

# Image-Based Multi-UAV Tracking System in a Cluttered Environment

Hsin-Ai Hung, Hao-Huan Hsu, and Teng-Hu Cheng

**Abstract**—A tracking controller for unmanned aerial vehicles (UAVs) is developed to track moving targets undergoing unknown translational and rotational motions. The main challenges are to control both the relative positions and angles between the target and the UAVs to within desired values, and to guarantee that the generated control inputs to the UAVs are feasible (i.e., within their motion capabilities). Moreover, the UAVs are controlled to ensure that the target always remains within the fields of view of their onboard cameras. To the best of our knowledge, this is the first work to apply multiple UAVs to cooperatively track a dynamic target while ensuring that the UAVs remain connected and that both occlusion and collisions are avoided. To achieve these control objectives, a designed controller solved based on the aforementioned tracking controller using quadratic programming can generate minimally invasive control actions to achieve occlusion avoidance and collision avoidance. Furthermore, control barrier functions (CBFs) with a distributed design are developed in order to reduce the amount of inter-UAV communication. Simulations were performed to assess the efficacy and performance of the developed CBF-based controller for the multi-UAV system in tracking a target.

**Index Terms**—Unscented Kalman filter, Estimation, Tracking of moving targets, UAV

## I. INTRODUCTION

The agile motion capabilities of unmanned aerial vehicles (UAVs) are being utilized to perform an increasingly diverse range of tasks, such as surveillance [1], [2], aerial photography [3], [4], and rescue [5]. One of the critical issues in completing these tasks is target tracking. Previous work [6] has demonstrated that a UAV is capable of tracking a dynamic and temporarily-occluded target undergoing unknown motions. However, tracking failure can occur when the occlusion period is too long to accurately predict the target motion. Using a single UAV to track a target can also lead to mission failure due to loss of tracking caused by occlusion. In contrast, combining multiple UAVs and inter-UAV communication can increase the robustness of tracking tasks in many applications, such as tracking a moving target from different angles [7] and tracking multiple moving targets [8].

Visual servo control has been developed for tracking objects, where knowledge of the target motion is important

for executing tracking control using UAVs. However, in many cases the motion of the target is unknown and difficult to predict. Many works have addressed this issue using visual errors for feedback control to keep the target within the field of view (FOV) [9]. For example, [10] and [11] determined the position of a target in an image frame based on its specific color and shape. Bounding boxes derived from an image tracking system (e.g., KCF-TLD) or a deep neural network (DNN) have also been utilized as a feature for feedback control [6], [12]–[14]. However, these previous works focused on either object detection or visual feedback control, and did not consider collision avoidance and occlusion avoidance.

It is crucial to avoid obstacle collision and occlusion in a cluttered environment during a mission. A collision can cause tracking failure and occlusion can cause loss of the target image that leads to failure of state estimation of the target undergoing unknown motion. In [15]–[17], collision avoidance was addressed by including a term in the cost function that considered distances from an obstacle. In [3] and [4], occlusion avoidance was addressed by solving a nonconvex optimization problem. However, collisions and occlusion may still occur since they were only considered as soft constraints [17]. This issue can be solved by designing control barrier functions (CBFs) [18], [19] as hard constraints in the optimization problem, which guarantees system safety as long as the control inputs and the optimization problem can be solved under the safety constraint. Additionally, the computational cost is important when designing controllers with the capability of obstacle avoidance, and this can be reduced by leveraging different mathematical characteristics. For example, a minimally invasive collision-avoidance controller was introduced in [20], where the controller can change the original control program only when it is absolutely necessary. However, the heavy computation load associated with solving the Hamilton-Jacobi-Bellman equations limits the applicability of that approach to large-scale multirobot systems.

CBFs are similar to control Lyapunov functions in the sense that they guarantee the forward invariance of the safety set without computing the system's reachable set. The use of CBFs for coordinating multirobot systems has been studied in recent years [18], [19]. CBFs were first applied to guarantee system safety in applications such as adaptive cruise control [21], legged robots [22], and the Segway vehicle [23]. CBFs have also been utilized to maintain the connectivity of a multirobot network [24], [25]. Multiple control objectives for robots were also addressed, such as

Department of Mechanical Engineering, National Yang Ming Chiao Tung University, Hsinchu 30010, Taiwan. Email: louise.me08g@nctu.edu.tw, water4202@gmail.com, tenghu@g2.nctu.edu.tw

This research is supported by the Ministry of Science and Technology, Taiwan (Grant Number MOST 110-2222-E-A49 -005 -) and by Pervasive Artificial Intelligence Research (PAIR) Labs, Taiwan (Grant Number MOST 110-2634-F-009 -018 -).

collision avoidance and connectivity maintenance [26], [27]. To increase the fault tolerance and guarantee the safety of robots based on information from neighboring robots, distributed CBFs were developed in [28], [29]. Motivated by the aforementioned works, distributed CBFs are designed in this work for multiple UAVs to cooperatively track a moving target while ensuring collision avoidance, connectivity maintenance, and occlusion avoidance. A target tracking controller is first developed based on image feedback, where You Only Look Once (YOLO) DNN is designed to generate a bounding box and determine the relative angle of the target as features for tracking control. To further meet the constraints of collision avoidance, occlusion avoidance, and connectivity maintenance, a modified controller designed based on a tracking controller using quadratic programming (QP) is developed to generate minimally invasive control actions to ensure occlusion avoidance and collision avoidance. The main contributions of this work can be summarized as follows:

- 1) In contrast to previous approaches, the relative rotation between the target and the UAV is modeled and can be controlled to a desired value.
- 2) Multiple UAVs are employed to cooperatively track a dynamic target by communicating its position among neighboring UAVs for feedback to ensure collision avoidance and connectivity maintenance.
- 3) A multiobjective optimization problem based on CBFs is developed to ensure not only collision avoidance and connectivity maintenance, but also avoiding occlusion by obstacles.
- 4) The velocity of the UAV not only restricted by the CBF, but also determined by the dynamics of the desired relative angle. Therefore, the interplay between the estimation, control, and constraints are determined by the developed optimal problem.
- 5) There are many ways to design the CBF functions, but to ensure that the controller is distributed, a sufficient condition needs to be developed and analyzed. For example, the condition  $\gamma_s$  defined in (28) is designed to ensure the designed controller is distributed.
- 6) Moreover, the CBFs and the controllers are co-designed to ensure the existence of the barrier functions and to ensure that the forward invariance property is preserved based on [18]. Therefore, as long as the constraints can be satisfied, the optimal solution can be solved from the cost function.
- 7) Slack variables are also added and used to determine if the constraints are too conservative.

The report is organized as follows: Section II formulates the multi-UAV tracking problem and describes the model for kinematics and dynamics between the UAVs and the target. Section III designs nonlinear model predictive control (NMPC) for target tracking. Section IV introduces CBFs for guaranteeing the safety of the multi-UAV system. In Section V, an optimization problem based on the CBFs

as constraints is constructed and solved for the minimally invasive controller. The simulation results are presented in Section VI, and conclusion are drawn and future work is summarized in Section VII.

## II. PRELIMINARIES

### A. Motivation

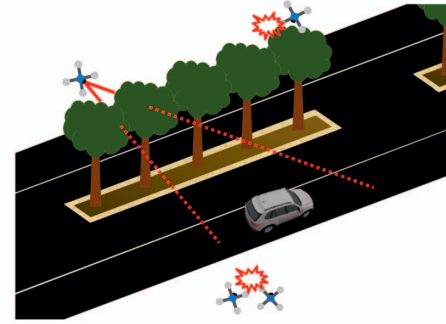


Fig. 1. Multiple UAVs tracking a moving target in a cluttered environment.

Fig. 1 illustrates the difficulties encountered when multiple UAVs are tracking a moving target in a cluttered environment. The tracking mission will fail if the UAVs collide with each other or with the static obstacles (e.g., trees). Moreover, occlusion of the target by obstacles will prevent the UAVs from detecting the target, and leading to failure of the tracking mission. To solve these problems, CBFs for (i) collision avoidance (ii) connectivity maintenance, and (iii) occlusion avoidance are introduced in this work.

### B. Kinematics Model

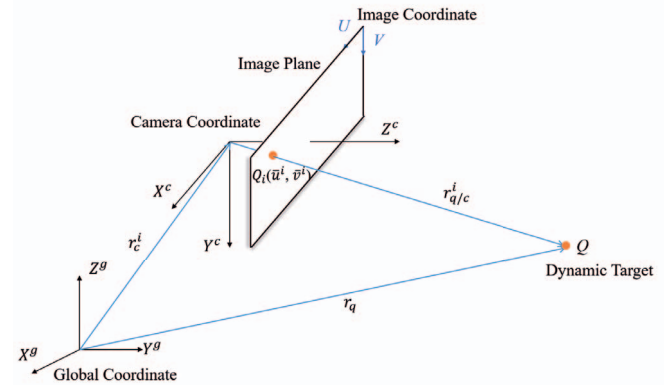


Fig. 2. Model for the kinematics between the camera and the target.

Consider a team of  $n$  UAVs in the system as depicted in Fig. 2, which shows the relationship between a moving target and an onboard camera fixed on the  $i^{\text{th}}$  UAV, where  $i \in N \triangleq \{1, 2, \dots, n\}$ .  $r_q^i = [x_q^i \ y_q^i \ z_q^i]^T$  denotes the position of the moving target estimated from the  $i^{\text{th}}$  UAV, and  $r_c^i = [x_c^i \ y_c^i \ z_c^i]^T$  denotes the position of the camera fixed on the  $i^{\text{th}}$  UAV, which can be measured by the embedded

GPS or motion-capture systems, all expressed in the camera frame. The relative position between the moving target and the camera is defined as

$$r_{q/c}^i = r_q^i - r_c^i, \quad (1)$$

where  $r_{q/c}^i = [X^i \ Y^i \ Z^i]^T$  is defined in the camera frame. Taking the time derivative on both sides of (1) yields the relative velocity as

$$\dot{r}_{q/c}^i = V_q^i - V_c^i - \omega_c^i \times r_{q/c}^i, \quad (2)$$

where  $V_q^i = [v_{qx}^i \ v_{qy}^i \ v_{qz}^i]^T$  is the translational velocity of the dynamic target, and  $V_c^i \triangleq [v_{cx}^i \ v_{cy}^i \ v_{cz}^i]^T$  and  $\omega_c^i \triangleq [\omega_{cx}^i \ \omega_{cy}^i \ \omega_{cz}^i]^T$  are the translational and angular velocities of the camera, respectively, both of which represent the control commands to be designed. Each of the UAVs is located at  $p^i = (r_c^i)^g \in \mathbb{R}^3$  with single-integrator dynamics

$$\dot{p}^i = (V_c^i)^g \in \mathbb{R}^3, \quad (3)$$

where the  $(\cdot)^g$  superscript denotes the vector defined in the global frame, and the stacked position of the team of UAVs is denoted as  $p = [p^1, \dots, p^n] \in \mathbb{R}^{3n}$ , along with control command  $(u^i)^g = [(V_c^i)^g \ (\omega_c^i)^g]^T \in \mathbb{R}^6$ .

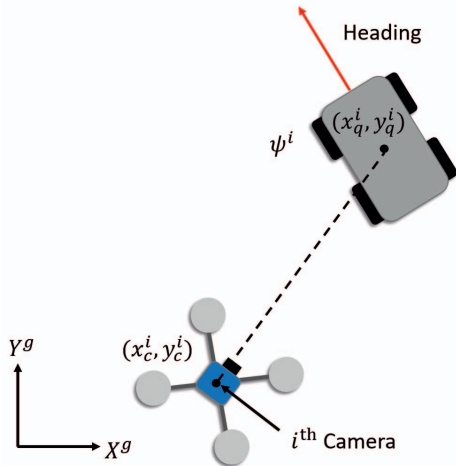


Fig. 3.  $\psi^i$  is defined as the angle between the heading of the target and the relative position from the target to the  $i^{th}$  UAV in the global frame.

Tracking a target undergoing 2D motion at a specific relative location requires knowledge of the relative angle of the target denoted as  $\psi^i \in \mathbb{R}$  (Fig. 3):

$$\psi^i = \arctan\left(\frac{(y_c^i)^g - (y_q^i)^g}{(x_c^i)^g - (x_q^i)^g}\right) - \arctan\left(\frac{(v_{qy}^i)^g}{(v_{qx}^i)^g}\right). \quad (4)$$

**Assumption 1.** The trajectory of the target is unknown but bounded; that is,  $r_q, \dot{r}_q, \ddot{r}_q \in \mathcal{L}_\infty$ .

To estimate the state of the target, a dynamics model and a measurement model are introduced in Section II-C and II-D.

### C. Dynamics Model

Since the dynamics of the camera and the target are coupled, the states of the overall system are defined as

$$\mathbf{x}^i = [x_1^i, x_2^i, x_3^i, \psi^i, (r_q^i)^T, (V_q^i)^T]^T, \quad (5)$$

where

$$[x_1^i, x_2^i, x_3^i] = \left[ \frac{X^i}{Z^i}, \frac{Y^i}{Z^i}, \frac{1}{Z^i} \right] \quad (6)$$

are the states defined to facilitate the subsequent analysis. The dynamics of  $\mathbf{x}^i$  can be obtained by taking the time derivative on both sides of (5):

$$\dot{\mathbf{x}}^i = \begin{bmatrix} v_{qx}^i x_3^i - v_{qz}^i x_1^i x_3^i + \zeta_1^i + \eta_1^i \\ v_{qy}^i x_3^i - v_{qz}^i x_2^i x_3^i + \zeta_2^i + \eta_2^i \\ -v_{qz}^i (x_3^i)^2 + v_{cz}^i (x_3^i)^2 - (\omega_{cy}^i x_1^i - \omega_{cx}^i x_2^i) x_3^i \\ \dot{\psi}^i \\ V_q^i \\ 0_{3 \times 1} \end{bmatrix}, \quad (7)$$

where (2) is used and  $\zeta_1^i, \zeta_2^i, \eta_1^i, \eta_2^i \in \mathbb{R}$  are defined as

$$\begin{aligned} \zeta_1^i &= \omega_{cz}^i x_2^i - \omega_{cy}^i - \omega_{cy}^i (x_1^i)^2 + \omega_{cx}^i x_1^i x_2^i \\ \zeta_2^i &= -\omega_{cz}^i x_1^i + \omega_{cx}^i + \omega_{cx}^i (x_2^i)^2 - \omega_{cy}^i x_1^i x_2^i \\ \eta_1^i &= (v_{cz}^i x_1^i - v_{cx}^i) x_3^i \\ \eta_2^i &= (v_{cz}^i x_2^i - v_{cy}^i) x_3^i. \end{aligned}$$

### D. Measurement Model

This section defines two measurements performed using a YOLO DNN. The bounding box and the relative angle measurement provide information to correct the states in the unscented Kalman filter (UKF) developed in Section II-E. A detailed description of the DNN can be found in [6].

a) **Bounding Box and Relative Angle:** Based on the pinhole model, states  $x_1^i$  and  $x_2^i$  can be obtained by projecting the position of the target onto the image plane as

$$\begin{aligned} x_1^i &= \frac{\bar{u}^i - c_u}{f_x} \\ x_2^i &= \frac{\bar{v}^i - c_v}{f_y} \\ x_3^i &= \frac{1}{d^i}, \end{aligned} \quad (8)$$

where  $f_x$  and  $f_y$  are the focal lengths,  $c_u$  and  $c_v$  represent the center of the image,  $\bar{u}^i$  and  $\bar{v}^i$  represent the center of the detected bounding box of the moving target along the  $U, V$  direction in the image frame (see Fig. 2), and  $d^i$  is the depth between the target and the  $i^{th}$  UAV. Therefore, based on (1), (6), and (8), the measurement model can be derived as

$$h(\mathbf{x}^i) = \begin{bmatrix} f_x x_1^i + c_u \\ f_y x_2^i + c_v \\ 1/x_3^i \\ \psi^i \\ r_c^i \end{bmatrix}, \quad (9)$$

where  $\psi^i$  is also measured by the YOLO as shown in Fig. 3.

*Remark 1.* Changes in the target velocity can result in the developed UKF making inaccurate state estimations. By considering the model mismatch between the real and constant-velocity models, defined in (7), as process noise in the UKF, the states can be estimated and updated to approach the ground truth.

### E. Unscented Kalman Filter

A UKF [30] is applied in this section to estimate the state of dynamic systems with noisy or intermittent measurements, since a UKF can outperform an extended Kalman filter in several ways: dealing with highly nonlinear systems, accurate to two terms of the Taylor expansion, and higher efficiency. By discretizing (7) and (9) using the forward Euler method and including noise, the nonlinear dynamic system can be expressed as

$$\mathbf{x}_{k+1}^i = F(\mathbf{x}_k^i, \mathbf{u}_k^i) + w_k \quad (10)$$

$$\mathbf{z}_k^i = H(\mathbf{x}_k^i) + v_k, \quad (11)$$

where  $w_k$  and  $v_k$  represent the process and measurement noise, respectively,  $\mathbf{x}_k^i$  is the discretized states of the overall system defined in (5),  $\mathbf{u}_k^i$  is the discretized control input,  $F(\cdot)$  is the dynamics model used for prediction,  $H(\cdot)$  is the measurement model used for updating, and  $\mathbf{z}_k^i$  is the discretized measurement defined as

$$\mathbf{z}_k^i = \begin{bmatrix} \bar{u}_z^i \\ \bar{v}_z^i \\ d_z^i \\ \psi_z^i \\ r_{c,z}^i \end{bmatrix},$$

where the  $z$  subscript represents the measured value of the denoted argument. When the bounding box cannot be detected, the state is predicted by the dynamics model using (10), which is reliable over a short period of time before the detection process recovers.

*Remark 2.* The UKF-based method developed in Section II-E utilizes a high-order polynomial to continuously estimate the unknown motion of a moving target; that is, the polynomial is found based on a QP method to fit the position and velocity data over the latest time period (cf. (27) of [6]).

## III. NONLINEAR MODEL PREDICTIVE CONTROL

An NMPC is designed in this section and is then utilized in Section V to compute the minimally invasive controller. The control input of the NMPC is calculated by optimizing the cost function designed in this section that considers not only the current and upcoming feature errors, but also the control input. Using the NMPC can decrease both the aggressiveness of the flying behavior and the energy consumption.

### A. Transformation of the Dynamics Model

Since computing the dynamics defined in (7) is too inefficient, the rotational dynamics is transformed to increase the computational efficiency. On the other hand, to avoid the UAVs reaching their maximum speed during tracking, slow time variations on the angular velocity of the target are also considered in Assumption 2.

**Assumption 2.** The target is initially located within the FOV of the camera, and the angular velocity of the target exhibits only slow time variations (i.e.,  $\omega_q \approx 0$ ). During the steady state, the UAV is controlled such that its camera is looking almost directly at the target (i.e.,  $V_q \approx V_c$ ), as shown in Fig. 4. In other words, the heading of the camera is parallel to the direction of the position vector from the camera to the target.

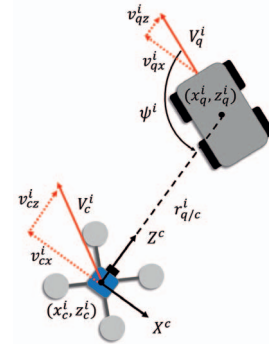


Fig. 4. Directions of velocities expressed in the camera frame.

Based on Assumption 2, velocities  $V_q^i$  and  $V_c^i$  defined in (2) satisfy  $V_q^i \approx V_c^i$  and  $\omega_q^i \approx 0$ , which implies  $\dot{\psi}^i \approx (\omega_{cz}^i)^g = -\omega_{cy}^i$ . Therefore, the dynamics of relative angle  $\psi^i$  can be expressed as

$$\dot{\psi}^i = -\omega_{cy}^i = \frac{-(v_{qx}^i - v_{cx}^i)}{\|r_{q/c}^i\|},$$

which is equivalent to

$$\dot{\psi}^i = \frac{(v_{cx}^i - v_{qx}^i)x_3^i}{\sqrt{1 + (x_1^i)^2 + (x_2^i)^2}} \quad (12)$$

based on the relation

$$\begin{aligned} \|r_{q/c}^i\| &= \sqrt{(X^i)^2 + (Y^i)^2 + (Z^i)^2} \\ &= \frac{\sqrt{1 + (x_1^i)^2 + (x_2^i)^2}}{x_3^i}. \end{aligned} \quad (13)$$

The dynamics model of the overall system defined in (7) can

be replaced with

$$\dot{\mathbf{x}}^i = \begin{bmatrix} v_{qx}^i x_3^i - v_{qz}^i x_1^i x_3^i + \zeta_1 + \eta_1 \\ v_{qy}^i x_3^i - v_{qz}^i x_2^i x_3^i + \zeta_2 + \eta_2 \\ -v_{qz}^i (x_3^i)^2 + v_{cz}^i (x_3^i)^2 - (\omega_{cy}^i x_1^i - \omega_{cx}^i x_2^i) x_3^i \\ \frac{(v_{cx}^i - v_{qx}^i) x_3^i}{\sqrt{1+(x_1^i)^2+(x_2^i)^2}} \\ V_q^i \\ 0_{3 \times 1} \end{bmatrix}. \quad (14)$$

*Remark 3.* Assumption 2 might not always be satisfied, in which cases the mismatches can be considered as process noise and be corrected using the relative angle as measured by the YOLO network.

### B. Nonlinear Model Predictive Control

Only part of the dynamics model in (14) is required to further design the controller; that is, given the UAV constraints  $\omega_{cx}^i \approx 0$  and  $\omega_{cz}^i \approx 0$  according to [10], the control input can be reduced to  $u^i \triangleq [v_{cx}^i, v_{cy}^i, v_{cz}^i, \omega_{cy}^i]^T \in \mathbb{R}^4$ , and the dynamics model of feature vector  $s_m^i = [x_1^i, x_2^i, x_3^i, \psi^i]^T \in \mathbb{R}^4$  used for prediction in the NMPC can be written as

$$\dot{s}_m^i = \begin{bmatrix} -x_3^i(v_{cx}^i - v_{qx}^i) + x_1^i x_3^i(v_{cz}^i - v_{qz}^i) - (1 + (x_1^i)^2)\omega_{cy}^i \\ -x_3^i(v_{cy}^i - v_{qy}^i) + x_2^i x_3^i(v_{cz}^i - v_{qz}^i) - \omega_{cy}^i x_1^i x_2^i \\ (x_3^i)^2(v_{cz}^i - v_{qz}^i) - \omega_{cy}^i x_1^i x_3^i \\ \frac{(v_{cx}^i - v_{qx}^i) x_3^i}{\sqrt{1+(x_1^i)^2+(x_2^i)^2}} \end{bmatrix}. \quad (15)$$

To deal with the modeling errors associated with using the transformed dynamics model and the imperfect actuation of the  $i^{\text{th}}$  UAV, an error signal  $\epsilon_k$  is designed as

$$\epsilon_k = s_k^i - s_{m,k}^i, \quad (16)$$

where feature vector  $s_k^i$  is estimated using the UKF, and feature vector  $s_{m,k}^i$  is predicted using the transformed model (15) at time  $t_k$ . The desired feature vector utilized in the NMPC can be defined as

$$s_{d,k}^i = s^{i,*} - \epsilon_k, \quad (17)$$

where  $s^{i,*} \triangleq [x_1^*, x_2^*, x_3^*, \psi^*]^T \in \mathbb{R}^4$  is the reference feature vector prescribed by the user.

With (15) and (17), the optimal control problem (OCP) of the NMPC at time  $t_k$  can be expressed as

$$\min_{s_m^i, u^i} \sum_{n=0}^{N_p-1} \left[ \|s_{m,k+n}^i - s_{d,k+n}^i\|_{Q_s}^2 + \|u_{k+n}^i - u_{d,k+n}^i\|_{R_u}^2 \right] + \|s_{m,k+N_p}^i - s_{d,k+N_p}^i\|_{W_s}^2, \quad (18)$$

s.t.

$$\begin{aligned} s_{m,k+n+1}^i &= F_m(s_{m,k+n}^i, u_{k+n}^i), \quad n = 0, \dots, N_p - 1 \\ u_{d,k+n+1}^i &= u_{d,k+n}^i + \left[ \hat{\Gamma}_k^i, \quad 0 \right]^T dt, \quad n = 0, \dots, N_p - 1 \\ s_{m,k+n}^i &\in \mathbb{S}, \quad n = 1, \dots, N_p \\ u_{k+n}^i &\in \mathbb{U}, \quad n = 0, \dots, N_p - 1 \\ s_{d,k+n}^i &= s^* - \epsilon_k, \quad n = 0, \dots, N_p \\ s_{m,k}^i &= s_k^i, \end{aligned} \quad (19)$$

where  $u_d^i \in \mathbb{R}^4$  is defined as the desired velocity that control input  $u^i$  needs to achieve, and can be updated by  $\hat{\Gamma}_k^i$  (as defined in Section 5 of [6]) over the prediction horizon, as depicted in Fig. 5.  $S_m^i = \{s_{m,k}^i, s_{m,k+1}^i, \dots, s_{m,k+N_p}^i\}$  denotes the predicted features at sampling time  $t_k$  based on control sequence  $U^i = \{u_k^i, u_{k+1}^i, \dots, u_{k+N_p-1}^i\}$  and estimation  $s_k^i$  obtained using the UKF.  $Q_s$ ,  $R_u$ , and  $W_s$  defined in (18) are positive weighting matrices, and  $N_p \in \mathbb{N}$  is the prediction horizon. Increasing  $N_p$  decreases the aggressiveness of the control inputs but also increases the computational effort.  $F_m(\cdot)$  is the model defined in (15), and  $\mathbb{S}$  and  $\mathbb{U}$  are the sets of constraints of the feature vector and the control input given by

$$\begin{aligned} \mathbb{S} &\triangleq \left\{ s^i \in \mathbb{R}^4 \mid \begin{bmatrix} x_{1,\min} \\ x_{2,\min} \\ x_{3,\min} \\ \psi_{\min} \end{bmatrix} \leq s^i \leq \begin{bmatrix} x_{1,\max} \\ x_{2,\max} \\ x_{3,\max} \\ \psi_{\max} \end{bmatrix} \right\} \\ \mathbb{U} &\triangleq \left\{ u^i \in \mathbb{R}^4 \mid \begin{bmatrix} v_{cx,\min} \\ v_{cy,\min} \\ v_{cz,\min} \\ \omega_{cy,\min} \end{bmatrix} \leq u^i \leq \begin{bmatrix} v_{cx,\max} \\ v_{cy,\max} \\ v_{cz,\max} \\ \omega_{cy,\max} \end{bmatrix} \right\}. \end{aligned}$$

Error signal  $\epsilon_k$  defined in (16) is assumed to be constant over the prediction horizon, and  $s_{d,k+n}^i$ ,  $n = 0, \dots, N_p$ , would be constant over the prediction horizon.

The direct multiple shooting method is employed to solve the problem since lifting the OCP to a higher dimension can usually speed up the rate of convergence. The OCP is repeated at every sampling time, with only the first vector  $u_k^i$  in control set  $U^i$  being adopted in the system. Several software libraries can be used to solve this kind of nonlinear programming problem (NLP); the CasADi tool is used in this work for formulating the NLP, and the Ipopt library is used to solve it.

*Remark 4.* A control architecture was developed for general target tracking without limitation to vehicles. Therefore, given a bounding box enclosing the target in the image, measurements of the feature vector and relative angle (i.e.,  $s^i$ ) can be used to predict the motion of target  $v_q^i$ , which can improve the tracking performance and, in turn, enhance the detection of bounding boxes.



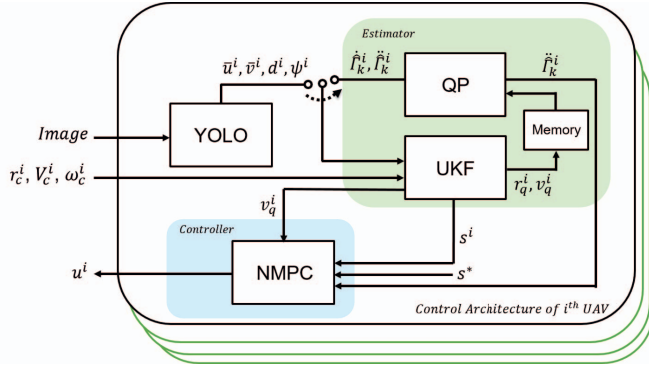


Fig. 5. Block diagram of the tracking controller for the  $i^{\text{th}}$  UAV. A switch is added to select one of the inputs to the UKF depending on the presence of the target in the image. When the measurement data  $[\bar{u}_z^i, \bar{v}_z^i, d_z^i, \psi_z^i]$  are not available due to occlusion, the motion obtained by QP (i.e.,  $\hat{u}_k^i$  and  $\hat{v}_k^i$ ) is used to predict the target motion for the UKF. Otherwise, the measured values obtained from the detected bounding box  $[\bar{u}_z^i, \bar{v}_z^i, d_z^i, \psi_z^i]$  are imported into the UKF for updating.

While using a decentralized NMPC controller  $u^i$  ensures the local optimal configuration of the  $i^{\text{th}}$  UAV, this approach might still lead to mission failure due to UAVs colliding or a disconnected network, since the UAVs need to communicate their positions for their controller to achieve collision avoidance. Therefore, CBFs are introduced to guarantee collision avoidance, occlusion avoidance, and connectivity maintenance when computing control command  $u^i$  developed in (18).

#### IV. CONTROL BARRIER FUNCTIONS

The CBF has been introduced as a tool for guaranteeing the safety of a system [21]. A typical CBF is designed to be a forward invariant set, so that the solved control command can always keep the closed-loop system in the safety set if it starts in that set [26]. To this end, CBFs for collision avoidance, connectivity maintenance, and occlusion avoidance are developed in this section.

##### A. CBF for Collision Avoidance

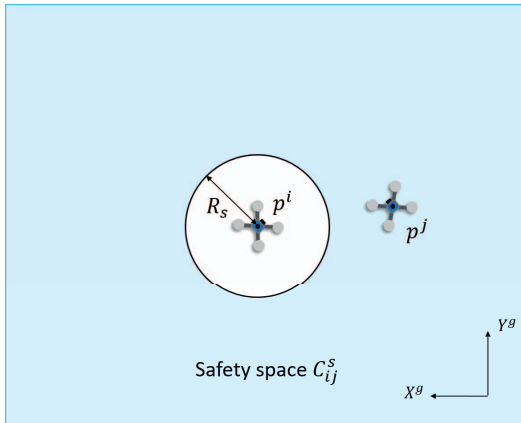


Fig. 6.  $R_s$  denotes the minimum safety range, and the blue zone is safety space  $C_{ij}^s$ , where a potential collision between the  $i^{\text{th}}$  and  $j^{\text{th}}$  UAVs can be avoided.

To avoid collisions between UAVs, a safety space denoted as  $C_{ij}^s$ ,  $\forall i, j \in N$ ,  $i \neq j$  is defined as

$$C_{ij}^s = \{p \in \mathbb{R}^{3n} \mid h_{ij}^s(p) = \|p^i - p^j\|^2 - R_s^2 \geq 0\}, \quad (20)$$

where  $C_{ij}^s$  is the set as shown in Fig. 6. Function  $h_{ij}^s$  is regarded as a CBF that satisfies

$$\dot{h}_{ij}^s(p) \geq -\gamma_s h_{ij}^s, \quad (21)$$

where  $\gamma_s$  is a user-defined positive constant. With (20) and the single-integrator dynamics defined in (3), (21) can be further expressed as

$$-2(p^i - p^j)(V_c^i)^g + 2(p^i - p^j)(V_c^j)^g \leq \gamma_s h_{ij}^s, \quad (22)$$

which can be separated into the following sufficient conditions [29]:

$$-2(p^i - p^j)(V_c^i)^g \leq \frac{1}{2}\gamma_s h_{ij}^s, \quad (23)$$

$$2(p^i - p^j)(V_c^j)^g \leq \frac{1}{2}\gamma_s h_{ij}^s. \quad (24)$$

Inequalities (23) and (24) can be transformed into the following linear inequality:

$$A_{ij}^s(u^i)^g \leq b_{ij}^s, \quad \forall i, j \in N, \quad i \neq j,$$

where

$$A_{ij}^s = \begin{bmatrix} \underbrace{-2(p^i - p^j)^T, 0, 0, 0}_{i^{\text{th}} \text{ UAV}} \end{bmatrix}, \quad (25)$$

$$b_{ij}^s = \frac{1}{2}\gamma_s h_{ij}^s.$$

In other words, the  $i^{\text{th}}$  and  $j^{\text{th}}$  UAVs can always avoid colliding if they are initially in safety space  $C_{ij}^s$  and control input  $(u^i)^g$  lies within admissible control set  $K_{ij}^{sN}$  defined as

$$K_{ij}^{sN} = \{(u^i)^g \in \mathbb{R}^6 \mid A_{ij}^s(u^i)^g \leq b_{ij}^s, \quad \forall i, j \in N, \quad i \neq j\}, \quad (26)$$

where  $K_{ij}^{sN}$  is the admissible control space obtained by the mapping of the safety space defined in (20). It is shown in [31] that if control inputs  $(u^i)^g$  stay in  $K_{ij}^{sN}(p)$ , the UAVs stay within safety space  $C_{ij}^s$  all of the time, and hence the forward invariance property of  $C_{ij}^s$  is guaranteed.

*Remark 5.* Admissible control set  $K_{ij}^{sN}$  defined in (26) ensures the safety of the entire UAV team; however, it requires position information from all of the UAVs (i.e.,  $j \in N$ ) in the network, and the amount of communication required to achieve this will increase with the number of UAVs. This makes it necessary to design a decentralized CBF based only on information about neighboring UAVs.

To this end, (26) is modified as

$$K_{ij}^s = \{(u^i)^g \in \mathbb{R}^6 \mid A_{ij}^s(u^i)^g \leq b_{ij}^s, \forall i \in N, j \in N_{si}\}, \quad (27)$$

$$\gamma_s \geq \frac{4\alpha_v R_s}{D_s^2 - R_s^2}, \quad (28)$$

where  $N_{si} = \{j \in N \mid \|p^i - p^j\| \leq D_s, \forall i \neq j\}$  is the set of the  $i^{\text{th}}$  UAV's neighbors that collisions start to be considered, and  $D_s$  is the distance, which satisfies  $R_s < D_s \leq R_c$ , where  $R_c \in \mathbb{R}$  is the maximum connectivity range. If  $\gamma_s$  satisfies (28), it is proven in [32] that safety is still guaranteed by admissible control set  $K_{ij}^s$  defined in (27), where  $\alpha_v$  is the upper bound of the translational velocities of the camera  $\|(V_c^i)^g\| \leq \alpha_v, \forall i \in N$ . The CBF can also be utilized to avoid collisions between the static obstacles and the UAVs.

### B. CBF for Connectivity Maintenance

In order to ensure inter-UAV communications, the distance between the  $i^{\text{th}}$  and  $j^{\text{th}}$  UAVs needs to be maintained to within maximum connectivity range  $R_c$ , which is defined by connectivity space  $C_{ij}^c$ ,  $\forall i \in N, j \in N_{ci}$  as

$$C_{ij}^c = \{p \in \mathbb{R}^{3n} \mid h_{ij}^c(p) = R_c^2 - \|p^i - p^j\|^2 \geq 0\}, \quad (29)$$

and the neighbor set of the  $i^{\text{th}}$  UAV is defined as

$$N_{ci} = \{j \in N \mid \|p^i - p^j\| \leq R_c, \forall i \neq j\}.$$

It is assumed that the  $i^{\text{th}}$  UAV can obtain the positions of its neighboring UAVs once they are inside connectivity range  $R_c$ . Function  $h_{ij}^c$  is considered to be a CBF that satisfies

$$\dot{h}_{ij}^c(p) \geq -\gamma_c h_{ij}^c, \quad (30)$$

where  $\gamma_c$  is a user-defined parameter for fine tuning the connectivity space defined in (29). With (29) and the single-integrator dynamics defined in (3), (30) needs to satisfy

$$2(p^i - p^j)(V_c^i)^g - 2(p^i - p^j)(V_c^j)^g \leq \gamma_c h_{ij}^c, \quad (31)$$

which can be separated into the following sufficient conditions [29]:

$$2(p^i - p^j)(V_c^i)^g \leq \frac{1}{2}\gamma_c h_{ij}^s, \quad (32)$$

$$-2(p^i - p^j)(V_c^j)^g \leq \frac{1}{2}\gamma_c h_{ij}^s. \quad (33)$$

Inequalities (32) and (33) can be further transformed into the following linear inequalities:

$$A_{ij}^c(u^i)^g \leq b_{ij}^c, \forall i \in N, j \in N_{ci},$$

where  $A_{ij}^c$  and  $b_{ij}^c$  are defined as

$$A_{ij}^c = \begin{bmatrix} \underbrace{2(p^i - p^j)^T}_{i^{\text{th}} \text{ UAV}}, 0, 0, 0 \end{bmatrix}, \quad (34)$$

$$b_{ij}^c = \frac{1}{2}\gamma_c h_{ij}^c.$$

Based on (34), the  $i^{\text{th}}$  and  $j^{\text{th}}$  UAVs will always stay connected if they are initially connected, and control input  $(u^i)^g$  inside admissible control space  $K_{ij}^c$  is defined as

$$K_{ij}^c = \{(u^i)^g \in \mathbb{R}^6 \mid A_{ij}^c(u^i)^g \leq b_{ij}^c, \forall i \in N, j \in N_{ci}\}. \quad (35)$$

### C. CBF for Occlusion Avoidance

In Fig. 7, the position of each obstacle is defined as  $p_o^j \in \mathbb{R}^3, j \in M \triangleq \{1, 2, \dots, m\}$ , and  $n^i$  is defined as the vector from the  $i^{\text{th}}$  UAV to the target in the global frame

$$n^i = (R_c^i)^g r_{q/c}^i, \quad (36)$$

where  $(R_c^i)^g$  is the rotation matrix of the  $i^{\text{th}}$  UAV from the camera frame to the global frame, and  $r_{q/c}^i$  is defined in (1). To ensure that the  $i^{\text{th}}$  UAV can always detect the target, occlusion angle  $\theta^i \in \mathbb{R}$  obtained by vectors  $p_o^j - p^i$  and  $n^i$  needs to be kept larger than a user-defined FOV (field of view)  $\theta^*$ . In this way, it creates an occlusion-free area in the image, and the CBF for occlusion avoidance  $h_{ij}^o$  can be designed as

$$h_{ij}^o(p^i, p_o^j) = \theta^i - \theta^* \geq 0, \quad (37)$$

where

$$\theta^i = \cos^{-1} \left( \frac{(p_o^j - p^i) \cdot n^i}{\|p_o^j - p^i\| \|n^i\|} \right). \quad (38)$$

The FOV  $\theta^*$  considers the motion of the target and creates a moving occlusion-free area centered at the target in the image. Obstacles are allowed to partially show up in the image as long as they are not inside the occlusion-free area, as shown in Fig. 7, which can decrease the aggressiveness of the flight control input for the UAVs.

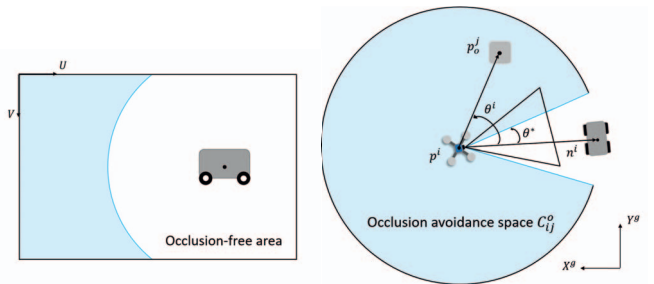


Fig. 7. Occlusion-free area and avoidance space.  $n^i$  is the relative position between the target and the  $i^{\text{th}}$  UAV. The blue zone is occlusion avoidance space  $C_{ij}^o$  where the  $j^{\text{th}}$  obstacle will not block the target in the image. The white zone is the occlusion-free area where the  $j^{\text{th}}$  obstacle is not allowed to exist.

Combining (37) and (38), occlusion avoidance space

$C_{ij}^o$ ,  $\forall i \in N$ ,  $j \in N_{oi}$  can be defined as

$$C_{ij}^o = \left\{ p^i \in \mathbb{R}^3 \mid h_{ij}^o(p^i, p_o^j) \triangleq \cos^{-1} \left( \frac{(p_o^j - p^i) \cdot n^i}{\|p_o^j - p^i\| \|n^i\|} \right) - \theta^* \geq 0 \right\}, \quad (39)$$

where  $N_{oi} = \{j \in M \mid \|p_o^j - p^i\| \leq \|r_{q/c}^i\|\}$  is the set of obstacles around the  $i^{\text{th}}$  UAV since it is possible that they will occlude the target. Function  $h_{ij}^o$  is considered to be a CBF that satisfies

$$\dot{h}_{ij}^o(p^i, p_o^j) \geq -\gamma_o h_{ij}^o, \quad (40)$$

where  $\gamma_o$  is also a user-defined constant for fine tuning invariant set  $C_{ij}^o$ .

With (39) and given that  $\frac{d}{dx} \cos^{-1}(x) = \frac{-1}{\sqrt{1-x^2}}$ , with  $x = \frac{(p_o^j - p^i) \cdot n^i}{\|p_o^j - p^i\| \|n^i\|}$ , (40) expressed as

$$\begin{aligned} \dot{h}_{ij}^o &= \frac{\partial \cos^{-1}(x)}{\partial x} \frac{\partial (x)}{\partial t} \\ &= \frac{-1}{\sqrt{1-(x)^2}} \left( \frac{n^i}{\|p_o^j - p^i\| \|n^i\|} \cdot (\dot{p}_o^j - \dot{p}^i) \right) \\ &\quad + \frac{-1}{\sqrt{1-(x)^2}} \left( \frac{-(p_o^j - p^i) \cos \theta^i}{\|p_o^j - p^i\|^2} \cdot (\dot{p}_o^j - \dot{p}^i) \right) \\ &\quad + \frac{-1}{\sqrt{1-(x)^2}} \left( \frac{(p_o^j - p^i)}{\|p_o^j - p^i\| \|n^i\|} \cdot \dot{n}^i \right) \\ &\quad + \frac{-1}{\sqrt{1-(x)^2}} \left( \frac{-n^i \cos \theta^i}{\|n^i\|^2} \cdot \dot{n}^i \right) \geq -\gamma_o h_{ij}^o, \end{aligned} \quad (41)$$

where  $\dot{n}^i$  can be expressed as

$$\begin{aligned} \dot{n}^i &= (\dot{R}^i)_c^g r_{q/c}^i + (R^i)_c^g \dot{r}_{q/c}^i \\ &= (\hat{\omega}_c^i)^g (R^i)_c^g r_{q/c}^i + (R^i)_c^g (V_q^i - V_c^i - \omega_c^i \times r_{q/c}^i), \end{aligned} \quad (42)$$

provided that the obstacles are static (i.e.,  $\dot{p}_o^j = 0$ ), where  $(\cdot)$  denotes the hat map, and  $r_{q/c}^i$  is defined in (2). Based on Assumption 2 (i.e.,  $V_q^i \approx V_c^i$ ), (42) can be further expressed as

$$\begin{aligned} \dot{n}^i &= (\hat{\omega}_c^i)^g (R^i)_c^g r_{q/c}^i - (R^i)_c^g (\omega_c^i \times r_{q/c}^i) \\ &= (\hat{\omega}_c^i)^g n^i - (\omega_c^i)^g \times n^i \\ &= 0. \end{aligned} \quad (43)$$

Based on (43), (41) can be rewritten as

$$\begin{aligned} \dot{h}_{ij}^o &= \frac{-1}{\sqrt{1-(x)^2}} \frac{n^i}{\|p_o^j - p^i\| \|n^i\|} (V_c^i)^g \\ &\quad - \frac{-1}{\sqrt{1-(x)^2}} \frac{(p_o^j - p^i) \cos \theta^i}{\|p_o^j - p^i\|^2} (V_c^i)^g \leq \gamma_o h_{ij}^o, \end{aligned} \quad (44)$$

which can be transformed into a linear inequality

$$A_{ij}^o (u^i)^g \leq b_{ij}^o \quad \forall i \in N, j \in N_{oi},$$

where

$$\begin{aligned} A_{ij}^o &= \underbrace{\left[ \frac{-1}{\sqrt{1-(x)^2}} \left( \frac{n^i}{\|p_o^j - p^i\| \|n^i\|} - \frac{(p_o^j - p^i) \cos \theta^i}{\|p_o^j - p^i\|^2} \right) \right]^T}_{i^{\text{th}} \text{ UAV}}, \quad 0^{1 \times 3} \\ b_{ij}^o &= \gamma_o h_{ij}^o. \end{aligned} \quad (45)$$

In other words, the occlusion of the target is avoided if control input  $(u^i)^g$  stays in admissible control space  $K_{ij}^o$  defined as

$$K_{ij}^o = \{ (u^i)^g \in \mathbb{R}^6 \mid A_{ij}^o (u^i)^g \leq b_{ij}^o \quad \forall i \in N, j \in N_{oi} \}. \quad (46)$$

## V. CONTROLLER DESIGN

To control both the relative positions and angles between the target and the team of UAVs to be within desired values, the objective of this work is to control the UAVs to achieve

$$s_m^i \rightarrow s^{i*}.$$

To guarantee the cooperative tracking by the UAVs is feasible, constraints of their interaction and the generated control inputs need to be enforced. Therefore, the following optimal control problem is defined.

Given optimal tracking control input  $u^i$  obtained from the NMPC designed in (18), its augmented controller  $\hat{u}^i = (R^i)_c^g [v_{cx}^i, v_{cy}^i, v_{cz}^i, 0, \omega_{cy}^i, 0]^T \in \mathbb{R}^6$  expressed in the global frame is defined. Then controller  $(u^i)^g$ , defined as a minimally invasive controller of  $\hat{u}^i$ , is solved as the following optimization problem using standard QP solvers to achieve the tracking performance while ensuring that the closed-loop system remains in the forward invariant set of the CBFs defined in Section IV:

$$\min_{(u^i)^g \in \mathbb{R}^6} \| (u^i)^g - \hat{u}^i \|^2, \quad (47)$$





in Table III. The parameters for deciding the safety ranges and the restrictiveness of the control space in CBFs are listed in Table IV.

TABLE I  
STATE CONSTRAINTS.

State	Minimum	Maximum
$x_1^i$	-0.84	0.84
$x_2^i$	-0.63	0.63
$x_3^i$	0.07	1
$\psi^i$	0°	180°

TABLE II  
CONTROL CONSTRAINTS.

Inputs	Minimum	Maximum
$v_{cx}^i, v_{cy}^i, v_{cz}^i$ (m/s)	-10	10
$w_{cy}^i$ (rad/s)	-0.6	0.6

TABLE III  
CONTROL PARAMETER SETTINGS.

Parameter	Value
Horizon length ( $N_p$ )	50
$Q_s$	$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 100 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$
$R_u$	$\begin{bmatrix} 0.02 & 1 & 1 & 1 \\ 1 & 0.03 & 1 & 1 \\ 1 & 1 & 0.01 & 1 \\ 1 & 1 & 1 & 0.3 \end{bmatrix}$

TABLE IV  
CBF PARAMETER SETTINGS.

Parameter	Value	Parameter	Value
$R_s$ (m)	2	$\gamma_s$	3
$R_c$ (m)	20	$\gamma_c$	1
$\theta^\circ$ (deg)	30	$\gamma_o$	0.1
$D_s$ (m)	20	$\alpha_v$ (m/s), $\alpha_\omega$ (rad/s)	10, 0.6

### B. Simulation Results

The performance of the feature tracking is quantified as the root mean square (RMS) tracking errors in Table V, where the values denote the position errors between the center of the bounding box and the reference location (i.e.,  $[x_1^*, x_2^*]$  defined in (17)) in the image frame. However, the baseline RMS errors are hard to obtain since the tracking errors can vary significantly in different scenarios and initial conditions.

TABLE V  
RMS ERRORS IN IMAGE FEATURES.

Agent	U-axis	V-axis	Unit
UAV1	10.26	63.23	pixel
UAV2	8.94	32.66	pixel
UAV3	23.22	86.28	pixel

a) *Distances Between UAVs* : Fig. 10 shows the distances between the UAVs. This figure indicates that the constraints of connectivity maintenance and collision avoidance

were satisfied among the UAVs, since they remained within connectivity  $R_c = 20$  m and collision range  $R_s = 2$  m.

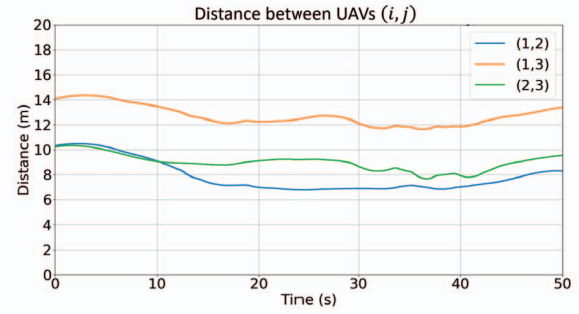


Fig. 10. Distances between the UAVs

b) *Distances Between UAVs and Obstacles*: Fig. 11 and 12 show the distances from the obstacles to UAV1 and UAV3, respectively. These figures indicate that the constraints between the obstacles and UAVs were also satisfied since they followed the rule of keeping a safe distance  $R_s = 2$  m from each other. UAV1 avoided colliding with obstacles 1 to 5 around the time period from 10 s to 30 s, while UAV3 avoided obstacles 6 to 10 from around 30 s to 50 s. Since UAV2 did not encounter any obstacles, its distances are not shown.

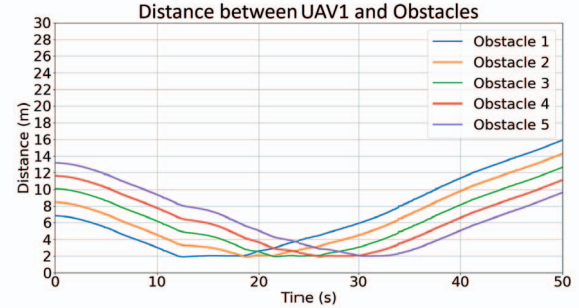


Fig. 11. Distance between UAV1 and the obstacles.

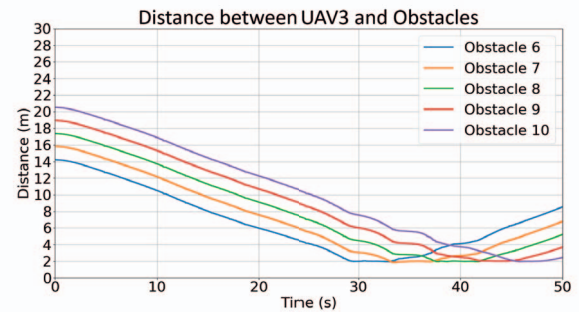


Fig. 12. Distance between UAV3 and the obstacles.

c) *Trajectories of UAVs* : The left figure in Fig. 13 shows the 2D trajectories of the UAVs in the  $X^g-Y^g$  plane.

This figure indicates that when UAV1 first approached obstacle 1, it moved forward to the obstacle to prevent the obstacle from showing up in the camera image; UAV3 similarly did that when it approached obstacle 6. The right figure in Fig. 13 shows the 2D trajectories of the UAVs in the  $X^g-Z^g$  plane. This figure indicates that when UAV1 approached obstacle 1, it moved upward to avoid colliding with the obstacles while also ensuring occlusion avoidance. UAV1 moved downward when it left the obstacles in order to keep the target within the FOV of the camera and thereby allow it to be tracked. The figure shows that UAV3 also satisfied the aforementioned behavior.

## VII. CONCLUSIONS

The simulation results presented here indicate that the developed CBFs can guarantee the safety of a multi-UAV tracking system in a cluttered environment. The modified controller solved based on the NMPC tracking controller using QP can generate minimally invasive control actions and meet the following objectives:

- All UAVs are proven to be safe since the constraints ensure that collisions are prevented, connectivity is maintained, and occlusion of the target in the camera image is avoided.
- The constraints are designed based on information about the neighboring UAVs, which decreases the communication burden and improves the performance of the controller.
- Relative angle  $\psi$  can be prescribed by the user for the desired angle of tracking and observations.

One of the contributions of this work is the relative angle  $\psi$  can be measured by the YOLO DNN, which enables the UAVs to track the target from the desired relative angle to the target. This capability is important in many applications. For example, to autonomously check the license plate of suspect vehicles, the UAV needs to fly to the front or back part of the vehicle based on relative angle feedback. Therefore, the relative angle is highly related to the aim of the task. This capability can also be extended to many applications such as formation, surveillance, etc. In addition, increasing coefficient  $\gamma_o$  will make the occlusion avoidance space more restrictive but will also increase the aggressiveness of the control input. This implies that the approximation error to the NMPC controller defined in (47) is compromised in order to meet the occlusion avoidance constraint. However, gimbals were not included in the current system, and new motion patterns of the UAVs can be expected once a gimbal with two to three degrees of freedom is installed in each UAV. The inclusion of a gimbal in each UAV will facilitate observations of the target with different attitudes at the same location, which implies that their accelerations would also be different. Future works could assume that the positions of the obstacles are unknown and so need to be detected using other algorithms, such as SLAM (simultaneous localization and mapping). Additionally, a CBF for occlusion avoidance could be developed that considers not only static obstacles but also

moving obstacles such as birds, humans, and bicycles. Moreover, investigating the interplay the air turbulence among the UAVs can also improve the instability of the UAVs.

## REFERENCES

- [1] J. Valasek, K. Kirkpatrick, J. May, and J. Harris, "Intelligent motion video guidance for unmanned air system ground target surveillance," *J. Aerosp. Inf. Syst.*, vol. 13, no. 1, pp. 10–26, 2016.
- [2] S. Papatheodorou, A. Tzes, and Y. Stergiopoulos, "Collaborative visual area coverage," *Robot. Autonom. Syst.*, vol. 92, pp. 126–138, 2017.
- [3] T. Nägeli, J. Alonso-Mora, A. Domahidi, D. Rus, and O. Hilliges, "Real-time motion planning for aerial videography with dynamic obstacle avoidance and viewpoint optimization," *IEEE Robot. Autom. Lett.*, vol. 2, no. 3, pp. 1696–1703, 2017.
- [4] R. Bonatti, Y. Zhang, S. Choudhury, W. Wang, and S. A. Scherer, "Autonomous drone cinematographer: Using artistic principles to create smooth, safe, occlusion-free trajectories for aerial filming," *CoRR*, vol. abs/1808.09563, 2018.
- [5] A. Macwan, G. Nejat, and B. Benhabib, "Target-motion prediction for robotic search and rescue in wilderness environments," *IEEE Trans. Syst. Man Cybern., Part B (Cybern.)*, vol. 41, no. 5, pp. 1287–1298, Oct. 2011.
- [6] J.-M. Li, C.-W. Chen, and T.-H. Cheng, "Motion prediction and robust tracking of a dynamic and temporarily-occluded target by an unmanned aerial vehicle," *IEEE Trans. Control Syst. Technol.*, pp. 1–13, 2020, doi:10.1109/TCST.2020.3012619.
- [7] R. Tallamraju, E. Price, R. Ludwig, K. Karlapalem, H. H. Bulthoff, M. J. Black, and A. Ahmad, "Active perception based formation control for multiple aerial vehicles," *IEEE Robotics and Automation Letters*, vol. 4, no. 4, pp. 4491–4498, 2019.
- [8] Y. Zhao, X. Wang, C. Wang, Y. Cong, and L. Shen, "Systemic design of distributed multi-uav cooperative decision-making for multi-target tracking," *Auton. Agents Multi-Agent Syst.*, vol. 33, no. 1-2, pp. 132–158, Jan. 2019.
- [9] M. Razzanelli, M. Innocenti, G. Pannocchia, and L. Pollini, *Vision-based Model Predictive Control for Unmanned Aerial Vehicles Automatic Trajectory Generation and Tracking*.
- [10] I. F. Mondragón, P. Campoy, M. A. Olivares-Mendez, and C. Martinez, "3D object following based on visual information for unmanned aerial vehicles," in *Proc. IEEE IX Latin Am. Robot. Symp. IEEE Colomb. Conf. Autom. Control*, Oct. 2011, pp. 1–7.
- [11] H. Wang, B. Yang, J. Wang, X. Liang, W. Chen, and Y. Liu, "Adaptive visual servoing of contour features," *IEEE/ASME Trans. Mechatron.*, vol. 23, no. 2, pp. 811–822, Apr. 2018.
- [12] D. Hulens and T. Goedemé, "Autonomous flying cameraman with embedded person detection and tracking while applying cinematographic rules," in *Conf. Comput. Robot Vis.*, Edmonton, AB, Canada, May 2017, pp. 56–63.
- [13] Y. Liu, Q. Wang, H. Hu, and Y. He, "A novel real-time moving target tracking and path planning system for a quadrotor UAV in unknown unstructured outdoor scenes," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 49, no. 11, pp. 2362–2372, Nov. 2019.
- [14] J. H. Lee, J. D. Millard, P. C. Lusk, and R. W. Beard, "Autonomous target following with monocular camera on uas using recursive-ransac tracker," in *Int. Conf. Unmanned Aircr. Syst.*, Dallas, TX, USA, Jun. 2018, pp. 1070–1074.
- [15] T. P. Nascimento, L. F. S. Costa, A. G. S. Conceição, and A. P. Moreira, "Nonlinear model predictive formation control: An iterative weighted tuning approach," *J. Intell. Robot. Syst.*, vol. 80, no. 3-4, pp. 441–454, 2015.
- [16] A. Mondal, C. Bhowmick, L. Behera, and M. Jamshidi, "Trajectory tracking by multiple agents in formation with collision avoidance and connectivity assurance," *IEEE Syst. J.*, vol. 12, no. 3, pp. 2449–2460, 2018.
- [17] H.-S. Shin, A. Garcia, and S. Alvarez, "Information-driven persistent sensing of a non-cooperative mobile target using UAVs," *J. Intell. Robot. Syst.*, vol. 92, Dec. 2018.
- [18] D. Pickem, P. Glotfelter, L. Wang, M. Mote, A. Ames, E. Feron, and M. Egerstedt, "The robotarium: A remotely accessible swarm robotics research testbed," in *2017 IEEE Int. Conf. Robot. Autom.*, 2017, pp. 1699–1706.

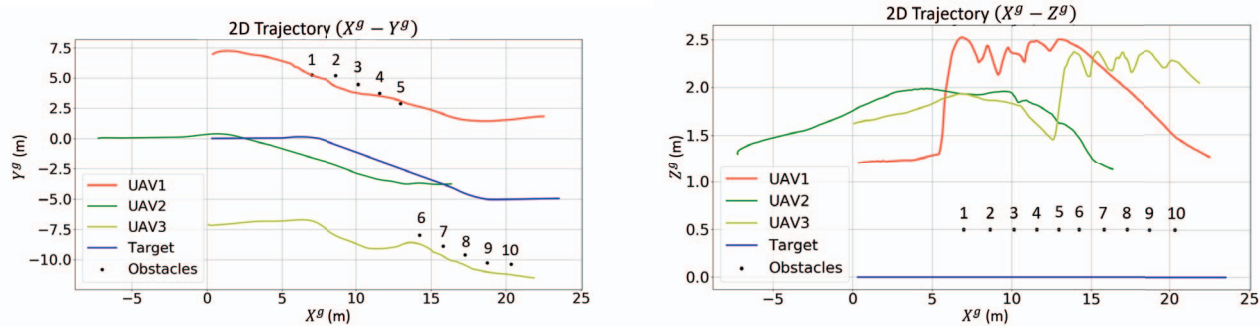


Fig. 13. 2D trajectories of the UAVs in the  $X^g$ - $Y^g$  plane (Left) and 2D trajectories of the UAVs in the  $X^g$  -  $Z^g$  plane (Right).

- [19] A. D. Ames, S. Coogan, M. Egerstedt, G. Notomista, K. Sreenath, and P. Tabuada, "Control barrier functions: Theory and applications," in *2019 Eur. Control Conf.*, 2019, pp. 3420–3431.
- [20] C. Tomlin, G. Pappas, and S. Sastry, "Conflict resolution for air traffic management: A study in multiagent hybrid systems," *IEEE Trans. on Autom. Control*, vol. 43, no. 4, pp. 509–521, 1998.
- [21] A. D. Ames, J. W. Grizzle, and P. Tabuada, "Control barrier function based quadratic programs with application to adaptive cruise control," in *IEEE Conf. Decis. Control*, 2014, pp. 6271–6278.
- [22] Q. Nguyen and K. Sreenath, "Safety-critical control for dynamical bipedal walking with precise footstep placement," *IFAC-PapersOnLine*, vol. 48, no. 27, pp. 147–154, 2015, analysis and Design of Hybrid Systems ADHS.
- [23] T. Gurriet, A. Singletary, J. Reher, L. Ciarletta, E. Feron, and A. Ames, "Towards a framework for realizable safety critical control through active set invariance," in *2018 ACM/IEEE Int. Conf. Cyber-Phys. Syst.*, 2018, pp. 98–106.
- [24] A. Li, L. Wang, P. Pierpaoli, and M. Egerstedt, "Formally correct composition of coordinated behaviors using control barrier certificates," in *2018 IEEE/RSJ Int. Conf. Intell. Robot. Syst.*, 2018, pp. 3723–3729.
- [25] W. Luo and K. Sycara, "Voronoi-based coverage control with connectivity maintenance for robotic sensor networks," in *2019 Int. Symp. Multi-Robot Multi-Agent Syst.*, 2019, pp. 148–154.
- [26] L. Wang, A. D. Ames, and M. Egerstedt, "Multi-objective compositions for collision-free connectivity maintenance in teams of mobile robots," in *2016 IEEE Conf. Decis. Control*, 2016, pp. 2659–2664.
- [27] D. Han and D. Panagou, "Robust multitask formation control via parametric lyapunov-like barrier functions," *IEEE Trans. Autom. Control*, vol. 64, no. 11, pp. 4439–4453, 2019.
- [28] U. Borrmann, L. Wang, A. D. Ames, and M. Egerstedt, "Control barrier certificates for safe swarm behavior," *IFAC-PapersOnLine*, vol. 48, no. 27, pp. 68–73, 2015, analysis and Design of Hybrid Systems ADHS.
- [29] L. Wang, A. Ames, and M. Egerstedt, "Safety barrier certificates for heterogeneous multi-robot systems," in *2016 Am. Control Conf.*, 2016, pp. 5213–5218.
- [30] E. A. Wan and R. van der Menve, "The unscented kalman filter for nonlinear estimation," in *Proc. IEEE 2000 Adaptive Syst. Signal Process., Commun., Control Symp.*, Lake Louise, Alberta, Canada, Oct. 2000, pp. 153–158.
- [31] X. Xu, P. Tabuada, J. W. Grizzle, and A. D. Ames, "Robustness of control barrier functions for safety critical control," *IFAC-PapersOnLine*, vol. 48, no. 27, pp. 54–61, 2015, analysis and Design of Hybrid Systems ADHS.
- [32] M. Endo, T. Ibuki, and M. Sampei, "Collision-free formation control for quadrotor networks based on distributed quadratic programs," in *2019 Am. Control Conf.*, 2019, pp. 3335–3340.



**Hsin-Ai Hung** received his M.S. degree from the department of Mechanical Engineering at the National Yang Ming Chiao Tung University, Taiwan, in 2021. Her current research interests include tracking control, optimal control, and robotics.



**Hao-Huan Hsu** is currently pursuing his M.S. degree in the department of Mechanical Engineering at the National Yang Ming Chiao Tung University, Taiwan. His current research interests include multi-agent systems, tracking control, and robotics.



**Teng-Hu Cheng** received the Ph.D. degree from the Department of Mechanical Engineering, University of Florida, Gainesville, FL, USA, in 2015. In 2016, he joined the Department of Mechanical Engineering, National Yang Ming Chiao Tung University, Hsinchu, Taiwan. He received the O. Hugo Schuck Best Paper Award from the American Automatic Control Council in 2015. His research interests include robotics, autonomous systems, networked system control, switched control, event-driven control, and non-

linear control.