

음절인식을 위한 회귀예측신경망에 관한 연구

한 학용, 김 주성, 허 강인
동아대학교 전자공학과

A study on the Recurrent Prediction Neural Networks for Syllables Recognition

Hag-Yong Han, Joo-Sung Kim, Kang-In Hur

Dept. of Electronic Eng, Dong-A Univ.

hyhan@electro.donga.ac.kr

요 약

본 논문에서는 MLP형 예측신경망, Jordan형과 Elman형 회귀예측신경망을 사용하여 예측차수와 은닉층의 유니트수의 변화에 따른 인식결과를 CHMM과 비교하였다. 음성데이터는 100음절데이터와 ETRI의 샘플이 숫자음을 사용하였다. 숫자음에서 신경망의 인식률은 98.5%로 5상대 CHMM의 99.0%와 유사한 인식결과를 나타내었다. 100음절에서는 은닉층의 유니트가 25개 이상인 Jordan형 회귀예측신경망이 5상대 CHMM의 85.6%보다는 향상된 인식성능을 보였으며, 6상대 이상의 CHMM보다는 다소 인식률이 낮게 나타났다.

1. 서 론

최근 다양한 구조의 신경망으로 패턴인식, 시스템 모델링, 비선형예측등과 같은 분야에 널리 사용되고 있다. 특히 MLP(Multi-layer Perceptron)는 정적패턴의 인식에는 우수한 성능을 보이지만, 음성과 같이 시간에 따라 특성이 변화하고 동일화지라도 발생속도에 따라 시간축 길이가 다른 동적패턴의 인식에는 적합하지 못하였다. 따라서 이러한 시계열 동적패턴을 음성인식에 적용하기 위한 몇가지 신경망이 제안되었다.

정적인 패턴분류를 위해서는 주로 정적구조의 MLP를 사용하여 구성하며, 메모리가 없는 처리요소를 사용한다. MLP는 패턴을 분류하기 위한 임의의 식별곡선을 형성할 수 있을 뿐만 아니라, 오차역전파(error back-propagation)라는 강력한 학습규칙이 존재하기 때문에 많은 분야에서 뛰어난 성능을 발휘하고 있다. 반면에 음성과 같은 시계열 동적패턴인 경우에는 동적신경망으로 망 내부에 어떠한 형태로든 메모리를 갖고 있어서 과거의 입력이 현재의 출력에 영향을 주게 하여 입력신

호의 문맥을 학습할 수 있도록 하여야 한다. 이것은 다시 전향에 의한 것과 회귀에 의한 것으로 나뉘어지며, 널리 알려진 TDNN(Time-Delay Neural Networks)의 경우는 전자에 해당한다. TDNN은 MLP로 구성된 전향신경망에 시간지연 개념을 도입한 것으로서 음소인식 및 신호예측에 적용한 예가 보고되고 있다. 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]

이와는 다르게 회귀신경망으로 동적신경망을 구성할 경우에는 출력층과 은닉층의 활성화 값을 회귀시키는 구조를 취하여 메모리를 구성하고 있다. 회귀신경망은 Hopfield망^[2]을 효시로 하여 Jordan망^[3]과 Elman망^[4]이 대표적이며, 이를 일반화시킨 Williams-Zipser망^[5]을 비롯한 많은 종류가 있다.

또한, 분류형신경망은 단일 네트워크로 구성되는 반면에 각 음절의 시간축상의 길이가 서로 다르므로 정규화 내지는 반드시 회귀연결과 같은 동적인 네트워크를 구성하여야 한다. 예측형신경망은 시간축상에서 정규화 과정이 필요 없으며 인식하는 음절수에 해당하는 네트워크가 필요한 반면에 대단위의 음절을 인식할 수 있으며 회귀연결을 할 경우에 보다 나은 예측성능을 얻을 수 있다.

본 논문에서는 서로 다른 길이를 가진 음성과 같은 동적패턴의 인식을 위해서 예측신경망(Predictive Neural Network)과 회귀예측신경망(Recurrent Prediction Neural Network)을 구성하여 패턴예측기로 사용하고 숫자음과 100음절에 대한 인식결과를 CHMM에 의한 인식결과와 비교하여 각 신경망 모델에 대한 성능을 서로 비교하였다.^{[6][7]}

2. 예측신경망^{[6][7]}

그림1은 MLP에 의한 음성패턴의 예측기를 보여주고 있다. 입력유니트에 음성특징벡터 $x_t, x_{t-1}, \dots, x_{t-r}$ 을 순차적으로 입력시킴으로써 예측된 특징벡터 \hat{x}_t 를 출력하도록 하는 것이다. 첨자는 음성특징벡터의 프레임수를 나타내며, r 는 입력특징벡터의 수를 나타낸다. 다층신경망의 입출력간의 관계는 식(1)(2)와 같다.

$$h_t = f\left(\sum_{p=1}^r w_p x_{t-p}\right) \quad (1)$$

$$\hat{x}_t = w_0 h_t \quad (2)$$

여기서, w_0, w_1, \dots, w_r 은 가중치 행렬을 나타내며, h_t 는 은닉층의 출력벡터, $f(x)$ 는 사그모이드함수로 벡터 x 의 각 요소에 대한 출력을 결정한다. 신경망의 학습규칙은 BP 알고리즘을 이용하였으며, 예측된 특징벡터인 \hat{x}_t 는 입력되는 음성특징벡터와 비교하여 이의 유클리디안 거리의 차승인 $\|x_t - \hat{x}_t\|^2$ 이 예측오차값이 된다. 인식결과를 이러한 예측오차값의 평관이 최소화 되는 모델로 한다.

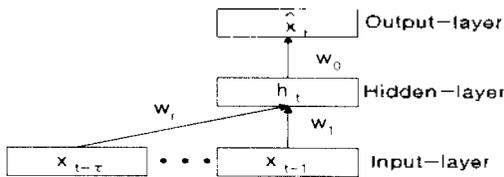


그림 1 MLP 예측신경망

3. 회귀예측신경망^[8]

회귀신경망은 회귀 구조에 따라 여러 가지 형태가 있으며, 학습 알고리즘도 시간에 따라 오차역전파 학습, RTRL(Real-Time Recurrent Learning)^[6]등이 있다. 또한 회귀형태에 따라 출력층의 활성치를 회귀시키기 위해 문맥층을 마련한 Jordan망^[3]과 은닉층의 활성치를 회귀시키기 위해 문맥층을 마련한 Elman망^[4]등이 있다.

3.1 Elman형 회귀예측신경망

그림 2는 Elman형 회귀예측신경망의 구조를 나타내고 있다. 이 신경망은 입력층과 출력층 사이에 은닉층을

삽입하고, 은닉층의 모든 유니트는 문맥층으로 회귀연결되는 다층구조이다. 출력층에는 선형출력함수, 은닉층에는 비선형함수인 시그모이드 함수를 사용한다.

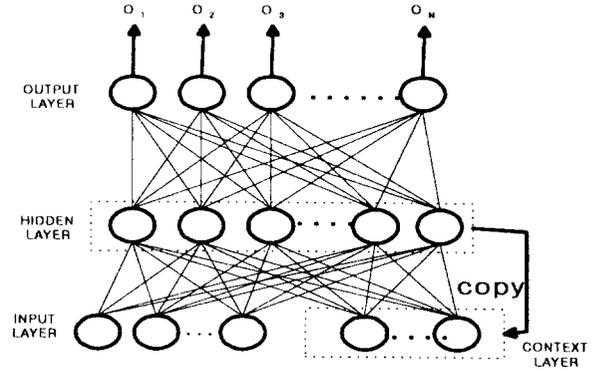


그림 2 Elman형 회귀예측신경망

$S^T = |s(1), \dots, s(T)|$ 를 음성의 발생시간 T 동안의 음성패턴이라 할 때 $s(t) = [s_1(t), \dots, s_M(t)]$ 는 시각 t 에서의 N 차원의 특징벡터이다. 시각 t 에서의 교사벡터를 $s(t)$ 라 할 때 예측차수가 r 인 경우, 입력벡터 $x(t) = [x_1(t), \dots, x_M(t)]$ 는 $s(t-1), \dots, s(t-r)$ 의 r 개의 벡터를 연결한 벡터이다. 은닉층의 전체입력값들로 이루어진 벡터를 $n(t) = [n_1(t), \dots, n_{H+N}(t)]$, 은닉층의 출력벡터를 $h(t) = [h_1(t), \dots, h_H(t)]$, 출력층의 출력벡터를 $o(t) = [o_1(t), \dots, o_M(t)]$ 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.

$$\begin{aligned} n_i(t) &= \sum_{j=1}^{H+N} w_{ij} z_j(t) & \begin{cases} i=1, \dots, H+N \\ j=1, \dots, H \\ k=1, \dots, N \end{cases} \\ h_j(t) &= f(n_j(t)) \\ o_k(t) &= \sum_{j=1}^N w_{kj} h_j(t) \end{aligned} \quad (3)$$

여기서 w_{ij} 는 i 번째 입력층 유니트에서 j 번째 은닉층 유니트로의 연결 가중치이고, w_{kj} 는 j 번째 은닉층 유니트와 k 번째 출력층 유니트로의 연결 가중치이다. $f()$ 는 시그모이드 함수를 나타낸다. 그리고 $z_j(t)$ 는 시각 t 에서 입력벡터 $x(t)$ 와 시각 $t-1$ 에서의 은닉층의 출력벡터 $h(t-1)$ 를 연결한 $H+N$ 차원 벡터의 i 번째 성분이다.

$$z_i(t) = [z_1(t), \dots, z_{H+N}(t)] \quad (4)$$

음절인식을 위한 회귀예측신경망에 관한 연구

$$= [h_1(t-1), \dots, h_H(t-1), x_1(t), \dots, x_{N_c}(t)]$$

$$= [C_1(t), \dots, C_H(t), x_1(t), \dots, x_{N_c}(t)]$$

여기서, $C_i(t)$ 는 문맥층의 벡터를 나타낸다.

시각 t 에서 교사벡터와 출력벡터 사이의 누적예측오차는 다음과 같다.

$$J(t) = \frac{1}{2} \sum_{k=1}^M [e_k(t)]^2 \quad (5)$$

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 시작 시각 t_0 에서 시작하여 최종 시각 T 에서 종료된다고 가정하면, 최종누적 예측오차 E 는 다음과 같다.

$$E = \sum_{t=t_0+1}^T J(t) \quad (6)$$

예측오차 제적을 최급강하법에 적용하면 가중치벡터, 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_{kj} 에 대한 전체 가중치 변화는 다음과 같다.

$$\Delta w_{ji}(t) = -\alpha \frac{\partial J(t)}{\partial w_{ji}} = \alpha \sum_{k=1}^M e_k(t) \frac{\partial h_k(t)}{\partial w_{ji}} \quad (7)$$

$$\Delta w_{kj}(t) = -\alpha \frac{\partial J(t)}{\partial w_{kj}} = \alpha e_k(t) o_j(t) h_j(t) \quad (8)$$

여기서 α 는 고정된 양(positive)의 학습률을 나타낸다.

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

$$\frac{\partial h_i(t_0)}{\partial w_{ji}} = 0 \quad (9)$$

그리고 회귀조건식은 다음과 같다.

$$\frac{\partial h_i(t+1)}{\partial w_{ji}} = f'(n_i(t)) \left[\sum_{l=1}^H w_{jl} \frac{\partial h_l(t)}{\partial w_{ji}} + \delta_{il} z_j(t) \right] \quad (10)$$

그러므로, 각 시각 t 에 대해서 가중치 변화식은 아래와 같다.

$$w_{ji}(t+1) = w_{ji}(t) + \Delta w_{ji}(t) \quad (11)$$

$$w_{kj}(t+1) = w_{kj}(t) + \Delta w_{kj}(t) \quad (12)$$

3.2 Jordan형 회귀예측신경망

그림 3는 Jordan형 회귀예측신경망의 구조를 나타내고

있다. 이 신경망은 출력층의 모든 유닛트가 문맥층으로 회귀연결되는 구조인 것과 문맥층에서 회귀계수를 적용한 것을 제외하면 Elman형 회귀예측신경망과 동일하다.

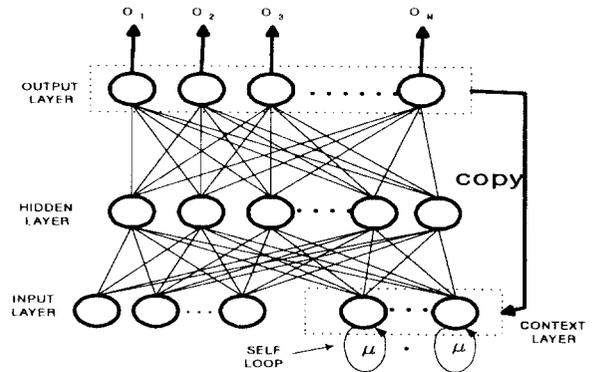


그림 3. Jordan형 회귀예측신경망

문맥층에서 회귀계수를 적용하는 수식은 식(13)과 같다.

$$C_i(t+1) = \mu C_i(t) + O_i(t)$$

$$= \mu(\mu C_i(t-1) - O_i(t-1)) + O_i(t) \quad (13)$$

$$= \sum_{\tau=0}^t \mu^\tau O_i(\tau) \quad (C_i(0) = 0)$$

4. 인식실험 및 조건

4.1 음성 DB 및 분석조건

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

100음절 데이터는 표 2와 같이 20대 남성화자 5명이 5회 발성한 100음절을 사용하였다. 음성DB는 신문사실과 초등학교 교과서중에서 사용 빈도수가 높은 음절을 조사하여 그 중 100개를 발래하였다. 5명의 남성화자가 5회 발성한 음성중에서 3회분은 학습용(1500개)으로, 나머지 2회분은 평가용(1000개)으로 사용하였으며 각각의 분석조건은 표1과 같다.

표 1. 음성데이터의 분석조건

	숫자음	100 음절
A/D 데이터	16 kHz, 16bit, signed	10 kHz, 12bit, signed
Filtering	LPF, 7 kHz	LPF, 5 kHz
프레임 길이	3.75 ms	5 ms
분석할 길이	16 ms	20 ms
특성파라미터	10차 LPC Melcepstrum	

표 2. 100음절 DB

가	간	갈	감	거	게	고	구	기
나	난	날	남	네	노	누	니	
다	단	달	담	대	도	두	디	
라	란	람	로	룽	루	리		
마	만	말	매	모	무	미		
바	반	발	보	부	비			
사	산	살	삼	세	소	수	시	
아	안	알	암	애	옥	우	이	
자	잔	잘	잠	제	조	주	지	
차	찬	찰	참	체	초	추	치	
키	킬							
타	탄	탈	탐	토	투	티		
파	판	팔	포	피				
하	한	할	할	호	후	히		

4.2. 예측차수에 따른 패턴의 설정

그림4은 예측차수에 따라 입력패턴과 교사패턴을 설정하는 방법을 나타낸다. 처음과 마지막 프레임이 교사패턴과 입력패턴으로 참여할 수 있게 하기 위해서 처음과 마지막 프레임을 복사하여 패턴을 설정한다. 이러한 진후 프레임에 대한 패턴설정 방법의 차이로 인하여 인식률에 큰 변화가 있음이 보고되어 있다.^[11]

이와 같이 설정되어진 패턴은 한 프레임씩 순차적으로 이동하면서 진음절구간에 대하여 신경망의 입력패턴과 교사패턴으로 사용된다.

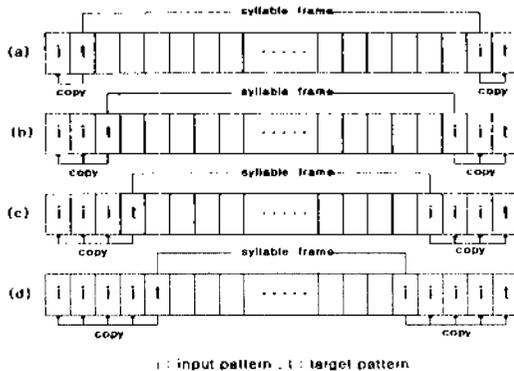


그림 4. 예측차수에 따른 패턴설정방법

4.3 인식실험

신경망의 구조는 숫자음에서는 예측차수가 1차, 2차, 3차, 4차 일때 입력층은 각각 10, 20, 30, 40 유닛으로 하고, 은닉층은 5,10,15,20 유닛 그리고 100음절에서는 예측차수가 2,3,4차로하고 은닉층은 5,10,15,20,25,30 유닛으로 하였으며 숫자음과 100음절 모두 출력층은 10 유닛으로 하여 한 프레임을 예측하도록 하였다. 그리고 숫자음에 대하여는 회귀연결하여 Elman형과 Jordan형 모두 문맥층에서 회귀계수를 적용한 경우의 실험을 추가

하였다.

- [방법 1] MLP형 예측신경망
- [방법 2] Jordan형 회귀예측신경망
- [방법 3] Elman형 회귀예측신경망
- [방법 4] Jordan형 회귀예측신경망
+ 문맥층에 회귀계수 적용(숫자음)
- [방법 5] Elman형 회귀예측신경망
+ 문맥층에 회귀계수 적용(숫자음)

이들 신경망 모델은 각 입력에 대한 교사패턴을 그 다음의 프레임으로 설정하고 예측레적을 학습한다. 매 순간의 회귀입력은 음성의 시변성에 따라 결정되고, 예측오차는 매 순간마다 누적된다. 인식실험은 모든 네트워크들에 입력을 인가하여 평균예측오차가 가장 최소가 되는 모델을 인식결과로 하였다.

5. 실험결과 및 고찰

표3과 표4는 숫자음에 대한 인식결과로 CHMM은 5상태 4출력분포의 경우 99.0%(198/200)로 인식하였고, Jordan형 회귀예측신경망은 은닉층 유닛20, 예측차수 4차에서 평가용데이터가 98.5%(197/200)로 CHMM과 거의 같은 인식결과를 나타내었다.

표5와 표6는 100음절에 대한 인식결과로 CHMM은 상태수가 증가할수록 인식률이 향상되었고 신경망 보다는 전반적으로 나은 결과를 나타내었으나 Jordan형 회귀예측신경망의 경우 은닉층 유닛 25이상일때는 오히려 CHMM의 5상태 4출력분포의 85.6%보다 향상된 인식결과를 보였다.

실험결과 각 신경망모델에 따른 인식성능은 예측차수와 은닉층 유닛의 개수에 따라 불규칙적인 변화를 보였다. 예측차수의 변화에 대해서는 예측차수가 증가함에 따라서는 두드러진 인식률의 향상을 보이지 않았다. 이는 입력층의 개수에 해당하는 예측차수에 따라서도 각 모델에 따라 서로 다른 예측능력을 가짐을 알 수 있지만 예측능력에는 커다란 영향이 없는 것으로 사료된다.

그리고 은닉층의 변화에 대해서는 예측신경망과 Jordan형 회귀예측신경망에서 은닉층의 유닛 수가 많을수록 인식률은 향상 되었지만 큰폭의 향상은 없었다. 그러나 Elman형 회귀예측신경망은 은닉층 유닛 수가 10일 때 가장 좋은 인식률을 보였지만 그 이상으로 증가하였을 때는 오히려 감소하였다.

음절인식을 위한 회귀예측신경망에 관한 연구

표 3. CHMM에 의한 숫자음 인식률(%)

상태수	학습	인식
4	98.7	98.5
5	100.0	99.0
6	99.3	98.0
7	99.8	98.0

표 4. 신경망에 의한 숫자음 인식률(%)

예측차수	은닉층	예측신경망		회귀예측신경망							
				회귀계수(λ)				회귀계수($\mu=0.5$)			
				Jordan형		Elman형		Jordan형		Elman형	
		학습	평가	학습	평가	학습	평가	학습	평가	학습	평가
1	5	89.5	86.0	88.5	85.5	87.5	87.0	88.5	86.5	87.7	85.5
	10	97.0	94.5	98.0	97.0	96.8	91.5	96.5	94.0	95.8	92.0
	15	98.0	96.0	98.0	95.0	98.7	94.5	98.7	95.5	98.3	96.0
	20	97.5	96.0	98.5	96.5	98.0	93.0	98.7	98.5	98.2	93.0
2	5	89.2	88.5	88.3	87.0	87.3	87.0	90.8	91.5	89.7	88.0
	10	97.2	94.0	97.2	93.0	96.8	94.0	97.2	97.0	98.7	96.0
	15	98.8	98.0	98.7	97.0	98.7	95.5	97.8	96.5	99.2	97.0
	20	99.3	97.0	99.2	97.5	99.0	95.0	98.3	96.5	99.2	96.5
3	5	89.5	88.0	90.2	87.5	90.0	89.0	90.2	91.5	89.5	97.5
	10	98.5	97.0	98.3	95.5	97.2	95.5	96.2	92.0	98.0	94.0
	15	98.7	97.5	98.5	97.0	99.0	95.5	99.3	97.0	96.7	94.5
	20	99.5	97.5	99.2	97.0	99.0	94.0	99.5	98.0	99.3	96.5
4	5	90.3	89.0	90.0	87.5	90.5	88.5	90.2	92.0	90.8	87.5
	10	97.8	96.5	97.2	94.5	98.2	97.5	98.5	96.5	97.5	94.5
	15	99.3	96.0	99.5	95.0	98.3	95.0	99.5	96.0	98.3	94.0
	20	97.8	95.0	99.7	98.5	99.3	95.0	99.3	96.0	98.7	94.5

표 5. CHMM에 의한 100음절 인식률(%)

상태수	학습	인식
4	96.92	82.8
5	98.10	85.6
6	98.62	87.7
7	98.89	87.4

표 6. 신경망에 의한 100음절 인식률(%)

예측차수	은닉층	예측신경망		회귀예측신경망			
				Jordan형		Elman형	
				학습	평가	학습	평가
2	5	77.6	67.2	80.3	68.6	80.1	66.3
	10	96.5	76.8	98.1	78.4	95.4	74.0
	15	98.1	80.2	98.9	83.1	76.5	57.1
	20	97.9	80.1	99.3	85.4	47.0	43.9
	25	97.9	80.4	99.6	86.0	23.6	22.5
30	97.9	79.5	99.5	83.9	15.2	22.0	
3	5	81.9	68.3	83.9	68.3	80.5	68.2
	10	98.1	78.6	98.1	77.9	94.8	75.0
	15	99.3	81.6	99.7	82.2	88.3	65.0
	20	99.3	80.8	99.9	84.9	47.5	41.2
	25	99.3	82.3	99.7	85.7	26.7	22.7
30	99.3	82.8	99.9	85.9	15.5	15.6	
4	5	86.4	69.9	85.7	70.9	80.7	66.9
	10	98.2	76.6	98.5	78.9	93.9	71.9
	15	99.5	80.9	99.9	80.7	88.3	65.0
	20	99.7	81.9	99.9	84.1	58.1	45.3
	25	99.6	82.4	99.9	86.0	24.9	22.0
30	99.6	82.6	99.9	86.2	16.1	15.9	

6. 결 론

본 논문에서는 MLP형 예측신경망, Jordan형과 Elman형 회귀예측신경망을 사용하여 예측차수와 은닉층의 유닛수와의 변화에 따른 인식성능을 연속분포 IIMM(CHMM)과 비교하였다. 신경망에 의한 숫자음 인식결과는 CHMM과 거의 유사한 인식성능을 보였고, 100음절에서는 은닉층의 유닛수가 25개 이상인 Jordan형 회귀예측신경망이 5상태 CHMM의 85.6%보다는 향상된 인식성능을 보였으며, 6상태 이상의 CHMM보다는 다소인식률이 낮게 나타났다.

100음절에서 신경망에서의 인식결과는 Jordan형 회귀예측신경망이 예측신경망보다 우수한 인식성능을 나타내었고, Elman형 회귀예측신경망은 오히려 인식성능이 저하되었다. 예측차수의 변화에 대한 인식성능은 큰 변화가 없었고, 은닉층 유닛 수의 변화에 따른 인식성능은 Jordan형에서는 향상되었으며, Elman형에서는 은닉층 유닛수가 10 이상에서는 오히려 저하되었다. 이는 은닉층의 과도적인 정보를 회귀시켜 현재의 예측에 반영하는 구조보다는 이전 출력층의 예측된 정보를 회귀시켜 현재의 예측에 반영하는 구조가 신경망의 예측능력향상에 유리하기 때문인 것으로 사료된다. 향후, 예측

음절인식을 위한 회귀예측신경망에 관한 연구

신경망구조의 대단위 음절인식을 바탕으로 연속음성인식에의 이용 가능성에 관하여 연구할 계획이다. 호, pp. 69-74. 1998.

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: Biocal, 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, San Diego, La Jolla, California, May 1986.
- [4] J. L. Elman, "Finding structure in time," Technical Report CRL-8801, Center for Reserch 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] 류재관, 라경민, 임재열, 성평모, 안수길, "회귀신경예측모델을 이용한 음성인식," 대한전자공학회 논문지, 제11권 B편, 제32권, 1995.
- [9] 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.
- [10] Soo Hoon Kim, Sang-Boum, Kang-In Hur, "The Recognition of Korean Syllables Using Neural Predictive HMM," ICASP, vol.1 of 2, pp.427-431, 1997.
- [11] 김주성, 한학용, 허강인, "패턴구성에 따른 RPNN의 음성인식 성능비교," 동아대학교 정보통신 연구소논문지, 제5권, 1호, pp.123-128, 1997
- [12] 한학용, 김주성, 허강인, 안점영, "회귀예측신경망을 이용한 음절인식," 부산경남 춘계합동학술 논문지 제7권1