A SYSTEM FOR ACTIVE IMAGE-BASED RENDERING

Cha Zhang and Tsuhan Chen

Dept. of Electrical and Computer Engineering, Carnegie Mellon University
5000 Forbes Avenue, Pittsburgh, PA 15213, USA
{czhang, tsuhan}@andrew.cmu.edu

ABSTRACT

In this paper, we develop a system for active image-based rendering (IBR). Active IBR is a framework that is capable of estimating the final rendering quality and capturing the next view at the position where the rendering quality can be improved the most. The result is a non-uniform capturing scheme for IBR. Experimental results on synthetic scenes have shown that active IBR outperforms uniformly captured IBR. In this paper, we set up an active IBR system that can capture real objects with off-the-shelf components. We show that our system can capture objects more intelligently than uniform capturing.

1. INTRODUCTION

Image-based rendering (IBR) has attracted much attention in the last few years. Unlike traditional rendering techniques where 3D geometry is dominantly used, IBR relies more on images. Depending on the availability of images and geometry, various IBR representations and rendering strategies have been proposed. For example, when dense image samples are available, new views can be rendered by light ray interpolation [1][2][3][4]. When some sort of geometry is known or obtainable, IBR can easily incorporate the geometry information into rendering, such as the work in [5][6][7][8][9]. When available images are limited and no geometry information is known, IBR can still be applied after geometry reconstruction from images [10][11].

Work in literature has largely focused on IBR representations and rendering. However, little work has been performed on the capturing process, which is as important as the rendering process. In [12], we proposed a framework called active image-based rendering. In active IBR, three stages are employed recursively, i.e., rendering quality estimation, capturing and geometry refinement. Compared with the traditional uniform capturing method such as that in [2][4], active IBR can capture the scenes intelligently by putting more cameras on where the rendering quality can be improved the most. Experimental results on synthetic data showed that active IBR outperforms traditional IBR based on uniform capturing.

In this paper, we develop an active IBR system for capturing real objects. The system is composed of a Pyro 1394 webcam, a Directed Perception pan-tilt (model PTU-46-17.5), a P4 2.4GHz computer and some accessories, as is shown in Figure 1. The benefit of using a pan-tilt instead of a normal turntable is that its rotation can be computer-controlled. We constrain the pan-tilt to rotate only horizontally. During the rendering, the above setup simulates the case where the environmental light and the camera are relatively static. When the environment has only ambient light, this is also the same as rotating the camera on a circle around the object.

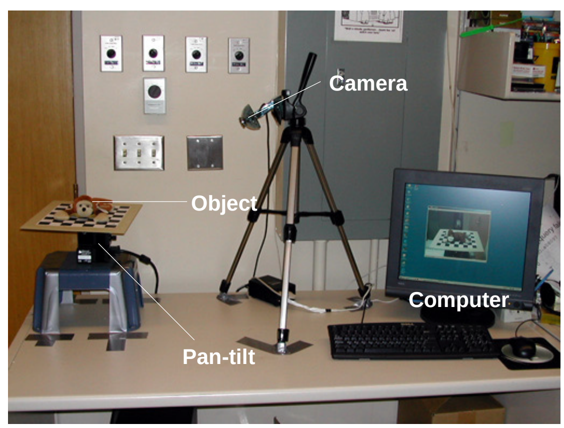

Figure 1 Our active IBR capturing system.

Given any object and the maximum number of images that the system is allowed to capture, our system will try to capture images at places where it thinks are the most beneficial by rotating the pan-tilt to certain angles. The positions are determined by our rendering quality estimator that can tell where the rendering quality can be improved the most. We show that active IBR will capture the object more intelligently than traditional uniform capturing.

The paper is organized as follows. Section 2 reviews the active IBR algorithm. The hardware setup is described in Section 3. Experimental results and conclusions are given in Section 4 and 5, respectively.

2. ACTIVE IMAGE BASED RENDERING

2.1. The color consistency criterion

Active IBR is based on one simple criterion: the *color consistency criterion*. For Lambertian scenes, color consistency criterion verifies that light rays from the same object surface point should have the same color (intensity). However, completely Lambertian scenes are rare in practice. In [12], we proposed a color consistency criterion for non-Lambertian scenes. That is, light rays from the same surface point should have the same color, as long as their angles of emission are close enough.

2.2. The rendering quality estimator

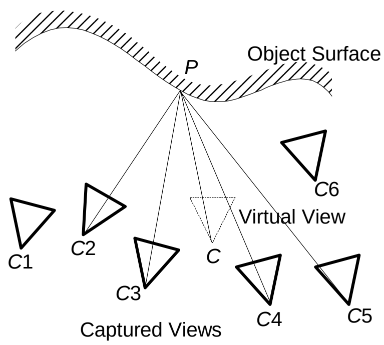

Figure 2 Depth-driven IBR rendering scheme.

In image-based rendering, to render the light rays that compose the virtual view, we need to interpolate them from captured light rays. For example, in Figure 2, C_1, C_2 etc. are captured images. To get the virtual light ray CP, we interpolate it from captured light rays such as C_2P, C_3P, C_4P and C_5P. If all the captured light rays satisfy the color consistency criterion, e.g., C_2P, C_3P, C_4P and C_5P do have similar colors, interpolation among them will not cause artifacts such as the ghosting effect. Unfortunately, color consistency may not be met due to multiple reasons, such as geometry inaccuracy, non-Lambertian reflection, occlusions and sensor noise. All these problems can be remedied by taking more images.

Therefore, color consistency is a good estimate of the rendering quality. Assume that we have some volumetric representation of the scene. For a certain neighborhood, the rendering quality can be estimated by scanning through all the voxels and measure the color consistency among images in the neighborhood. In [12], we captured the scene by putting cameras on a plane, similar to that in lightfield/Lumigraph [2][3]. Since cameras are on a 2D plane and each captured image is also 2D, the resultant IBR is 4D. *Quadruple* was used as the neighborhood unit. As shown in Figure 3, on the camera plane (s,t), each dot represents a camera position where an image was captured. We always find the quadruple that has the worst color consistency and split it by taking 5 new images at the middle positions marked as solid dots.

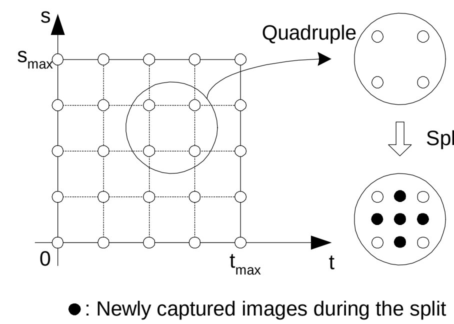

Figure 3 Active IBR on a capturing plane.

In the system setup described in Figure 1, cameras are on a circle instead of a plane. Therefore, the IBR dimension is reduced from 4 to 3. We consider each neighboring captured image pair as the basic unit, as shown in Figure 4. When a voxel is projected to a neighborhood (two images), the color consistency is measured by the color difference between corresponding pixels. Once the image pair with the worst color consistency is chosen, we split the pair by taking one more image at the center, as represented by a solid dot in Figure 4.

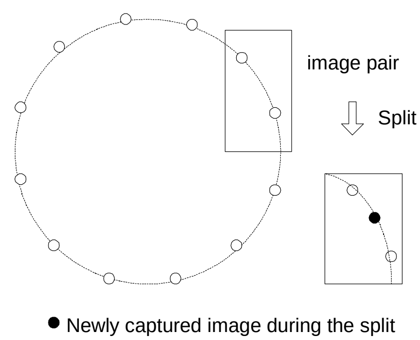

Figure 4 Active IBR in the current system.

2.3. Geometry reconstruction

In the rendering quality estimator, we assume that we have some volumetric geometry model to verify the color consistency criterion. In practice, we may or may not have such a model. If we do not have the model, we can obtain it through voxel coloring [11].

As the scene can be non-Lambertian, we use the extended color consistency criterion in Section 2.1 during the voxel coloring. In [12] we have shown that such criterion produces better geometry than the traditional criterion. In our setup, the camera projection center is always above the object, thus a voxel coloring algorithm can be applied by scanning the voxel planes from top to bottom.

Another option is to apply the space-carving algorithm [13] for getting the model, because if we want to scan the voxel planes from other directions (left to right, front to back), the cameras are not all on one side of the scanning plane, which violates the condition of voxel coloring. We choose the top-to-bottom voxel coloring because it has much less computational complexity.

2.4. The overall AIBR procedure

The overall AIBR procedure is shown in Figure 5. We start the algorithm by capturing an initial set of images uniformly. We then apply voxel coloring algorithm to obtain a 3D voxel model. With the voxel model and the color consistency criterion, we can locate which image pair to split with the algorithm in Section 2.2. After the splitting, we may continue capturing new images or applying voxel coloring again for geometry refinement (Since the computation of voxel coloring is heavy, we may prefer to do it for every several capturing steps). The whole process loops until the maximum number of images is reached, or all the images have been color consistent.

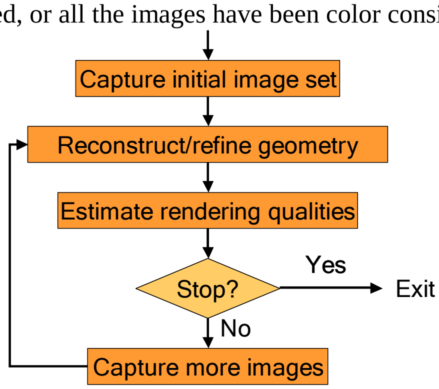

Figure 5 The flow of AIBR capturing.

3. THE ACTIVE IBR SYSTEM

3.1. The hardware

As explained in Section 1, we use off-the-shelf components to build our system. The pan-tilt can hold up to 4lbs and have a resolution of 3.086 arc minute (0.051 degree), which is good enough for our task. One limitation is that the pan-tilt can cover only about 310º pan rotation instead of 360º. The Pyro 1394 webcam is low-cost but has significant lens distortion and noise. We overcome the lens distortion by camera calibration and reduce the noise by smoothing the obtained image before processing.

One thing we notice is that voxel coloring is relatively sensitive to calibration errors. As we do not want to calibrate our cameras repeatedly, we have to fix the position of the pan-tilt with respect to the camera, and the object position with respect to the pan-tilt.

3.2. Camera calibration

For a real system, the most challenging task is camera calibration. We discuss the intrinsic parameter calibration and extrinsic parameter calibration separately.

3.2.1. Intrinsic parameter calibration

We calibrate the camera using the online calibration toolbox by Bouguet [14]. This toolbox allows us to capture a checkboard at random positions and obtain the intrinsic and extrinsic camera parameters. The constraint is that we have to specify the four corners of the checkboard manually for each image. Since intrinsic parameters are fixed for a given camera, such manual operation is affordable.

3.2.2. Extrinsic parameter calibration

The extrinsic parameters of an image describe the rotation and translation between the world coordinate and the camera coordinate. These parameters differ from image to image. To tell the extrinsic parameters of the captured images during active IBR, we need to find the correspondence between the pan-tilt pan position and the extrinsic parameters.

Although theoretically there is a unique relationship between the pan position and the extrinsic parameters, we found that such a model may introduce huge errors, as there is always mechanic inaccuracy for a real-world system. Instead, we adopt the following strategy.

We capture a checkboard on top of the pan-tilt at several pan positions (15 in our experiment). We calibrate them manually. For arbitrary pan position in-between these positions, we interpolate the existing parameters to get new parameters.

4. EXPERIMENTAL RESULTS

We have captured various objects with our system. Figure 6 shows some scenes we captured, their rendering results and reconstructed geometry. In all cases we start with 12 uniformly captured images and limit the maximum number of images to be 48. We are able to render the scene reasonably well even with very few images. Even though the reconstructed geometry is not perfect, for image based rendering, geometry distortion is allowable as long as the color consistency criterion is satisfied. Notice that we also assume the rendering virtual camera is close to the capturing camera trajectory, because the “maximal photo-consistent shape” [13] obtained through voxel coloring may be poor for views far from the captured views. Nevertheless, if the acquired geometry is accurate enough, viewing the object from other positions are straightforward.

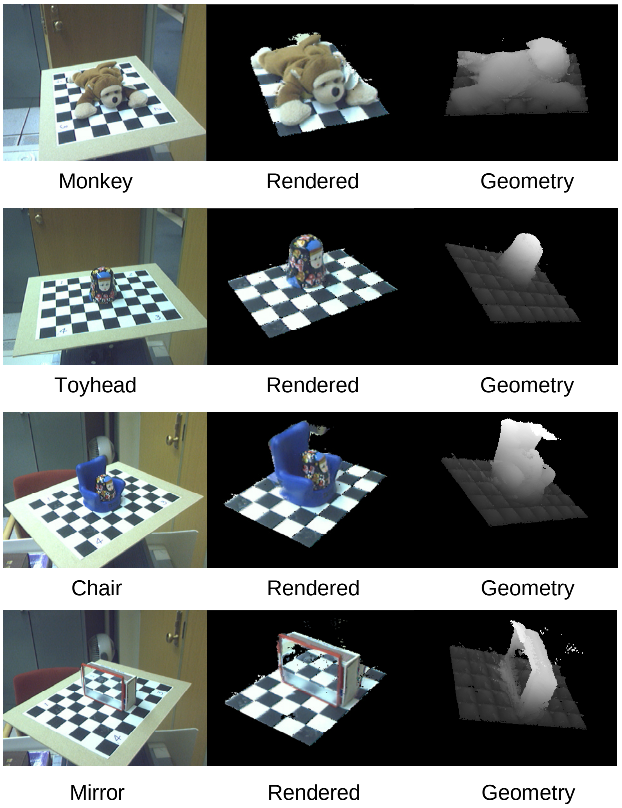

Figure 6 Some scenes captured by our system.

The effectiveness of active IBR compared to uniform capturing is shown in Figure 7. The object captured is a mirror bound with a matches box. Each dot in the center image represents one view captured by the active camera. The opening on the left of the circle is due to the limited rotation range of the pan-tilt. It can be seen that active IBR puts much more efforts on the mirror side of the object, which is highly non-Lambertion. By examining the geometry model obtained for the object (Figure 6), we found that due to the reflective property of the mirror, the voxel coloring algorithm failed to reconstruct any voxels there. Therefore the color consistency is very poor on the mirror side, which leads active IBR to capture more images there. This is a good example that active IBR can automatically recognize where the rendering quality can be improved the most and capture images more effectively.

5. CONCLUSIONS

We developed a system for capturing real-world objects with active IBR. We verified that active IBR is an intelligent way of capturing IBR scenes. We plan to extend our results on large scenes/environments captured with hand-held cameras.

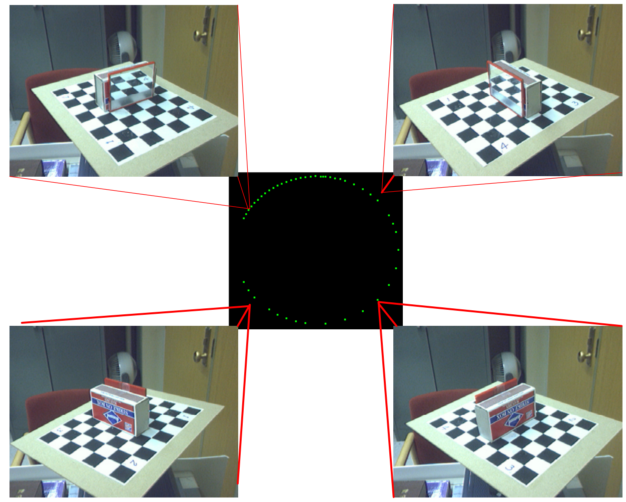

Figure 7 Camera positions for the Mirror scene.

REFERENCES

[1] L. McMillan and G. Bishop, “Plenoptic modeling: an image-based rendering system”, Computer Graphics (SIGGRAPH’95), pp. 39-46, Aug. 1995.

[2] M. Levoy and P. Hanrahan, “Light field rendering”, Computer Graphics (SIGGRAPH’96), pp. 31, Aug. 1996.

[3] S. J. Gortler, R. Grzeszczuk, R. Szeliski and M. F. Cohen, “The Lumigraph”, Computer Graphics (SIGGRAPH’96), pp. 43-54, Aug. 1996.

[4] H.Y. Shum and L.-W. He, “Rendering with concentric mosaics”, Computer Graphics (SIGGRAPH’99), pp.299-306, Aug. 1999.

[5] S. Chen and L. Williams, “View interpolation for image synthesis”, Computer Graphics (SIGGRAPH’93), pp. 279-288, Aug. 1993.

[6] S. M. Seitz and C.M. Dyer, “View morphing”, *Computer Graphics (SIGGRAPH’96)*, pp. 21-30, Aug. 1996.

[7] J. Shade, S. Gortler, L. W. He, and R. Szeliski, “Layered depth images”, *Computer Graphics (SIGGRAPH’98)*, pp. 231-242, July 1998.

[8] P. E. Debevec, C. J. Taylor, and J. Malik, “Modeling and rendering architecture from photographs: A hybrid geometry- and image-based approach”, *Computer Graphics (SIGGRAPH’96)*, pp. 11-20, Aug. 1996.

[9] C. Buehler, M. Bosse, L. McMillan, S. Gortler, and M. Cohen, “Unstructured Lumigraph rendering”, *Computer Graphics (SIG-GRAPH’01)*, pp 425-432, Aug. 2001.

[10] M. Pollefeys, "Self-calibration and metric 3D reconstruction from uncalibrated image sequences", *Ph.D. Thesis*, ESAT-PSI, K. U. Leuven, 1999.

[11] S. M. Seitz and C. R. Dyer, “Photorealistic Scene Reconstruction by Voxel Coloring”, *CVPR 1997*, pp. 1067-1073.

[12] C. Zhang and T. Chen, “Active Scene Capturing for Image-Based Rendering”, *Carnegie Mellon Technical Report: AMP03-02*.

[13] K.N. Kutulakos and S. Seitz, “A Theory of Shape by Space Carving”, *Technical Report TR692*, Computer Science Dept., U. Rochester, May, 1998.

[14] http://www.vision.caltech.edu/bouguetj/calib_doc/.