

Analysis of Visual Camera Applications in Navigation Systems

Gengrui Chen *

Newchannel Shenzhen School, Shenzhen, 518000, China

* Corresponding Author Email: Chengengrui0124@outlook.com

Abstract. This article elucidates that in the current development, the navigation systems are widely applied in a large number of equipment, and visual cameras are increasingly becoming the core component of such systems. Firstly, the article introduces the value of the application of visual cameras. Then, explain the hardware characteristics and advantages and disadvantages of monocular, binocular, and multi-camera vision systems. Subsequently, explores its complementary role in multi-sensor fusion with systems such as the Beidou Navigation Satellite System, as well as its integration with artificial intelligence. At last, the article looks ahead, explores the future development, and the future application of visual cameras. These cameras find extensive use in drones, assisted driving vehicles, and smart transportation systems. By integrating multiple sensors with artificial intelligence and deep learning, they significantly enhance navigation accuracy and reliability. Furthermore, future applications will expand into some new domains, such as utilizing wearable devices to guide people who have visual impairments.

Keywords: Visual camera, Navigation system, Multi-sensor fusion, SLAM; Deep learning.

1. Introduction

With technological development in assisted driving, drones, and other fields, the demand for navigation systems is increasing. Simultaneously, various devices now require higher precision from navigation systems, making traditional single-sensor solutions increasingly difficult to fulfill the requirements. For instance, GPS often has insufficient accuracy or even fails to locate in complex environments like tunnels and valleys. LiDAR is easily interfered with in adverse weather conditions such as heavy fog or sandstorms, leading to limited visibility. Therefore, navigation systems typically require the integration of multiple sensors, including visual cameras, LiDAR, and GPS, to enhance navigation accuracy and reliability through complementary data fusion. Visual camera, with its rich information content and cost-effectiveness, has become a crucial component in this multi-sensor fusion ecosystem.

Visual cameras have different configurations: monocular vision, binocular vision, and multi-camera vision. Monocular vision can be employed to recognize simple objects; binocular vision can estimate approximate distances through the angular difference between two cameras; multi-camera vision can simultaneously capture perspectives from different orientations. It is possible to construct an environmental map of the moment based on a multi-camera vision system. In existing relevant research, it is primarily focused on the application of various visual cameras within single scenarios. However, the integration of visual camera sensors with deep learning remains underdeveloped, has problems such as insufficient computational power or high AI decision-making latency.

This paper aims to systematically elucidate the core role of visual cameras within navigation systems. The study first outlines their practical applications, subsequently analyzes the hardware characteristics and shows the advantages and disadvantages of monocular, binocular, and multi-camera visual systems and the principles of SLAM systems. Then explores their complementary role within multi-sensor fusion alongside technologies such as GPS, whilst investigating how AI algorithms enhance the interpretative abilities of visual data. Finally, we think about the future development of visual cameras and their expanding range of application situations.

2. Application for Visual Camera Navigation Systems

Visual cameras, owing to their low cost and rich information content, have become essential sensors in navigation systems for complex scenarios such as drones, driver-assistance vehicles, and intelligent transport systems. Drones use visual cameras for obstacle avoidance and target acquisition; in driver-assistance vehicles, visual cameras can be the complement of the sensors such as LiDAR, GPS, and so on... let the navigation system have more precise guidance in adverse conditions; within intelligent transport systems, visual cameras identify vehicles, analyse traffic congestion in the moment, and enable dynamic adjustments to enhance traffic flow efficiency. Among these applications, visual cameras have the most widespread deployment in drones.

In drones, binocular vision cameras are employed for obstacle detection. By mounting two parallel cameras on the drone, data from the left and right cameras is acquired. Through calibration, noise is effectively eliminated to extract critical information. Subsequently, by applying matching algorithms to the perspective difference data within this information, the location of the safe obstacle avoidance zone can be determined [1]. Cameras can also identify drone landing pads, enabling precise landing without errors. The drone navigation system employs the ‘Aruco’ image recognition solution, which identifies the unique square code present on each landing pad to perform tasks like positioning, recognition, and attitude estimation [2].

3. The Various Forms of Visual Cameras

3.1. Monocular Vision, Binocular Vision, and Multi-Camera Vision

The application forms of visual cameras within navigation systems can be labeled into three types, which are based on the number of cameras and their configuration: monocular vision, binocular vision, and multi-camera vision. Different system configurations enable distinct functionalities; typically, more complex navigation environments require binocular or multi-camera vision to obtain more precise recognition.

Monocular vision systems only have a single camera, enabling fundamental environmental perception through captured two-dimensional imagery. It is typically employed as auxiliary sensors; monocular vision is commonly found in domestic robotic vacuum cleaners and small drones. Binocular vision camera systems emulate the structure of human eyes, operating on principles analogous to how humans judge distance through binocular parallax. Typically, letting two parallel cameras positioned at a distance (usually 5–20 centimetres), it captures positional deviations of the same target across two image frames, calculating depth information through algorithms. Multi-camera vision systems, an enhanced variant of binocular cameras, incorporate three or more cameras. By deploying cameras at different angles, they achieve more view coverage and greater redundancy.

3.2. Advantages and Disadvantages of Different Camera Types

The monocular camera offers the advantages of an extremely simple structure, small size, and low power consumption, making it suitable for small devices such as small drones and robotic vacuum cleaners. Monocular vision systems can perform fixed-path guidance within specific environments by recognizing pre-mapped routes. However, they are unable to meet the special circumstances and react immediately, such as sudden obstacles or alterations to the navigation path. Furthermore, monocular cameras cannot directly capture the absolute depth information. In the rapidly advancing field of monocular SLAM technology, the inability to obtain distance data may lead to scale uncertainty and scale drift in the acquired information [3].

Compared to monocular cameras, binocular and multi-camera systems can use the angular differences between cameras to compute depth information, enabling more precise navigation. Multi-camera systems additionally provide multi-directional perspective data. Picard Q et al. tested SLAM performance across various platform devices, demonstrating that power-constrained platforms (drones, small robots) necessitate trade-offs between accuracy, real-time capability, and

computational power consumption [4]. Simultaneously, achieving high-precision, low-latency guidance needs costly upgrades to computational hardware. Table 1 presents a comparative analysis of different camera types.

Table 1. Performance Comparison of Different Camera Types

| Type | Advantages | Disadvantages | Application situations |
|------------------------------------|---|---|---|
| Monocular camera | Simple structure, small size, low power consumption; only supports fixed-line guidance. | Difficult to handle unexpected situations; lacks absolute depth information, SLAM might have scale issues. | Micro-devices such as small drones and robotic vacuum cleaners |
| Binocular and multi-camera systems | Capable of calculating depth information for more precise guidance; multi-camera systems provide multiple viewing angles. | Power-constrained platforms require a trade-off between performance and energy consumption; high precision necessitates costly upgrades to computational power. | Equipment requiring depth and precision (where power consumption and cost must be considered) |

4. Use of Visual Cameras in Multi-Sensor Fusion

Visual cameras can be complementary sensors to others, such as when integrated with the aforementioned Beidou Navigation Satellite System. It is a high-precision positioning system based on global satellite positioning technology. By receiving signals from multiple satellites, it achieves positioning accuracy at the centimeter or even millimeter level. Yang, Y. et al. experimentally demonstrated that in the practical situation, they may encounter issues such as electromagnetic interference, signal obstruction, and data transmission delays [5]. Consequently, Yang, Y. et al. employed a BeiDou-Binocular Collaborative Navigation (BBCN) model, integrating BeiDou navigation with binocular vision. The binocular cameras can be a part of an auxiliary sensor: when BeiDou signals fluctuate, the binocular vision system can generate local environmental maps using techniques like visual SLAM, providing robots with short-term alternative positioning and path planning capabilities [5]. This approach compensates for potential positioning inaccuracies in the BeiDou system, enhancing the overall reliability of the navigation system.

Concurrently, in autonomous vehicle applications, visual cameras are typically integrated with LiDAR to form the primary sensor suite for intelligent driving systems. Visual cameras offer rich color and texture information but exhibit significant performance limitations in low-light conditions. LiDAR employs the time-of-flight (ToF) to generate high-precision 3D point clouds, offering advantages of high accuracy and reduced susceptibility to light intensity. In their research, Wang, H. et al. employed multiple algorithms to design a redundant fusion module based on millimeter-wave radar and cameras, compensating for LiDAR's performance decline in rain and snow scenarios [6]. In their research, Liu Bingqi et al. demonstrated that employing a combined laser-vision-inertial odometry (LVIO), which focuses on localization and mapping applications, enables the construction of more precise maps [7]. By compensating for the disadvantages of various sensors, these studies significantly enhance the reliability and safety of autonomous vehicle driving.

5. Deep Learning Empowers Visual Navigation Systems

Artificial intelligence (especially in deep learning) can compensate for the disadvantages of visual SLAM through learning from vast datasets, dynamic scene modelling, and multimodal information fusion. Zhao Z. et al. proposed a deep graph matching called LightGlue to replace conventional manual feature matching methods, significantly enhancing tracking stability in different light

conditions. This addresses the noticeable performance ability decline of traditional visual cameras in low-light environments [8]. In drones, satellite positioning maps and drone visual imagery during guidance processes can be matched using deep learning to achieve faster and more precise positioning. Gu, Mei-Ying et al. also used deep learning models such as Convolutional Neural Networks (CNNs) and Graph Convolutional Networks (GCNs) to enhance the reliability of visual navigation systems [9]. For instance, in Mohamed O et al.'s experiments, indoor robots achieved topology-based positioning and navigation using CNN-extracted visual features from a small number of image dataset trainings [10]. Alternatively, driver assistance systems may integrate deep learning frameworks such as E2E and VLM. These systems perceive the surrounding environment through sensors like cameras, then employ artificial intelligence to determine optimal guidance routes.

6. Outlook

Although now visual cameras are widely employed in navigation systems, technical deficiencies remain to be overcome. For instance, in deep learning, the high computational power and costly equipment required for such processing make it challenging to deploy on small devices. Relying on cloud computing may lead to significant latency and packet loss. Consequently, model compression or optimization is necessary to deploy complex algorithms directly into navigation terminals, enabling local computation and real-time decision-making. These accelerate the realization of localized deep learning navigation systems.

In practical applications, visual cameras rely on their lightweight and low-cost characteristics to be used in numerous domains. For visually impaired individuals, they can integrate with multi-sensor fusion navigation systems to provide help. Wearable devices, such as smart camera glasses, capture images of the surrounding environment. Utilizing algorithms, these systems can identify objects like people, doors, walls, and furniture, pinpointing targets or obstacles within the surroundings to offer comprehensive information for further guidance. Audio guidance systems and artificial skin tactile feedback can then provide directional feedback [11]. Alternatively, when navigating with mobile devices, integrating GPS with the phone's camera and LiDAR sensors enhances the accuracy, enabling users to reach destinations more easily.

7. Conclusion

This article demonstrates the extensive application value of visual cameras in scenarios such as drone obstacle avoidance, intelligent assisted driving, and smart traffic management. Within the drone sector, binocular vision technology enables critical functions including obstacle recognition and Aruco positioning. Monocular, binocular, and multi-camera visual systems each possess distinct characteristics: monocular vision offers simplicity of structure but lacks direct depth information; binocular and multi-camera systems can derive distance data through parallax calculation, though they face technical deficiencies of high computational demands and significant power consumption.

Multi-sensor fusion overcomes the limitations of individual sensors, with visual cameras playing a crucial complementary role. For instance, integration with the BeiDou-Binocular Collaborative Navigation model effectively solves positioning challenges in environments with absent of satellite signals; combination with LiDAR reduces the performance degradation caused by lighting conditions or adverse weather. The application of deep learning technology significantly enhances visual analysis abilities, improving the stability and accuracy of navigation systems in complex environments.

Looking ahead, visual camera technology must solve the computational restrictions in deep learning models and explore additional application scenarios. Its diverse configurations adapt to varied scenario requirements, while deep learning enhances navigation precision. This research holds reference value for deep learning researchers, researchers in the drone and intelligent driving sectors, smart transportation management departments, navigation system algorithm researchers, relevant

enterprises, and policy-making bodies. Simultaneously, the widespread adoption of navigation systems based on visual cameras can significantly enhance the convenience and happiness of people's daily lives.

References

- [1] Zhou Z X, Hao Y L, Yu P. UAV safety obstacle avoidance domain recognition method based on lightweight binocular vision. *Journal of Xinxiang University*, 2025, (06): 49–53.
- [2] Wang G Y. Design of a small lightweight UAV helipad system. *Science, Technology and Innovation*, 2025, (12): 62–64, 72.
- [3] Jiang L, Fang D J, Lei B, et al. Research status and trends of monocular vision navigation algorithms for mobile robots. *Computer Engineering and Applications*, 2021, 57 (05): 1–9.
- [4] Picard Q, Chevobbe S, Darouich M, et al. A survey on real-time 3D scene reconstruction with SLAM methods in embedded systems. *arXiv preprint*, 2023: arXiv:2309.05349.
- [5] Yang Y, Yu Z. A collaborative navigation model based on multi-sensor fusion of Beidou and binocular vision for complex environments. *Applied Sciences*, 2025, 15 (7912).
- [6] Wang H, Liu J, Dong H, et al. A survey of the multi-sensor fusion object detection task in autonomous driving. *Sensors*, 2025, 25, 2794.
- [7] Liu B Q, Zhang J S, Tang H, et al. A review of laser and visual SLAM algorithms. *Progress in Laser and Optoelectronics*, 2025: 1–23.
- [8] Zhao Z, Wu C, Kong X, et al. Light-SLAM: A robust deep-learning visual SLAM system based on LightGlue under challenging lighting conditions. *IEEE Transactions on Intelligent Transportation Systems*, 2025.
- [9] Gu M Y, Li H, Zhang J W, et al. A review of UAV localization and navigation methods based on vision. *Acta Electronica Sinica*, 2025, 53 (03): 651–685.
- [10] Mohamed O, Mohsen Z, Wageeh M, et al. Autonomous navigation in dynamic environments: Deep learning-based approach. *arXiv preprint*, 2021.
- [11] Tang J, Zhu Y, Jiang G, et al. Human-centred design and fabrication of a wearable multimodal visual assistance system. *Nature Machine Intelligence*, 2025, 7: 627–638.