Unmanned Aerial Vehicle's Obstacle Avoidance Research Based on Vision

Wenzhe Wang

School of Physics&Electronic Science, Changsha University of Science & Technology, Changsha, China
wo22ne@stu.csust.edu.cn

Abstract: With the increasingly widespread application scenarios of drones, higher requirements have been put forward for the autonomous flight capability of drones. The autonomous obstacle avoidance technology of unmanned aerial vehicles plays an important role in various environments that are inconvenient for pilots to operate. With the rise of image recognition technology, visual sensors have become the mainstream choice for obstacle avoidance in unmanned aerial vehicles. There are currently many autonomous obstacle avoidance systems based on visual design, and the chosen solutions are too broad. This article aims to introduce the current development status of visual obstacle avoidance, categorizing obstacle avoidance schemes based on the number of cameras used, and helping researchers choose appropriate obstacle avoidance schemes. This article provides a certain degree of analysis of obstacle recognition and avoidance solutions for drone obstacle avoidance, as well as the problems faced in development and possible future solutions. Visual obstacle avoidance technology still needs to overcome many obstacles.

Keywords: unmanned aerial vehicle, Obstacle Avoidance, Obstacle Detection, Vision

1 Introduction

The rapid development of the drone market and application scenarios in recent years brings tremendous changes to various civilian fields. In terms of express delivery logistics, Google launched Project Wing, and Amazon applied for a patent for the Prime Air delivery system; Several Chinese express delivery companies such as SF Express and Alibaba also start researching drone delivery systems [1]; The application of drones in power line inspection is also increasing. Zhou et al. designed a stereo vision guided automatic detection system for the maintenance of overhead high-voltage lines [2]; Aspragkathos et al. designed a hybrid model-based and data-driven framework for autonomous shoreline monitoring by unmanned aerial vehicle [3].

Free from the constraints of the ground environment to expand to move freely in three-dimensional space, unmanned aerial vehicles are able to perform many tasks that are not possible for ground robots and make unmanned aerial vehicle missions more diverse and complex. As technology evolves, researchers are demanding a higher level

P. Kar et al. (eds.), *Proceedings of the 2023 International Conference on Image, Algorithms and Artificial Intelligence (ICIAAI 2023)*, Advances in Computer Science Research 108,
https://doi.org/10.2991/978-94-6463-300-9_51

of autonomy and flight stability from unmanned aerial vehicles. An ideal unmanned aerial vehicle obstacle avoidance system would enable high-precision obstacle recognition and obstacle avoidance.

Early obstacle avoidance technologies are more based on simple recognition, the improvement of autonomous driving technology on automotive and robot platforms gave unmanned aerial vehicles more options for autonomous obstacle avoidance solutions. While miniaturization of electronic devices has allowed small devices like multi-rotor unmanned aerial vehicles to be equipped with more sophisticated sensors, the weight of system components still needs to be considered, and the limited weight-bearing capacity and battery capacity of small and medium-sized unmanned aerial vehicles can limit the use of some power-consuming, heavy-mass sensors. Unmanned aerial vehicle obstacle avoidance needs to rely on external sensors to obtain environmental information, common infrared sensors, radar sensors, laser sensors and vision sensors. Infrared sensors are costly and less resistant to interference [4]. In terms of weight and power requirements, structured light sensors and LIDAR are not suitable for small flight platforms [5]. In contrast, vision sensors have become the primary choice for commercial-grade unmanned aerial vehicles due to their relatively low cost and light weight. Visual obstacle avoidance is now a relatively mature industry in the field of unmanned aerial vehicle obstacle avoidance, and the DJI Marvic 3 series uses a binocular vision-based perception system.

The purpose of this paper is to summarize the current development status of visual obstacle avoidance systems for unmanned aerial vehicles to some extent. Section 2 will introduce the way to achieve obstacle perception through vision, Section 3 will introduce the implementation of obstacle avoidance function in conjunction with related literature, and Section 4 will summarize the problems faced in the development of visual obstacle avoidance and possible future solutions.

2 Obstacles Detection

2.1 Sensor Selection

The recognition of obstacles plays an important role in any unmanned machine navigation system. At present, there are many solutions designed with different sensors. According to the number of cameras used for navigation, solutions can be simply divided into monocular vision and binocular vision.

Monocular vision, as the name suggests, is the use of a single camera to perceive the presence of obstacles. Its principle is to use a single camera to capture each frame at a certain frequency and obtain obstacle position information by comparing the pixel changes or feature areas of the two frames of images before and after. However, the simple structure also leads to the disadvantage of monocular vision in directly measuring depth. To compensate for this deficiency, other methods need to be used to estimate depth information.

Some people choose to obtain depth information by processing visual odometer cues, common methods include optical flow, perspective, texture gradient, and relative size [6].

Binocular vision is a three-dimensional perception method that mimics the human eye. The foundation of binocular vision is triangulation. Figure 1 shows the sketch map of binocular vision.

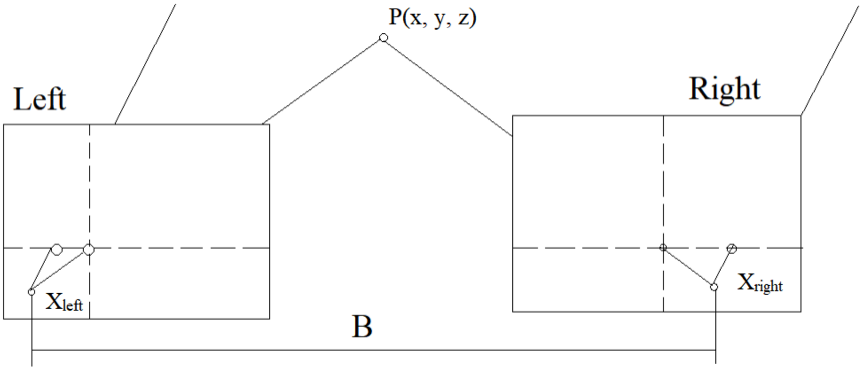

Fig. 1. The principle of binocular vision schematic

Figure 1 marks the position of P and its image points on the different view. B is the distance between the projection centers of two cameras.

Let a feature point P (x, y, z) in space, so set its image points in the binocular camera are

$$\text{Left}= (X_{left}, Y_{left}) \tag{1}$$

$$\text{Right}= (X_{right}, Y_{right}) \tag{2}$$

For the two cameras are on the same plane, the coordinates Y of all feature points P in the left and right camera images are the same, that is, their three-dimensional coordinates can be obtained from the triangular geometric relationship:

$$X=BX_{left}/d, \tag{3}$$

$$y=BY/d, \tag{4}$$

$$y=Bf/d, \tag{5}$$

$$d=X_{left}-X_{right}. \tag{6}$$

Among them, X and Y is P's horizontal and vertical coordinates. D is the parallax. F is the focal length of the camera. It is also common to use other sensors as auxiliary designs. Such as IMU [7], laser sensors [1], ultrasonic sensors [8], etc.

2.2 Obstacles Detection

Optical Flow. Optical flow refers to the motion of each pixel represented on a two-dimensional image plane when an object moving in three-dimensional space is projected onto it. There are various methods for estimating optical flow, with the Lucas Kanade algorithm [9] commonly used for sparse optical flow and the Farneback and Horn Schunck algorithms [10] commonly used for dense optical flow.

Cho et al. realized unmanned aerial vehicle's 3 direction obstacle detection in 3D texture environment through optical flow method [11].

Deep Learning. The images obtained by camera are also very suitable for deep learning processing to obtain relevant information. Among various common deep learning models, convolutional neural networks are commonly used in vision based autonomous flight perception and obstacle avoidance research for their ability to automatically detect input images or time series and extract image features from them. Mac et al. used openCV combined with EFK to process the body information obtained by IMU to achieve dynamic obstacle detection and estimation [12].

3 Obstacle Avoidance

The core of the unmanned aerial vehicle visual obstacle avoidance system is to achieve avoidance after obtaining obstacle information.

Obstacle avoidance algorithms are a very complex branch in computer science. Currently, obstacle avoidance algorithms are mainly divided into three categories: obstacle avoidance methods based on optimization, potential fields, and machine learning [13].

The optimization algorithms include nonlinear optimization methods, mixed integer linear/nonlinear programming, dynamic programming, etc.

Potential field algorithms mainly include artificial potential field method, velocity obstacle method, stream function method, etc.

Neural networks, on the other hand, include models such as DroNet, DNN, TrailNet, AlexNet, DenseNet, etc[14]. Among them, DroNet, DNN, TrailNet is verified in indoor environments. After testing, AlexNet and DenseNet are proofed effective outside the door. These models give researchers simple way to choose them.

There are also some other attempts. Some researchers try to build their own schemes to avoid collisions. Xiao et al. obtained a depth map of the scene through 3D reconstruction and stereo matching, and proposed a multi obstacle detection algorithm that applies a series of image processing algorithms to identify objects and provide their bounding boxes, achieving ideal measurement accuracy within a range of 15 meters [15]. Soria et al. used low-cost cameras to create local maps of the work area using depth information, construct point cloud representations of objects, and then use support vector machines to classify and evaluate them, returning accurate results [16]. Al Kaff et al. used the size expansion algorithm to simplify the drone model in path calculation to a single point and constructed an avoidance algorithm to establish a safety zone outside the drone model, when obstacles are too close to the drone, the drone will hover [4].

V-SLAM is a visual obstacle avoidance scheme first proposed for unmanned robots. It was first proposed by Smith, Self and Cheeseman in 1988[17]. Through years of development, SLAM has been developed to a considerable extent [18, 19]. SLAM technology is the sum of the many techniques involved in trying to solve the problem of determining one's own motion trajectory by observing the environment while constructing a map of the environment. Some researchers summarize some attempts in introducing SLAM to unmanned aerial vehicle's obstacle avoidance and V-SLAM's research hotspots and development trends [20].

4 Discussion

Current unmanned aerial vehicle visual obstacle avoidance systems face certain limitations in both sensors and obstacle avoidance solutions. First, battery technology and sensor technology make unmanned aerial vehicle platforms limited by issues such as their own range energy consumption and payload size limitations. Moreover, the verification of many obstacle avoidance systems is limited to indoors, and the recognition accuracy decreases due to changes in lighting and other conditions outdoors, which in turn affects the system accuracy. Though machine learning can make up some shortcomings, there still exists some troubles. Machine learning models and algorithms for visual technology require a large amount of data, among which data acquisition costs are high and data labeling is required, resulting in a long work time. In addition, the quality of data is also a key factor affecting models and algorithms. Using high-quality and accurate data to train models can make more accurate predictions. Bad predictions can even double the severity of the situation. Another issue faced by data is dataset bias, which refers to the lack of certain content in the dataset. When the data is not sufficiently representative, training data bias has a negative impact on the accuracy of the model. The currently available open-source datasets, such as ImageNet, CIFAR-10&CIFAR-100, MNIST, KITTI, nuScenes, etc., play an important role in the research and evaluation of deep learning algorithms.

By analyzing the use of various sensors in the current unmanned aerial vehicle obstacle avoidance system, it is easy to find that many studies have focused on the use of a single sensor.

Whether vision sensors (monocular vision or binocular vision) or non-vision sensors (millimeter wave radar, LIDAR, ultrasonic or infrared), all have their own limitations when used. For example: vision sensors are very sensitive to ambient light, infrared, ultrasonic detection distance is limited, the picture resolution is low and cannot obtain three-dimensional information about the object. LIDAR is affected by rain and snow, and the weight is expensive. Millimeter wave radar cannot obtain three-dimensional information about the object. The combination of multiple sensors has become the preferred way of unmanned aerial vehicle obstacle avoidance technology, such as vision-based and LIDAR way, vision-based and millimeter wave radar way, in which the combination of millimeter wave radar and vision has more advantages such as long-range detection, low cost and coping with bad weather, and the perception ability and safety performance of unmanned aerial vehicle has been greatly The sensing capability and safety performance of unmanned aerial vehicles have been improved. The detection results of vision sensors in low-texture, low-light environments are susceptible to interference, and the use of other sensors can be considered as auxiliary to the extent that the unmanned aerial vehicle's own load-bearing capacity allows, and as a backup solution when either sensor fails to provide accurate obstacle avoidance information.

There have been attempts to combine vision + vision sensors and vision + non-vision sensors, for example, RGB-D cameras and event cameras can facilitate the identification of obstacles in high-speed motion [21], and vision sensors + IMU can compensate for the disadvantages of IMU's tendency to drift and camera's poor performance in low-texture environments [22].

5 Conclusion

This article divides the obstacle avoidance system of unmanned aerial vehicles into two parts: obstacle detection and obstacle avoidance. In Section 2, the perception part, the selection of visual sensors is divided into two modes based on the number of cameras used for navigation: monocular vision and binocular vision, and commonly used obstacle detection schemes are summarized. In Section 3, the obstacle avoidance section, some common obstacle avoidance schemes are summarized. There are also some new schemes authors proposed in their articles. Finally, in section 4, considering the problems faced by the current development of visual obstacle avoidance systems, and a method of considering multi-sensor fusion was proposed, but no suggestions were made on how to achieve data fusion between various sensors.

References

1. Lee, J. Y., Chung, A. Y., Shim, H., Joe, C., Park, S., & Kim, H. UAV Flight and Landing Guidance System for Emergency Situations †. Sensors, 19(20), 4468. (2019).
2. Zhou, Y., Xu, C., Dai, Y., Feng, X., Ma, Y., & Li, Q. Dual-View Stereovision-Guided Automatic Inspection System for Overhead Transmission Line Corridor. Remote Sensing, 14(16), 4095. (2022).
3. Aspragkathos, S. N., Karras, G. C., & Kyriakopoulos, K. J. A Hybrid Model and Data-Driven Vision-Based Framework for the Detection, Tracking and Surveillance of Dynamic Coastlines Using a Multirotor UAV. Drones, 6(6), 146. (2022).
4. Abdulla, A. K. , García Fernando, Martín David, Arturo, D. L. E. , & Armingol José. Obstacle detection and avoidance system based on monocular camera and size expansion algorithm for uavs. Sensors, 17(5), 1061. (2017).
5. Zhang, Z. , Xiong, M. , & Xiong, H. . Monocular Depth Estimation for UAV Obstacle Avoidance. In: 2019 4th International Conference on Cloud Computing and Internet of Things (CCIOT). (2019).
6. Padhy, R. P., Choudhury, S. K., Sa, P. K., & Bakshi, S. Obstacle Avoidance for unmanned aerial vehicles: Using Visual Features in Unknown Environments. IEEE Consumer Electronics Magazine, 8(3), 74–80. (2019).

7. Gao, M. , Yu, M. , Guo, H. , & Xu, Y. . Mobile robot indoor positioning based on a combination of visual and inertial sensors. Sensors, 19(8). (2019).
8. Krmer, M. S. , & Kuhnert, K. D. . Multi-Sensor Fusion for UAV Collision Avoidance. In: the 2018 2nd International Conference. (2018).
9. Lucas, B. D. , & Kanade, T. . An Iterative Image Registration Technique with an Application to Stereo Vision. In: Proceedings of the 7th International Joint Conference on ArtificialIntelligence. Morgan Kaufmann Publishers Inc. (1997).
10. Horn, B. K. P. , & Schunck, B. G. . Determining optical flow. Artificial Intelligence, 17(1-3), 185-203. (1981).
11. Cho, G. , Kim, J. , & Oh, H. . Vision-based obstacle avoidance strategies for mavs using optical flows in 3-d textured environments. Sensors, 19(11), 2523-. (2019).
12. Mac, T. T. , Copot, C. , & Ionescu, C. M. . Detection and Estimation of Moving obstacles for a UAV. (Vol.52, pp.22-27). (2019).
13. Zhang, H. H. , Gan, X. S. , Mao, Y. , Zhang, C. L. , & Xie, X. W. . Review of UAV Obstacle Avoidance Algorithms (in Chinese). Aero Weaponry (05), 53-63. (2021).
14. Wang, C. B. , Zhang, A. S. , Yang, L. , Liang, G. Q. , & Zhang, B. . A Review of Deep Vision-Based Autonomous Flight Perception and Obstacle Avoidance for Quadrotor unmanned aerial vehicles (in Chinese). Radio Engineering. (2023).
15. Xiao, Y. , Lei, X. , & Liao, S. . Research on UAV Multi-Obstacle Detection Algorithm based on Stereo Vision. In: 2019 IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC). IEEE. (2019).
16. Ramon Soria, P., Bevec, R., Arrue, B. C., Ude, A., & Ollero, A. Extracting Objects for Aerial Manipulation on UAVs Using Low Cost Stereo Sensors. Sensors (Basel, Switzerland), 16(5), 700. (2016).
17. Smith, R. , Self, M. , & Cheeseman, P. . Estimating uncertain spatial relationships in robotics. Springer New York. (1990).
18. Davison, A. J., Reid, I. D., Molton, N. D., & Stasse, O. MonoSLAM: Real-Time Single Camera SLAM. IEEE Transactions on Pattern Analysis and Machine Intelligence, 29(6), 1052–1067 (2007).
19. Montiel, J., M. M. , Mur-Artal, Raul, Tardos, & Juan, D. . Orb-slam: a versatile and accurate monocular slam system. IEEE Transactions on Robotics: A publication of the IEEE Robotics and Automation Society. (2015).
20. Zhao, J. L. , Zhu, Y. Q. ,& Jin, R. . Review of Monocular V-SLAM for Multi-Rotor unmanned aerial vehicle (in Chinese). Aero Weaponry (02), 1-14. (2020).
21. Falanga, D. , Kleber, K. , & Scaramuzza, D. . Dynamic obstacle avoidance for quadrotors with event cameras. Science Robotics, 5(40), eaaz9712. (2020).
22. Brockers, R. , Fragoso, A. , & Matthies, L. . Stereo Vision-based Obstacle Avoidance for Micro Air Vehicles using an Egocylindrical Image Space Representation. In: Conference on Micro- and Nanotechnology Sensors, Systems, and Applications. (2016).