VIEW-DEPENDENT VIRTUAL REALITY CONTENT FROM RGB-D IMAGES

Chih-Fan Chen, Mark Bolas, Evan Suma Rosenberg
{cfchen, bolas, suma}@ict.usc.edu
MxR
USC Institute for Creative Technologies

Motivation

* With the recent proliferation of high-fidelity head-mounted displays, there is increasing demand for **photorealistic 3D content** in virtual reality environments.
* Creating photorealistic models is not only **difficult** but also **time consuming**.
* We present a complete **end-to-end pipeline** for the capture, processing, and rendering of **view-dependent** 3D models from a single consumer-grade RGBD camera.

3D model | Fixed Texture | View-Dependent Textures from three different viewpoints

Our Approach / System Overview

1. Offline Stage:

a. Geometry Reconstruction
* Using Kinect Fusion to reconstruct the 3D model

b. Keyframe Selection and Triangulation
* Select by space to maximize the covered area of viewpoints
* Use convex hull to triangulate selected frames

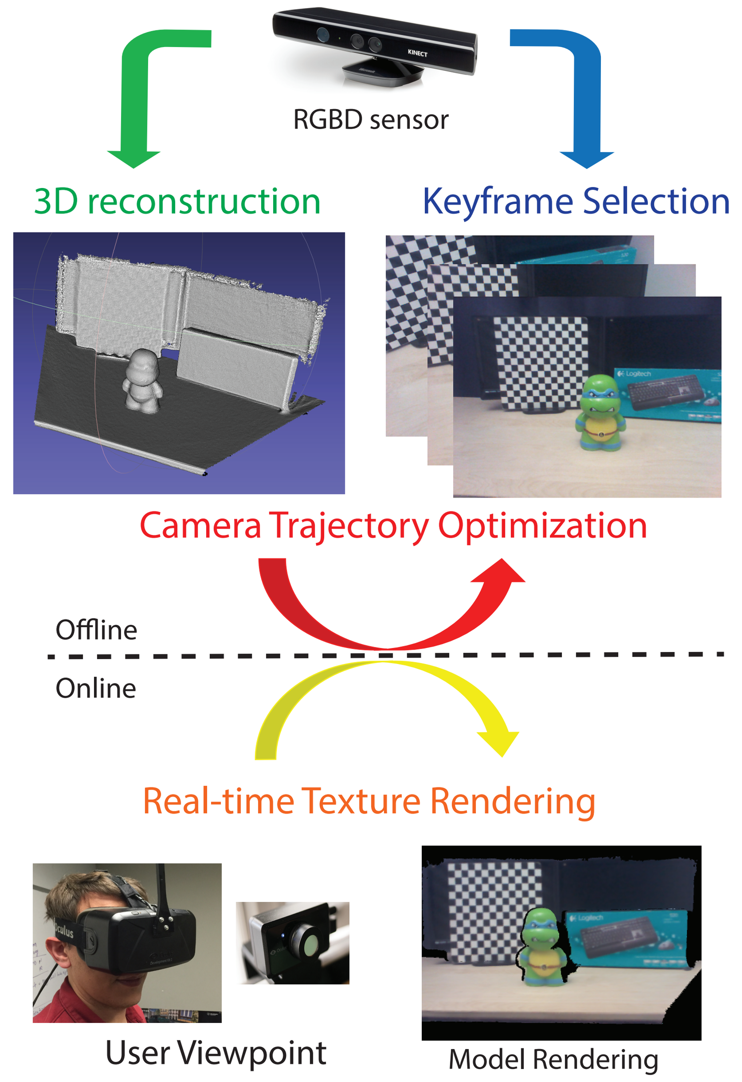
c. Camera Trajectory Optimization
* Find the best camera pose of each key frame to maximize the color and geometry agreement.

2. Online Stage:

a. Data Pre-processing
* Generate the visibility map for each keyframe to achieve real-time per-vertex rendering

b. Render Image Selection
* Select the triangle that intersects with the direction ray from HMD position to the model center

c. Real-time Texture Rendering:
* Find the corresponding RGB values from the keyframes of the selected triangle.
* Use barycentric coordinated to compute the weight of each frame and update the vertex color of the 3D model



Experimental Results

* The view-dependent 3D models captured by two depth sensors: Kinect and Intel Realsense
* The color images captured the light-burst effect and accurately reproduced it at run-time.
* Our method preserved the specularity of the object from different viewpoints
* Our method can render all the test models in only 10-15 milliseconds (i.e., 70 - 90 fps), which is sufficient for real-time high frequency rendering.

Kinect v1

Intel RealSense F200

Conclusion

* We proposed a pipeline for rapidly creating photorealistic virtual reality content with only single consumer-grade RGBD camera
* Our system can automatically reconstruct the geometry and generate a novel texture map for the model based on the HMD position in real-time.
* Our system can correctly reproduce the appearance of objects from the RGBD sequences without expert knowledge, making it a useful application for real-world 3D scanning.

Reference

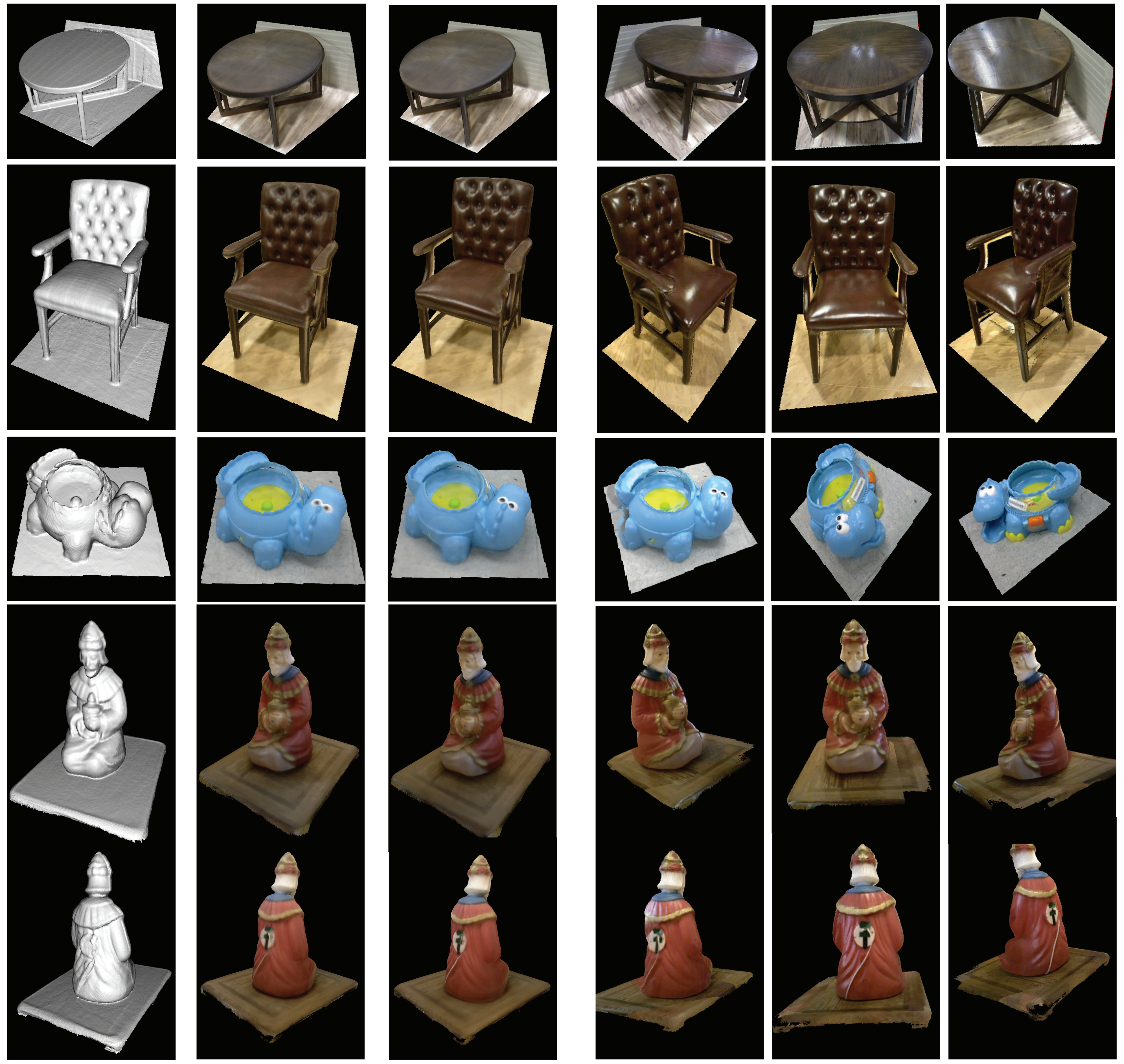
1. **Izadi, D. Kim, O. Hilliges, D. Molyneaux, R.` Newcombe, P. Kohli, J. Shotton, S. Hodges, D. Freeman, A. Davison, and A. Fitzgibbon. Kinectfusion: Real-time 3d reconstruction and interaction using a moving depth camera.** In Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology, UIST '11, pages 559–568, New York, NY, USA, 2011. ACM
2. **Q.-Y. Zhou and V. Koltun. Color map optimization for 3d reconstruc-tion with consumer depth cameras.** ACM Trans. Graph., 33(4):155:1–155:10, July 2014
3. **S. Choi, Q.-Y. Zhou, S. Miller, and V. Koltun. A large dataset of objectscans.** arXiv:1602.02481, 2016

More Results and Comparison

* We compared our results with Izadi et al. [1] and Zhou et al. [2] using a public RGB-D dataset from Choi et al.[3]
* All the RGBD sequences were captured by non-experts.
* Our system is able to handle the following conditions
  - various lighting conditions
  - various material (plastic, ceramics, leather, wood and metal)
  - various sizes (table, sculpture, toy)
  - 360-degree range (old man figurine)

Object	vertex	surface	images	color/depth images
Sculpture	208K	406K	108	3210 / 3225
Table	390K	763K	98	2906 / 2918
Chair	255K	495K	111	3299 / 3313
Toy	49K	97K	65	1936 / 1944
Figurine	99K	197K	96	2843 / 2856

Table 1: Information about the 3D models and images used in different examples.



Geometry | KinFu[1] | Zhou[2] | Our View-dependent Texture