Towards Eye-Perspective Rendering for Optical See-Through Head-Mounted Displays

Gerlinde Emsenhuber*
Salzburg University of Applied Sciences

Michael Domhardt
Salzburg University of Applied Sciences

Tobias Langlotz
University of Otago

Denis Kalkofen
Graz University of Technology

Markus Tatzgern
Salzburg University of Applied Sciences

Abstract

The optical see-through (OST) head-mounted display (HMD) is a typical platform for Augmented Reality (AR) and allows users to experience virtual augmentations in a wearable form factor. Utilizing information of the real-world background, visualization algorithms adapt the layout and representation of content to improve legibility. Typically, this background information is captured via built-in HMD cameras. However, HMD camera views of the real-world scene are distinctively different to the user's view through the OST display. In this work, we propose eye-perspective rendering (EPR) as a solution to synthesize high fidelity renderings of the user's view for mobile OST HMD to enable adaptation algorithms to utilize visual information as seen from the perspective of the user to improve placement, rendering and, thus, legibility of content.

1 Introduction

An Optical see-through (OST) head-mounted display (HMD) allows displaying virtual content to provide additional cues for users to solve a task or interact with the environment. The continuous overlay of annotations supports, for instance, maintenance or assembly tasks. OST HMDs achieve that graphical overlay using semitransparent displays placed in the users view. While research has proposed numerous improvements to OST HMDs [3], current applications still suffer from perceptual issues where the poor placement of content can lead to a lack of contrast between the augmentations and the real-world background (Fig. 1(Left)). Adaptive visualization techniques can improve the contrast by adjusting the placement [2, 6, 9] and appearance of content [10]. However, for both to work, algorithms require exact knowledge of the scene as seen by the user through the OST HMD.

Most approaches overcome this issue by using views captured by a camera integrated into the HMD itself. However, this camera view does not match the user's view (Fig. 1(Left)), resulting in poor placement due to incorrect assumptions made from the camera view of the HMD. Research overcame this issue by using software simulations instead of actual camera images or replaced the user's eyes with cameras allowing an application to capture the real-world scene. However, these approaches prevent a practical application where users look through the OST display. Recently, Langlotz et al. [4, 5] have presented a hardware-based solution that captures the user's view by introducing beam splitters that forward the view of the physical world to an eye camera. However, their approach requires extensive hardware modifications, is optimized for a specific eye position, and captures only a small field of view to limit the additional size and weight.

Instead of hardware-based solutions, we explore software approaches that approximate the user's view as seen through the OST HMD similarly to user-perspective rendering user-perspective rendering (UPR) for handheld Augmented Reality (AR) [7]. While UPR compensates for the viewpoint mismatch between the user's view of the real world and the handheld device, in this work focusing on eye-perspective rendering (EPR), we synthesize the user's view through the HMD to provide AR applications with information about the real-world background so that algorithms can adapt content to the respective background (Fig. 1(Right)). In this work, we demonstrate strategies for adaptive visualization for HMDs based the reliability of EPR information.

*e-mail: gerlinde.emsenhuber@fh-salzburg.ac.at

Figure 1: Eye-perspective rendering provides AR applications with information about the user's view of the scene by synthesizing their view through an optical see-through display. (Left) Using input from a built-in head-mounted display camera that films the real-world scene (small inset image), a view management algorithm places labels suitable for the respective background areas. Due to the viewpoint offset between camera and the user's eyes, the user actually perceives the label in a different location causing a lack of contrast. (Right) Eye-perspective rendering synthesizes the user's view through the display (small inset view) that is then used by view management algorithms to adapt labels in order to avoid visual interference with the background.

2 Analysis and Strategies

We identified two general approaches to synthesize the user's view through the display using EPR: a homography of the currently captured view [11], and using 3D proxy geometry and visual information of one [8] or more [1] captured views, e.g., based on RGBD camera input. We experimented with three EPR approaches: (1) a homography of the video based on an analysis of the scene that identifies the dominant plane, (2) view synthesis using proxy geometry and color information of a depth camera and reprojecting the information into the user's view, and (3) view synthesis using the same proxy and image-based rendering (IBR) for view synthesis.

Target Scene Complexity and Reconstruction Quality. Homography-based methods provide reliable information only for mostly planar scenes for which an accurate plane can be detected. Naturally, any 3D geometry is distorted leading to mismatches compared to the ground truth eye view. A clear advantage of homography-based methods is that they can be realized with limited hardware resources as depth cameras are not necessarily required. While the planarity requirement is a clear limitation of the technique, homographies are feasible for use cases where content is applied to

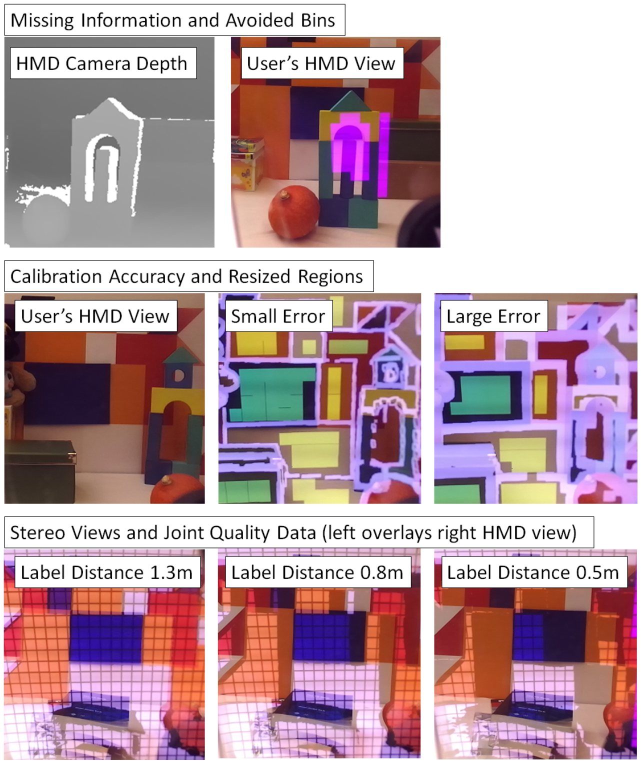

Figure 2: Adaptation algorithms must consider the reliability of EPR information. Reliability is influenced by (Top) reconstruction artifacts leading to missing information (purple) that must be avoided. (Middle) Calibration errors lead to an offset between the user's and the EPR view. Adaptation algorithms can classify information at region borders as unreliable using edge dilation based on estimated errors, thereby reducing the number of reliable regions. (Bottom) Adaptation algorithms must consider stereo vision of users to find homogeneous background regions for both eyes, thereby limiting the number of reliable regions.

planar surfaces such as walls, e.g., when placing virtual windows, or when scenes are distant from users.

The reprojection method using only RGBD information generally provides reliable information for the synthesized EPR views of 3D scenes. Missing information after reprojecting data into the user's view can be clearly identified and handled by any adaptation algorithm. Alternatively, an IBR-based inpainting strategy can be used to fill in missing information.

Reliable Image Information. Due to missing information in the EPR view, areas are marked in purple (Fig. 2(Top)). Adaptive visualization algorithms can avoid utilizing this incomplete information by defining thresholds that allow for a certain amount of missing information when analyzing EPR images.

Calibration Accuracy. The EPR view relies on estimating the user's eye locations utilizing calibration methods such as the single-point active alignment method (SPAAM) or eye trackers. These methods typically have issues with accurately estimating eye position and gaze. Such inaccuracies influence the reliability of the EPR information. An accurate SPAAM calibration leads to almost perfect registration in the HMD view. However, a larger error leads to less reliable information due to the offset between EPR view from the user's eye perspective. This visible error decreases with the distance. Hence, adaptation algorithms need to take into account potential calibration errors in their analysis of the EPR view.

An adaptive visualization algorithm can utilize information regarding calibration inaccuracies, e.g., the reprojection error of calibration methods, to adjust the reliability of EPR information. We implemented a strategy that dilates edges in the EPR image using a filter size based on the calibration error and scene distance from the user (Fig. 2 (Middle)). These dilated edges define unreliable information and restrict the reliable image regions.

Stereo View Differences. The issue of user's stereo vision becomes apparent when exploring EPR for adaptive visualization. Relying on the HMD camera, adaptation algorithms utilize only information from a single view. However, using EPR, algorithms can estimate the quality of the real-world background for each eye and adjust virtual content so that it lies within the same homogeneous color region. Advanced methods may adapt the representation of content separately for each eye to jointly enhance legibility (Fig. 2 (Bottom)). This avoids potential perceptual conflicts, where, e.g., text is fully legible in one eye, but partly overlaps with high frequency backgrounds in the other eye.

3 Conclusion

In this work, we investigate the often overlooked issues that arise from a lack of knowledge of the user's view of the physical scene when using an AR OST HMD. In particular, we demonstrate the necessity to utilize EPR for adapting virtual content in AR as image information from an HMD-integrated camera does not match the user's view. We discuss multiple issues of EPR that impact the reliability of its synthesized information in order to derive strategies to adapt content based on a reliability estimation.

Acknowledgments

This work was supported in part by a grant from the Austrian Research Promotion Agency (grant no. 877104). Tobias Langlotz is supported by the Marsden Fund Council from Government funding.

References

[1] D. Baricevic, T. Höllerer, P. Sen, and M. Turk. User-Perspective AR Magic Lens from Gradient-Based IBR and Semi-Dense Stereo. *IEEE Trans. on Visualization and Comp. Graphics*, 23(7):1838–1851, 2017.

[2] R. Grasset, T. Langlotz, D. Kalkofen, M. Tatzgern, and D. Schmalstieg. Image-driven view management for augmented reality browsers. In *IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 177–186, 2012.

[3] Y. Itoh, T. Langlotz, J. Sutton, and A. Plopski. Towards indistinguishable augmented reality: A survey on optical see-through head-mounted displays. *ACM Computing Surveys*, 54(6), July 2021.

[4] T. Langlotz, M. Cook, and H. Regenbrecht. Real-time radiometric compensation for optical see-through head-mounted displays. *IEEE TVCG*, 22(11):2385–2394, 2016.

[5] T. Langlotz, J. Sutton, S. Zollmann, Y. Itoh, and H. Regenbrecht. ChromaGlasses: Computational glasses for compensating colour blindness. In *Conf. on Human Factors in Computing Systems*, pp. 1–12, 2018.

[6] J. B. Madsen, M. Tatzqern, C. B. Madsen, D. Schmalstieg, and D. Kalkofen. Temporal Coherence Strategies for Augmented Reality Labeling. *IEEE Transactions on Visualization and Computer Graphics*, 22(4):1415–1423, 2016. doi: 10.1109/TVCG.2016.2518318

[7] P. Mohr, M. Tatzgern, J. Grubert, D. Schmalstieg, and D. Kalkofen. Adaptive user perspective rendering for Handheld Augmented Reality. In *IEEE Symposium on 3D User Interfaces (3DUI)*, 2017.

[8] T. Schöps, M. R. Oswald, P. Speciale, S. Yang, and M. Pollefeys. Real-Time View Correction for Mobile Devices. *IEEE Transactions on Visualization and Computer Graphics*, 23(11):2455–2462, 2017.

[9] M. Tatzgern, D. Kalkofen, R. Grasset, and D. Schmalstieg. Hedgehog Labeling: View Management Techniques for External Labels in 3D Space. In *IEEE Virtual Reality (VR)*, 2014.

[10] M. Tatzgern, V. Orso, D. Kalkofen, G. Jacucci, L. Gamberini, and D. Schmalstieg. Adaptive information density for augmented reality displays. In *IEEE Virtual Reality (VR)*, pp. 83–92, 2016.

[11] M. Tomioka, S. Ikeda, and K. Sato. Approximated user-perspective rendering in tablet-based augmented reality. *IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pp. 21–28, 2013.