ФЕДЕРАЛЬНОЕ АГЕНТСТВО ПО ОБРАЗОВАНИЮ

ГОСУДАРСТВЕННОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ
«САМАРСКИЙ ГОСУДАРСТВЕННЫЙ АЭРОКОСМИЧЕСКИЙ
УНИВЕРСИТЕТ имени академика С.П. КОРОЛЕВА»

В.В. КОТЛЯР, Л.Л. ДОСКОЛОВИЧ,
В.А. СОЙФЕР

ИТЕРАТИВНЫЕ МЕТОДЫ ПРОЕКТИРОВАНИЯ ПЛОСКОЙ ОПТИКИ

Утверждено Редакционно-издательским советом университета
в качестве учебного пособия

САМАРА
Издательство СГАУ
2006

УДК 535.4, 535.8
ББК 22.343
К 73

Инновационная образовательная программа "Развитие центра компетенции и подготовка специалистов мирового уровня в области аэрокосмических и геоинформационных технологий"

Рецензенты: д-р физ.-мат. наук, проф. В. В. Ивахник,
д-р физ.-мат. наук, проф. И. П. Завершинский

Котляр В.В.

К 73 **Итеративные методы проектирования плоской оптики:** учеб. пособие / *В.В. Котляр, Л.Л. Досколович, В.А. Сойфер*, – Самара: Изд-во Самар. гос. аэрокосм. ун-та, 2006. – 168 с.:58 ил.

ISBN 5-7883-0448-2

В данном пособии рассмотрены итеративные методы решения обратных задач скалярной параксиальной теории дифракции в применении к проектированию фазовых оптических элементов плоской оптики. Подробно описаны параметрические итеративные алгоритмы минимизации функционального критерия: алгоритм входа-выхода, адаптивно-аддитивный алгоритм, градиентный алгоритм. Рассмотренные алгоритмы применены для расчета дифракционных оптических элементов, формирующих заданное распределение интенсивности лазерного света в области наблюдения. Уделено внимание доказательству релаксации итеративных алгоритмов, то есть уменьшению среднеквадратичного критерия с ростом числа итераций. Обсуждается важность выбора начального приближения для успешной сходимости алгоритмов. Приведены конкретные примеры расчета фокусаторов лазерного излучения (фокусировка в круг, квадрат, кольцо), многофокусных линз, аксиконов.

УДК 535.4, 535.8
ББК 22.343

ISBN 5-7883-0448-2

© Котляр В.В., Досколович Л.Л.,
Сойфер В.А., 2006
© Самарский государственный
аэрокосмический университет, 2006

Оглавление

Введение 4
1. Алгоритм уменьшения ошибки 7
2. Алгоритм входа-выхода 12
3. Адаптивно-аддитивный алгоритм 15
4. Адаптивно-мультипликативный алгоритм 27
5. Адаптивно-регуляризационный алгоритм 34
6. Градиентный алгоритм расчета фазы ДОЭ 41
7. Применение итеративных алгоритмов для расчета ДОЭ 46
7.1. Расчет ДОЭ, фокусирующих в радиально-симметричные области Фурье-спектра 47
7.2. Расчет дифракционных аксиконов, формирующих осевые световые отрезки 53
7.3. Расчет радиально симметричных ДОЭ с квантованной фазой 58
7.4. Многопорядковые фазовые бинарные дифракционные решетки 63
7.5. Многоуровневые фазовые дифракционные решетки 75
7.6. Фазовые ДОЭ, фокусирующие в объем и на поверхность тел вращения 85
7.7. Фокусировка гауссового пучка в квадрат 102
7.8. Фокусировка в кольцо 114
7.9. Композиционные ДОЭ, формирующие контурные изображения 125
7.10. Квантованные ДОЭ для фокусировки в заданную двумерную область 134
7.11. Квантованные ДОЭ для формирования амплитудно-фазовых распределений 157
Заключение 164
Список специальных терминов 166
Список контрольных вопросов 167
Литература 168

Введение

В задачах обработки и восстановления изображений разработаны и широко применяются итерационные (итеративные) методы. Обзор большинства из них можно найти в /1/. Итеративные методы можно разделить на несколько классов: 1) методы на основе сжимающих или нерасширяющих операторов: алгоритм Бургера — Ван Циттерта, алгоритм сверхразрешения Герчберга — Папулиса; 2) методы на основе проекций на выпуклые множества — алгоритмы Старка и Юлы; 3) итеративные алгоритмы с псевдодифференциальными операторами ограничений: алгоритм Герчберга — Секстона и алгоритм Фиенапа; 4) методы нелинейного программирования или оптимизации: алгоритм наискорейшего спуска, алгоритм сопряженных градиентов, метод модифицированных функций Лагранжа.

Применение этих итеративных методов для решения задач расчета ДОЭ требует специального исследования. Это связано с тем, что, как правило, задачи восстановления искаженных изображений сводятся к решению линейного интегрального уравнения типа свертки с аддитивным шумом. А задачи синтеза ДОЭ сводятся к решению нелинейного интегрального уравнения, нелинейность которого связана с операцией взятия модуля от комплексной амплитуды когерентного света.

В этом пособии рассматриваются итеративные методы решения обратных задач скалярной теории дифракции применительно к синтезу оптических элементов. Решения, которые получаются с помощью итеративных методов, являются квазиоптимальными, так как они приводят к достижению локального минимума функционала-критерия или целевой функции. Да и сами итеративные методы получаются в результате решения вариационной задачи на минимум или максимум целевой функции. В качестве такой функции, как правило, используется среднеквадратичное отклонение заданной амплитуды светового поля в некоторой плоскости пространства от рассчитанной. Иногда вместо амплитуд сравниваются интенсивности, а вместо среднеквадратичного критерия выбираются критерии более высокого порядка. Итеративные

методы, которые используются в данном пособии, можно разделить на две группы: параметрические и градиентные. В параметрических алгоритмах один или два параметра, от которых зависит скорость сходимости алгоритма, остаются постоянными в течении нескольких итераций. В градиентных алгоритмах (сопряженного градиента или наискорейшего спуска) оптимальное значение шага вычисляется на каждой итерации. Кроме известных однопараметрических методов расчета ДОЭ, рассматриваются также двухпараметрические алгоритмы, полученные на основе минимизации функционала-критерия с регуляризирующим слагаемым. Пособие построено следующим образом. В разделах 1 — 6 описаны общие положения, касающиеся различных типов итеративных алгоритмов, применяемых для расчета ДОЭ в скалярном приближении. Это непараметрический алгоритм сокращения ошибки (раздел 1), однопараметрические итеративные алгоритмы (разделы 2 — 4), регуляризационный двухпараметрический метод (раздел 5) и градиентные алгоритмы расчета ДОЭ (раздел 6). В большом разделе 7 показаны многочисленные примеры расчета различных типов ДОЭ. Особенностью расчета ДОЭ с радиально-симметричной фазой (раздел 7.1) является использование преобразования Ханкеля. Задача расчета ДОЭ, формирующего продольные световые отрезки (раздел 7.2) сведена к одномерной задаче фокусировки в поперечный световой отрезок. Особый алгоритм получается при расчете радиально-симметричных ДОЭ, апертура которых разделена на несколько кольцевых субапертур (раздел 7.3). Применение градиентных и параметрических итеративных алгоритмов для расчета бинарных и многоградационных дифракционных решеток с заданным числом порядков и с требуемой интенсивностью рассмотрено в разделах 7.4 и 7.5. Расчету ДОЭ, выполняющих специальные функции, посвящены последние разделы пособия. В разделе 7.6 рассматриваются итеративные алгоритмы для расчета ДОЭ, фокусирующих лазерное излучение в трехмерные области пространства или на двумерные поверхности тел вращения: конус, цилиндр. Сравнению разных итеративных методов для выравнивания гауссовой интенсивности лазерного пучка с помо-

щью ДОЭ посвящен раздел 7.7. Показано, что наилучшим приближением для итеративного расчета ДОЭ, фокусирующего гауссовое распределение интенсивности в равномерный световой квадрат, является геометрооптическое решение задачи. Итеративные алгоритмы расчета ДОЭ, фокусирующих когерентное излучение в кольцо или контурное изображение, например, изображение буквы, рассмотрены в разделах 7.8 и 7.9 соответственно. В разделах 7.10 и 7.11 рассмотрены градиентные методы для расчета ДОЭ с квантованной фазой.

1. Алгоритм уменьшения ошибки

Ниже рассматриваются алгоритмы решения нелинейного интегрального уравнения Френеля, предназначенного для расчета фазовых оптических элементов, формирующих произвольное заданное распределение интенсивности когерентного монохроматического света в некоторой плоскости, перпендикулярной оптической оси.

Изучаемые алгоритмы являются адаптивными, т.к. новая оценка искомой функции на каждой итерации выбирается не только в соответствии с требуемой функцией интенсивности, но и в зависимости от предыдущей оценки.

Эти алгоритмы называются также параметрическими, т.к. скорость их сходимости зависит от выбора конкретных значений некоторых весовых или регуляризационных параметров.

В рамках скалярной теории дифракции комплексная амплитуда волны в плоскости оптического элемента

$$W(u,v) = A(u,v)\, e^{i\varphi(u,v)}$$

связана с комплексной амплитудой волны

$$F(\xi,\eta) = B(\xi,\eta)\, e^{i\psi(\xi,\eta)}$$

в плоскости наблюдения, в которой формируется требуемое распределение интенсивности $I_0(\xi,\eta)$, через интегральное преобразование /2/

$$F(\xi,\eta) = \frac{ik}{2\pi z} e^{ikz} \iint_{-\infty}^{\infty} W(u,v) H(u-\xi, v-\eta, z)\, du\, dv, \qquad (1)$$

где

$$H(u-\xi, v-\eta, z) = \exp\left[\frac{ik}{2z}\left\{(u-\xi)^2 + (v-\eta)^2\right\}\right] \qquad (2)$$

— функция импульсного отклика свободного пространства в приближении Френеля, z — это расстояние между ДОЭ и плоскостью на-

блюдения, $k = 2\pi/\lambda$. В уравнении (1) комплексная амплитуда $W(u,v)$ в приближении тонкого оптического элемента (приближение транспаранта), которое не учитывает рефракцию лучей, равна произведению комплексной амплитуды $W_0(u,v)$ на собственную функцию пропускания $T(u,v)$ ДОЭ:

$$W(u,v) = W_0(u,v)\, T(u,v). \tag{3}$$

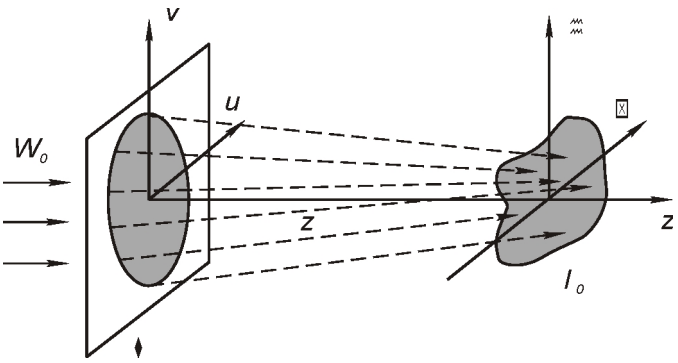


Рис. 1. Схема формирования изображения при помощи ДОЭ

Ввиду того, что в последующем мы предполагаем рассматривать только фазовые оптические элементы (если не указано иначе), функция пропускания ДОЭ выбрана в виде:

$$T(u,v) = \exp[\mathrm{i}\, g(u,v)], \tag{4}$$

где $g(u,v)$ — заданная фаза ДОЭ. Для иллюстрации вышеуказанных обозначений следует обратиться к рисунку 1. Задачу расчета фазовой функции ДОЭ $g(u,v)$ можно свести к решению нелинейного интегрального уравнения

$$I_0(\xi,\eta) = \left|F(\xi,\eta)\right|^2 = \left|\frac{k}{2\pi z}\int_{-\infty}^{\infty}\int A_0(u,v)\, e^{\mathrm{i}\varphi(u,v)} H(u-\xi, v-\eta, z)\,\mathrm{d}u\,\mathrm{d}v\right|^2, \tag{5}$$

где $I_0(\xi,\eta)$ — заданная интенсивность в области изображения, $A_0(u,v)$ — амплитуда освещающего пучка, $\varphi(u,v) = g(u,v) + g_0(u,v)$, где $g_0(u,v)$ — фаза освещающего пучка. Итеративный метод расчета фазы $\varphi(u,v)$, а вместе с ней фазы $g(u,v)$, состоит в решении уравнения (5) методом последовательных приближений. Алгоритм Герчберга — Секстона (ГС), или алгоритм уменьшения ошибки /3/, включает следующие шаги (см. его схему на рис. 2):

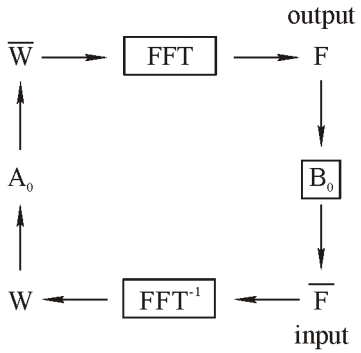


Рис. 2. Схема алгоритма уменьшения ошибки Герчберга-Секстона

1) выбирается начальная оценка фазы $\varphi_0(u,v)$;

2) осуществляется интегральное преобразование функции $A_0(u,v)\exp[i\varphi_0(u,v)]$ при помощи уравнения (1);

3) результирующая комплексная амплитуда $F(\xi,\eta)$ в плоскости формирования изображения заменяется на $\overline{F}(\xi,\eta)$ по правилу

$$\overline{F}(\xi,\eta) = B_0(\xi,\eta)\, F(\xi,\eta)\, \left|F(\xi,\eta)\right|^{-1}, \qquad (6)$$

где $B_0(\xi,\eta) = \sqrt{I_0(\xi,\eta)}$;

4) вычисляется преобразование, обратное (1) относительно функции $\overline{F}(\xi,\eta)$:

$$W(u,v) = \frac{\mathrm{i}k}{2\pi z} e^{-\mathrm{i}kz} \int_{-\infty}^{\infty}\int \bar{F}(\xi,\eta)\, H^{*}(\xi - u, \eta - v, z)\, \mathrm{d}\,\xi\, \mathrm{d}\,\eta \;; \tag{7}$$

5) полученная комплексная амплитуда $W(u,v)$ в плоскости ДОЭ заменяется на $\bar{W}(u,\mathrm{n})$ по правилу:

$$\bar{W}(u,v) = \begin{cases} A_0(u,v)\, W(u,v) \left| W(u,v) \right|^{-1}, & (u,v) \in Q, \\ 0, & (u,v) \notin Q, \end{cases} \tag{8}$$

где Q — форма апертуры ДОЭ;

6) переход к шагу 2.

Эта процедура повторяется до тех пор, пока ошибки δ_F и δ_W не перестанут значительно меняться:

$$\delta_F^2 = \frac{\int_{-\infty}^{\infty}\int \left[\left| F(\xi,\eta) \right| - B_0(\xi,\eta) \right]^2 \mathrm{d}\xi\, \mathrm{d}\eta}{\int_{-\infty}^{\infty}\int B_0^2(\xi,\eta)\, \mathrm{d}\,\xi\, \mathrm{d}\,\eta}, \tag{9}$$

$$\delta_W^2 = \frac{\int_{-\infty}^{\infty}\int \left[\left| W(u,v) \right| - A_0(u,v) \right]^2 \mathrm{d}\,u\, \mathrm{d}\,v}{\int_{-\infty}^{\infty}\int A_0^2(u,v)\, \mathrm{d}\,u\, \mathrm{d}\,v}. \tag{10}$$

Алгоритм ГС называют алгоритмом уменьшения ошибки потому, что было показано /4/, что ошибки (9) и (10) с ростом числа итераций не возрастают. Кроме того, было показано, что алгоритм ГС является вариантом градиентного метода /5/ или градиентного метода наискорейшего спуска /6/, при помощи которого можно минимизировать функционал среднеквадратичного отклонения амплитуды восстановленного изображения от заданного значения:

$$\varepsilon_0 = \iint_{-\infty}^{\infty} \left[\left| F(\xi,\eta) \right| - B_0(\xi,\eta) \right]^2 \mathrm{d}\xi \, \mathrm{d}\eta \,. \qquad (11)$$

Заметим однако, что процесс сходимости алгоритма ГС характеризуется эффектом стагнации: в ходе нескольких начальных итераций ошибка δ_F (или δ_W) быстро уменьшается, а все последующие итерации не приводят к ее заметному уменьшению. Эффект стагнации показывает, что алгоритм достигает локального минимума функционала (11), и в этом смысле он является квазиоптимальным. Чтобы увеличить скорость сходимости процедуры и частично избежать эффекта стагнации, используется целый ряд адаптивных алгоритмов, в которых вводятся некоторые параметры, контролирующие скорость сходимости. Поэтому иногда эти алгоритмы называют параметрическими.

2. Алгоритм входа-выхода

В ряде работ Фиенапом Дж. Р. /5,7,8/ разработан метод входа-выхода, в котором рассматривается проблема восстановления фазы по изменению одной интенсивности в области пространственного спектра. Рассмотрим этот алгоритм.

Функции комплексной амплитуды $W(u,v)$ в области ДОЭ и $F(\xi,\eta)$ в плоскости Фурье, (в задней фокальной плоскости сферической линзы), в которой формируется заданное изображение, связаны между собой при помощи двухмерного преобразования Фурье

$$F(\xi,\eta)=\frac{\mathrm{i}k}{2\pi f}\int_{-\infty}^{\infty}\int W(u,v)\exp\left[-\mathrm{i}\frac{k}{f}(u\xi+v\eta)\right]\mathrm{d}u\ \mathrm{d}v, \tag{12}$$

где f — фокальное расстояние тонкой линзы, формирующей Фурье-спектр.

Переход от преобразования Френеля (1), (2), к преобразованию Фурье (12) означает, что в последнем случае следует использовать комбинацию ДОЭ плюс линза.

Итеративное решение задачи восстановления фазы светового поля, т.е. аргумента функции $F(\xi,\eta)$ на основе известной амплитуды, т.е. модуля $|F(\xi,\eta)|$, при помощи алгоритма входа-выхода (одной из его модернизаций) имеет вид:

$$\overline{F}_{n+1}(\xi,\eta)=\overline{F}_n(\xi,\eta)+\beta\delta F_n(\xi,\eta), \tag{13}$$

где β — некоторый параметр $0{,}5\le\beta\le 1$,

$$\delta F_n(\xi,\eta)=\begin{cases}0, & (\xi,\eta)\notin\gamma,\\ -F_n(\xi,\eta), & (\xi,\eta)\in\gamma,\end{cases} \tag{14}$$

γ — область нарушения ограничений, накладываемых на заданную функцию поля,

$$F_n(\xi,\eta)=\Im D_A\Im^{-1}\left\{\overline{F}_n(\xi,\eta)\right\} \tag{15}$$

является выходной функцией, $\Im$ и $\Im^{-1}$ — прямое и обратное Фурье-преобразования, D_A — ограничивающий оператор в области ДОЭ, который может быть представлен в виде

$$D_A W_n = A_0 \frac{W_n}{|W_n|}. \tag{16}$$

Специально для решения проблемы синтеза ДОЭ предлагается выбирать приращение функции δF_n в алгоритме входа-выхода в следующем виде /8/:

$$\delta F = \left[B_0 F |F|^{-1} - F \right] + \left[B_0 F |F|^{-1} - B_0 \overline{F} |\overline{F}|^{-1} \right], \tag{17}$$

$B_0^2(\xi,\eta)$ — заданное распределение интенсивности в Фурье-плоскости, $\overline{F}$ и F — входная и выходная функции или функции комплексной амплитуды в плоскости Фурье-изображения после и до выполнения ограничений.

В работе /9/ сообщается о новом варианте итеративного алгоритма входа-выхода для синтеза ДОЭ, в котором предполагается, что фаза независима от модуля комплексной амплитуды света $F(\xi, \eta)$. Показано, что оптимальный выбор модуля и фазы на $(n+1)$-м шаге итерации принимает вид $(\overline{F}(\xi,\eta) = \overline{F})$:

$$\begin{aligned} |\overline{F}_{n+1}| &= B_0 + \beta \left[|F_n| - B_0 \right], \\ \arg \overline{F}_{n+1} &= \arg F_n + \alpha \left[\arg F_n - \arg \overline{F}_n \right]. \end{aligned} \tag{18}$$

На численных примерах в работе /9/ также показано, что лучшие значения α и β равны 2. Заметим, что итеративный алгоритм, который использует первое уравнение в (18) сходен с более ранним алгоритмом, описанным в работах /10,11/. Этот алгоритм авторами /9/ назван маятниковым. На наш взгляд он отличается от

начального варианта алгоритма входа-выхода, так как явно учитывает ограничение на амплитуду B_0 в Фурье-плоскости.

3. Адаптивно-аддитивный алгоритм

Первоначально алгоритм входа-выхода /5,7/ применялся для восстановления фазы светового поля на основе измерений одной интенсивности в плоскости пространственных частот. Такая проблема является характерной для звездной интерферометрии. Этот же алгоритм был успешно применен для расчета фазы ДОЭ /8/. Однако этот алгоритм недостаточно теоретически обоснован.

Ниже показано, что алгоритм (18) (только первое уравнение), который далее называется адаптивно-аддитивным алгоритмом (АА-алгоритмом), можно получить при минимизации некоторого функционала-критерия. Можно также показать, что для АА-алгоритма некоторая характерная среднеквадратичная ошибка с ростом числа итераций не возрастает.

Рассмотрим более подробно замену (6) для алгоритма ГС. Эта замена предполагает, что амплитуда света $|F_n(\xi,\eta)|$ в области наблюдения, полученная на n-м итеративном шаге заменяется на заданное значение амплитуды $B_0(\xi,\eta)$. Заметим, что, хотя преобразование Фурье (Френеля) от функции с ограниченной несущей есть аналитическая целая функция экспоненциального типа, функция $B_0(\xi,\eta)$ может не быть аналитической функцией и может быть задана произвольно.

Следовательно, имеет смысл вместо замены (6) попробовать использовать замену, в которой обе функции (заданная и рассчитанная аналитическая) участвуют в виде линейной комбинации с разными весами /12/:

$$\left|\overline{F}_n(\xi,\eta)\right| = \left|\gamma\, B_0(\xi,\eta) + (1-\gamma)\left|F_n(\xi,\eta)\right|\right|, \tag{19}$$

В этом случае замена (6) в итеративном ГС-алгоритме принимает вид

$$\overline{F}_n(\xi,\eta) = \left|\overline{F}_n(\xi,\eta)\right| F_n(\xi,\eta)\left|F_n(\xi,\eta)\right|^{-1}. \tag{20}$$

Диапазон значений параметра γ находится из условия, что среднеквадратичное отклонение ε_0, определяемое соотношением (11), не увеличивается при использовании замены (19). То есть мы предполагаем, что выполняется следующее условие:

$$\begin{aligned}
\overline{\varepsilon}_0 &= \int_{-\infty}^{\infty}\int \left[\left|\overline{F}_n\right| - B_0\right]^2 \mathrm{d}\xi\,\mathrm{d}\eta = \\
&= \int_{-\infty}^{\infty}\int \left[\gamma B_0 + (1-\gamma)\left|F_n\right| - B_0\right]^2 \mathrm{d}\xi\,\mathrm{d}\eta = \\
&= (1-\gamma)^2 \int_{-\infty}^{\infty}\int \left[\left|F_n\right| - B_0\right]^2 \mathrm{d}\xi\,\mathrm{d}\eta \le \varepsilon_0 = \\
&= \int_{-\infty}^{\infty}\int \left[\left|F_n\right| - B_0\right]^2 \mathrm{d}\xi\,\mathrm{d}\eta\,.
\end{aligned} \tag{21}$$

Из уравнения (21) следует, что весовой коэффициент следует выбирать из условия $0 \le \gamma \le 2$. Заметим, что для $\gamma = 1$ замена (19) переходит в уравнение (6) алгоритма ГС, в то время как $\gamma = 2$ приводит к замене "зеркального отражения":

$$\left|\overline{F}_n(\xi,\eta)\right| = \left|\, 2B_0(\xi,\eta) - \left|F_n(\xi,\eta)\right| \,\right|. \tag{22}$$

В последнем случае ($\gamma = 2$) амплитуда $|F_n(\xi,\eta)|$ в плоскости наблюдения, рассчитанная на n-м шаге итераций, заменяется на "зеркальное отражение" относительно заданного распределения амплитуды $B_0(\xi,\eta)$.

Заметим, что оригинальная замена "зеркального отражения" в уравнении (22) имела следующий вид /10,11/:

$$\overline{I}_n(\xi,\eta) = \begin{cases} \left|\, 2I_0(\xi,\eta) - I_n(\xi,\eta) \,\right|, & (\xi,\eta) \in \Omega, \\ I_n(\xi,\eta), & (\xi,\eta) \notin \Omega, \end{cases} \tag{23}$$

где $I_n(\xi,\eta)$ — распределение интенсивности, полученное на n-м шаге, $I_n = |F_n(\xi,\eta)|^2$, I_0 — заданное распределение интенсивности, $I_0 = B_0^2$,

Ω — область задания изображения. Заметим, что если $I_n \geq 2I_0$, то $\overline{I}_n(\xi,\eta) = 0$.

Покажем, что замена (19) в адаптивно-аддитивном алгоритме минимизирует функционал невязки ε_1, который представляет собой среднеквадратичное отклонение рассчитанной амплитуды от заданной в плоскости ДОЭ:

$$\varepsilon_1 = \int_{-\infty}^{\infty}\int \left[A_0(u,v) - \left|W(u,v)\right| \right]^2 \mathrm{d}u\ \mathrm{d}v , \tag{24}$$

где $A_0(u,v)$ — амплитуда света, падающего на ДОЭ. Вариация функционала ε_1 относительно функции W на n-м шаге определяется соотношением

$$\delta\varepsilon_1 = 2\,\mathrm{Re}\left\{ \int_{-\infty}^{\infty}\int \left[W_n - A_0 W_n \left| W_n \right|^{-1} \right] \delta W_n^* \,\mathrm{d}u\,\mathrm{d}v \right\} , \tag{25}$$

где Re{} означает действительную часть числа, * — знак комплексного сопряжения.

При получении уравнения (25) мы использовали соотношения

$$\delta\left\{ \int_{-\infty}^{\infty}\int A_0^2(u,v)\,\mathrm{d}u\ \mathrm{d}v \right\} = 0 , \tag{26}$$

$$\delta\left\{ \int_{-\infty}^{\infty}\int \left|W(u,v)\right|^2 \mathrm{d}u\ \mathrm{d}v \right\} \neq 0 , \tag{27}$$

$$\delta\left(WW^*\right) = W\delta W^* + W^*\delta W = 2\,\mathrm{Re}\left\{ W\delta W^* \right\} . \tag{28}$$

Используя равенства Парсеваля

$$\int_{-\infty}^{\infty}\int \left|W(u,v)\right|^2 \mathrm{d}u\,\mathrm{d}v = \int_{-\infty}^{\infty}\int \left|\overline{F}(\xi,\eta)\right|^2 \mathrm{d}\xi\,\mathrm{d}\eta , \tag{29}$$

$$\int_{-\infty}^{\infty}\int \left|\bar{W}(u,v)\right|^2 \mathrm{d}u\,\mathrm{d}v = \int_{-\infty}^{\infty}\int \left|F(\xi,\eta)\right|^2 \mathrm{d}\xi\,\mathrm{d}\eta\,, \tag{30}$$

где $\bar{W} = D_A W = AW\,|W|^{-1}$ — амплитуда света в области ДОЭ после осуществления замены (8), можно показать, что выполняется следующее соотношение:

$$\begin{aligned}&\mathrm{Re}\left\{\int_{-\infty}^{\infty}\int \left[W_n - A_0 W_n |W_n|^{-1}\right]\delta W_n^* \,\mathrm{d}u\,\mathrm{d}v\right\} = \\ &= \mathrm{Re}\left\{\int_{-\infty}^{\infty}\int \left[\bar{F}_n - F_n\right]\delta \bar{F}_n^* \,\mathrm{d}\xi\,\mathrm{d}\eta\right\}\end{aligned}\,, \tag{31}$$

где $\bar{F}_n = \Im\{W_n\}$, $\quad F_n = \Im D_A \Im^{-1}\{\bar{F}_n\} = \Im\left\{A_0 W_n |W_n|^{-1}\right\}$.

Приращение $\delta\varepsilon_1$ функционала ошибки ε_1 принимает максимальное отрицательное значение на n-м шаге при следующем параметрически выраженном условии:

$$\delta\bar{F}_n = \bar{F}_{n+1} - \bar{F}_n = \gamma\left(F_n - \bar{F}_n\right), \qquad \gamma > 0\,. \tag{32}$$

С учетом уравнения (32), вместо (25) мы получим

$$\delta\varepsilon_1 = -2\gamma\int_{-\infty}^{\infty}\int \left|\bar{F}_n - F_n\right|^2 \mathrm{d}\xi\,\mathrm{d}\eta < 0\,. \tag{33}$$

Из уравнения (32) следует параметрическое уравнение для интегрального расчета функции $F(\xi,\eta)$:

$$\bar{F}_{n+1} = (1-\gamma)\bar{F}_n + \gamma F_n = (1-\gamma)\bar{F}_n + \gamma\Im D_A \Im^{-1}\{\bar{F}_n\}\,. \tag{34}$$

Уравнение (34) предполагает, что на $(n+1)$-м шаге входная амплитуда представляет собой линейную комбинацию входной и выходной амплитуды на n-м шаге. Алгоритм (34) аналогичен алгоритму входа-выхода /5/, и предыдущие рассуждения следует рассматривать как теоретическое обоснование его оптимальности.

Для алгоритма (34) среднеквадратичная ошибка амплитуды не возрастает с ростом числа итераций при условии, что параметр γ удовлетворяет неравенству

$$0 < \gamma \le 2. \tag{35}$$

Действительно, введя обозначения

$$\begin{aligned} F_n &= B_n e^{\mathrm{i}\varphi_n}, \qquad \overline{F}_n = \overline{B}_n e^{\mathrm{i}\overline{\varphi}_n}, \\ W_n &= A_n e^{\mathrm{i}\psi_n}, \qquad \overline{W}_n = A_0 e^{\mathrm{i}\psi_n}, \end{aligned} \tag{36}$$

мы получим последовательность, в которой равенства 1, 3, 5 являются очевидными, равенства 2, 4 есть равенства Парсеваля, а неравенство представляет собой неравенство треугольника ($|a| - |b| \le |a - b|$):

$$\begin{aligned} \int_{-\infty}^{\infty}\int \left| A_0 - A_{n+1}\right|^2 \mathrm{d}u\ \mathrm{d}v &= \int_{-\infty}^{\infty}\int \left| A_0 e^{\mathrm{i}\psi_{n+1}} - A_{n+1} e^{\mathrm{i}\psi_{n+1}}\right|^2 \mathrm{d}u\ \mathrm{d}v \le \\ &\le \int_{-\infty}^{\infty}\int \left| A_0 e^{\mathrm{i}\psi_n} - A_{n+1} e^{\mathrm{i}\psi_{n+1}}\right|^2 \mathrm{d}u\,\mathrm{d}v = \\ &= \int_{-\infty}^{\infty}\int \left| B_n e^{\mathrm{i}\varphi_n} - \left[(1-\gamma)\,\overline{B}_n e^{\mathrm{i}\overline{\varphi}_n} + \gamma\, B_n e^{\mathrm{i}\varphi_n}\right]\right|^2 \mathrm{d}\xi\ \mathrm{d}\eta\ = \\ &= (1-\gamma)^2 \int_{-\infty}^{\infty}\int \left| B_n e^{\mathrm{i}\varphi_n} - \overline{B}_n e^{\mathrm{i}\overline{\varphi}_n}\right|^2 \mathrm{d}\xi\ \mathrm{d}\eta = \\ &= (1-\gamma)^2 \int_{-\infty}^{\infty}\int \left| A_0 e^{\mathrm{i}\psi_n} - A_n e^{\mathrm{i}\psi_n}\right|^2 \mathrm{d}u\ \mathrm{d}v = (1-\gamma)^2 \int_{-\infty}^{\infty}\int \left| A_0 - A_n\right|^2 \mathrm{d}u\ \mathrm{d}v. \end{aligned} \tag{37}$$

Из уравнения (37) следует, что при условии (35) имеет место релаксация среднеквадратической ошибки:

$$\int_{-\infty}^{\infty}\int \left| A_0 - A_{n+1}\right|^2 \mathrm{d}u\ \mathrm{d}v \le \int_{-\infty}^{\infty}\int \left| A_0 - A_n\right|^2 \mathrm{d}u\ \mathrm{d}v\ . \tag{38}$$

Отметим, что в алгоритме (34) учитывается, что комплексная амплитуда удовлетворяет ограничению в плоскости ДОЭ, но не учитывается, что модули входной и выходной функций комплексной амплиту-

ды в Фурье-плоскости, а именно $\overline{F}(\xi,\eta)$ и $F(\xi,\eta)$ должны стремиться к заданной функции $B_0(\xi,\eta)$. Чтобы использовать это условие в алгоритме (34) предположим, что входная функция на n-м шаге итерации подчиняется следующему ограничению

$$\overline{F}_n(\xi,\eta) = D_B F(\xi,\eta) = B_0(\xi,\eta) F(\xi,\eta) \left| F(\xi,\eta) \right|^{-1}, \tag{39}$$

и что на (n+1)-м шаге она подчиняется уравнению (34). Поэтому комбинация уравнений (34) и (39) приводит к адаптивно-аддитивному алгоритму.

$$\overline{F}_{n+1}(\xi,\eta) = (1-\gamma) B_0(\xi,\eta) \frac{F_n(\xi,\eta)}{\left| F_n(\xi,\eta) \right|} + \gamma F_n(\xi,\eta). \tag{40}$$

Заметим, что уравнение (40) отличается от (19) перестановкой весовых коэффициентов: $\gamma \to 1 - \gamma$.

Следует отметить также, что похожий прием был применен в работе /9/, чтобы получить алгоритм (40) из алгоритма (34).

В работе /13/ алгоритм (19) был получен как весовой алгоритм. Рассмотрим функционал невязки ε_0, но с весовой функцией:

$$\overline{\varepsilon}_0 = \int_{-\infty}^{\infty} \int S(\xi,\eta) \left[\left| F(\xi,\eta) \right| - B_0(\xi,\eta) \right]^2 \mathrm{d}\xi \, \mathrm{d}\eta, \tag{41}$$

где $S(\xi,\eta)$ — это реальная, положительно определенная весовая функция, которая задает, например, область определения синтезированного изображения

$$S(\xi,\eta) = \begin{cases} \gamma, & (\xi,\eta) \in \Omega, \\ 0, & (\xi,\eta) \notin \Omega, \end{cases} \tag{42}$$

где γ — постоянная.

Вариация функционала (41) имеет вид

$$\delta\overline{\varepsilon}_0 = 2\,\mathrm{Re}\int_{-\infty}^{\infty}\int S\left(\xi,\eta\right)\left[F - B_0\frac{F}{|F|}\right]\delta F^{*}\,\mathrm{d}\xi\,\mathrm{d}\eta\,. \tag{43}$$

Итеративное уравнение, которое позволяет получить функцию $F(\xi,\eta)$, минимизирующую функционал (41), следует из требования равенства нулю подынтегрального выражения в уравнении (43) и имеет вид:

$$\overline{F}_{n+1} = F_n + S\left[B_0\frac{F_n}{|F_n|} - F_n\right] = \left(1-S\right)F_n + SB_0\frac{F_n}{|F_n|}, \tag{44}$$

где F_n и $\overline{F}_{n+1}$ функции выхода на n-м шаге и входа на (n+1)-м шаге, соответственно. При получении уравнения (44) использовалось нетрадиционное понимание вариации функции. Традиционное понимание вариации- это изменение амплитуды на входе и выходе от итерации к итерации: $\delta F_n = F_{n+1} - F_n$. Мы понимаем вариацию функции, как переход с выхода на вход: $\delta F_n = \overline{F}_{n+1} - F_n$.

Преимущества АА-алгоритма можно проиллюстрировать на следующих примерах /14/.

Пример 1. Рассчитывается ДОЭ, который освещается лазерным пучком света с гауссовой интенсивностью

$$A_0\left(u,v\right) = \exp\left[-\frac{u^2+v^2}{r^2}\right], \tag{45}$$

и формирует изображение "мягкого квадрата" в дальней зоне дифракции. Это означает, что заданное распределение интенсивности в плоскости наблюдения выбирается в форме супергауссовой функции:

$$B_0\left(\xi,\eta\right) = \exp\left[-\frac{\xi^{2n}+\eta^{2n}}{a^{2n}}\right], \qquad n = 1, 2, \ldots\,, \tag{46}$$

где $2a$ — эффективная сторона квадрата, r — радиус гауссова пучка. Такая задача выравнивания интенсивности гауссового пучка актуальна для многих приложений: лазерной печати, лазерной термообработки поверхности. Подробнее эта задача рассматривается в разделе 7.7.

Преобразование Френеля (1) и обратное ему (7) рассчитывались при помощи алгоритма быстрого преобразования Фурье. Массив состоял из 256 × 256 пикселов. Интенсивность освещающего гауссова пучка вдоль границ квадратной апертуры ДОЭ составляла 10% от максимального значения в центре. Сторона изображения квадрата была равна примерно 10 диаметрам диска Эйри (минимальное дифракционное пятно). Число итераций равнялось 10.

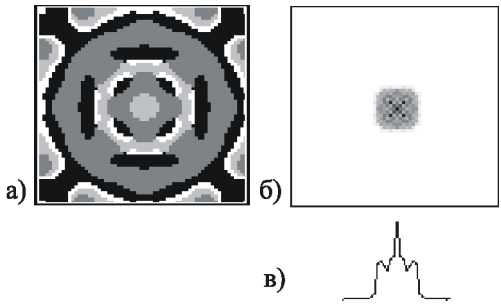


Рис. 3. Расчет ДОЭ при помощи ГС-алгоритма: фаза ДОЭ от 0 до 2π (а), сформированное изображение (б) , центральное горизонтальное сечение интенсивности изображения (в)

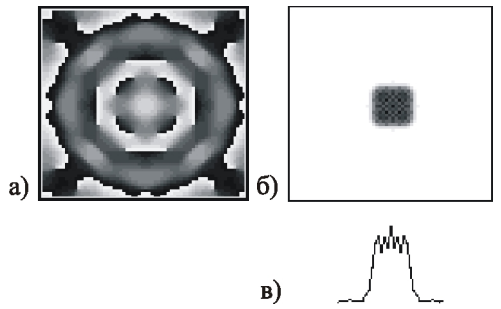


Рис. 4 Расчет ДОЭ при помощи АА-алгоритма: фаза ДОЭ от 0 до 2π(а), сформированное изображение (б), центральное горизонтальное сечение интенсивности изображения (в)

На рис. 3 показаны: фаза ДОЭ по модулю 2π (полутоновая, 16 уровней), рассчитанная при помощи ГС-алгоритма (а), изображение в дальней зоне дифракции (б) , горизонтальное сечение распределения интенсивности на изображении (в).

На рис. 4 показаны аналогичные результаты, полученные при помощи АА-алгоритма ($\lambda = 2$): фаза ДОЭ (а), сформированное изображение (б), горизонтальное сечение интенсивности изображения (в).

Среднеквадратичные ошибки, рассчитанные по формуле (9) для приведенных случаев, имели следующие значения: 27,2% (рис. 3) и 5,1% (рис. 4).

Введем понятие энергетической эффективности формирования заданного изображения. Это часть световой энергии, попадающей в заданную область определения изображения:

$$E = \frac{\iint\limits_{\Omega} \left|F(x,y)\right|^2 \mathrm{d}x\,\mathrm{d}y}{\int\limits_{-\infty}^{\infty}\int \left|F(x,y)\right|^2 \mathrm{d}x\,\mathrm{d}y}, \qquad (47)$$

где Ω — форма области определения изображения. В рассмотренном случае энергетическая эффективность была равна: 96,4% (рис. 3) и 91,5% (рис.4).

Из рис. 3 и 4 видно, что фазы, полученные этими двумя методами, слабо отличаются друг от друга. Однако незначительная разница между ними приводит к существенному различию в формируемых изображениях.

Моделирование показало, что скорость сходимости итеративного алгоритма увеличивается с ростом параметра релаксации γ от 0 до 2.

Пример 2. В данном примере рассчитывается оптический элемент, способный эффективно определять положение источника излучения /15/. Этот элемент является частью прибора для определения

координат центра поперечного сечения изображения точечного источника. Схема устройства представлена на рис. 5.

Точечный источник света S освещает сферическую линзу, за которой помещен ДОЭ (рис. 5). ДОЭ рассчитывается таким образом, что изображение точечного источника имеет распределение интенсивности в форме креста. Каждая сторона изображения креста пересекает линейный фотодатчик, как показано на рис. 6. В качестве линейного фотодатчика с характеристической стороной входного окна, равной 10 мм можно использовать кремниевую полупроводниковую структуру, у которой выходной аналоговый сигнал в форме падения напряжения пропорционален расстоянию между точкой максимальной интенсивности света и исходной точкой датчика. Технические параметры такого датчика следующие: изменение напряжения на 0,1 мВ эквивалентно смещению максимальной интенсивности света вдоль оси датчика на 0,2 мкм.

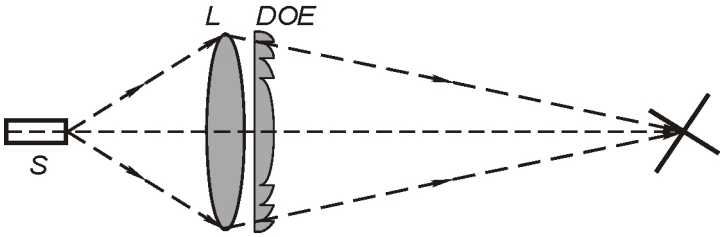


Рис. 5. Оптическая схема формирования крестообразного изображения

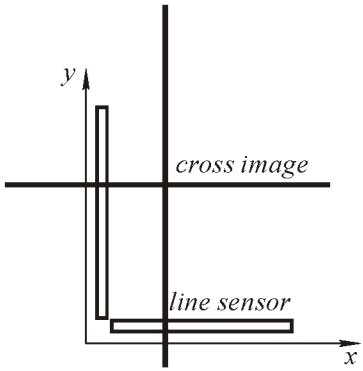


Рис. 6. Плоскость изображения прибора

Для расчета ДОЭ при указанных условиях использовалось параксиальное соотношение, связывающее распределение интенсивности в плоскости источника и изображения креста/16/:

$$F(\xi,\eta)=\frac{k^2}{ab}\exp\left[\mathrm{i}\frac{k}{2b}\left(\xi^2+\eta^2\right)\right]\times$$
$$\int_{-\infty}^{\infty}\int W_o(u,v)\exp\left[\mathrm{i}\frac{k}{2a}(u^2+v^2)\right]H\left(u+\frac{b}{a}\xi,v+\frac{b}{a}\eta\right)\mathrm{d}u\,\mathrm{d}v, \quad (48)$$

где a и b — это расстояния от источника до линзы и от линзы до изображения (рис 5), связанные соотношением для тонкой линзы

$$\frac{1}{a}+\frac{1}{b}=\frac{1}{f}, \quad (49)$$

где $H(u,v)$ — функция импульсного отклика, задаваемая в виде:

$$H(u,v)=\iint_{Q}P(\xi,\eta)\;T(\xi,\eta)\exp\left[-\mathrm{i}\frac{k}{b}(u\xi+v\eta)\right]\mathrm{d}\xi\;\mathrm{d}\eta. \quad (50)$$

Функция $P(\xi,\eta)$ в (50) называется функцией зрачка линзы, Q определяет форму апертуры диафрагмы, ограничивающей линзу. Функция $T(\xi,\eta)$ — это комплексная функция пропускания ДОЭ, являющаяся чисто фазовой функцией:

$$T(\xi,\eta)=e^{\mathrm{i}\varphi(\xi,\eta)}. \quad (51)$$

Фаза $\varphi(\xi,\eta)$ оптического элемента находится как решение интегрального уравнения

$$I(u,v)=|H(u,v)|^2, \quad (52)$$

где $I(u,v)$ — заданное распределение интенсивности для функции импульсного отклика или заданная функция рассеяния точки. Для креста эта функция интенсивности принимает вид:

$$I(u,v)=\begin{cases} I_0; & -l \le u \le l, \quad v=0 \quad u \qquad u=0, \quad -l \le v \le l, \\ 0; иначе, \end{cases} \tag{53}$$

где I_0 — постоянное значение, $2l$ — длина вертикального и горизонтального участков креста. Уравнение (52) решалось при помощи AA-алгоритма (19) при $\gamma = 2$.

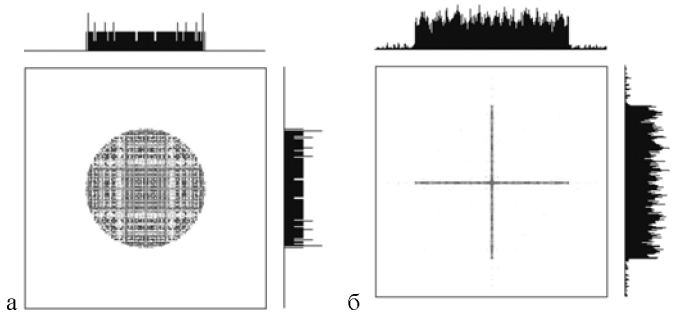


Рис. 7. Фаза ДОЭ, 5 градаций (а); крест в области изображения (б)

Параметры расчета: ДОЭ был ограничен круглой диафрагмой радиусом 64 пиксела, для выполнения преобразования Фурье был выбран массив размерностью 256 × 256 пикселов, каждая сторона креста состояла из 170 пикселов, ширина была равна одному пикселу. На рис. 7а показана фаза оптического элемента, рассчитанная за 25 итераций с 5 градациями. Квадрат модуля функции импульсного отклика для комбинации «линза плюс ДОЭ» показан на рис. 7б. Горизонтальный и вертикальный профили для полученных картин фазы и функции импульсного отклика также показаны на рис. 7.

4. Адаптивно-мультипликативный алгоритм

В этом разделе рассматривается еще один итеративный алгоритм для расчета ДОЭ, который также имеет скорость большую, чем скорость сходимости алгоритма ГС. Этот алгоритм может быть получен при помощи тождества, которое должно выполнятся в плоскости формирования изображения:

$$I_0(\xi,\eta) = F(\xi,\eta)F^*(\xi,\eta) \quad (54)$$

То есть заданную интенсивность $I_0(\xi,\eta)$ должно формировать световое поле с комплексной амплитудой $F(\xi,\eta)$.

Используя обратный оператор $\hat{L}^{-1}$, получим в плоскости ДОЭ

$$W(u,v) = \hat{L}^{-1}\left[\frac{I_0(\xi,\eta)}{F^*(\xi,\eta)}\right] = \hat{L}^{-1}\left[\frac{I_0 F}{|F|^2}\right], \quad (55)$$

где $\hat{L}^{-1}$ — обратное преобразование Френеля.

Уравнение (55) позволяет предложить итеративную процедуру поиска фазы ДОЭ (аргумент функции $W(u, v)$), которая очень похожа на процедуру, используемую в алгоритме ГС, но вместо уравнения (6) использует следующую замену:

$$\overline{F}(\xi,\eta) = I_0(\xi,\eta)F(\xi,\eta)\left|F(\xi,\eta)\right|^{-2}. \quad (56)$$

Уравнение (56) являетя некорректным, т.к. возможно равенство нулю модуля $F(\xi, \eta)$ в знаменателе. Корректная формулировка задачи нахождения фазы ДОЭ должна основываться на рассмотрении некоторого функционала невязки с регуляризацией /17/. С этой целью рассмотрим функционал

$$\varepsilon_2 = \int_{-\infty}^{\infty}\int\left[\,|F|^2 - I_0\right]^2 \mathrm{d}\xi\,\mathrm{d}\eta + \alpha\int_{-\infty}^{\infty}\int Q\,|F|^2\,\mathrm{d}\xi\,\mathrm{d}\eta. \quad (57)$$

Первый член в уравнении (57) — это среднеквадратичное отклонение формируемой интенсивности в плоскости наблюдения от заданной интенсивности $I_0(\xi,\ \eta)$.

В отличие от функционала (11), это слагаемое представляет собой функционал четвертого порядка относительно амплитуды светового поля. Второе слагаемое в уравнении (57) — стабилизирующее.

Отметим, что действительная, положительно определенная функция Q в уравнении (57) в методе регуляризации А.Н.Тихонова /18/ в виде:

$$Q(\xi,\eta)=\sum_{n=0}^{N}\left|C_n\right|\left(\xi^{2n}+\eta^{2n}\right), \tag{58}$$

где C_n — произвольные коэффициенты. Стабилизирующая константа $\alpha \geq 0$ определяет нижнюю границу изменения функционала ε_2. Минимизация ε_2 означает выбор самой гладкой функции из набора функций F наиболее близких по модулю к функции $\sqrt{I_0(\xi,\eta)}$. Итеративная процедура минимизации функционала (57) строится при помощи промежуточных функционалов, квадратичных относительно заданной функции $G_n(\xi,\eta)$

$$\varepsilon_{2n}=\int_{-\infty}^{\infty}\int\left|F_nG_n^*-I_0\right|^2\mathrm{d}\xi\,\mathrm{d}\eta+\alpha\int_{-\infty}^{\infty}\int Q\left|G_n\right|^2\mathrm{d}\xi\,\mathrm{d}\eta, \tag{59}$$

где $F_n(\xi,\eta)$ — комплексная амплитуда, рассчитанная на n-м шаге итераций в плоскости наблюдения, $G_n(\xi,\eta)$ — функция, минимизирующая функционал ε_{2n}. Вариация функционала (59) относительно функции G_n равна

$$\delta\varepsilon_{2n}=2\,\mathrm{Re}\left\{\int_{-\infty}^{\infty}\int\left[\left|F_n\right|^2G_n-I_0F_n+\alpha QG_n\right]\delta G_n^*\,\mathrm{d}\xi\,\mathrm{d}\eta\right\}. \tag{60}$$

Функционал (59) достигает минимума при условии, что вариация (60) становится равной 0. Как видно из (60), это становится возможным при условии, что

$$G_n = \frac{I_0 F_n}{\left|F_n'\right|^2 + \alpha\, Q}. \qquad (61)$$

Из сравнения уравнений (61) и (56) видно, что функция $G_n(\xi, \eta)$ является регулярным аналогом функции $\overline{F}(\xi, h)$. Замена (61) минимизирует промежуточные функционалы ε_{2n} на каждой итерации. Однако из-за сходимости итеративной процедуры разность $|F_n - G_n|$ уменьшается с ростом n, а функционалы ε_{2n} стремятся к функционалу ε_2. Следовательно, можно заключить, что в пределе замена (61) минимизирует исходный функционал (57). Приведенные рассуждения подтверждают возможность итеративного поиска фазы ДОЭ методом, сходным с алгоритмом ГС, но с использованием замены (61) вместо замены (6).

Данный алгоритм получил название адаптивно-мультипликативного (АМ) алгоритма, в отличие от АА-алгоритма. Его преимущества перед алгоритмом ГС видны из следующего примера.

Пример 3. Рассчитывается ДОЭ /17/, предназначенный формировать изображение "мягкого круга" в дальней зоне дифракции. Элемент освещается лазерным пучком с гауссовым распределением интенсивности. Заданная интенсивность в плоскости наблюдения имеет вид

$$I_0(\xi, \eta) = \exp\left[-2\left\{\frac{\xi^2 + \eta^2}{a^2}\right\}^n\right], \qquad (62)$$

где a — эффективный радиус "мягкого круга" $n = 1, 2, 3, \ldots$

Радиус a выбирался равным примерно двум радиусам диска Эйри. Массив пикселов был 256×256.

На рис.8 показаны: (а) фаза ДОЭ по модулю 2π (16 градаций), рассчитанная при помощи алгоритма ГС за 20 итераций, (б) результирующая дифракционная картина, (в) распределение интенсивности света в центральном сечении этой картины.

На рис. 9 показаны аналогичные результаты, полученные при помощи АМ-алгоритма: (а) фаза, рассчитанная за 20 итераций, (б) дифракционная картина, (в) центральное сечение картины в относительных единицах.

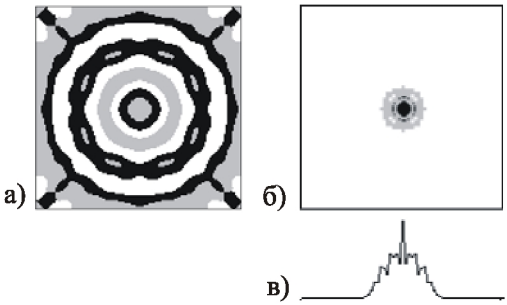


Рис. 8 Расчет при помощи алгоритма ГС: фаза ДОЭ 0 до 2π (а), дифракционная картина (б), распределение интенсивности в центральном сечении (в)

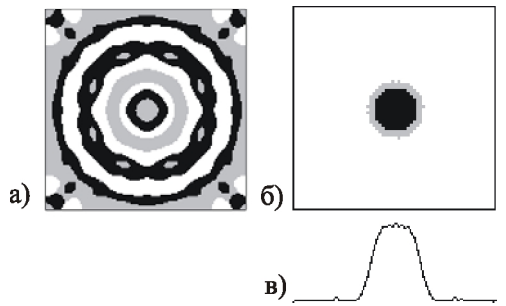


Рис. 9 Расчет при помощи АМ-алгоритма: фаза ДОЭ 0 до 2π (а), дифракционная картина (б), распределение интенсивности в центральном сечении (в)

Для данного случая энергетические эффективности, рассчитанные по формуле (47), были равны: 97,1% (рис.8), 91,8% (рис.9). Среднеквадратичные ошибки, рассчитанные по формуле (9), были равны: 22,9% (рис. 8) и 3,6% (рис. 9).

Из сравнения рис. 8 и 9 видно, что фазы отличаются друг от друга незначительно. Тем не менее эта незначительная разница приводит к формированию сильно различающихся дифракционных картин.

Отсутствие круговой симметрии в фазовых картинах, представленных на рис. 8а и 9а, можно объяснить нарушением круговой симметрии функции $I_0(\xi, \eta)$ в уравнении (62) для дискретных значений переменных ξ и η, при котором окружности переходят в многоугольники.

На рис. 10 показана зависимость значения функционала ε_2 в (57) в зависимости от числа итераций: кривая 1 — для ГС-алгоритма ($\alpha = 0$), кривая 2 — для АМ-алгоритма ($\alpha = 10^{-8}$). Из рис. 10 видно, что адаптивный алгоритм обладает лучшей сходимостью: кривая 2 достигает значения 1 после 4 итераций, в то время как кривая 1 достигает значения 1,2 после 20 итераций.

Заметим однако, что рис. 10 также выявляет недостаток АМ-алгоритма: его немонотонную сходимость. Но если брать значения функционала (57) через итерацию (см. рис. 10, кривая 2), они будут убывать монотонно. Для всех рассмотренных примеров начальная оценка фазы выбиралась случайной, и результаты практически не зависят от ее конкретной реализации.

Также оказывается трудным обосновать теоретически оптимальный выбор стабилизирующей константы α и степени полинома N для функции $Q(\xi, \eta)$, которые влияют на минимизацию функционала (57). Однако для ряда практических применений оказывается достаточным выбрать функцию $Q(\xi, \eta)$ как одночлен $(\xi^2+\eta^2)^n$ при $n = 1, 2$, и найти константу α методом подбора.

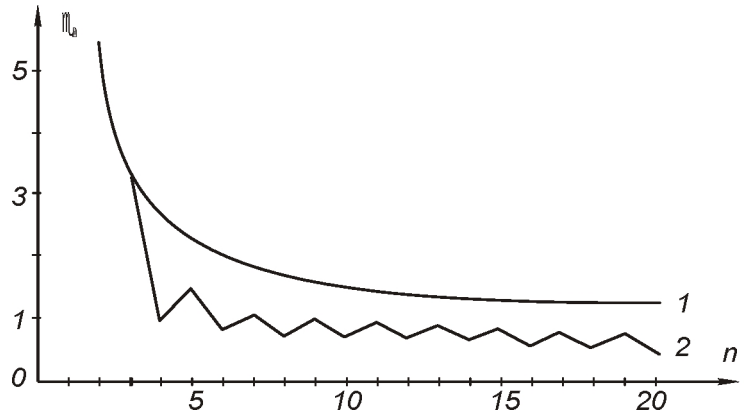


Рис. 10. Функционал ε2 как функция от числа итераций

Из вида функционала (57) при $\alpha = 0$ можно получить другую версию мультипликативного алгоритма. Рассмотрим функционал ошибки, заданный выражением

$$\varepsilon_{2p} = \int_{-\infty}^{\infty}\int \left[B_0^2(\xi,\eta) - \left|F(\xi,\eta)\right|^2 \right]^p \mathrm{d}\xi\, \mathrm{d}\eta, \qquad (63)$$

p — целое число.

Вариация этого функционала относительно функции $F(\xi,\eta)$ имеет вид

$$\delta\varepsilon_{2p} =$$
$$2\,\mathrm{Re}\iint p\left[B_0^2(\xi,\eta) - \left|F(\xi,\eta)\right|^2 \right]^{p-1} F(\xi,\eta)\,\delta F^*(\xi,\eta)\ \mathrm{d}\xi\, \mathrm{d}\eta, \qquad (64)$$

где $\delta F = \overline{F} - F$ разность между входной $\overline{F}$ и выходной F комплексными амплитудами в изображений Фурье. Из условия равенства нулю подынтегрального выражения в (64) следует итеративное уравнение для нахождения входной функции $\overline{F}$, минимизирующей функционал (63):

$$\overline{F}_{n+1} - F_n = \gamma p \left[B_0^2 - \left|F_n\right|^2 \right]^{p-1} F_n,$$

или

$$\overline{F}_{n+1} = \left\{1 + \gamma p\left[B_0^2 - \left|F_n\right|^2\right]^{p-1}\right\} F_n, \tag{65}$$

где γ — постоянная релаксации. Уравнение (65) нелинейно связывает функцию выходной амплитуды F_n на n-м шаге итераций с входной функцией $\overline{F}_{n+1}$. Заметим, что сама функция F_n рассчитывается при помощи уравнения

$$F_n = \Im D_A \Im^{-1}\left\{\overline{F}_n\right\}, \tag{66}$$

где ограничивающий оператор D_A в плоскости ДОЭ определяется уравнением (16).

В заключение укажем, что совместное применение АА-алгоритма, уравнения (40), АМ-алгоритма, и уравнения (61), порождает новый двухпараметрический итеративный алгоритм, который можно записать в виде:

$$\overline{F}_{n+1}(\xi,\eta) = \frac{(1-\gamma)\, I_0 F_n(\xi,\eta)}{\left|F_n(\xi,\eta)\right|^2 + \alpha Q(\xi,\eta)} + \gamma\, F_n(\xi,\eta). \tag{67}$$

5. Адаптивно-регуляризационный алгоритм

Так как ДОЭ формируют, как правило, изображения, ограниченные в пространстве, то появляется свобода в выборе поведения светового поля за областью определения изображения. Эта дополнительная степень свободы используется как в ранних алгоритмах Герчберга Р. /19/ и Папулиса А. /20/, так и в более поздних работах Вировского Ф. /21/ и Фиенапа Дж. /13/, и позволяет частично устранять эффект стагнации, присущий алгоритму Герчберга-Секстона /3/.

В данном разделе рассматривается алгоритм, являющийся обобщением перечисленных выше методов, так как он включает в себя релаксационную постоянную, учитывает пространственное ограничение области задания формируемого изображения и использует дополнительный параметр — параметр регуляризации.

Так как рассчитываемый ДОЭ должен формировать на некотором расстоянии заданное изображение в некоторой определенной области, то световое поле за пределами этой области может быть рассмотрено как некоторый шум. Энергетический вклад светового поля за пределами выделенной области может регулироваться согласно методу регуляризации А.Н. Тихонова /18/ стабилизирующим слагаемым. Сформулируем задачу расчета ДОЭ, фокусирующего падающее на него лазерное излучение в малую область Фурье-плоскости, как вариационную задачу минимизации функционала-критерия (для простоты обозначений рассмотрим одномерный случай):

$$\varepsilon(\alpha)=\int_{-\infty}^{\infty}\left[\,\left|F_0(u)\right|-D_A\left|F(u)\right|\,\right]^2\mathrm{d}u+\alpha\int_{-\infty}^{\infty}Q(u)\,\left|F(u)\right|^2\mathrm{d}u\,, \qquad (68)$$

где $Q(u)=\sum_{n=0}^{N}\left|c_n\right|u^{2n}$ — стабилизирующая функция, аналогичная функции (58),

$D_A f = \begin{cases} f, & u \in \Omega_A \\ 0, & u \notin \Omega_A \end{cases}$ — оператор пространственного ограничения, $|F_0(u)|$ — заданное распределение амплитуды в области Ω_A, α — параметр регуляризации, *F(u)* — комплексная амплитуда света в плоскости Фурье-спектра, в которой формируется требуемое изображение $|F_0(u)|$. Функционал (68) в отличие от функционала (57) квадратичный по амплитуде светового поля.

Вариация функционала (68) запишется в виде

$$\delta\varepsilon = 2\,\mathrm{Re}\left[\int_{-\infty}^{\infty}\left(-|F_0(u)|\frac{F(u)}{|F(u)|} + D_A F(u) + \alpha\, Q(u)F(u)\right)\delta F^*(u)\,\mathrm{d}u\right]. \quad (69)$$

Если выбрать вариацию δF как разность комплексных амплитуд до и после удовлетворения пространственным ограничениям

$$\delta F = \overline{F} - F, \quad (70)$$

то максимальная отрицательная вариация (69) функционала (68) достигается при условии

$$\overline{F} - F = \gamma G(u) = \gamma\left(|F_0(u)|\frac{F(u)}{|F(u)|} - D_A F(u) - \alpha\, Q(u)\, F(u)\right), \quad (71)$$

где $\gamma > 0$ — параметр адаптации алгоритма или константа релаксации. Действительно, при условии (71) вариация (69) будет иметь вид

$$\delta\varepsilon = -2\gamma\int_{-\infty}^{\infty}|G(u)|^2\,\mathrm{d}u \;\le 0. \quad (72)$$

Из условия (71) следует двухпараметрическое итеративное уравнение расчета ДОЭ /22/:

$$\overline{F}_{n+1} = \left[\,1 - \gamma\left(D_A + \alpha\, Q(u)\right)\,\right]F_n + \gamma|F_0(u)|\frac{F_n(u)}{|F_n(u)|}. \quad (73)$$

Заметим, что в уравнении (73) комплексная амплитуда на n-ом шаге итераций $F_n(u)$ является результатом удовлетворения ограничениям в плоскости ДОЭ и представима в виде

$$F_n(u) = \Im D_B W_B \Im^{-1} \overline{F}_n(u), \tag{74}$$

где $\overline{F}_n(u)$ — комплексная амплитуда на n-ом шаге, после замены (73), $D_B f(x) = \begin{cases} f(x), & x \in \Omega_B, \\ 0, & x \notin \Omega_B \end{cases}$ — оператор пространственного ограничения, Ω_B — форма апертуры ДОЭ, $W_B f(x) = |f_0(x)| \left(f(x) / |f(x)| \right)$ — оператор присвоения заданной амплитуды $|f_0(x)|$ освещающего пучка комплексной функции *f(x)*, описывающей световое поле в плоскости ДОЭ.

Итак, уравнения (73) и (74) описывают итеративный двухпараметрический алгоритм расчета ДОЭ, формирующего заданное распределение интенсивности (изображение) в ограниченной области Ω_B пространственного спектра. Назовем его адаптивно-регуляризационным (АР) алгоритмом.

Покажем далее, что некоторые известные алгоритмы расчета ДОЭ являются частными случаями АР-алгоритма.

При следующем выборе параметров $\gamma = 1$, $\alpha = 0$ вместо уравнения (73) получим следующее выражение:

$$\overline{F}_{n+1} = \left[1 - D_A\right] F_n + |F_0(n)| \frac{F_n(n)}{|F_n(n)|} = \begin{cases} |F_0(n)| \dfrac{F_n(n)}{|F_n(n)|}, & u \in \Omega_A, \\ F_n(n), & u \notin \Omega_A. \end{cases} \tag{75}$$

Алгоритм (75) сочетает в себе свойства алгоритма ГС /3/ и алгоритма Герчберга — Папулиса /19, 20/. Действительно, в соответствии с (75) в Фурье плоскости в заданной области Ω_A осуществляется замена рассчитанного модуля комплексной амплитуды $|F_n(u)|$ на заданный

$|F_0(u)|$, а вне области Ω_A рассчитанная функция $F_n(u)$ остается без изменений. Поэтому уравнение (75) можно интерпретировать как алгоритм ГС, обладающий сверхразрешающим свойством алгоритма Герчберга — Папулиса. Уравнение (75) рассматривалось в /21/ как реализация идеи дополнительной степени свободы для устранения стагнации итеративного алгоритма /3/.

При выборе параметров $\gamma \neq 1$, $\alpha = 0$ вместо уравнения (73) получим выражение

$$\overline{F}_{n+1} = \begin{cases} \gamma \left| F_0(u) \right| \dfrac{F_n(u)}{\left| F_n(u) \right|} + (1-\gamma) F_n, & u \in \Omega_A, \\ F_n(u), & u \notin \Omega_A. \end{cases} \tag{76}$$

Уравнение (76) является аналогом алгоритма входа-выхода /19/, метода ортогональных проекций /23/, адаптивного /10,11/ и маятникового /19/ алгоритмов, а также выражает идею “мягкого” кодирования комплексной амплитуды, высказанную в /24/. Релаксационный параметр γ в уравнении (76) позволяет регулировать скорость сходимости алгоритма, и при выборе $\gamma = 2$ эта скорость обычно бывает максимальной /9,10/.

При выборе параметров $\gamma = 1$, $\alpha \neq 0$ вместо уравнения (73) получим выражение

$$\overline{F}_{n+1} = \left[(1-\alpha) - D_A \right] F_n + \left| F_0(u) \right| \frac{F_n(u)}{\left| F_n(u) \right|}$$

$$= \begin{cases} \left| F_0(u) \right| \dfrac{F_n(u)}{\left| F_n(u) \right|} - \alpha F_n, & u \in \Omega_A, \\ (1-\alpha)\, F_n(u), & u \notin \Omega_A. \end{cases} \tag{77}$$

Алгоритм (77) аналогичен регуляризованному алгоритму Герчберга — Папулиса, предложенному в /25/ для задачи восстановления действительного объекта по зашумленному изображению. Параметр регу-

ляризации α в уравнении (77) регулирует вклад рассчитанной амплитуды света $F_n(u)$ внутри и вне заданной области Ω_A.

В общем случае АР-алгоритм (73) можно переписать в виде

$$\overline{F}_{n+1} = \begin{cases} \gamma \left| F_0(u) \right| \dfrac{F_n(u)}{\left| F_n(u) \right|} + (1 - \gamma - \gamma\alpha) F_n, & u \in \Omega_A \ , \\ (1 - \gamma\alpha)\ F_n(u), & u \notin \Omega_A \ . \end{cases} \tag{78}$$

Интересно отметить, что ни при каких значениях параметров γ и α нельзя перейти от АР-алгоритма (78) к алгоритму сокращения ошибки /3/, имеющего в данных обозначениях следующий вид:

$$\overline{F}_{n+1} = \begin{cases} \left| F_0(u) \right| \dfrac{F_n(u)}{\left| F_n(u) \right|}, & u \in \Omega_A \ , \\ 0 \quad , & u \notin \Omega_A \ . \end{cases} \tag{79}$$

Переход от уравнения (78) к уравнению (79) возможен только при условии рассмотрения релаксационной константы как двойной постоянной, принимающей разные значения внутри и вне заданной области:

$$\gamma = \begin{cases} (1 + \alpha)^{-1}, & u \in \Omega_A \ , \\ \alpha^{-1} \quad , & u \notin \Omega_A \ . \end{cases} \tag{80}$$

Пример 4 Преимущества алгоритма (73) или (78) по сравнению с базовым алгоритмом ГС (79) показаны для случая формирования в фокальной плоскости линзы распределения интенсивности вида $I(x) = I_0 \,\mathrm{rect}\,(x/a)$. Преобразование Фурье, связывающее комплексные амплитуды светового поля в плоскости ДОЭ и в фокальной плоскости линзы, вычислялось с помощью алгоритма быстрого преобразования Фурье на массиве отсчетов 256. Начальная фаза для итеративного процесса выбиралась в виде реализации псевдослучайной величины.

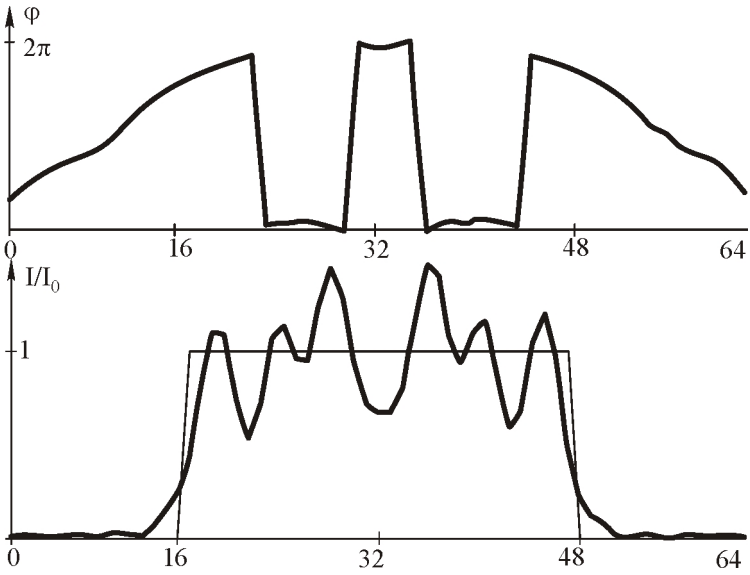


Рис. 11. Фаза ДОЭ (а), рассчитанная алгоритмом (79), и распределение интенсивности в Фурье-плоскости (б)

На рис. 11 показаны фаза ДОЭ (а), рассчитанная за 30 итераций алгоритмом (79), и нормированное распределение интенсивности (б) в фокальной плоскости линзы, сформированное ДОЭ с данной фазой. Оптический элемент занимал 64 из 256 пикселов. Пунктирная линия на рис.11*б* показывает требуемое распределение интенсивности, заданное на отрезке от 16 до 48 пикселов. В этот отрезок попадает 97,5% световой энергии, освещающей ДОЭ, а среднеквадратичное отклонение получившегося распределения интенсивности (рис.11*б*) от постоянного значения равно 27,2%. Если использовать фазу ДОЭ, показанную на рис.11а в качестве начальной фазы для алгоритма (78), то за следующие 100 итераций получится результат, показанный на рис.12. На этом рис. показана фаза ДОЭ (а) и сформированное им в фокальной плоскости нормированное распределение интенсивности (б). При этом выбирались параметры $\gamma = 1$, $\alpha = 0{,}1$. В отрезок отсчетов от 16 до 48 пикселов попадает 92,7% от всей световой энергии, а среднеквадратичное отклонение интенсивности от постоянного значения равно 2%.

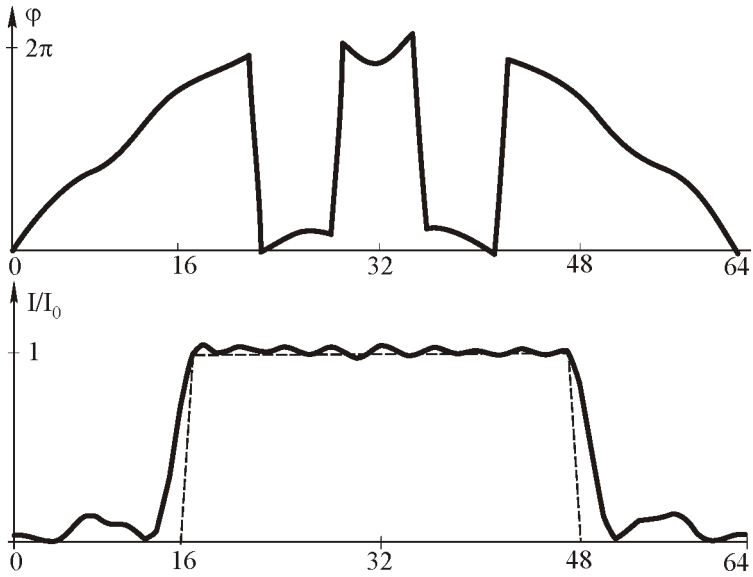


Рис. 12. Фаза ДОЭ (а), рассчитанная по алгоритму (78), и интенсивность в фокальной плоскости (б)

Таким образом, двухпараметрический АР-алгоритм (78) позволяет существенно уменьшить ошибку отклонения сформированной дифракционной картины от заданной (более чем в 10 раз в нашем примере) при несущественном снижении энергетической эффективности (на 5% в нашем примере).

6. Градиентный алгоритм расчета фазы ДОЭ

В работе /1/ можно найти обзор алгоритмов нелинейного программирования для задач восстановления изображений. Задача сводится к минимизации целевых функционалов с учетом ограничений, накладываемых на функции, входящие в задачу. Если результирующий функционал с учетом ограничений можно представить в виде суммы линейного и квадратичных функционалов, то решение задачи находится аналитически. В противном случае требуется создавать вычислительные алгоритмы. Среди них можно выделить следующие: метод прямой оптимизации, метод градиентного спуска, метод наискорейшего спуска, метод сопряженных градиентов. Последний из перечисленных методов имеет наилучшую сходимость. Еще более быструю сходимость демонстрирует метод модифицированных функций Лагранжа, согласно которому в функционал добавляются слагаемые, в которых функции ограничений взяты в квадрате. Однако для выполнения одной итерации градиентных методов требуется выполнить 3 — 5 интегральных преобразований типа свертки или Фурье-преобразования. Поэтому чем сложнее алгоритм, тем больше требуется временных затрат при его компьютерной реализации. Применение градиентных алгоритмов в задачах расчета ДОЭ имеет ряд особенностей, некоторые из которых будут рассмотрены в этом разделе.

Итак, следуя работе /5/ рассмотрим применение градиентного метода наискорейшего спуска для задачи расчета фазы ДОЭ.

Будем полагать, что переменные x и $u = 0,1,2,\ldots,N-1$ представляют собой массивы дискретных координат. Дискретизация переменных является необходимым шагом для проведения компьютерных расчетов. Тогда функционал (11) преобразуется к виду

$$\varepsilon = \sum_{u} \left(\left| G(u) \right| - \left| F(u) \right| \right)^2 , \tag{81}$$

где $|F(u)|$ и $|G(u)|$ — заданная и рассчитанная амплитуды света в плоскости пространственного спектра, а преобразование Фурье (12) примет вид дискретного преобразования Фурье (ДПФ):

$$G(u) = \Im\{g(x)\} = \sum_{u=0}^{N-1} g(x) \exp\left[-i 2\pi xu / N\right]. \tag{82}$$

Обратное ДПФ имеет вид:

$$g(x) = \Im^{-1}\{G(u)\} = \frac{1}{N}\sum_{u=0}^{N-1} G(u) \exp\left[i 2\pi xu / N\right]. \tag{83}$$

Градиент функционала (81) равен сумме частных производных данного функционала от каждого неизвестного параметра p (это может быть фаза ДОЭ, амплитуда освещающего пучка и т.д.):

$$\begin{aligned}
&\frac{\partial \varepsilon}{\partial p} = 2\sum_u \left[\,|G(u)| - |F(u)|\,\right] \frac{\partial |G(u)|}{\partial p} = \\
&= \sum_u \left[G^*(u) - \frac{|F(u)|}{|G(u)|} G^*(u) \right] \frac{\partial G(u)}{\partial p} + k.c. = \\
&= -\sum_u G^{w*}(u) \frac{\partial G(u)}{\partial p} + k.c.
\end{aligned} \tag{84}$$

где * — знак комплексного сопряжения, $k.c.$ означает комплексно сопряженное предыдущее слагаемое, а функция $G^w(u)$ имеет вид

$$G^w(u) = |F(u)| \frac{G(u)}{|G(u)|} - G(u). \tag{85}$$

В силу линейности преобразование Фурье можно найти связь производных в плоскости ДОЭ и плоскости Фурье-спектра:

$$\frac{\partial G(u)}{\partial p} = \frac{\partial}{\partial p}\left[\Im\{g(x)\}\right] = \Im\left\{\frac{\partial g(x)}{\partial p}\right\}. \tag{86}$$

Подставляя это выражение в (84) и производя некоторые преобразования, получаем

$$\frac{\partial \varepsilon}{\partial p} = -\sum_{u} G^{w*}(u)\ \Im\left\{\frac{\partial g(x)}{\partial p}\right\} + k.c. =$$
$$= -2\,\mathrm{Re}\left\{\sum_{x}\left[\frac{\partial g(x)}{\partial p}\right]^{*}\ \Im^{-1}\left\{G^{w}(u)\right\}\right\} = \qquad (87)$$
$$== -2\,\mathrm{Re}\left\{\sum_{x}\frac{\partial g(x)}{\partial p}\left(\Im^{-1}\left\{G^{w}(u)\right\}\right)^{*}\right\}$$

Вычисление выражения (87) предусматривает выполнение преобразования Фурье от $g(x)$, чтобы получить $G(u)$; расчет $G^{w}(u)$ по формуле (85); его комплексное сопряжение; выполнение обратного преобразования Фурье и скалярное перемножение результата с функцией $\partial g(x)/\partial p$.

Рассмотрим теперь конкретный случай, когда параметр p представляет собой отсчеты фазовой функции $\theta(x)$. Тогда имеем ($g(x) = |g(x)|\ \exp[\ \theta(x)\ ]$):

$$\frac{\partial g(x')}{\partial \theta(x)} = \mathrm{i}\, g(x)\ \ \delta(x', x), \qquad (88)$$

где

$$\delta(x', x) = \begin{cases} 1, & x = x', \\ 0, & x \neq x', \end{cases} \qquad (89)$$

— функция Кронекера. Подставляя выражение (88) в уравнение (84), получаем:

$$\frac{\partial \varepsilon}{\partial \theta(x)} = -2\,\mathrm{Re}\left\{\sum_{x}\mathrm{i}\, g(x')\delta(x', x)\, g^{w*}(x)\right\} = 2\,\mathrm{Im}\left\{g(x)\ g^{w*}(x)\right\}, \qquad (90)$$

где

$$g^{w}(x) = \Im^{-1}\left\{G^{w}(u)\right\}. \qquad (91)$$

Вычисленный градиент может быть использован в любом градиентном методе расчета фазовой функции $\theta(x)$. Например, в методе наискорейшего спуска. В этом методе, стартуя с какой-либо оценки фазы $\theta_0(x)$, каждое последующее приближение находится из итеративного уравнения

$$\theta_{n+1}(x) = \theta_n(x) + s\frac{\partial\varepsilon}{\partial\theta_n(x)}, \tag{92}$$

(n — номер итерации) путем одномерной минимизации по s целевой функции ε как функции от $\theta_{n+1}(x)$.

Чтобы избежать одномерной минимизации по s целевой функции ε, следуя /5/, получим оценку оптимального шага s, исходя из разложения Тейлора первого порядка ε как функции от $\theta(x)$ около точки $\theta_n(x)$:

$$\varepsilon \approx \varepsilon_n + \sum_x \frac{\partial\varepsilon}{\partial\theta_n(x)}\left[\,\theta(x) - \theta_n(x)\,\right]. \tag{93}$$

Это выражение равно нулю при

$$\theta(x) - \theta_n(x) = -\varepsilon_n \frac{\partial\varepsilon}{\partial\theta_n(x)}\left\|\frac{\partial\varepsilon}{\partial\theta_n(x)}\right\|^{-2}. \tag{94}$$

Таким образом, при допущении (93) оптимальная величина шага градиентного метода записывается в виде

$$s_n = -\varepsilon_n\left\|\frac{\partial\varepsilon}{\partial\theta_n(x)}\right\|^{-2}. \tag{95}$$

Для увеличения скорости сходимости алгоритма можно использовать и другие градиентные методы. Например, метод сопряженного градиента. Для его варианта Флетчера — Ривза каждое последующее приближение выглядит так /26/:

$$\theta_{n+1}(x) = \theta_n(x) + sd_n, \tag{96}$$

$$d_{n+1} = -\frac{\partial \varepsilon}{\partial \theta_{n+1}(x)} + \beta_n d_n, \tag{97}$$

$$\beta_n = \frac{\left\| \frac{\partial \varepsilon}{\partial \theta_{n+1}(x)} \right\|^2}{\left\| \frac{\partial \varepsilon}{\partial \theta_n(x)} \right\|^2}, \tag{98}$$

где $\left\| \frac{\partial \varepsilon}{\partial \theta_n(x)} \right\|^2 = \sum_{n=0}^{N-1} \left(\frac{\partial \varepsilon}{\partial \theta_n(x)} \right)^2$ — евклидова норма.

7. Применение итеративных алгоритмов для расчета ДОЭ

В этом разделе будут рассмотрены конкретные реализации итеративных алгоритмов, приведенных выше, для расчета различных типов ДОЭ.

В разделе 7.1 при расчете ДОЭ, формирующих дифракционные картины, обладающие радиальной симметрией (круг, набор колец) характерной особенностью итеративных алгоритмов является использование прямого и обратного преобразований Ханкеля m-го порядка. Рассмотрен быстрый алгоритм для численного выполнения преобразования Ханкеля.

В разделе 7.2 задачи расчета радиально-симметричных ДОЭ, формирующих продольные световые отрезки, и расчета многофокусных линз сведены к одномерным задачам фокусировки света в поперечный отрезок и расчета многопорядковой дифракционной решетки. В разделе 7.3 алгоритм расчета ДОЭ адаптирован к расчету фазы радиального элемента в виде кусочно-постоянной функции. При этом поверхность ДОЭ разбивается на кольца, в каждом из которых фаза считается постоянной. Бинарные и многоуровневые дифракционные решетки рассчитываются в разделах 7.4 и 7.5 с помощью градиентных методов.

В разделе 7.6 обсуждаются особенности итеративных алгоритмов для расчета ДОЭ, фокусирующих излучение в объем или на заданную поверхность. В отдельных разделах рассмотрены две важные с практической точки зрения задачи: выравнивание Гауссового профиля лазерного пучка (раздел 7.7) и фокусировка света в узкое и широкое кольцо (раздел 7.8).

Композиционные ДОЭ, предназначенные для фокусировки света в контурные изображения (например, набор букв для лазерной маркировки) рассматриваются в разделе 7.9. В разделах 7.10 и 7.11 обсуждаются модернизации градиентных итеративных методов для расчета ДОЭ с малым числом уровней фазы. Такие ДОЭ называются квантованными.

7.1. Расчет ДОЭ, фокусирующих в радиально-симметричные области Фурье-спектра

Чтобы рассчитать оптические элементы, формирующие в некоторой области пространства распределение интенсивности с круговой симметрией, оказывается полезным использовать полярные координаты (r,φ) и (ρ,θ) в плоскости ДОЭ и плоскости наблюдения соответственно:

$$\begin{aligned} u &= r\cos\varphi, & \xi &= \rho\cos\theta, \\ v &= r\sin\varphi, & \eta &= \rho\sin\theta. \end{aligned} \tag{99}$$

В этом случае интегральное преобразование Френеля (1) в полярных координатах будет иметь вид

$$F(\rho,\theta) = \frac{\mathrm{i}k}{2\pi z} e^{\mathrm{i}kz} \int_0^{\infty}\int_0^{2\pi} W(r,\varphi) \times \\ \times \exp\left[\frac{\mathrm{i}k}{2z}\left\{r^2+\rho^2-2r\rho\cos(\varphi-\theta)\right\}\right] r\,\mathrm{d}r\,\mathrm{d}\varphi. \tag{100}$$

Чтобы модуль комплексной амплитуды $F(\rho,\theta)$ в плоскости наблюдения не зависел от азимутального угла θ, ограничим наше рассмотрение набором функций, которые выражаются в виде

$$W(r,\varphi) = g(r)e^{\mathrm{i}m\varphi}, \qquad m = 1,2,3,\ldots \tag{101}$$

Тогда вместо уравнения (100) получим

$$F(\rho,\theta) = \mathrm{i}^{m+1} kz^{-1} e^{\mathrm{i}kz} e^{\mathrm{i}m\theta} \int_0^{\infty} g(r) e^{\frac{\mathrm{i}k}{2z}\left(r^2+\rho^2\right)} J_m\left(\frac{kr\rho}{z}\right) r\,\mathrm{d}r, \tag{102}$$

где

$$J_m(x) = \frac{(-\mathrm{i})^m}{2\pi}\int_0^{2\pi} e^{\,\mathrm{i}(mt+x\cos t)}\mathrm{d}t \tag{103}$$

— функция Бесселя первого рода m-го порядка.

Из уравнения (102) следует, что для нахождения фазы ДОЭ в рамках скалярной теории дифракции Френеля необходимо решить следующее интегральное уравнение:

$$I_0(\rho) = \left|F(\rho,\theta)\right|^2 = \left| kz^{-1} \int_0^R A_0(r)\, e^{i\varphi(r)}\, e^{\frac{ikr^2}{2z}} J_m\left(\frac{kr\rho}{z}\right) r\, \mathrm{d}r \right|^2, \qquad (104)$$

где $I_0(\rho)$ — радиальное распределение интенсивности изображения, которое должно быть сформировано на расстоянии z от плоскости ДОЭ, $A_0(r)$ — радиальное распределение амплитуды освещающего пучка, $\varphi(r)$ искомая фазовая функция ДОЭ. Ниже рассматриваются итеративные методы решения уравнения (104) для формирования заданного изображения в плоскости пространственного спектра.

Из уравнения (104) следует, что для итеративного поиска фазовой функции ДОЭ необходимо вычислить прямое и обратное преобразования Ханкеля:

$$H(\rho) = \frac{k}{z} \int_0^\infty h(r)\, J_m\left(\frac{k}{z} r\rho\right) r\, \mathrm{d}r, \qquad (105)$$

$$h(r) = \frac{k}{z} \int_0^\infty H(\rho) J_m\left(\frac{k}{z} r\rho\right) \rho\, \mathrm{d}\rho. \qquad (106)$$

Для быстрого расчета преобразований Ханкеля (105) и (106) можно воспользоваться методом экспоненциальной замены переменных /27/. Этот метод позволяет при помощи экспоненциальной замены переменных преобразование Ханкеля (ПХ) свести к корреляции, которую можно вычислить при помощи преобразования Фурье.

Действительно, после замены переменных:

$$r = r_0 \mathrm{e}^x, \quad \rho = \rho_0 \mathrm{e}^y, \qquad (107)$$

где r_0 и ρ_0 — постоянные, вместо (105) мы получим

$$\bar{H}(y)=\frac{k}{z}r_0^2\int_{-\infty}^{\infty}\bar{h}(x)S(x+y)\,e^{2x}\,\mathrm{d}\,x, \tag{108}$$

где

$$\bar{h}(x)=h\left(r_0\,e^{x}\right),\qquad S(x+y)=J_m\left(\frac{k}{z}r_0\rho_0\,e^{x+y}\right),$$
$$\bar{H}(y)=H\left(\rho_0\,e^{y}\right) \tag{109}$$

Функция Бесселя при $m = 0$ стремится к единице при $x \to -\infty$. Поэтому чтобы функция $S(x)$ стремилась к нулю при $x \to \pm\infty$, её можно умножить на exp $(x/4)$, а чтобы подынтегральное выражение в (108) не изменилось, следует также умножить функцию $h(x)$ на exp $(-x/4)$.

Т.к. функция пропускания $h(r)$ ДОЭ ограничена апертурой $r \in [0,R]$, где R — радиус апертуры, то для функции $\bar{h}(x)$ при $x \to \pm\infty$ не возникает расходимостей. После переобозначения получаются выражения

$$h_1(x)=\bar{h}(x)r_0^2\,e^{7x/4},\qquad S_1(x+y)=S(x+y)\exp\left(\frac{x+y}{4}\right),$$
$$H_1(y)=\bar{H}(y)e^{y/4}, \tag{110}$$

и вместо (105) получаем интеграл корреляции:

$$H_1(y)=\frac{k}{z}\int_{-\infty}^{a}h_1(x)S_1(x+y)\,\mathrm{d}\,x,\quad a=\ln\frac{R}{r_0}. \tag{111}$$

Интеграл (111) можно выразить через преобразование Фурье в виде:

$$H(\rho)=\frac{k}{z}\left[\frac{\rho}{\rho_0}\right]^{-1/2}\int_{-\infty}^{\infty}P(-w)U(w)\,e^{\mathrm{i}\,w\ln(\rho/\rho_0)}\,\mathrm{d}\,w, \tag{112}$$

где $P(w)$ — это Фурье-образ функций $h_1(x)$, $U(w)$ — Фурье-образ $S_1(y)$. Аналогичным образом можно представить и обратное преобразование

Ханкеля, уравнение (106). Отметим, что существуют и другие методы быстрого расчета ПХ /28,29/.

Возвращаясь к решению уравнения (104), объединим обе функции в показателе степени в одну функцию, которую следует найти

$$\overline{\varphi}(r) = \varphi(r) + \frac{k}{2z} r^2 . \qquad (113)$$

Процедура итеративного решения уравнения (104) реализуется аналогично алгоритму ГС (6) и последовательно использует пару преобразований, состоящую из прямого и обратного преобразований Ханкеля. При этом осуществляется ограничивающая замена переменных как в плоскости оптического элемента так, и плоскости наблюдения, на n-м шаге итераций. Функция $g_n(r)$, рассчитанная в плоскости ДОЭ, заменяется на функцию $\overline{g}_n(r)$ согласно правилу

$$\overline{g}_n(r) = \begin{cases} A_0(r)\, g_n(r) \left| g_n(r) \right|^{-1}, & r \in [0, R], \\ 0, & r \in [0, R]. \end{cases} \qquad (114)$$

в то время как рассчитанная амплитуда $F_n(\rho)$ в плоскости наблюдения заменяется на $\overline{F}_n(\rho)$, при помощи соотношения

$$\overline{F}_n(\rho) = \sqrt{I_0(\rho)}\, F_n(\rho) \left| F_n(\rho) \right|^{-1} . \qquad (115)$$

Между вышеуказанными заменами (114) и (115), и соответствующими заменами для двумерного алгоритма уменьшения ошибки, (6) и (8), нет существенной разницы. Чтобы увеличить скорость сходимости алгоритма, вместо уравнения (115) может быть использован АА-алгоритм (19), при соответствующем переходе от двумерного к радиальному случаю.

Данный алгоритм позволяет рассчитывать ДОЭ, предназначенные для формирования таких изображений, как круг, кольцо или набор колец /29,30/.

Пример 5. На рис. 13 представлена фаза (после 22 итераций) ДОЭ, способного преобразовывать лазерный свет с гауссовым распределением интенсивности в круг с однородной интенсивностью в фокусе линзы. Интенсивность гауссового пучка на границах апертуры ДОЭ составляет 0,1 от максимального её значения в центре пучка. Радиус ДОЭ равен R = 0,4 мм; число отсчетов вдоль радиуса 256, k/f = 100 мм$^{-2}$, f — фокальное расстояние линзы; радиус круга в плоскости изображения составляет 0,3 мм и примерно равен трем радиусам минимального дифракционного пятна (диск Эйри). На рис. 13б показано радиальное сечение фазы ДОЭ.

На рис. 14 показана дифракционная картина, которую формирует ДОЭ с такой фазой в фокальной плоскости (а) и её радиальное сечение (б). Среднеквадратичное отклонение интенсивности в сформированном круге, от постоянного значения составляет 6%, а энергетическая эффективность фокусировки в круг равна 91%.

В нашем случае расчет осуществлялся при помощи преобразования Ханкеля (для m = 0, преобразования Фурье — Бесселя).

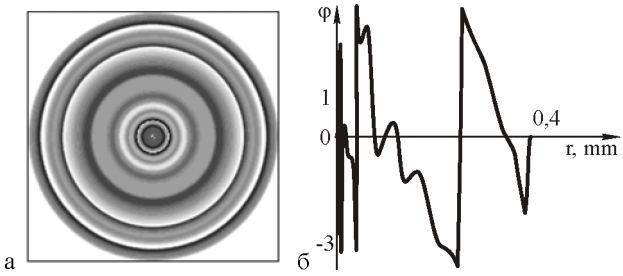


Рис. 13. Фаза фокусатора в круг: фаза ДОЭ (а), её радиальное сечение (б)

Если необходимо рассчитать ДОЭ с радиальной фазой, формирующий распределение интенсивности типа кольца, и если для такого расчета использовать преобразование Ханкеля нулевого порядка, то в центре кольца практически всегда будет формироваться пик интенсивности. Для его устранения надо использовать преобразование Ханкеля более высокого порядка.

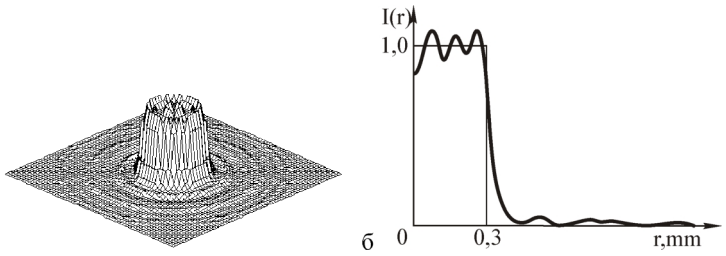


Рис 14. Сформированное изображение круга: дифракционная картина (а) и ее радиальное сечение(б)

Пример 6. На рис. 15 показан результат расчета ДОЭ в широкое кольцо постоянной интенсивности. В этом случае расчет проводился при помощи преобразования Ханкеля 8-го порядка (в уравнении (101) m = 8), радиус ДОЭ равнялся R =5 мм, радиусы кольца были r_1 = 0,1 мм и r_2 = 0,2 мм, число пикселов вдоль радиуса составило 256.

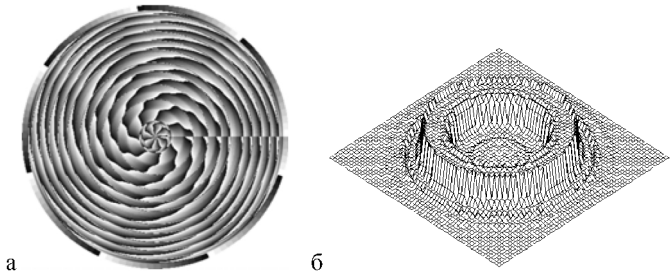


Рис. 15. Фаза ДОЭ в кольцо (а), дифракционная картина при m = 8 (б)

Энергетическая эффективность фокусировки в кольцо равнялась 83%, средне-квадратичное отклонение — 8%

Заметим, что азимутальный множитель в уравнении (101), в виде $e^{im\varphi}$, сам по себе представляет интерес как функция пропускания фазового оптического элемента, используя который, можно осуществлять множество оптических преобразований: преобразование Ханкеля m-го порядка /31/, формирование пучков Бесселя высших порядков /32/, оптическое дифференцирование радиально-

симметричных световых пучков /33/, а также радиальный аналог преобразования Гильберта /34/.

7.2. Расчет дифракционных аксиконов, формирующих осевые световые отрезки

В /35/ рассмотрен новый оптический элемент — аксикон, который представляет собой конус из прозрачного материала. Такой элемент используется для формирования узкого осевого светового пучка, диаметр поперечного сечения которого пропорционален углу при вершине конуса, а длина которого пропорциональна диаметру освещающего пучка и обратно пропорциональна углу при вершине конуса (световой пучок должен падать перпендикулярно основанию конуса). Вместе со сферической линзой аксикон используется для формирования узкого светового кольца /36/.

Ниже рассматриваются ДОЭ, по своим свойствам подобные аксиконам.

Фазовые оптические элементы, способные в комбинации со сферической линзой формировать световое поле с увеличенной глубиной резкости, называются фокусаторами в продольный отрезок /37/ и могут рассматриваться как обобщенные аксиконы /36/. Они нашли широкое практическое применение для изучения разрядов в газах /38/, для записи и считывания данных с оптических дисков /39/, для тестирования поверхностей /40/.

Ниже рассматриваются итеративные алгоритмы расчета ДОЭ (обобщенных аксиконов), которые могут формировать световые пучки заданного распределения интенсивности вдоль оптической оси.

Пусть требуется рассчитать радиальную фазовую функцию оптического элемента, который формирует осевой световой отрезок, ограниченный точками z_1 и z_2 и характеризующийся заданным распределением интенсивности вдоль оси $I_0(z)$, $z \in [z_1, z_2]$. В этом случае уравнение (102) примет вид:

$$F(\rho,z)=\frac{\mathrm{i}k}{z}e^{\mathrm{i}kz}e^{\mathrm{i}k\rho^2/2z}\int_0^R\exp\left[\mathrm{i}\varphi(r)+\mathrm{i}\frac{kr^2}{2z}\right]J_0\left(\frac{kr\rho}{z}\right)r\,\mathrm{d}r\,, \tag{116}$$

где R — радиус ДОЭ, $\varphi(r)$ — заданная фаза, $J_0(x)$ — функция Бесселя нулевого порядка. В уравнении (116) предполагается, что оптический элемент освещен плоской световой волной.

Так как известна интенсивность только вдоль оси z, можно положить в уравнении (116) $\rho = 0$ и получить (опуская тривиальный сомножитель exp (ikz)) уравнение:

$$F(0,z)=\frac{k}{z}\int_0^R\exp\left[\mathrm{i}\varphi(r)+\mathrm{i}\frac{kr^2}{2z}\right]r\,\mathrm{d}r\,. \tag{117}$$

Переобозначив

$$\xi=\frac{k}{z},\quad x=\frac{r^2}{2}\,, \tag{118}$$

перепишем уравнение (117) в виде

$$\frac{1}{\xi}F(\xi)=\int_0^a e^{\mathrm{i}\varphi(x)+\mathrm{i}x\xi}\,\mathrm{d}x,\qquad a=\frac{R^2}{2}\,. \tag{119}$$

Из уравнения (119) видно, что заданная функция пропускания ДОЭ-аксикона $\exp\left[i\varphi(r)\right]$ связана с осевой комплексной амплитудой света $F(\xi)$ через одномерное преобразование Фурье. Следовательно, проблема сводится к решению одномерного интегрального уравнения

$$\frac{1}{\xi^2}I_0(\xi)=\left|\int_0^a e^{\mathrm{i}\varphi(x)+\mathrm{i}x\xi}\,\mathrm{d}x\right|^2\,. \tag{120}$$

В этом случае применим алгоритм уменьшения ошибки и все его модернизации.

Если требуется рассчитать аксикон в качестве дополнения к сферической линзе, то фаза в уравнении (117) должна быть представлена в виде

$$\varphi(r) = \varphi_0(r) - \frac{kr^2}{2f}, \tag{121}$$

где f — фокальное расстояние сферической линзы. В этом случае вместо (117) можно записать

$$F(0,z) = \frac{k}{z}\int_0^R \exp\left[\mathrm{i}\varphi_0(r) + \mathrm{i}\frac{kr^2}{2}\left(\frac{1}{z} - \frac{1}{f}\right)\right] r\,\mathrm{d}r. \tag{122}$$

Уравнение (122) относительно $\varphi_0(r)$ будет эквивалентно уравнению (119), если использовать следующую замену переменных

$$\xi = k\left(\frac{1}{z} - \frac{1}{f}\right), \quad x = \frac{r^2}{2}. \tag{123}$$

Чтобы сформировать короткие световые отрезки (при условии, что $2d = z_2\text{-}z_1 << f$, где $2d$ — глубина фокуса линзы или длина заданного светового отрезка), вместо уравнения (122) можно использовать более простое выражение

$$F(\Delta z) = \frac{k}{f}\int_0^R \exp\left[\mathrm{i}\varphi_0(r) + \mathrm{i}\frac{k\Delta z}{2f^2}r^2\right] r\,\mathrm{d}r, \tag{124}$$

где $\Delta z \in [-d, d]$, $f = (z_1+z_2)/2$.

После замены переменных

$$\xi = \frac{k\Delta z}{f^2}, \qquad x = \frac{r^2}{2},$$

получается уравнение для поиска фазы ДОЭ

$$I_0(\xi) = \left|\frac{k}{f}\int_0^a e^{\mathrm{i}\varphi_0(x) + \mathrm{i}x\xi}\,\mathrm{d}x\right|^2. \tag{125}$$

Уравнение (125) может быть решено при помощи итеративной процедуры /41,42,43/, при этом степень близости оценочной функции $I_n(\xi) = |F_n(\xi)|^2$ к заданной функции $I_0(\xi)$ определяется из среднеквадратичного отклонения

$$\delta = \left[\int_{\xi_1}^{\xi_2} \left| I_n(\xi) - I_0(\xi) \right|^2 \mathrm{d}\xi \right]^{1/2} \left[\int_{\xi_1}^{\xi_2} I_0^2(\xi)\, \mathrm{d}\xi \right]^{-1/2}, \quad \xi_1 = \frac{k}{z_1}, \quad \xi_2 = \frac{k}{z_2}. \quad (126)$$

Энергетическая эффективность формирования осевого светового отрезка находится из выражения

$$E_k = \int_0^{\rho_0} I(\rho, z_k)\, \rho\, \mathrm{d}\rho \Big/ \int_0^{\infty} I(\rho, z_k)\, \rho\, \mathrm{d}\rho, \quad (127)$$

где $I_0(\rho, z_k)$ — интенсивность, рассчитанная в плоскости, находящейся на расстоянии z_k от ДОЭ, а ρ_0 — радиус первого локального минимума интенсивности $I(\rho, z_k)$. Функция E_k определяет порцию световой энергии, которая идет на формирование минимального дифракционного пятна в заданной плоскости наблюдения.

Заметим, что если заданная интенсивность $I_0(z)$ выражается в виде

$$I_0(z) = I_0 \sum_{n=1}^{N} \delta(z - z_n), \quad (128)$$

где $\delta(x)$ — функция Дирака, то решение уравнения (125), $\varphi_0(r)$, будет представлять собой фазу N-фокусной линзы с N продольными фокусами равной интенсивности I_0.

Пример 7. На рис. 16*а* представлена фаза ДОЭ, рассчитанная за 100 итераций, при этом для решения уравнения (125) был использован алгоритм ГС. Такой ДОЭ создает световой пучок с осевым распределением интенсивности, имеющим форму прямоугольного импульса (рис. 16*б*). Параметры расчета: $N = 256$ — общее число пикселов; $R = 6$ мм — радиус ДОЭ; $f = 400$ мм — фокусное расстояние линзы, которая в окрестности своей фокальной плоскости содержит сформированный отрезок длиной $2a = 40$ мм. Длина волны 0,63 мкм. Предполагается, что освещающий световой пучок плоский. Ошибка (126) равна 2%. Эффективность фокусировки (127) изменяется вдоль отрезка сложным образом и лежит в пределах от 15% до 85%. Длина отрезка

равна примерно восьми длинам Френеля. На рис.15в показано трехмерное распределение интенсивности.

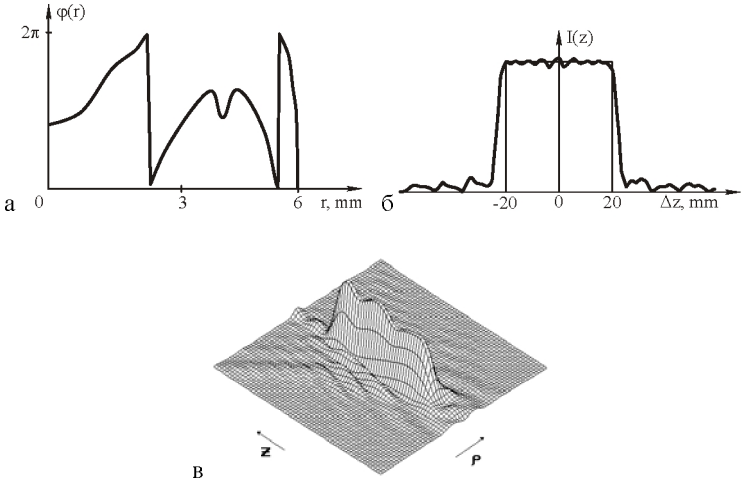


Рис. 16. ДОЭ, фокусирующий в осевой отрезок: радиальное сечение фазы (а), осевая интенсивность (б), распределение интенсивности в плоскости (z,ρ) (в)

На рис. 17а показана фаза ДОЭ в виде амплитудной маски, которая формирует 10 продольных фокусов равной интенсивности (рис. 17б).

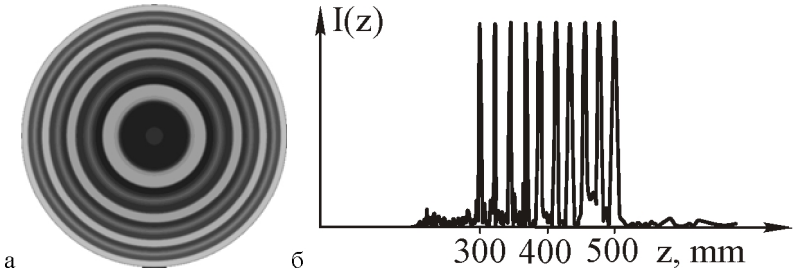


Рис. 17. Расчет многофокусной линзы: фаза ДОЭ(а), осевое распределение интенсивности (б)

Фаза многофокусной линзы приближенно представляет собой сумму фаз двух разных аксиконов или двух разных ДОЭ, формирующих продольные отрезки разной длины. В результате этого в пространстве распространяются две конических волны с разным наклоном к опти-

ческой оси. В результате интерференции этих волн и появляются продольные фокусы.

7.3. Расчет радиально симметричных ДОЭ с квантованной фазой

В данном пункте рассматривается итеративный алгоритм, который позволяет осуществлять расчет радиально-симметричных ДОЭ с небольшим числом уровней фазы, что дает существенное преимущество при изготовлении ДОЭ методами фото- или электронной литографии. На практике такие ДОЭ могут быть использованы для расширения минимального дифракционного пятна. Эта проблема актуальна для лазерной печати. Конечно, двукратное расширение дифракционного пятна (диска Эйри) может быть осуществлено без применения ДОЭ, путем двукратного увеличения фокальной длины линзы или путем уменьшения радиуса апертуры линзы в два раза. Однако, первый подход приводит к нежелательным изменениям в конструкции, а второй — уменьшает эффективность фокусировки в 4 раза.

Предлагается использовать комбинацию «ДОЭ+линза», которая обеспечивает более эффективное решение данной проблемы /44/. Предположим, что требуется рассчитать фазу ДОЭ $\varphi(r)$, которая зависит только от радиальной переменной, и что данный ДОЭ создает заданное распределение интенсивности $I_0(\rho)$ в фокальной плоскости линзы. В этом случае, вместо уравнения (104) можно записать следующее уравнение для расчета фазы $\varphi(r)$:

$$I_0(\rho) = \left| \frac{k}{f} \int A_0(r)\, e^{i\varphi(r)} J_0\left(\frac{kr\rho}{f}\right) r \, dr \right|^2, \qquad (129)$$

где $A_0(r)$ — амплитуда освещающего пучка, f — фокальное расстояние линзы. Предположим далее, что выполняется условие $A_0(r) = 1$. Будем искать функцию $\varphi(r)$ в кусочно-постоянном виде (рис. 18). Тогда можно представить комплексную амплитуду света $F(\rho)$ в фокаль-

ной плоскости линзы как сумму амплитуд, получающихся в результате дифракции плоской волны на кольцевой апертуре

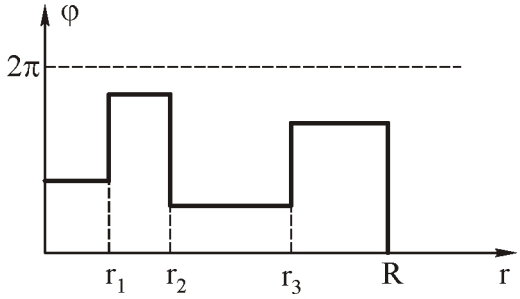


Рис. 18. Кусочно-постоянная фаза

$$F(\rho)=2\pi\sum_{n=0}^{N-1}e^{i\varphi_{n+1}}\left[r_{n+1}\frac{J_1(k\rho r_{n+1}/f)}{\rho}-r_n\frac{J_1(k\rho r_n/f)}{\rho}\right], \tag{130}$$

где φ_{n+1} — это значение $\varphi(r)$ при $r\in[r_n, r_{n+1}]$, $n=0, 1,..., N-1$; $r_0=0$, $r_N=R$, R — радиус ДОЭ, N — число точек деления радиуса ДОЭ. После соответствующего преобразования членов уравнения (130) получим

$$F(\rho)=\frac{2\pi}{\rho}\sum_{n=1}^{N}C_nJ_1\left(\frac{\gamma_n\rho}{a}\right), \tag{131}$$

где

$$\begin{cases}C_n=r_n\left[e^{i\varphi_n}-e^{i\varphi_{n+1}}\right],\\ C_N=\mathrm{R}e^{i\varphi_N}\end{cases}, \tag{132}$$

γ_n — нули функции Бесселя первого порядка: $J_1(\gamma_n)=0$. Из сравнения уравнений (130) и (131) следует, что точки деления r_n должны удовлетворять следующему условию

$$r_n=\frac{\gamma_n f}{ka},\quad a\ge R, \tag{133}$$

где a — это константа, контролирующая масштаб деления ДОЭ на кольца.

Введя обозначения $\overline{F}(\rho)=(2\pi)^{-1}\rho F(\rho)$, вместо (131) получим

$$\overline{F}(\rho)=\sum_{n=1}^{N}C_n J_1\left(\frac{\gamma_n\rho}{a}\right). \tag{134}$$

Используя свойство ортогональности функции Бесселя

$$\int_0^1 J_m(\gamma_p x)\,J_m(\gamma_q x)\,x\,\mathrm{d}x=\begin{cases}\left[J_m'(\gamma_p)\right]^2/2, & p=q,\\ 0, & p\neq q,\end{cases} \tag{135}$$

где $J_m'(\gamma_p)$ — производная функции Бесселя в точке γ_p, сумму (134) можно представить в виде

$$C_n=2\left[aJ_1'(\gamma_n)\right]^{-2}\int_0^a \overline{F}(\rho)\,J_1\left(\frac{\gamma_n\rho}{a}\right)\rho\,\mathrm{d}\rho\,. \tag{136}$$

Соотношения (134) и (136) позволяют организовать следующую итеративную процедуру для решения уравнения (129).

1. Начальная оценка $\overline{F}(\rho)$ комплексной амплитуды выбирается в виде

$$\overline{F}(\rho)=\sqrt{I_1(\rho)}\,e^{\mathrm{i}\psi_0(\rho)}, \tag{137}$$

где $I_1(\rho)=(2\pi)^{-2}\rho^2 I_0(\rho)$, $I_0(\rho)$ — заданная интенсивность света в фокальной плоскости линзы, $\psi_0(\rho)$ — случайная оценка фазы.

2. На k-м шаге итераций функцию $\overline{F}_k(\rho)$ представляют как сумму (134) с коэффициентами $C_n^{(k)}$, которые находятся из уравнения (136) и подвергаются замене

$$\overline{C}_n^{(k)}=B_n e^{\mathrm{i}\nu_{nk}}, \qquad n=\overline{1,N}, \tag{138}$$

где

$$B_n = 2r_n \sin\left[\frac{\varphi_n - \varphi_{n+1}}{2}\right], \tag{139}$$

$$\begin{cases} \varphi_n = \pi - \varphi_{n+1} + 2\nu_{nk}, & n = \overline{N-1,1}, \\ \varphi_N = \nu_{Nk}, \end{cases} \tag{140}$$

где $\nu_{nk} = \arg C_n^{(k)}$.

Фазы φ_n ищутся при помощи рекуррентных соотношений (140), вытекающих из уравнения (132).

3. Формируется (k+1)-я оценка функции $\overline{F}_{k+1}(\rho)$

$$\overline{F}_{k+1} = \sum_{n=1}^{N} \overline{C}_n^{(k)} J_1\left(\frac{\gamma_n \rho}{a}\right). \tag{141}$$

4. Рассчитанная функция $\overline{F}_{k+1}(\rho)$ заменяется функцией $\widehat{F}_{k+1}(\rho)$ по правилу

$$\widehat{F}_{k+1}(\rho) = \sqrt{I_1(\rho)}\, \overline{F}_{k+1}(\rho) \left|\overline{F}_{k+1}(\rho)\right|^{-1}. \tag{142}$$

5. Осуществляется переход к шагу 2, и т. д.

Данный алгоритм не сходится: с ростом числа итераций ошибка изменяется квазипериодически. На рис. 19 показана среднеквадратичная ошибка в зависимости от числа итераций для примера (8). Из рис.19 видно, что функция ошибки имеет несколько локальных минимумов.

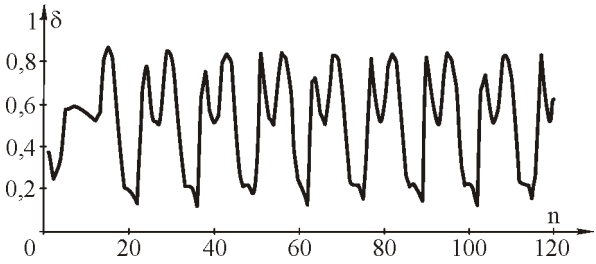


Рис. 19. Ошибка в зависимости от числа итераций

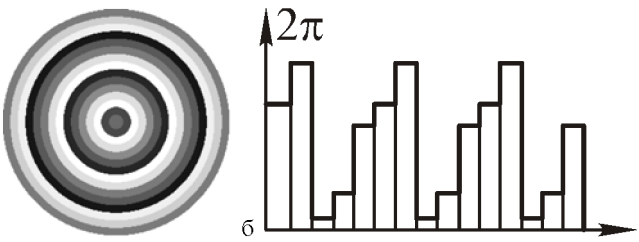
Пример 8. Рассчитаем ДОЭ с квантованной фазой, который фокусирует свет в кольцо /45/. Параметры расчета следующие: $k/f = 100$ мм^2, $R = 0{,}5$ мм — радиус ДОЭ, дискретность вдоль ρ равна 2 мкм, общее число отсчетов 256. Радиусы кольца равны $R_1 = 0{,}3$ мм, $R_2 = 0{,}5$ мм. На рис. 20 показан результат расчета при разбиении радиуса на несколько колец (N = 15) после 12 итераций: двумерное распределение фазы(а), её радиальное сечение (б) (оба с пятью уровнями квантования), дифракционная картина в фокальной плоскости линзы (в), её радиальное сечение (г) (кривая 2 соответствует рассчитанной интенсивности, кривая 1 — заданной интенсивности).

Значения энергетической эффективности фокусировки в кольцо 89% (рис. 20).

Отметим, что, в отличие от общепринятой практики /21/, в данном случае квантование фазы неэквидистантное. Действительно, обычно рассчитанную непрерывную функцию фазы $\varphi(x)$ подвергают эквидистантному равномерному квантованию по N уровням:

$$\varphi_n = \left\{ \frac{2\pi}{N} n; \quad \frac{2\pi}{N} n < \varphi(x) < \frac{2\pi}{N}(n+1), \qquad n = 0,1,..(N-1) \right\}$$

В алгоритме (140) уже имеются N отсчетов фазы φ_n, связанные с числом колец радиального ДОЭ и они, как правило, неравномерные.



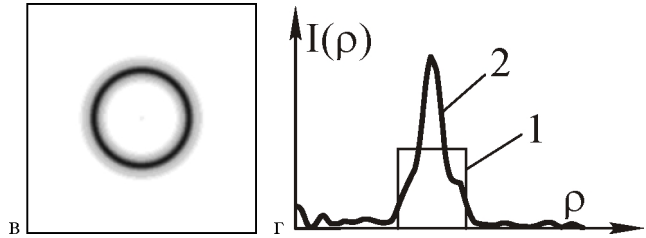


Рис. 20. Радиальный ДОЭ с пятью уровнями фазы: фаза ДОЭ (а), её радиальное сечение (б), дифракционная картина(в), её радиальное сечение (г)

7.4. Многопорядковые фазовые бинарные дифракционные решетки

Многопорядковые фазовые дифракционные решетки представляют собой дифракционную структуру с периодическим фазовым рельефом и предназначены для формирования одномерного или двумерного набора плоских пучков с заданным соотношением энергии между пучками. Многопорядковые дифракционные решетки применяются в оптических устройствах мультипликации изображений, в оптических соединителях для оптических волокон и для свободного пространства, в устройствах оптической связи и обработки информации, в когерентных оптических процессорах /46—51/.

Бинарные дифракционные решетки представляют особый интерес ввиду простоты их изготовления методами фотолитографии; бинарный рельеф формируется за одно химическое травление подложки.

Для расчета бинарных решеток используются, как правило, методы стохастической оптимизации, такие как прямой бинарный поиск /52, 53/, алгоритм симуляции отжига /54—58/ и итеративные алгоритмы /59—62/. В работах /63, 64/ расчет бинарных решеток основан на решении системы нелинейных уравнений, представляющих интенсивности порядков решетки через координаты фазового профиля. При этом расчет бинарной решетки с числом порядков $2N + 1$ сводится к решению нелинейной системы из N уравнений.

Для использования бинарных дифракционных решеток в качестве преобразователей лазерных пучков и формирователей пучков в сетях на оптических волокнах, является актуальным увеличение числа порядков бинарной решетки до нескольких сотен. С ростом числа порядков алгоритмы случайного поиска профиля дифракционной решетки требуют большого времени расчета и приводят к нестабильной сходимости, а методы расчета, основанные на решении системы нелинейных уравнений, становятся численно неустойчивыми. В данном пункте рассмотрен расчет многопорядковых бинарных решеток на основе градиентного алгоритма при специальном выборе начального приближения для профиля решетки. Выбор начального приближения основан на свойствах одномерного бинарного элемента для фокусировки в отрезок.

Фазовый профиль бинарной решетки состоит из K прямоугольных штрихов одинаковой высоты φ, но разной ширины (рис. 21). Координаты $x_1,\ldots,x_{2K}$ этих штрихов являются параметрами оптимизации и определяют значения интенсивности в дифракционных порядках.

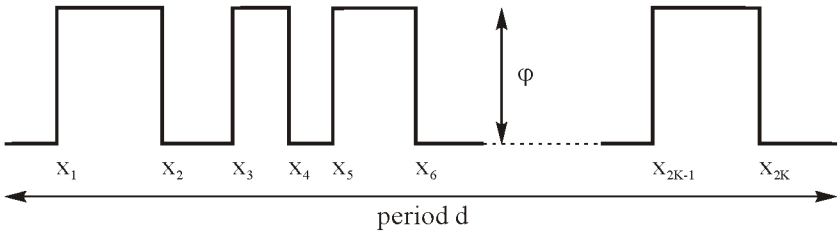


Рис. 21. Фазовая функция одного периода бинарной решетки

Для выбора начального приближения для профиля бинарной решетки, рассмотрим предварительно расчет одномерного фазового ДОЭ с апертурой d для фокусировки при $z = f$ в отрезок с распределением интенсивности $I(\xi)$ (см. рис. 22). Отрезок фокусировки $\xi \in [n_1\Delta, n_2\Delta]$, где $\Delta = \lambda f/d$ - размер дифракционного пятна, может быть представлен как совокупность $(n_2 - n_1 + 1)$ дифракционных пятен. Фазовая функция

$\varphi(x)$ $(0 \le x \le d)$ одномерного ДОЭ может быть рассчитана аналитически в приближении геометрической оптики /65—67/.

Для сходящегося цилиндрического освещающего пучка с фокусом f

$$W_0(x) = \exp\left(-\frac{\mathrm{i}kx^2}{2f}\right), \qquad k = \frac{2\pi}{\lambda},$$

геометрооптическая функция $\varphi(x)$ имеет вид:

$$\varphi(x) = \frac{2\pi}{d}(n_2 - n_1)\int_0^x \kappa(\alpha)\,\mathrm{d}\alpha + \frac{2\pi}{d}n_1 x, \tag{142}$$

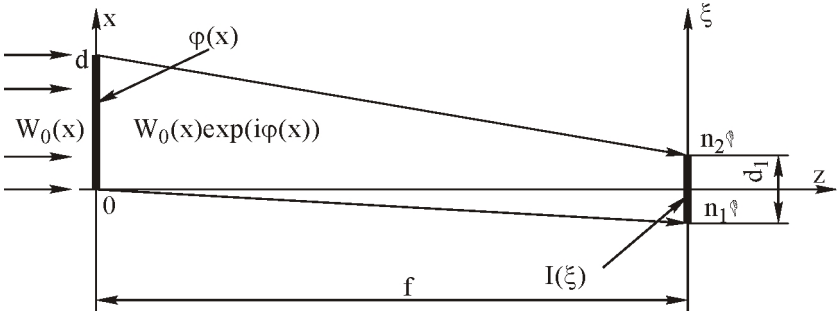


Рис. 22. Геометрия фокусировки в поперечный отрезок

Функция $\kappa(\alpha)$ в (142) ищется из решения следующего дифференциального уравнения

$$\frac{\mathrm{d}\kappa(\alpha)}{\mathrm{d}\alpha} = \left[I\left(\frac{\kappa(\alpha) - n_1\Delta}{(n_2 - n_1)\Delta}\right)\right]^{-1} \tag{143}$$

с граничными условиями: $\kappa(0) = 0$, $\kappa(1) = 1$. Функция $\varphi(x)$ соответствует фазовой функции небинарной дифракционной решетки, формирующей $(n_2 - n_1 + 1)$ дифракционных порядков. Действительно, в приближении Френеля-Кирхгофа значения интенсивности в точках $\xi = \xi_j = j\Delta$ в плоскости $z = f$

$$I\left(\xi_j\right)=\frac{1}{\lambda f}\left|\int_0^d \exp\left(\mathrm{i}\varphi(x)-\mathrm{i}\frac{2\pi j}{d}x\right)\mathrm{d}x\right|^2, \tag{144}$$

пропорциональны квадратам модулей коэффициентов Фурье в разложении функции $\exp(\mathrm{i}\varphi(x))$. Следовательно непрерывную функцию $\varphi(x)$ можно рассматривать как фазовую функцию дифракционной решетки с периодом d и интенсивностью дифракционных порядков

$$I_j=\begin{cases} I(j\Delta), & n_1\le j\le n_2, \\ 0, & \textit{иначе}. \end{cases} \tag{145}$$

Для расчета бинарной решетки, формирующей $2N + 1$ порядков в интервале $[-N\Delta, +N\Delta]$, определим бинарную фазовую функцию в виде /68, 69/:

$$\varphi_b(x)=\Phi[\overline{\varphi}(x)], \tag{146}$$

где $\overline{\varphi}(x)$ — непрерывная функция (142), (143) рассчитанная из условия формирования $(N+1)$ порядков в интервале $[0, N\Delta]$ и приведенная к интервалу $[0, 2\pi)$, а

$$\Phi[\eta]=\begin{cases} 0, & \eta\in[0,\pi), \\ \pi, & \eta\in[\pi,2\pi), \end{cases} \tag{147}$$

— бинарная функция, описывающая квантование фазы $\overline{\varphi}(x)$ по двум уровням. Для описания работы бинарной решетки (146), разложим функцию $\exp(\mathrm{i}\Phi[\eta])$ на интервале $[0, 2\pi)$ в ряд Фурье:

$$\exp\left(\mathrm{i}\Phi[\eta]\right)=\sum_{n=-\infty}^{\infty}c_n\exp(\mathrm{i}n\eta), \tag{148}$$

где

$$c_n = \begin{cases} \dfrac{1-(-1)^n}{\mathrm{i}\pi n}, & n \neq 0, \quad n = \pm 1, \pm 2, \ldots, \\ 0, & n = 0, \end{cases} \tag{149}$$

коэффициенты Фурье. Полагая $\eta = \bar{\varphi}(x)$ в (148), представим функцию комплексного пропускания бинарной решетки $\varphi_b(x)$ с учетом 2π-периодичности в виде:

$$\exp\left(\mathrm{i}\Phi\left[\bar{\varphi}(x)\right]\right) = \sum_{n=-\infty}^{\infty} c_n \exp\left(\mathrm{i}n\varphi(x)\right). \tag{150}$$

Уравнение (150) соответствует суперпозиции пучков с фазовыми функциями:

$$\varphi_n(x) = n\varphi(x). \tag{151}$$

Фазовые функции (151) порождены нелинейной операцией бинаризации (147) и также называются дифракционными порядками. Квадраты модулей коэффициентов Фурье (149) представляют распределение энергии между пучками (151). Поскольку $|c_1|^2 = |c_{-1}|^2 = 0{,}405$, то 81% энергии падающего пучка приходится на -1 и $+1$ дифракционные порядки. Изображения, формируемые в главных порядках -1 и $+1$, определяются функциями $-\varphi(x)$ и $+\varphi(x)$. Функции $-\varphi(x)$ и $+\varphi(x)$ являются непрерывными фазовыми функциями дифракционных решеток, формирующих по N порядков в интервалах $[-N\Delta, 0]$ и $[0, N\Delta]$, соответственно. Суммирование интервалов $[-N\Delta, 0]$ и $[0, N\Delta]$ дает $2N + 1$ порядков на интервале $[-N\Delta, N\Delta]$. Бинарная решетка (146) может рассматриваться лишь как приближенное решение. Действительно, функция $\varphi(x)$ была рассчитана в приближении геометрической оптики, а при анализе работы решетки (146) были рассмотрены только два главных дифракционных порядка без учета интерференции между ними. Поэтому, даже если эффективность, даваемая приближенным решением близка к 80% ($|c_1|^2 + |c_{-1}|^2 = 0{,}81$), то среднеквадратичная ошибка формирования заданной интенсивности порядков остается в пределах 40—50%. В дальнейшем приближенное ре-

шение (146) будет использоваться как начальное приближение профиля для градиентного алгоритма.

Согласно уравнениям (142) – (147), координаты начального профиля решетки для формирования $2N + 1$ порядков равной интенсивности, имеют вид:

$$x_n = \sqrt{\frac{n}{N}}, \quad n = \overline{1, N}. \tag{152}$$

Аналогично, для решетки с симметричным линейным распределением интенсивности в дифракционных порядках

$$I_n = I_0 + \alpha |n|, \qquad n = \overline{-N, N},$$

координаты начального профиля решетки определяются из решения следующего уравнения:

$$\frac{2N}{(b-1)(b^2-1)}\left[-3x_n(b^2-1) - 2 + 2\left(1 + (b^2-1)x_n\right)^{3/2}\right] = n, \quad n = 1, 2, \ldots, \tag{153}$$

где $b = 1 + \alpha N / I_0$.

Для оптимизации структуры решетки введем функцию ошибки, представляющую отличие между заданными значениями интенсивности в порядках и расчетными значениями I_j в виде:

$$\varepsilon(\mathbf{x}) = \sum_{j=-N}^{N} \left(I_j(\mathbf{x}) - \hat{I}_j\right)^2, \tag{154}$$

где $\mathbf{x} = (x_1, \ldots, x_{2K})$ — вектор координат границ штрихов. Без ограничения общности период решетки d будем считать равным единице. При $d = 1$ интенсивности порядков, соответствующие квадратам модулей Фурье-коэффициентов в разложении функции комплексного пропускания решетки, имеют вид /54, 64/:

$$\begin{cases} I_i(x) = \sin^2(\varphi/2)\,\dfrac{\left(C_i^2 + S_i^2\right)}{\left(\pi j\right)^2}, \quad j \neq 0, \quad j = \pm 1, \pm 2, \ldots \\ I_0(x) = 1 - 4Q(1-Q)\sin^2(\varphi/2), \end{cases} \tag{155}$$

где

$$C_j = \sum_{n=1}^{2K} (-1)^n \cos\left(2\pi j x_n\right), \quad S_j = \sum_{n=1}^{2K} (-1)^n \sin\left(2\pi j x_n\right),$$
$$Q = \sum_{n=1}^{2K} (-1)^n x_n. \tag{156}$$

Из уравнений (155) и (156) следует, что $I_j = I_{-j}$. Вследствие симметрии расчет бинарной решетки с числом порядков $2N + 1$ требует как минимум $N + 1$ свободных параметров; соответственно, число штрихов K на периоде должно быть не менее N /2 /54, 64/.

Для расчета профиля решетки будем использовать градиентный метод минимизации функции ошибки $\varepsilon(x)$. Основной проблемой, связанной с градиентными процедурами, является выбор начального профиля для получения стабильной сходимости. Как правило, при случайном выборе начального профиля, градиентные алгоритмы стагнируют при средне-квадратичной ошибке формирования заданных значений интенсивности порядков в 75—85%. Использование аналитического начального приближения (146), (152), (153) дает быструю и стабильную сходимость градиентных процедуры при высокой эффективности и низкой среднеквадратичной ошибке. Рассмотрим метод сопряженного градиента для минимизации функции ошибки $\varepsilon(x)$. Метод состоит в итерационной коррекции координат профиля решетки по правилу /70/:

$$\mathbf{x}_{n+1} = \mathbf{x}_n + \mathbf{h}\left(\mathbf{x}_n\right)t, \tag{157}$$

где $\mathbf{h}(\mathbf{x}_n) = -\mathbf{g}(\mathbf{x}_n) + \gamma_n \mathbf{h}(\mathbf{x}_{n-1})$, $\mathbf{g}(\mathbf{x}) = \left(\partial\varepsilon(\mathbf{x})/\partial x_1, \ldots, \partial\varepsilon(\mathbf{x})/\partial x_{2K}\right) = \left(g_1(\mathbf{x}), \ldots, g_{2K}(\mathbf{x})\right)$ — градиент функции $\varepsilon(\mathbf{x})$, t — шаг градиентного метода,

$$\gamma_n = -\frac{\left(\mathbf{g}(\mathbf{x}_n), \mathbf{g}(\mathbf{x}_n) - \mathbf{g}(\mathbf{x}_{n-1})\right)}{\left|\mathbf{g}(\mathbf{x}_{n-1})\right|^2}.$$

Согласно уравнениям (154) — (156) и соотношению симметрии, компоненты вектора градиента при $\varphi = \pi$ могут быть представлены в виде:

$$g_n(\mathbf{x}) = 16\frac{(-1)^n}{\pi}\sum_{p=1}^{N}\frac{1}{p}\left(S_p \cos(2\pi p x_n) - C_p \sin(2\pi p x_n)\right) \times \\ \times\left[I_p(\mathbf{x}) - \hat{I}_p\right] + 4(-1)^n(2Q-1)\left[I_0(\mathbf{x}) - \hat{I}_0\right], \tag{158}$$

где S_p, C_p и Q определены в (156).

Для расчета шага t градиентного метода, представим функцию ошибки вдоль сопряженного направления как функцию от t:

$$\varepsilon\left(\mathbf{x}_n + \mathbf{h}(\mathbf{x}_n)t\right) = \varepsilon_1(t).$$

Оптимальный размер шага t можно определить из разложения функции $\varepsilon_1(t)$ в ряд Тейлора до членов второго порядка, в окрестности точки $t = 0$:

$$\varepsilon_1(t) = \varepsilon_1(0) + \varepsilon_1'(0)t + \frac{\varepsilon_1''(0)t^2}{2}.$$

Далее значение шага t несложно получить из условия минимума разложения второго порядка в виде:

$$t = -\frac{\varepsilon_1'(0)}{\varepsilon_1''(0)}. \tag{159}$$

Аналитические формулы для производных $\varepsilon_1'(0)$ и $\varepsilon_1''(0)$ легко получить из уравнений (154) — (156).

Фазу штрихов φ можно рассматривать как дополнительный параметр в алгоритме оптимизации. Согласно уравнениям (155) и (156), значение $\varphi = \pi$ является оптимальным при расчете решетки без нулевого порядка. При расчете решетки с нулевым порядком, требуемое значение нулевого порядка может быть получено соответствующим выбором φ. В частности, при расчете решетки с равными интенсивностями порядков, определим значение φ в уравнении (155) в виде /52, 68/:

$$\varphi = 2\arcsin\left(\sqrt{\frac{2N}{E_0 + 8Q(1-Q)N}}\right), \tag{160}$$

где $E_0 = \sum_{\substack{j=-N \\ j\neq 0}}^{N} I_j$.

Величина E_0 является энергетической эффективностью и соответствует доле энергии, приходящейся на заданные порядки. Подставляя значение φ, определенное согласно уравнению (160), в (155) получим, что значение нулевого порядка становится равным среднему значению интенсивности в порядках. С учетом нулевого порядка энергетическая эффективность решетки увеличивается на величину:

$$\Delta E = E_0 \frac{2N + 1 - E_0 - \gamma N}{E_0 - \gamma N}, \tag{161}$$

где $\gamma = 8Q(1-Q)$.

Таблица 1.

Результаты расчета бинарных дифракционных решеток

Число порядков M	Число штрихов K	Энергетическая эффективность E(%)	средне-квадратичная ошибка(%)
1	***2***	***3***	***4***
5	2	77,5	0,7
7	2	81,2	1,4
9	3	78,2	2,6
11	4	84,6	0,5
15	4	84,6	0,5
19	5	82,4	1,1
25	7	80,3	3,8
31	8	83,4	3,1
41	11	82,0	4,7
51	14	82,3	3,7
61	16	80,7	2,1
71	19	80,4	1,3
81	22	82,8	3,7
91	25	81,0	2,5
101	27	81,4	2,4
151	40	82,8	4,9
181	44	82,7	3,9
201	54	81,4	4,6
251	69	82,2	4,7
281	74	82,4	4,0
301	81	83,3	3,7
451	118	85,1	3,2

В таблице 1 представлены результаты расчета бинарных решеток для различного числа порядков $M = 2N + 1$ равной интенсивности. Среднеквадратичная ошибка, представленная в колонке 4, имеет вид

$$\delta = \frac{1}{\overline{I}}\left[\frac{1}{M}\sum_{j=-N}^{N}\left(I_j - \overline{I}\right)^2\right]^{1/2},$$

где $\overline{I} = M^{-1}\sum_{j=-N}^{N} I_j$ — среднее значение. Высокая энергетическая эффективность E и низкая среднеквадратичная ошибка δ, полученные

для любых решеток с числом порядков до 451, подтверждает высокую эффективность градиентной процедуры с аналитически рассчитанным начальным профилем (152). Следует отметить, что разработанный метод можно без ограничений применять для расчета решеток с числом порядков до 1001 или более, оставаясь в пределах нескольких минут времени компьютерного расчета. Решетки с числом порядков $M > 15$ были рассчитаны при $\varphi = \pi$. В процессе расчета было получено, что при $M \le 15$ оказывается более удобным формировать нулевой порядок, выбирая значение φ из уравнения (160). Такой выбор φ позволяет на 2—4% увеличить энергетическую эффективность для решеток с числом порядков 5—11. Шаг t для градиентного метода (157) рассчитывался из условия минимума квадратичной аппроксимации функции ошибки вдоль сопряженного направления. Такой выбор t значительно увеличил скорость сходимости.

Важной технологической характеристикой решетки является размер минимального штриха, определяемый как минимальное расстояние между соседними координатами. В процессе расчетов было получено, что минимальная ширина штриха $\Delta_{\min} \approx 1/M$ первоначального бинарного профиля (152), остается примерно такой же после итерационной коррекции (157). Например, для решетки с числом порядков 101, отношение $\Delta_{\min}$/период равно 0,01.

На рис. 23 представлены рассчитанные профили для 101-порядковых решеток, формирующих постоянное и линейное распределение интенсивности в порядках дифракции. Распределение интенсивности в порядках для решеток на рис. 23 приведены на рис.24. Вследствие симметрии на рис. 24 показана только половина дифракционных порядков.

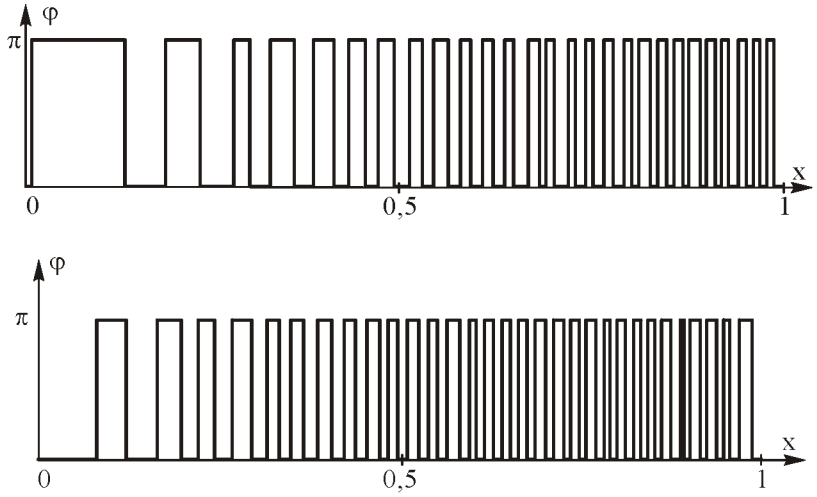


Рис. 23. Рассчитанный профиль 101-порядковой решетки, формирующей: постоянное распределение интенсивности (а), линейное распределение интенсивности (б)

В заключение отметим, что предложенный метод выбора начального профиля бинарной решетки может быть обобщен на двумерный случай. В двумерном случае начальный профиль может быть получен бинаризацией фазовой функции двумерного ДОЭ, фокусирующего в прямоугольную область /71/.

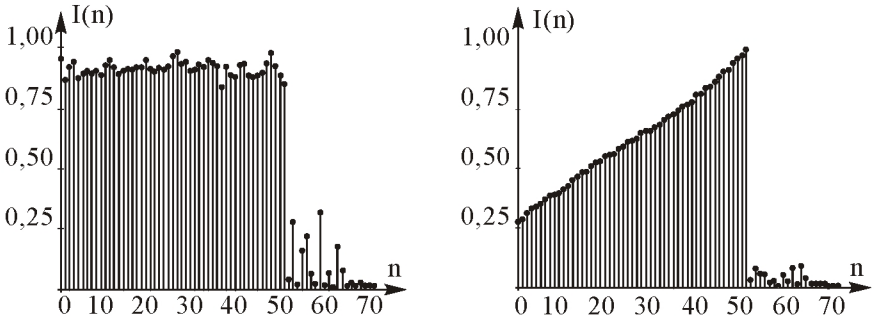


Рис. 24. Постоянное распределение интенсивности для 101-порядковой решетки (а), линейное распределение интенсивности для 101-порядковой решетки (б)

7.5. Многоуровневые фазовые дифракционные решетки

Согласно уравнениям (155) и (156), бинарные дифракционные решетки позволяют сформировать только симметричные распределения интенсивности дифракционных порядков. Дифракционные решетки с непрерывной фазовой функцией позволяют создавать несимметричные и сколь угодно сложные распределения интенсивности в дифракционных порядках. Для расчета дифракционных решеток с непрерывной фазовой функцией φ наиболее часто используется алгоритм Герчберга — Секстона (ГС). Для одномерных решеток, рассчитанных методом ГС энергетическая эффективность составляет более 90% ; при среднеквадратичной ошибке формирования заданной интенсивности $\hat{I}_j$ в дифракционных порядках — 10—20% /59, 60/. Следует отметить, что для применений дифракционных решеток в устройствах оптических вычислений и компьютерных сетях на оптических волокнах, среднеквадратичная ошибка должна быть на порядок меньше.

В работе /20/ было показано, что алгоритм ГС может быть рассмотрен как градиентный метод минимизации функционала ε(φ), представляющего квадрат ошибки:

$$\varepsilon(\varphi) = \sum_{j=-\infty}^{\infty} \left(\sqrt{I_j} - \sqrt{\hat{I}_j} \right)^2 . \qquad (162)$$

Интенсивности дифракционных порядков в уравнении (162) пропорциональны квадратам модулей коэффициентов Фурье в разложении функции комплексного пропускания решетки exp (*i*φ (*x*)):

$$I_j = \left| \frac{1}{d} \int_0^d \exp\left(\mathrm{i}\varphi(x) - \mathrm{i}\frac{2\pi}{d} jx \right) \mathrm{d}x \right|^2 , \qquad (163)$$

где d — период решетки. В дальнейшем будем считать что $d = 1$.

Для алгоритма ГС характерен эффект стагнации; после нескольких начальных итераций уменьшение ошибки (162) происходит очень медленно. Вид итерационного алгоритма и его сходимость существенно

зависят от выбора функционала ошибки. Скорость сходимости итерационного процесса можно увеличить, изменив форму функционала (162).

В связи с этим рассмотрим градиентный метод расчета фазовой функции $\varphi(x)$ решетки для функционала ошибки общего вида. Пусть $\varepsilon(\varphi)$ — некоторый функционал, характеризующий отличие расчетного распределения интенсивности в дифракционных порядках от требуемого распределения:

$$\varepsilon(\varphi) = \hat{\varepsilon}(\mathbf{I}, \mathbf{I}_d), \tag{164}$$

где $\mathbf{I}$ и $\mathbf{I}_d$ - вектора, компоненты которых соответствуют расчетным и требуемым интенсивностям дифракционных порядков. Согласно уравнению (163), компоненты вектора $\mathbf{I}$ являются функционалами от фазовой функции решетки $\varphi(x)$. Задача состоит в отыскании функции $\hat{\varphi}(x)$, минимизирующей значение функционала $\varepsilon(\varphi)$. В градиентном методе минимизации функционала $\varepsilon(\varphi)$, искомая функция $\hat{\varphi}(x)$ ищется как предел последовательности функций $\varphi_n(x)$, рассчитываемых по правилу:

$$\varphi_{n+1}(x) = \varphi_n(x) - t_n \varepsilon'(\varphi_n), \tag{165}$$

где $\varepsilon'(\varphi_n)$ — градиент функционала на n-том шаге процесса оптимизации, t_n — шаг градиентного метода. Для вычисления градиента функционала $\varepsilon'(\varphi)$ в (165), рассмотрим приращение функционала $\varepsilon(\varphi)$, вызванное малым изменением фазы $\Delta\varphi$. Согласно уравнению (164), приращение функционала имеет вид:

$$\Delta\varepsilon(\varphi) = \varepsilon(\varphi + \Delta\varphi) - \varepsilon(\varphi) =$$

$$= \sum_j \frac{\partial \hat{\varepsilon}(I, I_d)}{\partial I_j} \Delta I_j(\varphi). \tag{166}$$

Приращение интенсивности $\Delta I_j(\varphi)$ в (166), с точностью до членов второго порядка, может быть представлено в виде

$$\Delta I_j(\varphi) = 2\,\mathrm{Re}\left(F_j^*(\varphi)\,\Delta F_j(\varphi)\right), \tag{167}$$

где $F_j(\varphi)$ — комплексная амплитуда поля в j-ом дифракционном порядке;

$$F_j(\varphi) = \int_0^1 \exp\left(\mathrm{i}\varphi(x) - \mathrm{i}\,2\pi jx\right)\mathrm{d}x\,. \tag{168}$$

Согласно (168), приращение $\Delta F_j(\varphi)$ имеет вид:

$$\begin{aligned}\Delta F_j(\varphi) &= \int_0^1 \left(\exp\left(\mathrm{i}\varphi(x) + \mathrm{i}\,\Delta\varphi(x)\right) - \exp\left(\mathrm{i}\varphi(x)\right)\right)\exp\left(-\mathrm{i}\,2\pi jx\right)\mathrm{d}x \approx \\ &\approx \int_0^1 \mathrm{i}\,\Delta\varphi(x)\exp\left(\mathrm{i}\varphi(x) - \mathrm{i}\,2\pi jx\right)\mathrm{d}x\,.\end{aligned} \tag{169}$$

Подставляя уравнения (167) — (169) в (166), получаем приращение функционала в виде:

$$\Delta\varepsilon(\varphi) = 2\,\mathrm{Re}\left(\sum_j \frac{\partial\hat{\varepsilon}\left(\mathbf{I},\mathbf{I}_d\right)}{\partial I_j} F_j^*(\varphi)\int_0^1 \mathrm{i}\,\Delta\varphi(x)\exp\left(\mathrm{i}\varphi(x) - \mathrm{i}\,2\pi jx\right)\mathrm{d}x\right). \tag{170}$$

Согласно уравнению (170), градиент функционала $\varepsilon'(\varphi)$ имеет вид:

$$\varepsilon'(\varphi) = 2\,\mathrm{Re}\left(\mathrm{i}\sum_j \frac{\partial\hat{\varepsilon}\left(\mathbf{I},\mathbf{I}_d\right)}{\partial I_j} F_j^*(\varphi)\exp\left(\mathrm{i}\varphi(x) - \mathrm{i}\,2\pi jx\right)\right). \tag{171}$$

Для удобства последующего анализа уравнение (171) представим в виде:

$$\begin{aligned}&\varepsilon'(\varphi) = 2\,\mathrm{Re}\left(\mathrm{i}\exp\left(\mathrm{i}\varphi(x)\right)\Psi^*(x)\right) = \\ &= -2\left|\Psi(x)\right|\sin\left(\varphi(x) - \arg\left(\Psi(x)\right)\right),\end{aligned} \tag{172}$$

где $\Psi(x)=\int_{-\infty}^{\infty} f(\xi)\exp(\mathrm{i}2\pi x\xi)\mathrm{d}\xi$ — обратное преобразование Фурье от функции

$$f(\xi)=\sum_j \frac{\partial\hat{\varepsilon}(\mathbf{I},\mathbf{I}_d)}{\partial I_j}F_j(\varphi)\,\delta(\xi-j). \tag{173}$$

Расчет функции $\Psi(x)$ в (172) аналогичен первым трем шагам алгоритма ГС. Действительно, расчет $\Psi(x)$ в (172) сводится к выполнению следующих шагов: (1) вычисление прямого преобразования Фурье для расчета значений поля F_j в дифракционных порядках, (2) замена отсчетов поля F_j на функцию $f(\xi)$, (3) вычисление обратного преобразования Фурье от функции $f(\xi)$. Отличие алгоритма ГС от градиентной процедуры состоит в виде замены (173). В алгоритме ГС значения поля F_j заменяются не функцией $f(\xi)$ (173), а следующими значениями:

$$\hat{F}_j=\sqrt{\hat{I}_j}\cdot F_j/\left|F_j\right|.$$

Для расчета шага t градиентного метода, представим функционал $\varepsilon(\varphi)$ вдоль направления противоположного градиенту как функцию от t:

$$\varepsilon(\varphi)=\varepsilon\left(\varphi-t\varepsilon'(\varphi)\right)=\varepsilon_1(t).$$

Значение шага t можно определить из разложения функции $\varepsilon_1(t)$ в ряд Тейлора до линейных членов в окрестности точки $t=0$:

$$\varepsilon_1(t)=\varepsilon_1(0)+\varepsilon_1'(0)t. \tag{174}$$

Предполагая, что экстремальное значение функционала $\varepsilon(\varphi)$ равно нулю, определим шаг t из условия равенства нулю разложения (174):

$$t_f=-\frac{\varepsilon_1(0)}{\varepsilon_1'(0)}. \tag{175}$$

Производную $\varepsilon_1'(0)$ в (175) можно найти из (170) и (171) в виде

$$\varepsilon_1'(0) = -4\int_0^1 \left|\Psi(x)\right|^2 \sin^2\left(\varphi(x) - \arg\left(\Psi(x)\right)\right) \mathrm{d}x \,. \tag{176}$$

Предполагая функцию $\varepsilon_1(t)$ квадратичной вдоль направления сопряженного градиенту, несложно получить оценку для оптимального значения шага t_s в виде:

$$t_s = -\frac{\left(\varepsilon_1(0)\right)^2}{2\varepsilon_1\left(t_f\right)\varepsilon_1'(0)} \,. \tag{177}$$

Для расчета шага t_s необходим пересчет интенсивностей порядков для вычисления $\varepsilon_1(t_f)$. При стабилизации процесса сходимости функционал $\varepsilon(\varphi)$ меняется незначительно от итерации к итерации. Предполагая, что $\varepsilon_1(t_f) \approx \varepsilon_1(0)$, получаем оценку для шага t_s при стабилизации процесса сходимости в виде:

$$t_s = \frac{t_f}{2} = \varepsilon_1(0)\left(8\int_0^1 \left|\Psi(x)\right|^2 \sin^2\left(\varphi(x) - \arg\left(\Psi(x)\right)\right) \mathrm{d}x\right)^{-1} \,. \tag{178}$$

Отметим, что полученные выше уравнения для градиента функционала и величины шага градиентного метода являются общими и не зависят от вида функционала $\varepsilon(\varphi)$.

Рассмотрим связь между алгоритмом ГС и градиентной процедурой минимизации функционала (162). Минимизация функционала (162) эквивалентна максимизации следующего функционала:

$$\breve{\varepsilon}(\varphi) = 2\sum_j \sqrt{I_j \hat{I}_j} \,. \tag{179}$$

Градиент функционала (179) имеет вид (172) при функции $f(\xi)$, определенной в виде

$$f(\xi)=\sum_j \sqrt{\hat{I}_j}\,\frac{F_j(\varphi)}{\left|F_j(\varphi)\right|}\,\delta(\xi-j)\,. \qquad (180)$$

Функция (180) соответствует замене, выполняемой в Фурье-плоскости для алгоритма ГС. Соответственно расчет функции $\Psi(x)$ в уравнении (172) совпадает с первыми тремя шагами алгоритма ГС. Отличие градиентной процедуры (165), (172), (180) от алгоритма ГС состоит в расчете нового приближения фазы $\varphi_{n+1}(x)$ на n-м шаге итерационного процесса. Для алгоритма ГС новое приближение фазы вычисляется по правилу $\varphi_{n+1}(x) = \arg(\Psi_n(x))$, а для градиентной процедуры — по формуле (165). Рассмотрим ситуацию, при которой после выполнения первых итераций градиентный процесс стабилизировался, и функционал ошибки (179) стал изменяться медленно. В этом случае для большинства значений x в (172) имеем

$$\sin(\varphi_n(x) - \arg(\Psi_n(x))) \approx \varphi_n(x) - \arg(\Psi_n(x)). \qquad (181)$$

Из уравнений (165) и (172) следует, что для градиентного метода разность $\varphi_{n+1}(x)-\varphi_n(x)$ приблизительно пропорциональна величине $\varphi_n(x) - \arg(\Psi_n(x))$. Поскольку функция $\arg(\Psi_n(x))$ является новым приближением фазы для алгоритма ГС, то можно заключить, что градиентная процедура (165), (172) и (180) при $t = 1$ и $|\Psi_n(x)| \approx 1$ вычисляет новую фазу $\varphi_{n+1}(x)$, близкую к той, что получается в алгоритме ГС /5/.

В отличие от алгоритма ГС, градиентная процедура (165), (172) допускает выбор различных функционалов $\varepsilon(\varphi)$.

Вид функционала $\varepsilon(\varphi)$ оказывает существенное влияние на сходимость итерационного процесса. Целесообразно исследовать сходимость градиентной процедуры для функционала ошибки следующего вида:

$$\varepsilon(\varphi,p)=\sum_j\left(I_j-\hat{I}_j\right)^p\,. \qquad (182)$$

В отличие от функционала (162), функционал $\varepsilon(\varphi, p)$ (182) при $p > 1$ является более чувствительным к отклонению распределения интен-

сивности в дифракционных порядках от требуемого распределения. Градиент функционала (182) описывается уравнением (172) при функции $f(\xi)$, определенной в виде:

$$f(\xi) = p\sum_j \left(I_j - \hat{I}_j\right)^{p-1} F_j\, \delta(\xi - j). \tag{183}$$

Для выбора оптимального метода расчета многопорядковых дифракционных решеток, проведем сравнение работоспособности следующих итерационных процедур: алгоритма ГС, адаптивно-аддитивного алгоритма, рассмотренного в пункте 4 и градиентного метода (165), (172) при функционале ошибке (183).

Напомним, что отличие между алгоритмом ГС и адаптивно-аддитивным алгоритмом состоит только в форме замены расчетных значений F_j комплексной амплитуды поля в Фурье плоскости. В алгоритме ГС значения F_j заменяются на значения $\hat{F}_j$ по правилу

$$\hat{F}_j = \begin{cases} \sqrt{\hat{I}_j}\,\dfrac{F_j}{|F_j|}, & j \in Q, \\ 0, & j \notin Q, \end{cases} \tag{184}$$

где $Q = \{n_1, n_2, \ldots, n_M\}$ — индексы заданных дифракционных порядков.

В адаптивно-аддитивном алгоритме (40) замена (184) принимает вид

$$\hat{F}_j = \begin{cases} \left|\alpha\sqrt{\hat{I}_j} + (1-\alpha)\sqrt{I_j}\right|\dfrac{F_j}{|F_j|}, & j \in Q, \\ F_j, & j \notin Q, \end{cases} \tag{185}$$

где α — выбираемый коэффициент.

Для характеристики работоспособности дифракционных решеток будем использовать значения энергетической эффективности $E = \sum_{j \in Q} I_j$ и среднеквадратичной ошибки.

В таблице 2 приведены значения E и δ для рассчитанных дифракционных решеток с различным числом симметрично расположенных порядков равной интенсивности. Расчет производился с использованием алгоритма быстрого преобразования Фурье при числе элементов дискретизации фазовой функции решетки n = 64. Для учета дифракции на растре требуемые интенсивности порядков $\widehat{I}_j$ считались не постоянными, а обратно пропорциональными значениям $f(j) = \mathrm{sinc}^2(\pi j/n)$, $j = -n/2, n/2$ /42]. В качестве начальной фазы для итерационных процедур была использована случайная фаза, равномерно распределенная в интервале [0, 2π). В ходе расчета было получено, что для стабилизации ошибки в итерационных процедурах достаточно 80—100 итераций.

Таблица 2.

Результаты итеративного расчета многопорядковых дифракционных решеток

Число поряд-ков	Алгоритм ГС		АА-алгоритм		Градиентный метод для функционала (182) при p=2	
	E(%)	δ(%)	*E*(%)	δ(%)	*E*(%)	δ(%)
5	90,1	14,7	88,5	0,1	91,5	0,5
9	95,6	17,5	93,5	0,2	93,8	0,2
15	97,9	15,5	93,6	0,1	94,0	0,8
19	97,2	14,9	94,1	0,1	94,5	0,7
23	95,1	16,9	93,9	0,1	94,8	0,6
27	98,3	18,7	93,3	0,1	94,4	1,2
33	97,9	14,8	94,5	0,1	96,2	1,8
39	99,2	11,5	94,2	0,2	96,3	1,5
45	99,6	10,8	96,1	0,4	97,5	1,2
51	99,7	9,8	95,9	0,7	98,3	1,8

Левая часть таблицы 2 содержит значения *E* и δ для решеток, рассчитанных по алгоритму ГС (замена (184)). В средней части таблицы 2 даны значения E и δ для решеток, рассчитанных по адаптивно-аддитивному алгоритму (АА алгоритму). При расчете решеток по АА алгоритму, первые 15-20 итераций, пока отклонение расчетных интенсивностей порядков от заданных значений уменьшается быстро, вы-

полнялись по алгоритму ГС. После стабилизации ошибки в алгоритме ГС, осуществлялся переход к АА алгоритму (185) при α=2. В правой части таблицы приведены значения E и δ для решеток рассчитанных по градиентному методу (165), (172) и (173) для функционала (182) при p=2. Оптимальный размер шага t_n в (165) определялся по формулам (177) и (178). В процессе расчетов было получено, что в ходе первых 10-20 итераций шаг t_n должен определяться по более точной формуле (177). Использование упрощенной формулы (178) для величины шага t_n приводит к осцилляции критерия (182). Поскольку расчет t_n, по формуле (177) требует дополнительного вычисления преобразования Фурье, то для упрощения расчета следует выполнять первые 10-20 итераций по алгоритму ГС. Последующее применение градиентной процедуры приводит к гладким изменениям функционала $\varepsilon(\varphi, 2)$ и использование упрощенной формулы (178) для расчета шага t_n обеспечивает стабильную сходимость.

Данные таблицы 2 показывают, что решетки, рассчитанные по алгоритму ГС, имеют среднеквадратичную ошибку δ = 10—15% при энергетической эффективности E = 90—99%. АА алгоритм позволяет уменьшить ошибку δ до 0,1—0,2% при незначительном снижении энергетической эффективности E на 1—3%. Для решеток, рассчитанных по градиентному методу для функционала $\varepsilon(\varphi, 2)$, ошибка δ почти на порядок меньше, чем для решеток, рассчитанных по алгоритму ГС. В процессе расчетов было получено, что сходимость градиентного метода для функционала $\varepsilon(\varphi,p)$ наилучшая при p = 2, с ростом p скорость сходимости уменьшается.

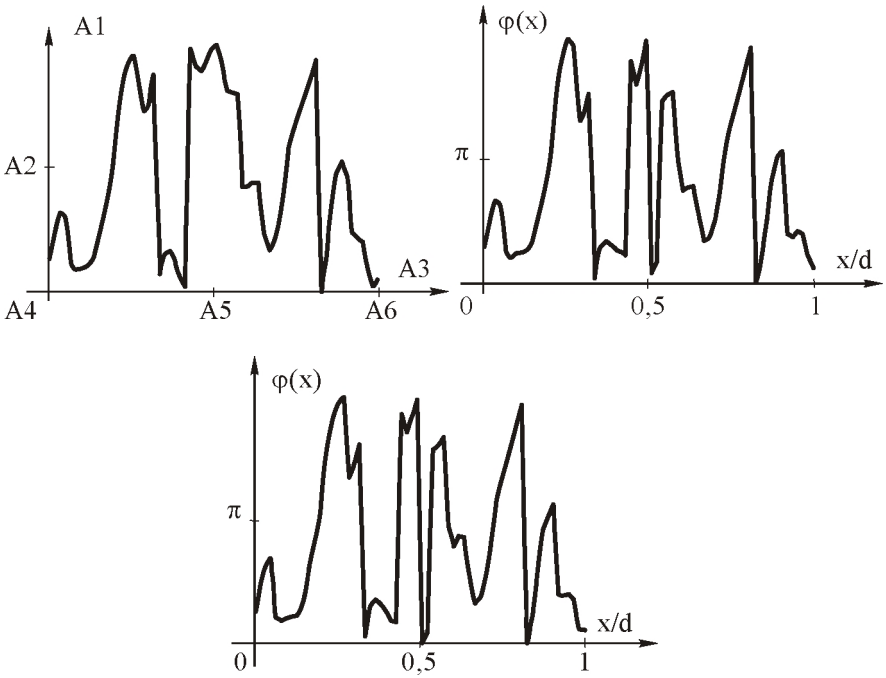


Рис. 25. Фазовые функции 11-порядковых решеток, рассчитанных по алгоритму ГС (а), по АА алгоритму (б), по градиентному методу (в)

Расчет проводился также для решеток с несимметричным расположением порядков. На рис. 25 показаны фазовые функции дифракционных решеток, формирующих 11 несимметричных порядков равной интенсивности. Распределения интенсивности в Фурье плоскости для решеток на рис. 25 представлены на рис. 26. Значения δ и E для решеток на рис. 25 составили 20,4% и 94,3% при расчете по алгоритму ГС (рис. 26а), 0,1% и 90,2% при расчете по АА-алгоритму (рис. 26б) и 0,6% и 91,2% при расчете градиентным методом для функционала $\varepsilon(\varphi, 2)$ (рис. 26в).

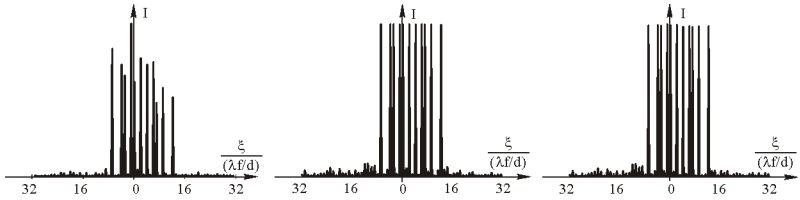


Рис. 26. Распределение интенсивности в Фурье-плоскости для 11-порядковых решеток, состоящих из 4 периодов: алгоритм ГС (а), АА алгоритм (б), градиентный метод (в)

Результаты расчетов (табл. 2 и рис. 26) показывают, что для расчета многопорядковых дифракционных решеток АА-алгоритм и градиентный метод для функционала $\varepsilon(\varphi,2)$ работают существенно лучше алгоритма ГС. При этом сходимость АА-алгоритма (185) несколько лучше по сравнению с градиентным методом. АА-алгоритм, по сравнению с градиентным методом, также проще в программной реализации (не требуется явного расчета градиента и величины шага на каждой итерации).

7.6. Фазовые ДОЭ, фокусирующие в объем и на поверхность тел вращения

Для расчета ДОЭ в объемные фигуры используют разбиение объема на N плоскостей и сведение задачи к расчету ДОЭ, формирующего плоские изображения. В /72, 73/ предложены различные подходы к решению этой задачи. В /72/ независимо рассчитываются амплитудно-фазовые функции $f_n(x, y)$, преобразование Френеля которых связано с заданным распределением интенсивности $I_n(\xi, \eta)$ на плоскостях $z = z_n$ (см. рис. 39) уравнением

$$I_n(\xi,\eta) = \left|\Im_n\{f_n(x,y)\}\right|^2, \qquad (186)$$

где $\Im_n\{\cdot\}$ — преобразование Френеля (1) на плоскость $z = z_n$. Это обозначение преобразования Френеля отличается от обозначения преобразования Фурье только нижним индексом n, который указывает номер плоскости (ξ,η), на которой вычисляется комплексная амплитуда света.

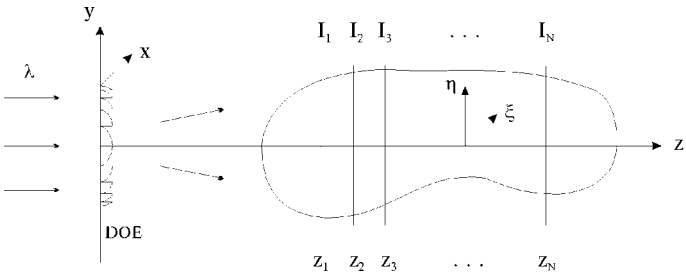


Рис. 27. Схема для фокусировки с помощью ДОЭ в 3- мерную область

Результирующая амплитудно-фазовая функция пропускания ДОЭ равна арифметическому среднему:

$$f(x,y)=\frac{1}{N}\sum_{n=0}^{N} f_n(x,y)\,. \qquad (187)$$

Далее для замены амплитудно-фазовой функции $f(x, y)$ на только фазовую функцию используют обычные процедуры кодирования цифровой голографии /74/.

В работе /73/ функции $f_n(x, y)$ рассчитываются взаимосвязанно, а не независимо, что позволяет в ходе итеративного процесса получить только фазовую функцию без дополнительных (малоэффективных) процедур кодирования.

В /75/ рассматривалась в приближении геометрической оптики задача расчета ДОЭ, фокусирующего на поверхность с осью вращения z.

В данном разделе, используя иной подход, рассматриваются итеративные методы расчета фазовых ДОЭ, фокусирующих лазерное излучение в объем и на поверхность тел вращения /76/.

Рис. 27 показывает, что ДОЭ, который далее рассматривается как тонкий фазовый транспарант, освещается плоской волной с длиной волны $\lambda=2\pi/k$ и формирует на расстояниях вдоль оптической оси z_n, $n=\overline{1,N}$, заданные распределения интенсивности $I_n(\xi,\eta)$.

Рассмотрение проводится в рамках скалярной волновой теории дифракции Кирхгофа в приближении Френеля.

В /73/ описан алгоритм расчета функции пропускания ДОЭ с помощью итеративной процедуры минимизирующей функционал

$$\mathbf{E}\varepsilon = \sum_{n=1}^{N} \int_{-\infty}^{\infty} \int \left[\left| F(\xi,\eta,z_n) \right| - A_n(\xi,\eta) \right]^2 \mathrm{d}\xi\,\mathrm{d}\eta\,, \tag{188}$$

где $A_n(\xi,\eta) = \sqrt{I_n(\xi,\eta)}, \qquad n = \overline{1,N}\,,$

$$F(\xi,\eta,z_n) = \Im_n \left\{ \mathrm{e}^{\mathrm{i}\varphi(x,y)} \right\} =$$

$$= \frac{k}{z_n} \iint_{\Omega} \exp\left[\mathrm{i}\,\varphi(x,y) + \frac{\mathrm{i}k}{2z_n} \left[(x-\xi)^2 + (y-\eta)^2 \right] \right] \mathrm{d}x\,\mathrm{d}y\,, \tag{189}$$

где $\varphi(x,y)$ искомая фаза ДОЭ, который формирует световое поле с комплексной амплитудой $F(\xi,\eta,z)$, модуль которой совпадает на заданных расстояниях z_n с требуемой амплитудой $A_n(\xi,\eta,z_n)$, Ω — форма апертуры ДОЭ.

Итеративный алгоритм, минимизирующий среднеквадратичное отклонение (88), имеет вид:

$$\varphi_p(x,y) = \arg\left[\sum_{n=1}^{N} U_n^{(p)}(x,y) \right], \tag{190}$$

$$U_n^{(p)}(x,y) = \Im_n^{-1} \left\{ A_n(\xi,\eta) \exp\left[\mathrm{i}\,Q_{n,p}(\xi,\eta) \right] \right\}, \tag{191}$$

$$Q_{n,p}(\xi,\eta) = \arg\left[\Im_n \left\{ \exp\left[\mathrm{i}\,\varphi_{p-1}(x,y) \right] \right\} \right], \tag{192}$$

где p — номер итерации, $Q_{n,p}(\xi,\eta)$ — фаза светового поля на плоскости, отстоящей от ДОЭ на расстоянии z_n и рассчитанная на p-ой итерации.

Ниже рассматриваются различные варианты итеративного алгоритма минимизирующего функционал среднеквадратичного отклонения вида

$$\varepsilon_1 = \iint_{\Omega} \left| \mathrm{e}^{\mathrm{i}\varphi(x,y)} - \sum_{n=1}^{N} C_n U_n(x,y) \right|^2 \mathrm{d}x\,\mathrm{d}y \,, \tag{193}$$

где C_n — комплексные весовые множители.

Далее, в зависимости от того, куда требуется сфокусировать излучение, будет зависеть вид функций U_n. Этим и будут отличаться рассматриваемые задачи.

Функционал (188), записанный для плоскостей, отстоящих от ДОЭ на расстояниях z_n, можно с помощью равенства Парсеваля записать в виде:

$$\varepsilon = \sum_{n=1}^{N} \iint_{\Omega} \left| \mathrm{e}^{\mathrm{i}\varphi(x,y)} - U_n(x,y) \right|^2 \mathrm{d}x\,\mathrm{d}y \,. \tag{194}$$

Функции $U_n(x,y)$ в уравнениях (193) и (194) являются результатом вычисления обратного преобразования Френеля:

$$U_n(x,y) = \Im_n^{-1}\left\{F_n(\xi,\eta)\right\}, \tag{195}$$

$$F_n(\xi,\eta) = A_n(\xi,\eta)\exp\left[\mathrm{i}\,Q_n(\xi,\eta)\right]. \tag{196}$$

Если переписать функционал (194) в обобщенной форме с весовыми коэффициентами

$$\widehat{\varepsilon} = \sum_{n=1}^{N} \iint_{\Omega} \left| \mathrm{e}^{\mathrm{i}\varphi(x,y)} - C_n U_n(x,y) \right|^2 \mathrm{d}x\,\mathrm{d}y \,, \tag{197}$$

то нетрудно получить уравнение, связывающее значения обоих функционалов

$$\varepsilon_1 = \widehat{\varepsilon} + \sum_{m\neq n}^{N}\sum^{N} C_m C_n^* \iint_{\Omega} U_n(x,y) U_m^*(x,y)\,\mathrm{d}x\,\mathrm{d}y \;-\; (N-1)\cdot W_0 \,, \tag{198}$$

где

$$W_0 = \iint_{\Omega} \left| \mathrm{e}^{\mathrm{i}\varphi(x,y)} \right|^2 \mathrm{d}x\,\mathrm{d}y \,. \tag{199}$$

—полная световая энергия в плоскости ДОЭ, S_Ω — площадь апертуры ДОЭ.

Из уравнения (198) видно, что при ортогональности функций $U_n(x,y)$, то есть при выполнении условия:

$$\iint\limits_{\Omega} U_n(x,y)U_m^*(x,y)\,\mathrm{d}x\,\mathrm{d}y = \int\limits_{-\infty}^{\infty}\int F_n(\xi,\eta)F_m^*(\xi,\eta)\,\mathrm{d}\xi\,\mathrm{d}\eta = W_0\delta_{mn}, \quad (200)$$

где δ_{mn} — символ Кронекера, функционалы ε_1 и $\widehat{\varepsilon}$ равны с точностью до несущественной постоянной $(N-1)\cdot W_0$. Функционал общего вида (193) имеет произвол в выборе функций $Q_n(\xi,\eta)$.

Условию ортогональности (200) можно наверняка удовлетворить, задавая амплитудные функции на плоскостях фокусировки в зоне дифракции Френеля $A_n(\xi,\eta)$ с пространственно разделенными носителями, то есть области G_n, в которых амплитуды $A_n(\xi,\eta)$ отличные от нуля, не пересекаются между собой:

$$G_n \cap G_m = \varnothing, \qquad m,n = \overline{1,N}. \quad (201)$$

Такое условие реализуется при расчете фазовых ДОЭ, формирующих моды Гаусса — Эрмита или Гаусса — Лагерра в различных порядках дифракции /77, 78/.

Условию (200) можно также удовлетворить, разделив апертуру ДОЭ Ω на N непересекающихся субапертур Ω_n, в каждой из которых задается функция $U_n(x,y)$:

$$\Omega_n \cap \Omega_m = \varnothing, \qquad m,n = \overline{1,N}. \quad (202)$$

Разбиение апертуры ДОЭ, например, на кольцевые субапертуры применяется при расчетах аксиконов и формирователей бесселевых мод /79/.

Минимизация функционала (193) эквивалентна поиску коэффициентов C_n, обеспечивающих выполнение следующего равенства

$$e^{i\varphi(x,y)} = \sum_{n=1}^{N} C_n U_n(x,y), \tag{203}$$

где функции $U_n(x,y)$ определяются уравнениями (195), (196) и в общем случае неортогональные. Уравнение (203) можно рассматривать как двумерную проекцию трехмерной задачи:

$$e^{i\varphi(x,y,z)} = \sum_{n=1}^{N} C_n U_n(x,y)\psi_n(z), \tag{204}$$

где функции $\psi_n(z)$ ортогональны:

$$\int_0^1 \psi_n(z)\psi_m^*(z)\,dz = \delta_{mn}. \tag{205}$$

Если, например, функции (205) выбрать в виде Фурье-гармоник

$$\psi_n(z) = \exp\left[i\,2\pi nz\right], \tag{206}$$

то коэффициенты суммы (204) вычисляются с помощью соотношений:

$$C_n = W_0^{-1} \iint_{\Omega} L_n(x,y)U_n^*(x,y)\,dx\,dy, \tag{207}$$

$$L_n(x,y) = \int_0^1 \exp\left[i\varphi(x,y,z) - i\,2\pi nz\right]dz. \tag{208}$$

С учетом уравнений (203) — (208) получим итеративный алгоритм минимизации функционала (193) /76/:

$$\varphi_p(x,y,z) = \arg\left[\sum_{n=1}^{N} B_n e^{iD_{n,p}} U_n(x,y)e^{i2\pi nz}\right], \tag{209}$$

$$D_{n,p} = \arg\left[\iint_{\Omega} L_{n,p}(x,y)U_n^*(x,y)\,dx\,dy\right], \tag{210}$$

$$L_{n,p}(x,y)=\int_0^1 \exp\left[\mathrm{i}\,\varphi_{p-1}(x,y,z)-\mathrm{i}\,2\pi nz\right]\mathrm{d}z\,, \tag{211}$$

где $B_n \geq 0$ — произвольные положительные числа, p — номер итерации. Искомая фазовая функция ДОЭ получается как проекция на ось $z=0$:

$$\varphi_p(x,y)=\varphi_p(x,y,z=0)\,. \tag{212}$$

Заметим, что система функций $U_n(x,y)\psi(z)$ не будет полной, и поэтому не всякую функцию $\exp\left[\mathrm{i}\varphi(x,y,z)\right]$ можно разложить в ряд по этой системе функций. Но так как алгоритм (209) — (211) используется для численного расчета и число N всегда конечное, то свойство полноты базисных функций не используется.

Далее ограничим рассмотрение радиально-симметричными относительно оси z телами. На рис.28 показана оптическая схема для расчета ДОЭ, фокусирующего излучение на ряд плоскостей, являющихся круглыми сечениями тела вращения. При этом амплитуда $A_n(\xi,\eta)$ на каждой из плоскостей зависит в полярных координатах $\xi = \rho \cdot \cos\theta$, $\eta = \rho \cdot \sin\theta$ только от радиальной переменной ρ, $A_n(\rho)$.

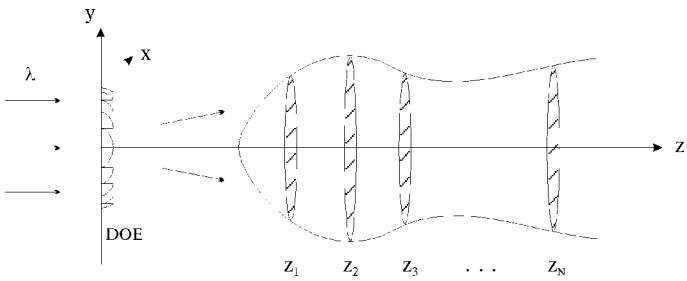


Рис. 28. Фокусировка в тело вращения с помощью ДОЭ

Тогда вместо уравнений (204) и (205) в данном случае можно записать:

$$\mathrm{e}^{\mathrm{i}\varphi(r,\varphi)} = \sum_{n=1}^{N} C_n U_n(r)\mathrm{e}^{\mathrm{i}n\varphi}, \tag{213}$$

$$C_n = \left(2\pi W_0\right)^{-1} \int_0^R \int_0^{2\pi} \mathrm{e}^{\mathrm{i}\varphi(r,\varphi)} U_n^*(r)\mathrm{e}^{\mathrm{i}n\varphi}\, r\,\mathrm{d}r\,\mathrm{d}\varphi, \tag{214}$$

$$U_n(r) = \Im_n^{-1}\left\{A_n(\rho)\mathrm{e}^{\mathrm{i}Q_n(\rho)}\right\} =$$

$$= \frac{k}{z_n}\exp\left[-\mathrm{i}\frac{kr^2}{z_n}\right]\int_0^{\infty} A_n(\rho)\exp\left[\mathrm{i}Q_n(\rho) - \mathrm{i}\frac{k\rho^2}{2z_n}\right] J_0\left(\frac{kr\rho}{z_n}\right)\cdot\rho\,\mathrm{d}\rho, \tag{215}$$

где (r,φ) — полярные координаты в плоскости ДОЭ, $J_0(x)$ — функция Бесселя нулевого порядка. Уравнение (215) является обратным преобразованием Френеля в полярных координатах.

Если интенсивность внутри сечений тела вращения постоянная, то амплитуду $A_n(\rho)$ можно выбрать в виде

$$A_n(\rho) = \sqrt{I_n}\,\mathrm{circ}\left(\frac{\rho}{\rho_0}\right), \tag{216}$$

где

$$\mathrm{circ}\left(\frac{\rho}{\rho_0}\right) = \begin{cases} 1, & \rho \le \rho_0, \\ 0, & \rho > \rho_0, \end{cases} \tag{217}$$

и I_n — постоянные значения интенсивности в круге на каждой плоскости.

Произвольные фазы $Q_n(\rho)$ в уравнении (215) можно выбрать квадратичными, т.к. параболическая волна является собственной функцией преобразования Френеля,

$$Q_n(\rho) = \frac{k\rho^2}{2z_n}, \tag{218}$$

а постоянные интенсивности I_n выбираются из условия сохранения энергии

$$\pi \rho_n^2 I_n = W_0, \qquad n = \overline{1,N} \,. \tag{219}$$

Подставив выражения (216) — (218) в уравнение (215), получим конкретный вид функций разложения:

$$U_n(r) = 2\pi \rho_n \sqrt{I_n} r^{-1} J_1\left(\frac{kr\rho_n}{z_n}\right) \exp\left[-\mathrm{i}\frac{kr^2}{2z_n}\right], \tag{220}$$

где $J_1(x)$ — функция Бесселя первого порядка.

Искомую функцию фазы ДОЭ $\varphi(r,\theta)$, фокусирующего в набор кругов с постоянными интенсивностями, с заданными радиусами и расположенных вдоль оптической оси на требуемых расстояниях z_n, можно искать в виде суммы с неполностью определенными коэффициентами:

$$\mathrm{e}^{\mathrm{i}\varphi(r,\theta)} = r^{-1} \sum_{n=1}^{N} C_n J_1(\alpha_n r) \exp\left[-\mathrm{i}\beta_n r^2 + \mathrm{i}n\theta\right], \tag{221}$$

где C_n — постоянные, модуль которых задается с учетом выбора интенсивностей I_n, а фаза является свободным параметром

$$\alpha_n = \frac{k\rho_n}{z_n}, \qquad \beta_n = \frac{k}{2z_n} \,.$$

Из уравнения (263) следует выражение для расчета коэффициентов:

$$C_n = W^{-1} \int_0^R \int_0^{2\pi} \mathrm{e}^{\mathrm{i}\varphi(r,\theta)} J_1(\alpha_n r) \exp\left[\mathrm{i}\beta_n r^2 - \mathrm{i}n\theta\right] \cdot r^2 \,\mathrm{d}r\,\mathrm{d}\theta \,, \tag{222}$$

$$W = 2\pi \int_0^R J_1^2(\alpha_n r) \cdot r \,\mathrm{d}r \,. \tag{223}$$

Уравнения (221) — (225) позволяют получить итеративный алгоритм расчета функции фазы ДОЭ, аналогичный алгоритму (209) — (211). Эти два алгоритма сходны по своей структуре и различаются только базисными функциями (220).

Далее рассмотрим частный случай этого алгоритма для расчета ДОЭ для фокусировки в набор осевых точек.

Нетрудно получить разложение аналогичное (221) для расчета многофокусных дифракционных линз, которые формируют вдоль оптической оси z заданное число фокусов на требуемых расстояниях и с заданной интенсивностью. Другие подходы к расчету ДОЭ типа многофокусных линз рассмотрены в /80, 81/. ДОЭ, рассчитанные методом (221), обладают свойствами голограмм. Например, малый участок такого ДОЭ способен восстановить требуемое изображение с минимальными искажениями, при прочих равных условиях.

В данном случае вместо уравнения (216) для фокусировки в набор кругов следует использовать уравнение для фокусировки в набор дельта-импульсов:

$$A_n(\rho) = \sqrt{I_n}\,\delta(\rho) \quad . \tag{224}$$

Тогда вместо уравнений (210) и (221) получим, соответственно

$$U_n(r) = \frac{k}{z_n}\sqrt{I_n}\exp\left[-\mathrm{i}\frac{kr^2}{2z_n}\right], \tag{225}$$

$$\mathrm{e}^{\mathrm{i}\varphi(r,\theta)} = \sum_{n=1}^{N} C_n \exp\left[-\mathrm{i}\beta_n r^2 + \mathrm{i}n\theta\right], \qquad \beta_n = \frac{k}{2z_n} \quad . \tag{226}$$

Вместо уравнения (222) коэффициенты суммы (226) вычисляются по формулам

$$C_n = \left(\pi R^2\right)^{-1} \int_0^R \int_0^{2\pi} \mathrm{e}^{\mathrm{i}\varphi(r,\theta)} \exp\left[+\mathrm{i}\frac{kr^2}{2z_n} - \mathrm{i}n\theta\right] \cdot r\,\mathrm{d}r\,\mathrm{d}\theta \quad . \tag{227}$$

Из уравнения (227) видно, что коэффициенты C_n вычисляются с помощью двумерного преобразования Фурье.

Заметим, что вместо суммы (226) для итеративного расчета многофокусных линз можно использовать только более простое радиальное уравнение:

$$e^{i\varphi(r)} = \sum_{n=1}^{N} C_n \exp\left[-i\beta_n r^2\right], \tag{228}$$

Но гауссовые экспоненты в уравнении (228) будут ортогональны только при условии

$$\int_0^R \exp\left[-i\left(\beta_m - \beta_n\right) r^2\right] \cdot r \, dr = \delta_{mn}, \tag{229}$$

которое выполняется при

$$z_n = n^{-1} z_0, \qquad z_0 = \frac{R^2}{2\lambda}. \tag{230}$$

Из (230) следует, что разложением (228) удобно пользоваться для вычисления фазы ДОЭ, если требуемые фокусы расположены только на определенных расстояниях z_n. .

Ниже получен алгоритм расчета ДОЭ, фокусирующих лазерное излучение на поверхность тел вращения. Расчет данного ДОЭ, ось которого совпадает с оптическою осью, был осуществлен методом геометрической оптики в /75/. Если тело вращения с осью z представить как набор его поперечных сечений, то его поверхность аппроксимируется набором колец. Поэтому для расчета ДОЭ, формирующего набор световых колец с заданными радиусами ρ_n и расположенных на требуемых расстояниях z_n , можно получить уравнение, аналогичное уравнению (221). Для этого вместо выражения (216) для требуемой амплитуды на n-й плоскости запишем:

$$A_n(\rho, \psi) = \sqrt{I_n} \cdot \delta(\rho - \rho_n) \cdot e^{in\psi}. \tag{231}$$

Подставив выражение (231) в уравнение (215), получим вместо соотношения (220) следующее выражение для комплексной амплитуды в плоскости ДОЭ:

$$U_n(r,\theta)=\frac{(-\mathrm{i})^{n_k}}{z_n}\sqrt{I_n}J_n\left(\frac{kr\rho_n}{z_n}\right)\exp\left[-\mathrm{i}\frac{k}{2z_n}\left(r^2+\rho_n^2\right)+\mathrm{i}n\theta\right]. \qquad (232)$$

Объединяя под знаком постоянной C_n сомножители в уравнении (274), которые не зависят от переменных r и φ, получим вместо уравнения (221), следующее уравнение:

$$\mathrm{e}^{\mathrm{i}\varphi(r,\theta)}=\sum_{n=1}^{N}C_nJ_n(\alpha_nr)\exp\left[-\mathrm{i}\beta_nr^2+\mathrm{i}n\theta\right], \qquad (233)$$

коэффициенты которого находятся по формулам, аналогичным (223) и (224):

$$C_n=W_n^{-1}\int_0^R\int_0^{2\pi}\mathrm{e}^{\mathrm{i}\varphi(r,\theta)}J_n(\alpha_nr)\exp\left[\mathrm{i}\beta_nr^2-\mathrm{i}n\theta\right]\cdot r\,\mathrm{d}r\,\mathrm{d}\theta, \qquad (234)$$

$$W_n=2\pi\int_0^R J_n^2(\alpha_nr)\cdot r\,\mathrm{d}r,$$

где $J_n(x)$ — функция Бесселя n-го порядка.

С целью ускорения расчетов вместо разложения (233) можно использовать следующее уравнение:

$$\mathrm{e}^{\mathrm{i}\varphi(r)}=\sum_{n=1}^{N}C_nJ_0(\alpha_nr)\exp\left[-\mathrm{i}\beta_nr^2\right]. \qquad (235)$$

Однако при этом слагаемые в уравнении (235) в общем случае не ортогональны между собой. Чтобы найти условие, которое надо наложить на параметры α_n и β_n для достижения ортогональности слагаемых в уравнении (235), можно использовать следующий справочный интеграл /82/:

$$\int_0^\infty e^{i\alpha x^2} J_\nu(bx) J_\nu(cx) x\,dx = \frac{i}{2a} J_\nu\left(\frac{bc}{2a}\right) e^{ih}, \qquad h = \frac{b^2+c^2}{4a} - \frac{\nu\pi}{2}. \tag{236}$$

Для функций, входящих в уравнение (235), вместо (236) получим

$$\int_0^\infty J_0(\alpha_n r) J_0(\alpha_m r) \exp\left[-i(\beta_n - \beta_m) r^2\right] r\,dr =$$

$$= i(\beta_n - \beta_m)^{-1} J_0\left(\frac{\alpha_n \alpha_m}{2(\beta_n - \beta_m)}\right) \exp\left[-i\frac{\alpha_n^2 + \alpha_m^2}{4(\beta_n - \beta_m)}\right], \qquad \beta_n > \beta_m. \tag{237}$$

Из уравнения (237) следует, что при удовлетворении параметров α_n и β_n условию

$$\frac{\alpha_n \alpha_m}{2(\beta_n - \beta_m)} = \gamma_p; \qquad n, m = \overline{1, N}, \qquad p = \overline{1, N^2}, \tag{238}$$

где γ_p — корни функции Бесселя: $J_0(\gamma_p) = 0$, слагаемые в сумме (235) будут ортогональны и коэффициенты можно рассчитывать по формулам:

$$C_n = W_n^{-1} \int_0^R e^{i\varphi(r)} J_0(\alpha_n r) \exp\left[i\beta_n r^2\right] \cdot r\,dr, \tag{239}$$

$$W_n = 2\pi \int_0^R J_0^2(\alpha_n r) \cdot r\,dr. \tag{240}$$

Условие (238) ограничивает выбор значений расстояний z_n и радиусов колец ρ_n. Однако, как показал численный эксперимент, можно получить хорошие результаты, считая слагаемые в уравнении (235) ортогональными и используя для расчета коэффициентов формулу (239).

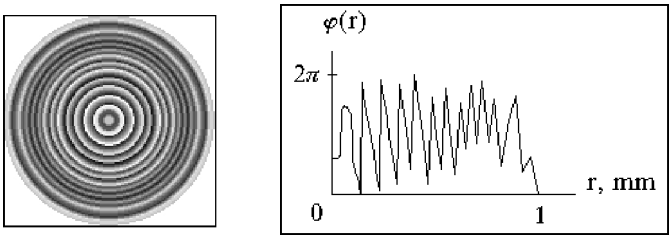
Пример 9. Численные результаты, приведенные ниже, относятся к случаю фокусировки на поверхность вращения. Фаза ДОЭ при этом рассчитывается с помощью итеративного алгоритма, основанного на уравнениях (235) и (239). Расчетные параметры: R = 1 мм — радиус

ДОЭ, $k = 10^4$ мм$^{-1}$, $n = 256$ — число отсчетов по радиальной переменной.

Коническая поверхность описывалась уравнением

$$\rho_n = \alpha\left(z_n - z_0\right) + \rho_0 \,, \tag{241}$$

где $\alpha = \pm 5\times10^{-3}$, $z_0 = 100$ мм, $\Delta z = 5$ мм — расстояние между плоскостями сечений конуса, $n = \overline{1,10}$. На рис. 29 показана полутоновая по уровню 2π фаза ДОЭ (а) и ее радиальное сечение (б), рассчитанные за 10 итераций алгоритмом на основе уравнений (235), (239).



а б

Рис. 29. Фаза (а) и ее радиальное сечение (б) для ДОЭ, фокусирующего на поверхность расходящегося конуса

На рис. 30 показаны распределения интенсивности (верхняя строка) и их сечения (нижняя строка), сформированные ДОЭ с фазой, показанной на рис. 29, и рассчитанные на разных плоскостях вдоль оси z в диапазоне [100 мм, 150 мм] с шагом 10 мм (z растет слева направо).

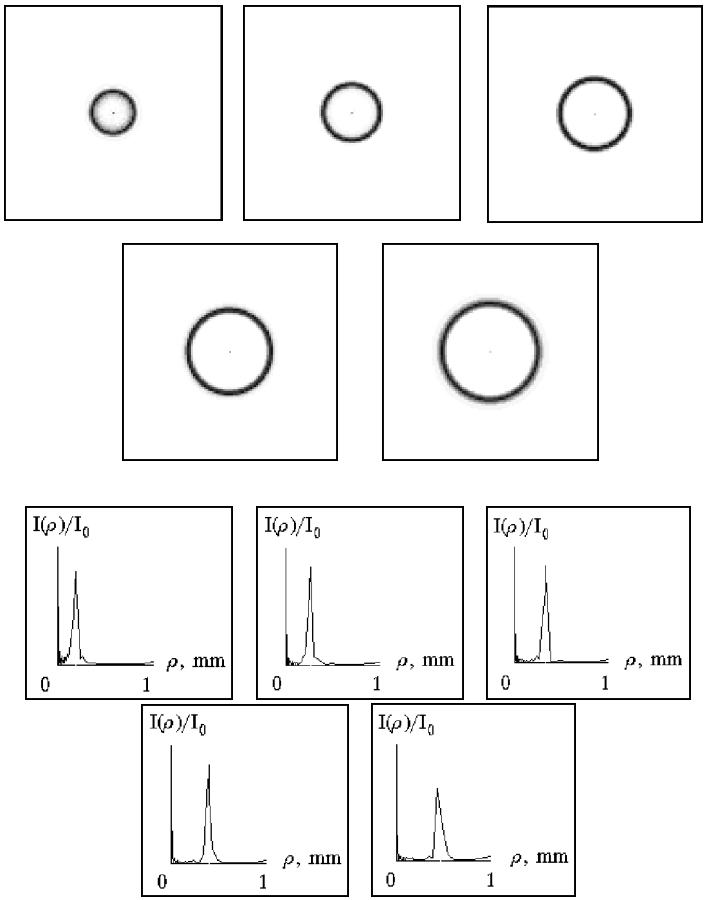


Рис. 30 Распределения интенсивности (верхняя строка) и их радиальные сечения (нижняя строка), сформированные ДОЭ с фазой, показанной на рис. 29

На рис. 31 показана зависимость нормированной интенсивности на поверхности конуса (рис. 30, $\alpha > 0$) вдоль оси z: кривая 1 — заданное распределение интенсивности, кривая 2 — рассчитанная.

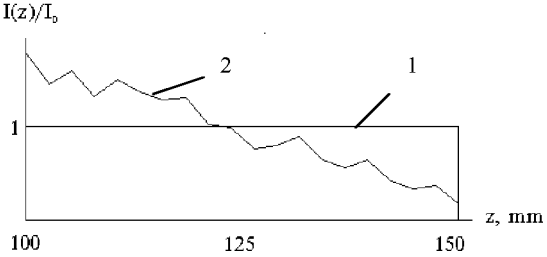
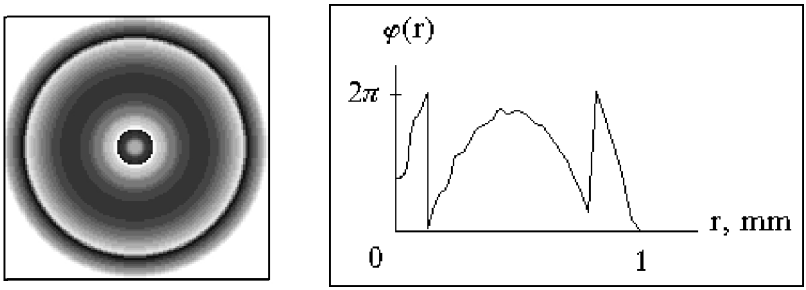


Рис. 31. Распределение относительной интенсивности на поверхности расходящегося конуса

Пример 10. На рисунках 32—34 приведены результаты численного эксперимента для фокусировки в цилиндр радиусом ρ_0 =0,5 мм. На рис. 32 показана фаза ДОЭ (а) и ее сечение (б), рассчитанные за 10 итераций по формулам (235) и (239) при $\alpha = 0$.



а б

Рис. 32. Фаза (а) и ее радиальное сечение (б) для ДОЭ, фокусирующего на поверхность цилиндра

На рис. 33 даны распределения интенсивности и их сечения, рассчитанные на тех же плоскостях, что и для рис. 32. Расчет проводился с помощью интегрального преобразования Френеля, которое, в свою очередь, рассчитывалось с помощью алгоритма быстрого преобразования Фурье.

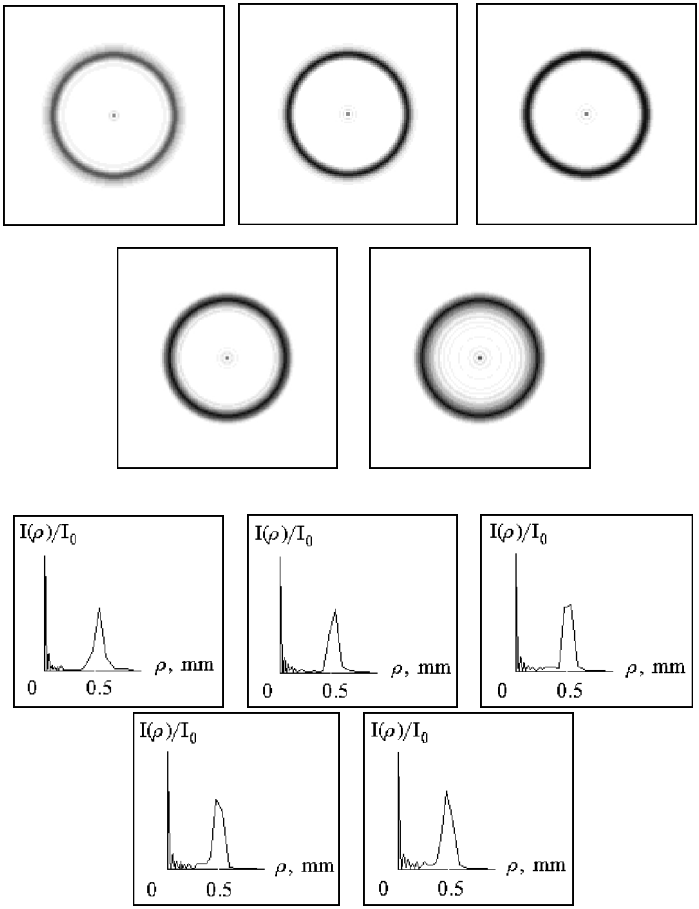


Рис. 33. Распределения интенсивности (верхняя строка) и радиальные сечения (нижняя строка), сформированные ДОЭ с фазой, показанной на рис. 32

На рис. 34 показан график зависимости относительной интенсивности света на поверхности цилиндра от расстояния z до оптического элемента.

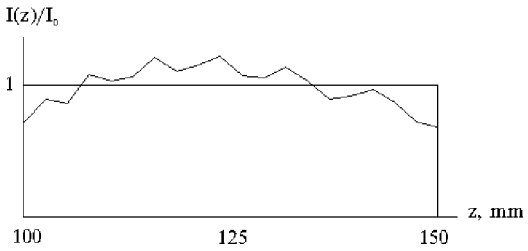


Рис. 34. Распределение относительной интенсивности на поверхности цилиндра

Приведенные результаты показывают возможность использования предложенных итеративных алгоритмов для расчета фазовых оптических элементов, фокусирующих лазерное излучение на поверхности тел вращения.

7.7. Фокусировка гауссового пучка в квадрат

Выравнивание плотности энергии в сечении когерентных световых пучков является важной задачей для технологических и информационных применений лазеров. Для этой цели используются дифракционные оптические элементы (ДОЭ), преобразующие волновой фронт таким образом, чтобы в заданной плоскости сформировалось требуемое распределение интенсивности.

История этого вопроса такова. В работе /83/ было предложено использовать бинарную дифракционную решетку для перераспределения энергии в пучке с целью ее выравнивания. Использование геометрооптического приближения для расчета ДОЭ, выравнивающего поперечное распределение интенсивности коллимированного гауссового пучка предложено в /84/. В /85/ с помощью киноформной линзы экспериментально осуществлено выравнивание интенсивности гауссового пучка в дальней зоне дифракции. Итеративный подход для расчета фазового ДОЭ, преобразующего гауссовое распределение интенсивности пучка в супергауссово в фокусе линзы предложен в /86/. При этом численно было показано, что преложенный алгоритм эффективно работает, если размер области фокусировки больше нескольких (4—5) минимальных дифракционных пятен.

Геометрооптическому расчету голографического фильтра для выравнивания профиля интенсивности пучка света посвящена работа /87/. Недостаток голографических фильтров в том, что они обладают низкой энергетической эффективностью. В /88/ голографический фильтр для выравнивания поперечной интенсивности пучка рассчитывался итеративным методом в приближении скалярной дифракции. Это позволило достичь более высокой точности формирования требуемого распределения интенсивности. Численный эксперимент с фокусатором, преобразующим гауссовый пучок в квадрат с постоянной интенсивностью проводился в /89/. Было, в частности, показано, что геометрооптический метод для расчета ДОЭ применим, если размер области фокусировки больше семи дифракционных пятен.

В /90/ экспериментально с помощью двух голографических фильтров, осуществленно преобразование эллиптического гауссового пучка в ближней зоне дифракции в квадрат с постоянной интенсивностью и фазой. В этой же работе, и независимо в работе /91/, получены аналитические формулы для фазы ДОЭ, фокусирующего гауссовый пучок в малый квадрат с постоянной интенсивностью. В /92/ приведен сравнительный численный эксперимент по расчету ДОЭ двумя способами: геометрооптическим и итеративным в приближении Френеля. Показано, что наилучшего результата по выравниванию гауссового пучка можно достичь, применяя геометрооптическое решение /90, 91/ как начальное приближение для итеративного процесса.

Выравнивание профиля интенсивности пучка полупроводникового лазера с помощью бинарного фазового ДОЭ продемонстрировано в /93/. В /94/ также с помощью бинарного отражающего ДОЭ с наклонным падением экспериментально показано выравнивание пучка гелий-неонового лазера в плоскости дефокусировки. Экспериментальные исследования по фокусировке гауссового пучка CO_2 лазера в прямоугольник с равномерной интенсивностью с помощью бинарного и многоградационного ДОЭ проведены соответственно в /95/ и /96/.

Во всех перечисленных работах, за исключением /85/, размер сфокусированного с помощью ДОЭ прямоугольного или круглого равно-

мерного распределения интенсивности составлял несколько минимальных дифракционных пятен. То есть полной фокусировки гауссового пучка не происходило. В /97/ кратко замечено, что выравнивать гауссовый пучок в фокальной плоскости в области минимального дифракционного пятна можно с помощью бинарной фазовой пластинки с цилиндрической ступенькой.

В работах /89, 90/ было показано, что для ДОЭ, формирующего однородное распределение интенсивности прямоугольной формы, задаваемое соотношением

$$I(x,y)=\begin{cases}1, & |x|\le d_1, \quad |y|\le d_2,\\ 0, & |x|>d_1, \quad |y|>d_2,\end{cases}$$

из плоского пучка с гауссовым профилем распределения интенсивности

$$I_0(u,v)=I_0\exp\left[-\frac{u^2+v^2}{\sigma^2}\right],$$

можно получить его фазовую функцию (при условии, что имеется линза, расположенная вплотную к ДОЭ) в виде суммы фаз по соответствующим координатам:

$$\varphi(u,v)=\varphi_1(u)+\varphi_2(v),$$

слагаемые которой удовлетворяют системе двух уравнений

$$\begin{cases}\dfrac{\mathrm{d}x}{\mathrm{d}u}=\dfrac{I_0(u)}{I(x)}\\ x=u+fk^{-1}\dfrac{\mathrm{d}\varphi(u)}{\mathrm{d}u}\end{cases}, \tag{242}$$

где f — фокальное расстояние линзы, в фокальной плоскости которой формируется световой прямоугольник, k — волновое число, (u,v), (x,y) — координаты в плоскостях ДОЭ и Фурье-спектра, соответственно, $2d_1$, $2d_2$ — размеры прямоугольника, σ — параметр гауссова пучка.

Первое уравнение в системе (242) устанавливает равенство плотности световой энергии одномерного ДОЭ с плотностью энергии соответствующих участков прямолинейного отрезка. Второе уравнение в системе (242) описывает стационарные точки в параксиальном приближении.

Дискретный вариант решения системы (242) можно записать в виде:

$$\varphi_{mn} = \ln 10 \left[M_x \left\{ \frac{6n}{N} \operatorname{erf}\left(\frac{6n}{N} \right) + \frac{1}{\sqrt{\pi}} e^{-\frac{36n^2}{N^2}} \right\} + \right.$$
$$\left. + M_y \left\{ \frac{6m}{N} \operatorname{erf}\left(\frac{6m}{N} \right) + \frac{1}{\sqrt{\pi}} e^{-\frac{36m^2}{N^2}} \right\} \right], \qquad (243)$$

где

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} \, dt, \qquad m = 0, \pm 1, \pm 2, \ldots, \pm \frac{N}{2}, \qquad n = 0, \pm 1, \pm 2, \ldots, \pm \frac{N}{2},$$

$$M_x = \frac{\pi N_x}{6\sqrt{\ln 10}}, \qquad M_y = \frac{\pi N_y}{6\sqrt{\ln 10}},$$

M_x и M_y — количество минимальных дифракционных пятен, попадающих в прямоугольник размерами $N_x \times N_y$. В этом случае квадратная апертура ДОЭ размером $N \times N$ имеет сторону квадрата длиной 6σ: $N\Delta = 6\sigma$,где Δ — расстояние между соседними пикселами в плоскости ДОЭ, а коллимированный гауссов пучок, освещающий ДОЭ, создает распределение света, задаваемое соотношением

$$I_{0mn} = \exp\left[-36N^{-2}\left(n^2+m^2\right)\right]. \qquad (244)$$

В дальнейшем изложении используется фаза (243) для численного моделирования работы геометрооптического ДОЭ.

Следуя работам /10, 11/, кратко опишем итеративные алгоритмы для расчета ДОЭ. Пусть даны комплексная амплитуда освещающего пучка $A(u)$ и требуемое распределение интенсивности $I(x)$ в фокальной

плоскости линзы. Комплексная амплитуда $F(x)$ в фокальной плоскости связана с комплексной амплитудой

$$f(u) = a(u) \exp [\mathrm{i}\varphi(u)]$$

сразу за ДОЭ преобразованием Фурье

$$F(x) = \int_{-b}^{b} f(u) e^{-\mathrm{i}\frac{k}{f}xu} \mathrm{d}x,$$

где $2b$ — размер апертуры ДОЭ. Уравнение для расчета ДОЭ принимает вид

$$|F(x)|^2 = I(x). \tag{245}$$

Чтобы найти итеративное решение уравнения (245) относительно фазы $\Phi(u)$, следует осуществить некоторую предварительную оценку фазы $\Phi_0(u)$ с последующим расчетом комплексной амплитуды света в фокальной плоскости. В этом случае комплексная амплитуда $F_n(x)$, рассчитанная на n-й итерации, заменяется на функцию $F_n^0(x)$ в соответствии с правилом

$$F_n^0(x) = \begin{cases} \sqrt{I_n(x)} \dfrac{F_n(x)}{|F_n(x)|}, & |x| \le d, \\ \sqrt{\alpha} F_n(x), & |x| > d, \end{cases} \tag{246}$$

где $I_n(x) = (1+\alpha)I(x) - \alpha|F_n(x)|^2$, $I(x)$ — заданное распределение интенсивности в интервале $[-d, d]$ фокальной плоскости, α — параметр, который контролирует скорость сходимости рассчитанной интенсивности к заданной. Для $\alpha = 0$, замена (246) переходит в стандартную замену в алгоритме ГС /3/.

Амплитуда света $f_n(u)$ в плоскости ДОЭ рассчитывается с помощью обратного преобразования Фурье и заменяется на функцию $f_n^0(x)$ согласно правилу

$$f_n^0(x)=\begin{cases} A(u)\dfrac{f_n(u)}{|f_n(u)|}, & |u|\le b, \\ 0, & |u|>b. \end{cases} \tag{247}$$

В отличие от алгоритмов условного градиента /98/, в данном случае параметр α вводится непосредственно в функцию интенсивности.

Скорость сходимости интенсивности $|F_n(x)|^2$ к заданной интенсивности проверяется по среднеквадратичному отклонению

$$\delta=\left[\frac{\int\limits_{-d}^{d}\left[I(x)-|F_n(x)|^2\right]^2 \mathrm{d}x}{\int\limits_{-d}^{d} I^2(x)\,\mathrm{d}x}\right]^{1/2}. \tag{248}$$

Ошибка (248) вычисляется для интенсивности в отличие от ошибок (9), (10), которые вычисляются для амплитуды. Кроме того, вводится еще параметр E, характеризующий энергетическую эффективность фокусировки:

$$E=\frac{\int\limits_{-d}^{d}|F_n(x)|^2\,\mathrm{d}x}{\int\limits_{-\infty}^{\infty}|F_n(x)|^2\,\mathrm{d}x}. \tag{249}$$

Формула (249) является одномерным аналогом формулы (47). Ниже применяется фаза, полученная из уравнений (245) — (247) для численного моделирования работы дифракционных ДОЭ /92/.

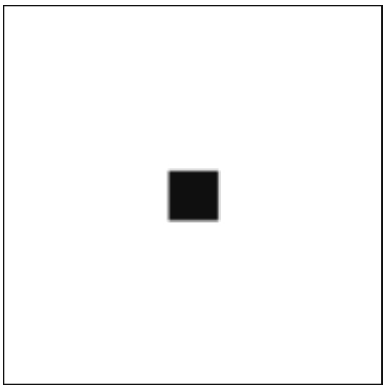


Рис. 35. Геометрооптическая фаза ДОЭ фокусирующего в квадрат

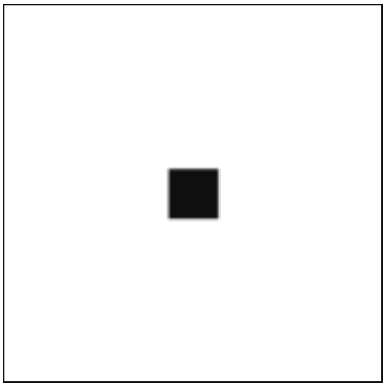


Рис. 36. Распределение интенсивности, полученное от геометрооптического ДОЭ в фокальной плоскости линзы

Пример 11. Рассмотрим ДОЭ, содержащий 32 × 32 пиксела и фокусирующий в квадрат, имеющий размеры, равные 10 минимальным дифракционным пятнам. Для ДОЭ, формирующего квадрат из коллимированного гауссова пучка, фаза, полученная из уравнения (243) и взятая по модулю 2π на массиве в 256 × 256 пикселов, находится в виде совокупности колец (линий равной фазы), переходящих в линии квадратного периметра (рис. 35). Рис. 36 иллюстрирует распределение

интенсивности света в фокальной плоскости линзы, рассчитанное как преобразование Фурье для амплитуды

$$f_{mn} = \sqrt{I_{0mn} e^{i\varphi_{mn}}},$$

где I_{0mn} взято из (244), а φ_{mn} из (243). Среднеквадратичное отклонение полученного распределения от равномерного распределения составляет 5%, а эффективность равна 91,6%.

Чтобы рассчитать ДОЭ, фокусирующий гауссов пучок в квадрат с однородной интенсивностью, в качестве начальной оценки выбирается геометрооптическая фазовая функция, описанная выше (см. рис. 35). Последующий итеративный расчет осуществлялся тремя способами. Первый подход к расчету фазы, основанный на стандартном варианте алгоритма ГС с заменой (246) для $\alpha = 0$, сначала дает увеличение ошибки δ на первых итерациях, а затем идет медленное уменьшение ошибки (рис. 37а, кривая 1). В этом случае в течение 10 итераций ошибка не уменьшается ниже 13%, хотя при этом эффективность E возрастает до 98.9% (рис. 37*б*, кривая 1).

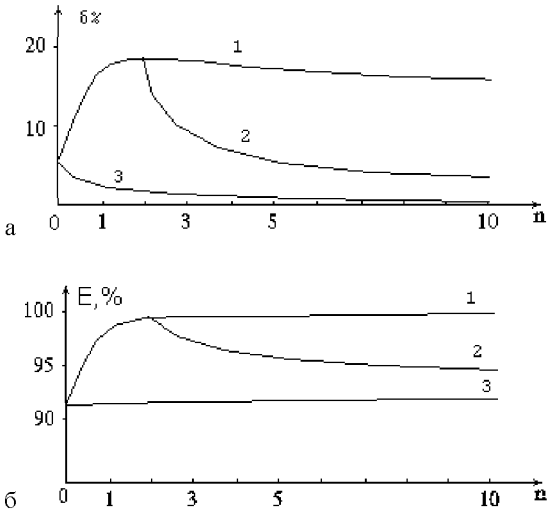


Рис. 37. Среднеквадратичное отклонение (а) и энергетическая эффективность (б) в зависимости от числа итераций для различных вариантов итеративного алгоритма: алгоритм Герчберга — Секстона (1), комбинированный (2), адаптивный (3)

Второй подход — комбинированный: первые три итерации осуществляются с заменой (246) для $\alpha = 0$, оставшиеся семь итераций для $\alpha = 1$ (рис. 37а, кривая 2). Этот метод после 10 итераций приводит к формированию квадрата равномерной интенсивности, с ошибкой 2,4% и с эффективностью 96,4% (рис. 37*б*, кривая 2). Видно, что этот подход более эффективен по сравнению с первым, т.к. он дает уменьшение ошибки с 13% до 2% (в 6 раз) без существенного уменьшения эффективности.

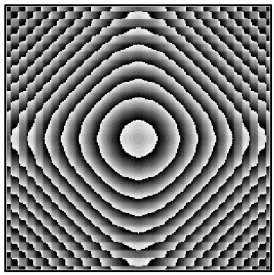


Рис. 38. Фаза ДОЭ, фокусирующего в квадрат, полученная после 10 итераций из первоначальной геометрооптической фазы

Третий подход — адаптивный, это означает, что замена (246) при $\alpha=1$ осуществляется на каждом шаге итераций (рис. 37а, кривая 3). Как видно в этом случае, ошибка уменьшается монотонно и после 10 итераций становится равной 0,1%, но при этом эффективность падает до 92,2% (рис. 37*б*, кривая 3). На рис. 38 показана фаза ДОЭ по модулю 2π, рассчитанная в течение 10 итераций на основе третьего подхода из геометрооптической фазы (243). Во-первых, видно, что применение итеративного алгоритма не приводит к существенному изменению начальной фазы. Выраженные изменения фазы (рис. 35) имеются только на краях фокусатора (рис. 38). Во-вторых, можно сделать вывод, что алгоритм, основанный на замене (246), при $\alpha = 1$ дает возможность улучшить геометрооптическую фазу таким образом, что ошибка в

формировании светового квадрата уменьшится более чем на порядок, при этом эффективность изменится несущественно.

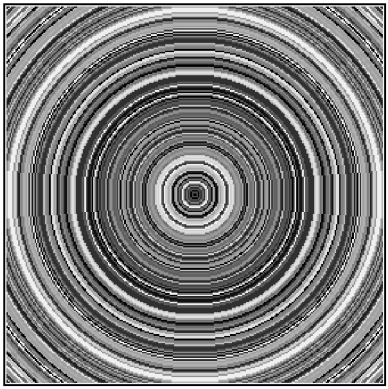


Рис. 39. Фаза со случайным радиусом, выбранная в качестве начальной оценки для итераций

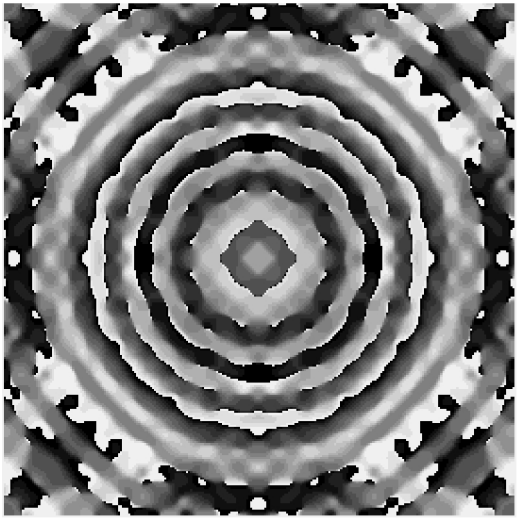


Рис. 40. Фаза ДОЭ, фокусирующего в квадрат, полученная после 10 итераций из случайной начальной фазы

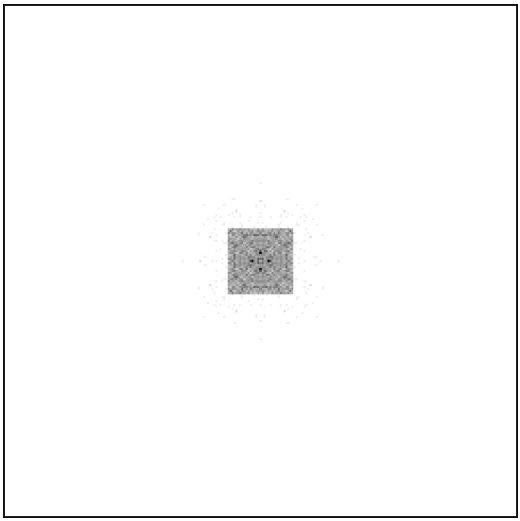


Рис. 41. Распределение интенсивности в фокусе линзы, полученное от ДОЭ, с фазой показанной на рис. 40

Для сравнения был проведен другой вычислительный эксперимент, в котором в качестве начального приближения использовалась фазовая функция со случайным радиусом (рис. 39). Распределение интенсивности, формируемое ДОЭ, имеющим вышеназванную фазу, представлено на рис. 41. При этом ошибка и эффективность составили, соответственно, $\delta = 6{,}4\%$ и $E = 80{,}8\%$,. Дальнейшие итерации не дали существенных изменений этих значений. Вышеприведенные данные показывают, что итеративный алгоритм, который начинается со случайной оценки фазы, приводит к нерегулярной структуре зон ДОЭ и дает значение точности и эффективности, несколько меньшие, чем при использовании геометрооптических ДОЭ.

В таблице 4 суммируются данные по всем ДОЭ, описанным в данном разделе. Из таблицы видно преимущество с точки зрения равномерности, эффективности, регулярности зонной структуры, которые мы получаем, используя физически обоснованное начальное приближение в виде геометрооптической фазы (сравни строки 1—4 со строкой 5). Из сравнения строк 1 и 2 таблицы видно, что метод ГС позво-

ляет получить увеличение энергетической эффективности на 7% при геометрооптическом приближении (строка 1), но приводит к значительной неоднородности интенсивности в квадрате (13% вместо 5%).

Таблица 4.

Сравнение ДОЭ.

ТИП ДОЭ И ФАЗЫ	Энергетическая эффективность E%	Относительная среднеквадратичная ошибка δ%
Геометрооптическая фаза	91,6	5,0
10 итераций алгоритма ГС на геометрооптической фазе	98,9	13,0
10 итераций адаптивно-итеративного алгоритма на геометрооптической фазе	92,2	0,1
10 комбинированных итераций на геометрооптической фазе	96,4	2,4
10 итераций на фазе со случайным радиусом	80,8	6,4

Адаптивный метод, напротив, обеспечивает высокую степень однородности интенсивности (0,1%), но он практически не улучшает эффективность по сравнению с геометрооптической аппроксимацией (строка 3). Комбинированный метод позволяет получить довольно высокую эффективность в сочетании с приемлемой однородностью интенсивности (строка 4). Из строки 5 видно, что, не зная геометрооптического решения, но основываясь на случайной начальной фазе в оценке фазы ДОЭ и используя итеративный метод, можно получить результаты, приемлемые с точки зрения точности и эффективности формирования заданного распределения интенсивности.

Таким образом, численно показано, что при расчете ДОЭ, фокусирующих свет в малые области пространственного спектра, наилучшего результата можно добиться при использовании адаптивно-итеративного алгоритма с начальной оценкой искомой фазы ДОЭ, полученной при решении обратной задачи в геометрооптическом при-

ближении. В ходе итераций регулярная структура зон фазы (рис. 47) видоизменяется, в основном на периферийной части апертуры ДОЭ.

7.8. Фокусировка в кольцо

Для фокусировки когерентного света в узкое кольцо обычно используются конические аксиконы в сочетании со сферическими линзами /36/ и бинарные аксиконы /99/. Если к фазовой функции, описывающей пропускание аксикона, добавить слагаемое, линейно зависящее от азимутального угла, полученный оптический элемент будет обладать новыми свойствами /100/.

Ниже для фокусировки в широкое кольцо с заданным распределением интенсивности вдоль радиуса кольца рассматривается метод расчета ДОЭ, фокусирующего в круглую область вне оси. Метод сочетает сведение задачи фокусировки в радиальную внеосевую область к одномерной задаче фокусировки в отрезок и использование итеративных способов решения одномерной задачи фокусировки в отрезок. Предложенный метод требует вычисления только двух преобразований Фурье на каждой итерации, в три раза сокращая время расчета по сравнению с итеративным расчетом при помощи преобразования Ханкеля, описанным в разделе 7.1.

Под узким световым кольцом в фокальной плоскости /30/ понимается такое кольцо, ширина которого меньше дифракционного предела Фраунгофера для конической волны с ограниченной апертурой, дифрагирующей на ДОЭ.

Для фокусировки в узкое кольцо можно использовать оптический элемент в виде конического аксикона с комплексной функцией пропускания в виде

$$\tau(r)=\exp(-i\alpha r), \qquad (250)$$

где α — характеристика «силы» аксикона, а r — радиальная координата. Если элемент, задаваемый уравнением (292), освещается плоским монохроматическим пучком света с волновым числом $k=2\pi/\lambda$ (где λ — длина волны), и если за этим элементом помещается сферическая лин-

за с фокальным расстоянием f, в задней фокальной плоскости линзы будет формироваться световое кольцо радиусом $R=\alpha f/k$ и шириной $b=2\lambda f/a$, где a — радиус апертуры аксикона. Комплексная амплитуда света в фокальной плоскости линзы задается дифракционным интегралом Фраунгофера, выраженным в полярных координатах:

$$F_0(\rho)=\frac{k}{f}\int_0^a e^{-i\alpha r}J_0\left(\frac{k}{f}r\rho\right)r\,dr, \tag{251}$$

где $J_0(x)$ — функция Бесселя нулевого порядка. Амплитуда пучка света на кольце рассчитывается по формуле

$$\begin{aligned}F_0(R)&=\frac{k}{f}\int_0^a e^{-i\alpha r}J_0(\alpha r)r\,dr=\\&=\frac{k}{f\alpha^2}e^{i\xi}\left\{\left[\xi J_1(\xi)-\frac{\xi^2}{3}J_2(\xi)\right]+i\frac{\xi^2}{3}J_1(\xi)\right\},\end{aligned} \tag{252}$$

где $\xi=\alpha a$, $J_1(x)$, $J_2(x)$ — функции Бесселя первого и второго порядков. Интенсивность света в центре кольца рассчитывается по формуле

$$I_0(0)=\left|\frac{k}{f}\int_0^a e^{-i\alpha r}r\,dr\right|^2=\left(\frac{ka^2}{2f}\right)^2\left[\left(\frac{\sin\nu}{\nu}\right)^2+\left(\frac{\cos\nu}{\nu}-\frac{\sin\nu}{\nu^2}\right)^2\right], \tag{253}$$

где $\nu=\alpha a/2$.

Как видно из последнего слагаемого в уравнении (253), интенсивность в центре принимает свое максимальное значение при $\alpha=0$,

$$I_0(0)\big|_{\max}=\left(\frac{ka^2}{2f}\right)^2. \tag{254}$$

Интенсивность в центре кольца асимптотически достигает нулевого значения, когда α стремится к бесконечности. Уравнение для определения локальных экстремумов интенсивности имеет вид:

$$\frac{dI_0}{d\nu}=\left(\frac{ka^2}{2f}\right)^2\left[\frac{-4}{\nu}\left(\frac{\cos\nu}{\nu}-\frac{\sin\nu}{\nu^2}\right)^2\right]=0 . \qquad (255)$$

Решение уравнения (255) имеет вид

$$\nu=\mathrm{tg}\,\nu . \qquad (256)$$

Энергетическая эффективность фокусировки в узкое кольцо на основе аксикона находится из соотношения

$$\varepsilon_0=2\pi RbI_0(R)[2\pi RbI_0(R)+\pi b^2 I_0(0)]^{-1} , \qquad (257)$$

где $b=2\lambda f/a$ - ширина кольца, практически равная радиусу центрального светового пятна, $I_0(R)$ берется из уравнения (252) и равняется

$$I_0(R)=\left(\frac{k}{f}\right)^2\frac{a^4}{9}\left[J_1^2(\xi)+J_2^2(\xi)\right]+$$
$$+\left(\frac{k}{f}\right)^2 a^4\frac{J_1(\xi)}{\xi}\left[\frac{J_1(\xi)}{\xi}-\frac{2}{3}J_2(\xi)\right] . \qquad (258)$$

Для фокусировки в узкое кольцо также можно использовать винтовой аксикон с функцией пропускания в виде /30, 32/:

$$\tau(r,\varphi)=\exp[-i\alpha r+im\varphi], \qquad (259)$$

где $m = 0, 1, 2, \ldots$, φ — азимутальный угол. Заметим, что α может принимать как положительные, так и отрицательные значения. Комплексная амплитуда света в фокальной плоскости линзы описывается в этом случае вместо уравнения (251) уравнением

$$F_m(\rho,\psi)=\frac{k}{f}i^m e^{im\psi}\int_0^a e^{-i\alpha r}J_m\left(\frac{k}{f}r\rho\right)r\,dr , \qquad (260)$$

где $J_m(x)$ — функция Бесселя m-го порядка, (ρ,ψ) — полярные координаты в плоскости Фурье. Из уравнения (260) видно, что независимо от значений α и a интенсивность в центре кольца равна нулю, $J_m(0) = 0$ для $m > 0$. С ростом порядка m интенсивность на кольце рассчитывается по формуле ($m = 1$):

$$I_1(R)=\left|F_1(R,\psi)\right|^2=\left(\frac{k}{f}\right)^2\frac{a^4}{9}\left[J_1^2(\xi)+J_2^2(\xi)\right]. \qquad (261)$$

Из сравнения уравнений (258) и (261) видно, что при использовании винтового аксикона (259) интенсивность света, падающего на кольцо, может быть как больше, так и меньше интенсивности, которая получается при использовании аксикона (250), и зависит от знака второго члена в выражении

$$I_0(R)=I_1(R)+\left(\frac{k}{f}\right)^2 a^4\frac{J_1(\xi)}{\xi}\left[\frac{J_1(\xi)}{\xi}-\frac{2}{3}J_2(\xi)\right].$$

Винтовой аксикон не дает значительного выигрыша в энергетической эффективности по сравнению с обычным аксиконом.

Рассмотрим далее приближенный метод расчета ДОЭ, фокусирующих в радиально-симметричные области. Предположим, что лазерный пучок с комплексной амплитудой

$$W_0(r)=\sqrt{I_0(r)}\exp\left(\mathrm{i}\varphi_0(r)\right),$$

где $I_0(r)$ — интенсивность освещающего пучка, $\varphi_0(r)$ — фаза пучка, падает на ДОЭ с комплексной функцией пропускания $\exp\left(i\tilde{\varphi}(r)\right)$, $r\le a$. Требуется найти фазовую функцию ДОЭ $\tilde{\varphi}(r)$, обеспечивающую формирование заданного кругового распределения интенсивности $I(\rho)$, $\rho_1\le\rho\le\rho_2$, в фокальной плоскости $z = f$ (см. рис. 42).

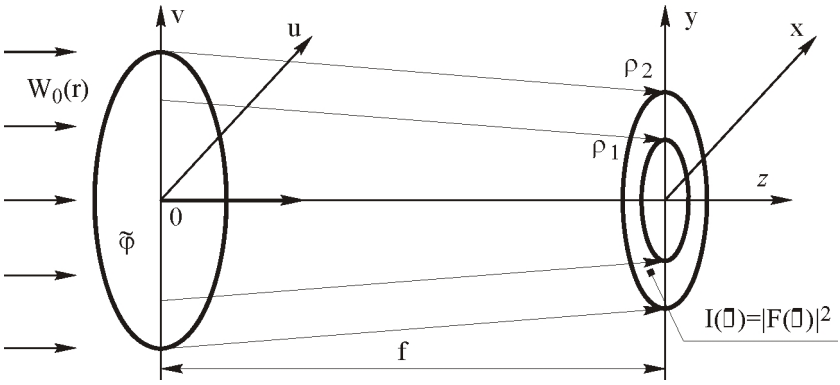


Рис. 42. Геометрия задачи фокусировки в кольцевую область

Ниже полагается, что фазовая функция ДОЭ имеет вид

$$\tilde{\varphi}(r) = \varphi(r) - \varphi_0(r). \tag{262}$$

Представление (262) позволяет рассчитывать ДОЭ без учета фазы освещающего пучка. Комплекная амплитуда поля в фокальной плоскости ДОЭ задается интегралом Френеля-Кирхгофа, выраженным в полярных координатах:

$$F(\rho) = \frac{k}{f}\exp\left(i\frac{k\rho^2}{2f}\right)\int_0^a \sqrt{I_0(r)}\exp\left(\mathrm{i}\varphi(r)\right)\exp\left(\mathrm{i}\frac{kr^2}{2f}\right)J_0\left(\frac{k}{f}r\rho\right)r\,\mathrm{d}r. \tag{263}$$

Используя асимптотическое представление функции Бесселя $J_0(\xi)$ при $\xi>>0$:

$$J_0(\xi) = \sqrt{\frac{2}{\pi\xi}}\cos\left(\xi - \frac{\pi}{4}\right),\ \xi \to \infty, \tag{264}$$

получим приближение следующего вида для $F(\rho)$ при $\rho>>0$:

$$F(\rho) = \exp\left(-\mathrm{i}\pi/4\right)\left(F_1(\rho)/\sqrt{\rho} + iF_2(\rho)/\sqrt{\rho}\right), \tag{265}$$

где

$$F_1(\rho)=\sqrt{\frac{k}{2\pi \mathrm{i} f}}\int_0^a \sqrt{I_0(r)r}\exp\left(\mathrm{i}\varphi(r)\right)\exp\left(\frac{\mathrm{i}k(\rho-r)^2}{2f}\right)\mathrm{d}r, \qquad (266)$$

$$F_2(\rho)=F_1(-\rho). \qquad (267)$$

Функция $F_1(\rho)$ соответствует комплексной амплитуде волны, формируемой одномерным ДОЭ, который освещается пучком с распределением интенсивности

$$\tilde{I}(r)=I_0(r)r. \qquad (268)$$

Задача фокусировки в радиальную область с распределением интенсивности $I(\rho)$, $\rho >> 0$, может быть сведена к одномерной задаче фокусировки в отрезок. Действительно, пусть $\varphi(r)$ есть фазовая функция одномерного ДОЭ, фокусирующего пучок (268) в отрезок с распределением интенсивности

$$I_1(\rho)=I(\rho)\rho, \quad \rho_1 \le \rho \le \rho_2 \qquad (269)$$

Заметим, что функции интенсивности $\tilde{I}(r)$ и $I_1(\rho)$ являются размерными и промежуточными функциями, которые отсутствуют в конечном результате.

Согласно уравнению (267) член $F_2(\rho)$ соответствует комплексной амплитуде поля при фокусировке в отрезок при $-\rho_1 < \rho < -\rho_2$ и не влияет на поле $F_1(\rho)$ при $\rho_1 \le \rho \le \rho_2$ (см. рис. 43).

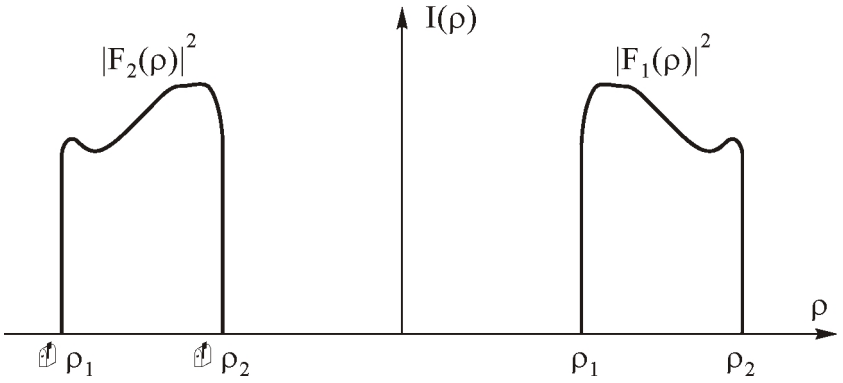


Рис. 43. Структура поля описываемая членами $F_1(\rho)$ и $F_2(\rho)$ уравнения (267)

Следовательно, аппроксимация $F(\rho)$ примет вид:

$$F(\rho) \approx \exp(-i\pi/4) F_1(\rho)/\sqrt{\rho} \,. \tag{270}$$

Соответственно, функция $|F_1(\rho)|^2 / \rho$, представляющая распределение интенсивности в фокальной плоскости радиального ДОЭ превращается в заданное распределение интенсивности $I(\rho)$ для $I_1(\rho) = |F_1(\rho)|^2$, определяемое уравнением (269).

Следует отметить, что при расчете интеграла (206) при помощи метода стационарной фазы, аппроксимация (270) приводит к известным формулам для фазовой функции фокусатора /30, 66, 67/:

$$\varphi(r) = -\frac{kr^2}{2f} + \frac{k}{f}\int_0^r \tilde{\rho}(\xi)\,\mathrm{d}\xi, \qquad r \in [0, a], \tag{271}$$

где функция $\tilde{\rho}(\xi)$ получена из следующего соотношения:

$$\int_0^r I_0(\zeta)\zeta\,\mathrm{d}\zeta = \int_{\rho_1}^{\tilde{\rho}(r)} I(\xi)\xi\,\mathrm{d}\xi \,. \tag{272}$$

В работах /30, 66/ уравнение (271) получено из уравнения эйконала, а уравнение (272) получено из уравнения переноса и соответствует интегральной форме закона сохранения светового потока.

Например, фазовая функция ДОЭ, фокусирующего плоский пучок в широкое кольцо с равномерным распределением интенсивности, может быть получена из уравнений (271) и (272) в виде:

$$\varphi(r) = -\frac{kr^2}{2f} + \frac{k\rho_1}{2fc}\left[cr\left(c^2r^2+1\right)^{1/2} + \ln\left(cr + \left(c^2r^2+1\right)^{1/2}\right)\right], \qquad (273)$$

где $c = \left(\rho_2^2 - \rho_1^2\right)^{1/2} / \left(a\rho_1\right)$, ρ_1 и ρ_2, соответственно, внутренний и внешний радиусы кольца.

Численные оценки позволяют определить фокальную внеосевую область, в которую фокусируется излучение, как область, внутренний радиус которой в 4—5 раз больше дифракционного размера $\Delta=\lambda f/a$.

Геометрооптический метод расчета $\varphi(r)$ приводит к известным уравнениям (271), (272) для фазовой функции фокусатора. Расчет $\varphi(r)$ при помощи одномерного итеративного алгоритма соответствует новому итеративному методу расчета ДОЭ, фокусирующего в круговую внеосевую область.

В разделе 7.1 рассматривается итеративный расчет ДОЭ, фокусирующих в радиальную область. При этом осуществляется пересчет поля между местоположением ДОЭ и фокальной плоскостью при помощи преобразования Ханкеля. Реализация преобразования Ханкеля сводится к вычислению трех преобразований Фурье при помощи экспоненциальной замены переменных. Следовательно, итеративный подход, описанный в разделе 7.1, требует вычисления шести преобразований Фурье на каждом итеративном шаге. Итеративный метод, рассмотренный в данном параграфе, требует вычисления двух преобразований Фурье на каждую итерацию. Ниже итеративный расчет, основанный на преобразовании Ханкеля, будет называться радиальным итеративным расчетом. Расчет радиального ДОЭ, основанный на итеративном расчете ДОЭ фокусирующего в отрезок, будет называться

линейным итеративным расчетом. Ниже проведено сравнение между ДОЭ, фокусирующими в широкое кольцо, которые рассчитывались при помощи геометрооптического подхода при выводе $\varphi(r)$ (см. уравнение (271)), при помощи линейного итеративного расчета $\varphi(r)$ и радиального итеративного расчета фазы.

Чтобы оценить характеристики ДОЭ, используем величины энергетической эффективности E и среднеквадратичного отклонения δ. Величина

$$E = \int_{\rho_1}^{\rho_2} I(\rho)\rho\,\mathrm{d}\rho \Big/ \int_{0}^{a} I_0(r) r\,\mathrm{d}r$$

характеризует долю энергии освещающего пучка, фокусируемую в заданной фокальной области. Величина

$$\delta = \frac{1}{\overline{I}} \sqrt{\frac{2}{\left(\rho_2^2 - \rho_1^2\right)} \int_{\rho_1}^{\rho_2} \left[I(\rho) - \overline{I}\right]^2 \rho\,\mathrm{d}\rho}$$

характеризует среднеквадратичное отклонение распределения интенсивности от среднего значения

$$\overline{I} = \frac{2}{\left(\rho_2^2 - \rho_1^2\right)} \int_{\rho_1}^{\rho_2} I(\rho)\rho\,\mathrm{d}\rho \,.$$

Пример 12. Линейный и радиальный расчет $\varphi(r)$ осуществлялись при помощи адаптивно-аддитивного алгоритма (40), описанного в разделе 3. Фазовая функция (271), полученная геометрооптическим методом выбиралась в качестве начальной в итеративном процессе. Пересчет поля между плоскостями производился при помощи быстрого преобразования Фурье с числом пикселов N = 256 для следующих параметров: λ = 1,06 мкм, a = 2,5 мм, f = 250 мм. В табл. 5 представлены значения E и δ в зависимости от параметра $S = \left(\rho_2 - \rho_1\right)/\Delta$, $\Delta = \lambda f / a$, для ρ_1 = 28Δ. Величина S характеризует

ширину фокального кольца по сравнению с дифракционным размером Δ.

Из данных, представленных в табл. 5 видно, что среднеквадратичная ошибка δ для ДОЭ, рассчитанных при помощи линейного итеративного алгоритма, в 3—5 раз меньше, чем для фокусатора, при этом энергетическая эффективность E меняется незначительно. Радиальный итеративный алгоритм не обеспечивает существенного улучшения E и δ по сравнению с ДОЭ, рассчитанными при помощи линейного алгоритма.

Таблица 5.

Параметры E и δ для фокусатора и ДОЭ, рассчитанного при помощи линейного и радиального итеративных алгоритмов

S	Геометрооптический метод		Линейный итеративный алгоритм		Радиальный итеративный алгоритм	
	E(%)	δ (%)	E(%)	δ (%)	E(%)	δ (%)
8	90,4	48,0	89,1	13,1	89,2	13,3
10	89,9	39,4	92,5	11,1	92,6	11,0
12	90,7	37,4	89,4	9,7	89,5	9,6
14	90,9	40,0	91,3	9,1	91,5	9,2
16	90,2	37,7	90,4	7,8	90,7	7,7
18	89,7	34,5	90,8	7,4	90,7	7,2
20	89,7	32,4	90,1	6,9	90,2	6,6

На рис. 44 показаны фазовые функции фокусатора и ДОЭ, рассчитанных при помощи линейного и радиального алгоритмов, соответственно, для $S = 10$. Расчетные значения распределений интенсивности на рис.45 выявляют практически полное отсутствие флуктуаций интенсивности на фокальном кольце для “линейного” и “радиального” ДОЭ. Фокусатор и линейный ДОЭ создают пики интенсивности при $\rho = 0$. Наличие пиков объясняется тем фактом, что как геометрооптический метод, так и линейный итеративный алгоритм не контролируют распределение интенсивности вблизи оптической оси. Для радиального ДОЭ центральный пик не появлялся.

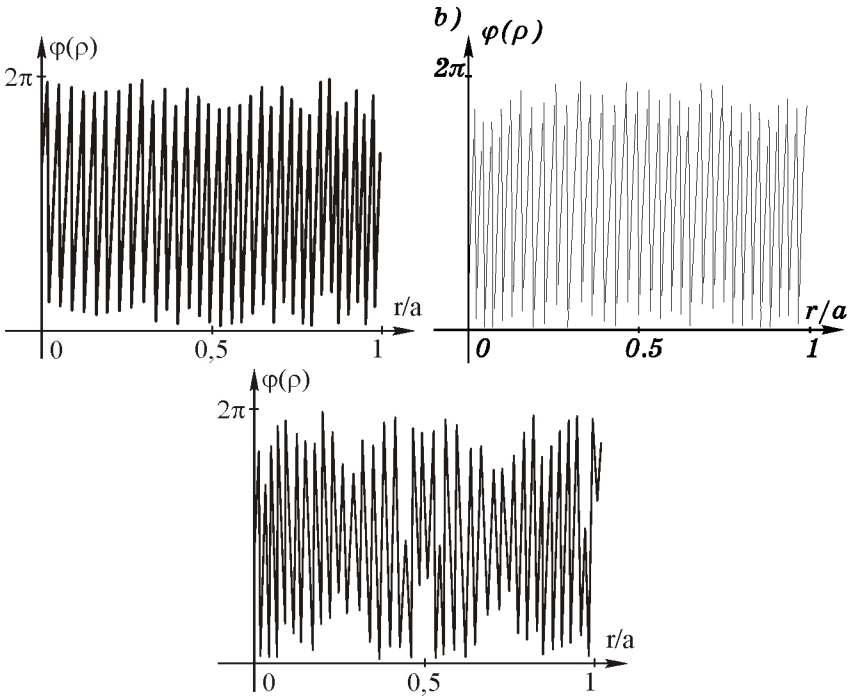


Рис. 44. Фазовые функции фокусатора (а), ДОЭ, рассчитанного при помощи линейного итеративного алгоритма (б), и ДОЭ, рассчитанного при помощи радиального итеративного алгоритма (в)

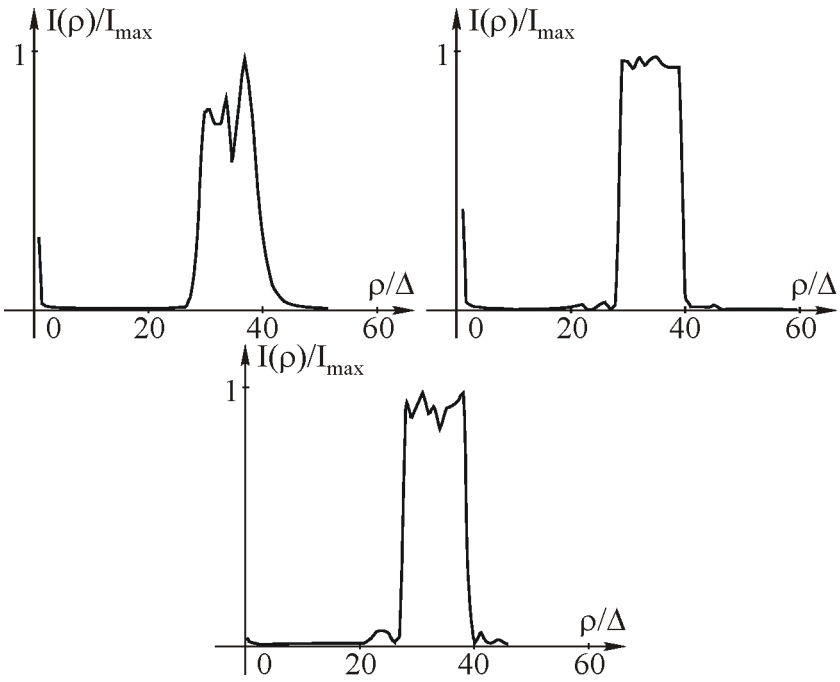


Рис. 45. Распределение интенсивностей в фокальной плоскости: фокусатора (а), ДОЭ, рассчитанного при помощи линейного итеративного алгоритма (б), ДОЭ, рассчитанного при помощи радиального итеративного алгоритма (в)

7.9. Композиционные ДОЭ, формирующие контурные изображения

В /101/ предложена фокусировка лазерного излучения в буквы и цифры с помощью “составных” фокусаторов. Апертура таких фокусаторов состоит из N непересекающихся субапертур, каждая из которых фокусирует свет либо в отрезок прямой линии, либо в узкое полукольцо, из которых состоят буквы. Недостатком таких “составных” ДОЭ является то, что ширина линий фокусировки определяется не размером апертуры ДОЭ, а размером субапертуры. Это приводит к уширению элементов букв и к уменьшению плотности световой энергии. Кроме того, наличие скачков фазы ДОЭ вдоль общих линий соседних

субапертур приводит к рассеянию света и уменьшению энергетической эффективности.

В /76/ предложен итеративный алгоритм расчета ДОЭ, фокусирующих лазерный свет в объем с заданным распределением интенсивности в N сечениях перпендикулярных к оси распространения света. Этот алгоритм основан на представлении функции пропускания ДОЭ в виде линейной комбинации комплексных амплитуд, являющихся результатом обратного преобразования Френеля от функций, описывающих световое поле в заданных сечениях объекта фокусировки.

В данном разделе рассматривается метод /102/ для расчета композиционных ДОЭ. Функция пропускания такого ДОЭ представляется в виде линейной комбинации комплексных амплитуд, являющихся Фурье-образами N непересекающихся изображений. Модули коэффициентов линейной комбинации характеризуют распределение световой энергии пучка между N изображениями, а аргументы коэффициентов являются свободными параметрами задачи. Фаза композиционного ДОЭ можно найти методом последовательных приближений с помощью итеративного алгоритма, аналогичного /3, 77, 79/ или градиентного алгоритма наискорейшего спуска /5/.

В /91/ приведены полученные в лучевом приближении выражения для фазовой функции ДОЭ, фокусирующего плоскую волну света в поперечный отрезок прямой линии с постоянной интенсивностью и расположенный на заданном расстоянии. Если ограничить рассмотрение формированием с помощью ДОЭ только контурных изображений, состоящих из участков прямых линий (например, букв), то лучевое решение из работы /91/ можно выбрать за базовое. Тогда любой отрезок в Фурье-плоскости можно получить путем растяжения (сжатия), смещения и поворота базового отрезка прямой. А фаза ДОЭ, фокусирующих в произвольный отрезок, получается путем растяжения (сжатия), смещения и поворота базового решения для фазы /91/.

Пусть плоский пучок лазерного света с длиной волны λ фокусируется ДОЭ с функцией пропускания $\exp[\mathrm{i}\varphi(x,y)]$ в контурное изобра-

жение на расстоянии f_0 в плоскости (ξ, η), состоящее из набора отрезков длиной d_n с центрами в точках (ξ_n, η_n) и расположенных под углами φ_n к оси ξ.

Пусть комплексная функция $f_n(\xi, \eta)$, $(\xi, \eta) \in \Omega_n$ описывает световое поле в плоскости (ξ, η). Квадрат модуля этой функции пусть задает распределение интенсивности в n-м изображении, заключенном в области Ω_n. И пусть набор таких областей не пересекается:

$\Omega_n \cap \Omega_m = 0$. Тогда функции $f_n(\xi, \eta)$ и их Френель-образы $F_n(x,y)$ будут ортогональными:

$$\int_{-\infty}^{\infty}\int f_n(\xi,\eta) f_m^*(\xi,\eta)\,\mathrm{d}\xi\,\mathrm{d}\eta = \int_{-\infty}^{\infty}\int F_n(x,y) F_m^*(x,y)\,\mathrm{d}x\,\mathrm{d}y = Q_n\delta_{nm}, \quad (274)$$

$$Q_n = \iint_{\Omega} |f_n(\xi,\eta)|^2\,\mathrm{d}\xi\,\mathrm{d}\eta = \int_{-\infty}^{\infty}\int |F_n(x,y)|^2\,\mathrm{d}x\,\mathrm{d}y, \quad (275)$$

$$F_n(x,y) = \frac{\mathrm{i}ke^{\mathrm{i}kf_0}}{2\pi f_0}\iint_{\Omega_n} f_n(\xi,\eta) e^{\mathrm{i}\frac{k}{2f_0}\left[(x-\xi)^2+(y-\eta)^2\right]}\,\mathrm{d}\xi\,\mathrm{d}\eta. \quad (276)$$

Формула (276) является преобразованием Френеля (1), записанным в других обозначениях. Исходя из уравнений (274) — (276) можно сформулировать задачу расчета фазовой функции $\varphi(x,y)$ дифракционного элемента, фокусирующего световое поле с известной амплитудой $A_0(x,y)$ в набор непересекающихся изображений $f_n(\xi, \eta)$ в плоскости (ξ, η), расположенной на расстоянии f_0 от ДОЭ:

$$A_0(x,y)e^{\mathrm{i}\varphi(x,y)} = \sum_{n=1}^{N} C_n F_n(x,y), \quad (277)$$

$$C_n = Q_n^{-1}\iint_W A_0(x,y)e^{\mathrm{i}\varphi(x,y)} F_n^*(x,y)\,\mathrm{d}x\,\mathrm{d}y, \quad (278)$$

где W — функция апертуры ДОЭ. В уравнении (277) квадраты модулей коэффициентов $|C_n|^2$ считаются заданными числами, пропорцио-

нальными части световой энергии, идущей на формирование n-го изображения $f_n(\xi,\eta)$. Аргументы коэффициентов ν_n=arg C_n являются искомыми параметрами задачи, также как и функция $\varphi(x,y)$.

Вместо уравнения (277) для расчета композиционных ДОЭ можно исходить из функционала среднеквадратичной невязки ε:

$$\varepsilon = \iint\limits_W \left[A_0(x,y) - \left| \sum_{n=0}^{N} C_n F_n(x,y) \right| \right]^2 \mathrm{d}x\,\mathrm{d}y \,. \tag{279}$$

Решение уравнения (277) эквивалентно минимизации функционала (279). Минимум критерия можно искать с помощью градиентного метода наискорейшего спуска. В данном случае, следуя работе /5/, градиентный метод можно представить в следующем виде:

$$\nu_n^{(p+1)} = \nu_n^{(p)} + t_p \frac{\partial \varepsilon}{\partial \nu_n^{(p)}} \,, \tag{280}$$

$$\frac{\partial \varepsilon}{\partial \nu_n^{(p)}} = 2 \left| C_n \overline{C}_n^{(p)} \right| \sin\left(\nu_n^{(p)} - \overline{\nu}_n^{(p)} \right), \tag{281}$$

$$\overline{C}_n^{(p)} = Q^{-1} \iint\limits_W R^{(p)}(x,y) F_n^*(x,y)\,\mathrm{d}x\,\mathrm{d}y \,, \tag{282}$$

$$R^{(p)}(x,y) = A_0(x,y) \frac{U^{(p)}(x,y)}{\left| U^{(p)}(x,y) \right|} - U^{(p)}(x,y) \,, \tag{283}$$

$$U^{(p)}(x,y) = \sum_{n=1}^{N} \left| C_n \right| e^{i\nu_n^{(p)}} F_n(x,y) \,, \tag{284}$$

где $\left|C_n\right|$ — произвольно заданные числа. Величина шага t_p алгоритма (280) на p-ой итерации может быть выбрана следующим образом /5/:

$$t_p = \varepsilon_p \overline{\varepsilon}_p^{-1} \,, \tag{285}$$

$$\overline{\varepsilon}_p = -\sum_{n=1}^{N}\left|\frac{\partial\varepsilon}{\partial\nu_n^{(p)}}\right|^2 . \qquad (286)$$

Заметим, что так как в данном алгоритме фазы изображений $f_n(x,y)$ задаются произвольно, то у задачи (277) может быть множество решений, из которых можно выбирать лучшее в смысле максимальной эффективности

$$E = \left[\sum_{n=1}^{N}|C_n|^2 \iint_{\Omega}|g_n(\xi,\eta)|^2\,\mathrm{d}\xi\,\mathrm{d}\eta\right]\left[\iint_{W}A_0^2(x,y)\,\mathrm{d}x\,\mathrm{d}y\right]^{-1} \qquad (287)$$

или минимальной ошибки формирования

$$\delta = \left[\sum_{n=1}^{N}\iint_{\Omega}\left[|g_n(\xi,\eta)| - \left|g_n^{(p)}(\xi,\eta)\right|\right]^2\mathrm{d}\xi\,\mathrm{d}\eta\right]\left[\sum_{n=1}^{N}\iint_{\Omega}|g_n(\xi,\eta)|^2\,\mathrm{d}\xi\,\mathrm{d}\eta\right]^{-1}, \qquad (288)$$

где $|g_n(\xi,\eta)|$ и $\left|g_n^{(p)}(\xi,\eta)\right|$ — заданный и рассчитанный на p-ой итерации модули изображений.

Если изображение $g(x,y)$ состоит из N подобных элементарных изображений $g_n(x,y)$, то все изображения можно получить из одного преобразованием подобия. Например, если контурное изображение состоит из N отрезков прямых, то все функции $g_n(x,y)$, $n=\overline{1,N}$ этих отрезков могут быть получены из некоторой базовой функции $g_0(x,y)$ путем смещения T, поворота R и масштабирования (сжатия и растяжения) D:

$$g_n(\xi,\eta) = DTRg_0(\xi,\eta), \qquad (289)$$

$$Rg_0(\xi,\eta) = g_0(\xi',\eta'), \qquad (290)$$

$$\begin{cases}\xi' = \xi\cos\varphi_n + \eta\sin\varphi_n \\ \eta' = -\xi\sin\varphi_n + \eta\cos\varphi_n\end{cases}, \qquad (291)$$

$$Tg_0(\xi,\eta) = g_0(\xi - \xi_n, \eta - \eta_n), \tag{292}$$

$$Dg_0(\xi,\eta) = Dg_0(\alpha_n\xi, \alpha_n\eta), \tag{293}$$

где φ_n — угол поворота, (ξ_n, η_n) — координаты центра, α_n — коэффициент масштабирования для n-го изображения.

Для контурного изображения, состоящего из набора отрезков, можно воспользоваться результатом работы /91/, в которой приведено выражение для фазы ДОЭ, фокусирующего в плоскую волну в отрезок прямой с постоянной интенсивностью:

$$\varphi_0(x,y) = -\frac{k}{2f_0}\left[x^2\left(1-\frac{a}{2A}\right)+y^2\right], \tag{294}$$

где a — длина отрезка, вытянутого вдоль оси ξ и имеющего минимальную дифракционную ширину по оси η, $2A$ — размер квадратной апертуры ДОЭ.

Считая фазовую функцию (294) за базовое решение нетрудно получить фазу ДОЭ, формирующего произвольный тонкий отрезок в плоскости (ξ, η) :

$$\varphi_n(x,y) = -\frac{k}{2}\left[A_n\left(x-x_n\right)^2 + B_n\left(y-y_n\right)^2 + C_n\left(x-x_n\right)\left(y-y_n\right)\right], \tag{295}$$

$$A_n = b_n^{-2}\cos^2\varphi_n + a^{-2}\sin^2\varphi_n,\quad B_n = b_n^{-2}\sin^2\varphi_n + a^{-2}\cos^2\varphi_n,$$

$$C_n = \left(ab_n\right)^{-2}\sin(2\varphi_n)\left(a^2+b_n^2\right),\quad a = f_0,\qquad b_n = f_0\left(1-\frac{a_n}{2R}\right)^{-1};$$

где R — радиус круглой апертуры ДОЭ.

В этом случае уравнение (277) для расчета фазовой функции ДОЭ $\bar{\varphi}(x,y)$, фокусирующего в контур из отрезков будет иметь вид:

$$A_0(x,y)e^{i\bar{\varphi}(x,y)} = \sum_{n=1}^{N} C_n e^{i\varphi_n(x,y)}, \tag{296}$$

где $\varphi_n(x,y)$ — функция, определенная уравнением (295). Аналитическая форма задания фазы элементарных изображений $\varphi_n(x,y)$ удобна с практической точки зрения. Действительно, решение задачи (296) с помощью итеративного градиентного алгоритма (280) — (284) можно реализовать численно, используя массивы чисел небольшой размерности (например 256 × 256), чтобы сократить временные затраты на поиск аргументов коэффициентов C_n . Для расчета функции $\overline{\varphi}(x,y)$ из уравнения (296) при генерации фотошаблона сумма, стоящая в правой части уравнения (296) вычисляется уже с использованием массива чисел большой размерности ($10^4 \times 10^4$ и более).

Решения (294) и (295) получены в лучевом приближении и обеспечивают высокую энергетическую эффективность фокусировки в требуемый световой отрезок, но не обеспечивают высокой точности постоянства интенсивности вдоль отрезка. Получить высокую точность фокусировки в отрезок можно с помощью скалярного дифракционного приближения. В этом случае вместо функции (294) следует выбирать в качестве базового решения комплексную амплитуду вида

$$f_0(\xi,\eta) = \mathrm{rect}\left(\frac{\xi}{a}\right)\mathrm{rect}\left(\frac{\eta}{b}\right), \tag{297}$$

из которой нетрудно получить функции, описывающие произвольный отрезок в плоскости (ξ,η) :

$$f_n(\xi,\eta) = \mathrm{rect}\left(\frac{\xi\cos\varphi_n + \eta\sin\varphi_n - \xi_n}{b_n}\right)\times$$
$$\times\,\mathrm{rect}\left(\frac{-\xi\sin\varphi_n + \eta\cos\varphi_n - \eta_n}{a}\right) \tag{298}$$

Тогда уравнение для расчета фазы ДОЭ, фокусирующего свет в контур из N отрезков будет иметь вид (вместо уравнения (296))

$$A_0(x,y)e^{i\overline{\varphi}(x,y)} = \sum_{n=1}^{N} C_n \operatorname{sinc}\left[\frac{b_n k}{2f_0}\left(x\cos\varphi_n + y\sin\varphi_n\right)\right]\times$$
$$\times\operatorname{sinc}\left[\frac{ak}{2f_0}\left(-x\sin\varphi_n + y\cos\varphi_n\right)\right]\cdot e^{-i\frac{k}{f_0}\left(x\xi_{2n}+y\eta_n\right)}, \qquad (299)$$

где $\operatorname{sinc}(x) = \sin(x)/x$, b_n — длина n-го отрезка, a — дифракционная ширина всех отрезков.

Пример 13. При численном моделировании в формуле (295) использовались следующие параметры: f_0=100 мм, R=10 мм, x_n, y_n, $a_n \in$ [0, 0,5 мм], $\varphi_n \in [0, \pi]$, размерность дискретной матрицы 256×256.

На рис.46—48 представлены результаты расчета ДОЭ (рис.46а—48а), фокусирующих в различные контурные фигуры: А (рис. 46), Y (рис. 47), О (рис. 48) и результаты моделирования работы таких ДОЭ, посредством вычисления распределения интенсивности светового поля на расстоянии f_0 (рис. 46б — 48б) с помощью преобразования Френеля.

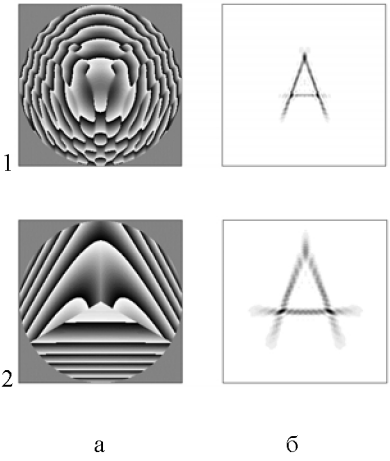


Рис. 46. Рассчитанные фазы ДОЭ для фокусировки в букву “А”: фазы композиционного (1а) и составного (2а) ДОЭ и сформированные ими картины дифракции (1б) и (2б) соответственно

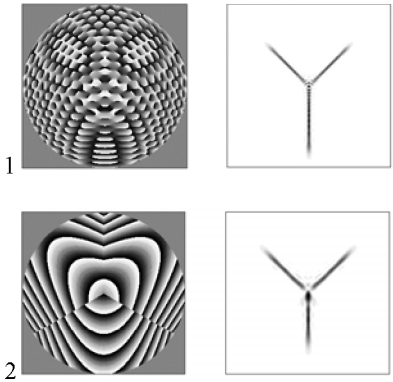


Рис. 47. Рассчитанные фазы ДОЭ для фокусировки в букву “Y”: фазы композиционного (1а) и составного (2а) ДОЭ и сформированные ими картины дифракции (1б) и (2б) соответственно

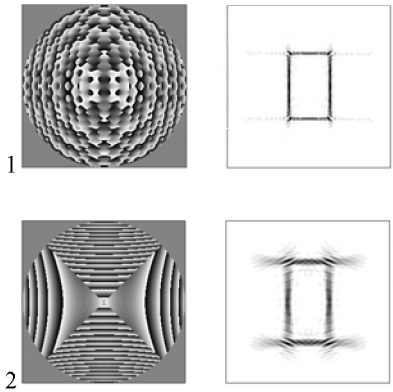


Рис. 48. Рассчитанные фазы ДОЭ для фокусировки в букву “O”: фазы композиционного (1а) и составного (2а) ДОЭ и сформированные ими картины дифракции (1б) и (2б) соответственно

При этом в верхней строке каждого рисунка показаны ДОЭ, рассчитанные по формуле (296), а в нижней строке—ДОЭ, являющиеся просто композицией угловых сегментов, рассчитанных по формуле (294). Из рисунка видно, что чем больше линий будет в контурном изображении, тем большее количество сегментов должен будет содер-

жать ДОЭ. А значит, как было замечено ранее, на каждую линию изображения будет приходиться меньший участок полезной площади ДОЭ, что приводит к уширению линий и уменьшению плотности световой энергии. Этот эффект можно наблюдать и в представленных изображениях, где используется всего 3—4 линии.

Кроме того, к недостаткам сегментного ДОЭ следует отнести необходимость подбора наилучшей композиции сегментов, в то время как формула (296) обеспечивает полный автоматизм.

7.10. Квантованные ДОЭ для фокусировки в заданную двумерную область

Фазовые ДОЭ для фокусировки в заданную двумерную область с требуемым распределением интенсивности важны в задачах преобразования и выравнивания распределения интенсивности лазерного пучка.

Технология фотолитографии, традиционно используемая для изготовления ДОЭ, предусматривает квантование фазовой функции ДОЭ по M уровням. Квантованные ДОЭ с малым числом уровней квантования ($M = 3, 4$) и особенно бинарные ДОЭ ($M = 2$) представляют особый интерес вследствие простоты их изготовления методами фотолитографии. В частности, бинарный рельеф формируется за одно травление подложки.

Хорошо известные итерационные алгоритмы, такие, как алгоритм уменьшения ошибки, алгоритм ГС и градиентные алгоритмы /3, 5, 7, 8, 98, 103/, предназначенные для расчета ДОЭ с непрерывными фазовыми функциями, оказываются неработоспособными в задачах расчета ДОЭ с малым числом уровней квантования ($M = 2, 3, 4$). В работах Ф. Вировского /21, 104—106/ представлена модификация итерационного алгоритма ГС для расчета квантованных ДОЭ при $M > 2$. Указанная модификация основана на введении в итерационный процесс операции частичного квантования и обладает достаточно хорошей сходимостью при фокусировке во внеосевые фокальные области даже при числе уровней квантования $M = 3, 4$. В случае расчета квантованных

ДОЭ для фокусировки в области вблизи оптической оси сходимость указанного алгоритма существенно хуже. В частности, при $M = 3, 4$ среднеквадратичная ошибка формирования заданного распределения интенсивности в области фокусировки составляет, как правило, 25—30%.

Расчет бинарных ДОЭ ($M = 2$) основан на методах, существенно отличных от быстрых итерационных алгоритмов типа алгоритма ГС. Например, для расчета многопорядковых бинарных дифракционных решеток обычно используются методы стохастической оптимизации такие как, прямой бинарный поиск и алгоритм симуляции отжига /52—58/. Градиентный метод расчета бинарных решеток, рассмотренный в п. 8.4, является эффективным лишь в задачах расчета одномерных решеток. Методы стохастической оптимизации в общем двумерном случае требуют больших вычислительных затрат и оказываются малопригодными при фокусировке в области размер которых на порядок превышает размер дифракционного пятна.

В данном параграфе рассматривается итерационный метод расчета квантованных ДОЭ. Метод по сложности и вычислительным затратам эквивалентен алгоритму ГС и состоит из следующих шагов: 1) аппроксимация функции комплексного пропускания (ФКП) квантованного ДОЭ непрерывной ФКП; 2) градиентный метод для оптимизации непрерывной ФКП; 3) обратный переход к квантованной ФКП.

По сравнению с алгоритмом, приведенным в работах Ф. Вировского, предлагаемый метод более общий и пригоден для расчета как квантованных ДОЭ, так и бинарных ДОЭ. По сравнению с известными методами /52—58/, ориентированными на расчет бинарных фазовых дифракционных решеток (ФДР), предлагаемый метод обладает следующими преимуществами: учет распределения интенсивности освещающего пучка, существенная двумерность, актуальная для нефакторизуемых областей фокусировки и для достижения высокой энергетической эффективности и высокая, по сравнению с методами стохастической оптимизации, скорость сходимости.

Пусть требуется рассчитать ДОЭ с квантованной по M уровням фазовой функцией $\varphi_M(\mathbf{u})$, который при освещении пучком с распределением интенсивности $I_0(\mathbf{u})$ и сходящимся сферическим волновым фронтом (фокус f) , формирует в плоскости $z = f$ область D с требуемым распределением интенсивности $I_D(\mathbf{x}), \mathbf{x} \in D$ (см. рис. 49). Жирным шрифтом здесь и далее обозначены вектора.

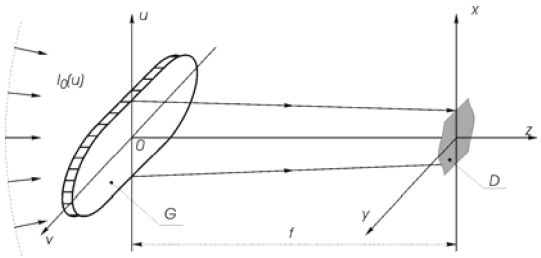


Рис. 49. Геометрия задачи фокусировки

Распределение интенсивности $I(\mathbf{x})$ в плоскости $z=f$ в скалярном приближении связано с фазовой функцией ДОЭ $\varphi_M(\mathbf{u})$ через интеграл Френеля-Кирхгофа:

$$I(\mathbf{x}) = \left| \frac{1}{\lambda f} \int_G \sqrt{I_0(\mathbf{u})} \exp\left(\mathrm{i}\varphi_M(\mathbf{u}) - \mathrm{i}\frac{k}{f}\mathbf{x}\mathbf{u} \right) \mathrm{d}^2\mathbf{u} \right|^2, \qquad (300)$$

где G — апертура ДОЭ. Уравнение (300) отличается от (5) только обозначениями.

Функция комплексного пропускания ДОЭ с квантованной по M уровням фазовой функцией $\varphi_M(\mathbf{u})$, принимающей значения $\varphi_j = j(2\pi/M)$, $j = \overline{0, M-1}$, может быть представлена в виде:

$$A_M(\mathbf{u}) = \exp\left(\mathrm{i}\Phi_M\left[\varphi(\mathbf{u})\right]\right), \qquad (301)$$

где $\varphi(\mathbf{u})$ — некоторая непрерывная функция, определенная в интервале $[0,2\pi)$, а функция

$$\Phi_M[\xi] = \frac{2\pi}{M}\,\mathrm{int}\left(\frac{\xi M}{2\pi}\right), \tag{302}$$

описывает квантование по M уровням. Функция int (x) в (302) является функцией выделения целой части x.

Таким образом, уравнение (301) представляет ФКП квантованного ДОЭ как квантованную функцию $\Phi_M[\xi]$ от непрерывного аргумента $\varphi(\mathbf{u})$.

Основная идея метода расчета квантованных ДОЭ состоит во введении непрерывной функции $\tilde{A}_M(\mathbf{u})$, аппроксимирующей разрывную ФКП $A_M(\mathbf{u})$ квантованного ДОЭ /107 —110/. При этом последующее использование быстрых градиентных алгоритмов для оптимизации непрерывной функции $\tilde{A}_M(\mathbf{u})$ позволяет предложить эффективный метод расчета квантованной ФКП (301).

Для построения непрерывной аппроксимации $\tilde{A}_M(\mathbf{u})$ разложим функцию $\exp(\mathrm{i}\Phi_M[\xi])$ в ряд Фурье:

$$\exp(\mathrm{i}\Phi_M[\xi]) = \sum_{n=-\infty}^{\infty} c_n \exp(\mathrm{i}n\xi), \tag{303}$$

где

$$c_n = \begin{cases} (-1)^{(n-1)/M}\,\mathrm{sinc}\left(\dfrac{\pi n}{M}\right)\exp\left(\mathrm{i}\dfrac{-\pi}{M}\right), & n = 1 + pM, \quad p = \pm 1, \pm 2, \ldots \\ 0, \quad \textit{иначе} \end{cases} \tag{304}$$

— коэффициенты Фурье. Полагая $\xi = \mathrm{mod}_{2\pi}[\varphi(\mathbf{u})]$ в (345), представим ФКП бинарного ДОЭ с учетом 2π периодичности в виде

$$\exp\left(\mathrm{i}\,\Phi_M\left[\varphi(\mathbf{u})\right]\right)=\sum_{n=-\infty}^{\infty}c_n\exp\left(\mathrm{i}\,n\varphi(\mathbf{u})\right). \tag{305}$$

Разложение (305) соответствует суперпозиции пучков с фазовыми функциями:

$$\varphi_n(\mathbf{x})=n\varphi(\mathbf{x}). \tag{306}$$

Квадраты модулей коэффициентов Фурье (304) представляют распределение энергии между пучками, а функции (306) определяют структуру изображений, формируемых в дифракционных порядках квантованного ДОЭ /111, 112/.

Введем вспомогательную функцию

$$\tilde{A}_M(\mathbf{u};L_1,L_2)=\sum_{p=-L_1}^{L_2}c_{1+pM}\exp\left(\mathrm{i}\left(1+pM\right)\varphi(\mathbf{u})\right). \tag{307}$$

С формальной точки зрения функция $\tilde{A}_M(\mathbf{u};L_1,L_2)$ является отрезком ряда (307), аппроксимирующего ФКП квантованного ДОЭ. При функции $\tilde{A}_M(\mathbf{u};L_1,L_2)$ (307) распределение интенсивности в плоскости $z=f$

$$I(\mathbf{x};M)=\left|\frac{1}{\lambda f}\int_G\sqrt{I_0(\mathbf{u})}\tilde{A}_M(\mathbf{u};L_1,L_2)\exp\left(-\mathrm{i}\frac{k}{f}\mathbf{x}\mathbf{u}\right)\mathrm{d}^2\mathbf{u}\right|^2. \tag{308}$$

учитывает вклад $K=L_2-L_1+1$ ненулевых порядков квантованного ДОЭ (301) и интерференцию между ними. Введение функции $\tilde{A}_M(\mathbf{u};L_1,L_2)$ (307) позволяет перейти от квантованной ФКП (301) к непрерывной функции. При этом (при разумном выборе параметров L_1,L_2) оказывается эффективным использование хорошо разработанных итеративных алгоритмов /3, 5, 7, 8, 98, 103/, предназначенные для оптимизации непрерывных функций.

Задачу расчета квантованного ДОЭ (301) предлагается заменить задачей расчета ДОЭ с непрерывной ФКП (307) для формирования

требуемого распределения интенсивности $I_D(\mathbf{x})$ при $\mathbf{x} \in D$ /107 — 110/. В этом случае расчет квантованного ДОЭ сводится к выполнению следующих шагов:

1. Замена ФКП квантованного ДОЭ функцией $\tilde{A}_M(\mathbf{u}; L_1, L_2)$;

2. Расчет $\varphi(\mathbf{u})$ из условия формирования ДОЭ с ФКП (307) заданного распределения интенсивности $I_D(\mathbf{x})$, $\mathbf{x} \in D$;

3. Обратный переход от функции $\tilde{A}_M(\mathbf{u}; L_1, L_2)$ (307) к квантованной функции (301).

Для расчета $\varphi(\mathbf{u})$ на шаге 2 могут быть использованы различные итерационные алгоритмы /3, 5, 7, 8/. Рассмотрим расчет $\varphi(\mathbf{u})$ с использованием градиентных алгоритмов /98/.

Введем некоторый интегральный функционал $\varepsilon(\varphi)$, характеризующий отличие распределения интенсивности (308), формируемого ДОЭ с непрерывной ФКП (307), от требуемого распределения интенсивности $I_D(\mathbf{x})$

$$\varepsilon(\varphi) = \bar{\varepsilon}\left(\tilde{I}_M(\mathbf{x}), I_D(\mathbf{x})\right). \tag{309}$$

Градиентная минимизация функционала (309) состоит в итеративной коррекции функции $\varphi(\mathbf{x})$ по правилу

$$\varphi_{n+1}(\mathbf{u}) = \varphi_n(\mathbf{u}) - \varepsilon \copyright(\varphi_n) \cdot t, \tag{310}$$

где индекс n — номер итерации, $\nabla\varepsilon(\varphi)$ — градиент функционала, t — шаг градиентного метода. Проводя вычисления, аналогичные рассмотренным в пунктах 7 и 8.5, несложно получить приращение функционала $\varepsilon(\varphi)$, вызванное малым изменением $\Delta\varphi$ фазы $\varphi(\mathbf{u})$ в следующем виде:

$$\Delta\varepsilon(\varphi)=\varepsilon(\varphi+\Delta\varphi)-\varepsilon(\varphi)=2\,\mathrm{Re}\left(\int_{-\infty}^{\infty}\sqrt{I_0(\mathbf{u})}\Psi^*(\mathbf{u})(\Delta\Phi)\,\mathrm{d}^2\mathbf{u}\right), \quad (311)$$

где $\Delta\Phi=\Delta\tilde{A}_M(\mathbf{u};L_1,L_2)$ — приращение ФКП ДОЭ, вызванное малым изменением фазы $\Delta\varphi$.

Расчет функции $\Psi(\mathbf{u})$ в (311) проводится в три этапа:

1) расчет комплексной амплитуды поля *W(***x***)* в плоскости $z = f$;

2) замена функции *W(***x***)* на функцию

$$W_1(\mathbf{x})=\frac{\mathrm{d}\,\overline{\varepsilon}\left(\tilde{I}_M,I_D\right)}{\mathrm{d}\,\tilde{I}_M}W(\mathbf{x}), \quad (312)$$

3) обратный пересчет поля в плоскость $z = 0$.

Пересчет поля между плоскостями $z = 0$ и $z = f$ сводится к расчету прямого и обратного преобразований Фурье от функций $W_0(\mathbf{u})=\tilde{A}_M(\mathbf{u};L_1,L_2)\sqrt{I_0(\mathbf{u})}$ и $W_1(\mathbf{x})$ соответственно. Используя символы $\Im\{\ \}$ и $\Im^{-1}\{\ \}$ для обозначения прямого и обратного преобразований Фурье, представим функцию $\Psi(\mathbf{u})$ в виде

$$\Psi(\mathbf{x})=\Im^{-1}\left\{\frac{\mathrm{d}\,\overline{\varepsilon}\left(\tilde{I}_M,I_D\right)}{\mathrm{d}\,\tilde{I}_M}\Im\left\{\sqrt{I_0(\mathbf{u})}\tilde{A}_M(\mathbf{u};L_1,L_2)\right\}\right\}. \quad (313)$$

Интересно отметить, что расчет функции $\Psi(\mathbf{u})$ совпадает с первыми тремя шагами алгоритма ГС с точностью до замены в плоскости $z = f$.

Согласно (307), приращение $\Delta\tilde{A}_M(\mathbf{u};L_1,L_2)$, вызванное малым изменением фазы $\Delta\varphi$, имеет вид

$$\Delta\tilde{A}_M(\mathbf{u};L_1,L_2)=\sum_{p=-L_1}^{L_2}c_{1+pM}\cdot\mathrm{i}(1+pM)\exp\left(\mathrm{i}(1+pM)\varphi(\mathbf{u})\right)\Delta\varphi(\mathbf{u}). \quad (314)$$

Подставляя (314) в (311), получим градиент функционала в виде

$$\varepsilon\,\mathbb{C}(\varphi) =$$

$$= 2\,\mathrm{Re}\left(\sqrt{I_0(\mathbf{u})}\Psi^*(\mathbf{u}) \sum_{p=-L_1}^{L_2} c_{1+pM} \cdot i(1+pM)\exp\left(i(1+pM)\varphi(\mathbf{u})\right) \right). \quad (315)$$

Уравнения (313) и (315) определяют градиент для непрерывного функционала (309), (308). Интересно отметить, что уравнение (315) при функции $\Psi(\mathbf{u})$, определенной в виде

$$\Psi(\mathbf{x}) = \Im^{-1}\left\{ \frac{\mathrm{d}\,\bar{\varepsilon}(I, I_D)}{\mathrm{d}\,I} \Im\left\{ \sqrt{I_0(\mathbf{u})} A_M(\mathbf{u}) \right\} \right\} \quad (316)$$

может быть рассмотрено как градиент разрывного функционала

$$\varepsilon(\varphi) = \bar{\varepsilon}\left(I(\mathbf{x}), I_D(\mathbf{x})\right), \quad (317)$$

представляющего отличие распределения интенсивности $I(\mathbf{x})$ в (300) для квантованного ДОЭ (301) от требуемого распределения интенсивности $I_D(\mathbf{x})$.

Действительно, функционал (317) является предельным случаем функционала (309) при $L_1 \to \infty$ и $L_2 \to \infty$. В этом предельном случае приращение функционала $\Delta\varepsilon(\varphi)$ также представимо в виде (311). При этом функция $\Psi(\mathbf{u})$ в (311) имеет вид (312), а приращение $\Delta\Phi$ соответствует приращению квантованной ФКП $A_M(\mathbf{u})$, вызванному малым изменением фазы $\Delta\varphi$;

$$\Delta A_M(\mathbf{u}) = \exp\left(\mathrm{i}\,\Phi_M\left[\varphi(\mathbf{u}) + \Delta\varphi(\mathbf{u})\right]\right) - \exp\left(\mathrm{i}\,\Phi_M\left[\varphi(\mathbf{u})\right]\right). \quad (318)$$

Используя в (301) непрерывную аппроксимацию $\Delta A_M(\mathbf{u}) \approx \Delta\tilde{A}_M(\mathbf{u}, L_1, L_2)$ (314) для приращения разрывной функции $A_M(\mathbf{u})$, получим выражение для градиента функционала в виде (315).

Следует отметить, что вычисление приращения $\Delta\tilde{A}_M\left(\mathbf{u};L_1,L_2\right)$ содержит операцию дифференцирования отрезка ряда Фурье

$$S\left(\xi;L_1,L_2\right)=\sum_{p=-L_1}^{L_2} c_{1+pM}\exp\left(\mathrm{i}\left(1+pM\right)\xi\right) \tag{319}$$

для функции $\exp\left(\mathrm{i}\Phi_M\left[\xi\right]\right)$. Формальное дифференцирование рядов Фурье может привести к значительным потерям в точности вычисления производной /113/. Для улучшения устойчивости градиентного метода, дифференцирование ряда (319) целесообразно проводить с применением методов регуляризации. Один из простых вычислительных приемов регуляризации, используемых для улучшения сходимости дифференцированного ряда Фурье, состоит в умножении коэффициентов формально дифференцированного ряда на весовые множители d_n, монотонно убывающие с ростом n. Методам выбора весовых множителей посвящено большое количество работ /113—115/. Для ряда (319) могут быть использованы следующие достаточно универсальные весовые множители /113/:

$$d_n=\operatorname{sinc}\left(\frac{\pi n}{m}\right),\qquad m=\max\left(L_1,L_2\right)+1\ , \tag{320}$$

При использование множителей регуляризации (320), приращение (314) принимает вид:

$$\begin{aligned}&\Delta\tilde{A}_M\left(\mathbf{u};L_1,L_2\right)=\\&=\sum_{p=-L_1}^{L_2} c_{1+pM}d_{1+pM}\cdot\mathrm{i}\left(1+pM\right)\exp\left(\mathrm{i}\left(1+pM\right)\varphi\left(\mathbf{u}\right)\right)\Delta\varphi\left(\mathbf{u}\right)\cdot\end{aligned} \tag{321}$$

При этом градиент функционалов (309) и (317) представляется в виде:

$$\varepsilon'(\varphi) = 2\,\mathrm{Re}\left(\sqrt{I_0(\mathbf{u})}\Psi^*(\mathbf{u}) \sum_{p=-L_1}^{L_2} c_{1+pM} \cdot d_{1+pM} \times \right.$$
$$\left. \times \mathrm{i}(1+pM)\exp\left(\mathrm{i}(1+pM)\varphi(\mathbf{u})\right) \right). \qquad (322)$$

Для иллюстрации влияния регуляризирующих множителей (320) на точность дифференцирования, рассмотрим дифференцирование отрезка ряда Фурье для функции

$$\exp\left(\mathrm{i}\,\Phi_2[\xi]\right) = -\mathrm{sgn}(\xi - \pi). \qquad (323)$$

Функция

$$\Phi_2[\xi] = \begin{cases} 0, & \xi \in [0, \pi), \\ \pi, & \xi \in [\pi, 2\pi), \end{cases}$$

в (323) описывает операцию бинаризации. Согласно (323), производная

$$\frac{\mathrm{d}\exp\left(\mathrm{i}\,\Phi_2[\xi]\right)}{\mathrm{d}\,\xi}$$

равна нулю в каждой точке интервала [0, 2π) за исключением точки $\xi = \pi$ где производной, понимаемой в классическом смысле, не существует. Отметим, что обобщенная производная периодически продолженной функции (323) соответствует знакочередующемуся ряду из δ-функций. Полагая в (303) и (304) $M = 2$, представим отрезок ряда Фурье для функции (323) при $L_1 = L_2$ в виде:

$$S(\xi; L_1, L_1) = \frac{2}{\pi} \sum_{p=1}^{L_1} \frac{1}{(2p-1)} \sin\left((2p-1)\xi\right). \qquad (324)$$

Формальное дифференцирование отрезка ряда Фурье (324) дает ряд

$$\frac{\mathrm{d}\,S(\xi; L_1, L_1)}{\mathrm{d}\,\xi} = \frac{2}{\pi} \sum_{p=1}^{L_1} \cos\left((2p-1)\xi\right), \qquad (325)$$

который при $L_1 \to \infty$ является расходящимся во всех точках интервала (0, 2π) /113/. Умножая ряд (325) на множители (320), получим регуляризованную производную виде:

$$\frac{\mathrm{d}S(\xi; L_1, L_2)}{\mathrm{d}\xi} = \frac{2}{\pi}\sum_{p=1}^{L_1} \frac{\sin(\pi(2p-1)/2L_1)}{\pi(2p-1)/2L_1}\sin\left((2p-1)\xi\right). \qquad (326)$$

Ряд (326), в отличие от ряда (325), при $L_1 \to \infty$ сходится к нулю в каждой фиксированной точке $\xi \neq \pi$ /113/. Интересно отметить, что ряд (326) является тригонометрическим рядом для δ-функции. Точность вычисления производной, для функции (323) с использованием ряда (325) и регуляризованного ряда (326) иллюстрируется графиком на рис.50. В силу симметрии относительно точки $\xi = \pi$, функции (325) и (326) на рис. 50 рассчитаны на отрезке [0, π]. Рисунок 50 ясно показывает, что для ряда (325) характерны существенно более сильные осцилляции в окрестности требуемого нулевого значения.

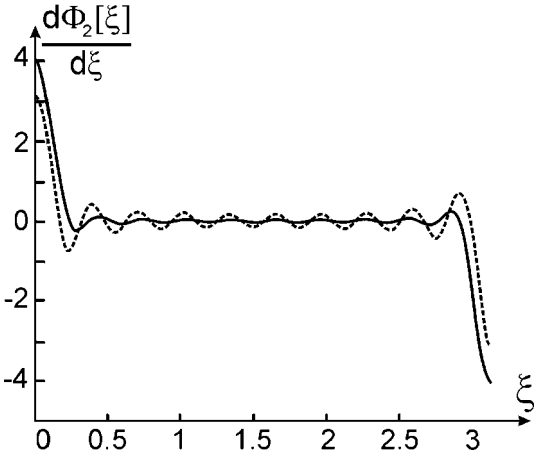


Рис. 50. Производные бинарной функции (365); пунктирная линия — ряд (325) при $L_1 = 10$, непрерывная линия — регуляризированный ряд (326) при $L_1 = 10$

Оптимальное значение шага выбирается в виде /5/:

$$t = -\varepsilon_1(0)/\varepsilon_1'(0). \tag{327}$$

Для вычисления производной $\varepsilon_1'(0)$ достаточно подставить в уравнения (301), (304) значение $\Delta\varphi = -(\Delta t)\cdot\varepsilon'(\varphi)$. В результате производная $\varepsilon_1'(0)$ в (327) примет вид:

$$\varepsilon_1'(0) = -4\,\mathrm{Re}\Big(I_0(\mathbf{u})(\Psi^*(\mathbf{u})\times \\ \times\sum_{p=-L_1}^{L_2}\gamma_{1+pM}\cdot \mathrm{i}(1+pM)\exp\big(\mathrm{i}(1+pM)\varphi(\mathbf{u})\big)\Big)^2\Big), \tag{328}$$

где

$$\gamma_{1+pM} = c_{1+pM} \tag{329}$$

при расчете без учета регуляризующих множитjелей (320) и

$$\gamma_{1+pM} = c_{1+pM}d_{1+pM} \tag{330}$$

при регуляризованном расчете приращения $\Delta\tilde{A}_M(\mathbf{u};L_1,L_2)$.

В заключение рассмотрим вопрос о выборе начального приближения для градиентной процедуры (310). Для расчета квантованных ДОЭ с числом уровней квантования $M > 2$ в качестве начального приближения для градиентной процедуры (310) можно использовать непрерывную функцию $\varphi_0(\mathbf{u}) \in [0, 2\pi)$, рассчитанную из условия формирования требуемого распределения интенсивности $I_D(\mathbf{x}), \mathbf{x} \in D$. Расчет непрерывной (неквантованной) функции $\varphi_0(\mathbf{u})$ не представляет сложности. Для расчета непрерывной функции $\varphi_0(\mathbf{u})$ могут быть использованы быстрые итерационные алгоритмы, рассмотренные в пунктах 2 — 8. Расчет $\varphi_0(\mathbf{u})$ может быть также проведен аналитически в приближении геометрической оптики /37, 66, 116/. Операция квантования (301), (302) приводит к снижению эффективности фокусировки и по-

явлению дополнительных дифракционных порядков — пучков с фазовыми функциями (306). Требуемая область фокусировки создается в первом порядке при $n = 1$. Паразитные порядки (306) при $n \neq 1$ накладываются на основное изображение и портят его.

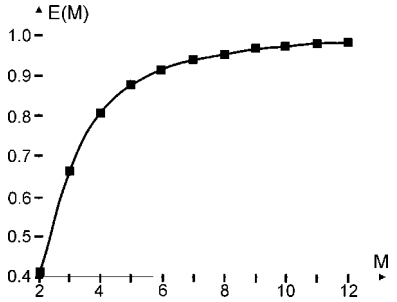


Рис. 51. Доля энергии освещающего пучка, фокусируемая в основном 1 порядке, при квантовании фазы ДОЭ по М уровням

Доля энергии освещающего пучка *E(M)*, фокусируемая в первом порядке, соответствует квадрату модуля коэффициента Фурье c_1 в (304):

$$E(M) = |c_1|^2 = \mathrm{sinc}^2(\pi/M) \quad . \tag{331}$$

График функции *E(M)* приведен на рис. 51. Согласно графику на рис. 51 эффекты квантования особенно существенны при малых $M = 2, 3, 4$, при $M > 4$ квантованный ДОЭ направляет уже более 85% энергии в требуемую область.

При расчете бинарных ДОЭ ($M = 2$) расчет начальной фазы $\varphi_0(\mathbf{u}) \in [0, 2\pi)$ следует модифицировать аналогично тому, как это было сделано в пункте 8.4 при расчете одномерных бинарных дифракционных решеток. Распределение интенсивности, формируемое бинарным ДОЭ, является центрально-симметричным; $I(\mathbf{x}) = I(-\mathbf{x})$. Поэтому область D и заданное распределение интенсивности $I_D(\mathbf{x})$ предполагаем центрально-симметричными. Представим область фокусировки D в виде объединения двух центрально-симметричных областей D_1 и

D_2. Начальную фазу $\varphi_0(\mathbf{u})$ при $M=2$ следует рассчитывать из условия фокусировки в половину области D, то есть в область D_1 или в область D_2. Действительно, пусть $\varphi_0(\mathbf{u})$ является фазовой функцией ДОЭ для формирования распределения интенсивности

$$I_{D_1}(\mathbf{x})=\begin{cases} I_D(\mathbf{x}), & x\in D_1, \\ 0, & x\notin D_1. \end{cases} \tag{332}$$

В этом случае интенсивности полей, формируемых бинарным ДОЭ с ФКП $\exp\left(\mathrm{i}\Phi_2\left[\varphi_0(\mathbf{u})\right]\right)$ в основных 1 и -1 порядках ($|c_1|^2=|c_{-1}|^2=0{,}405$), имеют вид:

$$I_1(\mathbf{x})=\left|\frac{c_1}{\lambda f}\int_G\sqrt{I_0(\mathbf{u})}\exp\left(\mathrm{i}\varphi_0(\mathbf{u})-\mathrm{i}\frac{k}{f}\mathbf{x}\mathbf{u}\right)\mathrm{d}^2\mathbf{u}\right|^2, \tag{333}$$

$$I_{-1}(\mathbf{x})=\left|\frac{c_{-1}}{\lambda f}\left(\int_G\sqrt{I_0(\mathbf{u})}\exp\left(\mathrm{i}\varphi_0(\mathbf{u})-\mathrm{i}\frac{k}{f}(-\mathbf{x})\mathbf{u}\right)\mathrm{d}^2\mathbf{u}\right)^*\right|^2. \tag{334}$$

Согласно (333), (334) распределения интенсивности являются центрально-симметричными:

$$I_1(\mathbf{x})=I_{-1}(-\mathbf{x}). \tag{335}$$

Из (335) следует, что если $I_1(\mathbf{x})$ является требуемым распределением интенсивности $I_D(\mathbf{x})$ при $\mathbf{x}\in D_1$, то в -1 порядке будет сформировано распределение интенсивности $I_{-1}(\mathbf{x})\equiv I_D(\mathbf{x})$ при $\mathbf{x}\in D_2$. Сложение интенсивностей полей, формируемых в 1 и -1 порядках, дает требуемое распределение интенсивности $I_D(\mathbf{x})$ при $\mathbf{x}\in D$. Поскольку $|c_1|^2+|c_{-1}|^2=0{,}81$, то при учете только главных ±1 порядков получим, что в требуемой области D фокусируется 81% энергии освещающего

пучка. Рассмотренный выбор фазовой функции $\varphi_0(\mathbf{u})$ дает приближенное решение $\Phi_2\left[\varphi_0(\mathbf{u})\right]$ задачи фокусировки при $M = 2$. Мы рассмотрели только два главных порядка без учета погрешности расчета функции $\varphi_0(\mathbf{u})$. Итерационные алгоритмы расчета $\varphi_0(\mathbf{u})$, как правило, позволяют сфокусировать в требуемой области 85—90% энергии освещающего пучка при среднеквадратичной ошибке (с.к.о.) формирования требуемого распределения интенсивности в 10—20%. В этом случае, даже при учете только 1 и –1 порядков, распределение интенсивности в области D будет отличаться от $I_D(\mathbf{x})$, не только вследствие погрешности расчета фазы $\varphi_0(\mathbf{u})$, но и вследствие интерференции между полями, формируемыми в 1 и –1 порядках. Взаимодействием порядков можно пренебречь лишь в случае фокусировки в специфические области, состоящие из двух разнесенных симметричных подобластей. Для практически интересных фокальных областей, таких как прямоугольник, круг, отрезок, бинарный ДОЭ, как правило, дает энергетическую эффективность 70—75% при большой с.к.о. формирования заданного распределения интенсивности в 45—55%.

Пример 14. Для иллюстрации работы разработанного метода был рассчитан бинарный ДОЭ для фокусировки пучка с гауссовым распределением интенсивности,

$$I_0(\mathbf{u}) = \exp\left(-\mathbf{u}^2/\sigma^2\right), \qquad |\mathbf{u}| \le \sigma, \tag{336}$$

в квадрат $2d \times 2d$ с постоянным распределением интенсивности при $2d = 13\Delta$, где $\Delta = \lambda f/\sigma$ —дифракционный размер. Значения энергетической эффективности E и ошибка δ вычислялись по формулам (47) и (248).

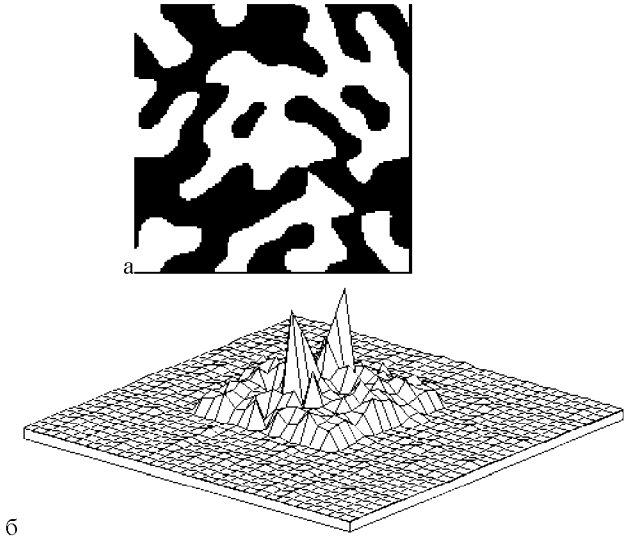


Рис. 52. Бинарная фаза ДОЭ для фокусировки в квадрат (приближенное решение) (а); распределение интенсивности при фокусировке в квадрат (б)

Как показано выше, приближенное решение для фазовой функции бинарного ДОЭ может быть получено прямой бинаризацией фазовой функции $\varphi_0(\mathbf{u})$, рассчитанной из условия фокусировки пучка (336) в "половину квадрата". На рис. 52а приведена бинарная фаза $\varphi_2(\mathbf{u}) = \Phi_2\left[\varphi_0(\mathbf{u})\right]$, полученная при $\varphi_0(\mathbf{u})$, рассчитанной по алгоритму ГС для фокусировки в половину квадрата.

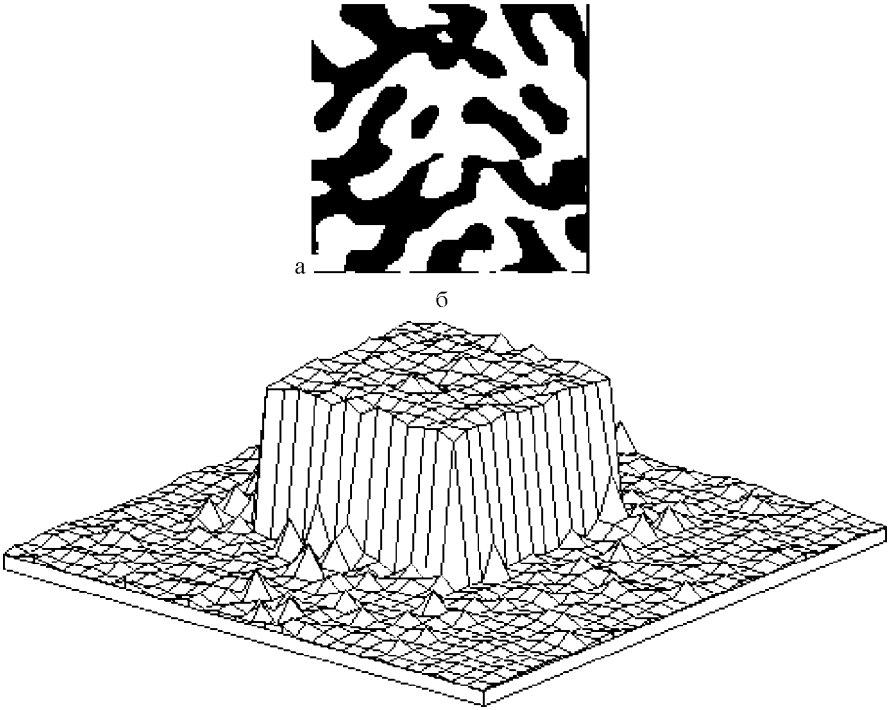


Рис. 53. Бинарная фаза ДОЭ для фокусировки в квадрат (а), рассчитанная итерационно при $L_1 = L_2 = 5$; Распределение интенсивности при фокусировке в квадрат (б)

Расчетное распределение интенсивности в плоскости $z = f$ для бинарного ДОЭ на рис. 52а приведено на рис. 52*б*. Рис. 52*б* показывает, что хотя энергия сфокусирована в пределах квадрата (E = 74,8%), распределение интенсивности является существенно неравномерным (δ = 54,5%). Для получения лучшего решения ФКП бинарного ДОЭ была аппроксимирована функцией $A_2(\mathbf{u};L_1,L_2)$ при L_1=L_2=5. Расчет $\varphi(\mathbf{u})$ в (307) проводился итерационно по формулам (310), (322) с использованием регуляризации (321). В качестве функционала ошибки (309) был использован следующий функционал квадрата ошибки:

$$\varepsilon(\varphi) = \int_{-\infty}^{\infty} \left(\sqrt{\tilde{I}_M(\mathbf{x})} - \sqrt{I_D(\mathbf{x})} \right)^2 \mathrm{d}^2\mathbf{x} . \qquad (337)$$

В качестве начального приближения для $\varphi(\mathbf{u})$ в (307) использовалась фаза $\varphi_0(\mathbf{u})$. Указанный выбор начального приближения позволяет в 2—3 раза сократить число итераций. На рис. 53,а приведена бинарная фаза $\varphi_2(\mathbf{u})$, полученная бинаризацией функции $\varphi(\mathbf{u})$ в (307), рассчитанной за 50 итераций. Расчетное распределение интенсивности, формируемое бинарным ДОЭ на рис. 53,а, приведено на рис. 53,б. Для ДОЭ с ФКП $A_2(\mathbf{u}; 5,5)$ и соответствующего ему бинарного ДОЭ значения E и δ отличались незначительно и составили 73,5%, 1,5% и 72,6%, 2,9% соответственно.

Пример 15. В качестве других примеров были рассчитаны ДОЭ для фокусировки пучка (336) в кольцо с радиусами $R_1 = 9\Delta$ и $R_2 = 18\Delta$ и в в букву "Н", состоящую из трех отрезков с длинами 12Δ, 12Δ и 7Δ. Рассчитанная бинарные фазы и распределения интенсивности, формируемые бинарными ДОЭ, приведены на рисунках 54 и 55. При фокусировке в кольцо (рис. 54*б*)) значения E и δ составили 73,2% и 3,2%, а при фокусировке в букву "Н" (рис. 55*б*) —- 70,2% и 2,1% соответственно.

В процессе расчетов получается, что аппроксимация ФКП бинарного ДОЭ функцией $A_2(\mathbf{u};L_1,L_2)$ является достаточно хорошей уже при $L_1=L_2>4,5$. Для приведенных примеров, включающих регуляризованный расчет приращения $\Delta\tilde{A}_2(\mathbf{u};L_1,L_2)$, обратный переход от функции $A_2(\mathbf{u};L_1,L_2)$ к бинарной функции при $L_1=L_2=4,5$ приводил к снижению энергетической эффективности E всего на 1—2% при незначительном увеличении δ на 1—2%. Отметим, что расчет приращения $\Delta\tilde{A}_2(\mathbf{u};L_1,L_2)$ без регуляризирующих множителей делает расчет менее устойчивым и приводит к большей (в 2—3 раза) ошибке в значениях E и δ при обратном переходе к бинарной функции.

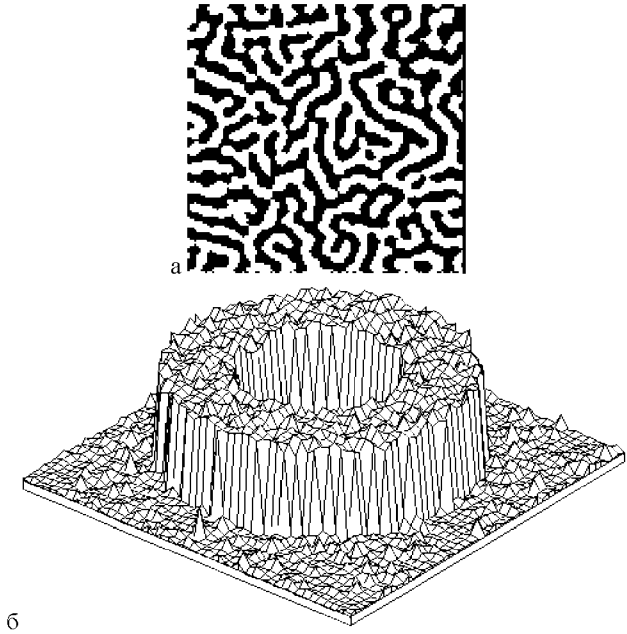


Рис. 54. Бинарная фаза ДОЭ для фокусировки в кольцо (а); Распределение интенсивности при фокусировке в кольцо (б)

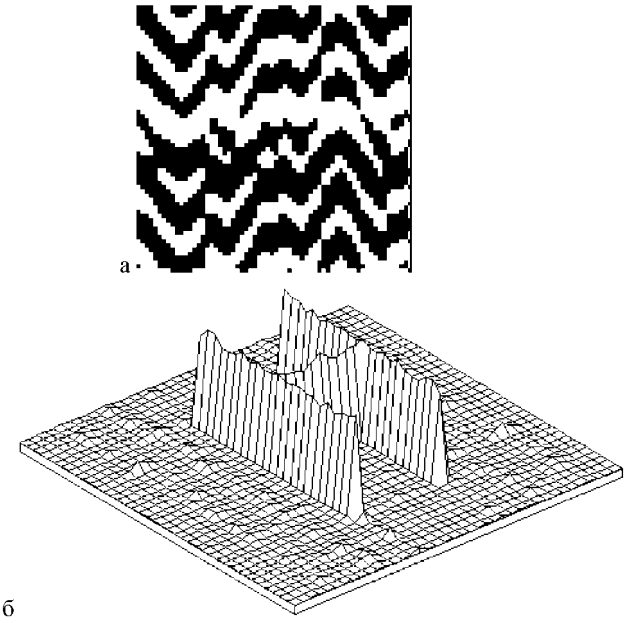


Рис. 55. Бинарная фаза ДОЭ для фокусировки в букву "Н" (а); Распределение интенсивности при фокусировке в букву "Н" (б)

Бинарные ДОЭ позволяют сформировать только центрально-симметричное распределение интенсивности $I_D(\mathbf{x})$. Для фокусировки в несимметричную область D число уровней квантования М должно быть больше 2. В качестве примера фокусировки в несимметричную область был рассчитан ДОЭ с четырьмя уровнями квантования ДОЭ для фокусировки в букву “F”, состоящую из трех отрезков с длинами 13Δ, 7Δ и 4Δ.

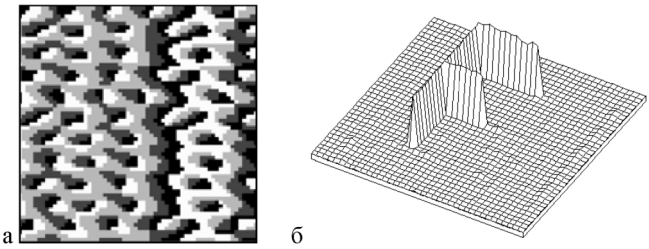


Рис. 56. 4-уровневая фаза ДОЭ для фокусировки в букву "F" (а); распределение интенсивности при фокусировке в букву "F" (б)

Рассчитанная фазовая функция и распределение интенсивности, формируемые четырехуровневым ДОЭ, приведены на рис. 56. При фокусировке в букву “F” (рис. 56,б)) значения E и δ составили 70,2% и 2,1% соответственно. Интересно отметить, что полученное значение энергетической эффективности E при $M = 4$ примерно такое же, как и для бинарных ДОЭ на рис.52—55. Это объясняется тем, что при расчете бинарных ДОЭ используются два основных порядка 1 и −1. При этом доля энергии $E = |c_1|^2 + |c_{-1}|^2 = 0,81$, фокусируемая в основных порядках 1 и −1 бинарного ДОЭ, равна доле энергии E (331) (см. рис.51), фокусируемой в основном первом порядке при $M = 4$.

Отметим, что расчет квантованного ДОЭ с квадратной апертурой G ($b \times b$) для фокусировки пучка с постоянной интенсивностью в прямоугольник $N\Delta \times K\Delta$, где $\Delta = \lambda f/b$ — размер дифракционного пятна, эквивалентен расчету фазовой дифракционной решетки (ФДР), формирующей $N \times K$ порядков. Действительно, для $\mathbf{x} = \mathbf{x}_{j,p} = (j\Delta, p\Delta)$ значения интенсивности в плоскости $z = f$

$$I_{j,p} = \left| \frac{1}{\lambda \boldsymbol{f}} \int_G \exp\left(\mathrm{i}\varphi_M(\mathbf{u})\right) \exp\left(-\mathrm{i}\frac{2\pi}{b}(ju + pv)\right) \mathrm{d}^2\mathbf{u} \right|^2, \qquad (338)$$

пропорциональны квадратам модулей коэффициентов Фурье в разложении функции $\exp\left(i\varphi_M(\mathbf{u})\right)$. Поэтому функция $\varphi_M(\mathbf{u})$ соответствует М-уровневой фазовой функции дифракционной решетки с периодом $b \times b$ и интенсивностью порядков:

$$I_{j,p} = \begin{cases} I_D(j\Delta, p\Delta), & |j| \le N/2, \quad |p| \le K/2, \\ 0, & \textit{иначе}. \end{cases} \qquad (339)$$

Пример 16. В таблице 6 представлены результаты расчета бинарных ФДР для различного числа порядков с равной интенсивностью. Энергетическая эффективность E=73—75% и среднеквадратичная ошибка δ=3—5%, полученные для ФДР с числом порядков до 51×51, подтверждают высокую работоспособность предложенного метода.

При итерационном расчете решеток табл. 6, в качестве начального приближения $\varphi_0(\mathbf{u})$ использовалась следующая фазовая функция

$$\varphi_0(\mathbf{u}) = \mathrm{mod}_{2\pi}\left(\frac{\pi N}{4}\left(\left(\frac{u}{b}\right)^2 + 2\left(\frac{v}{b}\right)^2 + 4\left(\frac{u}{b}\right)\right)\right), \quad (u,v) \in [-b/2, b/2] \tag{340}$$

Таблица 6.

Результаты расчета бинарных ФДР

Число порядков	Энергетическая эффективность E (%)	Среднеквадратичная ошибка δ (%)
5 × 5	74,4	3,1
7 × 7	73,7	3,1
9 × 9	72,6	2,9
11 × 11	73,8	3,6
15 × 15	74,7	3,4
21 × 21	73,8	3,7
33 × 33	75,1	4,8
51 × 51	74,2	4,9

Функция (340) является фазовой функцией фокусатора /71, 89/, рассчитанного в приближении лучевой оптики из условия фокусировки в прямоугольник

$$\mathbf{x} \in \left[0, \frac{N}{2}\Delta\right] \times \left[-\frac{N}{2}\Delta, \frac{N}{2}\Delta\right].$$

Фазовая функция (340) соответствует фазе периода дифракционной решетки с числом порядков $(N/2) \times N$ и обеспечивает фокусировку в половину исходной области из $N \times N$ порядков. Использование ана-

литического начального приближения (340) упрощает расчетную процедуру и дает более быструю сходимость итерационного процесса (310).

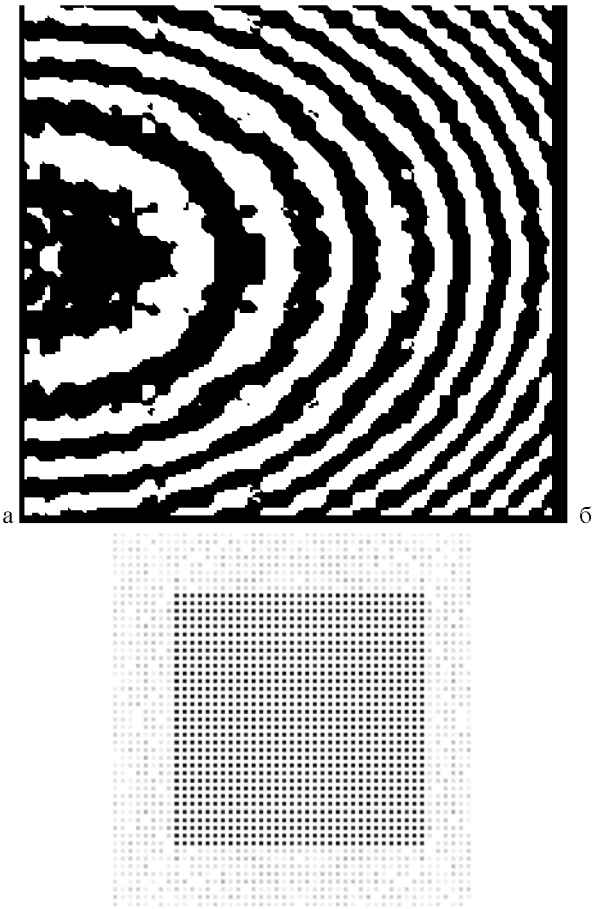


Рис. 57. Бинарная фаза ФДР с числом порядков 33 × 33 (а); распределение интенсивности в Фурье плоскости для решетки на рис. 68а (б)

На рис. 57а показана расчетная бинарная фаза дифракционной решетки, формирующей 33 × 33 порядков с равной интенсивностью. Квазирегулярный вид полученной бинарной фазы объясняется использованием в качестве начального приближения гладкой функции (340).

Распределение интенсивности в плоскости $z = f$ для ФДР на рис. 57*б* показано при числе периодов решетки 2 × 2.

7.11. Квантованные ДОЭ для формирования амплитудно-фазовых распределений

Применение рассмотренного в пункте 9.10 метода не ограничивается расчетом квантованных ДОЭ для фокусировки в заданную область. Метод может быть также применен к расчету бинарных и квантованных ДОЭ для формирования заданных амплитудно-фазовых распределений.

Пусть требуется рассчитать ДОЭ с квантованной по M уровням фазовой функцией $\varphi_{\mathrm{M}}(\mathbf{u})$ для формирования в области D при $z = f$ заданного распределения комплексной амплитуды $W_D(\mathbf{x})$. С учетом специфики задачи будем предполагать ДОЭ расположенным в передней плоскости Фурье-линзы (при $z = -f$). В этом случае комплексная амплитуда в выходной плоскости (при $z = f$) соответствует преобразованию Фурье входного поля:

$$W(\mathbf{x}) = \frac{1}{\lambda f}\int_G \sqrt{I_0(\mathbf{u})}\exp\left(\mathrm{i}\varphi_M(\mathbf{u}) - \mathrm{i}\frac{k}{f}\mathbf{x}\mathbf{u}\right)\mathrm{d}^2\mathbf{u}. \qquad (341)$$

Расчет квантованных ДОЭ для формирования полей с заданным распределением комплексной амплитуды также основывается на оптимизации непрерывной функции $\tilde{A}_M(\mathbf{u}; L_1, L_2)$ (307). Расчет функции $\varphi(\mathbf{u})$ в (349) проводится градиентным методом и состоит в итерационной коррекции функции $\varphi(\mathbf{u})$ по правилу (310). Функционал $\varepsilon(\varphi)$ в этом случае соответствует некоторому интегральному функционалу

$$\varepsilon(\varphi) = \bar{\varepsilon}\left(\tilde{W}_M(\mathbf{x}), W_D(\mathbf{x})\right),$$

характеризующему отличие распределения комплексной амплитуды, формируемого ДОЭ с непрерывной ФКП (307)

$$\tilde{W}_M(\mathbf{x}) = \frac{1}{\lambda f}\int_G \sqrt{I_0(\mathbf{u})}\,\tilde{A}_M(\mathbf{u};L_1,L_2)\exp\left(-\mathrm{i}\frac{k}{f}\mathbf{x}\mathbf{u}\right)\mathrm{d}^2\mathbf{u}, \tag{342}$$

от требуемого амплитудно-фазового распределения $W_D(\mathbf{x})$, $\mathbf{x}\in D$;

Градиент $\varepsilon'(\varphi)$ функционала также представляется формулами (315), (322) при функции $\Psi(\mathbf{u})$ определенной в виде:

$$\Psi(\mathbf{x}) = F^{-1}\left\{\frac{\mathrm{d}\,\overline{\varepsilon}\left(\tilde{W}_M, W_D\right)}{\mathrm{d}\,\tilde{W}_M}\right\}. \tag{343}$$

Обратный переход от непрерывной функции $\tilde{A}_M(\mathbf{u};L_1,L_2)$ к квантованной функции (301) дает M-уровневую фазовую функцию.

В заключении отметим, что метод может быть также обобщен на случай расчета ДОЭ с ФКП следующего общего вида:

$$A(\mathbf{u}) = \Phi\left[\varphi(\mathbf{u})\right], \tag{344}$$

где $\varphi(\mathbf{u})$ - непрерывная функция определенная в интервале [0, 2π), а $\Phi[\xi], \xi\in[0, 2\pi)$ - некоторая разрывная комплексная функция, определяющая поточечное преобразование функции $\varphi(\mathbf{u})$. ФКП квантованного ДОЭ (301), (302) является частным случаем представления (344) при функции $\Phi[\xi]$ определенной в виде

$$\Phi[\xi] = \exp\left(\mathrm{i}\,\Phi_M[\xi]\right) = \exp\left(\mathrm{i}\frac{2\pi}{M}\cdot\mathrm{int}\left(\frac{\xi M}{2\pi}\right)\right). \tag{345}$$

ФКП (344) также может быть представлена в виде ряда по дифракционным порядкам (305). В этом случае коэффициенты c_n будут коэффициентами Фурье функции $\Phi[\xi]$. Использование отрезка ряда по дифракционным порядкам для ФКП (344) также позволяет заменить задачу оптимизацию разрывной ФКП (344) задачей оптимизации не-

прерывного отрезка ряда (305). Конкретный вид функции $\Phi[\xi]$ в (344) зависит от расчетной задачи. Например, функция $\Phi[\xi]=(1/2\pi)\Phi_M[\xi]$ может быть использована для расчета квантованных амплитудных фильтров.

Пример 17. Работоспособность метода расчета квантованных ДОЭ для формирования полей с заданным распределением комплексной амплитуды проиллюстрируем на задаче расчета моданов — фазовых ДОЭ для формирования полей с комплексной амплитудой, соответствующей модовым функциям оператора распространения света /10, 117/. Метод был применен к расчету бинарных моданов для преобразования гауссового пучка (336) в унимодальные распределения Гаусса-Эрмита

$$M_{(p',l)}(\mathbf{x})=H_p\left(\sqrt{2}x/\sigma\right)H_l\left(\sqrt{2}y/\sigma\right)\exp\left(-\mathbf{x}^2/\sigma^2\right), \qquad (346)$$

где (p, l) — индекс моды, а $H_p(x)$ — полином Эрмита степени p. Расчет проводился для модовых распределений (1, 0) и (1, 1) и был основан на градиентной процедуре (310), (315), (321), (322), (343) при $L_1=L_2=5$ и следующем функционале ошибки:

$$\varepsilon(\varphi)=\int_{-\infty}^{\infty}\left|\tilde{W}_M(\mathbf{x})-W_D(\mathbf{x})\right|^2 \mathrm{d}^2\mathbf{x}. \qquad (347)$$

Расчет производился для следующих физических параметров: длина волны λ = 0,63 мкм, параметр гауссова пучка σ= 1,41 мм, фундаментальные радиусы мод (1, 0) и (1, 1) $\delta_{(0,\,1)}=\delta_{(1,\,1)}$= 0,123 мм, фокусное расстояние Фурье-линзы f = 500 мм, размер области D, в которой формируются модовые распределения, 0,2 мм × 0,2 мм.

Рассчитанные бинарные фазы ДОЭ для формирования мод Гаусса-Эрмита (0, 1) и (1, 1) приведены на рис. 58а и 58*б* соответственно.

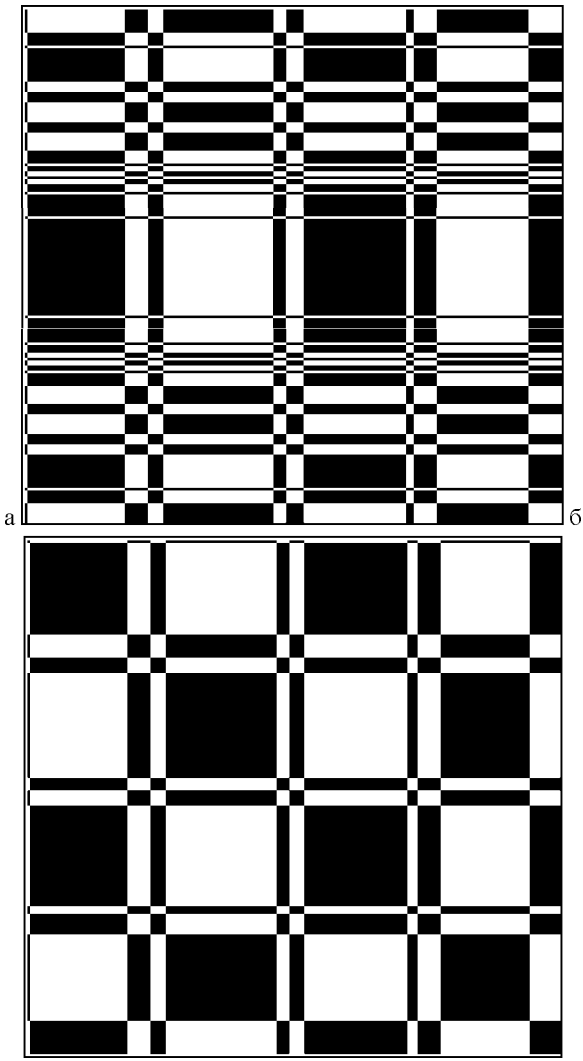
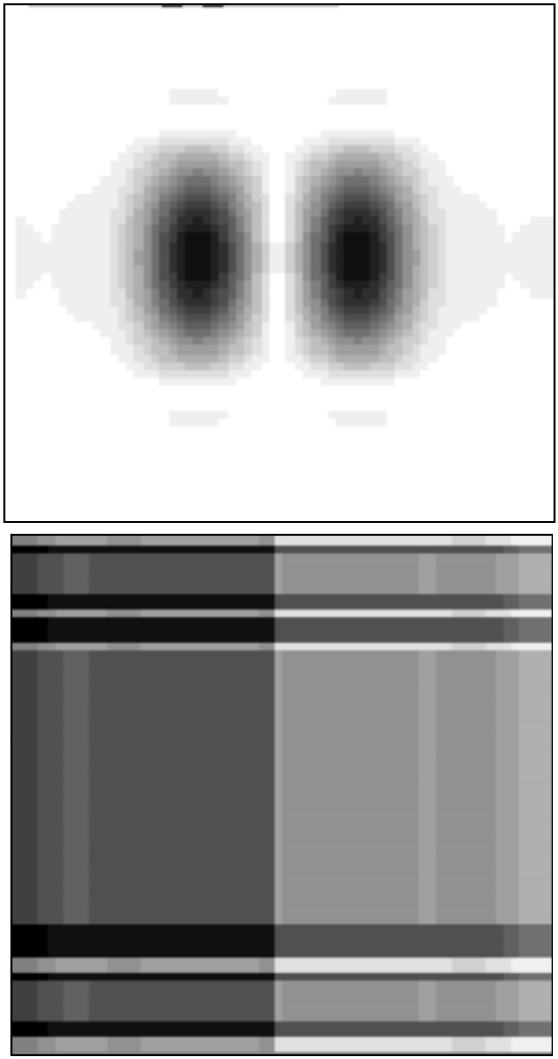


Рис. 58. Бинарная фазы моданов для преобразования гауссова пучка в модовые пучки: Гаусса-Эрмита (0, 1) (а), Гаусса-Эрмита (1, 1) (б)

Распределения амплитуды и фазы полей, генерируемых ДОЭ в выходной плоскости Фурье-линзы, приведены на рис. 59. Рисунок 59 ясно показывает структуру полей, характерную для модовых распределений (0, 1) (рис. 59 а, *б*) и (1, 1) (рис. 59 в, г).



а б

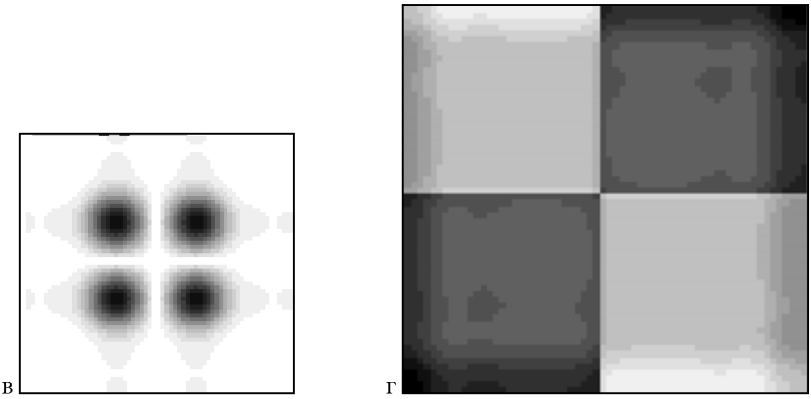


Рис. 59. Распределения амплитуды и фазы в Фурье-плоскости: для бинарного модана на рис. 58,а (а), (б); для бинарного модана на рис. 58б (в), (г)

Для характеристики работоспособности ДОЭ были рассчитаны значения энергетической эффективности

$$\mathrm{E}_m = \int\limits_D \left|W_D\left(\mathbf{x}\right)\right|^2 \mathrm{d}^2\,\mathbf{x} \Big/ \int\limits_G I_0\left(\mathbf{u}\right)\mathrm{d}^2\,\mathbf{u} \qquad (348)$$

и модового коэффициента

$$\gamma = \left|\int\limits_D W\left(\mathbf{x}\right)W_D^*\left(\mathbf{x}\right)\mathrm{d}^2\,\mathbf{x}\right|^2 \Big/ \int\limits_D \left|W\left(\mathbf{x}\right)\right|^2 \mathrm{d}^2\,\mathbf{x}\ , \qquad (349)$$

характеризующего содержание моды в полезной области D /117/. Рассчитанные значения $\mathrm{E_m}$ и δ составили 47,7%, 0,958 для моды (0, 1) и 43,9%, 0,947 для моды (1, 1).

Следует отметить, что обобщенный метод кодирования Кирка — Джонса, традиционно используемый для расчета моданов, не позволяет рассчитывать бинарные моданы. Для сравнения были рассчитаны моданы с непрерывной фазой для вышеуказанных параметров по обобщенному методу Кирка — Джонса /117/. Для рассчитанных по методу Кирка — Джонса моданов энергетическая эффективность более чем в 1,5 раза меньше, чем для бинарных моданов на рис. 59, что говорит о высокой работоспособности разработанного метода.

Заключение

Итеративные алгоритмы (ИА) решения обратных задач скалярной теории дифракции для расчета ДОЭ, рассмотренные в этом пособии, не охватывают весь спектр известных итеративных алгоритмов. Здесь не были рассмотрены такие популярные в прошлом алгоритмы, как метод прямого стохастического поиска бинарной фазы /52/, метод моделирования отжига /54/, метод диффузии ошибки /105/, метод Ньютона для расчета решеток Дамманна /61/ и другие. Эти методы менее эффективны по сравнению с параметрическими и градиентными ИА, рассмотренными в данном пособии.

Хотя ИА расчета ДОЭ имеют свои недостатки, например, стагнация ошибки, которая заключается в том, что из-за неудачного выбора начальной оценки искомой фазы ДОЭ, алгоритм может попасть в локальный минимум, когда с увеличением числа итераций ошибка не изменяется, несомненные достоинства ИА объясняют их широкое применение в дифракционной оптике. Укажем на некоторые из достоинств ИА.

Универсальность. С помощью ИА можно рассчитывать любые типы ДОЭ: фокусаторы, киноформы, аксиконы, дифракционные решетки, многофокусные линзы, компенсаторы, формирователи мод и бездифракционных пучков, пространственные фильтры.

Модульность. Различные типы ИА можно комбинировать и "стыковать" между собой так, чтобы выходные данные одного алгоритма поступали на вход другого.

Высокая эффективность. По отношению к неитеративным методам кодирования, применяемым для расчета цифровых голограмм, когда с помощью введения несущей пространственной частоты (явной, как в методе Кирка — Джонса, или неявной, как в алгоритме Ломана /74/) рассчитывается голограмма, формирующая в первом порядке дифракции требуемое изображение с эффективностью не более 10 — 30%, ИА позволяют рассчитывать ДОЭ, работающие в нулевом порядке с эффективностью 70—90%.

Физическая реализуемость. ИА типа «пинг-понг» или «маятника», основанные на последовательном применении пары прямого и обратного интегральных преобразований Фурье, Френеля, Ханкеля могут быть реализованы в оптической схеме с обратной связью /98/ или в резонаторе, образованном парой зеркал /118/.

Квантование фазы. ИА позволяют удобно выполнять такую нелинейную операцию над фазовой функцией ДОЭ как квантование (кусочно-постоянная аппроксимация фазы). При этом квантовать фазу можно прямым «жестким» способом на каждом шаге итеративного процесса или «мягким» способом, аппроксимируя ее конечной суммой ряда Гудмена — Сильвестри /112/.

Список специальных терминов

- Дифракционный оптический элемент (ДОЭ) 4, 10, 18, 30
- Скалярная теория дифракции 5, 8, 49, 88, 135
- Интегральное преобразование Френеля 22, 27, 47, 85, 86
- Интегральное преобразование Фурье 15, 38, 42, 49, 54, 78, 108
- Интегральное преобразование Ханкеля 48, 51, 52
- Некорректные обратные задачи в оптике 28
- Итеративный метод решения интегрального уравнения 9
- Метод (алгоритм) сокращения ошибки 5, 38
- Метод (алгоритм) входа-выхода12, 13, 14, 15, 18, 37
- Метод регуляризации А.Н. Тихонова 34
- Градиентный алгоритм расчета фазы ДОЭ 9, 15, 41, 58, 129, 146
- Пространственный Фурье-спектр 12, 35, 42, 47, 105
- Конический аксикон 113, 114
- ДОЭ с квантованной фазой 58, 62
- Бинарные оптические элементы 63, 132, 149
- Дифракционные решетки 46, 63, 75, 158
- Многофокусные линзы 158
- Дифракционные порядки 67
- Дифракционная эффективность 24, 57
- Гауссовый пучок 22, 103, 147
- Комплексная амплитуда светового поля 5, 8
- Интенсивность светового поля 5, 13
- Фаза светового поля 13, 16, 89
- Тонкая линза 13, 26
- Функция пропускания тонкого транспаранта 9,26
- Волновое число света 106, 115

Список контрольных вопросов

1. Какие световые поля описываются с помощью преобразования Френеля?
2. Что формируется в задней фокальной плоскости сферической линзы?
3. Какое основное свойство лазерного излучения?
4. Корректной или некорректной обратной задачей является расчет фазы ДОЭ?
5. Какие итеративные алгоритмы применяются для расчета фазы ДОЭ?
6. Как связана интенсивность света с функцией комплексной амплитуды?
7. Что понимают под эффектом стагнации итеративных алгоритмов?
8. На какие основные типы подразделяются ДОЭ?
9. Какими двумя параметрами контролируется процесс сходимости итеративного алгоритма?
10. Какой оптический элемент формирует в дальней зоне световое кольцо?
11. Какой оптический элемент формирует несколько дифракционных порядков?
12. Почему требуется рассчитывать ДОЭ с квантованной фазой, например бинарные ДОЭ?

Литература

1. Василенко Г И, Тараторин А М *Восстановление изображений* (М:Радио и связь, 1986)
2. Гудмен Дж. *Введение в Фурье-оптику* (М.:Мир, 1970)
3. Gerchberg R W, Saxton W O *Optik* **35**(2) 237-246 (1972)
4. Lin B, Gallagher N C *Appl. Opt* **13**(11) 2470-2471 (1974)
5. Fienup J R *Appl. Opt* **21**(15) 2758-2769 (1982)
6. Сивоконь В П *Формирование световых пучков с заданной структурой для задач лазерной технологии. Кандидатская диссертация* (М.:МГУ, 1986)
7. Fienup J R *Opt. Lett.* **3**(1) 27-29 (1978)
8. Fienup J R *Opt. Eng.* **19**(3) 297-305 (1980)
9. Lu G, Zhang Zh, Yu F T S, Tanone A *Opt. Eng.* **33**(2) 548-555 (1994)
10. Kotlyar V V, Nikolski I V, Soifer V A *Optik* **88**(1) 17-19 (1991)
11. Kotlyar V V, Nikolski I V *Opt. & Las. in Engin.* **15**(5) 323-330 (1991)
12. Kotlyar V V, Seraphimovich P G, Soifer V A *Proceedings of SPIE "Image Processing and Computer Optics"* **2363** 175-183 (1994)
13. Fienup J R *Appl. Opt.* **32**(10) 1737-1746 (1993)
14. Котляр В В, Серафимович П Г, Сойфер В А *Оптика и спектроскопия* **78**(1) 148-151 (1995)
15. Vasilyev E D, Kotlyar V V, Nikolsky I V, Soifer V A *Proceedings of SPIE* **1983** 1012-1013 (1993)
16. Борн М, Вольф Э *Основы оптики* (М.:Наука, 1973)
17. Котляр В В, Серафимович П Г *Оптика и спектроскопия* 77(4) 678-681 (1994)
18. Тихонов А Н, Арсенин В Я *Методы решения некорректных задач* (М.:Наука, 1979)
19. Gerchberg R W *Opt. Acta* **21** 709-720 (1974)
20. Papoulis A *IEEE Trans. Circuits Syst.* **22** 735-742 (1975)
21. Wyrowski F *J. Opt. Soc. Am.* 7(6) 961-963 (1990)
22. Kotlyar V V, Seraphimovich P G, Soifer V A *Optics and Laser in Engineering* **29**(4-5) 261-268 (1998)
23. Youla D C *IEEE Trans. Circuits. Syst.* **25** 694-702 (1979)

24. Aagedal H,Schmid M,Beth T, Teiwes S, Wyrowski F *J. Mod. Opt.* **43**(7) 1409-1421 (1996)
25. Abbis J B, Defrise M, DeMol C, Dhadwal H S *J. Opt. Soc. Am.* **73** 1470-1475 (1983)
26. Мину М *Математическое программирование* (М.:Наука, 1990)
27. Siegman A E *Opt. Lett.* **1**(1) 13-15 (1977)
28. Magni V, Gerullo G, Desilvestry S *J. Opt. Soc. Am. A* **9**(11) 2031-2033 (1992)
29. Khonina S N, Kotlyar V V, Soifer V *Optik* **88**(4) 182-184 (1991)
30. Doskolovich L L, Khonina S N, Kotlyar V V, Soifer V *Opt. & Quant. Electr.* **25** 801-814 (1993)
31. Березный А Е, Прохоров А М, Сисакян И Н, Сойфер В А *Доклады АН СССР* **274**(4) 802-805 (1984)
32. Khonina S N, Kotlyar V V, Shinkaryev M V, Soifer V A, Uspleniev G V *Opt. Commun.* **91**(3-4) 158-162 (1992)
33. Котляр В В, Сойфер В А *Письма в ЖТФ* **16**(12) 30-33 (1990)
34. Khonina S N, Kotlyar V V, Shinkaryev M V, Soifer V A, Uspleniev G V *J. Mod. Opt.* **39**(5) 1147-1154 (1992)
35. McLeod J H *J. Opt. Soc. Am.* **44**(8) 592-597 (1954)
36. Belanger P, Rioux M *Appl. Opt.* **17**(7) 1080-1086 (1978)
37. Голуб М А, Карпеев С В, Прохоров А М, Сисакян И Н, Сойфер В А *Письма в ЖТФ* 7(10) 618-623 (1981)
38. Tremblay R, D'Astons Y, Roy G, Blanshard M *Opt. Commun.* **28**(2) 193 (1979)
39. Brenden B B, Russel J T *Appl. Opt.* **23**(19) 3250 (1984)
40. Michaltsova I A, Nalivaiko V I, Soldatenkov I S *Optik* **67**(3) 267 (1984)
41. Khonina S N, Kotlyar V V, Soifer V A *J. Mod. Opt.* **40**(5) 761-769 (1993)
42. Khonina S N, Kotlyar V V, Soifer V A *Proceedings of SPIE* **1780** 263-272 (1993)
43. Khonina S N, Kotlyar V V, Soifer V A *Computer Optics* **13** 12-15 (1993)
44. Kotlyar V V, Khonina S N, Soifer V A *Opt. and Lasers in Technology* **27**(4) 235-240 (1995)
45. Khonina S N, Kotlyar V V, Philippov S V, Nikolsky I V, SoiferV A *Opt. Design Conf.* **22** 251-256 (1994)
46. Damman H, Klotz E *Opt. Acta* **24** 505-575 (1977)

47. Mait J N, Brenner K *Appl. Opt.* **27** 1692-1700 (1988)
48. Veldkamp W B, Leger J R, Swanson G J *Opt. Lett.* **11** 303-305 (1986)
49. Simpson M J *Appl. Opt.* **26**(9) 1786-1791 (1987)
50. Killat U, Rabe G, Rave W *Fiber Integ. Opt.* **4** 159-167 (1982)
51. Arsenault H H, Szoplik T, Macukow B *Optical processing and computing* (San Diego.:Academic Press, 1989)
52. Досколович Л Л,Сойфер В А, Шинкарев М С *Автометрия* **3** 104-108 (1990)
53. Seldowitz M A, Allebach J P, Sweeney D W *Appl. Opt.* **26** 2788-2798 (1987)
54. Feldman M R, Gest C C *Opt. Lett.* **14**(8) 479-481 (1989)
55. Turunen J, Vasara A, Westerholm J *Opt. Eng.* **28**(11) 1162-1167 (1989)
56. Morrison R L, Walker S L, Cloonan T *J Appl. Opt.* **32** 2512-2518 (1993)
57. Turunen J, Vasara A, Westerholm J *Opt. Commun.* **74** 245-252 (1990)
58. Mait J N *J. Opt. Soc. Am. A* 7(8) 1514-1528 (1990)
59. Березный А Е, Комаров С В, Прохоров А М, Сисакян И Н,Сойфер В А *Доклады АН СССР* **287**(3) 623-627 (1986)
60. Досколович Л Л, Котляр В В, Сойфер В А *Письма в ЖТФ* **17**(21) 54-57 (1991)
61. Dammann H, Gortler K *Opt. Commun.* **3**(5) 312-315 (1971)
62. Jahns J, Downs M M, Prise M E, Streibl N, Walker S. J *Opt. Eng.* **28** 1267-1275 (1988)
63. Бобров С Т, Туркевич Ю Г *Компьютерная оптика* **4** 38-45 (1989)
64. Bobrov S, Kotletsov B, Turkevich Y *Proc. SPIE* **1751** 154-165 (1992)
65. Pommet A D, Moharam M G, Grann E B *J. Opt. Soc. Am. A* **11**(6) 1827-1834 (1994)
66. Golub M A, Sisakyan I N, Soifer V A *Opt. & Las. in Engin.* **15**(5) 297-309 (1991)
67. Soifer V, Golub M *Proceedings of SPIE* **1751** 140-154 (1992)
68. Doskolovich L L, Soifer V A, Alessandretti G, Perlo P, Repetto M *Pure & Appl. Opt.* **3** 921-930 (1994)
69. Doskolovich L L, Repetto M, Perlo P, Lavagno L, Sinesi S *Proc. ISHM "The many aspects of microelectronics"* (1994)
70. Карманов В Г *Математическое программирование* (М.:Наука, 1986)

71. Голуб МА, Досколович Л Л, Казанский Н Л, Сисакян И Н, Сойфер В А, ХаритоновС И *Компьютерная оптика* 10-11 100-110 (1992)
72. Leseberg D *Appl. Opt.* **31**(2) 223-229 (1992)
73. Воронцов М А, Корябин А В, Шмальгаузен В И *Управляемые оптические системы* (М.:Наука, 1988)
74. Bryngdahl O, Wyrowski F *Progress in Optics* **28** 3-86 (1990)
75. Досколович Л Л,Казанский Н Л,Сисакян И Н, Сойфер В А, Харитонов С И *Компьютерная оптика* **12** 8-14 (1992)
76. Kotlyar V V, Khonina S N, Soifer V A *Journal of Modern optics* **43**(7) 1509-1524 (1996)
77. Kotlyar V V, Nikolsky I V, Soifer V A *Optik* **98**(1) 26-30 (1994)
78. Kazanskiy N L, Kotlyar V V, Soifer V A *Opt. Eng.* **33**(10) 3156-3166 (1994)
79. Khonina S N, Kotlyar V V, Soifer V A *J. Mod. Opt.* **42**(6) 1231-1239 (1995)
80. Golub M A, Doskolovich L L, Kazansky N L, Kharitonov S I, Soifer V A *J.Mod.Opt* **39**(6) 1245-1251 (1992)
81. Khonina S N, Kotlyar V V, Soifer V A *Proceedings of SPIE* **1780** 263-272 1993
82. Прудников А П, Брычков Ю А, Маричев О И *Интегралы и ряды. Специальные функции* (М.:Наука, 1983)
83. Veldcamp W P *Appl. Opt.* **21**(17) 3209 (1981)
84. Han C Y, Ishii Y, Murata K *Appl.Opt.* **22**(22) 3644-3647 (1983)
85. Коронкевич В П, Ленкова Г А, Михальцова И А, Пальчикова И Г, Полещук А Г, Седухин А Г, Чурин Е Г, Юрлов Ю И *Автометрия* **1** 4-25 (1985)
86. Воронцов М А, Матвеев А И, Сивоконь В П *Компьютерная оптика* **1** 74-78 (1987)
87. Roberts N C *Appl.Opt.* **28** 31-32 (1989)
88. Eismann M T, Tai A M, Cederquist J N *Appl.Opt.* **28**(15) 2641-2650 (1989)
89. Голуб М А, Досколович Л Л, Казанский Н Л, Сисакян И Н, Сойфер В А, Харитонов С И *Компьютерная оптика* **7** 42-49 (1990)
90. Aleksoff C C, Ellis K K, Neagle S D *Opt.Eng.* **30**(5) 537-543 (1991)
91. Golub M A, Sisakyan I N, Soifer V A *Opt. and Lasers in Engin.* **15**(5) 297-309 (1991)

92. Golub M A, Doskolovich L L, Kotlyar V V, Nikolsky I V, Soifer V A *Компьютерная оптика* **13** 30-33 (1993)
93. Cordingley J *Appl. Opt.* **32**(14) 2538-2542 (1994)
94. Седухин А Г, Чурин Е Г *Автометрия* **6** 75-81 (1995)
95. Duparre M, Golub M A, Ludge B, Pavelyev V S, Soifer V A, Uspleniev G V, Volotovskii S G *Appl.Opt.* **34**(14) 2488-2497 (1995)
96. Golub M A, Duparre M, Kley E, Kowarschik R, Ludge B, Rockstroh W, H.Fuchs *Opt.Eng.* **35**(5) 1400-1406 (1996)
97. Farn M W **22** 246-250 (1994)
98. Воронцов М А, Шмальгаузен В И *Принципы адаптивной оптики* (М.:Наука, 1985)
99. Fedotovsky A, Lehovec H *Appl. Opt.* **13**(12) 2919-2923 (1974)
100. Пальчикова И Г *Компьютерная оптика* **6** 9-19 (1989)
101. Doskolovich L L, Kasanskiy N L, Kharitonov S I,Uspleniev G V *Opt. and Lasers in Engineering* **15**(5) 311-322 (1991)
102. Khonina S N, Kotlyar V V, Lushpin V V, Soifer V A *Optical Memory and Neural Networks* **6**(3) 213-220 (1997)
103. Soifer V A, Kotlyar V V, Doskolovich L L. *Iterative methods for diffractive optical elements computation* (Taylor & Francis, 1997)
104. Wyrowski F. *Appl.Opt.* **28** 3864-3869 (1989)
105. Weissbach S, Wirowski F, Bringdahl O *Opt. Comm.* **72**(1,2) 37-41 (1982)
106. Wyrowski F, Bryngdahl O *In Rep.Prog.Phys.* **28** 1481-1571 (1991)
107. Doskolovich L L, Perlo P, Petrova O I,Repetto P, Soifer V A *J. Mod.Opt.* **44**(4) 685-695 (1997)
108. Doskolovich L L,Perlo P, Petrova O I., Repetto P, Soifer V A *Int. Jour. of Optoelectronics* **10**(4) 243-249 (1995)
109. Досколович Л Л, Казанский Н Л, Сойфер В А *Автометрия* **5** 42-50 (1995)
110. Repetto P, Doskolovich L, Bigliati C, Pierro P *Workshop on Diffractive Optics and Microoptics* **7** (1998)
111. Dallas J M *Appl. Opt.* **14** 674-676 (1971)
112. Goodman J M, Silvestri A M *IBM Journ. Res. Develop.* **14** 478-484 (1969)
113. Ланцош К *Практические методы прикладного анализа.* (М.:Гос. изд-во физ.-мат. Литературы, 1961)

114. Оппенгейм А В, Шафер Р В *Цифровая обработка сигналов* (М.:Связь, 1979)
115. Старк Г *Теория и измерение оптических фурье-спектров. В книге: Применение методов фурье-оптики* (М.: Радио и Связь, 1988)
116. Sisakyan I N, Soifer V A Infrared Physics 32 435-438 (1991)
117. Soifer V A, Golub M A Laser beam mode selection by computer gener*ated holograms* (CRC Press, Boca Raton, 1994)
118. Noel E, Khan R R, Dhadwal H S *Opto. Eng.* **32**(11) 2866-2871 (1993)

Учебное издание

Котляр Виктор Викторович,
Досколович Леонид Леонидович,
Сойфер Виктор Александрович

ИТЕРАТИВНЫЕ МЕТОДЫ ПРОЕКТИРОВАНИЯ ПЛОСКОЙ ОПТИКИ

Учебное пособие

Технический редактор *В. С. Павельев*
Редакторская обработка *Н. С. Куприянова, В. С. Телепова*
Корректорская обработка *А. А. Нечитайло*
Компьютерная верстка *Ю.В. Бондаренко*
Доверстка *О. Ю. Дьяченко*

Подписано в печать 15.12.06. Формат 60x84 1/16.
Бумага офсетная. Печать офсетная.
Усл. печ. л. 9,77. Усл. кр.-отт. 9,89. Печ.л. 10,5 .
Тираж 50 экз. Заказ ИП-64/2006

Самарский государственный аэрокосмический университет.
443086, Самара, Московское шоссе, 34.

Изд-во Самарского государственного аэрокосмического университета.
443086, Самара, Московское шоссе, 34.