Systematic Design of Phase-Shifting Masks with Extended Depth of Focus and/or Shifted Focus Plane

Yong Liu, Anton Pfau, and Avideh Zakhor, *Member, IEEE*

***Abstract*—We propose an optimization based algorithm for designing phase-shifting masks. Our approach is an extension of our previous work [1]–[3] in the sense that the intensity image is optimized at a number of optical planes rather than just the focus plane. In addition, our algorithm can be used to design masks with shifted focus plane and/or extended depth of focus. We also propose the concept of "dual mask" and show its consequences to practical phase-shifting mask design. Finally we show examples of our proposed design techniques for single line phase connectors, cross phase connectors, contact holes and bright lines. Simulation and experimental results verify the capability of our design technique to extend depth of focus and shift the focus plane.**

I. Introduction

PHASE-SHIFTING masks have shown a great deal of promise in extending the applicability of optical lithography to next generation of VLSI manufacturing. Many types of phase-shifting masks, such as Levenson type, self-aligned ring type, out-rigger type [4], [5] have been proposed. While these conceptual models exploit the intuition gained from optics, they are ad hoc and non-systematic. A systematic approach to phase-shifting mask design based on Simulated Annealing (SA) algorithm was proposed by Liu and Zakhor [2]. This approach is general in that, it can be used for arbitrary feature patterns, can be easily modified to include defocus and aberration effects, and can be adapted to any phase-shifting mask, such as out-rigger type and chromeless type. In this paper, we extend the algorithm in [2] to extend Depth of Focus (DOF) and show examples of phase-shifting mask design obtained via our proposed technique.

The basic idea is to optimize a cost function which depends on the quality of the light intensity at a number of optical planes rather than just the focus plane. Contrary to the conventional belief that higher contrast implies larger DOF, our optimization algorithm can trade-off contrast in focus plane with contrast and line width control at defocus planes. This, in turn, results in more uniform intensity profiles across various defocus planes.

The outline of this paper is as follows. In Section II, we formulate the phase-shifting mask design as an optimization problem. Specifically, in Section 2.2 we propose a new cost function for optimization across various planes. While our previous cost function in [2] is a piecewise linear function of error in image intensity, our new cost function is nonlinear. In Section 2.3, we present a new cooling schedule for the simulated annealing algorithm, in which we convert annealing into simple iterative algorithm when the cost function is near the basin of a minimum. We also propose a new rule for terminating Markov chains associated with the SA by considering that the states we could actually visit in a practical annealing process is only a small fraction of the total number of states.

In Section III, we introduce the concept of "dual mask." Specifically, we show that by straightforward modification of a given phase-shifting mask, we can flip the intensity pattern with respect to the optical focal plane. We will show consequences of the above result to practical phase-shifting mask design in Section 4.4.

In Section IV, we show simulation results for design examples such as phase connectors, contact holes, and bright lines using our proposed algorithm. The mask patterns for these examples are optimized at three optical planes in order to increase DOF. We also show that the focus plane can be shifted by appropriate choice of the location of the optical planes which are being optimized. Section V includes experimental results verifying performance improvements of our optimization based design techniques over conventional design techniques. Section VI includes conclusions.

II. Formulation with Defocus Effects

In this section, we formulate the optimization process by taking the defocus effects into consideration. Our approach is to optimize an objective function which depends upon the light intensity at various spatial points in various optical planes. Specifically, we begin by computing the local error functions and local cost functions at each spatial point on each optical plane. We then combine the local costs at each spatial point on different optical planes

Manuscript received April 13, 1992; revised July 27, 1992. This work has been supported in part by the SRC/SEMATECH contract Nos. 91-MC-500 and 92-DC-008, and in part by the National Science Foundation PYI grant MIP-9057466, IBM, Kodak, TRW and Joint Services Electronics Project. Y. Liu is partially supported by the K. C. Wong Education Foundation.

The authors are with the Department of Electrical Engineering and Computer Sciences, University of California at Berkeley, Berkeley, CA 94720.

IEEE Log Number 9205039.

0894-6507/93$03.00 © 1993 IEEE

to obtain the combined local cost at that spatial point. Finally, we sum up the combined local cost functions at different spatial points to obtain the total cost function. Minimization of the total cost function results in the optimal mask.

2.1 General Formulation

For simplicity, we use similar notation to [2]. Suppose the desired binary aerial image, shown in Fig. 1, is zero in domain C and one in domain D. Let A and B represent contours surrounding the boundaries of regions C and D respectively, where contour A is in region C and contour B is in region D. Furthermore, we assume the design specification on the light intensity $I(\vec{x})$ for all the optical planes to be identical and of the form:

$$I(\vec{x}) \le a \qquad \vec{x} \in A \tag{1}$$

$$I(\vec{x}) \ge b \qquad \vec{x} \in B \tag{2}$$

$$I(\vec{x}) \le c \qquad \vec{x} \in C \tag{3}$$

$$I(\vec{x}) \ge d \qquad \vec{x} \in D \tag{4}$$

where $\vec{x}$ represents spatial location (x, y), and a, b, c and d represent the corresponding intensity threshold values in regions A, B, C and D respectively. In practical situations, a, b, c and d are chosen in accordance with the parameters of the resist, shown in Fig. 2, where we have chosen $c = d$. We define an error function that measures the amount of violation of inequalities (1) through (4) at each spatial point $\vec{x}_k$ on each optical plane μ:

$$e_k^\mu(I^\mu(\vec{x}_k, \vec{q})) = \begin{cases} I^\mu(\vec{x}_k, \vec{q}) - a & \text{if } \vec{x}_k \in A \\ b - I^\mu(\vec{x}_k, \vec{q}) & \text{if } \vec{x}_k \in B \\ I^\mu(\vec{x}_k, \vec{q}) - c & \text{if } \vec{x}_k \in C \\ d - I^\mu(\vec{x}_k, \vec{q}) & \text{if } \vec{x}_k \in D \end{cases} \tag{5}$$

The superscript $\mu = 1, \cdots, P$ denotes the μth optical plane, k is the index for sampling points, $\vec{q}$ is the vector of unknown pixel variables which represent the mask pattern to be optimized.

Let the local cost function at spatial point k in region i on optical plane μ be denoted by $h_i^\mu(e_k^\mu)$. We combine the local cost functions at different optical planes in a pointwise fashion to obtain a combined local cost function Z at spatial point k,

$$Z(h_i^1(e_k^1), h_i^2(e_k^2), \cdots, h_i^P(e_k^P)) \tag{6}$$

where $i = A, B, C$ or D represents the region that the sampling point belongs to.

Finally, we define the total cost function to be the sum of combined local cost functions at each spatial point:

$$H(\vec{q}) = \sum_i \sum_{k \in i} Z(h_i^1(e_k^1(I^1(\vec{x}_k, \vec{q}))), h_i^2(e_k^2(I^2(\vec{x}_k, \vec{q}))), \cdots, h_i^P(e_k^P(I^P(\vec{x}_k, \vec{q})))) \tag{7}$$

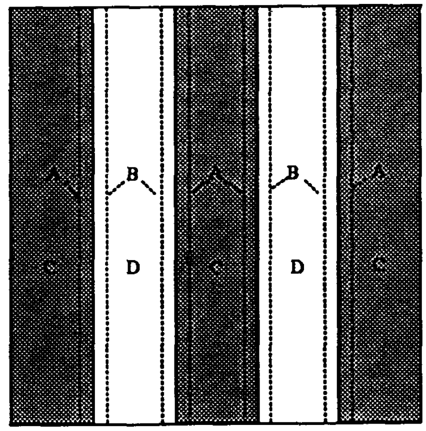


Fig. 1. Desired binary image of two infinitely long bars. The desired image is 1 in region D and 0 in region C. The regions between contours A and B are transition regions in the specifications.

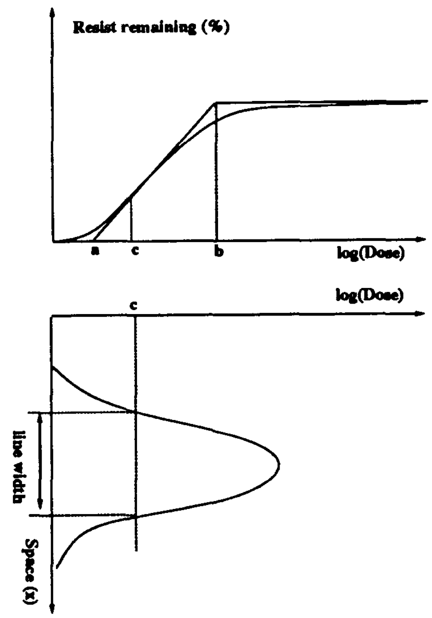


Fig. 2. A simple model of photoresist. As a rule of thumb, the line width corresponds to the 0.3 contour of light intensity when the light intensity is normalized by that of a large feature.

where $k \in i$ means that the sampling point k is in region i, with $i = A, B, C$ or D.

2.2 Design of Cost Function

As seen in (6) and (7), the total cost function H consists of a summation over the function Z, which combines local cost functions $h_i^\mu(e_k^\mu)$. It can be argued that rather than designing Z and h_i separately, one could directly design a total cost function W as a function of errors e_k^μ to combine errors at different sampling points and on different optical planes simultaneously. However, we have chosen to separate the total cost function into a local cost function and a combination function for different optical planes in order to trade-off intensity patterns at various optical planes easily. For example, by giving more weight to the error in the focal plane, we can ensure a more desirable pattern at the focal plane at the expense of less desirable ones in defocus planes. In the remainder of this section, we will describe our choice of local function $h_i^\mu(e_k^\mu)$ and the combination function Z.

2.2.1 The Local Cost Function

The local cost function is used to measure the degree of violation at different spatial points on the same optical plane. Ideally a local cost function should (a) penalize violations of constraints in inequalities 1 through 4, (b) detect intensity dips at phase transition areas, (c) allow the optimization algorithm to correct the dips, (d) result in a uniform violation of the constraints. To satisfy these properties, we devise the following guidelines for the construction of local cost function:

1. The user should not have to deal with constraints on a point by point basis. Rather, the relative importance of contours A and B, and regions C and D should be controllable.
2. There should be a large penalty for constraint violations but none or small penalty if the constraints are satisfied.
3. Local large violations such as dips in phase connectors should be less desirable than the sum of many small violations. As a consequence, given a fixed total amount of violations, a uniform violation among different sampling points is more desirable than nonuniform violations.
4. Linewidth control is more important than contrast enhancement if enough contrast is achieved, but less important than contrast if there is not enough contrast.
5. Since the developed resist edges are usually close to 0.3 intensity contours, there is less margin for violations in dark field region than in bright field region. For the pattern in Fig. 1, this implies less margin for violations in domain C and on contour A than in domain D and on contour B.

For convenience of discussion, we consider only one optical plane, $P = 1$, with two sampling points, and with a total cost function of two variables $H = H(e_1, e_2)$, where e_1 and e_2 are error functions at two different spatial points. We assume that functions H and h are differentiable and we note from (7) that $\partial H/\partial e_k = \partial h/\partial e_k$, for $k = 1, 2$. The desired properties of H and hence h are:

$$\frac{\partial h}{\partial e_k} \geq K(e_k), \qquad \text{for } e_k \leq 0 \qquad k = 1, 2; \tag{8}$$

$$\frac{\partial h}{\partial e_k} \ll K(0^+), \qquad \text{for } e_k \leq 0 \qquad k = 1, 2; \tag{9}$$

$$\left.\frac{\partial h(u, v)}{\partial v}\right|_{u = \text{constant}} \geq 0 \qquad \text{for } v > 0 \tag{10}$$

$$\left.\frac{\partial h(u, v)}{\partial v}\right|_{u = \text{constant}} \geq 0 \qquad \text{for } v < 0 \tag{11}$$

where $u - v$ are the Cartesian coordinates obtained by rotating the $e_1 - e_2$ coordinates 45° counter clockwise.

We now describe the way guidelines 1-5 are related to desired properties in inequalities (8)–(11). Guideline 2 requires h to be monotonically increasing, or equivalently $K(e_k)$ to be positive for all e_k. The "$\ll$" sign in inequality (9) also reflect guideline 2. Guideline 3 requires $K(e_k)$ to be an increasing function of e_k in regions C and D. Equations (10) and (11) enforce h to be at minimum when $e_1 = e_2$ and to increase with $|e_1 - e_2|$ for a fixed $e_1 + e_2$. Thus (10) and (11) are in compliance with guideline 3. Since contrast is related to error in regions C and D, and linewidth control is related to error on contours A and B, guideline 4 implies that for small e_k, we must have $h_A > h_C$ and for large e_k, we must have $h_C > h_A$.

One way to achieve guidelines 3 and 4 would be to choose $K(e_k)$ to be a positive constant for A and B and to be an increasing exponential for C and D. In fact, we have chosen $K(e_k)$ in this way to design a set of local cost functions as follows:

$$h_i(e) = e \qquad \forall e \leq 0; \qquad i = A, B, C, D \tag{12}$$

$$h_A(e) = Qem_c/t_{c2} \qquad \forall e \geq 0 \tag{13}$$

$$h_B(e) = Qem_d/t_{d2} \qquad \forall e > 0 \tag{14}$$

$$h_C(e) = a_c(e^{b_c e} - 1) \qquad \forall e > 0 \tag{15}$$

$$h_D(e) = a_d(e^{b_d e} - 1) \qquad \forall e > 0 \tag{16}$$

where $Q = aN_A + cN_C + (2 - b)N_B + (2 - d)N_D$ is a bound for negative contributions to the total cost function, N_A, N_B, N_C and N_D denote the number of points in regions A, B, C and D, respectively, the t parameters are directly related to the resist characteristics, m_c and m_d are parameters that control the growth rate of the cost functions, and a_c, b_c, a_d and b_d are functions of t and m parameters as will be explained shortly.

We impose the relationship shown in Fig. 3 between h_A, h_C, t_{c1}, t_{c2}, Q and m_c. As seen, the error for region C, $h_C(e)$, and the error on contour A in region C, namely $h_A(e)$, intersect at $e = t_{c2}$. For small values of error, namely $e < t_{c2}$, $h_A(e)$ is larger than $h_C(e)$ emphasizing the importance of linewidth control over contrast. For large values of error, namely $e > t_{C2}$, $h_C(e)$ is larger than $h_A(e)$, emphasizing contrast over linewidth control. Thus, the shapes of h_C and h_A shown in Fig. 3 are in compliance with the requirements of guideline 4. The relation between h_B and h_D are similar to that of h_A and h_C.

Given the growth rate m_c and the parameters t_{c1} and t_{c2} in Fig. 3, we can determine the parameters a_c and b_c in

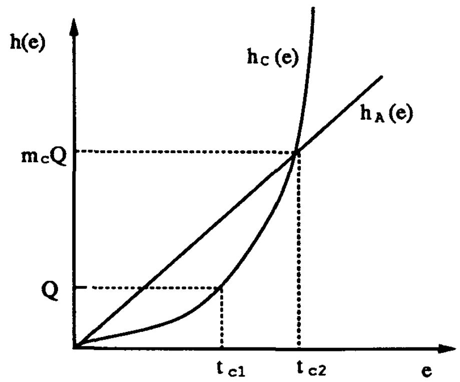


Fig. 3. The local cost function of contour A and region C. m_c is a parameter that controls the growth rate of the cost function. For contour B and region D, we simply replace the subscripts a by b and c by d.

the following way,

$$b_c = \frac{1}{t_{c2}} \ln \left[m_c(e^{b_c t_{c1}} - 1) + 1\right] \tag{17}$$

$$a_c = Q/(e^{b_c t_{c1}} - 1). \tag{18}$$

Equation (17) can be solved by an iterative scheme using the expression itself with initial guess $b_c = \epsilon$, where $0 < \epsilon << 1$. Under this condition, convergence to the root can be guaranteed. a_d and b_d in (16) are determined in a similar fashion to a_c and b_c.

The set of cost functions defined in (13)–(16) can be further modified according to different applications in order to allow trade-off between linewidth control and contrast, by multiplying different constants λ_i with different local cost functions $h_i(x)$. The parameters we use in most of our simulations in Section IV are $a = 0.25$, $b = 0.35$, $c = 0.15$ and $d = 0.75$, $m_c = m_d = 2.2$, $\lambda_a = \lambda_b = 2.2$ and $\lambda_c = \lambda_d = 1$, $t_{d1} = 0.1\,(d - c)$, $t_{d2} = 0.2\,(d - c)$ $t_{c1} = 0.07\,(d - c)$ and $t_{c2} = 0.14\,(d - c)$. The relative values of these t parameters reflect the requirement in guideline 5.

2.2.2 *Choice of Combination Function*

The choice of combination function is similar to that of local cost function. Since the local cost function can possess most of the desirable properties required for the total cost function H, we only need to be concerned about the uniformity property in choosing the combination function.

For convenience of discussion, we consider a combination function of two variables $Z = Z(h^1, h^2)$, where h^1 and h^2 are local cost functions at the same spatial point on different optical planes. It is natural to require Z to be symmetric with respect to variables h^1 and h^2. Clearly it is also desirable that Z be monotonically increasing with respect to h^1 and h^2. A necessary requirement for uniformity is

$$\left.\frac{\partial h^\mu(u, v)}{\partial v}\right|_{u = \text{constant}} \geq 0 \qquad \text{for } v > 0, \qquad \mu = 1, 2 \tag{19}$$

$$\left.\frac{\partial h^\mu(u, v)}{\partial v}\right|_{u = \text{constant}} \geq 0 \qquad \text{for } v < 0, \qquad \mu = 1, 2 \tag{20}$$

where $u - v$ are the Cartesian coordinates obtained by rotating the $h^1 - h^2$ coordinates 45° counter clockwise, as shown in Fig. 4. Equations (19) and (20) imply that a uniform image profile along different optical planes is favored and non-uniformity results in an increase in the cost function.

As shown in Fig. 4, for a given constant u corresponding to the point A on the u axis, we draw the line AB which is perpendicular to the h^1 axis, and the line AD which is perpendicular to the u axis. Since Z is a monotonically increasing function of h^2, it monotonically increases from point B to A. Furthermore, since the uniformity constraint of (19) requires Z to increase monotonically from A to D, the isocontour curve AC of combination function Z that passes through point A must lie within the triangle area ABD. In fact, the isocontour curve AC characterizes the combination function. When AC is close to AD, the combination function Z is close to $h^1 + h^2$, and uniformity is not emphasized. When AC is close to curve AB, the combination function nearly corresponds to the larger of h^1 and h^2, and uniformity of image along optical axis is emphasized. In fact, the AB curve corresponds to $Z = \max\{h^1, h^2\}$ and results in a min-max optimization problem, which is more difficult to solve than the resulting optimization problem due to AD curve.

For implementation purposes, we propose to use a combination function whose isocontours are straight lines between AB and AD. To be specific, we first sort $h_i^\mu(e_k^\mu)$ into descending order:

$$h_i^1(e_k^1) \geq \cdots, \geq h^{P_i}(e_k^P) \qquad \mu = 1, \cdots, P. \tag{21}$$

we then combine them as follows:

$$Z(e_k^1, e_k^2, \cdots, e_k^P) = \sum_{\mu=1}^{P} \kappa_\mu h_i^\mu(e_k^\mu) \tag{22}$$

where $\kappa_1 \geq \kappa_2 \geq \cdots, \geq \kappa_P$. If $\kappa_\mu = 0$ for all μ except for $\kappa_1 = 1$, then the above combination function corresponds to a min-max problem. If $P = 2$, the combination function in (22) is characterized by the line AE with slope $-\kappa_1/\kappa_2$.

The design examples shown in this paper are all obtained by using the AB curve. The min-max criterion in effect leads to more uniform intensity image distribution along the optical axis at the expense of contrast at focal plane, as shown in Fig. 5. This should change the traditional wisdom that higher contrast at focus plane implies larger depth of focus.

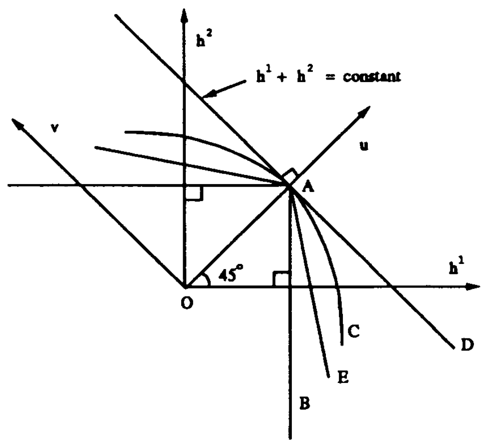


Fig. 4. Partition of the domain of the combination function Z, the function is symmetrical with respect to u-axis. The isocontour line of Z passing through A should fall in the triangular area ABD.

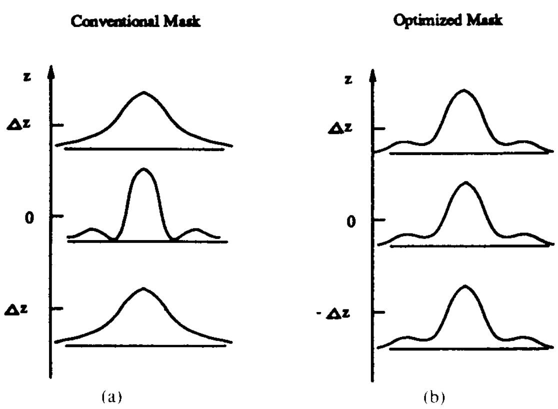


Fig. 5. Intensity profile variation at different optical planes. (a) Due to conventional binary or phase-shifting mask. (b) The ideal distribution enforced by our algorithm.

2.3 Fast Cooling Schedule

Similar to our approach in [2], we use the simulated annealing algorithm to solve our optimization problem. In this section, we will describe the particular cooling schedule we use in our optimization algorithm.

In SA algorithm, we view possible mask patterns as states of a thermal system, and the total cost function of mask patterns as the energy function associated with the states [2]. Mask patterns are randomly generated and accepted with a probability defined by Boltzmann distribution:

$$P\{\vec{\boldsymbol{q}} = \vec{q}\} = \frac{1}{U(T)} \exp\left(-\frac{H(\vec{q})}{T}\right), \qquad (23)$$

where $\vec{q}$ is the vector of pixel variables of a particular mask pattern to be optimized, $\vec{\boldsymbol{q}}$ represents the corresponding random variable, $U(T)$ is a normalized constant, or the partition function in statistical physics and T is the temperature that controls the distribution. We start with a high temperature and gradually cool the system down. At each temperature we allow the system to reach thermal equilibrium by updating the state of the system using the above acceptance/rejection process. The sequence of states experienced by the system is called a Markov chain [2]. The probability of accepting low energy states becomes larger as temperature is reduced. The system eventually settles down to a near global minimum energy state. The choice of initial and final temperatures, the rule for termination of Markov chain, and the rule for the reduction of temperature is referred to as a cooling schedule.

Optimization at different optical planes is computation intensive. Basic SA algorithms such as the one used in [2], [3] are primitive and have obvious drawbacks. To increase the speed, we propose two measures. Our first measure has to do with the fact that at very low temperature, the state bounces back and forth for a long time before it finally settles down to a local minimum. To avoid this problem, we compare the maximal change of total cost function in a single move at each temperature to the difference between the maximum and minimum cost of states at that temperature. When they are comparable, say less than 0.1% in difference, we set the temperature to zero and start a simple iterative algorithm. The transmittance of each of the mask pixels are then changed one by one; when all the pixels have been visited in a row without causing reduction in the cost, the optimization process is terminated.

Our second measure has to do with the fact that when using a simple cooling schedule, at high temperatures the state bounces back and forth, and therefore it is hard to detect the existence of equilibrium. Intuitively, there is no equilibrium unless the temperature reduction rate is very slow compared to the rate of visiting different states in the solution space. Since the simulation of imaging system in each iteration is time consuming, we can not afford to visit a large number of states, and any practically affordable annealing process cools "too fast." In view of this, we suggest the decision to reduce the temperature should be based on the position of the state on energy surface rather than the time spent on previous temperature. Hence we propose a new rule for termination of Markov chains at each temperature.

Let $\boldsymbol{H}$ denote the random variable for total cost function at certain temperature with expectation ν and the standard deviation σ. Clearly $\boldsymbol{H}$ follows Boltzmann distribution, and ν and σ are functions of temperature. Our proposed new rule terminates a Markov chain as follows: whenever the cost of the current state falls into within $\pm 0.5\sigma$ of the expectation ν of the cost $\boldsymbol{H}$ at current temperature, we conclude that a "good" position is reached and the Markov chain at that temperature is terminated. To avoid fluctuations of estimates of σ, σ_{l-1} for temperature T_{l-1} is used as an estimate of σ_l for temperature T_l. To estimate the expectation ν_l at temperature T_l, a running average is computed with a weight function that emphasizes recent observations more than earlier ones in order to reduce transient effects due to temperature change. To be specific, we define the running average as follows,

$$\overline{H}_n = \frac{\sum_{i=1}^{n} p_i H_i}{\sum_{i=1}^{n} p_i} \tag{24}$$

where n is the number of moves made at temperature T_l, H_n is the total cost after n moves under temperature T_l and the weighting coefficients are chosen to be $p_i = 1 - \rho^i$, with ρ being a parameter to be determined. By simple manipulation, we obtain a convenient form

$$\overline{H}_n = \frac{M_n}{D_n} = \frac{M_{n-1} + H_n(1 - \rho^n)}{D_{n-1} + (1 - \rho^n)} \tag{25}$$

where M_n and D_n are the numerator and denominator of (24), respectively. We denote the number of pixel variables by N and estimate the transient time to be $N_1 = \sqrt{\text{neighborhood size}} = \sqrt{\# \text{ of levels} \times N}$. The parameter ρ can be determined by setting the damping factor to be $1/\sqrt{2}$ at the end of transient time N_1, as shown in Fig. 6:

$$(1 - \rho^{N_1}) = \frac{1}{\sqrt{2}} \Rightarrow \rho = 0.29289^{1/N_1}. \tag{26}$$

To avoid termination of Markov chains prematurely due to lack of samples for meaningful statistics, we set up a buffer time before the test for "good" positions can be started. Specifically, we wait for N_1 cycles for the transients to die out and for N_2 cycles to obtain meaningful statistics. We estimate N_2 using Chebyshev's inequality. Let $\boldsymbol{H}_i$ be the random variable representing the total cost function of the ith accepted state at certain temperature. Strictly speaking, $\boldsymbol{H}_i$, $i = 1, 2, \cdots$ are not independent identically distributed (iid) random variables, but as a good approximation, we assume they are iid with expectation ν and standard deviation σ. Let $\overline{\boldsymbol{H}}_n$ denote the random variable representing a simple average of $\boldsymbol{H}_i$, $i = 1, 2, \cdots, n$, and $\overline{\mathbf{H}}_n$ denote a sample of $\overline{\boldsymbol{H}}_n$. Then the expectation and standard deviation of $\overline{\boldsymbol{H}}_n$ are ν and $\sigma/\sqrt{n}$, respectively. By Chebyshev's inequality

$$\epsilon = P\{|\overline{\boldsymbol{H}}_n - \nu| > \delta\} \leq \frac{\sigma^2}{n\delta^2} \tag{27}$$

Limiting the error in estimation of ν to half the standard deviation of $\boldsymbol{H}_i$, i.e. $\delta = \sigma/2$, then $\epsilon \leq 4/n$. If we want $\epsilon \leq 0.1$, we can choose $n \geq 40$ or $N_2 \geq 40$. We note that the running weighted average is different from a simple average, but they are close after $N_1 + N_2$ cycles.

In summary, we do not start testing whether a "good" position is reached until $n \geq N_1 + N_2$. After that time, if at any step $|H_n - \overline{H}_n| \leq 0.5\sigma_{l-1}$, then the Markov chain is terminated for temperature T_l and the temperature is reduced. On average, the effects of the two measures described in this section result in more than doubling the speed of the optimization process without deteriorating the quality of solutions.

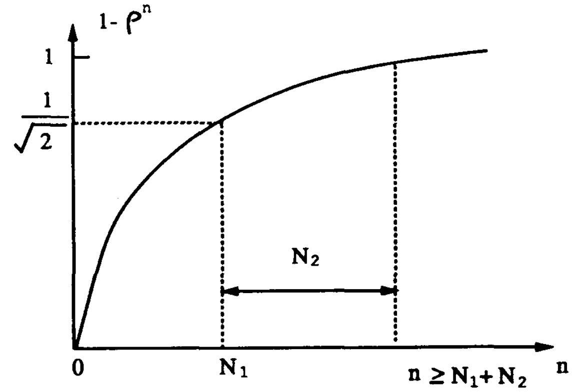


Fig. 6. The weighting function for the running average of total cost function, where N_1 represent the transient period after a new Markov chain is started, N_2 represent an estimate of the samples needed for a statistically meaningful running average.

III. Dual Mask System

In this section, we derive a relationship between mask phase-symmetry and the resulting intensity symmetry with respect to the optical focus plane. As we will see in Section 4.4, this property is of importance in designing focus shifted masks.

Let $g(x, y)$ and $f(x, y)$ denote the complex transmission coefficients of two phase-shifting masks. Let the superscript "*" denote complex conjugation. We introduce the concept of dual mask as follows:

Definition 1: Phase-shifting masks $g(x, y)$ and $f(x, y)$ are said to be dual to each other if there exists a real number θ such that $g(x, y) = e^{j\theta} f^*(x, y)$ for all x and y.

Since typical phase-shifting masks in practice have finite number of levels, the dual of an arbitrary phase-shifting mask may not have the same phase levels as the mask itself. For instance, if a phase-shifting mask has phase levels 0°, 60° and 180°, choosing $\theta = 180°$ in definition 1, then its dual has phase levels 0°, 120° and 180°. This means that the dual mask requires one more mask step to manufacture the 120° phase than the steps needed to manufacture the original mask. However, as we will see, if the phase levels of a mask belong to a "dual mask system," then the phase levels of its dual also belong to the same system. We now define a dual mask system as follows:

Definition 2: A phase-shifting mask system with m level phases $\{\phi_1, \phi_2, \cdots, \phi_m\}$, and identical amplitude of transmission for all the levels, is called a dual mask system if there exists at least one phase θ such that for any phase level ϕ_i, $i = 1, \cdots, m$, the phase $\phi_i^d = \theta - \phi_i$ also belongs to the phase levels $\{\phi_1, \phi_2, \cdots, \phi_m\}$. ϕ_i^d and ϕ_i are called the dual phase of each other.

Typical phase-shifting mask systems under investigation by researchers are all dual mask systems. For example, $\theta = 180°$ for the mask systems {0°, 180°}, {0°, 90°, 180°} and {0°, 60°, 120°, 180°}, $\theta = 0°$ for the mask system {−90°, 0°, 90°}, and $\theta = 0°$ or 120° or 240° for the mask system {0°, 120°, 240°}. It is interesting to note that for the $\{0°, 120°, 240°\}$ mask system, a dual mask can be obtained by keeping any one phase fixed and exchanging the other two phases for all spatial points of the original mask. Having defined the concept of dual mask, we now state the following symmetry property about it:

Theorem 1: If masks $g(x, y)$ and $f(x, y)$ are dual to each other, then, independent of the value of θ, the intensity image of one mask at an optical plane with Δz defocus is the same as the intensity image of the other mask at $-\Delta z$ defocus.

The proof consists of two steps, first we derive the relationship between Transmission Cross Coefficients (TCC) function at Δz defocus and $-\Delta z$ defocus, and the symmetry property of TCC function with respect to the origin; then we derive the sufficient condition under which the above statement is true. For details of proof, refer to the Appendix.

As an example of dual masks, the masks in Figs. 19 and 29 are dual to each other. In both figures, the dark areas represents opaque, and all other areas are transparent with different phases: white for 0°, light shade for 90° and dark shade for 180°. Their optical properties are symmetric with respect to the focus plane, as shown in Figs. 20 and 30.

IV. Simulation Results

Based on our proposed algorithm, we can design phase connectors with desirable contour shapes and contrast both at focal plane and defocus planes. More interestingly, we are able to design masks that shift focal plane of intensity image away from the optical focus plane. This capability allows us to design masks that adapt to the uneven topology of resist and hence results in an effective increase in depth of focus.

4.1 Phase-Connector Design

Among the proposed phase-shifting masks, Levenson type has been shown to result in significant performance improvements especially for patterns such as equal lines and spaces [4]. However it is used for alternating features and therefore when applied to more complex features, where opposite phase assignments may be required on different portions of a feature, phase-conflicts occur. Hence, a phase-connector, which can connect opposite phases, is needed in order to apply Levenson type mask to more general phase-shifting masks. An existing solution is to use segments of intermediate phases to connect the opposite phases; however, this results in light intensity dips at the junction of different phases. As minimum feature size reduces, this problem becomes more serious. In this section, we show the application of our algorithm to overcome this problem.

For all examples in this section, the resist is assumed to be positive, the wavelength is 0.365 μm, the numerical aperture is 0.32 and the degree of partial coherence is 0.50. All the masks are periodic with only one period

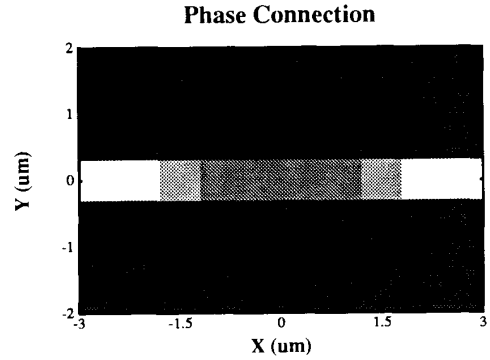


Fig. 7. Mask pattern for a bright line 0° − 90° − 180° phase connector designed by simple intuition. In the figure, dark area is opaque; all other areas are transparent with different phases: white for 0°, light shade for 90° and dark shade for 180°.

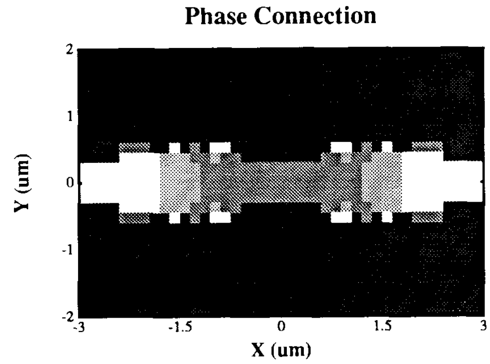


Fig. 8. Mask pattern for a bright line 0° − 90° − 180° phase connector designed by our proposed algorithm. In the figure, dark area is opaque; all other areas are transparent with different phases: white for 0°, light shade for 90° and dark shade for 180°.

shown in the figures. We use a phase-shifting mask system composed of opaque and transparent with 0°, 90° and 180° phases. In the figures, opaque is represented by dark areas, 0° phase by white area, 90° phase by light shade area and 180° phase by dark shade area.

Figs. 7 and 8 show designed single space phase connectors, using simple intuition and our proposed approach respectively. The feature size is 0.60 μm and the pixel size in Figs. 8 is 0.15 μm. Figs. 9 and 10 show the show intensity at 0.0, $\pm$ 1.0 μm defocus. As seen in Fig. 7, the simplest intuitive way of connecting a 0° phase region and a 180° phase region is to use a 90° segment in between. However, as seen in Fig. 9, this results in a light intensity dip at the transition area which becomes more serious as the transition length becomes shorter. The resulting mark using our proposed algorithm is shown in Fig. 8. Due to four-fold symmetry, we only optimize the mask in the first quadrant with the contributions from other quadrants taken into consideration. By taking advantage of the chrome area surrounding the bright line phase connector, our algorithm results in less intensity dip at both focus and defocus planes, without deteriorating the line width variation. The optimal design, though more diffi-

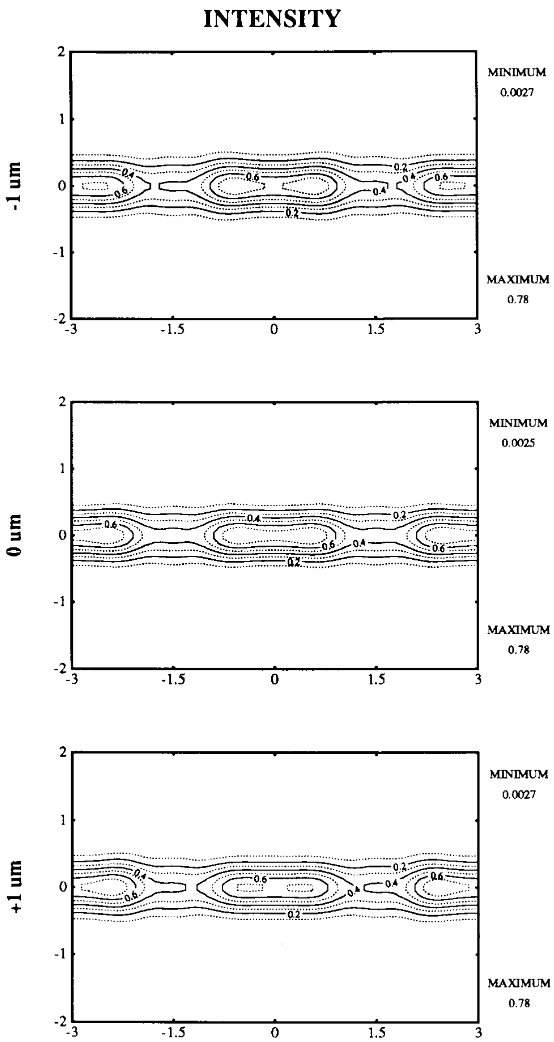


Fig. 9. Simulated intensity distribution of the intuitive line phase connector mask in Fig. 7 for different optical planes at 0.0, ± 1.0 μm defocus.

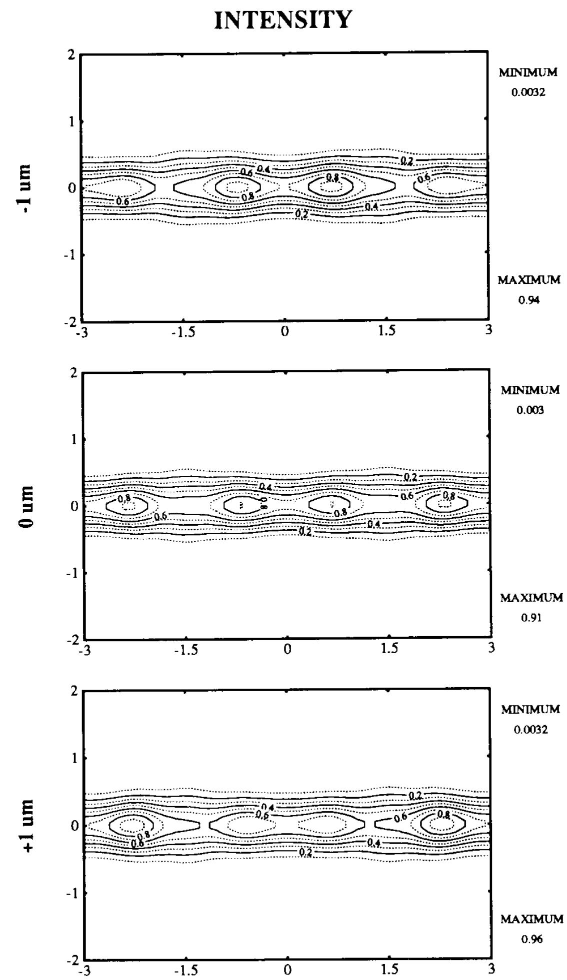


Fig. 10. Simulated intensity distribution of the optimized line phase connector mask in Fig. 8, for different optical planes at 0.0, ± 1.0 μm defocus.

cult to manufacture, leads to intuitive insight. For instance, it suggests that at a phase junction, we should make the transparent line wider and use neighboring opposite phases to cancel the light leaking into the chrome area.

Figs. 11 and 12 show a Levenson type mask for parallel bright lines which are to be connected to a vertical bright line. Specifically, Fig. 11 shows a simple intuitive design, while Fig. 12 shows our optimized design. The feature size is 0.64 μm and the pixel size in Fig. 12 is 0.16 μm. Phase-conflict occurs because the parallel lines with opposite phases have to join the same vertical line. The intuitive design is a clever arrangement for the phase connection, where all the phase transitions occur on vertical to horizontal line junction, so as to keep the advantages of Levenson type mask. As seen in Fig. 13, this mask results in a fair connection at the focus plane, but has problems at ± 1.0 μm defocus planes. We notice that the

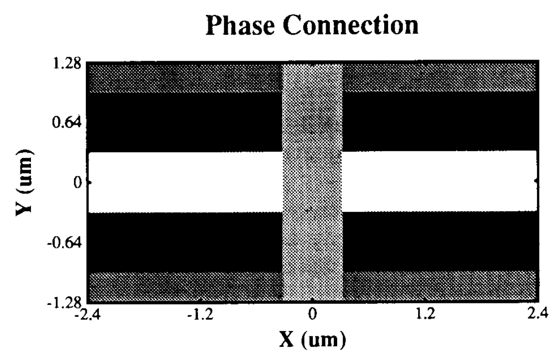


Fig. 11. Mask pattern for a cross connector designed by intuition. In the figure, dark area is opaque; all other areas are transparent with different phases: white for 0°, light shade for 90° and dark shade for 180°.

intensity images are different at positive and negative defocus planes. By modifying few chrome pixels at the junction, our designed mask in Fig. 12 results in improvement at both focus and defocus planes as seen in Fig. 14.

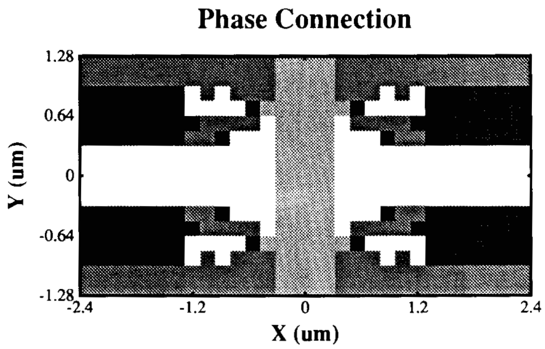


Fig. 12. Mask pattern for a cross connector designed by our proposed algorithm. In the figure, dark area is opaque; all other areas are transparent with different phases: white for 0°, light shade for 90° and dark shade for 180°.

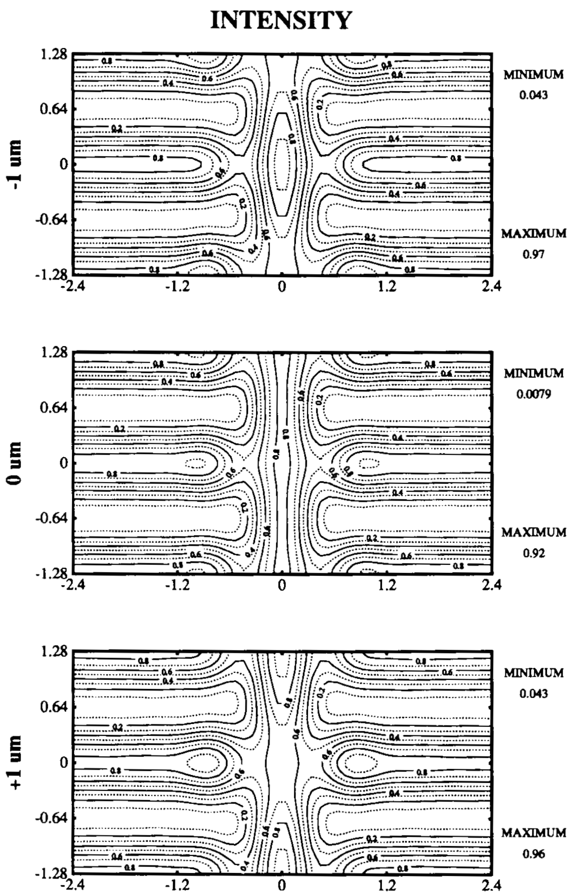


Fig. 13. Simulated intensity distribution of the intuitive cross connector mask in Fig. 11 for different optical planes at 0.0, ± 1.0 μm defocus.

Quantitative results on the improvement are hard to obtain because at these junctions there is no clear definition of linewidth, especially since there are distortions at defocus due to the phase transition. Nonetheless, our simulation results indicate that for medium features, the designed masks can compensate diffraction effects to result in better shaped intensity contours, and for small features, the designed masks can provide proper biases. Our contact hole examples in Section 4.2 show more clearly the way our proposed algorithm can increase depth of focus.

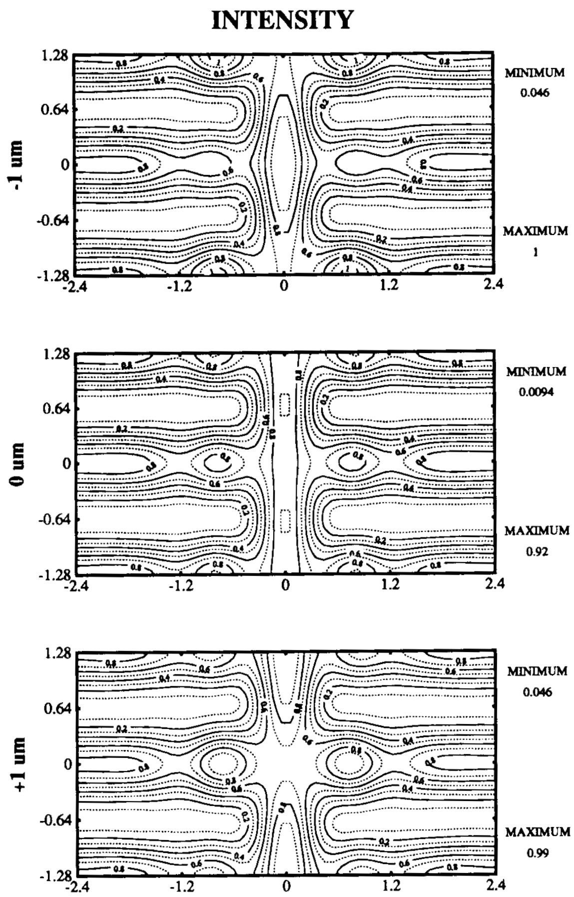


Fig. 14. Simulated intensity distributions of the optimized cross connector mask in Fig. 12, for different optical planes at 0.0, ± 1.0 μm defocus.

4.2 Contact Hole Design: Long DOF and Focus Shift

In this section, we present various contact hole designs with long depth of focus, with or without a shift in focus planes. In all the examples, we use a {0°, 90°, 180°} phase system. The dark region represents opaque, white region represents 0° transparent, light shade region represents 90° transparent, and dark shade region represents 180° transparent. The designed masks are compared to conventional binary masks, whose sizes are the same as the desired contact hole sizes. As a good approximation, the contact hole size is defined to be the diameter of normalized intensity contour at 0.3 level. Since the designed contact hole masks are frequently asymmetric with respect to $x = y$ line, we plot the contact hole sizes in both x and y directions.

4.2.1 Long Depth of Focus

In Figs. 15–18, i-line $\lambda = 0.365$ μm is used, the numerical aperture is 0.45, the degree of partial coherence

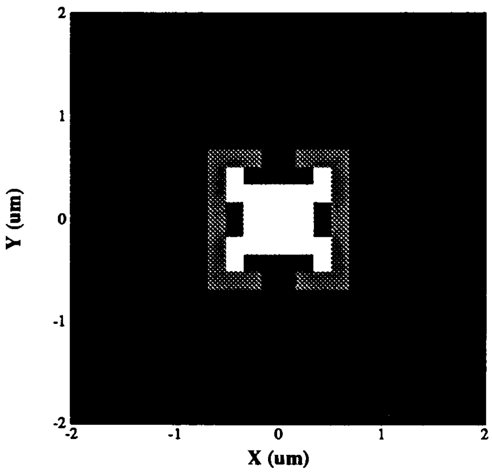


Fig. 15. An optimized mask for 0.68 μm contact hole, with no focus plane shift. In the mask, dark area is opaque; all other areas are transparent with different phases: white for 0°, light shade for 90° and dark shade for 180°.

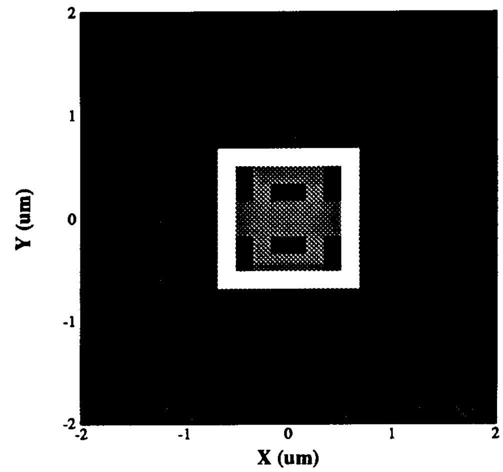


Fig. 17. Another optimized mask for 0.68 μm contact hole, with no focus plane shift. In the mask, dark area is opaque; all other areas are transparent with different phases: white for 0°, light shade for 90° and dark shade for 180°.

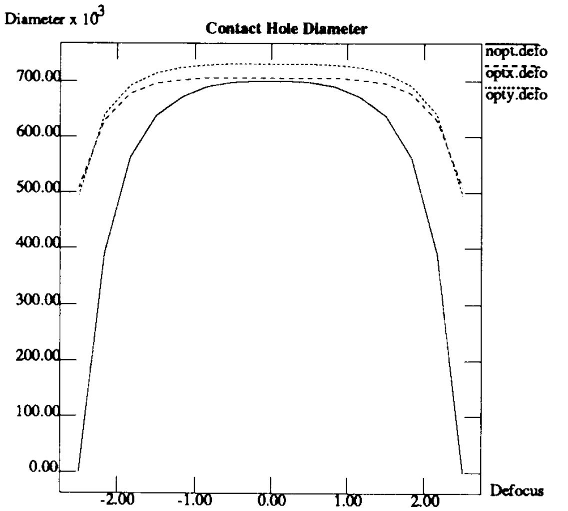


Fig. 16. Contact hole sizes for a conventional 0.68 μm binary mask is shown with the solid line marked with nopt.defo, and for the optimized mask in Fig. 15 is shown with the dashed lines marked with optx.defo for diameter in x-direction and opty.defo for diameter in y-direction. The contact hole sizes are measured using the diameter of 0.3 contour of normalized intensity images.

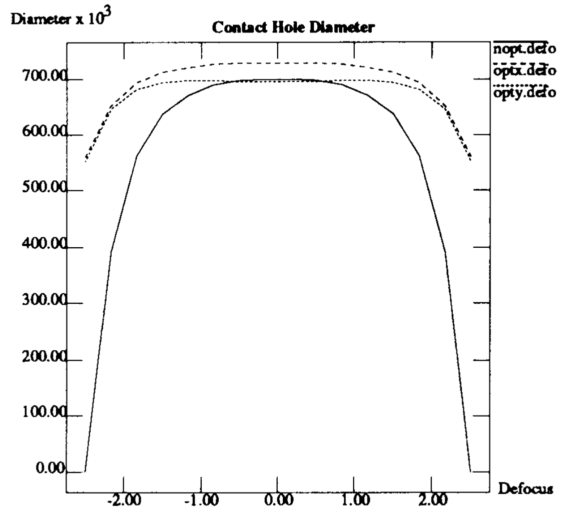


Fig. 18. Contact hole sizes for a conventional 0.68 μm binary mask is shown with the solid line marked with nopt.defo, and for the optimized mask in Fig. 17 is shown with the dashed lines marked with optx.defo for diameter in x-direction and opty.defo for diameter in y-direction. The contact hole sizes are measured using the diameter of 0.3 contour of normalized intensity images.

is 0.5, and the pixel size is $0.17 \times 0.17\ \mu m^2$. The desired image and the conventional binary mask is a 0.68 μm contact hole. The masks designed by our algorithm are optimized at optical planes $z = -2.0, 0, 2.0\ \mu$m. The vector $\vec{\kappa}$ is (1, 0, 0), that is, we choose to solve a min-max optimization problem.

Fig. 15 shows a designed mask pattern using our algorithm. Its corresponding contact hole size versus defocus is compared to a conventional binary mask in Fig. 16. Our approach clearly results in a significant increase in depth of focus. The reason for arriving at an asymmetric solution can be partially explained by considering a symmetric extension of the designed mask in Fig. 15. Specifically, we have found that the hole size of the symmetric mask deviates more from the desired size than the optimally designed asymmetric mask of Fig. 15.

The mask in Fig. 15 is not the only one with acceptable performance. By starting with a different random initial pattern for the simulated annealing algorithm, we can find another mask, shown in Fig. 17, that offers comparable performance, as shown in Fig. 18. Other comparable solutions are also possible. This phenomenon is typical of large "frustrated" systems, where conflicting requirements are to be satisfied. In such systems, with multiple

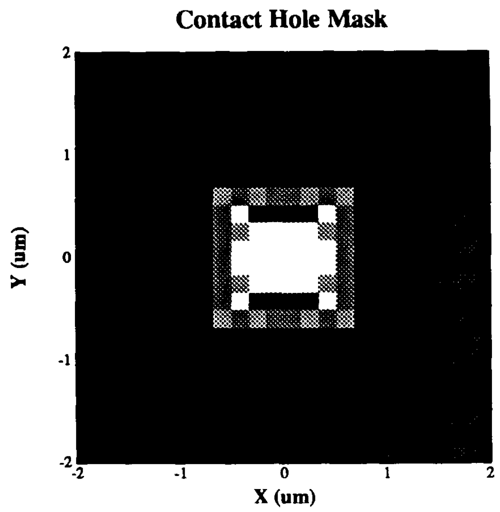


Fig. 19. An optimized mask for 0.68 μm contact hole, with a focus plane shift of 1.0 μm. In the mask, dark area is opaque; all other areas are transparent with different phases: white for 0°, light shade for 90° and dark shade for 180°.

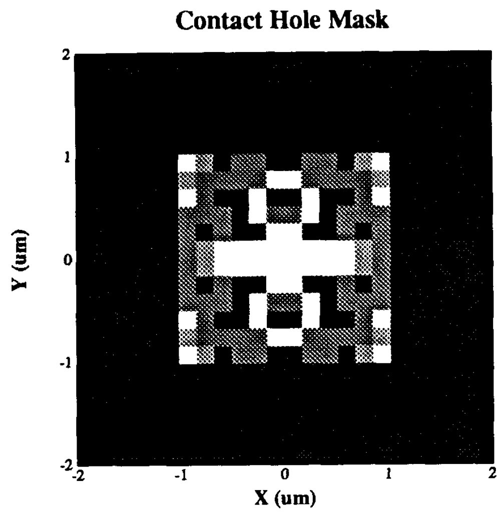


Fig. 20. Contact hole sizes for a conventional 0.68 μm binary mask is shown with the solid line marked with nopt.defo, and for the optimized mask in Fig. 19 is shown with the dashed lines marked with optx.defo for diameter in x-direction and opty.defo for diameter in y-direction. The contact hole sizes are measured using the diameter of 0.3 contour of normalized intensity images.

near global minima, it is not worthwhile to find the exact global minimum. This multiplicity of near global minima is referred to as *ultrametricity* [6].

Since the intensity specifications in the above examples are symmetric with respect to the focus plane, our algorithm chooses not to use the 90° phase. This means that 90° can not help in both directions of defocus, and the best masks for symmetric defocus behavior are composed of 0° and 180° phase shifters only. Another insight that can be gained from the designed masks in Figs. 15 and 17 is that, if we limit the mask area to a small region, then the optimal masks boost intensity by making the hole larger, and use a ring shifter to suppress the intensity at the edges.

4.2.2 Shifting the Focal Plane

Our algorithm can be used to design masks with shifted focal plane provided such shifts are not limited by the physics of the situation. In the following examples, the wavelength is 0.365 μm, the numerical aperture is 0.32, the degree of partial coherence is 0.5 and the pixel size is 0.17 μm $\times$ 0.17 μm. The desired contact hole size is 0.68 μm.

The designed mask in Fig. 19 is optimized for $z = -0.5$, 1.0, 2.5 μm optical planes with $\vec{\kappa} = (1, 0, 0)$ resulting in a desired shift of 1.0 μm in the focus plane. As seen in Fig. 20, the designed mask clearly accomplishes this shift. Besides the focus shift, the designed mask also provides proper bias and extended depth of focus. The conventional binary mask results in a contact hole of 0.64 μm instead of 0.68 μm.

Fig. 21 shows a mask with a larger area of optimization, optimized at $z = -1.0$, 0.5, 3.0 μm optical. As seen in Fig. 22, the designed mask accomplishes three tasks all at once: focus shift of 0.5 μm, extended DOF and

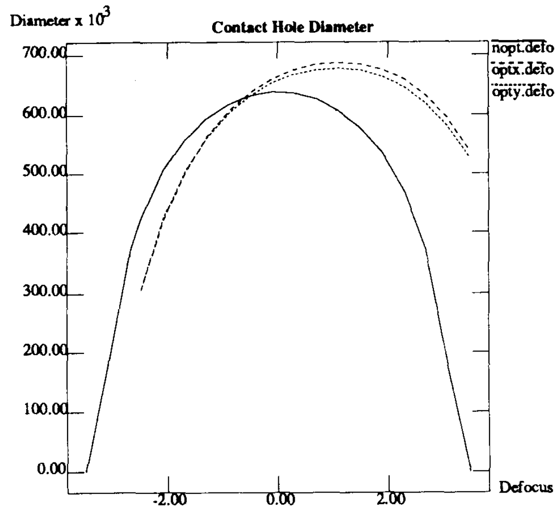


Fig. 21. An optimized mask for 0.68 μm contact hole, with a focus plane shift of 0.5 μm. The area of optimization is three times the size of a conventional mask. In the mask, dark area is opaque; all other areas are transparent with different phases: white for 0°, light shade for 90° and dark shade for 180°.

proper bias. The increase of DOF is about 100%, which is quite unusual. This can be explained by considering that a larger mask area results in more degrees of freedom and hence superior performance.

4.3 Bright Line Design: Long DOF and Focus Shift

In this section, we show two examples of bright line designs using our algorithm, one for extension of DOF and the other for focus shift. The numerical aperture is 0.48 and the degree of partial coherence is 0.38. Due to

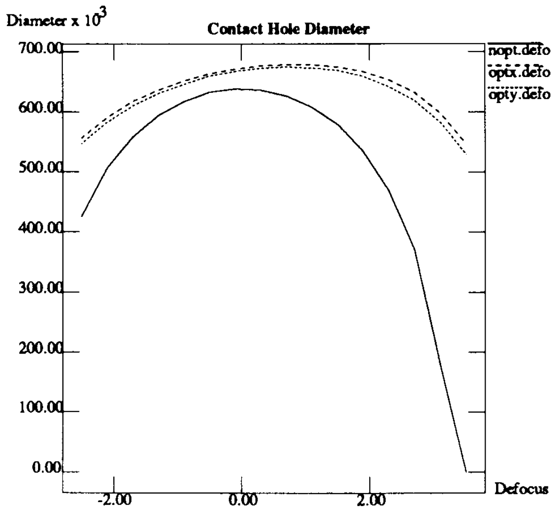


Fig. 22. Contact hole sizes for a conventional 0.68 μm binary mask is shown with the solid line marked with nopt.defo, and for the optimized mask in Fig. 21 is shown with the dashed lines marked with optx.defo for diameter in x-direction and opty.defo for diameter in y-direction. The contact hole sizes are measured using the diameter of 0.3 contour of normalized intensity images.

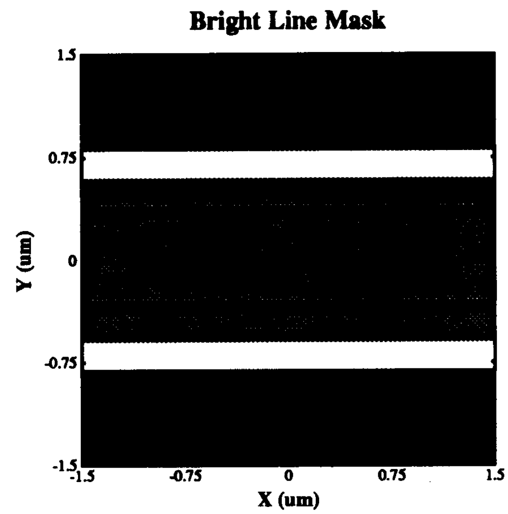


Fig. 23. An optimized mask for 0.8 μm space line, with long DOF. The area of optimization is twice the size of a conventional mask. In the mask, dark area is opaque; all other areas are transparent with different phases: white for 0°, and shade for 180°.

four-fold symmetry, we only optimize the first quadrant while taking the contributions of other quadrant into account. The pixel size is $1.5 \times 0.1\ \mu m^2$.

Fig. 23 shows a designed mask for a 0.8 μm bright line with extended DOF. The designed mask is optimized at optical planes $z = -2.5, 0.0, 2.5$. The vector $\vec{\kappa}$ is chosen to be (1.0, 0.5, 0.3). Since there is no focus shift, only two phases {0°, 180°} are necessary. The linewidth variation of the designed mask at different defocus is compared to a binary conventional mask in Fig. 24. As seen, the increase of DOF is tremendous, and the linewidth uniformity is excellent. Such performance improvement is possible, because the features size to be designed is above optical resolution, so there is more margin in contrast at focus plane to be traded-off with contrast at defocus planes. This will become more clear in the following discussion.

Figs. 25 and 26 show the intensity distributions along a line in y direction of the conventional mask and the designed mask in Fig. 23, respectively. The intensity of the designed mask is not as bright as the conventional binary mask at focus plane, but it maintains enough brightness along much wider range of defocus than the conventional mask. Physically, a conventional mask together with high numerical aperture lens results in a significant amount of energy carried by waves of large incident angles. These waves tend to destroy the coherency when they go out of focus. The assisting features in our designed mask re-focus some waves at defocus planes, thus inevitably reducing the coherency at the focus plane. It is interesting to note in Fig. 26 that outside the defocus region of interest, the intensity of the designed mask drops off rapidly and turns into two bright spots. One disadvantage of the designed mask is the existence of small ripples in the dark region; however these ripples are within specification in the range $\{-2.5, 2.5\}$ μm of defocus. Besides, these ripples are inherent to phase-shifting masks. The other disadvantage of the designed mask is the use of larger mask area than the conventional one.

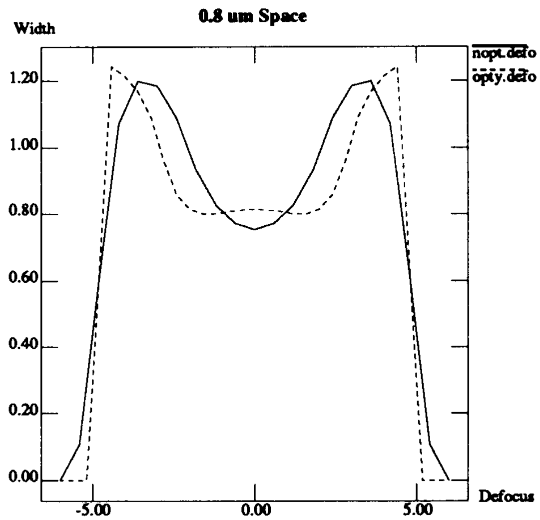


Fig. 24. Bright line linewidths of a conventional 0.8 μm binary mask is shown with the solid line marked with nopt.defo, and for the optimized mask in Fig. 23 is marked with opty.defo. The linewidths are measured using the 0.3 contour of normalized intensity images.

Fig. 27 shows a designed mask for a 0.4 μm bright line with focus plane shifted by about 0.76 μm. The corresponding linewidth as a function of defocus is shown in Fig. 28. The designed mask is optimized at optical planes $z = 0.0, 1.0, 2.0$ μm, which suggests a 1 μm shift in focus plane. We are not able to achieve the desired 1 μm

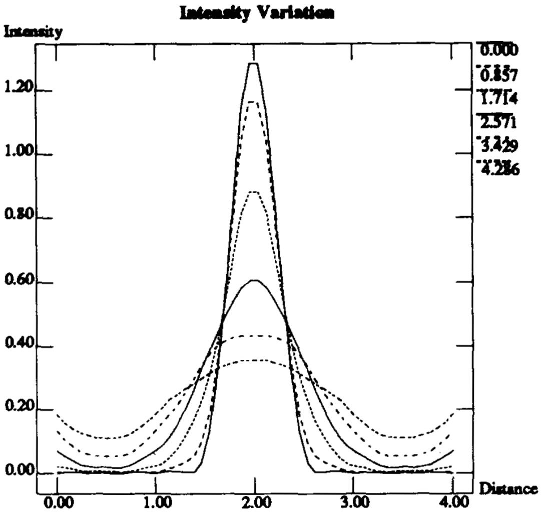


Fig. 25. Intensity distribution of a conventional mask with a width of 0.8 μm. The "distance" is measured along a line in y direction of the mask. The curves are intensity distribution along the line measured at different defocuses.

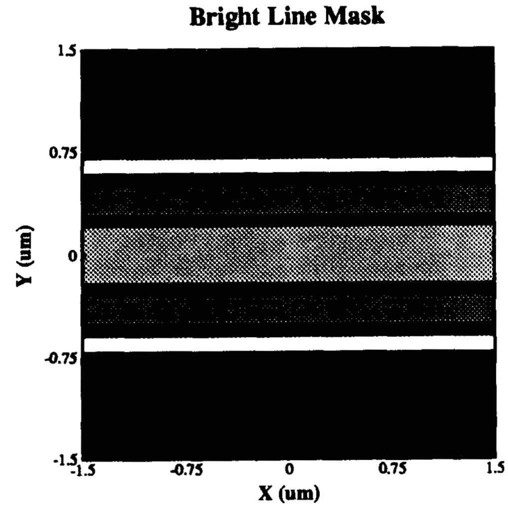


Fig. 27. An optimized mask for 0.4 μm space line, with a focus plane shift of 0.7 μm. The area of optimization is 3.5 times the size of a conventional mask. In the mask, dark area is opaque; all other areas are transparent with different phases: white for 0°, light shade for 60° and dark shade for 180°.

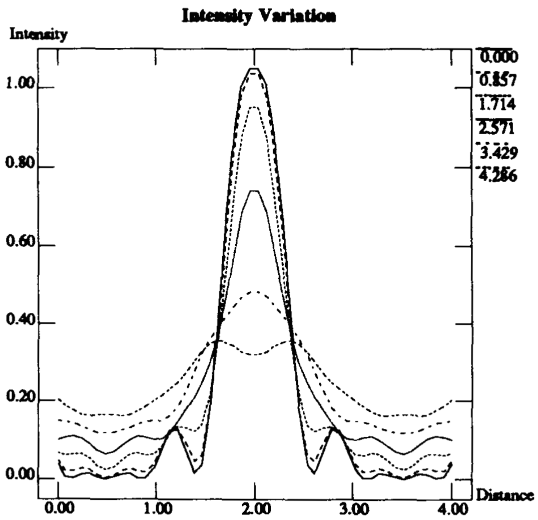


Fig. 26. Intensity distribution of the optimized mask for a 0.8 μm space, shown in Fig. 23. The "distance" is measured along a line in y direction of the mask. The curves are intensity distribution along the line measured at different defocuses.

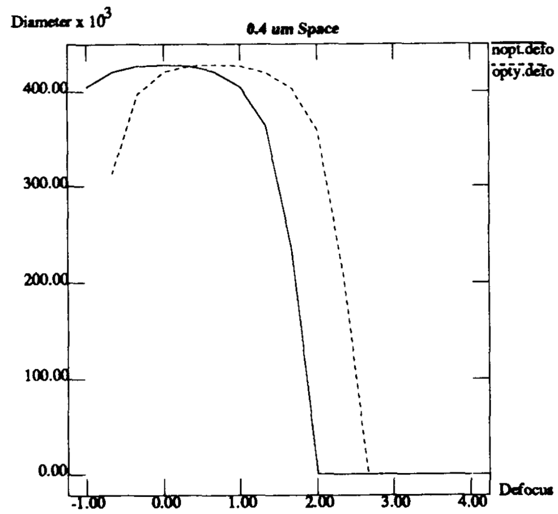


Fig. 28. Bright line linewidths of a conventional 0.4 μm binary mask is shown with the solid line marked with nopt.defo, and for the optimized mask in Fig. 27 is shown with the dashed line marked with opty.defo. The space linewidths are measured using the 0.3 contour of normalized intensity images.

shift because of physical limitations. The vector $\vec{\kappa}$ is chosen to be (1.0, 0.5, 0.3) and the available phases are 0°, 60° and 180°.

4.4 Application of Dual Mask Theorem

The examples shown in Section 4.2.2 all have positive focus shift. Applying the dual mask theorem, masks with negative focus shift plane can be obtained from those with positive focus shift by simply changing the 0° areas to 180°, 180° areas to 0°, and keeping the 90° areas unchanged. Fig. 29 shows the dual mask of the mask in Fig. 19. Fig. 30 shows the contact hole sizes versus defocus for a conventional binary mask and the designed mask of Fig. 29. Although neither of these two masks have an extraordinary long DOF, the two of them used properly for different resist topologies can cover 6 μm depth of focus, which is three times that of the conventional masks.

V. Experimental Results

In this section, we will show experimental results verifying the simulation results of our technique.

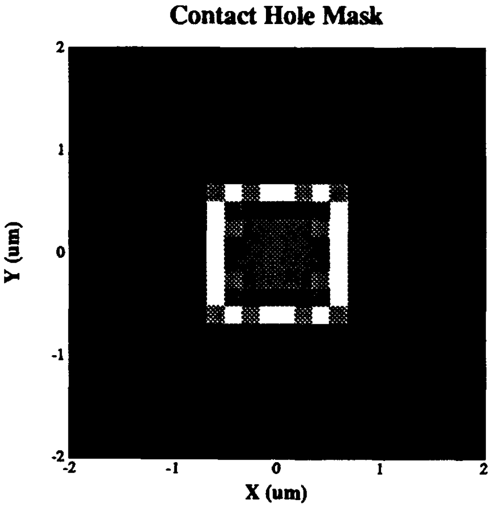


Fig. 29. The dual mask of the one in Fig. 19 or 0.68 μm contact hole, with a focus plane shift of -1.0 μm. In the mask, dark area is opaque; all other areas are transparent with different phases: white for 0°, light shade for 90° and dark shade for 180°.

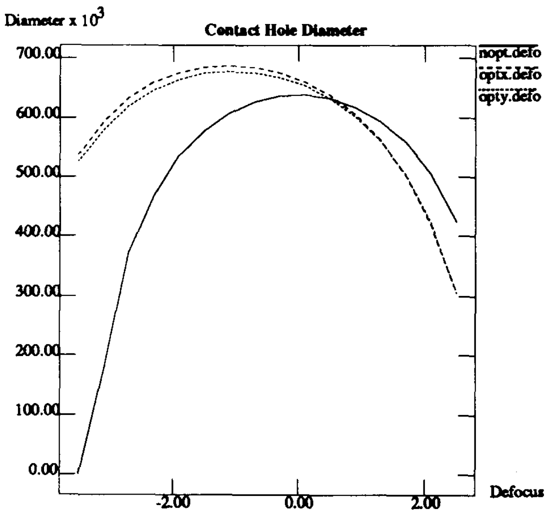


Fig. 30. Contact hole sizes for a conventional 0.68 μm binary mask is shown with the solid line marked with nopt.defo, and for the optimized mask in Fig. 29 is shown with the dashed line marked with optx.defo for diameter in x-direction and opty.defo for diameter in y-direction. The contact hole sizes are measured using the diameter of 0.3 contour of normalized intensity images.

5.1 Mask Fabrication

The predicted behavior of binary phase-shifting features optimized with the proposed algorithm was experimentally verified. A phase-shifting mask with three phase levels of 0°, 90° and 180° was fabricated in the Berkeley Microlab facility. The mask incorporated optimized features and intuitive designs for some of the cases described in Section IV of this paper, i.e., cross phase connectors, line phase connectors and contact holes with an extended range of DOF and a shifted focus plane.

The fabrication process of the test mask is based on an approach using dry RIE etching of the phase-shifter in CVD SiO_2 deposited on the chromium layer of the mask [7]. Quartz wafers of 4″ diameter with a 95 nm thick chromium layer were used as the substrate for the phase-shifting mask fabrication. The process started with delineating the 180° phase-shifting layer using i-line lithography with positive photoresist on a 10× reduction stepper. The use of the lithographic stepper with alignment capability and image demagnification was essential for this fabrication technique. The resist pattern was transferred into the chromium by wet etching.

Deposition of a SiO_2 layer of $\frac{1}{2}[\lambda/2(n-1)]$ thickness on the chromium was used to build up a first phase shifting layer of 90°. Next, the mask pattern of the 90° layer was delineated with the optical lithography stepper. The pattern was then transferred into the SiO_2 layer by RIE using the chromium, which was not yet removed for this layer, as an etch stop. After careful removal of the photoresist and of polymers created by the RIE etch, the chromium was etched in the SiO_2 windows. In a second SiO_2 deposition step the thickness of the 180° phase shifter was completed and the 90° layer was formed. The 0° phase areas were made by using i-line lithography for pattern delineation and by RIE etching the SiO_2 and wet etching the chromium.

Although alignment problems were greatly reduced by using the 10× reduction stepper, misalignment errors of up to 0.3 μm occurred. In cases, where areas of different phases were directly adjacent to each other, misalignment either caused a small trench according to the amount of overetch or a corresponding area of remaining chromium to be formed. The depth of the unwanted trenches was comparatively small for the used overetch of approximately 10%. These structures can be treated as phase-defects of less than 10° phase mismatch. Therefore no significant effect on the imaging properties are expected. However, the remaining chromium in the case of 0.3 μm misalignment causes a severe change in the transmission of the mask feature, strong enough to affect the printed photoresist pattern. The microscope pictures of Figs. 31 and 32 show the masks features of an intuitive and an optimized design, respectively of a phase cross connectors. The 0° layer is misaligned by 0.3 μm with respect to the 90° layer causing a small chromium bridge to remain at some transitions.

5.2 Resist Printing

Phase-shifting masks have been fabricated for the application to both a 0.32 NA i-line (λ = 365 nm) stepper with a 10× reduction factor and a 0.43 NA deep UV (λ = 248 nm) stepper with a 20× reduction lens. For the i-line system the phase mask features were printed in 1 μm thick positive photoresist (Waycoat HiPR 6512) on bare silicon. At deep UV wavelength negative 0.95 μm thick photoresist (Shipley SNR 248) on bare silicon was used. Since no anti-reflection coating was applied the small bandwidth light source leads to a strong standing wave effect which causes the exposure to vary vertically

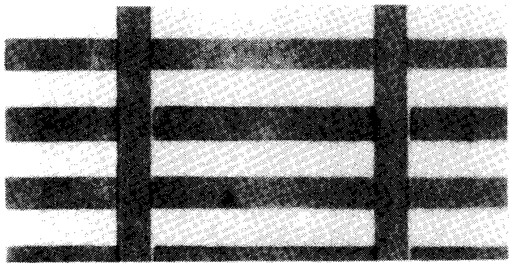


Fig. 31. Actual layout of the simple intuitive cross phase connector mask in Fig. 11.

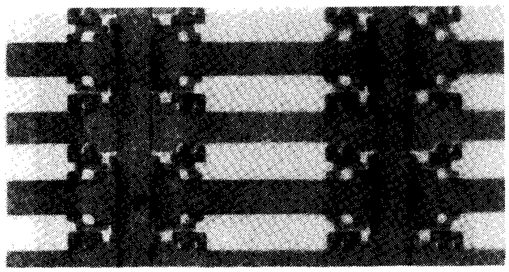


Fig. 32. Actual layout of our optimized cross phase connector mask in Fig. 12.

within the resist. Uniformity was achieved by applying a post exposure bake of 1 min at 120°C and 1 min at 140°C for the *i* line and the deep UV resist, respectively. However, the amount of energy coupled into the resist strongly depends on the reflection of the silicon resist surface. For the case of small bandwidth deep UV exposure with the Excimer laser stepper the energy coupled into the resist changes from 40% to 95% for a resist thickness variation of 50 nm [8]. The used resist process shows a thickness uniformity over the die size of better than 1%, however, absolute thickness measurements could only be performed to about 2% accuracy. Therefore, it was difficult to control the exposure dose to achieve a resist pattern matching the 0.3 intensity contours.

5.2.1 Cross Phase Connector

The cross phase connectors in Figs. 11 and 12 were printed on the *i*-line stepper in positive photoresist at various focus positions. The photoresist was assumed to remain for exposure levels above 0.3 in the optimization.

Figs. 31 and 32 show the actual masks corresponding to the designs of Figs. 11 and 12, respectively. Fig. 33 shows the corresponding SEM micrographs of the actual resist pattern for the printed cross phase connector of Fig. 31 at three focus positions of 0.0 μm and ± 1.0 μm. Throughout the three optical planes the resist pattern shows a high bridging tendency. The left hand side ends of the resist lines, representing the dark mask areas, exhibit a systematic bend. This effect is attributed to a small misalignment error in the mask fabrication process. The large misalignment of the 0° layer with respect to the 90° line causes the closed resist bridge to remain for all the shown optical planes. The aerial image contours show a significant linewidth decrease at the phase transition regions. The effect can still be seen, despite the superimposed influence caused by misalignment. The asymmetric behavior of the line width variation at different optical planes affects the amount of bending at the top ends of the resist lines. Although the effect is too small to compensate the artifact of the misalignment, an increase in the bending effect for 1.0 μm defocus and a decrease for -1.0 μm can be seen.

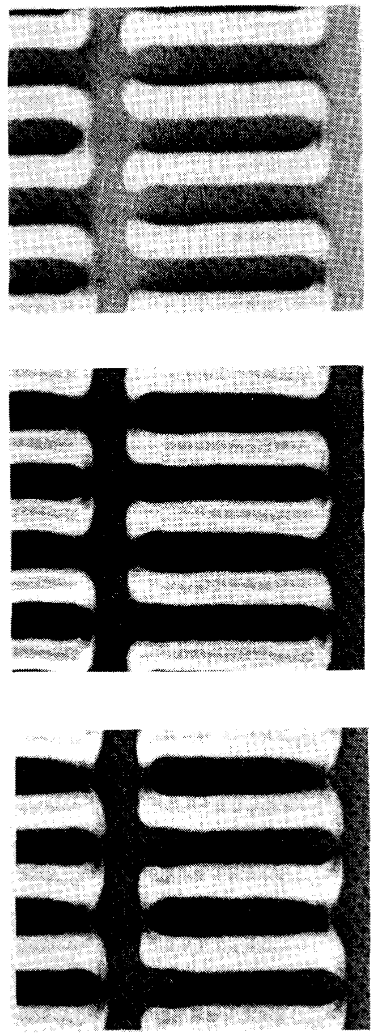


Fig. 33. Actual developed resist pattern of the intuitive cross connector mask in Fig. 31 for different optical planes at 0.0, ± 1.0 μm defocus.

In Fig. 34 SEM micrographs for the optimized cross phase connector are shown for 0.0 μm and ± 1.0 μm defocus. The optimization algorithm leads to a design with reduced linewidth variation at the transition regions as shown by the contour line plots in Fig. 14. Although both the intuitive and the optimized features were fabricated simultaneously on the same mask and were therefore suffering the same misalignment, the resist pattern for the

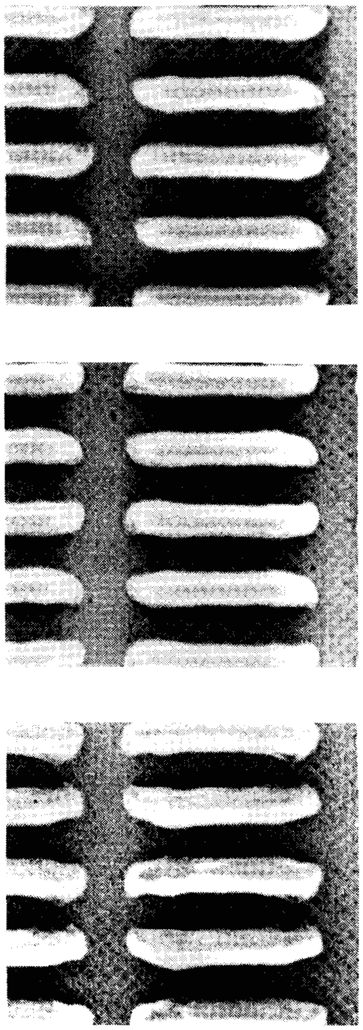


Fig. 34. Actual developed resist pattern of the optimized cross connector mask in Fig. 32 for different optical planes at 0.0, ± 1.0 μm defocus.

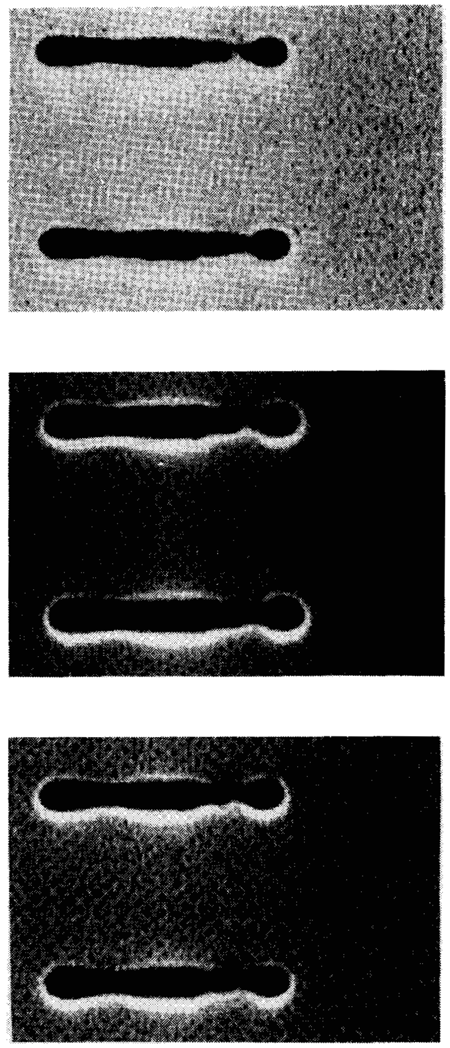


Fig. 35. Actual developed resist pattern of the intuitive line phase connector mask in Fig. 7 for different optical planes at 0.0, ± 1.0 μm defocus.

optimized case shows no increased bridging tendency. Obviously, the optimized design has a reduced sensitivity to misalignment errors. While for the intuitive case misalignment leads to a comparatively large chromium defect, the pixel structure of the optimized cross connector suppresses the formation of large area defects resulting in scattered nonprinting smaller defects. The misalignment still causes a small tendency for the resist line ends to bend. Besides the clear improvement in misalignment sensitivity, the resist pattern of the optimized cross connector shows less line width variation at the phase transition areas. Keeping in mind the effects caused by misalignment, the resist shape matches the corresponding 0.3 intensity contour lines at all given focus positions.

Thus, we have shown experimental evidence of the performance improvement and validity of the proposed pixel based phase-shifting mask feature optimization routine for the case of a cross phase connector. In addition, the optimized design showed a reduced sensitivity to misalignment errors.

5.2.2 Line Phase Connectors

Intuitive and optimized designs for line phase connectors shown in Figs. 7 and 8, were also printed on the *i*-line stepper in positive resist at focus positions of 0.0 μm and ± 1.0 μm. The misalignment errors of 0.3 μm as seen for the phase cross connectors were present both for the intuitive and the optimized design features.

The optimization for the phase line connector was performed with the goal to preserve the desired resist shape at 0.0 μm and ± 1.0 μm defocus position and thus to extend the DOF range. The resist profile was assumed to follow the 0.3 aerial image intensity contour.

In Figs. 35 and 36 SEM micrographs of resist patterns

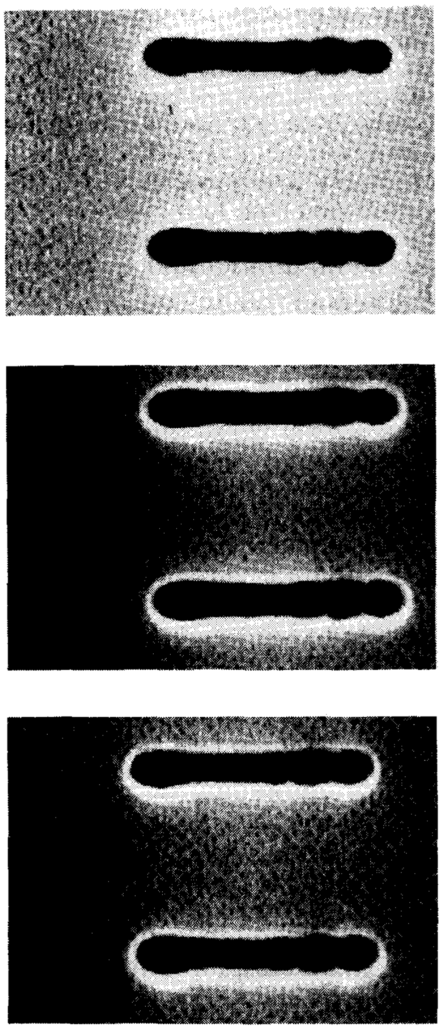


Fig. 36. Actual developed resist pattern of the optimized line phase connector mask in Fig. 8 for different optical planes at 0.0, ± 1.0 μm defocus.

printed with an intuitive and an optimized design mask feature are shown for optical planes of 0.0 μm and ± 1.0 μm. Again, the effect of the misalignment can be seen in both the optimized and, more pronounced, in the intuitive design pattern. The misalignment has a stronger effect at the right hand side of the resist pattern. Here, the mask has large chromium defects rather than phase defects. The optimized design shows reduced linewidth variation as compared to the intuitive design over the given defocus range. The results are in good agreement with the predicted shape obtained from the computed 0.3 intensity contours.

5.2.3 Contact Holes with Shifted Focus Plane and Extended DOF

Contact hole features with conventional design and with optimized design for shifted focus and extended DOF

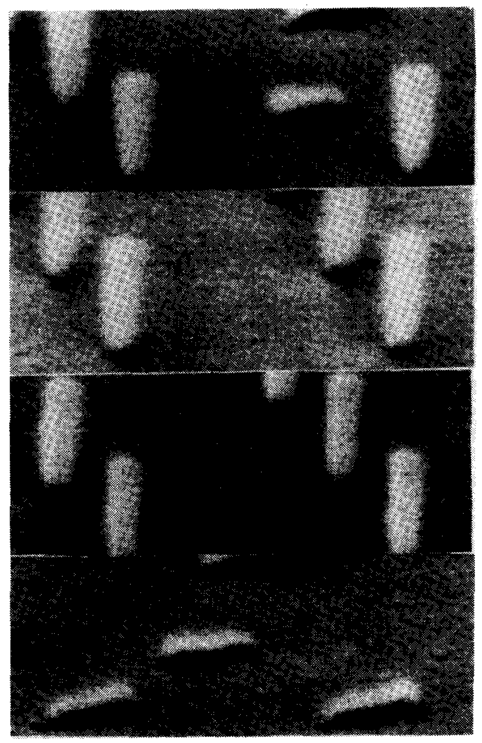


Fig. 37. Actual developed negative resist profile of a conventional, binary, contact hole mask at defocus levels −0.25, 0.0, +0.25, +0.5 λ/NA^2.

range were printed with the deep UV stepper in negative resist. The negative photoresist reversed the topography of the photoresist pattern. Therefore, the contact hole patterns appeared as posts of remaining resist rather than as holes in the photoresist. Negative resist polarity facilitates the SEM analysis of contact holes.

The fabrication process for the deep UV phase-shifting masks was the same as for the *i*-line stepper, except for an adjustment in thickness of phase shifting layers. The layout geometry of the phase-shifting mask features was identical for the *i*-line and the deep UV application. The *i*-line and the deep UV stepper had a 2 : 1 correlation for the resolution factor λ/NA and a compensating relation in the reduction values of 10 : 20, respectively. Therefore, designs for the intuitive and the optimized features with the same scale could be applied on both steppers.

The SEM micrographs in Figs. 37 and 38 show arrays of posts produced by the binary conventional contact hole features and by the optimized design feature of Fig. 19 for various focus positions. The focus steps are $0.25\lambda/\mathrm{NA}^2$. The SEM micrographs enable the determination of the threshold level for the onset of printing. A threshold value of approximately 0.6 was found by comparing the defocus range, where the conventional contact features print, with the computed values of the peak intensity as a function of defocus. The width of the resist posts at the various defocus planes could also be extracted. The accuracy of the width measurement was lim-

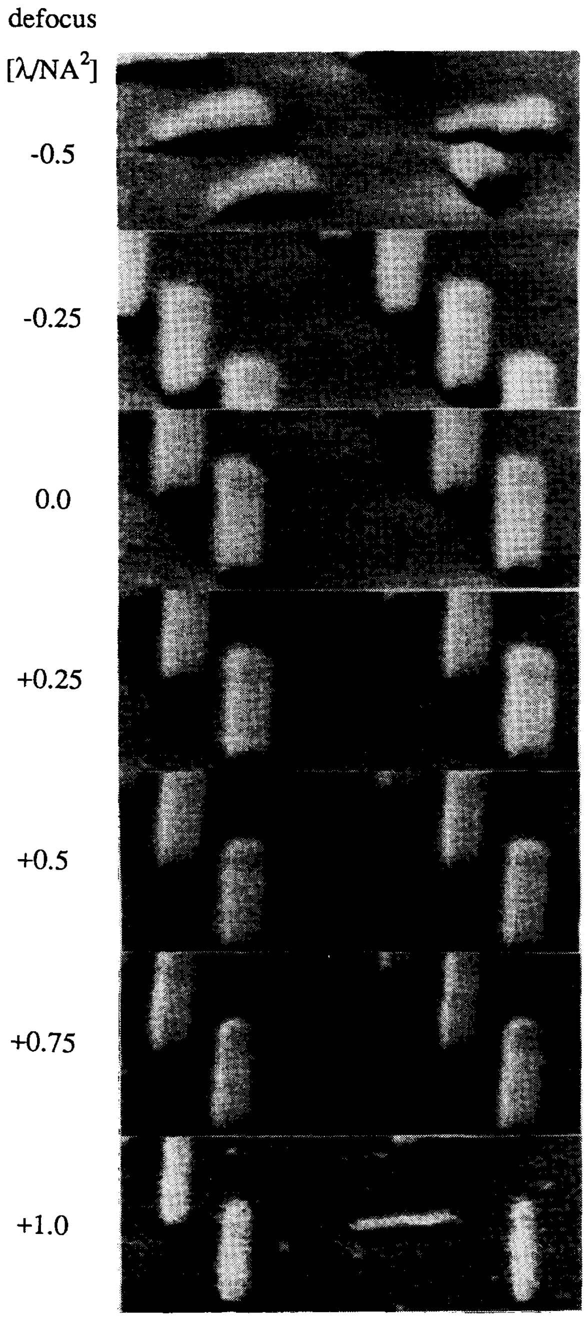


Fig. 38. Actual developed negative resist profile of the optimized contact hole mask in Fig. 19 at different defocus levels.

ited by the nonuniform thickness of the posts which was caused by the absorbing negative photoresist.

The optimized contact holes were designed to extend the DOF range and to shift the focus plane by $0.28\lambda/NA^2$. Although the design was originally performed for 0.3 resist threshold, the contact hole feature designed with our optimization algorithm shows improved performance even though the actual threshold was 0.6. The DOF range is considerably extended and a shift in focus plane was achieved. The image intensity of the optimized contacts reached the threshold over a defocus range of $1.5\lambda/NA^2$ as compared to $0.75\lambda/NA^2$ for the conventional contact holes design.

The experimental data verified the extended DOF range and the shift in focus plane for the optimized design contact hole feature. The quantitative agreement was limited by the measurement accuracy and the resolution in defocus steps.

VI. Conclusion

In this paper, we have proposed an automatic optimization technique that includes defocus effects into consideration. Equipped with this powerful tool, we are able to design masks that solve problems such as phase connection. Design examples on contact holes demonstrate that considerable increase in depth of focus can be achieved using our algorithm. Furthermore, desired focus plane shift can be achieved by specifying it implicitly in the input to the algorithm. Compared to existing approaches to phase-shifting mask design, our approach offers several advantages: (a) First, it is automated and systematic in that little expertise or intuition is needed to design fairly complex masks; (b) Second, our algorithm is general in that it can be applied to any mask pattern, and any phase-shifting mask systems with arbitrary number of phase levels; (c) Finally, our algorithm can accomplish multiple tasks, such as providing bias, compensating diffraction effects, extending DOF and shifting focus plane, all within a single optimization.

Our new concept of dual mask simplifies the design of focus plane shifted masks. The theorem we proved for dual masks is general enough to cover most phase-shifting mask systems under investigation by researchers.

The CPU times for the examples in this paper range from 20 minutes to several hours on a Sparc I Sun workstation, depending on the number of pixels and the "hardness" of the problem. By a hard problem we mean one in which physical limitations prevent the algorithm from achieving the user specified specification for intensity pattern. An example of a hard problem is one in which the user requests a 100% extension of the depth of focus. A promising direction of future research lies in applying parallel processing techniques in order to speed up application of our algorithm to "hard" problems.

To ensure obtaining acceptable solutions from our proposed algorithm, we have to pay attention to the interplay between physical limitations and the parameters of the optimization problem. For instance, if we try to optimize at optical planes which extend over a large range, we might jeopardize the quality of image at the focus plane.

One drawback of our algorithm is that it results in masks which may be expensive to manufacture. To circumvent this, future research should be directed towards modifying our algorithm so that it results in more practical, robust, manufacturable and inexpensive masks. Possible ways of achieving this include reducing the number of pixels, increasing minimum feature size for the mask, allowing pixels to be of different sizes and not on a fixed grid.

In summary, our algorithm provides a general, practical and user friendly optimization tool for phase-shifting mask designs.

APPENDIX

In this Appendix, we prove the dual mask theorem in Section III.

We start with Hopkins equation [9]–[11]:

$$I(f, g) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} T(f' + f, g' + g; f', g') \cdot F(f' + f, g' + g)F^*(f', g')\, df'\, dg' \tag{28}$$

where $I(f, g)$ is the Fourier transform of output image intensity $I(x, y)$, $F(f, g)$ is the Fourier transform of mask transmission function $F(x, y)$, which is usually 1 in transparent areas and 0 in opaque areas of the mask, and $T(f', g', f'', g'')$ in the Transmission Cross-Coefficient (TCC) of the optical system, which summarizes all the information about the imaging system and illumination. The TCC function is given by

$$T(f', g'; f'', g'') = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} J(f, g)K(f + f', g + g') K^*(f + f'', g + g'')\, df\, dg \tag{29}$$

where $J(f, g)$ is Fourier transform of mutual intensity of the light at the object and $K(f, g)$ is frequency response function of the imaging system.

For notational convenience, we define the normalized circular step function R:

$$R(f, g) = \begin{cases} 1 & f^2 + g^2 < 1^2 \\ 0 & f^2 + g^2 \geq 1^2 \end{cases}$$

If circular pupils are used, then the frequency response function on Δz defocus plane can be expressed as

$$K(f, g) = e^{i\phi(f,g)} R(f/r_k, g/r_k) \tag{30}$$

with $r_k = (\mathrm{NA}/\lambda M)^2$ and the phase function as

$$\phi = i\pi\, \Delta z \lambda (f^2 + g^2) \tag{31}$$

where NA is numerical aperture, λ is wavelength of light source, M is reduction factor of the lens system and Δz is the distance of the optical plane to focus plane with positive Δz representing above focus plane.

Assuming incoherent light source, the mutual intensity function J can be expressed by

$$J(f, g) = \frac{1}{\pi r_j^2} R(f/r_J, g/r_J) \tag{32}$$

where $r_J = SNA/\lambda$ and S, the partial coherence factor is: NA of condenser/NA of imaging optics.

First, we show that given a real $J(f, g)$, the TCC function has complex conjugate symmetry property with respect to the focal plane,

$$T_{\Delta z}(f'', g''; f', g') = T_{-\Delta z}(f', g'; f'', g'') = [T_{\Delta z}(f', g'; f'', g'')]^* \tag{33}$$

Since $K_{\Delta z} = e^{i\phi(f,g)}$, $K_{-\Delta z} = e^{-i\phi(f,g)} = K^*_{\Delta z}$, from (29), we have

$$\begin{aligned} T_{\Delta z}(f', g'; f'', g'') &= \\ &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} J(f, g) e^{i\phi(f+f', g+g')} \cdot e^{-i\phi(f+f'', g+g'')} R(f/r_J, g/r_J) \\ &\quad R\left(\frac{f+f'}{r_k}, \frac{g+g'}{r_k}\right) R\left(\frac{f+f''}{r_k}, \frac{g+g''}{r_k}\right) df\, dg \\ &= \Bigg[\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} J(f, g) e^{-i\phi(f+f', g+g')} e^{i\phi(f+f'', g+g'')} R(f/r_J, g/r_J) \\ &\quad R\left(\frac{f+f'}{r_k}, \frac{g+g'}{r_k}\right) R\left(\frac{f+f''}{r_k}, \frac{g+g''}{r_k}\right) df\, dg\Bigg]^* \\ &= T^*_{\Delta z}(f'', g''; f', g') \\ &= T^*_{-\Delta z}(f', g'; f'', g''). \end{aligned} \tag{34}$$

Next, we show that $T_{\Delta z}(f', g'; f'', g'')$ is symmetric with respect to origin if circular pupils are used,

$$T_{\Delta z}(f', g'; f'', g'') = T_{\Delta z}(-f', -g'; -f'', -g''). \tag{35}$$

From (34), we obtain

$$\begin{aligned} T_{\Delta z}(-f', -g'; -f'', -g'') &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} J(f, g) e^{i\phi(f-f', g-g')} e^{-i\phi(f-f'', g-g'')} \\ &\quad R(f/r_J, g/r_J) \\ &\quad R\left(\frac{f-f'}{r_k}, \frac{g-g'}{r_k}\right) R\left(\frac{f-f''}{r_k}, \frac{g-g''}{r_k}\right) df\, dg \end{aligned}$$

$$= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} J(f, g)e^{i\phi[(f+f'), (g+g')]} e^{-i\phi[(f+f''), -(g+g'')]} R(f/r_J, g/r_J) R\left(\frac{f+f'}{r_k}, \frac{+g+g'}{r_k}\right) R\left(\frac{-f-f''}{r_k}, \frac{-g-g''}{r_k}\right) df\,dg$$

$$= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} J(f, g)e^{i\phi[(f+f'), (g+g')]} e^{-i\phi[(f+f''), (g+g'')]} R(f/r_J, g/r_J) R\left(\frac{f+f'}{r_k}, \frac{g+g'}{r_k}\right) R\left(\frac{f+f''}{r_k}, \frac{g+g''}{r_k}\right) df\,dg$$

$$= T_{\Delta z}(f', g'; f'', g''). \tag{36}$$

In the above derivation, we have used the facts that $J(f, g) = J(-f, -g)$, $\phi(f, g) = \phi(-f, -g)$ and $R(f, g) = R(-f, -g)$.

Finally, we derive the conditions on the masks such that the intensity of one mask as Δz plane is the same as the intensity of the other mask at $-\Delta z$ plane. Denote $F(f, g)$ and $G(f, g)$ as the Fourier transform of the mask transmission functions $f(x, y)$ and $g(x, y)$ respectively. Then the corresponding intensity images are

$$I_{\Delta z}^{F}(f, g) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} T_{\Delta z}(f'+f, g'+g; f', g') \cdot F(f'+f, g'+g)F^*(f', g')\,df'\,dg' \tag{37}$$

$$I_{-\Delta z}^{G}(f, g) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} T_{-\Delta z}(f'+f, g'+g; f', g') \cdot G(f'+f, g'+g)G^*(f', g')\,df'\,dg'. \tag{38}$$

From (37), using (33) and (35), we obtain

$$I_{\Delta z}^{F}(f, g) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} T_{-\Delta z}(f', g'; f'+f, g'+g) \cdot F(f'+f, g'+g)F^*(f', g')\,df'\,dg'$$
$$= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} T_{-\Delta z}(f''-f, g''-g; f'', g'') \cdot F(f'', g'')F^*(f''-f, g''-g)\,df''\,dg''$$
$$= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} T_{-\Delta z}(-f''+f, -g''+g; -f'', -g'')F(f'', g'')F^*(f''-f, g''-g)\,df''\,dg''$$
$$= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} T_{-\Delta z}(f'+f, g'+g; f', g') \cdot F(-f', -g')F^*(-f'-f, -g'-g)\,df''\,dg'' \tag{39}$$

Comparing (39) and (38), we conclude that, if

$$G(f'+f, g'+g)G^*(f', g') = F(-f', -g')F^*(-f'-f, -g'-g) \tag{40}$$

then

$$I_{-\Delta z}^{G}(f, g) = I_{\Delta z}^{F}(f, g) \quad \text{or } I_{-\Delta z}^{G}(x, y) = I_{\Delta z}^{F}(x, y). \tag{41}$$

Condition (40) can be expressed as

$$\frac{F^*(-f'-f, -g'-g)}{G(f'+f, g'+g)} = \frac{G^*(f', g')}{F(-f', -g')} = k(f', g'). \tag{42}$$

Clearly the left term should not depend on f nor g because the middle term does not. Instead, they should be equal to a constant $k(f', g')$ depending only on f' and g'. From (42), we obtain $G(f', g') = k^*(f', g')F^*(-f', -g')$. Substituting this in (40), we obtain

$$k^*(f'+f, g'+g)k(f', g') = 1 \qquad \forall f, g, f', g' \tag{43}$$

Moving $k(f', g')$ to the right hand side, we immediately see that $k^*(f'+f, g'+g)$ can not depend on f and g. Hence, it must be independent of to f and g. Taking $f = g = 0$, (43) becomes

$$k^*(f', g')k(f', g') = 1 \quad \text{or } |k(f', g')|^2 = 1 \qquad \forall f', g'. \tag{44}$$

For (44) to be valid for any f' and g', we must have $k(f', g') = e^{i\theta}$ where θ is any constant independent of f' and g'.

A sufficient condition for (41) to be true is then

$$G(f, g) = e^{i\theta}F^*(-f, -g) \qquad \text{or } g(x, y) = e^{i\theta}f^*(x, y) \tag{45}$$

QED.

Acknowledgment

The authors are grateful to, Dr. Mark Levenson, Dr. John Nistler and Dr. Kenji Nakagawa for technical discussions, to Professor Andrew Neureuther for his generous help in many aspects of this project, and to Dr. Richard Lozes and Dr. Andrew Murry for their comments on practical aspects of this project. Thanks also extend to Dr. Kenny Toh for the lithography imaging program SPLAT and other graph programs that made this project easier.

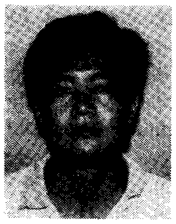
References

[1] Y. Liu and A. Zakhor, "Optimal binary image design for optical lithography," in *SPIE Optical/Laser Microlithography III*, vol. 1264, pp. 401–412, Mar. 1990.

[2] ——, "Binary and phase-shifting image design for optical lithography," in *SPIE Symp. on Optical/Laser Microlithography*, vol. 1263, 1991.

[3] ——, "Binary and phase-shifting mask design for optical lithogra-

phy," *IEEE Trans. Semicond. Manufact.*, vol. 5, no. 2, pp. 138-152, May 1992.
[4] M. D. Levenson, N. S. Viswanathan, and R. A. Simpson, "Improving resolution in photolithography with a phase-shifting mask," *IEEE Trans. Electron Devices*, vol. ED-29, no. 12, pp. 1828-1836, Dec. 1982.
[5] B. J. Lin, "Phase-shifting and other challenges in optical mask technology," in *SPIE 10th Ann. Symp. Microlithography*, vol. 1496, 1990.
[6] M. Mez, G. Paris, and M. A. Virasoro, *Spin Glass Theory and Beyond*. Singapore: World Scientific Publishing, 1987.
[7] A. K. Pfau, W. G. Oldham, and A. R. Neureuther, "Exploration of fabrication techniques for phase-shifting masks," in *SPIE Optical/Laser Microlithography IV*, vol. 1463, pp. 112-123, 1991.
[8] B. Kuyel and H. Sewell, "The effects of broadband 250 nm illumination on process latitude," *J. of Vacuum Science and Technology*, vol. 88, no. 6, pp. 1385-1391, Nov./Dec. 1990.
[9] H. H. Hopkins, "On the diffraction theory of optical images," in *Proc. Royal Soc. Series A*, vol. 217, no. 1131, pp. 408-432, 1953.
[10] M. Born and E. Wolf, *Principles of Optics*. London: Pergamon, 1980.
[11] K. K. H. Toh, "Two-dimensional images with effects of lens aberrations in optical lithography," Memorandum UCB/ERL M88/30, University of California, Berkeley, May 20, 1988.



Yong Liu was born in Yunnan province in China. He received the B.S. degree from Tsinghua University (China) in 1984 and the M.S. degree from University of California at Berkeley in 1989, both in electrical engineering.

He is currently studying toward his Ph.D. degree at the University of California at Berkeley in the department of electrical engineering and computer science. His current research interests include signal processing, microlithography, electromagnetics and microwave.

Anton Pfau was born in 1958 in Mühldorf, Germany. He received the Diploma in physics from the Technical University of Munich in 1986 with a thesis on proximity effects on GaAs MESFET.

He then worked with the Max-Planck-Institut for Quantumoptics in Garching/Munich and with Siemens Corp. in Munich on the application of deep-ultraviolet optical phase conjugation for photolithography. In 1990 he received the degree of a Dr. rer. nat. with a thesis on his research work. From 1990 to 1992 he was with the Electronics Research Laboratory of U.C. Berkeley where his research topics included deep-ultraviolet lithography and phase-shifting mask technology. His future research interests concern novel lithographic techniques, particularly for 256MB DRAM production.

Avideh Zakhor (S'87-M'87) received the B.S. degree from California Institute of Technology, Pasadena, and the S.M. and Ph.D. degrees from Massachusetts Institute of Technology, Cambridge, all in electrical engineering, in 1983, 1985, and 1987 respectively.

In 1988, she joined the Faculty at U. C. Berkeley where she is currently Assistant Professor in the Department of Electrical Engineering and Computer Sciences. Her research interests are in general area of signal processing and its applications to images and video, and biomedical data. She has been a consultant to a number of industrial organizations and holds four U.S. patents, including one on phase shift mask design.

Dr. Zakhor was a General Motors scholar from 1982 to 1983, received the Henry Ford Engineering Award and Caltech Prize in 1983, was a Hertz fellow from 1984 to 1988, received the Presidential Young Investigators (PYI) award, IBM junior faculty development award, and Analog Devices junior faculty development award in 1990, and Office of Naval Research (ONR) young investigator award in 1992. She is currently Associate Editor for IEEE TRANSACTIONS ON IMAGE PROCESSING, and a member of the Technical Committee for Multidimensional Digital Signal Processing.