Comparative Evaluation of Sensor-Based PDR and Visual SLAM for Smartphone-Based Indoor Positioning

Philipp Fiedler[1]

[1]*Federal Agency for Cartography and Geodesy (BKG), Richard-Strauss-Allee 11, 60598 Frankfurt a.M., Germany*

Abstract

In complex indoor environments, reliable positioning without Global Navigation Satellite Systems (GNSS) coverage poses a significant challenge, particularly for emergency responders operating in areas without installed building technology (e.g. Wi-Fi, BLE, UWB). This paper presents a comparative study of two infrastructure-free indoor localization approaches, both implemented on standard smartphones: inertial Pedestrian Dead Reckoning (PDR) and vision-based Simultaneous Localization and Mapping (SLAM). Each prototype is evaluated based on several criteria, including positioning accuracy, environmental dependencies, hardware requirements, and usability. While the PDR system offers robust, low-power tracking independent of visual conditions, the SLAM-based system achieves higher precision and supports augmented reality (AR) navigation under favorable lighting. A structured comparison is presented to highlight the strengths, limitations, and potential use cases for each approach, based on practical evaluations. The findings offer actionable insights for the future development of indoor positioning systems tailored for first responders in GNSS-denied environments.

Keywords

Pedestrian Dead Reckoning, Visual SLAM, Smartphone-based positioning, Infrastructure-free indoor positioning, augmented reality

1. Introduction

1.1. Motivation and Problem Statement

The positioning of first responders in complex indoor environments poses a particular challenge due to the insufficient strength of GNSS signals in such environments. Furthermore, the buildings in question are often unfamiliar upon their initial entry, and relying on preinstalled building infrastructure, such as Wi-Fi access points or Bluetooth beacons, is not feasible. Consequently, the Federal Agency for Cartography and Geodesy (BKG) is engaged in research to develop systems that will empower emergency services to determine their location in areas lacking GNSS coverage. Instead of existing location-based infrastructure, compact inertial navigation systems (INS) are implemented, which utilize integrated sensors - such as accelerometer, gyroscope, and magnetic field sensors — to detect the movement of the emergency responder within the room and derive their current position [1]. In addition, research is being conducted on camera-based systems that use Simultaneous Localization and Mapping (SLAM) to determine position based on visual information and a stored 3D map. The objective of this study is to provide a conceptual and practical comparison of both approaches in terms of their suitability for infrastructure-free indoor positioning.

1.2. Contribution of this Paper

This paper presents a comparative study of two infrastructure-independent approaches to indoor localization, both implemented as smartphone-based prototypes. The key contributions of this work are as follows:

IPIN-WCAL 2025: Workshop for Computing & Advanced Localization at the Fifteenth International Conference on Indoor Positioning and Indoor Navigation, September 15–18, 2025, Tampere, Finland
philipp.fiedler@bkg.bund.de (P. Fiedler)
https://www.bkg.bund.de/deniedAreas/ (P. Fiedler)

- Implementation of two distinct prototypes for infrastructure-free indoor positioning: one based on inertial Pedestrian Dead Reckoning (PDR), the other on visual SLAM.
- Qualitative comparison of both approaches on positioning accuracy, drift behavior, robustness, sensor requirements, and environmental dependencies.
- Identification of practical strengths and limitations of each approach based on real-world usage scenarios and developer observations.
- Presentation of a structured comparison table to support decision-making in future applications, such as emergency response, indoor navigation, or infrastructure-free localization.

The goal is to derive practical insights into the applicability of PDR and SLAM-based localization methods under constrained conditions, such as those encountered by first responders in unfamiliar indoor environments without prior mapping or network availability.

2. Related Work

2.1. Overview of Sensing Modalities

Table 1 presents a comparison of commonly used sensors and technologies for indoor localization, as compiled in [2]. It highlights key characteristics, including the availability on accuracy, cost, strengths, and weaknesses of different sensors.

Table 1
Comparison of Sensors for Indoor Positioning Systems. Adapted and extended from [2].

Sensor	Accuracy	Cost	Advantages	Disadvantages
WiFi	5–10m	\$10–\$100	Uses existing infrastructure, supported by smartphones	Low accuracy depending on signal strength and interference
BLE	5–10m	\$10– per 100m^2	Smartphone-compatible and energy-efficient	Requires beacon installation and has limited range
LiDAR	cm–dm	\$1,000–\$10,000	Very high measurement accuracy, dense point clouds	High cost, high power consumption, limited device support
Camera	cm–dm	\$10–\$100	Low cost, colored point clouds, available in all smartphones	Sensitive to lighting conditions and motion blur
HD Map	cm	\$100–\$1,000	High resolution and accuracy, includes infrastructure layers	Requires up-to-date maps, costly to create
5G	cm–100m	\$100–\$1,000	Uses existing network and is supported by smartphones	Accuracy varies with signal strength; signal loss indoors
INS	2% of distance	\$10–\$100	Fully self-contained with full motion state estimation	Drift over time, high-end sensors are expensive
GNSS	4–60m	\$1,000	Global positioning, smartphone support	Poor indoor coverage and inaccurate altitude estimation
UWB	2–50cm	\$100 per 100m^2	High accuracy in distance measurement	Requires infrastructure, suffers from signal issues

This comparison underscores the trade-offs between different sensing technologies. Inertial sensors such as accelerometers and gyroscopes are universally available but suffer from cumulative drift. Vision-based approaches offer high spatial accuracy, but are sensitive to lighting and require significant computational resources.

2.2. Sensor Fusion

One possibility that has been extensively discussed in the literature as a way of eliminating the aforementioned disadvantages of individual sensors is to combine different sensors within a single application [2], [3], [4]. The following figure illustrates that sensor fusion has a significant impact on the accuracy of indoor positioning with a smartphone.

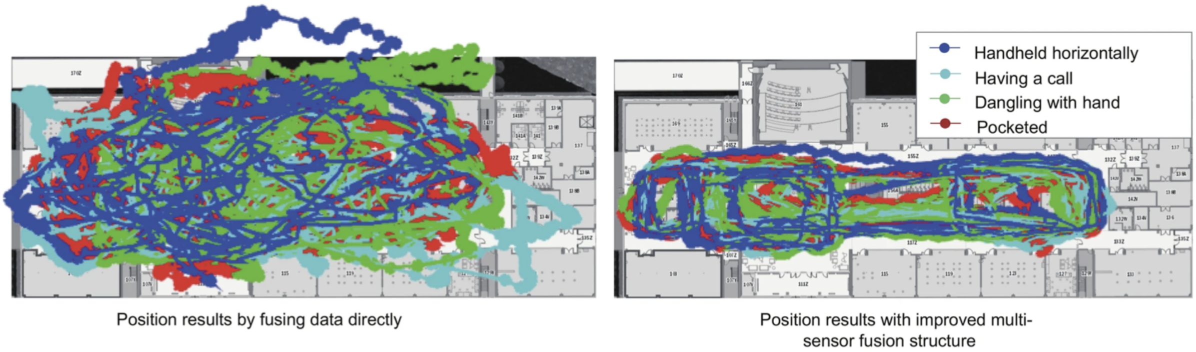

Figure 1: Effect of sensor fusion in terms of accuracy [4]

Kalman filters, particle filters, and machine learning-based models are commonly used to fuse sensor data, compensate for drift, and estimate user orientation and movement [5]. Particularly in PDR systems, fusion of accelerometer, gyroscope, and magnetometer data is essential for estimating heading and step length [6]. In vision-based systems, inertial data can support visual tracking and aid in loop closure detection [7].

This paper builds on these foundations by evaluating two sensor fusion-based prototypes that follow fundamentally different approaches: one primarily inertial (PDR) and one vision-based (SLAM), both implemented on commodity smartphones.

3. System Design and Implementation

3.1. System overview

In order to enable optimal indoor positioning that is independent of GNSS, the Internet, or known infrastructures such as Bluetooth beacons, two prototypes were developed based on different approaches: sensor-based PDR and visual SLAM. The two systems were developed for Android smartphones and are designed to be as universally applicable as possible in all buildings without external infrastructure. While the PDR prototype focuses solely on real-time positioning (Figure 2 A), the Visual SLAM prototype additionally provides navigation capabilities, offering a AR-based interface with route guidance for the user (Figure 2 B and C).

3.2. Prototype A: Sensor-Based PDR

3.2.1. Initialization and Heading Determination

Despite the challenges posed by the use of GNSS within structures due to the attenuation of signal strength, it remains a viable method to determine the initial starting point, if the user starts the application outside of the building. The system performs a variety of checks to assess the quality of the GNSS signal for positioning purposes, including the Carrier-to-Noise Ratio (CNR), the number of visible satellites, and the reported GNSS accuracy. If the user initiates the process within the building or encounters a disturbance in the GNSS signal, manual entry of a starting point is allowed. To determine the initial heading, the system uses a magnetometer to obtain an absolute orientation relative to the magnetic north. Given the potential for interference signals in this environment, GNSS direction or manual alignment to the north may be considered as alternatives for initial heading determination.

(A) (B) (C)

Figure 2: Screenshots of the developed prototypes: (A) Sensor-based PDR; (B) and (C) Visual SLAM.

3.2.2. Step Detection and Dead Reckoning

Once the initial pose (position and heading) is established, step detection is performed using accelerometer data. The step length $s(t)$ is dynamically estimated using the Weinberg model [8]. For each detected step, acceleration values are collected over the step interval. The difference between the maximum acceleration $a_{\text{peak}}(t)$ and the minimum acceleration $a_{\text{min}}(t)$ within this interval is then raised to the fourth root and multiplied by an empirical scaling factor k, which typically ranges between 0.25 and 0.5 depending on user characteristics such as body height and step frequency:

$$s(t) = k \cdot \sqrt[4]{a_{\text{peak}}(t) - a_{\text{min}}(t)} \tag{1}$$

The heading $\varphi(t)$ is continuously updated using fused gyroscope and magnetometer data. Using these values, the local coordinate differences for each step can be calculated starting from the initial position using formulas (2) and (3).

$$\Delta x = s(t) * \sin(\varphi(t)) \tag{2}$$

$$\Delta y = s(t) * \cos(\varphi(t)) \tag{3}$$

These differences are then converted into global coordinates:

$$\text{latitude} = \frac{\Delta y}{111{,}229} \tag{4}$$

$$\text{longitude} = \frac{\Delta x}{111{,}320 \cdot \cos\left(\frac{\text{latitude} \cdot \pi}{180}\right)} \tag{5}$$

The system also estimates the user's vertical position (floor level) by comparing the height delta to the starting point using the barometer. For this computation, a constant floor height of 3 m can be assumed.

3.2.3. Drift Correction and User Interface

To mitigate drift and cumulative positioning errors, several correction mechanisms are employed. The system continuously scans for usable GNSS signals. In some indoor environments, signal quality is sufficient to apply GNSS-based corrections. Additionally, users may temporarily leave the building, providing an opportunity for re-initialization via satellite data. Magnetometer readings are used to compensate gyroscope drift. Various filters determine which sensor should be given more weight based on magnetic field strength.

The smartphone interface shows the user's current position and orientation on an OpenStreetMap background, along with a GNSS accuracy circle. The minimum acceptable GNSS accuracy threshold can be adjusted using the slider at the bottom of the screen. Users also have the option to track and save their traveled path as GeoJSON. The operations center receives this information at 1 Hz intervals via WiFi Websockets, enabling real-time monitoring of the emergency responder's current position.

3.3. Prototype B: Visual SLAM-Based Positioning

3.3.1. Initialization via QR-Code

In contrast to sensor-based approaches, the Visual SLAM system utilizes a QR-Code scan to initialize the user's pose. The QR-Code is generally positioned at the building entrance, encoding both the absolute position and the orientation of the smartphone at the moment of scanning. This approach ensures a precise and unambiguous starting pose, which is essential for global map alignment and consistent localization in the building.

3.3.2. SLAM algorithm

After determining the initial pose, the system utilizes Google's augmented reality SDK ARCore to detect and track visual feature points from the camera stream. These features, including corners, edges, and texture regions, are extracted at a frequency of 60 Hz and tracked over time to estimate the smartphone's motion through space. When the same feature is observed across multiple frames from different viewpoints, its 3D position is triangulated.

To improve motion estimation, ARCore fuses data from the device's gyroscope, which helps infer rotation and movement direction, especially during rapid motion or in low-feature environments. The estimated trajectory is then aligned with a preloaded three-dimensional reference map of the building to provide global positioning. The 3D map can be derived from evacuation and emergency plans, which are available to the authorities for all public buildings and must be continually updated, making them ideal for SLAM-based navigation and matching. In the current implementation, this conversion is performed manually in Unity3D, a process that can take up to 30 minutes. Subsequently, a topological representation (NavMesh Surface) is automatically generated using Unity's AI Navigation tool.

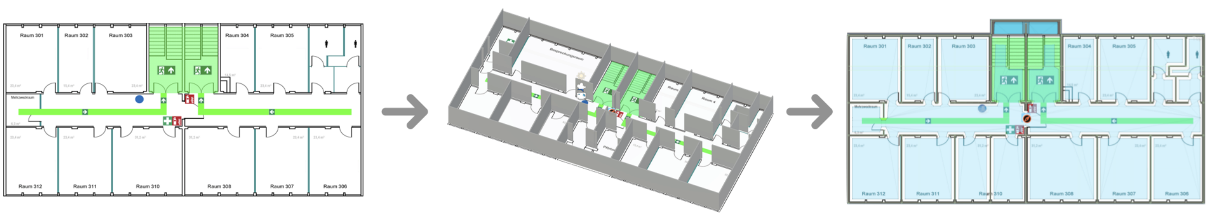

Figure 3: 2D emergency plan to 3D map to topological map conversion

The system is supported by additional sensors, including accelerometer for motion detection and a barometer for altitude detection, similar to prototype A. When the system revisits a previously recorded location, a process known as loop closing is initiated to address accumulated errors resulting from sensor noise or drift. This objective is accomplished by identifying and re-matching features that were recorded during the initial stages of the SLAM process.

3.3.3. Indoor Navigation and User Interface

Beyond mere positioning, the system provides comprehensive indoor navigation capabilities. According to the present position and the user-defined destination, the system continuously performs computations using an A* algorithm to ascertain the shortest path. In this process, multiple environmental constraints (walls, rooms, impassable areas) are taken into account, which are defined in the topological map.

Navigation instructions are displayed in Augmented Reality (AR), overlaid directly onto the smartphone camera view. Users have the option to select different routing preferences, such as avoiding stairs or choosing between two visualization modes (arrow-based overlay vs. floating path line). A minimap is displayed alongside the AR view to facilitate user orientation and provide an overview of the current route within the building structure.

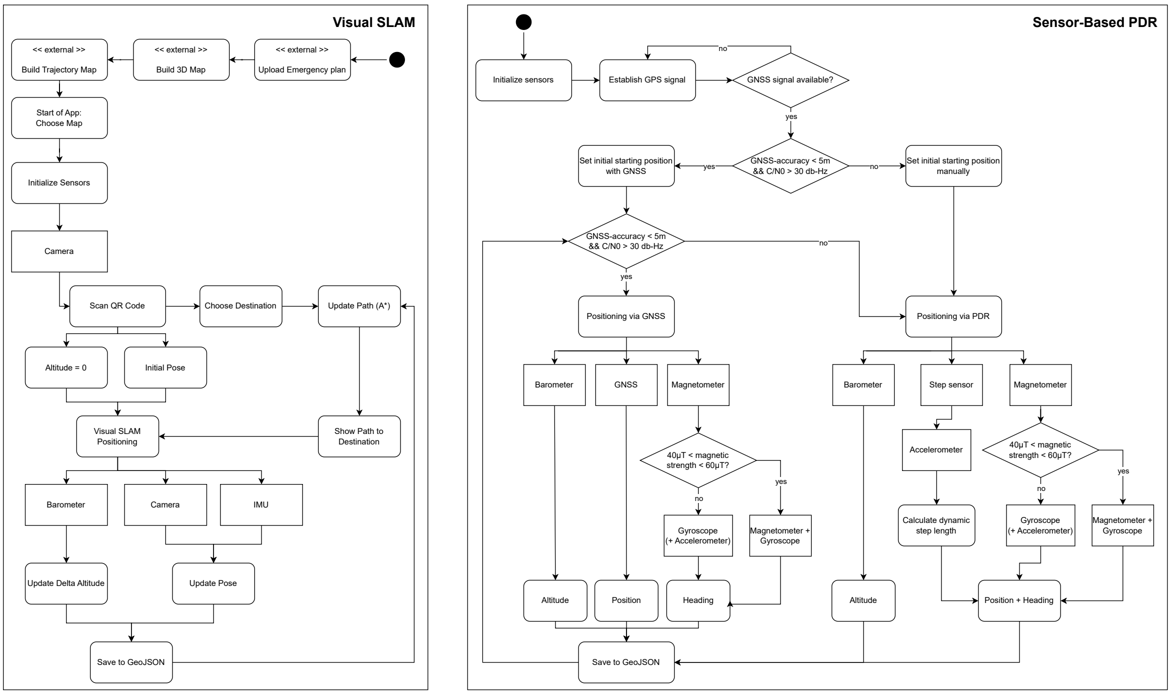

Figure 4: System Architecture of both prototypes

4. Comparative Performance Summary

Both approaches have been developed as prototypes and tested under realistic indoor conditions. This has provided valuable insights into the respective strengths and weaknesses of the two approaches, which will help to define the potential use cases and deployment scenarios for sensor- and vision-based indoor positioning.

The evaluation was conducted across diverse environments, including residential buildings, office complexes, and large-scale structures such as football stadiums. The predefined multi-level trajectories, varying in length from approximately 50 meters to several hundred meters, were traversed using both systems. At each step, the estimated position was recorded. These recorded positions were compared against the reference trajectory, and the mean positional accuracy was calculated. For a fair comparison, both systems were carried in a normal usage position, similar to how a smartphone is typically held during navigation tasks (e.g., in front of the user or in hand at waist level).

4.1. Sensor-Based PDR: Universal and Low-Cost

The primary advantage of the PDR-based system is its universality, which renders it applicable in the absence of any prior knowledge of the building environment. The system operates independently of infrastructure and can be deployed in nearly any indoor setting. Additionally, the smartphone does not need to be held in the user's hand: Due to its robustness against poor lighting or visual obstructions, the placement of the device in a pocket or on the body of the first responder is sufficient.

The PDR algorithm's low computational requirements and energy consumption make it suitable for execution on low-power microcontrollers (e.g. ESP32) equipped with low-cost inertial and barometric sensors. The BKG has already implemented several such prototypes.

However, PDR is susceptible to cumulative drift over time, particularly in scenarios where reliable GNSS or magnetometer data is not available. In such cases where only the gyroscope is utilized for heading estimation, the resulting drift is approximately 2% of the distance. The system's performance is contingent upon the precise initialization of both position and heading. Without a fix, even small initial errors can lead to significant positional errors as the user moves, leading to a mean accuracy tested of about 5m.

4.2. Visual SLAM: Accurate but Resource-Intensive

The visual SLAM approach addresses many of the limitations of PDR through continuous visual re-localization and loop closing, significantly reducing long-term drift. Under optimal visual conditions, characterized by a light intensity greater than 10 lux, the camera-based system achieves high accuracy of around 60 cm and is expected to be less sensitive to initialization errors, as the initial pose is determined via a fixed QR code position rather than GNSS.

However, the system is sensitive to low-light conditions and motion blur, which can impair feature tracking and stability. Moreover, measurements of processing power utilization and battery consumption indicate that visual SLAM demands approximately 3.3 times greater computational power and about 4.5 times higher energy consumption. For optimal performance, the smartphone must be actively held in the user's hand in a way that ensures an unobstructed camera view.

4.3. Summary Table of Key Differences

The following table 2 summarizes the key characteristics of both systems in direct comparison:

Table 2
Comparison of Sensor-Based PDR and Visual SLAM Approaches

Aspect	Sensor-Based PDR	Visual SLAM
Initialization	GNSS or manual input	QR code with known pose
Positioning Accuracy	~5 m (drift over time)	~0.6 m (under good lighting conditions)
Drift Handling	GNSS & Magnetometer fix	Loop closing reduces long-term drift
Lighting Dependence	Independent of lighting	Requires sufficient lighting (>10 lux)
Sensor Requirements	IMU, barometer	Camera, IMU, barometer
Processing Load	Low; suitable for microcontrollers	High; requires modern smartphone
Device Handling	Can be used in pocket or hand	Must be held in hand
Infrastructure Requirements	None	QR code at entry, escape plan / 3D map
Preparation Time	None	About 30 minutes (3D / trajectory map)
Navigation Support	Not available	AR navigation with A* pathfinding
Use Case Suitability	Universal deployment without prior map	Accurate navigation in known buildings

5. Discussion

5.1. Suitability for Different Use Cases

A comparison of the two prototypes reveals the distinct strengths of PDR and visual SLAM in terms of operational applicability. Depending on the specific requirements of the respective operation, each approach lends itself to distinct use cases.

The PDR-based system is particularly well suited for spontaneous operations where minimal setup time is available and first responders may not be able to hold a smartphone in their hands. In many of these scenarios, room-level accuracy (approximately 5 m) is sufficient to track and coordinate personnel from a command center. The following list contains some of the more common use cases:

- Tracking and coordination of first responders during dynamic or ad hoc operations
- Operations in darkness or low-light conditions (e.g., night-time or smoke-filled environments)
- Situations where responders require hands-free operation

In contrast, the camera-based SLAM approach is better suited for pre-planned deployments in public or complex buildings, where sufficient preparation time is available to generate a 3D map from official building evacuation plans. The system's enhanced accuracy and AR navigation capabilities support the following scenarios:

- Post-analysis and documentation of tactical training exercises
- Pre-mission preparation in structured public environments (i.e., train stations, malls, government buildings)
- Real-time indoor navigation and tracking during static or longer-term operations

5.2. Potential for Hybrid Integration

In order to eliminate the disadvantages of both systems as much as possible, research could be conducted in the future on a hybrid system that combines the advantages of both approaches. In conditions that allow for optimal visual perception, the potential system could leverage visual SLAM for high-precision positioning and use it to correct accumulated drift inherent in inertial-based tracking. In scenarios where visual tracking becomes unreliable due to occlusions, darkness, or motion blur, the PDR system could seamlessly take over, maintaining continuous position estimations in the absence of visual input.

This fusion would enable more robust and adaptive indoor localization, particularly in complex environments with rapidly changing conditions. Developing such a hybrid system would require careful synchronization of sensor data and intelligent switching or fusion strategies, but holds significant potential for improving overall reliability, particularly in mission-critical scenarios.

6. Conclusion and future work

This paper presented a comparative analysis of two infrastructure-independent approaches to smartphone-based indoor localization: one based on inertial PDR and the other on visual SLAM. While the PDR prototype demonstrated robustness and low dependency on environmental factors, the visual SLAM system offered significantly higher accuracy under favorable conditions.

The comparison indicated that each system is best suited to different operational contexts, ranging from hands-free emergency deployments to high-precision navigation in structured environments. A hybrid integration of both approaches could further improve adaptability and reliability, especially in dynamic or degraded settings.

Future work will focus on further improving both systems and exploring a hybrid solution that combines their respective strengths. The ongoing development includes a lightweight, ESP32-based version of the PDR system, investigations into alternative positioning technologies such as DAB+, and integration of LoRa for remote transmission of positional data. The PDR algorithm is undergoing enhancements that include improved dynamic step length estimation based on [9] and advanced sensor fusion. The SLAM-based prototype may also benefit from automated 3D map generation using computer vision techniques applied to official evacuation plans.

Declaration on Generative AI

During the preparation of this work, the author(s) used DeepL Write in order to: Grammar and spelling check. After using these tool, the author reviewed and edited the content as needed and takes full responsibility for the publication's content.

References

[1] BKG, Forschung - Denied Areas – Indoor-Positionierung, 2021. URL: https://www.bkg.bund.de/DE/Forschung/Projekte/DeniedAreas-IndoorPositionierung/DeniedAreas-IndoorPositionierung_cont.html.

[2] N. El-Sheimy, Y. Li, Indoor navigation: state of the art and future trends, Satellite Navigation 2 (2021) 7. URL: https://doi.org/10.1186/s43020-021-00041-3. doi:10.1186/s43020-021-00041-3.

[3] M. Elsanhoury, P. Mäkelä, J. Koljonen, P. Välisuo, A. Shamsuzzoha, T. Mantere, M. Elmusrati, H. Kuusniemi, Precision Positioning for Smart Logistics Using Ultra-Wideband Technology-Based Indoor Navigation: A Review, IEEE Access 10 (2022) 44413–44445. URL: https://ieeexplore.ieee.org/document/9761257?denied=. doi:10.1109/ACCESS.2022.3169267, conference Name: IEEE Access.

[4] Y. Li, Y. Zhuang, P. Zhang, H. Lan, X. Niu, N. El-Sheimy, An improved inertial/wifi/magnetic fusion structure for indoor navigation, Information Fusion 34 (2017) 101–119. URL: https://www.sciencedirect.com/science/article/pii/S1566253516300574. doi:10.1016/j.inffus.2016.06.004.

[5] M. Yang, X. Sun, F. Jia, A. Rushworth, X. Dong, S. Zhang, Z. Fang, G. Yang, B. Liu, Sensors and Sensor Fusion Methodologies for Indoor Odometry: A Review, Polymers 14 (2022) 2019. URL: https://www.mdpi.com/2073-4360/14/10/2019. doi:10.3390/polym14102019, number: 10 Publisher: Multidisciplinary Digital Publishing Institute.

[6] H. Mehrabian, R. Ravanmehr, Sensor fusion for indoor positioning system through improved RSSI and PDR methods, Future Generation Computer Systems 138 (2023) 254–269. URL: https://www.sciencedirect.com/science/article/pii/S0167739X22002874. doi:10.1016/j.future.2022.09.003.

[7] R. Mur-Artal, J. M. M. Montiel, J. D. Tardos, ORB-SLAM: a Versatile and Accurate Monocular SLAM System, IEEE Transactions on Robotics 31 (2015) 1147–1163. URL: http://arxiv.org/abs/1502.00956. doi:10.1109/TRO.2015.2463671, arXiv:1502.00956 [cs].

[8] H. Weinberg, Using the adxl202 in pedometer and personal navigation applications, Analog Devices AN-602 application note 2 (2002) 1–6.

[9] N.-H. Ho, P. H. Truong, G.-M. Jeong, Step-detection and adaptive step-length estimation for pedestrian dead-reckoning at various walking speeds using a smartphone, Sensors 16 (2016) 1423.