

# 에너지 상태와 VQ를 이용한 화자 독립 격리 단어 인식

## (Speaker-Independent Isolated Word Recognition Using Energy State and VQ)

黃 英 洙,\* 尹 大 熙,\* 車 日 煥\*

(Young Soo Hwang, Dae Hee Youn, and Il Whan Cha)

### 要 約

본 연구에서는, 시간 특성이 유사한 격리 단어 인식을 위하여 VQ와 피이드백(feedback) VQ방법에 에너지 파라미터를 이용한 격리 단어 인식방법을 제안한다.

제안된 방법은, 프레임 단위로 음성 에너지 변화를 추출한 후, 음성 구간을 세 종류의 구간(에너지가 상승하는 구간, 에너지 변화가 없는 구간, 에너지가 감소하는 구간)으로 구분한다. 표준 패턴은, 학습 데이터들을 세 종류의 구간으로 구분한 후, 각 구간에 대한 codebook과 상태함수(state function)을 구성한다. 인식시에는, 입력 음성의 프레임들을 세 종류의 구간중 한가지 구간으로 구분한 후, 해당된 구간의 표준 패턴과 비교하여 인식을 수행한다.

제안된 방법의 성능 평가를 위하여, 격리 숫자음을 화자 독립으로 인식 한 결과, 시간 배열 특성이 유사한 음성에 대한 인식을 향상시킬 수 있었으며, 계산량은 DTW방법에 비해 적은 것으로 나타났다.

### Abstract

In this paper, we suggest the algorithm of isolated word recognition in VQ and feedback VQ and energy state, in order to improve recognition performance of isolated word which has similar time-sequence information.

This method divides the speech frames into three part (energy ascend part, flat part, descend part) using speech energy information. After dividing the training data into three part, the reference patterns are consisted of codebook and state function in each part. In the recognition procedure, test speech frames are seperated by energy information and then are compared with a set of prestored reference patterns.

For speaker-independent recognition of Korean isolated digit, this algorithm shows the improvement of similar time-sequence information word and reduces the computational requirement of DTW.

### I. 서 론

음성 인식에서 널리 이용되는 패턴 매칭(Pattern matching) 방법으로는 VQ와 DTW 등이 있다.

DTW는, 입력 음성의 시간에 따른 변동 특성을 비선

\*正會員, 延世大學校 電子工學科  
(Dept. of Elec. Eng., Yonsei Univ.)  
接受日字: 1989年 8月 2日

형 시간 매칭 방법으로 인식하는 것으로서, 인식률은 높으나 시간이 많이 소요되는 단점이 있다.<sup>[1]</sup>

한편 VQ(2,12)를 격리 단어 인식에 이용할 경우에는, clustering기법으로 구한 codebook을 사용하게 되며, 이 codebook을 구성할 때의 평균 왜곡은 입력 단어 프레임의 시간 배열을 고려하지 않고 구성하였기 때문에, DTW와 달리 입력단어의 시간 배열 특성을 무시한 인식이 가능하다. 이와같이 시간 배열 특성을 고려하지 않고 음성 인식이 가능하기 때문에 codeword 수만 적으면 기억용량과 계산시간을 줄일 수 있고, 인식할 단어들의 구성 특성이 완전히 틀린 단어일 경우에는, VQ만으로도 인식률을 높일 수가 있다. 그러나 유사한 구성 특성을 갖는 단어들을 인식할 경우에는, 각 구성 특성을 구분하기 위하여, codebook 구성이나 인식방법을 달리 해야 한다. VQ의 인식방법을 변형시킨 것으로는 피이드백 VQ<sup>[2]</sup> 방법 등이 있다.

피이드백 VQ는, VQ에 입력 음성시간 특성을 고려하여 메모리 구조를 갖는 인식 방법으로, 보다 적은 계산량과 향상된 인식률을 얻고자 개발되었다. 그러나 피이드백 VQ는 입력음성의 시간 특성에 맞춰 고안되었기 때문에, 시간에 따른 특성이 비슷한 음성에서는 인식이 저하된다.

이러한 문제점을 보완하기 위하여, 본 연구에서는 인식시와 codebook구성시 에너지 파라미터를 이용하였다.

여기에서 각 단어의 codebook 구성은, 학습 데이터 프레임들을 에너지 파라미터를 이용하여 세 구간으로 구분한 후, 각 구간에 대해 codebook을 구성하게 된다. 즉, 각 단어에 대해 세 개의 codebook이 구성된다. 인식시에는 미지 입력 음성 프레임을 codebook 구성시와 같이 에너지 파라미터를 이용하여 구분한 후, 각 구간에 해당하는 codebook 내 codebook과 비교하게 된다.

이와같이 에너지 파라미터로 음성을 구분하여 codebook을 구성하고 인식함으로써 보다 세밀한 시간 특성을 갖는 codebook을 얻을 수 있어, 인식률을 향상시키는 동시에 인식시간은 일반 VQ나 피이드백 VQ와 거의 유사하다.

본 연구의 구성은 다음과 같다. II장에서는 VQ와 피이드백 VQ의 이론적 서술과 음성인식 방법을 설명하였고, III장에서는 제한한 방법의 구체적 사항들을 기술하였으며, IV장에서는 실험 및 결과에 대한 분석과 결론으로 끝을 맺었다.

II. VQ(Vector Quantization)

VQ는 DTW와 함께 음성인식에서 많이 사용하고 있

는 패턴 매칭 방법이다.

VQ를 격리 단어 음성인식에 이용할 경우에는, clustering기법을 사용하여, 각 단어에 대한 codebook을 구성한 후, 미지 입력 음성 프레임을 각 codebook에 대하여 최소화되는 codeword를 찾아내어, 전체거리가 최소화되는 단어를 미지의 입력음성이라고 판정하게 된다.

음성 인식에 적용된 VQ의 블록도를 그림에 나타내었다.

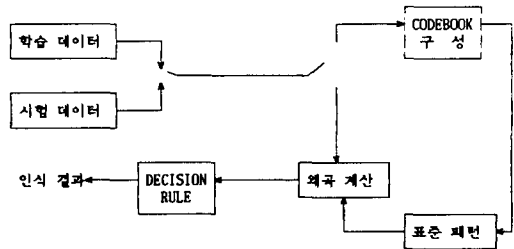


그림 1. VQ를 이용한 음성 인식 블록도  
Fig. 1. Speech recognition block diagram using VQ.

그림에서 codebook은 학습 데이터를 사용하여 작성하게 되며, 학습데이터로는 음성 프레임 단위의 샘플 데이터, FFT계수, 자기상관계수, 예측계수 등이 이용된다. 본 연구에서는 codebook 구성시, codebook의 주파수 특성이 좋다고 발표된 N.Sugamura 등이 사용한 SPLIT방법<sup>[3]</sup>으로 codebook을 구성하였다. 또한 그림의 왜곡 계산에 이용되는 search방법은 tree search방법과 full search방법 등이 있는데, 본 연구에서는 codebook의 크기가 적기 때문에 full search방법을 이용하였다.

VQ의 왜곡 계산식은 다음과 같다.

$$Du = \frac{1}{N} \sum_{n=1}^N \min_{1 \leq i \leq L} [d(A_n, b_i^n)] \quad (1)$$

- An : 입력 음성의 LPC벡터
- b<sub>i</sub><sup>n</sup> : U번째 단어 codebook내 I번째 codeword
- L : codebook의 크기
- N : 입력 음성의 프레임 수

윗 식에서 거리 d(.)는 Itakura-Saito 거리계산식을 사용하였다.

위와같은 VQ 패턴 매칭 방법은, codebook 수가 적을 경우 DTW에 비해 기억용량과 계산시간을 줄일 수 있다는 장점이 있지만, 음성의 시간 특성을 고려하지 않는다는 단점을 갖고 있다. 이러한 문제점을 보완하여 구성된 VQ의 변형방법으로 피이드백 VQ가 있다.

피이드백 VQ는 입력음성의 메모리 구조를 고안하여 제안된 것으로서, 각 입력 음성 벡터에 대하여 다른 codebook을 사용하고, 이 codebook은 과거의 입력음성 벡터에 의해 결정된다.

일반적인 VQ는 식에서 나타낸 것과 같이, 음성 인식 시 입력벡터와 가장 유사한 codebook를 구하기 위하여, codebook내의 모든 codebook와 비교해야하는 반면에, 피이드백 VQ는 음성 인식시, 입력음성 벡터와 비교되는 codewords가 전 프레임의 codewords에 의해 결정된다. 즉, 피이드백 VQ는 입력벡터들 사이의 상관 관계를 이용한 상태 함수에 의해 다음 프레임에서 비교될 codewords가 결정된다.

그림에 K개의 상태(state)를 갖는 유한상태 모델(finite state model)을 나타내었다.

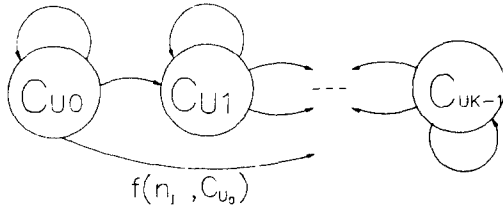


그림 2. K개의 상태를 갖는 유한상태 모델  
Fig. 2. Finite state model of K states.

그림에서 각 상태 Cui는 codeword를 나타내며, 가지(branch)는 다음 상태함수(next state function) f(nj, Cui), 즉 next code word를 나타내는 것으로서, 상태의 의미는 현재 입력 벡터에 대한 근사화이며, 다음 입력벡터에 대한 예측으로 볼 수 있다.

피이드백 VQ는 다음 프레임에서 비교될 next codeword가 적을수록 계산량이 적게 되며, 이러한 next codeword 상태함수 구성방법으로는 histogram방법, nearest neighbor방법, omniscient 구성방법<sup>[9,10]</sup> 등이 있는데, 본 연구에서는 histogram 구성방법을 사용하였다.

일반적인 피이드백 VQ를 이용한 음성인식 구성도를 그림에 나타내었다.

그림3의 피이드백 VQ의 왜곡 계산식은 다음과 같다.

$$D_{ui} = \min_k d(A_i, C_{uk}), K = f(n_0, C_{uj}), \dots, f(n_{i-1}, C_{uj}) \quad (2)$$

Dui: i라는 단어의 i번째 프레임의 최소왜곡

Ai: i라는 단어의 i번째 특징벡터

Cuj: (i-1)번째 프레임의 최소왜곡이 되는 codeword

N: next codeword수

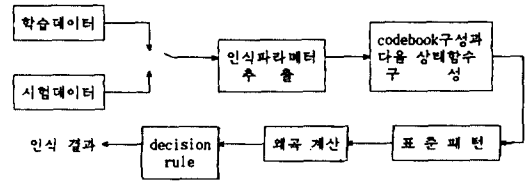


그림 3. 피이드백 VQ를 이용한 음성인식 구성도  
Fig. 3. Speech recognition block diagram using feedback VQ.

### III. VQ와 에너지 상태를 적용한 격리 단어인식

VQ의 codebook은 전체 학습 데이터를 대상으로 구성되기 때문에 codeword의 특성이 음성의 특징을 제대로 나타내지 못할 수 있고, feedback VQ는 VQ방법에 입력 음성시간 특성을 구조를 갖고 인식하기 때문에, 시간특성 배열이 유사한 음성(예: 일↔칠, 삼↔사 등)에서는 인식률이 저하된다.

위와 같은 codeword 구성상의 문제점을 해결하기 위하여, 본 연구에서는 음성구간의 에너지를 이용하여 피이드백 VQ와 VQ의 방법으로 격리 숫자음 인식을 하였다. 즉, 음성 구간중 에너지가 증가하고 있는 구간을 ASCEND부분으로, 에너지 변화가 거의 없는 구간을 FLAT부분으로, 에너지가 감소하고 있는 구간을 DESCEND부분으로 구분한 후, 에너지로 구분된 각 부분에 해당하는 학습 데이터들에 대해 SPLIT방법으로 음소형태 표준패턴을 작성한다. 이와같이 에너지로 음성 구간을 구분한 후, 세 구간에 대해 각각의 codebook을 작성하면, 전체음성 구간을 대상으로 구성한 codebook보다 정확한 특성을 갖는 codebook이 구성되어 인식률의 향상을 얻을 수 있다.

또한 각 에너지 구간(ASCEND, FLAT, DESCEND)의 codebook내 codeword 수를 일반 VQ에서 이용하는 codeword수와 같게 작성하고, 피이드백 VQ방법인 경우에는, next codeword수를 같게 하면, 미지의 입력음성 프레임과 비교할 codeword 수가 같게 되어, 인식시 계산량은 일반 VQ나 피이드백 VQ와 같게 된다.

그림에 피이드백 VQ방법에 에너지 파라미터를 적용시킨 유한상태 모델을 나타내었다

그림에서 Cuij는 에너지 i상태의 codeword를 나타내며, 가지는 next codeword의 결정을 나타내는 것으로서, 일반 피이드백 VQ와의 차이점은 codebook 구성이 각 에너지 부분으로 나뉘어져 구성되었기 때문에, 각 에너지 구간의 모델을 연결하는 가지가 있어야 한다. 이 가지를 살펴보면, 에너지 구간의 변환 가지가 ASCEND

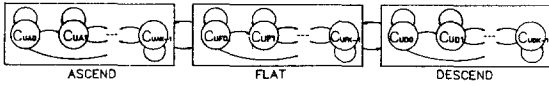


그림 4. 에너지를 이용한 유한상태 모델  
Fig. 4. Finite state model using energy.

부분과 DESCEND부분을 서로 연결하지 않고 있다. 이 이유는 음성에서 짧은 시간내 에너지의 급격한 변화는 없다고 생각되기 때문이다.

본 연구에서 제안한 격리단어 인식방법의 단계는 다음과 같다.

1) 에너지와 영교차율을 이용한 끝점 검출 알고리즘을 이용하여, 입력격리 단어의 음성 구간을 추출해 낸다.<sup>[6]</sup>

2) 음성 구간의 각 프레임에 대해 선형 예측 계수를 구하고, 현 프레임의 에너지, 두 프레임 이전의 에너지, 두 프레임 후의 에너지를 이용하여, 현 프레임이 ASCEND, FLAT, DESCEND 상태인지를 검토한다. 즉,

$$\frac{R_{i+2} - R_{i-2}}{R_i} < \theta 1, \dots \rightarrow \text{ASCEND}$$

$$\theta 2 \leq \frac{R_{i+2} - R_{i-2}}{R_i} \leq \theta 1, \dots \rightarrow \text{FLAT}$$

$$\frac{R_{i+2} - R_{i-2}}{R_i} < \theta 2, \dots \rightarrow \text{DESCEND}$$

$R_i$  :  $i$  번째 프레임의 0차 자기 상관 계수  
 $\theta$  : 임계치

이고, 그림5에 격리 숫자음에서의 에너지변화상태였다.

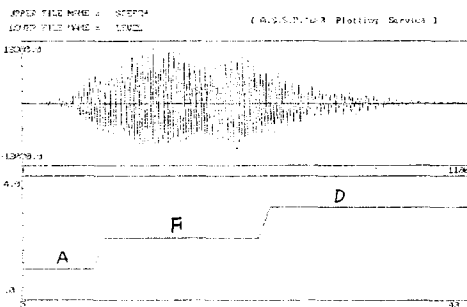


그림 5. 숫자음 '1'에서의 에너지 변화상태  
Fig. 5. Energy transition state of digit '1'

3) 각 에너지 구간에 해당하는 codebook내에서, 표준 패턴시 Histogram 방법으로 구한, 전 프레임에 해당하는 next codeword에 대해, 다음과 같은 Itakura-Saito distortion<sup>[6]</sup>을 이용하여 각 왜곡을 측정한다.

$$D = \frac{\hat{a}^T v \hat{a}}{a^T v a} - 1 \tag{3}$$

$\hat{a}$  : 시험 데이터의 선형 예측 계수  
 $a$  : 표준 데이터의 선형 예측 계수  
 $v$  : 시험 데이터의 자기 상관 계수

4) 단계3)에서 구한 왜곡값중 가장 적은 왜곡값을 현 프레임의 왜곡값으로 하며, 이에 해당하는 codeword를 다음 프레임 계산시 이용하기 위하여 저장한다. 즉,

$$D_{ui} = \min_k d(A_i, E', C_{uk}), k = f(n_0, E', E'', C_{uj}) \dots$$

$$\dots f(n_{N-1}, E, E, C_{uj}) \tag{4}$$

$D_{ui}$  :  $u$ 라는 단어의  $i$ 번째 프레임의 최소 왜곡  
 $A_i$  :  $u$ 라는 단어의  $i$ 번째 프레임의 특징 벡터 (에너지  $E'$  상태)  
 $C_{uj}$  :  $(i-1)$ 번째 프레임의 최소 왜곡이 되는 codeword (에너지  $E'$  상태)  
 $N$  : next codeword수

5) 모든 표준 패턴들에 대해 구한 왜곡값중 가장 적은 왜곡값을 갖는 표준패턴을 입력 미지의 입력단어라고 판정한다. 즉,

$$\text{미지의 입력 단어} = \min_u^{-1} D_u \tag{5}$$

$D_u$  :  $u$ 라는 단어의 평균 왜곡

위에서 서술한 에너지 파라미터를 이용한 피이드백 VQ의 격리단어 인식 과정에서 사용한 codebook 구성방법은 다음과 같다.

1) 표준 패턴을 구성시 사용되는 학습데이터 각각에 대해 ASCEND, FLAT, DESCEND부분으로 학습 데이터들의 프레임을 구분한다.

2) 단계 1)에서 구분된 각 에너지 부분에 해당하는 프레임들에 대해 SPLIT방법을 이용하여, 에너지 각 부분에 대한 codebook을 구성한다.

3) 학습 데이터들에 대해 Histogram방법을 이용하여, next codeword 확률분포를 구한다. 이 때 각 구간의 가능한 변화를 고려한다. 즉,

ASCEND → FLAT, ASCEND → ASCEND  
FLAT → ASCEND, FLAT → FLAT  
FLAT → DESCEND, DESCEND → FLAT  
DESCEND → DESCEND

위의 가능한 변화 고려 상태시, DESCEND →

ASCEND, ASCEND → DESCEND 에너지 변화상태는 고려하지 않는다. 그 이유는 음성이 상승에너지 상태에서 갑자기 에너지가 감소하는 경우와 그 반대의 경우는 발생하지 않기 때문이다.

본 연구의 인식시스템의 구성도를 그림6에 나타내었다.

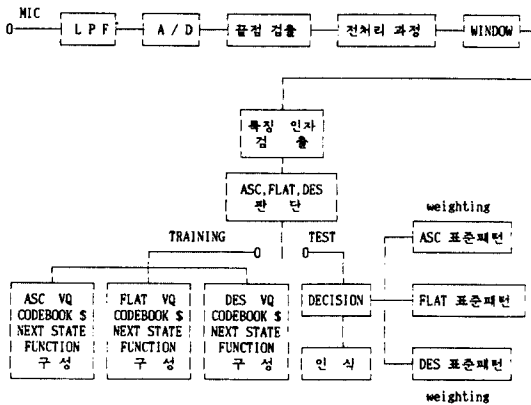


그림 6. VQ에 에너지를 적용시킨 음성인식 시스템  
Fig. 6. Speech recognition System using energy and VQ.

IV. 실험 및 결과 고찰

본 연구의 실험에서 사용된 데이터는, 방음 처리가 되지 않은 실내에서 20대 남성화자에 의해 자연스럽게 발음된 숫자음들을 사용하였다. 이 중 VQ와 feedback VQ의 표준패턴을 작성하기 위하여, 네 명의 화자가 각 숫자음을 열번씩 발음한 데이터를, DTW방법에서는 한명의 화자가 열번씩 발음한 음성중, MKM (Modified K-Means)방법<sup>[13]</sup>을 이용하여 한 개의 표준패턴을 작성하였다. 시험 데이터로는 표준패턴 구성시 포함되지 않은 다섯명이 다섯번씩 발음한 숫자음을 이용하였다. 그리고 DTW의 경우는 Sakoe가 제시한 경로를 사용하였다.

위와같이 구성된 데이터를 reel tape 녹음기에 녹음하여, 차단 주파수가 3.4KHz인 저역통과 필터에 통과시킨 후, 10KHz(16bit resolution)로 샘플링하였다. 음성신호 분석 구간은 25.6ms(256샘플), 분석 파라미터로 14차 LPC를 구하였고, 또한 음성의 각 구간을 구분하기 위한 에너지 파라미터로는 LPC계수를 구할 때 이용되는 0차 자기상관 계수를 사용하였다.

그리고 VQ와 피이드백 VQ에서는 각 숫자음에 대해 24개와 8개의 codeword로 구성된 codebook을 이용하여 인식결과를 구했고, 이 때 피이드백 VQ의 next codeword 수는 세개로 하였다. 또한, 본 연구에서 제안한 방법에서는, 각 숫자음에 대해, 한 에너지 구간당 8개의 codeword(즉, 각 숫자음에 대해 3\*8=24개의 codeword)를 사용하였으며, 피이드백 VQ의 next codeword 수는 세개로 하였다.

표에 에너지를 이용한 VQ, [표 2]에 에너지를 이용한 피이드백 VQ방법에 의한 인식률을 나타내었다.

VQ와 피이드백 VQ에서 24개의 codeword인 경우, 73.6%, 71.6%, 8개의 codeword인 경우 66%, 71%로 나타났으며, 표1과 2에 나타낸 에너지를 이용한 VQ와 피이드백 VQ에서는 87.6%, 88.4% 그리고 DTW에서는 85.6%의 인식률을 얻었다. 즉, VQ와 피이드백 VQ에 에너지 파라미터를 적용시켰을 경우, 모두 더 나은 인식률을 얻을 수 있었으며, 에너지를 이용한 피이드백 VQ가 에너지를 이용한 VQ에 비해 더 나은 인식률을 얻을 수 있었다.

오인식이 많이 발생한 숫자음들을 살펴보면, '오'와 '구'에서 많은 오인식이 발생하고 있는데, 이것은 두 음성의 모음 구간이 유사한 스펙트럼 특성을 갖고 있는 것으로 생각되며, 이 두 음성의 표준 패턴의 스펙트럼 특성을 그림 7에 나타내었다. 이 오인식에 대한 처리는, 후 처리 과정의 인식 단계가 필요할 것으로 생각된다.

본 연구에서 제안한 VQ와 피이드백 VQ에 에너지 파라미터를 적용한 숫자음 인식 결과를 다음과같이 요약할 수 있다.

표 1. 에너지를 이용한 피이드백 VQ인식결과  
Table 1. Speech recognition results of feedback VQ wing energy.

|   | 영  | 일  | 이 | 삼  | 사  | 오  | 육  | 칠  | 팔  | 구  |
|---|----|----|---|----|----|----|----|----|----|----|
| 영 | 24 |    |   |    |    |    |    |    |    |    |
| 일 |    | 19 | 7 |    |    |    |    |    |    |    |
| 이 |    |    | 4 | 18 |    |    |    |    |    |    |
| 삼 |    |    |   |    | 25 | 4  |    |    |    |    |
| 사 |    |    |   |    |    | 21 |    |    |    |    |
| 오 |    |    |   |    |    |    | 18 |    |    | 5  |
| 육 | 1  |    |   |    |    |    |    | 25 |    |    |
| 칠 |    |    | 2 |    |    |    |    |    | 25 |    |
| 팔 |    |    |   |    |    |    |    |    |    | 25 |
| 구 |    |    |   |    |    | 7  |    |    |    | 20 |

표 2. 에너지를 이용한 피이드백 VQ인식 결과  
Table 2. Speech recognition results of feedback VQ using energy.

|   | 영  | 일  | 이  | 삼  | 사  | 오  | 육  | 칠  | 팔  | 구  |
|---|----|----|----|----|----|----|----|----|----|----|
| 영 | 24 |    |    |    |    |    | 1  |    |    |    |
| 일 |    | 17 | 2  |    |    |    |    |    |    |    |
| 이 |    | 2  | 21 |    |    |    |    |    |    |    |
| 삼 |    |    |    | 25 | 3  |    |    |    |    |    |
| 사 |    |    |    |    | 22 |    |    |    |    |    |
| 오 |    |    |    |    |    | 18 |    |    |    | 5  |
| 육 | 1  |    |    |    |    |    | 24 |    |    |    |
| 칠 |    | 6  | 2  |    |    |    |    | 25 |    |    |
| 팔 |    |    |    |    |    |    |    |    | 25 |    |
| 구 |    |    |    |    |    | 7  |    |    |    | 20 |

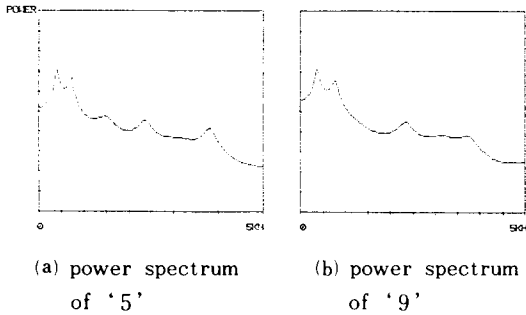


그림 7. 음성 '5'와 '9'의 FLAT부분의 파워 스펙트럼 (power spectrum)  
Fig. 7. Codeword spectrum of digit '5', '9'.

1) VQ와 피이드백 VQ에 에너지를 이용하여 인식할 경우, 보다 나은 인식률을 얻을 수 있었으며, DTW와 비슷한 인식률의 결과를 얻었다.

2) 계산량은 에너지를 이용한 피이드백 VQ와 일반 피이드백 VQ가, 에너지를 이용한 VQ와 일반 VQ 보다 적으며, 에너지를 이용한 피이드백 VQ와 일반 피이드백 VQ는 비슷하다.

3) 숫자음 인식률을 더 향상시키기 위해서는, 서로 거리가 비슷하게 나타난 음성(예: 오 ←→ 구) 등에, 후 처리 과정이 필요할 것으로 생각된다.

参 考 文 献

[1] H. Sakoe and S. Chiba, "Dynamic programming optimization for spoken word

recognition," *IEEE Trans. Acoust., speech, Signal Processing*, vol. ASSP-26, pp. 43-49, Feb. 1978.

[2] Y. Linde, A. Buzo and R.M. Gray, "An algorithm of vector quantizer design," *IEEE Trans. Commun.*, vol. COM-28, pp. 84-95, Jan. 1980.

[3] R.M. Gray, "Vector quantization," *IEEE ASSP Magazine*, vol. 1, pp. 4-29, Apr. 1984.

[4] N. Sugamura and S. Furui, "Isolated word recognition using strings of phoneme-like templates (SPLIT)," *J. Acoust. soc. Jpn. (E)* 5,4, pp. 243-252, 1984.

[5] L.R. Labiner and M.R. Sambur, "An algorithm for determining the endpoints of isolated utterances," *Bell syst. Tech. J.*, vol. 54, pp. 297-315, Feb. 1975.

[6] F. Itakura, "Minimum prediction residual principle applied to speech recognition," *IEEE Trans. Acoust. Speech, Signal Processing*, vol. ASSP-23, pp. 52-72, Feb. 1975.

[7] W.S. Youn and C.K. Un, "Isolated word recognition based on finite-state vector quantization," *Proc. ICASSP-86*, pp. 717-720, 1986.

[8] 황영수, 김원구, 차일환, 윤대희, "에너지와 VQ를 이용한 숫자음 인식", 한국음향학회 학술발표회 논문집, pp. 181~184, 1988.

[9] J. Foster, R.M. Gray and M.O. Dunham, "Finite-state vector quantization for waveform coding," *IEEE Trans. Inform. Theory*, vol. IT-31, pp. 348-359, May, 1985.

[10] M.D. Dunham and R.M. Gray, "Algorithm for the design of labeled-transition finite-state vector quantizers," *IEEE Trans. Commun.*, vol. COM-33, Jan, 1985.

[11] A. Haoui and D. Messerschmidt, "Predictive vector quantization," *Proc. ICASSP-84*, pp. 10.10.1-10.10.4, 1984.

[12] J.E. Shore and D.K. Burton, "Discrete utterance speech recognition without time alignment," *IEEE Trans. Inform. Theory*, vol. IT-29, pp. 473-491, July, 1983.

[13] J.G. Wilpon and L.R. Rabiner, "A modified K-Means clustering algorithm for use in isolated word recognition," *IEEE. Acoust., Speech, Signal Processing*, vol. ASSP-33, June, 1985.

著 者 紹 介



**黃 英 洙 (正會員)**  
 1960年 4月 9日生. 1982年 2  
 月 연세대학교 전자공학과 졸업.  
 1984年 2月 연세대학교 대학원  
 전자공학과 석사 학위 취득. 1984  
 年 3月~현재 연세대학교 대학  
 원 전자공학과 박사과정. 1989年  
 9月~현재 관동대학교 전임강사. 주관심분야는 음  
 향·음성 신호 처리.

**車 日 煥 (正會員)** 第25卷 第12號 參照  
 현재 연세대학교 전자공학과  
 교수



**尹 大 熙 (正會員)** 第25卷 第5號 參照  
 현재 연세대학교 전자공학과  
 조교수