Jie HUANG, Zhibin MO, Zhenyi ZHANG, Yutao CHEN, 2022. Behavioral control task supervisor with memory based on reinforcement learning for human–multi-robot coordination systems. *Frontiers of Information Technology & Electronic Engineering*, 23(8):1174-1188. https://doi.org/10.1631/FITEE.2100280

Behavioral control task supervisor with memory based on reinforcement learning for human–multi-robot coordination systems

Key words: Human–multi-robot coordination systems; Null-space-based behavioral control; Task supervisor; Reinforcement learning; Knowledge base

Corresponding author: Yutao CHEN
E-mail: yutao.chen@fzu.edu.cn
ORCID: https://orcid.org/0000-0001-6748-2866

Motivation

1. Existing human–multi-robot coordination systems (HMRCSs) suffer from high decision-making time cost and large task tracking errors caused by repeated human intervention, which restricts the autonomy of multi-robot systems (MRSs).

2. Most task supervisors in the null-space-based behavioral control (NSBC) framework need to formulate many priority-switching rules manually, which makes it difficult to realize an optimal behavioral priority adjustment strategy in the case of multiple robots and multiple tasks.

3. Deep reinforcement learning methods show great potential in HMRCSs.

Main idea

1. The novel HMRCS is designed under the NSBC-based framework.

2. A deep Q-network (DQN) with dueling structure is designed to address the problems of dynamic task priority adjustment in HMRCSs.

3. Long short-term memory (LSTM) knowledge base is designed within an NSBC framework to address the problems of repeated human intervention in HMRCSs.

4. An artificial potential field model is introduced to improve the state selection accuracy of the reinforcement learning task supervisor (RLTS).

Method

Behavioral control task supervisor with memory

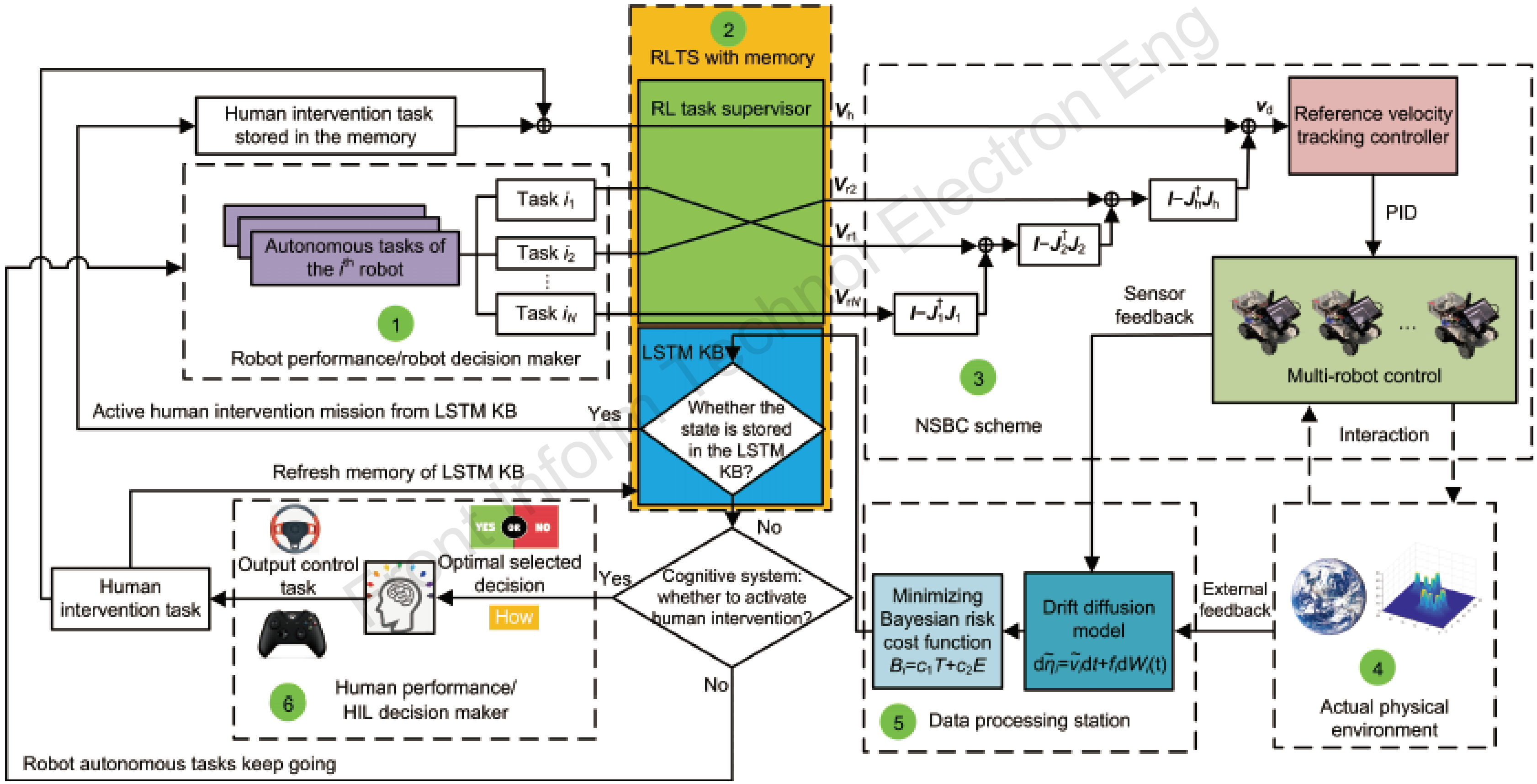

Fig. 1 The novel human–multi-robot coordination system (HMRCS) under the null-space-based behavioral control (NSBC) based framework (KB: knowledge base)

Major results

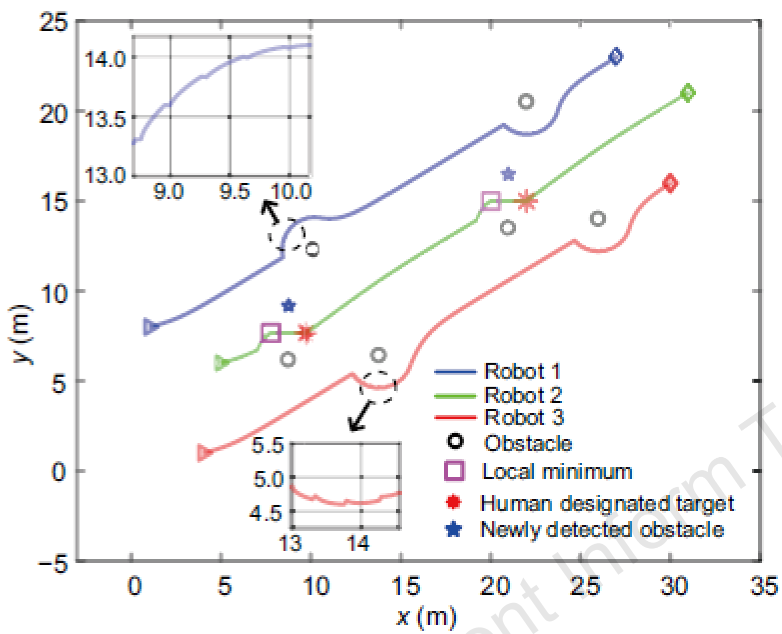

Fig. 3 Trajectories of the robots using the FSA-based task supervisor with two human intervention processes in the HMRCS

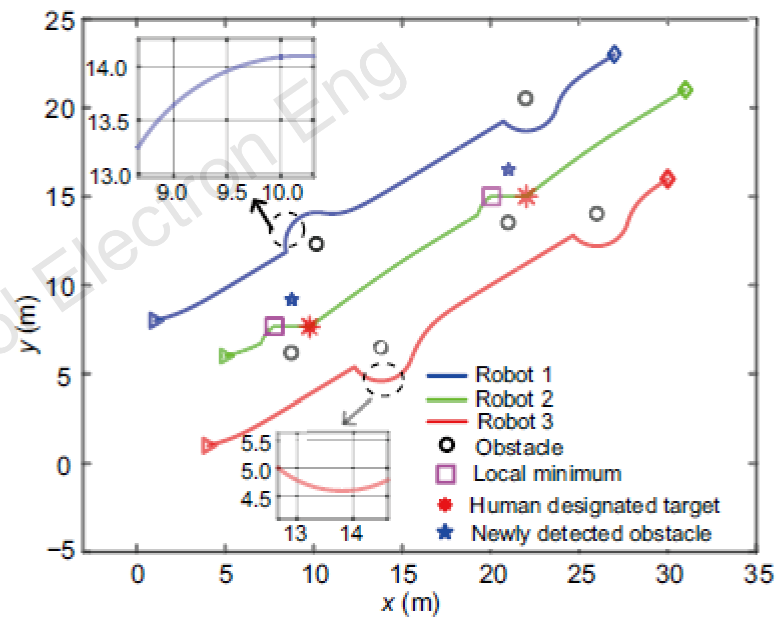

Fig. 4 Trajectories of the robots using the proposed RLTS with memory with one human intervention process and one reloading intervention process in the HMRCS

Major results (Cont'd)

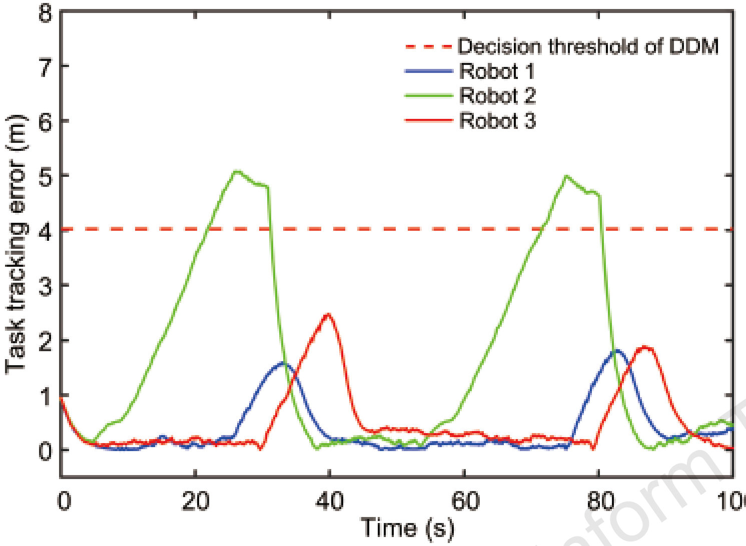

Fig. 5 Task tracking error of the robots using the FSA-based task supervisor with two human intervention processes in the HMRCS

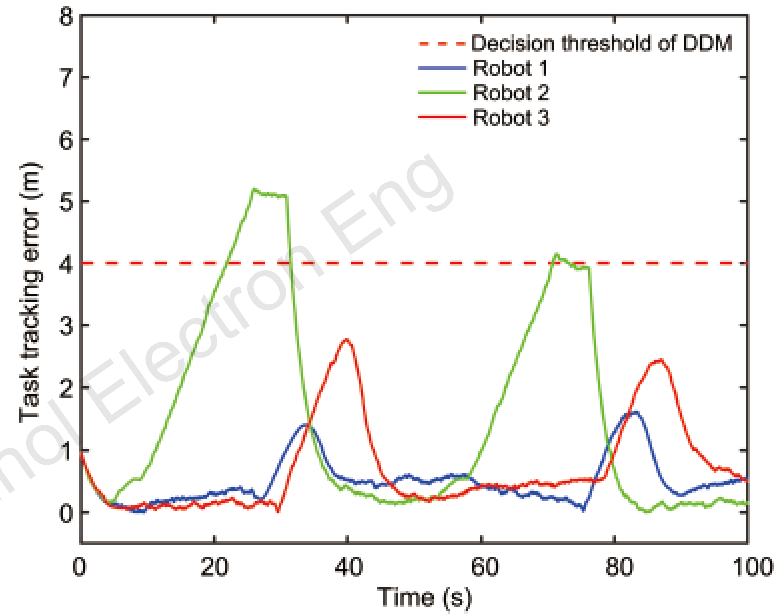

Fig. 6 Task tracking error of the robots using the proposed RLTS with memory with one human intervention process and one reloading intervention process in the HMRCS

Table 2 Performance comparison

Parameter	Value	
	FSA	RLTS
Frequency of task switch (Hz)	2.63	0.19
Number of task switches	28	4
Number of intervention tasks	2	1
Number of reloading tasks	0	1
Decision waiting time (s)	17.26	3.75

Major results (Cont'd)

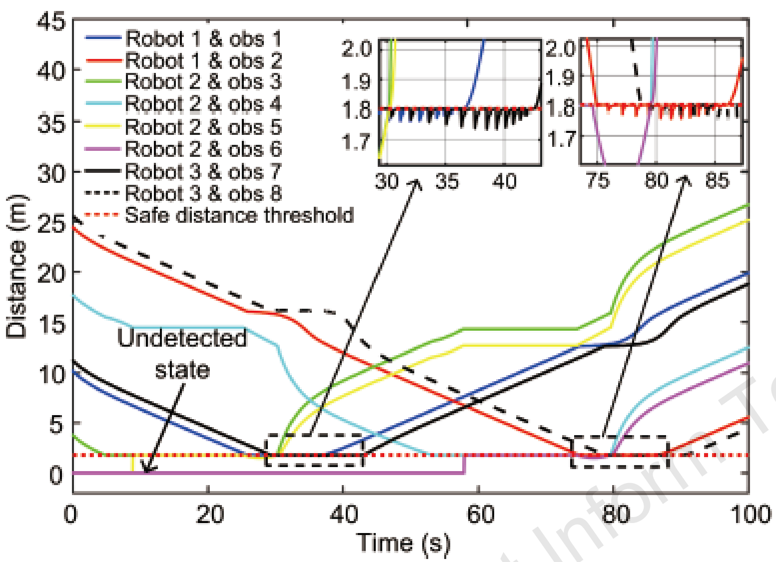

Fig. 7 Distance between the robots and the detected obstacles using the FSA-based task supervisor with two human intervention processes in the HMRCS

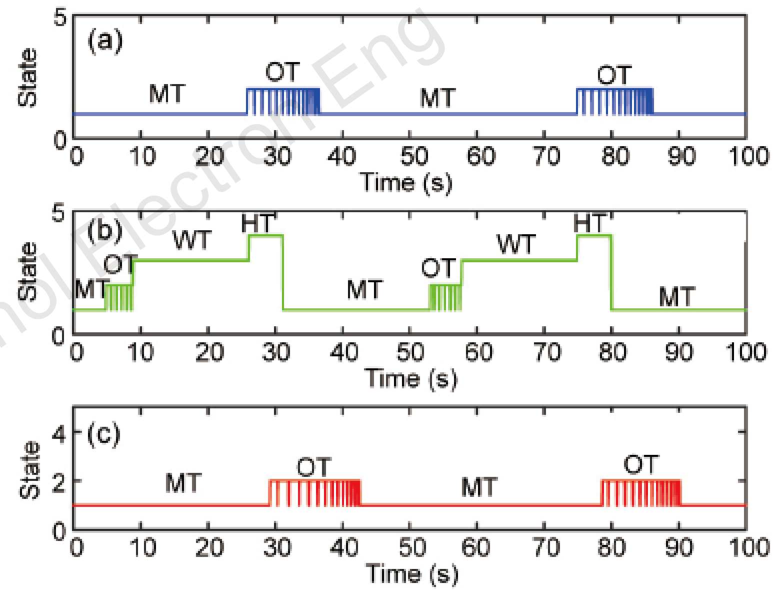

Fig. 8 Robot task mode using the FSA-based task supervisor in the HMRCS with two human intervention processes: (a) robot 1; (b) robot 2; (c) robot 3

MT: motion task; OT: obstacle avoidance task; WT: waiting for human intervention task; HT: human intervention task

Major results (Cont'd)

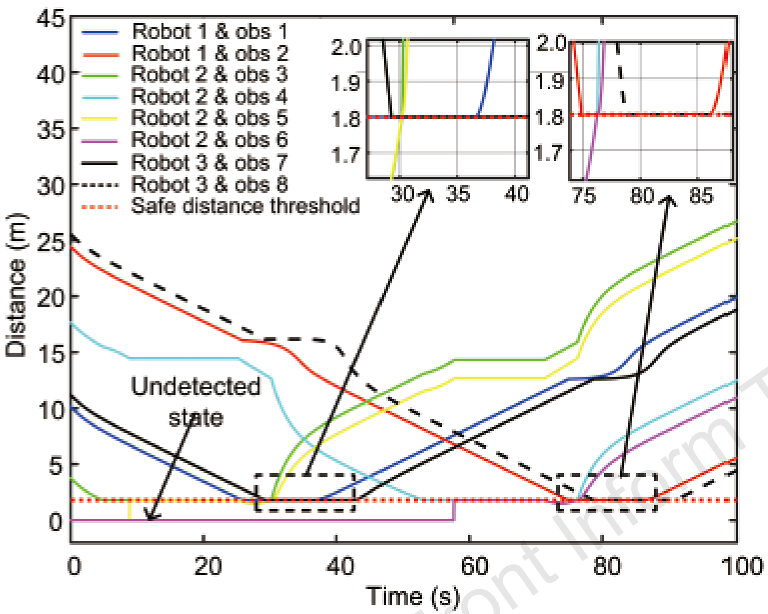

Fig. 9 Distance between the robots and the detected obstacles using the proposed RLTS with memory with one human intervention process and one reloading intervention process in the HMRCS

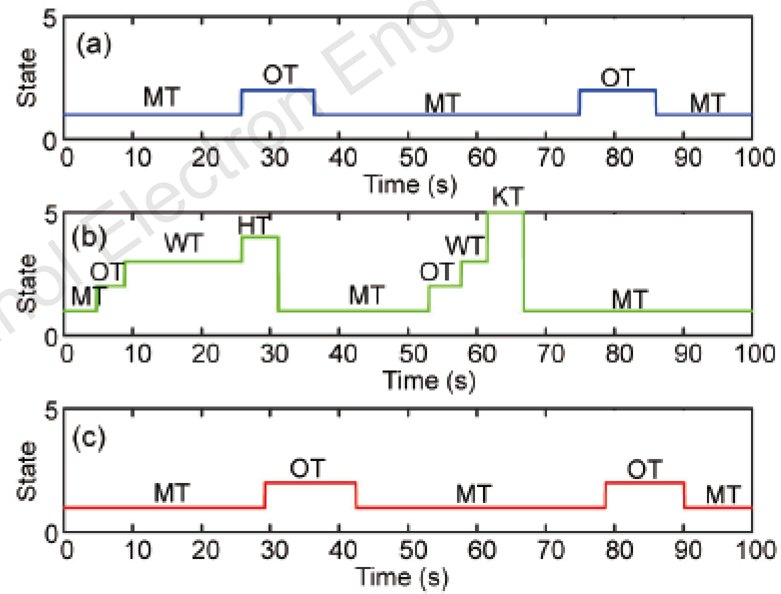

Fig. 10 Robot task mode using the proposed RLTS with memory in the HMRCS with one human intervention process and one reloading intervention process: (a) robot 1; (b) robot 2; (c) robot 3

MT: motion task; OT: obstacle avoidance task; WT: waiting for human intervention task; HT: human intervention task; KT: knowledge base reloading task

Major results (Cont'd)

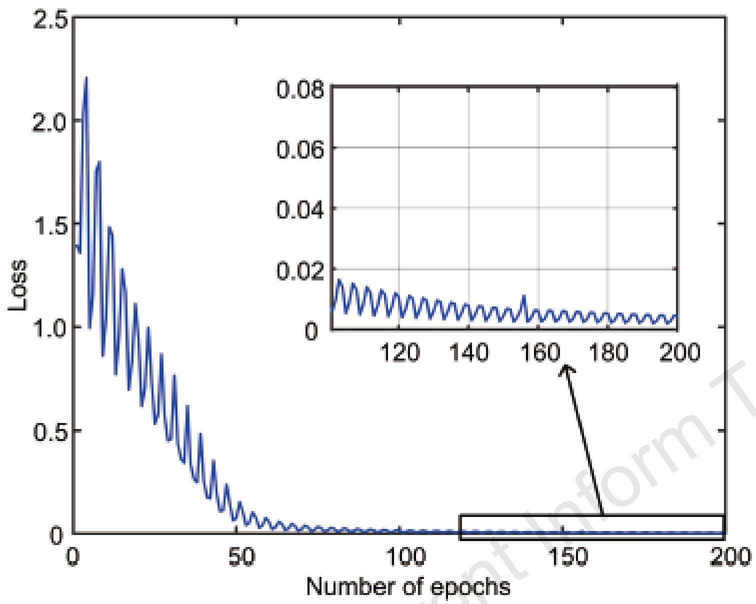

Fig. 11 LSTM network training loss of the RLTS with memory during 200 epochs (the training process is carried out after a human intervention task)

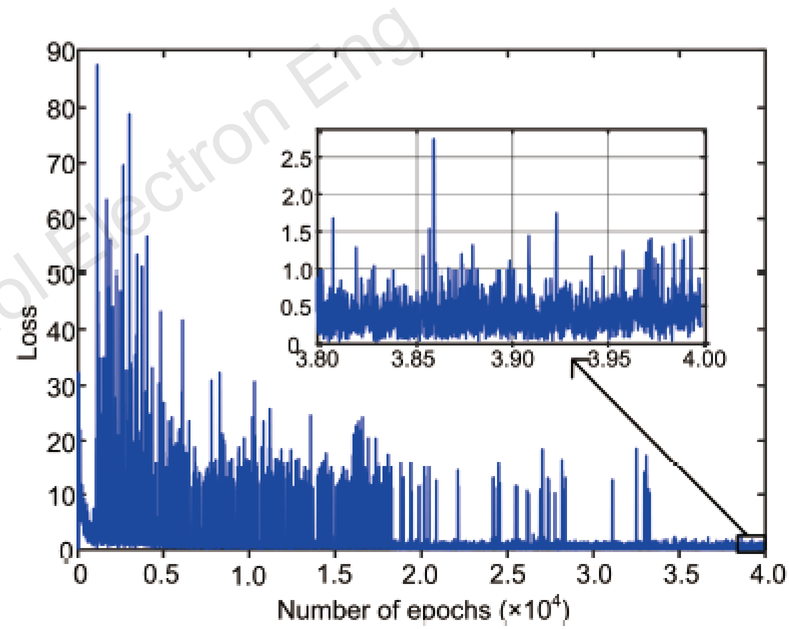

Fig. 12 Training loss of the RLTS with memory during 40 000 epochs

Major results (Cont'd)

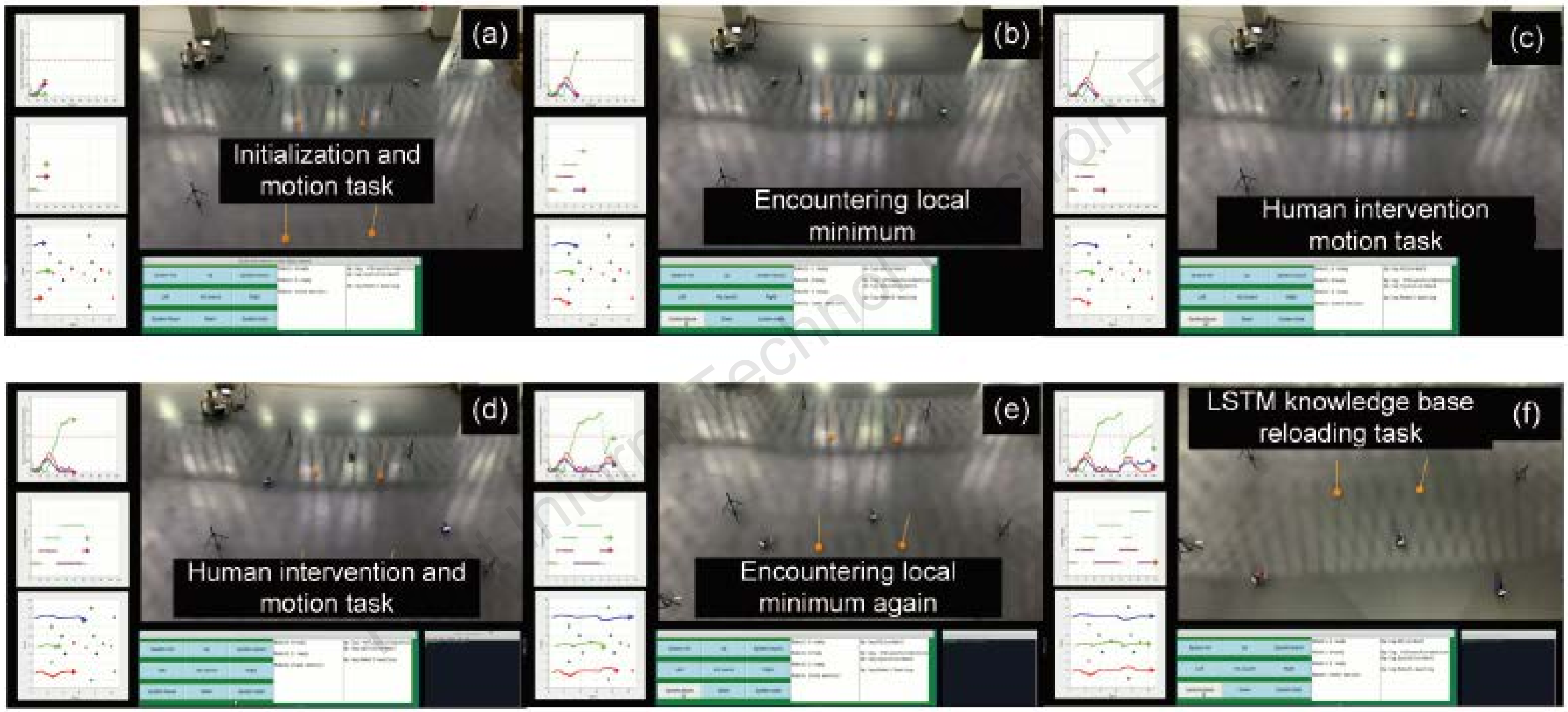

Fig. 16 Snapshots of the HMRCS experiment at 0 s (a), 36 s (b), 67 s (c), 78 s (d), 88 s (e), and 100 s (f)

Conclusions

1. The proposed RLTS with memory integrates DQN and LSTM knowledge base within an NSBC framework. Task switching frequency is reduced by 93%, and decision waiting time is reduced by 78%.

2. RLTS can successfully memorize the human intervention history and reload human control input when robots are not confident, greatly improving the robustness and flexibility of the HMRCS.

3. Simulations and experiments are conducted in an uncertain real-world environment to demonstrate the effectiveness of the proposed RLTS.

Yutao CHEN received his BE degree in automation from Hunan University, China, in 2012, MS degree from the School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, China, in 2014, and PhD degree from the Department of Information Engineering, University of Padova, Italy, in 2018. From 2019 to 2020, he was a postdoctoral researcher with the Department of Electrical Engineering, Eindhoven University of Technology, the Netherlands. He is currently an associate professor with the College of Electrical Engineering and Automation, Fuzhou University, China. His research interests include model predictive control algorithms and unmanned intelligent systems and applications.

Jie HUANG received his BE degree in electrical engineering and automation and ME degree in control engineering from Fuzhou University, China, in 2005 and 2010, respectively, and PhD degree in control science and engineering from Beijing Institute of Technology, Beijing, China, in 2015. From 2005 to 2015, he was a lecturer with Fujian Institute of Education, Fuzhou, China. From 2014 to 2018, he was a post-doctor and lecturer with the Faculty of Science and Engineering, University of Groningen, the Netherlands. He is currently a full professor of robotic and control with the College of Electrical Engineering and Automation, Fuzhou University, China. He is the vice-president of the Fujian Automation Association, Fujian Province, China. His research interests include autonomous robots, complex network dynamics, and multi-agent systems.