

준연속분포 HMM을 이용한 한국어 단어인식

Korean Word Recognition Using Semi-continuous Hidden Markov Models

조 병 서*, 이 기 영*, 최 갑 석**

(Byung Seo Cho, Ki Young Lee, Kap Seok Choi)

요 약

본 논문에서는 HMM의 이산분포를 연속분포로 근사시키는 준연속분포 HMM에 의한 한국어 단어인식에 관하여 연구하였다. 이 모델의 생성과정에서는 입력벡터의 출력확률을 혼합 다차원 정규분포로 가정하여 입력벡터의 확률함수와 코드워드의 심볼출력을 선형결합하므로써, 연속분포 모델로 근사시켰으며, 단어인식과정에서는 생성모델에 의해 이산분포 모델에서 발생하는 양자화 왜곡을 감소시키므로써 인식률을 향상시켰다.

이 방법을 평가하기 위하여 DDD 지역명을 대상으로 이산분포 HMM과 준연속분포 HMM의 비교실험을 수행하였다. 그 결과, 준연속분포 HMM에 의하여 이산분포 HMM보다 향상된 인식률을 얻을 수 있었다.

ABSTRACT

This paper studies on the Korean word recognition using the semi-continuous hidden Markov models(SCHMM) which approximate the discrete hidden Markov models(DHMM) to the continuous hidden Markov models(CHMM).

In generating of this model, the output probability of an input vector, which is assumed to the mixture continuous Gaussian distribution, is linearly combined with the probability function of each input vector and the symbol output of codewords in order to approximate this to the CHMM. In the word recognition, this model recognizes words more correctly than the DHMM because of reducing a quantized distortion due to the DHMM.

To evaluate this model, comparing experiments between DHMM and SCHMM are carried out for DDD area names. The results show that the SCHMM has a higher recognition rate than the DHMM.

I. 서 론

사람과 사람의 정보전달에서 음성은 대단히 중요한 역할을 하고있다.^[1] 그러나, 음성은 개인의 차이를 뿐만 아니라 같은 사람이 동일한 내용을 발성하여도 음성의 특징은 변동하며, 문장 중의 음성은 문맥

의 영향을 크게 받는다. 이에 따른 음성인식 기술에서 요구되는 과제는 (1) 불특정화자 (2) 잡음영향에 대한 강화 (3) 인식 대상어휘의 증가 (4) 연속음의 인식등이며, 이들을 목적으로 한 국가적 프로젝트^[2] ^{[3][4]}에서는 전체적으로 큰 진보가 필요로 하게 되었고, 새로운 인식방법이 요구되었다. 이와같은 배경에서 연구된 HMM은 음성변동을 통계적으로 처리하고, 이 통계량을 확률 형태의 모델에 반영하여 음성을 인식하는 방법이다. 이러한 HMM은 크게 이산적

*명지대학교 전자공학과

**명지대학교 정보통신공학과

접수일자: 1992년 10월 26일

인 심볼에 대해서 출력확률이 정의된 이산분포 HMM^[5]과 연속적인 심볼에 대해서 출력 확률이 정의된 연속분포 HMM^[6]으로 나눌 수 있다. 여기서, 이산분포 HMM 방법은 입력으로 사용하는 벡터 양자화가 연속적인 음성의 특징량을 이산화하기 때문에 양자화 왜곡이 발생하는 단점이 있는 반면에 연속분포 HMM 방법은 양자화 왜곡은 없앨 수 있으나, 계산량이 매우 많아진다는 단점이 있다. 이러한 두 모델에서의 문제점을 보완하기 위하여, 본 연구에서는 이산분포 HMM의 출력 확률에 보간계수를 도입하여 연속분포 HMM으로 근사시킨 준연속분포 HMM에 의한 단어인식 방법을 제시하고자 한다. 여기서의 준연속분포 HMM은 각 입력 특징벡터에 대한 출력확률을 각 코드워드를 평균벡터로 하는 다차원 정규분포로 가정하고, 이에 대한 입력의 함수값과 학습에 기초한 HMM 파라미터 추정에 의해서 얻어진 각 코드워드에 대한 심볼 출력확률을 선형결합하여 표현하며, 이 모델을 이용하여 음성인식을 수행한다.

본 연구의 준연속분포 HMM을 이용한 인식방법의 우수성을 확인하기 위해 28개의 DDD 지역명을 사용하여 코드워드의 수와 상태의 수를 변화시키면서 이산분포 HMM에 의한 인식실험을 하였으며, 가장 좋은 인식률을 나타내는 코드워드의 수와 상태의 수를 준연속분포 HMM에 적용하여 인식률을 비교 실험한 결과, 그 우수성을 확인하였다.

II. 음성신호의 특징추출

음성의 발생과정을 모델링하는 것은 사람이 음성을 발생시킬 때 각각의 발음이 성대로 부터 성도를 거쳐 입술에 이르는 과정을 모델링하고 이것으로부터 특징 파라미터를 구하는 방법이다. 이러한 방법 중 가장 대표적인 방법인 all-pole 모델^[7]로 모델링하는 것으로써 특징 파라미터가 잘 알려진 LPC(linear predictive coding)계수^[7]이다.

음성 특징 파라미터를 인식 시스템에 적용시킬 때 연속적인 값을 그대로 사용할 수 있지만, 이를 심볼로 변환하여 처리할 수도 있다. 이와같이 이산 심볼을 사용하면 음성 신호의 정보량을 감소시키게 되고, 결국 전체 계산량을 크게 줄일 수 있는 이점이 있다. 신호의 부호화로 많이 쓰이고 있는 데이터 압축 기법인 벡터 양자화방법^[8]을 이용하여 얻어진 이산 심볼을 본 연구에서 사용하는 인식방법인 HMM의

입력으로 사용하였다. 이 벡터 양자화 방법에서 초기 중심벡터를 잡는 방법은 여러가지 있으나, 본 연구에서는 계산 시간과 성능면에서 가장 효과적인 방법으로는 분할방법인 LBG 알고리즘^[9]을 이용하였다. 이 방법은 먼저 전체벡터의 평균을 취해 중심벡터를 계산하고, 중심벡터를 적당한 방법으로 두 개의 벡터로 분할하는 것이며, 그 다음 단계로는 K-means 알고리즘^[10]을 통해 새로운 중심벡터를 계산하게 된다. 실제로 학습벡터 $x(n)$ ($1 \leq n \leq M$)이 주어졌을 때, LBG 알고리즘에 의해서 코드북을 생성하는 과정은 그림 1과 같다.

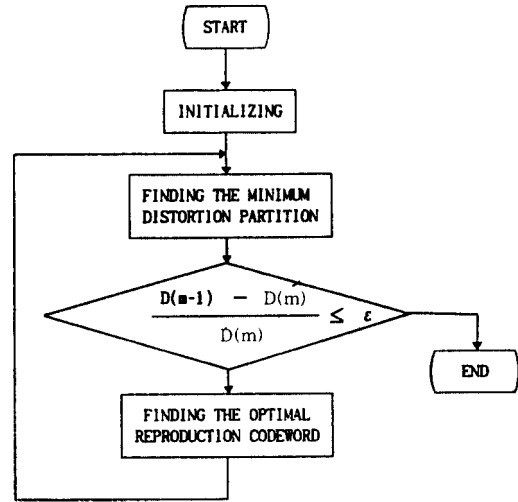


그림 1. LBG 알고리즘
Fig 1. The LBG algorithm

III. 준연속분포 HMM에 의한 단어인식

3-1. HMM의 정식화

HMM은 출력 심볼에 의해서 한결같이 상태 천이가 결정되지 않는 의미에서 비결정 유한상태 오토마타로서 정의할 수 있다. 이러한 HMM은 상태 천이

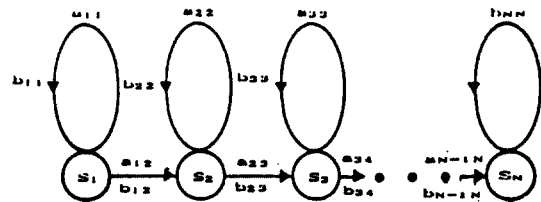


그림 2. left-to-right 모델
Fig 2. left-to-right model

가 유일한 것은 아니다. 즉, HMM은 출력심볼 시계열을 모델에 반영할수 있어도 상태전이 경로는 유일하게 결정되지 않는다. 관측이 가능한 것으로 출력심볼 시계열만이 존재하므로 HMM이라 부르며, 음성 인식에서 사용하는 HMM은 left-to-right^[5] 모델이라 부르며, 그림 2와 같이 나타낸다.

HMM 원소는 다음과 같이 특징지운다.

- 상태의 수 $N: S = \{S_1, S_2, \dots, S_N\}$
- 상태의 집합: $Q = \{q_1, q_2, \dots, q_N\}$
- 관측 심볼의 수 $M: V = \{V_1, V_2, \dots, V_M\}$
- 관측 시계열의 길이: $t = 1, 2, \dots, T$
- 상태 전이 확률 분포 $A = \{a_{ij}\}$ 가 다음과 같다.

$$a_{ij} = P\{q_{t+1} = S_j | q_t = S_i\}, \quad 1 \leq i, j \leq N \quad (4)$$

· 상태 j 에서 관측 심볼 확률 분포는 $B = \{b_j(k)\}$ 이다. 여기서,

$$b_j(k) = P\{V_k \text{ at } t | q_t = S_j\}, \quad 1 \leq j \leq N, 1 \leq k \leq M \quad (5)$$

- 초기 상태 분포 $\pi = \{\pi_i\}$ 는 다음과 같다.

$$\pi_i = P\{q_1 = S_i\}, \quad 1 \leq i \leq N \quad (6)$$

이러한 HMM을 정의하는데 세가지 기본적인 문제점을 해결하여야 한다. 그 첫째는 평가 문제인데, 이는 전향 알고리즘(forward algorithm)을 이용한다. 두번째는 관측 시계열이 주어졌을 때, 상태계열을 선택하는 복호화 문제로서, 이것은 Viterbi 알고리즘을 이용하여 해결할 수 있다. 세번째 문제는 학습에 관한 문제로서, 최적의 모델링을 하기 위하여 각 파라미터를 조정하는 것으로서, 이것은 Baum-Welch 알고리즘을 이용하여 수행한다. 또한, HMM을 실제로 모델링하는데는 또 다른 몇 가지 문제점이 제기된다. 첫째로, HMM의 파라미터를 어떻게 초기화할 것인가라는 문제이고, 두번째로, 관측의 수 T 가 증가함에 따라 확률의 값이 지수함수적으로 감소하여 언더플로우(underflow)가 발생하는 문제로서, 이것은 변수의 값들을 컴퓨터의 동적범위(dynamic range)에 있도록 스케일링을 하는 문제이며, 세번째는 훈련시키는 데이터 양이 충분히 많지 않을 경우, 훈련과정에서 어떤 심볼이 모델의 어느 상태에는 나타나지 않는 것으로 추정되었더라도, 시험과정 중에서 이 심볼이 나타나는 경우이다. 이에 대한 평활화 방법은 확률값들을 평활화함으로써 훈련되지 않은 심볼이 나

타날 가능성을 고려한다.

3-2. 준연속분포 HMM

준연속분포 HMM에서는 우선 HMM 파라미터의 추정에 앞서서, 인식단어 전체에서 최적으로 되는 것처럼 설계된 단어 공통의 코드북을 미리 준비한다. 각 입력 특징벡터에 대한 출력확률은 각 코드워드를 평균벡터로 하는 다차원 정규분포에 대한 입력의 함수값과 학습에 기초한 HMM 파라미터 추정에 의해서 얻어진 각 코드워드에 대한 심볼 출력확률과의 선형결합으로 주어진다. 각 코드워드를 평균벡터로 하는 다차원 정규분포의 공분산 행렬은 HMM 파라미터의 추정시에 각 단어에서 독립이고 각 모델내에서는 전체 상태전이에 공통인 것으로 해서 구한다.^[11]

일반적인 이산분포 HMM, 준연속분포 HMM에 있어서 입력 특징벡터에 대한 출력확률 산출법의 특징을 표 1에 나타냈다. 이 표에서 나타난 것 같이 준연속분포 HMM은 이산분포 HMM을 각 코드워드의 출력확률에 대한 모간계수의 도입에 의해서 확장한 것, 또는 혼합 정규분포에 의한 연속분포 HMM에서 평균벡터 및 공분산 행렬을 공통화한 간략화 방법으로 생각할 수 있다.^[12]

표 1. 각 HMM에서 출력확률 산출방법의 특징
Table 1. Characters of output probabilities for each HMM

모 델	출력확률 산출방법	모간 방법
이산분포 (DHMM)	$b_{ij}(O_k) : k = \operatorname{argmin}_m d(X_i, O_m)$	없음
준연속분포 (SCHMM)	$W_{ij}(X_i) = \sum_m N(X_i, \mu_m, \Sigma_m) b_{ij}(O_m)$	다차원 정규분포

본 연구에서는 HMM 파라미터의 추정과 코드북 설계를 분리하므로, 계산량의 증대를 최소한으로 억제하는 최적 코드북 설계법의 하나로써, 전체 단어에 영향을 미치는 입력벡터를 하나의 혼합 다차원정규분포에 따르는 것으로 가정하였으며, 이 가정 위에서 코드북을 maximum likelihood로 추정하는 방법인 혼합분포에 의한 최적 코드북 설계법^[14]을 도입하였다.

이 설계법에서는 평균벡터가 전체단어에 공통인 것에 중점을 두고, 전체 단어를 평균한 혼합 분포의 평균벡터를 구하기 때문에, 이 방법은 근사적인 의미

에서 준연속분포 HMM에 최적인 코드북을 설계하는 방법이다.

먼저, 어떤 코드북 B의 각 코드워드 V_k 를 평균벡터로 하는 다차원 정규벡터 $N(X_t, V_k, \Sigma_k)$ 과 혼합계수 λ_k 에 의한 혼합분포에 특징벡터의 집합 $X = \{X_1, X_2, \dots, X_T\}$ 가 나타나는 확률 $P(X|B)$ 는 식(7)과 같이 정의한다.

$$P(X|B) \equiv \prod_{t=1,T} P(X_t|B) = \sum_{k=1,M} P(X, k|B) \quad (7)$$

$$P(X_t|B) \equiv \sum_{k=1,M} \lambda_k N(X_t, V_k, \Sigma_k) \quad (8)$$

$$P(X, k|B) \equiv \prod_{t=1,T} \lambda_{kt} N(X_t, V_{kt}, \Sigma_{kt}) \quad (9)$$

$$\text{단, } \sum_{k=1,M} \lambda_k = 1, N_k \geq 0$$

$$k = \{k_1, k_2, \dots, k_T\}, 1 \leq k_t \leq M$$

$B = \{\lambda_k, V_k, \Sigma_k\}$: 코드북

M : 코드북 사이즈

이 likelihood 확률 $P(X|B)$ 를 최대로 하는 혼합분포, 또는 혼합계수 λ_k , 코드워드 V_k , 공분산 행렬 Σ_k 을 구한다면, 준연속분포 HMM에 최적한 코드북을 얻을 수 있다.

$$\lambda_k' = \frac{\sum_{t=1,T} \{\lambda_k N(X_t, V_k, \Sigma_k) / \text{Pr}(X_t|B)\}}{T} \quad (10)$$

$$V_k' = \frac{\sum_{t=1,T} \{\lambda_k N(X_t, V_k, \Sigma_k) X_t / \text{Pr}(X_t|B)\}}{\sum_{t=1,T} \{\lambda_k N(X_t, V_k, \Sigma_k) / \text{Pr}(X_t|B)\}} \quad (11)$$

$$\Sigma_k' = \frac{\sum_{t=1,T} \{\lambda_k N(X_t, V_k, \Sigma_k) [X_t - V_k][X_t - V_k]^T / \text{Pr}(X_t|B)\}}{\sum_{t=1,T} \{\lambda_k N(X_t, V_k, \Sigma_k) / \text{Pr}(X_t|B)\}} \quad (12)$$

전체 단어에 미치는 학습 데이터를 이용해서, 적당한 초기치에서 식 (10), (11), (12)를 이용하여 실행하는 것에 의해서 최적한 분포, 또는 평균 벡터 V_k' 를 코드워드로 하는 준연속분포 HMM에 적당한 코드북을 설계할 수 있다. 이와 같이 설계된 코드북을 전체 단어 공통의 심볼의 집합 $\{V_1, V_2, \dots, V_M\}$ 으로 하여 상태 i 에서 상태 j 로의 천이에 대해서, 상태 천이행렬을 $\{a_{ij}\}$, 코드워드 V_k 에 대한 심볼 출력확률들

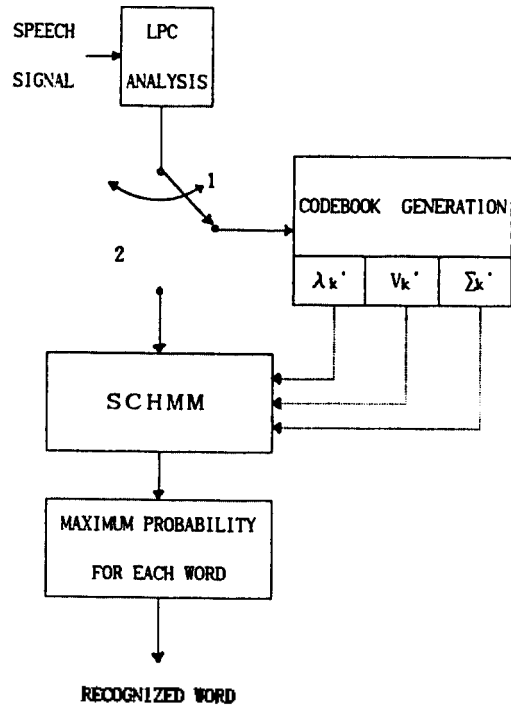
$\{b_{ij}(V_k)\}$, 시가 t 에서 입력특징벡터 X_t 에 대한 심볼 출력확률밀도를 $\{W_{ij}(X_t)\}$, (단, $i, j = 1, 2, \dots, N$)으로 하며, 이들 관계는 식(13)과 같다.

$$W_{ij}(X_t) = \sum_{k=1,M} R(X_t|V_k) b_{ij}(V_k) \quad (13)$$

이때, 준연속분포 HMM에서는 입력벡터 X_t 에 대한 심볼 출력확률 밀도를 상태 천이에 독립인 보간함수 $R(X_t|V_k)$ 로 하여 식(14)의 함수로서 정규분포를 가정하며, 식(13)과 같이 심볼출력 확률밀도를 구한다.

$$R(X_t|V_k) = \frac{1}{(2\pi)^{n/2} |\Sigma_k|^{1/2}} \exp\left\{-\frac{1}{2} (X_t - V_k)^T \Sigma_k^{-1} (X_t - V_k)\right\} \quad (14)$$

여기에서 V_k 는 코드워드이고, $(X_t - V_k)^T$ 는 전치행렬을 나타내며, Σ_k 는 공분산 행렬로서 $\Sigma_k = [(X_t - V_k)$



- 1 : Training
- 2 : Testing

그림 3 준연속분포 HMM에 의한 단어인식 구성도
Fig 3. Block diagram of word recognition using semi-continuous HMM

$(X_t - V_k)^2$ 이다.

보간함수로서 각 코드워드를 평균벡터로 하는 다차원 정규분포 $N(X_t, V_k, \Sigma_k)$ 를 가정한다면, 혼합 연속분포 모델의 경우와 유사한 식으로 되고, 공분산 행렬 Σ_k 를 Baum-Welch 학습 알고리즘에 의해서 식 (15)와 같이 근사적으로 추정할 수 있다.

$$\Sigma_k^+ = \frac{\left\{ \sum_{i=1}^N \sum_{k=1}^K \sum_{j=1}^J \alpha_{t-1}(i) a_{ij} b_{ij}(V_k) N(X_t, V_k, \Sigma_k) \beta_t(j) [X_t - V_k][X_t - V_k]^T \right\}}{\left\{ \sum_{i=1}^N \sum_{k=1}^K \sum_{j=1}^J \alpha_{t-1}(i) a_{ij} b_{ij}(V_k) N(X_t, V_k, \Sigma_k) \beta_t(j) \right\}} \quad (15)$$

단, α_t, β_t 는 전향, 후향확률.

이와 같은 보간함수는 각 단어내에서는 상태 천이에 의하지 않고 코드북에 의해서 결정되기 때문에, 인식에는 각 단어에서 입력 특징시계열에 대해서 1회만 보간계수를 산출해도 충분하며, 단어마다, 상태 천이마다 분포함수의 값을 산출할 필요가 있는 혼합 연속분포 모델보다 계산량을 줄일 수 있다. 그림 3은 준연속분포 HMM에 의한 단어인식 방법의 구성도이다.

IV. 실험 및 고찰

4-1. 음성 데이터 구성

본 연구의 음성인식시스템에서 사용할 음성 데이터 베이스는 학습에 사용될 학습 데이터와 인식시 사용될 시험 데이터로 분류된다.

데이터 베이스는 방음처리가 되지 않은 실내에서 남성 2인(MA, MC)과 여성 1인(WB)의 화자에 의해 발음한 DDD 지역명 28개 (직할시 이상 6개와 경기도 지역명 22개)를 3인×28개×5회=420개의 음성 데이터로 구성된다. 이 중에서 HMM에서 사용하기 위한 학습 데이터는 각 화자가 3회씩 발성한 지역명 3인×

표 2. 분석 조건
Table 2. Analysis conditions

Sampling frequency	10 [kHz]
Cut-off frequency of LPF	3.4 [kHz]
A/D resolution	12 [bits]
Analysis frame length	20 [ms]
Overlapping interval	10 [ms]
Window function	Hamming window
LPC order	14

28개×5회=252개를 사용하였으며, 나머지 2회씩 발성한 것을 HMM에서 시험 데이터로 하였다.

데이터 베이스의 구성을 위한 분석조건은 표 2에 나타내었다.

4-2. 인식 실험 및 고찰

본 절에서는 이산분포 HMM에 의한 인식 방법과 본 연구에서 제안한 준연속분포 HMM에 의한 인식 방법을 비교하기 위하여 DDD 지역명을 대상으로 인식 실험을 수행하였다. 또한, 코드워드의 수를 변화시키면서 인식 실험을 하였으며, HMM 상태의 수를 변화시키면서 실험하였다.

4-2-1. 이산분포 HMM에 의한 인식 실험

음성 인식 시스템의 표준 패턴을 구성하기 위한 각 단어에 대해 각 화자가 3번씩 발성한 음성을 이용하였다. 코드북의 생성은 LBG 집산화 알고리즘을 사용하였으며, 코드북 사이즈는 전체 단어에 대하여 코드워드의 수를 32, 64, 128로 하여 생성하였다.

각 단어에 대한 인식 알고리즘으로는 Viterbi 알고리즘을 사용하였으며, 각 화자에 대해 최저값 보정을 하여 심볼 출력확률의 평활화를 하여 인식 실험을 하였으며, 인식 실험에서 사용된 시험 데이터는 3명의 화자가 각 단어에 대해 두번씩 반복 발성한 음성을 사용하였기 때문에, 한 화자에 대해 56번씩 인식 실험한 것이 된다.

준연속분포 HMM은 이산분포 HMM을 연속분포 HMM에 근사시킨 모델이기 때문에 준연속분포 HMM에 적당한 상태의 수와 코드워드의 수는 이산분포 HMM의 경우와 같아야 할 필요가 있다. 따라서, 본 연구에서는 HMM의 입력으로 사용한 벡터 양자화 코드워드의 수를 변화시키면서 인식결과를 실험하였으며, 그 결과를 표 3에 나타내었다. 또한,

표 3. 코드워드에 따른 이산분포 HMM의 인식률(상태=8)
Table 3. Recognition rate of DHMM for codewords (state=8)

코드워드 수 화 자	[%]		
	32	64	128
MA	82.1	92.9	60.7
WB	78.6	87.5	48.2
MC	92.9	96.4	73.2
AVG.	84.5	92.3	60.7

표 4. 상태 수에 따른 이산분포 HMM의 인식률(코드워드=64)

Table 4. Recognition rate of DHMM for states (codeword=64)

상태 수 화 자	[%]			
	6	7	8	9
MA	89.3	91.1	92.9	92.9
WB	87.5	87.5	87.5	87.5
MC	96.4	96.4	96.4	96.4
AVG.	91.1	91.7	92.3	92.3

HMM의 상태의 수를 변화 시키면서 인식실험을 하여 표 4에 나타내었다.

표 3은 이산분포 HMM에 의한 인식결과이며, 상태 수를 8로 고정하고 코드워드의 수를 변화 시키면서 실험한 각 화자별 인식률을 보였다. 전체 화자에 대한 평균 인식률은 코드워드의 수가 32일 때 인식률은 85[%], 코드워드의 수가 64일 때 92[%], 코드워드의 수가 128일 때 61[%]를 얻었다. 여기서, 코드워드의 수가 32일 때 인식률이 떨어지는 것은 관측심분의 수가 부족하여 관측 심분의 시계열로 구성되는 유사패턴의 오인식이 발생한 것이며, 코드워드의 수가 128일 때는 28개의 지역명에 대하여 관측심분의 수가 너무 많아서 인식률이 저하된 것으로 생각된다. 본 연구에 사용한 데이터 베이스에서는 코드워드의 수가 64일 때 적합함을 알 수 있었다.

또한, 표 4는 코드워드의 수를 64로 고정하고 상태의 수를 변화시키면서 실험한 결과를 보이고 있다. 여기서, 전체 화자에 대한 평균 인식률은 상태의 수가 6일 때 92[%], 상태의 수가 7일 때 91[%], 상태의 수가 8과 상태의 수가 9일 때 92[%]를 얻었다. 이로부터, 코드워드의 수가 64이고, 상태의 수가 8일 때 가장 좋은 인식률을 얻을 수 있었다.

4-2-2 준연속분포 HMM에 의한 인식 실험

이산분포 HMM에서 가장 적당한 코드워드의 수는 64이고 상태의 수는 8일 때 이었으므로, 동일한 조건으로 준연속분포 HMM의 인식 실험을 하였으며, 그림 3에서 보인 준연속분포 HMM 단어인식 구성도에 의하여 단어인식을 수행하였다. 표 5에는 코드워드의 수를 64, 상태의 수를 8로 하였을 때, 이산분포 HMM과 준연속분포 HMM에 의한 인식률을 보이고 있다. 여기서, 준연속분포 HMM에 의한 인식률이

이산분포 HMM에 의한 인식률보다 향상된 것으로 나타났다. 이것은 준연속분포 HMM의 최우코드북 설계법에서 식(15)에 의한 보간 계수를 산출하므로, 이산분포 HMM보다 정확한 모델을 생성하기 때문으로 생각된다. 또한, 표 5에서 남성화자 대 여성화자의 인식률을 비교하면, 여성화자의 인식률이 남성화자에 비해 오인식이 많음을 볼 수 있다. 이것은 학습 과정에서 사용한 여성화자 1인 학습 데이터 수가 남성화자 2인의 학습 데이터의 1/2밖에 되지 않기 때문에 여성화자에 대한 HMM 모델이 학습과정에서 충분히 작성되지 않았기 때문이라 생각된다.

이상의 실험 결과로 부터, 그림 4에 코드워드의 수가 64이고 상태의 수가 8일 때, 이산분포 HMM과 준연속분포 HMM에 의한 전체 화자의 평균 인식률을 나타내었다.

표 5. 이산분포 HMM과 준연속분포 HMM의 인식률 비교
Table 5. Comparison of recognition rate for DHMM and SCHMM

실험방법 화 자	[%]	
	DHMM	SCHMM
MA	92.9	94.6
WB	87.5	91.1
MC	96.4	96.4
AVG.	92.3	94.1

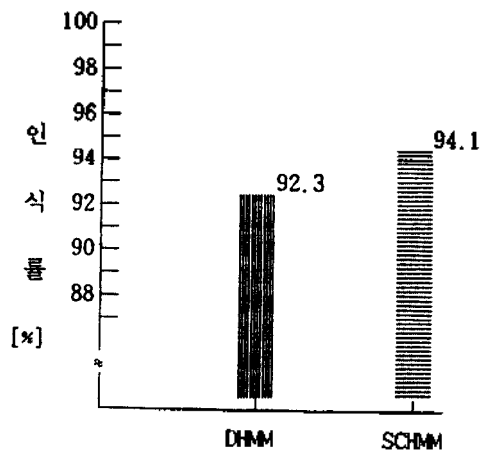


그림 4. 이산분포 HMM과 준연속분포 HMM의 인식률 비교

Fig 4. Comparison of recognition rates for DHMM and SCHMM

V. 결 론

본 연구에서는 HMM의 이산분포를 연속분포로 근사시키는 준연속분포 HMM에 의한 한국어 단어 인식 방법에 관하여 연구하였다. 이 방법을 평가하기 위하여 28개의 DDD 지역명을 대상으로 이산분포 HMM과 비교 실험한 결과, 다음과 같은 결론을 얻을 수 있었다.

첫째, 이산분포 HMM에 의한 인식 방법에서, 상태의 수가 8이고, 코드워드의 수가 64일 때 92%의 인식률을 얻었다.

둘째, 준연속분포 HMM에 의한 인식방법에서 이산분포 HMM과 동일한 상태의 수와 코드워드의 수로 하였을 때 94%의 인식률을 얻으므로써, 이산분포 HMM에 의한 인식률보다 향상된 인식률을 얻었다.

따라서, 준연속분포 HMM에 의한 음성인식 방법이 이산분포 HMM보다 우수한 방법임을 확인하였다.

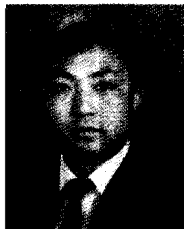
참 고 문 헌

1. J. L. Flanagan, "Voices of Men and Machine," J. Account. Soc. Am., Vol.51, Mar. 1971.
2. A. Averbuch, et al. "Experiments with the Tangora 20,000 Word Speech Recognizer," ICASSP-87, pp. 701~704, April, 1987.
3. Y. L. Chow, et al. "BYBLOS: The BBN Continuous Speech Recognition System," ICASSP 87, pp. 89~92, April, 1987.
4. Kai-Fu Lee, "Automatic Speech Recognition: The

Development of the SPHINX System," Kluwer Academic Pub., 1989.

5. L. R. Rabiner, "A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition," Proceedings of IEEE, Vol.77, No.2, Feb. 1989.
6. L. R. Rabiner, et al. "Recognition of isolated Digits Using Hidden Markov Models with Continuous Mixture Densities," BSTJ, Vol. 64, No.6, pp.1211~1234, Jul/Aug., 1985.
7. J. D. Markel, A. H. Gray, Linear Prediction of Speech, Springer Verlag, N. Y., 1976.
8. A. Gersho, "Speech Coding based upon Vector Quantizer," IEEE ASSP-28, Oct. 1980.
9. Y. Linde, A. Buzo, R. M. Gray, "An algorithm of vector quantization design," IEEE Trans. Commun., Vol.COM-28, pp.84~95, Jan. 1980.
10. J. G. Wilpon, L. R. Rabiner, "A Modified K-means Clustering Algorithm for Use in Isolated Word Recognition," IEEE Trans. ASSP, Vol. ASSP-33, pp.587~594, Jun. 1985.
11. X. D. Huang, et al., "Unified Techniques for Vector Quantization and Hidden Markov Modeling using Semi-continuous Models," ICASSP-89, pp. 639~642, April, 1989.
12. X. Huang, K. H. Lee, H. W. Hon, "On Semi-Continuous Hidden Markov Modeling," ICASSP-90, pp.689~692, April, 1990.
13. 今村, 浜田, 中津, "疑似連続分布 HMMによる不特定話者電話音聲認識," 信學技報, SP89-21, No.90, 1989.
14. 今村, "疑似連続分布 HMMにおけるコードブックの最適化," 信學技報, SP89-82, No.340, 1989.

▲趙炳瑞



1965年 12月 3日生
1989年 2月: 명지대 전자공학과
(학사)
1992年 2月: 명지대 전자공학과
(공학석사)