Iterative procedure for in-situ optical testing with an incoherent source

Ryan Miyakawa[a], Patrick Naulleau[b], Avideh Zakhor[c], and Ken Goldberg[b]

[a]Applied Science & Technology Graduate Group, University of California, Berkeley
Berkeley, CA 94720, USA;
[b]Center for X-ray Optics, Lawrence Berkeley National Laboratory
1 Cyclotron Road, Berkeley, CA 94720, USA
[b]Electrical Engineering and Computer Sciences, University of California, Berkeley
Berkeley, CA 94720, USA

ABSTRACT

Interferometry, the long-standing method for optical characterization, is difficult to perform at EUV wavelengths due to the lack of high power coherent EUV sources and difficult experimental setup. These problems are exacerbated by systematic errors from geometrical effects as EUV tools move to higher numerical apertures (NA) which require stricter tolerances for the optical elements involved in interferometry. In this paper we propose an iterative, image-based, in-situ method for optical characterization that is independent of the operating wavelength of light, illumination coherence, and NA of the system. In this method, a known pattern is imaged through focus, and matched to an experimental model with a trial set of pupil aberrations. The aberrations are then changed iteratively until the modeled images match the ones from the experiment. The Reduced Optical Coherent Sum (ROCS) decomposition for partially coherent aerial image calculation greatly reduces the computation time of each iteration which makes this method more computationally tractable.

Keywords: optical testing, EUV, aberrations, interferometry, iterative

1. INTRODUCTION

Characterizing the optical aberrations of the imaging system in microscopes and lithography systems is an important tool for optimizing their performance to operate at their maximum resolution. As EUV optical systems move to larger numerical apertures to achieve higher resolution, it is crucial to have a simple and reliable procedure for characterizing the aberrations present in the optics. Standard interferometric techniques are more difficult to perform at higher numerical apertures. Reference wave interferometry such as Phase-Shifting Point Diffraction Interferometry (PS/PDI) requires smaller pinholes that are difficult to fabricate and provide low photon flux that gives poor contrast fringes.[1] Grating-based interferometry such as lateral shearing interferometry (LSI) is promising, but has strict tolerances on the position and tilt of the optical elements which couple with aberrations much more prominently at higher numerical apertures. Our iterative procedure has the benefit of being independent of numerical aperture, making it much more experimentally feasible. It also has the advantage that it can be made to work with existing tools with no additional experimental setup.

2. SETUP AND PROCEDURE

Figure 1 shows a schematic representation of a typical EUV optical system. Light radiates from an extended incoherent source placed in the rear focal plane of a collector optic that illuminates a test pattern, which is imaged by the test optic onto the image plane. Here, the detector can either be a CCD camera in an microscope setup, or a resist-coated wafer in a lithography setup. We obtain a through-focus series of images by translating the detector stage along the optical axis between each exposure. Alternatively, the object or the lens system could be translated to generate an equivalent defocus. In a lithography tool, images are obtained by developing the

Further author information: (Send correspondence to Ryan Miyakawa)
E-mail: rhmiyakawa@lbl.gov

Extreme Ultraviolet (EUV) Lithography, edited by Bruno M. La Fontaine, Proc. of SPIE Vol. 7636,
76361K · © 2010 SPIE · CCC code: 0277-786X/10/$18 · doi: 10.1117/12.846235

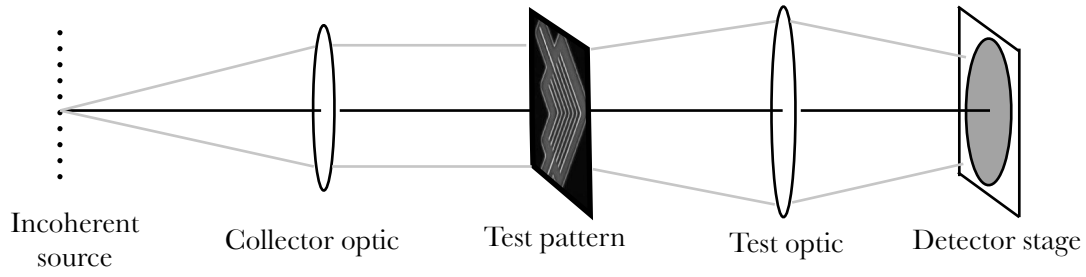

Figure 1. Schematic of a general partially coherent imaging system.

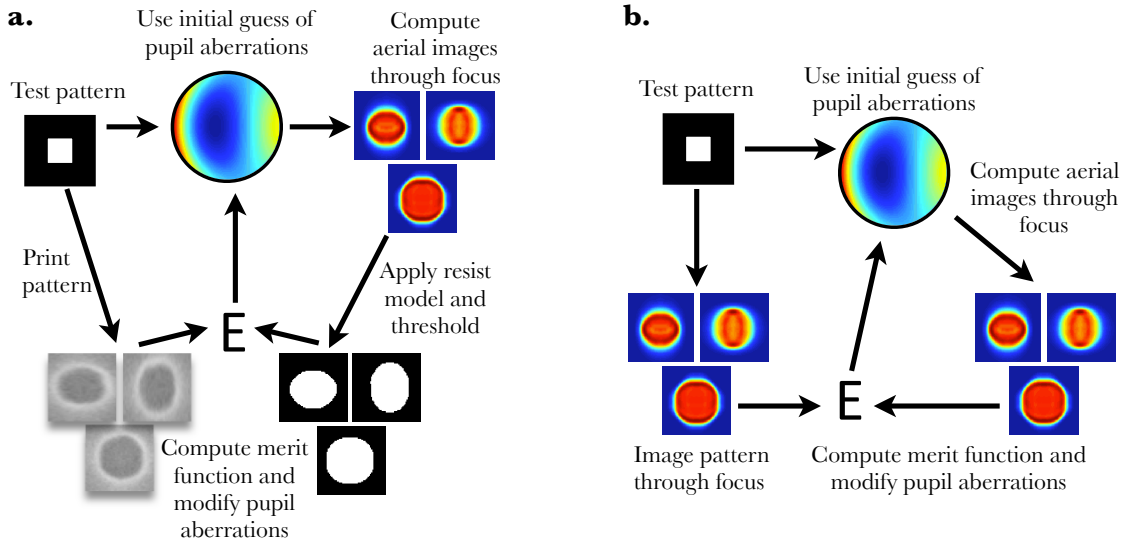

Figure 2. (a) Flow diagram of iterative reconstruction algorithm for lithography setup. (b) Flow diagram for imaging setup.

photoresist and viewing them by a scanning electron microscope (SEM), using the commercial software SuMMIT to process and convert the images into binary line-edge profiles.

Figure 2 shows a flow-diagram of the iterative reconstruction loop. A computer model of the optical system is generated using the known test pattern and source parameters. An aerial image through-focus series is generated via the Reduced Optical Coherent Sum (ROCS) decomposition detailed in the next section, using a initial guess vector of aberrations. In the imaging setup, a merit function is generated by comparing the image series to the experimental aerial images; in the lithography setup, the image series is convolved with the resist point-spread function, thresholded and compared with the experimental line-edges. The aberration vector and host function parameters are modified using a genetic simulated annealing algorithm whereby the best results from a set of independent trials is used as the initial guess of a subsequent generation of trials. A merit function is generated that serves to measure the proximity of the guess to the actual solution and the calculation is performed iteratively until the merit function reaches a desired tolerance chosen to accommodate the level of accuracy required by the test system.

3. FAST AERIAL IMAGE COMPUTATION VIA ROCS DECOMPOSITION

The robustness of the algorithm relies on the computation of many iterations, each of which involves the calculation of several aerial images. Aerial image modeling is normally computationally taxing because the Hopkins

equation which governs the partially coherent imaging system requires integration across four variables. In order to make the reconstruction algorithm more computationally feasible, we develop a new method for aerial image calculation that leverages the specific configuration of our experiment to optimize the calculation.

3.1. Motivation

In one dimension the Hopkins integral for aerial image calculation takes the following form:

$$\mathscr{I}(\nu) = \iint dp\, dz\ \mathscr{J}_0(p)\mathscr{K}(\nu)\mathscr{K}^*(\nu - z) \tag{1}$$

$$\times\ \mathscr{T}(\nu + p)\mathscr{T}^*(\nu - z + p) \tag{2}$$

Here $\mathscr{I}$ is the aerial image intensity Fourier transform, $\mathscr{K}$ is the coherent transmission function, $\mathscr{J}_0$ is the Fourier transform of mutual intensity of the light illuminating the mask and $\mathscr{T}$ is the mask Fourier transform. The Hopkins integral can be thought of as a system that takes three inputs, the mask, the source, and the pupil, and outputs the aerial image. In our experiment, the source and the mask are the same across all the calculations, so we can represent the Hopkins integral as a new system S that has the source and mask information integrated in the system. We can now think of the system S as mapping a pupil to an aerial image directly. Mathematically we can write this as

$$\mathscr{I}(\nu) = \int_{-\infty}^{\infty} dz\, \mathscr{K}(\nu)\mathscr{K}^*(\nu - z)S(\nu, \nu - z) \tag{3}$$

$$S(k, m) = \int_{-\infty}^{\infty} dp\ \mathscr{J}_0(p)\mathscr{T}(k + p)\mathscr{T}^*(m + p) \tag{4}$$

where S is the system cross-coefficient (SCC) matrix, that depends only on the source and mask. This formulation is in direct analogy with the more familiar transmission cross-coefficients (TCC) which depend on the source and the pupil.[3] Since S is a constant in each of the calculations, it can be computed once and stored, eliminating the need to evaluate it each time. We now show that by exploiting the mathematical structure of S, we can approximate it and significantly reduce the number of computations required to calculate the aerial image.

3.2. Spectral decomposition of the SCC matrix

By direct analogy to the TCC matrix in,[3] it can be shown that $S(k, m) = S^*(m, k)$ over the complex field so that S is Hermitian. Spectral theorem therefore guarantees us that S is diagonalizable and that we can write S as a sum over the outer products of its eigenvectors weighted by their corresponding eigenvalues,

$$S(k, m) = \sum_{i=1}^{N} \lambda_i \xi_i(k)\xi_i^*(m) \tag{5}$$

where $N = rank\{S\}$, which can be shown to be equal to the number of sampled source points. Plugging back into (3), and swapping the order of the sum and integral gives:

$$\mathscr{I}(\nu) = \sum_{i=1}^{N} \lambda_i \int_{-\infty}^{\infty} dz\, [\mathscr{K}(\nu)\xi_i(\nu)][\mathscr{K}^*(\nu - z)\xi_i^*(\nu - z)]$$

We recognize the integral as the autocorrelation of the product $\mathscr{K}\xi_i$ and take the inverse Fourier transform of both sides to get the space-domain aerial image intensity.

$$I(x) = \sum_{i=1}^{N} \lambda_i |\mathscr{F}^{-1}\{\mathscr{K}\xi_i\}|^2 \tag{6}$$

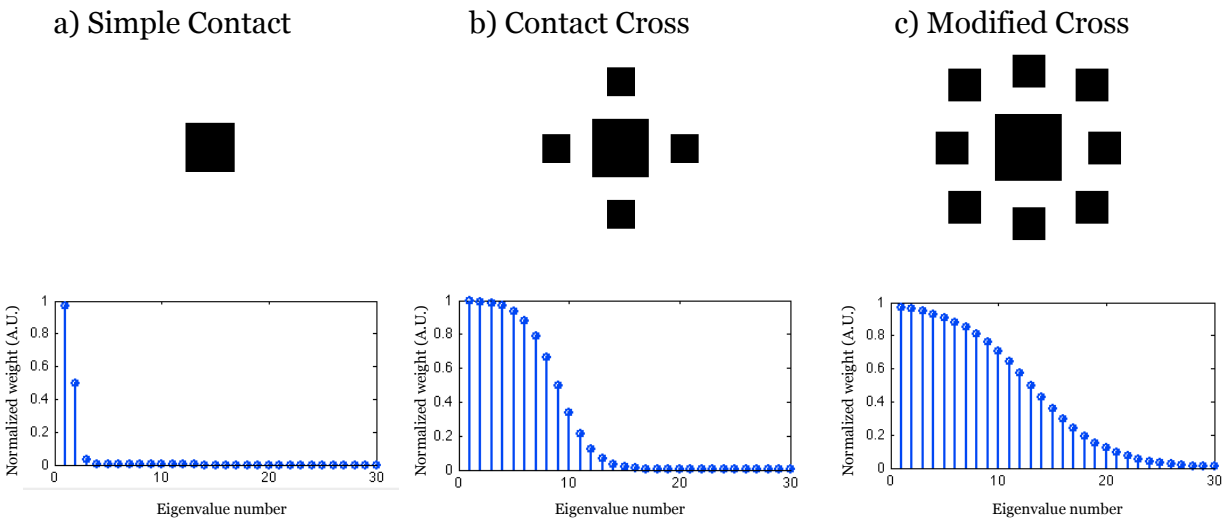

Figure 3. Different test patterns and the efficiency of the ROCS decomposition in each case.

3.3. Truncation of the spectral sum

Due to the energy compaction property of spectral decomposition,[4] the majority of the weight of S is represented in a relatively small number of terms. The aerial image intensity is therefore well-approximated by considering only the first $K^* < N$ terms of the sum. This truncation is known as a low-rank approximation.[5]

$$I(x) \approx \bar{I}(x) = \sum_{i=1}^{K^*} \lambda_i |\mathscr{F}^{-1}\{\mathscr{K}\xi\}|^2 \tag{7}$$

K^* is chosen to give a specified tolerance on the aerial image computation. The normalized error E between $I(x)$ and $\bar{I}(x)$ is bounded by the sum of the remaining eigenvalues:

$$E \leq \sum_{i=K^*+1}^{N} |\lambda_i| \tag{8}$$

We find that in a typical case, to achieve an error less than 1%, the value of $K^*/N \approx 0.1$ to 0.15, meaning that the aerial image computation is performed 7 to 10 times faster than with conventional methods.

3.4. Efficiency of decomposition

The net computational gains by the ROCS decomposition depend on the complexity of the source and the test pattern. Simple patterns such as a single contact under monopole illumination can achieve a ratio $K^*/N \approx 0.1$, where as a more complicated pattern that has features of different sizes with spatial frequencies that span a larger fraction of the pupil might have $K^*/N \approx 0.6$. Although they are slower to compute, more complex patterns will generally have a higher sensitivity to a larger number of aberrations. Therefore there is a natural tradeoff between the efficiency of the approximation and the precision of solution. Some common features are shown in Figure 3, and the table below lists the efficiency of the ROCS decomposition in each case.

4. SIMULATION AND RESULTS

We model a 0.35 NA imaging system operating at $\lambda = 13.5$ nm with a coherence factor $\sigma = 0.5$ in MATLAB. The test pattern is designed of a single pinhole of diameter $d = 50$ nm, equal to roughly twice the diffraction limit of the system, which is chosen to sweep through the full extent of the test optic pupil under the given illumination. Other test patterns were considered, including patterns with different size features, however for

Table 1. Comparison of ROCS decomposition on various pattern types

Feature Type	% Eigenvalues for 1 % error	% Eigenvalues for 2 % error	Computation Time in sec.	Aberration Sensitivity
Simple contact	12%	12%	.08	Moderate for primary aberrations
Contact cross	46%	41%	.28	Increased for comas
Modified contact cross	80%	70%	.53	Increased for secondary aberrations

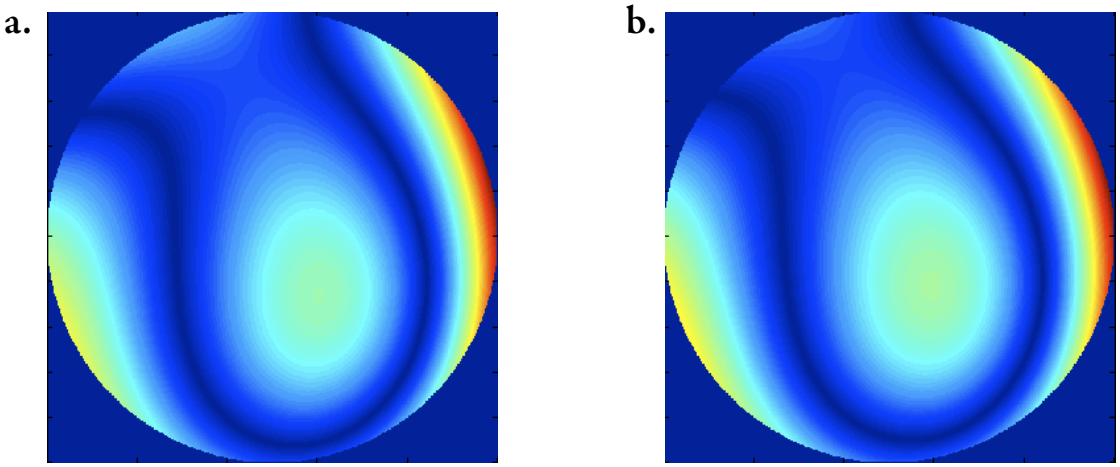

Figure 4. (a) Test wavefront. (b) Reconstructed wavefront via iterative algorithm. Total rms wavefront error $\mathcal{E}_{rms}$ = .04 waves

preliminary tests the simple pinhole pattern was chosen because it has a compact eigenvalue spectrum. An aberration vector of the first 8 Zernike polynomials is generated randomly using a Gaussian distribution with standard deviation $s_g = 0.25$ waves. Detector shot noise is modeled using Poisson photon statistics on a 16-bit camera. Aerial images at 3 focus steps are computed using ROCS and matched to the target image. A merit function is generated by integrating the absolute difference between the normalized image intensities.

Each subsequent aberration vector guess is generated via a genetic simulated annealing search algorithm. Simulated annealing is well suited for non-convex optimization because it contains a probabilistic component that helps resist the tendency to fall into local minima. More information regarding the details of the simulated annealing algorithm can be found in the literature.[7] The results in Figure 4 show the simulated test optic aberration map and the reconstructed aberration map to be in good agreement with total rms wavefront error $\mathcal{E}_{rms} = \lambda_{EUV}/25$. The simulation was performed on a Pentium-D 2.4 GHz dual-core processor and completed in 134 s.

To apply the method to real microscope data, we obtained through-focus data of a two-dimensional contact array of CD = 150 nm imaged at the SEMATECH Berkeley Actinic Inspection Tool (AIT) at Beamline 11.3.2 at the Advance Light Source operated at Lawrence Berkeley National Laboratory. The AIT operates at a wavelength $\lambda = 13.4$ nm, with coherence factors $[\sigma_x, \sigma_y] = [0.2, 0.1]$, and for our data uses a numerical aperture of 0.0875. Since the test pattern consists of an array of contacts rather than a single one, we are able to sample the aberrations across the field by applying the reconstruction to each contact independently. To prepare the data, we partition the field to isolate each contact and match it to its corresponding contact in the through-focus images. Preliminary results shown in Figure 6 compare the measured primary astigmatism with primary astigmatism measurements made independently by made independently by I. Mochi using a technique based on the observed ellipticity of the contacts through focus.

The two independent measurements are in good agreement with an error of $E = 0.09 \pm 0.08$ waves. We expect there to be slight discrepancies in the astigmatism measurement since the ROCS decomposition is fitting

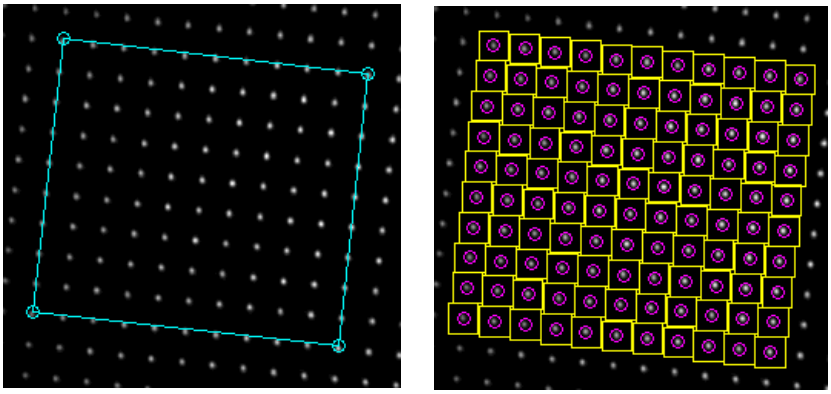

Figure 5. (a) Image of contacts with the region of interest highlighted. (b) Field partitioned to isolate contacts which are processed independently

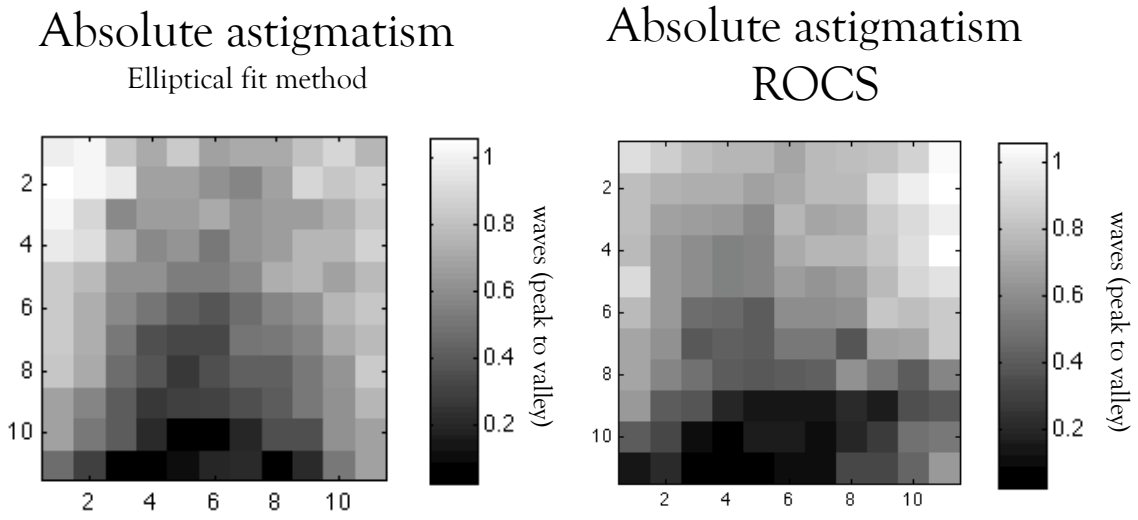

Figure 6. A comparison of the astigmatism measurements via an through-focus ellipticity fit and via ROCS

to 8 Zernike polynomials, whereas the ellipticity technique assumes that the contact broadening through focus is due to astigmatism alone. This assumption will accurately calculate the value of astigmatism when it is the only aberration present, but in the general case will couple with other aberrations. Additionally, neither method considers the effect of higher order astigmatism terms which will likely couple into the primary astigmatism measurements.

5. DISCUSSION

As with many iterative procedures, convergence of the algorithm can be susceptible to long computation times depending on the input parameters. Since non-convex searches like simulated annealing tend to converge much more slowly than convex algorithms, the computation of many generations may be required before a desirable tolerance is achieved. Additionally, since simulated annealing relies on random motion in the parameter space to step toward the solution, increasing the dimensionality of the space by including more Zernike polynomials can put further demands on the algorithm.

These drawbacks notwithstanding, we have found that feeding the algorithm a larger set of through-focus images or multiple illumination settings makes the contour of the merit function more convex, which helps the

algorithm converge more quickly. Another key advantage of simulated annealing is that it can be partitioned into independent tasks that can be run in parallel on several processors.[8]

In summary, our preliminary simulations verify the viability an iterative image-based approach to optical testing. Further work must be done in extending the simulation to a lithography setup to assess the feasibility of using the method to test lithography tools. As we continue to move toward higher resolution, iterative image-based optical testing is a promising alternative to interferometry and may play an important role in next-generation optical systems.

ACKNOWLEDGMENTS

The authors are grateful for support from the NSF EUV Engineering Research Center. Lawrence Berkeley National Laboratory is operated under the auspices of the Director, Office of Science, Office of Basic Energy Science, of the US Department of Energy.

REFERENCES

1. P. P. Naulleau, K. A. Goldberg, et al., Applied Optics 38 (35), 7252-63 (1999), "Extreme-ultraviolet phase-shifting point-diffraction interferometer: a wave-front metrology tool with subangstrom reference-wave accuracy"
2. C. Ahn and H. Kim and K. Baik, SPIE proc., Vol. 3334, pg. 752-763, 1998, "A novel approximate model for resist process
3. N. B. Cobb, Fast optical and process proximity correction algorithms for integrated circuit manufacturing, Ph.D. dissertation (Electrical Engineering and Computer Science, University of California, Berkeley, 1998).
4. Kenji Yamazoe, J. Opt. Soc. Am. A Vol. 25, No. 12, December 2008, "Computation theory of partially coherent imaging by stacked pupil shift matrix"
5. Jieping Ye, Machine Learning, Vol 61, Numbers 1-3, November 2005, "Generalized Low Rank Approximations of Matrices"
6. M. P. Rimmer, Applied Optics Vol. 13, No. 3, March 1974, "Method for Evaluating Lateral Shearing Interferograms"
7. S. Kirpatrick, Journal of Statistical Physics, Volume 34, Numbers 5-6 / March, 1984, "Optimization by simulated annealing: Quantitative studies"
8. Tarek M. Nabhan, Albert Y. Zomaya, "A Parallel Simulated Annealing Algorithm with Low Communication Overhead," IEEE Transactions on Parallel and Distributed Systems, vol. 6, no. 12, pp. 1226?1233, December, 1995.