The Application of Vision-Based Navigation Techniques on Drones

Xiaohong Gui

Shanghai Pinghe School, Shanghai, 201208, China
amyguixiaohong@hotmail.com

Abstract. With the development of unmanned aerial vehicles (UAVs), more and more of them are introduced into true working conditions, and they need a powerful navigation system to operate smoothly. However, the commonly used navigation systems, such as the global positioning system (GPS), are not always reliable since they depend highly on the satellites available in the region. As a result, the vision-based navigation technology is practiced on UAVs. This technique is much more flexible and fits the working conditions of UAVs. This article introduced three main techniques used in vision-based navigation: visual sensors, object detectors and navigation approaches. Various visual sensors are used to offer information about the environment. Object detection approaches enable drones to perceive their surroundings and notice obstacles. Then, navigation approaches are used to plan the flight path. Map-based navigation approaches let drones navigate themselves with a map of the environment. Map dependent system has access to the precise 3D map, while map building system requires generating maps while navigating. Path planning approaches can determine the best flight path. Global path planning is static, while local path planning uses local data to adjust the route.

Keywords: Vision-based Navigation, Drones, UAVs, Visual Sensors.

1 Introduction

As technology progresses, more and more unmanned robots are being applied to various industries and replacing humans to do some work since they can finish tasks with higher efficiency and safety. Thus, UAVs are introduced into various fields ranging from aerial photography to disaster response. In an industrial area, they can go on patrol and examine the potential dangers in the region. During the period of the COVID-19 pandemic, the UAV industry has gained increasing attention from the public. The drones are used to deliver things to people who are isolated. Other services like sanitization, temperature checking, surveillance, and transportation are carried out with deep learning algorithms and advanced vision systems [1]. Nowadays, drones are typically applied in areas of disaster response, military purposes, surveillance and delivery. In a disaster condition, UAVs perform a search and rescue mission. In the military domain, fixed wings drones are widely applied. They can fly longer distances and finish tasks

K. Subramanian (ed.), *Proceedings of the International Workshop on Advances in Deep Learning for Image Analysis and Computer Vision (IWADIC 2025)*, Advances in Computer Science Research 128,
https://doi.org/10.2991/978-94-6239-648-7_3

such as scouting and pinpoint bombing. They can also supervise work and deliver packages. Since the applications of UAVs are so diverse, UAV technologies are developing rapidly. The core aviation technologies of UAV systems primarily encompass the platform manufacturing technology, flight control and navigation systems, energy and power technologies, as well as mission-specific payload capabilities. [2]. Platform manufacturing technology mainly solves the mechanical structure of drones. Energy and power technologies study the fuel or batteries for drones. The mission-specific payload technology is mainly determined by the task type [2]. The flight control and navigation systems guide aircraft to travel from the starting point to the destination successfully without any collision. Generally, drones use navigation systems that depend on the global positioning system (GPS) or an internal navigation system (INS). However, those conventional navigation systems exhibit fatal vulnerabilities in special regions, including urban areas with little signals, indoor environments and fields with complex surroundings. The limitations of those traditional navigation systems have spurred intensive research into vision-based navigation. The vision-based simultaneous localization and mapping (SLAM) algorithm can generate a coherent environmental map while

concurrently determining the UAV's location within the map [3]. Since modern advancements in flight navigation technology largely focus on achieving greater autonomy in navigation and control, especially the vision-based navigation systems, the application of this techniques on drones will be introduced in this text. This article first reviews the visual sensors used and object detection approaches for different sensors. Then the navigation approaches are concluded, including map-based and path planning based ones. This article identifies the usage of various visual navigation on the flight of drones, outlines the process of vision-based navigation, and analyzes the advantages and disadvantages of different technologies used in the vision-based navigation on drones.

2 Visual Sensors

Visual sensors provide visual information about the surroundings. Currently, there are four types of cameras used on drones as their visual sensors.

2.1 Monocular Camera

Drones that use monocular vision SLAM only have one ordinary camera on them. Monocular cameras are easy-obtained and lightweight. Thus, they are the most widely used cameras on UAVs. However, the SLAM approaches are worth researching because the information provided by this kind of camera is limited.

For instance, Yang et al. applied monocular vision and SLAM approaches on UAVs in emergencies and unknown environments. They implemented a landing system with these algorithms [4]. They proposed a novel map representation approach that combines three-dimensional features and a mid-pass filter to remove noise and construct a grid map with different heights. A region segmentation method is proposed to identify boundaries between grid areas of varying heights, enhancing the efficiency and precision of later landing zone selection. This method was evaluated in terms of its ability to maintain scene reconstruction integrity and finding safe landing location. In both scenes, the autonomous landing is carried out successfully.

In addition, Ahmadi et al. used the UAV kinematics, the optical flow equations, and a low-cost monocular camera to develop a novel appropriate model for estimating the vertical distance and the velocity vector of a multirotor UAV for landing [5]. It demonstrated the capability of height and velocity estimation of a multirotor UAV in the landing phase of flight by just applying the low-cost camera information.

2.2 Stereo Camera

Stereo cameras are usually made up of two separate lenses. They can imitate human vision and obtain 3D images. Thus, they can solve some problems caused by the lack of information in a 2D vision condition. Duan et al. used stereo cameras to develop an innovative method for outlier rejection approach for feature-based visual odometry [6]. With the stereo cameras' known orientation, they derived bounds on each pose parameter, enabling explicit estimation of UAV pose uncertainty. Then, they made some calculations and formed a new method to estimate odometry. After implementing and testing the proposed method on UAV indoor navigation, they found their approach more reliable than the old one.

2.3 Red-green-blue-depth Camera

Red-green-blue-depth (RGB-D) cameras are aided with infrared cameras to offer both visual images and depth maps. It can offer both visual results and depth maps with the operation of infrared sensors. As a result, they can still work in the condition of a lack of light. For instance, Wang et al. presented a real-time obstacle perception method for UAVs with an RGB-D camera, which can overcome the difficulty of perceiving obstacles in real-time for UAVs in low-light environments [7]. They utilized the obstacle detector to construct a novel tracker and adapted it to more complex conditions such as dynamic obstacles and occlusion. The method worked well in both indoor and outdoor settings.

2.4 Fisheye Camera

Fisheye cameras have a larger viewing angle than any other type of camera. When capturing subjects at close range, fisheye lenses generate pronounced perspective distortion. The pictures taken by fisheye cameras undergo significant distortion and have a strong sense of perspective convergence. With a larger visual scope and a longer depth of field, UAVs equipped with a fisheye camera can easily handle complex environments since they can obtain more information about the environment. Since it is a relatively innovative technology, there are only a few recent studies. Donggeun Oh et al. applied a smart control system to a UAV with a fisheye camera [8]. They applied a dynamically adjusted map compression ratio based on the UAV's geographical position, and successfully operated the autonomous path planning algorithms on the UAV in their platform.

3 Object Detection Approaches

Object detection entails the identification and localization of object instances within an image [9]. In a UAV navigation system, object detection is used to detect obstacles and avoid collisions. Different object detection approaches are practiced for UAVs with different types of visual sensors, since they provide different kinds of information. Recently, researchers focused on object detection based on monocular and stereo vision as they are the most widely used ones.

3.1 For Monocular Camera

The object detection approaches for them are well-developed. For monocular cameras, there are mainly four types of detection algorithms: appearance-based, motion-based, depth-based, and expansion-based [10].

Appearance-based approaches distinguish the object by its color and shape. For instance, Ulrich et al. used one camera to provide continuous visual data [11]. It requires a relatively flat background, and the objects significantly differ from the ground. Thus, it judged the obstacles by picking out the pixel that differs in appearance. The classification of a pixel relies on visual attributes like intensity, color, edges, and texture within local regions [11], enabling it to detect small objects precisely. The algorithm is efficient and cheap, but it has numerous limiting conditions.

Motion-based approach is suitable for objects that are moving quickly. Two consecutive pictures are captured and compared. If the displacement of a pixel is faster than the ordinary displacement caused by the movement of UAVs, then the pixel is considered a part of the object. Jia et al. employ motion features to discriminate between obstacles, shadows, and pavement markings [12]. They introduced a model utilizing two consecutive video frames to distinguish obstacles from the ground plane through motion features [12]. They also utilized an adaptive mechanism to minimize false positives [12]. The experiment showed their method was practical in real-time.

UAVs using depth-based algorithms are usually equipped with a fisheye camera, since fisheye cameras can provide more information about depth. However, these approaches show defects in accuracy because of low image quality and high computational requirements [13]. Other researchers insist on utilizing depth-based approaches with one ordinary camera, although this has great difficulty. For example, Zhang et al. introduced a network called monocular detection with a depth-guided transformer (MonoDETR) [14]. They modify the vanilla transformer to be depth-aware by introducing a module that can predict a foreground depth map based on the single camera. They specialized a depth encoder to derive non-local depth embeddings as well. The 3D attributes of objects in their depth-guided regions were estimated adaptively by MonoDETR, and the model achieved state-of-the-art performance.

Expansion-based methods use means similar to humans to find objects. As objects get closer to the camera, objects would expand in a fixed proportion. The proportion is obtained by comparing two sequential pictures. When the enlargement proportion of a part of the image exceeds the normal value, it is considered part of the obstacle. Al-Kaff et al. applied this approach to UAVs [15]. The detection algorithm tracks the changes in the area size of the approaching obstacles and extracts boundary points of obstacles. After that, the method judges the collision risk by analyzing the relative area ratio between detected obstacles and the UAV's position, and initiating avoidance maneuvers when necessary. Experimental evaluation via both indoor and outdoor flight trials confirmed the algorithm's accuracy.

3.2 For Stereo Camera

Stereo vision-based object detection approaches require at least two cameras on one UAV to dynamically construct a 3D environmental map. For stereo cameras, there are mainly two types of detection algorithms: Inverse Perspective Mapping (IPM) - based and disparity histogram-based [16].

IPM-based methods have limited research and application because they are the most costly. It requires forming aerial view images from the images taken by stereo cameras on UAVs. The aerial view can eliminate perspective distortion, so that the sizes and positions of objects in the image are the same as those in the real world, making it easier for subsequent tasks such as object detection. Kim et al. utilized one wide field of view (FOV) camera and one narrow FOV view camera to detect obstacles [17]. They combined IPM between the input images of the stereo camera and IPM using consecutive input images from a camera with a narrow FOV [17]. Obstacle detection was performed through IPM-based region detection, while stereo matching provided spatial position coordinates of the obstacle [17]. Their experiment results show that this algorithm is precise and practical.

In disparity histogram-based approaches, two similar cameras are installed on one UAV. They construct a disparity map by taking pictures ceaselessly and calculating disparity for all image pixels. The data is then processed to generate a depth map of the surrounding objects [18]. However, the large number of calculations needed makes it hard to operate on UAVs. Therefore, McGuire et al. presented a highly efficient computer vision algorithm called Edge-FS for the application of this algorithm on micro aerial vehicles, which can determine velocity and depth [19]. The drone's velocity is calculated by the stereo-based distance estimates. These measurements enable drones only relying on on-board sensors to achieve fully autonomous flight, including control velocity and avoiding obstacles themselves [19].

4 Navigation Approaches

Besides the cameras and their object detection approaches, a vision-based navigation system also needs to plan the moving path for the UAV. Map-based approaches plan the required movement from the starting point to the ending point, and path planning approaches fine tune the route and enable UAVs to avoid collision and get to their destination successfully.

4.1 Map-based Approaches

Map-based navigation approaches use maps of the environment to guide the flight of UAVs. Based on various levels of contained detail of maps or the origin of maps, the approaches can be categorized as map dependent systems and map building systems.

UAVs applying map dependent systems rely on complete access to a 3D map of the entire environment they operate in. For instance, Saranya et al. established wireless communication to estimate real-time visual position of UAVs through continuous onboard camera processing [20]. They first utilized traditional means and the help of sensors to navigate the UAV. However, these means encounter errors, thus the direct visual location is applied as a remedy. The position estimation is achieved by utilizing the Google static map application programming interface. A computer at the ground station transmitted the UAV's position into GPS information.

However, providing a detailed 3D map for UAVs is rare and difficult. Moreover, the environment must be the same as the map, but this is hard to achieve. As a result, the UAV may not recognize its position precisely in those strategies, and the map building system is more widely preferred for autonomous navigation. For example, Bavle et al. [21] used stereo visual SLAM to provide real-time images to update the indoor map in a UAV. This method fuses low-level visual and visual-inertial odometry (VO/VIO) with geometrical information derived from detected semantic objects. Several state-of-the-art VO/VIO algorithms and object detectors were used in order to estimate the complete 6DoF pose of drones and create an annotated feature map of the environment [21].

4.2 Path Planning Based Approaches

Path planning is the process of determining the best route from start to finish point. The algorithms combine the results obtained previously and judge the path that both reaches the destination and avoids collision. There are two approaches to path planning: global path planning approaches and local path planning approaches.

Global path planning approaches depend on a global geographic map as the UAV's static map and find the most efficient path, such as the shortest route or the lowest cost of work. For instance, Belge et al. construct an algorithm called Harris hawk optimization (HHO)–grey wolf optimization (GWO), which is a hybrid metaheuristic optimization framework for UAVs that solves optimal path planning and trajectory tracking problems during payload delivery missions [22]. Its path is planned based on known maps, and the algorithm is used to make sure the path is the most energy-efficient while no collisions happen.

Local path planning approaches utilize data about the environment collected by UAVs. They usually involve complex computation because they require real-time dynamic calculations and planning of the route. Yang et al. proposed a 3D space path planning method for UAV atmospheric sampling [23]. The algorithm optimizes coverage density while satisfying onboard energy limitations, ensuring complete volumetric environment detection. Simulation results demonstrate that the newly proposed approach significantly enhances coverage density when compared to existing path generation techniques [23]. Since the energy is saved from its movements, it can collect more gas samples at one time and detect the atmospheric environment more precisely. Souza et al. introduce an enhanced potential field approach that not only enables obstacle avoidance, but also addresses issues of local minima and oscillations within the repulsive field's influence range [24]. They developed a strategy for every individual UAVs to autonomously determine optimal vortex field avoidance vectors based on each obstacle. Experiments show that it can generate feasible paths in real-time applications.

5 Conclusion

With the development of technology, UAVs have been developing rapidly in recent years. The vision-based navigation technology is an important part of the operation of drones. It includes obstacle detection approaches, map-based navigation approaches and path planning approaches. Visual sensors are necessary as the source of information about the environment. The monocular camera approach relies on only one camera. It is cheap and easy to obtain, but it cannot provide information about depth, so stereo cameras appear to make up for the disadvantage. They have a pair of visual sensors and can form 3D images. RGB-D cameras can work in low-light conditions with the help of infrared cameras. Fisheye cameras are cameras with larger viewing angle and longer depth of field, providing abundant depth information. For monocular cameras, the objects can be detected based on their appearance, motion, depth or expansion. For stereo cameras, there are IPM based and disparity histogram-based object detection algorithms.

The drones also need navigation approaches to find the flight path. For map-based approaches, map-dependent approaches require full access to the 3D map of the whole environment, but this is not very practical. Therefore, map-building strategies are more widely used. They involve building maps while moving and collecting information in the environment. With the rough route planned and obstacles discovered, UAVs require path planning approaches to obtain the best route and achieve other required performance.

This article outlines the process and means of the vision-based navigation of UAVs. Several studies of the related techniques are cited as examples and show the trend of development in this domain. Some similar technologies are compared and evaluated. This summary is not comprehensive enough to include all the state-of-the-art development on the techniques. Further research is essential to understand the development of this industry and improve the technologies used.

References

1. Aabid, A., Parveez, B., Parveen, N., Khan, S.A., Zayan, J.M., Shabbir, O.: Reviews on design and development of unmanned aerial vehicle (drone) for different applications. J. Mech. Eng. Res. Dev. 45(2), 53–69 (2022)
2. Fan, B., Li, Y., Zhang, R., Fu, Q.: Review on the technological development and application of UAV systems. Chin. J. Electron. 29(2), 199–207 (2020)
3. Arafat, M.Y., Alam, M.M., Moh, S.: Vision-based navigation techniques for unmanned aerial vehicles: Review and challenges. Drones 7(2), 2 (2023)
4. Yang, T., Li, P., Zhang, H., Li, J., Li, Z.: Monocular vision SLAM-based UAV autonomous landing in emergencies and unknown environments. Electronics 7(5), 5 (2018)
5. Nabavi-Chashmi, S.-Y., Asadi, D., Ahmadi, K.: Image-based UAV position and velocity estimation using a monocular camera. Control Eng. Pract. 134, 105460 (2023)
6. Duan, R., Paudel, D.P., Fu, C., Lu, P.: Stereo orientation prior for UAV robust and accurate visual odometry. IEEE/ASME Trans. Mechatron. 27(5), 3440–3450 (2022)

7. Wang, H., Wang, H., Liu, Y.-J., Liu, L.: Real-time obstacle perception method for UAVs with an RGB-D camera in low-light environments. Signal Image Video Process. 19(3), 256 (2025)
8. Oh, D., Han, J.: Fisheye-based smart control system for autonomous UAV operation. Sensors 20(24), 24 (2020)
9. Amit, Y., Felzenszwalb, P., Girshick, R.: Object detection. In: Computer Vision, pp. 875–883. Springer, Cham (2021)
10. Singh, Y., Kaur, L.: Obstacle detection techniques in outdoor environment: Process, study and analysis. Int. J. Image Graph. Signal Process. 9(5), 35–53 (2017)
11. Ulrich, I., Nourbakhsh, I.: Appearance-based obstacle detection with monocular color vision. In: AAAI-00 Proceedings (2000)
12. Jia, B., Liu, R., Zhu, M.: Real-time obstacle detection with motion features using monocular vision. Vis. Comput. 31(3), 281–293 (2015)
13. Aziz, M.W.A., Ramli, M.F.: A review on method of obstacle detection and avoidance systems for unmanned aerial vehicles: Monocular camera. J. Multi-Discip. Eng. Rev. 1(1), 1 (2024)
14. Zhang, R., et al.: MonoDETR: Depth-guided transformer for monocular 3D object detection. In: 2023 IEEE/CVF International Conference on Computer Vision (ICCV), pp. 9121–9132. IEEE, Paris (2023)
15. Al-Kaff, A., García, F., Martín, D., De La Escalera, A., Armingol, J.M.: Obstacle detection and avoidance system based on monocular camera and size expansion algorithm for UAVs. Sensors 17(5), 5 (2017)
16. Jung, H.G., Lee, Y.H., Kim, B.J., Yoon, P.J., Kim, J.H.: Stereo vision-based forward obstacle detection. Int. J. Automot. Technol. 8(4), 493–504 (2007)
17. Kim, C., Lee, T., Cho, D.: An application of stereo camera with two different FoVs for SLAM and obstacle detection. IFAC-Pap. 51(22), 148–153 (2018)
18. Badrloo, S., Varshosaz, M., Pirasteh, S., Li, J.: Image-based obstacle detection methods for the safe navigation of unmanned vehicles: A review. Remote Sens. 14(15), 15 (2022)
19. McGuire, K., de Croon, G., De Wagter, C., Tuyls, K., Kappen, H.: Efficient optical flow and stereo vision for velocity estimation and obstacle avoidance on an autonomous pocket drone. IEEE Robot. Autom. Lett. 2(2), 1070–1076 (2017)
20. Saranya, K.C., Naidu, V.P.S., Singhal, V., Tanuja, B.M.: Application of vision based techniques for UAV position estimation. In: 2016 International Conference on Research Advances in Integrated Navigation Systems (RAINS), pp. 1–5 (2016)
21. Bavle, H., De La Puente, P., How, J.P., Campoy, P.: VPS-SLAM: Visual planar semantic SLAM for aerial robotic systems. IEEE Access 8, 60704–60718 (2020)
22. Belge, E., Altan, A., Hacıoglu, R.: Metaheuristic optimization-based path planning and tracking of quadcopter for payload hold-release mission. Electronics 11(8), 1 (2022)
23. Yang, L., Fan, S., Yu, B., Jia, Y.: A coverage sampling path planning method suitable for UAV 3D space atmospheric environment detection. Atmosphere 13(8), 8 (2022)
24. Souza, R.M.J.A., Lima, G.V., Morais, A.S., Oliveira-Lopes, L.C., Ramos, D.C., Tofoli, F.L.: Modified artificial potential field for the path planning of aircraft swarms in three-dimensional environments. Sensors 22(4), 4 (2022)