Real-Time Video-Based Rendering from Uncalibrated Cameras Using Plane-Sweep Algorithm

Songkran Jarusirisawad, Vincent Nozick, Hideo Saito
Department of Information and Computer Science, Keio University
3-14-1 Hiyoshi, Kohoku-ku, Yokohama, 223-8522, Japan
songkran,nozick,saito@ozawa.ics.keio.ac.jp

I. Introduction

Image based rendering (IBR) is the process for creating new views from several input images of a static scene. Video based rendering (VBR) can be thought as an extension of IBR to the dynamic scene. We present a new VBR system that creates new views of the scene from uncalibrated cameras in real-time. The main different of our system from the other systems is that our method does not require the information of camera intrinsic parameters or the Euclidean information of a scene. For obtaining a geometrical relation among the cameras, we use Projective Grid Space (PGS)[7] which is 3D space defined by epipolar geometry between two basis cameras. The other cameras are registered to the same 3D space by trifocal tensors between these basis cameras. Our system simultaneously reconstructs and renders novel view using our proposed plane-sweep algorithm in Projective Grid Space. To achieve real-time performance, we implemented our plane-sweep method in a graphics processing unit (GPU). The results show the efficiency of our proposed system.

II. Previous Works

This section surveys previous recent works on real-time VBR techniques. Complex algorithms used for off-line methods are not suited for real-time implementation. Therefore, we can not expect the online methods have the same accuracy provided by off-line ones.

One of the most popular online VBR methods is the visual hulls algorithm. This method extracts the silhouette of the main object of the scene on every input image. The 3D shape of this object is then approximated by the intersection of the projected silhouettes. There exist several online implementations of the visual hulls described in [4]. The methods proposed by Li et al.[3] is probably the easiest to implement. The main drawback of the visual hulls methods is the impossibility to handle the background of the scene. Hence, only one main object can be rendered. Furthermore, the visual hulls methods usually require several computers, which make their use more difficult.

Among all these visual hulls methods, the image-based visual hulls presented by Matusik et al.[5] is an online VBR method from uncalibrated cameras. This method reconstructs visual hull of the object using epipolar geometry in an image space instead of 3D space. Thus, it does not suffer from quantization artifacts of voxels like in ordinary visual hull. This method can create new views in real-time from four cameras. Each camera is controlled by one computer and an additional computer creates the new views.

Finally, some plane-sweep implementations achieve online rendering using graphic hardware, graphics processing unit (GPU). The plane-sweep algorithm introduced by Collins[1] was adapted to on-line rendering by Yang et al. [8]. They computed new views in real-time from five cameras using four computers. Geys et al.[2] also used a plane-sweep approach to find out the scene geometry and rendered new views in real-time from three cameras and one computer. Nozick and Saito [6] introduced a plane sweep implementation for moving camera where all the input cameras are calibrated in real-time using ARToolkit markers.

Our method belongs to the real-time VBR using plane-sweep algorithm group. The main different is that in the previous works they assume that cameras are strongly calibrated. All reconstruction or depth recovery is done in the Euclidean space. This paper present a new method for online video-based rendering from uncalibrated cameras using plane-sweep algorithm.

III. Plane-Sweep Algorithm in PGS

To implement the plane-sweep algorithm, we need to project 3D points into image frame of each camera including the virtual one. Projective Grid Space (PGS)[7] allows us to define that 3D space and find the projection without knowing cameras intrinsic parameters or the Euclidean information of a scene. 3D space of PGS is defined by image coordinate of two arbitrarily selected cameras called basis camera 1 and basis camera 2. Fundamental matrix and trifocal tensors are estimated for relating all cameras to PGS.

In ordinary plane-sweep algorithms that work in the Euclidean space, near and far planes defining the the 3D space for recovering depth maps or rendering new views are selected from the Euclidean position of a scene. We proposed a new plane-sweep algorithm in PGS in which near and far planes are directly defined by image coordinate of basis camera 2 as illustrated in Figure 1.

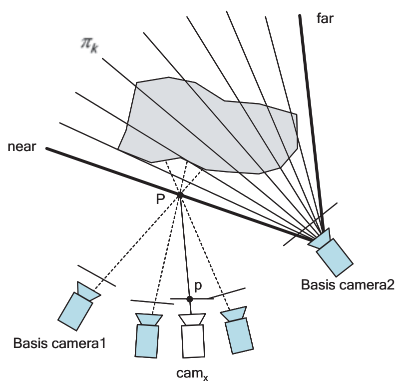

Fig. 1. Defining planes for doing plane-sweep in Projective Grid Space.

IV. System Configuration

We implemented our proposed method on a PC Pentium4 3.40 GHz with a NVIDIA Quadro FX 3400. Four webcam with a resolution 320x240 are used to capture input videos. We select two cameras for defining the Projective Grid Space as shown in Figure 2.

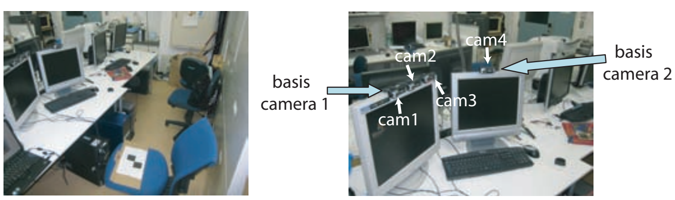

Fig. 2. Camera configuration.

Fundamental matrix between camera 1 and 4, two trifocal tensors defined by camera 1,4,2 and camera 1,4,3 are estimated for weakly calibrating cameras to PGS. 2D-2D correspondences for estimating fundamental matrix and trifocal tensors can be selected from feature points in a scene. In our implementation, we wave marker around a scene and track features for 2D-2D correspondence automatically.

V. Result

Figure 3 shows the resulting new view video from selected input frames. We use 50 planes in this experiment and our system can render new views at 15 fps. The ratio written under each figure is a virtual camera position relatives to two real reference camera. Camera 4 is selected as basis camera 2 that defines planes for depth recovery. In Figure 3, the green lines in camera 4 show all planes we used for plane-sweep algorithm in PGS. Even our method is fast and can run in real-time, we still get a good visual quality.

We also generate depth images in PGS to show the intermediate results. Figure 4 shows depth images created by plane-sweep algorithm in PGS using the same input frame as shown in Figure 3.

All video results in this paper are available for downloading at http://www.hvrl.ics.keio.ac.jp/∼songkran/psivt/index.html .

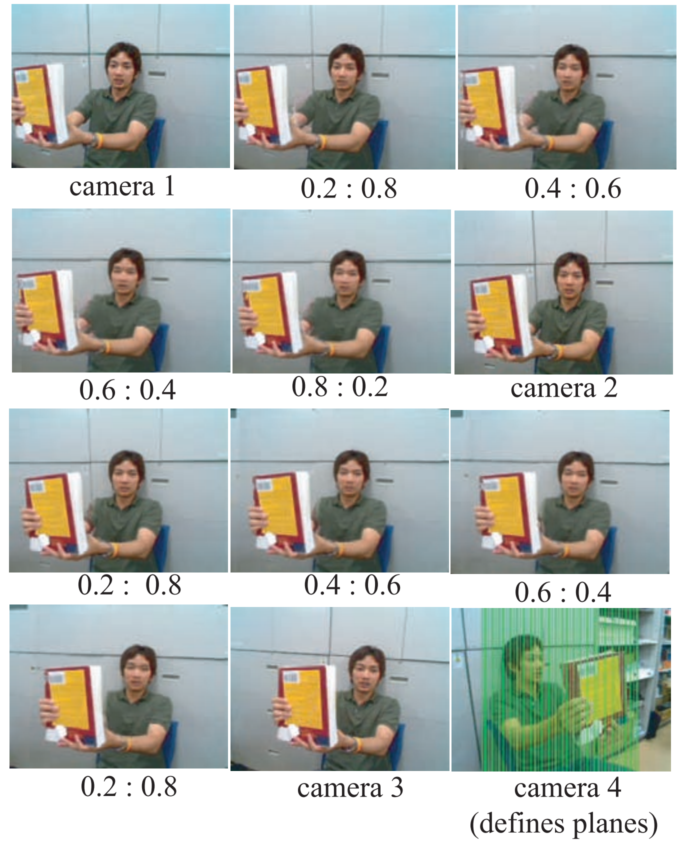

Fig. 3. Result new view images.

Fig. 4. Depth images in PGS from plane-sweep algorithm.

References

[1] R. T. Collins, "A space-sweep approach to true multi-image matching," in *Proceedings of CVPR*, 1996, pp. 358–363.

[2] I. Geys, S. D. Roeck, and L. V. Gool, "The augmented auditorium: Fast interpolated and augmented view generation," in *In Proceedings of the 2nd IEE European Conference on Visual Media Production*, 2005, pp. 94–103.

[3] M. Li, Magnor, and H. P. Seidel, "Online accelerated rendering of visual hulls in real scenes," *Journal of WSCG*, vol. 11, no. 2, pp. 290–297, 2003.

[4] M. A. Magnor, *Video-based Rendering*. A K Peters Ltd, 2005.

[5] W. Matusik, C. Buehler, R. Raskar, S. J. Gortler, and L. McMillan, "Image-based visual hulls," in *Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques*. ACM Press/Addison-Wesley Publishing Co., 2000, pp. 369–374.

[6] V. Nozick and H. Saito, "Real-time free viewpoint from multiple moving cameras," in *Advanced Concepts for Intelligent Vision Systems (ACIVS)*, 2007, pp. 72–83.

[7] H. Saito and T. Kanade, "Shape reconstruction in projective grid space from large number of images," in *Proceedings of CVPR*, June 1999.

[8] J. C. Yang, M. Everett, C. Buehler, and L. McMillan., "A real-time distributed light field camera," in *Proceedings of the 13th Eurographics Workshop on Rendering*, 2002, pp. 77–86.