

Robust Target Tracking using TDOA Probabilistic Image

이성호^a, 심재웅^{*}

서울과학기술대학교, 데이터사이언스학과^{*,a}

sean0310@seoultech.ac.kr^a, jaewoong@seoultech.ac.kr^{*}

Index

1. Introduction

- 1) Research background
- 2) Research motivation
- 3) Research purpose

2. Related Work

- 1) Indoor tracking
- 2) Convolutional-LSTM

3. Proposed Model

- 1) Model structure overview
- 2) TDOA to image conversion
- 3) Time series conversion & tracking

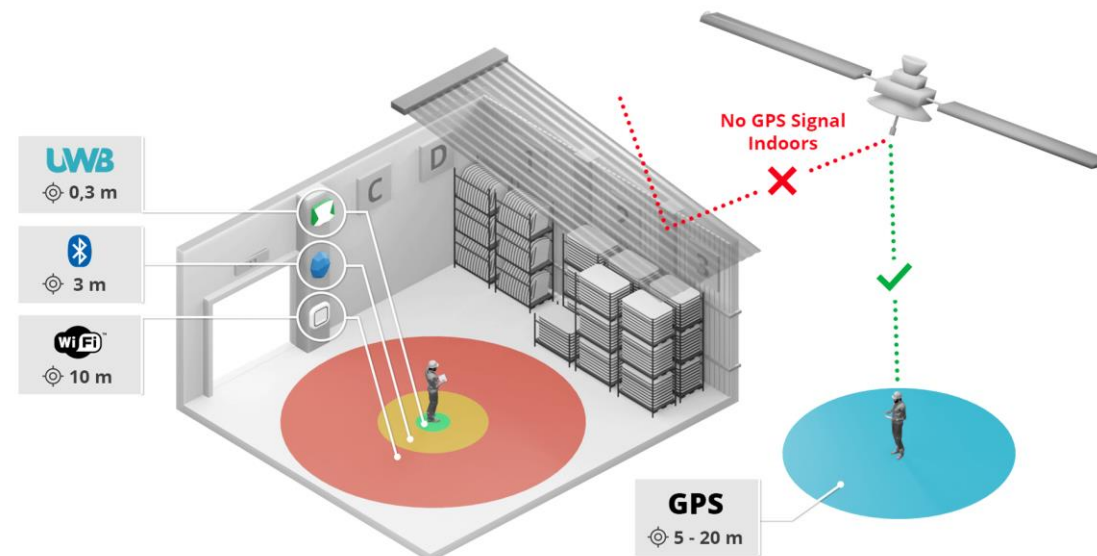
4. Experiment

- 1) Environment setting
- 2) Data generation
- 3) Model hyperparameter
- 4) Model comparison

5. Conclusion



- 소위 GPS라 불리는 글로벌 네비게이션 위성 시스템(Global Navigation Satellite System)은 인공위성을 사용해 위치를 추적하는 기술[1]
- 하지만 거리 오차 단위가 크며(수십 미터), 실내 혹은 지하 공간에서는 사용이 불가[2]
- 이러한 단점 때문에 실내에서 물체의 위치를 파악하기 위해 UWB, Wi-Fi, Bluetooth 등의 대체 기술들이 개발
- 그 중 다음과 같은 장점 때문에 최근 UWB 기술이 주로 사용됨[3-5]
 - ① 수십 cm 이내의 측위 정확성
 - ② Multipath 등의 간섭에 대한 저항성
 - ③ PHY layer와 난수를 활용한 보안성



- UWB : 짧은 시간의 낮은 출력의 펄스 신호를 사용하여 500MHz 이상의 넓은 주파수 대역으로 데이터를 송수신하는 100m 이내의 무선통신 기술
- 1970년대 군사 목적으로 개발되었으며 2000년대 이후 802.15.4a라는 이름으로 IEEE에서 발표
- 현재 다양한 분야에서 이와 같이 적용[6]
 - ① 실내 측위 – 편의점, 공항, 물류 창고, 스마트 팩토리
 - ② 스마트폰과 적용 – 에어드롭(파일 공유), 스마트싱스파인드(다른 디바이스 추적), 이즈렌(스마트홈)
 - ③ 보안 접속 – 스마트 키 및 원격 조정, 전자 결제
 - ④ 스포츠 – 운동량 측정, 장면 자동 추적 카메라
 - ⑤ 비 접촉 의료 모니터링



- 하지만 매점, 헬스장 같이 장애물이 많은 공간에서는 UWB 신호가 Additive White Gaussian Noise(AWGN)에 의해 변형되어 위치 추적 성능이 급격히 감소[7]
- AWGN에 대해 강건한 위치 추적 모델을 만들기 위해 딥러닝과 머신 러닝 기법을 활용한 연구 진행[8-15]
- 사전 연구들은 신호를 분해하여 AWGN이 제거된 깨끗한 신호를 얻고자 함
- 하지만 다음과 같은 단점을 가지고 있음
 - ① 많은 보정 프로세스를 거침
 - ▶ 실시간 위치 측위에 적합하지 않음
 - ② 실험 환경을 사전에 정의
 - ▶ 실험 환경 변화에 따른 성능 저하 극심



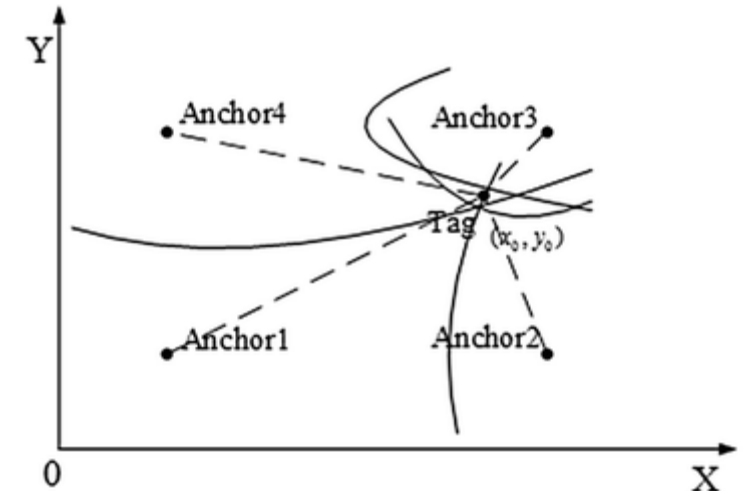
1

• 연구 목적

- 다양한 공간에서 발생하는 AWGN에 대해 범용적으로 사용 가능한 **강건한 모델 개발**
- 측정된 TDOA를 그대로 사용하는 **간단한 프로세스의 end-to-end 모델 개발**
 - ① 많은 보정 프로세스를 거침
 - ② 실험 환경을 사전에 정의
- 이동하는 물체를 실시간으로 추적하는 **시계열 모델 개발**



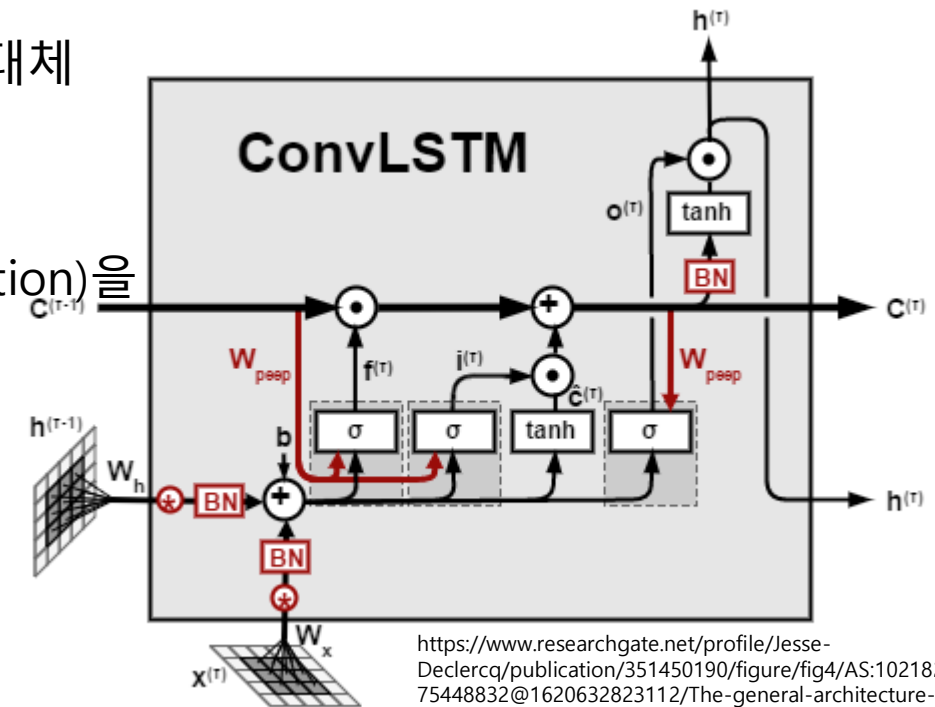
- UWB를 활용해 위치를 추정하는 방식은 time of arrival (ToA) 방법 존재[4]
- ToA는 송신기에서 빛을 쏘아 수신기에서 반사되어 돌아오는 시간을 측정한 뒤, 빛의 속도를 곱하여 거리를 구하는 방식
- ToA 방식은 태그(송신기)의 주소 바인딩이 필요해 통신할 수 있는 기기의 개수가 한정되어 있으며, 장치끼리 여러 번의 통신을 하기에 배터리 소모가 커 주로 측정된 신호 시간의 차이를 이용하는 TDOA(Time Difference of Arrival) 방식을 주로 사용[16]
- TDOA를 사용한 위치 추정 시스템의 순서는 다음과 같음
 - ① UWB 태그가 메시지를 송신
 - ② 주변의 UWB 앵커가 메시지를 수신
 - ③ 중앙 컴퓨팅 유닛(CCU)은 전체 UWB 앵커들의 데이터를 수집하여 각 앵커들 간 신호 도달시간 차이(TDOA)를 계산
 - ④ TDOA를 이용해 쌍곡선 방정식을 도출
 - ⑤ 방정식들의 해를 계산하여 UWB 태그의 위치를 추정



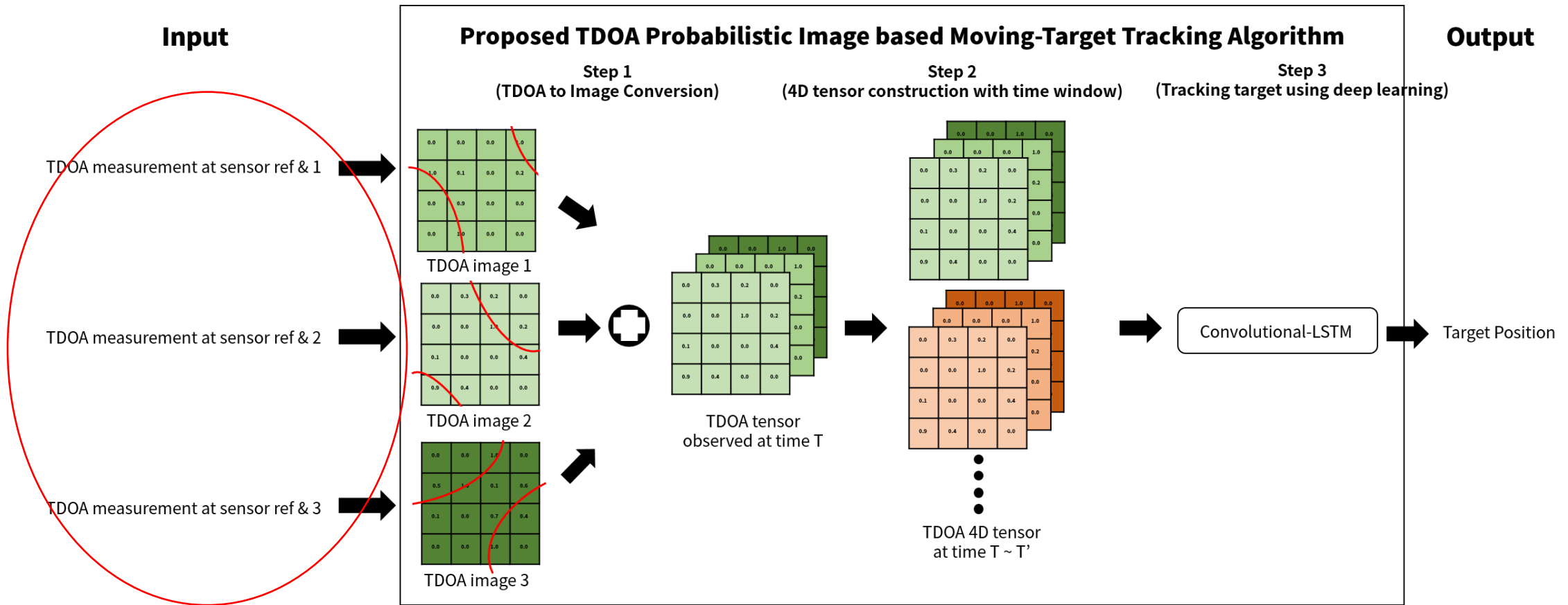
- 시계열 기반 TDOA 보정 사전 연구
 - GCC-PHAT을 사용해서 steering vector로 분리한 UWB 신호를 입력 값으로 하는 BLSTM을 사용하여 AWGN이 제거된 TDOA를 구함[11]
 - UWB 신호에서 cross correlation을 이용한 관측 값에서 신호들 간 분리 작업을 통해 TDOA를 구한 뒤, LSTM 모델을 사용하여 위치 추정[12]
 - LSTM 모델을 이용해 UWB 신호에 포함된 AWGN을 제거 후, weighted KNN 모델로 타겟의 위치 추정 [13]
- Convolution 기반 TDOA 보정 사전 연구
 - UWB channel impulse response 그래프를 입력 데이터로 하여 CNN을 통해 물체의 위치 추정[14]
 - 웨이블릿 변환(Wavelet Transform)을 통해 신호들을 이미지로 변환한 뒤, CNN을 통해 AWGN이 제거된 TDOA를 구함[15]

- 앞선 연구들의 보정 과정은 다음과 같은 문제를 지님
 - 여러 연산과 파이프 라인을 거침 ► 실시간 위치 측위에 적합하지 않음
 - 환경이 복잡할수록 신호 변형이 기하급수적으로 늘어남 ► 물체의 움직임과 관측 환경을 모두 고려해 정확한 TDOA를 보정할 수 있는 알고리즘 구현하기엔 너무 복잡함
- 해당 문제를 해결하고자 'TDOA image based target tracking(TITT)' 모델[17]이 등장
 - 관측한 TDOA를 그대로 사용해 이미지로 변환하여 물체의 위치 추정
 - 이미지 변환 과정은 무선 통신 분야에서 noise를 제거하기 위해 사용하는 'Time-Frequency Masking' 처럼 AWGN을 제거하는 과정에서 착안[18]
 - 간단한 프로세스 구조를 통해 빠른 연산 속도를 보였으며 AWGN이 많은 환경에서 안정적인 예측을 보임(예측 값의 분산이 작음)
 - 하지만 AWGN이 많이 발생하는 공간에 대해서만 한정적인 결과
 - 오히려 AWGN이 없는 환경에서는 물체 추적 성능이 다른 방법들에 비해 떨어짐

- Convolutional LSTM은 기존 LSTM 모델이 공간적 특성을 사용하지 못한다는 단점을 보완한 모델[19]
- 기존의 FC-LSTM 모델을 개선한 부분은 다음과 같음
 - 입력-망각-출력 게이트(i, f, o)와 셀 입력, 출력 상태(X, C, H)가 3차원 벡터로 계산
 - ▶ 3차원 채널 데이터의 손실없이 시퀀스 데이터의 학습이 가능
 - FC-LSTM에서 행렬곱이 사용되는 부분은 모두 합성곱으로 대체
 - ▶ 각 셀에서 학습할 weight의 수가 FC-LSTM보다 줄어듦
- 개선 과정을 통해 시간적, 공간적 상관성(Spatiotemporal correlation)을 동시에 학습 가능함

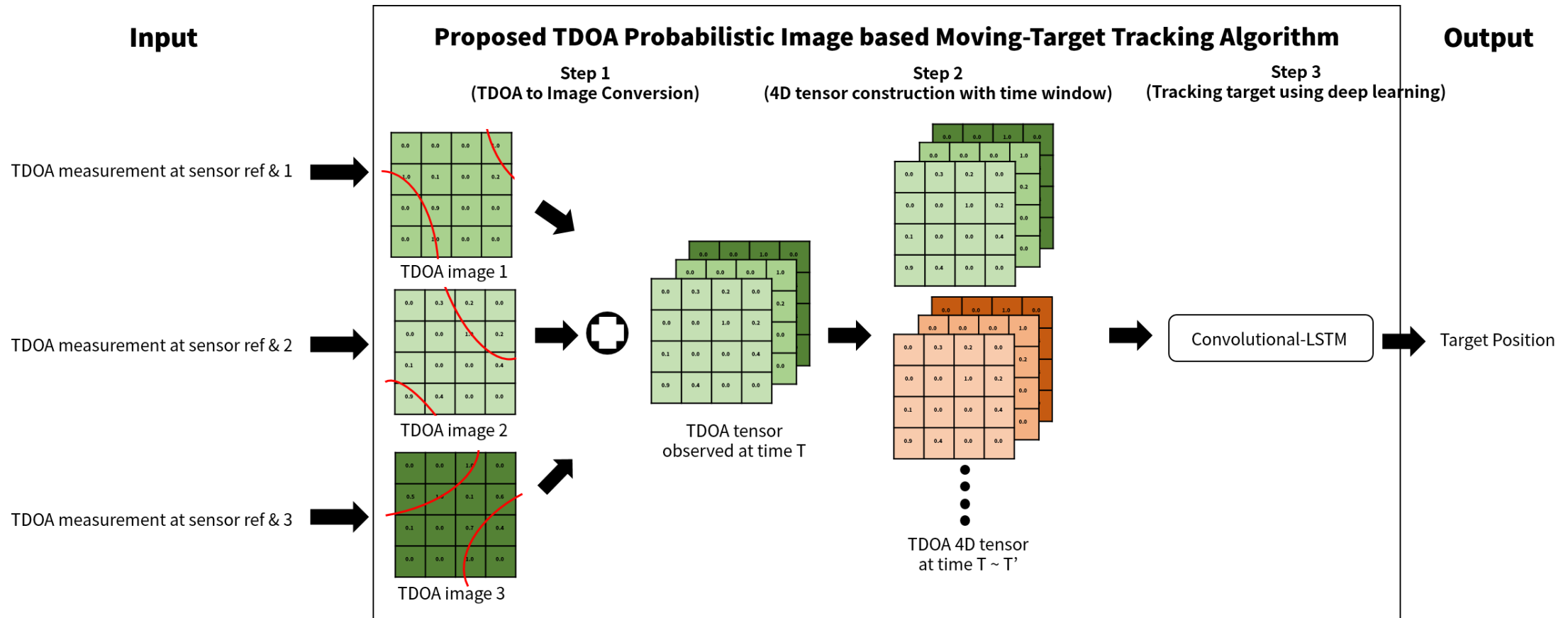


Proposed model – Model structure overview1



- 4개의 UWB 앵커와 1개의 태그를 사용해 얻은 TDOA를 입력 값으로 설정
- 입력 값 3개는 물체의 위치를 추정하기에 최소한으로 필요한 TDOA 개수

Proposed model – Model structure overview2

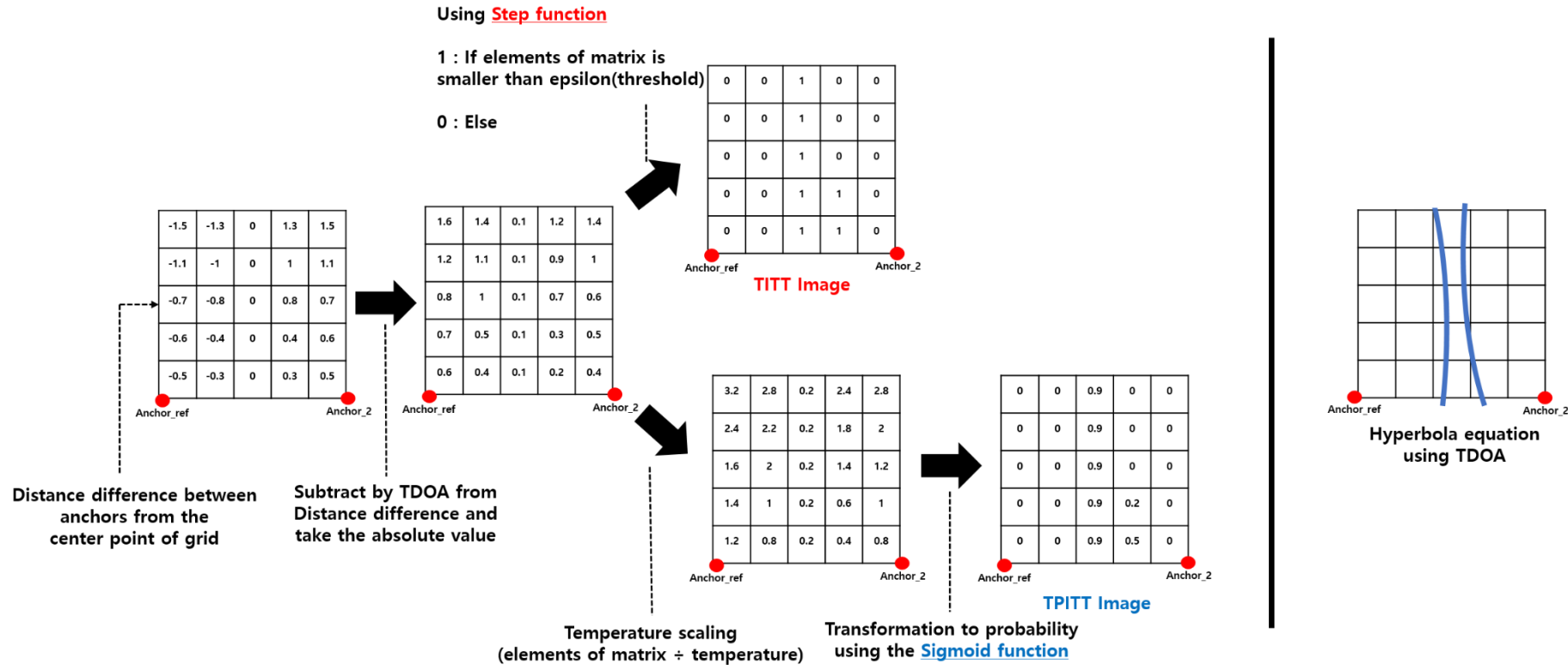


- TPITT(TDOA Probabilistic Image based moving-Target Tracking)를 통해 물체의 공간 좌표 벡터 추정
- TPITT 모델은 'TDOA 이미지 시계열 데이터 변환 뒤 conv-lstm을 통한 물체의 위치 추적'으로 표현 가능

Proposed model – TDOA to image conversion1

Estimate process

Real Process

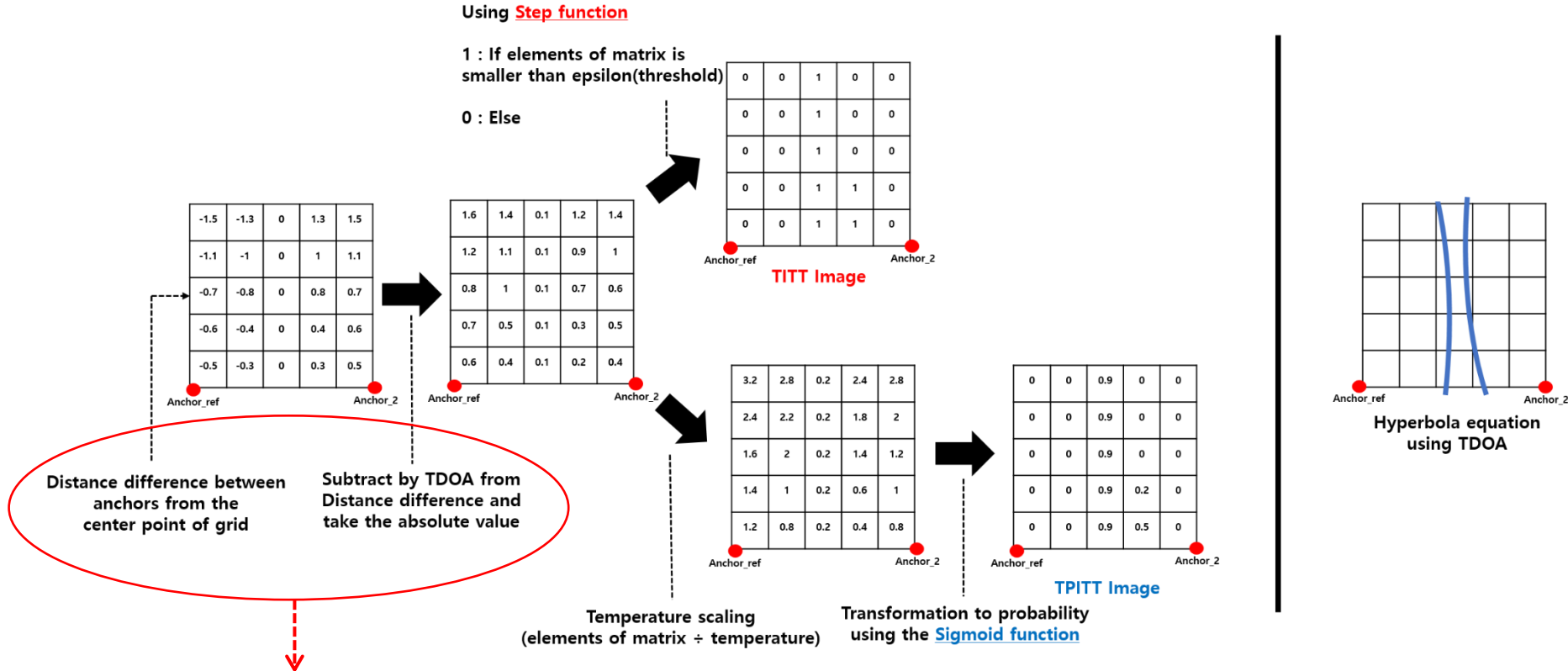


- 이미지 변환 과정을 위해 앵커 2개의 공간 좌표와 TDOA, 이미지 크기, 격자 크기, 최소 거리 차이 상수(ϵ), temperature(T) 하이퍼파라미터 필요
- ϵ : TDOA와 앵커와 격자 중점의 거리 차이 유사 비교 상수 / T : sigmoid 함수의 출력 값을 보정하는 양수

Proposed model – TDOA to image conversion2

Estimate process

Real Process



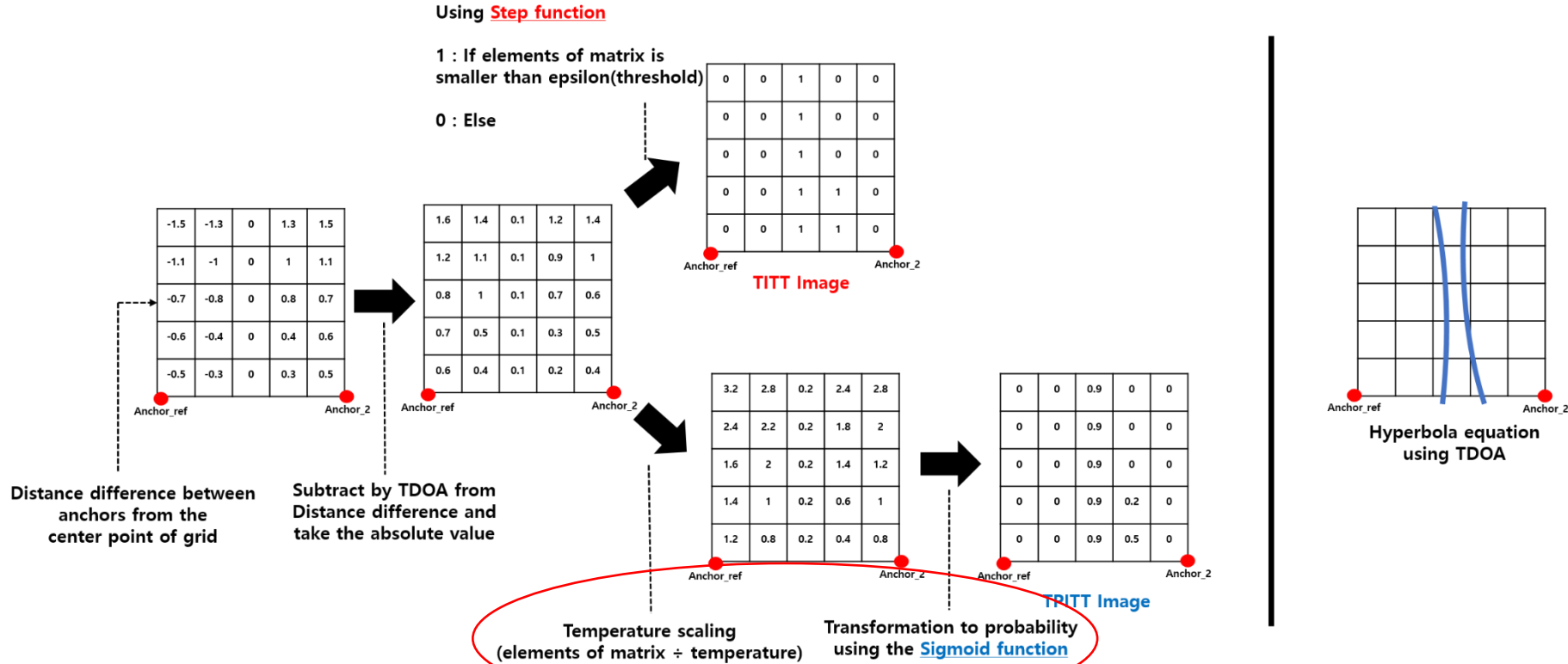
① 행렬의 원소가 격자의 중점과 기준 앵커(anchor_ref)의 거리와 격자의 중점과 나머지 앵커(anchor_n)의 거리 차이의 절대값인 행렬 M 생성

② sigmoid 함수를 이용해 행렬 M의 원소 smoothing 진행(이때, temperature 상수 T를 활용해 smoothing 정도를 조절 ($M_{i,j} = 1/1 + e^{\frac{|(C_{i,j}-A)-(C_{i,j}-A')-s|-\varepsilon}{T}}$))

Proposed model – TDOA to image conversion³

Estimate process

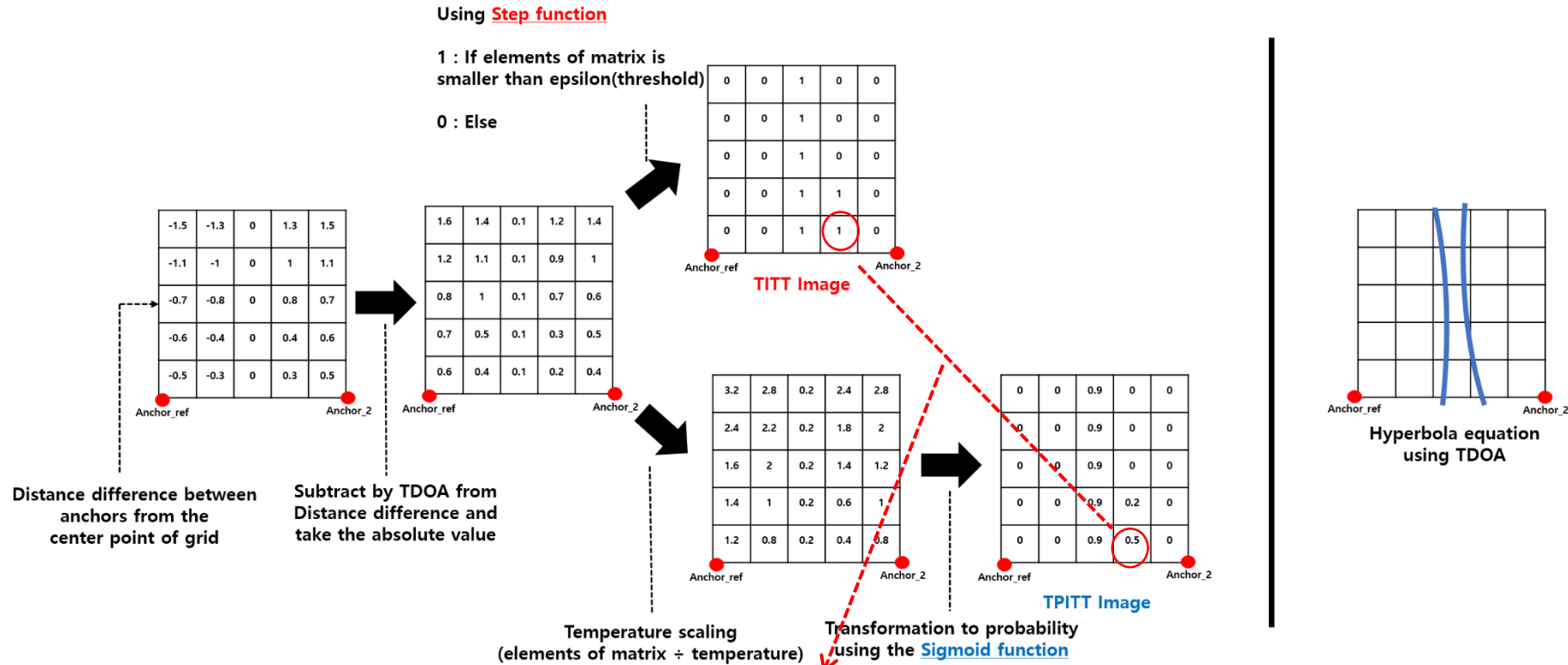
Real Process



- ① 행렬의 원소가 격자의 중점과 기준 앵커(anchor_ref)의 거리와 격자의 중점과 나머지 앵커(anchor_n)의 거리 차이의 절대값인 행렬 M 생성
- ② sigmoid 함수를 이용해 행렬 M의 원소 smoothing 진행(이때, temperature 상수 T를 활용해 smoothing 정도를 조절 ($M_{i,j} = 1/1 + e^{\frac{|(C_{i,j}-A)-(C_{i,j}-A')-s|-\epsilon}{T}}$))

Estimate process

Real Process

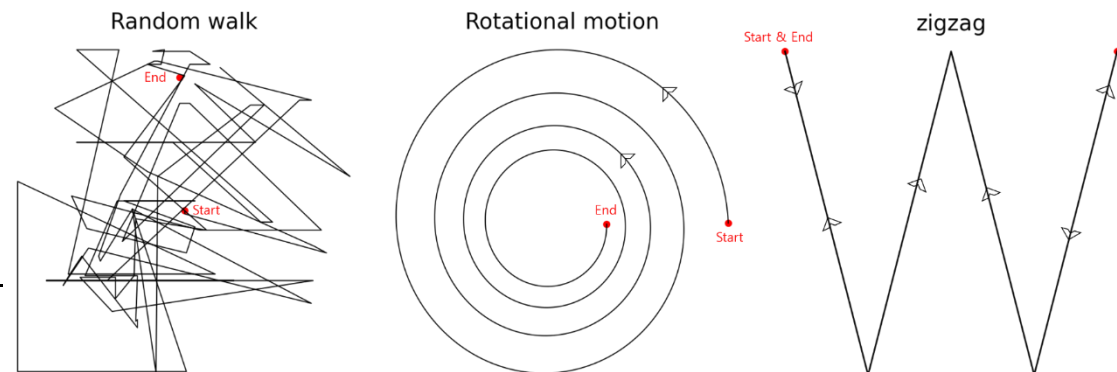


- epsilon을 threshold로 하는 구간 함수인 TITT 모델과 달리, TPITT 모델은 sigmoid 함수를 이용해 연속 함수로 표현하는 과정
- 이미지 변환 과정을 통해 실제 쌍곡선 방정식 그래프와 유사한 이미지가 생성됨

- 3차원 형태의 데이터를 Convolutional-LSTM에서 Convolutional filter의 입력 값으로 사용하기 위해 4차원의 시계열 데이터 형태로 변환하는 과정이 필요
- 본 실험에서는 window를 4로 설정
- Convolutional-LSTM 모델을 사용하여 물체의 공간 좌표에 대해 regression
- loss 함수로 Mean Square Error(MSE)를 사용
- $MSE = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2$ (N : Dataset의 크기, y : 물체의 실제 공간 좌표, \hat{y} : 예측한 물체의 공간 좌표)

- 제안한 모델을 검증하기 위해 컴퓨터 시뮬레이션을 진행하며 환경 정의는 다음과 같음
 - 단일 물체에 대해 위치를 추정하기 위해 크기가 **10m * 10m의 정사각형 공간**을 설정
 - 각 꼭지점에 4개의 센서가 위치하고 물체의 움직임을 추적하고 있으며, 그 중 하나가 기준 센서로 선택된다고 가정
 - 채널 모델은 IEEE 802.15.4a의 사무실 내 환경을 가정하며 “Bitcraze”사의 “loco positioning system” 모듈을 이용해 가상 수신 신호 생성
 - 각 센서는 TOA를 기반으로, TDOA는 기준 센서와 나머지 센서 간의 TOA 차이로 중앙 컴퓨팅 유닛 (CCU)에서 계산
 - TDOA 측정은 millisecond 단위로 이루어지며 각 센서와 CCU 사이의 통신 지연은 없다고 가정
 - **다양한 환경을 가정**하기 위해 환경에 따라 **각각 다른 측정 AWGN을 적용**
 - ① **AWGN이 극심한 환경(Noise extreme) : 평균 0, 표준 편차 2**
 - ② **AWGN이 높은 환경(Noise high) : 평균 0, 표준 편차 1**
 - ③ **AWGN이 적은 환경(Noise low) : 평균 0, 표준 편차 0.2**

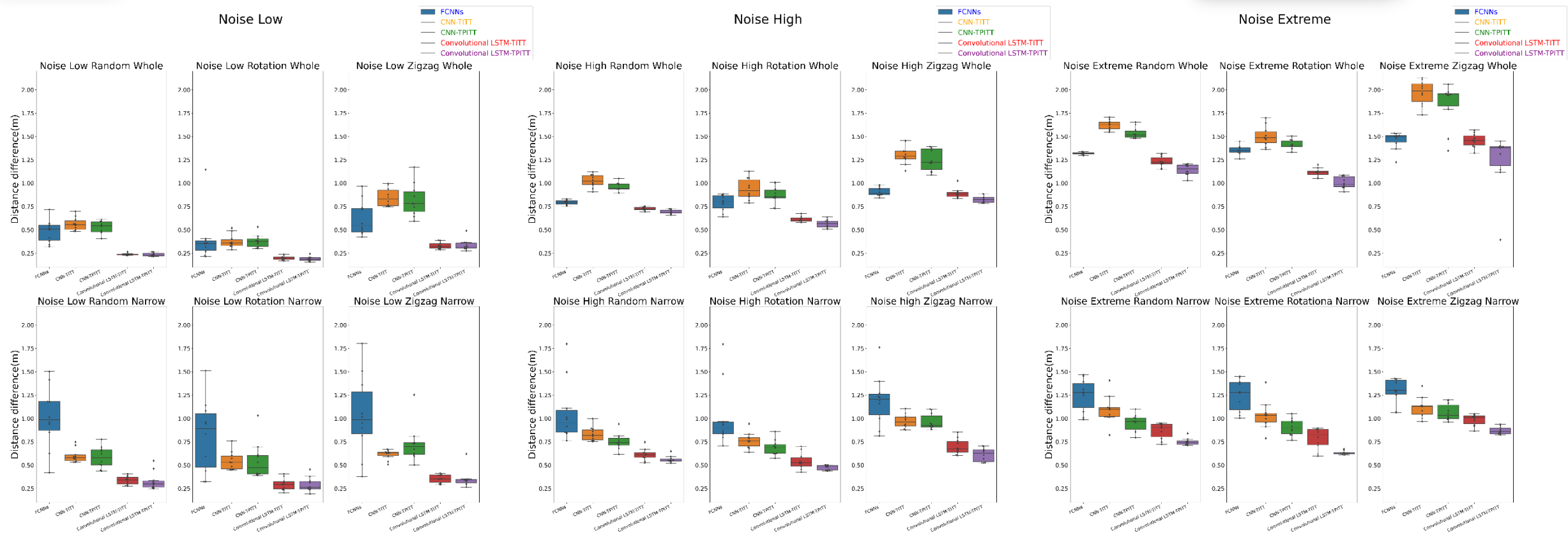
- 랜덤으로 운동하는 물체의 데이터 16,000개를 생성해 학습용으로 사용
- 물체는 가속도 운동을 진행하며 물체의 가속도는 표준정규분포를 통해 매초 변경
- 최대 속도는 인간의 평균 달리기 속도인 5m/s으로 가정
- 다양한 물체에 대해 검증하기 위해 검증용 데이터 생성 시 물체가 지그재그(zigzag), 랜덤(random walk), 회전 운동(Rotational motion)을 하며 전체 공간 혹은 기준 앵커 주변의 좁은 공간(기준 앵커 반경 2m)에서만 운동한다 가정
- 즉, 2가지의 제한된 공간(전체, 좁은)에서 3가지의 운동(랜덤, 지그재그, 회전)을 하는 검증용 데이터 각 4,000개 생성 후 6번의 경우에 대해 검증 실험 진행



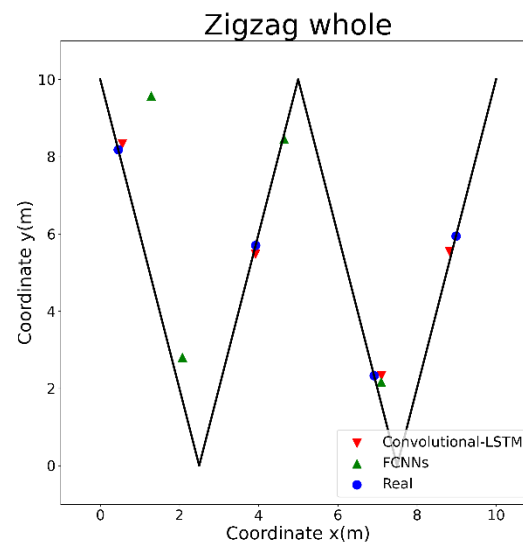
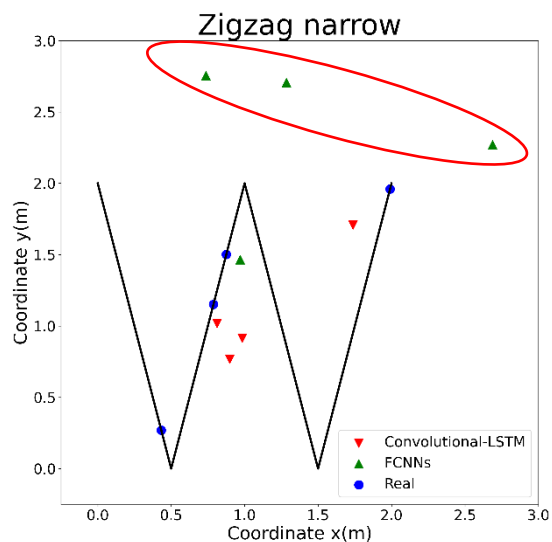
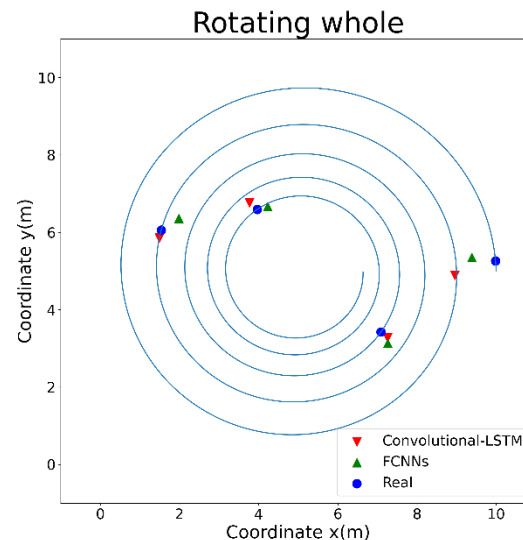
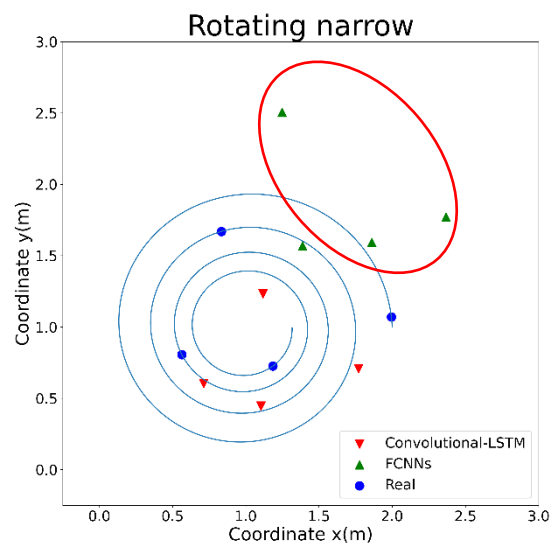
- 비교군 모델 4개(FCNNs, TPITT CNN, TITT CNN, TITT Convolutional LSTM) 설정
- Convolutional LSTM 모델의 구조는 3번의 convolution(Channel&Kernal: 4&11 / 8&8 / 16&4)과 max pooling(Kernal:2)를 거친 다음 batch normalization을 한 후, 3번의 dense layer(128 / 32 / 2)를 거친 구조
- CNN 모델의 구조는 Convolutional LSTM 층이 일반 Convolution 층으로 바뀐 것을 제외하고 모든 구조가 동일
- FCNNs 모델은 input이 TDOA값 3개이며 hidden layer는 16, 8 이고 output layer는 2(물체의 공간 좌표 벡터)인 구조
- 다른 하이퍼 파라미터는 표와 똑같이 모두 적용

Parameter	Value
Activation Function	Relu
Optimizer	Adam
Learning rate	1×10^{-3}
Loss function	MSE
Drop out	0.6
Mini batch size	20
Epoch	30
격자 크기	0.1m
Epsilon	0.05
Temperature	0.1

Experiment – Model comparison1

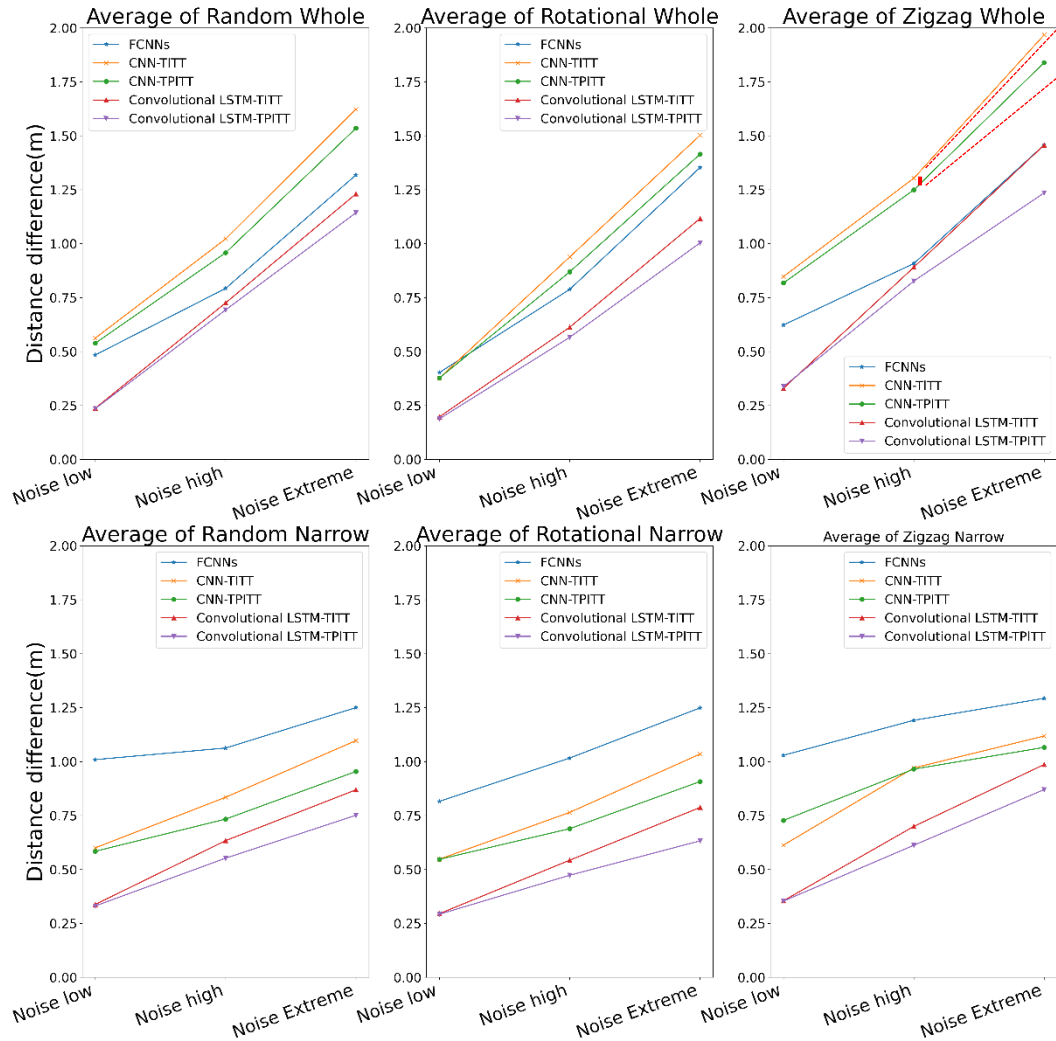


- 10번의 반복 실험을 진행 후 실험 결과를 boxplot & swarm plot으로 표현
- Convolutional –LSTM 기반 모델이 모든 상황에서 평균적으로 가장 좋은 예측을 하는 결과를 보임
- 랜덤으로 움직이는 물체를 학습하면서 다른 움직임을 보이는 물체 또한 운동성을 파악하여 물체의 움직임을 잘 예측할 수 있었던 것으로 추정



- Convolutional-LSTM과 FCNNs 모델이 예측한 결과를 좌표평면에 시각화
- 전체 공간에서 물체가 운동(그림 우측)에 비해 앵커 주변의 공간에서 운동 (그림 좌측)하는 상황일 경우 **FCNNs 모델의 성능 저하가 심함**
- 이미지 기반의 모델은 숫자 기반의 모델과 다르게 **보편적인 상황만 학습을 진행**하여도 다양한 상황에 대해 **섬세하고 강건한 예측 결과**를 보임

Mean change graph according to AWGN change



- AWGN의 변화와 모델 종류에 따른 실험 결과의 평균 변화 그래프
- 모든 상황에서 환경의 AWGN가 증가할수록 TITT와 TPITT의 성능 차이가 커지는 경향을 보임
- TPITT의 확률화 과정은 이미지 변환 방식의 주요 목적인 AWGN를 제거하는 역할을 더 잘 수행

- 연구 요약
 - AWGN이 많은 다양한 환경, 상황, 물체에 대해 보편적으로 사용할 수 있는 위치 추적 방법 제안
- 연구 의의
 - 왜 이미지 기반 모델인가 : TDOA를 그대로 사용하는 모델보다 일반적이지 않는 상황에서 강건한 성능을 보임
 - 왜 TPITT 방법인가 : AWGN이 많은 환경일수록 다른 방법들 보다 AWGN을 효과적으로 제거하여 예측 오차가 작음
 - 왜 Convolutional-LSTM인가 : 물체의 운동성을 학습하여 상황, 환경 상관 없이 움직이는 물체를 추적할 때 좋은 성능을 보임
- 한계 및 추후 연구
 - 본 연구는 실제 실험이 아닌 컴퓨터 시뮬레이션
 - 장애물로 인해 실내 위치 추적이 어려운 상황에서 TPITT 모델을 활용해 위치 추정이 안정적으로 진행되며 이러한 기술을 바탕으로 다양한 환경에서 위치 추적 서비스가 개발될 것이라 예상

- [1] Wells, D., Beck, N., Kleusberg, A., Krakiwsky, E. J., Lachapelle, G., Langley, R. B., ... & Delikaraoglou, D. (1987). Guide to GPS positioning. In Canadian GPS Assoc.
- [2] Rizos, C. (2013, March). Locata: A positioning system for indoor and outdoor applications where GNSS does not work. In Proceedings of the 18th Association of Public Authority Surveyors Conference (pp. 73-83).
- [3] Fontana, R. J., & Richley, E. A. (2007, September). Observations on low data rate, short pulse UWB systems. In 2007 IEEE International Conference on Ultra-Wideband (pp. 334-338). IEEE.
- [4] Coppens, D., De Poorter, E., Shahid, A., Lemey, S., & Marshall, C. (2022). An Overview of Ultra-WideBand (UWB) Standards (IEEE 802.15. 4, FiRa, Apple): Interoperability Aspects and Future Research Directions. arXiv preprint arXiv:2202.02190.
- [5] Kuvshinov, A., Knobloch, D., Külzer, D., Vardanyan, E., & Günnemann, S. (2022, June). Domain Reconstruction for UWB Car Key Localization Using Generative Adversarial Networks. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 36, No. 11, pp. 12552-12558).
- [6] 장병준. (2022). UWB 측위기술의 원리 및 동향. 한국전자과학회논문지, 33(1), 1-11.
- [7] J. Baek, J. Lee, H. Shim, S. Im and Y. Han, "Target tracking initiation for multi-static multi-frequency PCL system", IEEE Transactions on Vehicular Technology, vol. 69, no. 10, pp. 10 558-10 568, 2020.
- [8] Krishnan, S., Santos, R. X. M., Yap, E. R., & Zin, M. T. (2018, November). Improving UWB based indoor positioning in industrial environments through machine learning. In 2018 15th International Conference on Control, Automation, Robotics and Vision (ICARCV) (pp. 1484-1488). IEEE.

- [9] Musa, A., Nugraha, G. D., Han, H., Choi, D., Seo, S., & Kim, J. (2019). A decision tree-based NLOS detection method for the UWB indoor location tracking accuracy improvement. *International Journal of Communication Systems*, 32(13), e3997.
- [10] Graves, A. (2013). Generating sequences with recurrent neural networks. arXiv preprint arXiv:1308.0850.
- [11] Wang, Z. Q., Zhang, X., & Wang, D. (2018, August). Robust TDOA Estimation Based on Time-Frequency Masking and Deep Neural Networks. In *Interspeech* (pp. 322-326).
- [12] Poullose, A., & Han, D. S. (2020). UWB indoor localization using deep learning LSTM networks. *Applied Sciences*, 10(18), 6290.
- [13] Wang, H., Wang, X., Xue, Y., & Jiang, Y. (2020, June). UWB-based Indoor Localization Using a Hybrid WKNN-LSTM Algorithm. In *2020 IEEE 4th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)* (Vol. 1, pp. 1720-1725). IEEE.
- [14] Niitsoo, A., Edelh  user, T., & Mutschler, C. (2018, September). Convolutional neural networks for position estimation in tdoa-based locating systems. In *2018 International Conference on Indoor Positioning and Indoor Navigation (IPIN)* (pp. 1-8). IEEE.
- [15] Huang, L., Li, J., Hao, H., & Li, X. (2018). Micro-seismic event detection and location in underground mines by using Convolutional Neural Networks (CNN) and deep learning. *Tunnelling and Underground Space Technology*, 81, 265-276.
- [16] Zafari, F., Gkelias, A., & Leung, K. K. (2019). A survey of indoor localization systems and technologies. *IEEE Communications Surveys & Tutorials*, 21(3), 2568-2599.
- [17] Baek, J., Lee, S., Lee, C., & Park, S. (2021, October). A Deep Learning Approach for Robust Target Tracking in a Cluttered Environment. In *2021 International Conference on Information and Communication*

- [18] Yilmaz, O., & Rickard, S. (2004). Blind separation of speech mixtures via time-frequency masking. IEEE Transactions on signal processing, 52(7), 1830-1847.
- [19] Xingjian, S. H. I., et al. "Convolutional LSTM network: A machine learning approach for precipitation nowcasting." Advances in Neural Information Processing Systems. 2015

본 논문은 22년도 정부(산업통상자원부)의 재원으로 한국산업기술진흥원의 지원을 받아 수행된 연구입니다. (P0017123, 2022년 산업혁신인재성장지원사업)