ThirdEye: A Coaxial Feature Tracking System for Stereoscopic Video See-Through Augmented Reality

Yu-Xiang Wang Yu-Ju Tsai Yu-Hsuan Huang Wan-Ling Yang Tzu-Chieh Yu Yu-Kai Chiu Ming Ouhyoung

National Taiwan University

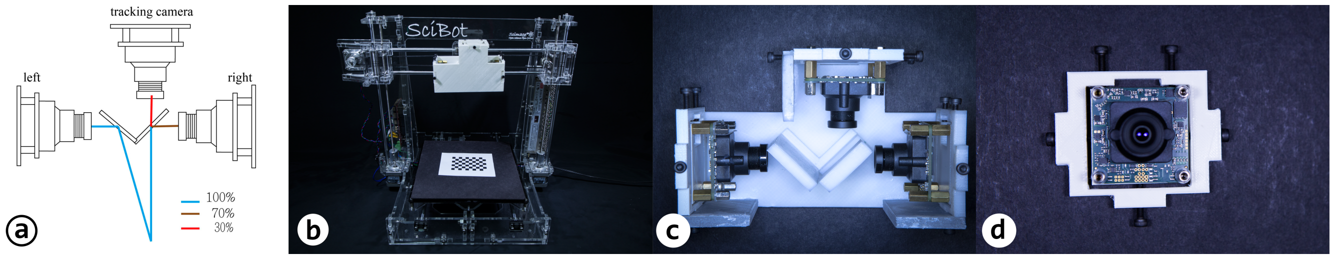

Figure 1: *(a) Optical path of the beam-splitter (b) Testing platform with camera calibration setting (c) Sectional view of ThirdEye (d) the fine-tuning screw structure in camera holding frame*

Concepts: •**Hardware** → **Emerging architectures;** •**Computing methodologies** → *Mixed / augmented reality;*

Keywords: augmented reality, beam-splitter, hardware design, stereo vision

Introduction

For stereoscopic augmented reality (AR) system, continuous feature tracking of the observing target is required to generate a virtual object in the real world coordinate. Besides, dual cameras have to be placed with proper distance to obtain correct stereo images for video see-through applications. Both higher resolution and frame rate per second (FPS) can improve the user experience. However, feature tracking could be the bottleneck with high resolution images and the latency would increase if image processing was done before tracking.

There are several solutions to conquer this conflict and also preserve stereo vision. In Scope+ [Huang et al. 2015], a video see-through AR microscope system, which chooses the resolution between the upper limit of the camera and the minimal requirement to enable tracking function, utilize the fusion ability of human to compromise the lower resolution images by another eye. It worked very well for general usage and gaming purposes. However, it would not be acceptable to the experts if there were a huge difference between the images for each eye, since the professional surgeons and researchers are very sensitive and deeply rely on good image quality.

Using an individual camera for tracking alone is another common solution with AR hardware design [Oda et al. 2015]. The camera is therefore off the visual axis, and the images would be different from those on the display panel. The offset between visual and tracking images would change based on the object distance, which becomes significant in near field below 1 meter.

In our approach, a coaxial camera is added for tracking through a splitter, which can provide a high resolution image for users to observe and also a reliable tracking result.

SIGGRAPH '16, July 24-28, 2016, Anaheim, CA,
ISBN: 978-1-4503-4371-8/16/07
DOI: http://dx.doi.org/10.1145/2945078.2945100

Experiment

Three identical cameras are fixed in the special case as shown in Figure 1(c). The resolution of the camera is 1600×1200 pixels per eye and is 800×600 pixels for the tracking camera.

Through the beam-splitter, the two cameras on the right side could obtain identical images at the same time ideally. In practical, we designed a special holding frame with fine-tuning screws for cameras (Figure 1d) to minimize the structure error, and also the camera calibration with chessboard was performed (Figure 1b) to achieve ideal results.

After calibration, three kinds of beam-splitters, including 3:7, 5:5, and 7:3 (Transmittance:Reflectance) were tested. With 3:7 beam-splitter, tracking successful rate was 92% within 0.5 second(46/50), which was acceptable in general applications but worse than 5:5(94%) and 7:3(98%). On the other hand, the images from 70% reflection could not be distinguished from those of left camera after properly adjusting the gain and shutter value of right camera. The results showed that 3:7 beam-splitter was the best choice to provide high resolution images with good visual quality and acceptable tracking successful rate.

To compare the performance of ThirdEye with previous solution, a simple AR test was done by overlaying a virtual cube onto a circuit board using Unity game engine and Qualcomm's Vuforia AR toolkit. The testing platform was equipped with i7-4790 CPU, 16GB RAM and GTX980 graphic card. Frame rate was recorded as the benchmark.

In solutions with 2 cameras, the frame rate was only 13 FPS when both cameras were set on high resolution (1600×1200). This result could be improved to 45 FPS while low resolution (800×600) was adopted on one of the cameras. With ThirdEye, the FPS is as good as the low resolution solution (45 FPS) and also has a better visual quality with acceptable tracking ability.

Conclusion

Both high resolution images and high frame rate are crucial in AR applications to provide a better user experience. We proposed a coaxial feature tracking module with three cameras which provide high resolution stereo images smoothly and perform object tracking quickly with acceptable successful rate at the same time. Also, ThirdEye could be integrated into other video see-through hardware such as head mounted displays or other wearable devices. That is, this module has many possible variants and applications in the future.

Acknowledgements

This work was partially supported by MOST 104-2622-8-002-002 (MediaTek Inc.), MOST 103-2218-E-002-025-MY3, MOST 103-2221-E-002-171-MY2, and MOST 104-2218-E-002-015.

References

HUANG, Y.-H., YU, T.-C., TSAI, P.-H., WANG, Y.-X., YANG, W.-L., AND OUHYOUNG, M. 2015. Scope+: A stereoscopic video see-through augmented reality microscope. In *SIGGRAPH Asia 2015 Emerging Technologies*, ACM, New York, NY, USA, SA '15, 21:1–21:3.

ODA, O., ELVEZIO, C., SUKAN, M., FEINER, S., AND TVERSKY, B. 2015. Virtual replicas for remote assistance in virtual and augmented reality. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology*, ACM, New York, NY, USA, UIST '15, 405–415.