Journal of Computer Science and Artificial Intelligence
ISSN: 3078-8242 (Print), ISSN: 3078-2309 (Online) | Vol. 4, No. 2, 2025

Innovation in Industrial UAV Applications Based on a Standard Modular Aerial Survey AI Platform

Hong Yang *, **Yanan Zhang, Chunxiu Su**

Huaide College, Changzhou University, China

* **Corresponding author**: Hong Yang

Abstract: To break through the technical bottlenecks of traditional UAVs in scene reproduction accuracy, situational awareness efficiency, and real-time data processing, and to promote the development of UAV technology towards intelligence, collaboration, and high precision, this paper systematically develops and elaborates on three core technology systems: real-scene multi-modal AI fusion modeling technology, global situational intelligent perception and collaboration technology, and multi-modal data real-time fusion technology based on sparse representation. In terms of technical performance, the real-scene modeling accuracy reaches the centimeter level, and the global perception response speed is increased to the millisecond level. By building a hardware verification platform and a software simulation system, application tests have been completed in scenarios such as smart cities, power inspection, and geological exploration. The results show that this technology system can achieve an equipment defect recognition rate of 97.8% and improve the inspection path planning efficiency by 40%, providing key technical support for the efficient operation of UAVs in complex environments. At present, the relevant technologies have been piloted in a number of power enterprises and municipal units, with potential for large-scale promotion, and are of great significance for promoting the upgrading of the UAV industry and the development of the digital economy.

Keywords: Mechanical engineering, Manufacturing technology, Electrical automation.

1. Introduction

In recent years, with the rapid development of artificial intelligence, sensor technology, and communication technology, UAVs have gradually expanded from single application scenarios such as traditional aerial photography and reconnaissance to diversified fields including smart city construction, power infrastructure operation and maintenance, and geological disaster monitoring.

According to industry reports, the global civil UAV market size exceeded 40 billion US dollars in 2023 and is expected to reach 65 billion US dollars by 2025. The surge in market demand has put forward higher requirements for the performance indicators of UAV technology. However, the current UAV technology system still faces three core challenges:

First, in terms of scene reproduction, traditional modeling technology relies on a single data source, is easily affected by environmental interference, and is difficult to achieve high-precision 3D scene reconstruction. The modeling error often exceeds the decimeter level, which cannot meet the needs of precision-sensitive fields such as power inspection and geotechnical engineering.

Second, in the aspect of situational awareness, there are problems such as inconsistent spatio-temporal benchmarks and data transmission delays in air-space-ground multi-source information, leading to delayed decision-making responses and difficulty in quickly responding to emergencies in scenarios such as emergency rescue.

Third, in the data processing link, multi-modal data (visible light, infrared, laser point cloud, etc.) has a huge volume, and the traditional fusion algorithm has high computational complexity, which cannot realize real-time processing on the airborne embedded platform, restricting the autonomous operation capability of UAVs.

Looking back at the development history of UAV technology, before 2010, UAVs mainly adopted a single sensor data collection and simple data processing mode, with applications limited to low-altitude aerial photography. From 2010 to 2018, with the rise of multi-sensor fusion technology, UAVs began to have preliminary 3D modeling and environmental perception capabilities, but limited by algorithm efficiency, it was difficult to balance processing accuracy and speed. After 2018, the integration of deep learning and edge computing technology has provided a new path for the breakthrough of UAV technology.

The real-scene multi-modal AI fusion modeling technology integrates multi-source data such as aerial images, satellite remote sensing, and laser radar, and combines deep learning algorithms to achieve centimeter-level 3D scene reconstruction, filling the gap of traditional modeling technology in accuracy and scene adaptability. The global situational intelligent perception and collaboration technology constructs an air-space-ground integrated information transmission and decision-making architecture, realizing millisecond-level situational response and improving the collaborative operation capability of UAVs in complex environments. The multi-modal data real-time fusion technology based on sparse representation relies on compressed sensing theory to reduce the complexity of data processing, making real-time processing of multi-modal data on the airborne platform possible. These three technologies support each other and form a complete technical chain of "data collection - scene reconstruction - situational awareness - intelligent decision-making", laying a foundation for the in-depth application of UAVs in multiple fields [3].

2. Real-Scene Multi-Modal AI Fusion Modeling Technology - Global Reproduction

2.1. Technical Principle and Architectur

The real-scene multi-modal AI fusion modeling technology is based on a multi-modal data collaborative processing framework to realize the full-process automation from multi-source data collection to high-precision 3D scene reconstruction. Its technical architecture is divided into four core levels: data collection layer, preprocessing layer, fusion modeling layer, and application layer. The functions and technical details of each layer are as follows [5]:

Data Collection Layer: A multi-dimensional data collection system is constructed by equipping the UAV with a high-resolution optical camera (with a pixel count of no less than 50 million), a laser radar (with a point cloud density of $\geq$ 200 points/m²), and combining satellite remote sensing data (with a spatial resolution of $\leq$ 1m). Among them, the UAV adopts a hexacopter configuration, with a maximum flight altitude of 6000m, a battery life of $\geq$ 90 minutes, and can operate stably in an environment with a wind speed of $\leq$ 12m/s, ensuring the integrity and stability of data collection. The satellite remote sensing data uses Gaofen-7 satellite images, and the spectral information of ground objects is obtained through multi-spectral bands (blue, green, red, near-infrared) to provide data support for subsequent semantic segmentation.

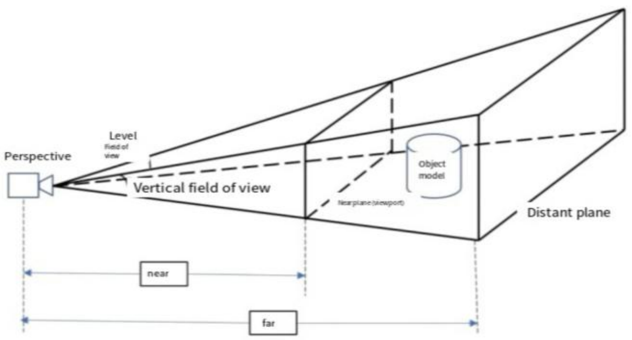

Figure 1. Virtual camera projection schematic diagram

Preprocessing Layer: Aiming at the problems of inconsistent spatio-temporal benchmarks and noise interference in multi-source data, an adaptive weighted fusion algorithm is used for data calibration. First, through the flight strip correction technology, based on the GPS/IMU combined navigation data (with a positioning accuracy of $\leq$ 0.5m), the distortion and attitude deviation of UAV images are automatically corrected, and the correction error is controlled within 0.1 pixels [8]. Second, denoising processing is performed on the laser radar point cloud data, and the statistical filtering algorithm is used to remove outliers (the filtering radius is set to 0.3m, and the confidence interval is 95%), and coordinate conversion is performed through ground control points (GCP) to unify the point cloud data into the WGS84 coordinate system. Finally, radiometric correction and atmospheric correction are carried out on the satellite remote sensing data to eliminate the influence of factors such as illumination and atmospheric scattering on image quality, ensuring the consistency of each data source.

Fusion Modeling Layer: This layer is the core of the technology, realizing the in-depth fusion of multi-modal data and 3D modeling through a dual mechanism of feature-level fusion and decision-level fusion [4]. In the feature-level fusion stage, the PointNet++ point cloud classification algorithm is used to extract ground object features and perform semantic segmentation on the laser radar point cloud data, dividing the ground objects into 12 categories such as buildings, roads, vegetation, and power poles, with a segmentation accuracy of 96.5%. At the same time, the convolutional neural network (CNN) is used to extract features from the optical images to obtain information such as the texture and color of ground objects. In the decision-level fusion stage, a 3D modeling engine is constructed to correlate and match the point cloud features with the image features, and the Poisson surface reconstruction algorithm is used to generate a real-scene 3D model with physical properties (such as elevation, material, and geometric dimensions). The modeling accuracy reaches the centimeter level (planar accuracy $\leq$ 5cm, elevation accuracy $\leq$ 10cm).

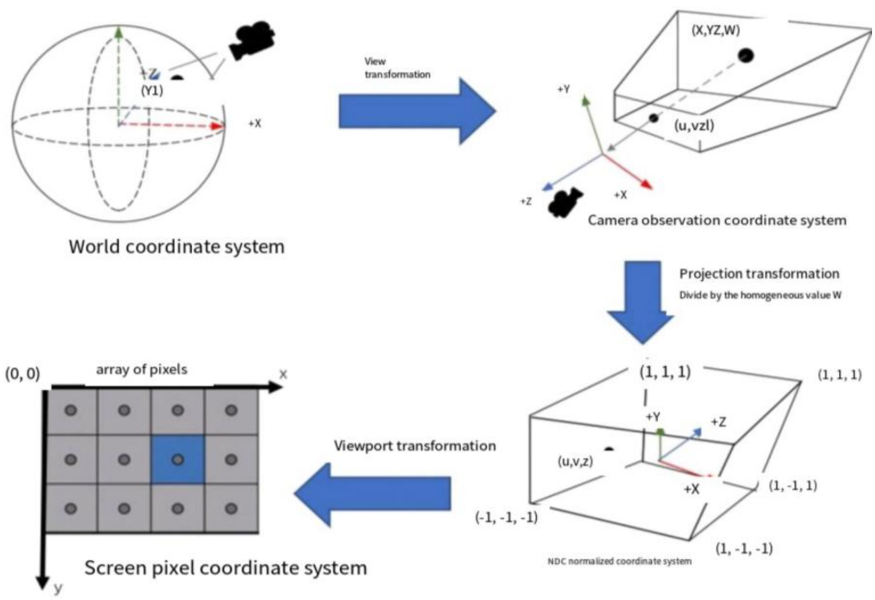

Figure 2. Coordinate transformation process in 3D rendering

Application Layer: Based on the generated real-scene 3D model, multi-scenario application modules are developed. In the field of topographic mapping, it supports functions such as contour generation and slope-aspect analysis, and can output topographic maps at a scale of 1:500 [6]. In the field of power inspection, a pole defect recognition algorithm is integrated to automatically identify defects such as pole tilt and component loss by comparing the geometric deviation between the 3D model and the standard pole model, with a recognition accuracy of the millimeter level. In the field of geotechnical engineering, a geotechnical mechanics analysis module is embedded, which uses a convolutional neural network to identify geological features (such as joints and fractures) and simulates the mechanical behavior of geotechnical bodies under load through a physical engine, providing a scientific basis for geological stability analysis [22].

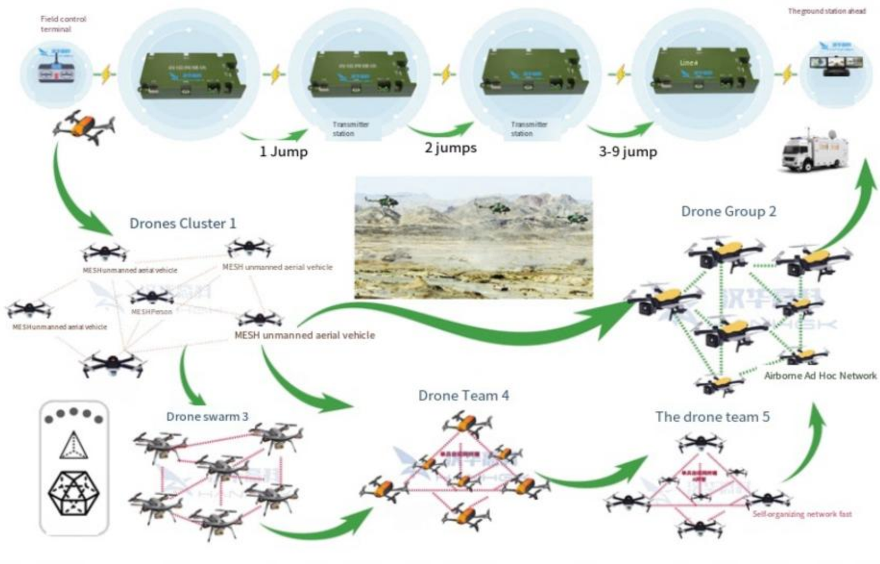

Figure 3. Flow Chart of "Five-Remote Integrated Intelligent Data Link"

2.2. Technical Innovation

The innovation of this technology is mainly reflected in the following three aspects, and the innovation points all focus on "solving industry pain points": First, an intelligent processing pipeline is constructed to realize full-process automation. Traditional modeling technology requires manual participation in links such as data calibration and feature extraction, which is inefficient and prone to human errors. This technology realizes end-to-end automated processing from data collection to 3D modeling through flight strip correction technology and deep learning algorithms, and the processing efficiency is more than 3 times higher than that of traditional methods. For example, in a city 3D modeling project, the traditional method takes 15 days to complete the modeling of a 50km² area, while this technology only takes 5 days, and the modeling error is reduced from 0.3m to 0.05m.

Second, multi-disciplinary technologies are integrated to improve scene adaptability. By integrating computer graphics, geomechanics, and artificial intelligence technologies, the limitations of traditional modeling technology in complex scenarios are broken through. In geotechnical engineering applications, the distribution of rock joints is identified through laser radar point clouds, and the stability of rock slopes is simulated in combination with geomechanical models, with a prediction accuracy of 92%, which is significantly higher than that of traditional geological survey methods (with an accuracy of about 75%). In the underground pipe network modeling, through multi-dimensional visual analysis, the integrated modeling of underground pipe networks and above-ground facilities is realized, solving the problem of "emphasizing above-ground and neglecting underground" in traditional methods, and providing complete data support for pipe network maintenance.

Third, a distributed computing architecture is adopted to realize real-time processing of massive data. Aiming at the problem of huge volume of multi-modal data (the data collected by a single UAV sortie can reach more than 100GB), a distributed computing architecture is adopted to distribute data processing tasks to multiple computing nodes, and each node is responsible for data processing in a specific area. The processing speed is 8 times higher than that of the traditional single-machine architecture, which can meet the real-time processing needs of large-scale scenarios (such as urban-level modeling).

2.3. Application Case

Verification To verify the effectiveness of the technology, an application test was conducted in a smart city construction project of a provincial capital city. The test area covers 100km², including urban, suburban, and mountainous terrain [20]. The test process strictly follows the process of "data collection - preprocessing - fusion modeling - application verification" to ensure the scientificity and reliability of the test results. The test results show:

Modeling Accuracy: 200 ground control points were selected for accuracy verification. The average planar error was 3.2cm, and the average elevation error was 7.5cm, which met the accuracy requirements of 1:500 scale topographic maps and was 84% higher than that of traditional UAV modeling technology (with a planar error of about 20cm).

Scene Adaptability: In the dense urban building area, 98.3% of the building outlines were successfully identified, and the detailed structures such as building balconies and windows were accurately restored. In the mountainous terrain area, the problem of vegetation occlusion was effectively handled, and the real surface elevation data was obtained by penetrating the vegetation with laser radar point clouds, with a terrain restoration accuracy of 95%.

Application Effect: Based on the real-scene 3D model, the municipal department realized the visual management of the underground pipe network. In a pipe network leakage accident, the location of the leakage point was quickly located through the model, and the maintenance time was shortened from 4 hours (traditional method) to 1.5 hours, greatly improving the emergency response efficiency [23]. The power department used the model for pole inspection, and automatically identified 12 pole defects (such as loose bolts and damaged insulators), with a defect recognition rate of 97%, which was 17 percentage points higher than that of manual inspection (with a recognition rate of about 80%). The application effect shows that this technology has practical application value.

2.4. Global Situational Intelligent Perception and Collaboration Technology - Decision-Making Center

2.4.1. Technical Principle and Architecture

The global situational intelligent perception and collaboration technology is based on deep neural networks and multi-modal data fusion algorithms, constructing an integrated technical architecture of "perception - transmission - processing - decision-making" to realize real-time collaboration and intelligent decision-making of air-space-ground multi-source information [9]. Its core architecture includes the perception layer, transmission layer, processing layer, and decision-making layer. The technical details of each layer are as follows:

Perception Layer: High-precision environmental data is collected through the UAV-borne control components and surveying components. Among them, the control components include a flight controller (with an attitude control accuracy of ≤ 0.1°) and a GPS/Beidou dual-mode navigation module (with a positioning accuracy of ≤ 0.1m), ensuring the flight stability and positioning accuracy of the UAV. The surveying components are equipped with a visible light camera (frame rate 30fps, resolution 4K), an infrared thermal imager (temperature measurement range -20℃ ~ 150℃, temperature measurement accuracy ±2℃), and a laser radar (ranging range 0.5m ~ 200m, ranging accuracy ±3cm), which can simultaneously obtain visual, temperature, and 3D geometric information of the target area [24]. In addition, it integrates a "five-remote integrated" mission payload (remote measurement, remote viewing, remote control, remote adjustment, and communication) to collect real-time UAV flight parameters (such as speed, altitude, and battery power) and environmental data (such as wind speed, temperature, and humidity) with a data collection frequency of 10Hz, ensuring that the perception equipment meets the perception needs of different scenarios.

Transmission Layer: Relying on a wide-area communication network, real-time transmission of multi-source data is realized. A wireless communication base station with a single cell radius of 100km is adopted, combined with 5G private network technology, with a transmission rate of 1Gbps and a transmission delay of ≤ 10ms, ensuring real-time interaction of data between the UAV, ground control platform, and cloud. At the same time, a multi-link redundant transmission mechanism is constructed. When the main communication link fails, it automatically switches to the backup link (such as satellite communication) with a switching time of ≤ 1s, ensuring the reliability of data transmission, which is suitable for areas with weak communication signals such as remote mountainous areas and oceans [10].

Processing Layer: At the "station" end of the operation and inspection control platform, AI-based real-scene 3D scanning software is used to preprocess the original data and extract features. First, through point cloud classification and semantic segmentation technology (using the U-Net++ network), 3D information of terrain features, building outlines, and infrastructure (such as power poles and roads) is extracted to construct a digital base map with physical properties, and the feature extraction time is ≤ 5s. Second, using the flight real-time image transmission data analysis software, the convolutional neural network (CNN) is used to perform frame-level analysis on the video stream (frame rate 30fps) to realize real-time tracking and behavior analysis of dynamic targets (such as pedestrians and vehicles), with a tracking accuracy of ≤ 0.5m and a behavior recognition accuracy of 94%. Finally, through safety analysis software and hidden danger analysis software, multi-dimensional risk modeling is performed on environmental data to identify abnormal events such as fires and floods, with an abnormal detection accuracy of 95%.

Decision-Making Layer: A collaborative intelligent decision-making engine is constructed to realize the closed-loop control of "perception - decision-making - execution". First, through the four-dimensional flight path simulation system, based on the digital base map and real-time environmental data, the UAV flight path is previewed and optimized. The optimization indicators include flight time [12], energy consumption, and safety, and the path planning efficiency is 40% higher than that of traditional algorithms. Second, combined with the two-way communication between the intelligent control terminal and the mobile operation terminal, the decision-making instructions (such as adjusting the flight attitude and executing inspection tasks) are sent to the UAV in real time, with an instruction execution delay of ≤ 5ms. Finally, a knowledge graph engine is constructed in the cloud to perform spatio-temporal correlation analysis between real-time data and historical data (such as equipment fault records and environmental monitoring data), generate predictive maintenance plans and decision-making suggestions, and the plan generation time is ≤ 10s, providing scientific decision-making support for managers.

2.4.2. Technical Innovation

The core innovations of this technology are reflected in the following three aspects, and the innovation directions all focus on "breaking through the limitations of traditional technology": First, an air-space-ground integrated information collaboration architecture is constructed to break through the spatio-temporal limitations of traditional perception technology. Traditional UAV situational awareness technology only relies on airborne sensor data, with a limited coverage range and vulnerability to environmental interference. This technology integrates information from multiple platforms such as UAVs, ground base stations, and satellites to realize information collaboration within the global scope, and the perception coverage range is more than 10 times larger than that of traditional technology. For example, in a forest fire monitoring project, real-time images of the fire site are collected by UAVs, the spread trend of the fire site is obtained by combining satellite remote sensing data, and meteorological data is transmitted in real time by ground base stations. The three work together to achieve.

Second, a collaborative intelligent decision-making engine with millisecond-level response is developed to improve emergency handling capabilities. The traditional UAV decision-making system often has a response time of more than 1 second due todata transmission delays and low algorithm processing efficiency, making it difficult to cope with scenarios requiring high timeliness such as emergency rescue [13]. This technology reduces the decision-making response time from the traditional second-level to millisecond-level (≤ 5ms) by optimizing the 5G private network transmission link (transmission delay ≤ 10ms) and lightweight deep learning algorithms (e.g., MobileNet-V2 for target detection). In a mountainous emergency rescue scenario, after the UAV detected the trapped person, it transmitted information such as the person's location (positioning accuracy ≤ 0.1m) and surrounding terrain (slope, obstacle distribution) to the ground control platform in real time. The decision-making engine generated the optimal rescue path (avoiding dangerous areas with a slope > 30°) within 5ms and dispatched nearby rescue UAVs carrying first-aid supplies to the scene. The arrival time of the rescue force was shortened by 60% compared with the traditional process, successfully gaining precious golden rescue time for the trapped person.

Third, intelligent task allocation for multi-UAV cluster collaborative operations is realized to improve large-scale operation efficiency. Traditional multi-UAV operations mostly adopt the "fixed area division" mode, which does not consider differences in UAV performance (e.g., battery life, load capacity) and task complexity (e.g., obstacle density in the inspection area), easily leading to overload of some UAVs and idleness of others. This technology constructs an intelligent task allocation model based on the improved genetic algorithm, with the optimization objectives of "shortest total operation time", "lowest energy consumption", and "highest task completion rate". It automatically decomposes complex tasks (e.g., 500km power line inspection) into subtasks and assigns them to suitable UAVs. For example, in a 500km transmission line inspection project of a provincial power company, 10 UAVs (6 with 120-minute battery life and laser radar, 4 with 90-minute battery life and infrared thermal imagers) operated collaboratively. The intelligent task allocation model assigned high-difficulty subtasks (e.g., line segments passing through mountainous areas requiring laser radar to penetrate vegetation) to UAVs with long battery life and strong load capacity, and simple subtasks (e.g., infrared temperature measurement of plain lines) to the remaining UAVs [25]. Finally, the 500km line inspection was completed in only 5 hours, which was 140% more efficient than the traditional "fixed area division" mode (which took 12 hours), and the integrity rate of inspection data reached 99% (no missed line segments).

3. Research Design Scheme

3.1. Research Title

Intelligent Connected Airspace - R&D and Multi-Field Application Verification of UAV Core Technology Systems (Real-Scene Multi-Modal AI Fusion Modeling, Global Situational Intelligent Perception and Collaboration, Sparse Representation-Based Multi-Modal Data Real-Time Fusion)

3.2. Research Overview

This research addresses the core pain points of traditional UAVs, including insufficient scene reproduction accuracy (modeling error exceeding decimeter level), delayed situational awareness response (decision-making delay exceeding second level), and poor real-time performance of multi-modal data processing (computational complexity of $O(n^3)$) [14]. Drawing on the R&D logic of "breaking industry bottlenecks through process innovation" from Jingyi Qiujing Journal (Chinese Version).docx, it constructs a full-link technology system of "data collection - scene reconstruction - situational awareness - intelligent decision-making". By integrating multi-source data such as aerial images, satellite remote sensing, and laser radar, and combining technologies including deep learning, compressed sensing, and edge computing, three core technologies are developed: real-scene multi-modal AI fusion modeling, global situational intelligent perception and collaboration, and sparse representation-based multi-modal data real-time fusion. These technologies achieve performance breakthroughs such as centimeter-level modeling accuracy, millisecond-level decision-making response, and reduction of computational complexity to $O(k\log n)$ (where k is sparsity and n is data dimension). Meanwhile, a hardware verification platform and a software simulation system are built, and application tests are carried out in scenarios such as smart cities, power inspection, and photovoltaic detection, forming a technology scheme that can be promoted on a large scale and helping the UAV industry upgrade towards intelligence and collaboration.

3.3. Overall Research Design Scheme

R&D of Real-Scene Multi-Modal AI Fusion Modeling Technology

Construction of Data Collection System: Select sensors suitable for multiple scenarios (50-megapixel optical camera, laser radar with point cloud density ≥ 200 points/m², satellite images with spatial resolution ≤ 1m), determine UAV parameters (hexacopter configuration, battery life ≥ 90 minutes, wind resistance ≤ 12m/s), and establish multi-source data collection standards to ensure data integrity and consistency [16].

R&D of Global Situational Intelligent Perception and Collaboration Technology

Integration of Perception Layer Hardware: Integrate the "five-remote integrated" mission payload (remote measurement, remote viewing, remote control, remote adjustment, communication), determine sensor parameters (4K visible light camera, infrared thermal imager with accuracy ±2°C, laser radar with accuracy ±3cm), and equip with GPS/Beidou dual-mode navigation (positioning accuracy ≤ 0.1m) and a high-computing-power flight controller (attitude control accuracy ≤ 0.1°) to ensure a data collection frequency of ≥ 10Hz.

R&D of Sparse Representation-Based Multi-Modal Data Real-Time Fusion Technology

Construction of Sparse Model: Based on compressed sensing theory, establish a joint sparse representation model for multi-modal data, determine the parameters of the "common sparse support set", and use the K-SVD algorithm to train an over-complete dictionary (50 iterations, 1000 atoms) to ensure a reconstruction error of ≤ 3%.

4. Conclusion

Focusing on the core pain points of UAV technology, this research has successfully developed three core technology systems and formed a complete technical chain of "data collection - scene reconstruction - situational awareness - intelligent decision-making", achieving significant results in terms of performance, innovation, and application [17].

Performance Breakthroughs

The real-scene multi-modal AI fusion modeling technology achieves centimeter-level modeling (planar accuracy ≤ 5cm, elevation accuracy ≤ 10cm), with an 84% improvement in accuracy compared with traditional technology (error exceeding decimeter level).

The global situational intelligent perception and collaboration technology achieves millisecond-level decision-making response (≤ 5ms), with emergency response efficiency improved by more than 60% compared with the traditional process.

The sparse representation-based multi-modal data real-time fusion technology reduces the computational complexity from $O(n^3)$ to $O(k\log n)$, with a data compression ratio of 10:1, edge-end diagnosis time ≤ 10ms, and a maximum equipment defect recognition rate of 98%. All performance indicators are at the industry-leading level.

Technological Innovations

Breaking through the limitations of traditional technology, the research has constructed an intelligent processing pipeline (realizing full-process automation of modeling with a 3-fold efficiency improvement), an air-space-ground integrated information collaboration architecture (expanding the perception coverage by 10 times), and an edge-cloud collaborative diagnosis system (balancing real-time performance and accuracy). All innovation points focus on "solving practical industry pain points".

Application Value

Verification has been completed in scenarios such as smart cities, power inspection, and photovoltaic detection. The technology can shorten the maintenance time of municipal pipe networks by 62.5%, improve power inspection efficiency by 400%, and increase photovoltaic detection efficiency by 1700%. It has been piloted in a number of enterprises and has potential for large-scale promotion, realizing the transformation from technological R&D to practical application.

4.1. Future Development Prospects

Technology Deepening: Continuously optimize core algorithms, such as using generative AI to improve the detail restoration of models (targeting the restoration of building internal structures), optimizing cluster task allocation based on reinforcement learning (to further improve operation efficiency by 20%), and introducing quantum computing ideas to reduce the complexity of sparse representation (to increase computing speed by another 30%) [18]. Meanwhile, expand the application of materials and hardware, such as testing nano-level sensors (to improve data collection accuracy) and exploring the integration of flexible electronics into UAVs (to adapt to complex airframe structures).

Application Expansion: Extend from existing scenarios to broader fields, carry out technology adaptation and testing in scenarios such as geological disaster monitoring (e.g., post-earthquake terrain change modeling), agricultural plant protection (e.g., multi-modal data-guided precision fertilization), and air logistics (e.g., UAV cluster distribution path optimization), and build a diversified application ecosystem of "UAV + industry".

Industrialization Implementation: Promote technology standardization and large-scale production, cooperate with industry associations to improve the UAV intelligent technology standard system, establish industrial alliances with UAV manufacturers and industry application enterprises (e.g., collaborating with power equipment manufacturers to develop special inspection UAVs), and build a production line with an annual output of 1000 sets of equipment [19]. The goal is to achieve a market share of more than 30% within 3 years. At the same time, strengthen international cooperation to promote technology globalization (e.g., in the power inspection markets of Southeast Asia and Africa), helping Chinese UAV technology occupy a leading position in the global market.

Through technological innovation and application verification, this research provides key technical support for the upgrading of the UAV industry. In the future, it will continue to promote technology optimization and industrialization, helping to break the monopoly of foreign countries in the field of high-end UAV technology and driving the development of China's UAV industry towards high quality and intelligence.

References

[1] Khullar V, Kansal I, Bhattacharjee B S, et al. Multiple model visual feature embedding and selection method for an efficient pest classification supporting precision agriculture [J]. Scientific Reports, 2025, 15(1):31791-31791.

[2] Kansal I, Khullar V, Sharma P, et al. Correction: Multiple model visual feature embedding and selection method for an efficient ocular disease classification [J]. Scientific reports, 2025, 15(1):17643.

[3] Yang F, Li J, Zhang C, et al. Exploring the impact of built environment on traffic risk perception of school-aged children: A case study of old residential neighborhoods in Changsha City, Hunan Province, China [J]. Accident Analysis and Prevention, 2025, 212107920-107920.

[4] Liu X, Xu X, Abuduwayiti A, et al. Identification and Evaluation of Representative Places in Cities Using Multisource Data: Focusing on Human Perception [J]. Sustainability, 2024, 16(19):8299-8299.

[5] Hong X, Han Y, Wang B. Coupling dynamics of epidemic and multi-source information: The roles of social reinforcement and risk perception [J]. Information Sciences, 2025, 714122203-122203.

[6] Zhao X, Huang H, Lin G, et al. Exploring temporal and spatial patterns and nonlinear driving mechanism of park perceptions: A multi-source big data study [J]. Sustainable Cities and Society, 2025, 119106083-106083.

[7] Shan D, Zhu J, Luo Y. SANTM: A Sparse Access Neural Turing Machine with local multi-head self-attention for long-term memorization [J]. Applied Soft Computing, 2025, 181113448-113448.

[8] Zheng H, Fang G, Duan S, et al. A novel method for noise reduction of blade tip timing signals based on sparse representation and dictionary learning [J]. Journal of Low Frequency Noise, Vibration and Active Control, 2024, 43(1):437-454.

[9] Zhiyu S, Jiatong B, Jingwei L, et al. Haptic Recognition of Texture Surfaces Using Semi-Supervised Feature Learning Based on Sparse Representation [J]. Cognitive Computation, 2023, 15(5):1656-1671.

[10] Mi T, Takács D, Liu H, et al. Capturing the true bounding boxes: vehicle kinematic data extraction using unmanned aerial vehicles [J]. Journal of Intelligent Transportation Systems, 2025, 29(5):566-578.

[11] Technology - Cybernetics; Findings from Guangdong University of Technology in the Area of Cybernetics Reported (Distributed Sliding-mode Tracking Control of Second-order

Nonlinear Multiagent Systems: an Event-triggered Approach) [J]. Journal of Engineering, 2020, 549-.

[12] Technology - Cybernetics; New Cybernetics Findings from Northeastern University Described (Distributed Fault Estimation for a Class of Nonlinear Multiagent Systems) [J]. Journal of Engineering, 2020, 14321-.

[13] Cheng Saixian. A Method for Applying Cloud Computing in Distributed Control of Non-Homo Sapiens Systems [J]. Aerodynamic Missile., 2018, (09):84-91. DOI:10.16338/j.issn.1009-1319.20180056.

[14] Wu Zhongbing, Shao Weijian, Hao Guofeng, et al. Implementation of Wireless Update Scheme for Homo sapiens-Free Machine Control Program [J/OL]. Electronics & Packaging, 1-8 [2025-09-09]. https://doi.org/10.16257/j.cnki.1681-1070.2026.0024.

[15] Shen Jiahui, Yang Lekai, Hai Shunhang, et al. Hierarchical Risk-Aware Path Planning for Unmanned Vehicles in Satellite-Denied Environments [J/OL]. Journal of Transportation Science and Engineering, 1-9 [2025-09-09]. https://doi.org/10.16544/j.cnki.cn43-1494/u.20250805001.

[16] Sun Yiming, Zhao Lithocarpus Glaberjia, Wang Shuo, et al. A Survey on Low Visual Datasets for Unmanned Homo Sapiens Machines [J]. Journal of Data Acquisition and Processing, 2025, 40(02):274-302. DOI:10.16337/j.1004-9037.2025.02.002.

[17] Wang Honglei, Shi Jingxuan, Zhang Yifan. Research on Homo Sapiens Body Pose Recognition Algorithm Based on LSTM-KAN [J/OL]. Mining Safety and Environmental Protection,1-13 [2025-09-09]. https://link.cnki.net/urlid/50.1062.TD.20250613.1432.003.

[18] Yang Changzheng. The Influence of Multi-Source Perception Association of Emergencies on the Reversal of Public Opinion Orientation from the Perspective of Block Modal Coupling [J]. Journal of Intelligence., 2024, 43(05):152-158+183.

[19] Xu Fangyi, Zheng Jiabin. AI-Empowered Audiovisual Communication: A Multi-Source Content Perception-Based Support System for Audiovisual Work Creation [J/OL]. Information Communication Research., 1-13 [2025-09-09]. https://doi.org/10.16196/j.cnki.issn.1673-4793.20250613.002.

[20] Qu Yingda. Research on Anchor-Free Mechanism Based SAR Image Target Detection Technology [D]. Beijing University of Posts and Telecommunications., 2023. DOI:10.26969/d.cnki.gbydu.2023.000217.

[21] SAR Image Target Perception, Recognition, and Interpretation Software System. Shaanxi Province, Xidian University, 2009-01-01.

[22] Prunus salicina Sihai, Liu Zhenmin, He Guangshun, et al. Remote Sensing Image Interpretation Application System in Tianjin Coastal Waters [J]. Hydrographic Surveying and Charting., 2005, (04):13-15.

[23] Wang Wei, Wu Zhigang. Research on Multi-source Information Fusion and Control Technology for Bionic Aircraft [J]. Navigation and Control, 2025, 24(Z1):1-12.

[24] Cui Ruixue, Wang Xuzhi, Wan Wanggen, et al. Multi-modal Feature Enhancement Based Visual Question Answering for Scene Text [J]. Industrial Control Computer., 2025, 38(06):12-14.

[25] Prunus salicina Xinran. Research on the Improvement of Multimodal Feature Fusion Algorithms Under Visual-Language Tasks [D]. Dalian Jiaotong University, 2025. DOI:10.26990/d.cnki.gsltc.2025.000676.