Received November 27, 2018, accepted January 2, 2019, date of current version February 8, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2894524

Hybrid Stochastic Exploration Using Grey Wolf Optimizer and Coordinated Multi-Robot Exploration Algorithms

KAMALOVA ALBINA AND SUK GYU LEE

Department of Electrical Engineering, Yeungnam University, Gyeongsan 38541, South Korea

Corresponding author: Suk Gyu Lee (sglee@ynu.ac.kr)

This work was supported by the Basic Science Research Program, through the National Research Foundation of Korea, Ministry of Science, under Grant 2017R1D1A3B04031864.

ABSTRACT Multi-robot exploration is a search of uncertainty in restricted space seeking to build a finite map by a group of robots. It has the main task to distribute the search assignments among robots in real time. In this paper, we proposed a stochastic optimization for multi-robot exploration that mimics the coordinated predatory behavior of grey wolves via simulation. Here, the robot movement is computed by the combined deterministic and metaheuristic techniques. It uses the Coordinated Multi-Robot Exploration and Grey Wolf Optimizer algorithms as a new method called the hybrid stochastic exploration. Initially, the deterministic cost and utility determine the precedence of adjacent cells around a robot. Then, the stochastic optimization improves the overall solution. It implies that the robots evaluate the environment by the deterministic approach and move on using the metaheuristic algorithm. The proposed hybrid method was implemented on simple and complex maps and compared with the Coordinated Multi-Robot Exploration algorithm. The simulation results show that the stochastic optimization enhances the deterministic approach to completely explore and map out the areas.

INDEX TERMS Multi-robot systems, robot sensing system, hybrid intelligent systems, optimization.

I. INTRODUCTION

The studies in the mobile robotics field involve a wide range of topics. Some of the well-known applications are path planning, navigation, localization, communication, and sensing, that work on preconditioned maps of environments. Exploration of an unknown area begins with having no knowledge regarding the arrangement of obstacles, nor the layout of terrain. The primary goal of exploration is to create a finite map.

In the present day and without less interest in the past, exploration is applied in search and rescue, reconnaissance, surveillance, data gathering, and simple indoor moving applications. It pursues to explore full space without supervised navigation using the autonomous multi-robot system. Compared to a single agent, a group of exploration robots can enhance the space coverage and decrease search time. However, the motion policy in the process requires efficient techniques to keep them safely for providing free driving.

Several well-known map forms are available to represent configuration space that fundamentally determines the selection of the algorithm. In the exploration task, maps are usually represented by occupancy grids, which can change the unknown to known modes of cells through every robot position upgrade. The relevant problem through the process is the scheduling of detection of the uncertainty of grid occupancy map using onboard sensors on robots.

Looking at the exploration methodologies, the frontier-based [1] and coordinated multi-robot explorations [2] are introduced solutions for one robot and a group of robots, respectively. The methods seek to reach the border of the known and unknown line with the least traversable cost. The utility allows consideration of the coordinated motion when a team of robots is applied. It targets to decrease the interest of other robots to move in the same direction.

In this paper, we aim to develop the hybrid method of the coordinated multi-robot exploration (CME) with the stochastic approach. As the deterministic method, CME may get into a circumstance when the task of full coverage is not achieved in a particular place. In this case, the solution will not be reached except as to change the objects of physical place or initial robot parameters, which is not possible in some hazardous conditions. Applying the metaheuristics, the probabilistic variations induce CME to search the values, for which the task to explore the place completely will be fulfilled.

By stochastic ones or so-called metaheuristics, the methods are classified into single and population solution-based. In comparison to the former, the population-based method has a number of solutions, which are updated iteratively until the termination condition is completed. One of the categories is Grey Wolf Optimizer (GWO) algorithm that was introduced by Mirjalili *et al.* [3]. It imitates the social behavior of grey wolves that is denoted as operators such as hunting, searching, encircling, and attacking for a prey. Abstracting from the nature wolf model, the prime principle of this optimizer relies on searching the best solutions and changing the current state according to them.

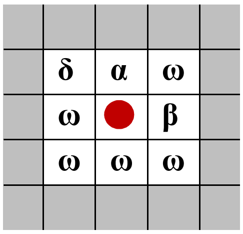

FIGURE 1. The representation of cell distribution around a robot into alpha, beta, and delta as optimized solutions. The rest of the cells is represented by omegas or other available solutions.

In the proposed hybrid stochastic exploration, the costs and utilities of cells around a robot provide knowledge over each robot step. The social hierarchy operator of GWO selects alpha, beta, and delta among the cells (Fig.1). Then, the hunting operator obliges the robot to move according to the value of the best cell that is formulated from the occupancy probability and the random parameters.

The remainder of this paper is organized as follows: section II discusses the previous studies of GWO, CME and some efforts of applying nature-inspired algorithms in mobile robotics. Section III and IV review the mathematical models of these two methods. Section V describes the proposed hybrid stochastic exploration, and section VI presents the simulation results and analysis. Finally, section VII sums up the study with the conclusion.

II. RELATED WORKS

This section discusses the previous researches applying GWO and CME in two separate subsections, and some studies of bio-inspired optimization techniques in cooperative exploration field using the multi-robot system at the end of the section.

A. GREY WOLF OPTIMIZER AND ITS APPLICATIONS

In the recent decade, researchers had imitated the numerous numbers of optimization models of various animals in metaheuristics. They gained knowledge of the distinctive behaviors of animals, that is peculiar to some species, in which a swarm uses the knowledge and experience obtained through evolutions, physical rules, searches of food, or emitted sounds. Some of the well-known methods are the genetic algorithm (GA), particle swarm optimization (PSO), ant colony optimization (ACO), gravitational search algorithm (GSA) and differential evolution (DE).

Assessing the nature of wolves, researchers were able to formulate mathematical expressions revealing their social behavior in terms of hierarchy distribution of roles in a pack, hunting, search for prey, and attacking strategies. It was observed that prey hunting is optimal at the expense of the wolves' collaboration among each other.

Liu *et al.* [4] first proposed the mathematical model that is called wolf colony algorithm (WCA). This algorithm describes the searching behavior and the besiege of quarry and updates the wolf colony according to the assignment rules of the colony. WCA was applied in the path planning motion for a mobile robot that computed the optimal length of a path in the shortest time compared to PSO and GA.

Another approach [5] is called wolf pack algorithm (WPA), which is an enhancement of WCA. It improved the foregoing drawbacks such as the efficiency and local optima fall and described two additional intelligent rules: winner-take-all generation of a lead wolf and stronger-survive renewing of a wolf pack. The former rule obliges the pack to follow the lead wolf. It can be replaced through iterations, but the tendency of obeying continues. The stronger-survive renewing rule is the distribution onto weak and strong wolves according to the objective function.

GWO differs from the WCA and WPA techniques in terms of the mathematical expressions. It demonstrates superior optimization capability over other metaheuristic algorithms. The benchmark function tests proved the advantage of GWO on the exploitation, exploration, and local optima avoidance analyses. This makes it possible to apply GWO for applications in the various domains, such as the clustering analysis [6], the distributed compressed sensing [7], the human recognition [8], the air quality classification [9], the generating schedules [10], the image thresholding [11], the PI controller for steam condenser [12], and the optimal power flow [13].

The modification of algorithms is a regular practice in metaheuristics. In a previous study [14], the random parameters of GWO were modified to achieve the balance in finding the global minimum with fast convergence speed. The next alternative modification [15] is a novel algorithm for multi-criterion optimization that was presented as Multi-Objective Grey Wolf Optimizer (MOGWO). The two functions were integrated into the original GWO that are retrieving the best non-dominated solution and a leader selection mechanism. MOGWO passed ten multi-objective benchmark tests and was compared with other two multi-objective metaheuristics where it demonstrated the highest convergence behavior. Another approach called modified discrete grey wolf optimizer (MDGWO) [11] updates the location of search agents by the introduction of weighting parameters that improves the optimal solution.

The hybrid variants of metaheuristics combine two or more existing methods to create new one. The hybrid nature-inspired algorithm that is called HPSOFWO [16] enhances the exploitation of PSO and the exploration of GWO. The low-level coevolutionary mixed hybrid updates three agents in the search space by the proposed inertia constant and velocity.

The hunting strategies by Muro *et al.* [17] asserted that knowing every agent position is enough for the formation of the wolf-pack hunting behavior. The study concluded that the communication and hierarchal division in a group is not necessary to achieve the hunting task. However, Rodríguez *et al*. [18] formed a contradicting conclusion about the hierarchical pyramid and introduced a new fuzzy hierarchical operator in GWO. They presented three variants to implement the hierarchical pyramid that affects the new position depending on the alpha, beta, and delta wolves: weighted average, weighted-based on the fitness, and fuzzy weights. The benchmark tests showed that the fuzzy weights improved the performance of the algorithm.

B. FRONTIER-BASED COORDINATED EXPLORATION

The proposed method considers the search uncertainty by a group of sensors. During the process, the multi-robot system tries to maximize knowledge over a given area. It moves to targets in each step, which are edges of sensor range between known and unknown space. The edge is the frontier, and the algorithm is called the frontier-based approach for autonomous exploration [1]. It was applied in many studies of the exploration field [19], [20].

Coordinated Multi-robot Exploration (CME) [2], [21] is a method which was hybridized in our study. It is an enhanced variant of Yamauchi's algorithm. The multiple robots collaborate in the exploration work as a group. It considers the cost of reaching frontiers and utility for robot divergence, which are described in section IV in details.

Benkrid and Achour [22] proposed a novel approach for the coordinated exploration that seeks to minimize the search time in terms of the limited energy consumption of the robots. In the original CME approach, robots share information about positions throughout the process, but in this novel approach, the remaining energy of the batteries is transferred as well. This allows all the individual robots to reach the assigned frontier points. Rappaport and Bettstetter [23] also worked on the coordinated recharging of multi-robots for exploration.

Senarathne and Wang [24] aimed for a balanced distribution of robots to explore unstructured environments. They applied two approaches for the exploration. The first one is the original CME and the other one is the proposed approach for the repositioning of robots over the environment. Kim *et al*. [25] proposed a similar hybrid variant to combine the sensor-based random tree method with the frontier-based coordinated exploration. The simulation demonstrated the efficient backtracking driving by decreasing the number of routes.

Puig *et al.* [26] had a similar purpose as our study, but they only applied the deterministic global optimization for CME. The exploration was improved by K-Mean clustering that gives each robot different assignments to travel separate places (K regions) simultaneously at the same speed. The approach resulted in the lowest variance of regional waiting time and the lowest variance of average waiting time of all regions consumed by the motion of multi-agents during the exploration process.

C. NATURE INSPIRED OPTIMIZATION APPROACHES FOR MULTI-ROBOT EXPLORATION

Bio-inspired heuristic methods relate to a stochastic global optimization class that does a random search of an optimal solution using swarm intelligence. Even though the swarm intelligence is widely known in theory, there is a specific research field called swarm robotics where a fixed number of mobile robots are controlled physically in real world in a certain coordinated way. The swarm robotics is used mainly in the formation control [27], [28].

By taking into consideration the terms of the bio-inspired optimization and the multi-robot exploration, some related works can be mentioned here as summarized in Table 1.

Fang *et al*. [29] applied the behavior-based method called social potential fields to obtain the course direction and move the robots toward unexplored areas. Then, they optimized the system by fine-tuning the angle and speed of each robot. The simulation showed that the most optimal coverage of unknown space by the multi-robot system with robot speed from 0.7 to 1.0 m/s and angle in the range −0.2 to 0.2 rad. This type of deterministic optimization has minimal effect compared to the present trends and does not take into account the other requirements of the robot system.

Wang, Y. *et al*. [30] applied the frontier-based method, A star and PSO in two stages. In the exploration stage, the method discovers available frontier points around each robot in its own subarea and selects the shortest distance using A star algorithm. In the walking stage, each robot is navigated through subareas according to the latest exploration information and the robot positions using PSO algorithm. It means that PSO is applied for the task assignment in the clustered area, like in the previously mentioned study [26].

PSO algorithm is widely used in mobile control systems because of its concept of upgrading the position and velocity of a swarm. Wang *et al*. [31] proposed and explored modified PSO algorithms such as Darwinian PSO (DPSO), robotic DPSO (RDPSO), fractional order RDPSO (FORDPSO), and fuzzy adaptive FORDPSO. The last two were proposed to adjust control coefficients, especially in multi-robot exploration problem.

The combined Clustering Based Distribution Factor (CBDF) and nature-inspired algorithm (NIA) method [32] is a hybrid approach to probe unknown areas. It is divided into the direction- and exploration-based movements. Each robot is assigned a direction by the CBDF to a specific subarea. For the exploration, three NIAs such as PSO, bacteria foraging optimization (BFO), and bat algorithm (BA), are compared to determine their efficiency.

TABLE 1. The efforts of optimization by nature-inspired approaches in the multi-robot exploration process.

Approaches	Description	Optimization	Exploration
Fang et al. [29]	Adjusted velocity and coarse direction parameters for the optimal solution	Numerical optimization	Social potential field
Wang et al. [30]	Applied two stages: exploration and walking. Robots explore frontier points in local sub-regions (exploration). Then, they can move other regions navigated by PSO (walking)	A star, PSO	Frontier-based
Fuzzy adaptive FORDPSO method [31]	Analysis and comparison of the influence of coefficients fractional order RDPSO and fuzzy adaptive FORDPSO in the exploration problem	Fuzzy adaptive FORDPSO	Reward and punishment mechanism as nature selection concept for swarm
CBDF and NIA method [32]	Applied directional and exploring based movements. Robot explores in the direction provided by CBDF and search short ways to reach up in the area using NIA	PSO, BFO, BA	CBDF

To sum up, individual or hybrid stochastic optimization algorithms, such as PSO, BFO, and BA, have been used for explorations to create maps but without the GWO, which is superior among existing methods based on benchmark testing. Likewise, deterministic optimization techniques such as terrain clustering, multitasking of robot position, increasing population, and swarm formation control have also been applied in the exploration processes. However, the significant drawback of the deterministic algorithms is their property to fall at a local optimum rather than search for a global optimum or the best solution. To unravel this issue, the GWO was combined with CME algorithm.

III. GREY WOLF OPTIMIZER

The main feature of GWO that makes its effective compared to other popular swarm intelligence algorithms is its hierarchical structure. The dominance hierarchy is formed according to a certain goal that is called the objective function. In turn, the objective function is classified into the cost function, evaluated lost, and the fitness function, which are used to summarize how accurate the final result is compared to the given design solution [33]. Whether the best solution is defined to be optimal among all available candidates, it should satisfy the fitness function and the cost function that are regarded as interchangeable functions of maximization and minimization.

The wolf pack is divided into four dominant ranks: alpha, beta, and delta wolves, which are the leading groups, maintaining the priority in the same sequence. The fourth group comprises omega wolves, which do not have any rights to make decisions in a swarm, although, their presence determines the swarm intelligence, *i.e.* high local optima avoidance. In terms of optimization, the swarm of solutions is filtering agreement with the objective functions that bring the fittest one as alpha (α), the second as beta (β), and the third as delta (δ). The strong point of the social hierarchy is that only leading wolves know the position of prey, and they guide omegas to perform the search.

The organization of collective behaviors in wolf pack can be described by operators such as social hierarchy, encircling prey, hunting, attacking prey (exploitation), and search for prey (exploration). GWO algorithm is mostly about hunting behavior, the technique of how wolves search collectively. It means that alpha, beta, and delta wolves occupy the best positions $\mathrm{X}_\alpha, \mathrm{X}_\beta, \mathrm{X}_\delta$ (1). They oblige omega wolves to accept the average distance between them (3).

$$\begin{aligned} \vec{D}_\alpha &= \left|\vec{C}_1 \cdot \vec{X}_\alpha - \vec{X}\right|, \\ \vec{D}_\beta &= \left|\vec{C}_2 \cdot \vec{X}_\beta - \vec{X}\right|, \\ \vec{D}_\delta &= \left|\vec{C}_3 \cdot \vec{X}_\delta - \vec{X}\right| \end{aligned} \tag{1}$$

$$\begin{aligned} \vec{X}_1 &= \vec{X}_\alpha - \vec{A}_1 \cdot \left(\vec{D}_\alpha\right), \\ \vec{X}_2 &= \vec{X}_\beta - \vec{A}_2 \cdot \left(\vec{D}_\beta\right), \\ \vec{X}_3 &= \vec{X}_\delta - \vec{A}_3 \cdot \left(\vec{D}_\delta\right) \end{aligned} \tag{2}$$

$$\vec{X}(t+1) = \frac{\vec{X}_1 + \vec{X}_2 + \vec{X}_3}{3} \tag{3}$$

The presence of vectors A and C makes GWO a stochastic algorithm. The vectors fluctuate randomly in defined ranges that helps avoid the local minima.

$$\vec{A} = 2\vec{a} \cdot \vec{r}_1 - \vec{a} \tag{4}$$

$$\vec{C} = 2 \cdot \vec{r}_2 \tag{5}$$

The A parameter defines the exploitation and exploration operators (4). If $|A| < 1$, wolves are obliged to attack the prey, otherwise, for $|A| > 1$, they turn to the exploration operator. The component of a is decreased linearly from 2 to 0 in each iteration and r_1, r_2 are random vectors in [0,1].

The C parameter gives the prey the weight that makes the exploration of the optimal solution more naturally obstructing the search (5). It varies randomly as well in range [0,2], where $C > 1$ simplifies and $C < 1$ complicates the process.

In this note, the hunting operator contains exploration and exploitation rules that can be adjusted by the A and C values. The α, β, and δ wolves lead the pack through the hunt, and the arbitrary parameters assist the search effectively.

GWO's performance was compared with other population-based algorithms using the benchmark functions on Table 1, Table 2, Table 3 in the original paper [3]. The aim of the optimization is killing the prey, which relates to its life reduction f_{min} by wolves attacking. Table 1 shows the scenario when thirty wolves attack the prey individually one after the other. When the last wolf kills the prey, then f_{min} gets a zero value. In Table 2 and Table 3, wolves attack collectively, and the injury dealt by the wolves on the prey varies. A value that is greater than zero means that the prey is still alive after the attacks. The negative values suggest that the wolves wasted needless powers for killing the prey. In Table 3, the fixed values for the dimension limits the number of wolves that can attack the prey.

Algorithm 1 Grey Wolf Optimizer algorithm

Initialization *population*, *iteration*, *search boundary*
Set random *position*
while *iteration* is not over **do**
    Calculate *costs* of wolf positions
    Find α, β, δ wolves
    for all *population*
        Find D_α, D_β, D_δ, X_1, X_2, X_3
        Change positions $X = (X_1 + X_2 + X_3)/3$
        Change costs
    end
    Change parameters a, A, C
end while

In our study, we tested GWO (Algorithm 1) with unimodal sphere function f_1 from Table 2 [3]. The results showed fast convergence for the first few iterations of the simulation run (Fig. 2 (a), (b), and (c)). In the two-dimension space with 5×5 bounds, the population of 10 reached optimal zero solution after 20 iterations. The optimization trend led by the alpha wolf is depicted in Fig. 2 (d).

The aim of killing a prey for wolves in nature demonstrates the high-efficient optimization that can be applied in various fields. In our study, we used GWO in the map coverage problem for the mobile robot system.

Aside from the basic merits of GWO which include ease of implementation due to its structure, lesser memory requirement compared to other techniques, and faster convergence rate because of the continuous reduction of search space and lesser judgment variables α, β, δ, the main advantage of this algorithm is that it evades the local optima when applied to composite functions and only two parameters need to be adjusted (A & C). The only disadvantage of this optimization technique is that, in case of unimodal problems, initially it hastens towards the optimal solution but starts slowing down soon due to diversity problems.

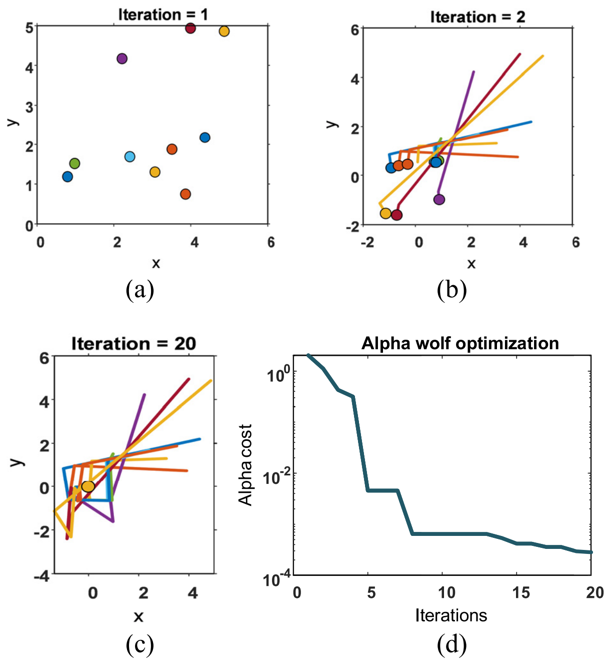

FIGURE 2. GWO performance that calculates the cost using sphere function in iteration 1 (a), in iteration 3 (b), and in iteration 20 (c). The alpha wolf trajectory in the search of the optimal solution (d). Population = 10. The alpha cost in iteration 1 is 2.0352, iteration 2 is 1.1173, and iteration 20 is 0.00028223. The runtime of 20 iterations is 665.467 s.

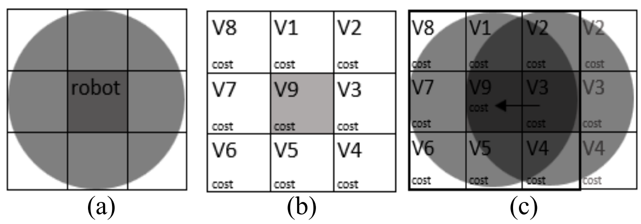

FIGURE 3. Representation of sensor observation in the occupancy grid map: a) the sensor touches eight neighbor cells, b) the neighbor cells have the name *V1, V2, V3, V4, V5, V6, V7, V8, V9* and cost, c) the costs *V6, V7, V8* do not have the intersecting sensor observations.

IV. COORDINATED MULTI-ROBOT EXPLORATION

Multi-robot exploration is the search process with the mobile robot team that begins from entire uncertainties to a finite map. Based on the obtained data, different algorithms are available for exploration. Two methods can be considered for the map building based on the communication between robots. The centralized exploration is when all robots have one common map. They sense the environment simultaneously that allows them to know about the progress of each other. The decentralized exploration is the individual map building [34]. The coordination of sharing data is needed only if robot locations intersect. In our study, the centralized strategy is applied, which computes the cost of travelling distance locally for each one in real time and the utility values, which all robots upgrade through iterations.

In the beginning, the robot is surrounded with complete unknown space of the indoor environment in the boundaries. The initial position and sensor vision with 360 degrees of view are given. The sensor range is limited for making a complete map at once and for planning the optimal motion of the robot team in advance. The cost-minimizing exploration computes the distance needed to reach the frontier point for each robot. The sensor view in terms of occupancy grid map covers several cells around a robot (Fig. 3). For these cells, the cost is computed using occupancy probability, Euclidean distance, and the sensor observation (6). If the cell has been explored before, then the cost of this cell in the previous step is appended to the cost of the new position. Otherwise, if the ray beams open the cell primarily, then the cell is denoted as a frontier cell without the backward costs of the previous steps (7). The occupancy probability value $P(occ_{x,y})$ of the unknown cell is 0.500, whereas the cell occupied by an obstacle has a value close to 1, and a value close to 0 represents the certainty that the cell is not occupied and obstacle-free. Depending on the probability that the sensor covers the cells at certain distance, the occupancy values decline whenever the sensors touch the cell, as shown by equation (2) in the original paper [2].

$$V_{x,y} = min\{V_{x+\Delta x,y+\Delta y}+\sqrt{\Delta x^2+\Delta y^2}\cdot P\left(occ_{x+\Delta x,y+\Delta y}\right)\} \tag{6}$$

$$V_{x,y} = \min\{\sqrt{\Delta x^2+\Delta y^2}\cdot P\left(occ_{x+\Delta x,y+\Delta y}\right)\} \tag{7}$$

where $\Delta x,\ \Delta y \in \{-1, 0, 1\} \wedge P(occ_{x+\Delta x,y+\Delta y}) \in [0, occ_{max}]$

The aim of the cost-minimization is to find the minimal value among the cell neighbors, which is the optimal next position of a robot. For a single mobile robot system, the search of minimal cost can be enough to determine the position. However, the multi-robot system requires the collective-organization interaction during the exploration. The CME approach introduced the utility for the arrangement of the tasks between robots.

The essence of the maximizing the utility is that initially each cell of the whole map had identical values. While the robots search, the utilities of their frontier cells decreases (8). The robots have less interest to visit the cells with low utilities. This is the reason why the robots try to search for new areas, which they have not yet explored to maximize utility values.

$$U_i^c = U_{i-1}^c - P(\left\|occ_{x,y}^c - occ_{x,y}^r\right\|) \tag{8}$$

The cell utility U_i^c equals the state of the previous modifications U_{i-1}^c, which can be changed by itself or other robots before, and the probability occupancy of the selected cell subtracting the current robot position. The utility U_i^c is opted as the maximum value by (9) in the iteration i.

$$(i, c) = max\left\{U_i^c - V_{x,y}\right\} \tag{9}$$

Towards cooperative operation, the robots should start to run so that their sensor scans reach each other at the first iteration (Fig. 4). It allows achieving the divergence of directions in the search by the decreasing utilities of selected targets. In the figure, the area is 20 m × 20 m in size, and the sensor ray length is 1.5 m. The CME with three robots completes the search after 100 iterations. The approach demonstrates the efficient exploration for a multi-robot system. Nevertheless, we found some points for improvement. First, the cost parameter seeks for the minimal waste; in the case when there are more than two identical minimal values, the last one is selected. Second, when the cost and utility consist of the same values, the robot can get stuck in one position on an explored space, even when other areas are still unexplored. Therefore, it is necessary to find the solution to enable the robot to keep searching for unexplored area.

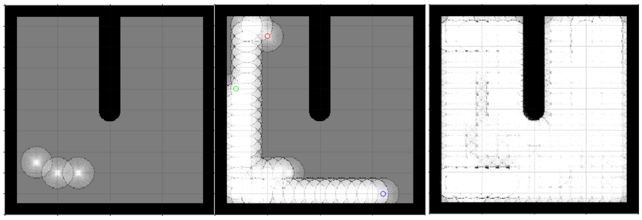

FIGURE 4. Coordinated exploration of three robots by CME approach at iterations 1, 15, and 100. The grey color defines the uncertainty, the black obstacles, and the white explored space.

In regards with CME, the robots have one best next position in each iteration, which is the maximum value of utility. This paper applied the hierarchical GWO optimization technique with CME, in which there are three best maximum positions. Depending on the stochastic parameters, the next robot position will be defined.

V. PROPOSED METHOD

In this study, the problem is formulated as the exploration of an unknown space by robot motion with sensor coverage because some studies focus only on research concerning static sensor coverage errors and robot motion in unknown environments [35], [36]. In the case where the sensors are mobile, the two terms coincide in meaning and scope. In other words, both studies have the common task to build a finite map.

Here, we optimize the process of constructing a map by a group of mobile robots. As the optimization technique, the GWO metaheuristic algorithm was used with modification on the task of sensor coverage. It generates random parameters for the maximum positions that change the order. Thus, the hierarchical optimizer reforms the robot position selection, which is the deterministic exploration approach. In cases where the optimization is not possible in real-time processes, a stochastic technique in the absence of a priori information about the environment is performed as one of the profound solutions of area coverage for the mobile sensor system.

Algorithm 2 describes the proposed hybrid exploration. Initially, space is unknown with the utility equal to 1. The eight cells V_c, where V_c is $V1_{x,y}$, $V2_{x,y}$, …, $V8_{x,y}$, around a robot are the candidates for the next position. The deterministic technique computes the cost and subtracts the utilities

Algorithm 2 Coordinated Multi-Robot Exploration With Grey Wolf Optimizer

1: Initialization the number of robots *nRbt* and iterations *t*, sensor range, initial position
2: Set utility of unknown space 1
3: **while** *t* is not over **do**
4: **for** all *nRbt*
5: Set coordinates of V_c
6: Calculate cost of V_c
7: Subtract U_c^i and V_c
8: Find α, β, δ wolves among values of step 7
9: Find D_α, D_β, D_δ, X_1, X_2, X_3
10: Find X(t + 1) as max (X_1, X_2, X_3)
11: Change robot position X(t+1)
12: Reduce U_c^i on X(t+1)
13: **end for**
14: Calculate a, A, C parameters
15: **end while**

from the cost for the eight cells (9). Then, the metaheuristic optimizer defines three maximum utility values that are assigned as alpha, beta, and delta candidates with priorities in the listed order (line 8). The hunting operator changes the priorities between them due to the random A (4) and C (5) parameters and the occupancy probability values of the dominated grid cells. The hunting operator for the area coverage problem is defined in (10) and (11). The robot's next position is either alpha, beta, or delta cell with the maximum value of $X_{1,i}$, $X_{2,i}$, $X_{3,i}$, where i is the number of robots.

$$\begin{aligned} D_{\alpha,i} &= \left|C_1 \cdot P_{\alpha,i}\left(occ_{x+\Delta x,y+\Delta y}\right) - P_i\left(occ_{x+\Delta x,y+\Delta y}\right)\right| \\ D_{\beta,i} &= \left|C_2 \cdot P_{\beta,i}\left(occ_{x+\Delta x,y+\Delta y}\right) - P_i\left(occ_{x+\Delta x,y+\Delta y}\right)\right| \end{aligned} \tag{10}$$

$$\begin{aligned} D_{\delta,i} &= \left|C_3 \cdot P_{\delta,i}\left(occ_{x+\Delta x,y+\Delta y}\right) - P_i\left(occ_{x+\Delta x,y+\Delta y}\right)\right| \\ X_{1,i} &= P_{\alpha,i}\left(occ_{x+\Delta x,y+\Delta y}\right) - A_1 \cdot D_{\alpha,i} \\ X_{2,i} &= P_{\beta,i}\left(occ_{x+\Delta x,y+\Delta y}\right) - A_2 \cdot D_{\beta,i} \\ X_{3,i} &= P_{\delta,i}\left(occ_{x+\Delta x,y+\Delta y}\right) - A_3 \cdot D_{\delta,i} \end{aligned} \tag{11}$$

In the original GWO, the solution is the mean value of X_1 X_2, and X_3 (3). It is related to the natural behavior of a wolf pack that uses the intelligence of dominant agents. However, it is not required in the target selection problem to find the average robot positions among the alpha, beta, and delta grid cells. By taking this into consideration, the next robot position *X(t+1)* is the maximum value among $X_{1,i}$, $X_{2,i}$, and $X_{3,i}$.

When the robot gets the next selected position (line 11), the utility values of the neighbor cells are reduced by (8). At the end of this, new values for the random parameters a, A, and C are generated for the next iteration.

The hybrid stochastic exploration seeks to search uncertainties through the exploration process in the same manner that CME does. This works because the unexplored cells have greater utility values than those of the explored cells. When the costs with minimal values are subtracted from the utilities of the unexplored cells, the maximal values become attractive targets for the next robot positions. This principle is valid for both methods. However, the proposed hybrid stochastic method has three best options that can change the hierarchical order according to the stochastic parameters. It means that the maximal value may have a beta or delta position, not only alpha, as it is in the CME.

The next section demonstrates the conditions when the two algorithms outperform one another.

VI. SIMULATION RESULTS AND ANALYSIS

This section presents the simulation of the hybrid coordinated exploration based on GWO in two maps: ordinary and complex. The proposed hybrid method was compared with the original CME, which we seek to outperform.

Considering that the robots can move in an arbitrary manner, the map coverage is the primary issue and the principal criterion for this type of system. With the aim of analyzing the simulation results for the two approaches, the following equation (12) computes the percentage of total explored grid cells (M_c):

$$M_c = \frac{U_{unexp} - U_{exp}}{U_{unexp}} \times 100\% \tag{12}$$

where U_{unexp} is the total unexplored utility values that are free from obstacles and U_{exp} is the total explored utility values. The comparison between the proposed hybrid stochastic exploration and CME can be done based on the M_c value after the simulation is completed.

The same map parameters were set for both methods in the ordinary and complex maps. The parameters are the number of iterations, obstacles, number of robots, map size, sensor range, and initial robot positions. By taking into consideration that we are comparing the deterministic and stochastic approaches, the former needs to run only once, *i.e.*, the trajectory of the robot's motion stays invariant when the map remains the same, whereas the proposed stochastic method requires finding the best result among the simulation runs.

A. ORDINARY MAP

An ordinary map is an environment with the minimum number of obstacles. It allows robots to have more freeways for motion that makes it easier for them to diverge from each other.

Fig. 5 illustrates the results of the explorations by two methods where three robots have different path colors. The CME achieved slightly better exploration (97.31%) than the proposed hybrid stochastic method on the ordinary map. With 100 iterations and 20 simulation runs, the stochastic approach gave different results each time due to the random A and C values. In Fig. 5, maps 2.a, 2.b, and 2.c show the map coverage of the proposed approach at 69.72% (worst), 85.98%, and 95.83% (best), respectively.

In the CME, the robots always move toward the maximum utility values (α). On the other hand, in the proposed hybrid stochastic exploration the robots can seek any of the three maximum utility values (α, β, and δ) depending on

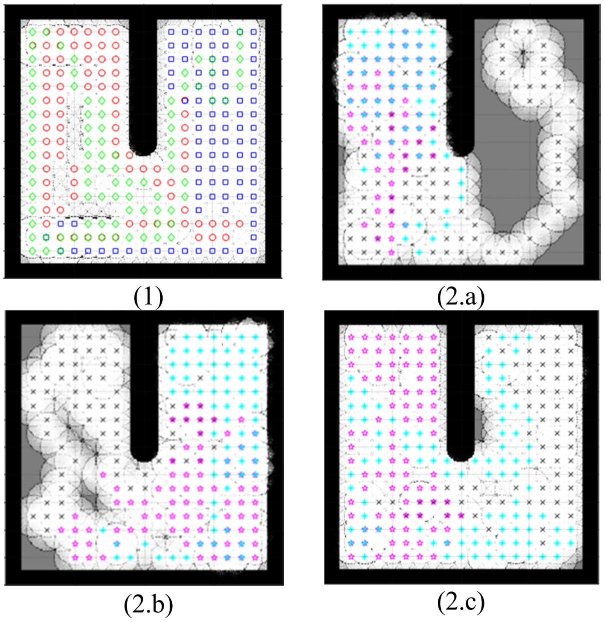

FIGURE 5. The map coverage for the CME (1) at 97.31% and the hybrid stochastic exploration based on GWO and CME at 69.72% (2.a), 85.98% (2.b), and 95.83% (2.c). The number of iterations is 100.

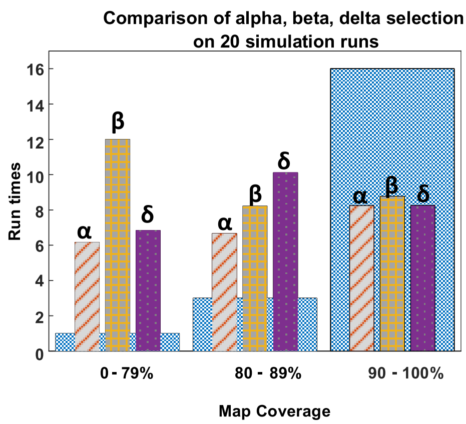

FIGURE 6. The histogram of the hybrid stochastic exploration based on GWO and CME using 20 simulation runs. The subcategories belong to the alpha, beta, and delta cell decisions.

the A and C parameters, which was observed through the 20 simulations mentioned above. The map coverage was classified into three categories: "0 to 79%", "80 to 89%", and "90% and higher". Fig. 6 shows that out of the 20 simulation runs, only one of them has the worst map coverage wherein the robots chose the β positions mostly. Likewise, the other 16 simulation runs in the "90% and higher" category showed that the robots favored mostly β positions instead of the α positions, which is typical of CME. In the second category, the δ positions dominated the runs.

The performance of the hybrid stochastic exploration may be improved when the number of iterations is increased to 130 (Fig. 7b). As seen, the hybrid stochastic method showed greater map coverage than the deterministic exploration in some runs.

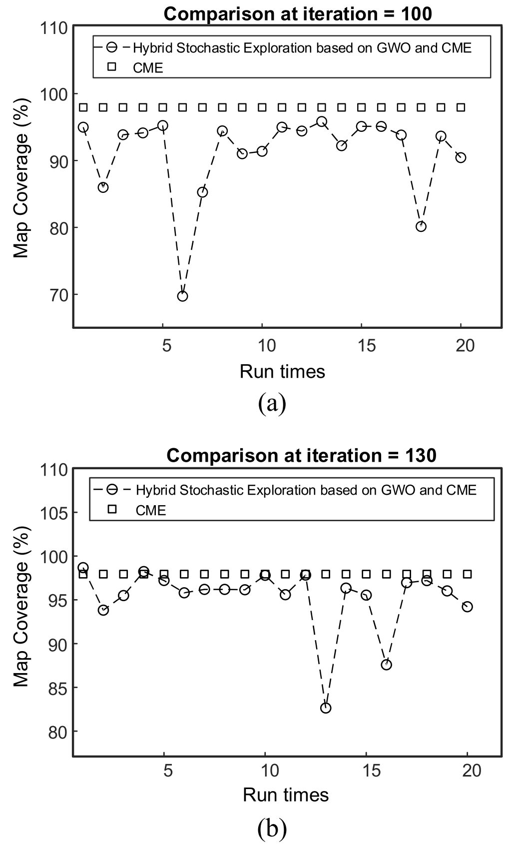

FIGURE 7. The map coverage comparison of CME and the hybrid stochastic exploration based on GWO and CME for the ordinary map with t = 100 (a) and t = 130 (b) iterations through 20 simulation runs. For 100 iterations, the maximum percentage belongs to the CME at 97.31%, while for 130 iterations, the hybrid stochastic exploration gives the maximum value at 98.68%.

B. COMPLEX MAP

Here we consider the algorithm performance in cluttered environments using four maps as shown in Fig. 8 and 9. Both the CME (Fig. 8) and hybrid stochastic algorithms (Fig. 9) were tested for the percentage of map coverage for complex areas.

The CME (Fig. 8) showed a high map coverage only in map 4. Its performance in the other maps was less efficient, and it did not allow improvement in the next iterations because of the invariant nature of the deterministic method. In contrast, the proposed hybrid stochastic exploration based on GWO and CME allowed the search for solutions even after relaunching the simulation runs. The four map results (Fig. 9) demonstrated the strong capability of the proposed hybrid method for the coverage of varying complex maps.

C. THE FEATURE OF THE HYBRID STOCHASTIC ALGORITHM

In using the hybrid approach, the random parameters as additional weights for utilities oblige the robots to estimate the next position, which can lead to obstacle collision problems in some situations. This problem occurs when a robot is

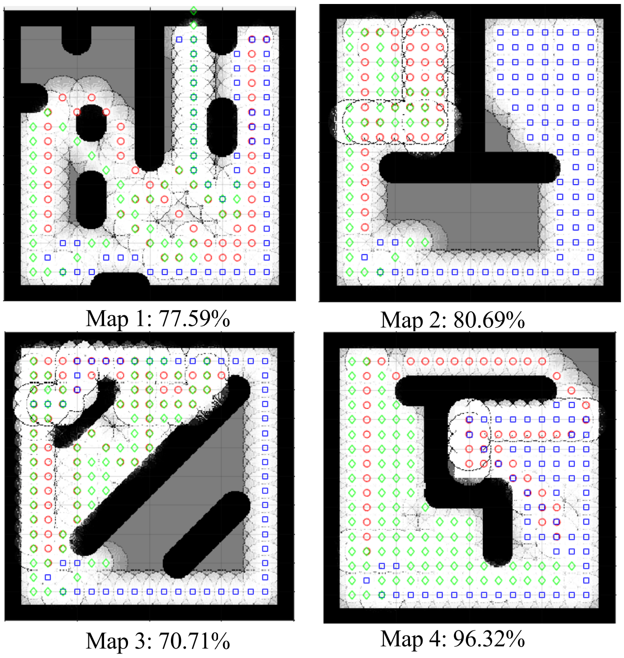

FIGURE 8. **The exploration results of the CME simulations with 100 iterations. Map 1 was interrupted on the 57th iteration because one robot was baffled by another robot on the upper right corner. On map 2, two of the robots could not get out from the left partition (bottleneck space), which limited the search of the unexplored space. Map 3 was explored with 70.71% coverage. The coverage is almost complete on map 4.**

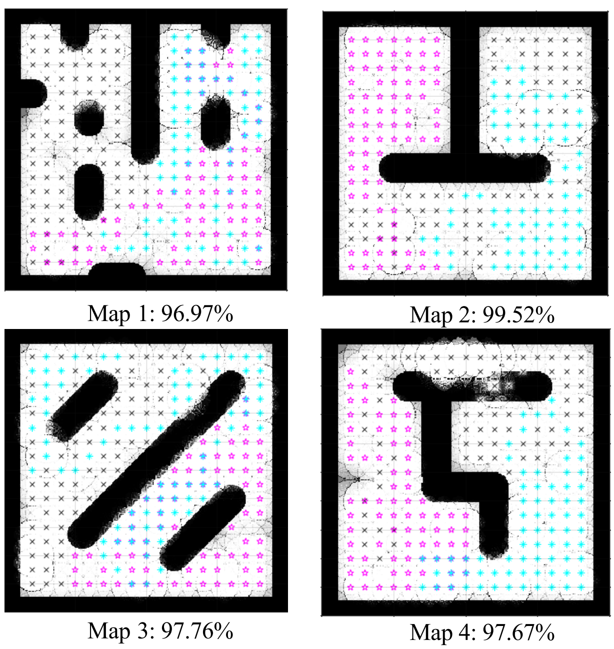

FIGURE 9. **The results of the hybrid stochastic exploration with 100 iterations. All maps have efficient exploration results, although the robot tried to go through the horizontal obstacle on the upper portion of map 4 to find a free path. It is noted that a narrow corridor may be a point of improvement for the method.**

surrounded with the same utility values equal to zero and/or obstacles of neighbor cells. In Fig. 8, map 1 showed that a similar circumstance also happens for CME. The difference between the two approaches, however, lies on the capability of the hybrid stochastic approach to search for a new solution with random values for the A and C parameters, which is not possible for CME or other deterministic approaches.

TABLE 2. **The number of aborted simulation runs of the hybrid stochastic exploration on four complex maps.**

Map 1	21
Map 2	0
Map 3	4
Map 4	14

The number of the simulation runs for the four maps in Fig. 9 is presented in Table 2. In order to obtain ten successful runs of map coverage passing 100 iterations, the total number of failed simulation runs are 21, 4, and 14 for maps 1, 3, and 4 respectively. The hybrid approach gave its best performance in map 2, where all 10 runs were successful (0 failed runs) and the map coverage is the highest at 99.52%.

VII. CONCLUSION

This paper proposed a hybrid approach for multi-robot exploration for unexplored space based on the CME, which is a deterministic approach, and GWO, which is a stochastic optimization approach. The hybrid approach allows searching for the impactful performance of the hybrid coordinated exploration by tuning random parameters. The technique performs comparatively well with CME on simple or ordinary maps but outperforms CME on complex maps with consideration of the number of iterations and simulations. For CME, it is worth emphasizing that if the method is not efficient under certain map conditions, there are no ways to find the optimal solution other than to change the map conditions, which is not always possible. In search and rescue operations where a human cannot change the obstacle locations, space size, or initial robot positions, the proposed hybrid stochastic approach can overcome the limitations and drawbacks of conventional or traditional exploration.

REFERENCES

[1] B. Yamauchi, "A frontier-based approach for autonomous exploration," in *Proc. IEEE Int. Symp. Comput. Intell. Robot. Automat.*, Jul. 1997, pp. 146–151.

[2] W. Burgard, M. Moors, C. Stachniss, and F. E. Schneider, "Coordinated multi-robot exploration," *IEEE Trans. Robot.*, vol. 21, no. 3, pp. 376–386, Jun. 2005.

[3] S. Mirjalili, S. M. Mirjalili, and A. Lewis, "Grey wolf optimizer," *Adv. Eng. Softw.*, vol. 69, pp. 46–61, Mar. 2014.

[4] C.-Y. Liu, X.-H. Yan, C. Liu, and H. Wu, "The wolf colony algorithm and its application," *Chin. J. Electron.*, vol. 20, no. 2, pp. 212–216, 2011.

[5] H.-S. Wu and F.-M. Zhang, "Wolf pack algorithm for unconstrained global optimization," *Math. Problems Eng.*, vol. 2014, Mar. 2014, Art. no. 465082.

[6] S. Zhang and Y. Zhou, "Grey wolf optimizer based on Powell local optimization method for clustering analysis," *Discrete Dyn. Nature Soc.*, vol. 2015, Oct. 2015, Art. no. 481360.

[7] H. Liu, G. Hua, H. Yin, and Y. Xu, "An intelligent grey wolf optimizer algorithm for distributed compressed sensing," *Comput. Intell. Neurosci.*, vol. 2018, Jan. 2018, Art. no. 1723191.

[8] D. Sánchez, P. Melin, and O. Castillo, "A grey wolf optimizer for modular granular neural networks for human recognition," *Comput. Intell. Neurosci.*, vol. 2017, Aug. 2017, Art. no. 4180510.
[9] A. Saxena and S. Shekhawat, "Ambient air quality classification by grey wolf optimizer based support vector machine," *J. Environ. Public Health*, vol. 2017, Aug. 2017, Art. no. 3131083.
[10] D. Gupta, C. Anand, and T. Dewan, "Enhanced heuristic approach for traveling tournament problem based on grey wolf optimizer," in *Proc. 8th Int. Conf. Contemp. Comput. (IC3)*, Aug. 2015, pp. 235–240.
[11] L. Li, L. Sun, J. Guo, J. Qi, B. Xu, and S. Li, "Modified discrete grey wolf optimizer algorithm for multilevel image thresholding," *Comput. Intell. Neurosci.*, vol. 2017, Jan. 2017, Art. no. 3295769.
[12] S.-X. Li and J.-S. Wang, "Dynamic modeling of steam condenser and design of PI controller based on grey wolf optimizer," *Math. Problems Eng.*, vol. 2015, Nov. 2015, Art. no. 120975.
[13] A. A. El-Fergany and H. M. Hasanien, "Single and multi-objective optimal power flow using grey wolf optimizer and differential evolution algorithms," *Electr. Power Compon. Syst.*, vol. 43, no. 13, pp. 1548–1559, 2015.
[14] N. Mittal, U. Singh, and B. S. Sohi, "Modified grey wolf optimizer for global engineering optimization," *Appl. Comput. Intell. Soft Comput.*, vol. 2016, Apr. 2016, Art. no. 7950348.
[15] S. Mirjalili, S. Saremi, S. M. Mirjalili, and L. D. S. Coelho, "Multi-objective grey wolf optimizer: A novel algorithm for multi-criterion optimization," *Expert Syst. Appl.*, vol. 47, pp. 106–119, Apr. 2016.
[16] N. Singh and S. B. Singh, "Hybrid algorithm of particle swarm optimization and Grey Wolf optimizer for improving convergence performance," *J. Appl. Math.*, vol. 2017, Nov. 2017, Art. no. 2030489.
[17] C. Muro, R. Escobedo, L. Spector, and R. P. Coppinger, "Wolf-pack *(Canis lupus)* hunting strategies emerge from simple rules in computational simulations," *Behav. Process.* vol. 88, no. 3, pp. 192–197, Nov. 2011.
[18] L. Rodríguez *et al.*, "A fuzzy hierarchical operator in the grey wolf optimizer algorithm," *Appl. Soft Comput.*, vol. 57, pp. 315–328, Aug. 2017.
[19] L. Freda and G. Oriolo, "Frontier-based probabilistic strategies for sensor-based exploration," in *Proc. IEEE Int. Conf. Robot. Automat. (ICRA)*, Apr. 2005, pp. 3881–3887.
[20] X. Li, H. Qiu, S. Jia, and Y. Gong, "Dynamic algorithm for safe and reachable frontier point generation for robot exploration," in *Proc. IEEE Int. Conf. Mechatron. Automat. (ICMA)*, Aug. 2016, pp. 2088–2093.
[21] W. Burgard, M. Moors, D. Fox, R. Simmons, and S. Thrun, "Collaborative multi-robot exploration," in *Proc. IEEE Int. Conf. Robot. Automat. Symposia (ICRA)*, vol. 1, Apr. 2000, pp. 476–481.
[22] A. Benkrid and N. Achour, "A novel approach for coordinated multi-robot exploration," in *Proc. 6th Int. Conf. Syst. Control (ICSC)*, May 2017, pp. 509–513.
[23] M. Rappaport and C. Bettstetter, "Coordinated recharging of mobile robots during exploration," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Sep. 2017, pp. 6809–6816.
[24] P. G. C. N. Senarathne and D. Wang, "A two-level approach for multi-robot coordinated exploration of unstructured environments," in *Proc. ACM 27th Annu. Symp. Appl. Comput.*, Mar. 2012, pp. 274–279.
[25] J. Kim, S. Bonadies, A. Lee, and S. A. Gadsden, "A cooperative exploration strategy with efficient backtracking for mobile robots," in *Proc. IEEE Int. Symp. Robot. Intell. Sensors (IRIS)*, Oct. 2017, pp. 104–110.
[26] D. Puig, M. A. García, and L. Wu, "A new global optimization strategy for coordinated multi-robot exploration: Development and comparative evaluation," *Robot. Auton. Syst.*, vol. 59, no. 9, pp. 635–653, 2011.
[27] D. Xu, X. Zhang, Z. Zhu, C. Chen, and P. Yang, "Behavior-based formation control of swarm robots," *Math. Problems Eng.*, vol. 2014, Jun. 2014, Art. no. 205759.
[28] R. Oikawa, M. Takimoto, and Y. Kambayashi, "Distributed formation control for swarm robots using mobile agents," in *Proc. 10th Jubilee Int. Symp. Appl. Comput. Intell. Inform. (SACI)*, May 2015, pp. 111–116.
[29] G. Fang, G. Dissanayake, and H. Lau, "A behaviour-based optimisation strategy for multi-robot exploration," in *Proc. IEEE Conf. Robot., Automat. Mechatron.*, vol. 2, Dec. 2004, pp. 875–879.
[30] Y. Wang, A. Liang, and H. Guan, "Frontier-based multi-robot map exploration using particle swarm optimization," in *Proc. IEEE Symp. Swarm Intell.*, Apr. 2011, pp. 1–6.
[31] D. Wang, H. Wang, and L. Liu, "Unknown environment exploration of multi-robot system with the FORDPSO," *Swarm Evol. Comput.*, vol. 26, pp. 157–174, Feb. 2016.
[32] S. Sharma, A. Shukla, and R. Tiwari, "Multi robot area exploration using nature inspired algorithm," *Biologically Inspired Cogn. Archit.*, vol. 18, pp. 80–94, Oct. 2016.
[33] K.-L. Du and M. N. S. Swamy, *Search and Optimization by Metaheuristics: Techniques and Algorithms Inspired by Nature*. Cambridge, MA, USA: Birkhaüser, 2016.
[34] A. J. Smith and G. A. Hollinger, "Distributed inference-based multi-robot exploration," *Auton. Robots*, vol. 42, no. 8, pp. 1651–1668, 2018.
[35] V. Gupta, T. H. Chung, B. Hassibi, and R. M. Murray, "On a stochastic sensor selection algorithm with applications in sensor scheduling and sensor coverage," *Automatica*, vol. 42, no. 2, pp. 251–260, Feb. 2006.
[36] E. Kaffashi, M. T. Shoorabi, and S. H. Bojnourdi, "Coverage optimization in wireless sensor networks," in *Proc. 4th Int. Conf. Comput. Knowl. Eng. (ICCKE)*, Oct. 2014, pp. 322–327.

KAMALOVA ALBINA received the B.S. degree in computer engineering from the Tashkent University of Information Technologies, Uzbekistan, in 2014. She is currently pursuing the M.S. and Ph.D. degree with Yeungnam University, Daegu, South Korea. Her research interests include mobile motion control, its approaches, and application solutions.

SUK GYU LEE received the B.S. and M.S. degrees in electrical engineering from Seoul National University, Seoul, South Korea, in 1979 and 1981, respectively, and the Ph.D. degree in electrical engineering from UCLA, USA, in 1990. Since 1982, he has been with the Department of Electrical Engineering, Yeungnam University, Daegu, South Korea, where he is currently a Professor. His research interests include robotics, SLAM, nonlinear control, and adaptive control.