

# GRU 기반 망각 인자 학습을 활용한 적응형 지수 가중 칼만 필터

## GRU-based Adaptive Exponentially Weighted Kalman Filter With Forgetting Factor Learning

장 선 호<sup>1</sup>, 권 소 연<sup>2</sup>, 권 보 규<sup>3,\*</sup>

(Sun-Ho Jang<sup>1</sup>, Soyeon Kwon<sup>2</sup>, and Bo-Kyu Kwon<sup>3,\*</sup>)

<sup>1</sup>Korea Institute of Robotics Technology Convergence

<sup>2</sup>Department of Electronics Engineering, Chungnam National University

<sup>3</sup>Department of Electronics·AI system Engineering, Kangwon National University

**Abstract:** This study proposes a Gated Recurrent Unit (GRU)-based Adaptive Exponentially Weighted Kalman Filter (GEWKF) to dynamically estimate the forgetting factor, addressing the limitations of conventional exponentially weighted Kalman filters (EWKFs) relying on fixed parameters. Unlike static EWKFs, the proposed GEWKF adaptively tunes the forgetting factor in real time, enhancing robustness and estimation accuracy during abrupt model changes. In contrast to previous deep learning-based adaptive filters that predict complex covariance matrices or full filter gains often at high computational cost and interpretability, GEWKF learns only a single scalar parameter while preserving the recursive framework of the standard Kalman filter. A GRU-based recurrent neural network leverages time-series data, including state prediction errors and measurement residuals, to compute the optimal forgetting factor efficiently and reliably. Simulations reveal that the GEWKF filter matches the performance of the standard Kalman filter under nominal conditions and significantly outperforms both the standard KF and fixed-parameter EWKF under sudden uncertainties. These results highlight that the GEWKF offers a practical and principled framework for state estimation, balancing adaptability, computational efficiency, real-time applicability, and interpretability.

**Keywords:** state estimation, Kalman filter, exponentially weighted Kalman filter, recurrent neural network, gated recurrent unit, adaptive filtering

### 1. 서론

상태 추정기는 로봇의 위치 추정, 자율 주행 차량의 경로 예측, 리튬-이온 배터리의 충전 상태 추정 등 다양한 응용 분야에서 시스템의 안정성과 성능을 보장하기 위한 핵심 기술이다[1]. 이러한 상태 추정 문제에 대해 가장 널리 활용되는 기법 중 하나는 칼만 필터(Kalman filter, KF)로, 시스템 모델과 잡음의 통계적 특성에 기반하여 최적의 상태 추정을 제공하며, 재귀적 계산 구조를 통해 높은 계산 효율성도 확보할 수 있어 다양한 공학 및 산업 응용에 적용되고 있다[2-7]. 그러나 KF는 시스템 모델의 정확성, 초기 조건, 그리고 잡음의 통계적 특성에 크게 의존한다는 근본적인 한계를 가진다. 실제 환경에서는 비정상적인 동작, 모델링 오차, 부정확한 초기 조건, 비가우시안 잡음 등 다양한 불확실성이 존재하며, 이로 인해 상태 추정 성능이 저하되고 누적된 계산 오차가 필터 안정성을 위협할 수 있다[8]. 또한, 특정 조건에 따라 시간이 무한히 증가하면 추정 오차 공분산이 영(0)에 수렴하며 상태 추정 결과가 발산할 수 있는 문제도 존재하고, 시스템 특성이나 잡음 분포의 급격한 변화에 대해서는 강인한 성능을 유지하기 어

렵다[9].

이러한 문제점을 보완하기 위해, 최근에는 과거 정보의 영향력을 줄이고 최근 정보에 높은 가중치를 부여하는 방식으로 설계된 지수 가중 칼만 필터(exponentially weighted Kalman filter, EWKF)가 제안되었다[8-12]. EWKF는 과거 오차가 누적되는 문제를 완화하고, 시스템이나 잡음 특성이 변하는 환경에서도 보다 안정적인 필터링 성능을 제공할 수 있으며, 변수형 가중치를 적용한 확장 연구에서는 비정상 시스템에서도 유효한 추정 성능을 확보할 수 있는 가능성을 보였다[10-12]. 하지만 기존 EWKF는 경험적으로 설정된 고정 가중치에 의존하므로, 시스템 환경이 변화하는 경우에도 동일한 가중치가 적용되어 적절한 적응성을 확보하기 어렵다[8]. 또한 다중 센서 환경에서의 통합 가중치 결정, 상황별 최적 가중치 설정 등에서도 구조적 제약이 존재한다[9].

이러한 배경에서 칼만 필터의 구조에 인공지능 기법을 접목하려는 다양한 시도가 이루어지고 있다. 최근 연구에서는 순환 신경망(recurrent neural network, RNN)을 이용하여 시스템의 동적 모델이나 측정 모델을 직접 학습하거나, 필터의 핵

\*Corresponding Author

Manuscript received August 31, 2025; revised October 5, 2025; accepted November 21, 2025

장선호: 한국로봇융합연구원 연구원(jang1229@kro.re.kr, ORCID<sup>®</sup> 0000-0003-2484-4630)

권소연: 충남대학교 전자공학과 학부생(mjsthdus04@o.cnu.ac.kr, ORCID<sup>®</sup> 0009-0006-4217-633X)

권보규: 강원대학교 교수(bkkwon@kangwon.ac.kr, ORCID<sup>®</sup> 0000-0002-0899-7824)

심 요소인 칼만 이득을 신경망으로 추정하는 방식이 제안되었다[13-15]. 이 중 KalmanNet과 같은 모델은 RNN이 시스템의 시계열 특성을 학습하여 추정 성능을 높일 수 있음을 보였다[13]. RNN은 시계열 데이터의 시간적 의존성을 효과적으로 학습할 수 있으며, 비선형 동적 시스템에서도 복잡한 특성을 모델링할 수 있다는 장점이 있다. 또한 최근에는 불확실성 추정까지 함께 고려하여 가중치를 조정하는 학습 기반 필터가 제안되고 있으며, 이러한 방법은 비정형 환경에서의 적응성을 확보하는 데 효과적임을 보였다[14-15]. 그러나 기존의 RNN 기반 칼만 필터 연구들은 주로 시스템 모델의 불확실성을 나타내는 잡음 공분산 행렬들을 추정하거나, 칼만 이득을 직접 학습하는 방식이 주를 이루었다. 이러한 방식은 높은 예측 정확도를 보일 수 있지만, 복잡한 학습 구조와 높은 계산량으로 인해 실시간 시스템에 적용하는 데 어려움이 있다.

따라서 본 연구에서는 이러한 높은 계산량으로 인해 실시간 시스템에 적용하기 어려운 문제를 해결하기 위해, 기존 EWKF의 재귀적 구조를 유지하면서도 GRU (Gated Recurrent Unit)를 이용하여 단일 스칼라 형태의 망각 인자를 시간에 따라 동적으로 학습·적용하는 적응형 필터 구조 (GEWKF, GRU-based EWKF)를 제안한다. 제안된 GEWKF는 구조의 단순성과 해석 가능성을 유지함과 동시에, 시스템 상태 및 환경 변화에 따라 실시간으로 망각 인자를 조정함으로써 기존 EWKF의 한계를 극복하고 보다 안정적인 상태 추정을 가능하게 한다.

본 논문의 구성은 다음과 같다. 2장에서는 KF 및 EWKF의 이론적 배경을 설명하고, 3장에서는 제안하는 GEWKF의 구조 및 GRU 학습 알고리즘을 기술한다. 4장에서는 다양한 시나리오에서의 시뮬레이션 실험을 통해 제안된 필터의 성능을 검증하며, 마지막으로 5장에서는 본 연구의 결론 및 향후 연구 방향을 제시한다.

## II. 본론

### 1. 칼만 필터 및 지수 가중 칼만 필터

이 절에서는 KF와 EWKF에 대한 알고리즘에 대해 설명하고, EWKF의 한계점 등을 소개한다.

먼저 다음과 같은 이산 선형 시스템을 고려한다.

$$x_k = Ax_{k-1} + Gw_k, \quad (1)$$

$$y_k = Cx_k + v_k, \quad (2)$$

여기서  $x_k \in R^n$ 는 상태 벡터,  $y_k \in R^m$ 는 측정값 벡터이다. 그리고 공정 잡음  $w_k$ 와 측정 잡음  $v_k$ 는 영-평균 가우시안 잡음으로 각각의 공분산은  $Q$ 와  $R$ 로 정의된다.

KF는 시스템 모델 (1), (2)에 대해 다음의 비용함수

$$J = (x_0 - \bar{x}_0)^T P_0^{-1} (x_0 - \bar{x}_0) + \sum_{\tau=0}^{k-1} w_\tau^T Q^{-1} w_\tau + \sum_{\tau=0}^{k-1} (y_\tau - Cx_\tau)^T R^{-1} (y_\tau - Cx_\tau), \quad (3)$$

의 최소화 해가 제공되도록 다음과 같이 설계된다[1].

#### • 예측 단계

$$\hat{x}_{k|k-1} = A\hat{x}_{k-1|k-1}, \quad (4)$$

$$P_{k|k-1} = AP_{k-1|k-1}A^T + GQG^T. \quad (5)$$

#### • 업데이트 단계

$$K_k = P_{k|k-1}C^T(CP_{k|k-1}C^T + R)^{-1}, \quad (6)$$

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k(y_k - C\hat{x}_{k|k-1}), \quad (7)$$

$$P_{k|k} = (I - K_kC)P_{k|k-1}. \quad (8)$$

여기서  $\bar{x}_0$ 는 초기 상태 기대치,  $P_0$ 는 초기 오차 공분산,  $\hat{x}_{\cdot|\cdot}$ 는 추정 상태 벡터,  $P_{\cdot|\cdot}$ 는 추정 오차 공분산,  $K_k$ 는 칼만 이득, 그리고  $I$ 는 단위 행렬이다.

EWKF는 KF의 오차 누적 문제를 완화하고 최근 정보에 더 높은 가중치를 부여하기 위한 필터로, 망각 인자(forgetting factor)  $\lambda (\geq 1)$ 를 사용하여 KF의 공분산 추정 단계에 망각 인자를 적용한다. EWKF는  $\lambda$ 를 예측 오차 공분산  $P_{k|k-1}$ 에 곱하여 불확실성을 높이고, 이를 통해 필터가 최근 측정값에 더 민감하게 반응하도록 다음과 같이 설계된다.

#### • 예측 단계

$$\hat{x}_{k|k-1} = A\hat{x}_{k-1|k-1}, \quad (9)$$

$$P_{k|k-1} = \lambda (AP_{k-1|k-1}A^T + GQG^T). \quad (10)$$

#### • 업데이트 단계

$$K_k = P_{k|k-1}C^T(CP_{k|k-1}C^T + R)^{-1}, \quad (11)$$

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k(y_k - C\hat{x}_{k|k-1}), \quad (12)$$

$$P_{k|k} = (I - K_kC)P_{k|k-1}. \quad (13)$$

수식 (10)에서  $\lambda$ 는 1보다 크거나 같은 값( $\lambda \geq 1$ )을 가지며,  $\lambda = 1$ 일 경우 EWKF는 KF와 동일하게 동작해 KF 대비 보다 일반적인 해를 제공한다.  $\lambda > 1$ 인 경우, 예측 오차 공분산  $P_{k|k-1}$ 의 값이 증가하여 필터 이득  $K_k$ 가 커지게되어 필터가 예측 상태보다 최근 측정값에 더 큰 가중치를 두어 과거 정보의 영향력을 감소시켜 오차 누적의 영향을 줄일 수 있다. 하지만 기존 EWKF는 고정된 망각 인자를 사용하므로, 시스템의 불확실성이 없거나 잡음 특성이 안정적인 환경에서도 불필요하게 필터 성능이 저하되며, 망각 인자를 시스템 변화에 따라 동적으로 조정하지 못하여 잡음 특성의 급격한 변화나 시스템 모델의 부정확성에 반응할 수 없는 구조적 한계를 가진다. 또한, 고정된 망각 인자는 대부분 경험에 의존하여 설계되어, 여러 불확실성 환경에 대해 일반화된 강인성을 확보하는 데도 한계를 가진다.

### 2. GRU 기반 적응 지수 가중 칼만 필터

딥러닝 기술은 기존의 전통적인 방법으로 해결하기 어려운 문제를 효과적으로 처리하며, 다양한 분야에서 혁신적인 성과를 창출하고 있다. 특히, 복잡한 패턴을 학습하고 예측하는 능력은 시스템의 불확실성을 모델링하고 적응적으로 대응하는 데 활용될 수 있다.

본 연구에서는 기존의 EWKF가 환경 변화를 고려하지 못하고 고정된 가중치를 적용하는 한계를 극복하기 위해, 딥러

닝 기반 접근법을 도입한다. 상태 추정 문제는 연속적인 시계열 데이터를 기반으로 하며, 과거의 측정값과 시스템 상태가 현재의 상태 결정에 중대한 영향을 미친다.

순환 신경망은 내부 순환 구조를 통해 시계열 데이터의 장기적인 패턴을 학습할 수 있으며, 이를 통해 최적의 망각 인자를 결정하는 것이 가능하다. 따라서, 본 연구에서는 기존 EWKF의 예측 및 업데이트 단계를 유지하면서, 고정된 망각 인자  $\lambda$  대신 순환 신경망을 활용하여 예측 단계에서 시간  $k$  마다 동적으로 조정되는 망각 인자  $\lambda_k$ 를 적용하는 그림 1과 같은 구조를 제안한다.

본 연구에서 망각 인자  $\lambda_k$ 를 동적으로 결정하기 위한 순환 신경망 구조로는 GRU를 고려한다. 순환 신경망의 대표적인 구조인 LSTM (Long Short-Term Memory)와 GRU는 모두 시계열 데이터의 장기 의존성을 학습할 수 있는 특성을 갖는다. 그러나 본 연구에서 제안하는 방법은 필터의 예측 단계에서 비교적 짧은 시간 내에 망각 인자를 결정해야 하는 실시간 적응성이 중요하다. LSTM 구조는 상대적으로 복잡한 게이트 구조로 인해 연산 비용이 크고 학습 속도가 느릴 수 있어 이러한 요구사항에 한계를 가질 수 있다. 반면, GRU는 LSTM에 비해

#### Information

Current Time :  $y_k, \hat{x}_{k|k-1}$

Previous Time :  $y_{k-1}, \hat{x}_{k-1|k-2}, \hat{x}_{k-1|k-1}$

#### GRU Neural Network

Input Vector :  $u_k = [\Delta_{x,k}, \Delta_{y,k}, d_{x,k}, d_{y,k}]$

GRU Layer 1

GRU Layer 2

$h_k$

Fully Connected Layer

$o_k$

Activation Function :  $ReLU(o_k)$

Exponential Weight  $\lambda_k = ReLU(o_k) + 1$

그림 2. 망각 인자 계산을 위한 GRU 기반 신경망의 구조.

Fig. 2. Structure of the GRU-based neural network for computing forgetting factor.

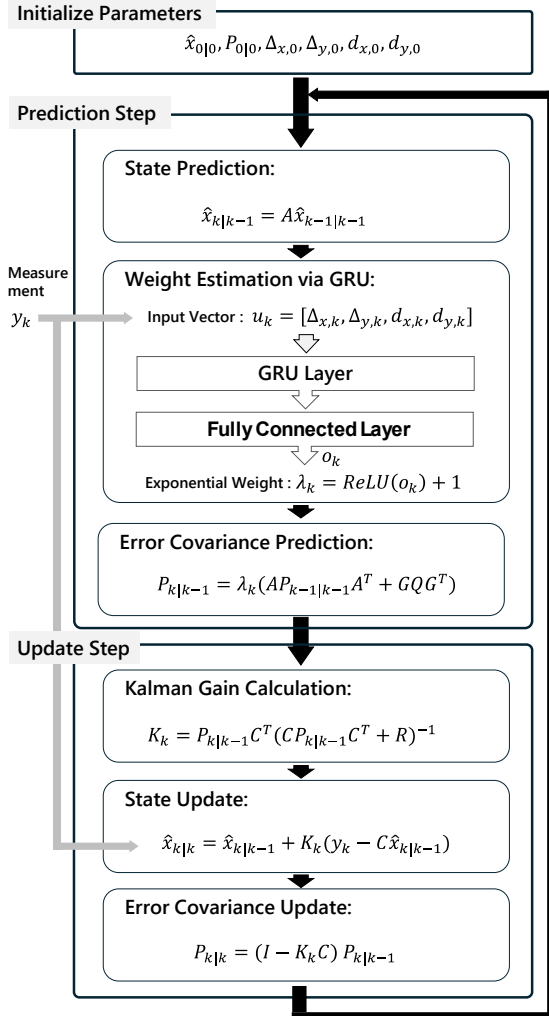


그림 1. 제안하는 필터의 전체 구조.

Fig. 1. Overall Architecture of the proposed filter.

구조가 간결하여 학습해야 할 파라미터 수가 적고 계산량이 감소하는 이점을 갖는다. 이는 실시간 상태 추정 및 적응형 파라미터 학습에 유리하며, [13]에서는 GRU가 상태 추정과 관련된 통계적 정보를 효과적으로 학습할 수 있음이 확인되었다. 또한, GRU는 별도의 셀 상태 없이 게이트 메커니즘만으로 정보 흐름을 제어하므로, 본 연구의 실시간 적응성 목표에 더 부합한다.

따라서 본 연구에서는 계산 효율성과 실시간 성능을 고려하여 그림 2와 같이 입력층, GRU 은닉층, 출력층으로 구성된 GRU 기반의 예측 모델을 설계한다. 이를 통해 예측 단계에서 동적으로 변하는 망각 인자를 효과적으로 학습하고 적용한다.

GRU 기반 신경망의 성능은 입력 벡터 구성에 크게 의존한다. 망각 인자  $\lambda_k$ 는 시스템의 동적 변화, 모델 오차, 잡음 특성의 불확실성 등 다양한 요인을 반영하여 결정되어야 한다. 따라서 그림 2와 같이, 연속된 두 시점의 시스템 상태 및 측정값과 관련된 정보를 포함하는 입력 벡터를 구성하며, 구체적으로 각 요소의 정의는 다음과 같다.

$$\Delta_{x,k} = \hat{x}_{k-1|k-2} - \hat{x}_{k|k-1}, \quad (14)$$

$$\Delta_{y,k} = y_k - y_{k-1}, \quad (15)$$

$$d_{x,k} = \hat{x}_{k-1|k-1} - \hat{x}_{k|k-1}, \quad (16)$$

$$d_{y,k} = y_k - C\hat{x}_{k|k-1}, \quad (17)$$

각 입력 요소는 망각 인자  $\lambda_k$ 의 예측을 위해 서로 다른 정보를 제공하도록 구성되었다. 먼저, 연속된 두 시점의 예측 상태 변화량  $\Delta_{x,k}$ 과 출력 변화량  $\Delta_{y,k}$ 는 시계열 데이터의 급격한 변화율을 반영하여, 불확실성이 급변하는 비정상 환경에 대한 적응에 필요한 동적 정보를 포함한다. 예측 상태 오차

$d_{x,k}$ 는 이전 시점과 현재 시점의 예측 상태 차이를 나타내며, 시스템 모델의 불완전성이 예측에 미치는 영향을 간접적으로 반영한다. 또한, 측정 잔차  $d_{y,k}$ 는 실제 측정값과 예측 출력 간의 직접적인 불일치를 의미하며, 모델링 오차나 관측 잡음의 영향을 나타내는 지표로 활용된다.

이러한 입력 구성은 추정 상태나 측정값의 절대적인 크기를 직접 사용하는 대신, 연속된 시점 간의 변화량 정보를 활용함으로써 시스템의 동적 특성을 보존하면서도 기존 오프셋 및 센서 편차 성분을 제거한다. 이로써 GRU 신경망은 정상 상태의 정적 패턴보다 변화 양상에 집중하여 학습 효율성을 높일 수 있다. 또한 변화량 입력은 평균이 0에 가까운 분포 특성을 가지므로 입력 자체에 정규화 효과가 내재되어 있으며, 별도의 정규화 과정을 거치지 않아도 학습 안정성이 확보된다. 이러한 입력 특성은 그라디언트 폭주(exploding gradient)와 수렴 불안정성을 완화하여 학습의 수렴 속도와 안정성을 동시에 향상시킨다. 결과적으로, 본 입력 구조는 학습 목표를 단순화하면서 망각 인자  $\lambda_k$ 의 동적 변화를 보다 안정적이고 효율적으로 학습할 수 있는 근거를 제공한다.

따라서 신경망 학습을 위한 입력 벡터  $u_k$ 는  $2(n+m)$  차원을 가지며 다음과 같이 정의된다.

$$u_k = [\Delta_{x,k} \ \Delta_{y,k} \ d_{x,k} \ d_{y,k}]. \quad (18)$$

입력층을 통과한 입력 벡터는 GRU 신경망의 핵심을 이루는 은닉층으로 전달된다. 본 연구에서는 복잡한 시계열 데이터의 특징을 효과적으로 학습하고 실시간성을 보장하기 위해 GRU 은닉층을 적층하는 구조를 채택한다. 은닉층의 개수가 증가할수록 복잡한 시스템의 동적 모델링에 유리하나 연산 복잡성도 증가시켜 실시간 처리 성능을 저하시킬 수 있는 trade-off 관계를 가진다. 따라서 본 연구에서는 그림 2와 같이 GRU 은닉층의 개수를 2개로 제한하여 네트워크를 구성함으로써 복잡한 데이터에 대한 모델 표현력과 연산 효율성의 균형을 유지하고자 한다. 각 GRU 은닉층의 크기는 공정 잡음과 측정 잡음의 공분산의 상대적 크기를 고려하여, 시스템 불확실성 수준을 반영할 수 있도록  $k(m^2+n^2)$  형태로 설정한다. 여러 정수  $k$ 에 대해 학습 안정성과 예측 오차를 다수의 실험적 경험에 의하여 비교한 결과,  $k=6$ 일 때 가장 우수한 수렴성과 일반화 성능을 보여, 본 연구에서는 은닉층의 크기를  $6(m^2+n^2)$ 개로 구성한다. GRU 셀(cell)은 업데이트 게이트(Update Gate)와 리셋 게이트(Reset Gate)의 두 게이트 메커니즘을 활용하여 입력 정보와 이전 은닉 상태를 조합하고 시간적 의존성을 학습한다. 업데이트 게이트는 과거 정보를 현재 상태에 얼마나 반영할지를 결정하며, 리셋 게이트는 이전 은닉 상태가 현재 입력과 얼마나 관련이 있는지를 판단하여 불필요한 정보를 망각하는 역할을 수행한다.

GRU 은닉층을 거친 후, 학습된 정보는 출력층을 거쳐 망각 인자  $\lambda_k$ 로 변환된다. 본 연구에서는 각 시점에서 단일 스칼라 값의 망각 인자를 도출하기 위해 하나의 뉴런으로 구성된 출력층을 사용하며, 이 출력층은 GRU 은닉 상태에 대해 선형 결합을 수행한 결과인  $o_k$ 를 출력한다. 이후, 출력값  $o_k$ 는 ReLU(Rectified Linear Unit) 활성화 함수를 통해 비선형 변환되며,

항상 0 이상의 값을 갖도록 한다.

마지막으로, 다음과 같이 망각 인자  $\lambda_k$ 를 설정하여 표준 KF에 해당하는  $\lambda_k=1$ 을 포함하고, 1 이상의 망각 인자값을 가질 수 있도록 한다.

$$\lambda_k = \text{ReLU}(o_k) + 1 = \max(0, o_k) + 1. \quad (19)$$

수식 (20)에서 출력  $o_k$ 가 0 이하일 경우  $\text{ReLU}(o_k)=0$ 이 되어  $\lambda_k=1$ 이 된다. 이는 신경망이 학습을 통해 KF가 최적이라고 판단될 경우 그에 해당하는 값을 출력할 수 있음을 의미한다. 반면,  $o_k > 0$ 일 경우에는  $\lambda_k = o_k + 1 > 1$ 이 되어 공정 잡음 및 측정 잡음 공분산 행렬의 크기를 조절하는 역할을 수행하며,  $\lambda_k$ 가 1보다 클수록 예측 오차 공분산 행렬의 크기가 KF 대비 커져, 최근 측정치의 영향을 높이는 효과를 가져온다. 이를 통해 제안하는 GEWKf는 최근 정보에 동적으로 더 높은 가중치를 부여하고 과거 정보의 영향력을 조절함으로써 변화하는 시스템 조건이나 잡음 환경에 빠르게 적응하고 오차 누적을 완화할 수 있다.

신경망 학습은 상태 추정 오차를 최소화하는 방향으로 진행된다. 본 연구에서는 손실 함수(loss function)로 평균 제곱 오차(mean squared error, MSE)를 사용한다. MSE는 추정된 상태 벡터와 실제 시스템 상태 벡터 간의 제곱 차이 평균으로, 상태 추정 정확도를 직접적으로 반영하며 경사 하강법 기반 최적화에 용이하다. 손실 함수는 다음과 같이 정의된다.

$$L = \frac{1}{N} \sum_{k=1}^N \|x_k - \hat{x}_{k|k}\|^2. \quad (20)$$

여기서,  $x_k$ 는 실제 상태 벡터이며,  $\hat{x}_{k|k}$ 는 제안된 GEWKf를 통해 추정된 상태 벡터, 그리고  $N$ 은 학습 데이터 시퀀스의 길이이다. 신경망의 내부 가중치 파라미터들은 표준적인 역전파 알고리즘을 사용하여 이 MSE 손실 함수를 최소화하도록 반복적으로 업데이트된다. 역전파는 손실 함수의 기울기를 계산하여 네트워크 가중치를 조정함으로써, GRU 네트워크가 입력 벡터  $u_k$ 로부터 최적의 망각 인자  $\lambda_k$ 를 예측하는 복잡한 비선형 관계를 학습하도록 한다.

제안하는 GEWKf의 성공적인 초기 학습과 실제 적용 환경에서의 강인한 성능을 보장하기 위해서는 학습 데이터셋의 품질과 다양성이 매우 중요하다. 신경망이 학습하지 못한 다양한 시스템 및 환경에 일반화된 성능을 갖추기 위해서는, 학습 데이터가 이러한 잠재적 변화를 폭넓게 반영해야 한다. 이를 위해 학습 데이터 생성 단계에서 시스템 모델의 불확실성과 시간적 변동 특성을 체계적으로 모사한다. 구체적으로, 시스템 행렬  $A$ 와 측정 행렬  $C$ 에 구조적 섭동을 가하여 모델 오차를 포함하고, 공정 잡음과 측정 잡음의 공분산  $Q$ 와  $R$ 의 크기를 균일 분포 내에서 스케일링하여 다양한 잡음 수준을 구현한다. 또한, 각 불확실성 매개변수의 변화 시점과 지속 시간을 무작위로 설정함으로써 시변 불확실성 조건을 반영한다. 이러한 다차원적 불확실성 조건에서 생성된 몬테카를로 기반 학습 데이터셋은 GRU 신경망이 불확실성의 시변 특성을 학습하고, 새로운 환경에서도 안정적이고 정확한 상태 추정을 수행할 수 있는 일반화 능력을 확보하는 기반이 된다.

안정적인 학습 과정은 학습률과 같은 핵심 파라미터의 적절한 설정에 크게 좌우되며, 이는 모델의 수렴 속도와 최종 성능에 직접적인 영향을 미친다. 특히 RNN 학습 시 나타날 수 있는 기울기 소실 또는 폭발 현상은 학습의 불안정성을 야기할 수 있으므로, 이를 완화하기 위한 학습률 스케줄링, 정규화, 그래디언트 클리핑 등의 기법 적용이 효과적일 수 있다. 따라서 본 연구에서도 이러한 점들을 고려하여 초기 학습 과정을 설계하고, 최종적으로 학습된 모델의 일반화 성능은 MSE와 같은 평가 지표를 통해 검증한다.

### III. 실험

이 장에서는 제안한 GEWKF의 성능을 표준 KF 및 고정된 가중치를 사용하는 지수 가중 칼만 필터와 비교 분석을 통해 검증한다. 시뮬레이션은 시스템 모델의 불확실성에 대해 제안한 GEWKF 필터의 상태 추정 정확도, 적응성, 그리고 강인성을 정량적으로 평가한다.

#### 1. 대상 시스템 및 실험 시나리오

다음과 같이 선형 상태 공간 모델로 표현되는 F404 엔진 모델을 고려한다[9].

$$x_{k+1} = \begin{bmatrix} 0.9305 & 0 & 0.1107 \\ 0.0077 & 0.9802 & -0.0173 \\ 0.0142 & 0 & 0.8953 \end{bmatrix} x_k + w_k, \quad (21)$$

$$y_k = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} x_k + v_k. \quad (22)$$

공정 잡음  $w_k$ 와 측정 잡음  $v_k$ 는 모두 평균이 0인 백색 가우시안 잡음으로 가정하며, 각각의 공칭(nominal) 공분산 행렬은  $Q_{nom} = 0.1^2 \cdot I$ 와  $R_{nom} = 0.05^2 \cdot I$ 로 설정한다.

제안된 GEWKF의 적응 성능과 강인성을 종합적으로 평가하기 위해 시스템 모델의 부정확성 시나리오를 설계하였다. 시뮬레이션은 250초 동안 진행되며, 전체 시뮬레이션 기간 중 50초부터 75초 사이의 특정 구간에서 시스템 모델에 의도적인 오차를 추가하였다. 이 모델 오차는 실제 시스템 행렬과 측정 행렬에 다음의  $\Delta A$ 와  $\Delta C$  행렬이 더해져 동적으로 변하는 환경을 모사한다.

$$\Delta A = \begin{bmatrix} 0.1 & 0 & 0 \\ 0 & 0.1 & 0 \\ 0 & 0 & 0.1 \end{bmatrix}, \Delta C = \begin{bmatrix} 0.01 & 0 & 0 \\ 0 & 0.01 & 0 \end{bmatrix}.$$

이 시뮬레이션 환경은 실제 시스템의 동특성이 필터에 사용된 공칭 모델과 다른 상황을 가정하며, 필터가 모델링 오차에 얼마나 효과적으로 대처하는지 평가한다.

필터의 상태 추정 성능에 대한 통계적 신뢰성을 확보하기 위해, 본 연구에서 제시하는 모든 실험 결과는 500회의 독립적인 몬테카를로 시뮬레이션의 앙상블 평균으로 산출되었다.

본 시뮬레이션은 AMD Ryzen 7 7800X3D CPU 및 NVIDIA GeForce RTX 4070 Ti GPU 기반 환경에서 MATLAB R2025a를 사용하여 수행되었다.

#### 2. 필터 설계

제안된 GEWKF의 성능 비교를 위한 표준 KF와 고정 EWKF의 설계는 다음과 같다. 표준 KF는 최적의 상태를 추정할 수

있는 공칭 잡음 공분산 값으로 설계하며, 고정 EWKF의 망각 인자  $\lambda$ 는 시뮬레이션을 통해 모든 시나리오에서 평균적으로 가장 낮은 평균 제곱근 오차(root mean squared error, RMSE)를 가지는  $\lambda = 1.05$ 를 적용하여 설계한다.

제안하는 GEWKF의 GRU 신경망은 MATLAB Deep Learning Toolbox를 이용하여 네트워크를 구현하였다. GRU 신경망의 학습 데이터셋은 필터가 다양한 형태의 불확실성에 효과적으로 적응할 수 있도록 총 2,000회의 몬테카를로 시뮬레이션을 통해 구축하였다. 각 시나리오에서는 망각 인자  $\lambda_k$ 의 동적 예측에 영향을 미치는 네 가지 핵심 불확실성 매개변수인 공정 잡음 공분산  $Q$ , 측정 잡음 공분산  $R$ , 시스템 행렬  $A$ , 측정 행렬  $C$ 를 무작위로 변동시켰다. 구체적으로, 잡음 공분산  $Q$ 와  $R$ 의 스케일링 인자는 균일 분포  $U[0.5, 3]$ 을 따르도록 설정하여, 기존 공분산 대비 50%에서 300% 범위 내에서 임의로 변화하도록 하였다. 또한, 모델 불확실성을 반영하기 위해, 시스템 행렬  $A$ 와 측정 행렬  $C$ 에 각각 균일 분포  $U[0, 1]$  기반의 섭동 계수를 적용하였다. 이때 섭동의 기본 크기는 시스템 행렬의 경우 0.03, 측정 행렬의 경우 0.005로 설정하였으며, 섭동 계수와 곱하여 다양한 수준의 모델 오차가 포함되도록 하였다. 이러한 불확실성 매개변수들의 변동은 각 시뮬레이션 내에서 무작위 시점에 발생하도록 설정하였으며, 변화의 지속 시간은 20~60 스텝 사이의 무작위 길이로 균일 분포를 따르도록 하였다. 결과적으로 구축된 학습 데이터셋은 다양한 잡음 강도, 모델 오차, 그리고 시간적 변동 패턴을 포함하여, GRU 신경망이 불확실성의 시변 특성에 대해 일반화된 망각 인자  $\lambda_k$ 를 학습할 수 있는 환경을 제공하였다.

각 시뮬레이션에서, 필터의 예측 오차와 측정 잔차를 기반으로 칼만 필터의 내부 상태를 나타내는 입력 벡터  $u_k$ 를 구성하였으며, 동시에 각 시점  $k$ 에서 실제 시스템의 불확실성을 반영하는 이상적인 망각 인자( $\lambda_k$ )를 오프라인으로 계산하여 정답 데이터로 활용하였다. 네트워크는 2개의 GRU 은닉층으로 구성되었고, 각 층은 78개의 유닛을 사용하였다. 훈련 과정에서는 과적합을 방지하고 네트워크의 일반화 성능을 검증하기 위해 데이터셋을 훈련, 검증, 테스트 세트로 나누어 활용

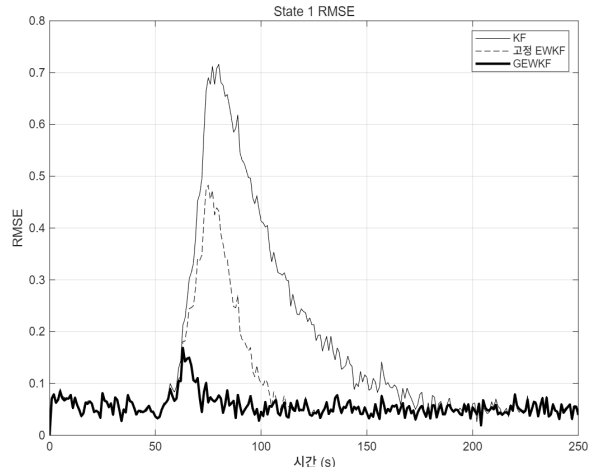


그림 3. 필터별 RMSE 비교(상태 1).

Fig. 3. Filter-specific RMSE comparison (State 1).

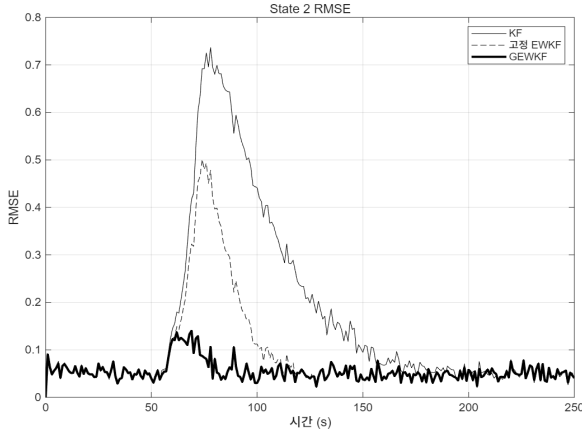


그림 4. 필터별 RMSE 비교(상태 2).

Fig. 4. Filter-specific RMSE comparison (State 2).

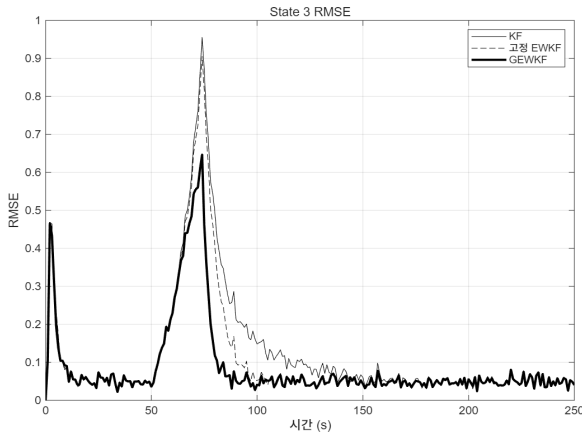


그림 5. 필터별 RMSE 비교(상태 3).

Fig. 5. Filter-specific RMSE comparison (State 3).

하였다. 네트워크는 Adam 옵티마이저를 사용하여 훈련되었으며, 초기 학습률은 0.01, 배치 크기는 256, 그리고 총 에폭(epoch)은 100으로 설정되었다.

### 3. 실험 결과 및 분석

본 시뮬레이션에서 각 필터의 상태별 추정 결과는 그림 3~5와 같다. 그림 3~5는 모델 부정확성 시나리오에서 KF, 고정 EWKF, 그리고 제안된 GEWKF의 시간에 따른 500회 앙상블 평균 상태 추정 오차(RMSE)를 보여준다.

또한, 표 1은 각 필터의 구간별 평균 RMSE를 수치로 보여주며, 이를 통해 필터의 성능을 정량적으로 비교할 수 있다. 그림 3~5와 표 1의 결과를 종합하여 상태 추정 오차를 분석해보면, 제안된 GEWKF가 시뮬레이션 전반에 걸쳐 가장 낮은 RMSE를 보이며 우수한 추정 성능을 나타내는 것을 확인할 수 있다.

먼저 불확실성 발생 이전인 10~49초 구간과 불확실성 종료 이후인 201~250초 구간을 살펴보면, 세 필터 모두 낮은 RMSE를 보이며 필터가 안정적으로 작동함을 보여준다. 이 구간에서는 모델 불확실성이 없으므로, 최적의 해를 제공하는 KF가 가장 좋은 결과를 보이며, 제안한 GEWKF가 동적 망각 인자 특성으로 고정 EWKF 대비 근소하게 좋은 성능을 보이는 것을 확인할 수 있다.

### 표 1. 필터별 구간 평균 RMSE 비교.

Table 1. Comparison of filter-specific sectional average RMSE.

| 상태   | 필터      | 10~49s        | 50~200s       | 201~250s      |
|------|---------|---------------|---------------|---------------|
| 상태 1 | KF      | 0.0549        | 0.2344        | 0.0514        |
|      | 고정 EWKF | 0.0558        | 0.1115        | 0.0503        |
|      | GEWKF   | <b>0.0552</b> | <b>0.0578</b> | <b>0.0503</b> |
| 상태 2 | KF      | 0.0523        | 0.2408        | 0.0501        |
|      | 고정 EWKF | 0.0525        | 0.1151        | 0.0498        |
|      | GEWKF   | <b>0.0524</b> | <b>0.0573</b> | <b>0.0500</b> |
| 상태 3 | KF      | 0.0523        | 0.1675        | 0.0497        |
|      | 고정 EWKF | 0.0528        | 0.1340        | 0.0494        |
|      | GEWKF   | <b>0.0527</b> | <b>0.1027</b> | <b>0.0494</b> |

그러나 불확실성 발생 구간인 50~200초 사이에서는 KF는 모델 부정확성을 보상할 수 없으므로 추정 오차가 가장 크게 발생한다. 고정 EWKF는 고정된 망각 인자를 사용하여 KF보다는 성능이 우수하지만, 동적으로 변화하는 불확실성에 완벽하게 적응하지 못한 결과를 보인다. 반면 GEWKF는 이 구간에서 고정 EWKF보다 훨씬 낮은 RMSE를 보인다. 이는 GEWKF가 불확실성을 실시간으로 감지하고, 적응적으로 망각 인자를 조절함으로써 오차를 효과적으로 억제하고 있음을 보여준다.

그림 6은 GEWKF가 추정한 망각 인자( $\lambda_k$ )의 평균적인 변화를 보여준다. 50초 이전인 정상상태에서는  $\lambda_k$  값이 1~1.18 사이의 안정적인 범위에 머문다. 그러나 모델 불확실성이 발생한 50초 이후,  $\lambda_k$  값이 2.4에 육박하는 높은 값으로 급증하는 것을 확인할 수 있다. 이는 GRU 네트워크가 불확실성을 정확하게 감지하고, 필터의 예측 오차 공분산을 크게 늘려 측정값에 대한 신뢰도를 높임으로써 오차를 빠르게 보정하고 있음을 명확하게 보여준다. 불확실성이 종료된 75초 이후에는  $\lambda_k$  값이 다시 1 근처의 범위로 안정화됨을 확인할 수 있다. 이와같이 제안한 GEWKF는 불확실성이 발생하는 환경에서 망각 인자  $\lambda_k$ 를 크게 늘려 추정 오차를 현저히 줄이며, 필터의 강인성을 크게 향상시키는 것을 확인할 수 있다.

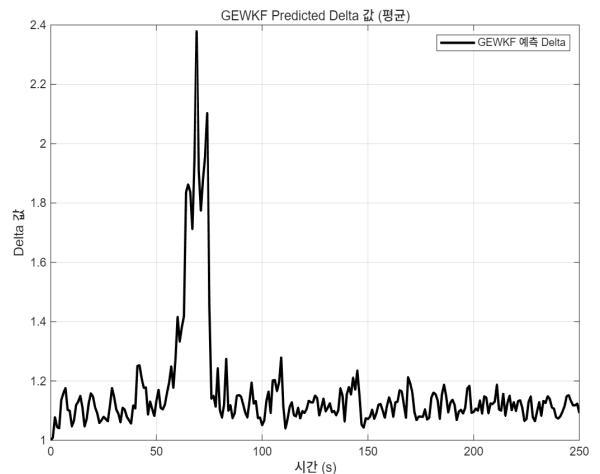
그림 6. GEWKF의 동적 망각 인자( $\lambda_k$ )의 변화.Fig. 6. Change of GEWKF's dynamic forgetting factor( $\lambda_k$ ).

표 2. GEWKF와 KalmanNet의 평균 RMSE 비교.

Table 2. Comparison of GEWKF &amp; KalmanNet average RMSE.

| 상태   | GEWKF RMSE | KalmanNet RMSE |
|------|------------|----------------|
| 상태 1 | 0.0522     | 0.0710         |
| 상태 2 | 0.0520     | 0.0675         |
| 상태 3 | 0.0523     | 0.4983         |

이 결과를 통해, 제안된 GEWKF가 정상 상태에서는 기존 필터와 대등한 성능을 유지하면서도, 모델 불확실성이 존재하는 비정상 환경에서는 표준 KF 및 최적으로 조정된 고정 EWKF보다 매우 우수한 적응성과 강인성을 제공함을 확인할 수 있다. 특히, 단일 고정 가중치는 모든 환경 변화에 최적으로 대응할 수 없는 반면, 제안된 GEWKF의 메커니즘은 실시간으로 필터 특성을 최적화하여 다양한 불확실성에 효과적으로 대처할 수 있음을 확인할 수 있다.

추가적으로, 제안된 GEWKF와 기존 RNN 기반 필터인 KalmanNet의 성능과 계산 복잡도를 비교하였다. 두 필터는 동일한 학습 데이터셋으로 학습되었으며, 성능 비교에 일시적 오차가 영향을 미치지 않도록 불확실성이 없는 구간으로 평가를 제한하였다. 상태 추정 성능은 각 상태별 평균 RMSE로, 계산 복잡도는 평균 추론 시간으로 비교하였다.

표 2에서 확인할 수 있듯이, 상태 1과 2에서는 두 필터 모두 안정적인 성능을 보였으나, KalmanNet은 각각 0.0710, 0.0675의 RMSE를 나타내어 제안된 GEWKF 대비 성능이 낮았다. 특히 상태 3의 경우, KalmanNet의 RMSE는 0.4983으로 제안 필터 대비 약 10배 높은 값을 보였다. 이는 제한된 학습 데이터 환경에서 KalmanNet이 시스템의 동적 특성을 충분히 학습하지 못한 결과로 해석된다. 평균 추론 시간 측정 결과, 제안된 GEWKF는 0.2624초로 KalmanNet의 0.3286초 대비 약 25% 빠른 속도를 보여, 보다 낮은 계산 복잡도를 가지는 것을 확인하였다. 결과적으로, 동일 학습 데이터셋 조건에서 제안된 GEWKF는 안정적인 성능을 유지하며, KalmanNet 대비 계산 복잡도가 낮아 빠른 추론 속도를 가지는 것을 확인할 수 있다.

#### IV. 결론

본 논문에서는 시스템 모델 불확실성과 잡음 특성 변화에 강인한 상태 추정을 위해, GRU 기반 순환 신경망을 활용하여 망각 인자를 동적으로 조정하는 적응형 칼만 필터인 GEWKF를 제안하였다. 제안된 GEWKF는 표준 KF의 재귀 구조를 유지하면서도, 시계열 데이터에 강한 GRU 신경망을 통해 시점별 망각 인자를 동적으로 학습하고 적용함으로써, 기존 고정된 가중치를 적용하던 고정 EWKF의 비적응성과 구조적 한계를 효과적으로 극복하였다. 특히, 단일 스칼라 형태의 파라미터만을 학습 대상으로 설정하여 계산 효율성과 해석 가능성을 확보하였으며, 수정된 ReLU 활성화 함수를 적용하여 표준 KF의 동작 조건까지 포함할 수 있는 일반화된 필터 구조를 제안하였다.

시스템 모델 부정확성 시나리오에 대한 시뮬레이션을 수행한 결과, 제안된 GEWKF는 정상 상태에서는 표준 KF 및 고정 EWKF와 유사한 추정 성능을 보였으며, 시스템 모델에 오차

가 존재하는 비정상 환경에서는 가장 낮은 RMSE를 보여 뛰어난 적응성과 추정 정확도를 확인하였다. 이는 GRU 신경망이 시스템의 불확실성 변화를 정확히 포착하고, 망각 인자를 적응적으로 증가시켜 최근 측정값의 신뢰도를 강화함으로써 오차를 효과적으로 보정했기 때문이다. 그리고 이 결과는 제안된 GEWKF가 기존 고정된 가중치 방식의 고정 EWKF가 가지는 비적응성과 성능 저하 문제를 효과적으로 해결했음을 보여준다. 또한, 기존 신경망 기반 KF 연구가 필터 이득이나 공분산 행렬 전체를 예측하여 높은 계산 복잡도와 해석상의 어려움을 가졌던 것과 달리, 본 연구는 망각 인자라는 단일 파라미터만을 학습함으로써 계산 효율성과 실시간 적용 가능성을 동시에 확보하였다는 점에서 의의를 가진다. 또한, 제안된 GEWKF는 동일 학습 데이터 조건에서 KalmanNet과 비교한 결과, 상태 추정 성능은 안정적으로 유지하면서 평균 추론 시간은 약 25% 빨라 계산 효율성 측면에서도 우수함을 확인하였다.

따라서 제안된 GEWKF는 계산 효율성과 해석 가능성을 확보하면서도 강인성과 적응성을 동시에 달성할 수 있는 실용적인 상태 추정 기법으로서, 불확실성이 존재하는 다양한 실제 시스템 환경에서 효과적으로 활용될 수 있을 것으로 기대된다.

#### REFERENCES

- [1] M. S. Grewal and A. P. Andrews, *Kalman Filtering: Theory and Practice Using MATLAB*, 4th ed., Wiley, 2014.
- [2] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Transactions of the ASME, Journal of Basic Engineering*, vol. 82, no. 1, pp. 35-45, 1960. doi : <https://doi.org/10.1115/1.3662552>
- [3] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*, Prentice-Hall, Englewood Cliffs, NJ, 1979.
- [4] D. Simon, "Kalman filtering with state constraints: a survey of linear and nonlinear algorithms," *IET Control Theory & Applications*, vol. 4, no. 8, pp. 1303-1318, 2010. doi: <https://doi.org/10.1049/iet-cta.2009.0032>
- [5] S. J. Julier and J. K. Uhlmann, "Unscented filtering and nonlinear estimation," *Proceedings of the IEEE*, vol. 92, no. 3, pp. 401-422, Mar. 2004. doi: <https://doi.org/10.1109/JPROC.2003.823141>
- [6] K. Na, Y. K. Kim, K. Choi and W. Kim, "A current-sensor error compensation algorithm for battery management systems based on sigma-point," *Journal of Institute of Control, Robotics and Systems (in Korean)*, vol. 30, no. 10, pp. 1131-1138, 2024. doi: <https://doi.org/10.5302/J.ICROS.2024.24.0095>
- [7] H. Suk and M. Won, "Docking control algorithm for autonomous mobile robot through sensor fusion," *Journal of Institute of Control, Robotics and Systems (in Korean)*, vol. 29, no. 9, pp. 717-725, 2023. doi: <https://doi.org/10.5302/J.ICROS.2023.23.0083>
- [8] S. Han, B. K. Kwon, and S. Han, "An extended finite impulse response filter for discrete-time nonlinear systems," *Journal of Institute of Control, Robotics and Systems*, vol. 21, no. 1, pp. 34-39, 2015.

doi: <https://doi.org/10.5302/J.ICROS.2015.14.0110>

- [9] B. K. Kwon, "An Exponentially weighted finite impulse response filter for discrete-time linear time-varying systems," *Journal of Institute of Control, Robotics and Systems (in Korean)*, vol. 26, no. 8, pp. 669-674, 2020.  
doi: <https://doi.org/10.5302/J.ICROS.2020.20.0088>
- [10] Q. Xia, M. Rao, Y. Ying and X. Shen, "Adaptive fading Kalman filter with an applications," *Automatica*, vol. 30, no. 8, pp. 1333-1338, 1994.  
doi: [https://doi.org/10.1016/0005-1098\(94\)90112-0](https://doi.org/10.1016/0005-1098(94)90112-0)
- [11] C. Hide, T. Moore, and M. Smith, "Adaptive Kalman filtering for low-cost INS/GPS," *Journal of Navigation*, vol. 56, no. 1, pp. 143-152, 2003.  
doi: <https://doi.org/10.1017/s0373463302002151>
- [12] K. H. Kim, G. I. Jee, C. G. Park, and J. G. Lee, "The stability analysis of the adaptive fading extended Kalman filter using the innovation covariance," *International Journal of Control, Automation and Systems*, vol. 7, no. 1, pp. 49-56, 2009.  
doi: <https://doi.org/10.1007/s12555-009-0107-x>
- [13] G. Revach, N. Shlezinger, X. Ni, A. L. Escoriza, R. J. G. van Sloun, and Y. C. Eldar, "KalmanNet: Neural network aided Kalman filtering for partially known dynamics," *IEEE Transactions on Signal Processing*, vol. 70, pp. 1532-1547, 2022.  
doi: <https://doi.org/10.1109/tsp.2022.3158588>
- [14] L. Xu and R. Niu, "EKFNet: Learning system noise covariance parameters for nonlinear tracking," *IEEE Transactions on Signal Processing*, vol. 72, pp. 3139-3152, 2024.  
doi: <https://doi.org/10.1109/tsp.2024.3417350>
- [15] X. Ni, G. Revach, and N. Shlezinger, "Adaptive KalmanNet: Data-driven Kalman filter with fast adaptation," *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Seoul, South Korea, Apr. 2024.  
doi: <https://doi.org/10.1109/ICASSP48485.2024.10447012>



### 장 선 호

2020년 강원대학교 제어계측공학과 (공학사). 2022년 고려대학교 전기전자공학부 (공학석사). 2023년~현재 동 대학 박사 과정. 2022년~현재 한국로봇융합연구원 주임연구원. 관심분야는 강화학습, 모방학습, 지능형 농업 로봇, 자율 주행 시스템.



### 권 소 연

2023년~현재 충남대학교 전자공학과 학사과정 재학 중. 관심분야는 자율형 지능 시스템, 지능형 영상처리, 선형시스템, 추정이론.



### 권 보 규

2002년 인하대학교 전자공학과 (공학사). 2005년 서울대학교 전기컴퓨터공학부 (공학석사). 2008년 동 대학 (공학박사). 2008년~2010년 삼성중공업 산업기술연구소 책임연구원. 2010년~현재 강원대학교 전자·AI시스템공학과 교수. 관심분야는 추정이론, 인공지능 기반 제어, 지능형 로보틱스.