

From Inside and Outside: Immersive 3D Video Generation

Hansung Kim*, Ryuuki Sakamoto*, Itaru Kitahara** and Kiyoshi Kogure*
Knowledge Science Lab, ATR

1 Introduction

We have introduced “Cinematized Reality,” the aim of which is to record unexpected moments and create movie-like footage by a virtual camera generated from eight environmental cameras [Sakamoto et al. 2005][Kitahara et al. 2005]. However, the system has problems such as phantom volumes and a trade-off between the size of modeling space and the resolution of textures, because the system uses only fixed environmental cameras. Therefore, we improved the system by inserting an omni-directional multi-camera into the modeling space. The camera can be moved to any place in the space to augment the quality of the 3D model and rendered video with high resolution. We also achieved minor improvement in segmentation and texture mapping so that enhanced the quality of final video from a virtual camera.

2 Exposition

We set up multiple environmental cameras on the wall and the ceiling to surround the target object and an omni-directional multi-camera inside the working space. When target objects are captured by cameras, each capturing PC segments the objects and transmits the masks and color textures to a 3D modeling server via network. The modeling server then reconstructs 3D models of each object from the gathered masks and generates a video at the designated point of view with the 3D model and texture information.

We used the Ladybug2™ camera as the omni-directional multi-camera system, which consists of six 1024×768 color CCDs, with five CCDs positioned in a horizontal ring and one pointing straight up. The extrinsic parameters of the omni-directional camera can be

changed by moving it inside the modeling space. Therefore, the system calibrates one of the sub-cameras using a chessboard pattern attached on ceiling in the modeling space, and calculates all parameters of the other sub-cameras using the geometrical relationship between sub-cameras in real time. When R_M and T_M are rotation and translation parameters of the reference sub-camera, respectively, and R_{Δ} and T_{Δ} are the geometrical relationship between sub-cameras, the extrinsic parameters of the other sub-cameras can be derived as follows.

$$\begin{cases} M_{C_i} = R_M M_w + T_M \\ M_{C_i} = R_{\Delta} (R_M M_w + T_M) + T_{\Delta} \end{cases} \Rightarrow \begin{cases} R_i = R_{\Delta} R_M \\ T_i = R_{\Delta} T_M + T_{\Delta} \end{cases} \quad (1)$$

In order to reconstruct the 3D shape of the object, we employ a shape-from-silhouette method. An intensity-based background subtraction method was used to segment the foreground and background regions in the input multiple images, and then we refined the silhouette in the HSV color domain to remove shadows.

In virtual view rendering, a micro-facet billboarding technique is employed. It approximates a 3D object’s shape to micro-facet billboards on the surface and expresses the object’s appearance by mapping its captured image. This technique has the advantage of not wasting the resolution of captured images, even if an accurate 3D shape is not available.

Figure 1 shows reconstructed 3D models and final rendered images without/with inner cameras, respectively. The model generated using both inner and outer cameras looks more natural because the redundancies in the models are carved out by the inner cameras, and the virtual camera can capture much more distinct textures because the inner cameras are placed nearer to the objects.

Figure 2 shows snapshots of the generated 3D free-viewpoint videos. It is clear that the system generally renders natural scenes from any point of view.

3 Conclusion

We improved the Cinematized Reality system by using multiple outer/inner cameras. Cooperation between environmental cameras and the omni-directional multi-camera solved the limitation of the previous system and provides more natural and vivid videos.

Acknowledgement

This research was supported by the National Institute of Information and Communications Technology.

References

- SAKAMOTO, R. et al. 2005. Cinematized Reality: Cinematographic Camera Control in 3D Videos, ACM SIGGRAPH (Sketch)
- KITAHARA, I. et al. 2005. Cinematized Reality: Cinematographic Camera Controlling 3D Free-Viewpoint Video, CVMP, 154-161



(a) Reconstructed 3D model (b) Texture mapping
Fig. 1. Final results (without / with inner cameras)



Fig. 2. Snapshots of the generated 3D free-viewpoint videos

*email: {hskim, skmt, kogure }@atr.jp

**email: kitahara@iit.tsukuba.ac.jp