Perspective

Can ChatGPT Boost Artistic Creation: The Need of Imaginative Intelligence for Parallel Art

By Chao Guo, Yue Lu, Yong Dou, Fei-Yue Wang, *Fellow, IEEE*

CHATGPT, one of the leading Large Language Models (LLMs), has acquired linguistic capabilities such as text comprehension and logical reasoning, enabling it to engage in natural conversations with humans. As illustrated in "What Does ChatGPT Say: The DAO from Algorithmic Intelligence to Linguistic Intelligence" [1], there are three levels of intelligence : 1) Algorithmic Intelligence (AI), 2) Linguistic Intelligence (LI), 3) Imaginative Intelligence (II). ChatGPT is a powerful demonstration of Linguistic Intelligence, another milestone after AlphaGo for Algorithmic Intelligence [1]–[3]. We believe the next breakthrough in intelligence should be Imaginative Intelligence for artistic creation.

This perspective prescribed a pathway to achieve II for artistic works through Parallel Art, in which LLMs like ChatGPT can serve as linguistics-based artistic knowledge foundation models and text-based human-machine interfaces for human-in-the-loop learning. Multi-modal artistic knowledge foundation models are constructed to perform linguistic, vision, and decision-making tasks of artistic creation in the human-cyber-physical hybrid creative systems. Besides, a case study of text-based painting imagination using ChatGPT is presented.

Parallel Art

To improve the creativity of machines and achieve mutual improvements and collaborations between humans and machines in artistic creation, we proposed the paradigm, framework, and methodologies of Parallel Art [4]–[6] based on Parallel Theory [7]–[13], and developed prototype systems to explore human-cyber-physical hybrid collaborative painting creation [14]–[17]. Parallel Systems and ACP methods [18]–[20] conduct computational experiments in artificial systems and perform virtual-real interactions and iterative optimizations between the artificial systems (virtual) and actual systems (real), expanding exploration space and improving machine learning efficiency [21]–[23].

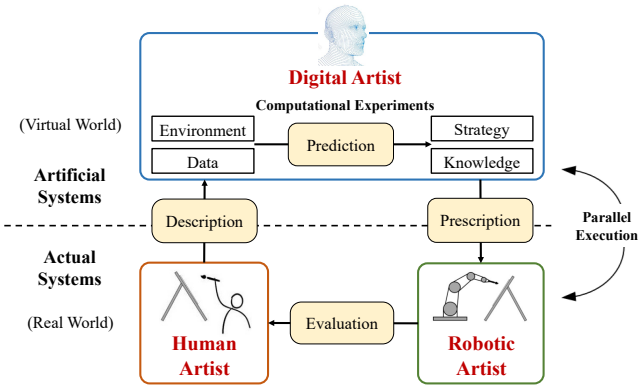

Fig. 1. Parallel Art framework. There are four stages (description, prediction, prescription, evaluation) in the iterative improvement between the artificial and actual systems, and three kinds of roles (human artist, digital artist in virtual world, robotic artist in real physical world) involved in the human-cyber-physical hybrid collaboration systems.

In the ACP-based Parallel Art framework, there are four stages, including description, prediction, prescription, and evaluation, as shown in Fig. 1. 1) In the description stage, data on actual human creative processes are collected and a virtual artificial system is established to reflect the actual creative process by learning from this data. 2) In the prediction stage, extensive painting experiments are conducted to explore creative plans in the established artificial environment and optimized painting strategies are recommended for execution in the actual painting process. Computational painting experiments can overcome the physical, temporal, and cost constraints in actual creation, thus expanding the exploration space and enhancing the efficiency of the creative process. 3) In the prescription stage, optimized strategies are performed with the guidance of virtual systems to improve the creative performance of the real painting system. 4) In the evaluation stage, the human artist will evaluate the computational experimental plans and actual creation results to provide human feedback for iterative optimization of the model, which is similar to the Reinforcement Learning from Human Feedback

Citation: C. Guo, Y. Lu, Y. Dou, and F.-Y. Wang, "Can ChatGPT boost artistic creation: The need of imaginative intelligence for parallel art," *IEEE/CAA J. Autom. Sinica*, vol. 10, no. 4, pp. 835–838, Apr. 2023.

C. Guo and Y. Lu are with The State Key Laboratory for Management and Control of Complex Systems, Chinese Academy of Sciences, Beijing 100190, China. Y. Lu is also with the School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing 100049, China (e-mail: chao.guo@ia.ac.cn; luyue2016@ia.ac.cn).

Y. Dou is with the Macao Institute of Systems Engineering, Macau University of Science and Technology, Macao 999078, China (e-mail: douyong@fortunevc.com).

F.-Y. Wang is with The State Key Laboratory for Management and Control of Complex Systems, Chinese Academy of Sciences, Beijing 100190 and also the Macao Institute of Systems Engineering, Macau University of Science and Technology, Macao 999078, China (e-mail: feiyue.wang@ia.ac.cn).

Digital Object Identifier 10.1109/JAS.2023.123555

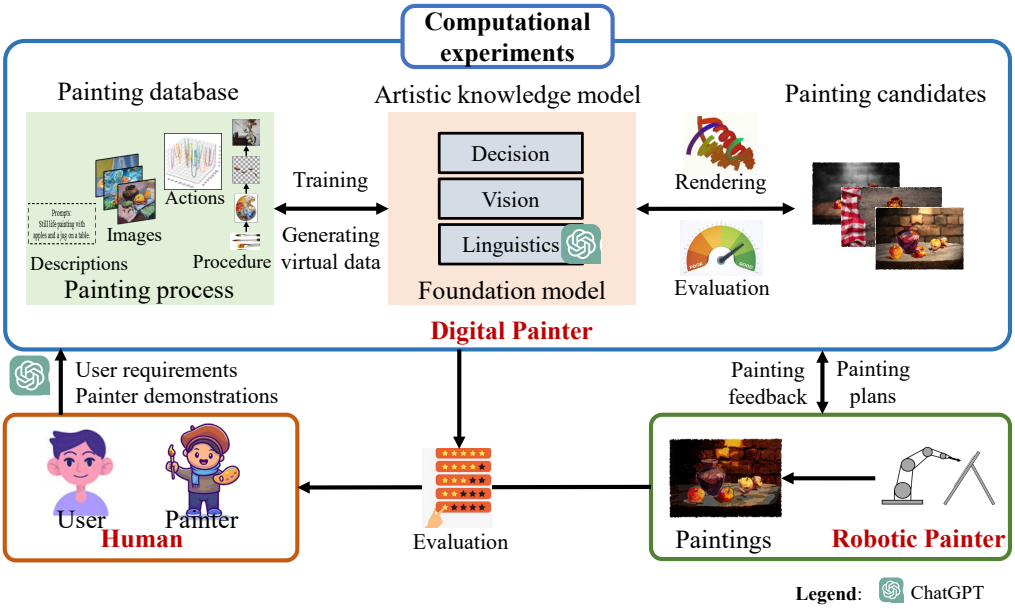

Fig. 2. ChatGPT in Parallel Art (consider painting creation as an example). ChatGPT can serve as linguistics-based artistic knowledge foundation models and text-based human-machine interaction interfaces in Parallel Art's creative computational experiment phase and human-in-the-loop learning phase.

(RLHF) applied in ChatGPT [24]. The artificial and actual systems are executed in parallel and optimized iteratively, which means that the artificial system will be updated based on observations of the real system, while the real system executes based on the practice and exploration of the artificial system.

There are three kinds of roles in Parallel Art to form a human-cyber-physical hybrid collaboration, i.e., human artist, digital artist, and robotic artist. 1) The digital artist is the counterpart of the human artist in cyberspace, containing art creation data and knowledge of human artists. It will perform large-scale creative experiments in artificial systems on behalf of the human artist. 2) The robotic artist is the counterpart of the human artist in physical space, replacing the human artist to perform physical creative processes, helping humans put their efforts where deep creativity is needed. 3) The human artist interacts with the digital artist and the robotic artist in two ways. On the one hand, human artists provide painting demonstrations for machines (digital and robotic painters) to imitate and learn from. On the other hand, human artists evaluate the creative processes and the results of digital and robotic artists, enabling them to improve skills based on human feedback and evolve to be companions of human artists to meet their personalized requirements. Human and robotic artists are involved in physical creation process, forming actual systems. Digital artists form artificial systems. Humans and machines collaborate to improve their capabilities in art creation.

The core of Parallel Art is to conduct computational experiments in virtual systems for creative exploration and perform machine learning with the help of human demonstrations and evaluation feedback through human-in-the-loop mechanisms. Computational experiments require artistic knowledge and skills to perform exhaustive artistic association, transformation, and combination in the creation process. As for human-in-the-loop learning, efficient human-machine interaction is required to achieve the collection of creative demonstrations and evaluations from human experts, as well as the precise conveying of human creative requirements.

ChatGPT for Parallel Art

ChatGPT, learning from large amounts of textual data, is capable of understanding, reasoning, association, and expression over creative content and even emotions and human-machine interaction in a text-based manner [25]. This demonstrates that ChatGPT masters sufficient theoretical and practical artistic knowledge and is able to communicate with humans in a natural conversational way, understanding and responding to their needs.

Therefore, ChatGPT and LLMs can serve two roles in Parallel Art: 1) linguistics-based artistic knowledge models required in the computational experiments, 2) text-based human-machine interaction interfaces required in human-in-the-loop learning, as shown in Fig. 2.

Based on the linguistics-based artistic knowledge learned from the web, ChatGPT can provide solutions according to the given painting requirements in the form of texts, enhancing the diversity and accuracy of painting computational experiments.

Besides, ChatGPT can serve as the human-machine interface between humans, algorithms, and robots with the help of prompt engineering [26]. ChatGPT is able to communicate with people in a conversational manner and understand their complex and abstract needs, such as emotions to convey. This enables the effective transmission of artistic requirements, concepts, and emotions that were initially difficult to express between humans and machines, which will further facilitate better artistic expression by digital artists in computational experiments and by robotic artists in actual executions. In addition, the experts' creative processes can be effectively recorded in the form of texts, forming a painting database

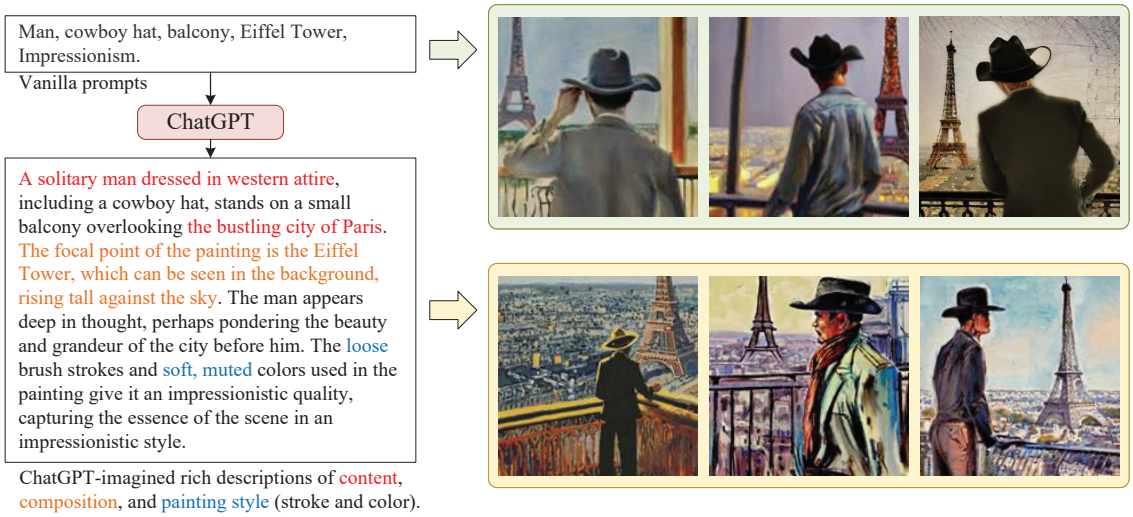

Fig. 3. Text-based painting imagination by ChatGPT can help art generation in Parallel Art using its artistic knowledge.

that can be further used to train digital artists.

Imaginative Intelligence in Parallel Art

In artistic creation, we argue that Imaginative Intelligence is not only embodied in the exploration of various creative plans and effects, but also in knowing how to complete them, i.e. imaging the creative process. In Parallel Art, we realize the Imaginative Intelligence of artistic creation by learning from the creative process and painting computational experiments.

As shown in Fig. 2, the creative data such as textual descriptions [27]–[29], visual states, and action demonstrations are collected, covering the whole human creative process containing conception, creating, results, evaluation, etc. General artistic knowledge towards the entire creation lifecycle is learned from these expert data, and multi-modal foundation models involving decision-making, vision, and linguistic tasks in artistic creation are constructed [30]–[32]. These artistic knowledge models representing the digital painter will conduct large-scale painting experiments in the artificial painting system, and update the painting knowledge according to the experimental evaluation results iteratively.

LLMs like ChatGPT contain only linguistics-based knowledge and are a special case of the artistic knowledge foundation models. Multi-modal foundation models of general artistic knowledge, which can handle language, vision, and decision-making tasks in artistic creation process, will be realized in the near future with the development of foundation models [33] and the increase of associated multi-modal data and computing power. Based on these models, digital painters can achieve comprehensive imagination from the painting process to the final painting and carry out exhaustive explorations of painting strategies to realize the Imaginative Intelligence of artistic creation.

Besides, emerging approaches of human-to-robot skill transfer and human-in-the-loop machine learning, such as RLHF, will promote human-cyber-physical hybrid learning and creation between humans, digital artists, and robotic artists in social space, cyberspace, and physical space respectively.

Text-based Painting Imagination by ChatGPT

Thanks to the artistic knowledge and logical association abilities possessed by ChatGPT, it is feasible to employ it in text-based painting imagination and explore the II initially at present. We use the text-to-image model Stable Diffusion [34] as the pipeline for generating painting images and introduce ChatGPT ahead of the generative model as the art knowledge model to produce a text-based painting plan according to the given painting requirements.

As shown in Fig. 3, ChatGPT produced rich descriptions of painting contents based on the given requirements, such as *"A solitary man dressed in western attire"* and *"the bustling city of Paris"*. It also imagined a reasonable relationship between painting elements, such as *"(A man) overlooking the bustling city of Paris"* and *"The focal point of the painting is the Eiffel Tower, which can be seen in the background, rising tall against the sky"*. Besides, based on the painting style requirements of *"Impressionism"*, ChatGPT points out that the strokes should be *"loose"* and the colors should be *"soft"* and *"muted"*.

The demonstrations show that ChatGPT can provide more explicit and detailed guidance on painting content and more reasonable organization of painting elements, forming clear and reasonable control over painting content. In addition, ChatGPT is able to understand abstract artistic expressions such as painting styles and emotions and relates these abstract concepts to specific painting techniques like brushstrokes and colors through text descriptions. Especially for the expression of emotion, which is difficult for existing AI artistic creation methods. LLMs typified by ChatGPT initially demonstrate the possibility to realize painting imagination based on the extensive linguistics-based artistic knowledge it has learned.

ACKNOWLEDGMENT

This work was supported in part by the Skywork Intelligence Culture and Technology LTD, the Science and Technology Development Fund, Macau Special Administrative Region

(SAR) (0050/2020/A1), and the National Natural Science Foundation of China (61533019).

REFERENCES

[1] F.-Y. Wang, Q. Miao, X. Li, X. Wang, and Y. Lin, "What does ChatGPT say: The DAO from algorithmic intelligence to linguistic intelligence," *IEEE/CAA Journal of Automatica Sinica*, vol. 10, no. 3, pp. 575–579, 2023.

[2] F.-Y. Wang, J. J. Zhang, X. Zheng, X. Wang, Y. Yuan, X. Dai, J. Zhang, and L. Yang, "Where does alphago go: From church-turing thesis to alphago thesis and beyond," *IEEE/CAA Journal of Automatica Sinica*, vol. 3, no. 2, pp. 113–120, 2016.

[3] J. Zhou, P. Ke, X. Qiu, M. Huang, and J. Zhang, "ChatGPT: Potential, prospects, and limitations," *Frontiers of Information Technology & Electronic Engineering*, pp. 1–6, 2023.

[4] F.-Y. Wang, "Parallel art: From intelligent art to artistic intelligence," The Alfred North Whitehead College, Technical Report, 2017.

[5] C. Guo, Y. Lu, Y. Lin, F. Zhuo, and F.-Y. Wang, "Parallel art: Artistic creation under human-machine collaboration," *Chinese Journal of Intelligent Science and Technology*, vol. 1, no. 4, pp. 335–341, 2019.

[6] C. Guo, Y. Dou, T. Bai, X. Dai, C. Wang, and Y. Wen, "Artverse: A paradigm for parallel human–machine collaborative painting creation in metaverses," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, pp. 1–9, 2023.

[7] F.-Y. Wang, "Parallel system methods for management and control of complex systems," *Control and Decision*, vol. 19, pp. 485–489, 2004.

[8] F.-Y. Wang, X. Wang, L. Li, and L. Li, "Steps toward parallel intelligence," *IEEE/CAA Journal of Automatica Sinica*, vol. 3, no. 4, pp. 345–348, 2016.

[9] L. Li, Y. Lin, N. Zheng, and F.-Y. Wang, "Parallel learning: A perspective and a framework," *IEEE/CAA Journal of Automatica Sinica*, vol. 4, no. 3, pp. 389–395, 2017.

[10] F.-Y. Wang, "Parallel intelligence in metaverses: Welcome to hanoi!" *IEEE Intelligent Systems*, vol. 37, no. 1, pp. 16–20, 2022.

[11] F.-Y. Wang, "The metaverse of mind: Perspectives on DeSci for DeEco and DeSoc," *IEEE/CAA Journal of Automatica Sinica*, vol. 9, no. 12, pp. 2043–2046, 2022.

[12] F.-Y. Wang, "The dao to metacontrol for metasystems in metaverses: The system of parallel control systems for knowledge automation and control intelligence in cpss," *IEEE/CAA Journal of Automatica Sinica*, vol. 9, no. 11, pp. 1899–1908, 2022.

[13] Q. Miao, Y. Lv, M. Huang, X. Wang, and F.-Y. Wang, "Parallel learning: Overview and perspective for computational learning across syn2real and sim2real," *IEEE/CAA Journal of Automatica Sinica*, vol. 10, no. 3, pp. 603–631, 2023.

[14] C. Guo, T. Bai, X. Wang, X. Zhang, Y. Lu, X. Dai, and F.-Y. Wang, "ShadowPainter: Active Learning Enabled Robotic Painting through Visual Measurement and Reproduction of the Artistic Creation Process," *Journal of Intelligent & Robotic Systems*, vol. 105, no. 3, p. 61, Jul. 2022.

[15] C. Guo, T. Bai, Y. Lu, Y. Lin, G. Xiong, X. Wang, and F.-Y. Wang, "Skywork-daVinci: A novel CPSS-based painting support system," in *2020 IEEE 16th International Conference on Automation Science and Engineering*, Aug. 2020, pp. 673–678.

[16] Y. Dou, C. Guo *et al.*, "Towards parallel intelligence based robotic painting," *Journal of Intelligent Science and Technology*, vol. 1, no. 3, pp. 27–31, 2021.

[17] C. Guo, Y. Dou *et al.*, "Human-cyber-physical hybrid painting creation in metaverses," *International Journal of Intelligent Control and Systems*, vol. 2, no. 1, pp. 29–34, 2022.

[18] F.-Y. Wang, "Artificial societies, computational experiments, and parallel systems: A discussion on computational theory of complex social-economic systems," *Complex Systems and Complexity Science*, vol. 1, no. 4, pp. 25–35, 2004.

[19] F.-Y. Wang, "The emergence of intelligent enterprises: From cps to cpss," *IEEE Intelligent Systems*, vol. 25, no. 4, pp. 85–88, 2010.

[20] K. Wang, C. Gou, N. Zheng, J. M. Rehg, and F.-Y. Wang, "Parallel vision for perception and understanding of complex scenes: methods, framework, and perspectives," *Artificial Intelligence Review*, vol. 48, pp. 299–329, 2017.

[21] P. Ye, X. Wang, W. Zheng, Q. Wei, and F.-Y. Wang, "Parallel cognition: hybrid intelligence for human-machine interaction and management," *Frontiers of Information Technology & Electronic Engineering*, vol. 23, no. 12, pp. 1765–1779, 2022.

[22] J. Lu, X. Wang, X. Cheng, J. Yang, O. Kwan, and X. Wang, "Parallel factories for smart industrial operations: From big ai models to field foundational models and scenarios engineering," *IEEE/CAA Journal of Automatica Sinica*, vol. 9, no. 12, pp. 2079–2086, 2022.

[23] F.-Y. Wang, J. Guo, G. Bu, and J. J. Zhang, "Mutually trustworthy human-machine knowledge automation and hybrid augmented intelligence: mechanisms and applications of cognition, management, and control for complex systems," *Frontiers of Information Technology & Electronic Engineering*, vol. 23, no. 8, pp. 1142–1157, 2022.

[24] L. Ouyang, J. Wu, X. Jiang, D. Almeida, C. L. Wainwright, P. Mishkin, C. Zhang, S. Agarwal, K. Slama, A. Ray *et al.*, "Training language models to follow instructions with human feedback," *arXiv preprint arXiv:2203.02155*, 2022.

[25] T. Brown, B. Mann, N. Ryder, M. Subbiah, J. D. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry *et al.*, "Language models are few-shot learners," in *Advances in Neural Information Processing Systems*, H. Larochelle, M. Ranzato, R. Hadsell, M. Balcan, and H. Lin, Eds., vol. 33. Curran Associates, Inc., 2020, pp. 1877–1901.

[26] K. Zhou, J. Yang, C. C. Loy, and Z. Liu, "Learning to prompt for vision-language models," *International Journal of Computer Vision*, vol. 130, no. 9, pp. 2337–2348, 2022.

[27] Y. Lu, C. Guo, X. Dai, and F.-Y. Wang, "Data-efficient image captioning of fine art paintings via virtual-real semantic alignment training," *Neurocomputing*, vol. 490, pp. 163–180, 2022.

[28] Y. Lu, C. Guo, Y. Lin, F. Zhuo, and F. Wang, "Computational aesthetics of fine art paintings: The state of the art and outlook," *Acta Automatica Sinica*, vol. 46, no. 11, pp. 2239–2259, 2020.

[29] Y. Lu, C. Guo, X. Dai, and F.-Y. Wang, "Artcap: A dataset for image captioning of fine art paintings," *IEEE Transactions on Computational Social Systems*, 2022.

[30] A. Radford, J. W. Kim, C. Hallacy, A. Ramesh, G. Goh, S. Agarwal, G. Sastry, A. Askell, P. Mishkin, J. Clark *et al.*, "Learning transferable visual models from natural language supervision," in *Proceedings of the International Conference on Machine Learning*. PMLR, 2021, pp. 8748–8763.

[31] J.-B. Alayrac, J. Donahue, P. Luc, A. Miech, I. Barr, Y. Hasson, K. Lenc, A. Mensch, K. Millican, M. Reynolds *et al.*, "Flamingo: a visual language model for few-shot learning," *arXiv preprint arXiv:2204.14198*, 2022.

[32] S. Reed, K. Zolna, E. Parisotto, S. G. Colmenarejo, A. Novikov, G. Barth-Maron, M. Gimenez, Y. Sulsky, J. Kay, J. T. Springenberg *et al.*, "A generalist agent," *arXiv preprint arXiv:2205.06175*, 2022.

[33] R. Bommasani, D. A. Hudson, E. Adeli, R. Altman, S. Arora, S. von Arx, M. S. Bernstein, J. Bohg, A. Bosselut, E. Brunskill *et al.*, "On the opportunities and risks of foundation models," *arXiv preprint arXiv:2108.07258*, 2021.

[34] R. Rombach, A. Blattmann, D. Lorenz, P. Esser, and B. Ommer, "High-resolution image synthesis with latent diffusion models," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, pp. 10 684–10 695.

ABOUT THE AUTHOR

Chao Guo is currently an Assistant Professor with the State Key Laboratory for Management and Control of Complex Systems, Institute of Automation, Chinese Academy of Sciences. His research focuses on computer vision for art, intelligent robotic systems, and machine learning.

Yue Lu is currently pursuing the Ph.D. degree with the State Key Laboratory for Management and Control of Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing, China, and the University of Chinese Academy of Sciences, Beijing. His research interests cover machine learning, few-shot learning, computational aesthetics, and style transfer.

Yong Dou is currently pursuing the Ph.D. degree with the Macao Institute of Systems Engineering, Macau University of Science and Technology, Macau, China. His research interests include robotic art, intelligent robotic systems, and machine learning.

Fei-Yue Wang (Fellow, IEEE) is chief scientist and the founding Director of the SKL-MCCS, CAS, and also with the Macau University of Science and Technology, Macao, China. His research interests include parallel intelligence, social computing, knowledge automation, blockchain and DAO, complex systems and complexity, and robotics and automation. Bio of Professor Fei-Yue Wang can be found at https://ieeexplore.ieee.org/author/37277656000.