
유전자알고리즘을 이용한 시그모이드 활성화 함수 파라미터의 최적화와 이중나선 문제의 입력공간 패턴인식 분석

Optimization of Sigmoid Activation Function Parameters using Genetic Algorithms and Pattern Recognition Analysis in Input Space of Two Spirals Problem

이상화

서원대학교 정보통신공학과

Sang-Wha Lee(swl@seowon.ac.kr)

요약

본 논문에서는 유전자알고리즘을 이용한 시그모이드 활성화 함수 파라미터의 최적화와 이중나선기준문제(two spirals benchmark problem)의 입력공간 패턴인식 상태를 분석 한다. 실험을 위하여 캐스케이드 코릴레이션 학습 알고리즘(Cascade Correlation learning algorithm)을 이용한다. 첫 번째 실험에서는 기본적인 시그모이드 활성화 함수를 사용하여 이중나선 문제를 분석하고, 두 번째 실험에서는 시그모이드 활성화 함수(sigmoidal activation function)의 파라미터 값이 서로 다른 함수를 사용하여 8개의 풀을 구성한다. 세 번째 실험에서는 시그모이드 함수의 변위를 결정하는 세 개의 파라미터 값을 유전자 알고리즘을 이용하여 얻고 이 파라미터 값들이 적용된 시그모이드 함수들은 후보뉴런의 활성화를 위해서 사용된다. 이러한 알고리즘의 성능평가를 위하여 각 학습단계 마다 입력패턴공간에서 인식된 이중나선의 형태를 보여준다.

■ 중심어 : | 캐스케이드코릴레이션 알고리즘 | 활성화함수 | 시그모이드함수 | 이중나선 문제 |

Abstract

This paper presents a optimization of sigmoid activation function parameter using genetic algorithms and pattern recognition analysis in input space of two spirals benchmark problem. To experiment, cascade correlation learning algorithm is used. In the first experiment, normal sigmoid activation function is used to analyze the pattern classification in input space of the two spirals problem. In the second experiment, sigmoid activation functions using different fixed values of the parameters are composed of 8 pools. In the third experiment, displacement of the sigmoid function to determine the value of the three parameters is obtained using genetic algorithms. The parameter values applied to the sigmoid activation functions for candidate neurons are used. To evaluate the performance of these algorithms, each step of the training input pattern classification shows the shape of the two spirals.

■ keyword : | Cascade Correlation Algorithm | Activation Function | Sigmoid Function | Two Spirals Problem |

1. 서론

패턴인식 분야에서 해결하기 어려운 문제 중의 하나가 움직이는 물체의 인식이다. 이러한 문제 해결의 시발점으로써 신경회로망의 새로운 알고리즘을 제안하고 움직이는 물체 인식 실험을 위한 기준문제(benchmark problem)로서 이중나선문제(two spirals problem)를 사용하였다[2][9]. 특히 이 문제는 움직이는 물체를 정확히 인식 할 수 있는 뉴로칩을 개발하고 제작하여 미사일 등에 탑재하면 목표물을 명중시킬 수 있는 미국방성 프로젝트의 일환으로 연구되었다. 이러한 문제의 해결을 위하여 Fahlman은 캐스케이드 코릴레이션(Cascade Correlation (CC))을 제안하게 되었고 패턴인식을 위한 학습은 퀵프로파게이션(quickpropagation) 알고리즘을 이용하였다[2].

CC 알고리즘은 초기에 은닉뉴런과 은닉층 없이 입력 뉴런과 출력뉴런의 완전한 연결에 의하여 시작한다. 그 후에 CC에서 은닉뉴런은 네트워크에 한 번에 한 개씩 추가되고 선택된 가중치의 값은 변화하지 않는다. 여기에서 추가할 뉴런의 출력과 네트워크의 잔여오차의 상호관계값(correlation value)의 극대화를 시도한다. 새로운 은닉뉴런의 생성을 위하여, 후보뉴런들로 구성된 하나의 풀(pool)에서 후보뉴런(candidate unit)들이 학습할 수 있는 시그널은 네트워크의 모든 입력과 이미 존재하고 있는 모든 은닉뉴런들과의 연결들을 통하여 전달된다. 네트워크는 주어진 패턴들에 대하여 한 번의 학습 후에 후보뉴런들의 입력가중치(input weight)를 수정한다. 이를 통하여 후보뉴런에서 생성된 값과 모든 출력뉴런에서 측정된 잔여출력오차의 상호관계(correlation)의 합계 중에서 가장 큰 값을 갖고 있는 후보뉴런이 선택되어 은닉뉴런을 구성하고 네트워크와 연결한다. 이러한 학습 과정은 주어진 학습오차의 범위에 도달할 때까지 반복되어 진행된다. [그림 1]은 학습을 통하여 n개의 은닉뉴런이 생성되어 구성되는 CC 네트워크를 보여준다. 그림에서 사각형으로 표시된 연결들은 고정되고(한번 선택된 가중치는 다음 학습단계에서도 수정 없이 계속 사용된다), x와 함께 표시한 연결들은 반복해서 학습한다(각 학습단계마다 가중치는 다

시 수정된다).

본 논문의 목적은 시그모이드 활성화함수의 사용 방법에 따라서 패턴인식 결과에 많은 영향을 미칠 수 있음을 보여주는 것이다. 이는 첫 번째 주어진 이중나선의 모양과 가급적 똑같은 형태의 인식된 패턴의 그림을 나타내야하고 두 번째는 가급적 작은 수의 은닉뉴런을 생성하여 네트워크를 구성하는 것이다. 이를 위하여 각 학습단계마다 입력패턴공간에서 인식된 이중나선의 상태를 그림으로 보여줌으로써 단계별 학습 과정의 인식도를 확인 할 수 있도록 하였고 또한 생성된 은닉뉴런의 수를 비교하였다.

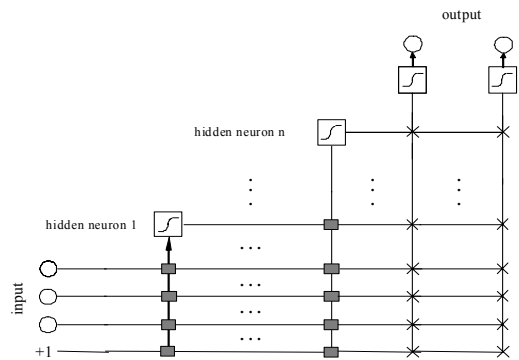


그림 1. CC의 구성

II. 이중나선문제(Two spirals problem)

이중나선문제(Two Spirals Problem)는 Alexis Wieland (MITRE Corp)의 기준문제(benchmark problem)이다[2][7]. Fahlman은 [2]에서 Moving-Target-Problem의 해결을 위하여 캐스케이드 코릴레이션 네트워크를 제안하면서 이 기준문제를 처음으로 실험에 사용하였다. 그 이후 이중나선문제는 신경망분야에서 중요한 기준문제로 대두되었다. 특히 움직이는 물체를 정확히 인식하기 위한 알고리즘을 테스트하기 위하여 많이 사용된다. 이중나선의 알고리즘은 연속된 값으로 이루어진 두 개의 입력을 갖고 한 개의 이원값의 출력을 갖는다. 서로 끼워 맞추어진 두 개의 나선은 학습을 통하여 구분하는 것이 목표이다. 각

나선은 97개의 측량점으로 이루어졌으며 3 회전 되었다. 따라서 32개의 측량점들과 함께 각 나선은 한번 회전 한다. 그리고 마지막 한 개의 점이 첨가된다. 각 점은 x축과 y축에 의하여 두 개의 실수값으로 나타낸다. 네트워크의 출력은 한 나선을 위하여 1.0 그리고 다른 나선을 위하여 0.0의 값을 나타낸다. [그림 2]는 이중나선문제를 위한 트레이닝 점들을 표시하였다.

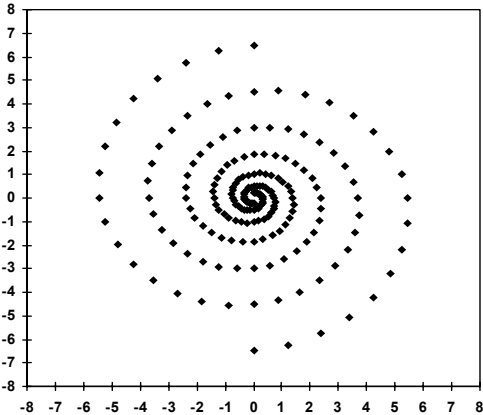


그림 2. 이중나선 문제를 위한 트레이닝점들

III. 기본 시그모이드 활성화 함수를 사용한 실험

[2]에서와 같이 CC는 이중나선문제의 패턴을 인식하기 위하여 하나의 풀(pool)을 구성하고 활성화 함수로는 기본 시그모이드 함수를 사용하였다. 보통 하나의 풀은 4~8개의 후보뉴런으로 구성하고 같은 활성화 함수를 사용한다. 한 풀에 10개 이상의 많은 후보뉴런을 형성하였다고 해서 반드시 항상 좋은 결과가 나오는 것은 아니다. 따라서 후보뉴런은 4~5개 정도가 적당하고 최대 8개 정도까지 풀을 구성하는 것이 일반적이다. 후보뉴런들은 모든 입력뉴런과, 바이어스와 그리고 이미 생성된 모든 은닉뉴런과 연결한다. 풀에 있는 모든 후보뉴런들은 같은 입력시그널을 받으며 각 패턴들로부터 원인이 된 같은 잔여오차를 갖고 있다. 그들은 서로 아무런 영향을 주고받지 않고 학습하는 동안 네트워크에도 영향을 주지 않기 때문에 풀에서 병렬로 학습한

다. 학습 중에 뉴런의 학습 횟수가 주어진 파라미터값에 도달 했을 때 가장 큰 상호관계값을 갖고 있는 한 후보뉴런이 은닉뉴런으로 선택되어 네트워크에 첨가되고 다른 뉴런들과 연결된다. [그림 3]은 실험에서 사용한 기본 시그모이드 활성화 함수를 보여준다.

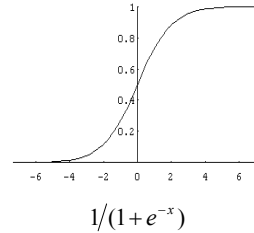
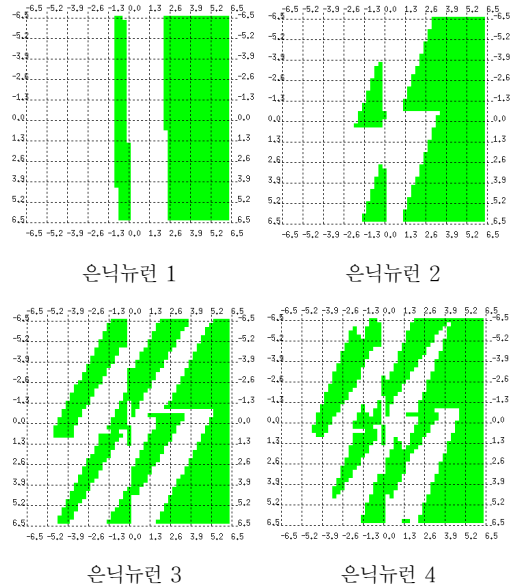


그림 3. 기본 시그모이드 활성화 함수

첫 번째 실험은 차후에 실행할 실험과 비교를 위하여 [2]에서와 같은 조건에서 실행한 실험이다. [그림 4]는 각 학습 단계에서 은닉뉴런이 생성 될 때 마다 입력패턴 공간에서 이중 나선의 인식 상태를 나타냈다. 그리고 [그림 5]는 학습오차와 생성된 은닉뉴런 수의 비교를 나타내는 그래프이다.



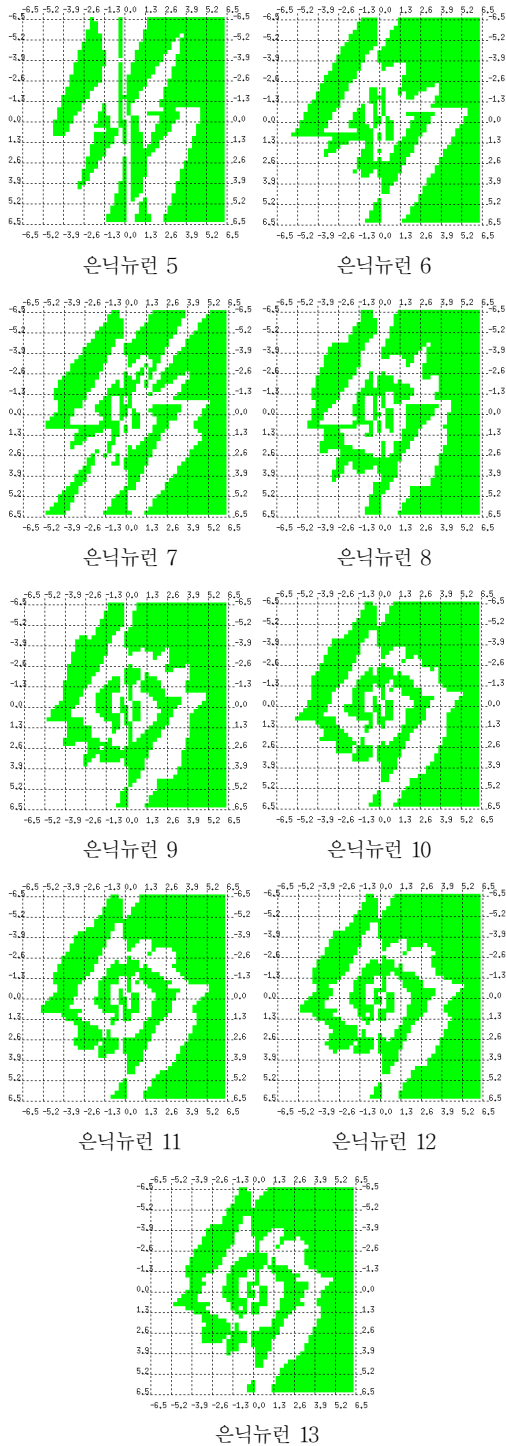


그림 4. 생성된 은닉뉴런의 숫자와 학습오류와의 관계

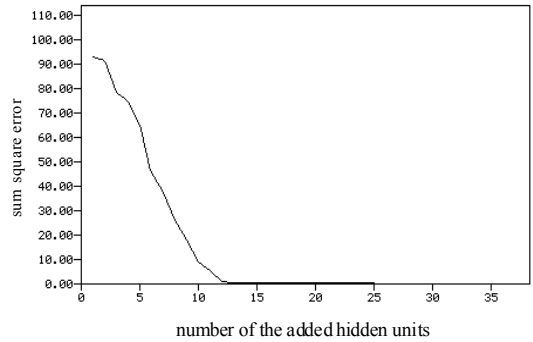
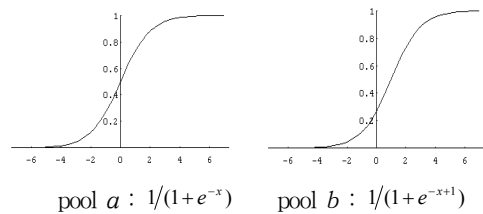


그림 5. 시그모이드 활성화 함수를 사용한 실험에서 발생한 학습오차와 생성된 은닉뉴런 수의 비교

CC에서 기본 시그모이드 활성화 함수를 사용한 경우에 이중나선 문제를 해결하기 위하여 13개의 은닉뉴런들이 필요 했다. 본래의 이중나선 형태는 등근데 각 은닉뉴런의 생성 마다 인식된 이중나선의 입력패턴은 상당히 각이진 형태를 보여준다. 비록 인식된 형상이 깨끗하지는 않지만 이중 나선의 형태라는 것을 인식하는 것에는 무리가 없다.

IV. 다양한 변형된 시그모이드 활성화 함수를 사용한 실험

이중나선 패턴 인식의 향상을 위하여 본 논문에서는 변형된 여러 형태의 다양한 시그모이드 활성화 함수를 사용한다. 이를 위하여 8개의 풀을 구성하고 각각의 풀에는 4개의 후보뉴런들을 나열한다. 각 풀에 구성된 시그모이드 함수는 아래 [그림 6]과 같다.



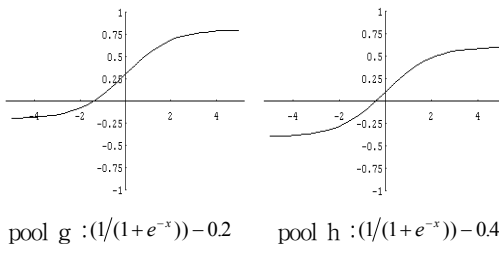
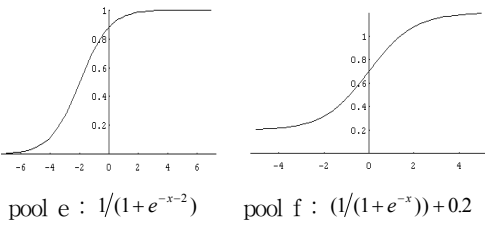
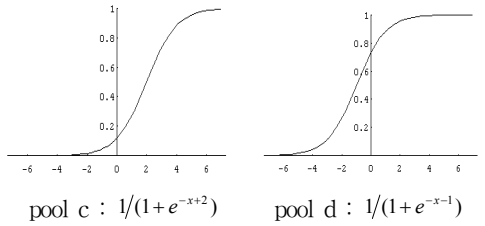


그림 6. 후보뉴런과 은닉뉴런의 활성화 함수로 사용한 변형된 시그모이드 활성화 함수

두 번째 실험에서는 가급적 적은 수의 은닉뉴런을 생성하여 학습을 완료하고 또한 입력패턴공간의 이중나선 인식 상태를 개선하는 것이 목적이다. [그림 7]은 각 학습 단계마다 인식된 입력패턴의 상태를 보여준다. 그리고 [그림 8]은 생성된 은닉뉴런의 수에 따른 학습 오차를 비교하였다.

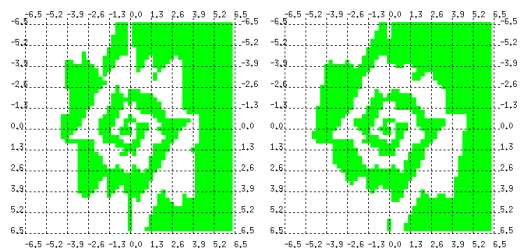
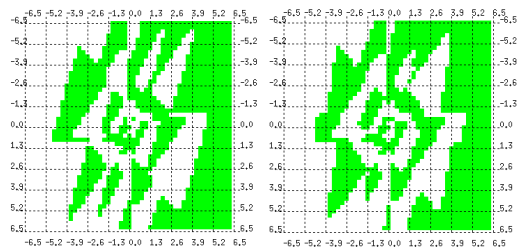
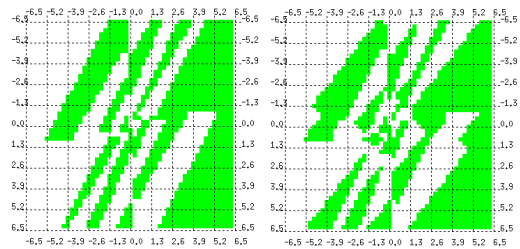
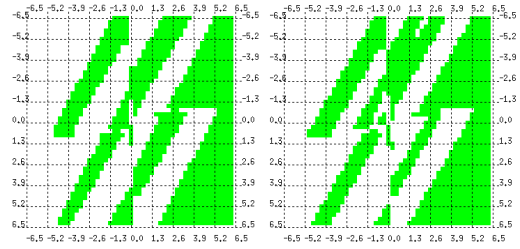
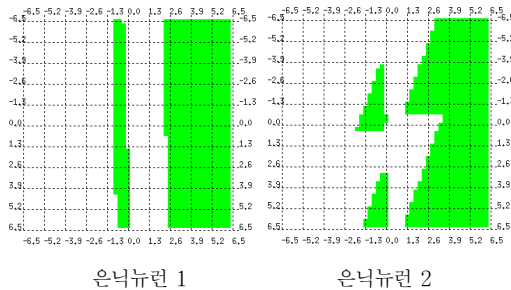


그림 7. 생성된 은닉뉴런의 숫자와 학습오류와의 관계

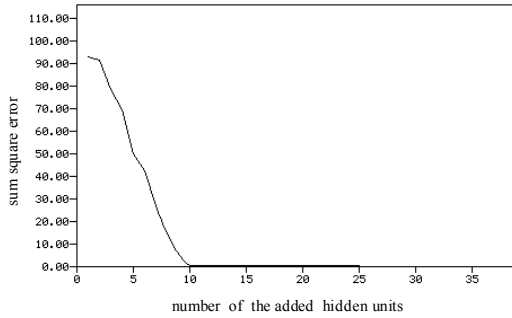


그림 8. 시그모이드 활성화 함수를 사용한 실험에서 발생한 학습오차와 생성된 은닉뉴런 수의 비교

두 번째 실험에서 필요한 은닉뉴런은 10개로서 첫 번째 실험에 비해서 3개 적게 소요 되었다. 그 만큼 빠른 학습 속도를 보였고 입력패턴 공간의 인식된 이중 나선을 보면 아직도 각이진 형태를 보이지만 전 실험에 비해서는 인식 상태가 조금 나아졌다. [표 1]에서는 각 학습 단계마다 선택된 은닉뉴런이 속한 풀을 기록 하였다.

표 1. 학습 중에 선택된 활성화 함수

선택된 활성화함수			
뉴런 1	뉴런 2	뉴런 3	뉴런 4
pool g	pool g	pool d	pool a
뉴런 5	뉴런 6	뉴런 7	뉴런 8
pool c	pool g	pool a	pool f
뉴런 9	뉴런 10		
pool g	pool a		

V. 유전자 알고리즘을 이용한 최적화

본 논문에서 시행한 세 번째 실험은 4개의 후보뉴런으로 하나의 풀을 구성한다. 이 후보뉴런들의 학습을 위하여 필요한 시그모이드 활성화함수의 세 개의 파라미터 값은 유전자 알고리즘(Genetic Algorithm(GA))을 이용하여 구한다. 즉 GA를 통해서 활성화함수의 파라미터 값을 구해서 CC에 적용하여 이중나선문제를 해결하는 최적화를 시도 한다.

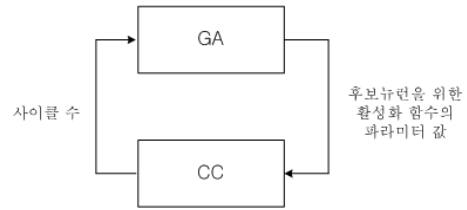


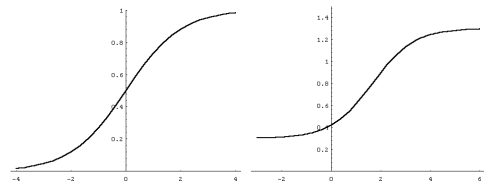
그림 9. 후보뉴런을 위한 파라미터 최적화

유전자 알고리즘의 실행을 위하여 주어진 파라미터 값들은 다음과 같다.

- 크로모솜(염색체)의 길이: 26
- 집단크기: 50
- 돌연변이비율: 0.04
- 크로스오버 포인트의 수: 1
- 세대 수: 100

유전자 알고리즘을 이용하여 시그모이드 활성화 함수의 파라미터를 최적화하기 위해서 (1)와 같은 식이 주어진다. [그림 10]은 파라미터 값에 따른 변형된 시그모이드 함수의 예를 보여준다. [그림 10]에서 (a)는 a=1, b=0, c=0, 그리고 (b)는 a=1.2, b=2, c=0.3을 갖고 있는 시그모이드 함수를 각각 나타낸다.

$$f(x) = \frac{1}{1 + e^{-ax+b}} + c \tag{1}$$



(a): $\frac{1}{1+e^{-x}}$ (b): $\frac{1}{1+e^{-1.2x+2}} + 0.3$

그림 10. 후보뉴런을 위한 파라미터 최적화

후보뉴런과 은닉뉴런을 위하여 사용할 시그모이드 활성화함수의 파라미터 a, b 와 c의 값을 유전자알고리즘을 이용하여 구한다. (1)의 주어진 일반형태의 시그

모이드 함수의 특성과 이중나선문제의 입력패턴공간에 따라서 시그모이드 함수의 파라미터 a의 값은 $0.5 \leq a \leq 1.5$, 파라미터 b의 값은 $-2.5 \leq b \leq 2.5$, 그리고 파라미터 c의 값은 $-2.5 \leq c \leq 2.5$ 의 구간으로 한정한다.

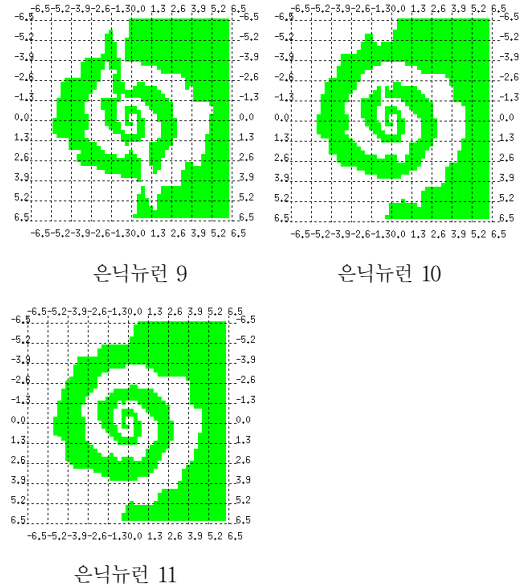
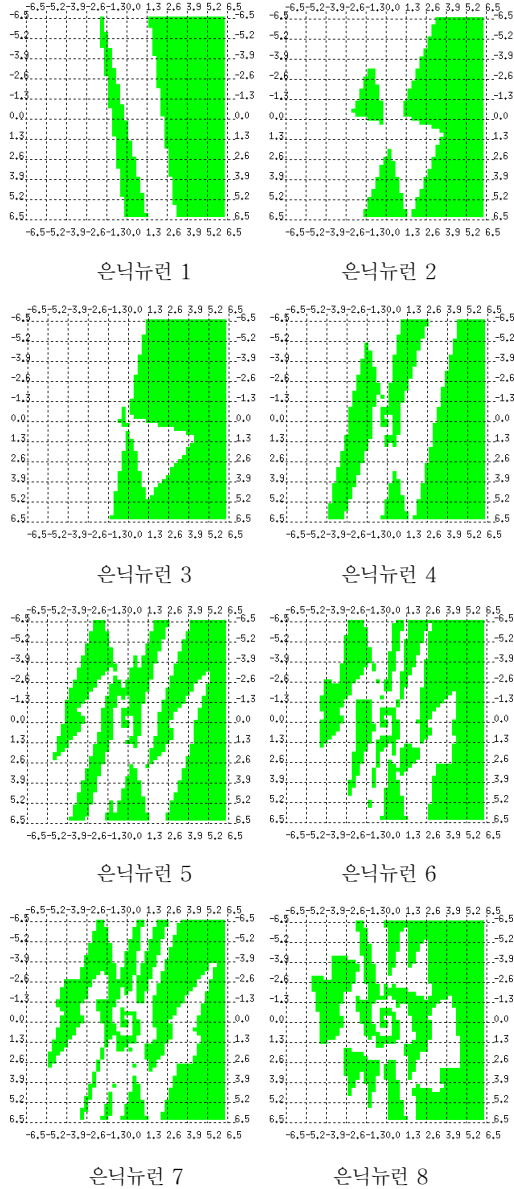


그림 11. 생성된 은닉뉴런의 숫자와 학습오류와의 관계

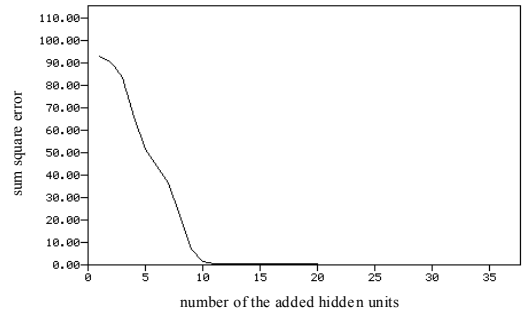


그림 12. 시그모이드 활성화 함수를 사용한 실험에서 발생한 학습오차와 생성된 은닉뉴런 수의 비교

유전자 알고리즘을 통해서 선택된 각 은닉뉴런을 위한 시그모이드 함수의 파라미터의 값은 [표 2]에 기술하였다. 주어진 값은 (a, b, c)의 순서로 기입하였다.

표 2. 학습 중에 선택된 활성화 함수

선택된 활성화함수의 파라미터		
은닉뉴런 1 (1.4, 0.7, 1.5)	은닉뉴런 2 (0.6, -1.5, -2.1)	은닉뉴런 3 (1.2, 2.4, 2.0)
은닉뉴런 4 (0.9, -1.3, -1.4)	은닉뉴런 5 (1.4, 1.0, -1.4)	은닉뉴런 6 (1.0, -1.0, 0.8)
은닉뉴런 7 (1.3, 2.4, 1.8)	은닉뉴런 8 (0.8, -0.9, 1.7)	은닉뉴런 9 (0.7, -1.5, 0.7)
은닉뉴런 10 (1.1, 1.2, -2.3)	은닉뉴런 11 (1.0, -0.2, 1.1)	

두 번째 실험과 비교하면 이중나선의 학습을 위하여 한개의 은닉뉴런이 더 필요했다. 두 번째 실험은 이중나선을 인식하는데 있어서 입력공간에서 각이진 형태로 패턴을 인식하게 되어 부자연스럽게 이중나선이 인식되었는데 반해서 세 번째 실험에서는 뚜렷하게 이중나선이 인식된 것을 확인할 수 있다. 본 논문에서는 시그모이드 활성화 함수의 사용 방법에 따라서 패턴인식의 결과에 많은 영향을 미치는 것을 보여주었다. 이를 극대화하기 위하여 유전자 알고리즘을 이용하여 시그모이드 활성화 함수의 변형을 위한 파라미터의 최적화를 시도하였고 실험을 통하여 보여 주었다.

VI. 결론 및 향후 과제

캐스케이드 코릴레이션 알고리즘을 이용한 이중나선 문제 해결의 최적화를 위하여 기본 시그모이드 함수만 사용하는 것 보다 여러 개의 다양하게 변형된 시그모이드 활성화 함수를 사용하는 것이 이중나선의 패턴 인식에 영향을 주는 것을 보여주었다. 이는 각 학습과정에서 패턴 분류(classification)를 통하여 패턴을 인식하는 과정을 단계별로 보여줌으로써 확인할 수 있었다. 따라서 결론적으로 입력패턴 공간에서 활성화 함수는 입력 패턴공간을 고려하여 파라미터 값에 변화를 주면 한 개의 시그모이드 활성화 함수를 사용하는 것보다 더 좋은 결과를 얻을 수 있다. 더 나아가서 패턴인식 학습을 위하여 유전자 알고리즘을 이용한다. 이러한 실험을 통하여 패턴 인식 능력을 최적화 시킬 수 있음을 본 논문에서 보여 주었다. 앞으로는 시그모이드 함수 이외의 함수를 이용하여 최적화 시키는 방법에 관하여 연구할 필요가 있다.

참 고 문 헌

- [1] F. Dandurand, V. Berthiaume, and T. R. Shultz, "A systematic comparison of flat and standard cascade-correlation using a student-teacher network approximation task," *Connection Science*, Vol.19, No.3, pp.223-244, 2007.
- [2] S. E. Fahlman and C. Lebiere, "The cascade-correlation learning architecture," *Advances in Neural Information Processing Systems 2*, Morgan Kaufmann, 1990.
- [3] S. E. Fahlman, "The Recurrent Cascade-Correlation Architecture," *Advances in Neural Information Processing Systems 3*, Morgan Kaufmann, pp.190-198, 1991.
- [4] X. Z Gao, X. Wang, and S. J. Ovaska, "A novel hybrid optimization method with application in Cascade-Correlation neural network training," *Proceedings, 8th International Conference on Hybrid Intelligent Systems*, Article number 4626728, pp.793-800, 2008.
- [5] B. Hammer, A. Micheli, and A. Sperduti, "Universal approximation capability of cascade correlation for structures," *Neural Computation*, Vol.17, No.5, pp.1109-1159, 2005.
- [6] T. D. Le, T. Komeda, and M. Takagi, "Knowledge-based recurrent neural networks in reinforcement learning," *Proceedings of the 11th IASTED International Conference on Artificial Intelligence and Soft Computing*, pp.169-174, 2007.
- [7] L. Prechelt, "PROBEN1-A Set of Neural Network Benchmark Problems and Benchmarking Rules," *Technical Report 21/94*, Department of Computer Science, University of Karlsruhe, 1999.
- [8] Stuttgart Neural Network Simulator (SNNS), *User Manual, Version 4.0*, Institute for Parallel and Distributed High Performance Systems (IPVR), University of Stuttgart, 1998.
- [9] A. Zell.: *Simulation neuronaler Netze*, Addison-Wesley, 1994.

[1] F. Dandurand, V. Berthiaume, and T. R. Shultz, "A systematic comparison of flat and standard

저 자 소 개

이 상 화 (Sang-Wha Lee)

정회원



▪ 1985년 10월 ~ 1998년 11월 :
도르트문트대학교(독일) 컴퓨터
공학과 (공학사, 공학석사, 공학
박사)

▪ 2000년 2월 ~ 2002년 8월 : 한
국전자통신연구원(ETRI) 선임
연구원

▪ 2002년 9월 ~ 현재 : 서원대학교 정보통신공학과 부
교수

▪ 2008년 7월 ~ 2009년 6월 : 한국전자통신연구원
(ETRI) 초빙연구원

<관심분야> : 지능형 시스템, RFID/USN, 정보통신