Choosing objective functions for inverse lithography patterning

Jue-Chin Yu and Peichen Yu
Department of Photonics and Institute of Electro-Optical Engineering,
National Chiao Tung University, 1001 Da-hsueh Rd., Hsinchu 30050, Taiwan
E-mail address: yup@faculty.nctu.edu.tw

ABSTRACT

Inverse lithography which generates model-based patterns theoretically has superior patterning fidelity comparing to conventional rule-based technique. Cost functions are the determinant of performance inverse lithography that is also an optimization problem. However, the design and know-how of cost functions have rarely been discussed. In this paper, we investigate the impacts of various cost functions and their superposition for inverse lithography patterning exploiting a steepest descent algorithm. We research the most generally used objective functions, which are the resist and aerial images, and also deliver a derivation for the aerial image contrast. We then discuss the pattern fidelity and final mask characteristics for simple layouts with a single isolated contact and two nested contacts. Moreover, the convergences which are expressed by edge-placement error (EPE) and contrast versus iteration numbers rapidly attain to steady sate in most hybrid cost functions. All in all, we conclude that a cost function composed of a dominant resist-image component and a minor aerial-image or image-contrast component can carry out a good mask correction and contour targets when using inverse lithography patterning.

Keywords: Microlithography, Inverse lithography, Model-based pattering, Gradient descent, Computational imaging.

1. INTRODUCTION

As Moore's law marches on and semiconductor manufacturers make the push towards the 22nm technology node and beyond, the challenges faced by microlithography are ever increasing. The need to print such small features using 193nm illumination -- which is way beyond the Rayleigh diffraction limit -- has made resolution enhancement techniques (RETs) mandatory. Optical proximity correction (OPC) is one of the RETs that modify design layouts in order to minimize their distortion when printed on silicon. Segment-based OPC, in particular, is the most-widely-used RET and has been a standard industry practice since the 90nm node. Because segment-based OPC only modifies edges already present in the design layout, it is relatively easy to implement, particularly in iterative algorithms. However, as the need for stronger OPC increases, this simplified process has encountered major limitations. As the reachable solution space lies strictly in the vicinity of the original layout, segment-based OPC is often not expressive enough and does not exploit the full range of possible mask configurations to get the best possible pattern fidelity and image contrast. To overcome this limitation, other approaches like inverse lithography (IL) have been proposed [1, 2, 3]. Inverse lithography is a technique that defines the search of mask corrections as an optimization problem, where the goodness criteria are solely determined by one or various objective functions. Many IL methods have been introduced over the past three decades [1-10], including pixel flipping [4, 5, 6] and other more refined gradient descent approaches [7, 8, 9, 10]. In general, IL techniques can handle more relaxed constraints and are often implemented as full-mask approaches, which enable higher pattern fidelity even if at a higher computational cost. In addition, the clever use of IL may enable a given technology node with traditional mask processes, averting the need of more expensive RET alternatives, such as alternating phase-shift masks or double patterning. Evidently, IL-based OPC offers clear advantages in terms of efficiency and manufacturing cost, and hence it is currently poised as the next-generation OPC for future lithography nodes.

As a general rule, the success of any particular IL algorithm depends strongly on the nature and form of its cost function, which is often constituted by various objective functions. A properly designed cost function paired with an appropriate algorithm can help avoid issues such as getting trapped in local minima, slow convergence, mask discontinuities, etc. In the literature, the most commonly used objective functions have traditionally been the aerial and resist images. The former compares the optical intensity distribution on the photoresist (after exposure but before developing) to a desired target intensity, while the latter compares the developed photoresist profile of a given mask to

Optical Microlithography XXIV, edited by Mircea V. Dusa, Proc. of SPIE Vol. 7973,
79731N · © 2011 SPIE · CCC code: 0277-786X/11/$18 · doi: 10.1117/12.879440

the desired outcome. Other cost functions, such as edge contours, aerial image contrast, and mask error enhancement factor (MEEF), have also been proposed and can be customized to satisfy strict technological specifications. In our experience, we have observed that given the same cost function and process model, two different optimization algorithms, like a wavefront-based pixel flipping [11], and gradient descent, tend to converge to nearly identical final mask patterns. This phenomenon further highlights the importance of using a favorable cost function. Curiously, even though cost functions play such an important role in IL patterning, their impacts and characteristics have rarely been reported explicitly. Therefore, it is beneficial to make a systematic analysis of their influence using a generic IL algorithm. In the present paper, we first employ the most commonly used objective functions, which are 1) the aerial and 2) resist images [7,8,9], and then derive an objective function for 3) the aerial image contrast. We next perform a side-by-side comparison of the three cost function components and examine their properties, including pattern fidelity and final mask characteristics using a gradient descent algorithm. It is important to note that while many IL algorithms presented in the past have made assumptions such as completely coherent or incoherent illumination sources in order to simplify the calculation, the approach presented in this work makes no such assumptions on the illumination source, and uses a partially coherent illumination model [12,13,14]. Therefore, the present study can draw more representative conclusions to contrast current industry practices than other simplified approaches.

2. METHODOLOGY

In this section, we describe the formulation of the studied objective functions and the computational flow of the gradient descent method. Three functions are considered in this work: 1) the aerial image, 2) the resist image, and 3) the aerial image contrast. The aerial image represents the optical intensity distribution formed by the projection system on the coated wafer. The resist image corresponds to the resist profile after removing the exposed resist. Finally, the aerial image contrast is highly related to the depth of focus (DOF), and hence determines the process window.

Optical microlithography simulations are made possible by the Köhler's illumination model [15, 16, 17]. Currently, the industry standard illumination source still employs an ArF excimer laser with a 193 nm wavelength. The quasi-monochromatism of the ArF excimer laser results in partially coherent images on the wafer, which are also referred as the aerial images. The optical intensity of an aerial image can be formulated by a Sum of Coherent Systems (SOCS) model [17, 18, 19], where the illumination system including the source and projectors is decomposed into eight kernels using Singular Value Decomposition (SVD) [20, 21]. Therefore, the electric fields are constructed by a convolution of different coherent kernels as indicated in Eq. (1). The total intensity at a given position (x, y) on the image plane is then the superposition of those kernels' intensities, which can be expressed as Eq. (2).

$$E_q(x,y) = \phi_q(x,y) \otimes \boldsymbol{o}(x,y), \tag{1}$$

$$I(x,y) = \sum_{q=1}^{n} \lambda_q \left| E_q(x,y) \right|^2, \tag{2}$$

where E and I represent the electrical field and image intensity respectively. ϕ_q is the q^{th} optical kernel, $\otimes$ denotes the convolution calculation, and $\boldsymbol{o}$ the mask function. λ_q is the eigenvalue of the q^{th} kernel with n kernels in total.

The aerial image represents the distribution of optical intensity on the wafer, which corresponds to the exposure condition of a photoresist. Some photoresist models employ a Constant Threshold Resist (CTR) [22], where the developed resist profile can then be described by a sigmoid transformation of the aerial image, [8, 9] that is,

$$T(I) = \frac{1}{1 + e^{-a(I - t_r)}}. \tag{3}$$

In Eq. (3), the parameter a describes the sensitivity of the photoresist reacting with the light, which depicts the slope of sidewall profiles, and t_r is the parameter of the constant threshold level. Here, the value is set and normalized to 0.5.

In optical microlithography, the aerial image contrast plays an important role in the development of photoresist, which is directly related to the number of molecular chemical-bonds that are broken or still linked after exposing. As the image intensity difference between the exposed and unexposed regions is desired to be as large as possible, increasing the image contrast can mitigate patterning damages when removing the resist. We derive and express the image contrast using a linear differential operator, defined as the following:

$$\nabla_T I = \frac{\partial I}{\partial x}\hat{x} + \frac{\partial I}{\partial y}\hat{y}, \tag{4}$$

where ∇_T is the differential operator to evaluate the aerial image contrast. Due to the discrete nature of the mask, the differential operator $\partial/\partial x$ and $\partial/\partial y$ is approximated by numerical differences. Therefore the operators can be represented in a matrix form and Eq. (4) re-written into Eq. (5):

$$\nabla_T I = (DI)\hat{x} + (ID^T)\hat{y}, \tag{5}$$

where $I \in R^{N\times N}$. $D \in R^{N\times N}$ is explicitly shown in Eq. (6):

$$D = \begin{bmatrix} 1 & -1 & & & & \\ & \ddots & \ddots & & \mathrm{O} & \\ & & 1 & -1 & & \\ & & & \ddots & \ddots & \\ & \mathrm{O} & & & 1 & -1 \\ & & & & -1 & 1 \end{bmatrix}_{N\times N}, \tag{6}$$

In Eq. (5), DI performs the operation that calculates the difference in the row direction, and ID^{T} computes the column difference.

In order to apply these objective functions in an optimization algorithm, we also need to define the target aerial image based on the input patterns. The target aerial image is parameterized as

$$I_t(r) = \begin{cases} \dfrac{\alpha}{1+e^{b(r-r_0)}}, & \textit{Geometric Center} \le r \le \textit{Boundary} \\ e^{-c(r-r_0)}, & r \ge \textit{Boundary} \end{cases}, \tag{7}$$

where I_t(r) is the target image as a function of the pixel index r. r is arranged in the radial direction from the geometric center of a drawn pattern, where r_0 is the index of the boundary pixels. The generation of the pixel index is adopted from a wavefront expansion technique previously developed by us [11]. For any particular layout, we construct a wavefront from the feature's edges both inwards and outwards. This way, we can construct a complete wavefront from a small region (or even a single point) inside each feature all the way to the edge of the image. Each wavefront, starting from the innermost, receives an index, and all the pixels on the layout are processed in order according to this index. A few examples of wavefront templates can be found in Ref. [11]. As a result, the indices of boundary pixels end up being the same for individual features. It is also worth noting that this indexing technique will clearly identify pixels that are equidistant from two or more features. Moreover, in Eq. (7), α, b and c are constants to adjust for the target aerial image profile. Here, we choose a target image that is unity within the drawn and zero outside, corresponding to α=1, b=-∞ and c=∞. Since the designed target is normalized to unity at the maximal, the optical model used for simulations must be calibrated with experimental values. Once the target aerial image is defined, the target resist image is obtained by the corresponding sigmoid transformation and the target image contrast is the differential operation of the target aerial

image. As a result, the three objective functions that evaluate the differences between the desired and the calculated profiles can be expressed as in Eq. (8), Eq. (9) and Eq. (10), respectively:

$$F_I = \left\| I_t - I \right\|^2, \tag{8}$$

$$F_R = \left\| T(I_t) - T(I) \right\|^2, \tag{9}$$

$$F_C = \left\| \frac{\partial}{\partial x}(I_t - I) \right\|^2 + \left\| \frac{\partial}{\partial y}(I_t - I) \right\|^2, \tag{10}$$

Here, F_I, F_R and F_C represent the costs for the aerial image, resist image, and aerial image contrast, respectively. The norm $\left\| \cdot \right\|^2$ denotes the square of Euclidean distance which denotes the inner product of the same vector. Moreover, we can further combine the components into a total cost function, F by assigning three weighting coefficients, γ_I, γ_R, and γ_C to the aerial image, resist image, and aerial image contrast, respectively. The final expression is then as follows:

$$F = \gamma_I F_I + \gamma_R F_R + \gamma_C F_C. \tag{11}$$

It is important to note that the numeric cost values of these three objective functions, i.e. the variations between their minimal and maximal values, are not always in the same range. The individual costs not only depend on the drawn layouts, but also the nature of different objective functions. For example, the cost of the aerial-image component is often larger than that of the resist-image for simple layouts, since the cost of the former is contributed by all of the pixels on mask, while that of the latter mainly arises from the contour pixels. The cost variation for the resist-image component is therefore very limited compared to that for the aerial image. The normalization approach is essential to remove such a dependency on different objective functions. Therefore, for each case studied, we normalize the costs of the individual functions to the same interval of [0,10] by performing a linear transformation for costs at the initial (maximal) and at the final (minimal) iterations. The normalization process is not absolutely stringent, but necessary to compare the characteristics of different objective functions on the same basis for different drawn layouts. Therefore F_I, F_R, and F_C in Eq. (11), represent normalized objective functions.

The goal of an inverse lithography algorithm is to find the optimal mask $\hat{\boldsymbol{o}}$ that minimizes a given constraint and therefore can be expressed as in Eq. (12):

$$\hat{\boldsymbol{o}} = \underset{o(x,y)}{\boldsymbol{argmin}}\, F(\boldsymbol{o}(x,y)), \tag{12}$$

where $\boldsymbol{o}$ is a mask function. A gradient-search method used to calculate $\hat{\boldsymbol{o}}$ is explained next. Since the gradient operation calculates the derivatives of the mask, the discrete binary mask described by $\boldsymbol{o}$, must be first parameterized by a continuous variable $\boldsymbol{\theta}$ in order to obtain an analyzable form. Here, a sinusoidal transformation is employed to convert a binary drawn mask in to a continuous grey-level mask [8]:

$$\boldsymbol{o} = \frac{1 + \cos(\boldsymbol{\theta})}{2}. \tag{13}$$

We note that $\boldsymbol{\theta}, \boldsymbol{o} \in R^{N \times N}$, $0 \le \boldsymbol{\theta} \le \pi$, $0 \le \boldsymbol{o} \le 1$. The converted mask then allows a continuous optical transmission value between zero and unity with $\boldsymbol{\theta}$ varying between zero and $\boldsymbol{\pi}$. Moreover, Eq. (13) can also be extended to a phase-shift mask (PSM) intuitively by multiplying a complex phase term [9].

The cost function gradient ∇F can be derived as shown in Eq. (14)

$$\nabla F = \gamma_I \nabla F_I + \gamma_R \nabla F_R + \gamma_C \nabla F_C. \tag{14}$$

The explicit expressions of the objective functions for the aerial image [7], resist image [8,9] and image contrast are listed in Eq. (15), Eq. (16) and Eq.(17), respectively:

$$\nabla F_I(\boldsymbol{\theta}) = \Big(\sum_{q=1}^{Q} \lambda_q \big[-2\big(I_t - I\big) \odot 2E_q \big] \otimes \phi_q^{flip} \Big) \odot \left(-\frac{\sin(\boldsymbol{\theta})}{2} \right), \tag{15}$$

$$\nabla F_R(\boldsymbol{\theta}) = \Big(\sum_{q}^{Q} \lambda_q \big[-2\big(T(I_t) - T(I)\big) \odot \big(a\big(1 - T(I)\big) \odot T(I)\big) \odot 2E_q \big] \otimes \phi_q^{flip} \Big) \odot \left(-\frac{\sin(\boldsymbol{\theta})}{2} \right), \tag{16}$$

$$\nabla F_C(\boldsymbol{\theta}) = \Big(\sum_{q=1}^{Q} \lambda_q \big[\big(-2D^T\big(D\big(I_t - I\big)\big) - 2\big(\big(I_t - I\big)D^T\big)D\big) \odot 2E_q \big] \otimes \phi_q^{flip} \Big) \odot \left(-\frac{\sin(\boldsymbol{\theta})}{2} \right), \tag{17}$$

where $\odot$ is the element-by-element multiplication operator and $\otimes$ is the convolution operator. Furthermore, ϕ_q^{flip} is the up-down and left-right flip of ϕ_q, i.e. $\phi_q^{flip}(i,j) = \phi_q(N - i + 1, N - j + 1)$ where i, j are integers and $\in [1, N]$.

Finally, we employ a steepest-descent approach [8, 23] to find a solution to the inverse problem. The step length is chosen to be 2 for the tradeoff between speed and convergence. The diagram in Fig. 1 shows the general procedure of the iterative calculation.

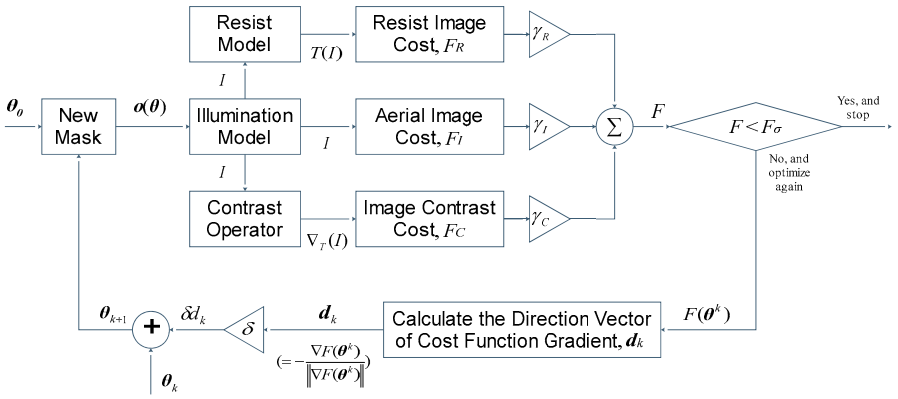

Fig. 1. Steepest descent algorithm. Three parameters γ_R, γ_I and γ_C are used to bias the relative cost of aerial image, resist image and image contrast.

The iterative algorithm can be summarized by the following pseudo-code:

. **Initialization :**

Assign an initial guess $\boldsymbol{\theta}^0$ and calculate the $\boldsymbol{d}_0 = -\dfrac{\nabla F(\boldsymbol{\theta}^0)}{\left\|\nabla F(\boldsymbol{\theta}^0)\right\|}$.

Initialize a constant step length δ.

. **Iterative Step:**

$$\boldsymbol{d}_k = -\frac{\nabla F(\boldsymbol{\theta}^k)}{\left\|\nabla F(\boldsymbol{\theta}^k)\right\|}.$$

$$\boldsymbol{\theta}^{k+1} = \boldsymbol{\theta}^k + \delta \boldsymbol{d}_k.$$

. **Stop condition:**

If the stop condition, $F < F_\sigma$ is satisfied, the algorithm terminates.

3. RESULTS AND DISCUSSION

In this section, we employ the gradient descent algorithm described previously to analyze two examples: a 190nm×190nm isolated contact and two horizontally-aligned contacts each with an area of 100nm×100nm and a spacing of 100nm. The test patterns are chosen such that the single contact is above the *Rayleigh* criterion: $R = 0.61\lambda/\mathrm{NA} \sim 170$nm, while for the pair, both the spacing and dimensions are below the limit. Both layouts are represented in 256×256 pixel tiles with a pixel dimension of 10 nm. An annular illumination source is employed with $\sigma_{in} = 0.4$ and $\sigma_{out} = 0.7$. The partially coherent illumination model contains eight kernels, where $\lambda = 193$nm, $NA = 0.7$. Moreover, the constant threshold for the aerial image intensity is set and normalized to 0.5. Therefore the threshold parameter t_r, required for the resist image transformation in Eq. (3) is set to 0.5, while a is chosen to be 90 to represent a conventional resist profile. We discuss the impact that the previously defined cost function components in various superposition configurations have on the resulting characteristics of the corrected grey-level masks, the contours and the aerial image intensities.

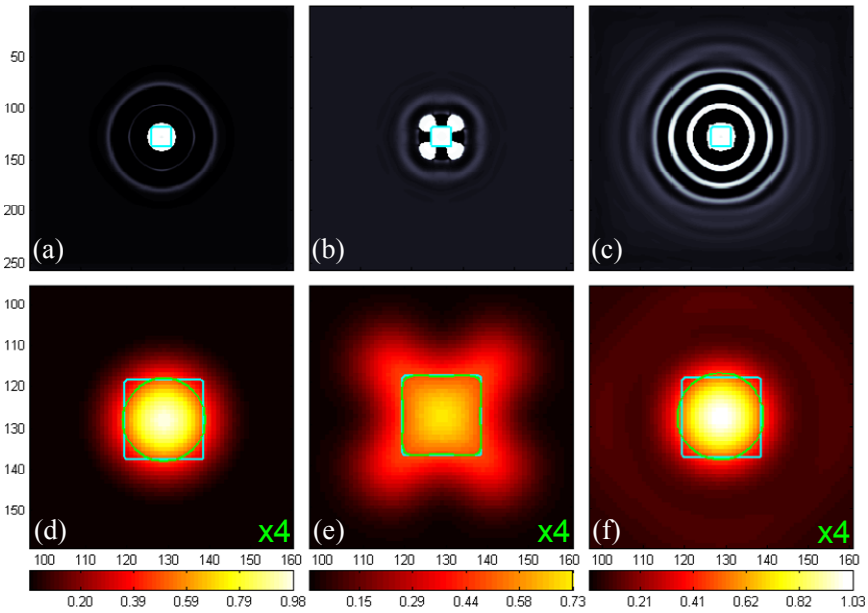

Fig. 2. The inverse results of a large isolated contact evaluated by every cost function component individually where (a) (γ_I, γ_R, γ_C)=(1, 0, 0), (b) (γ_I, γ_R, γ_C)=(0, 1, 0) and (c) (γ_I, γ_R, γ_C)=(0, 0, 1). The corresponding contours and aerial images of (a)-(c) are shown in (d)-(f) where the cyan and green contours respectively label the drawn pattern edges and aerial image threshold contours.

Figure 2(a)-(c) shows the generated gray-level masks using cost functions with a single component, that is, coefficients, (γ_I, γ_R, γ_C)=(1, 0, 0), (0, 1, 0), and (0, 0, 1), respectively. The corresponding aerial images and contours of the single large contact are shown in Fig. 2(d)-2(f). As shown in Fig. 2(a), the cost function based on only the aerial-image component results in mask features concentrated on the drawn pattern, while in Fig. 2(b) the correction shows very large serifs, and in Fig. 2(c) concentric rings. The mask features very much reveal the nature of each objective function as discussed below.

First, figure 2(a) shows that the resulting mask features only accentuate the drawn pattern with weak grey features around it. Such a correction is very similar to that obtained by segment-based OPC, where the edges of the drawn patterns are dissected and then repositioned inwards or outwards to minimize the EPEs. The similarity arises from the fact that the cost of the aerial-image is determined by its difference from the target aerial image in Eq. (8). Any mask features outside the drawn pattern increase the unwanted intensity distribution and hence the cost. Expectedly, the mask features are concentrated on the drawn and are very desirable for an eventual transformation to a binary mask. The corresponding contour shows negligible EPEs on all the edges, but with large rounded corners, as shown in Fig. 2(d). This behavior, with both the mask and optical intensities concentrated on the drawn has worked well for previous CMOS

generation nodes (45nm and above), however, it may become a hindrance for small features as shown in the next example.

Second, as shown in Fig. 2(b), the mask features show very large serifs. Since the sigmoid transformation can convert a number of different aerial images into resist images with sharp corners, as expressed in Eq. (3), the cost of the resist-image cost function mainly arises from the difference on the edges defined in Eq. (9). Therefore the nature of this objective function tends to drive the mask correction towards the threshold contour, in this case, a square. It is known that large serifs are signature structures to obtain contours with sharp corners. The aggressive corner correction is also confirmed by the green contour shown in Fig. 2(e). However, since there is a tradeoff between sharp corner contours and aerial image contrast, the overall optical intensity is weak. Moreover, because of the low intensity, this cost function may have limited performance for small features as well. From a mask characteristics point of view, the gray-level features surrounding the main pattern are not very significant, implying minimal EPEs arising from filtering out such features. Therefore, the mask resulting from the resist image is also adequate for manufacturing.

Finally, as shown in Fig. 2(c), the mask features arising from the aerial-image-contrast cost function appear as rings with a high grey-level contrast, which are very similar to a Fresnel zone plate in optics. The Fresnel zone plate functions as a lens and is used for focusing and imaging in optical systems. In other words, in order to obtain a high aerial-image contrast as dictated by the cost function, the resulting mask features converge to a pattern that can focus light. Due to the partial coherence of the illumination source, the resulting mask pattern does not have an analytical expression as a Fresnel zone plate which is often designed for a coherent source with a single wavelength [21]. However, based on the mask shown in Fig. 2(c), the dominant kernel that represents one of the coherent sources in the illumination model dictates the mask characteristics, making the zone-plate type correction still evident. Moreover, due to the distributed mask features, the spreading of light distribution on the resist, as shown in the red color of Fig. 2(f), is also higher than that of the other two cost functions. However, as long as the light spreading does not trigger the exposure of photoresist, i.e. remains below the threshold, the features have no side effects on patterning. Still, such mask features are not favorable in mask manufacturing, relatively speaking. Furthermore, the contour shows relatively large EPEs, which means that convergence is difficult. Overall, mask making is relatively feasible for the aerial-image and the resist-image cost functions, but not so much for the image-contrast. Finally, it is worth noting that the aerial image intensity on wafer is largest with the aerial-image-contrast, then the aerial-image, and lastly, the resist-image.

Fig. 3 shows the configurations of evaluation points for single contacts where the cut-line length is set to be 10 pixels. If the EPEs are over 10 pixels or without any developed contours, the EPEs are automatically saturated to 10 pixels. The contrasts are calculated based on the evaluation points. As shown in Fig. 4, the averaged EPE and contrast of cut-lines 2, 5, 8 and 11 are plotted as a function of the iteration. The curve of resist image only optimization has weak ripples after about 25 iterations. Such interesting phenomenon results from the sharp transform of aerial image by using eq 3. where the coefficient a is set to 90. To get rid of the vibration, a relatively small a can be applied. However the small a will deviate the simulated resist behavior.

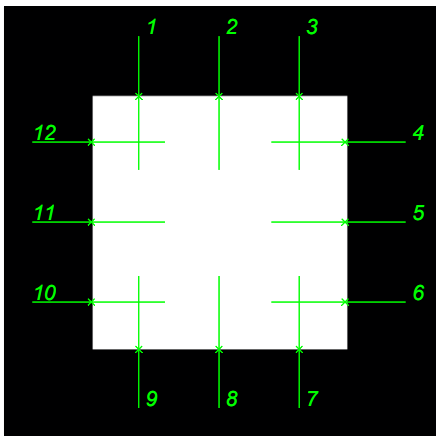

Fig. 3. The evaluation point configurations of single contact with 190 nm×190 nm feature size.

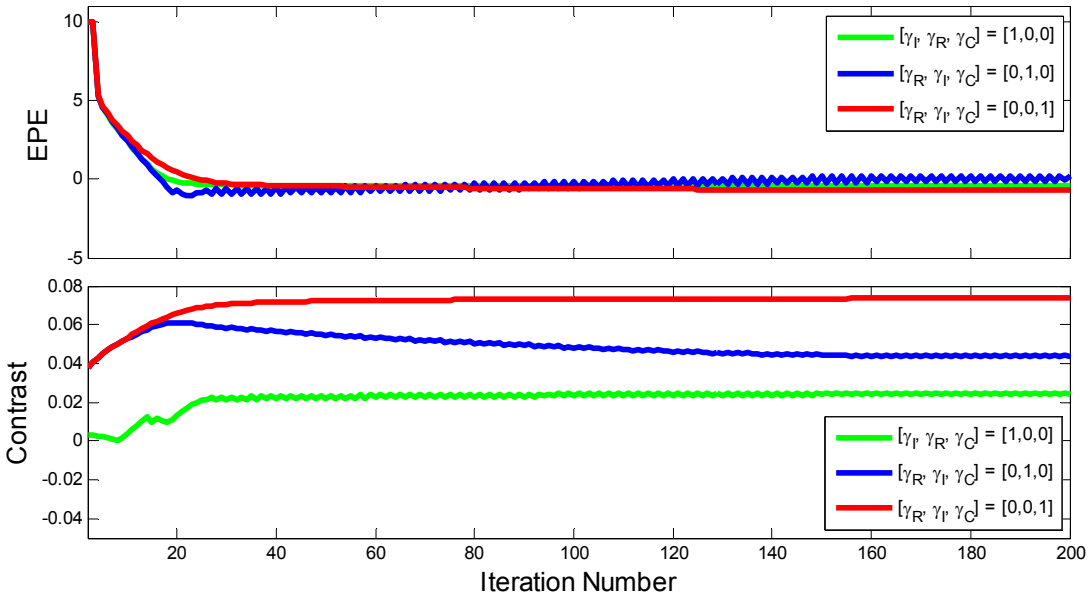

Fig. 4. The averaged EPE (Pixel) and contrast (AU) of cut-line 2, 5, 8 and 11 are plotted as a function of the iteration. The green, blue and red curves show that (γ_I, γ_R, γ_C) is under (1, 0, 0), (0, 1, 0) and (0, 0, 1) respectively.

The superposition of various objective functions is described next, as shown in Fig. 5(a)-(d) for (γ_I, γ_R, γ_C)= (1, 1, 0), (1, 0, 1), (0, 1, 1), and (1, 1, 1), respectively. When dealing with combined objective functions, we find that while the resulting mask features have the footprint from both components, one component always appears dominant to the others. So, in Fig. 5(a), 5(b), and 5(d) the aerial image dominates in the cost function, while in Fig. 5(c), the aerial image contrast dominates the resist image. The resulting contours and aerial image intensities are thus also dictated by the

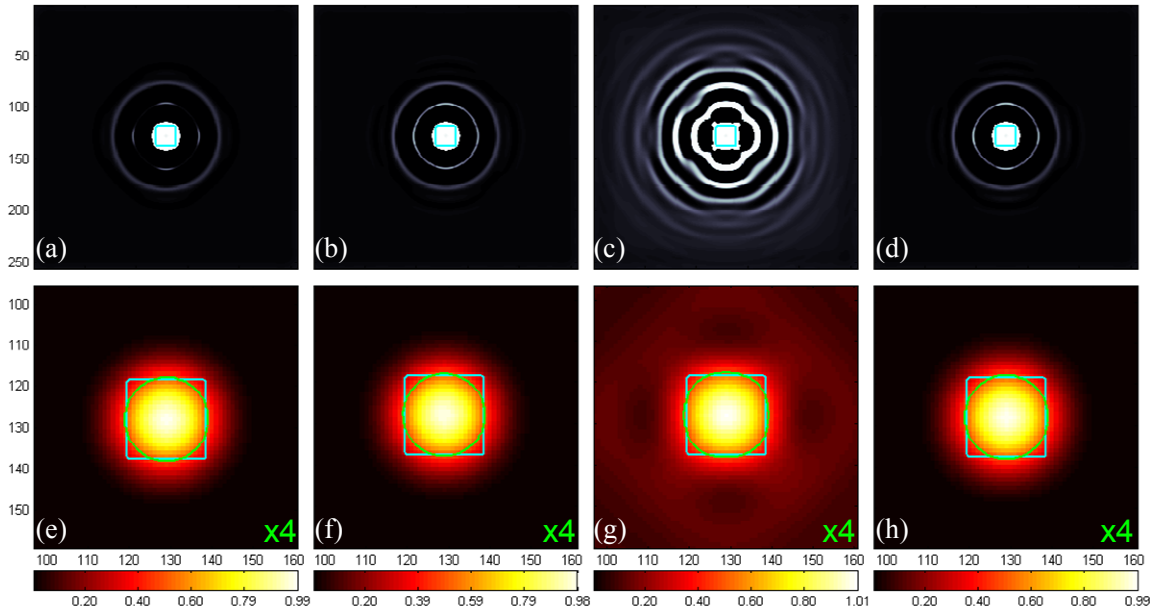

Fig. 5. The inverse results of a large isolated contact as with different combination of cost functions components where (a) (γ_I, γ_R, γ_C)=(1, 1, 0), (b) (γ_I, γ_R, γ_C)=(1, 0, 1), (c) (γ_I, γ_R, γ_C)=(0, 1, 1) and (d) (γ_I, γ_R, γ_C)=(1, 1, 1). The corresponding contours and aerial images of (a)-(d) are shown in (e)-(h) where the cyan and green contours respectively label the drawn pattern edges and aerial image threshold contours.

dominant objective function, as shown in Fig. 5(e)-5(h). However, since all the linear combinations of the cost function components can result in the exposure of photoresist for the large isolated contact, it is not conclusive as to what combination is best for inverse lithography patterning. In Fig. 6, the averaged EPE and contrast of cut-lines 2, 5, 8 and 11 are plotted as a function of the iteration. No ripples occur as the resist image cost function is applied. The reasonable explanation is that the other two cost function will alleviate the vibrations from resist image conversion function and reduce the ripples.

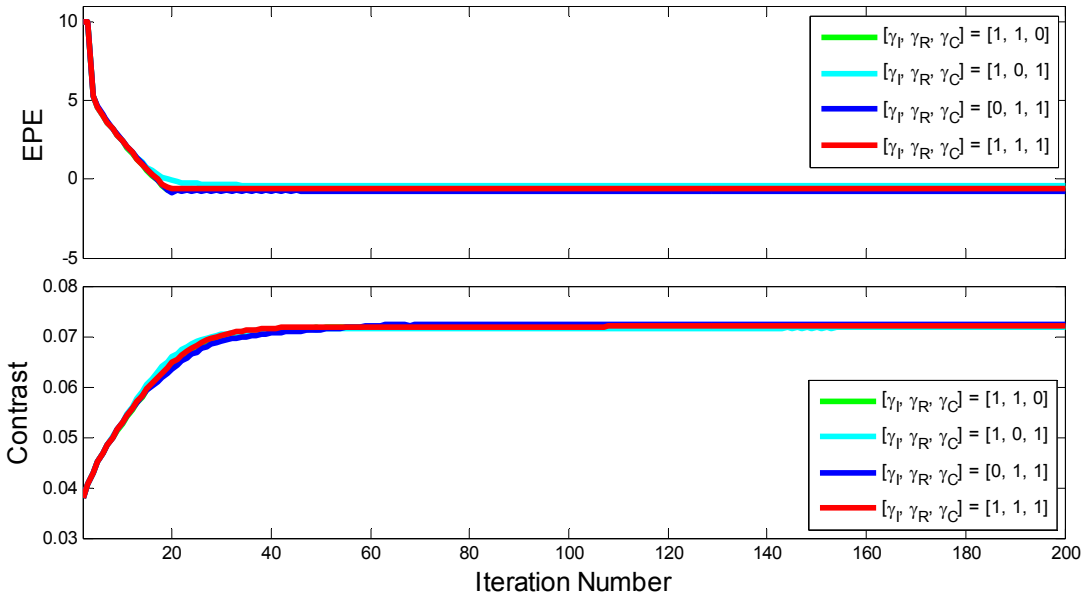

Fig . 6. The averaged EPE (Pixel) and contrast (AU) of cut-line 2, 5, 8 and 11 are plotted as a function of the iteration. The green, cyan, blue and red curves show that (γ_I, γ_R, γ_C) is under (1, 1, 0), (1, 0, 1), (0, 1, 1) and (1, 1, 1) respectively.

In followings, the same three individual cost function components are applied to the correction of two nested square contacts, where both the edge length and spacing are equal to 100 nm. This input pattern is challenging to the illumination model used in this work, and that reflects on the limitations of the different cost functions. As shown in Fig. 7(a), the corrected mask using the aerial-image component maintains similar mask characteristics as discussed previously, which only corrects for the main patterns. However, the corresponding aerial image shown in Fig. 7(d) clearly show that the cost function has hit a limitation for small features, as the photoresist is underexposed. The similarity of the corrected mask to a configuration achievable by segment-based OPC further highlights that such an approach is no longer sufficient for critical CMOS nodes.

Next, as shown in Fig. 7(b), the correction using the resist image component was trapped in a local minimum during optimization, and only the left contact was successfully exposed. Since the resist cost function allows a number of aerial images to be transformed into a similar resist image, as evidenced by Eq. (3), its solution space is the largest among these three cases. Therefore, the advantage of the resist image cost function is a more complete exploration of possible mask corrections, while the disadvantage is the higher probability of getting trapped in local minima. Still, this disadvantage may be mitigated with the assistance of other objective functions, as will be discussed later. Furthermore, the exposed contact has four large mask features at the corners, known as sub-resolution assist features (SRAFs). These features help the exposure of the drawn patterns by bring up the intensity level, but they themselves do not print. The corner SRAFs can be thought of as being evolved from the large serifs of a big contact, as previously shown in Fig. 2(b). Therefore, particular footprint of the resist image cost function is still present. While not complete, the contour of the exposed contact shown in Fig. 7(e) is also very round and symmetric in shape, distinguishing itself from those corrected by the aerial-image or image-contrast to be shown afterward.

Expectedly, the corrected mask in Fig. 7(c) using the image contrast results in zone-plate type patterns. The optical intensity from the image-contrast function shown in Fig. 7(f) over exposes the photoresist, in contrast to an insufficient exposure shown in Fig. 7(d). We found that the convergence of small mask features can rarely be obtained by single-component cost functions. The contours in Fig. 7 reflect the limitations of the individual objective functions. However, Fig. 7 also provides evidence that SRAFs are mandatory for pushing forward the resolution limit of this illumination model.

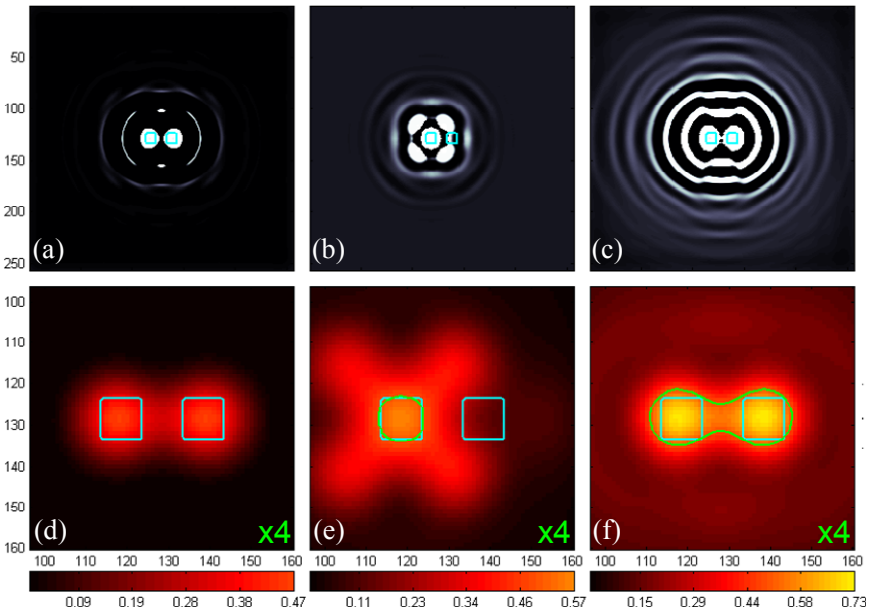

Fig. 7. The inverse results of two close contacts evaluated by every cost function component individually where (a) (γ_I, γ_R, γ_C)=(1, 0, 0), (b) (γ_I, γ_R, γ_C)=(0, 1, 0) and (c) (γ_I, γ_R, γ_C)=(0, 0, 1). The corresponding contours and aerial images of (a)-(c) are shown in (d)-(f) where the cyan and green contours respectively label the drawn pattern edges and aerial image threshold contours.

Fig. 8 shows the configurations of evaluation points for two closely-placed contacts where the cut-line length is set to be 10 pixels. As the same settings of Fig. 3, the EPEs are automatically saturated to 10 pixels as the EPEs are over 10 pixels or without any developed contours. Moreover the contrasts are also calculated based on the evaluation points.

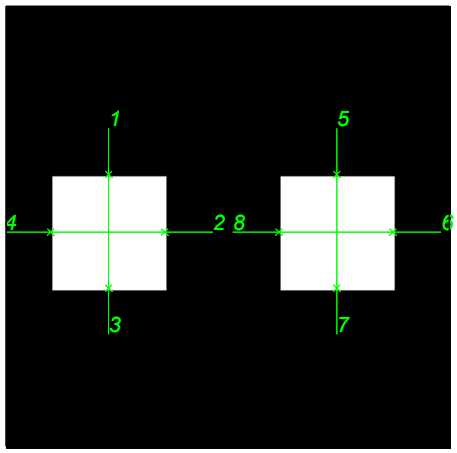

Fig. 8. The evaluation point configurations of two contacts with 100 nm×100 nm feature sizes and a 100-nm separation in a 256×256 pixel template.

Fig. 9. shows the convergence of EPE and contrast. Due to no developed contour in Fig. 7. (d), the green curve still holds in 10. The ripples also occur as only using resist image cost function. Furthermore the EPE records of (0, 1, 0) condition approximately stop in 5 because there is only one developed via in Fig. 7. (e).

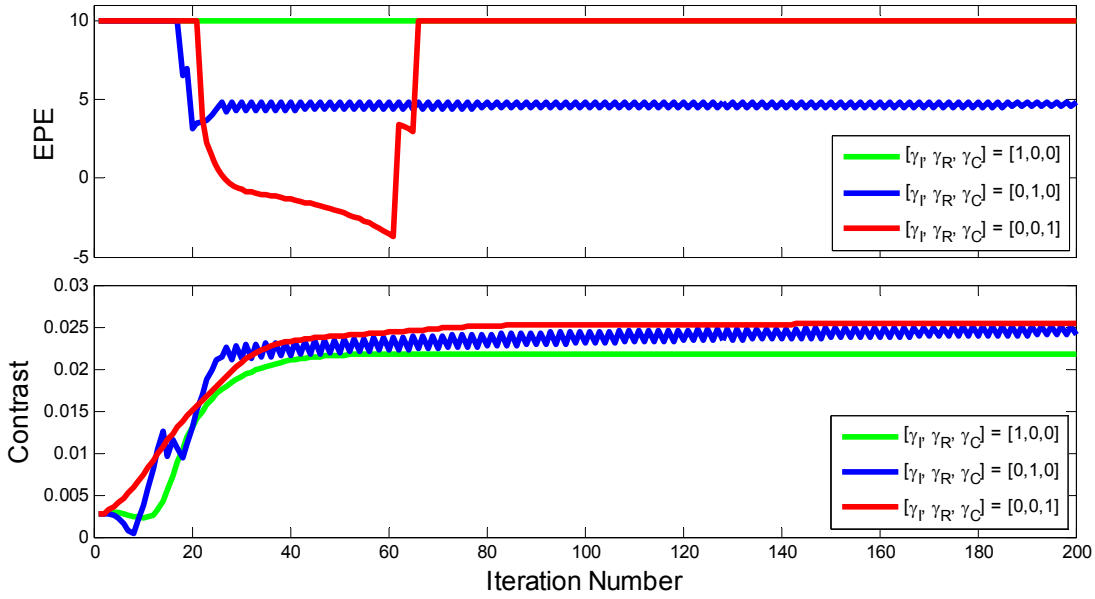

Fig. 9. The averaged EPE (Pixel) and contrast (AU) of cut-line 2 and 8 are plotted as a function of the iteration. The green, blue and red curves show that (γ_I, γ_R, γ_C) is under (1, 0, 0), (0, 1, 0) and (0, 0, 1) respectively.

Subsequently, we demonstrate that with a mix of cost function components, the contours can improve significantly, as shown in Fig. 10(a), 10(c), and 10(d). By combining different components, the solution space of good mask corrections is potentially increased while the risk of getting trapped in local minima is considerably alleviated. Therefore, it is relatively easy for the gradient descent algorithm to converge to a solution. However, we note that the contacts in Fig. 10(f) are still not exposed. As mentioned previously, the actual costs of three components are largest for the aerial image, then the aerial image contrast, and that of the resist image being the smallest. Therefore the mask features in Fig. 10(a)-10(d) are dominated by the larger cost component, just as in the single isolated contact. For example, due to the aerial image cost dominanting in Fig. 10(b), the influence of the aerial image contrast is evidently not large enough for bringing in enough light to raise the intensity above threshold. The superposition of two components at a time is further studied to obtain a better exposure. We also notice that the contours of the two square contacts are relatively pointy towards the nested space due to low image contrast. This phenomenon may result in bridging issues in contact layers, which have been observed in previous technologies. This characteristic is mostly dictated by the aerial-image objective function, which we show can be significantly relived when using the resist-image as the dominant component.

As shown in Fig. 11(a), the weighting coefficient of the image-contrast component can be gradually increased until contours appears. However, as seen in Fig. 11(a), after many iterations, the contours are still slightly deformed and the EPEs relatively large. In general, the superposition of the aerial image and image contrast results in unstable convergence for small features, as both of them have the same effect of increasing the intensity on input patterns, and therefore, a small change in the surrounding features could result in large intensity variations. Finally, since the influence of the resist image is the weakest among the three, in order to achieve a solution with the pattern characteristics of the resist image component, the coefficients of the other two have to be reduced and adjusted to achieve optimized EPEs. The masks obtained with a minor addition of aerial image and image contrast components are shown in Fig. 11(b) and 11(c), respectively. As seen in Fig. 11(b), the mask pattern contains two SRAFs near the corners of the central space, while in Fig. 11(c), there are six corner SRAFs. In both cases, the contours are very round in shape, unlike those obtained by the dominant aerial-image component by itself. The mask features are relatively adequate since the SRAFs are large in size, compared to those obtained by just the aerial-image component. Moreover, the mask features in Fig. 11(b) and 11(c) also coincide with general guidelines for SRAF placement in small contacts, which are usually determined by

either rule-based or design of experiments (DOE) techniques. Here, we show that the best mask patterns and contours are obtained by using the resist image component with the assistance of either the aerial image or the aerial image contrast to avoid convergence issues.

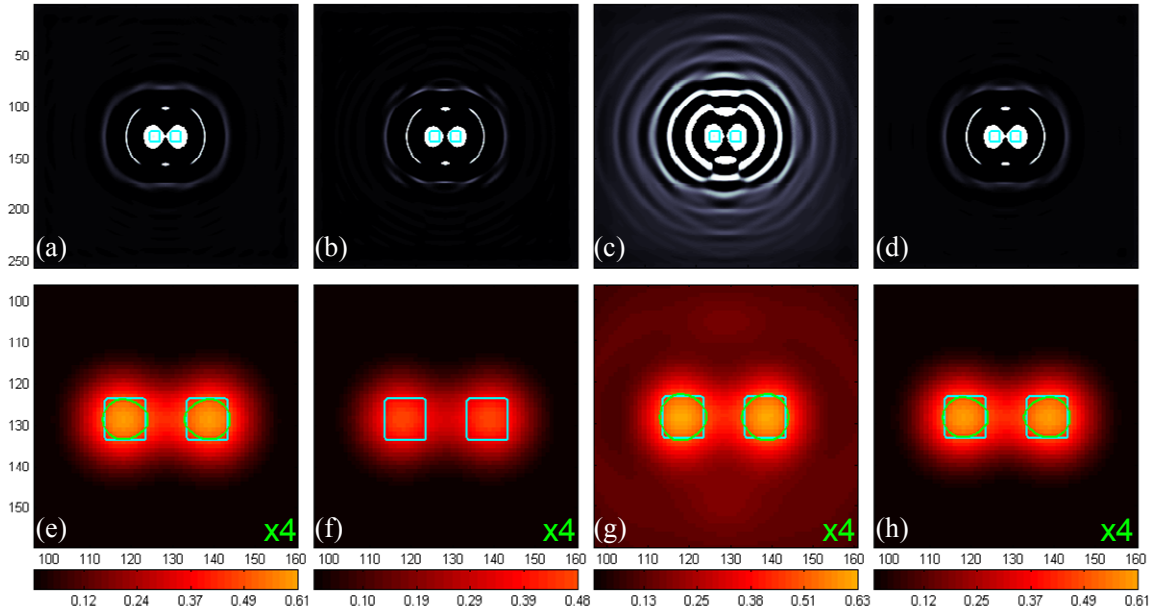

Fig. 10. The inverse results of two close contacts with different combination of cost functions components where (a) (γ_I, γ_R, γ_C)=(1, 1, 0), (b) (γ_I, γ_R, γ_C)=(1, 0, 1), (c) (γ_I, γ_R, γ_C)=(0, 1, 1) and (d) (γ_R, γ_I, γ_C)=(1, 1, 1). The corresponding contours and aerial images of (a)-(d) are shown in (e)-(h) where the cyan and green contours respectively label the drawn pattern edges and aerial image threshold contours.

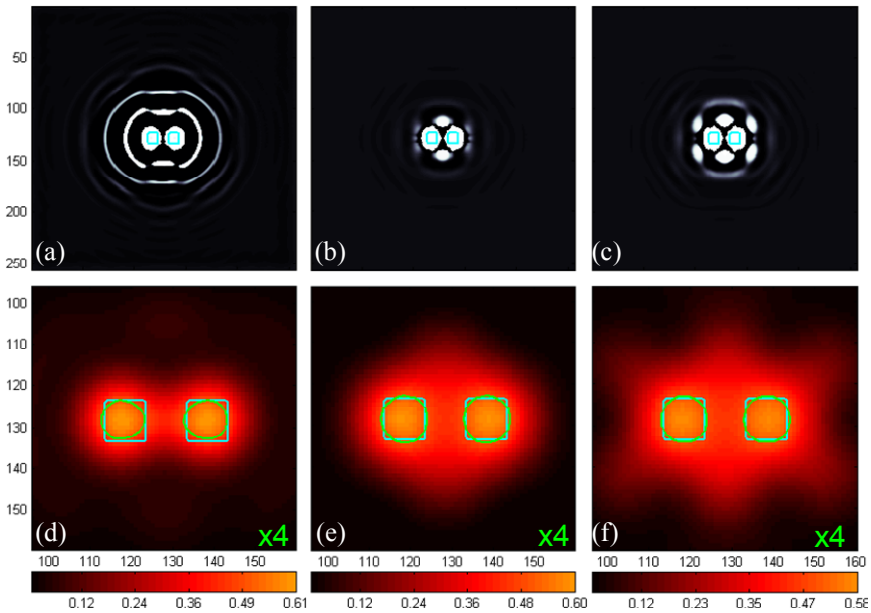

Fig. 11. The inverse results of two close contacts with different combinations of cost function components where (a) (γ_I, γ_R, γ_C)=(1, 0, 32.8947), (b) (γ_I, γ_R, γ_C)=(0.0041, 1, 0) and (c) (γ_I, γ_R, γ_C)=(0, 1, 0.0447). The corresponding contours and aerial images of (a)-(c) are shown in (d)-(f) where the cyan and green contours respectively label the drawn pattern edges and aerial image threshold contours.

As shown in Fig. 12 and 13, the averaged EPE and contrast of cut-lines 2 and 8 are plotted as a function of the iteration. The hybrid IL optimizations incorporating different cost functions exhibit better performance than those with individual cost functions only. However the image and image contrast combination still fail to corrected the mask, so the cyan curve holds in 10. Moreover the ripple phenomena are not eliminated when optimizing by hybrid cost function combination. The reason is that the over diffraction limit images, which induces the scattering light, will make resist image cost dominate the overall cost.

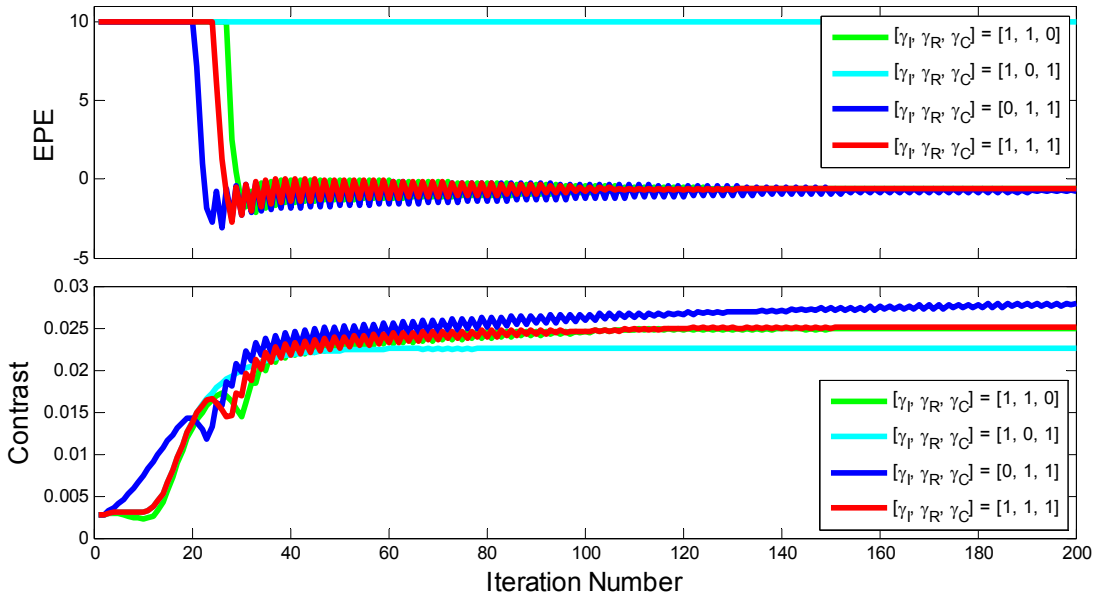

Fig. 12. The averaged EPE (Pixel) and contrast (AU) of cut-line 2 and 8 are plotted as a function of the iteration. The green, cyan, blue and red curves show that (γ_I, γ_R, γ_C) is under (1, 1, 0), (1, 0, 1), (0, 1, 1) and (1, 1, 1) respectively.

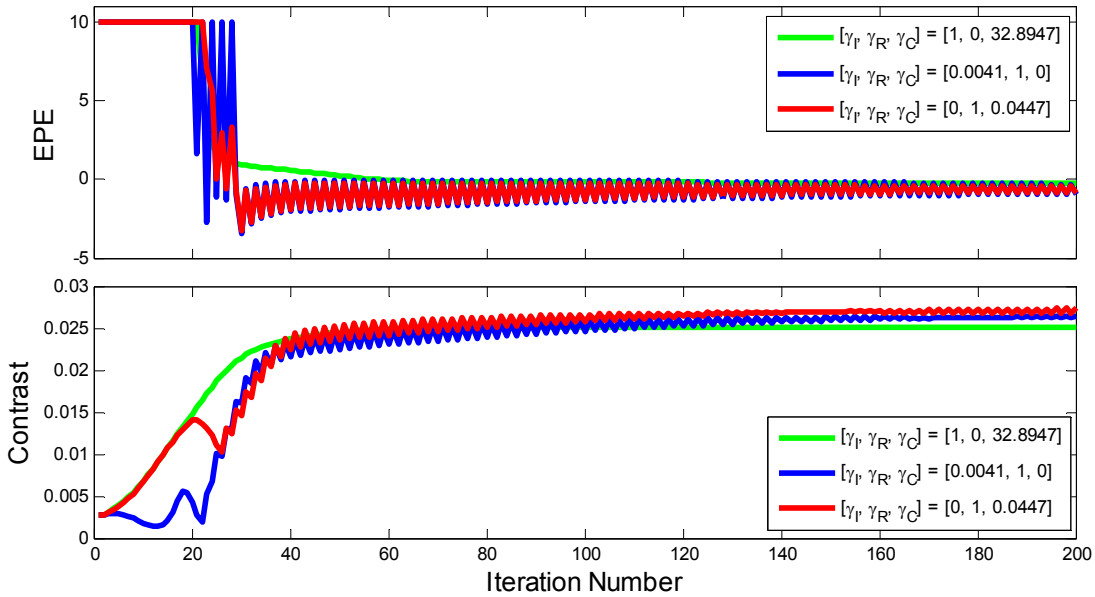

Fig. 13. The averaged EPE (Pixel) and contrast (AU) of cut-line 2 and 8 are plotted as a function of the iteration. The green, blue and red curves show that (γ_I, γ_R, γ_C) is under (1, 0, 32.8947), (0.0041, 1, 0) and (0, 1, 0.0447) respectively.

4. CONCLUSION

In conclusion, we developed a gradient descent approach to investigate three different objective functions and their combinations. All of them show very unique characteristics in the resulting mask patterns, aerial images, and contours. We demonstrate that a clever mix of the objective functions can push the resolution limits while maintaining manufacturing-friendly masks. In this work, a cost function composed of a dominant resist-image component and a minor aerial-image or image-contrast component achieves a good mask correction and contours close to the target for the resolution-challenging twin contacts. The intermediate results also validate the necessity of using sub-resolution assist features in advanced CMOS nodes. Furthermore the rapidly converging EPEs and contrast show the capability to quickly generate the reference mask for assisting current rule-based technique. We believe that these findings can provide informative guidance for the optimization of inverse lithography patterning in specific cases.

REFERENCES

[1] D. S. Abrams and L. Pang, "Fast inverse lithography technology," Proc. SPIE 6154, 534-542 (2006).
[2] C. Hung, B. Zhang, E. Guo, L. Pang, Y. Liu, K. Wang, and G. Dai, "Pushing the lithography limit: Applying inverse lithography technology (ILT) at the 65nm generation," Proc. SPIE 6154, 61541M (2006).
[3] L. Pang, Y. Liu, and D. Abrams, "Inverse lithography technology (ILT): What is the impact to the photomask industry?" Proc. SPIE 6283, 62830X (2006).
[4] S. Sherif, B. Saleh, and R. Leone, "Binary image synthesis using mixed linear integer programming," IEEE Trans. Image Process. Papers 4, 1252–1257 (1995).
[5] K. Nashold and B. Saleh, "Image construction through diffraction-limited high-contrast imaging systems: An iterative approach," J. Opt. Soc. Am. A Papers 2**,** 635–643 (1985).
[6] B. Saleh and S. Sayegh, "Reductions of errors of microphotographic reproductions by optical corrections of original masks," Opt. Eng. Papers 20**,** 781–784 (1981).
[7] Yuri Granik, "Fast pixel-based mask optimization for inverse lithography," J. Microlith., Microfab., Microsyst. Papers 5, 043002 (2006).
[8] A. Poonawala and P. Milanfar, "Mask design for optical microlithography — An inverse imaging problem," IEEE Trans. Image Process. Papers 16, 774–788 (2007).
[9] Xu Ma and G. R. Arce, "Generalized inverse lithography methods for phase-shifting mask design," Optics Express Papers 15, 15066-15079 (2007).
[10] Stanley H. Chan, Alfred K.Wong, and Edmund Y. Lam, "Initialization for robust inverse synthesis of phase-shifting masks in optical projection lithography," Optics Express Papers 16, 14746-14760 (2008).
[11] Jue-Chin Yu, Peichen Yu, and Hsueh-Yung Chao, "Model-based sub-resolution assist features using an inverse lithography method," Proc. SPIE 7140, 714014 (2008).
[12] Xu Ma and Gonzalo Arce, "Binary mask optimization for inverse lithography with partially coherent illumination," JOSA A, Papers 25, 2960-2970 (2008).
[13] Xu Ma and Gonzalo R. Arce, "PSM design for inverse lithography with partially coherent illumination," Optics Express Papers 16, 20126-20141 (2008).
[14] Xu Ma and Gonzalo R. Arce, "Pixel-based simultaneous source and mask optimization for resolution enhancement in optical lithography," Optics Express Papers 17, 5783-5793 (2009).
[15] M. Born and E. Wolf, [Principles of Optics], 7th(expanded) ed., Cambridge University Press, 598-599 (1999).
[16] J. W. Goodman, [Statistical Optics], John Wiley & Sons, New York, 306-307 (1985).
[17] A. K. Wong, [Optical Imaging in Projection Microlithography], SPIE Press, Bellingham Washington, 157-159 (2005).
[18] N. B. Cobb, [Fast optical and process proximity correction algorithms for integrated circuit manufacturing], University of California at Berkeley, Berkely California, 51-60 (1998).
[19] Edmund Y. Lama and Alfred K. K. Wong, "Computation lithography: virtual reality and virtual virtuality", Optics Express Papers 17, 12259-12268 (2009).
[20] Steven J. Leon, [Linear Algebra with applications], 6th ed., Prentice-Hall, 367-380 (2002).
[21] B. E. A. Saleh and M. Rabbani, "Simulation of partially coherent imagery in the space and frequency domains and by modal expansion," Applied Optics Papers 21, 15066-15079 (1982).

[22] W. Huang, C. Lin, C. Kuo, C. Huang, J. Lin, J. Chen, R. Liu, Y. Ku, and B. Lin, "Two threshold resist models for optical proximity correction," Proc. SPIE 5377, 1536–1543 (2001).
[23] M. Minoux, [Mathematical programming theory and algorithms], John Wiley & Sons, New York, 84-89 (1986).
[24] Eugene Hecht, [Optics], 4th ed, Addison Wesley, San Francisco, 495-497 (2002).

APPENDIX

As we mentioned in the manuscript, some sort of normalization is essential to compare the performance of objective functions on the same basis. Here, we perform a linear transformation for costs at the initial iteration and at the final iteration, when a steady state is reached, as shown in Eq. (a1) below.

$$y' = \left(\frac{10}{x_{\max.} - x_{\min.}} \right) x + \frac{10 x_{\min.}}{x_{\max.} - x_{\min.}}, \tag{a1}$$

where $x_{max.}$ and $x_{min.}$ denote the maximum and minimum of the original costs, which correspond to the initial and the final iterations, respectively. Fig. A1. shows the operation of the linear transformation on the single isolated contact used in Fig. 2 of section 3. As the same reason the two close contacts cases are under such operation as above description, but changing the mask for calculation.

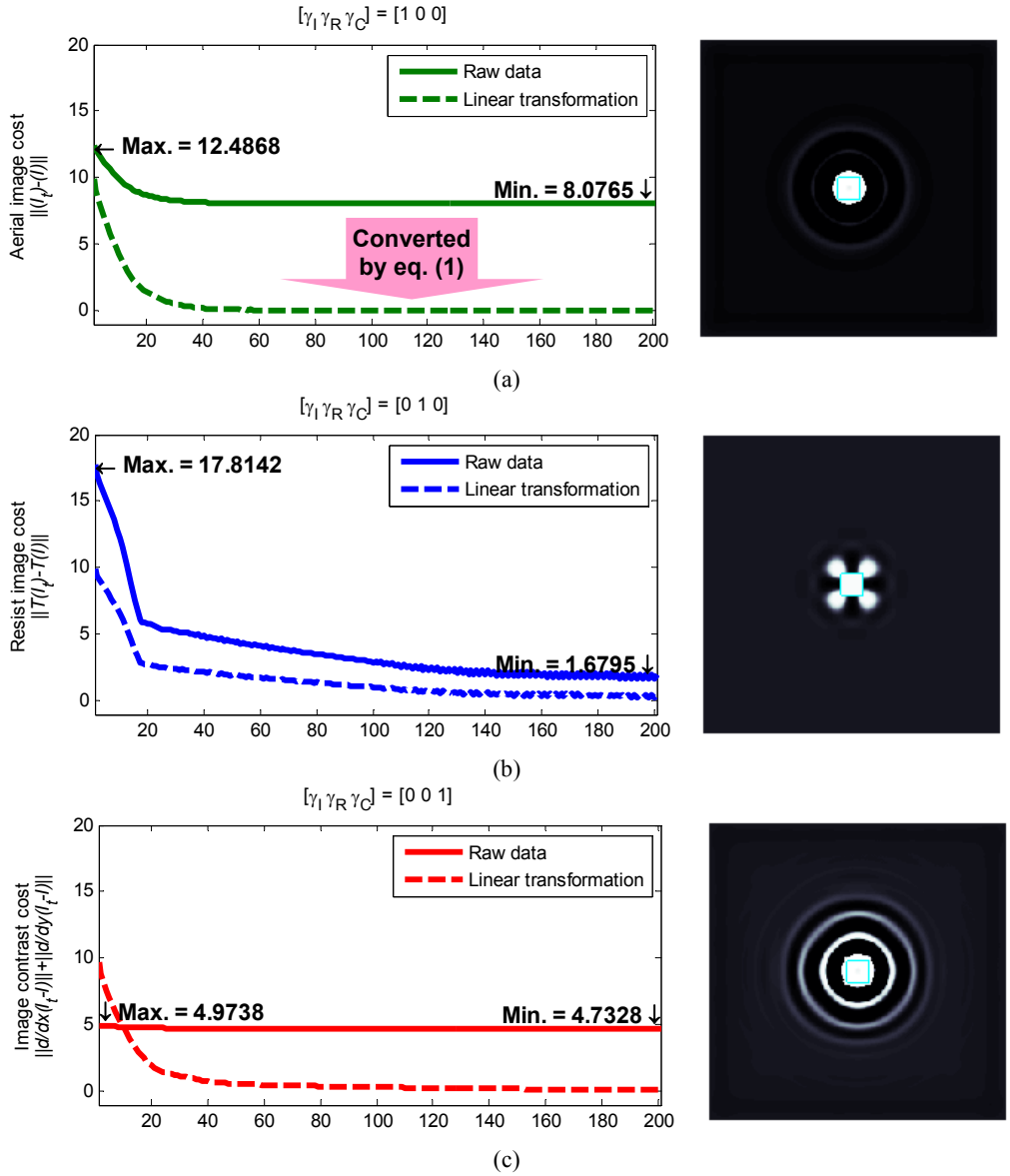

Fig. A1. The original and transformed cost curves of a large isolated contact as a function of iteration times: (a) (γ_I, γ_R, γ_C)=(1, 0, 0), (b) (γ_I, γ_R, γ_C)=(0, 1, 0) and (c) (γ_I, γ_R, γ_C)=(0, 0, 1). The cyan contour denotes the edge of the drawn features edges.