AdOpt: Analog VLSI Stochastic Optimization for Adaptive Optics

Marc Cohen
R. Timothy Edwards
Gert Cauwenberghs

Mikhail Vorontsov
Gary Carhart

Johns Hopkins University
Electrical and Computer Engineering
Baltimore, MD 21218
{marc,tim,gert}@bach.ece.jhu.edu

Army Research Laboratory
Intelligent Optics Laboratory
Adelphi, MD 20783
{vorontsov,gary}@iol.arl.mil

Abstract

Phase distortion in wavefront propagation is one of the key problems in optical imaging and laser optics applications. We present a hybrid VLSI and optical system for real-time adaptive phase distortion compensation. The system operates "model-free", independent of the specifics of the distorting optical medium and the compensation control elements. Our VLSI system implements stochastic parallel perturbative gradient descent/ascent so that we achieve fast optimization of the chosen performance metric to achieve real-time compensation. We include experimental results of the hybrid VLSI-optical system demonstrating successful operation for a laser-beam focusing/defocusing task.

Introduction

Many optical systems, such as imaging systems or laser beam communication systems, exhibit degradation due to phase distortions in the optical wavefront. As 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 propagation which leads to *phase distortions (aberrations)*.

The classical problem of compensating phase distortion originates in astronomy. The wave eminates from a point source (a distant star or planet) and travels through a phase distorting medium (Earth's atmosphere). This assumption of a point source is the foundation of the most widely used adaptive optics control algorithm, *phase conjugation* [1, 2, 3]. A significant problem with phase conjugation techniques is that phase modulation can not be directly measured and must be reconstructed from intensity information gathered from wavefront sensors.Calculating the phase requires extensive and time-consuming computations that prevent real-time phase compensation and significantly in-crease adaptive system cost and complexity.

There exists an alternative and more general approach that neither assumes a point source nor requires determining phase from intensity. In an adaptive system based on the direct optimization of a system performance metric, the control algorithm can be made independent of the system model ("model-free" or "blind" optimization) [4]. A schematic illustration of this method for atmospheric imaging systems is shown in Figure 1. The measured scalar quality metric $J = J(\mathbf{u})$ is a function of the control parameters $\mathbf{u} = \{u_1, u_2, \cdots, u_N\}$ of the wavefront corrector. In our present system, the u_n are voltages applied to electrodes of a 127 element liquid crystal spatial light modulator (SLM). Our system is scalable so that SLM or MEMs devices having many more control elements can also be used.

Methods

Among model-free optimization techniques that are in the literature, stochastic parallel perturbative gradient descent/ascent [5] is perhaps the most promising for adaptive optics applications [6, 7]. At worst, we can achieve a speedup of $\sqrt{N}$ over traditional sequential optimization techniques [8]. The parallel perturbation technique is well-suited to mixed-mode VLSI implementation [9, 10, 11].

In parallel stochastic optimization, a random ensemble of perturbations Δu_i are applied to all N control parameters simultaneously. The original state of the system is restored after each perturbation and the parameters u_i adapt along the direction of the perturbation vector by an amount proportional to the measured ΔJ and in a direction that descends/ascends the quality metric surface. In the simplest scenario, the perturbed metric ΔJ is given in terms of the one-sided perturbations Δu_n as

$$
\begin{aligned}
\Delta J \;=\; & J\left(u_1 + \Delta u_1, u_2 + \Delta u_2, \cdots, u_N + \Delta u_N\right) \\
& - J\left(u_1, u_2, \cdots, u_N\right).
\end{aligned} \tag{1}
$$
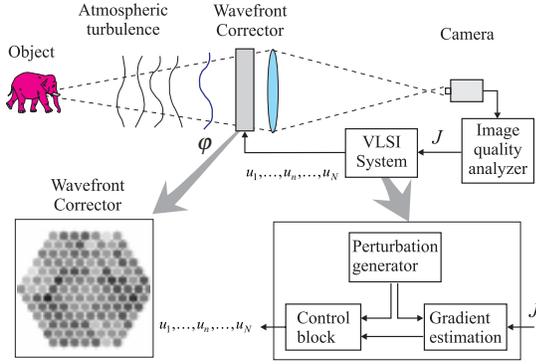
Figure 1: Block diagram for an adaptive hybrid VLSI-optical imaging system based on image quality metric optimization using stochastic parallel perturbative gradient descent.

from which, by Taylor series expansion,

$$\frac{\Delta J}{\Delta u_n} \approx \frac{\delta J}{\delta u_n} + \sum_{l \neq n}^{N} \frac{\delta J}{\delta u_l} \cdot \frac{\Delta u_l}{\Delta u_n}. \qquad (2)$$

When the perturbations Δu_n are random and statistically independent the second term in Equation (2) reduces in expectation to zero, yielding a good approximation to the true gradient for large N. This is true when Δu_n are Bernoulli distributed, of fixed amplitude, but random in sign, $\Delta u_n = \pm \sigma$ [5, 11].

More accurate gradient approximation is obtained with two-sided perturbations using a differential measurement under complementary perturbations $\pm \frac{1}{2} \Delta u_n$, yielding the following simplified implementation:

$$\begin{aligned}
u_n^- &= u_n^{(m)} - \sigma \pi_n^{(m)}, \quad \forall\, n, m \\
u_n^+ &= u_n^{(m)} + \sigma \pi_n^{(m)} \\
\Delta u_n &= u_n^+ - u_n^- = 2\sigma \pi_n^{(m)} \\
\Delta J &= J^+ - J^- \\
u_n^{(m+1)} &= u_n^{(m)} - \gamma \frac{\Delta J}{\Delta u_n}
\end{aligned} \qquad (3)$$

where the perturbation signals $\pi_n^{(m)}$ are generated from a Bernoulli random distribution:

$$\pi_n^{(m)} = \pm 1, \quad \Pr(\pi_n^{(m)} = +1) = 0.5, \quad \forall\, n, m \qquad (4)$$

with uncorrelated statistics across parameters and over time:

$$E(\pi_n^{(m)} \pi_p^{(q)}) = \delta_{np}\, \delta_{mq}, \quad \forall\, n, m, p, q. \qquad (5)$$

Considering that the pertubation amplitudes $|\Delta u_n|$ are identical (2σ) for all perturbations, the specification of the update (3) is further simplified:

$$u_n^{(m+1)} = u_n^{(m)} - \gamma' \pi_n^{(m)} \Delta J, \qquad (6)$$

where the constant $\gamma' = \gamma/2\sigma$ absorbs both learning rate and perturbation strengths.

Implementation

A mixed-mode (hybrid analog-digital) VLSI system interfaces with a liquid-crystal SLM to control in parallel all 127 elements which adjust the wavefront phase profile. Each chip controls 19 elements, 7 chips are needed for the 127 element SLM. For a detailed description of the mixed-mode VLSI architecture and circuits see Edwards et. al. [12]. Figure 2 shows a micrograph of a single chip fabricated through the MOSIS foundry service in $1.2\mu m$ CMOS technology.
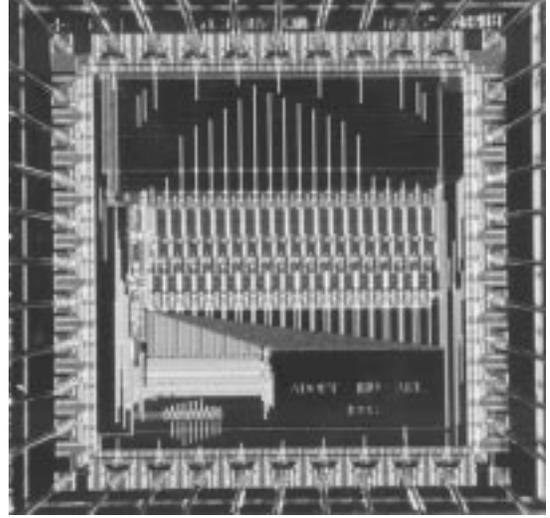
Figure 2: Micrograph of the 19 parallel channel mixed-mode VLSI stochastic gradient descent optical controller, a 2.2 × 2.25 sq. mm chip fabricated in 1.2 μm CMOS technology.

Each of the 19 channels consists of four main parts:

1. an analog memory circuit which maintains each control parameter $u_n^{(m)}$
2. a circuit which perturbs the parameter
3. a circuit which adapts the parameter
4. an output driver specific to the application

A feedback shift-register on each chip produces the 19 Bernoulli distributed pseudorandom perturbations in parallel at each iteration step.

On-chip CMOS circuitry performs parallel stochastic perturbative gradient descent/ascent of the externally supplied optimization metric J, e.g. a direct measure of image/beam quality. Parallel random perturbations of the parameters are generated locally, and the resulting differential performance measure ΔJ is locally correlated with the perturbations to generate parallel parameter updates, implementing a random-direction, stochastic approximation version of gradient descent/ascent. Additional on-chip circuitry provides

for liquid-crystal AC modulation and adaptive biasing of the mean phase of all 127 SLM elements. To set and maintain the aperture averaged mean phase value during adaptive operation, the control update algorithm is modified to include an additional penalty term to the metric J accounting for the drift of the mean as follows [6, 7]:

$$u_n^{(m+1)} = u_n^{(m)} - \gamma' \pi_n^{(m)} \Delta J - \eta(\overline{u}^{(m)} - u_0), \quad (7)$$

where the control parameter $\overline{u}^{(m)} \equiv \frac{1}{N} \sum_{n=1}^{N} u_n^{(m)}$ corresponds to the aperture averaged phase, and u_0 is a reference voltage with $u_0 = 2\,V$ corresponding to the middle of the liquid crystal phase *vs.* voltage characteristic. The quantity $\overline{u}^{(m)}$ is calculated on-and-across chips using compact charge-mode circuits.

Our system architecture is hierarchical; a personal computer is used to provide digital timing control signals to the VLSI system so that the chips can perform the parallel stochastic gradient descent/ascent on a fast timescale. The PC is equipped with A/D and D/A cards (Computer Boards CIO-DAS1602/12 and CIO-DAS08) to measure the performance metric and set learning-rate parameters. In this way, the PC acts as a "watchdog" over the system and can learn to dynamically adapt the learning-rate parameters based on measures of the performance metric.

Results

We characterized performance of the VLSI systems using a simple adaptive laser focusing system shown in Figure 3. The beam from an Argon laser ($\lambda = 514$ nm) was expanded to a diameter of 30 mm and then passed through the HEX-127 SLM. The optical axis of the SLM was set at an angle $\pi/4$ with respect to the direction of the input beam polarization. A polarizer P_1 was placed after the SLM and in order to separate the phase-only modulated component of the input wave, the optical axis of P_1 was set parallel to the liquid crystal optical axis. This phase modulated wave passed through the lens L_1 and the beamsplitter BS_1. A small pinhole of 25 μm diameter was placed in the focal plane of the lens L_1 (with a focal length $F_1 = 14$ in.) A photodetector placed behind the pinhole measured the laser beam power through the pinhole. The VLSI system used the photodetector output voltage (filtered with a simple lowpass antialiasing filter prior to sampling) as the performance metric J. A camera CCD_1 registered the intensity distribution of the laser beam in the focal plane of the lens. The video image was displayed on a monitor M_1.

The results of adaptive system performance are presented in Figure 4. The adaptive system was exercised with a repeating sequence of performance metric maximization and minimization. During the first half of the cycle (first 512 steps), the system performed beam quality metric maximization ($\gamma' > 0$), followed by another 512 iterations of
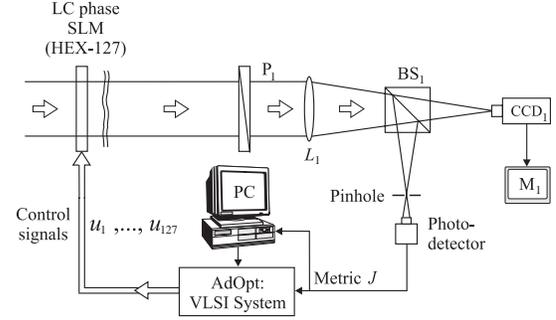
Figure 3: Schematic of the experimental setup for adaptive laser beam focusing using the mixed-mode VLSI system, designated "AdOpt".

beam quality metric minimization. During the minimization stage the adaptive system created random phase distortions resulting in the laser beam focal plane intensity spreading out.
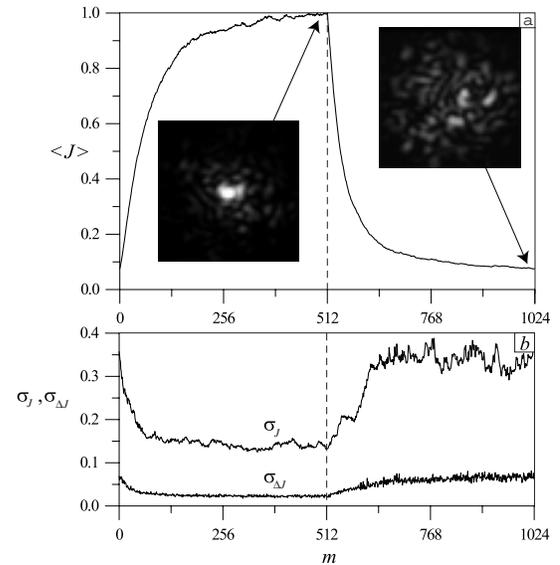
Figure 4: Averaged laser beam quality metric during the sequence of system performance metric J maximization ($m < 512$) and minimization ($m > 512$). The focal plane intensity distribution images from CCD_1 correspond to iteration step $m = 500$ (maximization) and to iteration step $m = 1000$ (minimization).

The dependencies $J(m), m = 1, \ldots, 1024$ (adaptation evolution curves) were averaged over 100 adaptation cycles. The normalized averaged evolution curve (metric values $\langle J(m) \rangle$) is shown in Figure 4 (*a*). The normalized standard deviation of the system performance metric $\sigma_J(m)$ and the standard deviation of the metric perturbation $\sigma_{\Delta J}(m)$ are shown in Figure 4 (*b*). The normalized standard devia-

tions were calculated using the following expressions:

$$\sigma_J(m) = \frac{\left\langle (J(m) - \langle J \rangle)^2 \right\rangle^{\frac{1}{2}}}{\langle J \rangle} \qquad (8)$$

and

$$\sigma_{\Delta J}(m) = \frac{\left\langle (\Delta J(m) - \langle \Delta J \rangle)^2 \right\rangle^{\frac{1}{2}}}{\langle J \rangle}. \qquad (9)$$

The evolution curves in Figure 4 (a) show the existence of two characteristic phases of the adaptation process: a relatively rapid convergence during first 100 to 150 iterations followed by a decrease in the convergence rate, a behavior also observed in numerical simulations [7]. The convergence occurs approximately 1.5 times faster for the metric minimization than for metric maximization (see Figure 4 (a)). This confirms a natural expectation that it is easier to create phase distortion than compensate for it. The adaptation behavior reflects the fact that the number of system states corresponding to a highly distorted beam (J at local minimum) is greater then the number of states corresponding to the (local) maximum; there are more ways to destroy quality of the beam then correct it. The presence of noise (laser beam intensity and photocurrent fluctuations) has more impact on the minimization phase of the adaptation process because the laser beam intensity after the SLM is low. Both factors—higher noise level and larger number of local minima of the performance metric—result in a significantly higher level of the performance metric fluctuations ($\sigma_J(m)$ in Figure 4 (b)) for the minimization than for the maximization process. The value of the standard deviation of the performance metric perturbation ($\sigma_{\Delta J}(m)$) was approximately 3% of the metric averaged value for the maximum $\langle J \rangle$ and about 7% for the minimum $\langle J \rangle$.

Conclusion

We have designed, built and tested a mixed-mode VLSI system for adaptive wavefront correction using stochastic parallel perturbative gradient ascent/descent. Although we have demonstrated its operation for an SLM with 127 elements, our design allows for expansion by addition of VLSI modules. The speed of operation of our system is presently limited by the dynamic response of the SLM (in the msec range). Our continued research in this area investigates alternative faster SLM technologies including high-speed and large-scale MEMS, integrated with CMOS circuitry to control parameters along with the adaptive processing elements.

This is the first demonstration of a VLSI system for adaptive phase wavefront correction. Since the implemented optimization is model-free and the objective measure can be arbitrarily specified, the results carry over to a large class of adaptive optics applications such as on-line corrective remote imaging through a turbulent atmosphere, laser communication through the atmosphere and adaptive focusing

for biomedical applications. Since our parallel architecture can be scaled to higher resolutions ($N \sim 10^3$ to 10^6 parameters) we are in a position to solve some of the most challenging problems in adaptive optics.

References

[1] R.K. Tyson, *Principles of Adaptive Optics*, Academic Press, Boston, 1991.

[2] M. A. Vorontsov and V. I. Shmalhauzen, *Principles of Adaptive Optics*, Nauka, Moscow, 1985.

[3] B.Y. Zeldovich, N.V. Pilipetsky, and V.V. Shkunov, *Principles of Phase Conjugation*, No. 42, Springer Series in Optical Sciences, Berlin Springer-Verlag, 1985.

[4] A. Dembo and T. Kailath, "Model-Free Distributed Learning," *IEEE Transactions on Neural Networks,* vol. **1** (1), pp 58-70, 1990.

[5] J.C. Spall, "A Stochastic Approximation Technique for Generating Maximum Likelihood Parameter Estimates," Proceedings of the 1987 American Control Conference, Minneapolis, MN, 1987.

[6] M.A. Vorontsov, G.W. Carhart and J.C. Ricklin, "Adaptive phase-distortion correction based on parallel gradient-descent optimization", Opt. Lett., 22, 907-909, 1997.

[7] M. A. Vorontsov and V. P. Sivokon, "Stochastic parallel gradient descent technique for high-resolution wavefront phase distortion correction," *J. Opt. Soc. Am.* A, **15**, pp. 2745–2758, 1998.

[8] G. Cauwenberghs, "A Fast Stochastic Error-Descent Algorithm for Supervised Learning and Optimization," in *Advances in Neural Information Processing Systems,* San Mateo, CA: Morgan Kaufman, vol. **5**, pp 244-251, 1993.

[9] D. Kirk, D. Kerns, K. Fleischer, and A. Barr, "Analog VLSI Implementation of Gradient Descent," in *Advances in Neural Information Processing Systems,* San Mateo, CA: Morgan Kaufman, vol. **5**, pp 789-796, 1993.

[10] G. Cauwenberghs, "An Analog VLSI Recurrent Neural Network Learning a Continuous-Time Trajectory," *IEEE Transactions on Neural Networks,* vol. **7** (2), March 1996.

[11] G. Cauwenberghs, "Analog VLSI Stochastic Perturbative Learning Architectures," *J. Analog Integrated Circuits and Signal Processing* **13** (1–2) pp. 195–209, 1997.

[12] R.T. Edwards, M. Cohen, G. Cauwenberghs, M.A. Vorontsov, and G.W. Carhart, "Analog VLSI Parallel Stochastic Optimization for Adaptive Optics," in G. Cauwenberghs and M. Bayoumi, Eds., *Learning on Silicon,* Norwell MA: Kluwer Academic, 1999.