Zhicheng WANG, Xin ZHAO, Meng Yee (Michael) CHUAH, Zhibin LI, Jun WU, Qiuguo ZHU, 2025. Efficient learning of robust multigait quadruped locomotion for minimizing the cost of transport. *Frontiers of Information Technology & Electronic Engineering*, 26(9):1679-1691. https://doi.org/10.1631/FITEE.2401070

Efficient learning of robust multigait quadruped locomotion for minimizing the cost of transport

Key words: Reinforcement learning; Locomotion; Motor learning; Energy efficiency

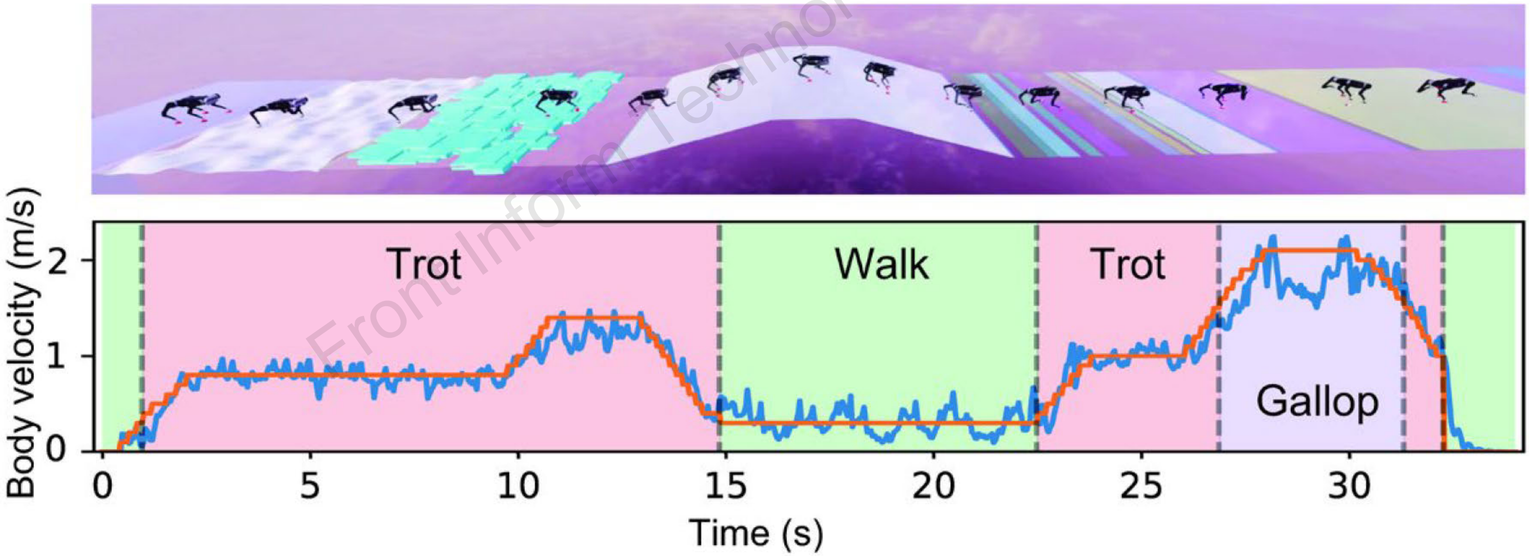
Corresponding author: Xin ZHAO
E-mail: pekingbit@163.com

Motivation

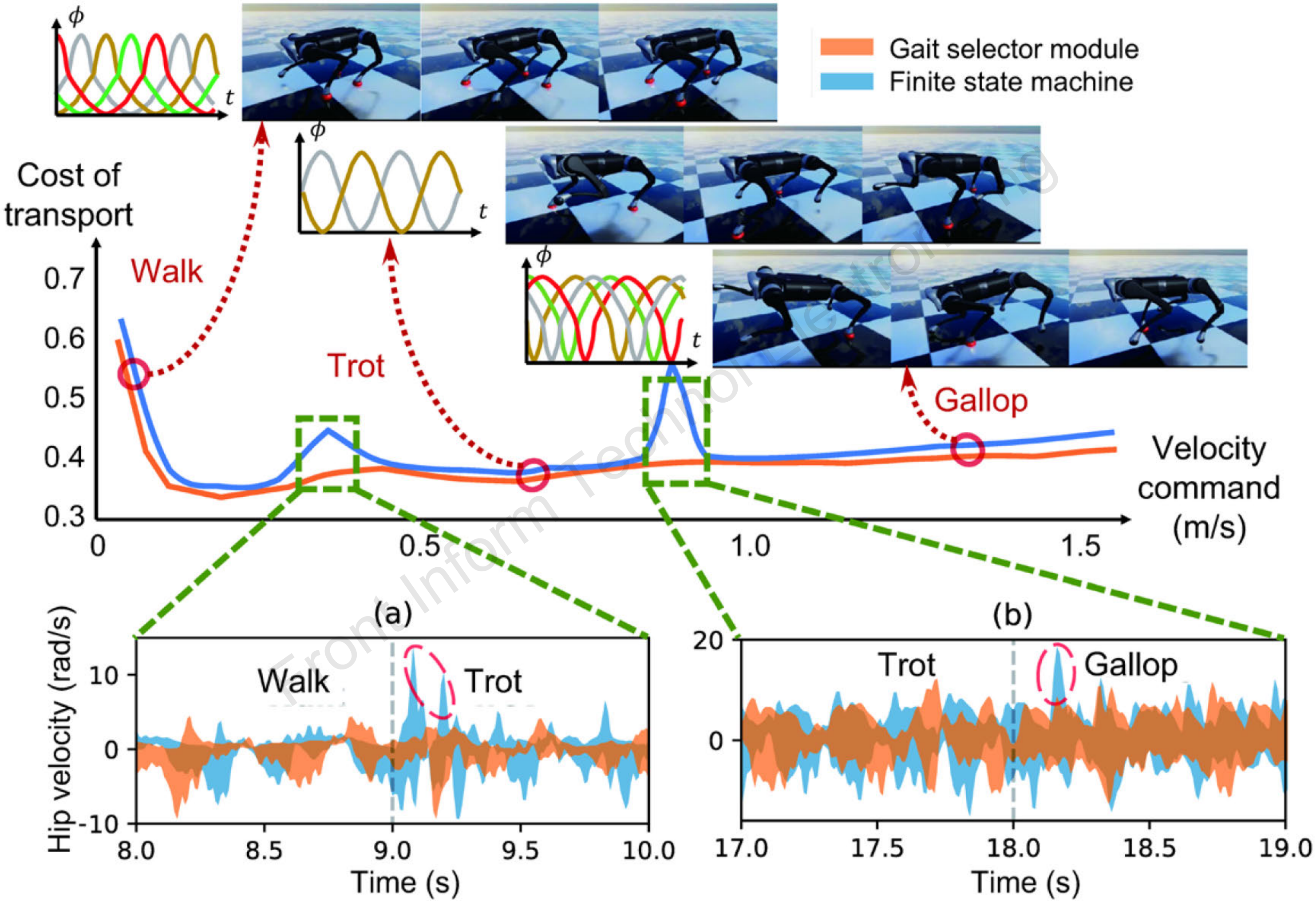
- Quadruped robots have been deployed to various application scenes. However, their locomotion behavior still lacks adaptability and energy efficiency compared to animals.
- Existing methods fix to one gait or leave gait choice uncontrolled, which leads to poor efficiency and controllability.
- We still need a framework for smooth, energy-efficient, and robust multigait locomotion.

Main idea

Develop a reinforcement-learning-based multigait locomotion framework that enables controllable, smooth, and energy-efficient gait switching.



Framework



Overview of the efficient gait-conditioned locomotion learning framework

Method

1. Gait pattern generator design

To achieve controllable gait patterns within a unified framework, we design a phase generator that yields periodic signals based on sine waves to specify gait phase commands. By modulating the period and offset of every leg, the robot can learn to perform any desired gait.

$$\phi_i = \begin{cases} \dfrac{\pi}{\gamma_i\, T_i} t + \Delta\phi_i & 0 \leq\ mod(t, T_i)\ < \gamma_i \\ \dfrac{\pi t}{(1-\gamma_i)T_i} + \dfrac{1-2\gamma_i}{1-\gamma_i}\pi + \Delta\phi_i & \gamma_i \leq mod(t, T_i) <\ T_i \end{cases}$$

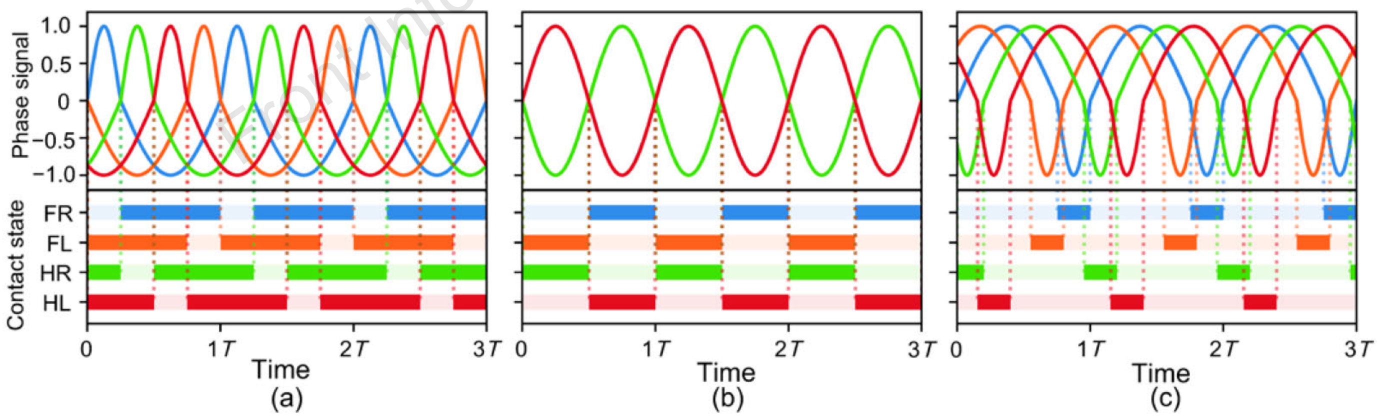


Fig. 3 Gait phase signals and foot contact sequence of different gaits: (a) quasi-static walk; (b) trot; (c) gallop

Method

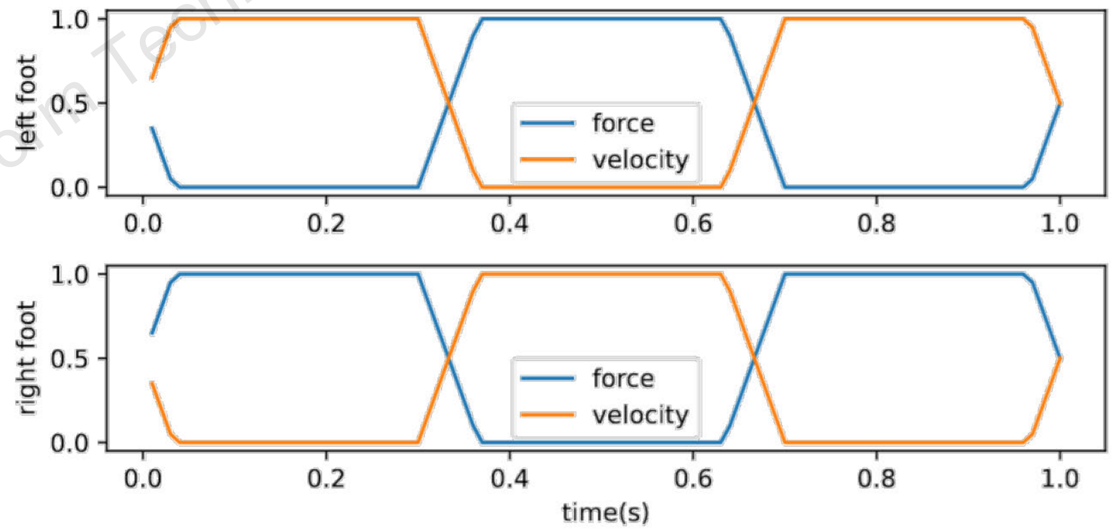
2. Gait-encouraging reward design

To make sure the robot perform accurate gait as commanded, a pair of rewards is designed. It penalizes contact force on end effectors during the expected swing phase and penalizes relative movement to the terrain during the expected stance phase.

$$R_{swing} = F_{\alpha_1,\beta_1}\left(\sum_i C_i \|F_i\|\right)$$

$$R_{stance} = F_{\alpha_2,\beta_2}\left(\sum_i (1 - C_i)\|v_i\|\right)$$

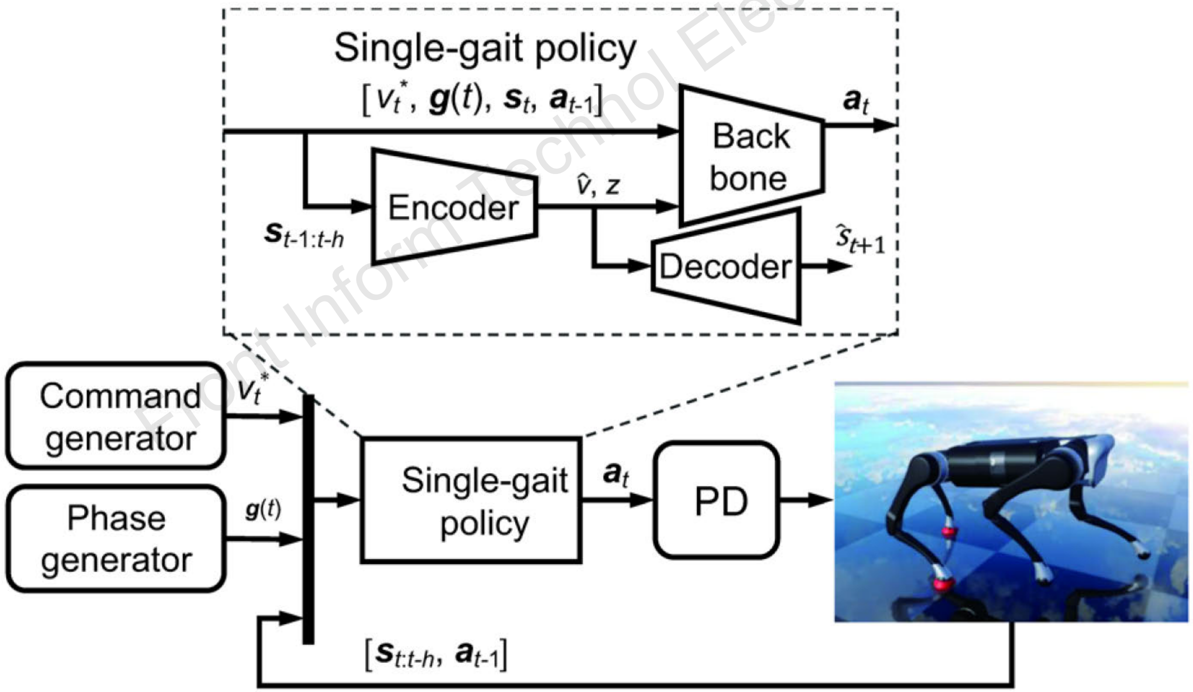
$$F_{\alpha,\beta}(x) = \alpha \exp\left(-\frac{x}{\beta}\right)^2$$



Method

3. Single-gait policy training

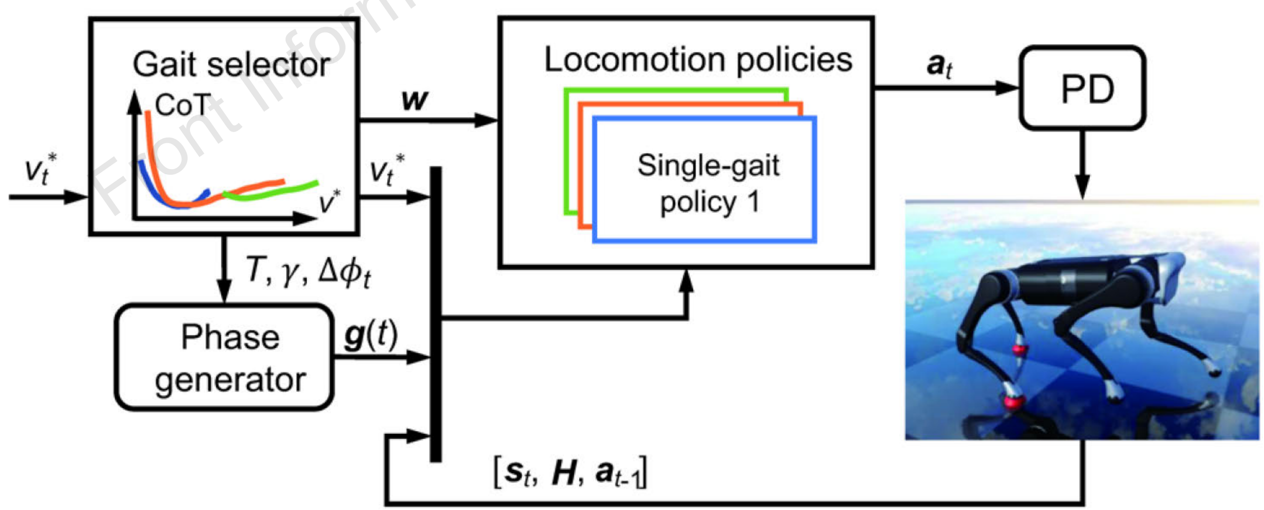
In single-gait training, each gait type is trained as an independent policy, conditioned on velocity commands and gait phase signals.



Method

4. Gait selector module training

The gait selector is trained with reinforcement learning while freezing the single-gait policies, taking as input the velocity command, proprioceptive history, and terrain information. Its reward encourages accurate velocity tracking, smooth parameter transitions, reduced joint velocity spikes, and low energy consumption, enabling adaptive and efficient gait switching.



Major results

Results in the table below demonstrate that all single-gait policies are capable of effectively tracking velocity and gait command signals, highlighting the robustness of our training framework in enabling the policies to master locomotion skills. Meanwhile, each single-gait policy exhibits trade-offs across different metrics.

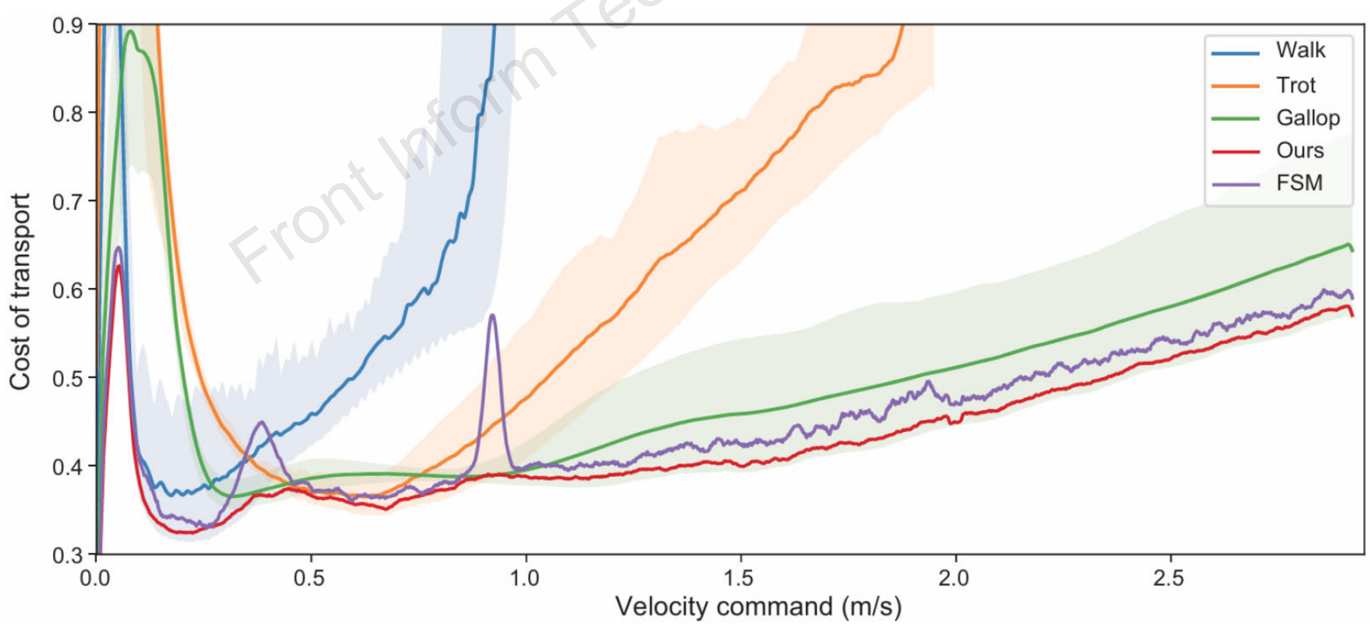
Table 6 Performance summary under different conditions

Parameter	Walk		Trot		Gallop		FSM	Ours
	0.8 Hz	1.6 Hz	1.2 Hz	2.4 Hz	2.0 Hz	3.2 Hz		
RMSE(Δv_{L})	0.1034	**0.0920**	0.0954	0.0982	0.0966	0.0941	0.1125	0.0929
RMSE(Δv_{M})	0.3768	0.1430	0.1365	0.0860	0.0621	0.0625	0.0827	**0.0620**
RMSE(Δv_{H})	0.8916	0.4997	0.3042	0.0880	0.0760	**0.0500**	0.0650	0.0522
$\mathrm{CoT}_{\mathrm{avg}}$	0.7241	0.4733	0.6886	0.6041	0.5675	0.4514	0.4531	**0.4306**
$\mathrm{TSR}_{\mathrm{fractal}}$	0.90	0.95	**1.00**	**1.00**	0.80	0.85	**1.00**	**1.00**
$\mathrm{TSR}_{\mathrm{slope}}$	0.85	0.85	**1.00**	**1.00**	0.65	0.70	**1.00**	**1.00**
$\mathrm{TSR}_{\mathrm{discrete}}$	0.80	0.95	**1.00**	0.95	0.70	0.60	0.75	**1.00**
$\mathrm{TSR}_{\mathrm{gaps}}$	0.80	0.80	0.90	**1.00**	0.80	0.80	**1.00**	0.95
$\mathrm{TSR}_{\mathrm{projectile}}$	0.95	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**

The best results are in bold

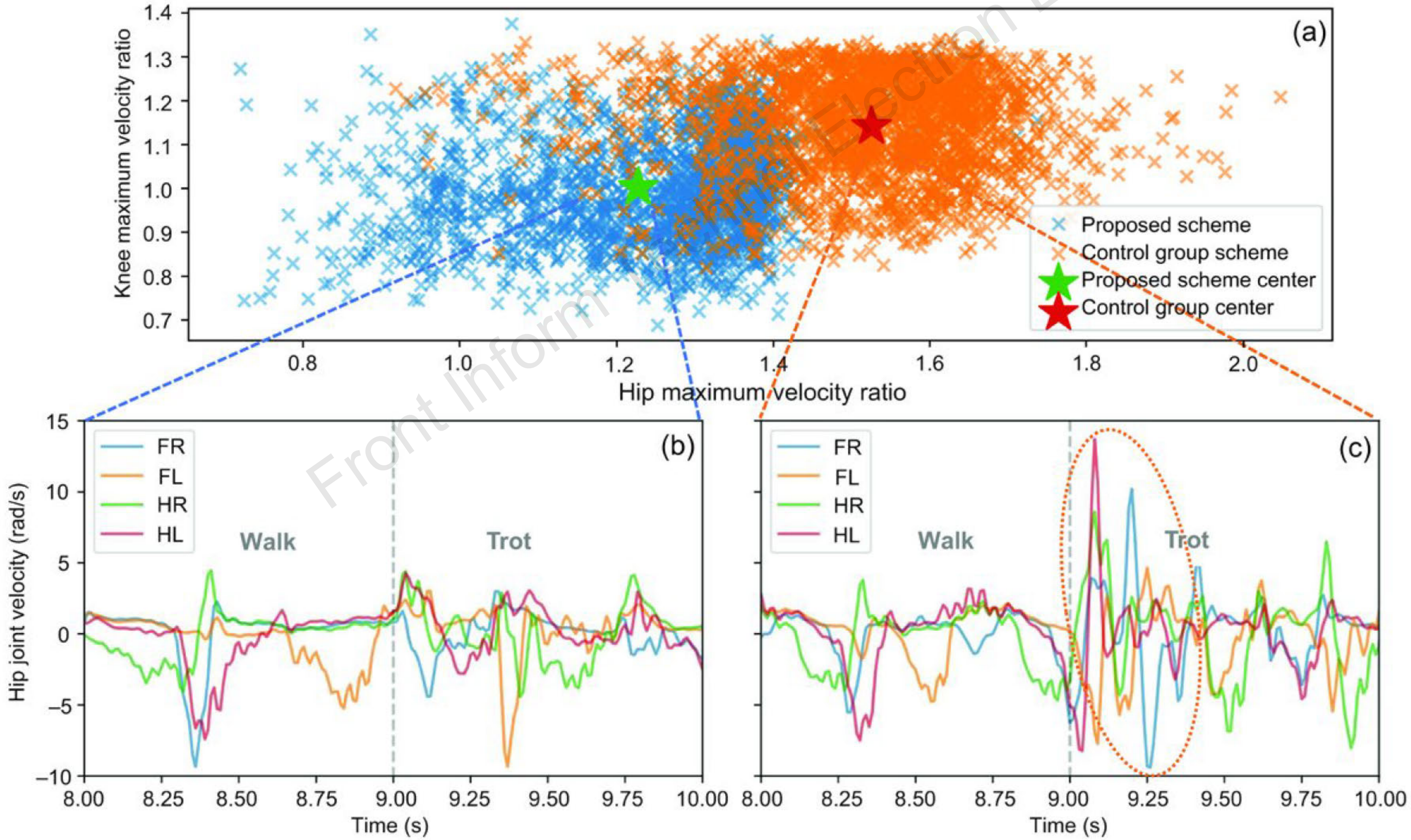
Major results

Both the proposed method and the finite state machine (FSM) effectively leverage the advantages of all single-gait policies. The velocity tracking accuracy of both groups is comparable to that of the best-performing single-gait policies, with the proposed method slightly outperforming FSM. Additionally, the proposed policy demonstrates superior adaptability to challenging terrains compared to FSM.



Major results

The proposed method is safer than FSM since the results show that the proposed method causes less joint fluctuation during gait switching.

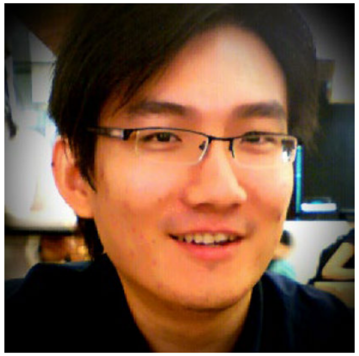


Conclusions

We present a learning framework for training and organizing robust gait-conditioned policies, revealing the relationship between velocity command and energy efficiency across different gaits. Based on these findings, an adaptive multimodal locomotion controller is developed, and it outperforms single-gait policies in terms of velocity tracking accuracy, cost of transport (CoT), and robustness.

Zhicheng Wang is a Ph.D. candidate with the Robotics & Machine Intelligence Lab at Zhejiang University (ZJU), China since 2020. She received the B.S. degree in Automation from ZJU in 2020. Between 2022 and 2023, she worked as a research assistant at the Institute for Infocomm Research (I2R), Agency for Science, Technology and Research (A*STAR), Singapore. Her research focuses mainly on learning-based locomotion for legged robots, especially humanoids. Supervised by Dr. Qiuguo Zhu, she participated in the Wukong project since 2019 and helped develop and build the Wukonmg-IV humanoid.

Xin Zhao is an Associate Researcher in the China North Vehicle Research Institute, Chinese Scholartree Ridge State Key Laboratory. She had been with Beijing Institute of Technology since 2010, where she obtained the B.S. and Ph.D. degrees. During 2017-2019, she worked as a visited Ph.D. candidate in Max-Planck Institute for Polymer Research. She worked as a post-doctoral researcher in Institute of Mechanics, Chinese Academy of Sciences. Her past work focused on modelling building, especially in boundary layer theory. Now she is interested in the combination of deep reinforcement learning and its application in fluid machinery, especially in robotics.



Meng Yee (Michael) Chuah is a Research Scientist with the Robotics & Autonomous Systems (RAS) department at the Institute for Infocomm Research (I2R), Agency for Science, Technology and Research (A*STAR), Singapore since 2018. He received his B.S. degree in Mechanical Engineering with a minor in Robotics from Carnegie Mellon University, Pittsburgh, PA, and both his M.S. and Ph.D. degrees in Mechanical Engineering from MIT, Cambridge, MA. As part of the research and development team in the Biomimetic Robotics Laboratory, Dr. Chuah helped design, develop, and build the MIT Cheetah robot with Prof. Sangbae Kim.

Zhibin Li is an Associate Professor in the University College London. He had been with the Italian Institute of Technology (IIT) since 2009, where he obtained the joint PhD degree from IIT and University of Genova, worked as a post-doctoral researcher, and led the locomotion group inside the Humanoid and Human Centred Mechatronics Lab at the Department of Advanced Robotics, IIT. His past work focused on the development of a variety of balancing and dynamic walking control of humanoid and other legged robots. His current research aims to control the dynamic behavior of a wide variety of robotic systems and to achieve human-comparable ability to move, manipulate, and grasp.

Qiuguo Zhu received the B.E. degree in mechanical engineering in 2008, the M.E. and Ph.D. degrees in control science and engineering from Zhejiang University, Hangzhou, China, in 2011 and 2020, respectively. He is currently an Associate Professor with the Institute of Cyber-Systems and Control, College of Control Science and Engineering, Zhejiang University. His research interests include humanoid robot and machine intelligence.

Jun Wu received the B.S. and Ph.D. degrees in industrial automation from Huazhong University of Science and Technology, Wuhan, China, in 1989 and 1994, respectively. Since 1994, he has been a Faculty Member with the College of Control Science and Engineering, Zhejiang University, where he is currently a Professor. He has authored or coauthored more than 60 papers in journals and conference proceedings. He was the recipient of the 1997/1998 IEE Heaviside Premium. His research interests include finite-precision digital controller design, networked control, model reduction, and robust control and optimization.