



## 一种基于图像融合和卷积神经网络的相位恢复方法

周静 张晓芳 赵延庚

### Phase retrieval wavefront sensing based on image fusion and convolutional neural network

Zhou Jing Zhang Xiao-Fang Zhao Yan-Geng

引用信息 Citation: *Acta Physica Sinica*, 70, 054201 (2021) DOI: 10.7498/aps.70.20201362

在线阅读 View online: <https://doi.org/10.7498/aps.70.20201362>

当期内容 View table of contents: <http://wulixb.iphy.ac.cn>

---

## 您可能感兴趣的其他文章

### Articles you may be interested in

基于遗传算法优化卷积长短记忆混合神经网络模型的光伏发电功率预测

A hybrid model for photovoltaic power prediction of both convolutional and long short-term memory neural networks optimized by genetic algorithm

物理学报. 2020, 69(10): 100701 <https://doi.org/10.7498/aps.69.20191935>

基于深度卷积神经网络的大气湍流相位提取

Extracting atmospheric turbulence phase using deep convolutional neural network

物理学报. 2020, 69(1): 014209 <https://doi.org/10.7498/aps.69.20190982>

一种基于双光束干涉和非线性相关的身份认证方法

Identity authentication based on two-beam interference and nonlinear correlation

物理学报. 2017, 66(4): 044202 <https://doi.org/10.7498/aps.66.044202>

基于单一分光棱镜干涉仪的双通路定量相位显微术

Dual-channel quantitative phase microscopy based on a single cube beamsplitter interferometer

物理学报. 2018, 67(14): 140704 <https://doi.org/10.7498/aps.67.20172722>

基于迭代重建算法的X射线光栅相位CT成像

Grating based X-ray phase contrast CT imaging with iterative reconstruction algorithm

物理学报. 2017, 66(5): 054202 <https://doi.org/10.7498/aps.66.054202>

基于相干调制成像的光学检测技术

Technique of detecting optical components based on coherent modulation imaging

物理学报. 2020, 69(17): 174202 <https://doi.org/10.7498/aps.69.20200224>

# 一种基于图像融合和卷积神经网络的相位恢复方法\*

周静 张晓芳† 赵延庚

(北京理工大学光电学院, 北京 100081)

(2020年8月18日收到; 2020年9月23日收到修改稿)

相位恢复法利用光波传输中某一(或某些)截面上的光强分布来传感系统波前,其结构简单,不易受震动及环境干扰,被广泛应用于光学遥感和像差检测等领域.传统相位恢复法采用迭代计算,很难满足实时性要求,且在一定程度上依赖于迭代转换或迭代优化初值.为克服上述问题,本文提出了一种基于卷积神经网络的相位恢复方法,该方法采用基于小波变换的图像融合技术对焦面和离焦面图像进行融合处理,可在不损失图像信息的同时简化卷积神经网络的输入.网络模型训练完成后可依据输入的融合图像直接输出表征波前相位的4—9阶Zernike系数,且波前传感精度均方根(root-mean-square, RMS)可达 $0.015\lambda$ ,  $\lambda = 632.8\text{ nm}$ .研究了噪声、离焦量误差和图像采样分辨率等因素对波前传感精度的影响,验证了该方法对噪声具有一定鲁棒性,相对离焦量误差在7.5%内时,波前传感精度RMS仍可达 $0.05\lambda$ ,且随着图像采样分辨率的提升,波前传感精度有所改善,但训练时间成本随之增加.此外,分析了实际应用中,当系统像差阶数与网络训练阶数略有差异时,本方法所能实现的传感精度,并给出了解决方案.

**关键词:** 相位恢复, 卷积神经网络, 图像融合, 传感精度

**PACS:** 42.15.Fr, 42.30.Kq, 42.30.Rx

**DOI:** 10.7498/aps.70.20201362

## 1 引言

空间遥感系统在轨工作时,由于受环境及系统自身像差的影响,需实时对系统波前进行监测以实现高分辨成像.在诸多波前传感方法中,相位恢复(phase retrieval, PR)法具有传感精度高、系统结构简单、不易受震动和环境干扰等优势,特别适宜空间应用<sup>[1-4]</sup>.近年来,已成为空间遥感系统波前传感的首选方案.

1972年Gerhberg和Saxton<sup>[5]</sup>提出G-S迭代算法,奠定了相位恢复方法的核心理论.随后出现了很多改进算法,如误差减少算法(ER)<sup>[6]</sup>、混合输入输出算法(HIO)<sup>[6]</sup>及杨-顾算法(Y-G)<sup>[7]</sup>等.但

这些迭代方法耗时长,难以满足实时性需求,且在一定程度上依赖于迭代转换或迭代优化过程中用到的迭代初值.

为克服传统迭代方法存在的问题,近年来许多学者将人工神经网络与波前传感结合<sup>[8-12]</sup>,提出了基于深度学习的波前传感新方法.2018年,Paine和Fienup<sup>[8]</sup>考虑到PR法对迭代初值的依赖性,提出利用卷积神经网络(convolutional neural-network, CNN)对迭代初值进行预测,提升了波前传感精度.但该方法应用神经网络后,仍要进行迭代计算,并未完全克服传统迭代方法的耗时问题.同年鞠国浩等<sup>[11]</sup>通过对焦面和离焦面图像提取切比雪夫矩特征,结合反向传播(back-propagation, BP)网络实现相位恢复,但切比雪夫矩特征提取方法

\* 国家自然科学基金(批准号: 61471039)资助的课题.

† 通信作者. E-mail: zhangxf@bit.edu.cn

复杂, 编程困难. 2019 年齐鑫等<sup>[12]</sup>通过对焦面和离焦面图像提取一种与目标无关的特征, 并结合长短期记忆 (long short-term memory, LSTM) 网络实现了相位变更波前传感方法, 但其用到的网络层数为 128 层, 结构复杂.

为提高波前传感精度, 并降低系统复杂度, 本文提出了一种基于图像融合和卷积神经网络的相位恢复方法, 该方法采用基于小波变换的图像融合技术对焦面和离焦面的点扩散函数 (point spread function, PSF) 图像进行融合处理, 可在不损失图像信息的同时简化 CNN 的输入. 训练好的网络模型, 可依据输入的融合图像直接输出表征波前相位的 Zernike 系数, 有效地提升波前传感精度的同时

简化了操作难度.

## 2 相位恢复法原理模型

PR 是一种间接测量波前方法, 即利用可直接测量的强度分布来恢复波前相位<sup>[13]</sup>. 由于传统 G-S 算法仅从一幅焦面图像恢复波前, 无法确保解的唯一性, 故本文采用引入了离焦相位变更的相位变更相位恢复 (phase-diversity phase retrieval, PDPR) 法<sup>[1,14]</sup>. 通过引入已知的离焦相位变更, 消除了光瞳面相位分布的不确定性. 图 1 为典型的离焦型 PDPR 法工作原理.

依据标量衍射理论, 焦面光场分布  $U(x, y)$  可表示为

$$U(x, y) = \frac{\exp(jkf)}{j\lambda f} \exp\left[j\frac{k}{2f}(x^2 + y^2)\right] \iint_{(-\infty, +\infty)} U_0(x_0, y_0) \exp\left[-j \cdot 2\pi \cdot \frac{(x_0x + y_0y)}{\lambda f}\right] dx_0 dy_0. \quad (1)$$

同理, 离焦面光场分布  $U'(x, y)$  可表示为

$$U'(x, y) = \frac{\exp[jk(f+d)]}{j\lambda(f+d)} \exp\left[j\frac{k}{2(f+d)}(x^2 + y^2)\right] \iint_{(-\infty, +\infty)} U_0(x_0, y_0) \cdot \exp\left[-j \cdot k \cdot d \frac{(x_0^2 + y_0^2)}{2(f+d) \cdot f}\right] \times \exp\left[-j \cdot 2\pi \cdot \frac{(x_0x + y_0y)}{\lambda(f+d)}\right] dx_0 dy_0, \quad (2)$$

其中  $\lambda$  为光波波长;  $f$  为焦距;  $d$  为已知离焦距离;  $j^2 = -1$ ;  $k$  为波矢数. 显然, 焦面及离焦面光场均可看作广义光瞳函数的傅里叶变换, 所不同的是, 对于离焦面光场, 光瞳函数的相位因子中引入了一离焦相位. 本文选用 Zernike 多项式表征相位  $\varphi(x, y)$ , 如 (3) 式所示:

$$\varphi(x, y) = \sum_{i=1}^N \alpha_i \cdot Z_i(x, y), \quad (3)$$

其中  $i$  为阶数;  $Z_i(x, y)$  为第  $i$  阶 Zernike 多项式的表达式;  $\alpha_i$  为第  $i$  阶多项式的系数. 可以看出, 焦面、

离焦面上的光场分布与系统波前存在非线性关系, 故可利用 CNN 建立两者联系, 由已知的光强信息反演出波前相位.

## 3 基于 CNN 模型的相位恢复方法

如前所述, 应用于波前传感的神经网络模型主要有 BP, CNN, LSTM 等. 由于相位恢复法是基于图像信息来实现波前传感, 故本文采用更擅长图片处理的 CNN 进行相位恢复训练<sup>[15-17]</sup>.

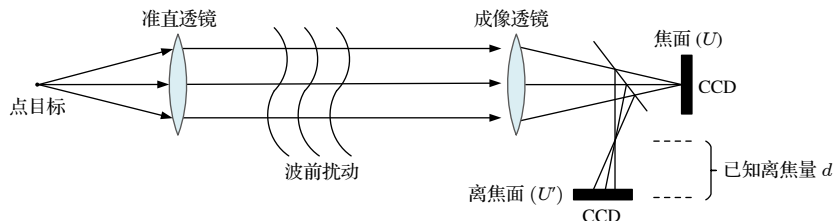


图 1 离焦型 PDPR 法工作原理图

Fig. 1. Schematic diagram of PDPR (defocus diversity).

本文提出了利用 CNN 拟合焦面、离焦面 PSF 融合图像与系统波前的非线性关系, 实现由已知的光强信息预测波前的传感方法. 其中, CNN 的输入为焦面、离焦面 PSF 图像经小波变换图像融合方法所得的融合图像, 输出为采用 Zernike 多项式表征的系统波前信息 (即各阶 Zernike 系数值). 由于 Zernike 多项式前三项分别表示平移,  $x, y$  方向的倾斜对图像像质没有影响, 故在相位恢复过程中不予考虑. 同时, 针对空间遥感系统以低空间频率像差为主的特点, 本文选择了 4—9 阶 Zernike 多项式表征系统波前, 分别对应离焦、与轴成  $0^\circ$  或  $90^\circ$  的像散、与轴成  $45^\circ$  的像散、 $X$  轴的三级彗差、 $Y$  轴的三级彗差和三级球差<sup>[18]</sup>.

### 3.1 CNN 结构

CNN 结构有很多变体, 但其基本结构大多包括输入层、卷积层、池化层、全连接层及输出层<sup>[19]</sup>. 笔者基于这些基本结构进行多次搭建尝试, 依据训练结果的不同表现, 对网络结构及参数进行调整, 最终确定了如图 2 所示的网络结构. 该结构共包含 22 层, 包括 1 层输入层, 13 层卷积层, 6 层池化层, 1 层平坦层和 1 层全连接层 (即输出层). 处理后的焦面和离焦面融合图像作为网络的输入, 经过卷积核大小为  $3 \times 3$ , 步长为 1 的第一层卷积层后得到 32 个大小为  $32 \times 32$  的特征图; 接着进入池化层, 采用最大池化, 池化核尺寸为  $2 \times 2$ , 输出特征图尺寸减半, 即此时特征图大小为  $16 \times 16$ , 特征图数量仍为 32 个; 之后经过 5 次相同的卷积层和池化层的交替结构后, 得到 64 个大小为  $1 \times 1$  的特征图; 再经过 7 层卷积核大小为  $3 \times 3$ , 步长为 1 的卷积层, 1 层平坦层及 1 层包含 6 个神经元的全连接层, 最终输出长度为 6 的一维数组, 对应 4—9 阶 Zernike 多项式系数. 训练过程中应用的学习率为 0.0001, 损失函数为均方误差 (mean square

error, MSE), 表达式为

$$MSE = \frac{1}{N} \sum_{i=1}^N (\alpha_{\text{true}} - \alpha_{\text{est}})^2, \quad (4)$$

其中  $\alpha_{\text{true}}$  为 Zernike 多项式系数真值;  $\alpha_{\text{est}}$  为多项式系数网络预测值; MSE 的值越小, 说明预测模型精确度越好. 应用的优化函数为 Adam, 该优化算法是一种对随机梯度下降法的扩展, 结合了自适应梯度算法 (AdaGrad) 和均方根传播算法 (RMSProp) 的优点<sup>[20]</sup>. 应用的激活函数为 ReLU (rectified linear unit), 表达式为

$$y = \begin{cases} x, & (x > 0), \\ 0, & (x \leq 0). \end{cases} \quad (5)$$

将 (5) 式中  $y$  对  $x$  求导, 则得到

$$\frac{\partial y}{\partial x} = \begin{cases} 1, & (x > 0), \\ 0, & (x \leq 0). \end{cases} \quad (6)$$

可见数据在网络内传播过程中, 正向输入  $x > 0$  时, 反向传播会将上游的值原封不动传给下游; 正向输入  $x \leq 0$  时, 反向传播中传给下游的信号将停在此处. 因此 ReLU 函数会使一部分神经元输出为 0, 致网络稀疏, 从而减少参数之间的依存性, 防止过拟合, 且与其他激活函数相比不会出现梯度消失问题.

### 3.2 基于小波变换图像融合方法生成 CNN 输入数据

为简化 CNN 输入, 降低系统复杂度, 本文采用基于小波变换的极值法对焦面、离焦面 PSF 图像进行融合. 其核心思想为挑选待融合图像中相应各像素点中像素值的极大或极小值, 作为融合图像在该点的像素值  $F(i, j)$ <sup>[21]</sup>, 如 (7) 式、(8) 式所示, 分别称为极大值法和极小值法:

$$F(i, j) = \max\{A(i, j), B(i, j)\}, \quad (7)$$

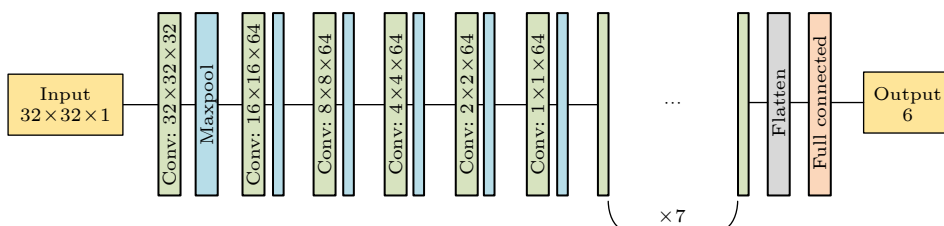


图 2 CNN 结构示意图

Fig. 2. The schematic of CNN structure.

$$F(i, j) = \min\{A(i, j), B(i, j)\}, \quad (8)$$

其中  $A(i, j)$  与  $B(i, j)$  分别表示待融合图像 A 和 B 在  $(i, j)$  点的像素值.

基于小波变换机理的图像融合原理如图 3 所示. 首先, 采用长度为 4 的多贝西小波对焦面、离焦面 PSF 图像进行 1 层小波分解 [22], 即采用低通 (L)、高通 (H) 滤波器分别对两幅源图像的水平、垂直方向进行滤波, 得到一个低频带 (水平低频垂直低频 LL) 和三个高频带 (水平低频垂直高频 LH、水平高频垂直低频 HL 及水平高频垂直高频 HH) 子图像, 且能量主要集中在低频部分 [23]; 接着, 对高频带和低频带图像分别采用极大值法和极小值法, 将不同频率信息分别进行融合, 既可不损失图像信息, 又可将焦面、离焦面 PSF 图像细节信息融合, 减少冗余数据; 最后, 采用第一步小波变换的重构算法对处理后的小波系数进行反变换, 实现图像重构 [24,25].

该融合方法计算简单、适宜实时处理. 本质上, 融合图像的过程可看作是数据压缩过程, 如图 4 所示, 在本文中, 两幅  $64 \times 64$  的焦面、离焦面图像 (图 4(a) 和图 4(b)), 经小波变换图像融合处理后合成了一幅  $64 \times 64$  的融合图像 (图 4(c)), 同时, 为了最大化排除冗余信息, 仅将融合图像中心  $32 \times 32$  的部分 (图 4(d)) 作为 CNN 的输入. 相比于将全部图像信息作为输入的传统方法, 采用图像融合技术可有效地减少冗余数据, 加快网络训练速度. 仿真结果表明, 若图像采样率为  $64 \times 64$ , 以融合图像作为网络的输入数据, 训练时间会缩短 1/2, 且随着图像采样率的增加时间优势愈发显著.

#### 4 仿真实验与精度分析

建立了网络结构及输入数据生成、处理方式后, 本文开展了大量的仿真实验, 对所提方法的波前传感精度及其影响因素进行了分析.

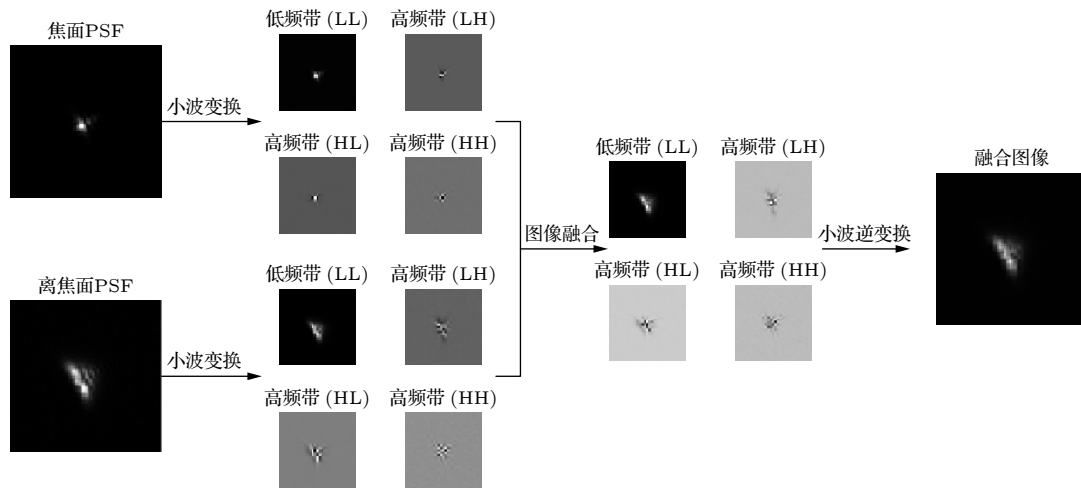


图 3 基于小波变换机理的图像融合原理

Fig. 3. Principle of image fusion based on wavelet transform.

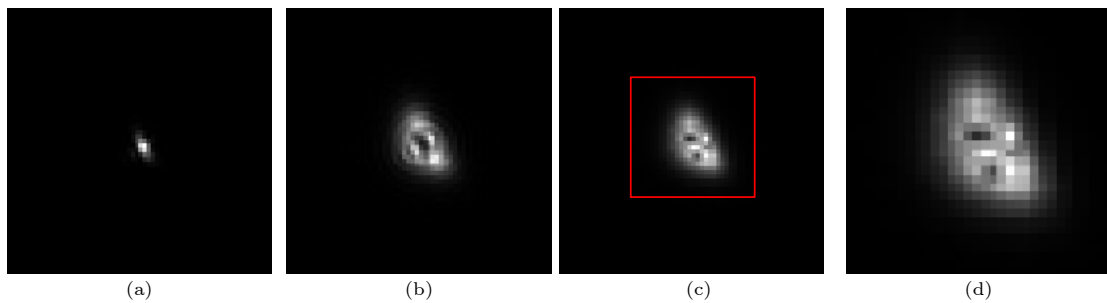


图 4 融合图像示意图 (a) 焦面 PSF 图像; (b) 离焦面 PSF 图像; (c) 融合图像; (d) CNN 的输入图像 (即图 (c) 中红框内部分)  
Fig. 4. Schematic diagram of fusion image: (a) PSF image of the focal plane; (b) PSF image of the defocal plane; (c) the fusion image; (d) the input image of CNN.



## 4.1 CNN 训练

首先确定系统参数,生成训练数据.本文仿真过程中使用单透镜成像系统,具体系统参数如表 1 所列.

在系统中引入 10000 组由 4—9 阶 Zernike 多项式表征的随机波前,其均方根 (root-mean-square, RMS) 值分布如图 5 所示.可以看出,波前 RMS 值在  $[0.1\lambda, 1.1\lambda]$  范围内,且近似符合高斯分布.

表 1 仿真系统参数

Table 1. Simulation system parameters.

透镜焦距/mm	入瞳直径/mm	波长/nm	离焦距离/mm
150	10	632.8	4

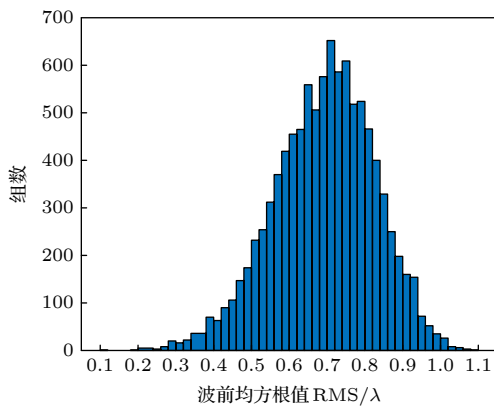


图 5 10000 组随机波前 RMS 分布图

Fig. 5. RMS distributions of 10000 random wavefront groups.

依据傅里叶光学原理生成相应的 10000 对焦点、离焦点 PSF 图,并在图中引入信噪比为 35 dB 的噪声来模拟实际噪声环境.然后利用小波变换图像融合方法将 10000 对 PSF 图像融合,得到 10000 张融合图像,作为网络训练的输入数据.而输出数据则为表征波前畸变的 10000 组对应的 4—9 阶 Zernike 系数.训练过程中将 10000 组训练数据集分成三部分:80% 用作训练集数据,用以拟合网络权值;15% 用作验证集数据,用以调整网络最终架构;5% 用作测试集数据,用以评估网络性能.网络训练环境为 python3.7, tensorflow-gpu2.1.0 及 keras-gpu2.3.1,训练所用计算机处理器为 Intel(R)Core (TM)i7-8750H CPU @2.20 GHz,显卡型号为 NVIDIA GeForce GTX 1060.

## 4.2 训练结果及分析

完成网络训练后,可得到一训练完成的网络

模型.采用 4.1 节所述 5% 的测试集数据对该网络模型进行测试,预测 4—9 阶 Zernike 系数值,实现波前传感.

### 4.2.1 网络预测系数与真实系数比较

500 组测试数据的真实系数与网络预测系数间的标准差统计直方图如图 6 所示,从图中数据可见,系数间的标准差不大于 0.025 的占 90%.故利用本文方法得到的 Zernike 系数与真实系数极为接近.

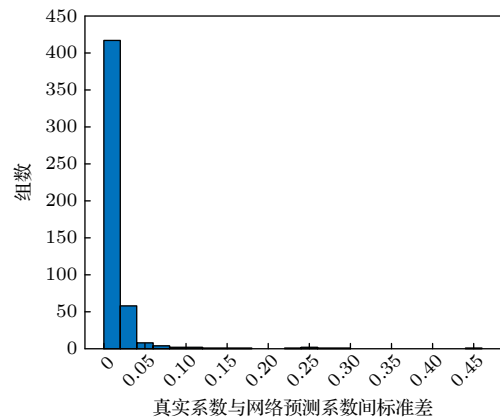


图 6 真实系数与网络预测系数间的标准差统计直方图

Fig. 6. Statistical histogram of standard deviation between the real coefficient and the network prediction coefficient.

### 4.2.2 波前传感精度

根据网络输出的各阶 Zernike 系数值可拟合出系统波前畸变  $\varphi_{est}$ ,将其与引入系统的真实波前畸变  $\varphi_{true}$  做差得到残余波面,并以该波面 RMS 值作为评价波前传感精度的指标.图 7 所示为 500 组测试数据波前传感精度 RMS 统计直方图.可见,

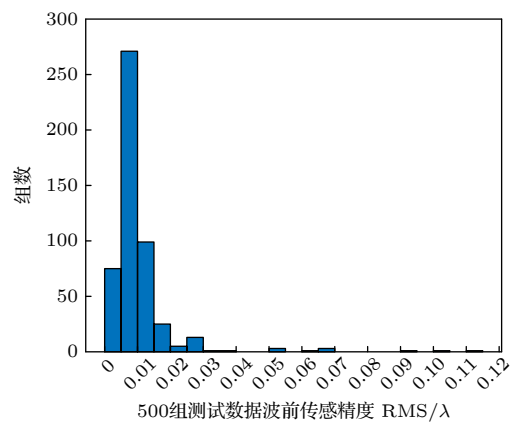


图 7 测试数据波前传感精度 RMS 统计直方图

Fig. 7. RMS statistical histogram of wavefront sensing accuracy of test data.

本文提出的方法传感精度最高可达 RMS  $0.0016\lambda$ , 最低为 RMS  $0.113\lambda$ , 均值为 RMS  $0.0104\lambda$ , 且  $\text{RMS} \leq 0.015\lambda$  占比 90%, 相比于文献 [11] 给出的传感精度  $0.025\lambda$  有所提高.

### 4.3 影响本方法波前传感精度的因素分析

#### 4.3.1 像差阶数的影响

如前文所述, 遥感系统在轨工作时, 系统像差以低空间频率误差为主, 故本文采用了 4—9 阶 Zernike 多项式表征波前畸变, 由于系统实际像差阶数未必局限于 4—9 阶, 考虑到 CNN 网络的普适性, 研究了系统存在略低或略高阶数像差时, 本文方法的波前传感精度.

假设系统实际像差阶数分别为 4—7 阶、4—8 阶、4—10 阶及 4—11 阶. 其中第 4—9 阶像差系数范围仍为  $[-0.5\lambda, 0.5\lambda]$ , 第 10, 11 阶的像差系数范围为  $[-0.1\lambda, 0.1\lambda]$ , 基于表 1 的系统参数, 产生各 500 组验证数据, 然后运用训练好的网络, 预测 4—9 阶 Zernike 系数值, 实现波前传感. 得到的传感精度如表 2 所列. 由表 2 可以看出, 当实际像差阶数小于 9 阶时, 本方法的传感精度 RMS 呈上升趋势; 当实际像差阶数在 10 阶以内时, 传感精度 RMS 仍可控制在  $0.015\lambda$  以内; 当实际像差阶数达到 11 阶时, 传感精度 RMS 降至  $0.1\lambda$ , 通常情况下, 遥感系统的传感精度需达到  $0.05\lambda$ , 故已很难满足遥感系统波前传感要求. 鉴于 CNN 网络的普适性, 此时, 可依据遥感系统像差空间频率特性, 基于本文所提方法增加网络输出层单元数, 并调整网络结构, 训练适用于更高阶像差的网络.

表 2 系统存在不同阶像差时本文方法的波前传感精度  
Table 2. The wavefront sensing accuracy of the proposed method when the system has different order of aberration.

像差阶数	4—7阶	4—8阶	4—10阶	4—11阶
传感精度RMS/ $\lambda$	0.010	0.012	0.015	0.100

#### 4.3.2 离焦量误差的影响

受实际图像采集过程中探测器定位精度的限制, 应用相位恢复法进行波前传感时, 很难保证预设的离焦量被精确引入系统. 为此, 笔者在传感系统中引入了一定量的相对离焦量误差, 研究了该误差对传感精度的影响.

相对离焦量误差  $D$  的定义为

$$D(\%) = \frac{\delta(\text{mm})}{d(\text{mm})} \times 100\%, \quad (9)$$

其中  $\delta$  为离焦量误差;  $d$  为离焦量. 仿真中, 将相对离焦量误差设置为 2.5%, 5.0%, 7.5% 和 10.0%, 应用本文方法进行波前传感, 所得传感精度如表 3 所列.

表 3 不同离焦量误差下, 本文方法的传感精度  
Table 3. The sensing accuracy of the proposed method under different defocusing errors.

相对离焦量误差/%	2.5	5.0	7.5	10.0
传感精度RMS/ $\lambda$	0.022	0.035	0.050	0.065

可以看出, 当相对离焦量误差在 2.5%—7.5% 时, 传感精度 RMS 值介于  $0.022\lambda$ — $0.050\lambda$  之间, 离焦量误差对传感精度影响不大. 当相对离焦量误差达到 10% 时, 传感精度 RMS 降至  $0.065\lambda$ , 已无法满足通常情况下遥感系统的传感要求. 可见, 本文所提方法对系统相对离焦量误差的容限为 7.5%, 随着离焦量误差的增大, 传感精度随之降低.

#### 4.3.3 噪声的影响

采用探测器获取焦面、离焦面图像时, 图像噪声会影响波前传感精度, 为此在 PSF 图像上叠加了不同程度的高斯噪声, 研究了不同信噪比情况下, 本文所提方法的传感精度. 其中图像信噪比 SNR 定义为

$$\text{SNR} = 10 \times \log \left( \frac{\sigma_s^2}{\sigma_n^2} \right), \quad (10)$$

其中  $\sigma_s^2$  和  $\sigma_n^2$  分别表示图像和噪声的方差.

研究结果如表 4 所列. 由表 4 可以看出, 当 SNR 为 30—50 时, 该方法的传感精度介于  $0.015\lambda$ — $0.020\lambda$  之间. 当 SNR 降至 25 时, 传感精度 RMS 降为  $0.06\lambda$ , 已不能满足通常情况下遥感系统波前传感要求.

表 4 噪声对传感精度的影响

Table 4. The influence of noise on the sensing accuracy.

信噪比/dB	50	40	35	30	25
传感精度RMS/ $\lambda$	0.015	0.015	0.015	0.020	0.060

#### 4.3.4 图像采样分辨率的影响

当采样分辨率提升时, 获取的 PSF 图像信息将更为丰富, 为此, 采用  $128 \times 128$  的图像训练网络并进行测试, 同时与上文采用的  $64 \times 64$  图像训练

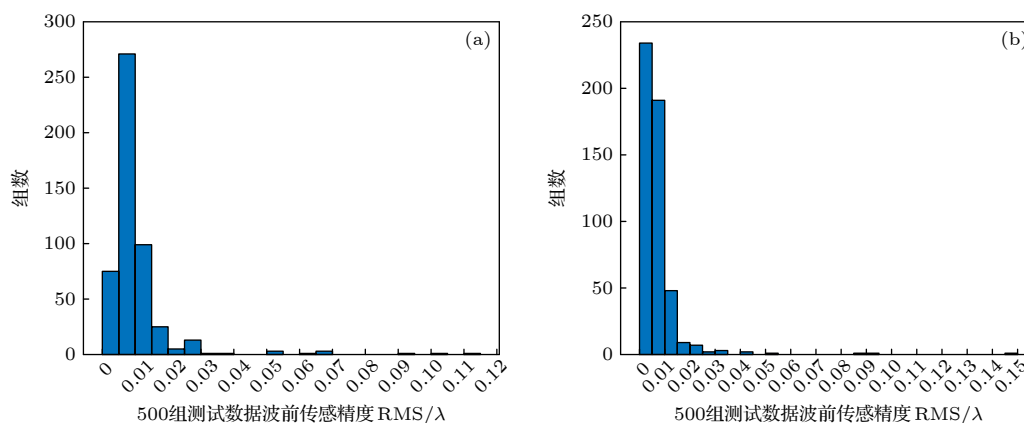


图 8 采用不同分辨率图像训练网络时,波前传感精度 RMS 统计直方图对比 (a) 图像采样率  $64 \times 64$ ; (b) 图像采样率  $128 \times 128$

Fig. 8. When image training networks with different resolutions are adopted, the RMS statistical histogram of wavefront sensing precision are compared: (a) Image sampling rate  $64 \times 64$ ; (b) image sampling rate  $128 \times 128$ .

结果进行对比. 不同采样率下, 500 组验证数据所得的波前传感精度 RMS 对比图如图 8 所示. 此时, 波前传感精度最高达 RMS  $0.0004\lambda$ , 最小为 RMS  $0.1500\lambda$ , 均值为 RMS  $0.0071\lambda$  且  $\text{RMS} \leq 0.015\lambda$  占比 92%. 可以看出, 随着图像采样分辨率的提升, 波前传感精度有所改善, 但网络输入数据亦随着采样率的升高而增加, 训练时间成本随之增加.

## 5 总 结

本文提出了一种基于小波变换图像融合和卷积神经网络的相位恢复方法, 采用小波变换图像融合技术对焦面、离焦面图像进行处理, 在不损失图像信息的同时简化网络输入. 将融合所得图像作为 CNN 的输入数据, 以相应的 4—9 阶 Zernike 系数作为 CNN 的输出数据进行训练. 训练结果表明当波前 RMS 值在  $1.1\lambda$  内时, 该方法波前传感精度 RMS 值可达  $0.015\lambda$ .

研究了噪声、离焦量误差、图像采样分辨率等因素对波前传感精度的影响, 验证了该方法对噪声具有一定的鲁棒性, 且相对离焦量误差在 7.5% 内时, 波前传感精度 RMS 仍可达  $0.05\lambda$ . 同时, 随着图像采样分辨率的提升, 波前传感精度有所改善, 但网络输入数据亦随着采样率的升高而增加, 训练时间成本随之增加.

此外, 在实际应用中, 为实现较高精度波前传感, 可依据系统像差特性, 基于本文方法, 改变网络输出层单元个数, 并调整网络结构, 训练适用于更高阶像差的新网络.

该方法与目前现有的相位恢复方法相比, 无需进行迭代处理, 更能满足实时性需求. 与现有的结合深度学习的相位恢复方法相比, 数据处理方法简单易行, 且精度更高, 除用于空间遥感系统外, 亦可用于天文望远镜波前传感, 大型光学镜面面形检测等领域.

## 参考文献

- [1] Roddier C, Roddier F 1993 *Appl. Opt.* **32** 2992
- [2] Lei W Z, Yuan L J, Su Z D, Kang Y, Wu Z H 2020 *Acta Optica Sin.* **40** 1312003-1 (in Chinese) [类维政, 袁吕军, 苏志德, 康燕, 武中华 2020 *光学学报* **40** 1312003-1]
- [3] Lu L H, Zhang W 2010 *J. Appl. Opt.* **31** 685 (in Chinese) [逯力红, 张伟 2010 *应用光学* **31** 685]
- [4] Wu Y L, Hu X J, Dai Y F, Li S Y 2009 *Chin. J. Mech. Eng.* **45** 157 (in Chinese) [吴宇列, 胡晓军, 戴一帆, 李圣怡 2009 *机械工程学报* **45** 157]
- [5] Gerehberg R W, Saxton W O 1972 *Optik* **35** 237
- [6] Fienup J R 1982 *Appl. Opt.* **21** 2758
- [7] Yang G Z, Gu B Y 1981 *Acta Phys. Sin.* **30** 410 (in Chinese) [杨国桢, 顾本源 1981 *物理学报* **30** 410]
- [8] Paine S W, Fienup J R 2018 *Opt. Lett.* **43** 1235
- [9] Nishizaki Y, Valdivia M, Horisaki R, Kitaguchi K, Saito M, Tanida J, Vera E 2019 *Opt. Express* **27** 240
- [10] Andersen T, Owner-Petersen M, Enmark A 2019 *Opt. Lett.* **44** 4618
- [11] Ju G H, Qi X, Ma H G, Yan C X 2018 *Opt. Express* **26** 31767
- [12] Qi X, Ju G H, Zhang C Y, Xu S Y 2019 *Opt. Express* **27** 26102
- [13] Mao H 2008 *Ph. D. Dissertation* (Beijing: Beijing Insitute of Technology) (in Chinese) [毛珩 2008 博士学位论文 (北京: 北京理工大学)]
- [14] Wang X 2010 *Ph. D. Dissertation* (Beijing: Beijing Insitute of Technology) (in Chinese) [王欣 2010 博士学位论文 (北京: 北京理工大学)]
- [15] Bai C T, Gao Z Q, Li A, Cui X L 2020 *Comput. Eng.* (in Chinese) [柏财通, 高志强, 李爱, 崔偻龙 2020 *计算机工程*]
- [16] Li Y D, Hao Z B, Lei H 2016 *J. Comput. Appl.* **36** 2508 (in Chinese)



- Chinese) [李彦冬, 郝宗波, 雷航 2016 *计算机应用* **36** 2508]
- [17] Xu Q W, Wang P P, Zeng Z J, Huang Z B, Zhou X X, Liu J M, Li Y, Chen S Q, Fan D Y 2020 *Acta Phys. Sin.* **69** 014209 (in Chinese) [徐启伟, 王佩佩, 曾镇佳, 黄泽斌, 周新星, 刘俊敏, 李瑛, 陈书青, 范滇元 2020 *物理学报* **69** 014209]
- [18] Shan B Z, Wang S Y, Niu H B, Liu S H 2002 *Opt. Precis. Eng.* **10** 318 (in Chinese) [单宝忠, 王淑岩, 牛憨笨, 刘颂豪 2002 *光学精密工程* **10** 318]
- [19] Wang C Y, Duan Q Q, Zhou K, Yao J, Su M, Fu Y C, Ji J Y, Hong X, Liu X Q, Wang Z Y 2020 *Acta Phys. Sin.* **69** 100701 (in Chinese) [王晨阳, 段倩倩, 周凯, 姚静, 苏敏, 傅意超, 纪俊羊, 洪鑫, 刘雪芹, 汪志勇 2020 *物理学报* **69** 100701]
- [20] Kingma D P, Ba J 2014 *Comput. Sci.* **1412** 6982
- [21] Yan S W 2012 *M. S. Thesis* (Lanzhou: Lanzhou University) (in Chinese) [闫胜武 2012 硕士学位论文 (兰州: 兰州大学)]
- [22] Sun A H 2014 *M. S. Thesis* (Qingdao: Ocean University of China) (in Chinese) [孙爱华 2014 硕士学位论文 (青岛: 中国海洋大学)]
- [23] Cai Z S, Chen M S 2015 *Las. Optoelect. Prog.* **52** 117 (in Chinese) [蔡植善, 陈木生 2015 *激光与光电子学进展* **52** 117]
- [24] Zhao L Y, Ma Q L, Li X R 2012 *Acta Phys. Sin.* **61** 194204 (in Chinese) [赵辽英, 马启良, 厉小润 2012 *物理学报* **61** 194204]
- [25] Yu S C, Lv Y Q 2009 *Appl. Res. Comput.* **26** 390 (in Chinese) [於时才, 吕艳琼 2009 *计算机应用研究* **26** 390]

# Phase retrieval wavefront sensing based on image fusion and convolutional neural network\*

Zhou Jing   Zhang Xiao-Fang<sup>†</sup>   Zhao Yan-Geng*(School of Optics and Photonics, Beijing Institute of Technology, Beijing 100081, China)*

( Received 18 August 2020; revised manuscript received 23 September 2020 )

## Abstract

The conventional phase retrieval wavefront sensing approaches mainly refer to a series of iterative algorithms, such as G-S algorithms, Y-G algorithms and error reduction algorithms. These methods use intensity information to calculate the wavefront phase. However, most of the traditional phase retrieval algorithms are difficult to meet the real-time requirements and depend on the iteration initial value used in iterative transformation or iterative optimization to some extent, so their practicalities are limited. To solve these problems, in this paper, a phase-diversity phase retrieval wavefront sensing method based on wavelet transform image fusion and convolutional neural network is proposed. Specifically, the image fusion method based on wavelet transform is used to fuse the point spread functions at the in-focus and defocus image planes, thereby simplifying the network inputs without losing the image information. The convolutional neural network (CNN) can directly extract image features and fit the required nonlinear mapping. In this paper, the CNN is utilized to establish the nonlinear mapping between the fusion images and wavefront distortions (represented by Zernike polynomials), that is, the fusion images are taken as the input data, and the corresponding Zernike coefficients as the output data. The network structure of the training in this paper has 22 layers, they are 1 input layer, 13 convolution layers, 6 pooling layers, 1 flatten layer and 1 full connection layer, that is, the output layer. The size of the convolution kernel is  $3 \times 3$  and the step size is 1. The pooling method selects the maximum pooling and the size of the pooling kernel is  $2 \times 2$ . The activation function is ReLU, the optimization function is Adam, the loss function is the MSE, and the learning rate is 0.0001. The number of training data is 10000, which is divided into three parts: training set, validation set, and test set, accounting for 80%, 15% and 5% respectively. Trained CNN can directly output the Zernike coefficients of order 4–9 to a high precision, with these fusion images serving as the input, which is more in line with the real-time requirements. Abundant simulation experiments prove that the wavefront sensing precision is root-mean-square(RMS)  $0.015\lambda$ , when the dynamic range of the wavefront is the aberration of low spatial frequency within  $1.1\lambda$  of RMS value (i.e. the dynamic range of Zernike coefficients of order 4–9 is  $[-0.5\lambda, 0.5\lambda]$ ). In practical application, according to the system aberration characteristics, the number of network output layer units can be changed and the network structure can be adjusted based on the method presented in this paper, thereby training the new network suitable for higher order aberration to realize high-precision wavefront sensing. It is also proved that the proposed method has certain robustness against noise, and when the relative defocus error is within 7.5%, the wavefront sensor accuracy is acceptable. With the improvement of image resolution, the wavefront sensing accuracy is improved, but the number of input data of the network also increases with the sampling rate increasing, and the time cost of network training increases accordingly.

**Keywords:** phase retrieval, convolutional neural network, image fusion, sensing accuracy**PACS:** 42.15.Fr, 42.30.Kq, 42.30.Rx**DOI:** 10.7498/aps.70.20201362

\* Project supported by the National Natural Science Foundation of China (Grant No. 61471039).

<sup>†</sup> Corresponding author. E-mail: [zhangxf@bit.edu.cn](mailto:zhangxf@bit.edu.cn)