Understanding resolution limit of displacement Talbot lithography

P. J. P. CHAUSSE, * E. D. LE BOULBAR, S. D. LIS, AND P. A. SHIELDS

Department of Electronic & Electrical Engineering, University of Bath, Claverton Down, Bath BA2 7AY, UK
*pierre.chausse@bath.ac.uk

Abstract: Displacement Talbot lithography (DTL) is a new technique for patterning large areas with sub-micron periodic features with low cost. It has applications in fields that cannot justify the cost of deep-UV photolithography, such as plasmonics, photonic crystals, and metamaterials and competes with techniques, such as nanoimprint and laser interference lithography. It is based on the interference of coherent light through a periodically patterned photomask. However, the factors affecting the technique’s resolution limit are unknown. Through computer simulations, we show the mask parameter’s impact on the features’ size that can be achieved and describe the separate figures of merit that should be optimized for successful patterning. Both amplitude and phase masks are considered for hexagonal and square arrays of mask openings. For large pitches, amplitude masks are shown to give the best resolution; whereas, for small pitches, phase masks are superior because the required exposure time is shorter. We also show how small changes in the mask pitch can dramatically affect the resolution achievable. As a result, this study provides important information for choosing new masks for DTL for targeted applications.

Published by The Optical Society under the terms of the Creative Commons Attribution 4.0 License. Further distribution of this work must maintain attribution to the author(s) and the published article’s title, journal citation, and DOI.

1. Introduction

Periodic organisations of structures are useful for the creation of devices in many different fields such as plasmonics [1], photonic structures [2] or metamaterials [3]. Existing techniques are capable of easily patterning periodic sub-micron features, but each has their advantages and disadvantages. Deep-ultraviolet immersion lithography using a 193 nm excimer laser is widely used in industry and is capable of achieving a resolution of 14 nm [4], and extreme ultraviolet (EUV) sources with a wavelength of 13.2 nm are on the horizon to further decrease the minimum feature sizes [5]. However, the very high cost of these techniques limits their penetration into lower volume industries and research organisations. Electron beam lithography is versatile and can achieve very high resolutions (< 10 nm), but the cost is prohibitive for full wafer patterning due to the long patterning time [6]. However, reaching such high resolutions is not necessary for all applications.

Alternative cheaper nanopatterning methods have become available in recent years. Nanoimprint lithography is a promising technology for large-area patterning of features below 10 nm [7]. Thanks to the mechanical pattern transfer, the resolution is not limited by an optical system. However, the main drawback is the lifetime of the 3D master mould. Another approach is to use interference lithography, in which coherent sources of electrons [8] or photons [9] interfere, creating a periodic array of intensity. Since maintaining control of the sources before they interfere with each other can be a challenge, a solution is to derive the multiple sources close to the region of interference through diffraction from a periodic mask.

Displacement Talbot lithography is a recently developed technique for patterning large areas with sub-micron periodic features [10]. It is an extension of Talbot lithography, which uses the three-dimensional interference pattern created when monochromatic light diffracts through a periodic mask. Coherent light passing through a mask patterned with a periodic
structure creates different diffraction orders that subsequently interfere causing a self-imaging of the mask. This phenomenon is well-known and called the Talbot effect [11]. Characteristic of the interference pattern is its repeating nature along the axis perpendicular to the mask, with a spatial period called the Talbot length. By itself, this interference pattern is difficult to use for photolithography directly due to the size and complexity of the pattern. However, introducing a displacement during a photolithography exposure along the axis perpendicular to the mask integrates the optical field and solves these problems. This technique is called Displacement Talbot Lithography (DTL) and has the advantage of a theoretical infinite depth of field [10]. The main disadvantages of this technique are: 1) the low contrast between exposed and unexposed regions on the sample due to the mixing of the self-image and other secondary constructive interference features, and 2) the restriction to simple periodic features. Nevertheless, the illumination process will not be sensitive to surface roughness, or imperfect parallelism between the mask and the sample, or the depth of field; all important parameters in conventional photolithography. Recently more complex periodic structures have also been obtained using DTL [12] as well as sub-wavelength patterning [13].

Applications of this new patterning approach include metamaterials [3], III-V semiconductor photonic materials in the form of core-shell structures [14] and nanotube cavities [15]; neuronal network formation [16] and nanoimprint master creation [17]. The minimum feature size that can be achieved is dependent on the source wavelength, with 125-300 nm features having been achieved for a near-UV laser source [18] and 75 nm features for a deep-UV source [19]. However, for any particular source wavelength, the resolution limit of this new lithography method has not yet been reported in the literature. Having a better understanding of such a limit will permit a greater use of this fast, cheap and flexible lithography technique.

In this paper, we analyse the resolution limit using the results from computer simulations of the intensity pattern seen by the sample, where the model has first been validated through a comparison of simulated results with existing experimental data. As a result, we determine the smallest feature size achievable as a function of the mask parameters, such as whether it is a phase or amplitude mask, the mask pitch and the feature diameter. Thanks to this comparison between experiment and modelling, the conditions required to optimize the resolution will be discussed.

2. Modeling of DTL

2.1 Simulation of aerial image

A MATLAB computer model was developed to simulate the operation of a DTL machine (PhableR 100, EULITHA) in which the light source is a 375 nm UV laser. The optical system generates a plane wave illuminating a conventional lithography mask at normal incidence so that the light arriving at the mask is homogenous, unpolarised, and in phase. The complex field distribution has been represented by a scalar field. It allows the calculation of any field distribution in any plane parallel to the mask, by the use of a free-space propagation method realised in Fourier space [20]. Both amplitude and phase masks are considered in the modelling, and any impact of the metal or phase shift layer thickness on the electric field has been neglected. Consequently, contributions from the different mask regions propagating behind the mask are given amplitudes of 1 and 0 for a chrome amplitude mask and 1 and $-1$ for a phase mask.

Experimentally, the integration of the three-dimensional light field has been realized 100 μm away from the mask, therefore Fraunhofer conditions can be applied in the subsequent modelling to calculate the three-dimensional light field behind the mask, known as the Talbot carpet. The Fourier transform of the mask generated is calculated as well as the 2D spatial wave vectors in the plane of the mask. All 3D wave vectors can be determined by knowing at each step of the integration the vertical position of the photo-sensitive layer. The electric field
amplitude at a specific coordinate is calculated by applying the inverse Fourier Transform to the combination of the 3D wavevectors, the Fourier transform of the mask and the depth positions [9]. The electric field is then multiplied by its conjugate to obtain the surface light intensity.

Figure 1 shows the results for the modelling of an amplitude grating with a periodicity of 600 nm and 200 nm openings. Figure 1(a) shows a cross-section in the x-z plane and illustrates the classical Talbot effect. By moving the sample through an integer number of Talbot lengths, \( L_T = \lambda / \left( 1 - \sqrt{1 - \lambda^2 / p^2} \right) \) (\( p \) is the pitch on the mask and \( \lambda \) is the laser wavelength), the intensity seen by the photoresist is integrated to remove the z-dependence as shown in Fig. 1(b). To ensure accurate integration within the simulation, a vertical step resolution of 1/100 of the Talbot length is sufficient for a small pitch mask (less than twice the laser wavelength), but this is insufficient for larger pitches due to the greater complexity of the Talbot carpet, since it is generated from a higher number of diffraction orders. Therefore, a step resolution of 1/200 has been chosen. A further increase in resolution only increases the computational time with no change in the simulated pattern. The resulting intensity in the x-y plane that will be transferred into the resist, called the aerial image, is shown in Fig. 1(c). The 600 nm pitch grating mask results in a 300 nm period grating on the sample.

![Fig. 1. Modelling of a 600 nm grating amplitude mask with 200 nm openings. Normalized figures of a) the Talbot carpet, b) the Talbot carpet after integration over the Talbot length, and c) the aerial image.](image)

### 2.2 Definition of parameters on aerial image

The computer simulations allow the aerial image to be determined for any mask, which is important since the results are not intuitive. As a function of the pitch, the filling factor and the nature of the mask, the aerial image evolves dramatically. This study has focused on two types of mask: square and hexagonal arrangements of circular features. Since the results from these two structures are quite similar, we primarily discuss the results for the hexagonal patterns, with the results from the square masks presented in the appendices. Figure 2 shows the simulation of the aerial image for a 1.5 µm pitch hexagonal amplitude mask with 800 nm openings. This mask pattern has the feature that it creates an aerial image with the same hexagonal pattern.

The aerial image is then transferred into the resist, and in order to understand the size of the resist features that can be created, we define three figures of merit in order to compare the aerial images from different masks: 1) the theoretical width of the pattern achievable, 2) the relative intensity of the background, and 3) the relative intensity of the maximum of the unwanted features within the images, which we call the secondary maxima. The latter two figures of merit are relatively easy to define by referring to the aerial image in Fig. 2(b): the
relative intensity of the background is the value corresponding to the minimum of the aerial image (e.g. at the location indicated by the orange circle in Fig. 2(b), and the relative intensity of the secondary maxima is the peak value of the unwanted patterns (e.g. the red circle in Fig. 2(b)). The width, however, is more difficult to define and is discussed below.

![Fig. 2](image)

**Fig. 2.** a) Hexagonal 1.5 μm pitch amplitude mask with 800 nm openings, b) simulated normalized aerial image, and c) cross section corresponding of the black line.

A cross-section through the primary intensity maxima, represented by the black line in Fig. 2(b), is shown in Fig. 2(c). In order to define a width, a threshold relative intensity level needs to be specified, representing the illumination dose at which any photoresist experiencing a lower dose will not be developed completely, whilst any receiving a higher dose will be totally developed. Increasing the total exposure dose illuminates more of the resist above the absolute threshold intensity and increases the size of the features. Conversely, in theory, by raising the threshold, the size of the features can be progressively reduced, with the highest resolution occurring when only the tip of the peak in Fig. 2(c) is above the threshold.

A primary factor in how high the threshold can be raised is the contrast of the photoresist, defined as the ratio of the range of doses for which the resist is partially developed to the dose when all resist is removed. Typical figures are ~10% for ULTRA-i 123 (Dow Corning) in MF CD26 developer. Higher contrast resists such as PFI-88 (Sumitomo Chemical Co.) can achieve lower values. A further consideration is the reduced steepness of the intensity curve near the intensity maxima. This causes the homogeneity of features close to the resolution limit to become highly dependent on the homogeneity of the optical system. Finally, the finite dispersion of the incident laser collimation (< 0.6 mR), which has not been incorporated into the model, will also impact the smallest feature achievable. Therefore, in order to determine a suitable threshold level that can be used as a basis to compare masks of different dimensions we turn to experimental measurements.

### 2.3 Experimental results for minimum feature size

Experiments using a hexagonal mask were performed to determine the minimum feature size achievable and shed light on a suitable figure for the threshold value to determine the minimum feature size in the modelling. The layers comprised 240 nm of ULTRA-i 123 on top of a bottom anti-reflection layer (WIDE 30C, Brewer Scientific) – see Experimental Section for further details. Secondary-electron SEM images of the resist features created from a hexagonal amplitude mask with a 1 μm and 1.5 μm pitch are shown in Figs. 3(a)-3(b) and Figs. 3(d)-3(e), respectively, for two different exposure doses. In Figs. 3(a)-3(d), for which the exposure dose was too low, a large dispersion of hole diameters and even closed holes can be observed in the resist. Higher exposure doses produce larger openings with better homogeneity (Figs. 3(b)-3(e)). The uniformity across a whole 2-inch wafer has been analysed for different illumination doses and is shown as a box plot in Figs. 3(c)-3(f). The uniformity
clearly improves above a threshold dose of between 70 and 80 mJ/cm² for the 1.5-µm-pitch pattern and 130-140 mJ/cm² for the 1-µm-pitch pattern.

Two dashed lines have been added to Figs. 3(c)-3(f) corresponding to values from the modelling of these masks of the diameter of the primary pattern achieved when the critical exposure dose needed to fully develop the resist matches 80% and 90% of the peak value. This means that the top 20% and 10%, respectively, of the aerial image, will be fully developed and not the rest. For a critical exposure dose matching the 90% threshold, we can observe that with both masks, we have a large dispersion and even closed holes. In contrast, an 80% threshold leads to a lower dispersion and a high homogeneity. Whilst a variation across a full 2-inch wafer was found, with larger openings on the edge of the wafer than in the center, likely due to an imperfect illumination or a non-uniform BARC thickness, locally, a dispersion lower than 10 nm was obtained.

As a result, the threshold of 80%, determined experimentally to balance high resolution and high homogeneity, will be used to calculate the minimum feature width from aerial images simulated through modelling. This lower threshold ensures that the imperfect collimation of the laser source, local mask variation, and illumination inhomogeneity are included in the modelling. Thus, masks with different geometries can be theoretically compared in order to understand their impact on the resolution limit of displacement Talbot lithography.

![Fig. 3. Openings in photoresist obtained for a 1.5 µm mask at a) 70 mJ/cm² b) 80 mJ/cm² and c) the size distribution as a function the illumination dose. Openings in photoresist obtained for a 1 µm mask at a) 130 mJ/cm² b) 140 mJ/cm² and c) the size distribution as a function of the illumination dose. The dashed lines represent the theoretical width for 80% and 90% thresholds.](image)
3. Impact of mask design on figures of merit

Using the computer model that has been developed, the influence of different parameters on the resolution has been analysed. Amplitude and phase masks have both been modelled for a range of pitches and mask feature sizes. Whilst amplitude masks are commonly used due to their cost and ease of manufacture, phase masks can improve the resolution of classical lithography. Therefore, it is important to understand whether the same is borne out for DTL.

3.1 Amplitude mask

Simulations of the aerial images were generated for mask pitches between 0.5 and 3.2 µm and mask opening diameters from 20 nm to 90% of the pitch with a resolution step of 20 nm in order to allow a good understanding of the impact of the different parameters on the aerial image. The integration was performed over four Talbot lengths to prevent artefacts arising from the grid definition. From the aerial images, the theoretical width, the relative background of intensity and the maximum of the secondary patterns were determined and these have been plotted in Figs. 4(a)-4(c), respectively, as a function of the pitch and feature diameter of the mask.

Figure 4(a) shows that the smallest features are obtained when the mask openings are smaller than the wavelength. With these feature sizes, the diffraction can be considered to arise from an array of point sources with hemispherical wavefronts. However, the transmission is highly reduced for subwavelength holes, where the transmission $T(\lambda)$ scales according to $T(\lambda) \approx (d / \lambda)^4$ [21-23]. A second valley giving small features also appears for a filling factor of 50% though the theoretical width achievable is not quite as small. Sharp vertical and horizontal lines are also apparent along this valley where the minimum feature size abruptly decreases with increasing size. These appear periodically at multiples of the laser wavelength as a result of the incremental addition of further diffraction orders as the wavelength is increased.

The plot of the relative intensity of the background in Fig. 4(b) shows that the background is lower for smaller mask openings. Thus, the highest resolution features can simultaneously be achieved with having a low background.

In contrast to the previous two figures of merit, the plot of the relative intensity of maximum of secondary patterns (Fig. 4(c)) is not as simple. Low maxima of the secondary patterns (dark blue regions in Fig. 4(c) can be obtained in two cases: for small mask openings and particular pitches, or for masks with a filling factor of 33%. This specific value can be explained by diffraction theory. Indeed, in the case of gratings, the disappearance of some diffraction orders can occur if the ratio of the grating width on the period is a rational fraction. The same phenomenon is appearing here with a 2D mask. By tuning the diameter opening, some diffraction orders can be canceled, and so reduce the importance of the secondary pattern. Furthermore, for small mask openings, all the diffraction orders are represented and more or less have the same amplitude. This can explain why the resolution is improved in Fig. 4(a).
The periodic nature of the plots can be seen in cross-sections through the colour plots of Figs. 4(a)–4(c) as shown in Fig. 5 for a fixed opening of 400 nm. In Fig. 5(a), the best resolution will be obtained when the pitch is small. The resolution then decreases until reaching a plateau. In this area, an improvement of 100 nm can be made by slightly changing the pitch. DTL resolution is not only dependent on the hole opening size but also on the pitch.

Figure 5(b) illustrates that at large pitches we have a high secondary pattern intensity (around 45%). However, by choosing a pitch of 1.25 or 2 µm, the secondary pattern intensity can be decreased. Furthermore, when the mask pitch is below but close to twice the wavelength (0.75 µm), the secondary pattern peak intensity drops significantly.

3.2 Phase mask

Corresponding calculations of the aerial image were made for a phase mask with a similar range of pitch and feature sizes, from which the figures of merit were extracted and displayed in Fig. 6. The most significant difference between the phase and amplitude mask is the much higher background level. Given that a large difference between the peak value and the background is essential for a good process, the results for the theoretical width and maximum of the secondary pattern were only extracted when the background was less than 70% of the maximum.
The modelling shows that the conditions to enhance the resolution are not the same as for amplitude masks. Indeed, larger openings lead to a better resolution (Fig. 6(a)). Nevertheless, the vertical and horizontal periodic features previously observed for the amplitude mask are also present.

Optimising all the figures of merit simultaneously is not possible. A dark blue valley appears in the relative intensity of the background for a constant ratio of feature size diameter to pitch of 66% (Fig. 6(b)), corresponding to a filling factor of 40%. This valley can be explained by the ratio between the two phases of the mask. Indeed, if one phase is predominant, the destructive interferences occurring in the background won’t be optimized. Similar to the case of the amplitude mask, there is also a minimum in the relative intensity of the secondary maxima for a filling factor of 33% (Fig. 6(c)). Whilst the positions of these two respective valleys are not perfectly matched, their broadness allows a simultaneous reduction in the relative intensities of both the background and the secondary maxima.

However, these regions do not coincide consistently with high-resolution features. And so the resolution is compromised.

Fig. 6. Theoretical a) width of the pattern achievable, b) relative intensity of the background and c) relative intensity of maximum of secondary patterns for a hexagonal phase mask.

4. Discussion

In this study, computer simulations of the DTL aerial image have been analysed for a laser wavelength source of 375 nm and assuming a specific photoresist and developer. Minimum feature sizes are predicted to be around 75-100 nm for an amplitude mask with openings smaller than the laser wavelength. It is expected that further decreasing the wavelength to 193 nm will improve the resolution to the sub 50 nm range. Furthermore, using a resist and developer with higher contrast will permit smaller features to be obtained.

Various assumptions have been made in the modelling, with the finite thickness of the metal layer on the mask, the dark erosion of the resist, the flow of the developer and the 3D shape of the structures having been omitted. The finite thickness could impact the accuracy of the results for very small mask openings and the simplified resist model would underestimate the rounding of the resist profile. In particular, the effect of dark erosion is likely to reduce the optimum development time to one that balances high contrast with low dark erosion.

Nevertheless, the modelling results agree with other modelling studies of the Talbot effect [9–11]. In addition, our previous work has successfully validated the model and predicted
experimental results qualitatively; for high illumination doses for certain hexagonal masks, the secondary patterns were seen to merge to create ring structures after development. The shape of the ring obtained experimentally was found to correspond well with the modelled shape [15] and could be used to create resonant cavity modes in axial InGaN/GaN nanotube microcavities.

One of the benefits of DTL is being able to vary the exposure dose to create a large range of feature diameters; a phenomenon that arises from the shape of the intensity peaks in the aerial image. For example, the 1.5 μm hexagonal amplitude mask with 800 nm openings allows a range of features in resist from 250 - 650 nm; thus demonstrating the flexibility of DTL. For this purpose having a low background and a less pronounced secondary pattern are more important factors than the resolution and these are better satisfied with phase masks.

By comparing Figs. 4-6, it can be seen that for mask pitches smaller than two wavelengths, both phase and amplitude masks achieve the same resolution. In this case, a phase mask would be preferred because of the higher transmission through the mask, leading to a shorter illumination time. For higher pitches, an amplitude mask would offer the smallest features, but at the expense of a long exposure time.

5. Conclusion

Displacement Talbot Lithography is a new lithography method that can pattern periodic features across a large area quickly and cheaply. The conditions to reach the smallest features with DTL are not trivial and in certain situations conflict with the optimisation of other figures of merit that influence the lithography process. For a specific i-line resist, the impact of the pitch and the size of the openings on the mask has been found and analysed not only on the minimum size feature achievable, but also the sensitivity of the resolution, the background, and the unwanted secondary pattern intensities. This study shows how sub-100 nm features can be achieved with DTL across large areas with conventional i-line resists and illumination. This limit is substantially below other classical photolithography methods using the same illumination wavelength.

6. Appendices

6.1 Experimental data

Silicon wafers were coated with a bottom antireflective layer (Wide 30, Brewer Science) prior to a 240 nm positive resist (Ultra i-123 diluted from 800 nm as supplied). The wafers were exposed via DTL using a hexagonal amplitude mask with either a 1.5 μm pitch and 800 nm mask openings or with a 1 μm pitch and 550 nm openings. The Talbot length associated with these masks are 8.80 μm and 3.80 μm, respectively (the section 6.2). A Gaussian velocity integration was applied and a travel distance of 8 Talbot lengths was chosen to ensure homogenous integration over several Talbot motifs. Multiple series of illuminations were realized to ensure a high reproducibility of the process: 70 to 100 mJ/cm² for 1.5 μm, and 130 to 180 mJ/cm² for 1 μm, with 10 mJ/cm² step. The wafers were immersed for 210 s in the developer (MF CD26). This development time was chosen to reach a 10% contrast in the resist profile calculated thanks to the Dills model [24–26].

The statistical data on the feature diameters in Fig. 3 was determined from image analysis of 3 SEM pictures per sample. The pictures were consistently taken around the center of the wafer and the magnification was chosen to observe around a hundred features in one picture.

6.2 Hexagonal mask modeling

The non-primitive unit cell size for a hexagonal organization has a specific characteristic: one cell edge is \(\sqrt{3}\) larger than the other. Due to this being an irrational number and the use of matrices in MATLAB, a rounding error will occur during the definition of the matrix size. This error can be mitigated by increasing the resolution without sacrificing the computational
time. Furthermore, for hexagonal masks, the distance used in the Talbot period calculation is not the nearest neighbour distance but this distance multiplied by $\sqrt{3}/2$ [27].

So the Talbot length calculation becomes:

$$L_T = \frac{\lambda}{1 - \sqrt{1 - \frac{\lambda^2/4}{p^2/3}}}$$

6.3 Square mask

The use of square mask with DTL allows the creation of a new square organisation with a smaller pitch. The distance between the pattern in the resist will be $\sqrt{2}$ smaller than that on the mask (Fig. 7). However, the intensity of alternate maxima in the pattern can differ; in Fig. 8(b), the maxima aligned with the holes and the additional maxima created between the openings are not exactly the same. As a result, the analysis of the difference of amplitude between the two patterns is important for homogenous patterning and uniformity.

![Fig. 7. Reduction of pitch for a 1 μm square array. a) Mask and b) aerial image.](image)

6.3.1 Amplitude mask

The conclusions with the square amplitude mask are similar to those of the hexagonal amplitude mask; small mask openings result in smaller resist features and the relative intensity of the background is low. However, the secondary pattern intensity is even higher than for the hexagonal mask. A valley with low values exists, but it occurs for mask openings and pitches not matching the optimum values for the two previous parameters. The intensity of the secondary patterns will be low for pitches smaller than $\sqrt{2}$ times the wavelength (around 530 nm). Figure 8(d) shows that a difference of 6% can appear between the patterns, and since the integration was realized over 4 Talbot lengths, this is unlikely to be an artefact in the modelling. In conclusion, to obtain a good homogeneity, small openings should be chosen for specific pitches, e.g. 1.05, 2 or 2.5 μm. Nevertheless, a high intensity of the secondary pattern is unavoidable.
6.3.2 Phase mask

The same conclusions can be drawn when using a square phase mask as for the hexagonal phase mask. Larger openings will give a higher resolution. The background will be quite low in this area, and the value of the secondary maxima is going to be smaller than for the hexagonal one.
In Fig. 9(d) the intensity ratio between neighbouring patterns shows similar values to the amplitude mask with some parts being almost the same. In order to follow a valley corresponding to a specific filling factor to obtain a good resolution, specific pitches and opening diameters should be targeted.

6.3.3 Discussion

For square mask, the smallest feature will be obtained with the amplitude mask but the secondary pattern intensity is significant and could even be critical. In square phase mask, for large openings, the secondary patterns are not going to as high as the hexagonal case, with reasonably small features being achievable. So for the patterning of square array, a phase mask will be a better choice for a high resolution with limited parasitic effect.

Funding

Engineering & Physical Science Research Council, UK (EPSRC) (EP/M015181/1, EP/M022862/1)

Acknowledgments

All data created during this research is openly available from the University of Bath Research Data Archive at [https://doi.org/10.15125/BATH-00570].

References


