Transformer-based Denoising of Microlens Array Shearing Interferogram

Ruoxin Wang, Wenkai Zhao, Zhanchen Zhu, Bo Wang, Chi Fai Cheung#

State Key Laboratory of Ultra-precision Machining Technology, Department of Industrial and Systems Engineering,
The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong
Corresponding Author / Email: benny.cheung@polyu.edu.hk, TEL: +852-2766-7905, FAX: +852-2362-5267

KEYWORDS: Transformer, Denoising, Microlens arrays, Shearing interferogram, Precision measurement

Microlens arrays are the key component in the next generation of 3D imaging systems, since it exhibits good optical properties such as extremely large field of view angles, low aberration and distortion, high temporal resolution, and infinite depth of field. During microlens array production, problems of mold making, and non-uniform expansion occur, caused by the heating cycle. Interferometry is an inherent high-precision and well-established optical technique for extracting the phase variation of an interferogram and has been appealing for characterizing optical phenomena in physics, chemistry, and engineering. Because of self-interference in shearing interferometry, it is more flexible for use in interferometer. However, a large amount of noise in the interference pattern seriously affects the phase reconstruction and measurement accuracy. Compared with traditional denoising methods, such as Gaussian filter, non-local mean filter, and bilateral filter, deep learning-based methods can automatically learn complex patterns and structures in images without manually designing features or filters. This paper attempts to present a transformer-based method to reduce the noises in interferograms of microlens arrays generated by shearing interferometry Restormer in which multiDconv head transposed attention and gated-Dconv feed-forward network are developed to perform linear computation complexity and controlled feature transformation. A series of comparison experiments are conducted to show the superiority of Restormer.

NOMENCLATURE

CNN = convolutional neural networks
MDTA = multi-Dconv Head Transposed Attention
GDFN = gated Dconv Feed-Forward Network

1. Introduction

Shearing interferometry [1] is a powerful optical technique widely used in the fields of engineering, materials science, and physics to measure and analyze wavefront distortions, surface deformations, and refractive index variations. This method leverages the principles of interference, where two or more light waves superimpose to produce a resultant wave, to provide high-precision measurements of optical path differences. Unlike traditional interferometry, which typically compares a test wavefront with a reference wavefront, shearing interferometry involves the comparison of a wavefront with a slightly shifted or "sheared" version of itself. This self-referencing nature makes it particularly robust against environmental disturbances and alignment errors.

The fundamental concept of shearing interferometry revolves around the generation of two closely spaced, parallel wavefronts from a single incident wavefront. These wavefronts interfere with each other, producing an interference pattern that encodes information about the phase differences between them. By analyzing this pattern, one can extract detailed information about the wavefront's shape and any deviations from the ideal form. This makes shearing interferometry an invaluable tool for applications such as optical testing, stress analysis in transparent materials, and the study of fluid dynamics. However, like many optical measurement techniques, shearing interferometry can be susceptible to various sources of noise that can affect the accuracy and clarity of the interference patterns, such as environmental vibrations, air turbulence, thermal fluctuations, optical component quality, coherence length of the light source, electronic noise and alignment error [2]. To improve the reliability and precision of measurement, denoising shearing interferogram is a very necessary and challenging task.

In recent years, transformer-based methods have revolutionized various fields of artificial intelligence, particularly in natural language processing (NLP) [3] and computer vision [4]. These methods, characterized by their ability to model long-range dependencies and capture intricate patterns in data, have shown remarkable success in tasks such as language translation [5], image classification [6], and text generation [7]. Building on this success, researchers have begun to explore the potential of transformer architectures for denoising applications, where the goal is to remove noise from signals or images to restore their original quality.

Denoising is a critical preprocessing step in numerous applications, including medical imaging, audio processing, and remote sensing. Traditional denoising techniques, such as wavelet transforms [8], Gaussian filtering [9], and non-local means [10], have been effective to some extent but often struggle with preserving fine details and textures while removing noise. Deep learning-based approaches, particularly convolutional neural networks (CNNs) [11], have shown significant improvements in denoising performance. However, CNNs are inherently limited by their local receptive fields, which can hinder their ability to capture global context and long-range dependencies in the data. Transformer-based methods, with their self-attention mechanisms, offer a promising alternative by enabling the modeling of global relationships within the data. The self-attention mechanism allows transformers to weigh the importance of different parts of the input, making them particularly adept at capturing both local and global features. This capability is crucial for effective denoising, as noise can be distributed across various scales and regions of the signal or image. In this paper, a transformer-based method is developed to conduct shearing interferogram denoising to improve the precision of measurements of microlens arrays.

2. Transformer-based Denoising Method

2.1 Overall framework

In the framework, Restormer [12] network is adopted to conducted shearing interferogram denoising to improve the measurement results. Given a noisy interferogram $I \in \mathbb{R}^{H \times W \times 3}$, a 4-level symmetric encoder-decoder is adopted to map interferogram into feature embedding $F_m \in \mathbb{R}^{H \times W \times 2C}$, in which each level of encoder-decoder contains multiple Transformer blocks. The number of blocks gradually increases from the top to bottom levels to maintain efficiency. The output feature embeddings of 4 level encoders are $F_1 \in \mathbb{R}^{H \times W \times C}$, $F_2 \in \mathbb{R}^{\frac{H}{2} \times \frac{H}{2} \times 2C}$, $F_3 \in \mathbb{R}^{\frac{H}{4} \times \frac{H}{4} \times 4C}$ and $F_4 \in \mathbb{R}^{\frac{H}{8} \times \frac{H}{8} \times 8C}$.

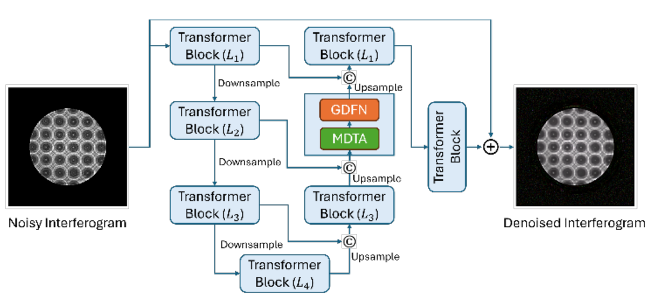

Fig. 1 Overall framework of transformer-based denoising method

These outputs are concatenated with the unsampled outputs of 4 level decoders by skip connections. Another Transformer block is used to further enrich features in the refinement stage operating at high spatial resolution. A convolution layer is used to the refined features to generate residual image $R \in \mathbb{R}^{H \times W \times 3}$. The original image is added to residual image to obtain the final denoised interferogram. In addition, to improve the model efficiency and accuracy, a Multi-Dconv Head Transposed Attention (MDTA) and Gated Dconv Feed-Forward Network (GDFN) is used in each Transformer block.

MDTA. In MDTA, a crucial aspect is the application of self-attention across channels, rather than spatial dimensions, to implicitly capture global context through cross-channel covariance. Moreover, depth-wise convolutions are a vital component, as they allow for emphasis on local context before calculating feature covariance, ultimately generating a comprehensive attention map.

GDFN. In GDFN, Restormer introduces two key innovations to enhance representation learning in feed-forward network: firstly, the integration of a gating mechanism, and secondly, the incorporation of depth-wise convolutions.

2.2 Training Settings

Loss function. In this paper, L1 loss function is used during training. It can be defined in Eq. (1)

$$l(y, \hat{y}) = \frac{1}{N}\sum_{n=1}^{N}|y_n - \hat{y}_n| \tag{1}$$

Optimizer. To optimize the loss function, the AdamW optimizer is adopted. The initial learning rate is set as $3*10^{-4}$ with weight decay of $1*10^{-4}$.

Architecture. The number of Transformer blocks in 4 levels are set as 4, 6, 6 and 8, respectively. In addition, the number of heads in multi-head self-attention of each level is set to 1, 2,4, and 8, respectively. The dimension C is set to 48. The last Transformer Block is regarded as refinement stage and the number of blocks is set to 4.

3. Experimental settings and Results

3.1 Microlens array measurement experiments

The working principle of the quadric wave lateral shearing interferometer technique is shown in Fig. 2. The wavefront to be measured is copied into four wavefronts with the same wavefront distortion information and different exit angles through the beam splitter, which are mutually displaced and interfere with each other on the CCD target surface of the optocoupler device; the collected interference pattern is processed by a computer, and the phase distribution of the wavefront to be measured is restored through the wavefront reconstruction algorithm.

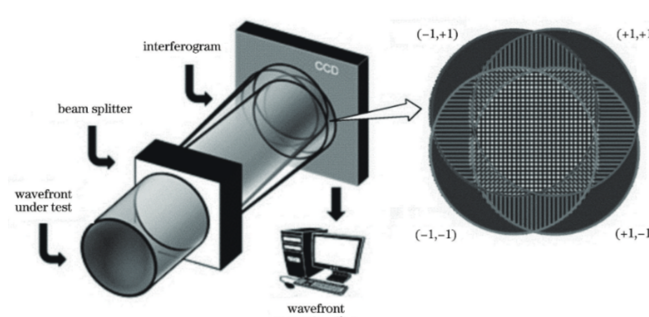

Fig. 2 The mechanism of shearing interferometry

The beam splitter is a key device of quadric wave lateral shearing interferometer technique and is currently implemented using diffraction grating. As shown in Fig. 2, under ideal circumstances, after the wavefront to be measured passes through the grating, there are only (+1, +1), (+1, -1), (-1, -1) and (-1, +1) order diffraction lights, forming four-wavefront transverse shear interference patterns. Using this technique, a microlens array is measured. The design of microlens array is shown in Fig. 3.

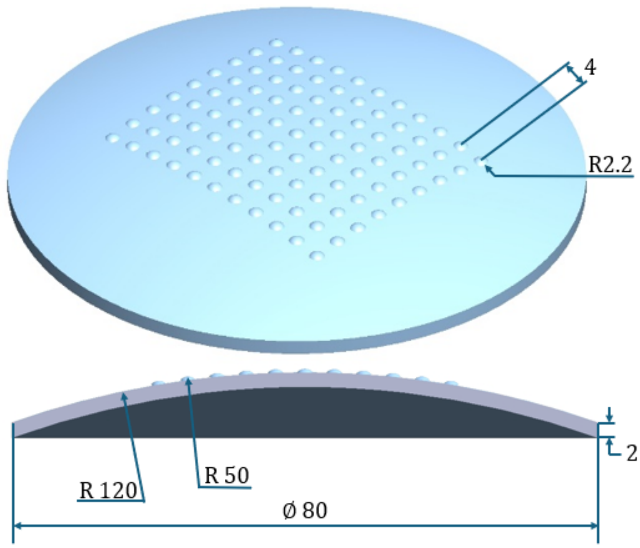

Fig. 3 The design of microlens array.

3.2 Results and discussion

The result of Restormer is shown in Fig. 4. Fig. 4 (a) shows the original interferogram, and Fig. 4(b) is the denoised interferogram. The bottom of Fig. 4 is the partially enlarged interferogram. Compared to the original interferogram, the denoised interferogram has much less noise, which helps the further phase unwrapping and reconstruction and improve the measurement accuracy.

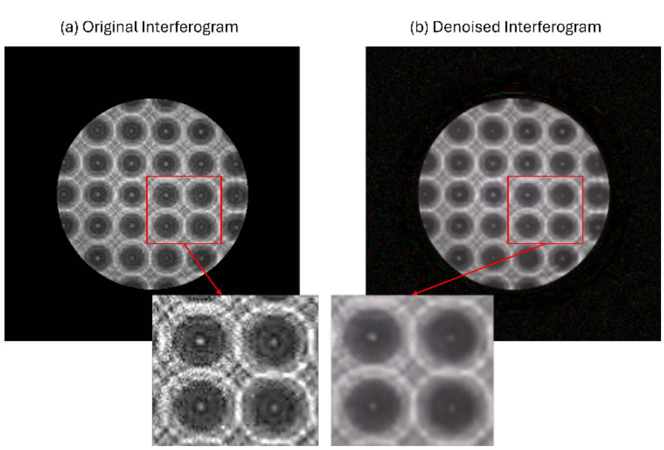

Fig. 4 The denoising result of Restormer

In addition, we compare the denoising performance on different noise, such as Gaussian color noise, Gaussian gray noise and real noise. The results in Fig. 5 show that conduct real noise denoising performs better, which means that in the original interferogram, there are some other real noises except for Gaussian noise. These experimental results also give proof that the Gaussian filter is not enough for interferogram.

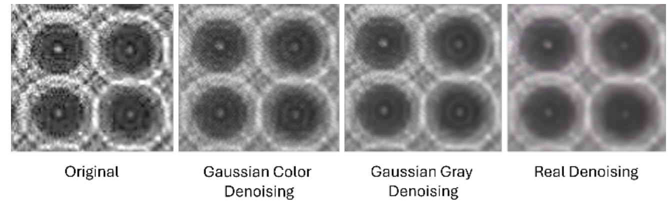

Fig. 5 The Restormer denoising performance on different noises

To show the superiority of Restormer, we compared the results with state-of-the-art models, such as MIRNet [13], MRPNet [14], CycleISP [15] and DANet [16]. The results in Fig. 6 show that the Restormer has better denoising performance and CycleISP has worst denoising performance. The MRPNet achieved competitive results compared to Restormer.

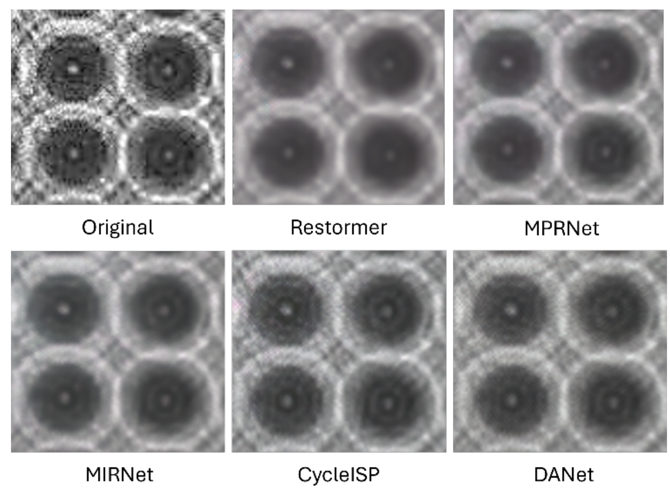

Fig. 6 The comparison results with state-of-the-art models.

4. Conclusions

In this paper, a transformer-based method, Restormer, is adopted to denoise the interferogram of microlens array. The results show that Restormer can perform effective denoising on interferogram. In addition, the different noises denoising experiments are conducted to show that Gaussian noises denoising is not enough for interferogram denoising tasks. The comparison experiments with other state-of-the-art models show that Restormer has better denoising performance on interferogram denoising tasks.

ACKNOWLEDGEMENT

The work described in this paper was mainly supported by a grant from the Research Grants Council (Project Code: C5031-22G) of the Government of the Hong Kong Special Administrative Region, China,

REFERENCES

1. Hardy, John W., and Alan J. MacGovern. "Shearing interferometry: a flexible technique for wavefront measurement." In *Interferometric Metrology*, vol. 816, pp. 180-195. SPIE, 1987.
2. Lopes, H., JV Araújo dos Santos, and P. Moreno-García.

"Evaluation of noise in measurements with speckle shearography." *Mechanical Systems and Signal Processing* 118 (2019): 259-276.

3. Gillioz, Anthony, Jacky Casas, Elena Mugellini, and Omar Abou Khaled. "Overview of the Transformer-based Models for NLP Tasks." In *2020 15th Conference on computer science and information systems (FedCSIS)*, pp. 179-183. IEEE, 2020.
4. Han, Kai, Yunhe Wang, Hanting Chen, Xinghao Chen, Jianyuan Guo, Zhenhua Liu, Yehui Tang et al. "A survey on visual transformer." *arXiv preprint arXiv:2012.12556* (2020).
5. Camgoz, Necati Cihan, Oscar Koller, Simon Hadfield, and Richard Bowden. "Sign language transformers: Joint end-to-end sign language recognition and translation." In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 10023-10033. 2020.
6. Bhojanapalli, Srinadh, Ayan Chakrabarti, Daniel Glasner, Daliang Li, Thomas Unterthiner, and Andreas Veit. "Understanding robustness of transformers for image classification." In *Proceedings of the IEEE/CVF international conference on computer vision*, pp. 10231-10241. 2021.
7. Deng, Yuntian, and Alexander Rush. "Cascaded text generation with markov transformers." *Advances in Neural Information Processing Systems* 33 (2020): 170-181.
8. Dautov, Çiğdem Polat, and Mehmet Siraç Özerdem. "Wavelet transform and signal denoising using Wavelet method." In *2018 26th Signal Processing and Communications Applications Conference (SIU)*, pp. 1-4. Ieee, 2018.
9. Singh, Amanjot, and Jagroop Singh. "Comparative analysis of gaussian filter with wavelet denoising for various noises present in images." *Indian Journal of Science and Technology* 9, no. 47 (2016): 112.
10. Buades, Antoni, Bartomeu Coll, and Jean-Michel Morel. "Non-local means denoising." *Image Processing On Line* 1 (2011): 208-212.
11. Ilesanmi, Ademola E., and Taiwo O. Ilesanmi. "Methods for image denoising using convolutional neural network: a review." *Complex & Intelligent Systems* 7, no. 5 (2021): 2179-2198.
12. Zamir, Syed Waqas, Aditya Arora, Salman Khan, Munawar Hayat, Fahad Shahbaz Khan, and Ming-Hsuan Yang. "Restormer: Efficient transformer for high-resolution image restoration." In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 5728-5739. 2022.
13. Zamir, Syed Waqas, Aditya Arora, Salman Khan, Munawar Hayat, Fahad Shahbaz Khan, Ming-Hsuan Yang, and Ling Shao. "Learning enriched features for real image restoration and enhancement." In *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XXV 16*, pp. 492-511. Springer International Publishing, 2020.
14. Zamir, Syed Waqas, Aditya Arora, Salman Khan, Munawar Hayat, Fahad Shahbaz Khan, Ming-Hsuan Yang, and Ling Shao. "Multi-stage progressive image restoration." In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 14821-14831. 2021.
15. Zamir, Syed Waqas, Aditya Arora, Salman Khan, Munawar Hayat, Fahad Shahbaz Khan, Ming-Hsuan Yang, and Ling Shao. "Cycleisp: Real image restoration via improved data synthesis." In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 2696-2705. 2020.
16. Yue, Zongsheng, Qian Zhao, Lei Zhang, and Deyu Meng. "Dual adversarial network: Toward real-world noise removal and noise generation." In *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part X 16*, pp. 41-58. Springer International Publishing, 2020.