AMATEUR DRONE SURVEILLANCE: APPLICATIONS, ARCHITECTURES, ENABLING TECHNOLOGIES, AND PUBLIC SAFETY ISSUES

Multiple Moving Targets Surveillance Based on a Cooperative Network for Multi-UAV

Jingjing Gu, Tao Su, Qiuhong Wang, Xiaojiang Du, and Mohsen Guizani

The authors propose a new cooperative network platform and system architecture of multi-UAV surveillance. They propose the design concepts of a multi-UAV cooperative resource scheduling and task assignment scheme based on the animal colony perception method. They provide the moving small target recognition technique and localization and tracking model using the fusion of multiple data sources.

ABSTRACT

With the development of better links, enhanced coverage, comprehensive data resources, and network system stability, the cooperative network formed by wireless sensor networks and unmanned aerial vehicles is envisioned to provide immediate and long-term benefits in military and civilian fields. Previous works mainly focus on how to use UAVs to assist WSNs in sensing and data collection jobs, or target localization with a single data source in surveillance systems, while the potential of multi-UAV sensor networks has not been fully explored. To this end, we propose a new cooperative network platform and system architecture of multi-UAV surveillance. First we propose the design concepts of a multi-UAV cooperative resource scheduling and task assignment scheme based on the animal colony perception method. Second, we provide the moving small target recognition technique and localization and tracking model using the fusion of multiple data sources. In addition, this article discusses the establishment of suitable algorithms based on machine learning due to the complexity of the monitoring area. Finally, experiments of recognition and tracking of multiple moving targets are addressed, which are monitored by multi-UAV and sensors.

INTRODUCTION

Unmanned aerial vehicles (UAVs), with the advantages of easy deployment and flexible usage, have been widely adopted in many military applications, including air reconnaissance, battlefield surveillance, target localization, tracking, damage assessment, and anti-terrorism arrests. They are also adopted in many civilian applications such as aerial photography, geophysical exploration, disaster monitoring, and coastal anti-smuggling. Recently, the applications that use UAVs to collect sensor data have received intensive attention because of wireless sensor networks’ (WSNs’) advantages like miniaturization of sensors, lower costs, availability of various types of data (temperature, humidity, angle and voltage, etc.), and fast and flexible deployment [1, 2]. A cooperative network formed by UAVs and sensors can provide support for a variety of studies. This kind of network can expand the coverage to a border area, thus effectively broadening the time and space coverage of monitoring and detection.

In this research and these applications of UAVs, the surveillance of complex environments or targets is a significant application whose core technology is recognition and tracking of multiple moving targets [3]. It is an interdisciplinary complex issue that involves multi-sensor information fusion, image processing, and artificial intelligence and control technology, due to organizing devices with various types of loads to carry out the task together. However, in practical applications, there are even more challenges for researchers to recognize and locate the multiple moving targets. In this article, we mainly solve the following:

1. In the field of image processing, there has been a lot of research on target recognition from which we can learn [4]. However, in the field of UAV applications, one of the common problems is recognition of small moving targets. Due to long distance, small targets in images often lack concrete shape, size, texture, and other features that can be used. It is difficult to detect a target only using gray information. Moreover, if the video image background has heavy noise and clutter, small targets are usually buried in a complex background with low signal-to-noise ratio (SNR).

2. In practical applications, especially localization and tracking, it is hard for the algorithms to meet the real-time requirement because UAVs and targets are likely to move at high speed. However, the state-of-the-art methods based on machine learning (ML) [5-6] could make the situation even worse by introducing more complex models to improve the accuracy. Therefore, how to balance speed and accuracy is another unsolved challenge in our problem.

3. From real implemented localization and tracking experiments, the UAVs and targets moving at high speed lead to unstable communication links and packet losses, which hinder the collection of enough data, such as received signal strength (RSS) and time of arrival (TOA), or data with various kinds of bias.

As far as we know, there has been no effective way to solve these problems in UAV applications. Taking into account these uncertain factors and challenges, the main contributions include:

Recognition of a small moving target based on video images. We mainly solve the challenge of small targets lacking clear features even under good light condition by detecting the moving targets through motion segmentation. The recognition problem is converted to an optimization
problem of recovering the sparse matrix and the low-rank matrix, which represent targets without background and a combination of neighboring frames, respectively. A cooperative task assignment based on a swarm intelligence optimization algorithm. With full consideration of the dynamic change in complex environment, multi-objective optimization of UAVs’ cooperative task assignment is performed based on the swarm intelligence optimization (SIO) algorithm, which is inspired by living group cooperative activities [7].

An efficient localization algorithm based on multiple data source fusion. To avoid significant errors or incomplete information on a single data source, and the consequent compromised results, localization and tracking of moving targets based on multiple data sources are employed. We mainly propose a localization model based on ML and data fusion methods to provide refined localization. For the calculation efficiency problem, the localization processing is divided into two phases: offline and online. In the offline phase, we build a model to reflect the characteristics and inherent network information of the whole surveillance region. In the online phase, we locate the targets by using the established model in the offline phase.

Experiments based on a real deployed multiple-UAV system. The results of the experiments of recognizing, localizing, and tracking air and ground moving targets are shown through the cooperative work of UAVs and the sensor network.

Network Environment and System Structure

UAVs are desired to work in human-hostile or unreachable environments for such tasks as performing search and rescue missions in explosion and fire scenes, as well as earthquake and disaster sites, or carrying out reconnaissance and arrest missions in mountains and forests. A UAV cluster system consists of multiple UAVs collaborating to complete tasks, and implement functions like task decomposition, cooperative scheduling, and segmented operations. It can coordinate multiple heterogeneous, low-cost UAVs to efficiently complete complex tasks. The data link networks of the UAV cluster have better fault tolerance and certain self-healing ability to provide usability assurance [3]. The multi-UAV system can perform data analysis and fusion on the results output with different devices and performance, which can not only monitor a wider range, but also improve the accuracy and precision of target recognition, localization, and tracking.

Figure 1 is an envisioned multi-UAV platform diagram based on a sensor network. The monitoring areas are supervised with ground deployed sensors, and UAVs equipped with different types of sensors, optical cameras, and infrared cameras for a multi-source dataset and distributed decision making [8]. The UAVs form a cluster with flying ad hoc networks (FANETs). The communication of UAV-UAV, UAV-sensor, UAV-ground station, and UAV-command center can be achieved by data links with short-range wireless communication technology (e.g., Bluetooth and Wi-Fi) [3] to perceive the real-time surrounding environment and collect data. The collected data includes not only the basic sensed information from the ground deployed sensors (e.g., signal strength, time, temperature, humidity, and wind speed direction), but also the data collected by UAVs (e.g., images, infrared data, radar data, and laser imaging data). These data are sent to a cloud data center through a suitable wireless network technology, like a fourth generation (4G)-LTE or 5G.
Our surveillance system includes recognizing targets, establishing localization model through multi-source data, and coordinating multi-UAV to determine that which UAV is assigned to perform the localization and tracking task. Moreover, we implement our surveillance module based on ML algorithms.

The collected data can be processed directly on the relevant UAVs or transmitted to the ground station and the server of a cloud data center for processing when energy resources are insufficient or computation consumption is large. The processed results could be used for tasks like resource scheduling, collaborative path planning, cooperative task planning, and target recognition, localization, and tracking, and form a unified platform for decision making.

The system architecture of a multi-UAV surveillance platform is shown in Fig. 2. The architecture can be divided into the service layer, support middleware layer, and physical layer.

1. The physical layer mainly involves three sublayers: the cloud data center equipment, the Internet of Things (IoT), and the infrastructure. The cloud data center equipment is mainly used to deal with the collected data, including relevant network equipment, such as the cluster server, network interconnection equipment, and storage devices. The IoT is mainly used to collect data and provide network connection between the UAVs/sensors and the cloud data centers, including facilities like access units, gateways, routers, sink nodes, and mobile terminals [10, 11]. The UAV platform is in the IoT sublayer and mainly includes data sensing and the acquisition subsystem, control subsystem, power subsystem, and communication subsystem. In the data sensing and acquisition subsystem, information on a monitored area is sensed and collected through various airborne devices, like an inertial navigator, an optical camera, an infrared camera, airborne radar, and a flight control sensor. The control subsystem usually includes flight control, flight status monitoring, take-off and landing control, and equipment management and control. The power subsystem is designed to satisfy power requirements of UAVs and the network. The communication subsystem includes communication between UAVs and other aircraft or UAVs and ground systems. The IoT sublayer also contains the sensor subsystem that refers to a variety of sensors deployed in the monitoring region, including a temperature sensor, RSS transmitter and receiver, an infrared sensor, a laser sensor, and so on. The infrastructure sublayer is the basis of realization of the overall architecture, including base stations, satellites, ground-based radars, ground stations, and related staff.

2. The support middleware layer is used to provide the interface between the service layer and the physical layer. It mainly provides all types of service support related to data processing, including data management, temporal and spatial alignment, feature extraction, real-time data processing, task assignment, data association, data fusion, and data storage.

3. The top-level service layer is mainly used to provide a variety of services, including target discovery, recognition, localization, tracking, assisted decision making, unified situation generation, situational consistency assessment, and map cloud services.

**MULTIPLE MOVING TARGETS SURVEILLANCE**

Our surveillance system, as shown in Fig. 3, includes recognizing targets, establishing a localization model through multi-source data, and coordinating multi-UAV to determine which UAV is assigned to perform the localization and tracking task. Moreover, we implement our surveillance module based on ML algorithms, which can optimize the performance by using example data or past experience. The detailed processes are discussed in the following subsections.

**MOVING TARGET RECOGNITION**

Moving target recognition based on UAVs in complex environments has always been a hot topic. A UAV equipped with a camera takes aerial
video images in the mission execution area and transmits real-time data to the ground station. After receiving the image data, the general recognition processes are:

1. **Image preprocessing**: Preliminarily detect images to remove low correlation or low-quality images.

2. **Motion segmentation**: Split moving objects of images from the background [5].

3. **Feature extraction**: Extract features of objects and background.

4. **Feature matching**: Match features with the feature template that has been stored to identify the same feature in multiple images.

5. **Perform target recognition and confirm the targets that need to be monitored or searched.**

It is generally known that in the field of UAV application, relevant target recognition research studies mainly include the following aspects [12].

**Motion Segmentation**: Effective segmentation of the motion area is important for feature extraction, feature expression and matching, and final recognition, since the postprocesses mainly take into account pixels in the image that correspond to the motion area. The typical methods are background subtraction (e.g., Gaussian of mixture and approximate median), optical flow, frame differencing, and so on.

**Feature Extraction**: The selected features can not only represent images, but also distinguish different categories of objects. Typical methods include shape-based, motion-based, texture-based, histogram orientation gradient (HOG)-based, algebraic-based, and geometric-based.

**Feature Matching**: Most of these works transform the feature matching problem into pattern classification problems by using supervised ML methods including support vector machine (SVM), AdaBoost, k-nearest neighbor (k-NN), conditional random field (CRF), sparse representation classification (SRC), artificial neural network (ANN), and so on.

Although the above methods have some advantages like mature algorithms and easy implementation, there are still some challenges when applying them to real UAV applications. Because UAVs work in a rapid flight state, collected images are vulnerable to environment change, such as light and shade change, target background changes, and disturbance of targets with similar shapes, which disturb the recognition. In addition, for UAV aerial videos or images, targets tend to become small targets that occupy only a small portion of the pixels of the image. These small targets are usually moving, and the background images probably change dynamically as well. Also, there may be large noise in these backgrounds. These would lead to small SNR and hinder the extraction and recognition of small targets.

Herein, we aim to solve this problem with the following processes.

**Step 1**: We first preprocess the distorted original video images by video stabilization technology [5]. Then we obtain the image sequences of the video frame by frame. If it is the first frame, we convert it into the initial background. If not, we convert it into a gray image.

**Step 2**: We extract the moving target from image sequences (by comparing the current frame with the previous one), and form the target image and background image. We then convert the background image into a background matrix $B$ and convert the target image into a target matrix $T$. If the size of each frame in the video is $m \times n$, where $m$ is the frame height and $n$ is the frame width (in our experiment, $m = 1080$, $n = 1920$), we obtain $B = \{b_{ij}\}_{m \times n}$ and $T = \{t_{ij}\}_{m \times n}$.

Here, $b_{ij}$ and $t_{ij}$ are presented as the pixel values of the $j$th ($j = 1, \ldots, n$) column and the $i$th ($i = 1, \ldots, m$) row in the background image and target image, respectively. Especially in the matrix, if the pixel is a background pixel, $t_{ij} = 0$. It is not difficult to find that $T$ is a sparse matrix when the target is relatively small with respect to the whole image.

**Step 3**: If the video contains $N$ frames (in our experiment, we do a motion segmentation every 300 frames, i.e., $N = 300$), we can stretch each
frame of the video into a vector, \( s_k \in \mathbb{R}^D \) (\( k = 1, \ldots, N \)), where \( D = m \times n \) and \( s_k \) consists of all pixels by scanning from left to right and from top to bottom in the \( k \)th frame. We can get a comprehensive image matrix \( S = [s_1, \ldots, s_k, \ldots, s_N] \in \mathbb{R}^{D \times N} \) by combining the vectors of all frames. In the same way, we can get a combined background matrix \( \hat{B} = [\hat{b}_1, \ldots, \hat{b}_N] \in \mathbb{R}^{D \times N} \) and a combined target matrix \( \hat{T} = [\hat{t}_1, \ldots, \hat{t}_N] \in \mathbb{R}^{D \times N} \). \( \hat{B} \) is a low-rank matrix because there are few changes in the background of two continuous frames.

**Step 4:** We extract the target by recovering sparse and low-rank matrices, more specifically, by solving the optimization problem

\[
\underset{\hat{B}, \hat{T}}{\text{argmin}} \|\hat{B} + \hat{T}\|_1 \tag{1}
\]

subject to \( S = \hat{B} + \hat{T} \). Here, \( \|\cdot\|_1 \) means the kernel-norm of the matrix (i.e., the sum of the singular values of the matrix), \( \|\cdot\|_1 \) means the 1-norm of matrix, and \( \lambda \) is a positive weighting parameter. We use the augmented Lagrange multiplier (ALM) [5] method to solve this problem.

**Step 5:** We solve the feature matching problem through SCR [12]. This has the advantages of low complexity and strong model generalization ability because of unbiased estimation of generalization error.

### Multi-UAV Cooperative Tasks

After targets are recognized, the next major work is to localize and track them in the surveillance system. Nowadays, some UAVs have local data storage, computing, and processing capabilities, which enable them to carry out tasks onboard. Thus, we need a more applicable approach in which the station chooses one specific UAV to perform localization and tracking according to the characteristics between each UAV and the targets.

The task assignment should not only satisfy the required constraints and scheduling objectives, but also provide a cooperative task assignment scheme. A good cooperative scheduling scheme can shorten the working hours, avoid occurrence of redundancy and conflict, reduce energy consumption, as well as improve the utilization of resources and success rate of the implementation of cooperative tasks.

After scheduling the cooperative work, UAVs not only need to know “what tasks” they should perform, but also need to specify how to “perform the tasks.” Such collaborative resource scheduling of multiple UAVs can be regarded as an optimization problem. We can establish the optimization model based on data collected by both sensor nodes and UAVs, such as distances, speeds, angles, and ranges.

To solve this optimization problem, we adopt swarm intelligence optimization (SIO) algorithms, which have been extensively studied in recent years [7]. The swarm intelligence is derived from macro collaboration intelligent behavior by a gregarious colony. It simulates the functions and behaviors of a biological system, which is distributed, decentralized, and self-organized. It also acts autonomously while searching for targets and relaying the information to all swarm members. Due to these characteristics, SIO has been applied for autonomous UAV control. The typical SIO algorithms that are applicable to UAVs are ant colony optimization (ACO), particle swarm optimization (PSO), artificial fish swarm optimization (AFSO), artificial bee colony (ABC), genetic algorithm (GA), pigeon-inspired optimization (PIO), and so on [7]. Each of them has its advantages to apply to different network scenarios. For instance, ACO, AFSO, and PIO have stronger robustness; PSO, ABC, and PIO have faster convergence speed; and GA is applicable to the optimization problem with multi-peak function.

In our work, we adopt PIO to build the optimization model, which is inspired by the natural behavior of the homing pigeon [13]. In the model, virtual pigeons are used to simulate the process of exploring the optimal solution path. The position and speed of the pigeons are initialized according to the landmark operator, map, and compass operator. In the multidimensional search space, the position and speed of the pigeons are updated in each iteration. This article, based on the PIO algorithm, calculates the global optimal solution and solves the collaborative optimization problem of task assignment. Taking the battery energy constraints into account, the goal of the collaborative scheduling is to assign different targets and resources to suitable UAVs to perform the entire task.

### Target Localization and Tracking Based on Multiple Data Sources

ML-based methods can take advantage of data’s inherent information, extract data’s topologic structure, optimize the nonlinear and noisy pattern, and build the mapping between data and environment. Consequently, during the last decade, many published research works focus on localization and tracking by utilizing ML algorithms in the wireless network field [6]. However, usually, the more accurate the model reflecting the characteristics of data and environment, the more prior knowledge might be induced. Furthermore, applying ML-based methods in UAV application has strict requirements of calculation speed and energy consumption. Therefore, in this article, we prefer an effective and feasible method of localization with two phases:

- **Model-estimating phase:** The localization model is built by as much collected data and experience information as possible.
- **An online localization phase:** The locations of targets are estimated by using the learned model.

In the first phase, an ML algorithm matched with the data distribution is selected and improved. Then it can be applied in the current scenario for better localization performance and stability. Due to digging up and reflecting both topology structures and characteristics in the whole network region, the model can make up the impact of environment conditions and has strong robustness. In this phase, a mapping between the position of targets and the acquired data is also built. Thus, in the second phase, we can estimate the position of a target according to its related data through the mapping. If the mapping is designed appropriately, it requires just a few steps of linear calculations whose cost is affordable by UAVs, and it can be used in real-time localization and tracking in the UAV network.
In the multi-UAV spatial cooperative network based on the sensor network, the types of data sources are diversified. This is not only because the collected information comes from the distributed multi-UAV, but also because the types of sensors are varied. For example, we use GPS to measure the position and velocity of each UAV, gyroscopes, and magnetometers to estimate the attitude, accelerometers, and acceleration of each UAV. All types of equipment loaded on the UAVs can collect different types of data, such as images, GPS, infrared, and laser signals. It is a great challenge to effectively perform data association, temporal and space alignment, and data fusion so that more accurate target localization and tracking results can be achieved.

In this article, we use the transmitted image information, signal data, time information, GPS and inertial sensors carried by UAVs, and laser range finder to comprehensively perform the effective target localization. Figure 3 shows how to build the model by different data sources. The following are the detailed processes.

**Images and GPS:** We analyze the images, measure the focal length of the image, and then calibrate the distortion image by the image focal length. After analyzing the calibrated image, we would obtain the coarse-grained coordinate position of the target. Then we use the GPS data and inertial sensors carried on UAVs to estimate the real-time altitude of UAVs [14], and use the laser range finder to measure the distance between UAVs and the target.

**Signal Strengths and Time Arrivals:** Sensors, deployed both on UAVs and on the ground, can measure the signal strengths and time differences of arrival between targets and sensors. Because of environmental influence, the collected signal values with noisy data may have large deviation. We adopt a Kalman filter to refine the data by utilizing the error characteristic data so that we can get more accurate data for data compensation [15].

After processing the data, we set each data source with a certain weighting factor according to the quality of the collected data. The weights are added to build a weight-based-fusion model by multi-source data association. Then we locate the targets, and the station assigns tasks to relevant UAVs under current locations of the moving targets. Then we get ready to collect information of the targets, and wake up the sensor that is related to the forward direction of targets. Therefore, we can constantly locate and track the moving targets with UAVs.

**Experiment Validation**

We use three four-rotor UAVs (DJI-M100, DJI-Innovations Inc.) to build a cooperative multiple-target localization and tracking system of our UAVs and sensor network. Each of them can carry 1.2 kg of equipment, including an airborne computer, "Manifold," with an embedded Linux system, a sensor node, and a camera. Figure 4b shows the UAV used in our experiment. The environment of our experiment is an approximate 110 m × 90 m open space, and 10 sensor nodes are deployed on the ground. In this scenario, two targets, including an air flight target and a ground moving target, carried on sensors need to be recognized and located. We use a small electric truck to stimulate the ground moving target and one UAV (DJI-Phantom series, DJI-Innovations Inc.) as the air target in our experiment, as shown in Fig. 4a.

The control software in the airborne computer mainly contains the flight control module, the target localization and tracking module, and the communication module. The flight control module has the landing, take-off, and flight trajectory control functions. The target localization and tracking module can locate and track the moving targets by using the established model, which has been described in the above parts. The communication module is mainly used for communication.
between sensor nodes and the PC on the control terminal. With a completely charged battery, the flight time of a UAV is around 30 minutes with a full payload.

Our experiments include two parts:
1. Recognition of moving small targets
2. Localization and tracking of both ground and air moving targets

Figure 5 shows some of the results of the first experiment, where Fig. 5a is an original image of a frame extracted from a video whose length is 300 frames. We can find that compared to the whole image, the size of the moving target is very small. There are a number of sensors whose sizes are similar to the target. And there are some large trucks, which are easy to confuse with the identified target. The key of this experiment is to separate the moving small target from the background image (including large trucks, sensors, houses, trees, grass, ground lines, etc.). Figure 5b is a grayscale image, and Figs. 5c and 5d are the background image and target image obtained by recovering the low-rank matrix and the sparse matrix (a detailed introduction can be found in earlier sections). From the experimental results, we can find that the background in Fig. 5c is completely extracted. In Fig. 5d (the target image), the target object is successfully extracted, and the small target is basically separated from the background. However, there is still some noise. The main reason is that the video background has dynamic changes. Also, the background images of two continuous frames sometimes change obviously, which would cause certain errors on the experimental results.

In the second experiment, the adopted sensors use a low-power and low-data-rate Zigbee protocol. At the beginning of the experiment, UAVs get their own position information according to their carried GPS, and then fly in accordance with a path that we set in advance. During the flight, UAVs periodically broadcast their own location information to the entire WSN through the onboard sensor node in 3 Hz. Figure 6 shows the tracking results of our experiments. It illustrates the GPS trajectories (denoted by red lines) planned in advance and their corresponding estimated trajectories (denoted by blue lines). Figure 6a shows the result of the ground target, while Fig. 6b shows those of the air targets. From these figures, we can find that our estimated trajectories are roughly consistent with the real values.

**CONCLUSION**

The research and application of UAVs are developing dramatically. When equipped with different monitoring and sensing devices, UAVs can integrate with WSNs and form integrated sky-ground cooperative networks. In this article, we describe the scenarios of such a network and design its system architecture. As an important application of this network, we first introduce the multi-UAV cooperative resource scheduling and task planning scheme. For data with multiple data sources based on multi-UAV and sensor collection, we propose a target recognition, location, and tracking method. This article also presents two real implemented experiments on moving target recognition and localization both on the ground and in the air. The results show that the proposed model and method have high accuracy.

**ACKNOWLEDGMENTS**

This work has been supported in part by the Aeronautical Science Foundation of China under grant no. 2016ZC52030.

**REFERENCES**


Figure 6. Targets' GPS trajectories and their estimation trajectories: a) result of the ground target; b) result of the air target.