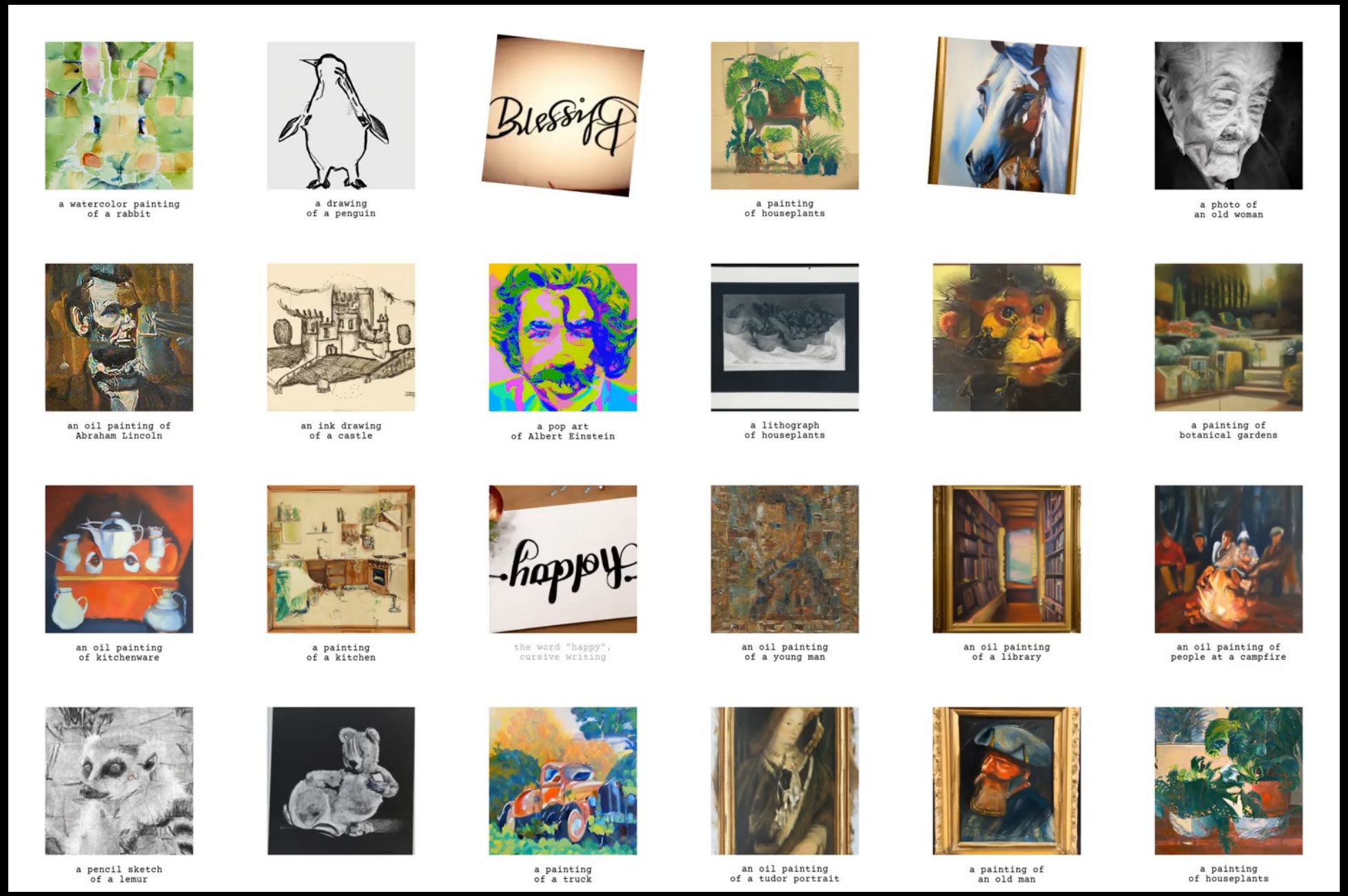
Modern Computer Vision Methods

Preliminary Meeting

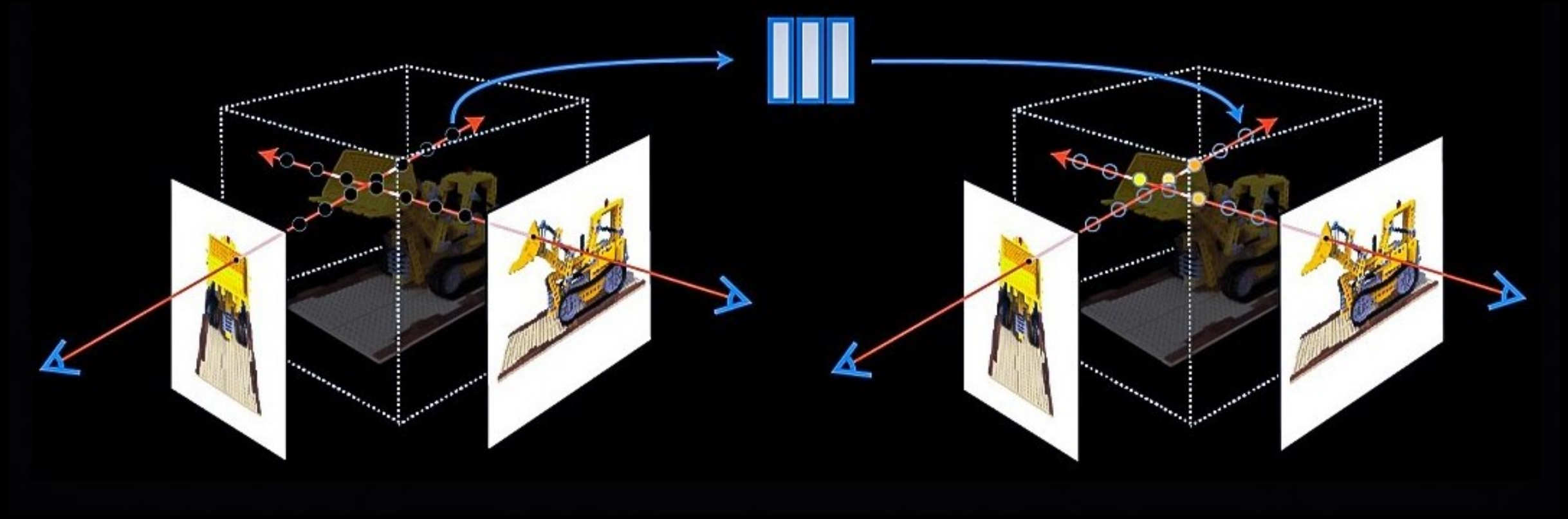
for WS 2024/25 [IN2107]

Dr. Benjamin Busam, Boody Elskhawy, Mert Kiray, Pengyuan Wang, Shishir Vutukur, Lennart Bastian, Junwen Huang, Ege Özsoy, Niko Brasch

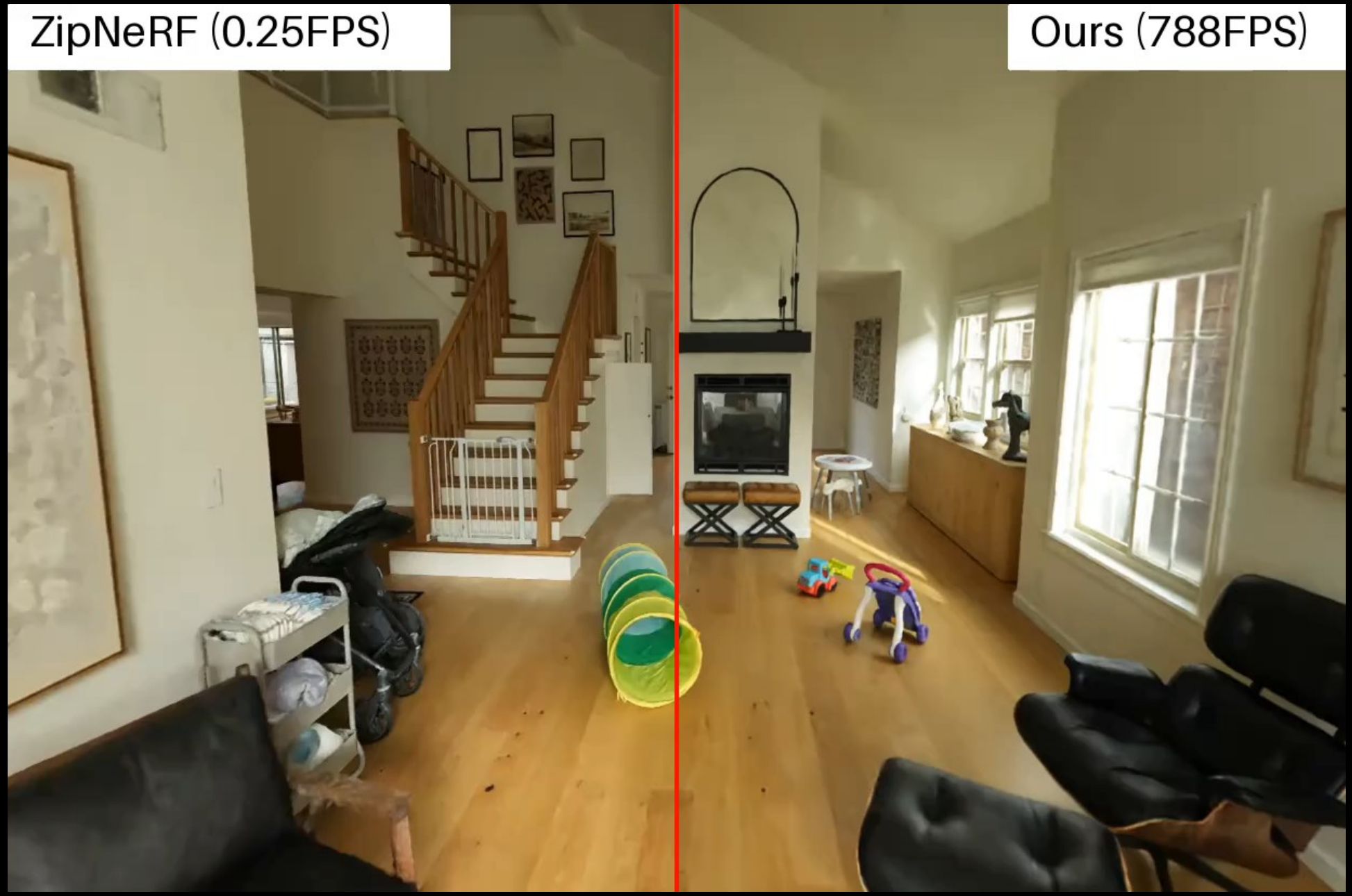


[Daniel Geng, Inbum Park, Andrew Owens. Visual Anagrams: Generating Multi-View Optical Illusions with Diffusion Models. CVPR 2024]

Novel View Synthesis



[Ben Mildenhall, Pratul P. Srinivasan, Matthew Tancik, Jonathan T. Barron, Ravi Ramamoorthi, Ren Ng. NeRF. ECCV 2022]



[Michael Niemeyer, Fabian Manhardt, Marie-Julie Rakotosaona, Michael Oechsle, Daniel Duckworth, Rama Gosula, Keisuke Tateno, John Bates, Dominik Kaeser, Federico Tombari. RadSplat: Radiance Field-Informed Gaussian Splatting for Robust Real-Time Rendering with 900+ FPS. arXiv 2024]

Goals

- Scientifically Learning about…
  - State-of-the-art Computer Vision
  - Current research challenges and applications
  - Communicate / discuss on most recent advantages with expert scientists
  - Hands-on experience with available code bases

- Skill training of…
  - Reading / understanding of a scientific work
  - Get overview of scientific field through literature research
  - Research talk in front of an audience, related Q&A

Seminar Contents

Most recent advances in Computer Vision field on

- Object Detection & Tracking
- 6D Object / Camera Pose Estimation
- Robotic Grasping / 3D Manipulation
- Generative Image / Video / Scene Synthesis
- 3D Scene Understanding / Reconstruction
- Multi-View Reconstruction
- Sensor Fusion / Multi-modal Imaging
- Universal Text & Vision Models

Presentation

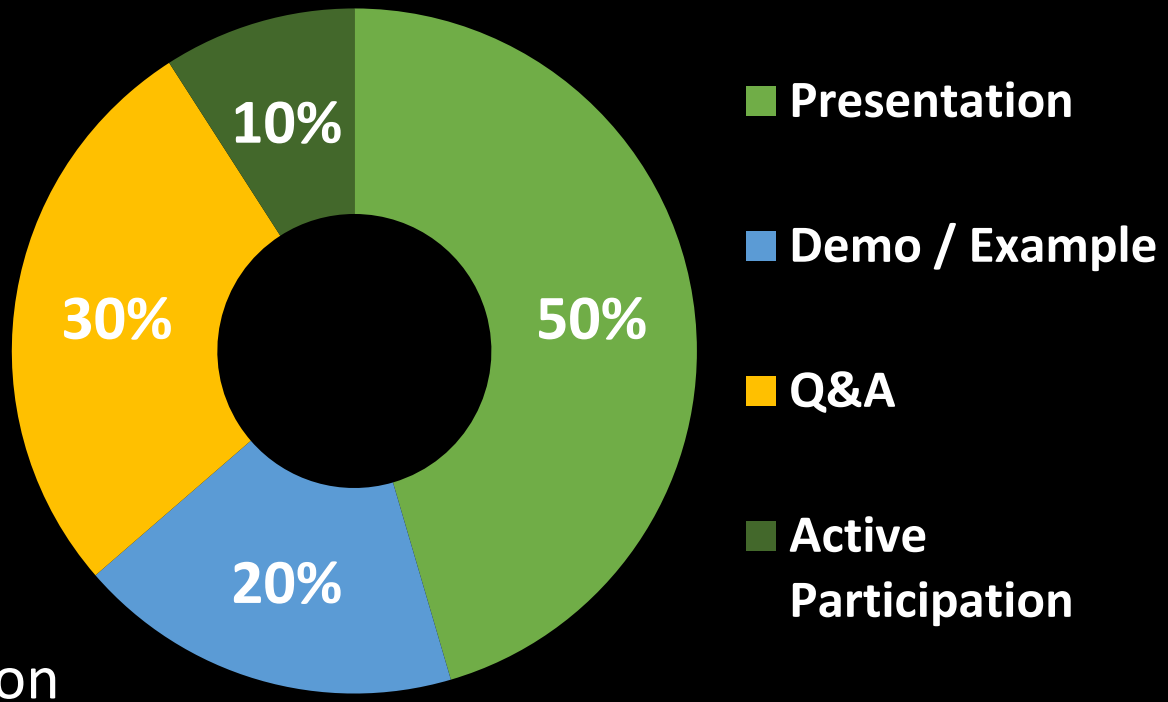
- Presentation: 20 minutes + 10-15 minutes Q&A
- Content should cover
  - Introduction / Relevance of Problem
  - Context / Related Work
  - Main Contribution(s)
  - Experimental Results
    - Hands-on experience with code
  - Discussion
  - Future Work
- Presentation should be self-contained
- Attend all talks + active participation in other discussions

Seminar Schedule

- 8 sessions (Tuesdays 16:00) + 1 intro + 1 presentation training
- 2 presentations per session (approx. 30-40 min each)
- Invited Talk(s): Renown computer vision researchers
- If necessary: hybrid meeting(s) via Zoom

- Topic assignment
  - Indicate preferences
  - Matching to maximize global happiness

Evaluation Criteria



- Quality of Presentation
  - Scientific Content of the Talk + Preparation
  - Quality of the Slides
  - Putting the Topic in Context (Related Work)
- Examples / Hands-on Code
- Scientific Discussion (Q&A)
- Independent Interaction / Active Participation in the Course

Some more Examples…

… with 3D computer vision applications

LangSplat: 3D Language Gaussian Splatting

- A 3D language field is learned by grounding CLIP language features into a set of 3D language Gaussians.

Rendered RGB Video

Visualization of Learned Language Feature[1]

[1]Different colors represent different language features.

[Minghan Qin, Wanhua Li, Jiawei Zhou, Haoqian Wang, Hanspeter Pfister. LangSplat: 3D Language Gaussian Splatting. CVPR 2024]

LangSplat: 3D Language Gaussian Splatting

- A 3D language field is learned by grounding CLIP language features into a set of 3D language Gaussians.

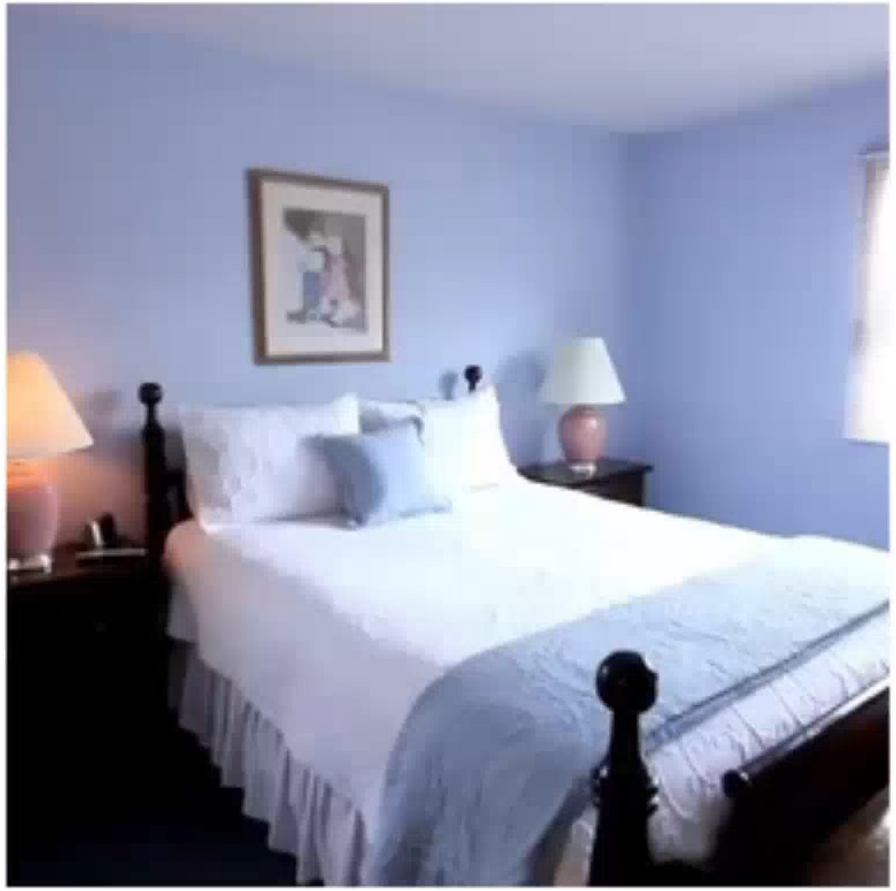
Rendered RGB Video

Visualization of Learned Language Feature[1]

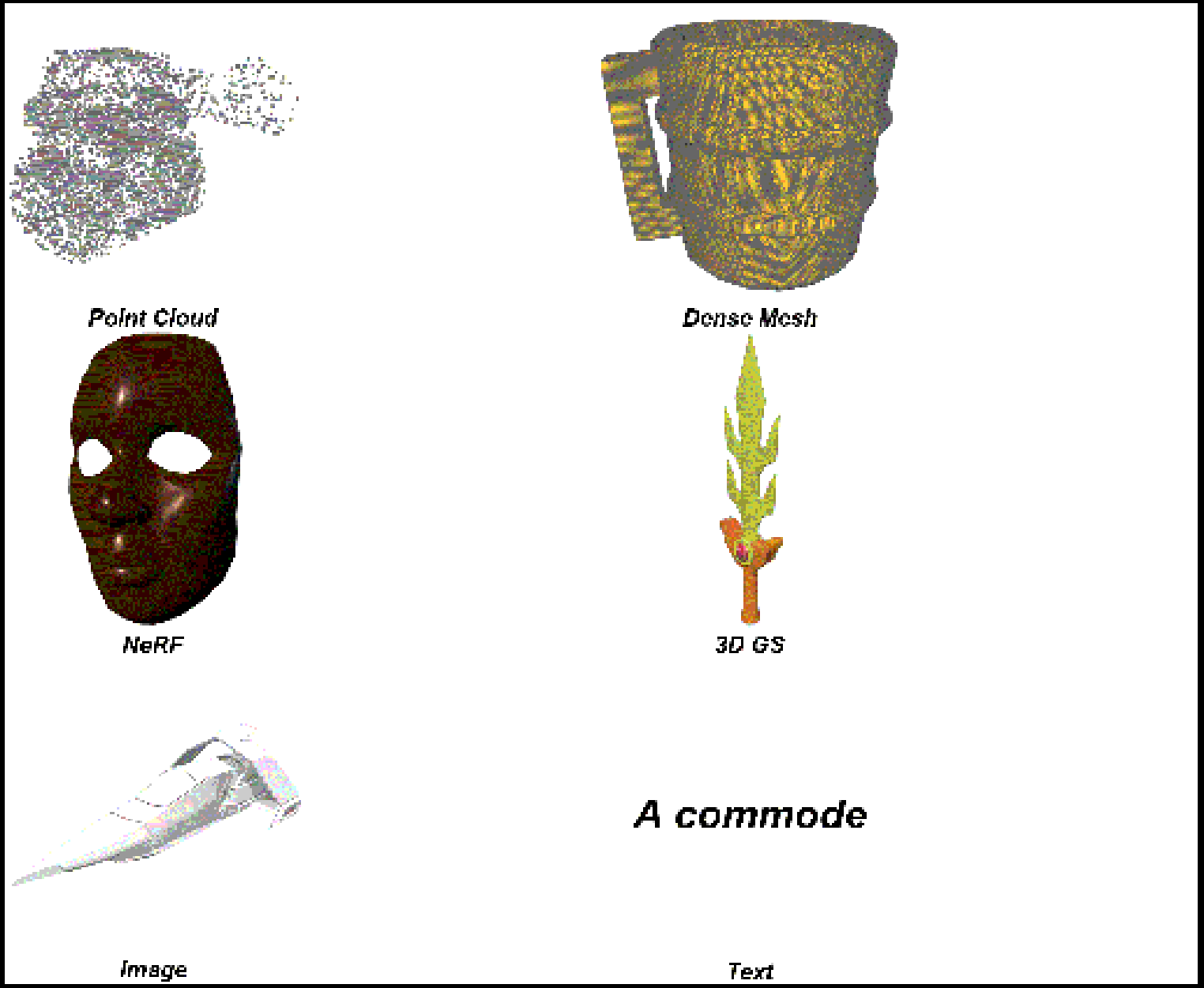
[1]Different colors represent different language features.

[Minghan Qin, Wanhua Li, Jiawei Zhou, Haoqian Wang, Hanspeter Pfister. LangSplat: 3D Language Gaussian Splatting. CVPR 2024]

2 Input Views



[David Charatan, Sizhe Li, Andrea Tagliasacchi, Vincent Sitzmann. pixelSplat: 3D Gaussian Splats from Image Pairs for Scalable Generalizable 3D Reconstruction. CVPR 2024]



[Yiwen Chen, Tong He, Di Huang, Weicai Ye, Sijin Chen, Jiaxiang Tang, Xin Chen, Zhongang Cai, Lei Yang, Gang Yu, Guosheng Lin, Chi Zhang. MeshAnything: Artist-Created Mesh Generation with Autoregressive Transformers. arXiv 2024]

Application

2 stage process

- Register in TUM Matching System
  http://docmatching.in.tum.de/index.php/schedule

- Submit motivation + background info to increase your chances
  mcvm@mailnavab.informatik.tu-muenchen.de

  Include:
  - Name, E-Mail, Study Program, Semester
  - Motivation + previous experience in Computer Vision (and related field)
  - (not mandatory): Submit your latest CV + transcript of records

- Deadline: 17th of July 2024

- 16 Students will be selected (usually 100+ applications)

Questions

E-Mail us on

mcvm@mailnavab.informatik.tu-muenchen.de

Your MCM Team:

Benjamin Busam, Boody Elskhawy, Mert Kiray, Pengyuan Wang,

Shishir Vutukur, Lennart Bastian, Junwen Huang, Ege Özsoy, Niko Brasch

Web:

https://www.cs.cit.tum.de/camp/teaching/seminars/modern-computer-vision-methods-ws-2024-25/