

## 自适应增益的 SPGD 算法

方 舟<sup>1</sup>, 徐项项<sup>2</sup>, 李 鑫<sup>2</sup>, 刘金龙<sup>2</sup>, 杨慧珍<sup>2\*</sup>, 龚成龙<sup>2</sup>

- (1. 中国矿业大学 信息与控制工程学院, 江苏 徐州 221116;
2. 江苏海洋大学 电子工程学院, 江苏 连云港 222005)

**摘 要:** SPGD 算法是一种应用广泛的无波前探测自适应光学控制算法。传统 SPGD 算法中增益系数保持某一固定值不变, 随着变形镜单元数的增加, 这将导致算法收敛速度变慢及陷入局部极值的概率增大。Adam 优化器是深度学习常用的一种优化随机梯度下降算法, 它具有增益系数自适应性调整的特点。将 Adam 优化器自适应调整增益系数的优势与 SPGD 算法结合起来用于自适应光学系统控制。分别以 32、61、97、127 单元变形镜作为波前校正器件, 不同湍流强度的波前像差作为校正对象, 建立了无波前探测自适应光学系统模型。结果表明, 优化后的算法收敛速度更快, 而且陷入局部极值的概率降低, 并且随着变形镜单元数的增加与湍流强度的增大, 算法的优势更加明显。以上研究结果为基于 Adam 优化的 SPGD 算法的实际应用提供了理论基础。

**关键词:** 自适应光学; SPGD; 波前校正; 变形镜; Adam

**中图分类号:** TP273.2 **文献标志码:** A **DOI:** 10.3788/IRLA20200274

## SPGD algorithm with adaptive gain

Fang Zhou<sup>1</sup>, Xu Xiangxiang<sup>2</sup>, Li Xin<sup>2</sup>, Liu Jinlong<sup>2</sup>, Yang Huizhen<sup>2\*</sup>, Gong Chenglong<sup>2</sup>

- (1. School of Information and Control Engineering, China University of Mining and Technology, Xuzhou 221116, China;
2. School of Electronic Engineering, Jiangsu Ocean University, Lianyungang 222005, China)

**Abstract:** SPGD is a control algorithm widely used in wavefront sensorless adaptive optics (AO) systems. The gain is commonly set to a fixed value in the traditional SPGD algorithm. With the increase of the number of DM elements, which can easily lead to the slow convergence speed of the algorithm and the increase of the probability of falling into the local extreme value. Adam optimizer is an optimized stochastic gradient descent algorithm commonly used in deep learning. It has the advantage of achieving adaptive learning rate. The advantages of Adam optimizer adaptive gain and SPGD algorithm were combined to realize adaptive gain for AO system control. The simulation model of wavefront sensorless AO system was established with 32, 61, 97 and 127 elements DM as wavefront correction devices respectively, wavefront aberrations with different turbulence intensities as correction objects. The results show that the optimized algorithm converges faster than basic SPGD algorithm and the probability of falling into local extremum decreases. As the number of DM elements increases and the turbulence intensity increases, the advantages of the optimized algorithm are more obvious. The above research results provide a theoretical basis for the practical application of the SPGD algorithm based on Adam

收稿日期:2020-07-10; 修订日期:2020-08-29

基金项目:国家自然科学基金(11573011);江苏省“六大人才高峰”高层次人才项目(KTHY-058);江苏省“333”高层次人才培养项目(BRA2019244);江苏省研究生科研与实践创新计划项目(KYCX20\_2906, KYCX20\_2961)

作者简介:方舟(1996-),男,硕士生,主要从事自适应光学系统控制算法方面的研究。Email: Fangzhou9629@163.com

导师简介:杨慧珍(1973-),女,教授,博士,主要从事自适应光学技术及其应用方面的研究。Email: yanghz@jou.edu.cn(通讯联系人)

optimization.

**Key words:** adaptive optics; SPGD; wavefront correction; DM; Adam

## 0 引言

随着自适应光学技术应用领域的拓展,小型化和低成本化成为趋势。相比于常规自适应光学系统,无波前探测自适应光学系统复杂性低、应用领域广,这一技术成为目前相关领域的研究热点<sup>[1-4]</sup>。在无波前探测自适应光学系统中,随机并行梯度下降算法(Stochastic Parallel Gradient Descent, SPGD)具有并行化、鲁棒性强、实现简单等优点,是目前最具有应用前景的一种无波前探测自适应光学系统控制算法<sup>[5-6]</sup>。在 SPGD 算法实现过程中,增益系数的大小决定控制算法的收敛速度和校正性能,但现有算法中大多凭借经验选取某一固定值作为增益系数。目前随着大口径望远镜的发展,变形镜所需单元数可达上千个。单元数越多,意味着控制参数维度增加,算法寻优范围变大,固定值增益容易导致算法收敛速度变慢及陷入局部极值的概率增大。

Adam (Adaptive Moment Estimation)<sup>[7]</sup> 优化器是深度学习里常用的一种优化随机梯度下降算法,具有实现简单、计算高效、所占内存较少和参数设定方便等优点,在应用深度学习的图像处理、机器视觉等领域使用广泛<sup>[8-9]</sup>。和传统的随机梯度下降算法保持固定的学习率不同,Adam 结合了动量<sup>[10]</sup>和 RMSProp (Root mean square Prop) 算法的优点。每一次算法迭代都利用梯度的一阶矩和二阶矩估计信息对学习率进行调整,通过调整学习率控制梯度下降的方向和步长,减小梯度下降的波动使其快速稳定在下降的最低点,从而提高算法的收敛速度和降低陷入局部极值的概率。

文中将 Adam 优化器自适应调整学习率的优势与 SPGD 算法结合起来用于自适应光学系统控制。分别以 32、61、97、127 单元变形镜作为波前校正器件,不同湍流强度的波前像差作为校正对象,建立了无波前探测自适应光学系统仿真模型。首先考察了现有 SPGD 算法随着控制参数维度的增多,收敛速度变慢及陷入局部极值的情况。其次分析了基于 Adam 优化的 SPGD 控制算法在自适应光学系统中的校正速度和陷入局部极值的情况,并将优化后的结果

与现有 SPGD 控制算法结果进行了比较。

## 1 算法介绍

### 1.1 SPGD 算法

SPGD 算法利用性能指标  $j$  的变化量  $\nabla j$  与控制参量  $u$  的变化量  $\nabla u$  对控制参量的梯度进行估计,利用迭代方式在控制参量的梯度下降方向上搜索。第  $k$  次迭代时,施加在变形镜  $N$  个驱动器上的电压向量  $u^k = \{u_1, u_2, \dots, u_N\}^k$  的计算公式为:

$$u^{(k)} = u^{(k-1)} + \alpha \nabla u^{(k)} \nabla j^{(k)} \quad (1)$$

式中:  $\alpha$  是增益系数;  $\nabla j^{(k)}$  是系统性能指标的变化量;  $\nabla u^{(k)} = \{\nabla u_1, \nabla u_2, \dots, \nabla u_N\}^k$  是第  $k$  次迭代时施加的随机扰动电压向量。具体过程如下,首先生成扰动电压  $\nabla u^{(k)}$ , 将电压  $u^{(k-1)} + \nabla u^{(k)}$  施加到变形镜上,然后采集 CCD 图像并计算性能指标  $j_+^{(k)}$ ; 同样再把电压  $u^{(k-1)} - \nabla u^{(k)}$  施加到变形镜上,采集图像计算出性能指标  $j_-^{(k)}$ ; 根据公式 (2) 计算性能指标变化量:

$$\nabla j^{(k)} = j_+^{(k)} - j_-^{(k)} \quad (2)$$

然后按照公式 (1) 计算控制电压向量  $\nabla u^{(k)} = \{\nabla u_1, \nabla u_2, \dots, \nabla u_N\}^k$  施加到变形镜上,采集图像得到校正效果完成第  $K$  次迭代。

### 1.2 基于 Adam 优化的 SPGD 算法

利用 Adam 的自适应学习率的优点,将 Adam 算法和 SPGD 算法相结合,把 SPGD 算法中固定增益系数优化为自适应增益系数,提高算法的收敛速度和减小局部极值的概率。

根据公式 (2) 求得第  $k$  次的梯度估计值为  $dw = \nabla u \nabla j$ 。然后再根据下面公式 (3)~(6) 进行梯度优化。

实际过程如下,首先初始化梯度的一阶动量  $v_{dw}$  和二阶动量  $s_{dw}$ 。在第  $K$  次迭代时,依据公式 (3) 和 (4) 计算梯度的一阶动量和二阶动量。

$$v_{dw} = \beta_1 v_{dw} + (1 - \beta_1) dw \quad (3)$$

$$s_{dw} = \beta_2 v_{dw} + (1 - \beta_2) dw^2 \quad (4)$$

由于在迭代初期,计算偏差太大需要修正,公式 (5) 和 (6) 是对梯度值进行偏差修正。

$$\hat{v}_{dw} = \frac{v_{dw}}{1 - \beta_1^k} \quad (5)$$

$$\hat{s}_{dw} = \frac{s_{dw}}{1 - \beta_2^k} \quad (6)$$

最后把以上求得的参数  $\hat{s}_{dw}$  和  $\hat{v}_{dw}$  代入公式 (7) 中, 计算出第  $K$  次迭代的控制电压。

$$u^k = u^{k-1} - \frac{\partial}{\sqrt{\hat{s}_{dw} + \varepsilon}} v_{dw} \quad (7)$$

$\frac{\partial}{\sqrt{\hat{s}_{dw} + \varepsilon}}$  代表了自适应增益, 每一次迭代都会进行调整。文中默认值  $\beta_1 = 0.9$ ,  $\beta_2 = 0.999$ ,  $\varepsilon = 10^{-8}$ 。 $\partial$  根据情况设定大小, 设为 0.01。算法的具体实现过程如图 1 所示。

## 2 系统介绍

如图 2 所示, 无波前探测自适应光学系统由波前校正器 (变形镜), CCD, 波前控制模块等组成。系统工作流程如下, 畸变的波前经过变形镜反射经透镜聚焦到 CCD 上, CCD 采集图像并传输到控制器中。通过控制算法的计算得到控制信号, 经过高压放大单元传送给变形镜, 从而驱动变形镜产生形变补偿像差。然后将残余波前作为待校正的波前继续进行闭环校正, 直到满足预设的条件为止。

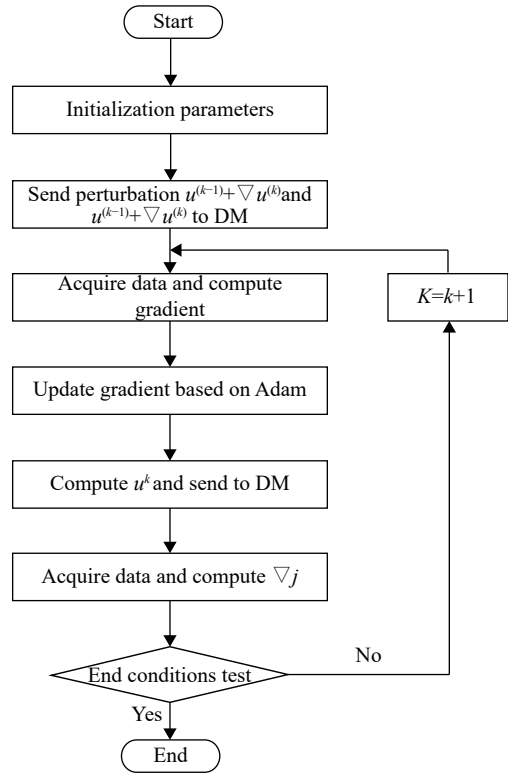


图 1 优化算法实现流程

Fig.1 Optimization algorithm implementation process

文中的无波前探测自适应光学系统仿真实验用到 4 种单元变形镜, 分别是 37、61、97、127 单元。根据变形镜校正波前的工作原理, 变形镜产生的

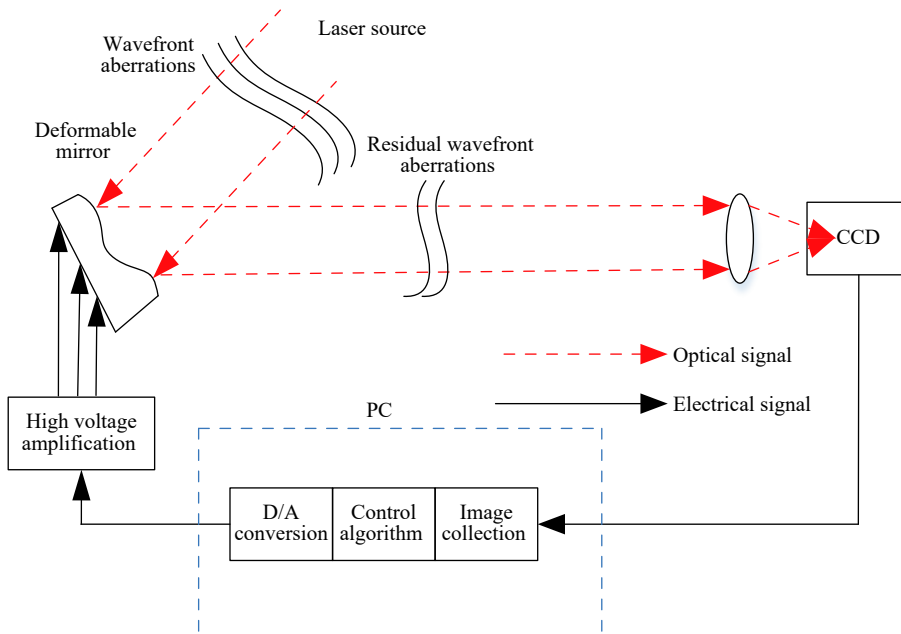


图 2 无波前探测自适应光学系统

Fig.2 Wavefront sensorless adaptive optics system

补偿面形 $\phi(x,y)$ 可以用各驱动器的影响函数线性组合表示:

$$\phi(x,y) = \sum_{i=1}^N v_i S_i(x,y) \quad (8)$$

式中:  $S_i(x,y)$  是每个驱动器的影响函数;  $v_i$  是对变形镜施加的控制电压,  $i$  是驱动器数量。

仿真实验中采用 Roddier<sup>[11]</sup> 提出的方法生成相屏模拟待校正正像差, 相屏由 3~104 阶泽尔尼克像差组成, 不包含倾斜项。生成的相屏统计属性符合 Kolmogrov 谱, 且相屏间不具有相关性。湍流强度大小用  $D/r_0$  表示,  $D$  为望远镜的口径,  $r_0$  为大气相干长度。斯特列耳比 (SR-Strehl Ration) 作为目标函数。以下如无特殊说明, 均为随机选择 500 帧相屏, 将统计平均值作为实验结果。

### 3 仿真结果与分析

#### 3.1 常规 SPGD 算法

基于建立的无波前探测自适应光学系统的仿真模型, 首先分析随着变形镜单元数的增加, 常规 SPGD 算法收敛性和陷入局部极值概率的情况。分别采用 4 种单元变形镜去校正湍流强度  $D/r_0=5$  情况下的波前像差。系统迭代 1000 次, 4 种单元变形镜各自收敛情况如下图 3 所示。图中“37”、“61”、“97”和“127”分别代表 4 种不同单元的变形镜。

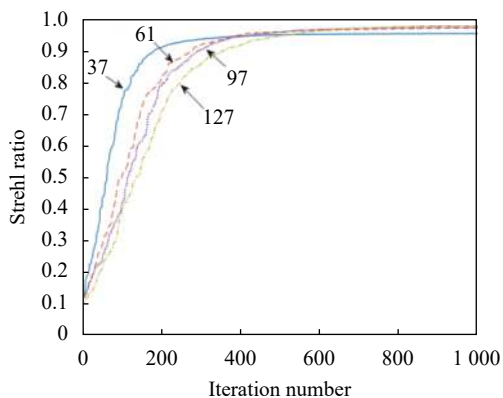


图 3 基于 SPGD 算法在 4 种变形镜情况下系统的收敛情况

Fig.3 System convergence based on SPGD algorithm under four DMs

从图 3 中可以明显看出随着变形镜单元数的增加, 系统收敛速度逐渐变慢。以 SR 达到 0.8 时的迭代次数作为校正速度的评判标准。37 单元变形镜需

要 107 次, 61 单元变形镜需要 146 次, 97 单元变形镜需要 190 次, 127 单元变形镜需要 220 次。

为了更加直观对比 4 种变形镜陷入局部极值的概率, 统计了系统陷入局部极值的帧数。37 单元变形镜有 47 帧, 61 单元有 72 帧, 97 单元有 98 帧, 127 单元有 136 帧。从数据中可以看出, 随着变形镜单元数的增加, 系统陷入局部极值的帧数逐渐增多, 意味着系统陷入局部极值的概率逐渐增大。

下面考察了 127 单元变形镜分别校正  $D/r_0=5$ 、10、15、20 这 4 种湍流强度下的波前像差的情况。收敛曲线如图 4 所示。

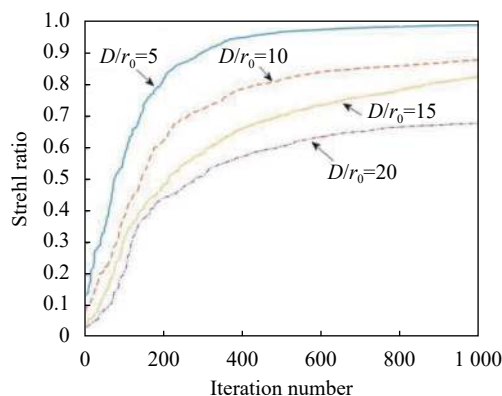


图 4 基于 127 单元变形镜 SPGD 算法在不同湍流强度下的收敛情况

Fig.4 Convergence of SPGD algorithm based on 127 elements DM under different  $D/r_0$

从图中可以看出随着湍流强度的增大, 系统的收敛速度逐渐变慢。算法迭代 1000 次后,  $D/r_0=15$  和 20 的情况下没有完全收敛。另外, 统计了 4 种湍流强度下陷入局部极值的帧数,  $D/r_0=5\sim 20$  分别为 136、158、200、240 帧。由数据可以看出随着湍流强度的增大, 系统陷入局部极值的概率也逐渐增大。

#### 3.2 基于 Adam 优化的 SPGD 算法与常规 SPGD 算法对比

##### 3.2.1 增益系数的对比

基于 Adam 优化的 SPGD 算法理论上具有增益系数自适应变化的特点, 随机选取  $D/r_0=5$  时的 50 帧不同相屏作为校正对象, 分析不同单元数变形镜对应的自适应光学系统增益系数变化情况, 如图 5 所示, 其中“SPGD”代表基于常规 SPGD 算法, 其它代表基于优化算法, 每条曲线为 50 帧相屏的平均结果。算法

迭代了 1000 次,但 300~1000 次迭代增益系数基本保持不变,为图像更加清晰,只给出了前 300 次迭代情况。

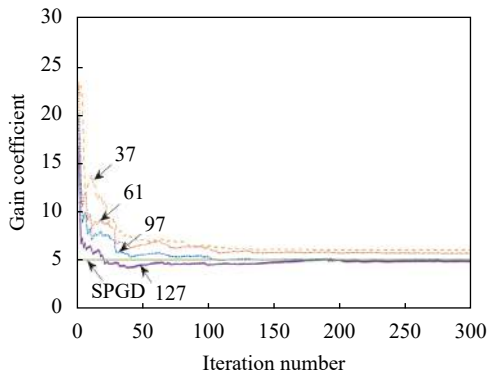


图 5 两种算法的增益系数变化情况对比

Fig.5 Comparison of gain coefficient changes of two algorithms

从图 5 中可以看出,4 种不同单元数的变形镜,优化后的增益系数在算法迭代初期都比固定增益值大得多,大的增益意味着朝优化方向的步长加大,收敛速度加快,并且收敛速度越快的增益初始值越大。经过一定的迭代次数后,优化后增益系数逐渐接近固定增益值的大小,且随着变形镜单元数的增加,接近固定增益值的速度越快。

### 3.2.2 收敛速度的对比

基于两种算法,首先对 4 种变形镜校正同一湍流强度像差的情况进行了对比分析。图 6 给出了基于优化后的 SPGD 算法在 4 种不同变形镜的情况下系统收敛情况。其中“37”、“61”、“97”和“127”分别代表 4 种单元变形镜。

对比图 3 可以看出,优化前后算法在 4 种变形镜

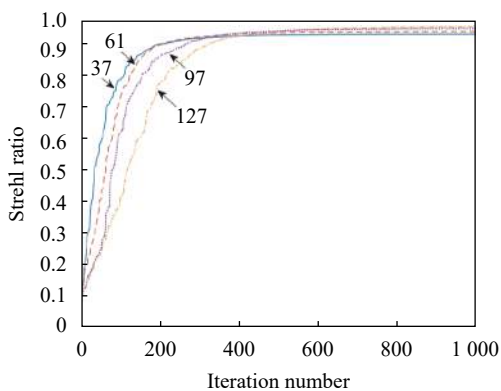


图 6 基于优化算法在 4 种变形镜情况下系统收敛情况

Fig.6 System convergence based on optimized algorithm under four DMs

的情况下迭代 1000 次都已经收敛。统计了优化后算法的 SR 达到 0.8 时所需迭代次数,在 4 种变形镜情况下分别需要 91、122、150、172 次。与常规 SPGD 算法的数据对比,优化后的算法收敛速度更快。

为了分析随着变形镜单元数的增加,优化前后算法的收敛趋势,将两种算法收敛时所需要的迭代次数进行对比。结果如图 7 所示,其中所示每条连接线的 4 个标记位置,对应的横坐标表示 4 种变形镜的单元数,纵坐标表示每种变形镜校正 SR 达到 0.8 时需要的迭代次数。“Adam”表示优化的 SPGD 算法,“SPGD”表示常规的 SPGD 算法。

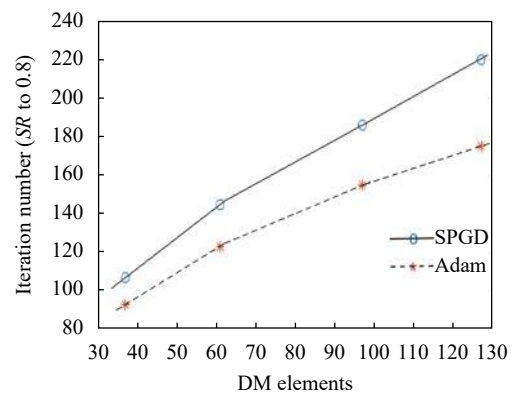


图 7 两种算法在 4 种变形镜情况下收敛时的迭代次数对比

Fig.7 Comparison of iteration number at convergence under four DMs based on two algorithm

从图 7 中可以看出,随着变形镜单元数的增多,优化前与优化后的算法收敛所需的迭代次数差距逐渐增大。表明了单元数越多,优化后算法的优势越明显。

图 8 给出了优化算法在 127 单元变形镜情况下对 4 种湍流强度下像差校正的对比情况。

与常规 SPGD 算法趋势相同,随着湍流强度的增大,系统收敛速度逐渐变慢。但算法迭代 1000 次后,在  $D/r_0=15$  和 20 情况下优化算法已经收敛,表明优化后算法收敛速度更快。

图 9 给出了使用 127 单元变形镜校正 4 种湍流强度像差时,优化前与优化后的算法收敛时所需迭代次数的对比情况。其中横坐标表示湍流强度,纵坐标表示算法收敛时的迭代次数。

从图 9 中可以看出,随着湍流强度的增大,优化前后算法收敛时所需的迭代次数差距变大。表明了湍流强度越大,优化后算法优势越明显。

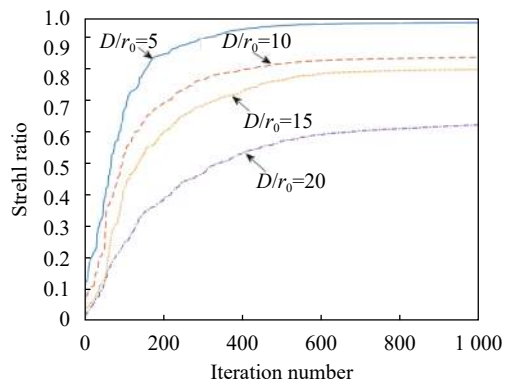


图 8 基于 127 单元变形镜的优化 SPGD 算法在不同湍流强度下的收敛情况

Fig.8 Convergence of optimized SPGD algorithm based on 127 elements DM under different  $D/r_0$

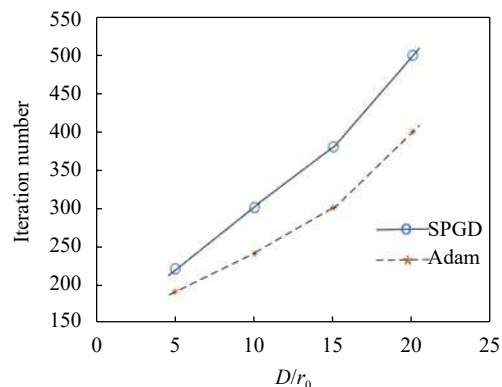


图 9 基于 127 单元变形镜不同湍流强下两种算法收敛时的迭代次数对比

Fig.9 Comparison of iteration number of two algorithms at convergence based on 127 elements DM under different  $D/r_0$

### 3.2.3 陷入局部极值情况的对比

为了对比优化前后算法陷入局部极值的情况,统计了基于 4 种变形镜在同一湍流强度和基于 127 单元变形镜在不同湍流强度校正像差时,两种算法陷入局部极值的帧数,见表 1 和表 2。

通过表 1 中两种算法的数据对比,优化后 37 单元变形镜陷入局部极值的概率降低了 30%, 61 单元变形镜收敛陷入局部极值的概率降低了 31.2%, 97 单元变形镜陷入局部极值概率降低了 32%, 127 单元变形镜陷入局部极值的概率降低了 35.3%。

根据表 2 中的数据对比,优化算法在 4 种湍流强度情况下陷入局部极值的概率分别降低了 30%、32%、33% 和 36%。

表 1 不同变形镜情况下陷入局部极值的帧数

Tab.1 Number of frames caught in local extremum in different DM

DM/elements	SPGD	Adam
37	48	30
61	72	50
97	98	67
127	136	88

表 2 不同湍流强度下陷入局部极值的帧数

Tab.2 Number of frames caught in local extremum under different  $D/r_0$

$D/r_0$	SPGD	Adam
5	136	88
10	158	110
15	200	135
20	240	154

由此可以看出优化后的算法陷入局部极值的概率有所降低。同时随着变形镜单元数的增加,湍流强度越大这种优势也越来越明显。

## 4 结论

在常规 SPGD 算法中随着控制参数的增多,算法收敛速度会逐渐变慢而且陷入局部极值的概率逐渐增大。在常规 SPGD 算法基础上,把求得的梯度估计值结合 Adam 优化器自适应学习率的更新公式,把固定增益变为自适应增益,将控制参量按照更新公式进行更新获得最优控制电压。以 4 种不同单元的变形镜作为校正器件,不同湍流强度下的像差作为校正对象,建立了无波前探测自适应光学系统模型。基于 Adam 优化的 SPGD 控制算法闭环校正结果表明,优化后的算法在每一次迭代中增益系数都自动进行更新,与现有的 SPGD 算法相比优化后的算法曲线波动较小、更加平滑,收敛速度更快并且陷入局部极值的概率降低。随着变形镜单元数增多,湍流强度的增大,优化后算法的优势更加明显。研究结果为基于 Adam 优化的 SPGD 算法的实际应用提供了理论基础。

### 参考文献:

[1] Song J, Li Y, Che D, et al. Coherent beam combining based on

- the SPGD algorithm with a momentum term [J]. *Optik*, 2020, 202: 163650.
- [2] Huang G, Geng C, Li F, et al. Adaptive SMF coupling based on precise-delayed SPGD algorithm and its application in free space optical communication [J]. *IEEE Photonics Journal*, 2018, 10(3): 1-12.
- [3] Zhou H Q, Hang L L, Wang Y T. Deep learning algorithm and its application in optics [J]. *Infrared and Laser Engineering*, 2019, 48(12): 1226004.
- [4] Yang Huizhen, Wang Bin, Liu Ruiming, et al. Analysis of anti-noise capability of model-based wavefront sensorless adaptive optics system[J]. *Infrared and Laser Engineering*, 2017, 46(8): 0817002. (in Chinese)
- [5] Wang Weibing, Wang Tingfeng, Guo Jin. Simulation on the law of wave-front shaping with stochastic parallel gradient descent algorithm for adaptive optic[J]. *Chinese Journal of Optics*, 2014, 7(3): 411-420. (in Chinese)
- [6] Yang Huizhen, Chen Bo, Li Xinyang, et al. Experimental demonstration of Stochastic Parallel Gradient Descent Control Algorithm for adaptive optics system[J]. *Acta Optica Sinica*, 2008, 2(6): 205-210. (in Chinese)
- [7] Diederik P Kingma, Jimmy Ba. Adam: A method for stochastic optimization[C]//3rd International Conference for Learning Representations, 2015.
- [8] Huang L H, Cao L H, Li N, et al. A state perception method for infrared dim and small targets with deep learning [J]. *Chinese Optics*, 2020, 13(3): 527-536.
- [9] Fan L L, Zhao H W, Zhao H Y. Survey of target detection based on deep convolutional neural networks [J]. *Optics and Precision Engineering*, 2020, 28(5): 1152-1164.
- [10] Yang G, Liu L, Jiang Z, et al. Incoherent beam combining based on the momentum SPGD algorithm [J]. *Opt Laser Technol*, 2018, 101: 372-378.
- [11] Roddier N A. Atmospheric wavefront simulation using Zernike polynomials [J]. *Optical Engineering*, 1990, 29(10): 1174-1180.