Pose estimation and data fusion algorithms for an autonomous mobile robot based on vision and IMU in an indoor environment

by

Mary Bosede Alatise

Submitted in partial fulfilment of the requirements for the degree

Philosophiae Doctor (Computer Engineering)

in the

Department of Electrical, Electronic and Computer Engineering

Faculty of Engineering, Built Environment and Information Technology

UNIVERSITY OF PRETORIA

July 2021

SUMMARY

POSE ESTIMATION AND DATA FUSION ALGORITHMS FOR AN AUTONOMOUS MOBILE ROBOT BASED ON VISION AND IMU IN AN INDOOR ENVIRONMENT

by

Mary Bosede Alatise

Supervisor:	Dr. G.P. Hancke
Co-Supervisor:	Prof. G.P. Hancke
Department:	Electrical, Electronic and Computer Engineering
University:	University of Pretoria
Degree:	Philosophiae Doctor (Computer Engineering)
Keywords:	Localisation, pose estimation, mobile robot, inertial sensor, vision, camera, data fusion, extended Kalman filter, particle filter, monocular

Autonomous mobile robots became an active research direction during the past few years, and they are emerging in different sectors such as companies, industries, hospital, institutions, agriculture and homes to improve services and daily activities. Due to technology advancement, the demand for mobile robot has increased due to the task they perform and services they render such as carrying heavy objects, monitoring, delivering of goods, search and rescue missions, performing dangerous tasks in places like underground mines. Instead of workers being exposed to hazardous chemicals or environments that could affect health and put lives at risk, humans are being replaced with mobile robot services. It is with these concerns that the enhancement of mobile robot operation is necessary, and the process is assisted through sensors. Sensors are used as instrument to collect data or information that aids the robot to navigate and localise in its environment. Each sensor type

has inherent strengths and weaknesses, therefore inappropriate combination of sensors could result into high cost of sensor deployment with low performance.

Regardless, the potential and prospect of autonomous mobile robot, they are yet to attain optimal performance, this is because of integral challenges they are faced with most especially localisation. Localisation is one the fundamental issues encountered in mobile robot which demands attention and the challenging part is estimating the robot position and orientation of which this information can be acquired from sensors and other relevant systems. To tackle the issue of localisation, a good technique should be proposed to deal with errors, downgrading factors, improper measurement and estimations. Different approaches are recommended in estimating the position of a mobile robot. Some studies estimated the trajectory of the mobile robot and indoor scene reconstruction using a monocular visual odmometry. This approach cannot be feasible for large zone and complex environment. Radio frequency identification (RFID) technology on the other hand provides accuracy and robustness, but the method depend on the distance between the tags, and the distance between the tags and the reader. To increase the localisation accuracy, the number of RFID tags per unit area has to be increased. Therefore, this technique may not result in economical and easily scalable solution because of the increasing number of required tags and the associated cost of their deployment. Global Positioning System (GPS) is another approach that offers proved results in most scenarios, however, indoor localization is one of the settings in which GPS cannot be used because the signal strength is not reliable inside a building. Most approaches are not able to precisely localise autonomous mobile robot even with the high cost of equipment and complex implementation. Most the devices and sensors either requires additional infrastructures or they are not suitable to be used in an indoor environment. Therefore, this study proposes using data from vision and inertial sensors which comprise 3-axis of accelerometer and 3-axis of gyroscope, also known as 6-degree of freedom (6-DOF) to determine pose estimation of mobile robot. The inertial measurement unit (IMU) based tracking provides fast response, therefore, they can be considered to assist vision whenever it fails due to loss of visual features. The use of vision sensor helps to overcome the characteristic limitation of the acoustic sensor for simultaneous multiple object tracking. With this merit, vision is capable of estimating pose with respect to the object of interest.

A singular sensor or system is not reliable to estimate the pose of a mobile robot due to limitations, therefore, data acquired from sensors and sources are combined using data fusion algorithm to estimate position and orientation within specific environment. The resulting model is more accurate because it balances the strengths of the different sensors. Information provided through sensor or data fusion can be used to support more-intelligent actions. The proposed algorithms are expedient to combine data from each of the sensor types to provide the most comprehensive and accurate environmental model possible. The algorithms use a set of mathematical equations that provides an efficient computational means to estimate the state of a process. This study investigates the state estimation methods to determine the state of a desired system that is continuously changing given some observations or measurements. From the performance and evaluation of the system, it can be observed that the integration of sources of information and sensors is necessary. This thesis has provided viable solutions to the challenging problem of localisation in autonomous mobile robot through its adaptability, accuracy, robustness and effectiveness.

DEDICATIONS

This thesis is dedicated to the Almighty God for the enablement to start and finish the project. I would also like to dedicate it to my family for their love and support.

ACKNOWLEDGEMENTS

I wish to express my deepest gratitude to my supervisors, Dr Gerhard Hancke and Prof Gerhard Hancke for their guidance, support, inputs, advice and constructive direction they gave me in the course of this study. Their patience as a virtue are worthy of emulation. It is whole-heartedly appreciated that your great advice proved monumental towards the success of this study. I would also like to recognize all the academic staffs for their precious assistance rendered to me towards the completion of my study.

I am grateful to the University of Pretoria and National Research Foundation for the financial support that has allowed me to conduct and complete this study.

To my darling husband, Mr TJ Olaniyi for his love, care and support, thank you. My appreciation also goes to my lovely sons, Mr Oluwasemilore and Mr Oluwabukunmi Olaniyi.

I will also want to appreciate my father, Prof MO Alatise for his great love, prayers and support. Indeed, you are the best daddy I could ever ask for. Thanks to my sister, Mrs IF Adetomiwa and brother, Mr OJ Alatise for their contributions and support. To my late mother, Mrs Adesola Alatise, your instilled words kept me going all through my research period. Continue to rest in the bosom of the Lord.

Finally, I will also like to thank my friends and family in Christ for their prayers, support and inputs.

LIST OF ABBREVIATIONS

AC	Alternating current
ACO	Ant colony optimisation
AMR	Autonomous mobile robot
APF	Auxiliary particle filter
AR	Augmented reality
BA	Bat algorithm
BFO	Bacterial foraging optimisation
BRISK	Binary robust invariant scalable keypoints
CCD	Charge coupled device
CMOS	Complementary metal oxide semiconductor
CoG	Centre of gravity
DC	Direct current
DCM	Direct cosine matrix
DoF	Degree of freedom
DoG	Difference of Gaussian
D-S	Dempster-Shafer
EKF	Extended Kalman filter
EMI	Electromagnetic interference
GPS	Global Positioning System
HNA	Hybrid navigation algorithm
HS	Harmony search
IDE	Integrated development environment
IMU	Inertial measurement unit
IoT	Internet-of-things
IR	Infrared

IWO	Invasive weed optimisation
KF	Kalman filter
LRF	Laser range finder
MEMS	Micro-electromechanical systems
NHNA	New hybrid navigation algorithm
ORB	Oriented fast and rotated BRIEF
PCA	Principal components analysis
PE	Probabilistic evolution
PDF	Probability density function
PF	Particle filter
RANSAC	Random sample consensus
RFID	Radio frequency identification
RMSE	Root mean square error
RPS	Revolution per second
SIFT	Scalable invariant feature transform
SIR	Sequential importance resampling
SIS	Sequential importance sampling
SLAM	Simultaneous localisation and mapping
SMC	Sequential Monte Carlo
SPI	Serial peripheral interface
SURF	Speeded up robust features
UART	Universal asynchronous receiver transmitter
UKF	Unscented Kalman filter
UWB	Ultra-wideband
VFH	Vector field histogram
VFF	Virtual field force
wrt	With respect to

TABLE OF CONTENTS

LISTS OF TABLES

LISTS OF FIGURES

INTRODUCTION

1.1 BACKGROUND

Autonomous mobile robots are becoming more prominent in recent time because of its relevance and applications to the world today. Their ability to navigate in an environment without need for physical or electro-mechanical guidance devices has made it more promising and useful for applications such as hospital, institutions, agriculture, industries etc [1]. According to the author in [2], mobile robot creation in 2020 will comprise of 26 Million mobile robots enabling autonomy in smart factories, unmanned transportation and connected homes. With their proficiency to navigate in a hazardous environment and even working beyond human capabilities using artificial intelligence [3], they are virtually applicable to most areas of operations [4, 5] and this has cause a huge demand in production. Ways by which robots support humans include: safety, productivity and efficiency. The main goal of this study is therefore to improve and implement existing technologies to have a reliable and feasible system, which allows to precisely localise autonomous mobile robot in indoor environment.

Despite the request to design and develop more mobile robots, some issues such as navigation and tracking [6], path planning, localisation, mapping [7], obstacle detection and avoidance [8] especially for an autonomous system in an uncertain and complex environment are encountered with. The major and crucial issue among the mentioned is localisation of mobile robot. Localisation is identified as a problem of estimating the position of a device or object such as aircraft, humans and robots, relative to a reference frame, based on sensor input [9, 10]. It is therefore more challenging to obtain current accurate estimate positioning of an object be it static or mobile. For the mobile robot location to be estimated, several devices or sensors are adopted, such devices include: inertial sensors, odometry, infrared, radio frequency identification (RFID), ultrasonic, Wi-Fi, Global positioning system (GPS) [11], and laser and sonar ranging sensors [12-14]. Inertial sensor information is used to facilitate localisation and navigation. However, the measurement derived from inertial sensors are corrupted by errors such as time-variant sensor biases and measurement noise. This problem is addressed by combining inertial sensor measurement with complementary information provided by additional sensor. Odometry on the other hand is used to determine the position of mobile robot by integrating the velocity of the robot over the period of its motion to obtain distance. This method is subjected to errors caused by uncertainty in robot and unevenness of the surface. A good sensor to detect motion of reliability is the use of infrared, but this device can only give appropriate measurement for target which are of short range because its performance degrades with longer distances. It requires direct-line of sight and thus it is not suitable for indoor location sensing. RFID is a very good system technology for indoor localisation because of reasonable system price and reader reliability, but it requires a large number of infrastructures to accurately determine the location. Another important device that has received a high rate of attention is Wi-Fi-based positioning system. This system is widely used to estimate location for indoor environment, but it requires Wi-Fi modules which could incur more cost. Ultrasonic method usually requires either a transmitter or receiver in the target objects. The technology utilizes Time-of-Flight (ToF) method to obtain the location information. However, it requires a great deal of infrastructure to accurately determine the location. Laser and sonar ranging sensor are considered as one of the most reliable and commonly used sensor for localisation of a mobile robot. Quite a number of studies have shown that this scheme satisfy practical requirements. However, it is a still challenging problem to employ the technology for indoor localisation because the measurements are erroneous in an environment surrounded by transparent or reflective objects such as glass walls or mirrors. GPS is considered inappropriate for indoor localisation because it does not work properly due to the signal from satellite that are attenuated and scattered by roofs, walls and other objects. Despite the efficiency and importance of using these devices they have their shortcomings such as limited power, high cost, noise, error, unreliability, limited range, inflexibility etc. In certainty, the use of relatively cheap sensors is important from a practical point of view; however, low-cost sensors seldom provide good performance due to measurement inaccuracies in various environments. Recently, augmented reality (AR) has been widely deployed to facilitate a new method for users to interact with their surroundings. Regardless of research carried out on current technologies for indoor environments to estimate the position and orientation of mobile devices, the high cost of deployment and inability to achieve minimised error in position and orientation are still major challenges. In recent times, with the Internet-of-things (IoT) and mobile devices enabling sensing [15, 16] for a variety of consumer, environmental and industrial applications [17-21], sensors and embedded intelligence have become cheaper and easier to integrate into systems, but they may not be suitable to present optimal performance [19, 22]. It is important to develop an effective system suitable to localise an autonomous mobile robot that works in an indoor or unstructured environment. Therefore, it is suggested that for a mobile robot to localise itself, the combination of sensors or other localisation devices will as well improve the efficiency. Recent research has proven that vision positioning techniques is becoming a promising approach that provides required information about positioning of object in an environment [23]. It aids to find an appropriate image process technique that is optimal for object detection within robot exploration application. The robot requires a vision system that employs some image analysis techniques, which are insensitive to environmental conditions such as lightning, texture etc. Though vision aid to computes object parameters from images acquired by the camera which is often uncomplicated and portable, however, it doesn't give depth information estimation and as such inaccurate information can be provided especially for fast navigation. Therefore, the following are the problem to be solved: to provide a viable solution on how information can be extracted from necessary devices through algorithms and mathematical expression to improve the overall performance of the system, to develop a model that is most suitable to combine two or more data together and to estimate the position of the mobile object and to adopt a low-cost positioning system and still give an optimal result which can also be used in a larger zone. The main motivation of this research work is to identify the significant challenges of autonomous mobile robot and to address them and on the other hand balance the trade-off between accuracy and performance.

1.2 PROBLEM STATEMENT

Autonomous mobile robots have not yet made much impact upon industrial and domestic applications, mainly due to the lack of robust, reliable, cheap and flexible localisation and behaviour mechanisms for robots operating in unmodified, semi-structured environments. There are sensors and devices that can be used by mobile robots to determine their location, but these sensors are not equipped with all it is required to satisfy accurate positioning because they are prone to noise, errors, vibrations and other factors that could degrade the performance. Therefore, these sensors and devices are required to be filtered, calibrated and fused to ascertain effective performance. More so, it is a challenging task to model a good, less complex and adaptive data fusion algorithm that could be updated in real time and have the capability to handle non-linearity.

1.2.1 Context of the problem

Localisation is very paramount in the performance of mobile robot navigation system but challenging. The ability for a mobile robot to estimate its position in each environment during autonomous movement is an issue that must be dealt with. Considerable studies have been carried out [13, 24, 25] in the development of various method of extracting information from the environment. Their works have employed different mechanisms to know the position of mobile robot in relative to position of objects with the consideration of robustness, efficiency, accuracy, deployment and reliability as a workflow for their paradigm.

Object identification is a fundamental workflow in computer vision to aid pose estimation. Most of the objects placed in the environment are of various sizes, shapes and colour. A high-level approach that can accurately detect or match objects in a clustered scene, unknown and known environment is required. Many robust local descriptors are stable under different viewpoints and lighting conditions; however, their computational requirement is stronger than methods that are dependent on artificial landmarks. Therefore, the use of artificial and natural landmarks is still a challenge for mobile devices [14]. To model a 3D displacement in motion that has a 2D transformation, an adaptive model is required to handle appearance change due to effects and deformation. Therefore, the construction of 3D models from images in the context of mobile robot exploring indoor environment is still an open research.

Sensors provide a robot with a capability of sensing its environment and handling environment uncertainty. No sensor works well in all situation and every sensor suffers from certain drawbacks [26]. The performance of sensors may degrade after a limited life span or under some conditions. It is therefore suggested that, the use of more than one sensor is encouraged while the limitation of one sensor could be compensated by another with better ability and improved performance [27].

It has been a huge challenge for researchers to develop diverse algorithms to interpret the data obtained by sensors especially in updating in real-time and handling non-linearity. This process can be established using sensor fusion algorithm. The algorithm fuses sensors together and thereby effectively reduce sensor inaccuracy and possible false sensor information. Without sensors and data fusion algorithm, the development of an autonomous and intelligent robotic system remains in the realms of science fiction, therefore no robot can function well in real world [26]. The main purpose of multi-sensor data fusion is to eradicate the limitations of individual sensors and as well produce accurate, robust, flexible and reliable estimate of the world based on multi-sensory information.

1.2.2 Research gap

The mobile robot may need to identify its absolute position but its relative position with respect to the target object is equally imperative. Sensors play an integral role in determining the position of the mobile robot, but the inexactness and incompleteness of these sensors poses difficult challenges in localisation. Noise from the sensors induces limitations as such they reduce the useful information contents and the consistency of the sensor readings. Data provided by sensors is always affected by some level of impreciseness as well as uncertainty in the measurement. The solution is to take multiple readings into account or multi-sensor fusion to increase the overall information content of the robot input.

For a mobile robot to know its location accurately in terms of coordinate, it is therefore suggested that using landmark-based method could be suitable to aid localisation. However, these methods are faced with some difficulties such as improper representations of markers, illuminate change, geometrical variations, lack of pattern quantity and computational complexity in the environment. Although landmark-based methods are simple and more efficient, but they still pose the problem of identifying robot position in relation to features [14].

The possibility of combining sensors to result in a basic system suitable for localisation is still open. Therefore, data fusion algorithm is the best method to be used to integrate the data acquired from diver's sources or devices. The algorithm should be able to express such inadequacy effectively and to exploit the data redundancy to reduce their effects. This study proposes a model to fuse data obtain from multi-sensors for estimation of mobile robot location. The method is based on an effort to develop a generic robust system for mobile robot location [26].

1.3 RESEARCH OBJECTIVES AND QUESTIONS

The objectives of this research study are:

- To investigate feasibility of autonomous robot localisation using light weight, low power consumption and low-cost sensors.
- To investigate methods for determining mobile robot orientation and location using multiple sensors.
- To develop a method for object identification and detection with the purpose of robot localisation.
- To investigate a data fusion algorithm with minimum computation time.

The research questions these study aims to answer are:

- Which sensors can be most reliably used for mobile robot localisation?
- How can sensors be combined to determine the pose estimation of a mobile robot?
- Does landmarks or objects place in the environment contribute to the effective estimation of robot localisation?
- How will the performance of the developed model and solutions be measured and validated?

1.4 HYPOTHESIS AND APPROACH

Localisation is a challenge in mobile robot and as such the efficiency and performance is dependent on the approach developed. This study seeks to presents a suitable methodology to estimate the optimal precise position and orientation of the mobile robot considering the data acquired by sensors and other source of information.

The use of a non-linear algorithm to fuse data obtained from sensors and devices tackles the issue of positioning. It is expected for the algorithm to show reduced error and overall improvement with good precision of localisation. This study can prove hypothetically that data acquired from sensors can determine mobile robot location and the use of fusion data algorithm to combine information from multiple sensor sources lead to accurate and efficient pose estimation.

1.5 RESEARCH GOALS

The goals that this research study seeks to achieve are to:

1. To investigate low power consumption, low cost, portable and easy to use sensors and devices to acquire data and determine accurate localisation [28].

2. To develop an algorithm that will consolidate the data collected from sensors and sources to adequately track and localise autonomous mobile robot in a dynamic indoor environment with the purpose of minimising errors and in all enhancing the performance of the system.

1.6 RESEARCH CONTRIBUTION

Recently, mobile robot is experiencing revolution and quite several autonomous objects are available to solve different tasks. One of the key problems that require intervention currently is the improvement of localisation of autonomous mobile robot. To this end, this work has addressed it. The following contributions have been made during this research work:

- A detailed study on sensor base localisation in autonomous mobile robot uncovered the most substantial shortcomings of its reliability and usefulness. The identified challenges are vital to be addressed, if not, they stand as potential barrier to mobile robot achieving its promise to providing a viable solution to the community at large. In the study, the limitations to achieving optimality in localisation for autonomous mobile robot were exposed, and as well solutions that could mitigate the effects were recommended. The solution models developed prevailed as great contributions to the body of knowledge for autonomous mobile robot.

- In addressing the enervating problem using IMU suffers from integration drift due to error and inaccuracy. To overcome this limitation, we propose the use of inertial system in combination with a vision-based system. These two sensors have shown to complement each other. For vision, both marker-based and markerless methods were revealed to be an effective approach for vision algorithm to aid localisation.

- The combination or fusion of two or more sensors or information acquired from a device have shown to be an effective method to have a reliable and efficient performance of determining localisation of a mobile robot, but such method have proven to be complex and difficult to understand because such systems are categorised as non-linear system which may require non-linear algorithms to tackle the problem. We proposed two simple data fusion algorithms that can be applicable to a non-linear system and the approach was able to estimate the pose estimation of a mobile robot which also can be applicable to other types of objects.

1.7 RESEARCH OUTPUTS

The outputs of this research study are as follows:

- M. Alatise and G. P. Hancke, "Pose estimation of a mobile robot based on fusion of IMU data and vision data using an extended Kalman filter," *Sensors (Basel)*, vol. 17, pp. 1-22, Sept. 2017.

- M. Alatise and G. P. Hancke, "Pose estimation of a mobile robot using monocular vision and inertial sensors data," in *Proc. IEEE AFRICON*, Cape Town, South Africa, Sept. 2017, pp. 1552-1557.

- M. Alatise and G. P. Hancke. "Indoor mobile robot localisation based on low-cost vision sensor" in *Proc. Southern Africa Telecommunication Network and Applications Conference*. (SATNAC), Ballito, South Africa, Sept. 2019, pp. 1-6.

- M. Alatise and G. P. Hancke, "A review on challenges of autonomous mobile robot and sensor fusion methods," *IEEE Access*, vol. 8, pp. 39830-39846, Feb. 2020.

1.8 DELINEATION AND LIMITATIONS

Delineation

The evaluation of this study was based on experimental setup carried out in an indoor environment and simulation using `Arduino IDE (integrated development` `environment)` and `MATLAB` as a software simulation for data analysis. The goal of this research is to improve the performance of mobile robot localisation most precisely to find the location and orientation of the mobile using a less complex algorithm to fuse data from different devices. In this study, the inertial sensor and vision are employed to estimate the movement of the mobile robot. Operations like obstacle avoidance, map analysis and predefined navigation pattern are not covered in this study.

Limitations

To ensure more precise location for the mobile robot, best quality hardware equipment need to be purchased. However, the acquisition of the equipment will jeopardize the goal of the research, therefore, cost is a constraint. Another limitation is that the experiment is required to be performed in a quite serenity that is, limited movement or else unnecessary noise and disturbance could affect the efficacy of the data collected.

1.9 STRUCTURE OF THE STUDY

The rest of the thesis is organised as follows: we commenced the work in Chapter 2 with a comprehensive review on challenges faced with autonomous mobile robot, more background about inertial sensors, vision sensors and the data fusion algorithms. Furthermore, the categories of landmarks (markers) method were explained with the specification and applications depending on what operation or task is to be carried out. Chapter 3 covers the necessary mathematical framework, definition of coordinate systems, general notations and transformation equations. The chapter further presents the layout for experimental set-up for autonomous mobile robot in an indoor environment. Chapter 4 further extend the work of improving the performance and minimizing sensor error by incorporating the use of image processing in combination with inertial sensor to track and estimate mobile robot. Finally, Chapter 5 covers concluding remarks, contributions and an outline of further work to be done in future.

CHAPTER 2 LITERATURE STUDY

The use of mobile robot is emerging in different sectors such as companies, industries and homes to improve services and daily activities. Due to technology advancement the demand for mobile robot has increased because of the task they perform and services they render. Various studies have been carried out by researchers on the importance of mobile robot, its composition and challenges. This chapter unravel the current literature, the challenges mobile robot is faced with. A comprehensive study of devices/sensors and prevalent techniques developed for tackling issues in mobile robot are presented as well. They are organised according to relevance, their strengths and weaknesses. The study therefore gives good direction for further investigation on developing methods to tackle the discrepancies faced with autonomous mobile robot.

The rest of the chapter is structured as follows: This chapter commences by providing an overview of autonomous mobile robot (AMR) in Section 2.1. This is followed by challenges mobile robot is faced with in Section 2.2. Section 2.3 highlighted the different types of sensors and techniques used to determine mobile robot positioning. Furthermore, related works were presented in Section 2.4 with the concentration on different types of sensor fusion methods addressing issues in mobile robot and their benefits. Future research areas and conclusion are given in Sections 2.5 and 2.6 respectively.

2.1 OVERVIEW OF AN AUTONOMOUS MOBILE ROBOT

An autonomous mobile robot is a system that operates in an unpredictable and partially unknown environment. This means the robot must have the ability to navigate without disruption and having the capability to avoid any obstacle placed within the confinement of movement. Autonomous mobile robots (AMR) have little or no human intervention for its movement and it is designed in such a way to follow a predefined path be it in an indoor or outdoor environment. For indoor navigation, the mobile robot is based on floor plan, sonar sensing, inertial measurement unit (IMU) etc. The first autonomous navigation was based on planar sensors such as laser rangefinder such that they navigate without human supervision. For an autonomous mobile robot to perform its task, it must have a range of environmental sensors. These sensors are either mounted on the robot or serve as an external sensor positioned somewhere in the environment. The number of different type of sensors mounted on the mobile robot to perform complex tasks and to determine estimation and localisation makes the design of the overall system extremely challenging [29-31].

Robot functionality is dependent on the composition of the robot. There are major components robot is made up with to perform some specified tasks such as movement and robot positioning. The basic parts are locomotion systems: to enable it to move through its environment either forward, backward, left, right, up and down [32]. Power supply unit is another paramount component. A direct current (DC) supply of 5V, 9V, 12V or more is required, and this is dependent on what the robot is to be used for. The most important thing to consider when choosing the voltage is that sufficient power should be supplied to drive all the loads connected [33]. Another cogent factor that ensures the effective performance of the robot is the selection of appropriate sensors. Sensors are one of the major components required in the configuration of robots. They perform a very significant role in localisation of robot be it controlled or uncontrolled robot. They measure robot condition and its environment and send such information to robot controller as electronic signals. Robot actuators utilise combinations of different electro-mechanical devices, synchronous motor, stepper motor, alternating current (AC) servo motor, DC servo motor and the microcontroller provides necessary intelligence to control the mobile robot. It processes the sensory information and computes the control commands for the actuators to carry out specific tasks.

Applications: Mobile robots attract attention more because of the increase in applications in various areas such as surveillance for security and monitoring, home for health and entertainment, research and education etc. [42]. Surveillance robots are now been installed in homes for monitoring activities, security purpose [43], by taking the responsibility of a security guard. Robot developed for domestic use is simple and easy to deploy, they are connected to Wi-Fi home network or smart environment to monitor and report activities going on in the environment. They have been designed further to engage in house cleaning, position objects where and when required. Recently, home robots are now been used by elderly people in a situation where emergency case arises. These robots have helped to promote technology that aids to detect and react to events that demand immediate response [44].

Another area where mobile robot is trending is the section of education. Educational robotics is primarily focused on creating a robot that will assist users to develop more practical, didactic, and cognitive skills. This approach is intended also to stimulate interest for research and science through set of different activities designed to support strengthening of specific areas of knowledge and skills. Introduction of mobile robot has increased not only on tertiary level and scientific research institutions, but also in lower grades such as secondary and primary schools [45]. These have therefore improved the knowledge of people about mobile robot worldwide.

Furthermore, mobile robot is gaining more interest in the area of mining industry [46]. The use of mobile robot has increased the efficiency and safety of miners. The robot can help to track people, robots and machines as well as monitor environmental conditions in mines. The mobile robotic platform is coupled with a set of range finders, thermal imaging sensors, and acoustic systems, all of which are functioned with neural networks. The robot navigates into different environments and identifies potential risk areas before the workers go in. Other use of mobile robots as shown in Figure 2.1 are not limited to the areas mentioned, other applications includes: fire-fighter, museum guides, planetary exploration, underwater and offshore structure inspection, underground pipe condition assessment, security, warehouse distribution application, space and ocean exploration, transmission tower inspection, patrolling, reconnaissance, petrochemical applications as well as for both domestic and industrial applications [47] etc. The construction industry is another sector of concern which is affected with low productivity and inefficiencies. Robots are now used to address the issues faced during the construction and maintenance of roads. Through the aid of robots, technologies are coupled into them to detect road defect and how to seal road cracks and potholes. The robotic system have proved to be effective in reducing labour, cost while improving productivity and quality. Moreover, robotics system can reduce injuries and free workers from conducting dangerous tasks. Beside construction sector, agriculture is another area where the application of robot is useful. Due to the rise in demand of food and human population growing rapidly, with same number of employees and the number even shrinking, a high demand is now placed on the use of robotic to ease the workload of workers. Robots are introduced to tap rubber, pick and harvest fruits and other farm produce.

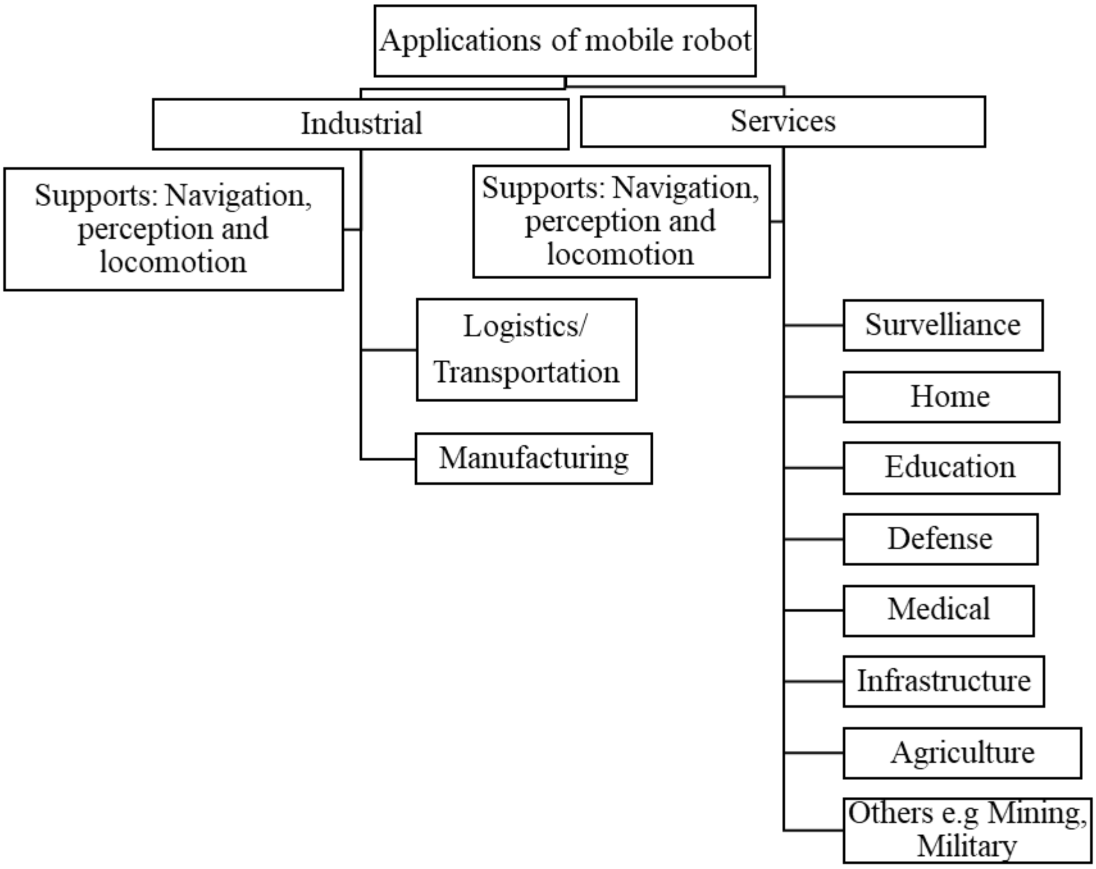


Figure 2.1. Applications of mobile robot taken from [1], © 2020 IEEE.

Locomotion system: Locomotion system is an important aspect of the mobile robot design which does not only rely on the medium in which the robot moves but also on other factors such as manoeuvrability, controllability, terrain conditions, efficiency, stability, and so on [34]. Locomotion of robots over surfaces and in various media can be based upon different principle. The types of locomotion are motions using special outer devices like wheels, legs, tracks and propellers. The design of mobile robot is dependent on the service to be rendered; therefore, a mobile robot can be designed to walk, run, jump, fly etc. With the requirement of the designed robot, they are categorised into stationary (arm and manipulator), land-based, air-based, water-based etc. Mobile robots especially autonomous are in high demand because of their ability and capacity to perform tasks that may seem difficult for humans. Examples of such mobile robots are wheeled, legged, walking or hybrid. Legged, wheeled, and articulated bodies are the main ways where mobile robot locomote [35]. The wheeled robots are one of the most vital systems of the robot locomotion and autonomous intelligent vehicle (AIVs) are part of the research field in mobile robotics which relies on principles such as pattern recognition, image process. They play an important role in logistics, transport and distribution. They are suited to ground either soft or hard ground while the legged and articulated bodies requires a certain degree of freedom and therefore greater mechanical complexity sets in [36]. The wheel has been by far the most popular locomotion mechanism in mobile robotics and in man-made vehicles in general. It can achieve very good efficiencies and does so with a relatively simple mechanical implementation. The use of wheels is simpler than using treads or legs and is easier to design, build, and program when the robot is moving on flat, non-rugged terrain. They also tend to be much cheaper than their legged counterparts. Wheel control is less complex, and they cause less wear and tear on the surface where they move in comparison with others. Another advantage is that they do not present any great difficulty in terms of balance issues, since the robot is usually in contact with a surface. The shortcoming of wheeled locomotion system is that they are not very good at navigating over obstacles, such as stony terrain, unsmooth surfaces, or areas with low friction [34]. The performance of a mobile robot can be improved by utilizing different modes in various terrain conditions. To design and develop the locomotion system, the terrain type for the mobile robot must be identified. The types of terrain are Uneven, Level Ground, Stair Up, Stair Down and Non-traversable [35].

Another factor to consider when designing a mobile robot is stability. Stability is not usually a great problem in wheeled robot, because they are designed in such that all the wheels are always in contact with the ground. The use of four-wheeled is more stable than three, two and one because the centre of gravity (COG) is located at the centre space of the wheels. These days, mobile robots are been designed to operate in two or more modes. In [37], the author proposed a mechanism structure for the mobile robot with the advantage of adaptability using hybrid modes with active wheels. On a rough terrain the robot locomote using the leg mode while for smooth terrain it makes use of the wheeled locomotion by roller-skating using the passive wheels. The challenging part is that the wheels are usually very heavy and bulky because they require driving actuators, steering and braking mechanisms. Therefore, installation of the active wheels usually increases the total weight of the walking vehicle which is already heavy enough limiting the versatility of the leg mechanism. To improve the localisation of a mobile robot irrespective of the terrain a technique is purpose to be employed. Dead reckoning has been already extended to the case of a mobile robot moving on uneven terrain. It provides the dynamic information of the pose for mobile robots by directly measuring the parameters such as position, velocity and orientation [38].

Perception: Perception refers to the ability of an autonomous system to collect information and extract relevant knowledge from the environment. Based on the sensors implemented, the environment perception task can be tackled by any kind of sensor or fusion of sensors. Fusion of information from different sensors enable to overcome limitation of individual sensor, to minimise overall uncertainty and to increase reliability in case of sensor failure. On the other hand, the fusion of sensory data has different effects in the different circumstances and sometimes sensors may fail to obtain accurate information. The performance of a multisensory is not only related to the sensor control, the accuracy of sensory measurements and fusion algorithm, but also to the perception planning. Perception is intended to design perception techniques and to perform high level uncertainty management, including sensor and sensory data selection according to the environment and sensors features, sensor parameter, selection of different fusion techniques and decision making in abnormal situations [40]. Perception is mostly concerned with map building and object recognition, specifically using LiDAR/laser range finders and vision sensors. For a map building problem, the challenge is how to construct a map with the minimum of time and energy, or minimum of the sensor readings so as to reduce the overload. A sensing strategy is added to acquire additional sensor data so as to maximize the amount of new information. In regard to object recognition, the problem is to develop a technique to sample the minimal sensor data that is sufficient to recognize and locate a known object. This decision involves sensor selection, view point determination as well as sensor parameter. In this regards, perception is being recognized as an important components and required to be focused more on mobile robotics.

It is very important for an autonomous mobile robot to acquire information from its environment, sense objects around itself, or its relative position. Perception contributes to this important aspect in mobile robot research. If a mobile robot cannot perceive its environment correctly and efficiently, then it will not be able to perform tasks such as estimating the position of an object accurately [39]. To achieve this, information are perceived by the use of sensors and other related devices [18]. Sensors make it possible to autonomously perform robot localisation through mechanism such as data collection, object identification, mapping and representation. Sensors used in the area of data collection is categorised into two major aspects: Proprioceptive/exteroceptive sensors and active/passive sensors. Proprioceptive sensors measure values internally to the system (robot), e.g. battery level, wheel position, joint angle, motor speed etc. These sensors can be encoders, potentiometers, gyroscopes, compasses, etc. Exteroceptive sensors are used to extract information from the environments or objects. Sonar sensors, Infrared (IR) sensitive sensors, ultrasonic distance sensors are some examples of exteroceptive sensors.

Active sensors emit energy into environment and then measure the environmental reaction. Because active sensors can manage more controlled interactions with the environment, they often achieve greater performance. However, for active sensor, the outbound energy may affect the very characteristics that the sensor is attempting to measure. Furthermore, an active sensor may suffer from interference between its signal and environment [40]. Examples of active sensors include sonar sensors, radars etc. While passive sensors receive energy to make observation like camera such as Charge Coupled Device (CCD) or Complementary Metal Oxide Semiconductor (CMOS) cameras, temperature sensors, touch sensors etc. These sensors are most applicable in relation to specificity and achievement in the design of an autonomous mobile robot. Table 2.1 gives types of sensors used by an autonomous mobile robot indoor localisation.

Pp= Proprioceptive, Ep= Exteroceptive, A=Active, P=Passive, A/P=Active and passive

Table 2.1. Classification of sensor systems [40].

Classification	**Sensor system**	**Category (Pp, Ep, A, P)**	**Purpose**
Tactile sensors:	Contact switches, bumpers Optical barriers Non-contact proximity sensors	Ep, P Ep, A Ep, A	They are designed to sense and determine the exact position of an object at a short distance via direct physical contact. They are also used to detect heat variations. Tactile sensors are mostly used to calculate the amount of force applied by the robot's end-effectors.
Wheel Encoders:	Optic Magnetic	Pp, A	It helps to measure the distance or speed the robot has driven. The wheel encoders also count the

			revolutions of each wheel and an orientation.
Optical sensor:	Infrared LiDAR	Ep, A Ep, A	These are light-based sensors which produces range estimate based on the time needed for the light to reach the target and return.
Heading sensors	Gyroscope	PP	It is a reliable rotation sensor that measures angular velocities and orientation.
Vision-based sensors	CCD/CMOS camera(s)	Ep, P	These sensors offer a vast sum of information about the environment and enables intelligent interaction in dynamic environments.
Active ranging sensors	Ultrasonic Laser rangefinder Optical triangulation	Ep, A Ep, A Ep, A	Active ranging sensors aid robot navigate and they usually originate as part of the localisation and environmental modelling. They are devices that generate highly precise distance measurement between sensor and target.

Navigation: Navigation is a fundamental problem in robotics and other important technologies. In order for the mobile robot to autonomously navigate, the robot has to know where it is at present, where the destination is, and how it can reach the destination [41]. The most important aspect in the design of a mobile robot is navigation abilities. The objective is for the robot to move from one place to another either in a known or unknown environment. Most of the time, the mobile robot cannot take the direct path from its initial position to the final goal, which means that motion planning techniques must be used. This means that the robot must rely on its other aspects, such as perception (the robot must use its sensors to obtain valuable data), localisation (the robot must know its position and configuration), cognition (the robot must decide what to do to achieve its goals), and motion control (the robot must calculate its input forces on the actuators to achieve the desired trajectory). Navigation in unfamiliar environment demand to give the robot the ability to generate its action plan and to track it. It is significant as they need to move through the environment to execute their tasks which maybe exploration, inspection transportation or any kind of interaction with the environment and the objects in it. There are basically two types of navigation problem. The local and global. The local navigation problem deals with navigation on the scale of a few distance, where the main issue is obstacle avoidance. A well-known solution to this problem is presented in where an occupancy grid map of the immediate surroundings of the robot is created and used to determine the navigation direction such that the robot is safely guided towards the goal. Since the map is local, and resembles a ‘sliding window’, mapping of the whole environment does not occur. As for the global navigation problem, it deals with navigation on a larger scale in which the robot cannot perceive the goal state from its initial position. A number of resolutions have been proposed in the literature [48] to address this problem. Most rely either on navigating using a pre-specified map or constructing a map on the fly. Most approaches also rely on some technique of localisation. Some work on robot navigation is landmark-based relying on topological maps which have a compact representation of the environment and do not depend on geometric accuracy. The downside of such approaches is that they suffer from sensors being noisy and the problem of sensor antialiasing (i.e. differentiating between similar landmarks) [52].

When the navigation goal is specified (either the robot requests to be guided to a certain place, or a sensor node requires the robot's assistance), the node that is closest to the goal triggers the navigation field computation. The navigation field provides the robot with the 'best possible' direction that has to be taken in order to reach the goal. Therefore, to have ascertain free collision path on the route of the robot to its destination, it is cogent to develop and implement a navigation system for a mobile robot with current technology sensors which clearly requires sophisticated algorithms which will handle uncertainty. More details on navigation are discussed in Section 2.2.1.

In robotics, another area to consider is the use of computer vision applications to aid navigation and localisation. In computer vision, object recognition and feature matching are a significant task to be performed for accurate positioning. Object recognition has long been adopted in mobile robot to detect or identify objects present in an image. The technique can either be used to determine coordinates of the object detected or calculate in relative to a proposed object identified in an image. Feature matching or image matching on the other hand performs the task of establishing correspondence between two images of the same scene/object. Examples of features associated between the images could be points, edges or lines, and these features are often called keypoints features [13, 14]. To perform the task of object recognition and feature matching, several algorithms were adopted and some of the algorithms will be mentioned and discussed later in the chapter.

2.2 AUTONOMOUS MOBILE ROBOT CHALLENGES

Autonomous mobile robots have proven to be a system that cannot be without as result of increase in demand for diverse applications. Regardless, the potential and prospect, they are yet to attain optimal performance, this is because of inherent challenges that they are faced with. These challenges (see Figure 2.2) have enabled more researchers to develop more interest in recent time. Some of the main challenges are listed below:

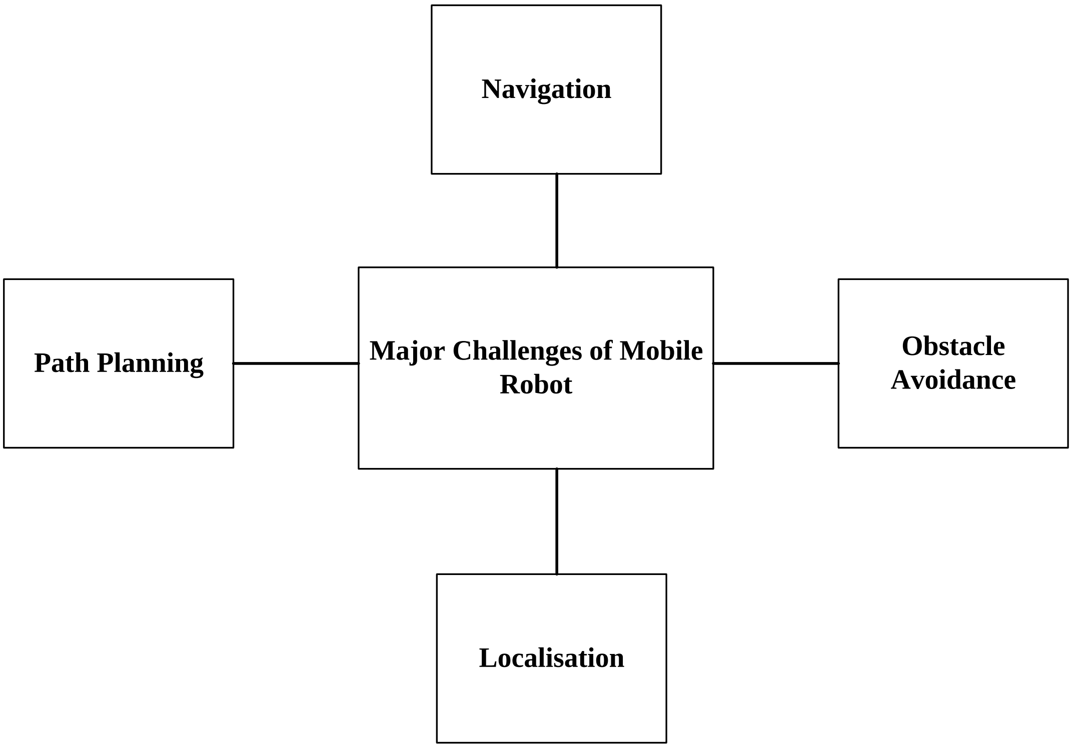


Figure 2.2. Challenges of mobile robot taken from [1], © 2020 IEEE.

2.2.1 Navigation

As earlier said, navigation of an autonomous mobile robot is an issue in robotics field. There are majorly two ways by which navigation problem is categorised into: local and global navigation. The local and the global navigation problem vary in terms of distances, scales and obstacle avoidance and in ability for the goal state to be observed. For local navigation, occupancy grid of map is used to determine the navigation direction and for global navigation, landmark approach based on topological map is used. This has a compact representation of the environment and do not depend on the geometric accuracy. The limitation of this approach is that they are downgraded by the noise generated from the sensor.

2.2.1.1 Environmental representation

Mobile robot navigation systems depend on the level of abstraction of the environment representation. To accurately determine the position and orientation of the mobile robot, it is imperative for the environment to be modelled in a simple and understandable structure. Three main techniques for representing the environment are given as: geometric, topological and semantic [48].

Geometric: The geometric representation is used to describe robot environment by parameterizing primitive geometric object such as curves, lines and points. The geometric representation of the environment is closer to the sensor and actuator world and it is the best one to perform local navigation. In [49], the author proposed the use of principal components analysis (PCA) - Bayesian based method with grid map representation to compress images and reduce computational resources. The PCA was also use to reduce dimensionality and model the parameter of the environment by considering the pixels of an image as feature vectors of the data set [50]. In [51] Markov localisation method was proposed to provide accuracy and multimodality to represent probability distribution of diverse kind but require significant processing for update, hence it is impractical for large environment.

Topological: A topological representation is characterized by defining reference elements of the environment according to the different relations between them. A conventional method for modelling the robot's environment is to discretize the environmental model by using a topological representation of the belief state, where each probable poses of the mobile robot is linked to a node in a topological map [52]. In [53], the proposed approach uses visual features extracted from a pair of stereo images as landmarks. While the new landmarks are fused into the map and transient landmarks are removed from the map over time. Topological representation of the environment uses graphs to model the environment and it is used in large navigation tasks.

Semantic: The current tendency in robotics is to move from representation models that are closest to the robot´s hardware such as geometric models to those models closer to the way how humans' reason, with which the robot will interact. It is intended to bring closer the models the way robots represent the environment and the way humans do. Robots that are provided with semantic models of the environments where they operate have a larger decision autonomy, and become more robust and more efficient [54]. An integrated approach for efficient online 3D semantic map building of urban environments and the subsequent extraction of qualitative spatial relationships between the different objects was presented, this enables efficient task planning [55]. Semantic information constitutes a better solution for interaction with humans [56], the representation is the most abstract representation model and adds concepts such as utilities or meanings of the environment elements in the map representation. Semantic navigation is considered as a navigation system that considers semantic information to model that includes conceptual and physical representation of objects and places, utilities of the objects, and semantic relation among objects and places. This model allows the robot to manage the environment and to make queries about the environment in order to do plans for navigation tasks [48].

Environmental model requires improved representation to enable successful result, better accuracy and as well reduce the computational cost [57]. For this to prevail, the environment must be well represented, simple technique must be adopted and be incorporated in to the robot's representation of its environment [58].

2.2.2 Localisation

Localisation is one the essential issues encountered in mobile robot which demands attention. The challenging part of localisation is estimating the robot position and orientation of which this information can be acquired from sensors and other systems. So, to tackle the issue of localisation, a good technique should be proposed to deal with errors, downgrading factors, improper measurement and estimations. The techniques are divided into two categories [59-62]: relative and absolute localisation.

Relative localisation techniques: This method estimate the position and orientation of the mobile robot by combining information produced by different sensors through the integration of information provided by diverse sensors, usually encoder or inertial sensors.

The integration starts from the initial position and continuously update in time. The relative positioning alone can be used only for a short period of time.

Absolute localisation techniques: This method allows the mobile robot to search its location directly from the mobile system domain. Their numerous methods usually depend on navigation beacons, active or passive landmarks, maps matching or satellite-based signals such as the global positioning system (GPS). For absolute localisation, the error growth is mitigated when measurements are available. The position of the robot is externally determined, and its accuracy is usually time and location independent. In other words, integration of noisy data is not required and thus there is no accumulation of error with time or distance travelled. The limitation is that one cannot keep track of the robot for small distances (barring exceptionally accurate GPS estimates).

2.2.3 Path planning

In autonomous robotics, path planning is a central problem in robotics. The typical problem is to find a path for a robot from a starting position to its target position. Safe and effective mobile robot navigation needs an efficient path planning algorithm since the quality of the generated path affects enormously the robotic applications [63-65].

In an environment with various obstacles, finding a path without collision with obstacles from the starting point to the destination becomes an issue such as shortness and simplicity of route are important criteria affecting on the optimality of selected routes. Considering the length of the path travelled by the robot, and energy consumption and its performance time, an algorithm that finds the shortest possible route [66] is the most appropriate. Basically, there are two types of environment: static and dynamic. While dynamic environment is divided into global and local path planning [64, 67]. Global navigation strategy deals with a completely known environment while local navigation strategy deals with the unknown and partially known environment.

Quite a number of studies have been investigated on path planning in dynamic environments. Authors in [68] proposed a new method to decide the optimum route of the mobile robot in an unknown dynamic environment, they used ant colony optimization (ACO) algorithm to decide the optimal rule table of the fuzzy system. Other related algorithms are bacterial foraging optimization (BFO) [64], and probabilistic cell decomposition (PCD) [69]. A new mathematical approach that is based on the concepts of

3-D geometry is proposed to generate the path of the mobile robot. The mobile robot decides its path in real time to avoid the randomly moving obstacles and to track randomly moving goal [70]. Other intelligent algorithms studied by researchers used by mobile robot to navigate in diverse environment are differential evolution (DE) algorithm [71, 72], harmony search (HS) algorithm [73], bat algorithm (BA) [74], and invasive weed optimization (IWO) [75].

2.2.4 Obstacle avoidance

Another challenging issue is obstacle avoidance, it is important that the mobile robot get to its destination without being obstructed by any obstacle or collision on its path. To this effect, collision free algorithm is a requirement of autonomous mobile robot, since it provides the safe trajectory and proves convergence [76]. Some of the main algorithms that can be used for obstacle avoidance are discussed in this section. Bug algorithm [77] is one of the earliest algorithms. It enables the robot to navigate the entire circumferences of the obstacle encountered and decide on the most appropriate point to leave toward the goal. The robot therefore moves to the best leaving position and later moves towards the object. The benefit of this algorithm is that it is easy to determine if an object is unreachable or not. However, the algorithm takes time to achieve its goal. Another algorithm is vector field histogram (VFH) [78] which is an improvement of the short coming of virtual force filed (VFF) algorithm [79]. VFH allows detection of unknown obstacle and avoids collision while simultaneously piloting the mobile robot towards the target. This algorithm employs a 2-stage data reduction process in order to compute the desired control command for the robot. This ensures accurate computation of the robot path to the target, but it consumes more resources like memory, processor and power. Hybrid navigation algorithm with roaming trails (HNA) [80] is an algorithm that is able to deal very efficiently with environments where obstacles are encountered by the robot during motion. During navigation the robot is allowed to deviate from its path to avoid obstacles on the basis of reactive navigation strategies, but it is never allowed to exit from the area. Since the robot is constrained to move within a convex area which includes the location of the target node, in presence of static obstacles it is guaranteed to reach the target by following a straight line. In some cases, the mobile robot has to either avoid the obstacles or simply stop in front of the obstacle. Another method that is similar to HNA is the new hybrid navigation algorithms (NHNA) [81]. The algorithm uses D-H bug algorithm (Distance Histogram bug) to avoid obstacle. It enables the robot to rotate freely at angle less than 90 degrees to avoid obstacle. If the rotation of 90 degrees or greater is required to avoid an obstacle, it acts as bug-2 algorithm [77] and start moving to destination when path is clear from obstacles. Conclusively, collision free algorithm is a requirement for autonomous mobile robot, since it provides safe trajectory.

In conclusion, challenges faced by mobile robot must be tackled to ensure effective performance. Navigation is one of the most important aspects to be considered when it comes mobile robot because it requires planning algorithms and appropriate information about robot's location. This will navigate the robot through its pre-defined path. In as much as navigation is important so also is trajectory planning. This will determine the path the robot must follow in order to reach its destination. Therefore, a path must be planned accordingly to avoid collision and obstacles. Different algorithms are considered for obstacle avoidance depending on the goal to be achieved. Finally, the robot must know its position and direction per time. In this regard, an effective localisation technique and reliable sensors are required to gather precise information.

2.3 SENSORS AND TECHNIQUES IN MOBILE ROBOT POSITIOING

To ensure accuracy in localisation, sensors and effective positioning system has to be considered. Objects positioning [14], robotics, and augmented reality (AR) tracking [82] have been of interest in the literature of recent. This section will discuss the existing technologies that aim at finding a mobile robot's position in its environment.

2.3.1 Inertial sensors

Inertial based sensor methods as also known as IMU (Inertial Measurement Units) which is a combination of accelerometers, gyroscopes and sometimes magnetometers. These sensors have become ubiquitous because many devices and system depend on them to serve a large sum of applications. They rely on measurement of acceleration, heading and angular rates, which can be acquired without external reference. Each of these sensors are deployed in robots, mobile devices and navigation systems [83]. The importance of using these sensors is solely to determine the position and orientation of a device and/or object.

Accelerometer: Accelerometer as a sensor measures the linear acceleration, which is the rate of change of velocity of an object. They measure in meters per second (m/s^2) or in gravity (g). They are useful for sensing vibration in system or for orientation applications [84]. Velocity is determined from it if integrated once and for position, integration is done twice. Using a standalone sensor like accelerometer could be simple and of low cost as stated by the author in [85], but the linear increasing error does not give a high-level of accuracy. Results produced by accelerometer for mobile robots have been unsuitable and of poor accuracy due to the fact that they suffer from extensive noise and accumulated drift. This can be compensated with the use of gyroscope.

Gyroscope: Gyroscope sensor measures the angular velocity in degrees per second $(^\circ/s)$ or revolution per second (RPS) and by integrating once, rotation angle can be calculated. Although gyroscope is small in size and inexpensive but run at a high rate in which they are able to track fast and abrupt movements. Another advantage of using gyroscope sensor is that it is not affected by illumination and visual occlusion [14]. However, they suffer from serious drift problem caused by accumulation of measurement errors for long period. Therefore, the fusion of both accelerometer and gyroscope sensor is suitable to determine the pose of an object and to make up for the weakness of one over the other.

Magnetometer: Magnetometer is another sensor used to determine the heading angle by sensing the earth magnetic field. They are combined with technologies to determine pose estimation [86]. However, magnetometer may not be so useful for indoor positioning because of the presence of metallic objects within the environments that could influence data collected through measurements [14]. Other methods that can be used to determine indoor localisation includes infrared, Wi-Fi, ultra-wideband (UWB), Bluetooth, wireless local area network (WLAN), fingerprinting etc. [87, 88]. However, these methods have their shortcomings, it is therefore necessary that two or more schemes be combined to achieve an accurate result.

Nowadays, accelerometer and gyroscope are integrated into a unit. Example of such IMU is Arduino 101. This has made measurement to be simple and reduction in error to be attained to a certain level. With the two sensors used, a six degree of freedom (6 DOF) were determined with the sides of X, Y and Z axes [89]. Acceleration can be measured independently in three axes: X, Y and Z as well as the angular rate by the gyroscope in X, Y and Z directions. Summary feature of accelerometer and gyroscope are given in Table 2.2. The table shows the ranges of measurement and with a digital resolution of 16 bits. The IMU unit combines the data captured from the tri-axis accelerometer and the tri-axis gyroscope. The 16-bit tri-axial accelerometer detects the linear motion and gravitational forces while the 16-bit tri-axial gyroscope measure the rate in space.

2.3.1.1 Calibration of IMU

Low cost micro-electromechanical (MEMS) used in many mobile robot applications still experiences, drifts, bias, repeatability, stability etc. [90, 91]. Therefore, it is necessary for IMU to be calibrated in the environments that they will operate in to improve performance. Several authors have proposed on how to calibrate sensor using different method. Traditionally the calibration of an IMU has been done by using special mechanical platforms such as a robotic manipulator, rolling shutter camera, moving the IMU with known rotational velocities in a set of precisely controlled orientations [92, 93]. However, the mechanical platforms used for calibration are usually very expensive; resulting in a calibration cost that often exceeds the cost of the IMU's hardware. In [94], the author proposed a model called turntable errors to reduce error in IMU by calibration. The result presented shows that with calibration, IMU accuracy improved despite using the traditional method. Also, David *et al.* [95] proposed an easy and robust method to calibrate IMU without the use of external equipment. This method only requires the sensor to be moved by hand and positioned in a different static location. The method is similarly applicable to our method. Here, calibration of IMU was done using the application programming interfaces (API) runtime made available in the CurieIMU library of the Arduino microcontroller. The CurieIMU library has made some values available to calibrate the IMU. The advantage of this method is that no additional equipment was required, and this has enabled the process to be simple and faster. The main reason for IMU calibration is to compensate the offsets values for accelerometer and gyroscope along each one of the reference axes. During this procedure the board should be placed on a flat surface and motionless. When finished, this procedure writes the proper offset value in the IMU register that can be read with getGyroOffset and getAccelerometerOffset.

2.3.2 Vision (Camera sensor)

Vision sensors are one of the most suitable devices to be used for positioning. Areas of application of vision sensors include indoor, outdoor and even in underwater [96, 97]. In recent times, mobile robot localisation researches are now based on vision. This is because information acquired through vision sensor from the environment aids to interpret using features descriptors and as such data is analyzed easily. In addition, the use of vision sensor is more accurate to estimate the relationship between 2D image and 3D scene. They also help to overcome the inherent limitation of the acoustic sensor for simultaneous multiple object tracking, while the acoustic sensor supports the estimation when the object is occluded.

Table 2.2. IMU feature summary [98].

	Accelerometer	**Gyroscope**
Digital Resolution	16-bit	16-bit
Measurement ranges (programmable)	$\pm 2g$, $\pm 4g$, $\pm 8g$, $\pm 16g$	$\pm 125°/s$, $250°/s$, $500°/s$, $1000°/s$, $2000°/s$
Sensitivity (calibrated)	$\pm 2g: 16384\ LSB/g$ $\pm 4g: 8192\ LSB/g$ $\pm 8g: 4096\ LSB/g$ $\pm 16g: 2048\ LSB/g$	$\pm 125°/s: 262.4\ LSB/°/s$ $\pm 250°/s: 131.2\ LSB/°/s$ $\pm 500°/s: 65.6\ LSB/°/s$ $\pm 1000°/s: 32.8\ LSB/°/s$ $\pm 2000°/s: 16.4\ LSB/°/s$
Zero-g offset	$\pm 40mg$	$\pm 10°/s$

2.3.2.1 Camera calibration

The issue of camera calibration is a known problem in computer vision. It is a necessary step in 3D computer vision to extract metric information from 2D image. There are several techniques and toolboxes available for such. The purpose of camera calibration is to find the intrinsic and extrinsic parameters of cameras. The basic methods of camera calibration can be divided into the traditional camera calibration methods and the camera self-calibration methods [99]. The traditional calibration method has a high calibration precision but need specific calibration reference substance. The self-calibration method does not rely on calibration reference substance, but the calibration results are relatively unbalanced. These procedures typically require images at several angles and distances of a known calibration object. The author in [100] proposed a self-calibration method based on active vision. The calibration method is simple and can get linear solution, but it is inflexible and has a high cost. With the developing of optimisation algorithms, many techniques are applied to camera calibration. References in [101-103] proposed schemes by which the problem of camera calibration can be solved using analytical solution and nonlinear optimisation techniques. It has fast convergence speed and good robustness. Zhang [73] proposed a camera calibration method based on planar template, which is flexible and very simple. A planar checkerboard pattern is a frequently used calibration object because it is very simple to produce, it can be printed with a standard printer, and has unique corners which are easy to detect [104].

2.3.3 Monocular vision positioning system

Monocular vision positioning uses a single camera to determine the pose estimation of a mobile device or static objects. Another type of vision positioning system is called binocular vision. Binocular stereo vision uses two cameras to estimate location of a mobile robot. Although it has the advantage of better performance in the regard of accuracy, but it is more expensive and complex to compute [105]. While monocular vision on the other hand is simple to set-up and of low cost. Information collected from the environment captured by the camera can be in form of an image or video. This information is therefore processed to estimate the position and orientation of the robot per time. This poses a spatial relationship between the 2D image captured and the 3D points in the scene. According to Navab [106], the use of marker in augmented reality (AR) is very efficient in the environment. It increases robustness and reduces computational requirement. However, there are exceptional cases where markers are placed in the area and they need re-calibration from time to time. Therefore, the use of scene features for tracking in place of markers is reasonable especially when certain parts of the workplace do not change over time. Placing fiducial markers [47] is a way to assist robot to navigate through its environments. In new environments, marker often need to be determined by the robot itself, using sensor data collected by IMU, sonar, laser and camera. Markers' locations are known, but the robot position is unknown, and this is a challenge for tracking a mobile robot. From the sensor readings, the robot must be able to infer its most likely position in the environment. With monocular vision (one camera), a good solution in terms of scalability and accuracy is provided. With the aid of other sensors such as ultrasonic sensor or barometric altimeter, the monocular vision can also provide the scale and depth information of the image frames. To calculate the pose of the mobile robot with respect to the camera based on the pinhole camera model, the monocular vision positioning system [107], was used to estimate the 3D camera from 2D image plane [108]. The relationship between a point in the world frame and its projection in the image plane can be expressed as:

$$\lambda p = MP \tag{2.1}$$

where λ is a scale factor, $p = [u, v, 1]^T$ and $P = [X_w, Y_w, Z_w, 1]^T$ homogenous coordinates of p and P , and M is a 3×4 projection matrix.

Equation (2.1) can further be expressed as:

$$\lambda \begin{bmatrix} u \\ v \\ i \end{bmatrix} = M(R_{wc}t_{wc}) \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \tag{2.2}$$

The projection matrix depends on both camera intrinsic and extrinsic parameters. The intrinsic parameters contain five parameters: focal length f , principal point u_0, v_0 and the skew coefficient between x and y axis and is often zero.

$$M = \begin{bmatrix} a_x & \gamma & u_0 \\ 0 & a_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} \tag{2.3}$$

Extrinsic parameters: R, T defines the position of camera center and the camera's heading in world coordinates. Camera calibration is to obtain the intrinsic and extrinsic parameters. Therefore, the projection matrix of a world point in the image is expressed as:

$$C = -R^{-1}T = -R^T T \tag{2.4}$$

where T is the position of the origin of the world coordinate, and R is the rotation matrix. For this research, camera calibration was done offline using MATLAB Calibration Toolbox [101]

2.3.4 Landmarks

Landmark is the feature information recognized through robot's sensors perception. For an autonomous robot localisation and navigation, how to identify landmarks quickly and accurately plays an important role. Robot navigation system based on landmarks, research areas include landmark selection, landmark design, landmark detection, landmark navigation, environmental characterization and path planning, etc. Generally, landmarks are classified into two types: markerless (also known as natural landmark) and marker-based (also known as artificial landmark) [109].

2.3.4.1 Artificial landmark

Artificial landmarks refer to the special designs of the objects or markers, placed in an environment that can be detected by laser, infrared, sonar and vision sensors. The peculiarity of the marker is important with the features for quick recognition and high reliability, these landmarks can be identified accurately at various visual conditions [109, 110]. Localisation based on artificial landmarks is used more widely than other methods because the artificial landmarks are easy to detect which could enhance precision. An artificial landmark could be any object whether static or mobile which could vary in size, shape feature of color as long as it is placed in the environment with purpose of robot localisation. The author in [99] use a sticker and LED array as an artificial landmark. These makers are easier to detect and describe because the details of the objects used are known in advance. These methods are used because of their simplicity and easy setup. However, the possibility of adopting the method in a large environment may not be feasible because of the numerous number of markers identified in the area.

2.3.4.2 Natural landmark

Natural landmarks are objects or features that are part of the environment and have a function other than robot navigation. Examples of natural landmarks are corridors, edges, doors, wall, ceiling light, lines, etc. The choice of features is vital because it will determine the complexity in the feature description, detection and matching [14]. Although the natural landmarks have little influence on the environment, it is rarely used in the practical applications for its low stability and bad adaptability. Visual features are divided into three categories: point feature, line feature, block feature. Point features are the easiest to extract, relatively stable and contain abundant information [111]. A lot of work has dealt with the issue of using natural landmarks to extract feature that will aid robot localisation using Scale-Invariance feature Transform (SIFT) features [112] and Speeded Up Robust Feature (SURF) features [113, 114].

2.3.5 Object recognition and feature matching

In this subsection we presented the proposed method of object recognition and matching features. Object recognition under uncontrolled, real-world conditions is of vital importance in robotics. It is an essential ability for building object-based representations of the environment and for the manipulation of objects. Different methods of scale invariant descriptors and detectors are currently being used because of their scale flexible and affine transformations to detect, recognize and classify objects. Some of these methods are Oriented Fast and Rotated BRIEF (ORB), Binary Robust Invariant Scalable Keypoints (BRISK), Difference of Gaussians (DoG), FERNS [115] SIFT [24] and SURF [113]. More details of these method can be found in reference [116]. Object detection and recognition can be done using computer vision whereby an object will be detected in image or video sequence. The recognised object is used as a reference to determine the pose of a mobile device. Basically, object detection can be categorised into three aspects: appearance based, color based and features based. All these methods have their advantages and limitations [117].

Appearance based objects are recognised based on the changes in color, size and shape. The techniques used are edge matching, divide and conquer search, greyscale matching, gradient matching etc. The color based techniques are based on the RGB features to represent and match images. They provide cogent information for object recognition. While the feature-based technique finds the interest points of an object in image and matches them to the find object in another image of similar scene. Features extracted are surfaces, patches, corners and linear edges. The methods used to extract feature are interpretations trees, hypothesize and test, pose consistency, geometric hashing, SIFT, and SURF.

Generally, finding the correspondences is a difficult image processing problem where two tasks have to be solved [118]. The first task consists of detecting the points of interest or features in the image. Features are distinct elements in the images; examples are corners, blobs, edges. The most widely used algorithm for detection includes the Harris corner detector [119]. It is based on the eigenvalues of the second moment matrix. Other types of detectors are correlation based: Kanade-Lucas-Tomasi tracker [120] and Laplace detector [121]. For feature matching, the two most popular methods for computing the geometric transformations are: Hough transform and RANSAC algorithm [113, 116, 122]. They could estimate parameter with a high degree of accuracy even when a substantial number of outliers are present in the data set.

2.3.5.1 Speeded-Up Robust Features (SURF)

SURF was first introduced by Herbert Bay *et al.* [113]. SURF outperforms formerly proposed scheme SIFT with respect to repeatability (reliability of a detector for finding the same physical interest points under different viewing conditions), distinctiveness, and robustness, yet can be computed and compared much faster. The descriptors are used to find correspondent features in the image. SURF detect interest points using (such as blob) using Hessian matrix because of its high level of accuracy (See (2.5) and (2.6)). This is achieved by relying on integral images for image convolutions; by building on the strengths of the leading existing detectors and descriptors (specifically, using a Hessian matrix-based measure for the detector, and a distribution-based descriptor); and by simplifying these methods is essential. This leads to a combination of novel detection, description, and matching steps. SURF is used to detect key points and to generate its descriptors. Its feature vector is based on the Haar Wavelet response around the interested features [117]. SURF is a scale-and rotation-invariant, that means, even with variations on the size and on the rotation of an image, SURF can find key points.

$$I(x) = \sum_{i=0}^{i \le x} \sum_{j=0}^{j \le y} I(x,\ y) \tag{2.5}$$

There is a point $X = (x,\ y)$ in an image I, Hessian matrix $H(X,\ \sigma)$ in X at scale σ is defined as:

$$H(X,\ \sigma) = \begin{bmatrix} L_{xx}(X,\ \sigma) L_{xy}(X,\ \sigma) \\ L_{xy}(X,\ \sigma) L_{yy}(X,\ \sigma) \end{bmatrix} \tag{2.6}$$

where $L_{xx}(x,\sigma)$ is the convolution of the Gaussian second order derivative $\frac{\partial^2}{\partial x^2} g(\sigma)$ with the image I in point X and derivative for $L_{xy}(x,\sigma)$ and $L_{yy}(x,\sigma)$.

2.3.5.2 Random Sample Consensus (RANSAC)

RANdom SAmple Consensus (RANSAC) is feature matcher which works well with SURF to match object detected by SURF in images. RANSAC was first published by Fischler and Bolles [122] in 1981 which is also often used in computer vision. For example, to simultaneously unravel the correspondence problem such as, fundamental matrix related to a pair of cameras, homograph estimation, motion estimation and image registration [123-128]. It is an iterative method to estimate parameters of a mathematical model from a set of observed data which contains outliers. Standard RANSAC algorithm of this method is presented as follows:

Assuming a 2D image corresponds to a 3D scene point. $(x_i,\ wX_i)$. Let us assume that some matches are wrong in the data. RANSAC uses the smallest set of possible correspondence and proceed iteratively to increase this set with consistent data.

- draw a minimal number of randomly selected correspondences S_k (random sample)
- compute the pose from these minimal set of point correspondences using () POSIT, () DLT
- determine the number C_k of points from the whole set of all correspondence that are consistent with the estimated parameters with a predefined tolerance. If C_k>C^* then

we retain the randomly selected set of correspondences S_k as the best one: $S*$ equal S_k and $C*$ equal C_k

- repeat first step to third step.

The correspondences that partakes to the consensus obtained from $S*$ are the inliers. The outliers are the rest. It has to be noted that the number of iterations, which ensures a probability p that at least one sample with only inliers is drawn can be calculated. Let p be the probability that the RANSAC algorithm selects only inliers from the input data set in some iteration. The number of iterations is denoted as [129-131]:

$$k = \frac{\log(1-p)}{\log(1-(1-w)^n)} \tag{2.7}$$

where w is the proportion of inliers and n is the size of the minimal subset from which the model parameters are estimated. Steps to detect and recognise object (marker) in a scene:

- Load training image
- Convert the image to grayscale
- Remove lens distortions from images
- Initialise match object
- Detect feature points using SURF
- Check the image pixels
- Extract feature descriptor
- Match query image with training image using RANSAC
- If inliers > threshold then
- Compute Homography transform Box
- Draw box on object and display

More details of the algorithm will be presented in Chapter 3 under result and discussion.

2.3.6 Rotation representations

Several representations can be used to obtain object (body) orientation. The rotation representations are axis angle, rotation matrix, Euler angles and quaternion [132-134].

Axis-angles: Axis angle is used to represent the rotation of an object in 3D. Rotation can be represented by a unit vector and an angle of revolution about that vector. Rodrigues' rotation formula is an efficient algorithm for rotating a Euclidean vector, given a rotation axis and an angle of rotation. It provides an algorithm to compute the exponential map without computing the full matrix exponential.

If v is a vector R^3 and e is a unit vector rooted at the origin, an axis of rotation about which v is rotated by an angle θ, the Rodrigues' rotation formula to obtain the rotated vector is given as:

$$v_r = (\cos\theta)v + (\sin\theta)(e \times v) + (1-\cos\theta)(e.v)e \tag{2.8}$$

Rotation matrix: Rotation matrix is a matrix whose multiplication with a vector rotates while preserving its length. The 3x3 rotation matric is denoted by $SO(3)$. Thus, if R$\in$ $SO(3)$, then det R=± 1 and $R^{-1} = R^T$

For three dimensions, using the right-hand rule, the rotation is given as

$$R_x(\theta) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix} \tag{2.9}$$

$$R_y(\theta) = \begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix} \tag{2.10}$$

$$R_z(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{2.11}$$

The general rotation can be obtained using matrix multiplication $R = R_z(\alpha)R_y(\beta)R_x(\gamma)$ where α, β, γ are represented as yaw, pitch and roll respectively.

Euler angles: Euler angles are intuitive to interpret and visualize and that's why that they are still widely used today**.** Euler angles represent an orientation as a series of three sequential rotations from an initial frame. Each rotation is defined by an angle and a single axis of rotation chosen among the axes of the previously transformed frame. Definition of Euler angles consists in giving the order of three successive rotation axes. For instance, XYZ, ZXZ. In particular, the angles of classification XYZ are also named roll (rotation about the x-axis), pitch (rotation about the x-axis) and yaw (rotation about the x-axis). As such, there are six possible Euler angle conventions. When the three axes are different, one refers to the angle triplet as Tayt-Briant angles. Euler angles are mostly used because of simplicity and ease of understanding. However, they have singularities when compared to quaternion when used to integrate incremental changes in attitude.

Quaternion: Quaternions are a number system that extends complex numbers applied to 3D space. When used to represent rotation, it is called rotation quaternion and if it is used for orientation it is called orientation quaternions or attitude quaternions. It can be expressed as:

$$a + bi + cj + dk \tag{2.12}$$

where a, b, c are real number and i, j, k are symbols as unit vectors.

Although quaternion have no singularities and the representation is well-suited to integrating the angular velocity, but the quaternion parameters do not have intuitive physical meaning and a quaternion must have unity norm to be a pure rotation [133]. They are used for a sensor that can be oriented anywhere in space. Table 2.3 shows the advantages and disadvantages of rotation representations.

Table 2.3. Characteristics and applications of rotation representations.

Rotation representation	Advantages	Disadvantages	Applications
Axis angle	Relatively simple, uses very little storage, efficient to compute	Still suffers from the “edge” and distance preserving problems of Euler Angles. The representation is very intuitive. Difficult to compose rotations and interpolate rotations.	Rigid body dynamics
Rotation matrices	Easy to understand and compute	Difficult to compose rotations and interpolate.	Materials field and texture community, aeroplanes
Euler angles	Simple and easy to comprehend. Widely used. Good for decomposing rotations into individual DoF.	Gimbal lock and ambiguity	Material science, aviation, rigid bodies
Quaternions	Representation of rotation is numerically convenient. They don’t suffer from ambiguity since they only represent a single rotation with a well-defined axis.	Complex. They don’t have an intuitive representation	Computer animation. Used to represent transformations of orientations of graphical object

2.4 RELATED WORK ON SENSOR FUSION TECHNIQUES

Several definitions of sensor fusion are given in the literature. Sensor fusion or data fusion as defined by Joint Directors of Laboratories (JDL) workshop [134] is a multi-level process dealing with the association, correlation, combination of data and information from single and multiple sources to attain refined position, identify estimates and complete timely assessments of situations, threats and their significance. Also, Hall and Llinas [135] provided the following well-known definition of data fusion: "data fusion techniques combine data from multiple sensors and related information from associated databases to achieve improved accuracy and more specific inferences than could be achieved by the use of a single sensor alone". According to the authors in [136, 137], sensor fusion was defined as the cooperative use of information provided by multiple sensors to aid on performing a function while several others authors [138, 139] defined data fusion algorithms as the combination of data from multiple sources in order to enhance the performance of mobile robot. Regardless of different definition given, sensor fusion is the integration of information from multiple sources to improve accuracy and quality content, also with the aim to reduce cost. The technique finds wide application in many areas of robotics such as object recognition, environment mapping, and localisation. Fusion techniques are therefore regarded as the most appropriate method to track objects and determine their locations. The advantages of sensor fusion are as follows: reduction in uncertainty, increase in accuracy and reduction of cost. It is therefore suggested by various researchers that to attain a level of accuracy, integration of more than one sensor is most suitable because the inadequacy of one sensor can be complemented by another. For example, the image captured by the camera was used to correct the abnormalities of inertial sensors [140, 141]. The data fusion techniques deployed is influenced by the objective of applications in which it aids in building a more accurate world model for the robot to navigate and behave more successfully. The three fundamental ways of combining sensor data are the following [138, 142]:

Competitive: The sensors are configured competitively to produce independent measurements of the same property i.e. diverse kinds of sensors are used to measure same environment characteristic. This means data from different sensors can be fused or measurement from a single sensor taken at different periods can be fused. A special case of competitive sensor fusion is fault tolerance. Fault tolerance requires an exact requirement of the service and the failure modes of the system. This configuration therefore reduces the risk of incorrect indication that could be caused by one of the sensors. Most importantly, this might result in an increase in the reliability, accuracy or confidence of data measured by the sensors. This technique can also provide robustness to a system by combining redundant information [105, 106]. However, the robust system provides degraded level of service in the presence of faults while this graceful degradation is weaker than the accomplishment of fault tolerance. The method performs better in terms of resource need and work well with heterogeneous data sources. Another name for competitive sensor configuration is also called a redundant configuration. An example of competitive is the reduction of noise by combining two overlaying camera images.

Complementary: This type of sensor configuration ensures that the sensors do not depend on each other but rather complement themselves with different measurements. This resolves the incompleteness of sensor data. This type is the most common for localisation. Example is when vision is complemented by the short coming of accumulated errors in IMU. Another example of complementary configuration is the employment of several cameras each observing different area of the mobile robot surrounding to build up a picture of the environment. Generally, fusing complementary data is simple, since the data from independent sensors can be appended to each other, but the disadvantage is that under certain conditions the sensors maybe ineffective, such as when camera used in poor visibility [107].

Cooperative: This method uses the information made available by the two separate sensors to originate data that would not be obtainable from the single sensors. An example of a cooperative sensor configuration is stereoscopic vision by combining two dimensional images from two cameras at slightly dissimilar viewpoints in which 3D of the detected scene is derived. According [107], cooperative sensor configuration is the most difficult system to design due to their sensitivity to imprecisions in all individual participating sensors. Thus, in contrast to competitive fusion, cooperative sensor fusion generally decreases accuracy and reliability.

Conclusively, competitive fusion combinations increase the robustness of the perception, while cooperative and complementary fusion provides extended and more complete views. The methods particularly used in the fusion level are subject to the availability of components. Furthermore, these three combinations of sensor fusion are not mutually exclusive. Therefore, many applications implement aspects of more than one of the three types.

2.4.1 Classification of sensor fusion algorithms

Sensor fusion algorithms are needed to translate the different sensory inputs into reliable estimates and environment models that can be used by other navigation subsystems. The methods usually adopt iterative algorithms to deal with linear and non-linear models. In order to localise robot, many sensors have been adopted and fusion methods developed. These algorithms are a set of mathematical equations that provide an efficient computational means to estimates the state of a process. Some of the sensor algorithms used are categorised into the following [143]:

2.4.1.1 State estimation method

The state estimation methods are used to determine the state of a desired system that is continuously changing given some observations or measurements. State estimation phase is a common step in data fusion algorithms because the target's observation could come from different sensors or sources, and the final goal is to obtain a global target state from the observations. Kalman filter, extended Kalman filter and particle filter are briefly discussed in this section and they are considered afterwards for this study.

Kalman filter

Kalman filter (KF) is an efficient estimator used in various fields to estimate the unknown state of the system. Several applications were developed with the implementation of Kalman filter such applications include navigation, localisation and object tracking. It involves using vision camera to perform real time image processing for robot tracking. Kalman filter is established to estimate the positions and velocities of vehicles or any moving object and provide tracking on such objects at a visible condition. Kalman filter is an algorithm that estimates the state of a discrete time-controlled process described by the linear stochastic equation. It processes all available measurements, regardless of their precision, to estimate the current value of the variables of interest. Kalman filters are well-known tools in theory of stochastic dynamic systems, which can be used to improve the quality of estimates of unknown quantities [144]. It is one of the most useful and common estimation techniques where it is easy to implement on linear systems. Equations for Kalman filter are given as follows [145]:

$$\hat{x}_k = F_k \hat{x}_{k-1} + B_k u_k \tag{2.13}$$

$$P_k = F_k P_{k-1} F_k^T + Q_k \tag{2.14}$$

where vector $\hat{x}_k$ is the estimate state of the system x_k. P_k is the predicted covariance matrix. F is the matrix that denote the dynamics of the system. B is the control matrix and Q is the noise covariance. The Kalman filter equations are used to generate new estimates with the addition of an external unit for correction. KF involve another stage to update the estimate. This is given by equations below:

$$\hat{x}'_k = \hat{x}_k + K'(z_k - H_k \hat{x}_k) \tag{2.15}$$

$$P'_k = P_k - k' H_k P_k \tag{2.16}$$

where $K' = P_k H_k^T (H_k P_k H_k^T + R_k)^{-1}$

From the above equations: z_k is the measurement vector which is a reading from the sensors. H is the transformation matrix, R is the covariance matrix of the measurement noise and k is the time interval. The Kalman gain (K) describes the amount of update needed at each recursive estimation which can be as the weighting factor that considers the relationship between the accuracy of the predicted estimate and the measurement noise. In order to use Kalman filtering it is important to analyze the statistical behaviour of the value to be measured. They are optimal estimators, which mean the initial uncertainty is Gaussian and the observation model and system dynamics are linear functions of the state. Most of the real time problem the systems may not provide linear characteristic, so we use extended Kalman filter, which will linearize the system. The main advantage of Kalman filter is its computational efficiency but it can represent only unimodal distributions. So Kalman filters are best when the uncertainty is not too high.

Extended Kalman Filter

The modified Kalman filter known as the extended Kalman filter (EKF) is an optimal approach for implementing nonlinear systems. It is used widely for state estimation because it can estimates the states with a slight computational load. EKF is the most efficient probabilistic solution to simultaneously estimate the robot position and orientation based on some interoceptive and exteroceptive sensor information. Comparing Kalman filter to EKF, author [146] proves that that the EKF algorithm improves the performance of robot localisation, reduces error in the calculation of mobile robot pose and the combination of all available sensors gives optimal result [147]. This filter is used for intermittent measurement since it provides adequate information for the estimation method. More information is given in Chapter 3.

Particle filter

It is known that the Kalman type filters are not suitable for state estimation for systems with non-Gaussian noises and/or strong nonlinearities since the Gaussian assumption on the state posterior is no longer valid. Particle filtering (PF), with the capability of approximating probability density functions (PDFs) of any form, has received considerable attention among researchers. PF method is a sequential Monte Carlo (SMC) technique solution of the state estimation problem, using the so-called Sequential Importance Sampling (SIS) algorithm and including a resampling step at each instant. This method builds the subsequent density function using numerous random samples called particles. Particles are propagated over time with a combination of sampling and resampling steps. At each iteration, sampling step is employed to discard some particles, increasing the relevance of regions with a higher posterior probability. The particle filter algorithm is comprised of the following steps [136, 148-152]:

Particle generation:

Generate $N\{x_1(0), x_2(0), x_3(0),..., x_N(0)\}$ initial particles according to the initial probability density function (PDF) $p(x(0))$

Prediction:

For each particle $x_i(k)$, propagate the $x_i(k+1)$ particle according to the transition PDF $p(x(k+1)|x(k))$. Here, each particle accounts for the sum of the random noise to simulate the noise effect.

Sampling:

For each particle $x_i(k+1)$, generate $w_i(k+1) = p[z(k+1)|x_i(k+1)]$ (2.17)

Normalization and rejected sampling:

Weights of the particles are normalized. Particles with low weight are removed and particles with high weight are duplicated such that each particle has the same weight.

PF is considered as an alternative for real-time applications, which are typically approached by model based traditional Kalman filter technique implementations. With the advantages of accuracy and stability, PF is currently being considered in the field of traffic control (car or people video monitoring), military field (radar tracking, air-to-ground passive tracking), mobile robot positioning and self-localisation.

KF is one of the most well-known and often-used tools for stochastic state estimation from noisy sensor measurements. Under certain assumptions, the KF is an optimal, recursive data processing or filter algorithm. The filter takes into account the different uncertainties and error sources that disturb the robot system and measurements. It processes all available measurements to estimate the state, both accurate and inaccurate measurements. It uses knowledge of the system and sensor dynamics, probabilistic descriptions of the system and measurement noises, and any available data about the initial values of the state. Kalman filter addresses the problem of estimating the state of a noisy system that can be described by a linear system and a linear measurement model. It is implemented for combining the sensory data for estimating the state of the robot which is the location and current direction of the robot. In an iterative manner, KF considers the prior information of the noise features to compensate and to filter out the noise. But issues arise during localization when trying to model the noise that is only an approximation and does not specify the noise real distribution. KF is well effective for linear system, however the implementation introduces delay in the systems processing and it greatly depends on the prediction models, an error in creating prediction model will surely cause the output data to deviate from the required reference data. For a nonlinear system, EKF is used to estimate the pose of the mobile robot employing the prediction and correction of a nonlinear system model. EKF has the benefits of simultaneously estimating the IMU sensor's systematics errors and corrects the positioning errors. The algorithm can accurately predict and correct its state estimates. Since the system function that describes state transition of the location is nonlinear, it linearizes the system function around recent state estimates. EKF is used to correct the state estimates of the robot and thus bring it closer to the true states. When EKF corrects the states estimates, uncertainty decreases. EKF linearizes the nonlinear system and measurement function. The algorithm integrates a new state estimates once a measurement has been incorporated. This permits the EKF to make future state and measurement predictions more accurate. Since the nonlinearities in the measurement model are caused by the orientation of the robot, the EKF suffers from significant estimation problems caused by EKF linearization. The EKF is difficult to tune and often provides unreliable estimates if the system nonlinearities are severe, the reason being EKF relies on linearization to propagate the mean and covariance of the state. For this purpose, particle filter was presented. Particle filter copes much better with non-linear models and has no limitations when it comes to the non-linearity of the application including non-Gaussian noise. PF is considered to solve the issue of measurement that is affected by non-Gaussian noise. Another benefit of PF is that it is easier to implement when compared to EKF, it allows the analysis of complex systems which are non-linear and non-Gaussian. The goal is to deal with probability distributions and model them correctly. Particle filters are a non-parametric approach for solving complex models. The algorithm tackles the errors as long there are sufficient particles available, an optimal solution can be obtained effectively. Because of uncertainties in sensor data and environmental factors there is always the risk that the error of the pose estimation increases significantly or that it diverges completely. Very often navigation and other performances of robots rely on a valid pose estimation. The filter is much better suited than Kalman filter to represent ambiguities and to cope with localization failures. The realization of PF does not require the process of linearizing non-linear models. Large number of samples guarantees sure convergence to the true probability density function. These filters are considered because of their simplicity, consistency and provision of sufficient information of estimation process compared to others.

2.4.1.2 Decision fusion method

Decision fusion is one form of data fusion that combines the decisions of multiple classifiers into a common decision about the activity that occurred. All the fusion methods in this group try to reduce the level of uncertainty by maximizing a measure of evidence [153]. These techniques frequently use symbolic information, and the fusion process requires to reason while accounting for the uncertainties and constraints. The two types of decision method discussed here are Bayesian Approach and Dempster-Shafer Approach.

Bayesian Approach

Bayesian approach is a basic method to deal with conditional probability more precisely it relates the condition probability of more than two events. They are practically used for more complex relationship description [154]. The method provides a theoretical framework for dealing with this uncertainty using an underlying graphical structure. They are ideal for taking an event that occurred and predicting the likelihood that any one of several possible known causes was the contributing factor. Bayesian method can be mathematically presented as [148]:

$$P(C \mid D) = \frac{P(C \mid D)P(C)}{P(D)} \tag{2.18}$$

where $P(C)$ is the probability of event C without any effect of any other event. $P(D)$ is the probability of the event D without any effect of any other event and $P(D \mid C)$ is the probability of event D given that event C is true. The result of $P(C \mid D)$ condition probability will be in range between zero and one [1 0]. Which means either the event $P(C \mid D)$ will occur or not. Bayesian method is computationally simpler and has higher probabilities for correct decision and it provides point estimates and posterior PDF [120]. However, they have the following demerits: difficulty in describing the uncertainty of decision, complexity when there are multiple potential hypotheses and a substantial number of events that depend on conditions, difficulty in establishing the value of prior probabilities. Bayesian method is applicable to solve image fusion, where no prior knowledge in available. Also, it is applied in robotics learning by imitation. The approach enables the robot to study internal models of their environment through self-experience and employ the model for human intent recognition, skill acquisition from human observation. [155].

Dempster-Shafer

Dempster-Shafer (D-S) has become very famous in which its application extends to pattern recognition methods which are widely used in signal solving and recognition. The method has a better adaptability of grasping unknown and uncertain problem when it is regarded as an uncertainty method. It also provides a vital formula which fuse diverse evident of different sources. Evidential reasoning based on Dempster-Shafer theory is used in data fusion, and cooperation strategies presented to avoid invalid sensing information. Dempster-Shafer theory has been considered for a variety of perceptual activities including sensor fusion, scene interpretation, object target recognition, and object verification. In [144], D-S theory was successfully used in building occupancy map to improve reliability. The approach is more robust to perturbations such as noise and imprecise prior information [155]. The method is based on concept of combining information from different sources such as sensors.

It is a fusion method that uses belief and plausibility values to represent the evidence and corresponding uncertainty [156] [157]. The method uses 'belief' rather than probability. Belief function is used to represent the uncertainty of the hypothesis [158]. The hypothesis is represented by a probability mass function 'm'. the amount of belief to a hypothesis (A) is denoted by a belief function [159]:

$$Bel(A) = \sum_{B \subseteq A} m(B) \tag{2.19}$$

Equation (2.19) is the sum of the mass probabilities assigned to all subsets of A by m. The availability of two or more evidence is integrated using the combination rule in (2.20).

$$m(A) = \frac{\sum_{\substack{i,j \\ B_i \cap C_j = A}} m_1(B_i).m_2(C_j)}{1-k} \tag{2.20}$$

where $1-k$ is a normalization factor in which k is the total of all non-zero values given to the null set hypothesis $\varnothing$. The decision on the class of a feature can be decided based on a maximum belief decision rule, which is assigned a feature to a class A if the total amount of belief supporting A is more than that supporting its negation:

$$Bel)A) \geq Bel(\bar{A}) \tag{2.21}$$

Table 2.4. Related work done on fusion algorithms.

Fusion Algorithm	Author	Classification of Fusion method	Contributions
Kalman filter	Mahood *et al.* [160]	State estimate	The fusion of visual navigation system and inertial navigation system are integrated using Kalman filters to provide accurate localization information about the mobile robots.
Extended Kalman filter	Faisal *et al.* [161]		The localisation system uses the extended Kalman filter combined

			with Infrared sensor to solve the problems of dead-reckoning.
	Andre *et al.* [162]		The localisation system proposed was based on the fusion of odometry and landmark detection to determine the position. The result shows a better result than using a single sensor in terms of minimised error.
	Sasiadek *et al.* [163]		EKF is applied to fuse odometry and sonar signals to solve the issue of navigation, control and guidance. The result shows the method was suitable to determine the estimation for autonomous mobile robot.
	Hoang *et al.* [164]		Using the proposed algorithm, the study shows that the estimated output values are close to the true value when the measurement from all sensors (encoder, compass sensor, laser range finder and omni-direction camera) are fused together to determine the robot position.
Particle filter	Raaj *et al.* [165]		Particle filter fused optical camera, sonar and odometry measurement data to track and localise object. This method was

			used to deal with issues such as poor lighting conditions and hazing over large distances with insignificant features to track.
	Wanfeng Ma *et al.* [166]		The approach fused the inertial navigation systems and light detection and ranging (LiDAR) data to correct robot position error, velocity error and orientation error.
	Jain *et al.* [167]		The application of PF was introduced to give reliable estimation of the state vector of a mobile robot with the fusion of odometry and laser range finder sensors for efficient control.
	Ren *et al.* [168]		The approach was used to estimate object pose parameter from the fusion of inertial and GPS measurement data. The method further improve the stability and the accuracy objects between virtual and real world. Using PF achieves high tracking accuracy, stability and robustness.
	Lee *et al.* [169]		The proposed algorithm combines the range information obtained from a low-cost IR

			scanner with the SIFT-based vision information obtained from a monocular to robustly estimate the robot pose.
Bayesian Network	Abdulhafiz *et al.* [170]	Decision	This approach was used to fuse sensor data to estimate the position of a mobile robot as well to handle the issue of data uncertainty and inconsistency.
	Motomura *et al.* [171]		The study developed a system that combines local information for localisation using Bayesian Network
	Hongjun *et al.* [172]		The author proposed the use of this method to represent and integrate the sensor information, the robot pose and the sensing actions.
	Vladareanu *et al.* [173]		The study propose the use of Bayesian approach of SLAM to avoid obstacles and maintain dynamic stable control for motion on rough terrain in a non-stationary and non-structured environment.
	Premebida *et al.* [174]		The problem of semantic place categorization in mobile robotics is addressed with focus on a

			probabilistic approach for classification using 2D laser scanner data. The method was successfully applied and from the performance it showed that the method has faster implementation and low complexity.
Demptser-Shafer	Soleimanpour *et al*. [175]	Decision	With the integration of vision and encoder, the approach was able to reduce localization error and improve the performance.
	Hyunki *et al*. [176]		To solve the problem of localisation, the author proposed the use of sensor fusion algorithm based on Dempster-Shafer to fuse laser range information and SIFT features. The uncertainty of the robot localisation decreases and the accuracy of the algorithm increased. Results from experiment confirms the usefulness and robustness of the fusion method.
	Gören *et al*. [177]		With the application of D-S to sensor data, it is expected to have more reliable sensor data. The obtained result can be useful for autonomous mobile robot to decide its optimal path to the

			target. Another contribution of the work is that it enables robot behaviour to be more stable during motion.
	Valente *et al.* [178]		D-S was used to model the environment perception by the sensor. The author combined a 2D laser scanner and a stereo camera to deal with the sensor uncertainty. From the result, the proposed method has more reliable representation of the environment.
	Carlson *et al.* [179]		The study considered using the D-S to present a potential solution to the problem of detecting the use of inappropriate sensors to an impaired or unknown environment.
	Erfani *et al.* [180]		D-S was used for the integration and processing of sensor data in robot location to achieve the best estimate of positioning according to the unstable environmental conditions.

2.4.2 Importance of sensor fusion techniques

Techniques that employ sensor fusion methods expect several benefits over single sensor systems. Joint information reduces the set of ambiguous interpretations of the measured value. The following advantages can be expected from the fusion of sensor data [142].

Reduction in uncertainty

Data provided by sensors is sometimes subjected to some level of uncertainty and discrepancy. Multi-sensor data fusion algorithms reduce the uncertainty by combining data from several sources [197]. It’s therefore imperative to compensate using other sensors by fusing their data together using data fusion algorithms. Authors in [198] was able to minimize uncertainty in robot localisation based on EKF and PF. The measurement from the kinetic sensor was used to correct the error accumulated by odometry in order to estimate the pose of the mobile robot.

Increase in accuracy and reliability

Multiple sensor suites have an inherent redundancy which enables the system to provide information even in case of partial failure.

Extended spatial and temporal coverage

One sensor can look where others cannot respectively or perform a measurement while others cannot. An example is inertial sensor such as accelerometer or gyroscope and vision. The coverage of a camera as vision sensor cannot be compared to the use of accelerometer which only takes measurement about the navigation route.

Improved resolution

When multiple independent measurements of the same property are fused, the resolution of the resulting value is better than a single sensor’s measurement.

Reduce system complexity

System where sensor data is preprocessed by fusion methods, the input to the controlling application can be standardized independently of the employed sensor types, thus facilitating application implementation and providing the possibility of modifications in the sensor system regarding number and type of employed sensors without modifications of the application software.

Table 2.5. Review of different sensor fusion algorithms.

Sensors	Fusion Method	Contributions
Visual and inertial [183]	UKF	The algorithm employs a fusion algorithm that provides accurate motion estimates of both calibration parameters and the local scene structure. The metric scene structure was recorded from the camera and the IMU alone. Their findings show that it is possible to accurately self-calibrate the sensors without using a known calibration target or another calibration object.
Inertial and magnetic sensor [184]	Variable State Dimension-EKF	The algorithm fused IMU and magnetic sensor data to tackle the difficulties in tracking abrupt magnetic distortions. With the algorithm better accuracy and effectiveness of the estimated quaternion rotation and good compensation for magnetic disturbances was achieved.
Visual, inertial and magnetic sensor [185]	EKF	The EKF algorithm was used to address the problem of estimating ego-motion of a hand-held IMU-camera system to track position and orientation of human body segments.

		The orientation was effectively achieved using EKF-based sensor fusion method based on inertial/magnetic measurement.
Encoder, inertial sensor, active beacon [186]	UKF	UKF was proposed for vision tracking system. The fusion algorithm accurately estimates the pose estimate of the mobile robot by combining information received from encoder, inertial sensor and active beacons. The algorithm was also used as an advance filter to minimize the position and orientation errors of the sensors.
Visual and inertial sensors [187]	EKF	The EKF algorithm integrated visual and inertial sensors measurement to tackle the issues of sensor noise and model inaccuracy in the area of human arm motion tracking system.
Visual and inertial sensors [188]	EKF &UKF	EKF and UKF fusion algorithms were compared to estimate simultaneous motion and structure estimate by integrating vision and inertial sensors data was tested in robotic motions and less smooth handheld motions. Their result shows that EKF and UKF have similar accuracy, but UKF has higher computational power.
Visual and inertial sensor [189]	PF	Computer vision and inertial sensors were combined to provide real-time position. Markov localisation using particle filter as a fusion algorithm was adopted to determine the location of the system.
INS and VNS [190]	KF	The work presented a seven error states for

		an accurate and stable data fusion filter which integrates the position of a mobile robot from VNS with the position from INS to accurately localize the robot using KF for data fusion.
Visual system and robot odometry [191]	KF	Kalman filter was implemented to combine visual system and odometry measurement to compute robot's actual position. The results proved more precise computation and improve localisation accuracy.

2.5 FUTURE RESEARCH AREAS

Navigation and localisation of a mobile robot in an arbitrary environment are a challenge due to the complexity and variety of environments, methods and sensors that are involved. It is therefore necessary to continue to research on new systems and new methods with the aim to solve specific sensor fusion problems for robot navigation and localisation. While not totally absent from the literature, several directions seem to call for further investigation.

3D Indoor Environmental modelling: 3D models of indoor environments are significant in many applications, but they usually exist only for newly constructed buildings [191]. For robot navigation purpose, 3D models are required in an indoor operation environment to ensure safe movement. The model is also expected to be used for recognition and location by robots. To develop a method to model a 3D simplicity, accuracy must first be put into consideration. A 3D model can convey more useful information than 2D maps used in many applications. For example, in an indoor environment where additional features are present and are also unresolved problems in modelling. This kind of environment requires more sophisticated models in order to determine the ability characteristics of the environment. Several methods are adopted in modelling the environment. Reference in [192] proposed a method of acquiring 3D models by a mobile robot with a laser scanner and a panoramic camera while Thrun *et al.* [193] developed a multi-planar model from dense range data and image data using an improved expectation-maximization (EM) algorithm. Some authors worked with generation of precise 3D models using a large amount of data and elaborate statistical and geometrical estimation technique. Environment models are required for localisation, object recognition, or manipulation. Recently, 3D models are usually obtained by hand-guided scanning which is very hard and time-consuming task for human operator. Therefore, a robotic system to obtain 3D models of environment is highly beneficial [194].

Landmarks and feature extraction: Perception-based localisation methods using vision are very active research areas, especially in topics related with the identification of objects and the pose estimation of the identified objects [194]. Another aspect to investigate is the appearance changes of target objects over time; this also as a research area has gained much attention in the literature but with the limitation of robust detection algorithm.

Distinct object: In order to cope with real situations, specific objects should be detected. This will also improve on localisation. Despite the work done, this is still and open problem.
There are various objects and features the robot uses to assist in localisation, their attributes must be taken into consideration; specifically, size, color, height, width, edges, contours etc. because it affects how object can be recognized. When the images are captured an appropriate algorithm is required to detect the object in the image. The robustness of the algorithm is vital. The addition of visual view planning techniques give rise to viable approach for object detection and localisation in indoor environment. The approach should have the ability to simultaneously search for the object with distinct feature.

Topological modelling and localisation: Most of the traditional localisation methods try to determine geometrically the position and the orientation of the robot. Recent approaches look for methods to build topological models once features and landmarks are detected and for the later topological estimation of the robot's state. Landmark-based methods which rely on the topology of the environment can better handle the issues because they only have to conserve topological global consistency. One of the advantage of topology modelling is that it does not suffer incremental drift just like metric approaches to determine localisation.

Though, the approach for topological modelling are less precise than metric due to discretization of the location space or computationally inflexible for autonomous robots. Therefore, it is suggested that a hybrid approach to combine topological and metric method could be an appropriate technique to compensate for the weakness of each single approach.

*Planning for world modelling***:** Important too, and still open is the problem of determining a plan for the robot in order to build a model of the environment that permits a safe motion planning and a good localisation of the robot into the operations area. In addition, the smoothness of planned robot trajectory and capability of responsive to dynamic environment must be well structured. It is important to develop a method that generate safe paths. The algorithm shouldn't only consider safe path, but collision-free paths.

Perception planning. Perception refers to the ability of autonomous system to gather information and extract important knowledge from the environment while planning means the process of making purposeful decisions in order to achieve the robot's goals. To overcome the uncertainty in robot position, some new approaches tend to determine motion plans which include the localisation requirements of the robot together with the path plan. Another important question seldom treated deals with determining what motion plan must be followed in case the robot gets lost (or the uncertainty becomes too big). A fundamental function to enable autonomous mobile robot is to provide the robot with vital information about the environment including area that are free of obstacles or to predict states of the future. In the planning goals, sensors or techniques to be implemented should consider using two or more devices to model the environment.

2.6 CONCLUSION

To determine the position and orientation of an autonomous mobile robot in an environment be it static or dynamic is a very challenging task. Accurate pose estimation is typically a requirement for robust robotic grasping and manipulation of objects placed in cluttered or dynamic environments. This chapter has been able to provide a background and identify the challenges aspect of an autonomous mobile robot. Using a single sensor to determine the pose of an object may not be reliable and accurate therefore, the use of multi-sensor is encouraged. Vision is considered as one of the best alternatives to compensate for errors acquired by other kinds of sensors such as accelerometer and gyroscope. Methods used to extract information from environment were also presented. These methods are used to detect/identify objects and match with the training image. How to represent object attitude in an environment was also highlighted in this chapter because it is more reasonable to consider the environment and what is to be achieved before system can be modelled. Strengths and weaknesses of rotation representations which can be used for rigid bodies were as well presented.

Furthermore, the chapter revealed the various sensor data fusion algorithm with the objective to multiple data sources to produce more consistent, accurate, and useful information than that provided by any individual data source. Exploring the conceptualizations and benefits, as well as existing methodologies, sensors are categorized into how they relate to one another, this is called configuration. They are cooperative, complementary and competitive. Finally, the chapter highlighted some of the research areas that can be investigated for further work.

CHAPTER 3 POSE ESTIMATION OF A MOBILE ROBOT BASED ON IMU AND VISION

3.1 INTRODUCTION

It is evident that mobile robots are currently gaining immense recognition because of their ever-increasing applications in recent years. As more attention is drawn on how mobile robot can improve our lives, it is therefore necessary to suggest appropriate mechanisms to improve challenges mobile robot are faced with [193]. Several studies have proposed different approaches on how to find an accurate mechanism to determine the pose estimation of a mobile robot by either developing new sensors to measure accurately or combine signals from several sensors to get information from different sources which can assist to estimate pose or by improving fusion algorithms [198]. The preliminaries on mobile robot have been well established and a fairly comprehensive overview has been provided in the previous chapter, as well as in references [130, 197-199]. Importantly from the literature study presented in the previous chapter, localization was identified as one of the cogent factor of mobile robot which therefore requires prompt intervention. There are certain work carried out in this regards [199]. However, there are still some challenges that are yet to be extensively addressed and one of such is measuring accurate quantities, but unfortunately there is no ideal sensor. Sensors are designed to work in specific environment with limited range as well as they are subject to noise. Due to the limitation of sensors, this chapter seek to address such aspect. For example, using only camera to solve the issue of pose might seem superfluous, the reason therefore IMU is introduced to produce more robust estimate. Any single camera system may experience issue during periods with uninformative or no vision data. This could happen due to occlusions or fast motion, an IMU will assist to bridge the gaps. Reducing computational demands for image processing. Accurate short time pose estimates are available using this information from the IMU reducing the need for fast vision updates. Furthermore, an area of concern is how to develop an accurate vision techniques that can track the location of a mobile robot with the identification of object in image. Different approaches have been suggested in the literature [194] but some of this techniques care less about the robustness and accuracy. Therefore, in this chapter an object detection algorithm is considered to improve the performance of vision to determine mobile robot pose.

It has been proposed that the use of vision and IMU can significantly help in mitigating the challenges of mobile robot [204]. For this reason, researchers are searching on applying sensor fusion. Sensor fusion is a way to estimate specific quantity by fusing measurement from multi-sensors. An ample information has been given in the previous chapter on sensor fusion algorithms. A fusion sensor algorithm is applied to give a stable and accurate signal out of noisy signals. The fusion of inertial sensors and vision has been used previously in literature. Ref. [216] give an introduction to field and its application. Sensor fusion algorithm can also be applied to minimise uncertainty and noise reduction [215, 198], helps to integrate multiple sensors and also to localise mobile robot. This chapter seeks to employ the use of fusion sensor techniques to determine the state of the mobile robot under movement given observation or measurement.

The main contributions of this chapter is summarized as follows:

- A method is put forward to achieve pose estimation for mobile robot by employing landmarks (both artificial and natural). Object detection and random sample consensus algorithms were integrated and used to recognize a sample object and features in several images taken. The proposed technique uses an iterative method to estimate the parameters of a mathematical model from a set of captured data which contains outliers.

- A fusion method which relies on information from inertial sensor and vision system becomes complementary
- Simple implementation and low cost technologies to obtain accurate localisation. The method adopted was evaluated and validated on a user-friendly software `Arduino environment` and `MATLAB`.

3.2 MODELING METHOD

When working with sensor unit containing a camera and an inertial measurement unit (IMU) several reference coordinate systems have to be presented and this is given as Figure 3.1:

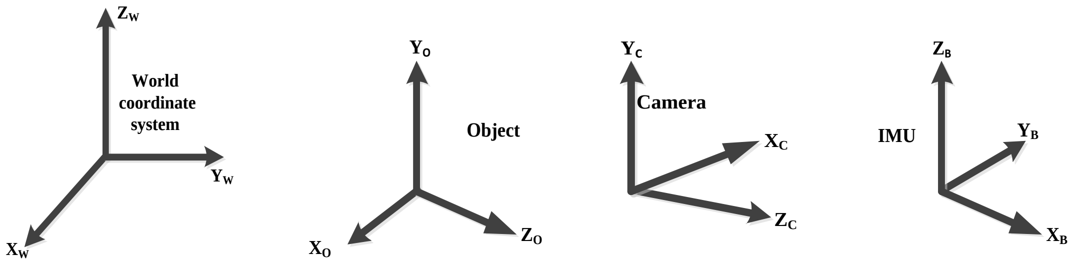


Figure 3.1. Reference coordinate system [205].

Reference frame for the system

Global frame/world frame {w}: This frame aids the user to navigate and determine the pose estimation in relative to IMU and camera frames.

IMU/body frame {b}: This frame is attached to the IMU (accelerometer and gyroscope) on the mobile robot.

Object coordinate frame {o}: This frame is attached to the object; in this study it is a 4WD mobile robot.

Camera frame {c}: This frame is attached to the camera on the mobile robot with the x-axis pointing to the image plane in the right direction and z-axis pointing along the optical axis and origin located at the camera optical center.

The IMU method provides orientation of the body *{b}* with respect to (wrt) world frame *{w}* R_{wb} and vision method provides orientation of the object*{o}* wrt to camera frame *{c}* R_{co} [205].

3.3 POSE ESTIMTAION

3.3.1 IMU-based pose estimation

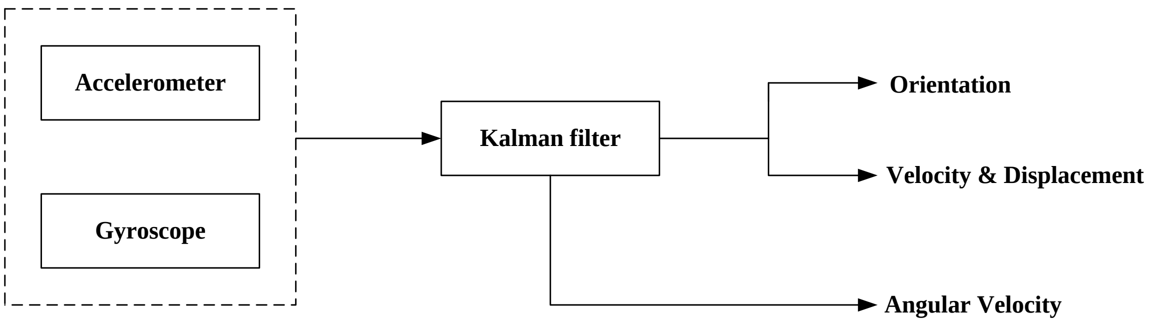


Figure 3.2. Block diagram for IMU [205].

Figure 3.2 above shows the block diagram of the inertial sensors after calibration, the data have been passed through the Kalman filter (KF) to reduce drifts and errors [206]. KF is a set of mathematical equations that provides efficient computation means to estimates the state of a process. It is a recursive filter that is based on Bayesian approach and is used for state prediction and update of the system. This filter is also capable of estimating accurate orientation of the system, but basically used for linear system. For KF, theoretically it is an ideal filter for combining noisy sensors to acquire accurate and estimated output. It is accurate because it takes known physical properties of the system into account. However, it is mathematically complex to compute, and code as compared to complementary filters. The calibrated accelerometer and gyroscope were used to determine orientation, angular velocity, linear velocity and displacement of the mobile robot with the use of KF. The KF was used as a prediction and correction model for the sensors.

To express an object or mobile robot orientation, several representations are proposed to be used. Examples are: axis angle, Euler angles, Direct Cosine Matrix (DCM) and quaternions [14, 132] as earlier discussed in Chapter 2. For this aspect of study, Euler angle was adopted to solve for roll, pitch and yaw angles because of its simplicity. The gravity in the world frame can be obtained using coordinate information from the body frame.

$$g_w = R_{wb} g_b \tag{3.1}$$

where g denotes the gravity and the subscripts b and w represents the body frame and world frame, respectively. To obtain the rotation matrix from the world frame *{w}* to the body frame *{b}*, (R_{wb}), the Euler angles, roll φ, pitch θ, and yaw ψ can be obtain as:

$$(R_{wb}) = \begin{bmatrix} c\theta c\psi & c\theta s\psi & -s\theta \\ s\varphi s\theta c\psi - c\varphi s\psi & s\varphi s\theta s\psi + c\varphi c\psi & s\varphi c\theta \\ c\varphi s\theta c\psi + s\varphi s\psi & c\varphi s\theta s\psi - s\varphi c\psi & c\varphi c\theta \end{bmatrix} \tag{3.2}$$

where *c* is defined as *cos ()*, and s is defined as *sin ()*.
The world frame provides the reference frame for the body frame, in which the x-axis and y-axis are tangential to the ground and the z-axis is in the downward direction (view direction). The initial gravity vector in the world frame is given as:

$$g_w = \begin{bmatrix} 0 \\ 0 \\ g \end{bmatrix} \tag{3.3}$$

The 3-axis accelerometer gives the components of the gravitational accelerations expressed in the object reference frame ($(g_b = [g_{bx} g_{by} g_{bz}]^T)$. Where the superscript T represents the transpose matrix. Hence, substituting the gravity vector is related through a rotation matrix, the relation is given as:

$$g_b = \begin{bmatrix} g_{bx} \\ g_{by} \\ g_{bz} \end{bmatrix} = R_{wb} g_w = R_{wb} \begin{bmatrix} 0 \\ 0 \\ g \end{bmatrix} = \begin{bmatrix} -g\sin\theta \\ g\cos\varphi\cos\phi \\ g\cos\varphi\cos\theta \end{bmatrix} \tag{3.4}$$

Equation (3.4), roll and pitch angles can be deduced from the gravity vectors as:

$$\theta = \arctan\left(\frac{g_{by}}{\sqrt{(g_{bx}{}^2 + g_{bz}{}^2)}}\right) \tag{3.5}$$

$$\phi = \arctan\left(-\frac{g_{bx}}{g_{bz}}\right) \tag{3.6}$$

Equations to calculate the position and velocity are given as:

$$V_{b(k+1)} = V_{bk} + a_{bk}\Delta t \tag{3.7}$$

$$S_{b(k+1)} = S_{bk} + V_{bk}\Delta t S_w = R_{wb} S_b \tag{3.8}$$

where a_b, V_b, S_b, k, $k+1$ and Δt are acceleration, velocity, position, time intervals and sampling time. The angular rate is integrated to determine the orientation from gyroscope [207].

3.3.2 Vision-based pose estimation

The 3D vision-based tracking approach tracks the pose of the mobile robot with camera in relative to the referenced object. For effective tracking, fast and reliable feature vision algorithm is vital. The process of vision localisation is categorised into four major steps: acquire images via camera, detect object in the current images, match the object recognised with those contained in the database and finally, calculate the pose as a function of the recognised object. A forward looking single camera (monocular) was used because it provides a high number of markers thus allowing good motion estimation accuracy, if the objects are closers to the camera [23, 188].

Projection of object reference points to image plane

With monocular vision (one camera), a good solution in terms of scalability and accuracy is provided [190]. The monocular vision demands less calculation than stereo vision (two cameras). With the aid of other sensors such as ultrasonic sensor or barometric altimeter, the monocular vision can also provide the scale and depth information of the image frames [108]. Vision method provides orientation of the object {o} with respect to (wrt) to camera coordinate frame {c}, $R_{co.}$ using the pinhole camera model [107] to calculate the pose of the mobile robot with respect to the camera [190]. The monocular vision positioning system was used to estimate the 3D camera from 2D image plane. The relationship between a point in world frame and its projection in the image plane can be expressed as:

$$\lambda p = MP \tag{3.9}$$

where λ is a scale factor, $p = [u, v, 1]^T$ and $P = [X_w, Y_w, Z_w, 1]^T$ are homogenous coordinates of p and P, and M is a 3×4 projection matrix. The above equation can further be expressed as:

$$\lambda \begin{bmatrix} u \\ v \\ i \end{bmatrix} = M(R_{wc} t_{wc}) \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \tag{3.10}$$

The projection matrix depends on both camera intrinsic and extrinsic parameters. The intrinsic parameters contain five parameters: focal length f, principal point u_0, v_0 and the skew coefficient between x and y axis and is often zero.

$$M = \begin{bmatrix} a_x & \gamma & u_0 \\ 0 & a_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} \tag{3.11}$$

Extrinsic parameters: R, T, defines the position of camera center and the camera's heading in world coordinates. Camera calibration is to obtain the intrinsic and extrinsic parameters. Therefore, the projection matrix of a world point in the image is expressed as:

$$C = -R^{-1}T = -R^T T \tag{3.12}$$

where T is the position of the origin of the world coordinate, and R is the rotation matrix. For this research, camera calibration used in this project is based on Zhang's calibration technique [101].

3.3.3 Fusion of IMU and vision

The objective of sensor fusion is to improve the performance acquired by each sensor taken individually and integrating their information. The use of vision alone fails to handle occlusion, fast motion and not all areas are covered due to the field of view of the camera. Therefore, with the shortcoming of each sensor, it is mandatory to combine their data sources to provide a better pose estimation result. Velocity, position, angular velocity and orientation are given by IMU and so also is the position and orientation given by vision. The fusion of vision and IMU is carried out using EKF. The fused EKF computed the overall pose of the mobile robot with respect to the world {w} frame. Figure 3.3 shows the overview stages of IMU and Vision fusion adopted. In the figure, the IMU which is 6-DoF is comprised of the gyroscope and accelerometer to determine the position and orientation of the robot. The IMU was first calibrated before mounted on the mobile robot. The orientation was measured using gyroscope and the accelerometer gives the vector parameters for the robot in terms of positions. The process model was use to model inertial sensor error for gyroscope and this is fused with the vector measured by the accelerometer using Kalman filter. This filter is also capable of estimating accurate orientation, angular velocity and displacement of the system. For the vision, camera (monocular vision) that was mounted on the mobile robot was used to capture several images of the experiment scene (environment). Thereafter, the images were saved on a computer to analyze and perform necessary operation to determine the vision parameters. SURF algorithm was adopted to detect interest key points and generate its descriptors on the images using Hessian matrix because of its high level of accuracy. With the algorithm, strongest feature points were extracted from the query image to match with the training image in order to have sufficient points when matching the images. RANSAC algorithm as the feature matcher which works well with SURF to match object detected by SURF algorithm in images. RANSAC algorithm was used to determine the pose estimation of the camera in relative to the object. This was use to estimate the position and orientation of the object identified in the image in respect to the mobile robot location. To get optimal result, EKF was used to fuse both IMU and vision data to give the desire pose estimation.

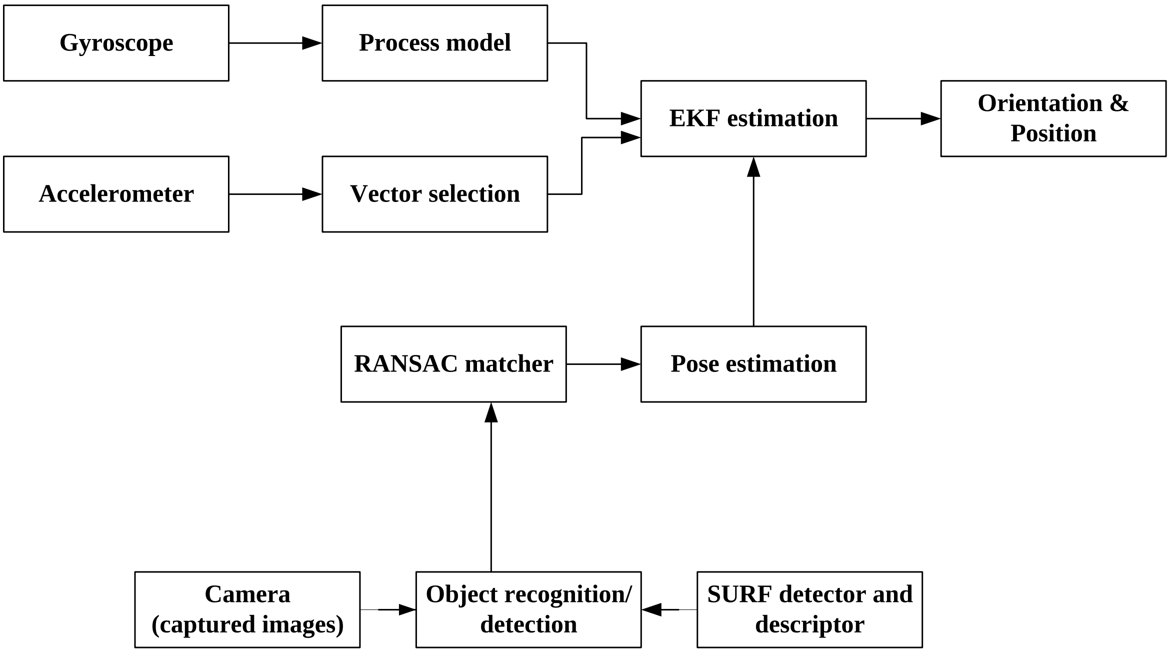


Figure 3.3. Overview of the stages in fusion of IMU and vision [205].

3.3.4 Homography estimation based on natural landmark

This section introduces a method for extracting natural landmark to estimate the position of mobile robot. The goal of homography estimation is to find an appropriate global transformation of images of the same scene taken at different viewpoints. Homography estimation can be classified into two categories: pixel-based approaches and feature-based approaches [212]. This study is based on features-based approach. The choice of features is vital because it will determine the complexity in the description, detection and matching. With the remarkable development of keypoint features such as SURF, feature-based approaches have gained great popularity in homography estimation. Using homography to estimate orientation of a camera relative to the planar surface of the image, from consecutives images, a point from the first image and second image is selected to represent the same object. A linear system is derived from homograph matrix that denotes a projective transformation between the first image and second image is provided and components of the normal vector in the camera coordinate system are estimated by solving the linear equation system and determining the orientation of the camera relative to the surface. The orientation of the camera to the image plane can be described by a homography which represents a projective transformation between two different images of the same plane captured from two different positions of the camera. The determination of the orientation of the camera relative to the image plane may comprise determining a pitch and roll angles of the camera relative to the surface. The pitch and roll angles are filtered by KF because the extraction angle information from homography matrix could comprise parameter ambiguities. Rotation and translation of the camera between capturing the first image and second image are based on RANSAC algorithm. The homography between the two images captured sequentially from the mobile robot includes information about camera's roll and pitch angles relative to image plane [213]. For the camera orthogonal coordinate x, y and z , the normal vector of the image is given as $n=\left[n_x, n_y, n_z\right]$.

The $Y-Z$ plane of the coordinate system for the pitch angle is defined as the angle between the projection of the normal vector n onto the $Y-Z$ plane and y-axis and the $X-Y$ plane represents roll angle. Therefore, the pitch and roll angle are defined by:

$$\theta = -\arctan\frac{n_z}{n_y} \tag{3.13}$$

$$\varphi = -\arctan\frac{n_x}{n_y} \tag{3.14}$$

Information regarding the pitch angle and roll angle are incorporated in a homography which characterizes a projective transformation between two images. The pose from extracted features on the image was estimated by homography matrix. Homography matrix was calculated using (3.15) [194].

$$H = K(R-\frac{tn^T}{d})K^{-1} \tag{3.15}$$

where H is the homography matrix, K is the intrinsic parameter d the distance between the camera based origin and the closet point on the image plane, n is normal vector, R is the rotation matrix, t is the translation of between two camera positions. The normal vector n is to estimate the pitch and roll angles and is conditioned by $\|n\|=1$.

The position of the points in the images is represented by homogenous vector $p=[p_1, p_2, p_3]^T$ for the first image and $q=[q_1, q_2, q_3]^T$ for the second image. Therefore, the homograph matrix H, for homogenous vector can be calculated as $q = Hp$ and the image coordinates (u, v) can be written as: $u = \frac{q_1}{q_3}$ and $v = \frac{q_2}{q_3}$.

Additional information is provided by the IMU and these sources are combined with extended Kalman filter. They are fused together to determine the position and orientation.

3.3.5 EKF Implementation

For sensor fusion, EKF was implemented to estimate position and orientation from IMU and vision. EKF is a classic approach for a nonlinear stochastic system; it uses discrete models with first-order approximation for nonlinear systems. The EKF algorithm enables complementary compensation for each sensor's limitations, and the resulting performance of the sensor system is better than individual sensors [82, 83, 214]. The motion model and the observation model in EKF are established using kinematics. EKF gives reasonable performance mostly in conjunction with a long iterative tuning process. More details are given for Kalman filter [20, 56], particle filter [57], unscented Kalman filter [58, 195, 196] for implementations and demonstrations of EKF.
The general EKF equations are given as:

$$x_{k+1} = f_k\left(\hat{x}_k, \mu_k, w_k\right) \quad w_k \sim N\left(0, Q_k\right) \tag{3.16}$$

$$y_k = h_k\left(x_k, v_k\right) \quad v_k \sim N\left(0, R_k\right) \tag{3.17}$$

where x_k is the state vector, u_k denotes a known control input, w_k denote the process noise, and v_k is the measurement noise. y_k is the measurement vector, h_k is the observation matrix all at time *k*. The process noise w_k has a covariance matrix *Q* and measurement noise v_k has a covariance matrix *R*, are assumed to be zero-mean white Gaussian noise processes independent of each other. EKF is a special case of Kalman filter that is used for nonlinear system. EKF is used to estimate the robot position and orientation by employing the prediction and correction of a nonlinear system model.

Time prediction update equation is given as:

$$\hat{x}^- = A\hat{x}_{k-1} + Bu_k \tag{3.18}$$

$$P_k^- = AP_{k-1}A^T + Q_{k-1} \tag{3.19}$$

where A is the transition matrix and B is the control matrix.

Measurement update equation is given as:

$$\hat{x}^+ = \hat{x}^- + K_k(z_k - h(\hat{x}_k^-)) \tag{3.20}$$

$$P_k^+ = (I - K_k H_k)P_k^- \tag{3.21}$$

where the Kalman gain is given as:

$$K_k = P_k^- H^T{}_k (H_k P_k^- H_k + R_k)^{-1} \tag{3.22}$$

The Jacobian matrix H_k with partial derivatives of the measurement function $h(\cdot)$ with respect to the state x is evaluated at the prior state estimate $\hat{x}_k^-$, the equation is given as:

$$H = \frac{\partial h}{\partial X} \mid X = x_{k-1,}\ v = 0 \tag{3.23}$$

For the fused filter method used in this study we adopted one of the model used in [83]. We used accelerometer data as a control input, gyroscope data and vision data were used as measurements. This model is extensively explained in reference above, but the process noise and covariance noise are suitably tuned. The state vector is given as:

$$x = [p \quad v \quad q \quad \omega]^T \tag{3.24}$$

where p and v stand for the state variables corresponding to the 3D position and velocity of the IMU in the world frame, q denotes the orientation quaternion corresponding to the rotation matrix R and ω is the angular velocity from gyroscope. The fused transition matrix used here is given as:

$$F_1^1 = \begin{bmatrix} F_1 & O_{6x7} \\ O_{7x6} & F^1 \end{bmatrix} \tag{3.25}$$

The state transition matrix can be written as:

$$F_1 = \begin{bmatrix} 1 & 0 & 0 & \frac{\Delta t}{2} & 0 & 0 \\ 0 & 1 & 0 & 0 & \frac{\Delta t}{2} & 0 \\ 0 & 0 & 1 & 0 & 0 & \frac{\Delta t}{2} \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{3.26}$$

$$F^1 = \begin{bmatrix} 1 & 0 & 0 & \frac{\Delta t}{2} & 0 & 0 & \frac{\Delta t^2}{2} \\ 0 & 1 & 0 & 0 & \frac{\Delta t}{2} & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & \frac{\Delta t}{2} & \cos^2 t \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & \sin^2 t \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \tag{3.27}$$

where Δt is the sampling time between images captured. The process noise covariance is taken from the acceleration and is given as:

$$Q_1^1 = \begin{bmatrix} Q_1 & O_{6x3} \\ O_{7x3} & Q^1 \end{bmatrix} \tag{3.28}$$

$$Q_1 = \begin{bmatrix} q_1 & O_{3x3} \\ O_{3x3} & q^1 \end{bmatrix} \tag{3.29}$$

$$Q^1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & \Delta t^2 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & \Delta t \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{3.30}$$

where $q^1 = I_3 \sigma_a{}^2$ and $q_1 = I \sigma_a{}^2$, are the process noise taken from accelerometer while the measurement noise is taken from the gyroscope and vision. I_n is the identity matrix dimension of n. R is the key matrix for sensor fusion, R_1 and R^1 are the covariance from gyroscope and vision.

$$R_1^1 = \begin{bmatrix} R_1 & O_{4x3} \\ O_{3x4} & R^1 \end{bmatrix}, \ R_1 = I_4 \sigma_g{}^2, \ R^1 = I_3 \sigma_v{}^2 T \tag{3.31}$$

Observation matrix is given as:

$H_1^1 = \begin{bmatrix} H_1 \\ H_1 \end{bmatrix}$, $H = [O_{3x3} \ I_{3x3}]$, is the observation matrix from gyroscope,

$H^1 = \begin{bmatrix} \frac{1}{2} t^2 R^T & R \end{bmatrix}^T$, observation matrix from vision. The parameters used for filter tuning and experiments are given in Table 3.1. Most of the IMUs have acceleration and gyroscope range of ±16 g and ±2000 deg/s, respectively, and sampling rates up to a few hundred Hz. Sufficient information for data analysis is presented by using high sampling rate, but the system may be burden due to the large data size and computation load. Conversely, applying low sampling rate may fail to capture inherent attributes of each activity. Therefore, a sampling rate around 100 Hz is adequate for capturing daily life human activities [201]. The requirement of sampling rate also depends on the type of analysis. For example, using a sampling rate higher than necessary will increase the computation load without improving the results. In a distinctive system, the accelerometer and gyroscope in the IMU run at relatively high sample rates. The complexity of processing data from those sensors in the fusion algorithm is relatively low. Conversely, the camera sensor runs at a relatively low sample rate and the complexity associated with processing is high. In this fusion algorithm the camera samples between frames are processed at a low rate, and the accelerometer and gyroscope samples are processed together at the same high rate. A sample rate of 100 Hz up to 20,000 Hz is allowed. For IMU the sample rate is selectable in intervals of 1 Hz to 200 Hz. To simulate this configuration, the IMU (accelerometer and gyroscope) is sampled at 100 Hz (16 bits/sample), and the camera is sampled at 25 Hz. For the noise for camera, we chose 0.9 to give a more stable filter in our experiment. The covariance measurement noise is referred to the variance of the measurement. If it is too high the filter will react slowly as the new measurement has more uncertainty, on the other hands if the measurement noise is too small, the output result becomes noisy. Therefore, measurement noise has significant role to increase estimation error in the data fusion process. Allan variance can be used to determine noise parameters of a MEMS gyroscope. The gyroscope measurement noise is assumed Gaussian with zero mean. The Gaussian distribution is more reasonable from several viewpoints because it is characterized by mean and standard deviation or variance. Refer to [202-203] on how to estimate noise parameters using power spectral density and Allan variance method.

Table 3.1. Parameters and their values for filter tuning.

Variables	Meanings
Sampling interval of IMU sensor	100 Hz
Gyroscope measurement noise variance, σ_g	0.001 rad2/s2
Accelerometer measurement noise variance, σ_a	0.001 m/s2
Camera measurement noise variance, σ_v	0.9
Sampling interval between image frames	25 Hz

3.4 SYSTEM HARDWARE AND EXPERIMENTAL SET-UP

Figure 3.4 shows the major hardware used to carry out the experiment. Besides other types of components such as IR sensors, ultrasonic sensor etc which aided robot navigation and validated the proposed method. The mobile robot used in this experiment is a four-wheel drive (4WD) as shown in Figure 3.4c with a working voltage of 4.8 V. Four servo motor controllers were used which allowed the robot to move up to 40 cm/s (0.4 m/s) with microcontroller (Arduino/Genuino 101) which has built-in of Inertial Measurement Unit of 3-axis accelerometer and 3-axis gyroscope, depicted in Figure 3.4a. To reduce the payload, the frame of the robot was built with aluminium alloy. The robot was equipped with a 6 V battery to power the servo motors and a 9 V battery for the microcontroller. The mobile robot is also installed with ultrasonic sensor to measure the object distance to the mobile robot in real time. The performance issues related to reflections, occlusions, and maximum emitting angles limit independent use of ultrasonic sensors [63]. The camera was also mounted on the mobile robot to take several images from the environments. The type of camera used for this experiment is the LS-Y201-2MP LinkSprite's new generation high-resolution serial port camera module. The pictorial representation is given in Figure 3.4b. Its resolution is 2 million pixels which can capture high resolution images using the serial port. The camera is a modular design that outputs JPEG images through universal asynchronous receiver transmitter (UART) and can be easily integrated into existing design. It has a default baud rate of serial port of 115,200. More of its specification can be found in [72]. The camera was connected to the programmed microcontroller Arduino 101 mounted on the robot to capture images with a resolution of 1600 × 1200 at 6 frame per second (fps). Images captured with the programme written on Arduino environment are stored in an SD card and corresponding IMU transmitted to the PC, via the USB cord which processes the images and locates the references points in the captured images. The marker (box) used as a reference object has a size of 15 × 24 cm and was placed at a known position. The object was used to calculate the pose estimation of the mobile robot relative to the camera. The image processing and pose estimation process were analysed offline using MATLAB software. The mobile robot trajectory is designed in such a way that it moves on a flat terrain in a forward, left and right directions. The work area for the experiment is 4 m × 5.2 m.

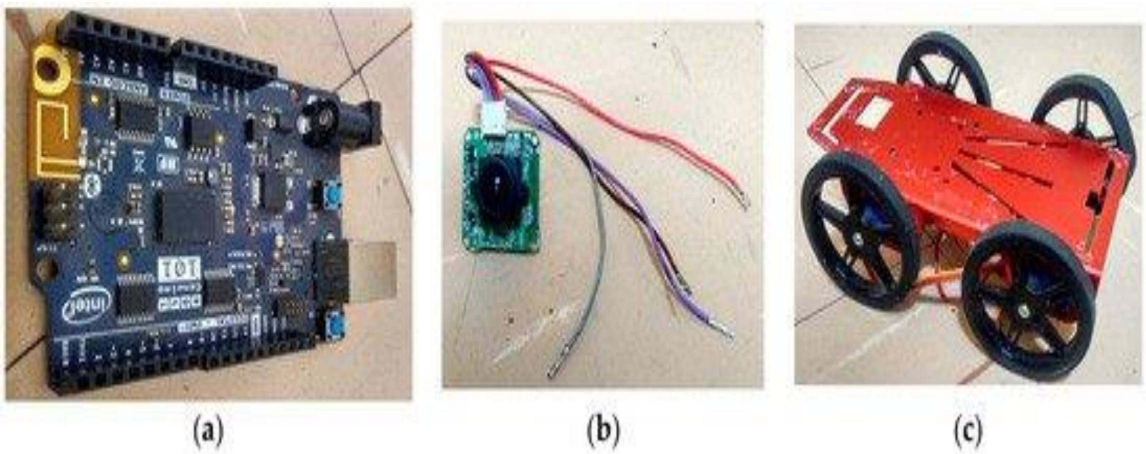


Figure 3.4. Hardware used for the experiment: (**a**) Arduino 101 microcontroller. (**b**) LinkSprite camera. (**c**) 4WD robot platform [205].

3.5 RESULTS AND DISCUSSION

In this section, the performance of the experiments and simulated results are evaluated and analysed. Firstly, the analysis of the images captured in MATLAB was presented. Secondly, the results of the experiments performed to determine the position and orientation of the mobile robot by fusing the inertial sensor and vision data was also presented afterwards.

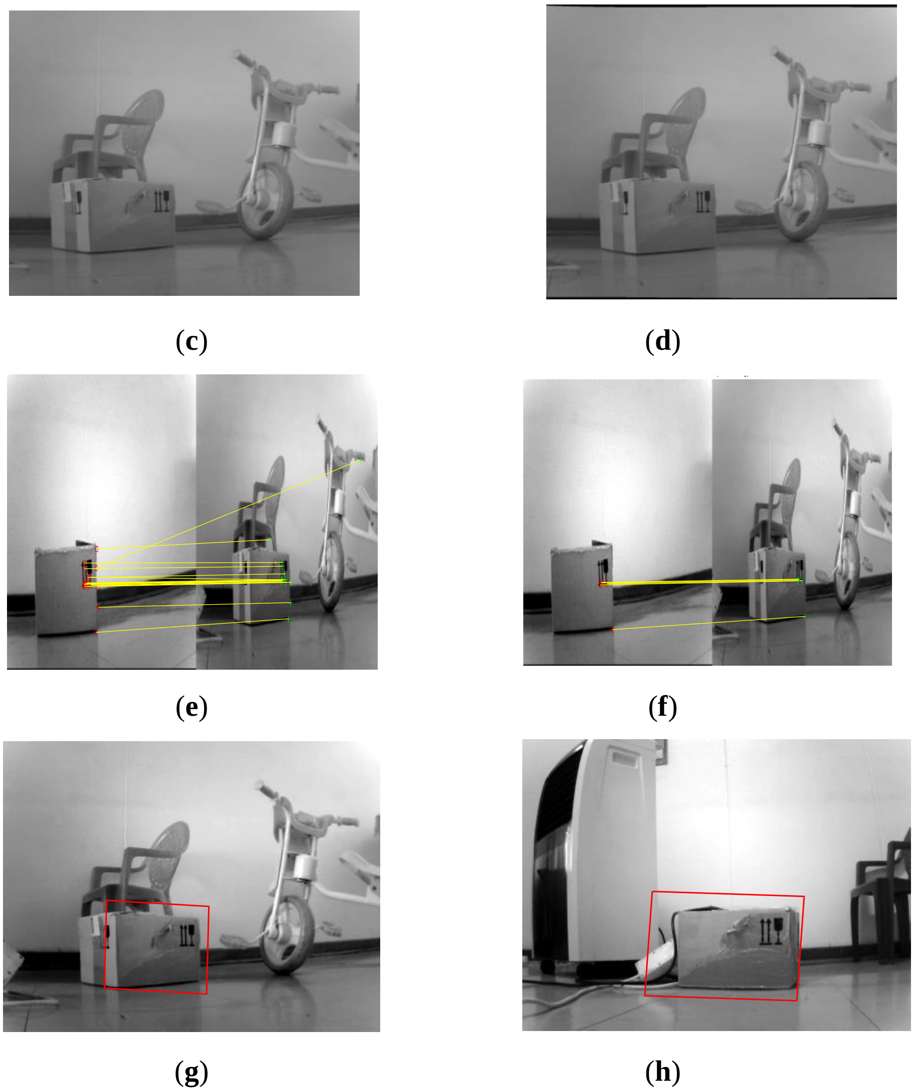
3.5.1 Simulated result of objects detection and recognition in an image

In this subsection, details of the vision techniques used for detection and recognition in an image was presented; this was implemented in MATLAB using the computer vision toolboxes following the steps given in Chapter 2, Section 2.3.5 and with the brief introduction given of how images were captured and stored on an SD card and transferred to MATLAB for simulation.

Figure 3.5 shows the detection of an object box placed in a known position to estimate the position of the mobile robot when moving in the confined area. The first step was to save the proposed object (which could also be called the query image); in this case a box was used, the image was saved in a database file. Thereafter the image was converted from RGB to grayscale after resizing the image so that it would not be too large to fit on a screen.

(a)

(b)



(c) (d) (e) (f) (g) (h)

Figure 3.5. A box detected from two different images but in similar scenes [205]. (**a**) Query image. (**b**) Training image. (**c**) Conversion of RGB to grayscale. (**d**) Removal of lens distortion. (**e**) Image including outliers. (**f**) Image with inliers only. (**g**,**h**) Images with display box around the recognised object.

The purpose of converting from RGB to grayscale is to acquire better results. Examples of such images are depicted in Figure 3.5b and c respectively. Figure 3.5b shows the RGB image, while Figure 3.5c shows the grayscale image. Some camera lenses are distorted, and therefore it is important that lens distortions are removed from images. The purpose of removing distortion in images is to correct any form of abnormalities and variations in the images to give a good quality output. Figure 3.5d shows an image in which distortion has been removed.

To detect features from images using SURF, Figure 3.5e shows a typical example of the outliers and inliers. For the simulation, 50 of the strongest feature points were extracted from the query image to match with the training image in order to have sufficient points when matching the images. The matching of images was done by RANSAC algorithm. With RANSAC algorithm, the inliers were computed in such that if the inliers points are more than the threshold then homograph transform will be estimated. This is shown in Figure 3.5f. The last step is for a bounding box to be designated and displayed around the recognised object as shown in Figure 3.5g, h.

3.5.2 Simulated and experimental results of object positioning

Figure 3.6 shows the experimental result of the Euler angles obtained from IMU and the filtered estimate. Various methods have been suggested to calculate Euler angles. Some methods considered using only data from a gyroscope to estimate Euler angles by integrating angular velocity to give orientation, while another uses only accelerometer data. Because gyroscope measures rotation and accelerometer does not, gyroscope seems to be the best option to determine orientation. However, both sensors have their limitations, and therefore it is suggested that the weakness of one sensor could be complemented by the other. For this study accelerometer and gyroscope data was integrated using Kalman filter. Figure 3.2 shows the block diagram of the stages. Equations (3.5) and (3.6) were used to calculate the pitch and roll angles, while the yaw angle was calculated as an integration of angular velocity. From Figure 3.6, it can be noted that, for about 49 s, roll and pitch angles maintained a close-to-zero angle until there was a change in direction. At the point where the robot turned 90 degrees to the right, the yaw angle was 91.25 degrees. The maximum values obtained for pitch and roll angles are 15 degrees and 18 degrees, respectively. From the experiment carried out on IMU, it can be concluded that Euler angles are a good choice for the experiment performed because the pitch angles did not attain ±90 degrees to cause what is known as Gimbal lock.

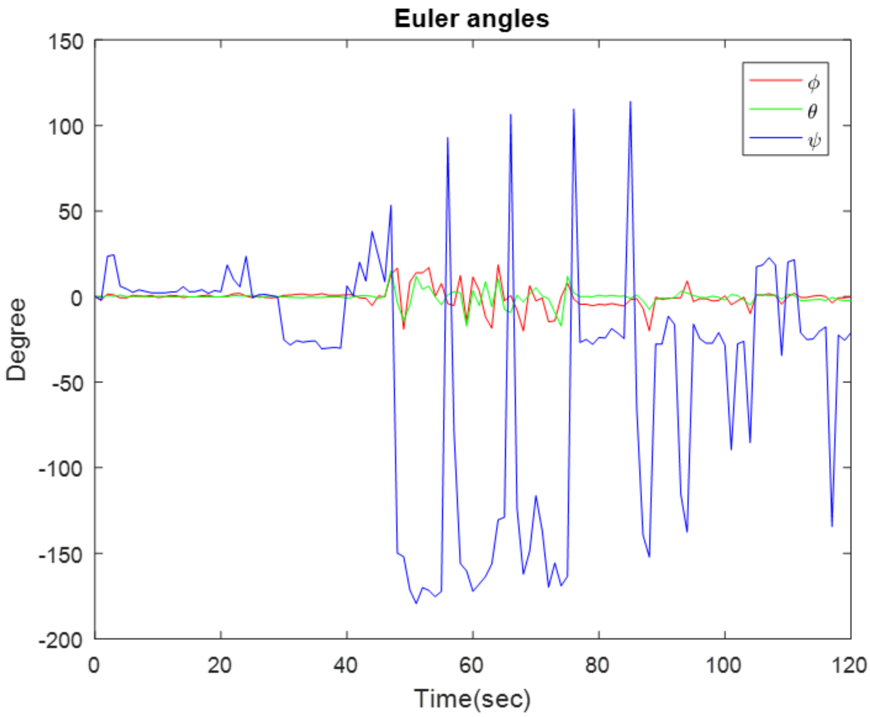
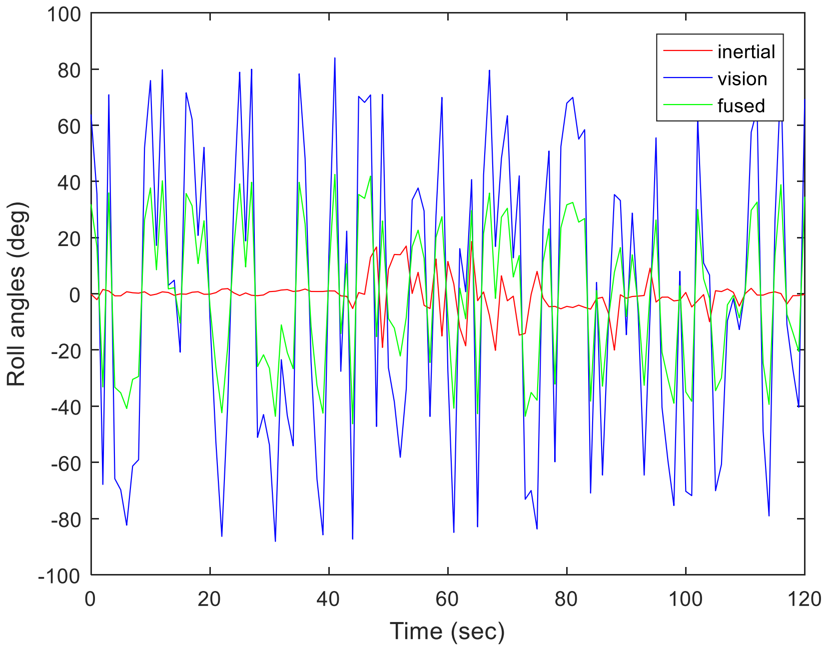


Figure 3.6. Euler angles from IMU. Roll: red; pitch: green; yaw: blue [205].

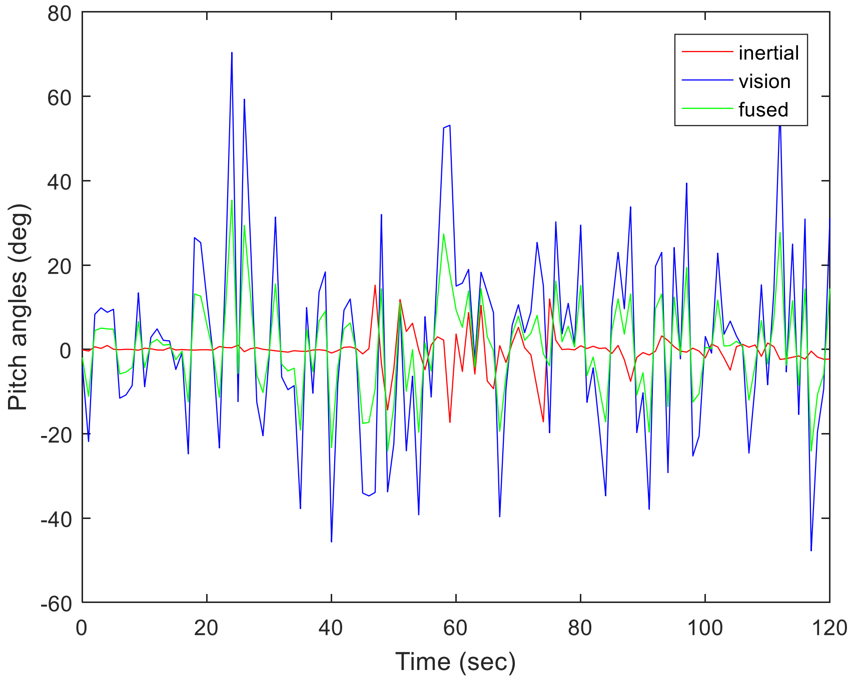
Figure 3.7a-c shows the orientation result of the fused data from inertial sensor and vision. The IMU was able to abruptly determine the direction of mobile, but the vision slowly captured the images to determine the orientation of the mobile robot. With different sampling frequencies, computation time did not allow both estimates to run at the same time. The IMU was able to determine the direction of the robot within a specific path, but with the camera, the rotational axis was extended to capture more views; therefore, the range of direction was widened and areas which could not be covered by IMU were captured by the camera, although vision-based tracking is more accurate for slow movement than IMU. However, using only computer vision, tracking is lost almost immediately; it is therefore obvious that the addition of IMU is beneficial. EKF was used to fuse the inertial and visual measurement to estimate the state of the mobile robot. With EKF, corrections for pose estimation were made; this shows that the filter is efficient, specifically when fusing two or more sensors together. Equations (3.10) -(3.12) were used to calculate the camera pose in reference to the image plane. From the equations, the intrinsic and extrinsic parameters were estimated through the camera calibration. It should be noted that the described system is very sensitive to calibration parameters. Errors in parameters used for calibration could deteriorate the tracking of the system. Hence, the design of accurate calibration methods is vital for proper operation. As observed from the figures, there is a slight difference between the data obtained from inertial sensor to that of vision. At the point where the robot made a 90 degrees right turn, the yaw value for IMU was 91.25 degrees, and 88 degrees for vision. Pitch and roll angles both have values of 1.2 degrees and 4 degrees. With the use of EKF data fusion, the proposed method was able to reduce accumulated errors and drifts, improvement was thereby achieved.

A comparison of the three directions of the mobile robot taken from vision only was presented as Figure 3.8. The figure shows a distinctive estimation of position of the mobile robot. The position estimation based on the reference object in the image is relative to the position of the mobile robot and the world coordinate, with the median vector of the planar object for Z-axis close to 1 and −1. This shows that the feature selection method used is effective. Therefore, SURF and RANSAC algorithms combination can be used to determine the accurate position of an object through vision.



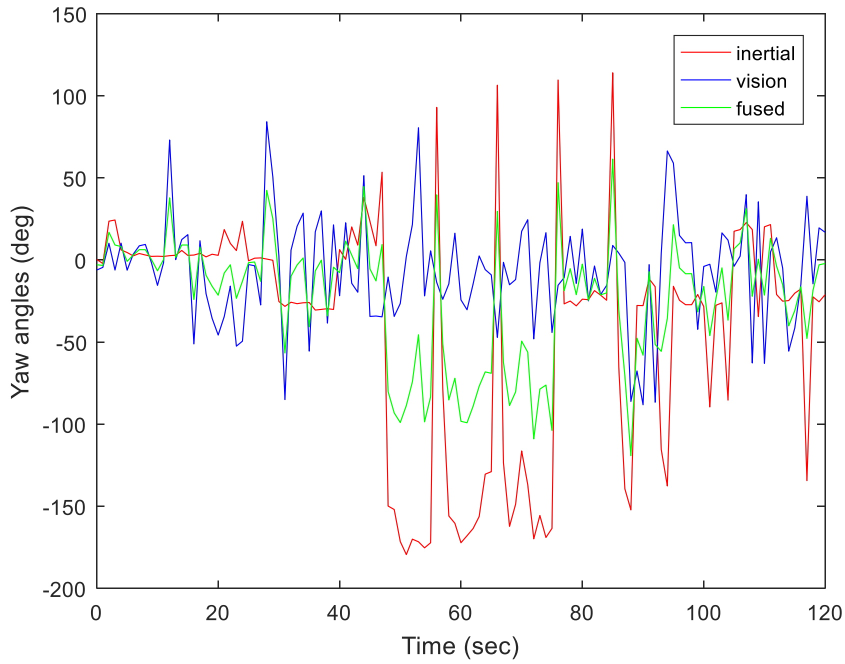
(a)

Figure 3.7a. Orientation results for fused sensors roll angles [205].



(b)

Figure 3.7b. Orientation results for fused sensors pitch angles [205].



(c)

Figure 3.7c. Orientation results for fused sensors yaw angles

Figure 3.7. Orientation results for fused sensors. Roll, pitch and yaw angles [205].

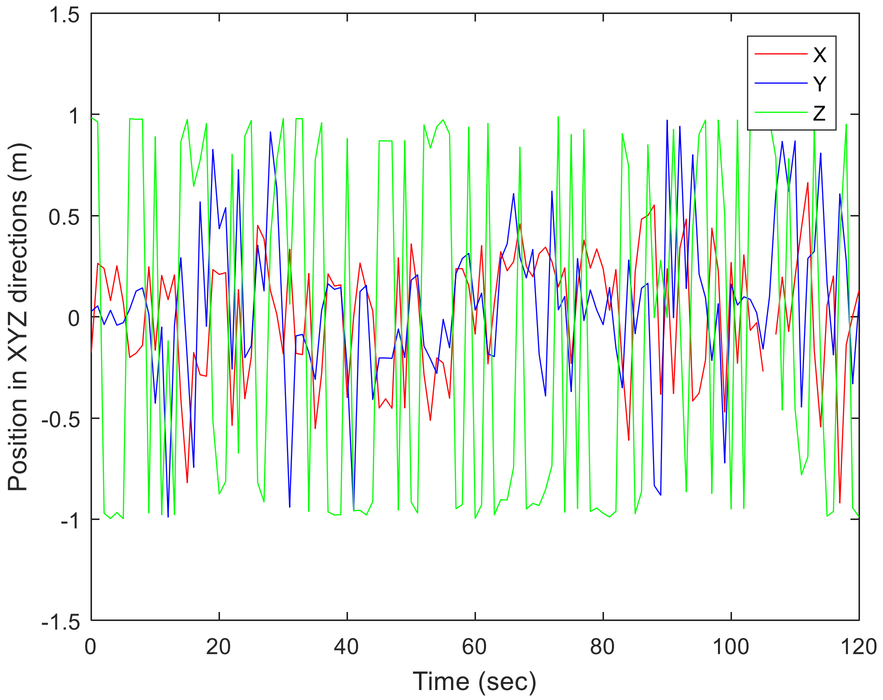


Figure 3.8. Experimental position in XYZ directions from vision data [205].

3.5.3 Performance: Accuracy

We estimated the performance of our algorithm using an external camera placed in the environment of experiment to capture data used as ground truth. The purpose of use of the camera was because of its availability and reliability to determine 6-DoF of position and orientation. The camera was positioned on a flat ground along with the mobile robot at a distance of 4.90 m between; the scenario is denoted as Figure 3.9. The camera adopted for the experiment wasn't a motion camera, but it was ensured that the experimental area was covered.

Figure 3.10a shows the trajectory of the mobile robot projected in the XY plane and Figure 3.10b shows the corresponding positions of the mobile robot trajectory. From the same figure, it can be observed that our proposed method was close to the ground truth which therefore exhibits good performance. Further improvement on ground truth data collected could suggest using a motion capture camera or a laser ranging sensor. Though expensive, but an accurate result is guaranteed.

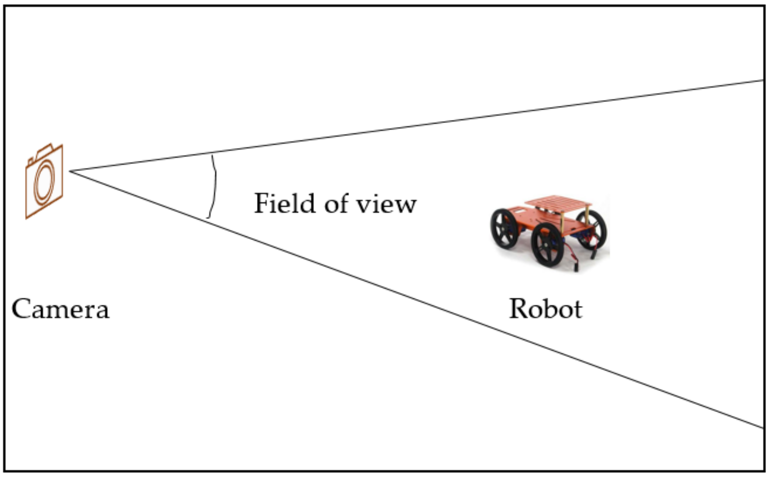
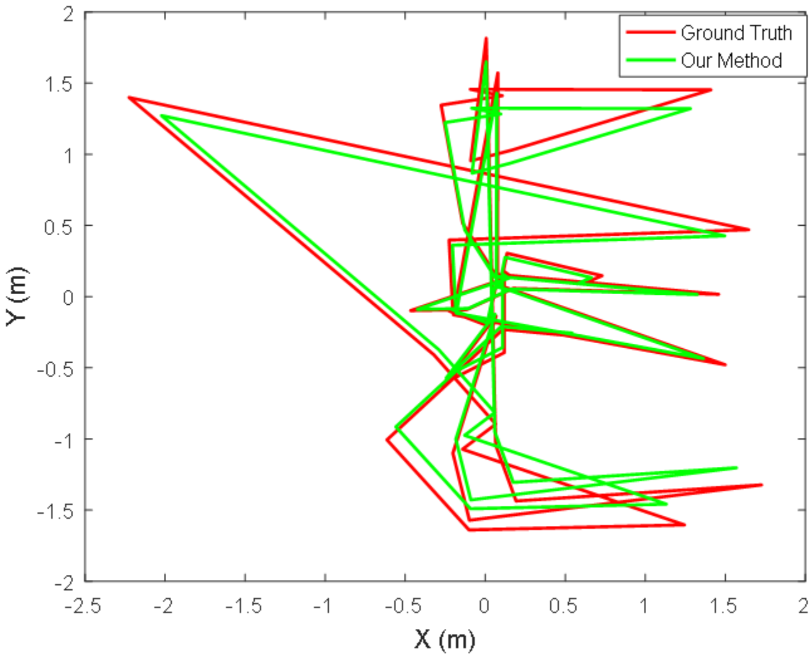
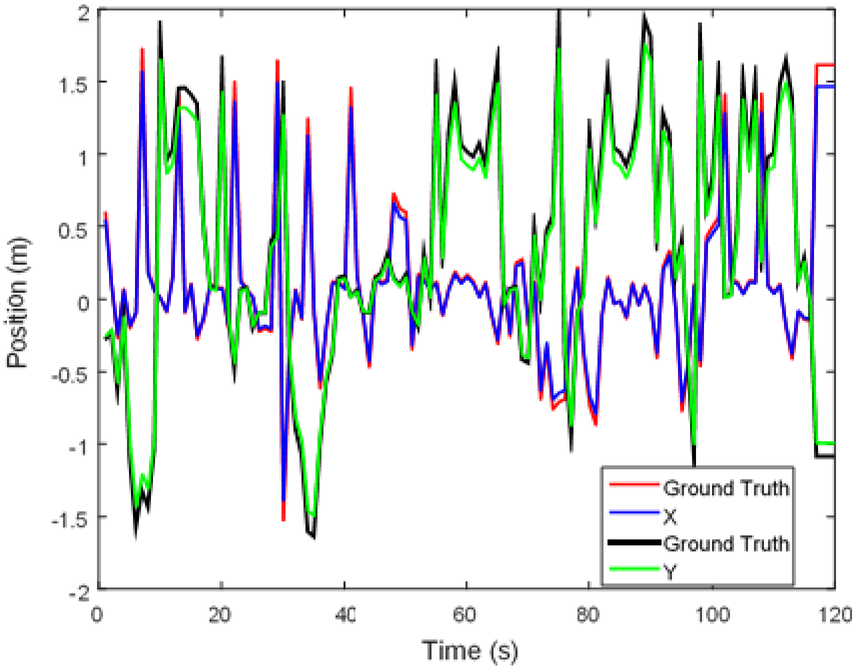


Figure 3.9. Ground truth system based on a camera [205].



(a)



(b)

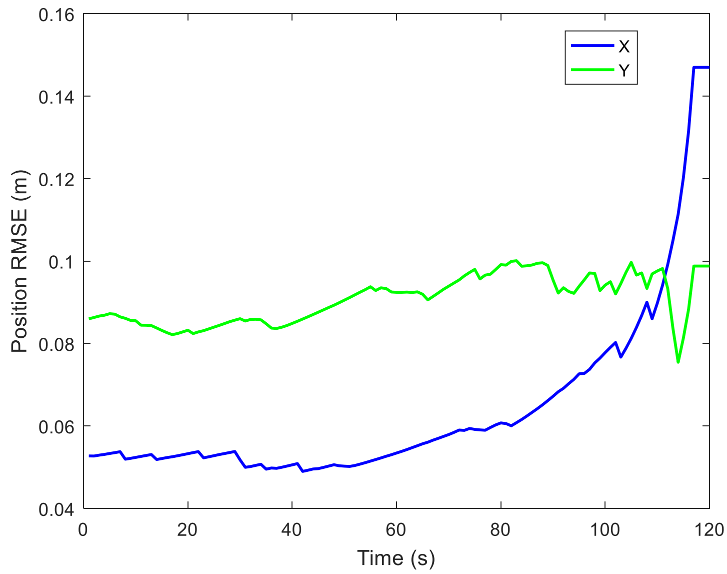
Figure 3.10. Comparing the proposed method with the ground truth. (**a**) Robot trajectory in the XY plane. (**b**) Position corresponding to the trajectory [205].

Furthermore, in order to investigate the performance and compare the accuracy of the proposed algorithm, we evaluated its effectiveness by the root mean squared error (RMSE) defined as:

$$RMSE(t) = \sqrt{\frac{\sum_{t=1}^{T}(x_t^{true} - x_t^{est})^2 + (y_t^{true} - y_t^{est})^2}{T}} \tag{3.32}$$

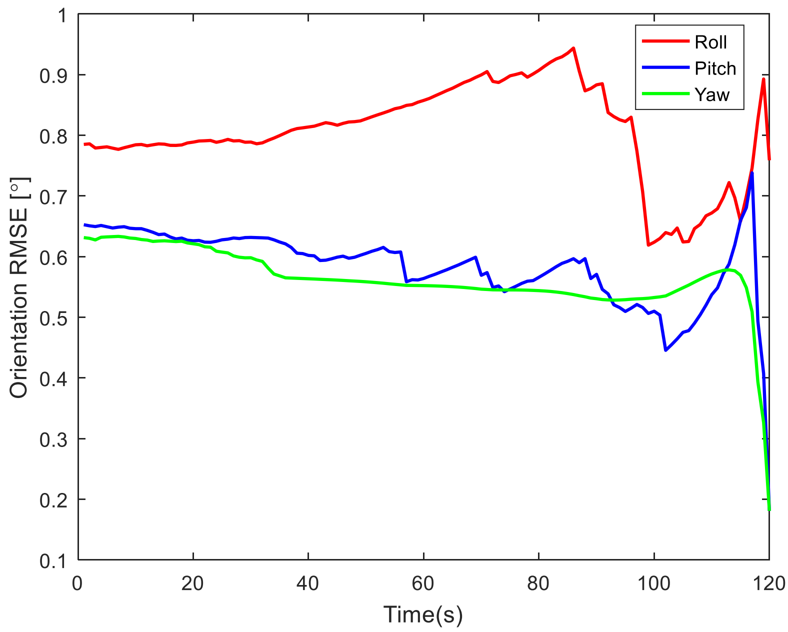
where x_t^{true}, y_t^{true} denotes the ground truth measurement and x_t^{est}, y_t^{est} represents the estimated filter algorithms and T is the total time variable.

Figure 3.11a, b shows the results of error for position and orientation, which is the difference between the ground truth and proposed method. From the graph, it can be deduced that the maximum error value for position and orientation are 0.145 m and 0.95° respectively. These error values are still reasonable for indoor localisation. In Table 3.2, RMSE position and orientation are further stated for specific periods. It can be observed from the table that the position error slightly increases with increase in time. For RMSE orientation, both pitch and yaw error angles decrease as time increases while for roll, error was gradually increasing from the start time to about 80 s and later decreases. The accuracy of the proposed method was improved, and better performances were achieved.



(a)

Figure 3.11a. Results of RMSE for position [185].



(b)

Figure 3.11b. Results of RMSE for orientation

Figure 3.11. Results of RMSE for position and orientation [185].

Table 3.2. RMSE of position and orientation.

Time (s)	Position Error (m)		Orientation Error (Degree)		
	X	Y	Roll	Pitch	Yaw
20	0.05	0.08	0.78	0.62	0.62
40	0.05	0.08	0.81	0.60	0.56
60	0.07	0.09	0.85	0.56	0.55
80	0.06	0.09	0.90	0.56	0.54
100	0.07	0.09	0.62	0.50	0.53
120	0.14	0.09	0.75	0.18	0.18

3.5.4 Results of pose estimation based on natural landmarks

In this section another experiment was performed based on natural landmarks using homography with EKF to estimate the pose estimation of a mobile robot. Figure 3.12 shows the experimental results for IMU, vision and fusion in the XY coordinates. From the results, it can be deduced that the coverage area of vision gave more detailed information than the IMU. The obvious spikes observed from the results, was due to the sudden drifts and imperfectness of the sensors, but it was corrected with the proposed scheme given minimised error when fused. Figure 3.13 depicts the orientation angles. From the same figure it can be seen that spines in vision could be as an effect of lighting changes and other changes in viewing angles. Figure 3.14 shows an illustration of how SURF feature points and FAST points of natural landmarks were extracted from the scene image. Using SURF and RANSAC algorithm, strongest points were extracted to calculate the pose estimation of the mobile robot using the homography estimation. The processing time for features extraction was done within 80 ms which runs more faster than other types of algorithms. This makes the method suitable for vision navigation and localisation.

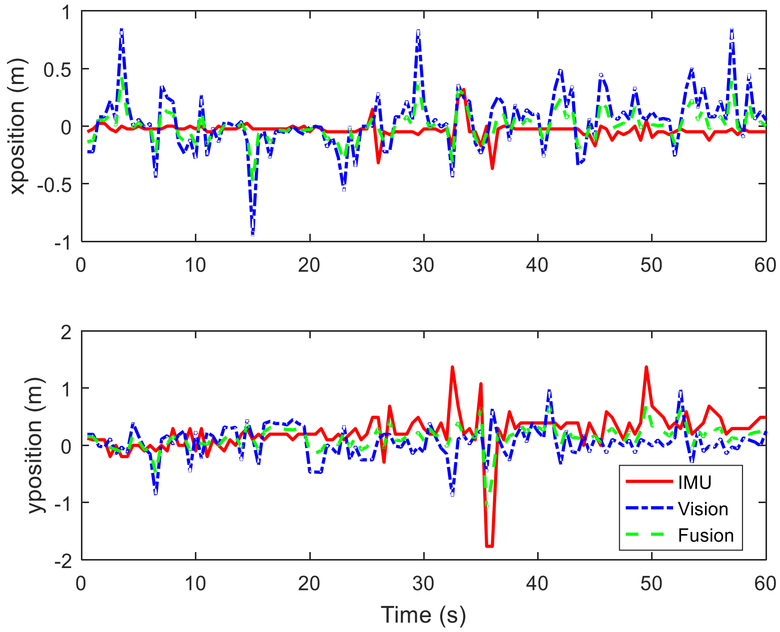


Figure. 3.12. Experimental result in the X and Y position for IMU, vision and fusion taken from [114], © 2017 IEEE.

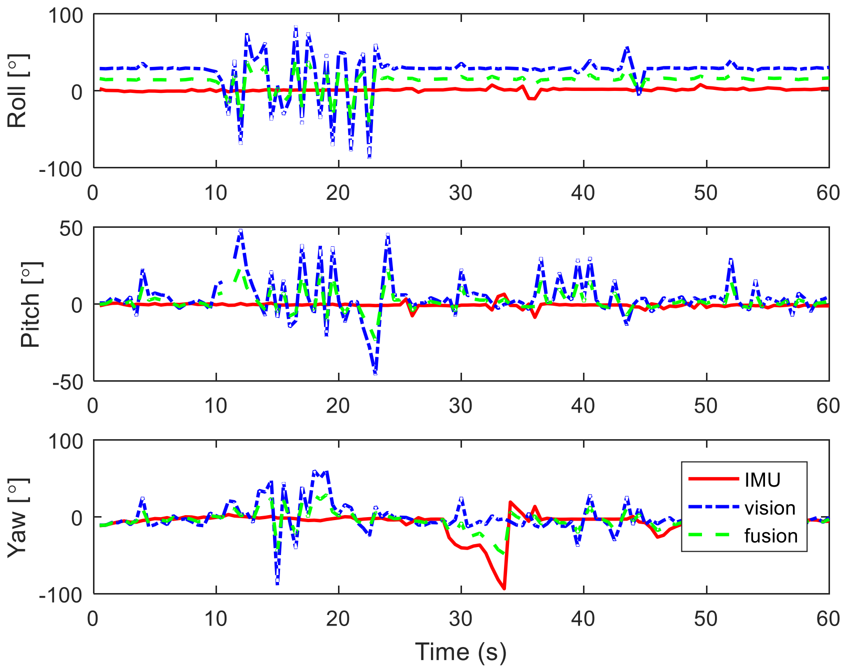


Figure 3.13. Experimental orientation of IMU, vision and fusion taken from [114], © 2017 IEEE.

RMS error for position is shown as Figure 3.15a. The RMSE mean value for IMU as a single sensor is 0.655 m and for two sensor units is 0.183 m. Therefore, fusing two or more sensors reduces the error and as well complements the weakness of a single sensor. Comparing the position error of artificial landmark and natural landmark to estimate position, it can be deduced from Table 3.3 that the use of artificial landmark is more reliable and effective. The RMSE values for orientation were shown in Figure 3.15b. All three angles, roll, pitch and yaw have less than 0.9° error which is reasonable for an indoor localisation.

Table 3.3. Comparison of artificial and natural landmark position error.

Max. position error (Artificial landmark)		Max. position error (Natural landmark)	
X	Y	X	Y
0.147 m	0.09 m	0.18 m	0.65 m

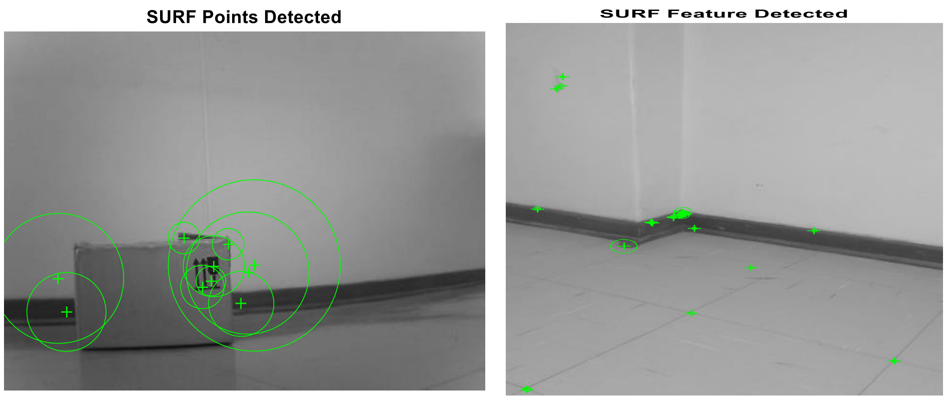


Figure 3.14. Illustration of SURF points extracted from scene image.

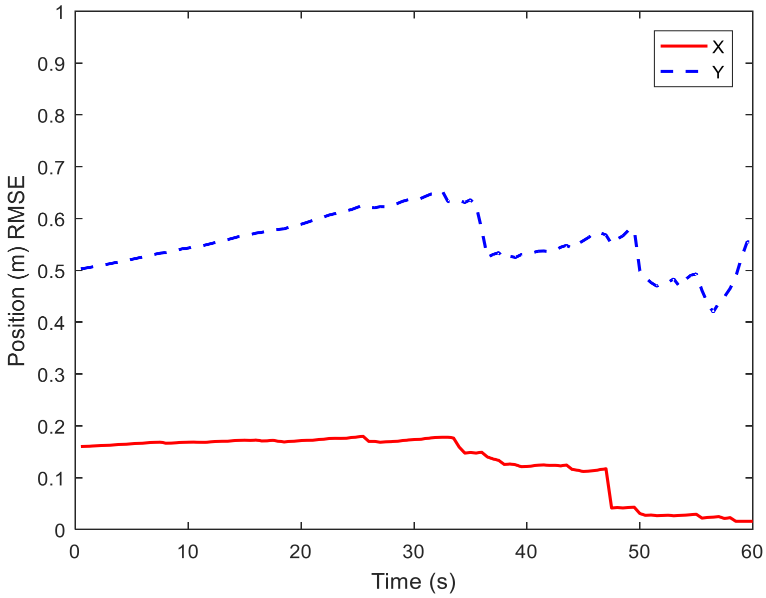


Figure 3.15a. RMSE of position.

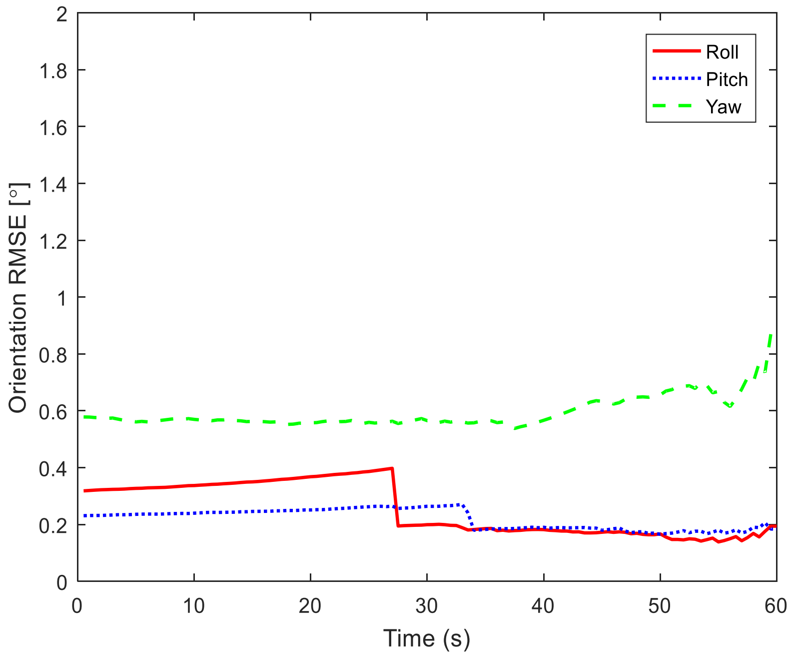


Figure 3.15b. RMSE of orientation.

Figure 3.15. RMSE of position and orientation taken from [114], © 2017 IEEE.

3.6 CONCLUSION

Fusion of vision and inertial measurements to obtain robust and accurate autonomous mobile robot pose estimation was presented for an indoor environment. The inertial sensor used is the 6-DoF, which was used to determine position and orientation. For the computer vision, a single forward-looking camera was used to generate 2D/3D correspondences. The purpose of data fusion is to produce reliable data that is not influenced by accelerometer noise and gyroscope drift. In respect to this, vision was proposed as the best fit to complement the weaknesses of inertial sensors. The inertial sensors and the camera were both mounted on the robot to give excellent performance of the robot estimate.

For object recognition, SURF and RANSAC algorithms were used to detect and match features in images. SURF is used to detect key points and to generate its descriptors. It is scale-and rotation-invariant, which means that, even with differences on the size and on the rotation of an image, SURF can find key points. In addition, RANSAC is an algorithm to estimate the homograph matrix of an image; therefore, the combination of SURF and RANSAC gives robust, fast computation and accurate results for vision tracking scenarios.

The experimental results have shown that a hybrid approach of using inertial sensors and vision is far better than using a single sensor. An extended Kalman filter was designed to correct each sensor hitches by fusing the inertial and vision data together to obtain accurate orientation and position. RMSE values for position and orientation were determined to evaluate the accuracy of the technique. As a result, the method shows reliable performance with improved accuracy.

The study further presented an experimental outcome of using natural landmarks to determine location of a mobile robot. The results confirmed the familiar fact that the effectiveness of the homography estimation is heavily dependent on the appropriate choice of the sensors and robust with fast computation algorithm to improve localisation performance. This type of system proposed can practically be considered for most indoor applications.

CHAPTER 4 DATA FUSION TECHNIQUE BASED ON PARTICLE FILTER

4.1 INTRODUCTION

One of the most recently researched areas in simultaneous localisation and mapping (SLAM) is pose estimation as it is one of the key issues to be tackled in autonomous mobile robot (AMR) operation [217]. AMR have a range of potential applications such as home monitoring [218], space exploration, agriculture, medicine, automotive [219] etc. Great large of AMR are equipped with sensing and measurement systems with the purpose of determining its position and orientation accurately and robustly. The major challenge in localisation is based on how to correct error associated with measurement via sensors and its environmental models. As mentioned in the previous chapters, inertial measurement unit (IMU) is considered as one of the commonly used unit to calculate positioning because of its self-independence leading to its validity in all environments and its capability to quickly obtain good real-time estimation. The IMU unit is comprised of accelerometer and gyroscope sensor [185, 224]. To compensate for the weaknesses of IMU, vision is considered as the most suitable alternative to provide appropriate source of information about the robot and its environment [225]. Information acquired from the environment with the use of camera which could be in form of images or video can be interpreted with the use of computer vision mechanism. The method assists in tracking object identified in the image accurately. Generally, vision based methods are categorized into artificial and natural landmarks [14] and the use of vision can either be monocular (single camera) or binocular vision (two cameras). A single camera [222] could be considered as appropriate because of its low cost and less complexity, however, the demerit of monocular vision is the absence of depth information when compared to the use of two or more cameras.

To have precise positioning system, it is vital that the sensors complement themselves through integration to present better improvement in localisation accuracy. Quite several studies on the use of IMU and vision have been researched by some authors [140, 223]. These studies are mainly concerned with either accelerometer or gyroscope combined with vision using Kalman filter (KF) [224]. The images captured by monocular camera were used in ref. [225] to correct inherent errors acquired by the inertial sensors using decentralized Kalman filter algorithm to fuse data. Conventional Kalman filter could estimate states correctly for a linear system because it uses linearized model, but not appropriate for nonlinear system. For a nonlinear system, it is important to consider data fusion algorithm that is nonlinear based [226]. For this study, particle filter was considered to estimate the state of the mobile robot. The purpose of this experiment is to introduce an effective method that would fuse sources of sensor and thereby estimate accurate position and orientation of mobile robot. The contributions of this chapter are summarized as follows:

- The particle filter based algorithm was adopted to effectively combine and estimate the pose estimation of an autonomous mobile robot using measurement from IMU sensor and camera.
- The proposed particle filter employed the use of increased sample of particles. The algorithm shows a moderate computation load with minimised position errors.
- At the end of the work, experimental results presented shows the benefit of the proposed algorithm.

4.2 ACQUISITION OF SENSOR DATA

This section explains how measurements are acquired from IMU and vision to determine the robot location and orientation.

4.2.1 IMU sensor

Data acquired from IMU was fused together using Kalman filter. The combination of accelerometer and gyroscope from IMU achieves better performance on positioning accuracy in a longer time span. For the definition of reference frame system refer to Chapter 3 Section 3.1. The IMU method provides orientation of the body with respect to (wrt) world frame {w} R_{wb} and vision method provides orientation of the object {o} wrt to camera frame {c} R_{co}, this was obtained through image processing method. As mentioned in Chapter 3, examples of object representations are axis angle, direct cosine matrix (DCM), Euler angles and quaternions [14, 132]. In this study, quaternion was used to solve for roll, pitch and yaw angles. The quaternion is an extension of complex numbers to a four dimensional manifold [227]. The definition is given as:

$$q = q_0 + q_1 i + q_2 j + q_3 k \tag{4.1}$$

where $q_0, q_1, q_2, q_3 \in \mathbb{R}$ and the three imaginary components i, j, k are defined as:

$$i^2 = j^2 = k^2 = -1 \tag{4.2}$$

With the following properties

$$ij = k, jk = i, ki = j, ji = -k, kj = -i, ik = -j \tag{4.3}$$

Therefore, the quaternion representation of rotation is written as a normalised four-dimension vector $\hat{q} = [q_0\ q_1\ q_2\ q_3]$.where $q_0^2 + q_1^2 + q_2^2 + q_3^2 = 1$.

Unit quaternion provides a convenient mathematical rotation for representing rotation matrix The rotation quaternion Q_{wb} in relative to the world frame can be obtained using coordinate information from the body frame. To determine the robot pose, quaternion representation of the coordinate transformation matrix can be calculated as follows:

$$Q_{wb} = \begin{bmatrix} q_0^2 + q_1^2 - q_2^2 - q_3^2 & 2(q_0 q_3 + q_1 q_2) & 2(q_1 q_3 - q_0 q_2) \\ 2(q_1 q_2 - q_0 q_3) & q_0^2 - q_1^2 + q_2^2 - q_3^2 & 2(q_0 q_1 + q_2 q_3) \\ 2(q_0 q_2 + q_1 q_3) & 2(q_2 q_3 - q_0 q_1) & q_0^2 - q_1^2 - q_2^2 + q_3^2 \end{bmatrix} \tag{4.4}$$

The world coordinate system is fixed in inertial space and the body is rigidly attached to the mobile robot whose attitude is defined. The orientation estimation is obtained using quaternion because of its simplicity, mathematical elegance and lack of singularities, therefore, a unit quaternion is used to present the attitude of a rigid body.

4.2.2 Image processing technique to obtain pose

Image processing is considered as another technique that can be used to determine mobile robot pose in the aspect of computer vision. In [228], the authors adopted the use of image processing algorithm to calculate the position and orientation with a proposition of triangulation method for positioning and the implementation of second-order moment for orientation. Here, same scheme was adopted but with a slight difference in positioning method. As shown in Figure 4.1, the camera is placed with a height of 120 cm from the ground to scan the movement of the mobile robot within an actual environment of specified dimension of 255 cm by 250 cm. With the geometrical information extracted from the images captured, accuracy and less complexity is ascertained with the use of image processing method for a mobile robot [229, 230]. To implement the process, it is required to remove background noise and distortion through calibration process [101] especially to cameras that are of low-cost, less quality and poor resolution. To remove distortion and improve the accuracy for the webcam to localise, the camera was calibrated before positioning operation was performed. The following steps below were followed:

- Webcam records the video of robot navigation
- Convert the video to image frames
- Convert original image to grey scale
- Background subtraction
- Binarization
- Detection of the object in the image using SURF algorithm [114].
- Determine the object coordinates.

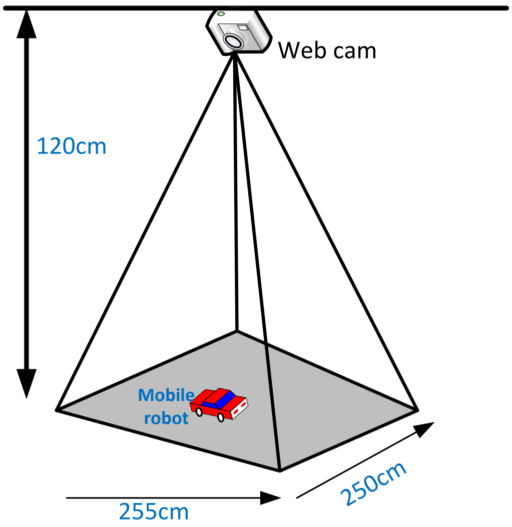


Figure 4.1. Illustration of how the experiment was performed.

For binary image, moment summarize an object given image as $I(x, y)$

$$M_{i,j} = \sum_x \sum_y x^i y^j I(x, y) \tag{4.5}$$

The central moments are translation invariant:

$$\eta_{pq} = \sum_x \sum_y (x - \bar{x})^p (y - \bar{y})^q I(x, y) \tag{4.6}$$

where $p, q = 0, 1, 2...$

To calculate the horizontal and vertical direction, the equation below was used.

$$\bar{x} = \frac{M_{1,0}}{M_{0,0}} \qquad \bar{y} = \frac{M_{0,1}}{M_{0,0}} \tag{4.7}$$

where $M_{0,0}$ is the total number of pixels in the object (object area), where $M_{1,0}$ and $M_{0,1}$ are the first order moment. Therefore, to determine covariance matrix of the binary object this is given as:

$$cov(Obj) = \begin{bmatrix} u'_{2,0} & u'_{1,1} \\ u'_{1,1} & u'_{0,2} \end{bmatrix} \tag{4.8}$$

$$u'_{2,0} = \frac{M_{2,0}}{M_{0,0}} - \bar{x}^2 \qquad u'_{1,1} = \frac{M_{1,1}}{M_{0,0}} - \bar{x}.\bar{y}$$
$$u'_{0,2} = \frac{M_{0,2}}{M_{0,0}} - \bar{y}^2 \tag{4.9}$$

where $M_{2,0}$, $M_{1,1}$ and $M_{0,2}$ are the second order moments. The eigenvectors of the covariance matrix correspond to the major and minor axes of the equivalent ellipse. To determine the orientation θ of the mobile robot [229]:

$$\theta = \frac{1}{2}.\tan^{-1}\left(\frac{2u'_{1,1}}{u'_{2,0} - u'_{0,2}}\right) \tag{4.10}$$

where the major and minor axes length is given as a and b respectively.

$$a = \sqrt{6(u'_{2,0} + u'_{0,2} + \sqrt{4u'^2_{1,1} + (u'_{2,0} - u'_{0,2})^2}} \tag{4.11}$$

$$b = \sqrt{6(u'_{2,0} + u'_{0,2} - \sqrt{4u'^2_{1,1} + (u'_{2,0} - u'_{0,2})^2}} \tag{4.12}$$

Algorithm 1: Systematic sampling [231]

$\left[\left\{\hat{x}_t^i\right\}\right]_{i=1}^N = \text{Resample}\left[\left\{x_t^{(m)}, w_t^{(m)}\right\}_{m=1}^{(m)}, N\right]$

$\left[\left\{Q_t^{(m)}\right\}_{m=1}^M\right] = \sum\left[\left\{w_t^{(m)}\right\}_{m=1}^M\right]$

$i = 0$

$m = 1$

$u_0 \sim U\left[0, 1/N\right]$

while $(i \leq N)$

$u = u_0 + i/N$

while $(Q_t^{(m)} < u)$

$m = m + 1$

end

$i = i + 1$

$\hat{x}_t^{(i)} = x_t^{(m)}$

End

4.3 DATA FUSION METHOD BASED ON PARTICLE FILTER

In recent times, data fusion in mobile robots has become a dominant pattern due to its potential advantages like reduction in uncertainty, increase in reliability and accuracy [232]. Several methods have suggested the fusion of data from different sources improved performances. According to the authors in ref. [233], extended Kalman filter algorithm is one of the most suitable algorithms to fuse data because it converges fast and it works well for nonlinear systems unlike Kalman filter that is more likely to work better for linear systems. KF also has limited abilities in providing accurate estimation of such system parameters, because it is delimited to use only Gaussian linear models for these sensors' stochastic errors. To ensure system availability and system's authenticity that define the exactness of the objects perceived data fusion is most suitable. Unscented Kalman filter is considered as another algorithm that works efficiently well for integration of data with the advantage of low computational cost, but its accuracy is not as high as particle filter. In robotics community, PF has been developed into one of the most effective algorithms for solving robot localisation problem and to deal with non-Gaussian distribution. The basic variants of PF are the extended version of the sequential importance sampling (SIS) algorithm with an added sampling step known as the sequential importance resampling (SIR). Particle filter was proposed by Gordon [237] which chose prior probability function as proposal probability function and use sequential particle sampling. Particles filters can be applied with data from almost any type of sensor. It has become very famous as it can deliver a quick location estimate, while handling uncertainties in sensor data and supporting re-localisation without adding algorithm complexity. This study seeks to implement on the basics of particle filter to determine the position and orientation of an autonomous mobile. The proposed particle filter is compared with three other existing methods which are EKF already explained in detail in the previous chapter, auxiliary particle filter (APF) and adaptive particle filter.

Auxiliary particle filter was suggested to improve the weakness of PF which provides enhanced solution for effective sampling when the observation model is known. The algorithm is reliable to determine optimally sample from posterior initially given by Pitt and

Shephard [239]. The same algorithm was implemented by the authors in [240] to tackle the problem of robot localisation. The use of APF is associated with the fact that optimal importance function, the weight at time t does not depend on the state x_t. Therefore, it seems wasteful to resample particles at the end of iteration $t-1$prior to consideration the observation y_t. The essence of APF was that the sampling step could be modified to sample, for each particle, an auxiliary variable corresponding to a particle index according to a distribution which weighs each particle in terms of it compatibility with the observation. When APF is considers as a sequence of weighting and sampling, it becomes apparent that it also has an interpretation as mutation selection algorithm. For a fully adopted APF, the importance weights by which estimation are made are uniform resampling is carried out both before and after auxiliary weighting as in the original implementations. APF is likely to yield more stable estimates provided that a good approximation of the procedure likelihood is available. APF in the prediction step, favor particle that are likely to get high likelihoods after incorporating the measurement in the update step. The availability of the latest measurement is exploited in the prediction step, instead of unnecessarily employing more sample from the prior. It uses resampling on predicted particles to select which particle to use in the prediction and measurement update. The pseudo code of the algorithm is given as Algorithm 2 below and the complexity is O(N^2). The steps below summaries the algorithm of auxiliary particle filter:

1. Compute N point estimate that are used to characterise $p(x_t \mid x_{t-1}^i) : \mu_t^i \sim p(x_t \mid x_{t-1}^i)$. Different characteristics are possible leading to a different variation of the particle filter. Compute weights for the characterisations: $w_t^{(i)} \propto p(y_t \mid u_t^i) w_{t-1}^i$. Thereafter, normalise the weights of each kernel. A kernel is a symmetric pdf and is a function of the particle state.
2. Use the weight w_t^i from step 1 in a resampling step. During resampling store the indices ij of the particle that would have been selected but do not perform the actual resampling. Each index ij refers to a particle at $t-1$ and the set if indices represent the set of particles that are expected to get high likelihoods.

3. Perform a prediction step for each of the N particle index from step 2.
4. Compute the weights for the propagated particles from step 3.

$$w_t^{(j)} \propto \frac{p(y_t \mid x_t^j)}{p(y_t \mid u_t^{ij})} \tag{4.13}$$

In the algorithm described above the characterisations are used to predict a particle as expected to get a high likelihood given the measurement.

Algorithm 2: APF algorithm [241, 248]

Initialization. At time $t=0$ draw N sample $x_t^{(i)}$ from the distribution $p(x_0)$

Set $w_0^{(i)} = 1/N,\ i,\dots,N$

Let $x(_{t-1}^{(i)})_{i=1}^{N}$ be the particles (samples) generated at $t-1$

Compute the mean of pdf $p(x_t \mid x_{t-1}^{(i)})$ as

$\mu_t^{(i)} = E_{p(x_t \mid x_{t-1}^{i})}.[x_t]$

Compute the normalised weight of each kernel in the mixture as

$\lambda_t^{(i)} \propto p(y_t \mid \hat{x}_t^{(i)}) w_{t-1}^{(i)}$

Let the importance sampling (IS) be

$q(x_t) = \sum_{i=1}^{N} \lambda_t^{(i)} p(x_t \mid x_{t-1}^{(i)})$

Draw N samples from $q(x_t)$ in two steps:

Select the indexes ij , $i = 1,\dots,N$

Simulate $x_t^{(i)} \sim p(x_t \mid x_{t-1}^{(ij)})$

Compute the weight as

$w_t^{(i)} = \dfrac{p(y_t \mid x_t^{(i)})}{p(y_t \mid \mu_t^{(ij)})}$

Normalise the weights

$\bar{w}_t^{(i)} = \dfrac{w_t^{(i)}}{\sum_{j=1}^{N} w_t^{j}}$

Adaptive particle filter: Adaptive particle filter is also consider to improve the challenging issues standard particle filter is faced with. This method suggests to ensure that the number of particles is kept minimal as possible. The idea of adaptive particle filter is to use a small number of particles in a situation where particles are based on a small part of the space and increased number of particles in case the particle filter uncertainty increases. For implementation purpose, the pseudo code is explain briefly.

- Sample a particle index proportional to its weight at time $t-1$
- Propagate the particle sample using the process model
- Update the number of required particles
- If the number of particle equal the number required particle (or the maximum number of particles), stop sampling new particle. For more information refer to [247].

4.3.1 Proposed particle filter

When using PF, the prediction of the robot state is established on a set of samples, which are initiated at time t and they are allocated weights individually depending on the initial value of each particle in relation to the nonlinear function. The particle filter algorithm adopts Sequential Importance Resampling (SIR) and the SIR algorithm uses the resampling method by implementing a threshold to disregard sampled particles which fall outside of the threshold [151, 235]. Resampling is an essential component of PF, because it combats sample degeneracy. The following advantages also have encouraged PF to be preferred to other types of filter methods: high accuracy and stability, fast convergence rate and not difficult to implement. The proposed method utilized particle filter to localise the mobile robot by fusing IMU and vision. The distinct feature of the proposed method is using more particles with a novel distribution function with a higher likelihood in which it has a better chance of surviving. To estimate robot localisation, particles filters are a great method to track the state of a dynamic system using Bayesian model. Bayes' rule is adopted to determine the state of a nonlinear dynamic system in respect to time. The robot seeks to estimate a posterior distribution over robot state space conditioned on the available sensor data. Therefore, the robot state at time t is denoted as, x_t and the observation is y_t. The system is represented as:

$$x_t = f(x_{t-1}, u_{t-1}) + Q \leftrightarrow p(x_t \mid x_{t-1}, u_{t-1}) \tag{4.14}$$

$$y_t = h(x_t) + R \leftrightarrow p(y_t \mid x_t) \tag{4.15}$$

where Q is the process noise and R is the measurement noise. x_t is estimated as the state vector and u_t is the control vector. f and h are assumed to be known functions. The fundamental idea of the system is to develop the posterior density by a random set of particles with associated weights and the estimate is evaluated based on samples. The samples drawn from the required distribution need no assumption to be made when the state space model is either nonlinear or linear [151]. Detailed explanation of particle filter is given in reference [236]. For this study, the implemented particle filter for data fusion is presented below:

The state vector x_t of the fused filter consists of angular velocity ω, quaternion for orientation q accelerometer a and velocity v.

$$x_t = [\omega\, a\, q\, v] \tag{4.16}$$

The measurement vector y_t consists of orientation θ and position p from vision.

$$y_t = [\theta\, p] \tag{4.17}$$

The relation between the measurement and the states is modelled by nonlinear function given in (4.14) and the initial a prior PDF $p(x_0)$ of the state vector is assumed to be known. Particle filters (PF) are capable of handling highly nonlinear models with any kind of noise distribution. The particle filter is a special version of the Bayesian filter, and is based on sequential Monte Carlo (SMC) sampling.

Initialisation: The particle filter initialised by drawing sample x_0^i, $i = 1, ..., N$, from the prior density function $p(x_0)$ and set weight w_0^i to $1/N$, N is the number of particles.

Sampling: samples are not drawn according to $p(x(t) \mid y_{0t})$ but are drawn from importance probability density function.

$$\hat{x}_t^i \sim q(x_t \mid x_{0:t-1}^i, y_{1:t}) \tag{4.18}$$

Prediction: Prediction is drawn from the proposal which is the motion model centered on IMU data and it is exploited to predict the current position x_t of the mobile robot in current step, given control input u_{t-1} and the previous posterior probability x_{t-1}.

$$x^{(i)}_t \sim p(x_t \mid x_{t-1}, u_{t-1}) \tag{4.19}$$

Measurement Update: Measurement update is performed via the observation model in which the sensor measurement (i.e. vision) is incorporated into predicted density which leads to posterior probability density function (PDF) of the state x_t .

Considering conditional probability distribution of future states, given the present state and all past states, depends only upon the present state and not on any past states and the measurements are assumed to be conditionally independent given the states, the following recursive general equation for weights updating can be obtained:

$$w_t^{(i)} = w_{t-1}^{(i)} \frac{p(y_t \mid x_t^{(i)})\, p(\hat{x}_t^i \mid x_{t-1}^{(i)})}{q(\hat{x}_t^{(i)} \mid x_{0:t-1}^{(i)}, y_{1:t})} \tag{4.20}$$

where w_{t-1}^i represents the weight at the previous time step for particle i, $p(y \mid x_t^i)$ is the likelihood function, $p(x_t^i \mid y_{t-1}^i)$ transition prior function.

The choice of importance density function $q(\hat{x}_t^{(i)} \mid x_{0:t-1}^{(i)}, y_{1:t})$ is one of the most critical issues in the design of a particle filter. The reason of this is that the samples are drawn from the proposed distribution, and the proposed distribution is used to evaluate importance weights. The most popular suboptimal choice is to use the conditional prior of the state vector as the proposed distribution for importance density function

$$q(\hat{x}_t^{(i)} \mid x_{0:t-1}^{(i)}, y_{1:t}) = p(\hat{x}_t^i \mid x_{t-1}^i) \tag{4.21}$$

Therefore, substituting (4.21) in (4.20) the weight's update equation will be

$$w_t^i = w_{t-1}^i . p(y_t \mid x_t^i) \tag{4.22}$$

The PF was invented to numerically implement the Bayesian estimator. Instead of applying prior probabilities in Bayes estimation, it employs a set of particle with values and weights to approximately represent $p(x_t \mid y_t)$ through the Monte Carlo Sampling approach. This permits the PF to be integrated into Bayesian solution to sums of weighted sample draw from posterior distribution. The probability density function $p(x_t \mid y_{1:t})$ of the estimated state can be represented by

$$p(x_t \mid y_{1:t}) \approx \sum_{i=1}^{N} w_t^{(i)} \partial(x_t - x_t^i) \tag{4.23}$$

This represents the discrete weighted approximation of the true posterior $p(x_t \mid y_{1:t})$, where $\partial(\cdot)$ is the Dirac delta function. The weights' values are always positive $w_t^{(i)} > 0$, and sum of weight is equal to 1.

$$\sum_{i=1}^{N} w_t^i = 1 \tag{4.24}$$

A common problem with the SIS particle filter is the degeneracy phenomenon, where after a few iterations all but one particle will have negligible weight. This degeneracy implies that a large computational effort is devoted to updating particles whose contribution to the approximation to $p(x_t \mid y_{1:t})$ is almost zero. Of course, the degeneracy problem is an undesirable effect in PF. An approximate effective sample size $\hat{N}_{eff}$ can be used to detect the occurrence.

$$\hat{N}_{eff} = \frac{1}{\sum_{i=1}^{N} (w_t^{(i)})^2} \tag{4.25}$$

If ($\hat{N}_{eff} < N_{thr}$), a little N_{eff} indicate severe degeneracy. N_{Th} is denoted as thresholds.

Resampling: Resampling is proposed to reduce the effect of degeneracy. Resampling is necessary to limit samples with non-zero weights. It is a scheme to reject particle with small weight to concentrate and replace on particle with large weights [237]. Systematic sampling was used as resampling method because it is simple to implement, has computational complexity O(N) and reduces the variance of importance weight variation. Systematic sampling is given as Algorithm 1, more details are given in ref. [238].

Figure 4.2 shows the block diagram of the sensor fusion. It can be deduced from the figure that IMU and vision were used to determine the mobile robot pose estimate. Webcam was used to capture the moment of the mobile robot within the area of experiment. The acquire images is required to go through processing for removing noise and unnecessary information. Thereafter, the object detection such as SURF algorithm was used to identify the robot in the images and orders of moment technique was used to calculate the direction and orientation of the object in image. IMU which is comprised of accelerometer and gyroscope were fused by KF to estimate their position and orientation which was integrated to obtain the direction of the mobile robot. In this case, the attitude of the body was presented using quaternion. KF is used to estimate the IMU measurement because the state-space model is linear and noise distribution is considered to be zero-mean Gaussian. Instead of using the direct measurement of position and orientation from IMU, the KF estimations are used to reduce effect of measurement noise. Thereafter, the estimation from the IMU and vision are fused together using proposed particle filter to give output estimate. PF is used to estimate orientation and position because it neither requires the state-space model to be linear nor assumes that the noise is Gaussian unlike KF. When the initial state are unknown, they will be converged to the correct values in PF. PF approximates posteriors with a set of state samples, called particles instead of assuming that the posteriors are Gaussian every time step.

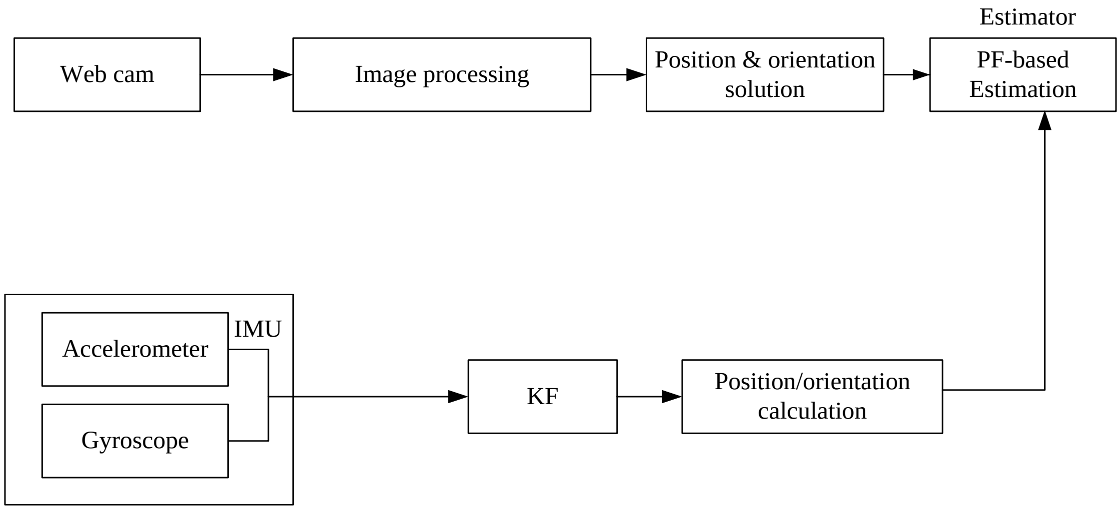


Figure 4.2. Block diagram of sensor fusion system.

4.4 RESULTS AND DISCUSSION

4.4.1 Components used and experimental setup

The components used to perform the experiment are: 4WD mobile robot, Arduino 101 IMU, webcam and pixy camera. The mobile robot used in this experiment is a four-wheel drive (4WD) with a working voltage of 4.8 V. Four servo motor controllers were used which allowed the robot to move up to 40 cm/s (0.4 m/s) with microcontroller (Arduino/Genuino 101) which has built-in of Inertial Measurement Unit with the integration of accelerometer and gyroscope sensors which is also called 6-DoF. The robot was equipped with a 6 V battery to power the servo motors and a 9 V battery for the microcontroller [63]. To capture the video of the mobile robot, Sparkfun USB webcam was used. It's a camera that can be plugged into computer or dev board. It has a resolution of 640x480 pixels, which is adequate for object detection and motion tracking. 8,000 frames were recorded at the above mentioned resolution at 30Hz. A software called `honestech` (ver 2) was installed on the personal computer (PC) to record video. The camera was elevated to a height as shown in Figure 4.1 in which it was connected to a PC for the webcam to record the moments of the mobile robot navigating through a pre-defined path. Thereafter, the image processing technique was adopted to filter unnecessary information and extract vital information that could aid the robot positioning. The measurement from IMU and webcam were both estimated using proposed particle filter to estimate the pose estimation of the mobile robot. For the ground truth, we adopted the use of a Pixy camera which was connected to the microcontroller through the serial peripheral interface (SPI). The Pixy camera was positioned beside the webcam with a distance of 10 cm so as to have a close and clear field of view like the webcam. The data collected were saved on an SD card and later analysed offline in `MATLAB`. The Pixy camera has the proficiencies to remember seven different color signatures, detect hundreds of objects at the same time and also able to quickly determine the location of the objects detected. Several authors [242-244] have worked on Pixy camera because of its simplicity, low cost and reliability in usage majorly for projects based on robotics. Image processing takes time and limits frequency at which position can be obtained from a camera.

The sample rates for IMU and camera capture are 100 Hz and 30 Hz respectively. Because of the instabilities of the values for both width and heights of the image frame during the mobility of the robot, using the method proposed in [245], the center of the image frame with the present position of the camera are used to determine the location of the mobile robot in relation to the world coordinate. The width and height are recalculated using (4.26) and (4.27).

$$W = (M_x \times A_d) / D \tag{4.26}$$

$$H = (M_y \times A_d) / D \tag{4.27}$$

where M_x and M_y are the distance moved by mobile robot to the x and y axes. A_d is the measurement of the marker used which in this case is the mobile robot itself, while the distance from the preceding position to present position of the mobile robot is represented by D as shown in Figure 4.3. This is given as (4.28). The robot has a fixed reference of (XOY) in which the centre of the robot has coordinates X and Y, and O is the centre of the robot. From the same figure x_r and y_r are defined as reference related to the robot in which x_r either moves backward or forward and y_r is perpendicular to x_r axis.

$$D = \sqrt{\left(X_2 - X_1\right)^2 + \left(Y_2 - Y_1\right)^2} \tag{4.28}$$

Therefore, to convert to the current position of the mobile robot to real world state, we used (4.29) and (4.30):

$$X_R = \left(X_c \times A_c\right) / W \tag{4.29}$$

$$Y_R = \left(Y_c \times A_c\right) / H \tag{4.30}$$

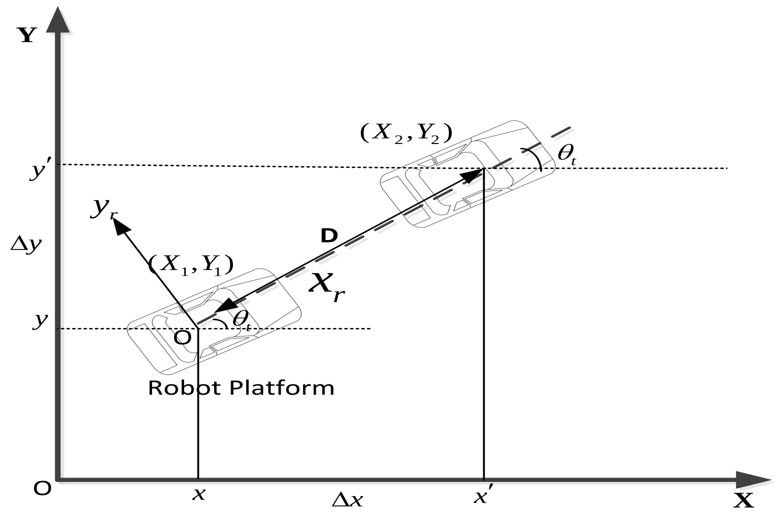


Figure 4.3. Arbitrary motion of mobile robot. [245], © 2015 IEEE.

To evaluate the performance of the proposed method we used (4.31) to calculate the position error as:

$$E = \sqrt{\left(x^{true} - x^{est}\right)^2 + \left(y^{true} - y^{est}\right)^2} \tag{4.31}$$

where $\left(x^{true},\ y^{true}\right)$ is the true measurement obtained and $\left(x^{est},\ y^{est}\right)$ represents the estimated coordinate of data fusion methods.

Table 4.1. Average error for data fusion algorithms.

Method	**Average position error (cm)**	**Average orientation error (deg)**
EKF	9.45	0.89
PF	5.82	0.614
APF	5.81	-
PPF	3.64	0.644
Adaptive	-	0.622

4.4.2 Performance analysis

We illustrated the performance of the algorithm by adopting a simple motion model taking into account by adding an amount of random Gaussian noise to IMU. While the observation model is a function that measures the likelihood of the current y_t if the robot is in the pose x_t. In the first experiment we obtained the 2D trajectory plot of the mobile robot in an indoor environment as illustrated in Figure 4.4. The position estimation of the x and y direction of the mobile robot were acquired through the integration of accelerometer parameters and the the rotation of the mobile robot was derived from gyroscope. The fusion of the IMU and vision was determined by combining the process using (4.15). The figure shows the result of IMU, ground truth and fused data from vision and IMU with PF. From the experiment, the IMU showed a quick response to the environment while the ground truth and fusion results had a slow response. This denote that IMU respond abruptly to factors such as errors or environmental noise. The second and the third experiments were performed to analyse the proposed scheme. The performance and the reliability of the algorithm can be assessed through the calculation of the positioning and orientation errors in the various schemes. The IMU was attached to the mobile robot to obtain the position along the x-axis in the body frame with the algorithm developed, if the IMU is unable to acquire the measurement, images of the robot location are captured to estimates the pose. This was implemented by estimating the set of normalized importance weights according to (4.22) basing on IMU measurement and vision measurement. The nonlinear measurement model $h(x_t + R)$ is directly used in PF, what is impossible for KF. In order to verify the performance of the proposed algorithm, this study compares with standard particle filter, auxiliary particle filter (APF) [236], EKF and adaptive particle filter. APF was considered to improve some deficiencies of SIR algorithm such as increase in variance. The algorithm is also is a nonlinear system model in which the particles are modified before resampling according to likelihood function. Unlike the standard particle filter, APF gets it particles from joint probability density.

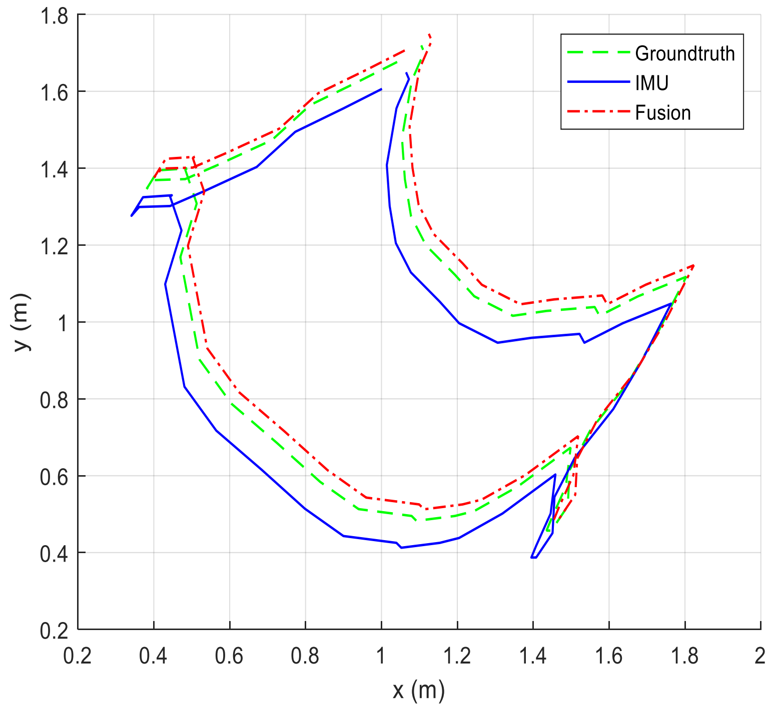


Figure 4.4. Trajectories of comparison robot's movement (ground-truth, IMU, fusion).

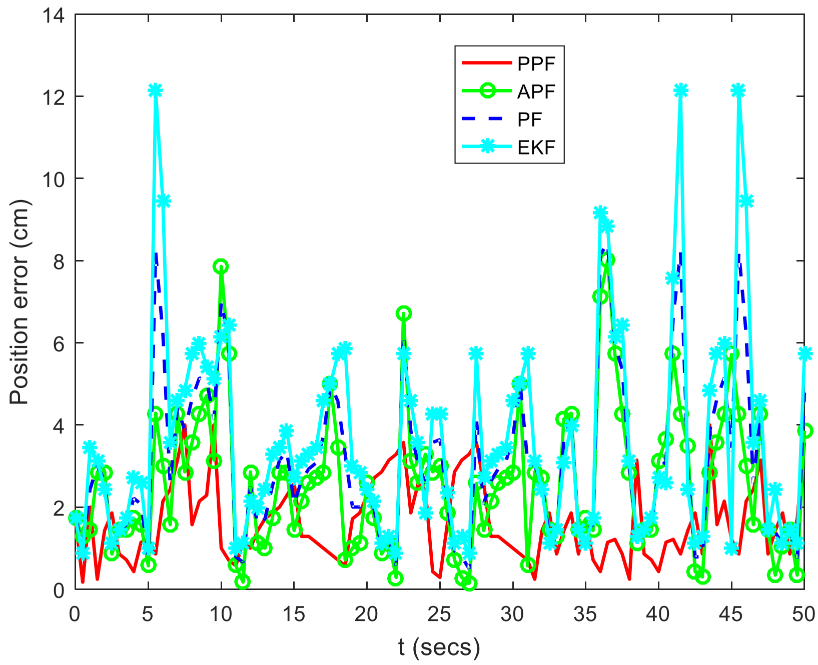


Figure 4.5. Position error (a) Proposed Particle filter (PPF) (red) (b) EKF (cyan) (c) APF (green) (d) PF (blue).

In Figure 4.5, to extract relevant information about the performance of the filters compared with the proposed method, $N = 1000$ particles was used, and the simulation was carried in MATLAB with the duration of 50 seconds for each run. The simulation was carried out to estimate the performance of each filter. It can be inferred from the figure that the proposed particle filter has minimised position error. This is due to increase in number of particles, application of weight and likelihood observation and sampling method against the particle filter which used SIR algorithm. As these results suggests, the position errors at the starts for all the methods were low, but a sudden spike was noticed, and this continued for EKF due to measurement noise likewise for PF and APF. This signifies that the methods are less robust due to the spikes present. From the same figure, proposed particle filter (PPF) with the least maximum position error of 4.02 cm, PF 8.01 cm, APF 7.85 cm and EKF 12.05 cm proved enhanced method over others. This means that PPF recovers from erroneous position while other changes more slowly.

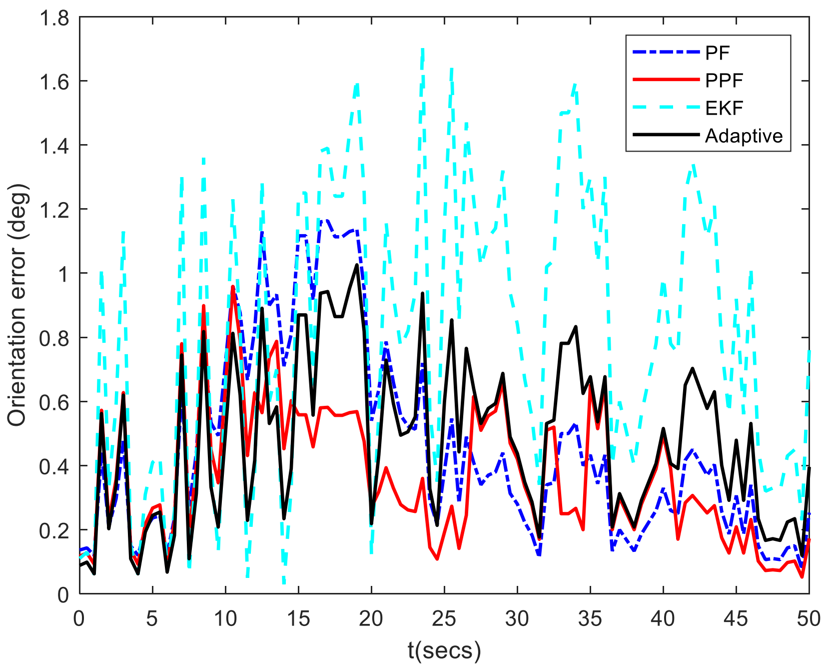


Figure 4.6. Orientation error (a) Proposed Particle filter (red) (b) EKF (cyan) (c) PF (blue) (d) Adaptive particle filter (black).

For Figure 4.6, one of the factors that affected the orientation error was the high number of particles and the sensor noise. A significant change with sudden drop in error was observed when the robot takes a turn at time instances. This means a significant change in direction could enable the robot to respond with an effect on measurement. It can also be inferred that EKF has the highest degree of orientation error compared to the other three methods (PF, PPF and adaptive). This relate to the fact that the use of number samples can improve the efficiency of estimating robot localization. The overall performance of the pose estimation was given in Table 4.1. It can be concluded that the PPF has minimised position and orientation error compared to other existing methods, but for orientation error PF showed a better performance over PPF with an average of 0.614 degrees which is still considerable for an indoor localization.

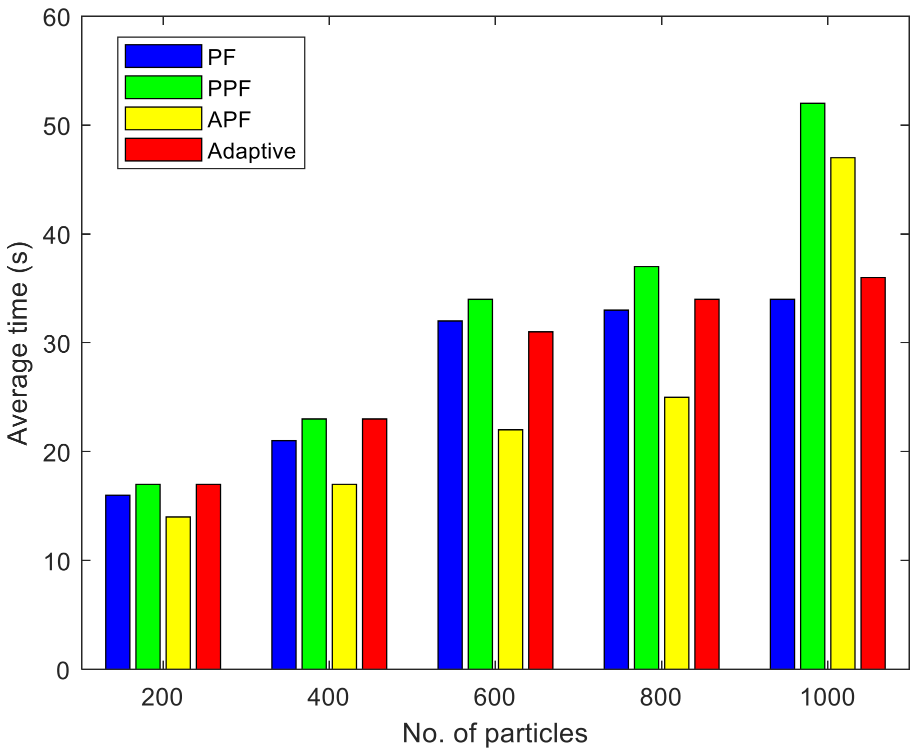


Figure 4.7. Comparison of processing time between PF, APF, Adaptive particle and PPF with different number of particles.

Figure 4.7 shows that as the number of particles increases the processing time also increases, however when the particle was 800 and above, the average time of PPF sporadically overshoot the average time of the other three methods. This indicates that more time is required to process more particles and resampling stages. Therefore, the number of particles should be chosen carefully to attain a tradeoff between accuracy and computational efficiency. This experiment proves that PF shows a better filter for data fusion than EKF especially for nonlinear system. To deal with nonlinear system subjected to non-Gaussian system and measurement noise characteristics, PF is an improved tool. From the results presented, the proposed approach was able to determine an accurate position and maintain a reduced error for localisation when compared to the other existing methods. From the experiment and simulation performed, it is clear that more number of samples could increase processing time, but it is essential for samples to denote ambiguous circumstances happening due to increased uncertainty.

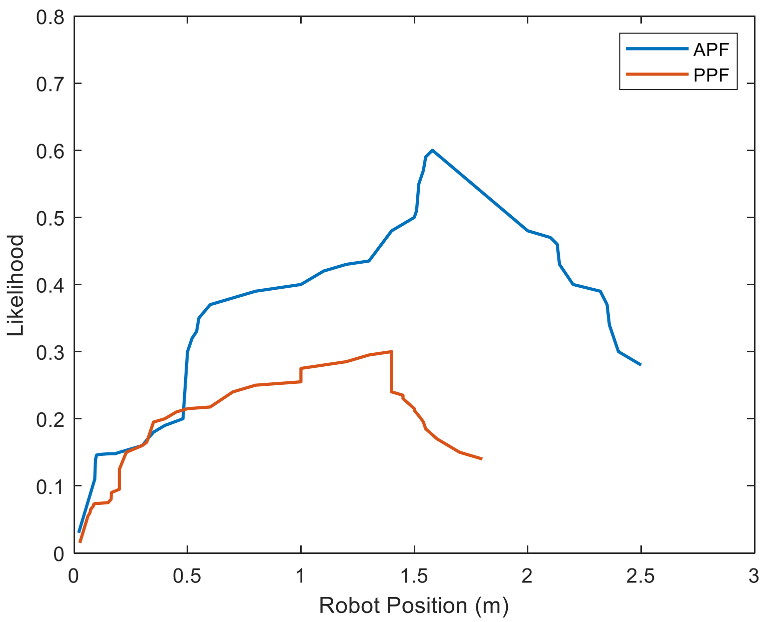


Figure 4.8. Relationship between likelihood and robot position.

Figure 4.8 shows the relationship between likelihood and robot position. They are simulated with 50 time steps with 1000 particles. The relationship between likelihood and pose estimation is shown in (4.20) and (4.31). We plot the distribution of likelihood $p(y_t \mid x_t^j)$ for auxiliary particle filter and particle filter. APF got high likelihood after incorporating measurement in the update step rather than drawing samples from prior. The likelihood $p(y_t \mid x_t)$ represent the conditional probability of a measurement given the predicted state $p(x_t \mid y_{1:t-1})$. It is important to note that the observation model is a function that measures the likelihood of the current measurement y_t if the robot is in the pose x_t . With increase in likelihood the figure shows that more distances are covered within a time frame using APF over the proposed particle filter algorithm. The simulation was performed on Windown7 Ultimate platform (64-bit operating system) with clock speed 2.20GHz and 2.00GB RAM.

4.5 CONCLUSION

This chapter investigated the issue of mobile robot localisation in an indoor environment. A system assembly that is comprised of low-cost monocular vision and IMU is considered. The contribution of this chapter is using image processing technique based on first and second order of moment to estimate the position and orientation of the mobile robot in the environment. Secondly, the method which was based on particle filter was used to fuse data from inertial sensor and vision sensor to improve accuracy of localisation. From the presented results, particle filter algorithm achieved reliable results when compared with other existing methods.

CHAPTER 5 CONCLUSIONS AND FUTURE WORK

5.1 CONCLUSION

Due to the services rendered to the community and globe at large, autonomous mobile robot (AMR) has attracted immense attention and recognition in recent times. Their application virtually in all sectors has set precedence for effectiveness and efficiency in their performances. It is with these requirements that the inherent challenges of autonomous robot mobile such as obstacle avoidance, navigation, localisation and path planning have to be tackled. As earlier mentioned in Chapter 1 that one of the most cogent issue is localisation. Therefore, a detailed investigation into essentials and intricacies of localisation in autonomous mobile robot is vital. Localisation or tracking is one of the fundamental competencies required by an autonomous robot as the knowledge of the robots' location is a vital precursor in making decision about the future of the robot. For adequate tracking and accurate estimation of mobile robot pose (position and orientation), it is therefore very paramount to ensure that an effective, reliable and robust scheme is implemented, and this is the aim of the study. To determine location, the use of sensors is of crucial importance for efficient and accurate robot operation. Robots equipped with sensors are used to obtain data and collect necessary information from the environment through some embedded computer processing. However, these sensors are limited in one area or the other. This study had three major objectives: to investigate the feasibility of using light weight, low power consumption and low-cost sensors to determine localisation, to investigate a robust and effective algorithm appropriate for object identification and detection and finally to investigate an appropriate data fusion method that is most suitable to fuse multiple sensors with less complexity.

There are several types of sensors that can used to estimate the position of a robot but for the purpose of low cost devices and simplicity, this study considered the combination properties of accelerometer and gyroscope sensors to give accurate and reliable measurement of orientation and position and this is achieved through sensor fusion which was done by using Kalman filter. Robot localisation techniques need to be able to deal with noisy observations and generate not only an estimate of the robot location but also a measure of the uncertainty of the location estimate. Having identified the gap in knowledge faced with AMR in that enough study has not be investigated and several possible solutions have been ignored, the challenge to undertake the research study presented in this thesis emerged. One of the possible solutions to determine pose of a mobile robot is using more than one sensor because a singular sensor or system may not be sufficient to estimate appropriate location of mobile robot.

In contrast to most sensory systems, visual system provides very rich information. To analyse the configuration of robotic vision system, it is therefore necessary to distinguish between possible placements of cameras. The camera can be placed in a fixed configuration where they are rigidly mounted in a specific place in the environment or in a mobile configuration where the camera is attached to a robot. In addition to the configuration, the number of cameras used can also contribute to the performance either monocular vision (single camera) or binocular vision (two cameras) can be adopted to determine robot location. Both configurations were considered in this study but using monocular vision method. Computer vision is another paradigm that can ensure reliable achievement for robot localisation. For object recognition and detection, speeded up robust features (SURF) and random sample consensus (RANSAC) algorithms were used to detect and match features in images. SURF is used to detect key points and to generate its descriptors. It is scale-and rotation-invariant, which means that, even with differences on the size and on the rotation of an image, SURF can find key points. In addition, RANSAC is an algorithm to estimate the homograph matrix of an image; therefore, the combination of SURF and RANSAC gives robust, fast computation and accurate results for vision tracking scenarios.

One such approach to localisation is the extraction and use of natural landmarks. Natural landmarks in mobile robot localisation poses a challenge due to the variances in shape, design and the effect of illumination circumstances. Therefore, we proposed the use of homograhies approach for the identification and recognition of natural landmarks to be implemented in mobile robots' localisation. Such landmarks include corners, edges and lines in a known indoor environment. The algorithm is used to determine pose estimation between two relative images. The effectiveness of landmark detection is related with the existence of markers in the location, sensor accuracy, frame rate and with the capability of the robot system processing. Our approach was able to distinguish natural landmarks in front of the robot and determine its positioning.

In order to maximize the working performance of the mobile robot, it is necessary to estimate and track the current pose of the mobile robot. Sensors and devices used to determine their poses are accumulated with errors and noises; therefore, the use of sensor fusion algorithms is applied to solve the problem. Fusion techniques are regarded as the most appropriate method to track objects and determine their locations. The advantages of using sensor fusion method include decrease in uncertainty, increase in accuracy and reduction of cost. Nonlinear sensor fusion based algorithm was considered as a state estimation method to determine mobile robot location and orientation in this study. The two most common techniques, extended Kalman filter (EKF) and particle filter (PF) was used to combine information from sensors (IMU and vision). These techniques are based on mathematical models used to describe the robot motion and observation from sensors. Proposed sensor fusion approaches correct the mobile robot pose and the error is limited within the best possible with the environmental model used. The EKF method produces high correction values, so that the mobile robot reacts more quickly to the increasing real and estimated pose incompatibility. The method contributes to a better performance regarding EKF convergence and stability to a more reliable feature matching process. The pose tracking and estimation performances of the non-linear model-based estimators were compared to each other. Although the EFK approach has less computational effort, the PF approach has been found to perform better in accuracy. Proposed particle filter had minimized pose estimation errors compared to other existing methods with increase in the number of samples and the use of systematic sampling. The experiment was carried out in a well-known indoor environment. We validated our approach with MATLAB by estimating the position and orientation of mobile robot. The core of our work, image processing, object recognition and detection and mathematical calculation on data fusion algorithms was done using MATLAB. MATLAB is a good fit for our research due to its simplicity and flexibility with embedded toolboxes and tutorial videos. The investigations conducted, as well as the findings presented in this thesis, thus form a cogent, concise and well-coordinated response to many glitches on localisation in mobile robot.

5.1.1 Summary of contributions

The major contributions of the study can be summarized as follows:

1. A detailed survey of recent literature, challenges and techniques was researched on in mobile robot. This study formed part of the survey paper that was published in IEEE Access Journal which contributed to the community of research. The purpose of the literature was to unravel some of the problem's autonomous mobile robot encounters and those factors that could degrade their performances. An extensive study was investigated on research gaps, issues associated with mobile robot and the suggestion and how to tackle the problem were presented. The study was able to present more information on the relevance, strengths and weaknesses on some of the method suggested to tackle challenges of mobile robot. The survey was also able to provide concise background information and the importance of each method through comparison.
2. To establish that the selection of components, hardware, software and techniques meet the goals of the proposed project. With the in-depth knowledge about the problem and the approaches that works best for the applications, this study ensured that the most appropriate devices and schemes was used and applied. For example, using GPS and magnetometer are considered inappropriate for indoor localisation because of the presence of objects that could attenuate signals and thus cause inaccuracies in measurements.

3. Information is perceived using sensors and other related devices. The combination of sensors and sources of information does require effective methodology. Therefore, linear and nonlinear algorithms are considered to estimate the position and orientation of the mobile robot. IMU alone is not sufficient due to noise and environmental errors therefore, vision was considered to be more appropriate source combine and localise robot. Computer vision was developed using a suitable algorithm to detect and match features in an environment to overcome challenges of IMU and aid pose estimation. A nonlinear data fusion algorithm, extended Kalman filter was used to fuse data collected from sources to improve the effectiveness of positioning performance. This aspect show case the use of low-cost devices estimates good accuracy of localisation. This contribution formed a part of the published journal in MDPI.
4. To establish that use of natural landmarks with homographies can also be used to determine the location a mobile robot in a known environment. This formed part of a published conference paper. The method achieved good results in certain scenes.
5. To propose a novel method which is based on particle filter to fuse data from IMU and vision precisely using image processing method to improve accuracy of localisation. The goal of image processing is to obtain numerical information from the image which provides a robust description of the object in the scene. The proposed method reduces the average errors when compared with other existing methods. The approach takes advantage of fast response and accurate measurements.

5.2 FUTURE RESEARCH WORK

Due to the increasing demand and diver services required by mobile robot, quite a lot of work to be done are highlighted in Chapter 2. Location information is extremely important for robot and this has been a key challenge in the field of mobile robotics. To have a positioning system low in cost and still provide satisfactory positioning accuracy requires more research investigation and the experiment should be performed on a large scale. Object detection and identification is another area of interest that necessitates further research. Information of objects could be identified in real time video using binocular whereby the two cameras are synchronized to capture more information about the environment through which more features can be extracted. Techniques on how features or information are extracted from sensors/devices demand more future work and further study should concentrate on the use of decision fusion, deep learning [246] or hybrid fusion method as a scheme to ensure robustness and improve pose estimation accuracy in addressing the issue of noisy and error measurements. 3D indoor environmental modelling is another aspect that requires further consideration of research.

REFERENCES

[1] M. B. Alatise and G. P. Hancke, "A review on challenges of autonomous mobile robot and sensor fusion methods," *IEEE Access*, vol. 8, pp. 39830-39846, Feb. 2020.

[2] S. Kumar, "Mobile robotics market forecast, trends to 2020," Grand View Research, Inc. California, United States, Nov. 2014. [Online]. Available: https://sites.google.com/site/electronicindustryhighlights/mobile-robotics-market.

[3] C. Supare and S. D. Adlinge, "Control of mobile robot using visual feedback and wireless communication," in *Proceedings of the 3rd International Conference for Convergence in Technology (I2CT)*, April 2018, pp. 1-5.

[4] Y. Kim, J. Kwak, D. Hong, J. Ahn, S. Wee and J. An, "Localization strategy based on multi-robot collaboration for indoor service robot applications," in *Proceedings of the 10th International Conference on Ubiquitous Robots Ambient Intelligence (URAI)*, Nov. 2013, pp. 225-226.

[5] W. Burgard, M. Moors, D. Fox, R. Simmons and S. Thrun, "Collaborative multi-robot exploration," in *Proceedings of the ICRA Millennium Conference. IEEE International Conference on Robotics and Automation Symposia Proceedings*, vol. 1, April 2000, pp. 476-481.

[6] P. Benavidez and M. Jamshidi, "Mobile robot navigation and target tracking system," in *Proceedings of the 6th International Conference on System of Systems Engineering*, June 2011, pp. 299-304.

[7] N. Kwok and K. Sam, "Path planning for mobile robot localisation and mapping," in *Proceedings of the 30th Annual Conference of IEEE Industrial Electronics Society (IECON)*, Nov. 2004, pp. 603-608.

[8] R. Lagisetty, N. K. Philip, R. Padhi and M. S. Bhat, "Object detection and obstacle avoidance for mobile robot using stereo camera," in *Proceedings of the IEEE International Conference on Control Applications (CCA)*, Aug. 2013, pp. 605-610.

[9] J. Röwekämper, C. Sprunk, G. D. Tipaldi, C. Stachniss, P. Pfaff and W. Burgard, "On the position accuracy of mobile robot localization based on particle filters combined with scan matching," in *Proceedings of the IEEE/RSJ International Conference on Intelligence Robots and Systems*, Oct. 2012, pp. 3158-3164.

[10] A. Alomari, W. Phillips, N. Aslam and F. Comeau, "Dynamic fuzzy-logic based path planning for mobility-assisted localization in wireless sensor networks," *Sensors*, vol. 17, no. 8, pp. 1-26, Aug. 2017.

[11] G. Dudek and M. A. Jenkin, "Inertial sensors, GPS and odometry," *Springer Handbook of Robotics*, B. Siciliano and O. Khatib, Eds., Berlin, Germany: Springer, 2008, pp. 446-490.

[12] J. Shen, D. Tick and N. Gans "Localization through fusion of discrete and continuous epipolar geometry with wheel and IMU odometry," in *Proceedings of the American Control Conference*, July 2011, pp. 1292-1298.

[13] E. N. G. Weng, R. U. Khan, S. A. Z. Adruce and O. Y. Bee, "Objects tracking from natural features in mobile augmented reality," *Procedia - Social and Behavioral Sciences,* vol. 97, pp. 753-760, Nov. 2013.

[14] J. Li, J. A. Besada, A. M. Bernardos, P. Tarrio and J. R Casar, "A novel system for object pose estimation using fused vision and inertial data," *Information Fusion,* vol. 33, pp. 15-28, Jan. 2017.

[15] C. H. Potter, G. P. Hancke and B. J. Silva, "Machine-to-Machine: Possible applications in industrial networks," in *Proceedings of the IEEE International Conference on Industrial Technology (ICIT),* Feb. 2013, pp. 1321-1326.

[16] C. A. Opperman and G. P. Hancke, "Using NFC-enabled phones for remote data acquisition and digital control," in *Proceedings of IEEE AFRICON*, Sept. 2011, pp. 1-6.

[17] A. Kumar and G. P. Hancke, "An energy-efficient smart comfort sensing system based on the IEEE 1451 standard for green buildings, " *IEEE Sensors Journal,* vol. 14, no. 12, pp. 4245-4252, Dec. 2014.

[18] B. Silva, R. M. Fisher, A. Kumar and G. P. Hancke, "Experimental link quality characterization of wireless sensor networks for underground monitoring," *IEEE Transactions on Industrial Informatics,* vol. 11, no. 5, pp. 1099-1110, Oct. 2015.

[19] C. P. Kruger, A. M. Abu-Mahfouz and G. P. Hancke, "Rapid prototyping of a wireless sensor network gateway for the internet of things using off-the-shelf components," in *Proceedings of the IEEE International Conference on Industrial Technology (ICIT)*, March 2015, pp. 1926-1931.

[20] K. S. E. Phala, A. Kumar and G. P. Hancke, "Air quality monitoring system based on ISO/IEC/IEEE 21451 standards," *IEEE Sensors Journal*, vol. 16, no. 12, pp. 5037-5045, April 2016.

[21] B. Cheng, L. Cui, W. Jia, W. Zhao and G. P. Hancke, "Multiple region of interest coverage in camera sensor networks for tele-intensive care units," *IEEE Transactions on Industrial Informatics*, vol. 12, no. 6, pp. 2331-2341, May 2016.

[22] R. Fisher, L. Ledwaba, G. P. Hancke and C. Kruger, "Open hardware: A role to play in wireless sensor networks?," *Sensors*, vol. 15, no. 3, pp. 6818-6844, March 2015.

[23] D. S. A. Ben-Afia, L. Deambrogio, A-C Esher, C. Macabiau, L Soulier V and Gay-Bellile, "Review and classification of vision-based localisation techniques in unknown environments," *IET Radar, Sonar and Navigation*, vol. 8, no. 9, pp. 1059-1072, July 2014.

[24] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91-110, Nov. 2004.

[25] G. Dudek and M. Jenkin, "Inertial sensors, GPS, and odometry," *Springer Handbook of Robotics*, B. Siciliano and O. Khatib, Ed., Berlin, Germany: Springer, 2008, pp. 477-490.

[26] H. Hu and J. Q. Gan, "Sensors and data fusion algorithms in mobile robotics," Department of Computer Science, University of Essex, Colchester, U. K., Technical Report CSM-422, pp. 1-12, Jan. 2005.

[27] Q. K. Dang and Y. S. Suh, "Trajectory generation of a mobile robot using inertial sensors and vision," in *Proceedings of the IEEE/SICE International Symposium on System Integration (SII)*, Dec. 2011, pp. 603-608.

[28] M. S. Ganeshmurthy and G. R. Suresh, "Path planning algorithm for autonomous mobile robot in dynamic environment," in *Proceedings of the 3rd International Conference on Signal Processing, Communication and Networking (ICSCN)*, March 2015, pp. 1-6.

[29] D. D. Paola, A. Milella, G. Cicirelli and A. Distante, "An autonomous mobile robotic system for surveillance of indoor environments," *International Journal of Advanced Robotic Systems,* vol. 7, no. 1, pp. 019-026, March 2010.

[30] V. M. Murthy, S. Kumar, V. Singh, N. Kumar and C. Sain, "Autonomous mobile robots designing," *Journal Global Research in Computer Science*, vol. 2, no. 4, pp. 126-129, April 2011.

[31] M. Köseoğlu, O. M. Çelik and Ö. Pektaş, "Design of an autonomous mobile robot based on ROS," in *Proceedings of the International Artificial Intelligence and Data Processing Symposium (IDAP)*, Nov. 2017, pp. 1-5.

[32] D. G. Lee, H. Kim, K. Jeong and T. Seo, "Development of a robotic platform for amphibious locomotion on ground and water surfaces," in *Proceedings of the 11th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI)*, Nov. 2014, pp. 417-418.

[33] W. L. J. Chen, Y. Wang and J. Guo, "Fusion of inertial and vision data for accurate tracking," in *Proceeding of the 4th International Conference on Machine and Vision, Image Processing, Pattern Analysis*, vol. 8349, Jan. 2012.

[34] F. Rubio, F. Valero and C. Llopis-Albert, "A review of mobile robots: Concepts, methods, theoretical framework, and applications," *International Journal of Advanced Robotic Systems*, vol. 16, no. 2, April 2019, Art. no. 172988141983959.

[35] A. Saudabayev, F. Kungozhin, D. Nurseitov and H. A. Varol, "Locomotion strategy selection for a hybrid mobile robot using time of flight depth sensor," *Journal of Sensors*, vol. 2015, pp.1- 14, March 2015.

[36] R. Siegwart, I. R. Nourbakhsh and D. Scaramuzza, *Introduction to Autonomous Mobile Robot*, 2nd ed. Cambridge, MA, USA: MIT Press, 2004, pp. 1-272.

[37] G. Endo and S. Hirose, "Study on roller-walker (multi-mode steering control and self-contained locomotion)," in *Proceedings of ICRA. Millennium Conference. IEEE International Conference on Robotics and Automation Symposia Proceedings*, vol. 3, April 2000, pp. 2808-2814.

[38] Y. Jin-xia, C. Zi-xing, D. Zhuo-hua and Z. Xiao-bing, "Design of dead reckoning system for mobile robot," *Journal of Central South University of Technology*, vol. 13, no. 5, pp. 542-547, Oct. 2006.

[39] D. Kortenkamp, "Perception for mobile robot navigation: A survey of the state of the art," NASA, Houston, TX, USA, Tech. Rep. 576344779, May 1994, pp. 446-453.

[40] M. Gilmartin, "Introduction to autonomous mobile robots," *Robotica*, vol. 23, pp. 271-272, March 2005.

[41] L. Tang and S. Yuta, "Indoor navigation for mobile robots using memorized omni-directional images and robot's motion," in *Proceedings of the IEEE/RSJ International Conference on Intelligence Robots and Systems*, Oct. 2002, pp. 269-274.

[42] G. Mester, "Motion control of wheeled mobile robot," in *Proceedings of the 4th Serbian-Hungarian Joint Symposium on Intelligent Systems*, April 2006, pp. 119-130.

[43] G. Marek and Š. Peter, "Design the Robot as Security System in the Home," *Procedia Engineering,* vol. 96, pp. 126-130, Jan. 2014.

[44] K. Berns and S. A. Mehdi, "Use of an autonomous mobile robot for elderly care," in *Proceedings of Advanced Technologies for Enhancing Quality Life,* Dec. 2010, pp. 121-126.

[45] B. Crnokic, T. Volaric and M. Grubisic, "Different applications of mobile robots in education," *International Journal on Integrating Technology in Education (IJITE),* vol. 6, no. 3, pp. 15-28, Sep. 2017.

[46] P. Corke, J. Roberts, J. Cunningham and D. Hainsworth, *Handbook of Robotics*. Berlin, Germany: Springer, 2008, pp. 1127-1150.

[47] A. Babinec, L. Jurisica, P. Hubinsky and F. Duchon, " Visual localization of mobile robot using artificial markers," *Procedia Engineering*, vol. 96, pp. 1-9, Dec. 2014.

[48] R. Barber, J. Crespo, C. Gomez, A. C. Hernamdez and M. Galli, *Mobile Robot Navigation in Indoor Environments: Geometric, Topological, and Semantic Navigation*. London, U. K.: IntechOpen, Nov. 2018, pp. 1-25.

[49] J J. Rodrigues, C. Cardeira, F. Carreira, J. M. F. Calado and P. Oliveira, "A Bayesian grid method PCA-based for mobile robots localization in unstructured environments," in *Proceedings of the 16th International Conference on Advanced Robotics (ICAR),* Nov. 2013, pp. 1-6.

[50] F F. Shamsfakhr, B. S. Bigham and A. Mohannadi, "Indoor mobile robot localization in dynamic and cluttered environments using artificial landmarks," *Engineering Computations*, vol. 36, no. 2, pp. 400-419, March 2019.

[51] D. Fox and W. Burgard, "Markov Localization for mobile robots in dynamic environments," *Journal Artificial Intelligence Research*, vol. 11, pp. 391-427, Nov. 1999.

[52] K. Konolige and K. Chou, "Markov localization using correlation," in *Proceedings of the 16th International Joint Conference on Artificial Intelligence*, vol. 2, Aug. 1999, pp. 1154-1159.

[53] J. A. Cetto and A. Sanfeliu, "Topological map learning for a mobile robot in indoor environments," in *Proceedings of the Spanish Symposium on Pattern Recognition Image Analysis (SNRFAI)*, Jan. 2001, pp. 221-226.

[54] J. Crespo, R. Barber and O. M Mozos, "Relational model for robotic semantic navigation in indoor environments," *Journal of Intelligent and Robotic Systems*, vol. 86, nos. 3-4, pp. 617-639, Jan. 2017.

[55] N. Mitsou *et al*., "Online semantic mapping of urban environments," in *Proceedings of the International Conference on Spatial Cognition*, Aug. 2012, pp. 54-73.

[56] L. Kostaveli and A. Gasteratos, "Semantic mapping for mobile robotics tasks: A survey," *Robotics and Autonomous Systems*, vol. 66, pp. 86-103, April 2015.

[57] A. C. Murtra, E. Trulls, J. M. M. Tur and A. Sanfeliu, "Efficient use of 3D environment models for mobile robot simulation and localization," in *Proceedings of the International Conference on Simulation, Modeling, and Programming for Autonomous Robots (SIMPAR)*, vol. 6472, 2010, pp. 461-472.

[58] E. Nelson, M. Corah and N. Michael, "Environment model adaptation for mobile robot exploration," *Autonomous Robots*, vol. 42, no. 2, pp. 257-272, Nov. 2017.

[59] J. Borenstein, H. R. Everett and L. Feng, *Navigating Mobile robots: Systems and Techniques*, Wellesley, MA, USA: A. K. Peters, 1996.

[60] S. Persa and P. P. Jonker, "Real-time computer vision system for mobile robot," *Intelligent Robots and Computer Vision XX: Algorithms, Techniques and Active Vision*, Oct. 2001, pp. 1-10.

[61] P. Goel, S. I. Roumeliotis and G. S. Sukhatme, "Robust localization using relative and absolute position estimates," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, Oct. 1999, pp. 1134-1140.

[62] F. Kong, Y Cheng. J. Xie, G. Zhang and Z. Zhou, "Mobile robot localization based on extended Kalman filter," in *Proceedings of the 6th World Congress on Intelligent Control and Automation*, Oct. 2006, pp. 9242-9246.

[63] A. Koubaa *et al.*, "Introduction to mobile robot path planning," in *Proceedings of Robot Path Planning and Cooperation, Studies in Computational Intelligence*, Jan. 2018, pp. 3-12.

[64] M. A. Hossain and I. Ferdous, "Autonomous robot path planning in dynamic environment using a new optimization technique inspired by bacterial foraging technique," in *Proceedings of the International Conference on Electrical Information and Communication Technlogy (EICT),* vol. 64, pp. 137-141, Feb. 2015.

[65] J-H. Park and U-Y. Huh, "Path planning for autonomous mobile robot based on safe space," *Journal Electrical Engineering and Technology*, vol. 11, no. 5, pp. 1441-1448, Sep. 2016.

[66] D. Xin, C. Hua-hua and G. Wei-kang, "Neural network and genetic algorithm based global path planning in a static environment," *Journal of Zhejiang University.-Science A*, vol. 6, no. 6, pp. 549-554, June 2005.

[67] A. Reshamwala and D. Vinchurkar, "Robot path planning using an ant colony optimization approach: A survey," *Internal Journal of Advanced Research Artificial Intelligence*, vol. 2, no. 3, pp. 65-71, 2013.

[68] F. K. Purian and E. Sadeghian, "Mobile robots path planning using ant colony optimization and fuzzy logic algorithms in unknown dynamic environments," in *Proceedings of the International Conference on Control, Automation Robotics and Embedded Systems (CARE)*, Dec. 2013, pp. 1-6.

[69] F. Lingelbach, "Path planning using probabilistic cell decomposition," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, vol. 1, July 2004, pp. 467-472.

[70] V. Aenugu and P-Y. Woo, "Mobile Robot path planning with randomly moving obstacles and goal," *International Journal Intelligent Systems and. Applications*, vol. 4, no. 2, pp. 1-15, March 2012.

[71] J. Chakraborty, A. Konar, U. K. Chakraborty and L. C. Jain, "Distributed cooperative multi-robot path planning using differential evolution," in *Proceedings of the IEEE Congress on Evolutionary Computation (IEEE World Congress on Computational Intelligence)*, June 2008, pp. 718-725.

[72] D. R. Parhi and S. Kundu, "Navigational control of underwater mobile robot using dynamic differential evolution approach," in *Proceeding of the Institution of Mechanical Engineers, Part M: Journal of Engineering for the Maritime Environment*, vol. 231, no. 1, Aug. 2016, pp. 284-301.

[73] S. Kundu and D. R. Parhi, "Navigation of underwater robot based on dynamically adaptive harmony search algorithm," *Memetic Computing*, vol. 8, no. 2, pp. 125-146, April 2016.

[74] G-G. Wang, H. E. Chu and S. Mirjalili, "Three-dimensional path planning for UCAV using an improved bat algorithm," in *Aerospace Science and Technology*, vol. 49, pp. 231-238, Feb. 2016.

[75] A. Sengupta, T. Chakraborti, A. Konar and A. Nagar, "Energy efficient trajectory planning by a robot arm using invasive weed optimization technique," in *Proceedings of the 3rd World Congress on Nature and Biological Inspired Computing*, Oct. 2011, pp. 311-316.

[76] M. Mouad, L. Adouane, D. Khadraoui and P. Martinet, "Mobile robot navigation and obstacles avoidance based on planning and re-planning algorithm," *IFAC Proceedings Volumes*, vol. 45, no. 22, pp. 622-628, Jan. 2012.

[77] K. Shahzad, S. Iqbal and P. Bloodsworth, "Points-based safe path planning of continuum robots," *International Journal of Advanced Robotic Systems*, vol. 12, no. 7, pp. 1-12, July 2015.

[78] W. J. Yim and J. B. Park, "Analysis of mobile robot navigation using vector field histogram according to the number of sectors, the robot speed and the width of the path," in *Proceedings of the 14th International Conference on Control, Automation and Systems (ICCAS)*, Oct. 2014, pp. 1037-1040.

[79] G. Li, A. Yamashita, H. Asama and Y. Tamura, "An efficient improved artificial potential field based regression search method for robot path planning," in *Proceedings of the IEEE International Conference on Mechatronics and Automation*, Aug. 2012, pp. 1227-1232.

[80] A. Sgorbissa and R. Zaccaria, "Planning and obstacle avoidance in mobile robotics," *Robotics and Autonomous Systems*, vol. 60, no. 4, pp. 628–638, April 2012.

[81] Y. Zhu, T. Zhang, J. Song, and X. Li, "A new hybrid navigation algorithm for mobile robots in environments with incomplete knowledge," *Knowledge-Based Systems*, vol. 27, pp. 302-313, March 2012.

[82] K. Kumar, A. Varghese, P. Reddy and N. Narendra, "An improved tracking using IMU and uision fusion for mobile augemented reality applications," *International Journal of Multimedia and its Applications (IJMA)*, vol. 6, no. 5, pp. 13-29, Oct. 2014.

[83] A. T. Erdem and A. O. Ercan, "Fusing inertial sensor data in an extended Kalman filter for 3D camera tracking," *IEEE Transactions on Image Processing*, vol. 24, no. 2, pp. 538-548, Feb. 2015.

[84] S. Borik, B. Babusiak and I. Cap, "Device for accelerometer and gyroscope measurements," *Information Technologies in Biomedicine*, vol. 4, pp. 139-146, Jan. 2014.

[85] C. Hsu and C. Yu, "An accelerometer based approach for indoor localization," in *Proceedings of the Symposia and Workshops on Ubiquitous, Autonomic Trusted Computing*, Nov. 2009, pp. 223-227.

[86] J J. Bird and D. Arden, "Indoor navigation with foot-mounted strapdown inertial navigation and magnetic sensors [emerging opportunities for localization and tracking]," *IEEE Wireless Communications*, vol. 18, no.2, pp. 28-35, April 2011.

[87] A. K. Nasir, C. Hille and H. Roth, "Data fusion of stereo vision and gyroscope for estimation of indoor mobile robot orientation," *IFAC Proceedings Volumes,* vol. 45, no. 4, pp. 163-168, Jan. 2012.

[88] F. N. Sibai, H. Trigui, P. C. Zanini and A. R. Al-Odail, "Evaluation of indoor mobile robot localization techniques," in *Proceedings of the International Conference on Computer Systems and Industrial Informatics*, Dec. 2012, pp. 1-6.

[89] N. Ahmad, R. A. Raja Ghazilla, N. Khairi and V. Kasi, "Reviews on various inertial measurement unit (IMU) sensor applications," *International Journal of Signal Processing Systems*, vol. 1, no. 2, pp. 256-262, Dec. 2013.

[90] A. M. Sabatini, "Estimating three-dimensional orientation of human body parts by inertial/magnetic sensing," *Sensors*, vol. 11, no. 2, pp. 1489-1525, Jan. 2011.

[91] A. Olivares, G. Olivares, J. M. Górriz and J. Ramírez, "High-efficiency low-cost accelerometer-aided gyroscope calibration," in *Proceedings of the International Conference on Test and Measurement*, Dec. 2009, pp. 354-360.

[92] V. Avrutov, M. D. Geraimchuk and X. Xiangming, "3D-calibration for IMU of the strapdown inertial navigation systems," *MATEC Web of Conferences*, vol. 114, July 2017, pp. 1-7.

[93] C.-R. Lee, J. H. Yoon and K.-J. Yoon, "Calibration and noise identification of a rolling shutter camera and a low-cost inertial measurement unit," *Sensors*, vol. 18, no. 7, pp. 1-24, July 2018.

[94] C.-M. Dong, S.-Q. Ren, X.-J. Chen and Z.-H. Wang, "A separated calibration method for inertial measurement units mounted on three-axis turntables, " *Sensors*, vol. 18, no. 9, pp. 1-14, Aug. 2018.

[95] D. Tedaldi, A. Pretto and E. Menegatti, "A robust and easy to implement method for IMU calibration without external equipments," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, June 2014, pp. 3042-3049.

[96] F. Sun, J. Yu and D. Xu, "Visual measurement and control for underwater robots: A survey," in *Proceedings of the 25th Chinese Control and Decision Conference (CCDC)*, May 2013, pp. 333-338.

[97] C. C. Liu, Z. M. Zhang and E. F. Sang, "A novel acoustical vision system design for automated underwater vehicles, ," in *Proceedings of the 7th World Congress on Intelligent Control Automation*, June 2008 pp. 7438-7443.

[98] Intel, "Intel Curie Module: Revision 1.3," Intel Curie datasheet, March 2017.

[99] L. Song, W. Wu, J. Guo and X. Li, "Survey on camera calibration technique," in *Proceedings of the 5th International Conference on Intellengent Human-Machine Systems and Cybernetics*, Aug. 2013, pp. 389-392.

[100] S. De Ma, "A self-calibration technique for active vision systems," *IEEE Transactions on Robotics Automation*, vol. 12, pp. 114-120, Feb. 1996.

[101] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 11, pp. 1330-1334, Nov. 2000.

[102] J.-Y. Bouguet, "Camera calibration toolbox for MATLAB," Jan. 2010. [Online]. Available: http://www.vision.caltech.edu/bouguetj/calib_doc. [Accessed: March 21, 2017]

[103] J. Kannala and S. S. Brandt, "A generic camera model and calibration method for conventional, wide-angle, and fish-eye lenses," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, pp. 1335-1340, Aug. 2006.

[104] Z. Wang, W. Wu, X. Xu and D. Xue, “Recognition and location of the internal corners of planar checkerboard calibration pattern image,” *Applied Mathematics and Computation,* vol. 185, no.2, pp. 894-906, Feb. 2007.

[105] Y. Yang, X. Meng and M. Gao, “Vision system of mobile robot combining binocular and depth cameras,” *Journal of Sensors*, vol. 2017, pp. 1-11, Sep. 2017.

[106] Y. Genc, S. Riedel, F. Souvannavong, C. Akinlar and N. Navab, “Marker-less tracking for AR: A learning-based approach,” in *Proceedings of the International Symposium on Mixed and Augmented Reality (ISMAR '02)*, Oct. 2002, pp. 1-10.

[107] A. Ben-Afia *et al.*, “Review and classification of vision-based localisation techniques in unknown environments,” *IET Radar, Sonar and Navigation,* vol. 8, no. 9, pp. 1059-1072, Dec. 2004.

[108] C. Wang, T. Wang, J. Liang, Y. Chen and Y. Wu, “Monocular vision and IMU based navigation for a small unmanned helicopter,” in *Proceedings of the 7th IEEE Conference on Industrial Electronics and Applications (ICIEA),* July 2012, pp. 1694-1699.

[109] L. Guanghui and J. Zhijian, "An artificial landmark design based on mobile robot localization and navigation,” in *Proceedings of the 4th International Conference on Intelligent Computation Technology and Automation*, March 2011, pp. 588-591.

[110] J. Seong, J. Kim and W. Chung, “Mobile robot localization using indistinguishable artificial landmarks,” in *Proceedings of the 10th International Conference on Ubiquitous Robots Ambient Intelligence*, Oct. 2013, pp. 222-224.

[111] W. Hong, H. Xia, X. An and X. Liu, “Natural landmarks based localization algorithm for indoor robot with binocular vision,” in *Proceedings of the 29th Chinese Control and Decision Conference (CCDC)*, May 2017, pp. 3313-3318.

[112] J. Liu, M. Wan and J. Zhang, “Monocular robot navigation using invariant natural features,” in *Proceedings of the 7th World Congress on Intelligent Control and Automation*, July 2008, pp. 5733-5738.

[113] H. Bay, A. Ess, T. Tuytelaars and L. V. Gool, “Speed-up robust feature (SURF),” *Computer Vision and Image Understanding,* vol. 110, no. 3, pp. 346-359, June 2008.

[114] M. Alatise and G. P. Hancke, “Pose estimation of a mobile robot using monocular vision and inertial sensors data,” in *Proceedings of IEEE AFRICON*, Sep. 2017, pp. 1552-1557.

[115] M. Ozuysal, P. Fua and V. Lepetit, “Fast keypoint recognition in ten lines of code,” in *Proceedings of the IEEE International Conference on Computer Vision and Pattern Recognition,* June 2007, pp. 1-8.

[116] P. Loncomilla, J. Ruiz-del-Solar and L. Martinez, “Object recognition using local invariant features for robotic applications: A survey,” *Pattern Recognition,* vol. 60, pp. 499-514, Dec. 2016.

[117] J. Farooq, “Object detection and identification using SURF and BoW model,” in *Proceedings of the International Conference on Computing, Electronic and Electrical Engineering (ICE Cube),* April 2016, pp. 318-323.

[118] J. D. Hol, T. B. Schon and F. Gustafsson, “Modeling and calibration of inertial and vision sensors,” *International Journal of Robotics Research,* vol. 29, no. 2-3, pp. 231-244, Jan. 2010.

[119] C. Harris and M. Stephens, “A combined corner and edge detector,” in *Proceedings of the 4th Alvey Vision Conference,* May 1988, pp. 147-151.

[120] C. Tomasi and T. Kanade, “Shape and motion from image streams: a factorization method,” in *Proceedings of the National Academy of Sciences of the United State of America,* vol. 90, no. 2, pp. 9795-9802, Nov. 1993.

[121] K. Mikolajczyk and C. Schmid, “Scale and affine Invariant interest point detectors,” *International Journal of Computer Vision,* vol. 60, no. 1, pp. 63-86, Jan. 2004.

[122] M. A. Fischler and R. C. Bolles, “Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography,” *Communications of ACM,* vol. 24, no.6, pp. 381-395, June 1981.

[123] E. Serradell, M. Ozuysal, V. Lepetit, P. Fua and F. Moreno-Noquer, “Combining geometric and appearance prioris for robust homography estimation,” in *Proceedings of the 11th European Conference on Computer Vision,* Sep. 2010, pp. 58-72.

[124] S. Fu, “Homography estimation from planar contours in image sequence,” *Optical Engineering,* vol. 49, no. 3, March 2010, Art. no. 037202.

[125] C-M. Cheng and S-H. Lai, “A consensus sampling technique for fast and robust model fitting,” *Pattern Recognition,* vol. 42, no. 7, pp. 1318-1329, July 2009.

[126] P. H. S Torr and D. W. Murray, “The development and comparison of robust methods for estimating the fundamental matrix,” *International Journal of Computer Vision,* vol. 24, pp. 271-300, July 1997.

[127] C-S. Chen, Y-P Hung and J-B. Cheng, “RANSAC-based DARCES: A new approach to fast automatic registration of partially range images,” *IEEE Transaction on Pattern Analysis and Machine Intelligence*, vol. 21, no. 11, pp. 1229-1234, Nov. 1999.

[128] D. Gonzalez-Aguilera, P. Rodriguez-Gonzalvez, D. Hermandez-Lopez and J. L. Lerma, “A robust and hierchical approach for the automatic co-registration of intensity and visible images,” *Optics and Laser Technology*, vol. 44, no. 6, pp. 1915-1923, Sep. 2012.

[129] Y. Lv, J. Feng, Z. Li, W. Liu and J. Cao, “A new robust 2D camera calibration method using RANSAC,” *Optik International Journal for Light and Electron Optics,* vol. 126, no. 24, pp. 4910-4915, Dec. 2015.

[130] H. Zhou, T. Zhang and W. Lu, “Vision-based pose estimation from points with unknown correspondences,” *IEEE Transactions on Image Processing*, vol. 23, no. 8, pp. 3468-3477, Aug. 2014.

[131] E. Marchand, H. Uchiyama and F. Spindler “Pose estimation for augmented reality: a hands-on survey,” *IEEE Transactions on Visualization Computer Graphics,* vol. 22, no. 12, pp. 2633-2651, Dec. 2016.

[132] J. Diebel, “Representing attitude: Euler angles, unit quaternions, and rotation vectors, ” *Matrix*, vol. 58, no. 15, pp. 1-35, Jan. 2006.

[133] K.-H. Yang, W.-S. Yu and X.-Q. Ji, “Rotation estimation for mobile robot based on single-axis gyroscope and monocular camera,” *International Journal of Automation and Computing,* vol. 9, pp. 292-298, June 2012.

[134] F. E. White, “Data fusion lexicon,” JDL, Technology Panel For C3, San Diego, CA, USA, Tech. Rep. 144275, Jan. 1991.

[135] D. L. Hall and J. Llinas, “An introduction to multisensor datafusion,” *Proceedings of the IEEE*, vol. 85, no. 1, pp. 6-23, Jan. 1997.

[136] F. Castanedo, “A Review of data fusion techniques,” *Scientific World Journal,* vol. 2013, pp. 1-19, Sep. 2013.

[137] K. S. Nagla, M. Uddin and D. Singh, “Multisensor data fusion and integration for mobile robots: A review,” *IAES International Journal of Robotics and Automation (IJRA)*, vol. 3, no. 2, pp. 131-138, June 2014.

[138] H. F. Durrant-Whyte, “Sensor models and multisensor integration,” The *International Journal of Robotics Research*, vol. 7, no. 6, pp. 97-113, Dec. 1988.

[139] R. C. Luo and M. G. Kay, “A tutorial on multisensor integration and fusion,” in *Proceedings of the 16th Annual Conference of IEEE Industrial Electronics Society (IECON)*, vol. 1, Nov. 1990, pp. 707-722.

[140] F. Coito, A. Eleuterio, S. Valtchev and F. Coito, “Tracking a mobile robot position using vision and inertial sensor,” *IFIP Advances in Information and Communication Technology*, vol. 423, April 2014, pp. 201-208.

[141] J. D. Hol, T. B. Schon, H. Luinge, P. J. Slycke and F. Gustafsson, “Robust real-time tracking by fusing measurements from inertial and vision sensors,” *Journal of Real-Time Image Processing*, vol. 2, pp. 149-160, Oct. 2007.

[142] W. Elmenreich, “An introduction to sensor fusion,” Institut für Technische Informatik, Vienna University of Technology, Austria, Research Report 47/2001, vol. 502, pp. 1-28, Nov. 2002.

[143] A. Sinha, H.Chen, D. G. Danu, T. Kirubarajan and M. Farooq, “Estimation and decision fusion: A survey,” *Neurocomputing,* vol. 71, nos. 13-15, pp. 2650-2656, Aug. 2008.

[144] J. Gu, M. Meng, A. Cook and P. X. Liu, “Sensor fusion in mobile robot: some perspectives,” in *Proceedings of the 4th World Congress on Intelligent Control and Automation*, vol. 2, June 2002, pp. 1194-1199.

[145] W. V. Drongelen, “Kalman filter,” in *Signal Processing for Neuroscientist*, 2nd ed. New York, NY, USA: Academic, 2018, ch. 19, pp. 361-374.

[146] T. T. Hoang, P. M. Duong, N. T. Van, D. A. Viet and T. Q. Vinh, “Multi-sensor perceptual system for mobile robot and sensor fusion-based localization,” in *Proceedings of the International Conference on Control, Automation and Information Sciences (ICCAIS)*, Nov. 2012, pp. 259-264.

[147] D. Jeon, H. Choi and J. Kim, “UKF data fusion of odometry and magnetic sensor for a precise indoor localization system of an autonomous vehicle,” in *Proceedings of the 13th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI)*, Aug. 2016, pp. 47-52.

[148] A. Alofi, A. Alghamdi, R. Alahmadi, N. Aljuaid and M. Hemalatha, “A review of data fusion techniques,” *Internatinal Journal of Computer Applications*, vol. 167, pp. 37-41, June 2017.

[149] A. Al-Fuqaha, M. Elbes and A. Rayes, "An intelligent data fusion technique based on the particle filter to perform precise outdoor localization," *International Journal of Pervasive Computing and Communications*, vol. 9, no. 2, pp. 163-183, June 2013.

[150] Y. Ren and X. Ke, "Particle filter data fusion enhancements for MEMS-IMU/GPS," *Intelligent Information Management*, vol. 2, no. 7, pp. 417-421, Jan. 2010.

[151] F. Cappello, R. Sabatini, S. Ramasamy and M. Marino, "Particle filter based multi-sensor data fusion techniques for RPAS navigation and guidance," *IEEE Metrology for Aerospace (MetroAeroSpace)*, June 2015, pp. 395-400.

[152] W. Li, Z. Wang, Y. Yuan and L. Guo, "Particle filtering with applications in networked systems: A survey," *Complex and Intelligent Systems*, vol. 2, no. 4, pp. 293-315, Oct. 2016.

[153] G. Texier, R. S. Allodji, L. Diop, J-B. Meynard, L. Pellegrin and H. Chaudet, "Using decision fusion methods to improve outbreak detection in disease surveillance," *BMC Medical Informatics and Decision Making*, vol. 19, no. 1, pp. 1-11, March 2019.

[154] S. Vechet and J. Krejsa, "Sensors data fusion via Bayesian network," *Recent Advances in Mechatronics*, T. Brezina and R. Jablonski, Eds., Berlin Germany: Springer, 2010, pp. 221-226.

[155] X. Liu, H. Leung, P. Valin and E. Bosse, "Multisensor joint tracking and identification using particle filter and Dempster-Shafer fusion," in *Proceedings of the 15th International Conference on Information Fusion*, July 2012, pp. 902-909.

[156] Y. Zhang, Q-A. Zeng, Y. Liu and B. Shen, "Integrated data fusion using Dempster-Shafer theory," in *Proceedings of the 1st International Conference on Computational Intelligence Theory, Systems and Applications (CCITSA)*, Dec. 2015, pp. 98-103.

[157] M. O. Oloyede and G. P. Hancke, "Unimodal and multimodal biometric sensing systems: A review," *IEEE Access*, vol. 4, pp. 7532-7555, Sep. 2016.

[158] N. Sghaier, R. B. Ayed, R. B. Marzoug and A. Rebai, "Dempster-Shafer theory for the prediction of auxin-response elements (AuxREs) in plant genomes," *BioMed Research International*, vol. 2018, pp. 1-13, Nov. 2018.

[159] K. Khoshelham, S. Nedkov and C. Nardinocchi, "A comparison of Bayesian and evidence-based fusion methods for automated building detection in aerial data," *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 37, pp. 1183-1188, Dec. 2008.

[160] A. Mahmood, A. Baig and Q. Ahsan, "Real time localization of mobile robotic platform via fusion of inertial and visual navigation system,"in *Proceedings of the International Conference of Robotics and Artificial Intelligence*, Oct. 2012, pp. 40-44.

[161] M. Faisal, M. Alsulaiman, R. Hedjar et al., "Enhancement of mobile robot localization using extended Kalman filter", *Advances in Mechanical Engineering*, vol., 8 no. 11, pp. 1–11, Nov. 2016.

[162] A. Sousa et al., "Localization of a mobile robot based in odometry and natural landmarks using extended Kalman filter", in *Proceedings of the 5th International Conference on Informatics, in control, Automation and Robotics*, May 2008, pp. 1-7.

[163] J. Z. Sasiadek and P. Hartana, "Odometry and sonar data fusion for mobile robot navigation," in *Proceedings of IFAC Volumes*, vol. 33, no. 27, pp. 411-416, Sep. 2000.

[164] T. T. Hoang, P. M. Duong, N. T. T. Van, D. A. Viet and T. Q. Vinh, "Multi-sensor perceptual system for mobile robot and sensor fusion-based localization", *in Proceedings of International Conference on Control, Automation and Information Sciences (ICCAIS)*, Nov. 2012, pp. 259-264

[165] Y. Raaj, A. John and T. Jin, "3D object localization using forward looking sonar (FLS) and optical camera via particle filter based calibration and fusion," in *Proceedings of OCEANS 2016 MTS/IEEE Monterey*, Sep. 2016, pp. 1-10.

[166] W. Ma, Y. Zhang, Q. Zhao and T. Liu, "Two particle filter-based INS/LiDAR-integrated mobile robot localization," in *Proceedings of International Conference on Multimedia Technology and Enhanced Learning*, June 2020, pp. 349-358.

[167] S. Jain, S. Nandy, R. Ray and S. N. Shome, "Application of particle filtering technique for sensor fusion in mobile robotics," in *Proceeding of IEEE International Conference on Mechatronics and Automation*, Aug. 2011, pp. 2285-2290.

[168] Y. Ren and X. Ke, "Particle filter data fusion enhancements for MEMS-IMU/GPS," *Intelligent Information Management*, vol. 2, no. 7, pp. 417-421, Jan. 2010.

[169] Y. Lee, B. Yim, and J. Song, "Mobile robot localization based on effective combination of vision and range sensors," *International Journal of Control, Automation, and Systems*, vol. 7, no. 1, pp. 97-104, March 2009.

[170] W. A. Abdulhafiz and A. Khamis, "Bayesian approach with pre- and post-filtering to handle data uncertainty and inconsistency in mobile robot local positioning," *Journal of Intelligent Systems*, vol. 23, no. 2, pp. 133–154, Jan. 2014.

[171] H. Asoh, Y. Motomura, I. Hara, S. Akaho, S. Hayamizu, and T. Matsui, "Combining probabilistic map and dialog for robust life-long office navigation," in *Proceedings International Conference Intelligent Robots System IROS*, vol. 2, Nov. 1996 pp. 880–885.

[172] H. Zhou and S. Sakane, "Sensor planning for mobile robot localization---A hierarchical approach using a bayesian network and a particle filter," *IEEE Transactions on Robotics*, vol. 24, no. 2, pp. 481-487, April 2008.

[173] L. Vladareanu, G. Tont, V. Vladareanu, F. Smarandache and L. Capitanu, "The navigation mobile robot systems using bayesian approach through the virtual projection method," in *Proceedings of International Conference on Advanced Mechatronic Systems*, Sep. 2012, pp. 498-503.

[174] C. Premebida , D. R. Faria and U. Nunes, "Dynamic bayesian network for semantic place classification in mobile robotics, " *Autonomous Robots*, vol. 41, pp. 1161-1172, July 2016.

[175] S. Soleimanpour, S. S. Ghidary and K. Meshgi, "Sensor fusion in robot localization using DS-evidence theory with conflict detection using mahalanobis distance," in *Proceedings of the 7th IEEE International Conference on Cybernetic Intelligent Systems*, Sep. 2008, pp. 1-6.

[176] H. Lee, X. Song and H. Cho, "Mobile robot self-localization and local map alignment with a dempster shafer sensor fusion algorithm," in *Proceedings of the 17th World Congress. The International Federation of Automatic Control*, July 2008, pp. 14663-14668.

[177] A. Gören, E. Uyar, Ö. Başer and Z. Dicle, "Sensor fusion using dempster-shafer theory of evidence in autonomous robot navigation," *Mechanics, Automatic Control and Robotics*, vol. 6, Nov. 2007, pp. 133 - 144.

[178] M. Valente, C. Joly and A. Fortelle, "Fusing laser scanner and stereo camera in evidential grid maps," *in Proceedings of International Conference on Control, Automation, Robotics and Vision (ICARCV)*, Nov. 2018, pp. 990-997.

[179] J. Carlson and R. R. Murphy, "Use of dempster-shafer conflict metric to detect interpretation inconsistency," in *Proceedings of the 21st Conference on Uncertainty in Artificial Intelligence* July 2012, pp. 1-10.

[180] S. Erfani, A. Jafari and A. Hajiahmad, "Localization and controlling the mobile robot by sensory data fusion," *Agricultural Engineering International: CIGR Journal*, vol. 21, no. 2, pp. 86-97, July, 2019.

[181] W. A. Abdulhafiz and A. Khamis, "Handling data uncertainty and inconsistency using multisensor data fusion," *Advances in Artificial Intelligence*, vol. 2013, no. 11, pp. 1-12, Sep. 2013.

[182] N. Ganganath and H. Leung, "Mobile robot localization using odometry and kinect sensor," in *Proceedings of the IEEE International Conference on Emerging Signal Processing Applications*, Jan. 2012, pp. 91-94.

[183] J. Kelly and G. S. Sukhatme, "Visual-inertial sensor fusion: localization, mapping and Sensor-to-Sensor self-calibration," *International Journal of Robotics Research,* vol. 30, no. 1, pp. 56-79, Jan. 2011.

[184] A. M. Sabatini, "Variable-state-dimension Kalman-based filter for orientation determination using inertial and magnetic sensors," *Sensors*, vol. 12, no. 7, pp. 8491-8506, June 2012.

[185] G. Ligorio and A. M. Sabatini, "Extended Kalman filter-based methods for pose estimation using visual, inertial and magnetic sensors: Comparative analysis and performance evaluation," *Sensors*, vol. 13, no.2, pp. 1919-1941, Feb. 2013.

[186] M. M. Shaikh *et al.*, "Mobile robot vision tracking system using Unscented Kalman Filter," in *Proceedings of the IEEE/SICE International Symposium on System Integration (SII)*, Dec. 2011, pp. 1214-1219.

[187] Y. Tao, H. Hu and H. Zhou, "Integration of vision and inertial sensors for 3D arm motion tracking in home-based rehabilitation," *International Journal of Robotics Research,* vol. 26, no. 6, pp. 607-624, July 2016.

[188] P. Gemeiner, P. Einramhof and M. Vincze, "Simultaneous motion and structure estimation by fusion of inertial and vision data," *International Journal of Robotics Research*, vol. 26, no. 6, pp. 591-605, July 2016.

[189] J. Durrie, T. Gerritsen, E. W. Frew and S. Pledgie, "Vision-aided inertial navigation on an uncertain map using a particle filter," in *Proceedings of the IEEE International Conference on Robotics and Automation*, May 2009, pp. 4189-4194.

[190] Z. Mikulová, F. Duchon, M. Dekan and A. Babinec, "Localization of mobile robot using visual system," *International Journal of Advanced Robotic Systems,* vol. 14, no. 5, Oct. 2017, Art no.172988141773608.

[191] K. Khoshelham and L. Diaz-Vilariño, "3D modelling of interior spaces: Learning the language of indoor architecture," *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 40, no. 5, pp. 321-326, June 2014.

[192] P. Biber, H. Andreasson, T. Duckett and A. Schilling, "3D modeling of indoor environments by a mobile robot with a laser scanner and panoramic camera," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots Systems (IROS)*, vol. 4, Oct. 2004, pp. 3430-3435.

[193] S. Thrun *et al.*, "A real-time expectation-maximization algorithm for acquiring multiplanar maps of indoor environments with mobile robots," *IEEE Transactions on Robotics and Automation*, vol. 20, no. 3, pp. 433-443, June 2004.

[194] S. Kriegel, C. Rink, T. Bodenmuller and M. Suppa, "Efficient next-best-scan planning for autonomous 3D surface reconstruction of unknown objects," *Journal of Real-Time Image Processing*, vol. 10, no. 4, pp. 611-631, Dec. 2013.

[195] X. Chen and Y. Jia, "Indoor localization for mobile robots using lampshade corners as landmarks: Visual system calibration, feature extraction and experiments," *International Journal of Control, Automation and Systems*, vol. 12, no. 6, pp. 1313-1322, Oct. 2014.

[196] G. Klančar, L. Teslić and I. Škrjanc, "Mobile-robot pose estimation and environment mapping using an extended Kalman filter," *International Journal of Systems Science,* vol. 45, no. 12, pp. 2603-2618, Dec. 2014.

[197] T. T. Q. Bui and K.-S. Hong, "A comparison of using probabilistic motion models for mobile robot pose estimation," in *Proceedings of the ICROS-SICE International Joint Conference 2009,* Sep. 2009, pp. 528-532.

[198] N. Kothari, M. Gupta, L. Vachhani and H. Arya, "Pose estimation for an autonomous vehicle using monocular vision, " in *Proceedings of the Indian Control Conference (ICC)*, Jan. 2017, pp. 424-431.

[199] L. F. Lupián, A. Romay and A. Espínola, "Vision based localization of humanoid robots by inverse pose-estimation using a small set of fixed landmark features," in *Proceedings of the IX Latin American Robotics Symposium and IEEE Colombian Conference on Automatic Control,* Oct. 2011, pp. 1-6.

[200] C. López-Franco, J. Gomez-Avila, N. Arana-Daniel and A. Y. Alanis, "Robot pose estimation based on visual information and particle swarm optimization," in *Proceedings of the World Automation Congress (WAC)*, Oct. 2014, pp. 768-773.

[201] A. Kim and M. F. Golnaraghi, "A quaternion-based orientation estimation algorithm using an inertial measurement unit," in *Proceedings of Position Location and Navigation Symposium*, April 2004, pp. 268-272.

[202] A. Foi, S. Alenius, V. Katkovnik and K. Egiazarian, "Noise measurement for raw-data of digital imaging sensors by automatic segmentation of non-uniform targets," I*EEE Sensors Journal*, vol. 7, no. 10, pp. 1456-1461, Oct. 2007.

[203] M. Matejcek and M. Sostronek , "New experience with Allan variance: Noise analysis of accelerometers, " in *Proceedings of Communication and Information Technologies (KIT)*, Oct. 2017, pp. 1-4.

[204] J. Park *et al.*, "High performance vision tracking system for mobile robot using sensor data fusion with Kalman filter," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, Oct. 2010, pp. 3778-3783.

[205] M. B. Alatise and G. P. Hancke, "Pose estimation of a mobile robot based on fusion of IMU data and vision data using an extended Kalman filter", *Sensors*, vol. 17, no. 10, pp. 1-17, Sep. 2017.

[206] N. B. Ahmad, R. A. R. Ghazilla, N. M. Khairi and V. Kasi, "Reviews on various inertial measurement unit (IMU) sensor applications," *International Journal of Signal Processing Systems*, vol. 1, no. 2, pp. 265-262, Dec. 2013.

[207] H. Zhao and Z. Wang, "Motion measurement using inertial sensors, ultrasonic sensors, and magnetometers with extended Kalman filter for data fusion," *IEEE Sensors Journal*, vol. 12, no.5, pp. 943-953, May 2012.

[208] S. Choi, J. H. Joung, W. Yu and J. I. Cho, "What does ground tell us? Monocular visual odometry under planar motion constraint," in *Proceedings of the 11th International Conference Control, Automation and Systems (ICCAS), 2011*, Oct. 2011, pp. 1480-1485.

[209] F. Caballero, L. Merino, J. Ferruz and A. Ollero, "Unmanned aerial vehicle localization based on monocular vision and online mosaicking," *Journal of Intelligent and Robotic Systems*, vol. 55, pp. 323-343, Jan. 2009.

[210] D. F. Dementhon and L. S. Davis, "Model-based object pose in 25 lines of code," *International Journal of Computer Vision*, " vol. 15, pp. 123-141, June 1995.

[211] Q. Zhou and X. Li, "Deep homography estimation and its application to wall maps of wall-climbing robots, " *MDPI Applied Science*, pp. 1-14, July 2019.

[212] J. Westerhoff, "A method and a device for estimating an orientation of a camera relative to a road surface," U.S. patent 3193306, Sep. 2017.

[213] E. Montijano and C. Sagues, "Fast pose estimation for visual navigation using homographies," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, Oct. 2009, pp. 2704-2709.

[214] J. Park, W. Hwang, H. Kwon, K. Kim and D. D. Cho, "A novel line of sight control system for a robot vision tracking system using vision feedback and motion-disturbance feed forward compensation," *Robotica*, vol. 31, no. 1, pp. 99-112, Jan. 2012.

[215] J. Delaune, G. L. Besnerais, T. Voirin, J. L. Farges and C. Bourdarias, "Visual-inertial navigation for pinpoint planetary landing using scale-based landmark matching," *Robotics and Autonomous Systems*, vol. 78, pp. 63-82, April 2016.

[216] G. Bleser and D. Stricker, "Advanced tracking through efficient image processing and visual–inertial sensor fusion," *Computer and Graphics*, vol. 33, pp. 59-72, Feb. 2009.

[217] Y. Miao, Y. Liu, H. Ma and H. Jin, "The pose estimation of mobile robot based on improved point cloud registration," *International Journal of Advanced Robotic Systems*, vol. 13, no. 2, pp. 1-10, March 2016.

[218] E. D. Zubiete *et al.*, "Evaluation of a home biomonitoring autonomous mobile robot," *Computational Intelligence and. Neuroscience*, vol. 2016, pp.1- 8, April 2016.

[219] B. C. Stahl and M. Coeckelbergh, "Ethics of healthcare robotics: Towards responsible research and innovation," *Robotics and Autonomous Systems,* vol. 86, pp. 152-161, Dec. 2016.

[220] R. Mardiyanto, M. F. R. Utomo, D. Purwanto and H. Suryoatmojo, "Development of hand gesture recognition sensor based on accelerometer and gyroscope for controlling arm of underwater remotely operated robot," in *Proceedings of the International Seminar on Intelligent Technology and its Applications (ISITIA)*, Aug. 2017, pp. 329-333.

[221] V. Sazdovski and P. M. G. Silson, "Inertial navigation aided by vision-based simultaneous localization and mapping," *IEEE Sensors Journal*, vol. 11, no. 8, pp. 1646-1656, Jan. 2011.

[222] E. Royer *et al.*, "Outdoor autonomous navigation using monocular vision, " in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots Systems*, Aug. 2005, pp.3395-3400.

[223] E. Royer, M. Lhuillier, M. Dhome and J.-M. Lavest, "Monocular vision for mobile robot localization and autonomous navigation," *International Journal of Computer Vision*, vol. 74, pp. 237-260, Jan. 2007.

[224] L. Cheng, Y. Dai, R. Peng and X. Nong, "Positioning and navigation of mobile robot with asynchronous fusion of binocular vision system and inertial navigation system," *International Journal of Advanced Robotic Systems*, vol. 14, pp. 1-16, Dec. 2017.

[225] D. I. B. Randeniya, S. Sarkar and M. Gunaratne, "Vision–IMU integration using a slow-frame-rate monocular vision system in an actual roadway setting," *IEEE Transactions on Intelligent Transportation Systems*, vol. 11, no. 2, pp. 256-266, Jan. 2010.

[226] D. H. Won, S. Sung and Y. J. Lee, "UKF based vision aided navigation system with low grade IMU," in *Proceedings of ICCAS 2010*, Oct. 2010, pp. 2435-2438.

[227] P. Hoang-Lan, V. Perdereau, B. V. Adorno and P. Fraisse, "Position and orientation control of robot manipulators using dual quaternion feedback," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, Oct. 2010, pp. 658-663.

[228] Taryudi and M. Wang, "3D object pose estimation using stereo vision for object manipulation system," in *Proceedings of the International Conference on Applied System Innovation (ICSAI)*, May 2017, pp. 1532-1535.

[229] L. Rocha, L. Velho and P. C. P. Carvalho, "Image moments-based structuring and tracking of objects," in *Proceedings of the XV Brazilian Symposium on Computer Graphics and Image Processing*, Jan. 2002, pp. 99-105.

[230] T. S. Rilwanu Bello and U. S. Tunga, "Determination of object position and orientation using image processing technique," *Research and Reviews: Journal of Physics*, vol. 7, no. 2, pp. 18–24, 2018.

[231] T. Li, M. Bolic and P. M. Djuric, "Resampling methods for particle filtering: Classification, implementation, and strategies," *IEEE Signal Processing Magazine*, vol. 32, pp. 70-86, April 2015.

[232] K. Nagla, M Uddin, D Singh and R Kumar. “Object identification in dynamic environment using sensor fusion,” in *Proceedings of the IEEE 39th Workshop on Applied Imagery Pattern Recognition*, Oct. 2010. pp. 1-4.

[233] T. T. Hoang, M. D. Phung, N. T. T. Van, D. A. Viet and T. Vinh, “Multi-sensor perceptual system for mobile robot and sensor fusion-based localization,” in *Proceedings of the International Conference on Control, Automation and Information Sciences (ICCAIS)*, Nov. 2012, pp. 259-264.

[234] N. J. Gordon, D. J. Salmond and A. F. M. Smith, “Novel approach to nonlinear/non-Gaussian Bayesian state estimation,” *IEEE Proceedings. F - Radar and Signal Processing*, vol. 140, no. 2, pp. 107-113, April 1993.

[235] Q.-B. Zhang, P. Wang and Z.-H. Chen, “An improved particle filter for mobile robot localization based on particle swarm optimization,” *Expert Systems with Applications,* vol. 135, pp. 181-193, Nov. 2019.

[236] X. Yu, J. Bai, T. Zhang and S. Wei, “One practical data fusion algorithm applied in auxiliary particle filtering,” in *Proceedings of the International Conference on Industrial Control and Electronics Engineering*, Aug. 2012, pp. 1724-1727.

[237] M. S. Arulampalam, S. Maskell, N. Gordon and T. Clapp, “A tutorial on particle filters for on-line nonlinear/ non-Gaussian Bayesian tracking,” *IEEE Transactions Signal Processing*, vol. 50, no. 2, pp. 174–88, Aug. 2002.

[238] R. Douc and O. Cappe, “Comparison of resampling schemes for particle filtering,” in *Proceedings of the 4th International Symposium on Image Signal Processing and Analysis (ISPA)*, Sep. 2005, pp. 64-69.

[239] M. K. Pitt and N. Shephard, “Filtering via simulation: Auxiliary particle filters,” *Journal of the American Statistical Association*, vol. 94, no. 446, pp. 590-599, June 1999.

[240] N. Vlassis, B. Terwijn and B. Krose, “Auxiliary particle filter robot localization from high-dimensional sensor observations,” in *Proceedings of the IEEE International Conference on Robotics and Automation*, vol. 1, May 2002, pp. 7-12.

[241] K. Brzozowska-Rup and A. L. Dawidowicz, “A new approach to the construction of the APF algorithm by applying the Pearson curves technique,” *Applied Mathematics and information Sciences*, vol. 10, no. 3, pp. 833-840, May 2016.

[242] J. Li, Y. Ho and J. Huang, "Line tracking with Pixy cameras on a wheeled robot prototype," in *Proceedings of the IEEE International Conference on Consumer Electronics-Taiwan (ICCE-TW)*, May 2018, pp. 1-2.

[243] M. F. Ahmad, H. J. Rong, S. S. N. Alhady, W. Rahiman and W. A. F. W. Othman, "Colour tracking technique by using pixy CMUcam5 for wheelchair luggage follower," in *Proceedings of the 7th IEEE International Conference on Control System, Computing and Engineering (ICCSCE)*, Nov. 2017, pp. 186-191.

[244] S. Gandhi and N. Gandhi, "A CMUcam5 computer vision based Arduino wearable navigation system for the visually impaired," in *Proceedings of the International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, Sep. 2018, pp. 1768-1774.

[245] S. Lee, G. S. Tewolde, J. Lim and J. Kwon, "Vision based localization for multiple mobile robots using low-cost vision sensor," in *Proceedings of the IEEE International Conference on Electro/Information Technology (EIT)*, May 2015, pp. 280-285.

[246] M. S. Bahraini, A. Rad and M. Bozorg, "SLAM in dynamic environments: A deep learning approach for moving object tracking using ML-RANSAC algorithm," *Sensors*, vol. 19, no. 17, pp. 1-20, Aug. 2019.

[247] D. Fox, "Adapting the sample size in particle filter through KLD-sampling," *International Journal of Robotics Research*, vol. 22, no. 12, pp. 985-1003, Dec. 2003.

[248] V. Elvira, L. Martino, M. F. Bugallo and P. M. Djuric, "In search for Improved Auxiliary Partial Filters," *In proceeding of European Signal Processing Conference (EUSIPCO)*, Sept. 2018, pp. 1637-1641.