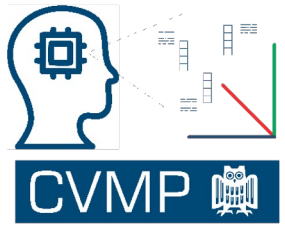
Computer Vision and Machine Perception Lab

Eddy Ilg

PhD Positions Available



Computer Vision and Machine Perception Lab at Saarland University

https://cvmp.cs.uni-saarland.de

About Me

M.Sc. Computer Science,
University of Southern California
Specialization in **Artificial Intelligence**

M.Sc. Computer Science
University of Freiburg
Specialization in **Robotics and Computer Vision**

PhD Computer Vision
University of Freiburg
Thesis: **Estimating Optical Flow with Convolutional Neural Networks**

Senior Research Scientist at Facebook Reality Labs
working in Augmented Reality
on **Object Reconstruction in the Wild**

Professor at Saarland Univeristy
for **Computer Vision**

Supported by:

DAAD
Deutscher Akademischer Austausch Dienst
German Academic Exchange Service

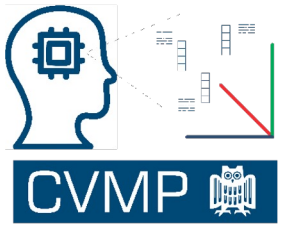
Studienstiftung des deutschen Volkes

Known for:

convolutional network

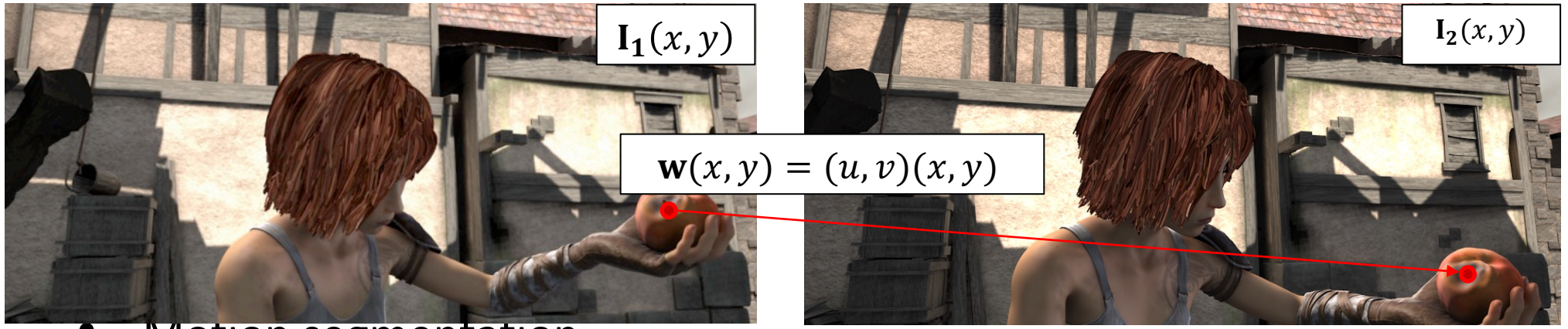
FlowNet and FlowNet 2.0

Past Work

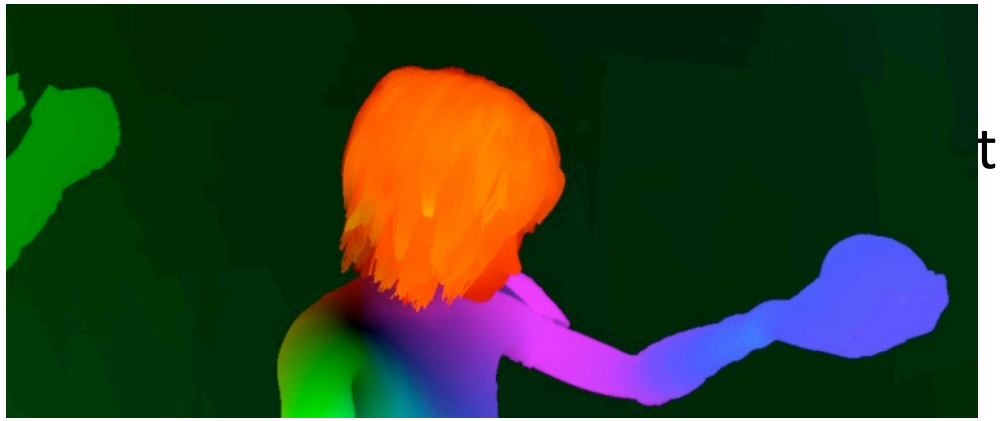


Computer Vision and
Machine Perception Lab
at Saarland University

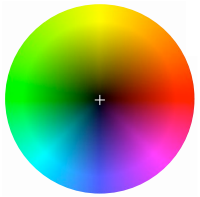
Optical Flow



- Motion segmentation



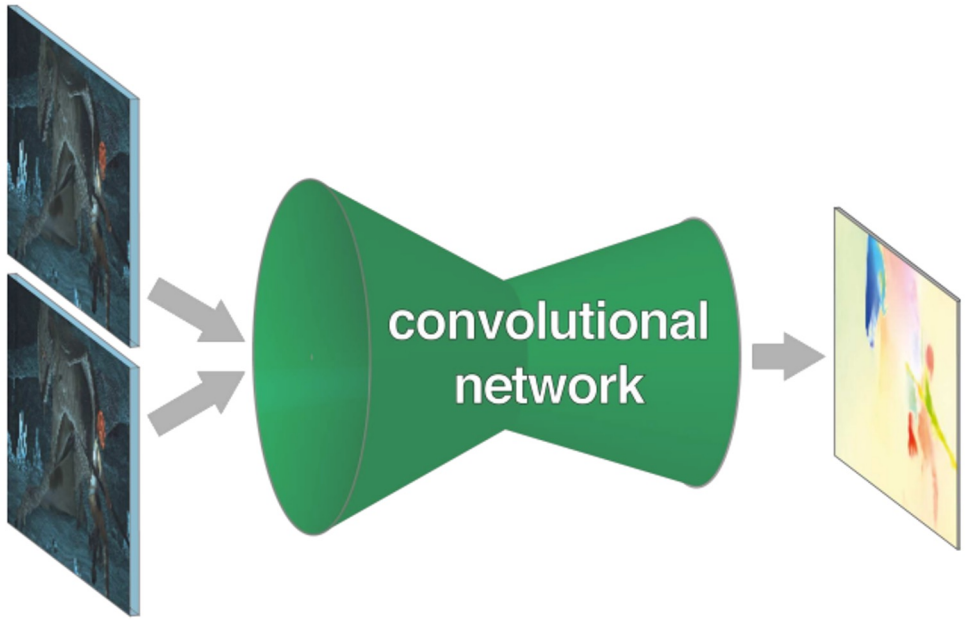
tion



FlowNet

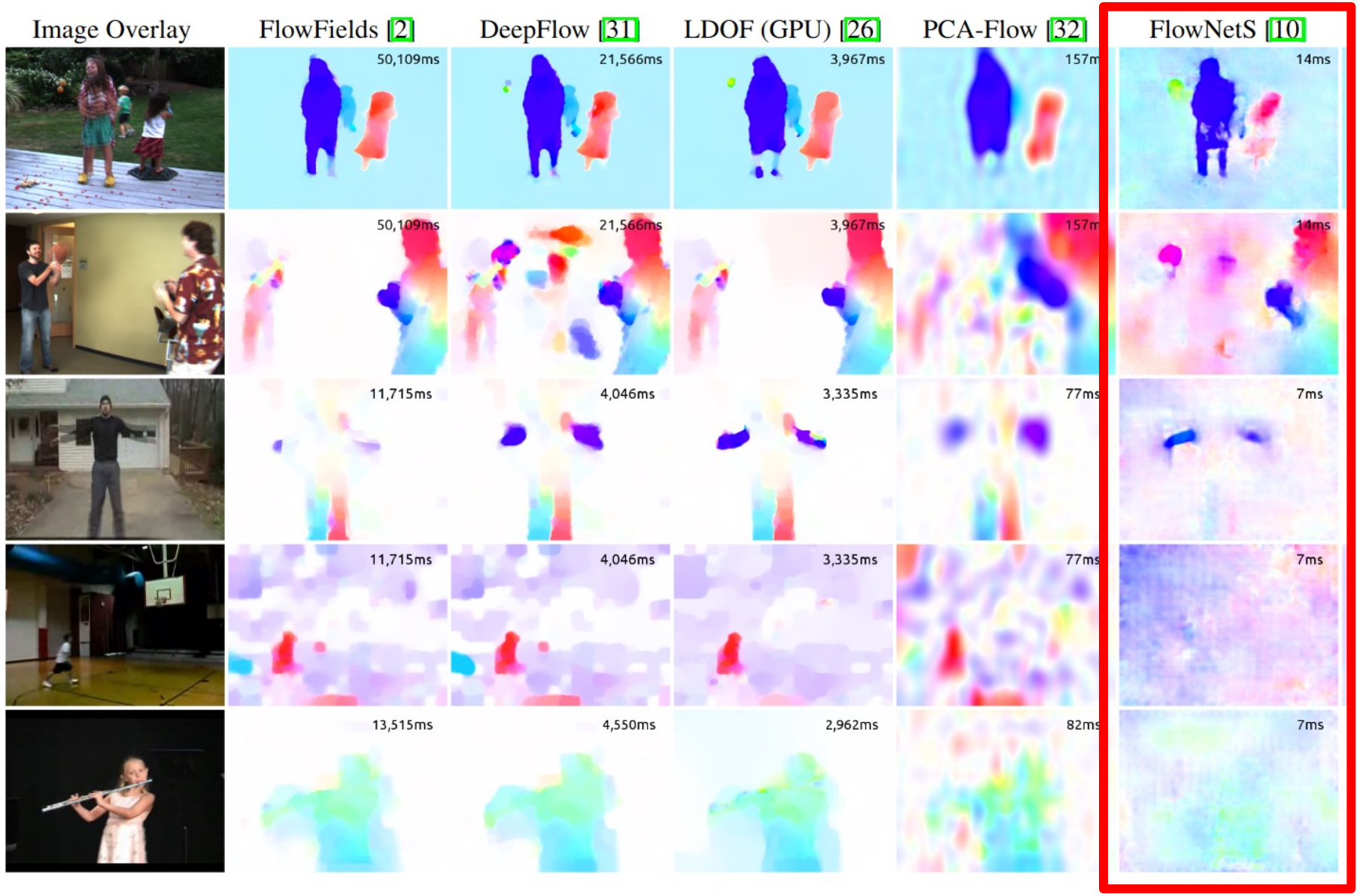
Idea:

- Learn to estimate optical flow end-to-end with a convolutional neural network



[FlowNet: Learning Optical Flow With Convolutional Networks]

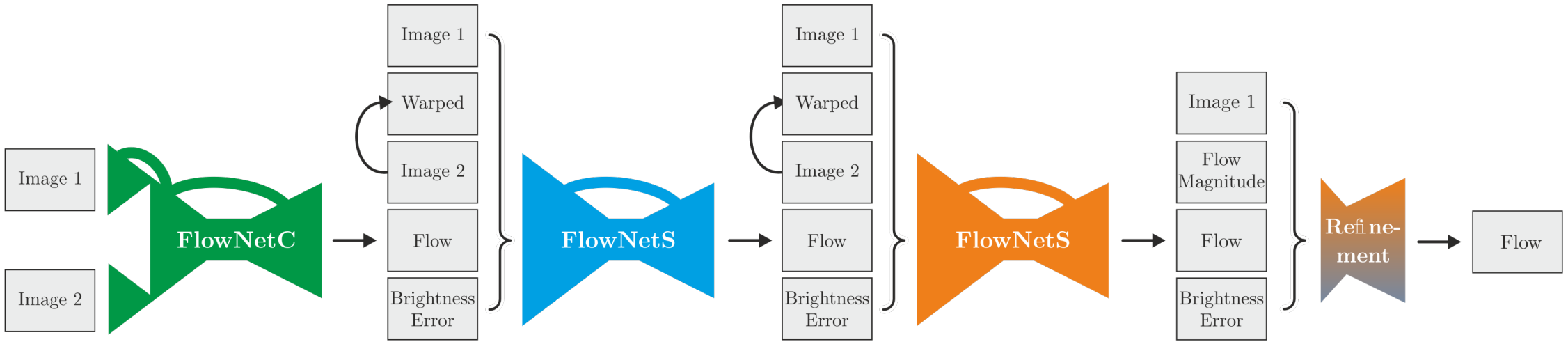
FlowNet 2.0 Results



FlowNet 2.0

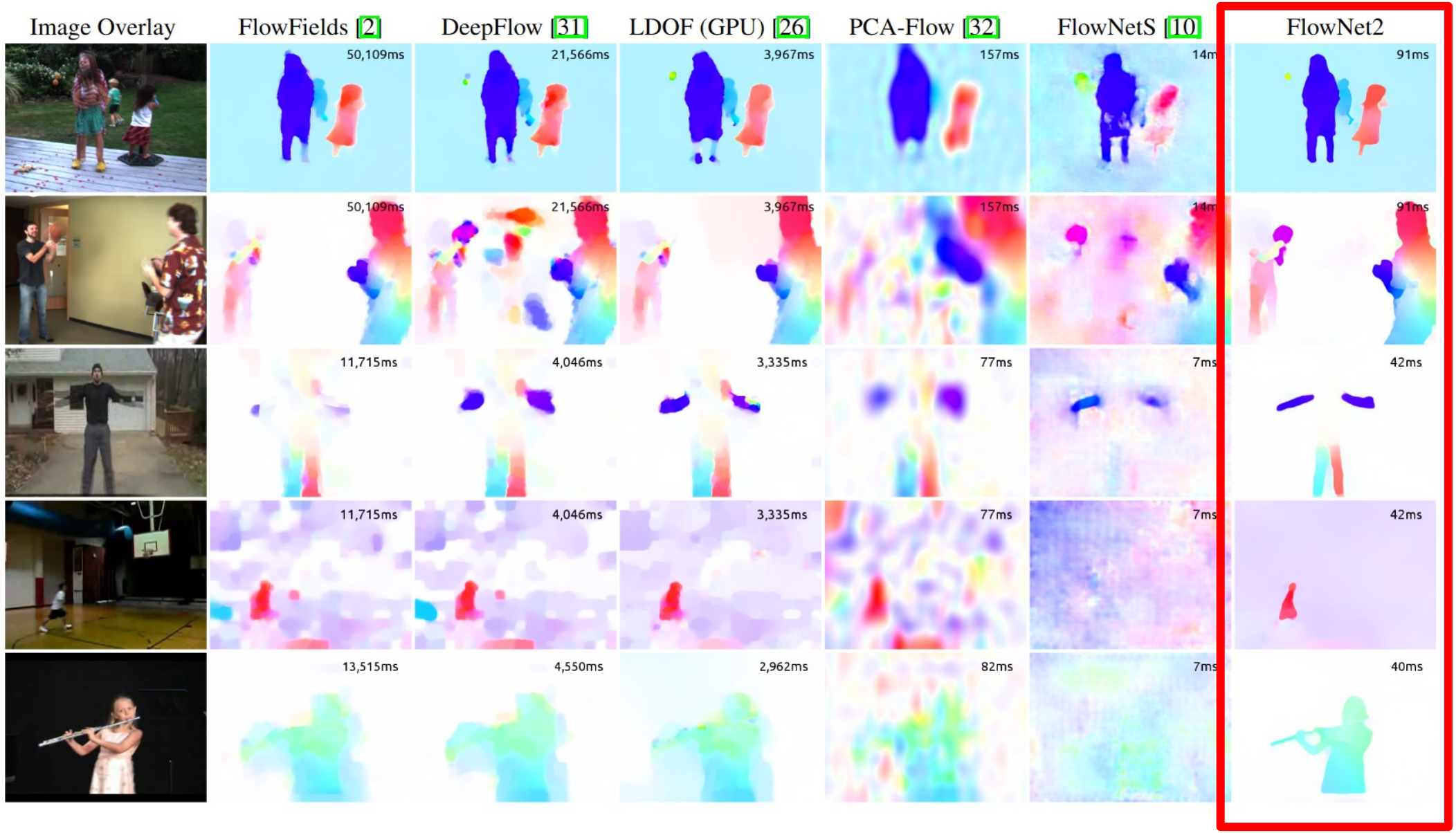
Idea:

- Design a bigger *generic* network
- Refine optical flow in serveral steps



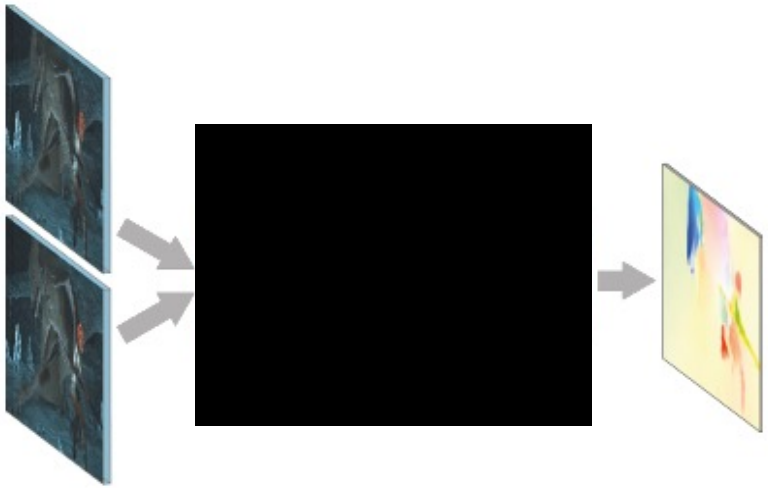
[FlowNet 2.0: Evolution of Optical Flow Estimation with Deep Networks]

FlowNet 2.0 Results

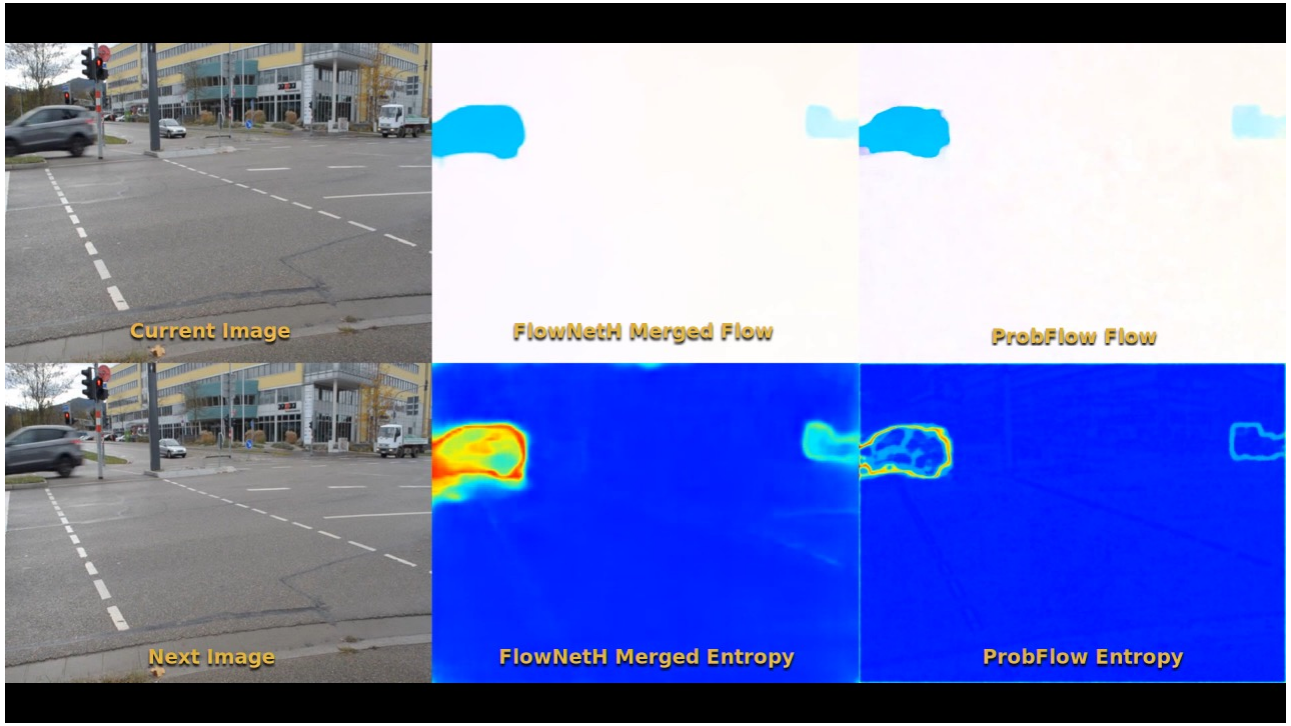


Uncertainty Estimation

Problem: CNN is a black box



Solution: Make CNNs aware of their own uncertainty

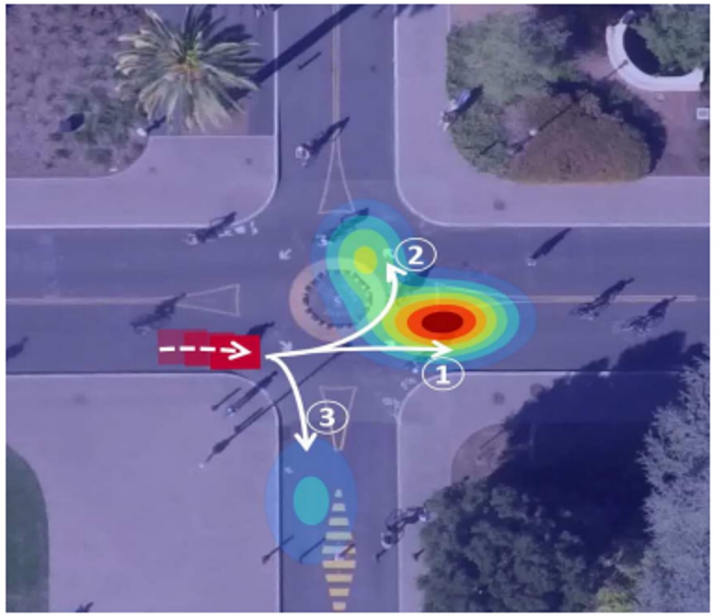


[Uncertainty Estimates and Multi-Hypotheses Networks for Optical Flow]

Extension To Mixture Distributions

Idea:

- Predict multiple hypotesis for the future



[Overcoming Limitations of Mixture Density Networks: A Sampling and Fitting Framework for Multimodal Future Prediction]

Facebook Reality Labs: AR Glasses

- See the real world with overlayed information on top („augmented“ real images)
- Example use-cases:
  - Navigation
  - Memory enhancement
  - Assistance
  - Virtual telepresence and Virtual spaces
- Log term goal: replace smart phone

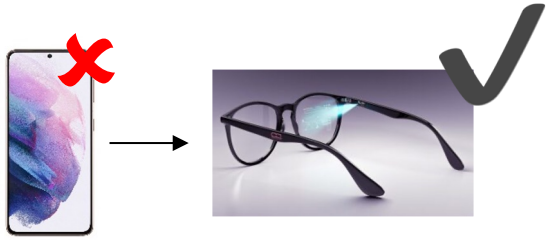
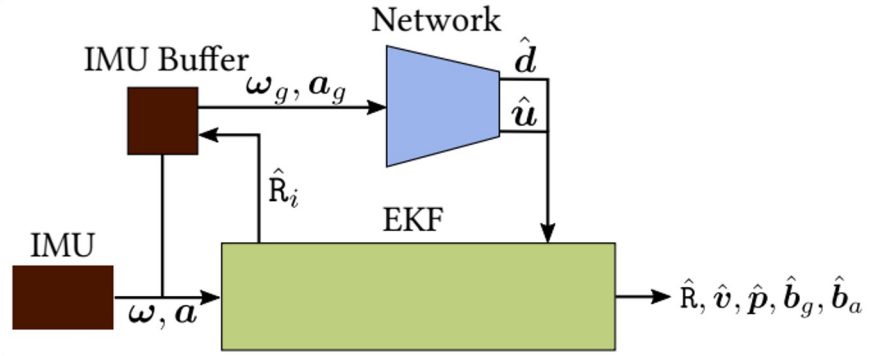
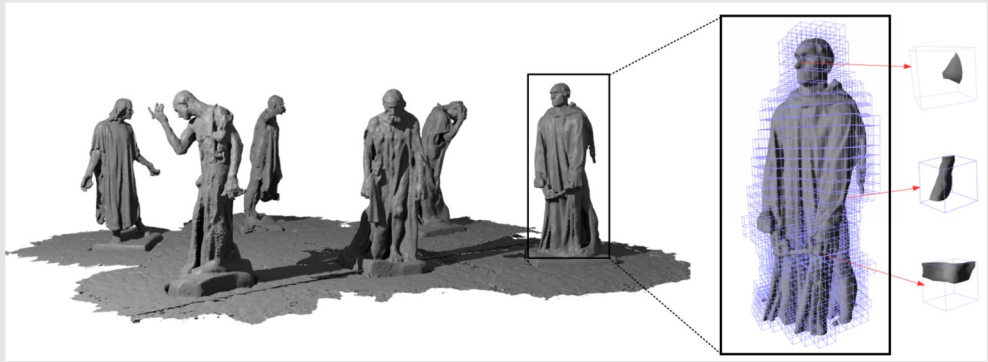


Image sources: www.nextpit.de, tradeinn.com, eyebizz.de

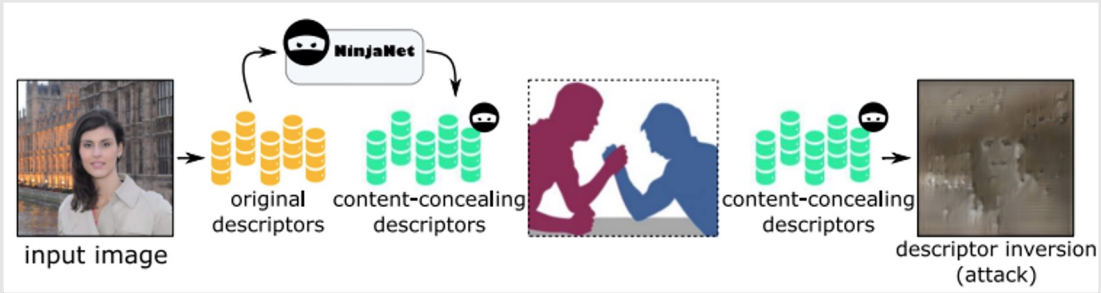
Publications – Facebook Reality Labs



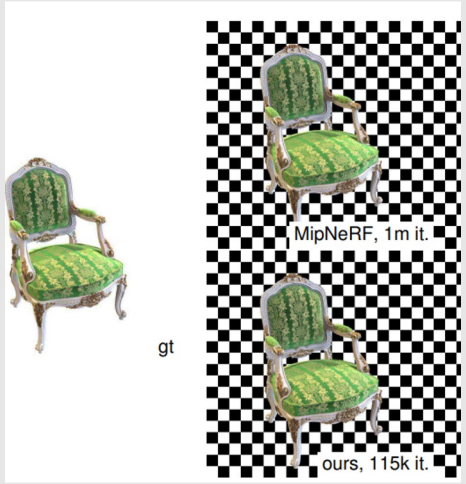
TLIO: Tight Learned Inertial Odometry



Deep Local Shapes: Learning Local SDF Priors for Detailed 3D Reconstruction



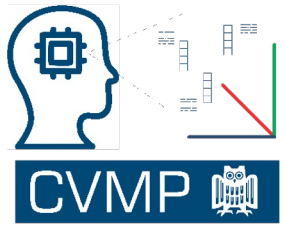
NinjaDesc: Content-Concealing Visual Descriptors via Adversarial Learning



ERF: Explicit Radiance Field Reconstruction From Scratch

Current Work of the CVMP Lab

PhD Positions Available



Computer Vision and Machine Perception Lab at Saarland University

https://cvmp.cs.uni-saarland.de

2D vs. 3D Models

2D:

2D:

2D:

3D:

Novel View Synthesis – Baseline Comparison

Shapenet v2 cars – training set objects

Tatarchenko et al.

Worrall et al.

Deterministic GQN

SRNs

Training on:

- 2434 cars
- 50 observations each

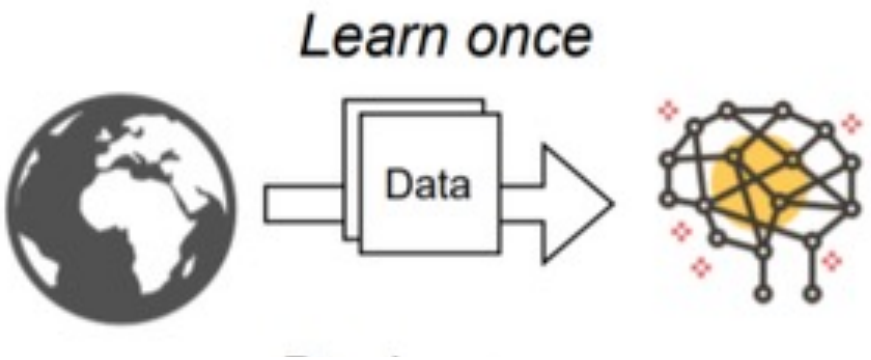
Testing on:

- 2434 cars from training set
- 250 novel views rendered in Archimedean spiral around each object

[Scene Representation Networks, Sitzman, Zollhöfer, Wetzstein]

Continual Learning

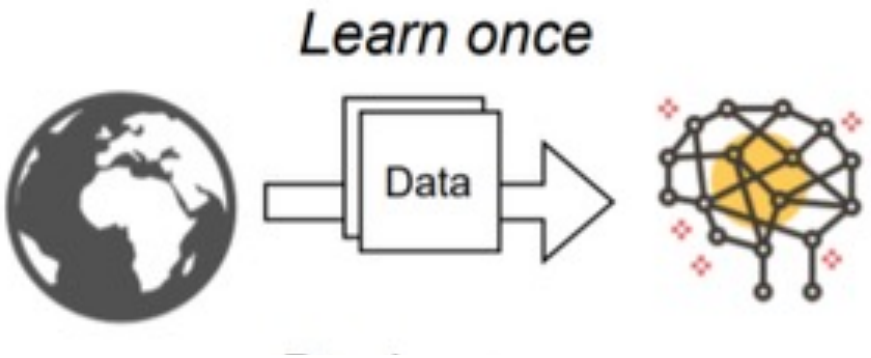
Machine Learning



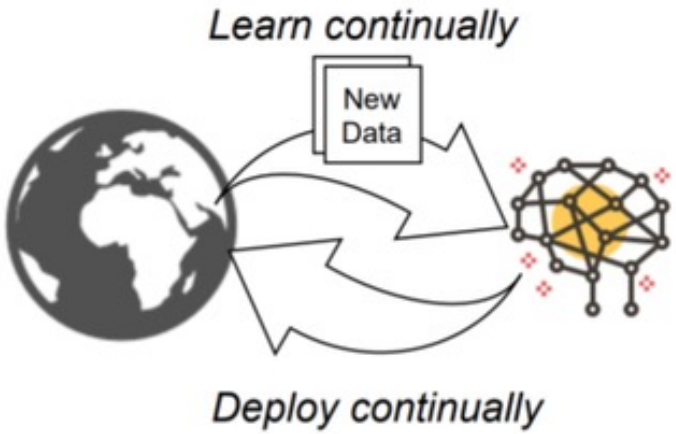
[Source: https://ai.kuleuven.be/stories/post/2021-05-10-continual-learning/]

Continual Learning

Machine Learning

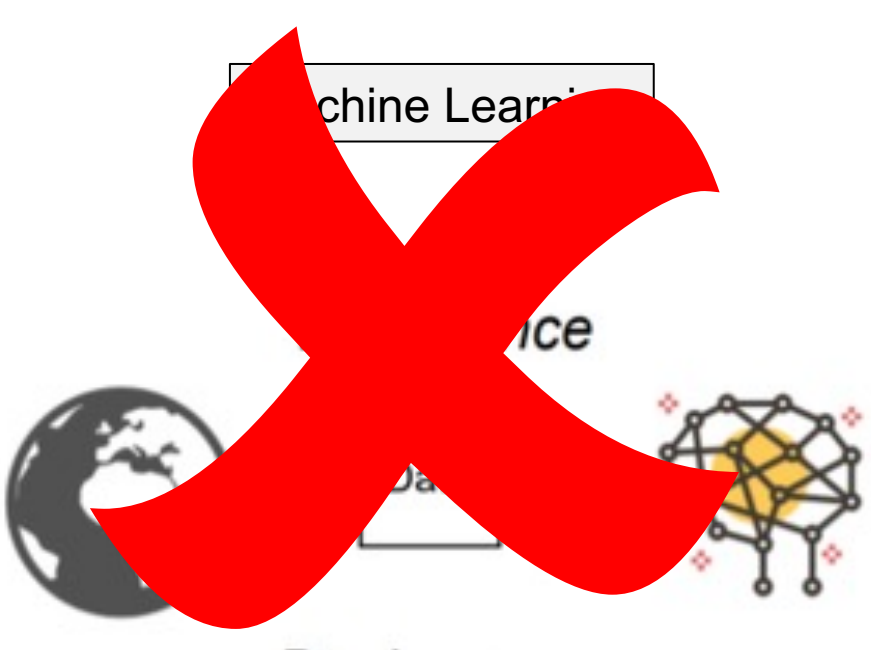


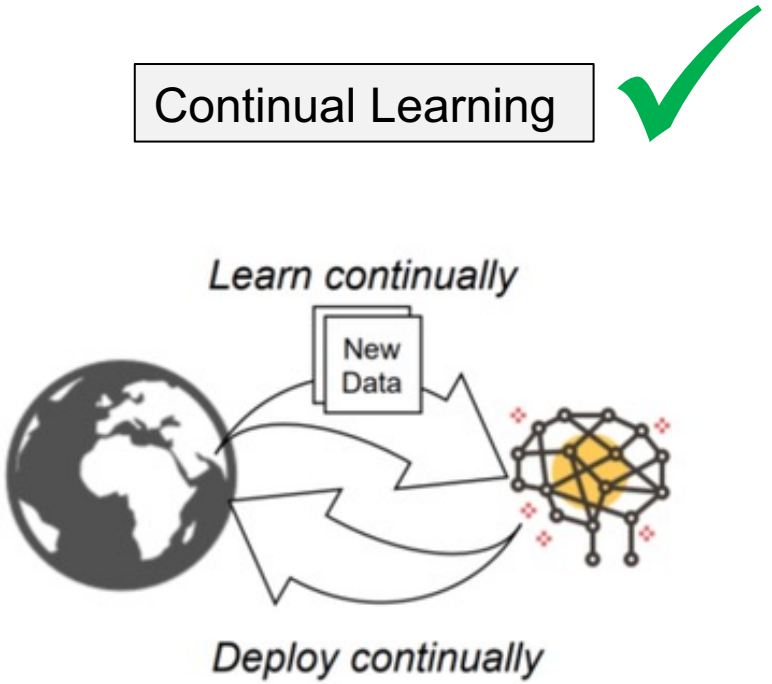
Continual Learning



[Source: https://ai.kuleuven.be/stories/post/2021-05-10-continual-learning/]

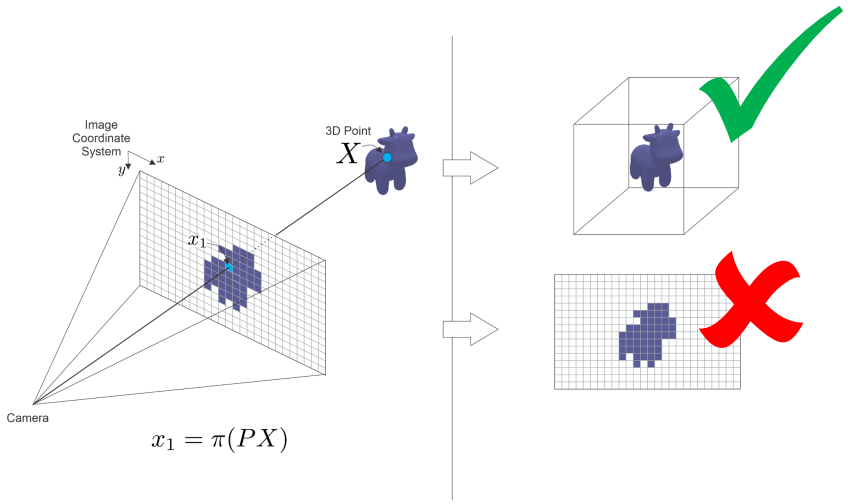
Continual Learning





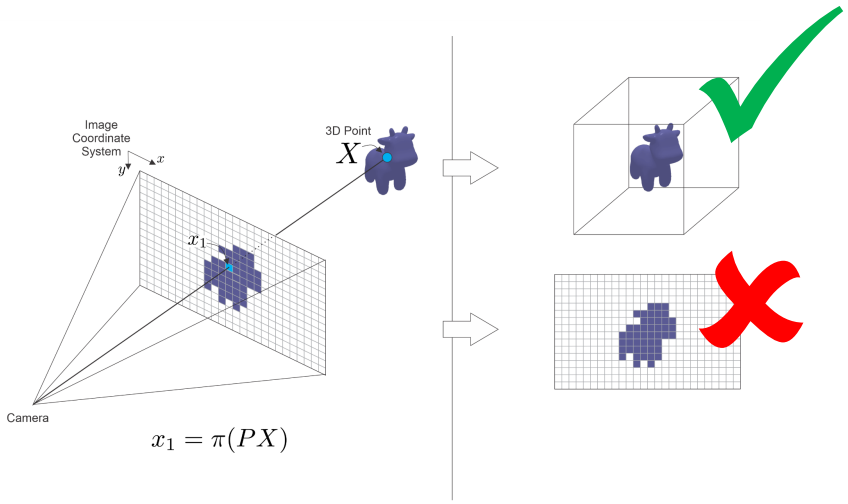
[Source: https://ai.kuleuven.be/stories/post/2021-05-10-continual-learning/]

Key Research Questions

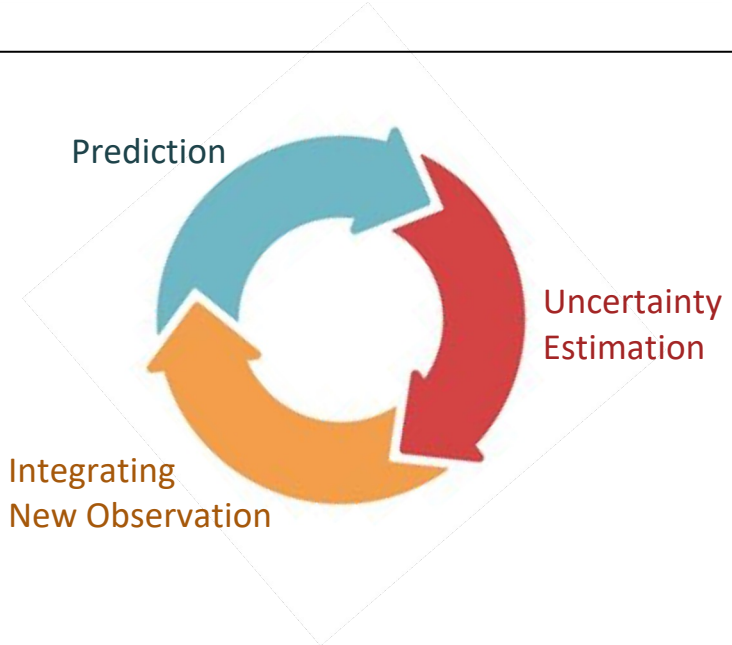


1. How can we build machine perception algorithms that understand our 3D world?

Key Research Questions



1. How can we build machine perception algorithms that understand our 3D world?



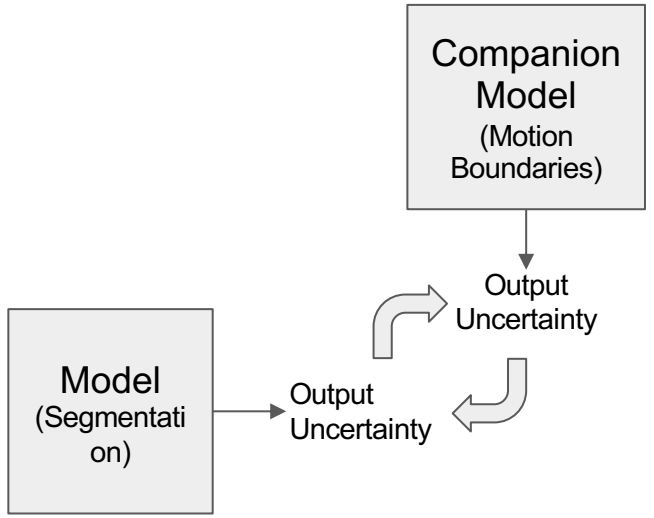
2. How can we build machine perception algorithms that evolve and continually adapt to their environment?

Project 1: Knowledge Transfer

Segmentation from Single Frame

+

Motion Boundaries from Video



[Source: https://neptune.ai/blog/image-segmentation]

Project 2: Object Discovery

Idea:

- Objects can be defined by the Gestalt Principle:
  „Things that move together are an object“

Project 2: Object Discovery

Idea:

- Objects can be defined by the Gestalt Principle:
  „Things that move together are an object“

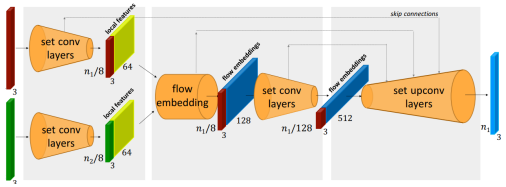
[Conditional Object-Centric Learning from Video]

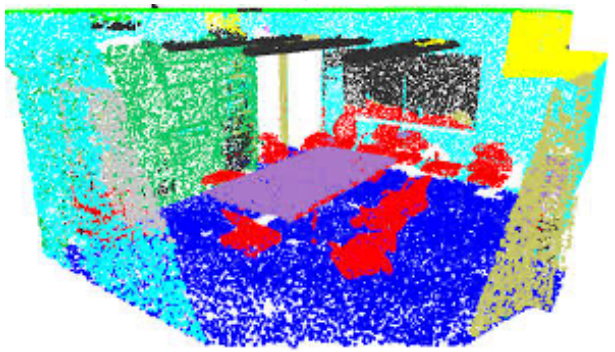
Project 2: Object Discovery

Idea:

- Objects can be defined by the Gestalt Principle:
  „Things that move together are an object“

Our work applies this to 3D Scenes:

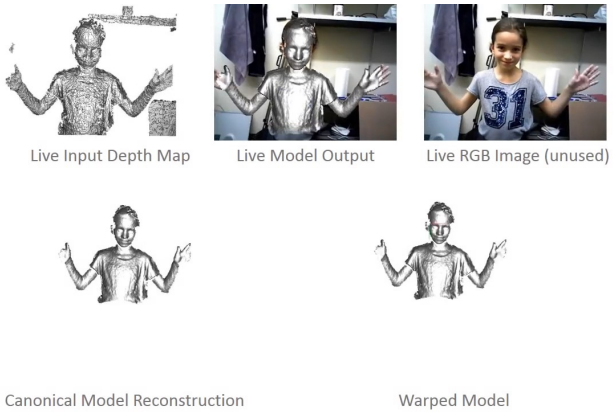




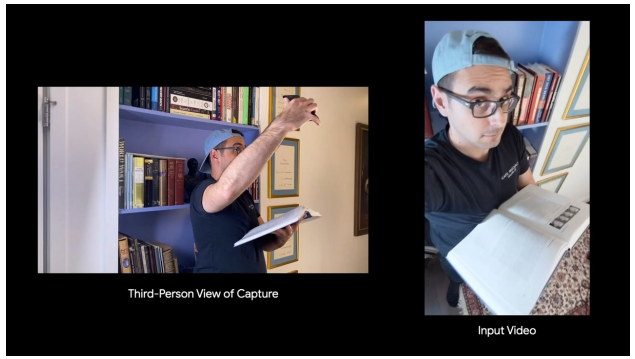
[Conditional Object-Centric Learning from Video]

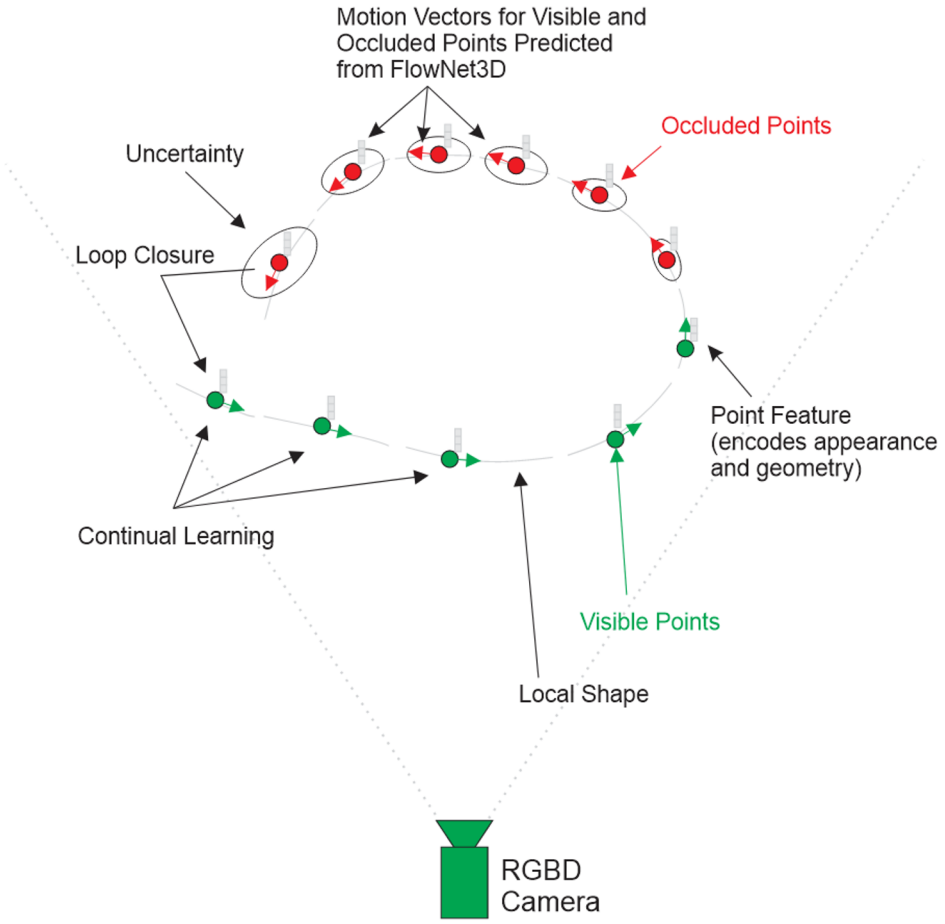
Project 3: Non-Rigid Reconstruction of Whole Scenes

DynamicFusion



Nerfies





[DynamicFusion: Reconstruction and Tracking of Non-rigid Scenes in Real-Time, Nerfies: Deformable Neural Radiance Fields]

Long Term Vision

Building Machine Learning Algorithms that Evolve from New Perceptions

Thank you!

CVMP Lab

LinkedIn: https://www.linkedin.com/in/eddy-ilg/

Website: https://cvmp.cs.uni-saarland.de

Office: Building E1.7, Room 1.05

E-Mail: ilg@cs.uni-saarland.de

PhD Positions Available

Current Projects

- Improving object segmentation from a raw video stream
- Relightable Point-NeRF
- Non-Rigid Reconstruction with Point Clouds
- Non-Rigid Object with Low-Rank Shape Representations
- Estimating Optical Flow with Event Cameras
- Estimating Optical Flow with Quantum Computing

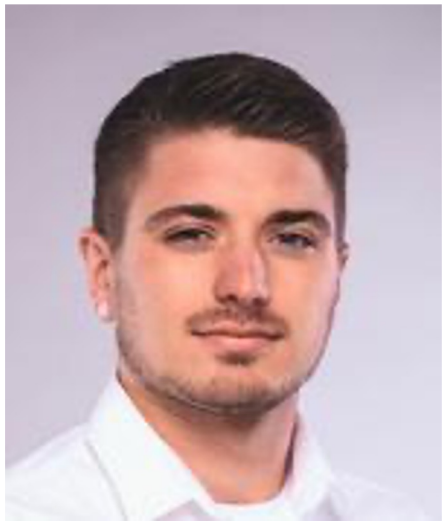
Collaboration Partners

Christian Theobalt

- 3D Computer Vision

Vladislav Golyanik

- Event Cameras
- Quantum Computing



Marc Habermann

- Non-Rigid Reconstruction
- SLAM

Jan-Eric Lenssen

- 3D Reconstruction
- Point Representations

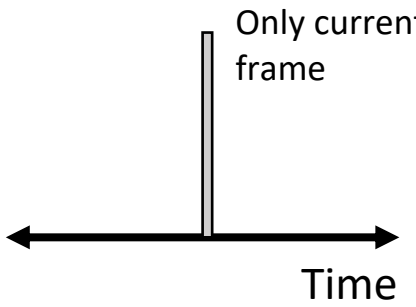
Collaboration with MPI

- **Large-Scale Capture Lab**
- **State-of-The-Art Light Stage**

Lab Resources

- **Workstation with A6000 GPU (48GB)**
- **16 dedicated A100 GPUs / 32 shared A100 GPUs**
- **Nearly unlimited resources available from MPI**

4D Non-Rigid Reconstruction

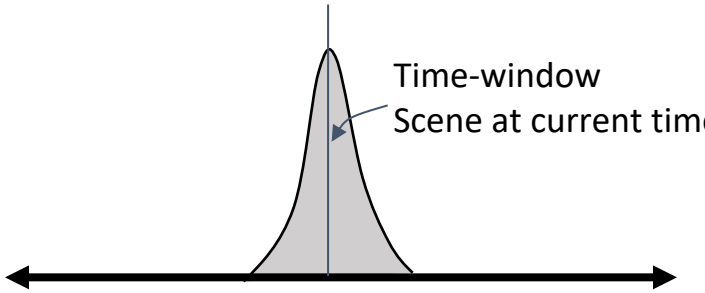


Scene Flow

✓Real time

✗Only visible part of the scene

Does not exist



Tracked Reconstruction

✓Real time

✓Complete Scene

✓Good accuracy / runtime trade-off

Non-Rigid Reconstruction (e.g. NeRFies, Unbiased 4D)

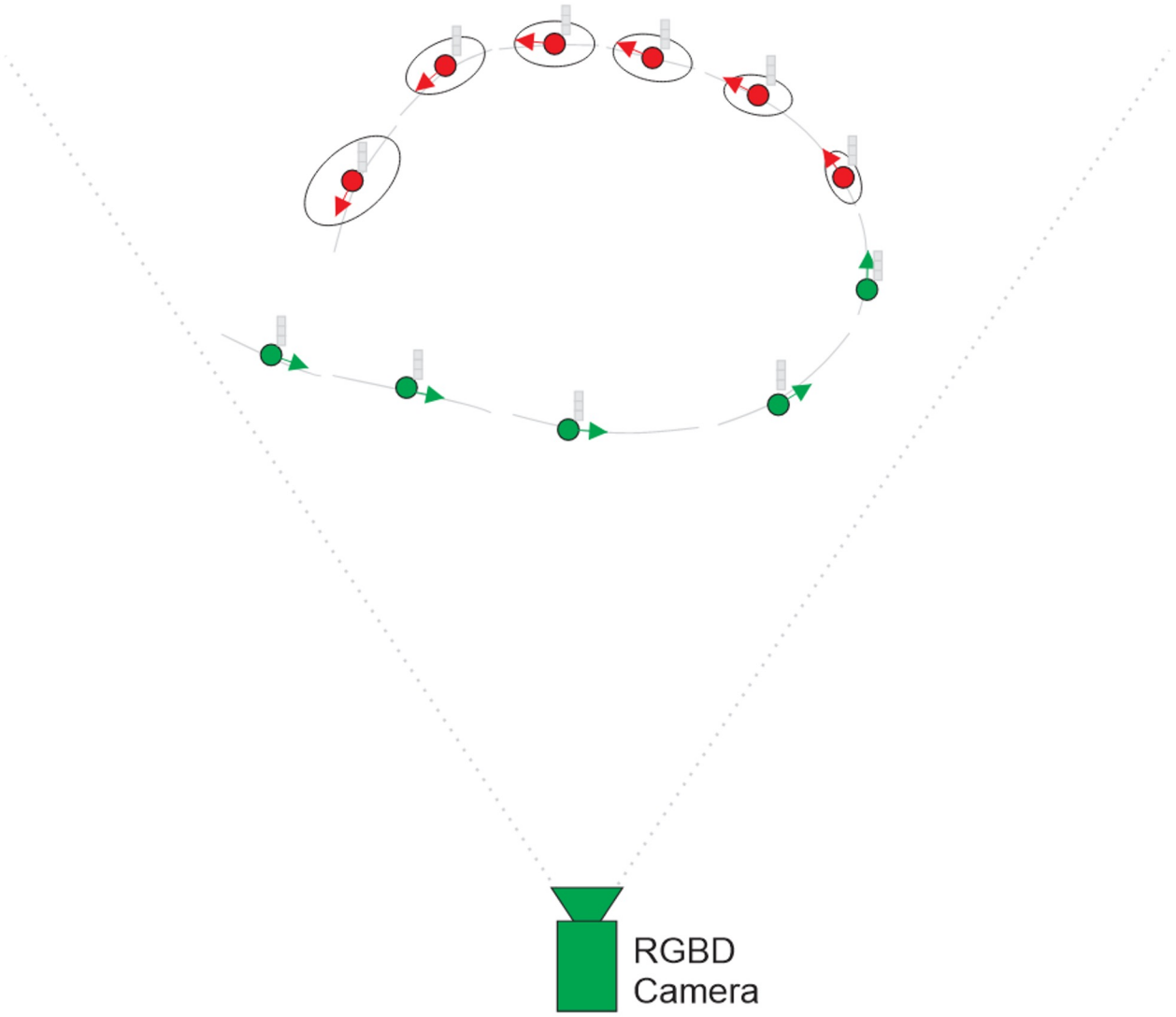
✓Most accurate

✗Post-hoc / slow

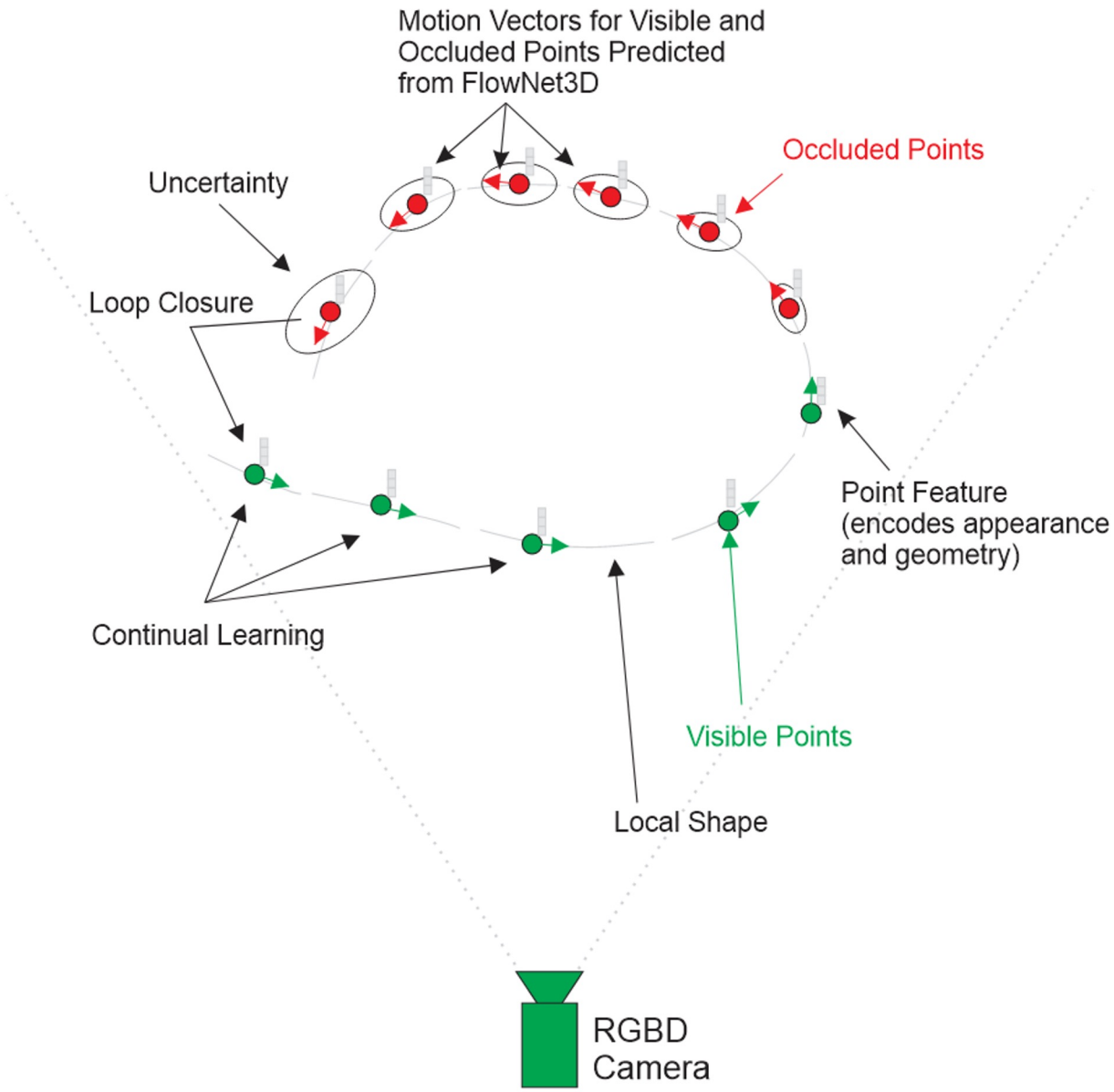
✗Single object

✗No topology changes

4D Non-Rigid Reconstruction



4D Non-Rigid Reconstruction



Potential Publications

ICCV 23 (March)

- Tracking Complete Scenes with Point Clouds from RGBD
- Lightweight Novel View Synthesis of Dynamic Scenes from RGBD

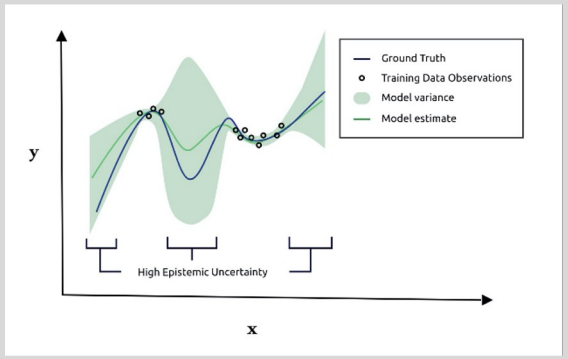
CVPR 24 (November)

- Realtime Reconstruction and Rendering with Local Shapes
- Realtime Non-Rigid Reconstruction of Complete Scenes from RGBD
- Extracting Objects from Non-Rigid Scenes with SLOT Attention from RGBD

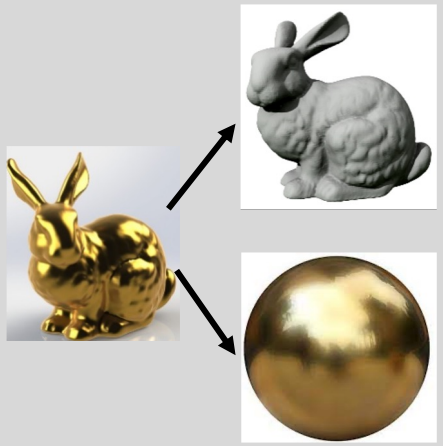
ECCV 24 (March)

- Realtime Non-Rigid Reconstruction of Complete Scenes from Images
- Continual Learning of Deformation Models in Non-Rigid Scenes
- Physically Realistic and Uncertainty-Aware Completion for Non-Rigid Reconstruction of Complete Scenes

Topics the lab works on

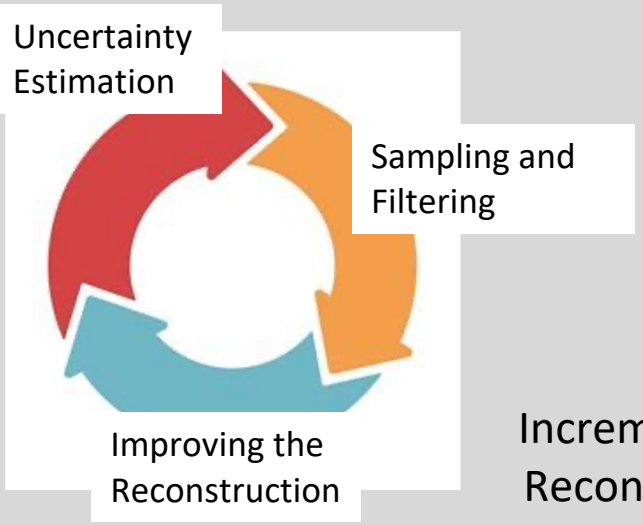


Uncertainty Estimation for 3D Reconstruction

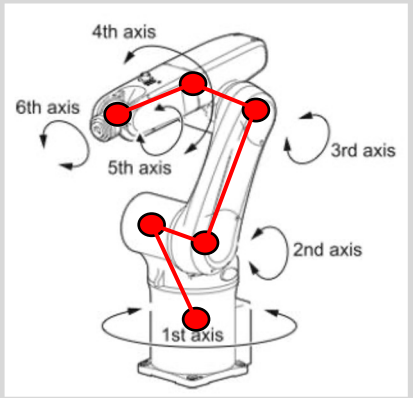


Material / Geometry Decomposition for Object Reconstruction in the Wild

Lightweight Realtime Non-Rigid RGB-D Reconstruction



Incremental 3D Reconstruction



Learning Object Skeletons from RGB-D