Improving Resolution in Photolithography with a Phase-Shifting Mask

MARC D. LEVENSON, MEMBER, IEEE, N. S. VISWANATHAN, MEMBER, IEEE, AND ROBERT A. SIMPSON, MEMBER, IEEE

***Abstract*—The phase-shifting mask consists of a normal transmission mask that has been coated with a transparent layer patterned to ensure that the optical phases of nearest apertures are opposite. Destructive interference between waves from adjacent apertures cancels some diffraction effects and increases the spatial resolution with which such patterns can be projected. A simple theory predicts a near doubling of resolution for illumination with partial incoherence $\sigma < 0.3$, and substantial improvements in resolution for $\sigma \leqslant 0.7$. Initial results obtained with a phase-shifting mask patterned with typical device structures by electron-beam lithography and exposed using a Mann 4800 10× tool reveals a 40-percent increase in usuable resolution with some structures printed at a resolution of 1000 lines/mm. Phase-shifting mask structures can be used to facilitate proximity printing with larger gaps between mask and wafer. Theory indicates that the increase in resolution is accompanied by a minimal decrease in depth of focus. Thus the phase-shifting mask may be the most desirable device for enhancing optical lithography resolution in the VLSI/VHSIC era.**

INTRODUCTION

WHEN LIGHT is transmitted through an aperture, the laws of diffraction cause the intensity to spread into the dark regions surrounding the transmitted beam. An imaging optical system can collect much the transmitted light and form an image of the aperture, but the finite numerical aperture of any such system causes the image to spread nonetheless [1]. The images of two neighboring apertures are said to be resolved when the intensity in the dark region between their images falls to some definite fraction of the intensity at the maxima [2]. When the illumination of the apertures is incoherent, the intensity between them can be calculated by summing the intensities diffracted from the individual apertures. The case of coherent and partially coherent illumination is less simple; for fully coherent light, the electric fields due to the waves diffracted from the individual apertures must be summed and then squared to yield the intensity. Constructive interference between the fields diffracted by the apertures in a transmission mask maximizes the intensity between the apertures, thereby reducing the resolution of any optical system with coherent illumination [3]. Fig. 1(a) illustrates this case which approximates the situation in most projection lithography exposure tools.

Manuscript received April 9, 1982; revised July 8, 1982.

M. D. Levenson is with IBM Research Laboratory, San Jose, CA 95193.

N. S. Viswanathan is with IBM General Products Division, San Jose, CA 95193.

R. A. Simpson is with IBM System Products Division, Hopewell Junction, NY 12533.

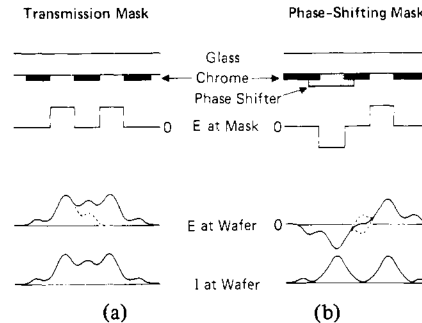


Fig. 1. Comparing the diffraction optics of an ordinary transmission mask with the phase-shifting mask. The transmission mask illustrated in (a) consists of a glass plate covered with an opaque chromium layer with apertures to define the desired intensity pattern. At the mask, the electric field corresponding to the intensity pattern has the same phase at every aperture. Diffraction and the limited resolution of the optical system spreads the electric field patterns. A single aperture would be represented at the wafer by the dotted electric field pattern. Constructive interference between waves diffracted by adjacent apertures enhances the field between them. The intensity pattern is proportional to the square of the electric field. The phase-shifting mask in (b) differs from the transmission mask in that a transparent phase-shifting layer covers adjacent apertures. This layer reverses the sign of the electric field corresponding to the covered aperture. The intensity pattern at the mask is unchanged. At the wafer, destructive interference between waves diffracted from adjacent apertures minimizes the electric field and intensity between the apertures.

If it can be arranged that the waves transmitted through adjacent apertures are 180° out of phase with one another, destructive interference minimizes the intensity between their images [3]. Such a situation occurs when a transparent layer of thickness $d = \lambda/2(n - 1)$ (where n is the index of refraction and λ is the wavelength) covers one aperture as in Fig. 1(b). Any given optical system will project the images of such a *phase shifting transmission object* with better resolution and higher contrast than a corresponding transmission object without phase shifts. Such an improvement in resolution and contrast can be highly valuable in fine line optical lithography.

We have shown that this phase shifting phenomenon can be applied to the complex geometries found in masks used for the production of microelectronic devices by fine line optical lithography. Such a mask we term a *phase-shifting* or *ϕ-mask* which contrasts with a conventional transmission or T-mask. We report here a series of experiments which demonstrates an improvement in resolution that can be obtained with conventional projection exposure tools and conventional partially coherent illumination. Also presented is a simple theory which estimates the contrast ratio of the projected image and predicts

0018-9383/82/1200-1828$00.75 © 1982 IEEE

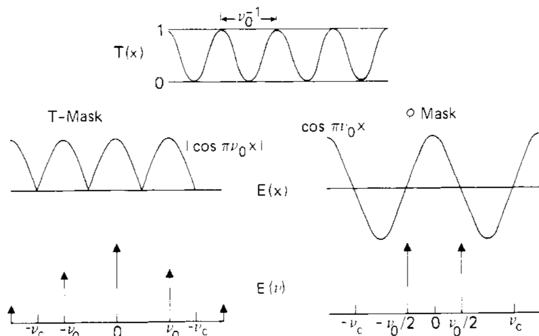


Fig. 2. Fourier optics analysis of phase-shifting and transmission masks. In each case, the intensity transmitted by the mask is T(x). The electric field of the transmission mask corresponding to this intensity pattern is at left, along with its Fourier spectrum $E(\nu)$. The electric field pattern and Fourier spectrum of the phase-shifting mask is at right. Optical systems transmit Fourier components up to a critical frequency ν_c. Since the phase-shifting mask requires less optical bandwidth to transmit the information in T(x), higher spatial frequencies can be transmitted.

a *doubling* of the useful resolution when a phase shifting mask is used with optimal illumination.

Theory

The conventional estimate of the resolving power of a projection system for transmission objects is the critical frequency

$$\nu_c = \frac{\text{N.A.}}{\lambda} \tag{1}$$

which is related to the wavelength λ and numerical aperture (N.A.) and assumes coherent illumination [4]. With incoherent illumination some spatial modulation is transmitted through such a system up to a frequency $2\nu_c$; but, for coherent illumination, the modulation transmission function is unity up to ν_c and zero for $\nu > \nu_c$ [3]. This critical frequency remains a useful parameter when the images of objects with phase and intensity modulations are projected, but the usual definition of the modulation transfer function becomes misleading.

For the purpose of calculation it is convenient to define a transmission function for the mask that has sinusoidal intensity transmission as at the top of Fig. 2

$$T = \tfrac{1}{2}\,(1 + \cos 2\pi\nu_0 x) \tag{2}$$

and shows 100-percent contrast. The spatial frequency for the intensity at the plane immediately past the mask is ν_0. The electric field at this plane can have two possible forms when the illumination is normally incident and coherent. A transmission mask would produce an electric field pattern proportional to $|\cos \pi\nu_0 x|$ as at the left of Fig. 2, while the phase-shifting mask yields the pattern on the right where $E \propto \cos \pi\nu_0 x$. Both patterns can be Fourier analyzed as shown at the bottom of the figure. The transmission mask yields a Fourier spectrum of the form

$$E_T(\nu) = \frac{4}{\pi}\left\{\frac{1}{2}\delta(\nu) + \frac{1}{6}\,[\delta(\nu - \nu_0) + \delta(\nu + \nu_0)] - \frac{1}{30}\,[\delta(\nu - 2\nu_0) + \delta(\nu + 2\nu_0)] + \cdots\right\} \tag{3}$$

where $\delta(\nu)$ is the Dirac delta function and higher harmonics have been ignored. The phase-shifting mask projects only two Fourier components.

$$E_\phi(\nu) = \tfrac{1}{2}\,[\delta(\nu - \nu_0/2) + \delta(\nu + \nu_0/2)]. \tag{4}$$

The same information—in the sense of an optical intensity pattern—is transmitted in each case, but the ϕ-mask requires less spatial bandwidth than the transmission mask. These Fourier spectra appear at the bottom of Fig. 2.

The optical system through which these Fourier components must propagate can be modeled as having a response function equal to unity for $\nu \leqslant \nu_c$ and zero for $\nu > \nu_c$. Thus the Fourier components of the electric field at the image plane are (except for a scale factor that can be set to unity) identical to those at the mask plane for $\nu < \nu_c$ and zero otherwise. From this treatment it is immediately evident that the critical frequency for transmission of modulation produced by a ϕ-mask is $\nu_0 = 2\nu_c$, twice that of a T-mask.

The contrast and resolution for an image produced by a ϕ-mask is most conveniently discussed in terms of a "contrast function" which must be defined for each intensity pattern. For the sinusoidal mask pattern of (2) the contrast function for the corresponding image can be defined as

$$CF(\nu_0) = \frac{I_{\max} - I_{\min}}{I_{\max} + I_{\min}} \tag{5}$$

where the intensity maxima occur for $x = 0,\ 1/\nu_0,\ 2/\nu_0$ and the minima are presumed to occur for $x = 1/2\nu_0,\ 3/2\nu_0$, etc. The intensity at the image plane is obtained by Fourier transforming the electric field pattern projected by the model optical system and squaring the resulting fields. For the transmission mask with coherent illumination the resulting intensity pattern for $1/2\nu_c < \nu_0 < \nu_c$ is

$$I_T(x) = \frac{16}{\pi^2}\left[\frac{1}{2} + \frac{1}{3}\cos 2\pi\nu_0 x\right]^2 \tag{6}$$

where for the phase-shifting mask the intensity pattern is

$$I_\phi(x) = \tfrac{1}{2}\,(1 + \cos 2\pi\nu_0 x) \tag{7}$$

for all $\nu_0 < 2\nu_c$.

The contrast function for coherent illumination can readily be calculated in each case. The top graph in Fig. 3 illustrates the difference between the T-mask and ϕ-mask. Since less optical bandwidth is required in the latter case, the contrast is always higher, and the contrast remains 100 percent for the sinusoidal intensity pattern until the cutoff at $2\nu_c$.

Partially coherent illumination is usually parameterized in terms of an incoherence factor σ which describes the fraction of the numerical aperture of the projection lens filled by the illuminating beam [5]. Totally coherent light corresponds to $\sigma = 0$ while true total incoherence occurs for $\sigma = \infty$. Most projection lithography exposure tools employ illumination with a partial incoherence of roughly $\sigma = 0.7$, and illumination with $\sigma \geqslant 1$ is often treated as totally incoherent [2].

The influence of reduced coherence on the formation of images of phase-shifting transmission objects can be modeled by employing the angular spectrum representation for the

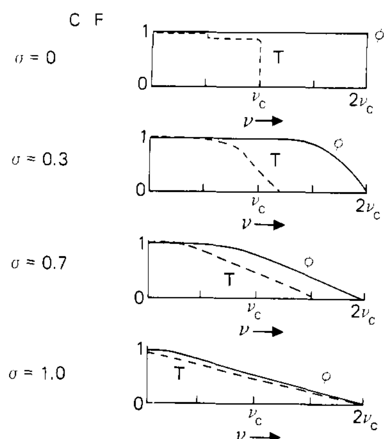


Fig. 3. An illustration of the contrast function defined in (5) for transmission and phase-shifting masks as a function of the input spatial frequency ν and the partial incoherence factor σ. The contrast function (*CF*) for phase-shifting masks is labeled ϕ, while for transmission masks it is labeled T. The intensity pattern employed for calculation is the sinusoid of (2). The phase-shifting mask transmits more contrast to higher frequency in each case. These plots illustrate the performance expected for ϕ-masks and T-masks under different illumination conditions, but these plots do not describe conventional modulation transfer functions.

illumination and allowing each plane wave component of the illumination to be incoherent with all the others [6]. Each of the plane wave elements of the illumination is diffracted independently by the object, producing Fourier components that are mutually coherent according to the laws of physical optics. At the image plane, the electric fields due to the Fourier components produced by a given element of the illumination angular spectrum are summed and squared to obtain the intensity due to a given illumination element, and the intensities due to the various illumination components are summed to give the total intensity.

Mathematically, the illuminating beam at the mask plane ($z = 0$) is then decomposed as

$$E_I = \int_{-\infty}^{\infty} E'(q)\, e^{i[qx+\phi(q,t)]}\, dq \tag{8}$$

where $\phi(q, t)$ is a random number and the illumination distribution function has the form

$$E'(q) = \begin{cases} \left(\dfrac{I_I}{2\sigma\nu_c}\right)^{1/2} & |q| \leqslant \sigma\nu_c \\ 0 & \text{otherwise} \end{cases} \tag{9}$$

and where the total incident intensity is

$$I_I = \int_{-\infty}^{\infty} |E'(q)|^2\, dq.$$

The mask acts independently on each element of this illumination, producing Fourier components of the form (3) and (4) with $\nu - q$ replacing ν and $E'(q)$ multiplying the expressions.

Partial coherence qualitatively alters the imaging process. The presence of spatial components of the illumination with many values of the transverse wavenumber q results in the shifting of Fourier components from their positions with coherent plane wave illumination ν_i to a variety of new positions given by $\nu_i + q$. These new positions can be either nearer or further from the transmittance cutoff ν_c. Spatial components that would have been completely transmitted with coherent illumination are thus transmitted only for illumination components with low values of q (i.e., those near the center of the angular spectrum of illumination). Their contrast is thus reduced when $\sigma \neq 0$. Fourier components of the mask that would not have been transmitted at all with coherent illumination can be transmitted due to illumination by large q value components in the angular spectrum. For a transmission mask, this effect extends the contrast function beyond ν_c. Thus the image contrast produced by a mask pattern depends upon the σ value of the illumination as well as upon the mask spatial frequency and the numerical aperture of the imaging system. The summing of intensities at the image plane due to different illumination components also averages out effects that depend critically on the details of the illumination. Thus our rather crude model may be used to approximate systems where the illuminating beam is not composed of mutually incoherent plane waves, where the angular spectrum of the illumination is complex and where the response function of the imaging system is other than as described.

The cutoff of the response function

$$\Theta(\nu') = \begin{cases} 1 & |\nu'| \leqslant \nu_c \\ 0 & \text{otherwise} \end{cases} \tag{10}$$

of the imaging system eliminates all Fourier components with $|\nu'| > \nu_c$. Thus the intensity patterns at the image plane differ for different elements of the illumination. Defining the intensity pattern at the image plane due to an illumination element at q with unit amplitude (i.e., $E'(q) = 1$)

$$I_q(x) = \left| \int_{-\infty}^{\infty} \Theta(\nu) E_{\underset{T}{\phi}}(\nu - q)\, e^{-i\,(\nu-q)x+\phi(q,t)}\, d\nu \right|^2 \tag{11}$$

where the appropriate form of the Fourier spectrum of the mask $E_{\underset{T}{\phi}}(\nu)$ is obtained from (3) or (4) as required. The total intensity at the image plane is then calculated by summing over illumination elements.

$$I(x) = \int_{-\infty}^{\infty} |E'(q)|^2\, I_q(x)\, dq. \tag{12}$$

The resulting intensity modulation can be estimated for sinusoidal intensity patterns projected by transmission and phase-shifting masks by noting the effects of deleting Fourier components on the intensity at the image plane resulting from a given spatial component of the illumination. In the case of a transmission mask, the deletion of one component at $\nu = \pm 2\nu_0$ reduces the intensity at maximum $I_{q_{\max}}$ from unity to 0.95 and increases the minimum intensity $I_{q_{\min}}$ from 0 to 0.016. The deletion of both components at $2\nu_0$ implies that $I_{q_{\max}} = 1.12$ and $I_{q_{\min}} = 0.045$. The further elimination of one component at $\nu = \pm\nu_0$ gives $I_{q_{\max}} = 0.72$ and $I_{q_{\min}} = 0.18$, while the elimination of all sidebands gives a constant intensity of illumination $I_{q_{\max}} = I_{q_{\min}} = 0.40$.

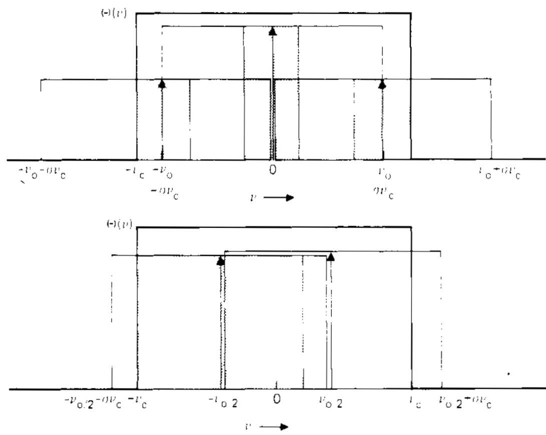


Fig. 4. An illustration of the reason for the improved performance of the phase-shifting mask when the partial incoherence is $\sigma = 0.8$ and the input spatial frequency is ν_0. The diagram at top shows the Fourier spectrum produced by a transmission mask in this case along with the response function $\Theta(\nu)$ for the optical system. Only in the portion of the spectrum shaded are both sidebands at $\pm\nu_0$ transmitted. The lower diagram illustrates the case for the corresponding phase-shifting mask. Again the broadening of the Fourier spectrum due to the partially incoherent light source prevents some components being transmitted by the optical system, but both sidebands are transmitted for a larger portion of the total illumination. Again the shaded region corresponds to the components where both sidebands are transmitted. For clarity, shading appears in the broadened spectrum of one sideband only.

The situation is different for the ϕ-mask case where only two Fourier components exist. When both sidebands are transmitted $I_{q_{max}} = 1$ and $I_{q_{min}} = 0$. Elimination of either one yields constant illumination with $I_{q_{max}} = I_{q_{min}} = 0.25$. The elimination of both sidebands reduces the illumination at the image plane to zero. The contrast function in (5) is then estimated by determining what proportion of the total illumination produces each degree of modulation and computing the appropriate weighted sum for both I_{max} and I_{min}.

An example may help clarify this procedure. Consider the case where $\nu_0 = 0.8\nu_c$ and $\sigma = 0.8$ that is illustrated at the top of Fig. 4. For $\frac{1}{4}$ of the illumination, the two sidebands at $\pm\nu_0$ are transmitted while for $\frac{3}{4}$ the illumination one is blocked. The maximum intensity is thus estimated as

$$I_{max} = \tfrac{1}{4}(1.12) + \tfrac{3}{4}(0.72) = 0.82$$

while the minimum intensity is

$$I_{min} = \tfrac{1}{4}(0.045) + \tfrac{3}{4}(0.18) = 0.15.$$

The contrast function is

$$CF(0.8\nu_c) = 0.69.$$

For the ϕ-mask, illustrated at bottom of Fig. 4, both sidebands are transmitted for $\frac{6}{8}$ of the illumination while one is blocked for $\frac{2}{8}$ of the illumination. The contrast function is then

$$CF(0.8\nu_c) = \frac{0.75(1) + 0.25(0.25) - [0.75(0) + 0.25(0.25)]}{0.75(1) + 0.25(0.25) + [0.75(0) + 0.25(0.25)]} = 0.86.$$

The improvement in contrast is clearly evident for the ϕ-mask even in this case. The results of a series of similar estimates are illustrated in Fig. 3 which shows that the phase-shifting transmission object projects higher contrast to higher spatial frequencies for every degree of partial coherence expected in projection lithography. Identical results are obtained using the F_1-F_2 plot technique [9]. The smaller the value of σ, the larger the improvement in contrast. In the totally incoherent case ($\sigma = \infty$) the phase-shifting aspects becomes immaterial, and the contrast functions become identical.

The simplicity of the Fourier spectrum produced by the phase-shifting mask and sinusoidal intensity pattern allows the contrast to be expressed as

$$CF(\nu_0) = \min\left(1, \frac{2\nu_c - \nu_0}{\nu_c + \sigma\nu_c - \nu_0/2}\right) \tag{13}$$

for all grating frequencies $\nu_0 < 2\nu_c$ where $\min(x, y)$ selects the minimum of the two arguments. No such widely applicable relation exists for the transmission mask sinusoid. However, when $\nu_c/2 < \nu_0 < (1 + \sigma)\nu_c$ the contrast function can be represented as

$$CF(\nu_0) \approx \min\left(0.92, \frac{(\nu_c - \nu_0) + \sigma\nu_c}{(\nu_c - \nu_0)/2 + \frac{5}{3}\sigma\nu_c}\right). \tag{14}$$

If we define the maximum useful spatial frequency as that at which $CF(\nu_0) = 0.6$, (13) and (14) imply that

$$\nu_{max} = (2 - \tfrac{6}{7}\sigma)\nu_c \tag{15}$$

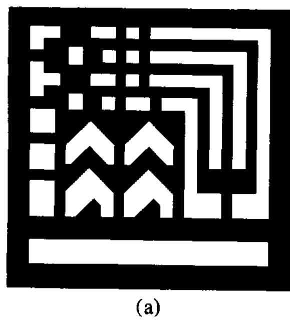
for the ϕ-mask, while $\nu_{max} = \nu_c$ for the transmission mask. Equations (13) and (14) apply for a wide variety of one-dimensional ϕ-mask structures for $\nu_0 > (1 + \sigma)\nu_c/3$. Equations (13), (14), and (15) breakdown for $\sigma \geqslant 1$.

The results for two-dimensional patterns are similar, but perhaps less dramatic. The main requirements for contrast improvement is that the contrast at the mask plane be 100 percent so that nearest apertures can have opposite phases and that each aperture have no nearest neighbors that are also nearest to one another. These requirements are often fulfilled in the masks employed in optical lithography. Further, Fermat's isochrony principle ensures that phase shifts introduced at the object will be propagated to the image by any imaging system [7]. Many of the details of this model were introduced for mathematical conveniences and are not essential to the argument.

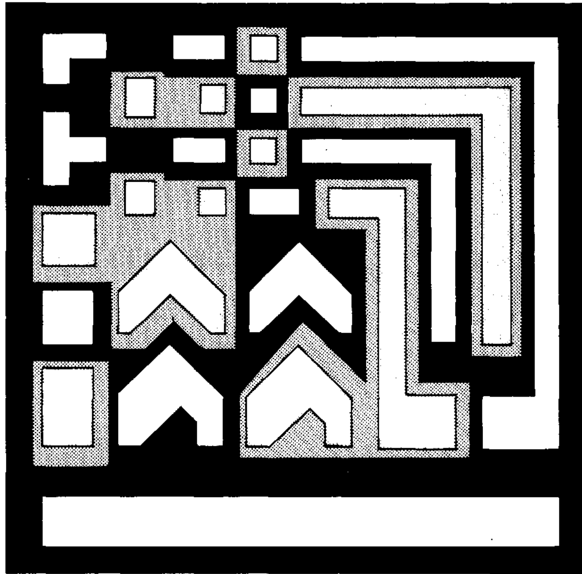
The Test Mask

The improved resolution predicted for a phase-shifting mask was demonstrated using the aperture pattern shown in Fig. 5(a). In this test chip all of the small lines and gaps have the same width—the minimum feature size. This pattern contains a spectrum of typical device geometries, but differed from more common test masks in its increased density. The test chip was patterned on the mask with a variety of magnifications. The overall repetitive pattern can be seen in Fig. 6. The largest test chip in the pattern had a minimum feature size intended to project to 1.5 μm on the wafer. Other chips in the pattern projected feature sizes of 1.0, 0.8, 0.7, 0.6, 0.5, and 0.4 μm.

The design of the pattern for the phase-shifting layer begins



(a)



(b)

Fig. 5. (a) The aperture pattern used in each chip of the test mask. All of the smallest lines and gaps have equal width—the minimum feature size. (b) A phase-shifting mask test chip. The aperture pattern is as in Fig. 5(a), and the transparent phase-shifting layer has been removed in the shaded regions. Regions with dark shading remain opaque, while those with light shading are transparent. Two pairs of adjacent apertures have the same phase; they appear in the pattern at upper left and near the bottom to right of center.

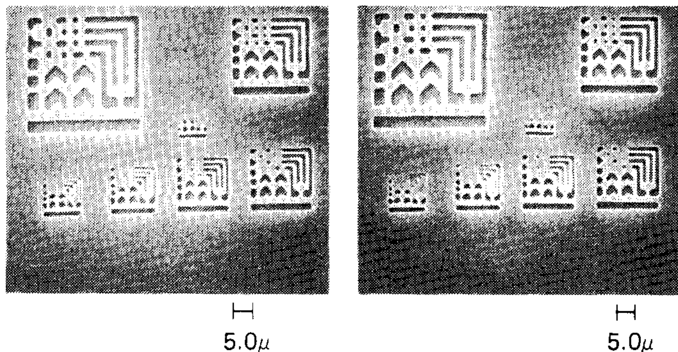


Fig. 6. Scanning electron micrographs of the developed images patterned by transmission and phase-shifting mask chip arrays. The ϕ-mask array is at left. The images were exposed and developed simultaneously and were located 200 μm apart on a bare silicon wafer. The large chip at upper left has a minimum feature size of 1.5 μm. Moving clockwise, the minimum feature size decreases to 1.0, 0.8, 0.7, 0.6, and 0.5 μm, respectively. A partially developed chip with 0.4-μm feature size is at center. The nominal resolution of the GCA Mann 4800 projection exposure tool used to print these patterns was 1.2 μm. The photoresist used was Shipley AZ1470, 1.2 μm thick.

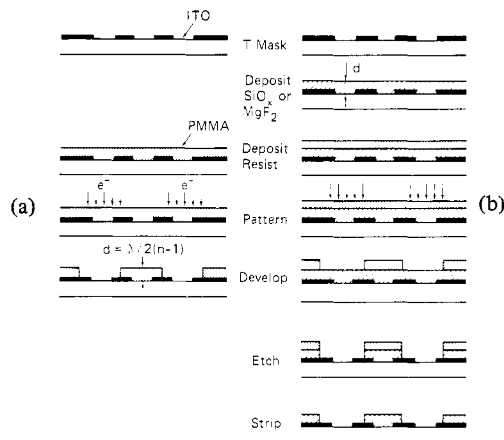


Fig. 7. Two ways of patterning the phase-shifting layer of a ϕ-mask. (a) A transmission mask with a conducting transparent layer of indium tin oxide in the apertures is coated with a transparent electron-beam resist. Subsequently the pattern is exposed by electron-beam lithography and developed to yield the desired phase-shifting mask. The after-development thickness of the phase-shifting layer is $\lambda/2(n-1)$. (b) The alternative process results in a phase-shifting layer of an inorganic transparent material. Any resist and lithography system can be used.

with the identification of the nearest neighbors of each aperture. The smallest gaps between apertures in Fig. 6 are all equal and equal to the minimum feature size. The phase-shifting layer must change thickness across these gaps between nearest neighbors. Several gaps—notably the gap between the large bar at bottom and the structure above—are larger than $1/(\nu_c + \sigma\nu_c)$. The phase shift across these gaps has no optical importance and can be chosen for overall convenience. Since sharp corners will not be faithfully reproduced in the image, the effective gap widths for diagonally adjacent apertures as at upper left will be larger than what appears on the mask. Even with the freedom implied by these design rules, it is not possible to arrange phase shifts for all the gaps in a ring formed by an odd number of apertures such as that which includes the chevrons at lower left. One gap necessarily lacks the optimum phase shift and thus prints differently from the others. The patterns chosen for the phase-shifting layer for a single chip appears in Fig. 5(b) overlain on the aperture pattern of Fig. 5(a).

The procedure for writing the phase-shifting test mask is illustrated in Fig. 7. Mask substrates were procured from Tau Laboratories coated with 750-Å bright chrome under which an indium tin oxide antistatic layer had been deposited. A 6000-Å resist layer of PMMA was coated over the chrome, and the pattern of apertures written by conventional electron-beam lithography. The resist was developed and the chrome etched according to standard procedures. The first layer of resist was then stripped and a second layer of resist spun on. The thickness of this layer was calculated to give one half wavelength of phase shift in the undeveloped regions *after* full processing. A measured with a film thickness analyzer, the initial resist thickness was 4450 Å for the masks used in this study. Two masks were prepared with PMMA second layers, and two with terpolymer layers.

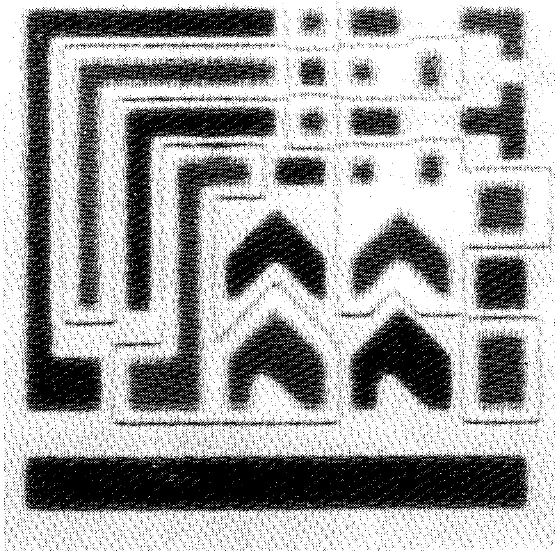


Fig. 8. Light micrograph of a 0.7-μm feature size ϕ-mask chip. The picture was taken in reflected light, thus the apertures appear dark. The narrow lines are the images of steps at the edges of phase-shifting regions. The microscope was focused on the phase-shifting step and consequently the apertures in the chrome mask are slightly defocused.

The pattern for the phase-shifting layer was written using an electron lithography exposure tool which is capable of the necessary overlay precision. In addition, one quarter of the chip patterns on the mask were blanket-exposed by the electron beam so that the phase-shifting layer would be completely removed. The apertures in these transmission or T-mask regions were used as references to facilitate side-by-side comparisons of the images projected by phase-shifting and conventional masks. One quarter of the test chips were not exposed at all in the second lithography state. These chips thus were covered by a continuous transparent layer after development and could be considered T-mask regions. Had the phase-shifting layer introduced absorption or some other extraneous optical effect, these coated T-mask regions would have exposed differently than the uncoated T-mask regions.

The polymeric phase-shifting layer was developed according to standard procedures, and the resist was allowed to remain in place. The measured thickness was 3670 Å to which the 750 Å of the chrome should be added to determine the actual phase difference between apertures. The available index of refraction data indicated $n \approx 1.51$ for both resists at 4360 Å. An alternative procedure for fabricating phase-shifting masks is illustrated in Fig. 7(b). In this case an inorganic phase-shifting layer is coated on the transmission mask before the resist. Exposure can then be performed using any conventional lithographic scheme, after which the resist is developed, the inorganic phase-shifting layer is etched, and the remaining resist removed. The result of this more elaborate procedure would be a phase-shifting mask with a hard, cleanable phase-shifting layer. For the purpose of our initial series of experiments, the polymeric layer has proved sufficiently sturdy.

Fig. 8 is a micrograph of a typical phase-shifting test chip. The narrow lines are the vertical walls at the edges of the regions where the resist has been developed. With two exceptions, such a wall lies between adjacent apertures separated by the minimum feature size. The two exceptions which can also be seen in Fig. 5(b) occur at the upper right of the pattern in Fig. 8 and near the bottom center. The two gaps between apertures having the same phase should project differently from the others in the pattern. It should also be noted that the required overlay precision is not excessive, especially on masks used in 10X step and repeat exposure tools.

Our masks were intended for use in a GCA Mann 4800 step and repeat exposure tool [8]. Each mask consisted of 119 arrays of 28 test chips with seven different minimum feature sizes. One half of the test chips were patterned as phase-shifting masks, and each had an adjacent transmission mask region for comparison. On the mask, the largest feature sizes were 15 μm, while the smallest were 4 μm nominal—well within the capability of present mask making tools. Excellent edge quality was achieved in both the chrome and polymeric patterns. The thickness of the phase-shifting layer was uniform to ±80 Å as measured using an IBM film thickness analyzer.

Exposure and Processing

The series of exposures reported here were performed using the GCA Mann 10X step and repeat camera which has a 0.28-numerical aperture lens with a nominal 1.2-μm resolution over a 5-mm square field. The brightness of the images were roughly 125 mW/cm^2 with monochromic illumination at 4358 Å; the nominal partial coherence was $\sigma = 0.7$.[1] Bare silicon wafer substrates were employed, coated with 1.2 μm of Shipley AZ1470 photoresist. The exposed images were developed using AZ developer diluted 3 : 1 with deionized water.

Once optimal focusing was achieved, a series of exposures was made with exposure times varying from 600 to 1200 ms in 30-ms increments. Optimum exposure on this system appeared to be 840 ms, but the larger structures appeared somewhat overexposed. The optimum developing time was found to be 4 min, followed by a 2-min deionized water rinse.

The developed wafers were diced, coated with 200 Å of sputtered palladium gold and inspected using a Hitachi S-500 scanning electron microscope at 30 kV. There was no attempt at quantitative metrology on these initial exposures.

Initial Results

Scanning electron micrograph pictures of the developed resist patterns resulting from the images formed by typical phase shifting and transmission mask regions on our test mask appear in Fig. 6. The large chip in the upper left of each figure has a minimum feature size of 1.5 μm. Moving around the pattern in a clockwise direction, the minimum feature size decreases to 1.0, 0.8, 0.7, 0.6, and 0.5 μm. A partially developed chip with a 0.4-μm minimum feature size appears at center. There is no perceptible difference between the 1.5 and 1.0 micrometer patterns exposed by ϕ-mask and T-mask regions. Because of the 840-ms exposure time, the corners are rounded in each case, and the resist has developed down to the substrate.

[1] While $\sigma = 0.3$ might have produced even more dramatic results, no 10X camera permitting such high coherence was available.

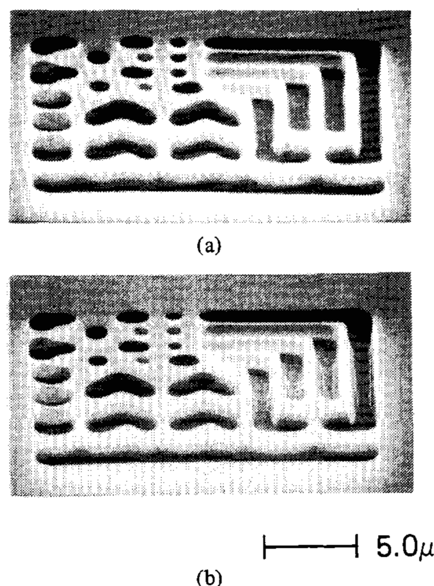


Fig. 9. Scanning electron micrographs of 0.8-μm chips projected from ϕ-mask and T-mask regions. The ϕ-mask patterned chip in (a) shows reduced residue in comparison to the T-mask patterned chip in (b).

The 0.8-μm chip shows some improvement in wall profile for the ϕ-mask case. Fig. 9 shows SEM pictures taken at an incident angle of 70°. Even though the exposure is nominally identical for the T-mask and ϕ-mask, the ϕ-mask pattern shows decreased resist residue in the bottom of the exposed regions. The reduced residue implies that the bright regions of the image are brighter for the ϕ-mask than for the T-mask. The argument of Fig. 1 demonstrates that the optical power that would be diffracted into dark regions of the image in the case of a transmission object is cancelled by destructive interference in the phase shift object. Thus the ϕ-mask makes dark regions darker. Energy is, however, conserved in physical optics, and the power that the phase-shifting masks directs away from the dark regions of the image appear in the bright regions [1]. Thus the bright regions must also be brighter when a phase-shifting mask is used. The reduced residue in Fig. 9 is one verification of this additional advantage for phase-shifting masks.

Fig. 10 shows 0.8-μm patterns exposed for 1020 ms in order to force the small contact holes at the top to develop down to the substrate. Since no allowance for windage was made in designing the mask and since no attempt at reducing the importance of standing wave patterns was attempted, these holes develop down to the substrate only at the center. Nevertheless, it is clear from the photographs at top that the ϕ-mask patterned holes develop down to the substrate over larger areas. The major difference, however, appears in the picture of the entire chip at the bottom of the figure. While the ϕ-mask patterned region shows substantial walls in spite of the overexposure, the region exposed by the T-mask contains walls that have developed though and others which are breaking off. Thus the phase-shifting mask widens the exposure window for 0.8-μm structures patterned using the Mann 4800 step and repeat tool.

Fig. 11 compares ϕ-mask and T-mask exposed chips with 0.7-μm minimum feature size for 960-ms exposure. The smallest contact holes fail to develop to the substrate in this case, but

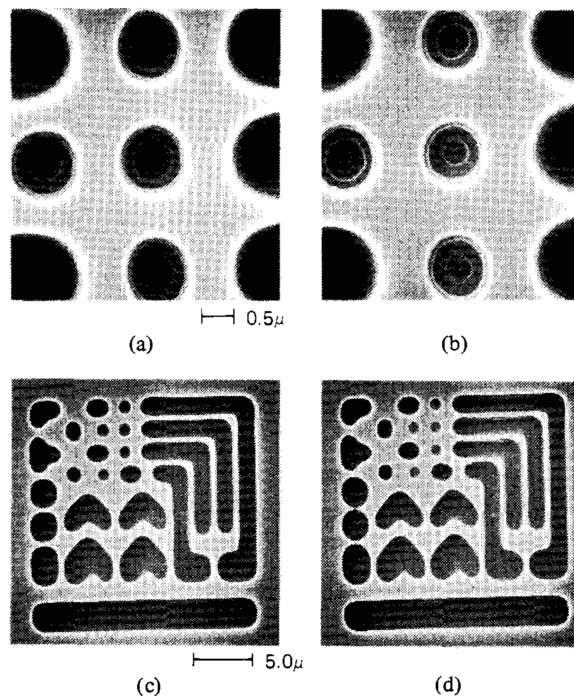


Fig. 10. Comparison of SEM pictures of 0.8-μm test chips exposed for 1020 ms using ϕ-mask and T-mask regions. The ϕ-mask chips in (a) and (c) shows improved development of the small contact hole at top with thicker continuous walls in the more heavily exposed regions of (c). In comparison the contact holes of the T-mask patterned chip shown magnified in (b) are less developed, but the walls between parallel lines at top right in (d) are flaking off while other walls have punched through. The extra exposure due to constructive interference has seriously degraded the wall profile.

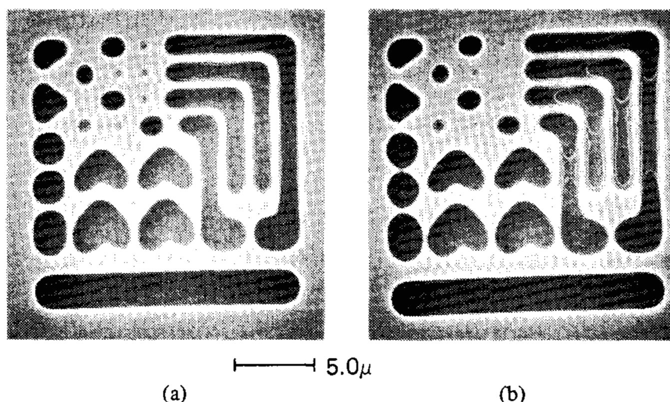


Fig. 11. Comparison of 0.7-μm minimum feature size chips. The ϕ-mask patterned chip in (a) is of a quality usable in photolithography. In contrast, the T-mask patterned chip in (b) shows a great deal of residue in unexpected regions as well as poor wall profiles.

the parallel lines and large pads do develop to the substrate without residue in the ϕ-mask case. The T-mask patterned region shows considerable residue in exposed areas and numerous thin and broken walls. The quality of the pattern obtained for 0.7-μm minimum feature size with the phase shifting is fully comparable to that achievable for 1.0 μm using a transmission mask. Thus in the first series of exposures we have demonstrated a 40-percent improvement in usable resolution when a ϕ-mask is employed.

The 1.2-μm resist thickness became a serious problem when

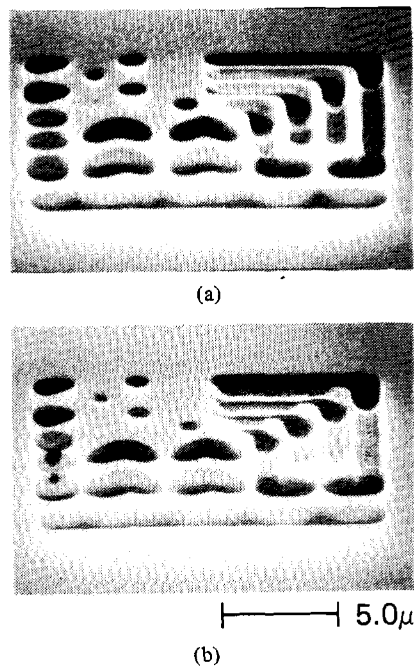


Fig. 12. SEM pictures of 0.6-μm chips taken at 70° angle. The great improvement of the φ-mask patterned chip in (a) is clearly evident in comparison with the T-mask patterned chip in (b). The 1.2-μm thick resist did not develop down to the substrate at these dimensions.

the minimum feature size was 0.6 μm or smaller. While we were able to obtain surface development with 0.6-μm line and gap dimensions with the φ-mask, we were never able to develop the patterns down to the substrate. There always remained some residue in the smaller features, although considerably less for the phase-shifting mask than for the transmission mask. Fig. 12 shows a pair of 70° SEM pictures for 0.6-μm chips exposed for 840 ms. The φ-mask patterned chip has continuous vertical walls of substantial thickness (except possibly for the gap at bottom center not protected by a phase shift) whereas the walls on the T-mask patterned chip have collapsed and punched through in many places. The shapes of the isolated bright regions have definitely degraded because of the limited resolution of the projection lens. The images of the more widely spaced apertures on this mask are formed independently and do not partake of the resolution enhancement due to destructive interference. Thus the closely spaced 0.6-μm lines at upper right more closely resemble the corresponding apertures than do the more widely spaced chevrons at bottom center.

Fig. 13 shows two views of the chips patterned with 0.5-μm minimum features. This represents the ultimate limit of the Mann 4800 exposure tool with a phase-shifting mask. Except for the one gap between apertures with the same phase on the mask, all of the walls on the φ-mask patterned chip are continuous. The T-mask patterned chip produced under the same exposure and development conditions shows few unperforated walls and no sign of the 0.5-μm line and gap structure at upper right. With thinner resist and control of standing wave effects, practical devices could be prepared at this resolution using the φ-mask.

Fig. 14. shows our best 0.4-μm results obtained with the Mann 4800 tool. Some walls between the images of large

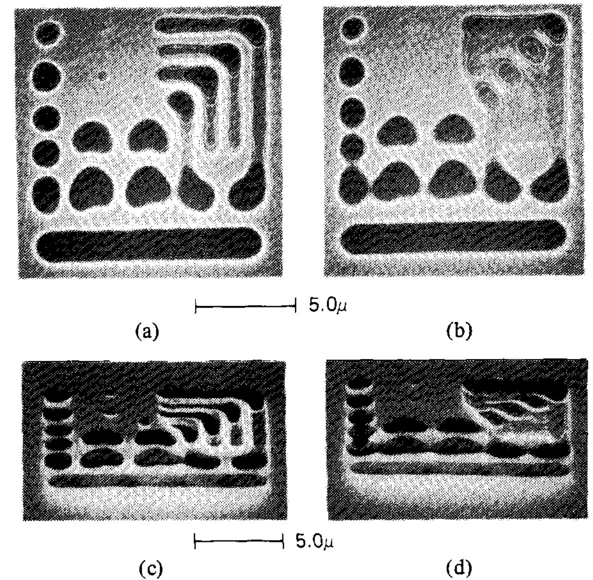


Fig. 13. Comparison of SEM pictures of φ-mask and T-mask patterned chips with 0.5-μm minimum feature size. The φ-mask patterned chip in (a) and (c) shows continuous walls except for the wall between two apertures of the same phase at right of center near the bottom. In contrast the T-mask patterned chip in (b) and (d) shows a near absense of 0.5-μm detail.

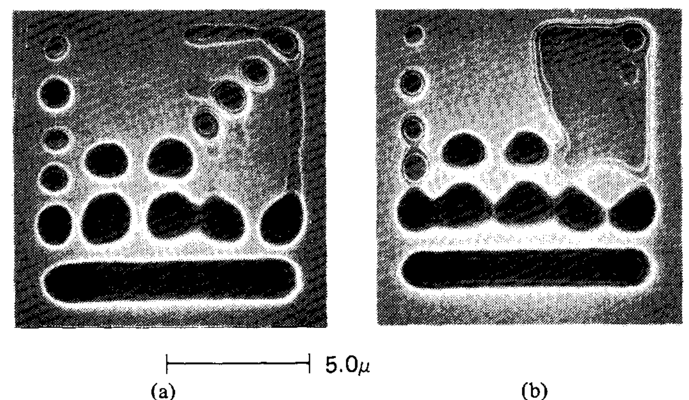


Fig. 14. Electron micrographs of 0.4-μm φ-mask and T-mask patterned chips. The Mann 4800 tool cannot pattern structures at the required 1250 line/mm frequency, even with a φ-mask, but the walls that are there have such improved profiles in the φ-mask patterned chip in (a) than in the T-mask patterned chip in (b). The two apertures at lower right of center in (a) have the same phase on the φ-plate.

apertures appear in the φ-mask patterned chip while the T-mask patterned region is mostly a useless blur. When the spatial frequencies produced by a mask exceed the maximum frequency that can appear in the image, a transmission mask projects uniform illumination. A phase-shifting mask projects a dark region; hence the increased exposure at upper right for the T-mask in comparison to the φ-mask region. In an initial series of exposures using a Censor 10X step and repeat tool, we were able to obtain 0.4-μm structures using the φ-mask nearly as good as the 0.5-μm structures projected by the Mann 4800.

These initial results demonstrate that the phase-shifting layer produces an improvement in resolution and a widening of the exposure window for the Mann 4800 exposure tool and realistic mask structures pretty much as predicted by the simple theory. Increasing the partial coherence of the illumination would

result in further improvement as would redesigning the mask to take proper account of processing windage and the limited lens resolution. With due care certain 0.7-μm structures can be patterned using this technology and an unmodified exposure tool.

Discussion and Conclusions

The phase-shifting mask has been demonstrated to improve the resolution for fine line photolithography with available 10X step and repeat exposure tools by about 40 percent. The exposure/processing window is also widened noticeably. The partial coherence of illumination in these devices is sufficient to realize a significant improvement in contrast at high resolution, but increasing the partial coherence to $\sigma = 0.3$ would yield additional improvement. When conventional light sources are employed, such an increase in coherence is accompanied by a decrease in brightness of illumination and a consequent increase in exposure time.

Improved contrast and resolution has now been demonstrated for 10X step and repeat exposure tools; similar improvements can be anticipated for 1X tools such as the Perkin-Elmer Micralign series and for proximity printing. In the latter case, the destructive interference between waves from neighboring apertures would allow the distance between wafer and mask to be increased, perhaps doubled. The fabrication of a phase-shifting mask for such applications is more of a challenge that the fabrication of a 10X mask, but making such a mask is within present electron-beam lithography technology [10].

The improved resolution achievable with phase-shifting masks should not be accompanied by a great reduction in depth of focus. Fig. 2 shows that the spatial bandwidth required to transmit an intensity pattern of given periodicity is halved when a phase-shifting mask is employed. This spatial bandwidth correlates with the numerical aperture of the lens needed to image the pattern, and the usual relation between numerical aperture and depth of focus applies. Thus the improved resolution achieved with phase-shifting masks is more useful than that obtainable by, for example, doubling the numerical aperture of the imaging system or reducing the illuminating wavelength.

In conclusion, the first experiments on phase-shifting mask structures confirm the advantages predicted by theory. Further improvement can be obtained with optimized projection tools. Considerable effort will, however, be necessary to document the properties of the phase-shifting mask and develop guidelines for mask design. Successful completion of such a project should extend the useful lifetime of familiar optical lithography technology well into the VLSI/VHSIC era.

Acknowledgment

The authors acknowledge stimulating and useful conversations with many colleagues, especially C. Cortellino, J. Zingerman, J. Wilczynski, D. Goodman, and G. Willson.

References

[1] M. Born and E. Wolf, *Principles of Optics*, 5th ed. Oxford, England: Pergamon Press, 1975, p. 370 ff.
[2] B. J. Lin, "Optical methods for fine line lithography," in *Fine Line Lithography*, R. Newman, Ed. Amsterdam: North Holland, 1980, pp. 107-230.
[3] J. W. Goodman, *Introduction to Fourier Optics.* San Francisco: McGraw-Hill, 1968, pp. 128-131.
[4] J. D. Cuthbert, *Solid State Technol.*, vol. 20, pp. 59-69, 1977.
[5] B. J. Lin, "Partially coherent imaging in two dimensions and the theoretical limits of projection printing in microfabrication," *IEEE Trans. Electron Devices*, vol. ED-27, pp. 931-938, 1980.
[6] C. J. Bouwkamp, *Rep. Progr. Phys.*, vol. 17, p. 39, 1954 and references therein.
[7] M. J. Born and E. Wolf, *Principles of Optics.* pp. 128-130.
[8] G.R.M. Thomas, H. L. Coleman, and M. Lanahan, "Use of 436 nm optical step-and-repeat imaging for wafer fabrication," in *SPIE Proc.*, vol. 174, *Developments in Semiconductor Microlithography IV*, pp. 15-21, 1979.
[9] D. S. Goodman, "Stationary optical projectors," Ph.D. dissertation, University of Arizona, 1979.
[10] G. R. Brewer, *Electron Beam Technology in Microelectronic Fabrication.* New York: Academic Press, 1980.