IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received September 24, 2019, accepted October 7, 2019, date of publication October 15, 2019, date of current version October 30, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2947541

Cooperative Navigation Based on Cross Entropy: Dual Leaders

LICHUAN ZHANG, YICHEN LI, LU LIU, AND XIUYE TAO
School of Marine Science and Technology, Northwestern Polytechnical University, Xi'an 710072, China
Corresponding author: Lichuan Zhang (zlc@nwpu.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 51979229.

ABSTRACT Cooperative navigation aims at improving positioning accuracy of Autonomous Underwater Vehicles (AUVs). In this paper, a dual leaders cooperative navigation method is proposed based on Cross Entropy (CE) algorithm. Since the trajectories of the slave AUVs are assumed to be predetermined, the Markov Decision Process (MDP) is also integrated in the proposed algorithm to generate optimal trajectories of master AUVs from the perspective of probability. Firstly, the navigation model and cost functions are established for the cooperative navigation system with multiple masters and slaves. Then, the CE algorithm is used to train the system with help of MDP to obtain the path of the master AUVs. In the simulation, the cooperative localization trajectories of the slave AUVs are obtained by Extended Kalman Filter (EKF) and are compared with other positioning methods. The results show that the trajectories of dual master AUVs obtained by the proposed algorithm can not only reduce the observation error of the slave AUVs in the system effectively, but also keep relative measurement distance between the master AUVs and the slave AUVs in a suitable range.

INDEX TERMS Cooperative navigation, cross entropy, Markov decision process, path planning.

I. INTRODUCTION

Autonomous Underwater Vehicles (AUVs) can expand scope of human marine activities. With the improvement of efficiency requirements for underwater operations, the usage of networked Autonomous Underwater Vehicle (AUV) for collaborative work has recently become a hot issue in the field of marine engineering. For example, when searching for the missing Malaysia Airlines Flight MH370, the Ocean Infinity Company used eight AUVs simultaneously, and was able to search more than 460 square miles of seabed per day.

During the operation, AUV needs to locate itself for navigation. Due to the barrier of water to electromagnetic waves and the complexity of underwater environment, the positioning and navigation methods that AUV can use are very limited [1]. Commonly used positioning methods for a single AUV are seabed terrain matching, dead reckoning (DR), visual navigation, acoustic navigation and so on. In order to improve positioning accuracy of the whole system, positioning accuracy of each AUV in the network can be improved separately. However, this will lead to a multiplication of the cost of navigation equipment, which is too expensive to achieve in most cases. Collaborative navigation can balance positioning accuracy and cost [2], [3].

For collaborative navigation, the positioning information is shared by each AUV in a multi-AUV system. One of the most commonly used structures is the master-slave structure, in which several master AUVs carry high precision navigation equipments to provide location service for the slave AUVs whose navigation equipments are less accurate. They communicate with each other using underwater acoustic devices.

Formation configurations of multi-AUVs cooperative navigation are mainly divided into single-leader formation and multi-leader formation [1], [4]. Mythrehee studied the optimal energy-saving distribution in the underwater mobile sensor network based on the single-leader and multi-follower formation through the Steinberger game theory [5]. Gao et al. proposed a new algorithm called Huber-based Iterated Divided Difference Filtering (HIDDF) and applied the algorithm to cooperative localization of AUVs supported by a single surface leader [6]. Huang et al. proposed a beacon network that does not depend on long baseline positioning system [7]. The cooperative navigation model of single leader AUV is used to verify the reliability of the algorithm. Yan solved the problem of longitude density in high latitudes by studing cooperative navigation in polar coordinates [8].

The associate editor coordinating the review of this manuscript and approving it for publication was Guangjie Han.

Wen proposed an Unscented Particle Filter (UPF) algorithm for cooperative navigation with two leaders, which can effectively reduce the impact of sensor noise [9]. Gao established a cooperative navigation algorithm for two leader AUVs, which solved the weakness of the observability of the system and reduced the requirement for the maneuverability of the leader AUVs [10]. Among literature [10], using multiple AUVs as leaders can not only improve the observability of the system, but also reduce the maneuverability requirements of the leaders.

In terms of error compensation, Xiao studied the characteristics of time delay in AUV cooperative navigation system in master-slave mode, and converted the time delay into measurement deviation from the observation equation of AUV platform [11]. Chen proposed a positioning error algorithm based on underwater acoustic propagation time compensation [12]. The author used the Extended Kalman Filter (EKF) to reconstruct the measurement equation through the position error state estimated by the inertial navigation system (INS). Therefore, the measurement equation and the system measurement became synchronized, and the error from the time delay was eliminated. Another Augmented Extended Kalman Filter (AEKF) algorithm was proposed by Qi to solve the positional failures due to time delay measurements [13].

Many researchers have proposed improved algorithms based on the principle of cooperative navigation. Li *et al.* proposed an algorithm based on Student's Extended Kalman Filter, which has strong robustness to abnormal points and measurement noise [14]. Chen studied the application of the Maximum Likelihood Method in cooperative navigation [15]. Huang proposed a new adaptive EKF algorithm [16]. Based on the online expectation maximization method, the prediction error covariance matrix and the measurement noise covariance matrix were adaptively estimated to solve the problem of unknown noise covariance matrix in cooperative localization. Allotta compared the performance of the EKF and the Unscented Kalman Filter (UKF) cooperative navigation algorithm and performed experiments on the Typhoon AUV [17]. The results showed that the UKF had better performance compared with EKF. Viegas studied the state estimation problem of vehicles with time-varying measurement topologies [18]. The proposed solution included implementing a local state observers on each vehicle and using the theory of switching system theory to study the effect of measurement topology changes on estimation error dynamics.

Generally, the operation path of the slave AUVs are planned before tasks. Therefore, it is the master AUVs' task to maneuver and improve the cooperative navigation positioning accuracy of the slave AUVs. In other words, the master AUVs need to plan optimal path to minimize the observation error of the slave AUVs. Since the Cross Entropy (CE) is suitable for measuring uncertain information [19], it is widely used for optimization problems, such as enterprise resource planning system selection [20] and information retrieval [21], etc. The property of CE can be used to design the path of the master AUVs in the cooperative navigation. In [22] and [23], the leader vehicles used zigzag path and rhombus path to improve the observability of the system. However, in [22], the author just made the leader AUV maintain a 45° zigzagging pattern, and did not consider the global optimization. Teck and Chitre [24] also had studied CE algorithm in cooperative navigation, but only one leader in the system is considered.

In this paper, a cooperative navigation system with dual master AUVs is proposed. A proper cost function to train the system based on CE algorithm is also established to reduce the observation error of the slave AUVs. In addition, EKF is applied to cooperative navigation system after trained by CE algorithm to obtain trajectories of slave AUVs. Morever, the proposed method is compared with other navigation methods to show its superiority.

The following sections are organized as follows: Section II gives a formulation of the model and problem. Section III is the proposed cooperative navigation method. In the Section IV, simulation results are analyzed. Then the conclusion is given, and some pending further research directions are proposed In section V.

II. PROBLEM FORMULATION

The core element of cooperative navigation is to allow the information exchange between nodes. Multi-AUV underwater cooperative navigation uses underwater acoustic equipment for information exchange, relative angle and distance measurement. The navigation process is shown in Fig. 1.

In Fig. 1, the master AUV is equipped with high-performance navigation equipment (with higher accuracy, smaller drift, higher sensitivity, etc.), the slave AUV is equipped with low-performance navigation equipment. We assume that the positioning accuracy of the master AUV is high enough to be a reference position. The slave AUV uses its own Doppler Velocity Log (DVL) and inertial measurement unit (IMU) for DR navigation. The master AUV performs underwater acoustic communication with the slave AUV every Δt time to transfer navigation message. The slave AUV corrects positioning error caused by DR using the message at every time acoustic communication performed. Then the slave AUV continues to be navigated by the DR, until next acoustic message is received to correct the positioning error again.

A. MATHEMATICAL MODEL

For the convenience of analysis, pitch of AUV is approximately considered as zero when AUV runs stably. It is also deemed that the depth of AUV is unchanged and the ocean current disturbance is neglected in the cooperative navigation process. The motion states of AUV are shown in Fig. 2 [25]–[27].

In Fig. 2, $X_tO_tY_t$ is the navigation coordinate system, and $X_bO_bY_b$ is the body coordinate system of AUV. $\boldsymbol{X}_k$ is the position of AUV, φ_k is the true heading angle, υ_k is the true longitudinal velocity of AUV at time k. The kinematic

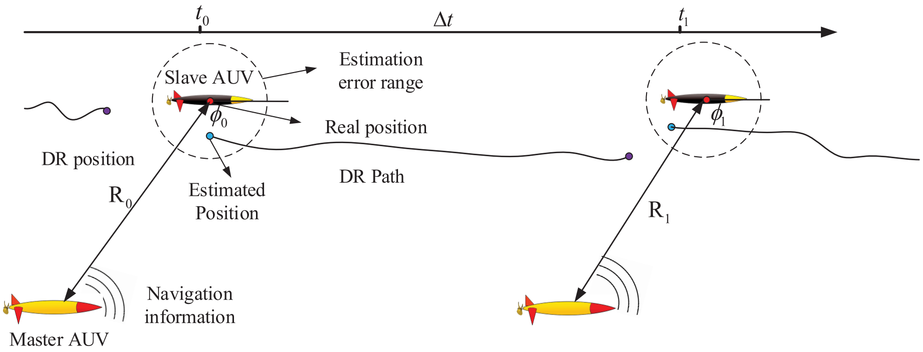


FIGURE 1. Two AUVs cooperative positioning process.

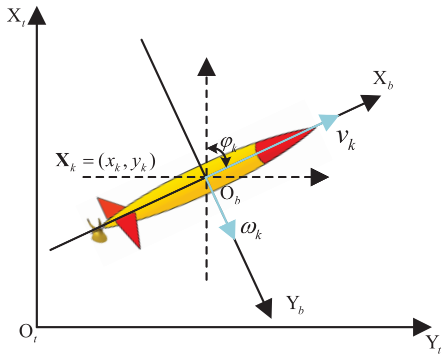


FIGURE 2. The motion states of the AUV.

equation of AUV can be expressed as:

$$\begin{cases} x_{k+1} = x_k + \Delta t \hat{\upsilon}_k cos\hat{\varphi}_k \\ y_{k+1} = y_k + \Delta t \hat{\upsilon}_k sin\hat{\varphi}_k \\ \hat{\varphi}_{k+1} = \varphi_{k+1} + w_\varphi \\ \hat{\upsilon}_{k+1} = \upsilon_{k+1} + w_\upsilon, \end{cases} \tag{1}$$

where x_{k+1} and y_{k+1} are the position of vehicle on the X and Y axes in the navigation coordinate system at time $k+1$; $\hat{\varphi}_{k+1}$ is the course angle measured by gyroscope and $\hat{\upsilon}_{k+1}$ is longitudinal velocity measured by DVL at time $k+1$, and w_φ, w_υ are the zero mean Gauss white noise caused by measurements of both the compass and DVL, respectively.

Equation (1) can be simplified to:

$$\begin{aligned} \boldsymbol{X}(k+1) &= \Phi(k+1,k)\boldsymbol{X}(k) + \Gamma(\boldsymbol{u}(k) + \boldsymbol{w}(k)) \\ &= f[\boldsymbol{X}(k), \boldsymbol{u}(k), \boldsymbol{w}(k)], \end{aligned} \tag{2}$$

where $\boldsymbol{X}(k+1) = (x_k \ y_k \ \theta_k)^{\mathrm{T}}$ is the system state matrix at time t_{k+1}, $\boldsymbol{\Phi}(k+1,k)$ is the state transition matrix of the system, $\Gamma(\boldsymbol{u}(k) + \boldsymbol{w}(k))$ is the nonlinear part of the system, $\boldsymbol{u}_k = (v_k \ \varphi_k)^{\mathrm{T}}$ and $\boldsymbol{w}_k = (w_v \ w_\varphi)^{\mathrm{T}}$ are the system process noises. Then, the measurement covariance matrix and state transition matrix can be described as

$$Q(k+1) = E[\boldsymbol{w}_{k+1} \quad \boldsymbol{w}_{k+1}^{\mathrm{T}}] = \begin{pmatrix} \sigma_{vk}^2 & 0 \\ 0 & \sigma_{\varphi k}^2 \end{pmatrix}, \tag{3}$$

$$\Phi(k+1,k) = \begin{bmatrix} 1 & 0 & -\Delta t \cdot \upsilon_k \cdot \sin\varphi_k \\ 0 & 1 & \Delta t \cdot \upsilon_k \cdot \cos\varphi_k \\ 0 & 0 & 1 \end{bmatrix}. \tag{4}$$

As the navigation and positioning error of INS will accumulate over time, underwater acoustic positioning system is required to continuously correct the positioning error. According to the principle of underwater acoustic localization, the observation equation is defined as

$$\begin{aligned} Z_{ij}(k+1) &= \left\| \boldsymbol{X}_i(k+1) - \boldsymbol{X}_j(k+1) \right\| + w_\rho(k+1) \\ &= h[\boldsymbol{X}_i(k+1), \boldsymbol{X}_j(k+1)] + w_\rho(k+1), \end{aligned} \tag{5}$$

where $Z_{ij}\,(k+1)$ represent the relative distance between AUV$_i$ and AUV$_j$ at time $k+1$; $\boldsymbol{X}_i(k+1) = [\,x_i(k+1)\ y_i(k+1)\,]$ and $\boldsymbol{X}_j(k+1) = [\,x_j(k+1)\ y_j(k+1)\,]$ respectively denote the position of AUV$_i$ and AUV$_j$ at time $k+1$; $w_\rho\,(k+1)$ is the system measurement noise with Gaussian distribution; $h[\,\boldsymbol{X}_i(k+1), \boldsymbol{X}_j(k+1)\,]$ is a nonlinear function of $\boldsymbol{X}_i(k+1)$ and $\boldsymbol{X}_j(k+1)$. Covariance matrix for systematic observation noise can be defined as

$$R_\rho(k+1) = E[\,w_\rho \quad w_\rho^{\mathrm{T}}\,] = \sigma_\rho^2. \tag{6}$$

B. OBSERVABILITY ANALYSIS

In order to avoid the loss of high-order terms after linearizing the nonlinear model in cooperative navigation, researchers in [25], [28]–[30] used the Lee derivative observability theory to analyze the observability of nonlinear systems. On this basis, Gao [25] has studied the observability analysis of multi AUVs cooperative navigation. First, the author built the state equation and observation equation of the navigation system. Then, as the nonlinear characteristics of the cooperative navigation model, observability analysis of nonlinear systems using the Lie Derivative weak observability theory was carried out. And the navigation model was based on the 2-Dimensional replaces the 3-Dimensional.

The following conclusions are derived: If the measurement angles γ between two contiguous times in underwater acoustic measurement direction are 0°, the system is unobservable at this time. The closer γ is to 0°, the lower observability of the system. When $\gamma = 90°$, observability of the system reaches to max. In other words, when the system is moving,

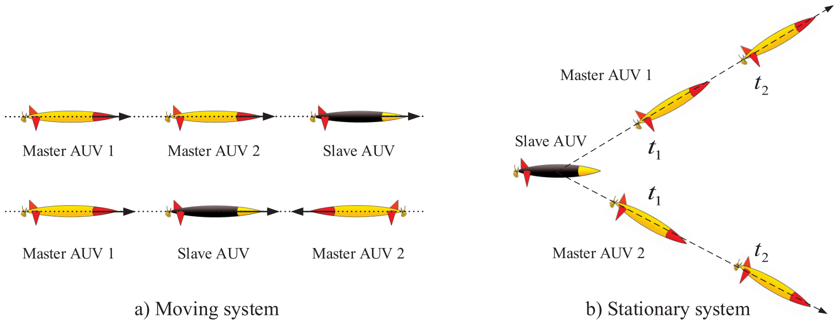


FIGURE 3. Unobservable motion states of the cooperative navigation system.

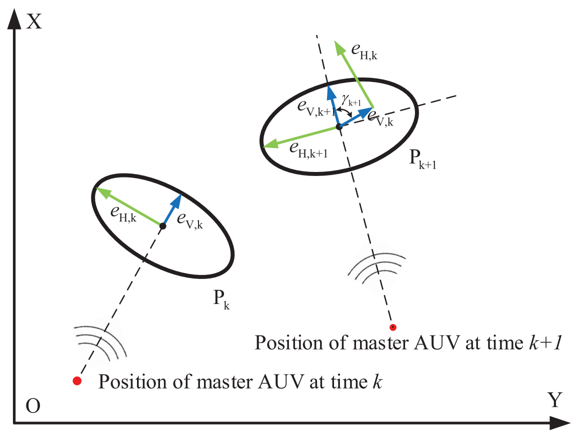


FIGURE 4. Error propagation in multi-AUV cooperative navigation.

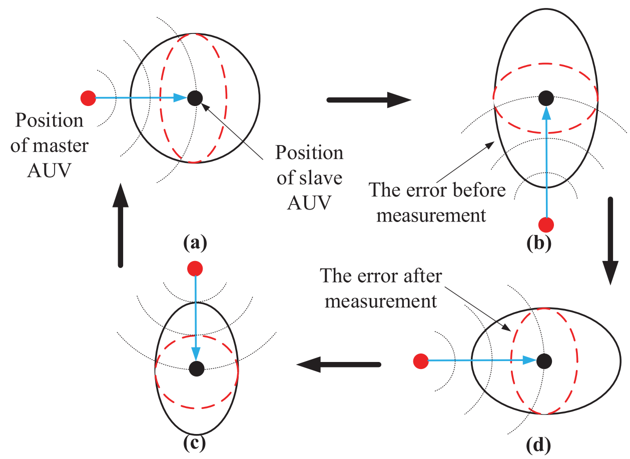


FIGURE 5. Error propagation process in underwater acoustic measurement.

the system is unobservable if and only if the master AUVs and the slave AUVs are moving along the same straight line; when the system is stationary, the system is unobservable if and only if the motion direction of the master AUVs and the slave AUVs along the measurement direction is constant, as shown in Fig. 3. Otherwise, the system is observable.

C. ERROR ANALYSIS

Maurice [31] analyzed the error of acoustic ranging. The position error of the slave AUV can be regarded as an oval error, the error in the direction of the underwater acoustic measurement is a fixed value. Fig. 4 indicates the propagation of errors during the two underwater acoustic measurements at adjacent times. γ_{k+1} is the angle between directions of underwater acoustic measurement at time $k+1$ and k. $e_{H,k}$ and $e_{V,k}$ respectively represent horizontal error and vertical error of underwater acoustic measurement between the master and the slave at time k.

$$e_{H,k+1} = \sqrt{\frac{e_{V,k}^2 e_{H,k}^2}{(e_{V,k}\cos\gamma_{k+1})^2 + (e_{H,k}\sin\gamma_{k+1})^2}} + \alpha \cdot \Delta t, \tag{7}$$

where α is a constant value, Δt is the length of communication interval.

Since $e_{V,k}$ is the error caused by underwater acoustic measurement, and $e_{H,k}$ is the error caused by DR with inertial equipment and DVL, and the inertial equipment and DVL of the slave AUVs are low accuracy equipment. So $e_{H,k}$ is far greater than $e_{V,k}$, and if $\gamma_{k+1} = 90°$, the position error of the slave AUVs is converging.

Figure 5 shows error change of the slave AUV. Master AUV first measures the distance and orientation between itself and the slave AUV by underwater acoustic as the basis for its decision-making process. Then the master AUV sends the message of its position to the slave AUV by acoustic system and the slave AUV corrects its own position with the message. Next, the slave AUV uses dead reckoning to navigate till the next communication node, and repeating the above process. In this way, the error can be limited to a fixed circumference.

III. MODEL OF PATH PLANNING

According to the above analysis, the purpose of this paper is to find the most effective navigation strategy, so that the adjacent two measurement angles of the master AUVs and the slave AUVs in the cooperative navigation process are as close as possible to 90° to improve the overall positioning accuracy. The key is to use CE algorithm to find the strategy to maximum change of measurement angle, and to keep the total observation error minimum, while maintaining the stability of the system. In order to facilitate the establishment of the system decision model, a multi-AUV collaborative navigation main framework is proposed based on MDP. The MDP framework does not require explicit system model data, which greatly facilitates the usage of CE algorithm.

A. CE ALGORITHM

CE is an important concept in the theory of Shannon information. The size of information can be measured by Information Entropy, and CE can be used to measure the degree of difference between two sets of information. From the perspective of probability selection, this paper selects the optimal action to make the change of each observation angle close to 90°. This is a sub-problem of the Optimal Combination Problem (COP). In COP, CE method involves an iterative procedure where each iteration can be broken down into two phase:

1) Generate a random data sample (trajectories, vectors, etc.) according to a specified mechanism.
2) Update the parameters of the random mechanism based on the data to produce a better sample in the next iteration.

When the COP is complex, not only is the model difficult to build, but even the equation cannot find the optimal solution. However, CE establish the problem under the probability model to estimate optimal solution through the analysis and selection of the probability.

We assume that $f(\boldsymbol{X})$ is the real probability density of $\boldsymbol{X}$, and $h(\boldsymbol{X})$ is the estimated probability density of of $\boldsymbol{X}$. CE value between $f(\boldsymbol{X})$ and $h(\boldsymbol{X})$ can be written as

$$\begin{aligned} D(f,g) &= \boldsymbol{E}_h \ln \frac{h(\boldsymbol{X})}{f(\boldsymbol{X})} \\ &= \int h(\boldsymbol{X}) \ln f(\boldsymbol{X}) d\boldsymbol{X} - \int f(\boldsymbol{X}) \ln h(\boldsymbol{X}) d\boldsymbol{X}, \end{aligned} \tag{8}$$

where $D(f,g)$ is also called the Kullback-Leibler (K-L) distance [32], which is used to measure the distance between h and f.

Equation (8) can be written as

$$\begin{cases} H(f) = \int f(\boldsymbol{X}) \ln f(\boldsymbol{X}) d\boldsymbol{X} \\ H(f,h) = -\int f(\boldsymbol{X}) \ln h(\boldsymbol{X}) d\boldsymbol{X}, \end{cases} \tag{9}$$

where $H(f)$ is the entropy of the true probability density $f(\boldsymbol{X})$, and $H(f,h)$ is the cross entropy of $f(\boldsymbol{X})$ and $h(\boldsymbol{X})$. In the optimization learning algorithm, the true probability density distribution of random variable X is invariant. Thus in (9) we ignore the influence of term $H(f)$ and focus on the influence of cross entropy term $H(f,h)$ on K-L distance. In information entropy, there is always

$$-\int f(\boldsymbol{X}) \ln f(\boldsymbol{X}) d\boldsymbol{X} \le -\int f(\boldsymbol{X}) \ln h(\boldsymbol{X}) d\boldsymbol{X}, \tag{10}$$

according to Gibbs' inequality. The equal sign is true if and only if $f(\boldsymbol{X}) = h(\boldsymbol{X})$. Therefore, $D(f,h) \ge 0$. In the case where the $H(f)$ term is fixed, in order to minimize the value of $D(f,h)$, it is necessary to make the value of $H(f,h)$ smaller. It can be obtained that

$$\max_{v} \int f(\boldsymbol{X}) \ln h(\boldsymbol{X}, v) d\boldsymbol{X}, \tag{11}$$

where v is sample data of the random variable $\boldsymbol{X}$, and $h(\boldsymbol{X}, v)$ is the probability density function estimated from the sample data v.

B. MARKOV DECISION PROCESS

The Markov decision process [33]–[35] is an optimal decision process of a stochastic dynamic system based on the Markov Process Theory. The MDP refers to decision-makers periodically or continuously observing a stochastic dynamic system with Markovian characteristics and making decisions in a sequential manner. That is, according to observed state at each moment, an action is selected from the available action set to make a decision. The next (future) state of the system is random, and its state transition probability is Markovian. Decision makers make a new decision based on the newly observed status, and then repeat this process. Markov's property refers to the nature of the probability that the stochastic process's future development is independent of the history before observation. Markov's property can be simply described as the non-post-effect of state transition probability.

MDP consists of four elements:

$$M = (S, A, P_{sa}, R), \tag{12}$$

where $S = \{s_1, s_2, s_3, \ldots, s_n\}$ indicates set of states. $A = \{a_1, a_2, a_3, \ldots, a_m\}$ indicates set of actions. P_{sa} is the state transition matrix, which indicates in the current state $s \in S$, the probability distribution of other states that will be transferred after action $a \in A$. R is the cost function, which indicates that the cost value of after performing the action a in the state of s, it can be marked as $r(s, a)$.

C. PATH PLANNING ALGORITHM

Based on the analysis above, state variables of the multi-AUV cooperative navigation system are determined first, including the heading angle $\varphi_k^{s,i}$ and the estimated position $(x_k^{s,i}, y_k^{s,i})$ of the slave AUV$_i$ at time k; the heading angle $\varphi_k^{m,j}$ and the estimated position $(x_k^{m,j}, y_k^{m,j})$ of the master AUV$_j$ at time k; the measured distance $r_k^{i,j}$ and the change of relative measured angle$\gamma_k^{i,j}$ between the slave AUV$_i$ and the master AUV$_j$ at time k, which can be expressed as

$$S \in \left\{ \varphi_k^{s,i}, (x_k^{s,i}, y_k^{s,i}), \varphi_k^{m,j}, (x_k^{m,j}, y_k^{m,j}), \gamma_k^{i,j}, \gamma_k^{i,j} \right\}. \tag{13}$$

S is a subset of the total state set and can describe the operating state of the multi-AUV collaborative navigation system at time k.

Because an AUV can directly control its heading angular velocity, we take the heading angular velocity $\omega_k^{m,j}$ of the master AUVs as an action. In order to avoid an infinite amount of computation, the angular velocity is discretized as:

$$A \in \{\omega_{\min}, \ldots, \omega_{\max}\}, \tag{14}$$

where A is a subset of the set of discrete heading angular velocities between the minimum heading angular velocity and the maximum heading angular velocity.

Let $p_{s_i,j}$ be the probability of action ω_j ($j = 0, 1, \ldots, n$) at state s_i ($i = 0, 1, \ldots, m$), then the state transition matrix can be written as:

$$P_{sa} = \begin{pmatrix} p_{s_1,1} & \cdots & p_{s_1,n} \\ \vdots & \ddots & \vdots \\ p_{s_m,1} & \cdots & p_{s_m,n} \end{pmatrix}. \tag{15}$$

The feedback cost function R is discussed in two parts. First, the influence relative measurement angle change is considered as

$$R_k^1 = \sum_{j=1}^{\mathrm{L}_n} K\left(1 - D_{j,k}\right),$$
$$D_{j,k} = \frac{2\zeta_k^j \left|\sin(\Delta\theta_k^j)\right|}{(\zeta_k^j)^2 + 1 + \sqrt{(\zeta_k^j)^4 + 2(\zeta_k^j)^2\cos(2\Delta\theta_k^j) + 1}}, \tag{16}$$

where L_n is the number of the slave AUVs. K is a constant coefficient which controls the proportion of angle change in the cost function. ζ_k^j is ratio of the distance between the slave AUV$_j$ and master AUV$_1$ to the distance between the slave AUV$_j$ and master AUV$_2$. $\Delta\theta_k^j$ indicates the angle measured by underwater acoustic equipments between the slave AUV$_j$ and the two master AUVs at time k.

Then the influence of relative distance is considered. A collision may happen if the distance is relatively small. Otherwise, the acoustic signal may be too weak to navigate if the distance is quite large. Assume that the minimum distance allowed between AUVs is $r_{\min}$, the maximum distance is $r_{\max}$, and the range suitable for navigation is $[r_{\min}, r_{\mathrm{fit}}]$. Since the slave AUVs has already planned their routes, collision between the slave AUVs is neglected. For the sake of convenience, it is assumed that the master AUVs are distributed at different depths, so the collision between the master AUVs is ignored. Therefore, only the collision avoidance problem between the master AUVs and the slave AUVs needs to be considered. The cost function of the distance can be expressed as:

$$R_k^2 = \begin{cases} C_k^1(r) \cdot \left(e^{r_{\min} - r} - 1\right) & 0 \le r < r_{\min} \\ 0 & r_{\min} \le r < r_{\mathrm{fit}} \\ C_k^2(r) \cdot \left(e^{r - r_{\mathrm{fit}}} - 1\right) & r_{\mathrm{fit}} \le r \le r_{\max}, \end{cases} \tag{17}$$

where $C_k^1(r)$ is a monotonic non-incremental function on interval $[0, r_{\min})$, $C_k^2(r)$ is a monotonic non-decreasing function on interval $[r_{\mathrm{fit}}, r_{\max}]$. Then the total cost function is defined as

$$R_k = R_k^1 + R_k^2. \tag{18}$$

After planning the path of the slave AUVs, the starting state of the master AUVs is determined. Then, the path of the master AUVs is planned according to CE algorithm. At each iteration, an action ω_i is randomly selected and executed according to the probability in the current system state S_k of the state transition matrix P_{sa}, and when the next measurement time comes, the above procedure is repeated until all the motions are completed.

The key of CE algorithm is picking the optimal sample to update the probability of samples iteratively. In one iteration process, total cost of the system under each trajectory is

$$\begin{aligned} l_1 : C_1 &= \sum_{k=1}^{N} R_k^{l_1}, \\ l_2 : C_2 &= \sum_{k=1}^{N} R_k^{l_2}, \\ &\vdots \\ l_L : C_L &= \sum_{k=1}^{N} R_k^{l_{\mathrm{L}}}, \end{aligned} \tag{19}$$

for $i = 1, \ldots, L$, l_i is the number of the master AUV trajectory. C_i is total cost of each trajectory. R_k^i is the cost of each step. Arrange C_i from small to large as

$$C = \left[C_1^{\min\_\max}, C_2^{\min\_\max}, \ldots, C_L^{\min\_\max}\right]. \tag{20}$$

The cutoff optimal sample value is determined as

$$\eta_m = C_{Co}^{\min\_\max}, \tag{21}$$

where $Co \in [1, 2, \ldots, \mathrm{L}]$, η_m represents the cutoff optimal sample value in iteration m. The optimal sample data in one iterative process can be obtained as

$$C_B = \left[C_1^{\min\_\max}, C_2^{\min\_\max}, \ldots, C_{Co}^{\min\_\max}\right]. \tag{22}$$

According to (22), times of the same state in the optimal sample are counted as

$$N_s = \sum_{i=1}^{C_o} \sum_{j=1}^{N} I_{\{S_{i,j}=s\}}, \tag{23}$$

where N_s represents the time of state $S_{i,j}$ and state s are same from the first step to the Nth step in the optimal sample path i. Then the number of identical actions in state s is counted as

$$N_{sa} = \sum_{i=1}^{C_o} \sum_{j=1}^{N} I_{\{(S_{i,j}=s)\cap(A_{i,j}=a)\}}, \tag{24}$$

where N_{sa} represents the time of identical actions a in state s from the first step to the Nth step in the optimal sample path i.

Then the state transition matrix P_{sa} is updated as

$$p_{sa} = \mu_{m,1} p_{sa} + \frac{\mu_{m,2} N_{sa}}{N_s}, \tag{25}$$

TABLE 1. Parameters of localization of the master AUVs and the slave AUVs.

AUV	Speed (m/s)	DVL Measurement Noise (m)	Heading Measurement Noise (rad)	Underwater Acoustic Noise (m)
Master AUV	3.0	$N(0, 0.5^2)$	$N(0, 0.1^2)$	$N(0, 8.0^2)$
Slave AUV	3.0	$N(0, 15^2)$	$N(0, 0.5^2)$	$N(0, 8.0^2)$

where $p_{sa} \in P_{sa}$. In order to avoid falling into local optimization, the probability of previous moment and the probability calculated at current moment are weighted. $\mu_{m,1}$ and $\mu_{m,2}$ is the weighting coefficient, we have

$$\mu_{m,1} + \mu_{m,2} = 1, \tag{26}$$

where $\mu_{m,1} < \mu_{m,2}$. As the number of update iterations increases, $\mu_{m,1}$ will decrease while $\mu_{m,2}$ will increase.

Finally, iteration stop condition is

$$|\eta_m - \eta_{m-1}| \leq \varphi_{\text{end}}. \tag{27}$$

In order to get the optimal path of the master AUVs, P_{sa} needs to be continuously optimized so that the total cost of the selected actions can be nearly minimized and the navigation accuracy can be near optimal.

The training process can be described as follows:

(a) Initialize the state transition matrix P_{sa}, cut-off cost value φ_{end}, etc.
(b) Generate the target path of the slave AUVs include N communication nodes.
(c) Initialize the state of the system S_1. Pick an action randomly from P_{sa} and reach a new state. Then calculate the cost value. Repeat this procedure until measurements of N communication nodes are complete. Total cost of this path is calculate according to (18).
(d) Repeat (c) till L master AUV paths are selected, and choose the optimal sample based on the (20) to (24).
(e) Update the state transition matrix according to (25) and verify the conditions of (27). If the condition is met, the iteration ends, otherwise return to (c) to continue iteration.

IV. RESULT

In this section, we design the simulation programs to analyze the performance of CE navigation algorithm. After setting trajectories of the slave AUVs as straight line and harvester line, the state transition matrix P_{sa} is finally obtained according to training process master AUVs using CE algorithm. In the process of cooperative navigation, the master AUVs choose actions from P_{sa} to maximize the change of observation angle. The EKF is applied to obtain the localization results of the slave AUVs.

The parameters of the master AUVs and the slave AUVs are shown in Table 1.

Three parallel curves are used as the path of the slave AUV, and two master AUVs are used for cooperative navigation. There are 100 communication nodes in the navigation process, and the communication time interval is 10 seconds.

The selection of system states and the discretization of state variables are shown in the table 2.

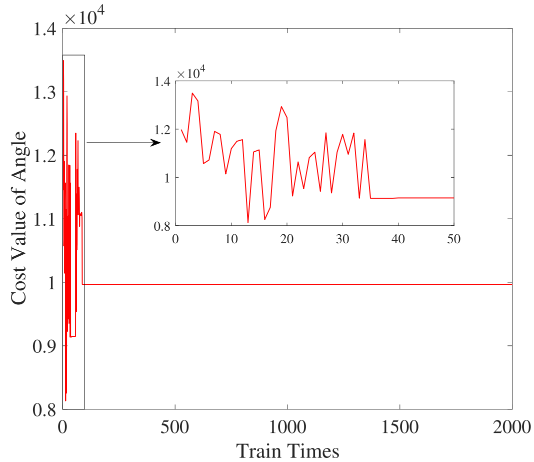


FIGURE 6. Cost of measured angle.

Actions of MDP are discrete heading angular velocity (rad/s), namely

$$A \in \{-0.08, -0.05, -0.03, 0.00, 0.03, 0.05, 0.08\}. \tag{28}$$

A. TRAINING PROCESS

In order to balance the observation angle change and the distance between the master and slave AUVs, after several rounds of trial and error, K is taken as 50. Set $r_{\min} = 100$, $r_{\text{fit}} = 700$, $r_{\max} = +\infty$, $C_k^1(r) = C_k^2(r) = 50.0$. In each train process, 300 master AUV trajectories is generated based on the state and their cost value is calculated. Then 30 paths with minimum cost are used as the optimal sample to update the probability values of P_{sa}. After 2000 times of training, the training process terminates. The main propose of the training is to get two state transition matrices of the master AUVs, so that they can choose certain action in the cooperative navigation.

The change of cost value caused by measured angle in the training process is shown in Fig. 6. It is shown that the cost is fluctuant during the first 35 sessions. After the 40th train process, the cost is stable at 9148.

The change of cost value caused by relative measured distance in the training process is shown in Fig. 7. During the first 4 sessions, the cost descends dramatically and is stable at 1186.

The trajectories of master AUVs selected by P_{sa} is shown in Fig. 8. Three slave AUVs move in sinusoidal curves and are in the same state of motion. The initial position of the slave AUV_1 is $(-300.0, 0.0)$, the slave AUV_2 is $(0.0, 0.0)$, and the slave AUV_3 is $(0.0, 300.0)$. Their initial heading is 45°. The initial position of the master AUV_1 is $(-150.0, -100.0)$, and

TABLE 2. State variables of MDP.

State variable	Discretization	Quantity
Heading of master AUV_1	$[0,30), [30,60), \ldots, [330,360)$	12
Heading of master AUV_2	$[0,30), [30,60), \ldots, [330,360)$	12
Measured distance between master AUV_1 and slave AUV_1	$[0,100), [100,700), [700,+\infty)$	3
Measured distance between master AUV_1 and slave AUV_2	$[0,100), [100,700), [700,+\infty)$	3
Measured distance between master AUV_1 and slave AUV_3	$[0,100), [100,700), [700,+\infty)$	3
Measured distance between master AUV_2 and slave AUV_1	$[0,100), [100,700), [700,+\infty)$	3
Measured distance between master AUV_2 and slave AUV_2	$[0,100), [100,700), [700,+\infty)$	3
Measured distance between master AUV_2 and slave AUV_3	$[0,100), [100,700), [700,+\infty)$	3
Measured distance between master AUV and slave AUV_1	$[0,30), [30,60), [60,90]$	3
Measured distance between master AUV and slave AUV_2	$[0,30), [30,60), [60,90]$	3
Measured distance between master AUV and slave AUV_3	$[0,30), [30,60), [60,90]$	3

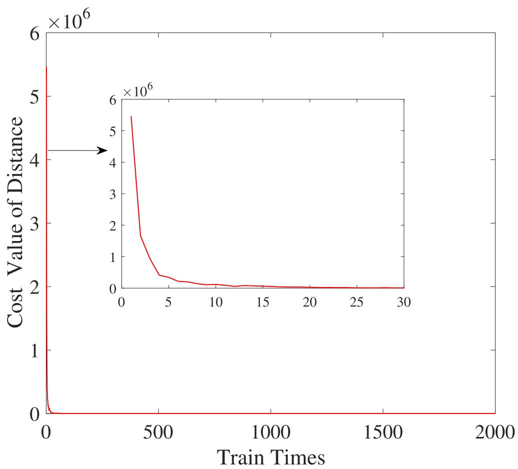


FIGURE 7. Cost of relative measured distance.

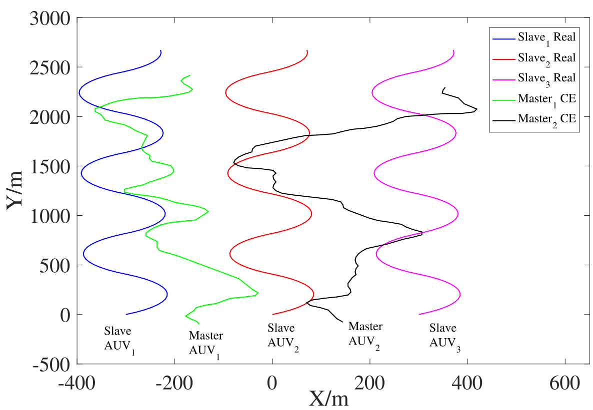


FIGURE 8. Trajectories of master AUVs based on CE algorithm.

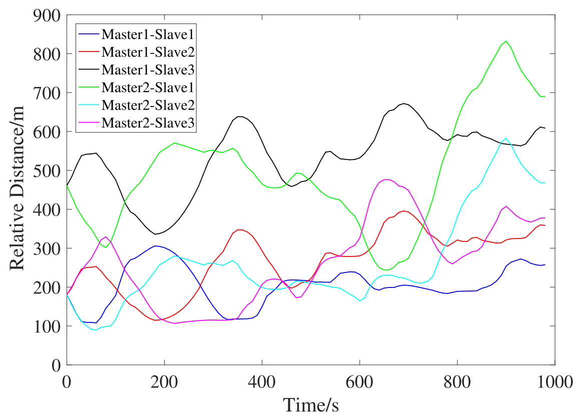


FIGURE 9. Relative distance between master AUVs and slave AUVs.

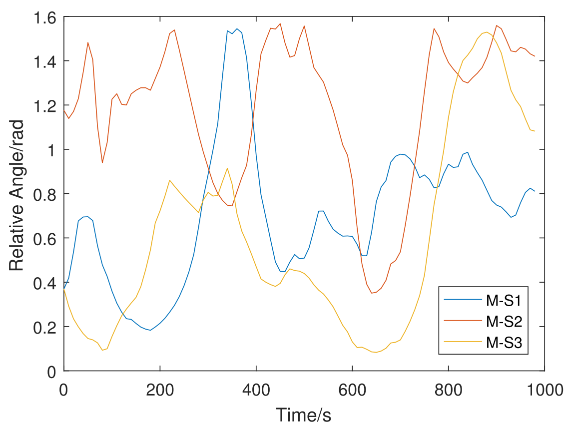


FIGURE 10. Measured angle between master AUVs and slave AUVs.

the master AUV_2 is $(150.0, -100.0)$. Their initial heading is 90°.

Relative distance between the master AUVs and the slave AUVs in Fig. 8 is shown in Fig. 9. It can be seen that distance between each master AUV and each slave AUV are in suitable range during the process.

Measured angle between the master AUVs and the slave AUVs in Fig. 8 is shown in Fig. 10. The probability change of action selection of master AUV_1 and AUV_2 is shown in Fig. 11, where the probability of master AUV_1 choosing action is 1 and the average of AUV_2 is 0.9899 and most of them are 1. This indicates that these state-action points are well-trained and are able to select proper action to execute according to the current state.

B. COOPERATIVE NAVIGATION SIMULATION

Based on the trajectories of the slave AUVs in section IV-A, DR paths of the slave AUVs is shown in Fig. 12. It demonstrates that the DR positions of the slave AUVs lag behind their real position and the error continues to increase.

TABLE 3. Positioning error of DR and CE.

Slave AUV	Average DR error	Max DR error	Average cooperative navigation error	max cooperative navigation error
AUV_1	160.8	314.1	33.48	65.15
AUV_2	167.8	324.8	25.14	66.62
AUV_3	168.9	336.8	49.82	142.2

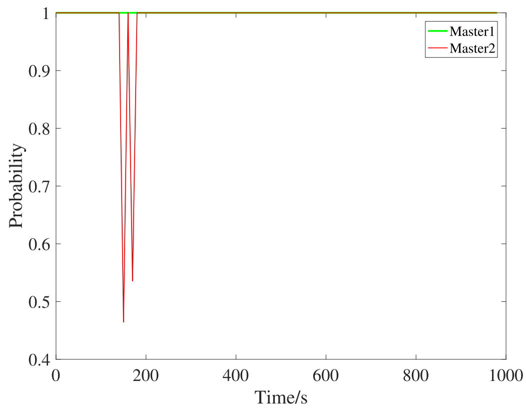


FIGURE 11. Probability of master AUVs selecting actions.

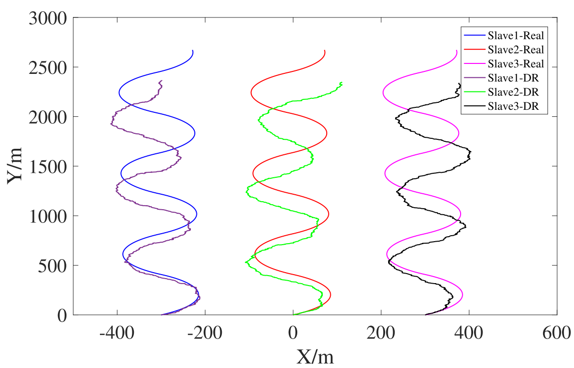


FIGURE 12. DR trajectories of slave AUVs.

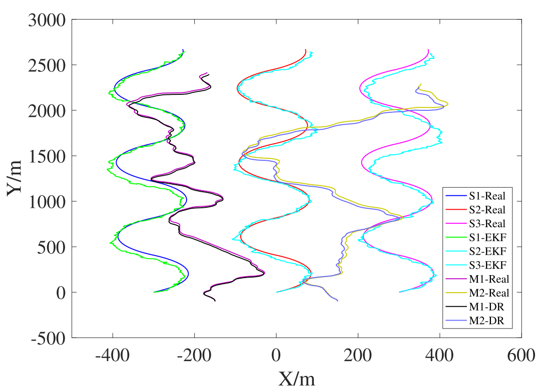


FIGURE 13. Cooperative navigation trajectories of the slave AUVs.

The navigation trajectories using cooperative navigation system with aforementioned CE algorithm and EKF are shown in Fig. 13. It can be seen intuitively from Fig. 13 that the trajectories obtained by cooperative navigation have only small difference between the real trajectories, and can follow every turn of the latter. The detailed data are listed in table 3

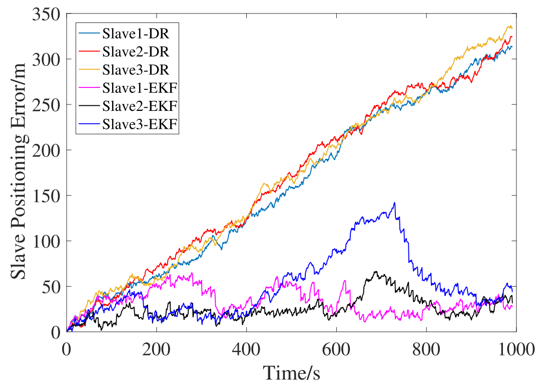


FIGURE 14. Localization error of slave AUVs.

Localization error of the aforementioned two kinds of navigation method is shown in Fig. 14. In this figure, it is obvious that the localization error of DR continues to increase while CE algorithm trained cooperative navigation is able to restrain the growth of error.

V. CONCLUSION

In this paper, the CE algorithm and MDP are used to generate a novel cooperative navigation method with dual master AUVs and multiple slave AUVs. Firstly, a cost function based on the mathematic model, observability and error model of the system is established, whose purpose is to quantify the advantages and disadvantages of the actions chosen by MDP according to the state transition matrices. During the training process, the system optimize the state transition matrices according to CE algorithm to minimize the cost of generated trajectories. In the end of the training process, probabilistic optimal state transition matrices are obtained. The simulation results show that the training process of CE algorithm is able to generate optimal state transition matrices of two master AUVs so that the best actions can be chosen at certain state. In addition with the help of the EKF, the proposed method can reduce the localization error efficiently.

There are still some aspects to be improved in the future. For example, adding the amounts of actions to choose such as changing the speed of master AUVs will improve the navigation accuracy. Considering the ocean current and time delay of acoustic communication may also help to improve the robustness of the algorithm.

REFERENCES

[1] L. Paull, S. Saeedi, M. Seto, and H. Li, "AUV navigation and localization: A review," *IEEE J. Ocean. Eng.*, vol. 39, no. 1, pp. 131–149, Jan. 2014.

[2] C. Sun, Y. Zhang, G. Wang, and W. Gao, "A maximum correntropy divided difference filter for cooperative localization," *IEEE Access*, vol. 6, pp. 41720–41727, 2018.

[3] B. Xu, S. Li, A. A. Razzaqi, and J. Zhang, "Cooperative localization in harsh underwater environment based on the MC-ANFIS," *IEEE Access*, vol. 7, pp. 55407–55421, 2019.

[4] B. Xu, J. L. Bai, Y. L. Hao, W. Gao, and Y. L. Liu, "The research status and progress of cooperative navigation for multiple AUVs," *Acta Automat. Sinica*, vol. 41, pp. 445–461, Mar. 2015.

[5] H. Mythrehee and A. Julian, "A cross layer UWSN architecture for marine environment monitoring," in *Proc. Global Conf. Commun. Technol.*, Apr. 2015, pp. 211–216.

[6] W. Gao, Y. Liu, and B. Xu, "Robust Huber-based iterated divided difference filtering with application to cooperative localization of autonomous underwater vehicles," *Sensors*, vol. 14, no. 12, pp. 24523–24542, 2014.

[7] X. Huang, L. Liu, T. Lv, and W. Li, "Cooperative navigation for multiple AUVs based on relative range measurements with a single leader," *Appl. Mech. Mater.*, vol. 556, pp. 3117–3123, May 2014.

[8] Z. Yan, L. Wang, T. Wang, Z. Yang, T. Chen, and J. Xu, "Polar cooperative navigation algorithm for multi-unmanned underwater vehicles considering communication delays," *Sensors*, vol. 18, no. 4, p. 1044, Mar. 2018.

[9] W. Xing, Y. Zhao, and H. R. Karimi, "Convergence analysis on multi-auv systems with leader-follower architecture," *IEEE Access*, vol. 5, pp. 853–868, 2017.

[10] G. Wei, L. Yalong, and X. Bo, "Multiple-AUV cooperative navigation based on two-leader alternated navigation," *J. Harbin Eng. Univ.*, vol. 35, no. 6, pp. 735–740, 2014.

[11] G. Xiao, B. Wang, Z. Deng, M. Fu, and Y. Ling, "An acoustic communication time delays compensation approach for master–slave AUV cooperative navigation," *IEEE Sensors J.*, vol. 17, no. 2, pp. 504–513, Jan. 2016.

[12] C. Jian-Hua, Z. Hai, G. De-Hong, and G. Zheng-Dong, "An AUV positioning error correct algorithm based on acoustic propagation time compensation," in *Proc. IEEE Int. Conf. Signal Process., Commun. Comput.*, Sep. 2015, pp. 1–5.

[13] Y. Qi, B. Wang, S. Wang, and M. Fu, "Cooperative navigation for multiple autonomous underwater vehicles with time delayed measurements," in *Proc. IEEE Chin. Guid., Navigat. Control Conf.*, Aug. 2016, pp. 295–299.

[14] Q. Li, Y. Ben, S. M. Naqvi, J. A. Neasham, and J. A. Chambers, "Robust student's*t*-based cooperative navigation for autonomous underwater vehicles," *IEEE Trans. Instrum. Meas.*, vol. 67, no. 8, pp. 1762–1777, Aug. 2018.

[15] Q. Chen, K. You, and S. Song, "Cooperative localization for autonomous underwater vehicles using parallel projection," in *Proc. 13th IEEE Int. Conf. Control Automat.*, Jul. 2017, pp. 788–793.

[16] Y. Huang, Y. Zhang, B. Xu, Z. Wu, and J. A. Chambers, "A new adaptive extended Kalman filter for cooperative localization," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 54, no. 1, pp. 353–368, Feb. 2018.

[17] B. Allotta, A. Caiti, L. Chisci, R. Costanzi, F. Di Corato, C. Fantacci, D. Fenucci, E. Meli, and A. Ridolfi, "An unscented Kalman filter based navigation algorithm for autonomous underwater vehicles," *Mechatronics*, vol. 39, pp. 185–195, Nov. 2016.

[18] D. Viegas, P. Batista, P. Oliveira, C. Silvestre, and C. L. P. Chen, "Distributed state estimation for linear multi-agent systems with time-varying measurement topology," *Automatica*, vol. 54, pp. 72–79, Apr. 2015.

[19] J. Ye, "Single valued neutrosophic cross-entropy for multicriteria decision making problems," *Appl. Math. Model.*, vol. 38, no. 3, pp. 1170–1175, 2014.

[20] G. Wei, "Picture fuzzy cross-entropy for multiple attribute decision making problems," *J. Bus. Econ. Manage.*, vol. 17, no. 4, pp. 491–502, 2016.

[21] V. L. Croft, and W. Bruce, "Relevance-based language models," *ACM SIGIR Forum*, vol. 51, no. 2, pp. 260–267, pp. 120–127, 2017.

[22] M. F. Fallon, G. Papadopoulos, J. J. Leonard, and N. M. Patrikalakis, "Cooperative AUV navigation using a single maneuvering surface craft," *Int. J. Robot. Res.*, vol. 29, no. 12, pp. 1461–1474, Oct. 2010.

[23] S. E. Webster, R. M. Eustice, H. Singh, and L. L. Whitcomb, "Advances in single-beacon one-way-travel-time acoustic navigation for underwater vehicles," *Int. J. Robot. Res.*, vol. 31, no. 8, pp. 935–950, 2012.

[24] T. Y. Teck and M. Chitre, "Single beacon cooperative path planning using cross-entropy method," in *Proc. OCEANS MTS/IEEE KONA*, Sep. 2011, pp. 1–6.

[25] W. Gao, Y. Liu, and B. Xu, "Observability analysis of cooperative navigation system for multiple AUV based on two-leaders," *Syst. Eng. Electron.*, vol. 35, no. 11, pp. 2370–2375, 2013.

[26] L. Zhang, D. Xu, and M. Liu, "Cooperative navigation for multiple autonomous underwater vehicles based on two hydrophones," *Syst. Eng. Electron.*, vol. 33, no. 7, pp. 1603–1606, 2011.

[27] L. Zhang, T. Wang, F. Zhang, and D. Xu, "Cooperative localization for multi-AUVs based on GM-PHD filters and information entropy theory," *Sensors*, vol. 17, no. 10, p. 2286, 2017.

[28] T. Paradowski, B. Tibken, and R. Swiatlak, "An approach to determine observability of nonlinear systems using interval analysis," in *Proc. Amer. Control Conf.*, May 2017, pp. 3932–3937.

[29] S. Heo, H. Kim, and C. G. Park, "Observability analysis of in-flight calibration of gyros and attitude sensors on orbit," in *Proc. IEEE/ION Position, Location Navigat. Symp.*, Apr. 2016, pp. 964–969.

[30] X. S. Zhou and S. I. Roumeliotis, "Robot-to-robot relative pose estimation from range measurements," *IEEE Trans. Robot.*, vol. 24, no. 6, pp. 1379–1393, Dec. 2008.

[31] M. F. Fallon, M. Kaess, H. Johannsson, and J. J. Leonard, "Efficient AUV navigation fusing acoustic ranging and side-scan sonar," in *Proc. IEEE Int. Conf. Robot. Automat.*, May 2011, pp. 2398–2405.

[32] H. Liang, R. C. Anderson-Sprecher, R. F. Kubichek, and G. Talwar, "A novel approach to approximate Kullback-Leibler distance rate for hidden Markov models," in *Proc. 39th Asilomar Conf. Signals, Syst. Comput.*, Oct./Nov. 2005, pp. 869–873.

[33] G. D'Angelo, M. Tipaldi, L. Glielmo, and S. Rampone, "Spacecraft autonomy modeled via Markov decision process and associative rule-based machine learning," in *Proc. IEEE Int. Workshop Metrol. Aerosp.*, Jun. 2017, pp. 324–329.

[34] S. K. Jayaweera, *Markov Decision Processes*. Hoboken, NJ, USA: Wiley, 2015.

[35] L. Liu and G. S. Sukhatme, "A solution to time-varying Markov decision processes," *IEEE Robot. Automat. Lett.*, vol. 3, no. 3, pp. 1631–1638, Jul. 2018.

LICHUAN ZHANG was born in Hebei, China, in 1982. He received the B.S. degree in automation and the Ph.D. degree in ocean engineering from Northwestern Polytechnical University, China, in 2004 and 2009, respectively.

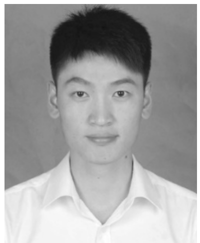
His research interests include underwater robot navigation and control technology, inertial and integrated navigation theory and application, and precision guidance, control, and simulation.

YICHEN LI was born in Shaanxi, China, in 1995. He received the B.E. degree in automation from Chang'an University, China, in 2017. He is currently pursuing the M.S. degree in shipbuilding and oceanography engineering with Northwestern Polytechnical University, China.

His research interests include underwater robot navigation and control technology, underwater formation control, and so on.

LU LIU received the Ph.D. degree in automation from Northeastern University, China, in 2017. From 2015 to 2017, she was a Visiting Ph.D. Student with the University of California at Merced, Merced, USA. She joined Northwestern Polytechnical University, China, in 2017.

Her current research interests include control system analysis, controller design, intelligent control, and so on.



XIUYE TAO was born in Xiaogan, Hubei, China, in 1992. He received the M.S. degree in weapons science and technology from Northwestern Polytechnical University, China, in 2019.

He majored in cooperative navigation of multiple AUVs. His research interests include decision and planning, and artificial intelligence.