Coordinated Complex-Valued Encoding Dragonfly Algorithm and Artificial Emotional Reinforcement Learning for Coordinated Secondary Voltage Control and Automatic Voltage Regulation in Multi-Generator Power Systems

LINFEI YIN¹, SHIKUI LUO¹, YAOXIONG WANG², FANG GAO¹, AND JUN YU³, (Member, IEEE)

¹College of Electrical Engineering, Guangxi University, Nanning 530004, China
²Institute of Intelligent Machines, Chinese Academy of Sciences, Hefei 230031, China
³Department of Automation, University of Science and Technology of China, Hefei 230026, China

Corresponding author: Fang Gao (fgao@gxu.edu.cn)

The work of Linfei Yin was supported by the Guangxi Natural Science Foundation under Grant AD19245001 and Grant 2020GXNSFBA159025. The work of Yaoxiong Wang and Fang Gao was supported by the National Natural Science Foundation of China under Grant 61773359 and Grant 61720106009. The work of Jun Yu was supported in part by the National Natural Science Foundation of China under Grant U1736123, and in part by the University of Science and Technology of China (USTC) Research Funds of the Double First-Class Initiative under Grant YD2350002001.

ABSTRACT This article proposes a coordinated optimization and control algorithm for coordinated secondary voltage control (CSVC) in multi-generator power systems. Firstly, to obtain a smaller voltage deviation and avoid the curse of dimensionality simultaneously, an artificial emotional reinforcement learning (AERL) is applied to automatic voltage regulation (AVR). Secondly, to obtain a smaller fitness value with lesser random for the decentralized independent variables optimization problem of the CSVC, a complex-valued encoding dragonfly algorithm (CDA) is proposed. Thirdly, the CDA and the AERL are coordinated for the CSVC and the AVR in multi-generator power systems. To verify the control performance of the AERL and the convergence of the proposed CDA, three simulation cases, i.e., IEEE 57-bus, 118-bus and 300-bus systems, are considered. The simulation results show that the CDA-AERL effectively obtains the smallest control objectives and the convergence for the CSVC in multi-generator power systems.

INDEX TERMS Coordinated secondary voltage control, artificial emotional reinforcement learning, complex-valued encoding dragonfly algorithm, automatic voltage regulation, multi-generator power systems.

I. INTRODUCTION

The conventional voltage control of power systems contains three levels, i.e., automatic voltage regulation (AVR) for real-time level, secondary voltage control (SVC) for middle-term level, and tertiary voltage control (TVC) for long-term level [1]–[3].

Generally, the AVR is controlled by a proportional-integral-derivative (PID) algorithm, which of parameters should be configured by operator experience or optimization algorithms. Numerous algorithms have been addressed to reduce the control errors of the PID, such as, modified differential evolution [4], [5], firefly algorithm [6], [7], and particle swarm optimization (PSO) [8], [9]. Nevertheless, the parameters of PID controllers should be reconfigured when the systemic parameters have changed. To obtain higher control performance for dynamic systems, reinforcement learning can be employed to a dynamic system [10]–[13]. To obtain a more accurate control strategy, more actions have been added to reinforcement learning [14]. Both the calculation memories of Q-value matrix and P-value matrix of reinforcement learning then will be increased when the number of actions

...
is increased. However, the programming of the reinforcement learning may out of memory, the curse of dimensionality then will occur [15]. As one branch of artificial intelligence, artificial emotion has been applied to increase the control performance of reinforcement learning [16], [17]. For example, the artificial emotional reinforcement learning (AERL) has been utilized in load frequency control of power systems in the previous work [18]. Consequently, the AERL can be introduced to the AVR. The major features of the AERL can be summarized as follows.

1) The agent based on the AERL consists of two parts, i.e., the artificial emotional part and reinforcement logic part.
2) Since the output actions can be updated by the artificial emotional part of the AERL, the curse of dimensionality can be mitigated and the control performance can be increased simultaneously.
3) The controller based on AERL can update the control strategy on-line for a dynamic system.

In a conventional framework of power systems, the SVC and the TVC are optimized independently [19]. The SVC in the conventional framework has one deficiency, the nodes of one control area are compacted, while the nodes of other control areas are loosed [20]. From power systems perspective, any control area should be a general node. To mitigate this deficiency, coordinated secondary voltage control (CSVC) has been applied in power systems [21]. The CSVC, which has been introduced by Paul et al. [22], aims to increase voltage stability in highly constrained areas [23]. Consequently, the CSVC can regulate the free variables of the reactive power flow [24]. Since the CSVC can be employed into system voltage control, the TVC is unnecessary for the power system [25]. Besides, since the CSVC can be applied to power grids, systemic voltage control contains only two levels, i.e., the CSVC and the AVR. Compared with the SVC, the major features of the CSVC can be summarized as follows [26], [27].

1) The CSVC aims to minimize the voltage deviation of the dominant node when the voltage stability margin of the power system is sufficient.
2) The CSVC can balance the reactive power flow of each control area and can maintain enough dynamic reactive power flow when the voltage stability margin of the power system is sufficient.
3) The CSVC aims to minimize the control cost of the power system when the voltage stability margin of the power system is insufficient.

Carbon dioxide emissions are related to all the sources of a power system. For example, carbon dioxide emissions have been considered in stochastic wind and solar power [28] and land transport infrastructure [29]; carbon dioxide emissions have been considered in natural gas and heat delivery system [30]; and carbon dioxide emissions have been considered in optimal power flow [31]. The carbon-energy combined-flow has been applied in multi-energy power systems [32], [33]. Therefore, the carbon-energy combined-flow is considered in the CSVC.

Recently, single-objective optimization problems, discrete optimization problems and multi-objective problems have been solved by a dragonfly algorithm (DA) [34]. The DA has been applied to numerous parameters optimization problems, such as the parameters optimization of support vector machine [35], numerical optimization problems [36] and improving the grey model [37]. Besides, an improved DA has been proposed for feature selection problem, which is an optimization problem with decentralized independent variables [38]. Therefore, the DA has been utilized for decentralized independent variable optimization problems. Since the CSVC optimization problem contains various decentralized independent variables, an improved DA is proposed for the optimization problem of the CSVC.

Numerous meta-heuristic optimization algorithms (e.g. grey wolf optimizer (GWO) [39]) have been proposed by Mirjalili, who proposed the DA [34]. In general, a meta-heuristic optimization algorithm can be improved by at least five types of operations: (1) more grouped operation: grouped grey wolf optimizer has been proposed for the parameters optimization of the maximum power point tracking of doubly-fed induction generator based wind turbine [40]; swarm moth–flame optimizer has been proposed for the racking of doubly-fed induction generator [41]; whale optimization algorithm has been grouped for standard benchmark functions [42]; (2) combined operation: the combined genetic algorithm (GA) and the PSO has been developed for hybrid wind-photovoltaic-battery system [43]; the GWO has been combined with a whale optimization algorithm for pressure vessel design [44] (3) adaptive parameters operation: an improved Jaya with self-adaptive weight has been applied for the parameters identification of photovoltaic models [45]; a teaching-learning-based optimization algorithm has been improved by adaptive inertia weights [46]; epsilon multi-objective genetic algorithm has been applied for PID parameters optimization [47]; (4) knowledge matrix based operation: knowledge matrix is employed to remember the optimization task [48]; transfer reinforcement learning with Q-value matrix has been proposed for reactive power optimizations [49]; a transfer matrix with Kriging model has been introduced into a multi-objective optimization algorithm [50]; (5) different coded operation: a real-coded GA has been applied into numerical optimizations [51]; binary coded GA has been employed to solve the path planning of mobile robots [52]; a binary operation has been added into the social minic optimization method [53]; hexadecimal coded optimization algorithm based on field-programmable gate-array has been applied for parallel computing [54]; complex-valued encoding operation has been employed to improve the optimal performance of the wind-driven optimization [55]. Furthermore, both more grouped operation and combined operation can obtain the global solution rather than a local solution; both adaptive parameters operation and knowledge matrix based operation can accelerate the convergence.
The voltage control of a power system contains three time slots, i.e., the AVR, the SVC, and the TVC. The AVR is often referred to as automatic voltage control, which is regulated by generators’ excitation or control algorithm. The control period of the AVR is several seconds. The SVC can be controlled by a closed-loop control algorithm, for example, PID, and reinforcement learning. The control period of the SVC sets a minute to five minutes. The TVC can be regulated by an optimal algorithm, such as GA, PSO, GWO, moth-flame optimization algorithm (MFO), and DA.

To mitigate the coupling structure of a power system with multiple control areas, the CSVC can be coordinated regulated by the AVR (FIGURE 2). For the \( i \)-th generator in FIGURE 2, \( E_{i\text{ref}} \) is the steady-state value of the excitation voltage; \( K_{Ai} \) means the gain constant of the voltage regulator; \( U_{\text{refi}} \) is the generator terminal reference voltage (standard value); \( T_{Ai} \) is the time constant of voltage regulator; \( E_{i} \) means excitation voltage; \( k_{i} \) is feedback gain, and \( k_{i} \in \{1, 100\} \); \( x_{d} \) means d-axis synchronous reactance; \( E_{qj} \) means the steady-state value of transient electromotive force; \( E_{qi} \) is transient electromotive force; \( x_{qi} \) is d-axis transient reactance; \( Q_{ei} \) means reactive power output; \( U_{li} \) is generator terminal voltage (standard value); \( V_{di} \) and \( V_{qj} \) are the d-axis and q-axis stator voltage, respectively.

The optimization objective of the CSVC that considering carbon emission flow can be described as,

\[
F = \min \left( \mu_{1} C_{\text{ds}} + \mu_{2} P_{\text{loss}} + \mu_{3} V_{d} \right)
\]

where \( \mu_{1}, \mu_{2}, \) and \( \mu_{3} \) are the weights of carbon-energy combined-flow, active power losses, and stable voltage component, respectively; and \( \mu_{1} + \mu_{2} + \mu_{3} = 1 \). The stable voltage component \( V_{d} \) is calculated as

\[
V_{d} = \sum_{j=1}^{n} \left| \frac{2V_{j} - V_{j}^{\text{max}} - V_{j}^{\text{min}}}{V_{j}^{\text{max}} - V_{j}^{\text{min}}} \right|
\]

where \( n \) is the number of generator nodes; \( V_{j}, V_{j}^{\text{max}}, \) and \( V_{j}^{\text{min}} \) are the real voltage, the maximum voltage, and the minimum voltage, respectively. The active power losses can
be presented as,

$$P_{\text{loss}} = \sum_{i,j \in \mathcal{E}_n} g_{ij} \left( V_i^2 + V_j^2 - 2V_iV_j\cos \theta_{ij} \right)$$  \hspace{1cm} (3)$$

where $n_L$ is the number of load nodes; $\theta_{ij}$ is the phase angle difference between the $i$-th and the $j$-th load node; $V_i$ and $V_j$ are the voltages of the $i$-th and the $j$-th load nodes, respectively; $g_{ij}$ is the conductance between the $i$-th and the $j$-th load nodes. The real carbon-energy combined-flow $C_{ds}$ can be calculated as,

$$C_{ds} = \sum_{i,j \in \mathcal{E}_n} \sum_{w \in \mathcal{W}} \left( A_{iw}^{-1} P_{iw} \delta_{sy} (1 - \alpha_s) \right) \Delta P_{ij}$$  \hspace{1cm} (4)$$

where $w$ means the $w$-th generator, $w \in \mathcal{W}$; $A_{iw}$ can be calculated as,

$$A_{iw} = \begin{cases} 1, & i = w \\ -P_{wi}/P_{ni}, & w \in i^+ \\ 0, & i \neq w \text{ and } w \notin i^+ \end{cases}$$  \hspace{1cm} (5)$$

where $P_{wi}$ means power flow value from the $w$-th node to the $i$-th node; $P_{wi} = 0$ when the $i$-th node and the $w$-th node are not connected; $P_{iw}$ is the power flow value from the $s$-th node to the $w$-th node; $\delta_{sy}$ means carbon emission intensity; $i^+$ means the connection of all the input bus of the $i$-th node; $P_{ni}$ means the power flow output; $\alpha_s$ and $\beta_s$ are the weights of the carbon emissions of the generator and the power grid, respectively (FIGURE 3).

All the inequality and equality constraints of the power flow equation can be described as,

$$\begin{align*}
P_{Gi} - P_{Di} - V_i \sum_{j \in \mathcal{N}_i} V_j (g_{ij}\cos \theta_{ij} + b_{ij}\sin \theta_{ij}) &= 0 \\
Q_{Gi} - Q_{Di} - V_i \sum_{j \in \mathcal{N}_i} V_j (g_{ij}\sin \theta_{ij} - b_{ij}\cos \theta_{ij}) &= 0 \\
V_{G\min} &\leq V_{Gi} \leq V_{G\max}, &i \in \mathcal{N}_G \\
Q_{G\min} &\leq Q_{Gi} \leq Q_{G\max}, &i \in \mathcal{N}_G \\
V_{L\min} &\leq V_{Li} \leq V_{L\max}, &i \in \mathcal{N}_L \\
P_{i\min} &\leq P_{Gi} \leq P_{i\max}, &i \in \mathcal{N}_G \\
Q_{i\min} &\leq Q_{Gi} \leq Q_{i\max}, &i \in \mathcal{N}_G \\
V_{i\min} &\leq V_{Li} \leq V_{i\max}, &i \in \mathcal{N}_L
\end{align*}$$  \hspace{1cm} (6)$$

where $S_i$ is the complex power value; $P_{Di}$ and $Q_{Di}$ mean the active and reactive power flow output, respectively; $P_{Gi}$ and $Q_{Gi}$ imply the active and reactive power flow input, respectively; $b_{ij}$ implies the susceptance between the $i$-th and the $j$-th load node; $N_i$ associates the node vector that connected to the $i$-th node; $N_G$, $N_B$, $N_C$, $N_k$ and $N_{L}$ are the generator, bus vectors, AVR device, on-load tap-changer and branch vectors, respectively; all the terminal voltage $V_{Gi}$, reactive power value $Q_{Gi}$, and transformer ratio $k_g$ are the variables need to be optimized by the optimization algorithm; the superscript “max” and “min” mean the maximum and the minimum values of related variables.

The CSVC aims to maintain the generator terminal voltage $U_{gi}$ to the generator terminal reference voltage (standard value) $U_{refi}$. Therefore, the input of the AVR controller is $E_{gi}$; the output of the AVR controller is voltage commands, which can be converted to sinusoidal pulse width modulation commands.

III. ARTIFICIAL EMOTIONAL REINFORCEMENT LEARNING

A. REINFORCEMENT LEARNING

A basic reinforcement learning consists of an agent and an “environment”. At the $(k+1)$-th control iteration, the environment provides reward value $R$ and system state $s_k$ to the agent; the agent then provides action $a_{k+1}$ to the environment. The agent can update its strategy at each iteration. Therefore, the state-action pair of reinforcement learning can be reinforced by these iterations, which means the agent is trained on-line. As a model-free algorithm, reinforcement learning is not based on an accurate system model.

The Q-value of reinforcement learning is reinforced as,

$$Q^{k+1}(s_k, a_k) = Q^k(s_k, a_k) + \alpha [R(s_k, s_{k+1}, a_k) + \gamma \max_{a' \in A} Q^k(s_{k+1}, a')] - Q^k(s_k, a_k)$$  \hspace{1cm} (7)$$

where $Q^k(s_k, a_k)$ is the Q-value of reinforcement learning; $s_k \in S$ is the state of the environment; $a_k \in A$ is the action of action set $A$; the ranges of learning rate $\alpha$ and discount coefficient $\gamma$ are set to be $0 < \alpha < 1$ and $0 < \gamma < 1$; respectively; the initialization of Q-value is set to $0$.

The action selection strategy in reinforcement learning is the key step for the control strategy. Generally, a greedy policy $\pi^*$, which means the action of maximum Q-value will be select as the output action $a_k$, is presented as,

$$\pi^*(s) = \arg \max_{a \in A} Q^k(s, a)$$  \hspace{1cm} (8)$$

Since the agent will select the action with the maximum Q-value at the $s_k$ state in the greedy policy, the other actions at the $s_k$ state may not be searched enough. To search all the actions at the $s_k$ state enough, a probability distribution selection policy is applied to select the action in this article. In the probability distribution selection policy, each Q-value corresponds to each P-value. Therefore, the computer memory of P-value and Q-value matrices are the same. The $P$ matrix at

---

FIGURE 3. Carbon emission ratios of generator, power grid and consumer.
the $s$ state can be updated as,
\[
\begin{align*}
P^k_{a}(s) &= P^k_{a}(s) + \beta (1 - P^k_{a}(s)) \\
\hat{P}^k_{a}(s) &= P^k_{a}(s)(1 - \beta), \quad \forall a \in A, \ a \neq a_g
\end{align*}
\tag{9}
\]
where $\beta$ is the probability coefficient of probability distribution selection policy, $0 < \beta < 1$; $P^k_{a}(s)$ implies the selected probability in the $s$ state; $a_g$ means selected action by the greedy policy. After enough searched on-line, one selected probability of the $\hat{P}^k_{a}(s)$ will be converged to 1, which means an optimal control strategy. The reward function of the AERL for the AVR can be designed as follows,
\[
R = -10\Delta v^2
\tag{10}
\]
where $\Delta v$ is the voltage deviation of the power system.

B. ARTIFICIAL EMOTION

Reinforcement learning belongs to the category of machine learning, which belongs to the category of artificial intelligence. Another major branch of artificial intelligence is artificial psychology, which includes artificial emotions, artificial consciousness, and artificial cognition. Furthermore, artificial emotion is the major branch of artificial psychology.

The artificial emotion should be quantified when the artificial emotion is applied to an engineering problem. The quantizer output $f_n$ and emotional coefficient $\eta$ (FIGURE 4) are simultaneously calculated as
\[
\begin{align*}
f_n &= \sum_{i=1}^{n} \lambda_i = \sum_{i=1}^{n} \theta_i \omega_i \\
\eta &= \begin{cases} 
\eta, & \text{if } 1/f_n \geq \eta_{\text{max}} \\
\eta/f_n, & \text{if } 1/f_n < \eta_{\text{max}}
\end{cases}
\tag{11}
\end{align*}
\]
where $\lambda_i$ is calculated by emotional weight $\omega_i$ and input information $\theta_i$; $f_n$ means the emotion value of the agent; $\eta_{\text{max}}$ means the configured maximum emotional coefficient; $\eta_{\text{max}}$ is the maximum agent emotion, and $\eta_{\text{max}}$ is set to 1 in this article.

The quantified emotion value $C_f(\eta)$ is calculated as
\[
C_f(\eta) = k_a \eta^2 + k_b \eta + k_c
\tag{13}
\]
$k_a$, $k_b$ and $k_c$ are the quadratic, linear, and constant coefficients, respectively. Then, the output of the AERL can be calculated as,
\[
a_{\text{AERL}} = C_f(\eta)a_{\text{Logicpart}}
\tag{14}
\]
where $a_{\text{Logicpart}}$ is the selected action from the logical part of the AERL by Eqs.(7)-(10).

C. ARTIFICIAL EMOTIONAL REINFORCEMENT LEARNING

To achieve smaller voltage deviation and avoid the curse of dimensionality simultaneously, the artificial emotion is applied to update the output action of the reinforcement learning.

The more intelligent agent of the AERL contains two parts, i.e., a logical part and an emotional part (FIGURE 5). The output action of the logical part is modified by the emotional part as,
\[
a_{k}^{\text{output}} = C_f(\eta) a_k
\tag{15}
\]

The major framework of the AERL is a framework of reinforcement learning. Compared with conventional reinforcement learning, the AERL can provide continuous actions rather than discrete actions for controlled systems. Since the AERL is a special reinforcement learning with a large number of actions, the AERL is a reinforcement learning based on the Markov decision process. More details of the proof of the Markov decision process have been presented in [56], [57]. Therefore, the convergence of the AERL can be proved by the Markov decision process as the convergence of reinforcement learning.

IV. COMPLEX-VALUED ENCODING DRAGONFLY ALGORITHM

A. DRAGONFLY ALGORITHM

The DA has been proposed by Mirjalili [34]. Two major processes of the dragonflies swarm (i.e., hunting and migration) are similar to exploration and exploitation, respectively.
Three primitive principles of the dragonflies swarm behaviors are separation $S_i$, alignment $A_i$ and cohesion $C_i$, which are calculated as,

$$S_i = -\sum_{j=1}^{N} (X - X_j)$$  \hspace{1cm} (16)

$$A_i = \frac{\sum_{j=1}^{N} U_j}{N}$$  \hspace{1cm} (17)

$$C_i = \frac{\sum_{j=1}^{N} X_j}{N} - X$$  \hspace{1cm} (18)

where $N$ means the number of individuals $X$; $X_j$ and $U_j$ are the position and velocity of individual, respectively.

The attraction $F_i$ and distraction $E_i$ behavior can be described as,

$$F_i = X^+ - X$$  \hspace{1cm} (19)

$$E_i = X^- + X$$  \hspace{1cm} (20)

where $X^+$ and $X^-$ represent target position and enemy position, respectively.

The position vectors at the $(t + 1)$-th iteration can be updated as,

$$X\,(t+1) = X\,^t + \Delta X\,(t+1)$$

$$= X\,^t + k_s S_t + k_a A_t + k_c C_t + k_t F_t + k_e E_t + k_s \Delta X\,^t$$  \hspace{1cm} (21)

where $k_s$, $k_a$, $k_c$, $k_t$, $k_e$, and $k_s$ imply the weight coefficients of the updated positions. Since all these weight coefficients should be configured for the convergence of the DA, an improved randomness stochastic behavior of the artificial dragonflies for exploration with a random walk (Lévy flight) can be employed to update the dragonflies’ positions at the $t$ iteration as,

$$X\,(t+1) = X\,^t + \text{Lévy}(d) \times X\,^t$$  \hspace{1cm} (22)

where $d$ means variables dimension; Lévy flight $\text{Lévy}(d)$ can be described as [58],

$$\text{Lévy}(d) = 0.01 \times \frac{r_1(d) \times \sigma}{|r_2(d)|^{1/\beta_L}}$$  \hspace{1cm} (23)

where $r_1(d)$ and $r_2(d)$ are two generated random numbers from 0 to $d$; $\beta_L$ is a constant, and $\beta_L$ can set to be 1.5; $\sigma$ can be presented as,

$$\sigma = \left( \frac{\Gamma(1 + \beta_L) \times \sin(\pi \beta_L/2)}{\Gamma((1 + \beta_L)/2) \times \beta_L \times 2^{(\beta_L - 1)/2}} \right)^{1/\beta_L}$$  \hspace{1cm} (24)

where $\Gamma(x)$ is a factorial function, $\Gamma(x) = (x - 1)!$.

**B. COMPLEX-VALUED ENCODING DRAGONFLY ALGORITHM**

To explore and exploit the solution with more spaces, the CDA is proposed in this article. Each individual in the CDA is recorded as,

$$X_p = R_p + iI_p, \quad p = 1, 2, \ldots, N$$  \hspace{1cm} (25)

where $R_p$ and $I_p$ are the real and imaginary values of the complex-valued position, respectively. The upper and lower absolute individual values are $|X_p^{\max}|$ and $|X_p^{\min}|$, respectively.

The complex-valued position of the $p$-th individual can be described as,

$$R_p + iI_p = \rho_p (\cos \theta_p + i \sin \theta_p)$$  \hspace{1cm} (26)

where the range of the radius and the angle of the complex-valued position are $\rho_p \in \left[0, \frac{|X_p^{\max}| - |X_p^{\min}|}{2}\right]$ and $\theta_p \in [-2\pi, 2\pi]$, respectively.

The real and imaginary values of the complex-valued position vectors can be updated as,

$$R^t+1 = R^t + \text{Lévy}(d) \times R^t$$  \hspace{1cm} (27)

$$I^t+1 = I^t + \text{Lévy}(d) \times I^t$$  \hspace{1cm} (28)

The fitness value is calculated by the real number, which can be converted from the complex-valued position, as follows.

$$X_p = \rho_p \text{sgn} \left( \sin \left( \frac{I_p}{\rho_p} \right) \right) + \frac{X_p^{\max} + X_p^{\min}}{2}$$  \hspace{1cm} (29)

where $\rho_p = \sqrt{R_p^2 + I_p^2}$, $p = 1, 2, \ldots, N$.

The convergence of the DA has been verified in [34]. Since the last step of the CDA can provide a real number position, the CDA has the same updated process with the DA. Consequently, the convergence of the CDA has the same convergence of the DA.

The computation complexities of the GA, the PSO, the GWO, the MFO and the DA are $O(n_{\text{iter}} \times n_{\text{size}})$; where $n_{\text{iter}}$ and $n_{\text{size}}$ are the number of maximum iteration and the population size of these optimization algorithms. Since the imaginary part is added into the CDA, the computation complexity of the CDA is $O(2n_{\text{iter}} \times n_{\text{size}})$.

**C. COORDINATION OF COMPLEX-VALUED ENCODING DRAGONFLY ALGORITHM AND ARTIFICIAL EMOTIONAL REINFORCEMENT LEARNING**

To obtain the highest control performance and the convergence performance simultaneously in the CSVC and the AVR, the proposed CDA and the AERL are coordinated in this article, i.e., a coordinated CDA-AERL. The proposed coordinated CDA-AERL (FIGURE 7) is employed to a coordinated framework with two voltage control levels, i.e., the CSVC and the AVR (FIGURE 8). Then, the coordinated CSVC-AVR framework can replace the three layers voltage control framework. With the imaginary part is added into the CDA, the proposed CDA can coordinate with the AERL.
Coordinated Complex-Valued Encoding DA and AERL for CSVC and AVR in Multi-Generator Power Systems

1) After the off-line training of the AERL, the AERL can effectively update the control strategy on-line for dynamic systems.

2) From the multi-agent system perspective, the game playing agent based on the AERL of each reactive power control device game with each other through the dynamic system.

3) The reactive power commands of all the agents are optimized by the CDA.

V. SIMULATION RESULTS
A. ARTIFICIAL EMOTIONAL REINFORCEMENT LEARNING

The proposed AERL is compared with PID and reinforcement learning (Q learning) in the simulation of a three-phase generator rated 200 MVA, 13.8 kV. The parameters of the PID (i.e., $k_p = 0.0001$, $k_i = 28.7652$, $k_d = 0.01$) are optimized by the GA with 1500 population size and 200 maximum iterations.

To achieve a smaller control error, the number of Q learning actions should be increased. In this article, the parameters of Q learning are set as that: action set $A = \{-0.68, -0.612, \ldots, 0.68\}$; $\alpha = 0.1$; $\beta = 0.05$; $\gamma = 0.09$; the number of states of the Q learning is set to 21; the parameters of the AERL are set as that: action set $A = \{-0.49, -0.392, \ldots, 0.49\}$; $\alpha = 0.1$; $\beta = 0.05$; $\gamma = 0.09$; the number of states of the AERL is set to 11; the weight of emotion $\omega = 0.0705$; the quadratic, linear, and constant coefficients of the quantified process of the emotional part of the AERL are set to 0.5, 0.4, and 0.01, respectively.

The results obtained by these simulated methods (FIGURE 9) show that: (i) the number of the simulated Q learning is larger than the number of the simulated AERL, i.e., $21 > 11$; while the terminal voltage obtained by the simulated AERL is nearer to the input voltage than the terminal voltage obtained by the simulated Q learning; (ii) compared with the terminal voltages obtained by the optimized PID and the configured Q learning, the terminal voltage obtained by the configured AERL with the same configuration as the for optimizing the CSVC; with the emotional part is introduced into the logical part, the designed AERL can provide a real-time strategy for the AVR of power systems.

Compared with conventional voltage control, the major characteristics of the coordinated framework are listed as follows.

FIGURE 7. Flow chart of coordinated complex-valued encoding dragonfly algorithm-artificial emotional reinforcement learning.

FIGURE 8. Framework of voltage control: (a) three voltage control levels; (b) coordinated two voltage control levels.

FIGURE 9. Voltage curves obtained by the simulated control algorithms.
simulated Q learning except the actions set is the nearest voltage to the input voltage. The calculation times of the PID, the reinforcement learning and the AERL are 0.000174 s, 0.000265 s and 0.000327 s, respectively.

The simulation results (FIGURE 9) obtained by the proposed AERL show that: (i) the agent based on the proposed AERL can effectively obtain the highest control performance when compared with the other simulated control algorithms; (ii) since the AERL can obtain higher control performance with lesser computation memory than Q learning, the curse of dimensionality can be mitigated. Therefore, with the emotional part in the agent, the agent based on the AERL is more intelligent than the agent based on conventional reinforcement learning. Besides, with the emotional part is added, the agent based on “emotional part + logical part” can obtain higher control performance than the agent only based on the logical part.

B. COMPLEX-VALUED ENCODING DRAGONFLY ALGORITHM

The CDA based on the AERL for the CSVC is applied into three cases, i.e., IEEE 57-bus, 118-bus and 300-bus systems. Each optimization algorithm simulated ten times for all these three power systems. All the cases are simulated on an Intel(R) Core(TM) i7-7820HK CPU 3.90 GHz and 64 GB RAM server with MATLAB R2019b.

1) IEEE 57-BUS SYSTEM

IEEE 57-bus system contains 7 generator nodes, 50 load nodes and 80 branches (FIGURE 10). This case contains nine variables, i.e., three AVR nodes (i.e., 18, 25, and 53 nodes) and six transformer tap changers (i.e., 4-18, 21-20, 34-32, 39-57, 7-29, and 9-55). Each variable has five options, i.e., -40%, -20%, 0%, 20%, and 40% for the AVR nodes; 0.80 p.u., 0.90 p.u., 1.00 p.u., 1.01 p.u., and 1.02 p.u. for the transformer tap changer.

The numbers of individuals and iterations of all the compared optimization algorithms (i.e., GA, PSO, GWO, MFO, DA, and CDA) in this case are configured as 1000 and 200, respectively. The parameters of all the simulated approaches in this case are set to similar values or default values, which are given in TABLE 1.

The convergence curves of one time of each compared optimization algorithm in this case are shown in FIGURE 11. FIGURE 11 shows that: (i) the CDA can obtain the minimum fitness function value when compared with the simulated methods; (ii) the CDA can convergence to the minimum fitness function value with the smallest iteration number. The optimal solutions obtained by all the compared methods are given in TABLE 2. The statistical results of these ten times simulations in this case are shown in FIGURE 12. FIGURE 12 shows that: compared with the simulated methods, since the CDA can obtain multiple approximate minimum fitness function values with multiple simulations, the CDA has a stable convergence feature, which means that an optimization algorithm can obtain a nearly similar solution with multiple random running. The statistical calculation times of these ten times simulations in this case are shown.

**TABLE 1. Parameters of all the compared optimization algorithms.**

| Algorithm | Parameters | Value |
|-----------|------------|-------|
| GA        | Maximum stall generations | 200   |
| GA        | Crossover fraction probability | 0.8   |
| GA        | Constraint tolerance | 0.001 |
| GA        | Elite probability | 0.05  |
| GA        | Function tolerance | 0.0001 |
| PSO       | Maximum stall generations | 200   |
| PSO       | Function tolerance | 0.0001 |
| PSO       | Cognitive attraction | 0.5   |
| PSO       | Social attraction | 1.25  |
| GWO       | Maximum stall generations | 200   |
| MFO       | Maximum stall generations | 200   |
| MFO       | Spiral shape constant determining | 0.5   |
| DA        | Maximum stall generations | 200   |
| DA        | Weight coefficients \(k_a, k_x, k_c, k_f, k_n\) | \(0.2, 0.2, 0.2, 2, 1\) |
| DA        | Constant coefficient \(\beta_d\) | 1.5   |
| CDA       | Maximum stall generations | 200   |
| CDA       | Weight coefficients \(k_a, k_x, k_c, k_f, k_n\) | \(0.2, 0.2, 0.2, 2, 1\) |
| CDA       | Constant coefficient \(\beta_p\) | 1.5   |
TABLE 2. Optimal solutions obtained by all the compared algorithms at IEEE 57-bus system.

| Algorithms | Solutions |
|------------|-----------|
| GA         | 0.4 0.4 0.4 1.0 0.9 0.9 1.0 0.9 1.0 |
| PSO        | 0.2 0.4 0.0 0.9 1.0 0.9 0.9 0.9 1.0 |
| GWO        | -0.2 0.4 0.0 0.9 1.0 0.9 0.9 0.9 1.0 |
| MFO        | -0.4 0.4 0.4 0.8 1.0 0.9 0.9 0.9 1.0 |
| DA         | -0.2 0.2 0.2 0.8 1.1 0.9 0.9 0.9 1.0 |
| CDA        | 0.2 0.4 0.4 0.9 1.0 0.9 0.9 0.9 1.0 |

in FIGURE 13. FIGURE 13 shows that the calculation time of the CDA is approximate to the calculation times of other compared methods.

The simulation results (FIGURE 11 and FIGURE 12) of this case show that:

1) Compared with other algorithms, the fitness value obtained by the proposed CDA is the smallest one (FIGURE 11).

2) From the statistical results of the fitness value with ten times (FIGURE 12), compared with other algorithms, the convergence of the proposed CDA is higher than that of other algorithms.

2) IEEE 118-BUS SYSTEM

A total of 25 variables of this case, which consists of 54 generators (FIGURE 14), are selected (TABLE 3). The numbers of individuals and iterations of all the compared approaches in this case are set to 2000 and 200, respectively. The parameters of all the compared optimization algorithms in this case are given in TABLE 1.

The convergence curves of one time of each compared optimization algorithm in this case are shown in FIGURE 15. FIGURE 15 shows that: compared with the convergence curve obtained by the DA, the CDA can obtain a convergence curve with smaller fitness function value. The statistical results of these ten times simulations in this case are shown in FIGURE 16. FIGURE 16 shows that: with the same configurations as the compared methods, the CDA has more opportunity to obtain a more optimal solution. The statistical calculation times of these ten times simulations in this case are shown in FIGURE 17. FIGURE 17 shows that the average calculation time of the CDA is smaller than the calculation times of other compared methods.
TABLE 3. Variables of IEEE 118-bus system.

| Variables          | Nodes/branches | Range of value  |
|--------------------|----------------|-----------------|
| AVR nodes          | 45, 79, 105    | -40%, 20%, 0%, 20%, 40% |
| Transformer tap changer | 8-5, 26-25, 30-17, 63-59, 64-61 | 0.98 p.u., 1.00 p.u., 1.02 p.u. |
| Generator voltage  | 10, 12, 25, 26, 46, 49, 54, 59, 61, 65, 66, 69, 80, 89, 100, 103, 111 | 1.00 p.u., 1.01 p.u., 1.02 p.u., 1.03 p.u., 1.04 p.u., 1.05 p.u., 1.06 p.u. |

TABLE 4. Variables of IEEE 300-bus system.

| Variables          | Nodes/branches | Range of value  |
|--------------------|----------------|-----------------|
| AVR nodes          | 117, 120, 154, 164, 166, 173, 190, 231, 238, 240, 248 | -40%, -20%, 0%, 20%, 40% |
| Transformer tap changer | 9001-9022, 9002-9024, 9023-9025, 9023-9026, 9007-9071, 9007-9072, 9003-9031, 9003-9032, 9003-9033, 9004-9041, 9004-9042, 9004-9043, 9003-9034, 9003-9035, 9003-9036, 9003-9037, 9003-9038, 213-214, 222-237, 227-231, 241-237, 45-46, 73-74, 81-88, 85-99, 86-102, 122-157, 142-175, 145-180, 200-248, 211-212, 223-224, 196-2040, 7003-3, 7003-61, 7166-166, 7024-24, 7001-1, 7130-130, 7011-11, 7023-23, 7049-49, 7139-139, 7012-12 | 0.98 p.u., 1.00 p.u., 1.02 p.u. |
| Generator voltage  | 84, 91, 92, 98, 108, 119, 124, 141, 143, 146, 147, 149, 152, 153, 170, 176, 177, 185, 186, 187, 190, 191, 198, 213, 220, 221, 222, 227, 230, 233, 236, 238, 239, 241, 242, 243, 7001, 7002, 7003, 7011, 7012, 7017, 7023, 7024, 7039, 7044, 7049, 7055, 7057, 7061, 7062, 7071, 7130, 7139, 7166, 9054 | 1.00 p.u., 1.01 p.u., 1.02 p.u., 1.03 p.u., 1.04 p.u., 1.05 p.u., 1.06 p.u. |

The simulation results (FIGURE 15 and FIGURE 16) of this case show that the CDA can obtain the optimal objective with a stability feature effectively for IEEE 118-bus system.

3) IEEE 300-BUS SYSTEM
A total of 111 variables in this case are selected (TABLE 4) in this case, which contains 69 generators (FIGURE 18). The numbers of individuals and iterations of all the compared approaches in this case are configured as 300000 and 200, respectively. The parameters of all the compared optimization algorithms in this case are given in TABLE 1.

The convergence curves of one time of each compared optimization algorithm in this case are shown in FIGURE 19. FIGURE 19 shows that: compared with the convergence curves obtained by the simulated methods, the CDA can obtain a convergence curve with the minimum fitness function value. FIGURE 11, FIGURE 15 and FIGURE 19 show that the CDA can obtain the minimum fitness function value with a complex optimization problem. The statistical results of these ten times simulations in this case are shown in FIGURE 20. FIGURE 12, FIGURE 16 and FIGURE 20 show that: with the same configurations as the compared methods, the CDA has more opportunity to obtain a more optimal solution for both simple and complex optimization problems. The statistical calculation times of these ten times simulations in this case are shown in FIGURE 21. FIGURE 21 shows that: for a complex optimization problem, the average calculation time of the CDA is less than the average calculation times of other compared methods except for the GA.

The simulation results (FIGURE 19 and FIGURE 20) of this case show that the proposed CDA can effectively obtain the economical operation.
considered in the DA, the CDA can fit optimization problems with decentralized independent variables more effectively than the DA; (iii) after simulating multiple times, the fitness values obtained by the CDA are the values that have one smaller range than that of other simulated algorithms under all the simulated cases.

The average fitness values obtained by all the simulated algorithms under IEEE 57-bus, 118-bus and 300-bus systems are given in TABLE 5, which shows that the proposed CDA can effectively fit the CSVC than other compared algorithms.

**C. DISCUSSIONS**

Although the AERL of the coordinated optimization and control algorithm can obtain higher control performance with nearer voltage curves to reference voltage curves than optimized PID and configured Q learning, the control performance obtained by the AERL could be changed with learning rate, probability distribution, discounted rate, the number of states, the weight of emotion and coefficients of the quantified process of the emotional part of the AERL. After numerous testing: (i) the ranges of learning rate, probability distribution and discount coefficient of AERL can be configured as $[0.001, 0.2]$, $[0.0001, 0.1]$ and $[0.01, 0.99]$, respectively; (ii) the number of states of the AERL can be set as $\{9, 11, \ldots, 51\}$; (iii) the weight of emotion of the AERL can be set as $[0.001, 0.1]$; (iv) the ranges of the quadratic coefficient, the linear coefficient and the constant coefficient of quantified process of the AERL can be set as $[0.3, 0.9]$, $[0.1, 0.9]$ and $[0.001, 0.01]$, respectively.

Since the CDA can fit the optimization problems with decentralized independent variables more suitable than other compared optimization algorithms, the final optimal objective values obtained by the CDA are smaller than other compared optimization algorithms. Generally, the convergence performances of the CDA are stable when the population size and the maximum iterations of the CDA are larger than 3 times optimized variables and 100, respectively. For example, the numbers of individuals and iterations of CDA should be large than 27 (or $3 \times$ nine variables) and 100 for
IEEE 57-bus system, respectively. Therefore, the maximum calculation times used by the CDA for IEEE 57-bus, 118-bus and 300-bus systems are 21.68 s, 67.20 s and 248.64 s, respectively, which are less than the optimization periods of the CSVC of these simulated systems. After numerous iterations in the optimization process of the CDA, the sensitivity of the configured weight coefficients of the CDA can be mitigated by the random features of the Lévy flight function. After the population size and the maximum iterations are respectively configured larger than 3 times optimized variables and 100, the weight coefficients of the updated positions of the CDA can be configured as random values with the range of (0, 1]. Graphics processing units and field programmable gate arrays can be applied to reduce the calculation time of the CDA.

The simulation results show that the improved optimization algorithm by the complex-valued encoding operation can fit optimization problems with decentralized independent variables. Numerous other similar optimization algorithms can be improved by the complex-valued encoding operation, such as the GWO, the MFO, etc. Besides, these five types of operations can be mixed together to improve optimization performance. For example, the more grouped operation can be integrated with the different coded operation. Although the CDA can fit optimization problems with decentralized independent variables, more complex optimization problems with complex constraint conditions could not be effectively fitted by the CDA.

VI. CONCLUSION AND PROSPECT

The AERL is applied to the AVR in this article. A novel complex-valued encoding optimization algorithm, which is named as CDA, is proposed in this article. Then, the CDA and the AERL are coordinated for the CSVC and AVR. To compare with the convergence of the CDA and the control performance of the AERL, three simulation cases (i.e., IEEE 57-bus, 118-bus and 300-bus systems) are simulated in this article. The major features of the coordinated CDA and AERL can be summarized as follows.

1) Compared with the simulation results obtained by the AERL and other algorithms, the AERL can effectively obtain the highest control performance. Since the agent based on the AERL has an artificial emotional part, the agent based on the AERL is more intelligent than the agent based on only reinforcement learning.
2) Since the output of the AERL can be modified by the artificial emotional part of the AERL, the AERL can update the control strategy on-line and can mitigate the curse of dimensionality.
3) Since the two voltage control levels are coordinated, the proposed coordinated CDA and AERL can effectively obtain the optimal convergence for the CSVC and AVR in multi-generator power systems.
4) The proposed CDA can effectively fit optimization problems with decentralized independent variables.

In the future works: (i) the proposed coordinated CDA and AERL could be employed into numerous optimization problems with decentralized independent variables, such as the coordinated economic dispatch and automatic generation control; (ii) the proposed coordinated CDA could be applied to mix integer nonlinear decentralized independent variables optimization problems, such as the unit commitment of power grids, energy and production efficiency optimization problems, etc.

REFERENCES

[1] Y. Lu and K. Tomsovic, “Wide area hierarchical voltage control to improve security margin for systems with high wind penetration,” IEEE Trans. Power Syst., vol. 33, no. 6, pp. 6218–6228, Nov. 2018.
[2] H. Tian, Y. Liu, Y. Zhang, and X. Qin, “An improved secondary voltage control method for half-wavelength transmission system,” Int. J. Electr. Power Energy Syst., vol. 115, Feb. 2020, Art. no. 105417.
[3] H. Liu, J. Su, J. Qi, N. Wang, and C. Li, “Decentralized voltage and power control of multi-machine power systems with global asymptotic stability,” IEEE Access, vol. 7, pp. 14273–14282, 2019.
[4] D. K. Sahoo, R. K. Sahu, G. T. C. Sekhar, and S. Panda, “A novel modified differential evolution algorithm optimized fuzzy proportional integral derivative controller for load frequency control with thyristor controlled series compensator,” J. Electr. Syst. Inf. Technol., vol. 5, no. 3, pp. 944–963, Dec. 2018.
[5] J. Luo and Z. Liu, “Novel grey wolf optimization based on modified differential evolution for numerical function optimization,” Int. J. Speech Technol., vol. 50, no. 2, pp. 468–486, Feb. 2020.
[6] K. Jagatheesan, B. Anand, S. Samanta, N. Dey, A. S. Ashour, and V. E. Balas, “Design of a proportional-integral-derivative controller for an automatic generation control of multi-area power thermal systems using firefly algorithm,” IEEE/CAA J. Automatica Sinica, vol. 6, no. 2, pp. 503–515, Mar. 2019.
[7] Y.-P. Huang, M.-Y. Huang, and C.-E. Ye, “A fusion firefly algorithm with simplified propagation for photovoltaic MPPT under partial shading conditions,” IEEE Trans. Sustain. Energy, vol. 11, no. 4, pp. 2641–2652, Oct. 2020.
[8] S. M. H. Mousakazemi, N. Ayoobian, and G. R. Ansarifar, “Control of the pressurized water nuclear reactors power using optimized proportional-integral-derivative controller with particle swarm optimization algorithm,” Nucl. Eng. Technol., vol. 50, no. 6, pp. 877–885, Aug. 2018.
[9] J. Liang, S. Ge, B. Qu, K. Yu, F. Liu, H. Yang, P. Wei, and Z. Li, “Classified perturbation mutation based particle swarm optimization algorithm for parameters extraction of photovoltaic models,” Energy Convers. Manage., vol. 203, Jan. 2020, Art. no. 112138.
[10] R. Xiong, J. Cao, and Q. Yu, “Reinforcement learning-based real-time power management for hybrid energy storage system in the plug-in hybrid electric vehicle,” Appl. Energy, vol. 211, pp. 538–548, Feb. 2018.
[11] H. Zhang, H. Jiang, Y. Luo, and G. Xiao, “Data-driven optimal consensus control for discrete-time multi-agent systems with unknown dynamics using reinforcement learning method,” IEEE Trans. Ind. Electron., vol. 64, no. 5, pp. 4091–4100, May 2017.
[12] L. Ma, S. Cheng, X. Wang, M. Huang, H. Shen, X. He, and Y. Shi, “Cooperative two-engine multi-objective bee foraging algorithm with reinforcement learning,” Knowl.-Based Syst., vol. 133, pp. 278–293, Oct. 2017.
[13] H. Zhang, Y. Liang, H. Su, and C. Liu, “Event-driven guaranteed cost control design for nonlinear systems with actuator faults via reinforcement learning algorithm,” IEEE Trans. Syst., Man, Cybern. Syst., early access, Oct. 24, 2019, doi: 10.1109/TCSM.2019.2946857.
[14] Y. Yuan, Z. L. Yu, Z. Gu, Y. Yeboah, W. Wei, X. Deng, J. Li, and Y. Li, “A novel multi-step Q-learning method to improve data efficiency for deep reinforcement learning,” Knowl.-Based Syst., vol. 175, pp. 107–117, Jul. 2019.
[15] F. Belletti, D. Haziza, G. Gomes, and A. M. Bayen, “Expert level control of ramp metering based on multi-task deep reinforcement learning,” IEEE Trans. Intell. Transp. Syst., vol. 19, no. 4, pp. 1198–1207, Apr. 2018.
[16] T. M. Moerland, J. Broekens, and C. M. Jonker, “Emotion in reinforcement learning agents and robots: A survey,” Mach. Learn., vol. 107, no. 2, pp. 443–480, Feb. 2018.

[17] A. Athar, K. Ahmed, S. Abbas, and Y. Saeed, “Modeling cooperative and competitive emotions using parallel SOM for artificially intelligent agents,” Int. J. Comput. Sci. Netw. Secur., vol. 17, no. 2, p. 270, 2017.

[18] L. Yin, T. Yu, L. Zhou, L. Huang, X. Zhang, and B. Zheng, “Artificial emotional reinforcement learning for automatic generation control of large-scale interconnected power grids,” IET Gener., Transmiss. Distrib., vol. 11, no. 9, pp. 2305–2313, Jun. 2017.

[19] G. Chen and Z. Guo, “Distributed secondary and optimal active power sharing control for islanded microgrids with communication delays,” IEEE Trans. Smart Grid, vol. 10, no. 2, pp. 2002–2014, Mar. 2019.

[20] H. Ma and D. J. Hill, “Sensitivity studies for adaptive coordinated voltage control: Scale and similarity of contingencies,” IEEE Trans. Power Syst., vol. 32, no. 5, pp. 3794–3802, Sep. 2017.

[21] H. Yassami, F. Bayat, A. Jalilvand, and A. Rabiee, “Coordinated voltage control of wind-penetrated power systems via state feedback control,” Int. J. Electr. Power Energy Syst., vol. 93, pp. 384–394, Dec. 2017.

[22] J. P. Paul, J. Y. Leost, and J. M. Tesseron, “Survey of the secondary voltage control of multi-area power grids using model predictive control,” vol. 210, pp. 914–924, Jan. 2018.

[23] A. Tharwat, T. Gabel, and A. E. Hassanien, “Parameter optimization of artificial neural networks based on water cycle algorithm for resolution enhancement of remote-sensing image classification,” Int. J. Remote Sens., vol. 38, no. 23, pp. 6970–6992, Dec. 2017.

[24] H. Mehrjerdi, S. Lefebvre, M. Saad, and D. Asber, “Coordinated control of multi-area power grids using model predictive control: Scale and similarity of contingencies,” IEEE Trans. Power Syst., vol. 93, pp. 384–394, Dec. 2017.
LINFEI YIN was born in Jiujiang, Jiangxi, China, in 1990. He received the B.S. and M.S. degrees in information engineering from Nanchang Hangkong University, Nanchang, China, in 2012 and 2015, respectively, and the Ph.D. degree from the College of Electric Power Engineering, South China University of Technology, Guangzhou, China, in 2018. He is currently an Assistant Professor with the College of Electrical Engineering, Guangxi University, Nanning, Guangxi, China. His major research interest includes artificial intelligence techniques in operation of power systems.

SHIKUI LUO was born in Qinzhou, Jiulong, China, in 1996. He received the B.S. degree in electrical and it’s automation from Guangxi University, Nanning, China, in 2019, where he is currently pursuing the M.S. degree with the College of Electrical Engineering. His major research interest includes artificial intelligence techniques in operation of power systems.

YAOXIONG WANG received the B.S. degree in physics and the M.S. degree in theoretical physics from Northwest University, Xi’an, China, in 2006 and 2009, respectively. He is currently an Assistant Researcher with the Institute of Intelligent Machines, Hefei Institutes of Physical Science, Chinese Academy of Sciences. His current research interests include industrial defects identification and marking based on visual information, and real-time tracking with machine learning.

FANG GAO was born in Zhenping, Henan, China, in 1983. He received the B.S. and Ph.D. degrees in chemical physics from the University of Science and Technology of China, in 2004 and 2010, respectively. After graduation, he has worked as an Assistant Professor and an Associate Professor with the Institute of Intelligent Machines, Chinese Academy of Sciences. He is currently a Professor with the College of Electrical Engineering, Guangxi University, Nanning, Guangxi, China. His major research interests include artificial intelligence techniques in the operation of power systems and machine learning methods in the control of quantum systems.

JUN YU (Member, IEEE) was born in Chuzhou, Anhui, China, in 1982. He received the B.S. degree in electronic engineering from the National University of Defence Technology, in 2003, and the Ph.D. degree in pattern recognition and intelligent system from the University of Science and Technology of China, in 2010. He is currently an Associate Professor with the Department of Automation, University of Science and Technology of China. His research interests include multimedia computing, multi-modal information synthesis, perception, and cognition. He has published more than 100 journal articles and conference papers, such as TIP, TMM, TASLP, TCSVT, TCYB, TITS, TOMM, TCDS, ACL, CVPR, MM, SIGGRAPH ASIA, VR, AAAI, and IJCAI. He has received two Best Paper Awards from premier conferences, namely ICME and FG, and won 12 Grand Challenge Champions and 1st Runner-Up Awards from premier conferences, such as CVPR, MM, ECCV, ICME, and FG.