Photogeometric Scene Flow for High-Detail Dynamic 3D Reconstruction

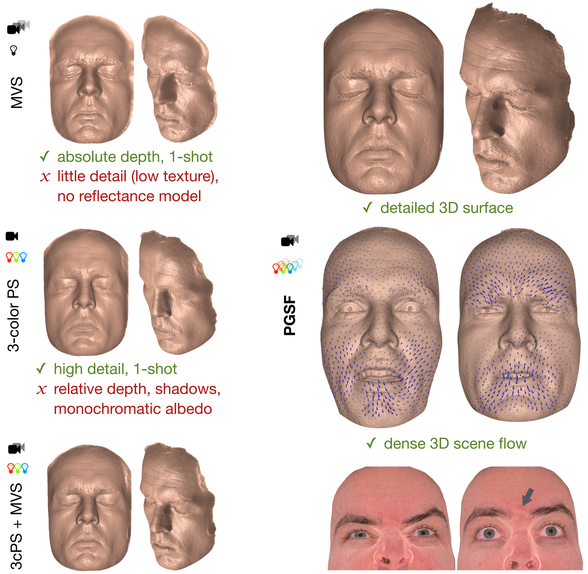
Disney Research[1]

Paulo Gotardo[1], Tomas Simon[2], Yaser Sheikh[2], Iain Matthews[1,2]

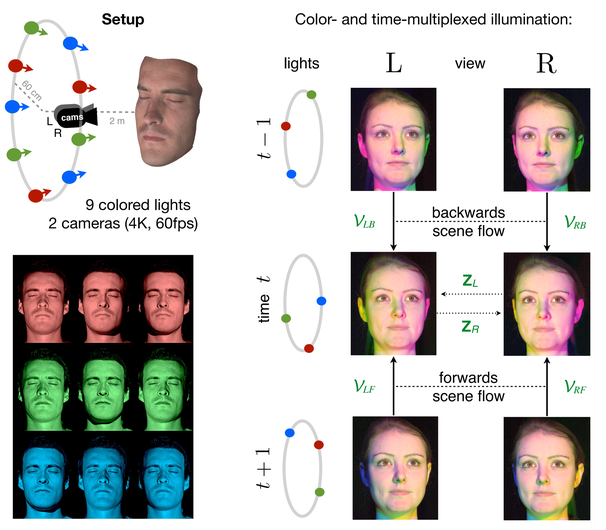
Carnegie Mellon University[2]
The Robotics Institute

Contribution

Photogeometric Scene Flow (PGSF): synergistic solution of photometric stereo (**PS**), multiview stereo (**MVS**) and optical flow (**OF**) under changing illumination



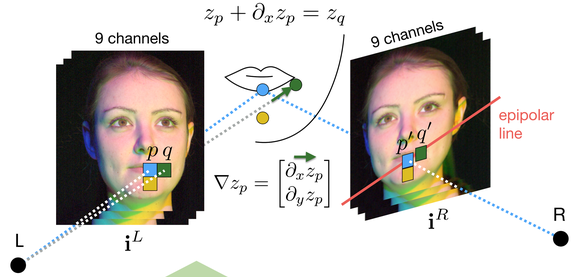
PGSF: reconstruction on 3-frame sliding window



MVS + PS: stereo matching with surface gradients

Update depthmaps with cross-view consistency (occlusion detection)

$$z_p + \partial_x z_p = z_q$$



$$\nabla z_p = \begin{bmatrix} \partial_x z_p \\ \partial_y z_p \end{bmatrix}$$

MVS data term

$$i^L(p) \approx i^R(p + \vec{d}_p(z_p))$$

Disparity: flow along epipolar line

$$\vec{d}_p(z_p) = \mathcal{P}_R\left(\begin{bmatrix} p \\ 1 \end{bmatrix} z_p\right) - p$$

PS spatial term (data driven)

$$\left[i_t(p)\mathbf{l}_{t'}^T - i_{t'}(p)\mathbf{l}_t^T \right] \mathbf{n}_p \approx 0$$

Linear constraint on depthmap

$$\mathbf{n}_p \propto \begin{bmatrix} 1 & & \\ & 1 & \\ \frac{u_p}{f} & \frac{v_p}{f} & \frac{1}{f} \end{bmatrix} \begin{bmatrix} \nabla z_p \\ z_p \end{bmatrix}$$

Linearized MVS term for **iterative updates**:

$$\nabla i^R(p')^T \frac{\partial \vec{d}_p}{\partial z_p} \Delta z_p \approx i^L(p) - i^R(p') \qquad p' = p + \vec{d}_p(z_p)$$

2 images with 9 aligned channels

surface (normals), calibrated lights

next pyramid level

DONE? → Compute albedo, exit

Matching pixels across views, time: **RGB albedo consistency**

input — Lambertian model — calibrated

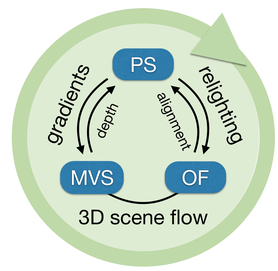
$$\begin{bmatrix} i_t(p) \\ i_t(p) \\ i_t(p) \end{bmatrix} = \begin{bmatrix} \alpha_{pt} & & \\ & \alpha_{pt} & \\ & & \alpha_{pt} \end{bmatrix} \begin{bmatrix} \mathbf{l}_t^T \\ \mathbf{l}_t^T \\ \mathbf{l}_t^T \end{bmatrix} \mathbf{n}_p$$

$$\alpha_{pt} \approx \alpha_{p't'} \qquad \alpha_{pt} \approx \alpha_{p't'} \qquad \alpha_{pt} \approx \alpha_{p't'} \qquad t' = t \pm 1$$

$$\frac{i_t(p)}{\mathbf{l}_t^T \mathbf{n}_p} \approx \frac{i_{t'}(p + \vec{v}_p)}{\mathbf{l}_{t'}^T \mathbf{n}_p}$$

basic PGSF constraint couples surface and flow

KEY IDEA



mutually beneficial solutions replace usual pipeline

Coarse-to-fine optimization using image pyramids

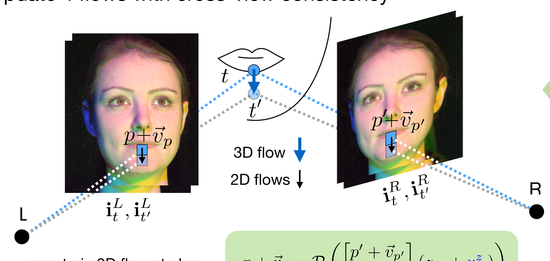
BEGIN

flow is nil at coarsest level

3D Scene Flow

Update 4 flows with cross-view consistency

fix extra degree of freedom in 2D flows



constrain 2D flows to be projections of 3D flow:

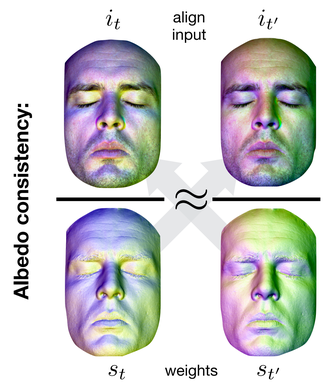
$$p + \vec{v}_p = \mathcal{P}_L\left(\begin{bmatrix} p' + \vec{v}_{p'} \\ 1 \end{bmatrix} (z_{p'} + v^z_{p'})\right)$$

2D pt — 3D pt — z-flow

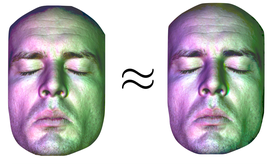
Cross-shaded images for OF + surface, L-R pixel correspondences

Relighting

Shading weights: $s_t(p) = \mathbf{l}_t^T \mathbf{n}_p$



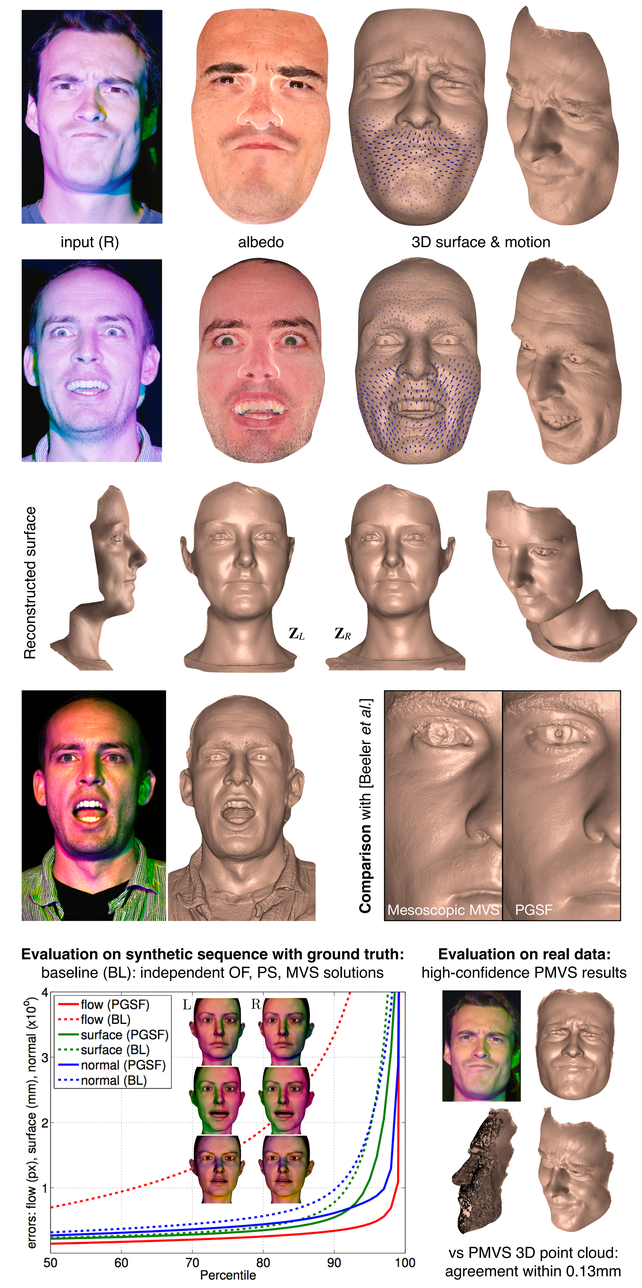
Cross-shaded (relit) images:



$$s_{t'}(p) i_t(p) \approx i_{t'}(p + \vec{v}_p) s_t(p)$$

Relit space-time gradient for TV-L1 OF update:

$$s_t(p) \nabla i_{t'}(p + \vec{v}_p) \Delta \vec{v}_p \approx s_{t'}(p) i_t(p) - i_{t'}(p + \vec{v}_p) s_t(p)$$

Results

see also www.disneyresearch.com/publication/pgsf