Supplementary Material
Joint Texture and Geometry Optimization for RGB-D Reconstruction

Yanping Fu[1] Qingan Yan[2] Jie Liao[1] Chunxia Xiao[1,3,4]
[1] School of Computer Science, Wuhan University, China
[2] Silicon Valley Research Center, JD.com, United States
[3] National Engineering Research Center for Multimedia Software, Wuhan University, China
[4] Institute of Artificial Intelligence, Wuhan University, China
{ypfu,liaojie,cxxiao}@whu.edu.cn, qingan.yan@jd.com

1. Evaluation on Public Datasets

Here we show more comparison results with related methods. First, we compare the texture and geometry optimization performance of our method and Intrinsic3D [4] on two other datasets provided by Intrinsic3D [4], as shown in Figure 1. Notable that the proposed method can restore more geometric detail than Intrinsic3D on the carving surface. In addition, we also compare our method with the methods of Zhou *et al.* [6], Fu *et al.* [2], Intrinsic3D [4] and Bi *et al.* [1] in texture optimization on the dataset provided by [6], as shown in Figure 2. The texture mapping results generated by our method are comparable to those methods focusing on texture optimization.

2. Application in Virtual Reality

Using the proposed method, we can reconstruct the objects in daily life and render the 3D reconstructed models via videos or images captured by mobile phone. We can also reconstruct the human body and insert them into another videos or images, and can share them with our friends on the social network. To this end, we first scan the target object and reconstruct the 3D model. Then we optimize the geometry and texture of the reconstruction model using the proposed method. Subsequently, we adopt the method based on [3] to render the reconstruction model in the video or image. Figure 3 demonstrates a sequence of frames where we rendered a virtual horse model optimized by our method in the realistic scene. After the texture and geometry optimization by the proposed method, the reconstruction model is realistic in the video.

References

[1] Sai Bi, Nima Khademi Kalantari, and Ravi Ramamoorthi. Patch-based optimization for image-based texture mapping. *ACM Transactions on Graphics*, 36(4):106–1, 2017.

[2] Yanping Fu, Qingan Yan, Long Yang, Jie Liao, and Chunxia Xiao. Texture mapping for 3D reconstruction with RGB-D sensor. In *CVPR*, 2018.

[3] Kevin Karsch, Kalyan Sunkavalli, Sunil Hadap, Nathan Carr, Hailin Jin, Rafael Fonte, Michael Sittig, and David Forsyth. Automatic scene inference for 3D object compositing. *ACM Transactions on Graphics*, 33(3):32, 2014.

[4] Robert Maier, Kihwan Kim, Daniel Cremers, Jan Kautz, and Matthias Nießner. Intrinsic3D: High-quality 3D reconstruction by joint appearance and geometry optimization with spatially-varying lighting. In *ICCV*, volume 3, 2017.

[5] Richard A Newcombe, Shahram Izadi, Otmar Hilliges, David Molyneaux, David Kim, Andrew J Davison, Pushmeet Kohli, Jamie Shotton, Steve Hodges, and Andrew W Fitzgibbon. KinectFusion: Real-time dense surface mapping and tracking. In *ISMAR*, 2011.

[6] Qianyi Zhou and Vladlen Koltun. Color map optimization for 3D reconstruction with consumer depth cameras. *ACM Transactions on Graphics*, 33(4):155, 2014.

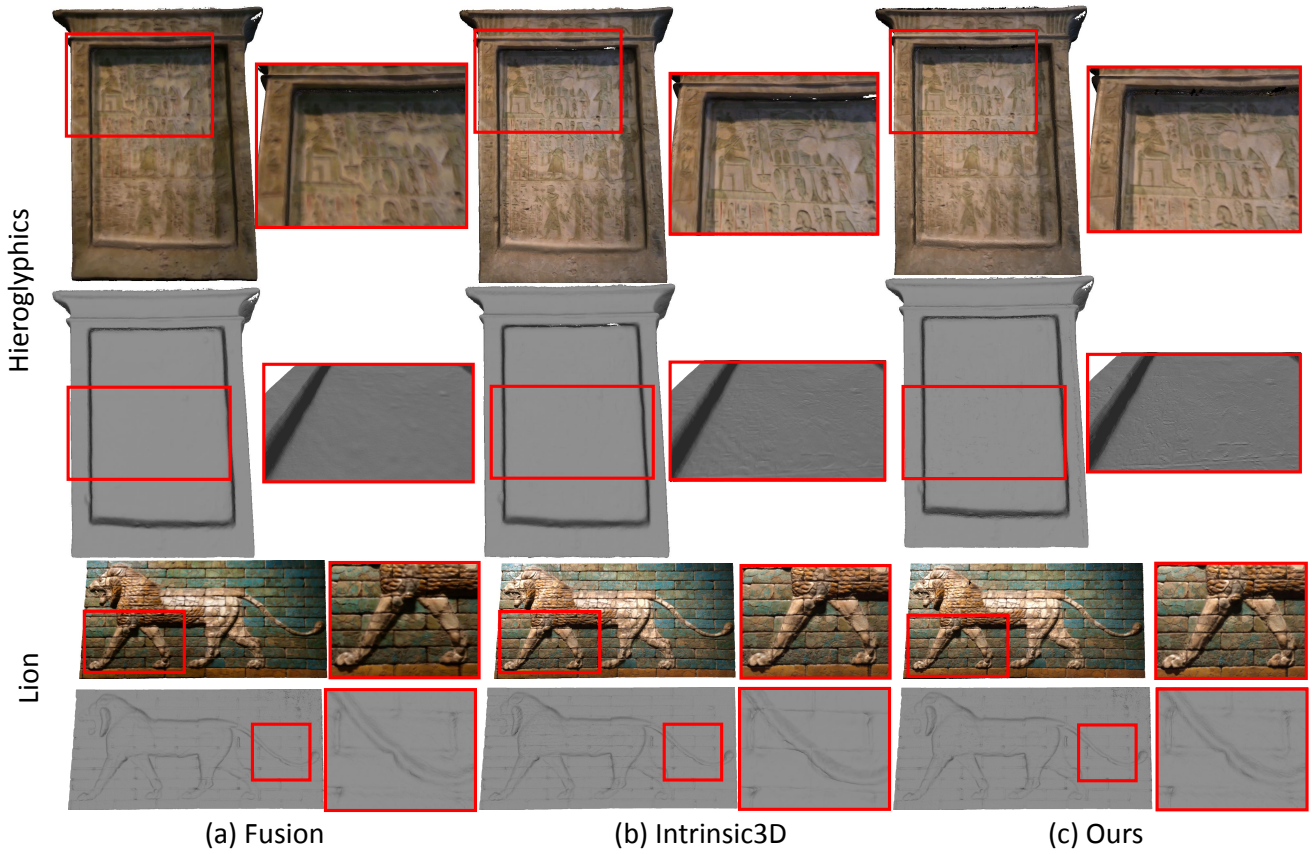

Figure 1. The comparison results with Intrinsic3D [4] on the datasets provided by Intrinsic3D [4]. (a) The reconstructed model and texture results. (b) The texture and geometry optimization results provided by Intrinsic3D [4]. (c) The texture and geometry optimization results of the proposed method.

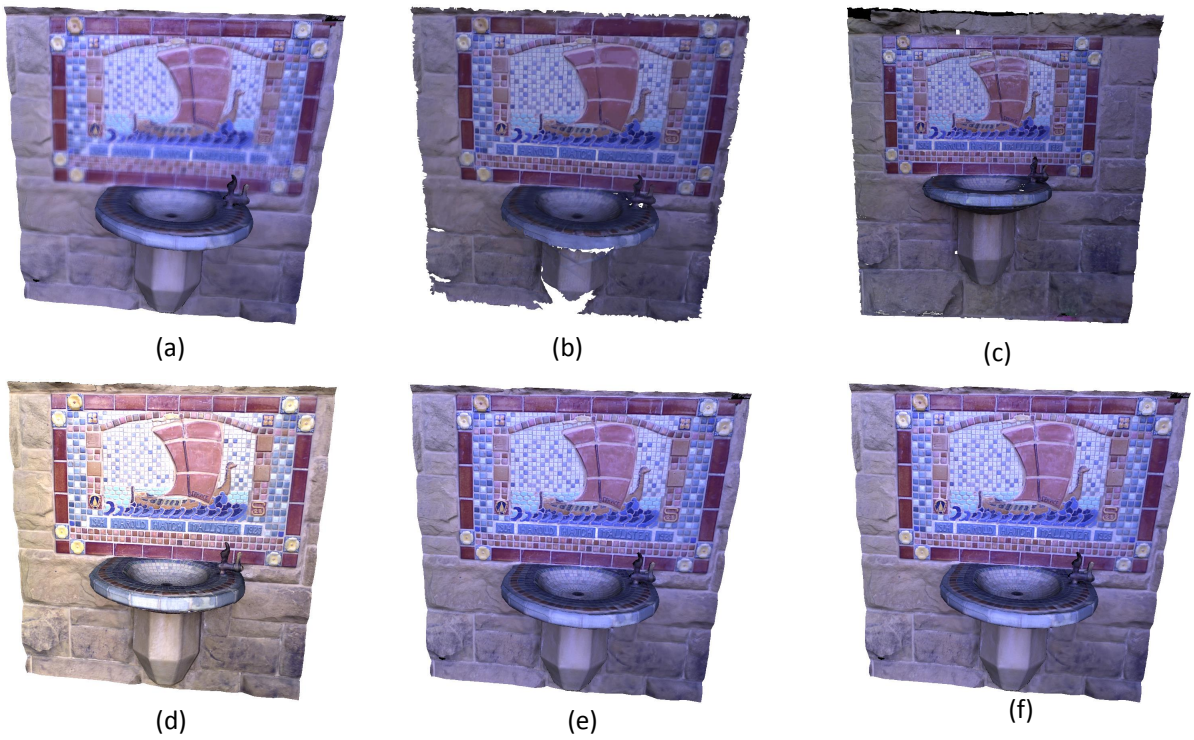

Figure 2. The qualitative comparisons with the state-of-the-art methods on public dataset *fountain* provided by [6]. (a) is the texture mapping result by KinectFusion [5] without any texture optimization. (b), (c), (d) and (e) are the texture optimization results by the methods of [1], [2], [6] and [4] respectively. (f) is the texture optimization result by the proposed method.

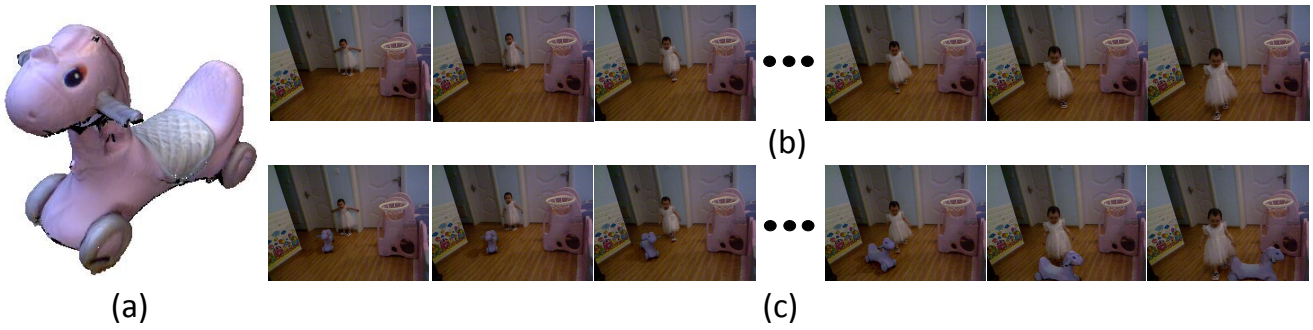

Figure 3. The application of the proposed method in video enhancement. (a) The textured 3D reconstruction model generated by the proposed method. (b) The original video frames. (c) The video frames enhanced with the 3D model generated by the proposed method.