Best Paper Nomination

1863 From Speaker to Dubber: Movie Dubbing with Prosody and Duration Consistency Learning

Zhedong Zhang, Liang Li, Gaoxiang Cong, Haibing YIN, Yuhan Gao, Chenggang Yan, Anton Hengel, Yuankai Qi

2569 A Picture Is Worth a Graph: A Blueprint Debate Paradigm for Multimodal Reasoning

Changmeng Zheng, DaYong Liang, Wengyu ZHANG, Xiaoyong Wei, Tat-seng Chua, Qing Li

2607 Not All Pairs are Equal: Hierarchical Learning for Average-Precision-Oriented Video Retrieval

Yang Liu, Qianqian Xu, Peisong Wen, Siran Dai, Qingming Huang

2782 U2UData: A Large-scale Cooperative Perception Dataset for Swarm UAVs Autonomous Flight

Tongtong Feng, Xin Wang, Feilin Han, Leping Zhang, Wenwu Zhu

4894 Cross-modal Observation Hypothesis Inference

Mengze Li, Kairong Han, Jiahe Xu, Li Yueying, Wu Tao, Zhou Zhao, Jiaxu Miao, Shengyu Zhang, Jingyuan Chen

5216 Navigating Beyond Instructions: Vision-and-Language Navigation in Obstructed Environments

Haodong Hong, Sen Wang, Zi Huang, Qi Wu, Jiajun Liu

2843 Gait Recognition in Large-scale Free Environment via Single LiDAR

Xiao Han, Yiming Ren, Peishan Cong, YUJING SUN, Jingya Wang, Lan Xu, Yuexin Ma

457 Hypergraph Multi-modal Large Language Model: Exploiting EEG and Eye-tracking Modalities to Evaluate Heterogeneous Responses for Video Understanding

Minghui Wu, Chenxu Zhao, Anyang Su, Donglin Di, Tianyu Fu, Da An, Min He, Ya Gao, Meng Ma, Yan Kun, Ping Wang

4318 Multi-Modal Diffusion Model for Recommendation

Yangqin Jiang, Lianghao Xia, Wei Wei, Da Luo, Kangyi Lin, Chao Huang

5592 Leveraging Knowledge of Modality Experts for Incomplete Multimodal Learning

Wenxin Xu, Hexin Jiang, xuefeng liang

981 LMM-PCQA: Assisting Point Cloud Quality Assessment with LMM

zicheng zhang, Haoning Wu, Yingjie Zhou, Chunyi Li, Wei Sun, Chaofeng Chen, Xiongkuo Min, Xiaohong Liu, Weisi Lin, Guangtao Zhai

3244 HKDSME: Heterogeneous Knowledge Distillation for Semi-supervised Singing Melody Extraction Using Harmonic Supervision

Shuai Yu, Xiaoliang He, Ke Chen, Yi Yu

3578 Improving Interaction Comfort in Authoring Task in AR-HRI through Dynamic Dual-Layer Interaction Adjustment

Yunqiang Pei, Kaiyue Zhang, Hongrong yang, Yong Tao, Qihang Tang, Jialei Tang, Guoqing Wang, Zhitao Liu, Ning Xie, Peng Wang, Yang Yang, Hengtao Shen

4352 VR-Mediated Cognitive Defusion: A Comparative Study for Managing Negative Thoughts

Kento Shigyo, Yi-Fan Cao, Kentaro Takahira, Mingming Fan, Huamin Qu

4896 Evolution-aware VAriance (EVA) Coreset Selection for Medical Image Classification

Yuxin Hong, Xiao Zhang, Xin Zhang, Joey Zhou

5 AV-Deepfake1M: A Large-Scale LLM-Driven Audio-Visual Deepfake Dataset

Zhixi Cai, Shreya Ghosh, Aman Pankaj Adatia, Munawar Hayat, Abhinav Dhall, Tom Gedeon, Kalin Stefanov

88 ResVR: Joint Rescaling and Viewport Rendering of Omnidirectional Images

Weiqi Li, Shijie Zhao, Bin Chen, Xinhua Cheng, Junlin Li, Li zhang, Jian Zhang

226 LoopGaussian: Creating 3D Cinemagraph with Multi-view Images via Eulerian Motion Field

Jiyang Li, Lechao Cheng, Zhangye Wang, Tingting Mu, Jingxuan He

2993 Investigating Conceptual Blending of a Diffusion Model for Improving Nonword-to-Image Generation

Chihaya Matsuhira, Marc A. Kastner, Takahiro Komamizu, Takatsugu Hirayama, Ichiro Ide

4096 4D Gaussian Splatting with Scale-aware Residual Field and Adaptive Optimization for Real-time rendering of temporally complex dynamic scenes

Jinbo Yan, Rui Peng, Luyang Tang, Ronggang Wang

4102 GPT4Video: A Unified Multimodal Large Language Model for Instruction-Followed Understanding and Safety-Aware Generation

Zhanyu Wang, Longyue Wang, Zhen Zhao, Minghao Wu, Chenyang Lyu, Huayang Li, Deng Cai, Luping Zhou, Shuming Shi, Zhaopeng Tu

1157 Reproducing the Past: A Dataset for Benchmarking Inscription Restoration

Shipeng Zhu, hui xue, Na Nie, Chenjie Zhu, Haiyue Liu, Pengfei Fang

1299 Video Bokeh Rendering: Make Casual Videography Cinematic

Yawen Luo, Min Shi, Liao Shen, Yachuan Huang, Zixuan Ye, Juewen Peng, Zhiguo Cao

2167 An In-depth Study of Bandwidth Allocation across Media Sources in Video Conferencing

Zejun Zhang, Xiao Zhu, Anlan Zhang, Feng Qian

4473 Loc4Plan: Locating Before Planning for Outdoor Vision and Language Navigation

Huilin Tian, Jingke Meng, Wei-Shi Zheng, Yuanming Li, Junkai Yan, Yunong Zhang

3624 Once-for-all: Efficient Visual Face Privacy Protection via Person-specific Veils

Zixuan Yang, Yushu Zhang, Tao Wang, Zhongyun Hua, Zhihua Xia, Jian Weng

Sponsors: