COVERAGE CONTROL FOR PTZ CAMERA NETWORKS USING SCENE POTENTIAL MAP

Liangliang Cai[†], *Hanyuan Ma*[†], *Zhuocheng Liu*[†], *Zhaoxin Li*[‡] *and Zhong Zhou*[†✉]

[†]State Key Laboratory of Virtual Reality Technology and Systems, Beihang University, Beijing, China
[‡]Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China
zz@buaa.edu.cn

ABSTRACT

Pan-Tilt-Zoom (PTZ) camera networks are pervasive in many applications, such as video surveillance, sports analysis and epidemic prevention. However, it is difficult to control several PTZ cameras observing a large scene due to the high complexity of PTZ camera networks. This problem, called the coverage control for PTZ cameras, draws broad attention of researchers. Virtual potential field (VPF) can enhance the coverage performance and reduce the overlap region of cameras, but it does not fully take the actual scenario into account. In this work, we introduce scene potential map (SPM) to the VPF method, investigating both the scene and the task in the optimization. The proposed scene potential map characterizes the importance of target region and evaluates the perception quality of PTZ camera. Then we propose a novel virtual force analysis method to optimalize poses of the PTZ camera network. We also develop a region partition method based on the perception quality measure to divide the target region and achieve better zoom level, realizing an excellent coverage performance of PTZ camera networks. Finally, the evaluation experiments clearly demonstrated that our proposed coverage control scheme can achieve remarkably better performance than state-of-the-arts.

Index Terms— Coverage control, PTZ camera network, Virtual potential field, Quality measure

1. INTRODUCTION

With decreasing hardware and setup costs, cameras have been used more than ever in many applications, such as situational awareness, security and surveillance. However, single camera is unable to monitor large scene due to the finite field of view (FoV). Pan-Tilt-Zoom (PTZ) camera networks deliver a bigger sight with controllable configuration, which attract people interests. However, increasing camera network size and vast amount of collected data make the coverage control of PTZ camera network be extremely difficult for human operation. Thus the coverage control in PTZ camera networks received a large amount of attention over the last several years [1,2]. The aim of coverage control for PTZ camera networks, also called as dynamic reconfiguration in the PTZ camera network [3], is to design an efficient control strategy to manipulate multiple cameras from initial poses to the desired one. An ideal coverage control method should eliminate the observing overlapping regions and coverage holes, achieving the entire coverage for a large region of interest.

Some early works treated dynamic reconfiguration in PTZ camera networks as the well-known art gallery problem [4], which has been proven to be NP-hard even under some simple circumstances. To mitigate this complexity, many gradient-based [5–7] and virtual potential field (VPF) -based algorithms [8–10] have been proposed to control PTZ cameras toward a local optimal deployment. A widely used class of gradient-based methods are to first assign each point to the closest camera, which generates the voronoi partition [11,12], and then controls cameras to the optimal configuration by solving a locational optimization problem [13]. However, PTZ cameras do not follow the law that the perception quality is generally modelled inversely proportional to the distance between the sampling point and the camera. Their perception quality measure relies on the relative pose between PTZ cameras and sampling points as well as the intrinsic parameters of PTZ cameras (e.g. image resolution and lens distortion) [6, 13]. The VPF is also widely used to address the issue of coverage enhancement. Tao et al. [14] translated the coverage-enhancing problem into a uniform distribution problem of centroid points by introducing the concept of coverage area ”centroid”. Ma et al. [8] and Li et al. [9] proposed a 3D PTZ camera coverage-control model and a VPF based coverage-enhancing scheme respectively, which result is closer to that of real-world. Ma and Kang [10] proposed a multi-detecting point based on virtual force-directed particle swarm optimization (MDPVF-PSO), where the particles are updated not only according to optimal solutions of population and individual, but also virtual force from neighboring sensors and multi-detecting points. Although the above methods can eliminate the observing overlapping regions by the repulsive interaction between virtural particles, they cannot find where the optimal coverage solution is reached due to lack of perception quality measure.

In this work, we introduce a perception quality measure into VPF methods and propose a novel virtual force analysis method based perception quality distribution, achieving the

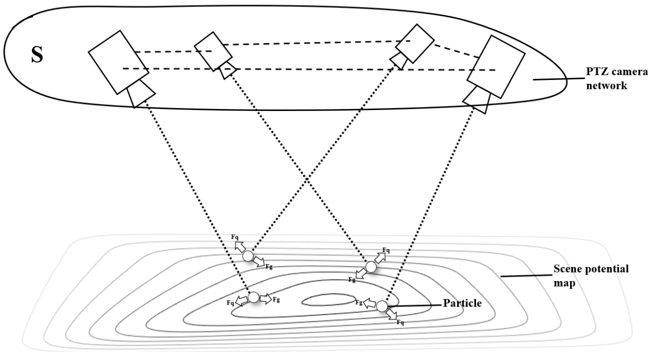


Fig. 1. The virtual force analysis method based on scene potential map.

optimal configuration of camera parameters (e.g. pan angle, tilt angle, zoom value) to cover the entire region of interests by the PTZ camera network. In our solution, the environment is represented by a discrete point map. First of all, we generate the scene potential map of each camera involving the target region information (e.g. region size, obstacles, building locations and heights) as well as the locations and orientations of PTZ cameras. Then we treat coverage area centroids of PTZ cameras as virtual charged particles. The virtual charged particles are subject to virtual gravity from the scene potential map and reputive force from other neighbor particles. These forces push each particle moving toward the position where perception quality is higher and the overlap regions between cameras are smaller. When all particles stop moving, each particle goes into a stable state. This step aims to maximize the coverage performance of target region and, meanwhile, improves the perception quality of the PTZ camera network. More specifically, we propose a novel perception quality measure for 3D PTZ camera model and a new region partition method based on perception quality measure to assign each point to the most acceptable camera. The above is shown in Fig. 1. Since our approach uses the centralized coverage control, the computation is only executed in the powerful sink node. Thus, each PTZ camera only needs to report its initial position and orientation, and receives the adjusted sensing orientation. As a result, our algorithm does not impose high communication and computation overheads as required by the other existing schemes. In summary, this work has three main contributions:

- A novel scene potential map generation method for each camera is presented to establish the relationship between the perception quality of PTZ cameras and the target region.
- A new virtual force analysis method based on scene potential map is developed to optimize camera coverage area centroids which are treat as virtual particles.
- A region partition method based on perception quality measure is designed to assign each point to the most acceptable camera. Then we can compute the best configuration of the PTZ camera network from these partitions.

2. PROBLEM FORMULATION AND EFFECTIVENESS EVALUATION OF PTZ CAMERAS

2.1. Problem formulation

To make the coverage enhancement problem for PTZ camera networks tractable, we first make the following assumptions.

(1) The target scene of the network is simplified as a 2D plane (i.e., $z = 0$).

(2) The location of each camera is fixed. And the center node knows each PTZ camera's location, internal parameters, and external parameters, which is necessary for computing the best pose of each camera with respect to the center node.

(3) The PTZ camera network can achieve complete coverage of the target area, ensuring the existence of feasible solutions.

Then, we can formulate the coverage control problem for PTZ camera networks as follows:

Let $W \subset \mathbb{R}^2$ with buildings (or obstacles) $B_i = (X_i, Y_i, Z_i, l_i, w_i, h_i), i = \{1, 2, 3, ..., M\}$, which is covered by a group of PTZ cameras $S = \{S_1, S_2, ..., S_n\}, S_j = \{X_j, Y_j, Z_j\}$. The position $\{X, Y, Z\}$ is the 3D point and the $\{l, w, h\}$ is expressed in the form of length, width,and height of each building (or obstacle). The goal of the coverage control is to find the optimal external parameter set of cameras $\hat{\Theta} = \{\hat{\theta}, \hat{\phi}, zo\hat{o}m\}$ from candidates of the external parameter space Θ, which maximize the coverage rate of the monitored region. Θ is expressed in the form of pan (θ), tilt (ϕ) and zoom value ($zoom$).

2.2. Effectiveness Evaluation of PTZ Cameras

Ineffective and low-quality cameras heighten the computational complexity while increasing the coverage redundancy. In order to minimize the influence of redundant cameras, we design an effectiveness evaluation method to measure effectiveness of PTZ cameras. The impact factors of PTZ camera effectiveness contain the maximal cover rate, the average projection angle, and the unique cover rate. The effectiveness evaluation of PTZ cameras is defined as

$$g = \frac{p \cdot cos(\gamma)}{1 - \omega} \quad p \geq 0, 0 \leq \omega \leq 1, 0 \leq \gamma \leq 1 \qquad (1)$$

where $p = s/S$ denotes the maximal cover rate, $\gamma = \frac{\sum_{i=1}^{n} < \overrightarrow{v}, \overrightarrow{cp_i} >}{n}$ denotes the average projection angle, and $\omega =$

s'/S denotes the unique cover rate. More specifically, p denotes the ratio of the maximal coverage area s and the total observing area S, γ is the average angle between the perpendicular vector $\overrightarrow{v}$ of surface and n lines $\overrightarrow{cp}$ between the camera center and n observing points. And ω is the ratio between the unique coverage area s' and the total ovserving area S.

Note that effectiveness evaluation of PTZ cameras is first step to address observation area coverage problems. Also, better quality of cameras is more beneficial for subsequent optimization.

3. COVERAGE CONTROL SCHEME BASED ON VIRTUAL POTENTIAL FIELD

According to potential field theory, each centroid can be treated as a virtual charged particle. However how to introduce perception quality into the potential field is a difficult problem. To this end, we introduce the concept of the scene potential map where virtual particles are subject to the force that push them to positions with higher perception quality. The force of scene potential map is called as the virtual gravity. Based on scene potential map, we propose a novel virtual force analysis method to address the coverage control problem. We first compute the pan angle and tilt angle of each PTZ camera by the final position of virtual particles. Then we involve the region partition method based on perception quality measure to divide the region, which helps us to compute the acceptable zoom value of each camera. Details of our method is described as follows.

3.1. Scene potential map generation

The proposed scene potential map (SPM) describes the distribution that contains both target region importance information and perception quality of single camera. Since the camera location is different, every camera has a unique scene potential map. As we know, targets of most observation tasks are often far away from buildings (or obstacles) and the edges of the scene. That is, the observation importance of the region far from the edge and buildings(or obstacles) is greater than that of the areas close to the edge and buildings (or obstacles). Besides, perception quality of PTZ cameras is relevant to the distance between the camera location and the target point. So we generate the scene potential map by:

$$m_i(P) = \{ \begin{array}{cc} -\frac{distance(P)}{\|P-C_i\|_2} & P \in D_i \\ 0 & P \notin D_i \end{array} \tag{2}$$

where P is the sampling point of the target scene 2D plane, C_i indicates the location of the i-th camera ($i \in \{1, 2, 3, ..., n\}$), and $distance(P)$ indicates the distance between the current point and the nearest "zero-potential point". D_i denotes the point set where the i-th camera can observe every point. To avoid getting trapped in local optima, every map is processed by Gaussian low-pass filtering.

3.2. Virtual force analysis method based on scene potential map

We treat the coverage area centroid of PTZ camera as a virtual charged particle. The simple force laws are defined as follows:

(1) Any particle is subject to the virtual gravity from the scene potential map. The particle under gravity will move to the region with greater scene potential energy.

(2) The repulsive force exists among multiple neighboring centroid points. The repulsive force pushes the particle far away from other particles.

Considering that the centroid represents the coverage area of PTZ camera, the gravity of every particle is the sum of all values of the scene potential map within $\Omega_i = \{P | \|P - C_{idi}\|_2 < r_c\}$, where C_{idi} is the centroid. The virtual gravity is defined as follows:

$$F_{gi} = \sum_{P \in \Omega_i} m_i(P) \quad i \in \{1, 2, 3, ..., N\} \tag{3}$$

The direction of F_{gi} points to the point, with the minimal value, of scene potential map.

The repulsive force between particles is defined as:

$$F_{qi} = \sum_{j \in N_i} k \frac{Q_i \cdot Q_j}{d_{ij}^2} \qquad i \in \{1, 2, ..., N\} \tag{4}$$

where d_{ij} indicates the distance between the two particles and k is the experience coefficient. Q represents the charge of particle. The greater the charge, the smaller the overlap area between the cameras. N_i is the neighboring set of the i-th camera within radius r_c. The direction of F_{qi} follows the parallelogram guideline.

Next, we introduce the virtual force analysis method. First of all, the initial camera orientation is randomly set and virtual particals are randomly distributed in the target scene. Then we compute the total force on every particle $F_i = r_1 \cdot F_{gi} + r_2 \cdot F_{qi}$. The direction of total force F_i is decided by the direction of gavity and repultive force. Next we update the position of every particle to new position $(x_i + F_{ix}, y_i + F_{iy})$ in the next iteration. F_{ix} is the component of F_i on the x-axis and F_{iy} is that on the y-axis. The virtual force analysis method stops when the number of iterations exceeds the preset value or the total force of every particle is smaller than the threshold.

After the final position of centroid is obtained, we can compute the pan angle θ and the tilt angle ϕ of every camera by:

$$\theta = \{ \begin{array}{cc} arctan(\frac{y_i}{x_i}) \cdot 180/\pi + 90 & x_i > 0 \\ arctan(\frac{y_i}{x_i}) \cdot 180/\pi + 270 & x_i < 0 \end{array} \tag{5}$$

$$\phi = 90 - arcsin(\frac{\sqrt{x_i^2 + y_i^2}}{\|C_{idi} - C_i\|_2}) \cdot 180/\pi \tag{6}$$

where (x_i, y_i) denotes the point in the projection plane, C_{idi} and C_i denote 3D positions of the centroid and the camera center, respectively. The initial oriention of the pan angle of PTZ camera is consistent with the negative y-axis.

3.3. Region partition method based on perception quality measure

The region partition method aims to search the finest coverage area of every area. Eqn. 2 roughly characterizes the perception quality of PTZ cameras. We further introduce accurate perception quality measure for region partition method. We extend the 2D camera perception quality measure proposed by Arslan [13] to that for 3D cameras. The 3D perception quality measure is a multiplicative combination of perspective and resolution quality measures as follows.

Perspective Quality: Visual perception quality is known to usually deteriorate away from the optical axis of a camera in the boundary of its field of view due to increased lens distortion, incomplete and nonpersistent visual data [15]. Arslan et al. [13] only considered the effort of the angle of view in horizontal axis. And, we propose a novel perspective quality for the 3D camera model:

$$q_t = \frac{(\cos(\alpha_p) - \cos(\alpha_{p0}))(\cos(\alpha_t) - \cos(\alpha_{t0}))}{(1 - \cos(\alpha_{p0}))(1 - \cos(\alpha_{t0}))} \tag{7}$$

where α_p and α_t are the horizontal and vertical angles between the center line of sight and the ray of the target point, respectively. α_{p0} and α_{t0} are the maximal horizontal and vertical angle of the view of field, respectively.

Resolution Quality: Similar to the resolution Quality proposed by Arslan, our resolution quality is expressed by:

$$q_r = e^{(-\frac{(d-D)^2}{2\sigma^2})} \tag{8}$$

where $\sigma = \frac{D}{3}$, and d represents the distance between the target point and the camera. $D = (zoom \cdot f)/W_{ccd}$ is the camera working distance, which represents the distance between the object and the camera when the camera clearly observes the target region. f is the focal length, W_{ccd} is the width of the camera charge coupled device (CCD), and $zoom$ is the level of Zooming. The perception quality measure is expressed as $q = q_t \cdot q_r$

The region partition method utilizes the perception quality measure to assign points of the target region to the finest camera. The rule is as follows.

$$V_i = \{P \in D_i | q_i(P) \geq q_j(P), \forall j \neq i\} \tag{9}$$

After obtaining the region partions, we try to further find the best zoom value that can improve the resolution of the coverage area. We first set the zoom values of all cameras to maximum. If all target points of the partion are in its camera's sight, the estimated zoom level will be the best zoom value. Otherwise, the zoom value minuses one. If the zoom value reaches the minimum, the iteration stops.

Table 1. The settings of the simulation scene

Information	Value	Information	Value
Scene size	$100 \times 100m^2$	Iteration	50
Camera number	20	Building number	3
Camera height	$10m$	Building height	$20m$
FoV	60°	Building Ratio	[0.03, 0.1]

4. EXPERIMENTAL SIMULATIONS

In this section, we provide numerical evidences demonstrating the effectiveness of the proposed coverage control schemes. We implemented our methods in Matlab R2018a and simulated the scene with parameter settings shown in Table 1. Note that "Building Ratio" in Tab. 1 denotes ratio range between building size and scene size and the area of simulation scene is $100 \times 100m^2$. Camera positions are randomly generated, where intial pan angles and tilt angles of cameras are all $(0°, 90°)$ respectively, which means that the horizontal orientation is parallel to the positive y-axis and the vertical orientation points to the ground. All cameras are simplified to conic view models. We compare our method with two competing methods: distributed coverage control scheme (DCCS) [3] and coverage-enhancing based on virtual force analysis (VFA-ACE) [8]. The former is a state-of-the-art gradient-based algorithm for 3D camera model. The latter is the classic method for 3D model in the VPF field. For these two methods, we adopt the same parameters as suggested in their papers.

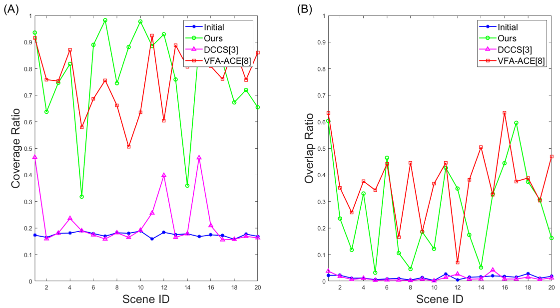


Fig. 2. Coverage performance and ovelap performance with different camera positions.

Coverage Performance and Overlap Performance. We first compare the coverage performance and overlap performance within the observation regions of three methods (see Fig. 2). It can be seen from Fig.2 that our method achieves higher coverage ratio and lower overlap ratio in comparison with the other methods. These results demonstrate that introducing the scene potential map can increase the coverage ratio and reduce the overlap ratio. Because cameras' initial orientation is vertically downward and the gradient generated after each iteration is close to 0, the gradient-based method usually leads to failure. So results of the DCCS is close to initial results.

Coverage Control with buildings. We demonstrate the validity of three methods in irregular-shape regions with buildings (or obstacles) (see Fig.3). According to Fig.3, our method also keeps valid in irregular-shape regions with buildings (or obstacles) and helps the improvement of coverage performance and decreases of overlap performance. The mean coverage ratio in Fig. 3 is lower than that in Fig. 2, which means that it is difficult to find the optimal camera network configuration in irregular-shape regions with buildings (or obstacles).

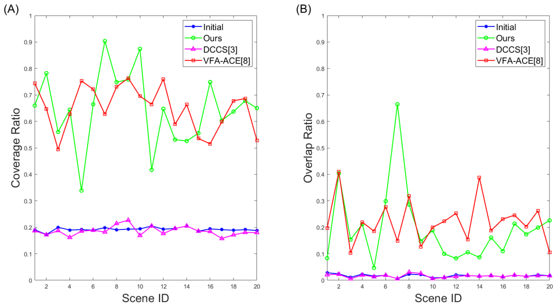


Fig. 3. The results of coverage control in irregular regions with buildings (or obstacles).

Coverage control with concentrated event distribution. In reality, we commonly control the PTZ camera network to cover some important regions such as exits, crossroads, sites of accident and so on. So we simulate the scenes using a Gaussian mixture distrubution determined by means $\{\mu_i | i = 1, 2, ..., K\}$ and variances $\{\sigma_i^2 | i = 1, 2, ..., K\}$ as

$$\phi(x) = \frac{1}{K} \sum_{i=1}^{K} \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left(- \frac{(x - \mu_i)^{\mathrm{T}} (x - \mu_i)}{2\sigma_i^2} \right) \quad (10)$$

where $K = 1, \sigma = 10$, and μ is generated randomly. Experimental results are shown in Tab 2. "Coverage Number" in Tab. 2 represents coverage frequencies of the regions in 20 groups scene data. The coverage ratio and coverage number for important regions of our method are higher than that of other methods and the overlap ratio of our method is close to the VFA-ACE. Therefore, our method can get a better balance between coverage performance and coverage for important regions. Comparing the results of our method with VFA-ACE, we find that our method has a better coverage performance for focal regions, which clearly verifies the effectiveness of the our proposed scene potential map.

Visual results of different methods in simulation environments: We reveal the visual results in simulation environments by VFA-ACE method and our method. To this end, We construct four simulation scenes: two simulation scenes with buildings (or obstacles) and two simulation scenes with concentrated event distributions, as shown in Fig.4. The VFA-ACE method considers only the view interval between cameras, leading to worse results whichever the simulation scenarios. Our method pushes the viewpoints of the PTZ cameras to move away from buildings (or abstacles) and secne edges (see Fig 4(a3, b3)). Also, in scenes with concentrated event distributions our method updates the viewpoints of PTZ cameras to the positions close to the event centers (see Fig 4(c3,d3)). The experimental results show that our method can realize a fine balance between coverage performance and different coverage tasks.

Table 2. The results of coverage control with concentrated event distribution

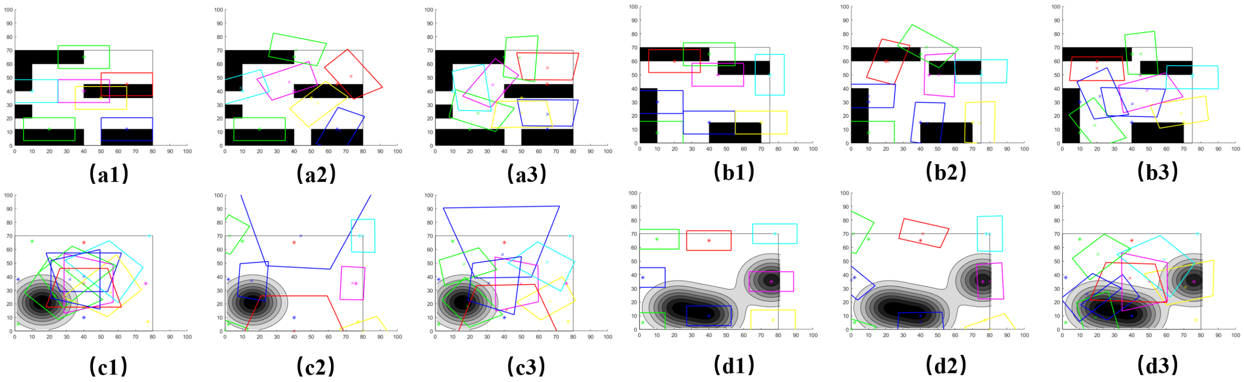
Method	Coverage Ratio	Overlap Ratio	Coverage Number
Initial	17.63%	1.23%	1
Ours	**84.42%**	42.16%	**18**
DCCS [3]	37.44%	**8.96%**	5
VFA-ACE [8]	77.72%	37.42%	15

5. CONCLUSION

We propose a novel virtual force analysis method for PTZ camera networks based on scene potential map, giving rise to a fine balance bewteen the coverage performance of PTZ camera networks and targets of different tasks. Firstly, we construct a scene potential map containing both the scene information and the event information. Then we design a new virtual force analysis method, in which viewpoints of PTZ cameras are treated as virtual charged particles that are prone to moving to more important regions. Furthermore we develop a region partition method based on perception quality measure, thereby resulting in better region partitions and zoom values. Finally, we demonstrate superior performance of our method in different simulation scenes. One limitation of this work is that we have not considered the situation that PTZ camera network may not cover the entire scene, and this issue will be the topic of our follow-up works.

6. ACKNOWLEDGEMENT

This work was supported by National Key R&D Program of China (Grant No.2018YFB2100601) and National Natural Science Foundation of China (Grant No.61872024). The authors would like to thank the anonymous reviewers for their critical and constructive comments and suggestions.



(a1) (a2) (a3) (b1) (b2) (b3)

(c1) (c2) (c3) (d1) (d2) (d3)

Fig. 4. Visual results of different methods in simulation environments. (a) and (b) represent scenes with buildings. (c) and (d) represent scenes with concentrated event distributions. (1) is the initial configuration of the PTZ camera network. (2) is the final configuration of the PTZ camera network by VFA-ACE [8] and (3) is that by our method. ”*” is the representative of a camera position and ”x” is the representative of a camera viewpoint. The shadow area represents the concentrated event distribution and the black blocks represent the buidings (or obstacles).

7. REFERENCES

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

[2] Claudio Piciarelli, Lukas Esterle, Asif Khan, Bernhard Rinner, and Gian Luca Foresti, “Dynamic reconfiguration in camera networks: A short survey,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 26, no. 5, pp. 965–977, 2015.

[3] Qier An and Yuan Shen, “Distributed coverage control for mobile camera sensor networks with anisotropic perception,” *IEEE Sensors Journal*, 2021.

[4] Joseph O’rourke et al., *Art gallery theorems and algorithms*, vol. 57, Oxford University Press Oxford, 1987.

[5] Katie Laventall and Jorge Cortés, “Coverage control by robotic networks with limited-range anisotropic sensory,” in *2008 American Control Conference*. IEEE, 2008, pp. 2666–2671.

[6] Bruno Hexsel, Nilanjan Chakraborty, and Katia Sycara, “Coverage control for mobile anisotropic sensor networks,” in *2011 IEEE International Conference on Robotics and Automation*. IEEE, 2011, pp. 2878–2885.

[7] Yiannis Kantaros and Michael M Zavlanos, “Distributed communication-aware coverage control by mobile sensor networks,” *Automatica*, vol. 63, pp. 209–220, 2016.

[8] Huadong Ma, Xi Zhang, and Anlong Ming, “A coverage-enhancing method for 3d directional sensor networks,” in *IEEE INFOCOM 2009*. IEEE, 2009, pp. 2791–2795.

[9] Li Yupeng, Ji Peng, Liu Hongchao, and Jiang Jingqi, “A virtual potential field based coverage-enhancing algorithm for 3d directional sensor networks,” in *2012 6th International Conference on New Trends in Information Science, Service Science and Data Mining (ISSDM2012)*. IEEE, 2012, pp. 225–230.

[10] Xuan Ma and Jingyan Kang, “A multi-detecting point based on virtual force-directed particle swarm optimization algorithm for coverage enhancement in directional sensor networks,” in *2018 13th IEEE Conference on Industrial Electronics and Applications (ICIEA)*. IEEE, 2018, pp. 1195–1200.

[11] Alexander Davydov and Yancy Diaz-Mercado, “Sparsity structure and optimality of multi-robot coverage control,” *IEEE Control Systems Letters*, vol. 4, no. 1, pp. 13–18, 2019.

[12] Yishan Su, Lei Guo, Zhigang Jin, and Xiaomei Fu, “A voronoi-based optimized depth adjustment deployment scheme for underwater acoustic sensor networks,” *IEEE Sensors Journal*, vol. 20, no. 22, pp. 13849–13860, 2020.

[13] Omur Arslan, Hancheng Min, and Daniel E Koditschek, “Voronoi-based coverage control of pan/tilt/zoom camera networks,” in *2018 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2018, pp. 5062–5069.

[14] Dan Tao, Hua-Dong Ma, and Liang Liu, “A virtual potential field based coverage-enhancing algorithm for directional sensor networks,” *Journal of Software*, vol. 18, no. 5, pp. 1152–1163, 2007.

[15] Dimitra Panagou, Dušan M Stipanović, and Petros G Voulgaris, “Distributed dynamic coverage and avoidance control under anisotropic sensing,” *IEEE Transactions on Control of Network Systems*, vol. 4, no. 4, pp. 850–862, 2016.