칼만-버슈 필터 이론 기반 미분 신경회로망 학습

Learning of Differential Neural Networks Based on Kalman-Bucy Filter Theory

조 현 철*, 김 관 형

(Hyun Cheol Cho¹ and Gwan Hyung Kim²)

¹Ulsan College

²Tongmyong University

Abstract: Neural network technique is widely employed in the fields of signal processing, control systems, pattern recognition, etc. Learning of neural networks is an important procedure to accomplish dynamic system modeling. This paper presents a novel learning approach for differential neural network models based on the Kalman-Bucy filter theory. We construct an augmented state vector including original neural state and parameter vectors and derive a state estimation rule avoiding gradient function terms which involve to the conventional neural learning methods such as a back-propagation approach. We carry out numerical simulation to evaluate the proposed learning approach in nonlinear system modeling. By comparing to the well-known back-propagation approach and Kalman-Bucy filtering, its superiority is additionally proved under stochastic system environments.

Keywords: differential neural networks, Kalman-Bucy filter, learning, augmented state estimation

I. 서론

인공 신경회로망 기술은 신호처리, 제어시스템, 영상인식 등과 같이 다양한 공학적 분야에서 널리 사용되고 있다. 신경회로망은 인간의 정보 메커니즘을 공학적 논리구조로 구성한 수학적 모델로서, 학습을 통해 주어진 문제를 해결하는 기능을 갖는다. 신경회로망 학습이란 시스템으로부터 주어진 입출력 패턴에 대하여 최적화 기법을 통해 목적함수를 만족하는 최적의 신경회로망 파라미터를 탐색하는 과정을 의미한다. 이러한 최적화 기법으로는 주로 Newton 기법이나 최급강하(steepest descent) 최적화 등을 이용하여 주어진 시스템에 적합한 알고리즘을 도출한다. 하지만 이러한 최적화 기반 학습 알고리즘에는 일반적으로 경도(gradient) 함수가 포함되어 있으므로 학습 과정 중에 파라미터의 국부 해(local solution)를 가질 수 있어 최적의 신경회로망 학습이 어려워 전체 시스템 성능을 저하시키는 결과를 초래한다. 최근에는 국부 해에 접근하지 않고 보다 빠른 수렴특성을 갖도록 칼만 필터 이론을 적용한 신경회로망 파라미터 학습법이 다양하게 발표되고 있다.

Li는 기존의 오차역전파(backpropagation) 학습 알고리즘과 칼만 필터를 적용한 신경회로망 학습기법 간의 성능을 학습 데이터 패턴에 따른 차이점을 분석하였으며[1], Takenga 등은 경도함수 기반 최적화, 일반 칼만 필터, decouple 칼만 필터 구조를 각각 신경회로망 학습에 적용하여 학습 성능을 서로 비교 분석하였다[2]. Wang 등은 다양한 활성화 함수에 대하여 다층 신경회로망의 확장된 칼만

필터 기법을 적용하여 XOR 패턴 분류, 패리티(parity) 발생, 숫자인식, piecewise 선형 함수의 근사화(approximation) 등과 같은 문제에 적용하였다[3]. 또한 unscented 칼만 필터[4]를 이용하여 신경회로망의 온라인 학습과 비선형 상태 추정 알고리즘을 제안한 논문이 각각 발표되었다[5,6]. 적용분야에 대해서도 칼만 필터 기반 신경회로망은 다양한 분야에 응용이 되었으며, Choi 등은 simultaneous localization & map building(SLAM) 문제에 적용하였으며[7], Kramer 등은 여러 형태의 패턴을 갖는 대상물의 추적 알고리즘을 개발하는데 효율적으로 적용하였다[8]. 최근에는 순환형(recurrent) 신경회로망 모델에 칼만 필터 기법을 적용한 사례가 많이 늘고 있다. Rubio 등은 순환형 신경회로망의 강인성을 위하여 칼만 필터 이론을 적용하였으며[9], Choi 등은 시변 통신 채널의 equalizer를 개발하기 위하여 순환형 신경회로망을 구성하였으며 unscented 칼만 필터링 알고리즘을 학습에 도입하였다[10]. 또한 Coelho 등은 고속 학습 시간을 위하여 칼만 필터 기반 complex real time recurrent learning (RTRL)을 제안하였으며[11], Mirikitani 등은 ensemble 칼만 필터 구조를 이용한 순환형 신경회로망의 적응형 학습 알고리즘을 제안하였다[12]. 보다 최근에는 Wang 등이 이러한 순환형 신경회로망에 확장형 칼만 필터 알고리즘을 적용하였을 때 그 수렴특성을 분석하였다[13]. 또한 이러한 학습의 수렴성을 개선하기 위하여 R 및 Q 적응형 학습 기법을 제안하였으며, 여기서 R과 Q는 칼만 필터 알고리즘에서 관측 노이즈 및 프로세서 노이즈의 공분산 행렬을 각각 의미한다.

현재까지의 칼만 필터 이론을 적용한 다양한 신경회로망 학습과 적용 사례가 발표가 되고 있지만, 대부분 이산 시간 모델의 신경회로망을 다루고 있으며 연속시간 신경회로망 모델에 적용한 사례는 거의 없다. 최근에 미분 신경회로망

* 책임저자(Corresponding Author)

논문접수: 2011. 5. 20., 수정: 2011. 6. 5., 채택확정: 2011. 6. 20.

조현철: 울산과학기술대 전기전자학부 조교수(hcjo@uc.ac.kr)

김관형: 동명대학교 컴퓨터공학과 전임강사(taichiboy1@gmail.com)

과 같은 연속시간 시스템의 모델링에 적합한 다양한 신경 회로망이 각광을 받고 있으므로 이와 같은 연속시간 신경 회로망의 학습 알고리즘의 연구가 절실히 요구되고 있다.

본 논문은 기존의 칼만-버쉬 필터 이론을 적용하여 연속 시간 미분 신경회로망 모델의 학습 알고리즘을 제안한다. 칼만-버쉬 필터링은 연속시간 시스템의 상태벡터 추정에 널리 적용되고 있는 알고리즘이다. 본 논문은 비선형 미분 신경회로망을 설계하며 신경회로망 파라미터 벡터와 상태 벡터를 함께 구성한 augment 형태의 비선형 상태방정식을 제안하여 칼만-버쉬 이론을 도입한 확률 시스템의 상태 추정 알고리즘을 제안한다. 컴퓨터 시뮬레이션을 통해 제안하는 학습 알고리즘의 타당성을 검증하였으며 기존의 오차역전과 학습의 다층 신경회로망 및 일반 칼만-버쉬 필터링과의 비교하여 성능의 우수성을 검증한다.

II. 칼만-버쉬 필터 알고리즘

칼만-버쉬 필터링은 연속시간 시스템의 상태추정을 위해 적용되며 플랜트의 출력에 대한 추정오차를 감소함으로써 상태를 추정하는 방식으로 행해진다[14]. 본 절은 기존의 잘 알려진 칼만-버쉬 필터링을 간략히 서술한다. 우선 다음과 같은 비선형 미분방정식으로 표현되는 연속시간 시스템 모델을 고려한다.

$$\begin{cases} \dot{x}(t) = f(x, u, w, t) \\ y(t) = h(x, \delta, t) \end{cases} \quad (1)$$

여기서 $x \in R^n$ 은 시스템 상태벡터, $u \in R^m$ 은 입력벡터, $w \in R^s$ 은 상태 외란벡터, $y \in R^l$ 은 출력벡터, 그리고 $\delta \in R^l$ 은 출력 외란벡터를 나타내며, $f: R^n \rightarrow R^n$ 및 $h: R^n \rightarrow R^l$ 은 비선형 상태 및 출력 함수를 각각 나타낸다. 칼만-버쉬 필터링 이론에 의하면 식 (1)의 시스템에 대한 필터의 상태 추정 방정식은 다음과 같다.

$$\dot{\hat{x}}(t) = f(\hat{x}(t), u, t) + K(t)(y(t) - h(\hat{x}(t))) \quad (2)$$

여기서 필터의 이득 행렬 $K(t)$ 와 공분산 행렬 $P(t)$ 에 대한 방정식은 각각

$$K(t) = P(t)H^T(t)R^{-1}(t) \quad (3)$$

$$\begin{aligned} \dot{P}(t) &= F(t)P(t) + P(t)F^T(t) + L(t)Q(t)L^T(t) \\ &\quad - P(t)H^T(t)R^{-1}(t)H(t)P(t) \end{aligned} \quad (4)$$

와 같으며 여기서 R 와 Q 은 확률벡터 w 와 δ 의 공분산(covariance) 행렬로서 각각 $R = E(\delta\delta^T)$ 와 $Q = E(ww^T)$ 로 정의되며 행렬 P 의 초기상태 조건은 실제 시스템의 초기값과 추정치의 초기값 간의 편차에 대한 공분산으로 주어지며 $P(0) = E\{[(x(0) - \hat{x}(0))][x(0) - \hat{x}(0)]^T\}$ 와 같으며 그 외의 행렬은 각각 다음과 같이 연산된다.

$$F(t) = \left[\frac{\partial f}{\partial x} \right]_{x=\hat{x}}^T \in R^{n \times n} \quad (5a)$$

$$L(t) = \left[\frac{\partial f}{\partial w} \right]_{x=\hat{x}}^T \in R^{n \times s} \quad (5b)$$

$$H(t) = \left[\frac{\partial h}{\partial x} \right]_{x=\hat{x}}^T \in R^{l \times n} \quad (5c)$$

식 (5)의 자코비안(Jacobian) 행렬은 해당 벡터에 대하여 비선형 상태 함수 f 와 출력 함수 h 를 편미분하여 구해진다.

III. 미분 신경회로망

일반적으로 미분 신경회로망은 연속 시스템 모델로 표현되면 시스템 상태가 미분변수로 표현되는 신경회로망을 일컫는다[15]. 본 논문에서 다루는 미분 신경회로망의 모델은 그림 1과 같으며 상태 및 출력벡터에 대한 방정식은 다음과 같다.

$$\begin{cases} \dot{\sigma}(t) = \phi(G\sigma(t) + Pu(t) + V\beta) \\ \zeta(t) = \Gamma\sigma(t) + \delta(t) \end{cases} \quad (6)$$

여기서 $\sigma \in R^N$ 은 신경회로망 상태벡터, $\beta \in R^q$ 은 바이어스 벡터, $\zeta \in R^l$ 은 신경회로망 출력벡터를 나타내며 $\delta \in R^l$ 은 확률 섭동벡터로 간주되며, $G \in R^{N \times N}$, $P \in R^{N \times m}$, $V \in R^{N \times q}$ 은 신경회로망의 파라미터 행렬, $\Gamma \in R^{l \times N}$ 은 출력 행렬을 각각 나타내며 $\phi: R^N \rightarrow R^N$ 은 비선형 활성화 함수를 나타낸다. 신경회로망 학습은 식 (6)에서 신경회로망의 파라미터로 주어지는 각 행렬 G , P , V 의 요소들을 최적화 기법을 통해 주어진 목적함수를 만족하는 최적의 값으로 결정하는 것을 말한다. 다시 말해, 주어진 플랜트의 입출력 신호의 매핑(mapping)을 통해 동일한 입력에 대하여 동일한 출력신호를 산출하는 기능을 가져야 한다.

IV. 칼만-버쉬 필터링 기반 미분 신경회로망 학습

칼만-버쉬 필터링 기법을 이용한 미분 신경회로망의 학습 알고리즘을 도출하기 위하여, 신경회로망의 상태 및 파라미터 벡터를 포함하는 augment 형태의 새로운 상태벡터를 다음과 같이 정의한다[16].

$$\sigma_a(t) = [\sigma \ \theta]^T \quad (7)$$

여기서 벡터 θ 는 식 (6)의 신경회로망의 모든 파라미터 요소들을 포함하기 위하여 다음과 같이 주어지며

$$\theta = [g \ p \ v]^T \in R^{N(N+m+q)} \quad (8)$$

그림 1. 미분 신경회로망 모델.

Fig. 1. A differential neural networks model.

여기서 $g = \text{col}(G) \in R^{N^2}$, $p = \text{col}(P) \in R^{Nm}$, $v = \text{col}(V) \in R^{Nq}$ 를 각각 나타낸다. 식 (6)을 바탕으로 식 (7)에 대한 상태 및 출력방정식을 다음과 같이 주어진다.

$$\begin{cases} \dot{\sigma}_a(t) = \phi_a(\sigma_a, u, \beta, t) \\ \zeta(t) = \Gamma_a(t)\sigma_a(t) + \delta(t) \end{cases} \quad (9)$$

여기서 $\Gamma_a = [\Gamma \ 0] \in R^{l \times (N(N+m+q))}$ 이다. 식 (9)는 신경회로망의 원 상태 벡터와 파라미터 벡터를 함수 ϕ 와 ψ 에 대하여 서로 분리하여 표현할 수 있으며 다음과 같다.

$$\begin{cases} \dot{\sigma}(t) = \phi(\sigma, \theta, u, \beta, t) \\ \dot{\theta}(t) = \psi(\theta, t) \end{cases} \quad (10)$$

마찬가지로 칼만-버쉬 필터링 이론에 의해 식 (9)에 대한 상태 추정식은 다음과 같이 주어진다.

$$\dot{\hat{\sigma}}_a(t) = \phi_a(\hat{\sigma}_a, u, \beta, t) + K_a(t)[y(t) - \Gamma_a \hat{\sigma}_a(t)] \quad (11)$$

이 추정식은 실제 시스템의 입력력 매핑과 동시에 신경회로망의 파라미터를 실시간으로 추정하는 연산 메커니즘을 갖는다. 또한 이득 행렬 $K_a(t)$ 및 공분산 행렬 $P(t)$ 에 대한 방정식은 각각 다음과 같다.

$$K_a(t) = P_a(t)\Gamma_a^T(t)R_a^{-1}(t) \quad (12)$$

$$\begin{aligned} \dot{P}_a(t) &= F_a(t)P_a(t) + P_a(t)F_a^T(t) + L_a(t)Q_a(t)L_a^T(t) \\ &\quad - P_a(t)\Gamma_a^T(t)R_a^{-1}(t)\Gamma_a(t)P_a(t) \end{aligned} \quad (13)$$

여기서 $R_a = \begin{bmatrix} R & 0 \\ 0 & I \end{bmatrix}$ 및 $Q_a = E(\beta\beta^T)$ 이며 일반적으로 신경회로망의 바이어스는 1이므로 Q_a 는 q 행 q 열의 단위행렬로서 $Q_a = I_q$ 로 표현된다. 최종적으로 식 (13)에서 자코비안 행렬은 각각 다음과 같다.

$$F_a = \begin{bmatrix} \frac{\partial \phi_a}{\partial \sigma_a} \end{bmatrix} = \begin{bmatrix} \frac{\partial \phi}{\partial \sigma} & \frac{\partial \phi}{\partial \theta} \\ \frac{\partial \psi}{\partial \sigma} & \frac{\partial \psi}{\partial \theta} \end{bmatrix} = \begin{bmatrix} \frac{\partial \phi}{\partial \sigma} & \frac{\partial \phi}{\partial \theta} \\ 0 & \frac{\partial \psi}{\partial \theta} \end{bmatrix} \quad (14a)$$

$$L_a = \begin{bmatrix} \frac{\partial \phi_a}{\partial \beta} \end{bmatrix}^T = \begin{bmatrix} \frac{\partial \phi}{\partial \beta} & \frac{\partial \psi}{\partial \beta} \end{bmatrix} \quad (14b)$$

그림 2는 본 논문에서 제안하는 미분 신경회로망의 학습 과정에 대한 개념도를 보여준다.

그림 2. 제안하는 미분 신경회로망의 학습알고리즘 개념도.

Fig. 2. A schematic diagram of the proposed differential neural network leaning algorithm.

V. 컴퓨터 시뮬레이션

본 논문에서 제안한 칼만-버쉬 필터 기반 미분 신경회로망의 모델링 성능을 검증하기 위하여 컴퓨터 시뮬레이션을 실시하였으며 기존의 오차역전파 학습 기능을 갖는 다층 신경회로망과 일반 칼만-버쉬 필터링 기술도 각각 적용하여 성능을 서로 비교 분석하였다. 본 시뮬레이션에 이용한 비선형 확률 시스템의 상태 및 출력방정식 모델은 각각 다음과 같다[15].

$$\begin{cases} \dot{x}_1 = x_2 - 0.05x_1 \cos(x_1) + w_1 \\ \dot{x}_2 = -\sin(x_1) + u - 0.05x_2 \sin(x_2) + w_2 \end{cases} \quad (15a)$$

$$\begin{cases} y_1 = x_1 + \delta_1 \\ y_2 = x_2 + \delta_2 \end{cases} \quad (15b)$$

여기서 제어입력변수는 $u = -5x_1$ 로 정의하였다. 첫 번째 시뮬레이션은 식 (15)에서 상태 및 출력 노이즈 변수를 각각 $w_{1,2} \sim N(0,1)$ 와 $\delta_{1,2} \sim N(0,0.5)$ 인 정지성(stationary) 가우시안 신호로 하여 시스템을 모델링을 실시하였다. 그림 3은 이와 같은 시스템 환경에서의 시스템 출력파형을 보여 준다. 여기서 출력 파형은 0을 기준으로 어느 정도 대칭적인 모습을 보여주며 어느 일정한 값으로 수렴하지 않고 비주기적으로 진동하는 특성을 나타낸다. 시뮬레이션에서 구성한 신경회로망의 구조는 식 (6)에서 $N=5$ 로 하였으며 따라서 $G \in R^{5 \times 5}$ 가 되며 $u, \beta \in R$ 이므로 $P, V \in R^5$ 가 된다. 결국 신경회로망 파라미터의 요소는 모두 35개가 된다. 모델링 시간은 20초로 하였으며 0.1초 간격으로 데이터를 샘플링하였으며 총 200개의 데이터를 획득하였다. 이러한 신경회로망 구조에 대하여 최적의 모델링 결과를 도출할

그림 3. 시스템 출력파형(시뮬레이션 1).

Fig. 3. Waveform of the system output(Simulation 1).

그림 4. 출력 y_1 에 대한 모델링 오차(시뮬레이션 1).

Fig. 4. Modeling error for the output y_1 (Simulation 1).

때까지 반복학습을 실시하였다. 그림 4와 5는 그림 3의 파형에 대하여 오차역전과 기반 다층 신경회로망(MLP), 기존의 칼만-버쉬 필터(KBF) 그리고 본 논문에서 제안한 미분 신경회로망(DNN)을 동일한 시뮬레이션 조건에서 각각 적용하여 모델링의 자승오차, 즉 $E(t) = \sqrt{e^2(t)}$ 를 파형으로 나타내었다. 그림 4와 5로부터 알 수 있듯이 기존의 다층 신경회로망의 경우 가장 큰 오차 파형을 발생하며 본 논문에서 제안한 미분 신경회로망의 가장 작은 오차를 보이고 있으며 기존의 일반 칼만-버쉬 필터의 경우도 비교적 양호한 결과를 보이고 있다. 한편 칼만-버쉬 필터는 정치성 가우시안 시스템의 상태추정에 효과적인 것으로 알려져 있으므로 이러한 모델링 성능은 다소 이상적인 것으로 볼 수 있다. 한편 이 3가지 추정기에 대한 모델링 성능의 정량적 비교를 위하여 샘플한 오차를 모두 합한 오차의 합과 이 값에 전체 샘플 횟수 200을 나눈 평균 오차를 표 1에 나타내었다. 표 1의 수치적 비교를 통해 본 논문의 미분 신경회로망의 경우 가장 우수한 성능을 보이고 있다. 그림 6은 신경회로망의 일부 파라미터의 파형을 보여주고 있으며 시간에 따라 그 값이 적응적으로 변화되는 것을 알 수 있다.

그림 5. 출력 y_2 에 대한 모델링 오차(시뮬레이션 1).

Fig. 5. Modeling error for the output y_2 (Simulation 1).

표 1. 3가지 추정기의 모델링 오차(시뮬레이션 1).

Table 1. Modeling error for the three estimators(Simulation 1).

추정기 \ 추정변수	오차의 합		평균오차	
	y_1	y_2	y_1	y_2
MLP	225	471	1.125	2.355
KBF	139	285	0.695	1.425
DNN	123	191	0.615	0.955

그림 6. 미분 신경망의 일부 파라미터(시뮬레이션 1).

Fig. 6. Partial parameters of the DNN(Simulation 1).

다음으로 앞서 구성한 신경회로망의 일반화(generalization) [17] 성능의 검증을 위하여 두 번째 시뮬레이션을 실시하였다. 시뮬레이션 환경은 식 (15b)에서 출력 노이즈를 분산이 1과 2사이의 항등분포를 갖는 비정치(nonstationary) 가우시안 신호로 하였으며 시스템 출력은 그림 7과 같다. 여기서 출력파형의 형태는 식 (3)의 경우와 매우 유사하지만 진폭이 전반적으로 약 3배 정도 증가하였음을 알 수 있다. 이러한 현상은 출력 노이즈의 분산 값이 1에서 2사이로서 첫 번째 시뮬레이션의 경우보다 그 값이 증가하였기 때문으로 분석된다. 앞서 실시한 시뮬레이션의 경우와 마찬가지로 3가지 추정기를 각각 적용하였으며 각각의 출력변수에 대한 오차파형을 그림 8과 9에 나타내었다. 또한 오차의 합과 평균치를 표 2에 나타내었다. 이러한 결과들을 바탕으로 우선 다층 신경회로망의 경우 모델링 성능이 현저히 저하가 되었음을 알 수 있다. 그 이유는 첫 번째 시뮬레이션의 학습

그림 7. 시스템 출력파형(시뮬레이션 2).

Fig. 7. Waveform of the system output(Simulation 2).

그림 8. 출력 y_1 에 대한 모델링 오차(시뮬레이션 2).

Fig. 8. Modeling error for the output y_1 (Simulation 2).

그림 9. 출력 y_2 에 대한 모델링 오차(시뮬레이션 2).

Fig. 9. Modeling error for the output y_2 (Simulation 2).

표 2. 3가지 추정기의 모델링 오차(시뮬레이션 2).

Table 2. Modeling error for the three estimators(Simulation 2).

추정기	오차의 합		평균오차	
	y_1	y_2	y_1	y_2
MLP	2,915	6,869	14.575	34.345
KBF	748	1,611	3.749	8.055
DNN	326	470	1.632	2.351

그림 10. 미분 신경망의 일부 파라미터(시뮬레이션 2).

Fig. 10. Partial parameters of the DNN(Simulation 2).

과정에서 경험하지 못한 데이터 패턴에 대하여 기존의 다층 신경회로망은 그다지 효과적이지 못하여 일반화 성능이 저하되는 것을 알 수 있다. 일반 칼만-버쉬 모델링의 경우 앞선 시뮬레이션에서는 미분 신경회로망의 성능과 매우 유사하게 나타났지만 비정치성 시스템의 경우 모델링 성능이 저하가 되었음을 알 수 있다. 일반적으로 칼만-버쉬 필터의 경우 정치성 가우시안 시스템에 대하여 모델링 및 상태 추정 성능이 효과적이나 비정치 신호에 대해서는 상대적으로 그 성능이 현저히 떨어지는 단점이 있다[4]. 한편 본 논문에서 제안한 미분 신경회로망의 경우 오차 파형이나 수치적인 결과를 통해 여전히 성능이 매우 양호한 것을 볼 수 있으며 이는 비정치 확률시스템의 경우에도 일반화 특성이 우수하다는 것으로 분석된다. 그림 10은 신경회로망의 일부 파라미터의 파형을 보여주며 대부분의 파라미터 값이 시간에 따라 첫 번째 시뮬레이션의 경우보다 상대적으로 큰 변화의 폭을 보여주며 있다. 이러한 현상은 비정치성 신호에 대하여 적응 강인한 신경회로망의 파라미터 학습이 효과적으로 이루어지고 있다는 것을 입증한다. 이 두 시뮬레이션 결과 및 분석을 통해 본 논문에서 제안하는 칼만-버쉬 필터링 기반 미분 신경회로망 학습 알고리즘이 매우 타당하다는 것을 보여주며 그 성능 또한 우수하다는 것을 검증한다.

VI. 결론

본 논문은 기존의 칼만-버쉬 필터 이론을 미분 신경회로망의 학습 알고리즘을 도출하는데 이용하였다. 우선 미분 신경회로망의 상태 및 파라미터 요소들을 augment 형태의 상태벡터로 구성한 후 칼만-버쉬 필터의 상태 추정 알고리즘을 적용하여 모델링을 실시함과 동시에 신경회로망 파라미터 추정을 연산하는 메커니즘으로 구성하였다. 제안한 학습 알고리즘의 타당성 및 성능의 우수성을 입증하기 위하

여 컴퓨터 시뮬레이션을 실시하였으며 기존의 다층 신경회로망 및 일반 칼만-버쉬 필터 기법도 각각 적용하여 성능을 비교 분석하였다. 시뮬레이션은 정치성 및 비정치성 확률 시스템의 두 가지 경우에 각각 적용하여 모델링 성능을 분석하였다. 기존의 신경회로망의 경우 일반화 성능이 매우 저조한 것으로 분석되었으며, 일반 칼만-버쉬 필터의 경우 정치성 시스템에서는 비교적 만족스러운 성능을 보였으나 비정치성의 경우에는 성능이 현저히 떨어지는 것을 알 수 있었다. 하지만 본 논문에서 제안한 미분 신경회로망의 경우 비정치성 확률 시스템에서도 상대적으로 대단히 우수한 모델링 성능을 보였으며 모델링 오차도 현저히 작은 것을 볼 수 있었다. 향후 연구로는 실제 시스템 구현을 바탕으로 실시간 실험을 통해 제안하는 신경회로망 기술의 실 적용 가능성을 검증하고자 한다.

참고문헌

- [1] S. Li, "Comparative analysis of backpropagation and extended Kalman filter in pattern and batch forms for training neural networks," *International Joint Conference on Neural Networks*, Washington, DC, pp. 144-149, 2001.
- [2] C. M. Takenga, K. R. Anne, K. Kyamakya, and J. C. Chedjou, "Comparison of gradient descent method, Kalman filtering and decoupled Kalman in training neural networks used for fingerprint-based positioning," *Proc. of the IEEE Vehicular Technology Conference*, pp. 4146-4150, 2004.
- [3] K.-W. Wong, C.-S. Leung, and S.-J. Chang, "Use of periodic and monotonic activation functions in multilayer feedforward neural networks trained by extended Kalman filter algorithm," *IEEE Image Signal Process*, vol. 149, no. 4, pp. 217-224, 2002.
- [4] S. J. Julier and J. K. Uhlmann, "Unscented filtering and nonlinear estimation," *Proc. of the IEEE*, vol. 92, no. 3, pp. 401-422, 2004.
- [5] B. Todorovic, M. Stankovic, and C. Moraga, "On-line learning in recurrent neural networks using nonlinear Kalman filters," *Proc. of the IEEE Int. Symposium on Signal Processing and Information Technology*, pp. 802-805, 2003.
- [6] R. Zhan and J. Wan, "Neural network-aided adaptive unscented Kalman filter for nonlinear state estimation," *IEEE Signal Processing Letters*, vol. 13, no. 7, pp. 445-448, 2006.
- [7] M. Choi, R. Sakthivel, and W. K. Chung, "Neural network-aided extended Kalman filter for SLAM problem," *IEEE International Conference on Robotics & Automation*, Roma, Italy, pp. 1686-1690, 2007.
- [8] K. A. Kramer and S. C. Stubberud, "Tracking of multiple target types with a single neural extended Kalman filter," *International Journal of Intelligent Systems*, vol. 25, pp. 440-459, 2010.

- [9] W. Yu and J. de J. Rubio, "Recurrent neural networks training with stable risk-sensitive Kalman filter algorithm," *Proc. of International Joint Conference on Neural Networks*, Montreal, Canada, pp. 700-705, 2005.
- [10] J. Choi, A. C. Lima, and S. Haykin, "Kalman filter-trained recurrent neural equalizers for time-varying channels," *IEEE Transaction on Communications*, vol. 53, no. 3, pp. 472-480, 2005.
- [11] P. H. G. Coelho and L. B. Neto, "Complex RTRL neural networks fast Kalman training," *International Conference in Intelligent Systems Design and Applications*, pp. 573-577, 2007.
- [12] D. T. Mirikitani and N. Nikolaev, "Dynamic modeling with ensemble Kalman filter trained recurrent neural networks," *International Conference on Machine Learning and Applications*, pp. 843-848, 2008.
- [13] X. Wang and Y. Huang, "Convergence study in extended Kalman filter-based training of recurrent neural networks," *IEEE Transaction on Neural Networks*, vol. 22, no. 4, pp. 588-600, 2011.
- [14] J. M. Mendel, *Lessons in Estimation Theory for Signal Processing, Communications and Control*, New Jersey, Prentice Hall, 1995.
- [15] A. S. Poznyak, E. N. Sanchez, and W. Yu, *Differential Neural Networks for Robust Nonlinear Control*, New Jersey, World Science, 2001.
- [16] H. C. Cho, J. W. Lee, Y. J. Lee, and K. S. Lee, "Online dynamic modeling of ubiquitous sensor based embedded robot systems using Kalman filter algorithm," *Journal of Institute of Control, Robotics and Systems (in Korean)*, vol. 14, no. 8, pp. 779-784, 2008.
- [17] S. Haykin, *Neural Networks and Learning Machines*, New Jersey, Prentice Hall, 2009.

조 현 철

1997년 2월 부경대학교 제어계측공학과 졸업(공학사). 1999년 2월 동아대학교 전기공학과 졸업(공학석사). 2006년 8월 미국 네바다 주립 대학교 전기공학과 졸업(Ph.D.). 2006년 9월~2009년 2월 동아대학교 전기공학과 포닥연구원.

2009년 3월~현재 울산과학기술대학교 전기전자학부 조교수. 관심분야는 자동제어 및 응용, 신경회로망, 신재생 발전 시스템.

김 관 형

2001년 2월 한국해양대학교 전자통신공학과 졸업(공학박사). 2000년 3월~현재 동명대학교 컴퓨터공학과 전임강사. 관심분야는 최적제어, 인공지능, 임베디드 시스템, 신호처리.