Super-resolution imaging of telescopic systems based on optical-neural network joint optimization (Postprint)

Authors: You-Hong Sun, Tao Zhang, Hao-Dong Shi, Qiang Fu, Jia-Nan Liu, Kai-Kai Wang and Chao Wang

Date: 2024-10-08T00:00:00+00:00

Abstract

Optical telescopes are essential instruments for acquiring optical information of distant objects, with resolution being a critical metric that measures the capability to discern object details. However, due to influences from system aberrations, atmospheric seeing, and other factors, images captured by ground-based telescopes often suffer from degradation, resulting in reduced resolution. This paper proposes an optical-neural network joint optimization approach to enhance the resolution of observed images through the co-optimization of the telescope system's point-spread function (PSF) and an image super-resolution (SR) network. To accelerate image reconstruction, we have designed a lightweight generative adversarial network (LCR-GAN) that operates significantly faster than current state-of-the-art unsupervised networks. To implement the network-trained PSF reconstruction within the optical path, a phase mask is introduced. This improves the image reconstruction performance of LCR-GAN by reconstructing the PSF that optimally matches the network. Simulation and experimental validation results demonstrate that, compared with pure deep learning methods, the SR image

Full Text

Preamble

Research in Astronomy and Astrophysics, 24:095015 (11pp), 2024 September

https://doi.org/10.1088/1674-4527/ad4fc1

Super-resolution Imaging of Telescopic Systems based on Optical-neural Network Joint Optimization

You-Hong Sun[1],[2], Tao Zhang[3], Hao-Dong Shi[1],[2], Qiang Fu[1],[2], Jia-Nan Liu[1],[2], Kai-Kai Wang[1], and Chao Wang[1],[2]

[1] Research Institute of Chongqing, Changchun University of Science and Technology, Changchun 130022, China; 1768811682@qq.com
[2] National and Local Joint Engineering Research Center for Space Photoelectric Technology, Changchun University of Science and Technology, Changchun 130022, China
[3] The Institute of Remote Sensing Satellites, China Academy of Space Technology (CAST), Beijing 100094, China

Received 2023 December 22; revised 2024 April 22; accepted 2024 May 7; published 2024 September 18

Abstract

Optical telescopes are essential tools for acquiring optical information about distant objects, and resolution is a critical metric that measures the ability to observe fine details. However, due to system aberrations, atmospheric seeing, and other factors, images observed by ground-based telescopes are often degraded, resulting in reduced resolution. This paper proposes an optical-neural network joint optimization method to improve the resolution of observed images by co-optimizing the point-spread function (PSF) of the telescopic system and the image super-resolution (SR) network. To accelerate image reconstruction, we designed a lightweight generative adversarial network (LCR-GAN) with significantly fewer parameters than state-of-the-art unsupervised networks. To physically implement the PSF learned by the network in the optical path, we introduce a phase mask that improves the image reconstruction performance of LCR-GAN by reconstructing the PSF that best matches the network.

Simulation and verification experiments demonstrate that compared with pure deep learning methods, the SR images reconstructed by our method are richer in detail and make it easier to distinguish stars or stripes.

Key words: Techniques: image processing – Telescopes – Stars: imaging

1. Introduction

Optical telescopes are widely used in astronomical observation, remote sensing, and optical surveillance as essential tools for obtaining optical information from distant targets \cite{Wang_{2015}, Wang_{2020}, Mikhail_{2019}, He_{2021}}. Resolution reflects the ability to distinguish two adjacent objects and serves as a crucial indicator for observing object details, with higher resolution typically enabling richer object detail. According to the Rayleigh criterion, a system's resolution can theoretically reach the diffraction limit.

However, due to system aberrations, atmospheric seeing, and other factors, images observed by ground-based telescopes are often degraded, resulting in reduced resolution \cite{Li_{2018}}. Super-resolution (SR) technology can reconstruct high-resolution (HR) images from one or more low-resolution (LR) observation images to improve image resolution \cite{Park_{2003}}. Among mainstream image SR algorithms, interpolation-based methods produce mediocre results when reconstructing fine textures \cite{Lehmann_{1999}}. Prior learning-based algorithms are sensitive to training sample selection and cannot simultaneously recover high- and low-frequency information in original astronomical images \cite{Yan_{2015}}. Iterative Back Projection \cite{Irani_{{{Park}}*{{1991}}}} and Projection Onto Convex Sets \cite{Youla*{{{Webb}}_{{1982}}}} struggle to reconstruct the texture features of astronomical images.

In recent years, with the vigorous development of deep learning, numerous neural networks have emerged. Through combinations of different convolutional layers, linear layers, activation functions, etc., deep convolutional neural networks can learn highly complex functional features and thus solve previously intractable problems across many fields \cite{Wang_{2023c}}. Deep learning is also widely applied in telescopic systems for tasks such as denoising and SR \cite{Rahman_{2020}, Sweere_{2022}}. In 2022, Shoubaneh et al. used an improved generative adversarial network model to enhance the resolution of ground-based images from the Subaru Hyper Suprime-Cam to match that of the Hubble Space Telescope \cite{Shoubaneh_{2022}}. Most convolutional neural network-based methods require paired data for supervised training. However, producing paired data in astronomical observation is typically costly, and algorithmically synthesized images may contain unrealistic details that affect model generalization. In contrast, unsupervised networks overcome the limitation of supervised learning networks that rely on LR-HR paired data to improve image resolution \cite{Chang_{{{Wetzstein}}_{{2019}}}}. They can learn distribution information from HR observation images with similar celestial environments and use this as prior information to enhance the resolution of distant galaxy observation images.

From the perspective of the overall optical imaging process, the imaging process of optical systems is essentially equivalent to optical computation \cite{Yuan_{2020}, Wang_{2023a}}. The goal of joint optimization methods is to achieve optimal matching between the optical system and image processing algorithm through cooperative optimization of the "optical system" and "digital image post-processing," thereby obtaining superior imaging performance. In 2019, Peng et al. from Stanford University proposed an end-to-end design method from optical system to image processing that achieves achromatic depth-of-field expansion and SR imaging \cite{Peng_{2015}, Peng_{2019}}. In parallel to our work, Sun Q. et al. used an end-to-end framework to simultaneously optimize a single-lens imaging system and a reconstruction network to reconstruct SR images from raw measurement data \cite{Sun_{2020}}.

This paper proposes an optical-neural network joint optimization method to improve the resolution of observation images from telescopic systems for known objects with basic outlines or shapes, such as stars and nebulae. First, a deep learning convolutional layer equivalent to the telescopic system point-spread function (PSF) is integrated into the front end of the neural network, and the PSF and network parameters are co-optimized through the training process. After training, a phase mask is designed to physically implement the learned PSF in the optical path. Considering the error between the actual PSF reconstructed by the phase mask and the PSF learned by the network, the phase mask is added to the telescopic system to collect its observation images, and the network is retrained to fine-tune the parameters. Finally, we use Peak Signal-to-Noise Ratio (PSNR) and other metrics to evaluate the quality of reconstructed images. In this paper, HR observation images can be reconstructed rapidly by designing a high-performance unsupervised network and co-optimizing the network parameters and the PSF of the telescopic system.

This paper is organized as follows: In Section 2, we present the theoretical background. In Section 3, we describe the unsupervised network structure and loss function. In Sections 4 and 5, we evaluate the network performance and the SR effect of the joint optimization method. In Section 6, we present our conclusions.

2.1. Incoherent Imaging Model and Cooperative Optimization

First, we introduce the relationship between the incoherent imaging model of telescope systems and the convolutional layer of deep learning. The illumination source of telescope systems is usually incoherent light, which belongs to passive imaging \cite{Yang_{2022}}. When an object is imaged linearly and spatially incoherently by the telescope system, the invariant imaging model can be expressed as \cite{Goodman_{2005}}:

$$I_g(u,v) = I_o(u,v) * |H(u,v)|^2$$

where u and v are the spatial coordinates of the image plane, $I_g(u,v)$ is the ideal geometric irradiated image, and $|H(u,v)|^2$ is often referred to as the PSF, where $*$ denotes two-dimensional convolution. The imaging of the telescope system can be modeled as a space-invariant convolution of the object with the system's PSF.

According to theory, the incoherent imaging model for monochromatic light illumination can be modeled as a convolutional layer of deep learning \cite{Chang_{2018}}. The imaging of the telescopic system for monochromatic illumination discussed in this paper can be modeled as a convolution layer: the flipped PSF is used as the convolution kernel, the number of input and output channels are both 1, and the input and output feature maps correspond to the object and image, respectively. Therefore, we use a single-kernel convolutional layer to simulate the PSF of the optical system, which is integrated into the front end of the neural network, as diagrammed in Figure 1 [Figure 1: see original paper].

To improve the observation capability of the telescopic system on distant celestial bodies, as depicted in Figure 1, HR stellar images with similar celestial environments are used as the original HR images. The HR image is input into the joint optimization network for training: the HR image is first convolved with the equivalent convolution layer to synthesize the LR image, which is equivalent to the incoherent imaging process of the telescope system. Then the synthesized LR image is input into the deep learning network to reconstruct the SR image. The loss function between the SR image output by the network and the HR image is calculated, and the parameters of the network and the equivalent convolution layer (PSF of optical systems) are updated simultaneously through gradient feedback. The training process is repeated until an optimal match between the optical system encoding and the neural network decoding is achieved. After training, the equivalent convolution layer parameters are derived as the PSF of the optical system.

Next, it is necessary to physically implement the PSF learned by the network in the optical path to realize joint optimization of the optical-neural network in the physical sense. In this paper, a phase mask is used to modulate the two-dimensional phase distribution in the optical field to obtain the ideal PSF in the image plane.

2.2. Phase Mask Solution

In this paper, a double-glued telescopic objective (DGTO) was used instead of the actual telescopic system to simplify the derivation process. As displayed in the lower-left corner of Figure 1, a phase mask is placed immediately behind the DGTO with the same net aperture as the DGTO. The detector is placed in the image plane of the DGTO.

Suppose that light emitted by a point source on the object surface propagates to the telescope system, forming an initial amplitude $U_1(x_1, y_1)$ and initial phase $\psi(x_1, y_1)$ on the surface behind the last lens. Let the modulation phase of the phase mask be $\phi(x_1, y_1) = \phi_0(x_1, y_1) + f(x_1, y_1)$, where $\phi_0(x_1, y_1)$ is the preset phase part. z_1 denotes the distance between the rear surface of the telescope system and the image plane. Considering that the phase mask is close to the rear surface of the telescope system and generally has a small thickness, (x_1, y_1) denotes the position coordinate of the rear surface of the telescopic system and also represents the position coordinate of the phase mask.

According to scalar diffraction theory \cite{Goodman_{2005}, Wang_{2023b}}, the complex amplitude of the light field in the image plane is:

$$U(x,y) = \frac{e^{ikz_1}}{i\lambda z_1} \iint U_1(x_1,y_1) e^{i\psi(x_1,y_1)} e^{i\phi(x_1,y_1)} e^{\frac{ik}{2z_1}[(x-x_1)^2+(y-y_1)^2]} dx_1 dy_1$$

Equation (3) is essentially a Fourier transform; therefore, the PSF of the system can be expressed as:

$$h(x,y) = \left| \mathcal{F}\{U_1(x_1,y_1) e^{i\psi(x_1,y_1)} e^{i\phi(x_1,y_1)}\} \right|^2$$

where λ is the wavelength of the point source, (x_1, y_1) denotes the spatial position of the phase mask, and $\mathcal{F}$ denotes the Fourier transform operator. According to Equation (4), we set the PSF of the telescope system as the PSF learned by the network. We set up the telescope system in the non-sequential mode of optical design software and trace one million grid rays from the point source. A detector records the coherent superposition of complex amplitudes of the plane waves corresponding to all rays to obtain the initial light field of Equation (2). Next, the Gerchberg–Saxton (GS) phase retrieval algorithm is used to solve the modulation phase $f(x_1, y_1)$. The GS phase retrieval algorithm is an iterative method to recover object phase with strong anti-disturbance capability and serves as an important tool for phase recovery \cite{Gerchberg_{{{Saxton}}_{{1972}}}}.

Finally, according to the equation that relates phase difference to the thickness of diffractive optical elements \cite{Wang_{2022}, Xu_{2022}}, we design a mask specifying the physical height on a substrate with refractive index n, which becomes the phase mask. The relationship is expressed as:

$$\phi(x_1,y_1) = k \cdot (n-1) \cdot t(x_1,y_1)$$

where k is the wavenumber and $t(x_1, y_1)$ denotes the thickness profile. Removing unnecessary parameters and simplifying yields the phase mask design.

Considering the error between the actually recovered PSF and the trained PSF, the manufactured phase mask is added to the telescopic system, the observed image of the distant object is acquired as LR data, and the network is retrained to fine-tune the parameters of the reconstruction network.

3. Unsupervised Image Super-resolution Network

The proposed joint optimization method based on deep learning uses the designed unsupervised network Lightweight Cascaded Residual Network Using Generative Adversarial Network (LCR-GAN) as the foundation for efficient reconstruction of SR images. LCR-GAN offers the advantages of lightweight parameters and fast reconstruction speed, making it easily applicable to real-time tasks on mobile low-cost devices and thus conducive to practical application of the joint optimization method. The LCR-GAN framework is illustrated in Figure 2 [Figure 2: see original paper], consisting of three components: generator, discriminator, and loss function, which are detailed in Sections 3.2 and 3.3.

3.1. Network Architecture

A generative adversarial network obtains realistic images through adversarial learning between a generator that produces images close to real ones and a discriminator that distinguishes false images, representing one of the most promising unsupervised learning methods \cite{Ledig_{2017}}. In this paper, a generative adversarial network model is adopted to build an unsupervised network, as depicted in Figure 2. "Bicubic" denotes the interpolated upsampling operation. The red boxes correspond to the mean absolute error and adversarial loss. To improve the generator's ability to extract detailed features, we add a detail discriminator that enables the network to pay more attention to high-frequency components in the input image, thus making the reconstructed image texture clearer. For this purpose, a Butterworth high-pass filter is introduced before the second discriminator to filter high-frequency components from the input image, which are then distinguished by the second discriminator to identify spurious details generated by the generator.

In LCR-GAN, the generator extracts features at different levels from the input image, fuses them through the upsampling layer, and reconstructs an image of the same size as the ground truth. The discriminator takes the reconstructed image and the real image as input to distinguish fake images generated by the generator. Both discriminators share the same structure.

3.2. Generator and Discriminators

SR Using a Generative Adversarial Network (SRGAN) has been widely investigated for image SR because it can reconstruct fine texture details when the upsampling factor is large \cite{Ledig_{2017}}. To this end, we build a new SR network based on the SRGAN framework, as illustrated in Figure 3 [Figure 3: see original paper], to achieve high-performance reconstruction of SR images. SRGAN is modified as follows: First, the 16-layer residual block in SRGAN is replaced with a 7-layer cascading block (CB). The CB can make full use of feature information extracted at all levels and optimize and extend the information propagation path, thereby reducing network parameters and improving model reconstruction efficiency. Second, the activation function PReLU is replaced with GELU, which has smoother nonlinear characteristics that can improve model performance and accelerate convergence.

As shown in Figure 3, the network is divided into three parts: SFE Module (Shallow Feature Extraction), DFE Module (Deep Feature Extraction), and Irec Module (Image Reconstruction). K denotes the number of CBs in the network, and PW denotes pointwise convolution. After the LR image is input into the network, deep feature information is extracted using seven CBs. Then the image reconstruction module fuses all extracted features to reconstruct the SR image.

The structure of the CB is depicted in Figure 3. The residual block (ResBlock) is responsible for feature map extraction, while the concatenation operator collects feature information extracted at different levels. The PW layer fuses the feature maps extracted from previous layers and compresses the number of feature channels, thereby reducing the training parameters of the residual block. The residual block first learns feature information on $C \times H \times W$ feature maps. It is then cascaded with feature graph F_1 to obtain feature image F_2 with size $2C \times H \times W$. After two pointwise convolution layers compress the feature channel number, the residual feature graph F_1 with input information is obtained. To learn deeper features, two tandem residual units are used to form a residual block, as shown in the lower right of Figure 3.

Considering that PatchGAN trains the discriminator by cutting the input image into small patches for separate discrimination, which enhances local texture details \cite{Phillip_{2017}}, we use PatchGAN as the discriminator, with both discriminators sharing the same architecture. The network model is visualized in Figure 4 [Figure 4: see original paper]. BN represents the batch normalization layer. k, n, and s denote the filter kernel size, number of output feature channels, and convolution layer stride, respectively. Each dark blue box integrates a convolution layer, BN layer, and LeakyReLU activation function. As demonstrated in Figure 4, the input passes through each convolution layer sequentially and is mapped into an $N \times N$ matrix, where each point represents a patch in the input image. Finally, a sigmoid activation function is used to discriminate the probability that different patches are real.

3.3. Loss Function

To accurately reconstruct SR images, we use the following loss function to train the generator:

$$\mathrm{Loss}_G = L_p + \lambda_1 L_G^1 + \lambda_2 L_G^2$$

where Loss_G is a weighted combination of pixel loss L_p, the generative adversarial loss L_G^1 of the first discriminator, and the generative adversarial loss L_G^2 of the second discriminator. Here λ_1 and λ_2 are the weights of the two losses, respectively.

The expression for pixel loss L_p is:

$$L_p = \|I_{SR} - B_{LR}\|_1$$

where I_{SR} represents the network output image and B_{LR} represents the bicubic upsampling result of the input LR image.

The expression for the generative adversarial loss L_G^1 of the first discriminator is:

$$L_G^1 = -\log D_1(I_{SR})$$

where D_1 represents the discriminator output.

The expression for the generative adversarial loss L_G^2 of the second discriminator is:

$$L_G^2 = -\log D_2(W_2 * I_{SR})$$

where D_2 represents the second discriminator output and W_2 represents the Butterworth high-pass filter convolution kernel.

To better distinguish between the generator's reconstructed image and HR images (label data), the first discriminator is trained using:

$$\text{Loss}_{D_1} = -\log D_1(I_{HR}) - \log(1 - D_1(I_{SR}))$$

where I_{HR} represents the HR images (label data).

To better distinguish spurious detail features reconstructed by the generator, the second discriminator is trained using:

$$\text{Loss}_{D_2} = -\log D_2(W_2 * I_{HR}) - \log(1 - D_2(W_2 * I_{SR}))$$

4. Simulation of Joint Optimization Method

To verify the performance of the joint optimization method, we performed SR reconstruction of degraded images using both deep learning and joint optimization approaches and evaluated image quality. Specifically, we first created corresponding star atlas datasets using two methods and used these datasets to train two state-of-the-art unsupervised image SR networks—Direct Unsupervised Super-Resolution Using Generative Adversarial Network (DUS-GAN; \cite{Prajapati_{2021}}), Metric Learning based Interactive Modulation for Real-World Super-Resolution (MM-realSR; \cite{Mou_{2022}}), LCR-GAN, and LCR-GAN with the joint optimization method—until the networks converged. Then, degraded star atlases from the test set were input into the trained networks to reconstruct deep learning SR images and jointly optimized SR images. Image quality evaluation metrics including PSNR, Structural Similarity Index Metric (SSIM), and Learned Perceptual Image Patch Similarity (LPIPS) were used to assess reconstruction quality.

4.1. Data Preparation

The actual astronomical telescopic system was used for simulation testing, which is utilized to identify and track space point targets. Its optical path diagram is shown in Figure 5 [Figure 5: see original paper]. The entrance pupil diameter is 250 mm, the focal length is 2990 mm, and the wavelength is 0.623 m. When the object height is set to 0.1°, the image surface size is 5.23 mm × 5.23 mm. With a pixel size of 192 × 192, the spacing between adjacent pixels is about 27.24 m. We selected 300 clear and easily distinguishable HR star atlases from the NASA website. These atlases are images of celestial objects captured by different astronomical telescopes, ensuring data authenticity. The data are publicly available and can be provided as an attachment. We converted them to grayscale images as the original HR images.

Second, we introduce two methods for generating degraded or LR images for deep learning and joint optimization methods. Traditional deep learning reconstruction methods require LR star atlases. Therefore, we constructed a 1:1 imaging optical path of the actual astronomical telescopic system in optical design software, set the corresponding basic parameters according to the actual imaging scene, and used the image simulation function to obtain the image plane image of the telescopic system. For the joint optimization method proposed in this paper, as described in Section 2.1, a learnable single-kernel convolution layer (convolution kernel size 23 × 23, stride 1) is used as the PSF of the actual telescopic system and convolved with the HR image to generate an LR image that matches the incoherent imaging process of the telescopic system.

Images obtained by both methods are downsampled 4× as the LR image to match potential low sampling rate detectors. Given the large object distance of the telescope system and the difficulty of reproducing astronomical stars in reality, we weighed data fidelity against production cost and used both methods to generate the degraded data required for this study.

According to the deep learning degraded data production method described above, 300 pairs of LR-HR star atlases can be easily obtained. To gather sufficient training data, we randomly rotated by 90° and flipped for data augmentation. From the obtained 1200 pairs of data, 960 pairs were extracted for the training set, 120 pairs for the validation set, and 120 pairs for the test set. We then randomly cropped 48 × 48 and 192 × 192 unpaired patches from the training dataset for training the SR networks DUS-GAN, MM-realSR, and LCR-GAN.

According to the joint optimization method's degraded data production method described above, the initial HR star atlas was randomly rotated by 90° and flipped for data augmentation. The 1200 obtained images were randomly cropped into 192 × 192 unpaired patches, with 960 used for HR training set data, 120 for HR validation set data, and 120 for HR test set data. The HR images in the training set are convolved with a learnable single-kernel convolution layer, and the resulting images are downsampled 4× to generate

LR images for training the LCR-GAN image reconstruction network of the joint optimization method. As mentioned in Section 2.2, after training, the PSF recovered by the GS algorithm is taken as the PSF of the actual telescopic system. The HR images in the test set are convolved with this PSF, and LR images of the test set are generated by 4× downsampling.

4.2. Simulation Result Analysis

Figure 6 [Figure 6: see original paper] displays the simulation results of the GS algorithm implementation described in Section 2.2. The average absolute error between the PSF generated by the algorithm and the PSF trained by the network converges quickly. The final Root Mean Square Error between the two is 0.0189, and the sum of squares due to error (SSE) is 0.0097, indicating that the modulation phase obtained by the algorithm can accurately realize the PSF trained by the network.

In the training phase, 960 sets of two kinds of unpaired images were used as input for their respective unsupervised networks, with 300,000 iterations performed. The learning rate is 0.0001, batch size is 1, and epoch is 300. Our implementation uses the PyTorch framework on a PC with a GeForce RTX 4090 (NVIDIA).

After training, 60 sets of data were randomly selected from both the deep learning and joint optimization test sets. The LR images were input into the corresponding networks to reconstruct the SR star atlas. Three quantitative image quality indices were used for performance evaluation: PSNR, SSIM, and LPIPS.

Table 1 lists the average metrics of reconstructed images by different methods. We observe that the PSNR and SSIM values of LCR-GAN are better than MM-RealSR and comparable to bicubic upsampling and DUS-GAN, indicating that LCR-GAN reconstructed images have less distortion, better structural similarity, and higher image quality. The image reconstruction time of LCR-GAN is much smaller than that of the other two networks, and the ultra-fast reconstruction speed is conducive to real-time reconstruction. Deep learning methods and bicubic upsampling methods are significantly lower than joint optimization methods in PSNR, SSIM, and LPIPS, indicating that SR images reconstructed by the joint optimization method are closer to real images.

We compare the training parameters of several main unsupervised networks, as shown in Table 2 . The parameters of LCR-GAN are smaller than those of other unsupervised networks. The simplified model not only improves image reconstruction speed but can also be deployed on low-cost computing devices, facilitating practical deep learning applications.

Figure 7 [Figure 7: see original paper] shows the effect of the 4× super-resolved star atlas reconstructed by different methods. According to the Rayleigh criterion, the diffraction limit of the system is about 9.12 m, much smaller than the pixel spacing of 27.24 m, meaning the system can theoretically distinguish adjacent pixels. In Figure 7, the image plane image was downsampled 4× (matching the low sampling detector) to generate LR images. Due to system aberrations and other factors, as well as detail information loss caused by 4× downsampling, adjacent weak star points in LR images are difficult to distinguish. The adjacent weak stars in LCR-GAN reconstructed images are easily distinguishable, indicating good SR performance. Compared with DUS-GAN and MM-realSR, the SR star atlas reconstructed by LCR-GAN has relatively clear star points and textures, with weak star points well preserved.

Compared with LCR-GAN, the star atlas reconstructed by the joint optimization method is richer in detail, with clearer nebular texture, and adjacent faint stars are distinguishable. We attribute this to the joint optimization method reconstructing the PSF of the telescopic system, enabling the system to retain optimal information and thus producing reconstruction results closer to the original HR image.

5.1. Experimental Setup

Using deep learning, Sun's method \cite{Sun_{2020}}, and the joint optimization method proposed in this paper, we conducted SR reconstruction experiments on a resolution plate to verify the performance of the joint optimization method. The optical path for the validation experiment is illustrated in Figure 8 [Figure 8: see original paper]. A light-emitting diode emits scattered light with a central wavelength of 623 nm, which is filtered to 623 nm monochromatic light by a narrow band filter. The beam intensity is adjusted using an attenuator plate, and the beam forms many scattered point sources through the aperture of the resolution plate, simulating light emitted by star points.

A single lens is placed at a focal length from the resolution plate to correct for infinite distance. The focal length of the lens is 100 mm. A beam splitter prism splits the light into two beams, with one transmitted to the surface of the spatial light modulator and reflected back to the prism after phase modulation. The prism then deflects the reflected light perpendicular to the incident light. The light is imaged by the DGTO in the back focal plane and recorded by a low-illumination camera. The focal length of the double-glued objective is 200 mm. The detector pixel size is 3.45 m. The spatial light modulator dimensions are approximately 13 mm × 8 mm.

In terms of waveform modulation, the telescopic system is equivalent to a lens with an arbitrarily modulated wavefront shape. In this paper, a non-ideal lens—a double-glued objective—is used instead of the telescopic system to verify the joint optimization method. As shown in Figure 8, since the light incident on the DGTO is a plane wave and the double-glued lens thickness is small, the spatial light modulator can be placed in front of the double-glued objective. Since the fringe width of the resolution plate is less than 30 m, to better characterize the performance of the joint optimization method, the resolution plate is used instead of the star atlas in the experiment. Additionally, as described in Section

4.1, we added two images of resolution plates to the original images and retrained the corresponding networks to help reconstruct the SR image of the resolution plate.

To verify the performance of the joint optimization method, we generated a modulation phase mask according to the method described in Section 2. Based on the joint optimization method using an ideal lens proposed by Sun Q., we solved the corresponding phase mask. The spatial light modulator modulates the phase of the light field according to the grayscale image of the phase mask. In both off and on states of the spatial light modulator, the detector collects images without phase mask, with Sun Q. method phase mask, and with our joint optimization method phase mask, as inputs for their respective networks. The SR images corresponding to deep learning, Sun's method, and our joint optimization method are all reconstructed.

5.2. Experimental Result Analysis

The experimental results are displayed in Figure 9 [Figure 9: see original paper]. Panel (a) shows the image of the standard resolution plate actually captured by the detector, with the red box indicating the reference position of the control group. Panels (b), (c), and (d) show images without phase mask, with Sun method phase mask, and with our joint optimization method phase mask, respectively. Panels (e), (f), and (g) show SR images reconstructed by LCR-GAN, Sun's method, and our joint optimization method, respectively.

According to the Rayleigh criterion, the lateral diffraction limit of the system is about 11.69 m, and the longitudinal diffraction limit is about 19 m. The fringe width is about 8.3 pixels, or 28.64 m, which is larger than the diffraction limit, meaning the system can theoretically distinguish between light and dark fringes. Comparing panels (b) and (e), due to system aberrations, noise, and other factors, it is difficult to distinguish light and dark fringes in images captured by the detector, though LCR-GAN demonstrates certain SR effects.

Comparing panels (d) and (g), the SR image reconstructed by our joint optimization method is rich in detail, with fringes clearly distinguishable, greatly improving the resolution of the observation image. Comparing panels (f) and (g), the SR effect of the image reconstructed by Sun's method is lower than that of the joint optimization method. We attribute this to the modulation phase of the double-glued objective lens being different from that of an ideal lens, causing the phase mask solved by Sun's method to have errors that result in loss of some high-frequency information in the image plane. These results verify the simulation findings.

To quantitatively measure the SR effect, we took tangents to the stripes in the reconstructed image at the position corresponding to the black line in Figure 9(a). The pixel position is taken as the X-axis, and the gray value along the tangent line is plotted. The legend in the upper right corner of Figure 10 [Figure 10: see original paper] shows curves corresponding to different reconstruction methods. For comparison, we performed 4× bicubic upsampling on the real image to match the upsampling ratio of other SR reconstruction methods. Since the stripes are equally spaced, we measured fringe contrast by calculating the average absolute error between seven pairs of clear crests and troughs in Figure 10. The mean absolute errors for bicubic upsampling, LCR-GAN, Sun's method, and our joint optimization method are 17, 17.7, 68.7, and 71.4, respectively. The joint optimization method achieves the highest contrast, with stripes in the reconstructed image being easier to distinguish.

6. Conclusions

In this paper, we propose an optical-neural network joint optimization method for telescopic systems. First, a deep learning convolutional layer equivalent to the telescopic system PSF is integrated into the front end of the neural network, and the PSF and network parameters are jointly optimized through training. After training, a phase mask is constructed to physically implement the learned PSF in the telescopic system; the manufactured phase mask is added to the system to collect observation images, and the network is retrained to fine-tune parameters. This paper also constructs an unsupervised network LCR-GAN, which offers lightweight parameters and fast reconstruction speed, making it suitable for real-time tasks on mobile computing devices.

Simulations and experiments show that the reconstructed star atlas is rich in detail and easy to distinguish, and the resolution of observed images can be improved without adding complex hardware equipment. It is worth noting that since the network in this paper is trained using a dataset composed of known star atlases, it can only be applied to observation images of known targets with basic outlines or shapes.

Acknowledgments

Funding is provided by the National Natural Science Foundation of China (NSFC, Grant Nos. 62375027 and 62127813), Natural Science Foundation of Chongqing Municipality (CSTB2023NSCQ-MSX0504), Natural Science Foundation of Jilin Province (YDZJ202201ZYTS411), and Jilin Provincial Education Department Fund of China (JJKH20240920KJ).

References

Chang, J., Sitzmann, Q., Dun, X., et al. 2018, NatSR, 8, 12324
Chang, J.-L., & Wetzstein, G. 2019, in IEEE/CVF Int. Conf. on Computer Vision (ICCV) (Piscataway, NJ: IEEE), 10192
Gerchberg, R. W., & Saxton, W. O. 1972, Optik, 35, 237
Goodman, J. W. 2005, Introduction to Fourier Optics (3rd ed.; Englewood, CO: Roberts & Co. Publishers)
He, Z.-Q., Yin, K., Wu, S.-T., et al. 2021, LSA, 134, 10
Irani, M., & Park, S. 1991, ITIP, 53, 231

Lehmann, T. M., Park, M. K., Kang, M. G., et al. 1999, IEEE Trans. Med. Imaging, 18, 1049
Ledig, C., Theis, L., Huszar, F., et al. 2017, in 2017 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR) (Piscataway, NJ: IEEE), 105
Li, Z., Peng, Q.-Y., Bhanu, B., et al. 2018, Ap&SS, 363, 92
Mikhail, K., Ma, Y.-Q., Chen, Y.-B., et al. 2019, LSA, 118, 8
Mou, C., Wu, Y.-Z., Wang, X.-T., et al. 2022, in Computer Vision – ECCV 2022, ed. S. Avidan, G. Brostow, M. Cissé, G. M. Farinella, & T. Hassner (Cham: Springer), 723
Park, S. C., Park, M. K., Kang, M. G., et al. 2003, IEEE Signal Process Mag., 20, 21
Peng, Y.-F., Fu, Q., Amata, H., et al. 2015, OExpr, 23, 31393
Peng, Y.-F., Sun, Q.-L., Dun, X., et al. 2019, ACM Trans. Graph, 38, 219:1
Phillip, I., Zhu, J.-Y., Zhou, T.-H., et al. 2017, in 2017 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR) (Piscataway, NJ: IEEE), 5967
Prajapati, K., Chudasama, V., Patel, H., et al. 2021, ITIP, 30, 8251
Rahman, S., Moon, Y. J., Park, E., et al. 2020, ApJL, 897, L32
Shoubaneh, H., Eric, H., Hooshang, N., et al. 2022, ApJ, 941, 141
Sweere, S. F., Valtchanov, I., Lieu, M., et al. 2022, MNRAS, 517,
Sun, Q.-L., Zhang, J., Dun, X., et al. 2020, ACM Trans. Graph, 39, 9:1
Wang, C., Shi, H.-D., Liu, Z., et al. 2020, ApOpt, 59, 7883
Wang, C., Xing, S.-Y., Xu, M., et al. 2022, Senso, 22, 2717
Wang, C.-T., Tang, D.-L., Wang, Y., et al. 2015, NatSR, 5, 18485
Wang, K.-K., Fu, Q., Shi, H.-D., et al. 2023a, InPhT, 2023, 105032
Wang, K.-K., Fu, Q., Wang, B., et al. 2023b, OExpr, 31, 30750
Wang, S., Jiang, X.-N., Guo, H.-J., et al. 2023c, OptCo, 545, 129634
Xu, M., Wang, C., Wang, K.-K., et al. 2022, Senso, 22, 9676
Yan, Q., Xu, Y., Yang, X., et al. 2015, ITIP, 24, 3187
Yang, M.-Y., Fan, X.-W., Wang, Y.-M., et al. 2022, RemS, 14, 2264
Youla, D. C., & Webb, H. 1982, IEEE Trans. Med. Imaging, 1, 81
Yuan, R., Zhao, C.-Z., Guo, Y., et al. 2020, LOP, 57, 232201

Note: Figure translations are in progress. See original paper for figures.

Source: ChinaXiv — Machine translation. Verify with original.