SIPENG ZHENG

✉ zhengsipeng27@gmail.com 🏠 zhengsipeng.github.io Ⓖ Sipeng Zheng 📞 +86 15905058181

I am a research scientist at Beijing Academy of Artificial Intelligence (BAAI). I received my PhD and bachelor degrees from Renmin University of China in 2023 and 2018 respectively under the supervision of Prof. Qin Jin. My research interest lies in human behavior understanding, vision-and-language learning, and embodied artificial intelligence.

EDUCATION

Renmin University of China   *2018.09 - 2023.06*
Ph.D in Computer Science   Supervisor: Qin Jin

Renmin University of China   *2014.09 - 2018.06*
B.S in Computer Science   Supervisor: Qin Jin

WORK EXPERIENCE

Beijing Academy of Artificial Intelligence   **Beijing, China**
Research scientist   *2023.07 - Present*
• Large-scale multi-modal pre-training to enable agents solving open-ended tasks in Minecraft.
Research intern   *2021.09 - 2022.02*
• Multi-lingual language-vision-audio pre-training.

Microsoft Research Asia   **Beijing, China**
Research intern   *2022.04 - 2022.10*
• Temporal sentence grounding for long-term videos.

PUBLICATION

1. Boshen Xu, **Sipeng Zheng**, and Qin Jin. Spaformer: Sequential 3d part assembly with transformers. Under review 2024

2. **Sipeng Zheng**, Bohan Zhou, Yicheng Feng, Ye Wang, and Zongqing Lu. Unicode: Learning a unified codebook for multimodal large language models. Under review 2024

3. BAAI Multimodal Interaction Group. Towards general computer control: A multimodal agent for red dead redemption ii as a case study. ICLR 2024 workshop

4. **Sipeng Zheng**, Jiazheng Liu, Yicheng Feng, and Zongqing Lu. Steve-eye: Equipped llm-based embodied agents with visual perception in open worlds. ICLR 2024

5. Yicheng Feng, Yuxuan Wang, Jiazheng Liu, **Sipeng Zheng**, and Zongqing Lu. Llama rider: Spurring large language models to explore the open worlds. Under review 2024

6. Qi Zhang, **Sipeng Zheng**, and Qin Jin. No-frills temporal video grounding: Multi-scale neighboring attention and zoom-in boundary detection. Under Review 2024

7. **Sipeng Zheng**, Boshen Xu, and Qin Jin. Open-category human-object interaction pre-training via language modeling framework. In *CVPR*, 2023

8. Boshen Xu, **Sipeng Zheng**, and Qin Jin. Pov: Prompt-oriented view-agnostic learning for egocentric hand-object interaction in the multi-view world. In *ACM MM*, 2023

9. Ludan Ruan, Anwen Hu, Yuqing Song, Liang Zhang, **Sipeng Zheng**, and Qin Jin. Accommodating audio modality in clip for multimodal processing. In *AAAI*, 2023

10. **Sipeng Zheng**, Shizhe Chen, and Qin Jin. Few-shot action recognition with hierarchical matching and contrastive learning. In *ECCV*, 2022

11. **Sipeng Zheng**, Shizhe Chen, and Qin Jin. Vrdformer: End-to-end video visual relation detection with transformers. In *CVPR*, 2022 (Oral)

12. **Sipeng Zheng**, Qi Zhang, and Qin Jin. Exploring anchor-based detection for ego4d natural language query. In *CVPR Ego4D workshop*, 2022

13. Bei Liu, **Sipeng Zheng**, Jianlong Fu, and Wen-Huang Cheng. Anchor-based detection for natural language localization in ego-centric videos. In *IEEC*, 2022

14. **Sipeng Zheng**, Shizhe Chen, and Qin Jin. Skeleton-based interactive graph network for human object interaction detection. In *ICME*, 2020

15. **Sipeng Zheng**, Shizhe Chen, and Qin Jin. Visual relation detection with multi-level attention. In *ACM MM*, 2019

16. **Sipeng Zheng**, Xiangyu Chen, Shizhe Chen, and Qin Jin. Relation understanding in videos. In *ACM MM*, 2019

AWARDS

⋆ National Scholarship for Ph.D Students.	2022
⋆ Ranked 3th in Facebook CVPR 2022 Ego4D Natural Language Query Challenge.	2022
⋆ Ranked 3th, NIST TRECVID 2021 Ad-hoc Video Search (AVS) Challenge.	2021
⋆ Ranked 4th in CVPR 2021 HOMAGE Scene-graph Generation Challenge.	2021
⋆ Ranked 2nd in ACM MM 2020 Video Relationship Understanding Grand Challenge.	2020
⋆ Ranked 2nd in ACM MM 2019 Video Relationship Understanding Grand Challenge.	2019
⋆ Best Method Prize in ACM MM 2019 Grand Challenge	2019
⋆ First Class Scholarship for Ph.D Students.	2018-2021
⋆ First Prize in National University Mathematical Modeling Competition.	2015

PROFESSIONAL ACTIVITIES

⋆ Conference Reviewer for CVPR, ICCV, ECCV, NeurIPS, AAAI, ACM MM, ICME.

⋆ Journal Reviewer for IJCV, TCSVT, TMM.