

LPC계수의 예측 피라미드 벡터 양자화 기술

정회원 조인환*, 강상원*

Predictive Pyramid Vector Quantization of LPC Parameters

Inwhan Cho*, Sangwon Kang* *Regular Members*

요 약

본 논문은 LPC계수의 효율적인 양자화기에 대한 연구로서, LSF 계수의 프레임간 높은 상관도와 예측 에러 신호의 기하학적 성질을 이용한 predictive pyramid vector quantization(PPVQ) 방식을 제안한다. LSF 계수의 예측을 위해 weighted distortion measure를 이용하여 설계한 4차 moving average(MA) 예측기를 사용하였고, Laplacian 분포를 가지는 예측 에러 신호의 효율적인 양자화를 위해 pyramid VQ(PVQ)를 이용하였다. PPVQ는 적은 계산량으로 23 bits/frame에서 transparent한 성능을 가진다. 또한 PPVQ는 계산량이 적으며 코드북 저장을 위한 메모리를 전혀 필요치 않는다.

ABSTRACT

In this paper, the predictive pyramid VQ(PPVQ) is proposed to quantize LSF parameters. The properties of the LSF parameters, whose prediction error signal is Laplacian and interframe correlation is high, are exploited. To predict the LSF parameters, 4th order moving average (MA) predictor is used, and to quantize the prediction error signal, the pyramid VQ (PVQ) is exploited. It is verified that the performance of PPVQ has the transparent quality at 23 bits/frame. PPVQ requires low complexity and does not need any memories for the codebook storage.

I. 서 론

고음질의 저속 음성 부호화기를 구현하기 위해서는 음성신호의 단구간(short term) 상관도를 나타내는 LPC 계수를 효율적으로 양자화하는 것이 매우 중요하다. LPC 계수는 단구간 합성필터의 안정도 검사가 어려우며 dynamic range가 커서 양자화에 적합하지 않기 때문에 line spectral frequencies (LSF) 계수^[1]로 변환하여 이를 양자화한다. 효율적인 음성 부호화기에 적합한 LSF 양자화는 적어도 프레임당 24 비트 이하에서 transparent한 성능을 가져야 한다.

LSF 계수의 양자화 방식은 크게 스칼라 양자화 (scalar quantization) 방식과 벡터 양자화 (vector quantization)방식이 있다. 스칼라 양자화 방식은 이

미 많은 연구가 수행되어 32~40 비트에서 평균 SD가 1 dB이하인 양자화기들이 많이 발표되었다^[2]. 그러나 30 비트 이하에서는 성능이 우수하지 못하므로 효율적인 저속 음성 부호화기의 LSF 양자화 방식으로는 적합하지 못하다. 벡터 양자화 방식으로는 Linde, Buzo, Gray(LBG) 알고리즘^[3]을 기초로 한 다양한 LSF 벡터 양자화 방식들이 제안되었다^[4,5]. 그러나 LBG 알고리즘을 기초로 한 LSF 양자화 방식은 벡터의 차수의 증가에 따라 코드북 탐색 복잡도가 지수적으로 증가하는 구조적인 단점을 가진다. 즉 LSF 파라미터와 같은 10차 이상의 벡터를 양자화 하기 위해서는 많은 양의 메모리와 계산량이 요구된다. 이러한 이유로 입력 벡터를 split하여 성능을 최적화하는 split VQ(SVQ)방식^[4,5]이나, multi-stage VQ(MSVQ)방식^[6] 등이 제안되었다. 하지만

* 한양대학교 제어계측공학과(swkang@selab.hanyang.ac.kr)

논문번호 : 98351-0811, 접수일자 : 1998년 8월 11일

※ 이 논문은 1996년도 한국학술진흥재단의 공모과제 연구비에 의해서 연구되었음

이러한 방식들은 입력 벡터 자체를 나누어서 양자화하거나 입력 벡터를 다단계로 나누어서 양자화함으로써 나누어진 벡터들간의 상관도를 이용하지 못하기 때문에 입력 벡터 전체를 벡터 양자화 했을 때에 비해 성능이 떨어진다.

이러한 단점을 극복하기 위해 본 장에서는 LSF 계수의 interframe 상관도를 이용하여 입력 LSF 벡터를 예측 에러 신호에 근거한 MA 예측기^[7]로 예측하고, 입력 벡터와 예측 벡터간의 차이 벡터를 PVQ로 양자화하는 predictive pyramid VQ(PPVQ) 방식을 제안한다.

II장에서는 LSF 계수 분석에 대해서 기술하였고, III장에서는 PVQ에 대해서, IV장에서는 MA 예측기에 대해 개략적으로 설명하였다. V장에서는 설계한 PPVQ의 구조에 대해서 기술하였고, VI장에서는 실험 및 결과에 대해 다루었다.

II. LSF 계수의 분석

저속에서 고 음질의 음성 부호화를 위해 분석/합성 구조를 갖는 혼합 부호화 방식이 많이 사용되는데, 이러한 방식들은 매 전송 프레임(약20msec)마다 음성의 포먼트 성분을 나타내는 LPC (linear predictive coding) 계수를 전송한다. 이러한 LPC 계수에 의한 LPC 합성필터의 전달 함수는 다음과 같이 주어진다.

$$H(z) = \frac{1}{A(z)} = \frac{1}{1 - a_1z^{-1} - a_2z^{-2} - \dots - a_pz^{-p}} \quad (1)$$

여기서, $\{a_i\}$ 는 LPC 계수이며, p 는 LPC 계수의 차수이다.

LPC 계수는 흔히 LPC 계수와 수학적으로 동가 이면서 양자화 특성이 좋고 합성필터의 안정도를 검사하기 쉬운 LSF 계수로 변환된 후 양자화 된다. LPC 계수로부터 LSF 계수를 구하기 위해서는 분석필터 $A(z)$ 의 회귀 방정식을 이용하여 다음과 같은 2개의 다항식을 만든다.

$$P(z) = A(z) - z^{-(p+1)}A(z^{-1}) \\ = (1 - z^{-1}) \prod_{i=2,4,\dots,p} (z - 2z^{-1} \cos \omega_i + z^{-2}) \quad (2)$$

$$Q(z) = A(z) + z^{-(p+1)}A(z^{-1}) \\ = (1 + z^{-1}) \prod_{i=1,3,\dots,(p-1)} (z - 2z^{-1} \cos \omega_i + z^{-2}) \quad (3)$$

여기서, $e^{j\omega}$ 는 $P(z)$ 와 $Q(z)$ 의 근이며, ω_i 는 LSF 계수값이다.

위의 $P(z)$ 와 $Q(z)$ 다항식은 다음과 같은 중요한 특성을 가진다.

- 1) $P(z)$ 와 $Q(z)$ 의 모든 근들은 단위원 상에 존재한다.
- 2) $P(z)$ 와 $Q(z)$ 의 근들은 서로 interlace된다. 위의 두가지 특성에 근거해서 LSF 계수의 몇가지 중요한 성질들을 살펴보면 다음과 같다.

$$0 < \omega_1 < \omega_2 < \dots < \omega_p < \pi \quad (4)$$

- 2) LSF 계수는 음성의 포먼트 주파수와 포먼트 대역폭을 나타낸다. 즉, LSF 계수가 가까이 위치할수록 더욱 좁은 포먼트 영역을 나타낸다.
- 3) 보간된 LSF 계수도 순서성질을 만족한다.
- 4) LSF 계수는 dynamic range가 적어 양자화에 유리하다.

LSF 양자화기의 성능평가를 위해 다음과 같이 표현되는 spectral distortion(SD)^[8]을 사용한다.

$$SD = \left[\frac{1}{M} \sum_{n=1}^M \left(\frac{1}{\pi} \int_0^{\pi} (\log S_n(\omega) - \log \hat{S}_n(\omega))^2 d\omega \right) \right]^{\frac{1}{2}} \text{ (dB)} \quad (5)$$

여기서, $S_n(\omega)$ 와 $\hat{S}_n(\omega)$ 는 각각 원음의 LPC 계수에 의한 스펙트럼과 양자화된 LPC 계수에 의한 스펙트럼을 나타내며, M 은 SD계산에 사용된 전체 프레임 수이다. Transparent한 성능의 LSF양자화기란 평균 SD가 1dB이하이며, 2~4 dB를 갖는 프레임의 개수가 전체 프레임 수의 2% 이내이고, 또한 4 dB 이상의 SD가 0% 로서 사람의 귀로 LSF 계수의 양자화 에러 영향을 거의 감지할 수 없는 수준의 양자화기를 의미한다.

III. Pyramid Vector Quantization (PVQ)

1. PVQ 개요

PVQ는 96년Fischer^[9]에 의해 제안된 양자화기로서 피라미드의 표면상에 존재하는 큐빅 격자점들을 코드워드로서 사용하는 일종의 격자 양자화 방식이다. 따라서 PVQ는 코드북을 저장할 필요가 없으므로

메모리가 요구되지 않으며, 벡터 차수가 증가할수록 인코딩의 복잡도가 선형적으로 증가하는 특성을 가지고 있으므로, 적은 계산량으로 입력 벡터 전체를 양자화할 수 있다. 특히 memoryless Laplacian 소스에 대한 PVQ 양자화의 성능은 입력 벡터 차수가 클 경우, 엔트로피 제한 스칼라 양자화와 비슷한 성능을 나타낸다^[10]. 표 1은 memoryless Laplacian 소스에 대한 PVQ의 성능을 Lloyd-Max 스칼라 양자화와 비교한 것이다.

표 1. Memoryless Laplacian 소스에 대한 SNR 인코딩 성능 비교

Rate (bits)	Lloyd-Max SQ (dB)	PVQ (dB) (dim.=10)
1	3.01	5.09
2	7.54	10.32
3	12.64	16.12

K 를 양의 정수라 할 때, 다음과 같은 피라미드 $S(L, K)$ 는 표면에 코드워드 이용되는 많은 정수 큐빅 격자점을 가진다.

$$S(L, K) = \left\{ c(n) : \sum_{i=1}^L |c_i(n)| = K \right\} \quad (6)$$

그림 1은 차수가 3이고 K 가 4인 피라미드 $S(3, 4)$ 의 기하학적인 형태를 나타낸 것이다. 이러한 피라미드의 표면상의 정수 격자점의 좌표값들이 PVQ에서 코드워드 사용된다.

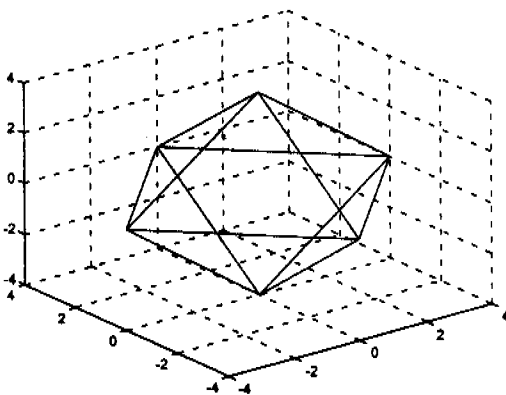


그림 1. 피라미드 $S(3,4)$ 의 격자 형태

만약 피라미드 $S(L, K)$ 상에 있는 전체 코드워드의 수를 $N(L, K)$ 라 한다면, K 는 다음 조건을 만족하는 가장 큰 정수로 선택된다.

$$N(L, K) \leq 2^{RL} \quad (7)$$

여기서,

$$N(L, K) = N(L-1, K) + N(L-1, K-1) + N(L, K-1) \quad (8)$$

이며, R 은 벡터 차수당 평균 비트율이다.

양자화기 입력 벡터가 Laplacian 분포를 가질 때 최적 코드북은 다음과 같이 표현되는 단일 피라미드상의 큐빅 격자점으로 설계될 수 있다^[9].

$$S(L, L/\lambda) = \left\{ c(n) : \sum_{i=1}^L |c_i(n)| = L/\lambda \right\} \quad (9)$$

여기서 $\lambda = 1/E[|c(n)|]$ 이며, L 은 양자화기 입력 벡터 차수이다.

양자화기 입력 벡터에 대한 PVQ 인코딩 알고리즘은 다음과 같다.

과정 1: 양자화기 입력 벡터 $c(n)$ 에 대해서, 피라미드 $S(L, L/\lambda)$ 에 가장 근접한 $\hat{c}(n) \in S(L, L/\lambda)$ 을 선택한다.

과정 2: $\hat{c}(n)$ 을 $K/\|\hat{c}(n)\|_1$ 로 scaling하고, scaling된 벡터 $\hat{v}(n)$ 이 피라미드 $S(L, K)$ 상에 존재하도록 한다.

과정 3: $\hat{v}(n)$ 과 가장 근접한 정수값을 갖는 최적 코드워드 $\hat{v}(n) \in S(L, K)$ 을 찾는다. 이것은 다음과 같은 방법으로 수행된다.

- $\hat{v}(n)$ 을 반올림하여 v 구한다.
- $\|\hat{v}(n)\|_1$ 을 계산한다. 만약 $\|\hat{v}(n)\|_1 = K$ 이면 이때의 $\hat{v}(n)$ 이 $\hat{v}(n)$ 의 최적 코드워드 사용되며, $\|\hat{v}(n)\|_1 > K$ 이면 올림한 요소값들 중에서 가장 오차가 큰 $(\|\hat{v}(n)\|_1 - K)$ 개의 요소값들을 하나씩 감소시키고, $\|\hat{v}(n)\|_1 < K$ 이면 내림한 요소값들 중에서 가장 오차가 큰 $(K - \|\hat{v}(n)\|_1)$ 개의 요소값들을 하나씩 증가시켜서 $\|\hat{v}(n)\|_1 = K$ 을 만족시킬때 $\hat{v}(n)$ 을 $\hat{v}(n)$ 의 최적 코드워드 사용한다.

입력 벡터 $c(n)$ 의 양자화값 $\hat{c}(n)$ 은 $\hat{v}(n)$ 의 최적 코드워드 $\hat{v}(n)$ 을 $\|\hat{c}(n)\|_1/K$ 로 scaling하여 구한다.

2. Product code PVQ (PCPVQ)

PVQ는 입력 벡터의 차수가 충분히 클 때 우수한 성능을 나타낸다. 즉 입력 신호가 Laplacian 분포를 가지고 벡터 차수가 약 20 이상일 때, 프레임에 상

관없이 입력 벡터의 l_1 norm값($= \|c(n)\|_1$)이 상수값에 가까워진다. 그러나 20 이하일 때에는 $\|c(n)\|_1$ 을 상수값으로 볼 수 없는 단점을 가진다. 그렇기 때문에 단일 피라미드를 사용하여 양자화 할 때 많은 오차가 발생하게 된다. 이러한 단점을 보완하기 위하여 product code PVQ^[9]가 함께 제안되었다.

Product code PVQ는 양자화기 입력 벡터를 normalize하고, normalize된 벡터를 단일 피라미드를 사용하여 양자화 한후, normalization 인자 $\hat{r} = Q(\|c(n)\|_1)$ 을 사용하여 양자화된 피라미드를 indexing하는 방식이다. 여기서 $Q(\cdot)$ 은 스칼라 양자화기를 의미한다. 만약 $\hat{c}(n) = PVQ(\hat{v}(n))$ 이 PVQ의 출력 벡터이고 $\hat{r} = Q(\|\hat{c}(n)\|_1)$ 이 스칼라 양자화기의 출력값이라면, product code PVQ의 출력 벡터 $\hat{c}_{PCPVQ}(n)$ 은 다음과 같이 구할 수 있다.

$$\hat{c}_{PCPVQ}(n) = \hat{r} \times \hat{c}(n) \quad (10)$$

이렇게 함으로써 스칼라 양자화기의 양자화 레벨수 만큼의 피라미드를 사용하는 효과를 얻을 수 있다. PVQ의 평균 벡터 차수당 비트율을 R_p , 스칼라 양자화기에 할당된 비트를 R_r 이라 하면 전체 비트율 R 은 다음을 만족한다.

$$R_p L + R_r = RL \quad (11)$$

IV. Predictive PVQ (PPVQ)

그림 2는 PPVQ의 구조를 나타내었다. MA 예측기 블록에서는 MA 벡터 예측기가 사용되었다. $f(n)$ 은 n번째 프레임의 10차 LSF 입력 벡터이며, $x(n)$ 은 벡터 $f(n)$ 으로부터 평균 벡터 \underline{f} 를 제거한 벡터이다.

예측 에러 벡터 $c(n)$ 은 다음 식과 같이 벡터 $x(n)$ 과 예측 벡터 $p(n)$ 과의 차이신호에 해당한다.

$$c(n) = x(n) - p(n) \quad (12)$$

예측 벡터 $p(n)$ 은 다음식과 같이 양자화된 예측 에러 벡터 $\hat{c}(n)$ 을 이용한 MA 예측기를 사용해서 구한다.

$$p(n) = \sum_{j=1}^M g_j \hat{c}(n-j) \quad (13)$$

여기서 $\hat{c}(k)$ 는 k번째 프레임에서 양자화된 예측

에러 벡터이며, g_j 는 j번째 LSF 벡터에 사용되는 MA 예측기의 계수 벡터, M은 사용된 MA 예측기의 차수이다.

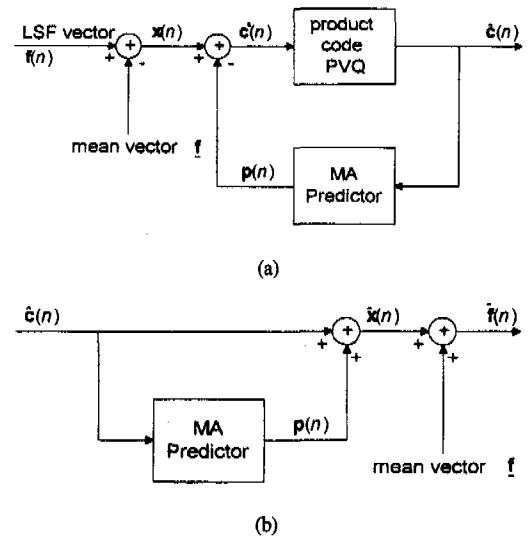


그림 2. Predictive PVQ의 구조 (a) 부호화기 (b) 복호화기

본 논문에서 사용된 MA 예측기는 예측에 이용되는 입력값이 예측기의 차수로 한정되어 있으므로 채널 에러에 효율적인 특성을 가진다. MA 예측기는 차수가 증가할수록, 예측 에러값은 감소하지만 늘어난 차수만큼 채널 에러에 영향을 많이 받게 된다. 그래서, 채널 에러에 강인하면서 적당한 왜곡값을 갖도록 MA 예측기 차수를 결정하는 것이 중요하다. 최적 MA 예측기 계수를 구하기 위해 Lloyd 알고리즘^[11]을 사용하여 training한다.

MA 예측기 계수를 구하기 위한 과정은 다음과 같다.

과정 1 : 초기 MA 예측기 계수 벡터와 training 음성 데이터를 이용하여 양자화된 예측 에러 벡터를 구한다.

과정 2 : 다음과 같은 가중 자승 왜곡 함수 D_w 를 최소화하는 MA 예측기 계수를 구한다.

$$D_w = \sum_{i=1}^N \sum_{n=1}^N w_i(n) \cdot (f_i(n) - \hat{f}_i(n))^2 \quad (14)$$

여기서, N은 training에 사용되는 데이터의 프레임 수이고, $w_i(n)$ 은 가중 값으로서 다음과 같다.

$$w_i(n) = \frac{1}{f_i(n) - f_{i-1}(n)} + \frac{1}{f_{i+1}(n) - f_i(n)}, \quad i = 1, \dots, 10$$

where $f_0(n) = 0, f_{11}(n) = 0.5$

(15)

과정 3 : 구해진 MA 예측기 계수 g_{ij} 을 사용하여 왜곡 감쇄율을 구한후 다음 조건을 만족하면 과정을 멈추고 이때의 계수 g_{ij} 를 최적 MA 예측기 계수로 결정하고, 그렇지 않을 경우에는 과정2로 돌아간다.

$$\left| \frac{D_w(\text{previous}) - D_w(\text{current})}{D_w(\text{current})} \right| \leq \epsilon = 0.001 \quad (16)$$

MA 예측기 차수에 따른 시뮬레이션 결과를 표 2에 나타내었다. MA 예측기 계수의 설계를 위해 남녀화자에 의해 발음된 20분 동안의 음성 신호를 사용하였다. 앞에서 서술한 바와 같이 MA 예측기는 차수가 증가할수록 적은 오차값을 나타내는 것을 실험을 통해 확인하였다. 그러나 MA 예측기 차수가 증가할수록 채널 에러에 대한 영향을 많이 받는 것을 고려하여, 본 논문에서는 평균 SD값이 포화하는 4차MA 예측기 계수를 선택하였다.

표 2. MA 예측기 차수에 따른 성능비교

MA 예측기 차수	평균 SD (dB)	% of Outlier	
		> 2 dB	> 4 dB
1	1.35	9.91	0
2	1.18	6.61	0
3	1.09	4.95	0
4	1.02	4.13	0
5	0.99	3.31	0
6	0.97	2.48	0

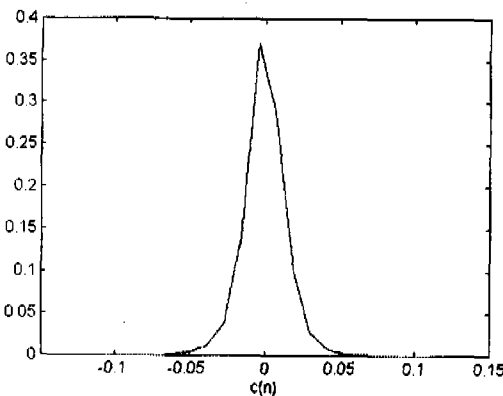


그림 3. 예측 에러 신호의 분포도

설계된 4차MA 예측기 계수를 표3에 나타내었다.

그림 3에서와 같이 예측 에러 벡터 $c(n)$ 은 통계적으로 Laplacian분포를 가지며, 양자화기 입력 벡

터의 차수가 20이하이므로 product code PVQ를 벡터의 효율적인 양자화기로 이용할 수 있다.

Product code PVQ에서는 예측 에러 벡터의 l_1 norm값($= \|c(n)\|_1$)의 양자화 값($\hat{r} = Q(\|c(n)\|_1)$)을 index로 사용한다. 효율적인 양자화를 위해 클러스터링 방식을 사용하여 스칼라 양자화기를 설계하였다. 스칼라 양자화기의 설계를 위해 20분동안의 음성 신호를 사용하였다. 먼저 앞에서 설계된 MA 예측기를 사용하여 60000개의 예측 에러 벡터를 추출하였고, 추출된 예측 에러 벡터를 training 데이터로 하여 Lloyd 알고리즘을 사용하여 스칼라 양자화기 레벨값을 설계하였다.

표 3은 프레임당 23비트를 사용하는 product code PVQ를 LSF양자화기로 이용시, PVQ와 스칼라 양자화기에 다르게 비트를 할당했을 때의 평균 SD 결과를 나타낸 것이다. 표에서 보는 바와 같이 PVQ에 21비트, 스칼라 양자화기에 2비트를 할당했을 때 가장 최적인 결과를 나타내었다.

표 3. 비트할당에 따른 성능 비교

비트할당 (PVQ, SQ)	평균 SD (dB)	% of Outlier	
		> 2 dB	> 4 dB
(23,0)	3.78	29.75	33.88
(22,1)	1.16	5.78	0.83
(21,2)	0.96	4.92	0.17
(20,3)	1.02	4.13	0.00
(19,4)	1.15	6.61	0.00
(18,5)	1.39	13.22	0.83

복호화 과정에서 양자화된 LSF 벡터 $\hat{f}(n)$ 은 다음식과 같이 양자화된 예측 에러 벡터 $\hat{c}(n)$ 에 예측 벡터 $p(n)$ 과 평균 벡터 f 를 더하여 구한다.

$$\hat{f}(n) = \hat{c}(n) + p(n) + f \quad (17)$$

복호화 과정시에 LPC 합성필터의 안정도를 위해 양자화된 LSF 계수를 사용하여 순서성질을 확인한 후 필요에 따라 다음과 같은 재배치 과정을 수행한다.

```

for i=1,...,9
  if ( $\hat{f}(n) - \hat{f}_{i-1}(n) < JP$ )
     $\hat{f}_i(n) = (\hat{f}_i(n) + \hat{f}_{i-1}(n) + JP) / 2$ 
     $\hat{f}_{i-1}(n) = (\hat{f}_i(n) + \hat{f}_{i-1}(n) - JP) / 2$ 
  end
end
    
```

여기서 JP는 재배치 계수이며, 실험적으로 구한

다. 최적의 재배치 계수의 결정을 위해 5분간의 남녀 화자에 의한 음성 신호를 사용하여 최적 재배치 계수를 결정하였다. 그림4는 프레임당 전송률을 23 비트로 고정하였을 때, 재배치 계수에 따른 평균 SD의 그래프이다. 그림에서와 같이 가장 최적인 재배치 계수는 0.0092임을 알 수 있다.

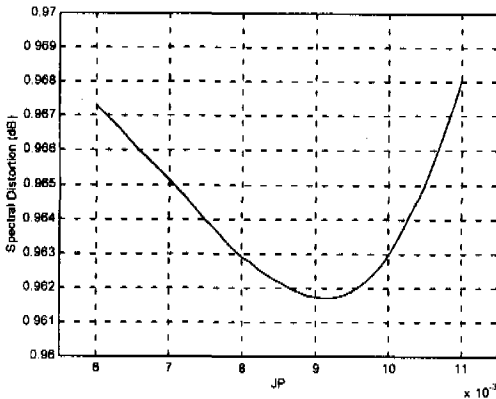


그림 4. 재배치 계수의 변화에 따른 평균 SD값 분포

V. 실험 및 결과

성능을 평가하기 위해 20 msec의 LPC 분석 주기를 사용하였으며, 30 msec Hamming window를 사용하였다. LPC 분석 차수는 10차 이며, 분석 방식으로는 autocorrelation 방식을 사용하였다.

실험에 이용된 모든 음성 데이터들은 FM 라디오를 통해 녹음한 음성신호를 3.4kHz로 대역제한시킨 후 8kHz로 샘플링하여 제작하였다. PPVQ 양자화기의 최적 계수값 설계를 위해 남성화자 10명과 여성화자 10명에 의해 발음된 20분 동안의 음성신호를 사용하였고, 설계된 PPVQ 양자화기의 성능을 평가하기 위한 테스트 데이터로는 최적 계수값 설계시 사용되지 않은 5명의 남성화자와 4명의 여성화자에 의해 발음된 9개의 단문을 사용하였다.

본 실험에서 PPVQ의 SD성능을 시플레이션을 통해 분석하였다. 표 4은 4가지 비트할당에 대한 PPVQ LSF 양자화기의 성능을 나타내었다. 그림 5는 프레임당 23 비트 할당시 테스트 데이터에 대한 PPVQ 양자화기의 SD값 분포를 나타내었다. 표4으로부터 PPVQ방식은 프레임당 23비트를 할당했을 때, 0.96 dB의 평균 SD값을 가짐을 알 수 있다. 그림6은 PPVQ방식과 split VQ(SVQ) 방식의 평균 SD성능을 비교하였다. 사용된 SVQ방식은 입력 벡

터를 3개의 부벡터로 나눈 three-split VQ방식이다. 그림 6으로부터 PPVQ방식은 기존의 SVQ방식보다 평균 SD값이 낮음을 알 수 있다. 또한 같은 SD값을 기준을 했을 때 PPVQ 방식은 SVQ방식에 비해 프레임당 대략 1 비트 절약할 수 있다.

PVQ는 일종의 격자 양자화 방식으로서 코드북 저장에 위한 메모리를 전혀 필요로 하지 않고 같은 비트율의 SVQ보다 적은 계산량을 가지는 큰 장점을 가지고 있다. 그래서 PVQ를 여러 벡터 양자화기로 사용하는 PPVQ방식은 메모리와 계산량 측면에서 매우 효율적이다.

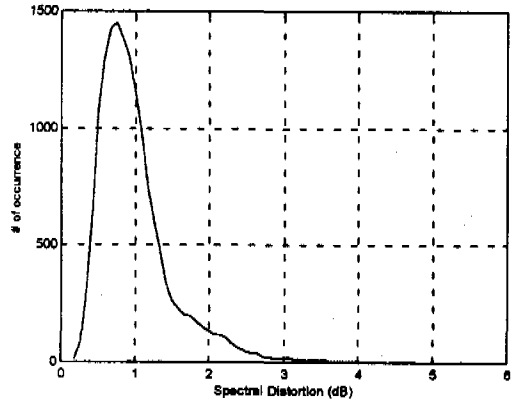


그림 5. 프레임당 23비트 할당시 테스트 데이터의 SD값 분포

표 4. 설계된 PPVQ 양자화기의 SD성능

비트할당	평균 SD(dB)	% of Outlier	
		> 2 dB	> 4dB
24	0.91	4.34	0.00
23	0.96	4.92	0.17
22	1.06	6.76	0.08
21	1.19	9.60	0.25

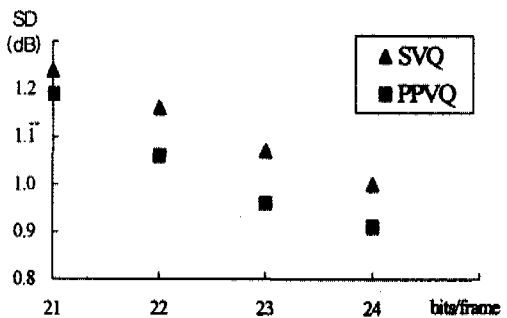


그림 6. PPVQ방식과 SVQ방식의 평균SD 성능 비교

VI. 결론

본 논문은 LPC 계수의 효율적인 양자화기에 대한 연구로서, 입력 신호의 기하학적인 특징을 이용하여 우수한 양자화 성능을 나타내는 pyramid vector quantization(PVQ)과 채널 에러의 영향을 적게 받는 MA 예측기를 사용한 predictive PVQ(PPVQ)를 제안하였다. MA 예측기를 이용하여 예측 에러 벡터를 추출하고, 추출된 예측 에러 벡터는 product code PVQ를 사용하여 양자화된다. 전송된 LSF 벡터는 순서성질을 만족하도록 재 배치되었다.

제안된 PPVQ 양자화기는 프레임당 23비트에서 0.96 dB의 평균 SD 성능을 보였다. 또한 PPVQ 양자화기는 적은 계산량과 코드북을 위한 메모리를 전혀 필요로 하지 않는 장점을 가지고 있다.

참고 문헌

[1] F. K. Soong and B. H. Juang, "Line spectrum pair(LSP) and speech data compression," Proc. IEEE ICASSP, pp. 1.10.1~ 1.10.4, Mar. 1984.

[2] N. Sugamura and N. Farvardin, "Quantizer design in LSP speech analysis-synthesis," IEEE J. Select. Areas in Commun., vol. 6, pp. 432-440, Feb. 1988.

[3] J. Grass and P. Kabal, "Methods of improving vector-scalar quantization of LPC coefficients." Proc. Int. Conf. Acoust., Speech, Signal Proc., Toronto, pp. 657-660, May 1991.

[4] K. K. Paliwal and B. S. Atal, "Efficient vector quantization of LPC parameters at 24 bits/frame," IEEE Trans. Speech and Audio Proc., pp. 3-14, Jan. 1993.

[5] A. Gersho and R. M. Gray, Vector Quantization and Signal Compression, Norwell, MA: Kluwer, 1992.

[6] W. Linde, A. Buzo and R. M. Gray, "An algorithm for vector quantizer design," IEEE Trans. Commum., vol. COM-28, pp. 84-95, Jan. 1980.

[7] H. Ohmuro, T. Moriya, K. Mano and S. Miki, "Coding of LSP parameters using

interframe moving average prediction and multi-stage vector quantization," IEEE Workshop Speech Coding for Telecom., Sainte-Adele, Canada, pp. 63-64, Oct. 1993.

[8] N. Sugamura and N. Farvardin, "Quantizer design in LSP speech analysis-synthesis," IEEE J. Select. Areas in Commun., vol. 6, pp. 432-440, Feb. 1988.

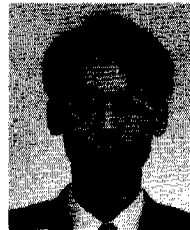
[9] T. R. Fischer, "A pyramid vector quantizer," IEEE Trans. on Information Theory, vol. 32, no. 4, pp. 568-583, July 1986.

[10] W. A. Finamore and W. A. Pearlina, "Optimal encoding of discrete-time continuous-amplitude memoryless sources with finite output alphabets," IEEE Trans. Inform. Theory, vol. IT-26, pp. 144-155, Mar. 1980.

[11] S. P. Lloyd, "Least squares quantization in PCM," IEEE Trans. Inform. Theory, vol. IT-28, pp. 129-137, 1982.

조인환(Inwhan Cho)

정회원



1974년 2월 26일생
 1996년 2월 : 한양대학교 제어계측공학과 졸업(공학사)
 1998년 2월 : 한양대학교 대학원 제어계측공학과 졸업(공학석사)

1998년 2월 현재 : 세트리 연구소 연구원

강삼원(Sangwon Kang)

정회원



1958년 3월 20일생
 1980년 2월 : 한양대학교 전자공학과 졸업(공학사)
 1982년 2월 : 서울대학교 대학원 전자공학과 졸업(공학석사)

1990년 7월 : Texas A&M 대학교 전기공학과 졸업(공학박사)

1982년12월~1994년2월 : 한국전자통신연구소

1994년 3월~현재 : 한양대학교 제어계측공학과 부
교수

<주관심 분야> 음성코딩, 이동통신, 신호처리