

論文94-31B-6-3

관성과 SOFM-HMM을 이용한 고립단어 인식

(Isolated word recognition using the SOFM-HMM and the Inertia)

尹 皙 鉉*, 鄭 光 宇*, 洪 光 錫**, 朴 炳 哲*

(Seok Hyun Yoon, Kwang Woo Chung, Kwang Seok Hong and Byung Chul Park)

要 約

본 논문은 한국어 단어 인식에 관한 연구로써, 특징 벡터 시퀀스에 관성을 도입하여 HMM(Hidden Markov Models)의 상태 천이를 안정화시키는 방법을 제안한다.

입력 특징 벡터의 확률분포를 고려하여 양자화 왜곡을 감소시키기 위해, 자율 학습 기능을 갖는 SOFM (Self Organizing Feature Map)을 벡터 양자화기로 사용하였다. 또한, 특징 벡터 시퀀스에 관성을 도입함으로써, SOFM상의 반응경로에서 확률분포가 중첩되는 것을 억제하여 HMM의 상태 천이를 안정화시킬 수 있었다.

본 논문에서 제안한 방법을 평가하기 위하여 50개의 DDD 지역명을 대상으로 인식 실험을 수행하였다. 그 결과, 특징 벡터의 시퀀스에 관성을 도입하므로써 7.4%의 인식율이 향상되었으며, 고정된 수의 뉴런을 갖는 SOFM에 대해 인식률의 큰 저하없이 인식 대상 단어의 모델수를 증가시킬 수 있었다.

Abstract

This paper is a study on Korean word recognition and suggest the method that stabilizes the state-transition in the HMM by applying the 'inertia' to the feature vector sequences.

In order to reduce the quantized distortion considering probability distribution of input vectors, we used SOFM, an unsupervised learning method, as a vector quantizer. By applying inertia to the feature vector sequences, the overlapping of probability distributions for the response path of each word on the self organizing feature map can be reduced and the state-transition in the HMM can be stabilized.

In order to evaluate the performance of the method, we carried out experiments for 50 DDD area names. The results showed that applying inertia to the feature vector sequence improved the recognition rate by 7.4% and can make more HMMs available without reducing the recognition rate for the SOFM having the fixed number of neuron.

* 正會員, 成均館 大學校 電子工學科
(Dept. of Elec. Eng., Seonggyungwan Univ.)

(Dept. of Information. Eng., Univ.)

接受日字 : 1993年 8月 17日

** 正會員, 濟州 大學校 情報工學科

I. 서론

현재 널리 이용되고 있는 음성인식 알고리즘으로는 크게 패턴 매칭방식을 이용한 DP 매칭 방식과 벡터 양자화 방식, 통계적 방식을 이용한 HMM 방식, 그리고 신경망 및 퍼지추론을 이용한 알고리즘등으로 분류될 수 있다. DP매칭의 경우, 인식률은 좋으나 계산량이 많아 대용량 어휘의 인식모델로는 적합하지 않으며, 신경망 모델의 경우는 Recurrent Network 과 최근들어 음소인식 모델로 많은 연구가 이루어지고 있는 TDNN이 대표적인 것으로 시변 패턴의 정보처리 능력은 뛰어나지만 학습시간이 길고 일단 학습이 끝나면 확장성이 없다는 단점이 있다. 또다른 형태의 신경망 모델로는 SOFM이 있는데, 2차원 배열의 SOFM은 음소 식별과 집단화에 효과적인 것으로 보고된 바 있으며^[1], 출력층 뉴런간 상호작용이 학습과정에 도입되므로써 SOFM을 벡터 양자화기로 사용할 경우 입력벡터의 확률분포를 고려할 수 있다는 장점을 가지고 있다. 기존에도 SOFM에 의한 음소인식의 연구가 이루어지긴 하였으나 음소식별의 처리능력에 비해 음소간 천이를 적절히 처리하지 못한 문제점을 나타내고 있는데^[2], 본 논문에서는 음소간 천이를 처리하기 위해 계산량의 부담이 적고 최근들어 인식률에서도 상당한 향상을 보이고 있는 HMM을 적용, SOFM을 벡터 양자화기로 사용한 HMM을 구성하였다. 이러한 인식 모델의 구성은 SOFM의 음소집단화 처리능력에 비교적 계산량이 적은 HMM의 잇점을 결합한 것으로 SOFM이 모든 음소의 특징들을 포함한다면 대용량 어휘의 인식기로 쉽게 확장될 수 있다.

문제는 “이러한 SOFM에 인식률의 저하없이 얼마나 많은 HMM모델을 형성할 수 있는가” 이다. HMM은 경쟁적 학습에 의한 모델이 아니기때문에 벡터 양자화기에 사용된 코드워드의 수를 증가시킨다 하더라도 한계가 있어서 더이상의 코드워드수의 증가는 계산량만 증가시킬뿐 구별 가능한 모델의 수를 인식률의 저하없이 증가시키기는 어렵다. 대개의 경우 특징 벡터의 시퀀스에 대한 벡터 공간상의 궤적은 국소구간에서 랜덤한 변동을 보이고 이러한 변동이 각각의 음소에 대한 확률분포를 중첩시키는 하나의 요인이 되기 때문에 특정 음성신호에 대한 특징 벡터 시퀀스의 안정성은 인식률에 중요한 인자로 작용하는데, 특히 상태간 천이 확률과 각 상태에서의 출력 확률로 특성화되는 HMM의 경우, 특징 벡터 시퀀스의 안정성은 일정수의 코드워드를 갖는 벡터 양자화기에 대해 얼마나 많은 모델을 인식률의 저하없이 형성할

수 있는가와 관련되기 때문에 보다 더 중요하다.

본 논문에서는 이러한 문제점을 보완하기 위하여, 입력 벡터의 확률분포가 고려되므로써 양자화 왜곡을 감소시키는데 적절한 SOFM을 벡터 양자화기로 사용하였으며, 특징 벡터의 시퀀스에 대한 궤적의 안정화를 위해 “관성”을 도입하여 특징 벡터 시퀀스에 대해 SOFM상의 확률 분포가 중첩되는 것을 억제하므로써, 일정수의 뉴런을 갖는 SOFM에 보다 많은 모델을 형성할 수 있도록 하였다.

본 논문에서 제안한 방법의 타당성을 검증하기 위하여 한국어 단어 인식을 수행하였다. 특징 파라미터로는 선형예측계수로 부터 얻어진 all-pole 모델의 스펙트럼상에서 나타나는 인접 pole들간의 중첩효과를 제거한 SGDS(smoothed group delay spectrum)을 이용하였고, 3인 발성한 53개의 단음절을 이용하여 SOFM을 학습하였으며, 50개의 DDD 지역명을 대상으로 HMM을 구성하여 한국어 단어인식을 수행하였다.

II. 음성신호의 특징파라미터 추출

음성인식에 있어서, 일반적으로 영점보다는 스펙트럼상에서 피크를 나타내는 극점이 더 중요한데, 선형 예측 분석으로부터 얻어진 all-pole 모델의 스펙트럼의 경우, 인접한 극점의 영향으로 극점을 독립적으로 추출해 내는데 어려움이 있다.

이러한 결점을 보완하기 위하여 SGDS(Smoothed Group Delay Spectrum)가 제시되었다.^[3] 이산적으로 정의된 스펙트럼의 위상 성분 $\Theta(\phi)$ ($1 \leq i \leq L$)에 대해 GDS(Group dDelay Spectrum)는 다음 식1과 같이 정의된다.

$$t(\phi_i) = -\frac{Q(\phi_i) - Q(\phi_{i-1})}{\phi_i - \phi_{i-1}} \quad (\text{식 1})$$

각 채널의 대역폭을 B 라할때 GDS τ_n 는 다음과 같다.

$$\tau_n(\phi, B) = -\sum_{k=1}^{\infty} \frac{\sin(kB)}{B} C_k \cos(k\phi) \quad (\text{식 2})$$

(C_k 는 캡스트럼 계수)

식2는 캡스트럼 계수에 sine window를 씌운 결과와 동일하다.^[4] 또한 캡스트럼계수의 푸리에 변환은 다음과 같다.

$$\log(H(\phi)) = \sum_{k=1}^{\infty} C_k \cos(k\phi) \quad (\text{식 3})$$

($H(\phi)$ 는 all-pole 모델의 전달함수)

식2를 식3에 나타난 캡스트럼 계수의 푸리에 변환과 비교할때, 식2는 캡스트럼의 낮은 Quefrequency 성분을 제거하므로써 스펙트럼의 피크가 강조되며 높은 Quefrequency 성분을 제거하므로써 스펙트럼의 작은 변동에 의한 과잉 강조가 제거될 수 있다.

본 논문에서는 16차 LPC Cepstrum에 $B = \pi / 32$ 를 적용하였고, 부가적 Smoothing 으로 다음과 같이 LPC 계수에 Exponential window 를 씌웠다.^[4]

$$\alpha_k = A^k a_k (a_k \text{는 LPC계수, } A=0.9) \quad (\text{식 4})$$

특징 파라미터 추출을 위한 블록 다이어그램은 그림 1과 같다.

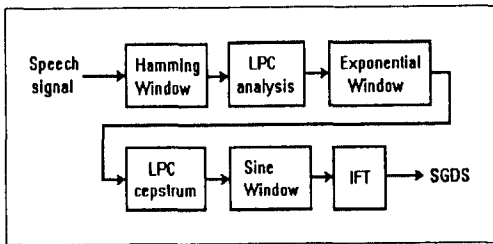


그림 1. SGDS의 계산
Fig. 1. Calculation of SGDS.

III. 인식 모델의 구성

본 논문에서 제안한 인식모델의 블록 다이어그램을 그림2에 나타내었다. 특징 추출부에서 계산된 특징 벡터의 시퀀스는 smoother에 의해 관성이 적용되어 smoothing처리 된후 SOFM에 입력되고, 그에 상응한 SOFM상의 반응경로를 확률적으로 처리하기 위해 HMM이 사용된다.

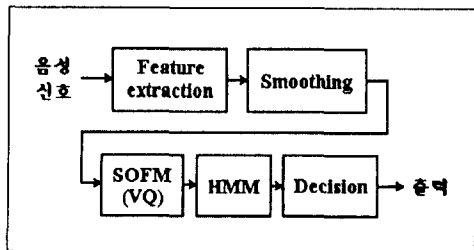


그림 2. 인식기의 블록 다이어그램
Fig. 2. Block diagram of the recognizer.

3-1. SOFM (Self Organizing Feature Map) Kohonen은 감각 신경계에서 흔히 볼 수 있는 뉴

런 간의 lateral interaction과 이에 관련된 receptive field의 개념을 단층 신경망 학습 알고리즘에 도입하므로써 자기조직적 ordering의 모델을 구성하였다.^[5] 즉, 일군의 감각 신경으로부터 입력된 패턴들은 자기 대뇌 피질의 특정 부위에 활성전위를 유발시키면서 lateral interaction에 의해 피질상 뉴런의 활성 패턴에서도 감각패턴들 간의 상관성이 그대로 유지되어 나타나는데 이러한 과정에 대한 모델로써 SOFM이 제시되었다.^[5] 실제로 receptive field내의 뉴런들은 그 중심 뉴런의 활성화에 대해 그림 3 과 같은 Mexican hat 형태의 lateral interaction을 하는데 이것을 학습 알고리즘에 적용하면 "Topology Preserved Mapping"의 효과로 특징 벡터의 궤적을 가시화할 수 있으며, 또한 입력 패턴의 확률분포가 고려되기 때문에 패턴의 집단화에 효과적인 수단을 제공한다.

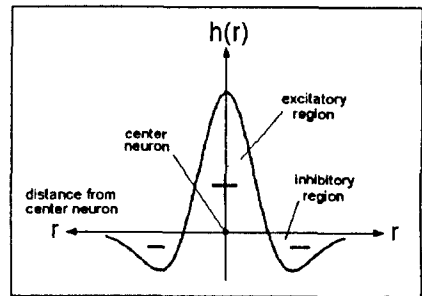


그림 3. Mexican hat 함수
Fig. 3. Mexican hat function.

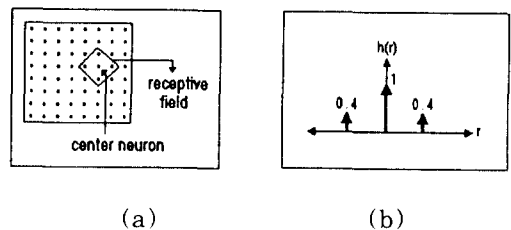


그림 4. 뉴런 배열(a)과 학습에 도입된 근사함수(b)
Fig. 4. Neural array(a) and approximation function(b).

본 연구에서는 그림4에 나타난 격자구조의 뉴런 배열에 lateral inhibition을 무시한 단순화된 lateral interaction을 가정하였다. 그림4(a)의 점들은 뉴런의 출력노드를 나타내며 입력노드와 입출력노드간 연결은 생략되어있다. 입력노드는 16개로 각 노드의 입력은 관성이 적용된 16채널의 SGDS를 사용하였으며, 점선으로

나타낸 사각형은 receptive field를 나타내는데, field 내의 뉴런간 상호작용은 그림3에 나타낸 Mexican cap 함수 대신 그림4(b)의 lateral inhibition을 무시한 단 순화된 근사 함수를 사용하였다.

SOFM의 학습 알고리즘은 다음과 같다.

초기화. 연결 강도를 랜덤한 값으로 초기화 한다.

단계 1. 주어진 입력 벡터 $X = \{X_1, X_2, \dots, X_N\}$ 에 대해 $C(i, j)$ 를 계산한다.

$$C(i, j) = \arg \min(\|X - W_{ij}\|) \quad (\text{식 } 5)$$

단계 2. $C(i, j)$ 를 중심 뉴런으로 갖는 receptive field N_c 내의 뉴런에 대해 다음과 같이 연결 강도를 조절한다.

$$W_{ij}(t+1) = W_{ij}(t) + \alpha(r, t) \cdot (X - W_{ij}(t)) \quad (\text{식 } 6)$$

for $C(i, j) \in N_c$

여기서, $\alpha(r, t) = \varepsilon(t) \cdot h(t)$
 $\varepsilon(t) = 0.8 \cdot \exp(-t/T)$
 $h(t)$: 그림4 b)의 근사함수

단계 3. 주어진 training set에 대해 단계1,2를 반복한다.

3-2. 벡터 양자화기로서의 SOFM

3-1절의 학습 알고리즘에서 단계2는 인식 모드에서의 동작을 나타내는데 lateral inhibition에 근거한 "winner takes all" 형태의 동작⁶⁾은, 매핑의 관점에서 볼때, 하나의 통계적 변수로 볼 수 있다. 즉 매핑 Φ_w 는 식7과 같이 표현한다.

$$\Phi_w : V \rightarrow A \quad (\text{식 } 7)$$

식7은 입력 벡터의 set로 구성되는 특징 벡터 공간 V 를 그에 상응한 코드워드로 구성되는 A 공간으로 매핑하는것으로 볼 수 있고, 여기서 특징 벡터 공간 상에서 벡터 시퀀스의 궤적은 SOFM상의 반응경로로 매핑된다.

입력 벡터의 확률분포를 $P(X)$ 라 하면 최적의 벡터 양자화기는 총 왜곡

$$\sum \|X_i - W_{s(i)}(t)\|^2 \cdot P(x_i) \quad (\text{식 } 8)$$

를 최소화하는 연결 강도에 의해 구현되는데, "gradient descent"에 근거하여 Linde, Buzo and Gray에 의해 제안된 학습 알고리즘⁷⁾에 관한 식은 다음과 같다.

$$W_{ij}(t+1) = W_{ij}(t) + \sum_{X \in S(x_i)} (X - W_{ij}(t)) \cdot P(X) \quad (\text{식 } 9)$$

일반적으로 입력벡터의 확률분포 $P(X)$ 는 알려지지 않기 때문에 식9는 계산될 수 없지만 충분히 작은 값의 ε 에 대해 식9를 다음의 식10으로 대치하여 반복하므로써 근사적 효과를 얻을 수 있다.

$$W_{ij}(t+1) = W_{ij}(t) + \varepsilon \cdot (X - W_{ij}(t)) \quad (\text{식 } 10)$$

이는 SOFM 학습 알고리즘의 단계2과 동일한 것으로, 박희의 청각 피질 형성에 대한 시뮬레이션에서 볼 수 있듯이 SOFM은 위와 같은 방식으로 입력 벡터의 확률분포가 고려되므로써 확률적으로 자주 나타나는 패턴의 범주에 대해서는 SOFM상의 영역이 확장되어 보다 많은 구획으로 나누어지는 효과를 볼 수 있다.⁸⁾

3-3. 관성(Inertia)

음성신호는 특징벡터들의 시퀀스로 표현되는데 각 프레임간 간격이 작을 경우 벡터 시퀀스는 벡터 공간 상에서 그 벡터가 가리키는 질점의 연속적 운동으로 볼 수 있다. 그러나 일반적으로 사용되는 특징벡터의 시퀀스에서는 안정구간의 인접한 프레임간에도 변동이 커서, 벡터 공간상에서 질점의 궤적에 대한 연속성을 가정하는 것에는 무리가 있다. 질점의 궤적에 대한 안정성은 HMM에 의한 인식률과 관련되므로, 질점의 운동에 관성을 주어 궤적의 랜덤한 변동을 제거하여 인식률의 향상을 가져올 수 있는데, 이는 특징 벡터의 시퀀스에 대한 궤적에 필터를 취하여 smoothing처리를 하는 것과 동일하다.

본 논문에서는 식11로 표현되는 autoregressive 필터를 사용하여 특징벡터의 시퀀스에 대한 궤적을 smoothing처리하였으나, 성도 파라미터의 변동에 비해 고주파로 분류될 수 있는 랜덤한 변동이 0의 평균을 갖는다고 가정하여 몇개의 인접프레임에 대해 moving average를 취하더라도 큰 차이는 없을 것으로 본다.

$$Y(t) = aX(t) + bY(t-1) + cY(t-2) \quad (\text{식 } 11)$$

여기서, $Y(t) = \{y_1(t), y_2(t), \dots, y_N(t)\}$

$$X(t) = \{x_1(t), x_2(t), \dots, x_N(t)\}$$

식11에서 $X(t)$ 는 필터의 입력으로 16채널의 SGDS이고, $Y(t)$ 는 관성이 적용된 SGDS이다.

필터링의 효과는 SOFM상에서 시퀀스의 각 벡터에 반응한 활성 뉴런을 추적하므로써 가시화할 수 있

는데 그림5에 관성을 적용했을 경우와 적용하지 않았을 경우의 SOFM상의 반응경로를 도시하였다.

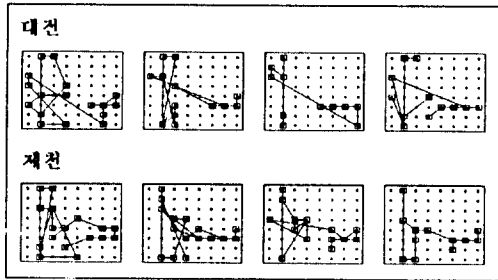
VI. 실험 및 결과

4-1. 음성 데이터 구성

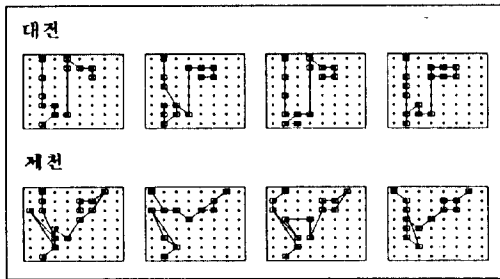
음성 데이터는 남성 화자 3인이 발성한 50개의 지역명 대해 각각 6회씩 발성하여 이중 3명 x 3회 x 50= 450개의 데이터를 HMM의 학습 데이터로 사용하였고 나머지 3 x 3 x 50= 450개의 데이터를 실험 데이터로 사용하였다.

SOFM의 학습은 학습시간의 단축을 위해 50개의 지역명에 포함된 단음절 53개를 따로 1회씩 발성하여 이를 SOFM의 학습 데이터로 사용하였다.

실험에 사용된 데이터는 DDD 지역명중 대도시명을 선정하였으며, 이를 표 1에 나타내었다.



(a) 관성을 적용하지 않은 경우



(b) 관성을 적용시킨 경우

그림 5. 입력 시퀀스에 대한 SOFM상의 반응경로
Fig. 5. Response path of input sequences on SOFM.

그림5는 “대전” 과 “제천”에 대해 각각 4회 발성한 음성 데이터의 SOFM상의 반응경로를 보여주는데, 그림5(a)는 관성을 주지 않은 시퀀스에 대한 반응경로이며 그림5(b)는 관성을 주었을 때의 시퀀스에 대한 반응경로를 나타낸다. 그림에서 알 수 있듯이, (a)의 경우 반응경로 자체가 불안할 뿐만 아니라 외관상으로 보더라도 “대전”과 “제천”의 반응경로에 대한 확률분포가 상당 부분 중첩되어 오인식률이 높게 나타난다. 그러나 (b)의 경우 같은 음성에 대해 반응경로가 뚜렷하게 구분되는 것을 볼 수 있다. 이러한 반응경로의 안정화는 특히 유사음성의 식별에 효과적인데, 안양-단양, 진해-김해, 영주-경주-청주 등과 같은 유사음성에 대해서도 반응경로의 안정화를 통한 인식률의 향상을 볼 수 있었다.

표 1. 음성 데이터 목록

Table 1. List of speech data.

온양	인천	원주	서울	구미
안양	김천	광주	수원	울진
단양	춘천	공주	성남	포항
밀양	제천	경주	평택	창원
부산	대전	영주	강릉	남원
마산	대구	청주	삼척	이리
서산	진해	충주	속초	여수
논산	김해	전주	목포	안동
군산	동해	진주	부여	천안
울산	해남	제주	홍성	정부

특징 파라미터 추출에 대한 분석 조건은 표 2와 같다.

표 2. 분석 조건

Table 2. Analysis condition.

샘플링 주파수	10 kHz
Cutoff Frequency of LPF	4.5 kHz
분석 프레임	25.6 msec
Overlapping Interval	12.8 msec
윈도우 함수	Hamming Window
SGDS 채널	16 채널

SOFM상의 반응경로를 smoothing처리하기 위해 사용한 필터의 계수값은 음소의 평균적 길이를 고려하여 적절하게 선택해야 하는데 본 논문의 분석조건에 대해, 초성자음의 경우 지속되는 프레임수는 3-8 프레임 정도이며 모음이나 종성자음(유성자음)의 경우 10-20 프레임 정도로 나타나기때문에 식11의 필

터 계수에 대해 $a=0.25$, $b=0.72$, $c=-0.1$ 로 각각 선택하였다. 필터계수의 이러한 선택은 실험적인 값으로, MA필터에 의한 반복실험에 의하면 필터의 탭수가 작을 경우 반응 경로의 안정화를 볼 수 없었으며 반면 필터 탭수가 너무 크면, 특히 짧은 음절에 대해, 음소간 천이가 불분명해져 인식률이 오히려 감소되었다. 그러므로 필터계수의 선택은 주어진 분석조건에 대해 반응경로의 안정화와 음소간 천이특성을 동시에 고려하여 적절히 선택되어야한다.

HMM에 의한 인식은 forward-backward 알고리즘을 사용하였고 zero의 확률을 방지하기 위해 HMM의 모든 파라미터 값에 대해 임의의 최저치 이하의 값을 최저치로 대치하여 계산하였으며 상태수는 단어당 음소수를 고려하여 7개로 하였다.

그림6에 본 논문에서 사용된 상태수가 7인 HMM의 구조를 나타내었다.

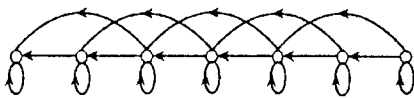


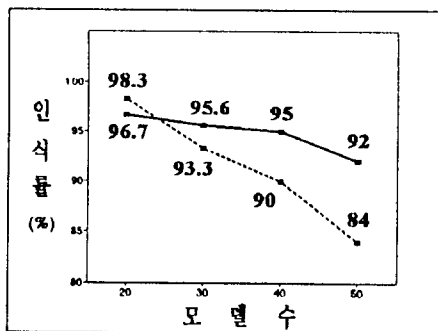
그림 6. HMM의 구조
Fig. 6. Structure of the HMM.

4-2. 실험 결과 및 고찰

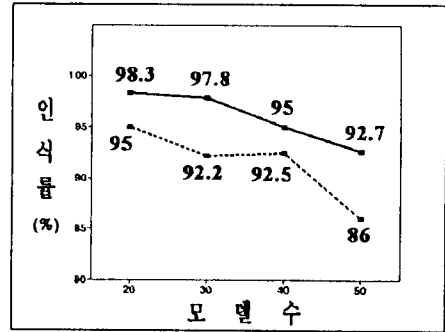
실험에서는 3명의 화자에 대해 각각 8x8의 SOFM으로 벡터 양자화기를 구성하여 비교될 HMM모델수를 변화시키면서 음성 파라미터에 관성을 적용했을 경우와 적용하지 않았을 경우의 인식률을 비교하였다.

모델수는 대도시 순으로 처음에 20개를 선택하였고, 이후 소도시 및 군 단위의 지역 순으로 10개씩 추가하면서 인식률을 비교하였다.

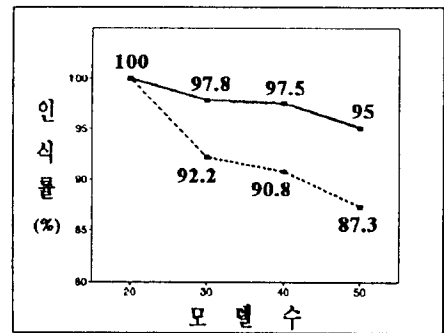
그림7에 실험 결과를 그래프로 표시하였다. 실선은 관성을 적용했을 때의 인식률을 나타내며 점선은 관성을 적용하지 않았을 때의 인식률을 나타낸다.



(a) 화자 A



(b) 화자 B



(c) 화자 C

그림 7. 모델수에 따른 인식 결과

Fig. 7. Recognition rate versus the number of Models.

표3에 대표적인 유사 단어군에 대해 관성을 적용했을 경우와 적용하지 않았을 경우의 오인식 횟수를 나타내었다.

표 3. 유사 단어군에 대한 오인식 예

Table 3. Examples of misclassification for similar words.

	관성을 적용하지 않은 경우	관성을 적용한 경우
영주 - 경주 - 청주	7	3
대전 - 제천	4	1
안양 - 단양	3	1
김해 - 진해	4	2
인천 - 김천	3	1

표1에 나타난 50개 지역명에 대해, 관성을 적용하지 않은 경우 평균 85.8 %의 인식률을 보인 반면,

관성을 적용한 경우 평균 인식률 93.2 %로 7.4 % 향상된 인식률을 보였다. 그리고 모델수의 증가에 따른 인식률의 감소를 볼때, 관성을 적용하지 않은 경우 0.4 %/model로 관성을 적용한 경우의 0.17 %/model과 비교 해 보면 관성을 적용하므로써 인식 대상 단어의 모델수가 증가함에 따른 인식률의 하락이 감소되었음을 알 수 있다. 이는 관성을 적용하지 않았을 경우 그림 5a)에서 볼 수 있듯이, 같은 음성 신호에 대해 SOFM상의 뉴런의 반응경로가 불안정하므로써 생기는 HMM모델간 확률분포의 중첩이 심하게 일어났기 때문이며, 반면에 관성이 적용되었을 경우 이러한 확률분포의 중첩을 억제시켜 인식률의 큰 저하없이 동일한 벡터 양자화기에 대해 더 많은 HMM모델을 형성 할 수 있었다.

V. 결 론

본 논문에서는 SOFM을 벡터 양자화기로 사용하여 HMM을 구성하였고, 음성 파라미터의 시퀀스에 대한 SOFM상의 반응경로를 안정화시키기 위해 "관성"을 도입하였는데, 인식 실험 결과 다음과 같은 결론을 얻을 수 있었다.

첫째, 관성의 도입을 통해 특징 벡터 시퀀스의 확률분포가 중첩되는 것을 억제하므로써 50개 지역명에 대해 7.4 %의 인식률의 향상을 가져왔으며, 둘째로는 이와 등가적으로 고정된 수의 뉴런을 갖는 SOFM에 대해 인식 대상 단어의 수가 증가함에 따른 인식률의 급격한 하락을 방지할 수 있었다.

이러한 구성의 인식모델은, 일단 SOFM이 모든 음소의 특징들을 포함한다면, 기존 SOFM에 새로운 단어의 HMM 모델을 쉽게 추가, 확장할 수 있으며, 또한 많은 학습 데이터를 필요로 하는 HMM의 학습

에 있어서, 반응경로의 안정화와 SOFM에 의한 출력 확률의 보간을 통해 적은 학습 데이터만으로도 신뢰성있는 모델을 구성할 수 있다는 장점이 있다.

參 考 文 獻

- [1] T. Kohonen "The Neural Phonetic Typewriter", Computer pp 11-22 march 1988.
- [2] 임은진, 안제성, 김기석, 황희용 "Self Organizing Feature Map과 IAC모델을 이용한 단어인식", 제9회 음성통신 및 신호처리 워크샵 논문집 pp 237-241. 1992.
- [3] F. Itakura and T. Umezaki, "Distance Measure for Speech Recognition based on the Smoothed Group Delay Spectrum", ICASSP vol.3 pp 1257-1260 1987.
- [4] Furui and Sondhi "Advances in Speech Signal processing", Dekker pp 419-452.
- [5] T. Kohonen "Self Organization and Associated Memory", 1987 pp 119-157.
- [6] Warsserman "Neural Computing" Van Nostrand Reinhold, pp 64-71.
- [7] Linde, Buzo and Gray, "An Algorithm for Vector Quantizer Design" IEEE Trans. Com. vol.com-28 no.1 January 1980.
- [8] H. Ritter, T. Martinitz and K. Schul-ten "Neural computation and Self Organizing Maps" Addison-Wesley, pp 81-98.

著者紹介



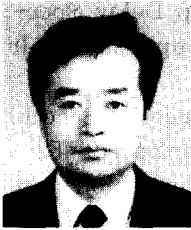
尹 哲 鉉(正會員)

1966年 12月 1日生. 1992年 2月 성균관대학교 전자공학과 공학사. 1994年 3月 ~ 현재 성균관대학교 대학원 전자공학과 석사과정. 주관심 분야는 음성 및 신호처리, 신경회로망 등임.



鄭 光 宇(正會員)

1966年 12月 1日生. 1989年 2月 성균관대학교 대학원 전자공학과 공학석사. 1992年 3月 ~ 현재 성균관대학교 대학원 전자공학과 박사과정. 주관심 분야는 음성 및 신호처리, 신경회로망 등임.



洪 光 錫(正會員)

1959年 2月 7日生. 1985年 2月 성균관대학교 전자공학과 공학사. 1988年 2月 성균관대학교 대학원 전자공학과 공학석사. 1992年 2月 성균관대학교 대학원 전자공학과 공학박사. 1990年 3月 ~ 1993年 2月 서울보건전문대학 전산정보처리학과 전임강사. 1993年 3月 ~ 현재 제주대학교 정보공학과 전임강사. 주관심 분야는 음성 및 신호처리, 패턴인식 등임.



朴 炳 哲(正會員)

1930年 4月 30日生. 1957年 9月 서울대학교 통신공학과 공학사. 1975年 2月 인하대학교 대학원 전기공학과 공학박사. 1980年 9月 일본 동경대학 외국인 연구원(1년간). 1972年 3月 ~ 현재 성균관대학교 전자공학과 교수. 주관심 분야는 ISDN, 음성 및 신호처리 등임.