Capturing and Editing Reality – Performance Capture of the Real World in Motion

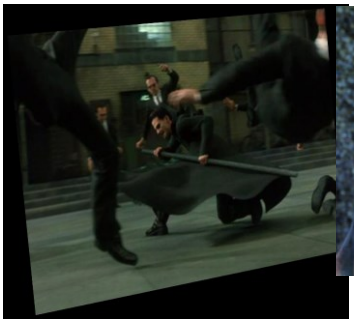
Christian Theobalt

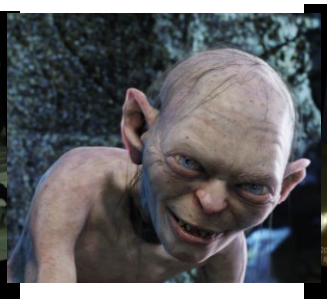
Graphics, Vision, and Video

max planck institut informatik

The Modeling Bottleneck

max planck institut informatik

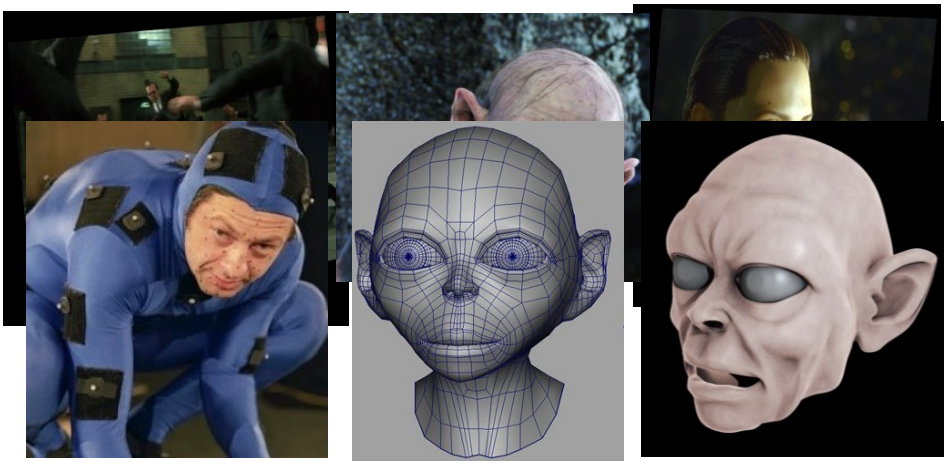




[Lord of the Rings]

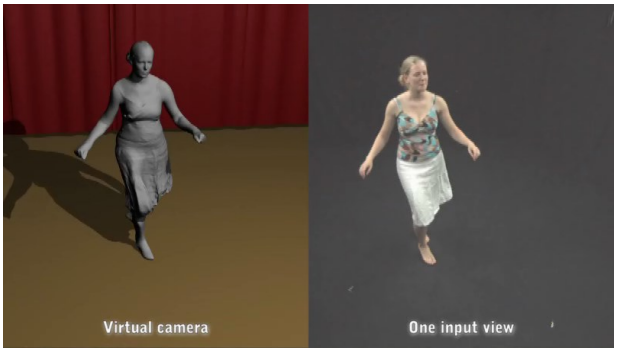
[Beowulf]

The Modeling Bottleneck



Motion Shape Textures

Performance Capture



[de Aguiar et al,, SIGGRAPH 2008]

- High-quality shape, motion and texture of people in general clothing from multi-view video
- Handful of cameras to single camera, no markers
- Exceed capabilities of existing Mocap technology by far

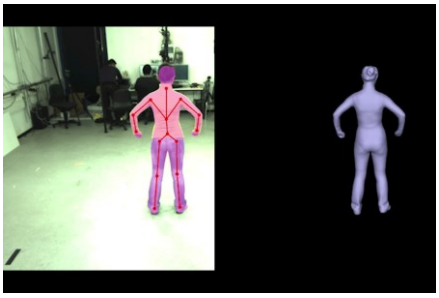
Skeleton + Surface

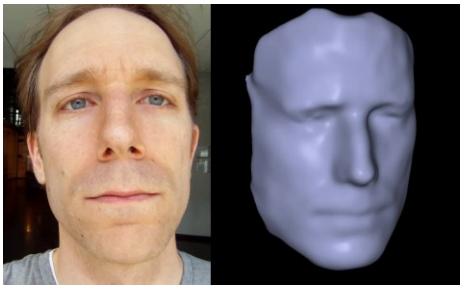
Template-based pose and shape optimization + multi-view segmentation

[Liu et al. CVPR 2011 oral, IEEE Trans. PAMI 2013]

Christian Theobalt

Important Limitations





- Limited accuracy
- Limited scene complexity
- Starkly restricted capture environment

Christian Theobalt

What We Want: Capture Reality

- Complex appearance
- Uncontrolled recording and lighting settings
- Tremendous scene complexity
- Simple and sparse camera setups
- Ability to edit captured footage

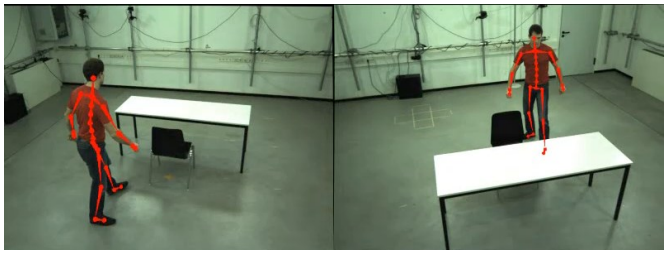
Christian Theobalt

Improving Reconstruction Quality and Ability

Christian Theobalt

Capturing in Less Constrained Environments

Occlusions,
fast motions,
dynamic backgroud...

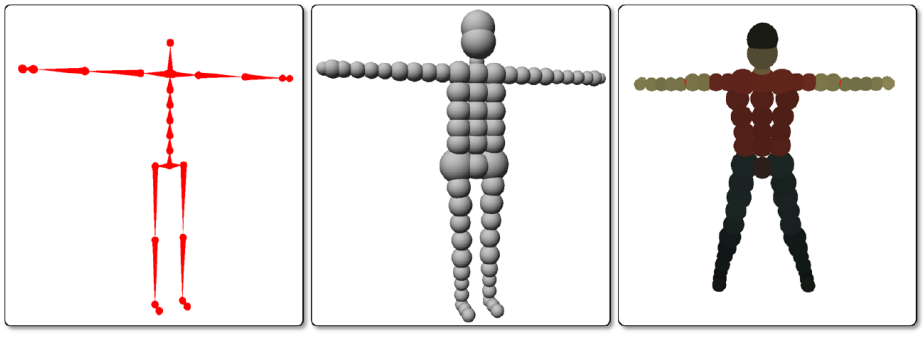


[Stoll et el. ICCV 2011]

Christian Theobalt

3D Sum of Gaussians Body Model

- Skeleton + Gaussians with color model



58 DoFs

63 Gaussians

Christian Theobalt

2D Sum of Gaussians Body Model

- Each video frame converted to 2D SoG

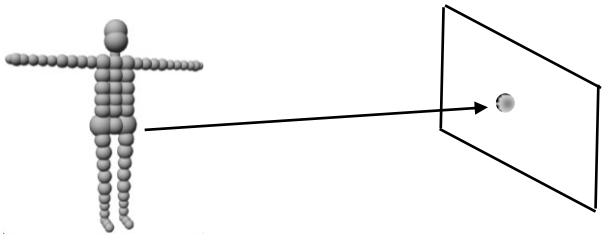
Video frame

Quad tree –
each region one 2D Gaussian

Christian Theobalt

3D-2D Similarity Function

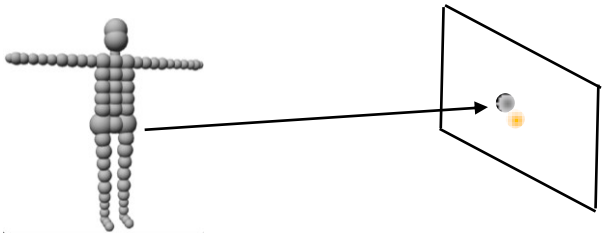
- Projection of 3D Gaussians into cameras



Christian Theobalt

3D-2D Similarity Function

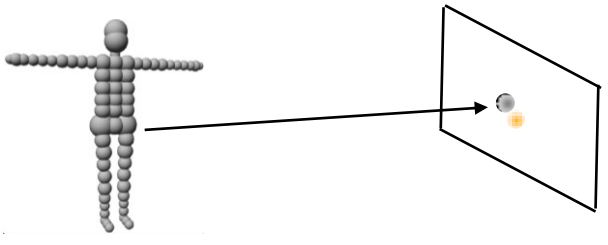
- Projection of 3D Gaussians into cameras



Christian Theobalt

3D-2D Similarity Function

- Projection of 3D Gaussians into cameras



- 2D-2D SoG similarity in one camera view

$$E_{ij} = d(\mathbf{c}_i, \mathbf{c}_j) \int_{\Omega} \mathcal{B}_i(x)\mathcal{B}_j(x)\mathrm{d}\mathbf{x}$$

Color similarity

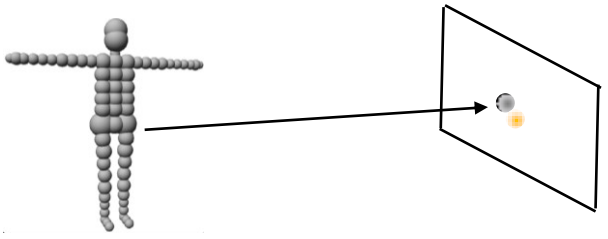
Projected model Gaussian

Image Gaussian

Christian Theobalt

3D-2D Similarity Function

- Projection of 3D Gaussians into cameras



- 2D-2D SoG similarity – closed form solution

$$E_{ij} = d(\mathbf{c}_i, \mathbf{c}_j) \int_{\Omega} \mathcal{B}_i(x)\mathcal{B}_j(x)\mathrm{d}\mathbf{x}$$

$$= d(\mathbf{c}_i, \mathbf{c}_j) 2\pi \frac{{\sigma_i}^2 {\sigma_j}^2}{{\sigma_i}^2 + {\sigma_j}^2} \exp\left(-\frac{\|\mu_i - \mu_j\|^2}{{\sigma_i}^2 + {\sigma_j}^2}\right)$$

Christian Theobalt

Pose Similarity Measure

- Find pose parameters Θ that maximize

$$\mathcal{E}(\Theta) = E(\Theta) - w_l E_{lim}(\mathcal{M}\Theta) - w_a E_{acc}(\Theta)$$

SoG similarity over all Gaussians and cameras

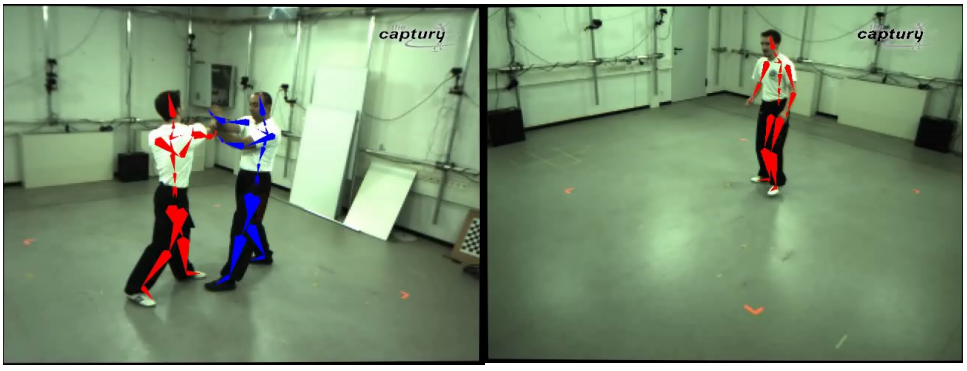
Joint limits

acceleration penalty

- Energy function smooth
- Analytic derivatives convenient to compute
- Occlusion conveniently handled
- Efficient gradient-based optimization

Christian Theobalt

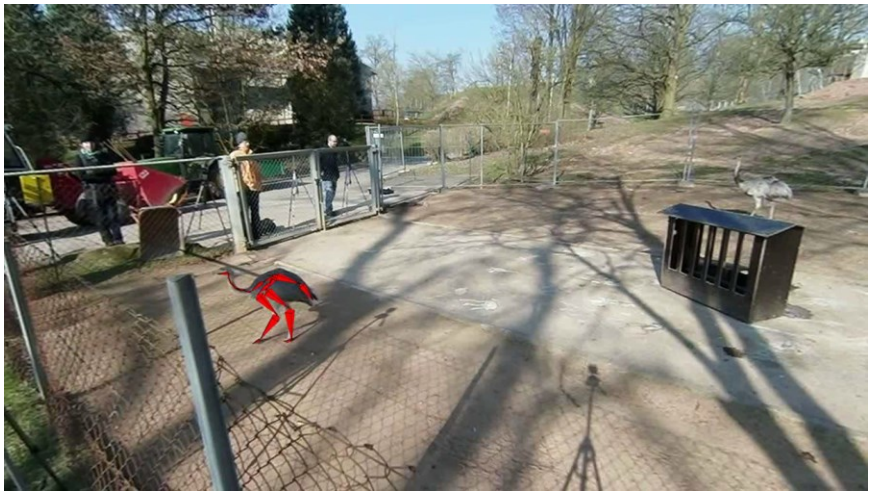
More Results



Min nr. of cameras: 5
Pose estimation > 30 fps

Christian Theobalt

Space-time Reconstruction

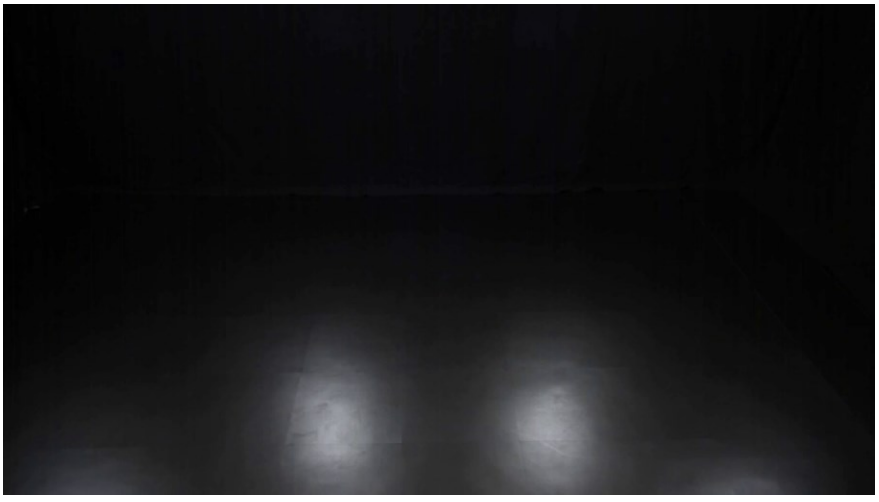


8 GoPros, 30 fps, not frame synchronized

[Elhayek et al., CVPR 2012]

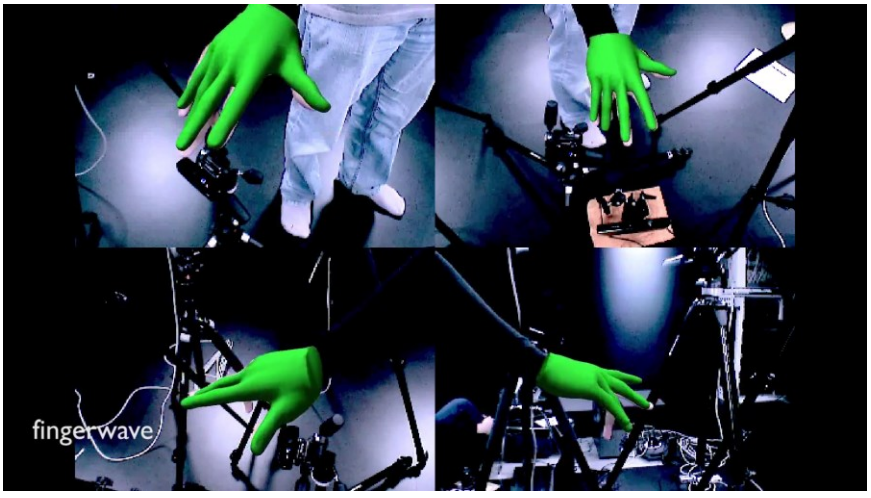
Christian Theobalt

www.thecaptury.com



Christian Theobalt

Extension to Hands



S. Sridhar, A. Oulasvirta, and C. Theobalt, *Interactive Markerless Articulated Hand Motion Tracking using RGB and Depth Data*, ICCV 2013

Christian Theobalt

Real-time Hand-Tracking: Setup

5 video cameras

1 Creative TOF Camera

Christian Theobalt

Real time Hand Tracking with Single Depth Camera

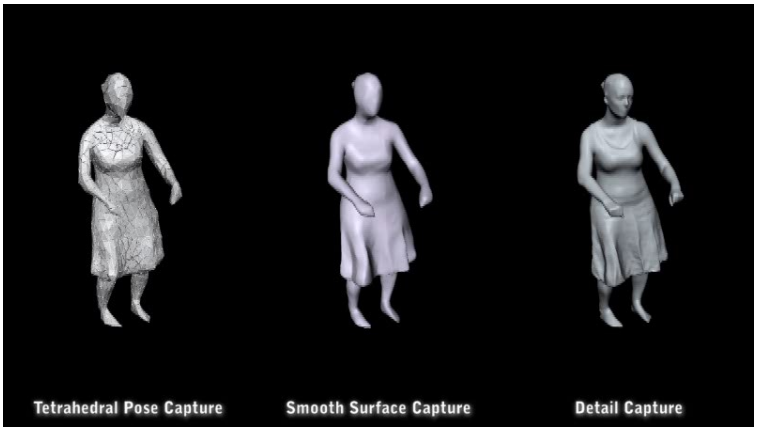
Christian Theobalt

How can we reconstruct more detail in uncontrolled environments ?

Christian Theobalt

Current Methods

- Limit to amount of captured detail



Christian Theobalt

Exploit Shading / Lighting Cues

- Shading / lighting effects around high frequency features

[http://www.verycoolphotoblog.com/2012/11/05/old-man-3]

Related Work

Photometric Stereo

[Woodham '80, Hernandez et al. 2007...]

Shape-from-Shading

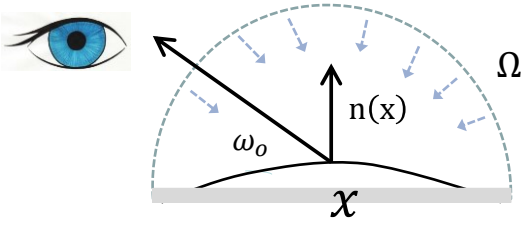
[Zhang '99, Prados et al. 2005...]

Dynamic Photometric Stereo

[Vlasic et al., SIGASIA '09]

Shape Refinement – Inverse Rendering

- Reflectance Equation (Kajiya'86)

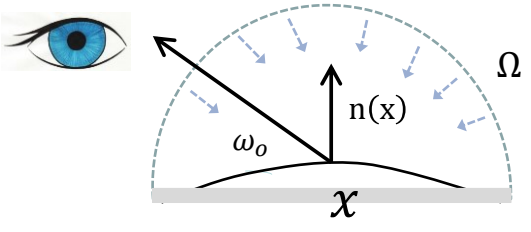


$$B(x,\omega_o) = \int_{\Omega} \rho(\omega_i,\omega_o) L(\omega_i) V(x,\omega_i)\ \max(\omega_i \cdot \mathrm{n}(\mathrm{x}), 0)\, \mathrm{d}\omega_i$$

Rendering

Shape Refinement – Inverse Rendering

- Reflectance Equation (Kajiya'86)



$$B(x,\omega_o) = \int_{\Omega} \rho(\omega_i,\omega_o) L(\omega_i) V(x,\omega_i)\ \max(\omega_i \cdot \mathrm{n}(\mathrm{x}), 0)\, \mathrm{d}\omega_i$$

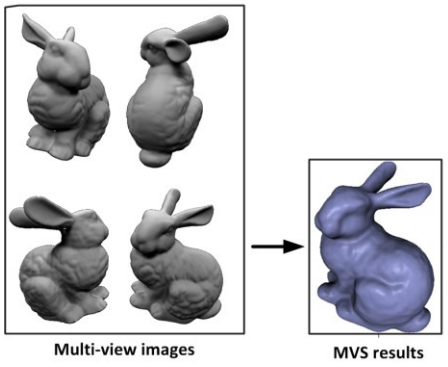
Rendering

Inverse rendering: find illumination, reflectance, shape

[Wu et al. CVPR2011, ICCV 2011, ECCV2012]

Christian Theobalt

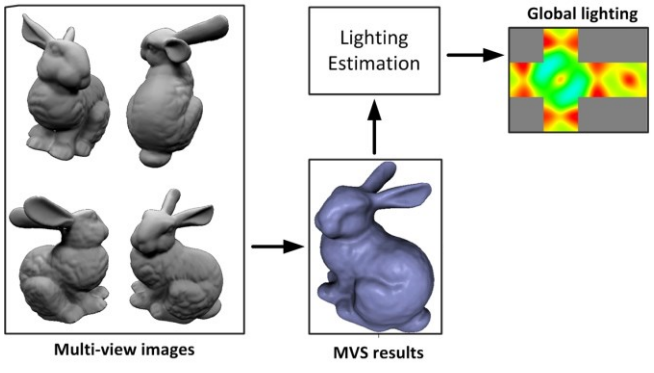
Shading-based Static Scene Refinement



[Wu et al. CVPR 2011]

Christian Theobalt

Shading-based Static Scene Refinement



[Wu et al. CVPR 2011]

Christian Theobalt

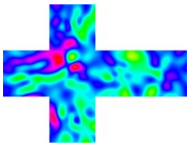
Step I: Lighting Estimation

$$\hat{l} = \arg\min_{l} \sum_{i} \sum_{c} \left| l \cdot t_i - I_c(P_c(x_i)) \right|$$

Coefficients of lighting

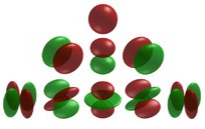
Rendering equation in spherical harmonics

image intensity



Estimated

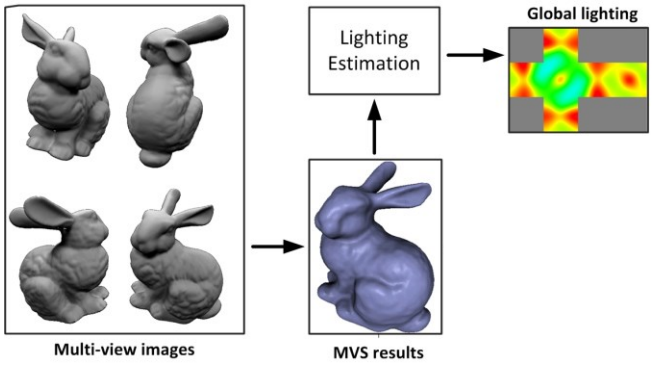
Environment map



Spherical harmonics ([Sloan et al. '02])

Christian Theobalt

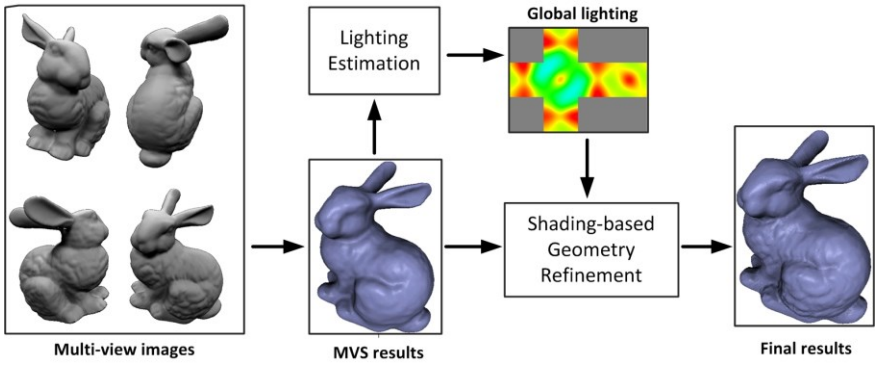
Shading-based Static Scene Refinement



[Wu et al. CVPR 2011]

Christian Theobalt

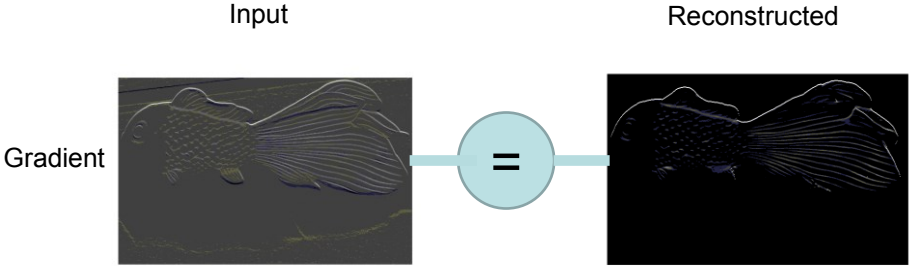
Shading-based Static Scene Refinement



[Wu et al. CVPR 2011]

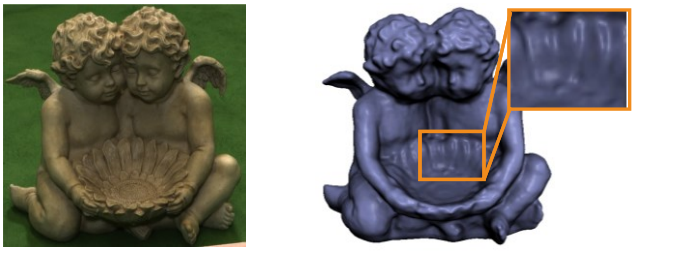
Christian Theobalt

Shading-based Geometry Refinement

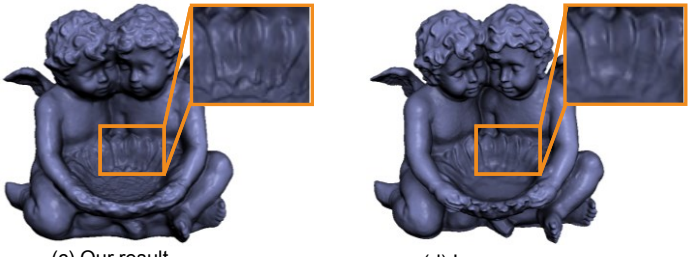


Christian Theobalt

Combine Stereo + Shading



(a) Captured image (b) MVS result

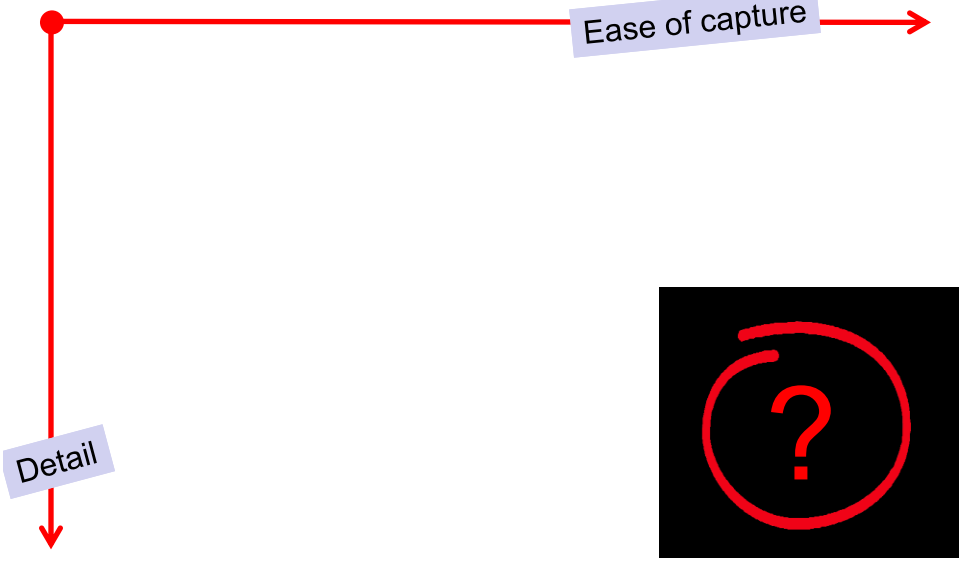


(c) Our result (d) Laser scan

[Wu et al. CVPR 2011]

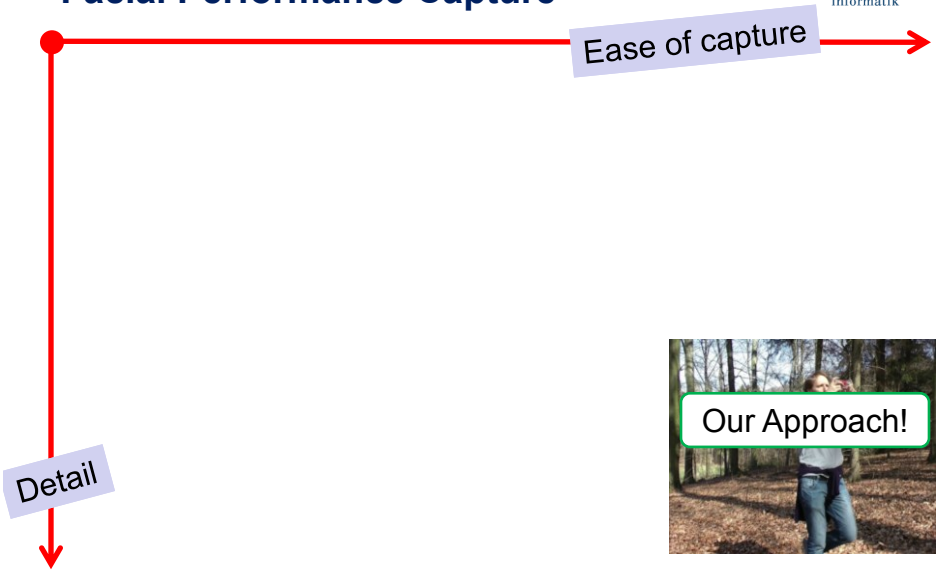
Christian Theobalt

Facial Performance Capture



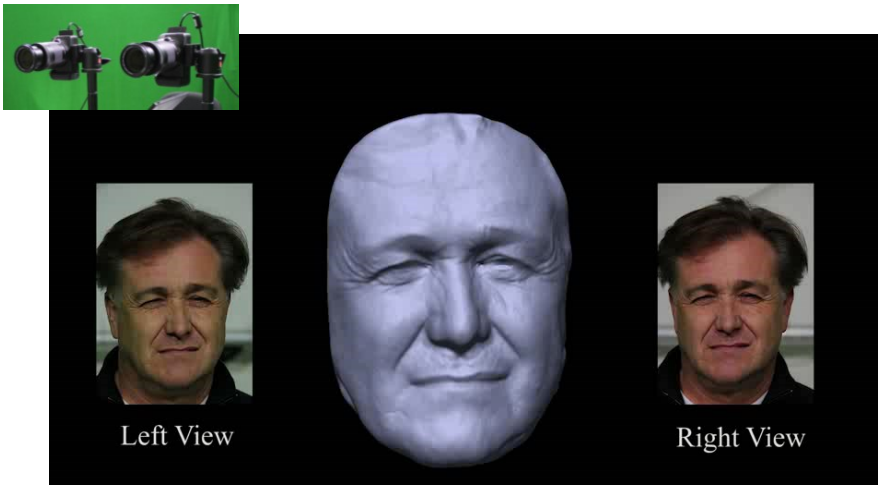
Christian Theobalt

Facial Performance Capture



Christian Theobalt

High-quality Binocular Face Capture



L.Valgaerts, C. Wu, A. Bruhn, H.-P. Seidel, C. Theobalt, Binocular Facial Performance Capture under Uncontrolled Illuminationm, SIGGRAPH Asia 2012

Christian Theobalt

Method Overview

max planck institut informatik

Left view

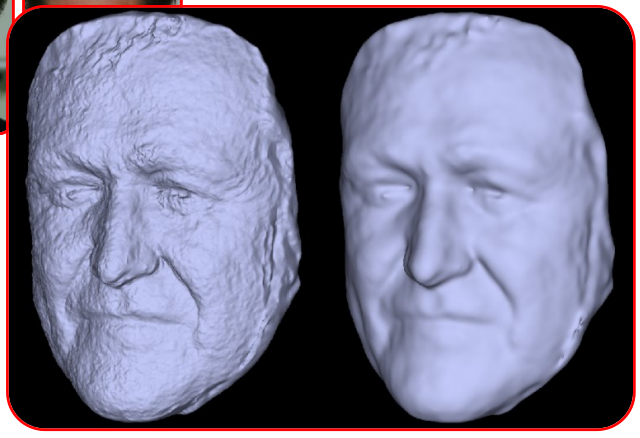
Right view

Christian Theobalt

Method Overview

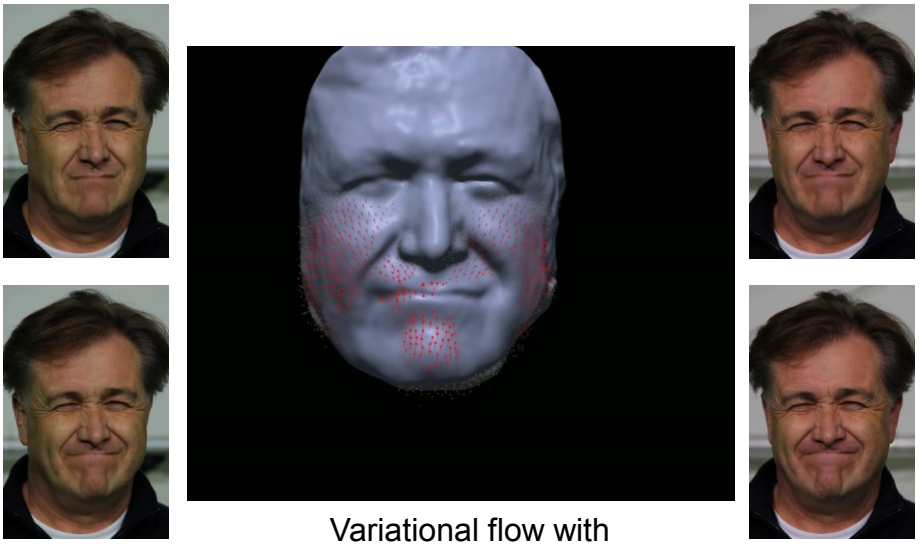
max planck institut informatik

Template reconstruction



Christian Theobalt

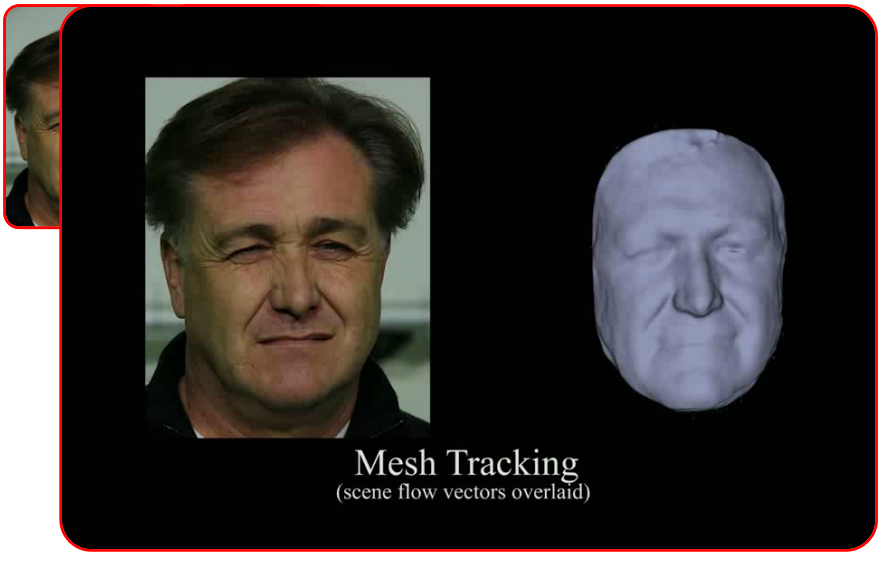
Scene Flow Result



Variational flow with
structure-aware regularization

Christian Theobalt

Method Overview



Christian Theobalt

Method Overview

Christian Theobalt

Dynamic Shape Refinement

Previous frame

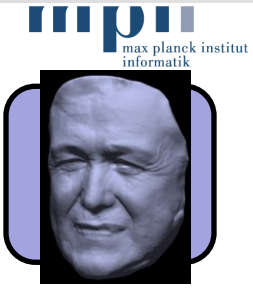
Fine Scale Shape?

Lighting?

Current frame

Christian Theobalt

Dynamic Shape Refinement



Previous frame

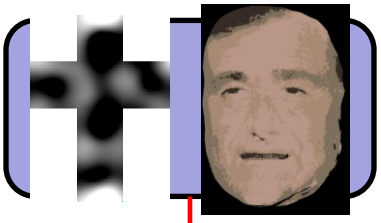
Current frame

⇒ Compute lighting + shape jointly (extends [Wu et al. 2011])

Christian Theobalt

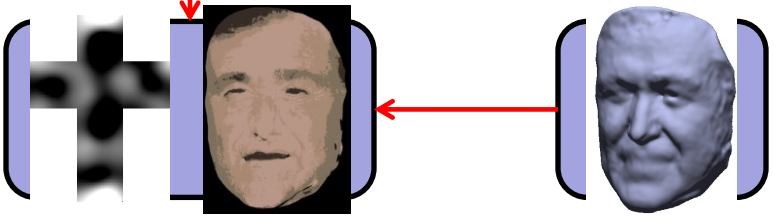
Dynamic Shape Refinement

Previous frame



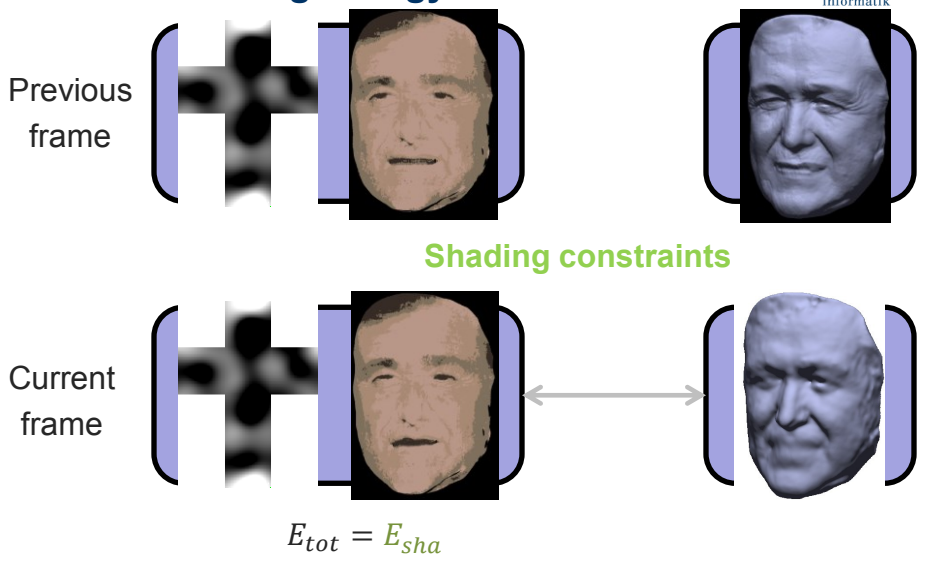
Lighting estimation

Current frame



Christian Theobalt

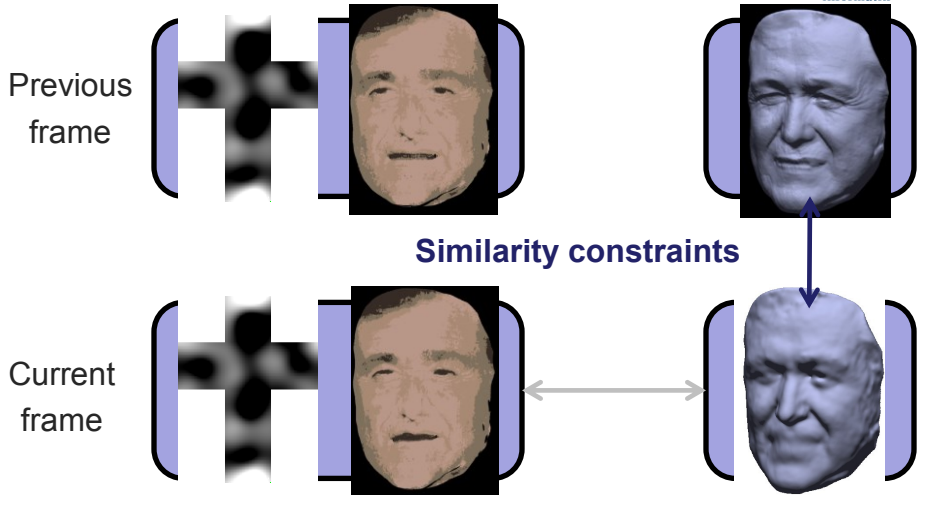
Novel Shading Energy



$$E_{tot} = E_{sha}$$

Christian Theobalt

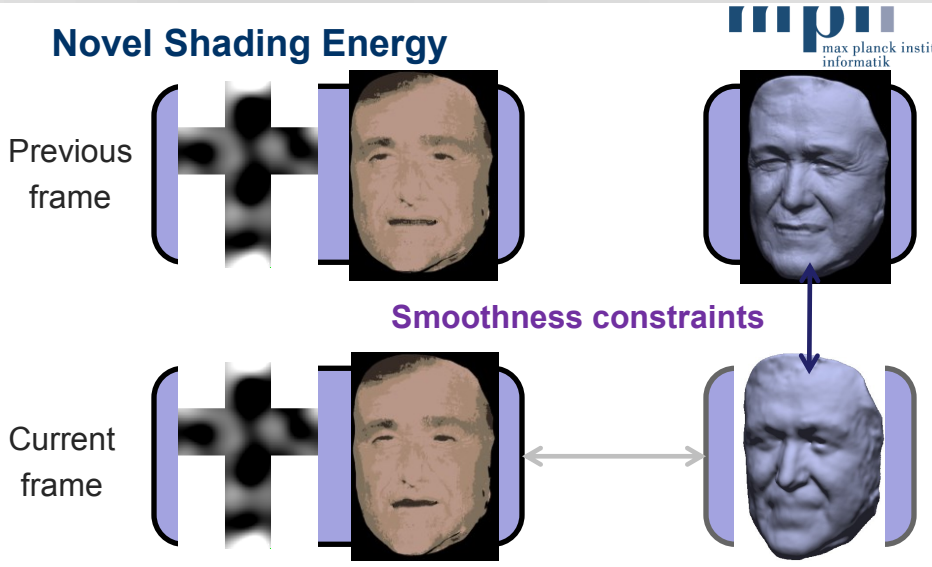
Novel Shading Energy



$$E_{tot} = E_{sha} + \alpha E_{sim}$$

Christian Theobalt

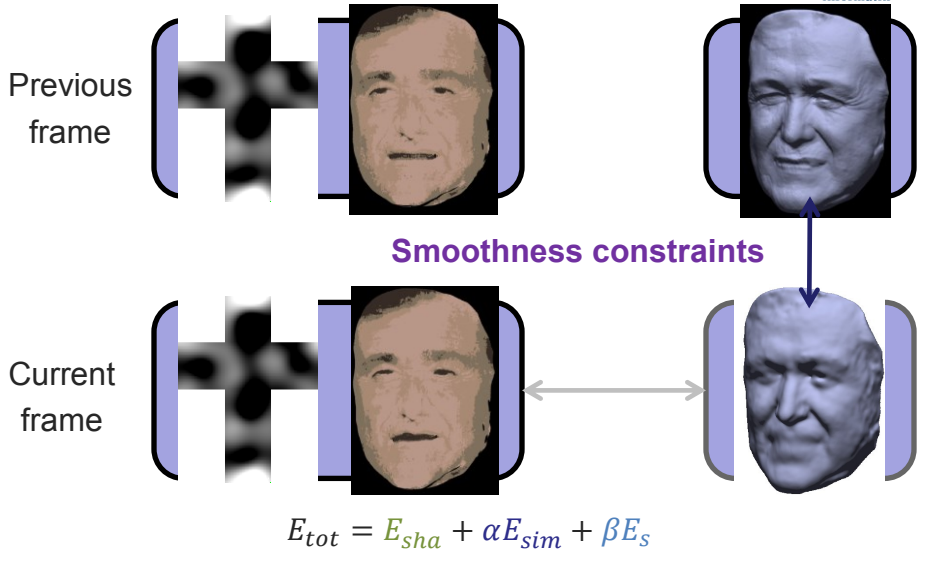
Novel Shading Energy



$$E_{tot} = E_{sha} + \alpha E_{sim} + \beta E_s$$

Christian Theobalt

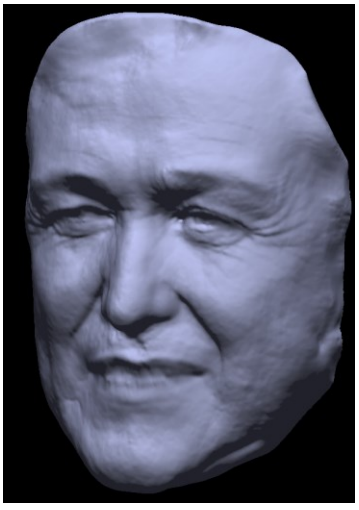
Novel Shading Energy



$$E_{tot} = E_{sha} + \alpha E_{sim} + \beta E_s$$

Christian Theobalt

Refinement Result



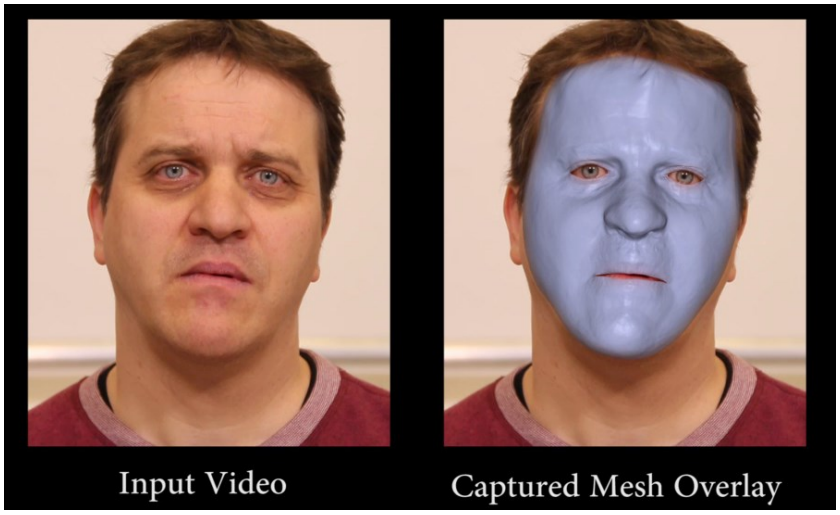
Christian Theobalt

Result - outdoors, changing lighting

Stereo pair of GoPro bike helmet cams, handheld

Christian Theobalt

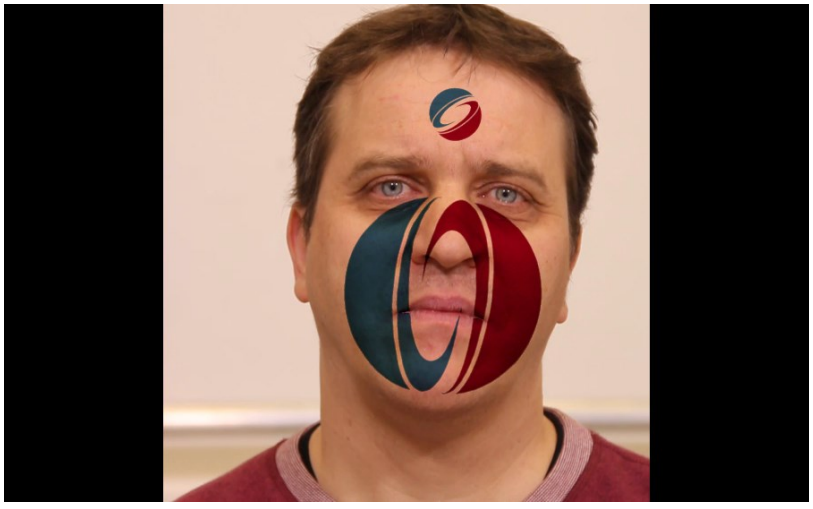
Monocular Face Reconstruction



P. Garrido, L. Valgaerts, C. Wu, C. Theobalt, Reconstructing Detailed Dynamic Face Geometry from Monocular Video, SIGGRAPH Asia 2013

Christian Theobalt

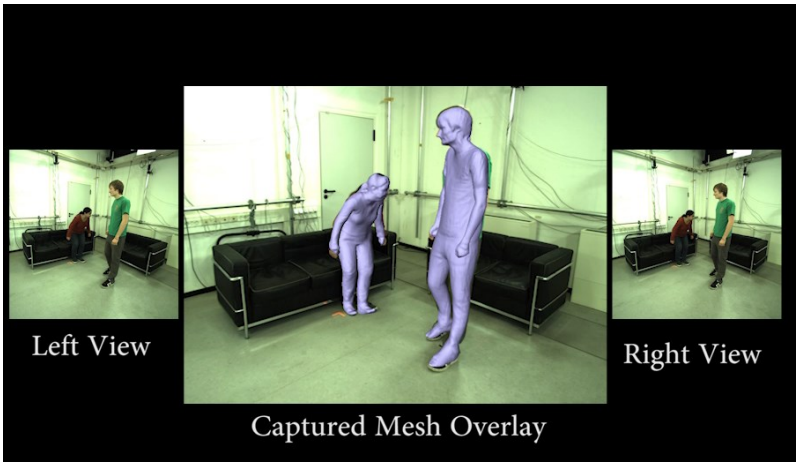
Monocular Face Reconstruction



P. Garrido, L. Valgaerts, C. Wu, C. Theobalt, Reconstructing Detailed Dynamic Face Geometry from Monocular Video, SIGGRAPH Asia 2013

Christian Theobalt

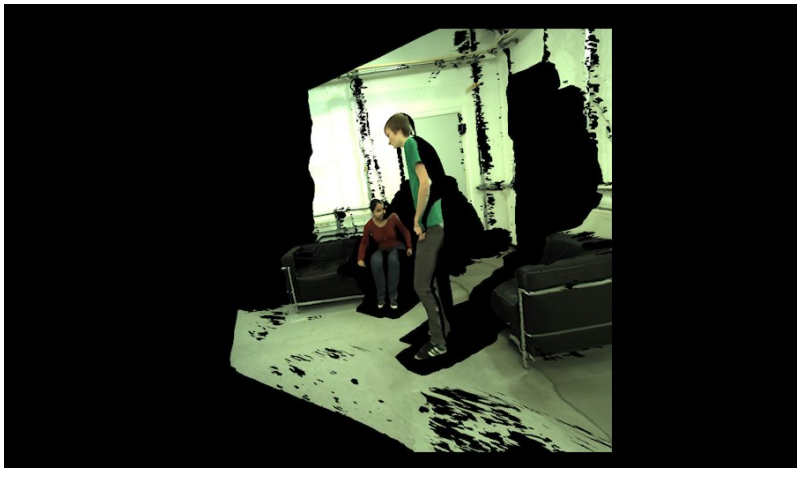
On-Set Performance Capture with a Stereo Camera



C. Wu, C. Stoll, L. Valgaerts, C. Theobalt, On-set Performance Capture of Multiple Actors with a Stereo Camera, SIGGRAPH Asia 2013

Christian Theobalt

On-Set Performance Capture with a Stereo Camera



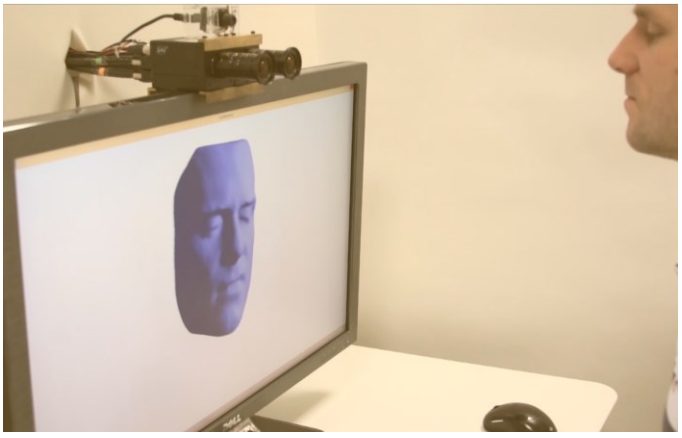
C. Wu, C. Stoll, L. Valgaerts, C. Theobalt, On-set Performance Capture of Multiple Actors with a Stereo Camera, SIGGRAPH Asia 2013

Christian Theobalt

Real-time Deformable Mesh Tracking

max planck institut informatik

Surface Deformation



M. Zollhoefer, M Niessner, S. Izadi, C. Rehmann, C. Zach, M. Fisher, C. Wu, A. Fitzgibbon, C. Loop, C. Theobalt, M. Stamminger, Real-time Non-rigid Reconstrution Using an RGB-D Camera, ACM TOG (SIGGRAPH 2014)

Christian Theobalt

Modifying Captured Content

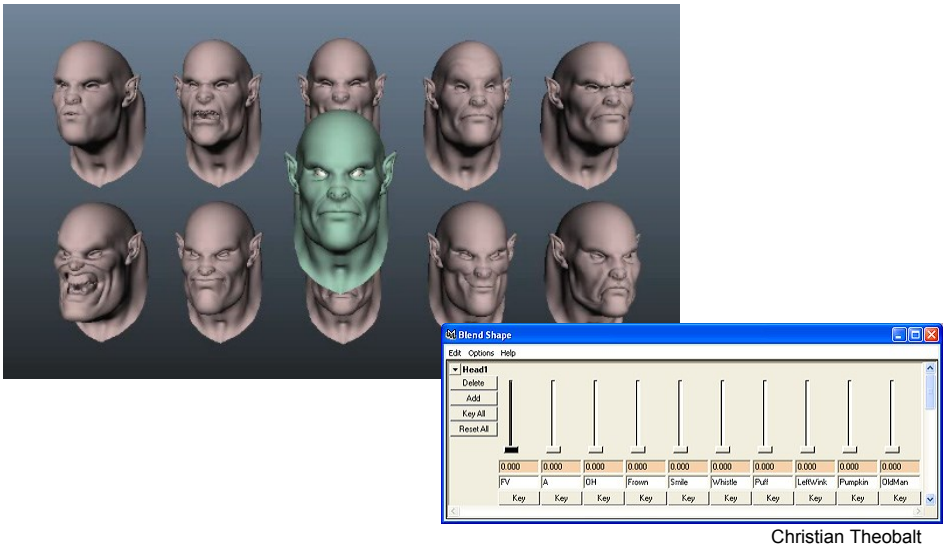
Christian Theobalt

Sparse Localized Deformation Components

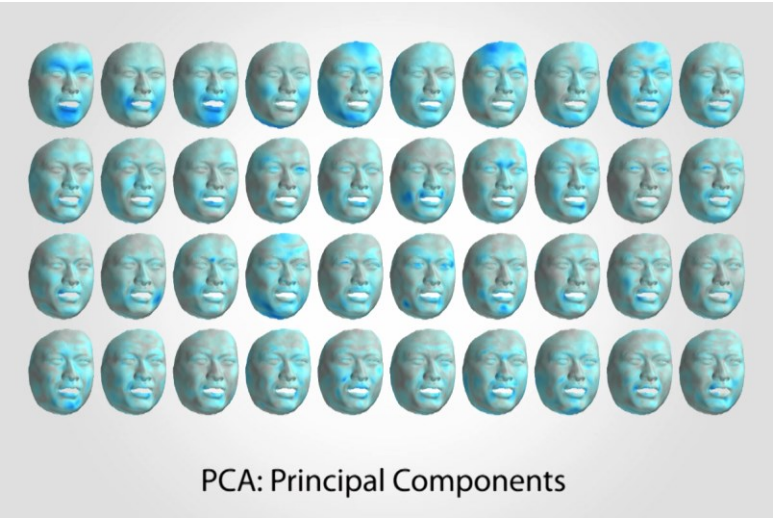
T. Neumann, K. Varanasi, S. Wenger, M. Wacker, M. Magnor, C. Theobalt,
Sparse localized deformation components, *ACM TOG (SIGGRAPH ASIA) 2013*

Blend Shape Model

- Linear blend of artist defined deformations from rest shape



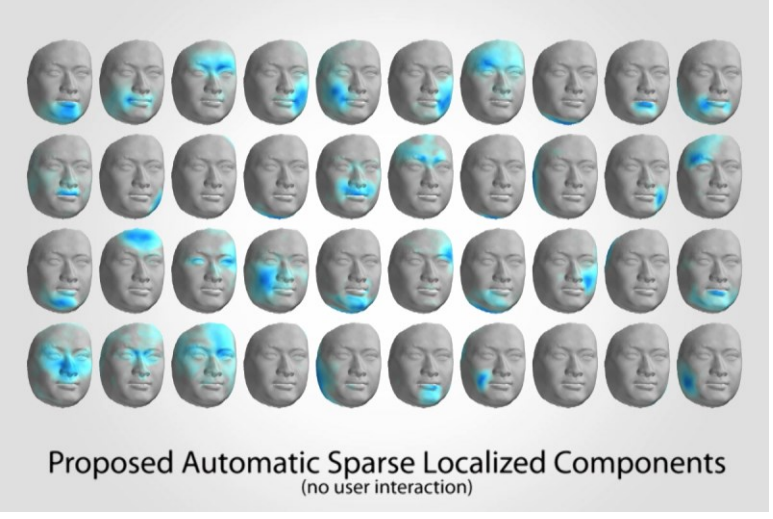
Sparse Localized Deformation Components



T. Neumann, K. Varanasi, S. Wenger, M. Wacker, M. Magnor, C. Theobalt,
Sparse localized deformation components, *ACM TOG (SIGGRAPH ASIA) 2013*

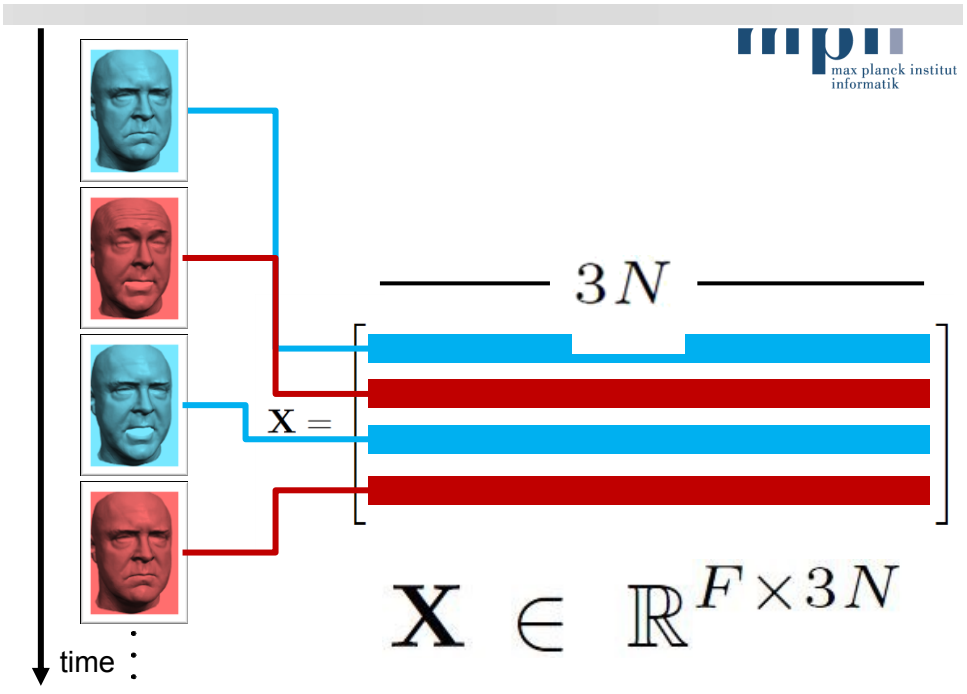
Christian Theobalt

Sparse Localized Deformation Components

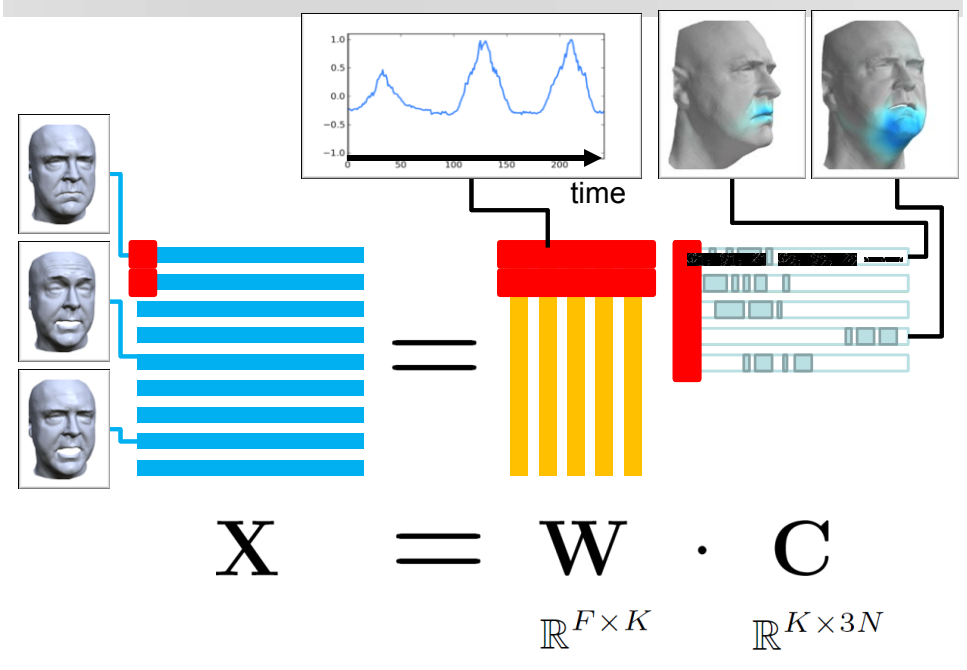


T. Neumann, K. Varanasi, S. Wenger, M. Wacker, M. Magnor, C. Theobalt,
Sparse localized deformation components, *ACM TOG (SIGGRAPH ASIA) 2013*

Christian Theobalt



Christian Theobalt



Christian Theobalt

$$\min_{\mathbf{W},\mathbf{C}} \|\mathbf{X} - \mathbf{W}\cdot\mathbf{C}\|_{\mathrm{F}}^2 + \sum_{k=1}^{K}\sum_{i=1}^{N} \|\mathbf{c}_k^{(i)}\|_2 \quad \text{subject to} \quad \max(|\mathbf{W}_{:,k}|) = 1$$

Data Term

Christian Theobalt

$$+ \sum_{k=1}^{K}\sum_{i=1}^{N} \|\mathbf{c}_k^{(i)}\|_2$$

Sparse, but **not localized**

Christian Theobalt

$$+\sum_{k=1}^{K}\sum_{i=1}^{N}\Lambda_k^{(i)}\|\mathbf{c}_k^{(i)}\|_2$$

Sparse and **localized**

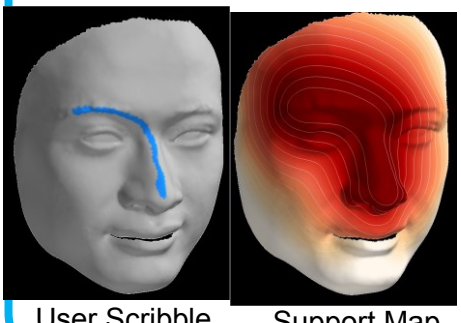
Christian Theobalt

Support Maps

- Geodesic distance to center vertices (on rest shape)

User Interaction

(rough binary mask)

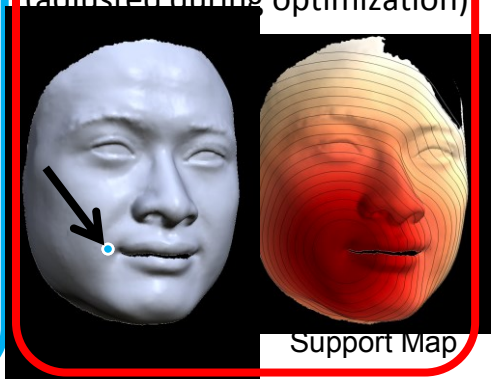


User Scribble

Support Map

Automatic

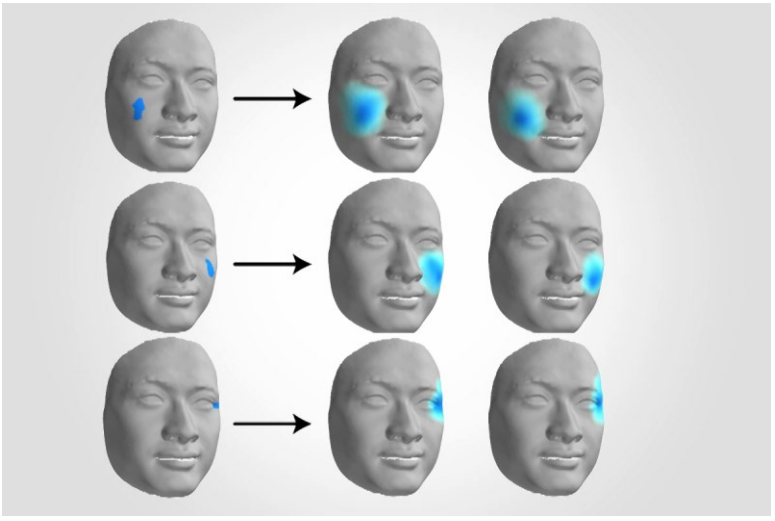
(adjusted during optimization)



Support Map

Christian Theobalt

Sparse Localized Deformation Components

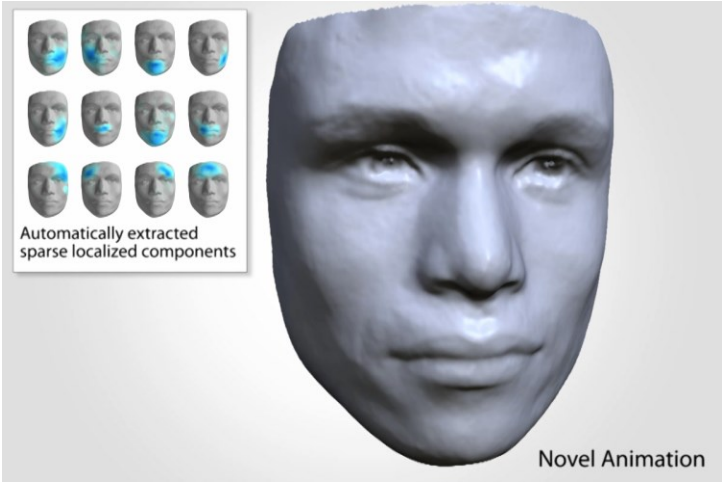


T. Neumann, K. Varanasi, S. Wenger, M. Wacker, M. Magnor, C. Theobalt,
Sparse localized deformation components, *ACM TOG (SIGGRAPH ASIA) 2013*

Christian Theobalt

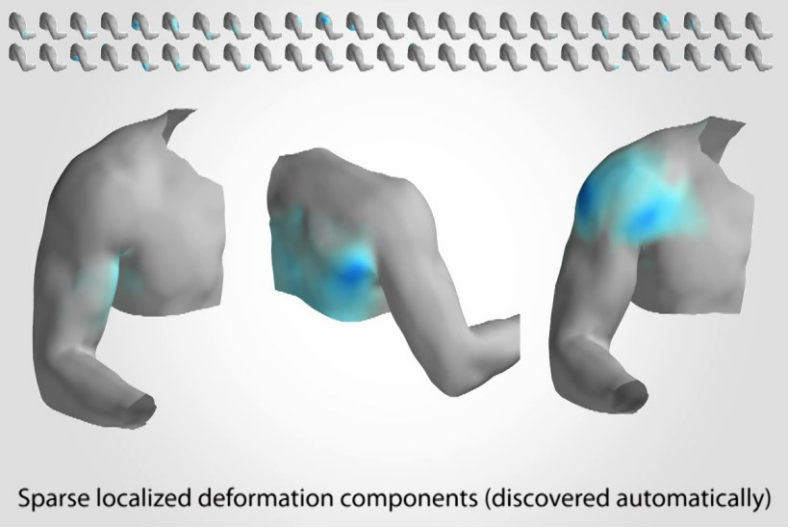
New Animation

Input



Christian Theobalt

“Physiological” Components



Conclusions

- 4D reconstruction has made great progress
- Many open challenges until we capturing reality
- Needed: Rethinking fundations of 4D reconstruction
- First steps:
  - Estimation and exploitation of more expressive light transport models
  - New parameterizations of dynamic scenes

Thank You !

gvv.mpi-inf.mpg.de

www.thecaptury.com

Thanks to:

ERC Starting Grant
CapReal

Christian Theobalt