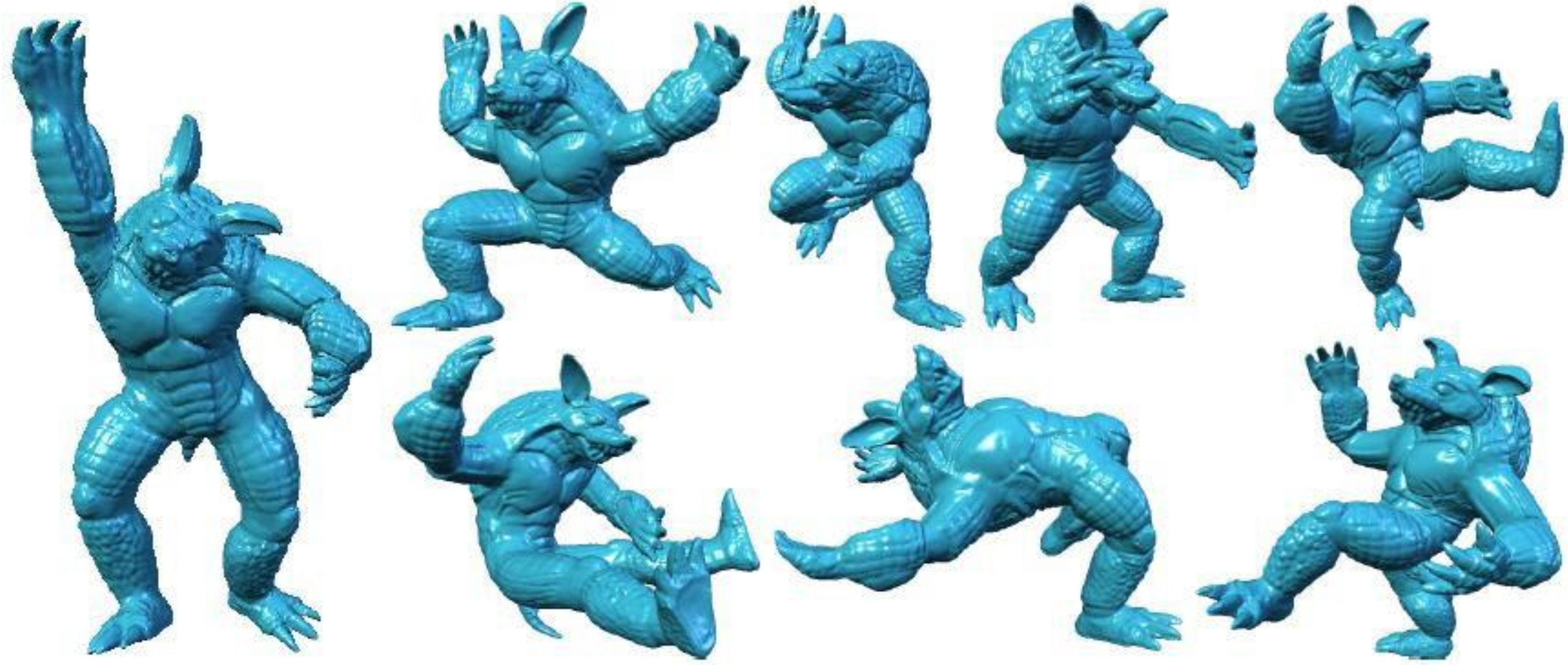
3D Scanning & Motion Capture

Non-Rigid Tracking and Reconstruction

Prof. Matthias Nießner

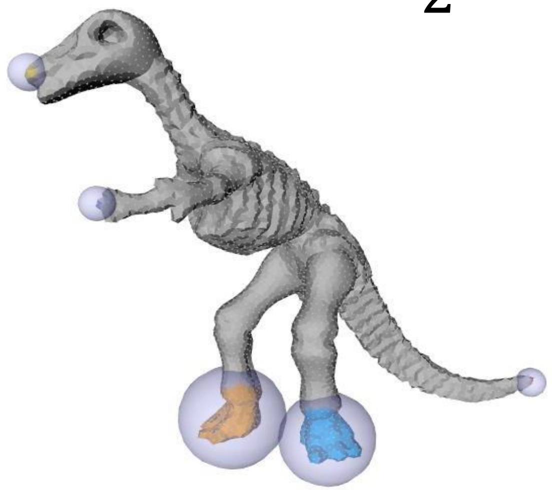
Last lecture: Mesh Deformation

- In graphics: e.g., character animation

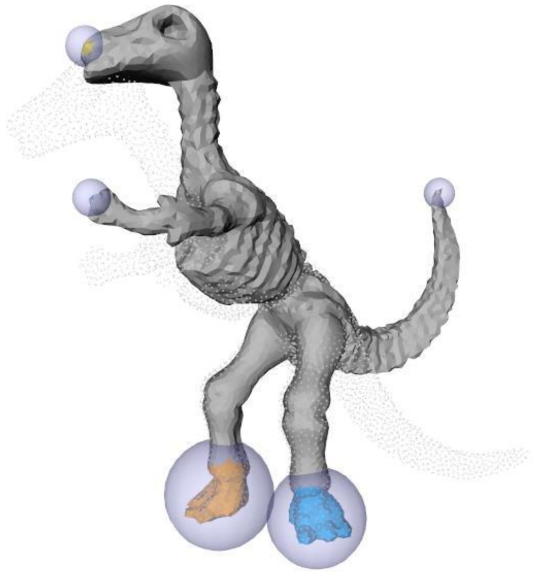


Last lecture: As-Rigid-as-Possible Deformation

- One rotation matrix per fan; sum up deviations from rigidity
- $E_{ARAP}(\boldsymbol{R}, \boldsymbol{V}) = \sum_i \sum_{j \in \mathcal{N}(i)} \left\| (v_i - v_j) - R_i(v_i' - v_j') \right\|_2^2$
- $E_{fit}(\boldsymbol{V}) = \sum_{i \in C} \left\| v_i{}' - v_i \right\|_2^2$



Initial mesh V'

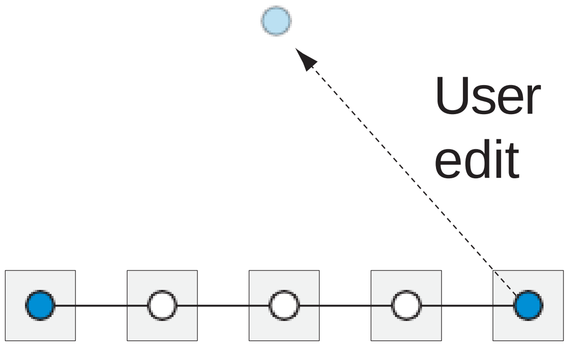


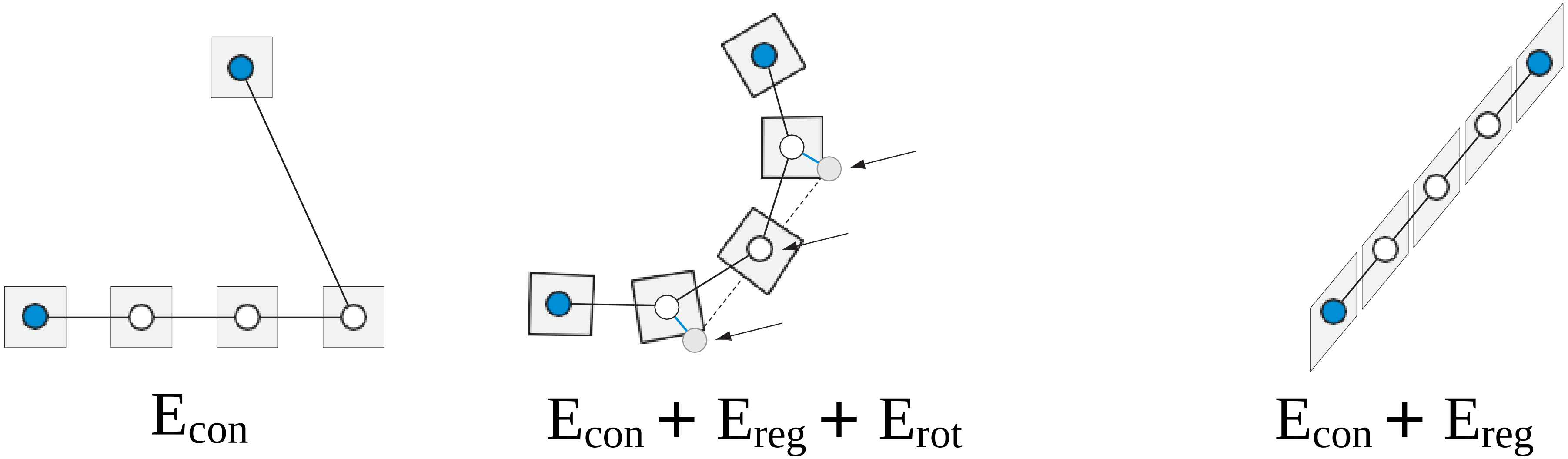
Deformed mesh V

[Sorkine et al. 08]

Last lecture: Embedded Deformation

$$E_{ED}(\boldsymbol{M},\boldsymbol{V}) = \sum_{i} \sum_{j \in \mathcal{N}(i)} \left\| (v_i - v_j) - M_i(v_i' - v_j') \right\|_2^2$$





[Sumner et al. 07]

Last lecture: ARAP vs ED

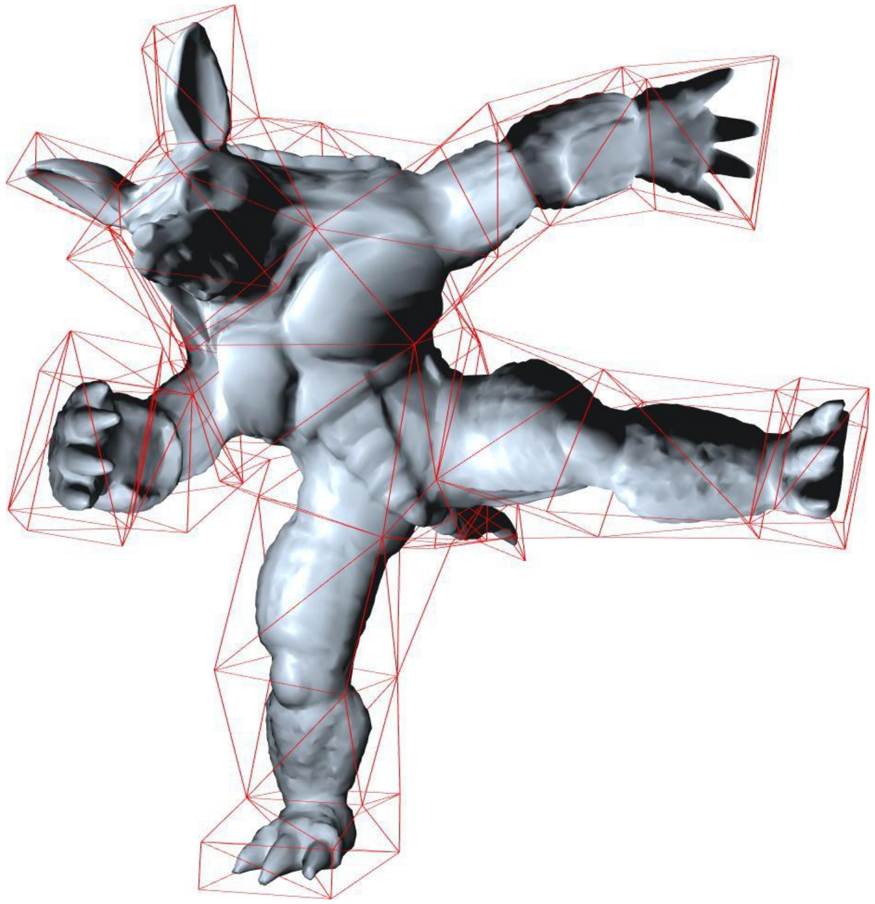
- ARAP: *hard-constraints* for rotations
  - Forced rotations matrixes -> 3 unknowns per local matrix (6 per vertex)
  - Can me efficiently optimized with flip-flop (local SVDs / Procrust)

- ED: *soft-constraints* for rotations
  - I.e., -> 9 unknowns per local matrix (12 per vertex)
  - Harder to optimize (actually a quartic problem)

Last lecture: Deformation Proxies

- Dimensionality reduction on deformations

- Directly on Mesh
- Cages
- Grids
- Skeletons (e.g., Linear Blend Skinning)
- Deformation Graphs
- Tetrahedral Mesh (full volume)

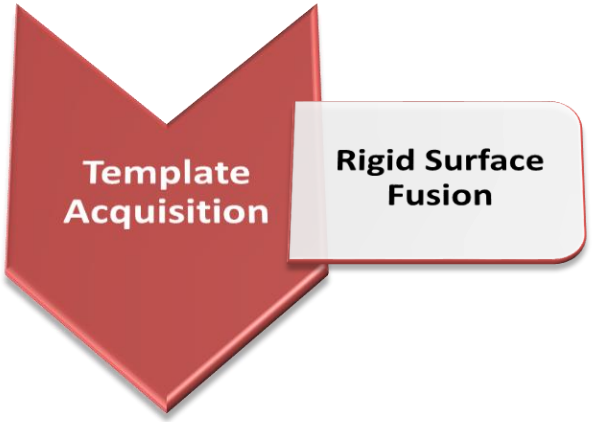


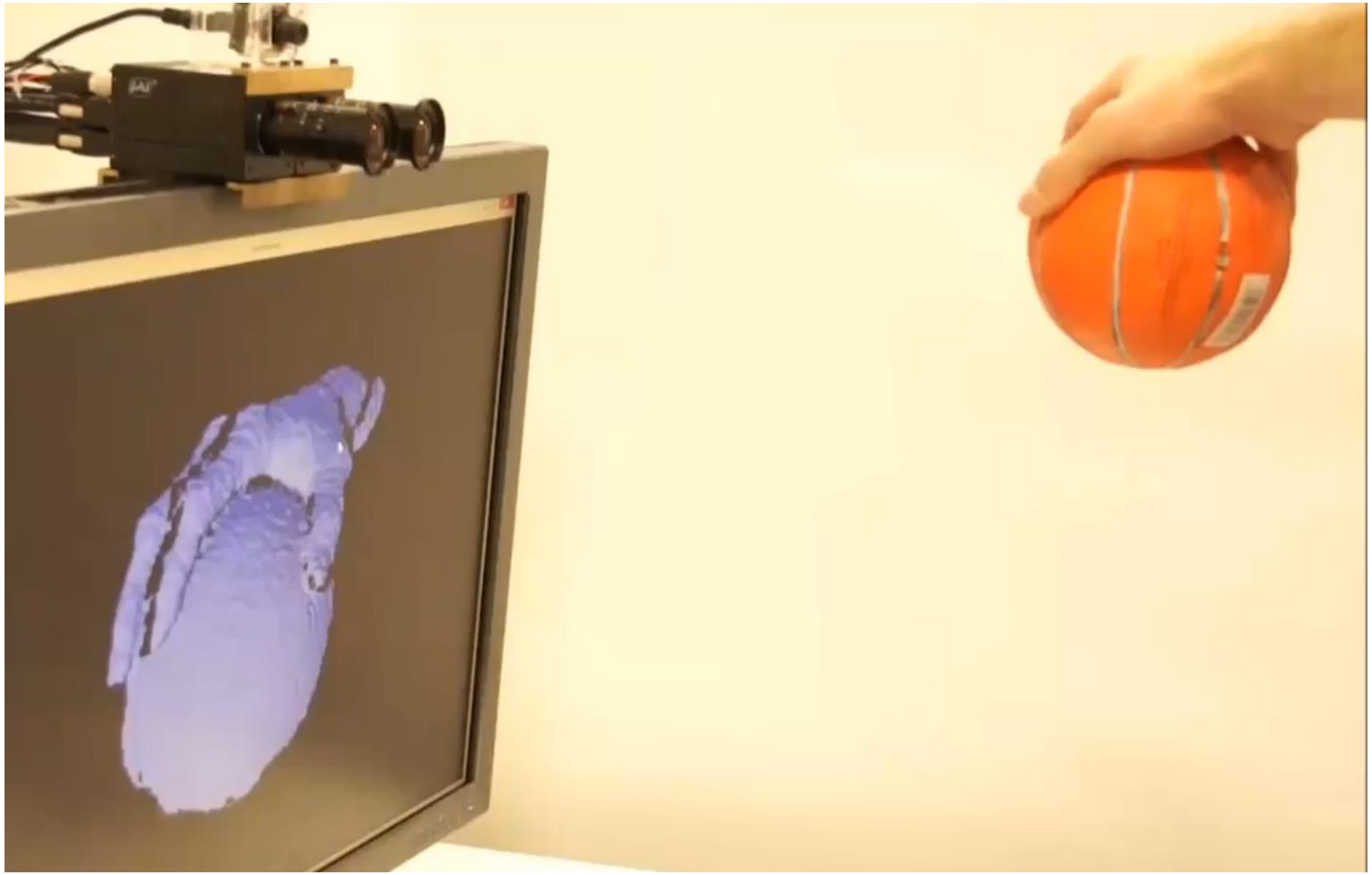
Last lecture: Non-rigid Tracking

- We know how to deform a mesh
  - Regularizers + Deformation proxies
  - User picks handles to edit the mesh (i.e., constraints for deformation)

- Tracking / Fitting:
  - Find these correspondences from data!
  - E.g., Non-rigid ICP
  - E.g., Sparse Features (e.g., SIFT, SURF)

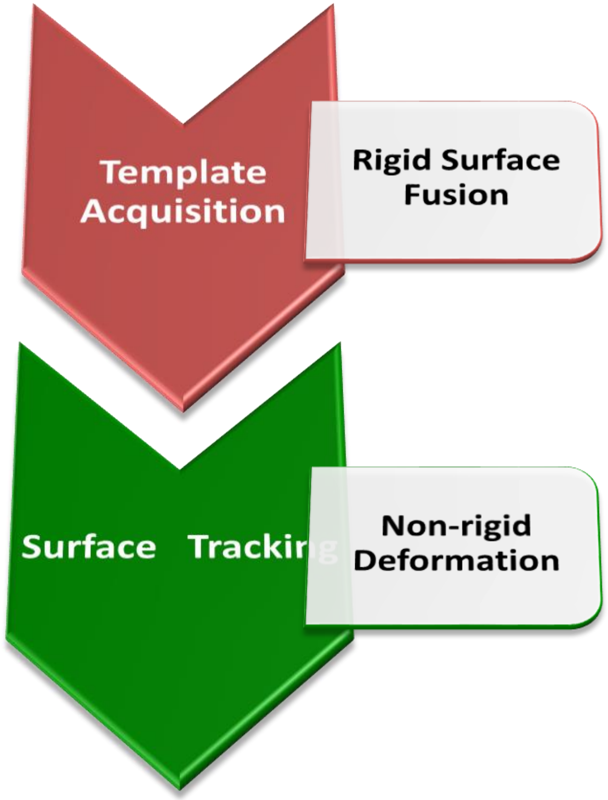
Non-rigid Tracking

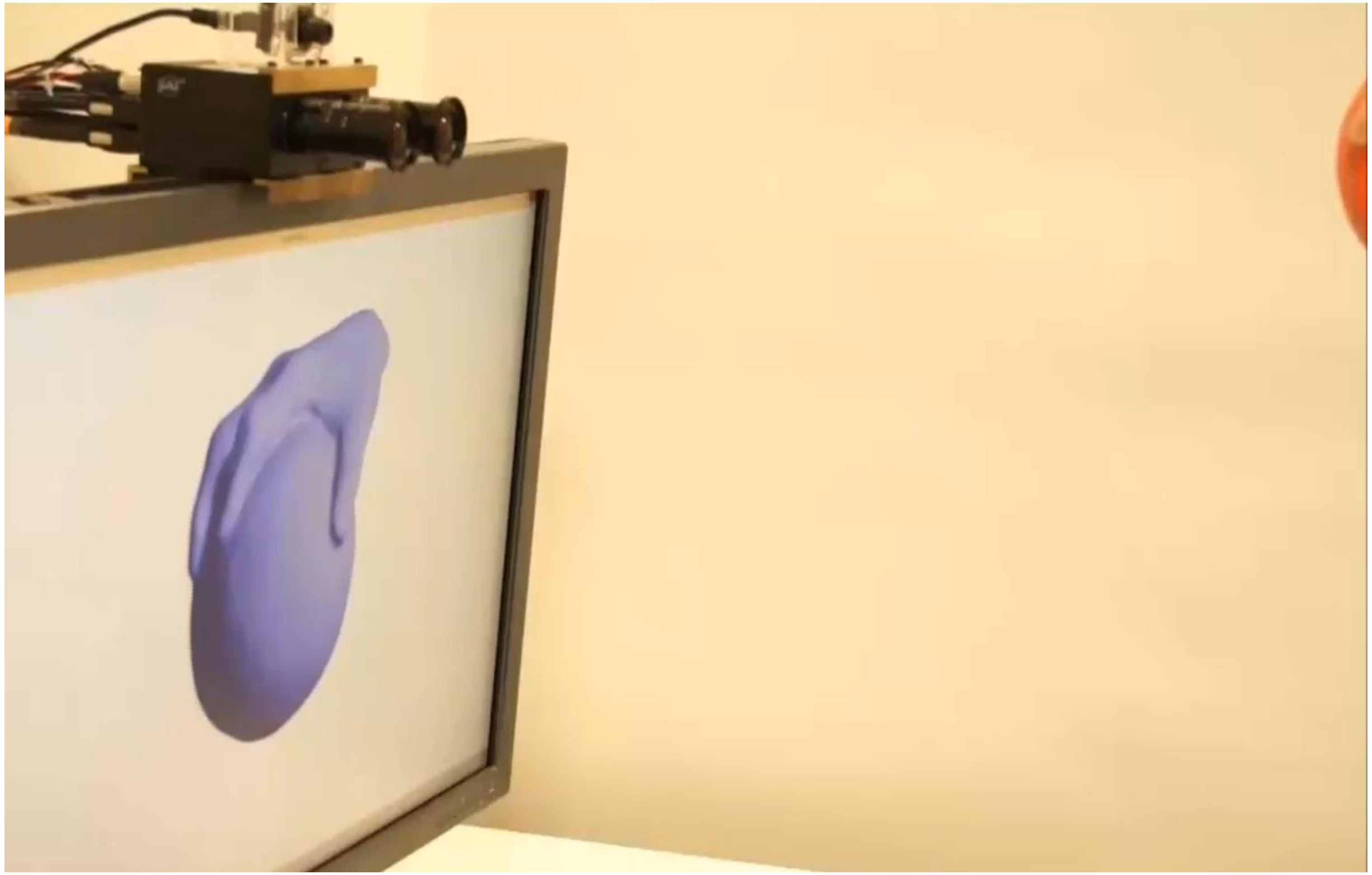




TOG'14 [Zollhöfer et al.]: Deformables

Non-rigid Tracking





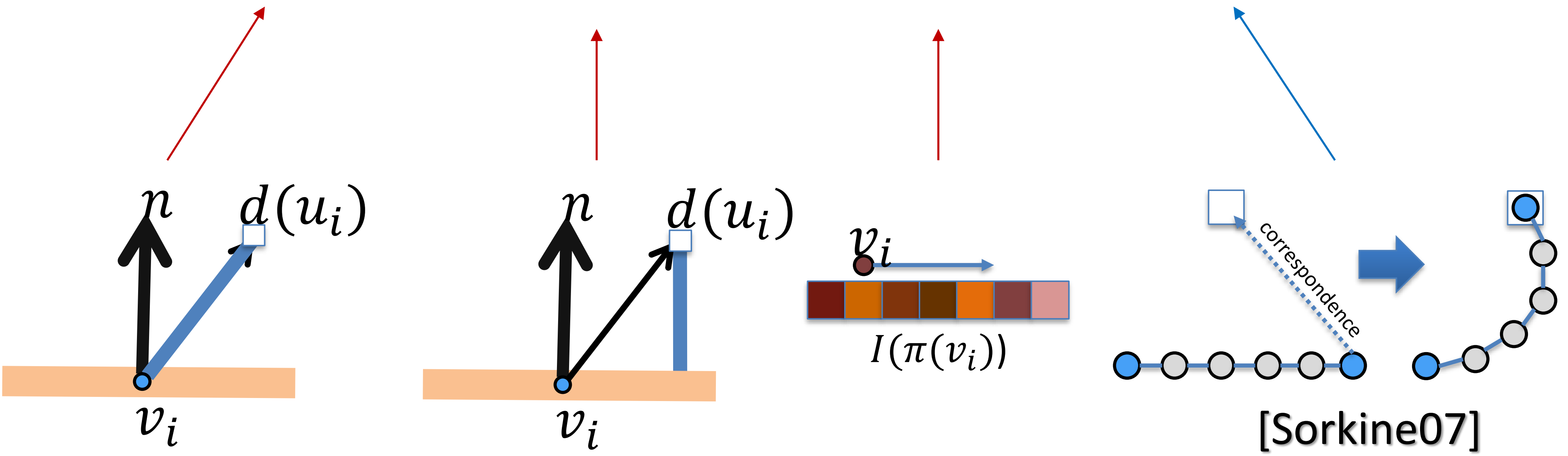
TOG’14 [Zollhöfer et al.]: Deformables

Non-rigid Tracking: Objective

Data Terms

Regularization

$$E = \lambda_{point} E_{point} + \lambda_{plane} E_{plane} + \lambda_{color} E_{color} + \lambda_{reg} E_{reg}$$

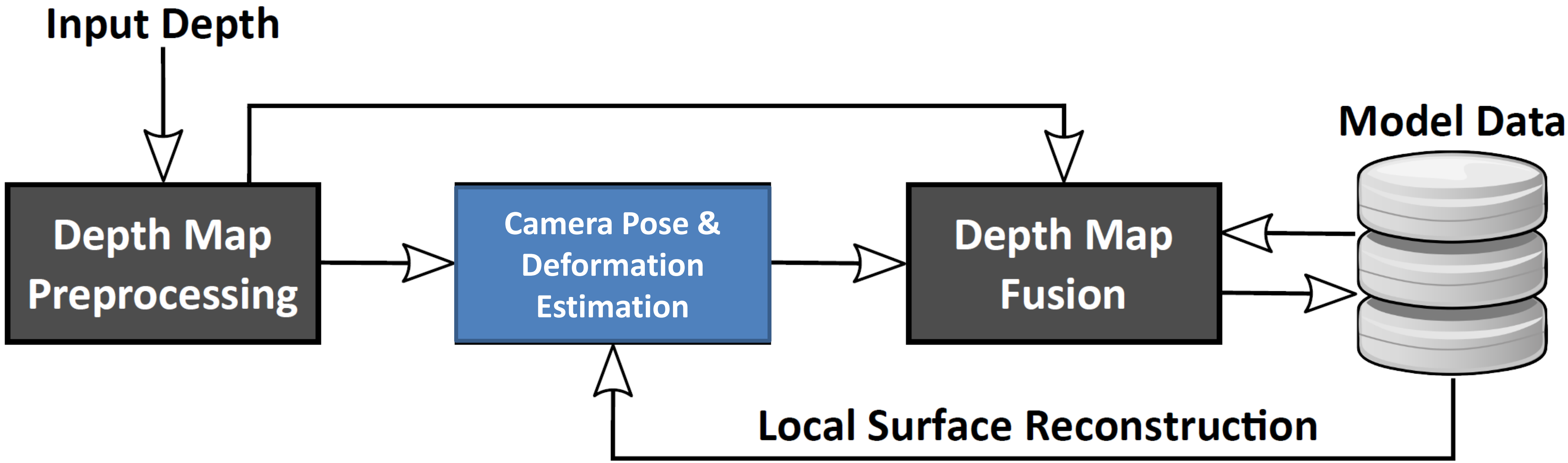


[Sorkine07]

Today: Non-Rigid Tracking and Reconstruction

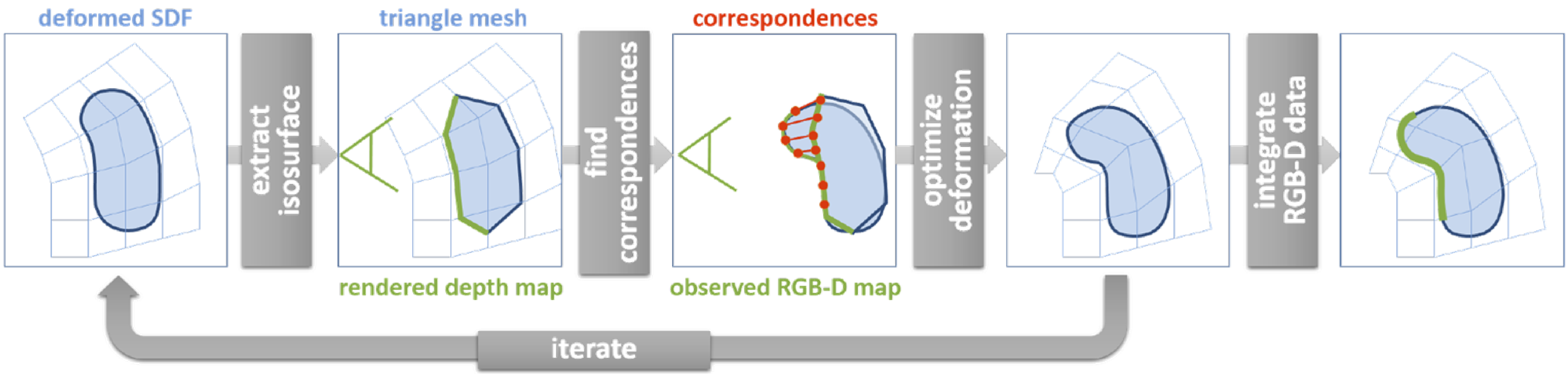
Dynamic Surface Reconstruction

- Non-rigid tracking + volumetric fusion

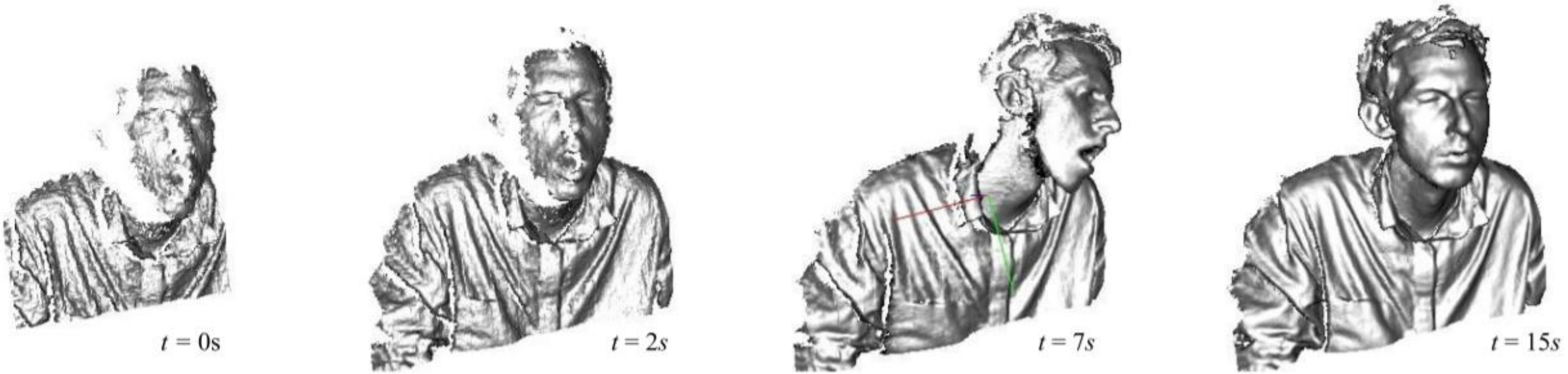


Dynamic Surface Reconstruction

- Non-rigid tracking + volumetric fusion



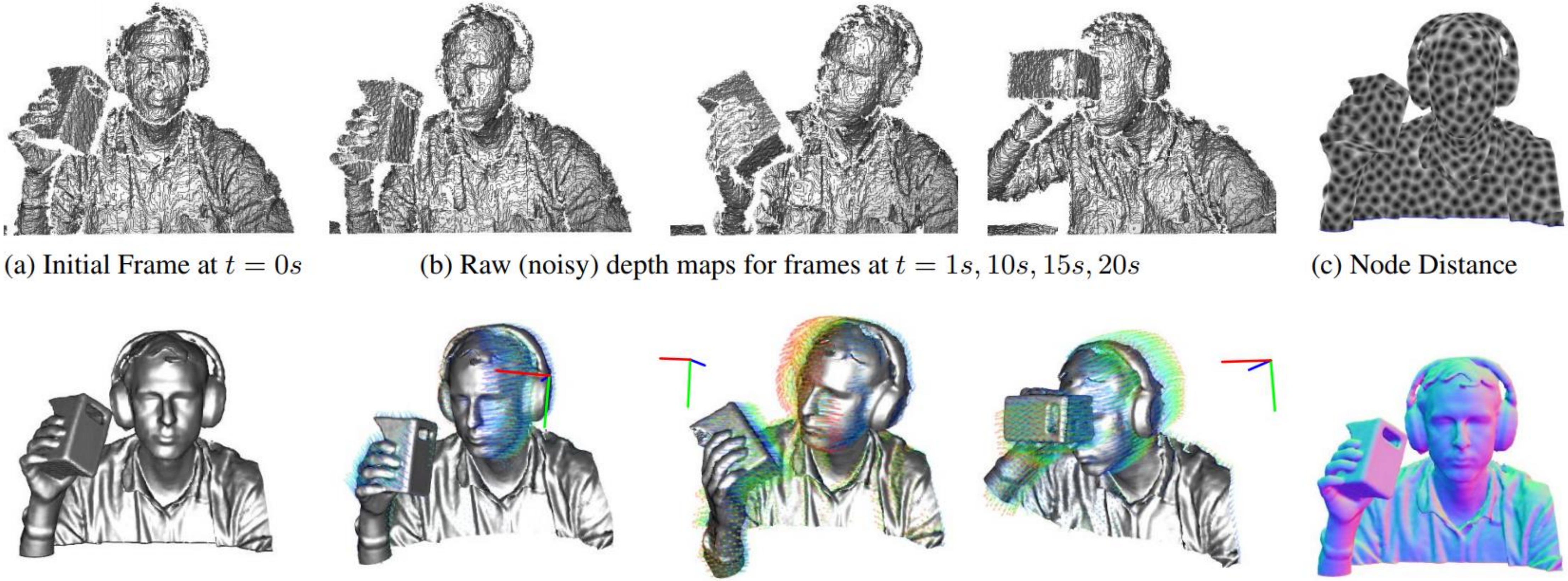
Dynamic Fusion



[Newcombe et al. 15]

Dynamic Fusion

- ED + (dense) non-rigid ICP + volumetric fusion



(a) Initial Frame at $t = 0s$ (b) Raw (noisy) depth maps for frames at $t = 1s, 10s, 15s, 20s$ (c) Node Distance

(d) Canonical Model (e) Canonical model warped into its live frame (f) Model Normals

[Newcombe et al. 15]

Dynamic Fusion

- Deformation graph (ED) + non-rigid ICP + volumetric fusion

- For each frame
  – Take depth frame
  – Extract depth map from current warped model (using Marching Cubes)
  – Align input frame with extracted frame using (dense) non-rigid ICP
  – Integrate current frame into canonical model
  – Go to next frame

[Newcombe et al. 15]

Dynamic Fusion

*Dynamic***Fusion**:

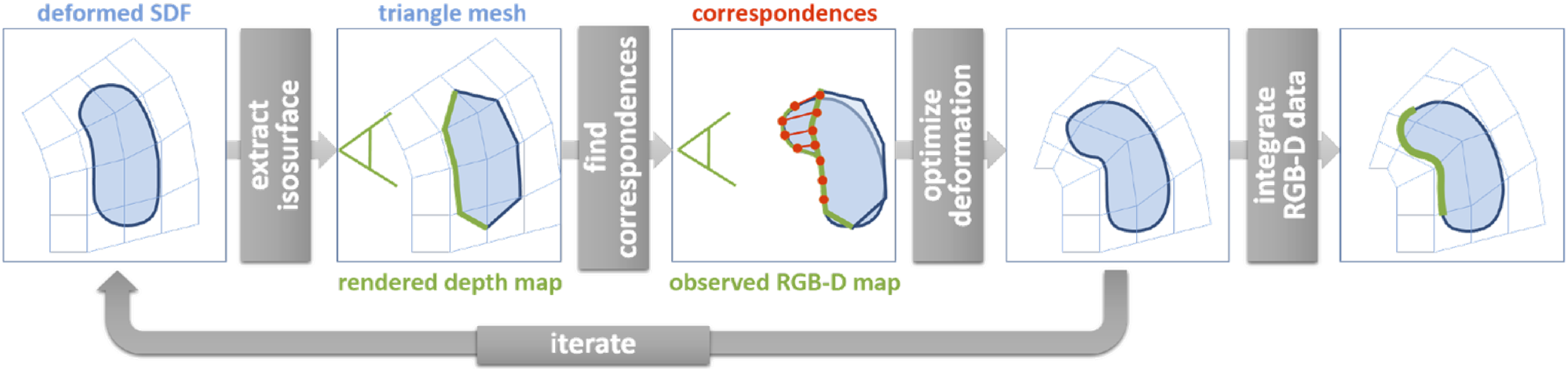
Reconstruction & Tracking of Non-rigid Scenes in *Real-Time*

Richard Newcombe, Dieter Fox, Steve Seitz

Computer Science and Engineering,
University of Washington

[Newcombe et al. 15]

Dynamic Surface Reconstruction



ECCV'16 [Innmann et al.]: VolumeDeform

VolumeDeform

ECCV'16 [Innmann et al.]: VolumeDeform

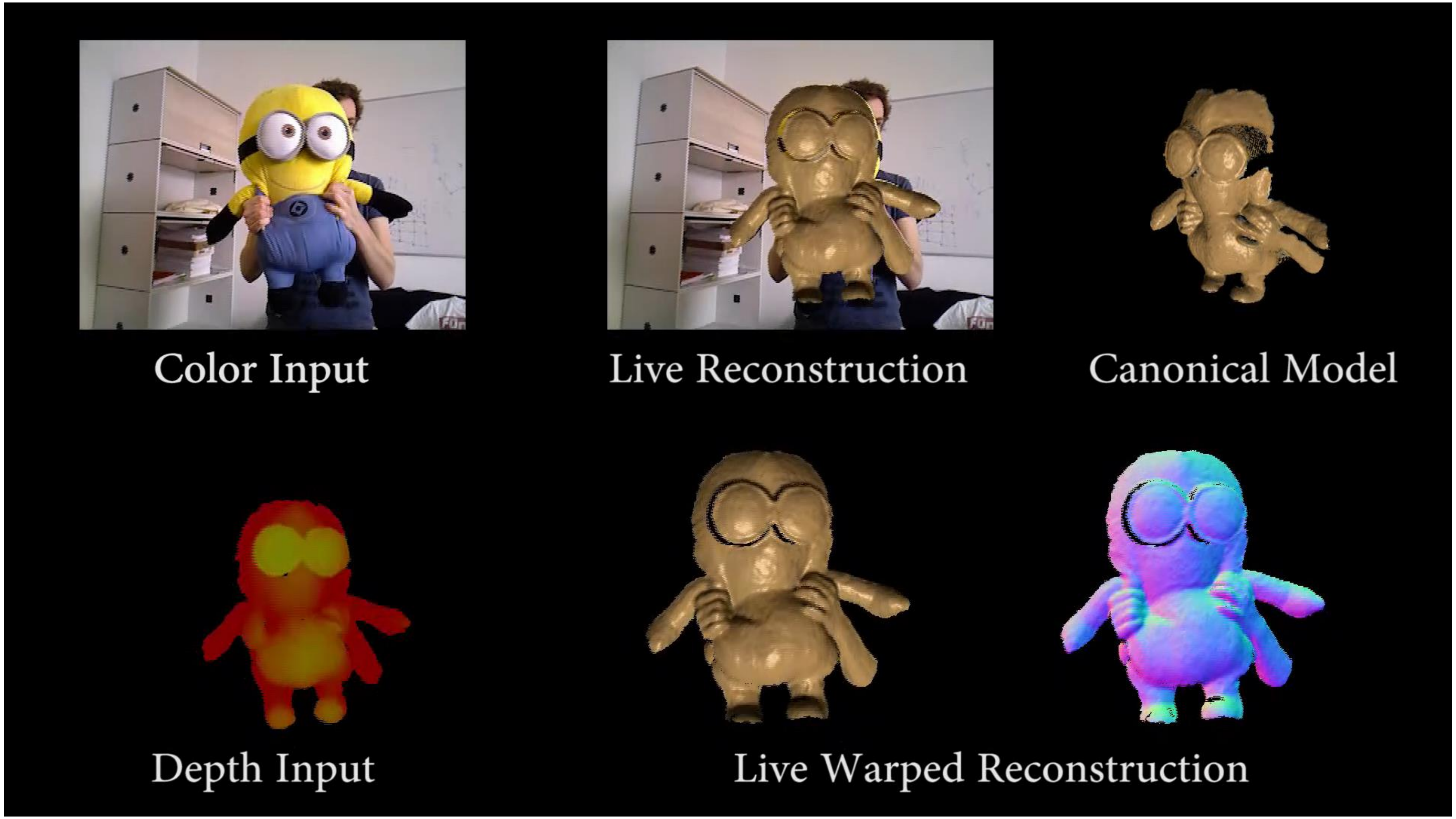
VolumeDeform

- Regular deformation grid (ARAP) + non-rigid ICP (dense depth + sparse SIFT matches) + Volumetric Fusion

- For each frame
  - Take depth frame
  - Extract depth map from current warped model (using Marching Cubes)
  - Align input frame with extracted frame using (dense) non-rigid ICP
  - Integrate current frame into canonical model
  - Go to next frame

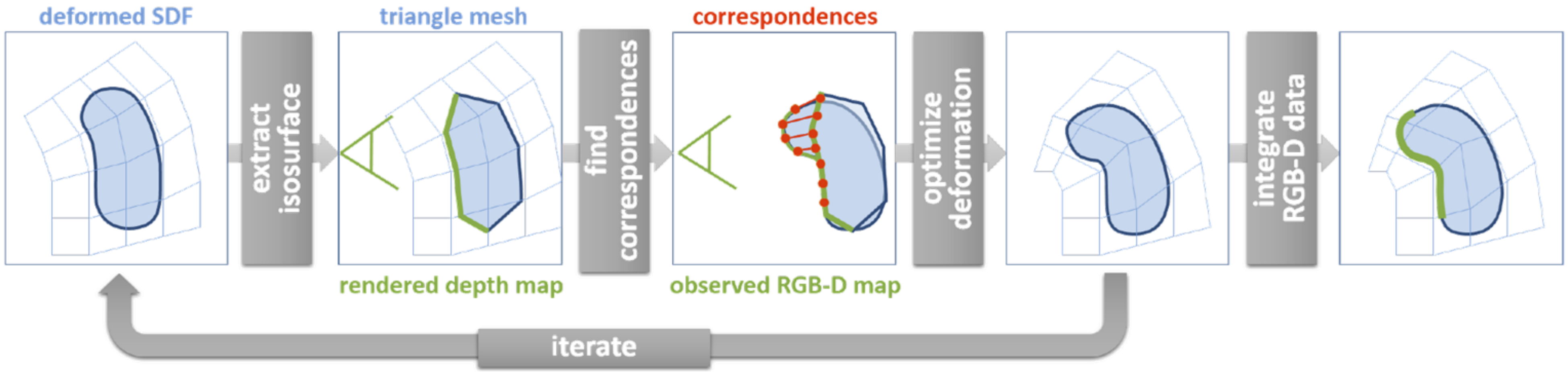
ECCV'16 [Innmann et al.]: VolumeDeform

Dynamic Surface Reconstruction



ECCV'16 [Innmann et al.]: VolumeDeform

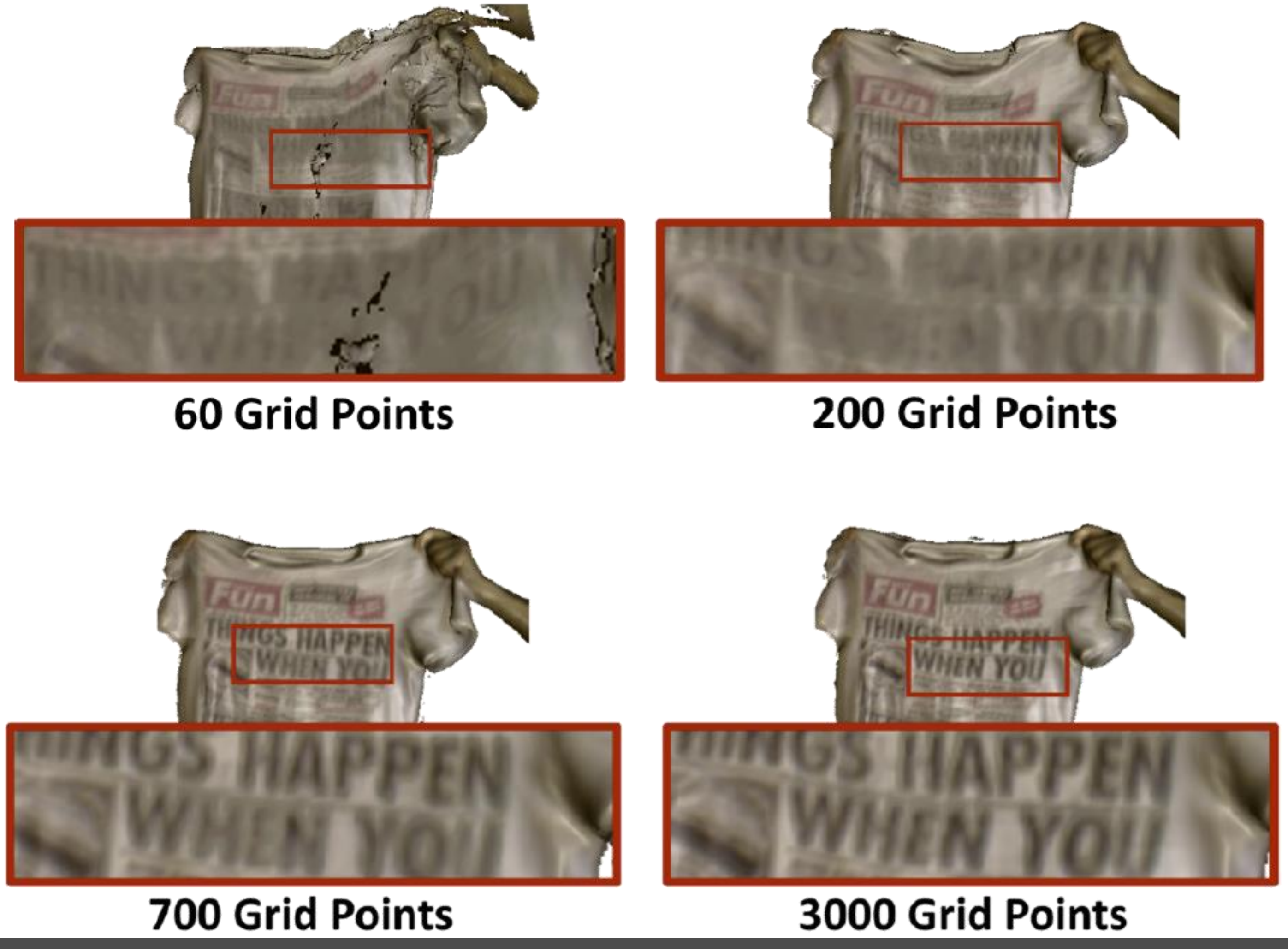
VolumeDeform



- Key differences to [Newcombe et al. 15]
  - Sparse RGB correspondences as global anchors
  - High-resolution deformation field (same as surface)

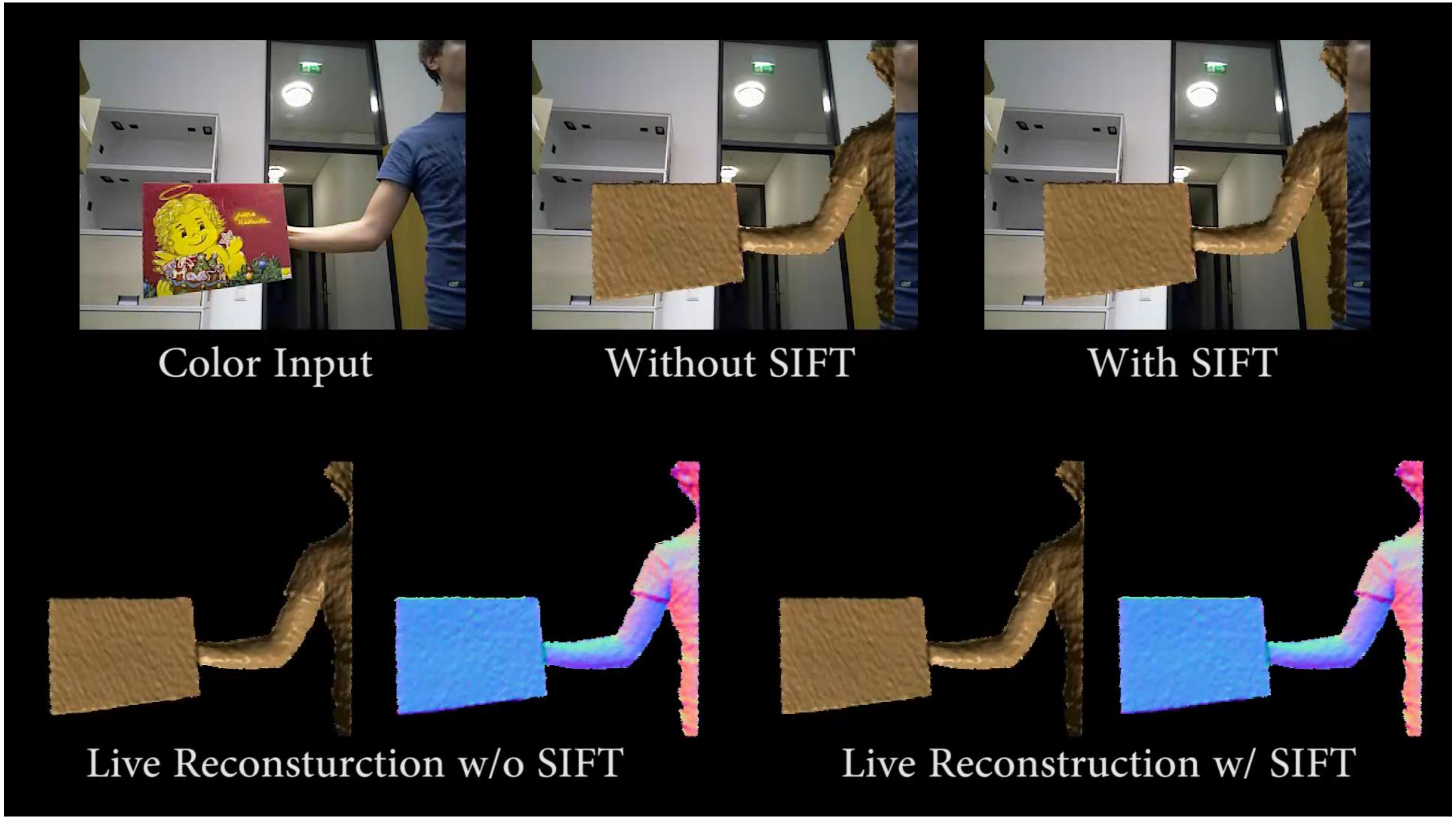
ECCV'16 [Innmann et al.]: VolumeDeform

Deformation Field Resolution



ECCV’16 [Innmann et al.]: VolumeDeform

Impact of SIFT Anchor Points



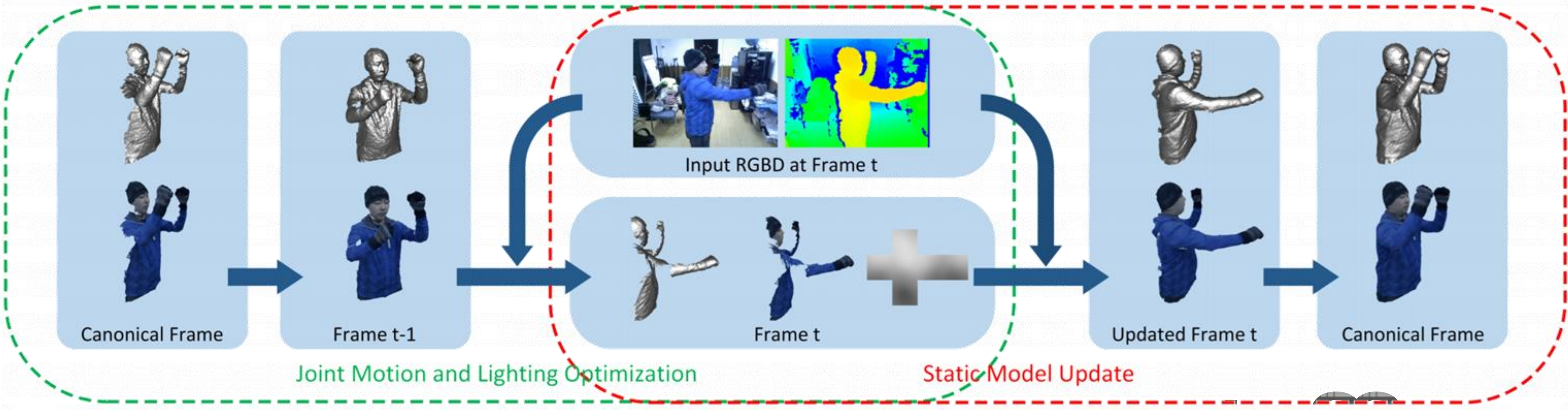
ECCV'16 [Innmann et al.]: VolumeDeform

Real-time Geometry, Albedo, and Motion Reconstruction

[Guo et al. 17]

Real-time Geometry, Albedo, and Motion Reconstruction

- Deformation graph (ED)
- Non-rigid ICP (dense depth + dense color)
- Illumination correction for color term (more later on lighting)



[Guo et al. 17]

Real-time Geometry, Albedo, and Motion Reconstruction

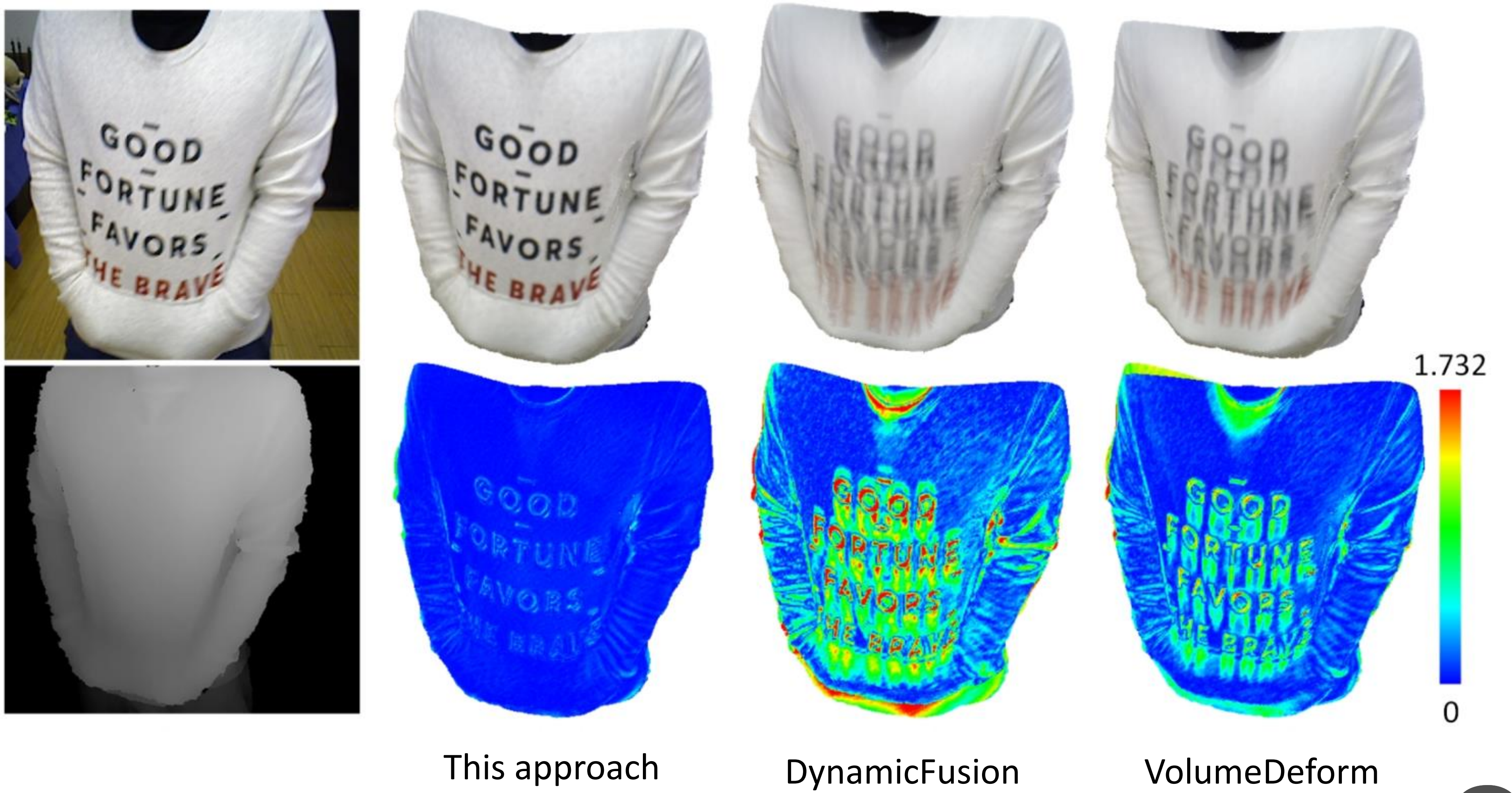
Real-time Geometry, Albedo and Motion Reconstruction
Using a Single RGBD Camera

Kaiwen Guo[1], Feng Xu[1], Tao Yu[1,2], Xiaoyang Liu[1], Qionghai Dai[1], Yebin Liu[1]
Tsinghua University[1]   Beihang University[2]

primary video
(this video contains audio)

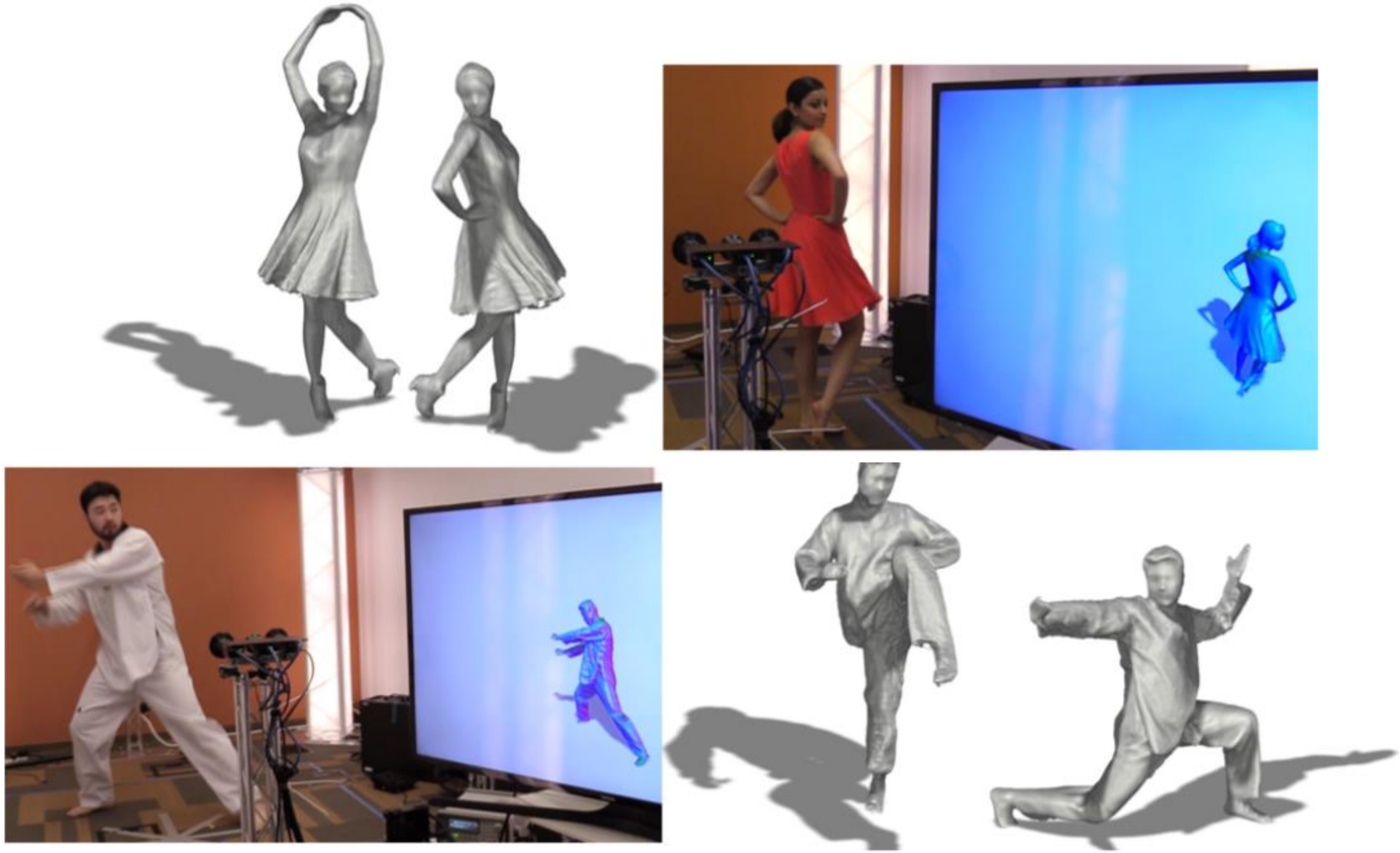
[Guo et al. 17]

Real-time Geometry, Albedo, and Motion Reconstruction



[Guo et al. 17]

Fusion4D / Holoportation



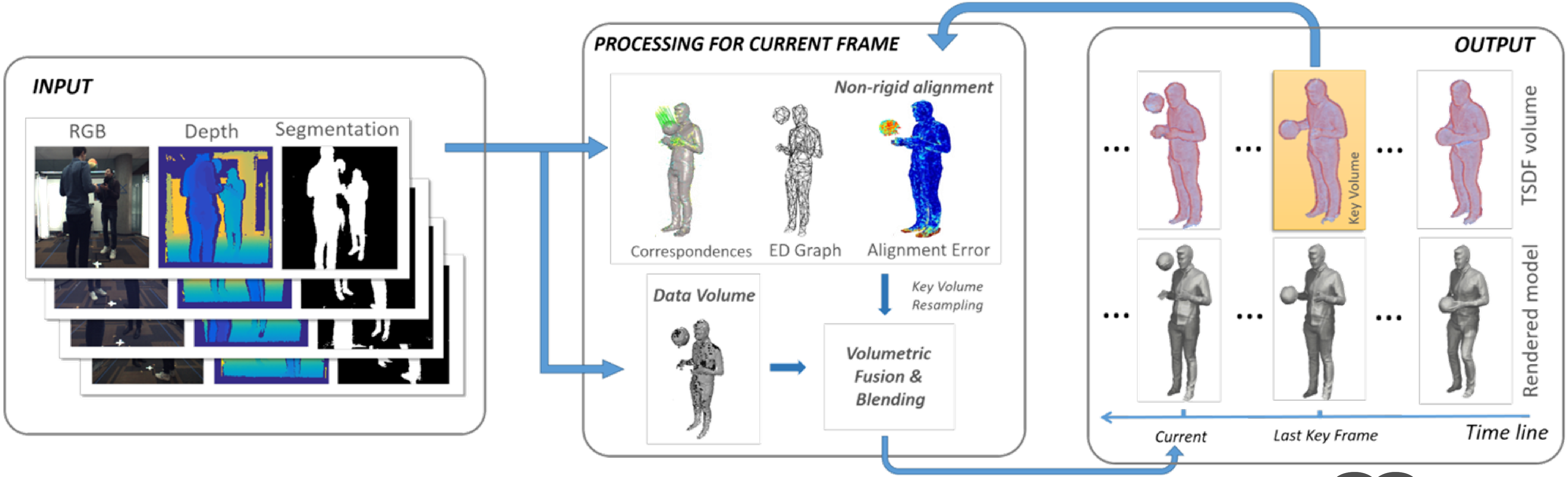
[Dou et al. 16]

Fusion4D / Holoportation

[Dou et al. 16]

Fusion4D / Holoportation

- 8 depth maps (16 IR cameras)
- Key volumes: multiple canonical poses



[Dou et al. 16]

Fusion4D / Holoportation

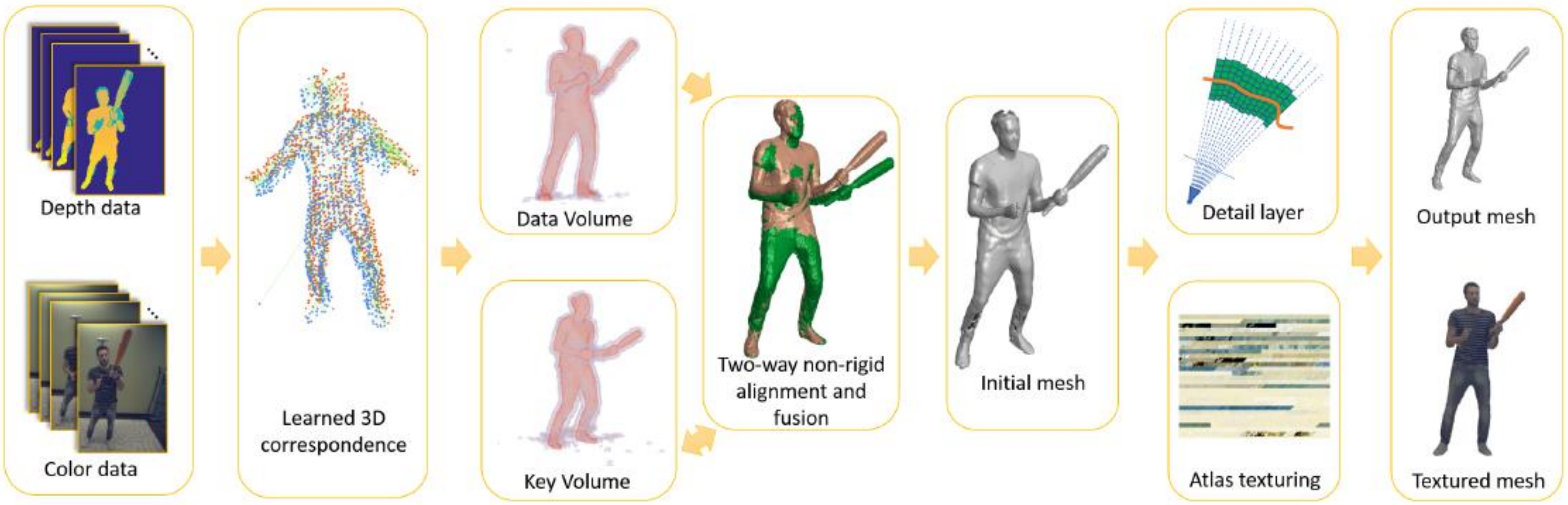
Motion2Fusion

[Dou et al. 17]

Motion2Fusion

[Dou et al. 17]

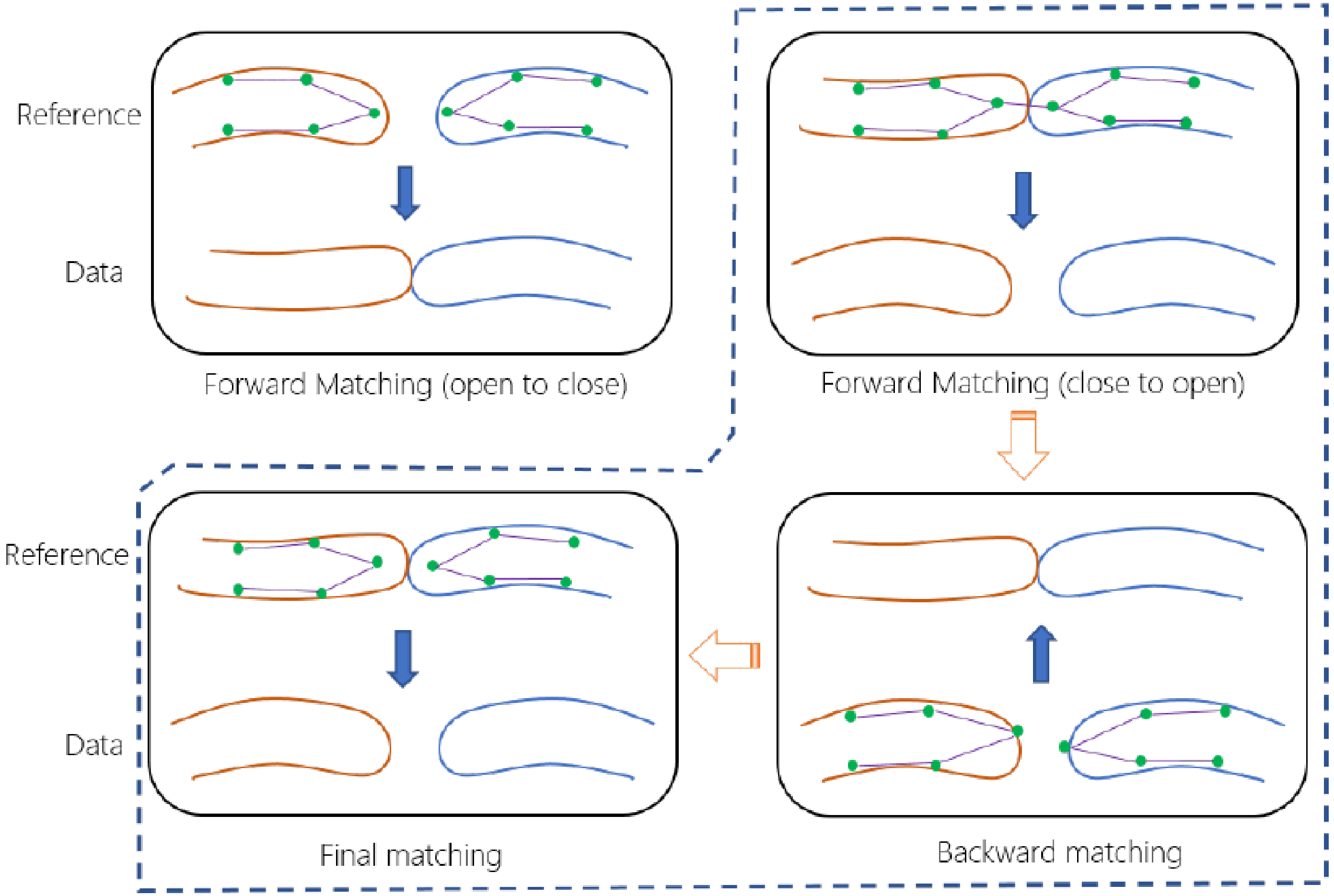
Motion2Fusion



- ED graph, key volumes, multiple depth cameras
- Much faster: 200 FPS => Faster motion tracking

[Dou et al. 17]

Motion2Fusion

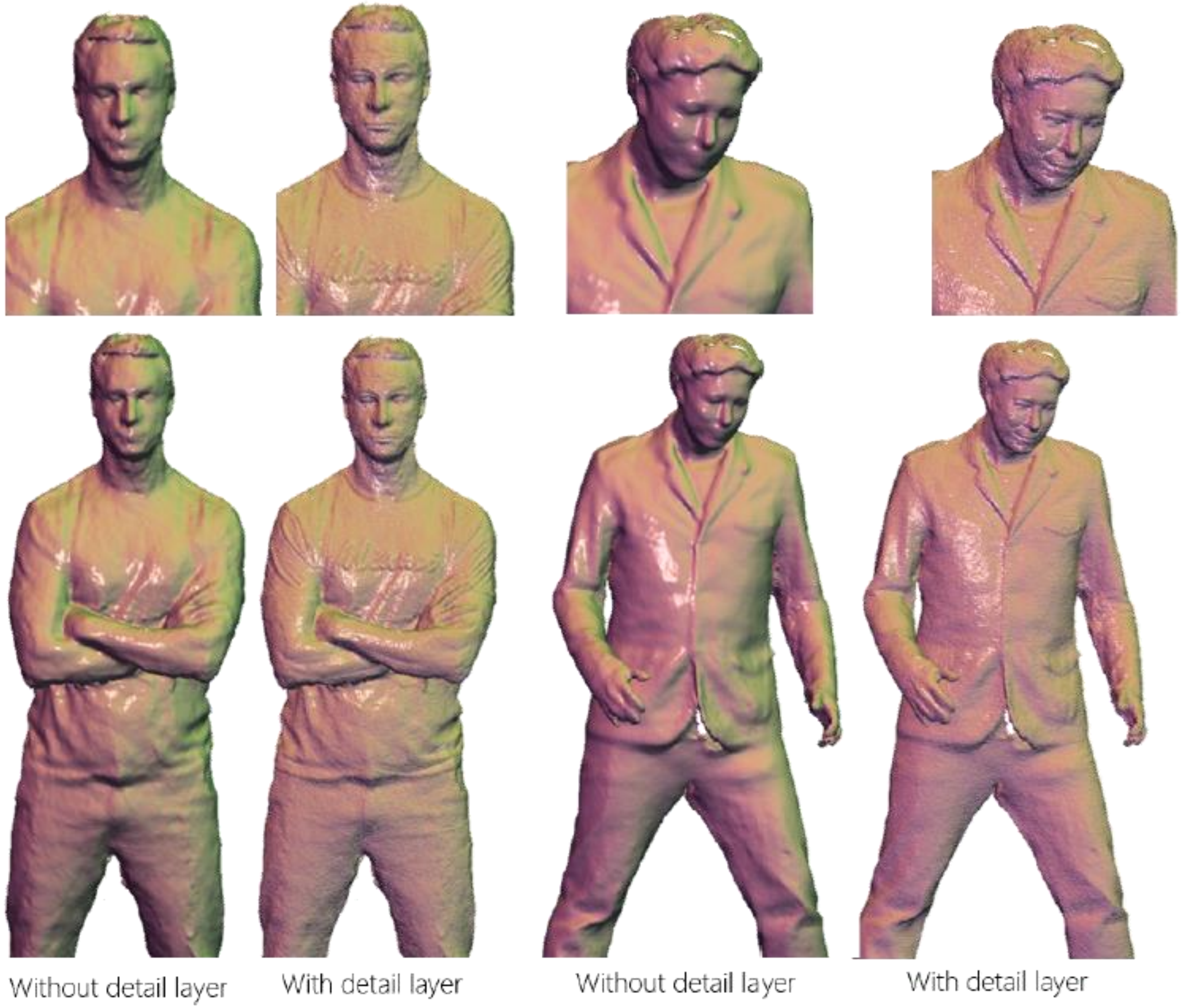


ED Graph based alignments fail surface splitting (from closed to open topology)

2 Way Backward and forward nonrigid alignment strategy

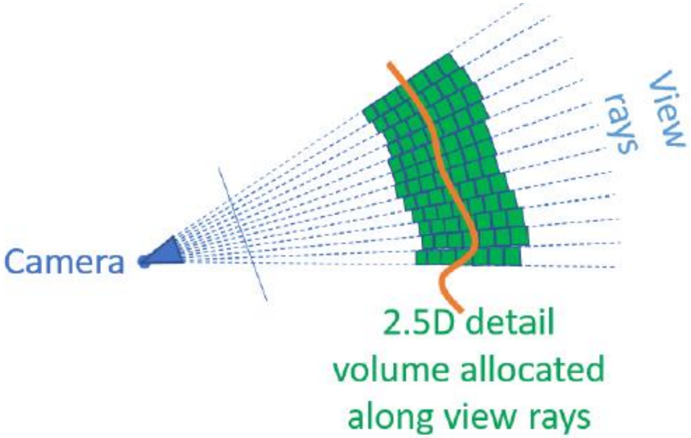
[Dou et al. 17]

Motion2Fusion



Detail layer

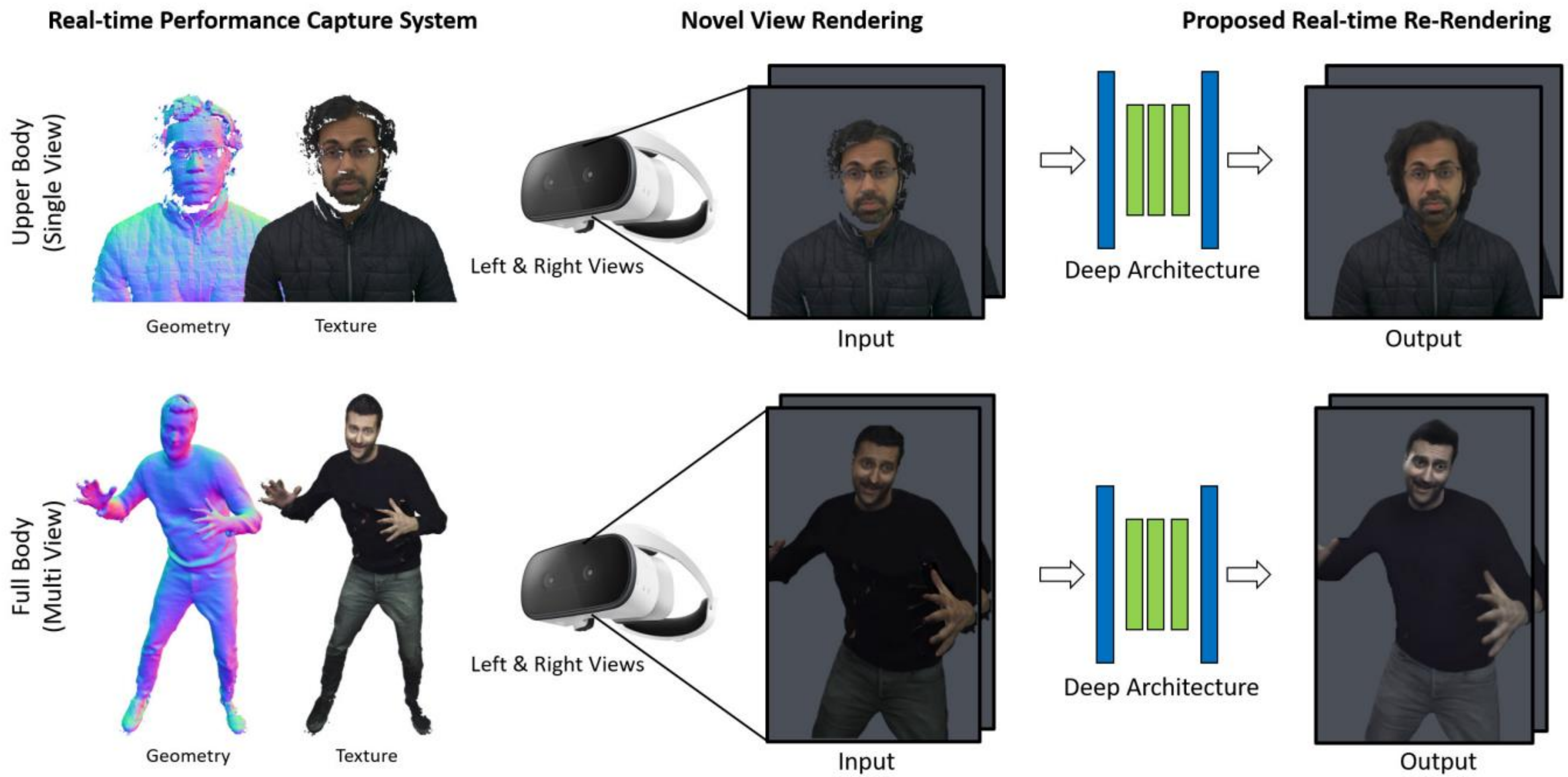
2.5D volume to capture the high frequency details that can be lost during the matching process



resolution of regular voxel grid: 5mm
resolution along the viewing ray : 2mm

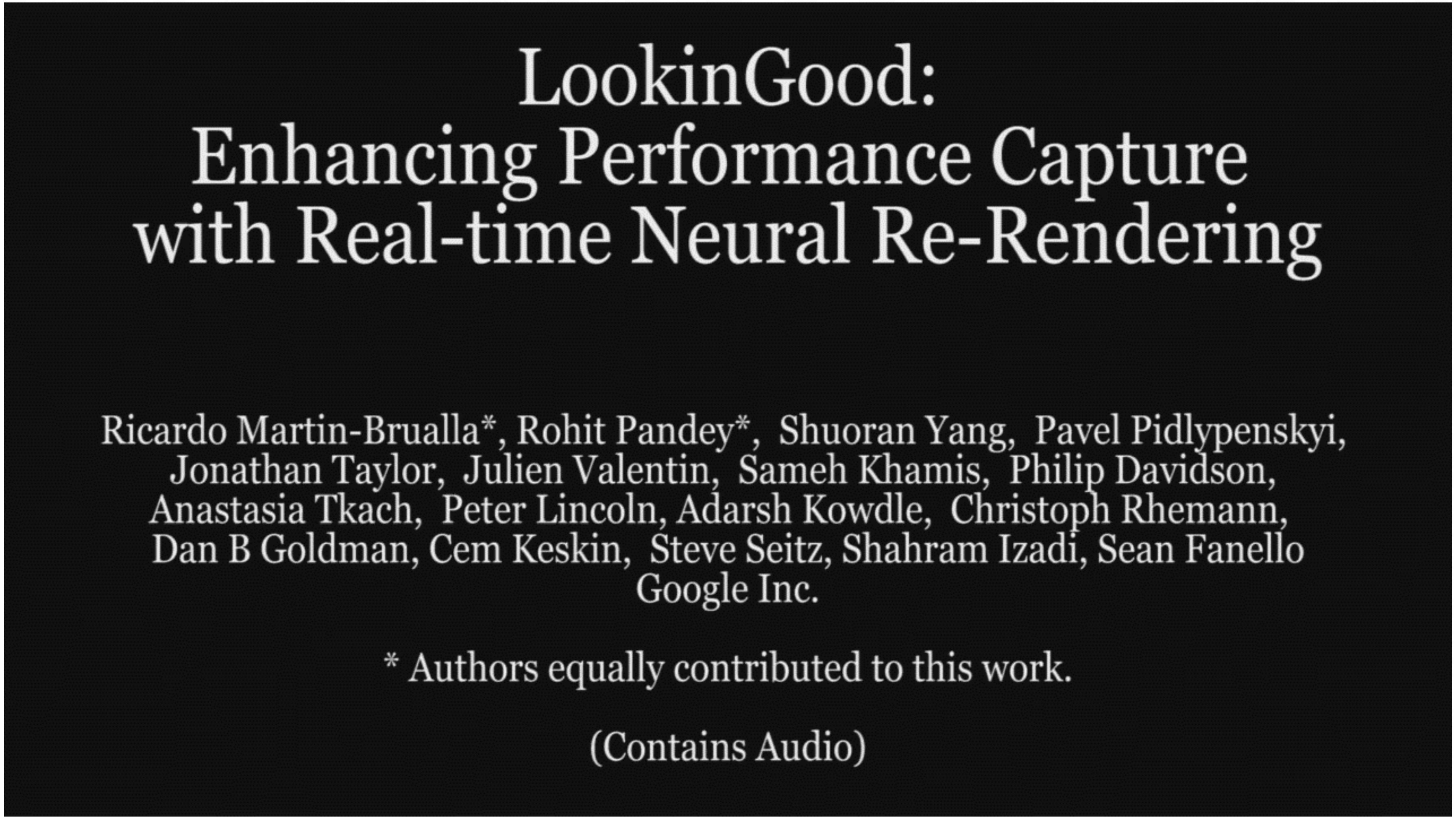
[Dou et al. 17]

LookinGood



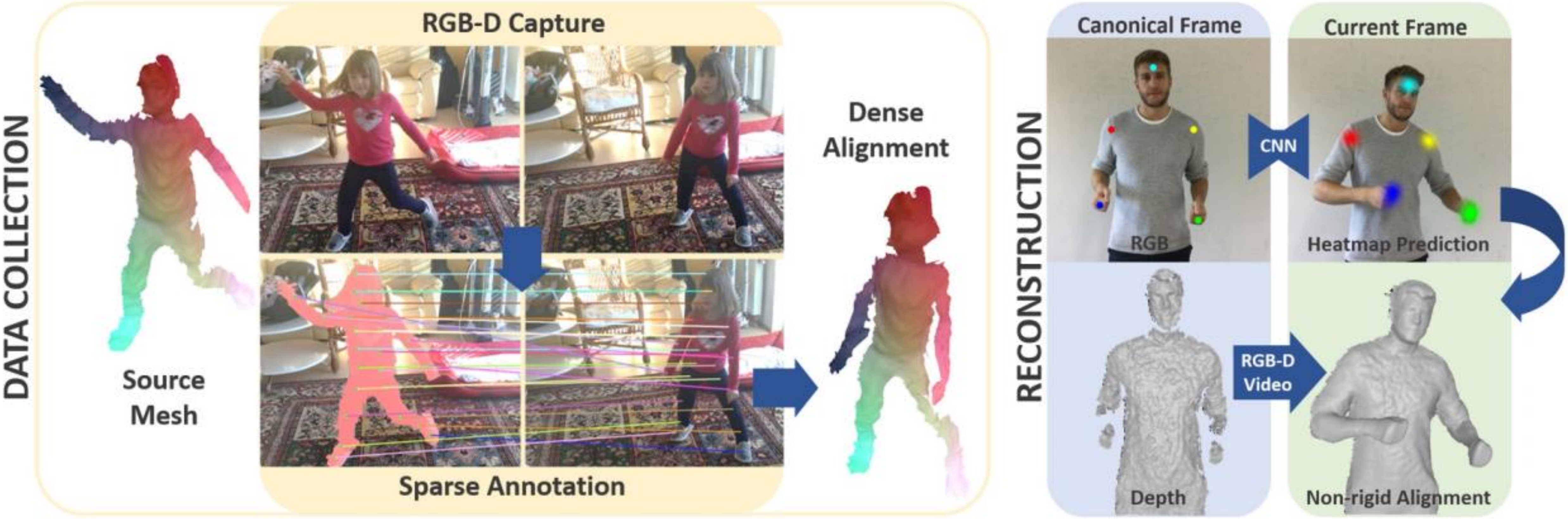
[Martin-Brualla et al. 18]
https://arxiv.org/abs/1811.05029

LookinGood



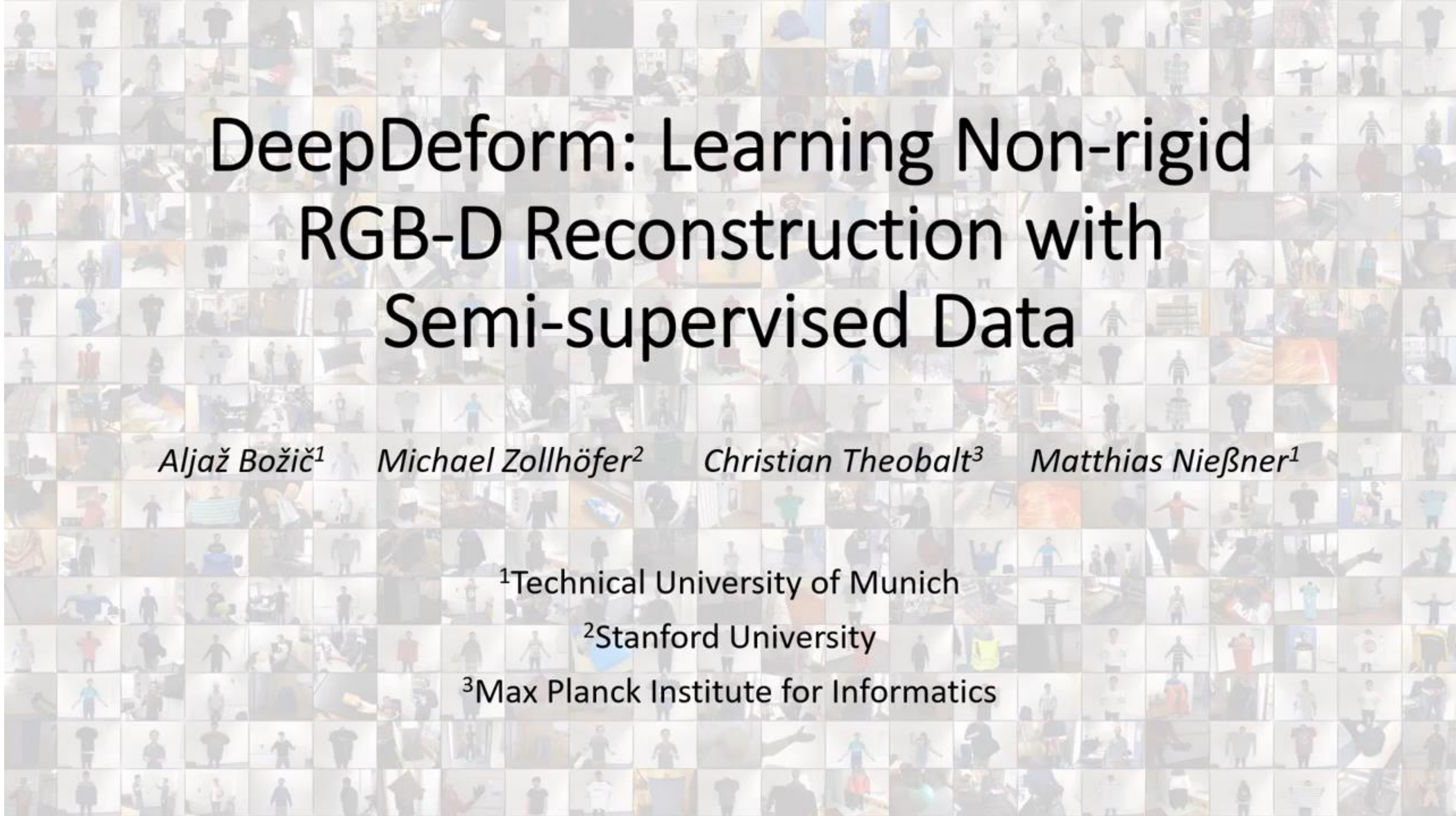
[Martin-Brualla et al. 18]
https://arxiv.org/abs/1811.05029

DeepDeform – Correspondence Learning



[Bozic et al. 19]
https://arxiv.org/abs/1912.04302

DeepDeform – Correspondence Learning



[Bozic et al. 19]
https://arxiv.org/abs/1912.04302

Summary: Non-rigid Tracking vs Non-rigid Reconstruction

- Tracking
  - Several degrees of freedom per vertex
  - Mitigates drift via regularization w.r.t. to base mesh
  - W/o correspondences → would snap back to base mesh

- Non-rigid Reconstruction
  - Several degrees of freedom per vertex
  - Accumulates drift: if tracking is wrong, surface integration is wrong
    - Can be reduced by resetting to new key volume

Summary: Non-rigids / Dynamic 3D Capture

- It's a challenging problem!
  – Often under-constrained
  – In particular, from a single depth (or even RGB) camera
  – Btw. RGB-only tracking does (somewhat work)
  – Joint reconstruction with RGB-only not so much
    - Something like non-rigid bundle adjustment / non-rigid multi-view stereo

- Need to hallucinate missing data through regularization
  – I.e., where we have no correspondences for deformation

Summary: Dynamic 3D Capture

Need more regularization with fewer DoF to make it practical!

Domain-Specific

- Human body
- Hands
- Faces

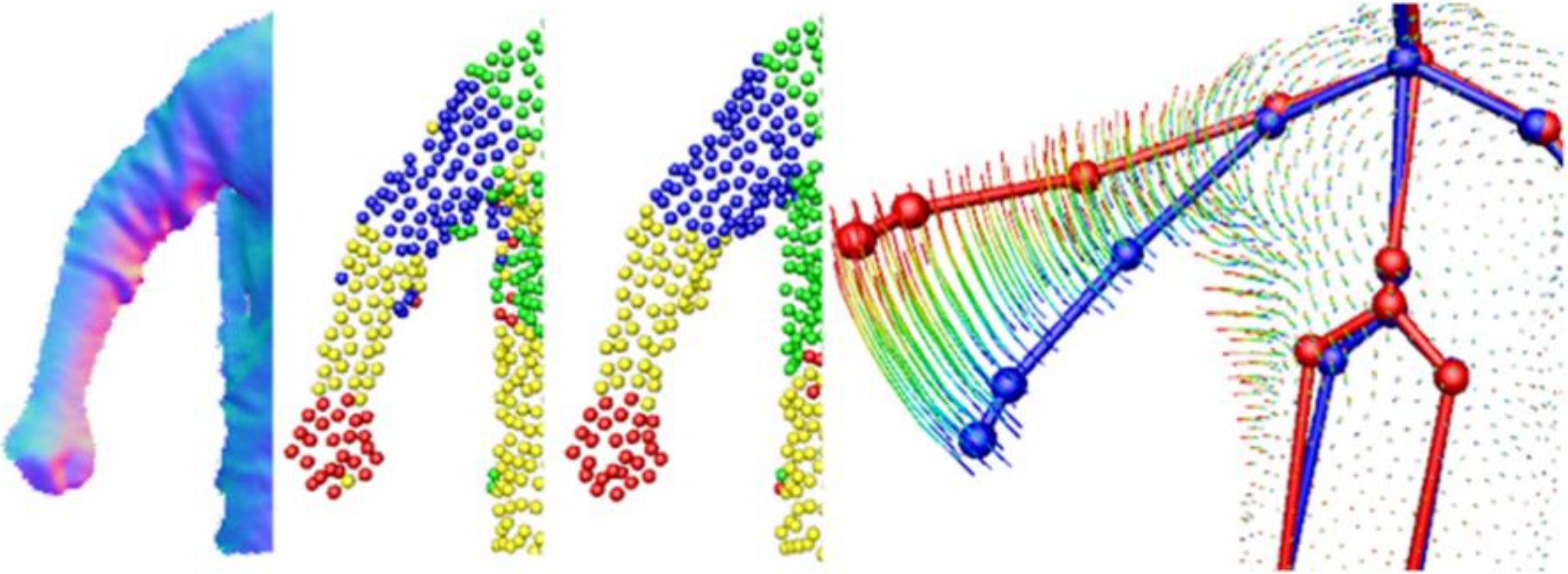
Free-form Reconstruction

- Joint reconstruction and tracking

Free-form Tracking

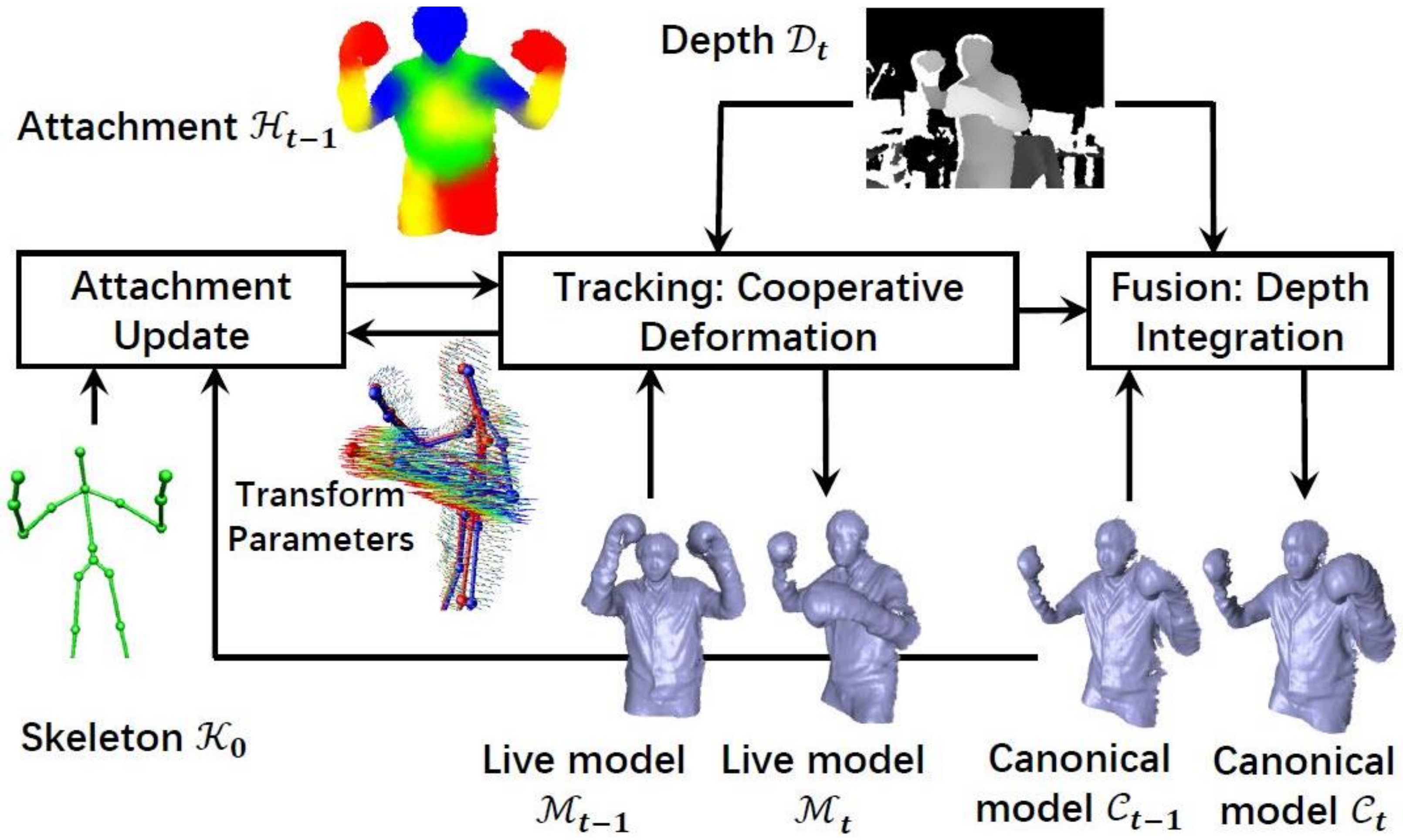
- Offline template reconstruction
- Online template tracking

BodyFusion



[Yu et al. 2017]

BodyFusion



[Yu et al. 2017]

BodyFusion

ICCV 2017

BodyFusion

Real-time Capture of Human Motion and Surface Geometry Using a Single Depth Camera

Tao Yu[1,2], Kaiwen Guo[2], Feng Xu[2], Yuan Dong[2], Zhaoqi Su[2], Jianhui Zhao[1], Jianguo Li[3], Qionghai Dai[2], Yebin Liu[2]

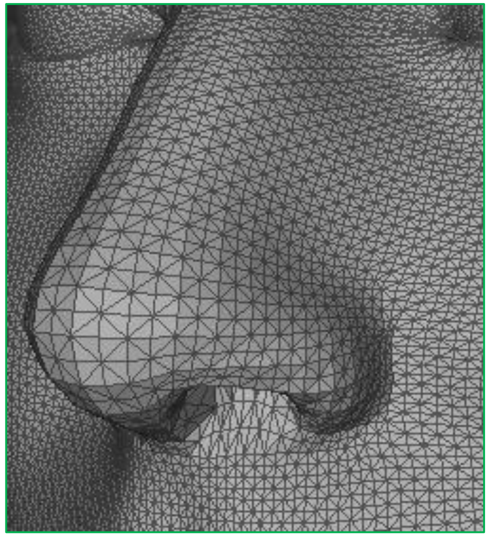
Beihang University, Beijing, China[1]
Tsinghua University, Beijing, China[2]
Intel Labs China, Beijing, China[3]

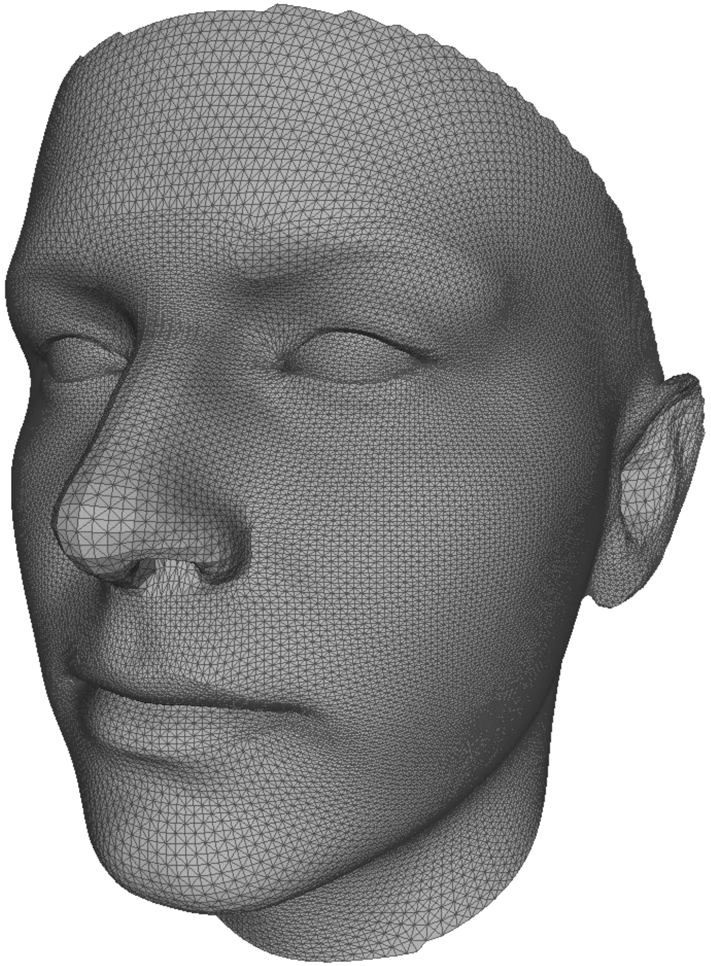
[Yu et al. 2017]

Next Lecture: Domain-specific Tracking & Reconstruction

- Small number of parameters change larger number of vertices
  - E.g., basis functions + skinning

- For Faces, Hands, Bodies





Administrative

- Reading Homework:
  - [Newcombe et al. 15] DynamicFusion: Reconstruction and Tracking of Non-Rigid Scenes in Real-Time
    https://grail.cs.washington.edu/projects/dynamicfusion/papers/DynamicFusion.pdf
  - DeepDeform: Learning Non-rigid RGB-D Reconstruction with Semi-supervised Data
    https://arxiv.org/pdf/1912.04302.pdf
- Next week(s):
  - Domain-specific Tracking & Reconstruction (Faces, Bodies, Hands, etc.)

Administrative

See you next week!