Mathematical problems of holographic mask synthesis

Borisov, M., Chelubeev, D., Chernik, V., Merkushov, L., Rakhovskiy, V., et al.

M. Borisov, D. Chelubeev, V. Chernik, L. Merkushov, V. Rakhovskiy, A. Shamaev, "Mathematical problems of holographic mask synthesis," Proc. SPIE 11324, Novel Patterning Technologies for Semiconductors, MEMS/NEMS and MOEMS 2020, 113241I (23 March 2020); doi: 10.1117/12.2551942

SPIE.

Event: SPIE Advanced Lithography, 2020, San Jose, California, United States

Mathematical problems of holographic mask synthesis

M. Borisov, D. Chelyubeev, V. Chernik, L. Merkushov, V. Rakhovskiy*, A. Shamaev
NANOTECH SWHL GmbH, Überlandstrasse 129, 8600 Dübendorf, Switzerland

ABSTRACT

SWHL Nanotech team have implemented effective algorithms and developed scalable software allowing to synthesize holographic masks for various lithography applications including MEMS, MOEMS and high-end IC production. Most of the technical problems of state-of-art projection photolithography such as 3D-imaging, quality optimization were stated and solved as a completely numerical problems in the case of holographic lithography approach. Nanotech SWHL team developed effective algorithms of the holographic mask synthesis based on FFT with the complexity of $O(N\,ln\,N)$, which allowed to synthesize holographic masks for any IC layer. We developed the continuous phase-shifting optimization method based on WFS, DFS and gradient descent, in which a hologram is synthesized not for the original pattern, but for a pattern with altered amplitude and phase distribution. Like in other RET, the holographic mask synthesized for the properly altered pattern provides a much better-quality image of the original pattern. Thus, today it is possible to use modern computing clusters for the synthesis of holographic masks and to implement them in inexpensive and sustainable devices for holographic photolithography.

Keywords: Computer Generated Hologram, Fast Fourier Transform, phase-shift, optimization, steepest descent

1. INTRODUCTION

The constant increase of the resolution requirements made the technological process of projection lithography very sensitive to external perturbations and increased the cost of mask manufacturing since the mask design utilized ever-declining transmissive zones (TZ). Resolution enchantment techniques (RET), e.g. OPC, PSM, SMO also significantly increase the mask costs and complicated the manufacturing process[1]. The RET modelling turned into an intricate inverse problem of electrodynamics, aimed to fight against manifestations of the wave nature of light (proximity effect, Gibbs phenomenon). Negative effects of the diffraction are strengthened while we move to advanced nodes.

Sub-Wavelength Holographic Lithography (SWHL) has been intensively developed in recent years[2-4]. This technology responds to several challenges that have emerged in the microelectronics industry over past decades. The main idea is to assimilate wave properties of light using holographic methods and not to suppress them. However, direct holographic mask synthesis is affordable for small patterns, but it is unfeasible for a real IC layer consisting of billions of elements even on modern computer clusters. Nanotech SWHL team developed effective algorithms of the holographic mask synthesis based on FFT with the complexity of $O(N\,ln\,N)$, where N is the image elements quantity, which allowed to synthesize holographic masks for any IC layer.

Moreover, we achieved a subwavelength resolution of about 0.5λ using optimization methods. We developed the continuous phase-shifting optimization method based on WFS, DFS and gradient descent, in which a hologram is synthesized not for the original pattern, but for a pattern with altered amplitude and phase distribution. Like in other RET, the holographic mask synthesized for the properly altered pattern provides a much better quality image. The continuous phase-shift method complexity is $O(N)$ and worth noting in terms of final mask cost.

The continuous transmission function of a hologram can be approximated using TZ with sizes of several wavelengths since the optical scheme based on lensless Fourier hologram provides a transmission function of smooth variation. Further, our mask synthesis and optimization methods for plain patterns were extended for the one-shot exposure from a single mask for 3D objects and patterns on piecewise-flat surfaces. The same result is impossible to achieve with a single mask in projection lithography due to the limited depth of focus. The modified algorithms provide promising results for 3D imaging, despite the exact reconstruction of 3D objects is an ill-posed inverse problem. We have reformulated the problem as an optimization one in which we try to find a transmission function reconstructing an image of sufficient quality. This is a promising way to solve inverse problems of electrodynamics with numerous applications.

*rakhvi@nanotech-swhl.com; phone +41 79 583 19 57; nanotech-swhl.com

Novel Patterning Technologies for Semiconductors, MEMS/NEMS and MOEMS 2020,
edited by Martha I. Sanchez, Eric M. Panning, Proc. of SPIE Vol. 11324, 113241I
 · doi: 10.1117/12.2551942

2. TYPES OF HOLOGRAPHIC MASKS

Holographic mask is a physical implementation of an amplitude modulation function of the light wave. This function can be obtained as the intensity of reference and object waves interference pattern using the classical principle proposed by D. Gabor[5,6]. The simplest design of the holographic mask is an opaque screen with openings, where the function of amplitude modulation is approximated with the density and size of openings. Elements on such mask should be large enough for the light to pass. It is also necessary to take into account that obtaining high-resolution light images with elements of subwavelength size is possible only with a high angular aperture of the illuminator. Large angles of incidence of the restoring wave and the complex multilayer structure of the mask require to develop an appropriate model of diffraction, which takes into account effects of re-reflection at layers boundaries, waveguide effects in openings. We will describe the types of holograms with different structures and point to their features and advantages.

2.1 Amplitude hologram

The Gabor's method of hologram generation is associated with the registration of interference of mutually coherent reference and object waves. Let (ξ, η) are coordinates on the hologram plane, $O(\xi, \eta)$ is the complex amplitude of reference wave on hologram plane, $\Pi(\xi, \eta)$ is the complex amplitude of object wave. Then, the function

$$T(\xi, \eta) = \frac{|O(\xi, \eta) + \Pi(\xi, \eta)|^2}{|O(\xi, \eta)|^2},$$

which takes only non-negative values, will be considered as the function of transmission for amplitude hologram. Note that

$$T(\xi, \eta)O^*(\xi, \eta) = \frac{(|O(\xi, \eta)|^2 + |\Pi(\xi, \eta)|^2)O^*(\xi, \eta)}{|O(\xi, \eta)|^2} + \frac{\Pi(\xi, \eta)O^{*2}(\xi, \eta)}{|O(\xi, \eta)|^2} + \Pi^*(\xi, \eta).$$

Thus, if we highlight the plate with the transmission function $T(\xi, \eta)$ by the restoring wave $O^*(\xi, \eta)$, which is a reversed reference wave, the resulting wave in the near field of plate can be decomposed into three components:

- $\Pi^*(\xi, \eta)$ is the field conjugated to the object wave and focuses in the object image at the object plane;
- $\frac{(|O(\xi,\eta)|^2+|\Pi(\xi,\eta)|^2)O^*(\xi,\eta)}{|O(\xi,\eta)|^2}$ is the zero order of diffraction, which in the case of a spherical reference wave and square holographic mask focuses into a bright cross image, which is the diffraction at the edges of the square mask;
- $\frac{\Pi(\xi,\eta)O^{*2}(\xi,\eta)}{|O(\xi,\eta)|^2}$ is the wave field, which focuses distorted object image into a symmetrical with respect to the focus of the restoring wave.

It is necessary to notice that for possibility of physical realization of the hologram it is required that the coefficient of amplitude modulation does not exceed one in each point. Therefore, making the normalization will receive

$$T(\xi, \eta) = \frac{|O + \Pi|^2 - (O^2 + \Pi^2)}{M_1 * |O|^2},$$

where $M = max\{|O(\xi, \eta) + \Pi(\xi, \eta)|^2 / |O(\xi, \eta)|^2\}$ is a constant value.

Realization of this transmission function is a quartz plate with a thin opaque chromium layer and a set of openings in it, the density and size of which corresponds to the transmission function.

2.2 Amplitude hologram with phase-shifting layers

The zero order of diffraction concentrates most of the restoring wave energy, therefore, interfering with the useful image, it creates significant disturbances. To eliminate a component that corresponds to the zero order, the corresponding real valued term must be subtracted from the transmission function. Then the function

$$T_1(\xi, \eta) = \frac{|O + \Pi|^2 - (O^2 + \Pi^2)}{M_1 * |O|^2} = \frac{O^*\Pi + O\Pi^*}{M_1 * |O|^2}$$

takes values from the interval $[-1;1]$ by selecting a normalization factor M_1 similar to the coefficient M in the previous case. Modulation of a wave by a negative factor means proportional change of its amplitude and shift of its phase by half of wavelength. Therefore, such transmission function can be implemented physically by applying an additional phase-shifting layer at points with negative modulation.

Note that in this case there is a gain in the diffraction efficiency of the hologram. This is a consequence of inequality

$$M_1 = max\left\{\frac{|O^*\Pi + O\Pi^*|}{|O|^2}\right\} < M = max\left\{\frac{|O + \Pi|^2}{|O|^2}\right\},$$

which is a consequence of triangle inequality. With the right choice of reference wave, the diffraction efficiency is increased by about five times compared to the amplitude mask. The effect is confirmed by numerical experiments.

2.3 Hologram with dimming layers

One of the possible problems of holographic image creation is the lack of light transmission by small openings. On a holographic mask with phase-shifting layer there are quite a lot of such small openings. As shown by previous studies based on the Kirchhoff approximation, on a holographic mask whose dimensions are about 5 times larger than the image size, the vast majority of the transmission zones are in the range $[0;0{,}7\lambda]$. In addition to that, mask elements with characteristic dimensions less than 1.7λ in a layer of chromium with a thickness of about 0.5λ exhibit waveguide properties. This leads to the fact that the calculation of the diffraction on mask elements with sizes in the range $[0{,}7\lambda;1{,}7\lambda]$ is complex, and, taking into account errors of mask manufacturing, is unreliable. During the passage through the waveguide, the wave is divided into modes that propagate in different directions (dispersion), that modes fade when moving along the waveguide with different decrement, reflect differently from the open ends of the waveguide, transmit energy to each other in the presence of local irregularities within the waveguide and also due to the finite conductivity of the waveguide material. Thus, the use of elements whose size is less than 1.7λ, on the mask, which is calculated and optimized in the framework of Kirchhoff's theory, is undesirable.

In order to increase the range of used elements size, it is necessary to significantly increase the maximum size of the transmission zone, and hence the size of the hologram. The idea of using a darkening layer is to replace small openings with larger ones with a filtering layer that reduces the final transmittance. Thus, it is possible to deduce the size of small transmission zones from the undesirable range and thus save a significant part of the information on the mask. You can go further and allow the use of multiple dimming layers with different absorption factors.

To calculate the thickness, refractive indices and absorption of phase-shifting and additional darkening layers, a diffraction model is required, which is fundamentally different from the Huygens-Kirchhoff formulas. It is important that the function describing the distribution of optical parameters of the mask layers may have discontinuities.

3. HOLOGRAPHIC IMAGE OPTIMISATION

We have developed and implemented fast algorithms for the quality of reconstructed images optimization. The computational complexity of the algorithms is $O(N)$, where N is the image elements quantity. Optimization algorithms are based on gradient methods and special graph coloring algorithms. These quality optimization algorithms are virtual counterparts of the OPC and the phase contrast technologies well-known in projection lithography.

3.1 Holographic mask variation method

Optimiztion problem can be formulated as follows: to make a mask, so that the known coherent source radiation passing through that mask will form imprint in photoresist sufficiently close to the target. Let S is the operator which assigns holographic mask M to any object O, image of which we need to produce. That is $S(O) = M$. Specification of this operator depends on the parameters of the optical system (reference and restoring wave, aperture, etc.) and the available mask design. It could be named as synthesis operator and represent a model of diffraction and interference from the synthesis of

the transmission function process described above. We have developed a special software module for this operator. It is based on fast Fourier transform (FFT) and its computational complexity is $O(N \ln N)$, where N is the image elements quantity. Also let D is an operator which assigns an image I to the holographic mask M, that is $I = D(M)$. This operator is an integral transformation of the mask transmission function with some kernel $K_D(A,B)$. We have developed a special software module for the operator D too with FFT implementation and same computational complexity. Consider the optimization problem of finding a holographic mask, which provides an image of the best possible quality:

$$\sigma(O, D(M)) \xrightarrow[M]{} min$$

where S is some metric showing the difference between obtained image and target obect. In the simplest case (in discrete form for the quadratic functional) this problem can be represented as

$$\sigma = \sum_{i,j}(b_{ij}b_{ij}^* - d_{ij})^2 \xrightarrow[\{a_{pq}\}]{} min, \quad b_{ij} = \sum_{p,q} K_D(i,j,p,q)m_{pq} \tag{1}$$

In problem (1) the set of numbers d_{ij} defines the ideal image that we want to create, the set of numbers m_{ij} describing the mask is selectable, $K_D(i,j,p,q)$ is the kernel $K_D(A,B)$ discretization. In this paper, it is proposed to use an optical scheme with a coherent light source in the absence of a lens between the mask and image, so the nature of the function $K_D(A,B)$ is significantly different from the above function $K(A,B)$ (**Fig. 1**).

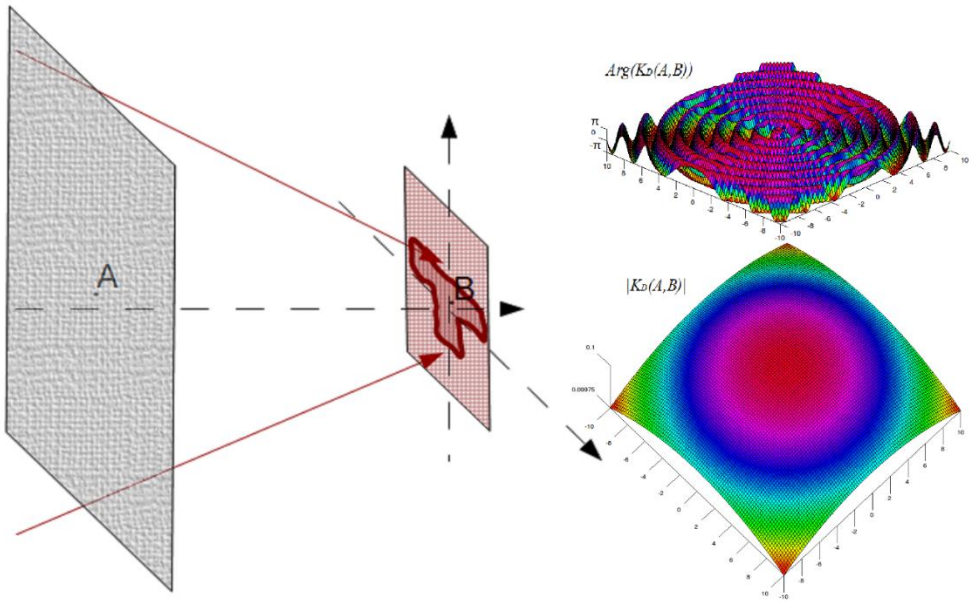


Figure. 1. Scheme of holographic lithography. The point A is on the mask, the point $\boldsymbol{B}$ **is** on the image plane. The graphs of the modulus and argument of the kernel $\boldsymbol{K_D(A,B)}$ are shown.

The image of the unit cell of the holographic mask on the plate will be blurred over the entire surface of the image. To approximate a useful image with acceptable quality by a combination of such blurred elementary images is a complicated problem, but if it is solved, the resulting holographic mask and its image quality will be stable to local defects on mask. This property of stability of image quality with respect to local defects on mask is widely known[2,5] and can become very important for applications in electronics, because, as we have already noted, it radically increases the service life of the mask, reduces the requirements for a clean room and eliminates the need for complex check and repair processes.

With a sufficiently small grid step on the hologram mask, the calculation of the useful image matrix can be reduced to the convolution of matrices:

$$\sigma = \sum_{i,j}(b_{ij}b_{ij}^{*} - d_{ij})^2 \xrightarrow[\{a_{pq}\}]{} min, \quad b_{ij} = \sum_{p,q} K_D(i-p, j-q)m_{pq}. \tag{2}$$

In the simplest case, the following compact analytic expression takes place for the gradient of the minimized quadratic difference functional

$$\nabla\sigma = 4[(B \times B^{*} - D) \times B^{*} * K]^{*},$$

where the asterisk in superscript denotes complex conjugation, the operation * is convolution, ×- cross is the Hadamard product of matrices.

3.2 Virtual object variation method

The previous formulation of the problem allows speeding up the calculations using FFT, however, it is very cumbersome in terms of the number of calculations. The dimension of the kernel matrix K_D in the case of images and masks, the dimensions of which may be of industrial interest, is too large. Another approach to image quality optimization is to implement variations of the object used in the synthesis of the hologram, so that the resulting image after recovery is as close as possible to the original object. To describe formally a mathematical problem, we introduce an operator H, which is a composition of operators S and D, which assosiates holographic image to the object used in the process of a holographic mask synthesis. Then the optimization problem can be set as follows:

$$\sigma(H(O^{*}), O) \xrightarrow[O^{*}]{} min.$$

Here O^{*} is the altered virtual object O.This formulation of the optimization problem fits into the above mathematical scheme (1). The field in the image plane restored with a hologram of one emitting element at the point A on the virtual object plane is introduced as $K_H(A, B)$ (**Fig. 2**). Such kernel, as in the case of a projection scheme with lens, will resemble the function $sin(ar)/ar$, the value of the coefficient a as well as the proximity of this function to the δ-function are determined by the angular aperture of the optical scheme, whereas in the case of projection lithography it is determined by the aperture of the lens.

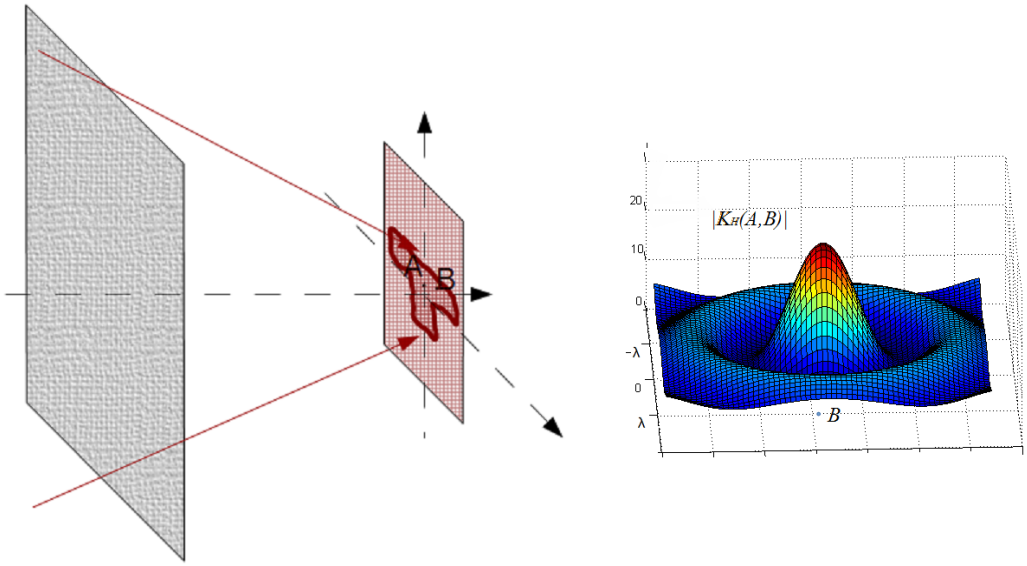


Figure. 2. Scheme of holographic lithography. Points A and B are on the image plane (virtual object is the same). The graphs of the modulus of the kernel $K_H(A, B)$ is shown.

Here it should be noted that the kernel $K_H(A,B)$ depends exactly on the difference between the coordinates of the points A and B on the image-object plane, only if the hologram is infinite. Therefore, for the physically realizable case of a finite hologram, the optimization problem takes the form

$$\sigma = \sum_{i,j}(b_{ij}b_{ij}^* - d_{ij})^2 \xrightarrow[\{a_{pq}\}]{} min, \quad b_{ij} = \sum_{p,q} K(i-p, j-q, p, q)a_{pq} \tag{3}$$

We can assume here that the matrix $\{a_{ij}\}$ describes the predistorted object O^*, and the matrix $\{d_{ij}\}$ is the ideal image of object O.

Therefore, it is possible to calculate the field in the image plane using convolutions, but only with a certain degree of accuracy. To increase the accuracy, we split an object into smaller parts and introduce several kernels describing the holographic image of one elementary pixel of objectfor each zone. The corresponding modification of the optimization problem (3) is obvious. Numerical experiments show that the picture quality is significantly improved if several convolution kernels are used.

Thus, the formulation of the optimization problem is determined by the technology used to make the mask and the choice of a variable object, namely:

1) the problem (2) corresponds to the amplitude holographic mask, which changes only the amplitude of the restoring wave passing through the plate, and $m_{ij} \in [0;1]$;
2) amplitude holographic mask with additional binary phase-shifting layer corresponds to task (2), $m_{ij} \in [-1;1]$;
3) complex hologram with variable phase and amplitude layers corresponds to the problem (2), where the values m_{ij} are selected from the unit circle on the complex plane;
4) the virtual object variations method can be implemented, which corresponds to the problem (3) with a δ-shaped kernel with values a_{pq} being chosen from the unit circle on the complex plane.

Solution algorithms based on gradient methods were implemented and tested on parallel computing systems for all variants of optimization problems.

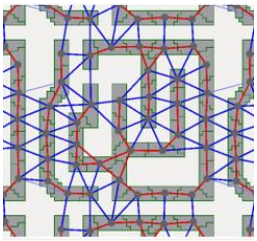
Some similar statements of optimization problems with corresponding changes for incoherent illuminator are used in projection methods[7-9].

3.3 Choice of the first approximation

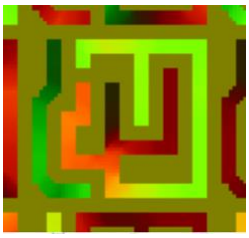
Since in real problems of microelectronics the number of elements on an object amounts to billions, the optimization problem is a problem of great computational complexity. The application of the gradient method tends to getting stuck in local extrema. Therefore, it is important to choose the initial approximation. The starting position should be as close as possible to the point of the global minimum. The technology of phase contrast allows to increase resolution in projection lithography. Some mask elements of sub-wavelength dimensions are covered with a transparent layer causing phase delay between the neighboring mask elements. As a result, it separates sub-wave elements in the image which would merge without the use of phase contrast.

In holographic lithography, this technique is accounted during the synthesis of a holographic mask. We just alter phase distribution on the elements of a virtual object whose image we want to encode on the mask (**Fig. 3**). The first step of the process is representing the virtual object as a two-level graph with colored edges. Its topology is split into sections, each of which corresponds to one node of the base level graph (**Fig. 3**, A). Blue edges connect fragments that should end with maximal possible phase contrast, red edges connect fragments that should end with minimal possible phase contrast. Modified version of breadth-first search (BFS) algorithm is used to do that. Further the phases are optimally arranged on the elements of the virtual object using the modified depth-first search (DFS) and annealing method (**Fig. 3**, B). The structure of the holographic mask does not change, the mask still consists of TZ with minimal dimensions 1.7λ, and there is no need to apply phase-shifting layers. Thus, the complex and expensive problem of local coating of a phase-shift layer

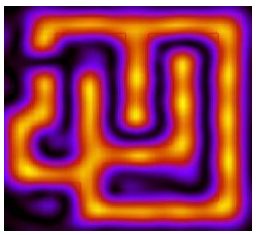
in DUVL is accounted during the holographic mask synthesis that naturally reduces the cost of the holographic mask compared to the projection one.



a. Blue edges connect elements that should end with max possible phase contrast.



b. Optimal phase distribution.



c. Reconstructed image.

Figure. 3. Analogue of phase contrast techniques in SWHL. The figures show the phase coloring of the elements of a virtual pattern.

4. SOFTWARE DESCRIPTION

Our mathematical group have developed software package for holographic masks synthesis and optimization. The simplified scheme of that package workflow is shown on the Figure 4. The first two modules may work independently. The HoloPhaseShift module produces the first approximation to the phase distribution on the object elements. The other module produces object area partition and calculates kernels described in the **Section 3.2** for each area. Then HoloOptimization procedure is starting, which consists in variations of the virtual object to obtain the highest possible quality image of the original object. Quality Functional may take into account many aspects of image quality, such as L2-norm of image intensity deviation from that of target object, deviations of the contour in photoresist, depth of focus and so on.

The software package consists of 34 modules with over 40000 strings on C++ and Python languages. All modules are scalable and can be run both on desktops and on cluster systems using OpenMP and MPI interfaces. Scalar and vector Fresnel-Kirchhoff diffraction models, two-threshold photoresist model are implemented in the software. For an example, we have shown results for the mask synthesis for an image of 1 cm by 1 cm with a resolution of 240 nm on a cluster consisting of 32 nodes, two sockets each, with Intel Xeon E5 2690v2 processors and 256GB of RAM on each node. All stages: synthesis, quality optimization and reconstruction took 2 hours. The software is currently used on the Piz Daint supercomputer at the Swiss National Supercomputer Center. The scalability of the algorithms was tested and verified.

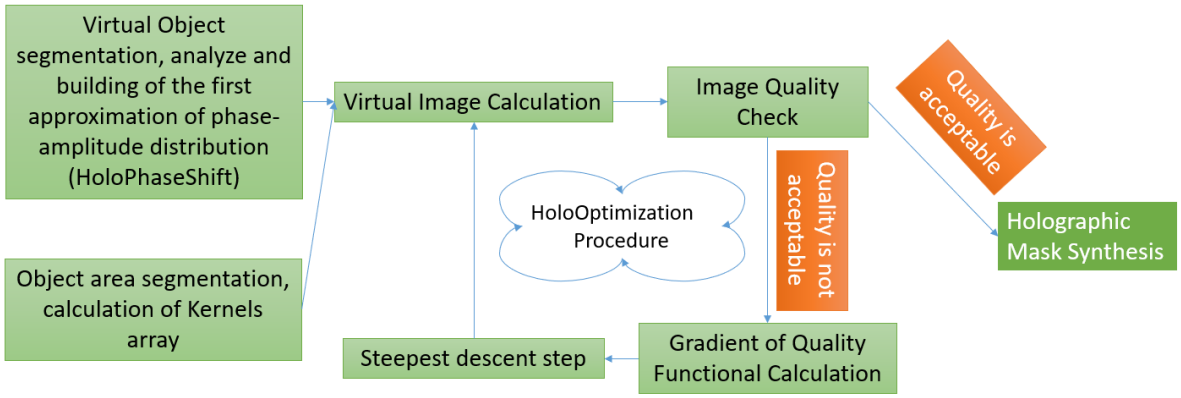


Figure. 4. Scheme of software package workflow (simplified).

The liability of the software was also verified experimentally. The lab tool “Demonstrator” was developed to demonstrate the principal capability of SWHL to generate images on non-planar surfaces and to test the software. An

example of a 3D surface was developed (**Fig. 5-6, Table 1**). The topologies are located on the top and bottom planes moved apart at distances between the planes is 100 microns. A holographic mask was generated, fabricated and exposed in the laboratory of EMPA, Dübendorf, Switzerland. The obtained experimental results have shown perfect conformity with the model.

Table 1. Synthesis parameters.

Wavelength (λ)	441.6 nm
Distance between the mask and the wafer	88.3 mm
Mask shape and size	Ø 21.75 mm
Mask grid step	2 μm (4.53 λ)
Mask type	Binary Amplitude
Min. size of opening	750 nm (1.7 λ)
Shift from optical axis	4,25 mm left along X-axis
Image size	2.5x2.5 mm^2
Depth of Focus	3 μm
Distance between planes	100 μm
CD	1 μm

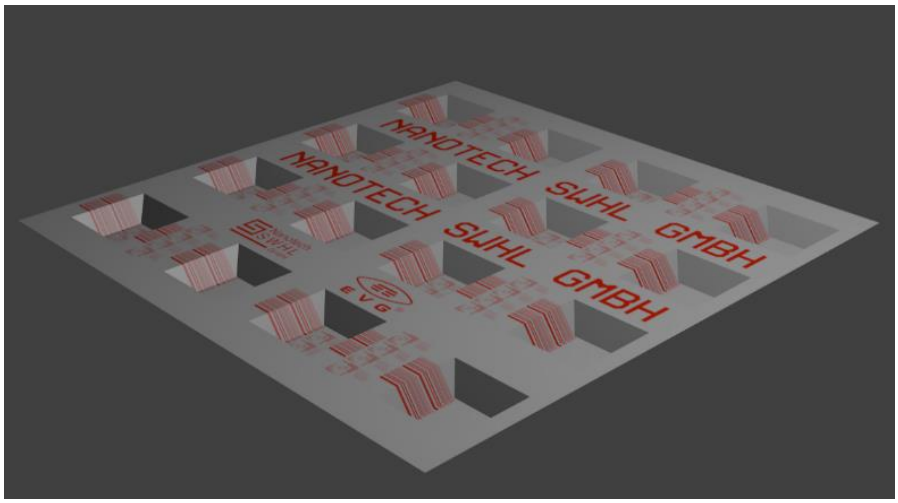


Figure 5. Piecewise flat surface and the pattern design.

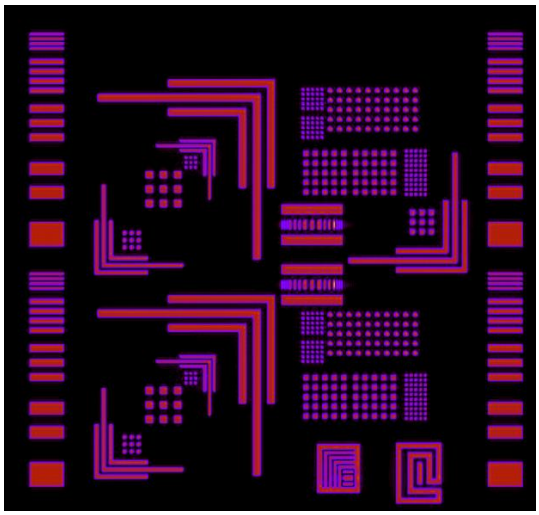


Figure 6. Fragment of the reconstructed image (computer simulation).

5. CONCLUSIONS

SWHL Nanotech team have implemented effective algorithms and developed scalable software allowing to synthesize holographic masks for various lithography applications including MEMS, MOEMS and high-end IC production. Most of the technical problems of state-of-art projection photolithography such as 3D-imaging and quality optimization were stated and solved as a completely numerical problems in the case of holographic lithography approach. Thus, today it is possible to use modern computing clusters for the synthesis of holographic masks and to implement them in inexpensive and sustainable devices for holographic photolithography.

REFERENCES

[1] Mack, C. A., [Fundamental Principles of Optical Lithography: The Science of Microfabrication], John Wiley & Sons, 417 (2007)

[2] Borisov M. V., Chelyubeev D. A., Chernik V. V., Gavrikov A. A., Knyazkov D. Yu., Mikheev P. A., Rakhovskiy V. I., Shamaev A. S., "Analysis of an Effect of Perturbations in SWHM and Illuminating Optical Scheme Parameters on an Aerial Image," Proc. of Advanced Semiconductor Manufacturing Conference, 23rd Annual SEMI, 165-169 (2012)

[3] Borisov M. V., Chelyubeev D. A., Chernik V. V., Gavrikov A. A., Knyazkov D. Yu., Mikheev P. A., Rakhovskiy V. I., Shamaev A. S., "Phase-shift at subwavelength holographic lithography (SWHL)," Proc. SPIE 8352, 28th European Mask and Lithography Conference, 83520P (2012)

[4] Borisov M. V., Chelyubeev D. A., Chernik V. V., Mikheev P. A., Rakhovskiy V. I., Shamaev A. S., "Experimental verification of sub-wavelength holographic lithography physical concept for single exposure fabrication of complex structures on planar and nonplanar surfaces," Proc. SPIE 10446, 33rd European Mask and Lithography Conference, 104460X (2017)

[5] Harry J. Levinson, H. J., [Principles of Lithography], Bellingham, USA: SPIE Press, 330-334 (2010).

[6] Gabor D. A., "New Microscopic Principle," Nature 161, 777–778 (1948)

[7] Poonawala, A., and Milanfar, P., "Mask Design for Optical Microlithography – An Inverse Imaging Problem," IEEE Transactions on Image Processing 16 (3), 774–788 (2007)

[8] Jia, N., and Lam, E.Y., "Pixelated source mask optimization for process robustness in optical lithography," Optics Express 19(20), 19384–19398 (2011)

[9] Robert Socha, Xuelong Shi, and David LeHoty "Simultaneous source mask optimization (SMO)", Proc. SPIE 5853, Photomask and Next-Generation Lithography Mask Technology XII, (28 June 2005)