

УДК 573.876.4

ЭФФЕКТИВНЫЕ МЕТОДЫ СИНТЕЗА И ОПТИМИЗАЦИИ ГОЛОГРАММНОЙ МАСКИ

© 2023 г. В. В. Черник^{1,*}

Представлено академиком РАН Ф.Л. Черноусько 03.04.2023 г.

Поступило 18.04.2023 г.

После доработки 18.04.2023 г.

Принято к публикации 07.05.2023 г.

Описаны эффективные алгоритмы синтеза голографических масок на основе БПФ со сложностью $O(M \ln N)$, где N — количество элементов изображаемого объекта. Разработан и внедрен масштабируемый программный комплекс, позволяющий синтезировать голограммные маски для различных приложений литографии, включая производство МЭМС, МОЕМС и микросхем высокого класса. Предложены решения проблем оптимизации качества голографических изображений и повышения дифракционной эффективности.

Ключевые слова: дифракция, компьютерно-синтезированные голограммы, оптимизация, фотолитография

DOI: 10.31857/S2686740023050036, **EDN:** OWHLPJ

Проекционная фотолитография сегодня является основным инструментом производства микроэлектронных изделий. Постоянное повышение требований к разрешению сделало технологический процесс фотолитографии очень чувствительным к внешним возмущениям и увеличило стоимость изготовления проекционной маски. По мере приближения к оптическому пределу разрешения дифракционные эффекты все интенсивнее негативно влияют на качество проекционного изображения. Это связано с тем, что исторически технология проекционной фотолитографии развивалась на основе принципов геометрической оптики. Методы повышения разрешающей способности (resolution enhancement techniques, RET), такие как OPC, PSM, SMO, по существу являются методами борьбы с дифракционными эффектами. Они значительно увеличивают стоимость маски и усложняют процесс изготовления [1–3]. Моделирование и проектирование RET стали весьма затратными теоретическими и, что даже более важно, инженерными задачами, направленными на борьбу с проявлениями волновой природы света (рассеяние, эффекты близости, явление Гиббса). Стоимость проекционных масок и соответствующего оптического оборудования непропорционально увеличивается по ме-

ре миниатюризации микроэлектронных элементов.

В последние годы интенсивно исследуется голографический подход к получению изображений с высоким разрешением. Группа авторов развивают технологию субволновой голографической литографии (СВГЛ) [1–3]. Эта технология отвечает нескольким вызовам, возникшим в индустрии микроэлектроники за последние десятилетия. Основная идея состоит в том, чтобы использовать волновые свойства света с помощью голографических методов, а не подавлять их. Многие инженерные задачи при этом перемещаются в область моделирования и численных расчетов. Были разработаны эффективные алгоритмы компьютерного синтеза голографических масок на основе БПФ со сложностью $O(M \ln N)$, где N — количество элементов изображения, что позволило синтезировать голографические маски для фотолитографии слоев современных микроэлектронных устройств.

Голограммная маска — это физическая реализация функции амплитудной, фазовой или амплитудно-фазовой модуляции световой волны. Эта функция может быть получена как интенсивность интерференционной картины опорных и объектных волн, в соответствии с классическим принципом, предложенным Д. Габором. Согласно этому принципу, объект освещается некоторым световым пучком, образуя объектную волну (рис. 1а). В то же время опорный световой пучок, взаимно когерентный с объектным, образует

¹Институт проблем механики им. А.Ю. Ишлинского
Российской академии наук, Москва, Россия

*E-mail: gungho424@gmail.com

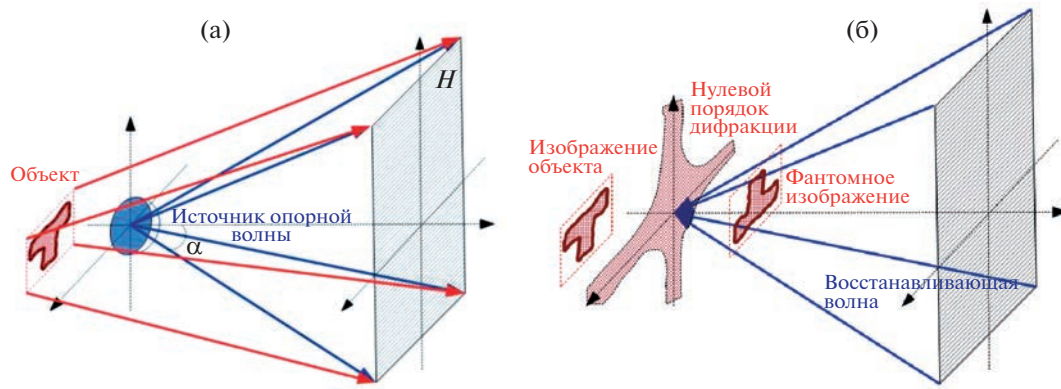


Рис. 1. Голографический подход к получению изображения объекта.

опорную волну [4, 5]. Результат интерференции этих волн записывается на плоскости H . Полученную интерференционную картину можно понимать как функцию амплитудной модуляции $T: H \rightarrow [0; 1]$. Если на место плоскости H установить дифракционный элемент, реализующий полученную функцию амплитудной модуляции для восстанавливающей волны, которая является обратной к опорной волне, то на месте объекта будет восстановлено поле, повторяющее поле объектной волны, т.е. голографическое изображение объекта (рис. 1б). Простейшая конструкция голографической маски представляет собой непрозрачный экран с отверстиями, где функция амплитудной модуляции аппроксимируется плотностью и размером отверстий. Такая конструкция, однако, обладает рядом недостатков, главным из которых является наличие высокоинтенсивного нулевого порядка дифракции. Опишем некоторые типы конструкций голограммных масок и укажем на их особенности и преимущества.

Пусть $O: H \rightarrow \mathbb{C}$ — комплексная амплитуда опорной волны на плоскости голограммы, $\Pi: H \rightarrow \mathbb{C}$ — комплексная амплитуда объектной волны. Тогда рассмотрим функцию

$$T = \frac{|O + \Pi|^2}{|O|^2}, \quad (1)$$

принимаящую действительные неотрицательные значения на области голограммы H как функцию амплитудной модуляции. Применив ее к обратной волне O^* , получим

$$TO^* = \frac{(|O|^2 + |\Pi|^2)O^*}{|O|^2} + \frac{\Pi O^{*2}}{|O|^2} + \Pi^*. \quad (2)$$

Результирующее поле раскладывается в три компоненты:

1) Π^* — поле обращенной объектной волны, которая создает изображение объекта;

2) $\frac{\Pi O^{*2}}{|O|^2}$ — поле искаженной волны, псевдосимметричной объектной волне относительно оптической оси, создающей искаженное изображение объекта (фантомное изображение);

3) $\frac{(|O|^2 + |\Pi|^2)O^*}{|O|^2}$ — нулевой порядок дифракции, фокусирующийся в окрестности фокуса в случае расходящейся опорной волны, либо создающей фоновый шум в случае плоской опорной волны.

Отметим, что для физической реализации голограммной маски потребуются ввести также нормировочный коэффициент M такой, что $\frac{T}{M}: H \rightarrow [0; 1]$.

Нулевой порядок дифракции концентрирует существенную часть энергии восстанавливающей волны, что негативно влияет на дифракционную эффективность схемы. Также излучение нулевого порядка, рассеянное элементами оптической схемы, является существенным источником шума, который вследствие когерентности используемого излучения обнаруживает себя в виде фазового спекл-шума, разрушающего качество изображения [5]. Чтобы уменьшить степень влияния нулевого порядка дифракции на изображение, используются оптические схемы со сходящимся восстанавливающим пучком, в которых нулевой порядок фокусируется и полезное изображение возможно разделить в пространстве с основным потоком компоненты нулевого порядка. Преимуществом такой голограммы является простота изготовления ее физической реализации в виде бинарной амплитудной маски, в которой амплитудная модуляция реализуется с помощью отверстий

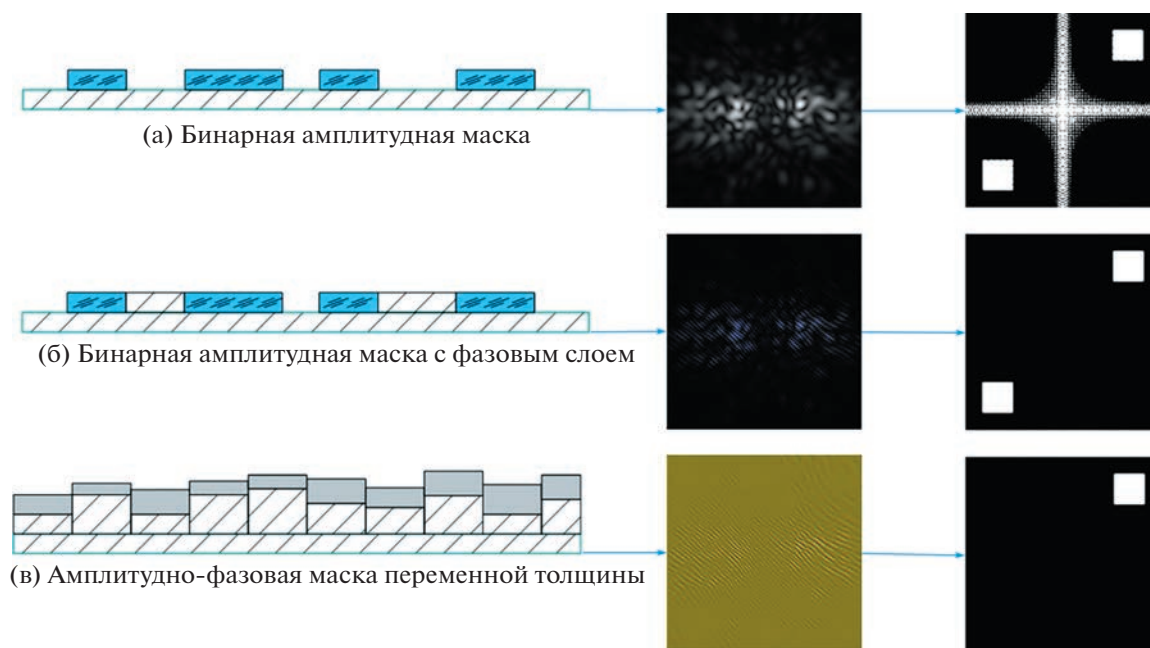


Рис. 2. Три типа конструкции голограммной маски. Соответствующая дифракционная картина при восстановлении: а — с помощью амплитудной маски состоит из нулевого порядка дифракции, изображения объекта и фантомного изображения; б — с помощью амплитудной маски с фазовым слоем состоит из изображения объекта и фантомного изображения; в — в случае амплитудно-фазовой маски это только изображение объекта.

переменной величины в непрозрачном экране (рис. 2а). Размер отверстий выбирается таким образом, чтобы локально в данной области реализовать требуемую амплитудную модуляцию.

Обратим внимание на то, что компонента функции T , отвечающая за нулевой порядок дифракции, является действительно-значной функцией. Рассмотрим следующую функцию модуляции:

$$T_1 = \frac{|O + \Pi|^2 - (|O|^2 + |\Pi|^2)}{M_1 |O|^2} = \frac{O^* \Pi + O \Pi^*}{M_1 |O|^2}. \quad (3)$$

Здесь T_1 принимает действительные значения, а нормировочный коэффициент M_1 выбран таким образом, чтобы $T_1 : H \rightarrow [-1; 1]$. Применяя такую функцию модуляции к O^* аналогично (2), получим

$$T_1 O^* = \frac{\Pi O^{*2}}{|O|^2} + \Pi^*,$$

т.е. восстанавливающая волна под действием такой функции модуляции будет иметь только две компоненты, соответствующие изображению объекта и фантомному изображению. Для физической реализации требуется накладывать дополнительный фазовый слой, поворачивающий фазу волны на 180° на элементах маски, соответствующих отрицательным значениям функции модуля-

ции (рис. 2б). Существенным минусом также является сдвиг среднего уровня пропускания маски к нулю. То есть если среднее значение функции T после нормировки равно 0.5, то среднее значение функции T_1 оказывается нулевым, причем большая часть маски оказывается темной. Маски такого типа обладают серьезным преимуществом за счет практически полного отсутствия нулевого порядка дифракции. Дифракционная эффективность такой маски существенно выше амплитудной (до 4 раз). Качество изображения также заметно повышается за счет лучшего соотношения сигнал—шум. Однако даже с учетом дополнительных процедур просветления для изготовления такой маски требуется большое количество элементов малого размера в диапазоне $(0; 0.25\lambda]$, где λ — длина волны опорного и восстанавливающего излучения [3]. Это сказывается не только на стоимости маски, но и на устойчивости к ошибкам моделирования и дефектам изготовления таких элементов. Для достоверного моделирования прохождения света через такие отверстия требуется точная модель дифракции.

Также возможно рассматривать и функцию модуляции T_2 , в которой присутствует только компонента, отвечающая за изображение объекта, т.е. комплексно-значную функцию модуляции

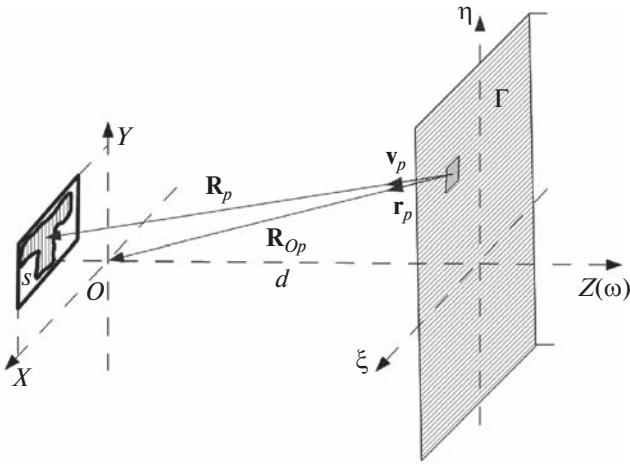


Рис. 3. Схема голографического подхода к получению изображений.

$$T_2 = \frac{OP^*}{M_2 |O|^2}. \quad (4)$$

Для физической реализации такой функции на голограммной маске требуются амплитудный и фазовый слои переменной толщины (рис. 2в). Также необходима строгая модель дифракции для моделирования прохождения световой волны через слоистую среду с переменным показателем преломления.

Рассмотрим вопрос качества создаваемого изображения. Основными объектами описанных выше методов восстановления голографического изображения, основанных на принципе Д. Габо-ра, являются опорная волна O , объектная или предметная волна Π и голограмма, математическим представлением которой является функция модуляции T , T_1 или T_2 в зависимости от предполагаемой структуры маски. Будем далее обозначать функцию модуляции буквой T , подразумевая один из вариантов. Введем оператор D , который функции модуляции T ставит в соответствие восстановленное голографическое изображение I , т.е. $I = D(T)$. Дефекты изображения обычно проявляются в виде неравномерной яркости разных элементов, колебаний яркости в пределах одного элемента, а также в виде недостаточно высокого контраста, что в результате экспонирования маски из фоточувствительного материала при производстве микрoeлектронных компонентов может приводить к неровности краев элементов (LER, LWR [6]) отпечатков в слое фоторезиста. Степень дефектности голографического изображения I можно выразить через отличие от некоторого идеального изображения A по некоторой метрике σ , в которой могут быть отражены различные аспекты качества от равномерности интенсивности

засветки по всему полю изображения до локальных дефектов итогового отпечатка в фоторезисте. Назовем эту метрику функционалом качества. Задачей оптимизации является минимизация этих отличий и дефектов изображения. Рассмотрим постановку задачи оптимизации качества изображения за счет вариаций функции модуляции T , т.е.

$$\sigma(A, D(T)) \rightarrow \min. \quad (5)$$

Для решения такой задачи численными методами разобьем голограмму прямоугольной сеткой на ячейки малого размера и представим функцию модуляции в виде $T(\xi, \eta) = \sum_p \alpha_p \psi_p(\xi, \eta)$ (рис. 3).

Здесь α_p — усредненный коэффициент модуляции p -й ячейки, $\psi_p(\xi, \eta)$ — функция-индикатор p -й ячейки, равная нулю вне ячейки и единице внутри, (ξ, η) — координаты на плоскости голограммы.

Оператор D , опираясь на подход Френеля—Кирхгофа, можно представить в виде интеграла

$$D(T) = \iint_{\Gamma} K(x - \xi, y - \eta, d) O(\xi, \eta) T(\xi, \eta) d\xi d\eta \approx \sum_p \alpha_p D_p(x, y), \quad (6)$$

где $O(\xi, \eta)$ — комплексная амплитуда поля восстанавливающей волны на плоскости голограммы,

$$K(x - \xi, y - \eta, d) = \frac{de^{-ikR_p}}{R_p^2}, k = \frac{2\pi}{\lambda}, R_p = (x - \xi, y - \eta, d),$$

Γ — область голограммы. С учетом представления функции модуляции $T(\xi, \eta)$ в виде линейной комбинации характеристических функций ячейки данный интеграл распадается в соответствующую линейную комбинацию интегралов по ячейкам, каждый из которых с помощью приближения дальней зоны [7] и в предположении о сферической восстанавливающей волне можно представить как

$$D_p(x, y) \approx \frac{2ik \exp(ik(R_p + R_{op}))}{\pi R_p^2} \times \frac{\sin(k(r_{p1} + v_{p1})\Delta) \sin(k(r_{p2} + v_{p2})\Delta)}{k(r_{p1} + v_{p1}) k(r_{p2} + v_{p2})},$$

где Δ — шаг сетки на голограмме, (v_{p1}, v_{p2}) и (r_{p1}, r_{p2}) — координаты единичных направляющих векторов $\mathbf{v}$ и $\mathbf{r}$ по осям X и Y , начало координат системы OXY — фокус восстанавливающей сферической волны.

Введем также сетку на плоскости изображения. Тогда интенсивность идеального изображения объекта A можно представить как набор значений $\{a_q \in [0; 1], q = 1, \dots, N\}$ в узлах данной сетки.

Назовем число N узлов сетки на изображении, достаточное для того, чтобы передать информацию об изображаемом объекте при дискретизации, сложностью объекта. Тогда задачу в дискретном виде можно представить в следующем виде:

$$\sigma = \sum_q (b_q b_q^* - c a_q)^2 \xrightarrow{\{\alpha_p\}, c} \min, \quad (7)$$

$$b_q = \sum_p K_D(p, q) \alpha_p.$$

Здесь в роли функционала качества выбрано отклонение интенсивности восстановленного изображения от идеального распределения интенсивности по l_2 -норме в узлах вычислительной сетки, коэффициент c имеет смысл нормировочного коэффициента, отвечающего за то, что восстановленное изображение должно быть близко к идеальному по форме. Коэффициент c подлежит вариации в процессе оптимизации. Матрица $K_D(p, q)$ является матрицей локальных откликов системы, т.е. $K_D(p, q) = D_p(x_q, y_q)$, где (x_q, y_q) — координаты соответствующего узла на плоскости изображения. Данная матрица имеет размерность $N \times M$, где N — сложность изображаемого объекта, а M — количество ячеек на голограмме, причем она не является разреженной, т.е. величины ее элементов равномерно распределены по всей матрице. Сложность голограммной маски соответствует сложности изображения, поэтому асимптотически вычислительная сложность такой задачи $O(N^2)$. Алгоритм оптимизации для данной постановки задачи был реализован [8] методом градиентного спуска, причем голограмма Габора (1) была использована в качестве первого приближения. Было показано, что качество изображения, восстановленное с помощью компьютерно-синтезированной голограммы, возможно существенно улучшить. В результате 10 итераций метода градиентного спуска l_2 -норма отклонения интенсивности некоторых изображений от идеального была улучшена более чем в два раза по сравнению с первым приближением. Однако в связи с квадратичной зависимостью вычислительной сложности задачи оптимизации от сложности изображаемого объекта в постановке (7) такой метод оптимизации возможно реализовать только для малых объектов. Реальные объекты, фотолитография которых представляет промышленный интерес, являются слоями микроэлектронных компонентов и могут иметь сложность $\sim 10^{12}$ и более.

Введем понятие виртуального объекта $V: S \rightarrow \Omega$, S — область, в которой определен объект (рис. 3), Ω — круг единичного радиуса на комплексной плоскости. Также введем оператор G , который каждому виртуальному объекту V ставит в соответствие функцию модуляции T голограммной

маски, полученной одним из описанных выше способов, т.е. $T = G(V)$. Фактически оператор G является композицией нелокального оператора типа (6), ставящего в соответствие каждому виртуальному объекту с подсветкой V распределение комплексной амплитуды предметной волны на области голограммы, и одного из трех операторов (1), (3) или (4), ставящего в соответствие каждой предметной волне некоторую функцию модуляции голограммной маски. Оператор T так же, как и оператор D , введенный ранее и каждой функции модуляции T ставящий в соответствие восстановленное голографическое изображение, имеет нелокальный характер. То есть каждый элемент виртуального объекта вносит вклад в определение функции модуляции T в каждой точке голограммы. Рассмотрим композицию операторов $H(V) = D(G(V))$. Интегральное представление оператора H будет иметь вид

$$H(V) = \iint_S K_H(x - x_0, y - y_0, x_0, y_0) V(x_0, y_0) dx_0 dy_0, \quad (8)$$

где (x_0, y_0) и (x, y) — декартовы координаты в области изображения, $K_H(\Delta x, \Delta y, x_0, y_0)$ — ядро оператора, описывающее влияние изменения значения комплексной амплитуды волны подсветки виртуального объекта в точке (x_0, y_0) на комплексную амплитуду восстанавливающей волны в точке (x, y) . Опишем некоторые качественные свойства ядра K_H . Оно зависит от оптической схемы и от выбора восстанавливающей волны. Его абсолютная величина быстро убывает при увеличении $r = \sqrt{\Delta x^2 + \Delta y^2}$, т.е. оператор H , в отличие от операторов D и G , композицией которых он является, имеет локальный характер. Для вычислительных задач имеет смысл ограничить носитель функции K_H по $(\Delta x, \Delta y)$ областью $r < 5\lambda$, а значениями K_H при $r \geq 5\lambda$ пренебречь. Наконец, если восстанавливающая волна $O(\xi, \eta)$ является идеально плоской, а область голограммы бесконечна по всем направлениям, то $K_H = K_H(\Delta x, \Delta y)$ не зависит от точки (x_0, y_0) . В практически значимом случае конечной голограммы и сходящейся восстанавливающей волны ядро K_H зависит от точки (x_0, y_0) , однако эта зависимость проявляет себя только на расстояниях порядка $10^4 \lambda$. То есть для вычислительных целей в пределах подобласти размером порядка $10^3 \lambda$ можно считать ядро K_H зависящим только от $(\Delta x, \Delta y)$. Это наблюдение дает возможность разбить область виртуального объекта на подобласти A_n , внутри каждой из которых оператор H представляется как оператор свертки с некоторым ядром K_H^n , являющимся приближением ядра K_H в данной подобласти.

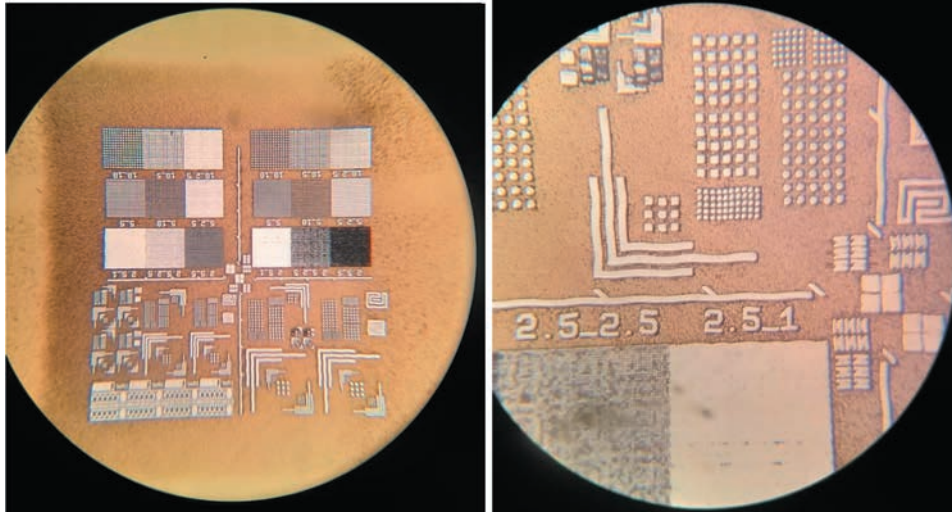


Рис. 4. Результат проявки экспозиции тестового голографического изображения $2.5 \times 2.5 \text{ мм}^2$ в фоточувствительном материале толщиной 30 мкм ($\text{NA} = 0.24$; $\lambda = 441.6 \text{ нм}$; разрешение 5 мкм).

Рассмотрим теперь задачу оптимизации качества изображения аналогично (5), однако объектом вариаций будем считать виртуальный объект:

$$\sigma(A, H(V)) \xrightarrow{V} \min.$$

Введем равномерную прямоугольную сетку с шагом Δ на каждой подобласти S_n . Тогда идеальное изображения объекта A можно представить как набор значений $\{a_{pq} \in [0;1]\}$ в узлах данной сетки. Виртуальный объект V также можно представить двумерным массивом $\{v_{pq}\}$, где каждое значение v_{pq} — это некоторое комплексное число из единичного круга с центром в нуле, а функционал качества как сумму ограничений данного функционала на каждую из подобластей $\sigma = \sum \sigma_n$. На практике не требуется находить глобальный минимум функционала качества. Достаточно достичь такого уровня качества, что выполняются заранее определенные допуски. Поэтому в каждой подобласти можно поставить задачу следующим образом. Найти $\{v_{pq} \in \mathbb{C}: |v_{pq}| \leq 1\}$ и действительную константу c , такие что

$$\sigma_n = \sum_{p,q} (b_{pq} b_{pq}^* - c a_{pq})^2 < \varepsilon,$$

где $b_{pq} = \sum_{lm} K_H((p-l)\Delta, (q-m)\Delta) v_{lm}$.

Для решения задачи синтеза голограмм на языке C++ с использованием библиотек MPI и OpenMP был разработан программный комплекс для кластерных систем. В нем реализованы алгоритмы оптимизации, описанные выше, а также аналогичные алгоритмы для векторной модели дифракции [9]. Вычисления распараллелены таким образом, что каждая подобласть S_n обрабаты-

вается на отдельном наборе вычислительных узлов методом градиентного спуска. Все градиенты и поля представлены с помощью операторов типа свертки, поэтому для ускорения вычислений используются быстрое преобразование Фурье и библиотека fftw3. Вычислительная сложность алгоритмов таким образом равна $O(N^* \ln N)$.

Независимость определений виртуального объекта и идеального изображения в предложенном подходе дает дополнительную гибкость и позволяет универсальным образом ставить более сложные задачи синтеза и оптимизации голограмм. Например, можно поставить задачу оптимизации качества изображения в нескольких параллельных плоскостях. То есть

$$\sigma(A, H(V)) = \sum_n \sigma_n(A_n, H(V)) \xrightarrow{V} \min,$$

где A_n — сечения идеального изображения плоскостями, σ_n — ограничение функционала качества на данную плоскость. Численные решения задач оптимизации в некоторых из таких постановок были реализованы в программном комплексе. Это позволяет решать задачи синтеза голограммных масок для сложных трехмерных изображений, а также задачу оптимизации глубины фокуса оптической системы. Для верификации программного комплекса и математических моделей, используемых в его основе, была проведена серия экспериментов [10, 11], в ходе которых голограммные маски синтезировались с его помощью, изготавливались на электронном литографе, устанавливались в оптической схеме с заранее спроектированными для засветки оптикой и осветителем. Одна из последних серий экспериментов была посвящена получению изображений

в толстых (10–50 мкм) фоторезистах при длине волны восстанавливающего излучения 441.6 нм (рис. 4).

СПИСОК ЛИТЕРАТУРЫ

1. Борисов М.В., Боровиков В.А., Гавриков А.А., Князьков Д.Ю., Раховский В.И., Челюбеев Д.А., Шамаев А.С. Методы создания и коррекции качества голографических изображений геометрических объектов с элементами субволновых размеров // ДАН. 2010. Т. 434. № 3. С. 332–336.
2. Rakhovsky V., Knyazkov D., Shamaev A., Chernik V., Gavrikov A., Chelyubeev D., Mikheev P., Borisov M. // Proc. European Mask and Lithography Conference. SPIE. 2012. V. 8352. 83520P.
3. Черник В.В. // Журнал Радиоэлектроники. 2017. № 1. С. 1–20.
4. Gabor D.A. // Nature. 1948. № 161. P. 777–778.
5. Колфилд Г. Оптическая голография. М.: Мир, 1982. 376 с.
6. Mack C.A. Fundamental Principles of Optical Lithography: The Science of Microfabrication.: John Wiley & Sons, 2007. 417 p.
7. Басс Ф.Г., Фукс И.М. Рассеяние волн на статистически неровной поверхности. М.: Наука, 1972. 424 с.
8. Черник В.В. // Труды VI международной конференции “Параллельные вычисления и задачи управления”. РАСО’2012. Т. II. 2012.
9. Михеев П.А. // Вестн. Моск. ун-та. Сер. 15: Вычислительная математика и кибернетика. 2014. Т. 1. С. 15–22.
10. Borisov M., Chernik V., Merkushev L., Shamaev A., Rakhovskiy V., Chelubeev D. // Proc. SPIE 11324. Novel Patterning Technologies for Semiconductors. MEMS/NEMS and MOEMS. 2020. XXXIV. 113241J (2020) <https://doi.org/10.1117/12.2552013>
11. Borisov M., Chernik V., Shamaev A., Rakhovskiy V., Chelubeev D. // Proc. SPIE 11324. Novel Patterning Technologies for Semiconductors, MEMS/NEMS and MOEMS. 2020. XXXIV; 1132417 (2020). <https://doi.org/10.1117/12.2551936>

EFFECTIVE METHODS OF SYNTHESIS AND OPTIMIZATION OF A HOLOGRAPHIC MASK

V. V. Chernik^a

^a*Ishlinsky Institute for Problems in Mechanics of the Russian Academy of Sciences, Moscow, Russia*

Presented by Academician of the RAS F.L. Chernous'ko

The paper presents several statements of the problems of synthesis of holographic mask in the form of optimization problems for quality of holographic images. An effective algorithm for the synthesis of holographic masks based on FFT with complexity $O(N \ln N)$, where N is the number of elements of the depicted object, is described. Based on this algorithm, a scalable software package has been developed and implemented that allows synthesizing holographic masks for various lithography applications, including the production of MEMS, MOEMS and high-end chips. Experimental results are presented.

Keywords: diffraction, computer-generated holography, optimization problems, photolithography