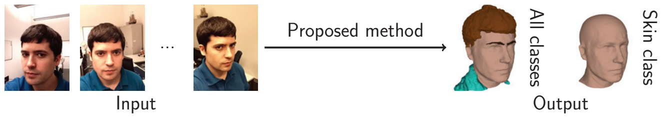
Semantic 3D Reconstruction of Heads

Fabio Maninchedda[1], Christian Häne[2,*], Bastien Jacquet[3,*], Amaël Delaunoy[*], and Marc Pollefeys[1,4]

[1]ETH Zürich — [2]UC Berkeley — [3]Kitware — [4]Microsoft — [*]Work done at ETH Zürich

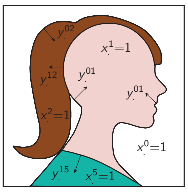
Introduction

We propose a system that automatically reconstructs and semantically segments human heads from images taken in uncontrolled environments using a multi-label shape prior



Formulation

- Labeling of a voxel space into L labels $x_s^i \in [0,1]$ and $\sum_i x_s^i = 1$
- Labels: free space, skin, hair, eyebrows, beard and clothing
- Unary cost contains depth information
- Smoothness cost depends on surface orientation and involved labels (contains shape prior and semantic information)



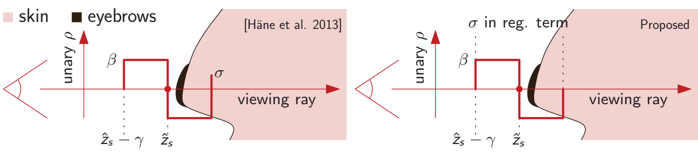
Optimization Problem

$$E(\mathbf{x}, \mathcal{T}) = \sum_{s\in\Omega}\left(\sum_i \rho_s^i(\mathcal{T})x_s^i + \frac{1}{\alpha^2}\sum_{i,j:i<j}\phi_s^{ij}(\mathcal{T}, x_s^{ij} - x_s^{ji})\right)$$

$$\text{s. t. } x_s^i = \sum_j (x_s^{ij})_k, \quad x_s^i = \sum_j (x_{s-e_k}^{ji})_k, \quad \sum_i x_s^i = 1, \quad x_s^i \geq 0, \quad x_s^{ij} \geq 0 \; .$$

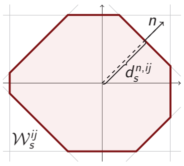
- $x_s^i \in [0,1]$: indicates whether label i is chosen at voxel s
- $y_s^{ij} = x_s^{ij} - x_s^{ji} \in [-1,1]^3$: presence of a surface and its direction
- ρ_s^i: unary term, cost for label i at location s
- ϕ_s^{ij}: convex smoothness term at voxel s for labels i and j
- $e_k \in \mathbb{R}^3$: k-th canonical basis vector
- $\mathcal{T}$: alignment between prior and input data
- α: normalization factor for scale
- Convex in $\mathbf{x}$, not convex in $\mathcal{T}$ (optimized using alternating optimization)

Unary Term



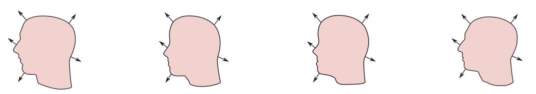
Data Dependent Regularization Term

- Smoothness term $\phi_s^{ij}(\cdot)$ in terms of a convex Wulff shape $\mathcal{W}_s^{ij}$, approximated as intersection of half spaces
- $d_s^{n,ij}$: cost for surface with normal direction n between labels i and j, corresponds to distance of half space boundary to origin
- We have $\phi_s^{ij} = d_s^{n,ij}$
- $d_s^{n,ij} = -\log(P(n_s^{ij}|\mathcal{T},\Gamma))$ is approximated with



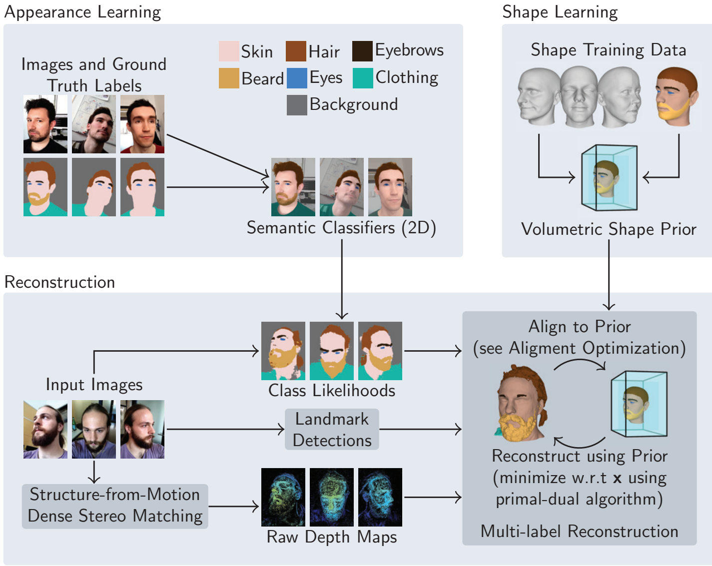
$$\tilde{d}_s^{n,ij} := \underbrace{w_s^{ij}(\mathcal{T},\Gamma)}_{\text{semantic weight}}\Big(\underbrace{-\log(P(n_s^{ij}|\leftrightarrow_s^{ij}))}_{\text{normal based shape prior}}\; \underbrace{-\log(P(\leftrightarrow_s^{ij}|\leftrightarrow_s))}_{\text{transition prior}}\; \underbrace{-\log(P(\leftrightarrow_s))}_{\text{surface prior}}\Big)$$

- Normal based shape prior: locally, surface normals are similar between different heads



- General Wulff shape replaced by surrogate when all normals point in similar direction

Overview

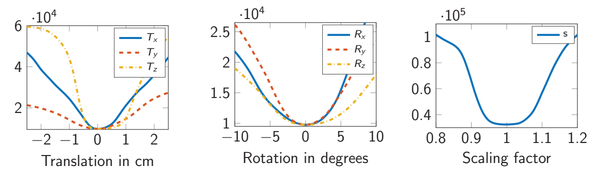


Alignment Optimization

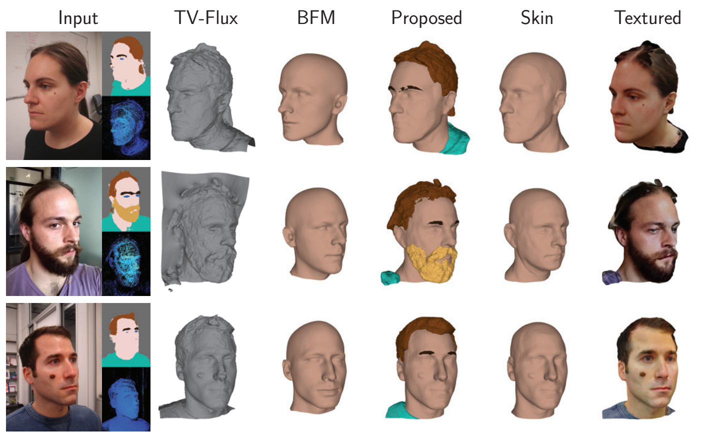
- Align visible geometry, transitions from free space to skin (more generally multiple classes)
- Formulated as energy over surface

$$E_{\text{mesh}}(\mathcal{T}) = \sum_{t\in\mathbb{T}}\phi_t(\mathcal{T}, n_t(\mathcal{T}))\frac{A_t(\mathcal{T})}{\alpha^2} = \sum_{t\in\mathbb{T}}\phi_t(\mathcal{T}, n_t(\mathcal{T}))A_t(\mathcal{I})$$

  - $\mathbb{T}$: set of triangles
  - $\phi_t(\mathcal{T}, n_t(\mathcal{T}))$: smoothness over triangle t with normal $n_t(\mathcal{T})$
  - $A_t(\mathcal{I})$: area of triangle t
- Gradient descent minimization with respect to $\mathcal{T}$
- Energy plots for translation, rotation and scaling for fixed geometry:



Results



Acknowledgment

This project is supported by Grant 16703.1 PFES-ES of CTI Switzerland and the Swiss National Science Foundation under Project Nr. 143422