Seamlessly Depth-Enhanced VR Display for 360 Wild Images

Yan-Jen Su
Industrial Technology Research Institute

Chia-Hsin Chien
Industrial Technology Research Institute

Ruen-Rone Lee
Industrial Technology Research Institute

Feng-Sheng Lin
Industrial Technology Research Institute

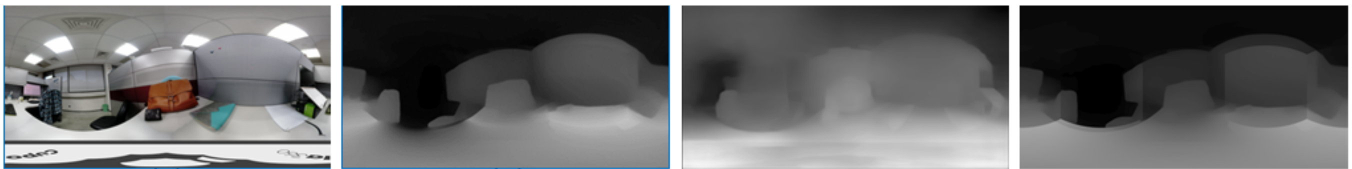

Figure 1: Comparison of depth estimation. (a) Original input; (b) Our result; (c) Direct NN estimation; (d) Depth cubepadding.

ABSTRACT

Due to the limited range of depth sensors, the depth estimators for wild images were trained by ordinal data. Without the need of re-train, we proposed an approach to create seamlessly ordinal depth for 360 wild images. The 360 depth was applied to VR display for improving the virtual perception.

CCS CONCEPTS

• **Computing methodologies** → **Virtual reality**.

KEYWORDS

360 wild image; depth reconstruction; VR

ACM Reference Format:
Yan-Jen Su, Chia-Hsin Chien, Ruen-Rone Lee, and Feng-Sheng Lin. 2020. Seamlessly Depth-Enhanced VR Display for 360 Wild Images. In *Special Interest Group on Computer Graphics and Interactive Techniques Conference Posters (SIGGRAPH '20 Posters), August 17, 2020*. ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3388770.3407413

1 INTRODUCTION

360 cameras become increasingly popular in daily life. It records the whole angle of view at the same time. A general display method is to project the 360 image onto a skydome and set the viewpoint at the center of the dome for display. It could be easily extended to VR interaction for advanced visual perception by using two fixed 360 cameras to approximate the omnidirectional stereo images [Matzen et al. 2017].

The approaches to estimate depth from a single image were commonly classified into indoor and outdoor solutions. The indoor depth data could be easily measured, so [Zioulis et al. 2018] projected the measured 3D points onto the space of 360 images and use the depth information to train a neural network for depth inference.

SIGGRAPH '20 Posters, August 17, 2020, Virtual Event, USA

ACM ISBN 978-1-4503-7973-1/20/08.
https://doi.org/10.1145/3388770.3407413

Besides, a cubepadding method [Wang et al. 2018] was introduced which projected 360 images to cubemaps and trained a neural network for depth estimation by each cube face. In order to preserve the depth coherence, they suggested to pad part of adjacent cube faces for training.

In the cases of outdoor depth estimation, the depth value was unrestricted and most of it could not be measured by current sensors. [Li and Snavely 2018] trained a neural network using the ordinal depth labeled by their automatic approach. [Lasinger et al. 2019] released MiDaS system that improved the network to accommodate the inconsistency between different large datasets.

According to [Didyk et al. 2011], the human visual system uses an interplay of many cues to estimate spatial configurations for understanding a scene. Inspired by the concept of bas-relief that enhances visual perception by related depth, we developed an algorithm to create a seamless 360 depth map from a wild 360 image and applied the ordinal depth to stretch the disparity for VR display.

2 APPROACH

Our approach is best illustrated by an example. Fig. 2 shows the processing steps of our algorithm. First, a 360 image, represented as an equirectangular image in Fig. 2(a), is converted to a cubemap (Fig. 2(b)) to avoid distortion. In order to retain more spatial information, the top and bottom cube faces are duplicated and rotated to match the border of the horizontal cube faces, as indicated by the yellow boxes in Fig. 2(b). Besides, the horizontal four cube faces are continuous and forming a panorama-like image. We then estimate the depth on two image patches centered at green and red triangles, each includes 3x3 cube faces as shown in Fig. 2(c). Linear blending is applied to the overlapping faces and created a horizontally continuous depth map except for the top and bottom faces as indicated by the green and red crosses in Fig. 2(d). In order to relief the depth gaps between horizontal cube faces and the top and bottom cube faces, we stitch the four horizontal cube faces to the respective adjacent boundaries of the top and bottom cube faces. Then, Poisson blending is applied to smooth out the boundaries and results as in Fig. 2(e). Finally, the depth map is derived by reprojecting the cubemap back to a 360 image as shown in Fig. 2(f) and the result is seamless with ordinal depth preserved.

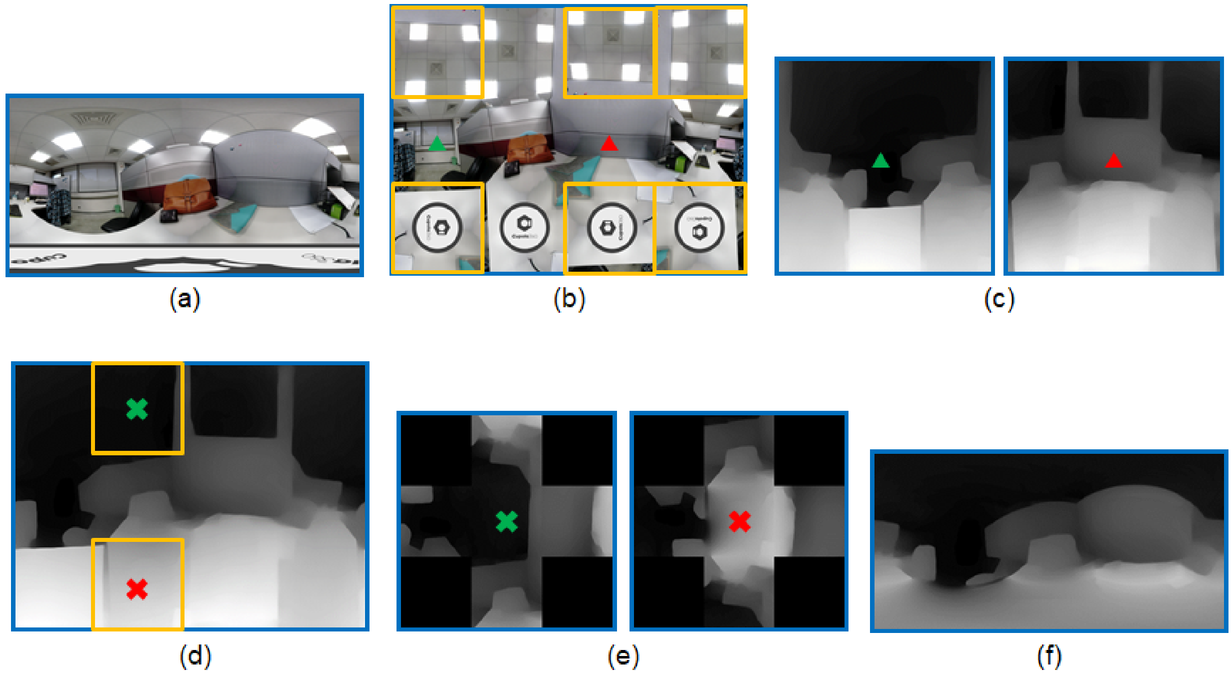

Figure 2: Steps of proposed approach. (a) Input 360 image; (b) Convert the 360 image to a cubemap; (c) Estimated depth maps; (d) Horizontally smooth processing; (e) Top and bottom cube faces smooth processing; (f) Final seamless 360 depth map.

Our approach is independent of depth estimation, so depending on the cases, a suitable algorithm could be applied to. Using MiDaS as the depth estimator, different approaches compare in Fig. 1. Fig. 1(a) is the input 360 image and Fig. 1(b) is the result derived by our approach. Fig. 1(c) is a direct depth estimation on the 360 image where the distortion of the equirectangular image largely increased the estimated error. Fig. 1(d) is the cubepadding method which estimates the depth by padding neighboring images. However, MiDaS was not trained for boundary coherence, gaps between different views are notable. For more results, please refer to the supplementary materials.

For better depth perception in VR display, the 360 depth map is used as a displacement map to displace the skydome alone in the direction to the center of the dome. However, if the viewing point located in the center of the skydome, as Fig. 3(a), the viewing rays were totally the same with the stretching directions such that the disparity was eliminated. Therefore, as 3(b), the viewing point was moved some distance in the opposite direction of view to enhance the visual perception in VR interaction.

3 CONCLUSIONS AND FUTURE WORKS

We proposed a seamless 360 depth image derivation from an input 360 color image. The resulting depth image can help to enhance the depth perception in VR display by displacing the skydome and offsetting the viewpoint out from the center. Our approach can adapt to the ordinal depth which covers both indoor or outdoor 360 images. Despite the depth enhancement is significant, further study is required to assess the effect of visual perception.

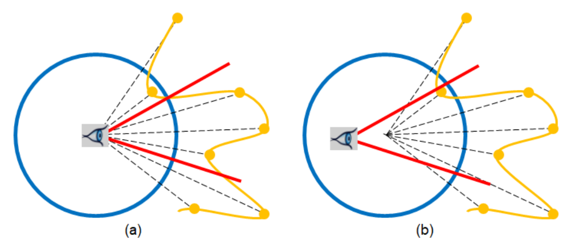

Figure 3: Deviate the viewpoint from the center of the dome for better VR depth perception.

REFERENCES

P. Didyk, T. Ritschel, E. Eisemann, K. Myszkowski, and H.P. Seidel. 2011. A perceptual model for disparity. *ACM Transactions on Graphics (TOG)* 30, 4 (2011), 1–10.

K. Lasinger, R. Ranftl, K. Schindler, and V. Koltun. 2019. Towards robust monocular depth estimation: Mixing datasets for zero-shot cross-dataset transfer. *arXiv preprint arXiv:1907.01341* (2019).

Z. Li and N. Snavely. 2018. Megadepth: Learning single-view depth prediction from internet photos. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2041–2050.

K. Matzen, M. F. Cohen, B. Evans, J. Kopf, and R. Szeliski. 2017. Low-cost 360 stereo photography and video capture. *ACM Transactions on Graphics (TOG)* 36, 4 (2017), 148.

F.E. Wang, H.N. Hu, H.T. Cheng, J.T. Lin, S.T. Yang, M.L. Shih, H.K. Chu, and M. Sun. 2018. Self-Supervised Learning of Depth and Camera Motion from 360 Videos. *arXiv preprint arXiv:1811.05304* (2018).

N. Zioulis, A. Karakottas, D. Zarpalas, and P. Daras. 2018. Omnidepth: Dense depth estimation for indoors spherical panoramas. In *Proceedings of the European Conference on Computer Vision (ECCV)*. 448–465.