IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received September 3, 2020, accepted September 18, 2020, date of publication September 29, 2020, date of current version October 8, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3027571

Mobile Robot's Sensorimotor Developmental Learning From Orientation and Curiosity

XIAOPING ZHANG[1], XIAOGANG RUAN[2], HONG ZHANG[3], LEI LIU[1], CUNWU HAN[1], AND LI WANG[1]

[1]School of Electrical and Control Engineering, North China University of Technology, Beijing 100144, China
[2]Faculty of Information Technology, Beijing University of Technology, Beijing 100124, China
[3]School of Automation, Xi'an University of Posts and Telecommunications, Xi'an 710121, China

Corresponding author: Xiaoping Zhang (zhangxiaoping369@163.com)

This work was supported in part by the Beijing Natural Science Foundation under Grant 4204096 and Grant 4202022, in part by the National Natural Science Foundation of China under Grant 61903006 and Grant 61573024, in part by the Beijing Municipal Great Wall Scholar Program under Grant CIT&TCD 20190304, in part by the National Key Research and Development Program of China under Grant 2017YFC0821102 and Grant 2017YFC0822504, and in part by the Youth Yuyou Talent Project and Research Initial Foundation of North China University of Technology.

ABSTRACT Simulating biological intelligence has been proved to be an effective way to design intelligent robots, and simultaneously can solve the problems existing in machine learning methods. For creatures, their motor skills achieving is the first stage of learning. By combining two important cognitive elements: orientation and curiosity, this article proposes a new neurobiologically-inspired sensorimotor developmental learning method for the mobile robot. In this method, curiosity promotes robot's exploration of the environment, while orientation enhances robot's exploitation knowledge of the environment. The orientation cognitive algorithm is designed based on Skinner's operant conditioning theory, and its rationality is proved. The balance of exploration and exploitation, which is a key problem for all the cognitive learning method, is solved in this method. The developmental learning process can avoid fixed sensorimotor mapping space problem, and help reduce learning waste as well as computing waste. All of the developmental learning method's characters are finally verified via simulations on a virtual mobile robot.

INDEX TERMS Artificial curiosity, autonomous robot, developmental learning, orientation, sensorimotor skill.

I. INTRODUCTION

Humans and animals could autonomously acquire knowledge and skills through their interaction with the environment, and all these derive from their strong sensorimotor ability. The psychologist Piaget pointed out that the first stage of human's cognitive development was the acquisition of their sensorimotor skills [1], which showed the importance of the sensorimotor system for biological learning. In the opinion of neurophysiology, the motor skills of humans and animals are gradually formed and developed during their continuous interaction with the environment through their organs. Research in robotics make much of the embodied cognition, whose core is motor [2]. All above from different perspectives point out the importance of sensorimotor system in cognitive learning process.

Research in biology related fields showing that sensorimotor learning is not simple reactive processes, but driven by orientation, curiosity, emotion and so on [3]. These elements all play important roles during sensorimotor learning but work in different ways. Understanding these elements' working principles and simulating them from artificial aspects will greatly improve our ways in designing cognitive models for robots.

On the basis of sensorimotor system, and combining related learning mechanism in cognitive psychology, designing cognitive models and duplicating them to robots has been a useful way to achieve autonomous robots. In this article, we will consider orientation and curiosity in artificial sensorimotor system's cognitive model designing, and finally realize the robot's developmental learning.

The associate editor coordinating the review of this manuscript and approving it for publication was Francisco J. Garcia-Penalvo.

II. RELATED WORKS

A. ARTIFICIAL SENSORIMOTOR SYSTEM

There have been a lot of research show that the realization of biological intelligence depends on the realization of biological nervous system, and it has been more than

20 years since the study of artificial sensorimotor system. As early as in 1995, Webb [4] established a set of auditory sensorimotor system for machine cricket. The cricket had an auditory sensor on each leg to act as ear, and two bump sensors (microswitches) as well as two infrared sensors to detect obstacles ahead at a range of about 30 m. All these together formed the perception unit of the machine cricket. The motor unit consisted of two wheels propelled by two independent motors, and a third ball-bearing castor wheel in front. The motor unit had three states: forward, backward, and stop. Thus, the sensorimotor cycle "perception-hearing the sound-considering in the context of action-finding the sound" was formed. Based on "if-then" rule, they successfully simulated the female cricket's tendency behavior to the male cricket through "calling song". In 2006, capturing essential features of biological systems, Kuniyoshi and Sangawa established a model of neuro-musculo-skeletal system, which consisted of skeleton, muscles, spindles, tendon organs, spinal circuits, medullar circuits (CPGs) and a basic cortical model. Through a series of experiments with a minimally simple body model, it was shown that the model had the capability of generating partially ordered behavior, a mixture of chaotic exploration and ordered entrained patterns. On the other hand, results showed the possibility that a rich variety of meaningful behavior could be discovered and acquired by the neural-body dynamics without predefined coordinated control circuits [5]. In 2015, Laflaquière *et al.* [6] simulated a redundant robotic arm with a retina installed at its endpoint, and thus formed a visual sensorimotor system. Then they designed algorithms from the perspective of control science, and enabled the robot to learn the configuration space of its retina and to gradually develop perceptual notions starting from scratch. In the brain, Gain-Field (GF) neurons in the parietal cortex are involved in computing the necessary spatial transformations for aligning the tactile, visual and proprioceptive signals. And in reaching tasks, these GF neurons exploit a mechanism based on multiplicative interactive for binding simultaneously touched events from the hand with visual and proprioception information. Based on GF neurons, in 2019, Pugach *et al.* [7] proposed a neural model to integrate tactile events with arm postures and visual locations for constructing hand-and target-centered receptive fields in the visual space.

At present, the research related to artificial sensorimotor system, as mentioned above, in the cycle of "perception-cognition-motion-perception", is more focused on the physical realization of the sensor (such as vision, hearing, touch, smell, etc.) or the motor (such as arm, foot, wheel, etc.), and the algorithms that used are mostly aimed at engineering realization, rarely involving cognitive factors.

B. SENSORIMOTOR COGNITIVE MECHANISMS

Research shows that the sensory-motor process is not a simple input-output information transfer, but contains a large number of cognitive processes, which relate to the cognitive neurons as well as cognitive mechanisms in the nerve center.

1) BEHAVIORIST MECHANISM

In 1938, American psychologist and behaviorist Skinner researched on animals' behaviors, and put forward the concept of "operating conditioning (OC)," which has been proved to be an important learning mechanism in human and animal's nervous system [8]. There are also a lot of application examples of OC in robot systems. In 2005, Itoh *et al.* put forward a new behavior model under the OC theory, which helped the robot WE-4RII learn the handshake skill [9]. In 2017, Zhang *et al.* designed a learning algorithm under OC mechanism, and with discrete motion spaces, a two-wheeled robot learned the skill of self-balancing [10]. In 2018, Arena *et al.* [11] proposed an insect-inspired body size learning algorithm, and adopted it to a humanoid robot and a control system who was composed of a series of layers developed using spiking neurons. The final processing layer was considered to be a gate to determine if an object is reachable or not depending on its estimated distance, and the correct decision was therefore learned through an operant conditioning method. To demonstrate the potential application of the learning method, a Darwin-OP robot equipped with an extended hand was used. The robot could freely move in an environment to discover objects. Both of the results in dynamic simulation environment and with the Darwin-OP robot proved that using operant conditioning, the learning scheme was able to enable the robot to learn which objects could be reached via estimating the objects' distances by varying the length of the equipped tool. Cyr's team has a deep accumulation in the artificial research of OC theory [12]. In 2019, Cyr's team proposed an artificial spiking neural network (SNN) sustaining the cognitive abstract process of spatial concept learning, and embedded it in virtual and real robots [13]. Results showed that based on OC theory, robots could learn the relationship of horizontal/vertical and left/right visual stimuli, regardless of their specific pattern composition or their location on the images. After acquisition learning phase, tests with novel patterns and locations were also successfully completed, which proved that the SNN could adapt its behavior in real time when the rewarding rule changes.

2) INTRINSIC MOTIVATION MECHANISM

OC theory builds up the relationship between skill learning and external rewards. However, research in biology shows that in addition to the external rewards, creature's sensorimotor learning process is also driven by intrinsic motivation, such as orientation, curiosity, emotion, belief and so on. Intrinsic motivation can drive creature to learn skills autonomously without external motivation, and is considered as the power and source of all cognition [3]. In 2004, Barto and Singh firstly presented computational study results of intrinsically motivated learning aiming at allowing artificial agents to construct and extend hierarchies of reusable skills that were needed for competent autonomy [14]. The core of the model was mainly based on reinforcement learning framework, later then, the research on intrinsic motivation became more and more diversified.

In terms of artificial curiosity, Oudeyer's team has achieved very representative research results. In 2007, after discussing related research coming from developmental psychology, neuroscience, developmental robotics, and active learning, they presented an intrinsic motivation system named Intelligent Adaptive Curiosity (IAC), which could push a robot towards situations in which it could maximize its learning progress [15]. Applying IAC to a physical robot which was placed on a baby mat with objects that the robot could learn to manipulate, experimental results showed that the robot first spent time in situations which were easy to learn, then shifted its attention progressively to situations of increasing difficulty, meanwhile avoiding situations in which nothing could be learned. In 2014, to achieve the robot's skill learning, Oudeyer's team continued to combine intrinsic motivation learning with imitation learning, and put forward an architecture called SGIM-D (Socially Guided Intrinsic Motivation by Demonstration), which allowed effective learning of high-dimensional continuous sensorimotor inverse models in the robot [16]. Experiments were taken on a robot arm with the learning task of using a flexible fishing line. Results proved that SGIM-D efficiently combined the advantages of social learning, intrinsic motivation and benefits from human demonstration properties to learn how to produce varied outcomes in the environment, while developing more precise control policies in large spaces. In 2019, Oudeyer *et al.* extended Universal Value Function Approximators (UVFA), and proposed CURIOUS which enabled intrinsically motivated agents to learn to achieve both of multiple tasks and multiple goals within a unique policy, leveraging hindsight learning [17]. Putting the robot in an open environment, then the robot needed to autonomously choose the object that it will practice, here the CURIOUS algorithm was used to influence UVFA and target task learning mechanism. Experiment results showed that the robot could realize self-organization of target learning by paying attention to the targets with different complexity in sequence and focusing on the forgotten targets again. In 2018, Ren *et al.* proposed a new neurobiologically-inspired cognitive computational model: C-DCCM (Curiosity-Driven Cognitive Computing Model), and implemented the effect of curiosity to the balance learning problem of the two-wheeled robot [18]. Verification was conducted via simulation on a computing system of sustained dopamine modulation mechanism and its effects on the cerebral cortex.

In the aspect of orientation, in 2016, Sadeghi *et al.* studied it in a plant-inspired robot named Plantoid with soft differential bending capabilities [19]. In 2018, this team presented a plant root behavior-based approach to define the control architecture of a plant-root-inspired robot, which was composed of three root-agents for nutrient uptake and one shoot-agent for nutrient redistribution [20]. With the previously proposed orientation-inspired control, the root-agents could ideally and autonomously grow at the best speed, exploit nutrient distribution and improve performance, in terms of exploration capabilities and exploitation of resources.

During agents' learning, both of external motivation such as in OC and intrinsic motivation play important roles. Integrating different cognitive mechanisms can help to improve the robot's learning ability more effectively, and there have been many researchers who put forward new research ideas based on this thought. In 2011, on the basis of Singh and Barto's work about intrinsically motivated reinforcement learning (IMRL), and combining with the emotional appraisal, Sequeira *et al.* introduced four common evaluation dimensions which are novelty, motivation, valence and control to IMRL, and proposed a new intrinsic motivation learning framework, and successfully simulated the foraging behavior [21]. In 2012, based on Distributed Adaptive Control (DAC), Mathews *et al.* proposed an artificial sensorimotor system named PASAR with a single framework [22]. PASAR integrated the elements of prediction, anticipation, sensation, attention, and response. Simulation experiment and real world task results verified PASAR's feasibility in solving complex real world problems and building world model with limited resources.

Simultaneously considering the external OC behaviorist mechanism and orientation as well as curiosity intrinsic motivation mechanisms, this article proposed a new cognitive method to help mobile robots learn its environment and fulfill the path planning task in a developmental way.

III. VIRTUAL ROBOT AND SIMULATION ENVIRONMENT

A. THE VIRTUAL ROBOT

A virtual robot as shown in Fig. 1 is used to show the properties of our developmental learning method. For simplicity, the robot is in circle shape with the radius of $r = 0.1$ m. Five sonar sensors are supposed uniformly distributed in the front of the robot to measure its distances to obstacles. The robot's moving is realized by differentially driving the two wheels on its both sides. While moving, the robot is supposed to turn firstly, and then move forward with a fixed speed of $v = 0.1$ m/s. So during the robot's sensorimotor process, the learning objective can be described as choosing different steering angles $\Delta\theta$ under different sensory situations. Supposing that the robot's positions and its heading angles to the x axis before and after each moving step are (x_o, y_o, θ_o) and (x_n, y_n, θ_n) respectively, then the robot's kinematic model can be described as in (1).

$$\begin{cases} \theta_n = \theta_o + \Delta\theta, \\ x_n = x_o + v \times t_s \times \theta_n, \\ y_n = y_o + v \times t_s \times \theta_n. \end{cases} \tag{1}$$

where $t_s = 0.5$ s is the sampling time, and the steering angle $\Delta\theta$ can be realized by:

$$\Delta\theta = \frac{v_r - v_l}{r} \times t_s. \tag{2}$$

In (2), we can see that for steering angle, counter-clockwise direction is the positive direction.

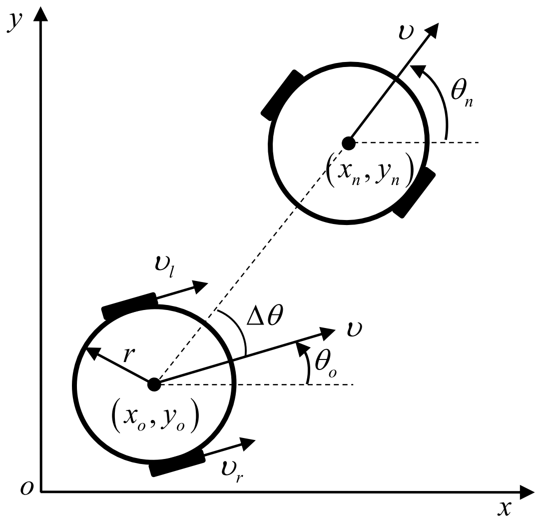


FIGURE 1. The schematic diagram of the mobile robot.

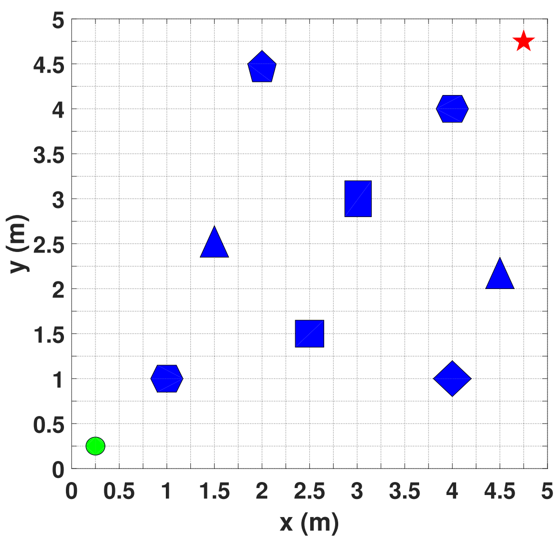


FIGURE 2. The simulation environment.

B. THE SIMULATION ENVIRONMENT

In simulations, the robot is situated in a rectangular region with the size of 5 m × 5 m as shown in Fig. 2. Obstacles scattered inside are blue, and with different shapes. The green circle represents the robot and the red star is the target. The robot's task is to find an effective path to the target point without collision all by its own learning.

IV. SENSORIMOTOR DEVELOPMENTAL LEARNING METHOD

The sensorimotor developmental learning method that designed aiming at mobile robots with the task of path planning works as bellow.

Step 1: The sensorimotor mappings definition.

Before learning, robot's sensorimotor mappings need to be defined firstly. At every learning time, the robot will sense its current situation S_i, including its position situation $S_{\text{pos}}(j)$ in the environment as well as its heading angle situation $S_{\text{ang}}(k)$. In the simulations of this article, the environment in Fig. 2 is divided into 20 equal parts along both of the horizontal axis x and vertical axis y. So the number of position situation S_{pos} is $N_{\text{pos}} = 20 \times 20 = 400$. For turning, we define the robot's steering angle space which is also its motion space as $M = \{-60°, -30°, 0°, 30°, 60°\}$. So the robot's heading angle situation S_{ang} can be divided every 30°, and the number of S_{ang} is $N_{\text{ang}} = 360°/30° = 12$. Finally the total number of sensory situations S for the robot is $N_S = N_{\text{pos}} \times N_{\text{ang}} = 4800$, and the number of the robot's motion space is $N_M = 5$.

In most autonomous learning methods, the sensorimotor mappings are one-to-one corresponding relations as in [9], and there is connection between every sensory situation and every motion. If in this way, the number of the sensorimotor mappings for the robot above will be $N_S \times N_M = 24000$. At every learning time, the robot in sensory situation S_i needs to update the connection between S_i and all the motions in M, which actually causes a lot of waste in both of calculating and learning, especially when one motion for the current situation has been proved to be unexpected. So in this article, we design the developmental learning method, which can effectively avoid the problem above, and greatly saves the calculating and learning cost. At the same time, we will find that its learning process is more biomimetic.

In this developmental learning methods, corresponding to $S_i(i = 1, 2, \cdots, N_S)$, M_i, $\bar{M}_i$ and RM_i are defined. The motions that the robot has explored from M will be placed either in M_i or $\bar{M}_i$, while the motions that the robot has not learnt will be in RM_i. The motions in M_i are the motions that have been proved effective for the situation S_i, and can lead the robot to the target. The motions in $\bar{M}_i$ are the motions that will lead the robot far from the destination, or some motions that not good enough although they also can lead the robot to the target. Obviously, we have:

$$M_i + \bar{M}_i + RM_i = M. \tag{3}$$

If the sizes of M_i, $\bar{M}_i$ and RM_i are n_i, $n_{\bar{i}}$ and n_{R_i} respectively, then

$$n_i + n_{\bar{i}} + n_{R_i} = N_M. \tag{4}$$

At one learning time t, the robot must be in one situation, supposing to be S_i, then the robot may: ① explore RM_i, or ② exploit M_i. While ①, if the motion m that the robot explores at time t is proved effective for S_i, then at time $t+1$, m will be moved from RM_i to M_i, or m will be moved to $\bar{M}_i$.

The sensorimotor mappings we defined in this article are between S_i and M_i as in Fig. 3 instead of between S_i and M. And between every S_i and M_i, we define the orientation vector $O_i(t) = [o_{i1}(t), o_{i2}(t), \cdots, o_{in_i}(t)]$, in which $o_{ij}(t)(i = 1, 2, \cdots, N_S, j = 1, 2, \cdots, n_i)$ represents the robot's preference in sensory situation S_i to choose the motion m_{ij} in M_i at the learning time t. Although all the motions in M_i can lead the robot to the target, we hope that the robot can learn a better one. So while ②, O_i will update according to the result at time $t+1$. Here, if some o_{ij} is lower than the threshold value, then

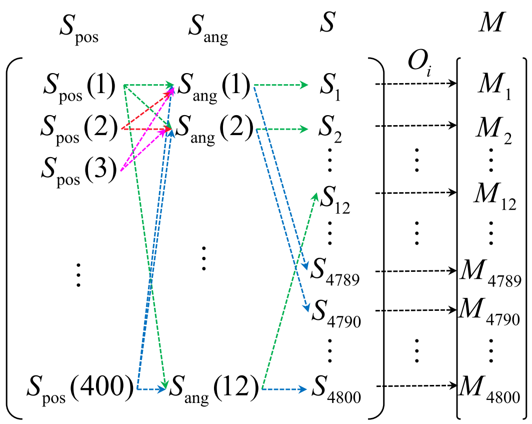


FIGURE 3. **The sensorimotor mappings for the developmental learning method.**

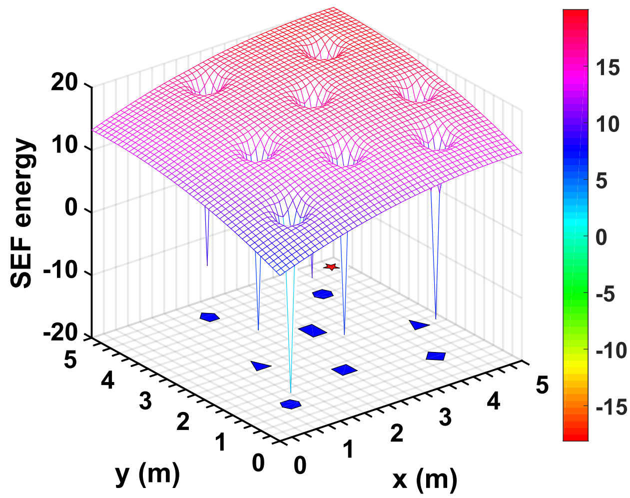


FIGURE 4. **The *SEF* energy diagram for the environment in Fig. 2.**

move the related motion m_{ij} from M_i to $\bar{M}_i$. We can find that M is fixed, but M_i can develop. O_i here is what the robot needs to learn, and for each O_i at any learning time t, $0 \leq o_{ij}(t) \leq 1$ and $\sum_{j=1}^{n_i} o_{ij}(t) = 1$.

For more efficient learning, the curiosity vector C_i is introduced to avoid the local optimal strategy problem which easily happens only based on orientation. Curiosity is a concept from neuropsychology, and there is already a lot of works proving its important role in biological learning. As the same as O_i, in this developmental learning method, $C_i(t) = [c_{i1}(t), c_{i2}(t), \cdots, c_{in_i}(t)]$ is also defined between S_i and M_i, and $c_{ij}(t)$ means the degree of the robot's curiosity to explore m_{ij} in M_i while the robot is in sensory situation S_i at the learning time t. As the number n_{ij} that the motion m_{ij} is learned increases, the robot's curiosity for m_{ij} will decrease. So the curiosity function is defined as:

$$c_{ij}(t) = \frac{1}{1 + e^{c_1 \times (n_{ij} - c_2)}}, \tag{5}$$

where the curiosity function parameters are set as $c_1 = 0.5$, $c_2 = 1$.

Step 2: State evaluation function definition.

For the orientations' updating, we define a function named *SEF* to evaluate the robot's sensory situation. In mobile robots' navigation task, $SEF(t)$ is with higher value while the robot is closer to the target, and with lower value while the robot is closer to the obstacles. Meanwhile, the robot takes precedence on obstacles avoiding, and then target navigation. With the principles above, $SEF(t)$ is defined as:

$$SEF(t) = (a_1 - a_2 \times d_{\text{g}}^2(t)) - a_3 \times e^{a_4 - a_5 \times d_{\text{o}}(t)}, \tag{6}$$

where $d_{\text{g}}(t)$ means the robot's distance to the target, and $d_{\text{o}}(t) = min(d_1(t), d_2(t), \cdots, d_n(t))$ represents the minimum distance of the robot to the obstacles, while $d_i(t)(i = 1, 2, \cdots, n)$ is the robot's distance to the i^{th} obstacle, and n is the number of the obstacles. In the environment in Fig. 2, $n = 8$. The *SEF* parameters are set as $a_1 = 20$, $a_2 = 0.3, a_3 = 0.2, a_4 = 5, a_5 = 12$ respectively. The energy diagram of *SEF* is then as shown in Fig. 4.

Step 3: Motion chosen and execution.

At the learning time t, the robot estimates its sensory situation supposing to be S_i. Then for motion chosen, the robot may: ① explore RM_i or ② exploit M_i. Considering the different conditions of M_i and RM_i, three cases may happen.

Case I. M_i is empty.

This case happens while the sensory situation S_i has not been experienced by the robot, or it has been experienced but no effective motions are learnt. In such case, the robot will explore a motion in RM_i randomly and execute it.

Case II. RM_i is empty.

That RM_i is empty means that all the motions in M has been explored under the current sensory situation S_i. In this case, the robot will exploit M_i. Here orientation and curiosity will take effect. Orientation means the preference of the robot for different motions, and the robot tends to choose the motion with higher value of orientation. Curiosity means the robot's will to explore one motion, and the degree of curiosity relates to the number n_{ij} that this motion has been learned under sensory situation S_i. As n_{ij} increases, c_{ij} will decrease. At one learning time t, the robot will show its curiosity to one motion supposing to be $m_{ik} (k \in \{1, \cdots, n_i\})$ randomly, then $c_{ik}(t)$ is activated and calculated according to (5). For simplicity, we define $C'_i(t) = [c'_{i1}(t), c'_{i2}(t), \cdots, c'_{in_i}(t)]$, where $c'_{ik}(t) = c_{ik}(t)$, and for the rest motions in M_i, $c'_{ij}(t) = 0 (j = \{1, \cdots, n_i\}$, and $j \neq k)$.

For motion chosen in motion space M_i, operate function $OP_i(t)$ is defined:

$$OP_i(t) = \sigma O_i(t) + (1 - \sigma) C'_i(t). \tag{7}$$

$0 < \sigma < 1$ is the $OP(t)$ function parameter, which decides how much that orientation and curiosity will affect the motion chosen. The bigger σ is, the greater influence orientation will make. Thus the robot will experience short exploration stage and enter exploitation stage rapidly. On the contrary, the smaller σ is, the greater influence curiosity will make. Then the robot will explore the motion space M_i more to get better choice, but the convergence procedure then will be slow. For the environment in Fig. 2, σ is set as 0.3. The balance of exploration and exploitation is an important problem faced by all intelligent learning methods. *OP* function here can achieve smooth as well as intelligent transition between exploration and exploitation stage, which improves the robot's learning property in some aspects. The motion chosen strategy in case II is winner-take-all, which means the robot will choose the motion with the biggest value of $op(t)$.

Case III. both of M_i and RM_i are not empty.

In this case, the robot may: ① explore RM_i space with probability $\frac{n_{R_i}}{n_i+n_{R_i}}$, or ② exploit M_i space with probability $\frac{n_i}{n_i+n_{R_i}}$. If ①, the robot will act as in case I, while ②, act as in case II.

Here it is easily proved, when M_i is empty, $n_i = 0$, the robot will explore the RM_i space with probability 1. The same, when RM_i is empty, $n_{R_i} = 0$, the robot will definitely exploit the M_i space. On the other hand, it can be found that more motions have not been explored, with greater probability the robot will explore the RM_i space.

Step 4: Sensorimotor mapping evaluation.

At the learning time t, the robot executes one motion under sensory situation S_i with $SEF(t)$, then at the learning time $t+1$, the robot will place itself in a new sensory situation with $SEF(t+1)$. Here we define the sensorimotor mapping evaluation function *SMF* as in (8) to evaluate whether this motion is good or not under S_i, and at the same time to realize the updating of O_i.

$$SMF(t+1) = SEF(t+1) - SEF(t). \tag{8}$$

If $SMF(t+1) \geq 0$, it means the motion that the robot executed at time t will lead the robot to a better sensory situation, in other words, this motion is an effective one for S_i. On the contrary, if $SMF(t+1) < 0$, this motion should be abandoned under S_i.

Step 5: Developmental learning.

Cases I-III in step 3 are stated separately in this part to more clearly clarify the developmental learning process. At time t, the robot is supposed to be in sensory situation S_i.

In **case I**, the robot chooses a motion m from RM_i. At the learning time $t+1$, if $SMF(t+1) < 0$, move m to $\bar{M}_i$, if $SMF(t+1) \geq 0$, do *Extension Development* as step 5.1.1 to step 5.1.4.

In **case II**, the robot chooses a motion supposed to be $m_{ik}(k \in \{1, 2, \cdots, n_i\})$ from M_i, then at time $t+1$, O_i will be updated according to (9).

$$\begin{cases} o_{ik}(t+1) = \dfrac{o_{ik}(t) + o_{ik}(t) \times (1 - e^{-\eta \times SMF(t+1)})}{1 + o_{ik}(t) \times (1 - e^{-\eta \times SMF(t+1)})}, \\ o_{ik'}(t+1) = \dfrac{o_{ik'}(t)}{1 + o_{ik}(t) \times (1 - e^{-\eta \times SMF(t+1)})}. \end{cases} \tag{9}$$

$0 < \eta < 1$ is the orientation parameter. In this article $\eta = 0.05$. At the same time, after updating O_i at the learning time $t+1$, if $o_{ib}(t+1) < \frac{1}{3 \times n_i}(b = \{1, 2, \cdots, n_i\})$, then do *Reduction Development* as step 5.2.1 to step 5.2.4.

In **case III**, case I and case II will happen in probability. If any case happens, just do the same as stated above.

The details of *Extension Development* and *Reduction Development* are as follows.

A. EXTENSION DEVELOPMENT

For extension development, a new motion m in RM_i will be moved to M_i. At the learning time $t+1$, M_i as well as its properties such as O_i and C_i are needed to be updated.

Step 5.1.1: extend m to be the $n_i + 1$ motion of M_i.

$$M_i = [M_i, m]. \tag{10}$$

Step 5.1.2: update the orientations of M_i.

$$\begin{cases} o_{i(n_i+1)}(t+1) = \dfrac{1}{n_i+1}, \\ o_{ij}(t+1) = \dfrac{n_i}{n_i+1} o_{ij}(t), j = \{1, \cdots, n_i\}. \end{cases} \tag{11}$$

Step 5.1.3: active the curiosity of the new motion $m_{i(n_i+1)}$.

$$\begin{cases} n_{i(n_i+1)} = 1, \\ c_{i(n_i+1)} = \dfrac{1}{1 + e^{c_1 \times (n_{i(n_i+1)} - c_2)}}. \end{cases} \tag{12}$$

Step 5.1.4: update the size of M_i.

$$n_i = n_i + 1. \tag{13}$$

B. REDUCTION DEVELOPMENT

For reduction development, a motion m_{ib} $(b \in \{1, 2, \cdots, n_i\})$ will be moved out of M_i. At the learning time $t+1$, M_i as well as its properties are updated as follows.

Step 5.2.1: update the orientations of M_i.

$$o_{ij}(t+1) = \frac{o_{ij}(t)}{1 - o_{ib}(t)}, (j = \{1, 2, \cdots, n_i\}, \quad and\ j \neq b). \tag{14}$$

Step 5.2.2: move m_{ib} out of M_i.

$$M_i = [m_{i1}, \cdots, m_{i(b-1)}, m_{i(b+1)}, \cdots, m_{in_i}]. \tag{15}$$

For $b = 1$, $M_i = [m_{i2}, \cdots, m_{in_i}]$, and for $b = n_i$, $M_i = [m_{i1}, \cdots, m_{i(n_i-1)}]$.

Step 5.2.3: update M_i.

$$\begin{cases} m_{ij} \leftarrow m_{ij}, & 1 < j < b; \\ m_{ij} \leftarrow m_{i(j+1)}, & b \leq j \leq n_i - 1 \end{cases} \tag{16}$$

The orientation as well as curiosity are the properties of the motions, and will change along with the motions.

Step 5.2.4: update the size of M_i.

$$n_i = n_i - 1. \tag{17}$$

Step 6: Learning termination judgement.

If the learning process fits the termination condition, then learning ends, or turns to Step 3. The flowchart of this developmental learning method can be shown in Fig. 5.

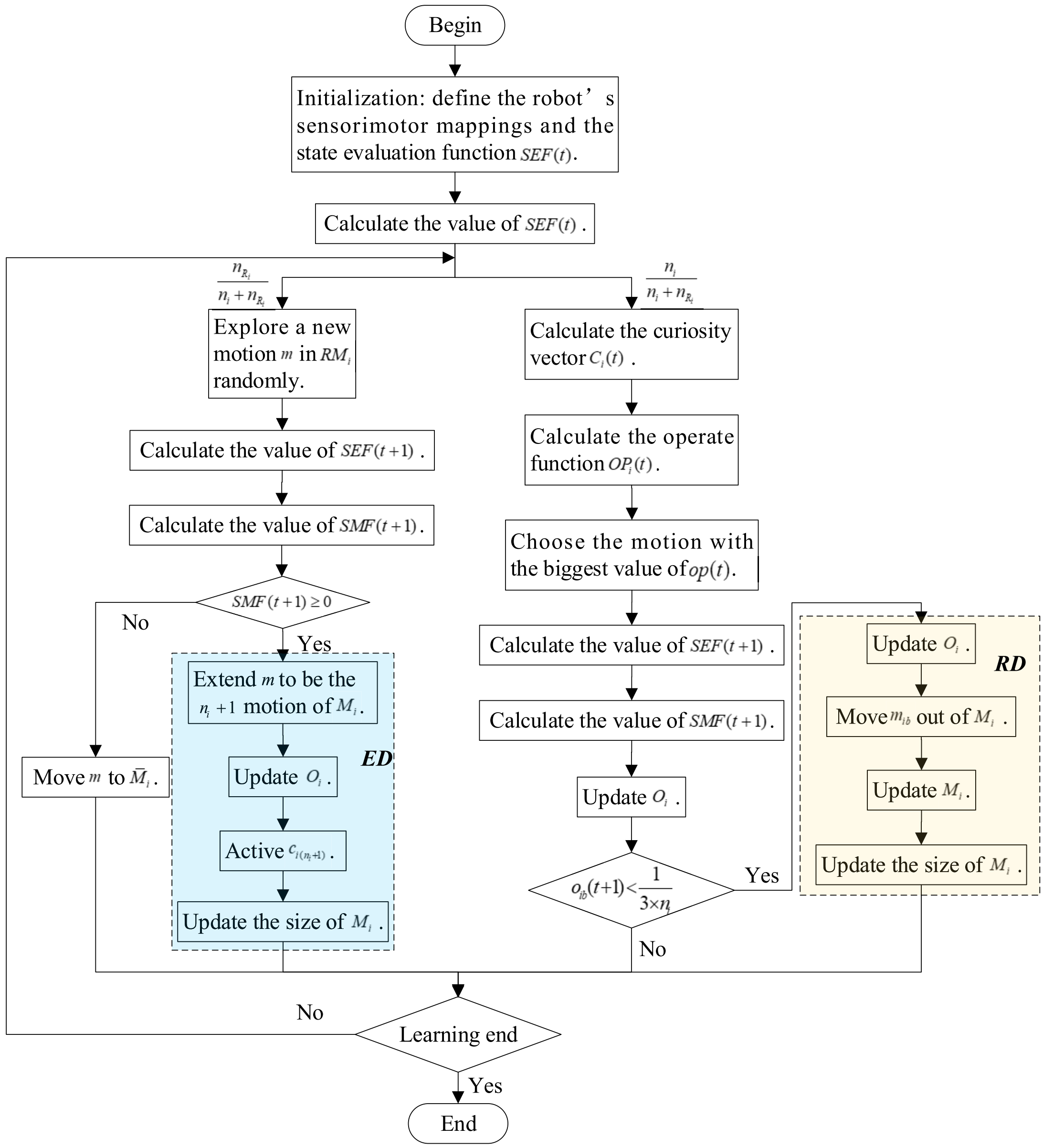


FIGURE 5. The flowchart of the developmental learning method (*ED* represents *Extension Development* and *RD* represents *Reduction Development*).

V. ALGORITHM ANALYSIS

The developmental characteristics of this method react in that the motion space M_i will change during the learning process according to *Extension Development* and *Reduction Development*. And the learning characteristics of the method react in the updating of the orientation vectors O_i. To guarantee the effectiveness of the learning, the orientations o_{ij} should satisfy condition ①: $0 \leq o_{ij} \leq 1$ as well as $\sum_{j=1}^{n_i} o_{ij} = 1$, and condition ②: updating according to operant conditioning theory during all the learning process. Here we give the proof that the method's developmental process can satisfy condition ①, and the updating of o_{ij} satisfies condition ① and ② simultaneously.

Proof I: *Extension Development* satisfies condition ①.

At one learning time, M_i gets its first motion m_{i1}. According to (11),

$$0 \leq o_{i1} = \frac{1}{0+1} = 1 \leq 1. \tag{18}$$

For O_i, supposing at the learning time t,

$$\begin{cases} 0 \leq o_{ij}(t) \leq 1, \quad (j = \{1, 2, \cdots, n_i\}), \\ \sum_{j=1}^{n_i} o_{ij}(t) = 1. \end{cases} \tag{19}$$

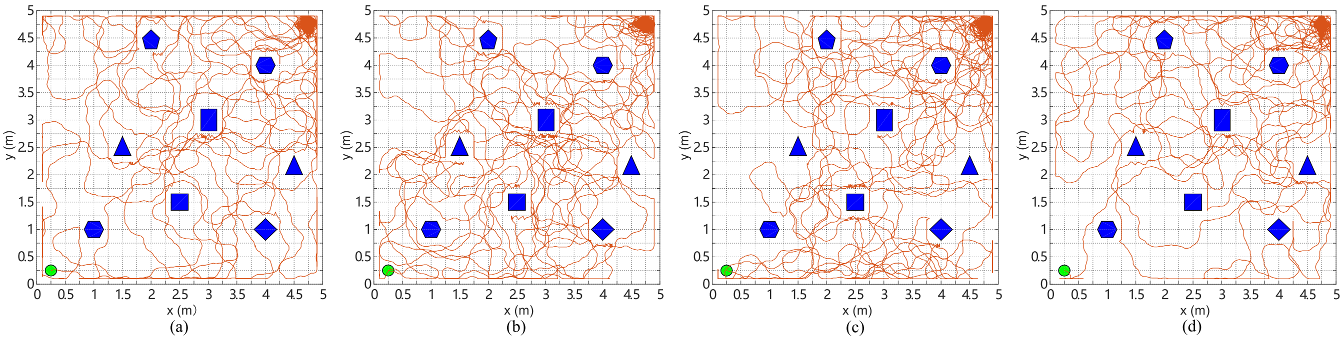


FIGURE 6. **Free learning processes.**

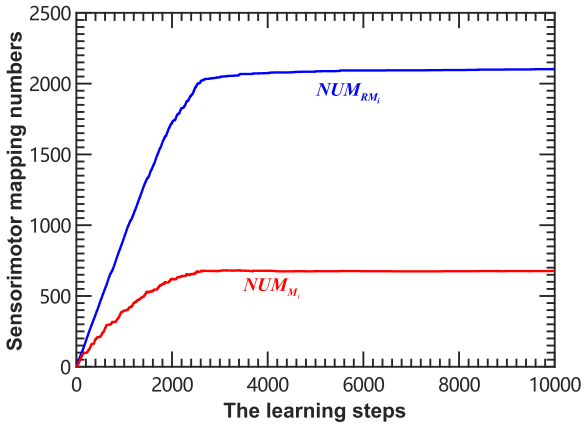


FIGURE 7. **The sensorimotor mapping numbers during the free learning process as in Fig. 6(a).**

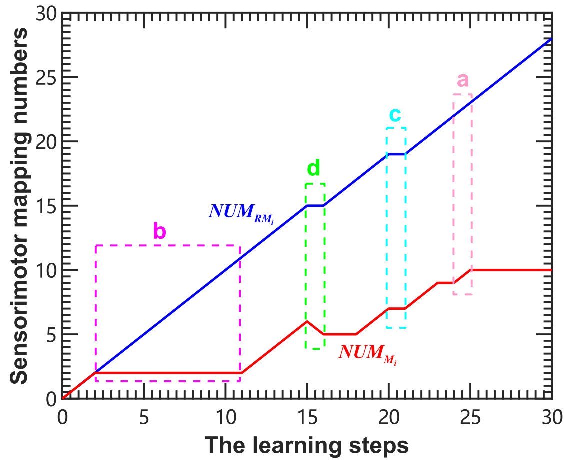


FIGURE 8. **The sensorimotor mapping numbers from step 1 to step 30 of the free learning process as in Fig. 6(a).**

Then at the learning time $t+1$, one new motion is added, the orientations are:

$$\begin{cases} 0 \le o_{i(n_i+1)}(t+1) = \dfrac{1}{n_i+1} \le 1, \\ 0 \le o_{ij}(t+1) = \dfrac{n_i}{n_i+1} o_{ij}(t) \le 1, \quad (j = \{1, \cdots, n_i\}), \end{cases} \tag{20}$$

and

$$\begin{aligned} \sum_{j=1}^{n_i+1} o_{ij}(t+1) &= o_{i(n_i+1)}(t+1) + \sum_{j=1}^{n_i} o_{ij}(t+1) \\ &= \frac{1}{n_i+1} + \frac{n_i}{n_i+1} \sum_{j=1}^{n_i} o_{ij}(t) \\ &= \frac{1}{n_i+1} + \frac{n_i}{n_i+1} \\ &= 1. \end{aligned} \tag{21}$$

By mathematical induction, *Extension Development* always satisfies condition ①.

Proof II: *Reduction Development* satisfies condition ②.

Before every time of *Reduction Development*, O_i is built up by *Extension Development*. According to Proof I, condition ① is satisfied at learning time t, then at the learning time $t+1$,

$$0 \le o_{ij}(t+1) = \frac{o_{ij}(t)}{1-o_{ib}(t)} = \frac{o_{ij}(t)}{o_{ij}(t)+(1-o_{ib}(t)-o_{ij}(t))} \le 1, \tag{22}$$

and

$$\begin{aligned} &\sum_{j=1}^{b-1} o_{ij}(t+1) + \sum_{j=b+1}^{n_i} o_{ij}(t+1) \\ &\quad = \sum_{j=1}^{b-1} \frac{o_{ij}(t)}{1-o_{ib}(t)} + \sum_{j=b+1}^{n_i} \frac{o_{ij}(t)}{1-o_{ib}(t)} \\ &\quad = \frac{1-o_{ib}(t)}{1-o_{ib}(t)} \\ &\quad = 1. \end{aligned} \tag{23}$$

By mathematical induction, *Reduction Development* always satisfies condition ①.

Proof III: the updating of o_{ij} satisfies condition ①.

M_i is built up by *Extension Development*, so before O_i updates according to (9), o_{ij} satisfies condition ①. Supposing at the learning time t, the robot chooses one motion m_{ik} from M_i, then at time $t+1$, O_i will update according to (9). As the

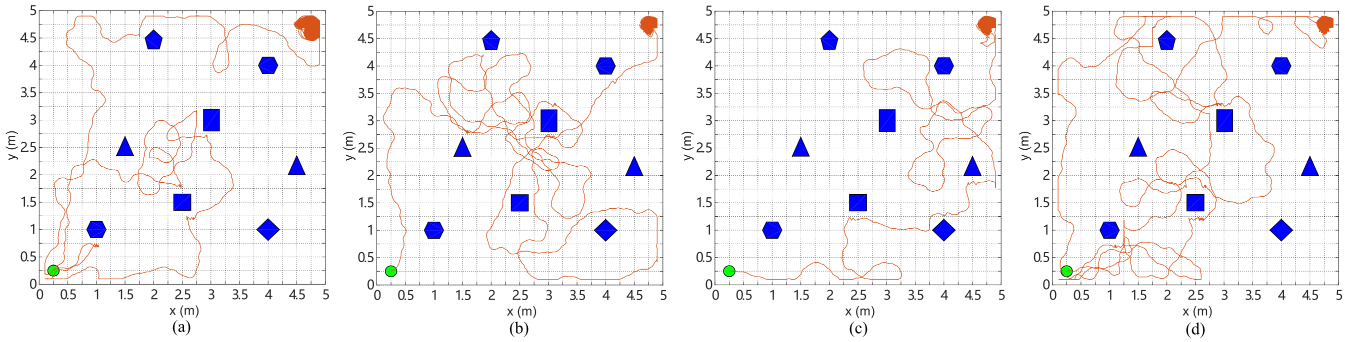


FIGURE 9. The 2nd round of free learnings.

motion m_{ik} is an effective motion, so $SMF(t+1) \geq 0$. While $0 < \eta < 1$, we have $0 < \Delta = (1 - e^{-\eta \times SMF(t+1)}) \leq 1$. Then,

$$\begin{cases} 0 \leq o_{ik}(t+1) = \dfrac{o_{ik}(t) + o_{ik}(t) \times \Delta}{1 + o_{ik}(t) \times \Delta} \leq 1, \\ 0 \leq o_{ik'}(t+1) = \dfrac{o_{ik'}(t)}{1 + o_{ik}(t) \times \Delta} \leq 1, \end{cases} \tag{24}$$

and

$$\begin{aligned} \sum_{j=1}^{n_i} o_{ij}(t+1) &= o_{ik}(t+1) + \sum_{j=1, j\neq k}^{n_i} o_{ij}(t+1) \\ &= \frac{o_{ik}(t) + o_{ik}(t) \times \Delta}{1 + o_{ik}(t) \times \Delta} + \sum_{j=1, j\neq k}^{n_i} \frac{o_{ik'}(t)}{1 + o_{ik}(t) \times \Delta} \\ &= \frac{o_{ik}(t) + \sum_{j=1, j\neq k}^{n_i} o_{ik'}(t) + o_{ik}(t) \times \Delta}{1 + o_{ik}(t) \times \Delta} \\ &= \frac{1 + o_{ik}(t) \times \Delta}{1 + o_{ik}(t) \times \Delta} \\ &= 1. \end{aligned} \tag{25}$$

By mathematical induction, the updating of o_{ij} always satisfies condition ①.

Proof IV: the updating of o_{ij} satisfies condition ②.

Operant conditioning has two meanings: (I) rewards help improve orientations, and (II) greater reward gets greater improvement in orientations.

For the updating of O_i, if one motion m_{ik} is chosen, (I) according to (24),

$$\begin{aligned} o_{ik}&(t+1) - o_{ik}(t) \\ &= \frac{o_{ik}(t) + o_{ik}(t) \times \Delta}{1 + o_{ik}(t) \times \Delta} - o_{ik}(t) \\ &= \frac{o_{ik}(t) \times \Delta \times (1 - o_{ik}(t))}{1 + o_{ik}(t) \times \Delta} \geq 0. \end{aligned} \tag{26}$$

So $o_{ik}(t+1) \geq o_{ik}(t)$, which means the orientation for m_{ik} will improve when getting a reward $SEF(t+1) \geq 0$.

(II) Based on (26),

$$\begin{aligned} &\frac{d(o_{ik}(t+1) - o_{ik}(t))}{d(SMF(t+1))} \\ &= o_{ik}(t) \times (1 - o_{ik}(t)) \times \frac{\eta e^{-\eta \times SMF(t+1)}}{(1 + o_{ik}(t) \times \Delta)^2} \geq 0. \end{aligned} \tag{27}$$

So the difference between $o_{ik}(t+1)$ and $o_{ik}(t)$ is proportional to $SMF(t+1)$, which means greater reward will get greater improvement in orientations.

VI. SIMULATION EXPERIMENTS

Here we give the simulation results as well as analysis to further elaborate the characteristics of the developmental learning method.

For most learning based methods, they always make a lot of trials and errors, which in practice will cause much waste in learning as well as calculation. Here learning waste means the repeated explorations of those ineffective sensorimotor mappings, and calculation waste happens for O_i updating when the sensorimotor mapping that ineffective or no longer expected is chosen. But for the method in this article, the robot extends only the effective motions to M_i, and at the same time deducts the motions that no longer expected out of M_i so as to improve the learning efficiency. In such way, the learning waste will happen only once for each ineffective sensorimotor mapping when it is explored from RM_i.

A. FREE LEARNING

For animals or humans, when they are in an unknown environment, they will wander around to familiarize themselves with the environment so as to finally fulfill the task. Free learning here means that the robot moves freely in the environment without collision, which is for the robot fully cognizing the environment. In free learning phase, the learning termination condition is the learning steps. Here we set as 10000. When the robot learn 10000 steps, then learning ends. Fig. 6 shows 4 different learning processes. Firstly it can be found that the paths that the robot moves are different. This is just the nature of biological learning. As for different learners, their learning "habits" are different. What's more, different learning processes will lead to different final paths. This will be shown in later *Path Learning* simulation experiments. Secondly we can see that after the robot learns to a degree, it will keep moving around the target point. This is because that the robot has learned the knowledge of the target point, and prefers to stay around to get rewards until achieving the learning steps that have been set.

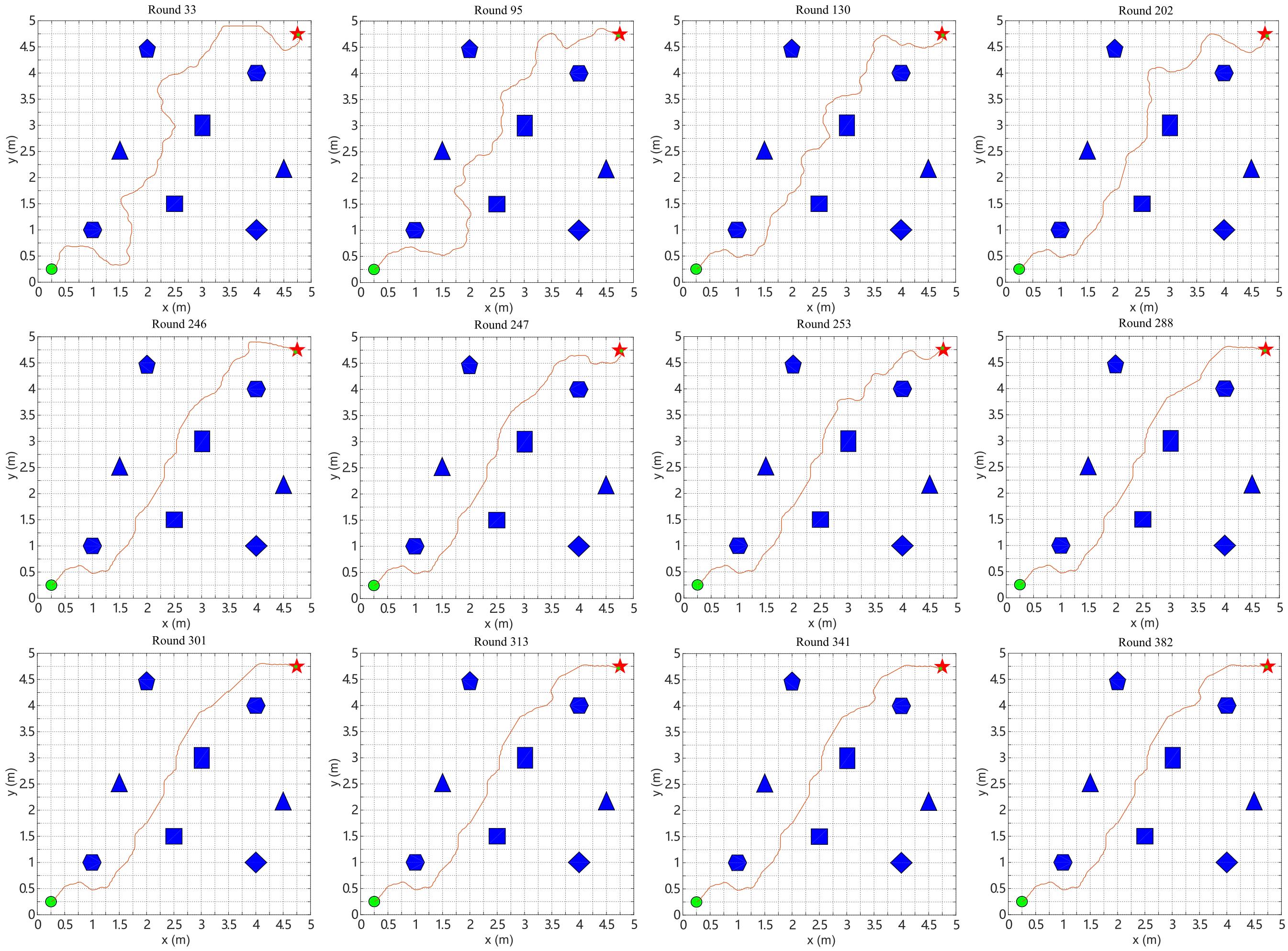


FIGURE 10. Path optimization process.

Taking the learning process of Fig. 6(a) as an example, we give the descriptions of the two developmental modes in this method. As shown in Fig. 7, the blue line is the total sensorimotor mapping number that the robot has explored from RM_i at every learning step, here we define it as NUM_{RM_i}. The red line is the sensorimotor mapping number that the robot has developed in M_i at each learning step, and is defined as NUM_{M_i}. From Fig. 7, it can be found that as the robot explores RM_i more and more, more effective motions will be developed to M_i. What's more, the blue line is always higher than the red line, which means during the exploration of RM_i, there exit many ineffective motions which are moved to $\bar{M}_i$. The difference between the blue line and the red line is that the blue line will not decrease all the time while the red line may decrease sometime because of *Reduction Development*.

To more elaborately clarify the two developmental modes, we take out the data from step 1 to step 30 as shown in Fig. 8. Four representative periods are chosen to explain the developmental learning process. In the rectangle a, we can see NUM_{RM_i} increases 1 from the learning step 24 to 25, and at the same time NUM_{M_i} also increases 1. This process means

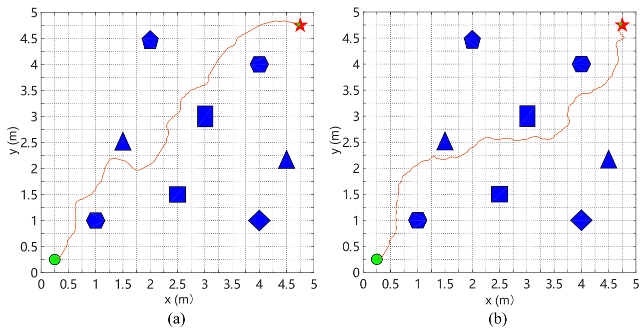


FIGURE 11. Different path learning results.

that at the learning step 24, the robot explores a new motion from RM_i, and at the learning step 25, this motion is proved to be an effective motion for the current sensory situation, so it is developed to M_i. In the rectangle b, we can see NUM_{RM_i} keeps increasing from the learning step 2 to 11, but NUM_{M_i} always stays the same. This means that during this process, although the robot explores RM_i at every learning step, all the motions are ineffective, so M_i is not developed. In rectangle

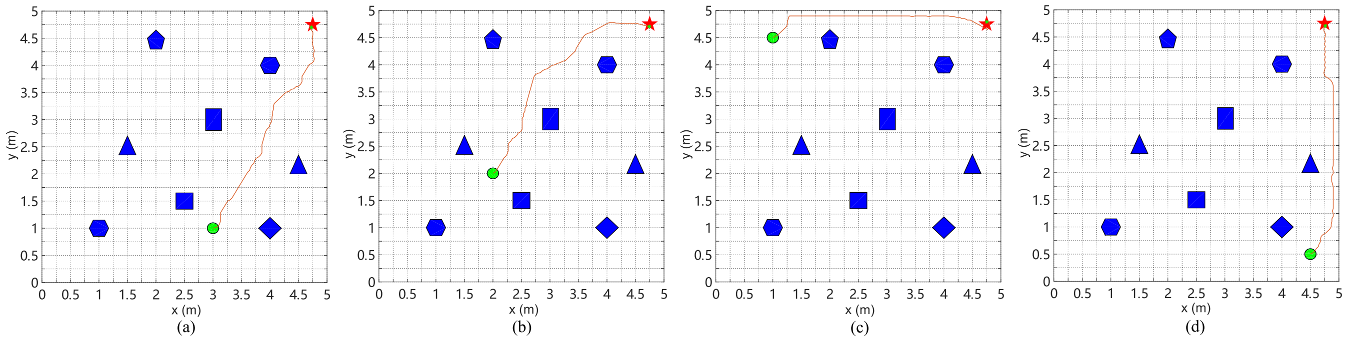


FIGURE 12. Learning ability test results.

c and d, we can see NUM_{RM_i} does not increase, which means the robot does not explore RM_i during this period, so it must exploit M_i and update O_i. The difference between rectangle c and d is that, in c, NUM_{M_i} does not change, which means the values of all the o_{ij} at next step are bigger than the threshold value. But in d, after updating, one motion's orientation value is lower than the threshold value, so the robot does reduction development and moves the motion out of M_i.

B. ROUNDS OF LEARNING

Rounds of learning means that the robot will begin a new round of learning from the start point based on the knowledge it has learnt in the previous round. From Fig. 7 we find that after about the 3000th learning step, the sensorimotor mapping numbers hardly change, that is because the robot basically stays around the target point then, and only learns limited sensorimotor mappings. To let the robot fully learn the environment so as to find more reasonable path, here we let the robot do "rounds of learning." In the phase, the learning termination condition is still the learning steps. Based on the 1st round of learning in Fig. 6, Fig. 9 shows their corresponding 2nd round of learning results. It can be found that, compared to the 1st round of learnings, in the 2nd round, the robot's wandering behaviors greatly reduce. Once the robot arrives at the target point, it will stay around. This in some aspects illustrates the learning stability of this developmental learning method, and this stability comes from the design of the effective motion space M_i. As in M_i, all the motions will lead the robot to the target. Once the robot finishes the exploration of space RM_i, all it does is exploiting M_i. In this way, the robot's exploration to those ineffective sensorimotor mappings will be only once, which greatly reduces the learning waste as well as computing waste.

C. PATH LEARNING

In Fig. 9, we can find that, compared to the 1st round of learnings, in the 2nd round of learnings, the robot experiences fewer steps to get to the target point. This proves that the robot does learn some knowledge. From the 3rd round on, we let the robot do "path learning." In "path learning" simulation experiment, robot's learning in every round is also based on the previous round, but the termination condition goes from "reach the learning steps that have been set" to "the robot reaches the target point". Fig. 10 records some representative rounds of path optimization process. The exact steps that the robot uses in each round of learning are as in TABLE 1. We can see that as the learning goes on, the steps that the robot uses are basically going down. That means the learnt path is better and better. Between round 247 and round 301, we find the steps sometimes increase, which is because the curiosity still plays a role, so the robot may explore some new sensorimotor mappings. From the round 341 on, the steps are always 149, and corresponding to Fig. 10, the paths are all the same since the round 341. This indicates that the robot has learned a stable path.

TABLE 1. Steps of Different Round of Learnings.

Round	33	95	130	202	246	247	253	288	301	313	341	382
Step	194	177	165	164	153	151	154	152	155	150	149	149

What's worth mentioning is that the learning results in Fig. 10 are just to show the robot's learning process. For the robot, its path learning result is not fixed, and may be different according to the robot's learning "habit". Fig. 11 are another two path learning results.

D. LEARNING ABILITY TEST

To show the learning ability of this developmental learning method, after the robot's learning as in Fig. 10, we change its start position as in Fig. 12. The test results show that after learning the environment, even though the robot is placed at new different positions, it can still find useful paths with just a few rounds of learning.

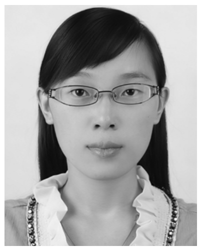
VII. CONCLUSION

In this article, a new developmental learning method is proposed to help mobile robots fulfill the path planning task during its sensorimotor process in the environment. In this method, orientation and curiosity are introduced from the perspective of cognitive psychology. Orientation decides the robot's preference of different motions, and helps the robot enhance the knowledge that it has learned, while the curiosity can promote the robot to explore new knowledge. The design of orientation and curiosity well balances the exploration and exploitation which is faced by all the intelligent learning methods. What's more, the sensorimotor mappings in this method are not fixed. The robot only exploits effective sensorimotor mappings which it develops to M_i. For the ineffective sensorimotor mappings, the explorations for them are at most once. This greatly reduces the learning waste as well as the calculation waste. The analysis of the algorithms of this method's every part is given. Simulation experiments are also done to show the method's properties of learning as well as development. In this method, the parameters of the *SEF* function are chosen based on the environment in Fig. 2. For the environments in which the obstacles as well as the target are with similar distribution, these parameters will also work. But for total different environments, the parameters of *SEF* are usually different. To enable the robot to adapt to dynamic environments, an adaptive *SEF* function will be designed in the following work.

This article introduces behavior mechanism and intrinsic motivation mechanism simultaneously in artificial learning method designing, and provides a way to realize autonomous robots. Although this development learning method is designed for mobile robots with path planning task, it can also guide other types of robots with specific tasks. What needs to do then is redesigning the *SEF* function and redefining the sensory situation spaces as well as motion spaces. So this developmental learning method is broadly applicable.

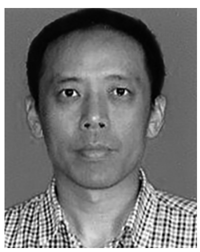
REFERENCES

[1] J. Piaget, *The Origins of Intelligence in Child*, New York, NY, USA: International Universities Press, 1952.
[2] D. J. Lewkowicz and R. Lickliter, "A dynamic systems approach to the development of cognition and action," *J. Cognit. Neurosci.*, vol. 7, no. 4, pp. 512–514, 1995.
[3] P.-Y. Oudeyer, "What is intrinsic motivation? A typology of computational approaches," *Frontiers Neurorobotics*, vol. 1, p. 6, Oct. 2007.
[4] B. Webb, "Using robots to model animals: A cricket test," *Robot. Auto. Syst.*, vol. 16, nos. 2–4, pp. 117–134, Dec. 1995.
[5] Y. Kuniyoshi and S. Sangawa, "Early motor development from partially ordered neural-body dynamics: Experiments with a cortico-spinal-musculo-skeletal model," *Biol. Cybern.*, vol. 95, no. 6, pp. 589–605, Dec. 2006.
[6] A. Laflaquière, J. K. O'Regan, S. Argentieri, B. Gas, and A. V. Terekhov, "Learning agent's spatial configuration from sensorimotor invariants," *Robot. Auto. Syst.*, vol. 71, pp. 49–59, Sep. 2015.
[7] G. Pugach, A. Pitti, O. Tolochko, and P. Gaussier, "Brain-inspired coding of robot body schema through visuo-motor integration of touched events," *Frontiers Neurorobotics*, vol. 13, p. 5, Mar. 2019.
[8] B. F. Skinner, *The Behaviour of Organisms: An Experimental Analysis*. New York, NY, USA: Appleton-Century, 1938.
[9] K. Itoh, H. Miwa, M. Matsumoto, M. Zecca, H. Takanobu, S. Roccella, M. C. Carrozza, P. Dario, and A. Takanishi, "Behavior model of humanoid robots based on operant conditioning," in *Proc. 5th IEEE-RAS Int. Conf. Hum. Robot.*, Tsukuba, Japan, May 2005, pp. 220–225.
[10] X. Zhang, X. Ruan, Y. Xiao, and J. Huang, "Sensorimotor self-learning model based on operant conditioning for two-wheeled robot," *J. Shanghai Jiaotong Univ.*, vol. 22, no. 2, pp. 148–155, Apr. 2017.
[11] P. Arena, L. Patane, D. Sanalitro, and A. Vitanza, "Insect-inspired body size learning model on a humanoid robot," in *Proc. 7th IEEE Int. Conf. Biomed. Robot. Biomechatronics (Biorob)*, Enschede, The Netherlands, Aug. 2018, pp. 1127–1132.
[12] A. Cyr, M. Boukadoum, and F. Theriault, "Operant conditioning: A minimal components requirement in artificial spiking neurons designed for bio-inspired robot's controller," *Frontiers Neurorobotics*, vol. 8, p. 151, Oct. 2014.
[13] A. Cyr and F. Thériault, "Spatial concept learning: A spiking neural network implementation in virtual and physical robots," *Comput. Intell. Neurosci.*, vol. 2019, pp. 1–8, Apr. 2019.
[14] A. G. Barto, S. Singh, and N. Chentanez, "Intrinsically motivated learning of hierarchical collections of skills," *Int. Conf. Develop. Learn. (ICDL),* La Jolla, CA, USA, 2004, pp. 112–119.
[15] P.-Y. Oudeyer, F. Kaplan, and V. V. Hafner, "Intrinsic motivation systems for autonomous mental development," *IEEE Trans. Evol. Comput.*, vol. 11, no. 2, pp. 265–286, Apr. 2007.
[16] S. M. Nguyen and P.-Y. Oudeyer, "Socially guided intrinsic motivation for robot learning of motor skills," *Auto. Robots*, vol. 36, no. 3, pp. 273–294, Mar. 2014.
[17] C. Colas, P. Fournier, O. Sigaud, and P.-Y. Oudeyer, "CURIOUS: Intrinsically motivated modular multi-goal reinforcement learning," in *Proc. Int. Conf. Mach. Learn.*, Long Beach, CA, USA, 2019, pp. 1331–1340.
[18] H. Ren, C. Liu, and T. Shi, "A computational model of cognitive development for the motor skill learning from curiosity," *Biologically Inspired Cognit. Archit.*, vol. 25, pp. 101–106, Aug. 2018.
[19] A. Sadeghi, A. Mondini, E. Del Dottore, V. Mattoli, L. Beccai, S. Taccola, C. Lucarotti, M. Totaro, and B. Mazzolai, "A plant-inspired robot with soft differential bending capabilities," *Bioinspiration Biomimetics*, vol. 12, no. 1, Dec. 2016, Art. no. 015001.
[20] E. Del Dottore, A. Mondini, A. Sadeghi, and B. Mazzolai, "Swarming behavior emerging from the Uptake–Kinetics feedback control in a plant-root-inspired robot," *Appl. Sci.*, vol. 8, no. 1, p. 47, Jan. 2018.
[21] P. Sequeira, F. S. Melo, and A. Paiva, "Emotion-based intrinsic motivation for reinforcement learning agents," in *Proc. Int. Conf. Affect. Comput. Intell. Interact.* Memphis, TN, USA, 2011, pp. 326–336.
[22] Z. Mathews, S. B. i Badia, and P. F. M. J. Verschure, "PASAR: An integrated model of prediction, anticipation, sensation, attention and response for artificial sensorimotor systems," *Inf. Sci.*, vol. 186, no. 1, pp. 1–19, Mar. 2012.



XIAOPING ZHANG was born in Shanxi, China, in 1991. She received the B.S. degree from the School of Automation, Xi'an University of Posts and Telecommunications, in 2011, and the Ph.D. degree from the Faculty of Information Technology, Beijing University of Technology, in 2018.

From 2015 to 2017, she was a Visiting Scholar with Michigan State University, USA. Since 2018, she has been a Teacher with the School of Electrical and Control Engineering, North China University of Technology. She is also the Project Leader of the National Natural Science Foundation of China and the Beijing Natural Science Foundation, China. Her research interests include intelligent robot, machine learning, path planning, and cognitive learning. She has deep understanding and experience in cognitive robot and artificial intelligence. She has strong ability in model design and algorithm analysis.



XIAOGANG RUAN was born in Sichuan, China, in 1958. He received the Ph.D. degree from Zhejiang University, Hangzhou, China, in 1992. He is currently a Professor with the Beijing University of Technology. His research interests include automatic control, artificial intelligence, and intelligent robot.

HONG ZHANG received the Ph.D. degree in pattern recognition and intelligent system from Xidian University, Xi'an, China, in 2015. She is currently an Associate Professor with the Xi'an University of Posts and Telecommunications, Xi'an. Her research interests include machine learning, image processing, and pattern recognition method.

LEI LIU was born in Shanxi, China, in 1984. She received the B.S. degree from the Tianjin University of Technology, in 2006, and the M.S. and Ph.D. degrees from Tianjin University, in 2008 and 2011, respectively. From 2011 to 2013, she was a Postdoctoral Fellow with the Department of Mechanics and Aerospace Engineering, College of Engineering, Peking University. She is currently an Associate Professor with the Beijing Key Laboratory of Fieldbus Technology and Automation, School of Electrical and Control Engineering, North China University of Technology. Her research interests include singularly perturbed systems, robust control, fault diagnosis, and fault-tolerant control.

CUNWU HAN received the B.S. degree in electrical engineering from Shanghai Jiaotong University, Shanghai, China, in 1982, and the M.S. and Ph.D. degrees in electrical engineering from Northeastern University, Shenyang, China, in 1991 and 1994, respectively. He is currently a Professor with the Beijing Key Laboratory of Fieldbus Technology and Automation, North China University of Technology, Beijing, China. His research interests include intelligent control, robot, networked control systems, wireless communication networks, and the Internet of Vehicles.

LI WANG received the B.S. and M.S. degrees from Yanshan University, in 1999 and 2002, respectively, and the Ph.D. degree from Beihang University, in 2006. From 2006 to 2009, he was a Postdoctoral Fellow with Beihang University. He is currently a Professor and the Dean of the School of Electrical and Control Engineering, North China University of Technology. His research interests include intelligent control systems, traffic control, and road traffic engineering design.