PROCEEDINGS OF SPIE

SPIEDigitalLibrary.org/conference-proceedings-of-spie

Computer-generated hologram compression with attention-based deep convolutional neural network

Shen, Zhelun, Yang, Guanglin, Xie, Haiyan

Zhelun Shen, Guanglin Yang, Haiyan Xie, "Computer-generated hologram compression with attention-based deep convolutional neural network," Proc. SPIE 11898, Holography, Diffractive Optics, and Applications XI, 118981Y (9 October 2021); doi: 10.1117/12.2602659

Event: SPIE/COS Photonics Asia, 2021, Nantong, Jiangsu, China

Computer generated hologram compression with attention-based deep convolutional neural network

Zhelun Shen[a], Guanglin Yang*[a], Haiyan Xie[b]
[a]Laboratory of Signal and Information Processing, Department of Electronics, Peking University, Beijing 100871, China
[b]China Science Patent and Trademark Agent, Beijing 100083, China

ABSTRACT

We propose an attention-based deep convolutional neural network for computer generated hologram (CGH) compression, where a channel attention mechanism is applied to both computer generated hologram compression and reconstruction. By applying deep convolutional neural networks in the compression process, we can extract more compact and representative information than bicubic interpolation. Additionally, a channel attention mechanism is applied to selectively emphasize informative features and suppress less useful ones for both CGH compression and reconstruction. By employing attention mechanisms to enhance the feature representation ability of deep convolutional neural networks, we can further improve the performance of the reconstructed computer generated hologram. Experimental results show our method can better recover the compressed computer generated hologram than only employing convolutional neural networks in the reconstruction process.
Keywords: computer generated hologram, image compression, convolutional neural network

1. INTRODUCTION

Computer Generated Holography (CGH) aims to record and reconstruct a desired complex light wavefront with both amplitude and phase information, which is fundamental to various applications such as three-dimensional dynamic holographic display[1], holographic projection[2], and optical information security[3]. Moreover, as the CGH typically contains a large number of pixels and occupies considerable storage space, the compression of CGH has drawn great attention from the community[4, 5].

Recently, the unprecedented advancement of deep learning in computer vision tasks has been extended to the CGH and some work has proposed to use the deep convolutional neural network[6] to compress the CGH. However, these works only transfer the method of compressing ordinary images to CGHs without considering the nature of holograms. In this paper, we propose an attention-based deep convolutional neural network for hologram compression. Our contribution can be divided into two parts. First, previous work only employs deep convolutional neural networks to recover the lost information while still use JPEG compression[6] or bicubic interpolation[7] to compress the image. For natural images, since we hope that the compressed image and reconstructed image can keep the structure similar to the original image, such operations are reasonable. However, for holograms, we only focus on the quality of the reconstructed hologram and do not care about the structure of the compressed image. Thus, we propose to apply deep convolutional neural networks to get more compact and representative information in the compression process. Second, we propose a channel attention mechanism to selectively emphasize informative features and suppress less useful ones of the compression and reconstruction process. By employing attention mechanisms to enhance the feature representation ability of deep convolutional neural networks, we can further improve the performance of the reconstructed CGH.

2. ATTENTION BASED DEEP CONVOLUTIONAL NEURAL NETWORK

We propose an attention-based deep convolutional neural network for hologram compression, which can be divided into three parts. 1) Hologram generation 2) Hologram compression and decompression 3) Image reconstruction.

*ygl@pku.edu.cn;

Holography, Diffractive Optics, and Applications XI, edited by Yunlong Sheng, Changhe Zhou, Liangcai Cao
Proc. of SPIE Vol. 11898, 118981Y · © 2021 SPIE · 0277-786X · doi: 10.1117/12.2602659

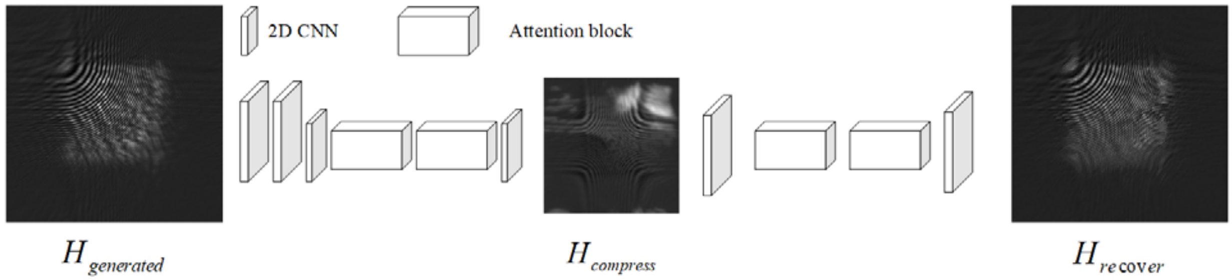


Figure 1. The architecture of our proposed network.

Hologram generation: Given a color image, we first employ the holographic principle of Leith and Upatnieks with regard to off-axis reference beam holograms[8] to generate the corresponding Fresnel hologram $H_{generated}$. The amplitude and phase transmittance of a hologram recorded can be defined as follows:

$$\begin{aligned} I(x,y) &= \left| R(x,y)\exp[j\phi(x,y)] + A(x,y)\exp[j\varphi(x,y)] \right|^2 \\ &= R(x,y)^2 + A(x,y)^2 + 2R(x,y)A(x,y) \times \cos[\phi(x,y) - \varphi(x,y)] \end{aligned} \quad (1)$$

Where $R(x,y)\exp[j\phi(x,y)]$ is the reference wave and $A(x,y)\exp[j\varphi(x,y)]$ denotes the object wave.

Hologram compression and decompression: Let us denote the generated hologram as $H_{recover}$, our goal is to generate a compressed hologram H_{compress} and then recover from H_{compress} a hologram H_{recover} that is as similar as possible to the original hologram $H_{generated}$. We propose an attention-based deep convolutional neural network to achieve this goal. Previously, some work has introduced using the deep convolutional neural network[6] to compress holograms. However, these works only transfer the method of compressing ordinary images to holograms without considering the nature of holograms. Instead, we propose a channel attention to better recover holograms and apply the attention mechanism to both compression and decompression processes.

Specifically, typical convolutional neural networks equally treat every channel of extracted features. However, intuition indeed tells us that a more reasonable way is to selectively emphasize informative features and suppress less useful ones. Thus, we propose a channel attention mechanism[9-11] to enhance such ability. Our channel attention mechanism can be divided into two steps:

Squeeze: Given an input feature map $u \in R^{H \times W \times C}$, we first propose a squeeze operation to extract the global spatial information into a channel descriptor. Specifically, the squeeze operation is implemented by global average pooling, which is defined as:

$$z_c = \frac{1}{H \times W} \sum_{i=1}^{H} \sum_{j=1}^{W} u_c(i,j) \quad (2)$$

Where u_c is the c^{th} channel of the input feature map.

Excitation: After generating the channel descriptor in the squeezing step, we further propose an excitation operation to aggregate the channel-wise dependencies. Specifically, the excitation operation is implemented by a sigmoid activation based gating mechanism, which is defined as:

$$s = \sigma(W_2 \delta(W_1 z)) \quad (3)$$

Where σ denotes the sigmoid activation function and δ is the Relu function. W_1 and W_2 denotes two fully connected (FC) layers. Finally, we can rescale the input feature map u with the activations s:

$$\widetilde{u_c} = s_c u_c \quad (4)$$

Intuitively, our channel attention mechanism can be regarded as a self-attention function, which reweights the weight of each channel to selectively emphasize informative features and suppress less useful ones.

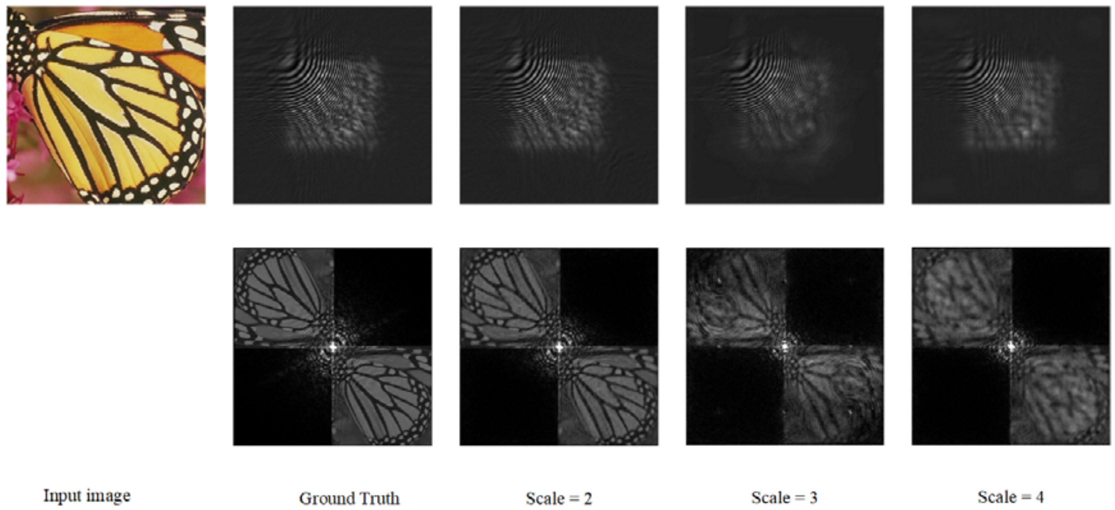


Figure 2. Visualization reconstructed images with different compression scales. The left panel shows the input image, and for each example, the first row shows the reconstructed hologram and the second row shows the corresponding recovered image.

All in all, the network architecture consists of two parts: hologram compression and hologram decompression. In the hologram compression, we first employ three convolutional layers to compress $H_{generated}$ to the desired resolution and extract shallow features. Then two residual blocks [12] are employed to further regularize the extracted features. As discussed before, we add channel attention to selectively emphasize informative features and suppress less useful ones at the end of each residual block. Finally, the compressed hologram is generated via one convolutional layer. In hologram decompression, we replace the first three convolutional layers with transposed convolution to upscale the compressed hologram to the original resolution. Then, for maintaining a uniform structure, we design similar network architecture to get the recovered hologram.

Image reconstruction: Finally, we employ fresnel transform techniques to reconstruct the original image from the recovered hologram.

3. EXPERIMENT AND RESULTS

Dataset: Following SRCNN[7], we employ a relatively small training set that consists of 91 images to train our network and use Set5[5, 7] as the validation set. All images are transformed to 256×256 the corresponding Fresnel hologram with the method proposed in the previous work[8].

Implementation Details: We use PyTorch to implement our attention-based deep convolutional neural network and employ Adam to train the whole network in an end-to-end way. The batch size is set to 16. We employ the L1 loss function to train our network for 300 epochs. The initial learning rate is set to 0.001 and is downscaled by 10 after epoch 200.

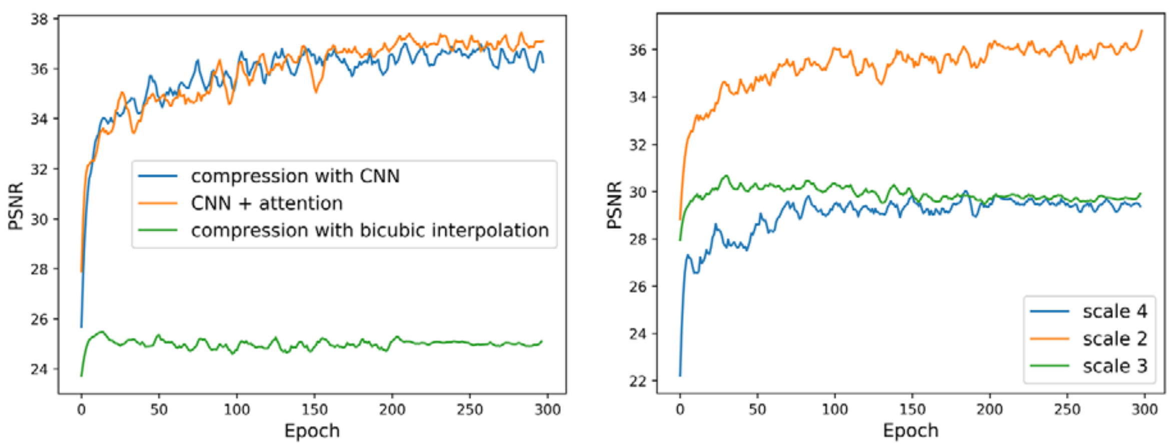
Ablation Study: We perform various ablation studies to show the effectiveness of each component in our network. Results are shown in Table 1. As shown, we compare using convolutional neural networks to compress holograms with the widely used bicubic interpolation. Results show our method can significantly improve the PSNR of reconstructed images from 25.60 to 36.56. In addition, we also test the impact of using the attention mechanism. As shown, the attention mechanism can further promote the PSNR of reconstructed images from 36.56 to 37.73. Moreover, as shown in Table 2, our method can adjust the downscaling factor (compression scale) to make a trade-off between reconstruction quality and compression ratio, which gives our approach more flexibility to adjust the reconstruction quality. Visualization results can be seen in Figure 2.

Table 1. Ablation study results of the proposed network on the validation set.

Compression with bicubic interpolation	Compression with CNN	Attention mechanism	Compression Scale	PSNR(dB)
√			2	25.50
	√		2	36.56
	√	√	2	37.73

Table 2. Ablation study results of the compression scale.

Compression Scale	PSNR(dB)
2	37.73
3	30.99
4	30.80



(a) (b)

Figure 3. Left: Plots comparing PSNR vs iteration number with different network architecture on the Set 5 dataset. Right: Plots comparing PSNR vs iteration number with different compression scales on the Set 5 dataset.

4. CONCLUSION

We propose an attention-based deep convolutional neural network for hologram compression, where a channel attention mechanism is applied to both hologram compression and reconstruction. By employing the channel attention mechanism to selectively emphasize informative features and suppress less useful ones, we can enhance the feature representation ability of deep convolutional neural networks. Experiments show our attention-based deep convolutional neural network can greatly improve the reconstructed image quality of the compressed CGH. In the future, we plan to extend our attention-based deep convolutional neural network to other terms of Computer Generated Holography.

5. ACKNOWLEDGEMENTS

The authors would like to thank the Spatial Image Processing Laboratory for their support. This work was supported by the National Science Foundation of China (No. 62071009).

REFERENCES

[1] F. Yaraş, H. Kang, and L. Onural, "Real-time phase-only color holographic video display system using LED illumination," Applied optics, 48(34), H48-H53 (2009).

[2] E. Buckley, "Holographic laser projection," Journal of display technology, 7(3), 135-140 (2011).

[3] Y.-Y. Wang, Y.-R. Wang, Y. Wang *et al.*, "Optical image encryption based on binary Fourier transform computer-generated hologram and pixel scrambling technology," Optics and lasers in engineering, 45(7), 761-765 (2007).

[4] M. Liu, G. Yang, and H. Xie, "Method of computer-generated hologram compression and transmission using quantum back-propagation neural network," Optical Engineering, 56(2), 023104 (2017).

[5] S. Hou, G. Yang, and H. Xie, "Optimized initial weight in quantum-inspired neural network for compressing computer-generated holograms," Optical Engineering, 58(5), 053105 (2019).

[6] S. Jiao, Z. Jin, C. Chang *et al.*, "Compression of phase-only holograms with JPEG standard and deep learning," Applied Sciences, 8(8), 1258 (2018).

[7] C. Dong, C. C. Loy, K. He *et al.*, "Image super-resolution using deep convolutional networks," IEEE transactions on pattern analysis and machine intelligence, 38(2), 295-307 (2015).

[8] E. N. Leith, and J. Upatnieks, "Reconstructed wavefronts and communication theory," JOSA, 52(10), 1123-1130 (1962).

[9] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 7132-7141 (2018).

[10] C. Peng, and J. Ma, "Semantic segmentation using stride spatial pyramid pooling and dual attention decoder," Pattern Recognition, 107, 107498 (2020).

[11] J. Fu, J. Liu, J. Jiang *et al.*, "Scene segmentation with dual relation-aware attention network," IEEE Transactions on Neural Networks and Learning Systems, (2020).

[12] K. He, X. Zhang, S. Ren *et al.*, "Deep residual learning for image recognition." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 770-778 (2016).