Real-time Dynamic Scene Reconstruction Using a Single RGBD Camera
基于单个深度相机的动态场景实时三维重建

于涛 北京航空航天大学

GAMES Webniar 2017.7.27

Research Background

Multiple sensors to get 2D information

KinectFusion: Real-time static 3D reconstruction

dynamic 3D reconstruction

Research Background

Skype/FaceTime

Automultiscopic Display (USC ICT Graphics Lab)

Holoportation (Microsoft Research)

Two Projects

- Real-time Geometry, Albedo and Motion Reconstruction Using a Single RGB-D Camera (SIGGRAPH17)
- BodyFusion: Real-time Capture of Human Motion and Surface Geometry Using a Single Depth Camera (ICCV17)



Real-Time Geometry, Albedo, and Motion Reconstruction Using a Single RGB-D Camera

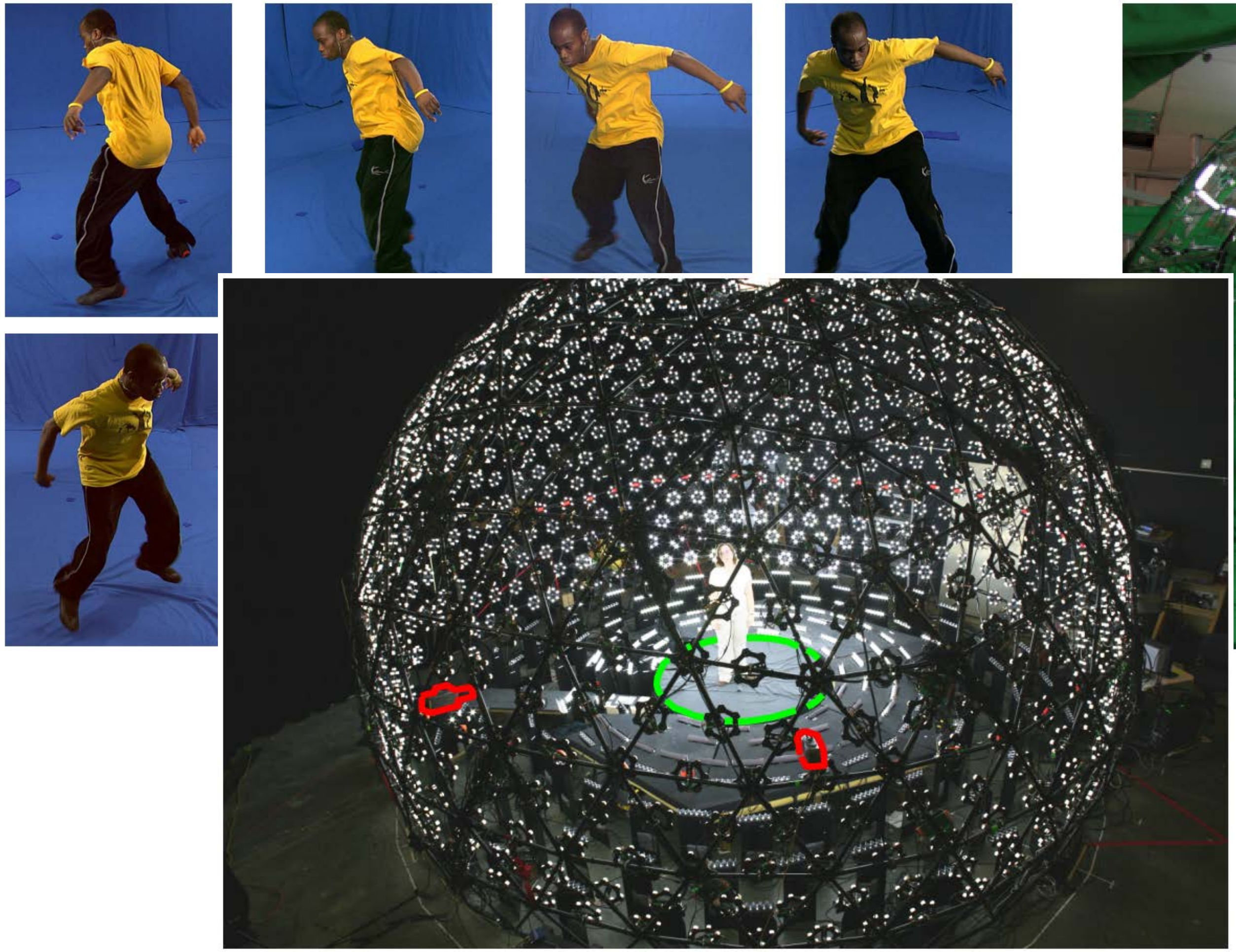
Kaiwen Guo[1], Feng Xu[1], **Tao Yu[1,2]**, Xiaoyang Liu[1], Qionghai Dai[1], Yebin Liu[1]

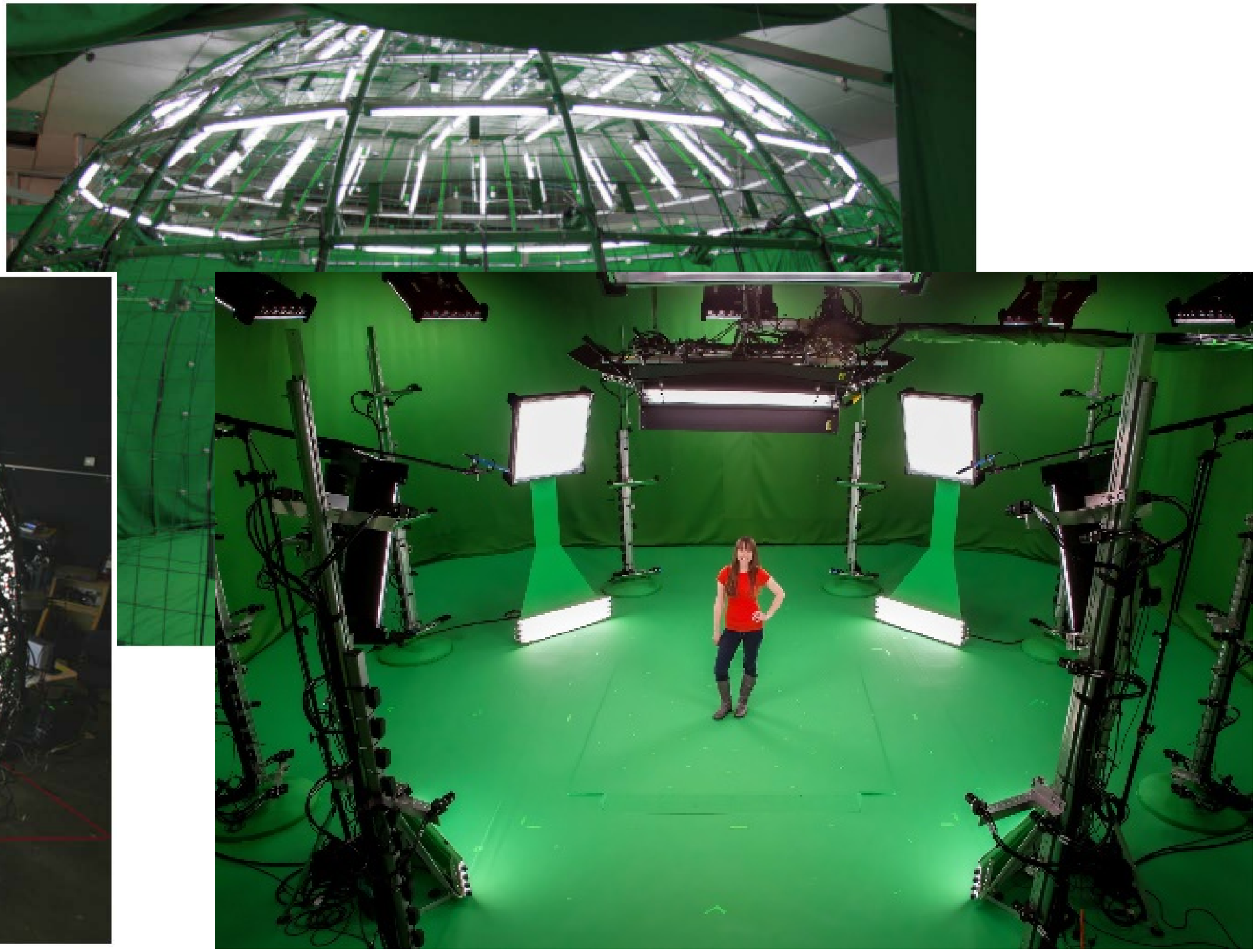
Tsinghua University[1] Beihang University[2]

清華大學 Tsinghua University

30 JULY–3 AUGUST Los Angeles
SIGGRAPH2017

Offline Multiview Volumetric Performance Capture





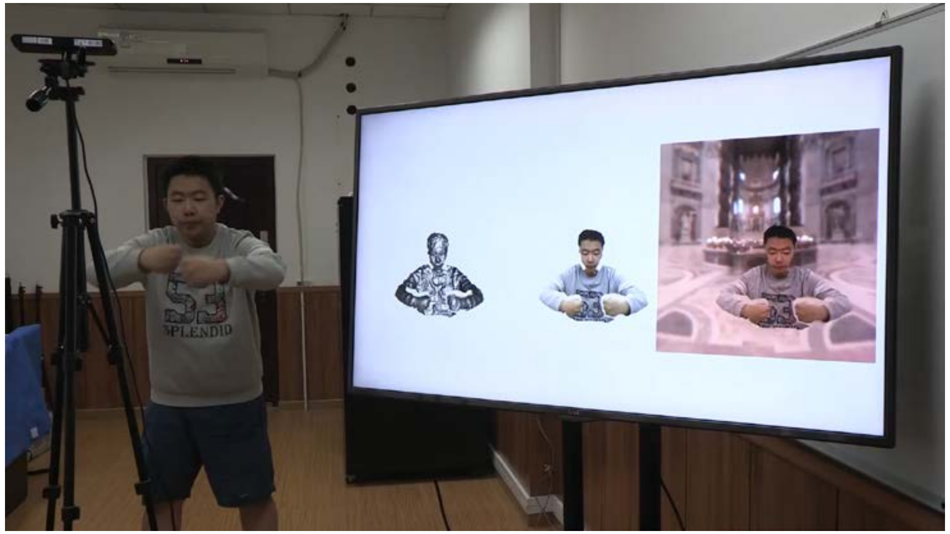
[Starck and Hilton, 2007], [Liu et al., 2009]

[Vlasic et al. 2009], [Debevec, Light Stage], [Collet et al. 2015]

Real-time Multiview Volumetric Performance Capture

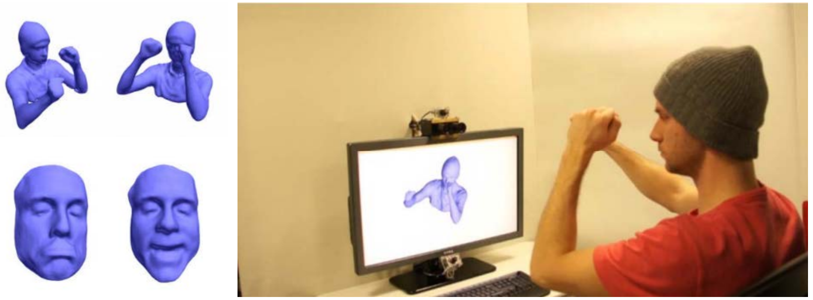
[Fusion4D, Dou et al. 2016], [Holoportation, Orts-Escolano et al. 2016]

Real-time Single-view Volumetric Capture



Related Works – Single-view Performance Capture

Single-view 3D Scanning
[Dou et al. CVPR15]



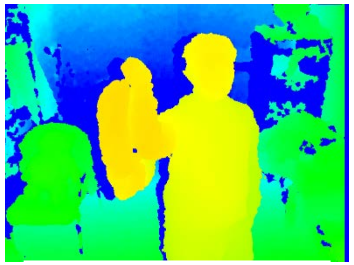
Real-time Templates Tracking
[Zollhofer et al. SIGGRAPH14]

DynamicFusion
[Newcombe et al. CVPR15]

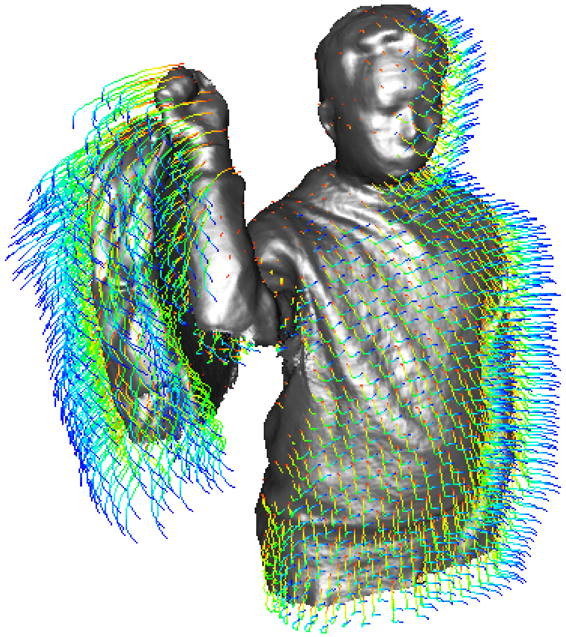
VolumeDeform
[Innmann et al. ECCV16]

System Overview

Color

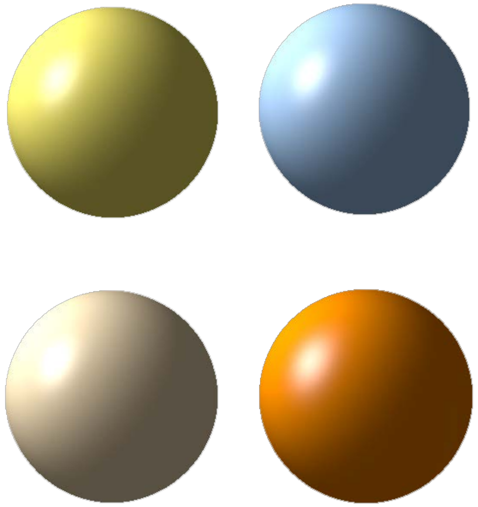


Depth



Motion

Geometry



Albedo

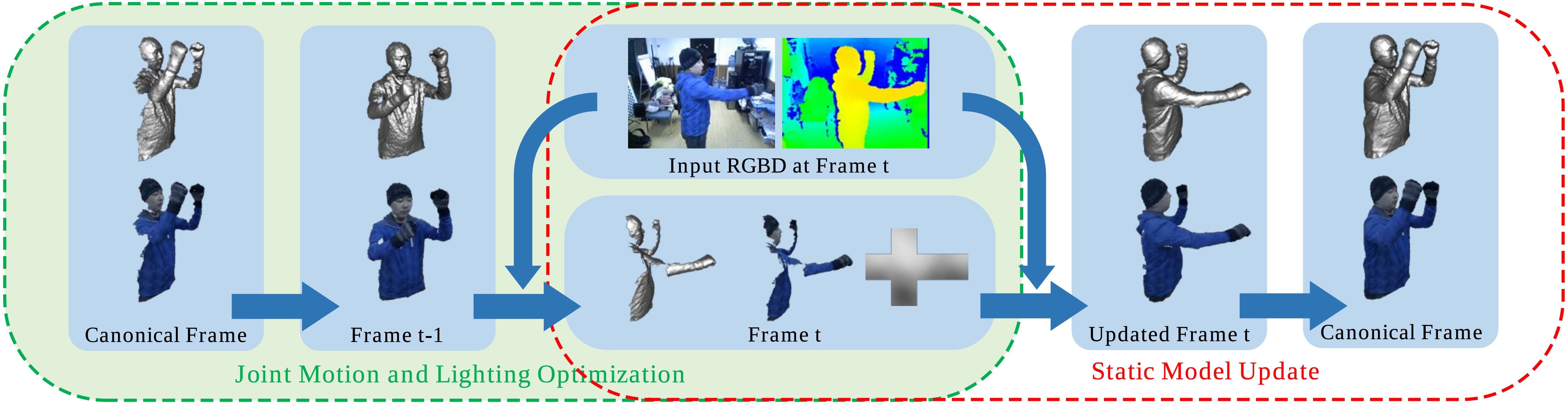
Input

Output

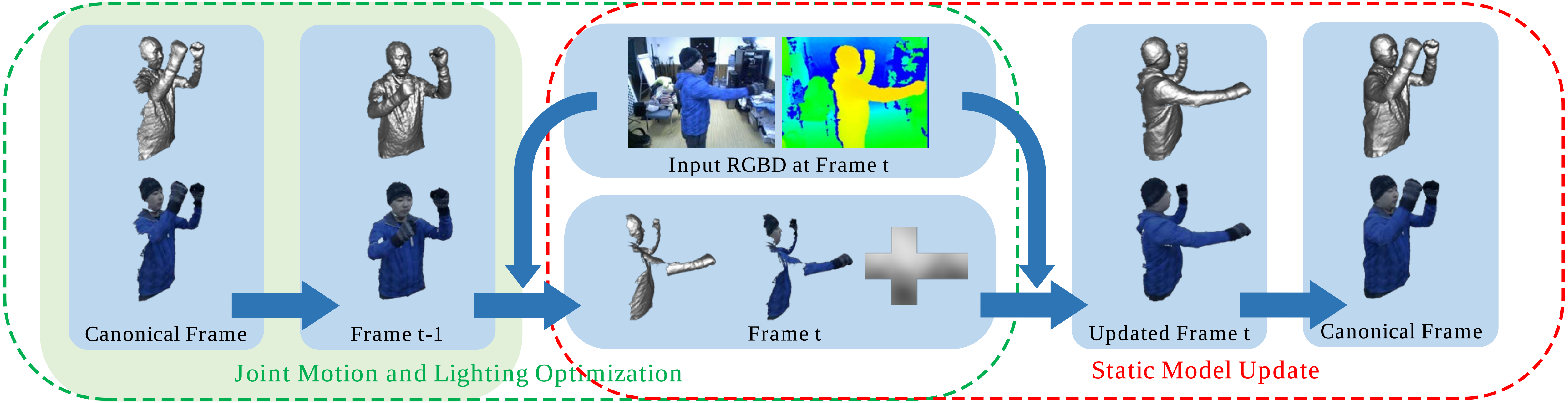
Key Idea


Good Albedo Results
Good Correspondences
Good Tracking Results

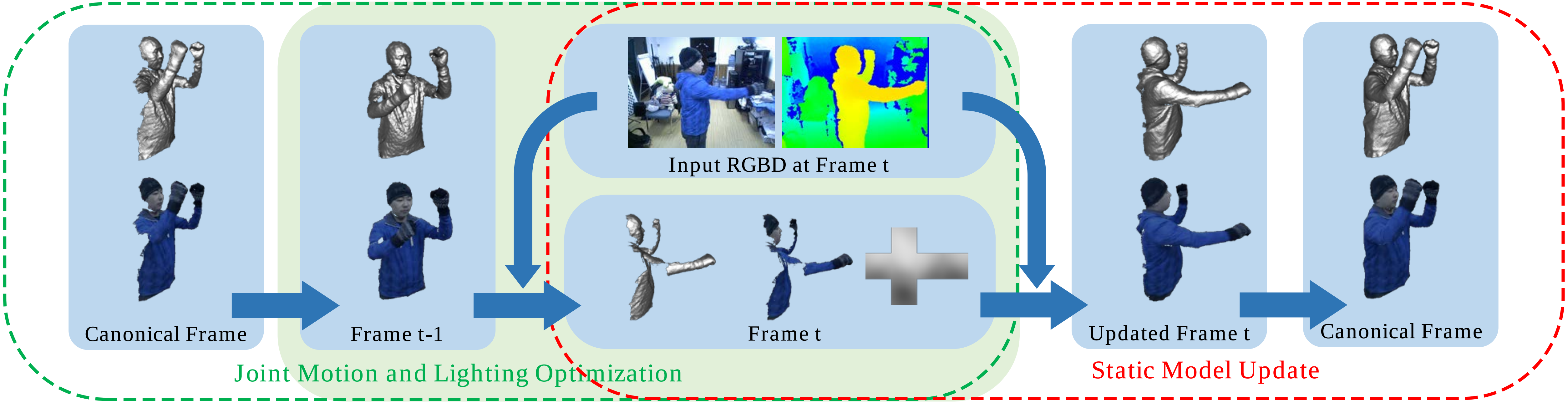
Algorithm Overview



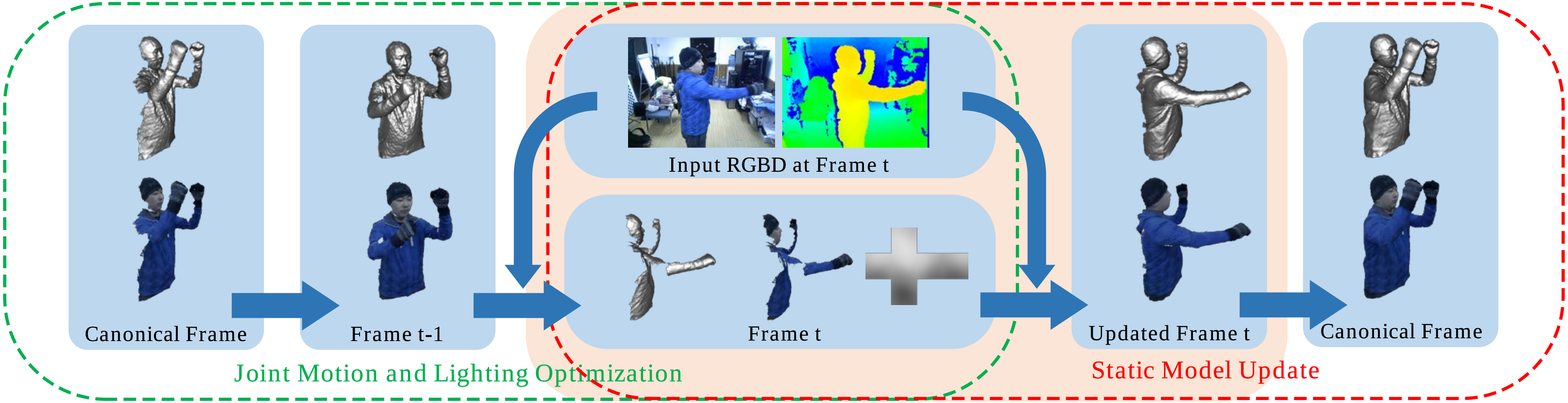
Algorithm Overview



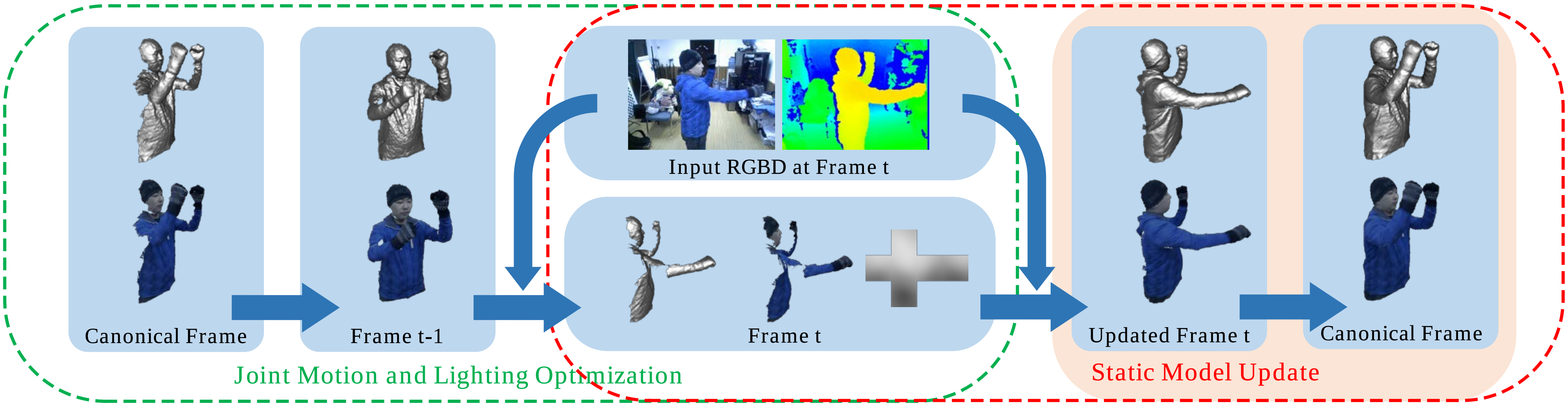
Algorithm Overview



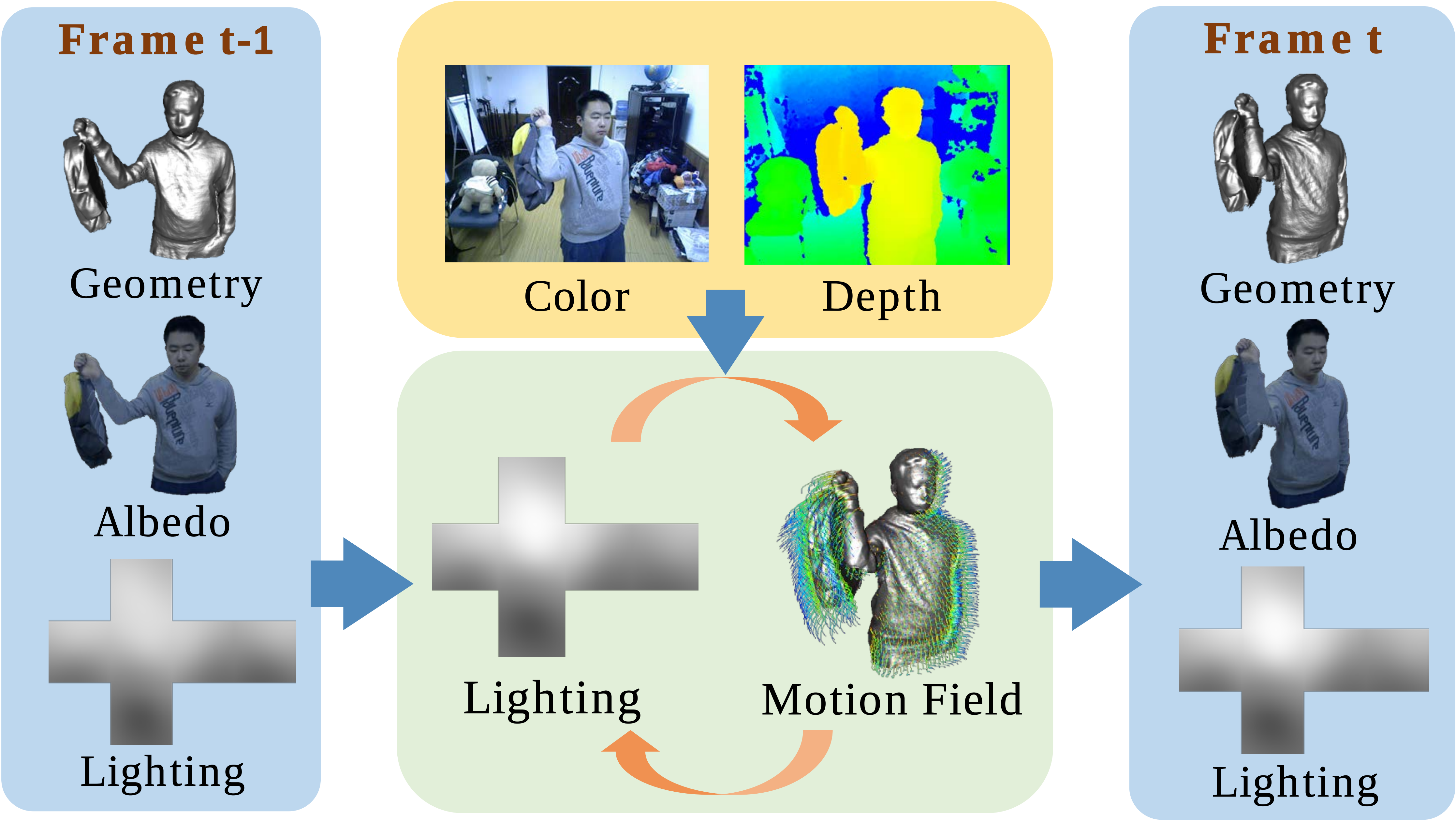
Algorithm Overview



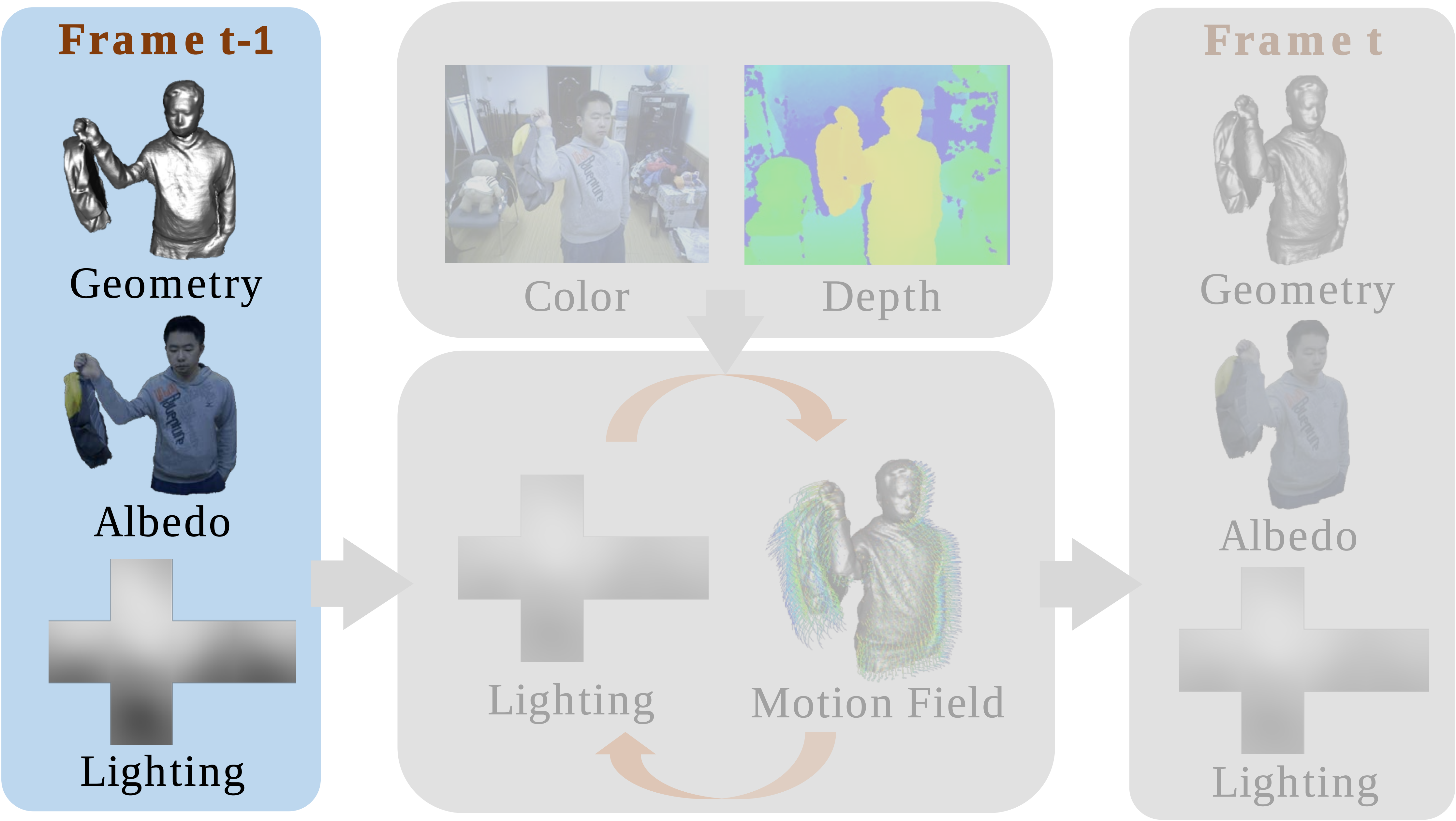
Algorithm Overview



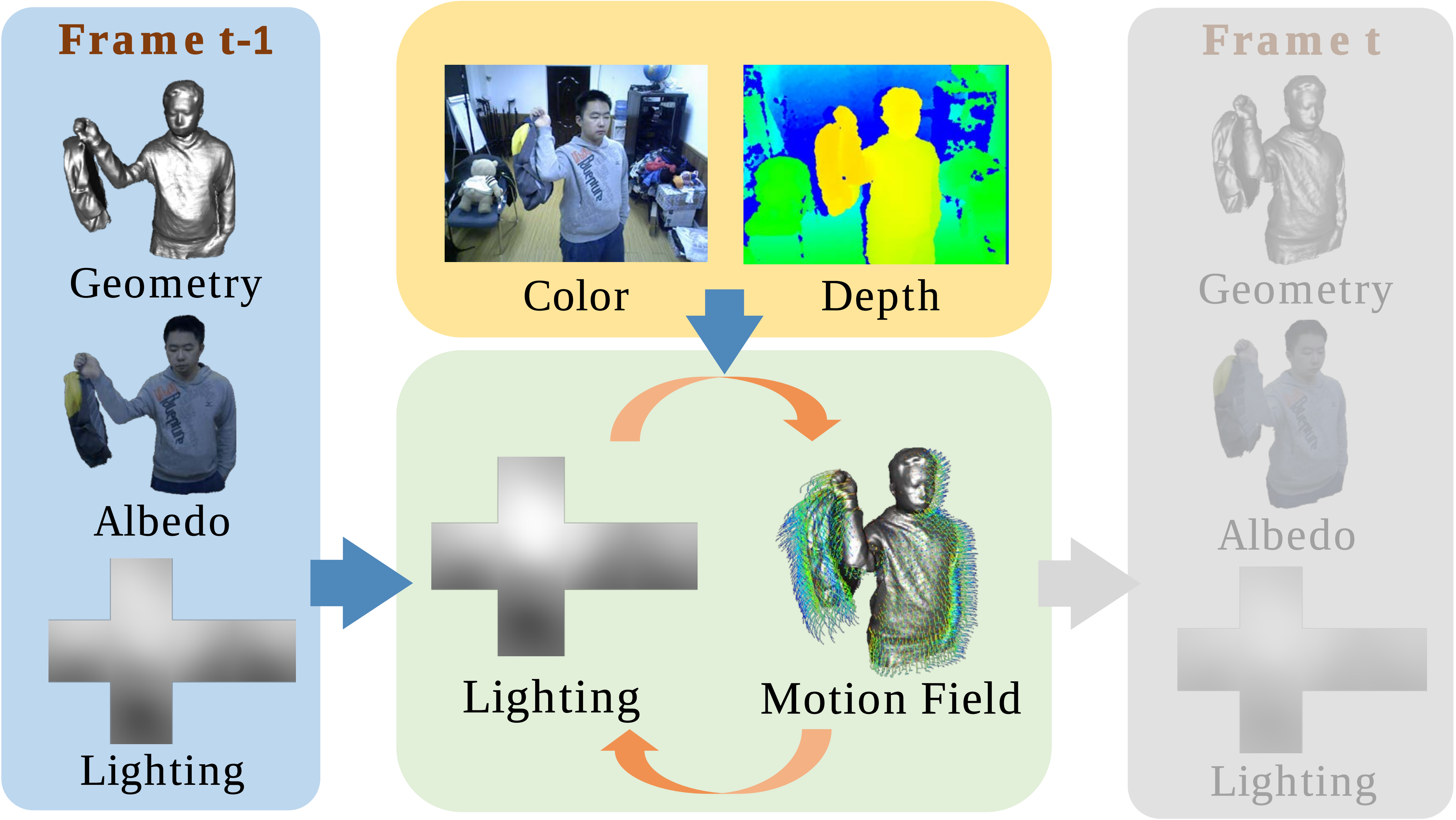
Joint Optimization of Motion and Illumination



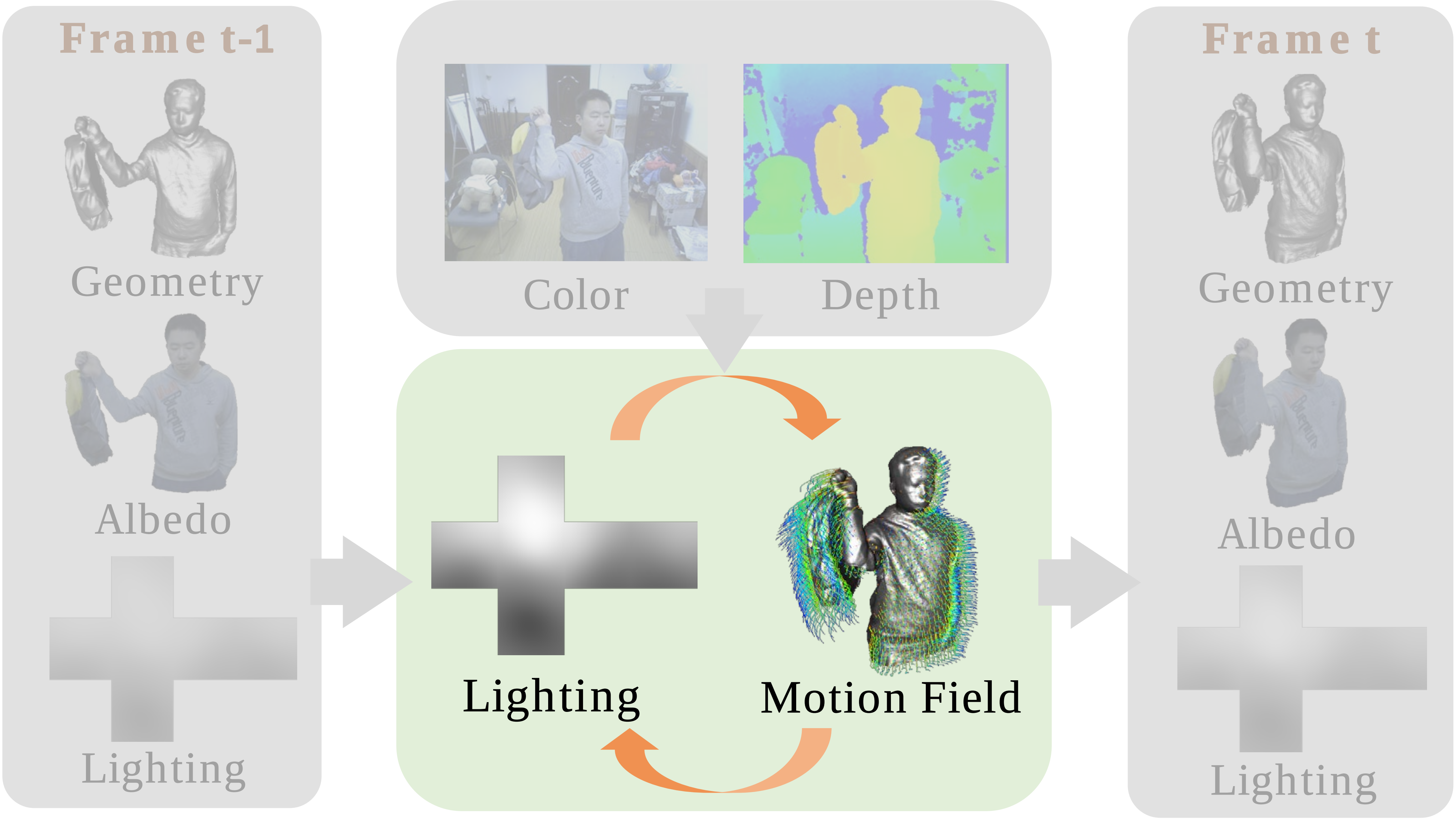
Joint Optimization of Motion and Illumination



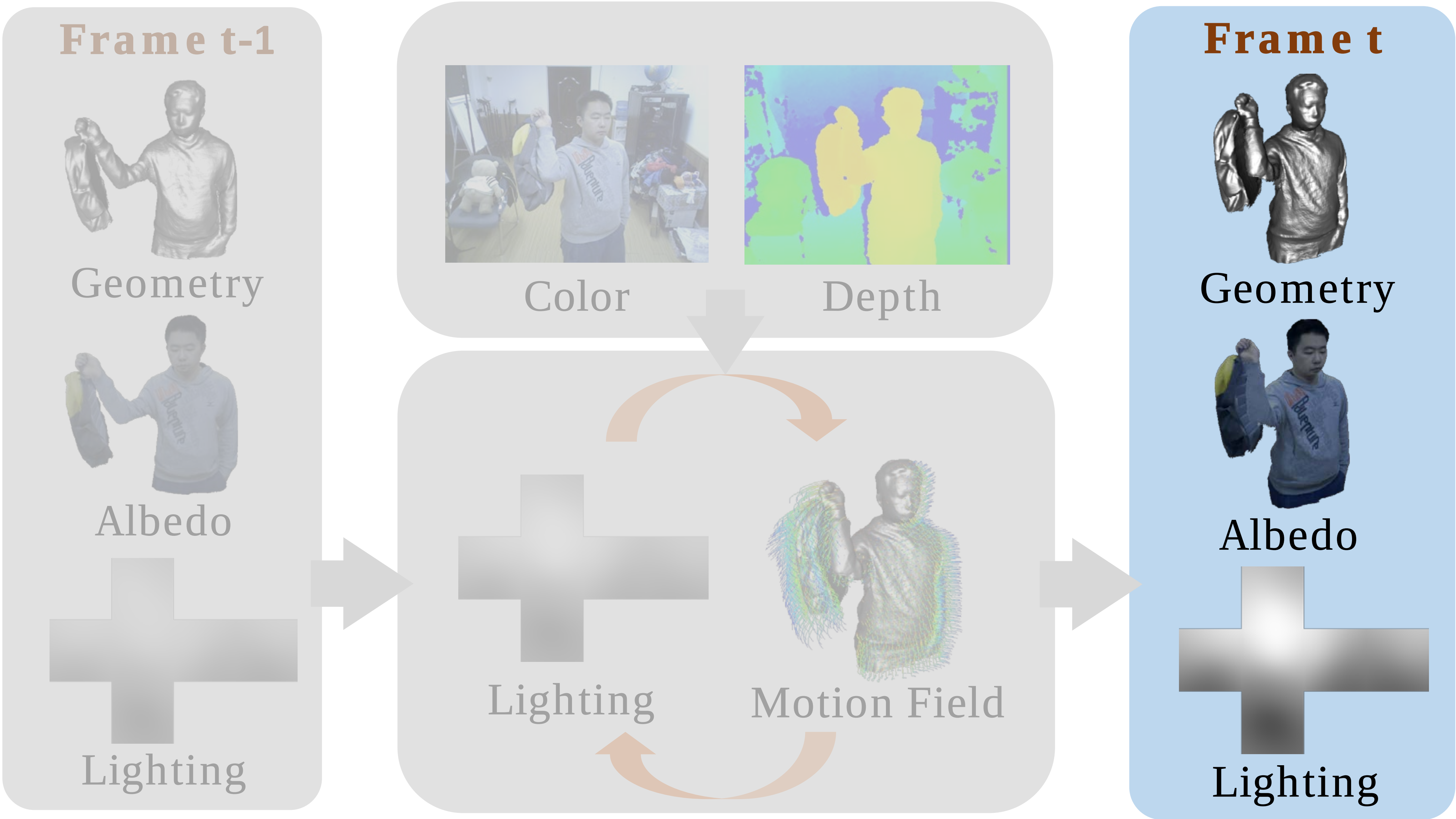
Joint Optimization of Motion and Illumination



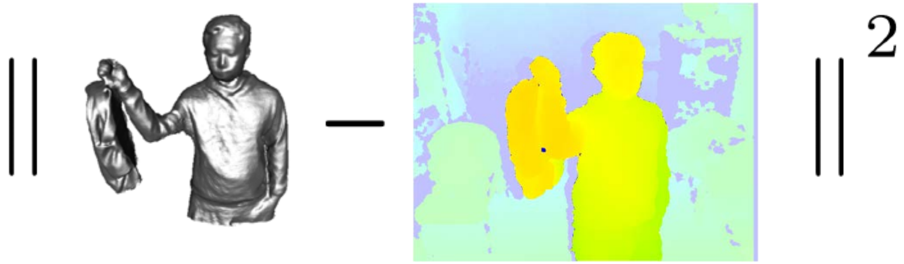
Joint Optimization of Motion and Illumination



Joint Optimization of Motion and Illumination

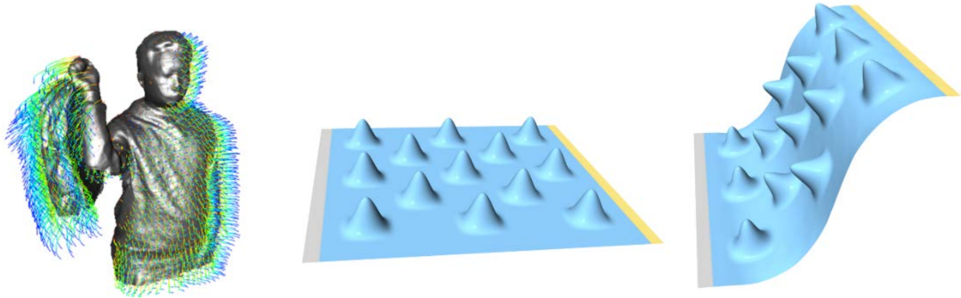


Joint Optimization of Motion and Illumination



$$E_{\text{depth}}(\mathcal{W}^t) = \sum_{(\mathbf{v},\mathbf{u}^t)\in\mathcal{P}} (\mathbf{n}_{\mathbf{u}^t}^{\text{T}}(\mathbf{v}' - \mathbf{u}^t))^2$$

$\mathcal{E}_{depth}$

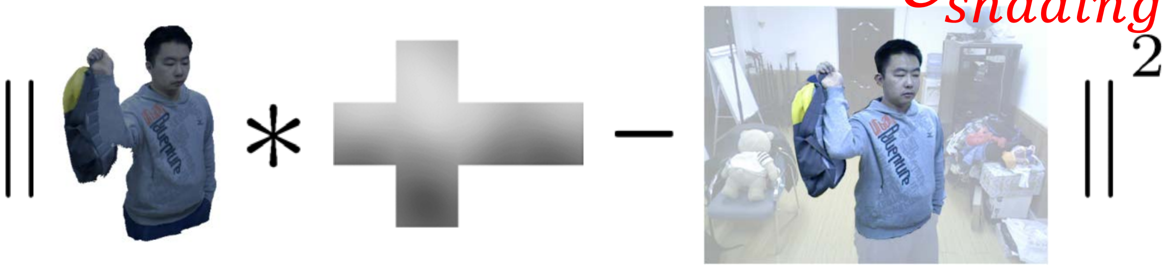


$$E_{\text{mreg}}(\mathcal{W}^t) = \sum_{j\in\mathcal{G}^t} \sum_{i\in\mathcal{N}_j} \|\mathbf{T}_j^t \mathbf{p}_j - \mathbf{T}_i^t \mathbf{p}_j\|_2^2$$

$\mathcal{E}_{arap}$

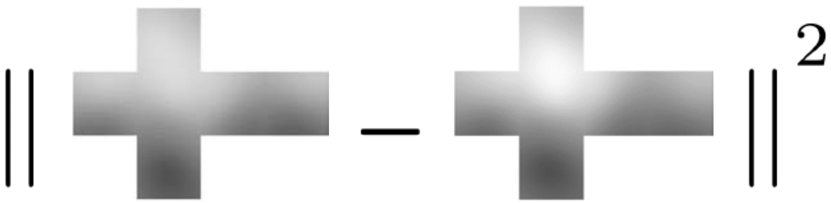
[Sumner et al. 2007]

$\mathcal{E}_{shading}$



$$E_{\text{shading}}(\mathcal{W}^t) = \sum_{\mathbf{v}\in\mathcal{V}_C^t} \|C^t(M_c(\mathbf{v}')) - B^t(\mathbf{v}')\|_2^2$$

$\mathcal{E}_{lreg}$



(a) Previous Lighting

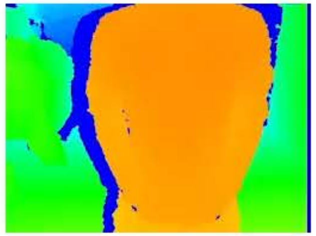
(b) Current Lighting

$$E_{\text{lreg}}(\mathcal{L}^t) = \sum_{i=1}^{b^2} \|\ell_i^{t-1} - \ell_i^t\|_2^2$$

Shading Based Motion Tracking

observation | our results | [Newcombe et al. 2015] | [Innmann et al. 2016]

Shading Based Motion Tracking

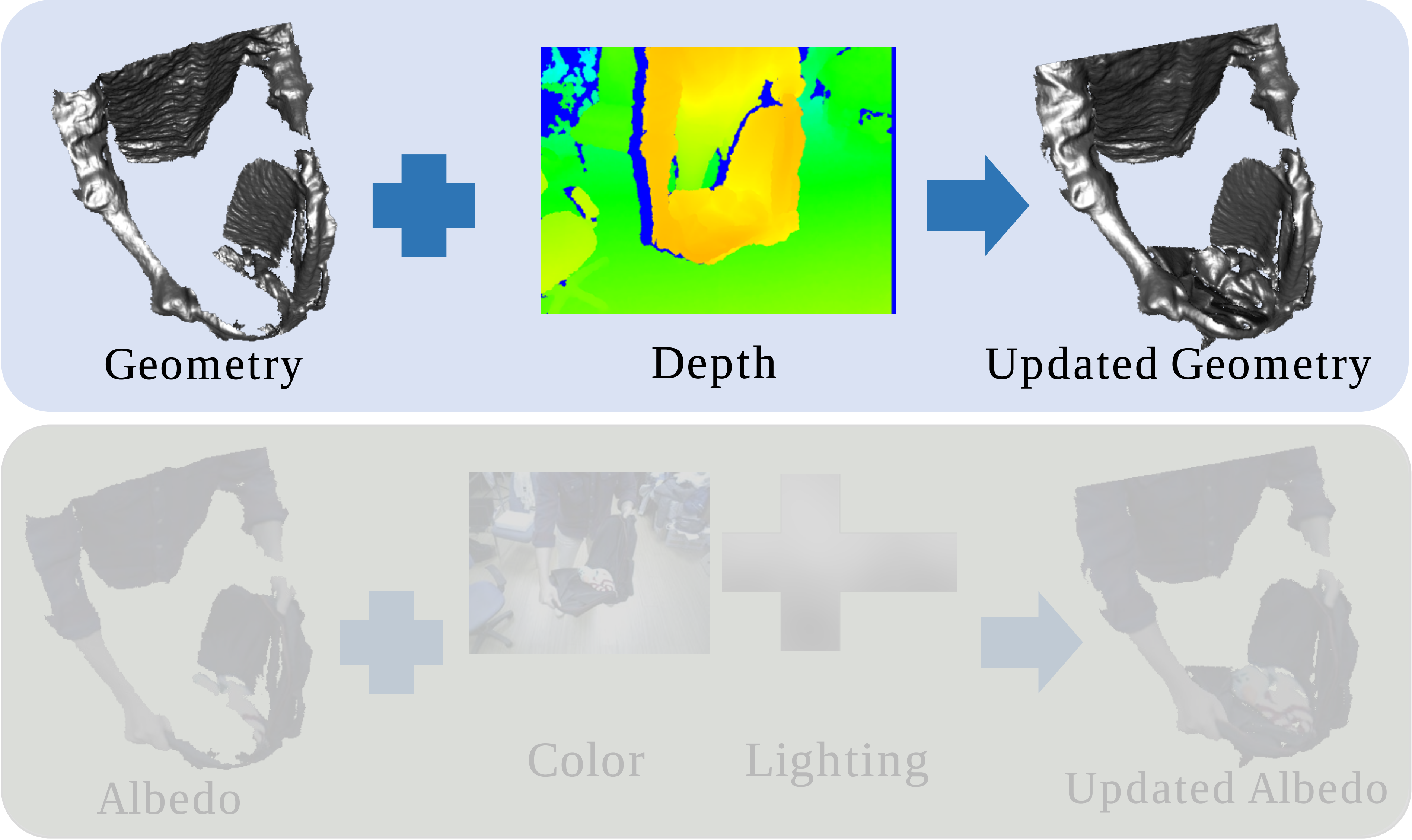


our result

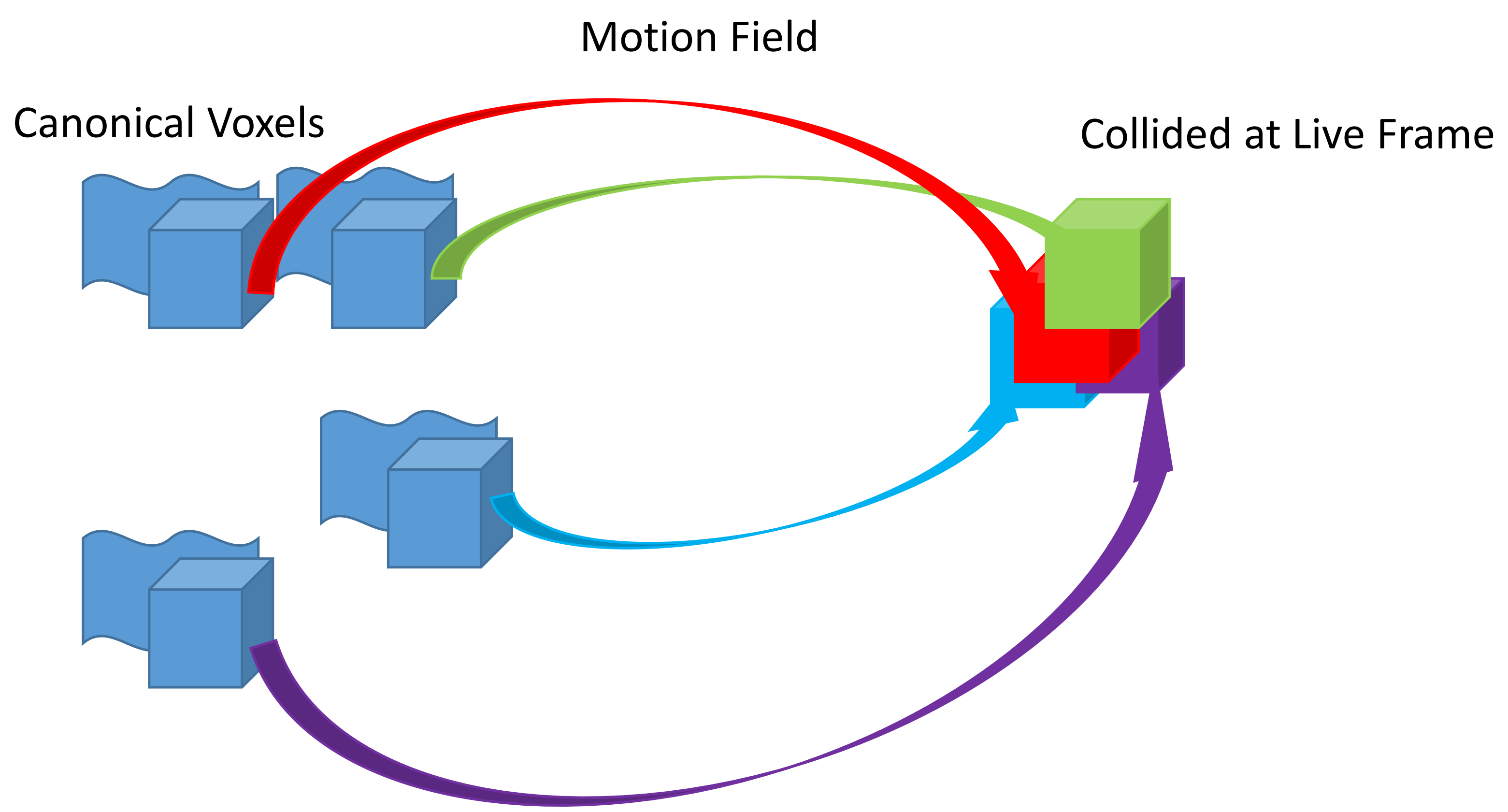
[Newcombe et al. 2015]
(our implementation)

[Innmann et al. 2016]
(our implementation)

Update Geometry and Albedo Model

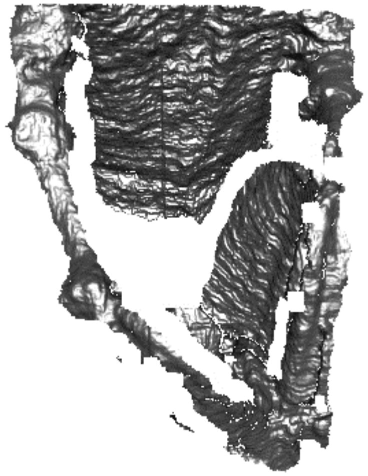


Resolving Voxel Collision



Resolving Voxel Collision

Collided Voxels



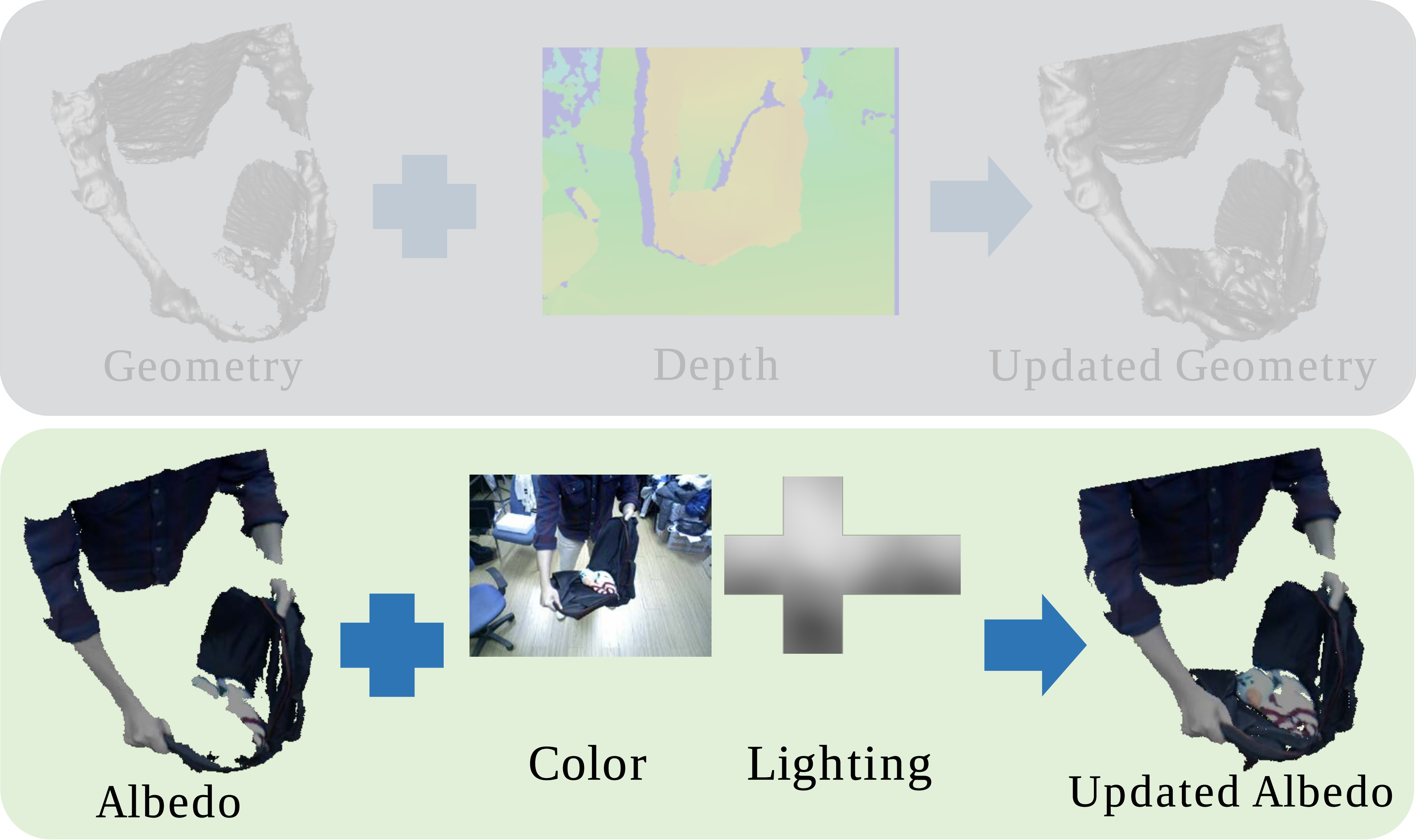
Erroneous Surface in Canonical Frame

Erroneous Surface in Live Frame

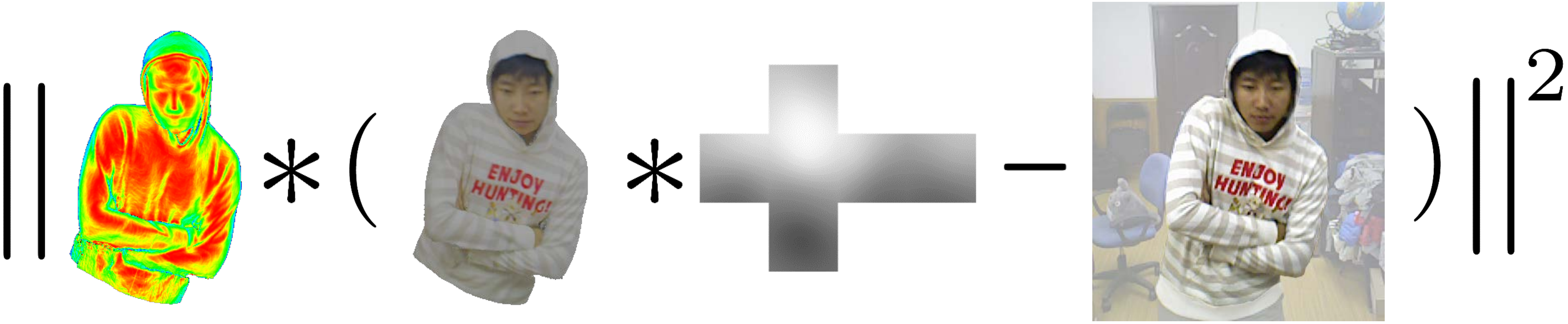
Correct Surface in Canonical Frame

Correct Surface in Live Frame

Update Geometry and Albedo Model



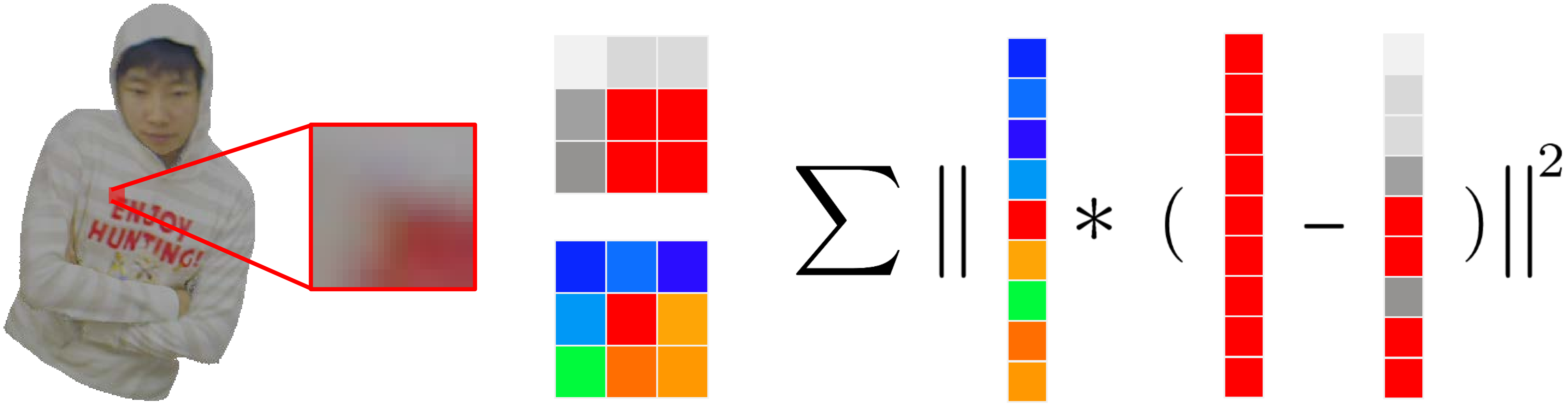
Volumetric Optimization of Albedo



$$E^v_{\text{shading}} = \sum_{\mathbf{x}\in V^t_M} \phi(\mathbf{n}(\mathbf{x}^t) - \mathbf{c}(\mathbf{x}^t)) \left\|C^t(M_c(\mathbf{x}^t)) - B^t(\mathbf{x})\right\|_2^2$$

$$E_{\text{albedo}}(\mathcal{A}) = E^v_{\text{shading}} + w_{as}E_{\text{asreg}} + w_{at}E_{\text{atreg}}$$

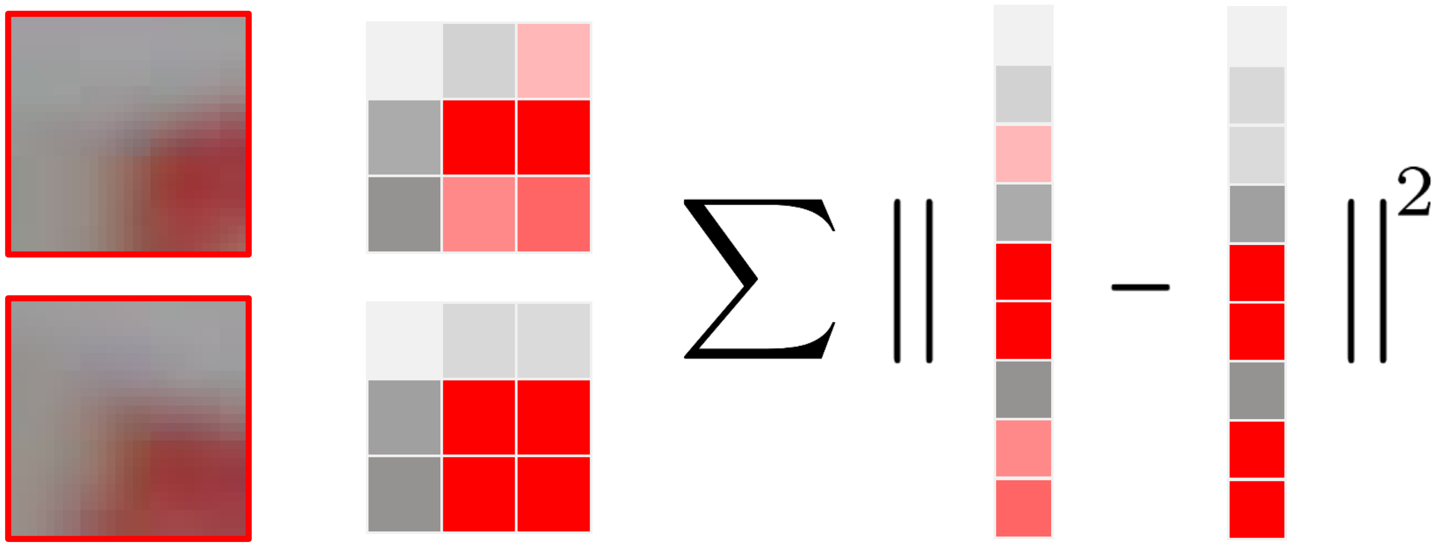
Volumetric Optimization of Albedo



$$E_{\text{asreg}} = \sum_{\mathbf{x} \in V_M^t} \sum_{\mathbf{z} \in N(\mathbf{x}) \cap V_M^t} \phi(\Gamma(\mathbf{x}^t) - \Gamma(\mathbf{z}^t)) \left\| \mathbf{a}^t(\mathbf{x}) - \mathbf{a}^t(\mathbf{z}) \right\|_2^2$$

$$E_{\text{albedo}}(\mathcal{A}) = E^v_{\text{shading}} + \boxed{w_{as} E_{\text{asreg}}} + w_{at} E_{\text{atreg}}$$

Volumetric Optimization of Albedo



$$E_{\text{atreg}} = \sum_{\mathbf{x} \in V_M^t \cap V_M^{t-1}} \left\| \mathbf{a}^{t-1}(\mathbf{x}) - \mathbf{a}^t(\mathbf{x}) \right\|_2^2$$

$$E_{\text{albedo}}(\mathcal{A}) = E_{\text{shading}}^v + w_{as} E_{\text{asreg}} + w_{at} E_{\text{atreg}}$$

Effects of Albedo Fusion

input color

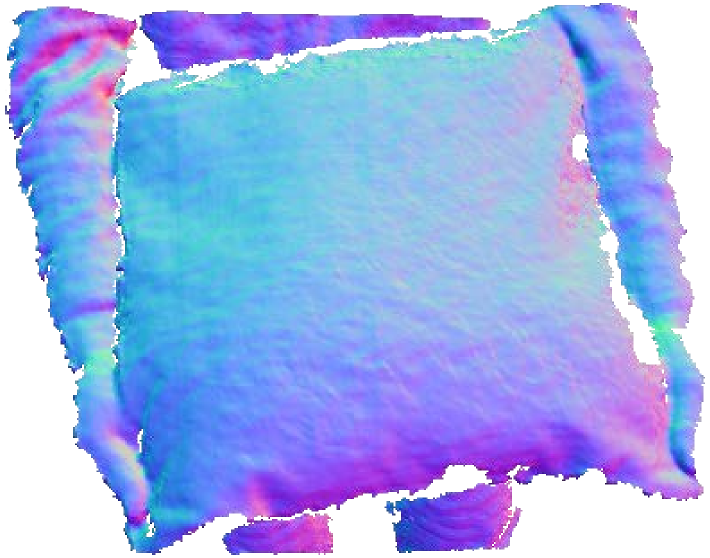
direct blending

our method

Effects of Albedo Fusion

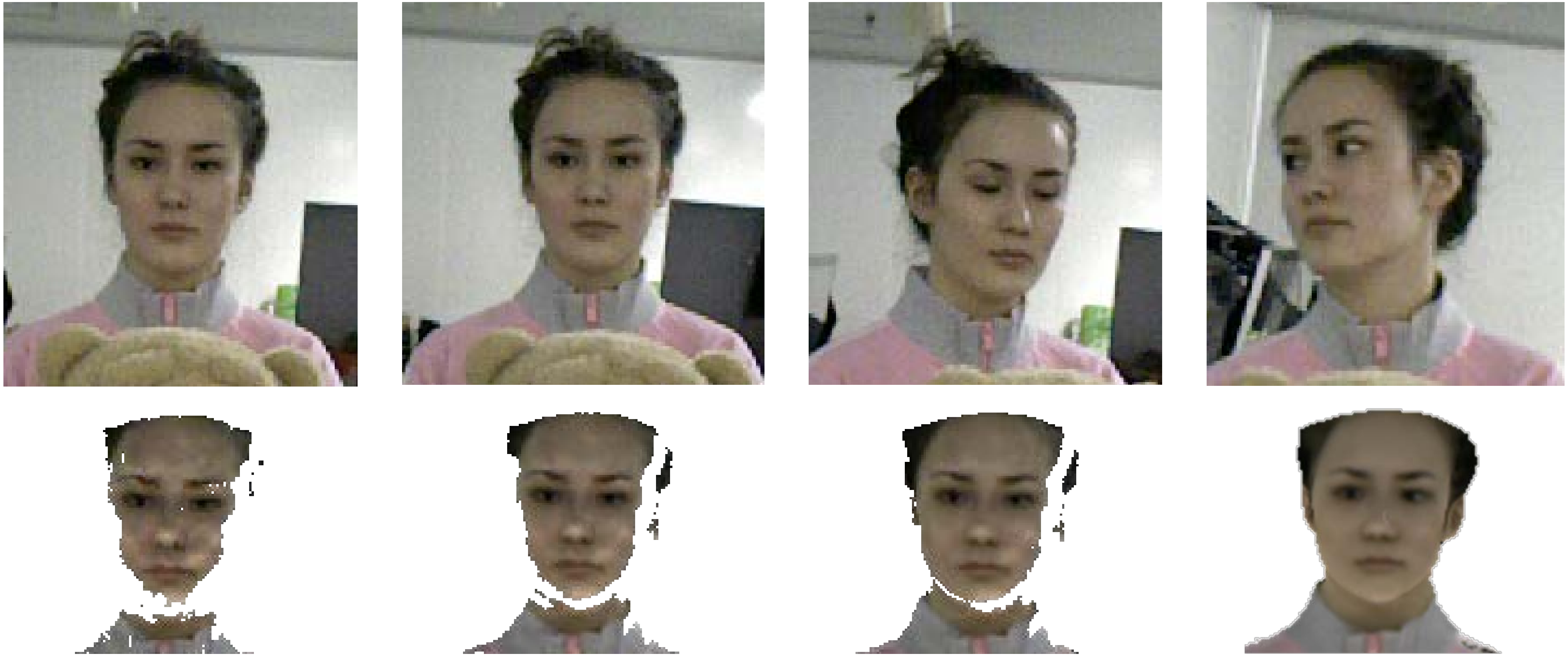
input color

albedo map

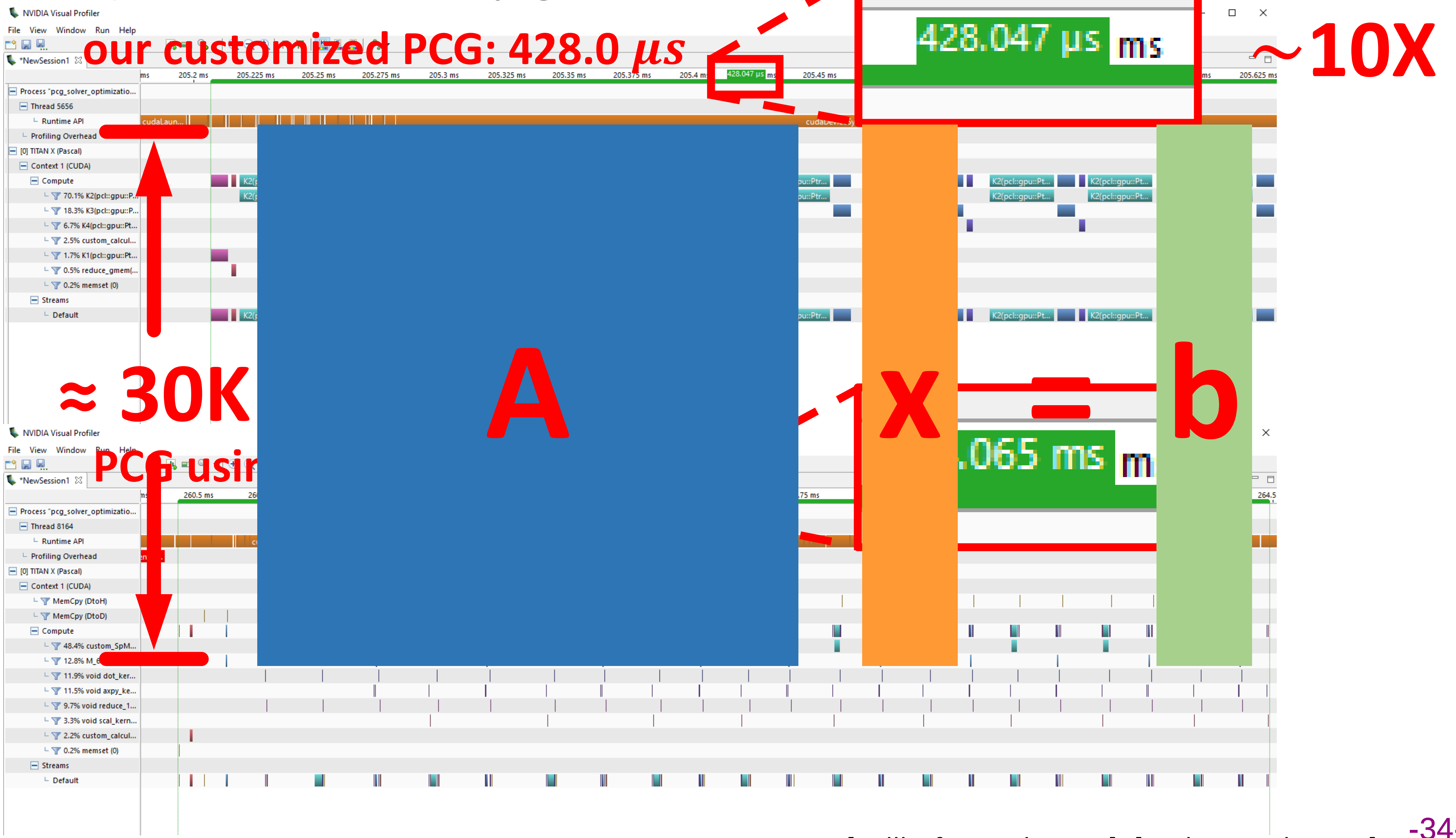


normal map

Effects of Albedo Fusion

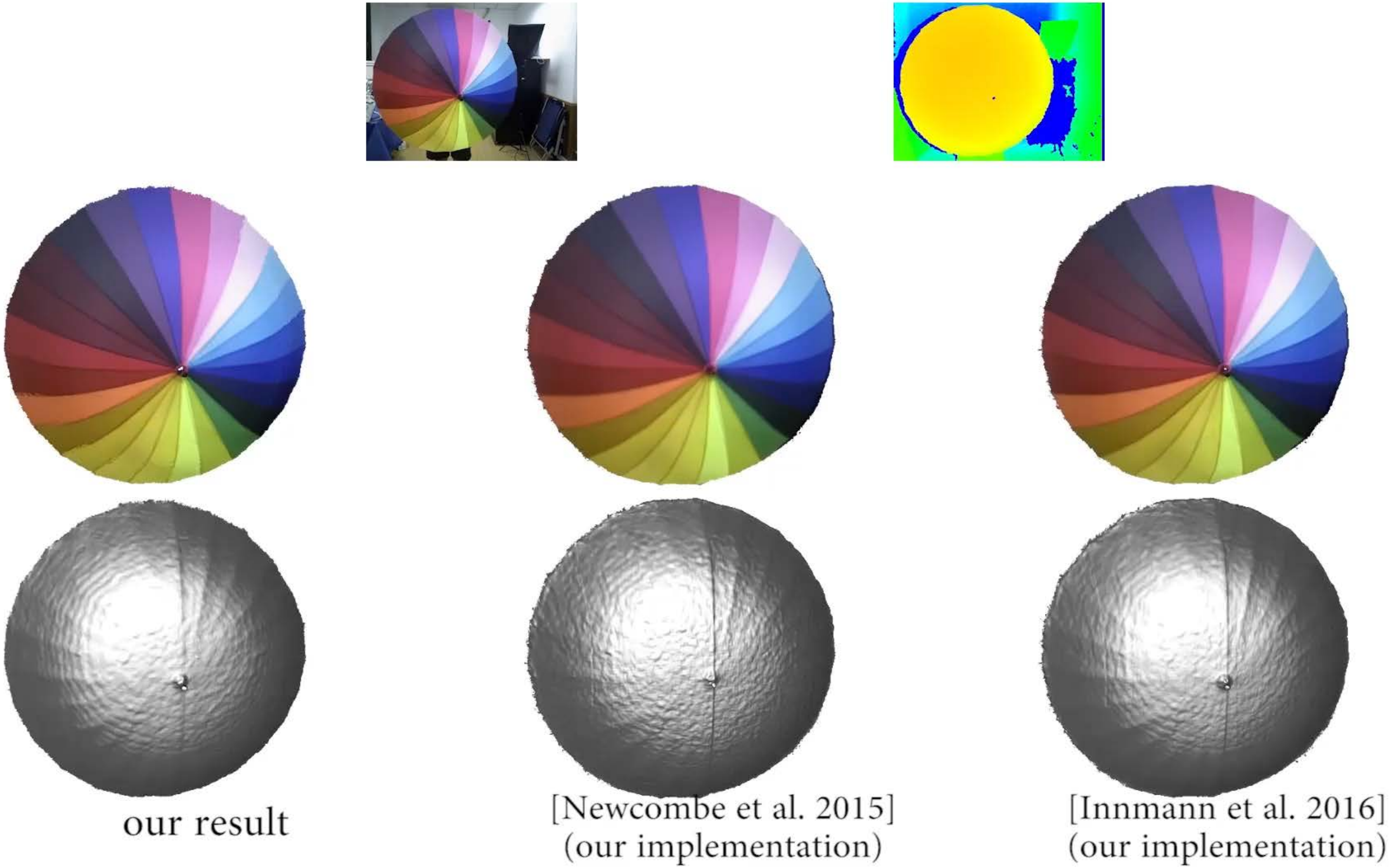


PCG （Preconditioner Conjugate Gradient） Solver Performance

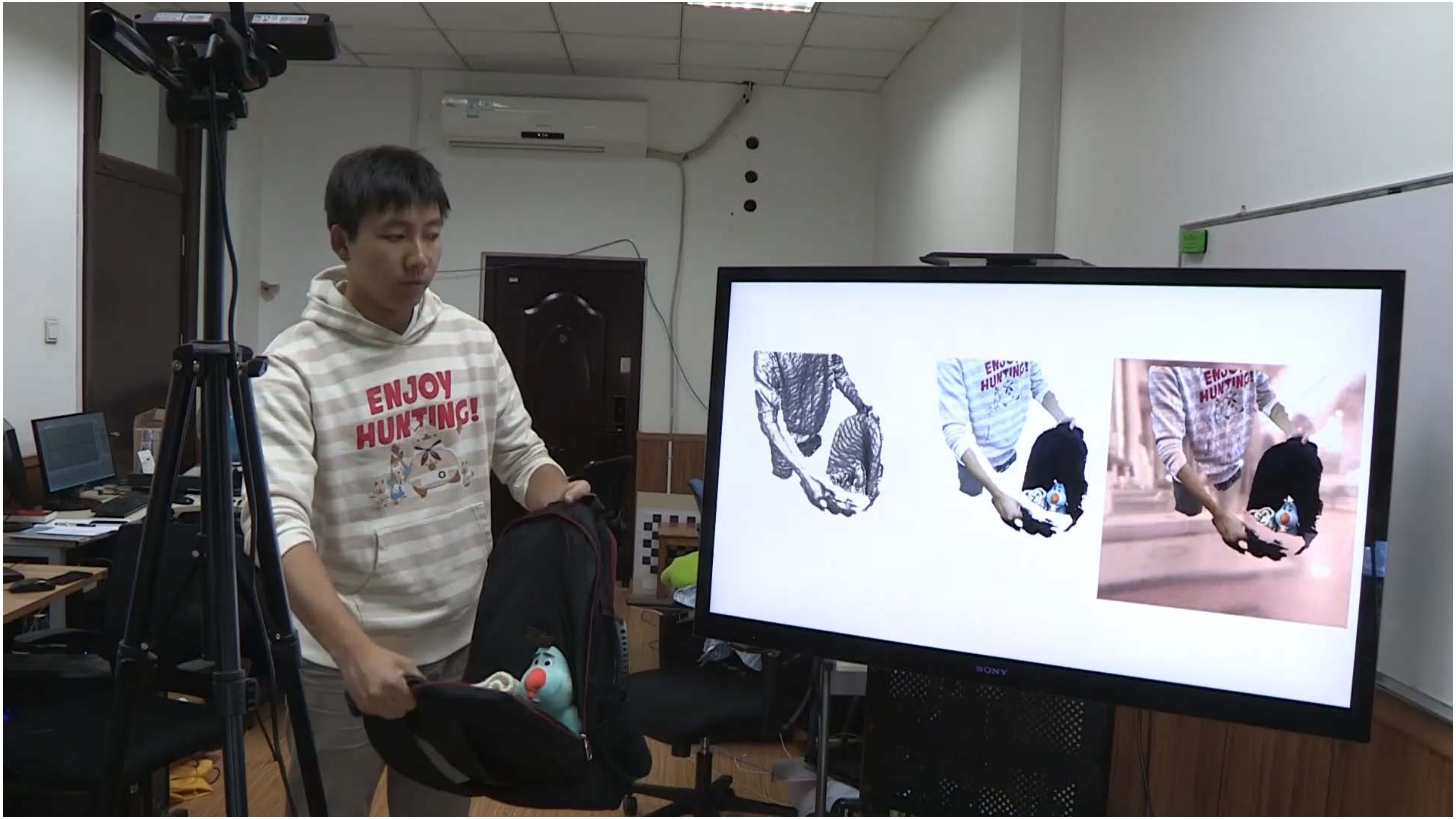


[Zollhofer et al. 2014], [Weber et al. 2013]

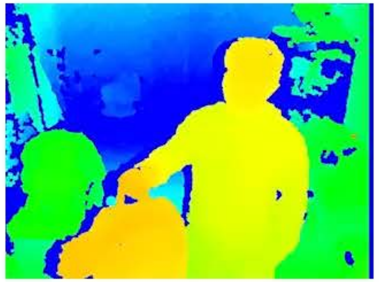
Comparison


our result
[Newcombe et al. 2015]
(our implementation)
[Innmann et al. 2016]
(our implementation)

Results

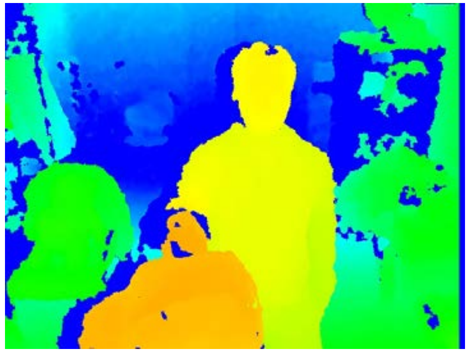


Application of Free Viewpoint Video



Virtual Camera 0 Virtual Camera 1 Virtual Camera 2

Application of Holographics

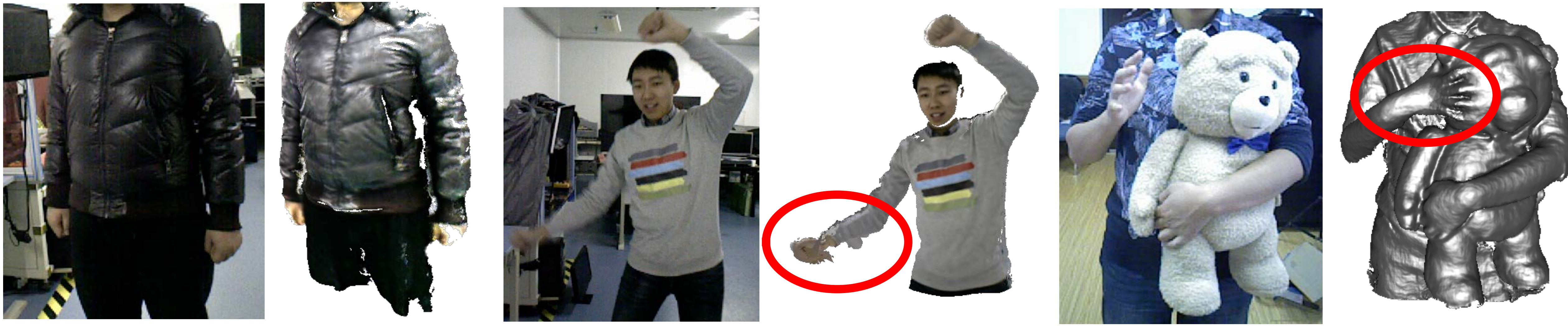


Application of Real-time 3D Self-portrait

Video Editing



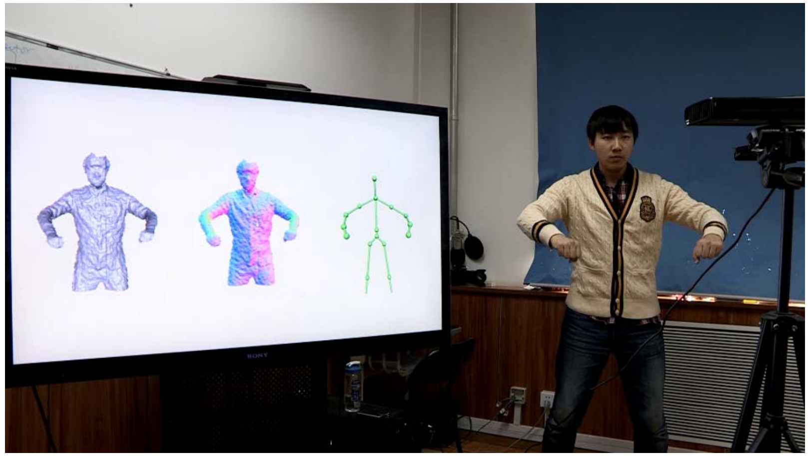

Limitations

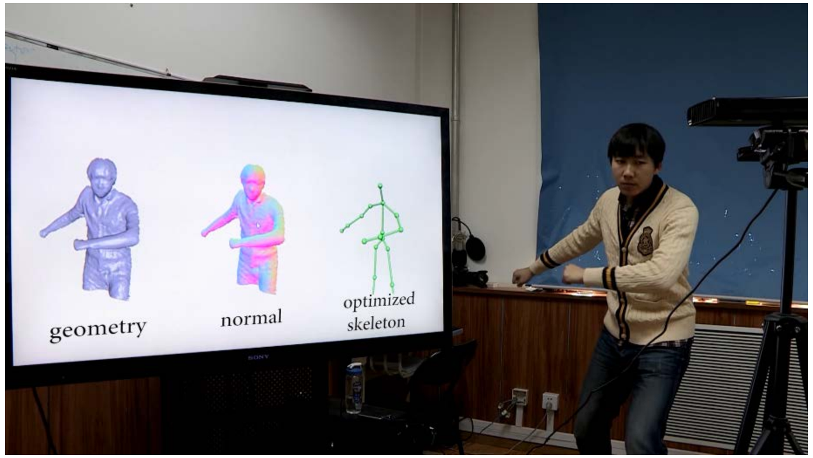


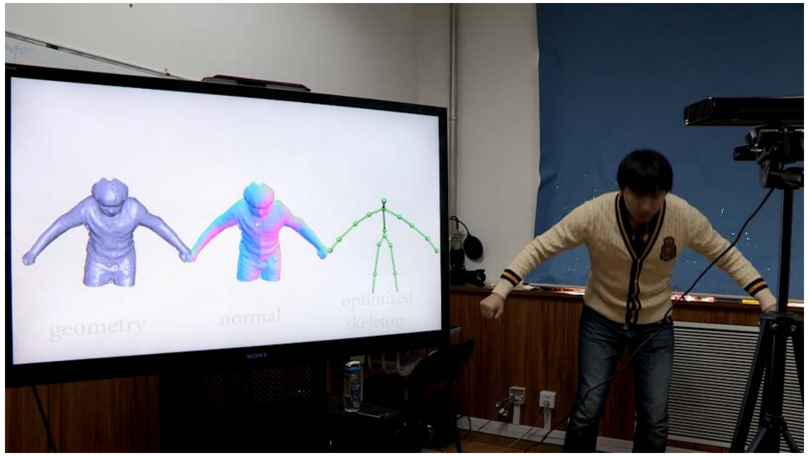
specular

fast motions

close-to-open





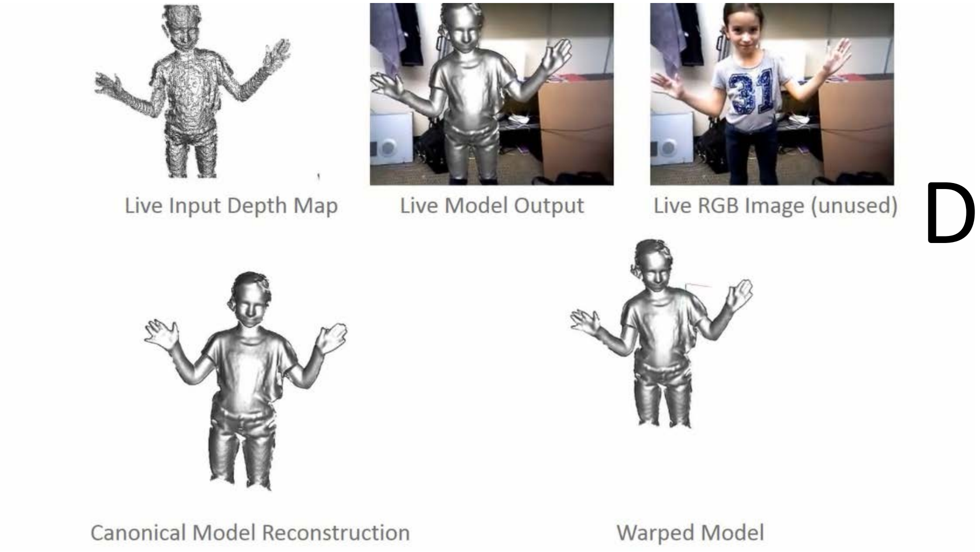


BodyFusion: Real-Time Capture of Human Motion and Surface Geometry Using a Single Depth Camera

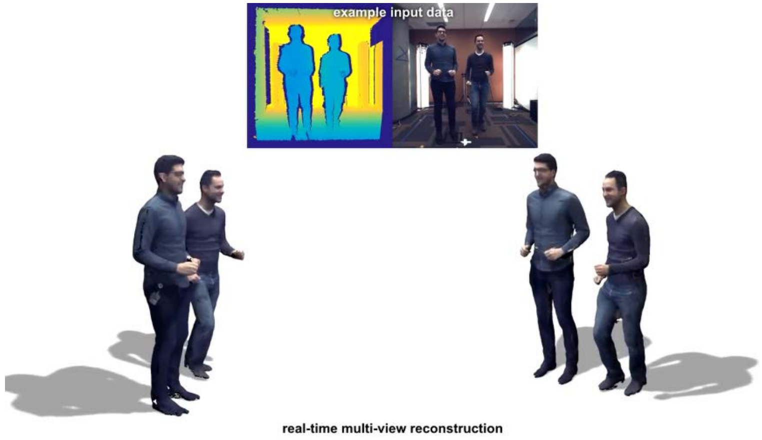
Tao Yu[1,2] , Kaiwen Guo[1], Feng Xu[1], Yuan Dong[1], Zhaoqi Su[1], Jianhui Zhao[2], Jianguo Li[3], Qionghai Dai[1], Yebin Liu[1]

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

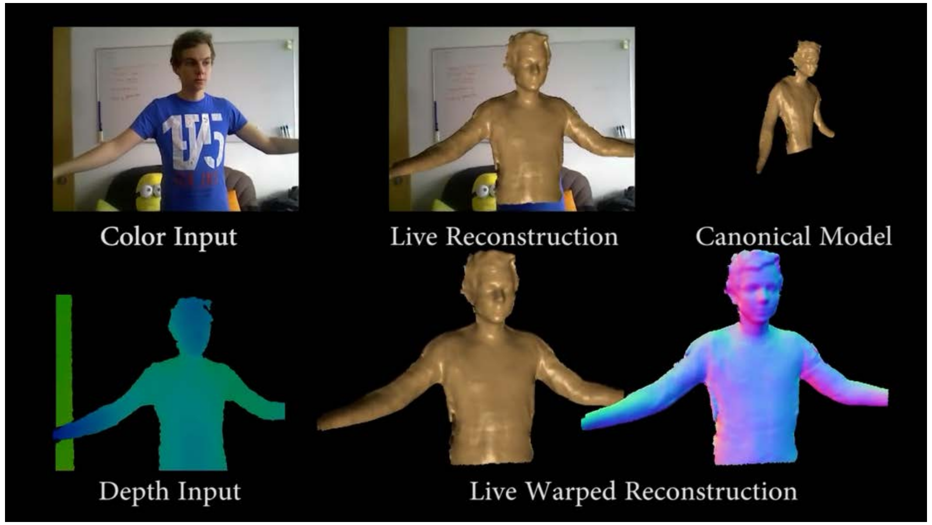
BodyFusion: Background & Related Works



DynamicFusion



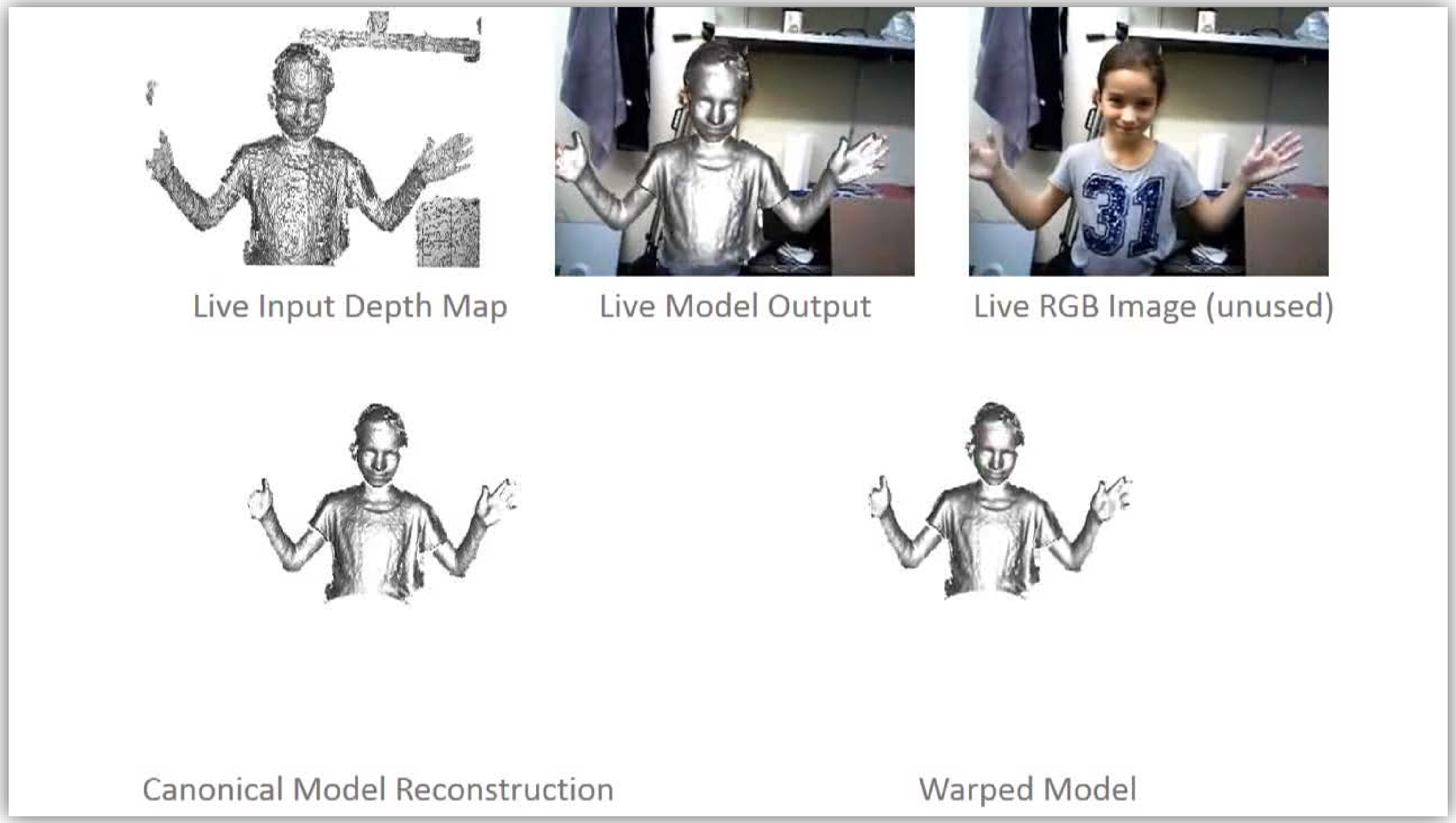
Fusion4D



VolumeDeform

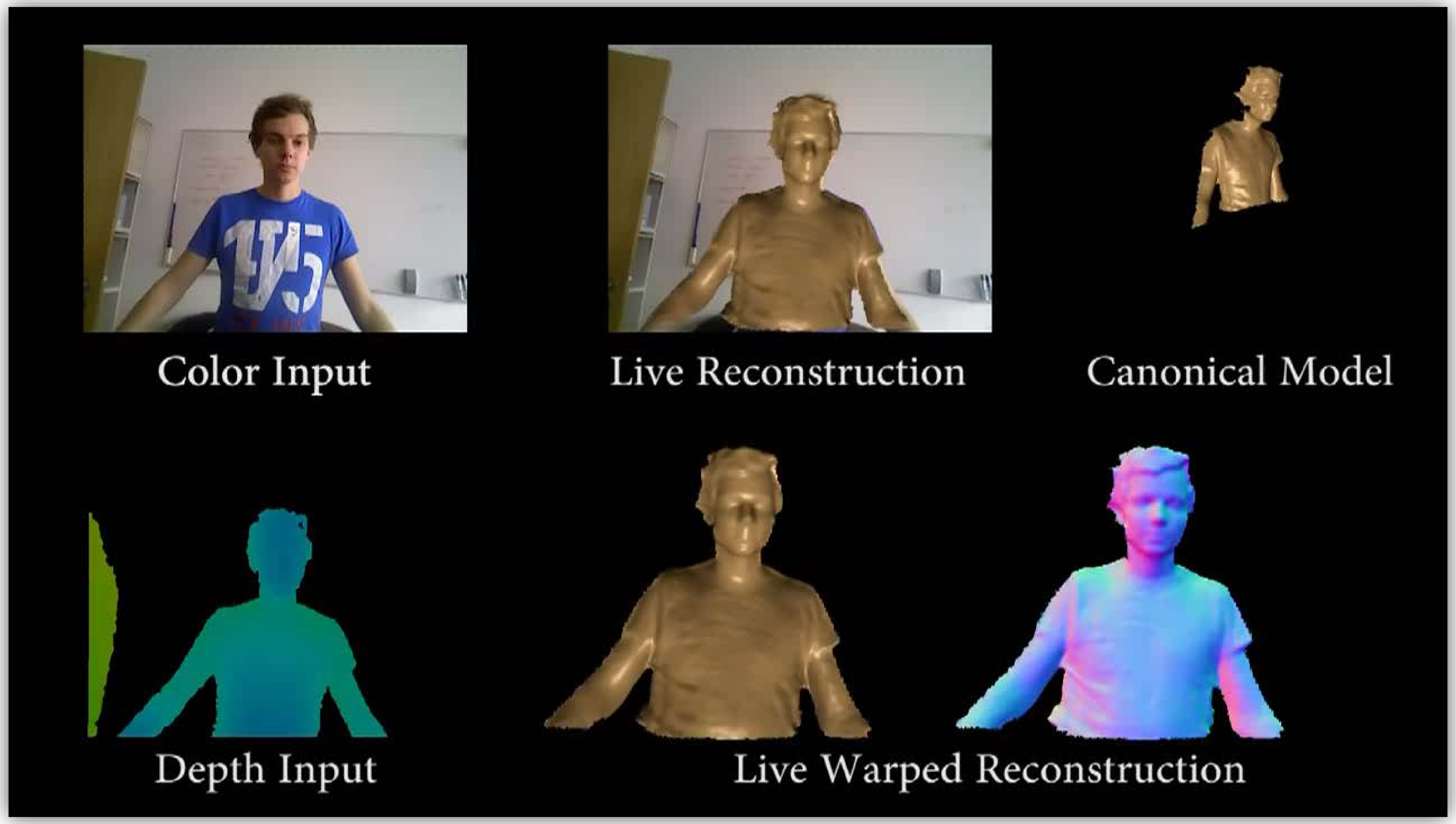
BodyFusion: Background & Related Works

DynamicFusion Human Body Reconstruction Result



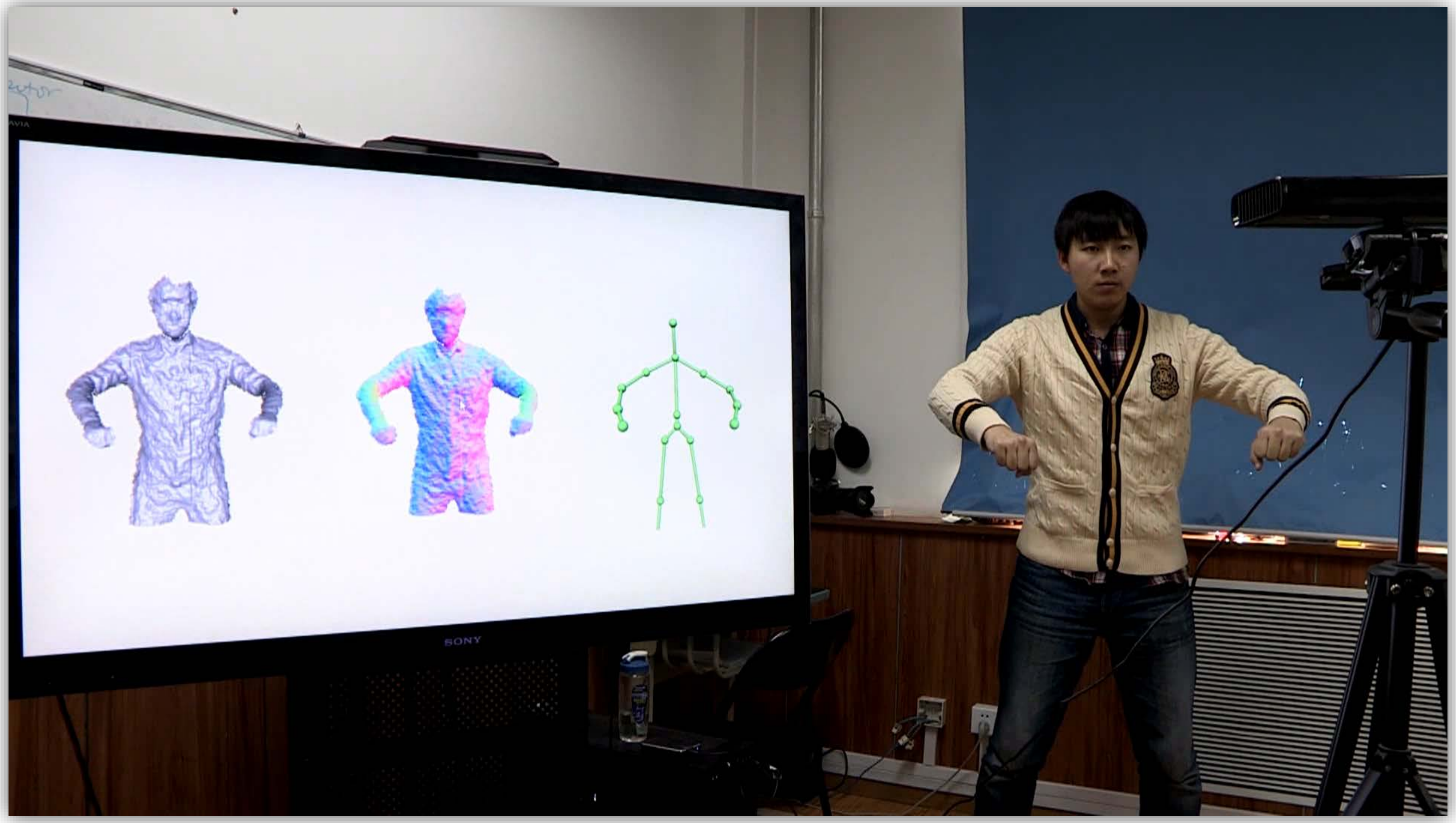
BodyFusion: Background & Related Works

VolumeDeform Human Body Reconstruction Result

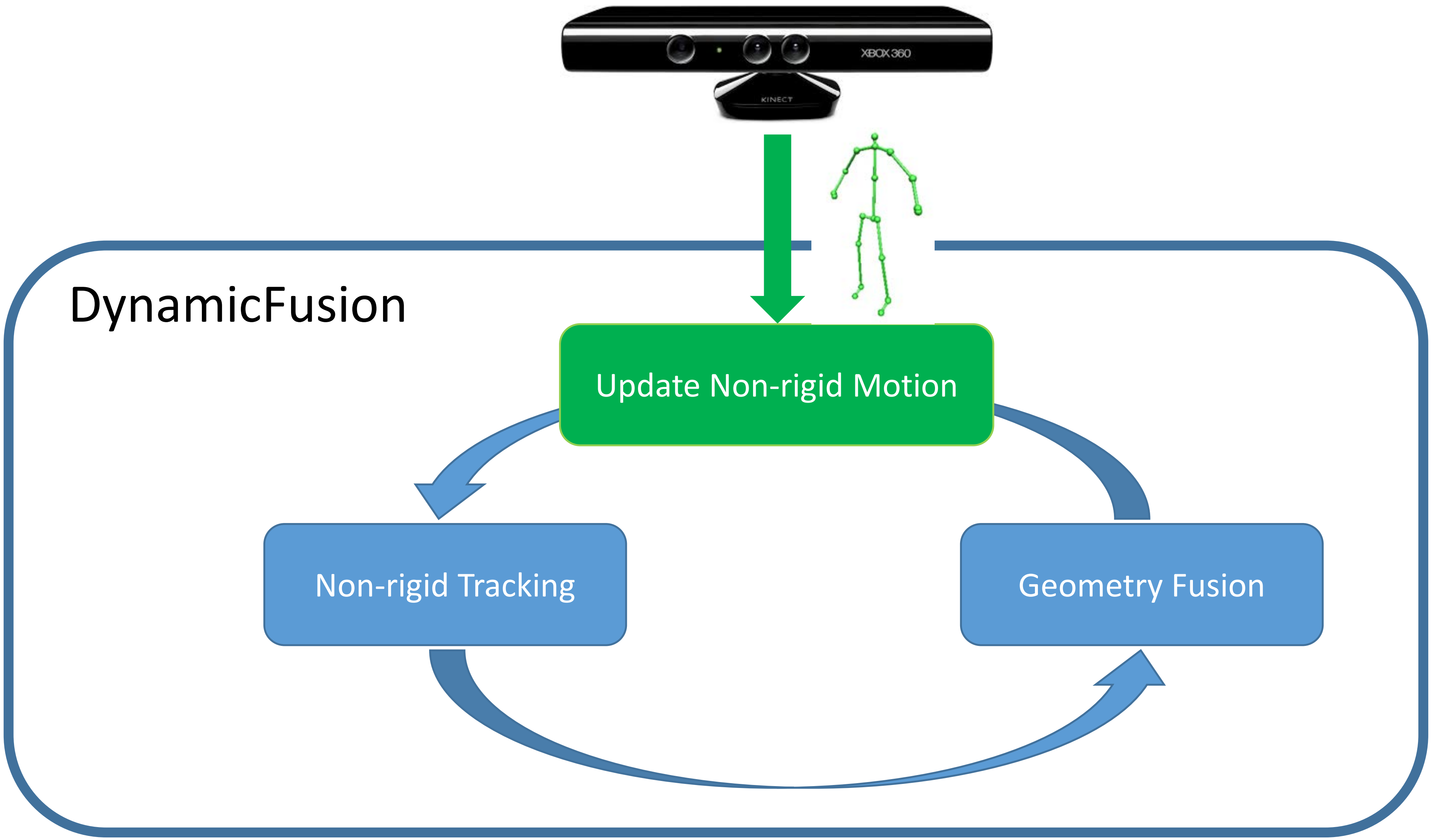


BodyFusion: Background & Related Works

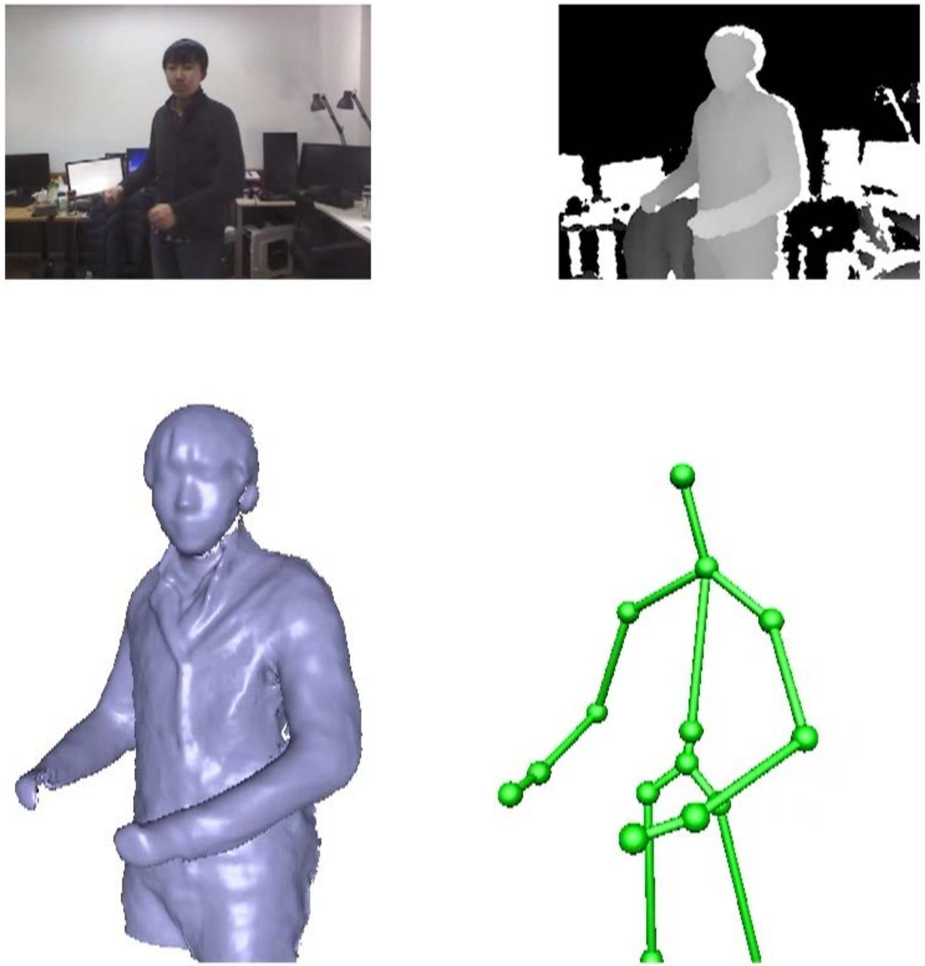
BodyFusion Human Body Reconstruction Result



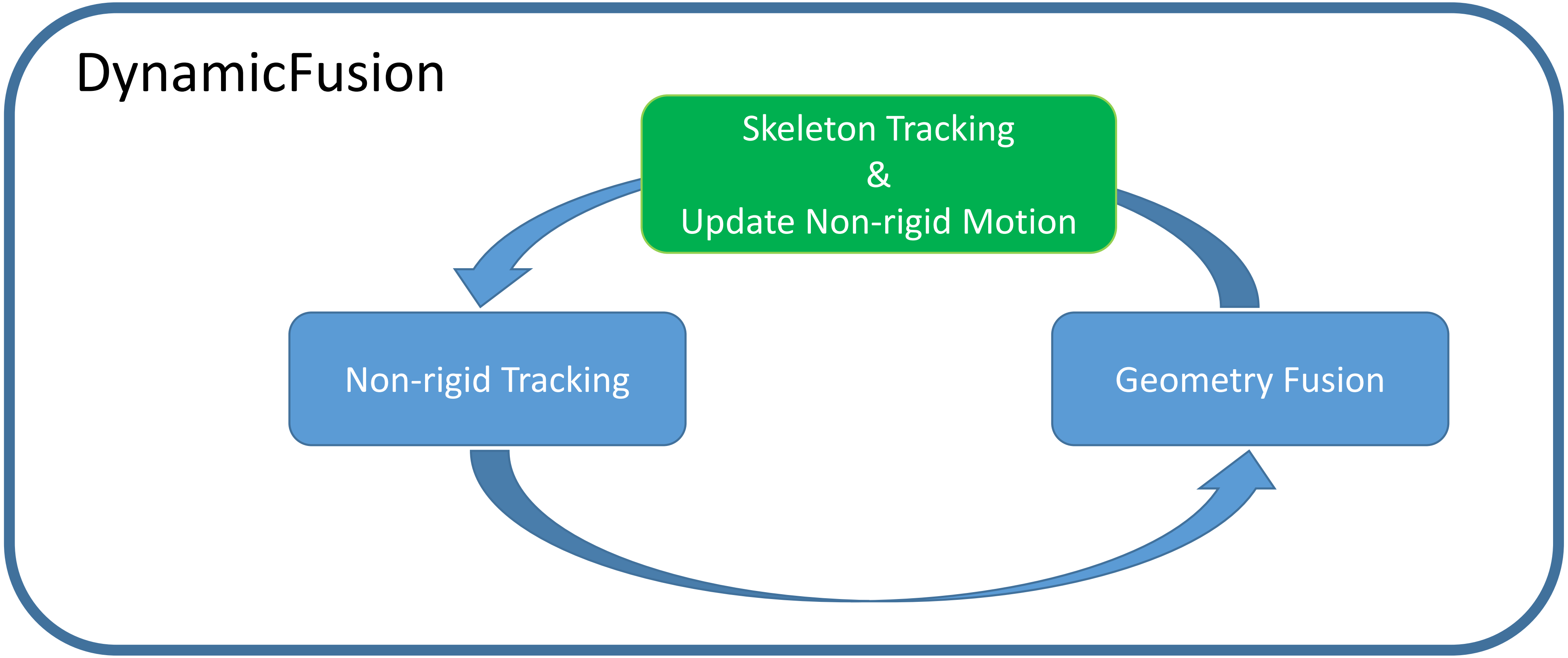
BodyFusion: History (Solution1)



BodyFusion: History (Solution1)

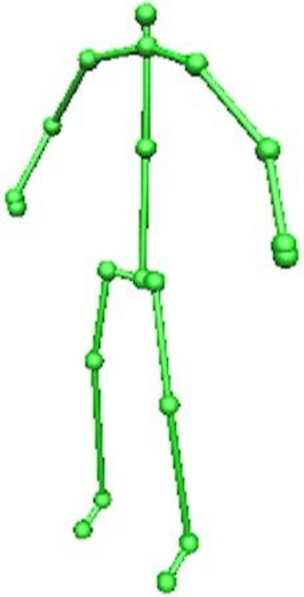


BodyFusion: History (Solution2)

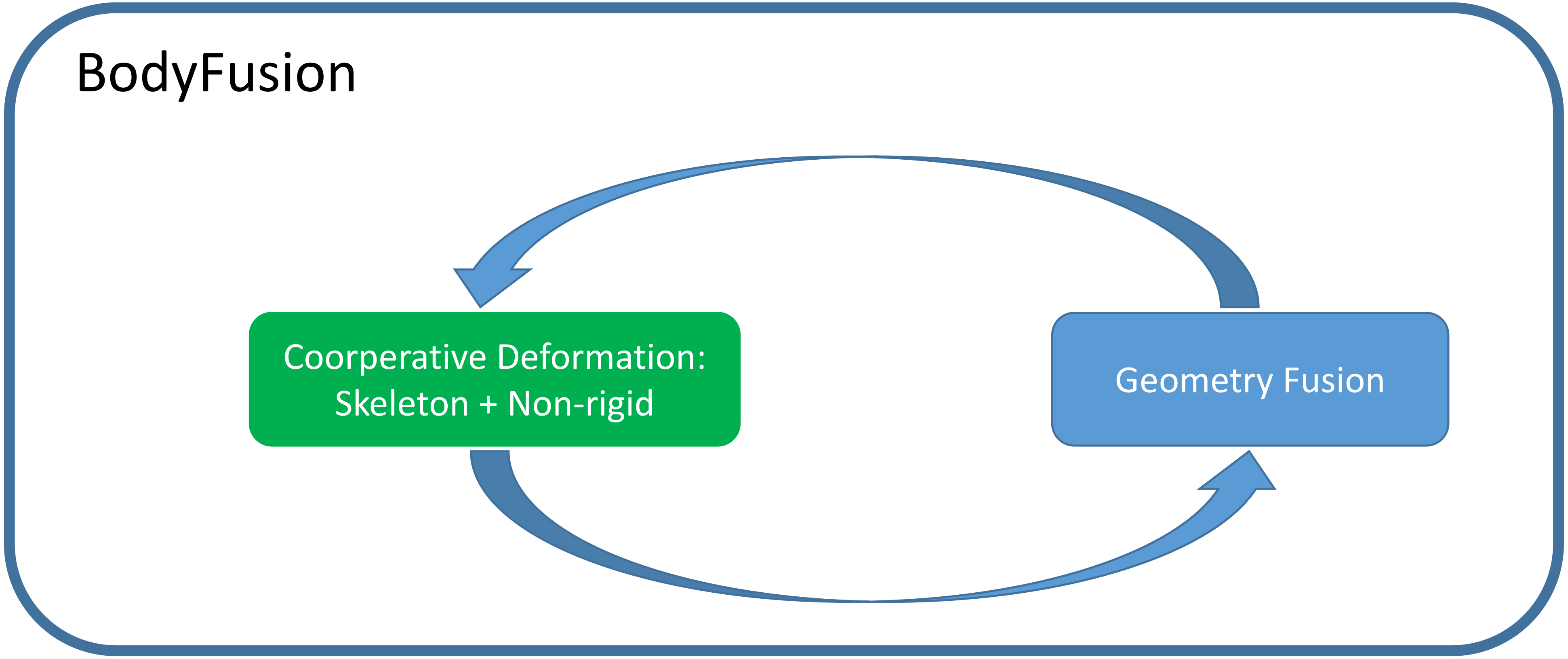


BodyFusion: History (Solution2)

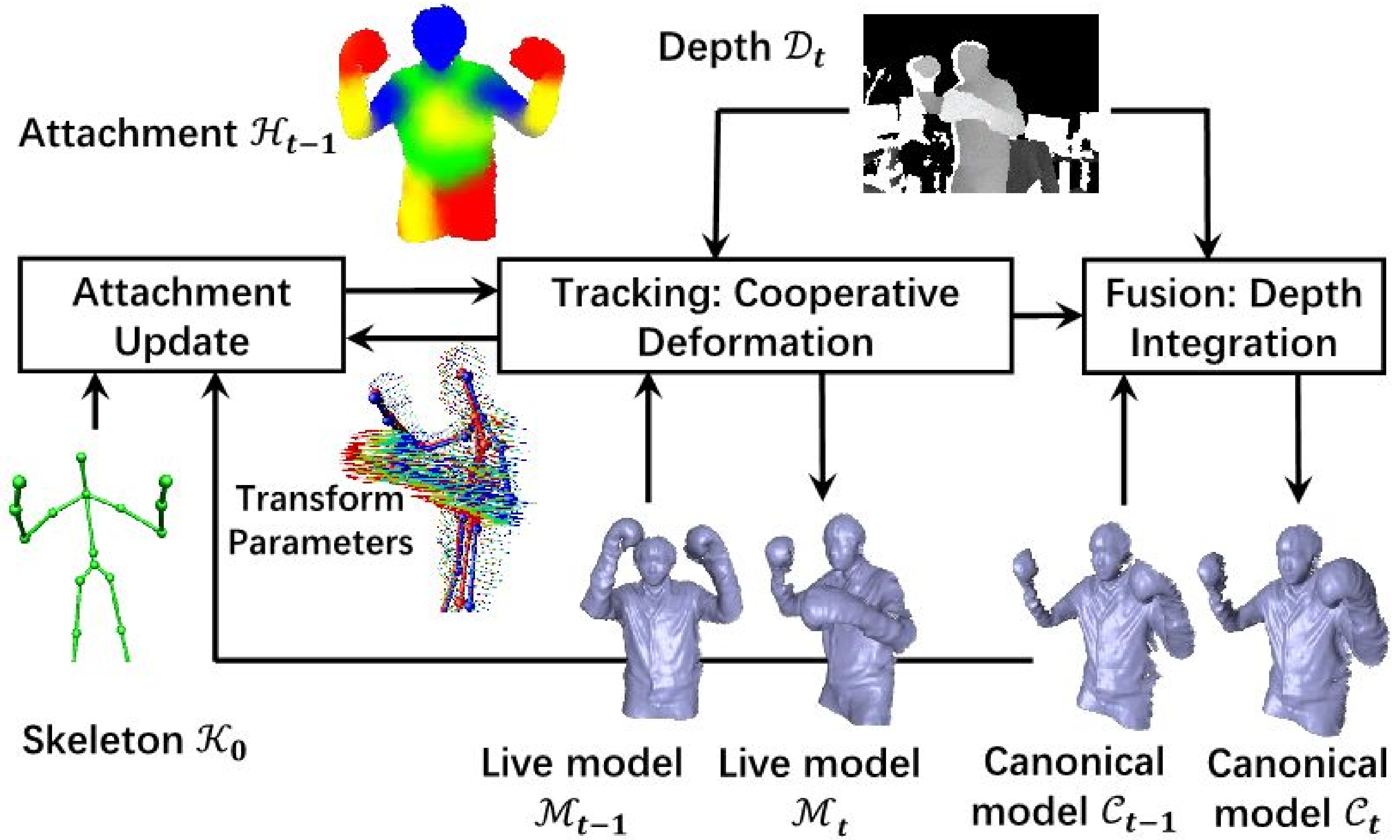


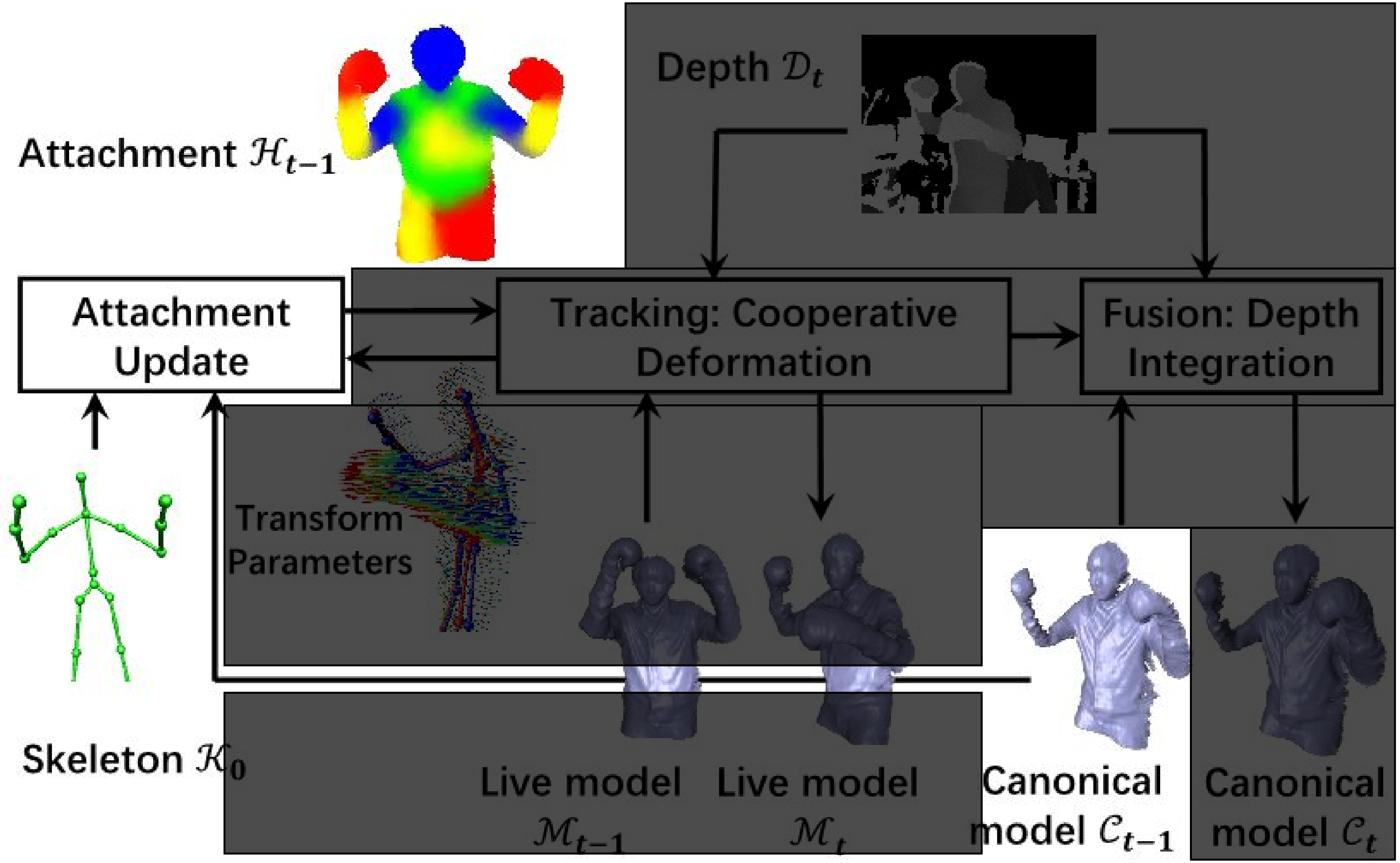
BodyFusion: History (Solution3)



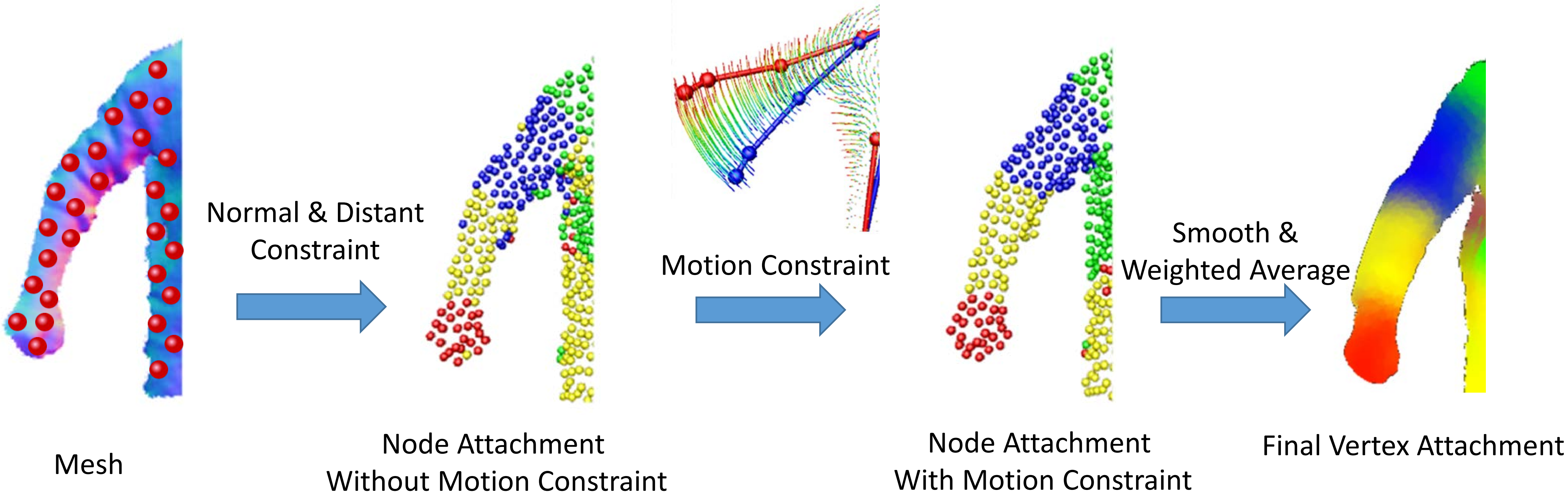
BodyFusion: Overview



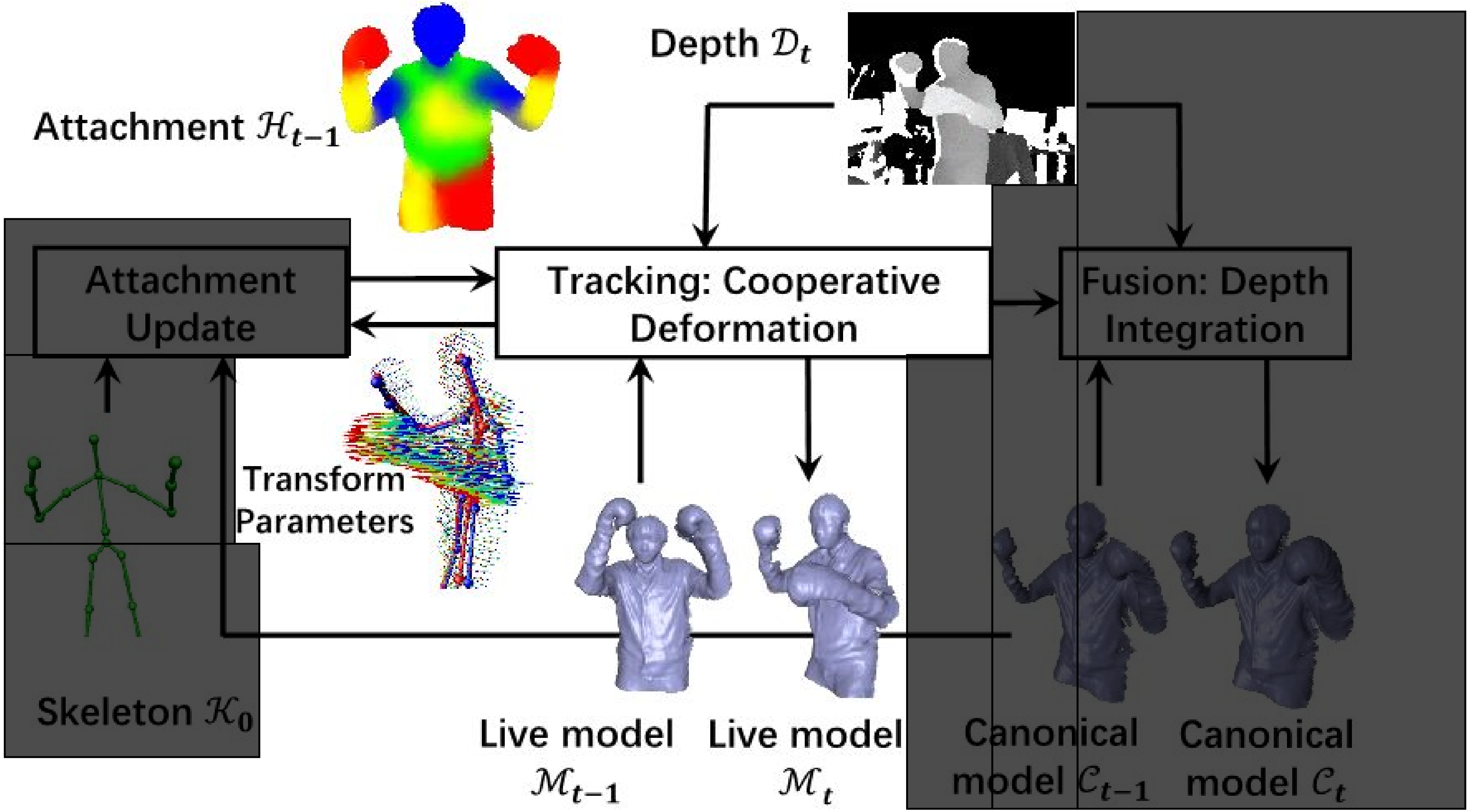
BodyFusion: Attachment Update



BodyFusion: Attachment Update



BodyFusion: Coorperative Deformation



BodyFusion: Coorperative Deformation

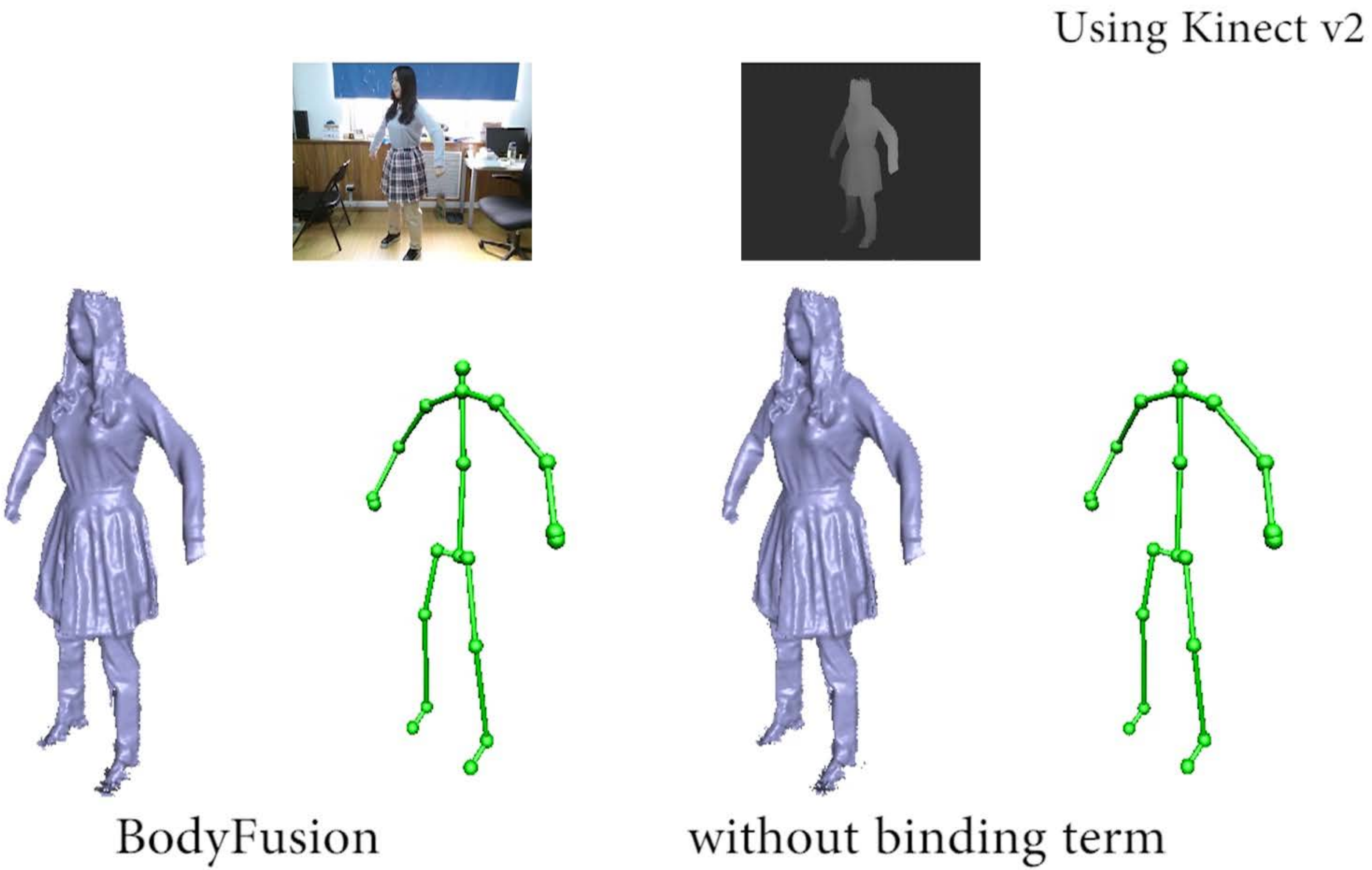
$$E_{\mathrm{t}} = \lambda_n E_{\mathrm{nonrigid}} + \lambda_s E_{\mathrm{skeleton}} + \lambda_g E_{\mathrm{graph}} + \lambda_b E_{\mathrm{binding}}$$

Non-rigid tracking data term &
Skeleton tracking data term

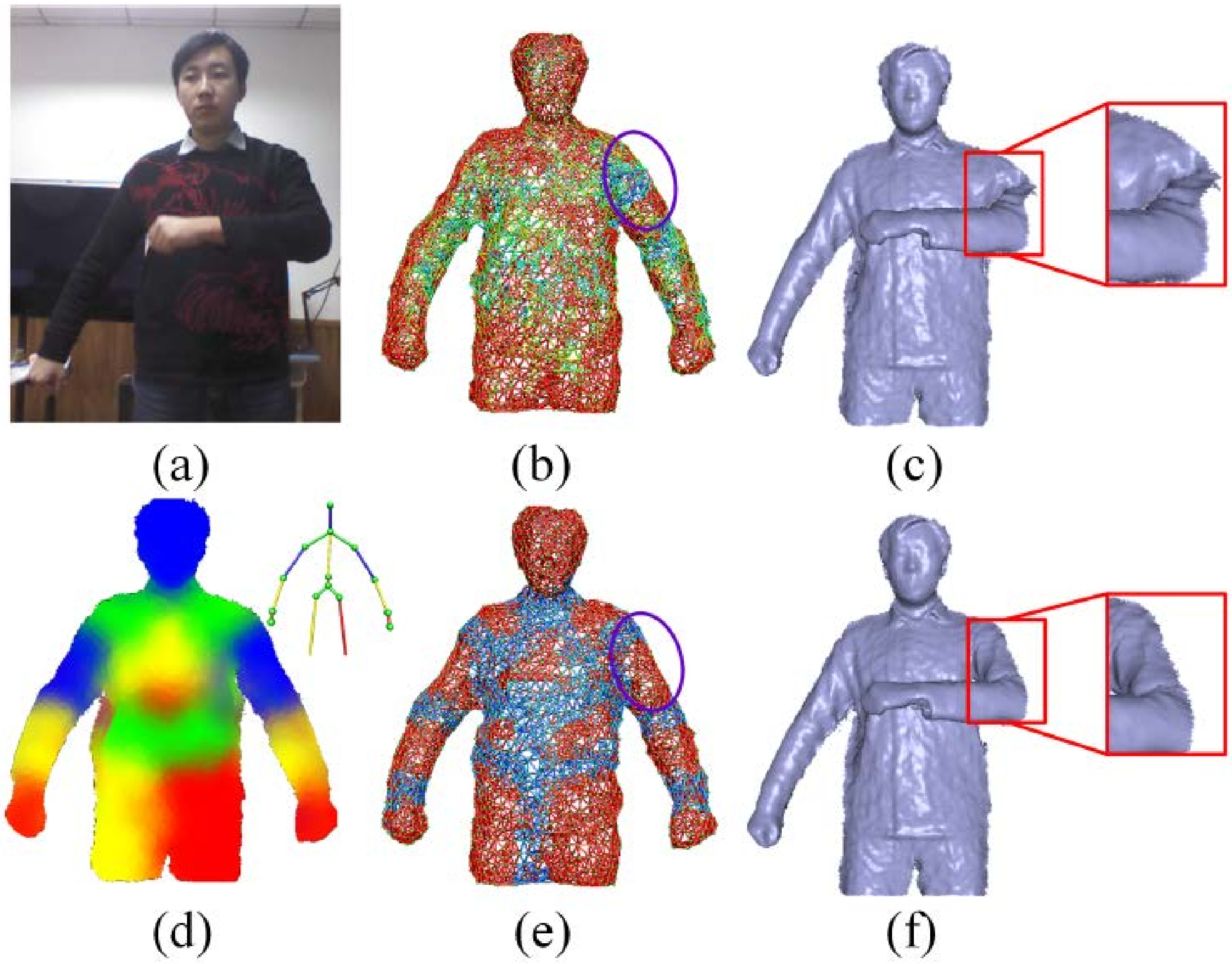
Node-graph local
as-rigid-as-possible term

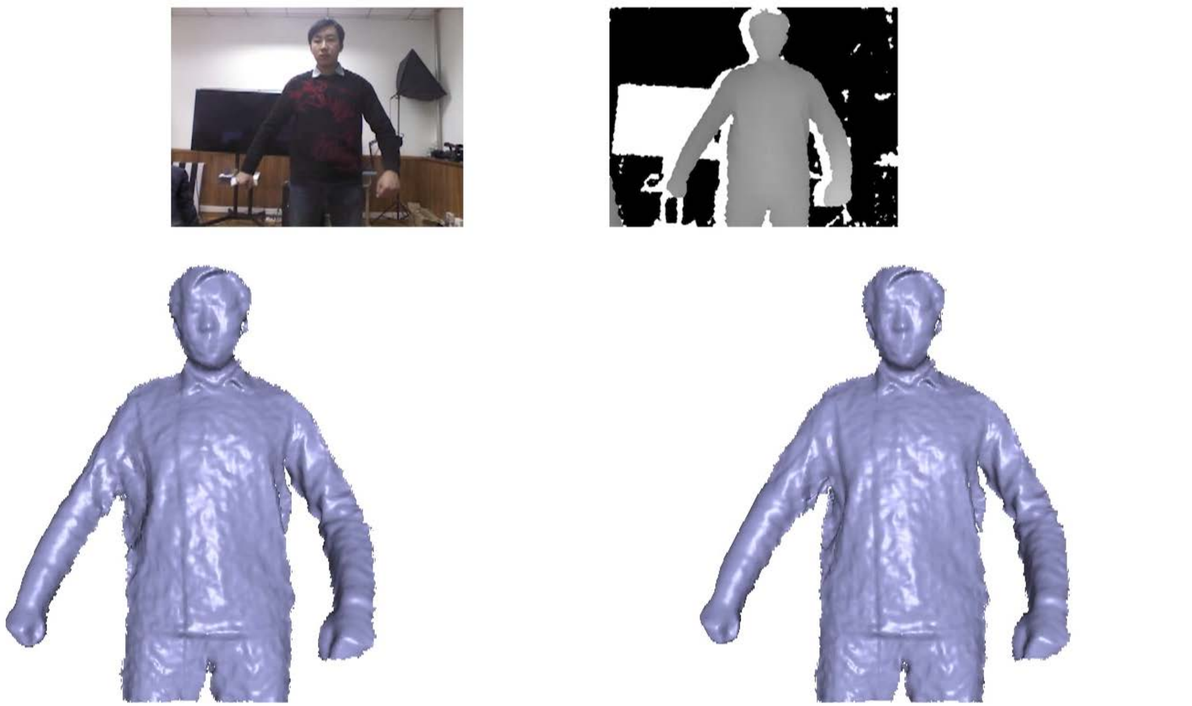
Skeleton tracking and Non-rigid
tracking Binding term

BodyFusion: Evaluation – Binding Term



BodyFusion: Evaluation – Graph Term

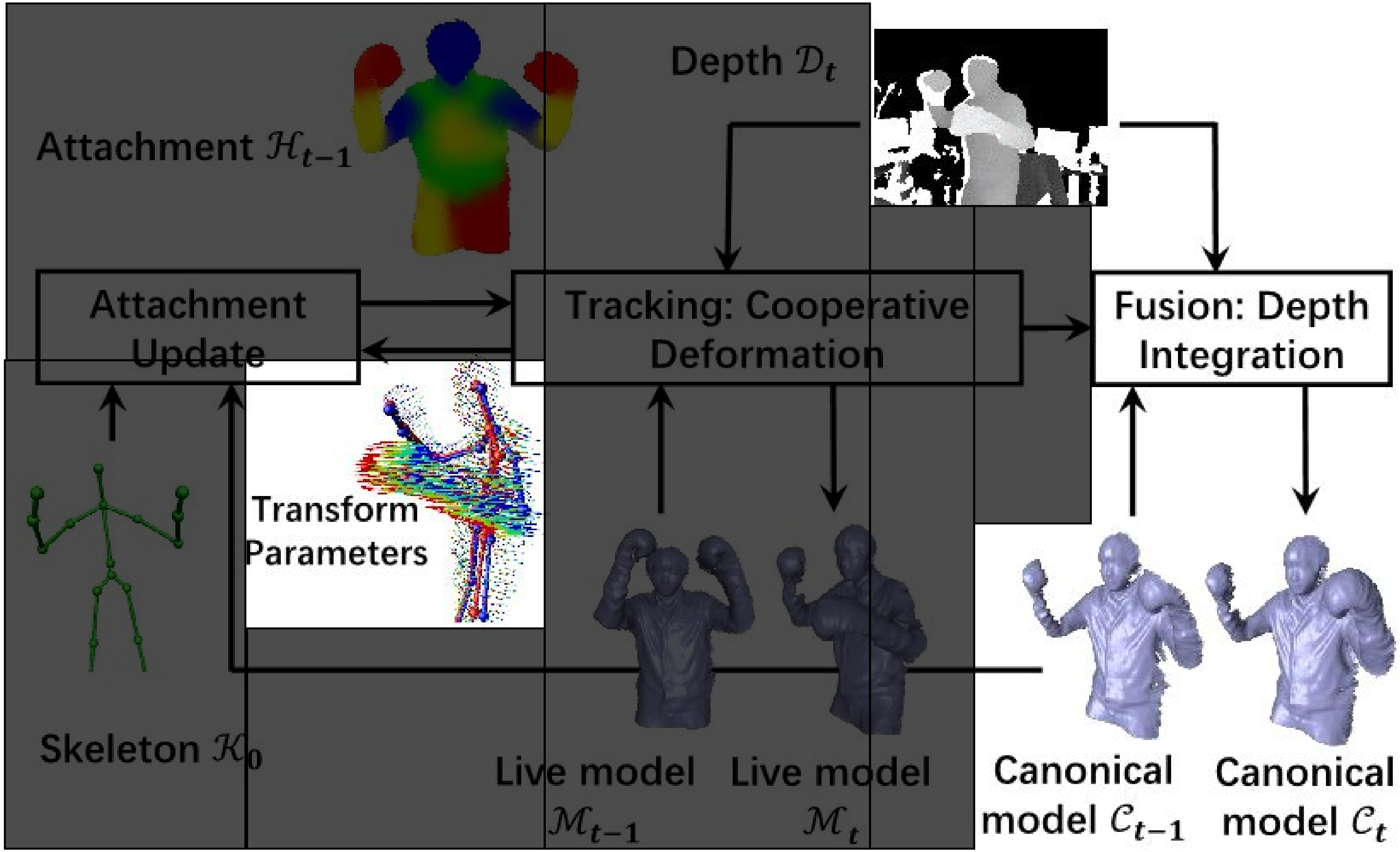




smooth term based on attachment

smooth term used in DynamicFusion

BodyFusion: Depth Integration



BodyFusion: More Results

Using Kinect v1

Boxing

BodyFusion: Comparisons

Dancing

BodyFusion: Limitations


(a) (b) (c) (d) (e)

Thank you