Modern Computer Vision Methods

Introduction Meeting

for WS 2023/24 [IN2107]

Dr. Benjamin Busam, Hyunjun Jung, Pengyuan Wang, Guangyao Zhai, Junwen Huang, Ege Özsoy, Felix Tristram, Hannah Schieber, Lars Heckler, Niko Brasch

MCVM Team


Benjamin
Busam


Ege
Özsoy


Felix
Tristram


Junwen
Huang


Lars
Heckler


Guangyao
Zhai


Pengyuan
Wang


Niko
Brasch


HyunJun
Jung


Hannah
Schieber


TEAM AWESOME!
memegenerator.net

MCVM

Course Structure

Course Dates

30.10.2023	**Introduction Session**
06.11.2023	**Invited Talk: Vasileios Belagiannis, FAU Erlangen-Nürnberg, GER**
13.11.2023	CVPR Break (Individual Preparation)
20.11.2023	Slot to meet with Supervisor
27.11.2023	**Presentation Training**
04.12.2023	**Data Distributions and Anomaly Detection**
11.12.2023	**Implicit 3D Reconstruction and Novel View Synthesis**
18.12.2023	**Invited Talk: Steven McDonagh, University of Edinburgh, UK**
15.01.2024	**Reconstructing a Dynamic World**
22.01.2024	**Implicit Representations for Humans**
29.01.2024	**Object Poses: Visual Ambiguities and Object Categories**
05.02.2024	**Generalizing to the Unseen with Text and Parts**

Course Dates

30.10.2023	**Introduction Session**
06.11.2023	**Invited Talk: Vasileios Belagiannis, FAU Erlangen-Nürnberg, GER**
13.11.2023	CVPR Break (Individual Preparation)
20.11.2023	Slot to meet with Supervisor
27.11.2023	**Presentation Training**
04.12.2023	**Data Distributions and Anomaly Detection**
11.12.2023	**Implicit 3D Reconstruction and Novel View Synthesis**
18.12.2023	**Invited Talk: Steven McDonagh, University of Edinburgh, UK**
15.01.2024	**Reconstructing a Dynamic World**
22.01.2024	**Implicit Representations for Humans**
29.01.2024	**Object Poses: Visual Ambiguities and Object Categories**
05.02.2024	**Generalizing to the Unseen with Text and Parts**

Course Dates

30.10.2023	**Introduction Session**
06.11.2023	**Invited Talk: Vasileios Belagiannis, FAU Erlangen-Nürnberg, GER**
13.11.2023	CVPR Break (Individual Preparation)
20.11.2023	Slot to meet with Supervisor
27.11.2023	**Presentation Training**
04.12.2023	**Data Distributions and Anomaly Detection**
11.12.2023	**Implicit 3D Reconstruction and Novel View Synthesis**
18.12.2023	**Invited Talk: Steven McDonagh, University of Edinburgh, UK**
15.01.2024	**Reconstructing a Dynamic World**
22.01.2024	**Implicit Representations for Humans**
29.01.2024	**Object Poses: Visual Ambiguities and Object Categories**
05.02.2024	**Generalizing to the Unseen with Text and Parts**

Course Dates

30.10.2023 **Introduction Session**

06.11.2023 **Invited Talk: Vasileios Belagiannis, FAU Erlangen-Nürnberg, GER**

13.11.2023 CVPR Break (Individual Preparation)

20.11.2023 Slot to meet with Supervisor

27.11.2023 **Presentation Training**

04.12.2023 **Data Distributions and Anomaly Detection**

11.12.2023 **Implicit 3D Reconstruction and Novel View Synthesis**

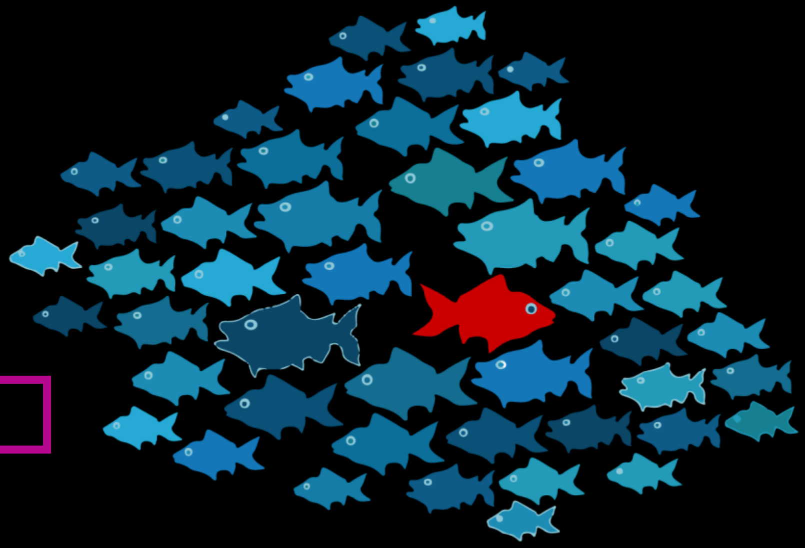
18.12.2023 **Invited Talk: Steven McDonagh, University of Edinburgh, UK**

15.01.2024 **Reconstructing a Dynamic World**

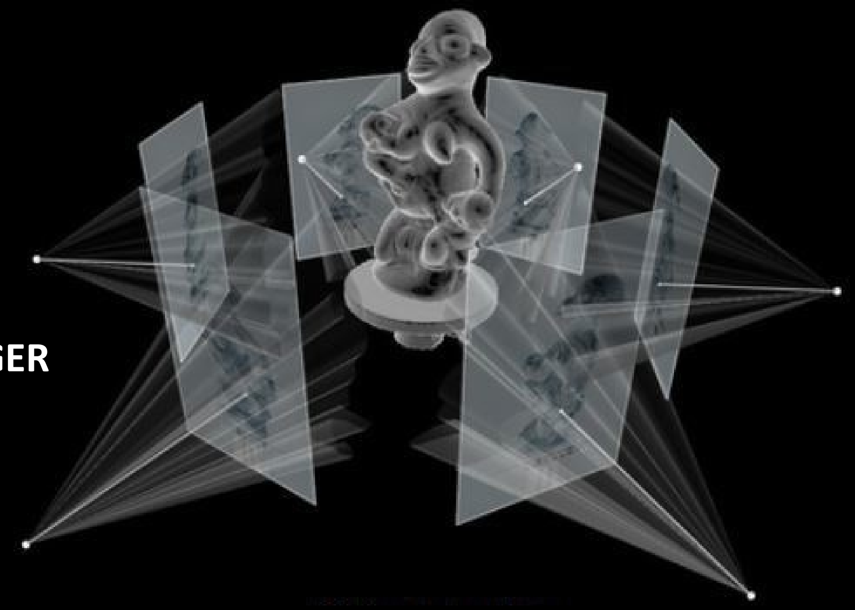
22.01.2024 **Implicit Representations for Humans**

29.01.2024 **Object Poses: Visual Ambiguities and Object Categories**

05.02.2024 **Generalizing to the Unseen with Text and Parts**



Course Dates



30.10.2023 **Introduction Session**

06.11.2023 **Invited Talk: Vasileios Belagiannis, FAU Erlangen-Nürnberg, GER**

13.11.2023 CVPR Break (Individual Preparation)

20.11.2023 Slot to meet with Supervisor

27.11.2023 **Presentation Training**

04.12.2023 **Data Distributions and Anomaly Detection**

11.12.2023 **Implicit 3D Reconstruction and Novel View Synthesis**

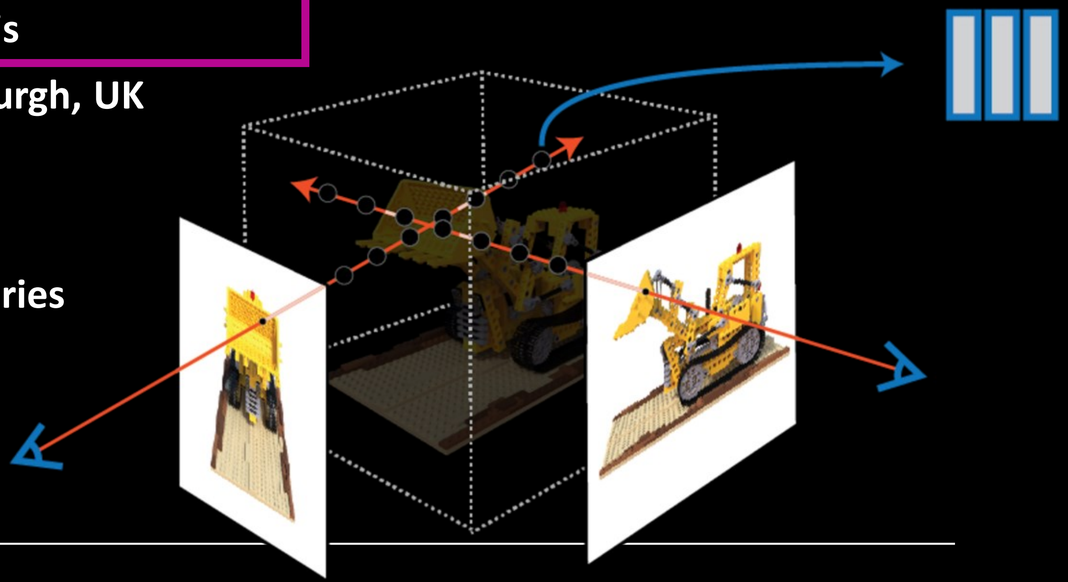
18.12.2023 **Invited Talk: Steven McDonagh, University of Edinburgh, UK**

15.01.2024 **Reconstructing a Dynamic World**

22.01.2024 **Implicit Representations for Humans**

29.01.2024 **Object Poses: Visual Ambiguities and Object Categories**

05.02.2024 **Generalizing to the Unseen with Text and Parts**



Course Dates

30.10.2023 **Introduction Session**
06.11.2023 **Invited Talk: Vasileios Belagiannis, FAU Erlangen-Nürnberg, GER**
13.11.2023 CVPR Break (Individual Preparation)
20.11.2023 Slot to meet with Supervisor
27.11.2023 **Presentation Training**
04.12.2023 **Data Distributions and Anomaly Detection**
11.12.2023 **Implicit 3D Reconstruction and Novel View Synthesis**
18.12.2023 **Invited Talk: Steven McDonagh, University of Edinburgh, UK**
15.01.2024 **Reconstructing a Dynamic World**
22.01.2024 **Implicit Representations for Humans**
29.01.2024 **Object Poses: Visual Ambiguities and Object Categories**
05.02.2024 **Generalizing to the Unseen with Text and Parts**

Course Dates

30.10.2023 **Introduction Session**
06.11.2023 **Invited Talk: Vasileios Belagiannis, FAU Erlangen-Nürnberg, GER**
13.11.2023 CVPR Break (Individual Preparation)
20.11.2023 Slot to meet with Supervisor
27.11.2023 **Presentation Training**
04.12.2023 **Data Distributions and Anomaly Detection**
11.12.2023 **Implicit 3D Reconstruction and Novel View Synthesis**
18.12.2023 **Invited Talk: Steven McDonagh, University of Edinburgh, UK**
15.01.2024 **Reconstructing a Dynamic World**
22.01.2024 **Implicit Representations for Humans**
29.01.2024 **Object Poses: Visual Ambiguities and Object Categories**
05.02.2024 **Generalizing to the Unseen with Text and Parts**

Course Dates

30.10.2023 **Introduction Session**

06.11.2023 **Invited Talk: Vasileios Belagiannis, FAU Erlangen-Nürnberg, GER**

13.11.2023 CVPR Break (Individual Preparation)

20.11.2023 Slot to meet with Supervisor

27.11.2023 **Presentation Training**

04.12.2023 **Data Distributions and Anomaly Detection**

11.12.2023 **Implicit 3D Reconstruction and Novel View Synthesis**

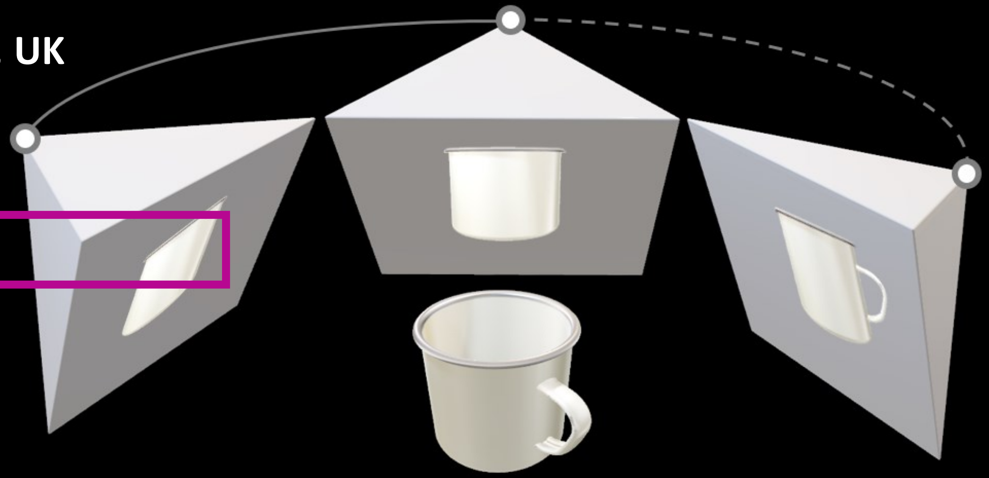
18.12.2023 **Invited Talk: Steven McDonagh, University of Edinburgh, UK**

15.01.2024 **Reconstructing a Dynamic World**

22.01.2024 **Implicit Representations for Humans**

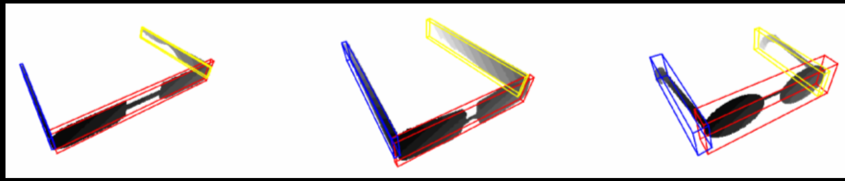
29.01.2024 **Object Poses: Visual Ambiguities and Object Categories**

05.02.2024 **Generalizing to the Unseen with Text and Parts**

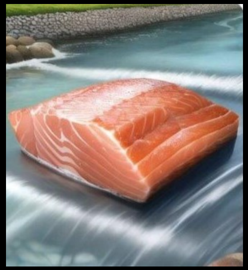


Course Dates

30.10.2023	**Introduction Session**
06.11.2023	**Invited Talk: Vasileios Belagiannis, FAU Erlangen-Nürnberg, GER**
13.11.2023	CVPR Break (Individual Preparation)
20.11.2023	Slot to meet with Supervisor
27.11.2023	**Presentation Training**
04.12.2023	**Data Distributions and Anomaly Detection**
11.12.2023	**Implicit 3D Reconstruction and Novel View Synthesis**
18.12.2023	**Invited Talk: Steven McDonagh, University of Edinburgh, UK**
15.01.2024	**Reconstructing a Dynamic World**
22.01.2024	**Implicit Representations for Humans**
29.01.2024	**Object Poses: Visual Ambiguities and Object Categories**
05.02.2024	**Generalizing to the Unseen with Text and Parts**



„Salmon swimming in river“



Course Dates

30.10.2023	**Introduction Session**
06.11.2023	**Invited Talk: Vasileios Belagiannis, FAU Erlangen-Nürnberg, GER**
13.11.2023	CVPR Break (Individual Preparation)
20.11.2023	Slot to meet with Supervisor
27.11.2023	**Presentation Training**
04.12.2023	**Data Distributions and Anomaly Detection**
11.12.2023	**Implicit 3D Reconstruction and Novel View Synthesis**
18.12.2023	**Invited Talk: Steven McDonagh, University of Edinburgh, UK**
15.01.2024	**Reconstructing a Dynamic World**
22.01.2024	**Implicit Representations for Humans**
29.01.2024	**Object Poses: Visual Ambiguities and Object Categories**
05.02.2024	**Generalizing to the Unseen with Text and Parts**

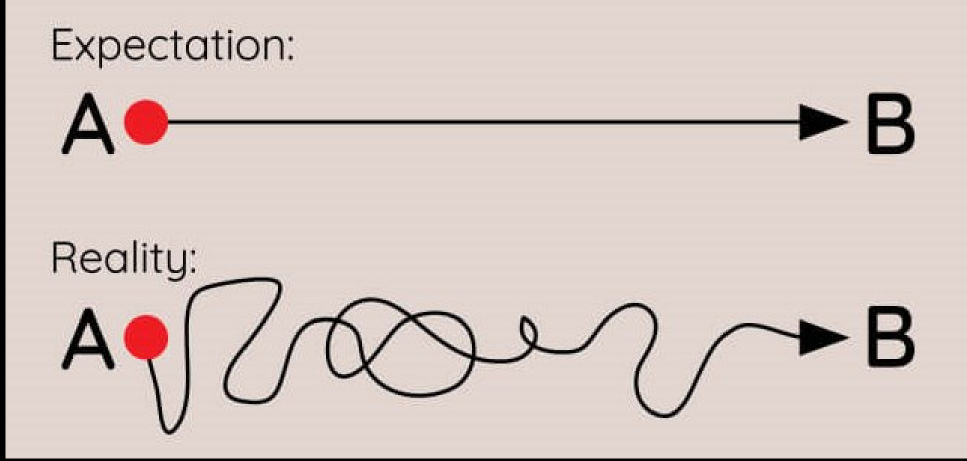
Paper Overview

Authors	Title	Conference	Year	Supervisor
Sheng, Jing, Jiao, Wang, Dong	MÆIDM: multi-scale anomaly embedding inpainting and discrimination for surface anomaly detection	MVA	2023	Lars
Jeong, Zou, Kim, Zhang, Ravichandran, Dabeer	WinCLIP: Zero-/Few-Shot Anomaly Classification and Segmentation	CVPR	2023	Lars
Truong, Rakotosaona, Manhardt, Tombari	SPARF: Neural Radiance Fields from Sparse and Noisy Poses	CVPR	2023	HyunJun
Li, Müller, Evans, Taylor, Unberath, Liu, Lin	Neuralangelo: High-Fidelity Neural Surface Reconstruction	CVPR	2023	Hannah
Li, Wang, Cole, Tucker, Snavely.	Dynibar: Neural dynamic image-based rendering	CVPR	2023	Felix
Qiu, Chen, Zhou, Xu	REC-MV: REconstructing 3D Dynamic Cloth from Monocular Videos	CVPR	2023	Niko
Jiang, Ren, Dou, Xue, Fu, Zhang	LoRD: Local 4D Implicit Representation for High-Fidelity Dynamic Human Modeling	ECCV	2022	Niko
Guo, Jiang, Chen, Song, Hilliges	Vid2Avatar: 3D Avatar Reconstruction from Videos in the Wild via Self-supervised Scene Decomposition	CVPR	2023	Niko
Haugaard, Buch	SurfEmb:Dense and Continuous Correspondence Distributions for Object Pose Estimation with Learnt Surface Embeddings	CVPR	2022	Junwen
Lin, Wei, Zhang, Jia	VI-Net: Boosting Category-level 6D Object Pose Estimation via Learning Decoupled Rotations on the Spherical Representations	ICCV	2023	Pengyuan
Zhang, Pan, Yao, Huang, Mei, Chen	Learning to Generate Language-supervised and Open-vocabulary Scene Graph using Pre-trained Visual-Semantic Space	CVPR	2023	Ege
Geng, Xu, Zhao, Xu, Yi, Huang, Wang	GAPartNet: Cross-Category Domain-Generalizable Object Perception and Manipulation via Generalizable and Actionable Parts	CVPR	2023	Guangyao

In Person / Virtual – Hybrid

- Mostly onsite. Following government / TUM regulations
- In exceptional cases: virtual via zoom
- Mondays at 12 noon in MI 03.13.010

What we expect from you



- Interest in Computer Vision
- Independent and pro-active participation
- Actively asking for help [supervisor meetings]
- Coding knowledge
- Illustrating methods with examples / demos

Goals

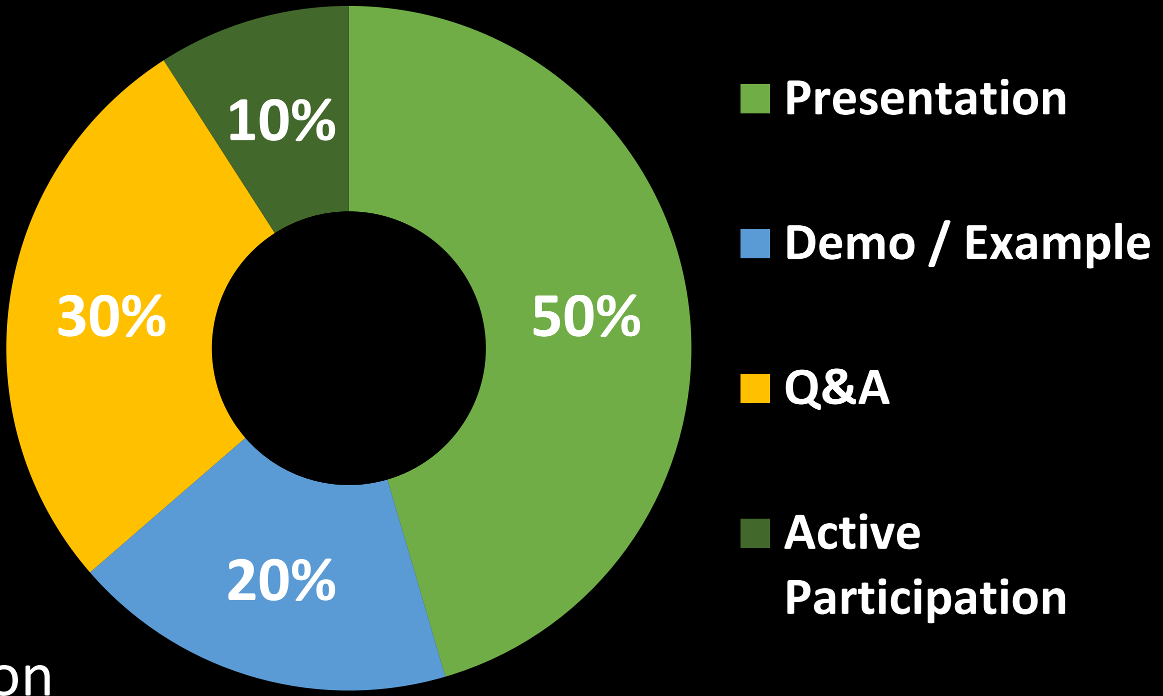
- Scientifically Learning about…
  - State-of-the-art Computer Vision
  - Current research challenges and applications
  - Communicate / discuss on most recent advantages with expert scientists
  - Hands-on experience with available code bases

- 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

Presentation

- Presentation: 20 ± 2 minutes talk + 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

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

Seed Paper Intros

MCVM WiSe 23/24

Anomaly Detection and Localization

Lars Heckler

Summary

- high quality products from modern production systems
- lack of defective samples
- mostly unsupervised approaches requiring only normal data
- anomaly detection as a crucial tool for quality assurance

Open Questions

- To what extend can LLMs be a useful tool (defect categories, applicability on site)?
- What are the chances and risks of using synthetic defects for training anomaly detection systems?

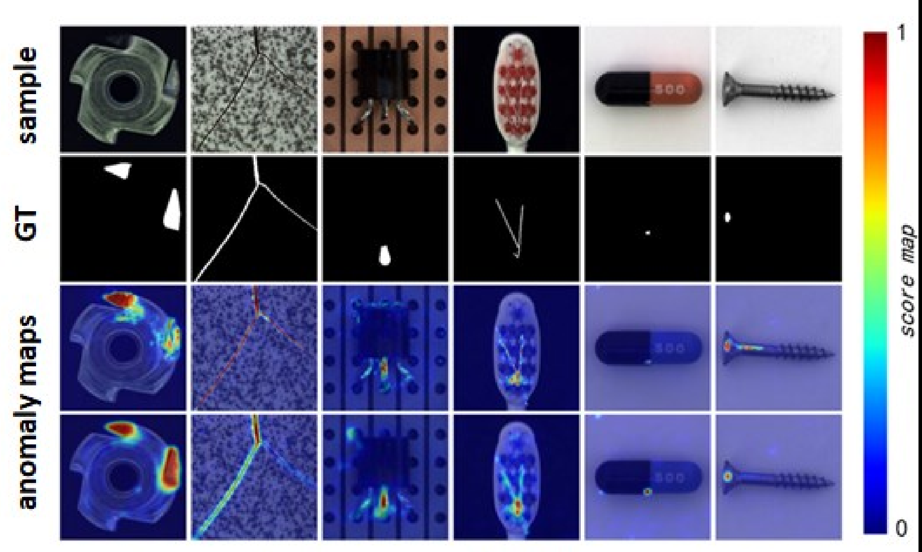


figure taken from Sheng et al. 2023 (captions modified)

Jongheon Jeong, Yang Zou, Taewan Kim, Dongqing Zhang, Avinash Ravichandran, Onkar Dabeer. WinCLIP: Zero-/Few-Shot Anomaly Classification and Segmentation. CVPR 2023.

Jing Siyu Sheng, Junfeng Jing, Xintian Jiao, Yafei Wang, Zhenyu Dong. MÆIDM: multi-scale anomaly embedding inpainting and discrimination for surface anomaly detection. Machine Vision and Applications 2023.

SPARF

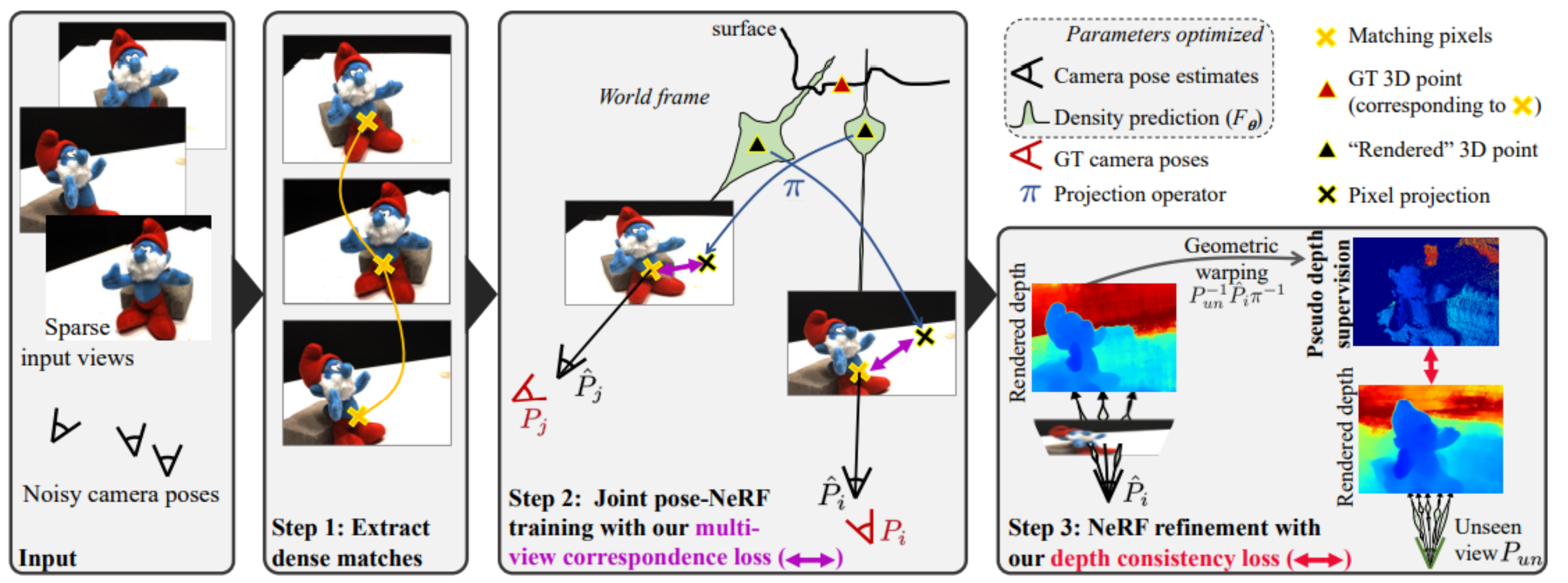
Hyun Jun Jung

Introduction

- Neural Radiance Field requires many images and good camera poses
- New NeRF pipeline that can be trained spare images and noisy camera pose

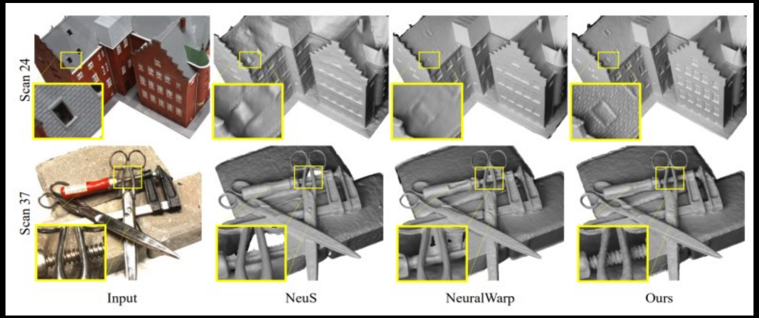
Method

- Enforce geometric loss with multiview dense matching
- Update camera pose together with NeRF geometry improve the result



P.Truong, M.J.Rakotosaona, F.Manhardt, F.Tombari. SPARF : Neural Radiance Fields from Sparse and Nosy Poses (CVPR 2023)

Neuralangelo



Hannah Schieber

Summary

- NeRF tackles mainly novel view synthesis. However, extracting surfaces meshes from NeRF results in low quality. Neus/Neuralangelo target surface reconstruction
- Neuralangelo combines the representation power of multi-resolution 3D hash grids with neural surface rendering --> enables recovering details
- Works with RGB and RGB-D

Key ingredients

- (1) numerical gradients for computing higher-order derivatives as a smoothing operation
- (2) coarse-to-fine optimization on the hash grids controlling different levels of details. Even without auxiliary inputs such as depth

Objectives

- Find the limitations of the work, under what conditions is Neuralangelo challenged (which are not mentioned in the paper), produce staysfying/unsatisfying reconstruction results

Code: https://github.com/nvlabs/neuralangelo, https://github.com/hugoycj/Instant-angelo

Related Work

- InstantNGP - Müller, T., Evans, A., Schied, C., & Keller, A. (2022). Instant neural graphics primitives with a multiresolution hash encoding. ACM Transactions on Graphics (ToG), 41(4), 1-15.
- Neus - Wang, P., Liu, L., Liu, Y., Theobalt, C., Komura, T., & Wang, W. (2021). Neus: Learning neural implicit surfaces by volume rendering for multi-view reconstruction. arXiv preprint arXiv:2106.10689.
- Dogaru, A., Ardelean, A. T., Ignatyev, S., Zakharov, E., & Burnaev, E. (2023). Sphere-Guided Training of Neural Implicit Surfaces. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 20844-20853).

Li, Z., Müller, T., Evans, A., Taylor, R. H., Unberath, M., Liu, M. Y., & Lin, C. H. (2023). Neuralangelo: High-Fidelity Neural Surface Reconstruction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 8456-8465).

DynIBaR

Felix Tristram

Background

- **AR/VR requires rendering a (dynamic!) scene from many different varying viewpoints**
- **Recent Approaches are very promising in a static setting but this targets dynamics**

Method

- **Feature Aggregation and Image-Based Rendering is used to create NVS of dynamic scenes**
- **Motion-Trajectory Fields represent scene motion over long video sequences**
- **Works even in very long video sequences and 'in the wild'**

Objectives

- **Understand the paper, find potential failure points, maybe run with self-captured video!**

Li, Z., Wang, Q., Cole, F., Tucker, R., & Snavely, N. (2023). Dynibar: Neural dynamic image-based rendering. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition

4D Human Reconstruction

Nikolas Brasch

- Background
  - 3D Human Pose -> 3D Body Shape -> Clothed humans
  - 4D Reconstruction with large deformations
  - Problem is constrained by space of human motions
- Applications
  - XR Streaming
  - Human avatars

- Open questions
  - Reconstruct detailed deformations (wrinkles, stretching and inertia of cloth, hair, ...)
  - Disentangle independent parts (clothing items, skin, hair)
  - Propagate temporal information to occluded parts to reduce number of views needed

Lingteng Qiu, Guanying Chen, Jiapeng Zhou, Mutian Xu REC-MV: REconstructing 3D Dynamic Cloth from Monocular Videos, CVPR 2023
Guo, C., Jiang, T., Chen, X., Song, J., & Hilliges, O. Vid2Avatar: 3D Avatar Reconstruction from Videos in the Wild via Self-supervised Scene Decomposition, CVPR 2023
Jiang, B., Ren, X., Dou, M., Xue, X., Fu, Y., Zhang, Y. LoRD: Local 4D Implicit Representation for High-Fidelity Dynamic Human Modeling, ECCV 2022

SurfEmb: Dense and Continuous Correspondence Distributions for Object Pose Estimation with Learnt Surface Embeddings

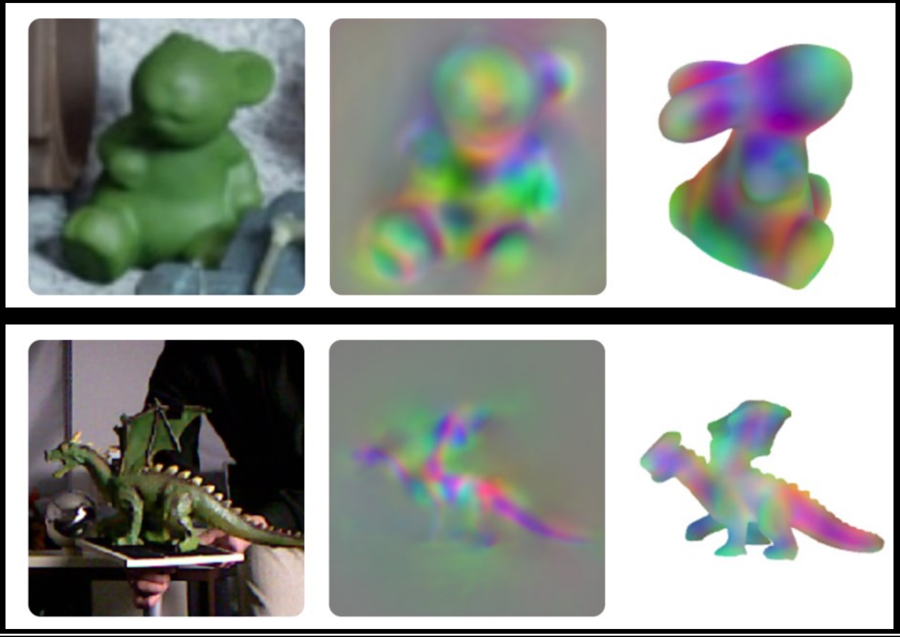
Junwen Huang

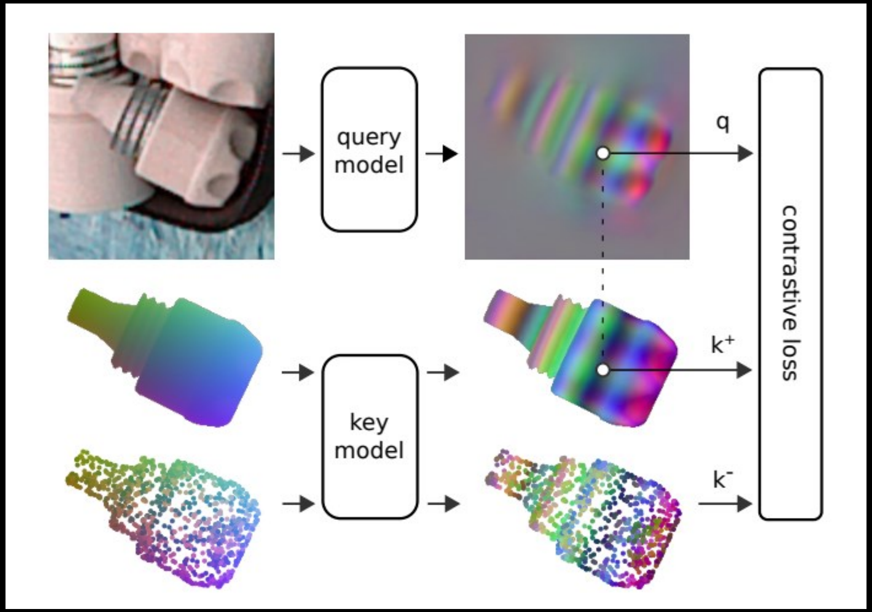
Introduction

- Instance-level object pose estimation
- Surface-pixel correspondence

Method

- Contrastive learning between 2D RGB pixel and 3D model surface
- PNP projection of 2D-3D correspondence





J. Lin, Z. Wei, Y. Zhang, K. Jia. VI-Net: Boosting Category-level 6D Object Pose Estimation via Learning Decoupled Rotations on the Spherical Representations, ICCV, 2023

VI-Net: Boosting Category-level 6D Object Pose Estimation via Learning Decoupled Rotations on the Spherical Representations

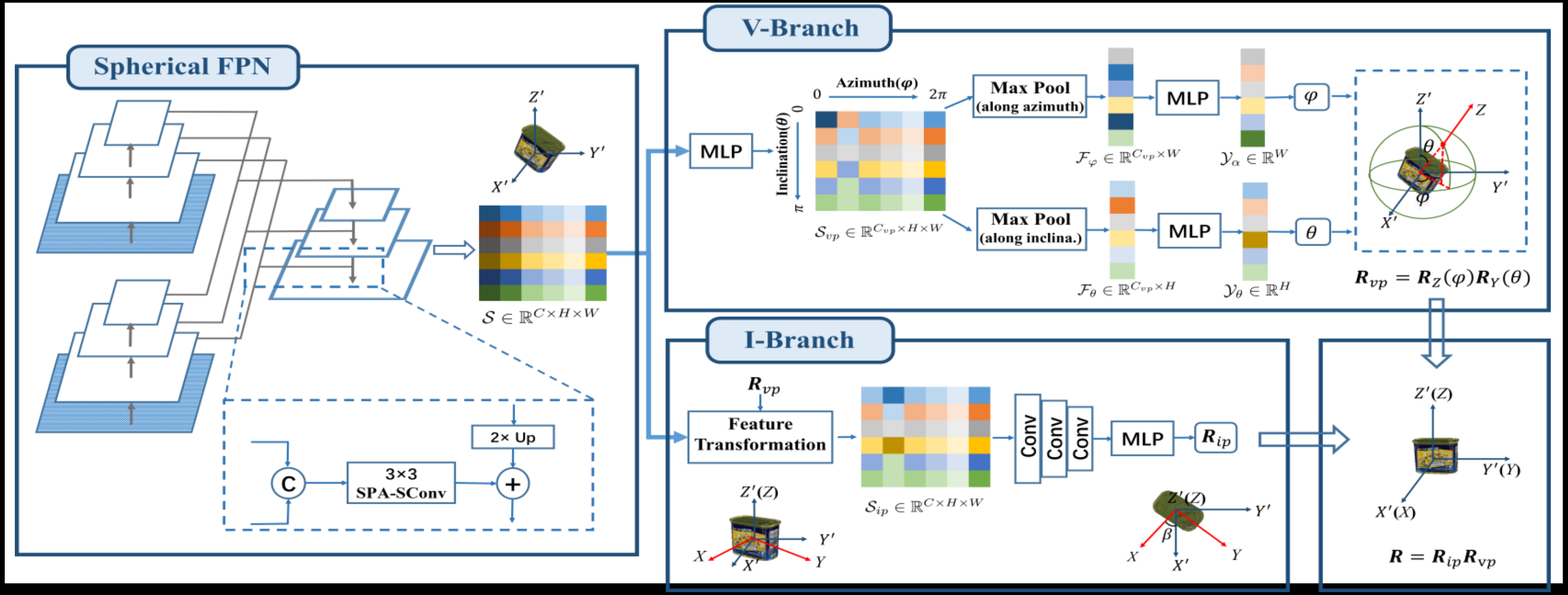
Pengyuan Wang

Introduction

- Category-level object pose estimation
- Accurate estimation of object rotations

Method

- Decoupled rotation estimation mechanism
- Spherical Feature Pyramid Network for feature extraction

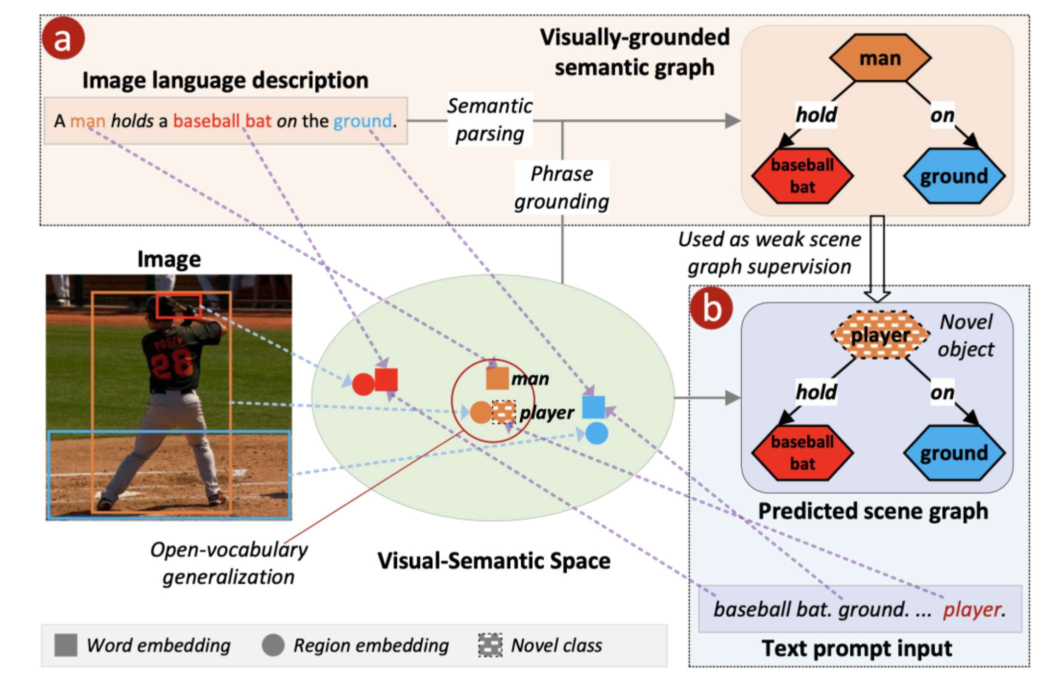


J. Lin, Z. Wei, Y. Zhang, K. Jia. VI-Net: Boosting Category-level 6D Object Pose Estimation via Learning Decoupled Rotations on the Spherical Representations, ICCV, 2023

Learning to Generate Language-supervised and Open-vocabulary Scene Graph using Pre-trained Visual-Semantic Space

Ege Özsoy

- Ideal Goal: Holistic Scene Understanding with Scene Graphs with limited data
- Seed Paper: Utilizes language supervision which is readily available for SGG
- Paves the way for the use of scene graphs for limited data domains as well as for novel objects and relation
- Potentially interesting additional preprint: ConceptGraphs https://arxiv.org/abs/2309.16650



Zhang, Yong, Yingwei Pan, Ting Yao, Rui Huang, Tao Mei, and Chang-Wen Chen. "Learning To Generate Language-Supervised and Open-Vocabulary Scene Graph Using Pre-Trained Visual-Semantic Space". CVPR. 2023.

GAPartNet

Guangyao Zhai

Problem

- For years, researchers have been devoted to generalizable object perception and manipulation, where cross category generalizability is highly desired yet underexplored.

Contributions

- (1) A large-scale interactive dataset, GAPartNet, with rich part semantics and pose annotations that facilitates generalizable part perception and part-based object manipulation.
- (2) A pipeline for domain-generalizable 3D part segmentation and pose estimation via learning domain-invariant features
- (3) A new solution to generalizable object manipulation by leveraging the concept of GAParts.

Objectives

- Find the limitations of the work, and run a demo.

Website: https://pku-epic.github.io/GAPartNet

Geng, H., Xu, H., Zhao, C., Xu, C., Yi, L., Huang, S. and Wang, H. Gapartnet: Cross-category domain-generalizable object perception and manipulation via generalizable and actionable parts. CVPR 2023.

Course Dates

30.10.2023	**Introduction Session**	
06.11.2023	**Invited Talk: Vasileios Belagiannis, FAU Erlangen-Nürnberg, GER**	
13.11.2023	CVPR Break (Individual Preparation)	
20.11.2023	Slot to meet with Supervisor	
27.11.2023	**Presentation Training**	
04.12.2023	**Data Distributions and Anomaly Detection**	**MÆIDM + WinCLIP**
11.12.2023	**Implicit 3D Reconstruction and Novel View Synthesis**	**SPARF + Neuralangelo**
18.12.2023	**Invited Talk: Steven McDonagh, University of Edinburgh, UK**	
15.01.2024	**Reconstructing a Dynamic World**	**Dynibar + REC-MV**
22.01.2024	**Implicit Representations for Humans**	**LoRD + Vid2Avatar**
29.01.2024	**Object Poses: Visual Ambiguities and Object Categories**	**SurfEmb + VI-Net**
05.02.2024	**Generalizing to the Unseen with Text and Part Understanding**	**OpenVoc Scene Graph + GAPartNet**

Next Steps

Paper Selection

https://forms.gle/XWDw6iaoERwn395U7

Deadline: November 05, 2023

- We optimize for global happiness

Next Meeting: Invited Talk by Vasileios Belagiannis

- Monday, November 06 at 12 noon in MI 03.13.010

Questions

E-Mail us on

mcvm@mailnavab.informatik.tu-muenchen.de

Your MCVM Team:

Benjamin Busam, Hyunjun Jung, Pengyuan Wang, Guangyao Zhai,

Junwen Huang, Ege Özsoy, Felix Tristram, Hannah Schieber, Lars Heckler, Niko Brasch

Web:

https://www.cs.cit.tum.de/camp/teaching/seminars/modern-computer-vision-methods-ws-2023-24/