

## 저비트속도음성부호화를 위한 고조파스펙트르진폭의 벡토르량자화방법

박 광 철

경애하는 최고령도자 김정은동지께서는 다음과 같이 말씀하시였다.

《정보통신을 최첨단수준으로 발전시키는데 힘을 넣어야 합니다.》

저비트속도의 음성부호화에서 고조파부호화방법[1, 2]은 려진신호를 고조파를 리용하여 구성함으로써 보다 높은 음질을 보장하는 효과적인 기술수법으로 되고있다.

론문에서는 여러가지 고조파진폭들로 구성된 스펙트르벡토르에 대한 무게벡토르량자화방법을 고찰하였다.

### 1. 차 원 변 환

#### 1) 스펙트르보간

고조파벡토르려진 음성부호화에서 고조파스펙트르포락(SE)을 이루는 점들의 수는 피치에 따라 변한다. 그것은 스펙트르포락이 매개 고조파의 추정된 진폭들의 모임이기 때문이다. 고조파개수는 대략 9부터 70사이의 범위에 놓여있다.

고정된 차원으로 스펙트르포락을 벡토르량자화하자면 변수개수를 고정된 개수로 변환시켜야 한다.

론문에서는 FIR저역려파기와 1차선형보간기를 결합한 차원변환기를 제안하였다.

첫번째 단계에서는 8배과표본화를 위한 FIR저역려파기를 다음과 같이 설계한다.

65차과표본화려파기의 결수  $coef[i]$ 는 다음식에 의하여 얻는다.

$$coef[i] = \frac{\sin \pi(i-32)/8}{\pi(i-32)/8} (0.5 - 0.5 \cos 2\pi i / 64) \quad (0 \leq i \leq 64) \quad (1)$$

두번째 단계에서는 8배과표본화된 자료에 대하여 1차선형보간을 진행하여 고정된 고조파주파수  $n \times \omega_0$  ( $1 \leq n \leq N$ )에서의 스펙트르진폭을 계산한다. 여기서  $\omega_0$ 은 기본과주파수 혹은 고정차원벡토르에서 성분들의 주파수간격이고  $N$ 은 고정차원스펙트르벡토르를 포함하는 고조파들의 수이다. 8배과표본화를 위한 FIR저역려파에 요구되는 곱하기와 더하기의 수는  $N \times 2 \times 8$ 로서 직접형FFT-IFFT방법보다 1/40배 작다. 이러한 방법으로 고정차원이  $N$ 인 스펙트르벡토르를 얻는다.

#### 2) 감각무계려파행렬의 구성

고정차원고조파스펙트르벡토르를 량자화하자. 고조파진폭의 량자화를 하기 전에 감각무계려파기의 전달함수  $W(z)$ 와 LPC합성려파기  $H(z)$ 를 정의하는데  $W(z)$ 와  $H(z)$ 의 주파수응답은 고조파스펙트르포락의 무게벡토르량자화에 리용할수 있다.

$H(z)$ 와  $W(z)$ 는 다음과 같이 정의한다.

$$H(z) = \frac{1}{1 + \sum_{i=1}^P \alpha_i z^{-i}}, \quad W(z) = \frac{1 + \sum_{i=1}^P \alpha_i \lambda_b^i z^{-i}}{1 + \sum_{i=1}^P \alpha_i \lambda_a^i z^{-i}}$$

여기서  $\alpha_i (1 \leq i \leq P)$ 는 현재프레임의 LPC결수들이다. 처리되고있는 음성프레임의 LPC결수들은 이미 얻어졌다.  $\lambda_a$ 와  $\lambda_b$ 는 량자화잡음의 형태를 결정하는 상수로서 각각 보통 0.5와 0.8로 준다. 고조파주파수점  $\omega = n \times \omega_0$ 에서의 진폭  $|H(e^{j\omega})|$ 와  $|W(e^{j\omega})|$ 를 표본화하여 무게행렬  $\mathbf{H}$ 와  $\mathbf{W}$ 의 대각요소들을 얻는다. 이 행렬을 벡토르량자화에서 리용한다.

## 2. 벡토르량자화

고정차원벡토르는 무게행렬  $\mathbf{H}$ 와  $\mathbf{W}$ 를 리용하여 량자화하는데 그것의 구성도는 그림과 같다. 2Kbit/s의 비트속도를 위한 부호화에서는 스펙트르포락에 대한 벡토르량자화와 리득에 대한 스칼라량자화기를 함께 적용한다. 형태부호책과 리득부호책 2개를 모두 설계하고 탐색하는데 무게이지러짐측도  $D$ 를 사용한다.

$$D = \left\| \frac{1}{\|\mathbf{x}\|} \mathbf{W} \mathbf{H} (\mathbf{x} - \mathbf{g}(s_0 + s_1)) \right\|^2 \quad (2)$$

여기서  $\mathbf{x}$ 는 고정벡토르차원을 가진 차원변환고조파스펙트르진폭차원모임,  $\mathbf{s}_0$ 은 스펙트르포락형태 0의 출력,  $\mathbf{s}_1$ 은 스펙트르포락형태 1의 출력,  $\mathbf{g}$ 는 스펙트르포락리득부호책의 출력이다.  $\mathbf{H}$ 와  $\mathbf{W}$ 는 앞에서 유도된 무게행렬들이다.

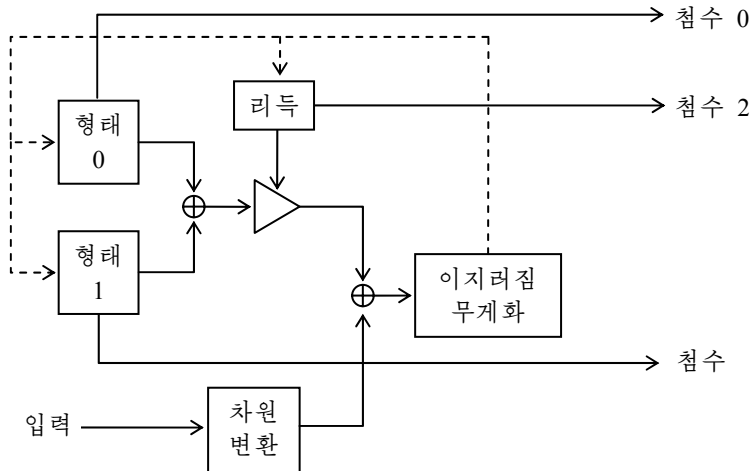


그림. 고조파진폭벡토르량자화의 구성도

부호책훈련은 다음과 같이 진행한다. 무게이지러짐측도 (2)를 리용하여 일반화된 로이드(Lloyd)알고리즘(GLA)에 따라 부호책에 대한 훈련을 진행한다. 스펙트르포락형태  $k (k=0, 1)$ 에서의  $j$ 번째 형태벡토르  $\mathbf{s}_{k,j}$ 를 리용하는 모든 프레임들의 이저리짐을 합하고  $\mathbf{s}_{k,j}$ 에 관하여 전체 무게이지러짐의 미분을 취한 다음 그 결과를 령으로 놓으면  $\mathbf{s}_{k,j}$ 에

대한 새로운 중심  $s_{k,j}$  를 얻는다.

$$u_{k,j} = \frac{\sum_{i \in j} g_i A_i^T A_i (x_i - g_i s_{1-k,i})}{\sum_{i \in j} g_i^2 A_i^T A_i}$$

여기서

$$A_i = \frac{W_i H_i}{\|x\|}$$

이며 첨수  $i$ 는 스펙트르포락형태  $k$ 로부터  $s_{k,j}$  를 선택한 프레임번호이며  $x_i, W_i, H_i$  들은 원천벡토르와 해당하는 무게들이며  $g_i, s_{1-k,i}$  들은  $i$ 번째 프레임에서 선택된 부호단어들이다.

마찬가지로  $j$ 번째 리득부호단어에 대하여 새로운 중심  $v_j$  는 다음과 같이 주어진다.

$$v_j = \frac{\sum_{i \in j} x_i^T A_i^T A_i (s_{0,i} + s_{1,i})}{\sum_{i \in j} \|A_i (s_{0,i} + s_{1,i})\|^2}$$

여기서 첨수  $i$ 는  $j$ 번째 리득단어를 선택하는 프레임번호이다.

### 3. 실험 결과

2.0Kbit/sHVXC의 성능평가는 직접 입력된 음성표본들을 가지고 남성과 여성음으로 분류하고 배경잡음을 추가하는 방법으로 진행하였다. 검사에서 부호기성능을 평가하는 95%신뢰도(CI)를 가지는 평균평가량(MOS) 분석결과를 표에 주었다.

표. 각이한 음성부호기들의 평균평가량 분석결과

부호기	남성			여성		
	조선어	영어	평균	조선어	영어	평균
LPC-10	2.29	2.44	2.36	2.23	2.44	2.33
AMR	3.32	3.18	3.25	3.02	2.91	2.96

### 맺 는 말

2.0Kbit/sHVXC부호기가 2.4Kbit/s LPC-10보다 낮은 속도를 보장하면서도 높은 성능을 나타낸다는것을 밝혔다.

### 참 고 문 헌

- [1] Y. Linde et al.; IEEE Trans. Commun., 28, 84, 2011.
- [2] T. Moriya et al.; IEEE Trans. Acous. Speech Signal Process, 36, 1223, 2013.

## **A Method of the Vector Quantization of Harmonic Spectral Magnitudes for Low-bit-rate Speech Coding**

*Pak Kwang Chol*

In this paper we proposed a new method of the weighted vector quantization for spectral vectors composed of the harmonic excitation magnitudes to ensure a better quality of speech in the low bit rate.

Key words: vector quantization, speech coding