



(19) 대한민국특허청(KR)  
(12) 등록특허공보(B1)

(45) 공고일자 2022년05월19일  
(11) 등록번호 10-2400610  
(24) 등록일자 2022년05월17일

(51) 국제특허분류(Int. Cl.)  
G06N 3/08 (2006.01) G06N 3/04 (2006.01)  
(52) CPC특허분류  
G06N 3/084 (2013.01)  
G06F 17/18 (2013.01)  
(21) 출원번호 10-2020-0116991  
(22) 출원일자 2020년09월11일  
심사청구일자 2020년09월11일  
(65) 공개번호 10-2022-0034517  
(43) 공개일자 2022년03월18일  
(56) 선행기술조사문헌  
Saeid Iranmanesh et al., A Differential Adaptive Learning Rate Method for Back-Propagation Neural Networks, World Academy of Science, Engineering and Technology 50, 285-288pages (2009. 2.)\*  
Matthew D. Zeiler, ADADELTA: AN ADAPTIVE LEARNING RATE METHOD, arXiv:1212.5701v1, 1-6pages (2012. 12. 22.)\*  
\*는 심사관에 의하여 인용된 문헌

(73) 특허권자  
포항공과대학교 산학협력단  
경상북도 포항시 남구 청암로 77 (지곡동)  
(72) 발명자  
박부견  
경상북도 포항시 남구 지곡로 155, 9동 1403호(지곡동, 교수아파트)  
김동우  
대전광역시 서구 정림로 55, 112동 802호(정림동, 우성아파트)  
(74) 대리인  
이철희

전체 청구항 수 : 총 5 항

심사관 : 양대경

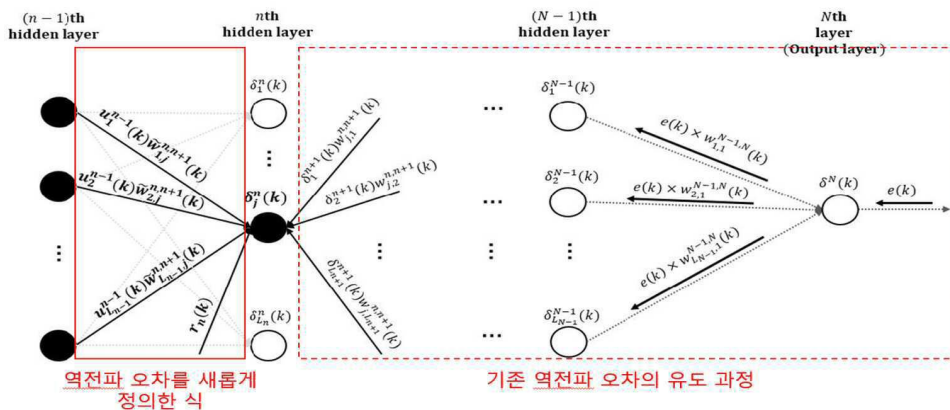
(54) 발명의 명칭 평균제곱편차 분석기반의 적응형 학습률을 갖는 역전파 신경망 학습방법

(57) 요약

본 발명은 평균 제곱 편차 분석 기반의 적응형 학습률을 갖는 역전파 신경망 학습방법에 관한 것으로, 각 층 내부 신경망의 역전파 에러 신호가 해당 층의 가중치 오차에 의해 발생한다는 새로운 개념을 도입하여 가중치 오차를 최소화하는 방향으로 실시간 에러에 따라 학습률을 적응적으로 업데이트하기 위한 것이다.

이를 위하여 본 발명은 신경망 알고리즘에서 필수적으로 유도되는 n번째 층의 j번째 뉴런에 존재하며, n-1번째와 n번째 층 사이에 연결되어 있는 가중치의 오차의 영향으로 발생하는 역전파 오차를 산출하는 제1과정, 현재 시간의 n-1번째 층과 n번째 층 사이에 존재하는 가중치의 평균제곱편차를 분석하여 이 분석결과를 바탕으로 가중치의 평균제곱편차를 최소화할 수 있는 현재 시간의 적응형 학습률을 업데이트하는 제2과정을 포함하여, 충격 소음이나 잘못된 학습 데이터에 대해 강인한 역전파 신경망을 구축할 수 있게 한다.

대표도 - 도1



(52) CPC특허분류  
**G06N 3/04** (2013.01)

**이민호**

대구광역시 서구 달구벌대로377길 5-7(내당동)

(72) 발명자

**박태수**

경상북도 포항시 남구 청암로 77, 포항공과대학교  
 대학원 아파트 4동 901호(지곡동)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711113185
과제번호	2019R1A4A1029003
부처명	과학기술정보통신부
과제관리(전문)기관명	한국연구재단
연구사업명	집단연구지원(R&D)
연구과제명	강인성과 보안성을 지닌 다개체 시스템에서의 편대 제어
기여율	1/2
과제수행기관명	포항공과대학교
연구기간	2020.03.01 ~ 2021.02.28

이 발명을 지원한 국가연구개발사업

과제고유번호	1711109130
과제번호	2020R1A2C2005709
부처명	과학기술정보통신부
과제관리(전문)기관명	한국연구재단
연구사업명	개인기초연구(R&D)
연구과제명	가상 물리 시스템을 위한 확장된 하이브리드 시스템의 제어 이론 연구
기여율	1/2
과제수행기관명	포항공과대학교
연구기간	2020.03.01 ~ 2021.02.28

공시예외적용 : 있음

---

## 명세서

### 청구범위

#### 청구항 1

- (a) 신경망 알고리즘에서 필수적으로 유도되는 n번째 층의 j번째 뉴런에 존재하며, n-1번째 층과 n번째 층 사이에 연결되어 있는 가중치의 오차의 영향으로 발생하는 역전파 오차를 산출하는 제1과정;
- (b) 현재 시간의 n-1번째 층과 n번째 층 사이에 존재하는 가중치의 평균제곱편차를 분석하여, 이 분석결과를 바탕으로 가중치의 평균제곱편차를 최소화할 수 있는 현재 시간의 적응형 학습률을 업데이트하는 제2과정;을 포함하되,

상기 역전파 오차( $\delta_j^n(k)$ )는,

n-1번째 층의 각 뉴런에서 발생하는 최종값으로 수렴된 입력값( $\mathbf{u}_t^{n-1,T}$ ), n-1번째와 n번째 사이에 연결된 가중치가 최종값으로 수렴되었을 때의 값( $\mathbf{W}_{j,t}^{n-1,n}$ ), n-1번째 층의 각 뉴런에서 발생하는 현재 시간 k에서의 입력값( $\mathbf{u}^{n-1}(k)^T$ ), 각 뉴런에서 발생할 수 있는 외란( $\eta_n(k)$ ), 및 일종의 외란의 일부로서  $u_t$ 를  $u(k)$ 로 변경할 때 발생하는 차이의 값( $\gamma_n(k)$ )으로 정의되는 파라미터들에 의하여,

아래의 수학적 식 1

[수학적 식 1]

$$\begin{aligned} \delta_j^n(k) &= \mathbf{u}_t^{n-1,T} \mathbf{W}_{j,t}^{n-1,n} - \mathbf{u}^{n-1}(k)^T \mathbf{W}_j^{n-1,n}(k) + \eta_n(k), \\ &= \mathbf{u}^{n-1}(k)^T \mathbf{W}_{j,t}^{n-1,n} - \mathbf{u}^{n-1}(k)^T \mathbf{W}_j^{n-1,n}(k) + r_n(k), \\ &= \mathbf{u}^{n-1}(k)^T \tilde{\mathbf{W}}_j^{n-1,n}(k) + r_n(k), \end{aligned}$$

로 산출되는 것을 특징으로 하는 평균 제곱 편차 분석 기반의 적응형 학습률을 이용한 역전파 신경망의 학습방법.

#### 청구항 2

제1항에 있어서, 상기 제2과정(b)은,

- (b1) n-1번째 층과 n번째 층 사이의 가중치를 업데이트하기 위한 현재시간의 적응형 학습률을 산출하는 제1단계;
- (b2) 상기 제1단계(b1)에서 산출된 현재 시간의 적응형 학습률과 현재 시간의 임의로 주어진 현재시간의 역전파 오차와 n-1번째 층의 입력값의 정규화값을 통해 n-1번째 층과 n번째 층 사이에 존재하는 현재 시간의 가중치를 이용하여 n-1번째 층과 n번째 층 사이에 존재하는 다음 시간의 가중치를 산출하는 제2단계;
- (b3) 상기 제1단계에서 산출된 현재 시간의 적응형 학습률을 이용하여 n-1번째 층과 n번째 층 사이에 존재하는 가중치를 위한 다음 시간의 평균제곱편차 분석치를 산출하는 제3단계; 및
- (b4) 상기 제1단계(b1) 내지 제3단계(b3)을 반복 수행하여 각 은닉층에 대한 가중치를 업데이트하는 제4단계;을 포함하여 이루어지는 것을 특징으로 하는 평균 제곱 편차 분석 기반의 적응형 학습률을 이용한 역전파 신경망의 학습방법.

#### 청구항 3

삭제

**청구항 4**

제2항에 있어서, 상기 제1단계(b1)에서 적응형 학습률( $\mu_j^n(k)$ )은,

1보다 큰 정수로 설정되는 임의의 값( $\alpha_{n-1}$ )과 n-번째 층의 뉴런개수( $L_{n-1}$ )와 최종 에러의 변화량( $\sigma_{rn}^2(k)$ ) 및 초기 평균제곱편차 분석치( $Tr\{P(1)\}$ )에 대하여,

아래의 수학적 식 2

[수학적 식 2]

$$\mu_j^n(k) = \frac{Tr\{\mathbf{P}_j^{n-1,n}(k)\}/(\alpha_{n-1}L_{n-1})}{Tr\{\mathbf{P}_j^{n-1,n}(k)\}/(\alpha_{n-1}L_{n-1}) + \sigma_{rn}^2(k)/(\|\mathbf{u}^{n-1}(k)\|^2 + \epsilon)}$$

에 의해 산출되는 것을 특징으로 하는 평균 제곱 편차 분석 기반의 적응형 학습률을 이용한 역전파 신경망의 학습방법.

**청구항 5**

제4항에 있어서, 상기 제2단계(b2)에서 다음 시간의 가중치( $\mathbf{w}_j^{n-1,n}(k+1)$ )는,

상기 제1단계(b1)에서 산출된 현재 시간(k)의 적응형 학습률( $\mu_j^n(k)$ )과 제1과정에서 설정된 임의로 주어진 현재 시간(k)의 역전파 오차( $\delta_j^n(k)$ )와 n-1번째 층의 입력값의 제1정규화값( $\frac{\mathbf{u}^{n-1}(k)}{\|\mathbf{u}^{n-1}(k)\|^2}$ )을 통해 n-1번째 층과 n번째 층 사이에 존재하는 현재 시간(k)의 가중치( $\mathbf{w}_j^{n-1,n}(k)$ )를 이용하여,

아래의 수학적 식 3

[수학적 식 3]

$$\mathbf{w}_j^{n-1,n}(k+1) = \mathbf{w}_j^{n-1,n}(k) + \mu_j^n(k)\delta_j^n(k)\frac{\mathbf{u}^{n-1}(k)}{\|\mathbf{u}^{n-1}(k)\|^2}$$

에 의해 산출되는 것을 특징으로 하는 평균 제곱 편차 분석 기반의 적응형 학습률을 이용한 역전파 신경망의 학습방법.

**청구항 6**

제2항에 있어서, 상기 제3단계(b3)에서 다음시간의 평균제곱편차 분석치( $Tr\{P_j^{n-1,n}(k+1)\}$ )는,

상기 제1단계(b1)에서 산출된 현재 시간(k)의 적응형 학습률( $\mu_j^n(k)$ )과, 1보다 큰 정수로 설정되는 임의의 값( $\alpha_{n-1}$ )과 n-번째 층의 뉴런개수( $L_{n-1}$ )와 최종 에러의 변화량( $\sigma_{rn}^2(k)$ )과 n-1번째 층의 입력값의 제2정규화값( $\|\mathbf{u}^{n-1}(k)\|^2$ ) 및 현재 시간의 평균제곱편차 분석치( $Tr\{P_j^{n-1,n}(k)\}$ )를 이용하여 아래의 수학적 식 4

[수학적 식 4]

$$Tr\{\mathbf{P}_j^{n-1,n}(k+1)\} = \left[1 - \frac{2\mu_j^n(k) - \mu_j^n(k)^2}{\alpha_{n-1}L_{n-1}}\right] Tr\{\mathbf{P}_j^{n-1,n}(k)\} + \frac{\{\mu_j^n(k)\}^2 \sigma_{rn}^2(k)}{\|\mathbf{u}^{n-1}(k)\|^2 + \epsilon}$$

에 의해 산출되는 것을 특징으로 하는 평균 제곱 편차 분석 기반의 적응형 학습률을 이용한 역전파 신경망의 학습방법.

**발명의 설명**

**기술 분야**

[0001] 본 발명은 평균제곱편차(Mean Square Deviation) 분석기반의 적응형 학습률을 갖는 역전파(backpropagation) 신경망 알고리즘에 관한 것으로서, 보다 상세하게는 인공지능(Artificial Intelligence) 알고리즘의 가장 기본적인 분야인 신경망 구조의 역전파알고리즘 분야에서 고정된 학습률(learning rate)값을 가지는 역전파 알고리즘이 아닌 실시간 에러에 따라 학습률이 변화하는 적응형 학습률 값을 갖는 역전파 신경망 학습방법에 관한 것이다.

**배경 기술**

[0002] 일반적으로 AI 알고리즘의 가장 기본적인 분야인 신경망 구조의 역전파(backpropagation) 알고리즘은 고정된 학습률(learning rate) 값을 가지는 역전파 알고리즘으로서, 기존 대부분의 적응형 역전파 신경망 알고리즘 분야가 에러 신호를 직접적으로 최소화하는 것에 기반하여 개발되어 있으므로 새로운 환경, 즉 충격 소음 발생시 또는 학습에 잘못된 데이터(bad data)가 섞일 경우 이러한 충격 소음이나 잘못된 학습 데이터에 의해서도 가중치를 업데이트하게 되는 문제점이 있었다.

**발명의 내용**

**해결하려는 과제**

[0003] 따라서 본 발명은 상기의 문제점을 해결하기 위한 것으로, 본 발명이 해결하고자 하는 기술적 과제는, 역전파 알고리즘에서 보편적으로 사용하는 각 층 내부 신경망의 역전파 에러 신호가 해당 층의 가중치 오차(weight deviation)에 의해 발생한다는 새로운 개념을 도입함으로써 이러한 가중치 오차를 최소화하는 방향으로 실시간 에러에 따라 학습률을 업데이트하는 평균제곱편차(MSD) 분석 기반의 적응형 학습률을 갖는 역전파 신경망 학습방법을 제공하고자 하는 것이다.

**과제의 해결 수단**

[0004] 상기 목적을 달성하기 위한 본 발명의 일 실시 형태는, 신경망 알고리즘에서 필수적으로 유도되는 n번째 층의 j 번째 뉴런에 존재하며, n-1번째 층과 n번째 층 사이에 연결되어 있는 가중치의 오차의 영향으로 발생하는 역전파 오차를 산출하는 제1과정, 현재 시간의 n-1번째 층과 n번째 층 사이에 존재하는 가중치의 평균제곱편차를 분석하여 이 분석결과를 바탕으로 가중치의 평균제곱편차를 최소화할 수 있는 현재 시간의 적응형 학습률을 업데이트하는 제2과정을 포함하는, 평균 제곱 편차 분석 기반의 적응형 학습률을 갖는 역전파 신경망 학습방법이다.

[0005] 상기 본 발명에 의한 역전파 신경망 학습방법에서 제2과정은, n-1번째 층과 n번째 층 사이의 가중치를 업데이트 하기 위한 현재 시간의 적응형 학습률을 산출하는 제1단계, 제1단계에서 산출된 현재 시간의 적응형 학습률과 현재 시간의 임의로 주어진 현재 시간의 역전파 오차와 n-1번째 층의 입력값의 정규화 값을 통해 n-1번째 층과 n번째 층 사이에 존재하는 현재 시간의 가중치를 이용하여 n-1번째 층과 n번째 층 사이에 존재하는 다음 시간의 가중치를 산출하는 제2단계, 제1단계에서 산출된 현재 시간의 적응형 학습률을 이용하여 n-1번째 층과 n번째 층 사이에 존재하는 가중치를 위한 다음시간의 평균제곱편차 분석치를 산출하는 제3단계, 및 제1단계 내지 제3단계를 반복 수행하여 각 은닉층에 대한 가중치를 업데이트하는 제4단계를 포함하여 이루어질 수 있다.

**발명의 효과**

[0006] 본 발명에 의하면, 현재시간의 n-1번째 층과 n번째 층 사이에 존재하는 가중치의 평균제곱편차를 분석하여 이 분석결과를 바탕으로 가중치의 평균제곱편차를 최소화할 수 있도록 현재시간의 적응형 학습률을 업데이트하게 되므로, 새로운 환경이나 충격 소음 발생시 또는 학습에 잘못된 데이터가 섞일 경우에도 이러한 충격 소음이나 잘못된 학습 데이터에 의해서 가중치를 업데이트하는 것을 방지할 수 있어 견고한 신경망을 구축할 수 있게 된다.

**도면의 간단한 설명**

- [0007] 도 1은 본 발명에 의한 평균제곱편차 분석기반의 적응형 학습률을 갖는 역전과 신경망 학습방법을 설명하기 위한 역전과 신경망의 토폴로지 예시도이다.  
도 2는 본 발명에 의한 평균제곱편차 분석기반의 적응형 학습률을 갖는 역전과 신경망 학습과정을 예시한 흐름도이다.

**발명을 실시하기 위한 구체적인 내용**

- [0008] 이하, 본 발명의 바람직한 실시 형태에 따른 평균제곱편차(MSD) 분석기반의 적응형 학습률을 갖는 역전과 신경망 학습방법의 구성 및 그에 의한 작용 효과를 첨부 도면을 참조하여 상세히 설명한다.
- [0009] 본 명세서 및 청구범위에 사용된 용어나 단어는 통상적이거나 사전적인 의미로 한정 해석되지 아니하며, 발명자는 그 자신의 발명을 가장 최선의 방법으로 설명하기 위해 용어의 개념을 적절하게 정의할 수 있다는 원칙에 입각하여 본 발명의 기술적 사상에 부합하는 의미와 개념으로 해석되어야만 한다. 따라서, 본 명세서에 기재된 실시 예와 도면에 도시된 구성은 본 발명의 가장 바람직한 일 실시 예에 불과할 뿐이므로, 본 출원시점에 있어서 이들을 대체할 수 있는 다양한 균등물과 변형예들이 있을 수 있음을 이해하여야 한다.
- [0010] 도 1은 본 발명에 의한 평균제곱편차 분석기반의 적응형 학습률을 갖는 역전과 신경망 학습방법을 설명하기 위한 역전과 신경망의 토폴로지 예시도이고, 도 2는 본 발명에 의한 평균제곱편차 분석기반의 적응형 학습률을 갖는 역전과 신경망 학습 과정을 예시한 흐름도로서, 본 발명에 의한 평균제곱편차(MSD) 분석기반의 적응형 학습률을 갖는 역전과 신경망 학습방법은 도면에 예시된 바와 같은 역전과 신경망 알고리즘에서 필수적으로 유도되는 n번째 층의 j번째 뉴런에 존재하며, n-1번째 층과 n번째 층 사이에 연결되어 있는 가중치의 오차의 영향으로 발생하는 역전과 오차를 산출하는 제1과정(a), 현재 시간의 n-1번째 층과 n번째 층 사이에 존재하는 가중치의 평균제곱편차를 분석하여, 이 분석결과를 바탕으로 가중치의 평균제곱편차를 최소화할 수 있는 현재시간의 적응형 학습률을 업데이트하는 제2과정(b;b1-b4)를 포함하여 구성될 수 있다. 여기서 제2과정은 n-1번째 층과 n번째 층 사이의 가중치를 업데이트하기 위한 현재시간의 적응형 학습률을 산출하는 제1단계(b1), 제1단계에서 산출된 현재 시간의 적응형 학습률과 현재 시간의 임의로 주어진 현재 시간의 역전과 오차와 n-1번째 층의 입력값의 정규화값을 통해 n-1번째 층과 n번째 층 사이에 존재하는 현재 시간의 가중치를 이용하여 n-1번째 층과 n번째 층 사이에 존재하는 다음 시간의 가중치를 산출하는 제2단계(b2), 제1단계에서 산출된 현재 시간의 적응형 학습률을 이용하여 n-1번째 층과 n번째 층 사이에 존재하는 가중치를 위한 다음 시간의 평균제곱편차 분석치를 산출하는 제3단계(b3), 및 제1단계 내지 제3단계를 반복 수행하여 각 은닉층에 대한 가중치를 업데이트하는 제4단계(b4)를 포함하여 이루어질 수 있다.

- [0011] 제1과정에서 산출되는 역전과 오차( $\delta_j^n(k)$ )는 n-1번째 층의 각 뉴런에서 발생하는 최종값으로 수렴된 입력값( $\mathbf{u}_t^{n-1,T}$ ), n-1번째와 n번째 사이에 연결된 가중치가 최종값으로 수렴되었을 때의 값( $\mathbf{W}_{jt}^{n-1,n}$ ), n-1번째 층의 각 뉴런에서 발생하는 현재시간 k에서의 입력값( $\mathbf{u}^{n-1}(k)^T$ ), 각 뉴런에서 발생할 수 있는 외란( $\eta_n(k)$ ), 및 일종의 외란의 일부로서  $u_t$ 를  $u(k)$ 로 변경할 때 발생하는 차이의 값( $\gamma_n(k)$ )으로 정의되는 파라미터들에 의하여, 아래의 수학적 식 1로 산출된다.

[0012] [수학적 식 1]

$$\begin{aligned} \delta_j^n(k) &= \mathbf{u}_t^{n-1,T} \mathbf{W}_{j,t}^{n-1,n} - \mathbf{u}^{n-1}(k)^T \mathbf{W}_j^{n-1,n}(k) + \eta_n(k), \\ &= \mathbf{u}^{n-1}(k)^T \mathbf{W}_{j,t}^{n-1,n} - \mathbf{u}^{n-1}(k)^T \mathbf{W}_j^{n-1,n}(k) + r_n(k), \\ &= \mathbf{u}^{n-1}(k)^T \tilde{\mathbf{W}}_j^{n-1,n}(k) + r_n(k), \end{aligned}$$

- [0013]
- [0014] 여기서 t는 그 변수들이 수렴해야 할 목표(target) 값이 된다.

- [0015] 제1단계에서 적응형 학습률( $\mu_j^n(k)$ )은 임의로 설정되는 정수값( $\alpha_{n-1}$ )과 n번째 층의 뉴런개수( $L_{n-1}$ )와 최종

에러의 변화량( $\sigma_{rn}^2(k)$ ) 및 초기 평균제곱편차 분석치( $Tr\{P(1)\}$ )에 대하여, 아래의 수학적 식 2에 의해 산출한다.

[0016] [수학적 식 2]

$$\mu_j^n(k) = \frac{Tr\{\mathbf{P}_j^{n-1,n}(k)\}/(\alpha_{n-1}L_{n-1})}{Tr\{\mathbf{P}_j^{n-1,n}(k)\}/(\alpha_{n-1}L_{n-1})+\sigma_{rn}^2(k)/(\|\mathbf{u}^{n-1}(k)\|^2+\epsilon)}$$

[0018] 여기서  $\alpha_{n-1}$ 은 1보다 큰 정수로 설정하며, 초기  $Tr\{P(1)\}$ 은 약 5~20 사이로 설정하고,  $\epsilon$ 는 고정된 정수값으로 설정한다.

[0019] 제2단계에서 다음 시간의 가중치( $\mathbf{w}_j^{n-1,n}(k+1)$ )는 제1단계에서 산출된 현재시간(k)의 적응형 학습률( $\mu_j^n(k)$ )과 제1과정에서 설정된 임의로 주어진 현재시간(k)의 역전과 오차( $\delta_j^n(k)$ )와 n-1번째 층의 입력값의 제1정규화값( $\frac{\mathbf{u}^{n-1}(k)}{\|\mathbf{u}^{n-1}(k)\|^2}$ )을 통해 n-1번째 층과 n번째 층 사이에 존재하는 현재 시간(k)의 가중치( $\mathbf{w}_j^{n-1,n}(k)$ )를 이용하여, 아래의 수학적 식 3에 의해 산출한다.

[0020] [수학적 식 3]

$$\mathbf{w}_j^{n-1,n}(k+1) = \mathbf{w}_j^{n-1,n}(k) + \mu_j^n(k)\delta_j^n(k)\frac{\mathbf{u}^{n-1}(k)}{\|\mathbf{u}^{n-1}(k)\|^2}$$

[0022] 제3단계에서 다음 시간의 평균제곱편차 분석치( $Tr\{P_j^{n-1,n}(k+1)\}$ )는,

[0023] 제1단계에서 산출된 현재 시간(k)의 적응형 학습률( $\mu_j^n(k)$ )과, 1보다 큰 정수로 설정되는 임의의 값( $\alpha_{n-1}$ )과 n-번째 층의 뉴런개수( $L_{n-1}$ )와 최종 에러의 변화량( $\sigma_{rn}^2(k)$ )과 n-1번째 층의 입력값의 제2정규화값( $\|\mathbf{u}^{n-1}(k)\|^2$ ) 및 현재 시간의 평균제곱편차 분석치( $Tr\{P_j^{n-1,n}(k)\}$ )를 이용하여, 아래의 수학적 식 4에 의해 산출한다.

[0024] [수학적 식 4]

$$Tr\{\mathbf{P}_j^{n-1,n}(k+1)\} = \left[1 - \frac{2\mu_j^n(k)-\mu_j^n(k)^2}{\alpha_{n-1}L_{n-1}}\right] Tr\{\mathbf{P}_j^{n-1,n}(k)\} + \frac{\{\mu_j^n(k)\}^2\sigma_{rn}^2(k)}{\|\mathbf{u}^{n-1}(k)\|^2+\epsilon}$$

[0026] 이상의 본 발명에 의한 평균제곱편차(MSD) 분석기반의 적응형 학습률을 갖는 역전과 신경망 학습방법에서는 적응형 학습률( $\mu_j^n(k)$ )을 수학적 식 2와 같이 업데이트할 수 있게 되므로 각 역전과 오차가 크게 되어 평균제곱편차 분석값( $Tr\{P\}$ )이 크게 되면 적응형 학습률의 값이 거의 1에 가까워져 신경망의 업데이트 속도가 매우 빨라지고, 역전과 오차가 작아지면 적응형 학습률이 0에 가까워져 정상상태 오차(steady state error)를 작게 만들어줄 수 있게 된다. 또한 갑작스러운 충격 소음이나 잘못된 학습 데이터가 들어오는 경우에는 최종 에러의 변화량( $\sigma_{rn}^2(k)$ )의 값이 매우 크게 변하기 때문에 적응형 학습률은 매우 작아져 잘못된 가중치의 업데이트를 방지할 수 있게 된다.

[0027] 이와 같은 본 발명에 의한 평균제곱편차(MSD) 분석 기반의 적응형 학습률을 갖는 역전과 신경망 학습방법에서는 역전과 알고리즘에서 보편적으로 사용하는 각 층 내부 뉴런의 역전과 에러 신호 발생 원인을 각 층에 존재하는 가중치가 최종적으로 학습되기를 원하는 가상의 목표 가중치와의 차이인 가중치 오차에 의해 발생한다는 새로운 개념을 도입하여 이러한 가중치 오차(deviation)를 최소화하는 방향으로 학습률을 실시간 업데이트할 수 있게 되므로, AI 알고리즘의 가장 기본적 분야인 신경망 구조의 역전과 알고리즘 분야에서 고정된 학습률(learning rate)값을 가지는 역전과 알고리즘이 아닌 실시간 에러에 따라 학습률이 변화하는 적응형 학습률을 갖는 역전과

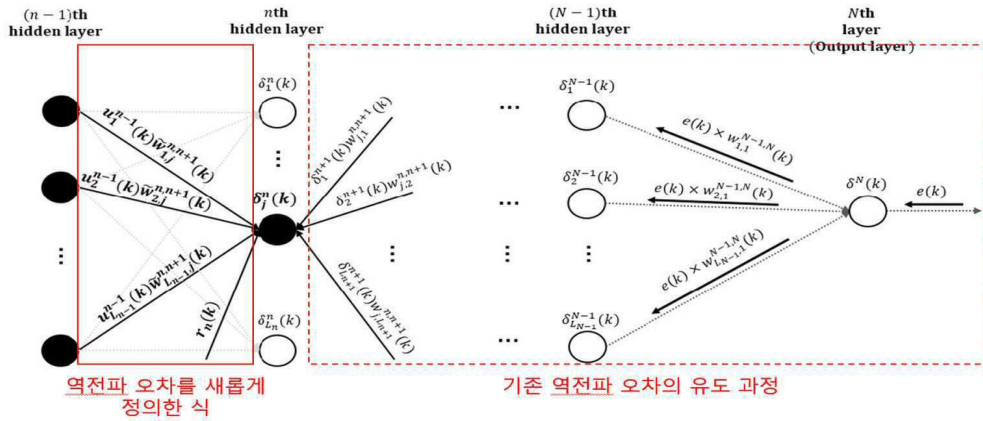
신경망을 구축할 수 있게 된다.

[0028]

이와 같은 본 발명의 적응형 학습률을 갖는 역전파 신경망 알고리즘이 적용된 역전파 신경망에서는 새로운 환경 (충격소음 발생시 또는 학습에 bad data가 섞일 경우)에도 강인한 AI 알고리즘들을 추가적으로 응용 및 개발할 수 있게 하는 이점도 제공할 수 있게 된다.

도면

도면1



도면2

