Panoramic UAV Surveillance and Recycling System Based on Structure-Free Camera Array

TAO YANG[1], (Member, IEEE), ZHI LI[1], (Student Member, IEEE),
FANGBING ZHANG[1], (Student Member, IEEE), BOLIN XIE[1],
JING LI[2], (Member, IEEE), AND LINFENG LIU[1]

[1]National Engineering Laboratory for Integrated Aero-Space-Ground-Ocean Big Data Application Technology, School of Computer Science, Northwestern Polytechnical University, Xi'an 710129, China
[2]School of Telecommunications Engineering, Xidian University, Xi'an 710071, China

Corresponding author: Tao Yang (tyang@nwpu.edu.cn)

ABSTRACT In recent years, unmanned aerial vehicles (UAVs) have rapidly developed, but the illegal use of UAVs by civilians has resulted in disorder and security risks and has increasingly triggered the community concern and worry. Therefore, the monitoring and recycling of UAVs in key regions are of great significance. This paper presents novel panoramic UAV surveillance and autonomous recycling system that is based on the unique structure-free fisheye camera array and has the capability of real-time UAV detection, 3D localization, tracking, and recycling capacity over a very wide field of view. The main characteristics of this paper include the following: 1) constructing a structure-free camera-array-based panoramic UAV surveillance and recycling system; 2) designing a robust dynamic near-infrared laser-source-based self-calibration algorithm for large-scale arbitrary layout of the camera array; and 3) presenting a set of UAV detection, 3D localization, tracking, and autonomous recycling algorithms based on a fish-eye camera array. The system has been tested in various challenging scenarios, including multiple UAVs with significant appearance and scale changes and even poor weather conditions. The extensive experimental results analyzed both qualitatively and quantitatively, and the analysis of the time performance demonstrate the robustness and effectiveness of the proposed system. In addition, we successfully conduct a recycling and landing experiment with the Parrot Beobop.

INDEX TERMS UAV, structure-free fisheye camera array, surveillance and recycling system.

I. INTRODUCTION

With the maturity of UAV technology, UAV manufacturing and use costs are decreasing, large numbers of UAVs are being researched and developed, and manufacturing and application companies have boosted the popularity of civilian UAVs [1]. Civilian UAVs have exhibited a worldwide upsurge, with a large number of applications [2]–[4] but also many problems and challenges. On the one hand, a large number of illegal uses cause security risks [5]–[7], such as stealing private information, releasing hazardous materials and even committing terrorist attacks, which have prompted increasing community concern and worry. Therefore, it is necessary to deploy and install UAV surveillance systems at airports and key monitoring points to maintain the safety of these areas. On the other hand, accurate autonomous recycling of UAVs and cooperative landing of UAV-UGVs in low-altitude and GPS-denied areas are very urgent and important, such as the application of UAV logistics transport and recycling on vehicle platforms and aircraft carriers. Consequently, UAV surveillance and autonomous recycling technology operating over wide areas, a relatively new and largely unexplored research topic, is of great value and significance.

In recent years, wide-area UAV surveillance has been widely realized in anti-UAV technology and systems. From a technical view, there are several major types of anti-UAV surveillance and defense technologies, such as signal interference, radar detection, and integrated technology. Signal jamming technology can affect the UAV GPS signal receiver, where the control and navigation information is disturbed

The associate editor coordinating the review of this manuscript and approving it for publication was Md. Moinul Hossain.

by strong interference. One example is the electromagnetic gun, which disrupts the auto-driving system or the communication system of UAVs. Radar detection technology has powerful detection ability for small and low speed targets. In the Swedish "GIRAFFE" radar system [8], the cross section is accurate to 0.001 square meters. Italy Finmeccanica Selex ES company has demonstrated its "falcon shield" UAV system [9], which can locate, identify and control long-range small UAVs in a short time. Integrated technology has the ability to be more powerful. The British anti-UAV defense system AUDS [10] can detect, track and destroy small and large UAVs. Although the existing UAV surveillance and defense systems can serve as suitable inspiration for research on UAV surveillance, there are still many issues requiring solutions. First, UAVs are generally small in size and weight, and they appear suddenly, which makes them hard to detect, particularly in low-altitude areas. Second, the integration of artificial intelligence and advanced materials into the design of UAVs will significantly improve their capability in stealth. Third, the majority of existing anti-UAV technologies are designed for specific UAVs or scenarios via few technical methods, which limits their applications and generalization.

To solve the UAV surveillance and recovery problem in low-altitude and GPS-denied areas, many high-precision measuring sensors and methods are used, such as laser range scanners [11], [12], ultra-wideband-based pose estimation and recycling methods [13]. Despite their effectiveness, these methods have some disadvantages, such as complex structures and the challenge of deployment for specific expensive sensors or devices. The rapid development of intelligent vision technology and visual sensors, such as monocular cameras [14], [15], stereo cameras [16], [17], and RGB-D sensors [18], [19], has become the core part of the UAV surveillance and recycling system [20], [21]. Martínez *et al.* [22] design a trinocular system, which is composed of three FireWire cameras fixed on the ground, to estimate the vehicle position and orientation by tracking the markers on the UAV. Pebrianti *et al.* [23] propose a ground-based stereo vision system to estimate the three-dimensional position of a quadrotor. Although those systems achieve satisfactory results, they can be used only in an indoor environment due to the limited distance measurement. To enlarge the surveillance range and improve the recycling capacity, various systems and solutions have been proposed in recent years [24]–[29]. Kim *et al.* [30] present an autonomous vision-based net recovery system for small fixed-wing unmanned aerial vehicles. Kong *et al.* [31], [32] construct a system that mounts two separate sets of pan-tilt units integrated with visible light cameras on both sides of the runway, which can detect a fixed-wing UAV at approximately 600 meters. Yang *et al.* [33] present a high-accuracy large-scale outdoor camera array calibration method in their work from 2015. By integrating an infrared camera and laser lights, their system can robustly detect UAVs and safely guide their landing within a range of 800 meters. Although those systems [31]–[33] achieve satisfactory performance on cooperative targets, they cannot

be applied to non-cooperative UAV surveillance. Moreover, those system are designed mainly for fixed-wing UAV recycling using long-focus lenses; due to the small field of view, they cannot be used for large-area UAV surveillance and recycling tasks directly.

For the UAV surveillance and recycling system, wide-area monitoring capacity and high-precision localization capacity are the two most important aspects. To solve these key technical problems, this paper propose a novel panoramic UAV surveillance and autonomous recycling system based on a fisheye camera array; the architecture of the system is shown in Fig. 1. The overall structure and algorithm flow of the proposed system is shown in Fig. 2, which mainly consists of two main parts. One is the offline part (Section II), including the structure-free fisheye camera array imaging module and the camera array self-calibration module. Another is the online part (Section III), including the camera array collaborative multi-UAV detection, 3D localization, and tracking and UAV autonomous recycling module. The contributions of this work are as follows:
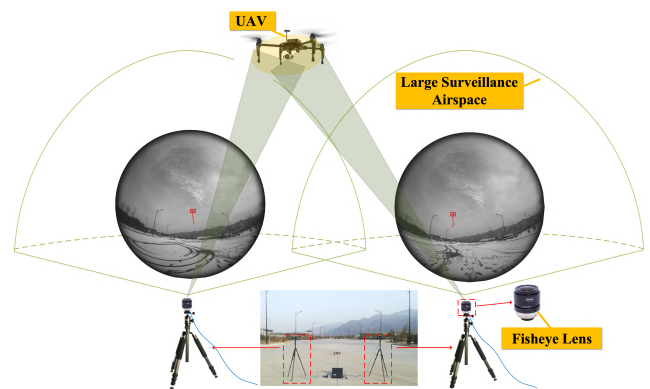

FIGURE 1. Architecture of the panoramic UAV surveillance and autonomous recycling system based on structure-free fisheye camera array.

- We construct a novel real-time panoramic UAV surveillance and autonomous recycling system based on a structure-free fisheye camera array for the first time, to the best of our knowledge. The fisheye camera array ensures a large monitoring FOV, and the unstructured camera array architecture is convenient for rapid deployment and flexible configuration.
- To address the large-scale unstructured camera array fast calibration problem, we design a robust dynamic near-infrared laser-source-based fast self-calibration algorithm. In this algorithm, an airborne laser source is used as a collaboration marker. First, many light correspondence points between cameras are generated by UAV flight, and then the camera pose is defined by these correspondence points. Extensive experiments prove its effectiveness and high efficiency for different camera array structures, particularly in environments with a weak texture or solid color background.

- We present an intelligent synergetic UAV detection, 3D localization, tracking and autonomous recycling algorithm based on a fisheye camera array. This algorithm can be used effectively for low-speed and low-strength UAVs, particularly in low-altitude and GPS-denied environments. The whole algorithm uses the ROS platform for parallel processing, which can run on a laptop and on an embedded computing platform in real time.
- We conduct extensive real experiments in complex and challenging scenarios using four different types of UAVs with discrepant appearance and size. Our algorithm exhibits strong performance under many challenging conditions, including the large-scale changes, the interleaving of UAVs and even low-speed flights, thus demonstrating the effectiveness and robustness of the proposed system.

This remainder of this paper is organized as follows. The structure-free fisheye camera array imaging system is presented in Section II.A, and the dynamic laser source based camera array self-calibration method is described in Section II.B. Section III.A-C describes the detailed algorithm of the proposed system. Section IV verifies the proposed system through several experiments. Finally, the conclusions are given in Section V.

II. PANORAMIC CAMERA ARRAY IMAGING SYSTEM AND SELF-CALIBRATION METHOD

This paper focuses on a panoramic UAV surveillance and recycling mission based on a structure-free fisheye camera array. This section mainly introduces our ground-based unstructured fisheye camera array imaging system and the dynamic laser-source-based self-calibration method.
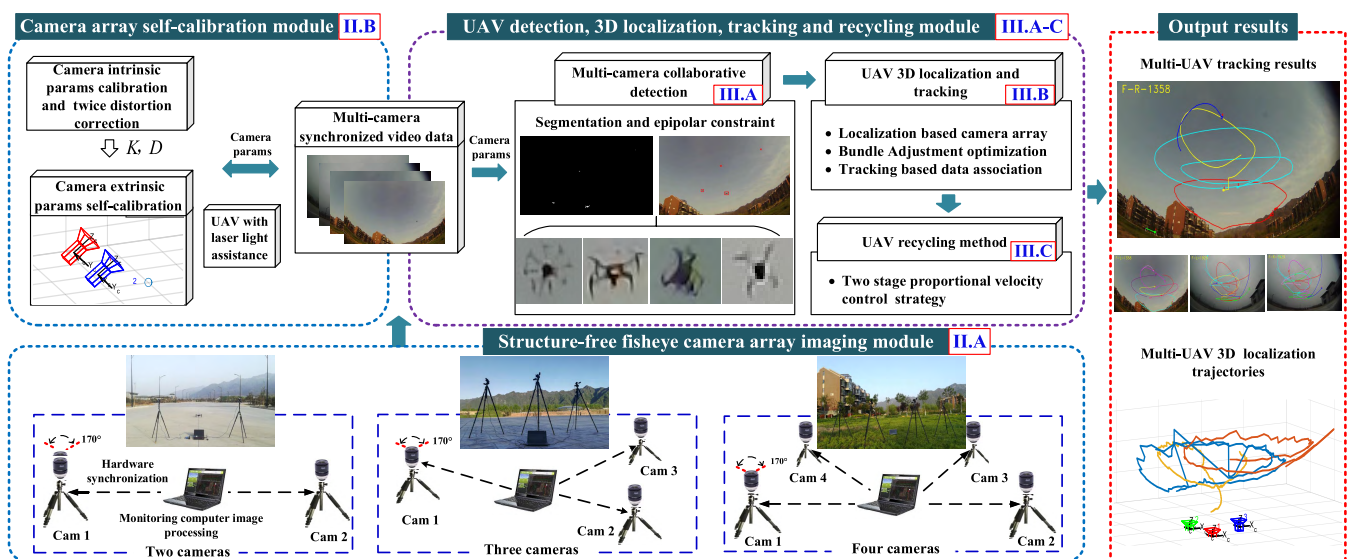
A. STRUCTURE-FREE FISHEYE CAMERA ARRAY IMAGING SYSTEM

1) FISHEYE OPTICAL IMAGING MODULE

To cover the larger monitoring area, we construct a novel panoramic UAV surveillance and autonomous recycling system based on a structure-free fisheye camera array. The system consists mainly of several fisheye cameras, which are fixed on the pan-tilt units. These units are mounted on tripods or the vehicle platform, and the entire device can be rotated, allowing quick adjustment of the camera angle. We select two kinds of cameras in the experiment, which are shown in Table 1. The first imaging module combines a PointGrey GS3-U3-41C6NIR-C near infrared camera and Kowa LM4NCL fisheye lens. The camera has a powerful frame rate acquisition capability of 1024×1024 pixels at 90 fps with a large FOV of $150°$. The second imaging module is the integrated USB module camera, of which focal length is 3.8 mm with a wide FOV of up to $170°$. The cameras use the USB3.0 data interface to satisfy the requirements of high-speed image synchronization acquisition, which captures the images at 30 fps with a resolution of 1280×720 pixels. The detailed camera and lens parameters are shown in Table 1.

2) STRUCTURE-FREE CAMERA ARRAY ARCHITECTURE

We design diversified camera array structures ranging from two to four cameras with random positions, which is shown in the lower left corner of Fig. 2. Each camera is coupled to the entire camera array network as an independent node, and the images of all cameras are transmitted to the ground processor by data synchronization. The camera array ensures a large FOV and the stability and robustness of the system, and the free structure is convenient for rapid deployment and configuration flexibility.

FIGURE 2. Overall structure and algorithm flow of proposed panoramic UAV surveillance and recycling system. This system contains two main parts: one is the offline part (Section II), including the structure-free fisheye camera array imaging module (Section II.A) and the camera array self-calibration module (Section II.B). The other is the online part (Section III), including camera array collaborative multi-UAV detection (Section III.A), 3D localization and tracking (Section III.B) and UAV autonomous recycling (Section III.C).

TABLE 1. Camera and fisheye lens parameters.

Camera/Lens	Specification	Parameter
	Sensor	CMV4000-3E12
	Resolution	1024×1024 pixels
	Max. frame rate	90 fps
PointGrey	Working voltage	DC 4.5v
	Focal length	3.5 mm
	Angle of view	150°
	TV distortion	-28.0%
Kowa LM4NCL	Weight	60 g
	Focal length	3.8 mm
	Resolution	1280×720 pixels
	Max. frame rate	30 fps
USB camera	Angle of view	170°

B. DYNAMIC LASER-SOURCE-BASED SELF-CALIBRATION METHOD

Camera array calibration methods have been widely studied, including visible-light and IR cameras. For example, Ellmauthaler *et al.* [34] propose a novel iterative calibration approach for thermal infrared cameras using a calibration device consisting of 81 light bulbs, arranged in a 9×9 matrix. These previous works are mostly based on calibration plates, designed markers or the structural information in the scene, which are aimed at small-scale camera array calibration. Considering that the proposed panoramic structure-free fisheye camera array imaging system focuses mainly on a wide airspace, it is challenging to place any available markers in the airspace. Therefore, the conventional methods are not valid for our proposed camera array imaging system.

In this paper, we propose a robust dynamic laser-source-based camera array fast self-calibration algorithm, which effectively solves the large-scale unstructured camera array calibration problem. The dynamic laser source is generated by the UAV flying in the air with a laser lamp. In this method, on the one hand, the near-infrared laser lamp has favorable illumination characteristics, large brightness, satisfactory parallelism, small divergence and a more concentrated imaging spot, which can effectively image over a long distance and wide angular range. On the other hand, a UAV with a laser lamp can generate dynamic laser sources in large airspace through free flight, which is suitable for large-scale monitoring of camera arrays. Moreover, this method is simple to use and can be quickly re-calibrated when the structure of the camera array changes. This calibration process is divided into three parts, including the fisheye lens double distortion correction, extraction of laser lamp from the image and extrinsic parameter calibration. These three parts are described in detail.

1) FISHEYE LENS AND DOUBLE DISTORTION CORRECTION

The fisheye lens is an ultra-wide lens that can provide a wide field of view both horizontally and vertically; the imaging model is shown in Fig. 3. The image captured by the lens is highly distorted. The most widely used lens is the equiangular fisheye lens, which features a constant angular resolution. These parameters are characterized by the equality

$$r = f\theta \tag{1}$$

where f is the focal length of the lens, θ is the incoming ray angle and r is the projection of the ray on the image plane. This relationship is a theoretical one. In reality, there is deviation from this equality due to imperfections in the lens production and assembly.
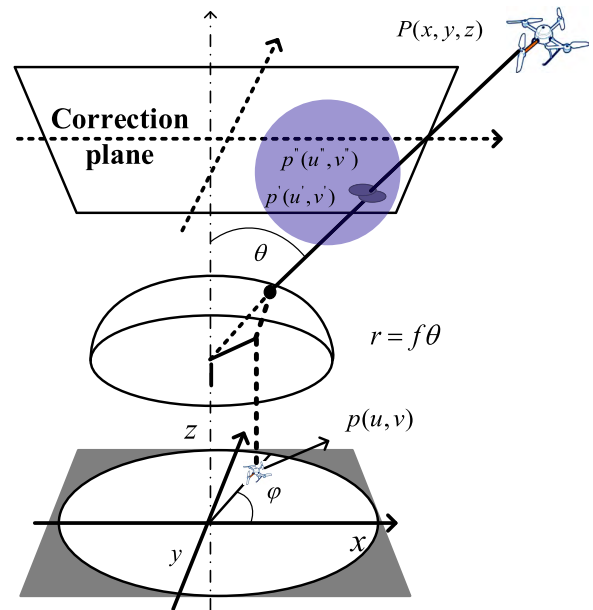

FIGURE 3. Fisheye camera model. The $P(x, y, z)$ is imaged at $p(u, v)$ by the fisheye camera, and $p'(u', v')$ and $p''(u'', v'')$ are the first and second correction results, respectively.

To accurately calibrate the intrinsic parameters of the fisheye camera and remove the distortion, the double distortion correction process is carried out. For the first calibration, we use the fisheye camera model, which derives a projection function that can accommodate any type of ultra-wide lens. Briefly,

$$\theta_d = \theta(1 + k_1\theta^2 + k_2\theta^4 + k_3\theta^6 + k_4\theta^8) \tag{2}$$

We can easily obtain the intrinsic parameters K_i and distortion coefficient $D_i(k_1, k_2, k_3, k_4)(i = 1, 2, 3...)$ of each camera. The first correction result

$$p'(u', v') = K \cdot \theta_d \begin{bmatrix} \cos\varphi \\ \sin\varphi \end{bmatrix} \tag{3}$$

where $p'(u', v')$ are the first corrected image coordinates, θ_d is defined in (2), φ is the angle between the radial direction and the x-axis, and K represents the camera intrinsic parameters.
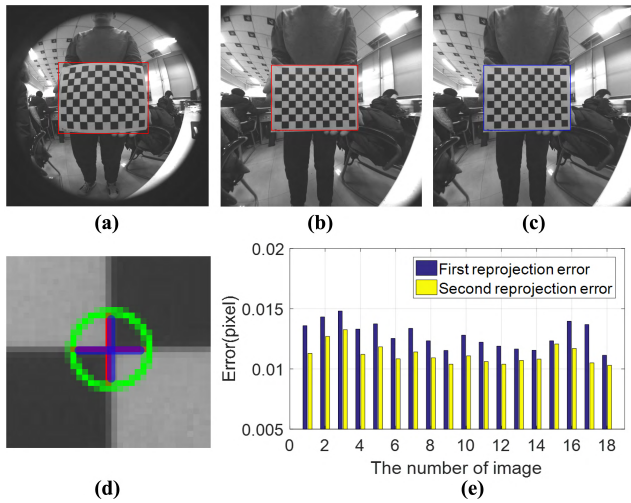
(a) **(b)** **(c)**

(d) **(e)**

FIGURE 4. Distortion correction of the fisheye camera. (a) Original distorted image. (b) First undistorted result. (c) Second undistorted result. (d) Corner reprojection result. (e) First and second reprojection pixel error.

As shown in Fig. 4(b), there are still some local residuals, so we carry out the second fine calibration. In the second calibration, we calibrate the images after the first correction using a pinhole camera model. $p''(u'', v'')$ can be defined as

$$\begin{cases} u'' = u' + (u' - c_x) \cdot k \cdot r^2 \\ v'' = v' + (v' - c_y) \cdot k \cdot r^2 \end{cases} \quad (4)$$

where $p''(u'', v'')$ defines the second correction image coordinate, k is the radial distortion coefficient by the second calibration, (c_x, c_y) is the optical center, and r is the distance to the optical center. The original image and double correction results are shown in Fig. 4; the final calibration accuracy is better than 0.015 pixel.

2) EXTRACTION OF LASER LAMP FROM IMAGE

After obtaining the intrinsic parameters and distortion coefficients of the fisheye lens, we need to calibrate the camera array with extrinsic parameters. This paper proposes a novel camera array self-calibration method based on a near-infrared laser lamp mounted on a UAV. In this method, the matching point pairs are obtained by extracting the laser lamp from the image to determine its position, and then the extrinsic parameters are calculated. To accurately extract the laser lamp on the image, we first eliminate the interference of other light sources in the scene by adjusting the exposure settings of the camera. The imaging result of the near-infrared laser lamp is shown in Fig. 5. Fig. 5(a) shows the imaging result under auto-exposure settings, where the red solid box represents the laser lamp and some interfering light sources are shown by the yellow dashed circles. Fig. 5(b) shows the imaging result under manual exposure settings; the interfering light sources are well filtered out.

The extraction of the light point is crucial, and its accuracy directly affects the accuracy of the calibration parameters. The imaging result of the near-infrared lamp is shown
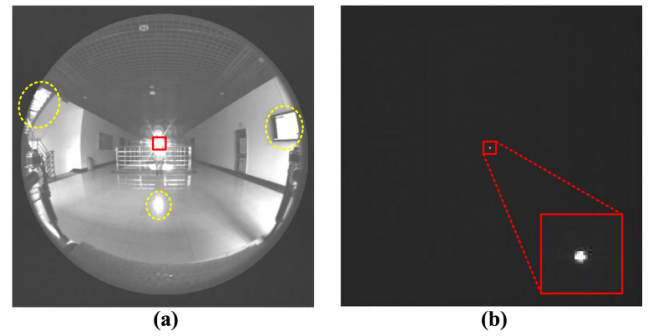

(a) **(b)**

FIGURE 5. Imaging result of the near-infrared laser lamp under the near-infrared camera. (a) Imaging result under the auto-exposure settings; the red solid box represents the laser lamp, and some interfering light source are shown by the yellow dashed circles. (b) Imaging result under the manual exposure settings; the interfering light sources are well filtered out.
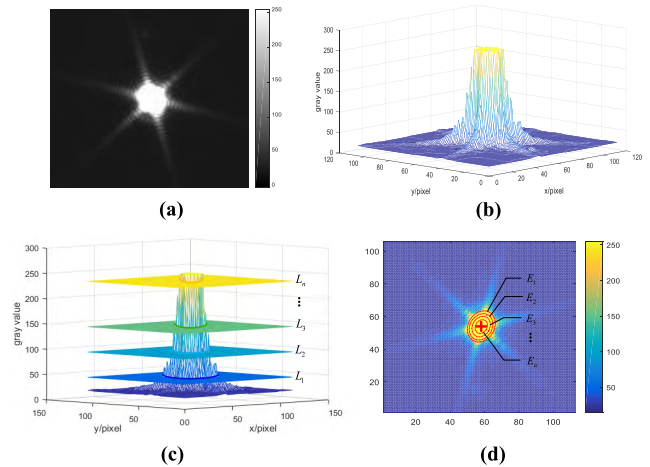

(a) **(b)**

(c) **(d)**

FIGURE 6. Energy distribution of the light point and the extraction method of the light point center. (a) Imaging result of the near-infrared laser lamp. (b) Light point energy distribution. (c) Different energy threshold planes. (d) Multi-layer ellipse fitting results and the final center of the light point.

in Fig. 6(a), which is more concentrated and more completely represented. The gray energy distribution of the light point in 3D is similar to a cone shape (shown in Fig. 6(b)) whose cross section is an approximate ellipse. Considering the characteristics of the laser point and the real-time requirements of the algorithm, we do not use surface fitting methods because of their complex computations, instead using a multi-layer ellipse fitting method to locate the light point center accurately. The idea of this method is to use a set of threshold planes to intercept the gray energy distribution of the light point, to then fit the contour points of each level to obtain the ellipse center, and finally to calculate the mean value to obtain the final light point center. The complete extraction process consists of several steps. First, we segment the light point regions by pixel level on the original image using a gray-scale threshold. Second, we choose a set of threshold planes $L_i (i = 1, 2, \cdots, n)$ to intercept the gray energy distribution of the light point shown in Fig. 6(c). A group of contour

points $P_j(x_j, y_j), j = 1, 2, \cdots, m$, is satisfied by the elliptic function:

$$f(x_i, y_i) = ax_i^2 + bx_iy_i + cy_i^2 + dx_i + ey_i + 1 \qquad (5)$$

and the elliptic parameters (a, b, c, d, e) are obtained by the least squares method. The ellipses fitted by different threshold planes are $E_i(i = 1, 2, \cdots, n)$ as shown in Fig. 6(d). Third, we calculate the center $E(x_e, y_e)$ of each ellipse layer by the elliptic equation

$$x_e = \frac{be - 2cd}{4ac - b^2}, \qquad y_e = \frac{bd - 2ae}{4ac - b^2} \qquad (6)$$

and take its mean as the final center of the light point. In the experiment, we take five different energy planes ($n = 5$) for ellipse fitting.

3) CAMERA ARRAY EXTRINSIC PARAMETER SELF-CALIBRATION

The self-calibration process does not depend on any device and information, using only a near-infrared laser lamp mounted on a UAV. The detailed algorithm flow is shown in Algorithm 1, which includes several steps. First, we can obtain the center position of the light points in each camera on the basis of the extraction of the light point, and then we correct the image coordinates of the light points using the distortion parameters synchronously, which ensures that the light points between cameras are a pair of corresponding points. Second, we obtain a number of pairs of light points between cameras through the UAV flight, which is shown in Fig. 7. $T1, T2, ..., Tn$ (Fig. 7) represent the different imaging results of the laser lamp at different times, and the lower half of the Fig. 7 shows the matching results, where the red points are the central positions of light points and the blue

FIGURE 7. Light point pair generation process and the matching results between cameras 1 and camera 2. *T1, T2, . . . , Tn* represent the different imaging results of the laser lamp at different times. The red points are the central position of light points, and the blue lines are the matching relationship among cameras.

Algorithm 1 Camera Array Self-calibration Algorithm.

Require: n cameras of the structure-free fisheye camera array; We define camera 1 as the reference camera; Offline measurements of the camera baseline between the camera $i(2 \leq i \leq n)$ and the reference camera, D_i; Number of corresponding light points between cameras, N_{points};

Ensure: Extrinsic parameters of all cameras, $T_n[R|t]$;
1: Synchronized original image of each camera at time t_j, I_i^j;
2: **while** ($j \leq N_{points}$) **do**
3:      **while** ($i \leq n$) **do**
4:         Extract the image coordinate of the light point center in camera i;
5:         Correct the image coordinate of the light point center by the double distortion correction, x_i^j;
6:      **end while**
7: **end while**
8: Obtain the 2D-2D corresponding point sets between camera $i(2 \leq i \leq n)$ and the reference camera;
9: Estimate the essential matrix E_i between camera $i(2 \leq i \leq n)$ and the reference camera by 2D-2D corresponding points using the RANSAC method;
10: Calculate the relative transform T_i of camera $i(2 \leq i \leq n)$ from E_i by SVD, and recover the scale by the D_i;
11: Triangulate the N_{inner} 3D coordinates of the light points X_{vision};
12: Optimizing the reprojection error between all cameras by bundle adjustment using the LM algorithm, as described in formula 7;
13: **return** T_n;

---

lines are the matching relationship between cameras. Finally, we estimate the extrinsic parameters between cameras by 2D-2D corresponding point sets using the RANSAC method, and the 3D coordinates of the light points are triangulated. We further optimize the final extrinsic parameters by minimizing the following function using bundle adjustment:

$$min \sum_{i=1}^{n} \sum_{j=1}^{N_{inner}} d(x_i^j, P_i X_{vision}^j)^2 \qquad (7)$$

where the projection matrix of each camera is calculated by $P_i = K_i \cdot T_i$ and where K_i and T_i are the intrinsic and extrinsic parameters of each camera, respectively.

III. PANORAMIC UAV SURVEILLANCE AND AUTONOMOUS RECYCLING SYSTEM

This section elaborates the detailed algorithm of our proposed system for UAV surveillance and autonomous recycling based on a structure-free fisheye camera array. The algorithm mainly consists of several modules. First, we realize the camera array collaborative multi-UAV detection and obtain the UAV correspondence between cameras using epipolar geometry constraints. Next, we measure the 3D coordinates
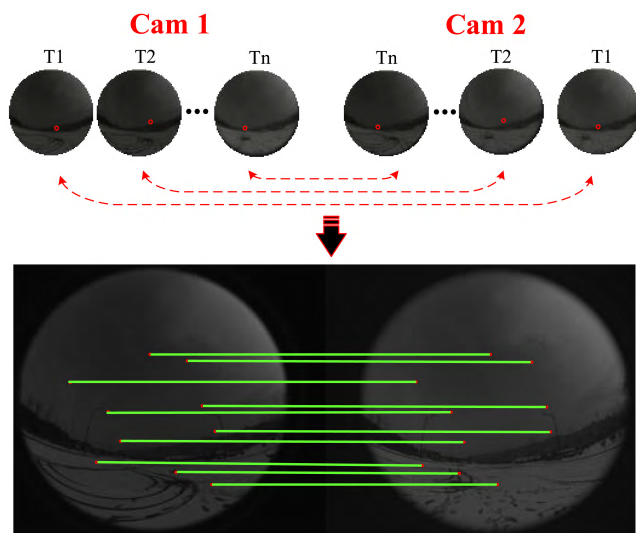
of UAVs using the projection matrices of cameras and achieve the data association by the historical tracking data. Finally, we convert the 3D measurement data of the UAV to control commands and conduct feedback autonomous recycling control using the wireless data link.

A. CAMERA ARRAY COLLABORATIVE MULTI-UAV DETECTION

This section mainly introduces multi-UAV detection based on fisheye camera array. As we mentioned above, the fisheye camera array are installed on the tripod with a static and upward camera view. To obtain suitable foreground targets (UAV), we use the ViBe [35] foreground extraction method in this paper, which performs robustly and produces complete segmentation results as shown in Fig. 8. The algorithm can extract multiple UAVs completely under different sizes, particularly if those are weak and small targets. We cluster the foreground areas and regard the clustering centers as the image coordinate of the candidate targets in the image. The pixel distance between clustering centers is defined as

$$f_{pd}(p_i, p_j) = \sqrt{(p_i^x - p_j^x)^2 + (p_i^y - p_j^y)^2} \quad (8)$$

where p_i and p_j are image pixels and where (p_i^x, p_i^y) and (p_j^x, p_j^y) are the image coordinates of p_i and p_j respectively.
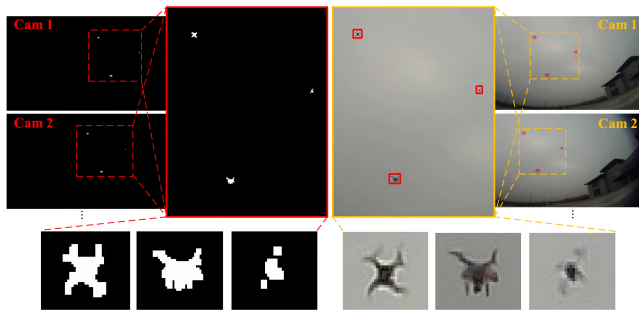

FIGURE 8. Camera array collaborative multi-UAV detection results based on the Vibe [35] foreground segmentation algorithm and epipolar constrain.

On the basis of obtaining the candidate targets in each camera, to determine the corresponding relationship of the candidate targets between the cameras, epipolar geometry constraints between the cameras are used. Epipolar geometry refers to the inherent projective geometry between the camera views; it depends only on the camera intrinsic parameters and the relative pose of the cameras. Thus, epipolar geometry constraints can be used to obtain data association in candidate targets between the cameras. In this way, the corresponding relationships of the candidate targets are confirmed, and portions of false targets are effectively removed.

We use two cameras as an example. Define $I_1 = \{x_1^1, x_2^1, ..., x_m^1\}$ and $I_2 = \{x_1^2, x_2^2, ..., x_n^2\}$ as the detection results of the first and second camera, respectively. The goal of data association is to find the corresponding relationship between x_i^1 and x_j^2. The distance measurement is obtained by

the symmetric transfer error between $x_i^1 (i = 1, 2, ..., m)$ and $x_j^2 (j = 1, 2, ..., n)$, which can be defined as

$$d(x_i^1, x_j^2) = d(x_i^1, F^T x_j^2) + d(x_j^2, F x_i^1) \quad (9)$$

where F is the fundamental matrix between the two cameras. The matching matrix between two images is

$$D = \begin{bmatrix} d(x_1^1, x_1^2) & d(x_1^1, x_2^2) & \cdots & d(x_1^1, x_n^2) \\ d(x_2^1, x_1^2) & d(x_2^1, x_2^2) & \cdots & d(x_2^1, x_n^2) \\ \vdots & \vdots & \ddots & \vdots \\ d(x_m^1, x_1^2) & d(x_m^1, x_2^2) & \cdots & d(x_m^1, x_n^2) \end{bmatrix} \quad (10)$$

The global optimal matching result is obtained by solving the matching matrix D using the Hungarian algorithm, which is taken as the final detection results.

B. 3D LOCALIZATION AND TRACKING

In Section II.B, we obtain the fisheye camera parameters, including the intrinsic parameters, distortion coefficient and extrinsic parameters of each camera, which are used to correct the image pixels and calculate the camera projection matrix. In this section, we explain the principles of the 3D localization and tracking methods.

1) 3D LOCALIZATION

The UAV 3D localization method based on a fisheye camera array is shown in Fig. 9. Let the number of the cameras and the UAV of the proposed system be $N (N \geq 2)$ and M, respectively. The process of the UAV 3D localization includes several steps. First, we correct the image coordinates of the UAV in all cameras. Second, we calculate the initial value of the 3D coordinates of the UAV by the projection matrices and the corrected 2D corresponding image coordinates of the UAV. Finally, we conduct further optimization using
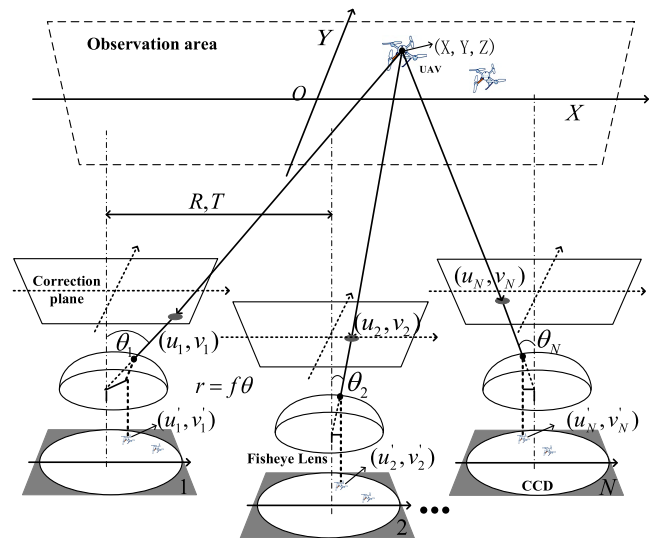

FIGURE 9. Architecture of the proposed fisheye camera array imaging system and 3D localization method.

all cameras by bundle adjustment. For clear explanation, we introduce the detailed 3D localization principle with two camera models ($N = 2$) and a UAV ($M = 1$) as an example, with a designated camera 1 and camera 2; the computing process for multiple targets is the same. We define the camera coordinate system of camera 1 as the world coordinate system, and the projection matrices of the two cameras are P_1 and P_2. The world coordinates of the UAV is (X, Y, Z), and its homogeneous form are defined as $X_{uav} = (X, Y, Z, 1)$; the original image coordinates of the UAV in this two cameras are (u'_1, v'_1) and (u'_2, v'_2), and the corrected values are (u_1, v_1) and (u_2, v_2) with homogeneous form defined as $x_1 = (u_1, v_1, 1)$ and $x_1 = (u_1, v_1, 1)$, respectively. Due to the measurement error, the projection function and epipolar geometry constraint cannot be fully satisfied. To obtain the precise initial values of x_1 and x_2, the projective invariant binocular location method is used to minimize the reprojection error by finding the minimum epipolar geometry constraint and reprojection error.

After obtaining the initial value of x_1 and x_2, we assume values for $x_1 \cong P_1 X_{uav}$ and $x_2 \cong P_2 X_{uav}$ with the corresponding homogeneous relations $x_1 \times P_1 X_{uav} = 0$ and $x_2 \times P_2 X_{uav} = 0$. These are linear equations with respect to X_{uav}, which can be written as $A X_{uav} = 0$. Each set of points corresponds to the three equations, of which only two are linearly independent; therefore, each point gives only two equations correspondingly with respect to X_{uav}, and the matrix A can be described as

$$A = \begin{bmatrix} u_1 p_1^{3T} - p_1^{1T} \\ v_1 p_1^{3T} - p_1^{2T} \\ u_2 p_2^{3T} - p_2^{1T} \\ v_2 p_2^{3T} - p_2^{2T} \end{bmatrix} \quad (11)$$

where p_1^{iT} and p_2^{iT} are the i-th row of matrix P_1 and P_2, respectively.

X_{uav} is the homogeneous coordinates, of which only three degrees of freedom are scale independent. The linear equation set $A X_{uav} = 0$ contains four equations; thus, this system is over-determined. To find the approximate solution of the equation $A X_{uav} = 0$, we can transform the problem into the following optimization problem:

$$\min_X ||A X_{uav}|| \quad (12)$$

subject to $\| X_{uav} \| = 1$. Then, we calculate the initial value of X_{uav} using the DLT algorithm.

Primarily, we obtain the initial value of X_{uav} and then optimize the value of X_{uav} using the method of minimizing the reprojection error between multiple cameras by bundle adjustment. We obtain the final X_{uav} by minimizing the following function using the LM iterative optimization algorithm.

$$\min \sum_{i=1}^{n} v_i d(x_i, P(X_{uav}, I_i))^2 \quad (13)$$

where I_i is the i camera frame. If there are imaging points on the I_i image, we set $v_i = 1$; otherwise, we set $v_i = 0$. $P(X_{uav}, I_i)$ is the projection on the i camera frame, and $d(x_i, P(X_{uav}, I_i))^2$ is the reprojection errors computed using the corresponding projection matrices.

2) MULTIPLE UAV TRACKING

Suppose that $T^t = \{T_1^t, T_2^t, ..., T_p^t\}$ is the UAV tracking result at time t and $O^{t+1} = \{X_1^{t+1}, X_2^{t+1}, ..., X_q^{t+1}\}$ is the 3D localization result at time $t + 1$. To track the targets on consecutive frames between times t and $t + 1$, we measure the similarity of targets by associating the history tracking result $T_i^t (i = 1, 2, ..., p)$ and the current 3D localization result $X_j^{t+1} (j = 1, 2, ..., q)$. First, the Euclidean distance is used as the distance measurement in the 3D space between the tracking result T_i^t and current localization result X_j^{t+1}, which defines $d(T_i^t, X_j^{t+1})$. Furthermore, the matching matrix D_t^{t+1} between the historical tracking results and current observations is computed as follows:

$$D_t^{t+1} = \begin{bmatrix} d(T_1^t, X_1^{t+1}) & d(T_1^t, X_2^{t+1}) & \cdots & d(T_1^t, X_q^{t+1}) \\ d(T_2^t, X_1^{t+1}) & d(T_2^t, X_2^{t+1}) & \cdots & d(T_2^t, X_q^{t+1}) \\ \vdots & \vdots & \ddots & \vdots \\ d(T_p^t, X_1^{t+1}) & d(T_p^t, X_2^{t+1}) & \cdots & d(T_p^t, X_q^{t+1}) \end{bmatrix} \quad (14)$$

In the last expression, the Hungarian algorithm is used to obtain the UAV tracking results from D_t^{t+1}. The Hungarian algorithm, a maximum matching algorithm based on the graph model, is widely used in data association problems.

The detailed multiple UAV detection, 3D localization and tracking algorithm flow is shown in Algorithm 2.

C. UAV AUTONOMOUS RECYCLING METHOD

In this section, we mainly discuss UAV recycling methods, including the transformation between the world frame (X_w, Y_w, Z_w) and the UAV body frame (X_b, Y_b, Z_b) and the UAV control strategy. Note that the whole control is based on the ground processing computer unit by sending control signals to the UAV to achieve the recycling task. The Fig. 10 shows the whole recycling process.

It is necessary to obtain the direction of motion of the UAV in the process of controlling a UAV landing, while our localization system can obtain only the real-time location. To solve the UAV direction problem, we determine the transformation between the world frame and the UAV body frame by referring to the geomagnetic north direction. As shown in Fig. 10, we set the world frame as $(O - X_w, Y_w, Z_w)$ and the UAV body frame as $(O - X_b, Y_b, Z_b)$. First, we record the northward angle $\theta_{North}^{Y_w}$ of the Y-axis for the world coordinate system. Second, we obtain the northward angle $\theta_{North}^{X_b}$ of the X-axis for the UAV body coordinate system by the magnetometer of the UAV itself. Third, to simplify the recycling, we control the UAV to rotate θ_{Yaw} so that the X-axis of the UAV body coordinate system is parallel to the negative direction of the Y

Algorithm 2 Multiple UAV Detection, 3D Localization and Tracking Algorithm.

Require: Synchronized camera image data and the projection matrices of all cameras

Ensure: Real-time 3D coordinates of the UAV and its consecutive tracking results

1: Extract the regions of the UAV from the background of each camera using the ViBe algorithm;

2: Cluster the regions in each camera, and regard the clustering centers as the candidate targets;

3: Associate the targets between cameras using the epipolar geometry constraint, and eliminate the false targets simultaneously, as described in formula 10;

4: Correct the image coordinates of the targets in each camera by the fisheye distortion model;

5: Calculate the initial 3D coordinates of the UAV by the DLT algorithm, as described in formula 11 and formula 12;

6: Optimize the 3D coordinates of the UAV by BA using the LM iterative optimization algorithm, as described in formula 13;

7: Track the multiple UAVs in the current frame using the Hungarian algorithm, as described in formula 14;

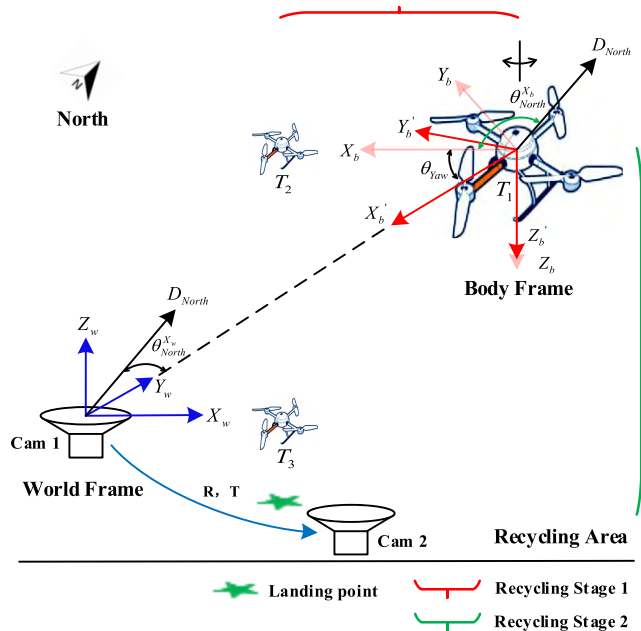8: **return** 3D coordinates of the UAV and tracking results.

FIGURE 10. UAV recycling method. The whole recycling process includes the transformation from the world frame (X_W, Y_W, Z_W) to the UAV body frame (X_b, Y_b, Z_b) and a two-stage UAV control strategy. In the system, we set the reference camera frame as the world frame.

axis of the world coordinate system, where θ_{Yaw} is calculated:

$$\theta_{Yaw} = 180° - \theta_{North}^{X_b} - \theta_{North}^{X_w} \tag{15}$$

and the new body frame is $(O - X_b', Y_b', Z_b')$ after rotation.

For the control strategy, we adopt a two-stage operation based on proportional velocity control, and we

set the landing point as $L(x_w, y_w, z_w)$, the initial UAV position as $I_{uav}(x_w, y_w, z_w)$ and current UAV position as $C_{uav}(x_w, y_w, z_w)$,. In the first stage ($T_1 \sim T_2$), keeping the altitude unchanged, the UAV is controlled to reach the upper part of the landing point. The velocity v_{uav} is calculated by the proportion of the position:

$$v_{uav} = (1 - D_2/D_1) * V \tag{16}$$

where D_1 and D_2 are

$$D_1 = \sqrt{(C_{uav}(x_w) - I_{uav}(x_w))^2 * (C_{uav}(y_w) - I_{uav}(y_w))^2}$$

$$D_2 = \sqrt{(I_{uav}(x_w) - L(x_w))^2 * (I_{uav}(y_w) - L(y_w))^2} \tag{17}$$

and V is the artificial setting value determined by the UAV type and scene size. We use the calculated v_{uav} to orthogonally decompose the velocity into X and Y directions of the UAV body frame, which are v_{uav}^x and v_{uav}^y, respectively. In the second stage ($T_2 \sim T_3$), a uniform speed control model is used to control the UAV to descend smoothly. When the UAV reaches the boundary of the camera monitoring blind area, we start the landing command directly to complete the recycling task.

IV. EXPERIMENTAL RESULTS

To evaluate the performance of the proposed panoramic UAV surveillance and recycling system, we conducted a series of field experiments. There are three main parts of the experiments, including the camera array self-calibration (Section IV.A); multi-UAV detection, 3D localization and tracking (Section IV.B); and UAV autonomous recycling (Section IV.C). There are four different experimental scenarios, including indoor and outdoor scenes. Moreover, the UAV used in the test has a variety of sizes and appearances. For example, DJI Matrix 100 is used in the self-calibration experiment. In multi-UAV detection, 3D localization and tracking experiment, four kinds of UAVs, such as the DJI Matrix 100 and DJI Phantom 2, are used. UAV autonomous recycling has security risks, so we choose the small UAV Parrot Beobop for the recycling experiment. In addition, to test the time performance of the system, we use two different computing platforms laptop OMEN (HP Laptop 15-dc0124TX) and NVIDIA Jetson TX2 to implement our system; the computation analysis of each module is shown in Section IV.D.

A. CAMERA ARRAY SELF-CALIBRATION

1) DYNAMIC LASER SOURCE FOR SELF-CALIBRATION

In our system, we propose a dynamic laser-source-based self-calibration method to obtain the extrinsic parameters of the camera array. The self-calibration algorithm details are explained in the Section II.B. To prove the validity and accuracy of this self-calibration method, we choose the DJI Matrix 100 to carry the laser lamp as a dynamic light source, as shown in Fig. 11. Note that the manifold and wireless datalink module carried by UAVs are used to acquire and transmit GPS information. These GPS data do not participate in the
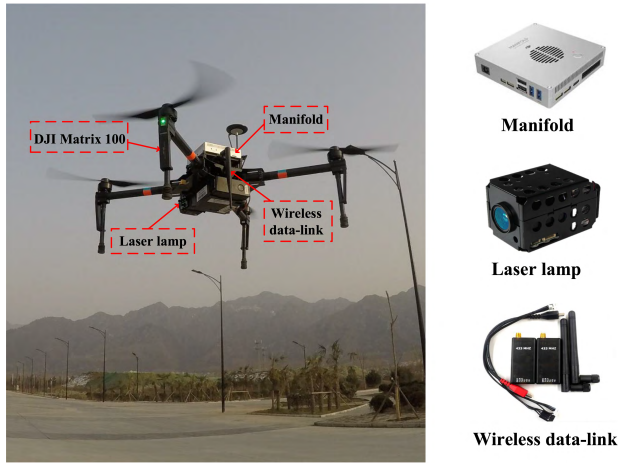
FIGURE 11. Complete UAV configuration for fisheye camera array self-calibration, including DJI Matrix 100, the near-infrared laser lamp, the on-board computer DJI Manifold and the wireless data-link module.

self-calibration process but provides only the true value for the experiment, which is used to compare with the system measurement data.

In the experiment, all the flight experiment data collection is performed via manual flight or remote control with the aid of our GPS waypoint navigation system. To determine the effective range of the dynamic light source in the calibration process, we test the imaging effect of the near-infrared laser lamp at different distances and angles. Because there may be other light sources in the scene, we use filters with the same band, while reducing the exposure time and brightness of the cameras to avoid interference from other light sources

in the scene. In testing, we find that the fisheye camera can effectively extract the light spots within 200 meters and 150 degrees FOV, as shown in Fig. 12. Therefore, to ensure the robustness of light spots detection, the UAV flying altitudes are less than 50 meters in the whole calibration process.

2) CAMERA ARRAY SELF-CALIBRATION RESULTS

We conducted a large number of real experiments to verify the proposed self-calibration method for the different structures of the fisheye camera array. In the experiment, most of the misclassified and outlier points are removed by the epipolar constraint, and to obtain a better calibration result, the UAV flies at different heights and distances to cover the whole camera FOV as much as possible. Fig. 13(a) shows the layout of the ground camera array and the matching result of the laser light points between cameras; the red points represent the imaging position of the laser lights, and the green lines are the matching result among cameras. The camera array self-calibration result of relative poses between two independent cameras is shown in Fig. 13(b). The calibration error analysis of 3D coordinates of light points between vision and GPS data is shown in Fig. 13(c), which reveals that the error is within 1 meter. The whole calibration accuracy meets our localization requirements.

B. MULTIPLE-UAV DETECTION, 3D LOCALIZATION AND TRACKING

This section mainly studies the performance of the multi-UAV detection, 3D localization and tracking algorithm in our system. To verify the stability and validity of these
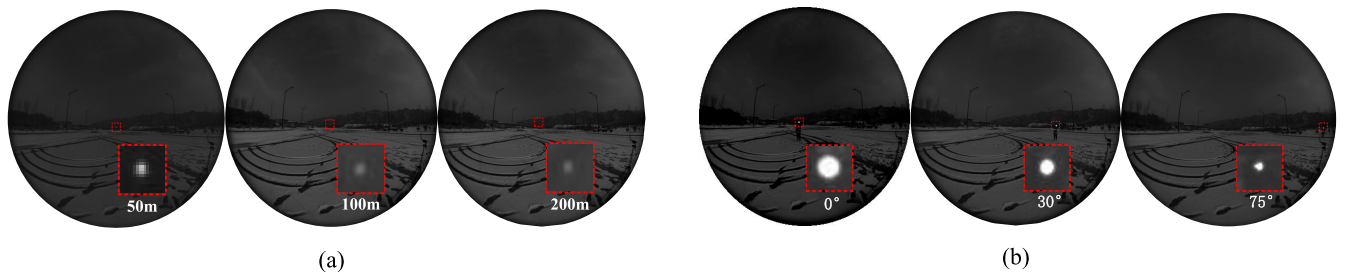

(a)

(b)

FIGURE 12. Laser light imaging. (a) and (b) compare the imaging at different distances and different angles, respectively.
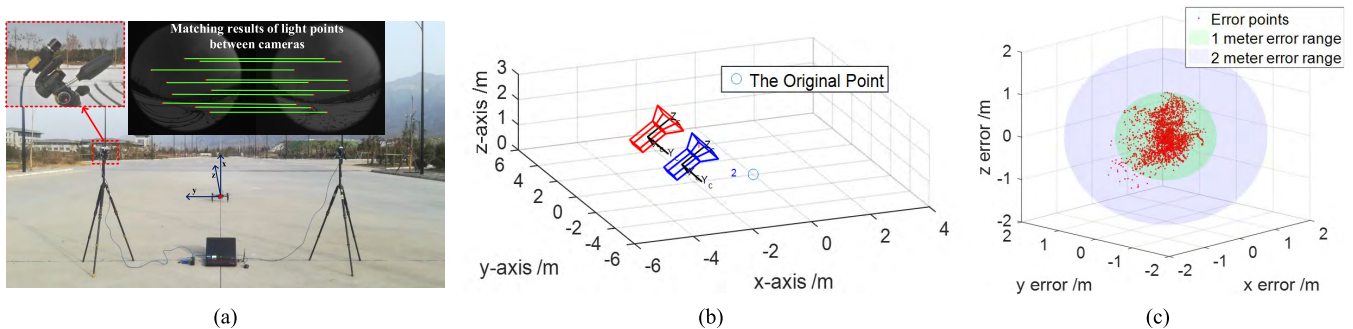

(a)

(b)

(c)

FIGURE 13. Camera array self-calibration result. (a) Layout of the ground camera array. (b) Camera array self-calibration result. (c) Calibration accuracy comparison of vision and GPS data.

TABLE 2. Four types of multiple-rotor aircraft.

UAV name	Eight-rotor	DJI Matrix 100	DJI Phantom 3	DJI Phantom 2
Wheelbase (mm)	1133	650	350	350
Weight (kg)	5.30	2.30	1.28	1.00
Max. speed (km/h)	65	20	20	15

FIGURE 14. Experimental UAVs with different appearances and sizes, including DJI Phantom 2, DJI Phantom 3, DJI Matrix 100 and eight-rotor UAV.

algorithms, we selected a number of UAVs with significant differences in appearance and size as targets and carried out experiments with a variety of camera array structures and in five different environments (such as a sunny or cloudy day and an indoor environment).

The experimental UAVs and their detailed parameter configurations are shown in Fig. 14 and Table 2, respectively. There are four different UAVs. One is an eight-rotor drone, the wheelbase of which is approximately 1133 mm with a 5.30 kg weight with two standard batteries. The maximum speed is 65 km/h, achieved with the DUAL SKY XM5015TE-5 motor. There are also three quad-rotor vehicles: a DJI Matrix 100, the wheelbase of which is approximately 650 mm with a 2.30 kg weight with one standard battery, a DJI Phantom professional 3 with a 350 mm wheelbase and 1.28 kg weight, and a DJI Phantom 2 with a 350 mm wheelbase and 1.00 kg weight.

1) MULTIPLE UAV DETECTION

The motor propeller movement is greater than the movement of the body when the vehicles are undergoing hover or low-speed flight in the experiments. The original ViBe algorithm causes over-segmentation and produces too many small pieces for an independent target. To solve the over-segmentation problem, we perform morphological processing and carry out further regional merging, which produces satisfactory detection performance. Selected detection results of UAVs are shown in Fig. 15. The first three columns are photographed using the USB camera and contains three scenes: a sunny, cloudy and rainy day. The last two columns are photographed by the PointGrey camera and include snowy and indoor scenes. The experimental results show that the target can be effectively detected, particularly for small, low-speed and scale-changing targets under some challenging conditions, such as Frame 1_#1897 and 2_#2214. Incorrect detection events may also occur. For example, in Frame 1_#305 and 2_#303 in Fig. 15, there are some false detection results in the lower left corner of the image. To improve detection performance, we mainly remove false targets from three aspects. In the process of multi-camera collaborative detection based on the epipolar constraint, the false targets can be
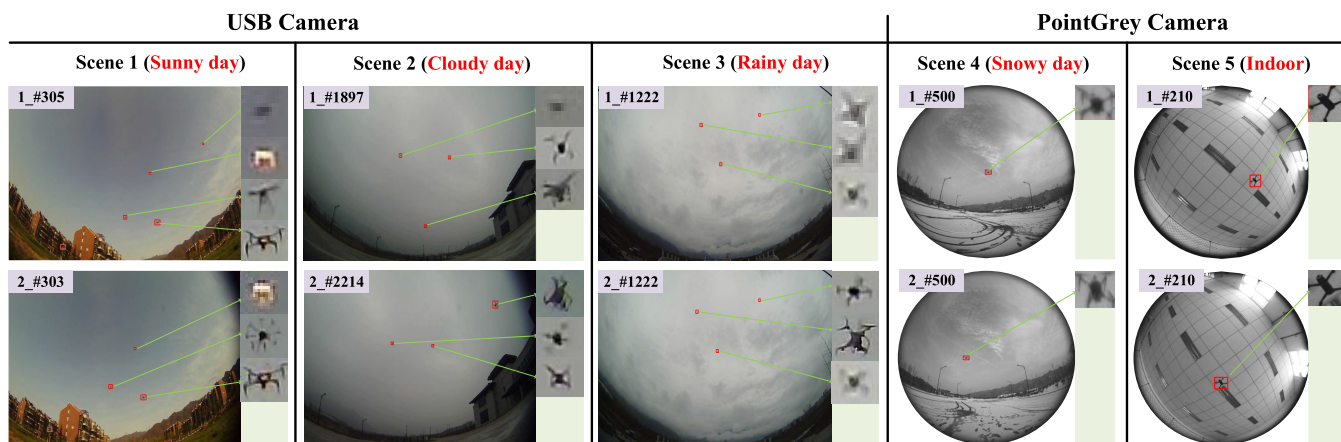

FIGURE 15. Detection results of the multiple UAVs. The first three columns are photographed using a USB camera, which contains three scenes, including a sunny, cloudy and rainy day. The last two columns are photographed by the PointGrey camera and show a snowy day and an indoor scene.

removed by the symmetric transfer error. In the process of the camera array vision 3D localization, the false targets can be removed by the space motion constraints of the UAVs. In the process of target tracking, the false targets can be removed by analyzing the motion directions and velocities of the candidate targets. In this way, the targets can be detected correctly. In addition, intermittently detected targets can also be identified by the following tracking algorithm. As shown in Fig. 15, the results of the detection bounding boxes are complete and accurate in various challenging scenarios, which effectively illustrates the robustness of the proposed method.

2) 3D LOCALIZATION AND TRACKING

On the basis of accurate UAV detection and correlation between cameras, we further accomplish UAV synergetic 3D localization and tracking based on a camera array. In the experiment, we take a UAV (DJI Matrix 100) as an example to analyze the accuracy of 3D localization results. We know

that the aircraft supports secondary development and can transmit real-time GPS data to the ground computer through the wireless data link, which helps us use the GPS data as ground truth to make a comparison with the calculated 3D localization. When making comparisons, the integration of coordinate systems between the GPS data and vision localization data is crucial, which includes several steps. First, we set the take-off point of the UAV as the origin of the GPS coordinate system and record the GPS data as G_0. Second, we calculate the position of UAV in GPS coordinates at each time by $P_{gps}^{j} = F(G_j, G_0)$ simultaneously, where G_j is the GPS data of the UAV and the function F is based mainly on the Earth's radius. Then, we obtain all the position of the UAV flight trajectory. Finally, we estimate the transformation by the 3D points of the vision and GPS data and integrate them by the transformation.

Fig. 16(a) shows a real experimental environment of a snowy day; we synchronously save the visual 3D localization
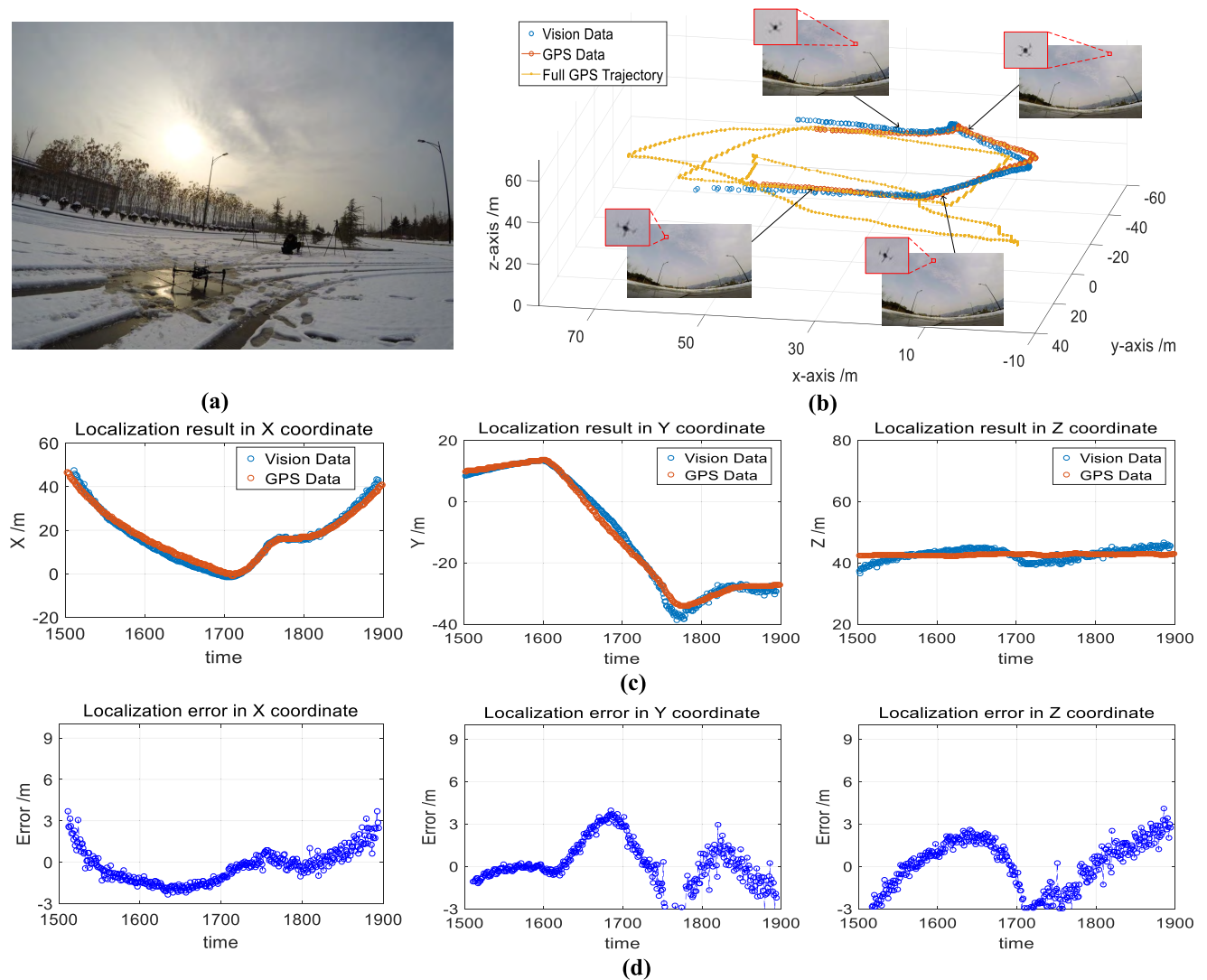

FIGURE 16. UAV 3D vision localization results and the accuracy analysis. (a) Real experimental environment. (b) Comparison of the vision localization data and GPS data of the UAV trajectory. (c) Comparison of vision localization results in the X, Y and Z directions with GPS data. (d) Vision localization errors in the X, Y and Z directions.
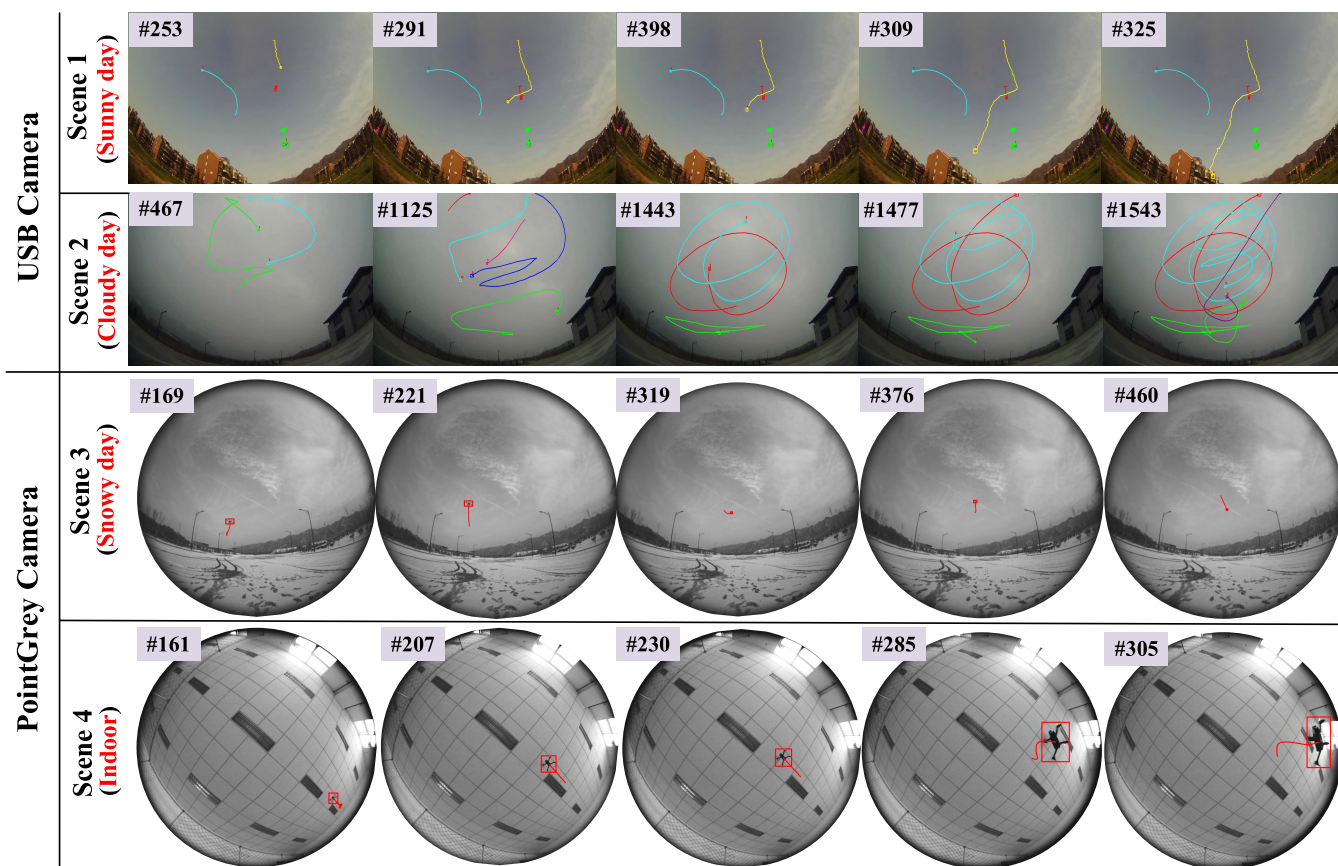
FIGURE 17. Tracking results of multiple UAVs, showing the experimental results of different scenarios using different sensors. The first two rows correspond to the USB fisheye camera and several UAVs, including several complete trajectories. The last two rows correspond to the PointGrey camera and include the outdoor and indoor experiment.

data and the real-time GPS data in the experiments. Fig. 16(b) shows the UAV fight 3D trajectory; clearly, the trajectory results of the visual localization coincide well with the GPS data. In addition, the accuracy comparison and error analysis on the X-, Y- and Z-axes are shown in Fig. 16(c) and Fig. 16(d), respectively. The maximum errors in the X-, Y- and Z-axes are less than 3 meters, and the vision localization errors gradually decrease as the UAV approaches. For the structure of the fisheye lens, the experimental results indicate that the 3D localization accuracy near the center of the camera is clearly higher than that of the surrounding area and that the localization error decreases gradually from far to near for the constraints of the baseline between the cameras. The multiple-UAV tracking experimental results are shown in Fig. 17, which includes various scenarios using different UAVs. Some false targets are effectively excluded by the tracking module. The experimental results show that our method can effectively track multiple UAVs in various scenarios, even in the challenging case of targets crossing.

C. BEOBOP AUTONOMOUS RECYCLING EXPERIMENTS

In this section, we conduct a real indoor UAV autonomous recycling experiment to further verify the performance of the proposed system. The indoor experimental environment is

a 10.0 m × 8.0 m × 3.2 m space, surrounded by a safety net, as shown in Fig. 18(a). The system settings are shown in Fig. 18(b), which mainly includes the Parrot Beobop as the UAV, a ground static vehicle as the recycling platform, a double fisheye camera array and NNIDIA Jetson TX2 as the computing platform. The Beobop has the characteristics of small size and light weight, so its flight safety is quite high. In addition, it communicates through WiFi, and the maximum communication distance is as high as 100 m. The recycling platform is based a ground static vehicle (height 0.31 m), which includes a 1.0 m × 1.0 m PVC landing pad. In the experiment, the algorithm runs with the NVIDIA Jetson TX2, which sends the control instructions to the UAV through WiFi in real time.

In the experiment, we set the initial height of the UAV to 3.0 m above the ground, and then the autonomous recycling system is implemented. There are two landing stages, T_{stage1} and T_{stage2}, in the whole recycling process as shown in Fig. 19(a). The T_{stage1} is mainly based on a proportional velocity control model to fly above the landing point. When the first stage is completed, the second stage is started, and the UAV is controlled by a uniform velocity model to land. The landing trajectory and the flying states at different times are shown in Fig. 19(a). Fig. 19(b) shows the velocity changes in the X, Y and Z directions of the UAV. In T_{stage1} (before t_1),
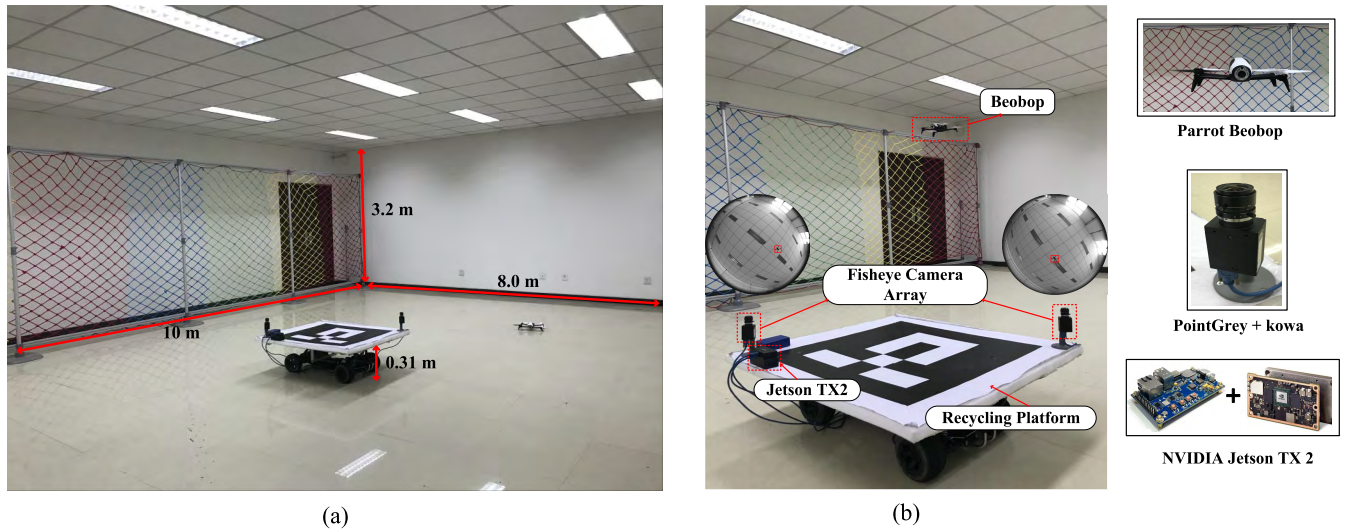
FIGURE 18. Indoor UAV autonomous recycling experiment used the proposed system. (a) Indoor experimental environment. (b) Complete UAV recycling experiment settings, including the Parrot Beobop, a double fisheye camera array, NVIDIA Jetson TX2 and static vehicle recycling platform.
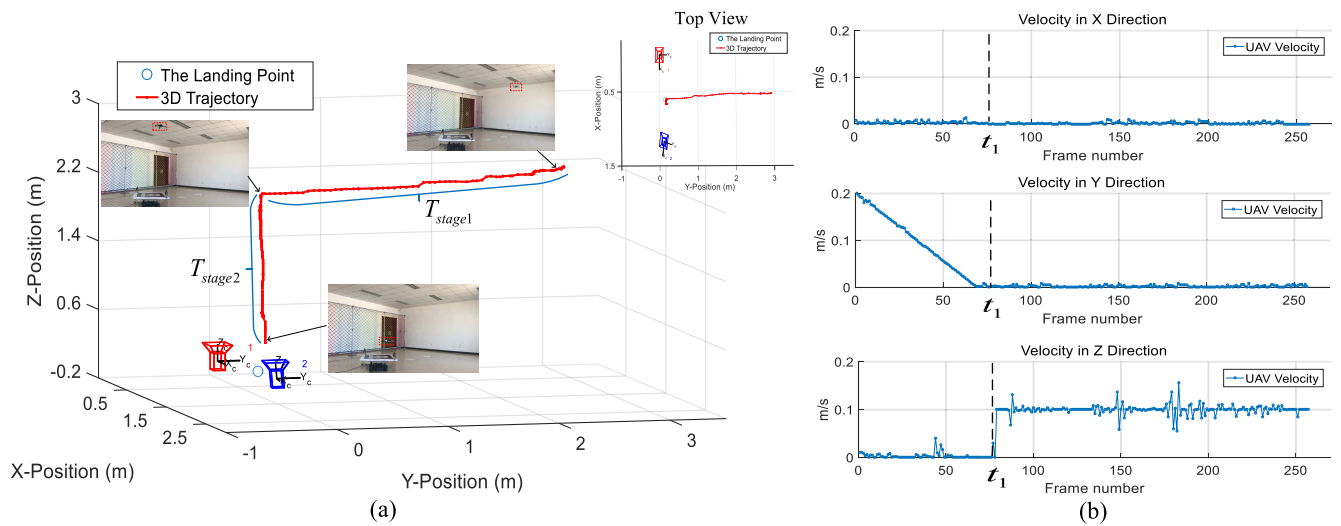

FIGURE 19. Results of the UAV recycling experiment built with Parrot Beobop, including (a) the landing trajectory and the flying states at different times. (b) Velocity changes in the X, Y and Z directions of the UAV.

TABLE 3. Computational analysis of the system with HP OMEN and NVIDIA Jetson TX2.

Content	HP OMEN		NVIDIA Jetson TX2		
	Mean	**Standard Deviation**	**Mean**	**Standard Deviation**	
Synchronous Image Acquisition	3.52	0.53	13.65	0.68	(ms)
Camera Array Self-calibration (**offline**)	3.56	0.74	8.34	0.84	(min)
Multi-UAV Detection	48.60	1.28	102.36	1.32	(ms)
3D Localization and Tracking	3.52	0.08	6.53	0.80	(ms)
Autonomous Recycling Control	6.68	0.06	11.68	0.11	(ms)
Total (Max. thread)	55.35	0.05	115.65	0.08	(ms)

the velocity of Y direction decreases continuously until it reaches zero as the UAV approaches, and the height of UAV is constant, so the velocity in the Z direction is almost zero.

In T_{stage2} (after t_1), the velocity is only in the Z direction and is maintained at approximately 0.1 m/s. Since the initial UAV is positioned directly in front of the recycling platform,

the velocity in the X direction is always zero. Note that the recycling system is GPS-denied and cannot obtain the ground truth for performance comparison.

D. COMPUTATIONAL ANALYSIS OF THE SYSTEM

In this part, we test and analyze the time performance of each module in our system. The system includes the offline camera array self-calibration module and online processing module. In the experiments, the resolution of the processed image is 1024×1024, and we use a high-performance laptop OMEN (HP Laptop 15-dc0124TX) and NVIDIA Jetson TX2 to conduct experiments. The running time of the system is shown in Table 3.

For the offline calibration module, the running time is less than 10 minutes, including the data acquisition process. The online processing module includes synchronous image acquisition, multi-UAV detection, 3D localization and tracking, and autonomous recycling, which is implemented on the ROS platform for parallel processing. We measure the runtime of each ROS node of the system on the laptop and Jetson TX2. As shown in the table, the UAV detection module is time consuming, and other modules are very fast. On average, the system takes approximately $55.35\ ms$ and $115.65\ ms$ in the HP OMEN and NVIDIA Jetson TX2, respectively. In practice, we find that the processing speed is sufficient to achieve effective UAV surveillance and recycling performance.

V. CONCLUSIONS

This article first constructs a novel real-time panoramic UAV surveillance and recycling system based on a structure-free fisheye camera array. Relative to the existing UAV surveillance technologies, the fisheye camera array ensures a large monitoring FOV, and the free system structure is convenient for rapid deployment and flexible configuration. In addition, we design a robust dynamic laser source based camera array self-calibration algorithm for large-scale arbitrary layout of a camera array without any other markers, particularly in environments with a weak texture or solid color background. Finally, we design an intelligent multi-UAV fast detection, 3D localization, tracking and recycling algorithm, which effectively addresses the issues of the scale changes and the interleaving of UAVs during the flight. The system has been tested in various challenging scenarios, including multiple UAVs with significant appearance and scale changes and even poor weather conditions. Extensive experimental results, both qualitative and quantitative, in addition to the time performance, demonstrate the robustness and effectiveness of the proposed system. Furthermore, we successfully conduct the recycling and landing experiment with Parrot Beobop, and the experimental results show that the proposed panoramic UAV surveillance and recycling system can effectively recycle the UAV autonomously.

In the future, we plan to continue studying the UAV autonomous recycling technology on a vehicle platform to achieve collaborative control and recovery tasks. In addition, on the basis of the proposed system, combined with simultaneous localization and mapping (SLAM) technology based on the UAV platform, we will build a fully autonomous UAV-UGV cooperative recycling system.

REFERENCES

[1] P. Boucher, "Domesticating the drone: The demilitarisation of unmanned aircraft for civil markets," *Sci. Eng. Ethics*, vol. 21, no. 6, pp. 1393–1412, 2015.

[2] M. García-Fernández *et al.*, "Antenna diagnostics and characterization using unmanned aerial vehicles," *IEEE Access*, vol. 5, pp. 23563–23575, 2017.

[3] P. Liu *et al.*, "A review of rotorcraft unmanned aerial vehicle (UAV) developments and applications in civil engineering," *Smart Struct. Syst*, vol. 13, no. 6, pp. 1065–1094, 2014.

[4] M. Erdelj, E. Natalizio, K. R. Chowdhury, and I. F. Akyildiz, "Help from the sky: Leveraging UAVs for disaster management," *IEEE Pervasive Comput.*, vol. 16, no. 1, pp. 24–32, Jan./Mar. 2017.

[5] R. Altawy and A. M. Youssef, "Security, privacy, and safety aspects of civilian drones: A survey," *ACM Trans. Cyber-Phys. Syst.*, vol. 1, no. 2, p. 7, 2017.

[6] R. Clarke, "Understanding the drone epidemic," *Comput. Law Secur. Rev.*, vol. 30, no. 3, pp. 230–246, 2014.

[7] M. J. Boyle, "The race for drones," *Orbis*, vol. 59, no. 1, pp. 76–94, 2015.

[8] (2014). *GIRAFFE Radar System*. [Online]. Available: https://en.wikipedia.org/wiki/Giraffe_radar

[9] (2015). *Selex ES launches Falcon Shield Counter-UAV System*. [Online]. Available: http://www.us.selex-es.com/-/falconshield

[10] (2017). *AUDS Anti UAV Defence System*. [Online]. Available: https://www.auds.com/

[11] J. F. Vasconcelos, C. Silvestre, P. Oliveira, and B. J. N. Guerreiro, "Embedded UAV model and laser aiding techniques for inertial navigation systems," *Control Eng. Pract.*, vol. 18, no. 3, pp. 262–278, 2010.

[12] J. Wu, Q. Li, M. He, and F. Zhang, "A terrain model generation method based on 2D plan laser scanner for micro UAV autonomous flight," in *Proc. Int. Conf. Inf. Technol. Softw. Eng.* Berlin, Germany: Springer, 2013, pp. 215–225.

[13] M. Strohmeier, T. Walter, J. Rothe, and S. Montenegro, "Ultra-wideband based pose estimation for small unmanned aerial vehicles," *IEEE Access*, vol. 6, pp. 57526–57535, 2018.

[14] S. Weiss, D. Scaramuzza, and R. Siegwart, "Monocular-SLAM–based navigation for autonomous micro helicopters in GPS-denied environments," *J. Field Robot.*, vol. 28, no. 6, pp. 854–874, 2011.

[15] Q. Fu, Q. Quan, and K.-Y. Cai, "Robust pose estimation for multirotor UAVs using off-board monocular vision," *IEEE Trans. Ind. Electron.*, vol. 64, no. 10, pp. 7942–7951, Oct. 2017.

[16] H. Khosravi and E. Fazl-Ersi, "Trackerbot: A robotic surveillance system based on stereo-vision and artificial neural networks," in *Proc. IEEE Int. Conf. Robot. Mechatronics*, Oct. 2016, pp. 449–454.

[17] D. Stepanov and I. Tishchenko, "The concept of video surveillance system based on the principles of stereo vision," in *Proc. IEEE 18th Conf. Open Innov. Assoc. Seminar Inf. Secur. Protection Inf. Technol. (FRUCT-ISPIT)*, Apr. 2016, pp. 328–334.

[18] S. Lange, N. Sünderhauf, P. Neubert, S. Drews, and P. Protzel, "Autonomous corridor flight of a UAV using a low-cost and lightweight RGB-D camera," in *Advances in Autonomous Mini Robots*. Berlin, Germany: Springer, 2012, pp. 183–192.

[19] A. Bachrach *et al.*, "Estimation, planning, and mapping for autonomous flight using an RGB-D camera in GPS-denied environments," *Int. J. Robot. Res.*, vol. 31, no. 11, pp. 1320–1343, 2012.

[20] K. Berger, R. Voorhies, and L. Matthies, "Incorporating polarization in stereo vision-based 3d perception of non-Lambertian scenes," *Proc. SPIE*, vol. 9837, May 2016, Art. no. 98370P.

[21] Y. Gui *et al.*, "Airborne vision-based navigation method for UAV accuracy landing using infrared lamps," *J. Intell. Robot. Syst.*, vol. 72, no. 2, pp. 197–218, 2013.

[22] C. Martínez, P. Campoy, I. Mondragón, and M. A. Olivares-Méndez, "Trinocular ground system to control UAVs," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Oct. 2009, pp. 3361–3367.

[23] D. Pebrianti, F. Kendoul, S. Azrad, W. Wang, and K. Nonami, "Autonomous hovering and landing of a quad-rotor micro aerial vehicle by means of on ground stereo vision system," *J. Syst. Des. Dyn.*, vol. 4, no. 2, pp. 269–284, 2010.

[24] T. Furukawa, C. Kang, B. Li, and G. Dissanayake, "Multi-stage Bayesian target estimation by uav using fisheye lens camera and pan/tilt camera," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Sep. 2017, pp. 4167–4172.

[25] A. Perez-Yus, G. Lopez-Nicolas, and J. J. Guerrero, "A novel hybrid camera system with depth and fisheye cameras," in *Proc. IAPR Int. Conf. Pattern Recognit. (ICPR)*, 2016, pp. 2789–2794.

[26] W. Gao and S. Shen, "Dual-fisheye omnidirectional stereo," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Sep. 2017, pp. 6715–6722.

[27] A. Gupta, D. Bessonov, and P. Li, "A decision-theoretic approach to detection-based target search with a UAV," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Sep. 2017, pp. 5304–5309.

[28] J. Moreau, S. Ambellouis, and Y. Ruichek, "Fisheye-based method for GPS localization improvement in unknown semi-obstructed areas," *Sensors*, vol. 17, no. 1, p. 119, 2017.

[29] F. Zhou, X. Chai, X. Chen, and Y. Song, "Omnidirectional stereo vision sensor based on single camera and catoptric system," *Appl. Opt.*, vol. 55, no. 25, pp. 6813–6820, 2016.

[30] H. J. Kim *et al.*, "Fully autonomous vision-based net-recovery landing system for a fixed-wing UAV," *IEEE/ASME Trans. Mechatronics*, vol. 18, no. 4, pp. 1320–1333, Aug. 2013.

[31] W. Kong, D. Zhang, X. Wang, Z. Xian, and J. Zhang, "Autonomous landing of an UAV with a ground-based actuated infrared stereo vision system," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Nov. 2013, pp. 2963–2970.

[32] W. Kong *et al.*, "A ground-based optical system for autonomous landing of a fixed wing UAV," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Sep. 2014, pp. 4797–4804.

[33] T. Yang *et al.*, "A ground-based near infrared camera array system for uav auto-landing in GPS-denied environment," *Sensors*, vol. 16, no. 9, p. 1393, 2016.

[34] A. Ellmauthaler, E. A. B. da Silva, C. L. Pagliari, J. N. Gois, and S. R. Neves, "A novel iterative calibration approach for thermal infrared cameras," in *Proc. IEEE Int. Conf. Image Process.*, Sep. 2013, pp. 2182–2186.

[35] O. Barnich and M. Van Droogenbroeck, "ViBe: A universal background subtraction algorithm for video sequences," *IEEE Trans. Image Process.*, vol. 20, no. 6, pp. 1709–1724, Jun. 2011.
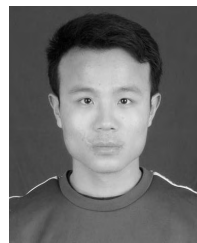
ZHI LI received the B.S. degree in computer science and technology from Northwestern Polytechnical University, in 2016, where he is currently pursuing the master's degree in computer science and technology. His research interests include camera array imaging, unmanned aerial vehicle navigation, multi-target detection and tracking, and visual simultaneous localization and mapping.

FANGBING ZHANG received the M.S. degree from the School of Communication and Information Systems, Xidian University, Xi'an, China, in 2018. She is currently pursuing the Ph.D. degree with the School of Computer Science and Technology, Northwestern Polytechnical University, Xi'an. Her research interests include multi-target detection and tracking based on a binocular stereo vision systems.

BOLIN XIE received the B.S degree in computer science and technology from Northwestern Polytechnical University, in 2017, where he is currently pursuing the master's degree in computer science and technology. His research interests include unmanned aerial vehicle landing, and multi-target detection and tracking.
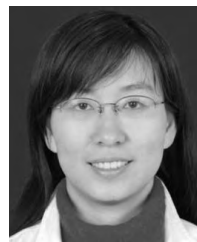
JING LI (M'14) received the Ph.D. degree in control theory and engineering from Northwestern Polytechnical University, Xi'an, China, in 2008. She was a Visiting Scholar with the National Laboratory of Pattern Recognition, Beijing, China, from 2004 to 2005, and a Research Assistant with the Department of Computing, The Hong Kong Polytechnic University, in 2008. She was a Visiting Scholar with the University of Delaware, USA, from 2013 to 2014. She is currently an Associate Professor with the School of Telecommunications Engineering, Xidian University, Xi'an, where she serves as the Leader of the Intelligent Signal Processing and Pattern Recognition Laboratory. She has published over 50 research papers in international journals and conference proceedings in the areas of computer vision and pattern recognition. Her research interests include image registration, matching and retrieval, and video content analysis and understanding.

TAO YANG (M'13) received the Ph.D. degree in control theory and engineering from Northwestern Polytechnical University, Xi'an, China, in 2008. He was a Visiting Scholar with the National Laboratory of Pattern Recognition, Intelligent Video Surveillance Group, Beijing, China, from 2004 to 2005, a Staff Member with Microsoft Research Asia, and a Research Intern with the FX Palo Alto Laboratory, Palo Alto, CA, USA, from 2006 to 2007. He was a Postdoctoral Fellow with the Shaanxi Provincial Key Laboratory of Speech and Image Information Processing, Northwestern Polytechnical University, from 2008 to 2010. He was a Visiting Scholar with the University of Delaware, USA, from 2013 to 2014. He is currently a Full Professor with the School of Computer Science, Northwestern Polytechnical University. His research interests include image and video registration in addition to video content analysis and understanding. He has published over 50 research papers in these fields. He also serves as a Reviewer for numerous international journals, conferences, and funding agencies.

LINFENG LIU received the B.S. degree from the School of Information Science and Engineering, Henan University of Technology, Zhengzhou, China, in 2018. He is currently pursuing the M.S. degree with the School of Computer Science and Technology, Northwestern Polytechnical University, Xi'an, China. His research interests include simultaneous localization and mapping based on a monocular vision systems.

• • •