Holographic 3D Portrait Created by Full-Parallax High-Definition Computer Holography and 3D Face Reconstruction

Sukai Ikeda* and Kyoji Matsushima
Department of Electrical and Electronic Engineering,
Kansai University, Yamate-cho 3-3-35, Suita, Osaka 564-8680, Japan.
**Corresponding author e-mail: k565888@kansai-u.ac.jp*

Abstract: A technique using the 3D Face Reconstruction technology is proposed to create 3D portrait CGHs from a single 2D portrait image. An actual full-parallax portrait CGH is demonstrated to verify the technique.

1. Introduction

For the last decade, the technology of large-scale computer holography has been developed. Several brilliant full-parallax high-definition computer-generated holograms (FPHD-CGH), such as “Sailing Warship II”, “Toy Train” and “Brothers”, are created using the technique [1]. These are composed of several tens or a hundreds billion pixels; the 3D images are very impressive because of a great sense of depth, and actually comparable to those by classical optical holography. One of the feasible applications of the FPHD-CGHs is producing 3D portraits for ID or memories. However, 3D data of live faces is required to create holographic portraits as a FPHD-CGH. 3D scanners and depth cameras are not suited for this work because there are inevitably unscannable parts such as a part under the chin. Another method to acquire 3D data of live face is to use multi-viewpoints images (MVI). However, capturing MVI is commonly too time-consuming to create a 3D portrait simply.

In this paper, we use 3D Face Reconstruction (3D-FR) technology [2] to obtain the 3D model of a live face from the single 2D portrait image and create the holographic 3D portrait as a FPHD-CGH.

2. 3D Face Reconstruction

As shown in Fig. 1, the 3D-FR consists of three convolutional neural networks (CNN): Reconstruction Network [3], Segmentation Network, and Depth Estimation Network [2]. The Reconstruction Network generates only the 3D model of the facial part of the head from the 2D portrait image, excluding ears and hair regions. We used BiseNet [4] as the Segmentation Network to perform accurate semantic segmentation in this work. Here, the original BiseNet uses Xception [5] for down sampling, but we use ResNet-18 [6] because of its considerable accuracy. The Depth Estimation Network generates the 3D model of the whole head from the three inputs, i.e., the 2D portrait image, the facial 3D model output by the Reconstruction Network, and the segment information output by the Segmentation Network.

3. Creation of FPHD-CGH by 3D Face Reconstruction

Figure 2 (a) shows the input portrait image and the 3D model generated by the 3D-FR. A full-color FPHD-CGH, whose size is 10.5 cm × 10.5 cm, was created from the 3D model using the RGB color filters [7]. The parameters and photograph of the created CGH are shown in Table 1 and Fig. 2 (b), respectively. The closeup photographs of optical reconstruction, taken from various viewpoints, are shown in Fig. 3. It is confirmed that the created FPHD-CGH properly reconstructs the human face with a sense of depth. However, 3D-FR does not generate any parts other than the head, such as the neck and chest. As a result, the reconstructed image of the created FPHD-CGH looks as if a human head is floating in the air, and thus, may cause a sense of fear to some viewers.

4. Conclusion

We proposed a novel technique to create 3D portrait CGHs from a single portrait photograph using the 3D-FR technique, which consists of three CNNs. However, since the 3D-FR does not support generation of the 3D model of the parts of the neck and chest, only the human head is reconstructed by the CGH. Thus, the future work is to create a complete 3D portrait like a photograph of bust shots that include not only the head but also the neck and chest.

References

[1] K. Matsushima, “*Introduction to Computer holography*,” Sect. 1.3 (Springer, 2020).
[2] S. Xu, J. Yang, D. Chen, F. Wen, Y. Deng, Y. Jia, X. Tong, “Deep 3D portrait from a single image,” IEEE/CVF Conference on Computer

Vision and Pattern Recognition, 7707-7717(2020).
[3] Y. Deng, J. Yang, S. Xu, D. Chen, Y. Jia, X. Tong, "Accurate 3D face reconstruction with weakly-supervised learning: From single image to image set," IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, 285-295(2019).
[4] C. Yu, J. Wang, C. Peng, C. Gao, G. Yu, N. Sang, "BiSeNet: Bilateral segmentation network for real-time semantic segmentation," Proc. of the European Conference on Computer Vision, 325-341(2018).
[5] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," IEEE Conference on Computer Vision and Pattern Recognition, 1800-1807(2017).
[6] K. He, X. Zhang, S. Ren, J, Sun, "Deep residual learning for image recognition," IEEE Conference on Computer Vision and Pattern Recognition," 770-778(2016).
[7] Y. Tsuchiyama, and K. Matsushima, "Full-color large-scaled computer-generated holograms using RGB color filters," Opt. Express **25**, 2016-2030 (2017).

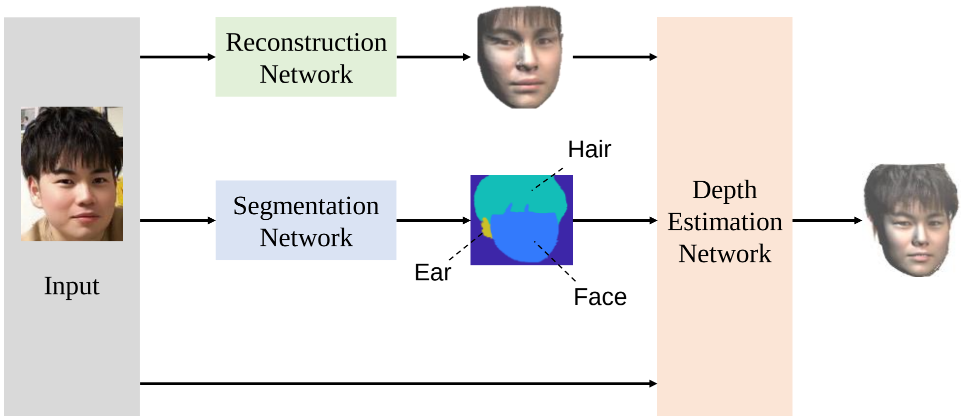
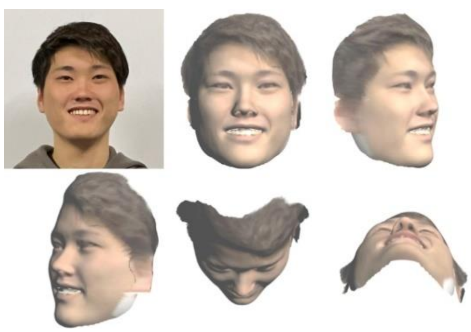


Fig. 1. Flow of the 3D Face Reconstruction.



(a) Input 2D photograph and generated 3D model.

(b) Photograph of the created FPHD-CGH.

Fig. 2. The 3D model generated by the 3D-FR and the created FPHD-CGH.

Table 1 Parameters of the created FPHD-CGH.

Number of pixels	131,072 × 262,144
Pixel pitches [μm]	0.8 × 0.4
Wavelength (R, G, B) [nm]	(635, 517, 443)
Number of polygons	82,834

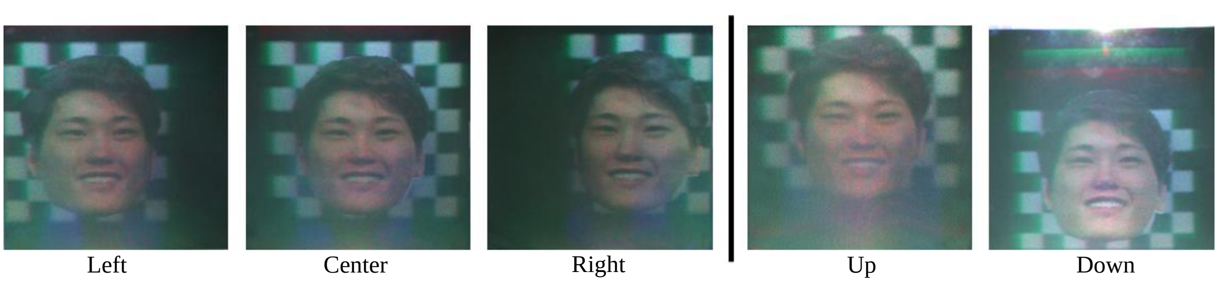


Fig. 3. Closeup photographs of optical reconstruction of the created FPHD-CGH.