

**40th AAAI Conference on Artificial Intelligence (AAAI-26), 38th Conference on Innovative Applications of Artificial Intelligence (IAAI-26), and 16th Symposium on Educational Advances in Artificial Intelligence (EAAI-26)**

Volume 22: AAAI Technical Tracks

- Intelligent Robots

Singapore  
20-27 January 2026

ISBN: 979-8-3313-3396-6

**Printed from e-media with permission by:**

Curran Associates, Inc.  
57 Morehouse Lane  
Red Hook, NY 12571



**Some format issues inherent in the e-media version may also appear in this print version.**

Copyright© (2026) by Association for the Advancement of Artificial Intelligence  
All rights reserved.

Printed with permission by Curran Associates, Inc. (2026)

For permission requests, please contact Association for the Advancement of Artificial Intelligence  
at the address below.

Association for the Advancement of Artificial Intelligence  
2275 East Bayshore Road  
Suite 160  
Palo Alto, California 94303  
USA

Phone: 1-650-328-3123  
Fax: 1-650-321-4457

<https://aaai.org/Press/press.php>

**Additional copies of this publication are available from:**

Curran Associates, Inc.  
57 Morehouse Lane  
Red Hook, NY 12571 USA  
Phone: 845-758-0400  
Fax: 845-758-2633  
Email: [curran@proceedings.com](mailto:curran@proceedings.com)  
Web: [www.proceedings.com](http://www.proceedings.com)

# TABLE OF CONTENTS

## AAAI TECHNICAL TRACK ON INTELLIGENT ROBOTICS

GRIM: Task-Oriented Grasping with Conditioning on Generative Examples .....	18118
<i>Shailesh , Alok Raj, Nayan Kumar, Priya Shukla, Andrew Melnik, Michael Beetz, Gora Chand Nandi</i>	
Dexterous Manipulation Transfer Via Progressive Kinematic-Dynamic Alignment.....	18126
<i>Wenbin Bai, Qiyu Chen, Xiangbo Lin, Jw L, Quancheng Li, Hejiang Pan, Yi Sun</i>	
H-RDT: Human Manipulation Enhanced Bimanual Robotic Manipulation.....	18135
<i>Hongzhe Bi, Lingxuan Wu, Tianwei Lin, Hengkai Tan, Zhizhong Su, Hang Su, Jun Zhu</i>	
Steering Visuomotor Policy in Open Worlds Via Cross-View Goal Alignment .....	18144
<i>Shaofei Cai, Zhancun Mu, Anji Liu, Yitao Liang</i>	
A Natural-Gradient Approach for Nonlinear Stochastic Systems with Parameter Uncertainty.....	18153
<i>Liang Cao</i>	
AerialVLA: A Vision-Language-Action Model for Aerial Navigation with Online Dialogue .....	18161
<i>Jinyu Chen, Hongyu Li, Zongheng Tang, Xiaoduo Li, Wenjun Wu, Si Liu</i>	
PIPHEN: Physical Interaction Prediction with Hamiltonian Energy Networks .....	18170
<i>Kewei Chen, Yayu Long, Mingsheng Shang</i>	
FT-NCFM: An Influence-Aware Data Distillation Framework for Efficient VLA Models.....	18180
<i>Kewei Chen, Yayu Long, Shuai Li, Mingsheng Shang</i>	
ManiLong-Shot: Interaction-Aware One-Shot Imitation Learning for Long-Horizon Manipulation.....	18189
<i>Zixuan Chen, Chongkai Gao, Lin Shao, Jieqi Shi, Jing Huo, Yang Gao</i>	
STOLA: Self-Adaptive Touch-Language Framework for Tactile Commonsense Reasoning in Open-Ended Scenarios .....	18198
<i>Ning Cheng, Jinan Xu, Jialing Chen, Bin Fang, Wenjuan Han</i>	
PEOD: A Pixel-Aligned Event-RGB Benchmark for Object Detection Under Challenging Conditions .....	18207
<i>Luoping Cui, Hanqing Liu, Mingjie Liu, Endian Lin, Donghong Jiang, Yuhao Wang, Chuang Zhu</i>	
RflyPano: A Panoramic Benchmark for Ultra-Low Altitude UAV Localization Powered by RflySim.....	18216
<i>Dun Dai, Ze Lu, Xunhua Dai, Quan Quan</i>	
History-Enhanced Two-Stage Transformer for Aerial Vision-And-Language Navigation.....	18225
<i>Xichen Ding, Jianzhe Gao, Cong Pan, Wenguan Wang, Jie Qin</i>	
NaVLA2: A Vision-Language-Audio-Action Model for Multimodal Instruction Navigation .....	18234
<i>Jugang Fan, Peihao Chen, Changhao Li, Qing Du, Jian Chen, Mingkui Tan</i>	
MHED-SLAM: Multi-Scale Hybrid Encoding-Based Decoupled SLAM .....	18243
<i>Dengfang Feng, Wenyang Qin, Zhongchen Shi, Wei Chen, Yanhui Duan, Liang Xie, Erwei Yin</i>	

VPN: Visual Prompt Navigation .....	18253
<i>Shuo Feng, Zihan Wang, Yuchen Li, Rui Kong, Hengyi Cai, Shuaiqiang Wang, Gim Hee Lee, Piji Li, Shuqiang Jiang</i>	
Learning Diffusion Policy from Primitive Skills for Robot Manipulation .....	18262
<i>Zhihao Gu, Ming Yang, Difan Zou, Dong Xu</i>	
Just Few States Are Enough: Randomized Sparse Feedback for Stability of Dynamical Systems .....	18271
<i>Zaid Hadach, Hajar El Hammouti, El Houcine Bergou, Adnane Saoud</i>	
SeqWalker: Sequential-Horizon Vision-And-Language Navigation with Hierarchical Planning.....	18279
<i>Zebin Han, Xudong Wang, Baichen Liu, Qi Lyu, Zhenduo Shang, Jiahua Dong, Lianqing Liu, Zhi Han</i>	
Learning Object-Centric Motion Priors from Human for Robotic Dexterous Manipulation.....	18288
<i>Zhengdong Hong, Guofeng Zhang</i>	
LOG-Nav: Efficient Layout-Aware Object-Goal Navigation with Hierarchical Planning .....	18297
<i>Jiawei Hou, Yuting Xiao, Xiangyang Xue, Taiping Zeng</i>	
Real Garment Benchmark (RGBench): A Comprehensive Benchmark for Robotic Garment Manipulation Featuring a High-Fidelity Scalable Simulator.....	18306
<i>Wenkang Hu, Xincheng Tang, Yanzhi E, Yitong Li, Zhengjie Shu, Wei Li, Huamin Wang, Ruigang Yang</i>	
UNeMo: Collaborative Visual-Language Reasoning and Navigation Via a Multimodal World Model .....	18315
<i>Changxin Huang, Lv Tang, Zhaohuan Zhan, Lisha Yu, Runhao Zeng, Zun Liu, Zhengjie Wang, Jianqiang Li</i>	
GraphCoT-VLA: A 3D Spatial-Aware Reasoning Vision-Language-Action Model for Robotic Manipulation with Ambiguous Instructions .....	18324
<i>Helong Huang, Min Cen, Kai Tan, Xingyue Quan, Guowei Huang, Hong Zhang</i>	
RENEW: Risk- And Energy-Aware Navigation in Dynamic Waterways.....	18333
<i>Mingi Jeong, Alberto Quattrini Li</i>	
Towards Autonomous UAV Visual Object Search in City Space: Benchmark and Agentic Methodology .....	18342
<i>Yatai Ji, Zhengqiu Zhu, Yong Zhao, Beidan Liu, Chen Gao, Yihao Zhao, Sihang Qiu, Yue Hu, Quanjun Yin</i>	
PanoNav: Mapless Zero-Shot Object Navigation with Panoramic Scene Parsing and Dynamic Memory .....	18351
<i>Qunchao Jin, Yilin Wu, Changhao Chen</i>	
PhyPlan: Learning to Plan Tasks with Generalizable and Rapid Physical Reasoning for Embodied Manipulation .....	18360
<i>Ankit Kanwar, Hartej Soin, Abhinav Barnawal, Mudit Chopra, Harshil Vagadia, Tamajit Banerjee, Shreshth Tuli, Rohan Paul, Souvik Chakraborty</i>	
Lightweight Adaptive Topological Layout and Semantic Mapping in Vision-And-Language Navigation on Websites .....	18370
<i>Pingrui Lai, Zihao Xie, Hua Yang</i>	
DiTEA: Mixture-Of-Experts for Vision-Language-Action Model in Robotic Manipulation.....	18379
<i>Chengxuan Li, Xingwan Wang</i>	

Towards Efficient and Robust Manipulation Via Multi-Frame Vision-Language-Action Modeling.....	18388
<i>Hao Li, Shuai Yang, Yilun Chen, Xinyi Chen, Xiaoda Yang, Yang Tian, Hanqing Wang, Tai Wang, Dahua Lin, Feng Zhao, Jiangmiao Pang</i>	
SemanticVLA: Semantic-Aligned Sparsification and Enhancement for Efficient Robotic Manipulation .....	18397
<i>Wei Li, Renshan Zhang, Rui Shao, Zhijian Fang, Kaiwen Zhou, Zhuotao Tian, Liqiang Nie</i>	
LiDARCrafter: Dynamic 4D World Modeling from LiDAR Sequences.....	18406
<i>Alan Liang, Youquan Liu, Yu Yang, Dongyue Lu, Linfeng Li, Lingdong Kong, Huaici Zhao, Wei Tsang Ooi</i>	
Cook and Clean Together: Teaching Embodied Agents for Parallel Task Execution .....	18415
<i>Dingkang Liang, Cheng Zhang, Xiaopeng Xu, Jianzhong Ju, Zhenbo Luo, Xiang Bai</i>	
A3D: Adaptive Affordance Assembly with Dual-Arm Manipulation.....	18425
<i>Jiaqi Liang, Yue Chen, Qize Yu, Yan Shen, Haipeng Zhang, Hao Dong, Ruihai Wu</i>	
Whole-Body Coordination for Dynamic Object Grasping with Legged Manipulators .....	18434
<i>Qiwei Liang, Boyang Cai, Rongyi He, Hui Li, Tao Teng, Haihan Duan, Changxin Huang, Runhao Zeng</i>	
Affordance-Guided Coarse-To-Fine Exploration for Base Placement in Open-Vocabulary Mobile Manipulation .....	18443
<i>Tzu-Jung Lin, Jia-Fong Yeh, Hung-Ting Su, Chung-Yi Lin, Yi-Ting Chen, Winston H. Hsu</i>	
TTF-VLA: Temporal Token Fusion Via Pixel-Attention Integration for Vision-Language-Action Models.....	18452
<i>Chenghao Liu, Jiachen Zhang, Chengxuan Li, Zhimu Zhou, Shixin Wu, Songfang Huang, Huiling Duan</i>	
FoAM: Foresight-Augmented Multi-Task Imitation Policy for Robotic Manipulation.....	18460
<i>Litao Liu, Wentao Wang, Yifan Han, Zhuoli Xie, Pengfei Yi, Junyan Li, Wenzhao Lian</i>	
Intention-Aware Diffusion Model for Pedestrian Trajectory Prediction.....	18469
<i>Yu Liu, Zhijie Liu, Xiao Ren, Youfu Li, He Kong</i>	
SIAM: Towards Generalizable Articulated Object Modeling Via Single Robot-Object Interaction .....	18478
<i>Yuyan Liu, Li Zhang, Di Wu, Yan Zhang, Anran Huang, Zhi Wang, Liu Liu, Dan Guo</i>	
DNOI-4DRO: Deep 4D Radar Odometry with Differentiable Neural-Optimization Iterations .....	18487
<i>Shouyi Lu, Huanyu Zhou, Guirong Zhuo, Xiao Tang</i>	
TouchFormer: A Robust Transformer-Based Framework for Multimodal Material Perception.....	18496
<i>Kailin Lyu, Long Xiao, Jianing Zeng, Junhao Dong, Xuexin Liu, Zhuojun Zou, Haoyue Yang, Lin Shu, Jie Hao</i>	
UrbanNav: Learning Language-Guided Embodied Urban Navigation from Web-Scale Human Trajectories .....	18505
<i>Yanghong Mei, Yirong Yang, Longteng Guo, Qunbo Wang, Ming-Ming Yu, Xingjian He, Wenjun Wu, Jing Liu</i>	
Autonomous Vehicle Path Planning by Searching with Differentiable Simulation.....	18514
<i>Asen Nachkov, Jan-Nico Zaech, Danda Pani Paudel, Xi Wang, Luc Van Gool</i>	

Coordinated Humanoid Robot Locomotion with Symmetry Equivariant Reinforcement Learning Policy.....	18523
<i>Buqing Nie, Yang Zhang, Rongjun Jin, Zhanxiang Cao, Huangxuan Lin, Xiaokang Yang, Yue Gao</i>	
MP1: MeanFlow Tames Policy Learning in 1-Step for Robotic Manipulation .....	18532
<i>Juyi Sheng, Ziyi Wang, Peiming Li, Mengyuan Liu</i>	
Real-Time Path Planning for UAVs in Windy Environments Without Computational Fluid Dynamics.....	18540
<i>Abhudaya Shrivastava, Shelly Gupta, Zoran Obradovic</i>	
ReconVLA: Reconstructive Vision-Language-Action Model as Effective Robot Perceiver .....	18549
<i>Wenxuan Song, Ziyang Zhou, Han Zhao, Jiayi Chen, Pengxiang Ding, Haodong Yan, Yuxin Huang, Feilong Tang, Donglin Wang, Haoang Li</i>	
ManipLVM-R1: Reinforcement Learning for Reasoning in Embodied Manipulation with Large Vision-Language Models.....	18558
<i>Zirui Song, Guangxian Ouyang, Mingzhe Li, Yuheng Ji, Chenxi Wang, Zixiang Xu, Zeyu Zhang, Xiaoqing Zhang, Qian Jiang, Fengxian Ji, Zhenhao Chen, Zhongzhi Li, Xiuying Chen</i>	
VirtualEnv: A Platform for Embodied AI Research.....	18567
<i>Kabir Swain, Sijie Han, Ayush Raina, Jin Zhang, Shuang Li, Michael Stopa, Antonio Torralba</i>	
FARM: Frame-Accelerated Augmentation and Residual Mixture-Of-Experts for Physics-Based High-Dynamic Humanoid Control .....	18575
<i>Jing Tan, Shiting Chen, Yangfan Li, Weisheng Xu, Renjing Xu</i>	
WorldAgen: Unified State-Action Prediction with Test-Time World Model Training .....	18584
<i>Chi Wan, Kangrui Wang, Yuan Si, Pingyue Zhang, Manling Li</i>	
LatentVLA: Taming Latent Space for Generalizable and Long-Horizon Bimanual Manipulation .....	18593
<i>Junming Wang</i>	
ODYSSEY: Open-World Quadrupeds Exploration and Manipulation for Long-Horizon Tasks.....	18602
<i>Kaijun Wang, Liqin Lu, Mingyu Liu, Jianuo Jiang, Zeju Li, Bolin Zhang, Wancai Zheng, Xinyi Yu, Hao Chen, Chunhua Shen</i>	
PLUM-Net: Prototype-Induced Label Structuring for Disentangled Multimodal Representation Network.....	18611
<i>Kehan Wang, Huan Zhao, Yong Wei, Xupeng Zha, Guanghui Ye, Cheng Zhu, Yiming Liu, Zixing Zhang</i>	
Expand Your SCOPE: Semantic Cognition Over Potential-Based Exploration for Embodied Visual Navigation .....	18620
<i>Ningnan Wang, Weihuang Chen, Liming Chen, Haoxuan Ji, Zhongyu Guo, Xuchong Zhang, Hongbin Sun</i>	
Lifelong Language-Conditioned Robotic Manipulation Learning .....	18629
<i>Xudong Wang, Zebin Han, Zhiyu Liu, Gan Li, Jiahua Dong, Baichen Liu, Lianqing Liu, Zhi Han</i>	
VLA-Adapter: An Effective Paradigm for Tiny-Scale Vision-Language-Action Model.....	18638
<i>Yihao Wang, Pengxiang Ding, Lingxiao Li, Can Cui, Zirui Ge, Xinyang Tong, Wenxuan Song, Han Zhao, Wei Zhao, Pengxu Hou, Siteng Huang, Yifan Tang, Wenhui Wang, Ru Zhang, Jianyi Liu, Donglin Wang</i>	

Self-Supervised Multiplex Consensus Mamba for General Image Fusion .....	18647
<i>Yingying Wang, Rongjin Zhuang, Hui Zheng, Xuanhua He, Ke Cao, Xiaotong Tu, Xinghao Ding</i>	
MMMamba: A Versatile Cross-Modal in Context Fusion Framework for Pan-Sharpening and Zero-Shot Image Enhancement .....	18656
<i>Yingying Wang, Xuanhua He, Chen Wu, Jialing Huang, Suiyun Zhang, Rui Liu, Xinghao Ding, Haoxuan Che</i>	
ForeDiffusion: Foresight-Conditioned Diffusion Policy Via Future View Construction for Robot Manipulation .....	18665
<i>Weize Xie, Yi Ding, Ying He, Leilei Wang, Binwen Bai, Zheyi Zhao, Chenyang Wang, F. Richard Yu</i>	
Firing Bits Where it Matters: Spiking-Guided Just Recognizable Distortion Modeling for Machine-Centric Video Coding .....	18674
<i>Wuyuan Xie, Zhenming Li, Yuwu Lu, Di Lin, Yun Song, Miaohui Wang</i>	
Zero-Shot Robotic Manipulation Via 3D Gaussian Splatting-Enhanced Multimodal Retrieval-Augmented Generation.....	18683
<i>Zilong Xie, Jingyu Gong, Xin Tan, Zhizhong Zhang, Yuan Xie</i>	
Actor-Critic for Continuous Action Chunks: A Reinforcement Learning Framework for Long-Horizon Robotic Manipulation with Sparse Reward.....	18692
<i>Jiarui Yang, Bin Zhu, Jingjing Chen, Yu-Gang Jiang</i>	
ReflexDiffusion: Reflection-Enhanced Trajectory Planning for High-Lateral-Acceleration Scenarios in Autonomous Driving.....	18701
<i>Xuemei Yao, Xiao Yang, Jianbin Sun, Liuwei Xie, Xuebin Shao, Xiyu Fang, Hang Su, Kewei Yang</i>	
Indoor Multi-View Radar Object Detection Via 3D Bounding Box Diffusion.....	18710
<i>Ryoma Yataka, Pu (Perry) Wang, Petros Boufounos, Ryuhei Takahashi</i>	
GraphGrasp: Lightweight and Efficient Graph-Guided 6-DoF Robotic Grasp Pose Estimation Network.....	18719
<i>Sheng Yu, Di-Hua Zhai, Yuanqing Xia</i>	
Learning from Human Gaze: Human-Like Robot Social Navigation in Dense Crowds .....	18728
<i>Zhecheng Yu, Yan Lyu, Chen Yang, Tao Chen, Yishuang Zhang, Bo Ling, Peng Wang, Guanyu Gao, Weiwei Wu, Brian Y. Lim</i>	
CorrectNav: Self-Correction Flywheel Empowers Vision-Language-Action Navigation Model .....	18737
<i>Zhuoyuan Yu, Yuxing Long, Zihan Yang, Chengyan Zeng, Hongwei Fan, Jiyao Zhang, Hao Dong</i>	
DIMM: Decoupled Multi-Hierarchy Kalman Filter Via Reinforcement Learning .....	18746
<i>Jirong Zha, Yuxuan Fan, Kai Li, Han Li, Chen Gao, Xinlei Chen</i>	
Balancing Signal and Variance: Adaptive Offline RL Post-Training for VLA Flow Models .....	18755
<i>Hongyin Zhang, Shiyuan Zhang, Junxi Jin, Qixin Zeng, Yifan Qiao, Hongchao Lu, Donglin Wang</i>	
MoLe-VLA: Dynamic Layer-Skipping Vision Language Action Model Via Mixture-Of-Layers for Efficient Robot Manipulation.....	18764
<i>Rongyu Zhang, Menghang Dong, Yuan Zhang, Liang Heng, Xiaowei Chi, Gaole Dai, Li Du, Dan Wang, Yuan Du, Shanghang Zhang</i>	

RaLD: Generating High-Resolution 3D Radar Point Clouds with Latent Diffusion.....	18773
<i>Ruijie Zhang, Bixin Zeng, Shengpeng Wang, Fuhui Zhou, Wei Wang</i>	
Grounding Actions in Camera Space: Observation-Centric Vision-Language-Action Policy.....	18782
<i>Tianyi Zhang, Haonan Duan, Haoran Hao, Yu Qiao, Jifeng Dai, Zhi Hou</i>	
Agent Journey Beyond RGB: Hierarchical Semantic-Spatial Representation Enrichment for Vision- And-Language Navigation .....	18791
<i>Xuesong Zhang, Yunbo Xu, Jia Li, Ruonan Liu, Zhenzhen Hu</i>	
Keep on Going: Learning Robust Humanoid Motion Skills Via Selective Adversarial Training.....	18800
<i>Yang Zhang, Zhanxiang Cao, Buqing Nie, Haoyang Li, Zhong Jiangwei, Qiao Sun, Xiaoyi Hu, Xiaokang Yang, Yue Gao</i>	
Bridging Scale Discrepancies in Robotic Control Via Language-Based Action Representations .....	18809
<i>Yuchi Zhang, Churui Sun, Shiqi Liang, Di Yuan Liu, Chao Ji, Weinan Zhang, Ting Liu</i>	
Towards Adaptive Humanoid Control Via Multi-Behavior Distillation and Reinforced Fine-Tuning .....	18818
<i>Yingnan Zhao, Xinmiao Wang, Dewei Wang, Xinzhe Liu, Dan Lu, Qilong Han, Peng Liu, Chenjia Bai</i>	
CoEvoer: Collaborative Evolution Transformer for Upper-Body Expressive Human Pose and Shape Estimation.....	18827
<i>Yuxiang Zhao, Wei Huang, Yujie Song, Liu Wang, Huan Zhao</i>	
DexGraspVLA: A Vision-Language-Action Framework Towards General Dexterous Grasping .....	18836
<i>Yifan Zhong, Xuchuan Huang, Ruochong Li, Ceyao Zhang, Zhang Chen, Tianrui Guan, Fanlian Zeng, Ka Nam Lui, Yuyao Ye, Yitao Liang, Yaodong Yang, Yuanpei Chen</i>	
Run, Ruminant, and Regulate: A Dual-Process Thinking System for Vision-And-Language Navigation .....	18845
<i>Yu Zhong, Zihao Zhang, Rui Zhang, Lingdong Huang, Haihan Gao, Shuo Wang, Da Li, Ruijian Han, Jiaming Guo, Shaohui Peng, Di Huang, Yunji Chen</i>	
Gentle Manipulation Policy Learning Via Demonstrations from VLM Planned Atomic Skills.....	18855
<i>Jiayu Zhou, Qiwei Wu, Jian Li, Zhe Chen, Xiaogang Xiong, Renjing Xu</i>	
Collaborative Representation Learning for Alignment of Tactile, Language, and Vision Modalities .....	18864
<i>Yiyun Zhou, Mingjing Xu, Jingwei Shi, Quanjiang Li, Jingyuan Chen</i>	
Effective Robotic Cloth Grasping Through Suppressing False Discoveries .....	18873
<i>Xingyu Zhu, Zhiwen Tu, Yan Wu, Shan Luo, Hechang Chen, Yixing Gao</i>	
H-GAR: A Hierarchical Interaction Framework Via Goal-Driven Observation-Action Refinement for Robotic Manipulation .....	18882
<i>Yijie Zhu, Rui Shao, Ziyang Liu, Jie He, Jizhihui Liu, Jiuru Wang, Zitong Yu</i>	
D <sup>2</sup> PPPO: Diffusion Policy Policy Optimization with Dispersive Loss.....	18891
<i>Guowei Zou, Weibing Li, Hejun Wu, Yukun Qian, Yuhang Wang, Haitao Wang</i>	

## Author Index