PROCEEDINGS OF THE INTERNATIONAL DISPLAY WORKSHOPS, VOL. 27, 2020

Synthesis of Holographic 3D Content based on Deep Learning

MinSung Yoon[1], Hyeonji Kim[1], and HwaSeon Yeo[1]

msyoon@etri.re.kr

[1] Next-Generation Digital Content Research Laboratory, Electronics and Telecommunications Research Institute, 218 Gajeong-ro, Yuseong-gu, Daejeon, 34129, Korea

Keywords: Deep learning, Depth map, CGH, holographic content, SLM.

ABSTRACT

After extracting depth map from colour images using deep learning, we demonstrate that its 3D content with all directions of 360˚ is applicable to holographic display. For obtaining holographic content, CGHs are synthesized through FFT algorithm from sets of colour image/depth map, and are encoded by Lee's scheme for amplitude-modulator.

1 INTRODUCTION

Three-dimensional (3D) displays have recently been a key subject to attract various industrial opportunities supported by mobile device applications such as virtual reality or augmented reality [1,2]. Especially, electro-digital holography is a promising field among 3D displays to visualize the most real-like 3D objects/scenes. The CGH (computer-generated hologram) method has been used for calculation of fringe patterns from the position of the 3D object [2]. The distance from an object can be captured either by a depth map camera or derived from multi-view images [3,4]. Depth map data represent distance information between surface of an object and observation position as the grey-scaled levels. Recent advances in deep learning have given rise to a new paradigm for techniques of depth map estimation as well as holographic reconstruction with real-time performance. The conventional works have featured the estimation of depth information using either single images or stereo images [3,4]. In the paper, we report a novel demonstration of a convolutional neural network (CNN)-based method to extract depth map from multi-view colour images, so that we can generate a holographic 3D content, namely, CGHs with the full 360° views.

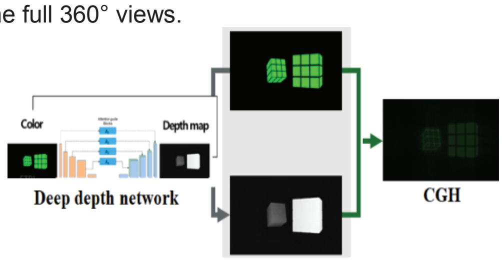

Fig. 1 Pipeline of preparation for the holographic content based deep learning.

In addition, for a test of the proposed model, using the holographic 3D content, we apply it for optical reconstruction through a reflective LCoS-SLM in order to observe holographic 3D scenes in real space as well as we make an observation through its numerical reconstruction.

2 EXPERIMENT AND RESULTS

The pipeline of preparation for the holographic 3D content based on the deep learning approach is shown in Figure 1. The proposed CNN model to estimate the depth map by a training set of multi-view colour images is composed of the encoder and the decoder, which is a modification of the DenseNet architecture [5]. We utilize the loss function between ground truth of depth map and estimated depth map to optimize the weights for each component.

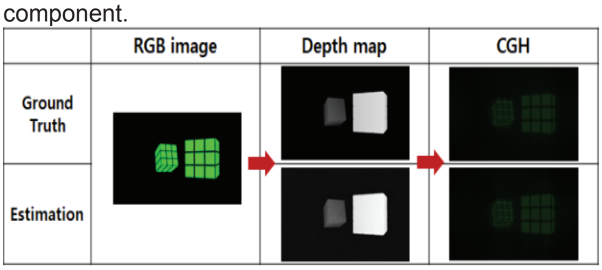

Fig. 2 Schematic of the CGH based on the set of RGB image & depth map estimated from deep learning.

For the preparation of learning data, the Maya™ program uses to obtain RGB-depth map data related with the full 360° range (1,024 views) from a given 3D scene. As shown in Figure 2, the choice scene that consists of two objects is designed to observe the accommodation from holographic reconstruction. Using a set of RGB image & depth map per view, a CGH is synthesized by the FFT (fast Fourier transform) algorithm [2].

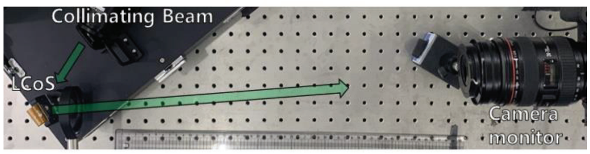

Fig. 3 Configuration of optical setup for the holographic 3D observation that comprises collimating optics, LCoS-SLM, and a field lens.

After CGHs of 1,024 views are prepared, an encoding process is adapted to upload them onto an amplitude-modulating, spatial light modulator (SLM). Here we choose a reflective LCoS-SLM (resolution of 3840 x 2160 pixels, diagonal size of 0.62", and pixel pitch of 3.6 um expressing the grey levels of 8 bits) as the

ISSN-L 1883-2490/27/0503

amplitude-modulating SLM for optical observation. The uniformly-illuminated, coherent beam for the 0.62” size of SLM is prepared after each light combined from a RGB laser sources (638 nm, 520 nm, 450 nm) passes through a collimating & expanding optics (10 times magnification). There is a lens (focal length: 50 cm) located just after LCoS (See Figure 3).

We choose a typical representation scheme of CGH onto amplitude modulation called Lee’s method; The Lee’s scheme decomposes a given complex-value field into four real and non-negative components [6, 7]. The decomposition by Lee’s method can be written as

$$H_L(x,y) = L_1(x,y)e^{i0} + L_2(x,y)e^{\frac{i\pi}{2}} + L_3(x,y)e^{i\pi} + L_4(x,y)e^{\frac{i3\pi}{2}}.$$

Here at least two among four coefficients are equal to zero. The relative phases for Lee’s encoding are physically implemented by the lateral displacement within a macro-pixel that consists of four sub-pixels.

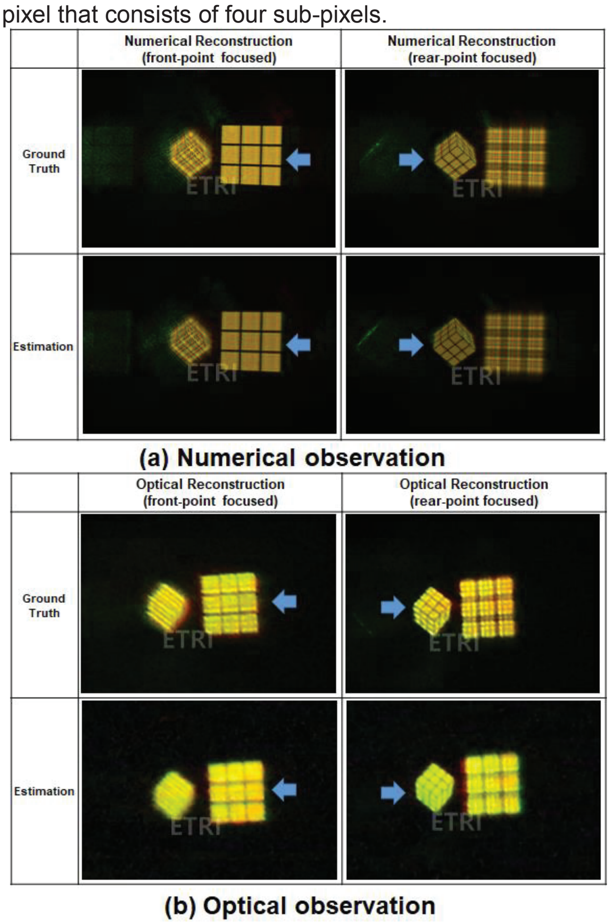

Fig. 4 Numerical and optical reconstruction from CGH

The experimental observations of optical reconstruction are made by using a DSLR camera (Canon EOD 5D Mark III). Photographs of holographic reconstructions are captured with the camera’s lens positioned near the focal position of the field lens (see Figure 3). The setting conditions of camera are given as follows: Shutter speed is 1.25 sec, aperture size F 2.8, and ISO 500. Figure 4 shows photographs of numerically and optically reconstructed scenes, supporting an accommodation effect of holographic 3D scenes; An object between two objects is within the camera’s focus and then its photographic image is sharp, while another object out of focus is clearly blurred as shown in Figure 4, which manifests the accommodation.

3 CONCLUSIONS

In conclusion, we demonstrate a novel, CNN-based model that estimates the depth map from multi-view colour images, so that we can successfully apply it to synthesize computer-generated holograms (CGHs). In order to prepare the training data for the holographic 3D content related to the full range of 360˚, a set of colour image and depth map per view obtains from the 3D scene by using the Maya™ software. Our proposed, deep learning method of depth map estimation is a kind of modified DenseNet. We analyze the quality property between the reference depth map and the depth map prepared from the proposed CNN-based model by directly demonstrating applicability for holographic 3D content, through both numerical reconstruction and optical reconstruction of the holographic 3D scene from each CGH. For our future work, we will reconstruct a complicated, detailed texture and further multi-coloured scene from a digital hologram, that we may make the qualitative analysis of holographic image quality by directly comparing CNN-based CGH from reference CGH.

Acknowledgment

This work was supported by the Institute of Information & Communication Technology Planning & Evaluation (IITP) grant funded by the Korea government (MIST) (No.2020-0-00954, Development of high efficiency high transparency large HOE screen)

REFERENCES

[1] MinSung Yoon, Hyeonji Kim, and HwaSeon Yeo, "Deep Depth Extraction of Multi-view Images for Holographic 3D Content," IMID2020 (2020).

[2] M. Yoon, K. Oh, H. Choo, and J. Kim, “A spatial light modulating LC device applicable to amplitude-modulated holographic mobile devices,” Industrial Informatics (INDIN), pp. 678-681 (2015).

[3] Ryoichi Horisaki *et. al.*, “Deep-learning-generated holography,” Appl. Opt. 57, 3859-3863 (2018).

[4] Sunghoon Im *et. al.*, “DPSNet: End-to-end Deep Plane Sweep Stereo,” ICLR (2019).

[5] Ibraheem Alhashim and Peter Wonka, “High Quality Monocular Depth Estimation via Transfer Learning,” https://arxiv.org/abs/1812.11941 (2018).

[6] Wai Hon Lee, “Sampled Fourier transform hologram generated by computer,” Applied Optics, Vol. 9, No. 3, 639-644 (1970).

[7] C. B. Burckhard, “A simplification of Lee’s method of generating holograms by computer,” Applied Optics, Vol. 9, No. 8, 1949-1949 (1970).