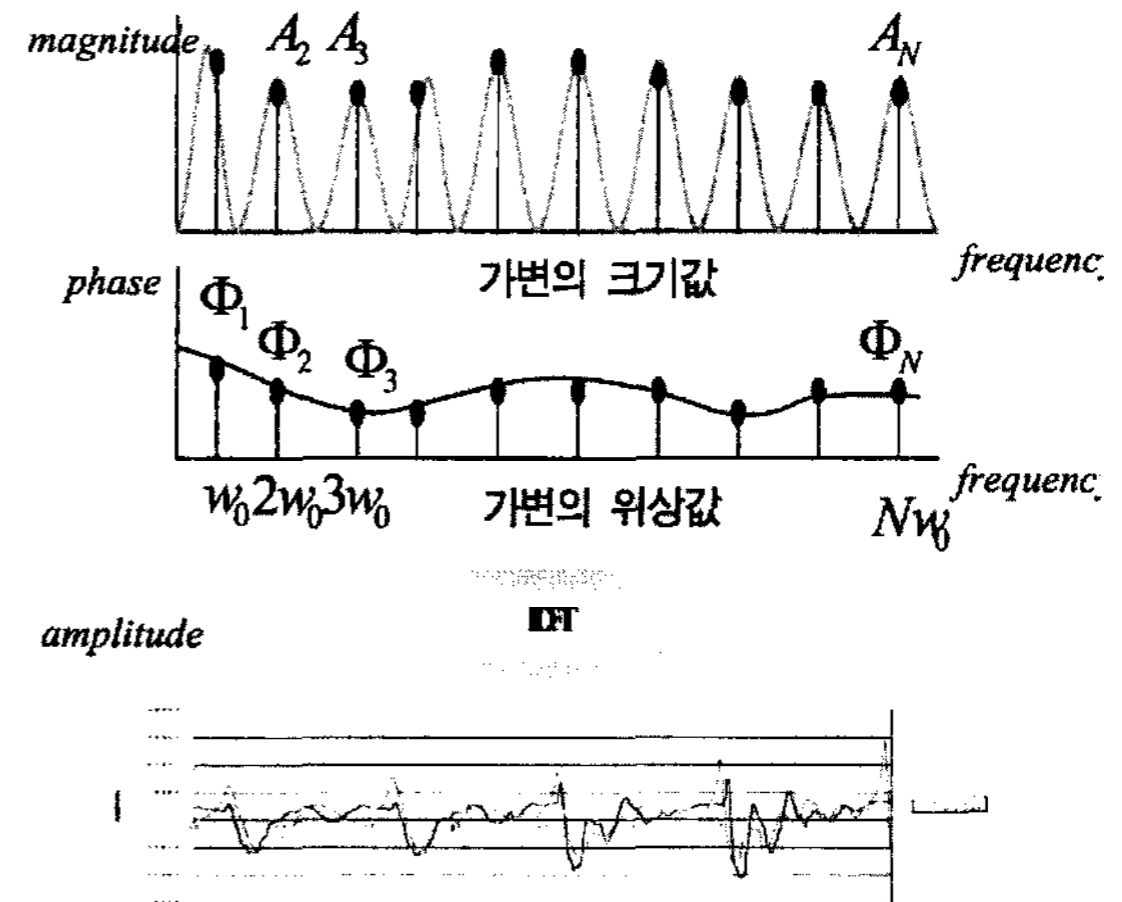


그림 1. magnitude 모델 파라미터 추출 및 합성파형  
Figure 1. Extraction of magnitude and Synthesized signal.

와 같이 측정 위상을 선형 위상 합성으로 대체하는 근사 모델을 사용할 수 있다.

$$s^k(n) = \sum_{m=1}^{L_k} A_m^k \cos(mw_0^k n + \psi^k(m, w_0^k, n) + \Phi_m^k) \quad (2)$$

여기서  $k$ 와  $L_k$ 는 프레임 번호와 각 프레임당 하모닉 개수를 나타내며,  $w_0$ 는 피치 각주파수를,  $w_0^k$ 는  $k$ 번째 프레임의 각주파수를,  $\Phi_m^k$ 는  $k$ 번째 프레임,  $m$ 번째 하모닉의 이산위상을 나타낸다. 위와는 달리 선형위상인  $\psi^k(m, w_0^k, n)$ 는 복호기에서 합성되는 정보로써, 다음과 같이 선형위상을 합성할 수 있다.

$$\psi^k(m, w_0, n) = \psi^{k-1}(m, w_0^{k-1}, n) + \frac{(w_0^{k-1} + w_0^k)}{2} n \quad (3)$$

위의 식에서 알 수 있듯이, 선형 위상은 이전 프레임과 현재 프레임의 시간에 따른 피치 각주파수를 선형보간하여 얻어진다. 인간의 청각 시스템이 위상 연속성이 보존되는 동안 선형 위상에 대해 감각적이지 못하며, 부정확하거나 혹은 완전히 판이한 이산 위상을 허용할 수 있다는 지각적 특성은 저 전송률 코딩에 있어 하모닉 모델의 연속성에 대한 중요한 조건이 된다. 이에 따라, 합성 위상은 측정된 위상을 대체할 수 있게 된다.

## III. 매칭 퍼슈잇을 이용한 하모닉 주파수에서의 크기 (Magnitude)값 검색

정현파 모델에서 신호는 정현파 성분들의 선형합으로

다음과 같이 정의된다.

$$s[n] \approx \hat{s}[n] = \sum_{m=0}^M A_m^k \cdot \cos(w_m^k \cdot n + \phi_m^k) \quad (4)$$

여기서  $A_m^k, w_m^k, \phi_m^k$ 는 k번째 프레임에서 m 번째 진폭, 주파수, 위상을 나타낸다.

정현파 파라미터의 예측을 위한 매칭 퍼슈잇 방법은 오류상쇄 (error concealment) 원리에 바탕을 둔다. 식 (5)는 매칭 퍼슈잇의 반복과정을 위한 왜곡 측정 함수이다.

$$MSE(A_m^k, \theta_m^k) = \sum_{n=1}^L (s_{m-1}^k(n) - A_m^k \cos(w_m^k n + \phi_m^k))^2 \quad (5)$$

여기서  $s_{m-1}^k(n)$ 은 k번째 프레임에서 m-1 번째 반복단계의 목적신호이고, 가중치 함수  $w(n)$ 은 해밍윈도우를 사용한다.

$$E_m^k = \sum_{n=0}^{N-1} [w(n) \{s_{m-1}^k(n) - A_m^k \cos(w_m^k n + \phi_m^k)\}]^2$$

$$= \sum_{n=0}^{N-1} [w(n) \{s_{m-1}^k(n) - A_m^k \cos(w_m^k n) \cos(\phi_m^k) + A_m^k \sin(w_m^k n) \sin(\phi_m^k)\}]^2 \quad (6)$$

where  $s_m^k(n) = s_{m-1}^k(n) - A_m^k \cos(w_m^k n + \phi_m^k)$

식 (6)에서  $w_m^k$ 를 기본 주파수의 배수로 고정시키면,  $w_m^k$ 은  $m \cdot w_0^k$ 이다. m은 정수값이고, 최대값은 전체 하모닉의 개수이고 피치주기를 2로 나누어 계산된다. 즉 파라미터 예측과정은 하모닉 개수만큼 반복되게 된다. 왜곡 측정 함수를 최소화하는 해를 구하면 스펙트럼 크기와 초기 위상을 아래 식과 같이 구할 수 있다.

$$A_m^k = \sqrt{(a_m^k)^2 + (b_m^k)^2}, \quad \phi_m^k = -\tan^{-1}\left(\frac{b_m^k}{a_m^k}\right) \quad (7)$$

where  $a_m^k = A_m^k \cos(\phi_m^k), b_m^k = -A_m^k \sin(\phi_m^k)$

일반적으로 파라미터 예측에 사용되는 스펙트럼 피크 검출방법보다 주파수 해상도가 높다. 매칭 퍼슈잇 방법은 윈도링 후에 512 포인트 FFT를 취하는 방법에 비해 뛰어난 성능을 나타내고, 2048 포인트 FFT를 취하는 방법과 비슷한 해상도를 갖는다 [7][8].

본 논문에서는 매칭 퍼슈잇을 이용하여 하모닉 주파수에서의 magnitude값을 추출하여 직선과 DCT를 적용하여 양자화 계수를 추출한다.

## IV. 스펙트럼 포락선을 직선으로 모델링 & 가변 DCT 적용하여 양자화 계수 추출의 결합된 방법

### 4.1. 주파수 분할

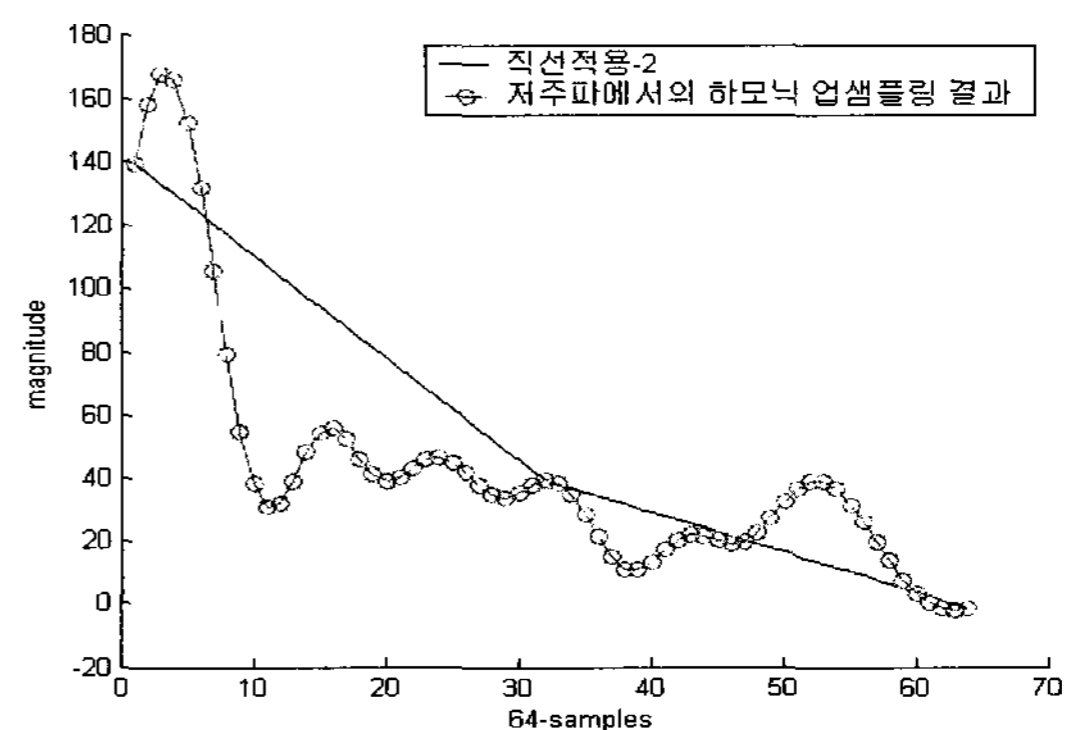
주파수 영역을  $\pi/2$ 를 중심으로  $\pi/2$ 이전까지의 하모닉과  $\pi/2$ 이후의 하모닉을 나누어서 양자화 계수를 추출한다.  $\pi/2$ 이전의 하모닉을 저주파 하모닉이라 하고  $\pi/2$ 이후의 하모닉을 고주파 하모닉으로 간주하고 저주파에는 업샘플링과 직선을 사용하고 고주파에는 DCT를 적용한다.

### 4.2. 스펙트럼 포락선을 직선으로 모델링

매칭 퍼슈잇을 이용하여 추출한 하모닉 크기 (magnitude)를 업샘플링하여 64개로 차원을 고정한다. 업샘플링 방법은 HVXC에서 사용한 Linear Interpolation을 이용한 오버샘플링 방법을 이용하였다. HVXC에서는 한프레임에 대해 44개로 차원을 고정하였으므로  $\pi/2$ 이하의 이것의 절반이 22개의 업샘플이 적용되었다고 간주한다.

따라서  $\pi/2$ 이하의 하모닉에 대해 22개보다 많은 업샘플링은 하모닉에 대한 더 많은 정보를 갖는다. 이는 업샘플링을 더 많이 적용할 경우 왜곡이 줄어드는 것으로 확인 할 수 있었다.

본 논문에서는 64개로 차원을 고정하였고 스펙트럼 포락선을 직선으로 대체하여 4의 배수가 되는 지점 (직선이 연결되는 지점)의 샘플값만을 양자화 계수로 사용한다. 고정된 차원의 처음과 끝의 크기와 주파수값을 알고 있으므로 그것을 이용하여 직선의 방정식을 적용하여 직선을 구한다. 그림 2에서 처럼 하나의 직선을 두 개로 나누는 과정을 반복하여 2의 배수로 직선이 많아 질수록 그림 2와 같이 고정된 차원의 스펙트럼과 점차 가까워짐을 알 수 있다. 이 때 직선의 개수는 1,2,4,8,16으로 증가하여



(a)

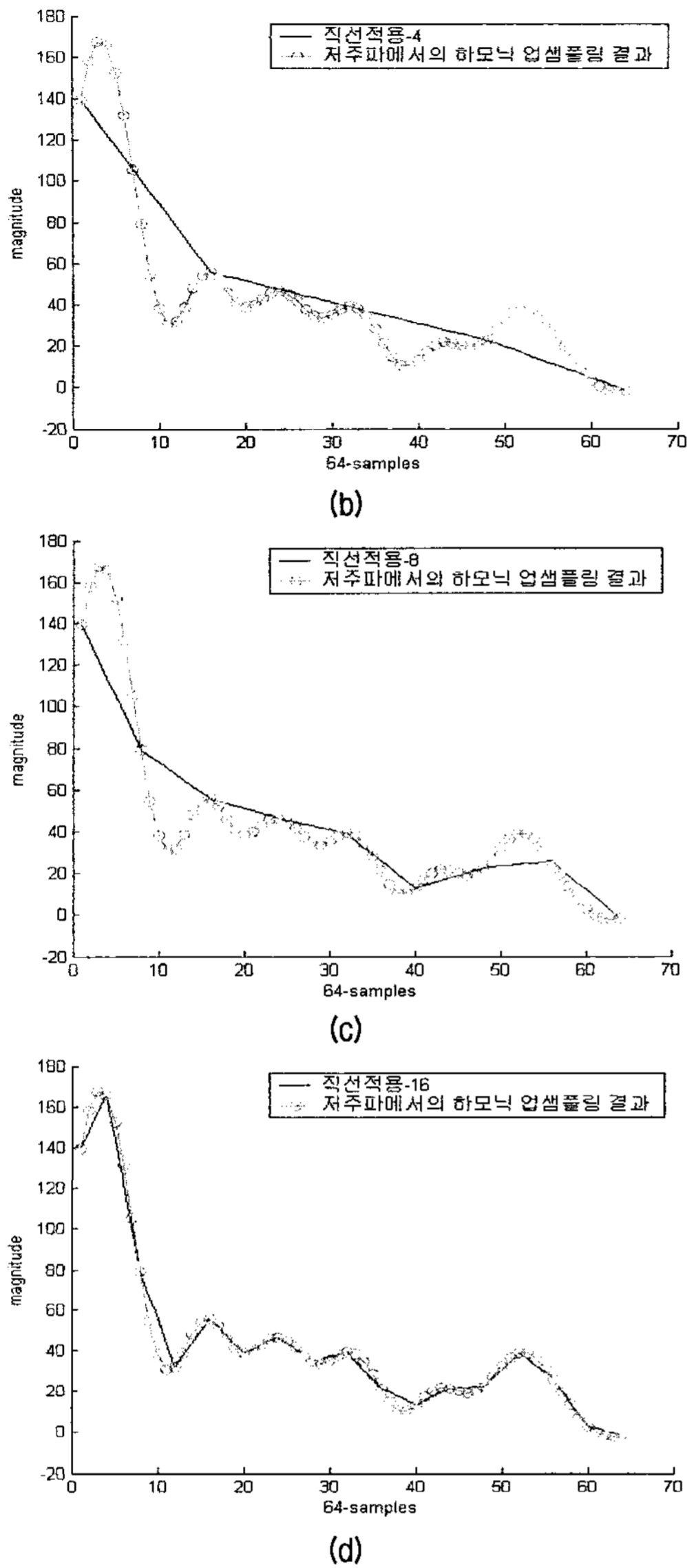


그림 2. 저주파 하모닉 업샘플링 후 직선 2의 배수로 증가하면서 스펙트럼 포락선 표현 (a)직선 2개 (b)직선 4개 (c)직선 8개 (d)직선 16개

Figure 2. Representation of spectral envelope in low frequency band (a) two straight-line (b) four straight-line (c) eight straight-line (d) sixteen straight-line.

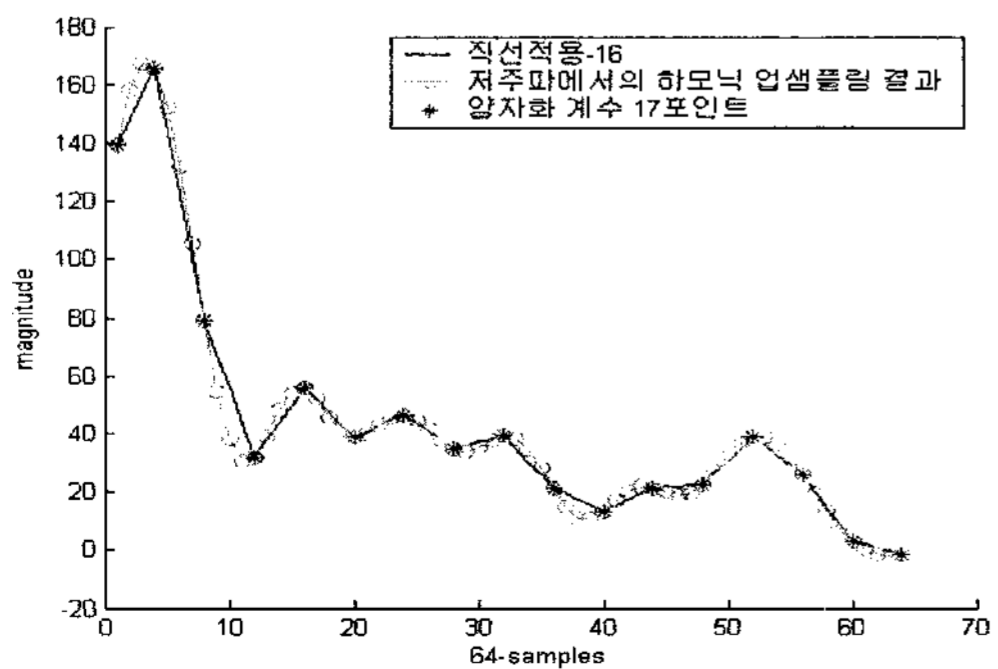


그림 3. 직선 적용시 추출되는 양자화 계수 17개  
Figure 3. The seventeen quantization parameters extracted from straight-lines.

스펙트럼 포락선을 표현하게 된다 [9].

직선방정식을 디코더에서 다시 재구성하기 위해서 직선의 양쪽 크기와 주파수 값을 알아야 하므로 총 17개의 점을 전송한다. 이때 주파수는 64개로 일정하므로 계산으로 구할 수 있다. 그림 3에서 직선에서 표시된 점이 전송될 값이다.

### 4.3. NSTVQ (Nonsquare Transform Vector Quantization)

고주파의 하모닉은 DCT를 적용하였다. 가변적인 하모닉에 대해 가변 DCT를 사용함으로써 가변 스펙트럴 파라미터들의 에너지 집중률을 높이는 장점을 갖는다. 하모닉이 제한된 DCT계수보다 적을 경우 나머지를 '0'으로 채우는 NSTVQ (Nonsquare Transform Vector Quantization) 방법을 양자화에 적용하였다 [6].

가변 차원인 스펙트럴 크기를 양자화 함에 있어 계수들에 대한 비트 효율을 높이고자 NSTVQ를 사용하여 계수들을 양자화 한다. NSTVQ 방법은 다음과 같이 수식으로 표현할 수 있다.

$$\widehat{z}_{opt} = A_p^T \cdot \widehat{y} \tag{8}$$

$$A_p = \begin{cases} (\widehat{a}_1 \widehat{a}_2 \widehat{a}_3 \dots \widehat{a}_M) & \text{if } N \geq M \\ (\widehat{a}_1 \widehat{a}_2 \widehat{a}_3 \dots \widehat{a}_N | 0) & \text{if } N < M \end{cases} \tag{9}$$

여기서 N은 가변 차원 계수의 개수를, M은 NSTVQ에 사용될 제한된 계수의 개수를,  $\widehat{z}_{opt}$ 는 트레이닝을 위한 데이터를,  $\widehat{y}$ 는 스펙트럴 크기 값들을,  $A_p$ 는 가변 차원인  $\widehat{y}$ 의 에너지 집중률을 높이기 위한 변환 행렬을 뜻하며, 각 원소들은 직교 특성의 기본 함수값을 나타낸다. 위 식에서 볼 수 있듯이  $A_p$ 는 M개의 고정 차원의 행렬이며, 기본 변환 함수 계수값의 차원 N이 M보다 작을 경우 나머지 부분에 0을 채워주는 행렬이다. 여기서 N은 스펙트럴 크기 값의 차원을 뜻하며, M은 훈련시킬 데이터의 차원을 뜻한다. 우리가 사용한 스펙트럴 크기 값들은 각 프레임에서의 피치 주기의 1/2만큼 추출하여 고주파 부분의 하모닉만을 양자화 하였다. 이때 고주파 하모닉 계수들의 최대값은 10으로 제한한다.

여기서 행렬  $A_p$ 를 통해 변수들의 변환이 이루어지는데, 이때 이산 코사인 기본 행렬을 사용한다. 이산 코사인 변환은 변수의 에너지 집중효율을 높임으로써 변수 특성을 잘 표현하게 해준다.  $A_p$ 의 사용을 통해 스펙트럴 크기값들을 이산 코사인 변환 (DCT) 하는데 사용된 식은 다음과 같다.

$$S(k) = \sum_{n=0}^{N-1} s(n)\lambda(k)\cos\left[\frac{(2n+1)\pi k}{2N}\right] \quad (10)$$

$$\lambda(k) = \begin{cases} 1 & ; k=0 \\ \sqrt{2} & ; otherwise \end{cases} \quad (11)$$

여기서  $s(n)$ 은 입력 신호를,  $S(k)$ 는 이산 코사인 변환 계수를 나타낸다. NSTVQ 방법에서는 가변차원 계수 값들의 훈련에 사용되는  $k$ 차 코드북의 중심 값인  $\hat{c}_k$ 를 계산하는데 있어 다음과 같은 식을 사용한다.

$$\hat{c}_k[n] = \frac{\sum_{l \in S^*} \frac{1}{N_l} z_l[n] p_l[n]}{\sum_{l \in S^*} \frac{1}{N_l} p_l[n]} \quad (12)$$

여기서  $n$ 은 1에서  $M$ 의 값을 갖고,  $z_l[n]$ 은  $l$ 차 훈련 벡터의  $n$ 차 요소이며,  $p_l[n]$ 은 다음과 같이 정의된다 [4].

$$p_l[n] = \begin{cases} 1 & \text{if } n \leq \min(N, M) \\ 0 & \text{otherwise} \end{cases} \quad (13)$$

#### 4.4. 직선과 DCT를 이용하여 추출한 계수값에 대한 Vector Quantization

직선과 DCT를 적용하여 추출한 양자화 계수를 그림 4와 같은 구조의 양자화기를 사용하여 양자화한다. 표 1은 4 kbit/s의 제안된 코더에 대한 비트할당을 나타내었고, 표 2는 표 1에서 Spectral Shape, gain, Spectral Shape (harmonic)에 할당된 코드북의 비트수와 차원을 나타낸다. 차원은 27개 (직선17개+DCT10개)로 고정하여 cb1 코드북으로 양자화한 후 그 결과의 오차를 한 번 더 cb2 코드북으로 양자화한다.

코드북2 (cb2)까지의 양자화 결과를 우선 원래의 하모닉 차원으로 다운샘플링 한다. 그리고 원래의 하모닉과 2-stage 양자화 결과의 하모닉 차이값을 그림 4와 같이 추가 양자화기를 사용하여 양자화되며, 이 경우에는 오직 하모닉 15이하의 값에 대해서만 Split VQ를 이용하여 양자화한다. 코드북은 그림 4와 같이 Harmonic\_cb3, Harmonic\_cb4, Harmonic\_cb5, 그리고 Harmonic\_cb6이 사용된다.

그림 5는 양자화기에서 차원변경, 직선 그리고 DCT 적용부분에 대한 내부구조를 나타내었다. 저주파 하모닉은 오버샘플링 후 직선으로 스펙트럴 커브를 표현하고 4의 배수가 되는 지점의 스펙트럴 크기값을 전송한다. 고주파 하모닉은 DCT로 변환하여 그 계수를 전송한다. 디코더에서 양자화 후 계수값을 복원시에는 반대로 적용한다.

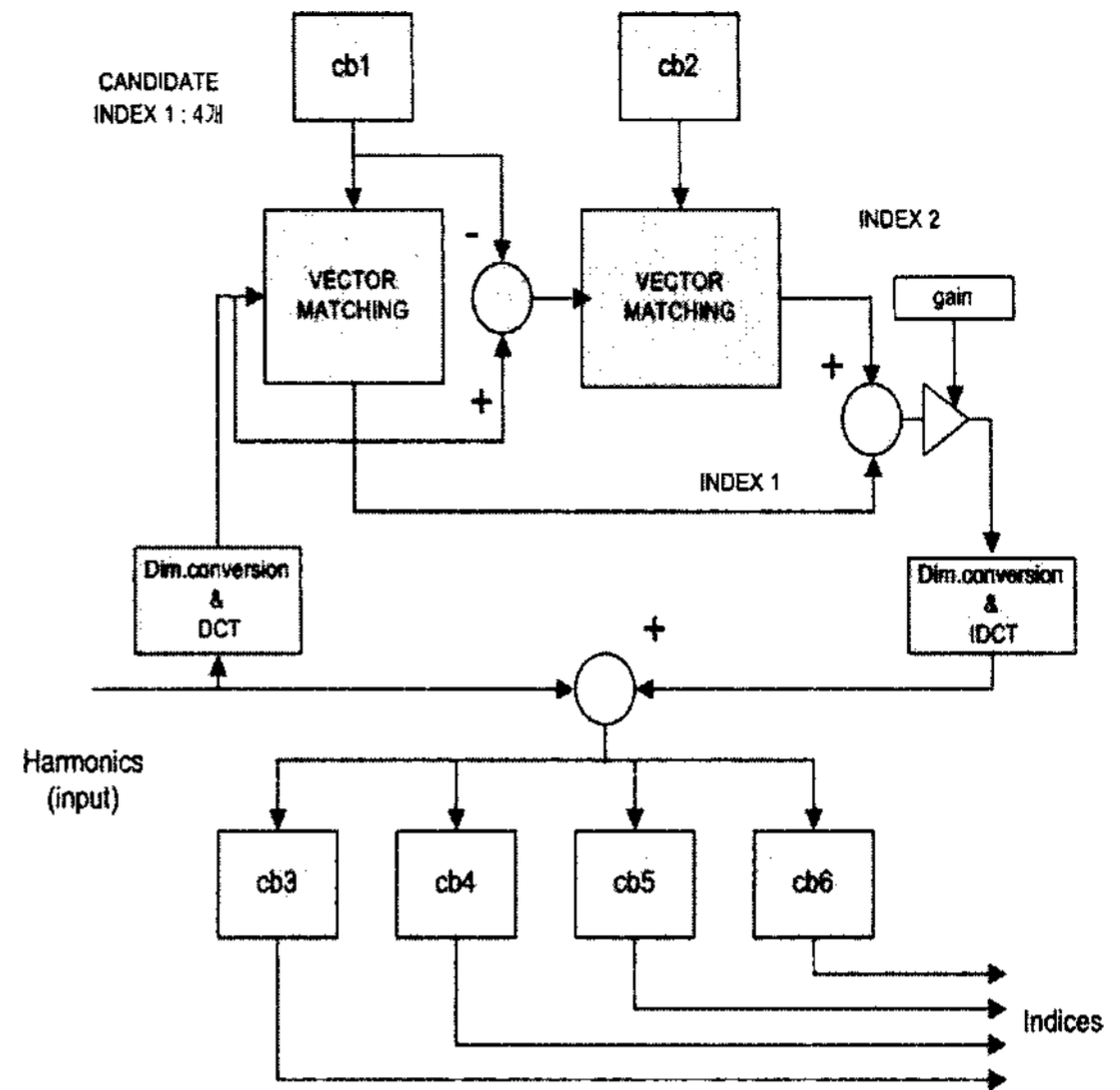


그림 4. 하모닉 스펙트럼 크기값의 양자화 구조  
Figure 4. The inside structure of quantization for harmonic spectrum magnitude.

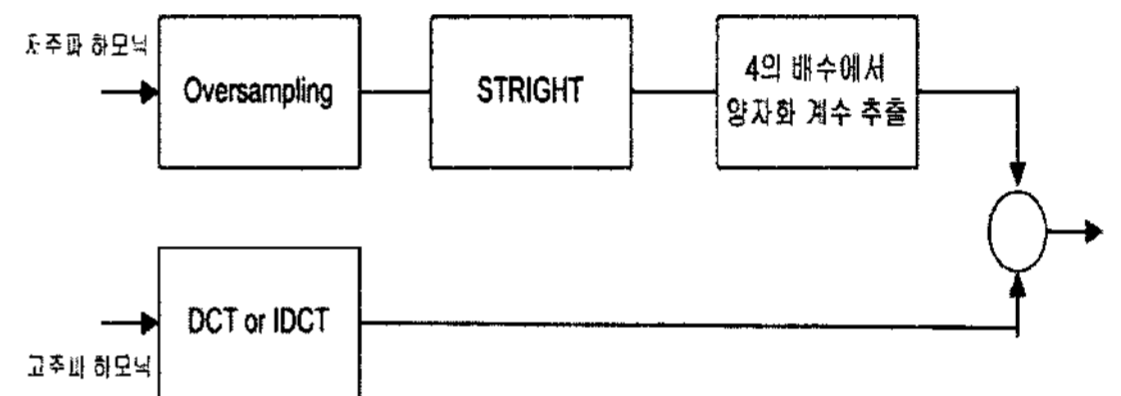


그림 5. 차원 변경 내부구조  
Figure 5. The inside structure of Dimension conversion.

표 1. 비트 할당 4 bits/s.  
Table 1. bit allocation 4 kbit/s.

	Voiced
LSP	24 bits/20ms
V/UV	2 bits/20ms
pitch	7 bits/20ms
Spectral shape	5 + 5 bits/20ms
gain	5 bits/20ms
Spectral shape (harmonic)	32 bits/20ms
Total	80 bits/20ms

표 2. 코드북 사이즈와 차원 (Spectral Shape).  
Table 2. Codebook Size and Dim (Spectral Shape).

Codebook	Dim.	Number of Bits
cb1	27	5
cb2	27	5
gain	1	5
Harmonic_cb3	2	5
Harmonic_cb4	4	10
Harmonic_cb5	4	10
Harmonic_cb6	5	7
Total		47

### V. 실험 결과

성능 평가는 다음 식을 이용하여 SD (Spectral Distortion)을 측정하였다 [3][4].

$$SD = \sqrt{\frac{1}{i_2 - i_1} \sum_{n=i_1}^{i_2-1} \left[ 10 \log_{10} \left( \frac{|m[n]|^2}{|m_q[n]|^2} \right) \right]^2}$$

여기서,  $i_1$ 과  $i_2$ 는 선택된 하모닉의 주파수 범위이며,  $m[n]$ 은 원래의 계수 값을,  $m_q[n]$ 은 양자화 되고 난 후의 계수 값을 의미한다. 실험 데이터는 한 프레임을 160샘플로 하고, 피치 주기의 1/2개 수 만큼 추출된 스펙트럴 크기값들을 사용하였다. 합성 후 Target 신호와 합성 신호의 스펙트럼을 비교하였다. 여기서 Target 신호는 양자화 하지 않은 신호를 합성한 신호이다.

스펙트럴 크기 양자화 기법의 성능 평가를 위해 ISO/IEC 국제 표준 HVXC 양자화 방법을 사용하였다. HVXC 음성 부호화기에서는 하모닉을 44개로 오버샘플링하여 고정하고 5비트 코드북으로 양자화한 후 그 오차를 5비트 코드북으로 한번 더 양자화 하는 2-stage 양자화방법이다. 여기에 추가 양자화 기를 사용하여 원래 하모닉 14개 이하의 양자화 에러값을 Split VQ를 이용하여 성능을 개선하는 방식이다.

이 논문에서 제안하는 양자화 기법에서는 비교할 HVXC와 비트할당의 형평성을 맞추기 위해 코드북에 사용된 비트수는 동일하며 2-stage 코드북과 추가 Split VQ 코드북의 차원수만 변경하였다 [1]. Split VQ의 코드북 차원과 비트수는 표 2에 나타내었다.

실험 결과 HVXC보다 제안한 양자화 기법의 SD가 0.3DB정도 개선되었다. 결과는 아래 표 3에 나타내었다. 그림 6은 HVXC와 제안된 양자화 기법을 사용하여 얻은 결과의 주파수 왜곡 측정값을 그래프로 나타내었다.

표 3. 4 Kbit/s HVXC와 제안된 양자화 기법의 주파수 왜곡 측정값  
Table 3. SD for 4 Kbit/s HVXC and Proposed method of Quantization (단위: dB).

bits \ Coder	HVXC	제안된 양자화 기법
cb1 (5)/cb2 (5) 4.0 Kbit/s	2.883	2.589
cb1 (6)/cb2 (6) 4.1 Kbit/s	2.766	2.492
cb1 (7)/cb2 (7) 4.2 Kbit/s	2.634	2.403
cb1 (8)/cb2 (8) 4.3 Kbit/s	2.489	2.287

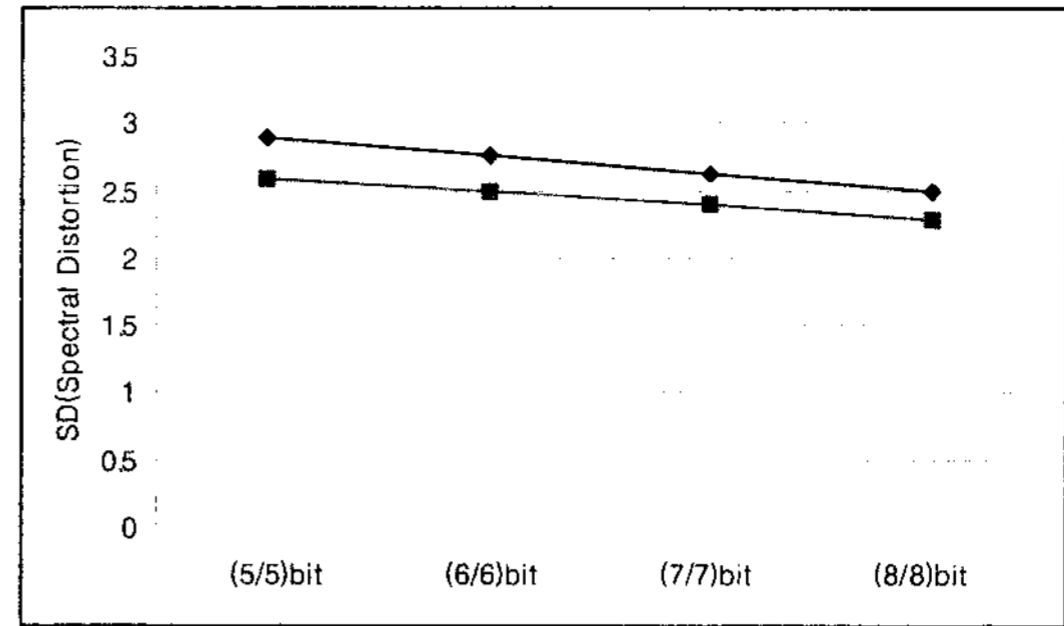


그림 6. 주파수 왜곡 측정값 (위)HVXC (아래)제안된 양자화기법  
Figure 6. Spectral Distortion (the upper part) HVXC (the lower part)the proposed method of quantization.

### VI. 결론

이 논문에서는 스펙트럼 크기값에 대한 효율적인 양자화 기법에 대해 제안하였다. 양자화 효율을 높이기 위해 주파수를 분할 하였고, 각각 직선과 DCT를 이용하여 양자화 계수를 추출하였다. 오버샘플까지는 HVXC와 유사한 방법이지만 이 논문에서는 저주파에서의 차원고정을 22에서 64로 늘리면서 더 많은 하모닉 정보를 포함시킬 수 있었으며 직선을 적용하므로써 64개 차원이 아닌 17개 차원만을 양자화 계수에 사용하였다. 또한 가변 DCT를 이용한 NSTVQ방식은 에너지 집중률을 높이고 가변차원에 대한 적합한 코드북 훈련 방법을 제공함으로써 비트효율을 증가시키므로 한 프레임을 44개로 고정시키는 HVXC보다 좋은 성능을 낼 수 있었다.

“이 논문은 2007년 정부 (교육인적자원부)의 재원으로 한국학술진흥재단의 지원을 받아 수행된 연구임” (지방연구중심대학육성사업/충북BIT연구중심대학육성사업단)

### 참고 문헌

1. M. Nishiguchi, "Weighted vector quantization of harmonic spectral magnitudes for very low-bit-rate speech coding," *Acoust. Sci. & Tech*, 27, 43-49, 2006.
2. Wai C. Chu, "Vector Quantization of Harmonic Magnitudes in Speech coding Applications-A Survey and New Technique," *EURASIP Journal on Applied Signal Processing* 2004:17, 2601-2613, 2004.
3. C. Li, P. Lupini, E. Shlomot, and V.Cuperman, "Coding of variable dimension speech spectral vectors using weighted nonsquare transform vector quantization," *IEEE Trans. Speech, and Audio Processing*, 9(6), 622-631,2001.
4. P.Lupini and V.Cuperman, "Nonsquare transform vector quantization," *IEEE Signal Processing Letters*, 3(1), 1-3, 1996.

5. E. Shlomot, V. Cuperman, A. Gersho, "Combined harmonic and waveform coding of speech at low bit rates", ICASSP' 98, 585-588.
6. 신 경진, 이 인성, "효율적인 가변차원 하모닉 크기 양자화기법", 한국 학회지 20(7) 47-54, 2001.
7. 안 영욱, 정 규혁, 김 종학, 양 용호, 이 인성, "정현파 모델 부호 화기를 위한 MP (Matching Pursuit) 알고리즘과 파라미터 양자 화기", 음향학회지 24(7), 402-409, 2005.
8. 정 규혁, 김 종학, 임 정우, 주 기호, 이 인성, "Damping 요소를 첨가한 매칭 퍼슈잇 정현파 모델링," 전자공학회 논문지 44(1), 105-113, 2007.
9. \* 0.8\* 1.2Vorbis I specification, Xiph.org Foundation, pp. 37-38, <http://xiph.org/vorbis/doc/>
10. A. V. McCree and T. P. Barnwell III, "A mixed excitation LPC vocoder model for low bit rate speech coding," IEEE Trans. Speech Audio Process., 3, 242-50, 1995.
11. W. B. Kleijn and J. Haagen, "A speech coder based on decomposition of characteristic waveforms," Proc. ICASSP 95, 1-508-11, 1995.
12. M. Nishiguchi, K. Iijima and J. Matsumoto, "Harmonic vector excitation coding of speech at 2.0 kbps," IEEE Workshop on Speech Coding, 1997.

---

## 저자 약력

---

### •최 지 욱 (Ji-Wook Choi)



2004년 2월: 충북대학교 전자공학과 (공학사)  
 2003년 12월~2004년 11월 유비랩(주)  
 2006년 3월~현재: 충북대학교 전자공학과 (공학석사)  
 ※주관심분야: 음성신호의 부호화, 음성 신호의 저장 및 전송

### •정 규 혁 (Gyu-Hyeok Jeong)



2004년 2월: 충북대학교 전기전자공학 (공학사)  
 2006년 2월: 충북대학교 전자공학과 (공학석사)  
 2006년 3월~현재: 충북대학교 전자공학과 (공학박사)  
 ※주관심분야: 음성/오디오 부호화, 신호처리, VoIP

### •이 인 성 (In-Sung Lee)



1983년 2월: 연세대학교 전자공학과 (공학사)  
 1985년 2월: 연세대학교 전자공학과 (공학석사)  
 1992년 2월: Texas A&M University 전기공학과 (공학박사)  
 1993년 2월~1995년 9월: 한국전자 통신연구원 이 동통신 기술연구단 선임연구원  
 1995년10월~ 현재: 충북대학교 전기전자공학부 정 교수  
 ※주관심분야: 음성/영상 신호 압축, 이동통신, 적응필터