

ML/MMSE를 이용한 HMM-Net 분류기의 학습에 대한 실험적 고찰

(An Empirical Study on the Learning of HMM-Net Classifiers Using ML/MMSE Method)

金商雲*, 愼晟孝*

(Sang-Woon Kim and Seong-Hyo Shin)

요 약

HMM-Net은 은닉 마르코프 모델(HMM)의 계산과정을 신경망 구조로 구현한 것으로, HMM이 갖고 있는 시계열 모델링 능력과 신경망이 갖고 있는 우수한 변별력을 결합한 것이다. HMM-Net 분류기를 학습하는 학습방법으로는 HMM의 학습에 이용되는 最尤法(ML)과 신경망 학습의 最小二乗誤差法(MMSE)를 적용할 수 있다. 이들 중 ML이 MMSE보다 안정된 학습을 보장하는 반면 초기 학습조건을 적절하게 설정하였을 경우에는 MMSE가 ML보다 우수하다고 알려져 있다^[3]. 따라서 이 논문에서는 먼저 ML을 이용하여 초기학습을 수행한 다음 보다 학습성능이 우수한 MMSE로 바꾸어 최적 또는 준최적으로 학습하는 하이브리드 학습법(ML/MMSE)을 제안한다. 실험용 시계열 패턴으로 /0/부터 /9/까지의 고립 숫자음을 이용하여 실험한 결과, 제안한 방법이 학습특성 및 인식률면에서 ML이나 MMSE만을 이용하는 기존의 방법보다 우수하였음을 확인하였다.

Abstract

The HMM-Net is a neural network architecture that implements the computation of output probabilities of a hidden Markov model (HMM). The architecture is developed for the purpose of combining the discriminant power of neural networks with the time-domain modeling capability of HMMs. Criteria of maximum likelihood(ML) and minimization of mean squared error(MMSE) are used for learning HMM-Net classifiers. The criterion MMSE is better than ML when initial learning condition is well established. However ML is more useful one when the condition is incomplete^[3]. Therefore we propose an efficient learning method of HMM-Net classifiers using a hybrid criterion(ML/MMSE). In the method, we begin a learning with ML in order to get a stable start-point. After then, we continue the learning with MMSE to search an optimal or near-optimal solution. Experimental results for the isolated numeric digits from /0/ to /9/, a training and testing time-series pattern set, show that the performance of the proposed method is better than the others in the respects of learning and recognition rates.

I. 서 론

은닉 마르코프 모델(hidden Markov model; HMM)

* 正會員, 明知大學校 컴퓨터學部

(Dept. of Computer Science & Eng., Myongji Univ.)

接受日字: 1998年12月19日, 수정완료일: 1999年6月1日

과 인공 신경망 모델(artificial neural network; ANN)은 서로 보완적인 특성을 갖는 계산모델이다. 즉, HMM은 패턴의 시계열 특성을 잘 반영할 수 있으나 변별력이 약하고, ANN은 변별력이 우수하고 패턴분포의 통계적 변화에 유연하나 시계열 패턴의 순서 특성을 모델링하기 어렵다는 특징이 있다^[1].

HMM-Net(hidden Markov model network)은

HMM과 ANN의 장점만을 결합하기 위하여, HMM의 출력확률 계산과정을 신경망으로 구현한 구조이다. 따라서 HMM-Net의 학습을 위하여 HMM학습에 이용되는 最尤法(maximum likelihood: ML), 最大相互情報法(maximum mutual information: MMI)이나 ANN 학습에 널리 쓰이는 最小二乘誤差法(minimization of mean squared error : MMSE) 등을 이용할 수 있다^[2]. ML은 클래스별 분류기를 학습할 때 자신의 클래스에 속하는 학습패턴(정인식 패턴)만을 이용하는 반면, MMSE는 다른 클래스에 속하는 패턴이 정인식에 미치는 영향까지도 고려할 수 있도록 자신 클래스의 패턴뿐만 아니라 다른 클래스의 학습패턴(오인식 패턴)까지도 이용하는 학습방법이다.

문헌 [3]은 HMM-Net 분류기를 학습하기 위하여 ML과 MMSE를 적용하기 위한 학습식을 유도하였고, 이를 이용하여 실제의 음성 데이터에 대한 학습 특성을 고찰하였다. 실험 결과 ML을 이용할 경우 학습이 빠른 반면 인식률이 저조하고, MMSE의 경우 초기조건이 적절할 경우 학습속도는 느리나 인식률은 양호하였음을 보고하고 있다. 또한 MMSE는 ML보다 정확한 결정경계를 제공하지만, 학습속도가 느리고 지역 최소값(local minima)에 빠지는 등의 단점이 있음을 보여 준다.

MMSE는 각 클래스별로 정인식의 학습패턴 뿐만 아니라 오인식 패턴까지도 학습에 이용하기 때문에 성공적으로 학습하였을 경우에는 ML보다 정확한 결정 경계를 제공하지만 학습속도가 느리게 되고 학습특성 곡선이 진동하는 등의 문제를 갖게 된다. 반면에 ML은 정인식 학습패턴만을 이용하기 때문에 비교적 빨리 안정하게 준 최적 상태로 학습할 수 있으나 학습 능력은 MMSE에 미치지 못하게 된다. 이러한 사실로부터 ML과 MMSE를 결합하면 두 학습방법의 단점을 보완하고 성능 향상을 기할 수 있다는 직관적 관찰이 가능하다.

따라서 본 논문에서는 먼저 ML로 적절한 초기상태를 학습한 다음 학습방법을 MMSE로 바꾸어 최적 또는 준최적으로 학습하는 ML/MMSE 학습방법의 유용성을 검토하기 위한 실험결과를 보고한다. 이 방법에서는 학습초기에는 정인식 학습 패턴만을 이용하는 ML로 안정하고 빠르게 학습한 다음, 학습 후반에서는 보다 학습 능력을 향상시키기 위하여 정인식 뿐만 아니라 오인식 클래스의 패턴까지도 이용하는 MMSE를

적용하여 수렴속도를 가속시킬 수 있도록 한다.

제안 방법은 앞에서 살펴본 바와 같이 직관적인 이해에 기반한 학습방법으로서, 이론적으로 최적해가 보장된다는 근거가 없기 때문에 이 방법을 실제의 시계열 패턴(전화 음성 데이터)에 적용하여 그의 유효성을 실험적으로 고찰하고자 한다.

먼저 제 II장에서는 HMM-Net의 구조를 정의하고, 제 III장에서는 기울기 강하법(gradient descent)으로 HMM-Net 분류기를 학습하는 방법에 대하여 설명한다. 그리고 제 IV장에서는 제안한 방법과 기존의 방법을 이용하여 고립된 숫자 음성 데이터를 대상으로 실험한 결과를 고찰하고, 제 V장에서 결론을 맺는다.

II. HMM-Net

HMM 모델 λ_M 이 관측열 $O = v_{y_1}, v_{y_2}, \dots, v_{y_T}$ 를 출력할 확률 $\lambda_M(y_1^T)$ 은 다음과 같다.

$$\lambda_M(y_1^T) = P_M(y_1^T) = \prod_{t=1}^T \alpha_t(t) \beta_t(t) \quad (1)$$

여기서 v_{y_t} : 시간 t 에서의 관측기호, $y_t = 1, 2, \dots, M$; N : 상태의 수; M : 관측기호의 수; T : 관측열의 길이이다. 또한 $\alpha_t(t)$ 와 $\beta_t(t)$ 는 각각 전향 알고리즘(forward algorithm) 및 후향 알고리즘(backward algorithm)으로 계산하며, 정규화하면 다음과 같다.

$$\tilde{\alpha}_i(t) = \frac{\sum_{j=1}^N \tilde{\alpha}_j(t-1) a_{ji} b_{iy}}{\sum_{j=1}^N \sum_{k=1}^N \tilde{\alpha}_j(t-1) a_{jk} b_{iy}}, \quad \tilde{\alpha}_i(0) = \pi_i \quad (2a)$$

$$\tilde{\beta}_i(t) = \frac{\sum_{j=1}^N a_{ji} b_{iy+1} \tilde{\beta}_j(t+1)}{\sum_{j=1}^N \sum_{k=1}^N \tilde{\alpha}_j(t-1) a_{jk} b_{iy}}, \quad \tilde{\beta}_i(T) = 1 \quad (2b)$$

여기서 π_i : i 상태의 초기 확률값; a_{ij} : i 상태에서 j 상태로 천이하는 확률값; b_{jk} : j 상태에서 k 번째 심볼 v_k 를 출력할 확률값이다. 로그를 취한 출력확률 $\Lambda_M(y_1^T)$ 은

$$\begin{aligned} \Lambda_M(y_1^T) &= \log P_M(y_1^T) \\ &= \Lambda_M(y_1^{T-1}) + \log \sum_{i,j} \tilde{\alpha}_j(t-1) a_{ji} b_{iy} \end{aligned} \quad (3)$$

이 된다. HMM-Net은 식(3)의 계산 과정을 ANN으로 구현한 것으로서, HMM이 파라미터를 학습하기 위하여 Baum-Welch 알고리즘^[1]을 이용하는 반면에 신경망 구조인 HMM-Net은 기울기 강하법을 사용할 수 있다. 식(3)에서 시간 1에서 t 까지의 관측열에 대한 출력확률 $\Lambda_M(y_i)$ 를 계산하기 위한 재귀 신경망 구조는 그림 1.과 같다.

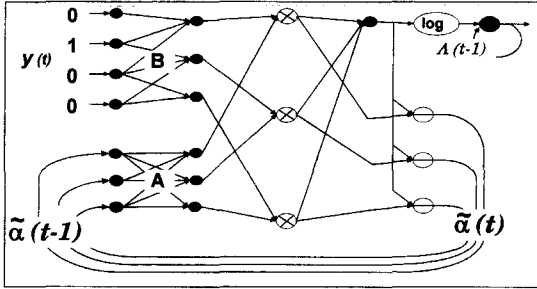


그림 1. HMM-Net 구조^[2]
Fig. 1. Structure of HMM-Net^[2].

여기서 ●는 각 링크의 값을 더하는 노드이고 ⊗는 두 입력을 서로 곱하는 노드이며, ⊖는 하나의 입력을 다른 입력으로 나누는 노드이다. 그리고 A와 B는 HMM의 상태전이확률과 출력기호확률에 대응하는 연결 가중치(connection weights) 파라미터이다.

III. 학습

기울기 강하법에 의한 HMM-Net의 학습이란 학습 기준함수 G 를 설정한 후 G 의 극대점에서의 가중치값을 찾기 위하여 $\partial G/\partial \theta = 0$ 를 만족하는 가중치값을 결정하는 방법이다. 즉, 네트워크의 가중치 파라미터 A 의 수정 전후의 값을 각각 $A(t)$, $A(t+1)$ 라고 하면 학습식은 다음과 같다.

$$A(t+1) = A(t) + \eta \nabla G \quad (4)$$

여기서 t 는 반복회수(epochs)이고, η 는 학습률이며, ∇ 는 G 의 기울기(gradient)를 구하는 연산자이다.

기울기 강하법에서 이용할 수 있는 기준함수로는 HMM에서 이용하는 G^{ML} 과 ANN에서 이용하는 G^{MMSE} , G^{MMI} 가 있으며 각각 다음과 같다^[2].

$$G^{ML} = \sum_n \log X_{C_n}(n) = \sum_n \Lambda_{C_n}(n) \quad (5a)$$

$$G^{MMSE} = -\frac{1}{2} \sum_n \sum_k (X_k(n) - \delta_{k,C_n})^2 \quad (5b)$$

$$G^{MMI} = \sum_n (\log X_{C_n}(n) - \log(\frac{1}{K} \sum_{k=1}^K X_k(n))) \quad (5c)$$

여기서 $X_k(n)$ 은 k 번째 모델에서 n 번째 학습 기호열에 대한 출력확률이고, C_n 은 n 번째 기호열의 클래스이며, K 는 클래스의 수, $\delta_{i,j}$ 는 크로벡터 델타함수이다.

G^{ML} 은 각 클래스의 학습패턴이 단지 자신의 클래스에 소속할 정도만을 고려하는 반면, G^{MMSE} 는 자신의 클래스뿐만 아니라 다른 클래스(오인식 클래스)에 미치는 영향까지 고려하는 기준함수이다. 따라서 성공적으로 학습되었을 경우 G^{MMSE} 는 G^{ML} 보다 정확한 결정정계를 제공하지만, 학습속도가 느리고 지역 최소값에 빠지는 등 효율적으로 학습시키 못하는 경우가 있다. 한편, G^{MMI} 는 자신의 클래스에 소속할 정도를 상대적인 값으로 나타낸 기준이다.

기울기 강하법으로 기준함수의 극점을 찾기 위한 파라미터 a_{ij} 에 대한 G 의 변화량은

$$\frac{\partial G}{\partial a_{ij}} = \sum_{n=1}^{N_t} \sum_{k=1}^K \phi_{n,k} \frac{\partial \Lambda_k(n)}{\partial a_{ij}} \quad (6)$$

이다. 여기서 $\phi_{n,k} = \partial G / \partial \Lambda_k(n)$ 는 식(5)로부터 구하고, $\partial \Lambda_k(n) / \partial a_{ij}$ 는 그림1.로부터 재귀적으로 계산한다^[3]. 즉, 식(6)의 결과를 식(4)에 대입하여 파라미터 a_{ij} 를 수정하는 과정을 모든 학습패턴에 대하여 반복 수행하여 a_{ij} 값을 학습한다. 같은 방법으로 파라미터 b_{jv} 에 대한 G 의 변화량 $\partial \Lambda_k(n) / \partial b_{jv}$ 를 계산하여 b_{jv} 값을 학습한다.

또한 학습기준을 G^{ML} 로부터 G^{MMSE} 로 바꾸기 위하여 다음과 같이 기준함수 G 의 기울기값의 변화 ΔG 를 이용한다.

$$\Delta G = |\nabla G(t+5) - \nabla G(t)| \quad (7)$$

여기서 ∇G 는 기준함수 G 의 기울기값이고, t 는 반복회수이다. 먼저 G^{ML} 로 학습을 시작한 후 기준함수 G^{ML} 의 기울기 변화 ΔG^{ML} 가 정해놓은 값 이하로 떨어질 경우 G^{ML} 을 G^{MMSE} 로 바꾸어 학습을 계속한다.

IV. 실험 및 고찰

1. 실험 데이터

HMM-Net분류기의 학습특성과 인식률을 고찰하기 위한 실험 데이터는 한국의 전자통신연구원(ETRI)^[5]에서 배포한 음성 데이터베이스의 데이터 중에서 실험 및 고찰이 용이한 전화 음성을 이용하였다. /0/부터 /9/까지의 음성 데이터 중 남자 9명이 네 번씩 발음한 360개의 데이터(Data Set 1; DS1)와 여자 9명이 네 번씩 발음한 360개의 데이터(Data Set 2; DS2) 그리고 남녀의 발음을 혼합한 360개의 실험 데이터 집합(Data Set 3; DS3)을 실험 데이터로 이용하였다.

실험 데이터의 전처리과정은 다음과 같다. 먼저 전화 음성을 3.4KHz로 필터링하여 8KHz로 샘플링하였다. 그리고 앞 뒤의 묵음 부분을 제거한 후 SOM (self-organization feature map)법^[6]을 이용하여 벡터 양자화(vector quantization: VQ)를 수행하였으며, 10차 LPC계수를 추출하여 패턴벡터를 구성하였다. 또 다양한 실험을 위하여 LPC계수를 추출할 때 벡터 클러스터의 수를 10, 12, 16으로 다르게 하여 구성하였다.

2. 실험 방법

실험은 전체 데이터를 학습 및 평가에 모두 이용하는 전표본 학습법(Resubstitution; R-법)과 실험 데이터를 두 부분으로 나누어 한 쪽으로 학습한 후 나머지로 평가하는 분할 학습법(Holdout; H-법)을 이용하였다. R-법에서는 360개의 실험 데이터를 학습 및 평가에 이용하였고, H-법에서는 720개 중에서 임의로 360개를 선택하여 학습하고 나머지로 평가하였다.

기울기 강하법의 학습률은 ML학습(기준함수가 G^{ML} 인 학습 방법)에서는 $\eta=0.01$ 로 하고, MMSE학습(기준함수가 G^{MMSE} 인 학습 방법)에서는 $\eta=0.02$ 로 하였다. 또한 반복 학습 횟수는 최고 $t=300$ 회까지로 하였으며, 기준함수 G 값이 떨어지지 않고 진동하여도 학습을 종료시켰다.

3. 결과 고찰

학습 실험은 ML 및 MMSE법으로 학습하는 방법과 먼저 ML로 초기학습한 후 MMSE로 바꾸어 학습하는 ML/MMSE법으로 실험하여 학습 특성을 비교 고찰하였다. 그리고 ANN에서 많이 이용되고 있는

MMI와도 비교하였다. 또한 학습패턴의 전처리과정이 학습에 미치는 특성을 고찰하기 위하여 VQ를 수행할 때 벡터의 클러스터 수를 달리하면서 생성한 학습패턴에 대한 인식률을 비교하였다.

먼저 ML 학습을 수행하다가 언제 학습방법을 MMSE로 바꿀 것인가를 결정하기 위하여 식(7)을 이용하여 기준함수 G^{ML} 의 기울기 변화 ΔG^{ML} 를 고찰하였다. 세개의 실험 데이터 집합 DS1, DS2, DS3에 대한 학습 기준함수의 기울기(ΔG^{ML}) 변화는 그림 2와 같다.

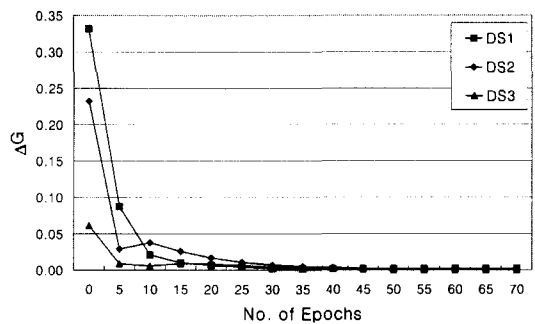


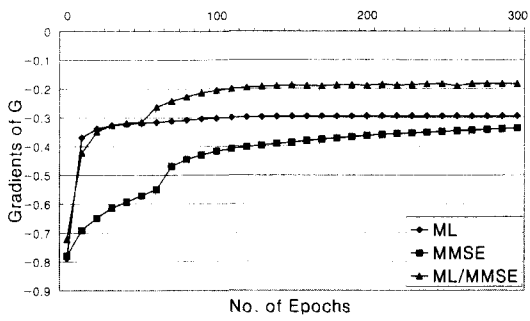
그림 2. 학습 기준함수의 기울기(ΔG^{ML})의 변화
Fig. 2. Variation of gradient of criteria function (ΔG^{ML}).

그림 2에서 ΔG^{ML} 는 학습초기에는 급격히 감소하지만 $t>50$ 에서는 거의 변화하지 않음을 보여준다. 따라서 본 논문에서는 $t=50$ 까지는 ML로 학습하고 그 후에는 MMSE로 바꾸어 학습하였다.

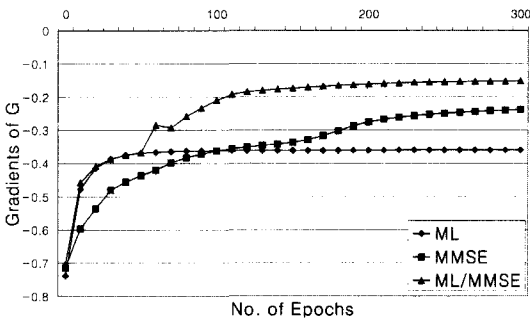
효과적인 학습법이란 빠르고 안정하게 기준함수의 극대점을 탐색해 갈 수 있는 학습법이다. 제안 방법의 학습특성을 고찰하기 위하여 학습 반복횟수(No. of Epochs)에 따른 기준함수의 변화(Gradients of G)를 실험하였다.

전처리 과정에서 클러스터의 수를 12로 벡터 양자화한 세종류의 학습 데이터 DS1, DS2, DS3에 대한 ML, MMSE, ML/MMSE의 학습 특성곡선은 그림 3과 같다.

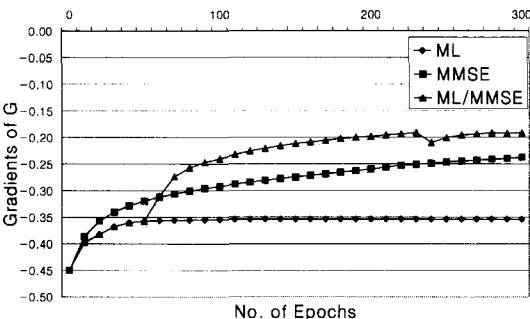
그림 3은 학습 데이터에 무관하게 ML/MMSE으로 학습하였을 때 가장 양호하게 학습이 됨을 보여준다. 즉 ML에서는 빠르게 학습이 되나 최적상태로 학습되지 못하였고, MMSE는 ML이나 MMI보다 느리게 학습되나 반복회수를 충분히 크게 하였을 경우 ML보다 우수하게 학습될 수 있음을 보여준다.



3(a)



3(b)



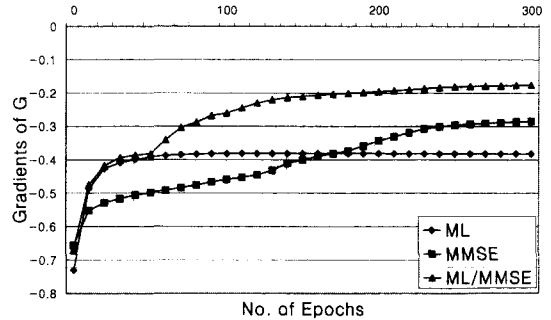
3(c)

그림 3. 학습 데이터 DS1, DS2, DS3에 대한 학습방법별 학습특성의 비교(VQ할 때의 클러스터 수: 12) (a) DS1에 대한 학습특성, (b) DS2에 대한 학습특성, (c) DS3에 대한 학습특성

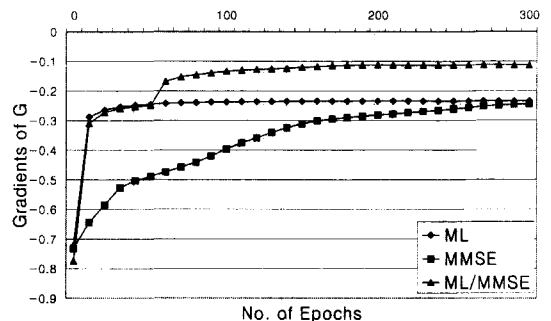
Fig. 3. Comparisons of learning curves for training data set DS1, DS2 and DS3(No. of clusters is 12). (a) Learning curves for DS1, (b) Learning curves for DS2, (c) Learning curves for DS3.

또한 학습방법은 학습데이터의 전처리 과정에 민감하다. 따라서 VQ에서 클러스터의 수를 다르게하여 준비한 학습 데이터에 대한 학습특성을 실험하였다. 전처리에서 클러스터 수를 10, 12, 16으로 바꾼 학습 데이터 DS1에 대한 ML, MMSE, ML/MMSE의 학습

특성곡선은 그림 4와 같다.



4(a)



4(b)

그림 4. Q의 클러스터 수를 10, 12, 16으로 하였을 경우의 학습방법별 학습특성의 비교(학습 데이터: DS1) (a) 클러스터 수가 10인 경우의 학습특성, (b) 클러스터 수가 16인 경우의 학습특성(클러스터 수가 12인 경우의 학습특성은 그림 3(b)와 같다)

Fig. 4. Comparisons of learning curves for the data quantized with clusters 10, 12 and 16(Training data set: DS1). (a) Learning curves for cluster 10, (b) Learning curves for cluster 16(Learning curves for cluster 12 are shown in Fig. 3(b)).

그림 4는 학습 데이터의 전처리 과정(클러스터의 수)과 무관하게 그림 3의 결과와 같이 ML/MMSE으로 학습하였을 때 가장 양호하게 학습이 됨을 보여준다.

또한 유용한 학습법이란 같은 학습 데이터로 학습하였을 경우 같은 결과를 학습할 수 있어야 한다. 따라서 본 논문에서는 같은 학습 데이터에 대하여 동일한 조건에서 세번씩 학습한 후 학습특성을 고찰하였다. 그림 5는 클러스터의 수를 16으로 한 학습 데이터 DS3에 대하여 동일한 조건에서 학습한 ML, MMSE, ML/MMSE의 학습 특성곡선이다.

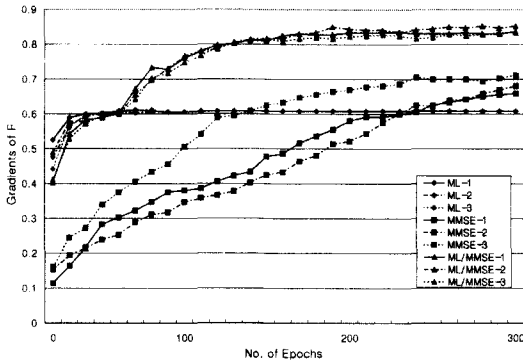


그림 5. 같은 학습을 세 번 수행하였을 경우의 학습방법별 학습 편차(학습 데이터: DS3; VQ의 클러스터 수: 16)

Fig. 5. Variations of learning curves for the learning methods carried out three times under same conditions(Training data set: DS3; No. of clusters in VQ: 16).

그림 5에서 MMSE는 편차(variation)가 큰 반면, ML에서는 작은 편차를 보여준다. 그리고 ML/MMSE에서는 MMSE의 높은 학습률을 유지하면서 편차가 ML수준으로 줄어 들었음을 보여준다.

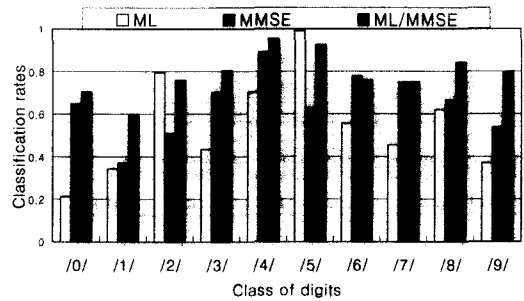


그림 6. 세 학습방법(ML, MMSE, ML/MMSE)의 클래스별 평균인식률의 비교(학습 데이터: DS1; VQ 클러스터의 수: 12)

Fig. 6. Comparisons of averaged classification rates of the ML, MMSE, ML/MMSE for each classes(Training data set: DS1; No. of clusters in VQ: 12).

높은 정인식률을 얻기 위해서는 패턴 분류기의 각 클래스별 고른 정인식률을 얻을 수 있어야 한다. 그림 6은 클러스터 수가 12인 학습 데이터 DS1을 이용하여 학습한 ML, MMSE, ML/MMSE의 클래스별 평균인식률을 보여준다. ML, MMSE에서는 클래스별 편차가 큰 반면, ML/MMSE에서는 편차가 작아 견고

표 1. 실험데이터 DS1, DS2, DS3에 대한 학습방법별 인식률(% ± s.d.)의 비교(10, 12, 16은 VQ할 때의 클러스터 수; MMI* : 학습기준 G^{MMI} 로 학습한 HMM-Net 분류기의 인식률; MMI/MMSE** : MMI와 MMSE를 결합하여 학습한 HMM-Net분류기의 인식률; HMM*** : 바움 웰치 알고리즘으로 학습한 HMM 분류기의 인식률)

Table 1. Comparisons of classification rates(% ± s.d.) for experimental data set DS1, DS2 and DS3(10, 12 and 16 are numbers of clusters in VQ; MMI* is classification rates of HMM-Net classifier trained with criterion G^{MMI} ; MMI/MMSE** is classification rates of HMM-Net classifier trained with a hybrid MMI/MMSE; HMM*** is classification rates of HMM classifier trained with Baum-Welch algorithm).

	10	12	16	10	12	16	10	12	16
ML	48.89 ± 0.00	60.28 ± 0.00	66.39 ± 0.00	49.44 ± 0.00	51.94 ± 0.00	61.39 ± 0.00	47.50 ± 0.00	50.56 ± 0.00	60.83 ± 0.00
MMSE	53.89 ± 7.63	61.39 ± 4.54	71.02 ± 8.39	59.91 ± 4.51	61.39 ± 7.67	75.19 ± 10.75	48.06 ± 3.37	58.33 ± 8.90	73.06 ± 3.64
ML/MMSE	77.13 ± 1.67	76.48 ± 1.05	88.33 ± 0.00	79.26 ± 0.57	74.56 ± 6.00	79.89 ± 4.55	71.39 ± 2.42	74.63 ± 3.58	85.74 ± 2.39
MMI*	74.72	68.89	76.25	76.39	73.75	83.05	64.86	67.78	74.54
MMI/MMSE**	71.94	77.22	88.06	79.44	83.33	87.78	64.44	72.78	81.11
HMM***	33.79	71.67	85.83	38.05	80.56	80.83	32.73	76.39	81.67

한 학습이 이루어졌음을 보여준다.

세 종류의 학습방법 ML, MMSE, ML/MMSE 및 MMI 학습 방법으로 학습한 HMM-Net 분류기와 Baum-Welch 알고리즘으로 학습한 HMM 분류기의 인식률을 실험하였다. 실험 데이터는 전처리 과정에서 클러스터 수를 10, 12, 16으로 바꾸어 준비한 DS1, DS2, DS3로 하였으며, 실험 방법은 학습 데이터를 그대로 측정 데이터로 이용하는 R-법으로 하였다. 각각의 실험 데이터에 대한 학습방법별 인식률(%)과 표준편차(\pm s.d.)는 표 1과 같다.

표 1은 데이터 종류나 전처리 과정에 무관하게 ML/MMSE의 인식률 및 표준편차가 ML이나 MMSE 및 MMI, HMM 중에서 가장 양호하고 견고함을 보여준다. 또한 MMI와 MMSE를 결합한 MMI/MMSE와 비교하여 비교적 양호한 특성을 보여준다.

실험방법 중 R-법은 전체 실험 데이터를 학습 및 평가에 똑 같이 이용하는 방법으로 낙관적인 결과를 도출하는 방법으로 알려져 있다. 따라서 비관적인 결과를 출력하는 H-법으로도 실험하였다. 실험 데이터 DS3의 720개 데이터 중에서 임의로 360개의 데이터를 선택하여 학습 데이터로 이용하였고 나머지 360개를 이용하여 인식률을 평가하였다. 각 실험 방법에 대한 학습방법별 인식률(%)과 표준편차(\pm s.d.)는 표 2와 같다.

표 2. 실험방법(R-법과 H-법)에 대한 인식률(% \pm s.d.)의 비교(실험 데이터: DS3; VQ의 클러스터 수: 16; MMI* : 학습 기준 G^{MMI} 로 학습한 HMM-Net 분류기의 인식률)

Table 2. Comparisons of classification rates(% \pm s.d.) for the experimental methods (R- and H-methods) (Experimental data set: DS3; No. of clusters in VQ is 16; MMI* is classification rates of HMM-Net classifier trained with criterion G^{MMI}).

Experimental Methods	R-Method	H-Method
ML	51.39 \pm 0.00	48.89 \pm 0.00
MMSE	62.78 \pm 1.47	51.30 \pm 2.51
ML/MMSE	78.43 \pm 0.32	60.46 \pm 0.70
MMI*	68.43 \pm 1.67	60.28 \pm 0.48

표 2는 모든 학습법에서 R-법의 인식률이 H-법보다 양호하여, R-법과 H-법에 대한 일반적 관측과 일치하는 결과를 보여준다. 또 R-법과 H-법 모두에서 ML/MMSE의 인식률이 가장 우수하고 표준편차는 ML보다 크나 MMSE보다는 적음을 보여준다.

끝으로 ML, MMSE, ML/MMSE의 학습 시간을 고찰하기 위하여 반복횟수별 평균 학습시간(minutes per epoch: m/e)을 고찰하였다. ML의 경우 13.49 m/e이었으며 MMSE에서 13.45 m/e이고 ML/MMSE는 15.28 m/e로, ML/MMSE 경우 ML이나 MMSE보다 처리시간이 다소 길었다.

V. 결론

이 논문에서는 HMM-Net 분류기를 학습할 때 먼저 ML로 학습한 다음 도중에 MMSE로 바꾸어 학습하는 ML/MMSE 학습법의 학습특성을 실험적으로 고찰하였다.

실험을 위한 시계열 패턴으로는 실험 및 실험결과의 고찰이 용이한 숫자음성 데이터를 이용하였다. 다양한 비교를 위하여 남성 여성 및 혼성의 세 종류 데이터를 준비하였으며, 전처리 과정이 학습에 미치는 효과를 관찰하기 위하여 벡터 양자화할 때 클러스터 수를 달리하면서 실험 데이터를 준비하였다.

ML/MMSE의 학습 특성을 고찰하기 위하여 서로 다른 학습 데이터를 이용하여 학습 실험한 결과, ML/MMSE가 ML이나 MMSE보다 양호한 학습특성을 보였다. 또한 이홉 종류의 데이터에 대하여 세 번씩 R-법으로 실험한 ML, MMSE, ML/MMSE의 평균 인식률 및 표준편차는 각각 55% \pm 0.0, 62% \pm 6.6, 79% \pm 2.5 으로, ML/MMSE가 ML이나 MMSE보다 우수하였다. H-법으로도 실험하였으며 R-법과 동일한 결론을 얻었다. 이러한 실험 결과로부터 ML과 MMSE를 결합한 ML/MMSE는 ML의 장점(안정되고 빠른 학습능력)과 MMSE의 장점(우수한 학습능력)을 모두 갖춘 학습방법임을 확인하였다.

한편 클러스터의 수에 따라 ML/MMSE의 인식률이 변하였다. 따라서 전처리 과정을 보강하는 방법으로 인식률을 더욱 향상시킬 수 있을 것으로 사료되며, 앞으로의 연구과제이다. 또한 ML/MMSE의 학습시간은 ML이나 MMSE보다 다소 길었다. 이는 정인식뿐만 아니라 오인식 패턴까지도 학습에 이용되기 때문

인 것으로, 학습시간을 줄이는 연구도 앞으로의 과제이다.

참 고 문 헌

- [1] H. Bourlard and C. J. Wellekens, "Links between Markov models and multilayer perceptrons", IEEE Trans. on PAMI, vol. 12, no. 12, pp. 1167 -1178, Dec. 1990.
- [2] L. T. Niles and H.F. Silverman, "Combining hidden Markov model and neural network classifiers", Proc. of ICASSP, pp. 417- 420, Apr. 1990.
- [3] 金商雲, 吳嶺煥, "HMM-Net 분류기의 학습", 大韓電子工學會論文誌, vol. 34-C, no. 9, pp. 703 - 709, 1997. 09
- [4] 김상운, 신성효, "하이브리드법에 의한 HMM-Net분류기의 학습", 대한전자공학회 추계종합학술대회 논문집, vol. 21, no. 2, pp. 1273-1276, 1998. 11
- [5] 韓國電子通信研究院, ETRI 音聲DB 配布說明書, 1995
- [6] [http : // www. cis. hut. fi / nnrc / nnrc-programs.html](http://www.cis.hut.fi/nnrc/nnrc-programs.html).

저 자 소 개

金 商 雲(正會員) 第 35卷 C編 第 2號 參照
현재 명지대학교 컴퓨터학부 교수

愼 晟 孝(正會員) 第 35卷 C編 第 2號 參照
MIND TELL, INC 선임연구원