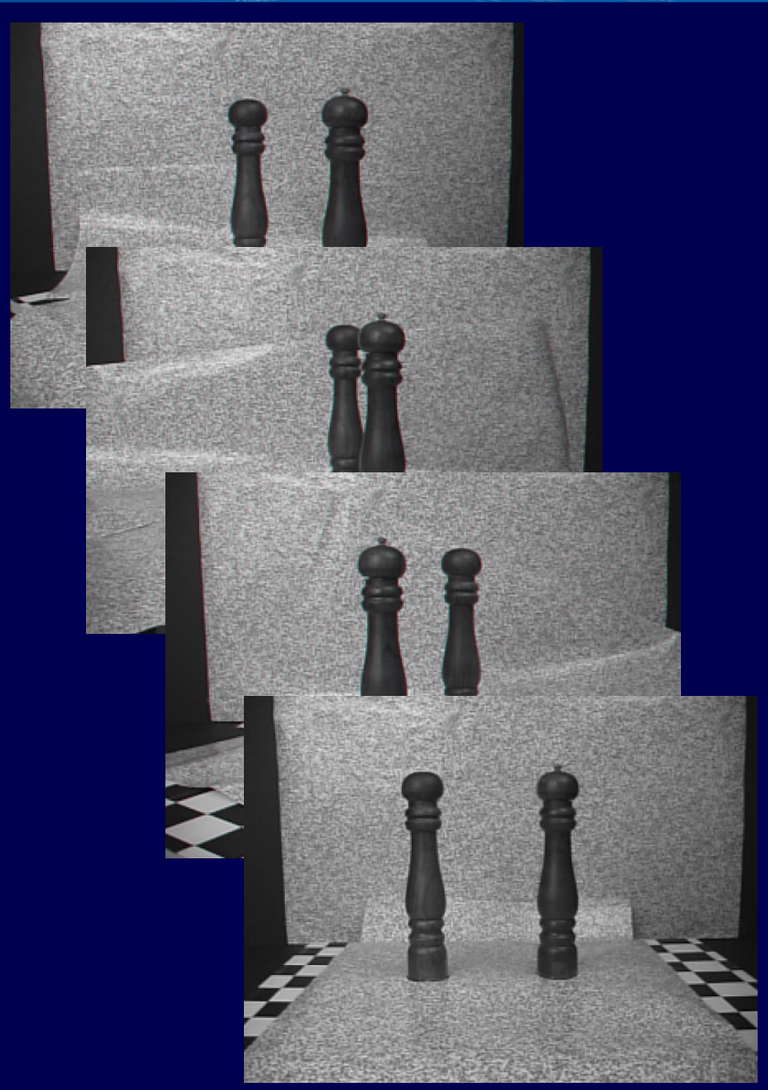
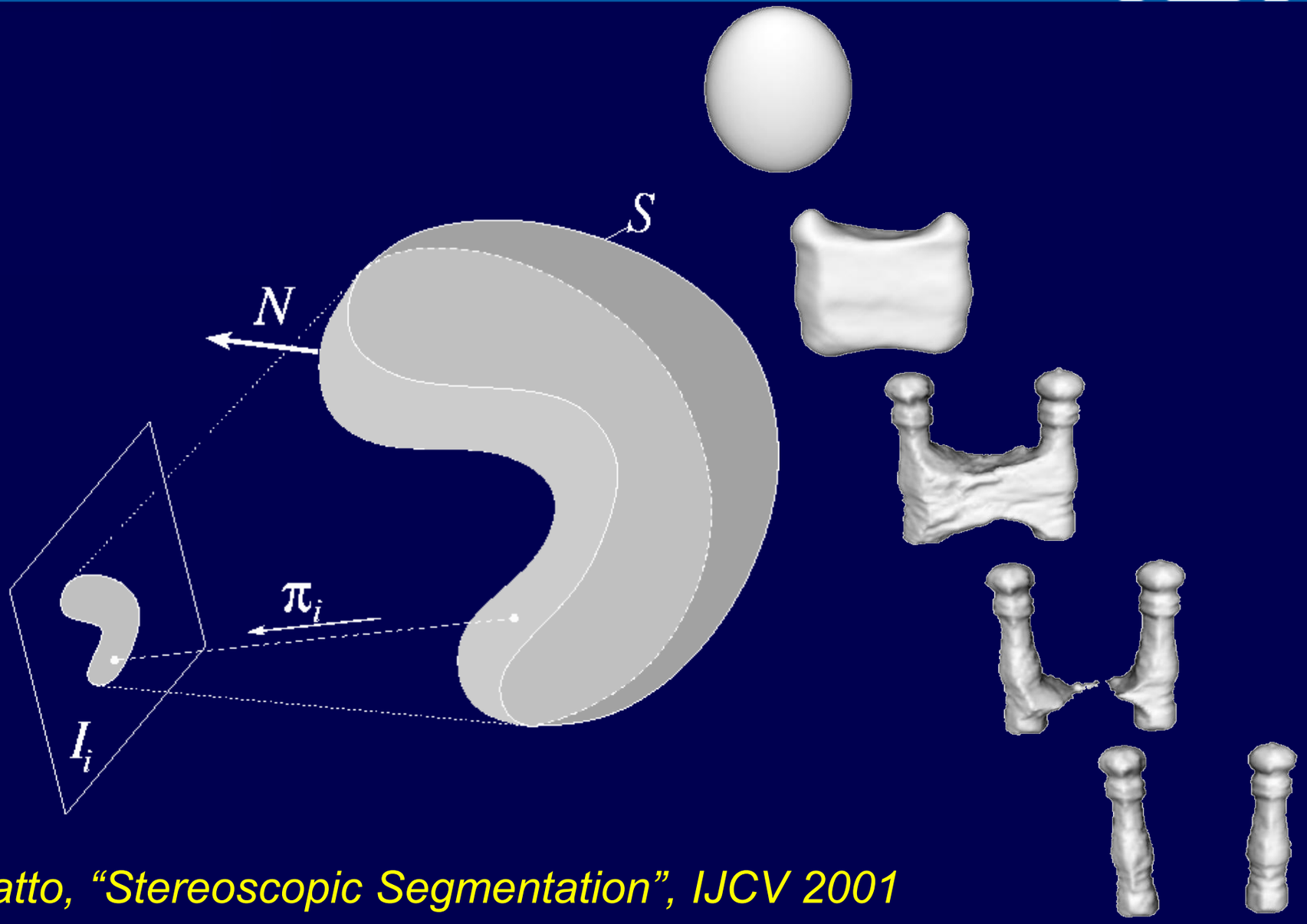
Classical and Learning-based Approaches to 4D Reconstruction

Daniel Cremers
Chair of Computer Vision and AI, TU Munich
Munich Center for Machine Learning

Reconstructing 3D Shape

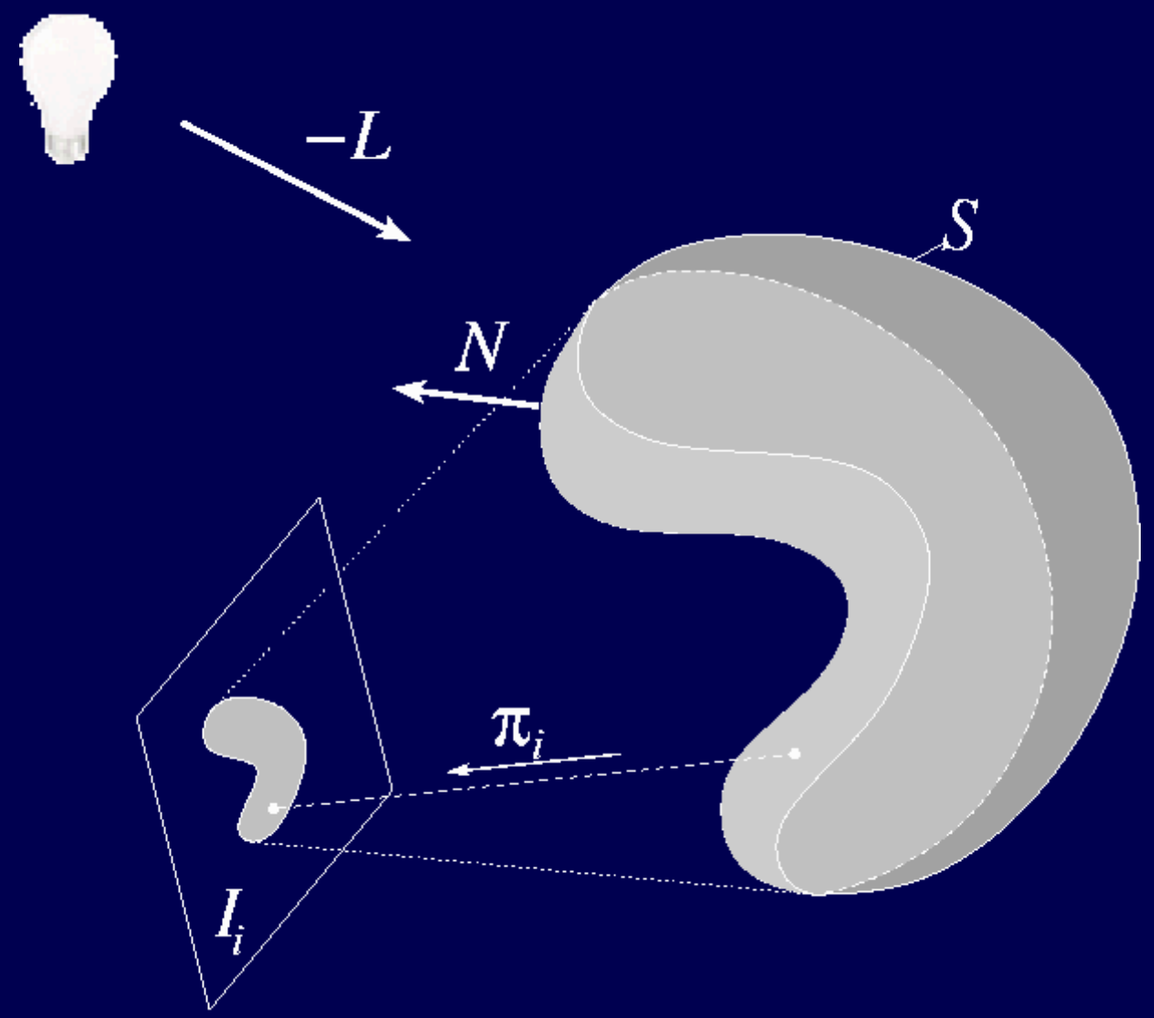
Reconstructing 3D Shape





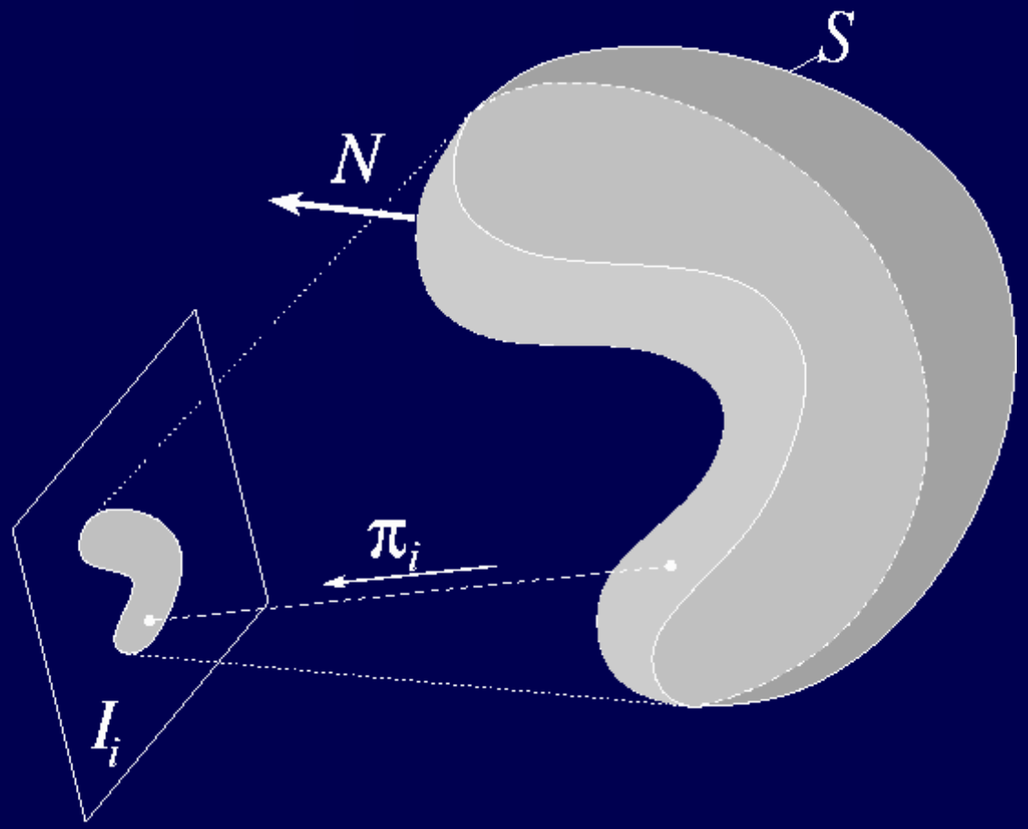
Yezzi, Soatto, "Stereoscopic Segmentation", IJCV 2001

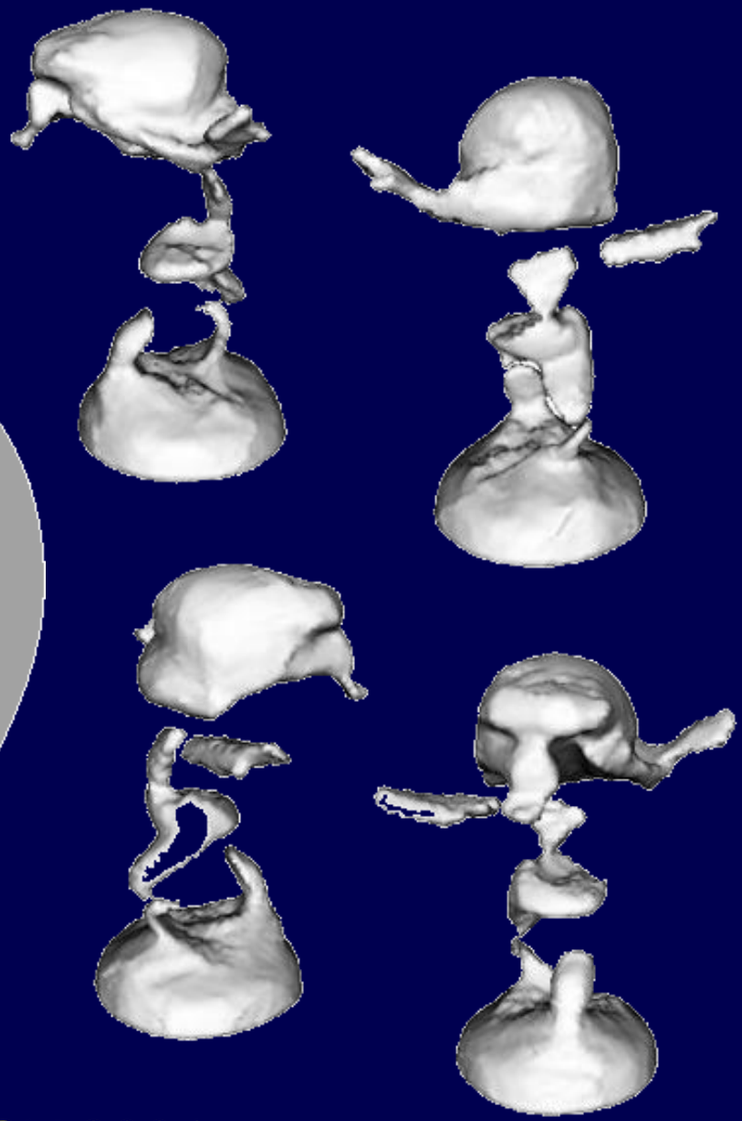
Reconstructing 3D Shape and Lighting



Jin, Cremers, Yezzi, Soatto, "Shedding Light on Stereoscopic Segmentation", CVPR 2004

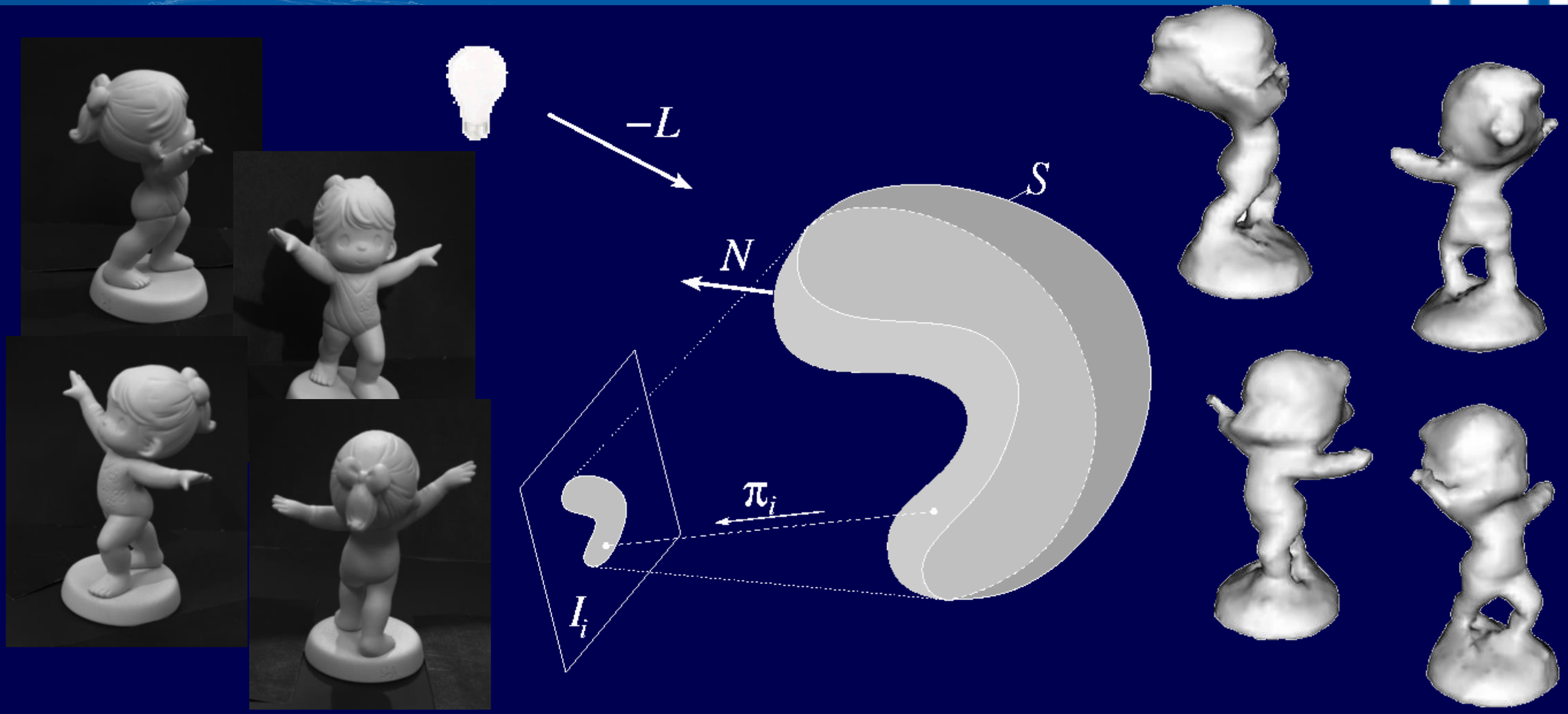
Reconstructing 3D Shape





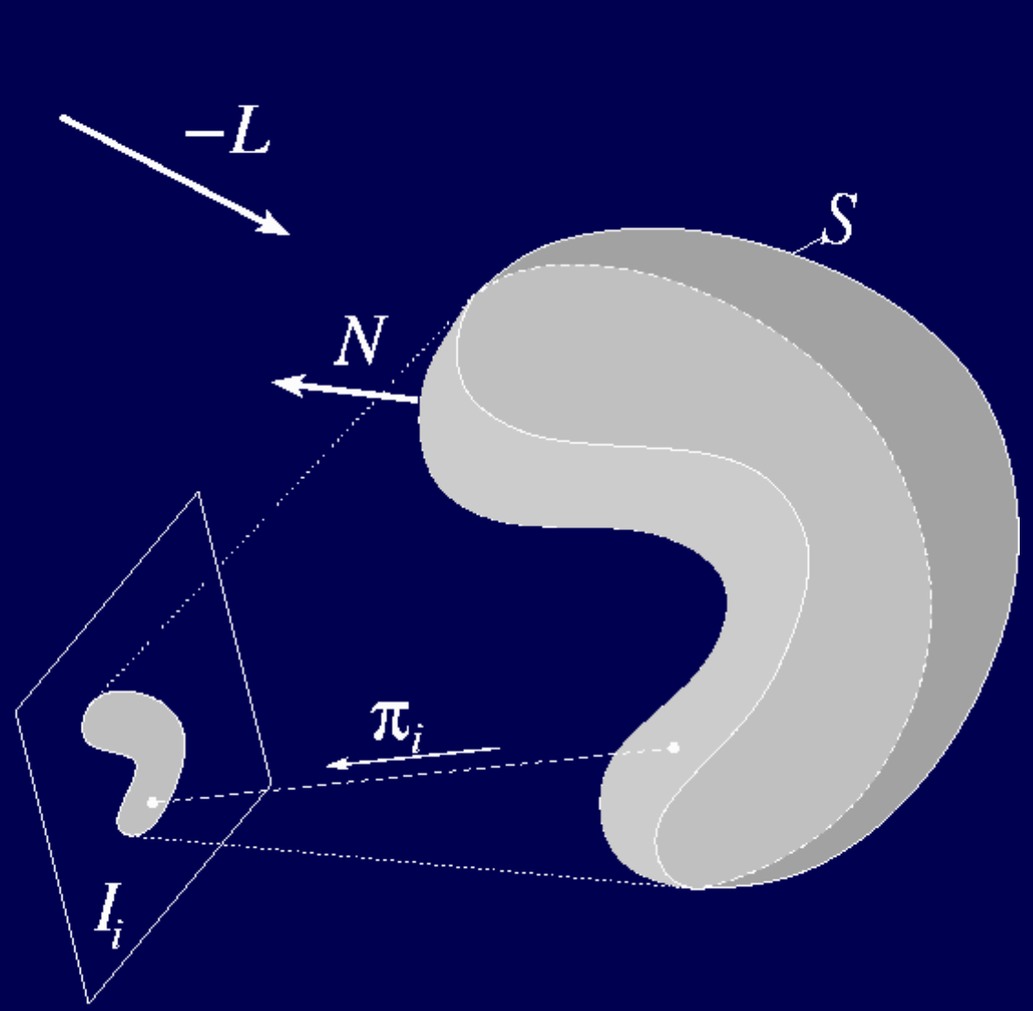
Yezzi, Soatto, "Stereoscopic Segmentation", IJCV 2001

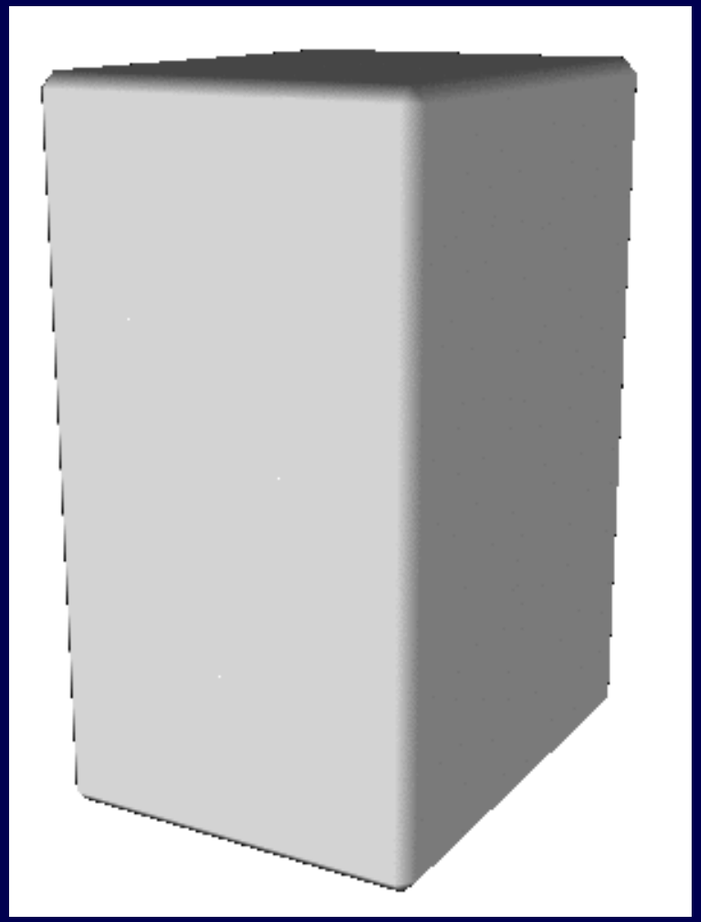
Reconstructing 3D Shape and Lighting



Jin, Cremers, Yezzi, Soatto, "Shedding Light on Stereoscopic Segmentation", CVPR 2004

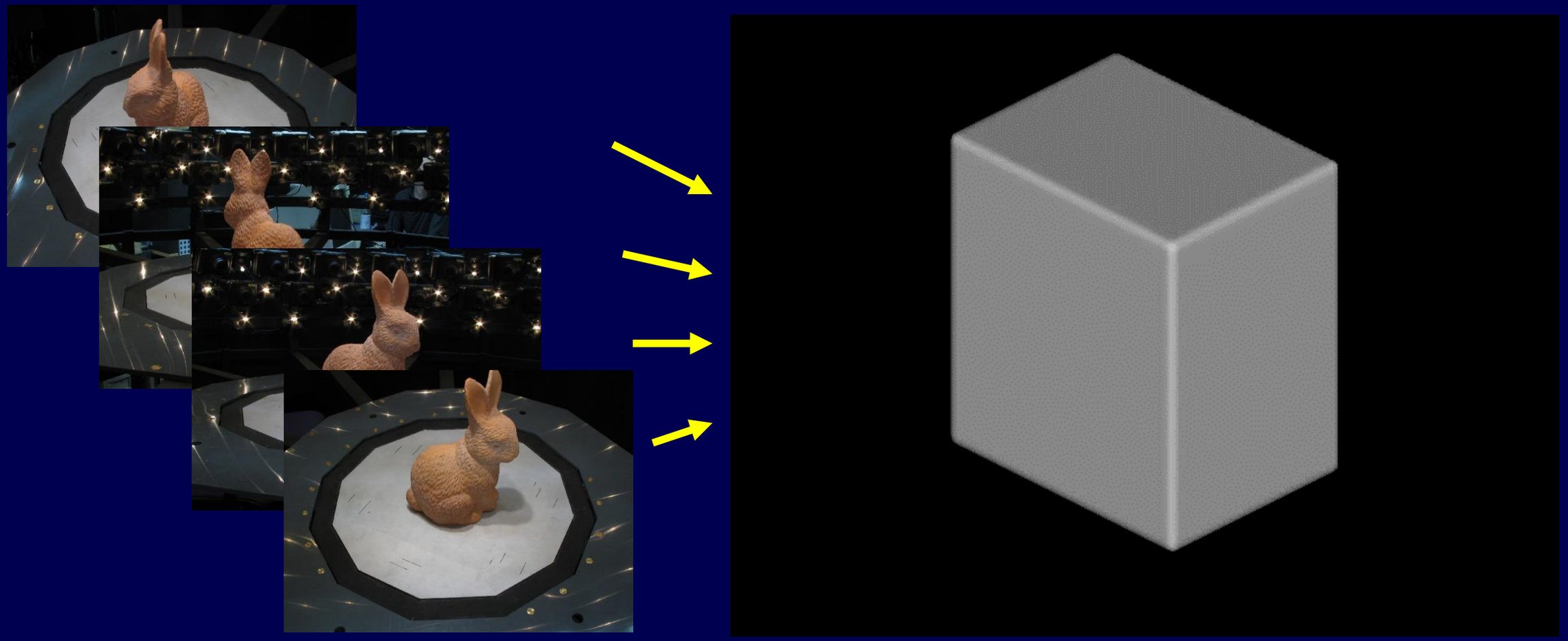
Reconstructing 3D Shape and Lighting





Jin, Cremers, Yezzi, Soatto, "Shedding Light on Stereoscopic Segmentation", CVPR 2004

Provably Optimal Reconstructions



Kolev, Klodt, Brox, Cremers, Int. J. of Computer Vision ’09:

Theorem: Optimal reconstructions can be computed via convex relaxation.

Super-Resolution Texture

*Goldlücke, Cremers, ICCV ’09, DAGM ’09**

** Best Paper Award*

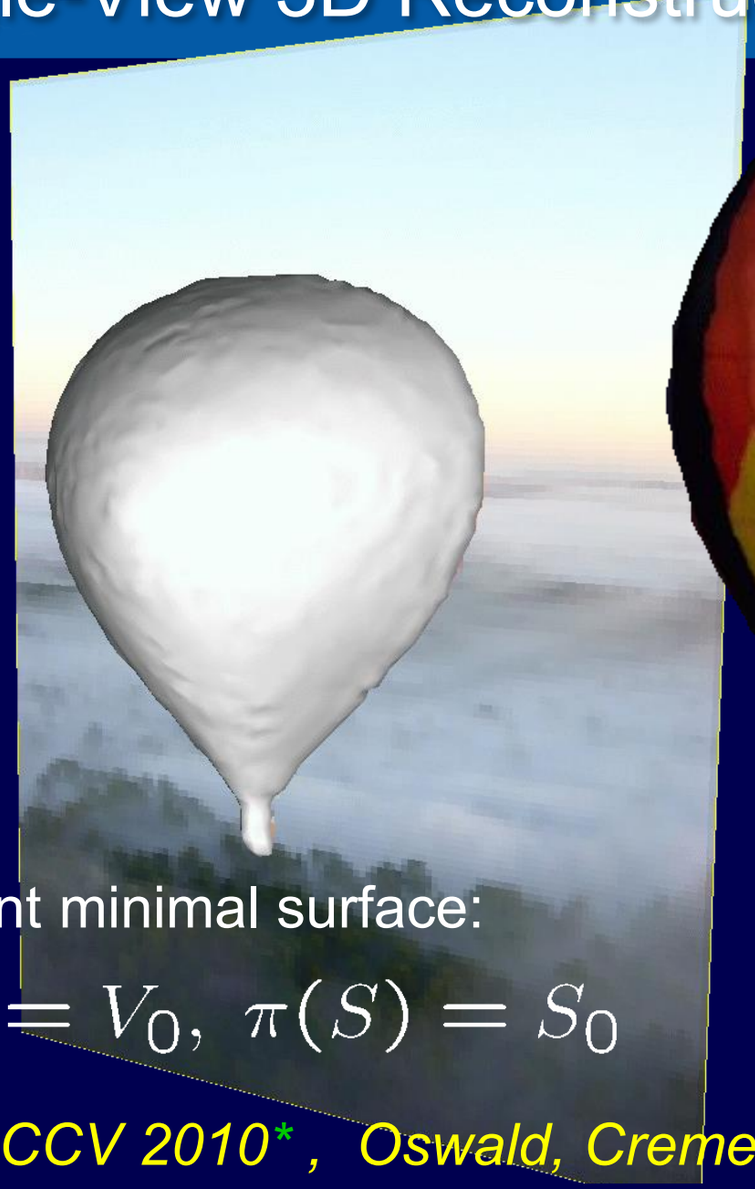
Super-Resolution Texture

Closeup of input image

Super-resolution texture

*Goldlücke, Cremers, ICCV '09, DAGM '09** ** Best Paper Award*

Single-View 3D Reconstruction



Fixed-volume silhouette-consistent minimal surface:

$$\min_S |S| \quad \text{s.t. } \text{Vol}(S) = V_0,\ \pi(S) = S_0$$

Toeppe et al., ACCV 2010 , Oswald, Cremers, CVPR 2012*

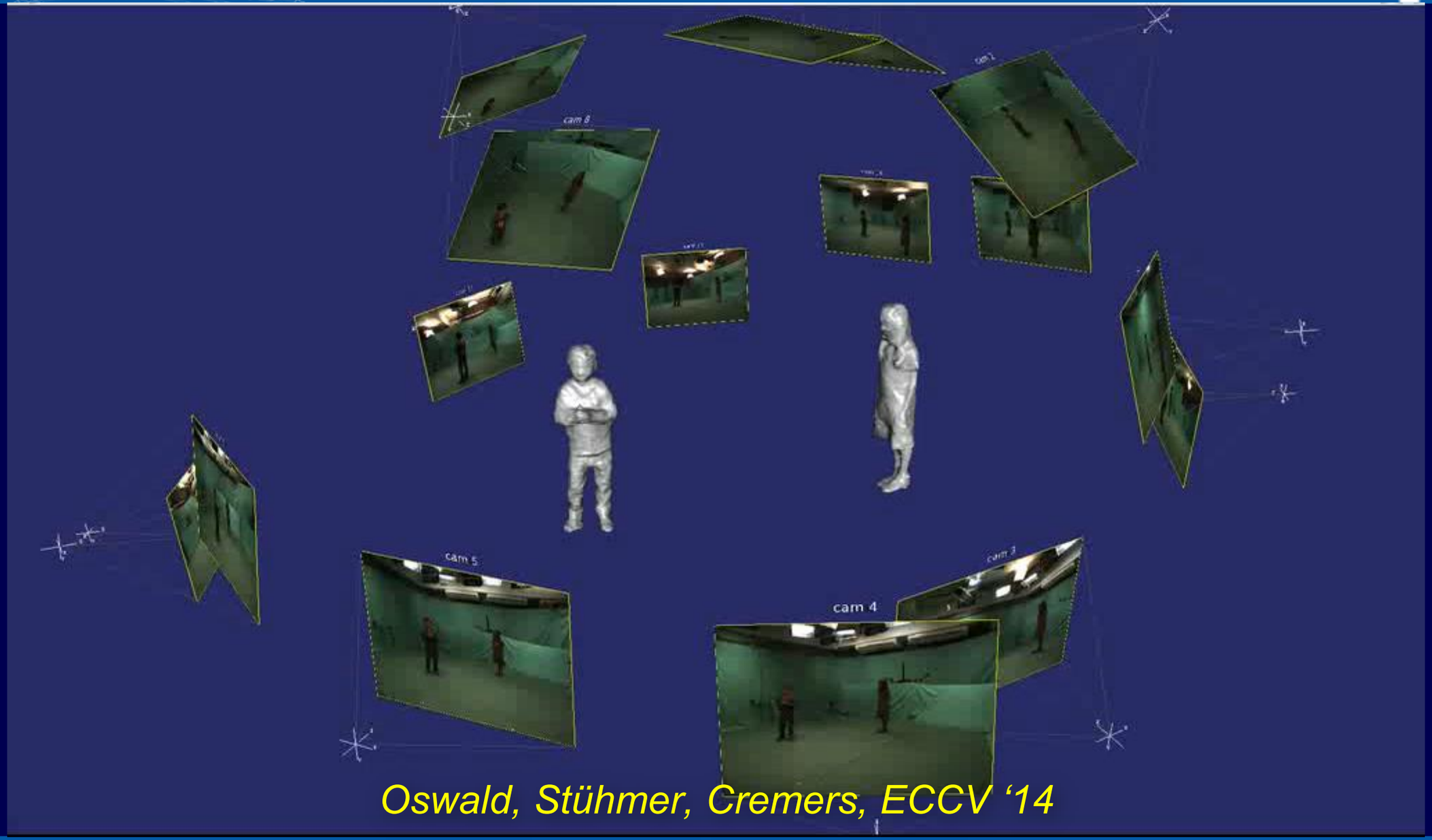
** Best Paper Honorable Mention*

Single-View 3D Reconstruction

Toeppe et al., ACCV 2010 , Oswald, Cremers, CVPR 2012*

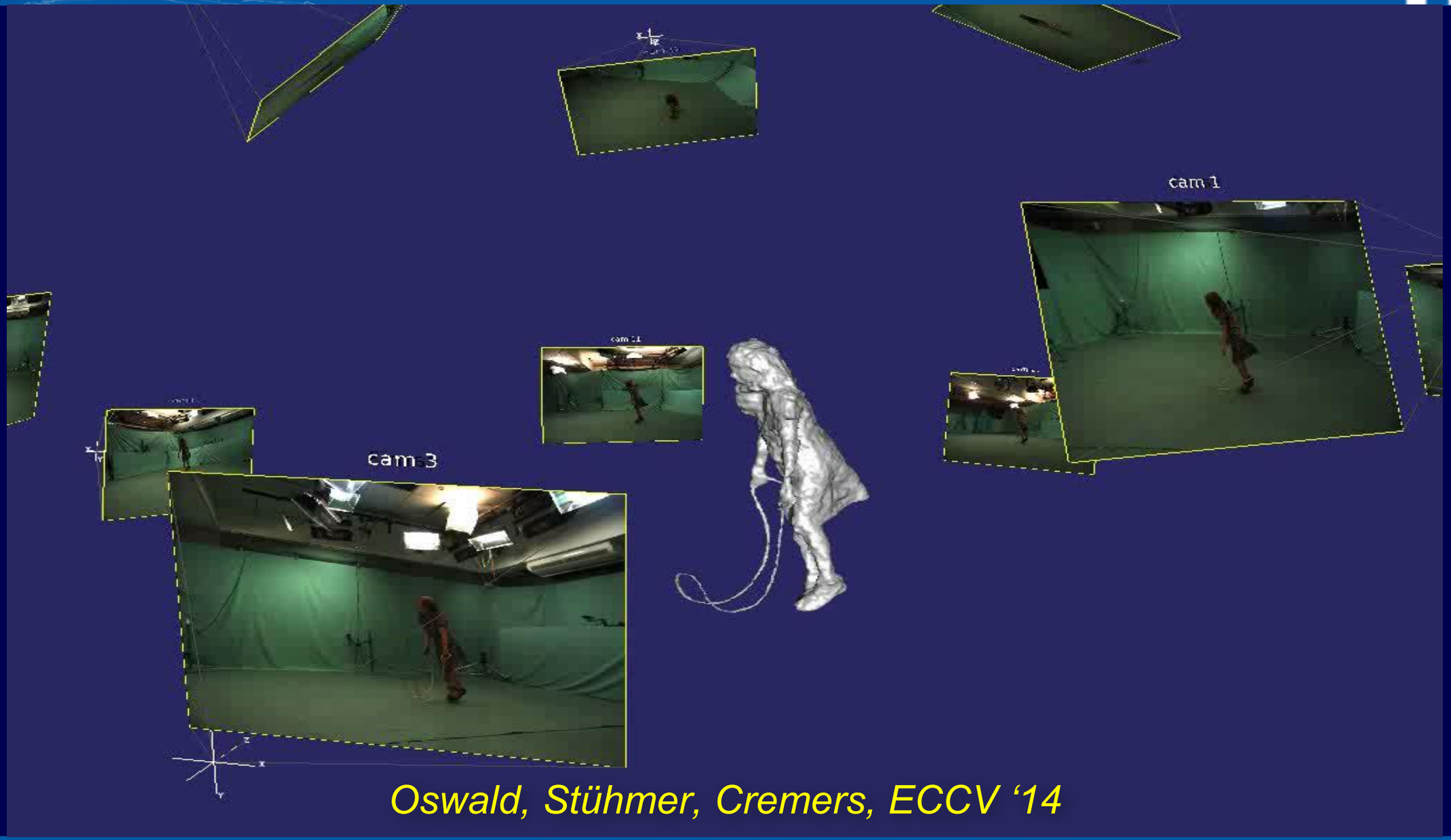
** Best Paper Honorable Mention*

4D Reconstruction from Multiview Video



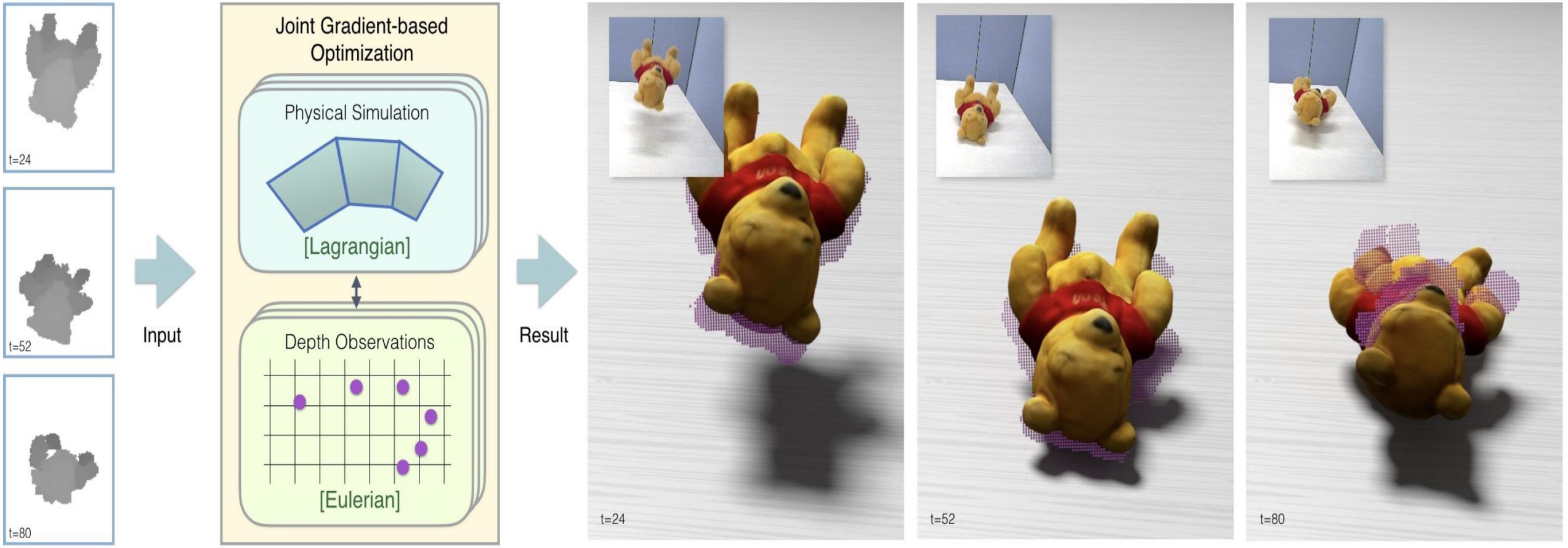
Oswald, Stühmer, Cremers, ECCV '14

4D Reconstruction from Multiview Video



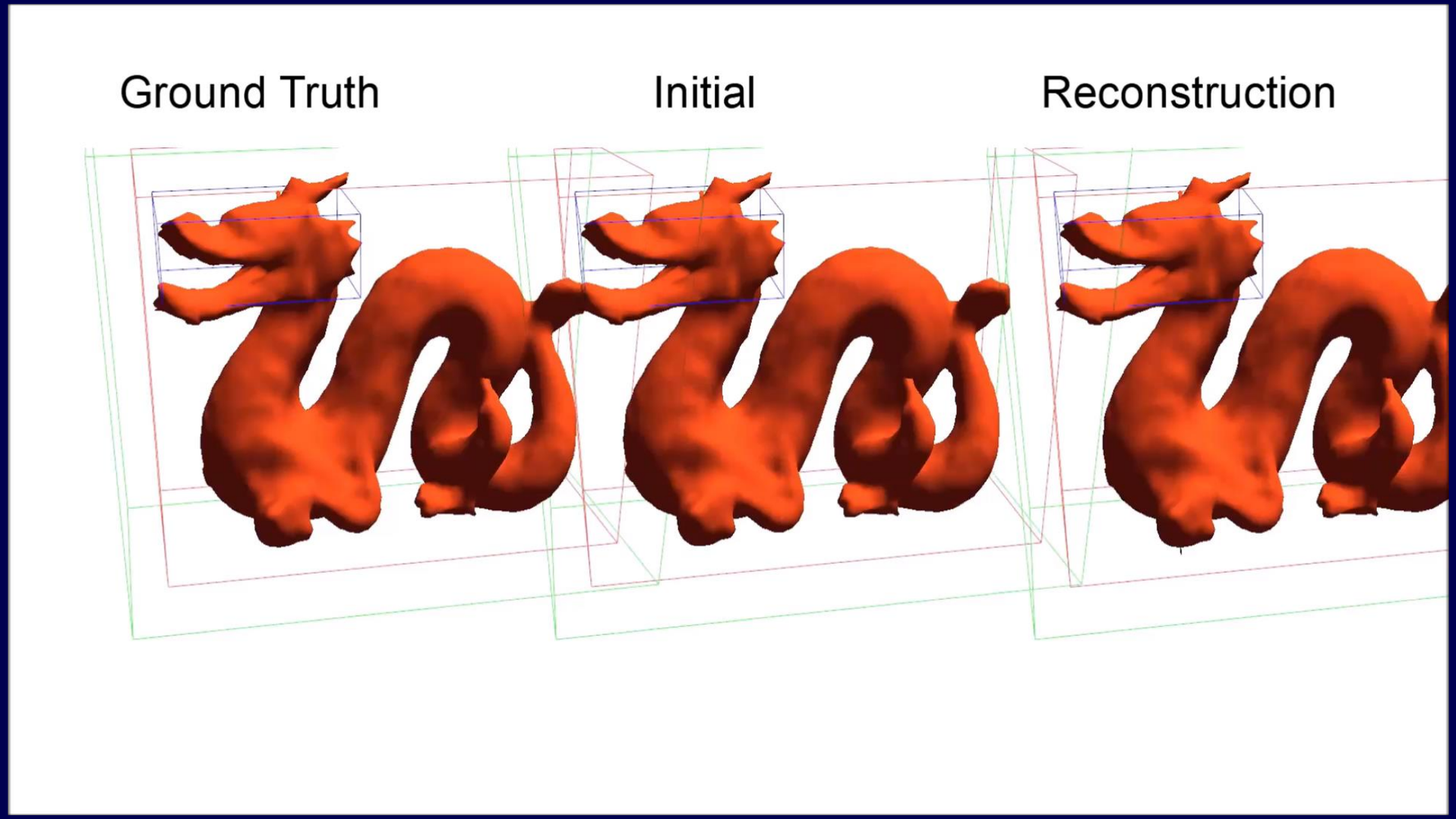
Oswald, Stühmer, Cremers, ECCV ‘14

Reconstructing Physical Simulations from Video



Weiss et al., CVPR 2020

Reconstructing Physical Simulations from Video



Weiss et al., CVPR 2020

Reconstructing Physical Simulations from Video

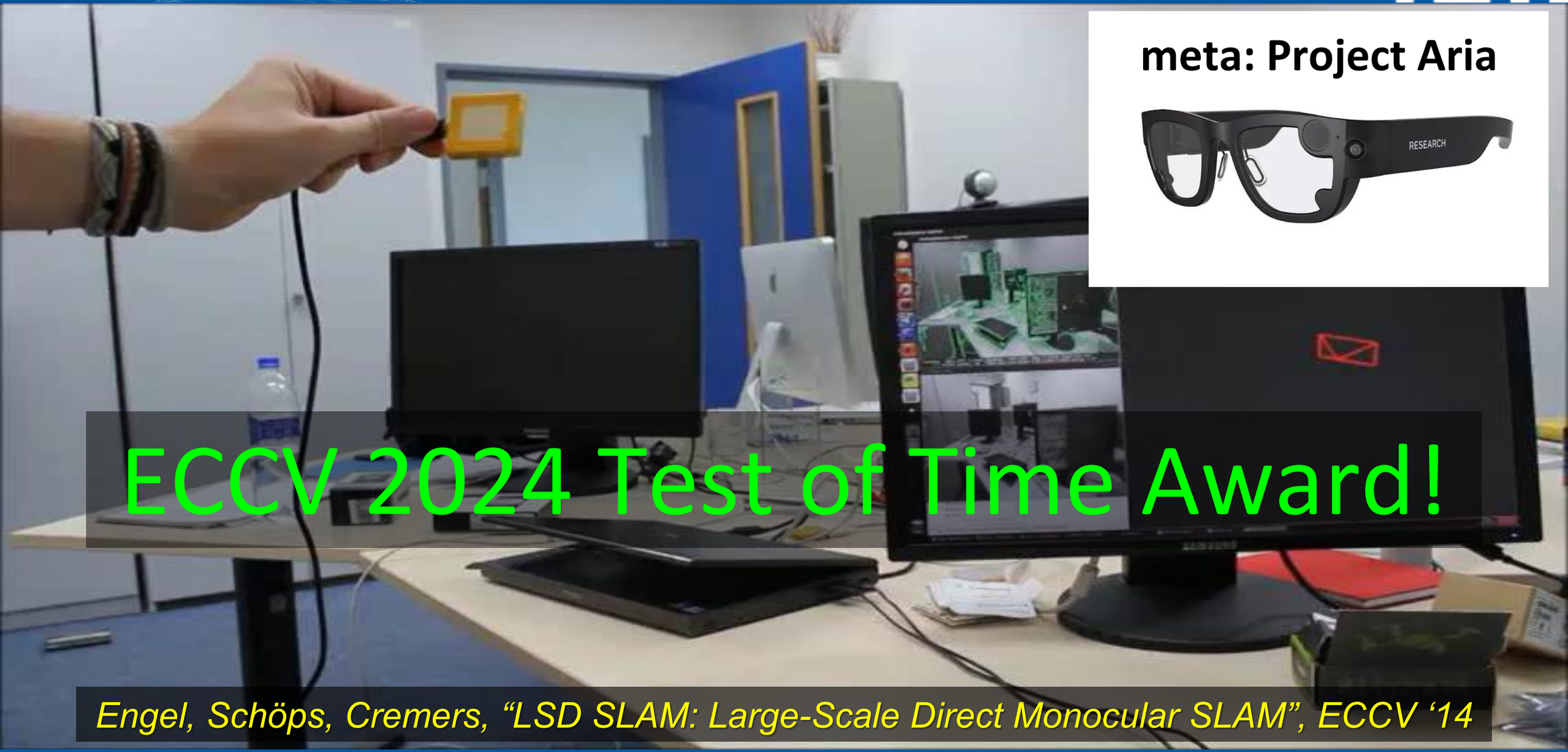
Observation with an RGB-D Camera

Color

Filtered Depth

Weiss et al., CVPR 2020

LSD SLAM: Large-Scale Direct Monocular SLAM



meta: Project Aria

ECCV 2024 Test of Time Award!

Engel, Schöps, Cremers, "LSD SLAM: Large-Scale Direct Monocular SLAM", ECCV '14

Dense Reconstructions from a Single Camera

Wimbauer et al., “MonoRec: Monocular Dense Reconstruction”, CVPR ‘21

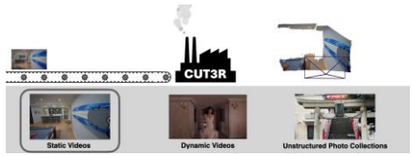
Reconstruction from Casual Videos

AnyCam: Reconstruction from Casual Videos

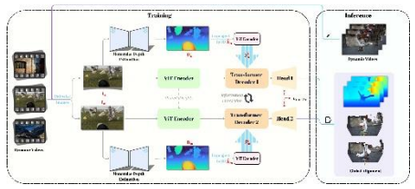
Monst3r (ICLR 2025)

MegaSaM (CVPR 2025)

Cut3r (CVPR 2025)



Align3r (CVPR 2025)



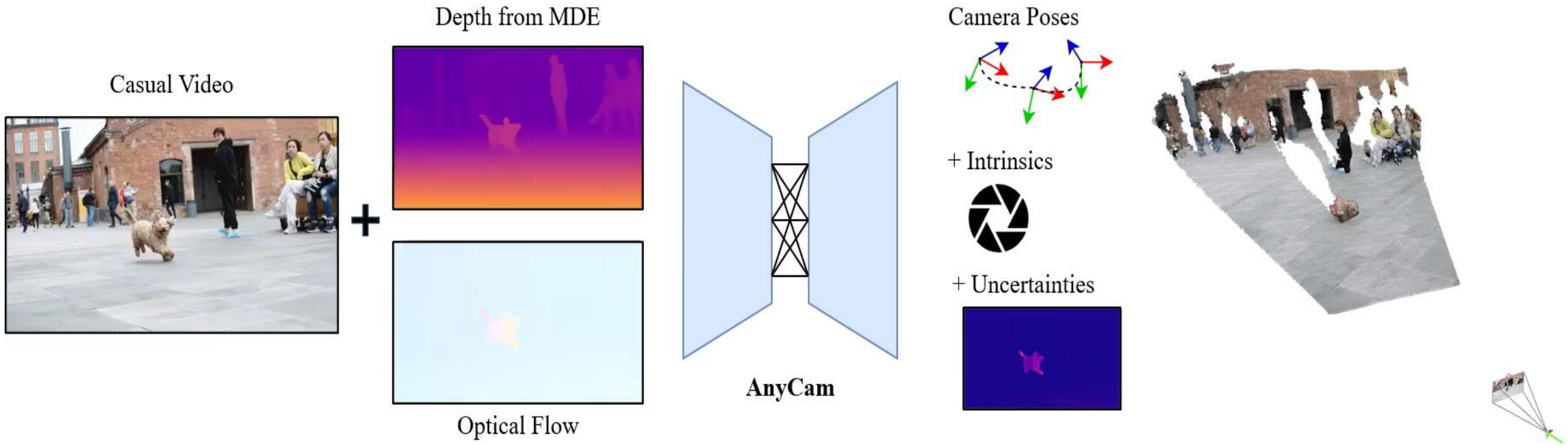
Impressive results

but

Supervised training

- ⚠ Expensive data collection
- ⚠ Limited datasets
- ⚠ Dataset biases
- ⚠ Sim-to-real gap

AnyCam: Reconstruction from Casual Videos



Felix Wimbauer[1,2,3] Weirong Chen[1,2,3] Dominik Muhle[1,2] Christian Rupprecht[3] Daniel Cremers[1,2]

[1]Technical University of Munich [2]MCML [3]University of Oxford

Wimbauer et al., "AnyCam: Learning to Recover Camera Poses and Intrinsics from Casual Videos", CVPR '25

AnyCam: Reconstruction from Casual Videos

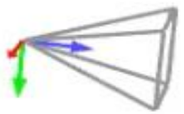
AnyCam is self-supervised on casual videos:

From YouTube

No ground truth data

Wimbauer et al., "AnyCam: Learning to Recover Camera Poses and Intrinsics from Casual Videos", CVPR '25

AnyCam: Reconstruction from Casual Videos

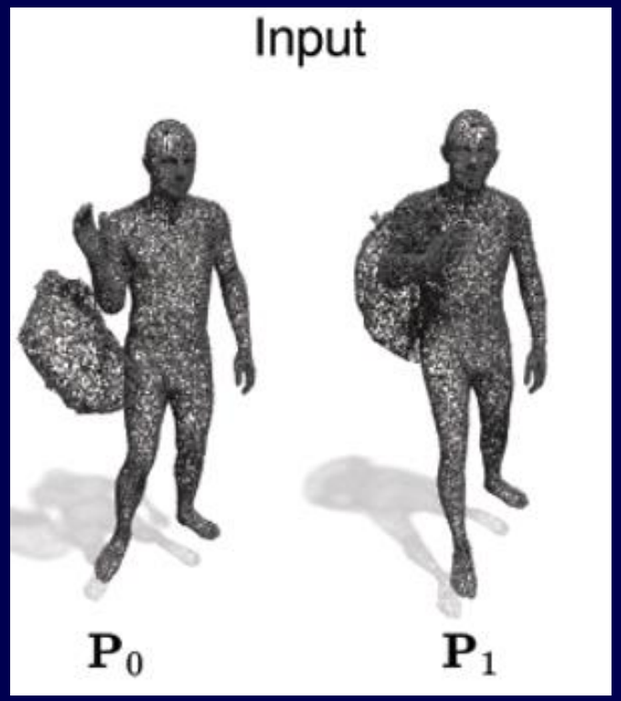


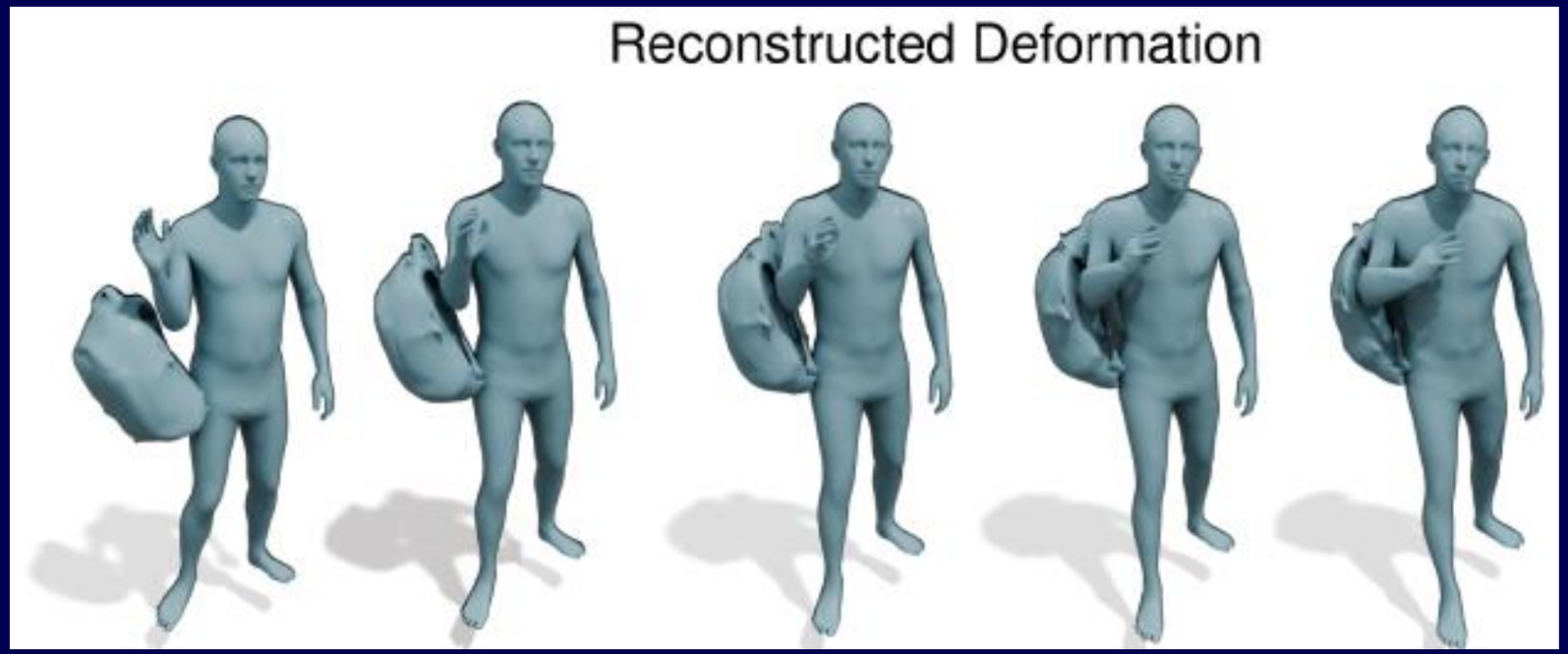
Wimbauer et al., “AnyCam: Learning to Recover Camera Poses and Intrinsics from Casual Videos”, CVPR ‘25

AnyCam: Reconstruction from Casual Videos

Wimbauer et al., "AnyCam: Learning to Recover Camera Poses and Intrinsics from Casual Videos", CVPR '25

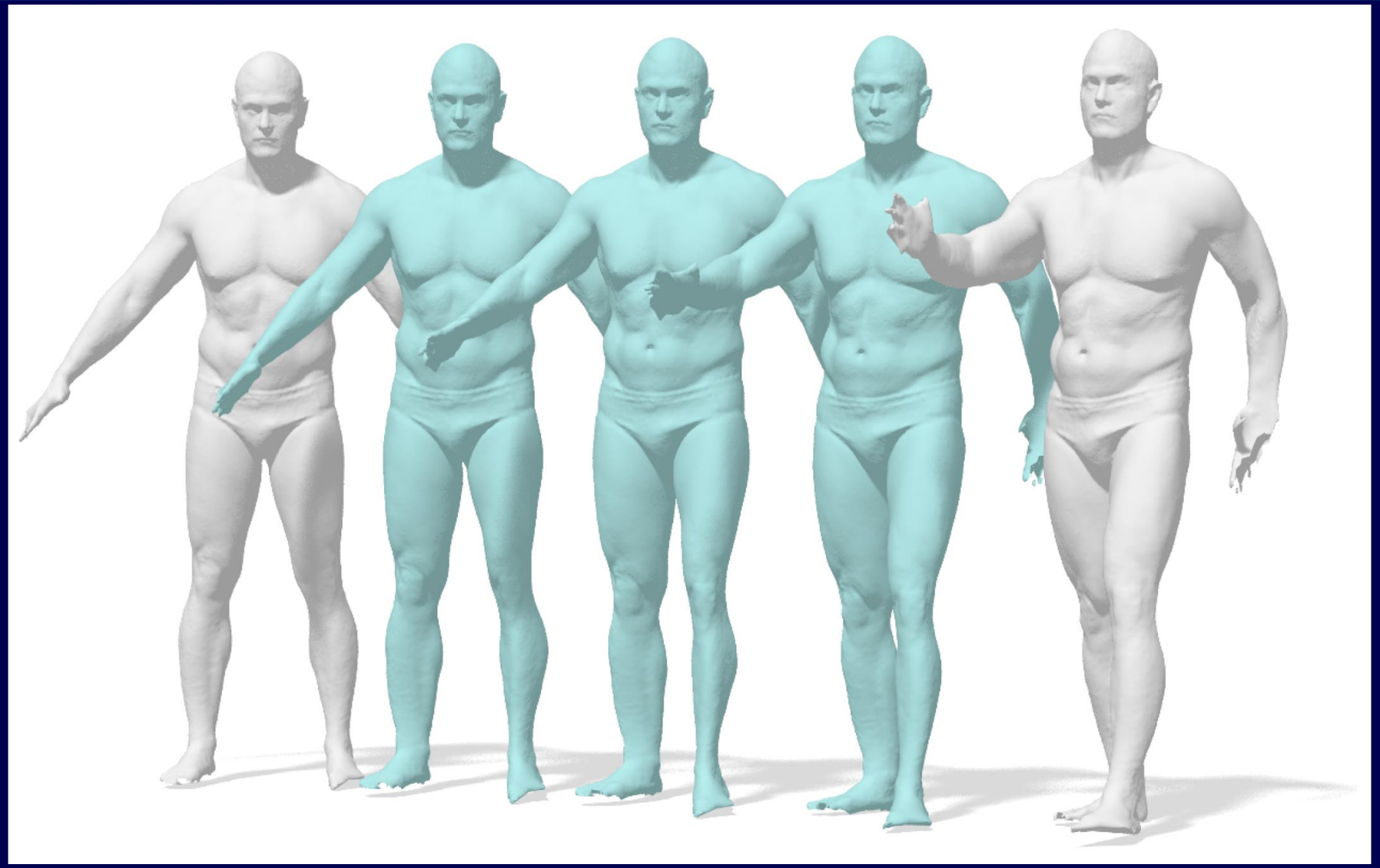
4D Reconstruction from Sparse Observations





Sang et al., "4Deform: Neural Surface Deformation for Robust Shape Interpolation", CVPR '25

Volume-Preserving Shape Interpolation



Eisenberger, Laehner, Cremers, SGP 2019

Volume-Preserving Shape Interpolation

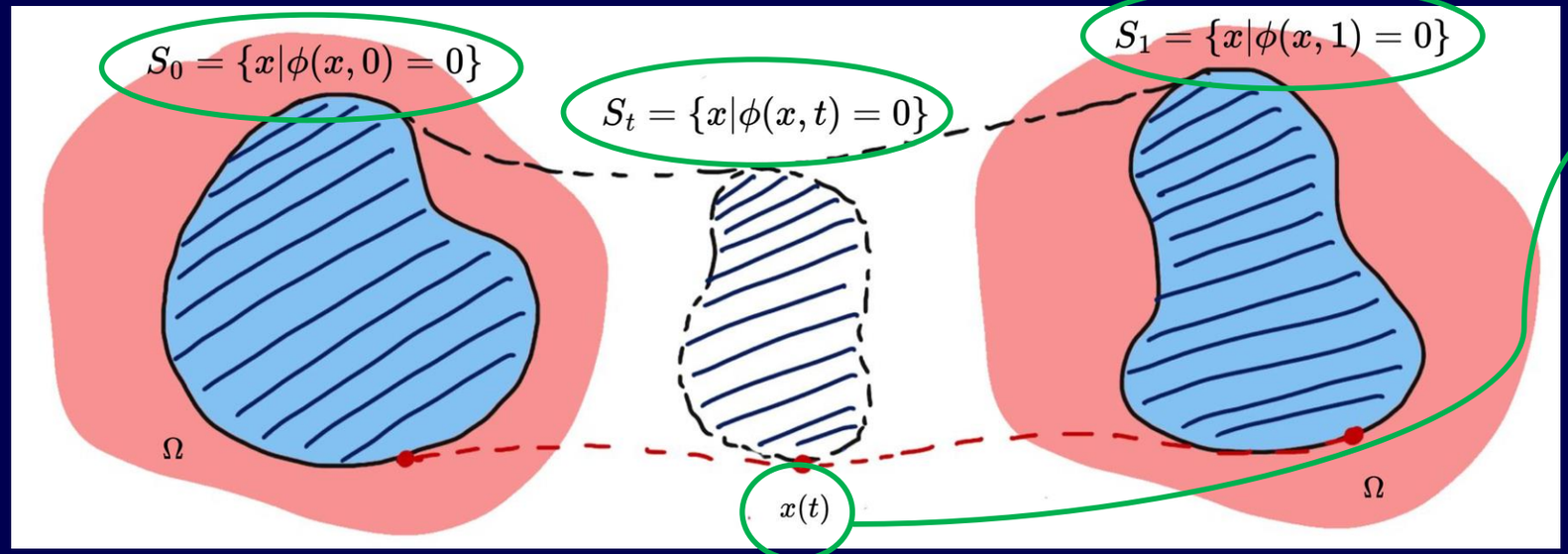
Eisenberger, Laehner, Cremers, SGP 2019

4D Reconstruction from Sparse Observations

keyframe point clouds keyframe kinect data generated intermediate frames

Sang et al., "4Deform: Neural Surface Deformation for Robust Shape Interpolation", CVPR '25

4D Reconstruction from Sparse Observations



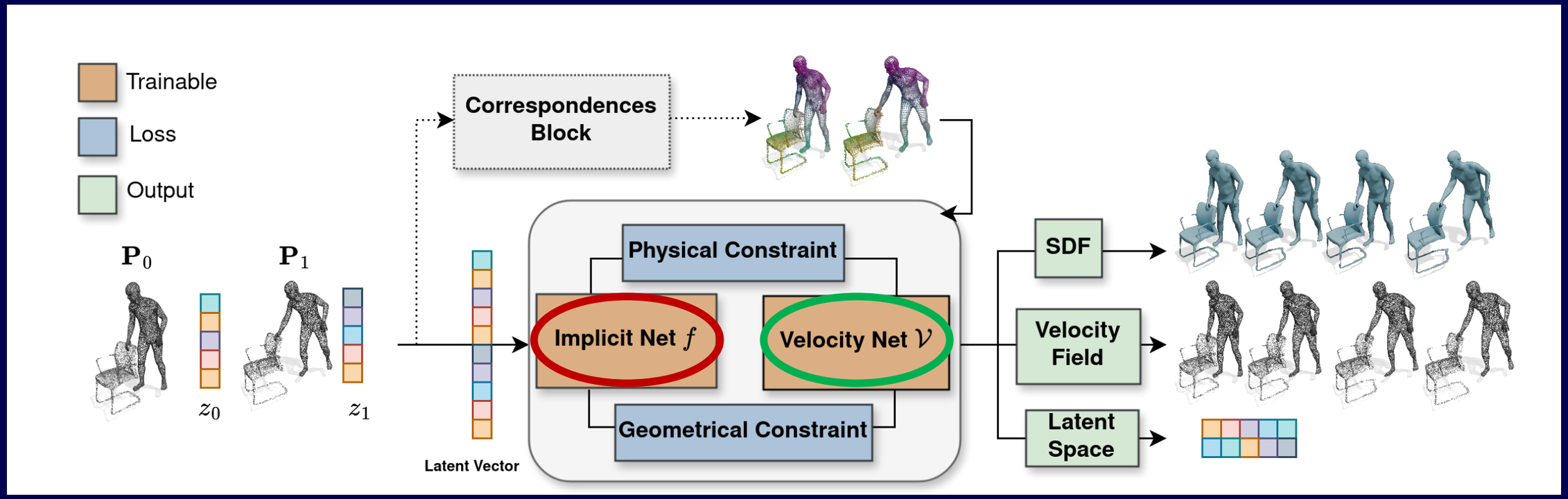
$$\phi(x(t), t) = 0$$

$$\frac{d}{dt}\phi(x,t) = \partial_t\phi + V^\top \nabla\phi = 0$$

$$\frac{d}{dt}x(t)$$

Sang et al., "4Deform: Neural Surface Deformation for Robust Shape Interpolation", CVPR '25

4D Reconstruction from Sparse Observations



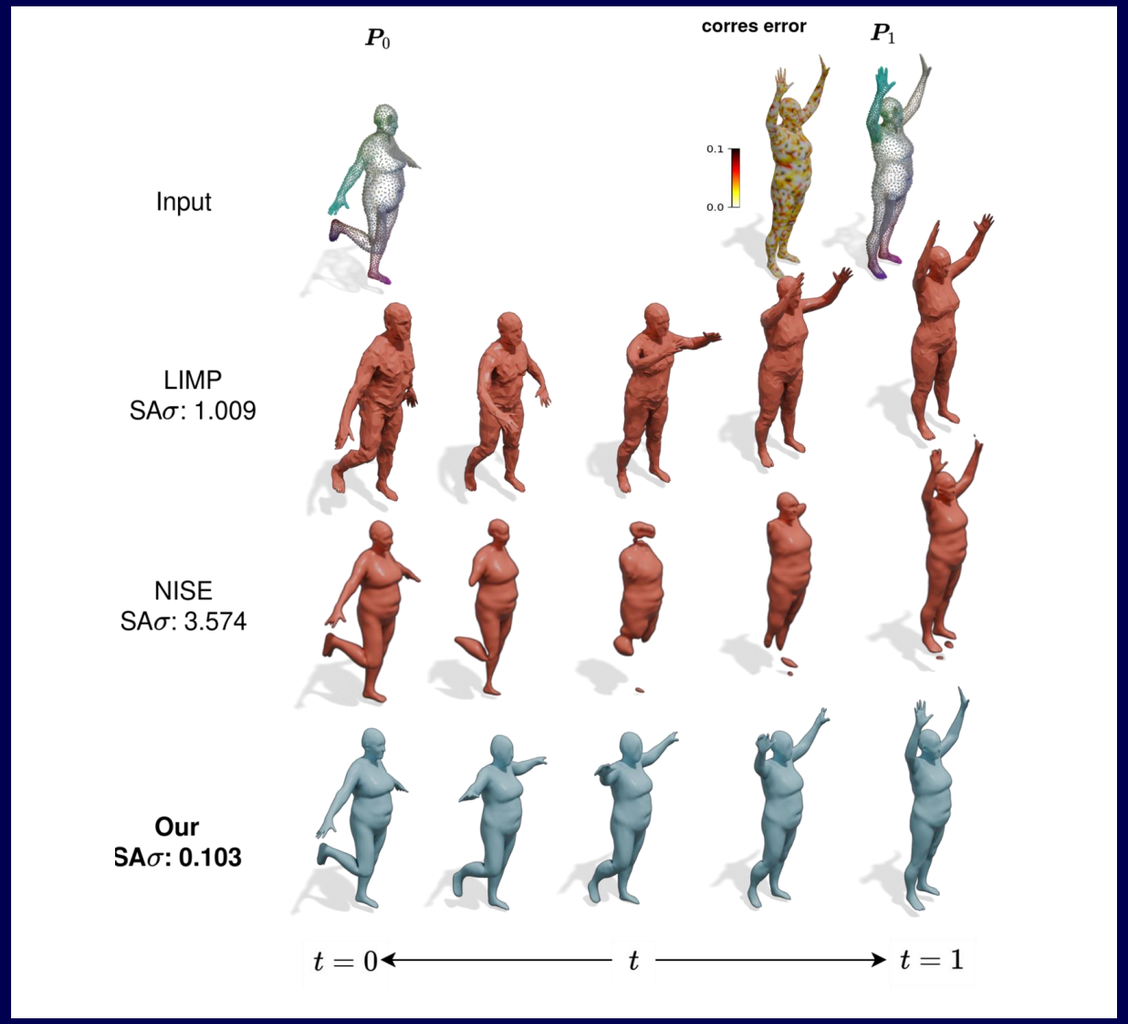
Geometrical Constraints
- Normal deformation constraint
- Level set equation constraint
- Matching loss

Physical Constraints
- Spatial smoothness velocity
- Volume preserving deformation
- Stretching constraint
- Distortion constraint

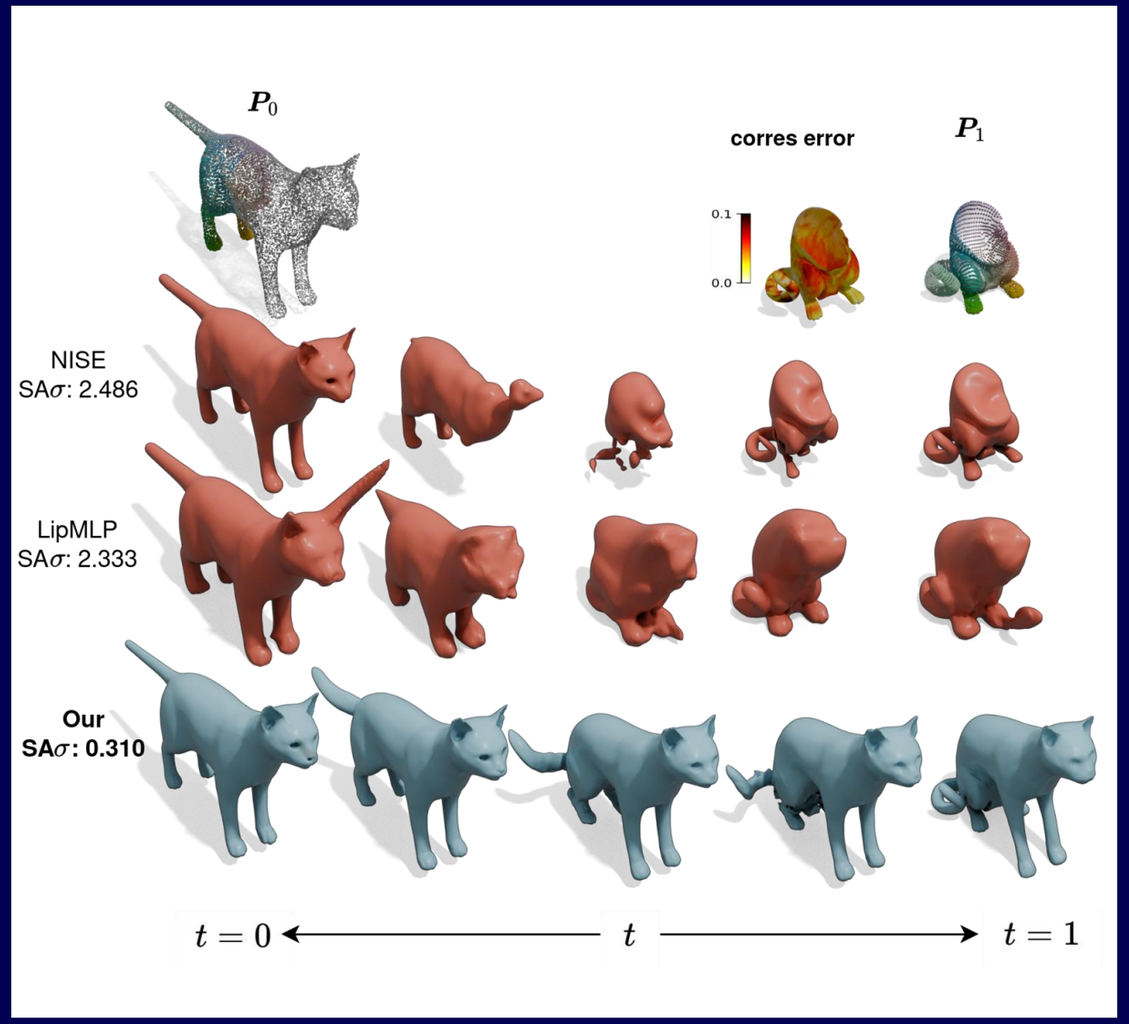
Sang et al., "4Deform: Neural Surface Deformation for Robust Shape Interpolation", CVPR '25

4D Reconstruction from Sparse Observations

Large deformation



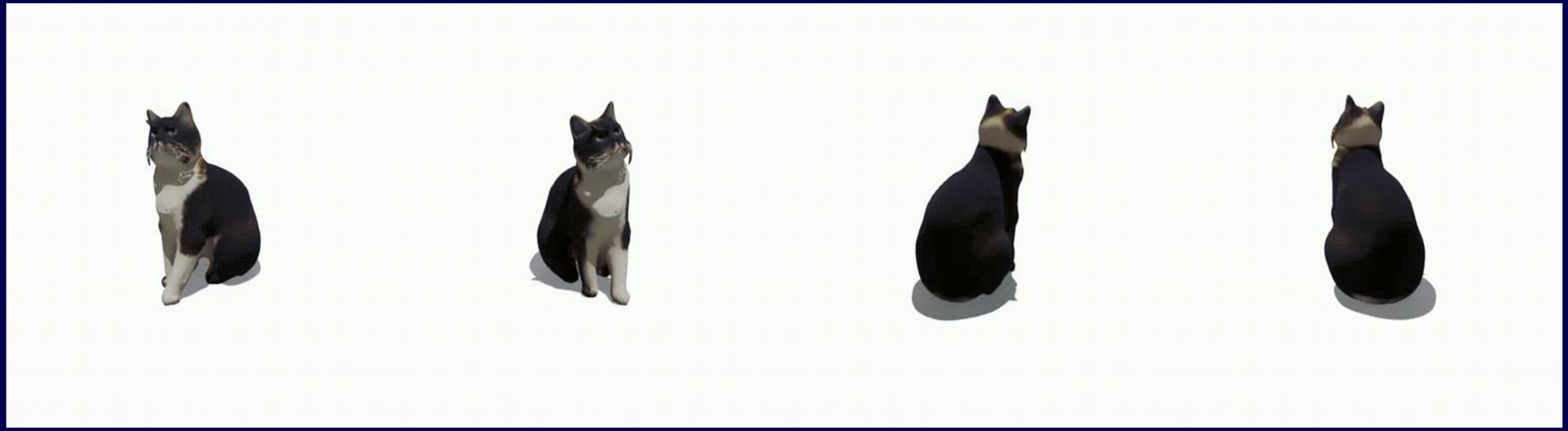
Partial shape deformation



Sang et al., "4Deform: Neural Surface Deformation for Robust Shape Interpolation", CVPR '25

TwoSquared: 4D Generation from 2D Image Pairs

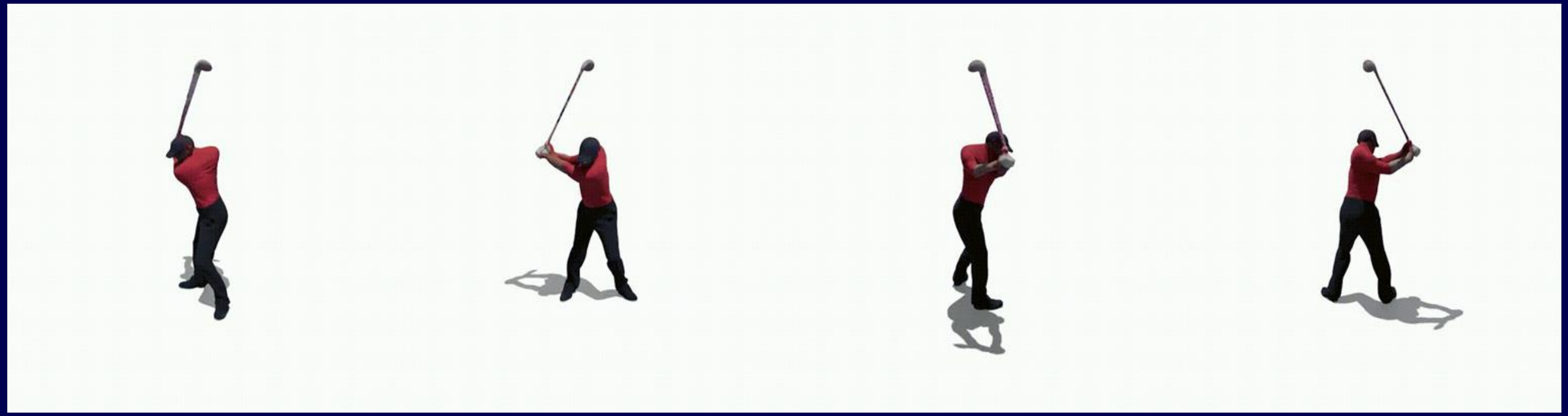
2D input images:



Sang et al., “TwoSquared: 4D Generation from 2D Image Pairs“, arxiv ‘25

TwoSquared: 4D Generation from 2D Image Pairs

2D input images:

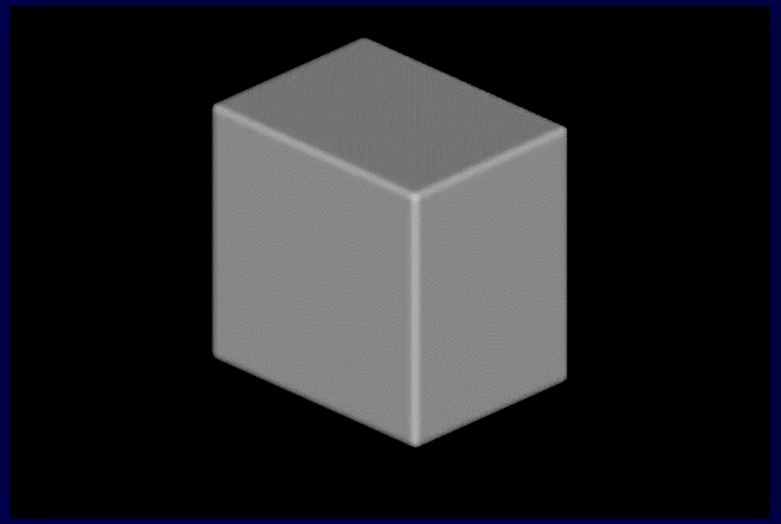


Sang et al., "TwoSquared: 4D Generation from 2D Image Pairs", arxiv '25

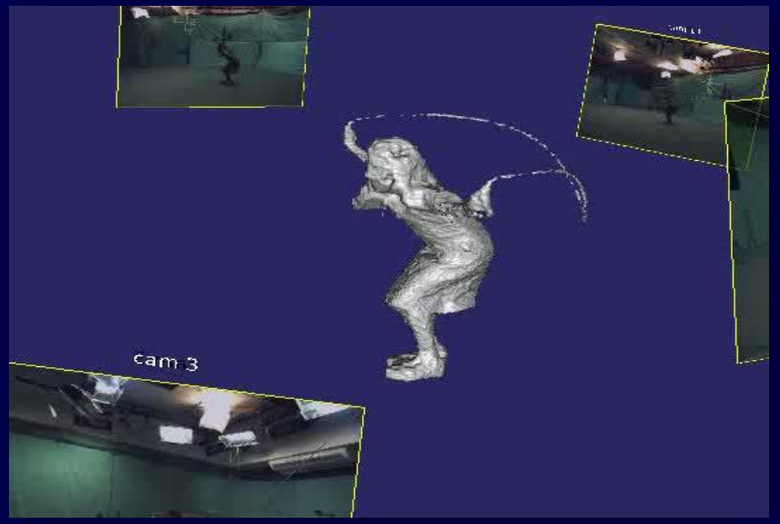
TwoSquared: 4D Generation from 2D Image Pairs

Sang et al., “TwoSquared: 4D Generation from 2D Image Pairs“, arxiv ‘25

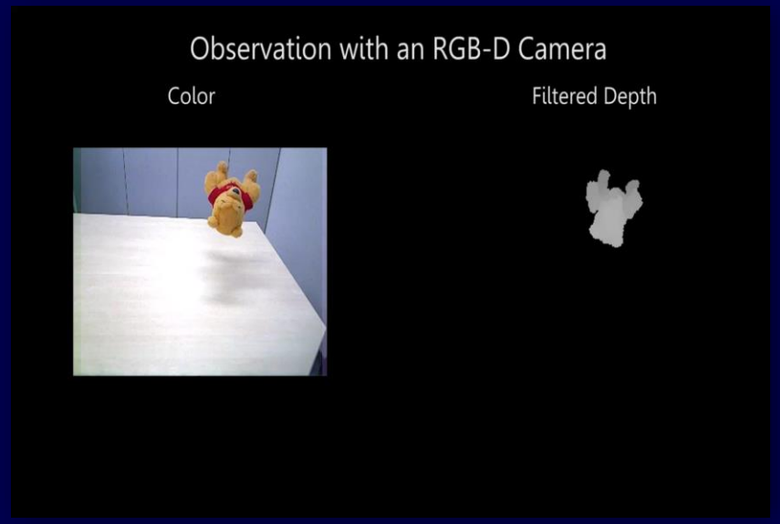
Summary



Convex multi-view 3D reconstruction

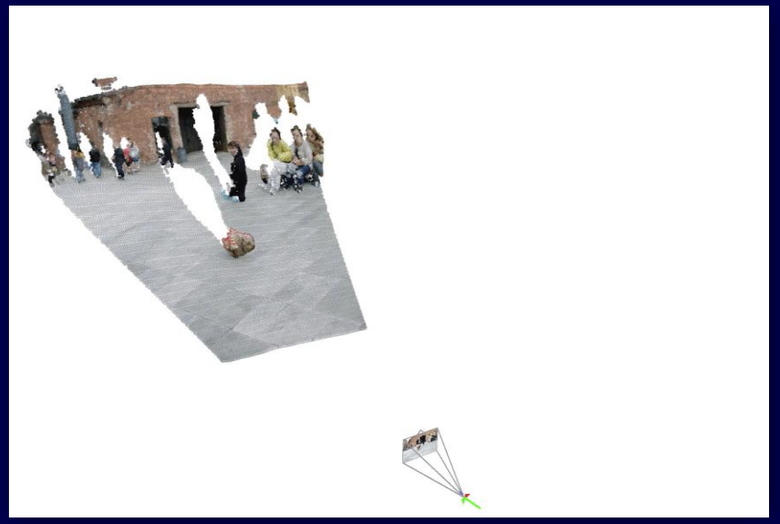


Convex multi-view 4D Reconstruction



Physical simulations from video

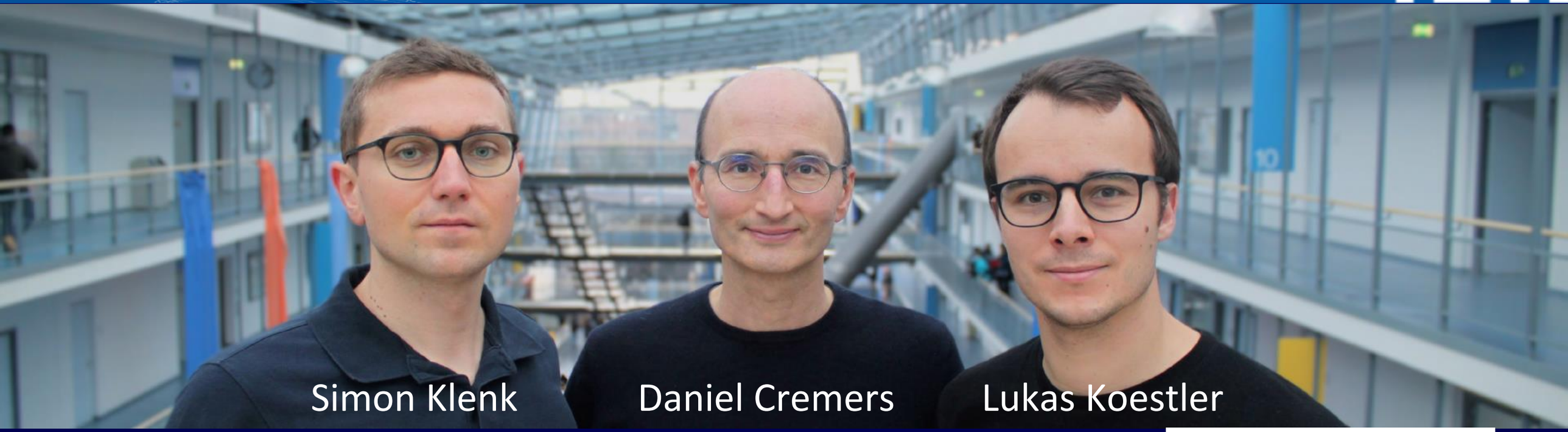
MonoRec for dense reconstruction



AnyCam for dynamic reconstruction

4Deform & TwoSquared

SE3 Labs: Startup on Spatial AI



Interested in joining
as PhD intern or full time?