

Indoor Path Recognition Based on Wi-Fi Fingerprints

Donggyu Lee¹, Jaehyun Yoo^{2†}

¹Department of Future Convergence Technology Engineering, Sungshin Women's University, Seoul 02844, Korea

²School of AI Convergence, Sungshin Women's University, Seoul 02844, Korea

ABSTRACT

The existing indoor localization method using Wi-Fi fingerprinting has a high collection cost and relatively low accuracy, thus requiring integrated correction of convergence with other technologies. This paper proposes a new method that significantly reduces collection costs compared to existing methods using Wi-Fi fingerprinting. Furthermore, it does not require labeling of data at collection and can estimate pedestrian travel paths even in large indoor spaces. The proposed pedestrian movement path estimation process is as follows. Data collection is accomplished by setting up a feature area near an indoor space intersection, moving through the set feature areas, and then collecting data without labels. The collected data are processed using Kernel Linear Discriminant Analysis (KLDA) and the valley point of the Euclidean distance value between two data is obtained within the feature space of the data. We build learning data by labeling data corresponding to valley points and some nearby data by feature area numbers, and labeling data between valley points and other valley points as path data between each corresponding feature area. Finally, for testing, data are collected randomly through indoor space, KLDA is applied as previous data to build test data, the K-Nearest Neighbor (K-NN) algorithm is applied, and the path of movement of test data is estimated by applying a correction algorithm to estimate only routes that can be reached from the most recently estimated location. The estimation results verified the accuracy by comparing the true paths in indoor space with those estimated by the proposed method and achieved approximately 90.8% and 81.4% accuracy in two experimental spaces, respectively.

Keywords: indoor path estimation, Wi-Fi fingerprinting, KLDA

1. 서론

최근 대형 건물의 증가와 스마트폰 보급의 증가로 실내 위치 기반 서비스에 대한 수요가 증가하고 있다. 실내 공간에서의 위치 정보는 쇼핑몰이나 터널 등에서 보행자의 위치인지 및 보행 경로 안내, 화재와 같은 긴급 상황 발생 시 고립된 인명 파악과 구조 등 여러 방면에서 활용할 수 있다. 이로 인해, 실내 위치 추정 방법에 대해 여러 산업과 연구로 많은 주목을 받고 있다. 위치 추정에 일반적으로 사용되는 Global Positioning System (GPS)은 수신된 전파의 강도 및 수신 대역 문제로 실내에서 사용하기에 적합하지 않다. 이러한 실내 환경에서 위치를 추정하기 위해 다양한 방법이 연구되고 있으며, 관성 센서를 이용한 실내 위치 추정 방

법 (Lu et al. 2019) 및 무선 Radio Frequency (RF) 신호를 이용한 방법 (Yamasaki et al. 2005, Xu et al. 2008, Kaune 2012), Radio-Frequency-Identification (RFID) (Jin et al. 2006), 센서 네트워크 (Chriki et al. 2017), Ultra-Wide Band (UWB) (Dabove et al. 2018), Bluetooth Low Energy (BLE) Beacon (Zhuang et al. 2016), Wireless Fidelity (Wi-Fi) (Hansen et al. 2010, Leu & Tzeng. 2012, So et al. 2013, Mendoza-Silva et al. 2022) 등이 있다.

대표적으로 관성 센서를 이용한 실내 위치 추정은 Inertial Measurement Unit (IMU)를 이용하여 보행자의 걸음 수, 보폭, 속도, 이동 거리, 보행 방향 등을 계산하여 보행자의 위치를 추정하는 방법으로, 비교적 정밀하고 즉각적인 위치 추정을 수행할 수 있지만, 센서 값이 연산 중 오차가 누적되어 위치를 추정하는 시간이 지속됨에 따라 추정된 위치 정보가 발산하는 문제가 있다. 그러므로 관성센서만을 이용하여 위치를 추정하는 것은 적절한 방법으로 볼 수 없고 누적 오차를 보정하는 알고리즘이나 다른 센서를 융합하여 보정하는 방법 등이 필요하다.

전파 모델 기반 무선 신호 실내 위치 추정 방법은 대표적으로 Time Difference of Arrival (TDOA), Time of Arrival (TOA),

Received Feb 24, 2023 Revised Mar 21, 2023 Accepted Apr 27, 2023

[†]Corresponding Author

E-mail: jhyoo@sungshin.ac.kr

Tel: +82-920-7695

Donggyu Lee <https://orcid.org/0009-0005-7199-5492>

Jaehyun Yoo <https://orcid.org/0000-0002-6167-2842>

Angle of Arrival (AOA) 등이 있고 실내에 설치된 송신기 신호와 전파 모델을 이용하여 송수신 시간, 각도, 거리 등을 추정하여 위치를 계산하는 방법이다. 이 방법은 수신 신호의 Line-Of-Sight 성분이 강한 경우에만 사용이 가능하고 정확한 동기화가 이루어져야 하며, 오프셋으로 인한 위치 오차가 발생할 수 있다. 그러므로 이 방법은 실내 위치를 추정하기에 적합하지 않다.

UWB를 이용한 위치 추정 방법은 기존 무선통신기술보다 대역폭이 매우 넓은 초광대역 펄스를 이용하여 실내 공간에서 1 m 이하인 비교적 낮은 오차로 위치를 추정하는 방법이다. 소비 전력이 적고 비교적 정밀하게 위치를 추정할 수 있지만, 다른 통신과의 간섭이 발생할 수 있으며, UWB를 이용하여 위치를 추정하기 위해 실내 공간에 UWB 네트워크를 구축해야 하므로, 추가적인 인프라 구축 비용이 필요하다. BLE를 이용한 위치 추정 방법은 추정 시 전력소모가 적고 대부분의 스마트폰에 내장되어 있으며, 장치가 비교적 저렴하다. 그러나, 장치의 전송 거리가 짧으며, 적용범위가 제한되며, 실내 공간에서 벽 또는 사람에 의해 쉽게 영향을 받고, UWB와 같이 실내 공간에서 위치를 추정하기 위해 추가적인 인프라 구축이 필요하다.

Wi-Fi를 이용한 위치 추정 방법은 Cell-ID 기반 추정 방법과 fingerprinting 기반 방법, 전파 예측 모델 기반 방법 등이 있다. 대표적으로 fingerprinting 기반 방법은 위치를 추정하려는 실내 공간에 일정한 간격으로 레이블 된 좌표를 정하고 각 좌표에서 주변 Wi-Fi Access Point (AP)로부터 수신되는 Received Signal Strength Indicator (RSSI)을 반복하여 수집하고 평균 낸 값을 Data-Base (DB)에 저장하고, 실내 공간에서 위치 정보를 얻고 싶은 곳에서 수신된 RSSI 값과 DB에 저장된 값을 비교하여 가장 근사한 특성을 갖는 좌표로 위치 정보를 제공하는 방법이다. 이 방법은 대부분의 건물과 공공 시설에 이미 무선 AP가 구축되어 있어 추가적으로 장비를 설치하지 않아도 실내 위치 추정을 수행할 수 있고, 상대적으로 서비스 범위가 넓으며, AP의 정확한 위치를 식별할 필요없이 Wi-Fi 신호 강도를 사용하므로 복잡한 실내 환경에서 도착 각도 및 시간과 같은 시공간 속성과 다르게 상대적으로 안정적이고 (Xia et al. 2017), 건물 내부 구조에 상관없이 적용이 가능하여 널리 이용된다. 그러나, 실내 공간에서 레이블을 정한 특정 간격의 위치 마다 AP로부터 수신되는 많은 양의 RSSI 데이터를 필요로 하고, 이는 많은 비용과 시간이 필요하다 (Yoo 2020). 또한, 실내공간의 배치에 따라 RSSI 값이 크게 변하므로, DB를 다시 구축해야 하는 번거로움이 있다.

기존 Wi-Fi fingerprinting을 이용한 실내 위치 추정 방법은 실내 공간을 일정 간격으로 레이블 된 좌표를 정하여 모두 수집해야 하고 이는 데이터 수집에 많은 비용이 요구되며, 규모가 큰 수집 공간에서는 수집이 불가능에 가깝다. 또한, 낮은 정확도를 가지므로, 좀 더 효율적인 방법이 필요하다 (Kamarungsi 2005, Mainetti et al. 2014). 본 논문은 기존 방법과 같이 특정 좌표로 위치를 추정하지 않고, 실내 공간에서 특정 경로 또는 교차로 (특징지역)로 위치를 추정하지만, 레이블 없이 이동하며 데이터를 수집하여 수집 비용을 크게 줄이고 정확도를 크게 높여, 효율적으로 위치를 추정하는 방법을 제안하였다. 제안하는 방법은 다음과 같다. 실내 공간에 임의로 특징지역 (Landmark)을 지정하여 특징지역의 위치 라벨과 Wi-Fi fingerprinting을 수집한

다. 특징지역을 제외한 실내 공간을 이동하며 라벨이 없는 Wi-Fi fingerprinting을 수집하고 이후 알고리즘을 통해 자동으로 라벨링을 하여 보행자의 이동 경로를 추정한다. 제안하는 방법과 기존 방법의 비교를 위해, 두 방법 모두 수행할 수 있는 비교적 규모가 작은 공간에서 두 방법을 비교하였고 기존 방법으로 추정거리의 불가능에 가까운 규모의 큰 공간에서도 제안하는 방법으로 경로 추정이 가능함을 보였다.

본 논문의 구성은 다음과 같다. 2장에서 기존 Wi-Fi fingerprinting 기반의 경로를 추정하는 방법에 대해 서술하고 3장에서 제안하는 경로 추정 방법으로 실험공간에서 Wi-Fi fingerprinting을 수집하는 방법과 전처리를 통한 DB 구성 방법을 서술하고, 4장에서 실험 도구 및 실험환경에 대해 서술하고, 5장은 실험 결과, 그리고 6장은 결론이다.

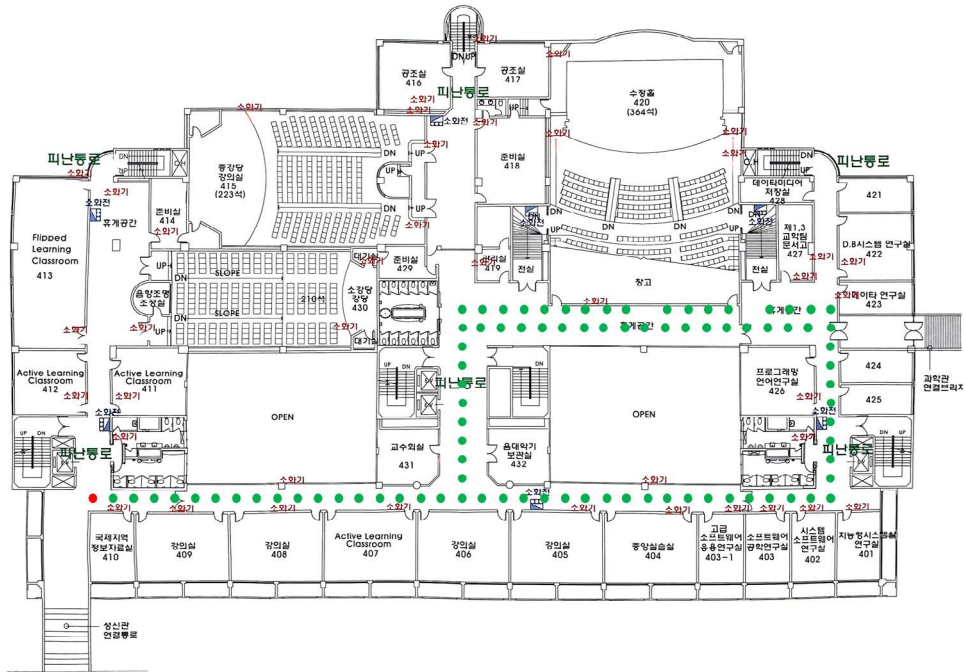
2. 경로 추정을 위한 기존 방법

본 장에서는 Wi-Fi fingerprinting 기반의 실내 위치 추정 시 가장 일반적으로 사용하는 방법인, Principal Component Analysis (PCA) (Daffertshofer et al. 2004)을 이용하여 데이터를 처리하고, K-Nearest Neighbor (K-NN) (Keller et al. 1985)을 이용하여 위치와 경로를 추정하는 방법을 이용하여 Fig. 1a의 위치를 추정하는 과정에 대해 서술한다.

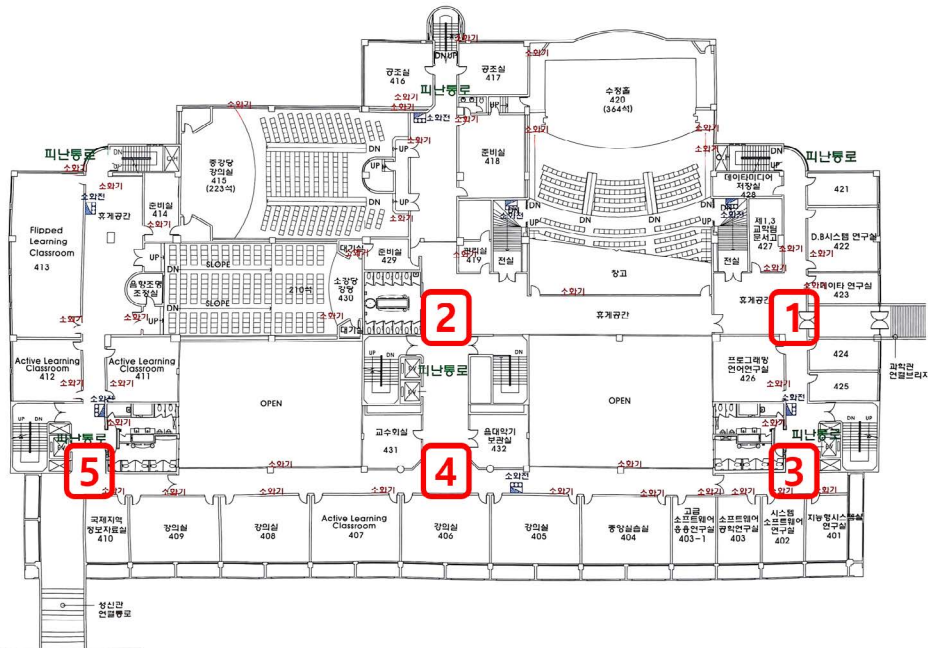
데이터 수집과정은 다음과 같다. 학습데이터는 수집 공간인 Fig. 1a에서 보행자가 수집 도구를 들고 미리 지정한 89개의 좌표에 서서 좌표와 RSSI, BSSID 쌍을 각각의 좌표에서 20개가 저장되도록 하여 얻었다. 테스트 데이터는 Fig. 1b에 나타난 특징지역 번호 1, 3, 4, 5, 4, 2, 1의 순으로 이동하며 데이터를 수집하였고, 총 60개가 수집되었다. 데이터 전처리 과정은 다음과 같다. 학습 데이터와 테스트 데이터 모두에서 수집된 BSSID를 중복 없이 나열하고, 나열된 순서에 맞게 학습데이터와 테스트 데이터 각각의 RSSI를 재배열한다. 재배열 시 특정 BSSID에 대해 수집된 RSSI가 없는 경우, 가장 작은 가중치 값인 -100을 배열한다. 재배열된 학습 데이터 중 같은 좌표에서 수집한 RSSI 값들은 평균 내어, $(89(\text{좌표의 개수}) \times \text{중복없이 나열한 BSSID의 수})$ 의 학습 데이터와 $(60 \times \text{중복없이 나열한 BSSID의 수})$ 의 테스트 데이터를 얻고, PCA를 이용하여 데이터 차원의 크기를 축소한다. 전 처리 과정을 거친 학습 데이터를 K-NN을 이용하여 학습하고 테스트 데이터의 위치와 경로를 추정한다.

3. 경로 추정을 위해 제안하는 방법

이번 장에서는 특징지역과 특징지역을 경유하며 라벨없이 수집한 데이터를 Kernel Linear Discriminant Analysis (KLDA)을 이용하여 클러스터링 (Clustering)하고 특정 특징지역 사이에 위치한 경로 데이터로 라벨링하는 과정을 거쳐 K-NN을 이용하여 경로를 추정하는 과정을 서술한다. 전체 과정은 데이터 수집 단계와 데이터 전 처리, 추정 단계로 나누었고, 각 단계에서 수행하는 것은 다음과 같다.



(a) An existing method



(b) Methods proposed in this paper.

Fig. 1. 4F of Sungshin Women's University, Sujeong Hal.

● 데이터 수집 단계

- 1) 수집 공간 내 복도 교차로 부근을 특징지역으로 설정.
- 2) 데이터 (D_1): 특징지역에 수집 도구를 배치하여 라벨 (특징 지역 번호)과 RSSI, BSSID 수집하여 구축.
- 3) 학습 데이터 (D_2): 각 특징지역을 최소 한 번씩 경유하며 라벨 없이 RSSI, BSSID 수집하여 구축.
- 4) 테스트 데이터 (D_3): 실내 공간을 무작위로 이동하며 RSSI,

BSSID 수집하여 구축.

● 데이터 전 처리 단계

- 1) 수집한 데이터 (D_1, D_2, D_3) 재배열.
- 2) KLDA를 이용하여 D_1, D_2, D_3 을 D_1, D_2, D_3 로 처리.
- 3) 특징 공간 (feature space) 내에서 D_1 의 각 클래스와 D_2 의 유클리드 거리 값을 구하고 값의 계곡 지점 (valley point)들을

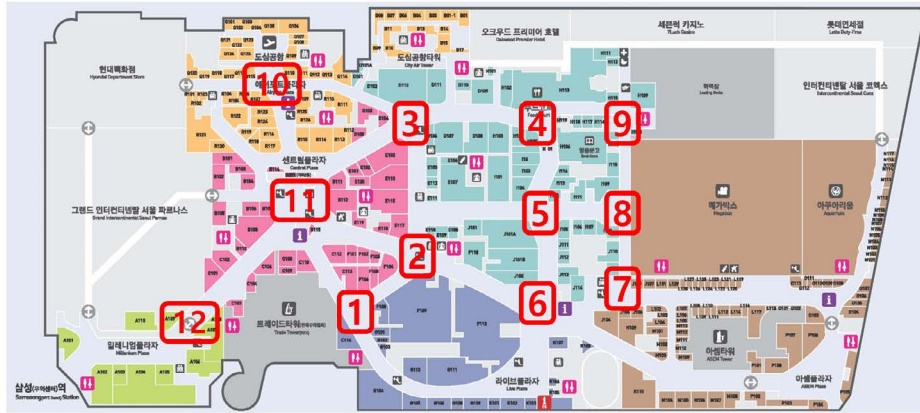


Fig. 2. Coex mall B1F.

구함.

- 4) 각 계곡 지점 사이 데이터 (D_2)를 해당 특징지역 사이 경로로 라벨링, 계곡 지점과 가까운 데이터 (D_2)를 특징지역의 라벨로 라벨링.

● 추정 단계

- 1) 라벨링한 학습 데이터 (D_2)를 K-NN을 이용해, 학습하고 테스트 데이터 (D_3)를 추정.

3.1 데이터 수집

데이터 수집 공간은 Fig. 1과 같이 성신여자대학교 수정관 4층과 Fig. 2의 삼성역 스타필드 코엑스몰 B1층이고, 일반적으로 사람이 보행할 수 있는 공간 중 교차로 부근을 특징지역으로 지정하였다. 데이터 수집 방법은 다음과 같다. Fig. 1b와 Fig. 2에서 동일한 방법으로 총 3종류의 데이터 (D_1, D_2, D_3)를 구축하였다. D_1 은 각각의 특징지역에서 사람이 움직이지 않고 정지한 상태로 스마트폰을 들고 있거나 삼각대를 이용하여 스마트폰을 고정한 상태로 각 특징지역을 나타내는 라벨과 RSSI, BSSID를 Fig. 1에서는 100개씩 총 500개, Fig. 2에서는 200개씩 총 2400개 수집하여 얻었다. D_2 는 각각의 특징지역들 간에 이동 가능한 모든 경로를 최소 한번 이상 경유하며 라벨없이 RSSI와 BSSID를 수집하여 얻었다. D_3 은 결과 분석을 위한 테스트 데이터로, 실내 공간을 무작위로 이동하며 RSSI와 BSSID를 수집하였고 이동 중 특징지역에 도착할 때마다 타임 스탬프 (time stamp)를 저장하여 이후 결과 분석에 사용하도록 하였다.

3.2 데이터 전 처리

3.2.1 데이터 재배열

수집한 D_1, D_2, D_3 의 BSSID에서 특정 빈도수 이하로 수집된 데이터를 제거하고 중복이 없도록 하여 하나의 벡터로 나열한다. 이 중복없이 나열한 BSSID에 대해 나열된 순서와 동일하게 D_1, D_2, D_3 의 RSSI를 재배열한다. 재배열 중 특정 BSSID에 대해 수집된 RSSI가 없는 경우, 가장 작은 가중치 값인 “-100”을 배치한다.

3.2.2 경로 데이터 라벨링을 위한 KLDA

KLDA는 비선형 함수를 이용하여 각 데이터 클래스의 분산을 최소화하고 데이터 클래스 간의 분산은 최대화하는 특징 벡터 ω 에 데이터를 투영하여 Sparse (Aarons 1993)한 데이터의 차원을 축소한다. 동시에 데이터를 각 클래스 별로 나누어 데이터를 클러스터링 한다. 본 장에서는 라벨이 없는 D_2 을 특징지역 사이에 위치한 경로로 라벨링하기 위해 D_1, D_2 순으로 KLDA를 이용하여 처리하고 처리한 D_2 와 D_1 의 특정 클래스 간의 유클리드 거리를 이용하여 D_2 을 경로 데이터로 클러스터링하는 과정을 서술한다.

앞서 처리한 D_1, D_2, D_3 을 식 (1)로 각각 정의한다. n 은 수집된 Wi-Fi AP 개수로, 앞서 말한 BSSID를 중복없이 나열한 벡터의 차원 수와 같다. $rssi_{dk}$ ($k=1, \dots, n$)는 n 개의 Wi-Fi AP 중 k 번째 AP에서 얻은 RSSI 스칼라 값이고 c 는 클래스 수로, 각 수집 공간에서의 특징지역 수와 같고 L_1, \dots, L_c 는 각 특징지역에서 얻은 데이터의 개수이고 l_2, l_3 은 D_2, D_3 을 각각 수집한 데이터 개수이다.

$$\begin{aligned}
 D_1 &= \begin{bmatrix} rssi_{l_{11}}, rssi_{l_{12}}, rssi_{l_{13}}, \dots, rssi_{l_{1n}} \\ \vdots \\ rssi_{l_{c1}}, rssi_{l_{c2}}, rssi_{l_{c3}}, \dots, rssi_{l_{cn}} \end{bmatrix} \in R^{L \times n}, \left(L = \sum_{i=1}^c L_i \right) \\
 D_2 &= \begin{bmatrix} rssi_{l_21}, rssi_{l_22}, rssi_{l_23}, \dots, rssi_{l_2n} \\ \vdots \\ rssi_{l_31}, rssi_{l_32}, rssi_{l_33}, \dots, rssi_{l_3n} \end{bmatrix} \in R^{l_2 \times n} \\
 D_3 &= \begin{bmatrix} rssi_{l_31}, rssi_{l_32}, rssi_{l_33}, \dots, rssi_{l_3n} \\ \vdots \\ rssi_{l_31}, rssi_{l_32}, rssi_{l_33}, \dots, rssi_{l_3n} \end{bmatrix} \in R^{l_3 \times n}
 \end{aligned} \tag{1}$$

KLDA의 커널은 가우시안 커널 (Gaussian Kernel)을 사용하였으며, Φ 는 특징 공간에 대한 비선형 매핑 함수이다. 특징 공간 (\mathcal{F} , feature space)에서 식 (2)를 최대화하는 벡터 ω 가 식 (2)의 최적 해가 된다. 여기서 $\omega \in \mathcal{F}$ 와 S_B^Φ, S_W^Φ 는 \mathcal{F} 의 대응 행렬이다. 특징 공간은 D_1 을 KLDA를 이용하여 차원을 축소하고 축소된 각 차원의 스칼라 값을 하나의 축의 값으로 하여 나타낸 공간을 말한다.

$$J(\omega) = \frac{\omega^T S_B^\Phi \omega}{\omega^T S_W^\Phi \omega} \tag{2}$$

식 (2)의 S_B^Φ 와 S_W^Φ 그리고 m_i^Φ 는 식 (3)과 같다.

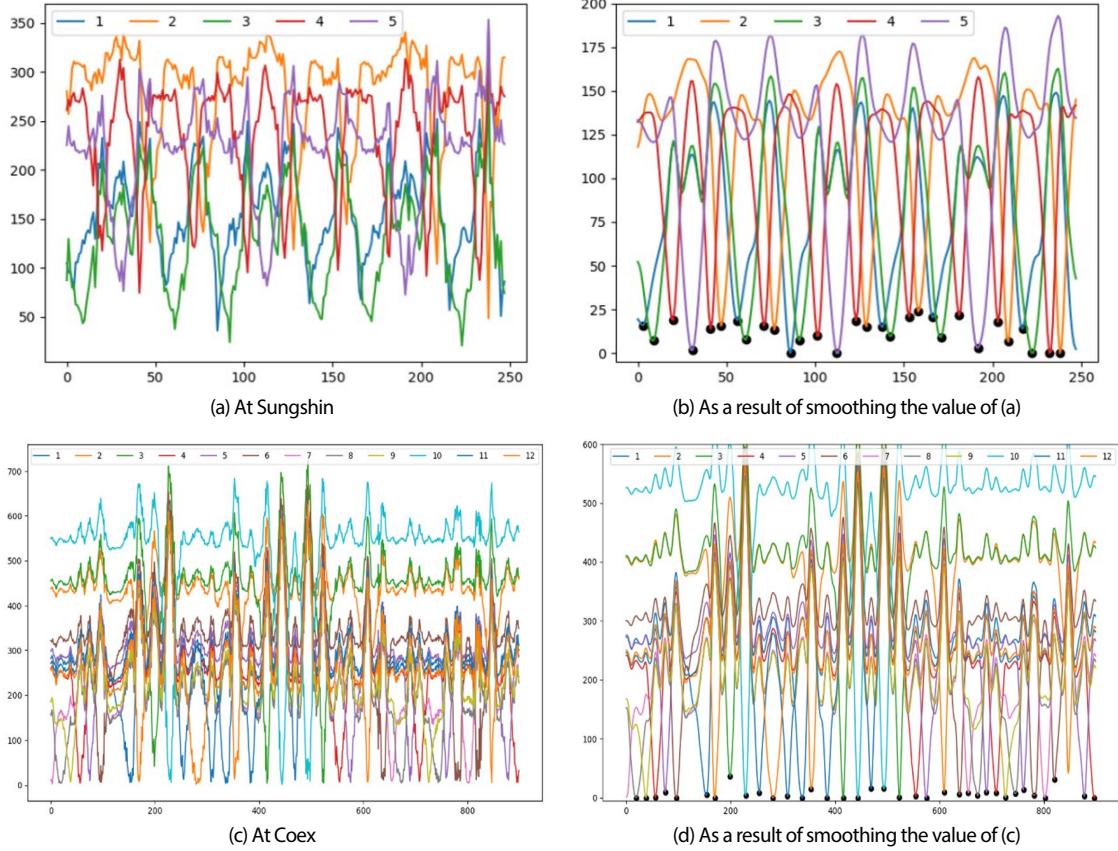


Fig. 3. Euclidean distance for each class of D_1 and D_2 in feature space.

$$\begin{aligned}
 m_i^\Phi &= \frac{1}{n} \sum_{x \in X} x_k, \text{ with } (X = D_1(i, \{1, 2, \dots, n\})) \\
 S_B^\Phi &= (m_1^\Phi - m_2^\Phi - \dots - m_c^\Phi)(m_1^\Phi - m_2^\Phi - \dots - m_c^\Phi)^T \\
 S_W^\Phi &= \sum_{i=1}^L \sum_{x \in X} (\Phi(x) - m_i^\Phi)(\Phi(x) - m_i^\Phi)^T \quad (3)
 \end{aligned}$$

Mika et al. (1999)의 계산과정에 따라 식 (3)을 이용하여 식 (2)의 최적해를 구하는 과정을 거쳐서 식 (4)를 얻을 수 있고, 식 (4)는 $Av = \lambda v$ 형태로 일반적인 고유 값 문제 (Eigenvalue problem) (Parlett 1998)의 풀이 과정으로 ω 를 구하여 식 (5)와 같이 ω 에 D_1, D_2, D_3 을 각각 투영하여 $\mathbb{D}_1, \mathbb{D}_2, \mathbb{D}_3$ 을 얻는다.

$$(S_W^\Phi)^T S_B^\Phi \omega = \lambda \omega \quad (4)$$

$$\begin{aligned}
 \mathbb{D}_1 &= (\omega \cdot \Phi(D_1)) = \sum_{i=1}^n \omega_i \cdot k(D_{1i}, D_1) \\
 \mathbb{D}_2 &= (\omega \cdot \Phi(D_2)) = \sum_{i=1}^n \omega_i \cdot k(D_{1i}, D_2) \\
 \mathbb{D}_3 &= (\omega \cdot \Phi(D_3)) = \sum_{i=1}^n \omega_i \cdot k(D_{1i}, D_3) \quad (5)
 \end{aligned}$$

3.2.3 경로 데이터 라벨링

이번 절은 KLDA를 이용하여 처리한 \mathbf{D}_2 을 특징지역 사이의

경로로 라벨링하는 과정을 서술한다. Fig. 3은 특징 공간 상에서 \mathbf{D}_1 의 각 클래스의 평균값과 \mathbf{D}_2 간의 유클리드 거리 (Euclidean distance)를 나타낸 그래프이다. Fig. 3a는 Fig. 1에서 수집한 데이터의 유클리드 거리를 나타내고 Fig. 3b는 Fig. 3a를 가우시안 필터 (Gaussian filter)를 이용해 값을 스무딩 (smoothing)하여 나타낸 그래프이다. Fig. 3c는 Fig. 2에서 수집한 데이터의 유클리드 거리를 나타내고, Fig. 3d는 Fig. 3c를 가우시안 필터를 이용하여 값을 스무딩한 그래프이다. 다음으로, Figs. 3b와 3d에 적절히 정한 임계 값 (threshold) 이하의 값들 중 식 (6)을 사용하여 계곡 지점 (valley point)을 찾아 표시하였다. 식 (6)은 특징 공간 상에서 \mathbf{D}_1 의 각 클래스와 \mathbf{D}_2 간의 유클리드 거리를 구한 값들의 집합을 순서대로 나열한 값을 Ed 로 정의하고 n 은 Ed 의 수, 계곡 지점의 데이터는 Ed_k 로 정의한다.

$$\begin{aligned}
 Ed &= (Ed_1, Ed_2, \dots, Ed_n) \\
 Ed_v &= Ed_k \text{ (when } Ed_{k-1} < Ed_k < Ed_{k+1} \text{ (} 1 < k < n \text{))} \quad (6)
 \end{aligned}$$

계곡 지점 데이터는 보행자가 하나의 특징지역과 가장 근접했을 때 수집된 데이터라 할 수 있다. 그리고 계곡 지점 사이의 데이터들은 보행자가 두 곳의 특징지역 사이를 이동할 때 수집된 데이터라고 할 수 있다. 그러므로, 각 계곡 지점 사이에 존재하는 \mathbf{D}_2 데이터들을 두 특징지역 사이 경로 데이터로 라벨링한다. 다음으로, 특징지역은 실내 공간 내에 특정 크기 공간을 아우르는 지

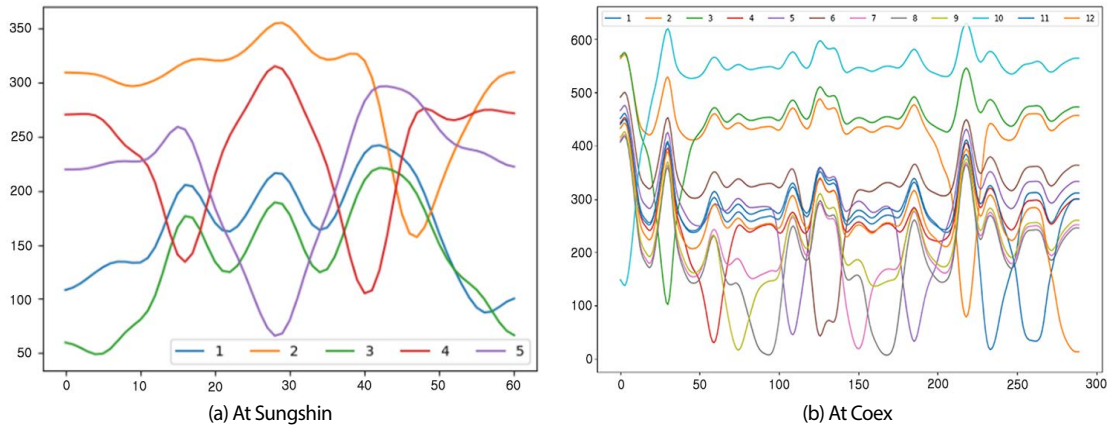


Fig. 4. Euclidean distance for each class of D_1 and D_3 in feature space.

Table 1. Answer standard for path estimation.

Step	True	Some of case that are considered correct						Some of cases wrong	
		case 1	case 2	case 3	case 4	case 5	case 6		
1	between 1 and 3(1,3)	(1,3)	(1,3)	(1,3)	(1,3)	(1,3)	(1,3)	(1,3)	(1,3)
2	(1,3)	(1,3)	(1,3)	(1,3)	(1,3)	(1,3)	(1,3)	Land 3	(1,3)
3	(1,3)	(1,3)	(1,3)	(1,3)	(1,3)	(1,3)	(1,3)	(3,4)	(1,3)
4	(1,3)	Land 3	(1,3)	(1,3)	(1,3)	Land 3	(1,3)	(3,4)	Land 3
5	(1,3)	(3,4)	Land 3	(1,3)	(1,3)	Land 3	Land 3	(3,4)	(3,4)
6	Land 3	(3,4)	(3,4)	Land 3	(1,3)	Land 3	Land 3	(3,4)	Land 3
7	(3,4)	(3,4)	(3,4)	(3,4)	Land 3	Land 3	(3,4)	(3,4)	(3,4)
8	(3,4)	(3,4)	(3,4)	(3,4)	(3,4)	Land 3	(3,4)	(3,4)	(3,4)
9	(3,4)	(3,4)	(3,4)	(3,4)	(3,4)	(3,4)	(3,4)	(3,4)	(3,4)
...

역이기 때문에 D_2 의 데이터 중 각 특징지역과 인접한 데이터들은 특징지역 라벨로 설정한다. Fig. 4는 특징 공간 상에서 D_1 의 각 클래스와 D_3 간의 유클리드 거리를 나타낸다. Figs. 4a,b는 각각 Figs. 1과 2에서 수집한 데이터의 유클리드 거리로, 가우시안 필터를 사용해 값을 스무딩하여 나타낸 그래프이다.

3.3 K-NN을 이용한 경로 추정

본 연구에서는 식 (7)와 같이 K-NN 알고리즘을 이용하여 D_2 를 학습하고 D_3 의 이동 경로를 예측하였다. N 은 하나의 D_3 데이터와 모든 D_2 데이터와의 유클리드 거리를 구해 나열한 벡터이다. N 의 값들을 오름차순으로 나열한 후 가장 작은 수부터 적절히 선정한 변수 K 까지의 원소들의 라벨 중 가장 빈도가 많은 라벨로 D_3 의 경로를 예측한다.

$$D_2 = \begin{bmatrix} a_{11}, a_{12}, \dots, a_{1M} \\ \vdots \\ a_{\ell_2 1}, a_{\ell_2 2}, \dots, a_{\ell_2 M} \end{bmatrix} \in R^{\ell_2 \times M}$$

$$D_3 = \begin{bmatrix} b_{11}, b_{12}, \dots, b_{1M} \\ \vdots \\ b_{\ell_3 1}, b_{\ell_3 2}, \dots, b_{\ell_3 M} \end{bmatrix} \in R^{\ell_3 \times M}$$

$$N_j = \sqrt{\sum_{i=1}^M (a_{ji} - b_{1i})^2}, (j = 1, 2, \dots, \ell_2)$$

$$N = (N_1, N_2, \dots, N_{\ell_2}) \in R^{\ell_2} \tag{7}$$

또한, 식 (7)를 통해 추정된 라벨이 특징지역일 때, 해당 특징지역에서 도달할 수 있는 경로로만 추정하거나 가장 최근에 추정된 특징지역에서 이동 가능한 경로만 추정하도록 보정하는 알고리즘을 사용하였다.

4. 실험 환경

이 논문에서는 두 곳의 실내공간에서 실험을 진행하였다. 첫 번째 실험 환경은 약 3,400 m²의 비교적 협소한 실내 공간으로 Fig. 1과 같다. Fig. 1a는 기존 추정 방법을 수행하기 위해, 실내공간에 약 2 m 간격으로 좌표를 설정하여 나타내었다. Fig. 1b는 본 논문에서 제안하는 방법을 수행하기 위해, 복도 교차로 부근에 임의로 총 5개의 특징지역을 설정하였다. 이동 경로는 Fig. 1b에서 각각의 특징지역과 특징지역 사이의 공간을 말하고, 총 10개이다.

두 번째 실험 환경은 약 119,008.8 m²의 비교적 넓은 실내 공간으로, Fig. 2와 같다. 두 번째 실험 공간은 기존 Wi-Fi

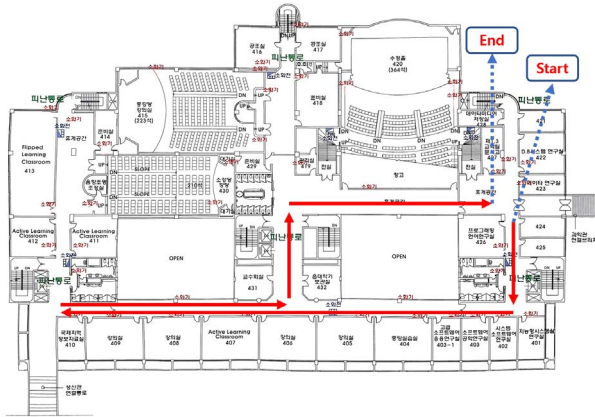
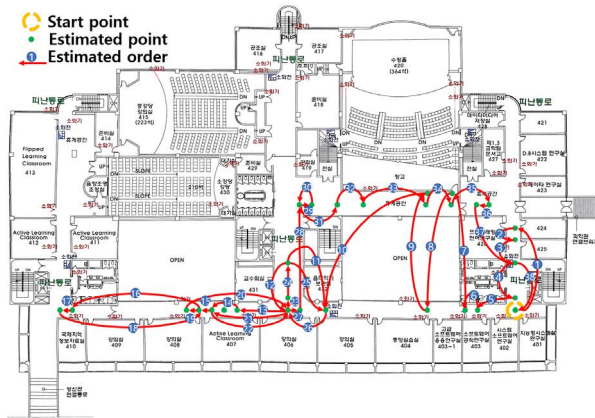
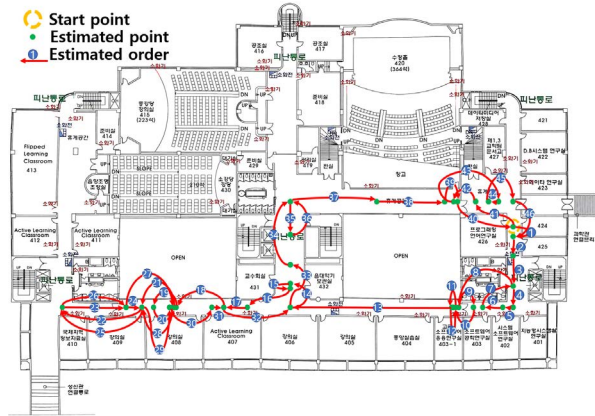


Fig. 5. True Path of test data at Sungshin.



(a) Paths estimated by the traditional method



(b) Paths estimated by the proposed method

Fig. 6. Result image at Sungshin.

fingerprinting 기반 위치 추정 방법을 사용하기에 거의 불가능에 가까운 규모의 실내 공간으로, 본 논문에서 제안하는 방법 추정 여부를 확인하기 위해 설정하였다. 실험을 위해 복도 교차로 부근에 임의로 총 12개 (1~12)의 특징지역을 설정하였다. 본 논문에서 추정하는 이동 경로는 Fig. 2에서 각각의 특징지역과 특징지역 사이의 공간을 말하고, 총 31개이다. 이 논문이 제안하는 추정 방법은 수집 공간 Figs. 1과 2에서 수행하였고, 기존 추정 방법은

Fig. 1에서만 수행하였다

5. 실험 결과

Table 1은 본 논문의 알고리즘을 이용하여 추정한 이동경로를 평가하기 위한 기준표로, Land 1 (특징지역 1)에서 Land 3, Land 4로 이동하며 수집한 테스트 데이터를 평가한 결과의 예시이다. Table 1의 Step은 데이터의 번호로, Step 1은 첫번째 스캔하여 얻은 데이터이다. True는 테스트 데이터 수집 시 실제 이동한 경로이다. 테스트 데이터 수집 시, 특징지역을 지나칠 때마다 타임 스탬프를 저장하여 특징지역에 도착한 것을 기록하였고, 이는 실제 특징지역 도착 시간과의 오차가 있으므로 예측 가능 구간을 설정하였다. 평가 시, Land 3 전후로 성신여자대학교에서는 2개, 코엑스몰에서는 3개의 Step을 예측 가능 구간으로 설정하였다. 예측 가능 구간 내에 Land 3을 예측 또는 언달아 예측한 후 다음 경로를 예측했을 때, 옳게 예측한 것으로 평가하였다. Table 1의 some of cases wrong은 틀리게 예측한 케이스의 예시로, Land 3 예측 가능 구간 외에 Land 3을 예측하거나 예측 가능 구간 내에서 Land 3을 불연속적으로 예측했을 때를 틀리게 예측한 것으로 평가하였다. Fig. 5는 Fig. 1에서 수집한 테스트 데이터의 실제 이동 경로이다. Fig. 6의 초록색 점은 추정한 좌표를 표시한 것이고, 파란색 동그란 점선은 첫 번째로 추정한 좌표를 표시한 것이며, 빨간 화살표와 번호는 추정한 좌표의 순서를 나타낸 것이다. Fig. 6a는 기존 방법으로 테스트 데이터의 경로를 추정한 결과이고, 초록색 점이 추정한 좌표에 해당한다. Fig. 6b는 제안하는 방법으로 테스트 데이터의 경로를 추정하고, 식 (5)의 D_1 과 식 (7)의 D_3 간의 유클리드 거리를 실내 지도 비율에 맞게 표시한 것이다. Fig. 7a는 Fig. 2에서 수집한 테스트 데이터의 실제 이동 경로를 나타내고, Fig. 7b는 Fig. 6b와 같이 제안하는 방법으로 추정한 테스트 데이터의 경로를 실내 지도에 표시한 것이다. Figs. 6a,b와 7b에서 수집한 데이터의 개수보다 표시된 좌표의 수가 적은 이유는 연속적으로 같은 곳을 추정한 좌표는 제외했기 때문이다. Table 2는 실험 환경과 각각의 다른 환경에서 기존 방법과 본 논문에서 제안하는 방법으로 실험을 진행했을 때의 결과를 나타낸다. Fig. 6a와 같이 기존 방법으로 추정한 결과 1,780개의 데이터를 수집하여 학습하고 59개의 테스트 데이터를 추정한 결과 47개를 옳게 예측하여 79.7%의 정확도를 얻었다. 또한, Fig. 6b와 같이 제안하는 방법으로 500개의 데이터를 수집하여 학습하고 61개의 테스트 데이터를 추정한 결과 55개를 옳게 예측하여 90.2%의 정확도를 얻었다. 최종적으로, Fig. 7b와 같이 제안하는 방법으로 290개의 테스트 데이터를 추정한 결과 236개를 옳게 예측하여 81.4%의 정확도를 얻었다.

6. 결론

본 논문에서 실내 공간에 임의로 특징지역을 정하고 특징지역과 특징지역을 이동한 경로의 Wi-Fi fingerprinting을 KLDA를 이용하여 전처리하고 라벨이 없는 경로 데이터를 라벨링하여

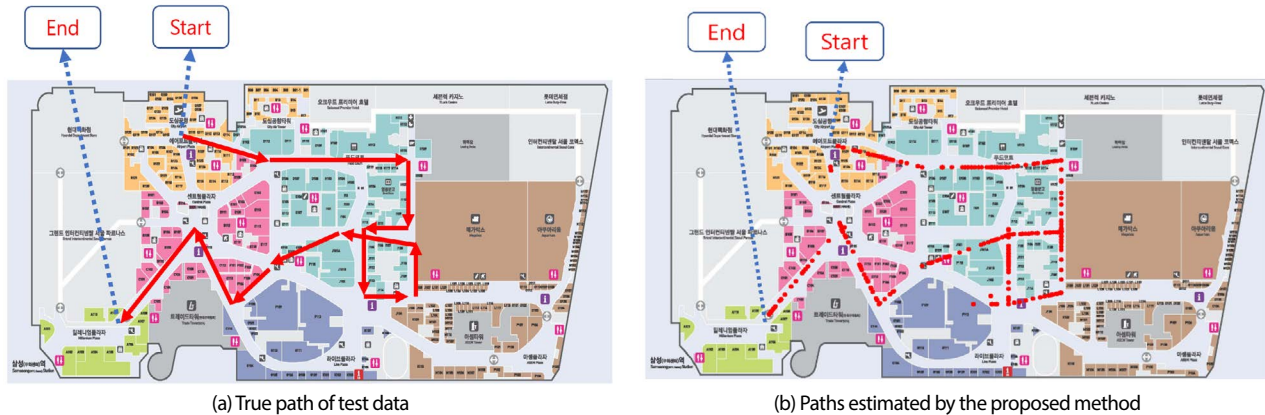


Fig. 7. Result image at Coex.

Table 2. Experimental environment and result.

Environment	Sungshin		Coex	
Area (m ²)	3,400		119,008.8	
Method	Existing	Proposed	Existing	Proposed
Number of data	1,780	500	64,000 (presumption)	2,400
Accuracy (%)	79.7	90.2	-	81.4

K-NN을 이용해 보행자의 이동 경로를 추정하는 방법을 제안하였다. 본 논문에서 제안하는 방법과 기존 추정 방법을 Fig. 1에서 각각 추정한 결과 약 3.5배의 더 많은 학습 데이터를 수집한 기존 방법이 더 낮은 정확도를 보였다. 또한, 기존 방법으로 추정이 불가능에 가까운 큰 규모의 Fig. 2에서 제안하는 방법으로 추정된 결과 81.4%의 높은 정확도로 추정이 가능함을 보였다. 본 실험에서 Fig. 2을 기존 방법으로 추정이 불가능에 가깝다고 가정할 이유는 실험공간, Fig. 1에서 통로의 일부만 좌표를 설정하고 실험하여, 89개의 좌표가 필요로 했는데, Fig. 2의 면적은 Fig. 1의 면적보다 약 35배 크므로, Fig. 2에서 실험을 진행하기 위해, Fig. 1의 약 35배인 3,200개의 좌표가 필요하다면, 한 좌표 당 20개의 데이터를 수집한다면, 64,000개의 데이터를 수집해야 하고, 수집 시간은 약 3,200분이 필요할 것이고, 실내 공간을 일정한 크기로 나누어 좌표를 정하고 라벨을 매기는 것을 더하면 실로 더 많은 시간이 필요할 것이다. 그리고 실험을 진행하더라도, 정확한 추정 결과를 기대할 수 없다. 본 논문에서 제안하는 이동 경로 추정 방법은 좌표로 위치 정보를 제공하지 않지만, 기존 방법에서 좌표와 같이 라벨을 설정하지 않아도 되고, 특징지역에 수집 도구만 비치하여 수집하고 라벨없이 이동하며 수집한 데이터만을 이용하여, 기존보다 훨씬 적은 수의 데이터를 수집해도 되며, 학습 데이터를 지속적으로 최신화하기에 용이하다. 그러므로, 본 논문에서 제안하는 이동 경로 추정 방법은 기존 방법에 비해 훨씬 효율적인 방법이라고 할 수 있다. 후속 연구에서는 본 논문에서 제안한 이동 경로 추정 방법과 지속적인 위치 추정이 가능하지만, 위치 추정 시 오차가 누적되는 IMU 센서를 상호 보완하여 실시간의 높은 정확도로 실내 보행자의 위치 정보를 제공하는 알고리즘을 연구할 계획이다.

ACKNOWLEDGMENTS

This work was supported by a National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT) (No. 2022R1F1A1070968).

AUTHOR CONTRIBUTIONS

Conceptualization, J.Y.; methodology, D.L.; software, D.L.; validation, D.L.; formal analysis, J.Y.; investigation, J.Y.; experiment, D.L.; writing—original draft preparation, D.L.; writing—review and editing, J.Y.; funding acquisition, J.Y.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

REFERENCES

- Aarons, L. 1993, Sparse data analysis. *European journal of drug metabolism and pharmacokinetics*, 18, 97-100. <https://doi.org/10.1007/BF03220012>
- Chriki, A., Touati, H., & Snoussi, H. 2017, SVM-based indoor localization in wireless sensor networks, 13th international wireless communications and mobile computing conference (IWCMC), IEEE, 26-30 June 2017, Valencia, Spain. <https://doi.org/10.1109/IWCMC.2017.7986446>
- Dabove, P., Di Pietra, V., Piras, M., Jabbar A. A., & Kazim, S. A. 2018, Indoor positioning using Ultra-wide band (UWB) technologies: Positioning accuracies and sensors' performances, IEEE/ION Position, Location and Navigation Symposium (PLANS), 23-26 April 2018, Monterey, CA, USA, pp.175-184. <https://doi.org/10.1109/PLANS.2018.8400000>

- org/10.1109/PLANS.2018.8373379
- Daffertshofer, A., Lamoth, C. J. C., Meijer, O. G., & Beek, P. J. 2004, PCA in studying coordination and variability: a tutorial, *Clinical biomechanics*, 19, 415-428. <https://doi.org/10.1016/j.clinbiomech.2004.01.005>
- Hansen, R., Wind, R., Jensen, C. S., & Thomsen, B. 2010, Algorithmic strategies for adapting to environmental changes in 802.11 location fingerprinting, in *Proc. IEEE Indoor Positioning and Indoor Navigation (IPIN)*, 15-17 September 2010, Zurich, Switzerland, pp.1-10. <https://doi.org/10.1109/IPIN.2010.5648270>
- Jin, G. Y., Lu, X. Y., & Park, M. S. 2006, An indoor localization mechanism using active RFID tag. *International Conference on Sensor Networks, Ubiquitous, and Trustworthy Computing (SUTC'06)*, IEEE, 05-07 June 2006, Taichung, Taiwan. <https://doi.org/10.1109/SUTC.2006.1636157>
- Kaemarungsi, K. 2005, Efficient design of indoor positioning systems based on location fingerprinting, IEEE, In *2005 International Conference on Wireless Networks, Communications and Mobile Computing*, 13-16 June 2005, Maui, HI, USA, pp.181-186. <https://doi.org/10.1109/WIRLES.2005.1549406>
- Kaune, R. 2012, Accuracy studies for TDOA and TOA localization, *2012 15th International Conference on Information Fusion*, 09-12 July 2012, Singapore, pp.408-415.
- Keller, J. M., Gray, M. R., & Givens, J. A. 1985, A fuzzy k-nearest neighbor algorithm, *IEEE transactions on systems, man, and cybernetics*, SMC-15, 580-585. <https://doi.org/10.1109/TSMC.1985.6313426>
- Leu, J.-S. & Tzeng, H.-J. 2012, Received signal strength fingerprint and footprint assisted indoor positioning based on ambient Wi-Fi signals, in *Proc. IEEE Veh. Technol. Conf. (VTC Spring)*, 06-09 May 2012, Yokohama, Japan, pp.1-5. <https://doi.org/10.1109/VETECS.2012.6239883>
- Lu, C., Uchiyama, H., Thomas, D., Shimada, A., & Taniguchi, R. 2019, Indoor Positioning System Based on Chest-Mounted IMU, *Sensors*, 19, 420. <https://doi.org/10.3390/s19020420>
- Mainetti, L., Patrono, L., & Sergi, I. 2014, A survey on indoor positioning systems, IEEE, *22nd international conference on software, telecommunications and computer networks (SoftCOM)*, 17-19 September 2014, Split, Croatia, pp.111-120. <https://doi.org/10.1109/SOFTCOM.2014.7039067>
- Mendoza-Silva, G. M., Costa, A. C., Torres-Sospedra, J., Painho, M., & Huerta, J. 2022, Environment-aware regression for indoor localization based on WiFi fingerprinting, *IEEE Sensors Journal*, 22, 4978-4988. <https://doi.org/10.1109/JSEN.2021.3073878>
- Mika, S., Ratsch, G., Weston, J., Scholkopf, B., & Mullers, K. R. 1999, Fisher discriminant analysis with kernels, *Neural networks for signal processing IX: Proceedings of the 1999 IEEE Signal Processing Society Workshop (Cat. No.98TH8468)*, 25 August 1999, Madison, WI, USA. <https://doi.org/10.1109/NNSP.1999.788121>
- Parlett, B. N. 1998, *The symmetric eigenvalue problem* (Philadelphia, PA: Society for industrial and Applied Mathematics). <https://doi.org/10.1137/1.9781611971163>
- So, J., Lee, J. Y., Yoon, C. H., & Park, H. 2013, An improved location estimation method for WiFi fingerprint-based indoor localization, *International Journal of Software Engineering, and Its Applications*, 7, 77-86. <https://www.earticle.net/Article/A208533>
- Xia, S., Liu, Y., Yuan, G., Zhu, M., & Wang, Z. 2017, Indoor fingerprint positioning based on Wi-Fi: An overview, *ISPRS International Journal of Geo-Information*, 6, 135. <https://doi.org/10.3390/ijgi6050135>
- Xu, J., Ma, M., & Law, C. L. 2008, AOA Cooperative Position Localization, *IEEE GLOBECOM 2008 - 2008 IEEE Global Telecommunications Conference*, 30 November - 04 December 2008, New Orleans, LA, USA, pp.1-5. <https://doi.org/10.1109/GLOCOM.2008.ECP.720>
- Yamasaki, R., Ogino, A., Tamaki, T., Uta, T., Matsuzawa, N., et al. 2005, TDOA location system for IEEE 802.11b WLAN, *IEEE Wireless Communications and Networking Conference*, 13-17 March 2005, New Orleans, LA, pp.2338-2343. <https://doi.org/10.1109/WCNC.2005.1424880>
- Yoo, J.-H. 2020, Semi-supervised Generative Adversarial Network for Wi-Fi Fingerprint Based Indoor Location Awareness, *Journal of Institute of Control, Robotics and Systems*, 26, 1116-1121. <https://doi.org/10.5302/J.ICROS.2020.20.0151>
- Zhuang, Y., Yang, J., Li, Y., Qi, L., & El-Sheimy, N. 2016, Smartphone-based indoor localization with Bluetooth low energy beacons, *Sensors*, 16, 596. <https://doi.org/10.3390/s16050596>



Donggyu Lee received the B.S. degree in the Department of Electronic Engineering, Hankyong National University, Gyeonggi-do, in 2021. He is currently a master's student in the Department of Future Convergence Technology Engineering, Sungshin Women's University. His research interests are machine learning and indoor localization.



Jaehyun Yoo received the Ph.D. degrees in the School of Mechanical and Aerospace Engineering, Seoul National University, Seoul, in 2016. He was a postdoctoral researcher at the School of Electrical Engineering and Computer Science, KTH Royal Institute of Technology, Stockholm, Sweden. He is currently a Professor at the School of AI, Sungshin Women's University. His research interests include machine learning, indoor localization, automatic control, and robotic systems.