

함수 근사화를 위한 방사 기저함수 네트워크의 전역 최적화 기법

이 종 석[†] · 박 철 훈^{**}

요 약

본 논문에서는 방사 기저함수 네트워크의 파라미터를 전 영역에서 최적화하는 학습 알고리즘을 제안한다. 기존의 학습 알고리즘들은 지역 최적화만을 수행하기 때문에 성능의 한계가 있고 최종 결과가 초기 네트워크 파라미터 값에 크게 의존하는 단점이 있다. 본 논문에서 제안하는 하이브리드 모의 담금질 기법은 모의 담금질 기법의 전 영역 탐색 능력과 경사 기반 학습 알고리즘의 지역 최적화 능력을 조합하여 전 파라미터 영역에서 해를 찾을 수 있도록 한다. 제안하는 기법을 함수 근사화 문제에 적용하여 기존의 학습 알고리즘에 비해 더 좋은 학습 및 일반화 성능을 보이는 네트워크 파라미터를 찾을 수 있으며, 초기 파라미터 값의 영향을 크게 줄일 수 있음을 보인다.

키워드 : 하이브리드 모의 담금질, 방사 기저함수 네트워크, 전역 최적화, 학습, 일반화 성능

A Global Optimization Method of Radial Basis Function Networks for Function Approximation

Jong-Seok Lee[†] · Cheol Hoon Park^{**}

ABSTRACT

This paper proposes a training algorithm for global optimization of the parameters of radial basis function networks. Since conventional training algorithms usually perform only local optimization, the performance of the network is limited and the final network significantly depends on the initial network parameters. The proposed hybrid simulated annealing algorithm performs global optimization of the network parameters by combining global search capability of simulated annealing and local optimization capability of gradient-based algorithms. Via experiments for function approximation problems, we demonstrate that the proposed algorithm can find networks showing better training and test performance and reduce effects of the initial network parameters on the final results.

Key Words : Hybrid Simulated Annealing, Radial Basis Function Network, Global Optimization, Training, Generalization Performance

1. 서 론

방사 기저함수 네트워크(RBFN: radial basis function network)는 인공 신경회로망의 한 종류로써 함수 근사화나 패턴 인식과 같은 문제에 성공적으로 적용되어 왔다. RBFN은 국소적 반응을 보이는 생물학적인 뉴런에 기초한 것으로서, 작은 주파수 대역에 선택적인 청각 시스템의 세포나 국소적인 특정 영상 특징에 반응하는 시각 피질의 세포와 같은 예에서 찾아볼 수 있다[1,2].

RBFN은 (그림 1)에서 보인 것과 같이 은닉층과 출력층으로 이루어진 전방향 구조를 가지며, 은닉층의 뉴런으로 가장 많이 쓰이는 것은 가우시안 함수이다. 주어진 입력 \mathbf{x} 에 대해 한 개의 출력 노드를 가지는 RBFN의 출력은

$$f(\mathbf{x}) = \sum_{i=1}^N w_i \exp\left(-\frac{\|\mathbf{x} - \mathbf{c}_i\|^2}{2\sigma_i^2}\right) + b \quad (1)$$

로 주어진다. 여기서 N 은 은닉 뉴런의 개수이고 \mathbf{c}_i , σ_i 및 w_i 는 각각 i 번째 은닉 뉴런의 중심, 폭 및 출력 가중치이며, b 는 출력 바이어스이다.

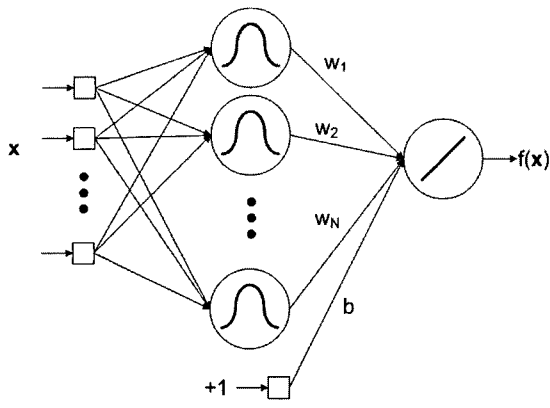
주어진 문제에 RBFN을 사용하기 위해서는 사전에 학습 데이터를 가지고 네트워크의 파라미터를 학습해야 한다. 전

※ 본 연구는 2007년도 한국과학기술원 BK21 정보기술사업단에 의하여 지원되었음.

† 정 회 원 : 한국과학기술원 전자전산학부 연수연구원

** 정 회 원 : 한국과학기술원 전자전산학부 교수

논문접수 : 2007년 2월 8일, 심사완료 : 2007년 7월 19일



(그림 1) RBFN의 구조

형적인 학습 과정으로써 은닉층과 출력층을 따로 학습하는 ‘2단계 학습’을 들 수 있다[3]. 이 학습 방법은 RBFN의 구조적 특성을 고려한 학습 규칙으로써, 가우시안 함수의 중심은 학습 데이터 중에서 무작위로 선택하여 정하거나 k-평균 군집(k-means clustering)과 같은 방식으로 정해지고, 가우시안 함수의 폭은 중심과 주변 데이터들 사이의 거리의 평균으로 얻어진다. 그리고 출력층의 가중치와 바이어스는 은닉층의 출력과 출력의 목표값을 가지고 유사역원(pseudo-inverse)에 의해 유일한 값을 얻을 수 있다. ‘3단계 학습’은 2단계 학습을 통해 얻은 네트워크를 경사(gradient)에 기반한 학습 알고리즘으로 추가 학습을 하는 것으로써, 이러한 추가적 학습을 통해 성능의 향상을 얻을 수 있다[3].

그러나, 이러한 학습 방식들은 네트워크의 파라미터 공간에서 지역 최적화만을 수행하는 기법들이다. 따라서, 파라미터의 초기값에 따라 지역 최적해에 머무르고 마는 경향이 있다. 전 영역에서의 최적해를 찾지 못함으로 인해 성능에 제한이 있게 되고, 최종 네트워크가 초기 파라미터값에 크게 의존하게 되는 단점이 있다.

본 논문에서는 이러한 단점들을 해결하기 위해 하이브리드 모의 담금질(HSA: hybrid simulated annealing) 기법을 이용하여 전 영역에서의 최적화를 수행할 수 있는 알고리즘을 제안한다. HSA는 원래 은닉 마르코프 모델(HMM: hidden Markov model)의 학습에 사용하기 위해 개발되었으며 [4], 본 논문에서는 이를 RBFN의 최적화에 적합하도록 설계한다. 제안하는 알고리즘은 전 영역에서 확률적인 최적화를 수행하는 모의 담금질(SA: simulated annealing) 기법과, 지역 최적화를 수행하여 현재보다 항상 더 나은 해를 찾을 수 있는 경사기반 알고리즘을 결합한 것이다. 이를 통해 알고리즘의 속도와 해의 질을 향상시키면서 전역 최적화를 수행할 수 있다. 결과적으로 RBFN의 일반화 성능을 향상시키고 초기 네트워크 파라미터의 영향을 최소화하여 알고리즘의 신뢰도를 높이도록 한다. 또한, 본 논문에서는 알고리즘이 좋은 성능을 낼 수 있도록 알고리즘의 파라미터를 정하는 기법을 보인다.

다음 장에서는 RBFN 학습과 관련한 기존의 연구를 소개한다. 3장에서는 제안하는 알고리즘의 자세한 과정을 기술

한다. 4장에서는 제안하는 알고리즘을 함수 근사화 문제에 적용한 결과를 보이고 기존의 방법에 비해 향상된 성능을 얻음을 보인다. 마지막으로 5장에서 결론으로 논문을 맺는다.

2. 기존의 학습 기법

RBFN을 학습하는 가장 쉬운 방법은 다층 퍼셉트론(multi-layer perceptron)의 학습에서와 같이 오차 역전파(EBP: error back-propagation) 방법과 같은 경사기반 학습 방법으로 모든 파라미터들을 한꺼번에 학습하는 것이다. Wettschereck과 Dietterich은 이러한 방법이 성공적으로 적용될 수 있음을 보고하였다[5].

이후 RBFN의 구조적 특징을 활용하는 학습 방법이 개발되었다. Broomhead와 Lowe는 가우시안 함수의 중심을 학습 데이터에서 무작위로 선택하여 정하는 방법을 개발하였다[6]. 이 방법은 어떤 데이터를 선택하느냐에 따라 그 결과가 민감하게 변한다는 단점이 있다. 이후 Moody와 Darken에 의해 제안된 방법은 k-평균 군집과 같은 군집화 기법을 사용하여 데이터를 몇 개의 군집으로 묶고 각 군집의 중심을 가우시안 함수의 중심으로 사용하는 기법이다[7]. 이 방법은 RBFN을 학습하는 방법으로 가장 널리 사용되는 기법 중 하나이다.

가우시안 함수의 너비는 그 중심으로부터 가까운 몇 개의 학습 데이터까지의 평균 거리로 정하는 기법이 제안되어 많이 사용되고 있다[7].

가우시안 함수의 중심과 너비를 결정한 후 출력층의 가중치와 바이어스는 유사역원(또는 최소자승)에 의해 구할 수 있으며 이렇게 구한 해는 정해진 가우시안 함수의 출력에 대해 유일하게 얻어진다.

이상과 같은 학습 기법은 은닉층과 출력층을 따로 학습하기 때문에 ‘2단계 학습’ 기법이라 한다. Schwenker 등은 이러한 2단계 학습 이후에 EBP 기법으로 모든 파라미터를 한꺼번에 추가적으로 학습하는 과정을 거치는 ‘3단계 기법’의 성능을 실험하였다[3]. 그 결과 성능이 크게 향상되는 것을 보고하였다.

머리말에서 언급한 것과 같이 2단계 학습이나 3단계 학습 모두 파라미터의 지역 최적해만을 찾으려 초기 파라미터 값에 따라 전역 최적해를 찾지 못할 가능성이 있다. 이러한 제약은 최종 네트워크의 성능에 영향을 미쳐서 일반화 성능에 제한이 생기고, 무작위값에 따라 그 결과가 크게 차이는 단점이 있다. 이를 보완하기 위한 전역 최적화 알고리즘을 다음 장에서 제안한다.

3. 제안하는 알고리즘

SA는 해공간의 전 영역에서 탐색이 가능한 확률적 탐색 알고리즘으로써, Kirkpatrick에 의해 제안되었다[8]. 금속이 액체 상태에서 담금질 과정에 의해 냉각될 때 최종적으로

최소 에너지를 갖는 결정 형태로 안정화되는 과정을 공학적으로 응용한 기법이다. SA는 현재의 해에서 새로운 해를 확률적으로 생성하여 평가하고 이 해를 받아들일 것인지를 확률적으로 결정하는 과정을 반복하여 전역 최적화를 수행한다. SA에서 '온도' 파라미터는 새로운 해를 받아들이는 것을 조절하는 역할을 한다. 온도가 높을 때는 새로운 해의 비용 함수(또는 목적함수)의 값이 현재보다 나쁜 경우에도 받아들일 확률이 높다. 알고리즘을 반복하는 동안 온도는 냉각 과정에 의해 점차 낮아진다. 온도가 매우 낮아지면 현재보다 나쁜 해가 받아들여질 확률은 극히 작아진다. SA의 중요한 특징은 현재의 해보다 비용함수값이 나쁜 해로 옮겨가는 것이 가능하여 지역 최적해를 벗어나 전 영역에서의 탐색이 가능하다는 것이다. 이러한 특징으로 인해 복잡한 문제들을 포함한 다양한 문제들에 성공적으로 적용되어 왔다[9].

SA의 전 영역 탐색 능력과 경사기반 학습에 의한 지역 최적화 능력을 조합하면 알고리즘의 속도를 향상시키면서 좋은 해를 찾는 기법을 얻을 수 있다. 기존의 연구에서는 HMM의 학습을 위해 SA와 기대-최대(expectation-maximization) 알고리즘을 결합하여 기존의 HMM 학습 기법들보다 향상된 결과를 얻을 수 있음을 보였다[4].

HSA 알고리즘은 기본적으로 빠른 모의 담금질(FSA: fast simulated annealing)[10]을 사용한다. FSA는 새로운 해를 생성하기 위해 코시(Cauchy) 확률분포를 사용하며 담금질에는 반복횟수의 역수에 비례하는 담금질 과정을 사용한다. HSA에서 확률적으로 탐색하는 해는 가우시안 함수의 파라미터인 c_i 와 σ_i 이며, 출력층의 w_i 와 b 는 유사역원에 의해 자동적으로 정해진다. 알고리즘의 성능에 중요한 영향을 미치는 초기온도와 비용함수의 계수를 자동적으로 정하기 위한 기법을 포함한다. 전체 알고리즘의 자세한 과정은 다음과 같다.

1단계) 데이터 정규화 및 해의 범위 정의: 학습데이터의 입력값을 [-1, 1] 범위가 되도록 정규화한다. 그리고 가우시안 함수의 중심과 너비 파라미터의 상한 및 하한값을 정의한다. 중심의 상·하한은 [-1.1, +1.1]로 하여 데이터의 범위 주변으로 10%의 여유를 허용한다. 너비의 상·하한은 [0.1, 1]로 한다.

2단계) 초기화: 2단계 학습 기법에 의해 RBFN을 초기화한다. 초기해 p_0 는 n 차원 벡터로써, RBFN의 중심 및 너비 파라미터를 [0, 1] 사이의 범위로 정규화한 것으로 구성된다. 여기서, 정규화는 1단계에서 정의한 상·하한값을 이용하며, 해 공간이 n 차원 공간에서 각 성분별로 [0, 1]까지를 범위로 하는 초입방체(hypercube)가 되도록 한다. 초기온도 T_0 는 알고리즘의 초기에 해공간의 전 영역을 탐색하기에 충분하도록 정해야 한다. 온도 파라미터는 3단계의 식 (4)에 의한 해의 생성 과정에서 해

의 변화량에 영향을 미친다. 즉, 온도가 높으면 큰 변화량이 생성될 확률이 높고 반대로 온도가 낮으면 변화량이 클 확률은 낮아진다. 따라서 초기온도는 알고리즘의 초반에 생성되는 해의 변화량이 클 확률이 높아서 현재의 해로부터 멀리 떨어진 해의 생성이 쉽고 결과적으로 해공간 전체에서 탐색을 수행하는 것이 수월하도록 충분히 높게 설정한다. 하나의 해와 그로부터 생성되는 해의 거리가 가장 먼 경우는 그 두 해가 해 공간 $[0,1]^n$ 의 대각선 양 끝에 위치하는 경우이다. 따라서 초기온도를 이 대각선의 길이에 비례하도록 하며, 이때 비례상수는 1로 하는 것이 적절함을 실험으로 확인하였다. 즉,

$$T_0 = \sqrt{n} \tag{2}$$

3단계) 새로운 해의 생성: 현재의 해 p_k 로부터 새로운 해 q_k 를 생성한다.

$$q_k = p_k + \Delta p_k \tag{3}$$

해의 변화량 Δp_k 는 다음과 같은 n 차원 코시(Cauchy) 생성함수가 사용된다.

$$g(\Delta p_k) = \frac{a_n T_k}{(\|\Delta p_k\|^2 + T_k^2)^{(n+1)/2}} \tag{4}$$

여기서 T_k 는 k 번째 반복에서의 온도, a_n 은 정규화를 위한 상수이다. 즉, 위 식과 같은 확률분포함수를 따르는 n 차원 무작위 벡터를 생성함으로써 Δp_k 를 얻는다[11]. 출력층의 가중치와 바이어스는 유사역원으로 결정된다.

4단계) 지역 최적화: 몇 번의 레벤버그-마퀴트(LM: Levenberg-Marquardt) 알고리즘[12]을 적용하여 전체 네트워크를 학습한다. LM 알고리즘은 신경 회로망의 경사기반 학습 알고리즘의 하나로써 수렴 속도가 빠르고 좋은 성능을 보이는 학습 기법으로 알려져 있다. 이 단계를 통해 알고리즘이 빨리 수렴하고 해의 질이 향상되도록 한다.

5단계) 해의 평가: 새로운 해의 비용함수값을 다음과 같이 계산한다.

$$C = M \log(\text{RMSE}) \tag{5}$$

여기서 RMSE는 평균제곱근 오차(root-mean-squared error)를 의미한다. M 은 비용함수값의 크기를 조절하기 위한 파라미터로써 2단계에서 결정하며, 100번의 해를 시도하

여 그들 중 가장 좋은 것으로부터 가장 나쁜 것이 선택될 확률(식 (6))이 0.5가 되도록 결정한다. 위 식에서 RMSE에 로그를 취하는 것은 알고리즘 수행 과정에서 RMSE의 변화가 로그 스케일로 나타나는 것을 관찰했기 때문이다.

6단계) 해의 선택: 새로운 해를 받아들일지 메트로폴리스(Metropolis) 규칙[13]에 의해 결정한다. 즉, 새로운 해를 받아들일 확률은

$$p_a = \min[1, \exp(-\Delta C/T_k)] \quad (6)$$

로 주어진다. 여기서 ΔC 는 현재 해와 새로운 해의 비용함수값의 차이이다. 즉, 새로운 해가 현재의 것보다 더 좋으면 항상 받아들이고, 그렇지 않으면 비용함수값의 차이에 의해 확률적으로 받아들인다.

7단계) 담금질: 다음의 식에 의해 온도를 낮춘다.

$$T_k = T_0/k \quad (7)$$

8단계) 종료조건 판단: 알고리즘이 수렴단계에 있다고 판단되면 종료하고, 그렇지 않으면 3단계로 돌아간다. 일정 반복횟수(s) 동안 비용함수값의 상대적인 변화량이 정해진 작은 값 ϵ 보다 작으면, 즉

$$\frac{C_{k-s} - C_k}{C_{k-s}} < \epsilon \quad (8)$$

이면 수렴단계로 판단한다. 여기서 C_k 는 k 번째 반복에서의 비용함수값이다.

4. 실험 및 결과

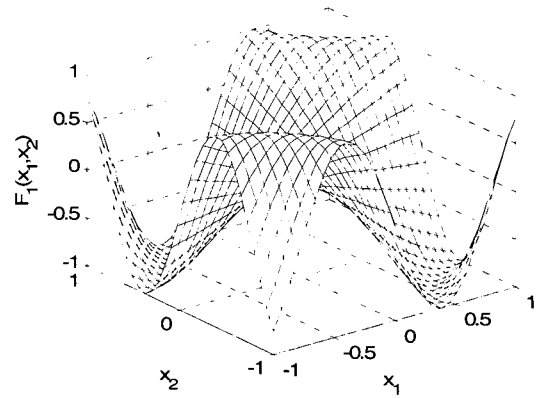
본 장에서는 함수 근사화 문제를 통해 제안하는 알고리즘의 성능을 알아본다. 신경회로망의 성능 비교를 위해 많이 사용되는 다음의 문제들을 사용한다[14-17].

$$F_1(x_1, x_2) = \sin(4x_1x_2) \quad (9)$$

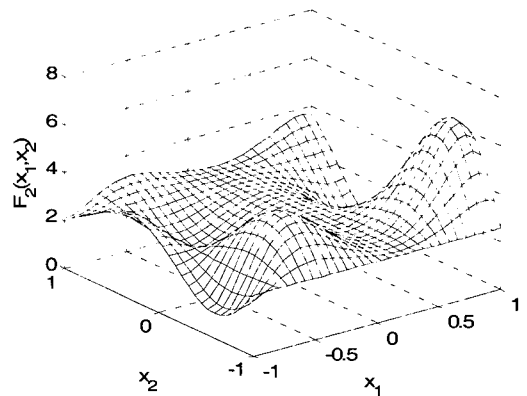
$$F_2(x_1, x_2) = 1.9(1.35 + e^{0.5(x_1+1)}) \sin(13(0.5x_1 - 0.1)^2) \epsilon^{-3.5(x_1+1)} \sin(3.5x_2 + 3.5) \quad (10)$$

$$F_3(x_1, x_2) = \sin(3x_1 + \sin(3x_2)) + 0.25[3(1 - 3x_1)^2 e^{-9x_1^2 - (3x_2+1)^2} - 10(1.5x_1 - 27x_1^3 - 81x_2^3) e^{-9x_1^2 - 9x_2^2} - e^{-(3x_1+1)^2 - 9x_2^2}/3] \quad (11)$$

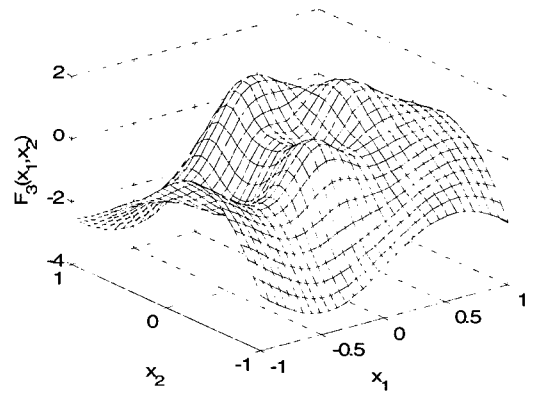
(그림 2)에 위의 함수들을 그림으로 나타내었다.



(a)



(b)

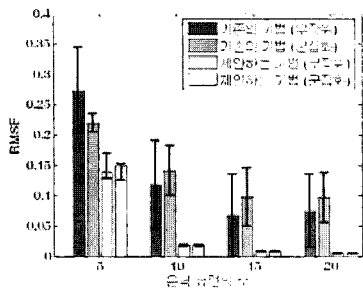


(c)

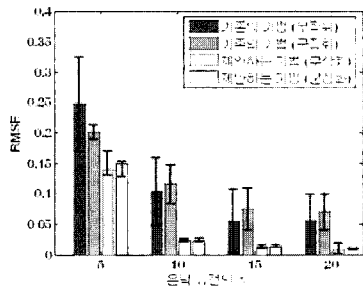
(그림 2) 성능 측정을 위한 함수 근사화 문제 (a) F_1 (b) F_2 (c) F_3

학습 데이터는 $[-1, 1]^2$ 영역에서 무작위로 선택한 225개와 네 모서리를 합한 229개의 데이터를 사용한다. 일반화 성능 분석을 위한 테스트 데이터로는 10,000개의 점을 균일하게 선택하였다. 실험은 서로 다른 무작위 값을 사용하여 20번 반복하였다.

제안하는 HSA 알고리즘을 기존의 3단계 학습 기법과 비교하였다. 기존의 3단계 학습 기법이 전체 네트워크의 학

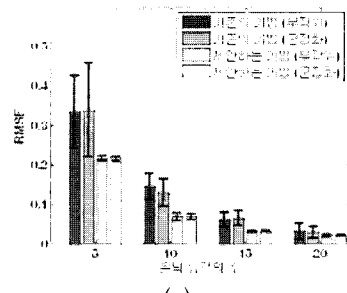


(a)

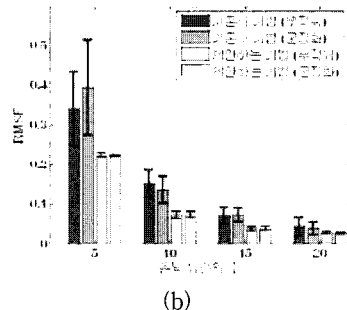


(b)

(그림 3) F_1 에 대한 기존의 알고리즘과 제안하는 알고리즘의 최종 RMSE의 평균 및 표준편차 (a) 학습데이터 (b) 테스트 데이터

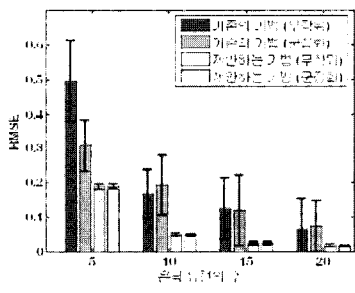


(a)

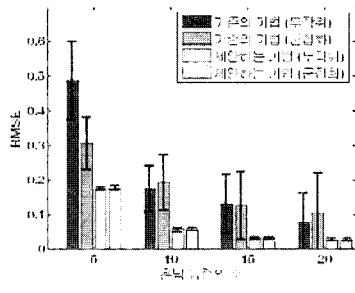


(b)

(그림 5) F_3 에 대한 기존의 알고리즘과 제안하는 알고리즘의 최종 RMSE의 평균 및 표준편차 (a) 학습데이터 (b) 테스트 데이터



(a)



(b)

(그림 4) F_2 에 대한 기존의 알고리즘과 제안하는 알고리즘의 최종 RMSE의 평균 및 표준편차 (a) 학습데이터 (b) 테스트 데이터

습을 위해 EBP 알고리즘을 사용하는 것과는 달리, 본 논문에서는 LM 알고리즘을 사용하였다. LM 알고리즘은 EBP 기법보다 더 좋은 성능을 보이는 것으로 알려져 있다. 따라서 이하의 결과에서 보이는 3단계 학습은 Schwenker[3]의 것보다 더 좋은 성능을 보이는 것이라 할 수 있다. 두 기법

에서 가우시안 함수의 중심의 초기값은 학습 데이터 중 무작위로 선택하여 결정하거나 k-평균 군집 알고리즘으로 결정하였으며, 너비의 초기값은 중심으로부터 가까운 5개의 데이터까지의 거리의 평균으로 결정하였다. 식 (8)의 종료조건을 사용하였으며, 3단계 학습 기법에 대해서는 $s = 100$ 을, HSA에 대해서는 $s = 500$ 을 사용하였다. 두 기법 모두 $\epsilon = 0.001$ 로 하였다. HSA에서 지역 최적화에는 30번의 LM 알고리즘을 사용하였다.

(그림 3), (그림 4) 및 (그림 5)는 각 문제에 대해 RBFN의 은닉 뉴런의 개수를 5에서 20까지 변화시켜가며 학습했을 때 학습 데이터 및 테스트 데이터에 대한 RMSE값을 나타낸다. 20번의 반복 실험에 대한 평균과 표준편차를 보여준다. 제안하는 기법이 모든 경우에 대해 학습 및 테스트 데이터에 대한 성능이 더 좋은 것을 볼 수 있다. 이러한 결과를 통해 전역 최적화에 의한 성능 향상을 확인할 수 있다. 또한, 제안하는 기법에서는 두 가지 초기화에 따른 결과의 차이가 거의 없으며 반복실험에 대한 표준편차 역시 크게 작아진 것을 볼 수 있다. 이러한 결과 역시 전 탐색 영역에 대한 최적화가 성공적으로 이루어졌음을 뒷받침한다.

제안하는 알고리즘이 이러한 우수성을 보이는 대신, 학습 과정에 필요한 시간을 더 많이 소요한다. <표 1>은 은닉 뉴런을 20개로 하였을 때 각 알고리즘이 학습에 소모하는 시간을 나타낸다. 제안하는 알고리즘은 평균적으로 약 13배의 시간을 더 필요로 한다. 그러나 앞에서 살펴본 것과 같은 성능의 향상을 고려할 때 이러한 추가적인 계산시간은 충분히 가치가 있다고 할 수 있다. 또한, RBFN을 문제에 적용하여 사용하는 과정은 두 알고리즘이 같으며 학습은 그 이전에 완료되어 있기 때문에 실제 적용 과정에서 HSA의 학

<표 1> 은닉 뉴런이 20개일 때 각 학습 알고리즘에 소요되는 시간(초)

문제	초기화	기존의 기법	제안하는 기법
F_1	무작위	33	361
	군집화	19	415
F_2	무작위	27	365
	군집화	20	351
F_3	무작위	30	328
	군집화	34	341
평균		27	360

습 시간은 문제가 되지 않는다.

5. 결 론

본 논문에서는 RBFN의 전역 최적화를 위한 HSA 알고리즘을 제안하였다. 전역 최적화를 수행할 수 있는 확률적 탐색 알고리즘인 SA와 지역최적화를 수행하는 경사기반 학습 알고리즘을 결합하여 전 영역에서 학습을 수행할 수 있는 알고리즘을 개발하였다. 최적의 성능을 위해 알고리즘의 파라미터를 정하는 기법들을 포함하였다. 함수 근사화 문제에 적용한 결과 기존의 3단계 학습 알고리즘에 비해 초기 파라미터의 영향을 줄이면서 더 좋은 성능의 RBFN을 생성함을 확인하였다.

추후 과제으로써 제안한 알고리즘을 패턴인식이나 예측 등의 다양한 문제에 적용하는 것이 필요하다.

참 고 문 헌

[1] L. Rabiner and B.-H. Juang, 'Fundamentals of Speech Recognition,' Englewood Cliffs, Prentice Hall, 1993.
 [2] T. Poggio and F. Girosi, "Regularization algorithms for learning that are equivalent to multilayer networks," Science, Vol. 2247, pp. 978-982, 1990.
 [3] F. Schwenker, H. A. Kestler, and G. Palm, "Three learning phases for radial-basis-function networks," Neural Networks, Vol. 14, No. 4-5, pp. 439-458, 2001.
 [4] J.-S. Lee and C. H. Park, "Training hidden Markov models by hybrid simulated annealing for visual speech recognition," in Proc. IEEE Int. Conf. Systems, Man, and Cybernetics, pp. 198-202, Taipei, Taiwan, Oct. 2006.
 [5] D. Wettschereck and T. Dietterich, "Improving the performance of radial basis function networks by learning center locations," in J. E. Moody, S. J. Hanson, and R. P. Lippman, eds., Advances in Neural Information Processing Systems, Vol. 4, Morgan Kaufmann, Los Altos, CA, pp. 1133-1140, 1992.
 [6] D. S. Broomhead and D. Lowe, "Multivariable functional interpolation and adaptive networks," Complex Systems, Vol. 2, pp. 321-355, 1988.
 [7] J. Moody and C. Darken, "Fast learning in networks of locally-tuned processing units," Neural Computation, Vol. 1, No. 2, pp. 291-294, 1989.

[8] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, "Optimization by simulated annealing," Science, Vol. 220, pp. 671-680, May 1983.
 [9] P. J. M. V. Laarhoven and E. H. L. Aarts, 'Simulated Annealing: Theory and Applications,' Kluwer Academic Publishers, Dordrecht, The Netherlands, 1987.
 [10] H. Szu and R. Hartley, "Fast simulated annealing," Phys. Lett. A, Vol. 122, No. 3-4, pp. 157-162, Jun. 1987.
 [11] D. Nam, J.-S. Lee, and C. H. Park, "n-dimensional Cauchy neighbor generation for the fast simulated annealing," IEICE Trans. Information and Systems, Vol. E87-D, No. 11, pp. 2499-2502, 2004.
 [12] C. Bishop, 'Neural Networks for Pattern Recognition,' Oxford Univ. Press Inc., New York, 1995.
 [13] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, "Equation of state calculations by fast computing machines," J. Chem. Phys., Vol. 21, No. 6, pp. 1087-1092, 1953.
 [14] L. Breiman, "The II method for estimating multivariate functions from noisy data," Technometrics, Vol. 3, No. 2, pp. 125-160, 1991.
 [15] J.-N. Hwang, S.-R. Lay, M. R. D. Martin, and J. Schimert, "Regression modeling in back-propagation and project pursuit learning," IEEE Trans. Neural Networks, Vol. 5, No. 3, pp. 342-353, 1994.
 [16] J.-S. Lee, H. Lee, J.-Y. Kim, D. Nam, and C. H. Park, "Self-organizing neural networks by construction and pruning," IEICE Trans. Information and Systems, Vol. E87-D, No. 11, pp. 2489-2498, Nov. 2004.
 [17] 이종석, 김재영, 정승범, 박철훈, "다층 신경회로망을 위한 자기 구성 알고리즘," 전자공학회논문지, 제41권 CI편 제3호, pp. 55-66, 2004년 5월.



이 종 석

e-mail : jslee@nnmi.kaist.ac.kr
 1999년 한국과학기술원 전기및전자공학과 학사
 2001년 한국과학기술원 전자전산학과 공학석사
 2006년 한국과학기술원 전자전산학과 공학박사

2006년~현재 한국과학기술원 전자전산학부 연수연구원
 관심분야: 시청각 음성인식, 멀티모달 인터페이스



박 철 훈

e-mail : chpark@kaist.ac.kr
 1984년 서울대학교 전자공학과 학사
 1985년 Caltech 전자공학과 공학석사
 1990년 Caltech 전자공학과 공학박사
 1991년~현재 한국과학기술원 전자전산 학부 교수

관심분야: 지능시스템, 신경회로망, 최적화, 지능제어