Intrinsic Image Decomposition using Focal Stacks

Supplementary Material : ICVGIP 2016

Saurabh Saini Parikshit Sakurikar P J Narayanan
Centre for Visual Information Technology, Kohli Center on Intelligent Systems, IIIT Hyderabad

1. SYNTHESIZING FOCAL STACKS

In order to synthesize a focal stack over a scene where the depth information for each pixel is available, we need to first estimate the correct focus distances for each of the slices in the focal stack. We do this by clustering the pixels in the depth space. We use the k-means clustering algorithm over the depth map to label pixels into depth clusters. The mean of each cluster refers to the mean depth value of the cluster and is thus used as the object-side focus distance for that focal slice. Using this known focus distance, we can compute the amount of defocus for all the other pixels in the scene based on their depth values.

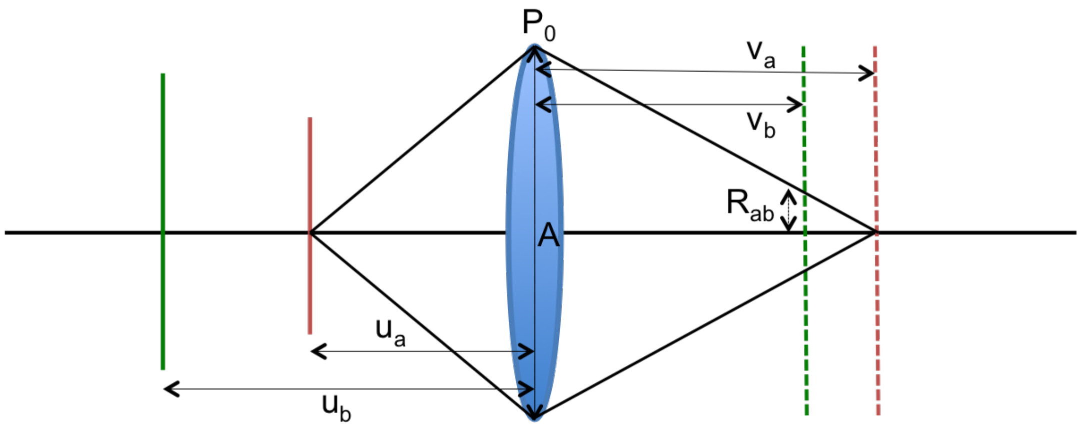

Figure 1: Defocus radius for scene objects at u_a when sensor is placed at u_b

As shown in Fig. 1, the defocus radius for pixels focused at sensor location v_a when the sensor is moved to v_b is given by

$$R_{ab} = \frac{A}{2}\frac{|v_a - v_b|}{v_a} \tag{1}$$

where A is a the aperture of the camera, and F is the focal length. The values for A and F are assumed independent of the scene. Using the thin-lens equation governing u, v and F, the above is equivalent to:

$$R_{ab} = \frac{A}{2}\left|1 - \frac{u_b(u_a - F)}{u_a(u_b - F)}\right| \tag{2}$$

Since the depth data is known, the u values are known for all the pixels in the scene. Thus, we iterate over the previously defined cluster centers and fix the focus distance to be the mean depth value of the cluster center. Thus, for each focal slice, u_b is derived from the cluster center, as u_b is the distance where the sensor is currently placed Figure 1. For all the other pixels having a depth u_a, we find the defocus radius using Eqn. Equation 2 and we apply a spatially invariant blurring operation as described in the supplementary material of Barron et al. [1].

2. CHROMATICITY CONSTRAINT

Chromaticty constraint is gven by :

$$f^R(R_B) = \sum_{\mathcal{P}}\sum_{\mathcal{N}_p} w^R_{pq}(R_B(p) - R_B(q))^2$$

Here w^R_{pq} is the angular distance based chromaticity difference.

$$w^R_{pq} = \exp\left(-\frac{1 - \langle C(p), C(q)\rangle^2}{\sigma_c^2}\right)\left(1 + \exp\left(-\frac{B(p)^2 + B(q)^2}{\sigma_i^2}\right)\right)$$

Here $C(p)$ represents normalized RGB value for 3 channels, $B(p)$ is the actual RGB value and σ's are free parameters.

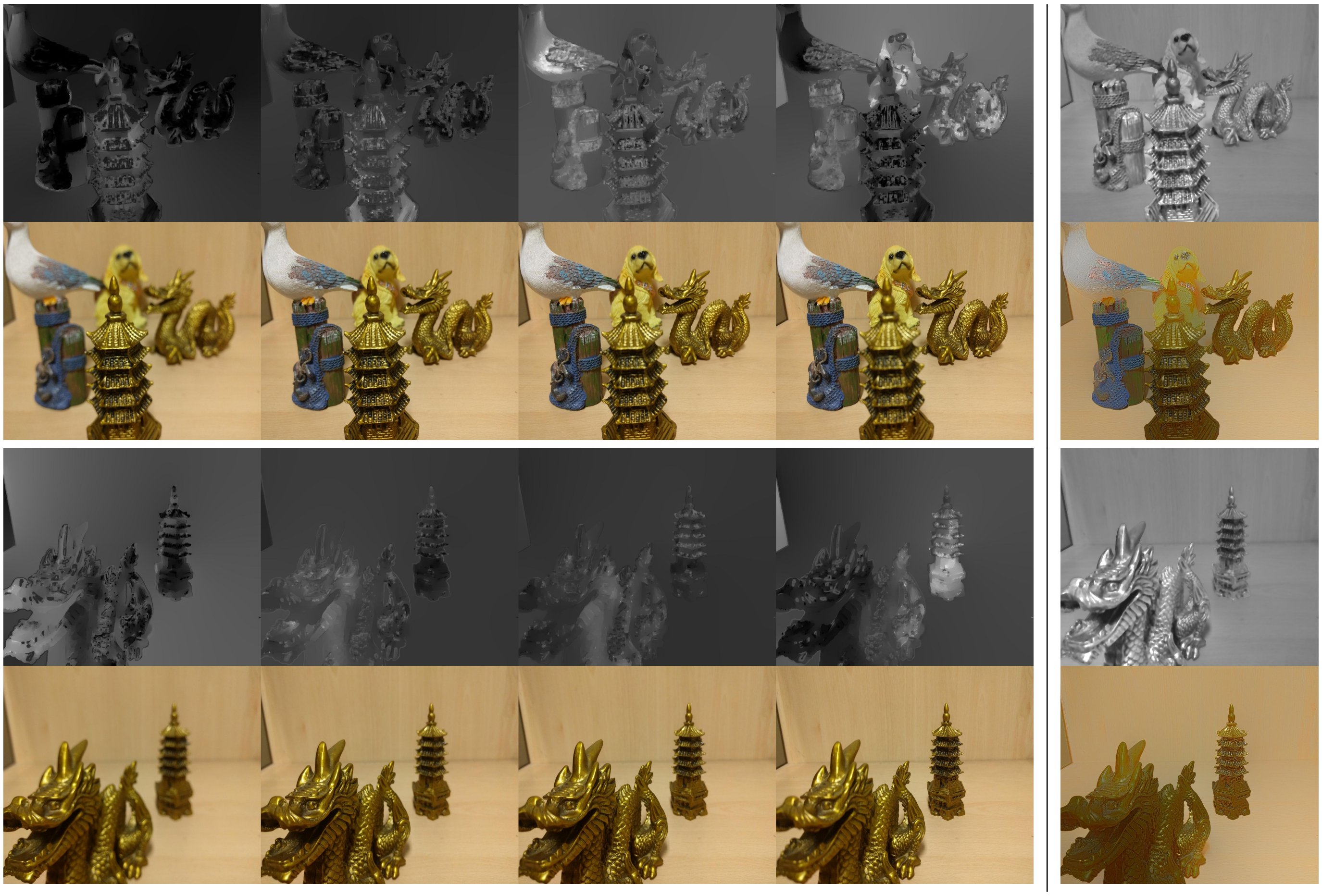

Figure 2: Top row first column : Focal slices. Bottom row : Probability Maps. Second column : Shading on top and Reflectance in bottom

3. SAMPLE FOCAL STACKS AND PROBABILITY MAPS

Here we show some sample probability maps from two focal stacks we captured using Nexus5x Figure 2.

4. NYU RESULTS

Some NYU dataset results. Each scene has two rows. First row : input image, reflectance from our proposed method fStackIID, textSepIID Jeon et al. [3] , depthCuesIID Chen and Koltun [2]. Second row : depth image, shading from fStackIID, textSepIID, depthCuesIID. Results shown in Figure 3, 4, 5, 6 and 7.

5. MPI RESULTS

Figure 8 and Figure 9 shows some extra results on MPI dataset. Each scene shows respectively original image, reflectance and shading.

References

[1] Jonathan T Barron, Andrew Adams, YiChang Shih, and Carlos Hernández. Fast bilateral-space stereo for synthetic defocus. In *CVPR*, pages 4466–4474, 2015.

[2] Q. Chen and V. Koltun. A simple model for intrinsic image decomposition with depth cues. In *2013 ICCV*, pages 241–248, 2013.

[3] Junho Jeon, Sunghyun Cho, Xin Tong, and Seungyong Lee. Intrinsic image decomposition using structure-texture separation and surface normals. In *ECCV 2014*, 2014.

Figure 3: NYU dataset results

Figure 4: NYU dataset results

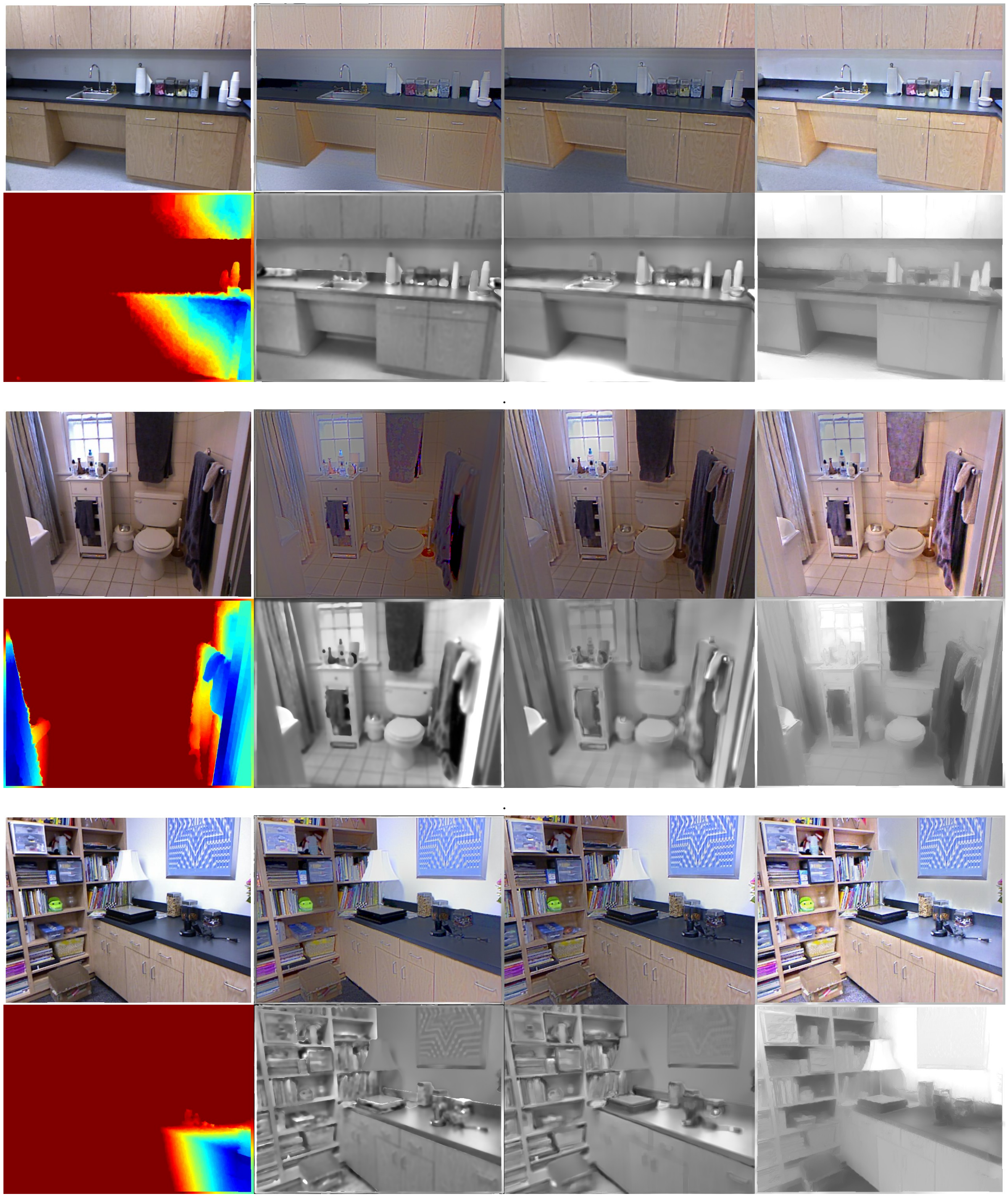

Figure 5: NYU dataset results

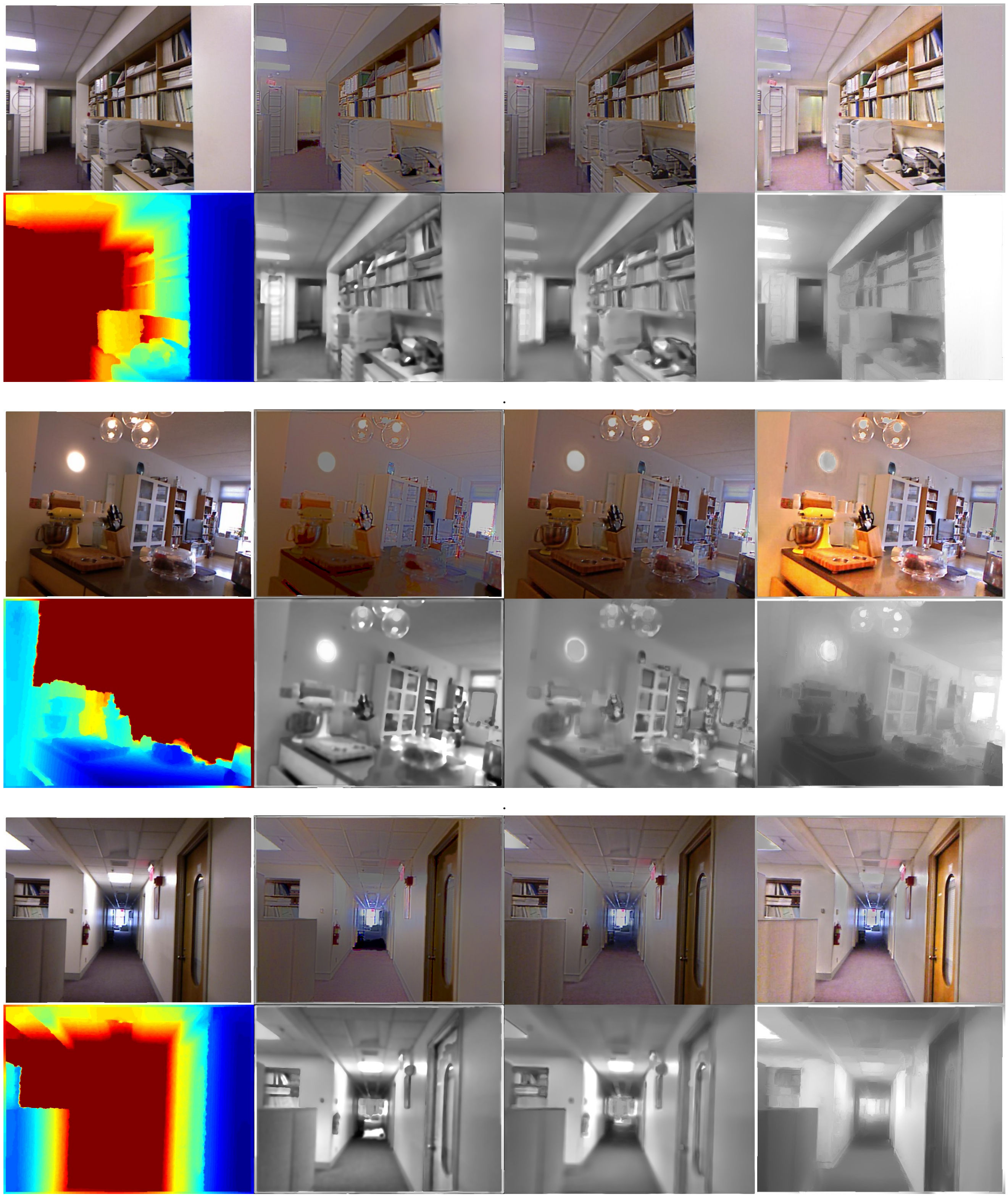

Figure 6: NYU dataset results

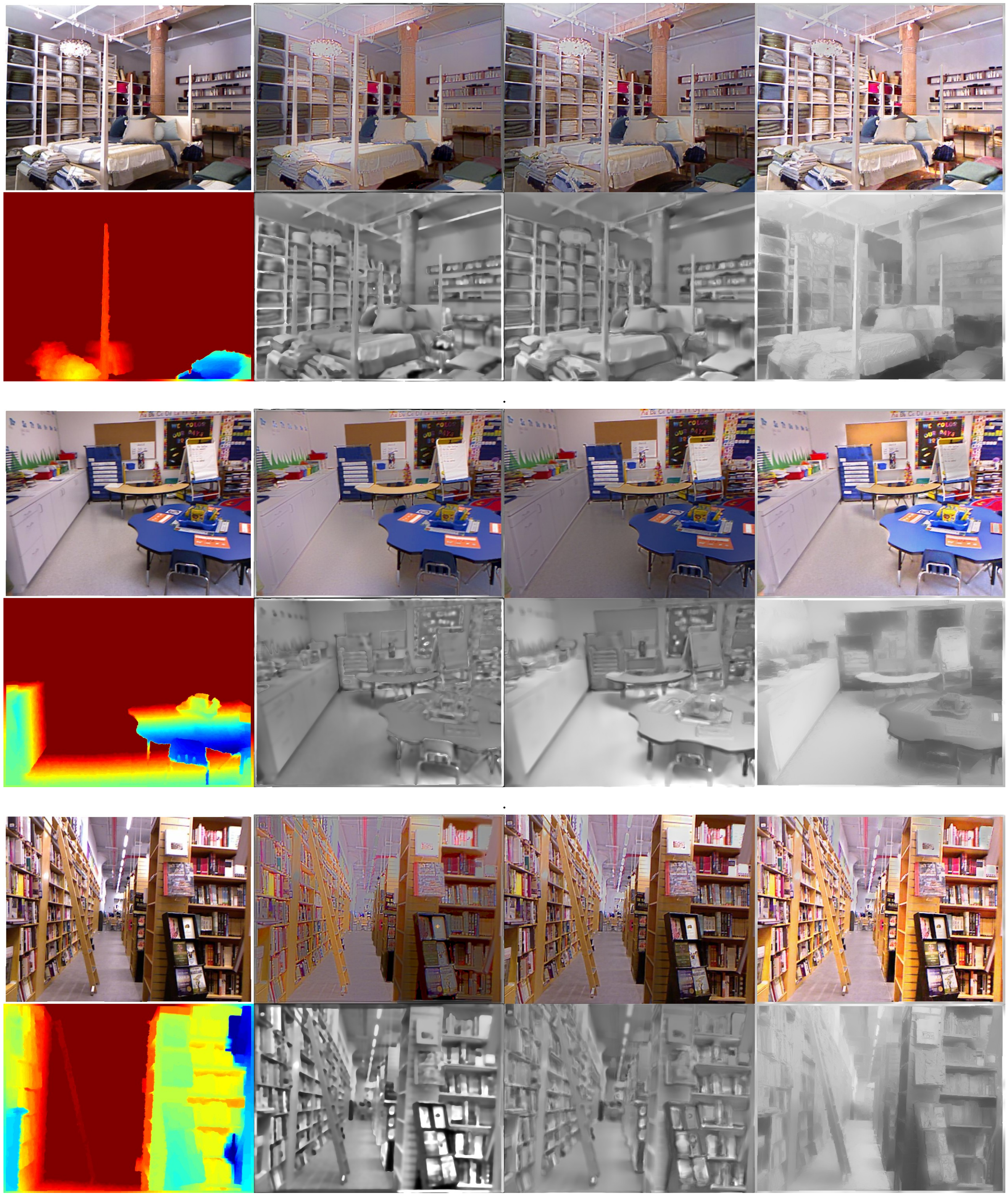

Figure 7: NYU dataset results

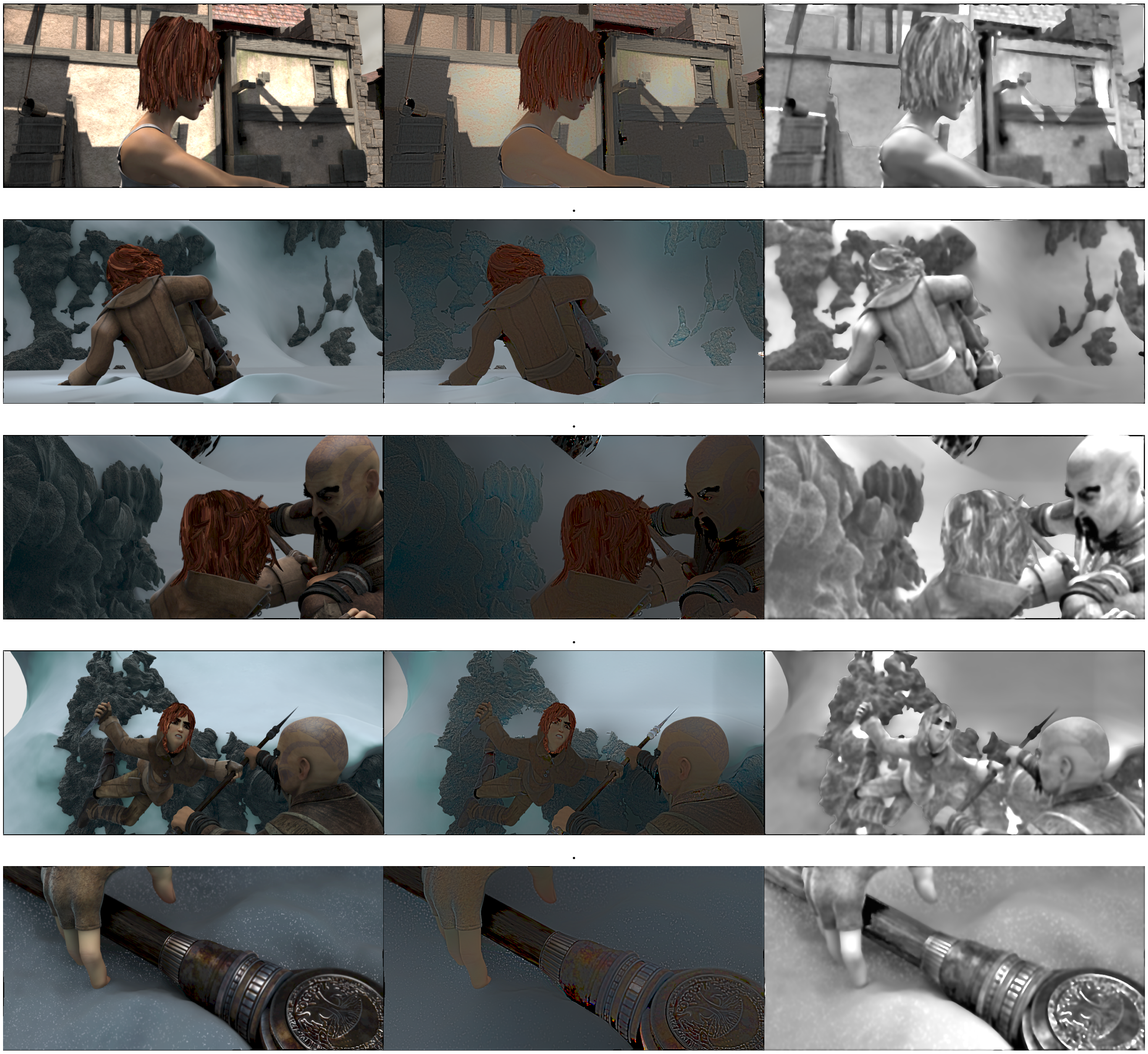

Figure 8: MPI dataset results

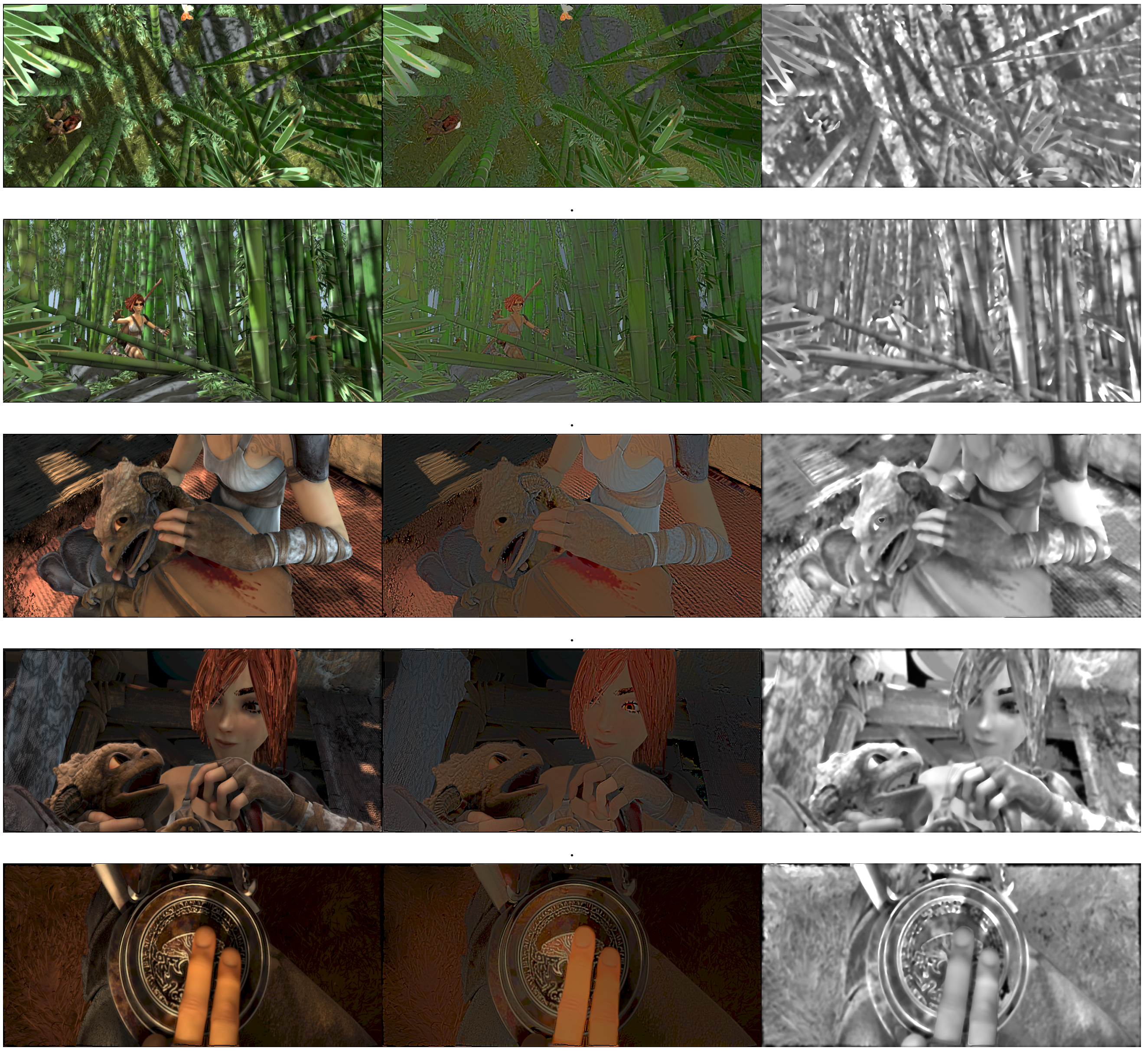

Figure 9: MPI dataset results