

경쟁학습 신경망의 환경 적응성

최두일[○] · 박양수
공주대학교 전기공학과

Circumstance Adaptability of Competitive Learning Neural Networks

Choi, Doo-Il[○] · Park, Yang-Su

Department of Electrical Engineering Kongju National University

Abstract - When input circumstance is changed abruptly, many nodes of Competitive Learning Neural Networks far from new input vector may never win, and therefore never learn. Various techniques to prevent these phenomena have been reported. We proposed a new technique based on Self Creating and Organizing Neural Networks, and which is compared to Self Organizing Feature Map and Frequency Sensitive Neural Networks.

1. 서 론

"winner take all"법칙을 이용하는 경쟁학습 신경회로망에서 가장 큰 문제점은 입력 벡터군과 거리가 먼 가중치 벡터의 노드는 경쟁에서 항상 지게되므로 학습되지 못하는 것이다. 이러한 죽은 노드의 발생 문제는 여러 가지 방법으로 막을 수 있다. Kohonen은 승리 노드의 이웃 노드까지 같이 학습시키는 방법을 제안하여 자기 구조화 특징도(Self Organizing Feature Map : SOFM)으로 명명하였다[1]. 이 방법은 학습의 초기에는 이웃의 범위를 크게하고 시간에 따라 점차적으로 이웃의 범위를 줄어나감으로서 죽은노드의 발생을 억제하고 있다. 그러나 입력이 비정상 시퀀스(nonstationary sequence)로 제시되는 경우는 다수의 죽은 노드를 발생시키게 된다.[2] DeSieno[3]는 Kohonen의 신경회로망에 빈도감지 기법을 추가하여 비정상 입력 시퀀스에 대해서도 최적의 벡터 양자화를 이룰 수 있는 빈도 감지 신경망(Frequency Sensitive Neural Network : FSNN)을 제안하였다. 이 방법은 비정상 시퀀스에 대해서도 최적의 양자화를 이루는 것으로 평가되고 있다. 그러나 본 연구에서는 FSNN이 매우 오랜 학습시간이 요하며, 오버플로우의 발생 가능성 또한 존재함을 보이고, 매우 빠르게 새로운 입력 환경에 적응해 가는 신경망 기법을 제안한다.

제안된 기법은 고속 화자 적응형 음성인식에 응용하였다..

2. 신경회로망의 환경 적응성

그림 1은 SOFM이 입력환경이 갑자기 바뀔 경우에 적응해

가는 과정을 보이고 있다. 입력이 $t < 5000$ 인 경우에는 좌하의 상자 내에서 균일 분포하고, $t > 5000$ 인 경우에는 우상의 상자 내에서 균일 발생하는 경우로서, 대다수의 신경세포들이 과거의 입력환경에 적응해 있는 채 학습되지 못함을 알 수 있다.

이 남아 있는 신경세포는 미래에 발생될지 모를 입력에 대비하고 있으나, 현재의 환경에선 죽은 노드로 남게 된다. 이와 같은 일련의 문제점 - Stability and Plasticity dilemma - 은 빈도 감지 신경망을 이용하거나 자기 생성 및 구조화 신경회로망(Self Creating and Organizing Neural Network : SCONN)[2]의 학습법을 변형하여 해결할 수 있다. 빈도 감지 신경망에서는 승리 노드(winner node)의 빈도 감지 계수를 증가시킴으로써 승리 노드가 또다시 승리하기 어렵도록 하는 방법을 이용한다. 따라서 죽어있는 노드들도 언젠가는 승리노드가 되게 되므로써 학습을 받을 기회가 생기게 된다. 이 방법은 Stability and Plasticity dilemma를 해결하는 최적의 수단으로 평가되고 있다. 그림 2는 FSNN의 비정상 입력 시퀀스에 대한 응답의 천이를 보이고 있다. 입력 조건은 위의 SOFM 경우와 같도록 하였다. FSNN은 새로운 입력환경에 잘 적응해 나가고 있으나 매우 오랜 학습시간을 요함을 알 수 있다. 또한 빈도 감지 계수가 계속적으로 증가되므로 오버플로우(overflow)를 발생할 위험이 있다.

표 1의 SCONN 알고리즘의 7단계를 보면 하나의 승리 노드만이 학습된다. 이를 변형하여 활성 노드들을 모두 학습함으로써 Stability and Plasticity dilemma를 해결할 수 있다. SCONN 알고리즘에서는 경쟁에서 진 노드들의 응답범위가 증가하므로 죽어있던 노드들도 언젠가는 활성 노드가 되어 학습 받을 수 있게 된다. 활성 노드가 되면 응답범위가 감소하기 때문에 오버플로우의 발생 위험은 전혀 없다. 그림 3은 SCONN의 비정상 입력 시퀀스에 대한 응답의 천이를 보이고 있다. 매우 빠르게 새로운 입력환경에 적응해 나감을 알 수 있다. 실험 조건은 위의 두 경우와 같도록 설정하였으며, 승리 노드의 학습율은 0.18, 활성 노드의 학습율은 0.02로 한 경우이다.

표 1 SCONN 알고리즘

Table 1 An algorithm for SCONN

Step 1. Initialize Weights
Step 2. Present New Input
Step 3. Calculate Distance to All Node(s)
Step 4. Find Active Node(s) and a Winner Node
Step 5. If Active Node Does not Exist, then go to step 8
Step 6. Decrease Response Ranges of Active Node(s) Increase Response Ranges of Inactive Node(s)
Step 7. Adapt Weights of Winner Node (or Winner node and its family nodes) go to Step 2
Step 8. Create a Son Node from an Inactive Winner (Mother) Node go to Step 2

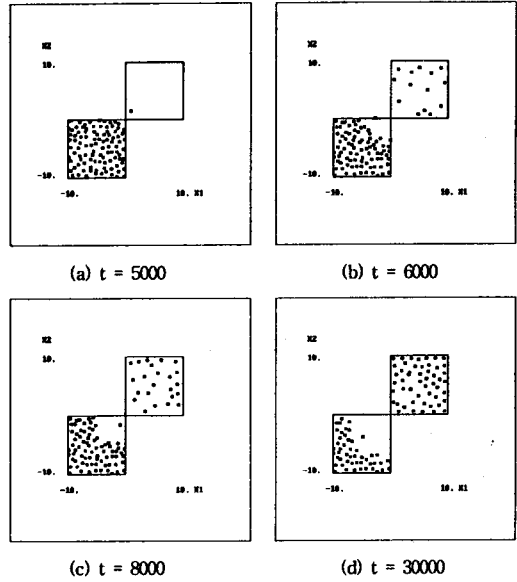


그림 2 FSNN의 비정상 입력 시퀀스에 대한 응답의 천이

Fig. 2 The transition FSNN's response for nonstationary input sequence

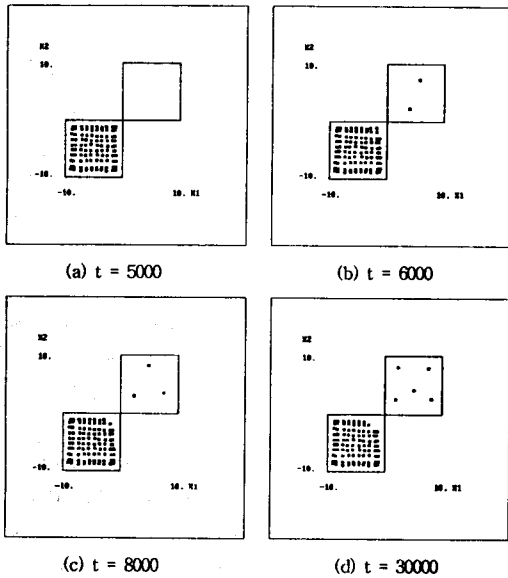


그림 1 SOMF의 비정상 입력 시퀀스에 대한 응답의 천이

Fig. 1 The transition SOMF's response for nonstationary input sequence

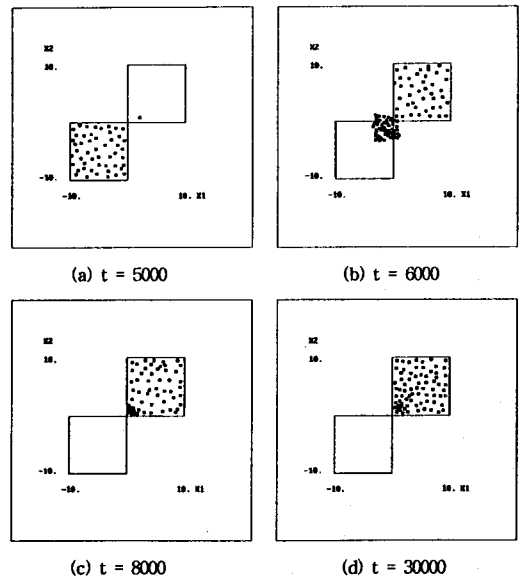


그림 3 SCONN의 비정상 입력 시퀀스에 대한 응답의 천이

Fig. 3 The transition SCONN's response for nonstationary input sequence

3. 고속 화자 적응형 음성인식에의 응용

3.1 음성데이터 수집

본 연구에서 수집한 27개의 단음절 음성 데이터는 다음과 같다.

13개 기본초성 : /가/, /나/, /다/, /라/, /마/, /바/, /사/,
/자/, /차/, /카/, /타/, /파/, /하/

7개 기본 중성 : /아/, /에/, /이/, /오/, /우/, /으/, /이/

7개 기본 종성 : /악/, /안/, /안/, /알/, /암/, /압/, /양/

수집된 음절 음성신호에서 음소 단위로의 절단(segmentation)은 특별한 알고리즘을 사용하지 않고 인공적으로 이루어졌다. 즉, 초성은 시작점부터 121.6msec동안을, 중성은 시작점부터 428.8msec 동안을, 그리고 종성은 끝점부터 224.0msec동안을 절단하였다. 단, 종성중 /악/, /안/, 과 /압/은 발음 구간이 매우 짧으므로 70.4msec 동안을 절단하였다.

3.2 전처리

개략적인 전처리 과정은 다음과 같다.

- 1) 마이크로폰.
- 2) 50 - 5KHz의 대역폭을 갖는 엔터 얼라이어징 대역통과 필터 (Anti-aliasing band pass filter).
- 3) 표본화율 10 KHz와 16비트의 A/D 변환기.
- 4) 25.6msec의 구간과 75%의 시간 중첩을 갖는 이동 해밍창 (moving hamming window).
- 5) $y(t) = x(t) - \alpha x(t-1)$ 을 이용한 Pre-emphasis
- 6) 256점 고속 푸리에 변환(256 point FFT)
- 7) 스펙트럼 파워 (spectral power)의 계산
- 8) bark scale을 이용한 16차원의 벡터로의 그룹핑 (grouping).

3.3 인식결과

화자 A의 각 음소에 대한 화자 종속 인식율은 13개 초성에 대하여 81.73%, 7개 중성에 대하여 98.21%, 7개 종성에 대하여 60.71%를 얻었다. 화자 A의 음성으로 학습된 시스템에 화자 B의 음성을 입력한 경우의 인식율은 13개 초성에 대하여 16.35%, 7개 중성에 대하여 57.14%, 7개 종성에 대하여 19.64%를 얻었다.

화자 A의 음성으로 학습된 시스템이 화자 B의 음성으로 재적용된 후의 인식율은 13개 초성에 대하여 63.44%, 7개 중성에 대하여 89.29%, 7개 종성에 대하여 60.71%를 얻었다.

4. 결 론

자기 조직형 신경망의 성능은 최적성, 배열성 및 적응성 등으로 평가될 수 있다. 본 연구에서는 SCONN의 학습 방법을 개선하면 최적성과 배열성 뿐 아니라 고속의 환경 적응성을 갖게됨을 보였다. 이를 위하여 비정상 입력환경에서 SOFM, FSNN과 SCONN이 새로운 환경에 적응해 가는 과정을 비교하였으며, SCONN에서 완벽한 환경 적응성이 나타남을 보였다. 또한, SCONN의 최적성과

빠른 적응성을 이용하여 음소 인식실험을 수행하였다. 특히 갑자기 화자가 바뀔 경우에, 바뀐 화자의 음성신호에 빠르게 적응해감을 보였다.

참고문헌

- [1] T. Kohonen, Self Organization and Associative Memory. 2nd edition, Springer-Verlag, ch. 5, pp.119-157, 1988.
- [2] D. Choi and S. Park, "Self-Creating and Organizing Neural Networks," IEEE Trans. on Neural Networks, Vol. 5, No. 4, pp. 561-575, July 1994.
- [3] D. DeSieno, "Adding a Conscience to Competitive Learning", Proc. IEEE ICNN, vol. 1, pp. 117- 124, San Diego, 1988.