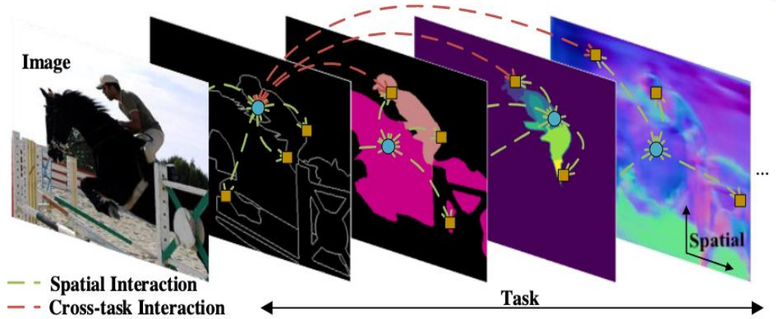
Research Background

THE DEPARTMENT OF
COMPUTER SCIENCE & ENGINEERING
計算機科學及工程學系

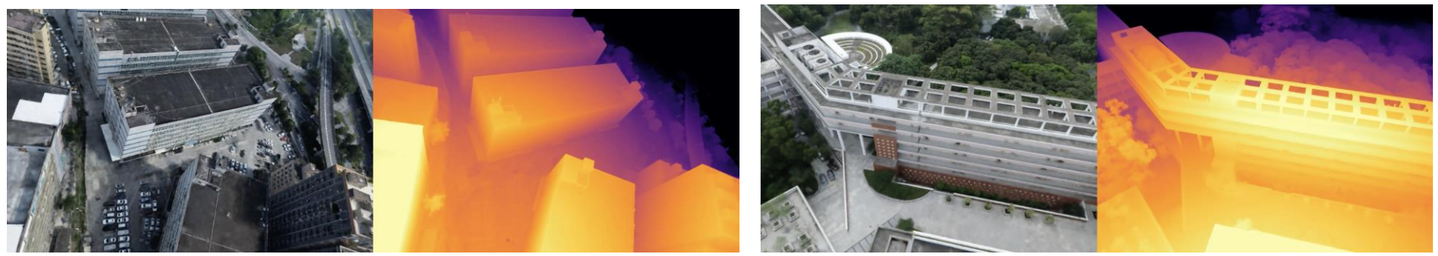
- Research Interests

Computer Vision & Deep Learning

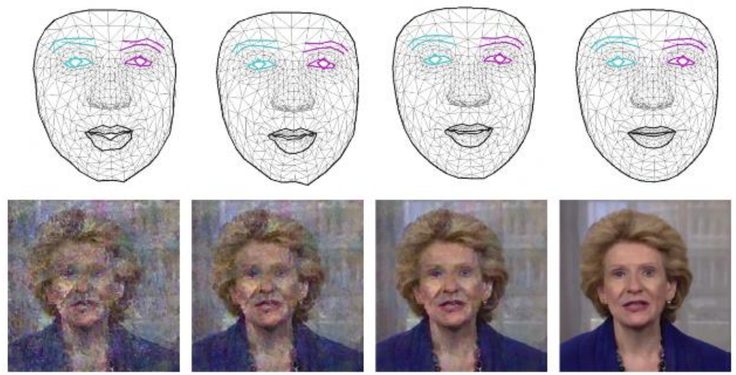
- Deep Structured Feature Learning and Prediction
- Deep Multi-Modal, Multi-Task Learning
- 2D/3D Holistic Visual Scene Understanding and Generation



#1 Joint Multi-Task Scene Understanding



#2 Large-scale Unsupervised 3D Urban Scene Modeling



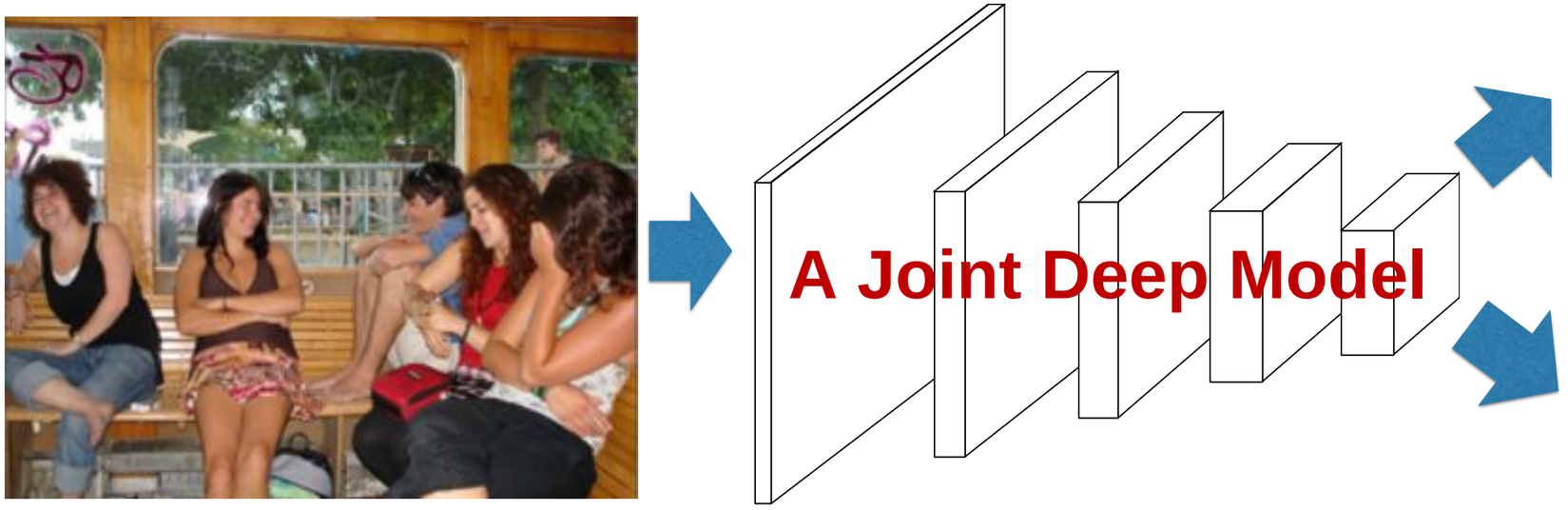
#3 Human, object, and scene-centric 3D/Video Generation

Research Highlights

THE DEPARTMENT OF
COMPUTER SCIENCE & ENGINEERING
計算機科學及工程學系

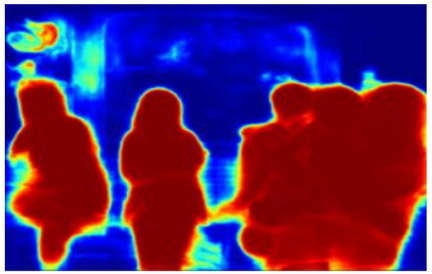
- Joint Deep Multi-Task Scene Understanding

- **Joint perception and reasoning** within a unified deep framework

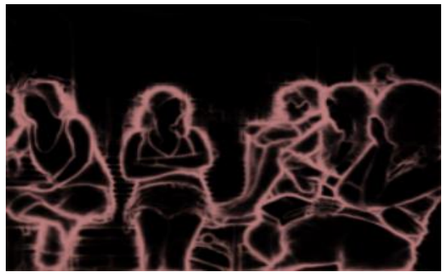


Boundary Detection

Surface Normal

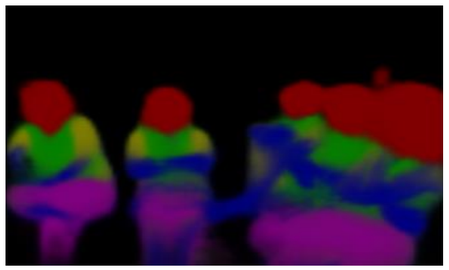


Segmentation



Sem. Boundaries

Object Detection



Human Parsing

Research Highlights

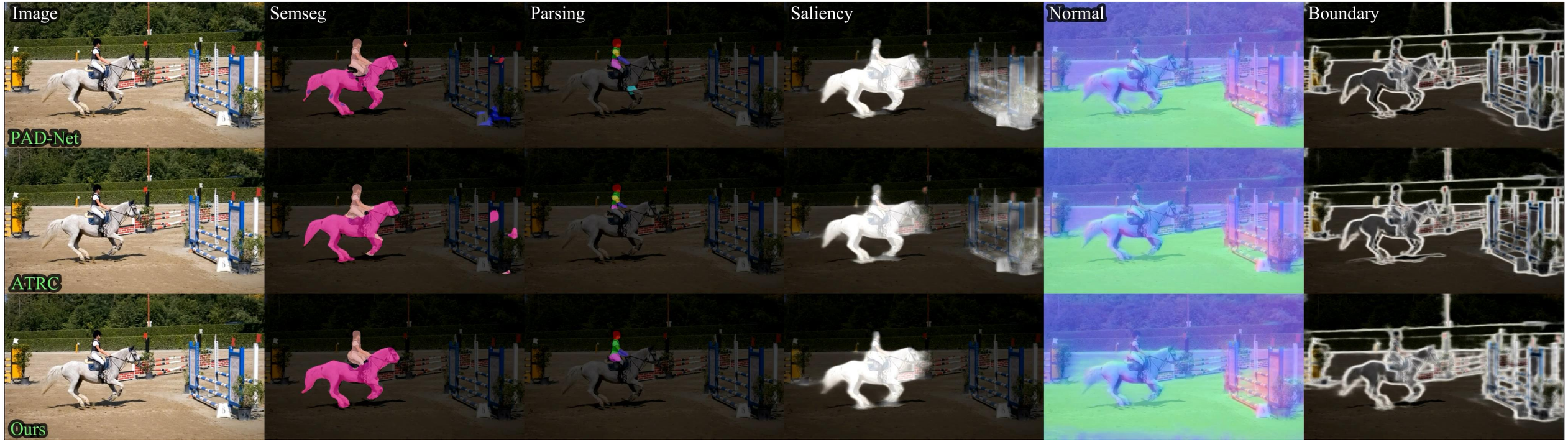
THE DEPARTMENT OF
COMPUTER SCIENCE & ENGINEERING
計算機科學及工程學系

- Joint Deep Multi-Task Scene Understanding

- **Joint perception and reasoning** within a unified deep framework

- Highly beneficial for the training **efficiency** and model **generalization**

- Publications: CVPR 24, ICLR 23, ECCV 22, ICCV’ 21, CVPR’ 20, CVPR’ 18, etc.



Experiment of Multi-Task Preidciton on Cityscapes-3D

3D Detection, Segmentation, and Depth

TaskPrompter

TaskPrompter: Spatial-Channel Multi-Task Prompting for Dense Scene Understanding

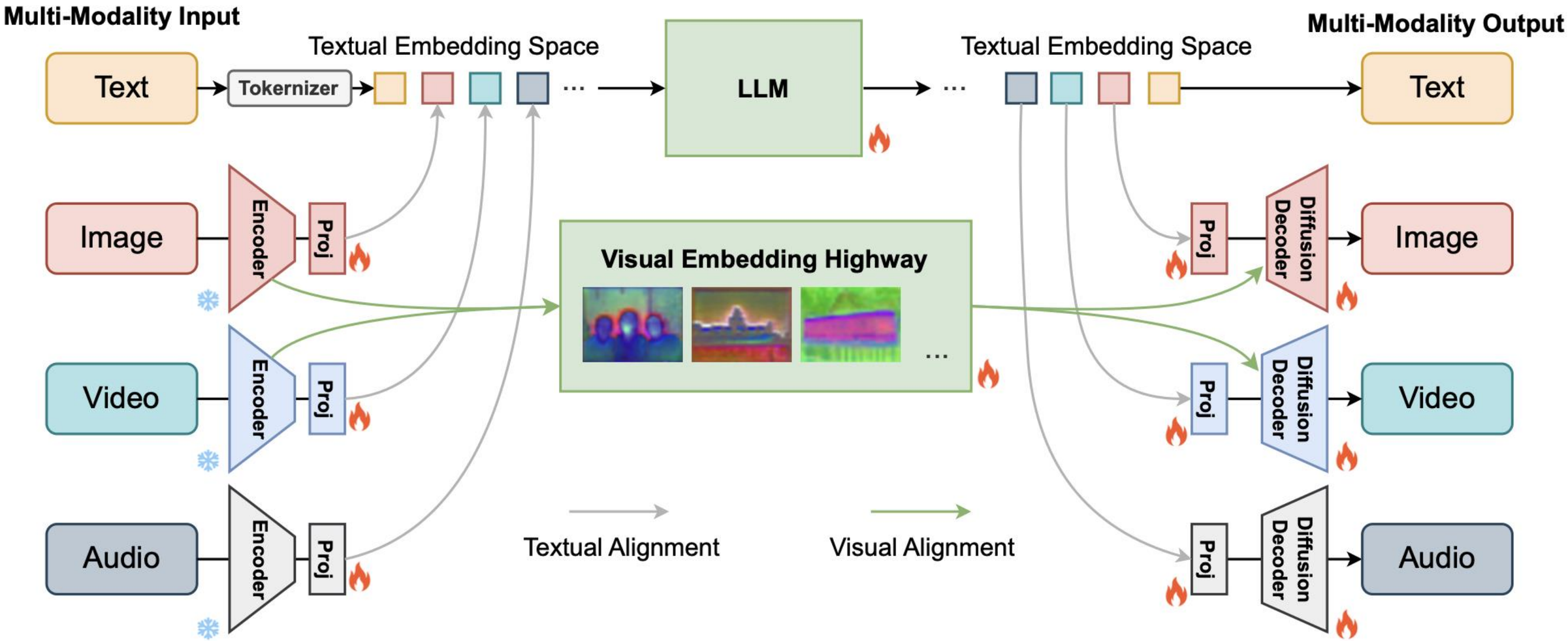
Hanrong Ye and Dan Xu, CSE, HKUST

ICLR 2023

Research Highlights

THE DEPARTMENT OF
COMPUTER SCIENCE & ENGINEERING
計算機科學及工程學系

- Joint Multi-Modal Multi-Task Framework via LLMs



Research Highlights

THE DEPARTMENT OF
COMPUTER SCIENCE & ENGINEERING
計算機科學及工程學系

- Large-Scale Unsupervised 3D Urban Scene Modeling

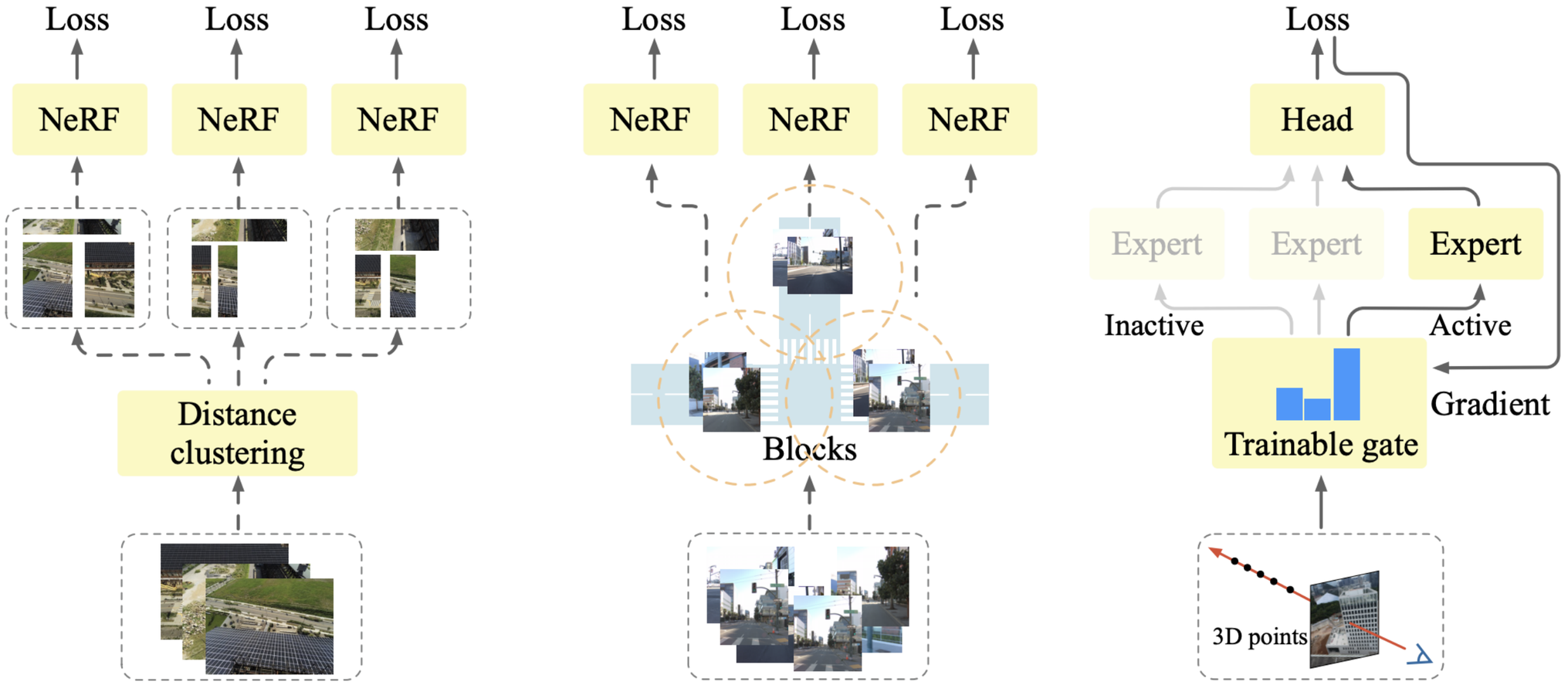
- Unsupervised: **no 3D sensor data** is available
- Very challenging if dealing with a large-scale scene
- NeRF is an effective **continuous** scene presentation (multi-view posed images)

Zhenxing Mi, Dan Xu, ‘Switch-NeRF: Learning Scene Decomposition with Mixture of Experts for Large-Scale Neural Radiance Fields’, ICLR 2023

Research Highlights

THE DEPARTMENT OF
COMPUTER SCIENCE & ENGINEERING
計算機科學及工程學系

- **Scene decomposition** is critical for efficiency and flexibility



(a) Learning after distance-based decomposition (*e.g.* Mega-NeRF)

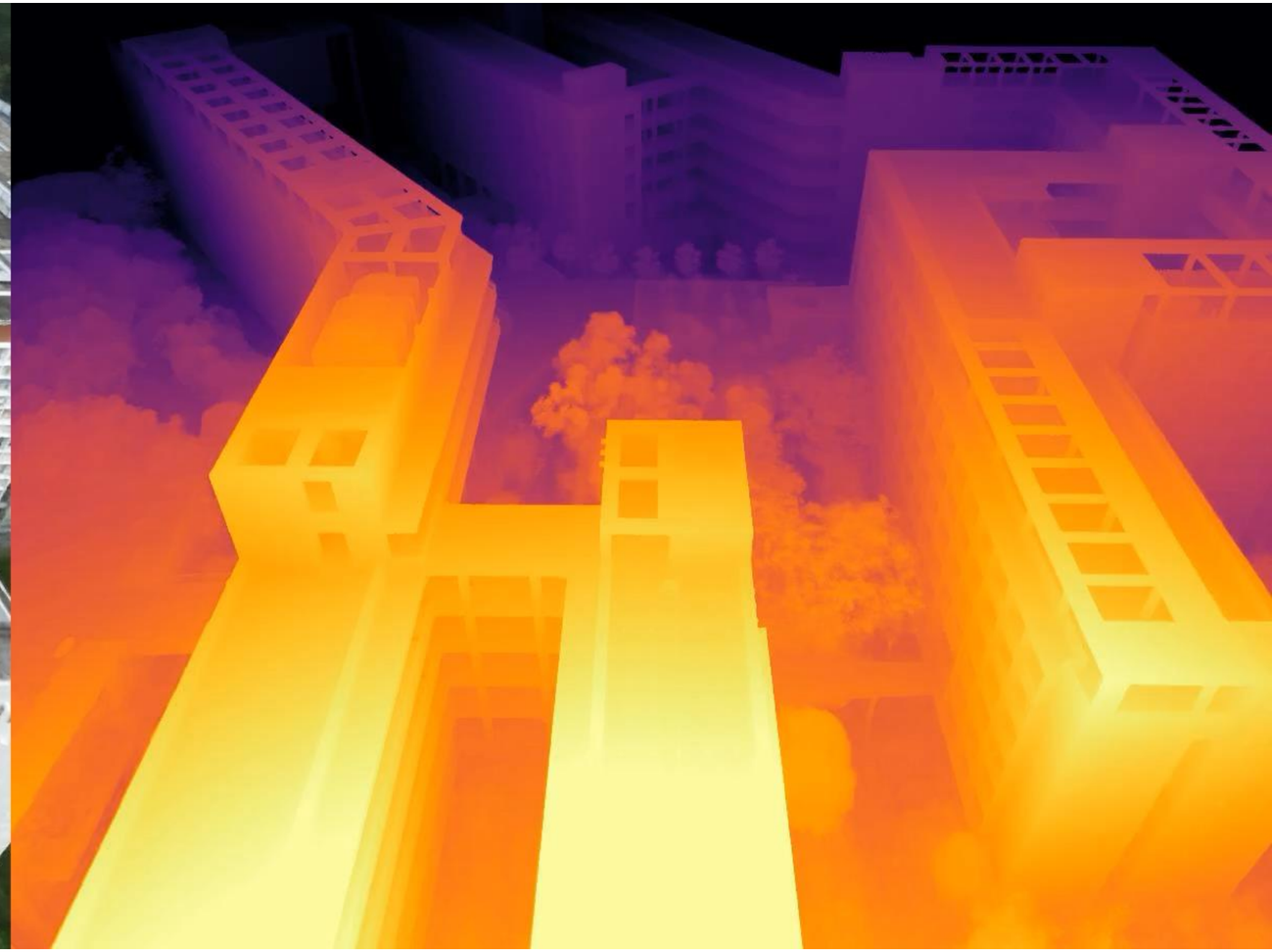
(b) Learning after physical-distribution-based decomposition (*e.g.* Block-NeRF)

(c) Learning with scene decomposition (Ours)

Research Highlights

THE DEPARTMENT OF
COMPUTER SCIENCE & ENGINEERING
計算機科學及工程學系

- 3D Rendering Results



Extend NeRF to 3D Gaussian Splatting

BungeeNeRF

PyGS (Ours)

Research Highlights

THE DEPARTMENT OF
COMPUTER SCIENCE & ENGINEERING
計算機科學及工程學系

- Human-Centric Video Generation
  - Video human body/head generation (**video**)
  - **Discover 3D geometry** from video -> Geometry-aware 3D generation

F. Hong, S. Li, D. Xu, 'DaGAN: Depth-Aware Generative Adversarial Network for Talking Head Video Generation', CVPR 2022.

Research Highlights

THE DEPARTMENT OF
COMPUTER SCIENCE & ENGINEERING
計算機科學及工程學系

- Human-Centric Video Generation

  - Video human body/head generation (**audio/text**)

  - **Hierarchical diffusion framework:**
    Input audio/text -> generation 3D facial landmarks -> video generation

We predict a sequence of accurate landmarks to provide the spatial cues for video synthesis.

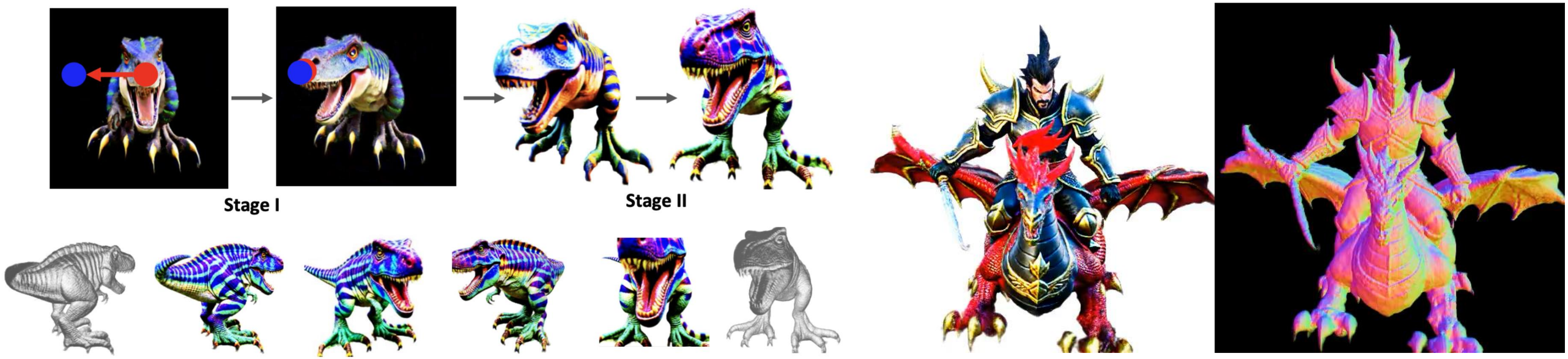
Research Highlights

THE DEPARTMENT OF
COMPUTER SCIENCE & ENGINEERING
計算機科學及工程學系

- Object-Centric Video Generation

- Interactive 3D generation:

- **edit during generation** -> generate objects that satisfy user expectation
- 3D Gaussian Splatting allows for flexible editing



S. Dong, L. Ding, T. Xue, D. Xu, ‘Interactive3D: Create What You Want by Interactive 3D Generation’, CVPR 2024.

Timestep = 9000

A King Kong.

A red axe.

A King Kong holding a red axe.

&

Original

Draged

Scene-Centric Video Generation

Hey kiddo! Here's your room! Decorate it by yourself!

Textured bedroom

Thanks!