

모음 인식과 벡터 양자화를 이용한 화자 인식

Speaker Identification Based on Vowel Classification and Vector Quantization

임 창 현* 이 황 수* 은 중 관*

(Chang Heon Lim, Hwang Soo Lee, Chong Kwan Un)

요 약

본 연구에서는, VQ(vector quantization)와 모음 인식에 기초한 화자 인식 알고리즘을 제안하고, 기존의 VQ를 사용한 화자 인식 알고리즘과 성능을 비교하였다. 제안된 화자 인식 알고리즘은 모음 분리, 모음 인식 그리고 평균 distortion 양을 계산하는 3개의 과정으로 구성되며, 이때 주어진 음성 신호로부터 모음 부분을 분리하기 위해 RMS 에너지, BTR (Back-to-Total cavity volume Ratio) 그리고 SFBR(Signed-Front-to-Back maximum area Ratio)이라는 3개의 parameter 를 사용하였다.

입력 음성 신호의 SNR이 20 dB이고 정확한 모음 분리가 수행되었을때, 제안된 화자 인식 알고리즘의 성능이 기존의 VQ를 사용한 화자 인식 알고리즘의 성능보다 대체로 좋았으며, 입력 신호가 전화선을 통과한 신호이고 잡음이 있는 경우에도 유사한 결과를 얻을 수 있었다.

ABSTRACT

In this paper, we propose a text-independent speaker identification algorithm based on VQ(vector quantization) and vowel classification, and its performance is studied and compared with that of a conventional speaker identification algorithm using VQ. The proposed speaker identification algorithm is composed of three processes: vowel segmentation, vowel recognition and average distortion calculation. The vowel segmentation is performed automatically using RMS energy, BTR(Back-to-Total cavity volume Ratio)and SFBR(Signed Front-to-Back maximum area Ratio) extracted from input speech signal.

If the input speech signal is noisy, particularly when the SNR is around 20dB, the proposed speaker identification algorithm performs better than the reference speaker identification algorithm when the correct vowel segmentation is done. The same result is obtained when we use the noisy telephone speech signal as an input, too.

* 한국과학기술원 전기 및 전자공학과

I. 서 론

음성 신호는 화자 고유의 음성 특징에 대한 정보와 음성 신호가 나타내는 의미를 동시에 내포하고 있다. 화자 인식은 이와 같은 사실에 기초하여 주어진 음성 신호로부터 화자의 음성 특징을 추출하여 음성 신호의 화자를 인식하는 것을 말한다.

화자 인식은 task의 종류에 따라 화자 확인(speaker verification)과 화자 식별(speaker identification)로 나눌 수 있다. 화자 확인은 입력 신호로서 음성 신호와 그 음성 신호를 제공한 화자의 identity가 주어졌을 때, 그 음성 신호와 identity가 서로 일치하는가를 검사한 후에 출력 신호로서 일치 또는 불일치라는 판정을 내리게 된다. 반면에 화자 식별은 입력 신호로서 음성 신호가 주어졌을 때, N명의 화자 중에서 그 음성 신호의 화자를 그 사람의 identity로 출력하는 것이다.

한편 화자 인식은 test 음성 신호의 text가 정해져 있으면 text-dependent 화자 인식이라고 하고, 정해져 있지 않으면 text-independent 화자 인식이라고 한다.

1977년 J.D. Markel⁽¹⁾ 등이 long-term averaging method라는 화자 인식 방법의 효용성에 대한 연구 발표를 하였다. 이 방법은 많은 음성신호를 가지고 test pattern과 reference pattern을 만들어야 비교적 좋은 성능을 얻을 수 있다는 단점을 가지고 있다.

이러한 문제점들을 해결하기 위해 제안된 방법이 VQ를 이용한 화자 인식방법이다.⁽²⁾⁽³⁾⁽⁴⁾ VQ는 일반적으로 data compression에 이용되는 방법으로서, 화자인식에서는 화자의 특성을 나타내는 data를 compression 하기 위해 사용된다고 할 수 있다. 이 방법에서는 VQ codebook으로 화자의 특성을 나타내는데, VQ codebook은 여러 개의 feature vector로 구성되기 때문에 test 음성 신호가 짧아서 생기는 variability를 고려할 수 있다. 따라서 VQ를 이용한 화자인식방법은 long-term averaging method에서 사용되는 test 음성 신호보다 짧은 음성 신호로 대등한 성능을 낼 수 있다는 장점을 가지고 있다.

본 논문에서는 VQ와 모음 인식을 이용한 text-

-independent 화자 인식 시스템을 제안하고, 기존의 VQ만을 이용한 text-independent 화자 인식 시스템과 성능을 비교하여 보았다. VQ와 모음 인식을 이용한 화자 인식 시스템은, 기존의 화자 인식 시스템이 test 음성 data의 모든 음소를 이용하여 인식하는데 반해, 모음 부분만을 사용하여 화자 인식을 수행하는 시스템이다.

2. VQ를 이용한 화자인식 시스템

앞서의 연구⁽²⁾에서 VQ를 화자 고유의 음성 특징을 모델링하는 방법으로 이용한 화자 인식 시스템이 제안된 바 있다. 본 논문에서는 reference 화자 인식 시스템으로 VQ를 이용한 화자 인식 시스템을 구성해 보았다.

화자의 음성 신호로 구성된 training data에 VQ를 사용하여 VQ codebook을 얻을 수 있다. 이 경우 VQ codebook은 화자 음성의 특징을 표현하는 feature의 집합이라고 할 수 있다. 본 논문에서 reference 화자인식 시스템으로 채택한 VQ를 이용한 화자 인식은 training 과정과 test 과정으로 구성된다. training 과정에서는 gain-normalized LPC distortion measure⁽⁵⁾와 modified K-means 알고리즘⁽⁶⁾을 사용하여, 각 화자에 대하여 1개의 VQ codebook을 구성한다. Test 과정에서는 training 과정에서 만든 VQ codebook을 이용하여 화자 인식을 수행한다.⁽²⁾

2.1 Test 과정

test 과정을 블록도로 나타내면 그림 2.1과 같다. 먼저, test 음성 신호가 주어지면 끝점 검출(end point detection)을 하고 각각의 frame에 대하여 LPC 계수를 구하게 된다. i번째 frame으로부터 구한 LPC 벡터를 a_i 라고 표시하자. 이러한 LPC 계수들을 VQ codebook을 이용하여 vector 양자화 한다. 이때 각각의 VQ codebook은 vector 양자화 과정에서 생기는 distortion들을 계속 더해 나간다. 이런 식으로 하여 음성 신호의 vector 양자화가 모두 끝났을 때, frame의 갯수를 L이라 하고 D_i 를 i번째 codebook의 평균 distortion이라 하면, 평균 distortion D_1 는

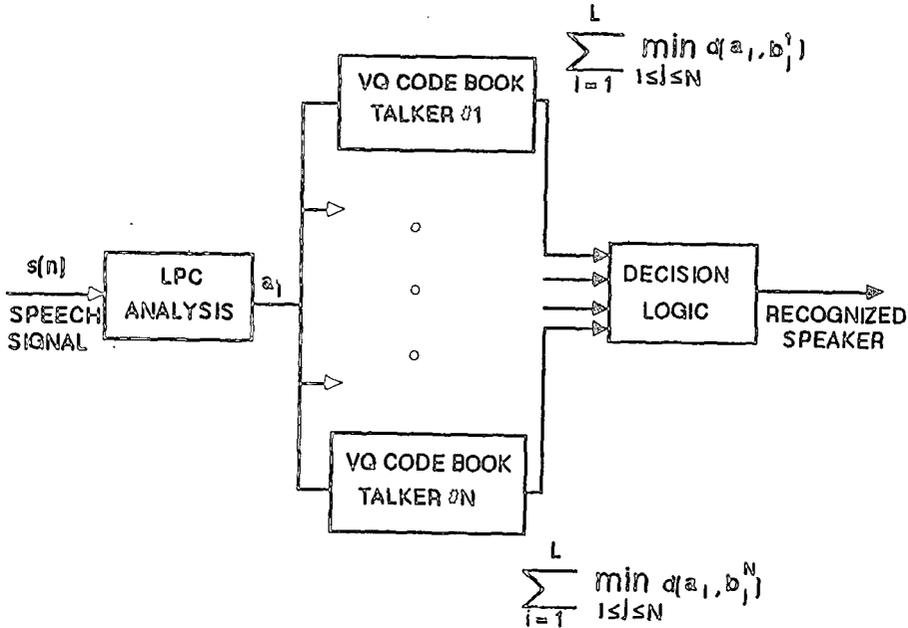


그림 2.1 VQ를 이용한 화자 인식 시스템의 블록도

다음과 같이 표시될 수 있다.

$$D_i = \frac{1}{L} \sum_{n=1}^L \min_j d(a_n, b_j^i)$$

위 식에서 b_j^i 는 i 번째 화자의 음성 특징을 나타내는 VQ codebook의 j 번째 codeword를 나타낸다.

test 음성 신호의 화자가 누구인가를 결정하기 위해서는, 먼저, codebook의 갯수가 N 이라고 할 때 D_1, \dots, D_N 중에서 최소값에 해당하는 D_i 를 고른후에, 평균 distortion으로 D_i 를 갖는 codebook이 대표하는 화자를 test 음성 신호의 화자로 선택하게 된다.

3. VQ와 모음 인식을 이용한 화자 인식 시스템

본 논문에서는, 음성 신호중에서, 화자의 음성 특징을 가장 많이 갖고 있다고 알려진 모음 부분만을 가지고 화자 인식 시스템을 구성해 보았다. 아울러 text-dependent 화자 인식의 장점을 이용하기 위해 모음 인식부분을 첨가하였다. 모음 부분만을 이용한 화자 인식 시스템을 제안한 이유를 정리하면 다음과 같다.

- 모든 단어는 적어도 하나의 모음을 포함하고 있기 때문에, 모음을 이용한 화자 인식 시스템은 tex-

t-independent 화자 인식 시스템으로 적합하다.

- 모음은 화자의 성도 모양과 glottal source의 성질과 밀접한 관계가 있으며, 다른 음소들에 비해 훨씬 많은 화자 음성 특징을 지니고 있다.⁽⁷⁾

- 음성 신호에서 모음 부분을 찾기가 비교적 수월하다.

- Energy level이 높기 때문에 다른 음소에 비해 noise의 영향이 적다.

- 주파수 영역에서 볼 때, 모음 부분은 다른 음소보다 전파선의 영향이 작다.

- 비음(nasal sound)은 생리적 상태에 매우 민감하여 unstable하기 때문에 본 논문에서는 이용하지 않았다.⁽⁸⁾

- 음성 처리 영역이 모음 부분으로 한정되기 때문에, 만약 모음을 분리하기 위한 overhead가 작다면 기존의 VQ를 이용한 화자 인식 시스템 보다 계산량을 줄일 수 있다.

그리고 벡터 양자화과정에서 서로 다른 음소의 LPC 계수가 matching되는 경우가 생기는 것을 방지하기 위해서 수동 모음 인식과정을 첨가하였다.

3.1 모음의 분리

본 논문에서 사용된 모음 분리 알고리즘은 다음과

같이 3개의 과정으로 구성된다.⁶⁹⁾

- 측정 과정 (Parameter extraction)
- 음성 특징 검출 과정 (Phonetic feature detection)
- 최종적인 모음/비모음 구간 결정 과정 (Vowel/nonvowel decision)

측정 과정에서는 주어진 음성 data로부터 RMS energy contour, BTR(Back-to-Total cavity volume Ratio), SFBR(Signed Front-to-Back maximum area Ratio)이라는 3개의 parameter를 추출한다.

음성 특징 추출 과정에서는 측정 과정에서 추출한 parameter를 이용하여, 주어진 음성 data를 묵음, turbulence noise, 모음/비모음의 4가지 부분으로 나눈다.

마지막으로 최종적인 모음/비모음 구간 결정 과정에서는 음성 특징 추출 과정에서의 결과를 토대로 최종적인 모음/비모음 구간을 결정한다.

그림 3.1은 이러한 모음 분리 알고리즘을 블록도로 나타낼 것이다.

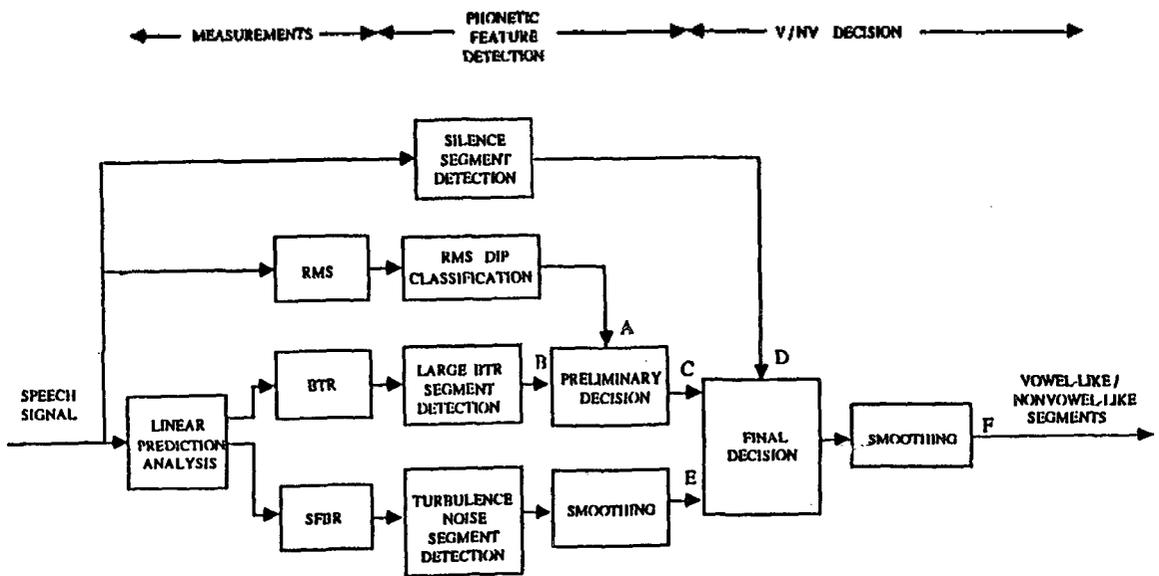


그림 3.1 자동 모음 분리 알고리즘의 블록도

3.2 VQ와 모음 인식을 이용한 화자 인식 시스템

3.2.1 Training과정

이 과정은 화자 인식에 사용되는 VQ codebook을 만드는 과정이다. VQ codebook을 구성하기 위해서는 training data가 필요하다. 이번 논문에서는 17개 격리 단어의 음성 신호에서 모음 부분을 분리하여 training 음성 data base를 만들었다. 이때 음성 신호파형을 직접 눈으로 보면서 모음 부분을 취했다. 17개의 격리 단어는 4개의 모음(아(와), 어, 오, 에(애))만을 포함한 단어를 선택하였기 때문에, 4개의 모음에 대한 training 음성 data만 생긴다. 4개의 모음을 선택한 이유는, 그 모음들이 일반단어에

서 나타나는 빈도수가 높고, 분리하기도 비교적 용이하기 때문이다.

VQ를 이용한 기존의 화자 인식 시스템에서는 화자 한 사람당 한개의 VQ codebook을 만들지만, 본 논문에서 사용한 모음 분리를 이용한 화자 인식 시스템에서는 한 화자의 각 모음에 대해서 VQ codebook을 한개씩 만들어야 한다. 본 논문에서는 7사람의 화자와 4개의 모음에 대해서 화자 인식 시스템을 구성하였으므로, 28개의 VQ codebook이 필요하다. VQ codebook은 gain normalized LPC distortion measure⁶⁵⁾와 modified K-means algorithm⁶⁶⁾을 사용하여 만든다.

3.2.2 test 과정

모음 분리를 자동으로 하고 모음 인식을 수동으로 하는 경우 수행과정은 다음과 같다. 그리고 시스템 블록도는 그림 3.2와 같다.

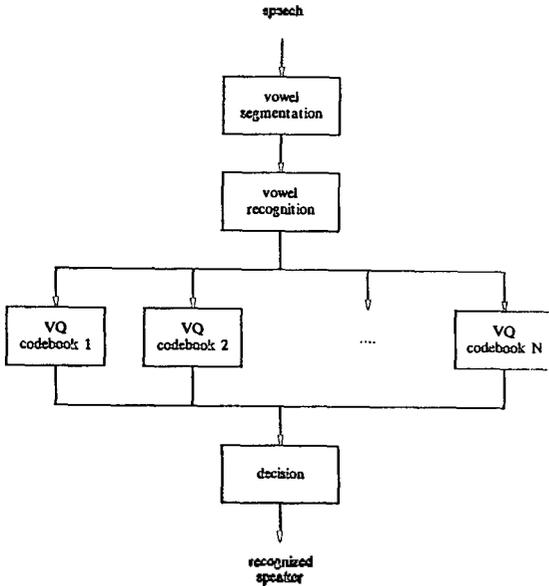


그림 3.2 VQ와 모음 인식을 이용한 화자 인식 시스템 블록도

- 1) Test 음성 data를 읽는다.
- 2) 자동 모음 분리 알고리즘(그림 3.1)에 따라 모음을 분리한다.
- 3) 분리된 모음 부분이 4가지 모음중 어느 것에 해당하는지 알아낸다.
- 4) 과정 3)에서의 모음인식 결과에 따라, 화자 인식에 사용될 VQ codebook을 선택한다.
- 5) 분리된 모음 부분에 대해서 LPC analysis를 하여 각 화자 마다 선택된 VQ codebook으로 벡터 양자화한다. 분리된 모음 부분 전체에 대해서 위 과정을 수행한 후에 각 VQ codebook의 평균 distortion을 계산한다.
- 6) 가장 작은 평균 distortion을 가진 VQ codebook이 대표하는 화자를 test 음성 신호의 화자로 결정한다.

4. 결과 및 토의

4.1 Data base구성

Data base는 서울의 공공 기관명을 나타내며 4가지 모음 (아(와), 어, 오, 에(애))만을 포함하는

33단어를 7명의 성인 남자가 각각 1번씩 발음한 총 231개의 격리 단어의 음성신호로 구성되어 있다. 화자 각각에게 주어진 33 단어중에서 17 단어의 음성 신호는 VQ codebook을 만드는데 사용하였다. 나머지 16단어의 음성 신호는 test 과정에서 사용하였다.

Reel tape에 기록된 analog 음성 data는 4.5KHz의 대역폭을 갖는 low pass filter를 통과시켜 고주파 성분을 제거한다. 이 때 low pass filter로는 8차의 Butterworth filter를 사용하였다. 그런 다음 NOVA 4/x minicomputer에 부착된 12 bit analog/digital (A/D) converter를 사용하여 녹음된 analog 음성 data를 10KHz의 clock으로 sampling하여 digital 음성 data로 변화시킨후, magnetic tape에 저장하여 data base로 사용했다.

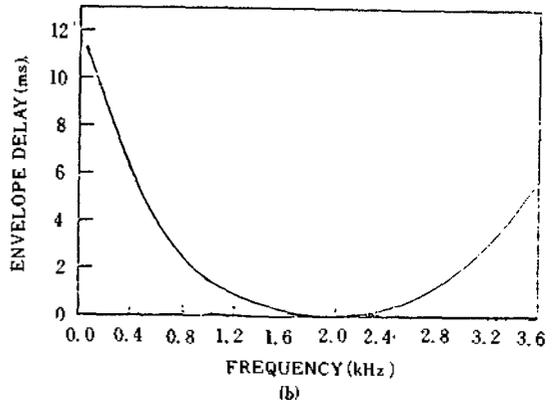
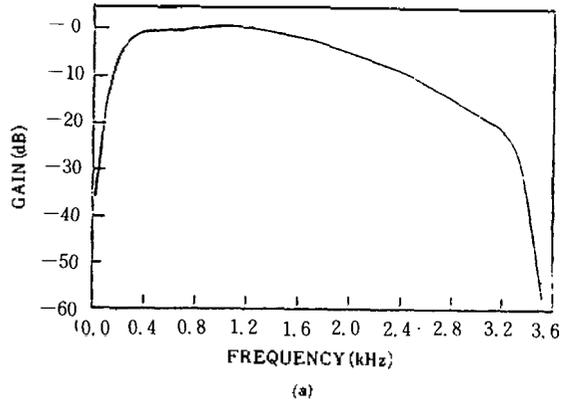


그림 4.1 (a) 전화선의 gain 특성 (b) 전화선의 delay 특성

4.2 실험 방법

Digital 음성 data로부터 LPC 계수를 추출할 때 spectral resolution을 높이기 위하여 Hamming window를 사용하였으며, preemphasis factor로는 0.95를 사용하였다. 그리고 frame 길이는 20ms이고 연결한 frame간의 겹치는 길이는 10ms로 하였다.

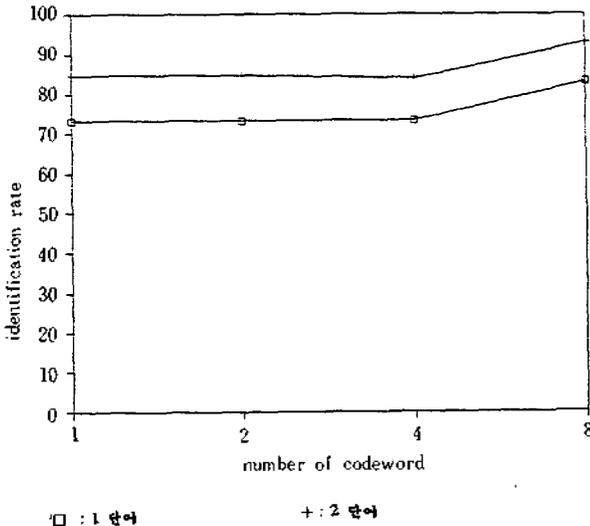


그림 4.2 Test 음성 data 길이 변화에 따른 인식율 변화

4.3 Computer simulation 결과

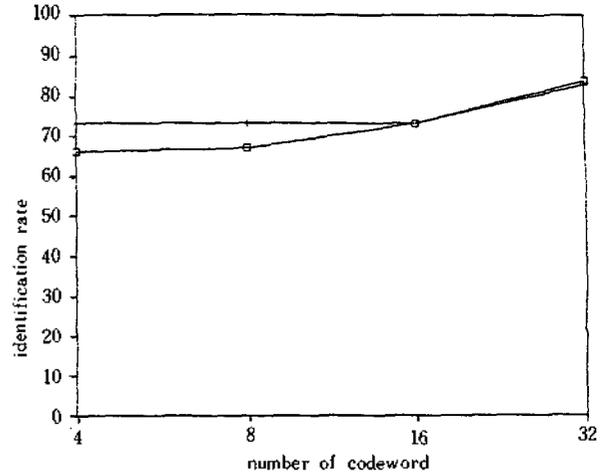
4.3.1 VQ와 모음 인식을 이용한 화자 인식 system의 인식 결과

그림 4.2에서는 모음 분리를 수동으로 했을 때, test 음성 data가 한 단어에서 두 단어로 변화함에 따라 인식율이 상승한다는 사실을 보여 주고 있다. 이러한 결과는, test 음성 data 양이 증가함에 따라 화자음성 특징에 관한 정보량이 증가하여 인식율이 상승함을 보여주고 있다.

모음분리를 수동으로 했을 때와 자동으로 했을 때의 인식율 차이를 표 4.1에 나타내었다. Codeword 수가 32이고 모음 인식을 수동으로 한 상태에서, 자동 모음 분리를 했을 때의 인식율이 75.5%이고, 수동 모음 분리의 경우 인식율은 83.0%이다. 자동

표 4.1 모음 분리 방법의 변화에 따른 인식율 변화

	자동 모음 분리	수동 모음 분리
인식율	75.5%	83.0%



□ : 모음 인식 없음 + : 모음 인식 있음

그림 4.3 모음 인식 과정의 유무에 따른 인식율 변화
모음 분리를 했을 때의 인식율이 수동 모음 분리의 경우보다 다소 떨어질 것을 볼 수 있다. 이러한 인식율의 하락은 자동 모음 분리의 정확도와 수동 모음 분리의 정확도 차이에 기인하는 것이다.

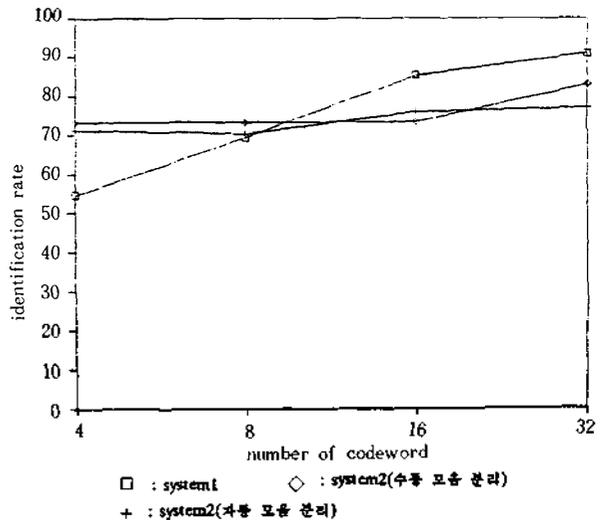


그림 4.4 system1과 system2의 인식율 변화

VQ와 모음 인식을 이용한 화자 인식 시스템과 VQ와 모음 분리를 이용하지만 모음 인식 과정이 없는 화자 인식 시스템의 성능을 비교한 것을 그림 4.3에 나타내었다. VQ와 모음 인식을 이용한 시스템을 system2이라고 하고, VQ와 모음 분리를 이용한 화자 인식시스템을 system3라고 하자. System3은 system2의 수동 과정에서 모음인식 부분이 인식율에

어느 정도 영향을 미치는가를 알아보기 위해서 만든 것이다. 그림 4.3에 나타난 바와 같이 system2와 system3의 인식율은 codeword 수가 작을 때는 다소 차이가 있지만 codeword 수가 증가함에 따라 거의 차이가 없는 것을 볼 수 있다. 이는 system2의 모음 인식 부분이 별로 큰 역할을 하지 못한다는 사실을 의미한다. 예를 들어 A라는 사람의 "아"모음이 B라는 사람의 "어"로 인식될 확률은 codeword 수가 증가함에 따라 매우 작아진다는 것을 의미한다.

4.3.2 VQ를 이용한 화자 인식 시스템과의 성능 비교

VQ를 이용한 화자 인식 시스템과 본 논문에서 제안한 화자 인식 시스템과의 인식율을 비교한 것이 그림 4.4에 나타나 있다. VQ를 이용한 화자인식 시스템을 system1이라고 하고, VQ와 모음 인식을 이용한 화자 인식 시스템을 system2라고 하자. System2는 codeword 수가 작을 때는 system1보다 성능이 좋다. 그러나 codeword 수가 증가함에 따라 system1보다 성능이 떨어진다. 이러한 현상이 나타난 이유는, system1의 경우 codeword 수가 작을 때는 그 만큼 음소의 표현 범위도 작을 수 밖에 없고 이 때문에 test 음성 data를 구성하는 음소중 상당수가 벡터 양자화하는 과정에서 short-time spectrum 특성이 많이 다른 codeword로 벡터 양자화되어 인식율이 상대적으로 낮은 것으로 생각된다. 그러나 codeword 수가 커짐에 따라 위의 mismatch 현상은 많이 줄어들고 모음에 있는 화자 음성 특징과 아울러 자음에 있는 화자 음성 특징을 이용하게 됨에 따라 인식율의 상승폭이 상대적으로 커진 것으로 생각된다. 반면 system2의 경우는 test 음성 data가 모음 부분으로 구성되어 있기 때문에 앞서 system1에서 발생했던 mismatch 현상은 상대적으로 적다. 따라서 codeword 수가 작을 때는 system1보다 높은 인식율을 기록했다. 그러나 test 음성 data가 모음 부분으로 한정되기 때문에 codeword 수가 증가하더라도 화자 음성 특징의 증가폭은 상대적으로 작을 수 밖에 없다. 따라서 codeword 수가 커지면 system2의 인식율은 system1에 비해 다소 떨어진다.

SNR이 20dB일때, system1의 인식율과 system2에 있어서 모음 분리를 자동으로 했을 때의 인식율과 수동으로 했을 때의 인식율을 비교한 것이 그림 4.5와 표 4.2에 나타나 있다. System1의 인식율과 모음 분리를 수동으로 했을 때의 system2의 인식율을 비교해보면 system2의 인식율이 전반적으로 높다는 것을 알 수 있다. 이러한 현상이 나타난 이유는 모음 부분이 다른 음소 부분에 비해 진폭이 커서 noise의 영향이 작기 때문이다. 그리고 그림 4.1에 표시된 것과 같은 전달 특성을 가진 전화선의 영향과 noise의 영향(SNR=20dB)을 고려할 때, system1의 인식율과 system2에 있어서 모음 분리를 자동으로 했을 때의 인식율과 수동으로 했을 때의 인식율을 비교한 것이 그림 4.6와 표 4.3에 나타나 있다. 앞서 언급한바와 같이, 여기에서 system1은 VQ만을 사용한 기존의 화자 인식 시스템을 의미한다. 그림 4.6으로부터 system2에서 모음분리를 수동으로 한 경우

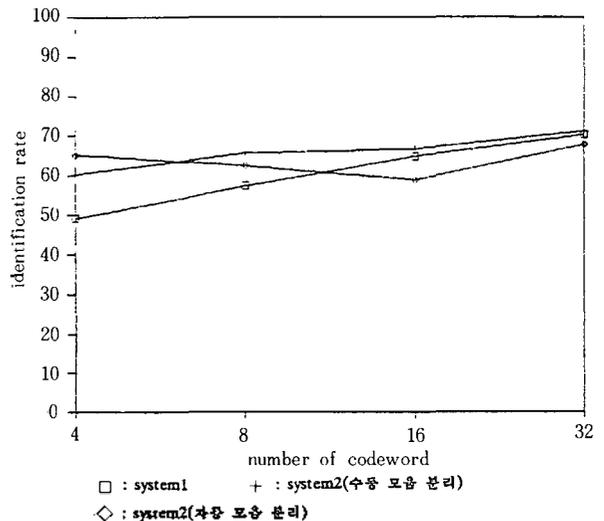


그림 4.5 SNR=20dB 일때, system1과 system2의 인식율 비교

표 4.2 SNR=20dB 일때, system1과 system2의 인식율 비교

	4	8	16	32
system1	49.1%	57.4%	64.8%	70.4%
system2 (수동 모음 분리)	60.2%	65.7%	66.7%	71.3%
system2 (자동 모음 분리)	64.8%	62.0%	58.3%	67.6%

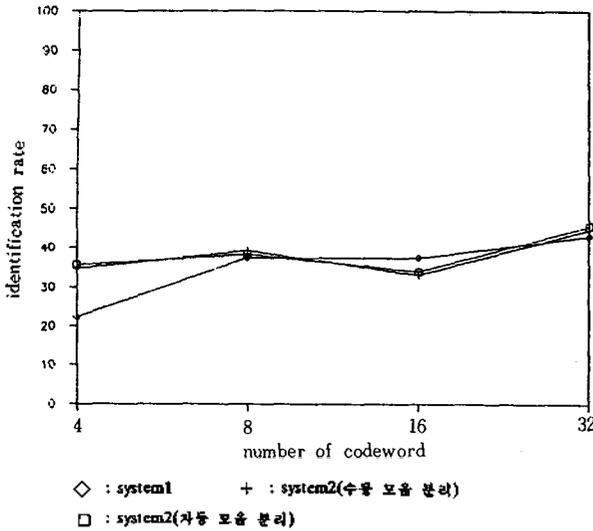


그림 4.6 전화선의 영향과 noise의 영향(SNR=20dB)을 고려할때, system1과 system2의 인식율 비교

표 4.3 전화선의 영향과 noise의 영향(SNR=20dB)을 고려할때, system1과 system2의 인식율 비교

	4	8	16	32
system1	22.3%	37.5%	37.5%	42.9%
system2 (수동 모음 분리)	34.8%	39.3%	33.0%	44.6%
system2 (자동 모음 분리)	35.7%	38.4%	33.9%	45.5%

의 인식율이 system1의 인식율보다 전반적으로 높다는 사실을 알 수 있다. 그리고 이 경우에는 system2에서 모음 분리를 자동으로 한 경우의 인식율도 system1의 인식율보다 높았다. 이와 같은 결과가 나타난 이유는 모음에 해당하는 부분이 다른 음소부분에 비해 전화선의 영향과 noise의 영향이 작기 때문이다.

5. 결 론

본 논문에서는 VQ와 모음 인식을 이용한 text-independent 화자 인식 시스템을 제안하고, 인식 결과를 얻었다. 그리고 기존의 VQ를 이용한 화자 인식 시스템과 여러가지 환경에서 인식 결과를 비교하였다.

VQ와 모음 인식을 이용한 화자 인식 시스템에서 모음 인식이 인식율에 어느 정도 영향을 미치는가를 알아보기 위해 VQ와 모음 분리만을 사용한 화자

인식 시스템과 인식결과를 비교하였다. 그 결과 codeword 수가 작을 때는 모음 인식과정이 효과가 있지만 codeword 수가 증가함에 따라 모음 인식은 별로 큰 영향을 주지 못함을 알 수 있었다. 그리고 기존의 VQ를 이용한 화자 인식 시스템과의 인식율 비교를 다음과 같은 3가지의 환경에서 수행해 보았다. Noise와 전화선의 영향이 없는 경우에, codeword 수가 작을 때는 기존의 VQ를 이용한 화자 인식 시스템의 성능이 좋지 않았으나, codeword 수가 커짐에 따라 성능이 좋아졌다. 그러나 noise의 영향(SNR=20dB)만을 고려하는 경우와 noise의 영향(SNR=20dB)과 전화선의 영향을 고려하는 2가지 경우에 대하여, VQ와 모음 인식을 이용한 화자 인식 시스템이 기존의 VQ를 이용한 화자 인식 시스템보다 전반적으로 대등하거나 좋은 성능을 나타내었다. 이것은 주어진 음성 신호에서 모음에 해당하는 부분이 다른 음소에 비해 noise의 영향이 작고, 전화선으로 인한 왜곡의 정도도 상대적으로 작기 때문이다.

VQ와 모음 분리 및 인식을 이용한 화자 인식 시스템은 기존의 시스템보다 음성 처리영역이 훨씬 작기 때문에, test 음성 data 양이 많아지고 codeword 수가 커지면 모음 분리와 모음 인식으로 인한 overhead를 고려하더라도 계산량이 작아 질 수 있을 것이다.

참 고 문 헌

1. J. D. Markel, B. T. Oshika, and A. H. Gray, "Long term feature averaging for speaker recognition," IEEE trans. on Acoust., Speech, and Signal Processing, Vol. 25, pp.330-337, August 1977.
2. F. K. Soong, A. E. Rosenberg, L. R. Rabiner, and B.H. Juang, "A vector quantization approach to speaker recognition," Proc. of the ICASSP 86, pp.17.5.1-17.5.4, 1986.
3. A. E. Rosenberg, and F. K. Soong, "Evaluation of a vector quantization talker recognition system in text independent and text dependent modes," Proc. of the ICASSP, pp.873-876, 1986.
4. J. T. Buck, D. K. Burton, J. E. Shore, "Text-dependent speaker recognition using vector quantization," Proc. of the ICASSP, pp.391-394, 1985.

5. A. H. Gray, Jr., J.D. Markel, "Distance measure for speech processing," IEEE trans. on Acoust., Speech, and Signal Processing, Vol. ASSP-24, No.5, pp. 380-391, October 1976.
6. J. G. Wilpon, L. R. Labiner, "A modified K-means clustering algorithm for use in isolated word recognition," IEEE trans. on Acoust., Speech, and Signal Processing, Vol. ASSP-3, pp. 587-594, No.3, June 1985.
7. H. Wakita, "Residual energy of linear prediction applied to vowel and speaker recognition," IEEE trans. on Acoust., Speech, and Signal Processing, Vol. 24, pp.270-271, June 1976.
8. M. R. Sambur, "Selection of acoustic feature speaker identification," IEEE trans. on Acoust., Speech, and Signal Processing, Vol. ASSP-23, No.2, pp.176-182, April, 1976.
9. M. Kasuya, H. Wakita, "An approach to segmenting speech into vowel and nonvowel-like intervals," IEEE trans. on Acoust., Speech, and Signal Processing, Vol. ASSP-27, No. 4, pp.319-327, August 1979.

△임 창 현



1963년 5월 18일생
 1986년 2월 서울대학교 공과
 대학 전자공학과 졸업
 1988년 8월 한국과학기술원
 전기및 전자공학과 졸
 업(M. S.)
 1989년 3월~현재 한국과학
 기술원 전기및 전자공
 학과 박사과정 재학중

△은종관 8권3호 참조

△이 황 수



1952년 9월 19일생
 1975년 2월 서울대학교공과
 대학 전기공학과 졸업
 1975년 1월~10월 현대조선
 중공업 주식회사 전장
 설계부
 1978년 8월 한국과학원 전
 기및 전자공학과(M. S.)

1983년 2월 한국과학원 전기및 전자공학과(Ph. D)
 1983년 3월~현재 한국과학기술원 전기및 전자공
 학과 조교수
 1984년 4월~1985년 5월 Post Doc. 연구원(Infor
 mation System Lab., Dept. of EE, Stan-
 dard University)