

Active Perception Based Formation Control for Multiple Aerial Vehicles

Rahul Tallamraju , Eric Price , Roman Ludwig, Kamalakar Karlapalem, Heinrich H. Bühlhoff , Michael J. Black , and Aamir Ahmad 

Abstract—We present a novel robotic front-end for autonomous aerial motion-capture (mocap) in outdoor environments. In previous work, we presented an approach for cooperative detection and tracking (CDT) of a subject using multiple micro-aerial vehicles (MAVs). However, it did not ensure optimal view-point configurations of the MAVs to minimize the uncertainty in the person's cooperatively tracked 3D position estimate. In this article, we introduce an active approach for CDT. In contrast to cooperatively tracking only the 3D positions of the person, the MAVs can actively compute optimal local motion plans, resulting in optimal view-point configurations, which minimize the uncertainty in the tracked estimate. We achieve this by decoupling the goal of active tracking into a quadratic objective and non-convex constraints corresponding to angular configurations of the MAVs w.r.t. the person. We derive this decoupling using Gaussian w.r.t. the person model assumptions within the CDT algorithm. We preserve convexity in optimization by embedding all the non-convex constraints, including those for dynamic obstacle avoidance, as external control inputs in the MPC dynamics. Multiple real robot experiments and comparisons involving 3 MAVs in several challenging scenarios are presented.

Index Terms—Visual Tracking, Aerial Systems: Perception and Autonomy, Multi-Robot Systems.

I. INTRODUCTION

AERIAL motion capture (mocap) of humans in unstructured outdoor scenarios is a challenging and important

Manuscript received February 24, 2019; accepted July 3, 2019. Date of publication August 1, 2019; date of current version October 24, 2019. This letter was recommended for publication by Associate Editor S. Julier and Editor E. Marchand upon evaluation of the reviewers' comments. This work was supported in part by the Max Planck Grassroots Project AirCap and in part by the internal budget of the Perceiving Systems Department, MPI for Intelligent Systems. (Corresponding author: Aamir Ahmad.)

R. Tallamraju is with the Max Planck Institute for Intelligent Systems, Tübingen, Germany, and also with the Agents and Applied Robotics Group, IIIT Hyderabad, Hyderabad, 500032, India (e-mail: rahul.tallamraju@tuebingen.mpg.de).

E. Price, R. Ludwig, M. J. Black, and A. Ahmad are with the Max Planck Institute for Intelligent Systems, Tübingen 72076, Germany (e-mail: eprice@tuebingen.mpg.de; roman.ludwig@tuebingen.mpg.de; black@tuebingen.mpg.de; aamir.ahmad@tuebingen.mpg.de).

K. Karlapalem is with the Agents and Applied Robotics Group, IIIT Hyderabad, Hyderabad, 500032, India (e-mail: kamal@iiit.ac.in).

H. H. Bühlhoff is with the Max Planck Institute for Biological Cybernetics, Tübingen 72076, Germany (e-mail: hhb@tuebingen.mpg.de).

This letter has supplementary downloadable material available at <http://ieeexplore.ieee.org>, provided by the authors. The first part of the material, i.e., a pdf file, shows the experimental derivation of the observation measurement model for a single MAV. The second is a video file that shows the experiment footage of our decentralized model predictive control approach for active target tracking using a team of three MAVs and also illustrates the results in simulation using eight MAVs and a human actor.

Digital Object Identifier 10.1109/LRA.2019.2932570

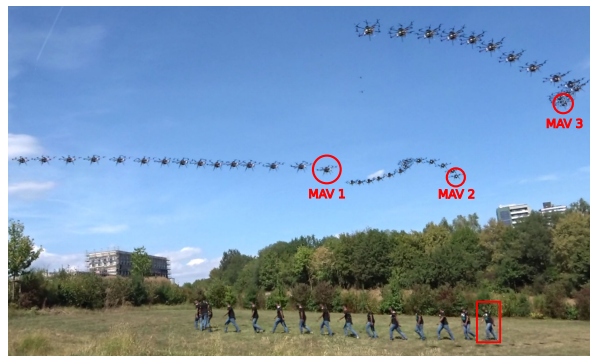


Fig. 1. Multi-exposure image of three octocopter MAVs cooperatively tracking a fast walking person while maintaining a perception-driven formation using the approach proposed in this letter. The person walks from right to left. The initial positions of the MAVs in their trajectories are circled.

problem [1]–[5]. It may directly facilitate applications like sports medicine, cinematography, or search and rescue operations. The front-end of our mocap system consists of a team of micro aerial vehicles (MAVs), autonomously detecting, tracking and following a person. The front-end is responsible for the online task of the system, which is to continuously estimate the 3D global position of the person and keep him/her centered in the field of view of their on-board camera, while he/she performs activities such as walking, running, jumping, etc. The online task is the core focus of this article. The offline task, which is not addressed in this article, is to estimate full body skeleton pose and shape using images acquired by the MAVs.

In order to develop the mocap front-end, in our previous work [1], we presented a marker-less deep neural network (DNN) based cooperative detection and tracking (CDT) algorithm. That work had two major shortcomings. First, the motion planner on each MAV did not facilitate uncertainty minimization of the person's 3D position estimate. Second, the local motion plans generated by the MAVs using a model-predictive controller (MPC) were subsequently modified using potential fields for collision avoidance. This further led to sub-optimal trajectories and robots getting easily stuck in field local minima, e.g., stuck behind each other or behind closely spaced obstacles.

In the current work, we solve both of the aforementioned problems through a novel convex decentralized formation controller based on MPC. This new MPC actively minimizes the joint uncertainty in the tracked person's 3D position estimate while following him/her. The key novelties are as follows.

- The first novel idea, which enables us to formulate this problem as a locally convex MPC, is the decoupling of the joint uncertainty minimization into i) a convex quadratic objective that maintains a threshold distance to the tracked person, and ii) constraints that enforce angular configurations of the MAVs with respect to (w.r.t.) the person. We derive this decoupling based on Gaussian observation model assumptions used within the CDT algorithm.
- To guarantee the safety of the motion plans, we incorporate collision avoidance constraints w.r.t. i) other MAVs, ii) the tracked person and iii) static obstacles, only as locally convex constraints. Collision avoidance and angular configuration constraints are inherently non-convex. We preserve convexity in our MPC formulation by converting them to external control input terms embedded inside the MPC dynamics, which are explicitly computed at every iteration of the MPC.

Real robot experiments and comparisons involving 3 MAVs in several challenging scenarios are presented. Fig. 1 shows a multi-exposure image of a short sequence from one of these experiments. Simulation results with up to 16 robots demonstrate the robustness of our approach and scalability w.r.t. to the number of robots. We provide open-source ROS-based¹ source code of our method.

II. STATE-OF-THE-ART

Multi-robot active tracking: In [6], multi-robot trajectories are analytically derived and uncertainty bounds are identified for optimizing the fused estimate of the target position. In [7], a non-convex optimization problem is solved analytically, to minimize the trace of the fused target covariance. However, their approach is centralized and uses Gauss-Seidel relaxation to compute optimal solutions for multiple robots. Centralized non-convex optimization [8], [9] is used to track a target in stochastic environments and locally minimize the fused estimate of the target position. In [10], active perception based formation control is addressed using a decentralized non-linear MPC. However, their method only identifies sub-optimal control inputs due to the non-convexity of optimization and the lack of collision avoidance guarantees. In [11], a perception-aware MPC generates real-time motion plans which maximize the visibility of a desired target. The motion plans are, however, generated only for a single aerial robot. A marker-based multi-robot aerial motion capture system is presented in [5], where the formation control is achieved using a decentralized non-linear MPC. Scalability in the number of robots and formation collision avoidance behaviors are not explicitly addressed in their approach.

Obstacle Avoidance: Sequential convex programming is applied to solve centralized multi-robot collision-free trajectory generation [12]. Due to non-convexities which arise from obstacle avoidance constraints, the obtained solutions are locally optimal albeit fast for a given time horizon. In [10] repulsive potential field functions are employed as one of the optimization objectives to avoid collisions. This is highly limiting as obstacle avoidance cannot be guaranteed. Recent work in distributed multi-agent obstacle avoidance [13] convexifies the reciprocal velocity obstacle constraint, to characterize velocities that do not lead to a collision in a perfectly localized environment. Motivated by multi-view cinematography applications, distributed non-linear model predictive control [14] is used to identify

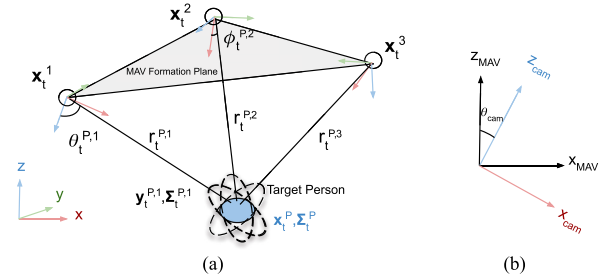


Fig. 2. Formation of 3 MAVs with the mathematical notations used in this letter.

locally optimal motion plans for aerial vehicles. In one of our previous works [15], we developed a convex optimization program to generate local collision-free motion plans, while tracking a movable pick and place static target using multiple aerial vehicles. This approach generates fast, feasible motion plans and has a linear computational complexity ($O(n)$) in the number of environmental obstacles. The work was validated only in a simulation environment and, moreover, obstacle avoidance was not guaranteed. In our current work, we consider real robots, stochasticity in the environment and bounds on repulsive potential fields to guarantee collision avoidance and generate safe local motion plans.

III. PROPOSED APPROACH

A. Preliminaries and Problem Statement

Let there be K MAVs tracking a person P . The pose (position and orientation) of the k^{th} MAV in the world frame at time t is given by $\xi_t^k = [(\mathbf{x}_t^k)^T (\Theta_t^k)^T]^T = [(\mathbf{x}_t^k)^T [\psi_t^k \theta_t^k \phi_t^k]^T]^T \in \mathbb{R}^6$. The MAV's pose uncertainty covariance matrix is denoted as $\Sigma_t^k \in \mathbb{R}^{6 \times 6}$. Each MAV has an on-board, monocular, perspective camera. It is important to note that the camera is rigidly attached to the MAV's body frame, pitched down at an angle of $\theta_{cam} = 45^\circ$. Hence, the pose of the camera in the MAV's frame is $[0 \ 0 \ 0 \ 0 \ \theta_{cam} \ 0]^T$ (see Fig. 2). The observation measurement of the person P 's position made by MAV k in its camera frame is given by a mean range $r_t^{P,k}$, bearing $\phi_t^{P,k}$ and inclination $\theta_t^{P,k}$ (see Fig. 2) in spherical coordinates. These are associated with a zero mean measurement noise, denoted by $\sigma_{r,t}^{P,k}$, $\sigma_{\phi,t}^{P,k}$ and $\sigma_{\theta,t}^{P,k}$ respectively. We assume that these measurements are uncorrelated. In camera frame's Cartesian coordinates these measurements are denoted as $\mathbf{y}_t^{P,k} \in \mathbb{R}^3$ and $\Sigma_t^{P,k} \in \mathbb{R}^{3 \times 3}$ respectively. The fused estimate of the person P 's position and uncertainty covariance in the world frame are denoted as \mathbf{x}_t^P and Σ_t^P . These estimates are computed by each MAV k fusing its own measurements, $\mathbf{y}_t^{P,k}$, with the measurements received from all other teammates ($\mathbf{y}_t^{P,j} \forall j = [1; j \neq k]$).

The MAVs operate in an environment with M known static obstacles (our approach is agnostic to how these obstacles are detected) and $K - 1$ neighboring MAVs as dynamic obstacles. The goal of each MAV is to cooperatively track the person using a replicated instance of the proposed formation control algorithm. This involves (a) minimizing the MAV's fused estimate Σ_t^P of the measurement uncertainty covariance and, (b) avoiding $M + K - 1$ environmental obstacles.

¹<https://github.com/AIRCAP/AIRCAP>

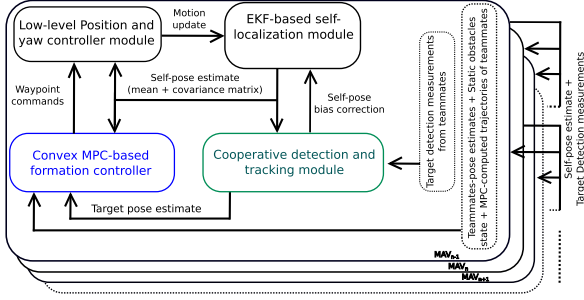


Fig. 3. Our aerial mocap front-end architecture. This letter focuses on the blue module.

B. Cooperative Detection and Tracking

Fig. 3 describes our mocap front-end architecture including i) the decentralized formation controller and ii) cooperative detection and tracking (CDT) modules. In [1], the focus was on developing the CDT algorithm (green block in Fig. 3). The work in the current paper focuses on developing the formation controller (blue block in Fig. 3), which utilizes information from CDT and other modules. In CDT, each MAV runs its own instance of an EKF for fusing detections $\in \mathbb{R}^3$ made by it and all its teammates. This results in a replicated state estimate of the person at each MAV, which is used to predict a region of interest (ROI) in future camera frames. The ROI guides the yaw of each MAV thereby ensuring that the tracked person is in the MAV's field of view. In Sec. III-D, we use these state estimates to drive our decentralized formation controller. At every time step, the controller of MAV k generates way-point positions and velocities for itself using (a) estimated state of the tracked person, (b) horizon motion plans communicated by teammates and (c) the positions of obstacles. The way-points guide MAV k 's low-level position and yaw controller.

C. Measurement Models and Joint Uncertainty Minimization

1) *Measurement Model for a Single MAV*: For a MAV k in this sub-section we drop the superscripts and subscripts. In a supplementary multimedia file attached with this letter, we empirically show that (a) the variance of the noise in the range measurement evolves quadratically with distance to the person, i.e., $\sigma_r^2 = C_1 r^2$, and (b) the variance of the noise in both bearing ϕ and inclination θ measurements are approximately constant w.r.t. the distance to the person, which we represent as $\sigma_\phi^2 = C_2$ and $\sigma_\theta^2 = C_3$. C_1, C_2 and C_3 are positive constants specific to the system.

We now convert the measurement noise variances from spherical to Cartesian coordinates. When using converted measurements (especially when measurement noise covariance is explicitly required) and performing estimation in Cartesian coordinates, most previous works, e.g., [16], [17], consider the measurement-conditioned conversion approach as derived in [18] or [19] as one of the best approximations. Using this conversion from [18] or [19] the noise covariances in Cartesian coordinates, i.e., the elements of $\Sigma_t^{P,k}$ can be written as

$$\begin{aligned}\sigma_x^2 &= A_1 f_1(\sigma_\theta^2, \sigma_\phi^2) g_1(\theta, \phi) r^2 + B_1 h_1(\theta, \phi, \sigma_\theta^2, \sigma_\phi^2) (r^2 + \sigma_r^2), \\ \sigma_y^2 &= A_2 f_2(\sigma_\theta^2, \sigma_\phi^2) g_2(\theta, \phi) r^2 + B_2 h_2(\theta, \phi, \sigma_\theta^2, \sigma_\phi^2) (r^2 + \sigma_r^2) \\ \sigma_z^2 &= A_3 f_3(\sigma_\theta^2) r^2 + B_3 h_3(\theta, \sigma_\theta^2) (r^2 + \sigma_r^2)\end{aligned}$$

$$\begin{aligned}\sigma_{xy}^2 &= A_4 f_4(\sigma_\theta^2, \sigma_\phi^2) g_4(\theta, \phi) r^2 + B_4 h_4(\theta, \phi, \sigma_\theta^2, \sigma_\phi^2) (r^2 + \sigma_r^2), \\ \sigma_{xz}^2 &= A_5 f_5(\sigma_\theta^2, \sigma_\phi^2) g_5(\theta, \phi) r^2 + B_5 h_5(\theta, \phi, \sigma_\theta^2, \sigma_\phi^2) (r^2 + \sigma_r^2), \\ \sigma_{yz}^2 &= A_6 f_6(\sigma_\theta^2, \sigma_\phi^2) g_6(\theta, \phi) r^2 + B_6 h_6(\theta, \phi, \sigma_\theta^2, \sigma_\phi^2) (r^2 + \sigma_r^2),\end{aligned}\quad (1)$$

where $A_1 \dots A_6$ and $B_1 \dots B_6$ are numerical constants, $f_1(\dots) f_6(\dots)$ are exponential functions of $\sigma_\theta^2, \sigma_\phi^2$, $g_1(\dots) g_6(\dots)$ are trigonometric functions of θ, ϕ , and $h_1(\dots) h_6(\dots)$ are functions containing trigonometric and exponential expressions involving θ, ϕ and $\sigma_\theta^2, \sigma_\phi^2$, respectively. Exact functions can be obtained from equations (10) of [18] or pp. 414, 417 of [19].

2) *Assumptions in Our Scenario*: Restoring the superscripts and subscripts for the spherical and Cartesian coordinate person measurements, we make the following assumptions on the MAV observation model.

Bearing Measurement: We employ a separate yaw controller [1] to guide the yaw angle of MAVs towards the tracked person. This is to ensure that the MAV's camera is always oriented towards the tracked person. Due to this, we can assume that the measurement component $\phi_t^{P,k}$ is almost always approximately zero ($\phi_t^{P,k} \approx 0$).

Inclination Measurement: We assume that the person's inclination measurement $\theta_t^{P,k}$ remains close to $\pi/2$ ($\theta_t^{P,k} \approx \pi/2$) as observed in Fig. 2. The rationale for this assumption is as follows. Due to our formation controller (as described in subsequent sub-sections) the MAVs track and follow the person while maintaining a desired altitude and distance to her/him. Consequently, the change in $\theta_t^{P,k}$ is negligible. Hence, we assume $\theta_t^{P,k} \approx \pi/2$.

Substituting these approximate values for $\phi_t^{P,k}$, $\theta_t^{P,k}$ and $C_1 (r_t^{P,k})^2$ for σ_r^2 (see beginning of this section for the latter substitution) in (1) and using expressions from either [18] or [19], the elements of $\Sigma_t^{P,k}$ are obtained in a compact form as,

$$\begin{aligned}\sigma_x^2 &= C_x (r_t^{P,k})^2, \quad \sigma_y^2 = C_y (r_t^{P,k})^2, \quad \sigma_z^2 = C_z (r_t^{P,k})^2, \\ \sigma_{xy}^2 &= \sigma_{xz}^2 = \sigma_{yz}^2 = 0,\end{aligned}\quad (2)$$

where C_x, C_y, C_z are constants involving exponential functions of constants C_1, C_2 and C_3 only.

3) *Minimizing Fused Uncertainty for MAV Formations*: The minimum uncertainty in the person's fused state estimate is achieved when the following conditions are met.

a) *The angles between the MAVs about the person's estimated position are $\frac{2\pi}{K}$* : Given measurements from observations of different robots (or sensors), [20] analytically derives optimal sensor geometries for unbiased and efficient passive target localization. Geometries that maximize the determinant of Fisher information matrix are identified. For $K(\geq 3)$ independent measurements, the angle between sensors resulting in minimum fused uncertainty is either $\frac{2\pi}{K}$ or $\frac{\pi}{K}$. We validated this analysis by computing Σ_t^P from $\Sigma_t^{P,k}$ for all possible angular configurations and using the recursive covariance merging equations from [21]. Fig. 4 illustrates some of the possible angular configurations. Out of these, Fig. 4(a), (b) are examples of configurations that do not minimize Σ_t^P . Fig. 4(c) is one of those configurations which minimizes Σ_t^P and which we use in this work. It is preferred over other optimal configurations (see Fig. 4 in [22]) because others do not ensure a uniformly distributed visual coverage of the tracked subject. This is important as our system is motivated

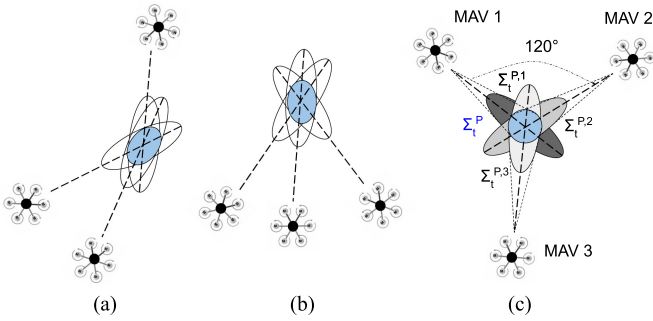


Fig. 4. (a), (b) represent arbitrary configurations with large joint uncertainties. (c) is the optimal configuration for 3 MAVs used in this work.

by outdoor motion capture of an asymmetric subject (at least in one axis).

b) The measurement uncertainty for each individual MAV is minimized: The previous angular constraint condition minimizes the fused uncertainty for any given set of measurements from the MAVs. However, the fused covariance remains a function of each individual MAV's measurement uncertainty, $\Sigma_t^{P,k} \forall k \in [1 \dots K]$. $\Sigma_t^{P,k}$ is a function of the relative position of MAV k w.r.t. the tracked person. Notice that each MAV's position remains controllable without changing the angular configuration of the whole formation. Therefore, the fused uncertainty is completely minimized by minimizing the trace of $\Sigma_t^{P,k}$ for each MAV k . How this is done is discussed in the following subsection.

D. Decentralized Quadratic Model Predictive Control

The trace of $\Sigma_t^{P,k}$ obtained using (2) is as shown below.

$$\text{tr}(\Sigma_t^{P,k}) = (C_1 + C_2 + C_3)(r_t^{P,k})^2 = \kappa(r_t^{P,k})^2, \quad (3)$$

where $\kappa = C_1 + C_2 + C_3$. Minimizing (3) ensures that each MAV improves its measurement uncertainty. However, $r_t^{P,k} \approx 0$ would cause the MAVs to collide with the tracked person and other MAVs. In order to ensure the safety of the person, we limit the MAV to reach a desired circular safety surface centered at the person with a radius d_{des} at a desired altitude h_{des} from her/him. Hence, instead of minimizing (3), which is equivalent to minimizing $(\mathbf{x}_t^k - \mathbf{x}_t^{P,k})^T \kappa \mathbf{I}_3 (\mathbf{x}_t^k - \mathbf{x}_t^{P,k})$, we now minimize $(\mathbf{x}_t^k - \tilde{\mathbf{x}}_t^k)^T \kappa \mathbf{I}_3 (\mathbf{x}_t^k - \tilde{\mathbf{x}}_t^k)$. The latter expression is in terms of the MAV's position (\mathbf{x}_t^k) and a desired position ($\tilde{\mathbf{x}}_t^k$) chosen on the aforementioned safety surface and lying in the direction from the MAV k to the tracked person. Here, \mathbf{I}_3 is a 3×3 identity matrix. We now define the objective of our MPC running on each MAV k as follows. Minimize,

- the distance to the desired target surface (safety surface) around the tracked person,
- the difference in velocity between the MAV and the fused velocity estimate of the tracked person and,
- MAV control effort.

Consequently, the optimization objective of our active tracking MPC is as given below.

$$J_{ACT} = \sum_{n=0}^N (\mathbf{u}_t^k(n)^T \mathbf{W}_E \mathbf{u}_t^k(n)) + \mathbf{X}_t^k{}^T \begin{bmatrix} \kappa \mathbf{I}_3 & 0 \\ 0 & \mathbf{W}_{\dot{\mathbf{x}}} \end{bmatrix} \mathbf{X}_t^k \quad (4)$$

where $\mathbf{X}_t^k = [\mathbf{x}_t^k(N+1)^T \dot{\mathbf{x}}_t^k(N+1)^T]^T - [(\tilde{\mathbf{x}}_t^k)^T (\dot{\tilde{\mathbf{x}}}_t^k)^T]^T$. The first part of this objective minimizes the input control signal and the second part ensures that the distance between desired terminal state (position and velocity) $[\tilde{\mathbf{x}}_t^k \dot{\tilde{\mathbf{x}}}_t^k]^T$ and the final horizon robot state $[\mathbf{x}_t^k(N+1) \dot{\mathbf{x}}_t^k(N+1)]^T$ is minimized. In order to ensure continuous and smooth person tracking, we minimize the difference in velocity of the MAV ($\dot{\mathbf{x}}_t^k$) and the fused velocity estimate of the person ($\dot{\tilde{\mathbf{x}}}_t^k$) available from the EKF. κ is as defined in (3) and is experimentally determined. $\mathbf{W}_E \in \mathbb{R}^{3 \times 3}$ and $\mathbf{W}_{\dot{\mathbf{x}}} \in \mathbb{R}^{3 \times 3}$ are custom-defined diagonal positive semi-definite weight matrices. Furthermore, the active tracking MPC is subjected to the following constraints.

- A non-convex constraint of maintaining a desired angular difference of $\frac{2\pi}{K}$ w.r.t. other MAVs about the fused estimate of the person's position $\tilde{\mathbf{x}}_t^P$.
- A non-convex constraint to maintain at least a distance d_{min} from other obstacles.
- Regional and control saturation bounds for MAV states and control inputs for a horizon.

The non-convex constraints are embedded into the MPC dynamics as repulsive potential field forces \mathbf{f}_t^k [15]. The computation of these forces is discussed in Sec. III-E. The convex MPC is thus defined by the following equations.

$$\begin{aligned} & \mathbf{x}_t(1)^{k*} \dots \mathbf{x}_t(N+1)^{k*}, \mathbf{u}_t^{k*}(0) \dots \mathbf{u}_t^{k*}(N) \\ & = \arg \min_{\mathbf{u}_t^k(0) \dots \mathbf{u}_t^k(N)} (J_{ACT}) \end{aligned} \quad (5)$$

$$\begin{aligned} & \text{s.t } [\mathbf{x}_t^k(n+1)^T \dot{\mathbf{x}}_t^k(n+1)^T]^T = \\ & \mathbf{A} [\mathbf{x}_t^k(n)^T \dot{\mathbf{x}}_t^k(n)^T]^T + \mathbf{B}(\mathbf{u}_t^k(n) + \mathbf{f}_t^k(n) + \mathbf{g}), \end{aligned} \quad (6)$$

$$[\mathbf{x}_{min}^T \dot{\mathbf{x}}_{min}^T] \leq [\mathbf{x}_t^k(n)^T \dot{\mathbf{x}}_t^k(n)^T] \leq [\mathbf{x}_{max}^T \dot{\mathbf{x}}_{max}^T], \quad (7)$$

$$\mathbf{u}_{min} \leq \mathbf{u}_t^k(n) \leq \mathbf{u}_{max} \quad (8)$$

The 3D translation motion of MAV k is governed partly using accelerations $\mathbf{u}_t^k(n) = \ddot{\mathbf{x}}_t^k(n)$ and partly by an external control input $\mathbf{f}_t^k(n)$, where n is the horizon time step. The state-space translational dynamics of the robot is given by (6). Dynamics ($\mathbf{A} \in \mathbb{R}^{6 \times 6}$) and control transfer ($\mathbf{B} \in \mathbb{R}^{6 \times 3}$) matrices are given by

$$\mathbf{A} = \begin{bmatrix} \mathbf{I}_3 & \Delta t \mathbf{I}_3 \\ \mathbf{0}_3 & \mathbf{I}_3 \end{bmatrix}, \mathbf{B} = \begin{bmatrix} \frac{\Delta t^2}{2} \mathbf{I}_3 \\ \Delta t \mathbf{I}_3 \end{bmatrix}, \quad (9)$$

where Δt is the sampling time. $\mathbf{f}_t^k(n)$ is a real-time computed external control input representing non-convex constraints of formation configuration and obstacle avoidance. The next section details the computation of these forces.

Note that in our formulation optimality is regarding the minimization of uncertainty in the person's 3D position jointly estimated by all MAVs. The way we formulate our MPC it cannot optimize the trajectories over the full belief space and hence transient behavior is not explicitly addressed. In our case, however, the MPC prediction time horizons are small enough (15 time steps or 1.5 seconds in total) to neglect the transient effects, as long as the MAVs are not very far away from the tracked person. Moreover, in the MPC we also minimize the difference between target person velocity and the MAV velocities, which further facilitates convergence to the desired view points.

Also, the optimal configuration may not always be feasible, especially, when the obstacles completely block or occlude the

desired safety surface around the tracked person. This is inherent to our formulation because we enforce that obstacle avoidance and feasibility of trajectories take precedence (by considering them as hard constraints in MPC) over the optimality of the view points. This is important for the safety of aerial systems.

E. Computation of External Control Input ($\mathbf{f}_t^k(n)$)

Inter-MAV angular configuration constraints involve inverse-tangent operations with positive or negative arguments, making the constraint non-convex. Moreover, avoiding collisions by maintaining a minimum distance w.r.t. other MAVs and obstacles in the environment, constraint the robot to operate in non-convex regions. In our previous work [15], we introduced the concept of converting non-convex constraints into external control inputs $\mathbf{f}_t^k(n)$ using repulsive potential fields. In this work, we use it to enforce formation angular configuration and collision avoidance constraints. The repulsive potential field function used in this work is referred to as cotangential field function (see (17) in [22]). The magnitude of the cotangential field force is a hyperbolic function denoted as $F^{k,j}(d)$. Argument d is either the Euclidean distance or the absolute angular difference between MAV k and entity j (another MAV or obstacle). Let d_{\max} and d_{\min} be threshold radii, where d_{\max} defines the region of influence of the potential field. d_{\min} is the distance below which the potential field value tends to infinity. Practically, this is clamped to a large positive magnitude F_{\max} which is derived to guarantee obstacle avoidance (see Sec. III-E3).

1) *Active Tracking Input*: To satisfy the inter-MAV angular configuration constraints, each MAV k computes its angle about the tracked person in the world frame ($\gamma_t^k(n)$) and the angle of neighboring MAV j about the tracked person ($\gamma_t^j(n)$). The absolute angular difference defined as $d_{act}(n) = |\gamma_t^k(n) - \gamma_t^j(n)|$, is used as the argument to compute the potential field magnitude. We compute for MAV k and teammate MAV j ,

$$F_{act}^{k,j}(n) = (|r_{des} - \|\mathbf{x}_t^k(n) - \mathbf{x}_t^P\|_2| + c) F^{k,j}(d_{act}(n)) \quad \forall j \quad (10)$$

The factor $|r_{des} - \|\mathbf{x}_t^k(n) - \mathbf{x}_t^P\|_2|$ in (10) helps avoid field local minima. This is because, if two robots have similar angles of approach (i.e., small $d_{act}(n)$), the MAV farther away from the desired distance (r_{des}) is repelled with a higher force than the MAV near the desired distance (see [15] for a detailed explanation on how these forces avoid field local-minima problems associated with potential field based planners). Factor c is a small positive constant which ensures that the force magnitude is non-zero at the target surface if the desired angular difference is not yet achieved. $F_{act}^{k,j}(n)$ acts in a direction $\alpha_{\perp}^{k,j}$ which is normal to the direction of approach to the tracked person $\alpha^{k,j} = \frac{\mathbf{x}_t^k(n) - \mathbf{x}_t^P}{\|\mathbf{x}_t^k(n) - \mathbf{x}_t^P\|_2}$. In the plane of approach towards the person, there are two direction choices for a MAV, namely, $\pm\alpha_{\perp}^{k,j}$. The direction pointing away from neighboring MAV j 's position is chosen. This choice ensures a natural deadlock resolution for robots having similar angles of approach to the tracked person. The total active tracking external control input of MAV k w.r.t. all other teammate MAVs is

$$\mathbf{f}_{act}^k(n) = \sum_j F_{act}^{k,j}(n) \alpha_{\perp}^{k,j} \quad (11)$$

For the active tracking control input $d_{\max} = (\frac{2\pi}{K})$.

2) Obstacle Avoidance Input.

Dynamic Obstacle Avoidance: All the teammates of MAV k are considered as dynamic obstacles. The argument d for $F^{k,j}(d)$ in case of dynamic obstacle avoidance is the euclidean distance between MAV k and teammate MAV j . It is defined as $d_{dyn}(n) = \|\mathbf{x}_t^k(n) - \mathbf{x}_t^j(n)\|_2, \forall n > 0$. Magnitude of $F^{k,j}(d)$ is computed using MAV k and teammate j 's horizon motion plans. This is enforced along the direction $\beta^{k,j} = \frac{\mathbf{x}_t^k(n) - \mathbf{x}_t^j(n)}{\|\mathbf{x}_t^k(n) - \mathbf{x}_t^j(n)\|_2}$, which is a unit vector pointing in the direction away from the teammate MAV j 's horizon motion plan. The total dynamic obstacle avoidance external input $\mathbf{f}_{dyn}^k(n)$ is as follows.

$$\mathbf{f}_{dyn}^k(n) = \sum_j F^{k,j}(d_{dyn}(n)) \beta^{k,j} \quad (12)$$

The choice for d_{\max} and d_{\min} for dynamic obstacle avoidance are discussed in the following sub-Section III-E3.

Static Obstacle Avoidance: The total external control input due to the M static obstacles is as given below.

$$\mathbf{f}_{sta}^k(n) = \sum_j F^{k,m}(d_{sta}(n)) \beta^{k,m}, \quad (13)$$

where $d_{sta}(n) = \|\mathbf{x}_t^k(n) - \mathbf{x}_t^m\|_2$. The tracked person is considered as an additional static obstacle for each MAV. An external control input $\mathbf{f}_{target}^k(n)$ is computed so as to enforce her/his safety. Further, multiple static obstacles placed closed together might cause the MAV k to get stuck in a field local minima (as shown in [15]). To avoid these scenarios, we compute a force $\mathbf{f}_{ang}^k(n)$ to penalize MAV approach angles ($\gamma_t^k(n)$) which have static obstacles in the line of approach. The computation of this force is similar to the active tracking force described in the previous section. The total external control input for obstacle avoidance $\mathbf{f}_{obs}^k(n)$ is given by the summation $\mathbf{f}_{obs}^k(n) = \mathbf{f}_{dyn}^k(n) + \mathbf{f}_{sta}^k(n) + \mathbf{f}_{target}^k(n) + \mathbf{f}_{ang}^k(n), \forall n$

Subsequently, the total external control input $\mathbf{f}_t^k(n)$ acting on MAV k is determined as follows.

$$\mathbf{f}_t^k(n) = \begin{cases} \mathbf{f}_{act}^k(n) + \mathbf{f}_{obs}^k(n) & \text{if } \|\mathbf{f}_{act}^k(n) + \mathbf{f}_{obs}^k(n)\|_2 < F_{\max} \\ F_{\max} \frac{\mathbf{f}_{act}^k(n) + \mathbf{f}_{obs}^k(n)}{\|\mathbf{f}_{act}^k(n) + \mathbf{f}_{obs}^k(n)\|_2} & \text{if } \|\mathbf{f}_{act}^k(n) + \mathbf{f}_{obs}^k(n)\|_2 \geq F_{\max} \end{cases} \quad (14)$$

Next, we describe the considerations for choosing d_{\min} , d_{\max} and F_{\max} to guarantee obstacle avoidance.

3) *Obstacle Avoidance Guarantee*: To ensure stable external control inputs and guarantee obstacle avoidance, we take into account the following important considerations.

- maximum tracking error (e_{\max}) of low-level controller w.r.t generated way-point (\mathbf{x}_t^k) for one time step Δt ,
 - maximum magnitude of the MAV velocity ($\|\dot{\mathbf{x}}_{\max}\|_2$),
 - self-localization uncertainty of the MAV (Σ_t^k),
 - localization uncertainty of teammate MAVs (Σ_t^j),
 - communication or packet loss with teammate MAVs.
 - maximum cotangential force magnitude $F_{\max} \geq \|\mathbf{u}_{\max}\|_2$.
- d_{\min} and d_{\max} corresponding to the external control input for obstacle avoidance are defined as follows.

$$d_{\min} = e_{\max} + \|\dot{\mathbf{x}}_{\max}\|_2 \Delta t, \quad d_{\max} = r_{\sigma_t}^k + r_{\sigma}^j + r_{comm,t}^{k,j} \quad (15)$$

where, $r_{\sigma_t}^k$ is the maximum eigenvalue of Σ_t^k . r_{σ}^j is the experimentally determined maximum possible eigenvalue of the localization uncertainty (Σ_t^j) of the teammate MAV j . Wireless communication delays are unavoidable in real-time implementations and cannot be ignored as they affect the MAVs knowledge of teammate MAV positions. $r_{comm,t}^{k,j}$ accounts for it by increasing in proportion with the communication delay between the MAVs. $r_{comm,t}^{k,j} \propto \|\dot{\mathbf{x}}_{max}\|_2 \Delta t$, which is the maximum distance a neighboring MAV can travel in Δt seconds.

IV. EXPERIMENTS AND RESULTS

A. Real Robot Experiments

1) *Setup*: We use three self-developed and customized octocopter MAVs in the field experiments. The MAVs are equipped with an Intel i7 CPU, an NVidia Jetson TX1 GPU and Open Pilot Revolution flight controller system on chip. Each MAV runs the CDT algorithm (Fig. III-B) on-board and in real-time, including the deep neural network-based single shot detector (SSD-Multibox). The MAVs use on-board GPS and IMU data, acquired at 100 Hz, for localization. We place Emlid Reach differential GPS receiver on each MAV and on either shoulder of the tracked person to acquire ground truth (GT) location. The data from the differential GPS is not used during real-time experiments and stored only for comparison. For a decentralized implementation, a ROS multi-master setup with communication over Wifi is used. Each MAV continuously captures and stores camera images acquired at 40 Hz. The MPC is evaluated at 100 Hz using the CVXGEN convex optimization library. The horizon length is $N = 15$ time steps corresponding to 1.5 seconds. The state and velocity limits of each MAV is $[-20, -20, 3] \leq \mathbf{x}_t^k(n) \leq [20, 20, 10]$ in m and $[-5, -5, -0.5] \leq \dot{\mathbf{x}}_t^k(n) \leq [5, 5, 0.5]$ in m/s . The control limits are defined as $[-3, -3, 3] \leq \mathbf{u}_t^k(n) \leq [3, 3, 11]$. The desired horizontal distance to the tracked person in all experiments is set as $d_{des} = 8$ m and the desired height as $h_{des} = 8$ m. The tracking error w.r.t. to the MPC generated way-point is $\epsilon_{max} \approx 1$ m, therefore we use $d_{min} = 1.5$ m (see subsection III-E3). d_{max} w.r.t. teammate MAVs varies with the change in self localization uncertainty and communication losses. The task for each MAV in all experiments is to track and follow the person using the proposed approach. We conducted 5 real-MAV experiments (RE), each with a 3-MAV formation as follows.

- **RE 1**: Approach of Price *et al.* [1] (duration 257s)
- **RE 2**: DQMPC (Tallamraju *et al.*) [15] (231s)
- **RE 3**: Our approach (269s)
- **RE 4**: Our approach with one MAV avoiding emulated virtual obstacles (278s)
- **RE 5**: Our approach with all MAVs avoiding emulated virtual obstacles (106s)

In each experiment, the person remains stationary at first. Data collection starts after all MAV's have acquired the visual line of sight to the person and converged on a stable formation around him/her. The person then walks randomly at moderate speeds, runs or performs standard exercise movements (see attached video or here² for details). In the end, MAVs are manually landed. The error in the tracked position estimate of the person is calculated as the 3D Euclidean distance between the estimated and corresponding GT provided by the

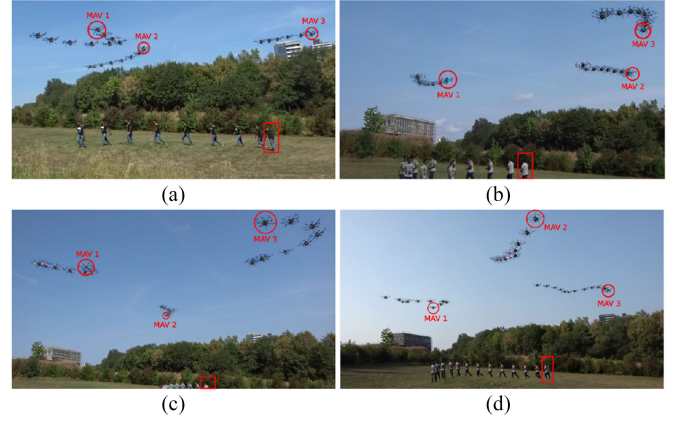


Fig. 5. Multi-exposure images of short sequences from (RE 1 to RE 4). (a) showcases the results based on our previous work [1]. Notice that the MAVs are close to the target person and never uniformly spread around the person's position. (b) shows the results based on our previous work [15]. Notice that the MAVs 2 and 3 are quite close to each other and the resulting formation is non-optimal for uncertainty minimization. In (c), (d) and the image in Fig. 1 the results of the approach presented in the current paper is showcased. These correspond to experiments RE 3 to RE 5, respectively. Notice that the MAVs are almost uniformly spread around the person's position and maintain an angular configuration with a difference of approx. $\frac{2\pi}{3}$ w.r.t. each other. Moreover, the MAVs successfully maintain the desired safe distance and altitude from the person. This configuration ensured minimization of fused uncertainty of the person's position in RE 3–RE 5. The experimental footage is provided in the accompanying video.

differential GPS system. Error for MAV pose is calculated similarly.

2) *Analysis of Results*: Images in Fig. 5 and Fig. 1 show short sequences from RE 1 – RE 5. Fig. 6(e) compares the corresponding GT with the tracked person's position estimate obtained by one of the MAVs using our proposed approach in RE 3. We analyzed the estimate's accuracy, uncertainty, and the MAV self-pose accuracy for all experiments in Fig. 6(a–b). RE 3 achieves significantly more accurate person track estimates (mean error of ~ 0.7 m) than using state of the art method in RE 1 (mean error of ~ 1.2 m), despite the worse self pose estimates of the MAVs in RE 3 (the self-localization errors were due to the high errors in the MAV's GPS localization and may vary arbitrarily over different experiments). This showcases the ability of our approach to attain high tracking accuracy despite bad GPS reception.

Even though the state of the art approach in RE 2 achieves similar accuracy to that of our approach in RE 3, we see that our approach outperforms the former in keeping the person always in the camera image and close to the image center as shown by bar and box plots in Fig. 6(c–d). Moreover, our approach achieves least tracking uncertainty in the tracked person estimate as compared to the other two state-of-the-art methods. The presence of obstacles that the MAVs need to navigate around affects the target tracking and self pose accuracy in RE 4 and RE 5. This is mainly due to the additional maneuvering overhead to avoid obstacles. However, the ability to keep the person close to the center of the image is only slightly affected. Fig 6(c–d) compares the different approaches ability to keep the tracked person centered in each MAV's camera view. Our approach in RE 3 not only reduces the average distance between the tracked person and each camera's image center but also ensures the person is completely covered by the camera image in 97.5%

²<https://youtu.be/0Al3MlW0R1I>

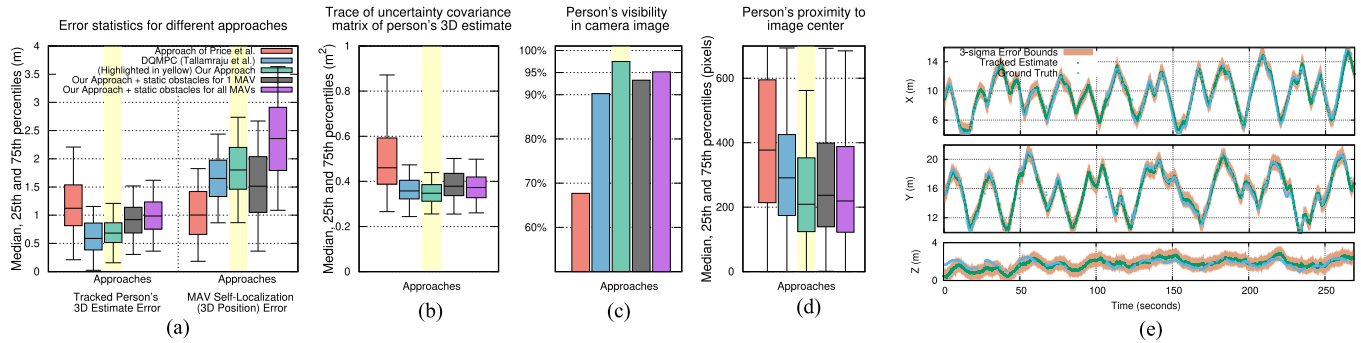


Fig. 6. Comparison of different methods in real world experiments: (a) accuracy of the tracked person and MAV self-pose estimates; (b) uncertainty in the tracked person estimate; (c) fraction of total experiment duration (jointly for all MAVs) the tracked person was completely in the camera image frame; (d) distribution of distance of tracked person from image center. (e) Real world experiment **RE 3**: Tracked person's estimated (blue) and ground truth (green) trajectory comparison of our active perception approach (without simulated stationary obstacles). Ground truth remains within the 3σ error bound of tracked estimate throughout the experiment duration except for a few initial seconds.

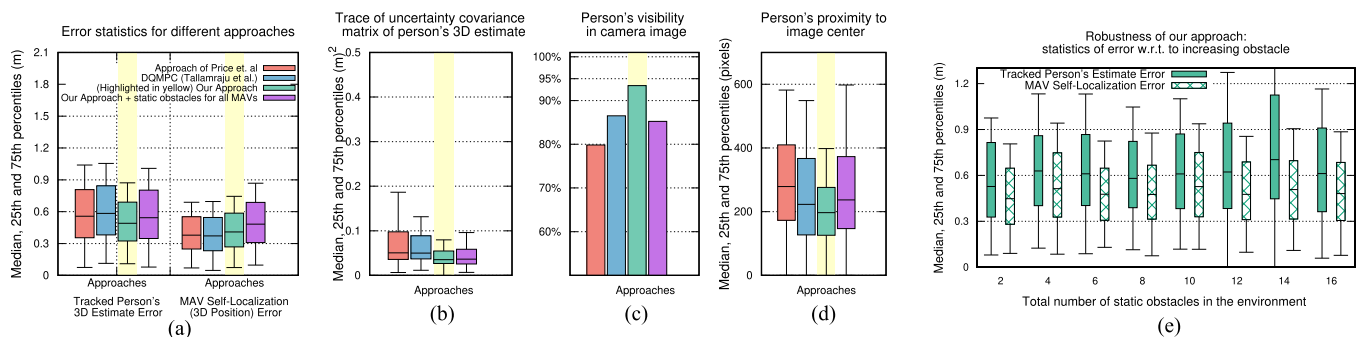


Fig. 7. Statistics of Simulation Experiment: (a)–(d) please refer caption of Fig. 6 for description of these plots; (e) robustness of our approach with increasing static obstacles.

of camera frames. This is a crucial feature for a system designed for aerial motion capture.

B. Simulation Experiments

1) *Setup*: The proposed algorithm was simulated in the Gazebo. The simulations were conducted on a standalone Intel i7-3970X CPU with an NVIDIA GTX 1080 GPU. We used AscTec Hexarotor Firefly MAVs in a world of $20\text{ m} \times 20\text{ m} \times 20\text{ m}$. Each MAV has a rigidly attached Asus Xtion camera with their parameters set to the real MAV camera parameters. We simulate GPS and IMU drift by imposing random-walk offset on the ground truth position. We simulate a human actor model and guide it along a pre-defined trajectory. The actor traverses on a randomly varying terrain interspersed with emulated trees. The trees create obstacles and occlusions for the hovering MAVs.

2) *Comparison of Methods*: For each of the following methods, we conducted 30 simulation runs with a 5 MAV formation: (i) **SE 1**: Approach of Price *et al.* [1], (ii) **SE 2**: DQMPC (Tallamraju *et al.*) [15], (iii) **SE 3**: Our approach, and (iv) **SE 4**: Our approach with all MAVs avoiding emulated virtual obstacles. First 10 runs were of 120 seconds each, and next 20 runs were 180 seconds each. Fig. 7 shows the statistics of this set of experiments. We clearly observe that our proposed approach with or without static obstacle avoidance outperforms both the state of the art approaches. Especially, when static obstacles are not present, our approach is able to keep the person fully in the camera image frame almost 94% of the

experiment duration. Also, it significantly outperforms all other methods in minimizing the joint uncertainty in the target position estimate.

3) *Scalability Experiments*: Fig. 8(a) shows the result of target tracking with our method for an increasing number of robots. We conduct 10 simulation trials for each number of robot configuration. Real-time factor for all runs in this experiment was set to 0.1. We validate the scalability of our model predictive controller by running these experiments in an environment with several static and known obstacles (in the form of trees) and perfect communication for a formation of 3 to 16 MAVs. We observe nearly linear improvement in tracking error with a higher number of robots. At the same time, we notice that computational requirements did not affect real-time performance for up to 16 robots.

4) *Experiments With Communication Failure*: In Fig. 8(b), the effect of inter-robot communication failure on tracking is demonstrated for our approach through experiments with 8 MAVs and simulated static obstacles. Communication losses varying from 10% to 100% is simulated. The results were averaged over 3 trials per communication loss percentage. It can be observed that the tracking gets progressively worse with a higher percentage of communication loss. At 100% communication loss, each robot relies only on its detection and does not cooperatively improve the target position estimate. Nevertheless, we observe that our approach is able to maintain an accurate target state estimate for up to 25% communication loss.

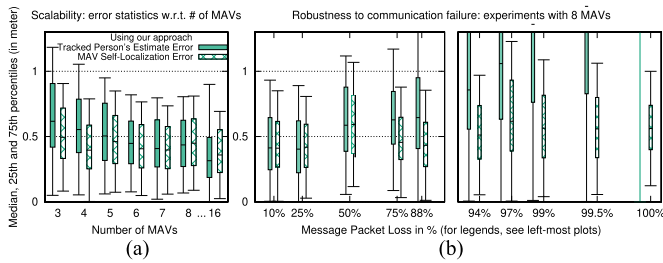


Fig. 8. Our method's scalability and robustness.

5) *Experiments With Increasing Number of Obstacles:* Fig. 7(e) shows the results of our approach with increasing number of environmental obstacles. For each map, we conducted 5 trials (180 seconds each) with 5 MAVs tracking a randomly walking human model. We observe that by increasing the number of obstacles, the tracking error only slightly deteriorates. Moreover, the error in tracking is close to that of the environment with no static obstacles (see Fig. 7(a)). This result further validates our approach as the robots attain the desired configuration around the target person by navigating around randomly placed environmental obstacles.

V. CONCLUSIONS AND FUTURE WORK

We proposed a decentralized convex MPC-based algorithm for the MAVs to actively track and follow a moving person in outdoor environments and in the presence of static and dynamic obstacles. In contrast to cooperatively tracking only the 3D positions of the person, the MAVs actively compute optimal local motion plans, resulting in optimal view-point configurations, which minimize the uncertainty in the tracked estimate. We showed how we embed all the non-convex constraints, including those for dynamic and static obstacle avoidance, as external control inputs in the MPC dynamics. We evaluated our approach through rigorous and multiple real robot field experiments in an outdoor scenario. These experiments validate our approach and show that it significantly improves accuracy over the previous methods. In simulations, we showed results with up to 16 MAVs. These demonstrate the scalability of our method as well as its robustness to communication failures. Future work includes addressing visibility issues, e.g., occlusions of the person and asymmetry of the target, which are not explicitly solved by our current approach. We also intend to extend our method to full body human pose detection.

ACKNOWLEDGMENT

The authors would like to thank I. Martinovic and M. Landry for their help with the field experiments and the anonymous reviewers and the editor for their highly constructive feedback.

REFERENCES

- [1] E. Price *et al.*, "Deep neural network-based cooperative visual tracking through multiple micro aerial vehicles," *IEEE Robot. Autom. Lett.*, vol. 3, no. 4, pp. 3193–3200, Oct. 2018.
- [2] C. Huang *et al.*, "ACT: An autonomous drone cinematography system for action scenes," in *Proc. IEEE Int. Conf. Robot. Autom.*, 2018, pp. 7039–7046.
- [3] X. Zhou, S. Liu, G. Pavlakos, V. Kumar, and K. Daniilidis, "Human motion capture using a drone," in *Proc. IEEE Int. Conf. Robot. Autom.*, Brisbane, Australia, 2018, pp. 2027–2033, doi: [10.1109/ICRA.2018.8462830](https://doi.org/10.1109/ICRA.2018.8462830).
- [4] L. Xu *et al.*, "FlyCap: Markerless motion capture using multiple autonomous flying cameras," *IEEE Trans. Visualization Comput. Graph.*, vol. 24, no. 8, pp. 2284–2297, Aug. 2018.
- [5] T. Nägele, S. Oberholzer, S. Plüss, J. Alonso-Mora, and O. Hilliges, "Flycon: real-time environment-independent multi-view human pose estimation with aerial vehicles," in *Proc. SIGGRAPH Asia 2018 Tech. Papers*, ACM, 2018, Art. no. 182.
- [6] F. Morbidi and G. L. Mariottini, "On active target tracking and cooperative localization for multiple aerial vehicles," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2011, pp. 2229–2234.
- [7] K. Zhou and S. I. Roumeliotis, "Multirobot active target tracking with combinations of relative observations," *IEEE Trans. Robot.*, vol. 27, no. 4, pp. 678–695, Aug. 2011.
- [8] K. Hausman, J. Müller, A. Hariharan, N. Ayanian, and G. S. Sukhatme, "Cooperative multi-robot control for target tracking with onboard sensing," *Int. J. Robot. Res.*, vol. 34, no. 13, pp. 1660–1677, 2015.
- [9] K. Hausman *et al.*, "Occlusion-aware multi-robot 3d tracking," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2016, pp. 1863–1870.
- [10] P. U. Lima *et al.*, "Formation control driven by cooperative object tracking," *Robot. Auton. Syst.*, vol. 63, pp. 68–79, 2015.
- [11] D. Falanga, P. Foehn, P. Lu, and D. Scaramuzza, "PAMPC: Perception-aware model predictive control for quadrotors," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Madrid, 2018, pp. 1–8, doi: [10.1109/IROS.2018.8593739](https://doi.org/10.1109/IROS.2018.8593739).
- [12] F. Augugliaro, A. P. Schoellig, and R. D'Andrea, "Generation of collision-free trajectories for a quadcopter fleet: A sequential convex programming approach," in *Proc. IEEE Intell. Robots Syst.*, 2012, pp. 1917–1922.
- [13] J. Alonso-Mora *et al.*, "Collision avoidance for aerial vehicles in multi-agent scenarios," *Auton. Robots*, vol. 39, no. 1, pp. 101–121, 2015.
- [14] T. Nägele, L. Meier, A. Domahidi, J. Alonso-Mora, and O. Hilliges, "Real-time planning for automated multi-view drone cinematography," *ACM Trans. Graph.*, vol. 36, no. 4, 2017, Art. no. 132.
- [15] R. Tallamraju, S. Rajappa, M. J. Black, K. Karlapalem, and A. Ahmad, "Decentralized MPC based obstacle avoidance for multi-robot target tracking scenarios," in *Proc. IEEE Int. Symp. Saf., Secur. Rescue Robot.*, 2018, pp. 1–8.
- [16] Zhanlue Zhao, T. X. Rong Li, and V. P. Jilkov, "Best linear unbiased filtering with nonlinear measurements for target tracking," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 40, no. 4, pp. 1324–1336, Oct. 2004.
- [17] X. Rong Li and V. P. Jilkov, "Survey of maneuvering target tracking: III. Measurement models," in *Proc. SPIE 4473, Signal and Data Process. Small Targets*, Nov. 2001, pp. 423–446, doi: [10.1117/12.492752](https://doi.org/10.1117/12.492752).
- [18] M. Longbin, S. Xiaoquan, Z. Yiyu, S. Zhong Kang, and Y. Bar-Shalom, "Unbiased converted measurements for tracking," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 34, no. 3, pp. 1023–1027, Jul. 1998.
- [19] M. D. Miller and O. E. Drummond, "Coordinate transformation bias in target tracking," in *Proc. SPIE 3809, Signal and Data Process. Small Targets*, Oct. 1999, pp. 409–424, doi: [10.1117/12.364039](https://doi.org/10.1117/12.364039).
- [20] A. N. Bishop, B. Fidan, B. D. Anderson, K. Doğançay, and P. N. Pathirana, "Optimality analysis of sensor-target localization geometries," *Automatica*, vol. 46, no. 3, pp. 479–492, 2010.
- [21] R. C. Smith and P. Cheeseman, "On the representation and estimation of spatial uncertainty," *Int. J. Robot. Res.*, vol. 5, no. 4, pp. 56–68, 1986.
- [22] R. Tallamraju *et al.*, "Active perception based formation control for multiple aerial vehicles," 2019, *arXiv:1901.07813*.