Fight Ill-Posedness with Ill-Posedness:

Single-shot Variational Depth Super-Resolution from Shading

Haefner et al.

Seminar: An Overview of Methods for Accurate Geometry Reconstruction

Fabian Schöttl

Garching, 27.11.2019

Motivation

- RGB-D cameras often output high-resolution color but low-resolution depth data

Motivation

- RGB-D cameras often output high-resolution color but low-resolution depth data
- Can we get high-resolution depth data?
  - How can we interpolate missing depth data to fit color data?
  - Using only one RGB-D image

Motivation

- RGB-D cameras often output high-resolution color but low-resolution depth data
- Can we get high-resolution depth data?
  - How can we interpolate missing depth data to fit color data?
  - Using only one RGB-D image

- Two solutions to this problem
  - Both are ill-posed…

Approach 1 – Single depth image super-resolution

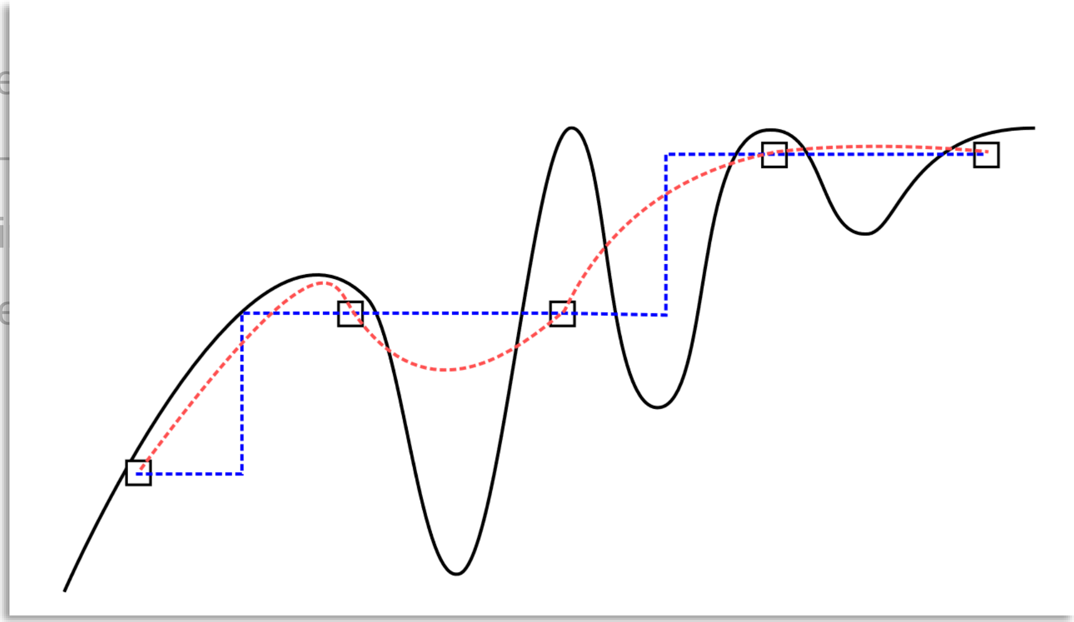
- Estimate high resolution depth map z from low resolution input z_0
- Solve $z_0 = K\, z + \eta_z$ for z

Approach 1 – Single depth image super-resolution

- Estimate high resolution depth map z from low resolution input z_0
- Solve $z_0 = K\,z + \eta_z$ for z
- Requires inverting K
  - Ill-posed since K maps from high-res domain to low-res

Approach 1 – Single depth image super-resolution

- Estimate high re
- Solve $z_0 = K\,z$ -
- Requires inverti
  - Ill-posed since



Approach 1 – Single depth image super-resolution

- Estimate high resolution depth map z from low resolution input z_0
- Solve $z_0 = K\,z + \eta_z$ for z
- Requires inverting K
  - Ill-posed since K maps from high-res domain to low-res

Approach 1 – Single depth image super-resolution

- Estimate high resolution depth map z from low resolution input z_0
- Solve $z_0 = K\,z + \eta_z$ for z
- Requires inverting K
  - Ill-posed since K maps from high-res domain to low-res

- **Use high-res RGB-image to guide interpolation between depth values**

Approach 2 – Shape from shading

- Invert image formation model
  - $I = R(z|l, \rho) + \eta_I$

Approach 2 – Shape from shading

- Invert image formation model
  - $I = R(z|l, \rho) + \eta_I$

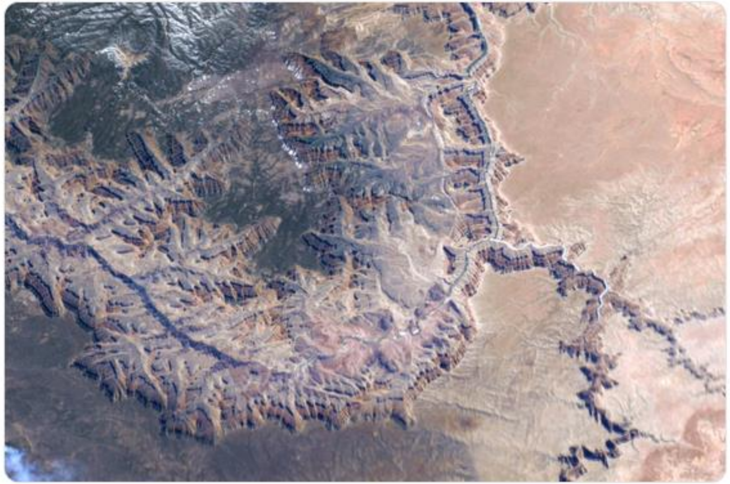
- Previous methods can only estimate magnitude of depth gradient
  - $|\nabla z| = \sqrt{\frac{1}{I^2} - 1}$
  - Ill-posed since gradient direction is ambiguous

App

- I
- F … itude of depth gradient
- … guous

Jack Fischer @Astro2fish — Folgen

Crazy--canyons look like plateaus from space. We should all remember a person's perspective can make the very same thing look very different

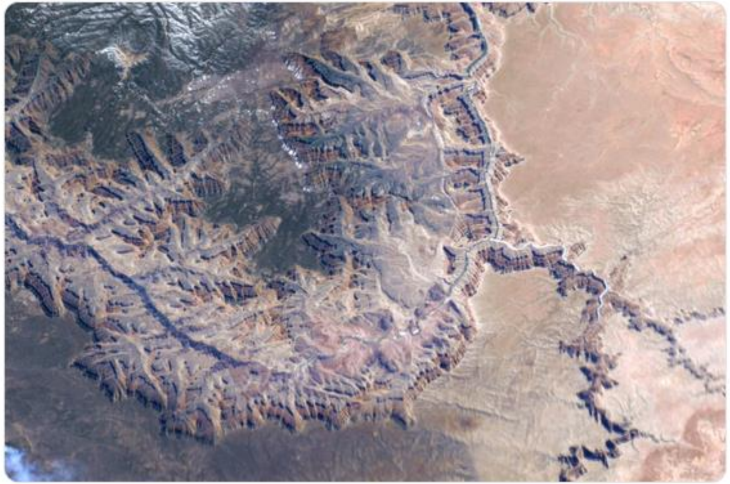


09:00 - 3. Mai 2017

Jack Fischer @Astro2fish

Folgen

Crazy--canyons look like plateaus from space. We should all remember a person's perspective can make the very same thing look very different



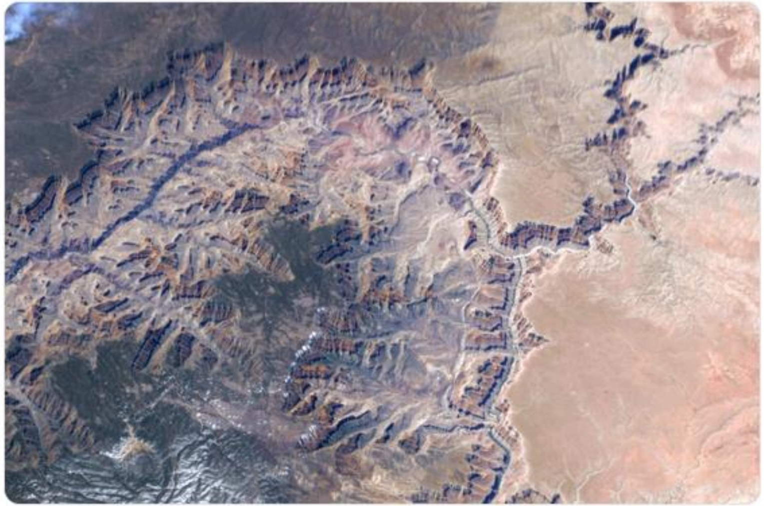
09:00 - 3. Mai 2017

Phil Plait @BadAstronomer

Folgen

Antwort an @Astro2fish

Flip the picture upside-down! @astro2fish twitter.com/Astro2fish/sta ...



09:02 - 3. Mai 2017

Approach 2 – Shape from shading

- Invert image formation model
  - $I = R(z|l, \rho) + \eta_I$

- Previous methods can only estimate magnitude of depth gradient
  - $|\nabla z| = \sqrt{\frac{1}{I^2} - 1}$
  - Ill-posed since gradient direction is ambiguous

Approach 2 – Shape from shading

- Invert image formation model
  - $I = R(z|l, \rho) + \eta_I$

- Previous methods can only estimate magnitude of depth gradient
  - $|\nabla z| = \sqrt{\frac{1}{I^2} - 1}$
  - Ill-posed since gradient direction is ambiguous
- **Use low-res depth map to guide shape from shading**

Fight ill-posedness with ill-posedness

Joint depth super-resolution and shape-from-shading

- Solve both problems at the same time
  - $z_0 = K\, z + \eta_z$
  - $I = R(z|l, \rho) + \eta_I$

Joint depth super-resolution and shape-from-shading

- Solve both problems at the same time
  - $z_0 = K\, z + \eta_z$
  - $I = R(z|l, \rho) + \eta_I$

$$\max_{z,\rho,l} P(z, \rho, l | z_0, I)$$

Joint depth super-resolution and shape-from-shading

- Solve both problems at the same time
  - $z_0 = K\,z + \eta_z$
  - $I = R(z|l,\rho) + \eta_I$

$$\max_{z,\rho,l} P(z,\rho,l|z_0,I)$$

$$= \max_{z,\rho,l} \frac{P(z_0,I|z,\rho,l)\,P(z,\rho,l)}{P(z_0,I)}$$

$$= \max_{z,\rho,l} P(z_0,I|z,\rho,l)\,P(z,\rho,l)$$

Joint depth super-resolution and shape-from-shading

- Solve both problems at the same time
  - $z_0 = K\, z + \eta_z$
  - $I = R(z|l, \rho) + \eta_I$

$$\max_{z,\rho,l} P(z, \rho, l | z_0, I)$$

$$= \max_{z,\rho,l} \frac{P(z_0, I | z, \rho, l)\, P(z, \rho, l)}{P(z_0, I)}$$

$$= \max_{z,\rho,l} \underbrace{P(z_0, I | z, \rho, l)}_{likelihood} \underbrace{P(z, \rho, l)}_{prior}$$

Likelihood

- $P(z_0, I|z, \rho, l) = P(z_0|z)P(I|z, \rho, l)$
  - RGB-D sensors capture depth and color independently

Likelihood

- $P(z_0, I|z, \rho, l) = P(z_0|z)P(I|z, \rho, l)$

Likelihood

- $P(z_0, I|z, \rho, l) = P(z_0|z)P(I|z, \rho, l)$
- Recall: $z_0 = K\, z + \eta_z$
- $P(z_0|z) \propto \exp\left\{-\frac{\|K\, z - z_0\|_{\ell^2}^2}{2\sigma_z^2}\right\}$

Likelihood

- $P(z_0, I|z, \rho, l) = P(z_0|z)P(I|z, \rho, l)$
- Image formation model
  - $I = \left(l \cdot m_{z,\nabla z}\right) \rho$
    - with $l \in \mathbb{R}^4$ and $m_{z,\nabla z} = \begin{bmatrix} n_z \\ 1 \end{bmatrix}$
    - Assume achromatic, directional (+ ambient) lighting
- $P(I|z, \rho, l) \propto \exp\left\{-\frac{\left\|\left(l \cdot m_{z,\nabla z}\right)\rho - I\right\|_{\ell^2}^2}{2\sigma_I^2}\right\}$

Likelihood

- $P(z_0, I | z, \rho, l) = P(z_0 | z) P(I | z, \rho, l)$

$$P(z_0 | z) \propto \exp\left\{ -\frac{\| K\, z \ - z_0 \|_{\ell^2}^2}{2\sigma_z^2} \right\}$$

$$P(I | z, \rho, l) \propto \exp\left\{ -\frac{\left\| \left(l \cdot m_{z, \nabla z} \right) \rho \ - \ I \right\|_{\ell^2}^2}{2\sigma_I^2} \right\}$$

Prior

- $P(z, \rho, l) = P(z)\, P(\rho)\, P(l)$
    - Assume depth, reflectance and lighting are independent

Prior

- $P(z, \rho, l) = P(z)\, P(\rho)\, P(l)$
  - Assume depth, reflectance and lighting are independent
- $P(l) = constant$

Prior

- $P(z, \rho, l) = P(z)\, P(\rho)\, P(l)$
  - Assume depth, reflectance and lighting are independent

- $P(l) = constant$
- $P(z) \propto \exp\left\{-\frac{\left\| dA_{z,\nabla z} \right\|_{\ell 1}}{\alpha}\right\}$, minimal surface prior

Prior

- $P(z, \rho, l) = P(z)\, P(\rho)\, P(l)$
  - Assume depth, reflectance and lighting are independent

- $P(l) = constant$
- $P(z) \propto \exp\left\{-\frac{\left\| dA_{z,\nabla z} \right\|_{\ell 1}}{\alpha}\right\}$, minimal surface prior
- $P(\rho) \propto \exp\left\{-\frac{\|\nabla \rho\|_{\ell 0}}{\beta}\right\}$, assume ρ to be piecewise constant

Joint depth super-resolution and shape-from-shading

- $\max_{z,\rho,l} P(z, \rho, l | z_0, I) = \max_{z,\rho,l} P(I | z, \rho, l)\; P(z_0 | z)\; P(z)\; P(\rho)$

Joint depth super-resolution and shape-from-shading

- $\max_{z,\rho,l} P(z,\rho,l|z_0,I) = \max_{z,\rho,l} P(I|z,\rho,l)\ P(z_0|z)\ P(z)\ P(\rho)$
- $\max_{z,\rho,l} P(z,\rho,l|z_0,I) = \min_{z,\rho,l}\ \left\|\left(l\cdot m_{z,\nabla z}\right)\rho\ -\ I\right\|_{\ell^2}^2 + \mu\ \|K\ z\ - z_0\|_{\ell^2}^2 +$

$$\nu\ \left\|dA_{z,\nabla z}\right\|_{\ell^1} + \lambda\ \|\nabla\rho\|_{\ell^0}$$

Joint depth super-resolution and shape-from-shading

- $\max_{z,\rho,l} P(z,\rho,l|z_0,I) = \max_{z,\rho,l} P(I|z,\rho,l)\, P(z_0|z)\, P(z)\, P(\rho)$
- $\max_{z,\rho,l} P(z,\rho,l|z_0,I) = \min_{z,\rho,l} \left\|\left(l \cdot m_{z,\nabla z}\right)\rho - I\right\|_{\ell^2}^2 + \mu \left\|K z - z_0\right\|_{\ell^2}^2 + \nu \left\|dA_{z,\nabla z}\right\|_{\ell^1} + \lambda \left\|\nabla\rho\right\|_{\ell^0}$
- $P(I|z,\rho,l) \propto \exp\left\{-\frac{\left\|\left(l\cdot m_{z,\nabla z}\right)\rho - I\right\|_{\ell^2}^2}{2\sigma_I^2}\right\}$

Joint depth super-resolution and shape-from-shading

- $\max_{z,\rho,l} P(z,\rho,l|z_0,I) = \max_{z,\rho,l} P(I|z,\rho,l)\ P(z_0|z)\ P(z)\ P(\rho)$

- $\max_{z,\rho,l} P(z,\rho,l|z_0,I) = \min_{z,\rho,l} \left\|\left(l \cdot m_{z,\nabla z}\right)\rho - I\right\|_{\ell^2}^2 + \mu \left\|K\,z - z_0\right\|_{\ell^2}^2 +$
  $\nu \left\|dA_{z,\nabla z}\right\|_{\ell^1} + \lambda \left\|\nabla\rho\right\|_{\ell^0}$

- $P(z_0|z) \propto \exp\left\{-\frac{\|K\,z - z_0\|_{\ell^2}^2}{2\sigma_z^2}\right\}$

Joint depth super-resolution and shape-from-shading

- $\max_{z,\rho,l} P(z,\rho,l|z_0,I) = \max_{z,\rho,l} P(I|z,\rho,l)\, P(z_0|z)\, P(z)\, P(\rho)$

- $\max_{z,\rho,l} P(z,\rho,l|z_0,I) = \min_{z,\rho,l} \left\|\left(l \cdot m_{z,\nabla z}\right)\rho - I\right\|_{\ell^2}^2 + \mu \left\|K z - z_0\right\|_{\ell^2}^2 + \nu \left\|dA_{z,\nabla z}\right\|_{\ell^1} + \lambda \left\|\nabla\rho\right\|_{\ell^0}$

- $P(z) \propto \exp\left\{-\frac{\left\|dA_{z,\nabla z}\right\|_{\ell^1}}{\alpha}\right\}$

Joint depth super-resolution and shape-from-shading

- $\max_{z,\rho,l} P(z,\rho,l|z_0,I) = \max_{z,\rho,l} P(I|z,\rho,l)\, P(z_0|z)\, P(z)\, P(\rho)$

- $\max_{z,\rho,l} P(z,\rho,l|z_0,I) = \min_{z,\rho,l} \left\|\left(l \cdot m_{z,\nabla z}\right)\rho - I\right\|_{\ell^2}^2 + \mu \left\|K z - z_0\right\|_{\ell^2}^2 +$
$$\nu \left\|dA_{z,\nabla z}\right\|_{\ell^1} + \lambda \left\|\nabla\rho\right\|_{\ell^0}$$

- $P(\rho) \propto \exp\left\{-\frac{\left\|\nabla\rho\right\|_{\ell^0}}{\beta}\right\}$

Joint depth super-resolution and shape-from-shading

- $\max_{z,\rho,l} P(z,\rho,l|z_0,I) = \max_{z,\rho,l} P(I|z,\rho,l)\, P(z_0|z)\, P(z)\, P(\rho)$

- $\max_{z,\rho,l} P(z,\rho,l|z_0,I) = \min_{z,\rho,l} \left\|\left(l \cdot m_{z,\nabla z}\right)\rho - I\right\|_{\ell^2}^2 + \mu\, \|K\, z - z_0\|_{\ell^2}^2 +$

$$\nu \left\|dA_{z,\nabla z}\right\|_{\ell^1} + \lambda\, \|\nabla\rho\|_{\ell^0}$$

- With $\mu = \frac{\sigma_I^2}{\sigma_z^2}$, $\nu = \frac{2\,\sigma_I^2}{\alpha}$ and $\lambda = \frac{2\,\sigma_I^2}{\beta}$

Joint depth super-resolution and shape-from-shading

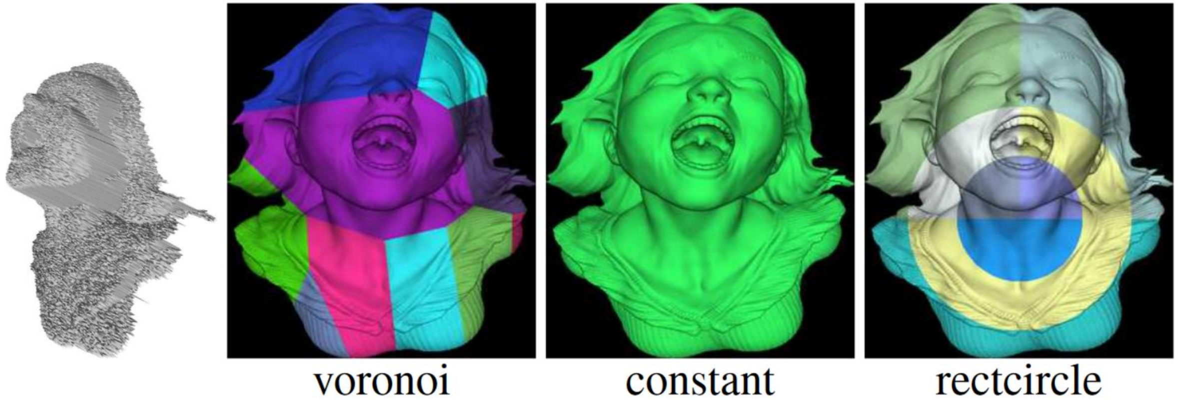
- $\max_{z,\rho,l} P(z,\rho,l|z_0,I) = \max_{z,\rho,l} P(I|z,\rho,l)\, P(z_0|z)\, P(z)\, P(\rho)$

- $\max_{z,\rho,l} P(z,\rho,l|z_0,I) = \min_{z,\rho,l} \left\|\left(l \cdot m_{z,\nabla z}\right) \rho \,-\, I\right\|_{\ell^2}^2 + \mu \,\|K\,z\, - z_0\|_{\ell^2}^2 +$
$$\nu \left\|dA_{z,\nabla z}\right\|_{\ell^1} + \lambda \,\|\nabla\rho\|_{\ell^0}$$

- With $\mu = \frac{\sigma_I^2}{\sigma_z^2}, \nu = \frac{2\,\sigma_I^2}{\alpha}$ and $\lambda = \frac{2\,\sigma_I^2}{\beta}$
- Problem is non-convex and non-smooth
  - But converges in practice using a numerical solver

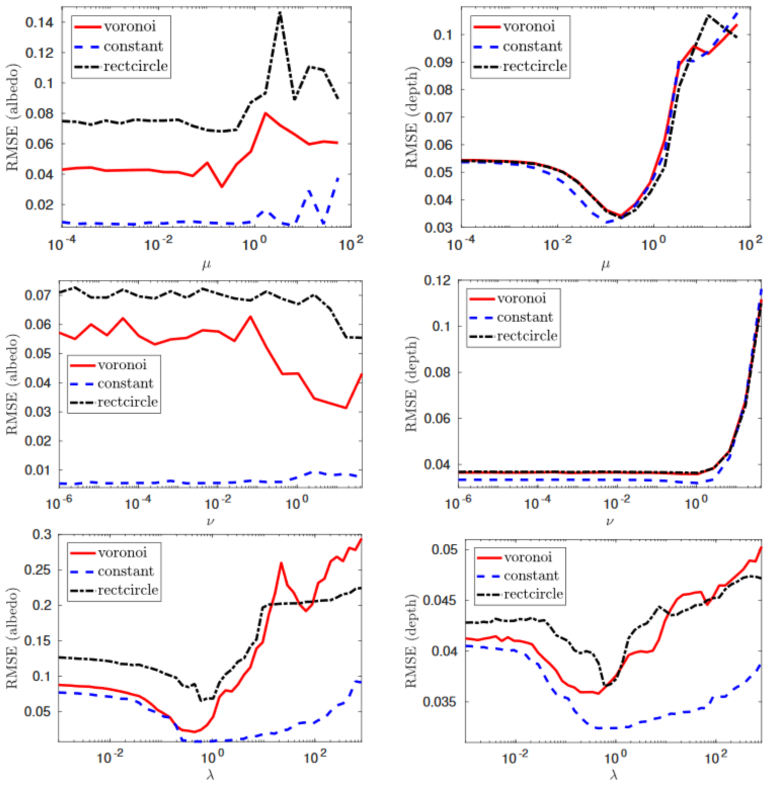
Experimental validation

- Render Test model (color + depth)
  - Single directional + ambient lighting
  - Three different reflectance maps
  - Gaussian noise on top of color/depth map



Experimental validation

- Evaluate root mean squared error between depth/albedo map and groundtruth maps
- Vary each hyperparameter μ, ν, λ to find optimum



Results

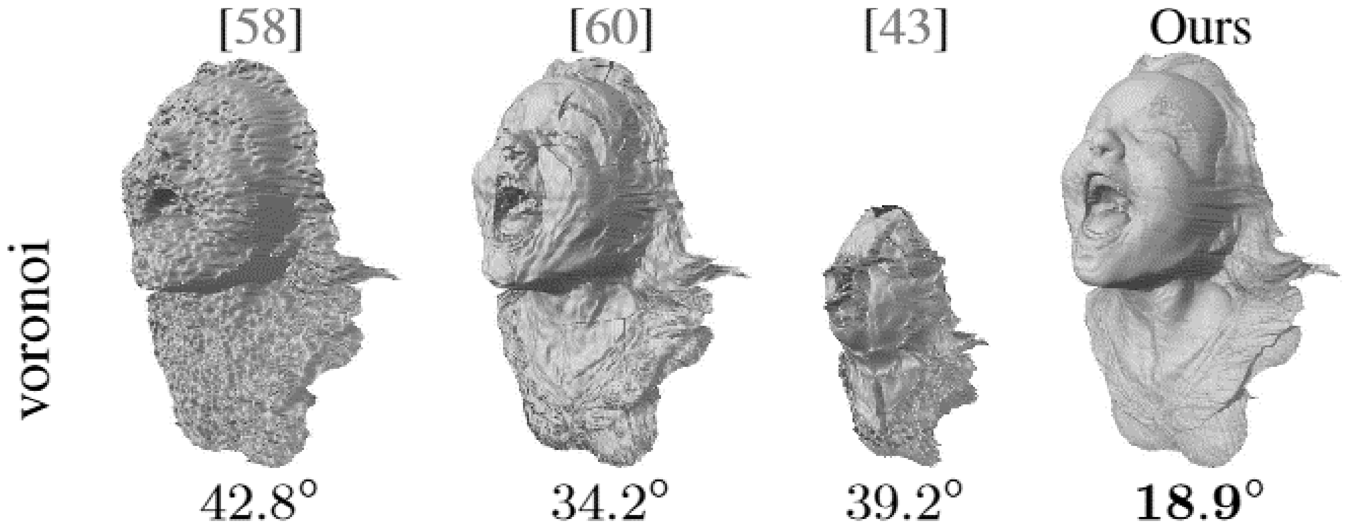
input depth input color

output depth output albedo

input depth input color

output depth output albedo

Results



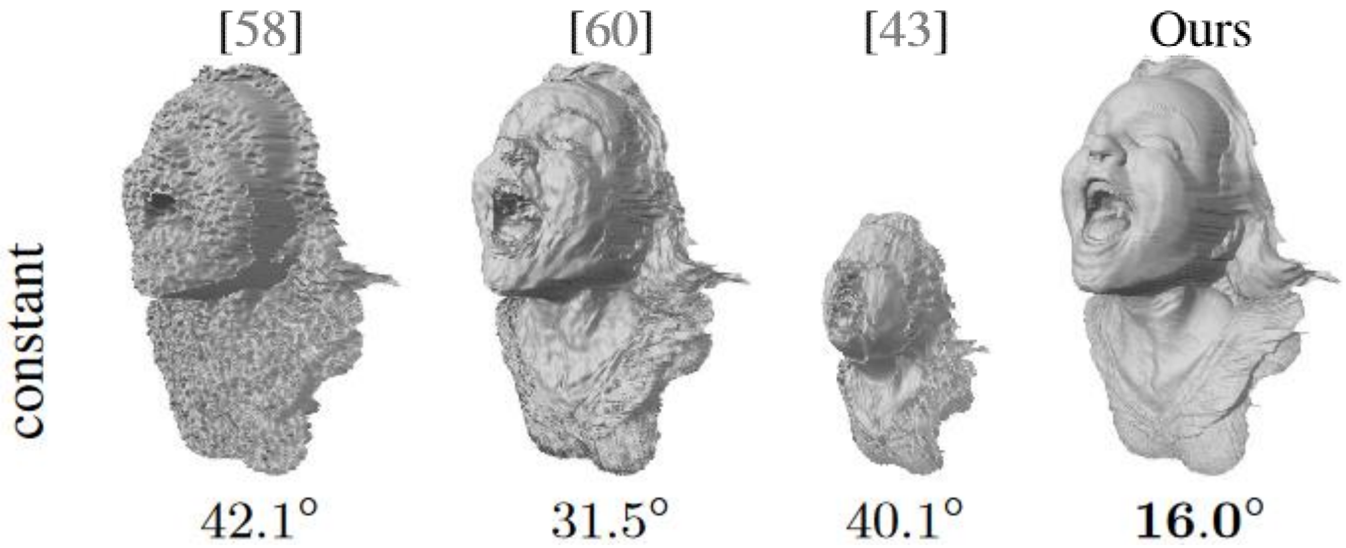
mean angular error on normals

[58] J. Xie, R. S. Feris, and M.-T. Sun.
Edge-guided single depth image super resolution

[60] Q. Yang, R. Yang, J. Davis, and D. Nist´er.
Spatial-depth super resolution for range image

[43] R. Or-El, G. Rosman, A. Wetzler, R. Kimmel, and A. Bruckstein
RGBD-Fusion: Real-Time High Precision Depth Recovery

Results



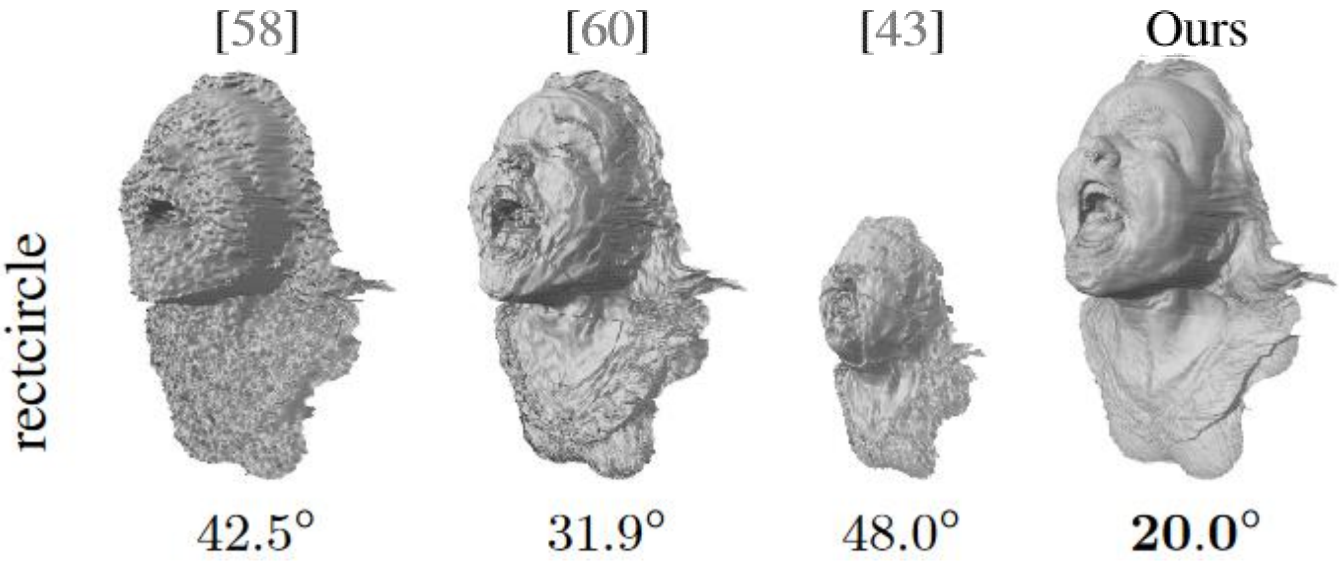
mean angular error on normals

[58] J. Xie, R. S. Feris, and M.-T. Sun.
Edge-guided single depth image super resolution

[60] Q. Yang, R. Yang, J. Davis, and D. Nist´er.
Spatial-depth super resolution for range image

[43] R. Or-El, G. Rosman, A. Wetzler, R. Kimmel, and A. Bruckstein
RGBD-Fusion: Real-Time High Precision Depth Recovery

Results



mean angular error on normals

[58] J. Xie, R. S. Feris, and M.-T. Sun.
Edge-guided single depth image super resolution

[60] Q. Yang, R. Yang, J. Davis, and D. Nist´er.
Spatial-depth super resolution for range image

[43] R. Or-El, G. Rosman, A. Wetzler, R. Kimmel, and A. Bruckstein
RGBD-Fusion: Real-Time High Precision Depth Recovery

Summary

- Combine two ill-posed problems and solve ambiguities
  - Single shot depth super-resolution is ill-posed
  - Shape from shading is ill-posed
  - Intuitive solution
    - Use high frequency color information to preserve detail in depth super-resolution
    - Use low frequency depth map as a baseline for shape from shading
- Formulate a variational problem
  - Maximize the posterior distribution of the input data
  - Can be split into likelihood and prior distribution
  - Solve numerically