RGBD-Fusion: Real-Time High Precision Depth Recovery

Roy Or - El[1], Guy Rosman[2], Aaron Wetzler[1], Ron Kimmel[1], Alfred M. Bruckstein[1],
[1]Technion, Israel Institute of Technology. [2]Computer Science and Artificial Intelligence Lab, MIT.

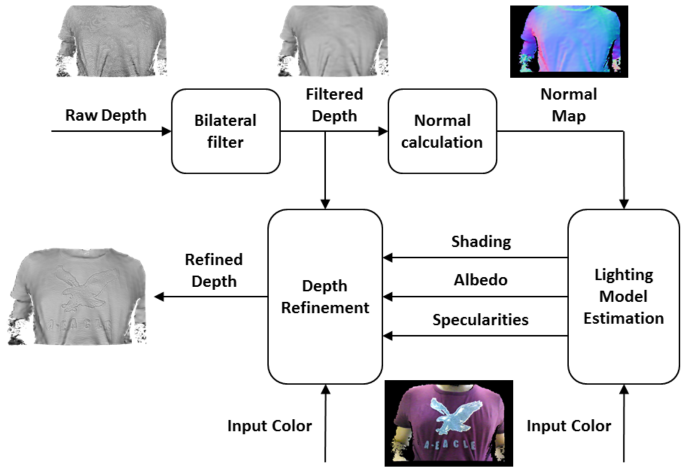


Figure 1: Algorithm's flowchart

Background: The availability of affordable depth scanners has sparked a revolution in many applications of computer vision, such as robotics, human motion capture, scene modeling and analysis. The increased availability of such scanners naturally raises the question of whether it is possible to exploit the associated intensity image to improve their lack of accuracy. Our goal is to fuse the captured data from the RGB-D scanner in order to enhance the accuracy of the acquired depth maps.

Contribution: We introduce a novel real-time method to directly enhance surface recovery that achieves state of the art accuracy. We apply a lighting model that uses normals estimated from the depth profile, and eliminates the need for calibration of the scene lighting. The lighting model accounts for light sources, multiple albedos, and local lighting effects such as specularities, shadows and interreflections. The high quality surface is reconstructed directly in a shape from shading fashion using a surface dependent cost functional without first finding and integrating its normals. In order to achieve fast convergence, we relinearize the variational problem. The main contributions of this paper are:

1. Presenting a novel robust depth enhancement method that operates under natural illumination and handles multiple albedo objects.
2. Showing that depth accuracy can be enhanced in real-time by efficiently fusing the RGB-D inputs.
3. 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.

Approach: Our algorithm accepts a depth map and a corresponding intensity image. Its main steps are:

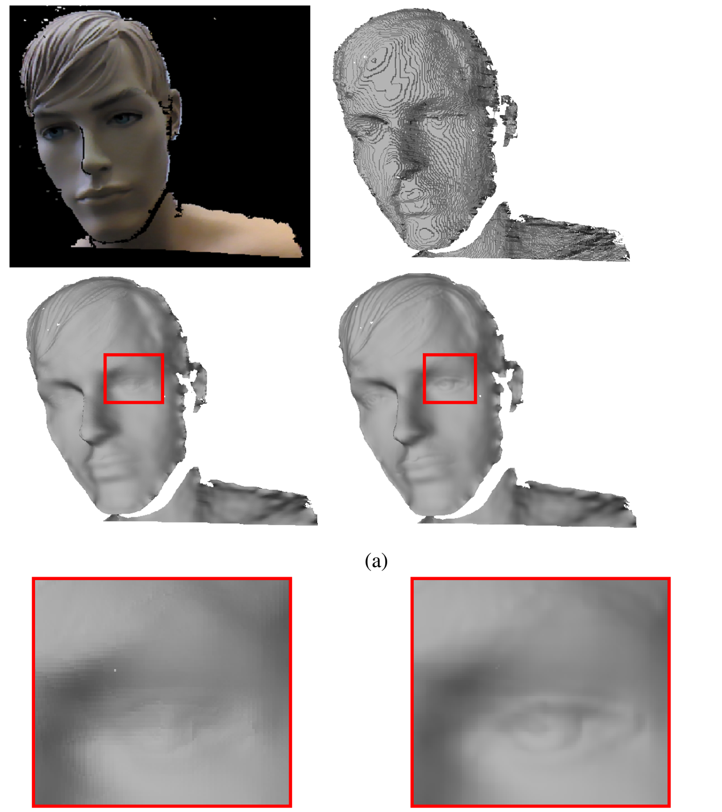
- **Pre-processing**. We obtain a rough version of the input surface by applying a bilateral filter on the input depth map and calculate initial surface normals corresponding to the smoothed surface.
- **Lighting model estimation**. We decompose the image intensity according to the model suggested in [1],

$$L(i,j,\vec{n}) = \rho(i,j)S(\vec{n}) + \beta(i,j). \quad (1)$$

  We recover the shading S from the initial normals and the intensity image using an affine illumination model. The non-uniform albedo map ρ is computed assuming piecewise smooth behaviour. To account for specularities, we estimate β which represents additional illumination artifacts.
- **Depth refinement**. To enhance the surface we minimize

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

This is an extended abstract. The full paper is available at the Computer Vision Foundation webpage.



(a)

(b) (c)

Figure 2: **Mannequin.** (a) Upper Left: Color Image. Upper Right: Raw Depth. Bottom Left: Result of Wu et al. [3] . Bottom Right: Our Result. (b), (c) Magnifications of the mannequin eye. The mannequin's hair and facial features can be easily recognized in our reconstruction.

To achieve real-time performance, we use an optimization scheme similar to IRLS where we solve a linear system and then update the non-linear terms at the end of each iteration.

Results: In the paper we show qualitative and quantitative visual and statistical evidence for depth image improvement. We compare the suggested method to state of the art approaches [2, 3] and demonstrate real-time performance. One such example can be seen in Fig. 2. A video describing the proposed framework can be found at `https://youtu.be/5nLnOTyrd-s`.

Acknowledgments: This research was supported by European Community's FP7-ERC program grant agreement no. 267414 and by the Broadcom Foundation. G.R is partially funded by VITALITE Army Research Office Multidisciplinary Research Initiative program, award W911NF-11-1-0391.

[1] Roger Grosse, Micah K. Johnson, Edward H. Adelson, and William T. Freeman. Ground-truth dataset and baseline evaluations for intrinsic image algorithms. In *International Conference on Computer Vision*, pages 2335–2342, 2009.

[2] Yudeog Han, Joon Young Lee, and In So Kweon. High quality shape from a single RGB-D image under uncalibrated natural illumination. In *IEEE International Conference on Computer Vision*, pages 1617–1624, 2013.

[3] Chenglei Wu, Michael Zollhofer, Matthias Niessner, Marc Stamminger, Shahram Izadi, and Christian Theobalt. Real-time shading-based refinement for consumer depth cameras. In *ACM Transactions on Graphics (Proceedings of SIGGRAPH Asia 2014)*, volume 33, December 2014.