

# Simulated annealing in ocular adaptive optics

S. Zommer, E. N. Ribak, S. G. Lipson, and J. Adler

*Department of Physics, Technion-Israel Institute of Technology, Haifa 32000, Israel*

Received October 3, 2005; accepted November 19, 2005; posted December 21, 2005 (Doc. ID 65139)

We present what is to our knowledge a first hardware realization of a simulated annealing algorithm in an adaptive optics system designed to image the retina of the human eye. The algorithm is applied to the retinal image itself without the need for wavefront sensors in the system. We find that this optimization algorithm can be an alternative to the traditional Hartmann-Shack sensing. We also compare the simulated annealing algorithm to the stochastic parallel gradient descent algorithm. © 2006 Optical Society of America  
OCIS codes: 010.1080, 170.3890.

Adaptive optics is a method designed to correct images deformed by nonideal optical systems in real time. In general, the distortion of the wavefront from an ideal point source is measured by a wavefront sensor, and a deformable mirror is used to compensate the aberrations and return the wavefront to a plane or spherical wave. An image of the object is then taken through the same corrected optics. This Letter utilizes methods that omit wavefront sensing hardware and software from the reconstruction process. Such methods use stochastic or directed algorithms to find the extremum of a certain sharpness function, thereby correcting the image without any information on the wavefront. Theoretical work and initial simulations<sup>1</sup> have shown that the optical problem can be mapped onto a model for crystal roughening that served as a motivation to implement the simulated annealing algorithm (SA).

Another random and iterative method is phase diversity, deduced from images taken at zero and a small defocus.<sup>2</sup> We did not pursue this avenue because of ocular speckle effects, even in polychromatic light.

The 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<sup>3</sup> and will be called henceforth the energy function. It 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 nonuniformities. The aim is to achieve a structure with long-range order that is as close as possible to the ground-state configuration.

The algorithm is a stochastic algorithm that generates random states and accepts not only the “good states,” which decrease the cost function (in case of minima search), but occasionally also accepts some “bad states,” which increase the cost function and can help the algorithm climb out from an occasional local minimum.

The basic probability distribution that governs a state's acceptance is proportional to the control parameter  $T$ , which is analogous to the temperature of a physical system, and to the difference between the previous state energy  $E_1$  and the new state energy  $E_2$ . The probability acceptance condition, also called the Metropolis criterion, is  $P = \exp[-(E_1 - E_2)/T]$  when  $(E_1 - E_2) < 0$ , and 1 otherwise. The algorithm

randomly chooses one control parameter and, according to the assigned temperature, changes its value. The new energy function is measured and the control parameter's new value is accepted by using the probability criterion above. The temperature is lowered each step  $i$  to  $T_{i+1} = \gamma T_i$ , where  $\gamma$  is the cooling rate factor.

A second search algorithm used in this research is the stochastic parallel gradient descent (SPGD) algorithm. This algorithm is basically a fast version of the well-known steepest descent algorithm,<sup>4</sup> capable of reducing the processing time by a factor of  $\sqrt{N}$ , where  $N$  is the number of control parameters or degrees of freedom (in our case, the number of deformable mirror actuators). It achieves this by applying small random perturbations to all control parameters simultaneously, and evaluating the gradient with two measurements, before and after the changes. Let  $\delta \mathbf{x}_i = \{\delta x_{i1}, \delta x_{i2}, \dots, \delta x_{iN}\}$  be the simultaneously applied small random perturbations on the control parameters  $\mathbf{x}$  at step  $i$ . The resulting difference in the energy function multiplied by the perturbation vector is an approximation to the true gradient,  $\delta \mathbf{x}_i \delta f \approx \delta \mathbf{x}_i [f(\mathbf{x}_i + \delta \mathbf{x}_i) - f(\mathbf{x}_i)]$ . Consequently, the control parameters are updated according to  $\mathbf{x}_{i+1} = \mathbf{x}_i - \alpha \delta \mathbf{x}_i \delta f(\mathbf{x}_i)$ , where  $\alpha$  is a positive weight coefficient for a minimum search and negative for a maximum search.

The two algorithms were first implemented numerically by computer on a simulated adaptive optics system. We have modeled the mirror as an apparatus composed of 57 square elements arranged in a circular formation on a  $15 \times 15$  grid. We restricted each element to values in the range of  $(-\lambda/2, \lambda/2]$  corresponding to its height  $h$  with discretization of  $\lambda/200$ . As the energy function for the optimization algorithms we have chosen the irradiance at the center of the mirror's point spread function (PSF), corresponding to the Strehl ratio.

Both algorithms performed roughly the same, reaching the global maximum in approximately 1000 steps. However, the two algorithms have different processing times. The SA requires only one energy evaluation per step whereas the SPGD requires two such evaluations. These energy evaluations are the most time-consuming operations, compared to calculating the next control parameters and updating them. Consequently, each SPGD step is twice as long

as the SA step and when the two algorithms are compared timewise, the apparent advantage of the SPGD algorithm is negated. Nevertheless the SPGD algorithm could again claim superiority as the number of actuators increase.

Aside from the Strehl criterion we have also used the sharpness function  $\Sigma I^2$  introduced by Muller and Buffington.<sup>5</sup> However, this energy function adds degeneracy to the state space with regard to tilt, as the central peak might shift around. Simulations show that the SA tends to find solutions with a high degree of tilt in them as opposed to solutions obtained with the SPGD algorithm, even though both start from a random state, where on average there is no initial tilt. SPGD searches for the closest maximum of the energy function, which will naturally be close to zero tilt. The SA, however, samples a wider state space and can easily find itself in a degenerate maximum state containing tilt.

Both algorithms were also implemented in an adaptive optics system designed to image the human retina. The deformable mirror used was an OKO (Flexible Optical B.V., Delft, The Netherlands) membrane mirror with 59 actuators.<sup>6</sup> The most limiting factor when using this mirror is its limited stroke that sometimes hindered a full search. Figure 1 shows the results of the SA on an artificial retina at the back of a glass eye model. The energy function used by the algorithm is the sharpness evaluated on a small portion of the frame (Figs. 1a–1d). The energy (Fig. 1e) converges in approximately 400 steps.

To speed up convergence of both algorithms we have changed the search space from 59 individual actuators to the 59 eigenmodes of the mirror. These eigenmodes were calculated by obtaining the actuators' response matrix and applying a singular value decomposition.<sup>6</sup> The results are shown in Fig. 2 where the algorithms were implemented on an artificial eye. Both algorithms managed to improve the image dramatically in both search spaces. However, if one looks at the energy graphs, it is evident that the SA found a better solution, whereas SPGD got trapped in a local minimum. Furthermore, it is evident that with the mode search scheme the SA converged faster to a superior solution.

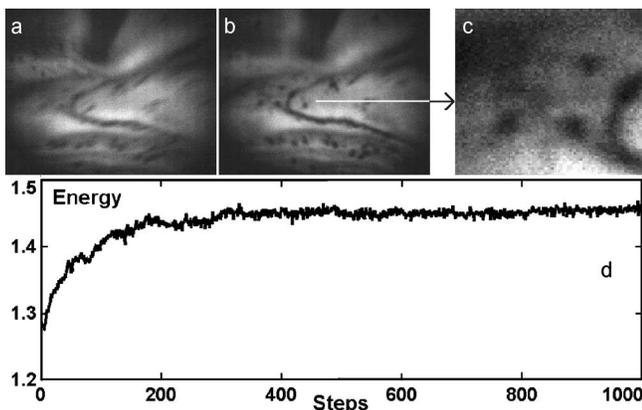


Fig. 1. Typical result of the SA algorithm. Initial and final frames (a and b). (c) Corresponding sections used to calculate energy. (d) Energy graph.

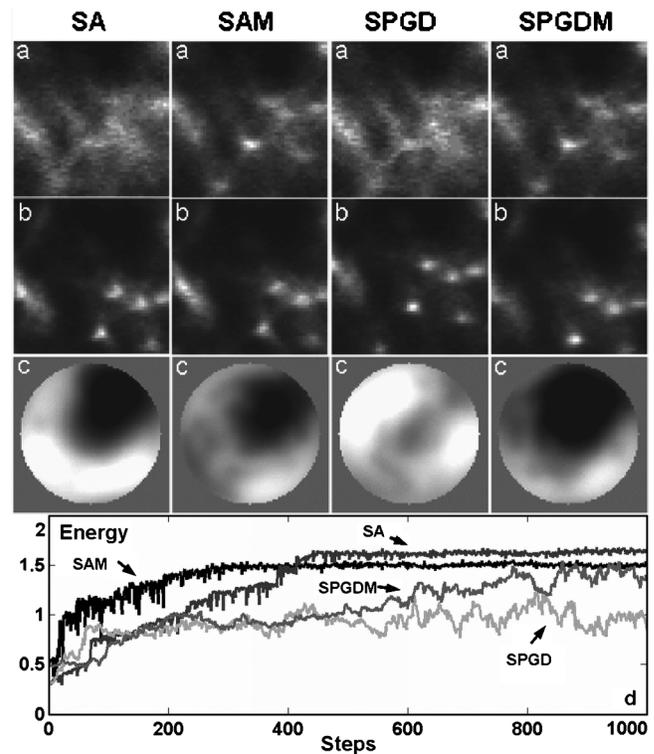


Fig. 2. Comparison of algorithms on a glass eye. SAM and SPGDM are the modal searches of simulated annealing and stochastic parallel gradient search, respectively. a, Initial images with all algorithms operating on the same retinal scene of  $10\ \mu\text{m}$  glass spheres. b, Final images. c, Final mirror configurations. d, Comparison of the rise in energy (sharpness).

The actuators' elevations (Fig. 2c) show, as in the simulations, the tilt that is present in the SA final state and missing in the SPGD final state. The same tilt can be observed in the retinal images as well. A closer look at the final frames reveals that the SPGD final state is somewhat better than the rest of the final frames, regardless of its inferior energy value. This illustrates the difficulty of choosing the best frame section to analyze and the correct energy function.<sup>7</sup>

Owing to the SA insensitivity to tilt, more often than not the algorithm found solutions that increase the energy by adding light into the frame, at the expense of the true image sharpening. This problem might be avoided if one is to use a different energy function such as spatial frequency analysis<sup>8</sup> rather than the irradiance-dependent sharpness function. Alternatively, if the illuminated spot size is much smaller than the full frame, its shift will not change the total light intensity.

Having achieved results in simulation and artificial eyes, we commenced experimenting on human subjects and obtained preliminary results (Fig. 3). As a reference source we used a superluminescent diode at  $0.8\ \mu\text{m}$ , whose maximum raw output was  $100\ \mu\text{W}$ , and after splitting the light into the eye always below  $40\ \mu\text{W}$  (we used an uneven beam splitter to capture as many retinal photons as possible). By analyzing the time scales in the system together with the laboratory and simulation results, it can be shown that

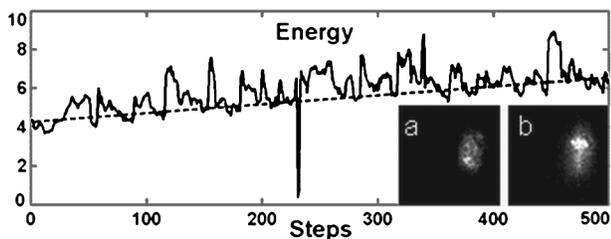


Fig. 3. Energy graph of a human aberrated eye (notice the blink recovery), and a, initial and b, final PSFs. The LED irradiance on the cornea was  $10 \mu\text{W}$ . The total time was 40 s, and can easily be reduced.

the SA can serve as a good alternative to the traditional wavefront sensing method.

There are four main time scales that govern any adaptive optics speed, including our system: (i) acquisition time of the frame, here at the standard video rate of 40 ms; (ii) wavefront calculation, or here energy evaluation of a subframe of  $100 \times 100$  pixels: 0.17 ms; (iii) actuator command calculation, for the SA: 7 ms; (iv) voltage application to the actuators; in our case, for 59 actuators, a total of 0.4 ms. As a result, the whole loop can be closed within 47.4 ms. In our experiments each SA loop was actually 80 ms long due to an additional, useless frame acquisition each step (we used a standard frame grabber on a personal computer running Windows XP).

We experimented with the SA and reached the optimum in  $\sim 400$  steps in artificial eyes by using a mode search, and  $\sim 500$  steps by using a poke search. This corresponds to a theoretical  $\sim 19$  s correction time, much longer than the required<sup>9</sup> 0.5 s. Figure 3 shows the SA performance on a human eye PSF obtained in a similar time. Despite the slow time scale of the system, there is definite improvement in the PSF that is validated in the energy graph.

The total loop time can be easily reduced if one uses a single photodiode to evaluate the Strehl ratio of a limited spot rather than a full CCD image. In this mode<sup>9</sup> all the light that passes through a hard or soft aperture is integrated. The acquisition step [(i)

above] can then be reduced to  $1\text{--}50 \mu\text{s}$ , with the total time for one algorithm step  $< 0.5$  ms, and with 1000 iteration steps,  $< 0.5$  s. In comparison, when a Hartmann–Shack wavefront sensor was used, the PSF was reconstructed in approximately 2–3 iterations, where each iteration was 0.2 s long (again due to the unnecessary loss of frames). This was checked both on a glass eye and on a living eye.

In conclusion, we simulated and ran experiments for comparison of optimization and direct wavefront sensing. The SA achieved good results in converging into a global optimum in simulations. Additional experiments are needed to evaluate its reconstruction accuracy and speed in comparison to the SPGD algorithm and a full adaptive optics system. However, the latter accuracy is only as good as its calibration, whereas in iterative algorithms no calibration is required.

We thank I. Perlman and E. Zemel for participation in some live experiments, the Israeli Ministry of Science for partial support, and the SHARP-EYE European Network. E. N. Ribak's e-mail address is eribak@physics.technion.ac.il.

## References

1. E. N. Ribak, J. Adler, and S. G. Lipson, *J. Phys. A* **23**, L809 (1990).
2. R. G. Paxman, T. J. Schulz, and J. R. Fienup, *J. Opt. Soc. Am. A* **7**, 1072 (1992).
3. S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, *Science* **220**, 671 (1983).
4. M. A. Vorontsov, G. W. Carhart, M. Cohen, and G. Cauwenberghs, *J. Opt. Soc. Am. A* **17**, 1440 (2000).
5. R. M. Muller and A. Buffington, *J. Opt. Soc. Am.* **64**, 1200 (1974).
6. E. J. Fernandez and P. Artal, *Opt. Express* **11**, 1056 (2003).
7. J. R. Fienup and J. J. Miller, *J. Opt. Soc. Am. A* **20**, 609 (2003).
8. M. A. Vorontsov, G. W. Carhart, D. V. Pruidze, J. C. Ricklin, and D. G. Voelz, *J. Opt. Soc. Am. A* **13**, 1456 (1996).
9. H. Hofer, P. Artal, B. Singer, J. L. Aragon, and D. R. Williams, *J. Opt. Soc. Am. A* **18**, 497 (2002).