ThRift: An omnidirectional image based virtual reality system

Abdoulaye Maiga*and Naoki Chiba†

Abstract

Virtual reality (VR) is collecting attention thanks to the progress of hardware such as head-mounted displays (HMD) and 3D graphics chips. Because most of the systems, however, are based on computer graphics (CG), they require a significant amount of time and energy to build 3D models manually. One solution is to build them automatically from images, but the display quality lacks realism due to low resolution and some artifacts. We propose a system based on image-based rendering that can provide the user with a photo realistic and immersive walk-through experience. Depending on the user's location, we switch images. Although the time consuming part is image acquisition, it takes less than an hour for 30 images by a consumer omnidirectional camera. The user is able to walk around in a room interactively in realtime, while looking in any direction with a wide-angle HMD.

1 Introduction and Previous Work

Virtual reality has long consisted in using a 3D environment in which the user can roam. However, designing 3D environments requires both expertise and time to obtain high levels of accuracy. It has been a challenge in Computer Graphics to retrieve 3D models from a set of images as described in [1] [2] and [3]. These methods try to provide an interactive experience: a sense of virtual presence or exploration of a real world scene. However, remaining artifacts and misplacements lacks of photo-realistic quality.

Fig. 1. Room used for demonstration

To solve this problem, Image based rendering (IBR) techniques are gaining considerable interests for over three decades. The first system dates back to Aspen Movie Maps in 1978 in [4], which can provide a virtual walkthrough on the streets. More recent works such as Apple's QuickTime VR in [5], Microsoft's interactive Exploration in [6] and Photo Tourism in [7] have proposed systems for user centered interactive guides. However, these systems require a specially designed imaging device or user's viewing directions are limited.

We aim at providing an immersive solution that is simple only with consumer-level components. Applications include navigation of hotel rooms or condos. Fig. 1 shows an example.

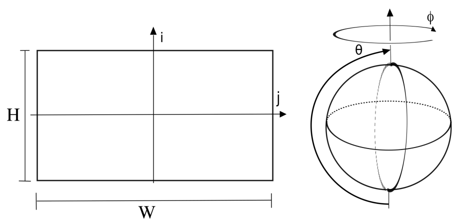

Fig. 2. Equirectangular and Spherical Notations

2 Capturing omnidirectional images for navigation

In order to simplify the image-capturing process, we use a consumer camera Theta made by Ricoh. This camera can capture a fully spherical 360^o image in Equirectangular format, which is generated from 2 fisheye lenses looking in opposite directions. In this section, we describe how we convert the image to a sphere for allowing the user to see an arbitrary viewing direction.

* Dept. of Information and Computer Science, Keio University
† Rakuten Institute of Technology, Rakuten, Inc.

2.1 Equiretangular Image and Spherical Warping

Each input image is an Equirectangular image with a height H and a width $W = 2*H$ since the coverage is fixed to 360^o horizontally and 180^o vertically as shown in Fig. 2. We can warp this image to the surface of a sphere using a forward projection. We can obtain spherical coordinates (θ, ϕ) from image coordinates (i, j):

$$\theta = i * \frac{\pi}{H}$$
$$\phi = j * \frac{2*\pi}{W}. \quad (1)$$

These spherical coordinates can then be used to create a unit sphere centered around the origin and the Cartesian values (x, y, z) can be computed by the following conversion:

$$x = cos(\theta) * cos(\phi)$$
$$y = cos(\theta) * sin(\phi) \quad (2)$$
$$z = sin(\theta).$$

This coordinate system is well adapted for providing omnidirectional views. These images are used as environment maps.

3 System Configuration

In order to realize a walkthrough experience, images are captured by the camera placed on an off-the-shelf tripod with an $80cm$ gap between them to simulate the human walking step distance. For this phase we suppose that the images are well aligned and ordered. For our demonstration, 30 images around the table were captured.

Fig. 3. Stereoscopic images of guiding arrow

The converted images in spherical coordinate system are then inputted via CG software (Unity) to create a 3D walkthrough allowing to navigate to different locations. Arbitrary views for user's viewing directions are displayed on an Oculus Rift DK2 HMD, which has a wide-field of view (110^o) equipped with positional and orientational sensors. As the user walks by using a gamepad, the system switches omnidirectional images according to the position.

To guide the user, 3D floating arrows are overlaid as shown in Fig. 3. The arrows appear depending on the direction the user is facing and the current position and point toward the following positions accessible to the user. As a result the users can easily orientate themselves in the room. These arrows also give a sense of depth.

4 Discussion

Our system enables the user to explore the room interactively by walking around the table while viewing arbitrary directions in realtime. One of the advantages of our system is its simplicity. A full room can be acquired in less than an hour in comparison to generating a full CG 3D model which could take weeks by a human operator.

Having images separated by $80cm$ seems to be highly efficient for human motion and to make users feel as if they were physically walking.

5 Conclusion

This work is a step toward integrating VR into daily lives in a simple and natural way. The next steps could involve incorporating sounds and videos to fully create an immersive world.

References

[1] Y. Furukawa et al., Reconstructing Building Interiors from Images, in *IEEE ICCV 2009.*

[2] Y. Zhang et al., PanoContext: A Whole-room 3D Context Model for Panoramic Scene Understanding, in *ECCV 2014.*

[3] R. Cabral et al., Piecewise Planar and Compact Floorplan Reconstruction from Images, in *IEEE CVPR 2014*

[4] A. Lippman Movie-maps: An application of the optical videodisc to computer graphics, in *ACM SIGGRAPH 1980.*

[5] S. E. Chen, Quicktime VR: An image-based approach to virtual environment navigation, In *ACM SIGGRAPH 1995.*

[6] M. Uyttendaele et al., High-quality Image-based Interactive Exploration of Real-World Environments, *Technical report MSR-TR-2003-61, Microsoft Research 2003.*

[7] N. Snavely et al., Photo tourism: Exploring photo collections in 3D, in *ACM SIGGRAPH 2006.*