Comprehensive Research and Discussion on Multi-Sensor Fusion Technology in UAV Vision System

Ziyang Liu *

Department of Aeronautical and Aviation Engineering, Hong Kong Polytechnic University, Hong Kong, China

* Corresponding Author Email: 22100364d@connect.polyu.hk

Abstract. This study comprehensively examines and analyses the application of multi-sensor fusion techniques within UAV vision systems. The study opens with an overview of current research advances in UAV vision systems and multi-sensor fusion technologies and specifies the methodological framework and experimental configuration of this study. Subsequently, the paper explains in detail the fundamental theoretical framework and constituent elements of UAV vision systems, as well as the basic definitions of multi-sensor fusion technologies, their classifications and their significant advantages. On this basis, the authors discuss in depth the specific application scenarios and realisation strategies of multi-sensor fusion technology in UAV vision systems, confirming the effectiveness of this technology in enhancing the effectiveness of UAV vision systems. In the epilogue, the key contributions and results of this study are summarised, and the innovations of the UAV vision system and the multi-sensor fusion technology are discussed, while the limitations are objectively pointed out, and a prospective outlook on potential future research paths is given.

Keywords: Unmanned aerial vehicle vision system; multi-source sensor data fusion technology; real-world application scenarios; efficacy validation methods.

1. Introduction

Against the backdrop of rapid technological evolution, the widespread penetration of UAV technology has brought innovative momentum to many industry sectors. In particular, the UAV vision system, as a key component of this technology, plays a central role in strengthening the UAV's sensitive response to the environment, realising precise real-time navigation, and improving target identification. By integrating data from multiple sensors, a large amount of information from cross-source sensors can be aggregated to enhance the accuracy and comprehensiveness of data analysis. Given the complexity and variability of environmental variables during missions, it is difficult for UAVs to rely on a single sensor configuration to maintain efficient and stable sensory performance. Therefore, the fusion of measurement data from multiple sensors greatly enhances the depth of understanding and adaptability of UAVs to the surrounding environment, ensuring their stable operation and high reliability in complex environments [1]. Specifically, the application of multi-sensor fusion technology in UAV vision systems is mainly manifested in several key aspects: firstly, the technology integrates the outputs of multiple sensing devices, such as visible cameras and infrared sensors, enabling UAVs to continuously track and accurately identify targets in various lighting environments. In addition, it also excels in improving the accuracy of UAV navigation and positioning. For example, when GPS signals are weak or blocked, the UAV can achieve a high degree of accuracy in self-positioning and route planning through the combined analysis of IMU (Inertial Measurement Unit) and visual odometry information. In addition, this technology effectively improves the UAV's obstacle avoidance capability, significantly reducing potential safety hazards during flight.

Multi-sensory system integration technology introduces an innovative strategy for the vision system of UAVs, which enables them to operate stably in complex and changing environments and promotes the broadening of application fields and the improvement of operational efficiency. However, there is still room for deepening and optimising the technology, aiming at adapting it to a wider range of applications and pursuing higher performance standards. Therefore, strengthening the exploration of multi-sensory system integration technology is not only an important driving force for the development of UAV technology but also has a positive radiation effect on other fields of engineering science and technology.

This study is dedicated to exploring the key role and practical value of multi-sensor fusion technology in the field of UAV vision systems and aims to introduce innovative concepts and strategic planning for the development of UAV technology through the integration of theoretical analyses, algorithm development, and field applications to deepen its application and expand its boundaries in a wide range of fields.

2. Research on multi-sensor fusion technology in UAV vision system

2.1. Unmanned Aerial Visions System

As a leading field at the forefront of today's technology, UAV vision systems are fast emerging as a focus of research, integrating the fine integration of multiple sensors and pioneering applications to enhance the optimisation of UAVs in key performance areas such as environmental recognition, target tracking and position determination. The core mechanism of the system's operation lies in integrating the outputs from various sensors and performing in-depth data fusion to create a comprehensive and highly accurate model of the environment for the UAV, which provides a solid foundation for flight safety and stability.

Within the visual sensing system of an unmanned aerial vehicle, the camera plays a pivotal role. It is not only capable of capturing visible light images to provide direct visual information to the operator or automatic navigation system but also capable of recognising and tracking objects with the help of image processing technology. There are many types of cameras, ranging from standard visible-light cameras, and ultra-high-definition cameras, to infrared cameras with night-vision functions [2]. These cameras highlight their irreplaceable and critical roles in multiple application scenarios, such as UAVs performing reconnaissance missions, search and rescue operations, and farmland monitoring.

In addition to camera devices, infrared sensors are also one of the key components in UAV vision systems. They can detect and receive infrared spectrum energy released by objects, ensuring that the UAV can still maintain a high level of environmental awareness at night or under adverse weather conditions. Especially in search and rescue missions, the sensors demonstrate their significant utility by recognising infrared radiation from the human body or other heat sources, thus accelerating the targeting process in complex situations.

LiDAR technology plays a central role within UAV vision systems, providing highly accurate distance measurement and 3D spatial construction. By emitting laser pulses and determining their round-trip time, the technology can create a detailed 3D topographical map of the surrounding environment, providing decisive information for automated UAV navigation and obstacle avoidance. Lidar technology has demonstrated its extensive and far-reaching application value in areas such as terrain mapping, environmental protection monitoring and precision agricultural production.

Overall, current UAV vision technology has made significant progress in in-depth perception and resolution in complex environments by integrating multiple sensing elements such as cameras, infrared sensors and LiDAR [2]. This multi-sensor synergy not only greatly enhances the air navigation safety performance of UAVs, but also opens up new avenues for exploring its applications in numerous fields

2.2. Multi-sensor fusion technology

Multi-sensor information integration (MSI) is a technology that aims to integrate information from multiple sensors to achieve a more detailed and comprehensive understanding of the target or environment. This technology plays a key role in the application of UAV vision systems, especially given the need for UAVs to make rapid and accurate judgements in dynamic and complex environments, where a single sensing device is often unable to meet the high level of demand. By integrating multiple sensor data, drones can better perceive their surroundings, greatly enhancing their decision-making accuracy and reliability.

Multi-sensor data integration algorithms, also known as Multi-sensor Information Fusion (MSI-F), represent a cutting-edge computer processing method. By aggregating data and information from multiple sensors or different information sources, the technology performs an automated analysis and synthesis process, aiming to achieve highly accurate judgement and decision-making and to harvest more refined and reliable insights. The core is to implement a multi-level, multi-oriented information complementary and optimal allocation scheme, to construct a comprehensive and global understanding of the monitored environment [3].

Data integration plays a central role in the process of multi-sensory integration, which essentially involves the collection of raw data from multiple sensory sources and its refinement, aiming at uncovering deeper information assets. Depending on the stage of data processing, the technique can be classified into three basic types: data-level integration, feature-level integration and decision-level integration. Data-level integration, often referred to as pixel-level integration, is at the beginning of the process and acts directly on the underlying data of each sensing device to execute fusion processing strategies, aiming to maximise the preservation of the granularity of the original information, but with the consequent increase in processing complexity, the expansion of the data volume, and the significant consumption of communication bandwidth. The step of feature-level integration, on the other hand, is to merge these features into a unified feature vector after extracting the signature features in the output data of each sensing source. This strategy effectively reduces the load of data processing, accelerates the processing process, and at the same time maintains a high accuracy standard. As for decision-level integration, each perceptual module will first complete the decision or classification task independently, and then converge these individual decision outputs to form a final comprehensive decision conclusion, this approach demonstrates a high degree of error tolerance and flexibility, although it has relatively little direct dependence on individual perceptual sources [4].

In the study of multi-sensing system integration, information from different sensors exhibits complementary qualities. This information creates a mutually complementary effect between different sensing devices, relying on the unique detection technology and perception breadth of each sensor. For example, LIDAR technology is excellent at achieving high accuracy in determining the distance to an object, but has difficulty in acquiring colour and speed of motion information; conversely, a camera can capture rich colours and texture details but may suffer degradation of performance due to adverse weather conditions. Thus, fusing the measurements from these two types of sensors can build a richer and more comprehensive description of the environment, thus enhancing the expressiveness and effectiveness of the overall system performance.

In addition, a significant advantage of multi-sensor fusion technology lies in the inherent redundancy of the system design. By deploying multiple sensors to observe the same target, even if a single sensor fails or the external environment interferes, the system can still maintain continuous data acquisition under the normal operation of the rest of the sensors, which greatly enhances the reliability and stability of the system [5]. This design also allows additional sensing information to be used to improve the accuracy and confidence level of the measurements and to effectively detect and reduce the effects of measurement bias and noise interference by comparing and analysing data from different sensors.

2.3. Main technical approaches to multi-sensor fusion technology

Multi-sensory data fusion techniques involve the integration of a wide range of algorithmic systems, of which weighted averaging and Kalman filtering algorithms have demonstrated a wide range of applications in visual navigation systems for unmanned aerial vehicles [6]. This programme aims to explore in detail the theoretical architecture underlying these key techniques and the mathematical principles that support their operation.

2.3.1. Weighted average method

At the heart of the weighted averaging method lies the integration of specific weight values and the performance of weighted totals. The criteria for establishing these weights implicitly consider elements such as the measurement accuracy of the sensors, the signal-to-noise ratio, and their historical performance, assigning a signature weight value to each sensor individually [7]. For example, in a temperature monitoring scenario, if multiple temperature sensors are deployed in different locations, their measurement accuracy will naturally vary. Through this process, the data provided by each sensor is weighted and averaged to obtain a more accurate overall temperature assessment.

Assuming that there are n sensor measurements z_1, z_2 ... z_n The corresponding weights arew_1 w_2...,w_n then the fusion result of the weighted average method can be expressed as:

$$\overline{z} = \frac{\sum_{i=1}^{n} w_i z_i}{\sum_{i=1}^{n} w_i} \tag{1}$$

2.3.2. Kalman filter

The Kalman filter algorithm achieves the goal of effective data fusion under its anticipatory prediction of future system states and continuous update mechanism. The theoretical cornerstone of the algorithm lies in the use of linear system theory and Gaussian noise assumptions to gradually optimise the judgement of the system state within the framework of Bayesian analysis. The recursive computation strategy adopted enables Kalman filtering to rapidly refresh the state prediction in changing environments, demonstrating excellent real-time data processing capabilities. In addition, the outstanding performance of this technique in reducing the influence of noise and merging multi-source data greatly improves the accuracy of system perception and the effectiveness of decision-making [6]. With the rapid changes in sensing technology and the rapid increase in computing power, it is expected that the application of Kalman filtering technology will usher in a wider range of expansion.

Given the specific application context and system characteristics, Kalman filtering technology is subdivided into several branches, including classical Kalman filtering, extended Kalman filtering, trackless Kalman filtering, information filtering methods, particle filtering technology, the square root form of Kalman filtering, adaptive Kalman filtering and distributed Kalman filtering and many other types [8]. In practice, the standard Kalman filter (KF), its extended form (EKF), the untraceable method (UKF), and the particle filter (PF) are the most common, and these techniques span the field of linear and nonlinear systems and can cope with both Gaussian and non-Gaussian noise scenarios and show a wide range of applicability in various real-world scenarios. In addition, Information Filtering (IF) and Distributed Kalman Filtering (DKF) are valuable in distributed system problems, while Adaptive Kalman Filtering (AKF) and Square Root Kalman Filtering (SRKF) highlight their unique performance advantages under specific constraints [9].

The standard Kalman filtering method is particularly well suited to cope with linear systems and is widely acclaimed for its computational efficiency. The mathematical description of the method is specifically presented as:

Equation of state:

$$x_k = F_{k-1} x_{k-1} + B_{k-1} u_{k-1} + w_{k-1} \tag{2}$$

Observation equation:

$$z_k = H_k x_k + v_k \tag{3}$$

Prediction Steps:

$$\overline{x}_{k|k-1} = F_{k-1} \overline{x}_{k-1|k-1} + B_{k-1} u_{k-1} \tag{4}$$

$$P_{k|k-1} = F_{k-1} P_{k-1|k-1} F_{k-1}^{T} + Q_{k-1} \tag{5}$$

Update Steps:

$$K_k = P_{k|k-1}H_k^T(H_kP_{k|k-1}H^T + R_k)^{-1} \tag{6}$$

$$\bar{x}_{k|k} = \bar{x}_{k|k-1} + K_k(z_k\text{-}H_k\bar{x}_{k|k-1}) \tag{7}$$

$$P_{k|k} = (I\text{-}K_kH_k)P_{k|k-1} \tag{8}$$

Where x_k is the state vector, F_{K-1} is the state transfer matrix, B_{k-1} is the control input matrix, u_{k-1} is the control input vector, w_{k-1} is the process noise (assuming zero-mean Gaussian white noise with covariance Q_{k-1}); and z_k is the observation vector, H_k is the observation matrix, v_k is the observation noise (assuming zero-mean Gaussian white noise with covariance R_k).

Compared to the traditional Kalman filtering methods, the extended Kalman filtering technique effectively adapts to the challenges of nonlinear system environments [8]. The technique implements a linear approximation strategy by employing a first-order Taylor series expansion of the nonlinear function, performing a linearisation operation at each point in time for both the system and the measurement model, and subsequently applying the classical Kalman filtering algorithm. The main differences from the standard Kalman filtering model are the formal differences in the formulation of the equations and the specific steps of the linearisation process, which can be summarised as follows:

Equation of state:

$$x_k = f\,(x_{k-1},u_{k-1}) + w_{k-1} \tag{9}$$

Observation equation:

$$z_k = h\,(x_k) + v_k \tag{10}$$

Where both f and h are nonlinear functions.

Linearization:

$$F_{K-1} = \frac{\delta f}{\delta x}\Big|_{\bar{x}_{k-1|k-1},u_k} \tag{11}$$

$$H_k = \frac{\delta h}{\delta x}\Big|_{\bar{x}_{k|k-1}} \tag{12}$$

3. Application of multi-sensor fusion technology in UAV vision system

3.1. Application of weighted average method in UAV vision system

3.1.1. Multi-camera fusion

Unmanned aerial vehicles are capable of carrying multi-sensor camera systems that enable concurrent recording of the same scene from multiple angles at the same time. The weighting parameters are differentially set according to the resolution of each sensor, current environmental observation conditions and other factors. Using the weighted integration algorithm, the visual information captured by each sensor can be integrated, and the images generated by this strategy show higher resolution and comprehensiveness of information [10]

3.1.2. Sensor Data Fusion

In implementing position estimation, we fused the position information provided by the Global Positioning System (GPS) with visual sensor data [5]. GPS demonstrates extremely high positioning accuracy in open areas, however, in urban environments, its signal is susceptible to interference, which affects the positioning results. To cope with this challenge, we adopt a weighted average method, which dynamically adjusts the weights based on the reliability of each sensor, aiming to achieve more accurate position estimation results under various environmental conditions [6].

3.2. Application of Kalman filtering in UAV vision system

3.2.1. State estimation

The Kalman filtering algorithm is used to achieve real-time estimation of the position, velocity and attitude parameters of a UAV during its flight. This algorithm achieves a highly accurate assessment of the dynamic state of the UAV by fusing the information from the IMU (Inertial Measurement Unit) and the visual sensing device [11]. Specifically, the IMU is responsible for monitoring trends in acceleration and angular velocity over short time intervals, while the visual sensors complement the information to discriminate between positioning coordinates and heading guidance. The Kalman filtering technique, by continuously iterating the state prediction model, can effectively correct for possible drift effects induced by the IMU and mitigate the influence of noise in the input from the visual sensors, thus enhancing the accuracy of the overall estimation [12].

3.2.2. Target tracking

In the process of performing the task of tracking the target, the application of Kalman filtering technology can realise the real-time acquisition and updating of the target position and velocity information. The visual sensor equipment on the UAV captures the target entity, and then uses the Kalman filter algorithm to predict and update the target trajectory operation, in the rapid evolution of the dynamic environment to show outstanding performance, to ensure that the target dynamic characteristics of the rapid response and appropriate adjustment [13].

3.3. Comparison and Synthesis

3.3.1. Comparative advantages

The weighted averaging method is widely recognised for its simplicity and ease of practice. In particular, the method demonstrates its unique value when faced with the integrated analysis of static multi-source sensor data, especially in scenarios where the quality of the data provided by the sensors varies. By intelligently assigning weights to each data source, it effectively facilitates the process of enhancing and optimising the overall data accuracy.

The Kalman filtering algorithm shows excellent performance in the field of coping with dynamic systems and is particularly effective in solving the noise problem of time-series data. In addition, it excels in the accuracy of continuous state estimation, especially in meeting the demanding requirements of real-time system applications.

3.3.2. Combination

The synergy of the two strategies is realised in the architecture of the UAV visual perception system. The weight averaging strategy was adopted in the initial stage to rapidly integrate the output information from multiple sensors and construct a preliminary state prediction model. Subsequent processes incorporate Kalman filtering algorithms for continuous state monitoring and fine-tuning, aiming to strengthen the stability and accuracy of the system [14]. This combined strategy significantly improves the effectiveness of the system in dealing with the challenges of complex environments, ensuring that the UAV demonstrates higher stability, reliability and operational efficiency when performing diverse tasks.

4. Summary

This research is dedicated to the in-depth exploration and wide application of multi-sensor fusion technology in UAV vision systems. Through rigorous analyses and practical tests, we have made several key discoveries and summarised a series of theoretical results.

The first step is to dig deeper into the theoretical level, where we have carefully analysed the underlying principles of multi-sensor fusion technology and explored in detail its potential application in the field of UAV vision systems. By efficiently integrating the information output from multiple sensors, this technology not only improves the accuracy and credibility of the data but also greatly enhances the UAV's ability to perceive the environment and make autonomous judgements. When embedded in the UAV vision system, this technology greatly improves the UAV's navigation accuracy, object recognition and tracking performance, and adaptability in complex situations.

Despite the above results, we are aware that multi-sensor fusion technology is facing a series of challenges and unanswered questions in the field of UAV vision system applications. These include: how to effectively reduce system complexity and cost investment to facilitate wider acceptance and deployment; how to adjust and optimise algorithmic architectures to speed up data processing and improve efficiency to meet the demanding application scenarios; and how to strengthen system stability and improve safety measures to ensure that UAVs can perform stable and continuous missions under various harsh environments. And safety measures to ensure that UAVs can continue to perform their tasks in various harsh environments. These topics will be the focus of our subsequent research.

In summary, this study extensively and deeply analyses the various aspects of multi-sensor fusion technology in UAV vision systems, which not only adds a new dimension to the development of the theoretical system but also lays a solid technical foundation for practical application. There is sufficient reason to believe that, with the continuous promotion of technological progress and the deepening of research, multi-sensor fusion technology will show more significant core performance advantages in the field of unmanned aerial vehicles (UAVs).

Reference:

[1] Research on Autonomous Flight Control of Foreign UAVs [J]. TANG Qiang, ZHU Zhiqiang, WANG Jianyuan. System Engineering and Electronic Technology, 2004(03):418-422.

[2] Multiple View Motion Estimation and Control for Landing an Unmanned Aerial Vehicle. Omid Shakernia. 2002

[3] Research on altimetry technology of shipborne UAV recovery system based on quaternion method [J]. Gao ZF. Aerospace Return and Remote Sensing. 2013(06)

[4] Research on data fusion methods on data level and feature level [D]. Zhang Baomei. Lanzhou University of Science and Technology, 2005.

[5] Data fusion techniques for super-resolution imaging [J]. Deepu Rajan,Subhasis Chaudhuri. Information Fusion. 2002 (1)

[6] A review of multi-sensor fusion [J]. WANG Jun,SU Jianbo,XI Yugeng. Data Acquisition and Processing, 2004(01):72-77.

[7] A mathematical method to improve the positioning accuracy of matched guidance [J]. Wang Gensheng. Tactical Missile Technology, 1992(04):30-34.

[8] An Introduction to the Kalman Filter. Greg Welch, Gary Bishop. 2006

[9] A Kalman filter model for an integrated land vehicle navigation system. Ramjattan, A. N. and Gross, P. A. The Journal of Neuroscience. 1995

[10] A review of multi-sensor image fusion techniques [J]. Mao Shiyi, Zhao Wei. Journal of Beijing University of Aeronautics and Astronautics, 2002(05):512-518.

[11] Fundamentals and applications of Kalman filtering [J]. Peng Dingcong. Software Guide, 2009, 8(11):32-34.

[12] Consideration of time-correlated errors in a Kalman filter applicable to GNSS [J]. M. G. Petovello, K. O’Keefe, G. Lachapelle, M. E. Cannon. Journal of Geodesy. 2009 (1)

[13] Research and application of target detection and tracking technology based on deep learning [D]. Wang Xiaolong. Hubei University of Technology, 2020.

[14] Multisensor Data Fusion Architecture Based on Adaptive Kalman Filter sand Fuzzy Logic Performance Assessment. Escamilla-Ambrosio P J, Mort N. Proceedings of the Fifth International Conference on Information Fusion. 2002.