On-Set Depth Capturing for VFX Productions using Time of Flight

Simon Spielmann
Filmakademie Baden-Wuerttemberg
simon.spielmann@filmakademie.de

Volker Helzle
Filmakademie Baden-Wuerttemberg
volker.helzle@filmakademie.de

Rahul Nair
HCI Heidelberg University
rahul.nair@iwr.uni-heidelberg.de

1. Introduction

Pixel accurate depth estimation of a dynamic scene, in addition to conventional recording on a film set, opens up new opportunities in modern filmmaking. Depth maps can be used for Keying, Set Reconstruction as well as for subsequent Depth Grading in Post Production [1].

Traditional methods using stereo image pairs to determine depth often provide inadequate results [2]. Structured backgrounds, object edges and fine details are the biggest issues for procedures that operate on stereo image pairs [2-4]. These approaches need at least a minimum of disparity to keep the calculation error at a reasonable level. The lack of information in homogeneous or regular structured images will lead to errors within the calculated depth map.

These problems particularly arise in film and media productions. Due to the increasing usage of blue and green screens, the amount of usable information in the raw picture material is more and more reduced. Moreover, on account of the small stereo basis used in movie productions [5] and the artistic use of motion blur, the needed disparity is often lower than the available resolution.

Within this talk, we present a use case of a time of flight (TOF) assisted method [3], which decisively improves the quality of the estimated depth maps. This method has been developed in the joint research project on exploration of *innovative software solutions for stereoscopic content creation*[1]. We also show the differences between controlled- and real live conditions on a film set, and suggest workflows to combine hardware and algorithms to optimize the quality of depth estimation.

2. On set depth capturing

The referenced approach has been tested under real conditions in a film production[2]. In addition to the stereo cameras, a TOF camera system[3] was mounted to a stereo mirror rig[4]. In comparison with the two RGB cameras[5] and the stereo rig, the utilized TOF system is small and lightweight. Therefore, the additional effort to integrate the camera into the rig is manageable. Finally, the camera was mounted on a mechanical linkage under the compendium of the mirror box.

The recording of the TOF data was achieved by a wired external PC. By using the already existing cable links from the RGB camera to the video village, the amount of work and time is minimal. Another solution could be the use of radio transmitters [6], which would allow wireless recording.

3. Stereo Conversion

Classical stereoscopic production have some drawbacks, often related to the need of a second camera and the bulky, additional equipment. Longer times are required for the technical preparation, since the two cameras and their lenses must be precisely calibrated to each other. Also, an additional processing of the depth perception in Post Production is limited as well as time and cost-intensive. The only practical applicable possibilities are manual- or algorithmic approaches [2-4].

In addition to the quality and resolution enhancement, the novelty of our work is the estimation of a high resolution depth map out of a low-res TOF depth map and a single high-res RGB image in a real live production environment.

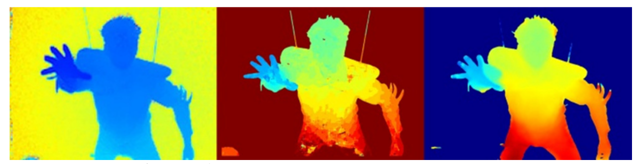

Figure 1. *Take Me Out to the Ballgame.*

3. Results

Figure 1 shows a comparison between TOF, a stereo matching- [2] and the introduced technique [3].With our approach typical artifacts like reconstruction errors in fine details, edge bleeding and noise are nearly eliminated. The resulting depth maps have the same resolution as the production plates and are less noisy than pure TOF images.

During production some system related restrictions occurred. Because of the functional principle of a TOF system, through phase correlation, the maximal measurable distance is limited by the modulation frequency. In our camera this is around 7-8 meters. Furthermore synchronization between the RGB and the TOF camera is necessary. This can be achieved by using dedicated hardware like a Genlock, or in Post, by warping and interpolating the footage.

4. Acknowledgements

This project is joint research with the Heidelberg Collaboratory for Image Processing (HCI). We thank the ministry of economy Baden-Württemberg and our partner companies Unexpected, Pixomondo, ScreenPlane, Bewegte Bilder and Tridelity for its funding under grant number 2-4225.16/380.

References

[1]PATEL, M. 2009. The New Art of Virtual Moviemaking. Audodesk Whitepaper.

[2]SCHARSTEIN A., SZELISKI D. 2002. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. IJCV 47

[3]NAIR, R. 2012. High Accuracy TOF and Stereo Sensor Fusion at Interactive Rates. *LNCS Vol. 7584*.

[4]PARK, J., KIM, H., TAI, Y., BROWN, S., KWEON, I. 2011. High Quality Depth Map Upsampling for 3D-TOF Cameras. *ICCV*

[5]Mendiburu, B. 2009. 3D Movie Making: Stereoscopic Digital Cinema from Script to Screen. *Focal Press* N.26

[6]ECMA INTERNATIONAL. 2008. High Rate Ultra Wideband PHY and MAC Standard. *Standard ECMA-368*, ISO/IEC 26907.

---

1 http://research.animationsinstitut.de/161.0.html

2 How To get a Girl in 60 seconds, Johannes Peter, Filmakademie 2013

3 pmd[vision]® CamCube 3.0

4 Screen Plane® Kite-Rig

5 Arri® ALEXA

SIGGRAPH 2013, July 21 – 25, 2013, Anaheim, California.

ACM 978-1-4503-2261-4/13/07