High Quality Photometric Reconstruction using a Depth Camera

Sk. Mohammadul Haque, Avishek Chatterjee, Venu Madhav Govindu

Department of Electrical Engineering, Indian Institute of Science, Bengaluru, INDIA

Introduction

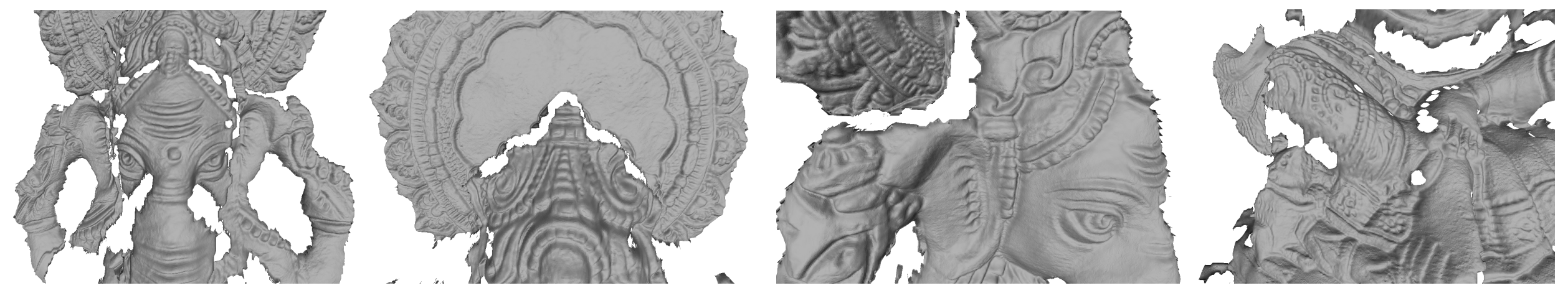
Consumer Depth Cameras (Kinect etc.):

- ☞ *Advantages:* Inexpensive, portable, easy to use, real-time depth acquisition
- ☞ *Disadvantages:* Highly noisy depth scans, fine-scale details are not captured
- ☞ *Solution:* Recover fine-scale details from photometry using same depth camera
- ☞ *Novelty:* Our is first method to use Kinect as depth *and* radiometric camera

Raw scan

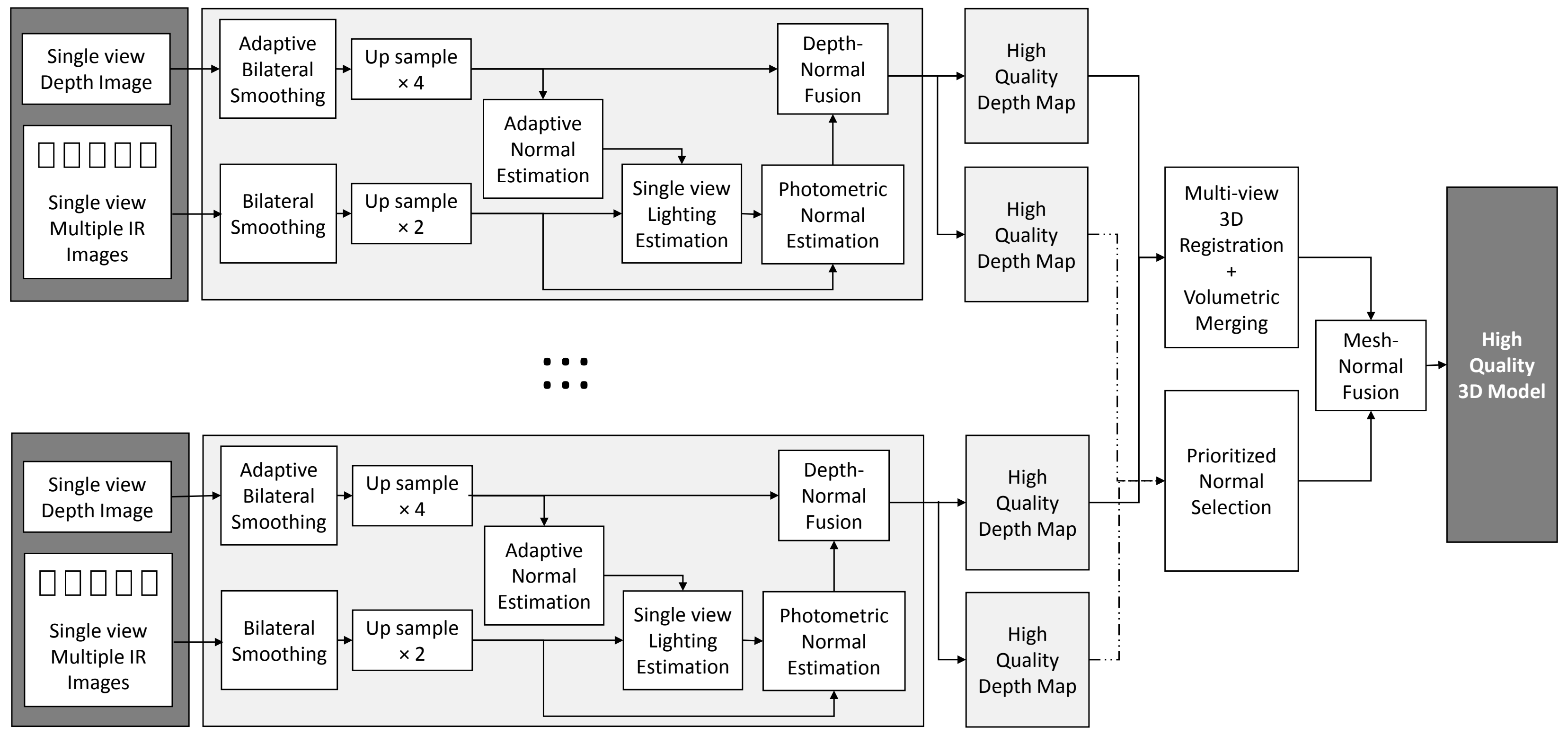
IR images

High quality scan



Novel rendered views of improved scan

Proposed Pipeline



Depth-map Guided Photometric-stereo

Linear lighting model:

I_p: intensity at pixel p; A: illumination strength; α: albedo; N_p: surface normal; V: lighting direction

For Lambertian reflectance

$$\begin{aligned} I_p &= A\alpha\, N_p \cdot V \\ \Rightarrow I_p^k &= N_p \cdot S^k \end{aligned}$$

with $S = A\alpha V$ (superscript k indexes different lighting conditions)

Linear lighting estimation:

Using estimate of surface normal N_p obtained from raw depth map

$$S^k = arg \min_{\mathbf{x}} \sum_p \rho \left(I_p^k - N_p \cdot x \right)$$

with $\rho(.)$ being a robust loss function

Photometric normal estimation:

$$N_p = arg \min_x \sum_k \rho \left(I_p^k - S^k \cdot x \right)$$

Advantages of using Kinect's Infrared Camera

- ☞ High resolution of 1280 × 960 pixels
- ☞ High dynamic range (10 bit)
- ☞ Almost linear photometric response
- ☞ Depth and IR image in same reference frame: no uncommon occlusion
- ☞ No interference due to indoor lighting and a halogen lamp acts as a good source of IR

A Novel Approach to Depth and Intensity Imaging

- ✍ Through software Kinect's IR projector turned on (for depth imaging) and off (for intensity imaging)
- ✍ Exposure of IR camera lowered through software for very close range (about 40cm) depth scanning
- ✍ Light sources placed far enough to achieve parallel lighting

Depth-Normal Fusion

$Z(x,y)$: raw depth map; $N(x,y)$: photometric normal map; $\widehat{Z}(x,y)$: estimated depth map

Relation between 3D points and depth map is given by

$$P(x,y) = \left[-\frac{x}{f} Z(x,y) \; -\frac{y}{f} Z(x,y) \; Z(x,y) \right]^T$$

with f being focal length of IR camera

Improved depth map is estimated as

$$\widehat{Z} = arg \min_{\bar{Z}} E\left(\bar{Z}\right) = arg \min_{\bar{Z}} \left[E_d\left(\bar{Z}\right) + \lambda_n E_n\left(\bar{Z}\right) + \lambda_s E_s\left(\bar{Z}\right) \right]$$

where

Depth penalty:

$$E_d\left(\bar{Z}\right) = \sum_p w_p \|\mu_p\|^2 \left(Z_p - \bar{Z}_p\right)^2$$

with $\mu_p = \left[-\frac{x}{f} \; -\frac{y}{f} \; 1 \right]^T$ and w_p is adaptive weighting, computed from the eigen values of local structure-tensor of photometric normal map

Normal penalty:

$$E_n\left(\bar{Z}\right) = \sum_p \left(N_p \cdot T_{x,p}\right)^2 + \left(N_p \cdot T_{y,p}\right)^2$$

where surface tangents T_x and T_y are given by

$$\begin{aligned} T_x = \frac{\partial P}{\partial x} &= \left[-\frac{1}{f}\left(\bar{Z} + x\frac{\partial \bar{Z}}{\partial x}\right) \; -\frac{1}{f} y \frac{\partial \bar{Z}}{\partial x} \; \frac{\partial \bar{Z}}{\partial x} \right]^T \\ T_y = \frac{\partial P}{\partial y} &= \left[-\frac{1}{f} x \frac{\partial \bar{Z}}{\partial y} \; -\frac{1}{f}\left(\bar{Z} + y\frac{\partial \bar{Z}}{\partial y}\right) \; \frac{\partial \bar{Z}}{\partial y} \right]^T \end{aligned}$$

where depth derivatives are computed by adaptively weighting the forward and backward differences depending on the similarity of the estimated normals (instead of estimated depths) of the pixels

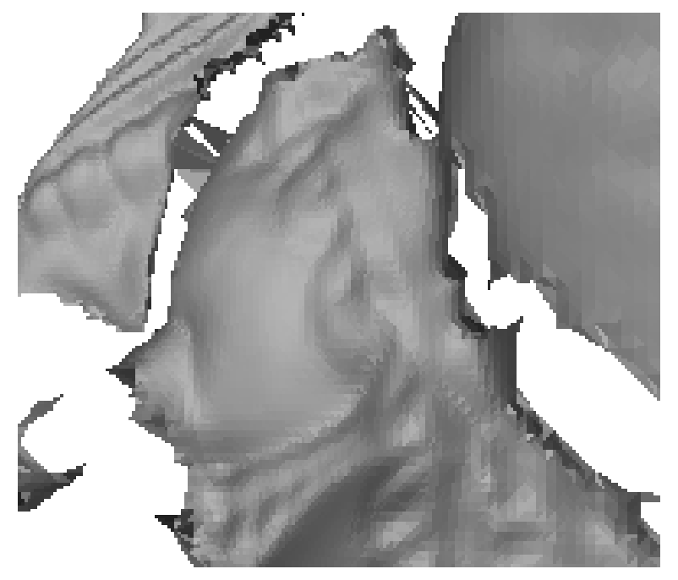
Similarity between p^{th} and q^{th} pixel is $w(p,q) = exp^{-\frac{1}{2\sigma^2}(1 - N_p^T N_q)}$

Smoothness penalty:

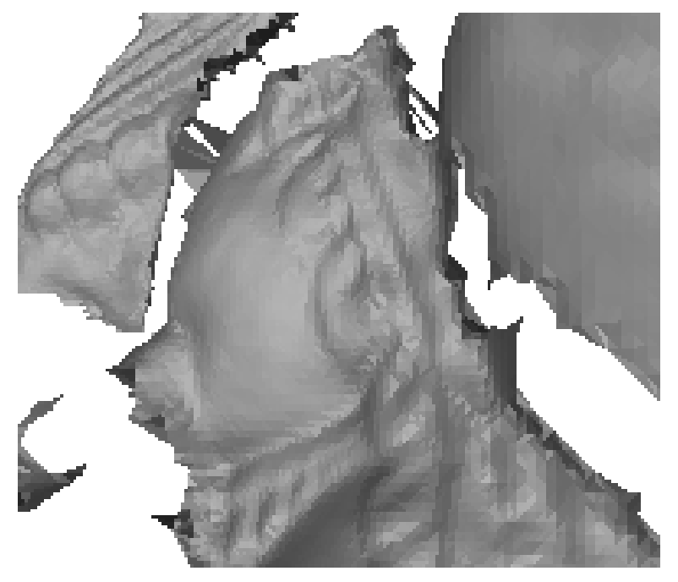
$$E_s(\bar{Z}) = \nabla^2(\bar{Z})$$

($\nabla^2(\bar{Z})$ is discrete version of the conventional Laplacian penalty)

Adaptive weighting



Depth based weighting



Our normal based weighting

Multi-view Reconstruction

Non-rigid deformation from depth-normal fusion

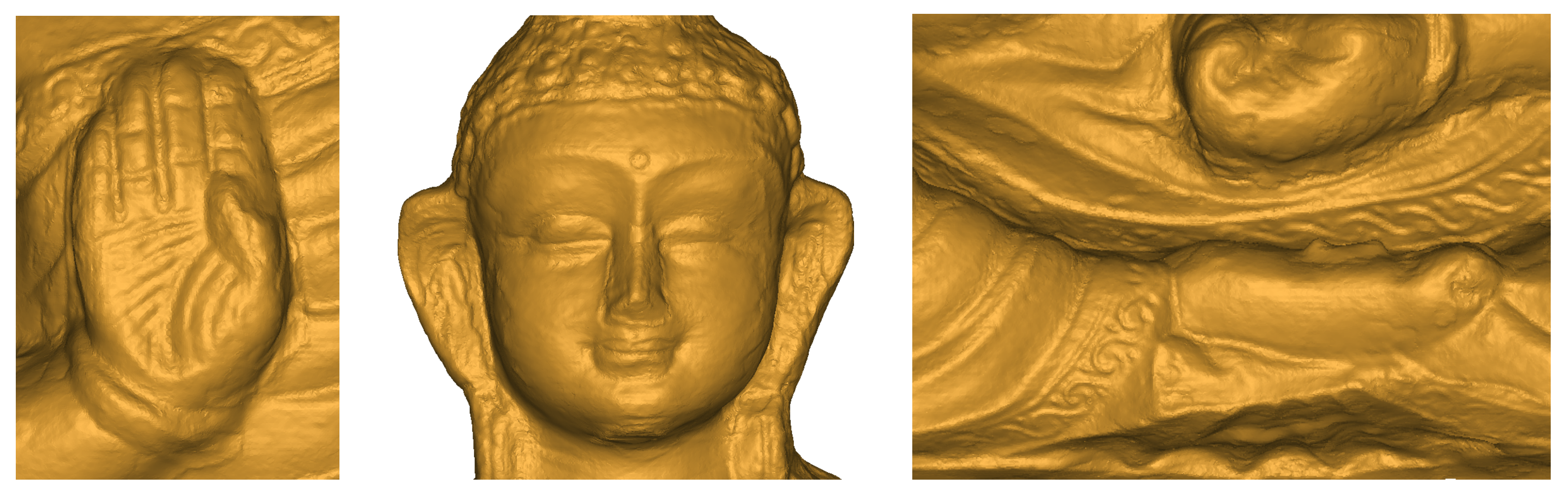
- ✍ Depth-Normal fusion causes independent non-rigid deformations of scans
- ✍ ICP-based scan registration also has small alignment error
- ✍ Both cause blurring and loss of details after merging
- ✍ Solved by recovering the normals from individual scans by ray shooting
- ✍ Averaging normal estimates from different scans also results in blurring
- ✍ Solved by selecting only one normal using a priority ordering to preserve local coherence

Result: Ganesha Model

Result: Buddha Model

Reconstruction with raw kinect depth map

Our reconstruction



Novel rendered views of our reconstruction

Conclusion

- ✍ An approach to combine a raw depth map with photometric normal map is presented
- ✍ A novel method to use the same depth camera for both depth and radiometric imaging is developed
- ✍ Automatic calibration of lighting directions is addressed
- ✍ Novel adaptive weighting and differentiation schemes are described
- ✍ A multi-view reconstruction method with coherent normal selection is developed
- ✍ Preserves fine-scale details after registration and merging

smhaque|avishek|venu@ee.iisc.ernet.in

http://www.ee.iisc.ernet.in/labs/cvl/cvpr2014