XIN CHEN

University of Chinese Academy of Sciences | H: +86 186 1639 5193 | chenxin2@shanghaitech.edu.cn
HomePage: chenxin.tech | GitHub | Google Scholar

About Me

I'm a Researcher Scientist T10 at Tencent, working with Dr. Gang Yu for Generative AI. I obtained my Ph.D. from Chinese Academy of Sciences, advised by Prof. Jingyi Yu. Before that, I was supervised by Prof. Youyi Zheng at ShanghaiTech University, working on 3D modeling.

My research interests focus on Computer Vision for Graphics. I have a great passion on new things and new ideas, my goal is to create Generative AI which is about humans, used for humans, and benefits humans.

Research Interests

- **Generative AI** Multi-modal Language Models 3D AIGC
- **Computer Vision** Human Performance Capture Motion Synthesis View Synthesis
- **Computer Graphics** Image-based Modeling Neural Avatar Neural Rendering

Education

Ph.D Degree, Computer Science 2018-2022
University of Chinese Academy of Sciences

Master Student, Computer Science 2016-2018
ShanghaiTech University

B.Sc, Electronic Information Science Rank: 1/172 2012-2016
Qingdao University of Technology

Experience

Conference/Journal Reviewer	CVPR, ICCV, AAAI/ TPAMI, TIP, IJCV, TMM, ...	Jan. 2020 to Present
Research Scientist	Tencent - QQ Image Lab	Feb. 2022 to Present
Research Scientist Intern	Tencent - Youtu Lab	Dec. 2020 to Mar. 2021

Projects

- **3D AIGC for Digital Avatar and Textured Mesh.** Feb. 2022 to Present
  Proposed a text-to-texture framework for creating diverse avatar appearances and a text-to-shape model, Michelangelo, to generate 3D objects using diffusion models. Accepted to NeurIPS'23.
- **Human Motion Generation via Language/Diffusion Models.** Feb. 2022 to Present
  Introduced **MotionGPT**, a unified motion-language model to learn the semantic coupling and generate both motions and languages on multiple motion tasks. Accepted to NeurIPS'23. Presented **Motion-Latent-Diffusion**, a fast and high-quality motion diffusion model. Accepted to CVPR'23.
- **Human Shape/Motion Reconstruction for Clothed Avatars.** Dec. 2018 to Apr. 2022
  Built a **Dome System using 80 cameras** for multi-view stereo. Proposed a GAN-based scheme for human reconstruction, clothing segmentation, and virtual fitting, using non-rigid deformation for alignment. Lead **the reconstruction project, 1000+ clothed humans**, accepted to SIGGRAPH'22. Lead **the MoCap project**. Collected a sports motion dataset in diving, boxing, and more. Published on IJCV'21.
- **Image-based Shape Generation.** Aug. 2017 to Aug. 2018
  Introduced a fully automatic framework with the learning-based instance semantic segmentation part and the graphics-based reconstruction part. Published on TVCG'18 (Graphics journal).
- **Early R&D Projects.** Before Aug. 2017
  Mobile Virtual Fitting. A single-view human body estimation and virtual fitting on Android in Real-time. Based on the front-facing RGBD camera (ToF). A Linear Blend Skinning (LBS) body model.
  Dynamic 4D Mesh Player. Stand-alone development work for free-view browsing on 4D scans, which supports mesh rendering, free-angle viewpoint change, and Poisson Surface Reconstruction.
  Gesture Interaction in VR. A two-hand controller. Leap Motion, HTC Vive for hardware support.

Selected Publications (complete list...)

About 30 papers, over 10 as first/co-first author(*) or project leader (†), 5000+ GitHub Stars in total.

- Paint3D: Paint Anything 3D with Lighting-less Texture Diffusion Models.
  Xianfang Zeng*, **Xin Chen***, Zhongqi Qi*, Wen Liu, Zibo Zhao, Zhibin Wang, Bin Fu, Gang Yu
  [Arxiv'23 | Project | Code | Paper | Github Stars 400+]
- AppAgent: Human Motion as a Foreign Language.
  Chi Zhang*, Zhao Yang*, Jiaxuan Liu*, Yuchen Han, **Xin Chen**, Zebiao Huang, Bin Fu, Gang Yu
  [Arxiv'23 | Project | Code | Paper | Github Stars 3k+]
- LL3DA: Visual Interactive Instruction Tuning for Omni-3D Understanding, Reasoning, and Planning.
  Sijin Chen, **Xin Chen**†, Chi Zhang, Mingsheng Li, Gang Yu, Hao Fei, Hongyuan Zhu, Jiayuan Fan, Tao Chen.
  [Arxiv'23 | Project| Code | Paper]
- M3DBench: Let's Instruct Large Models with Multi-modal 3D Prompts.
  Mingsheng Li, **Xin Chen**†, Chi Zhang, Sijin Chen, Hongyuan Zhu, Fukun Yin, Gang Yu, Tao Chen.
  [Arxiv'23 | Project | Code| Paper]
- Chartllama: A multimodal llm for chart understanding and generation.
  Yucheng Han*, Chi Zhang*, **Xin Chen**, Xu Yang, Zhibin Wang, Gang Yu, Bin Fu, Hanwang Zhang
  [Arxiv'23 | Project | Code | Paper]
- MotionGPT: Human Motion as a Foreign Language.
  Biao Jiang*, **Xin Chen***, Wen Liu, Jingyi Yu, Gang Yu, Tao Chen
  [NeurIPS'23 | Project | Code | Paper | Github Stars 1k+]
- Michelangelo: Conditional 3D Shape Generation based on Shape-Image-Text Aligned Latent Representation.
  Zibo Zhao, Wen Liu, **Xin Chen**, Xianfang Zeng, Rui Wang, Pei Cheng, Bin Fu, Tao Chen,Gang Yu, Shenghua Gao
  [NeurIPS'23 | Project | Code | Paper | Github Stars 100+]
- Executing your Commands via Motion Diffusion in Latent Space.
  Xin Chen*, Biao Jiang*, Wen Liu, Zilong Huang, Bin Fu, Tao Chen, Jingyi Yu, Gang Yu
  [CVPR'23 | Project | Code | Paper | Github Stars 400+]
- End-to-End 3D Dense Captioning with Vote2Cap-DETR.
  Sijin Chen, Hongyuan Zhu, **Xin Chen**, Yinjie Lei, Tao Chen, Gang Yu
  [CVPR'23 | Video | Code | Paper]
- A Large-Scale Outdoor Multi-modal Dataset and Benchmark for Novel View Synthesis and Implicit Scene Reconstruction.
  Chongshan Lu, Fukun Yin, **Xin Chen**, Tao Chen, Gang Yu, Jiayuan Fan
  [ICCV'23 | Project | Code | Paper]
- TightCap: 3D Human Shape Capture with Clothing Tightness Field.
  Xin Chen, Anqi Pang, Peihao Wang, Wei Yang, Lan Xu, Jingyi Yu
  [SIGGRAPH'22 | Project | Code | Paper | TOG Journal Track]
- SportsCap: Monocular 3D Human Motion Capture and Fine-grained Understanding in Challenging Sports Videos.
  Xin Chen, Anqi Pang, Yuexin Ma, Lan Xu, Jingyi Yu
  [IJCV'21 | Project | Code | Paper]
- ChallenCap: Monocular 3D Capture of Challenging Human Performances using Multi-Modal References.
  Yannao He, Anqi Pang, **Xin Chen**, Han Liang, Yuexin Ma, Lan Xu
  [CVPR'21 Oral | Project | Paper]
- Anisotropic Fourier Features for Neural Image-Based Rendering and Relighting.
  Huangjie Yu, Anpei Chen, **Xin Chen**, Lan Xu, Ziyu Shao, Jingyi Yu
  [AAAI'22 Oral | Project | Paper]
- Few-shot Neural Human Performance Rendering from Sparse RGBD Videos.
  Anqi Pang*, **Xin Chen***, Haimin Luo, Minye Wu, Jingyi Yu, Lan Xu
  [IJCAI'21 | Paper | Video]

- Neural Free-Viewpoint Performance Rendering under ComplexHuman-object Interactions.
  Guoxing Sun, **Xin Chen**, Yizhang Chen, Anqi Pang, Pei Lin, Lan Xu, Jingya Wang, Jingyi Yu
  [ACMMM'21 | Paper | Video]
- Pose2Body: Pose-Guided Human Parts Segmentation.
  Zhong Li*, **Xin Chen***, Wangyiteng Zhou,Yingliang Zhang, Jingyi Yu
  [ICME'19 Oral | Paper]
- AutoSweep: Recovering 3D Editable Objects from a Single Photograph.
  Xin Chen, Yuwei Li, Xi Luo, Tianjia Shao, Jingyi Yu, Kun Zhou, Youyi Zheng
  [TVCG'18 | Project | Code | Paper]
- Sparse Photometric 3D Face Reconstruction Guided by Morphable Models.
  Xuan Cao, Zhang Chen, Anpei Chen, **Xin Chen**, Shiying Li, Jingyi Yu
  [CVPR'18 | Paper | Video]

References

Prof. Jingyi Yu	Supervisor, IEEE Fellow	ShanghaiTech University	yujingyi@@shanghaitech.edu.cn
Prof. Youyi Zheng	Former Supervisor	Zhejiang University	youyi.zheng@zju.edu.cn