

Cooperative exploration based on supervisory control of multi-robot systems

Xuefeng Dai¹ · Laihao Jiang² · Yan Zhao³

Published online: 14 January 2016
© Springer Science+Business Media New York 2016

Abstract When multiple mobile robots cooperatively explore an unknown environment, the advantages of robustness and redundancy are guaranteed. However, available traditional economy approaches for coordination of multi-robot systems (MRS) exploration lack efficient target selection strategy under a few of situations and rely on a perfect communication. In order to overcome the shortages and endow each robot autonomy, a novel coordinated algorithm based on supervisory control of discrete event systems and a variation of the market approach is proposed in this paper. Two kinds of utility and the corresponding calculation schemes which take into account of cooperation between robots and covering the environment in a minimal time, are defined. Different moving target of each robot is determined by maximizing the corresponding utility at the lower level of the proposed hierarchical coordinated architecture. Selection of a moving target assignment strategy, dealing with communication failure, and collision avoidance are modeled as behaviors of each robot at the upper level. The proposed approach distinctly speeds up exploration process and reduces the communication requirement. The validity of our algorithm is verified by computer simulations.

Keywords Automaton · Coordinated algorithms · Future utility · Immediate utility · Multi-robot systems · Supervisory control

1 Introduction

There are many applications for mobile robots in areas such as surveillance, exploration and rescue, mowing, and cleaning. For the exploration and rescue application, an important goal is covering the whole environment to find possible victims in a minimal time. Compared with a single-robot systems, multi-robot systems (MRS) not only speed up the exploration process but also have advantages of robustness and redundancy. In addition, MRS can finish tasks that the single robot systems cannot do. In order to utilize the advantages of MRS, it is an essential to have a coordinated algorithm which realizes moving target selection for each robot whilst considering the environment status and the moving targets of other robots. During the past two decades, a great deal of results for coordinating motion of MRS was developed.

A primary method for MRS coordination is market economy based approach. The essence of the approach is that all robots in the team (teammate) trade tasks and resources with each other to maximize individual profits. In the results of early research, robots bid for tasks based on distances, or times [1–3]. A distributed bidding algorithm (DB-A) was proposed to coordinate the MRS exploring unknown environments [2]. One contribution of the approach is that a nearness measure was introduced in bidding to maintain all robots close to each other to overcome the shortage of communication range is limited. On the other hand, a decision-theoretic approach dispersed robots over unknown environment through a utility calculation scheme that a robot

✉ Xuefeng Dai
daiixuefeng203@163.com

¹ Centre of Networks and Information, Qiqihar University, Heilongjiang, 161006, China

² The School of Computer and Control Engineering, Qiqihar University, Heilongjiang, 161006, China

³ The School of Communication and Electronic Engineering, Qiqihar University, Heilongjiang, 161006, China

utility reduction was inversely related to distances to targets of other robots [27]. For a known environment, task allocation was realized by repeating the auction algorithm when a robot finished its task [3]. More recently, an energy consumption based bidding scheme was proposed [4]. In the proposed approach, a hierarchical coordinated architecture for MRS exploring unknown environment is proposed where the lower level is developed by maximizing two kind of new utility, respectively. A few of improvements for auction based approach were made [5–8]. For example, an adaptive bidding based on back propagation (BP) neural network in a foraging scenario was proposed [5]. Each robot adopted a neural network to learn environment structure regarding energy consumed, speed and damage to the robot. However, the result needed to be verified on actual robots. The other kind of improvement is to deal with heterogeneity in MRS. The method developed in [6] is that each robot offers three bids from the aspects of the shortest path length, moving speed and battery life, respectively. A BP neural network based approach was proposed to decide which robot was the winner. The method in [7] used vectors to express capabilities of robot and task, respectively. The auctioneer broadcasts a task and its capability vector, and each robot bids for the task based on its capability. The task allocation was realized within the framework of Contract Net Protocol. Finally, the authors of this paper proposed an index for describing exploration performance of each robot for a heterogeneous MRS [8]. However, the approaches mentioned above depend on a perfect communication among all teammate robots. Different to the existing results, the proposed algorithm improves the market economy based approach on the purpose of dealing with both target selection and heterogeneity.

Nowadays, both auction algorithm and frontier-based exploration strategies were still used to develop other coordination algorithms. For example, a multi-robot partially observable Markov decision process (POMDP) was emulated by combining individual behaviors that was represented by single-robot POMDPs, and cooperation between robots was realized by using POMDP behavior auctions [9]. Moreover, the authors of [10] proposed three coordination algorithms which are named as *Reserve*, *Divided and Conquer*, and *Buddy System*. All of them selected navigation goals utilizing frontier-based approach. The differences of the three algorithms lie in the levels of trading off availability and non-interference. The proposed approach selects navigation goals from frontier cells too.

Artificial potential field method is an efficient approach for mobile robot path planning [11]. In the approach, the motion direction of a robot is determined by the sum of attractive and repulsive force vectors which were result from goal and obstacle locations, respectively. Recently, the

approach was combined with fuzzy systems and searching algorithm to realize multiple robots path planning [12], and coordination [13, 14]. The work presented in [12] used a global path planner which selects paths using a potential field, and a local planner which modifies the global path using a fuzzy system. Similarly, the repulsive potential energy which is computed according to both distances between the robots and the potential collision points was used as cost map to estimate the collision risk [13]. Subsequently, the optimized coordination solution was found through searching the roadmap. On the purpose of hazardous spill detection and perimeter surveillance, a kind of hierarchical control architecture for multiple autonomous robotic agents was proposed [14]. The upper level model is a state transition graph which consists of three discrete states, and the lower level model is a unicycle one. Continuous control laws for each discrete state are designed based on the artificial potential field method. Furthermore, stability of the system is analyzed. In this paper, the idea of artificial potential field approach will be used to define a new kind of utility which is used to disperse robots over the environment.

Unlike the work in [14], a three-level optimization based hierarchical coordinated algorithm for MRS exploring an unknown environment was proposed in [15]. In order to overcome the greedy behavior the unknown workspace is partitioned into as many regions as there are robots by *K*-mean clustering at the uppermost level. The objective of this level is to reduce the variance of waiting time on those areas and to fulfill balanced and sustained exploration for each teammate of the MRS. The middle level realizes an on-line assignment of robots to targets based on linear programming, and the lowest level realizes path planning. In the proposed hierarchical coordinated architecture, the lower level is a variation of market economy algorithm, and the upper level is a supervisor which is described in the following.

The concept of discrete event system (DES) appeared in 1980s. Supervisory control theory (RW theory) of DES can effectively solve the control and synthesis problem of systems which cannot be modeled by differential/difference equations [16]. It is adept at dealing with the problem from a logical point of view.

There are two mathematical approaches for modeling and control synthesis of discrete event systems. The first is formal language and automata theory which was used in leader follower formation control for a team of nonholonomic mobile robots [17] and unmanned aerial vehicles [18], respectively, and the coordination of MRS [19]. In [17], a formation problem that the followers keep a predetermined geometric formation with the leader in an obstacle populated environment was discussed. Behaviors of leader robot and follower robots were modeled

by automata, respectively, for obstacle avoidance, navigation, and formation control. A supervisor accomplished the coordination between the continuous controllers. To have a fully trustable algorithm, a hybrid supervisory control architecture for formation control of aerial vehicles was proposed [18]. By applying a polar partitioning based state abstraction and utilizing the properties of multi-affine vector fields over partitioned space, the infinite space was reduced into finite states of an automaton model. Afterwards, two supervisors were designed; one for reaching and keeping the formation, and the other for collision avoidance. A common feature of the two approaches is that the final supervisor was achieved through the parallel composition of the individual supervisors. A shortcoming of parallel composition is state explosion of the product system which makes the final supervisor too complex to deal with. Another partitioning was implemented on trajectories of each robot in [19] where a series of trajectory sectors constituted members of the state set of an automaton model. Then, the collision and deadlock avoidances behaviors of MRS were discussed, respectively, by supervising the trajectory sectors. The results of temporal discrete event systems and fuzzy discrete event systems have been applied in coordination of MRS too [20, 21]. In present paper, formal language and automata are used to endow each robot with autonomy. Inspired by the state abstraction idea [18], a sectorzone definition is established to reduce the size of automata state set.

The other mathematical approach is Petri net which is good at modeling and analyzing concurrent behaviors of discrete event systems. In a disaster response scenario, colored Petri nets were used to realize the complex and concurrent conversations between agents [22]. According to whether installed task detection device, all robots are categorized into initiators or helpers, respectively. The later assisted the former to finish task when it matched the task requirements. The robots can change their teammate selection strategies to adapt to dynamic environment. The authors of [23] presented a language which allows for intuitive and effective behavior design of multi-robot based on Petri net. In [24], stochastic Petri nets were used to model coordinated behavior and search for optimal waiting time of bidding in an offline way.

The proposed approach will discuss how to speed up MRS exploration process in unknown environments. The remainder of this paper is organized as follows. After reviewing the related work, Section 2 presents the problem description of our coordination exploration and our motivation. Section 3 gives a few of definitions and a DES model. Section 4 is utility calculation schemes and upper supervisory control strategy. Section 5 is simulation results. And the paper is concluded by Section 6.

2 Problem description and motivation

The market economy approach is suitable for heterogeneous MRS of moderate size [1–8]. For a variety of market based approaches for unknown environment exploration, each robot bid for frontier cells based on the evaluations of utility and costs of the corresponding frontier cells. The frontier cell with maximal evaluation is assigned to the current robot as its target. However, there are a few cases where the targets chosen from frontier cells are not optimal in the sense of exploration efficiency.

First, a case where the target is selected from frontier cells detected by R_i is shown in Fig. 1 where there are frontier cells (shaded ones), free cells (white ones), and unknown cells (mesh ones). It is assumed that R_i has moved upwards, and it detected frontier cells f_1 to f_{11} at current position. In case of selecting f_6 as its moving target, it will move to f_6 consecutively before take part in next round of auctions. It is obvious that R_i is not necessary to reach at f_6 because it can detect the obstacle by moving upwards only one cell. If R_i turns moving direction to left or right after detected the obstacle, it will detect more frontier cells. So, the first goal of the proposed approach is making current exploration of each robot more efficient by supervisory control. To realize the above goal, the exploration process of each robot is taken as a discrete event system and a new supervisory control based coordinated algorithm is proposed that monitors and regulates the motion of R_i at each discrete time instant.

The other case is shown in Fig. 2 in Section 4.

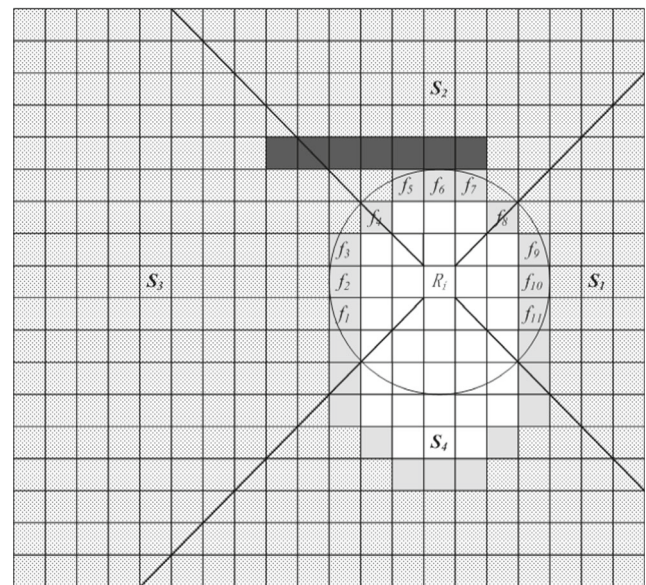


Fig. 1 Demonstration of frontier cells and partitioning of sector zones

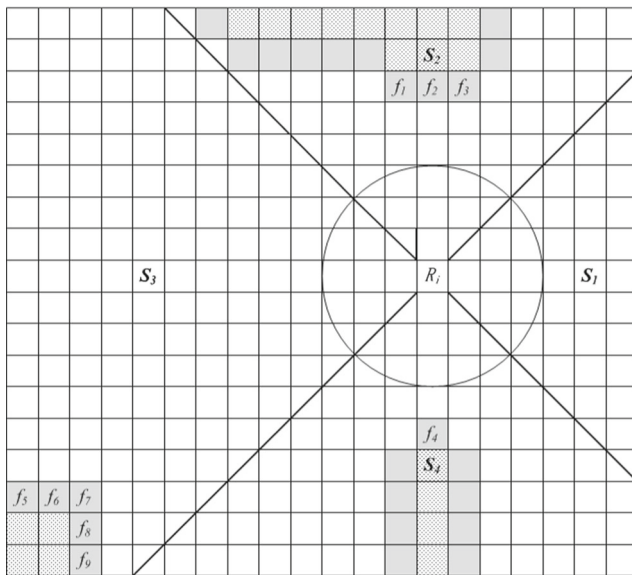


Fig. 2 Demonstration of NGFC_k and target selection based on future utility

Besides the above goal, the second goal is giving each robot more autonomous power, such as selecting a different target assignment strategy according to teammates and environment status, dealing with communication failure, collision avoidance ability, etc. The target assignment strategy behavior also includes forbidding undesired task switching action, i.e., R_i bids for and gets a frontier cell detected by R_j ($i \neq j$) as its moving destination. The switching phenomenon occurs often in MRS coordinated by traditional auction-based approaches. In fact, it may result in inefficient motion. For simplicity, the case shown in Fig. 1 that R_i can detect frontier cells from f_1 to f_{11} , is named as a complete detecting case. In other words, R_i can detect a series of congruent frontier cells at each scan in a complete detecting case. There is no case that R_i can detect more frontier cells than that in a complete detecting case. Obviously, it is not necessary for R_i to bid for the frontier cells detected by other teammate robots when it is in a complete detecting case. If there exists the other robot in a zone that R_i will explore, then R_i will be reassigned a new target. Else, R_i needs to continue its complete detecting case for succeeding exploration. The supervisor which is developed here will maintain R_i in complete detecting case and realize target selecting strategy switching. Correspondingly, the exploration is divided into two periods where each robot pursues two kinds of utility, respectively, as will be described in Section 3.

In order to realize above autonomous functions, each robot is modeled by an automaton at the upper level of the proposed two-layer coordination architecture. Autonomous abilities are defined as strings of actions included in desired behaviors. In order to model the state of each robot by a

finite state machine on a large environment, an environment partitioning scheme based on polar coordinate [18] is used. The exploration process of each robot is evaluated and controlled by an upper automaton at a fixed sample period, which is similar to the corresponding concept of sampled-data control systems. However, the sampled data are taken from robots and the environment.

In supervisory control theory, the modeling of current systems suffers from the problem of state explosion when constructing a product system. The third goal of the proposed approach is modeling the robot behavior with less state set members, and dealing with cooperation between teammate robots by a simple and efficient way. Also, each supervised robot in MRS has the ability of autonomous exploration in case of communication failure.

In traditional market economy approach, an auction action includes task announcement, bidding for tasks, notifying of winners, etc. All of these depend on a perfect communication. Moreover, the communication traffic concentrates on each auction period. It is shown by the analysis of Fig. 1 that some of the above communication traffic is not necessary when R_i is in a complete detecting case. So, the fourth goal of the proposed approach is spreading communication traffic over all sample rounds and saving waiting time in each auction period.

For a heterogeneous MRS conducting exploration, the detection ability (detecting range) of each robot is one of the factors which determines the exploration efficiency of the corresponding robot, even the exploration efficiency of MRS. So, the final goal of the proposed coordinated algorithm is taking the ability into account in the lower level task assignment strategy.

3 Preliminary

An MRS which consists of total n heterogeneous mobile robots is considered in this paper. These robots are named as R_1, \dots, R_n , respectively. The MRS carries out an exploration and mapping task in an unknown environment which is sparsely occupied by obstacles. It is assumed that each robot runs a simultaneous localization and mapping (SLAM) algorithm [28, 29] which produces a local occupancy grid map. All the robots have the same geometrical sizes which occupy a grid cell on the map, and they have the ability of communicating the environment information with no delay. It is also assumed that each robot has no knowledge about the environment except for the relative distances with other robots. In the following, R_i ($i = 1, \dots, n$) is taken for instance to discuss the proposed coordinated algorithm.

The objective of MRS exploration is covering the whole environment in a minimal time. All exploration tasks

can be finished by any teammate robot, and there is no inter dependence between tasks. According to [26], the exploration problem discussed in this paper belongs to the ND [ST-SR-IA] class, i.e., no dependence and single task robot and single robot task with an instantaneous assignment.

A few of definitions are now provided for the convenience of discussing moving target selections of robots.

Definition 3.1 Immediate utility is defined as the information that a robot will gain after it has moved to one of its adjacent cells from its current location.

Definition 3.2 Future utility is defined as the information that a robot will gain after it has moved at least two steps from its current location.

At a moment, the fact that R_i gains whether immediate utility or future utility is dependent on the distance between the target frontier cell and its current location. Even in the case that R_i cannot gain immediate utility, it has chance to gain future utility as long as the environment is not completely covered. A teammate robot can only belong to one of the following two exploration periods.

Definition 3.3 If R_i can gain immediate utility in the course of exploration, then it is said that R_i belong to an **independent exploration period**.

Definition 3.4 If R_i pursues future utility in the course of exploration, then it is said that R_i belong to an **cooperative exploration period**.

In the next section, the above definitions are going to be used to develop coordinated algorithm for the exploration process. In addition, the exploration process of each robot is going to be monitored and adjusted by the supervisor developed hereafter at a fixed time interval (sample period) to realize collective and desired behavior. According to RW theory, the behavior model of R_i is described by an automaton

$$G_i = \{Q_i, \Sigma_i, \delta_i, q_0\} \quad (1)$$

where Q_i is a state set, $\forall q \in Q_i$ represents a status of R_i , for example, arriving at a new target, or having finished communication with other teammate robots, etc., Σ_i is an alphabet, whose elements are listed in Table 1, $\delta_i : \Sigma_i \times Q_i \rightarrow Q_i$ represents a partial transition function which realizes a state transition from q_i to q'_i under the occurrence of σ_i , i.e., $q'_i = \delta_i(\sigma_i, q_i)$, and q_0 is an initial state. The behavior of R_i is described by the language generated by G_i

$$L(G_i) = \{s \in \Sigma_i^* \mid \delta_i(s, q_{i0})!\} \quad (2)$$

$L(G_i)$ represents all possible behavior when R_i conducting exploration. In order to realize a desired behavior $K \subseteq L(G_i)$, the necessary and sufficient conditions for a supervisor to exist is that K is prefix closed and controllable, i.e., $\bar{K} \subseteq K$ and $\bar{K} \Sigma_u \cap L(G_i) \subseteq \bar{K}$ [16]. The desired behavior is designed in advance according to the autonomous ability that each robot is given.

Fortunately, different with the coordination in discrete event systems [25], the proposed coordination [ST-SR-IA] has no logical restriction on the event occurrence sequence of different robots. So, the product system for all robots is not needed. In addition, cooperation between robots is realized by calculating and communicating the future utility which is calculated by a scheme inspired by the traditional potential field method for path planning. Therefore, only the supervisor and desired behavior for R_i is discussed hereafter.

4 Coordinated algorithm based on supervisory control

At first, to represent the moving target and current location of R_i with fewer variables in the whole environment, the following definition of surrounding environment of each robot is provided.

Definition 4.1 A **sector zone** of R_i is a part of the environment lying between two line segments that start from the geometrical centre of R_i and extend to the environment border.

Figure 1 shows an example where the surrounding environment of R_i is partitioned into four sector zones which are named S_1, S_2, S_3 , and S_4 . The surrounding environment may be known or unknown. In general, the intersection angles of all sector zones are the same, and most of the sector zones have no regular geometry shapes. As shown in Fig. 1, only the shape of S_1 is an equilateral triangle. The sector zones of R_i is dynamic, i.e., they are defined relative to the current location of R_i .

It must be remarked that each robot has its partitioning of the sector zones.

R_i calculates its immediate utility or future utility at each sample instant. If at least one frontier in sector zone S_k ($k = 1, \dots, 4$) is detected at current position, then the immediate utility of S_k is the number of the frontier cells.

$$U_i^{imm} = \sum_{j=1}^{n_{fi}} P_j \left(\|p_j^f - p_i^f\| \leq d_i^r \right) \quad (3)$$

where $P_j(\cdot)$ is the probability that frontier cell $f_j(i = 1, \dots, n_{fi})$ will be covered when R_i arrived

Table 1 ALPHABET LISTS

Symbol	QUANTITY	Controllability
α_0	sense surroundings	controllable
α_1	frontier cells are detected	uncontrollable
α_2	no frontier cells are detected	uncontrollable
α_3	obstacles or robots are detected	uncontrollable
η_0	communicate with others	controllable
η_1	communication succeed	uncontrollable
η_2	communication failed	uncontrollable
β_1	select a target under α_1	controllable
β_2	select a target under α_2 and η_1	controllable
β_3	select a target under α_2 and η_2	controllable
β_4	avoidance collision	controllable
τ	state update	controllable
μ	give up communication	controllable
λ	move to a target	controllable

at $p_i^r, p_i^r = [x_i^r \ y_i^r]^T$ is the moving target of $R_i, p_j^f = [x_j^f \ y_j^f]^T$ is the location of frontier cell j, n_{fi} is the number of frontier cells which are located in the sector zone, and d_i^r is the sensor detecting range.

If there are no frontiers detected in S_k at the current location, then the corresponding immediate utility is zero. If all immediate utility are zeros, then R_i will pursue future utility which is determined by known locations of frontier cells, obstacles, and other robots within the sector zone. The frontier cells mentioned above are the ones detected by all robots in previous exploration periods and stored in a unfinished job list which is used to remember the unexplored environment information. The list is maintained and updated by the achievements of all teammate robots during the course of exploration.

The future utility of S_k for R_i can be represented as

$$U_{ik}^{fut} = U_{ik}^{att} - \xi U_{ik}^{rep} \tag{4}$$

where U_{ik}^{att} and U_{ik}^{rep} are attractive and repulsive parts, respectively, and ξ is a weighting scalar.

For convenience to describe the future utility calculation scheme, the following definition is given.

Definition 4.2 A nearest group of frontier cells of S_k (NGFC_k) is defined as a series of frontier cells that satisfy the following conditions, i) one frontier cell is adjacent to at least one of other frontier cells, ii) a frontier cell has the shortest distance to current location of R_i , and iii) all of these frontier cells are located in S_k .

The attractive part U_{ik}^{att} of S_k is generated by all members of NGFC_k, and can be represented as,

$$U_{ik}^{att} = \sum_{j=1}^{n_{fi}} \frac{1}{d_j} \tag{5}$$

where $d_j = \|p_j^f - p_i^r\|$ is a Euclidean distance, p_i^r and $p_j^f (j=1, \dots, n_{fi}, j \neq i)$ are the same as in (3), and n_{fi} is the size of NGFC_k.

The assigned target locations which are located in S_k have repulsive affects to R_i . So, U_{ik}^{rep} can be calculated by

$$U_{ik}^{rep} = U_{ik}^{rep} (p_i^r, t_{ij}^r | j = 1, \dots, n_{ri}, i_j \neq i) \tag{6}$$

where, $t_{ij}^r = [x_{ij}^t \ y_{ij}^t]^T$ is the target locations of R_{ij} , and n_{ri} is the numbers of the corresponding robots. $U_{ik}^{rep}(\cdot)$ takes the role of preventing the occurrence of inefficiency exploration behavior of R_i , i.e., preventing more robots moving to the same area. If the target of R_j is located in S_k , then the reduction of the amount of unknown cells that R_i can explore after entered S_k is proportionate to the detecting radius of $R_j (d_j^r)$, and is inversely proportional to the distance between current location of R_i and the target location of R_j . If there are n_{ri} targets located in S_k , then the repulsive part can be represented as

$$U_i^{rep} = \sum_{j=1}^{n_{ri}} \frac{d_j^r}{d_j^t} \tag{7}$$

where d_j^t is an Euclidean from current location of R_i to the target location of R_j .

In the cooperative exploration period, R_i will select a sector zone with the maximal future utility to explore according to

$$k_{\max} = \arg \max_{k=1, \dots, 4} \{U_1^{fut}, \dots, U_4^{fut}\} \tag{8}$$

After deciding the target sector zone, the next step is to decide the target coordinates. It is realized by

$$\left. \begin{aligned} x_i^t &= \left(\sum_{j=1}^{n_{fi}} x_j^f \right) / n_{fi} \\ y_i^t &= \left(\sum_{j=1}^{n_{fi}} y_j^f \right) / n_{fi} \end{aligned} \right\} \tag{9}$$

Let the shortest distance from a frontier cell in S_k to the location of R_i is

$$d_{\min} = \min_{j=1, \dots, n_{fi}} \{d_1, \dots, d_{n_{fi}}\} \tag{10}$$

If $d_{\min} < \|p_i^t - p_i^r\|$ and the frontier cell is not an end of NGFC_k, where $p_i^t = [x_i^t \ y_i^t]^T$, then the target cell location (9) is replaced by $p_{i'}^t = [x_{i'}^t \ y_{i'}^t]^T$, where

$$i' = \arg \min_{j=1, \dots, n_{fi}} \{d_1, \dots, d_{n_{fi}}\} \tag{11}$$

The coordinated exploration for each robot is realized by monitoring immediate utility or future utility at every sample instant. This results in that each robot’s exploration process is an interleaved arrangement of independent exploration periods and cooperation exploration periods.

Proposition 4.1 *The exploring mode for each teammate robot in a sample period is that it pursues immediate utility first, and pursues future utility finally.*

The aim of the lower level target assignment is to maximize current exploration efficiency. Obviously, if R_i being

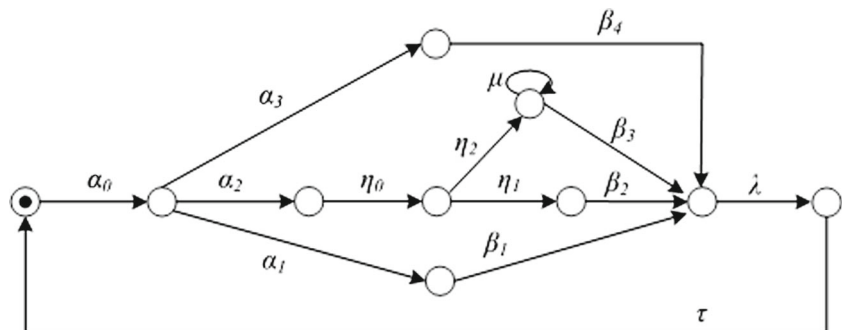
in the complete detecting case, then it belongs to independent exploration period, and has the most efficiency. The time complexity of the lower task assignment algorithm in the proposed approach is $O(n)$ and $O(m^3)$, where n is the numbers of frontier cells the involved robot detected at current period, and m are the size of the unfinished job list, when a robot pursues the immediate utility and the future utility, respectively.

In the proposed approach, R_i selects and moves to targets consecutively at each supervisory control sample period. The sample period is less than the auction period. An accessorial advantage of the proposed task assignment at lower level algorithm is communication volume being spread over all sample periods, not being concentrated on auction periods as in traditional market economy approach. In addition, waiting time in the traditional approach is saved.

The superiority of the above scheme for calculating utility is demonstrated by Fig. 2 where the color meanings of cells are the same as Fig. 1. In the figure, there are 25 frontier cells and 4 sector zones. In the market based approaches, if these frontier cells are considered as independent tasks, both of distanced based approach [2] or time based approach [3] will select f_4 as target. Although d_4^f is the shortest distance, f_4 is not the optimal target for R_i . It is obvious that R_i will explore 3 cells if it moves upwards for 2 cells and will explore only one cell if it moves downwards for one cell. According to Definition 4.2 and the sector zone partitioning in the figure, NGFC₁ = φ , NGFC₂ = { f_1, f_2, f_3 }, NGFC₃ = { f_5, \dots, f_9 }, and NGFC₄ = { f_4 }, respectively. The future utility of the four sector zones are calculated by (7), and the results are that $U_{i1}^{fut}, U_{i2}^{fut}, U_{i3}^{fut}$, and U_{i4}^{fut} are equal to 0, 0.49, 0.36, and 0.2, respectively. Then the proposed approach will select S_2 and f_2 as the target sector zone and the target cell based on (8) and (9), respectively.

Since the exploration mode and goal for each robot is identical, a universal and independent supervisor will be developed. In order to describe the behavior for R_i , the robot actions and action results are defined as events, respectively, as shown in Table 1.

Fig. 3 The desired behavior K of R_i



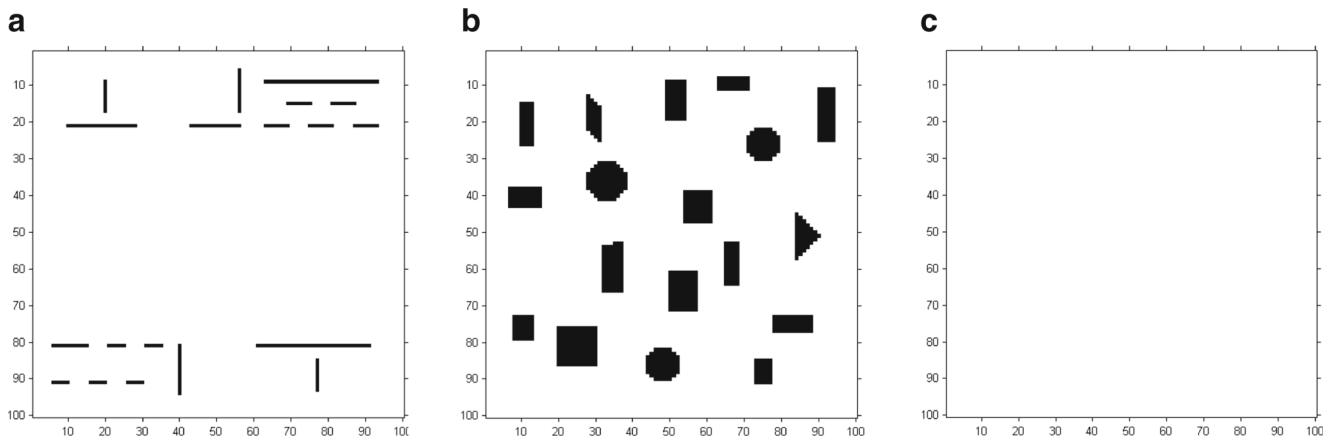


Fig. 4 Environments used for simulations: (a) line segments, (b) geometries, and (c) blank

The desired behavior K of R_i when realizing cooperative exploration is shown in Fig. 3. A supervisor S is constructed to make the closed-loop behavior

$$L(S/G_i) = K \tag{12}$$

In each supervisory control circle the actual behavior is a subset of the language described in Fig. 3.

The actions of R_i can be fired by the supervisor, and results of actions are uncontrollable. To deal with the occurrence of uncontrollable events, a few controllable events are introduced. For example, the detecting or sensing surrounding action (α_0) is controllable, i.e., its happening can be permitted or prohibited by the supervisor. However, the detecting results are uncontrollable, i.e., the results may be one of the following, frontier cells are detected (α_1), no frontier cells are detected (α_2), and obstacles are detected

(α_3). In other words, these result events cannot be prohibited from occurrence. In addition, for both communications to succeed (η_1) or communications to fail (η_2) between two robots rely on the environment and the communication infrastructure on two corresponding robots, not upon the supervisor. They are uncontrollable too.

The coordinated events are described as follows. τ is the state update event, i.e., after arrived at the target cell R_i updates the target cell as the new current state. Therefore, only one member of state set Q in (1) is needed to represent the location of R_i . β_1 selects a sector zone as moving target according to maximize immediate utility (3). Similarly, β_2 selects a sector zone which has the maximal future utility as target sector zone according to (8). β_3 is the behavior that R_i selects a sector zone based on limited available information. It must be remarked that β_1 and β_3 select their target

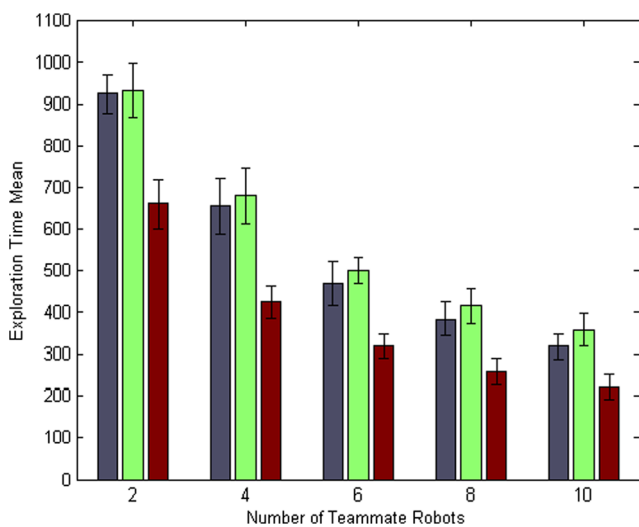


Fig. 5 Exploration time mean vs. robot team scale in the line segments environment. Grey, green and red bars represent the exploration time means of MRS under coordination of DB-A, RA-A and SC-A, respectively

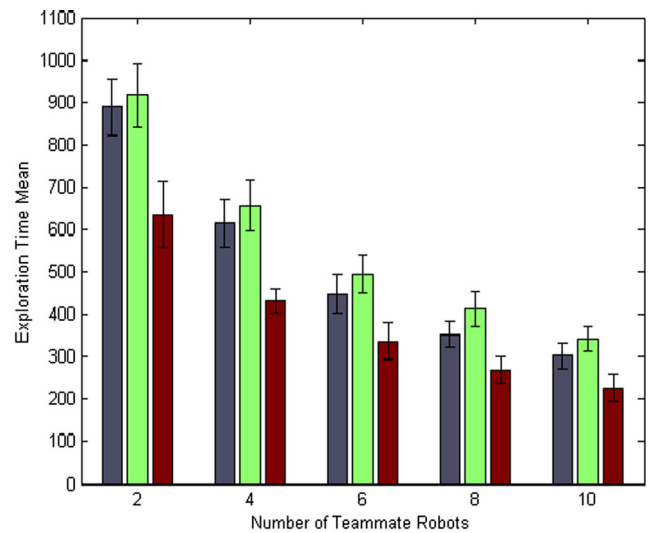


Fig. 6 Exploration time mean vs. robot team scale in the geometries environment. Grey, green and red bars represent the exploration time means of MRS under coordination of DB-A, RA-A and SC-A, respectively

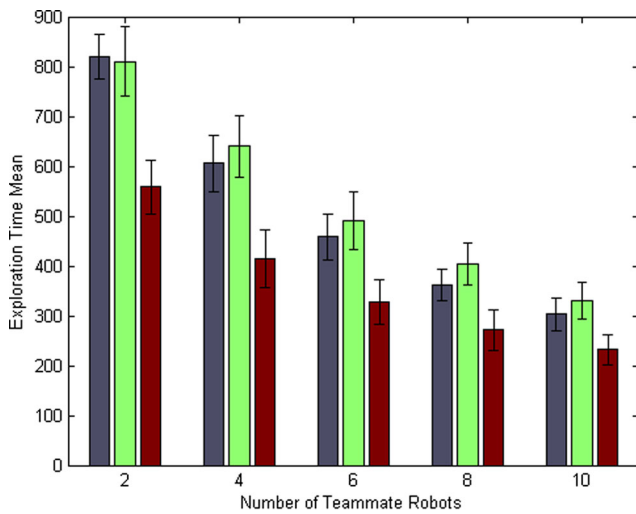


Fig. 7 Exploration time mean vs. robot team scale in the blank environment. Grey, green and red bars represent the exploration time means of MRS under coordination of DB-A, RA-A and SC-A, respectively

from adjacent cells, β_2 selects its target from the unfinished job list.

In the dependence exploration period, the behavior of R_i is $(\alpha_0\alpha_1\beta_1\lambda\tau)^*$. In the cooperative exploration period, the behavior consists of two strings: $(\alpha_0\alpha_2\eta_0\eta_1\beta_2\lambda\tau)^*$ and $(\alpha_0\alpha_2\eta_0\eta_2\beta_3\lambda\tau)^*$ for the cases of communication succeed and communication fail, respectively.

In general, the avoiding ability to obstacles and other robots cannot be guaranteed when R_i is going to the target generated by (9) or (11). The reason is that the repulsive part is fused with attractive part in future utility. In order to make R_i realizing reliable collision avoidance with other teammate robots or obstacles, a collision avoidance behavior $(\alpha_0\alpha_3\beta_4\lambda\tau)^*$ is included in the desired behavior. The avoidance behavior has the highest priority.

In the implementation, each robot has its own supervisor, so there is no central supervisor in the MRS. In other words, the supervisory control of MRS is realized in a distributed way. In addition, the synchronization of sampling

period for all teammate robots is not needed in both independent nor cooperative exploration periods. Each teammate robot decides its moving destination based on available information from the environment and all teammate robots.

5 Simulation results

The simulations are conducted on Matlab 7.1 on a personal computer with Intel[®] Core™i5-4460 CPU. The environments used for our simulations are shown in Fig. 4. In the figure, there are three different environments which are the same sizes. The first two environments are sparsely occupied by obstacles of different shapes, and the last is a blank one. The three environments are named as line segments, geometries, and blank, respectively. Each environment will be modeled as an occupancy grid of 100 cells multiply 100 cells after exploration. The objective of exploration is to cover the environment by a team of heterogeneous MRS in a minimal time.

It is assumed that all robots have the same speed of 1 distance unit per time unit. Also, all robots can move upward, downward, leftward, and rightward only. Finally, it is assumed that the sample period is 1 time unit.

The proposed supervisory control based algorithm (SC-A), DB-A and repeated auction algorithm (RA-A) are compared. In order to disperse all robots over each environment, the decision-theoretic utility [27] is adopted in DB-A, instead of the one with nearness measures. To apply RA-A, the explored part of each environment and frontier cells are taken as known world and tasks, respectively. Since the above tasks are dynamically changing, the RA-A applied hereafter has no commitment to unfinished tasks which were assigned to the corresponding robots. The pool of candidate robots is a set $R = \{R_1, \dots, R_{10}\}$. The teammate robots for each run are selected from R_1 to R_n in the set R , where n is a current team size for simulation, it is set as 2, 4, 6, 8, and 10, respectively. It is assumed that detecting ranges of R_2 and R_3 are 8 and 3, respectively, and those

Table 2 Auction times mean/exploration time mean of DB-A and RA-A

Algorithm	Number of Teammate Robots				
	2	4	6	8	10
DB-A	164/924	150/655	111/470	91/386	76/318
	152/889	142/615	104/449	82/354	71/302
	152/820	143/606	109/458	86/362	72/303
RA-A	145/931	135/679	105/500	89/415	77/359
	142/917	131/656	102/495	88/413	74/342
	134/810	129/640	103/491	87/404	72/330

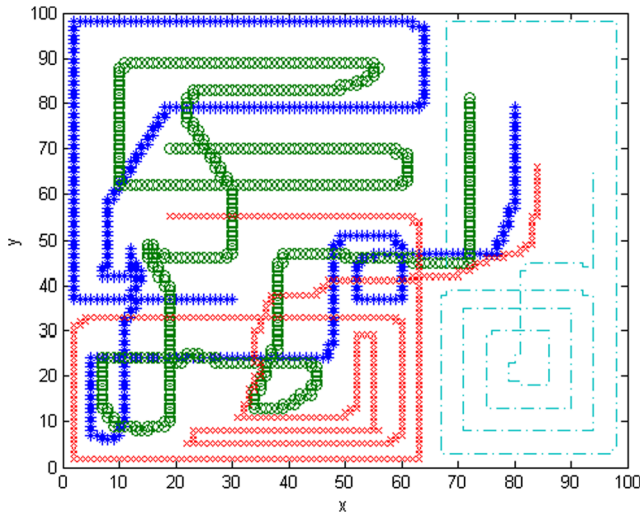


Fig. 8 Trajectories of four robots coordinated by DB-A in the blank environment. The correspondences are star - R_1 , circle - R_2 , x-mark - R_3 , and solid - R_4

of other robots are 4s. 30 runs for each algorithm per environment are conducted. The means and standard deviations of exploration times are shown in Figs. 5 to 7.

It is shown that SC-A improved exploration time efficiency compared with the other two algorithms for all cases of team sizes over the three environments. A common trend in the three figures is that the exploration time means descend when the number of teammate robots increase. The other trend is that exploration time means for the same team size in the three environments descend from Figs. 5 to 7. This is results from the fact that the obstacles in the first environment most limit the detecting ranges of each

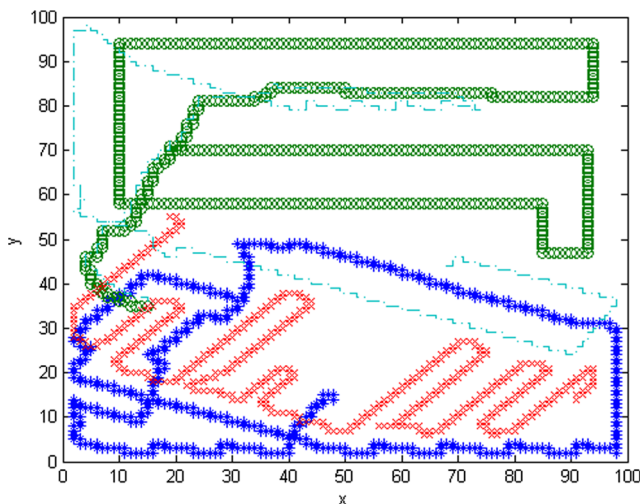


Fig. 9 Trajectories of four robots coordinated by RA-A in the blank environment. The correspondences are star - R_1 , circle - R_2 , x-mark - R_3 , and solid - R_4

robot. On the other hand, both detecting range and moving direction of each robot have no limitations in the third environment (Fig. 6).

In previous sections, Figs. 1 and 2 showed how our algorithm saves exploring time from the target selection point of view. Next, it will be demonstrated that the proposed algorithm improves exploration efficiency from the other two aspects. First, waiting time for each auction period during the exploration process is saved. In the presented simulation results, the waiting time in each auction round is set to be 1 time unit. In this time unit, all teammate robots do nothing except for waiting for better bids. Table 2 shows auction times means and exploration time means of MRS coordinated by DB-A and RA-A in the three environments, respectively. In the table, the three rows of data for each algorithm from up to down correspond to line segments, geometries and blank environments, respectively.

Finally, the exploration time efficiency improvement is demonstrated by comparing the trajectories of all teammate robots for the three algorithms. Due to limitations on paper space, only the trajectories of team size 4 for the three algorithms in the blank environment are shown in Figs. 8 to 10, respectively. The initial locations of the four robots for the three algorithms are identical which are randomly selected as (30, 37), (19, 70), (19, 55), and (68, 44), respectively. The initial orientations of the four robots are selected randomly which are upwards, leftwards, leftwards, and leftwards, respectively (Fig. 9).

It can be seen that the trajectories of robots coordinated by DB-A and RA-A are more abnormal and have more intersections than that of the corresponding robots coordinated by SC-A. The exploration times of the MRS coordinated by the three algorithms are 597, 684 and 427, respectively.

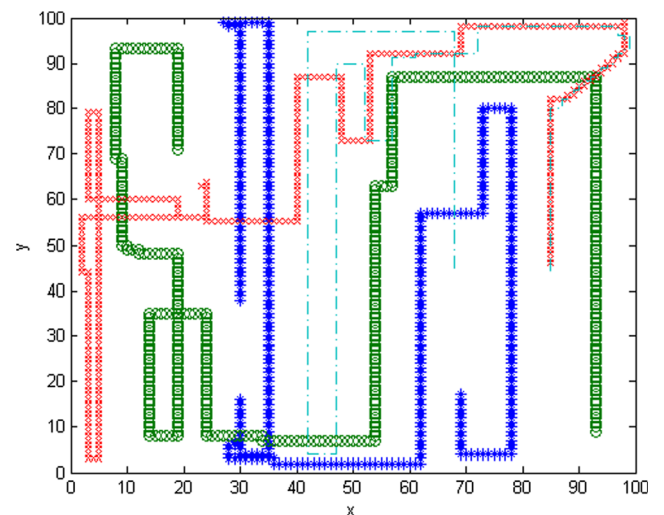


Fig. 10 Trajectories of four robots coordinated by SC-A in the blank environment. The correspondences are star - R_1 , circle - R_2 , x-mark - R_3 , and solid - R_4

6 Conclusion

We have proposed a hierarchical coordinated approach for MRS exploring unknown environments. The upper level emphasizes on endowing each robot autonomy based on supervisory control of discrete event systems, and the lower level emphasizes on navigation goal selection. By partitioning the environment into a few of sector zones, a simple upper automaton with a limited state set was established. The autonomous abilities are represented as symbol strings which are included in the desired behavior language. At the lower level, two kinds of utility which are used to realize efficient moving target selection and cooperation between robots are proposed. The proposed approach improved exploration efficiency compared with single robot systems. Moreover, it has the following advantages, selecting a target assignment strategy according to the MRS and the environment status, the ability of dealing with heterogeneity, less communication traffic, trajectory smoothness, etc. over the traditional market economy based approach. The validity of the proposed coordinated algorithm should be proved from a theoretic point of view. Moreover, the future work includes implementing the proposed algorithm on real robots.

Acknowledgments This work was supported in part by the Natural Science Fund of Heilongjiang Province, China under Grant F201331. The authors also gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved the presentation.

References

- Dias MB, Zlot R, Kalra N, Stentz A (2006) Market-based multi-robot coordination: a survey and analysis. *Proc. IEEE* 94(7):1257–1270
- Sheng W, Yang Q, Tan J, Xi N (2006) Distributed multi-robot coordination in area exploration. *Robot Auton Syst* 54(12):945–955
- Nanjanath M, M Gini (2010) Repeated auctions for robust task execution by a robot team. *Robot Auton Syst* 58(7):900–909
- Kaleci B, Parlaktuna O (2013) Performance analysis of bid calculation methods in multirobot market-based task allocation. *Turk. J Elec Eng & Comp Sci* 21(2):565–585
- Kensler JA, Agah A (2009) Neural networks-based adaptive bidding with the contract net protocol in multi-robot systems. *Appl Intell* 31(3):347–362
- Yuan Q, Guan Y, Hong B, Meng X (2013) Multi-robot task allocation using CNP combines with neural network. *Neural Comput Applic* 23(7–8):1909–1914
- Liu L, Ji X, Zheng Z (2006) Multi-robot task allocation based on market and capability classification. *Robot* 28(3):337–343. (in Chinese)
- Dai XF, Yao ZF, Zhao Y (2014) A discrete adaptive auction-based algorithm for task assignments of multi-robot systems. *J Robot Mech* 26(3):369–376
- Capitan J, Spaan MTJ, Merino L, Ollero A (2013) Decentralized multi-robot cooperation with auctioned POMDPs. *Int J Robot Res* 32(6):650–671
- Nieto-Granda Carlos, John G. Rogers III, Henrik IC (2014) Coordination strategies for multi-robot exploration and mapping. *Int J Robot Res* 33(4):519–533
- Xu D, Zou W (2008) Perception localization and control for indoor mobile service robots. China: Science Press, Beijing, pp 300–325. in Chinese
- Kim DW, Lasky TA, Velinsky SA (2013) Autonomous multi-mobile robot system: Simulation and implementation using fuzzy logic. *Int J Control Autom* 11(3):545–554
- Cui R, Gao B, Guo J (2012) Pareto-optimal coordination of multiple robots with safety guarantees. *Auton Robot* 32:189–205
- Zhang G, Fricke GK, Garg DP (2013) Spill detection and perimeter surveillance via distributed swarming agents. *IEEE/ASME Trans Mechatronics* 18(1):121–129
- Puig D, Garcia MA, Wu L (2011) A new global optimization strategy for coordinated multi-robot exploration: Development and comparative evaluation. *Robot Auton Syst* 59(9):635–653
- Ramadge PJ, Wonham WM (1989) The control of discrete event systems. *Proc. IEEE* 77(1):81–98
- Gamage GW, Mann GKI, Gosine RG (2009) Discrete event systems based formation control framework to coordinate multiple nonholonomic mobile robots. In: *Proceedings of IEEE/RSJ international conference on intelligent robotic system (IROS)*, pp 4831–4836
- Karimodini HL, Chen BM, Lee TH (2011) Hybrid formation control of the unmanned aerial vehicles. *Mechatronics* 21(5):886–898
- Roszkowska E (2007) DES-based coordination of space-sharing mobile robots. In: Moreno-Diaz R, et al. (eds) *EUROCAST 2007*, LNCS 4739, pp 1041–1048
- Koo TJ, Li R, Quottrup MM, Cliton CA, Izadi-Zamanabadi R, Bak T (2012) A framework for multi-robot motion planning from temporal logic specifications. *Sci China Inf Sci* 55(7):1675–1692
- Jayasiri GKIM, Gosine RG (2011) Tightly-coupled multi robot coordination using decentralized supervisory control of fuzzy discrete event systems. In: *Proceedings of 2011 IEEE international conference on robotics and automation*. IEEE, Shanghai, China, pp 3358–3363
- Ebadi T, Purvis M, Purvis M (2010) A framework for facilitating cooperation in multi-agent systems. *J Supercomput* 51(3):393–417
- Ziparo VA, Iocchi L, Lima PU, Nardi D, Palamara PF (2011) Petri net plans: A framework for collaboration and coordination in multi-robot systems. *Auton Agent Multi-Agent Syst* 23(3):344–383
- Sheng W, Yang Q (2005) Peer-to-peer multi-robot coordination algorithms: petri net based analysis and design. In: *Proceedings of 2005 IEEE/ASME international conference on advanced intelligent mechatronics*. Monterey, pp 1407–1412
- Lin F, Wonham WM (1990) Decentralized control and coordination of discrete-event systems with partial observation. *IEEE Trans Automatic Control* 35(12):1330–1337
- Korsah GA, Stentz A, Dias MB (2013) A comprehensive taxonomy for multi-robot task allocation. *Int J Robot Res* 32(12):1495–1512
- Burgard W, Moors M, Stachniss C, Schneider FE (2005) Coordinated multi-robot exploration. *IEEE Trans Robot* 21(3):376–386
- Durrant-Whyte H, Bailey T (2006) Simultaneous localization and mapping: Part I. *IEEE Robot Autom Mag* 13(2):99–110
- Durrant-Whyte H, Bailey T (2006) Simultaneous localization and mapping: Part I. *IEEE Robot Autom Mag* 13(3):108–117



Xuefeng Dai received the B.S. degree in electrical automation from Liaoning Technical University, Liaoning, China, in 1985, the M.S. degree in automatic control theory and applications from Harbin Engineering University, Heilongjiang, China, in 1992 and the Ph.D. degree in control theory and control engineering from Harbin Engineering University, Heilongjiang, China, in 2001.

From 1992 to 1997, he was a lecturer with the department of automation, Qiqihar University, Heilongjiang, China. From 2001 to 2004, he was a postdoctoral research associate with the Institute of Robotics, Harbin Institute of Technology, Heilongjiang, China. His research interest includes supervisory control of discrete event systems, neural networks, and coordinated control of multi-robot systems.



Laihao Jiang received the B.S. degree in agriculture electrification and automation from Northeast Agriculture University, Harbin, China, in 2009. He is currently pursuing the M. S. in control engineering, School of Computer and Control Engineering, Qiqihar University, Heilongjiang, China.

His research interest includes task allocation in multi-robot systems, and clustering algorithms and applications.

Yan Zhao received the B.S. degree in communication engineering from Northeast Dianli University, Jilin City, Jilin, China, in 1999, the M.S. degree in communication engineering from Qiqihar University, Qiqihar, Heilongjiang, China, in 2003. He has been a lecturer with School of Communication and Electronic Engineering, Qiqihar University, Heilongjiang, China, since 1999. His research interests include multi-robot systems and artificial intelligence. He is a member of China Institute of Communications.