Voronoi-Based Coverage Control of Pan/Tilt/Zoom Camera Networks

Omur Arslan, Hancheng Min, and Daniel E. Koditschek

***Abstract*— A challenge of pan/tilt/zoom (PTZ) camera networks for efficient and flexible visual monitoring is automated active network reconfiguration in response to environmental stimuli. In this paper, given an event/activity distribution over a convex environment, we propose a new provably correct reactive coverage control algorithm for PTZ camera networks that continuously (re)configures camera orientations and zoom levels (i.e., angles of view) in order to locally maximize their total coverage quality. Our construction is based on careful modeling of visual sensing quality that is consistent with the physical nature of cameras, and we introduce a new notion of *conic Voronoi diagrams*, based on our sensing quality measures, to solve the camera network allocation problem: that is, to determine where each camera should focus in its field of view given all the other cameras' configurations. Accordingly, we design simple greedy gradient algorithms for both continuous- and discrete-time first-order PTZ camera dynamics that asymptotically converge a locally optimal coverage configuration. Finally, we provide numerical and experimental evidence demonstrating the effectiveness of the proposed coverage algorithms.**

I. Introduction

With decreasing hardware and setup costs, (controllable) cameras have been used more than ever for situational awareness, security, surveillance and environmental monitoring. However, increasing camera network size and vast amount of collected data make camera control and data interpretation an impossibily difficult task for a human operator and thus necessitates automated camera network reconfiguration and data mining methods for increased efficiency, flexibility and functionality [1], [2]. In this paper, assuming a given event/activity distribution over a known convex environment, we consider the automatic reconfiguration problem of pan/tilt/zoom (PTZ) camera networks, and we propose a new simple greedy coverage control algorithm for a group of PTZ cameras in order to increase their collective coverage performance.

A. Motivation and Prior Literature

Finding a globally optimal coverage configuration of a sensor network is usually very difficult. For example, the art gallery problem [3] and many related facility localization (e.g., p-center and p-median) problems [4] are known to be NP hard. Hence, to mitigate this complexity, greedy gradient methods, usually combined with an expectation-maximization strategy, are utilized for finding a locally optimal sensor network configuration.

Omur Arslan, Hancheng Min, and Daniel E. Koditschek are with the Department of Electrical and Systems Engineering, University of Pennsylvania, Philadelphia, PA 19104, USA. E-mail: {omur, hanchmin, kod}@seas.upenn.edu.

This work was supported in part by AFRL grant FA865015D1845 (subcontract 669737-1).

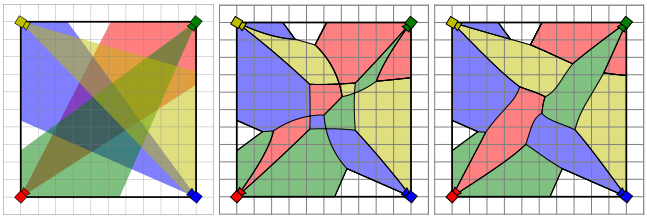


Fig. 1. Sensor allocation in camera networks via conic Voronoi diagrams, which determines where each camera should focus in its field of view given all the other cameras' configurations: (left) a camera network configuration, (middle) its sensor allocation based on perspective quality in (1), (right) its sensor allocation based on a multiplicative combination of perspective and resolution qualities in (4a) (with parameters $R = 0$, $\kappa = 0$ and $\sigma = 5$).

A widely used class of such gradient methods leverages Voronoi diagrams to decouple the problems of optimal sensory task assignment and sensor configuration optimization, and finds a locally optimal coverage configuration by following a (continuous) expectation-maximization approach. For example, consider a mobile sensor network consisting of isotropic sensors whose sensing cost of an event is given by the (squared) Euclidean distance between the sensor position and the event location. For such a mobile isotropic sensor network, a simple coverage strategy that continuously steers each sensor towards its associated Voronoi centroid is known to asymptotically converge to a locally optimal coverage configuration [5]. Our construction is inspired by this "move-to-centroid" coverage control law of mobile isotropic sensor networks [5], and has similar qualitative properties and an interpretation (i.e., we introduce new "move-to-centroidal-perspective" and "move-to-centroidal-angle-of-view" laws) for PTZ camera networks for controlling camera orientations and angles of view.

Extensions of Voronoi-based coverage control algorithms of mobile isotropic sensor networks for anisotropic sensors have received significant attention in the past literature in order to address the role of sensor orientation in sensing performance, because many physical sensors, such as cameras, radars, acoustic and ultrasonic sensors, have such a directional performance dependency [6]–[8]. A straightforward extension is to limit the standard Voronoi cell (constructed based on the Euclidean metric) of an anisotropic sensor to its directional (e.g., bounded conic) sensory footprint while still assuming an isotropic sensing performance within the sensory footprint [6]. A more elegant extension is to find a homeomorphism that maps a directional (ellipsoidal or bounded conic) sensory footprint to a circular sensory footprint and use the associated pull-back metric of the Euclidean metric to construct generalized Voronoi diagrams, such as anisotropic [7] or directional [8] Voronoi diagrams.

Although, these extensions take into account the effect of sensory orientation on sensing performance, such generic sensor models are not descriptive enough to model physical operation principles (e.g. resolution constraints, lens distortion) of cameras and performance limitations of computer vision algorithms, which usually don't have a simple (monotonic) functional relation with the Euclidean metric [9], [10]. As an alternative, in this paper we introduce a new form of generalized Voronoi diagrams, called *conic Voronoi diagrams*, based on accurate camera sensing quality measures, to solve the sensor allocation problem in camera networks, as illustrated in Fig. 1.

Finally, it also bears mentioning that most existing work on coverage control of mobile (e.g., flying) cameras [11]–[13] mainly concerns with optimal active camera placement and simplifies the directionality issues of camera sensors by assuming a certain (e.g., downward facing) camera orientation. More accurate and descriptive camera models are usually employed in the design of probabilistic coverage algorithms, based on joint detection probability [9], [14], [15] and Gaussian mixture models [16], [17]. However, these probabilistic methods are inherently centralized whereas Voronoi-based methods usually offer distributed computation. To best of our knowledge, this is the first work that shows how to leverage accurate camera spatial sensing models in Voronoi-based visual coverage settings. Our ultimate future goal is to enable optimal orientation and zoom control in mobile (e.g., flying) PTZ camera networks for active visual monitoring.

B. Contributions and Organization of the Paper

This paper proposes a new simple reactive coverage control algorithm for PTZ camera networks that continuously (re)configures camera orientations and zoom levels (i.e., angles of view) in order to maximize their collective coverage of a given event/activity distribution over a known convex environment. Our construction is based on newly introduced accurate camera spatial sensing quality measures that conform to the physical sensing (e.g., directionality, field of view, and resolution) constraints of camera sensors and limitations of visual sensing algorithms. Based on these sensing quality measures, we propose a new notion of *conic Voronoi diagrams* to determine optimal resource allocation in camera networks. We derive greedy gradient methods, for both continuous- and discrete-time first-order PTZ camera dynamics, that asymptotically bring cameras to an optimal coverage configuration where the total coverage quality of the cameras is locally maximized.

The rest of the paper is organized as follows. Section II presents coverage optimization of PTZ camera networks based on a new class of camera spatial sensing quality measures. Section III constructs and analyzes coverage control algorithms for continuous- and discrete-time PTZ camera dynamics. In Section IV and Section V, we illustrate the effectiveness of the proposed coverage algorithms using numerical simulations and experimental results, respectively. Section VI concludes with a summary of our contributions and a brief discussion of future directions.

II. Coverage Optimization of PTZ Cameras

In this section, we first introduce a new class of spatial sensing quality measures for PTZ cameras and present their use in coverage optimization of PTZ camera networks. In particular, we show that optimal sensor allocation is given by conic Voronoi diagrams constructed based on these quality measures, and we provide an explicit characterization of locally optimal coverage configurations.

A. Spatial Sensing Models for PTZ Cameras

For ease of exposition, we consider an ideal pan/tilt/zoom (PTZ) camera model, with a conic field of view in $n \geq 2$ dimensional Euclidean space, whose configuration is specified by its fixed position $\mathrm{p} \in \mathbb{R}^n$, variable optical-axis direction[1] $\mathrm{v} \in \mathbb{S}^{n-1}$ and adjustable angle of view $2\alpha \in (0, \pi)$ (corresponding to changeable zoom level). A nice feature of PTZ cameras is that they offer variable fields of view for optimizing their visual sensing performance. Thus, in order to model the effect of camera orientation and angle of view (i.e., zoom level) on the camera's spatial sensing performance, we consider the following two quality measures:

1) Perspective Quality: A key factor determining the performance of many visual detection, tracking, and localization algorithms is view perspective. Visual perception quality is known to usually deteriorate away from the optical axis of a camera towards the boundary of its field of view because of increased lens distortion, incomplete and nonpersistent visual data [10]. Accordingly, we define the camera's perspective sensing quality of an event at location $\mathrm{x} \in \mathbb{R}^n$ to be[2]

$$q_{\text{pers}}(\mathrm{x}) := \frac{1}{1-\cos\alpha}\left(\frac{(\mathrm{x}-\mathrm{p})^{\mathrm{T}}\mathrm{v}}{\|\mathrm{x}-\mathrm{p}\|} - \cos\alpha\right), \qquad (1)$$

where $\|.\|$ denotes the standard Euclidean norm, and $.^{\mathrm{T}}$ denotes the transpose operator. Note that the perspective quality measure is normalized to yield values in the interval $[-1, 1]$, where a negative value corresponds to an event location not being in the angle of view of the camera.

2) Resolution Quality: Similar to view perspective, the performance of many computer vision algorithms depends on the pixel density at which an event is observed by a camera. If our cameras have a circular image sensor with $N \in \mathbb{N}$ pixels along the diameter, then the pixel density (resolution) r, in pixel per unit length[3], of a visible event at location $\mathrm{x} \in \mathbb{R}^d$ is given by $r = \frac{N}{2\alpha\|\mathrm{x}-\mathrm{p}\|}$. Hence, the pixel density of an event increases with decreasing angle of view (i.e., zooming-in) and with decreasing relative distance to the camera position [10]. However, a large value of pixel density does not necessarily improve visual sensing performance, because the event can be very close to the camera or the zoom level can be very high, which might cause the event to occupy a significant fraction of the camera image [9]. Thus, assuming a desired sensing resolution $r^* > 0$, as natural

[1] $\mathbb{S}^n := \left\{\mathrm{x} \in \mathbb{R}^n \middle| \|\mathrm{x}\| = 1\right\}$ is the $(n-1)$-dimensional unit sphere embedded in n-dimensional Euclidean space $\mathbb{R}^n$. Moreover, $\mathbb{R}_{\geq 0}$ and $\mathbb{R}_{>0}$ denote the sets of nonnegative and positive real numbers, respectively.

[2] To resolve the indeterminacy, we set $\frac{\mathrm{x}}{\|\mathrm{x}\|} = 0$ for $\mathrm{x} = \mathbf{0}$.

[3] If the camera has a circular image sensor with N^2 pixels, then the pixel density (resolution) r, in pixel per unit area, of a visible event at location $\mathrm{x} \in \mathbb{R}^n$ is given by $r = \frac{N^2}{\pi\alpha^2\|\mathrm{x}-\mathrm{p}\|}$.

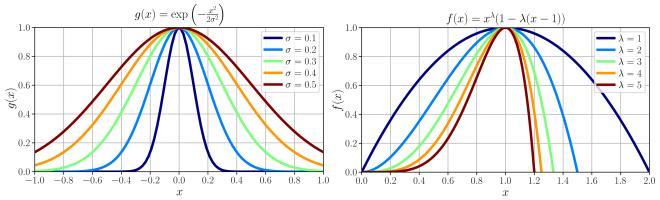


Fig. 2. Resolution quality distribution models: (left) Gaussian function $g(x) = \exp\left(-\frac{x^2}{2\sigma^2}\right)$, (right) beta function $f(x) = x^\lambda(1 - \lambda(x-1))$.

candidates in Gaussian and beta functional forms, illustrated in Fig. 2, one might consider the following *unlimited-* and *limited-range resolution quality measures* for quantifying sensing performance of an event at location $\mathrm{x} \in \mathbb{R}^n$, with respect to the desired sensing depth $\frac{N}{2r^*\alpha}$,

$$Q_{\text{res}}(\mathrm{x}) := \exp\left(-\frac{\left(\|\mathrm{x}-\mathrm{p}\| - \frac{N}{2r^*\alpha}\right)^2}{2\sigma^2}\right), \tag{2a}$$

$$\widehat{Q}_{\text{res}}(\mathrm{x}) := \frac{\|\mathrm{x}-\mathrm{p}\|^\lambda}{\left(\frac{N}{2r^*\alpha}\right)^{\lambda+1}}\left(\frac{N}{2r^*\alpha} - \lambda\left(\|\mathrm{x}-\mathrm{p}\| - \frac{N}{2r^*\alpha}\right)\right), \tag{2b}$$

where $\sigma > 0$ and $\lambda \geq 0$ are fixed scalar terms determining the spatial variance around the desired depth $\frac{N}{2r^*\alpha}$. Note that Q_{res} and $\widehat{Q}_{\text{res}}$ both take their maximum value of unity at the desired sensing depth, i.e., when $\|\mathrm{x}-\mathrm{p}\| = \frac{N}{2r^*\alpha}$. Also observe that while the unlimited-range quality measure Q_{res} always returns a nonegative value, the limited-range quality measure $\widehat{Q}_{\text{res}}$ is nonnegative if and only if $\|\mathrm{x}-\mathrm{p}\| \leq \frac{\lambda+1}{\lambda}\frac{N}{2r^*\alpha}$, which defines its effective range.

Although the quality measures in (2) are very natural selections with standard (Gaussian and beta) functional forms, their high nonlinearity complicates the design and analysis of gradient coverage algorithms for continuous- and discrete-time camera dynamics: in particular, it is not straightforward to explicit determine locally optimal coverage configurations (i.e., critical points) of the resulting gradient fields, and one needs to select a proper step-size to follow the resulting gradient dynamics while ensuring the positive invariance of half the camera's angle of view α in the interval $\left(0, \frac{\pi}{2}\right)$. To mitigate such technical issues, using the similarity of monotonic properties of $\frac{1}{\alpha}$ and $\cos\alpha$ over the interval $\left(0, \frac{\pi}{2}\right)$, we define the following *simplified unlimited-* and *limited-range resolution quality measures* for an event at $\mathrm{x} \in \mathbb{R}^n$,

$$q_{\text{res}}(\mathrm{x}) := \cos^\kappa\alpha \, \exp\left(-\frac{(\|\mathrm{x}-\mathrm{p}\| - R)^2}{2\sigma^2}\right), \tag{3a}$$

$$\widehat{q}_{\text{res}}(\mathrm{x}) := \frac{\|\mathrm{x}-\mathrm{p}\|^\lambda}{R^{\lambda+1}}\Big(R\cos\alpha - \lambda\big(\|\mathrm{x}-\mathrm{p}\| - R\cos\alpha\big)\Big), \tag{3b}$$

where $R > 0$ a fixed desired sensing range, and $\sigma > 0$, $\kappa > 0$ and $\lambda \geq 0$ are fixed scalar parameters that model spatial resolution variability. Note that q_{res} is nonnegative and has a maximum value of $(\cos\alpha)^\kappa$ when $\|\mathrm{x}-\mathrm{p}\| = R$, whereas $\widehat{q}_{\text{res}}$ is nonnegative iff $\|\mathrm{x}-\mathrm{p}\| \leq \frac{\lambda+1}{\lambda}R\cos\alpha$ and has a maximum value of $(\cos\alpha)^{\lambda+1}$ when $\|\mathrm{x}-\mathrm{p}\| = R\cos\alpha$. Hence, it is important to remark that the depth at which $\widehat{q}_{\text{res}}$ is maximized depends on the camera's angle of view (i.e., zoom level) as it is the case in practice, but q_{res} lacks modeling of such a dependency. This makes $\widehat{q}_{\text{res}}$ a more accurate spatial sensing model over q_{res}, especially for cameras with adjustable angle of view (i.e., zoom level).

Therefore, as illustrated in Fig. 3, the overall spatial sensing quality of a camera can be quantified by using a multiplicative combination of perspective and resolution quality measures, in (1) and (3), respectively, as follows,

$$q(\mathrm{x}) := q_{\text{pers}}(\mathrm{x})\, q_{\text{res}}(\mathrm{x}), \tag{4a}$$

$$\widehat{q}(\mathrm{x}) := q_{\text{pers}}(\mathrm{x})\, \widehat{q}_{\text{res}}(\mathrm{x}). \tag{4b}$$

Note that both perspective and resolution quality measures yield a nonnegative value at a visible point $\mathrm{x} \in \mathbb{R}^d$. Hence, the associated unlimited- and limited-range conic fields of view of a PTZ camera are, respectively, given by

$$C := \left\{\mathrm{x} \in \mathbb{R}^n \,\middle|\, q_{\text{pers}}(\mathrm{x}) \geq 0\right\} = \left\{\mathrm{x} \in \mathbb{R}^n \,\middle|\, \frac{(\mathrm{x}-\mathrm{p})^{\mathrm{T}}\mathrm{v}}{\|\mathrm{x}-\mathrm{p}\|} \geq \cos\alpha\right\}, \tag{5a}$$

$$\widehat{C} := \left\{\mathrm{x} \in \mathbb{R}^n \,\middle|\, q_{\text{pers}}(\mathrm{x}) \geq 0, q_{\text{res}}(\mathrm{x}) \geq 0\right\}, \tag{5b}$$

$$= \left\{\mathrm{x} \in \mathbb{R}^n \,\middle|\, \frac{(\mathrm{x}-\mathrm{p})^{\mathrm{T}}\mathrm{v}}{\|\mathrm{x}-\mathrm{p}\|} \geq \cos\alpha, \|\mathrm{x}-\mathrm{p}\| \leq \frac{\lambda+1}{\lambda}R\cos\alpha\right\}.$$

Note that quality measure q is strictly positive in the interior $\mathring{C}$ of its field of view cone C, and is zero on the boundary ∂C of its field of view, i.e., $q(\mathrm{x}) > 0$ for any $\mathrm{x} \in \mathring{C}$ and $q(\mathrm{x}) = 0$ for all $\mathrm{x} \in \partial C$, and this likewise holds for $\widehat{q}$ and $\widehat{C}$.

B. Optimal Sensor Allocation in PTZ Camera Networks

We now introduce a use of the proposed sensing quality measures for coverage optimization of PTZ camera networks.

Let W be a convex bounded environment in $\mathbb{R}^n$ and $\phi : W \to \mathbb{R}_{\geq 0}$ be a priori known time-invariant event distribution function that describes the probability of some event taking place in W. Suppose $m \in \mathbb{N}$ identical[4] PTZ cameras (i.e., they share the same fixed quality measure parameters R, σ, λ and κ) placed at locations $\mathbf{p} := (\mathrm{p}_1, \mathrm{p}_2, \ldots, \mathrm{p}_m) \in (\mathbb{R}^n)^m$ with optical-axis directions $\mathbf{v} := (\mathrm{v}_1, \mathrm{v}_2, \ldots, \mathrm{v}_m) \in \left(\mathbb{S}^{n-1}\right)^m$ and (halves the) angles of view $\boldsymbol{\alpha} := (\alpha_1, \alpha_2, \ldots, \alpha_m) \in \left(0, \frac{\pi}{2}\right)^m$. Further, let q_i and C_i denote ith camera's (unlimited- or limited-range) spatial sensing quality measure and field of view, respectively, defined as in (4) and (5).

1) Coverage Objective: Following a standard resource allocation approach widely used in facility localization [18], [19] and quantization [20], [21], we assume that each camera is assigned to observe events, distributed according to ϕ over W, based on a partition[5] $\mathcal{P} := \{P_1, P_2, \ldots, P_m\}$ of W. Accordingly, for a given PTZ camera network configuration, to determine the coverage quality of sensor allocation partition $\mathcal{P}$, we use the total spatial sensing quality of cameras, denoted by $H(\mathcal{P})$, as our utility function, defined as

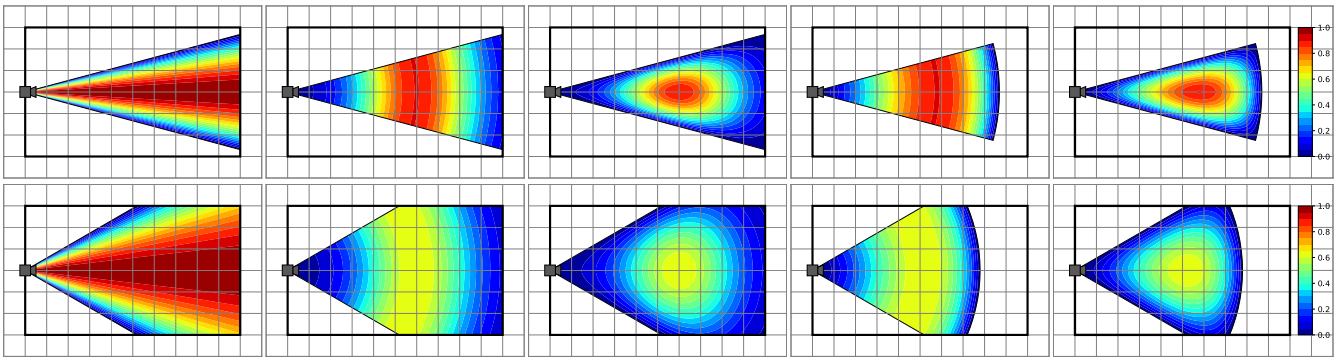
$$H(\mathcal{P}) := \sum_{i=1}^{m} \int_{P_i \cap C_i} q_i(\mathrm{x})\, \phi(\mathrm{x})\, \mathrm{dx}, \tag{6}$$

which is evidently tightly bounded above as

$$H(\mathcal{P}) \leq \int_{W \cap \bigcup_{i=1}^{m} C_i} \max_i q_i(\mathrm{x})\, \phi(\mathrm{x})\, \mathrm{dx}, \tag{7}$$

[4]This is only assumed for ease of presentation, our results can be directly applicable to both homogeneous and heterogeneous PTZ camera networks.

[5]By a slight abuse of terminology, we intend the term "partition" of a set to denote a collection of its subsets whose union returns itself and whose elements have mutually disjoint interiors.



(a) (b) (c) (d) (e)

Fig. 3. Spatial sensing quality measures. (top)-(bottom) 30-60 degree angle of view. (a) Perspective quality in (1), (b) Unlimited-range resolution quality in (3a), (c) Unlimited-range spatial sensing quality in (4a), (d) Limited-range resolution quality in (3b), (e) Limited-range spatial sensing quality in (4b). The figures are generated using $R = 6$, $\kappa = 3$, $\sigma = 2$, $\lambda = 2$. Please also refer to the accompanying video for an animated demonstration.

where $W \cap \bigcup_{i=1}^{m} C_i$ is the visible part of W by the cameras. Here, it is important to keep in mind that the quality measure q_i (and so the coverage objective H) is also a function of camera position p_i, optimal direction v_i and angle of view α_i, see (1)-(4), and we prefer to suppress them throughout the paper for ease of presentation.

2) Sensor Allocation via Conic Voronoi Diagrams: Thus, it directly follows from (7) that, for a given camera network configuration, the optimal camera allocation strategy is to assign each event location to the camera with maximal sensing performance. This defines a partition $\mathcal{V} := (V_1, V_2, \ldots, V_m)$ of the visible subset $W \cap \bigcup_{i=1}^{m} C_i$ of W, based on the camera's field of view and sensing quality measures, as

$$V_i := \left\{ \mathrm{x} \in W \cap C_i \middle| q_i(\mathrm{x}) \geq q_j(\mathrm{x}) \quad \forall j \neq i \right\}, \tag{8}$$

which we refer to as the *conic Voronoi diagram* of W, because it has the standard form of a generalized Voronoi diagram [18] whose generating objects are conic camera fields of view and the associated quality measures. Note that, since the quality measure q_i in (4) is a continuous and smooth function of event location x almost everywhere (the only exception is for the events happening at the camera positions, where the sensing quality is zero), the conic Voronoi cells have continuous and piecewise smooth boundaries.

Consequently, given optimal camera allocation, our utility function in (6) takes the following form

$$H_{\mathcal{V}} = \sum_{i=1}^{m} \int_{V_i} q_i(\mathrm{x})\, \phi(\mathrm{x})\, \mathrm{dx}. \tag{9}$$

In the following subsection, assuming optimal sensor allocation, we continue with important characteristic properties of locally optimal PTZ camera network configurations of $H_{\mathcal{V}}$.

C. Locally Optimal Coverage Configurations

Since the quality measures in (3) are zero at the boundary of their respective fields of view in (5), i.e., $q_i(\mathrm{x}) = 0$ for all $\mathrm{x} \in \partial C_i$, we have:

Lemma 1 *The partial derivatives of the utility function $H_{\mathcal{V}}$ in (9) with respect to the PTZ camera configuration variables are given by*

$$\frac{\partial H_{\mathcal{V}}}{\partial \mathrm{v}_i} = \int_{V_i} \frac{\partial q_i(\mathrm{x})}{\partial \mathrm{v}_i} \phi(\mathrm{x})\, d\mathrm{x}, \qquad \frac{\partial H_{\mathcal{V}}}{\partial \alpha_i} = \int_{V_i} \frac{\partial q_i(\mathrm{x})}{\partial \alpha_i} \phi(\mathrm{x})\, d\mathrm{x}. \tag{10}$$

Proof. It follows from the differentiation under the integral sign [22] (also see [21][Lemma 6.1]) that

$$\frac{\partial H_{\mathcal{V}}}{\partial \mathrm{v}_i} = \int_{V_i} \frac{\partial q_i(\mathrm{x})}{\partial \mathrm{v}_i} \phi(\mathrm{x}) \mathrm{dx} + \underbrace{\sum_{j=1}^{m} \int_{\partial V_j} q_j(\mathrm{x}) \phi(\mathrm{x})\, \dot{\mathrm{x}}^{\mathrm{T}} \mathrm{n}_j(\mathrm{x}) \mathrm{dx}}_{=0}, \tag{11}$$

where $\dot{\mathrm{x}}$ denotes the change in the boundary point $\mathrm{x} \in \partial V_j$ due to the change in v_i, and $\mathrm{n}_j(\mathrm{x})$ is the unit outward normal of ∂V_j at point x. Here, the second term is always zero, because either $\mathrm{x} \in \partial W$ where $\dot{\mathrm{x}} = 0$ or one of the following cases holds at each boundary point $\mathrm{x} \in \bigcup_{j=1}^{m} \partial V_j \setminus \partial W$:

i) if $\mathrm{x} \notin \partial V_i$, then $\dot{\mathrm{x}} = 0$.
ii) if $\mathrm{x} \in \partial V_i$ and $\mathrm{x} \notin \partial V_j$ for all $j \neq i$, then $q_i(\mathrm{x}) = 0$.
iii) if $\mathrm{x} \in \partial V_i \cap \partial V_j$ for some $j \neq i$ and $\mathrm{x} \notin \partial V_k$ for all $k \neq i$ and $k \neq j$, then $q_i(\mathrm{x}) = q_j(\mathrm{x})$ and $\mathrm{n}_i(\mathrm{x}) = -\mathrm{n}_j(\mathrm{x})$.
iv) Otherwise, $\mathrm{x} \in (\partial V_i \cap \partial V_j \cap \partial V_k) \setminus \partial W$ for some $i \neq j \neq k$ and $(\partial V_i \cap \partial V_j \cap \partial V_k) \setminus \partial W$ has measure zero.

Using a similar argument, one can also verify (10) for $\frac{\partial H_{\mathcal{V}}}{\partial \alpha_i}$. Thus, the result follows. ∎

1) Optimal PTZ Camera Configurations for Unlimited-Range Visual Sensing: In the case of unlimited-range visual sensing, defined as in (4a),

$$q_i(\mathrm{x}) = \frac{\cos^{\kappa} \alpha_i}{1 - \cos \alpha_i} \left(\frac{(\mathrm{x} - \mathrm{p}_i)^{\mathrm{T}} \mathrm{v}}{\|\mathrm{x} - \mathrm{p}_i\|} - \cos \alpha_i \right) e^{-\frac{(\|\mathrm{x} - \mathrm{p}_i\| - R)^2}{2\sigma^2}}. \tag{12}$$

the optimality of a PTZ camera configuration, as expressed by the gradients in (10), is strongly influenced by certain key terms, derived from the mass density function $\exp\left(-\frac{(\|\mathrm{x} - \mathrm{p}_i\| - R)^2}{2\sigma^2}\right) \phi(\mathrm{x})$, that we will now call out in isolation and refer to henceforth by the terms, respectively, *mass*, *centroidal perspective* and *centroidal aperture* of the camera's Voronoi cell, V_i, as follows:

$$\mu_{V_i} := \int_{V_i} e^{-\frac{(\|\mathrm{x} - \mathrm{p}_i\| - R)^2}{2\sigma^2}} \phi(\mathrm{x})\, \mathrm{dx}, \tag{13a}$$

$$\mathrm{v}_{V_i} := \frac{1}{\mu_{V_i}} \int_{V_i} \frac{\mathrm{x} - \mathrm{p}_i}{\|\mathrm{x} - \mathrm{p}_i\|} e^{-\frac{(\|\mathrm{x} - \mathrm{p}_i\| - R)^2}{2\sigma^2}} \phi(\mathrm{x})\, \mathrm{dx}, \tag{13b}$$

$$\delta_{V_i} := 1 - \mathrm{v}_{V_i}^{\mathrm{T}} \mathrm{v}_i. \tag{13c}$$

If $\mathring{V}_i = \varnothing$ (which corresponds to a redundant camera), then we set $\mu_{V_i} = 0$, $\mathrm{v}_{V_i} = (1 - \epsilon)\mathrm{v}_i$ and so $\delta_{V_i} = \epsilon$, where $\epsilon \in (0, 1)$ is a small fixed value that defines a small angle

of view below in (15), corresponding to zooming-in to some default view angle. Otherwise (i.e., $\mathring{V}_i \neq \varnothing$), we always have $\|\mathrm{v}_{V_i}\| \in (0,1)$ and $\delta_{V_i} \in (0,1)$ for $\alpha_i \in \left(0, \frac{\pi}{2}\right)$.

With these observations and the cell-specific parameters now in hand in (13), we can rewrite closed form expressions for the gradients in (10) as,

$$\frac{\partial H_{\mathcal{V}}}{\partial \mathrm{v}_i} = \mu_{V_i} \frac{\cos^{\kappa}\alpha_i}{1-\cos\alpha_i} \mathrm{v}_{V_i}, \tag{14a}$$

$$\frac{\partial H_{\mathcal{V}}}{\partial \alpha_i} = -\mu_{V_i} \frac{\tan\alpha_i \cos^{\kappa}\alpha_i}{(1-\cos\alpha_i)^2} \Big(\kappa\,(1-\cos\alpha_i)^2 - \big((\kappa-1)(1-\cos\alpha_i)+1\big)\,\delta_{V_i} \Big). \tag{14b}$$

We also find it convenient to define the camera's *centroidal angle of view* to be

$$\alpha_{V_i} := \arccos\left(1 - \frac{(\kappa-1)\delta_{V_i} + \sqrt{(\kappa-1)^2 {\delta_{V_i}}^2 + 4\kappa\delta_{V_i}}}{2\kappa}\right), \tag{15}$$

which is the only relevant zero of (14b) for $\alpha_i \in \left(0, \frac{\pi}{2}\right)$. Note that $\alpha_{V_i} \in \left(0, \frac{\pi}{2}\right)$, because the argument of $\arccos$ in (15) is in $(0,1)$ for $\delta_{V_i} \in (0,1)$.

Thus, it follows from (14a) and (14b) that a locally optimal PTZ camera network configuration satisfies:

Theorem 1 *For unlimited-range visual sensing (12), a PTZ camera network configuration is locally optimal in the sense of $H_{\mathcal{V}}$ in (9) if and only if all cameras look towards their centroidal perspectives with centroidal angles of view, i.e.,*

$$\mathrm{v}_i = \mathrm{v}_{V_i}/\|\mathrm{v}_{V_i}\|, \quad \textit{and} \quad \alpha_i = \alpha_{V_i}, \quad \forall i = 1,2,\ldots,m \tag{16}$$

Proof. The result trivially holds for redundant cameras with $\mathring{V}_i = \varnothing$. Otherwise (i.e., $\mathring{V}_i \neq \varnothing$), since $\mathrm{v}_i \in \mathbb{S}^{n-1}$, a valid infinitesimal change in the camera's optical direction v_i is specified by a tangent vector $\mathrm{u} \in T_{\mathrm{v}_i}\mathbb{S}^{n-1}$ from the tangent space $T_{\mathrm{v}_i}\mathbb{S}^{n-1}$ of $\mathbb{S}^{n-1}$ at point v_i, and the associated directional change of $H_{\mathcal{V}}$ along direction u is given by $\frac{\partial H_{\mathcal{V}}}{\partial \mathrm{v}_i}\mathrm{u}$. Hence, since ${\mathrm{v}_i}^{\mathrm{T}}\mathrm{u} = 0$ for all $\mathrm{u} \in T_{\mathrm{v}_i}\mathbb{S}^{n-1}$, it follows from (14a) that at a locally optimal configuration one should have $\mathrm{v}_i = \frac{\mathrm{v}_{V_i}}{\|\mathrm{v}_{V_i}\|}$ for $\alpha_i \in \left(0, \frac{\pi}{2}\right)$. Similarly, for a given camera perspective v_i, by construction (15), the centroidal angle of view α_{V_i} is the only relevant zero of (14b) for $\alpha_i \in \left(0, \frac{\pi}{2}\right)$ where $H_{\mathcal{V}}$ is locally maximal. Thus, the result follows. ■

2) Optimal PTZ Camera Configurations for Limited-Range Visual Sensing: In the case of limited-range visual sensing, as defined in (4b), i.e.,

$$q_i(\mathrm{x}) = \frac{\lambda+1}{1-\cos\alpha_i}\left(\frac{(\mathrm{x}-\mathrm{p}_i)^{\mathrm{T}}\mathrm{v}_i}{\|\mathrm{x}-\mathrm{p}_i\|} - \cos\alpha_i\right)\left(\cos\alpha_i - \frac{\lambda\|\mathrm{x}-\mathrm{p}_i\|}{(\lambda+1)R}\right)\frac{\|\mathrm{x}-\mathrm{p}_i\|^{\lambda}}{R^{\lambda}} \tag{17}$$

we redefine the *mass*, *centroidal perspective* and *centroidal aperture* of a camera's Voronoi cell, V_i, respectively, as

$$\widehat{\mu}_{V_i} := \int_{V_i} \frac{\|\mathrm{x}-\mathrm{p}_i\|^{\lambda}}{R^{\lambda}} \phi(\mathrm{x})\,\mathrm{d}\mathrm{x}, \tag{18a}$$

$$\widehat{\mathrm{v}}_{V_i} := \frac{1}{\widehat{\mu}_{V_i}} \int_{V_i} \frac{\mathrm{x}-\mathrm{p}_i}{\|\mathrm{x}-\mathrm{p}_i\|}\left(\cos\alpha_i - \frac{\lambda\|\mathrm{x}-\mathrm{p}_i\|}{(\lambda+1)R}\right)\frac{\|\mathrm{x}-\mathrm{p}_i\|^{\lambda}}{R^{\lambda}}\phi(\mathrm{x})\,\mathrm{d}\mathrm{x}, \tag{18b}$$

$$\widehat{\delta}_{V_i} := \frac{1}{\widehat{\mu}_{V_i}} \int_{V_i} \left(1 - \frac{(\mathrm{x}-\mathrm{p}_i)^{\mathrm{T}}\mathrm{v}_i}{\|\mathrm{x}-\mathrm{p}_i\|}\right)\left(1 - \frac{\lambda\|\mathrm{x}-\mathrm{p}_i\|}{(\lambda+1)R}\right)\frac{\|\mathrm{x}-\mathrm{p}_i\|^{\lambda}}{R^{\lambda}}\phi(\mathrm{x})\mathrm{d}\mathrm{x}. \tag{18c}$$

If $\mathring{V}_i = \varnothing$, then we set $\widehat{\mu}_{V_i} = 0$, $\widehat{\mathrm{v}}_{V_i} = (1-\epsilon)\mathrm{v}_i$ and $\widehat{\delta}_{V_i} = \epsilon$, where $\epsilon \in (0,1)$ is a small fixed scalar. Otherwise (i.e., $\mathring{V}_i \neq \varnothing$), we have $0 < \|\widehat{\mathrm{v}}_{V_i}\| < 1$ and $0 < \widehat{\delta}_{V_i} < 1$, because of the effective sensing range limit, i.e., $\|\mathrm{x}-\mathrm{p}_i\| \leq \frac{\lambda+1}{\lambda}R$ for any $\mathrm{x} \in V_i \subseteq \widehat{C}_i$, where $\widehat{C}_i$ is the camera's field of view defined as in (5b). Accordingly, the partial derivatives of $H_{\mathcal{V}}$ with respect to camera direction and angle of view, in (10), can be rewritten as

$$\frac{\partial H_{\mathcal{V}}}{\partial \mathrm{v}_i} = \widehat{\mu}_{V_i} \frac{\lambda+1}{1-\cos\alpha_i} \widehat{\mathrm{v}}_{V_i}, \tag{19a}$$

$$\frac{\partial H_{\mathcal{V}}}{\partial \alpha_i} = -\widehat{\mu}_{V_i} \frac{\sin\alpha_i}{(1-\cos\alpha_i)^2}\left((1-\cos\alpha_i)^2 - \widehat{\delta}_{V_i}\right), \tag{19b}$$

and so the associated *centroidal angle of view* is redefined as

$$\widehat{\alpha}_{V_i} := \arccos\left(1 - \sqrt{\widehat{\delta}_{V_i}}\right). \tag{20}$$

Therefore, similar to Theorem 1, one can conclude that:

Theorem 2 *For limited-range visual sensing in (17), at a locally optimal PTZ camera coverage configuration of $H_{\mathcal{V}}$ in (9), all cameras are directed at the centroidal perspectives of their respective Voronoi cells with the associated centroidal-angles-of-views, i.e.,*

$$\mathrm{v}_i = \widehat{\mathrm{v}}_{V_i}/\|\widehat{\mathrm{v}}_{V_i}\|, \quad \textit{and} \quad \alpha_i = \widehat{\alpha}_{V_i}, \quad \forall i = 1,2,\ldots,m \tag{21}$$

Proof. The proof follows the same pattern as of Theorem 1, and so it is omitted for the sake of brevity. ■

III. Coverage Control of PTZ Cameras

In this section, we consider a set of PTZ cameras placed at fixed locations $\mathbf{p} = (\mathrm{p}_1, \mathrm{p}_2, \ldots, \mathrm{p}_m) \in (\mathbb{R}^n)^m$ with variable optical directions $\mathbf{v} = (\mathrm{v}_1, \mathrm{v}_2, \ldots, \mathrm{v}_m) \in \left(\mathbb{S}^{n-1}\right)^m$ and variable angles of view $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_m) \in \left(0, \frac{\pi}{2}\right)^m$, and propose a simple reactive coverage control algorithm for maximizing the total coverage quality of the cameras. Here, one can freely choose unlimited- or limited-range visual sensing for modeling sensing performance of a PTZ camera.

A. Continuous-Time Camera Dynamics

Assuming continuous-time first-order PTZ camera dynamics, we propose greedy coverage control policies, respectively, called the *"move-to-centroidal-perspective"* and *"move-to-centroidal-angle-of-view"* laws, that continuously (re)configure camera orientations and angles of view to asymptotically reach a locally optimal coverage configurations of $H_{\mathcal{V}}$ in (9), as follows

$$\dot{\mathrm{v}}_i = K_{\mathrm{v}}\left(\mathbf{I} - \mathrm{v}_i{\mathrm{v}_i}^{\mathrm{T}}\right)\mathrm{v}_{V_i}, \tag{22a}$$

$$\dot{\alpha}_i = -K_{\alpha}\left(\alpha_i - \alpha_{V_i}\right), \tag{22b}$$

where $K_{\mathrm{v}} > 0$ and $K_{\alpha} > 0$ are fixed control gains, and $\mathbf{I}$ denotes the identity matrix of appropriate size. Here, v_{V_i} and α_{V_i}, respectively, denote the centroidal perspective and centroidal angle of view of the Voronoi cell V_i of ith camera, which are explicitly defined for unlimited- and limited-range sensing models in Section II-C. Also note that $\mathbf{I} - \mathrm{v}_i{\mathrm{v}_i}^{\mathrm{T}}$ defines the projection matrix onto the tangent space $T_{\mathrm{v}_i}\mathbb{S}^{n-1}$, which ensures the positive invariance of v_i in $\mathbb{S}^{n-1}$.

In brief, the coverage control strategy in (22) aim to asymptotically have each camera looking at its centroidal

perspective v_{V_i} with centroidal angle of view α_i where its coverage quality is locally maximized, see Theorem 1 and Theorem 2. Thus, some important qualitative properties of the proposed coverage laws can be summarized as:

Theorem 3 *The continuously differentiable "move-to-centroidal-perspective" and "move-to-centroidal-angle-of-view" laws in (22) leave camera's angles of view, $2\alpha_i$, positively invariant in $(0, \pi)$, and asymptotically bring a PTZ camera network to a locally optimal coverage configuration of $H_{\mathcal{V}}$ in (9) while strictly increasing the total coverage quality $H_{\mathcal{V}}$ along the way.*

Proof. The continuous differentability of the proposed coverage policies directly follows from the integral forms of Voronoi parameters in (13) and (18). The positive invariance of $\left(0, \frac{\pi}{2}\right)$ for α_i is guaranteed by construction, in (15) and (20), because we always have $\alpha_{V_i} \in \left(0, \frac{\pi}{2}\right)$. Finally, for the stability analysis, one can naturally consider the utility function $H_{\mathcal{V}}$ in (9) as a Lyapunov candidate function and verify from (14) and (19) that

$$\dot{H}_{\mathcal{V}} = \sum_{i=1}^{m} \underbrace{\frac{\partial H_{\mathcal{V}}}{\partial \mathrm{v}_i}\dot{\mathrm{v}}_i}_{\geq 0} + \underbrace{\frac{\partial H_{\mathcal{V}}}{\partial \alpha_i}\dot{\alpha}_i}_{\geq 0} \geq 0, \tag{23}$$

where the equality only holds at a locally optimal coverage configuration specified by Theorem 1 and Theorem 2. Thus, since $H_{\mathcal{V}}$ is bounded above for a bounded environment W, we have from LaSalle's Invariance Principle [23] that all cameras asymptotically converge an optimal coverage configuration where $H_{\mathcal{V}}$ is locally maximized. ■

B. Discrete-time Camera Dynamics

Denote by $\mathcal{V}(\mathbf{v}, \boldsymbol{\alpha}) = (V_1(\mathbf{v}, \boldsymbol{\alpha}), V_2(\mathbf{v}, \boldsymbol{\alpha}), \ldots, V_m(\mathbf{v}, \boldsymbol{\alpha}))$ the conic Voronoi diagram of W, defined in (8), associated with PTZ camera network configuration $(\mathbf{v}, \boldsymbol{\alpha})$, where fixed camera positions $\mathbf{p}$ are suppress for ease of notation. Accordingly, we propose an extension of the "move-to-centroid-perspective" and "move-to-centroidal-angle-of-view" laws in (22) for discrete-time first-order PTZ camera dynamics as follows:

$$\mathrm{v}_i[k+1] = \frac{\mathrm{v}_{V_i(\mathbf{v}[k], \boldsymbol{\alpha}[k])}}{\left\|\mathrm{v}_{V_i(\mathbf{v}[k], \boldsymbol{\alpha}[k])}\right\|}, \tag{24a}$$

$$\alpha_i[k+1] = \alpha_{V_i(\mathbf{v}[k+1], \boldsymbol{\alpha}[k])}, \tag{24b}$$

which sequentially optimize camera orientations and angles of view in order to asymptotically reach a locally optimal coverage configuration of $H_{\mathcal{V}}$, specified by Theorem 1 and Theorem 2 for unlimited- and limited-range sensing models, respectively. Therefore, we have:

Theorem 4 *The total coverage quality $H_{\mathcal{V}}$ in (9) of a PTZ camera network increases at each iteration of the "move-to-centroidal-perspective" and "move-to-centroidal-angle-of-view" laws in (24) until asymptotically reaching a locally optimal coverage configuration. Further, each iteration yields a valid camera angle of view, $2\alpha_i$, in $(0, \pi)$.*

Proof. By definition, in (15) and (20), the centroidal angle of view $\alpha_{V_i(\mathbf{v}, \boldsymbol{\alpha})}$ is always in $\left(0, \frac{\pi}{2}\right)$ for any $\alpha_i \in \left(0, \frac{\pi}{2}\right)$. Also recall from (13) and (18) that $0 < \left\|\mathrm{v}_{V_i(\mathbf{v}, \boldsymbol{\alpha})}\right\| < 1$ for any $\alpha_i \in \left(0, \frac{\pi}{2}\right)$. Thus, the iterations of our coverage optimization strategy in (24) yield valid PTZ camera configurations.

In the rest of the proof, we will continue with the generic form of the coverage object H in (6) defined for a partition of $\mathcal{P} = (P_1, P_2, \ldots, P_m)$ of W, and it is convenient to explicitly write its parametric dependency on variable PTZ camera configurations as

$$H = H(\mathbf{v}, \boldsymbol{\alpha}, \mathcal{P}) = \sum_{i=1}^{m} \int_{P_i \cap C_i(\mathrm{v}_i, \alpha_i)} q_i(\mathrm{x}; \mathrm{v}_i, \alpha_i)\, \phi(\mathrm{x})\, \mathrm{d}\mathrm{x}, \tag{25}$$

where q_i and C_i denote the ith camera's sensing quality measure and field of view cone, respectively, defined in (4) and (5). Further, it follows from 9 that

$$H_{\mathcal{V}} = H(\mathbf{v}, \boldsymbol{\alpha}, \mathcal{V}(\mathbf{v}, \boldsymbol{\alpha})). \tag{26}$$

We now show that $H_{\mathcal{V}}$ is strictly increasing at each iteration of (24) away from optimal coverage configurations.

First, since $\mathrm{v}_{V_i(\mathbf{v}[k], \boldsymbol{\alpha}[k])}$ is the only optimal critical point of $H_{\mathcal{V}}$ for given $\alpha_i[k]$ and $V_i(\mathbf{v}[k], \boldsymbol{\alpha}[k])$, as discussed in the proofs of Theorem 1 and Theorem 2, we have

$$H(\mathbf{v}[k], \boldsymbol{\alpha}[k], \mathcal{V}(\mathbf{v}[k], \boldsymbol{\alpha}[k])) \leq H(\mathbf{v}[k+1], \boldsymbol{\alpha}[k], \mathcal{V}(\mathbf{v}[k], \boldsymbol{\alpha}[k])), \tag{27}$$

where the equality only holds if $\alpha_i[k] = \mathrm{v}_{V_i(\mathbf{v}[k], \boldsymbol{\alpha}[k])}$ for all i. Further, since $q_i(\mathrm{x}; \mathrm{v}_i[k+1], \alpha_i[k]) < 0$ for all points $\mathrm{x} \in C_i(\mathrm{v}_i[k], \alpha_i[k]) \setminus C_i(\mathrm{v}_i[k+1], \alpha_i[k])$ and an event location is assigned to the camera with the maximal sensing quality via conic Voronoi diagrams in (8), we also have,

$$H(\mathbf{v}[k+1], \boldsymbol{\alpha}[k], \mathcal{V}(\mathbf{v}[k], \boldsymbol{\alpha}[k])) \leq H(\mathbf{v}[k+1], \boldsymbol{\alpha}[k], \mathcal{V}(\mathbf{v}[k+1], \boldsymbol{\alpha}[k])), \tag{28}$$

where the quality holds only if all the cameras are already directed at their respective centroidal perspective.

Similarly, due to the optimality of centroidal angle of view α_{V_i} (see Theorem 1 and Theorem 2) and the role of conic Voronoi diagrams in optimal camera allocation (see Section II-B), one can verify that:

$$\begin{aligned} H(\mathbf{v}[k+1], \boldsymbol{\alpha}[k], \mathcal{V}(\mathbf{v}[k+1], \boldsymbol{\alpha}[k])) &\leq H(\mathbf{v}[k+1], \boldsymbol{\alpha}[k+1], \mathcal{V}(\mathbf{v}[k+1], \boldsymbol{\alpha}[k])) \\ &\leq H(\mathbf{v}[k+1], \boldsymbol{\alpha}[k+1], \mathcal{V}(\mathbf{v}[k+1], \boldsymbol{\alpha}[k+1])), \end{aligned} \tag{29}$$

Thus, overall, we obtain

$$H(\mathbf{v}[k], \boldsymbol{\alpha}[k], \mathcal{V}(\mathbf{v}[k], \boldsymbol{\alpha}[k])) \leq H(\mathbf{v}[k+1], \boldsymbol{\alpha}[k+1], \mathcal{V}(\mathbf{v}[k+1], \boldsymbol{\alpha}[k+1])), \tag{30}$$

which completes the proof. ■

IV. Numerical Simulations

In this section, we provide numerical evidence demonstrating the descriptive strength of our spatial sensing quality measures for coverage optimization of PTZ camera networks and the effectiveness of the proposed coverage control laws.

In order to have a more intuitive understanding of the proposed spatial sensing quality measures in (4), we first consider coverage optimization of a single PTZ camera over a 10×10 planar environment, as illustrated at the top of Fig. 4. If one assumes a uniform event distribution over the environment, i.e., $\phi(\mathrm{x}) = 1$ for all $\mathrm{x} \in [0, 10]^2$, then our coverage control algorithm reconfigures the camera, as shown in Fig. 4(top)(b,d), to maximize area coverage

while maintaining a certain level of spatial sensing quality. Recall from (4) that our spatial sensing models define a trade-off between the camera's angle of view and sensing resolution quality, as illustrated in Fig. 3, and so our coverage control algorithm asymptotically finds an optimal coverage configuration which locally balances this trade-off. In the case of a concentrated event distribution, say,

$$\phi(\mathrm{x}) = e^{-\left\|\mathrm{x}-\begin{bmatrix}7\\3\end{bmatrix}\right\|^2}, \tag{31}$$

the camera, as expected, focuses on the region where the event occurs the most for both unlimited and limited range sensing models, as depicted in Fig. 4(top)(c,e).

To illustrate collective coverage behavior in multiple PTZ camera networks, we present at the bottom of Fig. 4 coverage optimization of four PTZ cameras placed at the corners of a 10×10 square environment. As one might naturally expect, multiple PTZ cameras can achieve better coverage performance, possibly with lower angles of view; for example, compare uniformly distributed event coverage performances in Fig. 4(top)(b,d) and Fig. 4(bottom)(b,d). As seen in Fig. 4(bottom)(b,d), in the case of a uniform event distribution, each camera observes the region around its adjacent corner (in circular order) at an optimal coverage configuration. It is important to observe that although the cameras in Fig. 4(bottom)(d) can observe every point in the environment at higher angles of view, because of the trade-off between view angle and view resolution, they do not cover the small region at around the center of the environment. In the case of a more complex event distribution function, for example,

$$\phi(\mathrm{x}) = 1 + 10\left(e^{-\frac{1}{9}\left\|\mathrm{x}-\begin{bmatrix}8\\8\end{bmatrix}\right\|^2} + e^{-\frac{1}{2}\left\|\mathrm{x}-\begin{bmatrix}8\\2\end{bmatrix}\right\|^2} + e^{-\frac{1}{2}\left\|\mathrm{x}-\begin{bmatrix}8\\4\end{bmatrix}\right\|^2} + e^{-\left\|\mathrm{x}-\begin{bmatrix}3\\7\end{bmatrix}\right\|^2}\right), \tag{32}$$

the cameras collectively focus in the region of the environment where events occur more often, see Fig. 4(bottom)(c,e).

V. Experimental Validation

For experimental validation of the proposed coverage control algorithm, we set up a $5m$ (width) $\times$ $13m$ (length) $\times$ $4m$ (height) rectangular box-shaped environment with intrinsically and extrinsically precalibrated six Bosch AUTODOME IP 4000 HD PTZ cameras[8], mounted on the ceiling equidistantly along the long edge, whose (vertical) angles of view ranges from 2.5 degrees to 28 degrees and pan and tilt angle ranges are 360 and 90 degrees, respectively. For the sake of clear visualization, we only consider coverage control of three cameras indicated in Fig. 5 in a $5m$ (width) $\times 6.5m$ (length) $\times 4m$ (height) environment. To indicate event locations, we use AprilTags [24], and cameras initially perform a scanning motion to locate them. Accordingly, we use a Gaussian mixture model to represent event distribution at the detected tag locations, denoted by $\mathrm{a}_i \in \mathbb{R}^3$ for $i = \{1, 2\}$, as follows:

$$\phi(\mathrm{x}) = 1 + \sum_{i=1}^{2} 10 e^{-\frac{\|\mathrm{x}-\mathrm{a}_i\|^2}{2(0.8)^2}}. \tag{33}$$

In Fig. 5, we present the initial and final coverage configurations of the PTZ cameras, and camera images at the locally

[6]For all simulations, we set $R = 7$, $\kappa = 3$, $\sigma = 2$ and $\lambda = 2$, and all simulations are obtained through numerical integration of the associated coverage control law, using a fixed step size, where the computation of the centroidal perspectives and centroidal angles of view are approximated by discretizing the entire space by a dense 200×200 grid. Since the proposed coverage laws are observed to yield similar optimal converage configurations for both continuous- and discrete-time PTZ camera dynamics, we present results only for continuous-time camera dynamics.

[7] In our experiments, we set $R = 4.0$, $\sigma = 1.2$, $\kappa = 15$, and perform coverage control using discrete-time camera dynamics, where centroidal perspectives and centroidal angles of view are approximately computed by discretizing the environment using a dense grid.

[8]https://us.boschsecurity.com/en/

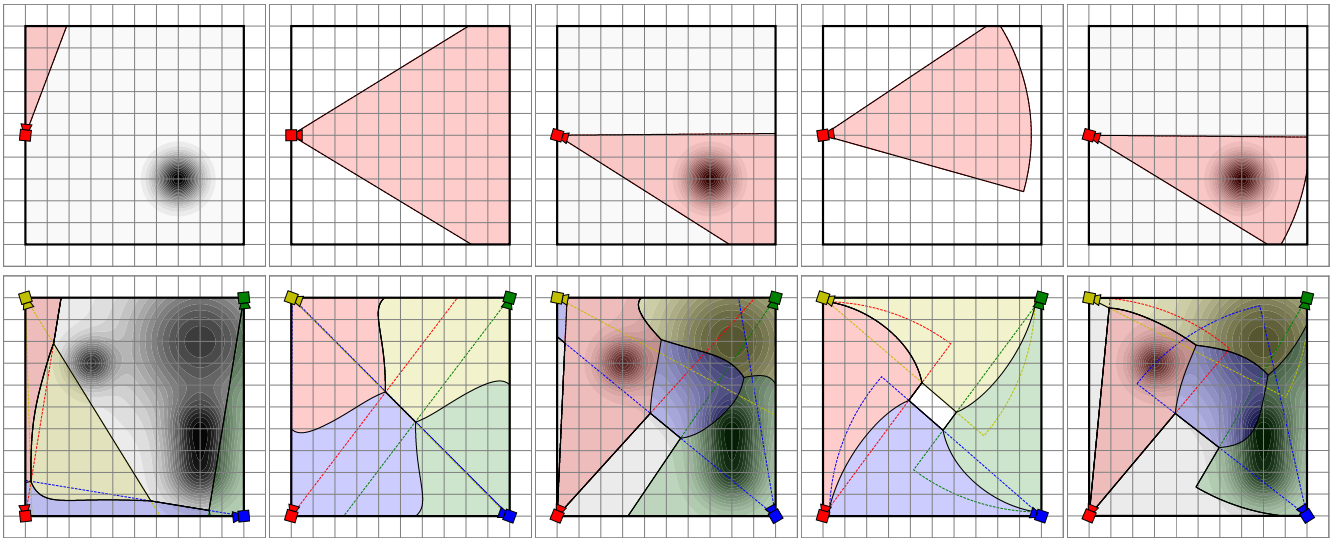


Fig. 4. Coverage optimization of (top) a single PTZ camera and (bottom) a multiple PTZ camera network. (a) Initial camera configuration and event distribution whose level sets are shaded darker with increasing event concentration, (b) Locally optimal area coverage configuration for unlimited range sensing in (4a), (c) Locally optimal event coverage configuration for unlimited-range sensing in (4a) , (d) Locally optimal area coverage configuration for limited range sensing in (4b), (e) Locally optimal event coverage configuration for limited range sensing in (4b). Please see the accompanying video submission for an animated demonstration.[6]

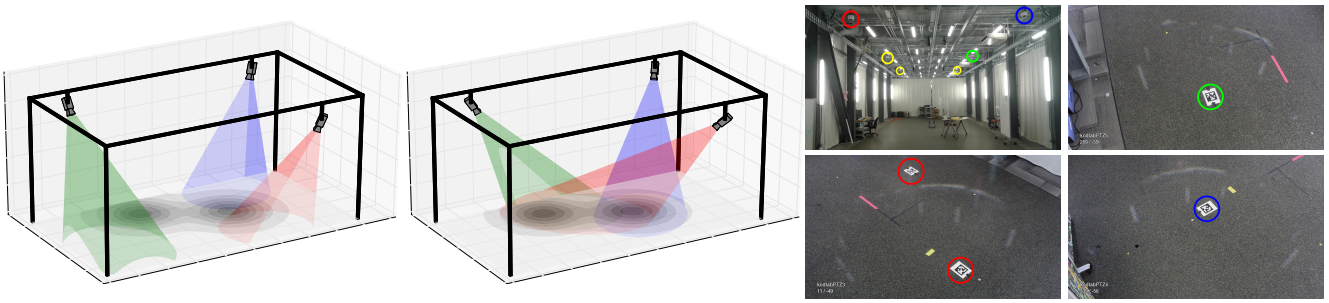


Fig. 5. Experimental demonstration of coverage optimization of a PTZ camera network. (left) Initial configuration and event distribution (middle) Locally optimal event coverage configuration for unlimited-range sensing in (4a) (right) Experimental setup: for ease of visualization, only three (indicated by red, blue and green) of six Bosch AUTODOME IP 4000 PTZ cameras are actively controlled, and the camera images are collected at the optimal coverage configuration, and the detected AprilTags are colored according to the color of the source camera. Please see the accompanying video for the full motion.[7]

optimal coverage configuration. In brief, while two cameras focus on the tag locations, the remaining third camera covers the region between them.

VI. Conclusions

In this paper, we consider the coverage optimization problem of PTZ camera networks and propose a new provably correct simple greedy coverage control algorithm that actively (re)configures cameras' orientations and angles of view (i.e., zoom levels) in order to optimally cover a given event distribution in a convex environment. We introduce two new visual quality measures for accurate modeling of spatial sensing performance of cameras. We show that the optimal camera allocation problem can be solved using conic Voronoi diagrams that are constructed based on the newly introduced quality measures, and we provide explicit conditions on locally optimal coverage configurations in terms of centroidal perspective and centroidal angle of the Voronoi cell of each camera. Accordingly, for both continuous- and discrete-time PTZ camera dynamics, we propose the "move-to-centroidal-perspective" and "move-to-centroidal-angle-of-view" laws that asymptotically reach a locally optimal coverage configuration where the collective coverage quality of cameras is locally maximized. The strengths of the proposed sensing quality measures and the effectiveness of the proposed coverage control laws are illustrated with numerical and experimental results.

Work now in progress targets online vision-based event/activity distribution estimation for coverage optimization of PTZ camera networks and balancing coverage and exploration in camera networks [25]. We are also exploring the extensions of our results for nonconvex environments and environments with visual occlusions. Our ultimate goal is to achieve optimal coverage control in mobile (e.g., flying) PTZ camera networks for active visual monitoring.

References

[1] B. Rinner and W. Wolf, "An introduction to distributed smart cameras," *Proceedings of the IEEE*, vol. 96, no. 10, pp. 1565–1575, 2008.

[2] C. Piciarelli, L. Esterle, A. Khan, B. Rinner, and G. L. Foresti, "Dynamic reconfiguration in camera networks: A short survey," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 26, no. 5, pp. 965–977, 2016.

[3] J. O'rourke, *Art gallery theorems and algorithms*. Oxford University Press Oxford, 1987, vol. 57.

[4] N. Megiddo and K. J. Supowit, "On the complexity of some common geometric location problems," *SIAM Journal on Computing*, vol. 13, no. 1, pp. 182–196, 1984.

[5] J. Cortés, S. Martınez, T. Karatas, and F. Bullo, "Coverage control for mobile sensing networks," *IEEE Transactions on Robotics and Automation*, vol. 20, no. 2, pp. 243–255, 2004.

[6] K. Laventall and J. Cortés, "Coverage control by multi-robot networks with limited-range anisotropic sensory," *International Journal of Control*, vol. 82, no. 6, pp. 1113–1121, 2009.

[7] A. Gusrialdi, S. Hirche, D. Asikin, T. Hatanaka, and M. Fujita, "Voronoi-based coverage control with anisotropic sensors and experimental case study," *Intel. Service Robot.*, vol. 2, no. 4, p. 195, 2009.

[8] X. Zhang, X. Chen, X. Liang, and Y. Fang, "Distributed coverage optimization for deployment of directional sensor networks," in *IEEE Conference on Decision and Control*, 2015, pp. 246–251.

[9] B. Hexsel, N. Chakraborty, and K. Sycara, "Coverage control for mobile anisotropic sensor networks," in *IEEE International Conference on Robotics and Automation*, 2011, pp. 2878–2885.

[10] D. Panagou, D. M. Stipanovic, and P. G. Voulgaris, "Distributed dynamic coverage and avoidance control under anisotropic sensing," *IEEE Trans. Control Netw. Syst.*, vol. PP, no. 99, pp. 1–1, 2016.

[11] M. Schwager, B. J. Julian, and D. Rus, "Optimal coverage for multiple hovering robots with downward facing cameras," in *IEEE International Conference on Robotics and Automation*, 2009, pp. 3515–3522.

[12] M. Schwager, B. J. Julian, M. Angermann, and D. Rus, "Eyes in the sky: Decentralized control for the deployment of robotic camera networks," *Proc. IEEE*, vol. 99, no. 9, pp. 1541–1561, 2011.

[13] S. Papatheodorou, A. Tzes, and Y. Stergiopoulos, "Collaborative visual area coverage," *Rob. Auton. Syst.*, vol. 92, pp. 126–138, 2017.

[14] W. Li and C. G. Cassandras, "Distributed cooperative coverage control of sensor networks," in *IEEE Conf. Decis. Cntl.*, 2005, pp. 2542–2547.

[15] A. Gusrialdi, T. Hatanaka, and M. Fujita, "Coverage control for mobile networks with limited-range anisotropic sensors," in *IEEE Conference on Decision and Control*, 2008, pp. 4263–4268.

[16] C. Piciarelli, C. Micheloni, and G. L. Foresti, "Ptz camera network reconfiguration," in *ACM/IEEE International Conference on Distributed Smart Cameras*, 2009, pp. 1–7.

[17] ——, "Automatic reconfiguration of video sensor networks for optimal 3d coverage," in *ACM/IEEE International Conference on Distributed Smart Cameras*, 2011, pp. 1–6.

[18] A. Okabe, B. Boots, K. Sugihara, and S. N. Chiu, *Spatial Tessellations: Concepts and Applications of Voronoi Diagrams*, 2nd ed., ser. Wiley Series in Probability and Statistics. John Wiley & Sons, 2000.

[19] H. W. Hamacher and Z. Drezner, *Facility Location: Applications and Theory*. Springer Science & Business Media, 2002.

[20] S. Lloyd, "Least squares quantization in PCM," *IEEE Transactions on Information Theory*, vol. 28, no. 2, pp. 129 – 137, 1982.

[21] Q. Du, V. Faber, and M. Gunzburger, "Centroidal voronoi tessellations: Applications and algorithms," *SIAM Review*, vol. 41, no. 4, pp. 637–676, 1999.

[22] H. Flanders, "Differentiation under the integral sign," *The American Mathematical Monthly*, vol. 80, no. 6, pp. 615–627, 1973.

[23] H. K. Khalil, *Nonlinear Systems*, 3rd ed. Prentice Hall, 2001.

[24] E. Olson, "AprilTag: Arobust and flexible visual fiducial system," in *IEEE Inter. Conf. on Robotics and Automation*, 2011, pp. 3400–3407.

[25] F. Pasqualetti, F. Zanella, J. R. Peters, M. Spindler, R. Carli, and F. Bullo, "Camera network coordination for intruder detection," *IEEE Trans. Control Syst. Technol.*, vol. 22, no. 5, pp. 1669–1683, 2014.