

МИНОБРНАУКИ РОССИИ

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ
БЮДЖЕТНОЕ ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ
ВЫСШЕГО ПРОФЕССИОНАЛЬНОГО ОБРАЗОВАНИЯ
«САМАРСКИЙ ГОСУДАРСТВЕННЫЙ АЭРОКОСМИЧЕСКИЙ
УНИВЕРСИТЕТ ИМЕНИ АКАДЕМИКА С.П. КОРОЛЕВА
(НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ УНИВЕРСИТЕТ)»

С.Н. Хонина

**Оптическая информатика.
Лабораторный практикум.**

Электронное учебное пособие

САМАРА
2011

Автор: **Хонина Светлана Николаевна**

Хонина, С. Н. Оптическая информатика. Лабораторный практикум. [Электронный ресурс] : электрон. учеб. пособие / С. Н. Хонина; Минобрнауки России, Самар. гос. аэрокосм. ун-т им. С. П. Королева (нац. исслед. ун-т). - Электрон. текстовые и граф. дан. - Самара, 2011.

Лабораторные работы предназначены для наглядного ознакомления и проведения исследований по итерационному расчету высокоэффективных фазовых дифракционных оптических элементов (ДОЕ), фокусирующих в радиально-симметричную область, а также формирующих многомодовые лазерные пучки Бесселя. Используемое для проведения расчетов программное обеспечение основано на быстрых алгоритмах для преобразований Фурье и Ханкеля, что позволяет выполнять расчеты для массивов большой размерности. Дано подробное теоретическое описание программно-реализованных методов и алгоритмов.

Лабораторные работы соответствуют программе курса «Оптическая информатика» 010400.62 Прикладная математика и информатика (бакалавриат), 6 факультет, кафедра технической кибернетики.

© Самарский государственный
аэрокосмический университет, 2011

1. Введение

1.1 Формирование волновых полей на основе дифракционных оптических элементов (ДОЕ).

Фазовые ДОЕ - это дифракционные элементы, осуществляющие требуемые преобразования волновых полей без потери световой энергии. То есть действие фазового ДОЕ сводится только к модуляции фазы падающего излучения. Везде в дальнейшем свет рассматривается в рамках скалярной теории дифракции, считается монохроматическим и полностью когерентным и описывается комплексной функцией, зависящей от двух или одной пространственных переменных. Модуль этой комплексной функции называется амплитудой светового поля, а аргумент - фазой.

Распространение света в свободном пространстве описывается в приближении дифракции Кирхгофа, а прохождение света через оптические элементы (линзы и ДОЕ) сводится к умножению комплексной амплитуды падающего света на комплексную функцию пропускания оптического элемента.

Фазовые ДОЕ представляют собой либо плоскую поверхность с микрорельефом, либо плоскопараллельную пластину, прозрачную для используемого света, на поверхности которой имеется микрорельеф. В данном пакете понятие "расчет ДОЕ" эквивалентно понятию "расчет фазовой функции ДОЕ". Результатом расчета являются функции двух переменных $\phi(u,v)$, имеющие смысл фазы ДОЕ. Аналитическая формула $\phi(u,v)$ определяет оптическую поверхность, которая теоретически должна быть изготовлена с помощью традиционных оптических технологий (резания, фрезерования, гравирования, шлифовки и т.п.). Однако здесь мы ориентируемся на технологию ДОЕ, близкую по сути и по используемому оборудованию, к микроэлектронным технологиям. Одной из основных идей этой технологии является учет периодичности фазы ϕ и приведение ее значений в диапазон от нуля до 2π . При этом традиционный оптический элемент толщиной до нескольких тысяч длин волн заменяется дифракционным оптическим элементом с высотой микрорельефа равного всего нескольким длинам волн.

Рельеф на поверхности ДОЕ может иметь регулярную зонную структуру. Например, у зонной пластинки Рэлея рельеф имеет кольцевую структуру, а у дифракционных решеток рельеф - это периодически расположенные канавки. В общем случае рельеф поверхности нерегулярный, как это имеет место у киноформов и цифровых голограмм. При этом зоны описываются более сложными кривыми, чем прямые и окружности. Различие в структуре зон зависит от того, каким методом рассчитывалась фазовая функция ДОЕ. Если решение задачи получено в виде аналитической функции - рельеф ДОЕ имеет регулярную структуру, а если решение задачи получено в ходе итеративного процесса со случайной начальной оценкой - рельеф ДОЕ и структура зон будут нерегулярными.

Глубина рельефа ДОЕ кратна длине волны света, для которого он рассчитывается.

Минимальный период модуляции рельефа ДОЕ (ширина минимальной зоны) оценивается аналогично тому, как оценивается расстояние между центрами колец соседних зон Френеля и определяется угловыми размерами изображения, которое формирует ДОЕ. Этот параметр определяет разрешение технологического оборудования, которое используется при изготовлении ДОЕ.

Качество расчета и изготовления ДОЕ характеризуется степенью отличия сформированного изображения от заданного и величиной световой энергии, идущей на формирование изображения. Качество ДОЕ определяют следующие характеристики:

- энергетическая эффективность - это отношение световой энергии, попавшей в область задания изображения, к световой энергии, падающей на ДОЕ;

- дифракционная эффективность - это отношение световой энергии, попавшей в область задания изображения, к световой энергии, прошедшей через ДОЕ (или отразившейся от него);

- ошибка восстановления изображения (среднее квадратичное отклонение) - это корень квадратный из суммы квадратов разности сформированной и заданной интенсивностей, а усредняющее суммирование осуществляется по всей области задания изображения.

Для задач обработки информации, в которых требуется высокая точность формирования волновых фронтов или распределений интенсивности, основной характеристикой качества ДОЕ является ошибка отклонения. А для задач обработки материалов мощным лазерным излучением требуется высокая степень концентрации энергии в области заданного изображения. Для таких задач требуются фазовые ДОЕ, обладающие высокой энергетической эффективностью (близкой к 100%). Поэтому различаются и методы, с помощью которых рассчитываются соответствующие ДОЕ: для задач оптической обработки информации используются, как правило, методы кодирования цифровых голограмм, а для задач лазерной технологии используются итеративные методы расчета киноформов или лучевые методы расчета фокусаторов. Суть технологических этапов кратко сводится к следующему.

После того как решена обратная задача и по требуемым характеристикам изображения рассчитана фазовая функция ДОЕ нужно отобразить ее на физический носитель в виде рельефа. Существует много способов перевода функции фазы в рельеф ДОЕ, но наиболее широко распространенным является метод литографии, позволяющий получать ступенчатый рельеф.

Исходным для литографии является фотошаблон, который получают с помощью фото или электронного генераторов изображений и который представляет собой бинарную амплитудную маску, состоящую из прозрачных и непрозрачных участков фотопленки. Далее на поверхность подложки наносится слой фоторезиста и освещается ультрафиолетовым светом через фоторезист. В освещенных местах фоторезист меняет свою структуру и при дальнейшей химической обработке, в зависимости от типа резиста, удаляются либо засвеченные, либо незасвеченные участки. В результате этого в слое фоторезиста образуются окна, через которые осуществляется травление материала подложки. После травления и удаления оставшегося резиста на подложке остается рельеф, соответствующий рисунку фазы на фотошаблоне. Глубина рельефа зависит от времени травления. Для получения многоступенчатого рельефа описанный выше процесс повторяется. При этом следующие бинарные фотошаблоны должны с высокой точностью совмещаться с уже имеющимся рельефом. Таким путем получают пропускающие фазовые ДОЕ. Если же на рельеф нанести напылением слой металла, то получится отражающий ДОЕ.

Необходимо также отметить, что процесс изготовления ДОЕ носит, как правило, итерационный характер, состоящий в том, что при необходимости каждый этап повторяется или осуществляется возврат к одному из предшествующих этапов.

1.2 Методы расчета ДОЕ.

Получение фазовой функции пропускания ДОЕ всегда связано с решением некорректной обратной задачи. Эта задача является обратной потому, что она относится к задачам синтеза, в которых требуется найти “источник” по заданным произвольно “данным”. То есть требуется рассчитать фазовую функцию ДОЕ, который должен формировать требуемые распределения амплитуды или фазы. Задача анализа - это прямая задача, в которой исследуется изображение, сформированное заданным “источником”.

Формально расчет фазы ДОЕ сводится к решению интегрального уравнения. Некорректность задачи синтеза заключается в том, что, во-первых, ее решения может не существовать, во-вторых, если решение существует, то оно может быть неединственное, и, в-третьих, решение может быть неустойчивым. Неустойчивость решения обратной задачи означает, что при малых отклонениях фазы ДОЕ от рассчитанной могут возникнуть значительные искажения формируемого изображения.

Что касается неединственности решения задачи синтеза ДОЕ, то, в отличие от задач реконструкции изображений и восстановления фазы, для которых неединственность решения вызывает дополнительные трудности, здесь неединственность оказывается полезной и дает возможность выбирать лучшее из нескольких решений.

В задачах реконструкции изображений причиной некорректности служит шум данных измерения, по которым восстанавливается объект. В задачах синтеза ДОЕ нет шума данных, так как характеристики изображения задаются однозначно. Но требуемое изображение может быть задано таким образом, что не будет существовать решения обратной задачи. То есть в рамках поставленной задачи нельзя рассчитать такого фазового ДОЕ, который бы сформировал требуемое изображение точно. Поэтому процедура регуляризации задачи синтеза ДОЕ сводится к замене требуемого изображения на близкое к нему такое, чтобы решение задачи существовало и было устойчивым. Обратная задача синтеза ДОЕ в общем случае является нелинейной задачей, так как требует решения нелинейного интегрального уравнения. Причина нелинейности в том, что оперировать приходится отдельно с амплитудой или фазой комплексной функции, а не с реальной и мнимой ее частями.

На первых шагах развития цифровой голографии эта нелинейная задача заменялась линейной, в которой амплитудно-фазовая функция голограммы получается как результат обращения линейного интегрального преобразования Френеля или Фурье. Чтобы далее привести амплитудно-фазовую функцию в только амплитудную или только фазовую потребовалось развитие методов кодирования: метод кодирования бинарных амплитудных голограмм Ломана, метод Ли, метод дополнительных элементов и т.д. Определенным обобщением методов кодирования является метод фазовой несущей Кирка-Джонса, который позволяет по амплитудно-фазовой функции рассчитать только фазовый ДОЕ с эффективностью около 30-40%.

Процедура кодирования амплитудно-фазовой функции только фазовой функцией осуществляется последовательно в ходе итеративного решения нелинейного интегрального уравнения. Впервые итеративный путь решения задачи синтеза ДОЕ был предложен в работе Хирша, Джордана и Лезема. Независимо от них аналогичный итеративный алгоритм для восстановления изображений разработали Герчберг и Сакстон. Этот алгоритм, который в дальнейшем будем называть алгоритмом сокращения ошибки или алгоритмом Герчберга-Секстона(ГС), дает базовое решение задачи синтеза фазовых ДОЕ.

Этот алгоритм получил свое дальнейшее развитие в целом ряде последующих работ. В работах Галлагера приведено доказательство сходимости ГС-алгоритма. Фиенап разработал параметрическое обобщение ГС-алгоритма и назвал его алгоритм входа-выхода. Он показал, что алгоритм сокращения ошибки является частным случаем градиентных методов. Наличие параметра (или шага) в итеративном алгоритме позволяет регулировать скорость сходимости и бороться с процессом стагнации, присущем ГС-алгоритму. В работах Брингдала и Вировского для расчета ДОЕ развит итеративный алгоритм диффузии ошибки и предложена итеративная процедура квантования фазы ДОЕ. В работах Сойфера и Котляра разработана адаптивная модернизация ГС-алгоритма, которая применена для расчета радиальных ДОЕ и для итеративного расчета фазовых формирователей мод Бесселя и мод Гаусса-Эрмита.

Альтернативным подходом к решению задачи синтеза ДОЕ является лучевой метод, основанный на аналитическом решении дифференциального уравнения эйконала и построения

хода лучей от точек поверхности ДОЕ к точкам заданного изображения. Этот подход развит в работах Голуба, Данилова, Сисакяна, Сойфера и других. Оптические элементы, рассчитанные этим методом были изготовлены для инфракрасного и видимого диапазонов спектра и прошли успешные испытания. Фазовые функции ДОЕ, рассчитанных лучевым методом имеют регулярную зонную структуру, так как дифференциальное уравнение эйконала решается аналитически в явном виде или с помощью рядов.

По сравнению с методами кодирования и с лучевым методом итеративные алгоритмы решения интегральных уравнений имеют следующие преимущества.

А. Возможность получения приближенного решения с заданной точностью.

В. Возможность получения нескольких решений одной задачи с помощью изменения начальной оценки в итеративном процессе.

С. Совмещение в одном алгоритме решения обратной и прямой задач. То есть текущая оценка фазовой функции ДОЕ, полученная в обратной ветви алгоритма, проверяется и уточняется в прямой ветви итеративного алгоритма.

Д. Универсальность и технологичность итеративных алгоритмов. Универсальность как независимость от вида требуемого изображения, технологичность как приспособленность для компьютерного моделирования, простота программирования.

Е. Возможность учета дополнительных ограничений, накладываемых на искомую функцию фазы ДОЕ: введение операторов квантования, дискретизации или других предсказаний фазы.

1.3 Особенности расчета разных видов ДОЕ.

Нет алгоритма, с помощью которого можно рассчитать все возможные ДОЕ. Для расчета каждого вида ДОЕ можно подобрать оптимальный метод. Фазовые ДОЕ условно можно разделить на следующие виды.

1. ДОЕ, формирующие полутонные изображения: цифровые голограммы, киноформы.
2. ДОЕ, фокусирующие в малую область пространства: “фокусаторы”.
3. ДОЕ, формирующие бинарные изображения: дифракционные решетки, матричные осветители.
4. ДОЕ, формирующие продольные изображения: многофокусные линзы, аксиконы.
5. ДОЕ, формирующие волновые фронты: корректоры, компенсаторы.
6. ДОЕ, генерирующие световые моды: моды Бесселя, моды Гаусса-Эрмита и Гаусса-Лагерра.

Кратко отметим особенности каждого из перечисленных типов ДОЕ.

1. Для цифровых голограмм, формирующих трехмерные или двумерные полутонные изображения, требуется расчет функций пропускания в виде матрицы с большим числом отсчетов $M \times M$. Так, чтобы изготовить голограмму для визуального наблюдения размером 100×100 мм с разрешением 1 мкм потребуется рассчитать матрицу отсчетов с $M=10$. В этом случае используются процедуры интерполяции или экстраполяции отсчетов базового массива голограммы. В простом случае экстраполяция сводится к повторению рассчитанного базового массива размером $N \times N$, $N \ll M$. При этом голограмма или киноформ имеет растровую структуру фазы, а изображение состоит из совокупности регулярной совокупности световых пятен (спеклов).

Также для расчета цифровых голограмм широко используются процедуры кодирования, которые также выполняют операцию интерполяции, потому что каждый отсчет голограммы представляется (кодируется) в виде прямоугольной площадки амплитудного или фазового раstra.

Итеративные методы для расчета голограмм и киноформов также широко применяются.

2. Фокусаторы - это ДОЕ, фокусирующие лазерное излучение в малые области пространства и с заданным распределением энергии в этой области. Такие оптические элементы используются в задачах лазерной технологии для резки, сварки, маркировки или упрочнения материалов. Качество фокусировки оценивается энергетической эффективностью ДОЕ, которая должна быть не ниже 85-95%. Высокая эффективность ДОЕ определяет выбор метода расчета. Лучевой метод позволяет за счет некоторого снижения точности формирования пятна фокусировки достичь рекордных показателей эффективности.

3. Для расчета ДОЕ, формирующих бинарные изображения, многопорядковых дифракционных решеток или матричных осветителей, которые используются для создания оптических нейронных сетей, применяются методы решения линейных систем алгебраических уравнений, стохастические методы или другие методы нелинейной оптимизации.

Выбор метода определяется требованием бинарности профиля таких ДОЕ. Кроме того, требуется, чтобы бинарная фазовая дифракционная решетка формировала много равных по амплитуде порядков (более 64×64) с эффективностью около 80% и с разбросом энергии между порядками не более 10%.

В последнее время растет интерес к синтезу дифракционных решеток с ультракоротким периодом, то есть с периодом равным или меньшим длины волны света. Расчет рельефа поверхности таких решеток основан на точном решении уравнений Максвелла. Постановка обратной задачи в данном случае крайне сложная из-за отсутствия явной аналитической интегральной связи между формой канавок дифракционной решетки и интенсивностью рассеянного излучения. Однако решетки с ультракоротким периодом находят применение как поляризационные расщепители лазерных пучков и для поворотов вектора поляризации излучения.

4. ДОЕ, формирующие продольные изображения, имеют, как правило, радиальную структуру зон рельефа, то есть уровнями равной фазы являются концентрические кольца. Для расчета таких ДОЕ требуется использовать итеративные методы с применением алгоритмов быстрого вычисления преобразования Ханкеля.

При расчете радиальных ДОЕ, формирующих кольцевые распределения интенсивности, в центре колец появляется нежелательный пик интенсивности, для устранения которого используются фазовые винтовые маски.

Кроме того задачу расчета радиальных ДОЕ можно свести, в ряде случаев, к задаче расчета одномерного ДОЕ.

Используются такие элементы для увеличения глубины резкости микрообъективов для лазерной записи и считывания с диска, для обострения лазерных пучков на больших расстояниях, в двухфокусных микроскопах и т.д.

5. Фазовые формирователи волновых фронтов используются как корректоры аберраций для стандартных объективов, как компенсаторы для контроля асферических зеркал, или как формирователи эталонных волновых фронтов в интерферометрах Тваймана-Грина.

Расчет таких ДОЕ итеративными методами является как бы обратной задачей к задаче расчета киноформов, так как требуется рассчитать ДОЕ по известному распределению фазы светового поля в некоторой плоскости, а не по заданному распределению интенсивности.

Качество корректоров оценивается не столько их энергетической эффективностью как степенью точности, с которой они формируют заданный волновой фронт. Степень точности должна быть высокой - десятые и сотые доли длины волны.

6. Фазовые формирователи световых мод применяются для параллельного ввода лазерного излучения в набор световых волокон или в качестве пространственных фильтров для анализа поперечно-модового состава лазерного излучения.

При расчете таких ДОЕ - моданов используются методы кодирования типа Кирка-Джонса. ДОЕ, рассчитанные таким способом, позволяют сформировать с высокой точностью сразу несколько мод Гаусса-Эрмита или Гаусса-Лагерра, распространяющихся в разных порядках дифракции. Однако уровень эффективности таких моданов остается невысоким - около 40%. Применение итеративных алгоритмов для расчета ДОЕ, формирующих гауссовы моды позволяет получить эффективность 90% за счет некоторого уменьшения точности формирования требуемых мод.

ДОЕ, формирующие моды Бесселя, используются для генерации бездифракционных пучков, распространяющихся в пространстве без изменения своей формы и для исследования свойств оптического разряда в газовых средах. Для расчета формирователей мод Бесселя применяются методы кодирования и итеративные методы, аналогичные ГС-алгоритму.

2 Итерационные методы расчета радиально-симметричных ДОЭ

2.1 Алгоритм уменьшения ошибки (алгоритм Герчберга-Секстона)

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

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

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 Френеля

$$F(\xi, \eta) = \frac{ik}{2\pi z} e^{ikz} \iint_{-\infty}^{\infty} W(u, \mathbf{v}) H(u - \xi, \mathbf{v} - \eta, z) du d\mathbf{v}, \quad (2.1)$$

где

$$H(u - \xi, \mathbf{v} - \eta, z) = \exp \left[\frac{ik}{2z} \{ (u - \xi)^2 + (\mathbf{v} - \eta)^2 \} \right] \quad (2.2)$$

— функция импульсного отклика свободного пространства в приближении Френеля, z — это расстояние между ДОЭ и плоскостью наблюдения, $k = 2\pi/\lambda$. В уравнении (2.1) комплексная амплитуда $W(u, \mathbf{v})$ в приближении тонкого оптического элемента (приближение транспаранта), которое не учитывает рефракцию лучей, равна произведению комплексной амплитуды $W_0(u, \mathbf{v})$ на собственную функцию пропускания $T(u, \mathbf{v})$ ДОЭ:

$$W(u, v) = W_0(u, v) T(u, v). \quad (2.3)$$

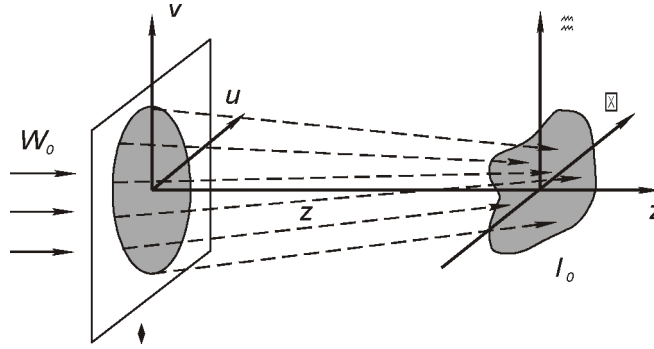


Рисунок 2.1 Схема формирования изображения при помощи ДОО

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

$$T(u, v) = \exp[i g(u, v)], \quad (2.4)$$

где $g(u, v)$ — заданная фаза ДОО. Для иллюстрации вышеуказанных обозначений следует обратиться к рисунку 2.1. Задачу расчета фазовой функции ДОО $g(u, v)$ можно свести к решению нелинейного интегрального уравнения

$$I_0(\xi, \eta) = |F(\xi, \eta)|^2 = \left| \frac{k}{2\pi z} \int \int_{-\infty}^{\infty} A_0(u, v) e^{i\varphi(u, v)} H(u - \xi, v - \eta, z) du dv \right|^2, \quad (2.5)$$

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 (2.5) методом последовательных приближений. Алгоритм Герчберга — Секстона (ГС), или алгоритм уменьшения ошибки, включает следующие шаги (см. его схему на рисунке 2.2):

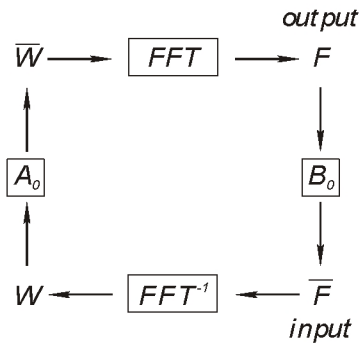


Рисунок 2.2 Схема алгоритма уменьшения ошибки Герчберга-Секстона

- 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 (2.1);

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

$$\bar{F}(\xi, \eta) = B_0(\xi, \eta) F(\xi, \eta) |F(\xi, \eta)|^{-1}, \quad (2.6)$$

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

4) вычисляется преобразование, обратное (2.1) относительно функции $\bar{F}(\xi, \eta)$:

$$W(u, v) = \frac{ik}{2\pi z} e^{-ikz} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \bar{F}(\xi, \eta) H^*(\xi - u, \eta - v, z) d\xi d\eta; \quad (2.7)$$

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

$$\bar{W}(u, v) = \begin{cases} A_0(u, v) W(u, v) |W(u, v)|^{-1}, & (u, v) \in Q, \\ 0, & (u, v) \notin Q, \end{cases} \quad (2.8)$$

где Q — форма апертуры ДОЭ;

6) переход к шагу 2.

Эта процедура повторяется до тех пор, пока ошибки δF и δW не перестанут значительно меняться:

$$\delta_F^2 = \frac{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [F(\xi, \eta) - B_0(\xi, \eta)]^2 d\xi d\eta}{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} B_0^2(\xi, \eta) d\xi d\eta}, \quad (2.9)$$

$$\delta_W^2 = \frac{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [W(u, v) - A_0(u, v)]^2 du dv}{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} A_0^2(u, v) du dv}. \quad (2.10)$$

Алгоритм ГС называют алгоритмом уменьшения ошибки потому, что было показано, что ошибки (2.9) и (2.10) с ростом числа итераций не возрастают. Кроме того, было показано, что алгоритм ГС является вариантом градиентного метода или градиентного метода наискорейшего спуска, при помощи которого можно минимизировать функционал среднеквадратичного отклонения амплитуды восстановленного изображения от заданного значения:

$$\varepsilon_0 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [F(\xi, \eta) - B_0(\xi, \eta)]^2 d\xi d\eta. \quad (2.11)$$

Заметим однако, что процесс сходимости алгоритма ГС характеризуется эффектом стагнации: в ходе нескольких начальных итераций ошибка δF (или δW) быстро уменьшается, а все последующие итерации не приводят к ее заметному уменьшению. Эффект стагнации показывает, что алгоритм достигает локального минимума функционала (2.11), 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.

2.2 Адаптивно-аддитивный алгоритм

Ниже предлагается алгоритм, который далее называется адаптивно-аддитивным алгоритмом (АА-алгоритмом), получаемый 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 (2.6) 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 (2.6) попробовать использовать замену, в которой обе функции (заданная и рассчитанная аналитическая) участвуют в виде линейной комбинации с разными весами:

$$|\bar{F}_n(\xi, \eta)| = |\gamma B_0(\xi, \eta) + (1-\gamma)|F_n(\xi, \eta)||, \quad (2.12)$$

В этом случае замена (2.6) в итеративном ГС-алгоритме принимает вид

$$\bar{F}_n(\xi, \eta) = |\bar{F}_n(\xi, \eta)| F_n(\xi, \eta) |F_n(\xi, \eta)|^{-1}. \quad (2.13)$$

Диапазон значений параметра γ находится из условия, что среднеквадратичное отклонение ε_0 , определяемое соотношением (2.11), не увеличивается при использовании замены (2.12). То есть мы предполагаем, что выполняется следующее условие:

$$\begin{aligned} \bar{\varepsilon}_0 &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [|\bar{F}_n| - B_0]^2 d\xi d\eta = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [\gamma B_0 + (1-\gamma)|F_n| - B_0]^2 d\xi d\eta = \\ &= (1-\gamma)^2 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [|F_n| - B_0]^2 d\xi d\eta \leq \varepsilon_0 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [|F_n| - B_0]^2 d\xi d\eta. \end{aligned} \quad (2.14)$$

Из уравнения (2.14) следует, что весовой коэффициент следует выбирать из условия $0 \leq \gamma \leq 2$. Заметим, что для $\gamma = 1$ замена (2.12) переходит в уравнение (2.6) алгоритма ГС, в то время как $\gamma = 2$ приводит к замене “зеркального отражения”:

$$|\bar{F}_n(\xi, \eta)| = |2B_0(\xi, \eta) - |F_n(\xi, \eta)||. \quad (2.15)$$

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 (2.15) имела следующий вид:

$$\bar{I}_n(\xi, \eta) = \begin{cases} |2I_0(\xi, \eta) - I_n(\xi, \eta)|, & (\xi, \eta) \in \Omega, \\ I_n(\xi, \eta), & (\xi, \eta) \notin \Omega, \end{cases} \quad (2.16)$$

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 $I_n > 2I_0$, то $\bar{I}_n(\xi, \eta) = 0$.

Покажем, что замена (2.12) в адаптивно-аддитивном алгоритме минимизирует функционал невязки ε_1 , который представляет собой среднеквадратичное отклонение рассчитанной амплитуды от заданной в плоскости ДОЭ:

$$\varepsilon_1 = \int \int_{-\infty}^{\infty} [A_0(u, \mathbf{v}) - |W(u, \mathbf{v})|]^2 du d\mathbf{v}, \quad (2.17)$$

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

$$\delta \varepsilon_1 = 2 \operatorname{Re} \left\{ \int \int_{-\infty}^{\infty} [W_n - A_0 W_n |W_n|^{-1}] \delta W_n^* du d\mathbf{v} \right\}, \quad (2.18)$$

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

При получении уравнения (2.18) мы использовали соотношения

$$\delta \left\{ \int \int_{-\infty}^{\infty} A_0^2(u, \mathbf{v}) du d\mathbf{v} \right\} = 0, \quad (2.19)$$

$$\delta \left\{ \int \int_{-\infty}^{\infty} |W(u, \mathbf{v})|^2 du d\mathbf{v} \right\} \neq 0, \quad (2.20)$$

$$\delta(WW^*) = W\delta W^* + W^*\delta W = 2 \operatorname{Re}\{W\delta W^*\}. \quad (2.21)$$

Используя равенства Парсеваля

$$\int \int_{-\infty}^{\infty} |W(u, \mathbf{v})|^2 du d\mathbf{v} = \int \int_{-\infty}^{\infty} |\bar{F}(\xi, \eta)|^2 d\xi d\eta, \quad (2.22)$$

$$\int \int_{-\infty}^{\infty} |\bar{W}(u, \mathbf{v})|^2 du d\mathbf{v} = \int \int_{-\infty}^{\infty} |F(\xi, \eta)|^2 d\xi d\eta, \quad (2.23)$$

где $\bar{W} = DA$, $W = AW$, $|W|^{-1}$ — амплитуда света в области ДОЭ после осуществления замены (2.8), можно показать, что выполняется следующее соотношение:

$$\operatorname{Re} \left\{ \int \int_{-\infty}^{\infty} [W_n - A_0 W_n |W_n|^{-1}] \delta W_n^* du d\mathbf{v} \right\} = \operatorname{Re} \left\{ \int \int_{-\infty}^{\infty} [\bar{F}_n - F_n] \delta \bar{F}_n^* d\xi d\eta \right\}, \quad (2.24)$$

где $\bar{F}_n = \mathfrak{Z}\{W_n\}$, $F_n = \mathfrak{Z} D_A \mathfrak{Z}^{-1} \{\bar{F}_n\} = \mathfrak{Z}\{A_0 W_n |W_n|^{-1}\}$.

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

$$\delta \bar{F}_n = \bar{F}_{n+1} - \bar{F}_n = \gamma (F_n - \bar{F}_n), \quad \gamma > 0. \quad (2.25)$$

С учетом уравнения (2.25), вместо (2.18) мы получим

$$\delta \varepsilon_1 = -2\gamma \int \int_{-\infty}^{\infty} |\bar{F}_n - F_n|^2 d\xi d\eta < 0. \quad (2.26)$$

Из уравнения (2.25) следует параметрическое уравнение для интегрального расчета функции $F(\xi, \eta)$:

$$\bar{F}_{n+1} = (1 - \gamma) \bar{F}_n + \gamma F_n = (1 - \gamma) \bar{F}_n + \mathcal{R} D_A \mathcal{F}^{-1} \{ \bar{F}_n \}. \quad (2.27)$$

Уравнение (2.27) предполагает, что на (n+1)-м шаге входная амплитуда представляет собой линейную комбинацию входной и выходной амплитуды на n-м шаге. Алгоритм (2.27) аналогичен алгоритму входа-выхода, и предыдущие рассуждения следует рассматривать как теоретическое обоснование его оптимальности.

Для алгоритма (2.27) среднеквадратичная ошибка амплитуды не возрастает с ростом числа итераций при условии, что параметр γ удовлетворяет неравенству

$$0 < \gamma \leq 2. \quad (2.28)$$

Действительно, введя обозначения

$$F_n = B_n e^{i\varphi_n}, \quad \bar{F}_n = \bar{B}_n e^{i\bar{\varphi}_n}, \quad W_n = A_n e^{i\psi_n}, \quad \bar{W}_n = A_n e^{i\psi_n}, \quad (2.29)$$

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

$$\begin{aligned} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |A_0 - A_{n+1}|^2 du dv &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |A_0 e^{i\psi_{n+1}} - A_{n+1} e^{i\psi_{n+1}}|^2 du dv \leq \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |A_0 e^{i\psi_n} - A_{n+1} e^{i\psi_{n+1}}|^2 du dv = \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |B_n e^{i\varphi_n} - [(1 - \gamma) \bar{B}_n e^{i\bar{\varphi}_n} + \gamma B_n e^{i\varphi_n}]|^2 d\xi d\eta = (1 - \gamma)^2 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |B_n e^{i\varphi_n} - \bar{B}_n e^{i\bar{\varphi}_n}|^2 d\xi d\eta = \\ &= (1 - \gamma)^2 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |A_0 e^{i\psi_n} - A_n e^{i\psi_n}|^2 du dv = (1 - \gamma)^2 \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |A_0 - A_n|^2 du dv. \end{aligned} \quad (2.30)$$

Из уравнения (2.30) следует, что при условии (2.28) имеет место релаксация среднеквадратической ошибки:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |A_0 - A_{n+1}|^2 du dv \leq \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |A_0 - A_n|^2 du dv. \quad (2.31)$$

Отметим, что в алгоритме (2.27) учитывается, что комплексная амплитуда удовлетворяет ограничению в плоскости ДОО, но не учитывается, что модули входной и выходной функций комплексной амплитуды в Фурье-плоскости, а именно $\bar{F}(\xi, \eta)$ и $F(\xi, \eta)$ должны стремиться к заданной функции $B_0(\xi, \eta)$. Чтобы использовать это условие в алгоритме (2.27) предположим, что входная функция на n-м шаге итерации подчиняется следующему ограничению

$$\bar{F}_n(\xi, \eta) = D_B F(\xi, \eta) = B_0(\xi, \eta) F(\xi, \eta) |F(\xi, \eta)|^{-1}, \quad (2.32)$$

и что на (n+1)-м шаге она подчиняется уравнению (2.27). Поэтому комбинация уравнений (2.27) и (2.32) приводит к адаптивно-аддитивному алгоритму.

$$\bar{F}_{n+1}(\xi, \eta) = (1 - \gamma) B_0(\xi, \eta) \frac{F_n(\xi, \eta)}{|F_n(\xi, \eta)|} + \gamma F_n(\xi, \eta). \quad (2.33)$$

Заметим, что уравнение (2.33) отличается от (2.12) перестановкой весовых коэффициентов: $\gamma \rightarrow 1 - \gamma$.

Рассмотрим функционал невязки ε_0 , но с весовой функцией:

$$\bar{\varepsilon}_0 = \int_{-\infty}^{\infty} \int S(\xi, \eta) \left[|F(\xi, \eta)| - B_0(\xi, \eta) \right]^2 d\xi d\eta, \quad (2.34)$$

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

$$S(\xi, \eta) = \begin{cases} \gamma, & (\xi, \eta) \in \Omega, \\ 0, & (\xi, \eta) \notin \Omega, \end{cases} \quad (2.35)$$

где γ — постоянная.

Вариация функционала (2.34) имеет вид

$$\delta \bar{\varepsilon}_0 = 2 \operatorname{Re} \int_{-\infty}^{\infty} \int S(\xi, \eta) \left[F - B_0 \frac{F}{|F|} \right] \delta F^* d\xi d\eta. \quad (2.36)$$

Итеративное уравнение, которое позволяет получить функцию $F(\xi, \eta)$, минимизирующую функционал (2.34), следует из требования равенства нулю подынтегрального выражения в уравнении (2.36) и имеет вид:

$$\bar{F}_{n+1} = F_n + S \left[B_0 \frac{F_n}{|F_n|} - F_n \right] = (1 - S) F_n + S B_0 \frac{F_n}{|F_n|}, \quad (2.37)$$

где F_n и $\bar{F}_{n+1}$ функции выхода на n -м шаге и входа на $(n+1)$ -м шаге, соответственно. При получении уравнения (2.37) 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: $\delta F_n = \bar{F}_{n+1} - F_n$.

Преимущества АА-алгоритма можно проиллюстрировать на следующих примерах.

2.3 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. и Папулиса А., так и в более поздних работах Вировского Ф. и Фиенапа Дж., и позволяет частично устранять эффект стагнации, присущий алгоритму Герчберга-Секстона.

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 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):

$$\varepsilon(\alpha) = \int_{-\infty}^{\infty} \left[|F_0(u)| - D_A |F(u)| \right]^2 du + \alpha \int_{-\infty}^{\infty} Q(u) |F(u)|^2 du, \quad (2.47)$$

где $Q(u) = \sum_{n=0}^N |c_n| u^{2n}$ — так называемая стабилизирующая функция,

$$D_A f = \begin{cases} f, & u \in \Omega_A \\ 0, & u \notin \Omega_A \end{cases} \quad \text{— оператор пространственного ограничения, } |F_0(u)| \quad \text{— заданное}$$

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 (2.47) квадратичный по амплитуде светового поля.

Вариация функционала (2.47) запишется в виде

$$\delta \varepsilon = 2 \operatorname{Re} \left[\int_{-\infty}^{\infty} \left(-|F_0(u)| \frac{F(u)}{|F(u)|} + D_A F(u) + \alpha Q(u) F(u) \right) \delta F^*(u) du \right]. \quad (2.48)$$

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

$$\delta F = \bar{F} - F, \quad (2.49)$$

то максимальная отрицательная вариация (2.48) функционала (2.47) достигается при условии

$$\bar{F} - F = \gamma G(u) = \gamma \left(|F_0(u)| \frac{F(u)}{|F(u)|} - D_A F(u) - \alpha Q(u) F(u) \right), \quad (2.50)$$

где $\gamma > 0$ — параметр адаптации алгоритма или константа релаксации. Действительно, при условии (2.50) вариация (2.48) будет иметь вид

$$\delta \varepsilon = -2\gamma \int_{-\infty}^{\infty} |G(u)|^2 du \leq 0. \quad (2.51)$$

Из условия (2.50) следует двухпараметрическое итеративное уравнение расчета ДОЭ:

$$\bar{F}_{n+1} = [1 - \gamma(D_A + \alpha Q(u))] F_n + \gamma |F_0(u)| \frac{F_n(u)}{|F_n(u)|}. \quad (2.52)$$

Заметим, что в уравнении (2.52) комплексная амплитуда на n -ом шаге итераций $F_n(u)$ является результатом удовлетворения ограничениям в плоскости ДОЭ и представима в виде

$$F_n(u) = \mathfrak{S} D_B W_B \mathfrak{S}^{-1} \bar{F}_n(u), \quad (2.53)$$

где $\bar{F}_n(u)$ — комплексная амплитуда на n -ом шаге, после замены (2.52), $D_B f(x) = \begin{cases} f(x), & x \in \Omega_B, \\ 0, & x \notin \Omega_B \end{cases}$ — оператор пространственного ограничения, Ω_B — форма апертуры ДОЭ, $W_B f(x) = |f_0(x)| (|f(x)| / |f_0(x)|)$ — оператор присвоения заданной амплитуды $|f_0(x)|$ освещающего пучка комплексной функции $f(x)$, описывающей световое поле в плоскости ДОЭ.

Итак, уравнения (2.52) и (2.53) 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 (2.52) получим следующее выражение:

$$\bar{F}_{n+1} = [1 - D_A] F_n + |F_0(n)| \frac{F_n(n)}{|F_n(n)|} = \begin{cases} |F_0(n)| \frac{F_n(n)}{|F_n(n)|}, & u \in \Omega_A, \\ F_n(n), & u \notin \Omega_A. \end{cases} \quad (2.54)$$

Алгоритм (2.54) сочетает в себе свойства алгоритма ГС и алгоритма Герчберга — Папулиса. Действительно, в соответствии с (2.54) 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 (2.54) можно интерпретировать как алгоритм ГС, обладающий сверхразрешающим свойством алгоритма Герчберга — Папулиса. Уравнение (2.54) рассматривалось как реализация идеи дополнительной степени свободы для устранения стагнации итеративного алгоритма.

При выборе параметров $\gamma \neq 1$, $\alpha = 0$ вместо уравнения (2.52) получим выражение

$$\bar{F}_{n+1} = \begin{cases} \gamma |F_0(u)| \frac{F_n(u)}{|F_n(u)|} + (1 - \gamma) F_n, & u \in \Omega_A, \\ F_n(u), & u \notin \Omega_A. \end{cases} \quad (2.55)$$

Уравнение (2.55) является аналогом алгоритма входа-выхода, метода ортогональных проекций, адаптивного и маятниковых алгоритмов, а также выражает идею “мягкого” кодирования комплексной амплитуды. Релаксационный параметр γ в уравнении (2.55) позволяет регулировать скорость сходимости алгоритма, и при выборе $\gamma = 2$ эта скорость обычно бывает максимальной.

При выборе параметров $\gamma = 1$, $\alpha \neq 0$ вместо уравнения (2.52) получим выражение

$$\bar{F}_{n+1} = [(1-\alpha) - D_A] F_n + |F_0(u)| \frac{F_n(u)}{|F_n(u)|} = \begin{cases} |F_0(u)| \frac{F_n(u)}{|F_n(u)|} - \alpha F_n, & u \in \Omega_A, \\ (1-\alpha) F_n(u), & u \notin \Omega_A. \end{cases} \quad (2.56)$$

Алгоритм (2.56) аналогичен регуляризованному алгоритму Герчберга — Папулиса, предложенному для задачи восстановления действительного объекта по зашумленному изображению. Параметр регуляризации α в уравнении (2.56) регулирует вклад рассчитанной амплитуды света $F_n(u)$ внутри и вне заданной области Ω_A .

В общем случае АР-алгоритм (2.52) можно переписать в виде

$$\bar{F}_{n+1} = \begin{cases} \gamma |F_0(u)| \frac{F_n(u)}{|F_n(u)|} + (1-\gamma-\gamma\alpha)F_n, & u \in \Omega_A, \\ (1-\gamma\alpha) F_n(u), & u \notin \Omega_A. \end{cases} \quad (2.57)$$

Интересно отметить, что ни при каких значениях параметров γ и α нельзя перейти от АР-алгоритма (2.57) к алгоритму сокращения ошибки, имеющего в данных обозначениях следующий вид:

$$\bar{F}_{n+1} = \begin{cases} |F_0(u)| \frac{F_n(u)}{|F_n(u)|}, & u \in \Omega_A, \\ 0, & u \notin \Omega_A. \end{cases} \quad (2.58)$$

Переход от уравнения (2.57) к уравнению (2.58) возможен только при условии рассмотрения релаксационной константы как двойной постоянной, принимающей разные значения внутри и вне заданной области:

$$\gamma = \begin{cases} (1+\alpha)^{-1}, & u \in \Omega_A, \\ \alpha^{-1}, & u \notin \Omega_A. \end{cases} \quad (2.59)$$

2.4 Фокусировка лазерного излучения в радиально-симметричную область

Для фокусировки когерентного лазерного света в радиально-симметричную область подходит метод стационарной фазы. Например, фазовая функция фокусатора в кольцо с постоянной интенсивностью рассчитывается из соотношения

$$S(r) = \frac{k}{f} \frac{R_1}{A} \left\{ Ar \left(A^2 r^2 + 1 \right)^{1/2} + \ln \left[Ar + \left(A^2 r^2 + 1 \right)^{1/2} \right] \right\}, \quad (2.60)$$

где A - радиус апертуры фокусатора, R_1 и R_2 - внутренний и внешний радиусы кольца.

Однако следует отметить, что оптический элемент, рассчитанный при помощи уравнения (2.60), создает в плоскости Фурье распределение интенсивности, которое, как правило, значительно отличается от заданного (не менее чем на 40%). Для того, чтобы получить более точное распределение интенсивности, используется итеративный алгоритм Герчберга-Сакстона. На первом этапе, начальную фазовую функцию можно выбрать случайной или при помощи уравнения (2.60).

Итеративный алгоритм основан на применении преобразования Ханкеля нулевого порядка, которые связывает радиальную амплитуду света $F(r)$ в фокальной плоскости линзы с функцией пропускания фокусатора $t(r)$:

$$F(\rho) = \frac{k}{f} \int_0^a t(r) \cdot J_0 \left(\frac{k}{f} r \rho \right) \cdot r dr \quad (2.61)$$

При помощи экспоненциальной замены переменных

$$r = r_0 \exp(x), \quad \rho = \rho_0 \exp(y),$$

где r_0 и ρ_0 - постоянные, мы можем свести интегральное уравнение к интегралу свертки

$$P(y) = \int_{-\infty}^{\infty} f(x) g(x+y) dx \quad (2.62)$$

где

$$P(y) = F(\rho_0 e^y) e^{y/4}, \quad f(x) = t(r_0 e^x) r_0^2 e^{7x/4},$$

$$g(x) = J_0 \left(\frac{k}{f} r_0 \rho_0 e^x \right) e^{x/4}, \quad \omega = \ln \left(\frac{a}{r_0} \right)$$

С учетом преобразования Фурье уравнение (2.62) примет вид

$$F(\rho) = \left(\frac{\rho_0}{\rho} \right)^{1/4} \int_{-\infty}^{\infty} F(-v) G(v) \exp \left(-i v \ln \frac{\rho_0}{\rho} \right) dv \quad (2.63)$$

где $F(v)$ - Фурье-образ от $f(x)$, $G(v)$ - Фурье-образ от $g(x)$. Обратное преобразование Фурье вычисляется аналогично.

Использование итеративного алгоритма Герчберга-Сакстона основано на последовательном применении прямого и обратного преобразования Ханкеля с наложением соответствующих ограничений. Тогда, на n -м шаге итераций найденная функция пропускания $t_n(r)$ оптического элемента заменяется на функцию $t'_n(r)$ по правилу

$$t'_n(r) = \begin{cases} \sqrt{I_0(r)} t_n(r) |t_n(r)|^{-1}, & r \in [0, a] \\ 0, & r \notin [0, a] \end{cases} \quad (2.64)$$

а вычисленная функция комплексной амплитуды в фокальной плоскости заменяется на функцию $F'_n(\rho)$:

$$F'_n(\rho) = \begin{cases} \sqrt{I_1(\rho)} F_n(\rho) |F_n(\rho)|^{-1}, & \rho \in [R_1, R_2] \\ 0, & \rho \notin [R_1, R_2] \end{cases} \quad (2.65)$$

Таким способом применяем стандартный алгоритм Герчберга-Сакстона. Однако для того, чтобы улучшить сходимость, можно использовать модификации замены (2.65), введя некоторый адаптивный коэффициент α :

$$F'_n(\rho) = \begin{cases} \left[\sqrt{I_1(\rho)} - \alpha(|F_n(\rho)| - \sqrt{I_1(\rho)}) \right] F_n(\rho) |F_n(\rho)|^{-1}, & \rho \in [R_1, R_2] \\ 0, & \rho \notin [R_1, R_2] \end{cases} \quad (2.66)$$

$$F'_n(\rho) = \begin{cases} \left[\sqrt{I_1(\rho)} - \alpha(|F_n(\rho)| - \sqrt{I_1(\rho)}) \right] F_n(\rho) |F_n(\rho)|^{-1}, & \rho \in [R_1, R_2] \\ F_n(\rho), & \rho \notin [R_1, R_2] \end{cases} \quad (2.67)$$

Чтобы получить радиальное распределение интенсивности с нулевой энергией в центре кольца в точке $r=0$, необходимо рассчитать фокусатор с винтовой функцией пропускания

$$\tau_m(r, \varphi) = \tau(r) \exp(im\varphi) \quad (2.68)$$

В этом случае вместо уравнения (2.61) следует использовать преобразование

$$F(\rho, \psi) = \frac{k}{f} i^m e^{im\psi} \int_0^a \tau(r) \cdot J_m\left(\frac{k}{f} r \rho\right) \cdot r dr \quad (2.69)$$

Алгоритм включает следующие шаги:

1. Присвоение значения начальной фазы (случайное или используя уравнение (2.60)).
2. Прямое преобразование Ханкеля.
3. Замена распределения интенсивности значением, заданным (2.65), (2.66) или (2.67).
4. Обратное преобразование Ханкеля.
5. Амплитуда фокусатора заменяется известным значением на основании уравнения (2.64).
6. Проверяется условие завершения алгоритма. Если оно выполняется, следует выход из алгоритма.
7. Переход к шагу 2.

3. Описание оболочки пользователя

Оболочка предназначена для облегчения работы с программным обеспечением. В нее входят меню, окно ввода параметров, окна для визуализации работы программного обеспечения, информационное окно, окно сообщения.

Для правильного отображения получаемых двумерных распределений интенсивности света и фазы ДОЭ необходимо перед запуском программы в свойствах экрана устанавливать качество цветопередачи в 16-бит.

Основное меню служит для выбора операции, которую пользователь желает выполнить.

Если требуется ввод параметров, то появляется окно ввода параметров (Task Parameters). Оно содержит несколько строк параметров, содержащих наименование параметра и предлагаемое значение.

Если около значения параметра находится выпадающий список, то этот параметр предусматривает ограниченный набор значений. Просмотреть весь набор и выбрать одно из значений можно с помощью стрелок “вверх” и “вниз”. Если значение параметра не устраивает пользователя, он должен его отредактировать.

Информационное окно (protocol) используется для вывода информации пользователю. При проведении расчета содержимое окна записывается в текстовый лог-файл, расположенный в том же каталоге, что и основная программа.

Графические окна используются для вывода графической информации в процессе исполнения программы. Информация может визуализироваться различными способами: двумерный полутоновой (серая палитра или негативная серая палитра), в трехмерном виде, в виде одномерных сечений. Также выводятся графики ошибок.

Окно сообщения используется для извещения пользователя о какой-либо особой ситуации.

4. Использование программного обеспечения: общее описание

При запуске программы появляется головное меню, содержащее следующие поля:

- “Action / Execute” – рассчитать ДОЭ с выбранными значениями параметров;
- “Action / Save mask Data” – сохранить в файл рассчитанную фазу ДОЭ;
- “Action / Quit” - выход.
- “Windows / ...” – скрыть / показать окна;

Для выбора пункта меню необходимо активизировать нужное поле нажатием соответствующей выделенной квадратными скобками буквы.

В зависимости от активности того или иного окна, в главном меню программы могут появляться дополнительные пункты:

- Protocol - для информационного окна, содержит подпункты:
 - Clear – очистка журнала.
 - Change File Name – изменение имени файла журнала.
- Graph - для графических окон с сечениями, содержит подпункты:
 - Change Color for Graphs – задание цвета, толщины и типа линии.
 - Zoom & Scale Parameters – задание диапазонов по координатным осям, характеризующим видимую область графика (в относительных координатах).

- Range Information for Graph – информация о видимой области графика (в абсолютных координатах), только для просмотра, внесение изменений невозможно.
- Zoom In / Zoom Out – увеличение / уменьшение.
- Surface - для графических окон с полутоновыми изображениями, содержит подпункты:
 - 2D View – переключение к двумерному виду.
 - 3D View – переключение к трехмерному виду.
 - Parameters -> Surface Viewing Part Parameters -
 - Parameters -> 3D View Parameters – задание некоторых параметров для трехмерного вида (например угол поворота).
 - Parameters -> 3D View Colors – задание цветов для трехмерного вида.

При запуске программы появляется основное окно параметров, запрашиваемых для расчета ДОЭ, фокусирующего в радиально симметричную область:

- Beam Parameters (параметры пучка)
 - Radius (радиус апертуры для плоской волны, радиус перетяжки для гауссового пучка, радиус перетяжки для супергауссового пучка)
 - Type (падающий на ДОЭ пучок: ограниченная плоская волна, гауссовый пучок, или супергауссовый пучок)
 - Cut Out level (уровень обрезания гауссового пучка)
- Ring Radii (радиусы фокусируемого кольца)
 - Internal (внутренний радиус кольца)
 - External (внешний радиус кольца)
- Physical Parameters (физические параметры)
 - Focus (фокусное расстояние сферической линзы)
 - Wave Length (длина волны светового излучения)
- Process Parameters (параметры алгоритма)
 - Init Phase, Type (фаза на нулевой итерации: случайная или полученная с помощью геометрической оптики)
 - Method (метод расчета: simple - алгоритм уменьшения ошибки, см. 2.1, cut adaptive - адаптивно-аддитивный алгоритм, см. 2.2, uncut adaptive - адаптивно-регуляризационный алгоритм, см. 2.3)
 - Iteration Number (число итераций)
 - Bessel Order (порядок функции Бесселя)
 - Radial Dimension (число отсчетов для построения радиальных сечений)
 - Output Mask Dimension (число отсчетов для построения полутоновых изображений)

Расчет начинается по нажатию клавиши 'Execute'. При этом в правой части экрана дисплея визуализируется в двумерном полутоновом виде рассчитываемый ДОЭ, а в нижней части экрана - рассчитанная интенсивность в заданных плоскостях. Данную процедуру можно повторять с различным числом итераций, пока пользователя не удовлетворит достигнутый результат.

В информационном окне выводятся заданные параметры.

5. Тестовые примеры

Рассмотрим пример расчета фокусатора в кольцо. После нажатия кнопки “Execute” проводится расчет с использованием параметров, заданных в окне “Task Parameters”. Пример завершенного расчета фокусатора приведен на рис. 5.1.

В окне “Radial Section of Beam Intensity” показано радиальное сечение интенсивности во входной плоскости. Красным цветом показано исходное распределение (на рис. 5.1 это усеченная гауссовская функция). Синим цветом показано распределение, получающееся в конце каждой итерации, т.е. после выполнения обратного преобразования Ханкеля.

В окне “Radial Section of DOE Phase” показано радиальное сечение фазы рассчитанного ДОО.

В окне “Radial Section of Spectrum” показано радиальное сечение интенсивности в фокальной плоскости линзы. Красным цветом показано идеальное распределение интенсивности, которое требуется получить, синим цветом показано полученное распределение.

В окне “Phase of DOE” показана фаза рассчитанного ДОО.

В окне “Spectrum” показана интенсивность в фокальной плоскости линзы.

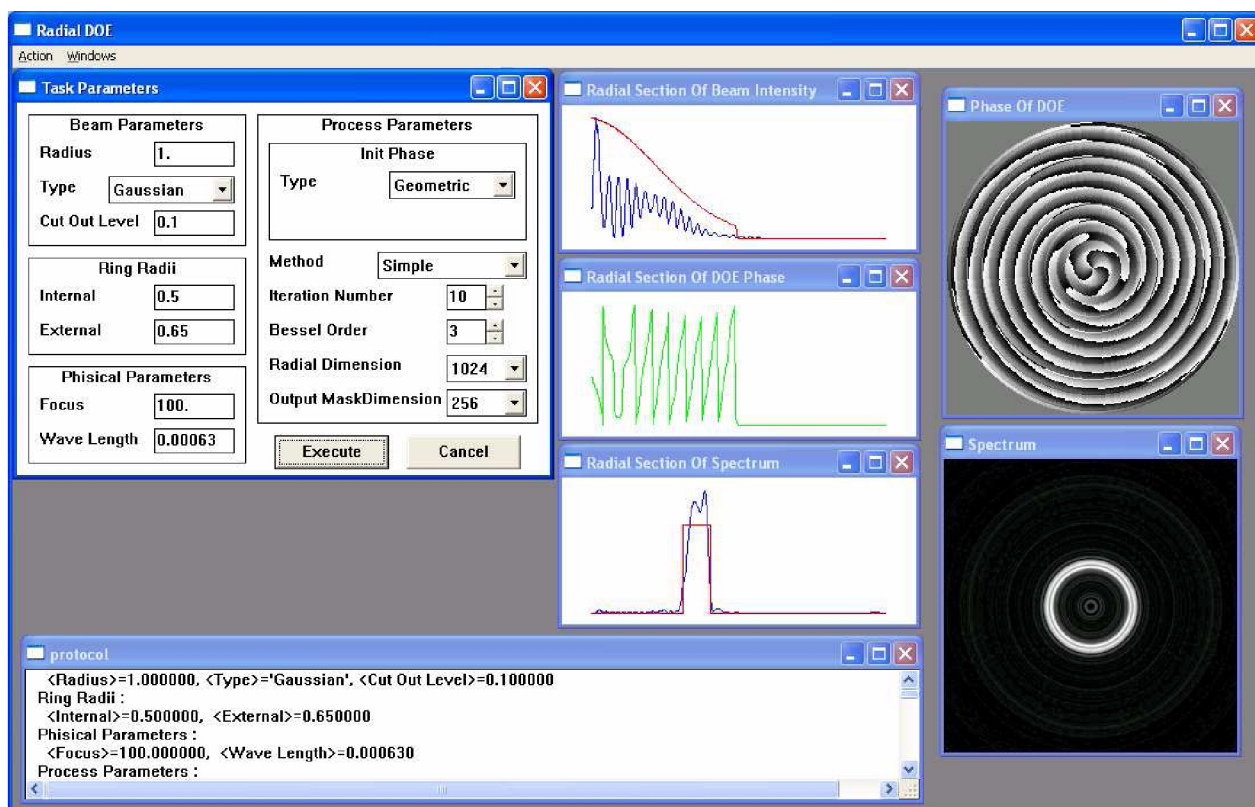


Рис. 5.1

Задания на лабораторные работы

1. Рассчитать с помощью пакета RadialDOE фазовый дифракционный элемент размерностью 256x256, формирующий в частотной плоскости (задней фокальной плоскости сферической линзы) кольцо с постоянной интенсивностью с внутренним и внешним радиусом, равным соответственно 0.5 мм и 0.65 мм. Освещение ДОЭ выбрать гауссовым (спад гауссового пучка на краях должен составлять 90% от максимального значения, т.е. параметр “Cut Out Level” равен 0.1). Сравнить результаты, полученные при случайной начальной фазе и при фазе, полученной с помощью геометрической оптики.

2. Рассчитать с помощью пакета RadialDOE фазовый дифракционный элемент размерностью 1024x1024, формирующий в частотной плоскости (задней фокальной плоскости сферической линзы) кольцо с постоянной интенсивностью с внутренним и внешним радиусом, равным соответственно 0.5 мм и 0.8 мм. Начальную фазу считать полученной с помощью геометрической оптики. Сравнить результаты, полученные при освещении ДОЭ плоской волной, гауссовым пучком и супергауссовым пучком (спад гауссового пучка на краях варьировать).

3. Рассчитать с помощью пакета RadialDOE фазовый дифракционный элемент размерностью 256x256, формирующий в частотной плоскости (задней фокальной плоскости сферической линзы) кольцо с постоянной интенсивностью с внутренним и внешним радиусом, равным соответственно 0.5 мм и 0.8 мм. Начальную фазу считать полученной с помощью геометрической оптики. ДОЭ освещается ограниченной плоской волной, радиус апертуры 1 мм. Сравнить результаты, полученные при различных порядках функции Бесселя. Варьируя этот порядок, необходимо добиться уменьшения центрального пика.

4. Условия и цель те же, что и в предыдущей задаче, но вместо плоской волны использовать гауссов пучок (спад гауссового пучка на краях должен составлять 90% от максимального значения, т.е. параметр “Cut Out Level” равен 0.1).

5. Рассчитать с помощью пакета RadialDOE фазовый дифракционный элемент размерностью 256x256, формирующий в частотной плоскости (задней фокальной плоскости сферической линзы) кольцо с постоянной интенсивностью с внутренним и внешним радиусом, равным соответственно 0.5 мм и 0.8 мм. Начальную фазу считать полученной с помощью геометрической оптики. ДОЭ освещается ограниченной плоской волной, радиус апертуры 1 мм. Сравнить результаты, полученные при использовании различных алгоритмов: уменьшения ошибки, адаптивно-аддитивного, адаптивно-регуляризационного. Варьируя порядок использования алгоритмов, необходимо добиться наименьшей ошибки фокусировки. Сравнить с последовательностью от простых алгоритмов к сложным.

6. Условия и цель те же, что и в предыдущей задаче, но вместо плоской волны использовать гауссов пучок (спад гауссового пучка варьировать).

Литература

- [1] Методы компьютерной оптики //Под ред. В.А. Сойфера, М.: Физматлит, 2003.
- [2] Хонина С.Н. Методы расчета дифракционных оптических элементов на основе функционального представления // Учебное пособие, Самара, СГАУ, ИСОИ РАН, 2006, 160с.