International Journal of Robotics and Automation, Vol. 39, No. 10, 2024
An Open Access Paper

VISION-BASED ROBOT INDOOR-POSITIONING AND NAVIGATION METHOD RESEARCH

Junfu Qiao*,** Jinqin Guo*,** and Yongwei Li*,**

Abstract

Indoor positioning and navigation of robots have become increasingly important for various applications, including manufacturing, logistics, and home automation. In this study, we propose a vision-based method for robot indoor positioning and navigation. Our approach utilises computer vision techniques to extract features from the environment and estimate the robot's position and orientation. Key facets explores encompasses visual sensing, feature extraction, mapping, localisation, path planning, and control strategies. An emphasis is placed on how theme components synergise to form a holistic vision-based approach tailored for indoor robot navigation. The integration of theme methods with existing robotic platforms is underscored, facilitating their seamless incorporation into a wide array of applications. This review paper consolidates the current state of knowledge, portraying vision-based indoor navigation as a transformative technology poised to revolutionise robotic operations within indoor environments.

Key Words

Robot, accuracy, indoor

1. Introduction

In recent years, propelled by remarkable strides in sensor technology, chip development, and artificial intelligence, robots have undergone a transformative evolution, emerging as remarkably intelligent and adaptable entities [1]–[4]. This progression has ignited a widespread integration of robots across an array of sectors, each demonstrating the potential of these machines. For instance, Zhu [5] engineered a cutting-edge robotic system, replete with pliable manipulators, stereo cameras, and three-dimensional (3D) vision capabilities, specifically tailored for handling delicate fabrics [5].

Venturing into the realm of healthcare, Anthon, a forward-thinking establishment rooted in Pittsburgh, Philadelphia, harnessed innovation to conceive a mobile medical robot, poised to revolutionise patient care within wards. This ingenious creation deftly manages tasks like ferrying essential sustenance, precisely dispensing prescribed medications, and ensuring timely delivery of vital medical provisions to nursing stations [6].

The hospitality sector, too, has warmly embraced the robotic wave, witnessing a proliferation of automaton assistance in sundry functions, spanning from meticulous cleaning to efficient item transportation, adept information dissemination, seamless order processing, and secure payment handling, thus enhancing guest experiences within hotel premises. Parallel, the literature delves into the realm of culinary robotics, where a novel fusion of ultra wideband positioning technology, an odometer, a cost-effective gyroscope, and an accelerometer converge to give rise to a trackless, navigation-enabled food delivery robot, as detailed in reference [7]–[9].

Depicted in Fig. 1, an illustrative array of application scenarios showcases the integral role of factory handling robots, hospital medical robots, hotel service robots, and restaurant delivery robots. An overarching narrative underscores the pivotal significance of equipping these automated entities with advanced positioning and navigation capabilities, enabling them to not only comprehend their present coordinates but also seamlessly navigate and fulfil assigned tasks across the diverse landscapes they traverse [7].

The widely used global positioning system (GPS) Saracoglu and Sanli [10] is renowned for its exceptional outdoor positioning accuracy. While civilian applications achieve a precision of around 3 m, military use can attain an astonishing 0.3 m accuracy. GPS technology finds extensive employment in the navigation and positioning of mobile objects, such as vehicles, ships, aircraft, and handheld devices. Nonetheless, the challenge arises when applying GPS to indoor environments, particularly complex ones, where the majority of robots operate. The signal strength of GPS weakens significantly upon encountering indoor

* Department of Automation, Taiyuan Institute of Technology University, Taiyuan, China 030008
** Intelligent Detection and Control Engineering (Technology) Research Center, Taiyuan, China 030008; e-mail: qiaojunfu@163.com; guojq@tit.edu.cn; liyongwei27@163.com
Corresponding author: Junfu Qiao

Recommended by Jianjun Ni
(DOI: 10.2316/J.2022.206-1059)

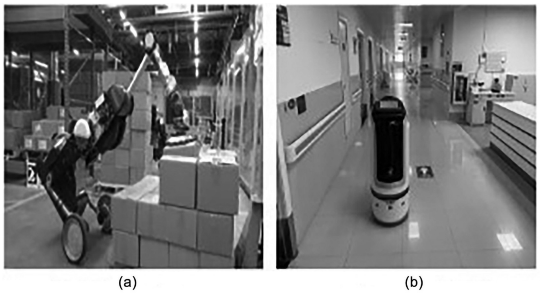


Figure 1. Robots in different fields: (a) factory handling robot and (b) hospital medical robot.

structures and diverse objects, consequently leading to a sharp rise in positioning errors [11]. Thus, a pressing need exists to identify an indoor positioning technology that combines high accuracy and affordability. Integrating such a technology into robot indoor positioning and navigation holds significant research value.

The proliferation of wireless local area networks has rendered Wi-Fi devices omnipresent within indoor spaces. Exploiting indoor Wi-Fi signals to locate robots presents several advantages, including cost-effectiveness, precise positioning, extensive coverage, robust indoor penetration, and strong communication capabilities. Nevertheless, the existing traditional Wi-Fi fingerprint-based positioning algorithms often fall short in delivering the required precision for indoor robot positioning. Addressing this, this article enhances the conventional Wi-Fi fingerprint-based positioning algorithm to meet the stringent indoor robot positioning accuracy criteria. To ensure seamless navigation of robots to designated points, the author proposes a grid-based navigation algorithm that leverages the characteristics of the enhanced positioning approach. This navigation algorithm stands out from the majority of counterparts due to its simplicity in implementation, reduced computational demands, and elevated accuracy [6].

Mobile robots have the potential to replace humans in dangerous tasks, making them significant in areas, such as military reconnaissance, anti-terrorism, explosion proof, and anti-nuclearisation. They also have numerous applications in daily life, such as guiding visitors in large museums to find exhibits they are interested in. To achieve this, mobile robots would require a highly portable and high-frequency indoor positioning system (IPS). Simultaneous localisation and mapping (SLAM) is applied for self-reliant navigation and indoor positioning of the robot. SLAM involves the use of various sensors, such as vision, laser, and odometer to help the robot locate itself and build maps in unfamiliar environments. Its cooperative operation of sensors results in relatively high positioning accuracy [12].

However, while monocular visual SLAM is simple and inexpensive, its reliability is poor. Binocular visual SLAM solves this problem but is expensive, complicated in design, and limited by distance. Achieving real-time positioning is difficult due to computational requirements, and flexibility is also reduced. Indoor positioning and navigation of robots is a rapidly developing field with numerous potential applications in various fields. Among the different approaches to robot positioning and navigation, vision-based methods have attracted big interest in latest years because of their excessive accuracy and efficiency. Vision-primarily based totally indoor positioning and navigation techniques use cameras and different imaging gadgets to recognise and locate objects in the environment, allowing robots to navigate autonomously in unfamiliar indoor environments. The SLAM technique is widely used in vision-based indoor positioning and navigation methods. It involves constructing a map of the environment while simultaneously determining the robot's location within that map [13]. By integrating data from different sensors, such as vision, laser, and odometer, SLAM enables accurate and robust positioning and navigation of robots in indoor environments.

Despite the many advantages of vision-based indoor positioning and navigation methods, there are still some challenges that need to be addressed. These include poor lighting conditions, occlusion, and complex indoor environments. Researchers are constantly developing new techniques and algorithms to overcome these challenges and improve the accuracy and reliability of vision-based indoor positioning and navigation systems. The development of vision-based indoor positioning and navigation methods has significant potential to revolutionise industries, such as manufacturing, healthcare, and retail. With the increasing demand for more efficient and automated processes, the use of vision-based indoor positioning and navigation methods is likely to become even more widespread in the future [14].

2. Literature Review

In recent years, extensive research has been dedicated to indoor location technologies. Technologies like UWB, RF, Zigbee, Bluetooth, and others have been proposed, but they often require the installation of additional devices, which can disrupt people's daily lives. As a result, WiFi-based indoor location technology has garnered significant attention due to its ability to utilise existing Wi-Fi infrastructure while providing highly accurate real-time indoor location information [15].

SLAM has become a widely adopted technique for autonomous navigation and location of robots, particularly in unfamiliar environments. These applications often rely on a combination of sensors, such as lasers, cameras, and odometers. The popularity of SLAM in indoor robot applications lies in its ability to accurately locate robots even in unfamiliar settings, with positioning accuracy reaching the centimeter level. Visual SLAM has played a pivotal role in creating indoor maps, enabling robots to operate without the needs for specialised localisation infrastructure. However, visual SLAM does have stability challenges, and binocular vision SLAM, while addressing some of these issues, can be cost-prohibitive. Additionally, environmental factors like lighting conditions and line-of-sight (LOS) obstructions pose significant challenges to visual SLAM [16].

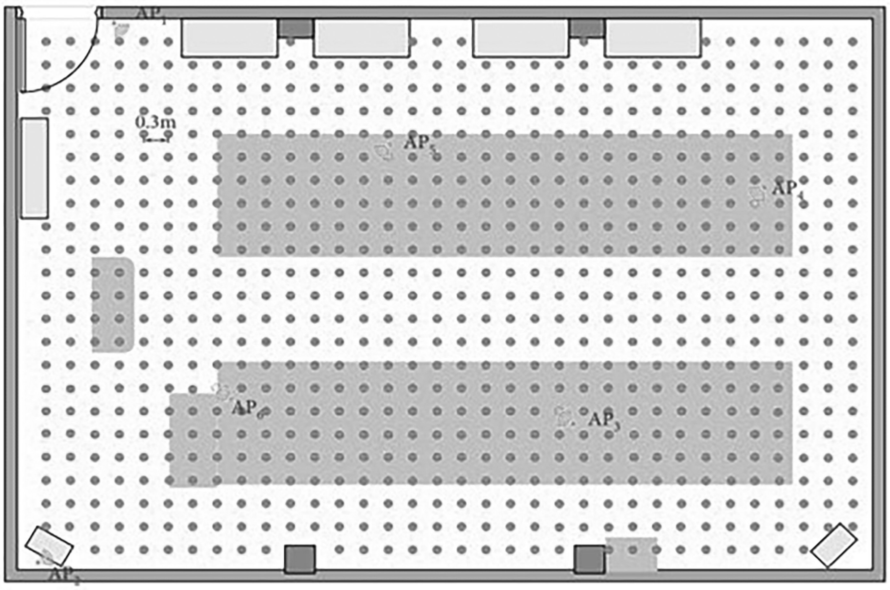


Figure 2. Fingerprint based algorithm for indoor positioning.

Fingerprint-based methods have also been extensive explores to achieve high-accuracy indoor localisation. In this approach, each location within an indoor environments is characterised, and the mobile device's position is determined through nearest neighbour matching. Various techniques have been integrated into fingerprinting, including Bayesian inference, K-nearest neighbour (KNN) inference, and machine learning methods, such as neural networks, back-propagation (BP), support victor machine for regression, compressed sensing, factor graphs, kernel estimation, and more [15].

Notably, One machine learning algorithm, extreme learning machine (ELM), has gained significant attention due to its rapid learning capabilities and ease of implementation. For instance, an RFID-based IPS employing ELM has been reported to offer improved efficiency and localisation accuracy. Online sequential ELM (OS-ELM), which adapts to changing environmental dynamics through online sequential learning, consistently provides higher localisation accuracy compared to traditional approaches.

3. Case Study Research for Indoor Positioning Method

Beijing Institute of Technology researchers have proposed a novel strategy to improve positioning and navigational capabilities. Their approach makes use of the combination of camera and lidar technology to produce a revolutionary 3D semantic map that smoothly combines precise camera and lidar data. This innovative map makes it easier to comprehend traffic scenes in real time. The use of a convolutional neural network (CNN) to perform semantic segmentation of images, thereby creating an environment-oriented semantic map, is a key component of this process. Harmonising time and spatial synchronization is a key component of their method, allowing the synthesis of semantic landmark frames from the point cloud data provided by lidar and camera sensors. The result of this fusion is the development of a semantic map that is based on attitude awareness [17].

A relevant work that delves further into the area of robotic positioning and mapping proposes a novel strategy that combines lidar with an inertial measurement unit. This combination creates a multisensory system that carefully calculates the motion trajectory of a mobile robot using rank Kalman filtering. This trajectory estimation's foundation is formed by the interaction between inertial measurement unit and lidar readings [18].

Indoor-positioning systems rely on devices to determine their own locations, while infrastructure positioning systems estimate device positions using other devices in the environment. In assisted architecture, an external system calculates and sends the user's position in response to a request. IPSs use a range of technologies to determine the positions of indoor environments, along with radio frequency (RF), inertial, audible and non-audible sound, light-primarily based totally, and vision-primarily based totally technologies. This paper classifies IPSs into two categories based on the use of RF signals, and then subdivides them into five groups based on the predominant technology used: RF-based, inertial, sound, light-based, and computer vision-based systems [19]. There is also a sixth category, hybrid systems, that combines two or more technological approaches. The Fig. 2 shows a complete classification of IPSs, excluding hybrid systems due to the many possible combinations that make a general classification infeasible [20].

4. Navigation System

The sensitive and valuable data of user position is widely used in various real-life applications, such as broadcasting, advertising, recommending, and navigation. While outdoor positioning is possible using GPS, indoor positioning, and navigation systems have been a focus of research for decades. However, due to the complexity of the problem, there is no perfect solution yet. To solve this problem, researchers from different fields have come up with various algorithms, such as triangulation, angle of arrival (AOA), and time of flight (TOF). Most current methods rely on hardware deployment using signals, such as Wi-Fi, beacon, or RFID. However, these hardware-dependent systems have several disadvantages, including high construction and maintenance costs, low accuracy, and frequent battery replacements due to energy consumption [21].

When it comes to navigation systems, the standard 2D mode is the norm. However, we believe that there should be a more intuitive way to guide users, such as a virtual tour guide that walks ahead and shows the way. With advances in hardware and algorithm development, heavy computing processes like 3D rendering, 3D reconstruction, visual inertial odometry (VIO), and camera pose estimation can now be performed in real-time on mobile devices. As a result, augmented reality (AR) has garnered significant interest in both research and business domains. Google recently launched an AR outdoor navigation system trial that displays direction arrows, street names, and distances in AR.

Based on our viewpoint, it confirms that our proposal to offer an AR indoor navigation system is a viable one. Additionally, we suggested an AR virtual tour guide to replace traditional navigation methods that use direction arrows. In another research paper, we conducted an experiment and a subsequent survey, which demonstrated that the AR virtual tour guide is more intuitive and practical than the conventional method [22].

5. ELM Algorithm

The ELM algorithm, introduced by Huang, offers a novel approach to tackle the challenges of training single-hidden layer feedforward neural networks (SLFNs). Traditional BP algorithms often demand meticulous tuning of numerous training parameters, leaving them vulnerable to local optima. In contrast, ELM streamlines the training process by employing a large number of hidden layer nodes, each contributing to a unique optimal solution. Furthermore, ELM permits the initialisation of input weights and biases using random functions while upholding learning accuracy. This algorithm distinguishes itself with its speed and applicability in the context of robot localisation [16].

SLFNs serve as the foundational model for feedforward neural networks, striving to approximate optimal training samples. ELM's remarkable speed proves invaluable in solving neural networks that comprise only a single hidden layer. In Fig. 2, we present the network framework. Consider a scenario with M arbitrarily independent samples denoted as (ξ, τ), where $\xi = [\xi 1, \xi 2, \ldots, \mathrm{and}]^{\mathrm{T}} \in \mathbb{R}^{\mathrm{n}}$ and $\tau = [\tau 1, \tau 2, \ldots, 3M]^{\mathrm{T}} \in \mathbb{R}^{\mathrm{m}}$. For a single-hidden layer neural network featuring N hidden nodes, the relationship can be expressed as:

$$\sum i = 1N\lambda\phi(\omega \cdot \xi + b) = \omega, j = 1, 2, \ldots, M$$

Here, λ represents the activation function, $\omega = [\omega 1, \omega 2, \ldots, \omega \mathrm{N}]^{\mathrm{T}}$ stands for the connection weight matrix between the input neuron and the ith layer neuron, $\Phi = [\Phi 1, \Phi 2, \ldots, \Phi \mathrm{N}]^{\mathrm{T}}$ denotes the weight matrix between the ith hidden layer neuron and the output neuron, b signifies the bias of neurons in the hidden layer i, and ω symbolises the corresponding output vector. The term $\omega \cdot \xi$ signifies the inner product of ω and ξ [23].

6. Methodology

The enhances Wi-Fi location fingerprinting algorithm comprises two distinct phases: an offline phase for constructing a comprehensive location fingerprint database and a real-time localisation phase for determining the robot's position. Let's deep into each of these phases with detailed explanations of the improvements made.

- **Offline Phase**

In this initial phase, the indoor environments is systematically divided into equally spaced raster points. Crucially, the spacing between these raster points is smaller compared to traditional algorithms, enhancing the granularity of our database. At each of these grid points (known as references points) within the room, the robot performs Wi-Fi signal acquisition. A notable departure from traditional approaches, the improved algorithm employs varying Wi-Fi signal acquisition methods over different time intervals [16].

One of the limitations of traditional algorithms lies in the low specificity of the location fingerprint, which can lead to matching errors. To rectify this, the collected Wi-Fi signal data undergoes a normalisation process. This normalisation stop significantly enhances the specificity of the location fingerprint, thereby increasing the accuracy of subsequent matching processes [24].

- **Real-Time Localisation Phase**

Moving to the real-time localisation phase, the robot seek to determined its position by collecting Wi-Fi signals at an unknown point within the room. These collected Wi-Fi signals are theme averaged to obtain an representative signal. The Mahalanobis distance formula, is employed in the matching algorithm. This formula calculates the Mahalanobis distance between the data obtained from the unknown point and every location fingerprint stored in the location fingerprint database. Importantly, the Mahalanobis distance serves as a preferences for similarity, with smaller distances indicating grater similarity.

To further enhances the accuracy of the localisation estimation, the improved algorithm integrates an adaptive-weight WKNN algorithm. This algorithm selectively incorporates information from the database while filtering out noisy data points, thereby refining the accuracy of the robot's position estimation.

One the estimated coordinates of the unknown point are determined, they are promptly transmitted to a mobile or PC terminal. This allows users to visualise the robot's position within the room on the control platform, represented by a distinguishable red dot. This user-friendly feature enhances the algorithm's practicality and usability.

6.1 Fingerprint-Based Algorithm

Traditional location fingerprinting algorithms have typically employed a methodology involving equidistant raster division of indoor spaces, with preferences point intervals typically set at 1 meet or more. However, through experimentation, we discovered that the localisation accuracy of the same algorithm can be notably enhanced by gradually reducing the spacing between reference points. In light of optimising the algorithm's localisation accuracy while taking into account the practicality of data collection, we have decided to implement a preferences point spacing of 0. Three motors in our improved Wi-Fi location fingerprinting algorithm [15].

Figure 2 provides a top-down view of the room, illustrating the distribution of preferences points at this 0.3 m spacing. All forthcoming experiments will be conducted within this specific environments, as it offers the best compromise between localisation precision and the workload associated with data collection. This adjustment allows us to better capture the nuances of the indoor environments, ultimately lading to improved accuracy in determining the robot's position [25].

7. Technologies Used in Indoor Positioning and Navigation

7.1 VLC – Visual Light Communication

It works by encoding information and transmitting it through the flickering of LED light, which can then be received and decoded by photosensitive sensors. VLC can utilise various localisation techniques, such as RSSI, TOA, and AOA to locate the position and orientation of LEDs. Moreover, it can also use a camera to receive images of LEDs, similar to computer vision. LED lights are cost-effective and environmentally friendly, making them an ideal option for indoor environmental modification.

They can also be used for both lighting and communication, transmitting information without interfering with illumination. However, VLC can only communicate within the LOS range and requires prior calibration. The direction of improvement in VLC localisation lies in optimising the signal strength mapping position. VLC can be fused with other localisation methods like the IMU and encoder using corresponding fusion algorithms, such as the EKF and particle filter [21].

Various methods have been developed by researchers to identify the location of LEDs, including the use of the Cartographer algorithm in LiDAR SLAM or vision processing algorithms. Although VLC is mainly used for one-way data transmission, bidirectional data transmission has been achieved by researchers through the use of parity bits to minimize bit errors. Additionally, some researchers have created position and orientation sensors based on LEDs, while others have suggested the implementation of fuzzy logic systems for VLC localization.

7.2 Vision-Based Indoor – Positioning and Navigation System

The construction of a VBIP indoor map can be done in multiple ways, each with its own strengths. While the

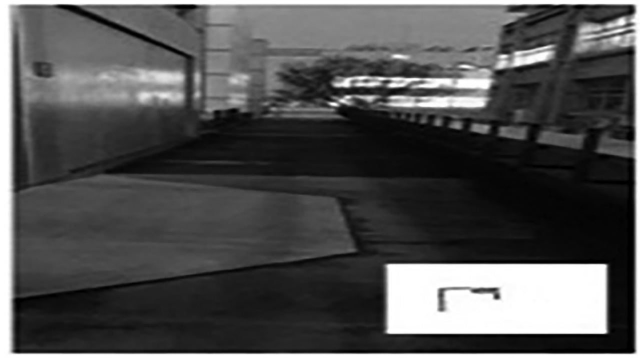


Figure 3. VSIP.

professional version may require a 360° camera and a floor plan, resulting in better accuracy and extra functions, we opted for the easiest and fastest way of using only one smartphone, due to lack of time and a floor plan. To produce a VBIP map, VIO is used to concurrently collect magnetic field and capture visual features, resulting in a virtual coordinate system that matches the real coordinate system. Each of these visual cues is captured with its colour, 3D positions, and RETOP descriptor for cross-comparison, and is saved as a sparse point cloud. For quick localisation, the magnetic field that is gathered throughout the procedure is recorded along with its coordinates and magnetic strengths. Adoption of VBIP occurs in two stages: tracking and localisation.

The localisation state uses a hybrid tracking method to increase positioning precision and two cascaded search algorithms to quickly localise the user's position. But in the first experiment, which we conducted during the 2018 Science and Technology Festival, we employed QR codes instead to quickly localise the whereabouts of the user. After that, we changed our VBIP system to use the suggested two cascaded search instead of QR codes, and we ran a second experiment to evaluate the improvement. The initial cascade search employs a pose estimation mechanism by employing magnetic strengths and a particle filter to focus the possible user position into a more limited range. Figure 5 depicts the cascade search's system flow [12].

8. New Techniques Used Based Current Trends

There are lots of vision based naviagtion methods are used in motion capturing. Navigation method with map method and withou map methods. The most common map method is sub-goal and one of the

8.1 Technique 1: Resilence

Resilience is a concept that refers to the ability of a system to recover its original function after external and internal disturbances, with its own resources and energy within a required period of time and an affordable cost. A resilient system is useful in situations where a system is vulnerable and/or impossible to repair and/or has an unaffordable cost to repair. Two similar concepts with the concept of resilience are robustness and reliability [26].

Figure 4. Cascade search of VSIP.

Robots are dynamic systems with various conceptual structures, including degrees of freedom (DOF), serial or open-loop structure, parallel structure, and closed-loop structure. They can include active, passive, fixed, servomotor, different distributions of actuators or motors over joints, and elements that play an active in actuating motion [27].

Resilience can be achieved by three approaches: changing the principle that governs the motion behaviour of a robot, changing the topology or configuration of a robot, and changing the state of a robot. In this paper, we consider the robot that may be under-actuated with a closed loop structure, fully-actuated with an open loop structure, or under-actuated with some components that have no definite motion.

Self-reconfigurable robots are with the same capability of self-changing their configurations. For example, CHOBIE, designed by Michihiko Koseki, is a lattice structure robot with sliding mechanisms. Conro, designed by Andres Castano, is a self-reconfigurable chain robot, but its self-reconfiguration is complex and slow. SUPERBOT, designed by Behnam Salemi of the University of Southern California, uses a hybrid type structure with two sub-modules with an anode joint and a cathode joint [26].

The resilience of a robot is improved by allowing the passive joint and passive link, and subsequently closed loop with passive joint, in the robot. This feature improves the robot's resilience, as if an active joint is broken to a passive joint, the whole robot may still work. This feature is expected to be more cost-effective than self-reconfigurable robots with all their joint modules active, meaning expensive servo-motors.

8.2 Technique 2: Computer Vision Technique

Computer vision is a crucial field that uses cameras to capture environmental imagery, enabling mobile robots to accurately determine their location. There are two main approaches to localisation in computer vision: beacon-based absolute localisation and visual odometry-driven SLAM techniques. Visual SLAM, particularly in indoor mobile robot positioning, closely mirrors LiDAR SLAM by employing multiple cameras or stereo cameras to gather feature-rich points with depth information from the surrounding environment. These systems rely on geometric features like points, lines, and planes as landmarks to construct comprehensive maps [28].

In the dynamic realm of semantic SLAM, researchers are working towards assigning explicit meaning to objects. Semantic SLAM takes this further by constructing maps enriched with semantic entities, combining spatial structure information with object-oriented semantics within the workspace. This approach yields accelerated closed-loop detection, fosters human-computer interaction, and facilitates the execution of complex tasks [29].

Dynamic scenes are tackled through innovative techniques like PSPNet, which partitions video frames into static, latent dynamic, and prior dynamic regions. Deep learning has propelled object detection and recognition to the forefront of computer vision, with techniques like YOLOv3 and RoomSLAM achieving enhanced precision in positioning. Sensor data fusion, including LiDAR, IMU, wheel odometers, IMU and UWB, LiDAR and Odometer, LiDAR Odometer and IMU, and IMU with wheel odometers, is also employed [30].

Despite the potential of visual SLAM, computational demands pose a significant challenge. Innovative solutions, such as Corotan *et al.*'s mobile robot system based on Google's ARcore platform and Zheng *et al.*'s cloud-based visual SLAM framework, aim to optimise processing efficiency by strategically utilising resources and offloading computation when possible [28].

9. Discussion

In recent years, the field of robotics has experienced a significant transformation, driven by advancements in sensor technology, chip development, and artificial intelligence.

Robots have evolved into highly intelligent and adaptable entities. This transformation has led to their widespread integration across various sectors, demonstrating their potential in various applications. For example, robots have been engineered with pliable manipulators, stereo cameras, and 3D vision capabilities, tailored for tasks like handling delicate fabrics. In healthcare, mobile medical robots have been developed to revolutionise patient care within medical facilities. These robots manage tasks, such as transporting sustenance, dispensing medications, and delivering medical provisions to nursing stations. The hospitality sector has also seen a surge in robot assistance, from cleaning and item transportation to information dissemination and secure payment handling. In the culinary industry, robots equipped with ultra-wideband positioning technology, odometers, gyroscopes, and accelerometers have emerged, enabling trackless navigation for food delivery.

While the GPS excels in outdoor positioning accuracy, indoor environments pose unique challenges. GPS signals weaken significantly indoors, leading to increased positioning errors. This has spurred the need for indoor positioning technologies that combine high accuracy and affordability. Leveraging Wi-Fi signals within indoor spaces offers advantages, such as cost-effectiveness, precise positioning, extensive coverage, robust indoor penetration, and strong communication capabilities.

Traditional Wi-Fi fingerprint-based positioning algorithms have faced challenges in achieving the required precision for indoor robot positioning. To address this, the article discusses enhancements to the conventional Wi-Fi fingerprint-based algorithm. These improvements aim to meet stringent indoor robot positioning accuracy criteria by introducing a grid-based navigation algorithm that leverages the enhanced positioning approach. This navigation algorithm stands out for its simplicity, reduced computational demands, and elevated accuracy.

Furthermore, mobile robots have significant potential in various applications, including military reconnaissance, anti-terrorism, and guiding visitors in museums. Achieving highly portable and high-frequency indoor positioning is crucial for these applications, with SLAM being a key technique. SLAM utilises various sensors like vision, laser, and odometer to enable accurate self-reliant navigation and positioning of robots in unfamiliar environments.

Vision-based indoor positioning and navigation methods have gained attention due to their accuracy and efficiency. These methods use cameras and other imaging devices to identify and locate objects in the environment, allowing robots to navigate autonomously in unfamiliar indoor environments. The SLAM technique, which involves constructing a map of the environment while determining the robot's location within that map, is widely used in vision-based indoor positioning and navigation methods. By integrating data from different sensors, such as vision, laser, and odometer, SLAM enables accurate and robust positioning and navigation of robots in indoor environments.

Despite the advantages of vision-based indoor positioning and navigation methods, challenges remain, including poor lighting conditions, occlusion, and complex indoor environments. Researchers are continually developing new techniques and algorithms to overcome these challenges and improve the accuracy and reliability of vision-based indoor positioning and navigation systems. The development of these methods has significant potential to revolutionize industries, such as manufacturing, healthcare, and retail.

10. Conclusion

The recent evolution of robotics has been marked by significant advancements in sensor technology, chip development, and artificial intelligence. Robots have transformed into highly intelligent and adaptable entities, finding applications in various sectors. Notable examples include robots designed to handle delicate fabrics in the textile industry, mobile medical robots revolutionising healthcare, and robots enhancing guest experiences in the hospitality sector by performing tasks like cleaning and item transportation. Additionally, the culinary industry has witnessed the emergence of navigation-enabled food delivery robots.

However, achieving accurate indoor positioning remains challenging due to the weakening of GPS signals indoors. To address this challenge, the article discusses the integration of robust indoor positioning technologies. Wi-Fi-based positioning, SLAM, and vision-based methods are highlighted as promising solutions for indoor robot positioning and navigation. These technologies offer cost-effective, precise, and extensive coverage, making them essential for various applications. Vision-based methods, in particular, are gaining traction due to their accuracy and efficiency. However, challenges, such as poor lighting conditions and complex indoor environments need to be addressed.

In conclusion, the field of indoor positioning and navigation for robots is rapidly advancing, driven by the increasing integration of robots into various industries. Advancements in positioning technologies are critical for the successful deployment of robots in real-world scenarios. As research and development efforts continue, we can expect further improvements in indoor positioning accuracy and reliability, opening up new possibilities for robots in industries ranging from manufacturing to healthcare and beyond.

Acknowledgement

This study was supported by the Basic Research Program of Shanxi province (free exploration type) (No. 202203021212335), the Scientific and Technologial Innovation Programs of Higher Education Institutions in Shanxi (2019L0920), and Taiyuan Academician Workstation Building Unit: Taiyuan Institute of Technology (TYSYSGZZ201903).

References

[1] X. Cai, H. Ning, S. Dhelim, R. Zhou, T. Zhang, Y. Xu, and Y. Wan, Robot and its living space: A roadmap for robot development based on the view of living space, *Digital Communications and Networks*, 7(4), 2021, 505–517.

[2] D. Kijdech and S. Vongbunyong, Using yumi robot and RGB-D camera with Yolov5 for pick-and-place application, *International Journal of Robotics and Automation*, *38*(10), 2023, 1–10. DOI: 10.2316/J.2023.206-0927.

[3] L. Deng, X. Ma, J. Gu, Y. Li, Z. Xu, and Y. Wang, Artificial immune network-based multi-robot formation path planning with obstacle avoidance, *International Journal of Robotics and Automation*, *31*(3), 2016. DOI: 10.2316/Journal.206.2016.3.206-4746.

[4] Y. Zhu and B. Jin, Compliance control of a legged robot based on improved adaptive control: Method and experiments, *International Journal of Robotics and Automation*, *31*(5), 2016. DOI: 10.2316/Journal.206.2016.5.206-4536.

[5] Y. Zhu, Further perspective of machine vision in industrial robot systems, *Highlights in Science, Engineering and Technology*, *39*, 2023, 909–914.

[6] H. Ye and J. Peng, Robot indoor positioning and navigation based on improved wifi location fingerprint positioning algorithm, *Wireless Communications and Mobile Computing*, *2022*.

[7] Y. Sun, L. Guan, Z. Chang, C. Li, and Y. Gao, Design of a low-cost indoor navigation system for food delivery robot based on multi-sensor information fusion, *Sensors*, *19*(22), 2019, 4980.

[8] T. Ren, Q. Liu, Y. Chen, and S. Ji, Variable pitch helical drive in-pipe robot, *International Journal of Robotics and Automation, 31(*3), 2016. DOI: 10.2316/Journal.206.2016.3.206-4774.

[9] C. Deniz and M. Cakır, A novel designed interactive training platform for industrial robot offline programming and robotics education, *International Journal of Robotics and Automation,32*(6), 2017. DOI: 10.2316/Journal.206.2017.6.206-5139.

[10] A. Saracoglu and D.U. Sanli, Accuracy of GPS positioning concerning Köppen-Geiger climate classification, *Measurement*, *181*, 2021, 109629.

[11] L. Qin, B. Niu, B.S. Li, X.L. Hu, and Y.X. Du, High precision indoor positioning algorithm of single LED lamp based on A-Bayes, *Optik*, *241*, 2021, 167190.

[12] H. Liu, H. Darabi, P. Banerjee, and J. Liu, Survey of wireless indoor positioning techniques and systems, *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, *37*(6), 2007, 1067–1080.

[13] G.B. Huang and C.K. Siew, Extreme learning machine: RBF network case, *Proc. ICARCV 2004* 8^{th} *Control, Automation, Robotics and Vision Conf.*, Kunming, 2004, 1029–1036.

[14] G.B Huang and C.K. Siew, Extreme learning machine with randomly assigned RBF kernels, *International Journal of Information Technology*, *11*(1), 2005, 16–24.

[15] Y. Kang, Z. Li, and T. Wang, Application of PID control and improved ant colony algorithm in path planning of substation inspection robot, *Mathematical Problems in Engineering*, *2022*.

[16] W. Cui, Q. Liu, L. Zhang, H. Wang, X. Lu, and J. Li, A robust mobile robot indoor positioning system based on Wi-Fi, *International Journal of Advanced Robotic Systems*, *17*(1), 2020, 1–10.

[17] J. Li, X. Zhang, J. Li, Y. Liu, and J. Wang, Building and optimization of 3D semantic map based on Lidar and camera fusion, *Neurocomputing*, *409*, 2020, 394–407.

[18] P. Jiang, L. Chen, H. Guo, M. Yu, and J. Xiong, Novel indoor positioning algorithm based on Lidar/inertial measurement unit integrated system, *International Journal of Advanced Robotic Systems*, *18*(2), 2021. DOI: 10.1177/1729881421999923.

[19] F.J. Aranda, F. Parralejo, F.J. Álvarez, and J.A. Paredes, Performance analysis of fingerprinting indoor positioning methods with BLE, *Expert Systems with Applications*, *202*, 2022, 117095.

[20] H. Cao, Y. Wang, J. Bi, S. Xu, M. Si, and H. Qi, Indoor positioning method using WiFi RTT based on LOS identification and range calibration, *ISPRS International Journal of Geo-Information*, *9*(11), 2020, 627.

[21] H.Y. Tsai, Y. Kuwahara, Y. leiri, and R. Hishiyama, Vision-based indoor positioning (VBIP)-an indoor AR navigation system with a virtual tour guide, *Proc. Collaboration Technologies and Social Computing:* 25^{th} *International Conf., CRIWG+ CollabTech 2019, Kyoto, Japan, September 4–6, 2019,* 96–109.

[22] C. Yang and H.R. Shao, WiFi-based indoor positioning, *IEEE Communications Magazine*, *53*(3), 2015, 150–157.

[23] X. Wang, D. Wang, M. Du, K. Song, Y. Ni, and Y. Li, A two-layer trajectory tracking control scheme of manipulator based on ELM-SMC for autonomous robotic vehicle, *IEEE Transactions on Automation Science and Engineering*, 2023.

[24] X. Lin, J. Gan, C. Jiang, S. Xue, and Y. Liang, Wi-Fi-based indoor localization and navigation: A robot-aided hybrid deep learning approach, *Sensors*, *23*(14), 2023, 6320.

[25] Z. Zhang, Y. Yu, L. Chen, and R. Chen, Hybrid indoor positioning system based on acoustic ranging and Wi-Fi Fingerprinting under NLOS environments, *Remote Sensing*, *15*(14), 2023, 3520.

[26] F. Wang, Z. Qian, Z. Yan, C. Yuan, and W. Zhang, A novel resilient robot: Kinematic analysis and experimentation, *IEEE Access*, *8*, 2019, 2885–2892.

[27] A. Prorok, M. Malencia, L. Carlone, G.S. Sukhatme, B.M. Sadler, and V. Kumar, Beyond robustness: A taxonomy of approaches towards resilient multi-robot systems, 2021, *arXiv:2109.12343*.

[28] J. Huang, S. Junginger, H., Liu, and K. Thurow, Indoor positioning systems of mobile robots: A review, *Robotics*, *12*(2), 2023, 47.

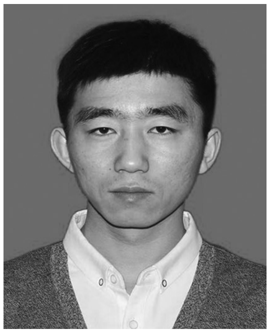
[29] G. Lee, B.C. Moon, S. Lee, and D. Han, Fusion of the SLAM with Wi-Fi-based positioning methods for mobile robot-based learning data collection, localization, and tracking in indoor spaces, *Sensors*, *20*(18), 2020, 5182.

[30] T. Ran, L. Yuan, and J.B. Zhang, Scene perception based visual navigation of mobile robot in indoor environment, *ISA Transactions*, *109*, 2021, 389–400.

Biographies

Junfu Qiao, male, Associate Professor, born in Jinzhong, Shanxi, P.R. China, in 1979. He received the master's degree from Taiyuan University of Technology, P.R. China. Now, he works with the Department of Automation, Taiyuan Institute of Technology. His research interests include automatic test and control technology, Intelligent robot design and application research, machine vision and artificial intelligence technology applications.

Jinqin Guo, female, Professor, born in December 1975. She received the master's degree from North University of China. Now, she works with Taiyuan Institute of Technology. Her main research interests include detection technology and automation devices, mainly engaged in test and measurement technology, intelligent instruments, embedded vision and other aspects of research work.



Yongwei Li, male, Associate Professor, born in Lvliang, Shanxi, P.R. China, in 1989. He received the doctoral degree from North University of China, P.R. China. Now, he works with the Department of Automation, Taiyuan Institute of Technology. His research interests include advanced sensor technology and intelligent control technology.