

무향 적응성 칼만 필터를 이용한 데이터 동화 접근 기반 자세 추정 활성화 출력의 정제 방법

양진모¹, 유동영², 김영철³
홍익대학교 소프트웨어융합학과

¹ymd2222@g.hongik.ac.kr, ^{2,3}bob}@hongik.ac.kr

Refining Activation Outputs of Pose Estimation based on Data Assimilation with Unscented Adaptive Kalman Filter

Jinmo Yang¹, Dong-Young Yoo², and R. Young Chul Kim³
Department of Software and Communications Engineering

요약

최근 대규모 언어 모델(LLM)을 포함한 인공지능 기반의 서비스가 대두되고 있다. 하지만 전통적인 통계학적 방법이 수학이나 수치 데이터 처리에 있어서는 여전히 강점을 가진다. 본 연구는 무향 칼만 필터를 적응성 방식으로 사용해서 자세 추정(pose estimation)의 정확도 성능을 높이기 위한 방법을 제안한다. 특히 매 시점마다 자세 추정 모델의 추정 신뢰도 정보를 측정 공분산 행렬에 반영하여 필터의 적응성을 확보했다. 또한 이 방법은 자세 추정 모델의 중간 단계와 최종 단계에서 자세 추정의 불확실성을 감소시키고자 했다. 적용 사례로, 낙상 감지에서는 0.933의 F1-Score를 기록했다. 향후 적응성을 측정 외 기타 영역으로 확장시키고, 범용적인 적용을 위해 필터 기반의 복합 모델의 성능을 면밀히 검토하고자 한다.

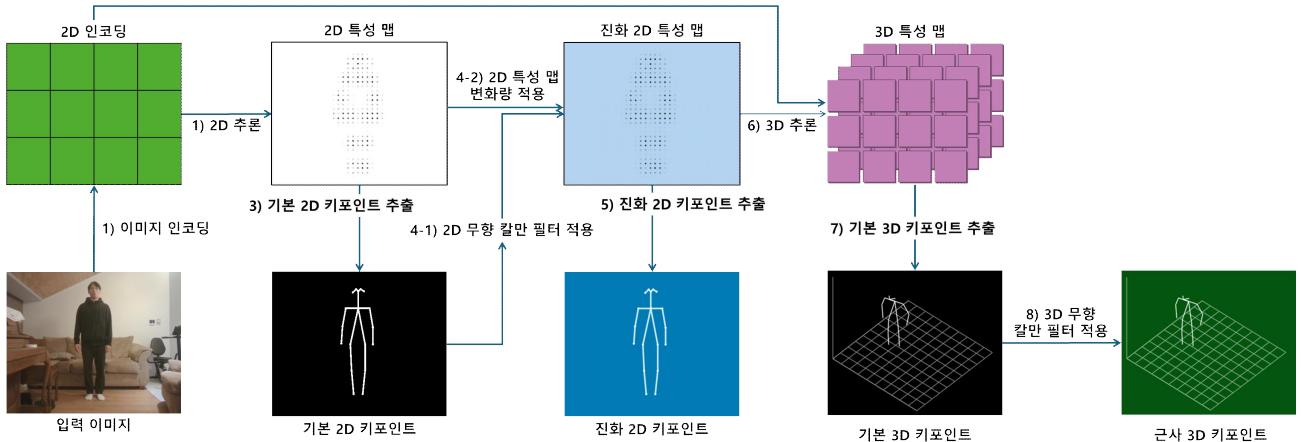
1. 서론

최근 인공지능에 대한 연구와 서비스가 활성화됨으로 인해 결과의 정확성이 화두가 되고 있다. 대규모 언어 모델(LLM) 기반 생성형 인공지능의 경우 결과를 개선하기 위해 프롬프트 엔지니어링을 적극적으로 활용한다[1]. 하지만 데이터 수준에서의 인공지능 활용에서 LLM은 시간적 측면에서 효율이 저조할 수 있으며, 특히 숫자의 경우 수학에 특화된 모델이 아닌 경우 환각현상이 쉽게 사라지지 않을 수 있다[2]. 따라서 이러한 데이터 정제에는 아직도 전통적(통계학적)인 방식이 사용되는데, 그 칼만 필터를 예로 들 수 있다[3]. 그리고 칼만 필터를 인공지능과 결합하여 결과를 개선하고자 하는 연구의 진행도 활발하다[4]. 본 논문은 사람의 관절 좌표를 감지하는 자세 추정(pose estimation) 모델에 칼만 필터 중 무향 칼만 필터에 매 상황에 대한 적응성을 더해 감지 성능을 향상시키는 방법을 제안한다. 2장에서는 관련 연구, 3장에서는 무향 적응성 칼만 필터 기반 자세 추정 방법, 4장에서는 낙상 감지 모델 적용 사례, 마지막으로는 결론을 언급한다.

2. 관련 연구

인공지능을 접목한 낙상 감지 모델은 종종 상황으로부터 핵심 특성만 추출하여 별도 예측 모델에 입력하여 수행한다[5]. 예를 들어, 자세 추정 기법이 이미지상 사람의 모습에서 각 관절 부위인 키포인트(landmark)의 좌표를 감지하고, 낙상 감지 모델이 키포인트로부터 낙상을 감지할 수 있다[6]. Chang 등(2021)은 자세 추정 모델을 통해 추출한 사람들의 자세 추정 좌표들을 시계열 데이터로 사용해 딥러닝 LSTM 모델의 입력값으로 사용했다[7]. 또한 넘어지는 사람을 추적(track)하기 위해 사람의 머리에 중심 점 트래커[8]를 사용했는데, 머리가 가려져 보이지 않을 수 있는 점은 고려하지 않았다.

칼만 필터는 잡음이 섞인 측정값으로부터 실제 상태값을 근사(estimate)하는 알고리즘이다[9]. 이는 기존 정보로부터 수행하는 예측(prediction) 단계와 측정이 이루어졌을 때 수행하는 갱신(update) 단계로 구성된다. 불확실성을 공분산 개념으로 모델링하는데, 필터 과정을 통해 최종적으로 공분산 값을 감소시켜 불확실성을 줄인다. Buizza 등(2020)은 자세 추정의 키포인트에 전통적인 칼만 필터를 적용하여



(그림 1) 무향 적응성 칼만 필터를 적용한 자세 추정 과정 도식화.

더 정확한 키포인트값을 추출하고자 했다[10]. 또한 칼만 필터를 적용했을 때 추정 속도를 30 FPS로 측정했는데, 이를 실시간 성능으로 간주할 수 있음을 강조했다. 하지만 측정 공분산 행렬을 하나의 벤치마크 데이터셋(PoseTrack 2018[11])으로 계산했는데, 실시간 측정은 이 벤치마크 데이터셋이 아닌 점을 간과하여, 실시간에 대한 적응성을 배제하였다.

3. 무향 적응성 칼만 필터 기반 자세 추정 방법

위 문제점들을 해결하기 위해 무향 적응성 칼만 필터를 낙상 감지 모델에 적용하는 방법을 제시한다. 이를 위해 자세 추정 모델 3D_Lightweight_OpenPose와 무향 칼만 필터를 사용한다[12, 13].

3D_Lightweight_OpenPose는 2D 이미지에서 2D 키포인트를 추정하고, 차원 확장(lift)을 통해 최종적으로 3D 키포인트를 추정한다[12]. 중간 결과와 최종 결과의 불확실성을 줄이기 위해 중간 필터 역할의 2D 필터와 최종 필터 역할의 3D 필터를 사용한다.

무향 칼만 필터는 비선형적 시스템에서 효과적인 근사를 위해 사용되는 필터의 일종이다[13]. 선행 연구로, 칼만 필터 없는 단순 낙상 감지 모델[14]과 전통적인 칼만 필터[15]가 있다. 하지만 기본 사람의 움직임과 낙상은 일부분 비선형적 움직임을 포함하고 있으므로 이를 근사할 수 있는 무향 칼만 필터의 사용이 더욱 적합하다.

머리가 보이지 않는 문제점은 기본 키포인트가 아닌 칼만 필터로 근사한 키포인트를 추적하는 것으로 완화할 수 있다. 추정이 이루어진 후 모든 좌표값들이 매 타임스텝마다 추정되지는 않아 개신은 할 수 없더라도, 예측 단계를 수행함으로써[16] 기존 불

확실한 정보로부터 근사를 진행할 수 있다. 즉, 보이지 않는 머리(머리뿐만 아니라 다른 관절 포함)의 키포인트를 공분산 추적을 통해 예측-근사 한다면, 이 근사값을 추적 가능하다.

또한 매 상황에 새롭게 필터가 적용해야 하는 측면에서는 매 상황에 적합한 정보가 제공되는 것으로 간소화될 수 있다. 이는 자세 추정 모델이 추정을 진행할 때 내포하는 불확실성을 이용하는 것으로 해결할 수 있다. 매 추정시 각 키포인트에 대한 신뢰값을 반환하는데[17], 이를 측정 공분산 행렬에 입력시키면 각 상황과 추정에 대한 필터의 적응성이 확보된다.

그림 1은 자세 추정에 무향 적응성 칼만 필터를 적용한 과정을 도식화한 것이다. 전체 과정을 다음과 같다.

Step 1) 이미지 인코딩. 입력 이미지를 자세 추정 모델에 입력하여 특성 추출(feature engineering)이 진행된 2D 인코딩 결과를 얻는다.

Step 2) 2D 추론(inference). 2D 인코딩으로부터 자세 추정 2D 특성 맵을 얻는다.

Step 3) 기본 2D 키포인트 추출. 2D 특성 맵으로부터 기본 2D 키포인트를 추출한다.

Step 4) 2D 특성 맵 진화(evolution). 추출한 기본 2D 키포인트에 2D 무향 칼만 필터를 적용하여 키포인트를 근사한다. 이 근사값과 기존값 사이의 변화량을 특성 맵에 적용하여 특성 맵을 또한 진화시킨다.

Substep 4-1) 2D 무향 칼만 필터 적용을 나타낸다. 무향 칼만 필터는 주어진 시스템의 상태를 잘 나타낼 수 있는 대표 시그마포인트(sigma points)를 선택하여 전이(예측)와 측정 함수에 통과시켜 평균 근사치와 공분산을 계산한다[12]. 함수가

반드시 선형일 필요가 없으므로, 전통적인 칼만 합수보다 비선형성이 포함된 사람의 움직임을 효과적으로 모델링할 수 있다.

Substep 4-2) 2D 특성 맵에 변화량을 적용한 것을 나타낸다. 다음의 식은 이를 수학적으로 나타낸다[13].

$$\begin{aligned}\vec{x}_{t,z} &= \vec{x}_t^+ - \vec{z}_t \\ H_t^+ &= H_t + X_{t,z}^{(H)} \\ P_t^+ &= P_t + X_{t,z}^{(P)}\end{aligned}\quad (1)$$

타임스텝 t 에서 2D 키포인트의 변화량 $\vec{x}_{t,z}$ 는 개신된 2D 키포인트 \vec{x}_t^+ 와 기본 2D 키포인트 \vec{z}_t 의 차로 나타난다. 특성 맵(히트맵 H 와 part affinity fields[15] P)에 대해서는 기본 2D 히트맵 H_t 와 P_t 에 키포인트 변화량 $\vec{x}_{t,z}$ 을 매핑($X_{t,z}^{(H)}$ 와 $X_{t,z}^{(P)}$)한 변화맵을 더해 진화한 특성 맵 H_t^+ 와 P_t^+ 를 얻을 수 있다.

Step 5) 진화 2D 키포인트 추출. 진화한 2D 특성 맵으로부터 진화 2D 키포인트를 추출한다. 이는 키포인트 추적에 사용될 수 있다.

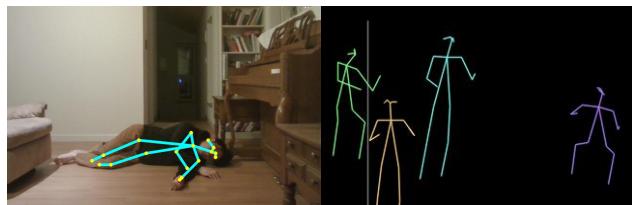
Step 6) 3D 추론. 2D 인코딩과 진화 2D 특성 맵으로부터 3D 특성 맵을 얻는다.

Step 7) 3D 키포인트 추출. 3D 특성 맵으로부터 3D 키포인트를 추출한다.

Step 8) 3D 무향 칼만 필터 적용. 3D 무향 칼만 필터를 적용하여 3D 키포인트를 근사한다.

4. 낙상 감지 모델 적용 사례

제안한 자세 추정 방법을 낙상 감지 모델[15]에 적용하여 모의 낙상과 실제 요양시설에서의 낙상 감지를 시도했다. 모의 낙상의 경우는 Fall detection dataset for SFDM[18]을 사용했으며, 0.933의 F1-Score를 기록했다. 요양병원에서의 낙상은, 실제 낙상사고가 발생하지 않아서 낙상이 검출되지 않았다. 그럼 2는 모의 환경과 요양시설에서의 적용을 보여준다.



(그림 2) a) 모의 환경과 b) 요양시설에서의 낙상 감지 적용 사례.

표 1은 모의 낙상에 대한 모델 성능 비교를 나타낸다. 기존 SFDM[14] 연구 0.925에서 0.933으로 성

능이 향상한 것을 확인할 수 있다.

(표 1) 다양한 낙상 감지 모델과의 성능 비교[14]

방법	정확도
PEFDM[19]	0.981
BP-2RF[20]	0.8999
Conv2D[21]	0.9983
Conv3D[22]	0.805
ST-GCN[23]	0.924
SFDM[14]	0.925
본 모델	0.933(마이크로-평균)

5. 결론

본 연구는 자세 추정의 성능을 높이기 위해 무향적응성 칼만 필터를 자세 추정의 활성 결과에 부분적으로 사용했다. 이를 실사용에 적용하기 위해 낙상 감지 모델에 적용했다. 비록 측정에 대한 적응성 필터를 사용했지만, 측정을 제외한 부분에서의 적합성을 검증해야 한다. 따라서 엄밀한 알고리즘의 검증을 위해 하이퍼파라미터 조정을 위한 반복 실험이 필요하다. 향후 3D_Lightweight_OpenPose에 대한 반복 실험을 진행하고, 다른 자세 추정 모델에도 일 반화할 수 있는지 확인할 예정이다.

Acknowledgment

본 연구는 2025년도 문화체육관광부의 재원으로 한국콘텐츠진흥원(과제명: 인공지능 기반 대화형 멀티모달 인터랙티브 스토리텔링 3D장면 저작 기술 개발, 과제번호: RS-2023-00227917, 기여율:100%) 지원과 한국연구재단의 4단계 두뇌한국21사업(과제명: 초분산 자율 컴퓨팅 서비스 기술 연구팀, 과제번호: 202003520005)의 지원을 받아 수행된 연구임.

참고문헌

- [1] 김유성, 김준삼, 이호용, 김상온. “퀵패스 AI 프롬프트 활용능력 2급”. 퀵패스. 2024.
- [2] B. Zhang, C. Li, and K. Fan, “MARIO Eval: Evaluate Your Math LLM with your Math LLM – A mathematical dataset evaluation toolkit,” arXiv, pp. 1–9, 2024.
- [3] C. Wang, H. Dai, W. Yang, and X. Yue, “High-efficiency unscented Kalman filter for multi-target trajectory estimation,” *Aerospace Science and Technology*, vol. 159, no. 109962, pp. 1–12, 2025.
- [4] N. Shlezinger *et al.*, “AI-Aided Kalman Filters,” *arXiv*, pp. 1–40, 2025.

- [5] A. Raza, M. Yousaf, W. Ahmad, S. Velastin, and S. Viriri, "Human fall detection using pose estimation: From traditional machine learning to vision transformers," *Engineering Applications of Artificial Intelligence*, vol. 143, no. 109809, pp. 1–18, 2025.
- [7] C. Zheng *et al.*, "Deep learning-based human pose estimation: A survey," arXiv, pp. 1–37, 2023.
- W. Chang, C. Hsu, and L. Chen, "A Pose Estimation-based Fall Detection Methodology Using Artificial Intelligence Edge Computing," *IEEE Access*, vol. 9, pp. 129965–119976, 2021.
- [8] J. Nascimento, A. Abrantes, and J. Marques, "An algorithm for centroid-based tracking of moving objects," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process*, Phoenix, AZ., USA, 1999, pp. 3305–3308.
- [9] J. Mochnac, S. Marchevsky, and P. Kocan, "Bayesian filtering techniques: Kalman and extended kalman filter basics," in *1009 19th Int. Conf. Radioelektronika*, pp. 119–122.
- [10] C. Buizza, T. Fischer, and Y. Demiris, "Real-time multi-person pose tracking using data assimilation," in *Proc. IEEE/CVF Winter Conf. on Appl. Of Computer Vision*, 2020, pp. 438–447.
- [11] M. Andriluka *et al.*, "Posetrack: A benchmark for human pose estimation and tracking," in *Proc. IEEE conf. on Computer Vision and Pattern Recognition*, 2018. Pp. 5167–5176.
- [12] D.Osokin, "Real-time 3d multi-person pose estimation demo," GitHub, Online, <https://github.com/Daniil-Osokin/lightweight-human-pose-estimation-3d-demo.pytorch>, Accessed April 10, 2025.
- [13] R. Labbe Jr., *Kalman and Bayesian Filters in Python*, GitHub, <https://github.com/rlabbe/Kalman-and-Bayesian-Filters-in-Python/tree/master>, Online, 2015. Accessed April 10, 2025.
- [14] J. Yang and R. Kim, "Simple Single-Person Fall Detection Model Using 3D Pose Estimation Mechanisms," *IEEE Access*, vol. 12, pp. 174640–174653, 2024.
- [15] J. Yang, Y. Jin, and R. Kim, "Multi-Person Fall Detection Using Data Assimilation Method with Kalman Filter," submitted for publication, 2025.
- [16] J. Hu, Z. Wang, H. Gao, and L. Stergioulas, "Extended Kalman filtering with stochastic nonlinearities and multiple missing measurements," *Automatica*, vol. 48, issue 9, pp. 2007–2015, 2012.
- [17] D. Mehta *et al.*, "Single-Shot Multi-Person 3D Pose Estimation From Monocular RGB," arXiv, pp. 1–16, 2018.
- [18] J. Yang, "Fall detection dataset for SFDM," *IEEE Dataport*, 2024.
- [19] W. Chang, C. Hsu, and L. Chen, "A pose estimation-based fall detection methodology using artificial intelligence edge computing," *IEEE Access*, vol. 9, pp. 129965–129976, 2021.
- [20] Y. Wang and T. Deng, "Enhancing elderly care: Efficient and reliable real-time fall detection algorithm," *Digit. HEALTH*, vol. 10, pp. 1–11, Feb. 2024.
- [21] L. Wu *et al.*, "Video-based fall detection using human pose and constrained generative adversarial network," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 34, no. 4, pp. 2179–2194, Apr. 2024.
- [22] A. Osigbesan *et al.*, "Vision-based fall detection in aircraft maintenance environment with pose estimation," in *Proc. IEEE Int. Conf. Multisensor Fusion Integr. Intell. Syst. (MFI)*, Sep. 2022, pp. 1–6.
- [23] C. Zampino *et al.*, "Fall detection using NAO robot pose estimation in RoboCup SPL matches," in *Proc. 9th Italian Workshop Artif. Intell. Robotics (AIRO)*, Udine, Italy, 2022. pp. 88–95.