

예측형과 분류형 신경망을 이용한 한국어 숫자음 인식

정회원 한 학용*, 김주성*, 고시영**, 허강인*, 안점영*

Recognition of Korean Isolated Digits Using Classification and Prediction Neural Networks

Hag-Yong Han*, Joo-sung Kim*, Si-Young Koh**, Kang-In Hur*, Jeom-Young Ahn*

Regular Members

요 약

본 논문은 기존 분류형 신경망의 인식성능을 향상시키기 위하여 프레임 정규화와 비선형 사후확률 추정법(N-APPEM)을 제안하고 한국어 숫자음에 대하여 예측형과 분류형 신경망으로 인식성능을 평가하였다. 실험결과 예측형 신경망에서 최고 98.0%의 인식률을 얻었다. 예측형 신경망은 네트워크가 입력패턴의 카테고리 수만큼 마련되는 복잡한 네트워크를 가지는 반면에 분류형 신경망은 단일 네트워크로 구성되며 프레임 정규화와 비선형 사후확률 추정법으로 85.5%까지 인식률을 향상시킬 수 있었으며 이는 기존의 방법보다 인식률이 12.0% 향상된 것이다.

ABSTRACT

This paper proposes a N-APPEM(Nonlinear A Posteriori Probability Estimation Method) with a frame normalization method to conventional classification network to increase speech recognition ability. It also tests the recognition ability of the classification and prediction neural networks for the Korean isolated digits. From the experimental results, the prediction network with MLP(Multi-Layer Perceptron) achieves the highest recognition ability of 98.0%. The prediction requires very complicated networks increased linearly with the number of incoming speech categories. However, the classification network with the N-APPEM and the normalization improves the recognition ability up to 85.5% with a single network, which is almost 12.0% improvement.

I. 서 론

MLP(Multi-Layer Perceptron)는 정적패턴의 인식에는 우수한 성능을 보이지만, 음성과 같이 시간에 따라 특성이 변화하고 동일화자라도 발생속도에 따라 시간축 길이가 다른 동적인 패턴의 인식에는 적합하지 못하였다. 따라서 이러한 음성과 같은 시계열 동적 패턴에 대한 분류를 음성인식에 적용하기

위한 몇 가지 신경망이 제안되어졌다.^[1]

정적인 패턴분류를 위해서는 정적구조 MLP를 사용하여 구성하며, 메모리가 없는 처리요소를 사용한다. 이러한 정적구조 MLP는 패턴을 분류하기 위한 임의의 식별곡선을 형성할 수 있을 뿐만 아니라, 오류역전파(Error Back-Propagation) 학습규칙으로 많은 분야에서 뛰어난 성능을 발휘하고 있다. 그러나 음성과 같은 시계열 동적 패턴분류를 위해서는 망 내부에 어떠한 형태의 메모리를 갖게 하여 과거의

* 동아대학교 전기·전자공학부,
논문번호: 99268-0703, 접수번호: 1999년 7월 3일

** 경일대학교 전자정보공학과

입력이 현재의 출력에 영향을 주도록 하여 입력신호의 문맥을 학습하는 동적 신경망을 구성하여야 한다. 이러한 동적 신경망의 구성은 전향에 의한 것과 회귀에 의한 것이 있으며, TDNN(Time-Delay Neural Networks)의 경우는 전자에 해당한다. TDNN은 MLP로 구성된 전향신경망에 시간지연 개념을 도입한 것으로서 유소인식 및 신호예측에 적용되어 왔다. 이와 유사한 것으로 최근에 보고된 ATNN(Adaptive Time-Delay Neural Networks)은 TDNN이 고정된 지연을 갖고 있음에 착안하여 가변지연을 구현한 것으로서, 신호예측에 적용하여 TDNN보다 향상된 결과를 얻고 있다^[1]

회귀를 이용하여 동적 신경망을 구성할 경우에는 출력층과 은닉층의 활성화 값을 회귀시키는 구조로서 메모리를 구성한다. 회귀신경망은 Hopfield망^[2], Jordan망^[3], Elman망^[4]이 대표적이며, 이를 일반화시킨 Williams - Zipser망^[5]을 비롯한 많은 종류가 있다.

II. 분류형 신경망

분류형 신경망으로는 음소와 같은 일정한 정적구간의 기준프레임에 대한 분류를 사후확률 추정법^[6]으로 이용하여 왔으며 본 논문에서는 이를 보완하여 단일 네트워크로 구성되는 분류형 신경망의 학습시 교사신호의 설정 방식을 인간의 인지모델을 적용한 비선형적 교사신호로 설정함으로써 인식물의 향상을 기하는 비선형적 사후확률 추정법을 제안한다. 또한, 분류형 신경망으로 음절과 같은 다양한 길이를 가진 시계열 동적패턴의 분류에 사용하기 위해서는 신경망의 정적인 구조를 동적인 구조로 바꾸어 시계열 동적패턴의 시변성을 신경망 내부에 포함시켜야 한다. 그러므로 본 논문에서는 회귀구조를 갖는 분류형 신경망을 구성하여 회귀구조가 분류에 미치는 영향을 고찰하였다.

2.1 사후확률 추정법(APPEM)^[10]

분류형 신경망은 입력신호와 교사신호간의 집합인 학습패턴으로 학습을 행하여 출력신호와 교사신호간의 자승오차가 최소가 되는 신경망의 파라미터를 결정한다. 학습된 신경망은 시그모이드 함수를 이용하여 비선형의 사상을 재현한다. 최적인 사상은 다음 식으로 주어진다.

$$y(\bar{x}) = \int \bar{i} \cdot P(\bar{i} | \bar{x}) d\bar{i} \quad (1)$$

여기서 \bar{x} 는 입력벡터, \bar{i} 는 교사벡터, $P(\bar{i} | \bar{x})$ 는 입력벡터가 \bar{x} 일 때의 교사벡터 \bar{i} 의 확률밀도함수이다. 패턴인식의 문제에서 교사벡터의 요소 t_1, t_2, \dots, t_k 를 각 카테고리에 대응한 값으로 한 경우의 최적인 변환은 식(2)와 같다.

$$y(x) = \sum t_k \cdot P(t_k = 1 | x) \quad (2)$$

여기서, $P(t_k/x)$ 는 입력이 x 일 때의 카테고리 t_k 에 속하는 확률로서 Bayes의 정리에 의한 사후 확률에 해당한다. 교사패턴의 각 요소를 입력벡터가 속하는 카테고리에 해당하는 경우에는 1, 다른 카테고리의 경우는 0으로 하면 입력벡터 x_i 에 대응하는 출력벡터의 k 번째의 요소 y_{ik} 는 식(3)이 된다.

$$y_{ik} = \sum_{t_k=1} P(t_k | x_i) = P(t_{ik} | x_i) \quad (3)$$

이것은 입력벡터 x_i 의 값으로서 교사벡터의 k 번째의 요소가 1이 되는 확률이다. 그러므로 어느 카테고리에서 발생하는지 알고 있는 입력샘플을 몇 개 준비하고 입력데이터를 발생하는 카테고리에 해당하는 출력 유니트에는 교사신호 1을, 그 외의 카테고리에는 0을 부여하여 학습패턴을 만든다.

이 학습패턴을 신경망에 부여하여 역전파학습법으로 학습한다. 학습된 신경망에 평가 샘플로서 입력 데이터를 부여하면 이 데이터가 각 카테고리에 어느 정도의 사후 확률에 속하는지를 출력하게 되며 이 확률값이 가장 큰 카테고리를 인식 결과로 하는 방법이다.

2.2 비선형적 사후확률 추정법(N-APPEM)

인간이 음성을 청취하는 경우에 처음에는 무엇이라고 하는지 알 수 없지만, 시간이 지남에 따라 발성된 음성이 무엇인지 알 수 있다. 이러한 시간의 경과에 따라 그 음성의 의미를 파악하는 인지모델을 신경망에 적용하여 시간이 진행됨에 따라 교사신호값도 서서히 증가하는 방식으로 사후확률적 의미를 지닌 비선형적 교사패턴을 설정한다. 이는 기존의 사후확률 추정법에서 입력벡터에 속하는 카테고리에 일괄적으로 사후확률 "1"을 설정하는 것보다 좀더 인간의 지각모델에 가까운 처리를 위한 것이다.

즉, 사후확률의 추정법에 의한 방법에서 교사패턴의 각 요소를 입력벡터가 속하는 카테고리에 해당하는 경우에 시각 t 에서의 확률값 $Target_{on}(t)$, 다른

카테고리의 경우에 $Target_{off}(t)$ 로 하면 입력벡터 x_t 에 대응하는 출력벡터의 k 번째의 요소 y_{tk} 는 식(4)와 같다

$$y_{tk} = \sum_{i=1}^n P(t_i|x_i) = P(t_{ik} = Target_{on}|x_i) \quad (4)$$

이것은 입력벡터 x_t 의 값으로서 교사벡터의 k 번째 요소가 $Target_{on}(i)$ 이 되는 확률이다. 이 때 설정하는 비선형 교사신호는 식(5)(6)에 의한다.

$$Target_{on}(t) = 0.5 + 0.5(1 - e^{-at/L}) \quad (5)$$

$$Target_{off}(t) = 0.5 - 0.5(1 - e^{-bt/L}) \quad (6)$$

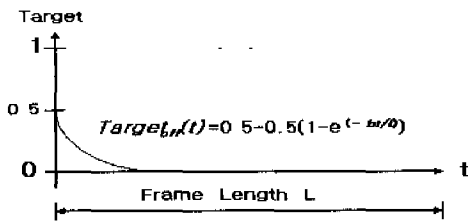
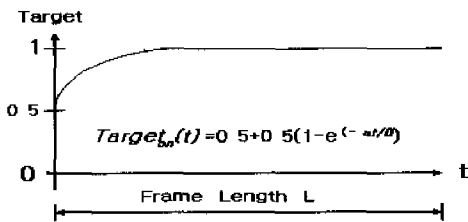


그림 1 N-APPEM의 교사신호 함수

그림1은 이러한 비선형계수의 변화에 따른 교사신호의 변화를 나타내며 이후의 학습방법과 인식방법은 사후확률 추정법과 동일하다.

여기서, L 은 음성의 전체길이에 해당하며, a, b 는 각각 $Target_{on}$ 과 $Target_{off}$ 의 비선형 계수이다

III. 예측형 신경망^[6]

예측형 신경망은 인식할 카테고리 수만큼의 네트워크를 구성하여 음성의 전구간에 걸쳐 순차적으로 전후프레임을 각각 입력신호와 교사신호로 사용하고, 학습된 각각의 네트워크가 뒤에 올 신호를 예측하도록 한다. 학습시에 다양한 길이를 가진 음성에 대하여 신경망의 입력 유니트에 전체 프레임을 순

차적으로 인가함으로써 특별한 프레임길이에 대한 정규화과정이 필요 없으며, 인식할 카테고리 수만큼의 네트워크가 마련되어야 하는 단점이 있는 반면에 대단위의 음절을 인식할 수 있다.

예측형 MLP는 과거의 특징벡터로부터 다음 시간의 벡터를 예측하므로 서로 인접한 특징벡터 사이의 상관 관계를 잘 반영할 수 있다 그림2는 예측형 MLP를 나타낸 것이며 입력층에 특징벡터 $y(t-\Delta), \dots, y(t-1)$ 을 순차적으로 입력시켜 예측된 특징벡터 $\hat{y}(t)$ 를 출력한다. 첨자 Δ 는 예측차수에 따른 입력벡터의 프레임 수를 나타낸다.

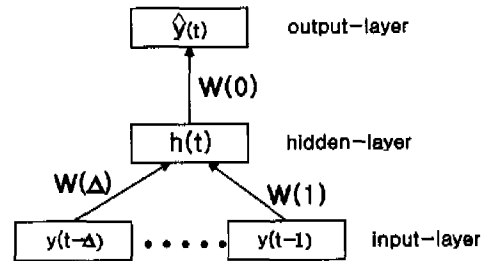


그림 2 예측형 MLP

신경망의 입출력 관계는 식(7), 식(8)과 같다.

$$h(t) = f \left(\sum_{s=1}^{\Delta} w(s)y(t-s) \right) \quad (7)$$

$$\hat{y}(t) = w(0)h(t) \quad (8)$$

여기서, $w(0), w(1), \dots, w(\Delta)$ 은 기중치 행렬을 나타내며, $h(t)$ 는 중간층유니트의 출력벡터, $f(\cdot)$ 는 시그모이드 함수로 입력벡터 y 의 각 요소에 대한 출력을 결정한다. 신경망의 학습규칙은 BP 알고리즘을 이용하며, 예측된 특징벡터인 $\hat{y}(t)$ 는 입력되는 특징벡터와 비교하여 유클리디안 거리의 자승 $\|y(t) - \hat{y}(t)\|^2$ 이 예측 오차값이 된다. 따라서 전체 길이가 T 프레임인 입력벡터에 대한 누적 예측 오차는 식(9)과 같다.

$$E = \sum_{t=1}^T \|y(t) - \hat{y}(t)\|^2 \quad (9)$$

IV. 회귀예측신경망

1. Jordan형 회귀예측신경망

그림3은 Jordan형 회귀예측신경망의 구조를 나타

내고 있다. 이 신경망은 출력층의 모든 유닛이 문맥층으로 회귀 연결되는 구조이다. 출력층에는 선형 출력함수, 은닉층에는 비선형함수인 시그모이드 함수를 사용한다.

$S^T = |s(1), \dots, s(T)|$ 를 음성의 발생시간 T 동안의 음성패턴이라 할 때 $s(t) = [s_1(t), \dots, s_M(t)]$ 는 시각 t 에서의 N 차원의 특징벡터이다.

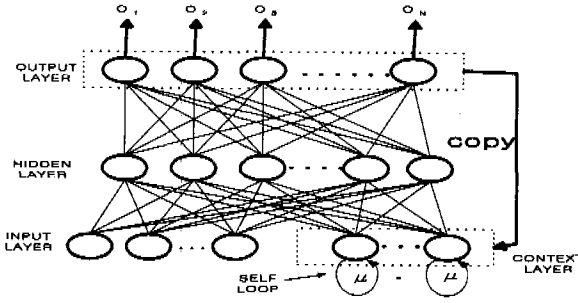


그림 3. Jordan형 회귀신경망

시각 t 에서의 교사벡터를 $s(t)$ 라 할 때 예측치수가 τ 인 경우, 입력벡터 $x(t) = [x_1(t), \dots, x_N(t)]$ 는 $s(t-1), \dots, s(t-\tau)$ 의 τ 개의 벡터를 연결한 벡터이다. 은닉층의 전체 입력값들로 이루어진 벡터를 $n(t) = [n_1(t), \dots, n_H(t)]$, 은닉층의 출력벡터를 $h(t) = [h_1(t), \dots, h_H(t)]$, 출력층의 출력벡터를 $o(t) = [o_1(t), \dots, o_N(t)]$ 라 하면, 교사벡터 $s(t)$ 에 대한 예측벡터 $\hat{s}(t)$ 는 출력벡터 $o(t)$ 와 같다. $n(t), h(t), o(t)$ 의 j, k 번째 성분 $n_j(t), h_j(t), o_k(t)$ 은 각각 다음과 같다.

$$\begin{cases} n_j(t) = \sum_{i=1}^{H+N} w_{ji} z_i(t) & i=1, \dots, H+N \\ h_j(t) = f(n_j(t)) & j=1, \dots, H \\ o_k(t) = \sum_{i=1}^N w_{ki} h_i(t) & k=1, \dots, N \end{cases} \quad (10)$$

여기서 w_{ji} 는 i 번째 입력층 유닛에서 j 번째 은닉층 유닛으로의 연결 가중치이고, w_{ki} 는 j 번째 은닉층 유닛과 k 번째 출력층 유닛으로의 연결 가중치이다. $f(\cdot)$ 는 시그모이드 함수를 나타낸다. 그리고 $z_i(t)$ 는 시각 t 에서 입력벡터 $x(t)$ 와 시각 $t-1$ 에서의 출력층의 출력벡터 $o(t-1)$ 를 연결한 $N+N_c$ 차원 벡터의 i 번째 성분이다.

$$\begin{aligned} z_i(t) &= [z_1(t), \dots, z_{N+N_c}(t)] \\ &= [o_1(t-1), \dots, o_N(t-1), x_1(t), \dots, x_N(t)] \quad (11) \\ &= [C_1(t), \dots, C_N(t), x_1(t), \dots, x_N(t)] \end{aligned}$$

여기서, $C_k(t)$ 는 문맥층의 벡터를 나타낸다.

시각 t 에서 교사벡터와 출력벡터 사이의 누적예측 오차는 다음과 같다.

$$J(t) = \frac{1}{2} \sum_{k=1}^N [e_k(t)]^2 \quad (12)$$

여기서, $e_k(t) = s_k(t) - \hat{s}_k(t) = s_k(t) - o_k(t)$ 는 k 번째 유닛의 예측오차, 즉 교사값과 실제 출력값 사이의 오차를 나타낸다.

한편, 신경망 학습이 시작시각 t_0 에서 시작하여 최종시각 T 에서 종료된다고 가정하면, 최종누적 예측오차 E 는 다음과 같다.

$$E = \sum_{t=t_0+\tau}^T J(t) \quad (13)$$

예측오차 제적을 최급강하법에 적용하면 가중치 벡터, W 는 ∇E 의 음(negative)의 값으로 수정되어진다. 이 기울기는 각 단계마다 예측오차 제적을 따라 $\nabla J(t)$ 의 값을 누적함으로 계산되어진다. 신경망에서 어떤 특정 가중치 w_{ji} 와 w_{ki} 에 대한 전체 가중치 변화는 다음과 같다.

$$\Delta w_{ji}(t) = -\alpha \frac{\partial J(t)}{\partial w_{ji}} = \alpha \sum_{k \in H} e_k(t) \frac{\partial h_j(t)}{\partial w_{ji}} \quad (14)$$

$$\Delta w_{ki}(t) = -\alpha \frac{\partial J(t)}{\partial w_{ki}} = \alpha e_k(t) o_i(t) h_i(t) \quad (15)$$

여기서 α 는 고정된 양(positive)의 학습률을 나타낸다.

신경망의 초기상태는 가중치에 대해서 함수적으로 독립이라고 가정하기 때문에 초기조건식은 다음과 같다.

$$\frac{\partial h_j(t_0)}{\partial w_{ji}} = 0 \quad (16)$$

그리고 회귀조건식은 다음과 같다.

$$\frac{\partial h_j(t+1)}{\partial w_{ji}} \quad (17)$$

그러므로, 각 시각 t 에 대해서 가중치 변화식은

아래와 같다.

$$w_n(t+1) = w_n(t) + \Delta w_n(t) \quad (18)$$

$$w_b(t+1) = w_b(t) + \Delta w_b(t) \quad (19)$$

문맥층에서 회귀계수를 적용하는 수식은 식20과 같다

$$\begin{aligned} C_i(t+1) &= \mu C_i(t) + O_i(t) \\ &= \mu(\mu C_i(t-1) - O_i(t-1)) + O_i(t) \quad (20) \\ &= \sum_{\tau=0}^t \mu^{t-\tau} O_i(\tau) \quad (C_i(0) = 0) \end{aligned}$$

V. 실험결과 및 고찰

1. 음성 데이터 및 분석조건

실험에 사용된 음성 데이터는 ETRI의 샘플이 데이터중에서 “영, 일, 이, 삼, 사, 오, 육, 칠, 팔, 구” 10개의 단독 숫자음을 사용하였다. 이는 남성화자 20명이 4회 발성한 총 800개의 데이터중에서 처음 3회분은 학습용(600개)으로 나머지 1회분은 평가용(200개)으로 사용하였으며 분석조건은 표1과 같다.

표 1. 음성 데이터의 분석조건

A/D	16kHz, 16bit, signed
Filtering	LPF , 7 kHz
Step Size	3.75 ms (프레임 정규화한 경우에는 가변)
Window Length	16 ms
Feature Parameter	10차 LPC Melcepstrum

2. 실험방법

신경망의 구조는 입력층 유니트수가 20, 은닉층 유니트수가 15, 출력층 유니트수가 10개로 하였다. 회귀구조일 경우에 Jordan형으로 하고 문맥층의 자기 회귀계수는 0.5로 하였다.

학습방법은 분류형 신경망의 경우, 사후확률추정에 의한 방법과 비선형 사후확률 추정법에 의한 방법으로 실험을 행하고, 5.3절에서 제안한 프레임 정규화한 경우의 실험 결과를 추가하였다. 예측형 신경망의 경우, 각 입력에 대한 교사패턴을 그 다음의 프레임으로 설정하고 예측제출을 학습한다. 매 순간

의 회귀입력은 음성의 시변성에 따라 결정되고, 예측오차는 매 순간마다 누적된다. 인식결과는 평균에 예측오차가 가장 최소가 되는 카테고리로서 한다. 실험 방법은 다음과 같다.

- ① MLP와 회귀신경망 모델의 분류를 통한 인식 (아래의 각 방법에 대하여)
 - 사후확률의 추정에 의한 분류
 - 비선형적 사후확률 추정법에 의한 분류
 - 프레임 정규화, 비정규화한 경우의 분류
- ② MLP 예측신경망과 회귀예측신경망에 의한 인식

3. 프레임 정규화

그림4는 평균 프레임 수를 보여주고 있다. 실험결과 숫자음 “육”의 경우 다른 숫자음의 평균 프레임 수보다 짧아 낮은 인식률을 보였다. 분류형 신경망의 경우 학습용 데이터의 수가 학습시 결정적인 영향을 미치기 때문에 이를 보상하기 위하여 본 논문에서 제안하는 정규화과정을 거친 후 실험을 하였다.

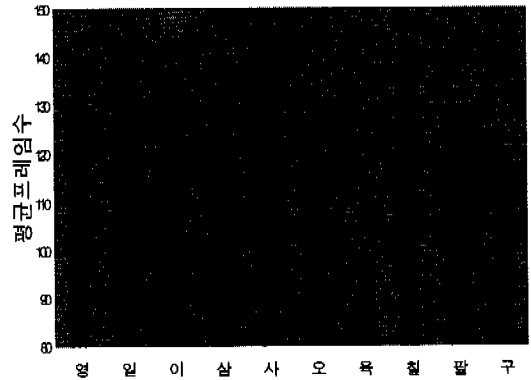


그림 4 숫자음에 대한 프레임수 비교

프레임 정규화는 실제음성에서 특징 파라미터를 추출하는 단구간 음성분석과정에서 프레임 간격을 두어 중첩되게 띄우는 창 의 간격을 실제음성의 총 길이에 따라 가변 시키면서 분석함으로써 프레임길이를 일정하게 하는 정규화 방법이다. 이 때 가변 시키는 창 의 간격은 실제음성의 총 길이를 정규화하고자 하는 프레임수로 나눈 값에 해당하며 이 간격이 창 의 크기를 넘지 않게 하여야 한다.

4 고찰

분류형 신경망과 예측형 신경망의 실험결과를 표

2. 표3에 보인다. 실험결과 예측형 신경망인 경우, 평가용 데이터에서 예측치수가 2이고 은닉층수가

15개인 경우 최고 98%의 인식률을 보였으며, 분류형인 경우, 오인식률의 분포가 MLP인 경우에는 단모음인 경우보다 이중모음이 포함되어 있는 “영”과 “육”인 경우에서 대부분의 오인식이 나타났으며, 회귀신경망의 경우에는 유사한 음인 “삼”과 “사”에서 나타났다. 이를 통하여 음성과 같은 시제열 패턴의 경우, 시변성을 신경망의 내부에 흡수시킨 회귀신경망이 분류능력은 우수하지만 이러한 유사음에 대한 인식률의 감소로 전체평균 인식률은 회귀구조에서 낮게 나타났다.

표 2. 학습용데이터의 실험결과[%]

NN Type		MLP	RNN
Normalization			
Frame non-normalization	APPEM	80.50	71.67
	N-APPEM	86.33 (3)	84.50 (6)
Frame normalization	APPEM	90.17	88.00
	N-APPEM	90.83 (7)	91.67(10)

표 3. 평가용 데이터의 실험결과[%]

NN Type		MLP	RNN
Normalization			
Frame non-normalization	APPEM	73.50	64.50
	N-APPEM	79.50 (3)	75.00 (6)
Frame normalization	APPEM	84.50	81.50
	N-APPEM	85.50 (7)	83.50(10)

*APPEM : A Posteriori Probability Estimation Method

*N-APPEM: Nonlinear APPEM

* ()안은 최적 비선형 계수(a)값을 나타낸다.

표 4. 예측형 신경망 실험결과

예측 차수	은닉층 수	인식률			
		회귀신경망 (Jordan)		예측형 MLP	
		학습	평가	학습	평가
1	5	88.3	85.0	89.5	86.0
	10	96.7	94.0	97.0	94.5
	15	96.8	92.5	98.0	96.0
	20	97.7	96.5	97.5	96.0
2	5	90.8	87.0	89.2	88.5
	10	97.2	97.0	97.2	94.0
	15	97.7	95.5	98.8	98.0
	20	99.0	98.0	99.3	97.0
3	5	90.2	87.5	89.5	88.0
	10	96.2	92.0	98.5	97.0
	15	99.2	97.0	98.7	97.5
	20	99.5	96.5	99.5	97.5
4	5	90.2	88.5	90.3	89.0
	10	98.5	96.5	97.8	96.5
	15	99.5	96.0	99.3	96.0
	20	99.5	96.5	97.8	95.0

VI. 결론

본 논문에서는 MLP와 회귀신경망을 이용하여 사후확률의 추정에 의한 방법(APPEM)과 비선형 사후확률 추정법(A-APPEM)으로 한국어 숫자음절에 대한 분류를 통한 인식실험을 행하고 예측형 신경망과 결과를 서로 비교하였다. 인식할 음성의 카테고리 수 만큼의 네트워크를 마련해야 하는 예측형 신경망의 경우, 평가용 데이터에서 최고 98%의 높은 인식률을 보였다. 반면 단일 네트워크로 구성되는 분류형 신경망의 경우에는 비선형 사후확률 추정법(79.5%)이 사후확률 추정법(73.5%)에 의한 방법보다 분류능력이 우수하였다. 또한, 신경망의 학습시에 결정적인 영향을 미치는 입력 프레임 수를 균일하게 하기 위해서 프레임 정규화 기법을 이용하여 정규화 한 후의 실험결과와 비교한 결과, 정규화한 경우(85.5%)가 정규화하지 않은 경우(79.5%)보다 우수하며 정규화 하지 않은 경우에는 비선형적 사후확률 추정법이 사후확률 추정법에 의한 방법보다 더욱 더 큰 효과를 보임을 알 수 있었다. 특히, MLP에서 오인식을 보이지 않는 유사한 음을 지닌 카테고리 회귀 신경망에서는 오인식의 요인이 되는 것으로 보아 음성의 시변성이 회귀구조의 동적인 신경망에 흡수되어 있음을 확인하였다.

참고 문헌

- [1] D. T. Lin, J.E. Dayhoff, and P. A. Ligomenides, "Adaptive time-delay neural network for temporal correlation and prediction," SPIE Intelligent Robots and Computer Vision XI: Bioloical, Neural Net, and 3-D Methods, vol. 1826,(Boston, November), pp.170-181, 1992

[2] J. J. Hopfield, "Neurons with Graded Response Have Collective Computational Properties Like Those of Two-State Neurons," Proceedings of the National Academy of Sciences USA, Vol. 81, pp.3088-3092, 1984

[3] M. I. Jordan, "Serial Order: A parallel distributed processing approach," Technical Report ICS-8604, Institute for Cognitive Science, University of California, Sab Diego, La Jolla, California, May 1986.

[4] J. L. Elman, "Finding structure in time," Technical Report CRL-8801, Center for Research in Language, University of California, San Diego, La Jolla, california, April 1988.

[5] R. J. Williams and D. Zipser, "A learning algorithm for continually running fully recurrent neural networks," Neural Computation, vol. 1, pp. 270-280, 1989.

[6] Ken-ichi Iso and Takao Watanabe, "Speaker-Independent Word Recognition Using A Neural Prediction Model," Proc.ICASSP'90, pp.441-444, 1990.

[7] K. Iso and T. Watanabe, "Large vocabulary speech recognition using neural prediction model," Proc. ICASSP'91, pp. 57-60, 1991.

[8] Joo-Sung Kim, Kwang-Suk Lee, Kang-In Hur, "The Recognition of Korean Syllables using Recurrent Prediction Neural Networks," ICASP, vol.1 of 2, pp.395-399, 1997.

[9] Soo Hoon Kim, Sang-Boum, Kang-In Hur, "The Recognition of Korean Syllables Using Neural Predictive HMM," ICASP, vol1 of 2. pp.427- 431, 1997.

[10] Seichi Nakagawa, Yoshiyuki Ono, Kang-In Hur, "Estimation of Probability Density Function and Evaluation by Vowel Recognition," Proc.IJCNN '93, pp.2223-2226, 1993.

[11] 한학용, 김주성, 허강인, "회귀신경망을 이용한 음성인식에 관한 연구", *한국음향학회지* 18권 3호 pp. 62-67. 1999.

한 학 용(Hag-yong Han)

정회원



1994년 2월: 동아대학교
전자공학과(공학사)
1994년~1997년: 경남에너지(주)
근무
1998년 2월: 동아대학교
전자공학과 (공학석사)
1999년~현재: 동아대학교
전자공학과 박사과정

동아대학교부설 정보통신연구소 특별연구원
<주관심 분야> 음성신호처리, DSP응용

김 주 성(Joo-sung Kim)

정회원



1990년 2월: 동의대학교
전자공학과(공학사)
1993년 2월: 동아대학교
전자공학과(공학석사)
2000년 2월: 동아대학교 전자공
학과 공학박사(예정)

<주관심 분야> 신경회로망, 음성신호처리

고 시 영(Si-young Koh)

정회원



1979년 2월: 영남대학교
전자공학과 (공학사)
1983년 2월: 영남대학교
전자공학과(공학석사)
1992년 8월: 동아대학교
전자공학과(공학박사)
1986년~현재: 경일대학교전자정
보공학과 교수

<주관심 분야> 음성신호처리, 생체신호처리

허 강 인(Kang-in Hur)

정회원



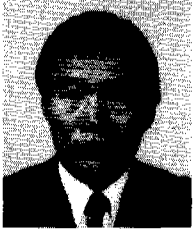
1980년 2월: 동아대학교
전자공학과 (공학사)
1982년 2월: 동아대학교
전자공학과(공학석사)
1990년 8월: 경희대학교
전자공학과(공학박사)
1984년 9월~현재: 동아대학교

전기·전자·컴퓨터공학부 교수

1988년 9월~1989년 8월: 일본 筑波大學 객원연구원
1992년 9월~1993년 8월: 일본 豊橋大學 객원연구원
<주관심 분야> DSP, 음성인식·합성, 신경회로망

안 점 영(Jeom-young Ahn)

정회원



1964년 2월 : 한국항공대학교
응용전자공학과(공학사)

1979년 2월 : 동아대학교
전자공학과 (공학석사)

1986년 8월 : 동아대학교
전자공학과 (공학박사)

1987년 3월~현재 : 동의대학교
전기·전자공학부 교수

<주관심 분야> 음성인식, 신경망, 디지털신호처리