Interactive Stereoscopic Authoring in MR-Based Pre-Visualization for 3D Filmmaking

Shohei Mori, Fumihisa Shibata, Asako Kimura, and Hideyuki Tamura

Graduate School of Information Science and Engineering, Ritsumeikan University
1-1-1 Noji-Higashi, Kusatsu, Shiga, 525-8577, Japan

mori@rm.is.ritsumei.ac.jp

Abstract. This paper presents a mixed reality (MR) based system for on-set pre-visualization (PreViz) of stereoscopic 3D (S3D) films. PreViz refers to simple computer-generated shots created in the preproduction phase (before shooting). In terms of stereography, this preparation enables stereographers to design S3D effects on the screen and camera-work using rather cheap computer graphics (CG). However, how a scene actually looks like in S3D are unknown until the production phase on-site. Therefore, we propose a system, which enables users to visualize expected results in the preproduction phase instead of the usual production phase.

Keywords: Mixed Reality, Stereoscopic 3D, Filmmaking, Camera Registration

1 Introduction

Pre-Visualization (PreViz) occupies one of the most important positions in the film industry as the process becomes more complicated. PreViz refers to the rough CG-animated movie created in the early stage of filmmaking to visualize scenes before the actual filming. Such planning is regarded to be important especially for live action shooting in stereoscopic 3D (S3D). However, the actual looks of the scenes are still unknown until the production phase on-site.

To solve this problem, we proposed a mixed reality (MR) based PreViz system named S3D MR-PreViz [1], which is based on its former self, MR-PreViz [2]. The system superimposes CG-animated characters onto scenes from a cinematographic stereo camera (stereo rig (**Fig. 1**)) and displays it in S3D (**Fig. 2**). One of the important requirements of this system is to accomplish camera registration as quickly as possible because venue and labor cost are incurred at all time during preparation such as 3D reconstruction and feature description.

On the assumption that camerawork is roughly determined before shooting, we developed a feature point-based camera tracking method named rehearsal path method to improve speed of offline setup and robustness of realtime camera tracking. Furthermore, in this paper, we propose to integrate the setup into the realtime camera tracking using some additional assumption of stereo rig and S3D cinematography for improving system interactivity.

C. Stephanidis (Ed.): Posters, Part II, HCII 2013, CCIS 374, pp. 718–722, 2013.

Fig. 1. Stereo rigs (Top: Beamsplitter rig. Bottom: Parallel rig)

Fig. 2. Image of on-set S3D MR-PreViz. Left: CG-animated characters are superimposed in realtime and it enables stereographers to confirm the results in the early stage. Right: Results of disparity adjustment in S3D MR-PreViz system.

2 Related work

Koppal *et al.* [3] proposed a view-centric editor for disparity adjustment in S3D shooting. It enables a user to change S3D effects virtually (increase or decrease depth of the scene by changing the disparity of a scene) based on their mathematical framework. While this tool provides users a way to edit captured stereo image sequences on a PC with GUI, the aim of our study is to develop a PreViz system.

Regarding stereo camera registration, we propose an approach for realtime 3D matchmoving without any prior knowledge. Existing approaches suffered from computational speed mainly due to trying to widen the range of applicability and using computationally expensive image feature descriptors for improving robustness [4, 5]. Our approach to the problem is to limit the range of application, to simplify the feature description based on the setup of the stereo rig used for filming, and to actively use multi-core processing. We also assume that a camerawork is roughly determined in advance and our camera passes close to the path repeatedly for improving the robustness.

3 S3D MR-PreViz system

3.1 System configuration

The proposed system is composed of a stereo rig, 3D displays, and a workstation (**Fig. 3**). We assume that the stereo rig is calibrated in advance using [1]. First, two 1280x720 (720p/30) YUV video streams are sent to the workstation from two cameras (Sony PMW-EX3) in the stereo rig (Nac 3D rig II). The input video stream is captured by Blackmagic DeckLink HD Extreme 3D+, a video capture card installed in the workstation. The proposed system uses HD SDI as input interface. Second, input video streams are converted into a grayscale and RGB stereo images using GPGPU [1]. The former is used for vision-based camera tracking and the latter is used for the final result (MR composited image). After camera pose estimation, which is described in the later section, CGI is superimposed onto the left and right images. Finally, the two images are converted into one of the S3D video format using GPGPU and the result is displayed on a 3D display (JVC GD-463D10).

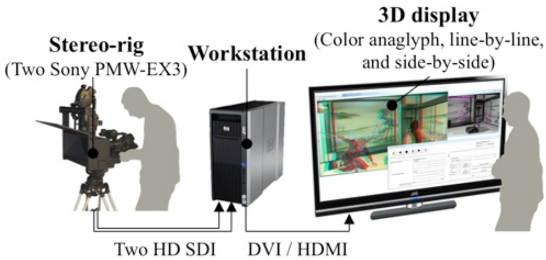

Fig. 3. System configuration

3.2 Procedures of S3D MR-PreViz shooting

S3D MR-PreViz shooting is done as follows.

(1) Setting up the stereo rig: Change convergence and baseline length between cameras in the rig. Assume that the baseline length is determined by the $1/30^{th}$ rule [6].
(2) Calibrate the camera using the camera calibration system in [1]
(3) Camerawork examination: Perform realtime 3D matchmoving by the combination of range data from the stereo rig and iterative closest point (ICP) algorithm [7].
(4) Repeat (1) – (3) for disparity examination by trial and error

3.3 Stereo camera tracking

In the proposed system, the stereo rig is used as a capture device and a vision-based range sensor. 3D positions of feature points are given to the system and the stereo camera pose is estimated using ICP algorithm. Given stereo correspondences, range data is calculated by a well-known formula,

$$Z = \frac{fB}{d} \tag{1}$$

where, Z is the depth, f is the focal length, B is the baseline length, and d is the disparity given by block matching between left and right images. To accelerate this process, we propose to limit its search range ($d_{\min}$, $d_{\max}$) by stereo uncertainty and native pixel parallax (NPP) respectively, which is the maximum disparity in S3D to avoid generating stereo images harmful to the eyes (Eq. (2)).

$$(d_{\min}, d_{\max}) = \left(\sqrt{\frac{\partial d f B}{2 \partial Z}}, \left(\frac{W_{interocular}}{W_{screen}} \right) W_{px} \right) \tag{2}$$

Here, $W_{\text{interocular}}$ is the interocular distance (= 2.5 [inches]), W_{screen} is the screen size in inches, and W_{px} is the width of input image in pixel. In the proposed system, W_{screen} is set as default to 41.06 [inches] because S3D MR-PreViz utilizes JVC GD-463D10 (46 inches 3D display).

We calculate camera pose from the first frame to the current frame. First, Lucas-Kanade (LK) tracker [8] continuously tracks feature points tracked in the previous frame. Second, newly observed points are matched using template matching instead of computationally expensive feature point descriptors. These results are given to ICP algorithm to calculate current camera pose.

Position of a feature point is refined by the following procedure. If the sum of the number of tracked feature points and that of newly matched feature points is small, range data in the current frame is added to the map. Range data is also used for refining the stored range data on the assumption that the same feature points are observed repeatedly while running S3D MR-PreViz because the rehearsal path is roughly determined in advance (Eq. (3)).

$$\mathbf{p}_i' = \frac{n_{ref}\mathbf{p}_i^{\text{map}} + \mathbf{p}_i}{n_{ref} + 1} \tag{3}$$

Here, n_{ref} is the number of reference, $\mathbf{p}_i$ is the range data in the current frame, and $\mathbf{p}_i^{\text{map}}$ is the range data in the stored range data. Because the above processes have to be repeated for each point, they are assigned to multiple threads for parallel processing. Multi-core processing is implemented using OpenMP. **Fig. 4** shows the results of tracking and 3D reconstruction.

Fig. 4. Examples of tracking result. Left: Map reprojection. Three maps are reconstructed and each of them is colored just for visibility. Right: Bird's eye view of the 3D reconstruction and camera trajectory.

4 Results

Fig. 5 shows the results of the proposed system. An S3D MR-PreViz was shot interactively using the stereo rig. The stereo rig moved 900 [mm] from right to left and panned on a rail facing at the robot. Distance between the rig and the closest miniature house to the rig was 1,800 [mm]. Baseline length was set to 50 [mm] by 1/30th rule and stereo search range was limited to 7 – 40 [px] according to the baseline length. Number of images was 450 frames and image resolution was 720x405 [px]. Tracking worked in 14.9 [FPS] when average number of matched points was 122.4 points. In the proposed system, S3D MR-PreViz shooting and disparity examination are repeated quickly because there is no preparation except stereo camera calibration.

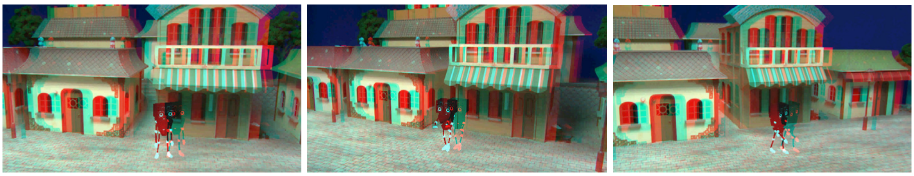

Fig. 5. Results of S3D MR-PreViz using the proposed system (Red-cyan color anaglyph)

5 Conclusion and future work

In this paper we presented a new MR-based PreViz shooting system for S3D films, which includes real-time 6-DOF stereo camera tracking using ICP algorithm. Our previous work required offline preparation but the process is accomplished in real-time using stereo vision in the proposed method. MR composite video was displayed in S3D properly. However, the proposed system suffered from tracking in a full-scale filming set due to short baseline setting for shooting scenes in S3D and convergence angle is not considered for simplifying stereo matching. Therefore, future work will include accuracy improvement of the tracking method and implementation of convergence examination.

References

1. S. Mori, *et al.*: "Enabling on-set stereoscopic MR-based previsualization for 3D filmmaking," SIGGRAPH ASIA 2011, Technical Sketch, 2011.
2. R. Ichikari, *et al.*: "Mixed reality pre-visualization for filmmaking: On-set camera-work authoring and action rehearsal," The Int. J. Virtual Reality, Vol. 7, No. 4, pp. 25 – 32, 2008.
3. Koppal, *et al.*: "A viewer-centric editor for 3D movies," *IEEE Trans. on Computer Graphics and Applications*, Vol. 31, No. 1, pp. 20 – 35, 2011.
4. S. Se, *et al.*: "Local and global localization for mobile robots using visual landmarks," Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, pp.414 – 420, 2001.
5. P. Elinas, *et al.*: "σSLAM: Stereo vision SLAM using the Rao-Blackwellised particle filter and a novel mixture proposal distribution," Proc. IEEE Int. Conf. on Robotics and Automation, pp.1564 – 1570, 2006.
6. B. Mendiburu: "3D movie making: Stereoscopic digital cinema from script to screen," Focal press, 2009.
7. P. Besl and N. McKay: "A method for registration of 3-D shapes," *IEEE Tran.s on Pattern Analysis and Machine Intelligence*, Vol. 14, No. 2, pp. 239 - 256, 1992.
8. B. Lucas and T. Kanade: "An iterative image registration technique with applications in stereo vision," Proc. DARPA Image Understanding Workshop, pp. 121 - 130, 1981.