JIANMAN LIN

linjianmancjx@gmail.com | Google Scholar

EDUCATION

South China University of Technology (SCUT) Guangzhou, China
Master of Engineering in Electronic Information *Sep. 2024 – Present*
Supervisors: Prof. Tianshui Chen (Research) & Prof. Chunmei Qing (Academic)
GPA: 80/100

Guangdong University of Technology (GDUT) Guangzhou, China
Bachelor of Engineering in Industrial Engineering *Sep. 2020 – June 2024*
GPA: 87.2/100

RESEARCH SUMMARY

Research Interest: Computer Vision, Generative AI, and Controllable Image Synthesis.
Research Trajectory: Progressed from **Speech-Preserving Facial Expression Manipulation** (addressing data scarcity via representation learning) to **Generalizable Diffusion-based Editing** (addressing the trade-off between editability and visual consistency by disentanglement learning).

SELECTED PUBLICATIONS

Neural Scene Designer: Self-Styled Semantic Image Manipulation [PDF]
Jianman Lin, Tianshui Chen, Chunmei Qing, Zhijing Yang, Shuangping Huang, Yuheng Ren, Liang Lin.
*IEEE Transactions on Image Processing (**TIP**), Vol. 34, pp. 6577-6588, 2025.*

Learning Adaptive Spatial Coherent Correlations for Speech-Preserving Facial Expression Manipulation [PDF]
Tianshui Chen, **Jianman Lin**, Zhijing Yang, Chunmei Qing, Liang Lin.
*IEEE/CVF Conference on Computer Vision and Pattern Recognition (**CVPR**), 2024.* **Highlight (Top 2.8%)**

Contrastive Decoupled Representation Learning and Regularization for Speech-Preserving Facial Expression Manipulation [PDF]
Tianshui Chen, **Jianman Lin***, Zhijing Yang, Chunmei Qing, Yukai Shi, Liang Lin. **(*Co-first author)**
*International Journal of Computer Vision (**IJCV**), Vol. 133, pp. 3822-3838, 2025.*

Learning Spatial-Temporal Coherent Correlations for Speech-Preserving Facial Expression Manipulation
Tianshui Chen, **Jianman Lin***, Zhijing Yang, Chunmei Qing, Liang Lin. **(*Co-first author)**
*Major Revision to IEEE Transactions on Pattern Analysis and Machine Intelligence (**T-PAMI**).*

Geometry-Editable and Appearance-Preserving Object Composition [PDF]
Jianman Lin, Haojie Li, Chunmei Qing, Zhijing Yang, Liang Lin, Tianshui Chen.
*Submitted to International Journal of Computer Vision (**IJCV**).*

RESEARCH EXPERIENCE

Topic 1: Speech-Preserving Facial Expression Manipulation via Representation Learning 2023 – 2024
Focus: Addressing the challenge of **Data Scarcity** (lack of paired data) by mining intrinsic priors and learning decoupled representations.

- Learning adaptive spatial coherent correlations for speech-preserving facial expression manipulation
  - **Problem:** Existing SPFEM models struggle to preserve fine-grained mouth animations due to the scarcity of paired training data (identical content, different emotion).
  - **Insight:** Discovered a novel spatial-coherent correlation between local visual disparities across emotions, enabling us to use a minimal paired data set to provide crucial pair-wise supervision.
  - **Methodology:** Proposed ASCCL, which formulated correlations into learnable metrics using a Difference-Aware Adaptive Strategy (DAAS) to resolve the paired data dependency.
  - **Outcome:** Selected as a **CVPR 2024 Highlight (Top 2.8%)**; extended version currently under review at **IEEE T-PAMI**.

- Contrastive decoupled representation learning and regularization for speech-preserving facial expression manipulation
  - **Problem:** The intrinsic intertwining of content and emotion hinders the extraction of ideal, independent supervisory signals, causing inaccurate expression transfer and poor lip synchronization in SPFEM.
  - **Insight:**Proposed CDRL to explicitly disentangle intertwined content/emotion features by leveraging distinct Audio and VLM-derived priors as guidance.
  - **Methodology:** Designed CCRL/CERL with Contrastive Learning to extract decoupled content/emotion representations, serving as explicit regularization signals for model constraint.
  - **Outcome:** Published in **International Journal of Computer Vision (IJCV 2025)**.

Topic 2: Controllable Image Editing via Diffusion-based Disentangled Learning 2024 – Present
Focus: Addressing the inherent trade-off between editability and visual consistency in controllable image editing within the Diffusion Model framework.

- Neural Scene Designer: Self-Styled Semantic Image Manipulation
  - **Problem:** Conventional inpainting methods disrupt stylistic consistency of the scene when manipulating semantic content.
  - **Insight:** Identified that intra-image regions share consistent style codes, providing a robust self-supervised signal for fine-grained style modeling.
  - **Methodology:**Proposed NSD, a Diffusion framework using a Dual Parallel Cross-Attention mechanism for text/style control, integrated with PSRL for self-style representation learning.
  - **Outcome:** Published in **IEEE Transactions on Image Processing (TIP 2025)**.
- Geometry-Editable and Appearance-Preserving Object Compositon
  - **Problem:** The coupling of structure and appearance in object composition leads to the loss of fine-grained texture details when applying geometric transformations.
  - **Insight:** Proposed a novel "Geometry-First, Appearance-Follows" decoupling paradigm, leveraging the Diffusion model's strong spatial perception to first establish geometry before mapping fine-grained appearance.
  - **Methodology:** Developed DGAD, which disentangles the process into two stages: implicit geometric modeling (Encoding) via semantic embeddings, and explicit appearance alignment (Decoding) via Dense Cross-Attention Retrieval.
  - **Outcome:** Manuscript currently under review at a top-tier venue, demonstrating flexible geometry-editable and high-fidelity composition.

Work Experience

Algorithm Researcher Intern Jul. 2024 – Present
ExpanderaAI – AI Interior Design Lab *Guangzhou, China*

- **MLLM-Driven Intelligent Interior Design System**
  - **VLM-based structured room reconstruction system:** Independently led VLM-based structured floor plan reconstruction from sparse image inputs (4 views). Established a 500k+ image data processing pipeline for paired data construction. **Progressively activated the VLM's spatial reasoning ability through multi-stage curriculum reinforcement learning, achieving a centimeter-level dimension error (< 3 cm)**, which demonstrates accurate geometric shape and size reconstruction for a single-space floor plan.
  - **VLM-Driven Pipeline for Intelligent Interior Design:** As a core algorithm contributor, designed and implemented the MLLM-driven pipeline for intelligent interior design, including: (1): VLM-based floor plan perception, enabling the parsing of structured information from 2D floor plans. (2):VLM-based automated layout generation, leveraging the powerful perception and reasoning capabilities of multimodal large models, integrated with physical rules priors and design aesthetics to achieve precise and aesthetically pleasing interior arrangements.