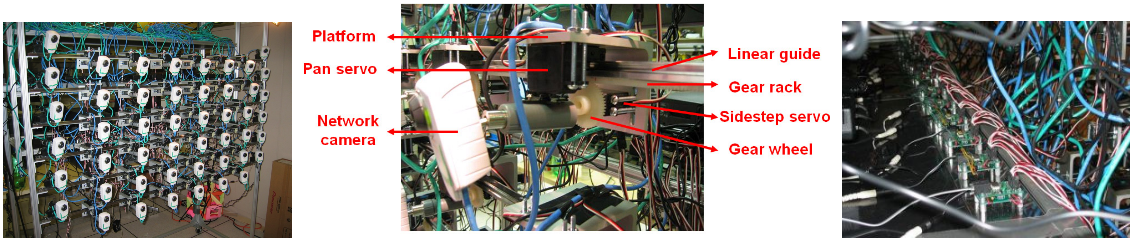
A Self-Reconfigurable Camera Array

Cha Zhang* Tsuhan Chen†
Dept. of Electrical and Computer Engineering, Carnegie Mellon University



(a) Overview. (b) Each mobile camera unit. (c) The controllers.

Figure 1: The self-reconfigurable camera array (for more information please visit http://amp.ece.cmu.edu/projects/MobileCamArray/)

1 Introduction

We present a self-reconfigurable large camera array system, as shown in Figure 1(a), which captures video sequences from an array of mobile cameras, renders novel views in real time and reconfigures the camera positions to improve the rendering quality.

Our system is composed of 48 Axis 205 network cameras mounted on mobile platforms. Each camera captures images at 320×240 and sends them to a central computer for rendering. There have been several large camera arrays on similar scale built in the literature, such as the distributed light field camera (DLFC) in [Yang et al. 2002a] and the Stanford multi-camera array [Wilburn et al. 2002]. Albeit having so many cameras, DLFC rendering exhibits ghosting or aliasing artifacts due to the constant depth assumption. On the other hand, work in [Yang et al. 2002b], which used 5 cameras, showed promising results through on-the-fly geometry reconstruction. In this sketch, we propose an efficient and flexible view-dependent geometry reconstruction algorithm for the rendering of images captured by our large camera array, which is implemented purely in software. The rendering speed approaches 5-10 frame per second (fps) on a single Intel Xeon 2.4GHz processor.

A unique property of our camera array is its self-reconfigurability. The cameras are mounted on mobile platforms composed of servos and controllers, as shown in Figure 1(b) and (c). We show that by allowing the cameras to move around, the rendering quality can be improved. This work is an extension and real-world system demonstration of our *view-dependent* non-uniform sampling theory recently proposed in [Zhang and Chen 2004].

2 Rendering

We propose to reconstruct the geometry of the scene as a 2D multi-resolution mesh (MRM) with depths on its vertices. The 2D mesh is positioned on the imaging plane of the virtual view, thus the geometry is view-dependent. The algorithm begins by constructing an initial sparse and regular 2D mesh on the imaging plane of the virtual view. For each vertex of the initial 2D mesh, we obtain the depth of the corresponding light ray by a plane sweeping algorithm, similar to that in [Yang et al. 2002b]. For each light ray, we loop through the hypothesis depth planes and calculate its color consistency score using normalized inner product. The depth plane that results in the best consistency score (the lower the better) are selected as the depth of the light ray. If the depths of the vertices of a certain triangle in the mesh bear large variation, subdivision is performed to obtain more detailed depth information. After the depth reconstruction, the novel view can be synthesized through multi-texture blending.

The above MRM solution saves about 95% of the computational cost compared with reconstructing a per-pixel depth map (which was done in [Yang et al. 2002b]) for a normal scene, and can be implemented in software only. Moreover, more complex algorithms can be adopted during the plane sweeping algorithm, such as using the more robust normalized inter product or adaptive window, etc.

3 Self-Reconfiguration

We make use of the color consistency score obtained during the plane sweeping algorithm for the self-reconfiguration of the cameras. We first locate the camera plane and the linear guides in the world coordinate. The vertices of the mesh model are then back-projected to the camera plane. For each pair of neighboring images on the linear guides, the consistency scores are calculated. If the score is high, it means either the baseline between the two views is too wide, or parts of the image are obscured in one view compared with the other. By moving the cameras closer, we should be able to improve the consistency thus also improve the rendering quality. The cameras are calibrated in real time during the motion. Such reconfiguration improves the rendering quality, as shown in the companion video.

References

WILBURN, B., SMULSKI, M., LEE, H.-H. K., AND HOROWITZ, M. 2002. The light field video camera. In *Proceedings of Media Processors 2002*.

YANG, J. C., EVERETT, M., BUEHLER, C., AND MCMILLAN, L. 2002. A real-time distributed light field camera. In *Eurographics Workshop on Rendering 2002*, 1–10.

YANG, R., WELCH, G., AND BISHOP, G. 2002. Real-time consensus-based scene reconstruction using commodity graphics hardware. In *Proc. of Pacific Graphics 2002*.

ZHANG, C., AND CHEN, T. 2004. View-dependent non-uniform sampling for image-based rendering. In *IEEE Int. Conf. Image Processing (ICIP) 2004*.

---

*e-mail: czhang@ieee.org
†e-mail: tsuhan@cmu.edu