Chao Wang

chaowang.info linkedin.com/in/chaowang15 chaowang15@gmail.com

Experience

Meta Reality Labs | *Research Engineer* Burlingame, CA, 08/2024 – Present
Multi-modal human video generation, talking head generation, generative modeling

Bytedance/Tiktok | *Research Engineer* San Jose, CA, 02/2020 – 08/2024
Audio/video driven portrait animation and body motion synthesis, talking head generation, real-time facial animation for virtual avatars, 3D face reconstruction

Matterport | *3D Vision Engineer* Sunnyvale, CA, 10/2018 – 02/2020
Large-scale 3D scene reconstruction, 3D geometry processing (meshing, simplification, texturing and rendering)

Education

The University of Texas at Dallas 2012 – 2018
Ph.D. in Computer Science. Advisor: Prof. Xiaohu Guo

Tsinghua University 2009 - 2012
M.S. in Computer Science. Advisor: Prof. Jun-hai Yong and Prof. Yu-shen Liu

Tsinghua University 2005 - 2009
B.S. in Automation

Peking University 2008 - 2011
B.S. in Economics (double major)

Highlighted Projects

Doubao Avatar: One-shot Real-time Chatting Avatar on Mobile Devices Bytedance, 2023–2024

- Developed efficient audio-driven single image portrait animation algorithm, utilizing LSTM-based autoregression for audio-to-neural-keypoints and neural-motion model for keypoints-to-talking-video. Integrated entire pipeline in Doubao App and plan to launch very soon this year. [Douyin demo1, demo2]
- This real-time running avatar only needs one single portrait image input and can chat with users with precise lipsync, authentic expressions and natural head movements. It connects LLM to get response text, utilizes TTS to convert text to audio, and leverages our model to create talking video instantaneously.

Diffusion-based Realistic Face Animations from Single Image Bytedance, 2023–2024

- Developed **DREAM-Talk**: a diffusion-based single image audio-driven emotional talking face pipeline which generates accurate lip motion, emotional talking face and diversified head movements all together. [Project Page]
- Developed **X-Portrait**: a video-driven single image animation method with conditional diffusion model for generating expressive and temporally coherent portrait animations, with highly dynamic and subtle facial expressions along with wide-range head movements. [Project Page]

Tiktok Avatar: Real-time Face-driven 3D Avatar Bytedance, 2022

- Team leader and core developer for video-driven animation avatar pipeline: developed efficient and robust video-to-blendshape model to predict blendshape weights and animate avatars. Integrated the model to the Tiktok-Avatar effect used in 1M+ new videos with 50M+ video plays. [News, Tiktok Avatar Videos]

Segment Everything for Celebration: Image Segmentation for Douyin Effects Bytedance, 2021

- Developed object segmentation algorithm on images based on saliency detection, image-to-mesh algorithm to convert segmented object to mesh, and rig animation algorithm to drive mesh to move and render on mobile devices. Integrated all features to a Douyin effect used in 10M+ new videos with 2B+ video plays. [Douyin Videos]

Matterport Scene: Large-scale Scene Reconstruction and Immersive Touring Matterport, 2018-2020

- Enchanced 3D scene reconstruction system significantly in: 1) increasing texture reconstruction efficiency for large mesh rendering on the web; 2) improving reconstruction quality, mesh alignment and simplification accuracy and efficiency; 3) developing new RGB-D reconstruction system on mobile devices. [Gallery Example]

Publications

[1] You Xie, Hongyi Xu, Guoxian Song, **Chao Wang**, Yichun Shi, Linjie Luo. X-Portrait: Expressive Portrait Animation with Hierarchical Motion Attention. In *Arxiv Preprint*, 2024 (Under review at *SIGGRAPH 2024*)

[2] Chenxu Zhang*, **Chao Wang***, Jianfeng Zhang, Hongyi Xu, Guoxian Song, You Xie, Linjie Luo, Yapeng Tian, Xiaohu Guo, Jiashi Feng. DREAM-Talk: Diffusion-based Realistic Emotional Audio-driven Method for Single Image Talking Face Generation. In *Arxiv Preprint*, 2023 (*Equal Contribution)

[3] Chenxu Zhang, **Chao Wang**, Yifan Zhao, Shuo Cheng, Linjie Luo, Xiaohu Guo. DR2: Disentangled Recurrent Representation Learning for Data-efficient Speech Video Synthesis. In *IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, 2024

[4] Shuo Cheng, Guoxian Song, Wan-Chun Ma, **Chao Wang**, Linjie Luo. Using Augmented Face Images to Improve Facial Recognition Tasks. In *AI-Generated Characters: Putting Deepfakes to Good Use (CHI Workshop)*, 2022

[5] Qiaomu Miao, Sinhwa Kang, Stacy Marsella, Steve DiPaola, **Chao Wang**, Ari Shapiro. Study of detecting behavioral signatures within DeepFake videos. In *Arxiv Preprint*, 2022

[6] **Chao Wang**, Xiaohu Guo. Efficient Plane-Based Optimization of Geometry and Texture for Indoor RGB-D Reconstruction. In *SUMO Workshop - 360° Indoor Scene Understanding and Modeling, IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2019

[7] **Chao Wang**, Xiaohu Guo. Plane-based optimization of geometry and texture for RGB-D reconstruction of indoor scenes. In *International Conference on 3D Vision (3DV)*, 2018

[8] **Chao Wang**, Xiaohu Guo. Feature-based RGB-D camera pose optimization for real-time 3D reconstruction. In *Computational Visual Media (CVM)*, 2017

[9] **Chao Wang**, Yang Liu, Xiaohu Guo, Zichun Zhong, Binh Le, Zhigang Deng. Spectral animation compression. In *Computational Visual Media (CVM)*, 2015

[10] **Chao Wang**, Chao Wang, Yu-Shen Liu, Min Liu, Jun-Hai Yong, Jean-Claude Paul. Robust shape normalization of 3D articulated volumetric models. In *Computer-Aided Design (CAD)*, 2012

Professional Services

- **Conference Reviewer**: SIGGRAPH'24, ECCV'24, CVPR'23-24, ICCV'23, AAAI'23, SIGGRAPH Asia'21, ACM MM'21-24, WACV'24, GMP'24, PacificVis'22, SPM'18-19, CVM'17-18, etc
- **Journal Reviewer**: TMM, GMOD, CASA, CAD, CAGD, TVCJ, CAVW, GM, WEVJ, Measurement Science and Technology (MST), Remote Sensing, Sensors, Drones, Electronics, Mathematics, Biomimetics, Forests, Symmetry, etc

Patents

[1] Wanchun Ma, Shuo Cheng, **Chao Wang**, Michael Leong Hou Tay, Linjie Luo. Neural network architecture for face tracking . US11803996B2, Granted in 2023

[2] Shuo Cheng, Guoxian Song, Wanchun Ma, **Chao Wang**, Linjie Luo. Using augmented face images to improve facial recognition tasks. US20230377368A1, 2022

[3] Michael Leong Hou Tay, Wanchun Ma, Shuo Cheng, **Chao Wang**, Linjie Luo. Asymmetric facial expression recognition. US20230046286A1, 2021