

文章编号: 0253-2239(2007)08-1355-6

自适应光学系统随机并行梯度下降控制算法 仿真与分析*

杨慧珍^{1,2} 李新阳¹ 姜文汉¹

(1 中国科学院光电技术研究所, 成都 610209)
(2 中国科学院研究生院, 北京 100039)

摘要: 随机并行梯度下降算法能不依赖波前传感器直接对系统性能进行优化。以 32 单元变形镜为校正器, 采用随机并行梯度下降算法建立了自适应光学系统仿真模型。通过分析该系统对静态波前畸变的校正能力, 验证了随机并行梯度下降算法的收敛性; 讨论了算法增益系数、随机扰动幅度与收敛速度的关系, 并指出通过算法增益系数的自适应调整可以改进算法的收敛速度。

关键词: 自适应光学; 随机并行梯度下降算法; 数值仿真; 变形镜; 收敛

中图分类号: TP273+.2 文献标识码: A

Simulation and Analysis of Stochastic Parallel Gradient Descent Control Algorithm for Adaptive Optics System

Yang Huizhen^{1,2} Li Xinyang¹ Jiang Wenhan¹

(1 Institute of Optoelectronics, the Chinese Academy of Sciences, Chengdu 610209)
(2 Graduate University of Chinese Academy of Sciences, Beijing 100039)

Abstract: The stochastic parallel gradient descent (SPGD) algorithm can optimize the system performance directly, while being independent of wave-front sensor. Based on SPGD algorithm, an adaptive optics system model with a 32-element deformable mirror was simulated. Convergence of SPGD algorithm was verified through analyzing correction capabilities for static wave-front aberrations. The relationship of algorithm gain coefficient, stochastic perturbation amplitude and convergence rate were discussed. Convergence rate can be improved by adaptive adjustment of algorithm gain coefficient.

Key words: adaptive optics; stochastic parallel gradient descent algorithm; numerical simulation; deformable mirror; convergence

1 引 言

自适应光学系统通常采用波前传感器探测畸变波前相位信息, 用波前校正器校正畸变波前, 得到接近衍射极限的校正效果, 该技术在天文自适应成像领域已得到成功应用^[1]。随着自适应光学技术应用领域的扩展, 如自由空间光通讯, 光波穿过的路径上湍流较强或传输距离较长时, 闪烁现象导致波前测量困难, 致使依赖波前测量的常规自适应光学系统不能正常工作, 因此常规自适应光学控制技术在这些情况下的应用就显得无能为力。而无需波前传感器的像清晰化自适应光学技术有可能解决这个问题。

1974 年, Richard A. Muller 等^[2] 提出像清晰化自适应光学技术, 利用优化算法优化系统性能指标, 无需使用波前传感器, 降低了系统复杂性。像清晰化技术独立于系统模型, 又称“盲优化”技术, 适用于大多数类型的性能质量指标以及多数波前校正器。中科院光电技术研究所于 1987 年实现了多元高频振动自适应光学系统^[3,4]。多元高频振动对系统的带宽要求高、信噪比低且硬件实现复杂, 而爬山法收敛速度太慢。随着随机优化算法理论的发展, 20 世纪 90 年代后期研究人员开始尝试一些新的优化算法, 如随机并行梯度下降 (SPGD) 算法^[5,6]、遗传算法^[7]、模拟退

* 国家高技术激光领域资助课题。

作者简介: 杨慧珍(1973-)女, 河南人, 博士研究生, 主要从事自适应光学技术等方面的研究。E-mail: yanghz526@126.com

导师简介: 李新阳(1971-)男, 四川人, 博士, 研究员, 主要从事自适应光学技术等方面的研究。E-mail: xyli@ioe.ac.cn

收稿日期: 2006-11-03; 收到修改稿日期: 2007-01-17

火算法^[8,9]等。这些算法中随机并行梯度下降算法实用性最高,曾在实际自适应光学系统上得到应用^[6]。随机并行梯度下降算法是在同时扰动随机近似控制算法^[10]基础上开发的自适应光学校正技术,和多元高频振动及爬山法相比,具有实现容易、控制参量的所有维并行计算的特点。

国内的自适应光学系统大都基于传统的波前测量,而对基于随机并行梯度下降算法的自适应光学技术的研究目前仍处于起步阶段。M. A. Vorontsov等^[5,6]针对不同单元数、不同校正器类型比较了随机并行梯度下降算法对波前畸变的校正能力。本文以32单元变形镜作为校正器建立自适应光学系统仿真模型,研究随机并行梯度下降算法的收敛性及其对静态波前畸变的校正能力,重点考察算法增益系数和随机扰动幅度之间的关系及其对校正效果和收敛速度的影响,为把随机并行梯度下降算法应用于实际自适应光学系统进行了初步的探索。

2 自适应光学系统随机并行梯度下降算法仿真模型

2.1 仿真模型的建立

图1为采用随机并行梯度下降控制算法的32单元自适应光学系统仿真模型,由32单元变形镜、光源、成像系统、性能指标分析模块及随机并行梯度下降算法模块组成。图2为32单元变形镜驱动器位置排布图。图1中 $\varphi(r)$ 为原始畸变波前 $u(r)$ 为变形镜引入的校正相位, $\phi(r)=\varphi(r)+u(r)$ 为残余相位, $r=\{x,y\}$ 为正交于光轴的平面中的向量, $\varphi(r)$ 、 $u(r)$ 都是连续函数。 J 为系统性能指标, $u=\{u_1, u_2, \dots, u_{32}\}$ 为控制变形镜32个驱动器的电压向量。

系统工作原理:首先由性能指标分析模块计算性能指标 J 的变化量 ΔJ ,随机并行梯度下降模块

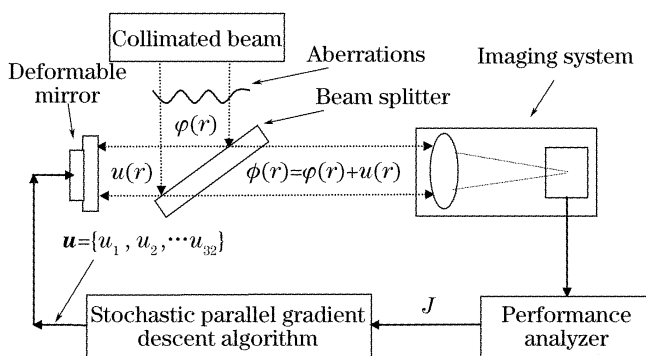


图1 仿真模型框图
Fig. 1 Block diagram of simulation

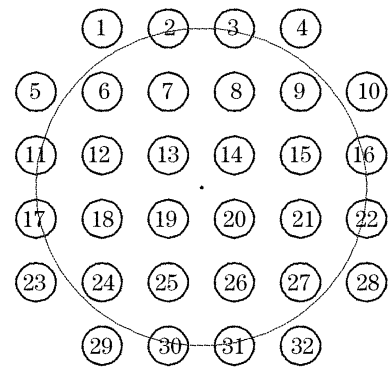


图2 32单元变形镜驱动器位置排布

Fig. 2 Location of 32-element deformable mirror actuator
根据 ΔJ 计算出 $u=\{u_1, u_2, \dots, u_{32}\}$,并施加到变形镜各驱动器;再次由指标分析模块计算校正相位后残余波前的系统性能指标;如此迭代进行,直至满足系统对指标的要求为止。迭代的过程则是在参量空间中寻找最佳电压向量 $u=\{u_1, u_2, \dots, u_{32}\}$,使得变形镜逐步生成一个趋于原始波前 $\varphi(r)$ 反相的 $u(r)$,从而使残余波前 $\phi(r)$ 最小,系统性能指标趋于最优。下面讨论每个模块的具体实现。

2.2 变形镜模块

32单元变形镜各驱动器成正方形排布,仿真中按照通光口径和驱动器间距对各驱动器坐标在单位圆内进行归一化,驱动器归一化间距为0.383。据测试32单元变形镜影响函数近似为高斯形式:

$$S_j(r) = S_j(x, y) = \exp\{\ln \omega [\sqrt{(x-x_j)^2 + (y-y_j)^2}/d]^a\}, \quad (1)$$

其中 $S_j(r)$ 为第 j 个驱动器的影响函数, (x_j, y_j) 为第 j 个驱动器的位置, ω 为驱动器交连值, d 为驱动器间距, a 为高斯指数。实际测试 $\omega = 0.08, a = 2$ 。变形镜面形变化 $u(r)$ 可用各个驱动器影响函数 $S_j(r)$ 的线性组合表示:

$$u(r) = \sum_{j=1}^{32} u_j S_j(r), \quad (2)$$

其中 u_j 为第 j 个驱动器所加电压。

2.3 性能指标分析模块

由于施特列耳比是自适应光学领域一个通用的性能评价标准,而且不管是成像系统还是激光发射系统,施特列耳比的大小都与具体应用中的性能评价标准相一致,因此选取施特列耳比作为系统性能指标 J 。将畸变波前和理想平面波前分别定义在 100×100 网格内的单位圆上,通过傅里叶变换计算波前对应的远场光斑。畸变波远场光斑中心光强与理想平面波远场光斑中心光强之比即为峰值施特列耳比。 J 与变形镜引入的校正相位之间的关系为

$$J \propto J[\phi(r)] = J[\varphi(r) + u(r)], \quad (3)$$

校正过程中原始畸变 $\varphi(r)$ 不变, J 可以看作 $u(r)$ 的非线性函数。再由(2)式知, J 是 32 个驱动器电压变量的非线性函数, 即 $J = J(\mathbf{u}) = J(u_1, u_2, \dots, u_{32})$ 。在实际的应用中, 性能指标分析所需数据可通过 CCD 或针孔获取, 并针对不同的应用, 定义不同的性能指标函数。

2.4 随机并行梯度下降算法模块

随机并行梯度下降算法实现如下:

性能指标 J 的变化量:

$$\Delta J^{(k)} = \Delta J_+^{(k)} - \Delta J_-^{(k)}, \quad (4)$$

其中 $\Delta J_+^{(k)}$ 和 $\Delta J_-^{(k)}$ 的计算公式:

$$\begin{aligned} \Delta J_-^{(k)} &= J[\mathbf{u}^{(k)} - \Delta \mathbf{u}^{(k)}] - J[\mathbf{u}^{(k)}], \\ \Delta J_+^{(k)} &= J[\mathbf{u}^{(k)} + \Delta \mathbf{u}^{(k)}] - J[\mathbf{u}^{(k)}], \end{aligned} \quad (5)$$

电压参量 u 的迭代计算公式:

$$\mathbf{u}^{(k+1)} = \mathbf{u}^{(k)} + \gamma \Delta \mathbf{u}^{(k)} \Delta J^{(k)}, \quad (6)$$

其中上标 (k) 表示第 k 次迭代, $\Delta \mathbf{u}^{(k)} = \{\Delta u_1, \Delta u_2, \dots, \Delta u_n\}^{(k)}$ 为第 k 次迭代时施加的扰动电压向量, γ 为增益系数。在实际应用中, 如使目标函数向极大方向优化, γ 取正值; 反之, γ 取负。在本文仿真中以施特列耳比极大化为目标, γ 取正值。

随机并行梯度下降算法的执行过程(第 k 次迭代时):

1) 随机生成扰动向量 $\Delta \mathbf{u}^{(k)} = \{\Delta u_1, \Delta u_2, \dots, \Delta u_n\}^{(k)}$, 各 Δu_i 相互独立且同为伯努利分布^[10], 即各分量幅值相等 $|\Delta u_j| = \sigma$, $\Pr(\Delta u_j = \pm \sigma) = 0.5$;

2) 使用(4)式, (5)式计算目标函数的变化量 $\Delta J^{(k)}$;

3) 利用(6)式更新控制参量, 进行第 $k+1$ 次迭代, 直至满足算法结束条件。

3 仿真结果与分析

依据仿真模型使用 Matlab 语言编写随机并行梯度下降算法的仿真程序, 变形镜不用于倾斜校正, 因此未考虑倾斜项。待校正像差由前 10 阶泽尼克多项式随机生成, 从 3 到 10 各阶系数如下:

$$\begin{aligned} Z(3) &= 1.223; & Z(4) &= 0.7; & Z(5) &= -0.5; \\ Z(6) &= 0.3; & Z(7) &= -0.4; & Z(8) &= -0.3; \\ Z(9) &= 0.3; & Z(10) &= -0.1; \end{aligned}$$

初始施特列耳比为 0.1。原始畸变波前和校正前远场光斑如图 3 所示。

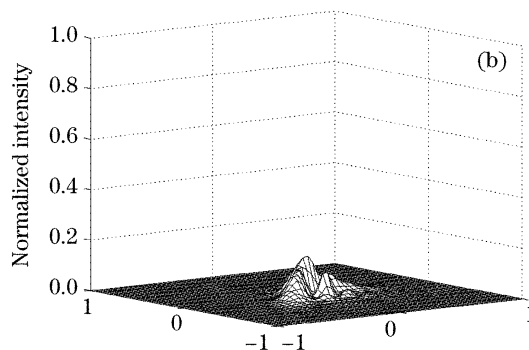
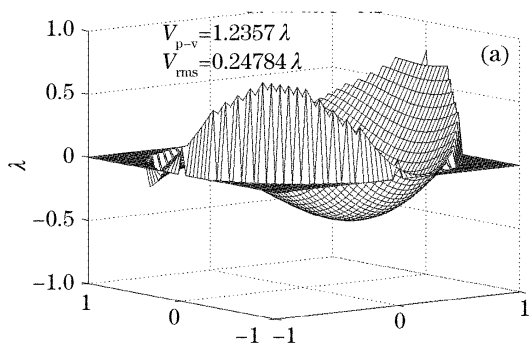


图 3 校正前波前(a)及远场分布(b)

Fig. 3 Wavefront (a) and farfield intensity distribution (b) before correction

3.1 算法收敛性验证

当使用不同的初始随机扰动向量时, 算法的收敛性会略有不同。增益系数、随机扰动幅度均固定, 随机运行程序 50 次的施特列耳比变化曲线如图 4。1000 次迭代后, 46 条收敛到全局极值 0.97 附近, 3 条分别收敛到 0.86、0.803 和 0.73, 一条未完全收敛, 证实了随机并行梯度下降算法陷入局部极值的可能, 同时也表明随机并行梯度下降算法至少能够以 90% 的概率收敛到全局极值。从虚线交点之间的曲线分布可以看出, 收敛到全局极值的 46 条曲线收敛速度又有不同。施特列耳比达到 0.8 时, 最快(左边交点)仅需 81 次, 最慢(右边交点)需 336 次。

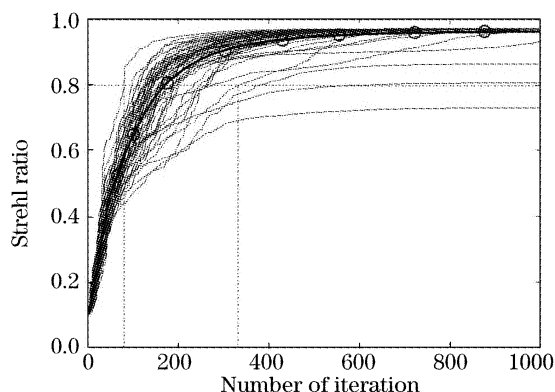


图 4 不同初始随机扰动向量的施特列耳比变化曲线
Fig. 4 Strehl ratio evolution curves of different initial perturbations

为得到更合乎统计意义的评价标准,把 50 条施特列耳比曲线取平均作为仿真结果,如图 4 中带圈的粗线所示,施特列耳比达 0.8 时需 171 次迭代。算法

收敛到全局极值时,残余波前及其远场光斑如图 5 所示可见随机并行梯度下降算法的收敛性及其对静态像差的校正效果都很好。

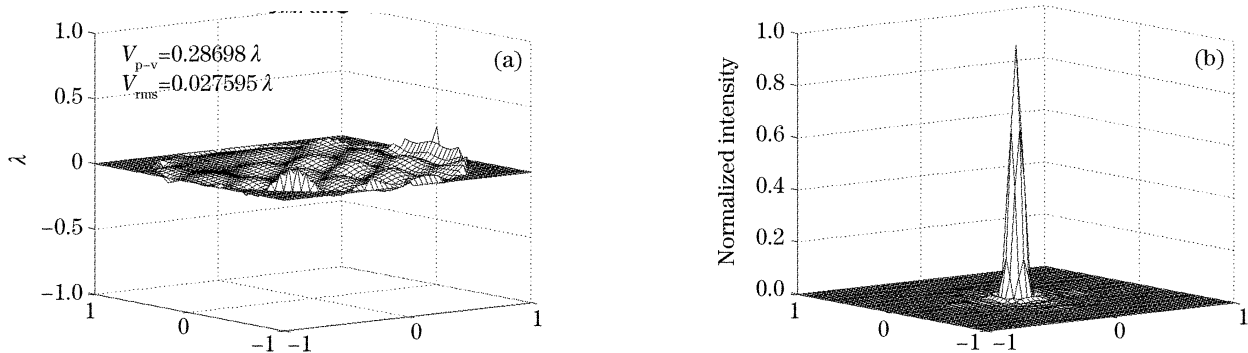


图 5 校正后波前(a)及远场分布(b)

Fig. 5 Wavefront (a) and farfield intensity distribution (b) after correction

3.2 对同一种畸变取不同扰动幅度和不同增益系数

图 6 给出了增益系数固定在 0.85、取不同扰动幅度时的施特列耳比变化曲线,曲线所标数据为随机扰动的幅值。从图 6 可以看出扰动幅值从 0.1 到 0.5 再到 1,收敛速度有逐步加快的趋势;而从 1 到 1.5 再到 2 收敛速度有逐步变慢的趋势。可见针对仿真中的畸变类型,在 $\gamma=0.85$ 时,合适的随机扰动幅值范围应在 0.5~1.5。

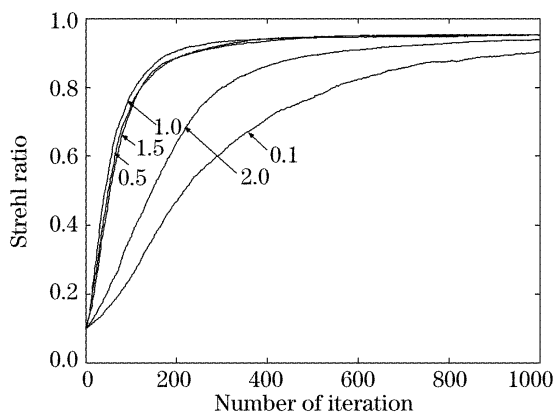


图 6 不同扰动幅度时的施特列耳比变化曲线

Fig. 6 Strehl ratio evolution curves for different perturbation amplitudes

图 7 给出了扰动幅度固定在 1、不同增益系数时的施特列耳比变化曲线,曲线所标数据为增益 γ 的大小。 γ 太小,收敛速度慢,如 $\gamma=0.08$ 时,1000 次迭代施特列耳比仅至 0.64; γ 过大,开始阶段收敛较快,但随后算法陷入局部极值,如 $\gamma=2.25$ 时施特列耳比收敛到 0.63; $\gamma=0.85$ 时,则是一种比较理想的情况,171 次迭代,施特列耳比达到 0.8,既保证了收敛速度又保证了系统校正能力。比较 $\gamma=1.85$ 和 $\gamma=2.25$ 时的施特列耳比曲线可以看出,随着 γ 增大施特列耳比曲线开始出现抖动,且随着 γ 的增大

抖动现象越明显。通过多次仿真得到的结果分析,可以认为增益系数最佳取值范围在 0.65~1 之间,而且这个范围同样适用于其它类型初始施特列耳比大于 0.1 的畸变。

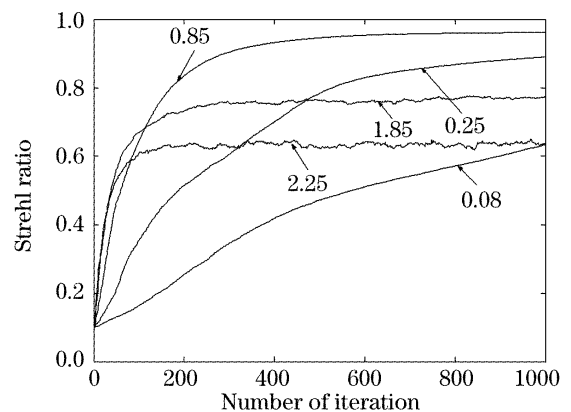


图 7 不同增益系数时的施特列耳比变化曲线

Fig. 7 Strehl ratio evolution curves for different gain coefficients

从以上分析可以看出,固定的增益系数存在着合适的扰动幅度范围,而固定的扰动幅度也存在合适的增益系数取值范围。因此就扰动幅度和增益系数而言,在具体应用中只需把其中一个固定在合适的范围内,调整另外一个就可以了。

3.3 固定增益、固定随机扰动幅度对不同程度畸变波前的适应情况

将图 3(a)中畸变放大或缩小一定倍数而得到具有不同初始施特列耳比的畸变,考察算法对这些畸变的校正能力。图 8 给出这些畸变的施特列耳比变化曲线,曲线所标数据为施特列耳比初始值,增益系数 γ 设定在 0.85,扰动幅度设定在 1。初始施特列耳比大于 0.1,算法收敛到 0.97 以上;小于 0.1,则收敛到相对较低的施特列耳比。表 1 为各曲线的

初始施特列耳比与算法结束的施特列耳比对比。

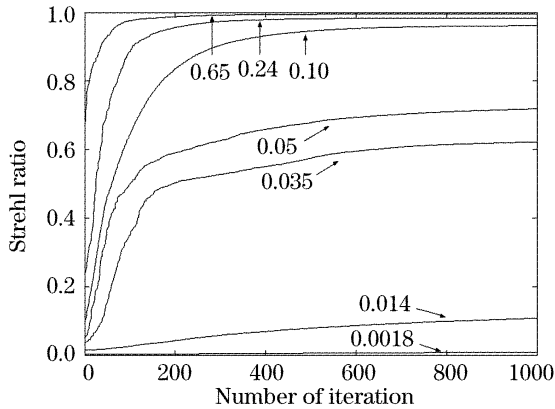


图 8 固定增益对不同畸变的适应情况

Fig. 8 Adaptation curves of constant gain for different aberrations

表 1 不同初始施特列耳比的提高倍数

Table 1 Amplification of different initial Strehl ratio

Original Strehl ratio	Final Strehl ratio	Amplification
0.653	0.995	1.51
0.24	0.98	4.08
0.1	0.97	9.7
0.05	0.74	14.8
0.035	0.64	18.28
0.014	0.11	7.85
0.0018	0.008	4.44

从表中最后一行数据可以看出,当波前畸变很大,初始施特列耳比降至千分之一量级时,采用随机

并行梯度下降算法的自适应光学系统仍具有一定的校正能力。

3.4 自适应增益

图 6~图 8 中的增益系数 γ 在算法迭代过程中不发生变化,即增益固定。从图 7 可以看出,随机扰动幅度固定在 1 时, γ 较大时,算法收敛速度较快,但很快收敛于局部极值; γ 较小时,算法收敛速度慢。根据这个结论,可以考虑在算法的迭代过程中进行 γ 的自适应调整。在迭代初始阶段,系统向着优化方向前进需要大的步长时选择大的增益,以增加收敛速度;而迭代中后期,选择小增益以获得更好的校正效果。

可以根据距当前最近的几次迭代信息(即参量的变化量以及参量的变化对目标函数的影响)来确定当前增益^[11]。第 k 次迭代时,增益系数 γ 的更新过程为

$$\gamma^{(k+1)} = \gamma^{(k)} + \alpha[\gamma_0 - \gamma^{(k)}] + \xi \bar{H}_1^{(k)} \bar{H}_2^{(k)} \bar{H}_3^{(k)}, \quad (7)$$

其中 $H_1^{(k)} = |\text{sign}(\Delta J_+^{(k)}) - \text{sign}(\Delta J_-^{(k)})|$, $H_2^{(k)} = |\Delta J^{(k)}|$, $H_3^{(k)} = \sum_{i=1}^l |J^{(k)} - J^{(k-l)}|$, 为降低噪声影响,将上述各式取一定迭代次数内的平均: $\bar{H}_i^{(k)} = \frac{1}{M} \sum_{m=0}^M H_i^{(k-m)}$, 其中 α 和 ξ 为常量。仿真中各参量分别

取值如下: $\alpha = 0.25, \xi = 30, \gamma_0 = 1.35, l = 5, M = 5$ 。取固定增益中收敛速度和收敛性能最好的情况(增益为 0.85)与自适应增益的施特列耳比变化曲线进行比较。结果如图 9(a)。

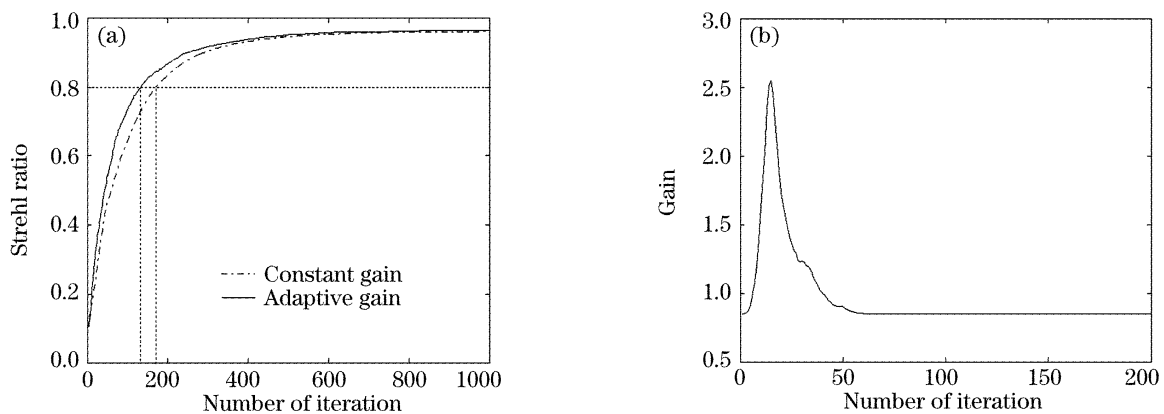


图 9 自适应增益与固定增益的比较。(a) 施特列耳比变化曲线, (b) 增益变化曲线

Fig. 9 Comparison of adaptive gain with constant gain. (a) Strehl ratio evolution curve, (b) gain evolution curve

固定增益时,施特列耳比达到 0.8 需 171 次迭代;自适应增益只需 116 次迭代,收敛速度明显提高。图 9(b)给出了自适应增益的前 200 次迭代增益变化曲线,初始值设定在 0.85。增益主要变化发生在前 50 次,最大值时为 2.54,增加将近 3 倍,80 次后固定

在 0.85。自适应增益实现了初始迭代所需的大增益,算法趋于收敛时所需小增益,从而改进了算法的收敛速度。同时,将自适应增益应用到不同类型像差以及不同大小像差的校正实验中,得到类似仿真结果,说明本文的研究结果具有一定普遍性。

4 结 论

本文介绍了自适应光学系统随机并行梯度下降控制算法的基本原理。以 32 单元变形镜为校正器,建立了采用随机并行梯度下降算法的自适应光学系统仿真模型。仿真结果验证了随机并行梯度下降算法的收敛性及该类型的自适应光学系统对静态波前畸变的校正能力,发现通过增益系数的自适应调整可以改进算法收敛速度,从而为基于系统性能指标直接优化技术的自适应光学系统的进一步研究提供了理论依据。

参 考 文 献

- 1 Robert K. Tyson. *Principle of Adaptive Optics* [M]. San Diego: Academic Press, 1991
- 2 Richard A. Muller, Andrew Buffington. Real-time correction of atmospherically degraded telescope images through image sharpening[J]. *J. Opt. Soc. Am.*, 1974, **64**(9): 1200~1210
- 3 Jiang Wenhan, Huang Shufu, Wu Xubin. Hill-climbing adaptive optics wavefront correction system[J]. *Chin. J. Lasers*, 1988, **A15**(1): 17~21 (in Chinese)
姜文汉,黄树辅,吴旭斌. 爬山法自适应光学波前校正系统[J]. *中国激光*, 1988, **A15**(1): 17~21
- 4 Zhang Yudong, Jiang Wenhan. Multiple maxima problem in image-sharpening adaptive optics systems[J]. *Chin. J. Lasers*, 1990, **A17**(4): 193~197 (in Chinese)
张雨东,姜文汉. 像锐化自适应光学系统中的多极值问题[J]. *中国激光*, 1990, **A17**(4): 193~197
- 5 M. A. Vorontsov, G. W. Carhart. Adaptive phase-distortion correction based on parallel gradient-descent optimization [J]. *Opt. Lett.*, 1997, **22**(12): 907~909
- 6 Mikhail A. Vorontsov, Gary W. Carhart. Adaptive optics based on analog parallel stochastic optimization; analysis and experimental demonstration[J]. *J. Opt. Soc. Am. A.*, 2000, **17**(8): 1440~1453
- 7 R. Mukai, K. Wilson, V. Vilnrotter. Application of genetic and gradient descent algorithms to wave-front compensation for the deep-space optical communications receiver [R]. *The Interplanetary Network Progress Report*, 2005. 1~21
- 8 S. Zommer, E. N. Ribak, S. G. Lipson *et al.*. Simulated annealing in ocular adaptive optics[J]. *Opt. Lett.*, 2006, **31**(7): 1~3
- 9 R. El-Agmy, H. Bulte, A. H. Greenaway *et al.*. Adaptive beam profile control using a simulated annealing algorithm[J]. *Opt. Exp.*, 2005, **13**(16): 6085~6091
- 10 James C. Spall. Multivariate stochastic approximation using a simultaneous perturbation gradient approximation [J]. *IEEE Trans. on Automatic Control*, 1992, **37**(3): 332~341
- 11 Thomas Weyrauch, Mikhail A. Vorontsov, Thomas G. Bifano *et al.*. Microscale adaptive optics: wave-front control with a μ -mirror array and a VLSI stochastic gradient descent controller [J]. *Appl. Opt.*, 2001, **40**(24): 4243~4253