RGBD-Fusion: Real-Time High Precision Depth Recovery

Roy Or - El, Guy Rosman, Aaron Wetzler, Ron Kimmel,
Alfred M. Bruckstein

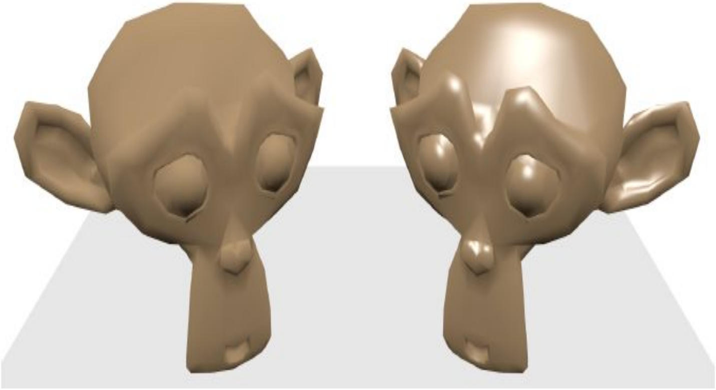
(Presented in Conference on Computer Vision and Pattern Recognition 2015)

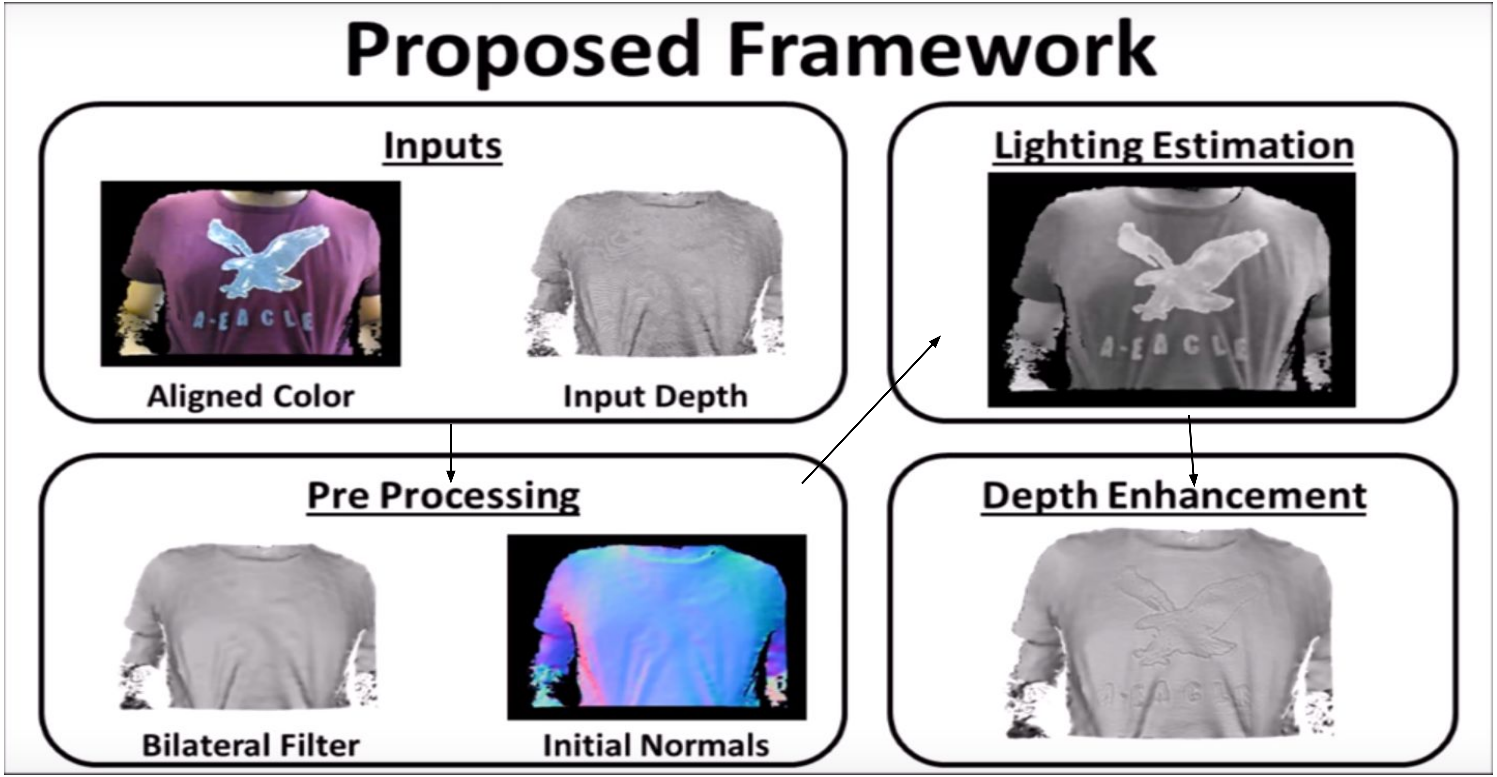
Ankith B V
Technical University of Munich
Department of Informatics
Chair for Computer Vision and AI
16th July 2019, Garching

Introduction

The main contribution of this paper:

- Presenting a novel robust method of depth enhancement under natural illumination and handles multiple albedo objects
- From a single RGB-D shot improve the underlying depth map
- Showing that improved depth maps can be acquired directly using shape from shading technique that avoids the need to first find the surface normals and then integrate them.





Proposed Framework

Assumption : The RGB-D scanner is stationary and its calibration parameters are known

Lighting Estimation

- The image is taken under natural illumination where there is no single point light source. Hence, the lambertian model cannot be used for lighting estimation.
- Defining a more lighting model :

$$L(i, j, \vec{n}) = \rho(i, j) S(\vec{n}) + \beta(i, j),$$

$\beta(i, j)$ - Accounts for local lighting variations such are interreflections and specularity.

Intensity I Albedo ρ Light $\mathbf{l}$ Surface normal $\mathbf{n}$

Shading Computation

- “Lambertian Reflectance and Linear Subspaces” by Basri and Jacobs states irradiance of diffuse objects in natural illumination scenes can be well described by the low order spherical harmonics components.

$$S(\vec{n}) = \vec{m}^T \tilde{n}$$

m is a vector of the four first order spherical harmonics coefficients

- Set reflectance term to 1 and beta to 0 in the lighting model to estimate shading.

$$\vec{m} = \underset{\vec{m}}{\operatorname{argmin}} \|\vec{m}^T \tilde{n} - I\|_2^2.$$

Multiple Albedo Recovery

- Now freeze the shading component and optimize reflectance to account for scene albedos and account for shadows
- In order to avoid overfitting, a penalty term is added.

$$\min_{\rho} \|\rho S(\vec{n}) - I\|_2^2 + \lambda_\rho \| \sum_{k \in \mathcal{N}} \omega_k^c \omega_k^d (\rho - \rho_k)\|_2^2$$

$$\omega_k^c = \begin{cases} 0, & \|I_k - I\|_2^2 > \tau \\ \exp\left(-\frac{\|I_k - I(i,j)\|_2^2}{2\sigma_c^2}\right), & \text{otherwise,} \end{cases} \qquad \omega_k^d = \exp\left(-\frac{\|z_k - z(i,j)\|_2^2}{2\sigma_d^2}\right)$$

Lighting Variation Recovery

- Similar function to the one used in albedo recovery is used to recover lighting variation.
- But the energy of the lighting variation term is limited to be consistent with the shading model.

$$\min_{\beta} \|\beta-(I-\rho S(\vec{n}))\|_2^2+\lambda_\beta^1\|\sum_{k\in\mathcal{N}}\omega_k^c\omega_k^d(\beta-\beta_k)\|_2^2+\lambda_\beta^2\|\beta\|_2^2.$$

Refining the Surface

- The surface normals are represented as a function of depth.

$$\vec{n} = \frac{(z_x, z_y, -1)}{\sqrt{1 + \|\nabla z\|^2}} \qquad z_x = \frac{dz}{dx}, \quad z_y = \frac{dz}{dy}$$

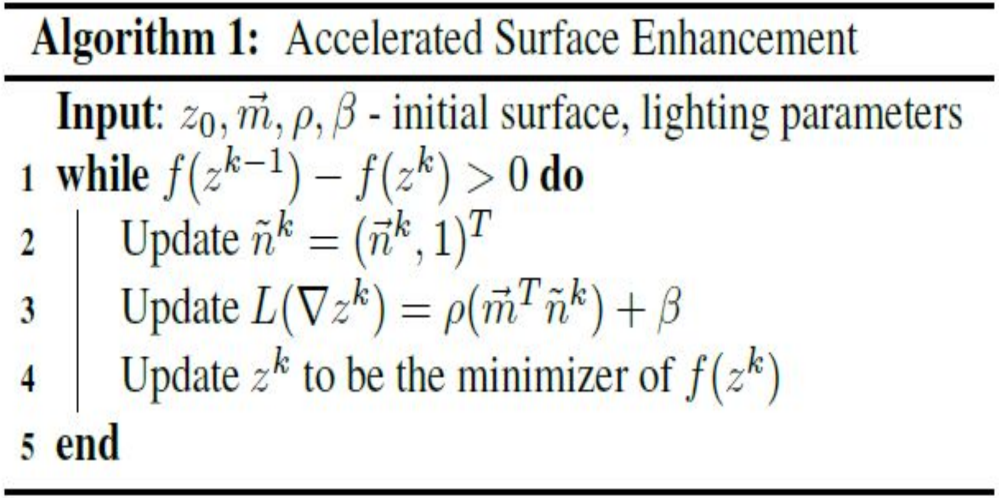
- Now, by fixing the lighting model parameters and allowing surface gradient to vary, subtle geometry recovered by the following l2 minimization with regularization.

$$f(z) = \|L(\nabla z) - I\|_2^2 + \lambda_z^1 \|z - z_0\|_2^2 + \lambda_z^2 \|\Delta z\|_2^2$$

Algorithm

- Freeze nonlinear terms inside the shading model.Solve the linear system in each iteration and then update the non-linear terms.

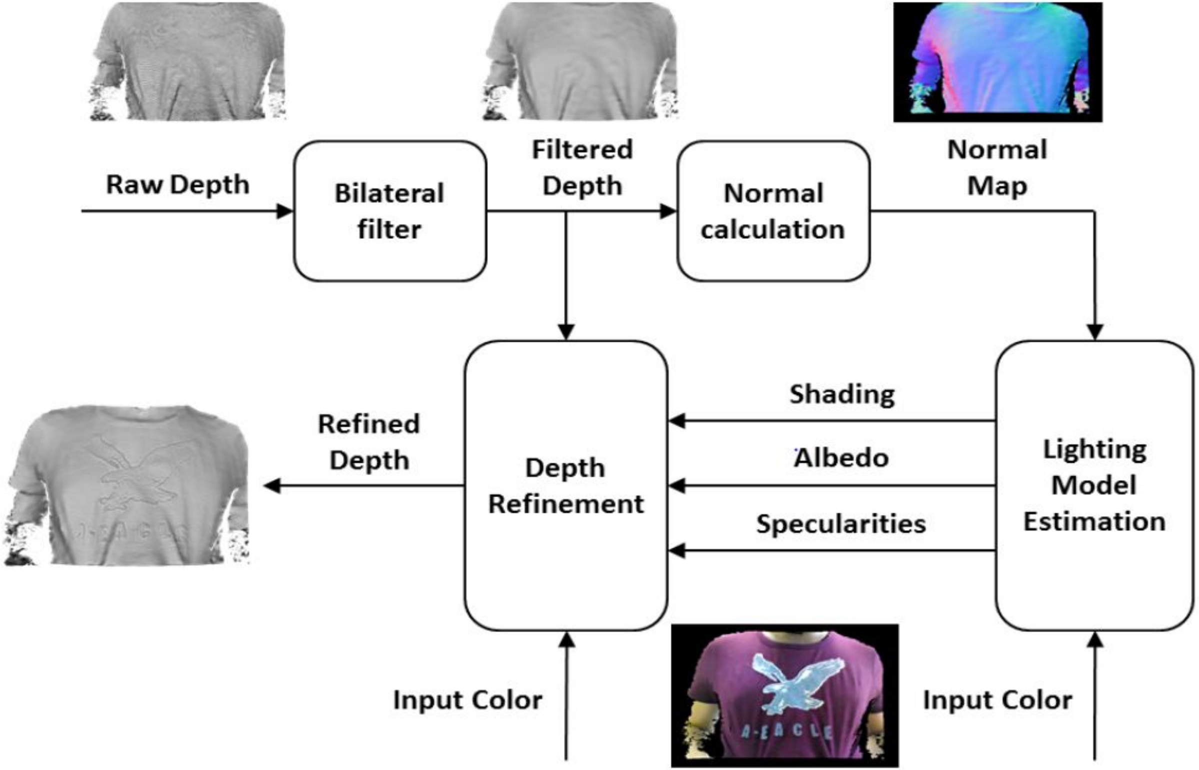
$$f(z^k) = \|\rho(\vec{m}^T\tilde{n}^k) - (I - \beta)\|_2^2 \\ + \lambda_z^1\|z^k - z_0\|_2^2 + \lambda_z^2\|\Delta z^k\|_2^2$$



Algorithm 1: Accelerated Surface Enhancement

```
Input: $z_0, \vec{m}, \rho, \beta$ - initial surface, lighting parameters
while $f(z^{k-1}) - f(z^k) > 0$ do
    Update $\tilde{n}^k = (\vec{n}^k, 1)^T$
    Update $L(\nabla z^k) = \rho(\vec{m}^T\tilde{n}^k) + \beta$
    Update $z^k$ to be the minimizer of $f(z^k)$
end
```

Summary of the proposed Algorithm



Results

- Proposed algorithm’s performance on synthetic data (Obtained from Stanford 3D repository and Smithsonian 3D archive) with simulated complex lighting environment using Blender as well as real data.
- The algorithm parameters were set to :

$$\lambda_\rho = 0.1,\ \lambda_\beta^1 = 1,\ \lambda_\beta^2 = 1,\ \tau = 0.05,\ \sigma_c = \sqrt{0.05},\ \sigma_d = \sqrt{50},\ \lambda_z^1 = 0.004,\ \lambda_z^2 = 0.0075$$

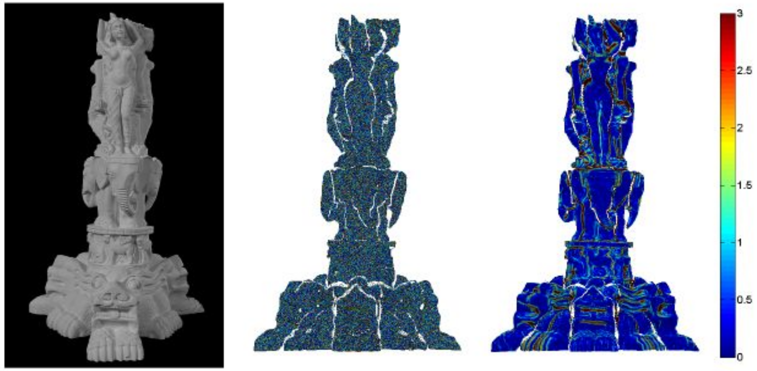


Figure 1: Thai Statue error analysis. From left to right: Input color image. Error image of the raw depth map. Error image of the final result. Note how the algorithm reduces the initial surface errors.

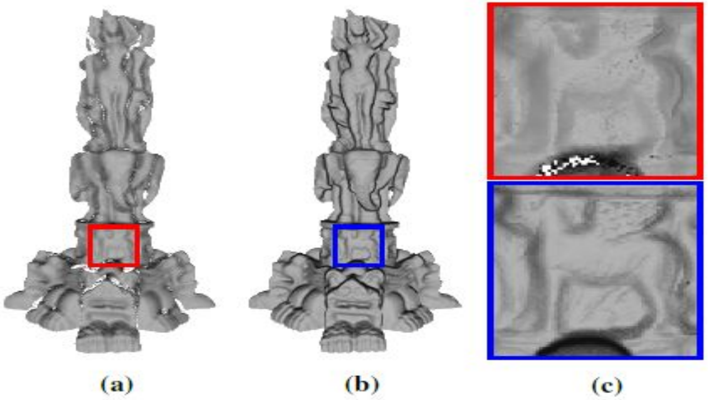
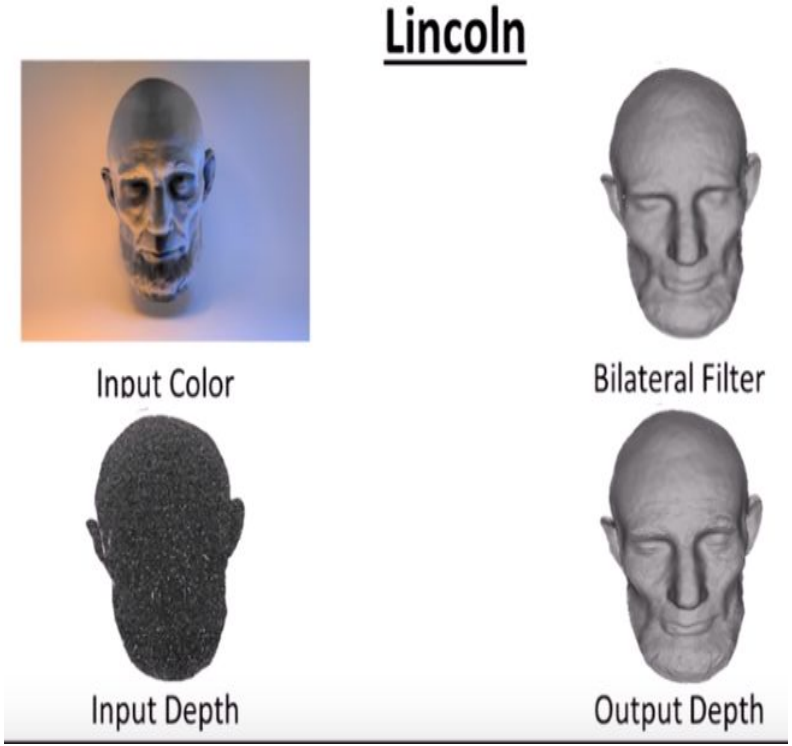
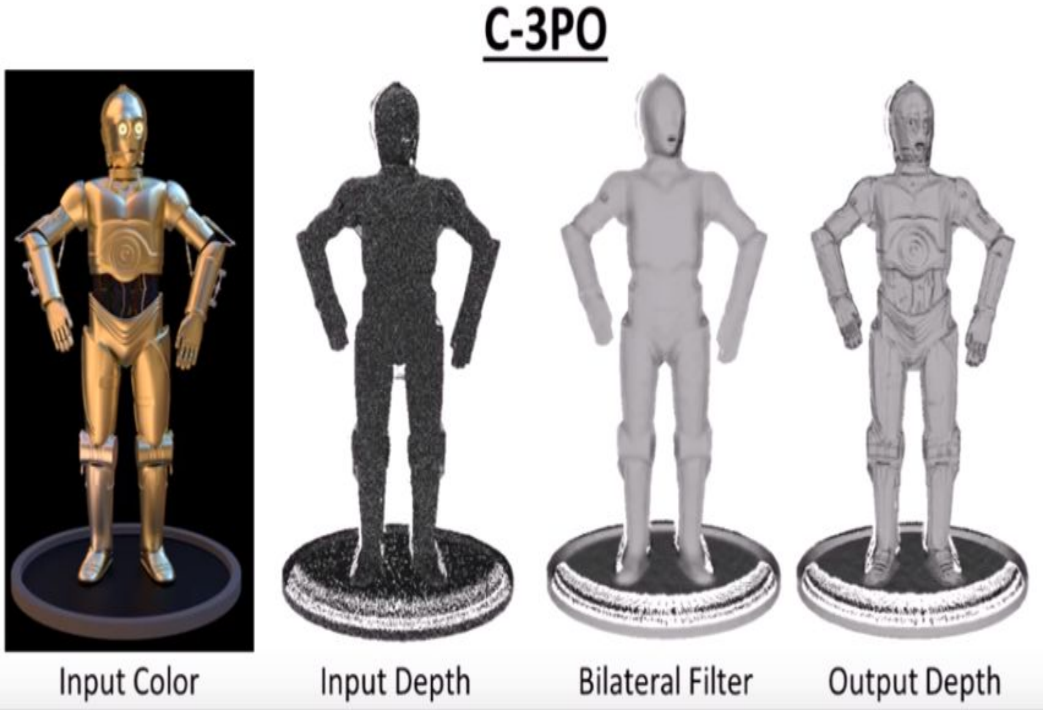

(a) (b) (c)

Figure 5: Robustness to normal outliers: Left to right: HLK reconstruction with the entire depth map normals (a). Our method reconstruction with the entire depth map normals (b). Magnification of the results is presented in (c). The proposed method yield accurate reconstruction despite the distorted shading.



Lambertian object in a complex lighting environment with multiple casted shadows



Non-Lambertian object with a point light source.

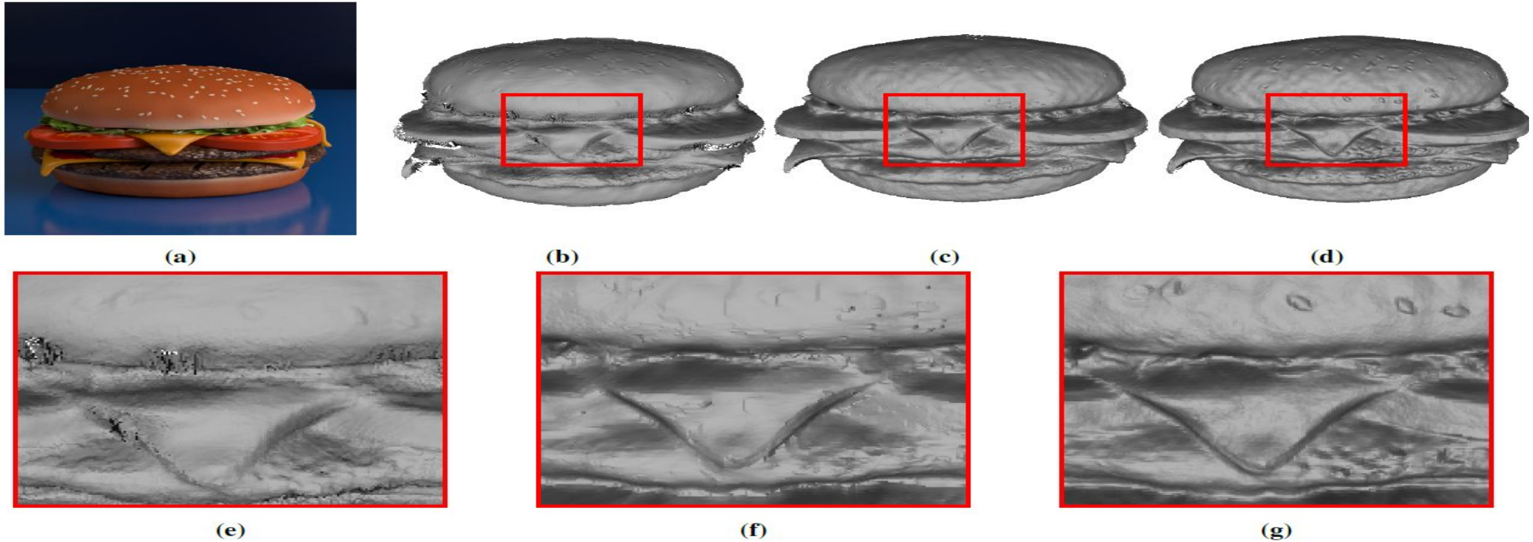
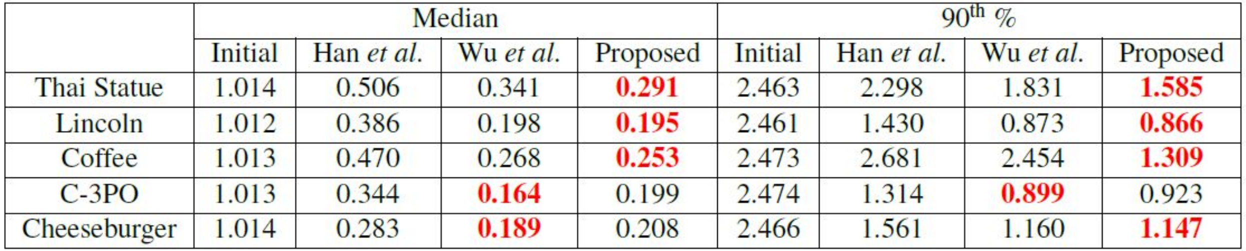


Figure 6: Handling a multiple albedo object. (a) Color Image. (b) HLK Reconstruction. (c) WZNSIT Reconstruction. (d) Our Reconstruction. (e) - (g) Magnifications of HLK, WZNSIT and Our Method respectively. Note how the proposed framework sharply distinguish albedo changes.



	Median				90^{th} %			
	Initial	Han *et al.*	Wu *et al.*	Proposed	Initial	Han *et al.*	Wu *et al.*	Proposed
Thai Statue	1.014	0.506	0.341	**0.291**	2.463	2.298	1.831	**1.585**
Lincoln	1.012	0.386	0.198	**0.195**	2.461	1.430	0.873	**0.866**
Coffee	1.013	0.470	0.268	**0.253**	2.473	2.681	2.454	**1.309**
C-3PO	1.013	0.344	**0.164**	0.199	2.474	1.314	**0.899**	0.923
Cheeseburger	1.014	0.283	**0.189**	0.208	2.466	1.561	1.160	**1.147**

Table 1: Quantitative comparison of depth accuracy on simulated models.

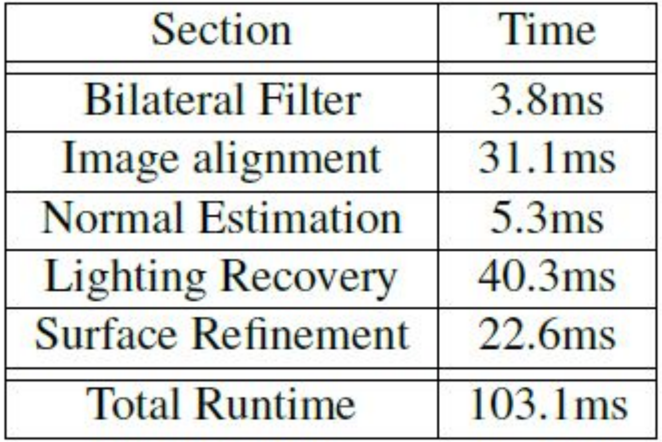
(a) Shirt

(b) Baseball Cap

Figure 4: Results of shape enhancement of real world multiple albedo objects. Left to right: Color Image, Raw Depth, Bilateral Filtering and the Proposed Method. Note how surface wrinkles and small surface protrusions are now visible.

Real Time Depth Enhancement

An unoptimised implementation of the algorithm is tested on Intel i7 3.4GHz processor with 16GB RAM and Nvidia Geforce GTX TITAN GPU at 10fps for a 640x480 depth profiles. Time-breakdown is as follows



Section	Time
Bilateral Filter	3.8ms
Image alignment	31.1ms
Normal Estimation	5.3ms
Lighting Recovery	40.3ms
Surface Refinement	22.6ms
Total Runtime	103.1ms

Thank you