A Multi-View Camera-Based Diminished Reality for Work Area Visualization

Momoko Maezawa, Shohei Mori, and Hideo Saito
Keio University

Introduction and Objectives

In handcrafting, the workers sometimes cannot see the work area from where they want because of occlusions caused by holding tools and the hands. This situation reveals a problem that workers are forced to move their head or the obstructions and mentally reconstruct their desired views. To deal with this issue, we propose a diminished reality (DR) method that provides a view in which the occluding objects look transparent. We use multiple cameras partially capturing the work area from different viewpoints and synthesize the background images in the worker's perspective in real time.

Approach

We take a similar approach to synthetic aperture photography (SAP) [1, 2]. SAP makes objects invisible using regularly arranged cameras by shallow refocusing with a simulated wide aperture or by masking the objects at every camera. The proposed method is a light field rendering, in which light rays detour points in the light fields that we refer to penalty points. Therefore, putting a penalty point on a user's hand or tools avoid light rays to be reconstructed in a synthesized view.

To achieve this, we re-designed camera blending fields (CBF) of Buehler *et al.*'s method [3], which is a generalized form of free-viewpoint image generation methods using unstructured cameras. A CBF is a map of blending weights of M background observers (i.e., data cameras in [3]), $D_m(m = 1,2, \dots, M)$, in a virtual view C. Our DR-specialized CBF is named detour light field rendering (DLFR) [4]. We use an RGB-D camera and multiple conventional cameras to determine where to focus at each reconstructed point and capture hidden light rays respectively (**Fig. 1**).

Implementation and Consideration

We present experimental results to show that the proposed system can visually removes an undesirable tool in real time. We used a Microsoft Kinect sensor and three USB cameras (640×480 resolution) facing work areas (**Fig. 2**). The proposed system worked on a Windows 10 64-bit laptop with an Intel Core i7-6567U 3.30 GHz CPU, Intel Iris Graphics 550 GPU, and 16.0 GB memory. The CBF calculation [4] and view synthesis based on the CBF is implemented on CPU and GPU respectively.

Fig. 3 show examples of DR results. In the results, the tool is automatically detected and visually removed in the view. Average frame rates of the proposed method were about 40.1 fps and thus the proposed method works at above video rates. These results demonstrate that our system visually removes a tool in real time.

Reference

[1] V. Vaish, *et al.*, “Using plane + parallax for calibrating dense camera arrays,” Proc. CVPR, pp. 2 - 9, 2004.
[2] M. Levoy, “Light fields and computational imaging,” *Computer*, 39(8): 46 - 55, 2006.
[3] C. Buehler, *et al.*, “Unstructured lumigraph rendering,” Proc. SIGGRAPH, pp. 425 - 432, 2001.
[4] S. Mori, *et al.*, “Detour light field rendering for diminished reality using unstructured multiple views,” Proc. Int. Workshop on Diminished Reality, pp. 292 - 293, 2016.

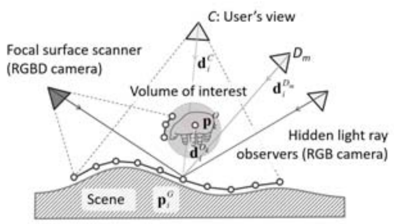

Fig. 1 Schematic illustration of our system

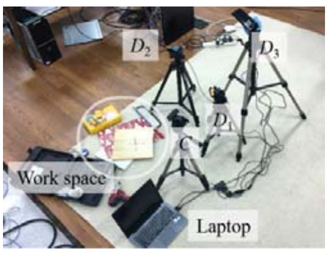

Fig. 2 Setup example

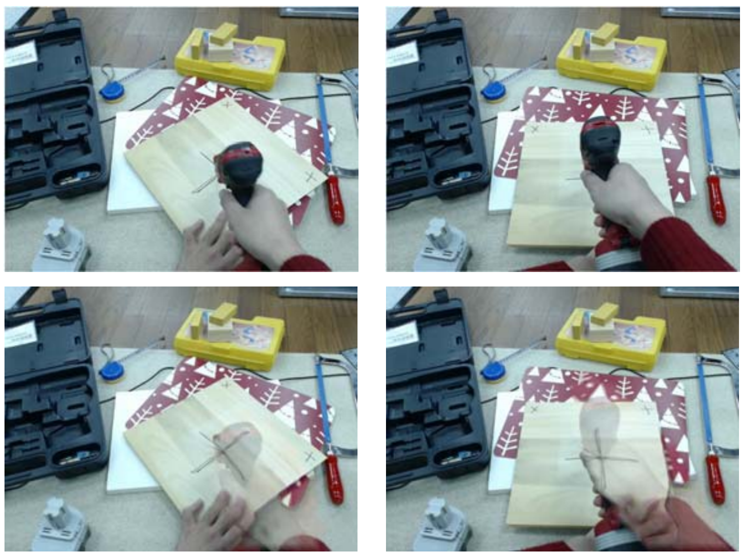

Fig. 3 Input frames (top row) and DR results (bottom row)