JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

An Overview of Vision-based methods for Autonomous UAV Navigation

Adeeba Ali
Interdisciplinary Centre for Artificial Intelligence
Aligarh Muslim University
Aligarh, India
aliadeeba98@gmail.com

Rashid Ali
Interdisciplinary Centre for Artificial Intelligence
Aligarh Muslim University
Aligarh, India
rashidaliamu@rediffmail.com

M.F. Baig
Interdisciplinary Centre for Artificial Intelligence
Aligarh Muslim University
Aligarh, India
mfbaig.me@amu.ac.in

Abstract* - The Development of different navigation strategies for Unmanned Aerial Vehicles has been under scientific research for the last few decades and resulted in the construction of a variety of aerial platforms. Currently, the main challenge that the scientific community is facing is the design of fully autonomous unmanned aerial vehicles having the competency of safely carrying out missions without any human intervention. In order to improve the guidance and navigation skills of these aerial platforms, the control pipeline of the UAVs is integrated with visual sensing techniques. The objective of this survey is to demonstrate an extensive review of the literature on vision-aided techniques for autonomous UAV navigation. Particularly on vision-based localization and UAV state estimation, collision avoidance, path planning, and control. In addition to this, throughout this article, we will provide an insight into the challenges to be addressed, current developments, and trends in autonomous UAV navigation.

***Keywords* - Unmanned aerial vehicles(UAVs), obstacle avoidance, path planning, visual SLAM.**

I. Introduction

An Unmanned Aerial Vehicle (UAV) can be defined as an aircraft that can navigate without the intervention of a human pilot. Nowadays, the deployment of UAVs is increasing more and more for regular applications, particularly for infrastructure inspection, and surveillance due to its high flexibility and mobility. Although, in several intricate environments, the UAV is unable to sense the local environment exactly due to the shortcomings of traditional sensors like poor perception ability and lack of stable communication with one another. In order to overcome these limitations a lot of effort has been made, however, a more effective and efficient method is still required to be developed. Hence, for the development and deployment of UAVs, a high-performance capability of autonomous navigation is of great significance.

A. UAV navigation

UAV navigation is a process of planning a path to navigate quickly and safely to the target location using only the information about the current location and environment. In order to carry out the entire programmed mission successfully, a UAV must be completely sensible of its position, pose, turning direction, forward speed, starting position, and location of the target. Until now, several methods for UAV navigation have been put forward by researchers and can be categorized into three types: satellite navigation, inertial navigation, and vision-based navigation. Yet, these approaches alone are not ideal for fully autonomous UAV navigation. Therefore, it is difficult to go for the most relevant method for autonomous UAV navigation depending upon the particular task. Although, after reviewing the proposed methods for UAV navigation under each category, vision-based navigation is a promising and primary research area with the growing research and development in the field of computer vision. The visual sensors that are used in vision-based autonomous navigation have several advantages over traditional sensors. First, the on-board cameras can provide a good amount of online information about the surrounding environment; Second, they are perfect for the perception of the rapidly changing environment because they possess a valuable anti-inference ability; Third, generally, most of the visual aided sensors are unassertive, that is, they don't allow the detection of sensing system. The United States and Europe have already developed research institutions for navigation of aerial vehicles, such as Johnson Space Center of NASA [32]. Massachusetts Institute of Technology [44], the University of Pennsylvania [34] and several other top-ranked institutions are also rapidly developing research in the field of vision-based autonomous UAV navigation and have embodied this automation technique into transport systems of next-generation such as NextGen [80] and SESAR [35].

The illustration of vision-based navigation is presented in Fig 1.

Through the usage of inputs from proprioceptive and exteroceptive sensors, a UAV would be able to steer safely towards the location of the target after completing the tasks of state estimation, collision detection and avoidance, and motion planning.

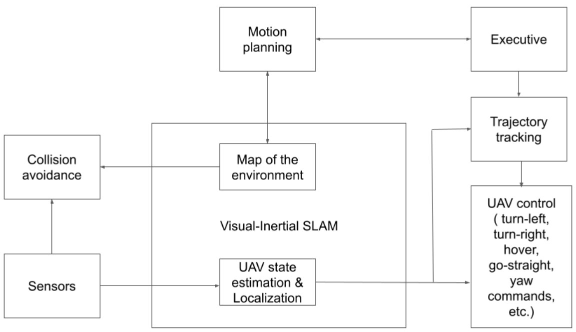

Fig 1: Vision-based UAV navigation system

B. UAV classification

After going through all the benchmark studies we classified UAVs into the following four types:

Fixed-Wing: These kinds of aircraft are incorporated with a rigid wing and a predetermined airfoil that generates lift due to the forward airspeed of the UAV, thereby making the flight possible. The forward airspeed of a UAV is produced by the thrust generated by a propeller in the forward direction. In addition to this, these aircraft are generally characterized for their high speed of voyage and high endurance mainly utilized for long-range, long-distance, and high altitude flights.

Rotatory-Wing: These aircraft are characterized by their ability to carry out tasks that need hovering of the flight. They have rotors made up of blades in continuous motion, which are required to produce lift by generating airflow. These aerial platforms are also known as vertical take-off and landing (VTOL) rotorcraft and are capable of a heavier payload, easier take-off and landing, and finer control than fixed-wing aircraft.

Flapping-Wing: These micro-UAVs are generally known for reproducing the flight of insects or birds. They have extremely low endurance and low payload capability due to their reduced size. These UAVs consume less power and are capable of performing vertical take-offs and landings with flexibility.

Airship: An airship also known as a dirigible is a “lighter-than-air” UAV that is propelled and driven across the air by utilizing propellers and rudders or other thrusts. By cushioning a large cavity with a lifting gas like a balloon, these aircraft can fly upwards. Non-rigid(or blimps), semi-rigid, and rigid are the three main types of Airship. A non-rigid airship or blimp is a kind of “pressure-airship”, that is the lighter-than-air vehicle that can be steered, powered, and its shape remains preserved by the pressure of gases within its envelope. These air vehicles are suitable for long-duration flights as no energy is required for lifting them, so that saved energy can be leveraged as a source of power for the movement of actuators, thus allowing flights of long-duration. Furthermore, these aircraft are capable of navigating with safety at low levels, close to people and buildings of the local environment. Fig 2 represents the discussed four classes of UAVs.

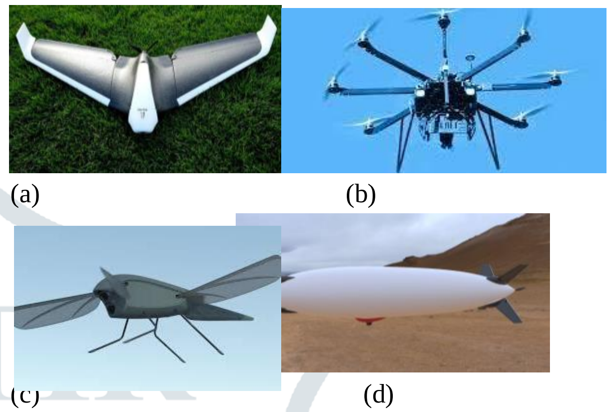

Fig 2: Classification of Unmanned Aerial Vehicles(UAVs): (a) Fixed-Wing UAV(image source: Jorge et al. (2017)[94]); (b) Rotatory-Wing UAV (image source: (https://www.shoghicom.com/unmanned-air-vehicle.php); (c) Flapping-Wing UAV(image source: (https://www.asme.org/topics-resources/content/engineering-a-robotic-bird-like-no-other)); (d) Airship(image source: (https://aerovehicles.net/avi-airship/av-10-airship))

C. On-board UAV sensors

Usually, Unmanned Aerial Vehicles acquire information of surroundings as well as states of their own through both proprioceptive and exteroceptive sensors. The conventional sensing devices used for UAV navigation are generally gyroscope, Inertial navigation system (INS), axis accelerometer, and global positioning sensor (GPS). However, these sensors are likely to affect the localization and navigation of UAVs. One of the biggest drawbacks of the Inertial navigation system (INS) is the generation of bias error which is caused by the problem of integral drift and results in decrement of accuracy to a certain extent. Also, the Global positioning system has limited reliability, since in some areas such as indoor environments, its precision is too low or it is not present at all. Furthermore, minute errors of angular velocity and acceleration are continuously engulfed into the errors (either linear or quadratic in nature) of velocity and position.

Therefore, one of the reasons for the limited use of UAVs in applications of both life and production is the improper functioning of traditional sensors. In order to cope up with these issues, the researchers are being more bothered about the use of state-of-the-art techniques for the enhancement of both the robustness and accuracy of the pose estimation. Ideally, one can acquire a much better estimation of the UAV’s state by using a fusion of multi-sensor data [11], which integrates the advantages of a variety of multiple sensors. However, restricted to particular environments, such as areas where GPS signal is unavailable, it is assumed that the incorporation of multiple sensors in small UAVs would be unnecessary and impractical. Therefore, a more typical approach is needed for the enhancement of UAVs' perception ability of the environment.

After comparing visual sensors to ultrasonic sensors, laser lightning, GPS, and other traditional sensors, we conclude that visual sensors capture much better information of surroundings with texture, the color of surrounding objects, and other visual details. Additionally, they can be deployed easily as well as they are cheaper, hence, this is the reason why vision-aided autonomous UAV navigation is tending to be a relevant topic in current research. The visual sensors which are commonly used for vision-based navigation are divided into four types as shown by Fig 3: monocular, stereo, RGB-D, and fisheye.

Monocular visual sensors are generally used in operations where minimum weight and compactness are required. Furthermore, they are cheap and can be easily deployed into UAVs. However, they are not capable of obtaining a depth map of the surrounding environment. When two monocular cameras of the same configuration are mounted on a rig then this system becomes a stereo camera. Therefore, a stereo camera is not only capable of providing information that a single monocular camera can give but also offers some further advantages of dual views. Additionally, it can be used to obtain depth maps using the principle of parallax without the support of infrared sensors. RGB-D cameras are capable of simultaneously generating depth maps and visible images using infrared sensors. They are mostly used in indoor environments due to the restricted scope of infrared sensors. Fisheye visual sensors are adapted from monocular visual sensors with the additional capability of providing a broad viewing angle and additional capability of obstacle detection in complex environments.

Fig 3: Typical visual sensors. (a) monocular camera; (b) stereo camera; (c) RGB-D camera; (d) fisheye camera

D. Previous Review Studies

Some of the researchers have carried out a survey of different techniques that could be used for autonomous UAV navigation and presented their valuable remarks and conclusions on those methods.

Kanellakis and Nikolakopoulos (2017) [40] presented a survey on vision-based methods that can be used for autonomous UAV navigation focusing mainly on current developments and trends. They discussed three steps that are required for autonomous flight: Navigation, Guidance, and Control. Navigation is further divided into the task of vision-based localization and mapping, obstacle avoidance, and aerial target tracking. Similarly, the guidance includes path planning, mission planning, and exploration. And then in last, the control task constitutes estimation of Attitude, position, velocity, and acceleration. Furthermore, the authors presented certain challenges that one have to face in implementing a vision-based autonomous UAV navigation system, such as lack of solid experimental evaluation in the integration of visual sensors in the UAV ecosystem, In spite of the development of elaborated SLAM algorithms for applications of vision-based systems, most of them cannot be used directly for UAV because of certain limitations posed by their processing power and architecture and the challenge imposed due to the incapability of UAVs to just stop operating in the state of great uncertainty like ground vehicles. Additionally, they presented their views on future trends in vision-based autonomous UAV navigation. According to them in the near future UAVs could get turned into the evolving elements in various applications as they possess some powerful characteristics like versatile movement, lightweight chassis, and onboard sensors. Also, they predicted that the techniques for mapping the surroundings will be further researched and revised for dynamic environments. Furthermore, they talked about the ongoing research on incorporating robotic arms/tools into UAVs so as to increase their competencies in aerial maneuver for several tasks, such as regular inspection, maintenance, and service. However, the authors don't discuss the vision-based navigation techniques that completely rely upon the front camera of UAV for self localization, mapping, and control. Furthermore, in the section visual localization and mapping the authors have mentioned about the techniques that utilize both the monocular camera and Inertial measurement unit (IMU) of the UAV. In these techniques the data from IMU sensors and camera are fused together and fed to an Extended Kalman filter (EKF) that would further estimate the pose, attitude, and velocity of the vehicle for flight control. The drawback in these techniques is that these techniques use additional IMU sensors for autonomous UAV navigation. The recent state-of-the-art works solve this issue by using ORB SLAM for pose estimation of the UAV and creating the map of the environment by leveraging only a monocular camera. Therefore, this review lacks the discussion about the latest state-of-the-art techniques that utilize only the front camera of the UAV for the fully autonomous navigation in GPS-denied areas.

Lu et al. (2018)[95] present a comprehensive review of the vision-based UAV navigation methods. In this review the process of UAV navigation is described in three steps: visual localization and mapping, obstacle avoidance, and path planning. In visual localization and mapping authors discussed various methods that can estimate the position, attitude, and velocity of the UAV. These methods may or may not use the map of the environment. The methods that utilize the map of the environment are further divided into two types: map-based and map building systems. Map-based systems require a priori map of the environment according to which they plan the navigation and control strategy of the UAV. Map-building systems leverage vision-based SLAM algorithms like ORB SLAM and PTAM that can build the map of the environment as a UAV moves forth, and localize the UAV in the environment by estimating its position coordinates. The second category of visual localization methods doesn't require any map of the environment for localizing the UAV in the environment. These methods include optical flow, and feature tracking techniques. The authors haven't discussed the state-of-the-art deep reinforcement learning methods under this category for localizing and navigating UAV in the given environment. For obstacle detection and avoidance the review lacks the discussion on the latest developments in this field. The algorithms that are discussed by the authors are optical-flow based and SLAM-based. However, these algorithms take much more processing time, hence not efficient for on-board utilization. Instead, many recent AI-based obstacle avoidance algorithms are proposed that can be executed during flight and take less computational time for sending response to the UAV. Furthermore, the authors of the review have discussed only classical path planning algorithms like A-star, RRT, Ant-colony optimization, etc. They don't present their review on the state-of-the-art AI-based path planning algorithms that perform much better than the traditional algorithm in terms of complexity, and computational time in unknown, dynamic, and large-scale environments. Along with this, the authors don't mention about the papers who have conducted real world experiments with the fully autonomous UAV by incorporating all the steps of autonomous navigation.

Later on, Belmonte et al. (2019) [6] carried out extensive research on the vision-based autonomous navigation method. He divides the UAV vision-based task into four subtasks: navigation, control, tracking or guidance, and sense-and-avoid. According to their review, the navigation task includes figuring out the position and orientation of an aircraft using visual sensing devices. Then, in the next step authors discussed the control of aircraft position on the basis of the data of surroundings captured by visual sensors. They further state that in the past few years several methods have been proposed for vision-based control of UAV. One of these methods is control using visual servoing. When the visual data of the surrounding environment is directly incorporated into the control loop, this is known as visual servoing. Therefore, in this technique, the control law depends on the error signal ascertained from data of the surrounding environment obtained through visual sensors. After this, the authors discussed the next subtask, that is, vision-aided tracking or guidance. According to them, the control system of UAVs leverages relative navigation for achieving a flight with respect to a target. Finally, vision-aided sense-and-avoid (SAA) is needed for a fully autonomous UAV system in order to sense and avoid obstacles in both static and dynamic environments. In the benchmark works reviewed by authors, the researchers used one or more visual sensors for detecting possible collisions with obstacles including other UAVs in the environment, and for determining the necessary actions required for achieving control over the flight which is free from collisions. Then, in the end, the authors concluded that the best system that could be used for autonomous navigation is the monocular UAV system as it is easier to install and allow the users to reduce the payload of UAV. Furthermore, the authors suggested that virtual reality would be an important aspect in the development process of personal UAVs as it would help in conducting experiments with UAVs in realistic indoor environments containing different kinds of obstacles, as well as in outdoor environments.

E. Motivation of this review

The objective of this survey is to present an outline of the most important vision-based navigation methods that can be used for autonomous UAV navigation. Furthermore, a collection of benchmark studies is provided that could act as a guideline for future research towards vision-based autonomous aerial exploration. In [83] authors proposed a fact that with the rapidly developing popularity of small-sized commercial UAVs and the huge development of computer vision, the combination of both of them, that is, vision-based UAV navigation has been a working research area. Seeing that, the field of Computer vision for aerial vehicles is emerging as a popular trend in autonomous navigation, hence the presented work will focus only on reviewing the fields of localization and mapping, obstacle detection and avoidance, and path planning. The essence of this work is to provide rich insight into the entire task of autonomous UAV navigation collecting all pieces together.

UAV state estimation includes the extraction of distinct features from the surrounding environment with the aid of visual sensors, so as to determine the next step of UAV in the environment. We categorized this phase into three categories: optical flow and machine learning based systems, sensor based systems, and Visual SLAM based systems. Where the Visual SLAM based methods are divided into feature based SLAM methods, intensity based methods, hybrid methods, and Visual inertial SLAM based techniques. As the size of the inertial measurement unit (IMU) is getting cheaper and smaller, hence Shen et al. 2014 [75] and Leuteneggar et al. 2015 [45] proposed and developed a navigation system in which they fused inertial measurement unit (IMU) and visual measurements together for obtaining better performance results. In order to circumvent the limitations on power consumption and perception ability that make a single UAV incapable of completing certain types of tasks, In [29,57] it is shown that with the enhancement of autonomous navigation, multiple UAVs can carry out such tasks together. Then in the next step, the task of obstacle detection and avoidance is discussed. The basic principle of this task is to sense and avoid obstacles and determine the distances between those obstacles and a UAV. When a UAV reaches an obstacle then it is required to avoid or turn around under the directions provided by the obstacle avoidance module. We provide a review of two kinds of obstacle avoidance techniques: obstacle avoidance using optical flow and obstacle avoidance using the visual SLAM approach. In [17,39,55,58,63,85] researchers have proposed different solutions which exploit cameras as the only visible sensing devices for obstacle avoidance in composite, unstructured, dynamic, and extensive environments. Then finally, we provide a review on path planning. In this step, local path planning and global path planning (trajectory generation) are incorporated for application-wise decision making, motion planning, or exploration of areas that are new and unknown. Leveraging all these steps, Wang et al. 2020 [89] proposed and developed a system for autonomous exploration with impressive progress in the task of vision-aided autonomous navigation, however, the development of fully autonomous navigation systems is still a challenge due to various unsolved problems in their designing process, such as autonomous obstacle avoidance, generation of an optimized path in non-static situations, and dynamic assignment of operations.

The remaining paper is categorized as follows: First, in section II, we present three different types of vision-based UAV state estimation techniques. Next in section III, we introduce a review of collision detection and avoidance methods in autonomous navigation. After that, in Section IV, we focus on the path planning and vision-based control approaches for autonomous UAV. Then, in Section V we mentioned the experiments that have been carried out with a physical UAV for applying vision-based autonomous navigation strategy to real world problems. Finally, in section VI we present a conclusion with additional analysis on difficulties and directions for future research in the field of vision-aided autonomous UAV navigation.

II. VISION-BASED UAV STATE ESTIMATION

On the basis of the techniques that can be used to determine the state of the UAV and create the map of the surrounding environment for autonomous navigation, visual localization and mapping systems can be categorized into three classes: Optical flow and machine learning based system, Visual SLAM based system, and Sensor based (Inertial) system [15], as shown by Fig 4.

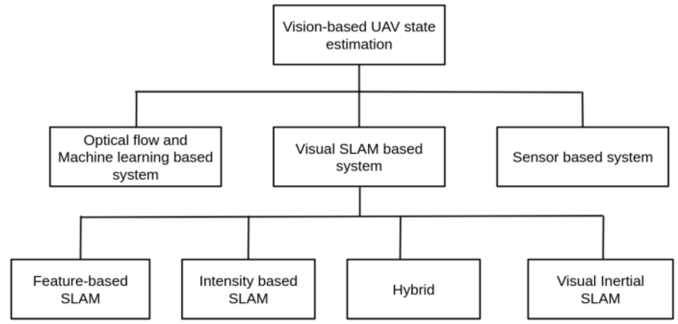

Fig 4: Visual localization and mapping systems.

A. *Optical flow and Machine learning based systems*

A UAV localization system, developed using optical flow and machine learning based techniques is capable of navigating without a prior map of the environment, and aerial vehicles following this approach navigate only through the extraction of distinct features from the observed environment. Nowadays, the most commonly used mapless-system-based methods for localization and mapping are feature tracking methods, optical flow methods, and machine learning techniques. Optical flow methods can be divided into two classes: local methods [50] and global methods [33]. In [72] a technique that can imitate the flying style of a bee through the estimation of the movement of objects by using visual sensors on the two sides of an autonomous aerial vehicle has been discussed. First, it estimates the ocular speed of both the visual sensors that are parallel to the barrier, respectively. If the optical velocities come out to be equal then the robot proceeds in the direction of the central line; otherwise, it goes forward with the pace of small places. Although, it might not have good performance while navigating in environments with texture-less walls. However, in recent years a significant growth of optical flow-based techniques has been made with some advancements in the field of detection, inspection, and tracking. Recently, in [62] authors proposed a novel approach for scene change detection using and describing the optical flow. Furthermore, in [31] researchers have worked on flight landing and hovering maneuvers upon a proceeding ground vehicle by fusing measurements obtained through optical flow techniques together with an inertial measurement unit (IMU). Even various challenging tasks, like surveillance and tracking of the target, can be achieved with systems using heavy optical flow computation as they can identify the maneuver of all the objects in motion [54]. In the field of localization and mapping, the feature tracking method [13] has become a standard and robust approach. This method can determine the maneuver of an obstacle through the detection of invariant features including corners, lines, and so on, and their corresponding movement in sequential images. The unvarying attributes of the surrounding environment that have been noticed previously are needed to be observed again from multiple viewing angles, illumination conditions, and distances during the task of robot navigation [84]. Basically, sparse features that can be used in localization and mapping are not suitable for obstacle avoidance as they are not dense enough. A behavioral method for navigation that utilized a system capable of recognizing visual landmarks combined with a fuzzy-based obstacle avoidance system was proposed by authors of [48]. Hui et al. (2018)[37] proposed a vanishing point-based approach for localization and mapping of autonomous UAVs to be used for safe and robust inspection of power transmission lines. In this proposed method authors regarded the transmission tower as an important landmark by which continuous surveillance and monitoring of the corridor can be achieved. Similarly, deep reinforcement learning based methods have also been used for estimating the state of the UAV and localizing it in the unknown environments without the need of a prior map of the environment. Following this approach, Huang et al. (2019)[116] proposed a Deep-Q network based method for autonomous UAV navigation. In the proposed approach the deep neural networks are integrated with reinforcement learning to overcome the limitations of reinforcement learning as it cannot be used to solve a complex problem alone. Similarly, Grando et al. (2020)[118] proposed a mapless navigation strategy for UAVs. The proposed approach leveraged two deep reinforcement learning based algorithms, Deep deterministic policy gradient (DDPG) [96] and Soft actor critic (SAC) [97]**.** The information about the relative position of the vehicle from target and vehicle's velocity are obtained through the reading of laser sensor mounted on UAV and localization data. This navigation approach is completely mapless and doesn't need any heavy softwares like ORB SLAM and LSD SLAM to create the map of the environment. For both DDPG and SAC methods, the network consists of 24 inputs and 2 outputs. The 24 inputs include 20 laser readings, 2 readings of previous linear velocity and yaw angle, and 2 readings are from the relative position of the UAV and angle to the target. And, 2 outputs are values of the linear velocity of the UAV and the yaw angle which are required to control the UAV. The authors of this work have verified the performance of their technique by comparing their navigation strategy of UAV with geometry based tracking controller on the Gazebo simulator. In the simulation environment without obstacles the geometry based tracking controller performs better whereas, in the presence of obstacles the proposed performs better as in that case the UAV navigating with the tracking controller collides with the obstacles. Some researchers have also applied deep neural networks to UAV localization problem. Li and Hu (2021)[98] proposed a mapless, deep learning based localization technique for the auto landing of unmanned aerial vehicles at a specified target. The proposed approach is based on ground-vision based methods in which two stereo cameras are installed or fixed on two independent Pan-tilt units (PTUs) symmetrically to capture the images of the vehicle at the time of landing. Furthermore, this work is based on deep learning, in which two deep neural networks are leveraged, Bbox-locate-Net and Point-Refine-Net. The images captured by the ground visual system are fed as an input to the Bbox-Net which predicts the coordinates for the bounding box of the target. Then, after combining these coordinates with the PTUs' parameters, the spatial position of the UAV can be determined. On the basis of Extended Kalman Filter (EKF) based object spatial localization algorithm, the motion continuity of the UAV can be checked. In case, the coordinate is not correctly estimated, then it would be given as an input to the Point-Refine-Net for precise prediction of the key points coordinates for the bounding box of the target. Furthermore, in order to improve the state estimation accuracy of the UAV, some researchers have also proposed techniques which integrate deep neural networks, and reinforcement learning. Following this idea, Afifi and Gadallah (2021)[99] proposed a 3-D localization method for a UAV using deep learning and reinforcement learning. In this research, authors leverage a 5-G cellular network of 4 ground base stations, and in order to determine the location of UAV with minimum error the problem is reduced to optimization problem where objective is to minimize the error in the measurement of RSSI (Radio received signal strength index) readings of the 4 ground base stations. In order to estimate location of the UAV in real-time a deep learning model is trained using the correlation between RSSI readings and UAV localization, further to improve the real-time performance of the proposed model, the authors also introduced reinforcement learning to their work and compare the localization results with those obtained from deep learning-based approach. After carrying out the analysis of the results, the authors conclude that for a small region with less dynamic obstacles, deep reinforcement learning gives better results but it is more computationally expensive, however for larger and crowded environments with dynamic obstacles, the deep learning approach performs better. Along with this, there are several mapless localization techniques which determine the location of UAV using feature matching between the UAV-satellite image pairs. Goforth and Lucy (2019)[100] proposed the CNN based method to estimate the position of the UAV in an unknown environment. In this method a convolutional neural network extracts feature representations from UAV-satellite image pairs given as an input to the network. By leveraging these extracted feature representations, the images of UAV are aligned with the ground-referenced satellite images and the location of the UAV is estimated. Similarly, Hou et al. (2020)[101] proposed a deep learning based technique for mapless UAV localization which integrates the digital evaluation model (DEM) and the deep learning architecture D2-Net. The D2-Net architecture is responsible for the extraction of keypoints and matching of satellite-UAV images. The idea behind this is to combine the obtained geological map of the area from DEM with the keypoint features to get the 3-D position coordinates of the key points in the UAV image. Subsequently, Bianchi and Barfoot (2021)[8] demonstrated a robust and fast method for localization of a UAV using satellite images which were further utilized for training autoencoders. In this work, the authors collected Google Earth (GE) images, and then an autoencoder model is trained to compact these images in order to represent them as a low-dimensional vector without distorting important features. The trained autoencoder model is further trained to reduce the size of a real-time image captured by UAV which is then compared with the pre-collected autoencoded GE images by using an inner-product kernel. Hence, localization of a UAV can be achieved by distributing weights over the relative GE poses. The evaluation results of this work have shown that the proposed approach is able to achieve root mean square error (RMSE) of less than 3m in real-time experiments.

B. Sensor based systems

These UAV systems leverage inertial unit, bioradar, lidar sensor, etc to explore the environment with divergence and ability of movement planning after defining the spatial configuration of the surroundings in a map. Therefore, with the aid of these sensors a static map of the environment can be obtained prior to navigation. Typically, environmental maps can be classified into two types: occupancy grids and octree. Each and every detail of the environment, varying from the 3D illustration of the environment to the interdependencies between environmental elements are included in these maps. In [22] authors presented a method that used a 3D volumetric sensor that can enable an autonomous robot to efficiently explore and map urban environments. In their work, they constructed the 3D model of the environment using a multi-resolution octree. Later on, for the depiction of the 3D model of the environment, an open-source framework was developed [33]. Basically, the approach here is to render the 3D representation of the environment octree. Although not only using the preoccupied area but also unknown and unoccupied space. In addition to this, by using an octree map compression technique, the represented model is made to occupy less space which makes the system capable of storing and updating the 3D models efficiently. Authors of [27] collected and processed data of the surroundings using a stereo camera, which can be leveraged further to produce a 3D map of the environment. The key of this approach is the extended scan line grouping technique that the authors used for precise segmentation of the range data into planar segments. This approach can also effectively confront noise present in the depth estimated by the stereo vision-based algorithm. Dryanovski et al. (2010)[16] proposed a method that represents a 3D environment by using a multi-volume occupancy grid which is a cable of explicitly storing information about both free space and obstacles. In addition to this, by fusing and filtering in new positive or negative sensor data incrementally, the proposed method can correct previous sensor readings that were potentially erroneous.

C. Visual SLAM based system

Many times, it is difficult for a UAV to traverse in the air with a previously known map of the surroundings obtained through UAV sensors due to certain environmental limitations. In addition to this, it would be unfeasible to acquire information of the target location in extreme cases (such as in calamity-hit areas). Therefore, in such situations, it would be a more efficient and attractive solution to go for simultaneous localization and mapping at the time of navigation. SLAM based robot navigation has been tremendously used in both semi-autonomous and fully autonomous disciplines, and this technique for state estimation and building map of the environment is getting popular rapidly along with the expeditious growth of visual simultaneous localization and mapping (visual SLAM) methods [4,79]. The size of the UAVs available in the market nowadays is shrinking, which restricts their competency of lifting payload in some measure. Hence, researchers and developers have been more focused on the usage of manageable and lightweight visual sensors rather than the conventional heavy sonar and laser radar, etc. Stanford CART Robot [59] carried out the first attempt at single camera-based map-building techniques. Later on, the detection of 3D coordinates of images was improved using a modified version of an interest operator algorithm. The result that was expected from this system was the demonstration of 3D coordinates of obstacles, which were placed on a mesh with each square cell of length 2m. However, this technology is capable of reconstructing the obstacles in the environment but still, it cannot model a large-scale world environment. Afterward, in order to recover states of cameras and structure of the environment simultaneously, vision-aided SLAM techniques have made a good amount of progress and resulted in four kinds of Visual SLAM techniques: feature-based, intensity-based, hybrid, and visual-inertial according to the method of image processing by visual sensors. Wang et al. (2020) [89] proposed and developed an efficient system for autonomous 3D exploration with UAV. The authors put forward a systematic approach towards robust and efficient autonomous UAV navigation. For localization of UAV, the authors leveraged a map of the environment which was constructed incrementally and preserved during the process of inspection and sensing. This road map is responsible for providing the reward and cost-to-go for a candidate region that is to be investigated and these are the two measures of the next best view (NBV) based evaluation. The authors verified their proposed framework in various 3D environments and the results obtained exhibit the typical attributes in NBV selection and much better performance of their approach in terms of the efficiency of aerial analysis as compared to other methods.

1) Feature-based SLAM

Feature-based SLAM systems first detect and extract attributes from multi-media data and then use these features as inputs for localization and estimation of motion instead of using the images directly. Generally, the extracted features are assumed to be uniform to viewpoint changes and rotation, as well as robust to noise and blur. In the past few years, thorough research on detection and description of features has been carried out and several types of attribute identifiers and descriptors have been proposed [47,86]. Therefore, most of the recently proposed SLAM algorithms are supposed to operate under this attribute-based framework. In [14] authors proposed a monocular vision-based localization along with the mapping of a sparse set of natural attributes based on a top-down Bayesian framework for achieving real-time performance. This work is a landmark for monocular vision-aided SLAM and has a considerable influence on subsequent research. Klein and Murray (2007) [41] proposed an algorithm for parallel tracking and mapping. This is the first algorithm that divides the SLAM system into two independent parallel threads: tracking and mapping which is the standard of current feature-based SLAM systems. A state-of-the-art approach for large-scale navigation was proposed by the authors of [53]. Common issues in large-scale environments, such as relocation of trajectories are being corrected by visual loop-closure detections. Celik et al. (2009) [12] proposed a visual SLAM-based system for navigation in unknown indoor environments without the aids of GPS. UAVs have to predict the state and range using only a single onboard camera. Representing the environment using energy-based straight architectural lines and feature points in the heart of the navigation strategy. In [30] researchers presented a vision-based attitude estimation method that leveraged multi-camera parallel tracking and mapping (PTAM) [41]. PTAM first integrates the estimates of the 3D motion of multiple visual sensors within an environment and then correlates the mapping modules and state estimation. A state-of-the-art approach is also demonstrated by the authors to calibrate external parameters for systems using multiple visual sensors with well-separated fields of view.

Although, many of the feature-based SLAM methods are able to reconstruct only a specific set of points as they pull out only sharp feature points from images. These types of methods are called sparse feature-based methods. So, researchers have been anticipating that more enhanced and dense maps of the environment could be reconstructed by the dense feature-based methods. A dense energy-based method is exploited by Valgaerts et al. (2012) [88] for the estimation of an elementary matrix and for further calibration of similarities by using it. Ranftl et al. (2016) [66] used a segmented optical flow field for producing a rich depth map of the surroundings from two successive frames, which means that adhering to this framework through optimization of a convex program could result in a dense reconstruction of the scene.

2) Intensity based SLAM

SLAM techniques that leverage the features of captured images for building a map of the environment seem to exhibit better performance in only ordinary and simple environments. However, they faced difficulties in texture-less environments. Therefore, the intensity based SLAM came into play. Visual SLAM algorithms based on this approach optimize geometric parameters by exploiting all the information regarding intensity present in the images of the surroundings, which can give strength to geometric and photometric deformations present in images. Furthermore, this strategy can find dense resemblances so they are capable of reconstructing a dense map at an additional price of computation. In [76] researchers presented a state-of-the-art approach for the estimation of scene structure and pose of the camera. In this work the authors directly utilized intensities of the image as observations, thereby, leveraging all data present in images. Hence, for the surroundings with few feature points, the proposed approach is proved to be much better and robust than feature based methods. Authors of [61] proposed a monocular visual SLAM algorithm that can perform in real-time, DTAM, using intensity of the images can also predict the 6 DOF motion of a camera. At a frame rate derived from the predicted comprehensive textured depth maps, the proposed algorithm is capable of generating dense surfaces. To provide the approximation for semi-dense maps authors of [18] proposed and developed an efficacious stochastic intensity based method, which can be utilized further for the calibration of images. Rather than optimizing parameters without using a scale, LSD SLAM [18] uses a different approach that exploits pose graph optimization, which allows loop closure detection and scale drift correction in real-time by explicitly taking scale factors into account. Krul et al. (2021) [42] presented and developed a visual SLAM-based approach for the localization of a small UAV with a monocular camera in indoor environments for farming and livestock management. The authors compared the performance of two visual SLAM algorithms: LSD-SLAM and ORB-SLAM and found that ORB-SLAM based localization suits best at those workplaces. Further, this algorithm was tested through several experiments including waypoint navigation and generation of maps from the clustered areas in the greenhouse and a dairy farm.

3) Hybrid method

The hybrid method is a combination of both feature based and intensity based SLAM methods. The first step in the hybrid method includes initialization of feature correspondences by using intensity based SLAM, then in the next step, the camera poses are refined continuously using feature descriptor SLAM algorithm. In [20] authors proposed an innovative semi-direct algorithm (SVO), for the estimation of the states of a UAV. Similar to PTAM, this algorithm also implements motion estimation and mapping of point clouds in two different threads. A more accurate pose estimation can be obtained using SVO by combining gradient information and pixel brightness with the calibration of feature points and loss in reprojection error. Subsequently, for real-time landing spot detection and 3D reconstruction of the surroundings, the authors of [21] proposed and developed an algorithm that is computationally efficient and capable of providing the best results in real-time performance. In order to carry out exploration of the real-world environment, a high frame rate visual sensor is required for the execution of a semi-direct algorithm, as it has limited resources to carry out the heavy computation.

4) Visual Inertial SLAM

Bonin-Font et al. (2008) [9] developed a system for the navigation of ground mobile robots in which they utilized laser scanners for accessing 3D point clouds of relatively good quality. Also, nowadays UAVs can be equipped with these laser scanners as their size is getting smaller. Though different kinds of measurements from different types of sensors can be fused together and this can enable a more robust and accurate estimation of the UAV state. A typical SLAM system, extended Kalman filter (MSF-EKF) can deal with multiple delayed measurement signals for different types of sensors and provide a more robust and accurate prediction of the UAV attitude for robust control and navigation [51]. Magree and Johnson (2014) [52] proposed and developed a navigation system that exploits the fusion of an EKF-based inertial navigation system with both laser SLAM and visual SLAM. The monocular visual SLAM is responsible for finding data association and estimating the pose of the UAV, whereas the laser SLAM system is liable for scan-to-map matching utilizing a Monte Carlo framework. Hu and Wu (2020) [36] proposed a multi-sensor fusion method based on the correction of an adaptive error using Extended Kalman Filter (EKF) for localization of UAV. In their presented approach first, a multi-sensor system for localization is fabricated using acceleration sensors, gyroscopes, mileage sensors, and magnetic sensors. Then, the data obtained from these sensors is adjusted and compared in order to minimize the error from the estimated value. Finally, measurement noise and system noise covariance parameters in EKF are optimized through the transformative iteration mechanism of the genetic algorithm. Then, the authors figure out the adaptive degree by obtaining the absolute value of the difference between the predicted and the real value of EKF. Table 1 summarizes the algorithms for UAV state estimation and localization cited in this review.

Table 1: Summary of important methods in vision-based UAV state estimation.

Techniques used	**Sensors / algorithms used**	**References**
Optical Flow	Monocular camera	[33],[50],[54],[62]
	Stereo Camera	[72]
	Multi-sensor fusion	[31]
Feature tracking	Monocular camera	[48],[84],[13],[37],[8]
Machine learning	Reinforcement learning	[116],[117]
	Deep learning	[98],[99],[100]
	UAV-satellite image pairs matching using deep learning	[101]
Inertial odometry	Stereo camera	[27]
	Depth camera	[22],[16]
Visual SLAM	Feature-based SLAM	[14],[41],[53],[12],[30],[88],[66],[89]
	Intensity-based SLAM	[76],[61],[18],[42]
	Hybrid method	[20],[21]
	Visual-Inertial SLAM	[59],[51],[52],[36]

III. Collision avoidance

Avoidance of collision with obstacles in the navigation path is a crucial step in the process of autonomous navigation since with the help of this capability a robot can detect, avoid collision with nearby objects, and navigate safely without any risk of a crash. Thus, this method plays a key role in increasing the level of automation of UAVs. The main objective of obstacle detection and avoidance is to estimate the distances between the aerial vehicle and obstacles after detecting them. When the UAV is getting closer to obstacles, then it is required to stay away or take about-turn as per the directions of the collision avoidance technique used. Among those approaches that can be used to solve this problem, one is to use laser range finders such as radar, IR, and ultrasonic, etc for measuring the distance between a UAV and an obstacle. However, these laser range finders are unable to get enough information in complex environments, since they have limited measurement range and field of view. In order to overcome these issues, laser sensors could be replaced with visual sensors that can provide an ample amount of data, which can be further refined for obstacle avoidance. Typically, there are two types of approaches for the avoidance of obstacles: Optical-flow based techniques and SLAM-based techniques. Authors of [26] used the image processing technique for the avoidance of obstacles. By exploiting optical flow, this technique is capable of generating local information flow and depth maps of images. A novel approach for detecting the change in the size of obstacles was proposed in [1]. Their proposed method is based on the principle of how human eyes perceive objects. According to this mechanism, the obstacles in the field of view are becoming larger as the eyes get nearer to them. Using this concept, the authors of this work proposed an algorithm that can detect obstacles by comparing and contrasting the successive images, and figure out whether the object present in the navigation path is closer or not. Various optical flow methods based on bionic insect vision have also been proposed for obstacle avoidance. Authors of [77] were inspired by vision of bees and presented a simple and non-continual optical flow technique for estimating the self-motion of the system and global optical flow. After being persuaded by the optic nerve structure of insects, In [28] authors proposed and developed a basic unit for the detection of local motion. Similarly, a sensor based on the compound structure of the visual system of flies and a flow strategy were designed in [70]. These algorithms are based on binocular vision of insects and can be employed in UAVs. In the work of [7], a student of physics, Darius Mark put forward a technique that leveraged only the speed of light for measuring distance between the objects. This approach is simple but efficient as many insects can detect the obstacles in the surrounding environment using light intensity. At the time of flight, the motion of the image produces a light flow signal on the retina, this visual flow for the vision-aided navigation of insects provides a variety of information of attributes present in space. Hence, insects can find out quickly whether they are able to pass by the obstacles safely or not, based on the intensity of light going through the aperture in the leaf. However, obstacle avoidance approaches based on the optical-flow method are not able to acquire precise distance due to which they are not used in several particular missions. In contrast, SLAM-based techniques can come up with a precise estimation of metric maps using an advanced SLAM algorithm, thereby making UAVs capable of navigating safely with a collision avoidance feature exploiting more environment information [23]. A novel technique of mapping a prior unknown environment using a SLAM system was proposed in [60]. Furthermore, the authors leveraged a state-of-the-art artificial potential field technique to avoid collision with both static and dynamic obstacles present in the surrounding environment. Later on, in order to cope up with the mediocre illumination of indoor environments and holding on the count of feature points, authors of [5] proposed a procedure for creating adjustable feature points in the map, which is based on the PTAM (Parallel Tracking and Mapping) algorithm. Subsequently, In [19] researchers proposed an approach based on oriented fast and rotated brief SLAM (ORB-SLAM) for UAV navigation with obstacle detection and avoidance aspect. The proposed method first processes data from video by figuring out the locations of the aerial vehicle in a 3D environment and then generates a sparse cloud map. After this, it enhances the density of the sparse map, and then finally, it generates an obstacle-free map for navigation using potential fields and rapidly exploring random trees (RRT). Yathirajam et al. (2020) [90] proposed and developed a real-time system that exploits ORB SLAM for building a 3D map of the environment and for continuously generating a path for UAV to navigate to the goal position in the shortest time. The main contributions of this approach are: implementation of chain-based path planning with built-in obstacle avoidance feature and generation of the path for UAV with minimum overheads. Subsequently, Lindqvist et al. (2020) [49] proposed and demonstrated a Nonlinear Model Predictive Control (NMPC) for obstacle avoidance and navigation of a UAV. In this work, the authors proposed a scheme to identify differences between different kinds of trajectories to predict positions of future obstacles. The authors conduct various laboratory experiments in order to illustrate the efficacy of the proposed architecture and to prove that the presented technique delivers computationally stable and fast solutions to obstacle avoidance in dynamic environments. Subsequently, Guo et al. (2021)[102] proposed geometry based collision avoidance techniques for real-time UAVs navigating autonomously in 3-D dynamic environments. In the proposed obstacle avoidance technique first, the onboard detection system of the UAV obtains the information about the irregular obstacles encountered in the navigation path, then that obtained data is transferred to the regular convex bodies, which is further used for generating avoidance trajectories in the shape of circular-arc. Then, on the basis of geometric correlation between UAV and obstacle modelling, the real-time obstacle avoidance algorithm is developed for UAV navigation. Hence, the authors of this work formulate the obstacle avoidance problem as a trajectory tracking strategy. Following the idea of geometrical approach for collision detection and avoidance, Aldao et al. (2022)[103] presented and developed a collision avoidance algorithm for UAV navigation. The authors have proposed this method mainly for the purpose of inspection and monitoring in civil engineering applications. So, by keeping this use in mind, the collision avoidance strategy is developed for a UAV navigating between the waypoints following a predefined trajectory. Furthermore, in order to avoid the collision with dynamic obstacles like people, other UAVs, etc, the 3-D onboard sensors are also utilized to determine the position of these dynamic objects which are not considered in the original trajectory. The proposed technique is based on sense and avoid obstacle avoidance in which first after reaching the waypoint of predefined trajectory UAV takes the sample pictures and inertial measurements of the room to detect the presence of obstacles. If the obstacle is detected then in that case the position of that obstacle is registered in the point cloud of the inspection room at that moment of time otherwise the scheduled path will be followed. If a new obstacle is detected after completing the trajectory then its position will be determined using on-board 3-D sensors. Table 2 encapsulates the techniques used for collision detection and avoidance.

Table 2: Summary of important methods in collision avoidance.

Collision avoidance techniques	Methods used	References
Vision based	Optical flow	[77],[70],[28],[26],[1]
	NMPC	[49]
Potential field	ORB-SLAM	[19],[90]
	PTAM	[5]
Geometrical algorithm	Trajectory tracking	[102]
Sense and avoid	Trajectory tracking; Multi-sensor fusion	[103]

IV. MOTION PLANNING

Motion planning refers to the process of navigating safely from one place to another in presence of obstacles. It comprises two tasks: path planning and control. Path planning is responsible for finding an efficient obstacle free path from source to destination which a UAV can follow, whereas control provides necessary commands to UAV for moving through that path without colliding with the obstacles. In this section, first we discuss different types of existing path planning approaches and then we demonstrate the review of the existing literature on different strategies of vision-based control for UAV.

A. Path planning

Planning of a collision-free path for safe navigation of an UAV is an essential step in the task of autonomous UAV navigation. Path planning is required for finding a best path from the initial point to the location of the target, on the basis of certain performance indicators like the less price of work, the shortest time of flight, and the shortest flying route. At the time of path planning, the UAV is also required to avoid obstacles. Based on the environmental information that is to be utilized for computing an optimal path, the problem of path planning can be classified into two types: global path planning and local path planning. The aim of the global path planning algorithm is to figure out an optimal path using a geographical map of the environment deduced initially. Although, for controlling a real-time UAV the task of global path planning is not enough, particularly when there are several other tasks that need to be carried out forthwith or unexpected hurdles emerging at the time of flight. Hence, there is a need for local path planning for estimating the path free from collisions in real-time by using data of the sensors taken from surrounding environments.

1) Global path planning

Global path planner requires a priori geographical map of the surrounding environment with locations of starting and target points to compute a preliminary path, therefore the global map is also known as non-dynamic, that is static map. Generally, there are two kinds of algorithms that are used for global path planning: Heuristic searching techniques and a sequence of intelligent algorithms. In [46] authors proposed and developed an algorithm for planning the trajectory of multi-UAV in static environments. The proposed algorithm includes three main stages: the generation of initial trajectory, the correction of trajectory, and the smooth trajectory planning. The first phase of the proposed algorithm can be achieved through MACO which incorporates metropolis measure into the node screening method of ant colony optimization (ACO) that can efficiently and effectively avoid cascading into the stagnation and local optimal solution. Then in the next phase of the algorithm the authors of this work proposed three different trajectory correction schemes for solving the problem of collision avoidance. Finally, the discontinuity resulting from the acute and edged turn in trajectory planning is resolved by using the inscribed circle (IC) smooth method. Results obtained through various laboratory experiments demonstrate the high effectiveness and feasibility of the proposed solution from perspectives of obstacle avoidance, optimal solution, and optimized trajectory in the problem of trajectory planning for UAVs. Yathirajam et al. (2020) [90] proposed a chain-based path planning approach for the generation of a feasible path for UAV in the ORB-SLAM framework with dynamic constraints on the length of the path and minimum turn radius. The presented path planning algorithm enumerates a set of nodes that could move in a force field, thereby permitting the rapid modifications of the path in real-time as cost function changes. Subsequently, Jayaweera and Hanoun (2020) [38] demonstrated a path planning algorithm for UAVs that enables them to follow ground moving targets. The proposed technique utilized a dynamic artificial potential field (D-APF), the trajectory generated by this algorithm is smooth and feasible to non-static environments with hindrance and capable of handling motion contour for the target moving on the ground considering the change in their direction and speed. The existing path planning techniques, such as graph-based algorithms and swarm intelligence algorithms are not capable of incorporating UAV dynamic models and flying time into evolution. In order to overcome these limitations of existing methods Shao et al, (2021) [74] proposed a hierarchical scheme for trajectory optimization with revised particle swarm optimization (PSO) and Gauss pseudospectral method (GPM). The proposed scheme is a two-layered approach. In the first layer, the authors designed a better version of PSO for path planning, then in the second layer after utilizing the waypoints in path generated by improved PSO, a fitted curve is constructed and used as the starting values for GPM. After comparing these initial values with the ones generated randomly, the authors conclude that the designed curve can improve the efficiency of GPM significantly. Further, the authors validate their presented scheme through plenty of simulations and the results obtained demonstrate that the proposed technique achieves much better efficiency as compared to existing path planning methods.

i. Heuristic searching methods

The well-known heuristic search method A-star algorithm is the advanced form of the typical Dijkstra algorithm. In the past few decades, the A-star has been significantly evolved and improved, hence lots of enhanced heuristic search methods have been derived from it. A modified A-star algorithm and an orographic database were used in [87] for searching the best track and building a digital map respectively. The heuristic A-star algorithm was used by the authors of [69], who first dissected the entire region into square mesh and then leveraged the A-star heuristic algorithm for finding the best path that is based on the value function of multiple points on the grid along the estimated path. In [82] authors proposed the sparse A-star search (SAS) for computing an optimal path, the presented heuristic search successfully minimizes the computational complexity by imposing additional impediments and limitations to the procedure of searching within space at the time of path planning. The D-star algorithm, another name for the dynamic A-star algorithm, was proposed and developed in [78] for computing an optimal path in a partially or completely unspecified dynamic environment. The algorithm keeps on improving and updating the map obtained from completely new and unknown environments and then amending the path on detection of new obstacles on its path. Authors of [91] proposed sampling-based path planning similar to rapidly exploring random trees (RRT) that can generate an optimal path free from collision when no information of the surrounding environment is provided initially. In [71] authors proposed and developed a 3D path planning solution for UAVs which makes them capable of figuring out a feasible, optimal, and collision-free path in complicated dynamic environments. In the proposed approach authors exploit a probabilistic graph in order to test allowable space without considering the existing obstacles. Whenever planning is required, then the A-star discrete search algorithm explores the generated probabilistic graph for obtaining an optimal collision-free path. Authors validate their proposed solution in the V-REP simulator and then incorporate it into a real-time UAV. As a kind of common obstacle in complex 3D environments, U-type obstacles might be responsible for confusing a UAV thus leading to a collision. Therefore, in order to overcome this limitation Zhang et al. (2019) [92] proposed and developed a state-of-the-art Ant-Colony Optimization (ACO)-based technique called Self Heuristic Ant (SHA) for the generation of the optimal, collision-free trajectory in unstructured 3D environments with solid U-type obstacles. In this approach authors first construct the whole space using a grid model of workspace and then a novel optimized method for path planning of UAV is designed. In order to present ACO deadlock, that is, trapping of ants in U-type obstacles in the absence of an ideal successor node, authors designed two different search approaches for selecting the succeeding path node. Furthermore, Self Heuristic Ant (SHA) is used for improving the efficacy of the ACO-based method. Finally, results obtained after conducting several deeply investigated experiments illustrate that the probability of deadlock state can be reduced to a great extent with the implementation of proposed search strategies.

ii. Intelligent algorithms

In the past few decades, researchers tried a lot to work out trajectory planning problems using intelligent algorithms and proposed a variety of intelligent searching techniques. However, the most renowned intelligent algorithms are the simulated anneal arithmetic (SAA) algorithm and genetic algorithm. In [93] authors used the SAA methods and genetic algorithm in the computation of an optimal path. Crossover and mutation operations of genetic algorithm and criterion of Metropolis are used for the evaluation of adaptation function of the path, thereby improving the effectiveness of trajectory planning. Authors of [2] proposed and developed an optimized global path planning technique using the improved conjugate direction method and the simulated annealing algorithm.

2) Local path planning

Local path planning methods exploit information of the local environment and state estimation of UAV to plan a collision-free local path dynamically. Path planning in dynamic environments might become computationally expensive due to some unexpected aspects, such as the motion of obstacles in the dynamic environments. In order to overcome this problem, algorithms for local path planning need to be feasible and adaptive to the changing attributes of the surrounding environment, by using valuable data such as the shape, size, and position about different sections of the environment obtained from multiple sensing devices.

Conventional local path planning techniques include artificial potential field methods, neural network methods, fuzzy logic methods, and spatial search methods, etc. Some general methods for local path planning are discussed below. In [81] authors proposed an artificial field method to navigate a robot from the local environment into the domain of a metaphysical artificial potential field. The final point has the “attraction” as well as an object with “repulsion” to the navigating robot, hence these two forces are responsible for moving the robot towards the target location. An example of the usage of the artificial gravitational field method for computing the local path through the threat area of radar was given in [10]. Genetic algorithms can provide a typical framework for solving typical and complex problems of optimization, especially the ones that are related to the computation of an optimal path. These algorithms are inspired by the evolution and inheritance concepts of biological phenomena. Problems should be solved by leveraging the principle of “survival of the fittest and survival competition,” in order to obtain the best solution. A genetic algorithm consists of five main components: initial population, chromosome coding, genetic operation, fitness function, and control parameters. In [65] authors proposed a path planning solution based on a genetic algorithm for an aircraft. Under the confession of biological functions, neural networks are established as computational methods. An example of a local path planning technique implemented using Hopfield networks was given in [25]. In [64] researchers proposed an ant colony algorithm which is a new type of binocular algorithm inspired by the ant activity. It is a probabilistic optimization method that resembles the behavioral attributes of ants. This technique could accomplish good results by solving a series of onerous combinatorial optimizations. Table 3 provides an outline of different methods discussed for the task of path planning.

B. Vision-based control

Shirai and Inoue(1973) [104] introduced the concept of vision-based control in robotics in which the data collected from visual sensors is used for manipulating and controlling the motion of the autonomous robot. In that era, the performance of autonomous robots based on vision-based control was not achieved up to the desire, mainly due to the issue of extracting information from visual sensors. Since 1990, with the reasonable improvement in the computing power of personal computers, there has been a high growth of research and development in the field of computer vision-based techniques for robotics applications. After that, the vision-based control for unmanned aerial vehicles had been considered as an important research problem that needed to be worked upon and till now this is under rapid development with the objective of attaining complete autonomy in UAVs. The several challenges that still exist in the vision based control of UAVs are obstacle recognition, collision avoidance, and delay in the transfer of information to the UAV regarding action that it needs to take. In the past few years, in order to overcome these limitations several strategies and solutions based on reinforcement learning, visual-inertial techniques, and hybrid approaches have been proposed. Few of them are mentioned here.

Kendoul et al. (2008)[105] proposed an adaptive controller based vision-based control approach in which the UAV is capable of hovering at a certain height and tracking a given trajectory autonomously. The measurements from the inertial measurement unit are merged with the optical flow data for vehicle's pose estimation and prediction of depth map with unknown scale factor use in obstacle detection. Then, finally the presented controller integrates all the measurements according to the control algorithm and provides the necessary commands to the UAV for autonomous navigation. In the progress of their previous work the Kendoul et al. (2009)[117] proposed a real time vision-based control for the UAV. The data from the downward looking camera equipped in the UAV was integrated with IMU measurements through an EKF, and then the obtained data was integrated with the designed non-linear controller for achieving the desired performance control of the UAV. Similarly, Carillo et al. (2012)[106] has developed a quadrotor UAV capable of carrying out autonomous take-off, navigation, and landing. The stereo camera and IMU sensors are installed in the UAV. The data from these sensors is merged using a Kalman filter for improving the accuracy of the quadrotor state estimation. The stereo visual odometry technique is used for determining the 3D motion of the camera in the environment. In order to leverage the integration of the data obtained through depth camera and inertial unit in controlling the UAV. Steggano et al. (2013)[107] proposed the design of a UAV platform equipped with a RGB-D camera and IMU sensors to achieve autonomous vision-based control and stabilization, especially in the teleoperations. And, the integration of the information from IMU and RGB-D sensors is used for the estimation of the UAV velocity, which is then used to control the UAV. Later on, the technique that utilized only a monocular camera for obtaining the control commands for UAV was proposed by Mannar et al. (2018)[108]. The authors of this work proposed and developed a vision-based control system to avoid aerial obstacles in a forest environment for a UAV. The proposed algorithm is the enhancement of the existing algorithm initially proposed by Michel [109] for vision-based obstacle avoidance in ground vehicles. The authors have used this algorithm for UAV control for the first time. The idea behind this algorithm is that the images captured by the monocular camera of UAV are first divided into various longitudinal strips then from each strip the texture features are extracted and the weighted combination of these texture features is used for the estimation of distance to the nearest obstacle. The weights of the prediction network are pre-computed at the time of supervised training of the network on the correspondences of the features to the ground truth distance extracted from the image frames captured from a simulated forest environment. Then, corresponding to the longitudinal strip with minimum distance to the UAV, the real angle is calculated which is then processed further to obtain the appropriate yaw/velocity commands for UAV control. Subsequently, Hu and Wang (2018)[110] proposed the hand-gesture based control system for a UAV. The deep learning method is trained on the dataset of gesture inputs to predict a suitable command for UAV control. First, the random hand gestures generated by the user are fed as an input input to the leap motion controller which consists of 2 optical sensors and 3 infrared lights and it is responsible for converting the dynamic gestures information to the hand skeleton data and provide this skeleton data to the data preprocessing unit which performs the task of feature selection and scaling on the output of the leap motion controller and creates the feature vectors. The composed feature vectors are delivered to the deep learning module which recognize the gesture and predict the appropriate command. Finally, the predicted control command is delivered to the UAV control module and it processes the prediction to extract the movement command and send it to the UAV over WiFi. In recent years, some researchers have also introduced reinforcement learning based strategies in the vision-based control problem of aerial vehicles in dynamic environments. Following this approach, Li et al. (2018)[111] proposed a machine learning based strategy for the autonomous tracking of a given target. The strategy combines UAV perception and control for following a specific target. The proposed approach leverages the integration of model-free policy gradient method and a PID controller for achieving stabilized navigation. The deep convolutional neural network is trained on the set of raw images for the classification of target and then model-free reinforcement learning technique is used for the generation of high level control actions. The high level actions are then converted to the low level UAV control commands by a PID controller which is then finally transferred to the quadrotor for real world navigation and control. Table 4 summarizes the above discussed vision-based UAV control methods.

Table 3: Summary of important methods in path planning

Types of Path planning	**Methods used**	**References**
Global	Potential field	[38],[90]
	Optimization	[46],[74]
	Intelligent	[2],[93]
	Heuristic search	[87],[82],[78],[91],[92]
Local	Hopfield	[25]
	Artificial potential field	[81],[10]

	Ant colony optimization	[64]

Table 4: Brief description of discussed vision-based control methods for UAV.

UAV state estimation techniques	**Control methods used**	**References**
Optical flow	Adaptive controller	[105],[107]
Extended Kalman Filter	Nonlinear Controller	[117],[106]
Deep learning	Feature-based control	[108]
	Hand-gesture based control	[110]
Reinforcement learning	PID controller	[111]

V. Real-time experiments with autonomous UAV

In this section we will describe some real time applications of vision-based autonomous UAV navigation in which researchers have designed a fully autonomous UAV system capable of navigating autonomously in 3-D dynamic environments and can be deployed for real-world applications like search and rescue, surveillance, target tracking, and for delivery products to the customers. Hulens et al. (2017)[112] designed a UAV system for autonomous navigation in a fruit orchard using an algorithm that estimates the centre point and end point of the orchard lane and then provides the necessary stabilization to the UAV by keeping it in the centre of the orchard and abstain it from colliding with the trees and walls of the orchard autonomously without any connection with the ground station. In order to attain complete autonomy the UAV is equipped with a front camera and on-board processing unit. The image stream from the camera passes to the on-board processing unit which processes the images and uses the obtained information for controlling the UAV and avoiding collision with surrounding obstacles. In order to keep the UAV in the centre of the orchard lane the authors proposed an algorithm that can estimate the vanishing point of the lane in which the UAV is navigating. Later on, in order to deploy the autonomous UAV in emergency services, Mittal et al. (2019)[113] proposed and developed a vision-based autonomous navigation strategy for a UAV to carry out the operation of search and rescue in urban areas after disastrous calamities like earthquakes, landslides, and floods, etc. In the proposed mechanism the UAV which is used for navigation in post disaster areas is equipped with several sensors like bioradar, GPS, IMU, barometer, stereo camera, and attitude controller. The proposed strategy consists of four steps: localization, mapping, detection of landing site, and landing trajectory estimation. In order to localize the UAV in previously unknown environments a robot-centric visual-inertial framework, robust visual-inertial odometry (ROVIO) is used. The data from the downward looking camera of the UAV and IMU are fed to the ROVIO module for pose estimation. Furthermore, in order to avoid drift the state of the UAV estimated by the ROVIO unit is fused with data from other sensors such as barometer and GPS through Extended Kalman Filter (EKF). Then, the 3-D map of the environment is created using the depth obtained through the stereo camera of the UAV and the planning in the obtained occupancy grid is done using sampling based path planning algorithm, that is, rapidly exploring random trees (RRT) algorithm. Then, the next step is to find a flat platform free from obstacles upon which the UAV can land. This step is completed with the help of cost maps which are completed using the estimated pose and the depth map obtained from the stereo camera of the UAV. Then finally, collision free and minimum-jerk landing trajectory is designed using the RRT-star algorithm. Subsequently, Lin and Peng (2021)[114] proposed a real-time vision-based autonomous navigation approach for exploring outdoor environments. In this work, the authors leverage the static map-based offline path planning for generating an initial path using RRT for the UAV to follow and an online optical flow based method for avoiding dynamic obstacles. First the stream of images captured by the monocular camera of the UAV is fed to the preprocessing module which is responsible for performing tasks like image down-sampling, grayscaling, and noise removal, etc. Then, a red bounding-box, defined as the region of interest (region with motion vectors) is computed. After the determination of the region of interest the optical flow in the sequence of captured images is used to determine the motion vectors for obstacle detection and avoidance. These algorithms are implemented on the single onboard computer which is the control platform for real-time processing of this problem. However, several vision-based strategies for real-time autonomous UAV navigation suffer from various illumination factors. As, in low illumination areas it would be very difficult for vision based algorithms to obtain the depth map of the environment for estimating the distance of the UAV from obstacles and for detecting the landing platform. In order to solve these issues, Lin et al. (2021)[115] presented a vision based system that can detect landing markers in low illumination areas for safe and efficient landing of UAVs. To improve the resolution quality and luminance of captured images and then the hierarchical learning based method consists of a rule-based classifier, decision tree and Convolutional neural networks for the localization of landing marker, the extracted feature information of the marker is used for state estimation and controlling the UAV during landing. The proposed scheme is verified with the real-time UAV during the nighttime where the quadrotor is required to land on the marker placed in the field.

VI. CONCLUSIONS

This paper presented a discussion on autonomous vision-based UAV navigation mainly from three facets: localization and mapping, obstacle avoidance, and path planning. Localization and mapping are the core of autonomous UAV navigation, which are responsible for providing information about the environment and location. Collision avoidance and path planning are vital in order to make the UAV capable of reaching the target location safely and quickly without any collision. Several visual SLAM algorithms have been developed by the computer vision society, still, many of them cannot be leveraged directly for navigation of UAVs due to shortcomings posed by their processing power and their architecture.

Nonetheless, UAVs share similar navigation solutions with mobile ground robots but still, researchers are facing many challenges in the implementation of vision-based autonomous UAV navigation. The aerial vehicle is required to process a variety of sensors' data in order to achieve safe and steady flight in real-time, especially the processing of visual data which increases the computational cost to a great extent. Thus, autonomous navigation of UAVs under limited consumption of power and computing resources has become a major challenge in the research field. It should also be noted that in contrast to ground vehicles UAVs are not able to just stop navigating in the state of great uncertainty, that is generation of incoherent commands can make the UAV unstable. Along with this, UAV could exhibit unpredictable behavior whenever the computational requirements are not sufficient to update attitude and velocity in time or in the case of hardware-mechanical failure. Therefore, researchers must put efforts in developing computer vision algorithms that possess the capability of responding quickly to the dynamic behavior of the environment. The development of such algorithms will help in improving the native ability of UAVs to navigate smoothly in various attitudes and orientations with sudden appearance and disappearance of targets and obstacles. Furthermore, autonomous UAV navigation needs a local or global 3D representation of the surrounding environments, thereby increasing the computation and storage overhead. Hence, there are many challenges for long-time UAV navigation in complex environments. Besides that, tracking and localization collapse can happen in the course of UAV flight due to obscure motion caused by fast rotation and movement. Algorithms used for tracking objects should be robust against illumination, vehicle disturbances, noise, and occlusions. Otherwise, it will be very difficult for the tracker to figure out the trajectory of the target, and to operate in consonance with the controllers of UAV. Hence, highly sophisticated, erudite, and robust control schemes must exist for closing the loop optimally using visual data. Therefore, we are expecting research on loop detection and relocalization of UAV in the near future. In addition to this, we also found that a partial or complete 3D map of the environment is not enough to figure out an obstacle-free path along with the optimization of the energy consumption or the length of the resulting path. In contrast to 2D path planning, the challenges, and difficulties in 3D map construction of the environment increase rapidly as the intricacy of varying impediments and kinematics of UAVs increase. Thus, no efficient solutions to this NP-hard problem exist, even contemporary path planning algorithms undergo the same problem of local minimum. Therefore, researchers are making a lot of efforts to discover a more efficient and robust algorithm for achieving global path optimization. The research work presented in this survey illustrates that few techniques are proved experimentally but many of the vision-based SLAM and obstacle avoidance techniques are not yet fully incorporated in the navigation controllers of autonomous UAVs, since the demonstrated methods either operate under some constraints in simple environments or their working is proved only through simulations. Therefore, influential engineering is necessary to move the current state-of-the-art a step ahead and for the evaluation of their performance in real-world environments. Another crucial finding from this review is that most of the flight tests discussed in the presented work were carried out on small commercial UAVs with increased payload for onboard processing units and multiple sensors. Yet, it can be understood from this discussion that trending research is focusing more on micro aerial vehicles that are capable of navigating indoors, outdoors and inspecting and maintaining target infrastructure utilizing their acrobatic maneuvering competencies.

REFERENCES

[1] Al-Kaff, A., Q. Meng, D. Martín, A. de la Escalera, and J.M. Armingol. 2016. "Monocular Vision-based Obstacle Detection/Avoidance for Unmanned Sweden, June 19-22.

[2] Andert, F., and F. Adolf. 2009. "Online World Modeling and Path Planning for an Unmanned Helicopter." Autonomous Robots 27 (3):147–164.

[3] Araar, O., and N. Aouf. 2014. "A New Hybrid Approach for the Visual Servoing of VTOL UAVs from Unknown Geometries." Paper Presented at the Control and Automation, 22nd Mediterranean Conference of, Palermo, Italy,June 16–19.

[4] Aulinas, J., Y. R. Petillot, J. Salvi, and X. Lladó. 2008. "The SLAM Problem: A Survey." Conference on Artificial Intelligence Research and Development, 11th International Conference of the Catalan Association for Artificial Intelligence, Spain, October 22–24.

[5] Bai, G., X. Xiang, H. Zhu, D. Yin, and L. Zhu. 2015. "Research on Obstacles Avoidance Technology for UAV Based on Improved PTAM Algorithms." Paper Presented at the Progress in Informatics and Computing IEEE International Conference on, Nanjing, China, December 18-20.

[6] Belmonte, L. M., Morales, R., & Fernández-Caballero, A. (2019). Computer Vision in Autonomous Unmanned Aerial Vehicles—A Systematic Mapping Study. Applied Sciences, 9(15), 3196. doi:10.3390/app9153196.

[7] Bertrand, O. J., J. P. Lindemann, and M. Egelhaaf. 2015. "A Bio-Inspired Collision Avoidance Model Based on Spatial Information Derived from Motion Detectors Leads to Common Routes." PLoS Computational Biology 11 (11):e1004339.

[8] Bianchi, M., & Barfoot, T. D. (2021). UAV Localization Using Autoencoded Satellite Images. IEEE Robotics and Automation Letters, 6(2), 1761-1768. doi:10.1109/lra.2021.3060397.

[9] Bonin-Font, F., A. Ortiz, and G. Oliver. 2008. "Visual Navigation for Mobile Robots: A Survey." Journal of Intelligent and Robotic Systems 53 (3): 263–296.

[10] Bortoff, S. A. 2000. "Path Planning for UAVs." Paper Presented at the American Control Conference, IEEE, Chicago, IL, USA, June 8–30.

[11] Carrillo, L. R. G., A. E. D. López, R. Lozano, and C. Pégard. 2012. "Combining Stereo Vision and Inertial Navigation System for a Quad-Rotor UAV." Journal of Intelligent & Robotic Systems 65 (1–4): 373–387.

[12] Celik, K., S. J. Chung, M. Clausman, and A. K. Somani. 2009. "Monocular Vision SLAM for Indoor Aerial Vehicles." Paper Presented at the Intelligent Robots and Systems, IEEE/RSJ International Conference on, St. Louis, USA, October 10–15.

[13] Cho, D., P. Tsiotras, G. Zhang, and M. Holzinger. 2013. "Robust Feature Detection, Acquisition and Tracking for Relative Navigation in Space with a Known Target." Paper Presented at the AIAA Guidance, Navigation, and Control Conference, Boston, MA, August 19–22.

[14] Davison, A. J. 2003. "Real-Time Simultaneous Localisation and Mapping with a Single-Camera." Paper Presented at the Computer Vision, IEEE International Conference on, Nice, France, October 13–16.

[15] Desouza, G. N., and A. C. Kak. 2002. "Vision for Mobile Robot Navigation: A Survey." IEEE Transactions on Pattern Analysis and Machine Intelligence 24 (2): 237–267.

[16] Dryanovski, I., W. Morris, and J. Xiao. 2010. "Multi-Volume Occupancy Grids: An Efficient Probabilistic 3D Mapping Model for Micro Aerial Vehicles." Paper Presented at the Intelligent Robots and Systems, IEEE/RSJ International Conference on, Taipei, China, October 18–22.

[17] Elmokadem, T., & Savkin, A. V. (2021). A Hybrid Approach for Autonomous Collision-Free UAV Navigation in 3D Partially Unknown Dynamic Environments. Drones, 5(3), 57. doi:10.3390/drones5030057.

[18] Engel, J., T. Schöps, and D. Cremers. 2014. "LSD-SLAM: Large-scale Direct Monocular SLAM." Paper presented at the European Conference on Computer Vision, Zurich, Switzerland, September 6–12, 834–849.

[19] Esrafilian, O., and H. D. Taghirad. 2016. "Autonomous Flight and Obstacle Avoidance of a Quadrotor by Monocular SLAM." Paper Presented at the Robotics and Mechatronics, 4th IEEE International Conference on, Tehran, Iran, October 26–28.

[20] Forster, C., M. Pizzoli, and D. Scaramuzza. 2014. "SVO: Fast Semi-direct Monocular Visual Odometry." Paper Presented at the Robotics and Automation, IEEE International Conference on, Hong Kong, China, May 31–June 7.

[21] Forster, C., M. Faessler, F. Fontana, M. Werlberger, and D. Scaramuzza. 2015. "Continuous On-board Monocular-vision-based Elevation Mapping Applied to Autonomous Landing of Micro Aerial Vehicles." Paper Presented at the Robotics and Automation, IEEE International Conference on, Seattle, USA, May 26–30.

[22] Fournier, J., B. Ricard, and D. Laurendeau. 2007. "Mapping and Exploration of Complex Environments Using Persistent 3D Model." Paper Presented at the Computer and Robot Vision, Fourth Canadian Conference on, IEEE, Montreal, Canada, May 28–30.

[23] Gageik, N., P. Benz, and S. Montenegro. 2015. "Obstacle Detection and Collision Avoidance for a UAV with Complementary Low-Cost Sensors." IEEE Access 3: 599–609.

[24] Gaspar, J., N. Winters, and J. Santos-Victor. 2000. "Vision-Based Navigation and Environmental Representations with an Omnidirectional Camera." IEEE Transactions on Robotics and Automation 16 (6): 890–898.

[25] Gilmore, J. F., and A. J. Czuchry. 1992. "A Neural Network Model for Route Planning Constraint Integration." Paper Presented at the Neural Networks, IEEE International Joint Conference on, Baltimore, MD, USA, June 7–11.

[26] Gosiewski, Z., J. Ciesluk, and L. Ambroziak. 2011. "Vision-Based Obstacle Avoidance for Unmanned Aerial Vehicles." Paper Presented at the Image and Signal Processing, 4th IEEE International Congress on, Shanghai, China, October 15–17.

[27] Gutmann, J., M. Fukuchi, and M. Fujita. 2008. "3D Perception and Environment Map Generation for Humanoid Robot Navigation." The International Journal of Robotics Research 27 (10): 1117–1134.

[28] Haag, J., W. Denk, and A. Borst. 2004. "Fly Motion Vision is Based on Reichardt Detectors regardless of the Signal-to-Noise Ratio." Proceedings of the National Academy of Sciences of the United States of America 101 (46): 16333–16338.

[29] Han, J., and Y. Chen. 2014. "Multiple UAV Formations for Cooperative Source Seeking and Contour Mapping of a Radiative Signal Field." Journal of Intelligent & Robotic Systems 74 (1–2): 323–332.

[30] Harmat, A., M. Trentini, and I. Sharf. 2015. "Multi-camera Tracking and Mapping for Unmanned Aerial Vehicles in Unstructured Environments." Journal of Intelligent & Robotic Systems 78 (2): 291–317.

[31] Herissé, B., T. Hamel, R. Mahony, and F. Russotto. 2012. "Landing a VTOL Unmanned Aerial Vehicle on a Moving Platform Using Optical Flow." IEEE Transactions on Robotics 28 (1): 77–89.

[32] Herwitz, S. R., L. F. Johnson, S. E. Dunagan, R. G. Higgins, D. V. Sullivan, J. Zheng, and B. M. Lobitz. 2004. "Imaging from an Unmanned Aerial Vehicle: Agricultural Surveillance and Decision Support." Computers and Electronics in Agriculture 44 (1): 49–61.

[33] Horn, B. K., and B. G. Schunck. 1981. "Determining Optical Flow." Artificial Intelligence 17 (1–3): 185–203. Hornung, A., K. M. Wurm, M. Bennewitz, C. Stachniss, and W. Burgard. 2013. "OctoMap: An Efficient Probabilistic 3D Mapping Framework Based on Octrees." Autonomous Robots 34 (3): 189–206.

[34] How, J. P., B. Behihke, A. Frank, D. Dale, and J. Vian. 2008. "Real-time Indoor Autonomous Vehicle Test Environment." IEEE Control Systems 28 (2): 51–64.

[35] Hrabar, S. 2008. "3D Path Planning and Stereo-based Obstacle Avoidance for Rotorcraft UAVs." Paper Presented at the Intelligent Robots and Systems, IEEE/RSJ International Conference on, Nice, France, September 22–26.

[36] Hu, F., & Wu, G. (2020). Distributed Error Correction of EKF Algorithm in Multi-Sensor Fusion Localization Model. IEEE Access, 8, 93211-93218. doi:10.1109/access.2020.2995170.

[37] Hui, X., Bian, J., Zhao, X., & Tan, M. (2018). Vision-based autonomous navigation approach for unmanned aerial vehicle transmission-line inspection. International Journal of Advanced Robotic Systems, 15(1), 172988141775282. doi:10.1177/1729881417752821.

[38] Jayaweera, H. M., & Hanoun, S. (2020). A Dynamic Artificial Potential Field (D-APF) UAV Path Planning Technique for Following Ground Moving Targets. IEEE Access, 8, 192760-192776. doi:10.1109/access.2020.3032929.

[39] Jones, E. S. 2009. "Large Scale Visual Navigation and Community Map Building." PhD diss. University of California at Los Angeles.

[40] Kanellakis, C., & Nikolakopoulos, G. (2017). Survey on Computer Vision for UAVs: Current Developments and Trends. Journal of Intelligent & Robotic Systems, 87(1), 141-168. doi:10.1007/s10846-017-0483-z.

[41] Klein, G., and D. Murray. 2007. "Parallel Tracking and Mapping for Small AR Workspaces." Paper Presented at the Mixed and Augmented Reality, 6th IEEE and ACM International Symposium on, Nara, Japan, November 13–16.

[42] Krul, S., Pantos, C., Frangulea, M., & Valente, J. (2021). Visual SLAM for Indoor Livestock and Farming Using a Small Drone with a Monocular Camera: A Feasibility Study. Drones, 5(2), 41. doi:10.3390/drones502004.

[43] Kümmerle, R., G. Grisetti, H. Strasdat, K. Konolige, and W. Burgard. 2011. "g 2 o: A General Framework for Graph Optimization." Paper Presented at the Robotics and Automation, IEEE International Conference on, Shanghai, China, May 9–13.

[44] Langelaan, J. W., and N. Roy. 2009. "Enabling New Missions for Robotic Aircraft." Science 326 (5960): 1642–1644.

[45] Leutenegger, S., S. Lynen, M. Bosse, R. Siegwart, and P. Furgale. 2015. "Keyframe-Based Visual–Inertial Odometry Using Nonlinear Optimization." The International Journal of Robotics Research 34 (3): 314–334.

[46] Li, B., Qi, X., Yu, B., & Liu, L. (2020). Trajectory Planning for UAV Based on Improved ACO Algorithm. IEEE Access, 8, 2995-3006. doi:10.1109/access.2019.2962340.

[47] Li, J., and N. M. Allinson. 2008. "A Comprehensive Review of Current Local Features for Computer Vision." Neurocomputing 71 (10): 1771–1787.

[48] Li, H., and S. X. Yang. 2003. "A Behavior-based Mobile Robot with a Visual Landmark-recognition System." IEEE/ASME Transactions on Mechatronics 8 (3): 390–400.

[49] Lindqvist, B., Mansouri, S. S., Agha-Mohammadi, A., & Nikolakopoulos, G. (2020). Nonlinear MPC for Collision Avoidance and Control of UAVs With Dynamic Obstacles. IEEE Robotics and Automation Letters, 5(4), 6001-6008. doi:10.1109/lra.2020.3010730.

[50] Lucas, B. D., and T. Kanade. 1981. "An Iterative Image Registration Technique with an Application to Stereo Vision." Paper Presented at the DARPA Image Understanding Workshop, 7th International Joint Conference on Artificial Intelligence, Vancouver, Canada, August 24–28.

[51] Lynen, S., M. W. Achtelik, S. Weiss, M. Chli, and R. Siegwart. 2013. "A Robust and Modular Multi-Sensor Fusion Approach Applied to Mav Navigation." Paper Presented at the Intelligent Robots and Systems, IEEE/RSJ International Conference on, Tokyo, Japan, November 3–7.

[52] Magree, D., and E. N. Johnson. 2014. "Combined Laser and Vision-Aided Inertial Navigation for an Indoor Unmanned Aerial Vehicle." Paper Presented at the American Control Conference, IEEE, Portland, USA, June 4–6.

[53] Mahon, I., S. B. Williams, O. Pizarro, and M. Johnson Roberson. 2008. "Efficient View-based SLAM Using Visual Loop Closures." IEEE Transactions on Robotics 24 (5): 1002–1014.

[54] Maier, J., and M. Humenberger. 2013. "Movement Detection Based on Dense Optical Flow for Unmanned Aerial Vehicles." International Journal of Advanced Robotic Systems 10 (2): 146.

[55] Márquez-Gámez, D. 2012. "Towards Visual Navigation in Dynamic and Unknown Environment: Trajectory Learning and following, with Detection and Tracking of Moving Objects." PhD diss., l'Institut National des Sciences Appliquées de Toulouse. Matsumoto.

[56] Y., K. Ikeda, M. Inaba, and H. Inoue. 1999. "Visual Navigation Using Omnidirectional View Sequence." Paper Presented at the Intelligent Robots and Systems, IEEE/RSJ International Conference on, Kyongju, South Korea, October 17 –21.

[57] Maza, I., F. Caballero, J. Capitán, J. R. Martínez-de-Dios, and A. Ollero. 2011. "Experimental Results in Multi-UAV Coordination for Disaster Management and Civil Security Applications." Journal of Intelligent & Robotic Systems 61 (1): 563–585.

[58] Mirowski, P., R. Pascanu, F. Viola, H. Soyer, A. Ballard, A. Banino, and M. Denil. 2017. "Learning to Navigate in Complex Environments." The 5th International Conference on Learning Representations, Toulon, France, April 24–26.

[59] Moravec, H. P. 1983. "The Stanford Cart and the CMU Rover." Proceedings of the IEEE 71 (7): 872–884.

[60] Moreno-Armendáriz, M. A., and H. Calvo. 2014. "Visual SLAM and Obstacle Avoidance in Real-Time for Mobile Robots Navigation." Paper Presented at the Mechatronics, Electronics and Automotive Engineering (ICMEAE), IEEE International Conference on, Cuernavaca, Mexico, November 18–21.

[61] Newcombe, R. A., S. J. Lovegrove, and A. J. Davison. 2011. "DTAM: Dense Tracking and Mapping in Real-time." Paper Presented at the Computer Vision (ICCV), IEEE International Conference on, Washington, DC, USA, November 6–13.

[62] Nourani-Vatani, N., P. Vinicius, K. Borges, J. M. Roberts, and M. V. Srinivasan. 2014. "On the Use of Optical Flow for Scene Change Detection and Description." Journal of Intelligent & Robotic Systems 74 (3–4): 817.

[63] Padhy, R.M., Choudhry, S.K., Sa, P.K., & Bakshi, S. (2019). Obstacle Avoidance for Unmanned Aerial Vehicles. IEEE Consumer Electronics Magazine, 19, 2162-2248. doi: 10.1109/MCE.2019.2892280.

[64] Parunak, H. V., M. Purcell, and R. O'Connell. 2002. "Digital Pheromones for Autonomous Coordination of Swarming UAVs." Paper Presented at the 1st UAV Conference, Portsmouth, Virginia, May 20–23.

[65] Pellazar, M. B. 1998. "Vehicle Route Planning with Constraints Using Genetic Algorithms". Paper Presented at the Aerospace and Electronics Conference, NAECON, 1998, IEEE National, Dayton, USA, July 17.

[66] Ranftl, R., V. Vineet, Q. Chen, and V. Koltun. 2016. "Dense Monocular Depth Estimation in Complex Dynamic Scenes." Paper Presented at the Computer Vision and Pattern Recognition, IEEE Conference on, Las Vegas, USA, June 27–30.

[67] Roberge, V., M. Tarbouchi, and G. Labonté. 2013. "Comparison of Parallel Genetic Algorithm and Particle Swarm Optimization for Real-time UAV Path Planning." IEEE Transactions on Industrial Informatics 9 (1): 132–141.

[68] Rogers, B., and M. Graham. 1979. "Motion Parallax as an Independent Cue for Depth Perception." Perception 8 (2): 125–134.

[69] Rouse, D. M. 1989. "Route Planning Using Pattern Classification and Search Techniques." Aerospace and Electronics Conference, IEEE National, Dayton, USA, May 22–26.

[70] Ruffier, F., S. Viollet, S. Amic, and N. Franceschini. 2003. "Bio-inspired Optical Flow Circuits for the Visual Guidance of Micro Air Vehicles." Paper Presented at the Circuits and Systems, IEEE International Symposium on, Bangkok, Thailand, May 25–28.

[71] Sanchez-Lopez, J. L., Wang, M., Olivares-Mendez, M. A., Molina, M., & Voos, H. (2018). A Real-Time 3D Path Planning Solution for Collision-Free Navigation of Multirotor Aerial Robots in Dynamic Environments. Journal of Intelligent & Robotic Systems, 93(1-2), 33-53. doi:10.1007/s10846-018-0809-5.

[72] Santos-Victor, J., G. Sandini, F. Curotto, and S. Garibaldi. 1993. "Divergent Stereo for Robot Navigation: Learning from Bees." Paper Presented at the Computer Vision and Pattern Recognition, IEEE Computer Society Conference on, New York, USA, June 15–17.

[73] Seitz, S. M., B. Curless, J. Diebel, D. Scharstein, and R. Szeliski. 2006. "A Comparison and Evaluation of Multi-view Stereo Reconstruction Algorithms." Paper Presented at the Computer Vision and Pattern Recognition, IEEE Computer Society Conference on, New York, USA, June 17-22.

[74] Shao, S., He, C., Zhao, Y., & Wu, X. (2021). Efficient Trajectory Planning for UAVs Using Hierarchical Optimization. IEEE Access, 9, 60668-60681. doi:10.1109/access.2021.3073420.

[75] Shen, S., Y. Mulgaonkar, N. Michael, and V. Kumar. 2014. "Multi-Sensor Fusion for Robust Autonomous Flight in Indoor and Outdoor Environments with a Rotorcraft MAV." Paper Presented at the Robotics and Automation (ICRA), IEEE International Conference on, Hong Kong, China, May 31–June 7.

[76] Silveira, G., E. Malis, and P. Rives. 2008. "An Efficient Direct Approach to Visual SLAM." IEEE Transactions on Robotics 24 (5): 969–979.

[77] Srinivasan, M. V., and R. L. Gregory. 1992. "How Bees Exploit Optic Flow: Behavioural Experiments and Neural Models [and Discussion]." Philosophical Transactions of the Royal Society of London B: Biological Sciences 337 (1281): 253–259.

[78] Stentz, A. 1994. "Optimal and Efficient Path Planning for Partially-Known Environments." Paper Presented at the IEEE International Conference on Robotics and Automation, San Diego, CA, USA, May 8–13.

[79] Strasdat, H., J. M. Montiel, and A. J. Davison. 2012. "Visual SLAM: Why Filter?" Image and Vision Computing 30 (2):65–77.

[80] Strohmeier, M., M. Schäfer, V. Lenders, and I. Martinovic. 2014. "Realities and Challenges of Nextgen Air Traffic Management: The Case of ADS-B." IEEE Communications Magazine 52 (5): 111–118.

[81] Sugihara, K., and I. Suzuki. 1996. "Distributed Algorithms for Formation of Geometric Patterns with Many Mobile Robots." Journal of Robotic Systems 13 (3): 127–139.

[82] Szczerba, R. J., P. Galkowski, I. S. Glicktein, and N. Ternullo. 2000. "Robust Algorithm for Real-time Route Planning." IEEE Transactions on Aerospace and Electronic Systems 36 (3): 869–878.

[83] Szeliski, R. 2010. Computer Vision: Algorithms and Applications. London: Springer Science & Business Media.

[84] Szenher, M. D. 2008. "Visual Homing in Dynamic Indoor Environments." Ph.D. diss. The University of Edinburgh.

[85] Trujillo, J., Munguia, R., Urzua, S., Guerra, E., & Grau, A. (2020). Monocular Visual SLAM Based on a Cooperative UAV–Target System. Sensors, 20(12), 3531. doi:10.3390/s20123531.

[86] Tuytelaars, T., and K. Mikolajczyk. 2008. "Local Invariant Feature Detectors: A Survey." Foundations and Trends in Computer Graphics and Vision 3 (3): 177–280.

[87] Vachtsevanos, G., W. Kim, S. Al-Hasan, F. Rufus, M. Simon, D. Shrage, and J. Prasad. 1997. "Autonomous Vehicles: From Flight Control to Mission Planning Using Fuzzy Logic Techniques." Paper Presented at the Digital Signal Processing, 13th International Conference on, IEEE, Santorini, Greece, July 2–4.

[88] Valgaerts, L., A. Bruhn, M. Mainberger, and J. Weickert. 2012. "Dense versus Sparse Approaches for Estimating the Fundamental Matrix." International Journal of Computer Vision 96 (2): 212–234.

[89] Wang, C., Ma, H., Chen, W., Liu, L., & Meng, M. Q. (2020). Efficient Autonomous Exploration With Incrementally Built Topological Map in 3-D Environments. IEEE Transactions on Instrumentation and Measurement, 69(12), 9853-9865. doi:10.1109/tim.2020.3001816.

[90] Yathirajam, B., S.M, V. & C.M, A. (2020). Obstacle Avoidance for Unmanned Air Vehicles Using Monocular-SLAM with Chain-Based Path Planning in GPS Denied Environments. Journal of Aerospace System Engineering, 14(2), 1-11.

[91] Yershova, A., L. Jaillet, T. Simeon, and S. M. Lavalle. 2005. "Dynamic-Domain RRTs: Efficient Exploration by Controlling the Sampling Domain." Paper Presented at the IEEE International Conference on Robotics and Automation, Barcelona, Spain, April 18–22.

[92] Zhang, C., Hu, C., Feng, J., Liu, Z., Zhou, Y., & Zhang, Z. (2019). A Self-Heuristic Ant-Based Method for Path Planning of Unmanned Aerial Vehicle in Complex 3-D Space With Dense U-Type Obstacles. IEEE Access, 7, 150775-150791. doi:10.1109/access.2019.2946448.

[93] Zhang, Q., J. Ma, and Q. Liu. 2012. "Path Planning Based Quadtree Representation for Mobile Robot Using Hybrid-Simulated Annealing and Ant Colony Optimization Algorithm." Paper Presented at The Intelligent Control and Automation (WCICA), 10th World Congress on, IEEE, Beijing, China, July 6–8.

[94] González-Jorge, H. et al., 2017. Unmanned aerial systems for Civil Applications: A Review. *Drones*, 1(1), p.2.

[95] Lu, Y. et al., 2018. A survey on vision-based UAV navigation. *Geo-spatial Information Science*, 21(1), pp.21–32.

[96] T. P. Lillicrap, J. J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D. Silver, and D. Wierstra, "Continuous control with deep reinforcement learning," in ICLR, 2015.

[97] T. Haarnoja, A. Zhou, P. Abbeel, and S. Levine, "Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor," in ICML, vol. 80, 2018, pp. 1861–1870.

[98] LI, M. & HU, T., 2021. Deep learning enabled localization for UAV autolanding. *Chinese Journal of Aeronautics*, 34(5), pp.585–600.

[99] Afifi, G. & Gadallah, Y., 2021. Autonomous 3-D UAV localization using cellular networks: Deep supervised learning versus reinforcement learning approaches. *IEEE Access*, 9, pp.155234–155248.

[100] H. Goforth and S. Lucey, "GPS-denied UAV localization using pre-existing satellite imagery," Proceedings - IEEE International Conference on Robotics and Automation, vol. 2019-May, pp. 2974–2980, 2019.

[101] H. Hou, Q. Xu, C. Lan, W. Lu, Y. Zhang, Z. Cui, and J. Qin, "UAV Pose Estimation in GNSS-Denied Environment Assisted by Satellite Imagery Deep Learning Features," IEEE Access, vol. 9, pp. 6358–6367, 2020.

[102] Guo, J. et al., 2021. Three-dimensional autonomous obstacle avoidance algorithm for UAV based on circular arc trajectory. *International Journal of Aerospace Engineering*, 2021, pp.1–13.

[103] Aldao, E. et al., 2022. UAV obstacle avoidance algorithm to navigate in Dynamic Building Environments. *Drones*, 6(1), p.16.

[104] Shirai, Y., Inoue, H., 1973. Guiding a robot by visual feedback in assembling tasks. Patt. Recogn. 5, 99–108.

[105] Kendoul, F., Fantoni, I., Nonami, K.: Optic flow-based vision system for autonomous 3d localization and control of small aerial vehicles. Robot. Autonom. Syst. 57(6), 591–602 (2009)

[106] Carrillo, L.R.G., López, A.E.D., Lozano, R., Pégard, C.: Combining stereo vision and inertial navigation system for a quad-rotor uav. J. Intelli. Robot. Syst. 65(1-4), 373–387 (2012).

[107] Stegagno, P.S., 2013. Vision-based Autonomous Control of a Quadrotor UAV using an Onboard RGB-D Camera and its Application to Haptic Teleoperation. In the 2nd IFAC Workshop on Research, Education and Development of Unmanned Aerial Systems, France..

[108] Mannar, S., 2018. Vision-based control for aerial obstacle avoidance in Forest environments. In the *International Federation of Automatic Control*. Bangalore: ELSEVIER, pp. 2405–8963.

[109] Michels, Jeff, Ashutosh Saxena, and Andrew Y. Ng. "High speed obstacle avoidance using monocular vision and reinforcement learning." In Proceedings of the 22nd international conference on Machine learning, pp. 593-600. ACM, 2005.

[110] Hu, B. & Wang, J., 2018. Deep Learning Based Hand Gesture Recognition and UAV Flight Controls. In *Proceedings of the 24th International Conference on Automation & Computing*. Newcastle University, UK.

[111] Li, S., 2018. Learning Unmanned Aerial Vehicle Control for Autonomous Target Following. In *Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence (IJCAI-18)*. HKUST, Hong Kong.

[112] Hulens, D., 2017. Real-time Vision-based UAV Navigation in Fruit Orchards. In *Proceedings of the 12th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2017)*. Belgium, pp. 617–622.

[113] Mittal, M., Mohan, R., Burgard, W., Valada, A. (2022). Vision-Based Autonomous UAV Navigation and Landing for Urban Search and Rescue. In: Asfour, T., Yoshida, E., Park, J., Christensen, H., Khatib, O. (eds) Robotics Research. ISRR 2019. Springer Proceedings in Advanced Robotics, vol 20.

[114] Lin, H.-Y. & Peng, X.-Z., 2021. Autonomous quadrotor navigation with vision based obstacle avoidance and path planning. *IEEE Access*, 9, pp.102450–102459.

[115] Lin, S., Jin, L. & Chen, Z., 2021. Real-time monocular vision system for UAV autonomous landing in outdoor low-illumination environments. *Sensors*, 21(18), p.6226.

[116] Huang, H., 2020. Deep Reinforcement Learning for UAV Navigation Through Massive MIMO Technique. In *IEEE Transactions on Vehicular Technology* , pp. 1117–1121.

[117] Grando, R.B., 2020. Deep Reinforcement Learning for Mapless Navigation of Unmanned Aerial Vehicles. In *2020 Latin American Robotics Symposium (LARS), 2020 Brazilian Symposium on Robotics (SBR) and 2020 Workshop on Robotics in Education (WRE)*. Natal, Brazil.

[118] Kendoul, F., Fantoni, I., Lozano, R.: Adaptive vision-based controller for small rotorcraft uavs control and guidance. In: Proceedings of the 17th IFAC world congress, pp. 6–11 (2008)