Yanhong Zeng

+8613146281900 | zengyh1900@gmail.com | zengyh1900.github.io | Yanhong Zeng

Biography

Yanhong Zeng is currently a researcher at Shanghai AI Laboratory. Before that, she obtained her computer science Ph.D. degree in the joint doctoral program between Sun Yat-sen University (SYSU) and Microsoft Research Asia (MSRA) in 2022, supervised by Prof. Hongyang Chao and Dr. Baining Guo.

Her research interest is developing high-quality and controllable generative models for open-domain multi-modality generation (e.g., image, video, audio, etc.). She has published papers in top international conferences and journals, such as CVPR/ECCV/NeurIPS/TVCG.

Education

Sun Yat-sen University, PhD in Computer Science and Technology Aug. 2017 – Jun. 2022

- Recipient of the National Scholarship Award
- Thesis: Research on Image and Video Inpainting by Generative Adversarial Networks

Sun Yat-sen University, BS in Software Engineering Aug. 2013 – Jun. 2017

- GPA: 3.9/4.0. Recipient of the National Scholarship Award, Outstanding Undergraduate Award

Experience

Researcher, Shanghai AI Laboratory – Shanghai, China Jul. 2022 – present

- **Poems of Timeless Acclaim (R&D).** This AI-driven animation series, created with China Media Group (CCTV), was broadcast in 10+ languages across 70+ platforms, reaching nearly 100 million viewers in two weeks. I developed the workflow for controllable image generation and human-centric animation.
- **MagicMaker (Project Owner).** It is an AI platform for effortless image generation, editing, and animation. I initiated and lead the project while leading a small R&D team in developing its AI models.
- **MMagic (Lead Core Maintainer).** MMagic is an open-source PyTorch toolbox for image and video editing as well as generation. I am responsible for feature development and community maintenance.

Research Intern, Microsoft Research Asia – Beijing, China Aug. 2018 – Dec. 2021

- Mentored by Dr. Jianlong Fu, conducting cutting-edge research on GAN and its applications in image and video inpainting, as well as video super-resolution. Delivered image inpainting models to Microsoft Office Team.

Research Intern, Microsoft Research Asia – Beijing, China Jun. 2016 – Jun. 2017

- Mentored by Dr. Richard Cai, conducting cutting-edge research in 3D human body reshaping.

Published

[1] Junshu Tang, **Yanhong Zeng**, Ke Fan, Xuheng Wang, Bo Dai, Kai Chen, and Lizhuang Ma. "Make-It-Vivid: Dressing Your Animatable Biped Cartoon Characters from Text". In: *CVPR*. 2024, pp. 6243–6253.

[2] Zhenzhi Wang, Yixuan Li, **Yanhong Zeng**, Youqing Fang, Yuwei Guo, Wenran Liu, Jing Tan, Kai Chen, Tianfan Xue, Bo Dai, et al. "Humanvid: Demystifying training data for camera-controllable human image animation". In: *NeurIPS (Datasets and Benchmarks)*. 2024.

[3] Jianzong Wu, Xiangtai Li, **Yanhong Zeng**, Jiangning Zhang, Qianyu Zhou, Yining Li, Yunhai Tong, and Kai Chen. "MotionBooth: Motion-Aware Customized Text-to-Video Generation". In: *NeurIPS (Spotlight)*. 2024.

[4] Yiming Zhang*, Zhening Xing*, **Yanhong Zeng**+, Youqing Fang, and Kai Chen+. "Pia: Your personalized image animator via plug-and-play modules in text-to-image models". In: *CVPR*. 2024, pp. 7747–7756.

[5] Junhao Zhuang, **Yanhong Zeng**, Wenran Liu, Chun Yuan, and Kai Chen. "A task is worth one word: Learning with task prompts for high-quality versatile image inpainting". In: *ECCV*. 2024, pp. 195–211.

[6] **Yanhong Zeng**, Jianlong Fu, Hongyang Chao, and Baining Guo. "Aggregated contextual transformations for high-resolution image inpainting". In: *IEEE TVCG* 29.7 (2022), pp. 3266–3280.

[7] **Yanhong Zeng***, Hongwei Xue*, Tiankai Hang*, Yuchong Sun*, Bei Liu, Huan Yang, Jianlong Fu, and Baining Guo. "Advancing high-resolution video-language representation with large-scale video transcriptions". In: *CVPR*. 2022, pp. 5036–5045.

[8] **Yanhong Zeng**, Huan Yang, Hongyang Chao, Jianbo Wang, and Jianlong Fu. "Improving visual quality of image synthesis by a token-based generator with transformers". In: *NeurIPS*. 2021, pp. 21125–21137.

[9] **Yanhong Zeng**, Jianlong Fu, and Hongyang Chao. "Learning joint spatial-temporal transformations for video inpainting". In: *ECCV*. 2020, pp. 528–543.

[10] Heliang Zheng, Jianlong Fu, **Yanhong Zeng**, Jiebo Luo, and Zheng-Jun Zha. "Learning semantic-aware normalization for generative adversarial networks". In: *NeurIPS (Spotlight)*. 2020, pp. 21853–21864.

[11] **Yanhong Zeng**, Jianlong Fu, Hongyang Chao, and Baining Guo. "Learning pyramid-context encoder network for high-quality image inpainting". In: *CVPR*. 2019, pp. 1486–1494.

[12] **Yanhong Zeng**, Jianlong Fu, and Hosngyang Chao. "3D human body reshaping with anthropometric modeling". In: *ICIMCS*. 2017, pp. 96–107.

Preprints

[1] Yicheng Chen, Xiangtai Li, Yining Li, **Yanhong Zeng**, Jianzong Wu, Xiangyu Zhao, and Kai Chen. *Auto cherry-picker: Learning from high-quality generative data driven by language*. 2024.

[2] Junyao Gao, Yanchen Liu, Yanan Sun, Yinhao Tang, **Yanhong Zeng**, Kai Chen, and Cairong Zhao. *Styleshot: A snapshot on any style*. 2024.

[3] Baiang Li, Sizhuo Ma, **Yanhong Zeng**, Xiaogang Xu, Youqing Fang, Zhao Zhang, Jian Wang, and Kai Chen. *Sagiri: Low Dynamic Range Image Enhancement with Generative Diffusion Prior*. 2024.

[4] Jianzong Wu, Chao Tang, Jingbo Wang, **Yanhong Zeng**, Xiangtai Li, and Yunhai Tong. *DiffSensei: Bridging Multi-Modal LLMs and Diffusion Models for Customized Manga Generation*. 2024.

[5] Zhening Xing, Gereon Fox, **Yanhong Zeng**, Xingang Pan, Mohamed Elgharib, Christian Theobalt, and Kai Chen. *Live2diff: Live stream translation via uni-directional attention in video diffusion models*. 2024.

[6] Yiming Zhang, Yicheng Gu, **Yanhong Zeng**+, Zhening Xing, Yuancheng Wang, Zhizheng Wu, and Kai Chen+. *Foleycrafter: Bring silent videos to life with lifelike and synchronized sounds*. 2024.

[7] Fuzhi Yang, Huan Yang, **Yanhong Zeng**, Jianlong Fu, and Hongtao Lu. *Degradation-guided meta-restoration network for blind super-resolution*. 2022.

Patents

[1] Jianzong Wu, Xiangtai Li, **Yanhong Zeng**, Jiangning Zhang, Qianyu Zhou, Yining Li, Yunhai Tong, and Kai Chen. "Customized object dynamic video generation method based on MotionBooth frames". CN118869902A. 2024.

[2] **Yanhong Zeng**, Junshu Tang, Ke Fan, Chuheng Wang, Bo Dai, Lizhuang Ma, and Kai Chen. "Three-dimensional cartoon image texture generation method and device based on text prompt words". CN118096979A. 2024.

[3] **Yanhong Zeng**, Yiming Zhang, Yicheng Gu, Zhening Xing, Yuancheng Wang, Zhizheng Wu, and Kai Chen. "Soundless video sound simulating method, electronic equipment and storage medium". CN118828050A. 2024.

[4] **Yanhong Zeng**, Yiming Zhang, Zhening Xing, Youqing Fang, and Kai Chen. "Content generation method, device and medium based on text prompt word and image drive". CN117911584A. 2024.

[5] **Yanhong Zeng**, Junhao Zhuang, Wenran Liu, Chun Yuan, and Kai Chen. "Multipurpose image redrawing method, device and medium based on prompt word learning". CN117710524A. 2024.

Services and Activities

Conference Reviewer

- International Conference on Learning Representations (ICLR): 2022, 2023, 2024, 2025
- The IEEE / CVF Computer Vision and Pattern Recognition Conference (CVPR): 2023, 2024, 2025
- The Conference on Neural Information Processing Systems (Neurips): 2021, 2022, 2023, 2024

- The International Conference on Computer Vision (ICCV): 2023
- The European Conference on Computer Vision (ECCV): 2024
- The International Conference on Machine Learning (ICML): 2022 (outstanding reviewer)
- The AAAI Conference on Artificial Intelligence (AAAI): 2022, 2023, 2024
- The ACM SIGGRAPH Conference: 2021
- International Conference on Artificial Intelligence and Statistics (AISTATS): 2025
- IEEE International Conference on Multimedia & Expo (ICME): 2021,2022,2023

Journal Reviewer

- IEEE Transactionson Image Processing (TIP)
- IEEE Transactions on Visualization and Computer Graphics (TVCG)
- IEEE Transactions on Multimedia (TMM)
- IEEE Transactions on Circuits and Systems for Video Technology (TCSVT)
- Pattern Recognition (PR)

Conference Committee

- Co-organizer, OpenMMLab: Open-source Platform for Vision, Language and Generative AI, Tutorial@ICCV2023
- Co-organizer, Boosting Computer Vision Research with OpenMMLab and OpenDataLab, Tutorial@CVPR2023
- Co-organizer, OpenMMLab: A Foundational Platform for Computer Vision Research and Production, Tutorial@AAAI 2023