2024 IEEE 63rd Conference on Decision and Control (CDC)
December 16-19, 2024. MiCo, Milan, Italy

Target Search and Tracking in an Unknown Environment by a Fleet of UAVs using a Set-Membership Approach

Maxime Zagar*, Luc Meyer**, Michel Kieffer* and Hélène Piet-Lahanier**

Abstract— **This paper addresses the problem of search and tracking an unknown number of mobile ground targets within an uncharted Region of Interest (RoI) using a fleet of cooperating Unmanned Aerial Vehicles (UAVs). Each UAV embeds a Computer Vision System providing images with labeled pixels, depth maps, and bounding boxes around identified targets. This information is used by a set-membership estimator to characterize sets guaranteed to contain the locations of already identified targets and a set containing the locations of all targets remaining to detect. A map of the unknown environment is constructed during search to favor exploration of areas previously occluded by obstacles.**

I. Introduction

The target search and tracking in some unknown Region of Interest (RoI) is an important problem which has given rise to a large research activity, see [1] and the references therein. To address this problem, fleets of cooperating UAVs collecting measurements from different point of views have been shown to be more efficient than a single UAV.

This paper presents a set-membership approach to search and track ground targets evolving in an unknown but structured environment. Each UAV embeds a Computer Vision System (CVS) providing images with labeled pixels, depth maps, and bounding boxes around identified targets. These CVS outputs are directly exploited, contrary to previous set-membership approaches [2], where simplified sensing models have been considered. This requires assumptions and models to relate the CVS outputs with the environment to allow their exploitation by a set-membership estimator [3]. One obtains sets guaranteed to contain the locations of identified targets and a set containing all possible locations of not yet detected targets. A set proved to contain no target is also provided. Based on this map and on the sets previously calculated, the trajectories of the fleet are then designed using a Model Predictive Control (MPC) approach [2] to minimize the target location uncertainties. Compared to [3], this paper exploits an occupancy elevation map (OEM) [4] to obtain a rough characterization of the obstacles. This map is constructed during the target search. The guidance law of the UAVs uses the OEM to predict the parts of the RoI which may be occluded by obstacles. UAVs are then driven into areas previously occluded by obstacles. This improves the search and tracking performance as illustrated in the simulations.

Some related work is introduced in Section II. Section III formulates the target search and tracking problem. Section IV introduces some hypotheses and models of the targets, the UAVs, and the CVS. Section V describes the way CVS information can be used in the set-membership estimator described in Section VI. Section VII presents the evaluation of the guidance law of each UAV. Simulation results are provided in Section VIII before the conclusion in Section IX.

II. Related work

Various estimation techniques for the localization of targets have been proposed such as stochastic approaches with grid-based probability map [5] or probability hypothesis density filter with random finite sets [6], or with alternative set-membership techniques [2] assuming bounded measurement noise. When the RoI is cluttered, trajectories which avoid collisions while maximizing the information collected have to be designed [7]. Without available map, environment mapping and target search and tracking have to be done simultaneously as in [8], where the RoI is mapped with an occupancy grid. Nevertheless, accurate 3D maps of some unknown RoI with an occupancy grid raises storage and computation issues for UAVs [9]. In [4], an OEM is used. It requires less storage but the representation of obstacles is less accurate. Alternatively, mapping is avoided in [10], by considering groups of UAVs, each observing a part of the RoI with complementary viewpoints. This requires, however, a much larger number of UAVs.

Many prior works addressing the Cooperative Search, Acquisition, and Track (CSAT) problem [10]–[12] assume that UAVs get directly a noisy measurement of the state of targets present in the Field of View (FoV) of their sensor. The way these measurements are obtained is often overlooked. Consequently, the characteristics of the assumed measurement noise are difficult to justify.

Several techniques have been proposed for target detection using a CVS system. For instance, [13]–[15] exploit some specific pixels belonging to a 2D bounding box containing the pixels associated to the target. The bounding box is provided by target detection techniques [16], [17], but the pixel selection is often heuristic and may lead to relatively large localization errors. Deep learning algorithms are able to estimate the location and orientation of vehicles [18] from images and depth measurements. Nevertheless, these algorithms may provide erroneous estimates. Moreover, several approaches introduce probabilities of non-detection and of false alarm [6], [19], [20], but their value is difficult to characterize, since the performance of target detection

* Université Paris-Saclay, Centralesupelec, CNRS, L2S 91192, Gif-sur-Yvette, France first name.last name @ l2s.centralesupelec.fr

** DTIS, ONERA Université Paris-Saclay, Palaiseau, France first name.last name @ onera.fr

979-8-3503-1632-2/24/$31.00 ©2024 IEEE

techniques depends on the observation conditions and on the efficiency of the embedded CVS [7], [21].

III. Problem formulation

Consider an environment equipped with a frame $\mathcal{F}$, in which a fleet of N^{u} UAVs, each equipped with a CVS, is tasked to search and track a fixed but unknown number N^{t} of ground targets located in a RoI $\mathbb{X}_0 \subset \mathbb{R}^2 \times \mathbb{R}^+$ with a flat ground $\mathbb{X}_{\mathrm{g}} = \{\boldsymbol{x} \in \mathbb{X}_0 \mid x_3 = 0\}$. The RoI contains an unknown number N^{o} of static obstacles. $\mathbb{S}^{\mathrm{o}}_m \subset \mathbb{X}_0$ is the shape of the m-th obstacle, *i.e.*, the part of $\mathbb{X}_0$ it occupies. The sets $\mathcal{N}^{\mathrm{u}}$, $\mathcal{N}^{\mathrm{t}}$, and $\mathcal{N}^{\mathrm{o}}$ respectively gather the indexes of UAVs, targets, and obstacles. The time is sampled with a period T and k refers to the time index.

At time $t_k = kT$, assume that the knowledge UAV i has about its environment consists of a list $\mathcal{L}^{\mathrm{t}}_{i,k-1}$ of indexes of targets already identified. For each $j \in \mathcal{L}^{\mathrm{t}}_{i,k-1}$, UAV i has access to a set estimate $\mathbb{X}^{\mathrm{t}}_{i,j,k-1} \subset \mathbb{X}_{\mathrm{g}}$ containing the location of target j at time t_{k-1}. A set $\overline{\mathbb{X}}^{\mathrm{t}}_{i,k-1} \subset \mathbb{X}_{\mathrm{g}}$ containing the locations of targets still to be identified at time t_{k-1} has also been evaluated. Finally, UAV i has a map $\mathcal{M}_{i,k-1}$ of the RoI, approximating $\mathbb{X}_{\mathrm{g}} \cup \bigcup_{m \in \mathcal{N}^{\mathrm{o}}} \mathbb{S}^{\mathrm{o}}_m$. All these sets exploit the information collected or exchanged up to time t_k.

Accounting for the dynamics of targets, a prediction $\mathbb{X}^{\mathrm{t}}_{i,j,k|k-1}$ of $\mathbb{X}_{i,j,k}$, $j \in \mathcal{L}^{\mathrm{t}}_{i,k-1}$ and a prediction $\overline{\mathbb{X}}^{\mathrm{t}}_{i,k|k-1}$ of $\overline{\mathbb{X}}^{\mathrm{t}}_{i,k}$ can be evaluated.

At time t_k, the camera of UAV i acquires an image $\mathbf{I}_{i,k}$ of a part of its environment. Its embedded CVS generates then a depth-map $\mathbf{D}_{i,k}$ [22], an array of pixel labels $\mathbf{L}_{i,k}$ [23], a list $\mathcal{D}^{\mathrm{t}}_{i,k}$ of indexes of identified targets, and a list $\mathcal{B}^{\mathrm{t}}_{i,k} = \{[\mathcal{Y}^{\mathrm{t}}_{i,j,k}]\}_{j \in \mathcal{D}_{i,k}}$ of bounding boxes around identified targets in $\mathbf{I}_{i,k}$ [14], [16], [17].

Using this new information, UAV i evaluates $\mathcal{L}^{\mathrm{t}}_{i,k|k} = \mathcal{L}^{\mathrm{t}}_{i,k-1} \cup \mathcal{D}^{\mathrm{t}}_{i,k}$ and updates $\mathbb{X}^{\mathrm{t}}_{i,j,k|k-1}$ to get $\mathbb{X}^{\mathrm{t}}_{i,j,k|k}$ for all targets $j \in \mathcal{L}_{i,k|k}$ as well as $\overline{\mathbb{X}}^{\mathrm{t}}_{i,k|k-1}$ to get $\overline{\mathbb{X}}^{\mathrm{t}}_{i,k|k}$. An updated version of the map $\mathcal{M}_{i,k|k}$ is also obtained.

Then, UAV i broadcasts the updated set estimates and the map to its neighbors with indices in $\mathcal{N}_{i,k} \subset \mathcal{N}^{\mathrm{t}}$. From similar information received from its neighbors, UAV i can evaluate $\mathcal{L}^{\mathrm{t}}_{i,k}$, $\mathbb{X}^{\mathrm{t}}_{i,j,k}$ for all $j \in \mathcal{L}^{\mathrm{t}}_{i,k}$, $\overline{\mathbb{X}}^{\mathrm{t}}_{i,k}$ as well as $\mathcal{M}_{i,k}$.

The aim of this paper is to design the trajectories and guidance laws of UAVs minimizing the target location estimation uncertainty evaluated as $\phi(\overline{\mathbb{X}}^{\mathrm{t}}_{i,k} \cup \bigcup_{j \in \mathcal{L}^{\mathrm{t}}_{i,k}} \mathbb{X}^{\mathrm{t}}_{i,j,k})$, where $\phi(\mathbb{X})$ is the area of the set $\mathbb{X} \subset \mathbb{R}^2$.

A distributed solution to evaluate the guidance laws of UAVs is proposed in what follows. Hypotheses and models related to the targets and UAVs are first introduced in Sections IV-A and IV-B. Then, hypotheses on the relation between the output of the CVS and the elements of the UAV environment are introduced in Section IV-C. These hypotheses are instrumental in the characterization of the set estimates and of the map managed by the UAVs.

IV. Hypotheses and models

In what follows, vectors with no superscript related to a frame are implicitly expressed in $\mathcal{F}$.

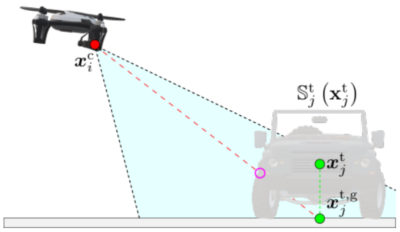
$\boldsymbol{x}_i^{\mathrm{c}}$
$\mathbb{S}^{\mathrm{t}}_j(\mathbf{x}^{\mathrm{t}}_j)$
$\boldsymbol{x}^{\mathrm{t}}_j$
$\boldsymbol{x}^{\mathrm{t,g}}_j$

Fig. 1. Unknown shape $\mathbb{S}^{\mathrm{t}}_j(\mathbf{x}^{\mathrm{t}}_j)$, center of mass $\boldsymbol{x}^{\mathrm{t}}_j$, and location $\boldsymbol{x}^{\mathrm{t,q}}_j$ of target j. The purple dot represents the part of $\mathbb{S}^{\mathrm{t}}_j(\mathbf{x}^{\mathrm{t}}_j)$ seen by the UAV.

A. Target model

At time t_k, the state of target $j \in \mathcal{N}^{\mathrm{t}}$ is $\mathbf{x}^{\mathrm{t}}_{j,k}$. It contains the vector $\boldsymbol{x}^{\mathrm{t}}_{j,k}$ of coordinates of the center of mass of target j. The targets are assumed to evolve on the ground $\mathbb{X}_{\mathrm{g}}$ and never leave $\mathbb{X}_0$ nor enter in any obstacle. The location $\boldsymbol{x}^{\mathrm{t,g}}_{j,k} = \boldsymbol{p}_{\mathrm{g}}(\boldsymbol{x}^{\mathrm{t}}_{j,k})$ of target j is the projection of $\boldsymbol{x}^{\mathrm{t}}_{j,k}$ on $\mathbb{X}_{\mathrm{g}}$ and is assumed to evolve as

$$\boldsymbol{x}^{\mathrm{t,g}}_{j,k+1} = \mathbf{f}^{\mathrm{t}}(\boldsymbol{x}^{\mathrm{t,g}}_{j,k}, \boldsymbol{v}^{\mathrm{t}}_{j,k}), \tag{1}$$

where $\mathbf{f}^{\mathrm{t}}$ is known and $\boldsymbol{v}^{\mathrm{t}}_{j,k}$ represents some unknown control input only assumed to belong to a known box $[\boldsymbol{v}^{\mathrm{t}}]$.

The space occupied by target j is $\mathbb{S}^{\mathrm{t}}_j(\mathbf{x}^{\mathrm{t}}_{j,k}) \subset \mathbb{X}_0$ and depends on $\mathbf{x}^{\mathrm{t}}_{j,k}$. We assume that $\boldsymbol{x}^{\mathrm{t}}_{j,k} \in \mathbb{S}^{\mathrm{t}}_j(\mathbf{x}^{\mathrm{t}}_{j,k})$. Nevertheless, the target location $\boldsymbol{x}^{\mathrm{t,g}}_{j,k}$ does not necessarily belong to $\mathbb{S}^{\mathrm{t}}_j(\mathbf{x}^{\mathrm{t}}_{j,k})$, as in the case of cars, see Figure 1. The type of targets to be localized is usually known and some information about their dimensions is assumed available: For any target state $\mathbf{x}^{\mathrm{t}}_{j,k}$, $\mathbb{S}^{\mathrm{t}}_j(\mathbf{x}^{\mathrm{t}}_{j,k}) \subset \mathbb{C}^{\mathrm{t}}(\boldsymbol{x}^{\mathrm{t,g}}_{j,k})$, where $\mathbb{C}^{\mathrm{t}}(\boldsymbol{x}^{\mathrm{t,g}}_{j,k})$ is the vertical circular right cylinder of known height h^{t} and radius r^{t} with basis centered in $\boldsymbol{x}^{\mathrm{t,g}}_{j,k}$.

Moreover, we assume that targets always remain at a distance larger than r_{s} to other targets and to obstacles of the environment.

B. UAV model

At time t_k, the state of UAV $i \in \mathcal{N}^{\mathrm{u}}$ is $\mathbf{x}^{\mathrm{u}}_{i,k}$. It is assumed perfectly known by UAV i and to evolve as

$$\mathbf{x}^{\mathrm{u}}_{i,k+1} = \mathbf{f}^{\mathrm{u}}\left(\mathbf{x}^{\mathrm{u}}_{i,k}, \mathbf{u}^{\mathrm{u}}_{i,k}\right), \tag{2}$$

where $\mathbf{f}^{\mathrm{u}}$ is known and $\mathbf{u}^{\mathrm{u}}_{i,k}$ is the control input constrained in the set $\mathbb{U}$. The vector $\mathbf{x}^{\mathrm{u}}_{i,k}$ contains the location of the center of mass $\boldsymbol{x}^{\mathrm{u}}_{i,k}$ of UAV i. The space occupied by UAV i is $\mathbb{S}^{\mathrm{u}}(\mathbf{x}^{\mathrm{u}}_{i,k})$. A body frame $\mathcal{F}^{\mathrm{b}}_i$, with origin $\boldsymbol{x}^{\mathrm{u}}_{i,k}$ is attached to UAV i. The rotation matrix from $\mathcal{F}^{\mathrm{b}}_i$ to $\mathcal{F}$ is $\mathbf{M}^{\mathcal{F}}_{\mathcal{F}^{\mathrm{b}}_i}$.

$\mathcal{N}_{i,k}$ is the set of indexes of UAVs with which UAV i is able to communicate at time t_k. We consider that $i \in \mathcal{N}_{i,k}$.

C. Model of the CVS

Each UAV is assumed equipped with identical camera and CVS. The time index k is omitted in this section.

1) Camera model: A pinhole model without distortion [24] is considered, where $\mathcal{F}^{\mathrm{c}}_i$ is the camera frame and $\boldsymbol{x}^{\mathrm{c}}_i$ is the vector of coordinates of its optical center. The FoV of UAV i is described by a half-cone of apex $\boldsymbol{x}^{\mathrm{c}}_i$ and with four unit vectors $\boldsymbol{v}^{\mathcal{F}^{\mathrm{c}}_i}_\ell$, $\ell = 1, \ldots, 4$, *i.e.*,

$$\mathbb{F}\left(\mathbf{x}^{\mathrm{u}}_i\right) = \left\{ \boldsymbol{x}^{\mathrm{c}}_i + \textstyle\sum_{\ell=1}^4 a_\ell \mathbf{M}^{\mathcal{F}}_{\mathcal{F}^{\mathrm{c}}_i} \boldsymbol{v}^{\mathcal{F}^{\mathrm{c}}_i}_\ell \mid a_\ell \in \mathbb{R}^+ \right\}, \tag{3}$$

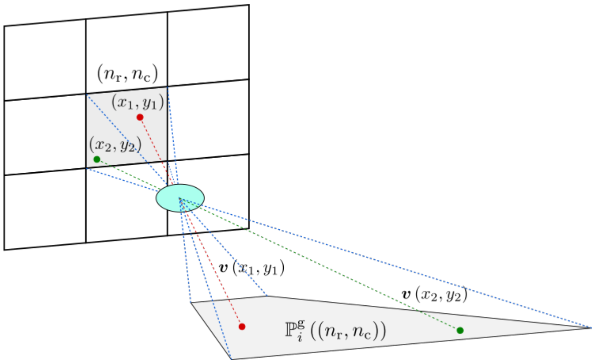
$(n_\mathrm{r}, n_\mathrm{c})$
(x_1, y_1)
(x_2, y_2)
$\boldsymbol{v}(x_1, y_1)$
$\boldsymbol{v}(x_2, y_2)$
$\mathbb{P}_i^\mathrm{g}((n_\mathrm{r}, n_\mathrm{c}))$

Fig. 2. Pinhole model of the camera and several light rays (green, red and blue) contributing to the illumination of the pixel $(n_\mathrm{r}, n_\mathrm{c})$ highlighted in gray. The quadrangle $\mathbb{P}_i^\mathrm{g}((n_\mathrm{r}, n_\mathrm{c}))$ is defined in Section V.

where $\mathbf{M}^{\mathcal{F}}_{\mathcal{F}_i^\mathrm{c}} = \mathbf{M}^{\mathcal{F}}_{\mathcal{F}_i^\mathrm{b}} \mathbf{M}^{\mathcal{F}_i^\mathrm{b}}_{\mathcal{F}_i^\mathrm{c}}$ and $\mathbf{M}^{\mathcal{F}_i^\mathrm{c}}_{\mathcal{F}_i^\mathrm{b}}$ is the rotation matrix from $\mathcal{F}_i^\mathrm{b}$ to $\mathcal{F}_i^\mathrm{c}$. A function $\boldsymbol{p}_\mathrm{c}(\mathbf{x}_i^\mathrm{u}, \boldsymbol{x})$ is assumed available to provide the coordinates in the CCD array of a point $\boldsymbol{x} \in \mathbb{F}(\mathbf{x}_i^\mathrm{u})$, see, *e.g.*, [25]. The notation $\boldsymbol{p}_\mathrm{c}(\mathbf{x}_i^\mathrm{u}, \boldsymbol{x}) \in (n_\mathrm{r}, n_\mathrm{c})$, with $\boldsymbol{x} \in \mathbb{F}(\mathbf{x}_i^\mathrm{u})$, indicates that the projection of $\boldsymbol{x}$ on $\mathbf{I}_i$ belongs to the pixel with coordinates $(n_\mathrm{r}, n_\mathrm{c}) \in \mathcal{N}^\mathrm{I}$, with $\mathcal{N}^\mathrm{I} = \{1 \dots N_\mathrm{r}\} \times \{1 \dots N_\mathrm{c}\}$.

Considering the pinhole model, each light ray passing through $\boldsymbol{x}_i^\mathrm{c}$ and illuminating the CCD array at $(x, y) \in [0, N_\mathrm{c}] \times [0, N_r]$ can be modeled by a half-line of direction

$$\boldsymbol{v}^{\mathcal{F}_i^\mathrm{c}}(x, y) = \frac{1}{\nu(x, y)} \begin{pmatrix} \left(\frac{N_\mathrm{c}}{2} - x\right)/f_\mathrm{c} \\ \left(\frac{N_\mathrm{r}}{2} - y\right)/f_\mathrm{r} \\ 1 \end{pmatrix}, \tag{4}$$

where f_c and f_r are the camera focal lengths (in pixels) and $\nu(x, y)$ is a normalization coefficient. Figure 2 illustrates the illumination of a pixel by two different light rays.

2) Depth map and bounded-error range measurements: The part $\mathbb{S}_\mathrm{RoI} = \mathbb{X}_\mathrm{g} \cup \bigcup_{m \in \mathcal{N}^\mathrm{o}} \mathbb{S}_m^\mathrm{o} \cup \bigcup_{j \in \mathcal{N}^\mathrm{t}} \mathbb{S}_j^\mathrm{t}(\mathbf{x}_j^\mathrm{t}) \cup \bigcup_{\ell \in \mathcal{N}^\mathrm{u}} \mathbb{S}^\mathrm{u}(\mathbf{x}_\ell^\mathrm{u})$ of the environment that UAV i is able to perceive consists of the ground, the obstacles, the targets, and other UAVs. The distance between $\boldsymbol{x}_i^\mathrm{c}$ of the camera and $\mathbb{S}_\mathrm{RoI}$ along $\boldsymbol{v} \in \mathcal{V}_i(n_\mathrm{r}, n_\mathrm{c})$, for any $(n_\mathrm{r}, n_\mathrm{c}) \in \mathcal{N}^\mathrm{I}$, is $\rho(\boldsymbol{x}_i^\mathrm{c}, \boldsymbol{v}) = d_{\boldsymbol{v}}(\boldsymbol{x}_i^\mathrm{c}, \mathbb{S}_\mathrm{RoI})$,where $d_{\boldsymbol{v}}(\boldsymbol{x}, \mathbb{S})$ is the distance from a point $\boldsymbol{x} \in \mathbb{X}_0$ to the intersection of the set $\mathbb{S}$ along the half-line of origin $\boldsymbol{x}$ and direction $\boldsymbol{v}$.

We assume that each element $\mathbf{D}_i(n_\mathrm{r}, n_\mathrm{c})$ of $\mathbf{D}_i$ satisfies

$$\mathbf{D}_i(n_\mathrm{r}, n_\mathrm{c}) = \mathbf{D}_i^0(n_\mathrm{r}, n_\mathrm{c})(1 + w) \tag{5}$$

where $\mathbf{D}_i^0(n_\mathrm{r}, n_\mathrm{c}) = \rho(\mathbf{x}_i^\mathrm{u}, \boldsymbol{v})$ is the distance between $\boldsymbol{x}_i^\mathrm{c}$ and $\mathbb{S}_\mathrm{RoI}$ along some unknown direction $\boldsymbol{v} \in \mathcal{V}_i(n_\mathrm{r}, n_\mathrm{c})$. The noise w is assumed to belong to the known interval $[\underline{w}, \overline{w}]$. Consequently, $\mathbf{D}_i^0(n_\mathrm{r}, n_\mathrm{c})$ is contained in the interval

$$[\mathbf{D}_i](n_\mathrm{r}, n_\mathrm{c}) = \left[\frac{1}{1 + \overline{w}}, \frac{1}{1 + \underline{w}}\right] \mathbf{D}_i(n_\mathrm{r}, n_\mathrm{c}), \tag{6}$$

and represents a bounded-error range measurement associated to the pixel $(n_\mathrm{r}, n_\mathrm{c})$.

3) Classifier: The array of labels $\mathbf{L}_i$ is obtained at the output of a classifier. Each element $(n_\mathrm{r}, n_\mathrm{c})$ of $\mathbf{L}_i$ is assumed to belong to one of the following classes: Ground (G), Target (T), Obstacle (O), and Unknown/Not Labeled (U). The last class is for unclassified pixels due to a lack of confidence. Then, four sets may be deduced from $\mathbf{L}_i$, namely $\mathcal{Y}_i^\mathrm{g}$, $\mathcal{Y}_i^\mathrm{t}$, $\mathcal{Y}_i^\mathrm{o}$, and $\mathcal{Y}_i^\mathrm{n}$ gathering coordinates of pixels respectively labeled as Ground, Target, Obstacle, and Unknown/Not Labeled.

Moreover, we assume that if a pixel $(n_\mathrm{r}, n_\mathrm{c}) \in \mathcal{Y}_i^\mathrm{g}$, then all light rays illuminating pixel $(n_\mathrm{r}, n_\mathrm{c})$ stem from the ground $\mathbb{X}_\mathrm{g}$, *i.e.*,

$$\forall \boldsymbol{v} \in \mathcal{V}_i(n_\mathrm{r}, n_\mathrm{c}), \rho(\mathbf{x}_i^\mathrm{c}, \boldsymbol{v}) = d_{\boldsymbol{v}}(\boldsymbol{x}_i^\mathrm{c}, \mathbb{X}_\mathrm{g}). \tag{7}$$

Similarly, if a pixel $(n_\mathrm{r}, n_\mathrm{c}) \in \mathcal{Y}_i^\mathrm{t}$, then there exists a target $j \in \mathcal{N}^\mathrm{t}$ such that all light rays illuminating pixel $(n_\mathrm{r}, n_\mathrm{c})$ stem from $\mathbb{S}_j^\mathrm{t}(\mathbf{x}_{j,k}^\mathrm{t})$, *i.e.*,

$$\exists j \in \mathcal{N}^\mathrm{t}, \forall \boldsymbol{v} \in \mathcal{V}_i(n_\mathrm{r}, n_\mathrm{c}), \rho(\mathbf{x}_i^\mathrm{c}, \boldsymbol{v}) = d_{\boldsymbol{v}}\left(\boldsymbol{x}_i^\mathrm{c}, \mathbb{S}_j^\mathrm{t}\left(\mathbf{x}_{j,k}^\mathrm{t}\right)\right). \tag{8}$$

Finally, if a pixel $(n_\mathrm{r}, n_\mathrm{c}) \in \mathcal{Y}_i^\mathrm{o}$, then there exists an obstacle $m \in \mathcal{N}^\mathrm{o}$ such that all light rays illuminating pixel $(n_\mathrm{r}, n_\mathrm{c})$ stem from $\mathbb{S}_m^\mathrm{o}$, *i.e.*,

$$\exists m \in \mathcal{N}^\mathrm{o}, \forall \boldsymbol{v} \in \mathcal{V}_i(n_\mathrm{r}, n_\mathrm{c}), \rho(\mathbf{x}_i^\mathrm{c}, \boldsymbol{v}) = d_{\boldsymbol{v}}(\boldsymbol{x}_i^\mathrm{c}, \mathbb{S}_m^\mathrm{o}). \tag{9}$$

4) Target detection and identification: When $\mathcal{Y}_i^\mathrm{t}$ is not empty, at least one target located within $\mathbb{F}(\mathbf{x}_i^\mathrm{u})$ has been detected. In such case, the CVS may also provide a list $\mathcal{D}_i^\mathrm{t} \subset \mathcal{N}^\mathrm{t}$ of identified target indexes and an axis-aligned box $\left[\mathcal{Y}_{i,j}^\mathrm{t}\right]$ for each $j \in \mathcal{D}_i^\mathrm{t}$. Even if $\mathcal{Y}_i^\mathrm{t}$ is not empty, $\mathcal{D}_i^\mathrm{t}$ may be empty when $\mathcal{Y}_i^\mathrm{t}$ does not provide enough information to identify a target.

Consider the set $\mathcal{Y}_{i,j}^\mathrm{t} \subset \mathcal{Y}_i^\mathrm{t}$ of all pixels associated to target j only. Even if the classifier is unable to provide $\mathcal{Y}_{i,j}^\mathrm{t}$, we assume that if $j \in \mathcal{D}_i^\mathrm{t}$, then $\mathcal{Y}_{i,j}^\mathrm{t} \neq \emptyset$ and that $\left[\mathcal{Y}_{i,j}^\mathrm{t}\right]$ contains $\mathcal{Y}_{i,j}^\mathrm{t}$, *i.e.*,

$$j \in \mathcal{D}_i^\mathrm{t} \Rightarrow \mathcal{Y}_{i,j}^\mathrm{t} \subset \left[\mathcal{Y}_{i,j}^\mathrm{t}\right]. \tag{10}$$

D. Assumption on observed targets

Consider some target j such that $\boldsymbol{x}_j^\mathrm{t,g} \in \mathbb{F}(\mathbf{x}_i^\mathrm{u})$. We assume that the half-open segment $[\boldsymbol{x}_i^\mathrm{c}, \boldsymbol{x}_j^\mathrm{t,g}[$ intersects $\mathbb{S}_j^\mathrm{t}(\mathbf{x}_j^\mathrm{t})$, *i.e.*,

$$\boldsymbol{x}_j^\mathrm{t,g} \in \mathbb{F}(\mathbf{x}_i^\mathrm{u}) \implies \left[\boldsymbol{x}_i^\mathrm{c}, \boldsymbol{x}_j^\mathrm{t,g}\right[\cap \mathbb{S}_j^\mathrm{t}\left(\mathbf{x}_j^\mathrm{t}\right) \neq \emptyset. \tag{11}$$

If $\boldsymbol{x}_j^\mathrm{t,g} \in \mathbb{F}(\mathbf{x}_i^\mathrm{u})$, then some point on $\mathbb{S}_j^\mathrm{t}\left(\mathbf{x}_j^\mathrm{t}\right)$ reflects a light ray that illuminates the CCD array of the camera (in absence of obstacles), see Figure 1. Therefore, if $\boldsymbol{x}_j^\mathrm{t,g} \in \mathbb{F}(\mathbf{x}_i^\mathrm{u})$, then there exists a pixel $(n_\mathrm{r}, n_\mathrm{c})$ such that $\boldsymbol{p}_\mathrm{c}\left(\mathbf{x}_i^\mathrm{u}, \boldsymbol{x}_j^\mathrm{t,g}\right) \in (n_\mathrm{r}, n_\mathrm{c})$. Then according to (11) and (7), $(n_\mathrm{r}, n_\mathrm{c}) \notin \mathcal{Y}_i^\mathrm{g}$.

The assumption (11) is instrumental to characterize parts of $\mathbb{X}_\mathrm{g}$ that cannot contain any target location. Its validity mainly depends on the camera orientation and on the target characteristics.

V. Exploiting CVS information

This section describes the way information from the CVS is used to evaluate set estimates of the location of targets, sets free of targets, and to build the map managed by UAVs. The time index k is also omitted in this section.

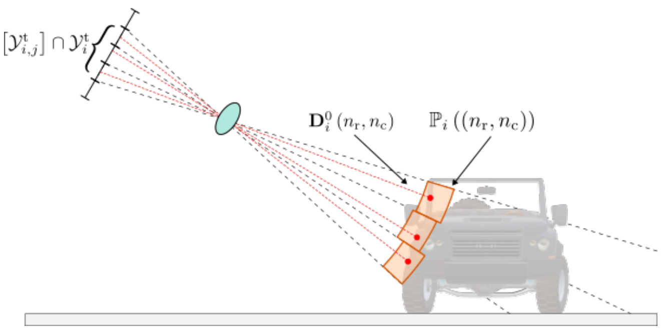
$[\mathcal{Y}^{\mathrm{t}}_{i,j}] \cap \mathcal{Y}^{\mathrm{t}}_i$
$\mathbf{D}^0_i(n_{\mathrm{r}}, n_{\mathrm{c}})$
$\mathbb{P}_i((n_{\mathrm{r}}, n_{\mathrm{c}}))$

Fig. 3. Sets $\mathbb{P}_i((n_{\mathrm{r}}, n_{\mathrm{c}}))$ for different $(n_{\mathrm{r}}, n_{\mathrm{c}}) \in \left[\mathcal{Y}^{\mathrm{t}}_{i,j}\right]$

A. Estimation of the location of a target

To estimate the location $\boldsymbol{x}^{\mathrm{t,g}}_j$ of an identified target $j \in \mathcal{D}^{\mathrm{t}}_i$, UAV i exploits the indexes of pixels in $\mathcal{Y}^{\mathrm{t}}_i$ (labeled as Target), the bounding box $\left[\mathcal{Y}^{\mathrm{t}}_{i,j}\right]$, and the depth map $\mathbf{D}_i$. A set estimate $\mathbb{X}^{\mathrm{t,m}}_{i,j}$ is evaluated such that $\boldsymbol{x}^{\mathrm{t,g}}_j \in \mathbb{X}^{\mathrm{t,m}}_{i,j}$.

One first characterizes a subset of $\mathbb{X}_0$ intersecting the shape $\mathbb{S}^{\mathrm{t}}_j\left(\mathbf{x}^{\mathrm{t}}_j\right)$ of target j. Consider the set

$$\mathbb{P}_i((n_{\mathrm{r}}, n_{\mathrm{c}})) = \{\boldsymbol{x} \in \mathbb{F}(\mathbf{x}^{\mathrm{u}}_i) \cap \mathbb{X}_0 \mid \exists \boldsymbol{v} \in \mathcal{V}_i(n_{\mathrm{r}}, n_{\mathrm{c}}), \\ d_{\boldsymbol{v}}(\boldsymbol{x}^{\mathrm{c}}_i, \{\boldsymbol{x}\}) \in [\mathbf{D}_i](n_{\mathrm{r}}, n_{\mathrm{c}})\} \tag{12}$$

containing all points in $\mathbb{F}(\mathbf{x}^{\mathrm{u}}_i) \cap \mathbb{X}_0$ that may have contributed to the illumination of a pixel $(n_{\mathrm{r}}, n_{\mathrm{c}}) \in \mathbf{I}_i$ while being at a distance from UAV i consistent with $\mathbf{D}_i(n_{\mathrm{r}}, n_{\mathrm{c}})$. Figure 3 represents $\mathbb{P}_i((n_{\mathrm{r}}, n_{\mathrm{c}}))$ for three different pixels $(n_{\mathrm{r}}, n_{\mathrm{c}}) \in [\mathcal{Y}^{\mathrm{t}}_{i,j}]$.

As $j \in \mathcal{D}^{\mathrm{t}}_i$, according to Section IV-C.4, $\mathcal{Y}^{\mathrm{t}}_{i,j} \subset \mathcal{Y}^{\mathrm{t}}_i$ is not empty and $\mathcal{Y}^{\mathrm{t}}_{i,j} \subset \left[\mathcal{Y}^{\mathrm{t}}_{i,j}\right]$. Since only $\mathcal{Y}^{\mathrm{t}}_i$ is available, the set

$$\mathbb{P}^{\mathrm{t}}_{i,j} = \bigcup\nolimits_{(n_{\mathrm{r}}, n_{\mathrm{c}}) \in [\mathcal{Y}^{\mathrm{t}}_{i,j}] \cap \mathcal{Y}^{\mathrm{t}}_i} \mathbb{P}_i((n_{\mathrm{r}}, n_{\mathrm{c}})), \tag{13}$$

is introduced and shown in Corollary 2 of [3] to have a non-empty intersection with $\mathbb{S}^{\mathrm{t}}_j(\mathbf{x}^{\mathrm{t}}_j)$. As $\mathbb{S}^{\mathrm{t}}_j\left(\mathbf{x}^{\mathrm{t}}_j\right)$ is unknown, the outer-approximation $\mathbb{C}^{\mathrm{t}}\left(\boldsymbol{x}^{\mathrm{t,g}}_j\right)$ of $\mathbb{S}^{\mathrm{t}}_j\left(\mathbf{x}^{\mathrm{t}}_j\right)$ is considered. If some $\boldsymbol{x} \in \mathbb{P}^{\mathrm{t}}_{i,j} \cap \mathbb{S}^{\mathrm{t}}_j\left(\mathbf{x}^{\mathrm{t}}_j\right)$ would be available, exploiting the fact that $\mathbb{S}^{\mathrm{t}}_j\left(\mathbf{x}^{\mathrm{t}}_j\right) \subset \mathbb{C}^{\mathrm{t}}\left(\boldsymbol{x}^{\mathrm{t,g}}_j\right)$, one would have $\boldsymbol{x}^{\mathrm{t,g}}_j \in \mathbb{C}^{\mathrm{t}}\left(\boldsymbol{p}_{\mathrm{g}}(\boldsymbol{x})\right)$. As $\mathbb{P}^{\mathrm{t}}_{i,j}$ is only known to intersect $\mathbb{S}^{\mathrm{t}}_j\left(\mathbf{x}^{\mathrm{t}}_j\right)$, the set estimate $\mathbb{X}^{\mathrm{t,m}}_{i,j}$ of $\boldsymbol{x}^{\mathrm{t,g}}_j$ introduced in Proposition 1 accounts for the fact that $\boldsymbol{x}$ is only known to belong to $\mathbb{P}^{\mathrm{t}}_{i,j}$.

Proposition 1: If $j \in \mathcal{D}^{\mathrm{t}}_i$, then $\boldsymbol{x}^{\mathrm{t,g}}_j \in \mathbb{X}^{\mathrm{t,m}}_{i,j}$ with

$$\mathbb{X}^{\mathrm{t,m}}_{i,j} = \mathbb{X}_{\mathrm{g}} \cap \bigcup\nolimits_{\boldsymbol{x} \in \mathbb{P}^{\mathrm{t}}_{i,j}} \mathbb{C}^{\mathrm{t}}\left(\boldsymbol{p}_{\mathrm{g}}(\boldsymbol{x})\right). \tag{14}$$

The proof for Proposition 1 is detailed in [3], Section IV-A. The evaluation of $\mathbb{X}^{\mathrm{t,m}}_{i,j}$ is detailed in [3], Appendix A and illustrated in Figure 4.

B. Estimation of the space free of target

Using pixels labeled as Ground and Obstacle, UAV i is able to characterize subsets of $\mathbb{X}_{\mathrm{g}}$ free of targets.

1) Using pixels labeled as Ground: According to (7), pixels in $\mathcal{Y}^{\mathrm{g}}_i$ have only been illuminated by light rays stemming from the ground $\mathbb{X}_{\mathrm{g}}$. Using (11), from $\mathcal{Y}^{\mathrm{g}}_i$, UAV i can evaluate a subset of $\mathbb{X}_{\mathrm{g}}$ which does not contain any target location. For that purpose, consider pixel $(n_{\mathrm{r}}, n_{\mathrm{c}})$ and the set

$$\mathbb{P}^{\mathrm{g}}_i((n_{\mathrm{r}}, n_{\mathrm{c}})) = \{\boldsymbol{x} \in \mathbb{F}(\mathbf{x}^{\mathrm{u}}_i) \cap \mathbb{X}_{\mathrm{g}} \mid \boldsymbol{p}_{\mathrm{c}}(\mathbf{x}^{\mathrm{u}}_i, \boldsymbol{x}) \in (n_{\mathrm{r}}, n_{\mathrm{c}})\}$$

of all points of $\mathbb{X}_{\mathrm{g}}$ which may have contributed to the illumination of $(n_{\mathrm{r}}, n_{\mathrm{c}})$. According to (7) and (11), for all $j \in \mathcal{N}^{\mathrm{t}}$, $\boldsymbol{x}^{\mathrm{t,g}}_j \notin \mathbb{P}^{\mathrm{g}}_i((n_{\mathrm{r}}, n_{\mathrm{c}}))$. Now, consider all $(n_{\mathrm{r}}, n_{\mathrm{c}}) \in \mathcal{Y}^{\mathrm{g}}_i$.

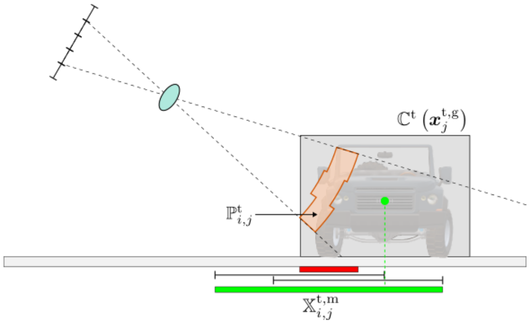
$\mathbb{C}^{\mathrm{t}}\left(\boldsymbol{x}^{\mathrm{t,g}}_j\right)$
$\mathbb{P}^{\mathrm{t}}_{i,j}$
$\mathbb{X}^{\mathrm{t,m}}_{i,j}$

Fig. 4. Estimation of the location of target j: The set $\mathbb{P}^{\mathrm{t}}_{i,j}$ is projected on the ground (red set). Then $\boldsymbol{p}_{\mathrm{g}}(\mathbb{C}^{\mathrm{t}})$ (black lines) is used to build $\mathbb{X}^{\mathrm{t,m}}_{i,j}$ (green set). The green dot represents $\boldsymbol{x}^{\mathrm{t}}_{j,k}$.

$$\mathbb{P}^{\mathrm{g}}_i\left(\mathcal{Y}^{\mathrm{g}}_i\right) = \bigcup\nolimits_{((n_{\mathrm{r}}, n_{\mathrm{c}})) \in \mathcal{Y}^{\mathrm{g}}_i} \mathbb{P}^{\mathrm{g}}_i((n_{\mathrm{r}}, n_{\mathrm{c}})). \tag{15}$$

Proposition 2 states that $\mathbb{P}^{\mathrm{g}}_i\left(\mathcal{Y}^{\mathrm{g}}_i\right)$ is free of target.

Proposition 2: For all $j \in \mathcal{N}^{\mathrm{t}}$. $\boldsymbol{x}^{\mathrm{t,g}}_j \notin \mathbb{P}^{\mathrm{g}}_i\left(\mathcal{Y}^{\mathrm{g}}_i\right)$.
The proof for Proposition 2 is in [3], Section IV-B. The characterization of $\mathbb{P}^{\mathrm{g}}_i\left(\mathcal{Y}^{\mathrm{g}}_i\right)$ is detailed in [3], Appendix C.

2) Using pixels labeled as Obstacle: Since the distance between a target and any obstacle is at least r_{s}, UAV i can also characterize parts of $\mathbb{X}_{\mathrm{g}}$ free of targets using pixels labeled as Obstacle. For that purpose, we introduce the r_{s}-neighborhood of a set $\mathbb{S} \subset \mathbb{X}_0$ as $\mathbb{N}(\mathbb{S}, r_{\mathrm{s}}) = \{\boldsymbol{x} \in \mathbb{X}_0 \mid d(\boldsymbol{x}, \mathbb{S}) \leqslant r_{\mathrm{s}}\}$ with $d(\boldsymbol{x}, \mathbb{S}) = \min_{\boldsymbol{y} \in \mathbb{S}} \|\boldsymbol{x} - \boldsymbol{y}\|$.

According to (9), for all $(n_{\mathrm{r}}, n_{\mathrm{c}}) \in \mathcal{Y}^{\mathrm{o}}_i$, we have $\mathbb{P}_i((n_{\mathrm{r}}, n_{\mathrm{c}})) \cap \mathbb{S}^{\mathrm{o}}_m \neq \emptyset$ for some obstacle $m \in \mathcal{N}^{\mathrm{o}}$. As for $\mathbb{P}^{\mathrm{t}}_{i,j}$, $\mathbb{P}_i((n_{\mathrm{r}}, n_{\mathrm{c}})) \cap \mathbb{S}^{\mathrm{o}}_m$ is only known to be included in $\mathbb{P}_i((n_{\mathrm{r}}, n_{\mathrm{c}}))$. Therefore, consider the set

$$\mathbb{S}^{\mathrm{o}}_i((n_r, n_c), r_{\mathrm{s}}) = \bigcap\nolimits_{\boldsymbol{x} \in \mathbb{P}_i((n_{\mathrm{r}}, n_{\mathrm{c}}))} \mathbb{N}(\{\boldsymbol{x}\}, r_{\mathrm{s}}). \tag{16}$$

defined as the intersections of the balls $\mathbb{N}(\{\boldsymbol{x}\}, r_{\mathrm{s}})$ for all $\boldsymbol{x} \in \mathbb{P}_i((n_{\mathrm{r}}, n_{\mathrm{c}}))$. If $(n_{\mathrm{r}}, n_{\mathrm{c}}) \in \mathcal{Y}^{\mathrm{o}}_i$, then $\mathbb{S}^{\mathrm{o}}_i((n_r, n_c), r_{\mathrm{s}})$ is an inner-approximation of the r_{s}-neighborhood of an unknown obstacle $m \in \mathcal{N}^{\mathrm{o}}$, *i.e.*,

$$\mathbb{S}^{\mathrm{o}}((n_r, n_c), r_{\mathrm{s}}) \subset \mathbb{N}(\mathbb{S}^{\mathrm{o}}_m, r_{\mathrm{s}}). \tag{17}$$

Now, the set

$$\mathbb{S}^{\mathrm{o}}_i(\mathcal{Y}^{\mathrm{o}}_i, r_{\mathrm{s}}) = \bigcup\nolimits_{(n_r, n_c) \in \mathcal{Y}^{\mathrm{o}}_i} \boldsymbol{p}_{\mathrm{g}}\left(\mathbb{S}^{\mathrm{o}}_i((n_r, n_c), r_{\mathrm{s}})\right) \tag{18}$$

is an inner-approximation of the r_{s}-neighborhood of the projection on $\mathbb{X}_{\mathrm{g}}$ all obstacles located within the FoV of UAV i. $\mathbb{S}^{\mathrm{o}}_i(\mathcal{Y}^{\mathrm{o}}_i, r_{\mathrm{s}})$ can be proved to be free of any target.

Proposition 3: If $\mathcal{Y}^{\mathrm{o}}_i \neq \emptyset$, then $\boldsymbol{x}^{\mathrm{t,g}}_j \notin \mathbb{S}^{\mathrm{o}}_i(\mathcal{Y}^{\mathrm{o}}_i, r_{\mathrm{s}})$, for all $j \in \mathcal{N}^{\mathrm{t}}$.

The proof for Proposition 3 is in [3], Section IV-B. The evaluation of $\underline{\mathbb{X}}^{\mathrm{o}}_i$ is detailed in [3], Appendix B.

C. Estimation of the hidden area

The pixels labeled Obstacle or Target result from the observation of an obstacle or a target which may have hidden another target. Moreover, pixels labeled Unknown may also correspond to a target. Consequently, as will be seen in Section VI, it is useful to characterize parts of the ground which where hidden, or not classified. This hidden area is evaluated as

$$\mathbb{P}_i^{\mathrm{g}}\left(\mathcal{Y}_i^{\mathrm{o}} \cup \mathcal{Y}_i^{\mathrm{t}} \cup \mathcal{Y}_i^{\mathrm{n}}\right) = \bigcup_{(n_{\mathrm{r}},n_{\mathrm{c}})\in\mathcal{Y}_i^{\mathrm{o}}\cup\mathcal{Y}_i^{\mathrm{t}}\cup\mathcal{Y}_i^{\mathrm{n}}} \mathbb{P}_i^{\mathrm{g}}\left((n_{\mathrm{r}},n_{\mathrm{c}})\right). \quad (19)$$

D. Estimation of the occupancy elevation map

UAV i maintains an OEM representation $\mathcal{M}_i$ of the obstacles inspired from [4]. $\mathcal{M}_i$ consists of a covering of $\mathbb{X}_{\mathrm{g}}$ with N^{M} non-overlapping square cells. $\mathcal{N}^{\mathrm{M}}$ is the set of cell indexes, $\boldsymbol{c}(n)$, $s_i(n)$, and $h_i(n)$ represent respectively the center, the status, and the estimated height of the cell n. The cell centers are the same for all UAVs. The status takes value in $\{\mathrm{U},\mathrm{O},\mathrm{G}\}$ where U stands for *Unexplored*, O for *Obstacle*, *i.e.*, an obstacle has been detected on that cell, and G for *Ground*, *i.e.*, no obstacle is located on that cell. Then $\mathcal{M}_i = \{(\boldsymbol{c}(n), s_i(n), h_i(n))\}_{n\in\mathcal{N}^{\mathrm{M}}}$.

The status and the estimated height of a cell are updated using CVS information.

For each pixel $(n_{\mathrm{r}}, n_{\mathrm{c}}) \in \mathcal{Y}_{i,k}^{\mathrm{o}}$ (labeled Obstacle), according to (9), there exists $m \in \mathcal{N}^{\mathrm{o}}$ such that $\mathbb{P}_i\left((n_{\mathrm{r}}, n_{\mathrm{c}})\right)$ introduced in (12) intersects $\mathbb{S}_m^{\mathrm{o}}$. As this intersection is difficult to characterize, we consider that $\boldsymbol{m}\left(\mathbb{P}_i\left((n_{\mathrm{r}}, n_{\mathrm{c}})\right)\right)$, the barycenter of $\mathbb{P}_i\left((n_{\mathrm{r}}, n_{\mathrm{c}})\right)$, belongs to $\mathbb{S}_m^{\mathrm{o}}$ and change the status of the cell n containing the projection on $\mathbb{X}_{\mathrm{g}}$ of $\boldsymbol{m}\left(\mathbb{P}_i\left((n_{\mathrm{r}}, n_{\mathrm{c}})\right)\right)$ to Obstacle, *i.e.*, $s_i(n) = \mathrm{O}$. Moreover, the estimated height of that cell n is updated using $m_3\left(\mathbb{P}_i\left((n_{\mathrm{r}}, n_{\mathrm{c}})\right)\right)$, the height of the barycenter of $\mathbb{P}_i\left((n_{\mathrm{r}}, n_{\mathrm{c}})\right)$ as

$$h_i(n) = \max\{h_i(n), m_3\left(\mathbb{P}_i\left((n_{\mathrm{r}}, n_{\mathrm{c}})\right)\right)\}. \quad (20)$$

For each pixel $(n_{\mathrm{r}}, n_{\mathrm{c}}) \in \mathcal{Y}_i^{\mathrm{g}}$ (labeled Ground), the status and estimated height of all cells $n \in \mathcal{N}^{\mathrm{M}}$ such that $c_i(n) \in \mathbb{P}_i^{\mathrm{g}}\left((n_{\mathrm{r}}, n_{\mathrm{c}})\right)$, are updated as $s_i(n) = \mathrm{G}$ and $h_i(n) = 0$.

This approach leads to an approximate map of the obstacles used to estimate the parts of the FoV that will be hidden by obstacles, see Section VII.

VI. Estimation algorithm

This section describes a distributed set-membership target location estimator exploiting the information provided by the CVS and the set estimates introduced in Section V. The structure of the set-membership estimator is adapted from [2]. Only the parts related to the exploitation of the CVS information are detailed. For UAV i, the set estimates are initialized at time t_0 as $\mathcal{L}_{i,0}^{\mathrm{t}} = \emptyset$, $\mathcal{X}_{i,0}^{\mathrm{t}} = \emptyset$, and $\overline{\mathbb{X}}_{i,0}^{\mathrm{t}} = \mathbb{X}_{\mathrm{g}}$. In addition, UAV i also maintains an inner aproximation $\underline{\mathbb{S}}_{i,k}^{\mathrm{o}}$ of the r_{s}-neighborhood of the projection on $\mathbb{X}_{\mathrm{g}}$ of all obstacles, with $\underline{\mathbb{S}}_{i,k}^{\mathrm{o}} = \emptyset$. Finally, the map $\mathcal{M}_{i,0}$ is initialized as $s_{i,0}(n) = \mathrm{U}$ and $h_{i,0}(n) = 0$ for all $n \in \mathcal{N}^{\mathrm{M}}$.

A. Prediction step

At t_k, the predicted sets $\mathbb{X}_{i,j,k|k-1}^{\mathrm{t}}$ of possible future locations of identified target $j \in \mathcal{L}_{i,k}^{\mathrm{t}}$ and $\overline{\mathbb{X}}_{i,k|k-1}^{\mathrm{t}}$ of possible locations of targets still to be identified are evaluated using (1) and the fact that $\boldsymbol{v}_{j,k} \in [\boldsymbol{v}^{\mathrm{t}}]$ as follows

$$\mathbb{X}_{i,j,k|k-1}^{\mathrm{t}} = \mathbf{f}^{\mathrm{t}}\left(\mathbb{X}_{i,j,k-1}^{\mathrm{t}}, [\boldsymbol{v}^{\mathrm{t}}]\right) \cap \mathbb{X}_{\mathrm{g}}, \forall j \in \mathcal{L}_{i,k}^{\mathrm{t}} \quad (21)$$

$$\overline{\mathbb{X}}_{i,k|k-1}^{\mathrm{t}} = \mathbf{f}^{\mathrm{t}}(\overline{\mathbb{X}}_{i,k-1}^{\mathrm{t}}, [\boldsymbol{v}^{\mathrm{t}}]) \cap \mathbb{X}_{\mathrm{g}}. \quad (22)$$

B. Correction step using CVS information

The CVS of UAV i provides the list $\mathcal{D}_{i,k}^{\mathrm{t}}$, and the sets $\mathbb{X}_{i,j,k}^{\mathrm{t,m}}$, $\mathbb{S}_{i,k}^{\mathrm{o}}\left(\mathcal{Y}_i^{\mathrm{o}}, r_{\mathrm{s}}\right)$, and $\mathbb{P}_{i,k}^{\mathrm{g}}(\mathcal{Y}_{i,k}^{\mathrm{g}})$.

Based on Section V-B.2, the inner approximation of the r_s-neighborhoods of all obstacles is updated as $\underline{\mathbb{S}}_{i,k|k}^{\mathrm{o}} = \mathbb{S}_{i,k}^{\mathrm{o}}\left(\mathcal{Y}_i^{\mathrm{o}}, r_{\mathrm{s}}\right) \cup \underline{\mathbb{S}}_{i,k}^{\mathrm{o}}$.

From Sections V-B.1 and V-B.2, as the set $\mathbb{P}_{i,k}^{\mathrm{g}}(\mathcal{Y}_{i,k}^{\mathrm{g}}) \cup \underline{\mathbb{S}}_{i,k|k}^{\mathrm{o}}$ is proved to be free of targets the set $\overline{\mathbb{X}}_{i,k|k-1}^{\mathrm{t}}$ is updated as

$$\overline{\mathbb{X}}_{i,k|k}^{\mathrm{t}} = \overline{\mathbb{X}}_{i,k|k-1}^{\mathrm{t}} \setminus \left(\mathbb{P}_{i,k}^{\mathrm{g}}\left(\mathcal{Y}_{i,k}^{\mathrm{g}}\right) \cup \underline{\mathbb{S}}_{i,k|k}^{\mathrm{o}}\right), \quad (23)$$

For identified targets, three cases are considered. For each target $j \in \mathcal{L}_{i,k-1}^{\mathrm{t}} \setminus \mathcal{D}_{i,k}^{\mathrm{t}}$ already identified before time t_k and not identified at time t_k one has

$$\mathbb{X}_{i,j,k|k}^{\mathrm{t}} = \mathbb{X}_{i,j,k|k-1}^{\mathrm{t}} \setminus \left(\mathbb{P}_{i,k}^{\mathrm{g}}\left(\mathcal{Y}_{i,k}^{\mathrm{g}}\right) \cup \underline{\mathbb{S}}_{i,k|k}^{\mathrm{o}}\right). \quad (24)$$

If target j is identified at time t_k, *i.e.*, $j \in \mathcal{D}_{i,k}^{\mathrm{t}}$, then according to Section V-A, $\boldsymbol{x}_{j,k}^{\mathrm{t,g}} \in \mathbb{X}_{i,j,k}^{\mathrm{t,m}}$. If $j \in \mathcal{L}_{i,k-1}^{\mathrm{t}}$, then $\boldsymbol{x}_{j,k}^{\mathrm{t,g}} \in \mathbb{X}_{i,j,k|k-1}^{\mathrm{t}}$, and $\mathbb{X}_{i,j,k|k}^{\mathrm{t}}$ is updated as

$$\mathbb{X}_{i,j,k|k}^{\mathrm{t}} = \left(\mathbb{X}_{i,j,k|k-1}^{\mathrm{t}} \cap \mathbb{X}_{i,j,k}^{\mathrm{t,m}}\right) \setminus \left(\mathbb{P}_{i,k}^{\mathrm{g}}\left(\mathcal{Y}_{i,k}^{\mathrm{g}}\right) \cup \underline{\mathbb{S}}_{i,k|k}^{\mathrm{o}}\right), \quad (25)$$

Otherwise, if $j \notin \mathcal{L}_{i,k-1}^{\mathrm{t}}$, then $\boldsymbol{x}_{j,k}^{\mathrm{t,g}} \in \overline{\mathbb{X}}_{i,k|k-1}^{\mathrm{t}}$, and $\mathbb{X}_{i,j,k|k}^{\mathrm{t}}$ is updated as

$$\mathbb{X}_{i,j,k|k}^{\mathrm{t}} = \left(\overline{\mathbb{X}}_{i,k|k-1}^{\mathrm{t}} \cap \mathbb{X}_{i,j,k}^{\mathrm{t,m}}\right) \setminus \left(\mathbb{P}_{i,k}^{\mathrm{g}}\left(\mathcal{Y}_{i,k}^{\mathrm{g}}\right) \cup \underline{\mathbb{S}}_{i,k|k}^{\mathrm{o}}\right). \quad (26)$$

The list of identified targets becomes $\mathcal{L}_{i,k|k}^{\mathrm{t}} = \mathcal{L}_{i,k-1}^{\mathrm{t}} \cup \mathcal{D}_{i,k}^{\mathrm{t}}$.

The map $\mathcal{M}_{i,k|k}$ is updated from $\mathcal{M}_{i,k-1}$ as described in Section V-D.

C. Update after communication with neighbors

Using the information received from its neighbors, UAV i updates the list of identified targets $\mathcal{L}_{i,k}^{\mathrm{t}} = \bigcup_{\ell\in\mathcal{N}_{i,k}} \mathcal{L}_{\ell,k|k}^{\mathrm{t}}$ the r_s-neighborhoods of all obstacles $\underline{\mathbb{S}}_{i,k}^{\mathrm{o}} = \bigcup_{\ell\in\mathcal{N}_{i,k}} \underline{\mathbb{S}}_{\ell,k|k}^{\mathrm{o}}$ and the set of possible locations of targets still to be detected $\overline{\mathbb{X}}_{i,k}^{\mathrm{t}} = \bigcap_{\ell\in\mathcal{N}_{i,k}} \overline{\mathbb{X}}_{\ell,k|k}^{\mathrm{t}}$. Then, if $j \in \mathcal{L}_{i,k}^{\mathrm{t}}$, one has $\mathbb{X}_{i,j,k}^{\mathrm{t}} = \bigcap_{\ell\in\mathcal{N}_{i,k}} \mathbb{X}_{\ell,j,k|k}^{\mathrm{t}}$.

UAV i also updates its OEM as follow. For all $n \in \mathcal{N}^{\mathrm{M}}$, if there exists $\ell \in \mathcal{N}^{\mathrm{u}}$, such that $s_{\ell,k|k}(n) = \mathrm{G}$, then $s_{i,k}(n) = \mathrm{G}$ and $h_{i,k}(n) = 0$. For all $n \in \mathcal{N}^{\mathrm{M}}$, such that there does not exist any $\ell \in \mathcal{N}^{\mathrm{u}}$ with $s_{\ell,k|k}(n) = \mathrm{G}$, but such that there exists $\ell' \in \mathcal{N}^{\mathrm{u}}$ with $s_{\ell',k|k}(n) = \mathrm{O}$, then $s_{i,k}(n) = \mathrm{O}$ and $h_{i,k}(n) = \max_{\ell\in\mathcal{N}^{\mathrm{u}}} h_{\ell,k|k}(n)$. A cell considered as Ground by at least one of the UAVs is updated to Ground. For all other cells with status Obstacle for at least one UAV, the height is set to the maximum observed height.

VII. Design of the UAV guidance law

This section introduces a distributed MPC approach [26] based on that of [2]. Compared to [2], previously occluded areas and predicted occluded areas are taken into account in the design of the control inputs of the UAVs.

For a sequence $\mathbf{u}_{i,k:h} = (\mathbf{u}_{i,k}, \ldots, \mathbf{u}_{i,k+h-1}) \in \mathbb{U}^h$ of control inputs over a horizon of h steps, the state $\mathbf{x}^{\text{u,P}}_{i,k+h}$ of UAV i can be predicted using (2). Then, two types of occluded areas are considered: those which were previoulsy occluded and those which are likely to be occluded after applying $\mathbf{u}_{i,k:h}$ on UAV i. At time t_k, the occluded parts of the RoI for UAV i correspond to the set $\mathbb{P}^{\text{g}}_{i,k}(\mathcal{Y}^{\text{o}}_{i,k} \cup \mathcal{Y}^{\text{t}}_{i,k} \cup \mathcal{Y}^{\text{n}}_{i,k})$ introduced in (19). Accounting for the occluded parts for its neighbors, UAV i evaluates the set

$$\mathbb{H}^{\text{CVS}}_{i,k} = \bigcup\nolimits_{\ell \in \mathcal{N}_{i,k}} \mathbb{P}^{\text{g}}_{\ell,k}\left(\mathcal{Y}^{\text{o}}_{\ell,k} \cup \mathcal{Y}^{\text{t}}_{\ell,k} \cup \mathcal{Y}^{\text{n}}_{\ell,k}\right) \tag{27}$$

of all occluded parts. In the MPC approach, the sets $\mathbb{H}^{\text{CVS}}_{i,k}$ obtained during the last h time steps are gathered as

$$\mathbb{H}^{\text{CVS}}_{i,k-h+1:h} = \bigcup\nolimits_{t=k-h+1}^{k} \mathbb{H}^{\text{CVS}}_{i,t}, \tag{28}$$

and are considered as areas requiring further exploration. Predicting the map is difficult. Nevertheless, the version $\mathcal{M}_{i,k}$ of the map at time t_k can be used to evaluate the part $\mathbb{H}^{\text{m}}_{i,k}(\mathbf{x}^{\text{u,P}}_{i,k+h})$ of the FoV that may be hidden by obstacles if the predicted state of UAV i is $\mathbf{x}^{\text{u,P}}_{i,k+h}$ using [27].

For UAV i, considering $\mathbf{u}_{i,k:h}$, a prediction $\mathbf{x}^{\text{u,P}}_{i,k+\tau+1}$ of $\mathbf{x}^{\text{u}}_{i,k+\tau+1}$ at step $\tau \in \{0, \ldots, h-1\}$ can be performed using (2) and $\mathbf{u}_{i,k+\tau}$.

Iterative predictions and corrections are then performed for all $\tau \in \{1, \ldots, h\}$. Predicted versions $\mathbb{X}^{\text{t,P}}_{i,j,k+\tau|k+\tau-1}$, $j \in \mathcal{L}^{\text{t}}_{i,k}$, and $\overline{\mathbb{X}}^{\text{t,P}}_{i,k+\tau|k+\tau-1}$, of the set estimates $\mathbb{X}^{\text{t}}_{i,j,k+\tau|k+\tau-1}$, $j \in \mathcal{L}^{\text{t}}_{i,k}$ and $\overline{\mathbb{X}}^{\text{t}}_{i,k+\tau|k+\tau-1}$ are evaluated using (1).

To account for the future CVS information, a predicted FoV is then evaluated as $\mathbb{F}^{\text{P}}_{i,k+\tau} = \mathbb{F}(\mathbf{x}^{\text{u,P}}_{i,k+\tau}) \setminus \mathbb{H}^{\text{m}}_{i,k}(\mathbf{x}^{\text{u,P}}_{i,k+\tau})$. Then, as described in [2], the predicted sets $\mathbb{X}^{\text{t,P}}_{i,j,k+\tau}$ and $\overline{\mathbb{X}}^{\text{t,P}}_{i,k+\tau}$, for all $\tau \in \{1, \ldots, h\}$, are evaluated as

$$\mathbb{X}^{\text{t,P}}_{i,j,k+\tau} = \mathbb{X}^{\text{t,P}}_{i,j,k+\tau|k+\tau-1} \setminus \mathbb{F}^{\text{P}}_{i,k+\tau} \tag{29}$$

$$\overline{\mathbb{X}}^{\text{t,P}}_{i,k+\tau} = \overline{\mathbb{X}}^{\text{t,P}}_{i,k+\tau|k+\tau-1} \setminus \mathbb{F}^{\text{P}}_{i,k+\tau}. \tag{30}$$

The aim of the considered MPC approach is to evaluate a sequence $\mathbf{u}^{\text{MPC}}_{i,k:h}$ that minimizes

$$J\left(\mathbf{u}^{\text{MPC}}_{i,k:h}\right) = \Phi^{\mathbb{X}}_{i,k+h|k} + \lambda \Phi^{\text{d}}_{i,k+h|k} + \kappa \Phi^{\text{h}}_{i,k+h|k}, \tag{31}$$

where λ and κ are tuning parameters. The components of (31) are detailed in what follows.

$\Phi^{\mathbb{X}}_{i,k+h|k}$ is the area of the predicted set estimates $\mathbb{X}^{\text{t,P}}_{i,j,k+h}$, $j \in \mathcal{L}^{\text{t}}_{i,k}$, and $\overline{\mathbb{X}}^{\text{t,P}}_{i,k+h}$ expressed as

$$\Phi^{\mathbb{X}}_{i,k+h|k} = \phi\left(\overline{\mathbb{X}}^{\text{t,P}}_{i,k+h} \cup \bigcup\nolimits_{j \in \mathcal{L}^{\text{t}}_{i,k}} \mathbb{X}^{\text{t,P}}_{i,j,k+h}\right). \tag{32}$$

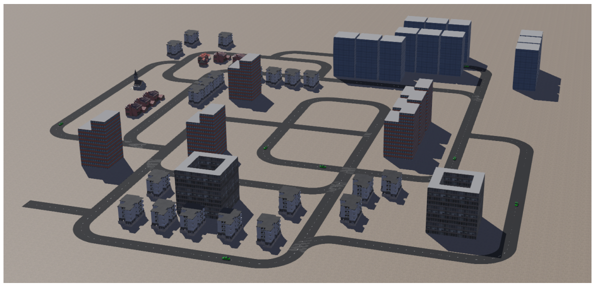

Fig. 5. Simulated urban environment in Webots

$\Phi^{\text{d}}_{i,k+h|k}$ is the distance between the predicted location $\boldsymbol{x}^{\text{u,P}}_{i,k+h}$ of UAV i and the closest predicted set estimate

$$\begin{aligned}\Phi^{\text{d}}_{i,k+h|k} =& d\left(\boldsymbol{p}_{\text{g}}\left(\boldsymbol{x}^{\text{u,P}}_{i,k+h}\right), \overline{\mathbb{X}}^{\text{t,P}}_{i,k+h} \cup \bigcup\nolimits_{j \in \mathcal{L}^{\text{t}}_{i,k}} \mathbb{X}^{\text{t,P}}_{i,j,k+h}\right.\\ &\left.\setminus \mathbb{H}^{\text{m}}_{i,k}(\mathbf{x}^{\text{u,P}}_{i,k+h})\right)\end{aligned}$$

When $\Phi^{\mathbb{X}}_{i,k+h|k}$ remains constant whatever $\mathbf{u}_{i,k:h}$, $\Phi^{\text{d}}_{i,k+h|k}$ is used to drive UAV i toward the closest set estimate. Taking into account $\mathbb{H}^{\text{m}}_{i,k}(\mathbf{x}^{\text{u,P}}_{i,k+h})$ avoids UAV i being attracted by a set estimate that would be hidden by an obstacle. Finally,

$$\Phi^{\text{h}}_{i,k+h|k} = \phi\left(\overline{\mathbb{X}}^{\text{t,P}}_{i,k+h} \cup \bigcup\nolimits_{j \in \mathcal{L}^{\text{t}}_{i,k}} \mathbb{X}^{\text{t,P}}_{i,j,k+h} \cap \mathbb{H}^{\text{CVS}}_{i,k-h+1:h}\right).$$

accounts for the parts of $\overline{\mathbb{X}}^{\text{t,P}}_{i,k+h}$ that were previously occluded and favors exploration of such area.

A round-robin evaluation of the guidance law of each UAV is considered. The first UAV evaluates independently its control input and broadcasts it as well as information related to its predicted FoV $\mathbb{F}^{\text{P}}_{1,k+\tau}$, $\tau = 1, \ldots, h$. The following UAV may then account for the information sent by the first UAV to evaluate $\mathbb{X}^{\text{t,P}}_{i,j,k+\tau}$ and $\overline{\mathbb{X}}^{\text{t,P}}_{i,k+\tau}$ using its own field of view $\mathbb{F}^{\text{P}}_{2,k+\tau}$ and $\mathbb{F}^{\text{P}}_{1,k+\tau}$, $\tau = 1, \ldots, h$...

VIII. Simulations

The simulations are performed using Webots[1] to generate the measurements collected by the UAVs. Webots is interfaced with Matlab to evaluate all set estimates and sequences of UAV control inputs.

A. Simulation conditions

The RoI $\mathbb{X}_0 = [-250\,\text{m}, 250\,\text{m}] \times [-250\,\text{m}, 250\,\text{m}] \times \mathbb{R}^+$ contains several buildings covering about 5 % of the area of the ground $\mathbb{X}_{\text{g}}$, see Figure 5. $N^{\text{t}} = 8$ target cars have to be localized. The dimension of each car is 4.6 m × 1.8 m × 1.5 m. A cylinder $\mathbb{C}^{\text{t}}$ of radius $r^{\text{t}} = 2.5$ m and height $h^{\text{t}} = 2$ m contains the unknown shape of the cars. The safety distance r_{s} between targets and obstacles is taken equal to 3 m. The target dynamic (1) is such that the cars stay on the road. Their speed is randomly chosen but is bounded by a known maximal speed. The UAVs do not exploit any knowledge related to the roads during the search. The OEM $\mathcal{M}_{i,k}$ consists of square cells of 5×5 m^2.

A fleet of 4 identical quadcopters is considered. Their dynamics is detailed in [3]. The CVS of each UAV provides

[1] https://cyberbotics.com/doc/guide/index

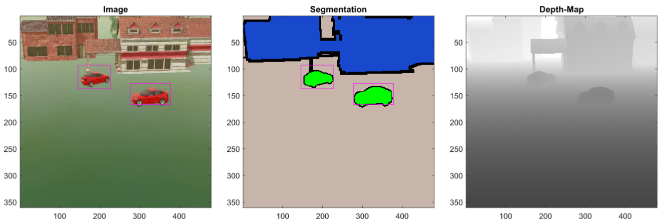

Fig. 6. Example of CVS information provided by Webots: Image and bounding boxes (left), image labels and bounding boxes (middle), and depth map (right)

labeled images of $N_\mathrm{r} \times N_\mathrm{c} = 360 \times 480$ pixels and a depth-map of the same size. The camera aperture angle is $\pi/4$ rad. The bounds of the noise w in (5) are $[\underline{w}, \overline{w}] = [-0.01, 0.01]$. Figure 6 shows an example of the information provided by the CVS of each UAV.

For the MPC approach of Section VII, $h = 2$ and only yaw angle increments are considered in $\mathbb{U} = \{-\pi/3, -\pi/6, 0, \pi/6, \pi/3\}$. The weights in the cost function (31) are taken as $\lambda = 10^{-5}$ (as in [2]) and $\kappa = 1$. The yaw angle increments provided by the MPC algorithm serve to deduce a reference state for a low-level PID controller also designed to maintain a flight at a constant speed module and at a constant altitude. To avoid collisions, the flight height is different for each UAV, and is larger than the height of all obstacles. Furthermore, the communication between UAVs is assumed perfect, and no restriction regarding the communication range is considered. The UAV speed and car speed are chosen such as the speed ratio between the UAVs and the targets is equal or above 5.

The CVS provides new information with a period $T = 0.5$ s. The estimation algorithms of Sections V and VI are applied with the same period. Nevertheless, the guidance law updated is set at 3 s to ensure that the UAVs have reached their yaw angle reference provided by the MPC algorithm.

B. Results

Results are averages over 10 independent simulations lasting 300 s of UAV flight. The video[2] illustrates a typical evolution of the sets $\mathbb{X}^{\mathrm{t}}_{i,j,k}$ gathering all the possible locations of identified targets (in green), the set $\overline{\mathbb{X}}^{\mathrm{t}}_{i,k}$ of possible locations of targets still to be identified (in yellow), and the set $\mathbb{S}^{\mathrm{o}}_{i,k}$ representing the inner-approximation of the r_s-neighborhood of the encountered obstacles (in black). The set $\overline{\mathbb{H}}^{\mathrm{CVS}}_{i,k}$ (in blue), represents the points of $\mathbb{X}_\mathrm{g}$ that have been occluded at some previous time instant due by an obstacle and were not observed up to time t_k.

Figure 7 shows $\mathcal{M}_{i,300}$ at the end of one simulation of 300 s. The perspective is the same as in Figure 5. $\mathcal{M}_{i,300}$ provides an approximation of obstacle shapes and heights sufficient to predict the occluded parts of the environment.

The metrics considered to evaluate the performance of the proposed algorithm are the area explored $\Phi(\mathbb{X}^{\mathrm{g}}_k) = \phi(\bigcup_{t=1}^{k} \bigcup_{\ell \in \mathcal{N}_{i,t}} \mathbb{P}^{\mathrm{g}}_{\ell,t}(\mathcal{Y}^{\mathrm{g}}_{\ell,t}))$, the Unexplored Area (UA) $\phi(\overline{\mathbb{X}}^{\mathrm{t}}_{i,k})$, the Occluded Area (OA) $\phi(\overline{\mathbb{X}}^{\mathrm{t}}_{i,k} \cap \overline{\mathbb{H}}^{\mathrm{CVS}}_{i,k})$, the average area of target set estimates $\Phi(\mathcal{X}^{\mathrm{t}}_{i,k}) = \sum_{j \in \mathcal{L}^{\mathrm{t}}_{i,k}} \phi(\mathbb{X}^{\mathrm{t}}_{i,j,k})/\mathrm{card}(\mathcal{L}^{\mathrm{t}}_{i,k})$, and the localization error $e(\mathcal{X}^{\mathrm{t}}_{i,k}) = \sum_{j \in \mathcal{L}^{\mathrm{t}}_{i,k}} ||\boldsymbol{x}^{\mathrm{t,g}}_{j,k} - c(\mathbb{X}^{\mathrm{t}}_{i,j,k})||/\mathrm{card}(\mathcal{L}^{\mathrm{t}}_{i,k})$ between the actual target location and the barycenter of its set estimate.

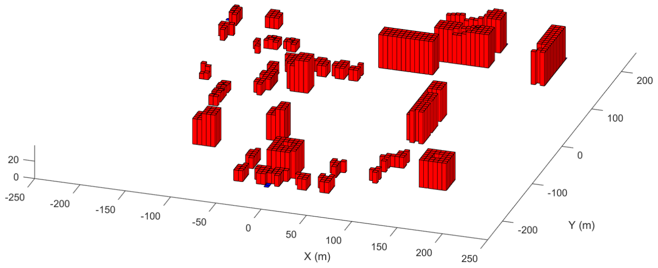

Fig. 7. Illustration of $\mathcal{M}_{i,300}$. The blue squares are the unexplored cells. The red rectangles are the occupied cells with their estimated height.

The evolution of these metrics is shown in Figure 8. $\Phi\left(\mathbb{X}^{\mathrm{g}}_k\right)$ (in purple) is the area of the map proved to be ground. Since around 5% of $\mathbb{X}_\mathrm{g}$ is occupied by obstacles, the maximum value of $\Phi\left(\mathbb{X}^{\mathrm{g}}_k\right)$ is 95%. After 300 s, $\Phi\left(\mathbb{X}^{\mathrm{g}}_k\right)$ reaches 94 %, meaning that the fleet has explored nearly all the RoI. The remaining 1 % corresponds to parts of $\mathbb{X}_\mathrm{g}$ located around buildings at the border of the RoI.

$\Phi(\mathcal{X}^{\mathrm{t}}_{i,k})$ (in green) stabilizes around $0.23\,\%$ of $\phi\left(\mathbb{X}_\mathrm{g}\right)$ with a standard deviation of $0.14\,\%$. This corresponds to the surface of a disc with radius between 8.3 m and 17.2 m. The area is quite large: As targets cannot be permanently observed, the uncertainty of their estimated location increases. Nevertheless, as seen in Figure 8-B, the localization error $e\left(\mathcal{X}^{\mathrm{t}}_{i,k}\right)$ (in red) is about $1.0\,\mathrm{m} \pm 0.3\,\mathrm{m}$. Thus, despite knowing only the target maximal speed, the UAV fleet manages to accurately localize the identified targets.

The area $\phi(\overline{\mathbb{X}}^{\mathrm{t}}_{i,k})$ (in orange in Figure 8-A) of the set estimate $\overline{\mathbb{X}}^{\mathrm{t}}_{i,k}$ containing the locations of targets still to be detected, decreases and reaches $19.1\,\% \pm 2.6\,\%$. At the end of the simulation, almost 60 % of $\overline{\mathbb{X}}^{\mathrm{t}}_{i,k}$ belong to $\overline{\mathbb{H}}^{\mathrm{CVS}}_{i,k}$ (in blue), and thus correspond to a portion of $\overline{\mathbb{X}}^{\mathrm{t}}_{i,k}$ that cannot be eliminated because of the presence of obstacles. This part of $\overline{\mathbb{X}}^{\mathrm{t}}_{i,k}$ continues to grow until observed from another point of view. These results illustrate the limitations of the criterion (31) in the design of a guidance law that would bypass obstacles to observe previously occluded regions.

To evaluate the benefits provided by the OEM in the design of the guidance law and the components of the cost function (31), the results of the proposed approach are compared to those obtained when $\kappa = 0$ in (31) (the occluded parts are not considered) and to those in [3] (the map is not exploited in the MPC algorithm), see Table I. Exploiting the OEM in the design of the guidance laws improves the exploration performance compared to [3]. Moreover, accounting for the previously occluded parts in (31) ($\kappa = 1$) improves the reduction of $\phi(\overline{\mathbb{X}}^{\mathrm{t}}_{i,k})$ and results in a better exploration of the environment since $\phi(\overline{\mathbb{X}}^{\mathrm{t}}_{i,k} \cap \mathbb{X}^{\mathrm{h}}_{i,k})$ is also reduced.

[2] https://nextcloud.centralesupelec.fr/s/eYD2CDgr5qxpee3

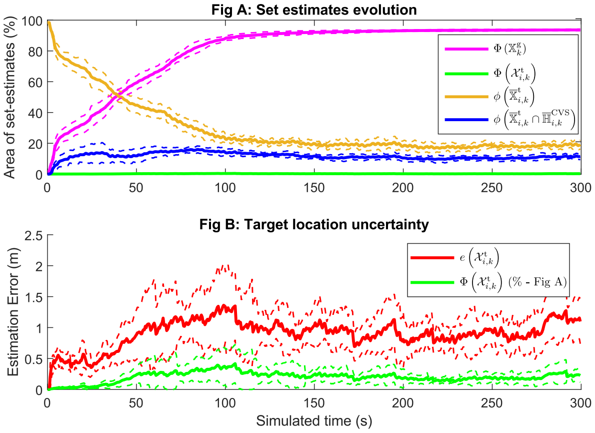

Fig. 8. Mean value with standard deviation of $\Phi(\mathbb{X}^{\mathrm{g}}_{k})/\phi(\mathbb{X}_{\mathrm{g}})$ (purple), $\Phi(\mathcal{X}^{\mathrm{t}}_{i,k})/\phi(\mathbb{X}_{\mathrm{g}})$ (green), $\phi(\overline{\mathbb{X}}^{\mathrm{t}}_{i,k})/\phi(\mathbb{X}_{\mathrm{g}})$ (orange), $\phi(\overline{\mathbb{X}}^{\mathrm{t}}_{i,k}\cap\overline{\mathbb{H}}^{\mathrm{CVS}}_{i,k})/\phi(\mathbb{X}_{\mathrm{g}})$ (blue) and $e(\mathcal{X}^{\mathrm{t}}_{i,k})$ (red).

	[3]	Here	
		$(\kappa = 0)$	$(\kappa = 1)$
$\phi(\overline{\mathbb{X}}^{\mathrm{t}}_{i,k})$ (%)	25.5 ± 2.7	20.9 ± 2.8	$\mathbf{19.1 \pm 2.6}$
$\phi(\overline{\mathbb{X}}^{\mathrm{t}}_{i,k} \cap \overline{\mathbb{H}}^{\mathrm{CVS}}_{i,k})$ (%)	16.8 ± 1.9	12.8 ± 2.1	$\mathbf{11.2 \pm 1.8}$

TABLE I
PERFORMANCE COMPARISON (AVERAGE OVER 10 INDEPENDENT SIMULATIONS)

IX. CONCLUSION

This paper proposes an approach to search and track ground targets evolving in an unknown but structured environment using a fleet of cooperating UAVs. Information provided by a CVS are exploited by a set-membership estimator to get sets containing the location of detected targets, a set containing the locations of targets still to detect, and a set proved to contain no target.

In the proposed approach, each UAV builds an OEM of the environment in parallel with the target search, to approximate the obstacles present in the environment. The map is used in a MPC algorithm to design guidance laws for the UAVs. The portion of the environment that has been occluded in the past is also used to favor the exploration of these areas.

The proposed approach, evaluated in a simulated urban environment, yields better search and tracking performance than a variant where no map is used. Future work will consider adaptations to more complex environments such as urban canyons, requiring to avoid obstacles.

REFERENCES

[1] M. Lyu, Y. Zhao, C. Huang, and H. Huang, "Unmanned aerial vehicles for search and rescue: A survey," *Remote Sensing*, vol. 15, no. 13, p. 3266, 2023.

[2] J. Ibenthal, M. Kieffer, L. Meyer, H. Piet-Lahanier, and S. Reynaud, "Bounded-error target localization and tracking using a fleet of UAVs," *Automatica*, vol. 132, pp. 109809–109824, 2021.

[3] M. Zagar, L. Meyer, M. Kieffer, and H. Piet-Lahanier, "Set-membership target search and tracking within an unknown cluttered area using cooperating UAVs equipped with vision systems," *ArXiv preprint arXiv:2403.15113*, 2024.

[4] A. Souza and L. M. Gonçalves, "Occupancy-elevation grid: an alternative approach for robotic mapping and navigation," *Robotica*, vol. 34, no. 11, pp. 2592–2609, 2016.

[5] A. Goldhoorn, A. Garrell, R. Alquézar, and A. Sanfeliu, "Searching and tracking people with cooperative mobile robots," *Autonomous Robots*, vol. 42, no. 4, pp. 739–759, 2018.

[6] P. M. Dames, "Distributed multi-target search and tracking using the PHD filter," *Autonomous robots*, vol. 44, no. 3, pp. 673–689, 2020.

[7] A. A. Meera, M. Popovic, A. Millane, and R. Siegwart, "Obstacle-aware adaptive informative path planning for uav-based target search," in *Proc. IEEE ICRA*, pp. 718–724, 2019.

[8] Y. Ji, Y. Zhao, B. Chen, Z. Zhu, Y. Liu, H. Zhu, and S. Qiu, "Source searching in unknown obstructed environments through source estimation, target determination, and path planning," *Building and Environment*, vol. 221, pp. 109266–109306, 2022.

[9] J. A. Placed, J. Strader, H. Carrillo, N. Atanasov, V. Indelman, L. Carlone, and J. A. Castellanos, "A survey on active simultaneous localization and mapping: State of the art and new frontiers," *IEEE Trans. Robotics*, vol. 39, no. 3, pp. 1686–1705, 2023.

[10] J. Ibenthal, L. Meyer, H. Piet-Lahanier, and M. Kieffer, "Localization of Partially Hidden Moving Targets Using a Fleet of UAVs via Bounded-Error Estimation," *IEEE Trans. Robotics*, vol. 39, no. 6, pp. 4211–4229, 2023.

[11] L. Zhao, R. Li, J. Han, and J. Zhang, "A distributed model predictive control-based method for multidifferent-target search in unknown environments," *IEEE Trans. Evol. Comput.*, vol. 27, no. 1, pp. 111–125, 2022.

[12] B. Zhang, X. Lin, Y. Zhu, J. Tian, and Z. Zhu, "Enhancing Multi-UAV Reconnaissance and Search Through Double Critic DDPG With Belief Probability Maps," *IEEE Trans. Intelligent Vehicles*, vol. 9, no. 2, pp. 3827–3842, 2024.

[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.*, vol. 49, no. 11, pp. 2362–2372, 2018.

[14] D. Liu, W. Bao, X. Zhu, B. Fei, Z. Xiao, and T. Men, "Vision-aware air-ground cooperative target localization for UAV and UGV," *Aerosp. Sci. Technol.*, vol. 124, pp. 107525–107540, 2022.

[15] T. M. Di Gennaro and J. Waldmann, "Sensor Fusion with Asynchronous Decentralized Processing for 3D Target Tracking with a Wireless Camera Network," *Sensors*, vol. 23, no. 3, pp. 1194–1228, 2023.

[16] X. Luo, Y. Wu, and L. Zhao, "YOLOD: A target detection method for UAV aerial imagery," *Remote Sensing*, vol. 14, no. 14, p. 3240, 2022.

[17] P. Jiang, D. Ergu, F. Liu, Y. Cai, and B. Ma, "A Review of Yolo algorithm developments," *Procedia computer science*, vol. 199, pp. 1066–1073, 2022.

[18] Z. Fan, Y. Zhu, Y. He, Q. Sun, H. Liu, and J. He, "Deep learning on monocular object pose detection and tracking: A comprehensive overview," *ACM Computing Surveys*, vol. 55, no. 4, pp. 1–40, 2022.

[19] B. Allik, "Tracking of multiple targets across distributed platforms with fov constraints," in *Proc. IEEE CDC*, pp. 6044–6049, 2019.

[20] Y. Hou, J. Zhao, R. Zhang, X. Cheng, and L. Yang, "UAV Swarm Cooperative Target Search: A Multi-Agent Reinforcement Learning Approach," *IEEE Trans. Intelligent Vehicles*, vol. 9, no. 1, pp. 568–578, 2023.

[21] A. Symington, S. Waharte, S. Julier, and N. Trigoni, "Probabilistic target detection by camera-equipped UAVs," in *Proc. IEEE ICRA*, pp. 4076–4081, 2010.

[22] L. Madhuanand, F. Nex, and M. Y. Yang, "Self-supervised monocular depth estimation from oblique UAV videos," *ISPRS journal of photogrammetry and remote sensing*, vol. 176, pp. 1–14, 2021.

[23] S. Minaee, Y. Boykov, F. Porikli, A. Plaza, N. Kehtarnavaz, and D. Terzopoulos, "Image segmentation using deep learning: A survey," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 7, pp. 3523–3542, 2021.

[24] O. Faugeras, *Three-dimensional computer vision: a geometric viewpoint*. MIT press, 1993.

[25] I.-F. Kenmogne, V. Drevelle, and E. Marchand, "Cooperative localization of drones by using interval methods," *Acta Cybernetica*, pp. 1–16, 2019.

[26] P. D. Christofides, R. Scattolini, D. M. De La Pena, and J. Liu, "Distributed model predictive control: A tutorial review and future research directions," *Comput Chem Eng*, vol. 51, pp. 21–41, 2013.

[27] L. Reboul, M. Kieffer, H. Piet-Lahanier, and S. Reynaud, "Cooperative guidance of a fleet of UAVs for multi-target discovery and tracking in presence of obstacles using a set membership approach," *IFAC-PapersOnLine*, vol. 52, no. 12, pp. 340–345, 2019.