

# XIN CHEN

University of Chinese Academy of Sciences | H: +86 186 1639 5193 | [chenxin2@shanghaiitech.edu.cn](mailto:chenxin2@shanghaiitech.edu.cn)  
[HomePage](#) | [GitHub](#) | [Google Scholar](#)

## RESEARCH INTERESTS

---

- |                            |                             |                  |                  |
|----------------------------|-----------------------------|------------------|------------------|
| • <b>Generative AI</b>     | Multi-modal Language Models | 3D AIGC          |                  |
| • <b>Computer Vision</b>   | Human Performance Capture   | Motion Synthesis | View Synthesis   |
| • <b>Computer Graphics</b> | Image-based Modeling        | Neural Avatar    | Neural Rendering |

## EDUCATION

---

<b>Ph.D Degree, Computer Science</b>		2018-2022
	University of Chinese Academy of Sciences	
<b>Master Student, Computer Science</b>		2016-2018
	ShanghaiTech University	
<b>B.Sc, Electronic Information Science</b>	Rank: 1/172	2016
	Qingdao University of Technology	

## EXPERIENCE AND SERVICE

---

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

## SELECTED PUBLICATIONS ([COMPLETE LIST...](#))

---

- MotionGPT: Human Motion as a Foreign Language.  
Biao Jiang\*, **Xin Chen\***, Wen Liu, Jingyi Yu, Gang Yu, Tao Chen  
[[Arxiv'23](#) | [Project](#) | [Code](#) | [Paper](#) | Github Stars 600+]
- 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  
[[Arxiv'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 300+]
- End-to-End 3D Dense Captioning with Vote2Cap-DETR.  
Chongshan Lu, Fukun Yin, **Xin Chen**, Tao Chen, Gang Yu, Jiayuan Fan  
[[CVPR'23](#) | [Project](#) | [Code](#) | [Paper](#)]
- A Large-Scale Outdoor Multi-modal Dataset and Benchmark for Novel View Synthesis and Implicit Scene Reconstruction.  
Sijin Chen, Hongyuan Zhu, **Xin Chen**, Yinjie Lei, Tao Chen, Gang Yu  
[[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 Complex Human-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**, Shiyong Li, Jingyi Yu [[CVPR'18](#) | [Paper](#) | [Video](#)]

## 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. Under review.
- **Human Motion Generation via Language/Diffusion Models.** Feb. 2022 to Present  
Introduced **MotionGPT**, a unified motion-language model to learn the semantic coupling of two modalities and generate both motions and languages on multiple motion tasks. 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 Feb. 2018  
**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.

## TECHNICAL SKILLS

---

Languages:	Python, Pytorch, Pyrender C#, C++ (OpenGL, OpenCV, Qt, Eigen, PCL, CUDA ...), Matlab
Softwares:	Unity3D, Blender, Adobe Photoshop, Premiere Visual Studio, Pycharm, Jupyter Notebook, Android Studio
System:	Multi-view Dome and Light Stage System for <b>Object, Hand, and Body</b> reconstruction Leap Motion, HTC Vive

## REFERENCES

---

<a href="#">Prof. Jingyi Yu</a>	Supervisor, IEEE Fellow	SHANGHAI TECH UNIVERSITY	<a href="mailto:yujingyi@shanghaitech.edu.cn">yujingyi@shanghaitech.edu.cn</a>
<a href="#">Prof. Youyi Zheng</a>	Former Supervisor	ZHEJIANG UNIVERSITY	<a href="mailto:youyi.zheng@zju.edu.cn">youyi.zheng@zju.edu.cn</a>