

Simulated Annealing for Control of Adaptive Optics System

Huizhen Yang¹ and Xinyang Li²

¹*Huaihai Institute of Technology,*

²*Institute of Optics and Electronics, Chinese Academy of Science
China*

1. Introduction to adaptive optics system

Many optical systems, such as imaging systems or laser communication systems, suffer performance degradation due to distortions in the optical wave-front. An optical wave propagates through an optically inhomogeneous medium such as the atmosphere, differences in the index of refraction along the propagation path cause variations in the speed of light propagation, which lead to phase distortions. Adaptive Optics (AO) techniques are often used to compensate these static or dynamic aberrations of a light beam after propagation through a distorting medium (Hardy, 1998). Although originally proposed for astronomical telescopes in 1953 (Babcock, 1953), adaptive optics did not become a reality until the 1970s, when it was developed for national defence applications, specifically laser beam compensation and satellite imaging. It consists of using an active optical element such as a deformable mirror to correct the instantaneous wavefront distortions. These are measured by a device called a wavefront sensor which delivers the signals necessary to drive the correcting element. The first adaptive optics system able to sharpen two-dimensional images was built at Itek by Hardy and his co-workers (Hardy et al, 1977).

AO provides a means to perform real-time correction of aberrations imposed on light waves as they travel from the source to the imaging system. While AO has its roots the field of astronomy it is currently used in a wide variety of medical, military and industrial applications. The papers by Milonniand (Milonni, 1999) and Parenti (Parenti, 1992) provide an excellent introduction to the use of AO in Astronomy. A comprehensive review of the medical and industrial applications of AO can be found in the technology tracking report by Greenaway and Burnett (Greenaway, & Burnett, 2004).

The most common conventional adaptive optics systems (Fig. 1) is implemented with a wave-front corrector to correct the distorted wave-front, a wave-front sensor to measure the aberrations present in the incoming beam, and a feedback control algorithm to link these two elements in real time. Although the technique based on rapid wave-front measurement has been found useful in astronomical applications, this approach to the control problem is much difficult to be used in situations where wave-front distortions can not be measured

directly, for example in atmospheric laser communications or anisoplanatic imaging conditions.

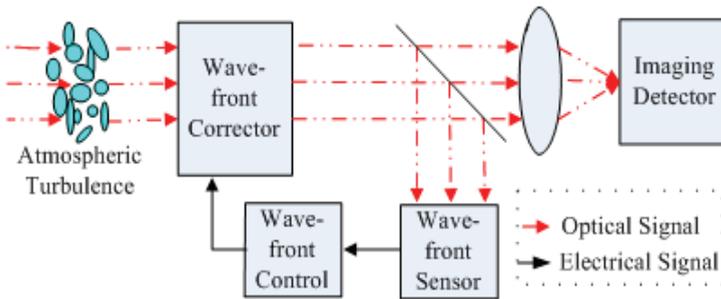


Fig. 1. Block diagram of conventional adaptive optics systems.

When a performance metric can be defined, stochastic optimization methods provide an alternative approach to the control problem that does not require the use of any a prior knowledge of a system model (Muller & Buffington, 1974). A common strategy used by model-free optimization techniques in adaptive optics systems (Fig.2) is to consider the performance metric as a function of the control parameters and then use certain optimization algorithm to improve the performance metric, which means that a wave-front sensor is no longer necessary. More and more researchers on adaptive optics system control are attaching importance to this kind of control technique because of its simpleness in system architecture and adaptability to the complicated conditions.

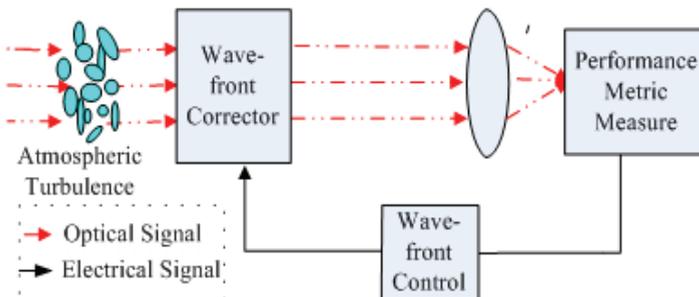


Fig. 2. Block diagram of adaptive optics systems without a wave-front sensor.

An appropriate optimization control algorithm is the key to correcting distorted wave-fronts successfully for this kind of adaptive optics system. Because most systems are multi-channel and real time systems, the control algorithm must meet the following requirements: rapid convergence so that the adaptive optics systems can keep up with changes of distorted wave-fronts, high correction capability for wave-front aberrations, and being carried out easily. The optical problem can be mapped onto a model for crystal roughening that served as a motivation to implement the Simulated Annealing algorithm (SA). In other words, SA is an optimization algorithm designed to find the extremum of a certain cost function, which can be regarded as the energy of the system and will be called hence forth the energy

function. The energy function is also called as the system performance metric in application of adaptive optics.

2. Simulated Annealing

Simulated annealing (Kirkpatrick et al 1983) is based on the physical annealing process by which a solid is heated to a temperature close to its melting point, after which it is allowed to cool slowly so as to relieve internal stresses and non-uniformities. The aim is to achieve a structure with long-range order that is as close as possible to the ground-state configuration. Presenting an optimization technique, SA can: (a) process cost functions possessing quite arbitrary degrees of nonlinearities, discontinuities, and stochasticity; (b) process quite arbitrary boundary conditions and constraints imposed on these cost functions; (c) be implemented quite easily with the degree of coding quite minimal relative to other nonlinear optimization algorithms; (d) statistically guarantee finding an optimal solution.

Generally, SA algorithm consists of three functional relationships per iteration: probability density of state-space of control parameters to create perturbation vector $\Delta u^{(k)} = \{\Delta u_i\}^{(k)}$; acceptance probability $p^{(k)} = \exp(-\Delta J^{(k)} / T^{(k)})$ to adjudge whether the new solution is accepted, which is also called as the Metropolis criterion (Metropolis et al 1953); and schedule of "annealing" in annealing-time steps $T^{(k)}$. In this text, we use the standard simulated annealing.

In order to investigate correction ability of the simulated annealing for adaptive optics system, we compare it with some other stochastic parallel optimization algorithms, such as Stochastic Parallel Gradient Descent (SPGD)(Vorontsov & Carhart, 1997), Genetic Algorithm (GA)(Goldberg, 1989), and Algorithm Of Pattern Extraction (Alopex)(Harth & Tzanakou, 1974). These algorithms optimize control parameters in a parallel way, which can accelerate the convergence of algorithms, and have some stochastic specialty, which can help the algorithm escape away from local extremums to some extent. The main advantage of these stochastic parallel optimization algorithms over traditional adaptive optics correction algorithms is that wavefront sensing is no longer required. The reduction in complexity, cost, and size is extremely beneficial. Even though the absence of a wavefront sensor makes the algorithm less efficient, advantages from increased speed, parallelism, and simplicity make it attractive in certain applications. These algorithms are both model-free as well as independent of the deformable mirror characteristics. This independence, as well as being a simple straightforward algorithm to implementation, allows a great deal of latitude in system design. The reason why we compare these algorithms with the simulated annealing is that these algorithms have ever been used to control the adaptive optics system and obtained some valuable research results, such as GA (Yang et al 2007), Alopex (Zakynthinaki & Saridakis, 2003), SA(Zommer et al 2006), SPGD (Vorontsov & Carhart, 2000).

3. The model of AO system with the simulated annealing

3.1 AO System Model

The AO system model is shown in Fig. 3, where $\varphi(r)$ is the distorted wavefront, $u(r)$ is the compensation phase, $\phi(r) = \varphi(r) + u(r)$ is the residual phase, J is the performance metric and $u = \{u_1, u_2, \dots, u_{61}\}$ is the control signal of actuators of 61-element deformable mirror.

$\varphi(r)$ and $u(r)$ are continuous functions ($r = \{x, y\}$ is a vector in the plane orthogonal to the optical axes). The AO system mainly includes a 61-element deformable mirror to correct the wave-front aberrations $\varphi(r)$, an imaging system to record the focal spot, a performance metric analyzer to calculate the system performance metric J from the data of focal spot, the simulated annealing algorithm to produce control signals $u = \{u_1, u_2, \dots, u_{61}\}$ for the 61-element deformable mirror according to changes of the performance metric J .

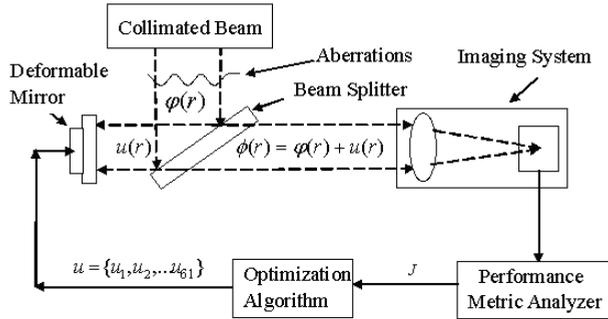


Fig. 3. Block diagram of simulation.

3.2 Specifications of 61-element deformable mirror

The phase compensation $u(r)$, introduced by the deformable mirror, can be combined linearly with response functions of actuators:

$$u(r) = \sum_{j=1}^{61} u_j S_j(r) \tag{1}$$

Where u_j is the control signal and $S_j(r)$ is the response function of the j 'th actuator. On the basis of real measurements, we know the response function of 61-element deformable mirror actuators is Gaussian distribution approximately (Jiang & etal 1991):

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

Where (x_j, y_j) is the location of the j 'th actuator, ω is the coupling value between actuators and is set to 0.08, d is the distance between actuators, and α is the Gaussian index and is set to 2. Fig. 4 gives the location distribution of 61-element deformable mirror actuators. The circled line in the figure denotes the effective aperture and the layout of all actuators is triangular.

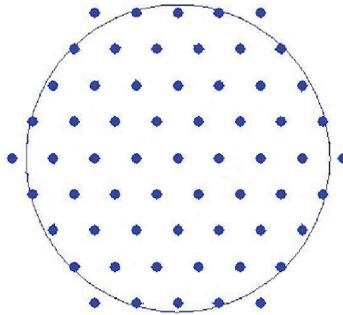


Fig. 4. Distribution of actuators location of 61-element deformable mirror.

3.3 Atmospheric turbulence conditions

We use the method proposed by N. Roddier, which makes use of a Zernike expansion of randomly weighted Karhunen-Loeve functions, to simulate atmospherically distorted wavefronts (Roddier, 1990). Considering that the low-order aberrations (tilts, defocus, astigmatism, etc) have the most significant impact on image quality, we use the first 104 Zernike polynomial orders. Different phase screens generated according to this method are not correlated to each other and represent the Kolmogorov spectrum. The phase screens are defined over 128×128 pixels which is also the grid of the wave-front corrector and don't include the tip/tilt aberrations. The tip/tilt aberrations are usually controlled by another control loop and are considered as being removed completely in our simulation.

Atmospheric turbulence strength for a receiver system with aperture size D can be characterized by the following two parameters: the ratio D/r_0 and the averaged Strehl Ratio (SR) of phase fluctuations, where r_0 is the turbulence coherence length and SR is defined as the ratio of the maximum intensities of the aberrated point spread function and the diffraction-limited point spread function. Phase screens of different atmospheric turbulence strength can be obtained through changing r_0 in the simulation program. The correction capability of the AO system based on simulated annealing is analyzed when D/r_0 is 10 and corresponding averaged SR is about 0.1.

3.4 Considerations for the performance metric J

Possible measures of energy spread in the focal plane that can be used as the energy function of simulated annealing J are:

- (1). The Strehl ratio(SR):

$$SR = \max(I(r)) / \max(I^{dl}(r)), \tag{3}$$

where $I(r)$ is the intensity distribution of focal plane with the turbulence and $I^{dl}(r)$ is the diffraction-limited laser intensity distribution achievable in the absence of the turbulence. This quantity is not difficult to measure on the focal plane. It does not seem to be very informative because it does not account for the whole intensity distribution. We use the SR as the reference of correction capability in the text.

(2). The encircled energy (EE):

$$EE = \int_{\mathfrak{R}} I(r) dr \quad (4)$$

where \mathfrak{R} is a region with a laser hot spot where maximum energy is to be collected. This metric also does not necessarily take into account the whole intensity distribution. In addition, it depends on the choice of area \mathfrak{R} .

(3). Image sharpness (IS_{mn}) used in various active imaging applications:

$$IS_{mn} = \int \left| \frac{\partial^{m+n} I(r)}{\partial^m x \partial^n y} \right|^2 dr, \quad (m+n) = 0, 1, \dots \quad (5)$$

This quantity is relatively simple to measure, and is intuitively appealing since smaller tighter intensity distributions have wider spatial frequency spectrum and, therefore, larger sharpness.

(4). The mean radius (MR):

$$MR = \frac{\int |r - \bar{r}| I(r) d^2 r}{\int I(r) d^2 r}, \quad \bar{r} = \frac{\int r I(r) d^2 r}{\int I(r) d^2 r} \quad (6)$$

where \bar{r} is the intensity distribution centroid. MR can be easily measured either by a single photodetector with a special mask attached to it or by postprocessing a matrix detector output. This measure appears to be the most attractive one for it gives straightforward mathematical meaning to the idea of energy spread, it is nonparametric, and it accounts for the whole intensity distribution.

For imaging applications metrics 1-3 are proven (Muller & Buffington, 1974) to attain their global maxima for the diffraction-limited image. It is clear that the global minimum of the MR metric corresponds to the smallest energy spread. It is also possible to invent other functions, including vector functions, as well as to create compound cost functions with additional penalty terms. All these possibilities deserve thorough investigation. However, only the MR metric is used in our simulations.

The relationship between the performance metric J and control parameters $\{u_j\}$ is

$$J \propto J[\phi(r)] = J[\varphi(r) + u(r)] \quad (7)$$

$J = J(u_1, u_2, \dots, u_{61})$ can be considered as the non-linear function of 61 control signals because the $\varphi(r)$ keeps unchanged during a relatively short time. In real applications, we can get the performance metric data from the photoelectric detector, for example from a CCD or a pinhole, and then define different performance metrics based on different applications.

3.5 Descriptions of other stochastic parallel optimization algorithms

(1). SPGD (Vorontsov & Carhart, 2000) control is a “hill-climbing” technique implemented by the direct optimization of a system performance metric applied through an active optical component. Control is based on the maximization (or, with equal complexity minimization) of a system performance metric by small adjustments in actuator displacement in the mirror array. SPGD requires small random perturbations $\Delta u = \{\Delta u_1, \Delta u_2, \dots, \Delta u_{61}\}$ with fixed amplitude $|\Delta u_j| = \sigma$ and random signs with equal probabilities for $\Pr(\Delta u_j = \pm\sigma) = 0.5$ (Spall, 1992), to be applied to all 61 deformable mirror control channels simultaneously. Then for a given single random Δu , the control signals are updated with the rule:

$$u^{(k+1)} = u^{(k)} + \gamma \Delta u^{(k)} \Delta J^{(k)} \tag{8}$$

where γ is a gain coefficient which scales the size of the control parameters. Note that non-Bernoulli perturbations are also allowed in the algorithm, but one must be careful that the mathematical conditions (Spall, 1992) are satisfied.

SPGD follows the rule during the iteration of algorithm:

$$\Delta J^{(k)} = J_+^{(k)} - J_-^{(k)} \tag{9}$$

where

$$J_+^{(k)} = J(u^{(k)} + \Delta u^{(k)}) \tag{10.a}$$

$$J_-^{(k)} = J(u^{(k)} - \Delta u^{(k)}) \tag{10.b}$$

From the introduction to SPGD, we know there are only two parameters to be adjusted in algorithm: one is perturbation amplitude σ and the other is gain coefficient γ .

(2). GA is a kind of evolutionary computation, which represents a class of stochastic search and optimization algorithms that use a Darwinian evolutionary model, adopts the concept of survival of the fittest in evolution to find the best solution to some multivariable problem, and includes mainly three kinds of operation in every generation: selection, crossover and mutation. GA works with a population of candidate solutions and randomly alters the solutions over a sequence of generations according to evolutionary operations of competitive selection, mutation and crossover. The fitness of each population element to survive into the next generation is determined by a selection scheme based on evaluating the performance metric function for each element of the population. The selection scheme is such that the most favourable elements of the population tend to survive into the next generation while the unfavourable elements tend to perish.

The control vector $\{u_j\}$ was considered as the individual to be evolved and the performance metric is called as fitness function. After the initial population is made according to the roulette selection principle, excellent individuals are selected from the population with a ratio r_s . Then new individuals are obtained by randomly crossing the chromosomes of the old individuals with a probability of P_c . Finally, some chromosome positions of individuals are mutated randomly with a mutation rate of P_m for introducing a new individual. By going

through above process, GA will gradually find the optimum mirror shape that can yield the best fitness. Parameter r_s , P_c and P_m are set at 0.2, 0.65 and 0.65 accord to the corresponding reference value (Chen, etal 1996). The population size N and the number of evolving generation L are needed to adjust.

(3). Alopex is a stochastic correlative learning algorithm which updates the control parameters by making use of correlation between the variations of control parameters and the variations of performance metric without needing (or explicitly estimating) any derivative information. Since its introduction for mapping visual receptive fields (Harth & Tzanakou, 1974), it has subsequently been modified and used in many applications such as models of visual perception, pattern recognition, and adaptive control, learning in neural networks, and learning decision trees. Empirically, the Alopex algorithm has been shown to be robust and effective in many applications.

We used a two timescale version Alopex, called as 2t-Alopex (Roland, etal 2002). The control signals are updated according to the rules:

$$u^{(k+1)} = u^{(k)} + \eta \Delta u^{(k)} \quad (11)$$

$$\Delta u^{(k)} = \begin{cases} 1 & \text{probability } p^{(k)} \\ -1 & \text{probability } 1-p^{(k)} \end{cases} \quad (12)$$

$$p^{(k)} = p^{(k-1)} + \theta(\beta^{(k)} - p^{(k-1)}) \quad (13)$$

$$\beta^{(k)} = 1/(1 + \exp(\Delta u^{(k)} \Delta J^{(k)} / \eta T^{(k)})) \quad (14)$$

$T^{(k)}$ in equation (12) is a "temperature" parameter updated every M iterations(for a suitably chosen M) using the following annealing schedule:

$$T^{(k)} = \begin{cases} T^{(k-1)} & \text{if } k \text{ is not a multiple of } M \\ \frac{\eta}{M} \sum_{k'=k-M}^k |\Delta J^{(k')}| & \text{otherwise} \end{cases} \quad (15)$$

Where η and θ are step-size parameters such that $\eta = o(\theta)$. There are at least two parameters to be adjusted: η and θ .

4. Results and Analysis

We perform the adaptation process over 100 phase realizations. The averaged evolution curves, the standard deviation evolution curves of the metric and corresponding SR evolution curves are the recorded simulation results.

4.1 Selection of different algorithm parameters

Every algorithm has its rational limit of parameters for a given application. We select the most optimal parameters of every algorithm through large numbers of simulation tests when D/r_0 is 10.

In SA, the adjustment coefficient δ and the cooling rate λ are main factors for convergence rate and correction effect and we set δ was 0.15 and λ was 0.98. The key parameters of Alopex are step-size parameters η and θ , and we set η was 0.03 and θ was 0.55.

The amplitude σ and the gain coefficient γ are two main factors which affect convergence rate and correction capability of SPGD. For a fixed σ , there exists an optimal range for γ . Too small γ will cause too slow convergence rate, while too big γ will make the algorithm trap into local extrimums and the evolution curve of performance metric appears dither. We find the effective range of σ is within 0.01-1.5 for SPGD. We fixed the same σ at 0.2 for SPGD, r is set at 5

After probability parameters r_s, P_c and P_m in GA are established on the basis of experience, the convergence rate is affected by the population size N and the number of evolution generation L . For the same correction effect, L will be fewer if N is bigger, while the algorithm will need more times of evolution when N is smaller. If GA not only converge rapidly but also has good correction effect, it's necessary to balance N and L . We set N at 100 and L at 500 in simulation.

4.2 Adaptation process

In order to converge completely, we set the iteration number of algorithms respectively. SA is set at 4000 times, GA 500 generations, Alopex 4000 times and SPGD 1500 times. The averaged evolution curves, the standard deviation (SD) evolution curves of the metric MR over 100 phase realizations and corresponding averaged SR evolution curves are given in Fig. 5, Fig. 6, Fig. 7 and Fig. 8, in which the value of MR is normalize by that of diffraction-limited focal plane and the standard deviation(SD) is calculated as follows:

$$SD = \frac{\langle (J - \langle J \rangle)^2 \rangle^{1/2}}{\langle J \rangle} \tag{16}$$

Fig. 5 to Fig. 8 show simulation results when SA, GA, SPGD and Alopex are use to control the AO system respectively. Averaged curves of MR are given in Fig. 5(a) to Fig. 8(a), in which averaged evolution curves are normalized to be 1 in the optimal case. Corresponding standard deviation curves over 100 different phase realizations and averaged SR curves are presented in Fig. 5(b) to Fig. 8(b) and Fig. 5(c) to Fig. 8(c). All MR curves have converged after complete iterations in Fig. 5(a) to Fig. 8(a). From Fig. 5(b) to Fig. 8(b), we can see that SA, GA and Alopex have relatively smaller standard deviations than SPGD, which shows that SA, GA and Alopex have stronger adaptability to different turbulence realizations than SPGD. The averaged SR's of these four different control algorithms are very close to each other in Fig. 5(c) to Fig. 8(c), which indicates SA, GA, SPGD and Alopex have almost equal correction ability under $D/r_0 = 10$.

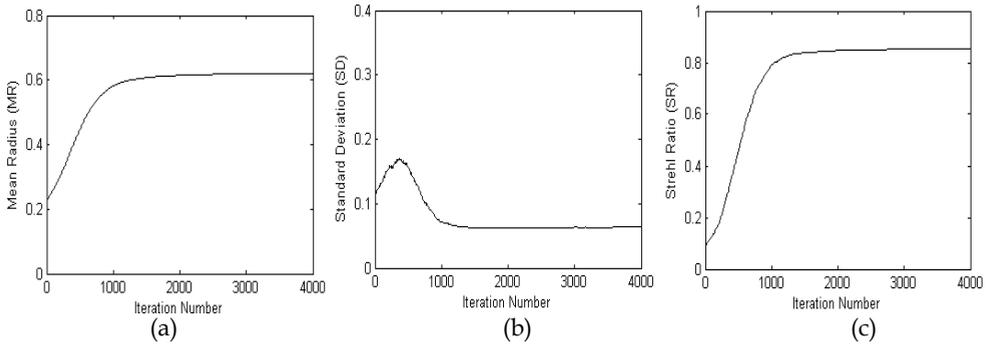


Fig. 5. Adaptation process of adaptive optics system when SA is used as the control algorithm. (a): averaged curves of MR, (b): the standard deviation curve of MR over 100 different phase realizations and (c): averaged SR curves during 4000 iterations.

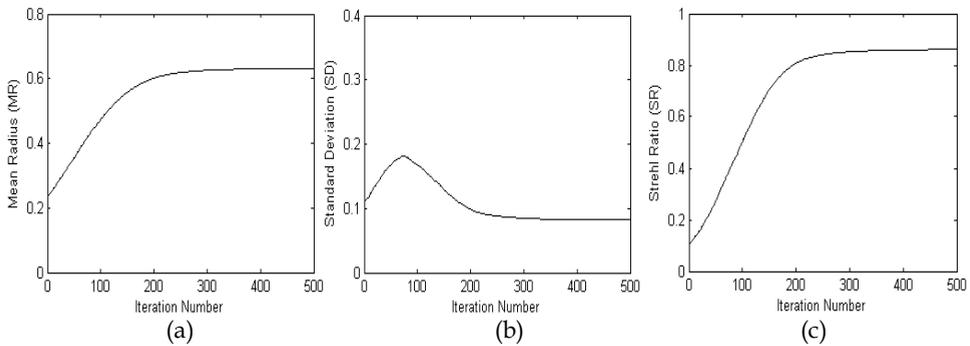


Fig. 6. Adaptation process of adaptive optics system when GA is used as the control algorithm. (a): averaged curves of MR, (b): the standard deviation curve of MR over 100 different phase realizations and (c): averaged SR curves during 500 generations.

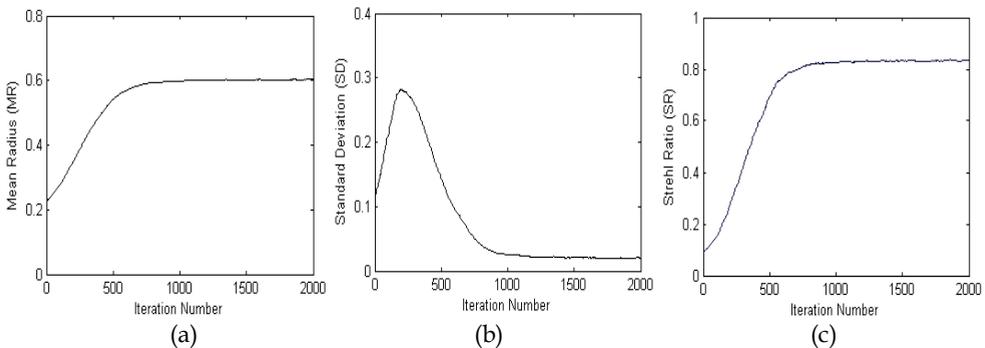


Fig. 7. Adaptation process of adaptive optics system when SPGD is used as the control algorithm. (a): averaged curves of MR, (b): the standard deviation curve of MR over 100 different phase realizations and (c): averaged SR curves(c) during 2000 iterations.

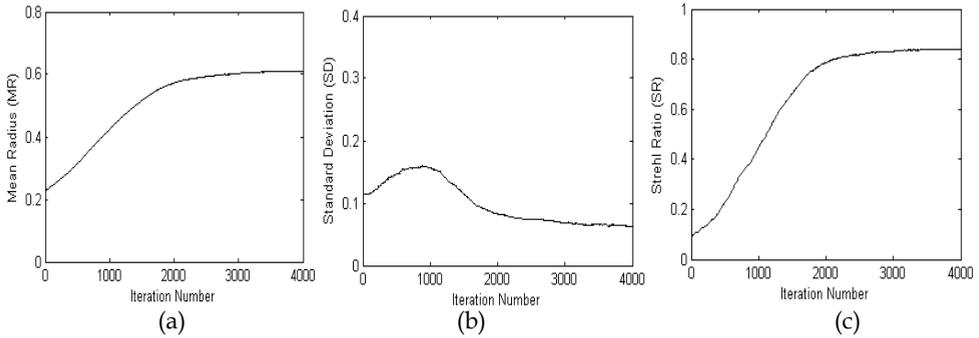
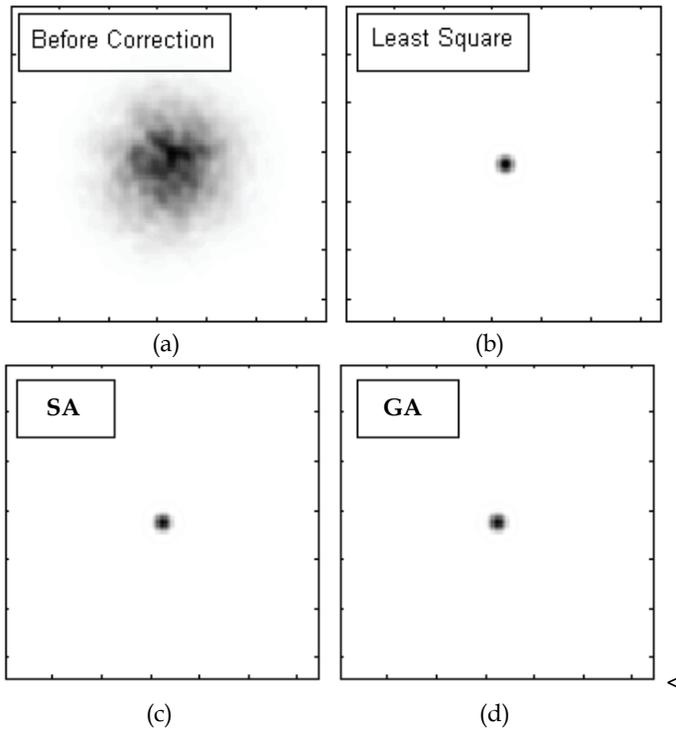


Fig. 8. Adaptation process of adaptive optics system when Alopex is used as the control algorithm. (a): averaged curves of MR, (b): the standard deviation curve of MR over 100 different phase realizations and (c): averaged SR curves during 4000 iterations.

Fig. 9 gives the averaged focal spot when SA, GA, SPGD and Alopex are use as the control algorithm of the AO system respectively. For purposes of comparison, we also fit the 61-element deformable mirror figure to the phase screens using least squares to obtain the best correction achievable with the given 61-element DM.



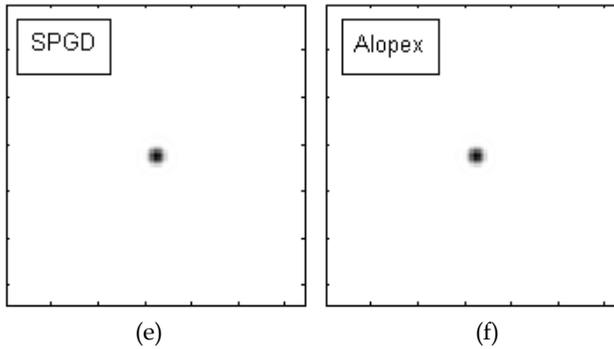


Fig. 9. Comparison of focal spots before correction (a) and after correction with SA (c), with GA (d), with SPGD (e) and with Alopex (f); (b) is the averaged focal spot of the residual wave-front with the least squares fitting.

From Fig. 9, we can get these four different algorithms have strong ability to atmospheric turbulence when D/r_0 is 10. Compared with the least squares fitting, they almost obtain the best correction achievable for the 61-element DM.

4.3 Analysis of averaged convergence speed

The convergence speed is an important criterion on which the algorithm can be applied to real-time adaptive optics system. Fig. 5 (a) to Fig. 8(a) give the averaged curves of MR over 100 different phase realizations. The abscissa in Fig. 5(a) to Fig. 8(a) is the iteration number of algorithm for SA, SPGD and Alopex and the number of evolution generation for GA. It seems that GA has the rapidest speed from the averaged curves of MR because of its fewer evolution generation. This result is not true because the number of small perturbations sent to the system per iteration is different for different algorithms. From the introduction to the basic idea of several algorithms in section 3.5, we know that SA and Alopex need one perturbation per iteration; SPGD needs two perturbations per iteration and GA needs 100 perturbations per generation. Note that the number of perturbation in GA bears on the number of the population size. To reduce the number of perturbation, one can choose a relative small population size but the convergence of system will need more generations. The related analysis can refer to section 4.1. We use the number of small perturbations not the number of iteration or generation to estimate the averaged convergence speed of different algorithms.

Consulting results in Fig. 5(a) to Fig. 8(a) and above analysis, we make use of the number of small perturbations needed by achieving the 80% of the range of MR during the adaptation process under control of different algorithms. Corresponding data are in Table 1.

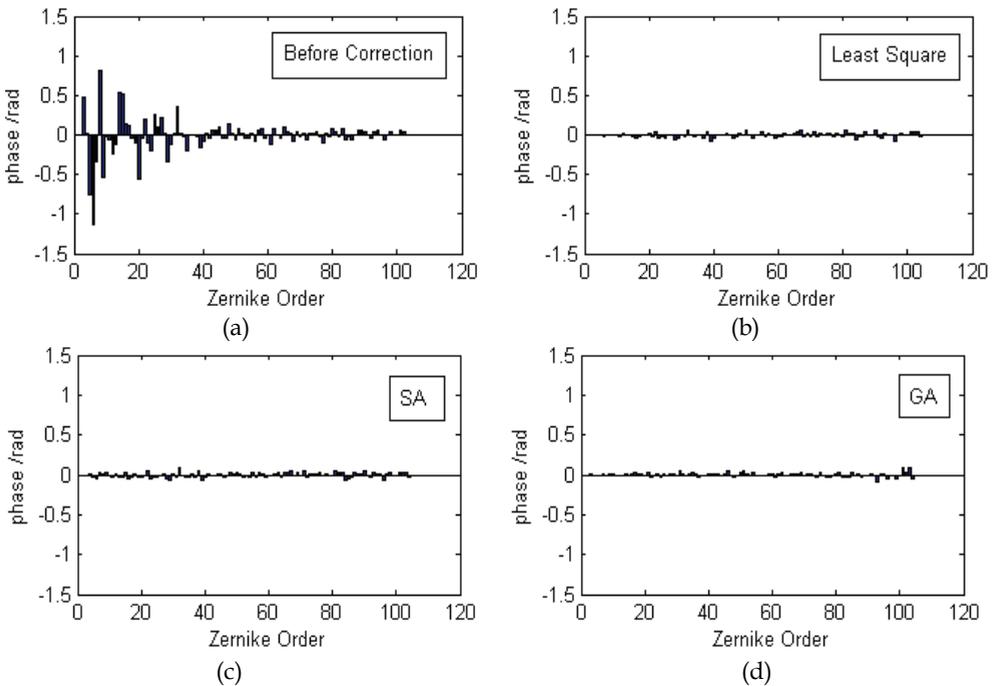
	Value of 80% MR Range	Iterations or Generations	Perturbations
SA	0.54	767	767
GA	0.548	143	143*100=14300
SPGD	0.524	464	464*2=928
Alopex	0.532	1609	1609

Table 1. Comparison of the number of small perturbations sent to the system when the adaptive optics system achieves the 80% of the range of MR during the adaptation process under control of different algorithms.

The value of MR is 0.54 for SA, 0.548 for GA, 0.524 for SPGD and 0.532 for Alopex respectively in Table 1. These data show the range of MR of different algorithms are close to each other because their start value of MR is the same. The number of iterations or generations is 767 for SA, 143 for GA, 464 for SPGD and 1609 for Alopex but the number of small perturbation is 767 for SA, 14300 for GA, 928 for SPGD and 1609 for Alopex respectively. These data show GA is the slowest algorithm and the number of perturbations is almost as 20 times as that of SA, 15 times as SPGD and 9 times as Alopex, while SA is the fastest algorithm because of its the fewest perturbations. The advantage of GA is that it is far more likely that the global extremum will be found, as the data shown in second column in Table 1; the disadvantage is that it often takes a long time to converge. Above simulation results express relative differences of these algorithms in convergence rate, which can offer us some references in choosing stochastic parallel optimization algorithm for real applications.

4.4 Zernike order and wavefront of the same single frame phase screen

Fig. 10 gives Zernike coefficients 3-104 decomposed from the same phase screen when SA, GA, SPGD and Alopex are used as control algorithm of adaptive optics system respectively. Corresponding wavefronts are shown in Fig. 11.



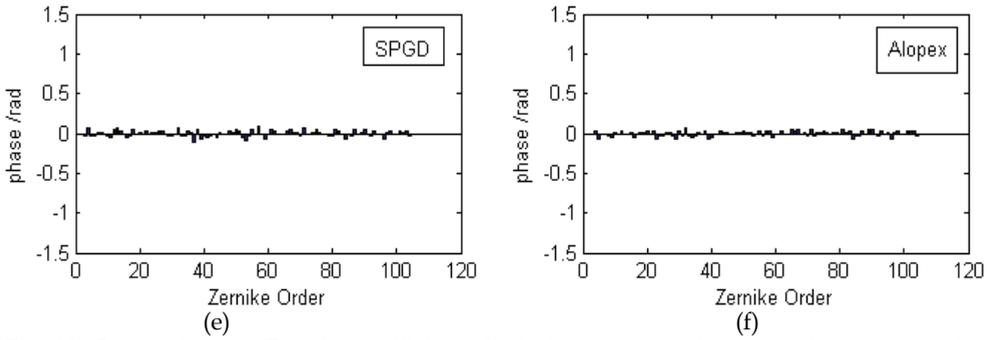
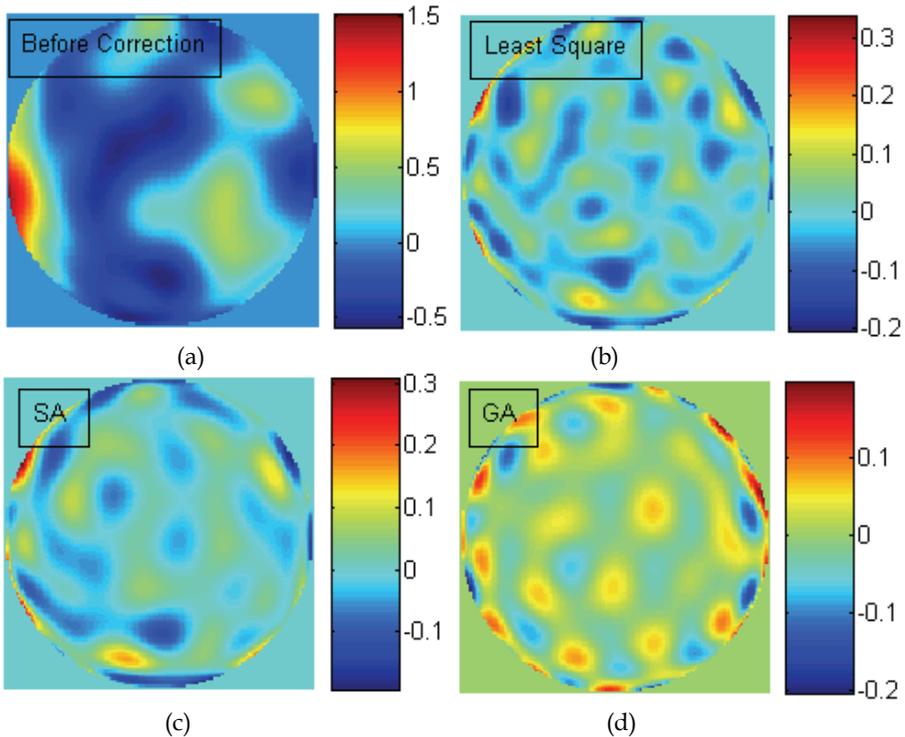


Fig. 10. Comparison of Zernike coefficients 3-104 before correction (a) and after correction with SA (c), GA (d), SPGD (e) and Alopex (f) ; (b) is the Zernike coefficients of the residual wave-front with the least squares fitting.

We also fit the DM figure to the phase screen using least squares to obtain the best correction achievable with the given 61-element DM. The fitting results are also shown in Fig. 10 and Fig. 11. The unit in Fig. 10 is rad and wavelength λ in colour bar of Fig. 11.



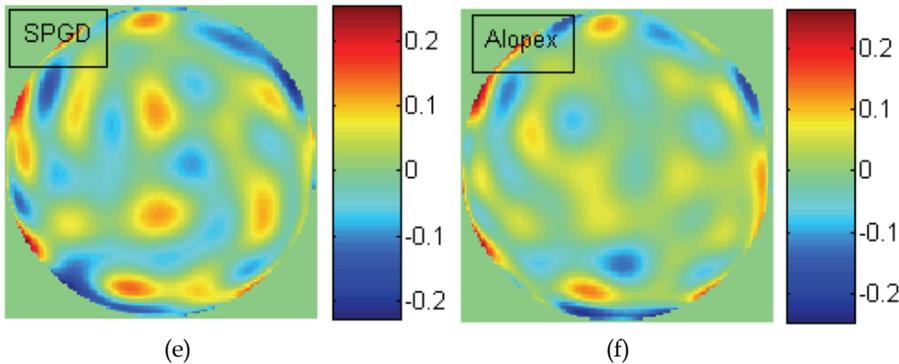


Fig. 11. Comparison of wavefronts before correction (a) and after correction with SA (c), GA (d), SPGD (e) and Alopex (f); (b) is the residual wave-front with the least squares fitting. The unit of colour bar is wavelength λ .

From Fig. 10 and Fig. 11, we can obtain that these four different algorithms have strong ability to atmospheric turbulence when D/r_0 is 10. Compared with the least squares fitting, they almost obtain the best correction achievable for the 61-element DM.

5. Conclusion

We presented basic principles of Simulated Annealing, Genetic Algorithm, Stochastic Parallel Gradient Descent, and Algorithm of pattern extraction in control application of adaptive optics system. Based on above stochastic parallel optimization algorithms, we simulated an adaptive optics system with a 61-element deformable mirror and compared these algorithms in convergence speed, correction capability.

From section 4.2 and 4.4, we can get these four different algorithms have strong ability to atmospheric turbulence when D/r_0 is 10. Compared with the least squares fitting, they almost obtain the best correction achievable for the 61-element DM. The correction effect of GA is litter better than other algorithms and SA is the secondly better algorithm. But SA, GA and Alopex have stronger adaptability to different turbulence realizations than SPGD because of its relatively big standard deviation.

From section 4.3, we can conclude SA is the fastest and GA is the slowest in these algorithms. The number of perturbation by GA is almost as 20 times as that of SA, 15 times as SPGD and 9 times as Alopex. GA begins with a population of candidate solutions (individuals) and evolves towards better solutions through techniques inspired by evolutionary biology (such as natural selection or mutation). Perhaps the main problem of GA is the time cost of it. The algorithm may converge, and it may be a guaranteed global extremum, however, if this requires excessive a mounts of computer equipment or if it takes an unreasonable length of time to provide the solution, then it will not be suitable. But if the real-time is not required by adaptive optics system in some special application fields, GA is the best choice.

In real applications, after the deformable mirror is established, the correction time of AO system is mainly affected by the read-out and computation time of performance metric,

which occupies the most part time of control algorithm. This point is the same in both simulation test and real AO systems. Above simulation results express relative differences of these algorithms in convergence rate, which can offer us some references in choosing stochastic parallel optimization algorithm for specific applications.

In conclusion, we can get that each algorithm has its advantages and disadvantages from above simulation results and discussions. For static or slowly changing wavefront aberrations, these algorithms all have high correction ability. For dynamic wavefront aberrations, convergence rates of these algorithms are slow relative to the change rate of atmosphere turbulence. They can be applied to real-time wavefront correction if being combined with high speed photo-detector, high speed data processing and high response frequency wave-front corrector. More research is necessary to this problem. Such as, how about these algorithms applied to much stronger turbulence conditions and much more elements deformable mirror or other kinds of wavefront corrector?

6. References

- J. W. Hardy (1998). *Adaptive optics for astronomical telescopes*, Oxford University Press, New York
- H. W. Babcock (1953). The possibility of compensating astronomical seeing, *Pub. Astr. Soc. Pac.* Vol. 65, pp. 229-236
- J. W. Hardy, J. E. Lefebvre, & C. L. Koliopoulos (1977). Real-time atmospheric compensation, *J. Opt. Soc. Am.* Vol. 67, pp. 360-369
- P. W. Milonni (1999). Adaptive optics for astronomy, *American Journal of Physics*, Vol. 67, No. 6, pp. 476-485
- P. R. Parenti (1992). Adaptive optics for astronomy, *The Lincoln Laboratory Journal*, vol. 54, No. 1, pp. 93-113
- A. Greenaway, & J. Burnett (2004). *Technology tracking : Industrial and medical applications of adaptive optics*, Institute of Physics Publishing Ltd, London
- R. A. Muller, & A. Buffington (1974). Real-time correction of atmospherically degraded telescope images through image sharpening, *J. Opt. Soc. Am. A.* Vol. 64, No. 9, pp. 1200-1210
- S. Kirkpatrick, C. D. Gelatt, & M. P. Vecchi (1983). Optimization by simulated annealing, *Science*, Vol. 220, pp. 671-680
- N. Metropolis, A. W. Rosenbluth, & M. N. Rosenbluth (1953). Equation of state calculations by fast computing machines, *J. Chem. Phys.* Vol. 21, pp. 1087-1092
- M. A. Vorontsov, & G. W. Carhart (1997). Adaptive phase-distortion correction based on parallel gradient-descent optimization, *Opt. Lett.* Vol. 22, No. 12, pp. 907-909
- D. E. Goldberg (1989). *Genetic algorithms in search, optimization and machine learning*, 1st Edition, Addison-Wesley Publishing Company, Boston
- E. Harth, & E. Tzanakou (1974). Alopex: a stochastic method for determining visual receptive fields. *Vis. Res.* , Vol. 14, pp. 1475-1482
- P. Yang, Y. Liu, W. Yang, & et al (2007). An adaptive laser beam shaping technique based on a genetic algorithm. *Chinese Optics Letters*, Vol. 5, No. 9, pp. 497-500
- M. S. Zakyntinaki, & Y. G. Saridakis (2003). Stochastic optimization for adaptive real-time wave-front correction, *Numerical Algorithms*, Vol. 33, pp. 509-520

- S. Zommer, E. N. Ribak, S. G. Lipson, & et al (2006). Simulated annealing in ocular adaptive optics, Vol. 31, No. 7, pp. 1-3
- M. A. Vorontsov, & G. W. Carhart (2000). Adaptive optics based on analog parallel stochastic optimization: analysis and experimental demonstration, *J. Opt. Soc. Am. A*, Vol. 17, No. 8, pp. 1440-1453
- W. H. Jiang, N. Ling, X. J. Rao, & et al (1991). Fitting capability of deformable mirror, *Proceedings of SPIE*, Vol. 1542, pp. 130-137, ISBN :9780819406705
- N. Roddier (1990). Atmospheric wavefront simulation using Zernike polynomials, *Optical Engineering*, Vol. 9, No. 10, pp. 1174-1180
- J. C. Spall(1992). Multivariate stochastic approximation using a simultaneous perturbation gradient approximation, *IEEE Trans. On Automatic Control*, Vol. 37, pp. 332-341
- G. L. Chen, F. X. Wang, & Z. Zhuang (1996). *Genetic algorithm and its application (in Chinese)*, Post & Telecom Press, Beijing
- P. S. Sastry, M. Magesh, & K. P. Ummikrishnan (2002). Two timescale analysis of the alopex algorithm for optimization, *Neural Computation*, Vol. 14, pp. 2729-2750

