CHENXU ZHANG

682-246-1040 | chenxu.zhang@utdallas.edu | github.com/zhangchenxu528

EDUCATION

The University of Texas at Dallas, Richardson, TX Jan. 2019 - Present
Ph.D in Computer Science

Beihang University, Beijing, China Sep. 2015 - Mar. 2018
M.S in Computer Science

Beihang University, Beijing, China Sep. 2011 - Jul. 2015
B.S in Software Engineering

RESEARCH INTERESTS

My research interests include Computer Graphics, Computer Vision, and AI, with a focus on Talking Face Generation, Conversational Gestures Synthesis, Deblur-NeRF with Human Motion, and Emotional Talking Avatar.

RESEARCH EXPERIENCE

Audio-driven personalized emotional talking avatar generation May 2022 - present

- Propose the first emotional talking avatar generation framework with the disentangled speech content and emotion training architecture.
- Design a mouth energy loss to guarantee the mouth accuracy with limited emotion datasets.
- Develop talking avatar applications, including emotion transfer, emotion exaggeration and personalized emotion.

Deblur-NeRF: moving people rendering from blurry videos Jan. 2022 - present

- A dataset of rendered human motions with paired realistic and blurred images for quantitative comparison of human motion deblur.
- Referring to the camera's photography process, we design the human motion deblur framework, which can improve the motion blur for any existing nerf-based human motion rendering framework (eg, Neural body, Human-NerRF).

Conversational gestures synthesis from single short video May 2021 - Jan. 2022

- Propose a novel Disentangled Recurrent Representation (DR^2) Learning framework to synthesize the long diversified gesture sequence with short training video data.
- Propose a disentangled module with a state bank to encourage the learning of unpaired pose and audio embedding, resulting in diverse "one-to-many" mappings in pose generation.
- Develop a recurrent inference module to feed back the last generation as initial pose prior, resulting in long-term diverse videos with full consistency..

Audio-driven talking face video generation Jan. 2020 - May 2021

- Synthesize photo-realistic talking face videos with audio-synchronized lip motion, personalized and natural head motion, and realistic eye blinks.
- Design a FACIAL-GAN module to encode the contextual information with the phonetic information to model the implicit attributes needed for synthesizing natural head motions.
- Embed eye blinking into an eye-attention map of rendered faces, which achieves realistic eye blinks in the resulting video produced by the Rendering-to-Video module.

3D talking face generation Jan. 2019 - Jan. 2020

- Construct a person-specific head motion dataset.
- Propose a unified audio-inspired approach to endow 3D talking face with personalized pose dynamics.

Modeling garment seam from a single image Mar. 2017 - Jan. 2018

- Establishment of a garment seam image database and the parametric seam models.
- Use the Deformable Parts Model (DPM) image detection method for precise seam detection.

Thread-level fabric modeling based on a single macro image Sep. 2016 - Jun. 2017

- The production of thread-level fabric model and the final rendering of the development work.

PUBLICATIONS

Chenxu Zhang, Chao Wang, Yifan Zhao, Shuo Cheng, Linjie Luo, Xiaohu Guo. DR^2: Disentangled Recurrent Representation Learning for Data-efficient Speech Video Synthesis. (Under Review in AAAI)

Chenxu Zhang, Yifan Zhao, Yifei Huang, Ming Zeng, Saifeng Ni, Madhukar Budagavi, Xiaohu Guo. FACIAL: Synthesizing Dynamic Talking Face with Implicit Attribute Learning. International Conference on Computer Vision (ICCV), 2021.

Chenxu Zhang, Saifeng Ni, Zhipeng Fan, Hongbo Li, Ming Zeng, Madhukar Budagavi, Xiaohu Guo. 3D Talking Face with Personalized Pose Dynamics. IEEE Transactions on Visualization and Computer Graphics (TVCG), 2021.

Yifei Huang, Chenhui Li, Xiaohu Guo, Jing Liao, **Chenxu Zhang**, Changbo Wang. DeSmoothGAN: Recovering Details of Smoothed Images via Spatial Feature-wise Transformation and Full Attention. Proceedings of the 28th ACM International Conference on Multimedia (MM), 2020.

Guihong Wan, Crystal Maung, **Chenxu Zhang**, Haim Schweitzer. Fast Distance Metrics in Low-dimensional Space for Neighbor Search Problems. IEEE International Conference on Data Mining (ICDM), 2020.

Hongyu Wu, Xiaowu Chen, **Chenxu Zhang**, Bin Zhou, Qinping Zhao. Modeling yarn-level geometry from a single micro-image. Frontiers of Information Technology & Electronic Engineering, 2019.

Chenxu Zhang, Xiaowu Chen, Hongyu Wu, Bin Zhou. Modeling Garment Seam from a Single Image. Journal of Computer Science and Technology (JCST), 2018.

WORK EXPERIENCE

Research Intern in ByteDance May 2021 - Aug. 2021, May 2022 - Present

- Developed the talking face and gesture video generation method, which can generate realistic human video when giving the speech.
- Developed the emotional talking avatar generation framework with emotion transfer, emotion exaggeration and personalized emotion applications.

Research Assistant Jan. 2020 - May 2021, Aug. 2021 - May 2022

The University of Texas at Dallas Richardson, TX

Teaching Assistant Jan. 2019 - Jan. 2020

The University of Texas at Dallas Richardson, TX

- Assist in the teaching of following courses: Computer Graphics, Computer Animation.

AWARDS

Lars Magnus Ericsson Graduate Fellowships (University of Texas at Dallas) 2019-2020

First Class Scholarship (Beihang University) 2017-2018

First Class Scholarship (Beihang University) 2016-2017

First Class Scholarship (Beihang University) 2015-2016