



**МОНОГРАФИИ  
НГТУ**

Серия основана в 2001 году

**РЕДАКЦИОННАЯ КОЛЛЕГИЯ  
СЕРИИ «МОНОГРАФИИ НГТУ»**

д-р техн. наук, проф. (председатель) *А.А. Батаев*  
д-р техн. наук, проф. (зам. председателя) *А.Г. Вострецов*  
д-р техн. наук, проф. (отв. секретарь) *В.Н. Васюков*

д-р техн. наук, проф. *А.А. Воевода*  
д-р техн. наук, проф. *В.И. Денисов*  
д-р физ.-мат. наук, проф. *А.К. Дмитриев*  
д-р физ.-мат. наук, проф. *В.Г. Дубровский*  
д-р филос. наук, проф. *В.И. Игнатьев*  
д-р физ.-мат. наук, проф. *О.В. Кибис*  
д-р филос. наук, проф. *В.В. Крюков*  
д-р социол. наук, проф. *Л.А. Осьмук*  
д-р техн. наук, проф. *Н.В. Пустовой*  
д-р техн. наук, проф. *Г.И. Расторгуев*  
д-р физ.-мат. наук, проф. *В.А. Селезнев*  
д-р техн. наук, проф. *Ю.Г. Соловейчик*  
д-р техн. наук, проф. *А.А. Спектор*  
д-р техн. наук, проф. *А.Г. Фишов*  
д-р экон. наук, проф. *М.В. Хайруллина*  
д-р техн. наук, проф. *В.А. Хрусталеv*  
д-р техн. наук, проф. *А.Ф. Шевченко*

**В. И. ГУЖОВ**

# **КОМПЬЮТЕРНАЯ ГОЛОГРАФИЯ**



**НОВОСИБИРСК  
2018**

УДК 621.383.8:004.942 + 004.353.256

Г 936

Рецензенты:

д-р техн. наук, профессор *Ю.В. Чугуй*  
канд. техн. наук, доцент *Е.Н. Денежкин*  
д-р физ.-мат. наук, профессор *Л.А. Борыняк*

**Гужов В.И.**

Г 936      Компьютерная голография: монография / В.И. Гужов. – Новосибирск: Изд-во НГТУ, 2018. – 270 с. (Серия «Монографии НГТУ»).

ISBN 978-5-7782-3718-6

В монографии излагаются основы получения оптических голограмм, методы регистрации голограмм и цифровые алгоритмы восстановления волновых полей из голограмм; рассматриваются вопросы дискретизации голограмм с помощью обобщенных функций; показаны реализация процесса восстановления изображений при расшифровке полученных классическим способом голограмм и методы получения и расшифровки цифровых голограмм с помощью методов пошагового фазового сдвига. Компьютер при этом используется не только для обработки, но и для управления процессом получения голографических картин.

Монография предназначена для научных работников и студентов, специализирующихся в области оптики, интерферометрии и голографии, а также для широкого круга специалистов, в сферу интересов которых входит обработка изображений.

УДК 621.383.8:004.942 + 004.353.256

ISBN 978-5-7782-3718-6

© Гужов В.И., 2018

© Новосибирский государственный  
технический университет, 2018

V.I. GUZHOV

# COMPUTER HOLOGRAPHY



NOVOSIBIRSK  
2018

УДК 621.383.8:004.942 + 004.353.256

Г 936

Reviewers:

Prof. Yu.V. Chugui, D.Sc. (Eng.)

Assoc. Prof. E.N. Denezhkin, PhD (Eng.)

Prof. L.A. Borynyak, D.Sc. (Phys. & Math.)

**Guzhov V.I.**

Г 936 Computer holography: monograph / V.I. Guzhov. – Novosibirsk: NSTU Publisher, 2018. – 270 p. (“NSTU Monographs” series).

ISBN 978-5-7782-3718-6

The monograph states the principles of receiving optical holograms, methods of recording holograms and digital algorithms of reconstructing wave fields from holograms. It also addresses the issues of hologram sampling by means of generalized functions. The implementation of the process of image restoration when interpreting holograms received by the classical method and methods of receiving and interpreting digital holograms by means of step-by-step phase shift methods are also described. In so doing the computer is used not only for processing but also for managing the process of receiving holographic patterns.

The monograph is intended for researchers and students specializing in optics, interferometry and holography as well as for a wide range of specialists interested in image processing.

УДК 621.383.8:004.942 + 004.353.256

ISBN 978-5-7782-3718-6

© Guzhov V.I., 2018

© Novosibirsk State

Technical University, 2018

## ВВЕДЕНИЕ

Наука начинается с тех пор, как начинают измерять.  
Точная наука немислима без меры

*Д.И. Менделеев*

Нет ни одной области практической деятельности человека, где можно было бы обойтись без количественных оценок, получаемых в результате измерений. Исследование и разработка новых методов измерений и создание на основе этих методов измерительных систем является актуальной задачей современных исследований. Развитие вычислительной техники позволило создавать информационно-измерительные системы, которые выводят промышленный контроль на новый уровень, делая его более точным, быстрым и дешевым. Исследование и разработка новых методов неразрушающего контроля и создание на основе этих методов информационно-измерительных систем нового поколения является актуальной задачей современных исследований в области измерительной техники.

По характеру воздействия на объект различают разрушающий и неразрушающий контроль.

При разрушающем контроле изделия не сохраняют своего качества. Типичными примерами здесь могут служить испытания на сжатие и растяжение, на удар, на высокие температуры. Результатом разрушающего контроля не всегда будет являться физическое разрушение объекта, но возможны ухудшение или потеря каких-либо свойств, важных при дальнейшем использовании изделия.

Неразрушающий контроль позволяет определять соответствие различных агрегатов качественным и количественным показателям, не причиняя им ущерба в процессе измерений. При неразрушающем кон-



троле изделия, в которых нет дефектов, сохраняют свои качества и свойства, несмотря на то что на них было оказано определенное воздействие, т. е. такие изделия можно использовать в дальнейшем.

Можно выделить девять видов неразрушающего контроля: 1) магнитный, 2) электрический, 3) вихретоковый, 4) радиоволновой, 5) тепловой, 6) оптический, 7) радиационный, 8) акустический, 9) контроль проникающими веществами [1]. Среди множества методов неразрушающего контроля стоит выделить ряд методов, получивших название оптических. Современные методы, основанные на использовании свойств волновых полей, позволяют проводить измерения с точностью до сотых и тысячных долей длины волны.

Методы оптического контроля появились и эффективно использовались задолго до появления других методов. Человеческий глаз способен без каких-либо приспособлений обнаруживать ряд явных дефектов, а с использованием оптических приборов еще большее их количество. В классической оптике изображение формируется с помощью оптических устройств, состоящих из набора линз и зеркал, а регистрируется с помощью датчиков, реагирующих на интенсивность оптического излучения. При помощи оптических методов можно оценивать поверхностные нарушения объектов (пустоты, расслоения, поры, трещины, инородные включения для оптически прозрачных материалов), геометрические параметры (шероховатость, толщина пленок и др.), физико-химические свойства (внутренние напряжения, структура материала). Эти методы применяются в настоящее время и будут использоваться в будущем.

Большое количество оптических методов измерений образовалось благодаря широкому спектру исследуемых объектов. Для оптически грубых поверхностей хорошо себя показывают проекционные методы и эффект Муара [2, 3], так как они менее чувствительны к среде и диапазон их измерений существенно больше интерференционных методов.

Однако если поверхность объекта является оптически гладкой, то нужны более точные методы, основанные на интерференции световых волн [4]. Основное достоинство интерференционных методов – высо-



кая точность измерений. Это вызвано тем, что в качестве меры используется длина волны света, используемого для освещения. В случае когерентного, монохроматического освещения длина волны является физической константой.

Световой (в общем случае электромагнитный) пучок разделяется на два или большее количество когерентных пучков. Пучки проходят разные оптические пути, попадают на фотоприемник и интерферируют. По полученной интерференционной картине можно установить фазовое смещение пучков. Разность оптического хода, создаваемая неровностями поверхности исследуемого объекта, позволяет по полученному смещению фаз установить физическую величину этой неровности.

Классическая оптическая интерферометрия имеет дело с прозрачными или хорошо отражающими (зеркальными) поверхностями. Для прозрачных объектов применение интерферометрических методов диагностики допустимо, если максимальная рефракция на объекте не превышает 0,1 мрад, а влиянием поглощения можно пренебречь. Такие прозрачные объекты получили название фазовых, т. е. изменяющих только фазу зондирующего излучения [5].

С появлением голографии появилась возможность проведения интерферометрических измерений не только прозрачных и зеркальных (т. е. объектов с оптически чистой поверхностью), но и диффузно отражающих объектов. Голография – это метод регистрации произвольного колебательного процесса, позволяющий зафиксировать как амплитуду, так и фазу колебания, а затем воспроизвести их в любой удобный момент времени. Так как полная информация о форме объекта содержится в оптическом поле, которое рассеивается исследуемым объектом при его освещении, голографический процесс позволяет регистрировать эту форму в неизменном виде. Сохраненную таким образом форму можно в любой момент времени воспроизвести и использовать в качестве шаблона. На этом принципе основано подавляющее большинство методов оптического неразрушающего контроля, использующих голографию.

Голографическая интерферометрия является одним из наиболее перспективных методов для контроля напряженно-деформированного



состояния объектов. Голографическая интерферометрия – получение и интерпретация интерференционных картин, образующихся при нескольких состояниях объекта, при этом одно или несколько состояний объекта записываются и восстанавливаются с помощью голографии. По результатам анализа голографических интерферограмм можно определить разницу между двумя состояниями объекта, а значит, построить модель изменений, произошедших с объектом. Голографическая интерферометрия существенно расширила функциональные возможности оптической интерферометрии, позволив проводить измерения диффузно отражающих объектов. Методы голографической интерферометрии оказались пригодными для исследования промышленных деталей и конструкций в заводских условиях [6–13].

Однако развитие этих методов ограничивается необходимостью применения фотоматериалов для регистрации промежуточных голограмм. Использование голографических фотопластинок и пленок связано со сложными химическими процессами при обработке. Поэтому измерительные системы на основе этих методов используются в основном в лабораторных условиях. Новый виток в развитии голографических измерительных систем появился в результате развития методов цифровой голографии.

Цифровая голография – это анализ, синтез и моделирование волновых полей средствами вычислительной техники. Цифровая голография существенно повышает производительность и практичность методов голографии (в частности, методов голографической и спекл-интерферометрии, голографической микроскопии).

Первые работы по генерации цифровых голограмм появились в 1967–1980 гг. Ряд работ был выполнен А. Ломаном и В. Ли за рубежом, а также Б.Ф. Федоровым и Р.И. Эльманом в нашей стране. За работами советских авторов внимательно следил Ю. Денисюк и периодически предоставлял исследователям возможность выступать на научных семинарах, руководителем которых он являлся. Он же предложил опубликовать монографию с результатами работ в издательстве «Наука», к которой написал предисловие. Она вышла в 1976 г. и была первой книгой, в которой на основе фактического материала излага-



лись методы синтезирования голограмм простейших объектов и восстановление изображений [14].

Большим вкладом в развитие цифровой голографии было открытие способа непосредственной записи голограмм с помощью фотоматриц, предложенного У. Шнарсом и В. Юптнером в 1994 г. [16]. Как отмечают У. Шнарс и В. Юптнер в своей книге [16], усовершенствование методов голографической интерферометрии, а также электронной спекл-интерферометрии, связанное с избавлением от сложных химических работ по обработке голографических пластин и пленок, подтолкнуло их к цифровой голографии. Этот способ позволил избавиться от каких-либо промежуточных шагов и подготовительных работ в процессе записи голограммы. На сегодняшний день этот метод часто ассоциируют с самим понятием цифровой голографии, хотя изначально метод получил название «прямая голография».

Р. Лоренц и Д. Гудман впервые предложили цифровой метод восстановления еще в 1967 г. [17]. Большой вклад в эту область внесли российские ученые Л. Ярославский и Н. Мерзляков. В своей книге [18] они дают подробную теоретическую основу преобразования, представления оптических волновых полей, а также описывают алгоритмы их обработки, основанные на дискретном преобразовании Фурье. Их идеи были развиты Л. Онуралом и П. Скоттом и перенесены на область измерения частиц в 1987 г. [19], а в 1992 г. В. Хаддад вместе со своими коллегами предложил схему голографического микроскопа, основанного на численном восстановлении [20].

С совершенствованием вычислительной техники методы оптического контроля получили новый толчок в своем развитии. Цифровая голография стала возможной с появлением устройств ввода оптических изображений с достаточной пространственной разрешающей способностью и появлением компьютеров, обеспечивающих приемлемую скорость вычислений. Основные принципы цифровой голографии такие же, как и в классической голографии. По-прежнему весь процесс разбивается на два этапа: фазу записи и фазу восстановления. Однако в цифровой голографии при записи роль фоточувствительной среды выполняет фоточувствительная матрица. Данные, полученные с матрицы



фотоприемников, передаются в компьютер, где сохраняются в цифровом виде. В процессе восстановления эта цифровая голограмма численно восстанавливается при помощи виртуальной опорной волны, которая вычисляется математически и по характеристикам совпадает с реальной опорной волной, использованной при записи. Так как процедура восстановления выполняется на компьютере, скорость работы зависит исключительно от характеристик алгоритмов и вычислительной мощности устройства.

Процесс переноса методов классической голографии на цифровую голографию развивался по двум направлениям: генерация искусственных голограмм с последующим оптическим восстановлением и численное восстановление оптически сформированных голограмм. Термин «цифровая голография» во многих работах связан именно с получением искусственных голограмм.

В этой книге рассматриваются вопросы получения оптических голограмм, методы регистрации голограмм и цифровые алгоритмы восстановления волновых полей из полученных обычными оптическими методами голографических изображений. Компьютер в голографических измерительных системах используется не только для обработки цифровых данных, но и для управления процессом получения голографических картин. Для этого задаются управляющие воздействия на различные элементы оптической схемы.

Для того чтобы убрать разночтения при использовании термина «цифровая голография», далее мы будем использовать другое выражение – «компьютерная голография».

## БИБЛИОГРАФИЧЕСКИЙ СПИСОК К ВВЕДЕНИЮ

1. ГОСТ 18353–79. Контроль неразрушающий. Классификация видов и методов. – Москва : Стандартинформ, 2004. – 12 с.
2. *Гужов В. И.* Методы измерения 3D профиля объектов. Контактные, триангуляционные системы и методы структурированного освещения : учебное пособие / В. И. Гужов. – Новосибирск : Изд-во НГТУ, 2015. – 82 с.
3. *Гужов В. И.* Методы измерения 3D профиля объектов. Фазовые методы : учебное пособие / В. И. Гужов. – Новосибирск : Изд-во НГТУ, 2016. – 83 с.



4. *Гужов В. И.* Оптические измерения. Компьютерная интерферометрия / В. И. Гужов, С. П. Ильиных. – 2-е изд. – Москва : Юрайт, 2018. – 258 с. – ISBN 978-5-534-06855-9.

5. *Шарп Р. С.* Методы неразрушающих испытаний : физические основы, практические применения, перспективы развития / Р. С. Шарп. – Москва : Мир, 1972. – 495 с.

6. *Александров Е. Б.* Исследование поверхностных деформаций с помощью голограммной техники / Е. Б. Александров, А. М. Бонч-Бруевич // ЖТФ. – 1967. – Т. 37, вып. 2. – С. 360–365.

7. *Ennos A. E.* Measurement of in-plane surface strain by hologram interferometry / A. E. Ennos // Journal of Physics E: Scientific Instruments. – 1968. – Vol. 58, N 1. – P. 731–734.

8. *Sollid J. D.* Holographic interferometry applied to measurements of small static displacements of diffusely reflecting surfaces / J. D. Sollid // Applied Optics. – 1969. – Vol. 8. – P. 1587–1595.

9. *Козачок А. Г.* Голографические методы исследования в экспериментальной механике / А. Г. Козачок. – Москва : Машиностроение, 1984. – 176 с.

10. *Островская Г. В.* Голографическая интерферометрия физических процессов / Г. В. Островская // Журнал технической физики. – 2016. – Т. 86, вып. 6. – С. 1–16.

11. *Островский Ю. И.* Голографическая интерферометрия / Ю. И. Островский, М. М. Бутусов, Г. В. Островская. – Москва : Наука, 1977. – 339 с.

12. *Вест Ч.* Голографическая интерферометрия / Ч. Вест. – Москва : Мир, 1982. – 503 с.

13. *Stetson K. A.* A brief history of holographic interferometry / K. A. Stetson // Proceedings of Frontiers in Optics : OSA technical digest / Optical Society of America. – Rochester ; New York, 2006.

14. *Федоров Б. Ф.* Цифровая голография / Б. Ф. Федоров, Р. И. Эльтман. – Москва : Наука, 1976.

15. *Schnars U.* Direct recording of holograms by a CCD-target and numerical reconstruction / U. Schnars, W. Jueptner // Applied Optics. – 1994. – Vol. 33 (2). – P. 179–181.

16. *Schnars U.* Digital holography: digital hologram recording, numerical reconstruction, and related techniques / U. Schnars, W. Jueptner. – Berlin : Springer-Verlag, 2005. – 164 p.

17. *Goodman J. W.* Digital image formation from electronically detected holograms / J. W. Goodman, R. W. Lawrence // Applied Physics Letters. – 1967. – Vol. 11. – P. 77–79.



- 
18. Ярославский Л. П. Цифровая голография / Л. П. Ярославский, Н. С. Мерзляков. – Москва : Наука, 1982. – 219 с.
  19. Onural L. Digital decoding of in-line holograms / L. Onural, P. D. Scott // Optical Engineering. – 1987. – Vol. 26 (11). – P. 1124–1132.
  20. Fourier-transform holographic microscope / W. Haddad, D. Cullen, J. Solem, J. Longworth, A. McPherson, K. Boyer, K. Rhodes // Applied Optics. – 1992. – Vol. 31 (24). – P. 4973–4978.

---

---

# ГЛАВА 1

## ОСНОВЫ ГОЛОГРАФИИ

---

---

Точные измерения стоят дорого,  
а неточные обходятся значительно дороже

*Генри Форд*

Голография – это один из способов регистрации информации, с помощью которой можно записывать, а потом воспроизводить не только амплитуду, но и фазу световой волны. Существует множество способов получения голографических изображений, но основные принципы их образования одинаковы. В этой главе кратко рассмотрена история возникновения и основные принципы оптических схем для получения и расшифровки голограмм, необходимые для понимания математических алгоритмов, изложенных в этой книге.

*Голография* – это процесс регистрации интерференционной картины (голограммы), которая образована волной, отраженной некоторым объектом, освещаемым источником света (объектная волна), и когерентной с ней волной, идущей непосредственно от источника света (опорная волна). Голограмма используется для восстановления объемного изображения.

Процесс голографии имеет сходство с процессом фотографии. И в том и в другом случае используется светочувствительная среда, регистрирующая интенсивность светового поля. Однако есть ключевое различие: на фотографии регистрируется только интенсивность объектной волны, а на голограмме – информация о полной объектной волне как результат интерференции с другой опорной волной. В записанной интерференционной картине содержится информация как об интенсивности, так и о фазе объектной волны, что невозможно в фотографии.

Появление голографии связывают с именем Д. Габора. Именно он впервые ввел термин «голография» в 1948 г. (от греч. *holos* – «весь,



полный» и *grapho* – записывать). Таким образом, термин «голография» переводится как «полная запись».



*Денеш Габор родился в еврейской семье в Будапеште 5 июня 1900 г. Д. Габор получил специальность инженера в 1920 г. в Техническом университете Будапешта. В 1921–1924 гг. продолжил обучение в Техническом университете Берлина. После защиты диссертации в 1927 г. начал работу в компании Siemens&Halske AG (сегодня компания Siemens).*

*В 1933 г. покинул Германию в связи со становлением режима нацистов и выехал на работу в Великобританию в компанию Томсона–Хоустона (Thomson–Houston).*

Принципиальная особенность нового метода состоит в том, что голография использует волновые свойства света, а не корпускулярные, как в традиционной фотографии. Для сохранения волнового фронта нужно регистрировать не только интенсивность, но и фазу отражаемых объектами волн. Д. Габор придумал способ, как это можно сделать, добавив так называемые эталонные волны для регистрации фазы. Благодаря наложению «предметной волны» от объекта с эталонной или «опорной волной» возникает интерференционная картина, которая и фиксируется на фоточувствительной пластине.

Идея голографического метода возникла в процессе усовершенствования электронной микроскопии, как отмечал Д. Габор, «двухступенчатого процесса, в котором предмет регистрируется при помощи пучка электронов, а восстанавливается при помощи светового пучка» [1]. В электронном микроскопе роль света выполняют электроны. Разреша-

ющая способность электронного микроскопа намного выше, чем оптического, поскольку длина волны быстрых электронов меньше длины световой волны. Однако на практике разрешающая способность первых электронных микроскопов оказывалась значительно ниже расчетной, и виной тому было несовершенство электронных линз, призванных фокусировать пучки быстрых частиц. В электронном микроскопе электронная линза исполняет ту же роль, что и оптическая в обычном: фокусирует «лучи» электронов. Д. Габор столкнулся с необходимостью улучшить качество изображения, которое сильно искажалось сферической аберрацией электронных линз. На первый взгляд эта трудность не кажется значительной: известно, что сферическая аберрация обычных линз исправляется достаточно просто. Однако в электронной оптике действуют несколько иные законы, и оказалось, что сферическую аберрацию в этом случае невозможно исправить по принципиальным соображениям.

Для того чтобы преодолеть возникшую трудность, Д. Габор предложил построить точную модель поля электронных волн, а затем исправить у этой модели сферическую аберрацию методами обычной световой оптики. Результаты своей работы Д. Габор опубликовал в журнале *Nature* в мае 1948 г. в коротком письме, озаглавленном «Новый принцип микроскопии» [1].

Идея Д. Габора состояла в том, чтобы, используя имеющийся в его распоряжении когерентный источник электронов, записать голограмму микроскопического объекта. Если затем эту голограмму в процессе восстановления осветить когерентными световыми волнами, то теоретически должно получиться значительно увеличенное объемное изображение данного микроскопического объекта. Увеличение обусловлено тем, что голограмма создается волнами (электронными), длина которых намного меньше длины световых волн, используемых для ее восстановления. Д. Габор надеялся, что подобный процесс позволит сделать видимыми более мелкие детали объекта, чем это прежде удавалось при помощи электронного микроскопа. К сожалению, проблема упиралась в создание когерентного источника электронов, и на самом деле качество такого микроскопа никогда не достигало теоретически возможного уровня.

На рис. 1.1 показана схема записи голограммы Габора. По этой схеме образец освещается сфокусированным электронным пучком света, на фотографической пластине записывается результат интерференции между первичной и вторичной волнами от объекта. Можно показать, что, по крайней мере, во внешних частях пластины интерферен-



ционные максимумы будут возникать очень близко к месту, где фазы первичной волны и вторичной волны совпали. Если освещать фотографическую запись оптическим подобием первичной волны, возникнет оптический фронт, который напоминает вторичную волну.

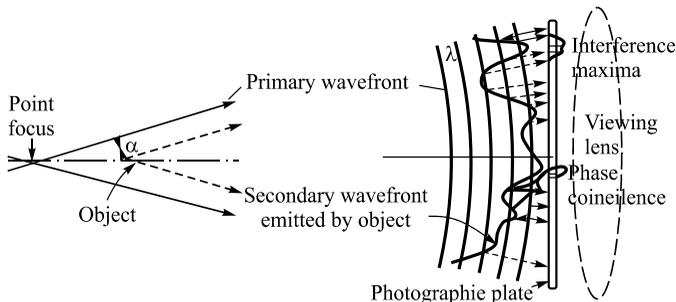


Рис. 1.1. Схема записи голограммы Габора [1]

Этот принцип был проверен Д. Габором с помощью обычной оптической системы. В оригинальной оптической установке, в которой моделировались результаты интерференции электронных пучков, опорная и объектная волны расположены вдоль оси, нормальной к фотографическая пластинке (рис. 1.2).

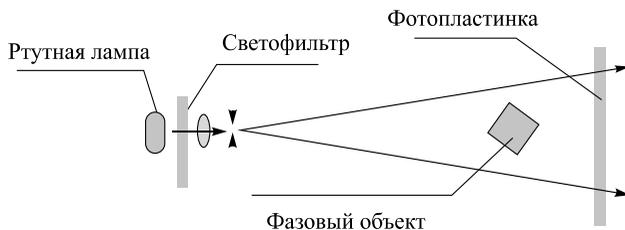


Рис. 1.2. Оптическая схема записи голограмм Габора

На рис. 1.3 показана схема получения и восстановления волнового фронта.

Как видно из рис. 1.3, б, при восстановлении возникает наложение волновых фронтов 0 и  $\pm 1$  порядков дифракции, поэтому голограммы отличались очень низким качеством.

Д. Габор писал о своем открытии, что «электронная микроскопия так до сих пор и не извлекла существенной пользы из восстановления волн, тогда как мои оптические опыты (которые были задуманы как модельные) положили начало голографии».

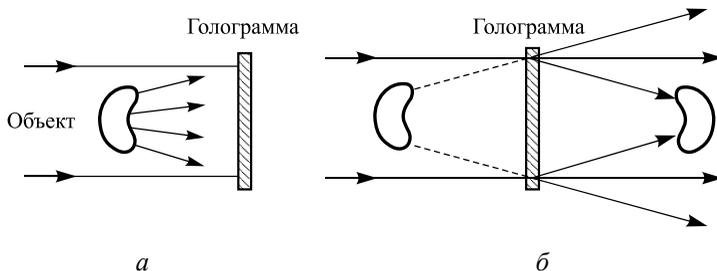


Рис. 1.3. Схема записи и восстановления голограмм:

*a* – получение голограммы; *б* – восстановление волнового фронта по Д. Габору

Однако были и более ранние работы в этой области [2]. Идея голографического метода получения изображения была выдвинута и экспериментально проверена польским физиком М. Вольфке и опубликована в 1920 г., за 28 лет до работы Д. Габор. Вольфке сформулировал принцип двухступенчатого восстановления изображения объекта, при котором используется дифракционная картина, образуемая этим объектом. Свою идею М. Вольфке сформулировал применительно к рентгеновским лучам и проверил на опыте в видимом свете.



*М. Вольфке (1883–1947) – польский физик, член Польской АН (1932).*

*Учился в 1901–1907 гг. в Льежском и Парижском университетах. В 1907–1910 гг. работал во Вроцлавском университете, в 1914–1922 гг. –*



в политехникуме и университете в Цюрихе, в 1922–1939 гг. – профессор Варшавского политехнического института, организатор Института низких температур (Варшава).

Будучи еще 15-летним гимназистом, в 1898 г. послал в Петербург патентную заявку на изобретение «телектроскопа», где впервые предложил передавать сигнал телевизионного изображения не по проводам, а по радио. Основные работы по физике низких температур и оптике (в частности, по теории дифракционного изображения).

В 1911 г. сконструировал новую ртутно-кадмиевую лампу, в 1927 г. совместно с В. Кеезом открыл две разновидности жидкого гелия (гелий I и гелий II), исследовал свойства жидкого гелия, жидкого и твердого водорода.

В 1920 г. в работе «О возможности оптического изображения молекулярной решетки» выдвинул и экспериментально проверил идею голографического метода получения изображений.

В своей статье «О возможности оптического изображения молекулярной решетки»\* он рассматривал возможность получения оптического изображения кристаллической решетки, используя дифракционные картины, полученные при прохождении рентгеновских лучей через кристалл. Для этого ученый предлагал использовать первичную рентгенограмму как дифракционную решетку для световых волн и сформулировал теорему: «При монохроматическом, параллельном, перпендикулярном освещении дифракционное поле дифракционной картины симметричного объекта без фазовой структуры тождественно изображению этого объекта». Кроме того, М. Вольфке экспериментально подтвердил свою теорему на различных структурах, помещенных в параллельных лучах желтой спектральной линии ртути. Эти исследования опирались на труды знаменитого оптика Э. Аббе и его теорию микроскопа. Работа М. Вольфке не нашла отклика среди физиков и была забыта.

М. Вольфке указывал, что к идее восстановления изображения подходил в 1913 г. и Э. Хупка при исследовании отражения рентгеновских лучей. Не зная о предшественниках, подобную идею высказал и немецкий физик Х. Берш. Спустя 18 лет он выступил с предложением о создании метода вторичного изображения, не упоминая работ М. Вольфке. Х. Берш предложил способ получения изображения ре-

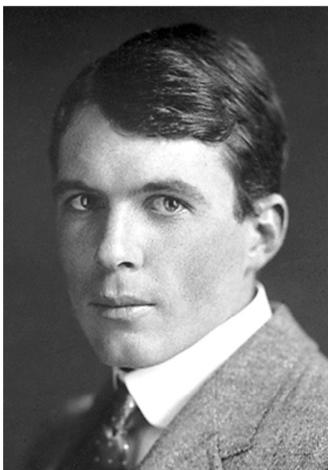
---

\* *Wolfke M. Über die Möglichkeit der optischen Abbildung von Molekulargittern / M. Wolfke // Physikische Zeitschrift. – 1920. – No 5(2). – P. 495–497.*



шетки в микроскопе без помещения ее на предметный столик. Для этого необходимо было создать в задней фокальной плоскости объектива микроскопа распределение светового потока, соответствующее дифракционной картине, полученной с помощью решетки.

Д. Габор тоже ничего не знал о трудах М. Вольфке, однако ссылаясь на работу Г. Берша «К образованию изображения в микроскопе», изданную в 1938 г. [H. Voersch, Z. techn. Phys. 19, 337 (1938)]. Как отмечает сам Д. Габор, общая идея метода голографии в электронной микроскопии как двухступенчатого процесса, в котором «предмет регистрируется с помощью пучка электронов, а восстанавливается с помощью светового пучка», возникла как модификация идеи У.Л. Брэгга, изложенного в 1942 г. в статье «Рентгеновский микроскоп» [W.L. Bragg, Nature 149 (3782), 470 (1942)], в которой предлагался метод визуализации кристаллической решетки с помощью процесса дифракции на дифракционной картине, полученной в рентгеновских лучах.



*Сэр Уильям Лоренс Брэгг (1890–1971) – австралийский физик, лауреат Нобелевской премии по физике за 1915 г. Самый молодой нобелевский лауреат по физике за всю историю премии (ему было 25 лет).*

*Родился в Австралии, в 14 лет поступил в университет. Окончил его в 1908 г. В том же году переехал в Англию. Наибольшую известность У.Л. Брэггу принес его закон о дифракции рентгеновских лучей на кристаллах. Закон Брэгга позволяет рассчитать положение атомов в кристалле по дифракционной картине, которую образуют рентгеновские лучи, проходя*



*сквозь кристаллическую решетку. Он сделал это открытие в 1912 г., в течение первого года обучения в качестве студента-исследователя в Кембридже.*

*В промежутке между войнами с 1919 по 1937 г. работал в Университете Манчестера в качестве профессора физики. С 1939 по 1943 г. – президент Института физики. Посвящен в рыцари в 1941 г.*

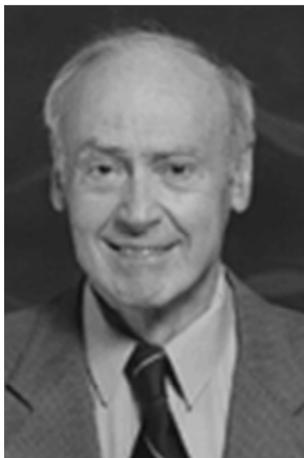
*После Второй мировой войны У.Л. Брэгг вернулся в Кембридж, разделив при этом Кавендишскую лабораторию на исследовательские группы. Он считал, что идеальной исследовательской единицей является группа от одного до 6–12 ученых и нескольких ассистентов.*

Еще раньше, в 1939 г., эта идея была сформулирована У.Л. Брэггом в статье «Новый тип рентгеновского микроскопа» [W.L. Bragg, Nature 143 (3625), 678 (1939)].

Новизна метода Д. Габора состоит в том, что он предложил к объектной волне добавлять опорную волну для регистрации не только амплитуды, но и фазы объектной волны. Однако голограммы, полученные Д. Габором с помощью ртутной лампы, хотя и подтверждали в целом концепцию, однако имели чрезвычайно низкое качество и для реального применения не годились. Несколько лет оживленных экспериментов вокруг явно перспективной идеи закончились ничем, и к 1955 г. интерес к оптической голографии практически пропал.

Однако после создания в 1960 г. импульсного рубинового и непрерывного гелий-неонового лазеров голография начала интенсивно развиваться. Первое применение лазеров в голографии связано с именами Э. Лейта и Ю. Упатниекса из Мичиганского технологического института. Э. Лейт и Ю. Упатниекс впервые создали голографические изображения, имеющие объем и параллакс, недостижимые ранее. При этом, в отличие от осевой схемы Д. Габора, Э. Лейт и Ю. Упатниекс использовали так называемую внеосевую схему записи голограмм, что позволило избавиться от проблемы наложения мнимого и действительного изображений. В 1961–1962 гг. они опубликовали идею и теорию голограмм с боковым опорным пучком [3].

Благодаря использованию лазера, обладающего несравнимо более высокой степенью пространственно-временной когерентности излучения и одновременно с этим высокой мощностью, а также введению в схему записи наклонного опорного пучка, широко используемого в классической интерферометрии, Э. Лейт и Ю. Упатниекс сумели получить объемные изображения ряда сцен.



*Эммет Лейт (12 марта 1927, Детройт – 23 декабря 2005, Энн-Арбор) – профессор электротехники в Университете Мичигана. Лейт получил степень бакалавра наук по физике в Государственном университете Уэйна в 1949 г. и степень магистра по физике в 1952 г. Докторскую степень в области электротехники получил в Государственном университете Уэйна в 1978 г.*

*Большая часть голографической работы Э. Лейта была следствием его исследования по радиолокационному синтезированию апертуры (РСА) во время работы в Radar Laboratory of the University of Willow Run Laboratory Michigan's.*

*В начале 1952 г. Э. Лейт присоединился к Университету Мичигана в качестве младшего научного сотрудника, стал научным сотрудником в 1956 г., инженером-исследователем в 1960 г., доцентом в 1965 г. и профессором в 1968 г.*

Успех этой группы был обусловлен не только появлением лазера, но также многолетней работой Э. Лейта над военным проектом разработки радиолокатора с когерентной антенной.

Длина электромагнитных волн, используемых в радиолокации, примерно в 100 000 раз больше длины волны света. Уже в 1950-е гг. на основе голографического принципа в США сумели успешно создать радиолокатор с синтезированной (когерентной) антенной, размещаемый на борту самолета и дающий качественные снимки местности в условиях плохой оптической видимости. По сути дела новая технология «боковой радиолокации» представляла собой двумерную голографию для микроволнового диапазона частот, аналогичную электронной голографии Д. Габора.



*Юрис Упатниекс (родился 7 мая 1936 г. в Риге). Американский физик латвийского происхождения. В конце Второй мировой войны семья переехала в Германию. В 1951 г. семья эмигрировала в Соединенные Штаты. Он учился в средней школе в городе Акрон, штат Огайо, и изучал электротехнику в Университете Акрона, где получил степень бакалавра в 1960 г., после этого он учился в Институте наук и технологий Университета Мичигана, где получил степень магистра в области электротехники в 1965 г.*

*Ю. Упатниекс имеет 19 патентов, среди них патент на голографический прицел (Juris Upatnieks, Ann Arbor, Mich. Holographic light line sight, Patent USA № 4012150 of Mar. 15, 1977).*

Через два года Э. Лейт и Ю. Упатниекс реализовали свою схему в экспериментах по записи и воспроизведению первых внеосевых голограмм. Сначала это были голограммы плоских пропускающих транспарантов с диффузной подсветкой.

С появлением газовых лазеров Э. Лейт и Ю. Упатниекс отказались от получения голограмм с транспарантов и стали работать с трехмерными объектами. Эти голограммы были совершенны, а изображения, восстанавливаемые в них, настолько точно передавали наблюдаемый предмет, что у зрителя возникал эффект присутствия. Высокое качество и натуральность трехмерных изображений вызвали всеобщий интерес к идее восстановления волнового фронта. Голографию стали воспринимать как реальный метод. Началось триумфальное шествие голографии по всему миру.



Э. Лейт и Ю. Упатниекс решили проблему перекрытия пучков, поставившую Д. Габора в трудное положение, следующим образом. Во-первых, пучок света, образующий когерентный фон, они направили на фотопластинку под довольно большим углом с пучком света, дифрагированным объектом. Это позволило получить изображения, которые при наблюдении не перекрываются. Во-вторых, в качестве источника излучения они использовали лазер, высокую интенсивность излучения и когерентность которого обеспечивают высокое качество восстановленного изображения.

Схема записи голограмм, предложенная этими учеными, теперь используется повсеместно (рис. 1.4).

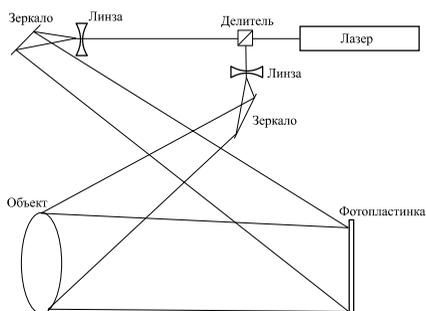


Рис. 1.4. Эммет Лейт и Юрис Упатниекс со своей оптической установкой

На рис. 1.5 показана схема получения и восстановления волнового фронта в схеме Лейта–Упатниекса. Из рисунка видно, что при определенном выборе угла между интерферирующими пучками можно преодолеть проблему перекрытия 0 и  $\pm 1$  порядков дифракции.

Значительный вклад в развитие голографии внес сотрудник Государственного оптического института им. С.И. Вавилова Ю.Н. Денисюк. Он использовал в качестве опорного пучка встречный. В результате полосы были локализованы не на плоскости фотопластинки, а в объеме фотослоя. Это усовершенствование позволило получать высококачественные изображения (монохромные и цветные) при освещении голограмм некогерентным белым светом, что значительно расширило область применения таких голограмм в изобразительных целях. На рис. 1.6, а приведена схема записи трехмерной голограммы.

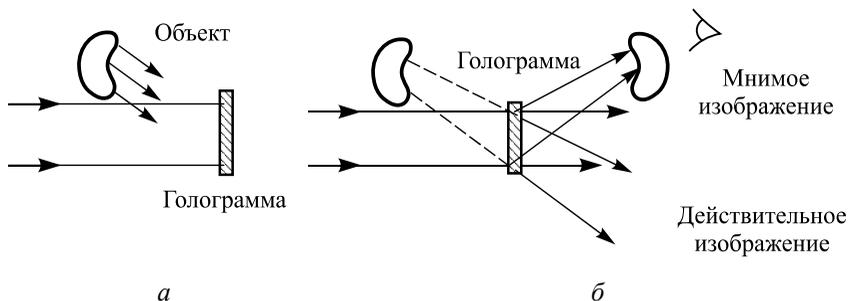


Рис. 1.5. Получение голограммы и восстановление волнового фронта по Лейту–Упатниексу:

*a* – регистрация голограммы; *б* – восстановление действительного и мнимого изображений

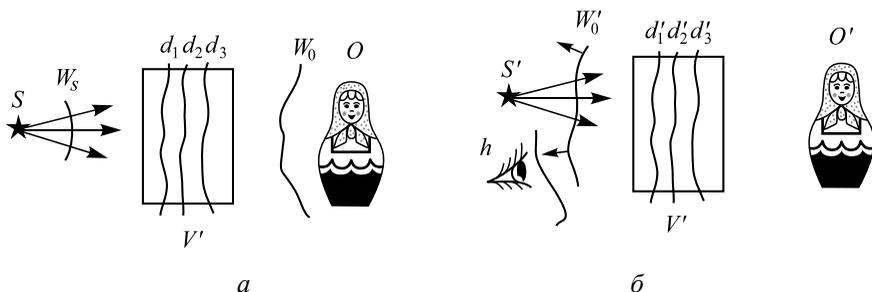


Рис. 1.6. Схема записи (*a*) и восстановления (*б*) трехмерной голограммы

Монохроматический свет источника  $S$  падает на произвольный объект  $O$ . Свет, рассеянный объектом (волна  $W_0$ ), интерферирует с падающим светом (волна  $W_s$ ), в результате чего возникает стоячая волна. Поверхности пучностей этой стоячей волны (места, в которых интенсивность света максимальна) обозначены  $d_1, d_2, d_3$ . В поле стоячей волны помещен некоторый объем  $V'$ , заполненный прозрачным светочувствительным материалом. После экспозиции и последующей химической обработки в объеме  $V'$  возникает материальная модель стоячей волны, при этом каждая зарегистрированная голограммой поверхность пучностей превращается в своеобразное кривое зеркало.

Процесс реконструкции трехмерной голограммы показан на рис. 1.6, б. Сферическая волна от обычного источника белого света  $S'$

падает на голограмму  $V'$ . Каждое образовавшееся на месте поверхности пучностей сферическое зеркало, отражая эту волну, трансформирует ее в объектную волну  $W_0'$ . Это означает, что трехмерная голограмма выбирает из сплошного спектра и отражает назад только те монохроматические компоненты, которые были на ней зарегистрированы. В результате всех этих процессов наблюдатель  $h$ , воспринимающий волны излучения, реконструированного голограммой, видит единственное цветное трехмерное изображение объекта, которое невозможно отличить от реального [4].

Принципиальное отличие метода Денисюка состояло в том, что голограмма восстанавливалась обычным белым светом, при этом, используя методы, аналогичные применяемым в цветной фотографии (производя съемку лазерами трех цветов – синим, зеленым и красным), голографические изображения можно было получить цветными, что было недостижимо для всех используемых ранее методов. Абсолютно иной подход к получению голограмм Денисюка объяснялся тем, что он просто ничего не знал о подобных разработках за рубежом и развивал идеи волновой фотографии самостоятельно. В основу оптической схемы были положены интересные, но в ту пору большинством давно уже забытые идеи французского физика Г. Липпмана.

Г. Липпман разработал оригинальный метод цветной фотографии на основе обычной фотоэмульсии черно-белых снимков и грамотного использования интерференции света. За свою технологию Липпман в 1908 г. был удостоен Нобелевской премии, однако последующее развитие фотографии пошло по направлению более практичных многоцветных эмульсий.

Главная находка Г. Липпмана состояла в том, чтобы использовать не плоский двумерный, а объемный или, как еще говорят, толстый слой фотоэмульсии. Благодаря этому фотография может сначала регистрировать, а затем и воспроизводить не только информацию об интенсивности, но и о спектральном составе волн света. Иначе говоря, передавать не только степень яркости каждого из участков картинка, но и их цветовую окраску. Достигал этого Г. Липпман тем, что с помощью ртути подложку каждой фотографии делал зеркальной, из-за этого в объемном слое фотоэмульсии фиксировалась (в виде параллельных слоев разной толщины) интерференция волн, рассеянных объектом и отраженных зеркалом подложки.



*Габриэль Ионас Липпман (фр. Gabriel Lippmann; 16 августа 1845 г. – 3 июля 1921 г.) – французский физик, лауреат Нобелевской премии по физике в 1908 г. за создание метода фотографического воспроизведения цветов на основе явления интерференции.*

*Работы Г. Липпмана по фотографии в настоящее время не используются из-за технической сложности реализации предложенного им процесса. В то же время эти работы получили свое развитие при создании голографии.*

Г. Липпман сумел объединить явление интерференции света с возможностями черно-белой галоген-серебряной фотографии. Изображение в камере Липпмана фокусировалось на стеклянную фотопластинку, положенную эмульсией на поверхность ртути, играющей роль зеркала (рис. 1.7).

Свет сначала проходил через фотоэмульсию сверху вниз, а затем, отражаясь от ртути, образовывал в фотоэмульсии стоячую волну, период которой зависел от длины волны света. Так, на рис. 1.7, а слева синие лучи образовали стоячую волну, расстояние между нулями (узлами) которой меньше, чем соответствующее расстояние для красных лучей справа. Разумеется, в узлах стоячей световой волны скрытое изображение не образовывалось, а в местах максимума стоячей волны (пучностях) происходила максимальная засветка, превращающаяся после проявления в систему параллельных полос металлического серебра с промежутками между полосами, соответствующими волне того цвета, который облучал в данном месте фотоэмульсию.

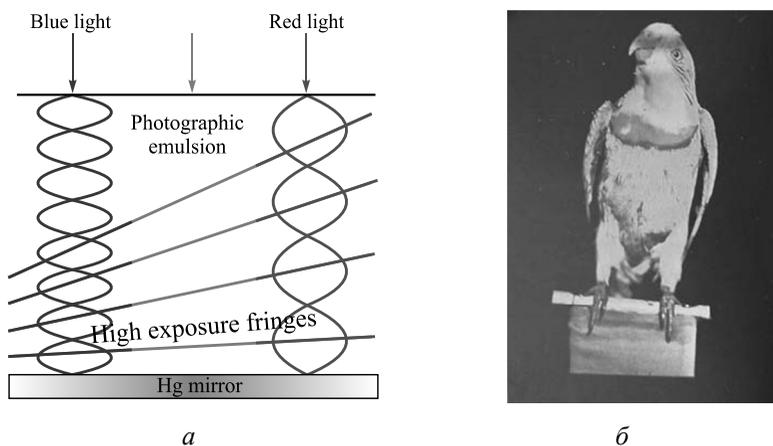


Рис. 1.7. Получение цветных фотографий Г. Липпмана:

*a* – схема записи цветных фотографий Липпмана; *б* – оригинальная цветная фотография Липпмана

При разглядывании снимка в белом свете происходит отражение световых волн от полос, при этом конструктивная интерференция возникает только для волн тех цветов, которые имеют ровно ту же самую длину волны, что и волны, вызвавшие образование стоячей волны при фотографировании. В итоге возникает сочное цветное изображение, представление о котором дает воспроизведение одной из оригинальных цветных фотографий Г. Липпмана.

Ю.Н. Денисюк, в свою очередь, предположил, что если вместо отраженных зеркалом волн использовать более регулярный монохромный фон, то, быть может, удастся зафиксировать не только амплитуду и спектральный состав, но еще и фазовые сдвиги волн. Фактически цветная фотография по Г. Липпману – это получение голограммы Ю.Н. Денисюка для плоского объекта, каковым является картинка, фокусируемая на поверхности ртути. Последнее становится очевидным, если внимательно посмотреть на две схемы, поясняющие получение голограммы по Ю.Н. Денисюку (рис. 1.8, *a*) и восстановление объекта с помощью голограммы Ю.Н. Денисюка (рис. 1.8, *б*).

Если голограмму по Ю.Н. Денисюку изготовить трижды, используя первичное монохроматическое излучение трех основных цветов, то возникнет цветная голограмма, причем в фотоэмульсии будет записан только



черно-белый результат интерференции. Для записи таких голограмм были разработаны специальные высокоразрешающие (до 5000 штрихов на миллиметр) фотоэмульсии.

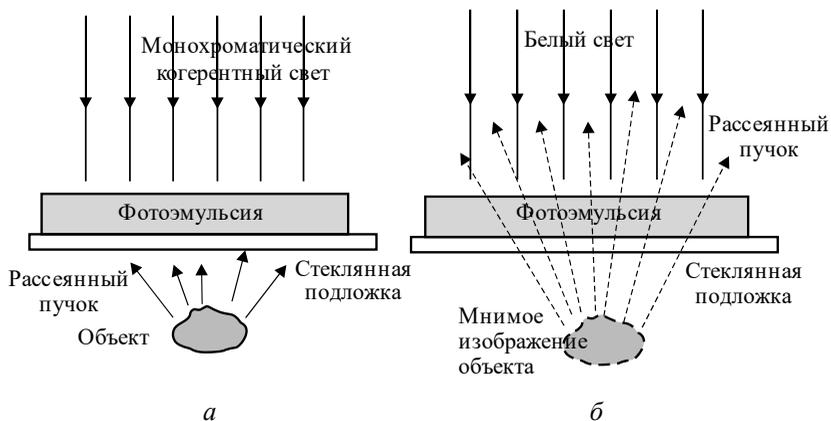


Рис. 1.8. Схема получения и восстановления голограмм по Денисюку:

а – получение голограммы; б – восстановление объекта с помощью голограммы

Экспериментальная установка была создана из подручных материалов в конце 1958 г. Поскольку в то время лазеров не было, для создания монохроматического когерентного пучка света пришлось воспользоваться схемой (рис. 1.9), которая применялась физиками еще в XIX в.

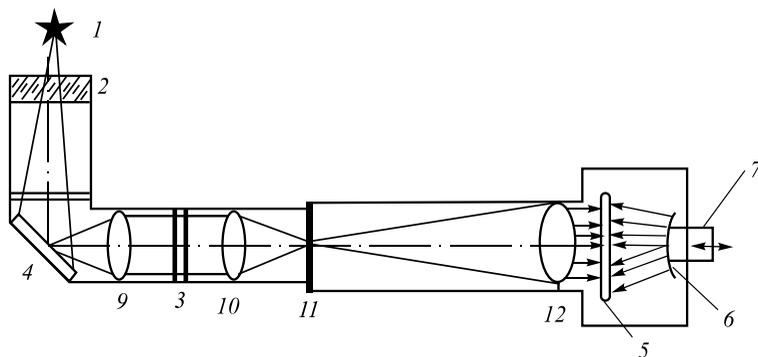


Рис. 1.9. Схема установки записи голограммы выпуклого зеркала во встречных пучках (1959 г.)



Источником света была ртутная лампа 1 (рабочая температура которой может достигать 2000 °С). Длина когерентности ртутной лампы составляла несколько миллиметров, поэтому глубина рельефа объекта должна быть ограничена до минимума. Это требование существенно ограничило выбор объектов. В результате Ю.И. Денисюк остановился на выпуклых зеркалах с большими радиусами кривизны (от  $2 \times 10^3$  до  $2 \times 10^2$  мм).

Через тепловой фильтр 2 (защита от теплового излучения ртутной лампы) свет попадал на зеркало 4 и направлялся на систему линз 9 и 10. Между линзами устанавливался фильтр 3, выделявший из широкого спектра излучения лампы 1 только одну волну (один цвет). В фокусе линзы 10 размещалась диафрагма (маленькое отверстие в непрозрачной пластине) 11. На выходе этого отверстия возникал когерентный свет, монохроматичность которого обеспечивал фильтр 3 (выделялась составляющая спектра 546 нм). Поскольку диафрагма 11 находилась в фокусе линзы 12, на ее выходе образовывался параллельный пучок света. Он направлялся через прозрачную фотопластинку 5 на выпуклое зеркало 6. Положение зеркала 6 можно было регулировать механизмом 7. Встречные пучки света, образованные падающим пучком линзы 12 и рассеянным пучком от зеркала 6, интерферировали в объеме фотопластинки 5, создавая голографическое изображение зеркала 6.

Ю.Н. Денисюк изготовил на своей установке образ вогнутого зеркала, и этот образ вел себя как настоящее вогнутое зеркало, сводил лучи света в точку.

Следует отметить, что первая из перечисленных работ Э. Лейта и Ю. Упатниекса не вызвала какого-либо резонанса и осталась незамеченной. Поворотным пунктом в истории голографии явилась работа этих авторов, опубликованная в 1964 г. В ней Э. Лейт и Ю. Упатниекс, используя свой метод и новый монохроматичный источник излучения – лазер, получили объемное изображение натурального объекта – шахматной доски с расположенными на ней фигурками. На одной из фотовыставок Э. Лейт и Ю. Упатниекс в присущей американцам манере устроили сенсацию. Они показали трехмерное изображение шахматной доски, записанное на плоской фотопластинке.

В дальнейшем Ю.Н. Денисюк сожалел, что в качестве объекта сразу не выбрал рельеф монеты. Если бы он это сделал, то сразу бы дока-



зал возможность использования голографии для создания иллюзии существования объемных объектов, и его приоритет в развитии голографии был бы более прочным.



*Ю.Н. Денисюк родился 27 июля 1927 г. в Сочи. Его научная деятельность началась в 1954 г. после окончания Ленинградского института точной механики и оптики, когда он начал работу в Государственном оптическом институте им. С.И. Вавилова. В 1958 г. он стал проводить собственные эксперименты, в которых использовал излучение лампы на парах ртути и впервые продемонстрировал трехмерную голограмму. С 1971 г. Ю.Н. Денисюк возглавлял голографическую лабораторию ГОИ, а позднее – и целый отдел, занимавшийся голографической тематикой. С 1988 г. Ю.Н. Денисюк руководил лабораторией голографии в Физико-техническом институте им. А.Ф. Иоффе.*

*На фотографии Ю.Н. Денисюк с собственным голографическим портретом.*

Обосновав теоретически и экспериментально свой метод трехмерной записи голограмм, Ю.Н. Денисюк показал, что информацию о форме и свойствах освещаемого предмета содержит не только плоская теневая проекция объекта с двумя присущими ей пространственными измерениями, но и все трехмерное пространство, которое окружает объект [4–8]. Редакция журнала «Оптика и спектроскопия» отказалась публиковать его работу, а один из известных академиков – членов редколлегии сказал, что «этот бред опубликуют в журнале только через его труп». Статья была опубликована в этом журнале только после активного вмешательства главного редактора журнала, чл.-корр. АН С.Э. Фриша [4]. По рекомендации академика В.П. Линника результаты, полученные



Ю.Н. Денисюком, были опубликованы и в журнале «Доклады Академии наук СССР» [5]. Публикация других статей, которые были отправлены одновременно с первой, существенно затянулась. Когда последняя из этих статей была опубликована в 1965 г. [6], редакция намеренно изменила дату ее поступления [7].

В феврале 1962 г. Ю.Н. Денисюк подал заявку на регистрацию открытия в Государственный комитет по делам изобретений и открытий при Совете министров СССР. Но в полученных отзывах эксперты отказывали работе в оригинальности, ссылаясь на результаты Д. Габора и Г. Липпмана, отрицалось и само существование описанного в заявке явления. Только в 1970 г. появилось решение Госкомитета о регистрации открытия за № 88 с приоритетом от 1 февраля 1962 г.

Ю.Н. Денисюк не был знаком с работами Д. Габора и введенным им термином «голография», поэтому он присвоил этому способу название «волновая фотография». Ю.Н. Денисюк получил первые открытия в области голографии раньше Э. Лейта и Ю. Упатниекса. Нобелевскую премию Ю.Н. Денисюк не получил по техническим причинам. Дело в том, что согласно Уставу Нобелевского комитета авторский коллектив, выдвигаемый на премию, не должен превышать трех человек. На премию по голографии могли претендовать четверо – Д. Габор, Э. Лейт, Ю. Упатниекс и Ю.Н. Денисюк. Работы последних трех ученых появились практически одновременно в 1962–1963 гг., и исключить кого-то одного из них было бы несправедливо. Поэтому Комитет принял решение – присудить Нобелевскую премию только Д. Габору, чей приоритет в области голографии (1948 г.) был бесспорным, несмотря на существование более ранних работ других авторов в этой области. В 1971 г. Д. Габор получил Нобелевскую премию по физике за изобретение и развитие голографического принципа.

«Наука есть достояние общее, а потому справедливость требует не тому отдать наибольшую научную славу, кто первый высказал известную истину, а тому, кто сумел убедить в ней других, показал ее достоверность и сделал ее применимой в науке» (Д.И. Менделеев)\*.

Для записи голограмм по схеме Ю.Н. Денисюка требуется наличие объемного слоя фотоземлюльсии. При вводе в компьютер анализ таких

---

\* *Виноградова Т.* Великие люди, которые изменили мир // Т. Виноградова, Д.А. Григорова. – Москва: АСТ, 2016. – С. 76.



голограмм вызывает большие сложности. Поэтому основной схемой для получения тонких голограмм в компьютерной голографии является схема с опорным наклонным пучком, предложенная Э. Лейтом и Ю. Упатниексом.

### БИБЛИОГРАФИЧЕСКИЙ СПИСОК К ГЛАВЕ 1

1. *Gabor D.* A new microscopic principle / D. Gabor // Nature. – 1948. – Vol. 161. – P. 777–778.

2. *Шушурин С. Ф.* К истории голографии / С. Ф. Шушурин // Успехи физических наук. – 1971. – № 105 (9). – С. 145–148.

3. *Leith E. N.* Reconstructed wavefronts and communication theory / E. N. Leith, J. Upatnieks // Journal of the Optical Society of America. – 1962. – Vol. 52. – P. 1123–1130.

4. *Денисюк Ю. Н.* Об отображении оптических свойств объекта в волновом поле рассеянного им излучения / Ю. Н. Денисюк // Оптика и спектроскопия. – 1962. – № 15. – С. 523–532.

5. *Денисюк Ю. Н.* Об отображении оптических свойств объекта в волновом поле рассеянного им излучения / Ю. Н. Денисюк // Доклады Академии наук СССР. – 1962. – Т. 144, № 6. – С. 1275–1278.

6. *Денисюк Ю. Н.* Об отображении оптических свойств объекта в волновом поле рассеянного им излучения / Ю. Н. Денисюк // Оптика и спектроскопия. – 1965. – Т. 18, № 2. – С. 275–283.

7. *Денисюк Ю. Н.* Мой путь в голографии / Ю. Н. Денисюк // Всероссийский семинар «Юрий Николаевич Денисюк – основоположник отечественной голографии»: сборник трудов. – Санкт-Петербург, 2007. – С. 7–14.

*Денисюк Ю. Н.* Об отображении оптических свойств объекта в волновом поле рассеянного им излучения : дис. ... канд. физ.-мат. наук / Ю. Н. Денисюк ; Гос. оптич. ин-т им. С. И. Вавилова. – Ленинград, 1963. – 133 с.

---

---

## ГЛАВА 2

# МАТЕМАТИЧЕСКАЯ МОДЕЛЬ ГОЛОГРАФИЧЕСКОГО ПРОЦЕССА

---

---

Для того чтобы исследовать новые методы анализа цифровых голограмм, необходимо построить математическую модель голографического процесса. Голографический процесс можно разделить на несколько этапов: получение волнового поля, отраженного от объекта; определение волнового поля в плоскости получения голограммы; получение голограммы в результате интерференции с опорным фронтом; восстановление комплексного волнового фронта из плоской голограммы; определение изображения в плоскости наблюдения.

Эта глава посвящена основным математическим алгоритмам, которые описывают физические процессы при распространении световых волн на каждом этапе.

### 2.1. ОСНОВНЫЕ ЭТАПЫ ГОЛОГРАФИЧЕСКОГО ПРОЦЕССА

На каждом этапе необходимо описать математические преобразования, которые происходят над волновым фронтом. Пусть объект освещается когерентной и монохроматической волной, тогда необходимо выполнить следующие действия [1–3]:

- определить распределение световой волны в плоскости, касательной к объекту;

- рассчитать математическое преобразование, которое произойдет с отраженной от объекта световой волной при распространении волнового поля от плоскости объекта  $(x, y)$  до плоскости регистрации голограммы  $(\xi, \eta)$ ;

- получить голограмму, которая образуется в результате интерференции при сложении объектного волнового фронта, отраженного от объекта, с известным опорным пучком;



- определить волновой фронт сразу же за голограммой, умножив голограмму на волновой фронт, сопряженный с опорным пучком;
- получить вид преобразования, которое будет описывать распространение волнового поля, восстановленного из голограммы, в плоскости сразу же за голограммой  $(\xi, \eta)$  перед плоскостью изображения  $(x', y')$ .

На рис. 2.1 показаны координатные системы на различных этапах. В плоскости  $(x, y)$ , касательной к объекту и параллельной плоскости наблюдения, необходимо задать комплексную функцию  $b(x, y)$ , полученную в результате пересчета амплитуды и фазы поля, отраженного объектом.

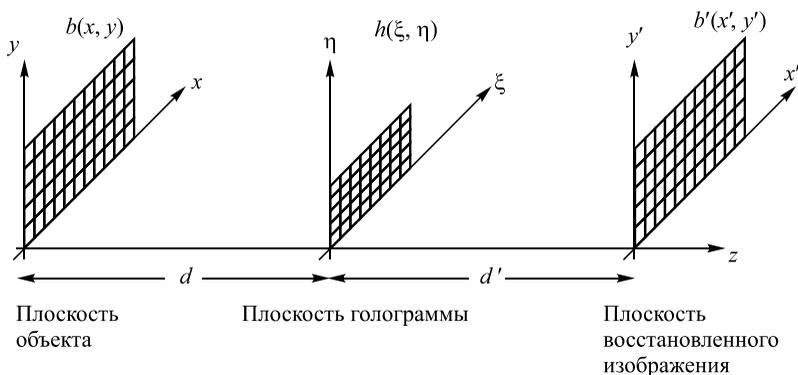


Рис. 2.1. Координатная система при получении голограмм и голографическом восстановлении

В плоскости  $(\xi, \eta)$  необходимо найти комплексное распределение  $h(\xi, \eta)$ , которое образуется при распространении волнового поля в результате прохождения расстояния  $d$  от плоскости объекта  $(x, y)$  до плоскости регистрации голограммы  $(\xi, \eta)$ . Затем в этой же плоскости необходимо найти голограмму – распределение интенсивностей по полю, которые будут образовываться в результате интерференции объектного  $h(\xi, \eta)$  и опорного пучков.

В плоскости формирования голограммы  $(\xi, \eta)$  определить комплексный волновой фронт  $h'(\xi, \eta)$  сразу же за голограммой, умножив плоскую голограмму на волновой фронт, сопряженный с опорным пучком.



В плоскости  $(x', y')$ , расположенной на расстоянии  $d'$  от плоскости  $(\xi, \eta)$ , нужно получить комплексное распределение  $b'(x', y')$ , которое должно соответствовать исходному распределению  $b(x, y)$ .

Свет – это определенный диапазон электромагнитных волн. Поэтому вид и распространение световой волны определяются уравнениями Максвелла.

## 2.2. ВИДЫ ЭЛЕКТРОМАГНИТНЫХ ВОЛН

Все электромагнитные волны принято делить на радиоволны, инфракрасное, видимое, ультрафиолетовое, рентгеновское излучения и гамма-излучение (рис. 2.2). Между ними нет резких переходов, сами участки перекрываются, а границы условны. Различные участки электромагнитного спектра отличаются друг от друга длиной волны  $\lambda$  и частотой  $\nu$ . Видимое, инфракрасное и ультрафиолетовое излучения составляют так называемую оптическую область спектра.

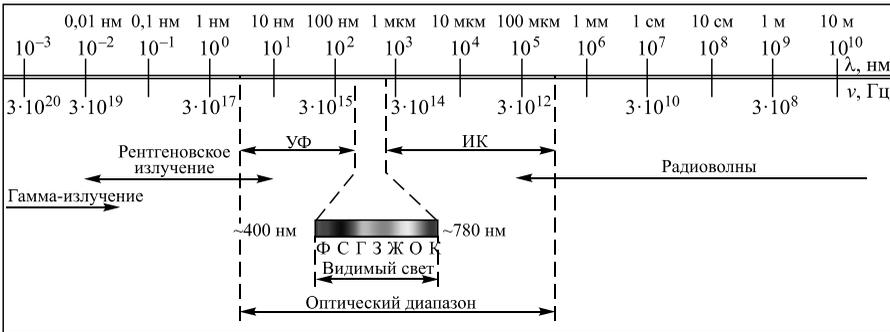


Рис. 2.2. Шкала электромагнитных волн

Выделение такой области обусловлено не только близостью соответствующих участков спектра, но и сходством методов и приборов, применяющихся для ее исследования (линзы и зеркала для фокусировки излучения, призмы, дифракционные решетки, интерференционные приборы для исследования спектрального состава излучения и пр.).

Оптический спектр занимает сравнительно небольшой диапазон от условной границы инфракрасного излучения до условной коротковолновой границы ультрафиолета:



- инфракрасное излучение

$$\lambda \sim 2 \text{ мм} \dots 760 \text{ нм} \quad (\nu = 1,5 \cdot 10^{11} \dots 4 \cdot 10^{14} \text{ Гц});$$

- видимое излучение

$$\lambda \sim 760 \dots 400 \text{ нм} \quad (\nu \sim 4 \cdot 10^{14} \dots 7,6 \cdot 10^{14} \text{ Гц});$$

- ультрафиолет

$$\lambda \sim 400 \dots 10 \text{ нм} \quad (\nu \sim 7,6 \cdot 10^{14} \dots 3 \cdot 10^{16} \text{ Гц})$$

Фотоны с точки зрения квантовой теории представляют собой волновой пакет, локализованный в пространстве и времени. Фотоны имеют длительность около  $0,1 \text{ пс}^*$ . Энергия светового кванта определяется выражением  $\varepsilon = h\nu$ , где  $h = 66,63 \cdot 10^{-27} \text{ эрг} \cdot \text{с} = 4 \cdot 10^{15} \text{ эВ} \cdot \text{с} = 6 \cdot 10^{-34} \text{ Дж} \cdot \text{с}$  – постоянная Планка, фундаментальная величина в современной физике [4]. Энергию квантов в физике принято выражать в электрон-вольтах. Это внесистемная единица измерения энергии. Один электрон-вольт (1 эВ) равен энергии, которую приобретает электрон, когда разгоняется электрическим полем напряжением 1 В. Это очень небольшая величина, в единицах системы СИ  $1 \text{ эВ} = 6 \cdot 10^{-18} \text{ Дж}$ , но в масштабах атомов и молекул электрон-вольт – достаточно значимая величина. Для длины волны  $\lambda = 1 \text{ мкм}$  энергия соответствующего кванта составляет  $\varepsilon = 1,23 \text{ эВ}$ , или приблизительно 1 эВ. На концах видимого спектра для энергии кванта формула для энергии светового кванта дает  $\varepsilon \sim 1,6 \dots 3 \text{ эВ}$ .

В области радиоволн квантовые свойства излучения практически не сказываются, поскольку энергия квантов в этой области очень мала.

В оптической области спектра частоты  $\nu$  уже перестают быть малыми по сравнению с собственными частотами атомов. Благодаря этому в такой области наряду с волновыми проявляются и квантовые свойства света. Именно поэтому эта область спектра настолько привлекательна для исследователей.

В области рентгеновского и гамма-излучения на первый план выступают уже квантовые свойства излучения. Энергия рентгеновских

---

\*  $1 \text{ пс} = 10^{-12} \text{ с}$ .



квантов лежит в пределах от 20 эВ до 1 МэВ ( $\lambda \sim 50 \dots 10^{-3}$  нм), а энергия гамма-квантов – больше 0,1 МэВ ( $< 10^{-2}$  нм).

Волновые и квантовые явления – общие для всего спектра электромагнитного излучения. Однако только в оптическом диапазоне можно измерять и волновые, и квантовые процессы. Поэтому оптике принадлежит ведущая роль в исследовании многих фундаментальных явлений.

Характеристики электромагнитных полей исчерпывающим образом описываются уравнениями Максвелла, которое представляют собой однородные дифференциальные уравнения первого порядка. Они получены феноменологическим обобщением ряда законов, описывающих электрические и магнитные явления, известные уже к середине XIX в.



*Джеймс Клерк Максвелл (1831–1879) – британский физик, математик и механик. Происходил из знатной шотландской семьи. Свою первую научную работу Дж. Максвелл выполнил еще в школе, придумав простой способ вычерчивания овальных фигур. Эта работа была представлена на заседании Королевского общества.*

*В 1855 г. Дж. Максвелл стал членом совета Тринити-колледжа, где занимался экспериментами по теории цветов. В экспериментах по смешиванию цветов Дж. Максвелл применил особый волчок, диск которого был разделен на секторы, окрашенные в разные цвета (диск Максвелла). В 1860 г. за работы по восприятию цвета и оптике Дж. Максвелл был награжден медалью Румфорда.*



*В 1860–1865 гг. Дж. Максвелл создал теорию электромагнитного поля, которую сформулировал в виде системы уравнений (уравнения Максвелла), описывающих основные закономерности электромагнитных явлений.*

Уравнения Максвелла в первоначальном виде состояли из системы 20 дифференциальных уравнений с 12 переменными [5]. В 1880 г. О. Хевисайд переписал результаты Максвелла в терминах векторного анализа. В настоящее время они записываются в виде восьми (два векторных с тремя компонентами каждое и два скалярных) линейных дифференциальных уравнений в частных производных первого порядка для 12 компонент четырех векторных функций ( $D$  – электрическая индукция,  $E$  – напряженность электрического поля,  $H$  – напряженность магнитного поля,  $B$  – магнитная индукция).

Распространение электромагнитных волн в веществе в тех случаях, когда длина волны велика по сравнению с межатомными расстояниями, можно рассматривать феноменологически, т. е. без учета атомистического строения среды. Для этого уравнения Максвелла дополняются материальными уравнениями, в которых свойства среды учитываются введением соответствующих параметров.

Доказал и обнаружил существование электромагнитных волн Г. Герц только через 25 лет после записи уравнений.



*Генрих Рудольф Герц (нем. Heinrich Rudolf Hertz) (1857–1894) – немецкий физик. Окончил Берлинский университет, где его учителями были Герман фон Гельмгольц и Г. Кирхгоф. С 1885 по 1889 г. профессор физики Университета в Карлсруэ. С 1889 г. – профессор физики Университета в Бонне.*



*Основное достижение – экспериментальное подтверждение электромагнитной теории света Дж. Максвелла. Генрих Герц доказал существование электромагнитных волн. Он подробно исследовал отражение, интерференцию, дифракцию и поляризацию электромагнитных волн и доказал, что скорость их распространения совпадает со скоростью распространения света и что свет представляет собой не что иное, как разновидность электромагнитных волн. Результаты, полученные Г. Герцем, легли в основу создания радио.*

*В 1886–1887 гг. Г. Герц впервые наблюдал внешний фотоэффект и дал ему описание. Герц разрабатывал теорию резонансного контура, изучал свойства катодных лучей, исследовал влияние ультрафиолетовых лучей на электрический разряд.*

Уравнения Максвелла не выводятся, они получены обобщением опытных фактов, и не стоит ожидать, что они угаданы в окончательной форме. Уравнения Максвелла, как говорил Г. Герц, «живут самостоятельной жизнью и иногда кажутся даже умнее человека, создавшего их».

Распространение электромагнитных волн со скоростью света первоначально интерпретировалось как возмущения некоторой среды, так называемого эфира. Были предприняты многочисленные попытки обнаружить движение Земли относительно эфира, однако они неизменно давали отрицательный результат. Поэтому А. Пуанкаре высказал гипотезу о принципиальной невозможности обнаружить подобное движение (принцип относительности). Ему же принадлежит постулат о независимости скорости света от скорости его источника. Эти две гипотезы легли в основу статьи А. Эйнштейна, в которой излагались эти результаты [6].

Практически любое физическое явление можно описать в виде дифференциальных уравнений, но получить их аналитическое решение можно только для ограниченного класса уравнений. Поскольку свет – это электромагнитная волна, то можно попробовать получить аналитические решения дифференциальных уравнений Максвелла при некоторых граничных условиях.

Уравнения Максвелла являются дифференциальными уравнениями первого порядка по координатам и времени. Для вакуума (при отсутствии токов и зарядов) уравнения Максвелла имеют простой вид:

$$\nabla \vec{B} = 0; \quad (2.2.1)$$

$$\nabla \times \vec{E} = -\frac{\partial \vec{B}}{\partial t}; \quad (2.2.2)$$



$$\nabla \vec{E} = 0; \quad (2.2.3)$$

$$c^2 [\nabla \times \vec{B}] = \frac{\partial \vec{E}}{\partial t}, \quad (2.2.4)$$

где  $\vec{E}$  – напряженность электрического поля (в единицах СИ – В/м);  $\vec{B}$  – магнитная индукция (в единицах СИ – Тл = Вб/м<sup>2</sup> = кг с<sup>-2</sup>; Тл – тесла, Вб – вебер, кг – килограмм (массы), с – секунда);  $c = \sqrt{1/\epsilon_0\mu_0}$  – электродинамическая постоянная;  $\epsilon_0$  – диэлектрическая постоянная ( $\epsilon_0 = 8,8541878210^{-12}$  Ф/м; Ф – фарад);  $\mu_0$  – магнитная постоянная ( $\mu_0 = 1,25663706 \cdot 10^{-6}$  Гн/м; Гн – генри). Электродинамическая постоянная имеет размерность скорости. Измерения показали, что ее значение совпадает со скоростью света в вакууме.

Уравнения Максвелла записаны в операторной форме. Ротор вектора обозначает векторное произведение векторного дифференциального оператора  $\nabla$  на вектор  $\vec{A}$ :

$$\begin{aligned} \text{rot} \vec{A} = [\nabla \times \vec{A}] &= \begin{bmatrix} \vec{e}_x & \vec{e}_y & \vec{e}_z \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ \vec{A}_x & \vec{A}_y & \vec{A}_z \end{bmatrix} = \\ &= \left( \frac{\partial \vec{A}_z}{\partial x} - \frac{\partial \vec{A}_y}{\partial z} \right) \vec{e}_x + \left( \frac{\partial \vec{A}_x}{\partial z} - \frac{\partial \vec{A}_z}{\partial x} \right) \vec{e}_y + \left( \frac{\partial \vec{A}_y}{\partial x} - \frac{\partial \vec{A}_x}{\partial y} \right) \vec{e}_z. \end{aligned} \quad (2.2.5)$$

Дивергенция вектора обозначает скалярное произведение векторного дифференциального оператора  $\nabla$  на вектор  $\vec{A}$ :

$$\text{div} \vec{A} = [\nabla \vec{A}] = \frac{\partial A_x}{\partial x} \vec{e}_x + \frac{\partial A_y}{\partial y} \vec{e}_y + \frac{\partial A_z}{\partial z} \vec{e}_z, \quad (2.2.6)$$

где  $\vec{e}_x, \vec{e}_y, \vec{e}_z$  – единичные векторы, направленные вдоль осей координат  $x, y, z$  соответственно.



При отсутствии зарядов и токов можно перейти к уравнениям второго порядка, каждое из которых зависит только от одного поля – электрического или магнитного. Такие уравнения называются *волновыми*.

Умножим обе части уравнения (2.2.2) векторно на  $\nabla$  и, раскрывая двойное векторное произведение получим:

$$\nabla \nabla \vec{E} = -\nabla \frac{\partial \vec{B}}{\partial t}. \quad (2.2.7)$$

Учитывая, что

$$\vec{A}\vec{B}\vec{C} = (\vec{B}(\vec{A}\vec{C}) - \vec{C}(\vec{A}\vec{B})), \quad (2.2.8)$$

получим

$$\nabla(\nabla \vec{E}) - \nabla^2 \vec{E} = -\frac{\partial(\nabla \vec{B})}{\partial t}. \quad (2.2.9)$$

Так как  $\nabla \vec{E} = 0$  (2.1.3) и  $\nabla \vec{B} = \frac{1}{c^2} \frac{\partial \vec{E}}{\partial t}$  (2.2.4), получим уравнение, которое будет учитывать только электрическое поле:

$$\nabla^2 \vec{E} = \frac{1}{c^2} \frac{\partial^2 \vec{E}}{\partial t^2}. \quad (2.2.10)$$

Аналогично для магнитного поля

$$\nabla^2 \vec{B} = \frac{1}{c^2} \frac{\partial^2 \vec{B}}{\partial t^2}. \quad (2.2.11)$$

Уравнения имеют решения в виде бегущих электромагнитных волн, которые распространяются с некоторой скоростью  $c$ .

Наиболее распространенные датчики реагируют только на электрическую составляющую электромагнитного поля. Если интересоваться только интенсивностью волн, отвлекаясь от их поляризации, то распространение света в вакууме можно записать в виде скалярного волнового уравнения

$$\nabla^2 E(r, t) - \frac{1}{c^2} \frac{\partial^2 E(r, t)}{\partial t^2} = 0, \quad (2.2.12)$$

где  $E(r, t)$  – длина вектора  $\vec{E}$ .



Нас интересует решение этого уравнения для случая монохроматических волн. Будем искать решение в виде

$$E(r,t) = E(r)e^{-\omega t}, \quad (2.2.13)$$

где  $E(r)$  – амплитуда;  $\omega$  – круговая частота волны. Ее удобно выражать через волновое число в вакууме, которое вводится для удобства как большой размерный параметр:

$$k = \frac{\omega}{c}. \quad (2.2.14)$$

Подставляя (2.2.13) в (2.2.12), получим

$$\nabla^2 E(r) + k^2 E(r) = 0 \quad (2.2.15)$$

– это *уравнение Гельмгольца* (*Ludwig Ferdinand von Helmholtz*).

При исследованиях волновых электромагнитных процессов обычно от уравнений Максвелла [2.2.1–2.2.4] и Гельмгольца [2.2.15] переходят к более простым выражениям, если на поле наложены некоторые дополнительные условия.



*Гельмгольц Герман Людвиг Фердинанд (1821–1894). Немецкий физик, математик, физиолог и психолог, родился в семье учителя гимназии. Несмотря на интерес к физике, Г. Гельмгольц не смог из-за недостатка средств посту-*



пить в университет. Подписав обязательство прослужить восемь лет военным хирургом, он был принят в Военно-медицинский институт в Берлине. В 1842 г. Г. Гельмгольц защитил диссертацию по физиологии, в 1843–1848 гг. служил военным врачом в Потсдаме. В 1848 г. Г. Гельмгольца освободили от военной службы, и он занял должность экстраординарного профессора физиологии и общей патологии в Университете Кёнигсберга. Работы Г. Гельмгольца увели его далеко за пределы физиологии. В 1870 г. Г. Гельмгольца пригласили в Берлин, где возглавляемая им кафедра и лаборатория стали неформальным центром физики в Германии. «Произошло неслыханное дело: медик и профессор физиологии занял главную физическую кафедру Германии». В 1888 г. он возглавил Физико-технической университет.

В 1870–1880 гг. Г. Гельмгольц много занимался проблемами электродинамики. Под его влиянием Г.Р. Герц провел исследования, приведшие к обнаружению электромагнитных волн. Большую роль в развитии электромагнетизма сыграли и собственные опыты Г. Гельмгольца, поставленные им еще в 1869 г.

Ему удалось стать первым врачом среди ученых и первым ученым среди врачей.

Решив уравнение Гельмгольца относительно  $E(r)$ , получим выражения для некоторых типов монохроматических волн.

**Плоская волна.** Решение для плоской волны с бесконечно большим поперечным размером и постоянной амплитудой  $E_0$  может быть записано в виде

$$E(\vec{r}) = E_0 e^{i\vec{k}\vec{r}}. \quad (2.2.16)$$

Согласно формуле Эйлера

$$e^{ia} = \cos(a) + i \sin(a). \quad (2.2.17)$$

В оптике используется только действительная часть, поэтому запись  $E(x, t) = E e^{i(\omega t - kx + \varphi_0)}$  означает то же, что и

$$E(x, t) = E_0 \cos(\omega t - kx + \varphi_0). \quad (2.2.18)$$

Выражение (2.2.18) описывает плоскую монохроматическую волну, распространяющуюся в направлении  $x$ . Аргумент  $\omega t - kx + \varphi_0$  называется фазой волны.



Значения, при которых  $\omega t - \vec{k}\vec{r} = \text{const}$ , определяют в пространстве плоскость, перпендикулярную вектору  $\vec{k}$ , называемому волновым вектором. Эти значения задают уравнение поверхности постоянной фазы (или волновой поверхности).

Волна называется плоской, если ее волновые поверхности являются параллельными плоскостями, перпендикулярными направлению распространения (рис. 2.3). Эта плоскость распространяется вдоль направления вектора  $\vec{k}$  со скоростью  $v = \omega/k$ , где  $k$  – модуль волнового вектора, называемый волновым числом.

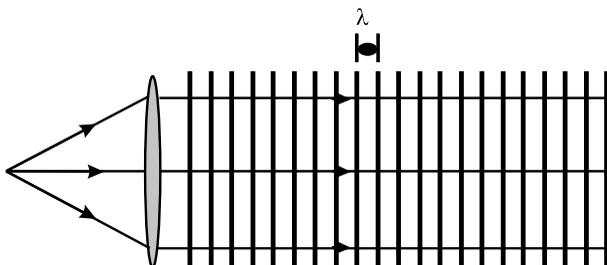


Рис. 2.3. Плоская монохроматическая волна, распространяющаяся в направлении  $x$

Скорость перемещения поверхности постоянной фазы в пространстве называется *фазовой скоростью волны*. Для плоских монохроматических волн в вакууме фазовая скорость равна скорости света  $c$ .

Период изменения напряженности поля в пространстве – это длина волны  $\lambda$  :

$$\lambda = \frac{2\pi}{k} = \frac{2\pi v}{\omega} = vT, \quad (2.2.19)$$

где  $v$  – частота колебаний.

Длина волны представляет собой расстояние, на которое перемещается плоскость постоянной фазы за время, равное периоду колебаний  $T = \frac{2\pi}{\omega}$ .

Из (2.2.19) следует, что модуль волнового вектора равен

$$k = \frac{2\pi}{\lambda}. \quad (2.2.20)$$



Для реальных волн существуют следующие ограничения:

- волна ограничена в поперечном размере. Это ограничение может быть вызвано оптическими элементами системы или физическими ограничениями;
- амплитуда  $E_0$  и начальная фаза  $\varphi_0$  плоских монохроматических волн не являются постоянными, а зависят от  $\vec{r}$  и  $t$ .

Волна с бесконечно большим поперечным размером и постоянной амплитудой  $E_0$ , распространяющаяся бесконечно, только математическая абстракция. Никакие реальные волны этим свойством не обладают, но они могут быть представлены в виде суммы таких простых волн (благодаря линейности уравнений Максвелла сумма любых решений также является решением).

**Принцип суперпозиции.** Электромагнитные волны (как и звуковые) подчиняются принципу *суперпозиции*. Пусть у нас есть несколько источников волн. Все источники независимы друг от друга. Каждый источник испускает свои волны, и в пространстве, окружающем источники, образуется сложное волновое поле. Принцип суперпозиции волн гласит, что волны от различных источников не взаимодействуют друг с другом и что сложное волновое поле от двух или большего числа источников находится путем геометрического сложения волн от отдельных источников.

Этот принцип позволяет не только складывать волны, но и раскладывать их, например, на независимые синусоидальные волны. Это означает, что любую волну, т. е. волну произвольного профиля, всегда можно представить как сумму синусоидальных волн с различными амплитудами, с различными фазовыми скоростями, с различными частотами и с различными начальными фазами.

Таким образом, произвольную (не обязательно плоскую) волну всегда можно представить в виде суммы плоских волн, движущихся в различных направлениях и имеющих разные частоты. Этой возможностью разложения волн широко пользуются в оптике.

**Сферическая волна.** Другим решением уравнения Гельмгольца является

$$E(\vec{r}) = \frac{A}{r} e^{ikr}, \quad (2.2.21)$$

где  $A$  – константа, которая показывает величину амплитуды при единичном расстоянии. Это решение соответствует сферической волне.



Плоская волна распространяется в определенном направлении, тогда как сферическая волна не имеет определенного направления распространения.

Уравнение (2.2.21) можно записать в виде

$$E(x, y, z) = \frac{A}{\sqrt{x^2 + y^2 + z^2}} \exp \left[ ik\sqrt{x^2 + y^2 + z^2} \right]. \quad (2.2.22)$$

Оно определяет амплитуду в точке  $(x, y, z)$  на расстоянии  $z$  от источника. Это выражение часто используется при описании распространения света через оптическую систему.

Сферическая волна расходится из точки источника или может сходиться к ней (рис. 2.4).

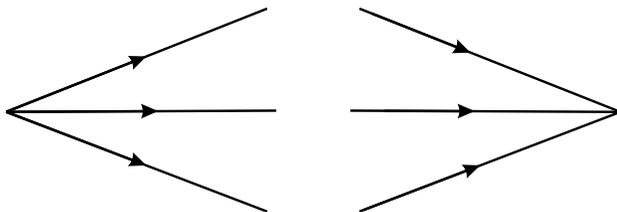


Рис. 2.4. Сферическая волна (слева расходящаяся, справа сходящаяся)

**Гауссовы пучки.** В узких световых пучках поле сконцентрировано около одной координаты (оси пучка  $z$ ) и быстро падает до нуля в поперечных координатах. Учесть это можно следующим образом. Любую из компонент поля можно представить в виде произведения  $\Psi(x, y, z) = F(x, y, z)e^{ikz}$ , где  $F$  – медленно меняющаяся с ростом  $z$  комплексная функция. Подставляя это выражение в уравнение Гельмгольца (2.2.15) и пренебрегая членом  $\frac{\partial^2 F}{\partial z^2}$  по сравнению с другими членами, получим

$$\frac{\partial^2 F}{\partial x^2} + \frac{\partial^2 F}{\partial y^2} + 2ik \frac{\partial F}{\partial z} = 0. \quad (2.2.23)$$



Уравнение (2.2.23) называется параболическим или уравнением *Леонтовича–Фока*, которое В.А. Фок и М.А. Леонтович впервые получили в 1946 г., изучая распространение радиоволн вокруг Земли.



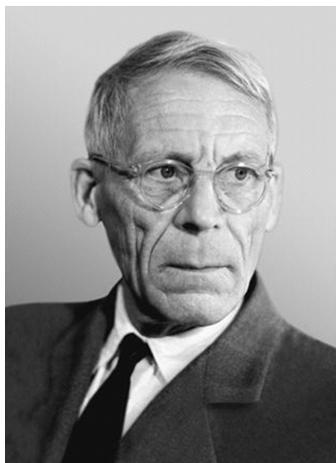
*Фок Владимир Александрович (1898–1974) – профессор Санкт-Петербургского госуниверситета, академик.*

*Для Фока характерно сочетание тонкого физического чутья и необычайной математической мощи. Многие впервые им полученные результаты вошли в учебники и научную литературу под его именем и стали основой современной физики и ее применений в сопредельных науках.*

*Метод параболического уравнения Фока–Леонтовича открыл новые возможности в изучении распространения радиоволн. В.А. Фок первым решил труднейшую проблему распространения радиоволн вокруг земного шара, сделав тем самым революционный шаг в теории дифракции. Современные работы по дифракции радиоволн опираются на эти результаты.*

В оптике функция  $F(x, y, z)$  описывает поведение когерентного светового поля с длиной волны  $\lambda$  в вакууме в параксиальном приближении, т. е. в предположении, что при распространении изменение поля  $F(x, y, z)$  вдоль осей  $x$  и  $y$  значительно слабее, чем вдоль оси  $z$ .

В этом случае  $z$  называется осью распространения, а  $k = \frac{2\pi}{\lambda}$  – волновым числом. Функция  $F(x, y, z)$  в оптике называется комплексной амплитудой.



*Леонтович Михаил Александрович (1903–1981) – академик АН СССР, лауреат Ленинской премии (1958).*

*В 1935 г. ему присуждена степень доктора физико-математических наук без защиты диссертации. В совместных работах М.А. Леонтовича и В.А. Фока (1944–1945) была развита теория распространения радиоволн вокруг Земли.*

*С 1951 г. Леонтович заведует теоретическими исследованиями по управляемому термоядерному синтезу в Институте атомной энергии, заняв этот пост по предложению И.Е. Тамма. Его идеи об уравнивании тороидального растяжения плазмы с током при помощи проводящего кожуха и о стабилизации плазменного витка сильным магнитным полем легли в основу системы «Токамак».*

Уравнению (2.2.23) будет удовлетворять гауссов пучок [7, 8], амплитуда которого будет меняться по гауссову закону

$$\Psi(x, y, z) = a \exp \left[ i \left( p + \frac{k}{2q} (x^2 + y^2) \right) \right], \quad (2.2.24)$$

где  $p$  – комплексный фазовый сдвиг при распространении света вдоль оси  $z$ ;  $q$  – комплексный параметр пучка, определяющий гауссово распределение поля по координате  $r$ , где  $r = \sqrt{x^2 + y^2}$  – расстояние до оси  $z$ .



Подставив (2.1.24) в (2.1.23) получим

$$2k \left( \frac{\partial p}{\partial z} + \frac{i}{q} \right) + \left( \frac{kr^2}{q} \right) \left( 1 - \frac{\partial q}{\partial z} \right) = 0. \quad (2.2.25)$$

Уравнение (2.2.25) эквивалентно двум уравнениям:

$$\frac{\partial q}{\partial z} = 1; \quad (2.2.26 \text{ а})$$

$$\frac{\partial p}{\partial z} = -\frac{i}{q}. \quad (2.2.26 \text{ б})$$

Интегрируя (2.2.26 а), получим

$$q(z) = q_0 + z. \quad (2.1.27)$$

Это уравнение устанавливает простое соотношение между параметром пучка  $q(z)$  в разных сечениях, отстоящих друг от друга на расстояние  $z$ . Постоянная интегрирования  $q_0$  находится из начальных условий. Предпочтительнее выбирать начало координат в том месте, где фазовая поверхность электромагнитного поля может быть плоской. В этом случае можно положить, что при  $z = 0$  амплитуда  $\Psi(x, y, z)$  должна иметь вид

$$\Psi(x, y, z) = \exp \left( -\frac{x^2 + y^2}{w_0^2} \right), \quad (2.2.28)$$

где  $w_0$  – радиус пучка в поперечнике. Распределение поля в этой плоскости, как видно из рис. 2.5, подчиняется закону Гаусса, и  $w_0$  равно расстоянию, на котором амплитуда поля убывает в  $e$  раз по сравнению с полем на оси ( $x = 0$ ,  $y = 0$ ,  $z = 0$ ).

Сравнивая (2.2.28) с (2.2.24), видим, что в начале координат смещение фазы  $p(z) = 0$ . При этом постоянная интегрирования равна чисто мнимой величине

$$q_0 = \frac{\pi w_0^2}{i\lambda}. \quad (2.2.29)$$

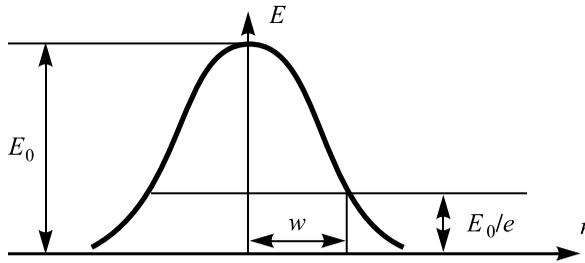


Рис. 2.5. Распределение поля  $E$  в поперечной плоскости

Действительная величина  $z_R = iq_0 = \frac{\pi w_0^2}{\lambda}$  называется длиной Рэлея.

Но иногда (в теории дифракции для фазированных решеток и т. п.) предпочтительнее считать  $q_0$  некоторым заданным комплексным числом. Ниже мы ограничимся простейшим случаем гауссовых пучков, когда  $q_0$  является чисто мнимой и определяется выражением (2.2.29). В этом случае

$$q(z) = iz_R + z. \quad (2.2.30)$$

Для расчета комплексного фазового сдвига  $p$  воспользуемся соотношениями (2.1.26 б):

$$\frac{\partial p}{\partial z} = -\frac{i}{q(z)} = -\frac{i}{iz_R + z}. \quad (2.2.31)$$

Интегрирование (2.2.31) дает

$$ip(z) = \ln(iz_R + z) + c. \quad (2.2.32)$$

Учитывая, что в начале координат  $p(0) = 0$ , находим, что  $c = -\ln(iz_R)$ . Следовательно,

$$ip(z) = \ln\left(\frac{iz_R + z}{iz_R}\right) = \ln\left(1 - \frac{iz}{z_R}\right). \quad (2.2.33)$$



Далее

$$\exp[ip(z)] = \sqrt{1 + \frac{z^2}{z_R^2}} \left( i \operatorname{arctg} \frac{z}{z_R} \right). \quad (2.2.34)$$

Таким образом, мы фактически определили функции  $p(z)$  и  $q(z)$ , следовательно, и функцию  $\Psi(x, y, z)$  (2.2.24).

Рассмотрим свойства гауссова пучка с длиной волны  $\lambda$  более подробно. Для этого выразим комплексный параметр  $q$  через два действительных параметра  $R$  и  $w$ :

$$\frac{1}{q} = \frac{1}{R} + i \frac{\lambda}{\pi w^2}. \quad (2.2.35)$$

Физический смысл этих параметров становится ясным из соотношений (2.2.31) и (2.2.36):

$$R(z) = z \left( 1 + \frac{z^2}{z_R^2} \right); \quad (2.2.36)$$

$$w^2(z) = w_0^2 \left( 1 + \frac{z^2}{z_R^2} \right). \quad (2.2.37)$$

Подставляя значения (2.2.34–2.2.37) в (2.2.24), получим

$$\Psi(r, z) = \frac{w_0}{w} \exp \left[ -i \left[ kz + \frac{k}{2R} (x^2 + y^2) - \Phi \right] - \frac{(x^2 + y^2)}{w^2} \right], \quad (2.2.38)$$

где

$$\Phi(z) = \operatorname{arctg}(z / z_R) = \operatorname{arctg} \left( \frac{\lambda z}{\pi w_R^2} \right). \quad (2.2.39)$$



Из формулы (2.2.39) видно, что  $\Phi(z)$  растет с увеличением  $z$  и уменьшением минимального радиуса пучка  $w_0$ . Максимальное значение  $\Phi(z)$  равно  $\pi/2$ .

Фазовая поверхность пучка образуется приравнением чисто мнимого показателя экспоненты (в 2.2.38) некоторой постоянной  $kz_1$ :

$$kz_1 = kz + \frac{k}{2R}(x^2 + y^2) - \Phi. \quad (2.2.40)$$

Поскольку смещение фазы  $\Phi$  мало по сравнению с другими членами, им можно пренебречь, и тогда

$$\frac{(x^2 + y^2)}{2R(z)} = z_1 - z \quad (2.2.41)$$

– это уравнение поверхности вращения. Параметр  $R(z)$  является радиусом кривизны этой поверхности для точек, лежащих на оси  $0z$ .

При рассмотрении гауссовых пучков принято определять форму пучка видом поверхности, которую образуют точки в пространстве, в которых амплитуда поля в  $e$  раз меньше, чем в точках на центральной оси пучка соответствующего поперечного сечения. То есть это граница области, значения которой больше  $e$ . Уравнение этой границы можно получить при приравнении единице вещественного показателя экспоненты (2.2.38):

$$\frac{(x^2 + y^2)}{w^2} = 1, \quad (2.2.42)$$

или

$$(x^2 + y^2) = w_0^2 \left( 1 + \frac{z^2}{z_R^2} \right) \quad (2.2.43)$$

– это уравнение однополосного гиперboloида вращения (рис. 2.6).

Образующая гауссова пучка (рис. 2.7) представляет собой гиперболу, асимптота которой наклонена к оси под углом  $\theta = \frac{\lambda}{\pi w_0}$ . Этот угол равен

углу дифракции в дальней зоне. Общая угловая расходимость  $\theta = 2\theta$ .

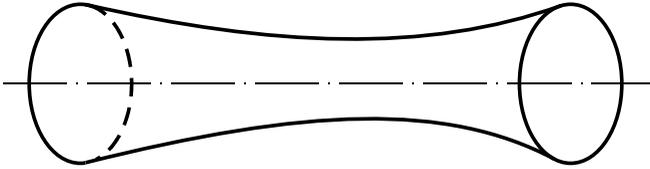


Рис. 2.6. Пространственная форма гауссова пучка

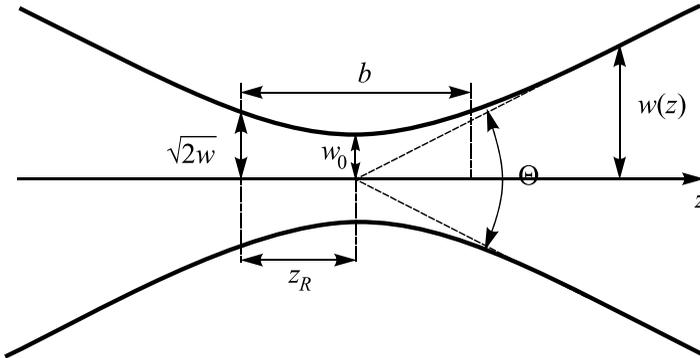


Рис. 2.7. Образующая гауссова пучка ( $b$  – глубина резкости)

В некоторой плоскости, называемой перетяжкой, гауссов пучок стягивается к минимальной ширине  $w_0$ . В этой плоскости, от которой целесообразно отсчитывать расстояние  $z$ , фазовый фронт является плоским и комплексный параметр пучка является чисто мнимым.

Видно, что  $R$  есть радиус кривизны волнового фронта (рис. 2.8), а  $w$  характеризует изменение поля  $E$  в поперечной плоскости.

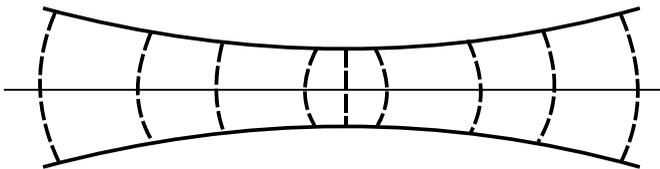


Рис. 2.8. Гауссов световой пучок (сплошными линиями изображена пространственная форма пучка, штриховыми – фазовая поверхность)



Окончательно гауссовый пучок можно представить в виде

$$\Psi(r, z) = A_0 \frac{w_0}{w(z)} \times \exp \left[ -i \left[ kz + \frac{k}{2R(z)}(x^2 + y^2) - \Phi(z) \right] - \frac{(x^2 + y^2)}{w^2(z)} \right]; \quad (2.2.44)$$

$$\Psi(r, z) = A_0 \frac{w_0}{w(z)} \exp \left[ -\frac{(x^2 + y^2)}{w^2(z)} \right] \times \exp \left[ -i \left[ kz - \Phi(z) + \frac{k}{2R(z)}(x^2 + y^2) \right] \right]. \quad (2.2.45)$$

Интенсивность светового пучка равна  $I(r) = |\Psi(r, z)|^2$ , откуда

$$I(x, y, z) = I_0 \left( \frac{w_0}{w(z)} \right)^2 \exp \left[ -\frac{2(x^2 + y^2)}{w^2(z)} \right], \quad (2.2.46)$$

где

$$w_0 = \sqrt{\frac{\lambda z_R}{\pi}}, \quad (2.2.47a)$$

$$w(z) = w_0 \sqrt{\left( 1 + \frac{z^2}{z_R^2} \right)}; \quad (2.2.47b)$$

$$I_0 \left( \frac{z_R^2}{(z_R^2 + z^2)} \right) \exp \left[ -\frac{2\pi(x^2 + y^2)z_R^2}{\lambda z_R(z_R^2 + z^2)} \right]. \quad (2.2.48)$$

Параметр  $z_R$  определяется из граничных условий.

Теперь рассмотрим первый этап получения голограмм – определим вид волнового фронта в плоскости голограммы при отражении от объекта.



### 2.3 МАТЕМАТИЧЕСКАЯ МОДЕЛЬ РАСПРОСТРАНЕНИЯ СВЕТА В ПРОСТРАНСТВЕ

Пусть объект освещается световой волной. Нам необходимо определить распределение световой волны в плоскости объекта и рассчитать математическое преобразование, которое произойдет при распространении светового волнового поля от плоскости объекта до плоскости регистрации голограммы (рис. 2.9).

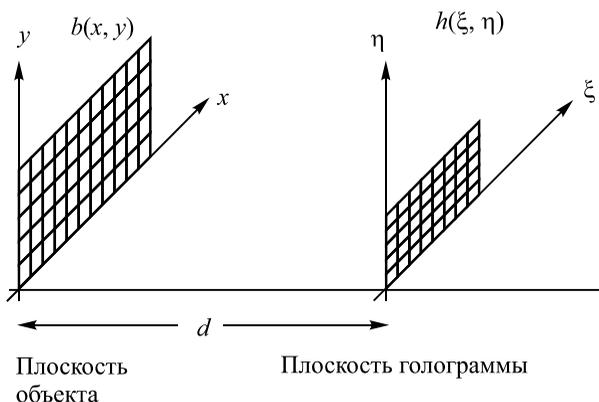


Рис. 2.9. Координатные системы плоскости объекта и плоскости объекта до плоскости формирования голограммы ( $d$  – расстояние между двумя плоскостями)

Существующая теория описывает распространение поля в соответствии с принципом Гюйгенса. Задолго до открытия электромагнитной природы света Х. Гюйгенс предложил трактовать распространение света с использованием понятия волновых поверхностей [9] (рис. 2.10).

Пусть известна какая-либо волновая поверхность. Тогда каждую точку этой поверхности можно рассматривать как вторичный источник колебаний, а затем следующие волновые поверхности находить как огибающие сферических волн, испускаемых этими точечными источниками. О.Ж. Френель дал математическую формулировку принципа Гюйгенса и предложил учитывать фазы колебаний, пришедших от различных точек, при их сложении.

Г. Киргоф придал этой идее строгий математический вид и показал, что принцип Гюйгенса–Френеля можно считать приближенной фор-



мой интегральное выражения. Проблемы, возникающие при изучении дифракционных явлений, являются наиболее трудными в оптике, и их редко удается довести до строгого математического решения. Первое такое решение было получено только в 1986 г. А. Зоммерфельдом, рассмотревшим вопрос о дифракции плоской волны на идеально проводящем полубесконечном плоском экране [10]. В большинстве случаев, представляющих практический интерес, из-за математических трудностей приходится прибегать к приближенным методам.

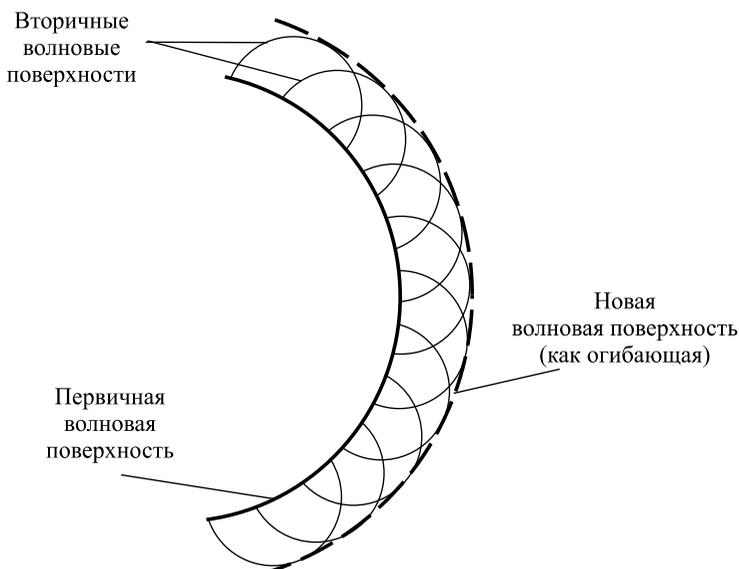


Рис. 2.10. Принцип распространения света Гюйгенса

Свойства объекта, определяющие его способность отражать и рассеивать падающее на него излучение, можно описать коэффициентом отражения излучения по интенсивности  $B(x, y, z)$  или амплитуде  $b(x, y, z)$  – функциям координат на поверхности объекта. Коэффициент отражения по амплитуде является комплексной функцией, которая может быть представлена в виде [11]

$$b(x, y, z) = |b(x, y, z)| \exp(i\beta(x, y, z)). \quad (2.3.1)$$



*Христиан Гюйгенс ван Зейлихем (Christiaan Huygens, 14 апреля 1629 г. – 8 июля 1695 г.) – нидерландский механик, физик, математик.*

*В 1678 г. он выпустил «Трактат о свете» – набросок волновой теории света. В 1690 г. изложил качественную теорию отражения, преломления и двойного лучепреломления в исландском шпате в том самом виде, в котором она излагается теперь в учебниках физики.*

*Сформулировал принцип Гюйгенса, позволяющий исследовать движение волнового фронта, впоследствии развитый О.Ж. Френелем и сыгравший важную роль в волновой теории света и теории дифракции.*

*В 1657 г. Х. Гюйгенс получил голландский патент на конструкцию маятниковых часов. В 1675 г. запатентовал карманные часы.*

*Первый призвал выбрать всемирную натуральную меру длины, в качестве которой предложил 1/3 длины маятника с периодом колебаний 1 с (около 8 см).*

Ее модуль  $|b(x, y, z)|$  и фаза  $\beta(x, y, z)$  показывают, во сколько раз изменяется амплитуда и, соответственно, насколько изменяется фаза излучения в точке  $(x, y, z)$  поверхности объекта после отражения. Функция  $B(x, y, z)$  определяется как

$$B(x, y, z) = |b|^2 = bb^*, \quad (2.3.2)$$

где \* – знак комплексного сопряжения.



Зная функцию  $b(x, y, z)$ , уравнение поверхности тела  $F(x, y, z) = 0$  и распределение амплитуды и фазы падающего на объект света, можно вычислить распределение амплитуды и фазы рассеянного света в произвольной точке пространства.

Пусть  $A(x, y, z) \exp(i\alpha(x, y, z))$  – распределение амплитуды и фазы освещения на поверхности объекта. Тогда поле на некоторой поверхности наблюдения можно описать с помощью интегрального соотношения Кирхгофа [10]

$$\Gamma(\xi, \eta, \zeta) = \int_{F(x, y, z)} A(x, y, z) |b(x, y, z)| \exp(i[\alpha(x, y, z) + \beta(x, y, z)]) T(x, y, z, \xi, \eta, \zeta) dx dy dz, \quad (2.3.3)$$

где интегрирование производится на поверхности объекта  $F(x, y, z) = 0$ . Вид ядра этого преобразования  $T(x, y, z, \xi, \eta, \zeta)$  зависит от пространственного расположения объекта и поверхности наблюдения.

Это преобразование в принципе обратимо:

$$\beta(x, y, z) = \int_S \Gamma(\xi, \eta, \zeta) \bar{T}(x, y, z, \xi, \eta, \zeta) d\xi d\eta d\zeta, \quad (2.3.4)$$

где  $\bar{T}$  – это оператор, взаимный  $T$ , а интегрирование происходит по поверхности наблюдения  $S$ .

Функцию  $\Gamma(\xi, \eta, \zeta)$  можно назвать математической голограммой, а выражение (2.3.4) описывает восстановление волнового поля на поверхности объекта [11].

Вычисление выражений (2.3.3) и (2.3.4) в общем случае представляет довольно сложную задачу. Ее удастся решить только для очень простых объектов, заданных небольшим количеством отдельных линий или точек. В общем случае приходится прибегать к некоторым упрощениям.

**Первое упрощение**, к которому можно прибегнуть, состоит в сведении трехмерной задачи к двумерной. Для этого поверхность наблюдения считается плоской, а распределение амплитуды и фазы волны на поверхности объекта заменяется по законам геометрической оптики



распределением амплитуды и фазы на плоскости, касающейся объекта или достаточно близкой к нему и параллельной плоскости наблюдения (рис. 2.11).



*Густав Роберт Кирхгоф родился 12 марта 1824 г. в Кёнигсберге; с 1842 по 1846 г. изучал математику и физику в Кёнигсбергском университете, а в 1847 г. уже выступил в качестве приват-доцента в Берлине; в 1850–1854 гг. в качестве экстраординарного профессора читал лекции в Бреславле, затем до 1874 г. занимал должность ординарного профессора в Гейдельберге, откуда в 1875 г. перешел в Берлин; в 1875 г. избран членом Берлинской академии наук, с 1862 г. состоял членом-корреспондентом Санкт-Петербургской академии наук.*

*Г. Кирхгоф, будучи прекрасным знатоком математики, обладал редким умением плодотворно прилагать эти знания к труднейшим вопросам математической физики.*

*Умер в Берлине 17 октября 1887 г.*

В этом случае при пересчете амплитуды и фазы можно пренебречь дифракцией и пользоваться геометрической оптикой.

Тогда (2.3.3) запишется в виде

$$\Gamma(\xi, \eta) = \int_{(x, y)} b_1(x, y) T_d(x, y, \xi, \eta) dx dy, \quad (2.3.5)$$

где  $b_1(x, y)$  – комплексная функция, полученная в результате пересчета амплитуды и фазы поля, отраженного объектом на плоскость  $(x, y)$ ,



касательную к нему и параллельную плоскости наблюдения  $(\xi, \eta)$ ;  $d$  – расстояние между этими двумя плоскостями (рис. 2.12).



Рис. 2.11. Аппроксимация поверхности наблюдения

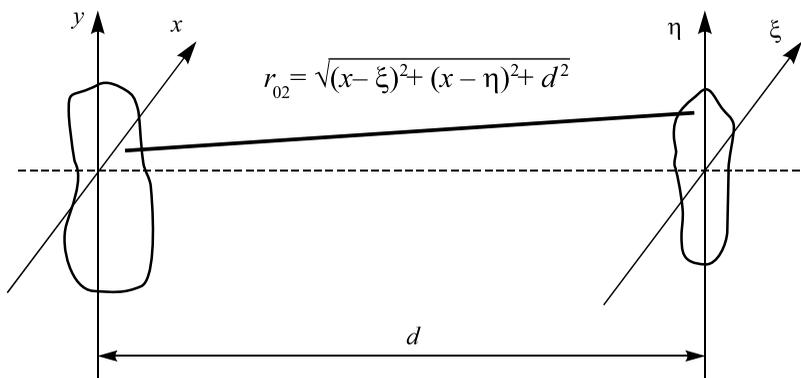


Рис. 2.12. Плоскость, касательная к объекту  $(x, y)$ , и плоскость дифракции  $(\xi, \eta)$

Очевидно, если угол, под которым виден объект с поверхности наблюдения (угол охвата), и площадь наблюдения малы – это естественная аппроксимация. Для задач, где угол охвата должен быть велик, такой подход означает необходимость сведения их к задаче расче-



та при малом угле охвата. При этом для реализации больших углов охвата поверхность наблюдения можно разбить на небольшие фрагменты, аппроксимируемые плоскостями, и рассматривать голограммы для отдельных фрагментов, каждая из которых представляет часть общего угла и воспроизводит объект под своим ракурсом. То есть объект нужно представить набором граней, и для каждой грани просчитать свою голограмму.

В соответствии с принципом Гюйгенса поле в точке  $P$  новой волновой поверхности находится как суперпозиция полей, приходящих в эту точку от различных точек первичной поверхности.

Ядро преобразования (2.3.5), связывающее распределение света на двух параллельных плоскостях, имеет вид

$$T_d(x, y, \xi, \eta) = \frac{\exp\left(-i \frac{2\pi}{\lambda} r(x, y, \xi, \eta, d)\right)}{r(x, y, \xi, \eta, d)} K(\alpha). \quad (2.3.6)$$

Интеграл (2.3.5) с ядром (2.3.6) является математической формулировкой принципа Гюйгенса и называется дифракционным интегралом или интегралом Гюйгенса–Френеля.

Множитель  $\frac{\exp\left(-i \frac{2\pi}{\lambda} r(x, y, \xi, \eta, d)\right)}{r(x, y, \xi, \eta, d)}$  описывает распространение

элементарной вторичной сферической волны от поверхности  $(x, y)$  к поверхности  $(\xi, \eta)$ .

Множитель  $K(\alpha)$  описывает зависимость амплитуды излучаемых вторичных волн от угла дифракции  $\alpha$ . Угол дифракции – это угол между нормалью к плоскости наблюдения и вектором  $r_{01}$  (рис. 2.13).

Независимо от формы дифракционного интеграла начальное значение этой функции  $K(0) = \frac{1}{i\lambda} = -\frac{i}{\lambda}$ . Френель полагал, что  $K(\alpha)$  убывает от начального значения до нуля при изменении угла от 0 до  $\frac{\pi}{2}$ ,

но окончательное выражение не определил. Кирхгоф рассмотрел распространение сферических волн через отверстия, решая уравнение



Гельмгольца с учетом граничных условий, описывающих свойства экрана и отверстия, и пришел к выводу, что

$$K(\alpha) = \frac{1}{i\lambda} \left( \frac{1 + \cos(\alpha)}{2} \right). \quad (2.3.7)$$

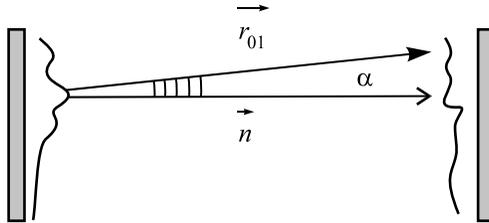


Рис. 2.13. Угол дифракции  $\alpha$  между нормалью к плоскости наблюдения  $\vec{n}$  и вектором  $r_{01}$

В формулировке Рэля–Зоммерфельда множитель  $K(\alpha)$  определяется следующим образом:

$$K(\alpha) = \frac{1}{i\lambda} \cos(\alpha). \quad (2.3.8)$$

С учетом (2.3.8) ядро преобразования будет выглядеть следующим образом:

$$T_d(x, y, \xi, \eta) = \frac{1}{i\lambda} \frac{\exp \left[ i2\pi\lambda^{-1} \sqrt{(x-\xi)^2 + (y-\eta)^2 + d^2} \right]}{\sqrt{(x-\xi)^2 + (y-\eta)^2 + d^2}} \cos(\bar{n}, \overline{r_{01}}), \quad (2.3.9)$$

где  $\lambda$  – длина волны излучения,  $r_{01} = \sqrt{(x-\xi)^2 + (y-\eta)^2 + d^2}$ .

Применение дифракционного интеграла с ядром (2.2.9) для описания распространения поля имеет свои ограничения. Для анализа дифракционных картин в области, прилегающей к касательной плоскости (рис. 2.11), когда расстояние между плоскостями меньше длины волны, следует непосредственно искать решение волнового



уравнения с учетом жестких граничных условий, описывающих свойства объекта.

Для расчета дифракционных картин на достаточном удалении от экрана с отверстием, когда расстояние намного больше длины волны, широко применяются приближенные и более простые решения задачи дифракции: приближение Френеля и приближение Фраунгофера.

**Второе упрощение.** Будем предполагать, что геометрические размеры объекта намного меньше, чем расстояние  $d$  до плоскости наблюдения. По мере удаления от поверхности предмета кривизна вторичных сферических волн уменьшается (рис. 2.14). На некотором расстоянии  $d$  сферические волны можно с высокой точностью аппроксимировать параболическими волнами.

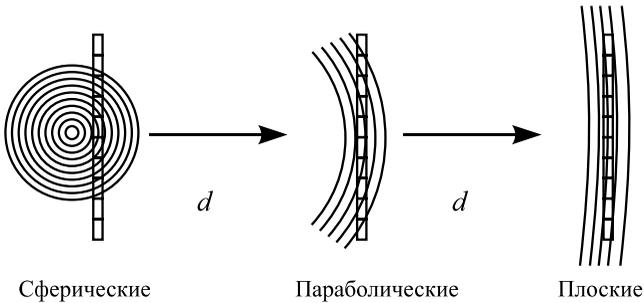


Рис. 2.14. Распространение сферической волны

Кроме того, предположим, что в плоскости наблюдения рассматривается только конечная область вблизи оси  $z$  и расстояние  $d$  много больше максимального размера этой области. С учетом этих допущений можно считать, что с точностью не менее 5 %  $\cos(\bar{n}, \bar{r}_{01}) \approx 1$ , если угол  $(\bar{n}, \bar{r}_{01}) \approx 1$  не превышает  $18^\circ$ :

$$T_d(x, y, \xi, \eta) \approx \frac{1}{i\lambda d} \exp \left[ \frac{i2\pi}{\lambda} \sqrt{(x-\xi)^2 + (y-\eta)^2 + d^2} \right]. \quad (2.3.10)$$

Величину  $r_{01}$  в знаменателе (2.3.9) можно заменить на  $d$ . Однако величину  $r_{01}$ , входящую в экспоненту, нельзя просто заменить на  $d$ , поскольку число  $K = 2\pi\lambda^{-1}$ , которое входит как множитель в экспо-



ненту, очень большое. Поэтому используем аппроксимацию корня разложением его в ряд.

Точная формула для расстояния имеет вид [3]

$$\begin{aligned} r_{01} &= \sqrt{(x-\xi)^2 + (y-\eta)^2 + d^2} = \\ &= d \sqrt{\left(\frac{x-\xi}{d}\right)^2 + \left(\frac{y-\eta}{d}\right)^2 + 1}. \end{aligned} \quad (2.3.11)$$

Для аппроксимации удобно воспользоваться разложением квадратного корня в ряд:

$$\sqrt{1+b} = 1 + \frac{1}{2}b - \frac{1}{8}b^2 + \dots, \quad |b| < 1. \quad (2.3.12)$$

Предположим, что квадратный корень достаточно хорошо аппроксимируется первыми двумя членами этого разложения. Тогда

$$\begin{aligned} r_{01} &\approx d \sqrt{\left(\frac{x-\xi}{d}\right)^2 + \left(\frac{y-\eta}{d}\right)^2 + 1} \approx \\ &\approx d \left[ 1 + \frac{1}{2} \left(\frac{x-\xi}{d}\right)^2 + \frac{1}{2} \left(\frac{y-\eta}{d}\right)^2 \right]. \end{aligned} \quad (2.3.13)$$

Это приближение называется приближением Френеля. С учетом этого приближения выражение (2.3.9) для ядра преобразования примет вид

$$T_d(x, y, \xi, \eta) \approx \frac{1}{i\lambda d} \exp\left[\frac{i2\pi d}{\lambda}\right] \exp\left\{\frac{i\pi[(x-\xi)^2 + (y-\eta)^2]}{\lambda d}\right\}. \quad (2.3.14)$$

Когда расстояние  $d$  достаточно велико для того, чтобы это выражение было достаточно точным. Говорят, что наблюдатель находится в области *дифракции Френеля*.

Приближение О. Френеля описывает параксиальное распространение пучков, т. е. пучков, амплитуда которых незначительно меняется при удалении от оси распространения. Амплитуда такого пучка в поперечном распределении хорошо аппроксимируется функцией Гаусса.



*Огюстен Жан Френель (фр. Augustin-Jean Fresnel; 10 мая 1788 г. – 14 июля 1827 г.), французский физик, один из создателей волновой теории света.*

*В 1811 г. стал самостоятельно изучать физику и вскоре начал экспериментальные исследования по оптике. В 1815 г. переоткрыл принцип интерференции, проделав по сравнению с Т. Юнгом несколько новых опытов. В 1816 г. дополнил принцип Гюйгенса, введя представление о когерентной интерференции элементарных волн, излучаемых вторичными источниками (принцип Гюйгенса–Френеля). Исходя из этого принципа, в 1818 г. разработал теорию дифракции света, на основе которой предложил метод расчета дифракционной картины, основанный на разбиении фронта волны на зоны (так называемые зоны Френеля). В 1821 г. независимо от Т. Юнга доказал поперечность световых волн. В 1823 г. установил законы изменения поляризации света при его отражении и преломлении (формулы Френеля).*

*Умер в возрасте 39 лет от туберкулеза.*

После разложения квадратичных членов в экспоненте (2.3.14) получим запись математической голограммы (2.3.5) в виде

$$\Gamma(\xi, \eta) = \frac{1}{i\lambda d} \exp\left[i\frac{2\pi d}{\lambda}\right] \exp\left[i\frac{\pi(\xi^2 + \eta^2)}{\lambda d}\right] \times \\ \times \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} b_1(x, y) \exp\left[i\frac{\pi(x^2 + y^2)}{\lambda d}\right] \exp\left[-i\frac{2\pi(x\xi + y\eta)}{\lambda d}\right] dx dy. \quad (2.3.15)$$



В (2.3.15) интегрирование ведется по бесконечным пределам, поскольку предполагается, что функция  $b_1(x, y)$  конечна.

Таким образом, с точностью до фазового множителя функцию  $\Gamma(\xi, \eta)$  можно найти как фурье-образ функции

$$b_1(x, y) \exp\{i\pi(\lambda d)^{-1}[x^2 + y^2]\}$$

для частот  $v_x = \frac{\xi}{\lambda d}$  и  $v_y = \frac{\eta}{\lambda d}$ , что обеспечит правильный масштаб в области наблюдения.

Если в разложении квадратного корня (2.3.13) сохраняем два члена ряда, то величина ошибки при вычислении  $r_{01}$  не превосходит по модулю третьего члена ряда:

$$\Delta r_{01} = d \left\{ \frac{1}{8d^4} \left[ (x - \xi)^2 + (y - \eta)^2 \right]^2 \right\}. \quad (2.3.16)$$

При аппроксимации корня сферические волны заменены поверхностями второго порядка. Такая аппроксимация накладывает достаточно большие ограничения на размер объекта, расстояние  $d$  и размеры области наблюдения.

Можно потребовать, чтобы максимальное изменение фазы при рассеивании света от объекта, вносимое членом более высокого порядка, было много меньше, чем 1 рад:

$$\begin{aligned} & \exp \left[ \frac{i2\pi d}{\lambda} \left\{ \frac{1}{8d^4} \left[ (x - \xi)^2 + (y - \eta)^2 \right]^2 \right\} \right] = \\ & = \exp \left[ \frac{i\pi}{4\lambda d^3} \left[ (x - \xi)^2 + (y - \eta)^2 \right]^2 \right] \ll 1. \end{aligned} \quad (2.3.17)$$

Это условие будет выполнено при

$$d^3 \gg \frac{\pi}{4\lambda} \left[ (x - \xi)^2 + (x - \eta)^2 \right]_{\max}^2. \quad (2.3.18)$$



Однако это требование обычно не является необходимым. Для справедливости приближения Френеля требуется только, чтобы члены более высокого порядка не меняли значения интеграла в (2.3.15). Для этого необязательно, чтобы максимальное значение фазовых множителей были много меньше, чем 1 рад. В случае, если  $d$  мало и условие (2.3.18) не выполняется, величина  $\pi(\lambda d)^{-1}$  будет такой большой, а колебания квадратичного множителя в (2.3.14) такими быстрыми, что основной вклад в интеграл будут давать только точки, расположенные в центре, где скорость изменения фазы минимальна. В окрестностях таких точек величиной фазового члена следующего более высокого порядка можно пренебречь.

**Третье упрощение.** Расчет дифракционной картины становится еще проще, если принять еще более жесткие ограничения. В этом случае сферические волны можно аппроксимировать плоскими волнами.

В области дифракции Френеля распределение волнового фронта можно найти как фурье-образ функции  $b_1(x, y)$ , умноженный на квадратичную фазовую функцию  $\exp\left[i\frac{\pi(x^2 + y^2)}{\lambda d}\right]$ . Если принять более строгое предположение Фраунгофера

$$d \gg \frac{\pi}{\lambda}(x^2 + y^2)_{\max}, \quad (2.3.19)$$

то квадратичный фазовый множитель становится примерно равным единице. Таким образом, в области *дифракции Фраунгофера* можно записать

$$\Gamma(\xi, \eta) = \frac{1}{i\lambda d} \exp\left[i\frac{2\pi d}{\lambda}\right] \exp\left[i\frac{\pi(\xi^2 + \eta^2)}{\lambda d}\right] \times \\ \times \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} b_1(x, y) \exp\left[-i\frac{2\pi(x\xi + y\eta)}{\lambda d}\right] dx dy. \quad (2.3.20)$$



Без учета фазовых множителей, стоящих перед интегралом, это выражение представляет собой фурье-образ функции  $b_1(x, y)$  для ча-

стот  $v_x = \frac{\xi}{\lambda d}$  и  $v_y = \frac{\eta}{\lambda d}$  :

$$\mathfrak{F}_{b_1}(v_x, v_y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} b_1(x, y) \exp\{-i2\pi[xv_x + yv_y]\} dx dy. \quad (2.3.21)$$



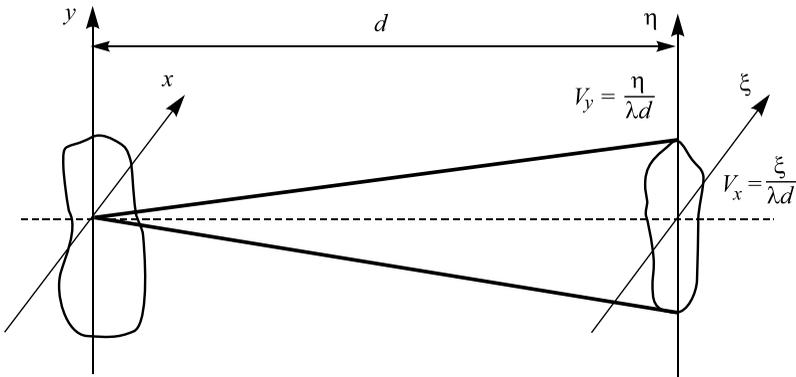
*Фраунгофер Йозеф (Fraunhofer, Joseph von) (1787–1826), немецкий физик-оптик. В возрасте 14 лет провел несколько дней под завалом дома, благодаря этому несчастному случаю приобрел некоторую популярность. В частности, получил от властей Баварии денежную компенсацию, на которую открыл собственное стекольное дело. С 1806 г. работал в оптической мастерской, в 1809 г. стал ее управляющим, в 1818 г. – руководителем и владельцем. С 1823 г. – профессор Мюнхенского университета. Исследовал явление дисперсии и достиг успехов в изготовлении ахроматических линз. Изобрел метод точного определения формы линз и машину для шлифования ахроматических линз; сконструировал спектрометр, ахроматический микроскоп, окулярный микрометр и гелиометр. Впервые наблюдал, исследовал и объяснил темные линии в солнечном спектре и измерил их длину волны (1814–1815, независимо от Волластона, фраунгоферовы линии). Изучил дифракцию в параллельных лучах (дифракция Фраунгофера). Широко использовал дифракционные решетки.*



Й. Фраунгофер скончался 7 июня 1826 г. на 39-м году жизни. На его могиле сделана короткая, но выразительная надпись: «*Approximavit sidera*» («Приблизил звезды»).

Таким образом, при прохождении оптического волнового фронта через свободное пространство на некотором расстоянии с точностью до фазовых множителей с ним происходит преобразование Френеля, которое при увеличении расстояния переходит в преобразование

Фурье с частотами  $v_x = \frac{\xi}{\lambda d}$  и  $v_y = \frac{\eta}{\lambda d}$  (рис. 2.15).



$$\Gamma(\xi, \eta) = F_b \left( V_x = \frac{\xi}{\lambda d}, V_y = \frac{\eta}{\lambda d} \right)$$

Рис. 2.15. Пространственные частоты в области формирования голограммы

С учетом (2.3.21) преобразование в области Френеля можно записать как

$$\Gamma(\xi, \eta) = \frac{1}{i\lambda d} \exp \left[ i \frac{2\pi d}{\lambda} \right] \exp \left[ i \frac{\pi(\xi^2 + \eta^2)}{\lambda d} \right] \times \mathfrak{F} \left[ b_1(x, y) \exp \{ i\pi(\lambda d)^{-1} [x^2 + y^2] \} \right]. \quad (2.3.22)$$



Преобразование в области Фраунгофера

$$\Gamma(\xi, \eta) = \frac{1}{i\lambda d} \exp\left[i\frac{2\pi d}{\lambda}\right] \exp\left[i\frac{\pi(\xi^2 + \eta^2)}{\lambda d}\right] \mathfrak{F}[b_1(x, y)], \quad (2.3.23)$$

где  $\mathfrak{F}_b(v_x, v_y) = \mathfrak{F}[b_1(x, y)]$  – фурье-образ функции  $b(x, y)$ .

Таким образом, волновой фронт в плоскости формирования голограммы на расстоянии  $d$  от объекта может быть получен с помощью двумерных преобразований Френеля или Фурье, над волновым полем, рассеиваемым объектом, в касательной плоскости (рис. 2.16).

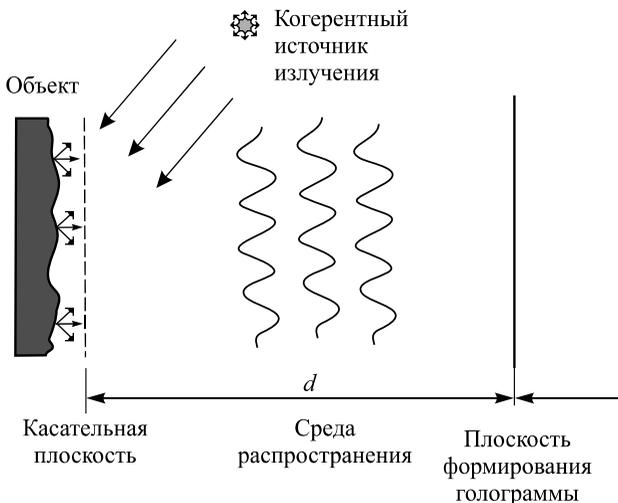


Рис. 2.16. Распространение света от плоскости объекта до плоскости формирования голограммы

В области оптических частот условия, необходимые для выполнения дифракции Фраунгофера, могут быть достаточно жесткими [12]. Например, при длине волны  $0,6 \cdot 10^{-6}$  м (красный свет) и ширине отверстия 2,5 см расстояние  $d$  должно удовлетворять условию  $d \gg 1600$  м. Тем не менее в ряде случаев условия выполняются при меньших расстояниях (например, если объект освещается волной, сходящейся к наблюдателю, или если между плоскостью наблюдения и объектом расположена собирающая линза).



Голограмма возникает в результате интерференции между волновым фронтом, отраженным от объекта и опорной волной. В результате интерференции когерентных монохроматических волн возникает устойчивая во времени картина, интенсивность которой удастся зафиксировать.

## 2.4. ИНТЕРФЕРЕНЦИЯ ВОЛН

**Интенсивность света.** Средняя частота колебаний электромагнитного поля в оптической области спектра составляет около  $10^{15}$  Гц. Всем приемникам света присуща определенная инерционность. Ее можно охарактеризовать временем установления или разрешения приемника. Так, для глаза это около 0,1 с. Глаз не замечает быстрых миганий света, если они следуют друг за другом через малые по сравнению с этим временем промежутки. Существуют приемники, время разрешения которых гораздо меньше. В ячейках Керра это время может быть доведено до  $10^{-8} \dots 10^{-9}$  с. Наиболее быстродействующие фотоэлектрические приемники имеют время разрешения порядка  $10^{-10}$  с. [13]. Однако в настоящее время ни один приемник света не позволяет измерить мгновенное значение напряженности электрического или магнитного поля в световой волне. Все приемники могут измерять только некоторые величины, усредненные за время, не меньшее времени разрешения приемника. Наблюдаемой величиной является поток энергии, который прямо пропорционален квадрату амплитуды электрического поля волны. Физическую величину, пропорциональную квадрату амплитуды электрического поля волны, принято называть интенсивностью:  $I \sim E_0^2$ .

Численно интенсивность равна усредненной за период колебаний волны мощности излучения, проходящей через единичную площадку, расположенную перпендикулярно направлению распространения энергии. В математической форме это может быть выражено следующим образом:

$$I(t) = \frac{1}{T} \int_t^{t+T} \frac{dP}{dS} dt, \quad (2.4.1)$$

где  $T$  – период волны;  $dP$  – мощность, переносимая волной через площадку  $dS$ .



Электромагнитные волны переносят энергию электромагнитного поля, поток которой определяется величиной вектора Пойнтинга [14]. Интенсивность электромагнитного излучения равна усредненному за период значению модуля вектора Пойнтинга:

$$I(t) = \frac{1}{T} \int_t^{t+T} |\vec{S}(t)| dt, \quad (2.4.2)$$

где вектор Пойнтинга выражается через напряженность магнитного поля и электромагнитную индукцию:

$$\vec{S}(t) = \frac{c}{4\pi} [\vec{E}(t)\vec{B}(t)]. \quad (2.4.3)$$

Для монохроматической линейно поляризованной волны с амплитудой напряженности электрического поля  $E_0$  интенсивность равна

$$I(t) = \frac{cE_0^2}{8\pi}. \quad (2.4.4)$$

Единицей измерения интенсивности в Международной системе единиц (СИ) является Вт/м<sup>2</sup>.

*Интерференция волн* – явление усиления или ослабления амплитуды результирующей волны в зависимости от соотношения между фазами складывающихся в пространстве двух (или нескольких) волн.

При рассмотрении интерференции ниже предполагается, что распространяющиеся от источников волны являются строго монохроматическими.

Интерференция может наблюдаться только при выполнении условия когерентности (пространственной и временной стабильности интерферирующих волновых фронтов). *Когерентность* (от лат. *cohaerens* – находящийся в связи) – согласованное протекание во времени и в пространстве нескольких колебательных или волновых процессов, проявляющееся при их сложении. Колебания называются *когерентными*, если разность их фаз остается постоянной (или закономерно изменяется) во времени и при сложении колебаний определяет амплитуду суммарного колебания. Тогда характерное распределение амплитуд с чередующимися максимумами и минимумами остается неподвижным в пространстве. Нарушение когерентности может происходить вследствие несовпадения частот складывающихся волн или вследствие того, что фазы обеих волн меняются независимо в резуль-



тате происходящих в самих излучателях процессов (например, при возникновении в двух излучателях колебаний с различными начальными фазами) либо на путях распространения (например, при изменении скорости распространения волн из-за флуктуаций плотности среды).

Если свет испускается реальным источником, то за время измерения (усреднения) колебания могут случайно обрываться или может происходить хаотическое изменение их фазы. Интерференция излучения от независимых источников излучения (например, от электрических лампочек) недоступна наблюдению глазом. Возбужденный атом испускает цуг волн в течение времени, характерная длительность которого около  $10^{-8}$  с. В таком цуге содержится  $10^6 \dots 10^8$  волн. В результате различных процессов (например, столкновений с другими атомами или ударов электронов) атом может снова вернуться в возбужденное состояние и начать излучать новый цуг волн. Таким образом, получится последовательность цугов, испускаемых атомом через малые и нерегулярно изменяющиеся периоды времени. В течение секунды система интерференционных полос будет меняться десятки и сотни миллионов раз. Глаз или другой приемник света фиксирует усредненную равномерную освещенность.

Смещения интерференционных полос при замене одной пары цугов на другую происходить не будет, если у цугов начальные фазы меняются, хотя и нерегулярно по времени, но одинаково. Чтобы этого достигнуть, надо излучение от одного и того же источника расщепить на два или несколько пучков и заставить попадать их на экран различными путями. Если при этом будут накладываться пучки из одного и того же цуга, то интерференционные полосы на экране будут сохранять свое положение и глаз зафиксирует устойчивую (неподвижную) картину.

Введение оптической разности хода эквивалентно задержке одной из волн во времени, поэтому способность световых колебаний к интерференции после разделения исходной волны на две и последующему их соединению с некоторой разностью хода называется *временной когерентностью*.

Максимальная разность хода, при которой возможна интерференция, называется *длиной когерентности* излучения, а соответствующее ей время запаздывания одного из пучков – *временем когерентности*. Условием интерференции волн одной и той же частоты является их



когерентность, т. е. сохранение неизменной разности фаз за время, достаточное для наблюдения (т. е.  $t \gg T$ ).

Когда в пространстве распространяются две волны, то в случае, если среда, в которой они распространяются, линейна (т. е. отсутствует зависимость свойств среды от происходящих в ней процессов), результирующее колебание представляет собой геометрическую сумму колебаний, соответствующих каждой из складываемых волн. Рассмотрим интерференцию двух когерентных волн с напряженностью поля  $\vec{E}_r$  и  $\vec{E}_p$ . Напряженность результирующего поля определяется выражением

$$\vec{E}_s = \vec{E}_p + \vec{E}_r. \quad (2.4.5)$$

Основные оптические детекторы реагируют лишь на усредненную интенсивность принимаемого излучения, которая определяется как

$$I \sim \vec{E}_s \vec{E}_s^*. \quad (2.4.6)$$

Напряженность электрического поля является комплексной величиной, которая зависит от пространственных координат и времени

$$\vec{E}(x, y, t) = \vec{A}(x, y) \exp(-i\omega t). \quad (2.4.7)$$

Комплексная пространственная амплитуда

$$\vec{A}(x, y) = A(x, y) \exp(-i\varphi(x, y)), \quad (2.4.8)$$

где  $A(x, y)$  – модуль напряженности поля, а  $\varphi(x, y)$  – фаза.

Подставляя (2.4.5) в (2.4.6), интенсивность может быть записана как

$$I = \vec{E}_s \vec{E}_s^* = (\vec{E}_p + \vec{E}_r) (\vec{E}_p^* + \vec{E}_r^*). \quad (2.4.9)$$

Так как пучки когерентны, они имеют одну и ту же зависимость от времени, выражение (2.3.8) примет вид (поскольку  $\exp(-i\omega t) \times \exp(i\omega t) = 1$ )

$$I = (\vec{A}_p + \vec{A}_r) (\vec{A}_p^* + \vec{A}_r^*). \quad (2.4.10)$$



Подставляя (2.3.8) в (2.3.10), получим выражение для результирующей интенсивности двух когерентных монохроматических пучков в плоскости  $(x, y)$

$$\begin{aligned} I &= \left( A_p e^{-i\varphi_p} + A_r e^{-i\varphi_r} \right) \left( A_p e^{i\varphi_p} + A_r e^{i\varphi_r} \right) = \\ &= A_p^2 + A_p A_r e^{i(\varphi_p - \varphi_r)} + A_p A_r e^{-i(\varphi_p - \varphi_r)} + A_r^2. \end{aligned} \quad (2.4.11)$$

Выражение (2.3.11) можно преобразовать в вид

$$\begin{aligned} I(x, y) &= A_p^2(x, y) + A_r^2(x, y) + \\ &+ 2A_p(x, y)A_r(x, y) \cos(\varphi_p(x, y) - \varphi_r(x, y)). \end{aligned} \quad (2.4.12)$$

После подстановки выражений для средней интенсивности  $I_0 = A_p^2 + A_r^2$  и для модуляции интенсивности (видности полос)

$$V = \frac{2A_p A_r}{A_p^2 + A_r^2} \text{ получим}$$

$$I(x, y) = I_0(x, y) \left( 1 + V(x, y) \cos(\varphi_p(x, y) - \varphi_r(x, y)) \right). \quad (2.4.13)$$

Выражения (2.4.12) и (2.4.13) представляют базовое уравнение двухлучевой интерферометрии. Распределение интенсивности в плоскости  $(x, y)$  называется *интерференционной картиной*, или *интерферограммой*.

Впервые явление интерференции было независимо обнаружено Р. Бойлем (1627–1691) и Р. Гуком (1635–1703). Они наблюдали возникновение разноцветной окраски тонких пленок (интерференционных полос), подобных масляным или бензиновым пятнам на поверхности воды. В 1801 г. Т. Юнг (1773–1829), введя «принцип суперпозиции», первым объяснил явление интерференции света, ввел в научный обиход термин «интерференция» (1803 г.). Он также выполнил первый демонстрационный эксперимент по наблюдению интерференции света, получив интерференцию от двух щелевых источников света (1802 г.).



Томас Юнг (1773–1829) – английский ученый, один из основоположников волновой теории света. В двухлетнем возрасте он научился читать, в девятилетнем – изучил латинский и греческий языки, а к 14 годам в совершенстве знал около десяти языков, в том числе древнееврейский, персидский и арабский.

Волновая теория света впервые была сформулирована Юнгом в 1801 г. Ее сущность он кратко выражает следующим предложением: «Излучаемый свет состоит из волнообразных движений светоносного эфира».

Впервые объяснил аккомодацию глаза изменением кривизны хрусталика (1793), сформулировал принцип интерференции (1801), ввел термин «интерференция» (1802), измерил длины волн разных цветов (1803), объяснил кольца Ньютона по волновой теории, высказал идею о поперечности световых волн (1817), разработал теорию цветового зрения, исследовал деформацию сдвига и ввел характеристику упругости – модуль Юнга.

Всей своей жизнью ученый оправдал прозвище «феноменальный», данное ему в годы учебы в Кембриджском университете.

Пусть интерферируют две плоские волны, которые имеют одинаковые частоты и одинаково линейно поляризованы:

$$E_1(x, t) = E_0 \cos(\omega t - \vec{k}_1 \vec{r} + \varphi_{01}) = E_0 \cos(\omega t + \varphi_1); \quad (2.4.14a)$$

$$E_2(x, t) = E_0 \cos(\omega t - \vec{k}_2 \vec{r} + \varphi_{02}) = E_0 \cos(\omega t + \varphi_2). \quad (2.4.14b)$$



Разность фаз будет равна

$$\varphi_1 - \varphi_2 = (\vec{k}_1 \vec{r} - \vec{k}_2 \vec{r}) + (\varphi_{01} - \varphi_{02}) = \vec{K} \vec{r} + (\varphi_{01} - \varphi_{02}), \quad (2.4.15)$$

где вектор  $\vec{K} = \vec{k}_1 - \vec{k}_2$ ,  $\alpha$  – угол между интерферирующими пучками (рис. 2.17).

Поверхности равных разностей фаз  $\varphi_1 - \varphi_2 = \text{const}$  – параллельные плоскости, перпендикулярные к вектору  $\vec{K}$ , на рис. 2.17 обозначены пунктирными отрезками. Вдоль этих плоскостей интенсивность результирующего колебания будет постоянна. Расстояние между двумя соседними плоскостями максимальной или минимальной освещенности найдется из условия  $K \Delta x = 2\pi$ . Так как длины обоих волновых векторов одинаковы и  $k = 2\pi/\lambda$ , то  $K = 2k \sin(\alpha/2)$ , и расстояние между интерференционными полосами определяется из условия

$$\Delta x = \frac{2\pi}{K} = \frac{2\pi}{2k \sin(\alpha/2)} = \frac{\lambda}{2 \sin(\alpha/2)}. \quad (2.4.16)$$

Если поставить плоский экран, то он пересечет плоскости равной интенсивности вдоль параллельных прямых. На экране появятся светлые и темные интерференционные полосы (рис. 2.18).

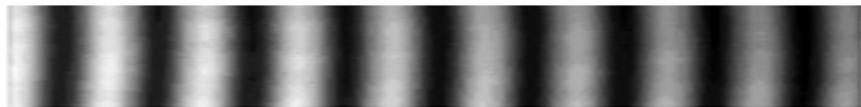


Рис. 2.18. Интерференционные полосы

Расстояние между серединами соседних светлых или темных полос называется *шириной интерференционной полосы*. Например, если угол между интерферирующими плоскими волнами равен  $30^\circ$  и  $\lambda = 0,6$  мкм, то ширина интерференционной полосы составит около 1 мкм.

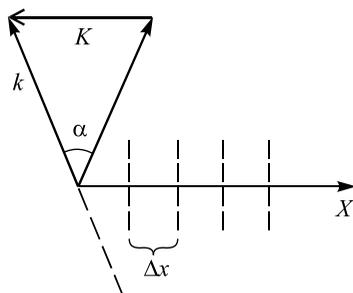


Рис. 2.17. Интерференция двух плоских волн



Когерентность лазерного излучения обеспечивает устойчивую во времени интерференционную картину, что важно при записи. Для того чтобы полосы были видны невооруженным глазом, необходимо выбирать схемы интерферометров, обеспечивающих небольшие углы между интерферирующими пучками.

Явления интерференции находят важные применения при решении измерительных задач как в научных исследованиях, так и в технике. Они основаны на том, что между длиной волны, разностью хода интерферирующих волн и расположением интерференционных максимумов и минимумов существует определенная связь. Поэтому, зная разность хода интерферирующих волн, можно по интерференционной картине определить длину волны или, наоборот, зная длину волны, определить разность хода лучей. Связь между распределением интенсивностей в интерференционной картине и разностью хода интерферирующих волн зависит от конкретной оптической схемы – интерферометра. Для обеспечения условия когерентности луч лазера делится на два пучка. Первый пучок попадает на светочувствительную среду без изменений и называется *опорным*, а второй отражается от объекта и называется *объектным*.

Теперь можно рассмотреть следующие этапы: получение голограмм, их регистрацию и восстановление из них волновых фронтов, соответствующих действительному и мнимому изображениям.

## 2.5. РЕГИСТРАЦИЯ ГОЛОГРАММ

Основные приемные устройства оптического сигнала реагируют лишь на интенсивность принимаемого излучения. Голограмма – это результат интерференции двух волновых фронтов. Один из фронтов – объектный – отражается от объекта, другой – опорный – проходит по другому пути.

В результате прохождения некоторого пути после отражения от поверхности объекта над объектным фронтом происходит некоторое преобразование, которое описывается интегралом Френеля или Фурье, в зависимости от расстояния  $d$ . Этот волновой фронт будет складываться с опорным (рис. 2.19) в плоскости голограммы.

В результате интерференции объектного и опорного фронтов образуется голограмма.

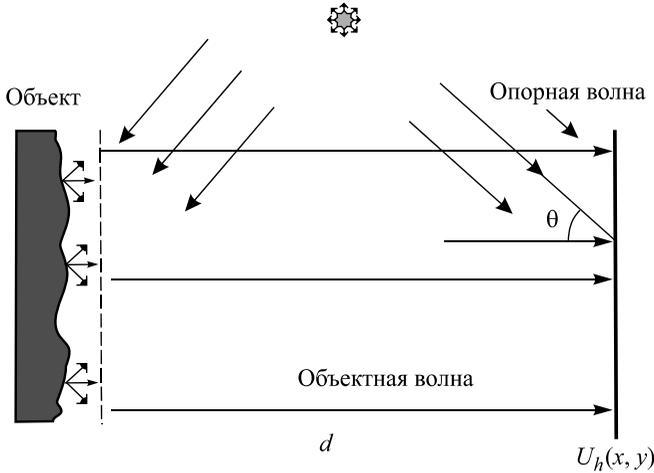


Рис. 2.19. Образование голографической картины как результат интерференции объектного и опорного волновых фронтов

Пусть координатная система голографической пластинки будет двумерной и обозначена  $(x, y)$ , тогда комплексная амплитуда объектной волны может быть записана как

$$U_0(x, y) = E_0(x, y) \exp[i\varphi_0(x, y)], \quad (2.5.1)$$

где  $E_0$  – амплитуда, а  $\varphi_0$  – фаза объектной волны. Если расстояние  $d$  соответствует зоне дифракции Френеля, то мы можем вычислить распределение  $U_0(x, y)$  с помощью интеграла Френеля, если зоне Фраунгофера, то с помощью интеграла Фурье.

Аналогично для опорной волны можно записать

$$U_r(x, y) = E_r(x, y) \exp[i\varphi_r(x, y)], \quad (2.5.2)$$

где  $E_r$  – амплитуда, а  $\varphi_r$  – фаза опорной волны. В качестве опорной волны наиболее часто используются плоские или сферические волны.

Когда объектная и опорная волны интерферируют на поверхности светочувствительной среды, образуется интерференционная картина, которая может быть представлена следующим образом:



$$\begin{aligned}
 U_h(x, y) &= |U_0(x, y) + U_r(x, y)|^2 = |U_0(x, y) + U_r(x, y)| \times \\
 &\quad \times |U_0(x, y) + U_r(x, y)|^* = |U_0(x, y)|^2 + |U_r(x, y)|^2 + \\
 &\quad + U_0(x, y)U_r^*(x, y) + U_0^*(x, y)U_r(x, y) = E_0^2(x, y) + E_r^2(x, y) + \\
 &\quad + U_0(x, y)U_r^*(x, y) + U_0^*(x, y)U_r(x, y), \quad (2.5.3)
 \end{aligned}$$

где  $*$  обозначает комплексно-сопряженную величину. Это выражение можно представить в следующем виде:

$$\begin{aligned}
 U_h(x, y) &= E_0^2(x, y) + E_r^2(x, y) + E_0(x, y)E_r(x, y)e^{i[\varphi_0(x, y) - \varphi_r(x, y)]} + \\
 &\quad + E_0(x, y)E_r(x, y)e^{-i[\varphi_0(x, y) - \varphi_r(x, y)]} = E_0^2 + E_r^2 + E_0E_r \exp \times \\
 &\quad \times \{i[\varphi_0(x, y) - \varphi_r(x, y)]\} + E_0E_r \exp \{-i[\varphi_0(x, y) - \varphi_r(x, y)]\}. \quad (2.5.4)
 \end{aligned}$$

Значение интенсивности, формирующее голограмму, будет пропорционально  $U_h(x, y)$ . Предположим, что фотопластинка обработана так, что ее амплитудное пропускание пропорционально регистрируемой интенсивности. Голограмма может быть рассмотрена как синусоидальная дифракционная решетка, попадая на которую опорная волна в результате дифракции формирует восстановленное изображение.

Светочувствительная среда обычно плоская, несмотря на то что содержит информацию об объекте в трехмерном пространстве. В классической голографии в качестве светочувствительной среды используют пластинку, покрытую слоем фотоэмульсии. При формировании голограммы выдерживается экспозиция, зависящая от интенсивности света и свойств эмульсии, затем голограмма проявляется. Процесс проявления схож с процессом проявления фотографий, занимает длительное время и требует наличия определенных химических реагентов.

В цифровой голографии регистратором излучения являются светочувствительные CCD или CMOS матрицы.

Получение и регистрация голограмм производится на оптической установке. Для расшифровки голограмм, т. е. восстановления из нее изображений, используется компьютерная техника.



## 2.6. ВОССТАНОВЛЕНИЕ ВОЛНОВОГО ФРОНТА ИЗ ГОЛОГРАММЫ

Для восстановления голограмму освещают световой волной, аналогичной той, что была использована при записи. В этом случае голограмма работает как дифракционная решетка, и в результате процесса интерференции дифрагированных световых пучков образуется виртуальное изображение записанного объекта.

Изображение находится в той же позиции, что и реальный объект, имеет схожие перспективные искажения и глубину. Вместе с виртуальным изображением объекта (его также называют мнимым) формируются сопряженное ему действительное изображение и постоянная составляющая. От угла отклонения опорного пучка  $\theta$  зависит перекрытие областей разных дифракционных порядков.

Значение интенсивности, формирующее голограмму, будет пропорционально

$$\begin{aligned}
 U_h(x, y) &= E_0^2(x, y) + E_r^2(x, y) + E_0(x, y)E_r(x, y)e^{i[\varphi_0(x, y) - \varphi_r(x, y)]} + \\
 &+ E_0(x, y)E_r(x, y)e^{-i[\varphi_0(x, y) - \varphi_r(x, y)]} = E_0^2 + E_r^2 + E_0E_r \times \\
 &\times \exp\{i[\varphi_0(x, y) - \varphi_r(x, y)]\} + E_0E_r \exp\{-i[\varphi_0(x, y) - \varphi_r(x, y)]\} = \\
 &= E_0^2 + E_r^2 + 2E_0E_r \cos(\varphi_0(x, y) - \varphi_r(x, y)). \quad (2.6.1)
 \end{aligned}$$

Светочувствительная среда обычно плоская. В классической голографии в качестве светочувствительной среды используют пластинку, покрытую слоем фотоэмульсии. При формировании голограммы выдерживается экспозиция, зависящая от интенсивности света и свойств эмульсии, затем голограмма проявляется.

Предположим, что фотопластинка обработана так, что ее амплитудное пропускание равно регистрируемой интенсивности.

Как видно из выражения (2.6.1), голограмма имеет вид плоской синусоидальной амплитудной решетки.

Рассмотрим процесс восстановления для одномерного случая [15]. Пусть полосы распределены в направлении  $y$  (рис. 2.20).

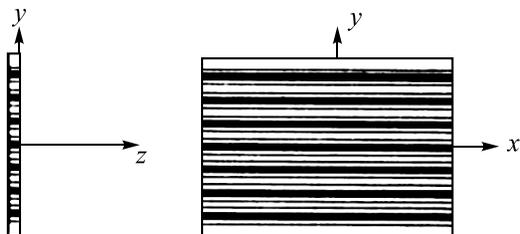


Рис. 2.20. Плоская синусоидальная решетка

Учитывая, что функцию косинус можно представить в виде суммы двух экспонент, для одномерного случая выражение (2.6.1) можно переписать в виде

$$\begin{aligned} U_h(x) &= t_0 + t_1 \cos[2\pi\xi_r y] = \\ &= t_0 + \frac{1}{2}t_1 \exp[-i2\pi\xi_r y] + \frac{1}{2}t_2 \exp[i2\pi\xi_r y]. \end{aligned} \quad (2.6.2)$$

Пусть на решетку падает плоская монохроматическая волна, распространяющаяся в направлении  $z$ . Комплексная амплитуда этой волны определяется как

$$U_0(x, y, z) = A_{r1}(x) e^{-i\frac{2\pi z}{\lambda}} = A_{r1}(x) e^{-i2\pi\xi_r z}. \quad (2.6.3)$$

В этом случае дифракция на плоской синусоидальной решетке разделяется на три компонента (рис. 2.21). Сразу же за голограммой образуются три волны:

$$\begin{aligned} U_0(x, y, 0) &= A_{r1}t_0 + \frac{1}{2}A_{r1}t_2 \exp[i2\pi\xi_y y] + \\ &+ \frac{1}{2}A_{r1}t_1 \exp[-i2\pi\xi_y y]. \end{aligned} \quad (2.6.4)$$

Одна из них распространяется в прямом направлении, а две другие – по направлениям, их угол распространения задается периодом решетки. Далее три волновых фронта распространяются в направлении  $z$ .

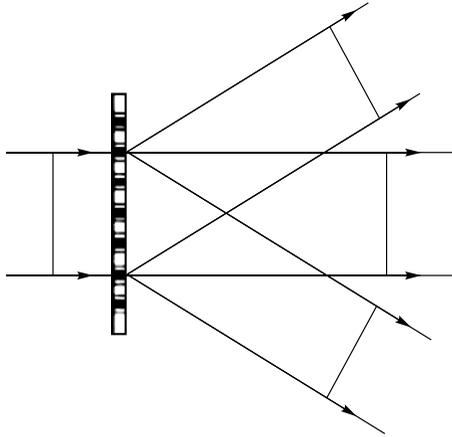


Рис. 2.21. Дифракция на синусоидальной амплитудной решетке (падающая плоская волна разделяется на три компонента)

В общем случае математически процесс дифракции можно записать в виде умножения голограммы  $U_h(x, y)$  (2.6.1) на сопряженный опорный пучок  $U_0^*(x, y) = E_{r1}(x, y) \exp[-i\varphi_r(x, y)]$ :

$$\begin{aligned} U(x, y) &= U_h(x, y)U_0^*(x, y) = U_h(x, y)E_{r1}(x, y) \exp[-i\varphi_r(x, y)] = \\ &= E_0^2 E_{r1} \exp(-i\varphi_r(x, y)) + E_r^2 E_{r1} \exp(-i\varphi_r(x, y)) + \\ &\quad + U_0 E_r E_{r1} \exp\{i[-2\varphi_r(x, y)]\} + U_0^* E_r E_{r1}. \end{aligned} \quad (2.6.5)$$

Если голограмму просветить плоской монохроматической волной, *сопряженной с опорной*, то получим сразу же за плоскостью голограммы (плоскость 1) четыре волны (рис. 2.22).

- Первая и вторая волны распространяются в направлении восстанавливающего пучка. Эти составляющие являются помехой в восстановленном сигнале (нулевой пучок).
- В направлении, перпендикулярном плоскости голограммы, распространяется исходная волна света, которая соответствует мнимому изображению.
- В направлении  $2\varphi_r$  распространяется волна, являющаяся точной копией исходной (действительное изображение).

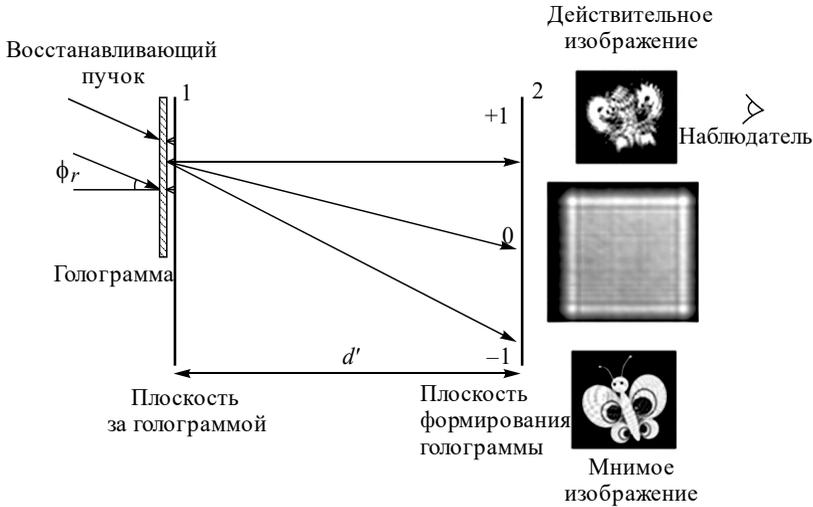


Рис. 2.22. Формирование мнимого и действительного изображений

Для восстановления исходных волновых фронтов в плоскости изображения можно воспользоваться преобразованием Френеля и на некотором расстоянии  $d'$  восстановить точную копию исходной волны. На расстоянии  $d'$  (рис. 2.22) получим действительное, мнимое и изображение нулевого пучка (плоскость 2 на рис. 2.20).

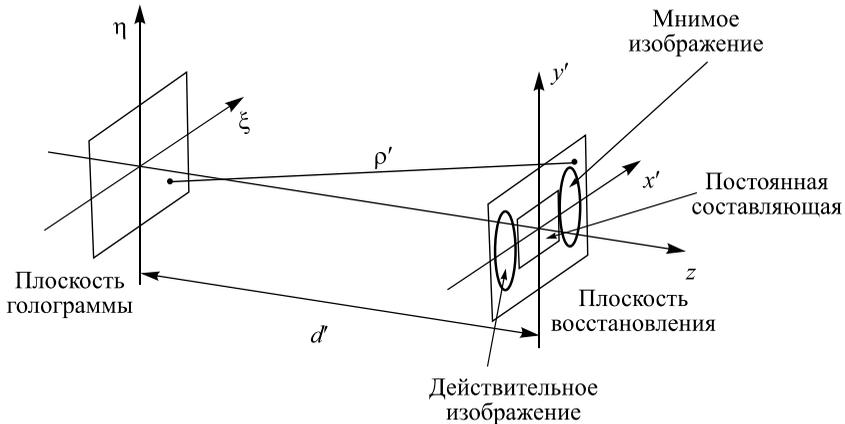


Рис. 2.23. Координатная система при голографическом восстановлении



Если расстояние  $d'$  соответствует расстоянию  $d$  от объекта до голограммы, то волновой пучок, соответствующий действительному изображению, будет точной копией волнового фронта от объекта.

Если опорный пучок падает перпендикулярно плоскости голограммы, а полосы на голограмме расположены в направлении  $x$ , то можно использовать координатную систему, приведенную на рис. 2.23, где плоскость восстановления находится на расстоянии  $d'$  от плоскости голограммы.

Мнимое изображение на расстоянии в области Френеля может быть размытым, а действительное – выглядеть четко (рис. 2.23).

Для того чтобы поменять местами мнимую и действительную части, необходимо осветить голограмму пучком сопряженным опорной волне.

## 2.7. НЕОБХОДИМОЕ ПРОСТРАНСТВЕННОЕ РАЗРЕШЕНИЕ ПРИ РЕГИСТРАЦИИ ГОЛОГРАММЫ

При получении голограмм с опорным наклонным пучком для полного разделения волновых полей регистрирующие среды должны иметь очень высокое разрешение. Пусть мы хотим зарегистрировать интерференционную картину от двух плоских единичных волн, одна из которых перпендикулярна плоскости регистрации, вторая имеет угол  $\alpha$  с нормалью к регистрирующей плоскости, тогда расстояние между полосами будет равно

$$\Delta x = \frac{\lambda}{2 \sin(\alpha / 2)}. \quad (2.7.1)$$

Допустим, что интерферируют две плоские волны под углом  $30^\circ$ , тогда при длине волны  $0,6$  мкм расстояние между полосами  $\Delta x$  будет примерно  $1$  мкм.

По теореме Найквиста–Котельникова для записи голограммы необходим регистрирующий материал с разрешающей способностью, как минимум в два раза большей. Таким образом, при выбранных параметрах необходимо обеспечить не менее  $2000$  лин/мм для регистрации только несущей частоты. В случае записи объектов с диффузной поверхностью разрешение должно быть еще больше.



Для того чтобы отделить волны, формирующие действительное и мнимое изображения от пучка нулевого порядка, необходимо иметь достаточно большой угол между предметным и опорным пучками. Величина этого угла будет определяться размером спектра исходного объекта.

Рассмотрим процесс восстановления для одномерного случая [11, 16]. В этом случае

$$\begin{aligned} U_h &= |E_0(x) \exp[i\varphi_0(x)] + E_r(x) \exp[i2\pi\xi_r x]| \times \\ &\times |E_0(x) \exp[i\varphi_0(x)] + E_r(x) \exp[i2\pi\xi_r x]|^* = \\ &= E_0^2 + E_r^2 + U_0 E_r \exp[-i2\pi\xi_r x] + U_0^* E_r \exp[i2\pi\xi_r x], \end{aligned} \quad (2.7.2)$$

где  $U_0(x) = E_0(x) \exp[i\varphi_0(x)]$  – объектная волна и  $U_r(x) = E_r(x) \exp \times [i2\pi\xi_r x]$  – опорная волна.

Далее зависимость от  $x$  подразумевается и убрана для упрощения визуального представления.

При восстановлении голограмму умножаем на сопряженный опорный пучок:

$$\begin{aligned} U &= U_h E_r \exp[-i2\pi\xi_r x] = \\ &= E_0^2 E_r \exp[-i2\pi\xi_r x] + E_r^3 \exp[-i2\pi\xi_r x] + \\ &+ U_0 E_r^2 \exp[-i4\pi\xi_r x] + U_0^* E_r^2. \end{aligned} \quad (2.7.3)$$

Пусть спектр объектного пучка конечен. Обозначим границы спектра объектного пучка  $S(\xi, \eta)$  и представим его графически на рис. 2.24.

Тогда спектр световой волны после восстановления можно графически представить следующим образом (рис. 2.25).

Первое и второе слагаемые в формуле (2.7.3) образуют постоянную составляющую, находящуюся в нулевом порядке дифракции. Первый член в (2.7.3) формирует спектр

$$E_0^2 E_r \exp[-i2\pi\xi_r x] \rightarrow A^*(\xi, \eta) * A(\xi, \eta),$$



где  $*$  – оператор автокорреляции. Ширина спектра  $A(\xi, \eta)$  определяется спектром  $S(\xi, \eta)$ , а ширина функции автокорреляции составит  $4\xi_{\max}$  по оси  $\xi$  и  $4\eta_{\max}$  по оси  $\eta$ . Центр функции совпадает с пространственной частотой, соответствующей наклону освещающего пучка. Второй член в (2.7.3) преобразуется в  $\delta$ -функцию в точке  $(\xi = \xi_r, \eta = \eta_r)$ .

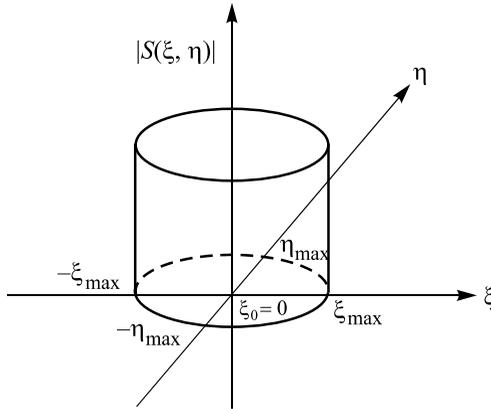


Рис. 2.24. Спектр объектного пучка

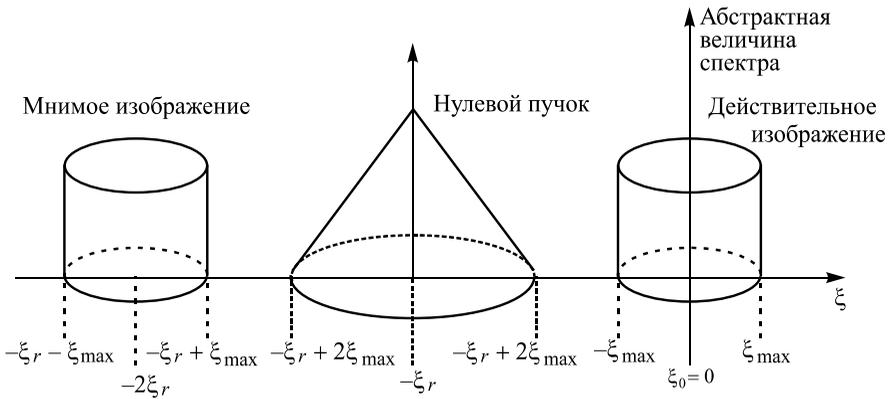


Рис. 2.25. Спектр световой волны после восстановления



Третье слагаемое с точностью до известного множителя  $E_r^2$  формирует собой точную копию исходного поля – мнимое изображение объекта. Величина ее спектра пропорциональна спектру  $S(\xi, \eta)$ . Ее центр находится в точке  $(\xi = -2\xi_r, \eta = -2\eta_r)$ , что соответствует первому порядку дифракции.

Последний член в (2.7.3) формирует волновой фронт, комплексно-сопряженный объектному волновому фронту в плоскости голограммы. Это слагаемое образует еще одно изображение объекта, которое называется действительным изображением (или парным изображением) и спектр которого равен  $S(\xi, \eta)$ . Действительное изображение образуется в  $-1$ -м порядке дифракции.

Использование опорной волны с высокой пространственной частотой (с большим углом наклона) обеспечивает угловое разделение спектров, образующихся в  $0$ ,  $1$  и  $-1$ -м порядках дифракции. Чтобы избежать наложения волны нулевого порядка на волны, образующие изображение, опорная пространственная частота  $\xi$  должна удовлетворять соотношению

$$|\xi_r - \xi_0| \geq 3\xi_{\max}, \quad (2.7.4)$$

где  $\xi_0$  – центральная пространственная частота спектра объекта (в приведенном примере равна нулю).

Из этого выражения видно, что для записи голограммы необходим регистрирующий материал, способный записывать частоты полос, которые в несколько раз превосходят наивысшую пространственную частоту объекта.

В отличие от обычных интерференционных картин голограмма состоит из огромного количества темных и светлых полос. Однако эти полосы не видны даже при сильном увеличении. Причиной этого являются спеклы (англ. *speckle* – крапинка, пятнышко).

При освещении диффузной поверхности лазерным излучением в отраженном свете наблюдается картина спеклов [18–20]. Спекл-структуры возникают из-за пространственного ограничения спектров интерферирующих волн. Такие ограничения возникают или в резуль-



тате конечного размера апертуры, используемой для наблюдения или формирования интерферирующих пучков [17], в результате недостаточной разрешающей способности регистратора или в результате ограничения размера освещаемой диффузной поверхности [21–23].

Рассмотрим рис. 2.26.

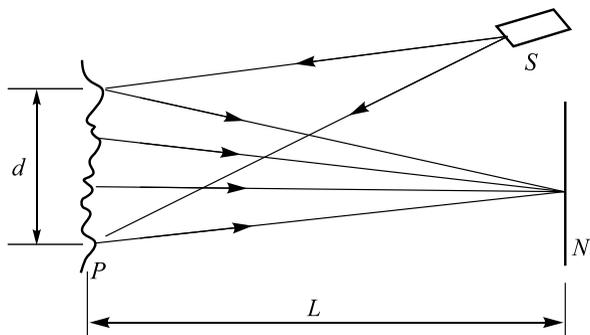


Рис. 2.26. Получение спекл-структуры при отражении от диффузной поверхности:

( $S$  – когерентный источник освещения;  $P$  – диффузная поверхность;  $N$  – плоскость наблюдения)

Средний размер спеклов можно определить с помощью выражения [23]

$$b = 1,22 \frac{\lambda L}{d}, \quad (2.7.5)$$

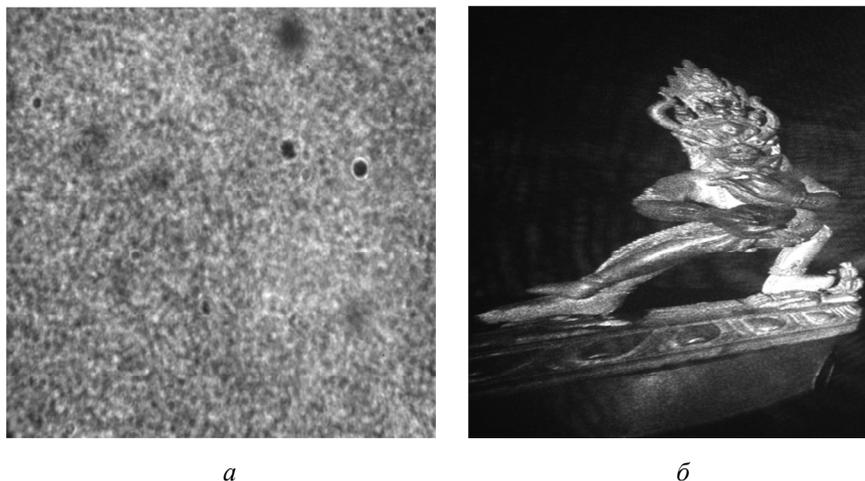
где  $L$  – расстояние от рассеивающей поверхности до плоскости наблюдения;  $d$  – диаметр освещенной поверхности. Спекл-картину, наблюдаемую на некотором расстоянии от диффузно-рассеивающей поверхности, называют объективной (в отличие от субъективной, формируемой в плоскости изображения оптической системы) [24].

Размер спеклов может существенно превышать период полос. В результате проследить интерференционные полосы на голограмме явно не удастся (рис. 2.27).

Цифровые матрицы, используемые в цифровой голографии, не могут на текущий момент обеспечить достаточное пространственное разрешение, которое требуется при больших углах между интерфериру-



ющими пучкам. Поэтому приходится уменьшать угол для того, чтобы можно было оцифровать голограммы с необходимым разрешением, что неизбежно приводит к перекрытию спектров в разных дифракционных порядках.



*Рис. 2.27.* Реальная голограмма и восстановленное изображение:

- a* – увеличенный под микроскопом фрагмент голограммы (0, 25 мм × 0,25 мм);  
*б* – восстановленное с голограммы действительное изображение

Для восстановления изображений из голограмм, полученных при небольших углах между опорной и объектной волнами, необходима разработка новых математических алгоритмов.

## БИБЛИОГРАФИЧЕСКИЙ СПИСОК К ГЛАВЕ 2

1. *Onural L.* Digital decoding of in-line holograms / L. Onural, P. D. Scott // *Optical Engineering*. – 1987. – Vol. 26 (11). – P. 1124–1132.
2. *Kim M. K.* Principles and techniques of digital holographic microscopy / M. K. Kim // *SPIE Reviews*. – 2010. – Vol. 1. – P. 018005-1–018005-50.
3. *Гужов В. И.* Математическая модель распространения света в пространстве / В. И. Гужов, П. А. Бердников, Р. А. Кузнецов // *Автоматика и программная инженерия*. – 2013. – № 1 (3). – С. 121–127.
4. *Детлаф А. А.* Курс физики / А. А. Детлаф, Б. М. Яворский. – 5-е изд. – Москва : Academia, 2005. – 720 с.



5. *Максвелл Дж. К.* Трактат об электричестве и магнетизме : в 2 т. – Москва : Наука, 1989.
6. *Einstein A.* Zur Elektrodynamik bewegter Körper / A. Einstein // *Annalen der Physik.* – 1905. – Vol. 17. – P. 891–922.
7. *Короленко П. В.* Оптика когерентного излучения / П. В. Короленко. – Москва : МГУ им. М. В. Ломоносова, 1997. – 222 с.
8. *Гончаренко А. М.* Гауссовы пучки света / А. М. Гончаренко. – Минск : Наука и техника, 1977. – 144 с.
9. *Домненко В. М.* Моделирование формирования оптического изображения : учебное пособие / В. М. Домненко, М. В. Бурсов, Т. В. Иванова. – Санкт-Петербург : НИУ ИТМО, 2011. – 141 с.
10. *Борн М.* Основы оптики / М. Борн, Э. Вольф. – Москва : Наука, 1973. – 719 с.
11. *Ярославский Л. П.* Цифровая голография / Л. П. Ярославский, Н. С. Мерзляков. – Москва : Наука, 1982. – 219 с.
12. *Гудмен Дж.* Введение в фурье-оптику / Дж. Гудмен. – Москва : Мир, 1970. – 363 с.
13. *Сивухин Д. В.* Общий курс физики. Оптика / Д. В. Сивухин. – Москва : Наука, 1980. – 752 с.
14. *Фейнман Р.* Плотность энергии и поток энергии в электромагнитном поле // Фейнмановские лекции по физике. Т. 6. Электродинамика / Р. Фейнман, Р. Лейтон, М. Сэндс. – Москва : Мир, 1965. – Гл. 27, § 3. – С. 286–290.
15. *Вест Ч.* Голографическая интерферометрия / Ч. Вест. – Москва : Мир, 1982. – 503 с.
16. *Мерзляков Н. С.* Некоторые особенности использования дискретного преобразования Фурье при цифровом восстановлении голограмм / Н. С. Мерзляков, Н. Р. Попова // *Автометрия.* – 1987. – № 5. – С. 17–22.
17. *Коронкевич В. П.* Формирование изображения в оптических системах : учебное пособие / В. П. Коронкевич. – Новосибирск : Изд-во НГТУ, 2005. – 76 с.
18. *Гужов В. И.* Измерение деформаций диффузных поверхностей методом цифровой спекл-интерферометрии / В. И. Гужов, А. Г. Козачок, В. Г. Нечаев // *Оптический журнал.* – 1996. – № 10. – С. 73–76.
19. *Gushov V. I.* Measuring the strains of diffuse surfaced by digital speckle-interferometry / V. I. Gushov, A. G. Kozachok, V. G. Nechaev // *Journal of Optical Technology.* – 1996. – Vol. 63, N 10. – P. 786–789.
20. *Нечаев В. Г.* Разработка методов цифровой спекл-интерферометрии с расширенным динамическим диапазоном : дис. ... канд. техн. наук : 05.11.07 / В. Г. Нечаев. – Новосибирск, 1998.



21. *Майоров Е. Е.* Расчет параметров сканирования интерферометрической системы контроля формы диффузно отражающих объектов / Е. Е. Майоров, В. Т. Прокопенко, Л. А. Ушверидзе // Приборы. – 2012. – Т. 145, № 7. – С. 23–25.

22. *Майоров Е. Е.* Исследование влияния спекл-структуры на формирование интерференционного сигнала и погрешность измерений / Е. Е. Майоров, В. Т. Прокопенко // Научное приборостроение. – 2013. – Т. 23, № 2. – С. 38–46.

23. *Клименко Н. С.* Голография сфокусированных изображений и спекл-интерферометрия / Н. С. Клименко. – Москва : Наука, 1985. – 224 с.

---

---

## ГЛАВА 3

### ДИСКРЕТИЗАЦИЯ ГОЛОГРАММ

---

---

Компьютерная обработка производится над цифровыми сигналами. Преобразование непрерывного сигнала в цифровой вид можно трактовать как отображение непрерывного множества в счетное множество. В результате преобразования в общем случае изображение представляется в виде некоторого массива чисел.

Необходимо соблюдать соответствие между непрерывными и цифровыми преобразованиями. Непрерывное и цифровое преобразования сигнала соответствуют друг другу, если одинаковые входные сигналы они преобразуют в одинаковые выходные.

Для преобразования непрерывных двумерных сигналов в цифровые используют две элементарные процедуры – *дискретизацию* и *квантование*. Дискретизация относится к классу линейных преобразований, поэлементное квантование – к классу нелинейных преобразований.

Дискретизация – это замена непрерывного аналогового сигнала последовательностью чисел, представляющих разложение этого сигнала по какому-либо конечномерному базису. Дискретизация заключается в получении отсчетов (*sampling*) сигнала. Наиболее удобным способом дискретизации является представление сигнала в виде выборок значений (отсчетов) в отдельных, регулярно расположенных точках, отделенных друг от друга интервалом, который называется интервалом дискретизации.

Воспроизведение непрерывного сигнала по выборкам может проводиться на основе как ортогональных, так и неортогональных базисных функций.

В 1807 г. Ж. Фурье изобрел метод решения уравнения, описывающего распространение тепла в твердом теле, – интегральное преобразование Фурье. Если ряды, которые теперь называют по имени Фурье, были хорошо известны к началу его работы по проблеме гармонического анализа, то интеграл Фурье является открытием Фурье.



### 3.1. ОСНОВНЫЕ СВОЙСТВА ПРЕОБРАЗОВАНИЯ ФУРЬЕ

Преобразование Фурье вводится как некоторая интегральная операция, позволяющая для функции  $f(x)$ , отвечающей определенным условиям, как правило, выполняющимся в физических экспериментах, найти спектр  $F(\omega)$  :

$$F(\omega) \rightarrow \mathfrak{F}(f(x)) = \int_{-\infty}^{\infty} f(x) e^{-i\omega x} dx, \quad (3.1.1)$$

где  $\omega$  – частота. Функцию  $F(\omega)$  называют фурье-образом функции  $f(x)$ .

В оптике имеют дело с волновыми полями, являющимися функциями пространственных координат. В соответствии с этим под спектром понимают спектр пространственных частот. Преобразование Фурье обратимо:

$$f(x) \rightarrow \mathfrak{F}^{-1}(F(\omega)) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega) e^{i\omega x} d\omega. \quad (3.1.2)$$

Это выражение описывает разложение неперiodической функции  $f(x)$  в непрерывную совокупность экспонент (плоских волн), весовым множителем которых является спектр  $F(\omega)$ . Отметим, что в общем случае  $f(x)$  и  $F(\omega)$  являются комплексными функциями переменных  $x$  и  $\omega$ .

Рассмотрим некоторые наиболее важные свойства преобразования Фурье.

**1. Линейность преобразования Фурье.** Пусть функции имеют фурье-образы

$$f(x) \rightarrow \mathfrak{F}(f(x)) = F_1(\omega) \text{ и } \varphi(x) \rightarrow \mathfrak{F}(\varphi(x)) = F_2(\omega),$$

тогда

$$\mathfrak{F}(af(x) + b\varphi(x)) = aF_1(\omega) + bF_2(\omega). \quad (3.1.3)$$



**2. Теорема о переносе.** Пусть  $F(\omega)$  – фурье-образ функции  $f(x)$ , тогда фурье-образом функции  $f(x-x_0)$  будет  $F(\omega)e^{-i\omega x_0}$ . Таким образом, при смещении функции по координате на интервал  $x_0$  ее фурье-образ умножается на гармонический множитель  $e^{-i\omega x_0}$ . Справедливо и обратное утверждение: умножение функции на гармонический множитель  $e^{-i\omega_0 x}$  приводит к смещению по частоте фурье-образа на интервал  $\omega_0$  (теорема о сдвиге по частоте):  $F(f(x)e^{-i\omega_0 x}) = F(\omega - \omega_0)$ .

**3. Теорема подобия.** Пусть  $F(\omega)$  – фурье-образ функции  $f(x)$ . Тогда фурье-образом функции  $f(ax)$ , у которой в  $a$  раз изменен масштаб по оси ординат, будет функция  $\frac{1}{a}F\left(\frac{\omega}{a}\right)$ .

#### 4. Преобразование Фурье от свертки двух функций

В оптике широко применяются функции корреляции и свертки. Это интегральные операции, которые двум заданным функциям  $f(x)$  и  $\varphi(x)$  ставят в соответствие некоторую третью функцию. Соответствующая интегральная операция для нахождения функции взаимной корреляции имеет вид

$$f(x)\varphi(x) = \int_{-\infty}^{\infty} f(\xi)\varphi^*(x+\xi)d\xi, \quad (3.1.4)$$

где  $\varphi^*$  – функция, комплексно-сопряженная к функции  $\varphi$ .

Функцию свертки находят через интеграл

$$f(x) \otimes \varphi(x) = \int_{-\infty}^{\infty} f(\xi)\varphi(x-\xi)d\xi = \int_{-\infty}^{\infty} f(x-\xi)\varphi(\xi)d\xi. \quad (3.1.5)$$

Подынтегральные выражения представляют собой произведения двух сдвинутых относительно друг друга исходных функций, причем величина сдвига  $x$  является параметром при интегрировании и значение интеграла зависит от этого параметра.



Преобразование Фурье от свертки двух функций равно

$$\mathfrak{F}(f(x) \otimes \varphi(x)) = F_1(\omega)F_2(\omega). \quad (3.1.6)$$

Это получаем из следующих соображений.

Пусть  $h(x) = f(x) \otimes \varphi(x) = \int_{-\infty}^{\infty} f(\xi)\varphi(x-\xi)d\xi$ , тогда

$$\mathfrak{F}(f(x) \otimes \varphi(x)) = \iint f(\xi)\varphi(x-\xi)e^{-i\omega x}d\xi = \int f(\xi) \int \varphi(x-\xi)e^{-i\omega x}dx d\xi.$$

Учитывая, что по теореме о запаздывании  $\int \varphi(x-\xi)e^{-i\omega x}dx = F_2(\omega)e^{-i\omega\xi}$ , получим

$$\mathfrak{F}(f(x) \otimes \varphi(x)) = \int f(\xi)F_2(\omega)e^{-i\omega\xi}d\xi = F_1(\omega)F_2(\omega).$$

Соответственно, преобразование Фурье переводит произведение двух функций в свертку их фурье-образов:

$$\mathfrak{F}(f(x)\varphi(x)) = F_1(\omega) \otimes F_2(\omega). \quad (3.1.7)$$

Если одна из функций берется с комплексным сопряжением, тогда

$$\mathfrak{F}(f(x)\varphi^*(x)) = F_1(\omega)F_2(\omega). \quad (3.1.8)$$

Эти свойства можно расширить и на функции, для которых определено скалярное произведение.

На сессии Французской академии наук Ж. Фурье сделал сообщение, ставшее историческим. Он заявил, что произвольную функцию, заданную на конечном интервале, всегда можно представить в виде суммы синусоидальных и косинусоидальных составляющих. Как утверждал Ж. Фурье, функция может состоять из произвольного числа «кусков», аналитическое выражение которых совершенно различно. Не требуется и непрерывность функций.

Многие выдающиеся математики, такие как Ж.Л. Лагранж, П.-С. Лаплас, А.М. Лежандр, С.Д. Пуассон, не могли принять утверждение Фурье о том, что любое исходное распределение можно разложить на составляющие в виде главной гармоники и гармоник более высоких частот. Л. Эйлер также считал идеи Фурье ошибочными.



В подходе Фурье основное возражение вызывало утверждение о том, что по существу разрывная функция может быть представлена суммой синусоидальных функций, являющихся непрерывными. В качестве примера можно привести функцию, называемую ступенькой Хевисайда, значение которой равно нулю слева от разрыва и единице – справа от разрыва. Если Ж. Фурье был прав в своих предположениях, то сумма бесконечного ряда тригонометрических функций должна сходиться к точному представлению ступенчатой функции. В то время это утверждение казалось совершенно невозможным.

Тем не менее академия не могла игнорировать значение результатов, полученных Ж. Фурье, и удостоила его премии за математическую теорию законов теплопроводности и сравнение результатов его теории с точными физическими экспериментами. Однако награда была присуждена со следующей оговоркой: «Исходя из новизны предмета исследований и его важности, мы решили присудить премию, отмечая в то же время, что путь, которым автор приходит к своим уравнениям, не свободен от затруднений и что анализ, проведенный им при их интегрировании, оставляет желать несколько большей общности, равно как и строгости».



*Фурье Жан Батист Жозеф (21.3.1768, Осер – 16.5.1830, Париж), французский математик, член Парижской АН (с 1817 г.). Ж. Фурье был 12-м из 15 детей в семье портного. Окончив военную школу в Осере, в 1796–1798 гг.*



преподавал в Политехнической школе. Уже в лекциях в 1796 г. он изложил теорему о числе действительных корней алгебраического уравнения, лежащих между данными границами, названную его именем. Основной областью занятий Фурье была математическая физика. В 1807 г. и 1811 г. он представил Парижской АН свои первые открытия по теории распространения тепла в твердом теле, а в 1822 г. опубликовал известную работу «Аналитическая теория тепла», сыгравшую важную роль в последующей истории математики. В ней Ж. Фурье вывел дифференциальное уравнение, разработал для решения уравнения теплопроводности при тех или иных заданных граничных условиях метод разделения переменных. В основе этого метода лежит представление функций тригонометрическими рядами. Ж. Фурье привел первые примеры разложения в тригонометрические ряды Фурье функций, которые заданы на различных участках различными аналитическими выражениями. Его попытка доказать возможность разложения в тригонометрический ряд Фурье любой произвольной функции была неудачна, но положила начало большому циклу исследований, посвященных проблеме представимости функций тригонометрическими рядами.

Вопрос о сходимости рядов Фурье для тех или иных классов функций привел к появлению новых областей в математике. Один из примеров этого – теория обобщенных функций.

### 3.2. ОСНОВНЫЕ СВОЙСТВА ОБОБЩЕННЫХ ФУНКЦИЙ

История обобщенных функций восходит к концу XIX в., когда О. Хэвисайд ввел систему символических правил для анализа процессов в электрических цепях. Ему пришлось ввести так называемую функцию включения – функцию Хэвисайда. Этим было положено начало операционному исчислению.

Позднее, в конце 20-х гг. XX в., физик П. Дирак в исследованиях по квантовой механике ввел еще одну новую функцию, названную его именем. Были определены также правила дифференцирования дельта-функции Дирака и функции Хэвисайда.

Основы математической теории обобщенных функций были заложены С.Л. Соболевым при решении задачи Коши для гиперболических уравнений (1937 г.).



*Оливер Хэвисайд (Oliver Heaviside, 18 мая 1850 г. – 3 февраля 1925 г.) – английский ученый-самоучка, инженер, математик и физик. Впервые применил комплексные числа для изучения электрических цепей, разработал технику применения преобразования Лапласа для решения дифференциальных уравнений, переформулировал уравнения Максвелла в терминах трехмерных векторов, напряженностей электрического и магнитного полей и электрической и магнитной индукций и независимо от других математиков создал векторный анализ.*

*Между 1880 и 1887 гг. О. Хэвисайд разрабатывал операционное исчисление, метод решения дифференциальных уравнений с помощью сведения к обыкновенным алгебраическим уравнениям, который вызвал бурную полемику из-за отсутствия строгого обоснования. Тогда он произнес известную фразу: «Математика есть наука экспериментальная, определения появляются последними». Эта фраза была ответом на критику математиков за использование еще не вполне определенных операторов.*

С другой стороны к теории обобщенных функций подошел С. Бохнер, который построил теорию преобразований Фурье функций степенного роста. Эти преобразования являлись по существу обобщенными функциями.

В 1950-х годах французский математик Л. Шварц дал систематическое изложение теории обобщенных функций, построил их теорию преобразования Фурье [1].



*Сергей Львович Соболев родился 6 октября 1908 г. в Петербурге. Программу средней школы освоил самостоятельно. В 1923 г. поступил в последний класс школы и в 1924 г. окончил ее с отличием. В 1925 г. был зачислен на физико-математический факультет Ленинградского университета.*

*В 1932–1934 гг. предложил новый метод решения задачи Коши для гиперболического уравнения с переменными коэффициентами, основанный на обобщении формулы Кирхгофа. Работы, связанные с гиперболическими уравнениями, привели его к пересмотру классического понятия решения дифференциального уравнения. Рассмотрение С.Л. Соболевым решений в пространствах функционалов ознаменовало начало теории обобщенных функций.*

*В конце 1950-х гг. академики М.А. Лаврентьев, С.Л. Соболев и С.А. Христианович выступили с инициативой организации нового крупного научного центра – Сибирского отделения Академии наук. Именем С.Л. Соболева назван Институт математики СО РАН, основателем и директором которого он был с 1957 по 1983 г. Скончался 3 января 1989 г.*

После выхода «Теории распределений» Л. Шварца обобщенные функции приобрели широкую популярность. Однако особенности их использования в инженерной деятельности, в частности в оптике, являются до сих пор недостаточно понятными.

Понятие обобщенной функции удастся ввести благодаря тому, что обобщенные функции никогда не входят непосредственно в результаты измерений. Они всегда фигурируют под знаком интеграла, описывающего стадию наблюдения или регистрации. Всегда, когда применя-



ется обобщенная функция, имеется в виду, что в дальнейшем будет производиться интегрирование по тем переменным, от которых она зависит. Поэтому достаточно знать только результат действия данной обобщенной функции на интегральное преобразование.



*Лоран Шварц (Loran Schwartz) родился в Париже 5 марта 1915 г. в семье хирурга. После окончания Высшей нормальной школы в 1937 г. был призван в армию. Во время оккупации Франции фашистами, будучи евреем, скрывался от гестапо по подложным документам. В 1943 г. тем не менее защитил диссертацию в Страсбурге. После войны был профессором в Сорбонне и Политехнической школе.*

*В 1950 г. получил Филдсовскую медаль за теорию распределений, а затем появился его знаменитый двухтомник *Theories des Distributions*.*

*Л. Шварц писал: «Чтобы совершить открытие в математике, надо преодолеть сдержанность и традицию. Нельзя двигаться вперед, не будучи подрынным элементом». Скончался 4 июля 2002 г. в Париже.*

Формально обобщенная функция  $\varphi(x)$  определяется как линейный непрерывный функционал над тем или иным векторным пространством достаточно «хороших» функций (так называемых основных функций)  $f(x)$ , которые имеют непрерывные производные всех порядков и финитны, т. е. отличны от нуля только в конечной области независимой переменной  $x$ . Все «хорошие» функции допускают пре-



образование Фурье. Поэтому все свойства преобразования Фурье полностью сохраняются для «хороших» функций.

Будем обозначать действие обобщенной функции  $\varphi(x)$  на основную функцию  $f(x)$  следующим образом [2]:

$$(f, \varphi) = \int_{-\infty}^{\infty} f(x)\varphi(x)dx. \quad (3.2.1)$$

Выражение (3.2.1) можно рассматривать как скалярное произведение двух функций: «хорошей» функции  $f(x)$  и обобщенной  $\varphi(x)$ . Запись скалярного произведения «хорошей» функции и обобщенной в виде интеграла является чисто условной, так как значение интеграла зависит от последовательного приближения к пределу, и поэтому интеграл, строго говоря, не имеет смысла.

Перемножить две произвольные обобщенные функции в общем случае нельзя. Чаще всего определяется произведение обобщенных функций  $\varphi(x)$  на обычные функции  $f(x)$ .

К понятию обобщенной функции не следует обращаться всякий раз, когда функция не является «хорошей». Например, функция

$$f(x) = \frac{\sin \omega_0 x}{\omega_0 x}$$
 не является «хорошей», так как область ее задания не

ограничена, но она имеет фурье-образ и допускает обратное преобразование Фурье. Поэтому к понятию обобщенной функции необходимо обращаться в следующих случаях [2]:

- когда исходная функция абсолютно неинтегрируема;
- исходная функция входит в класс сингулярных функций;
- необходимо провести дифференцирование функций абсолютно интегрируемых, но имеющих разрывы.

Отметим основные свойства скалярного произведения [3].

1. Если скалярное произведение существует, то операция скалярного умножения двух функций является коммутативной и ассоциативной по отношению к сложению:

$$(f, \varphi) = (\varphi, f), \quad (3.2.2a)$$

$$(a_1 f_1 + a_2 f_2, \varphi) = a_1 (f_1, \varphi) + a_2 (f_2, \varphi), \quad (3.2.2b)$$

где  $a_1$  и  $a_2$  – постоянные комплексные числа.



2. Если скалярное произведение функции  $f(x)h(x)$  на функцию  $\varphi(x)$  существует, то

$$(f(x)h(x), \varphi(x)) = \int_{-\infty}^{\infty} f(x)h(x)\varphi(x)dx = (f(x), \varphi(x)h(x)). \quad (3.2.3)$$

3. При переносе одной функции на величину  $x_0$  получаем

$$\begin{aligned} (f(x-x_0), \varphi(x)) &= \int_{-\infty}^{\infty} f(x-x_0)\varphi(x)dx = \\ &= \int_{-\infty}^{\infty} f(x)\varphi(x+x_0)dx = (f(x), \varphi(x+x_0)). \end{aligned} \quad (3.2.4)$$

4. При сжатии одной функции в один раз получим

$$\begin{aligned} (f(ax), \varphi(x)) &= \int_{-\infty}^{\infty} f(ax)\varphi(x)dx = \\ &= \frac{1}{|a|} \int_{-\infty}^{\infty} f(x)\varphi\left(\frac{x}{a}\right)dx = \frac{1}{|a|} \left( f(x), \varphi\left(\frac{x}{a}\right) \right). \end{aligned} \quad (3.2.5)$$

5. Свойства фурье-образов обобщенных функций.

Пример. Пусть  $e^{-i\omega x}$  – обобщенная функция [15]. По определению

$$(f(x), e^{-i\omega x}) = \int_{-\infty}^{\infty} f(x)e^{-i\omega x}dx = G(\omega) = \mathfrak{F}(f(x)), \text{ т. е. обобщенная функ-}$$

ция  $e^{-i\omega x}$  определяется функционалом, который ставит в соответствие любой «хорошей» функции  $f(x)$  ее фурье-образ  $G(\omega)$ .

Основные соотношения, задающие свойства фурье-образов обобщенных функций, имеют вид

$$(f(x), \varphi(x)) = \frac{1}{2\pi} (F(\omega), G(-\omega)), \quad (3.2.6a)$$

$$(f(x), \varphi^*(x)) = \frac{1}{2\pi} (F(\omega), G^*(\omega)), \quad (3.2.6b)$$



где  $F(\omega)$  и  $G(\omega)$  – фурье-образы функций  $f(x)$  и  $\varphi(x)$  соответственно.

Выражение (3.2.6) – основное соотношение, задающее свойства фурье-образа функции, называется теоремой Парсевалея. Теорема Парсевалея устанавливает равенство между энергией сигнала и энергией его спектра. Следует заметить, что общий вид теоремы Парсевалея часто называют теоремой Планшереля или обобщенной формулой Рэлея. Теорема была доказана для рядов М.-А. Парсевалем в 1799 г. и позднее была применена к рядам Фурье. Одновременно это соотношение является определением фурье-образа обобщенной функции.

Выражение (3.2.6) применимо не к любой обобщенной функции. Не все обобщенные функции имеют фурье-образ. Выделяется класс «умеренных» обобщенных функций, которые имеют фурье-образ [2]. В этот класс входят:

- абсолютно интегрируемые функции;
- ограниченные функции;
- обобщенные функции, имеющие ограниченную область задания.

Фурье-образ умеренной обобщенной функции является умеренной обобщенной функцией.

В анализе обобщенных функций большое значение имеет операция свертки двух функций. Определение операции свертки в области обобщенных функций опирается на понятие прямого произведения обобщенных функций [4]. Возьмем основную функцию  $f(x, y)$ , тогда прямое произведение двух обобщенных функций  $\varphi_1(x)$  и  $\varphi_2(x)$  можно определить как

$$\varphi_1(x)\varphi_2(y) = (\varphi_1(x), (\varphi_2(y), f(x, y))). \quad (3.2.7)$$

Пример. Прямое произведение двух дельта-функций  $\delta(x) \times \delta(y)$  есть  $\delta(x, y)$ .

Прямое произведение имеет следующие свойства.

1. Коммутативность:

$$\varphi_1(x) \times \varphi_2(y) = \varphi_2(y) \times \varphi_1(x). \quad (3.2.8a)$$

2. Ассоциативность:

$$\varphi_1(x) \{ \varphi_2(y) \varphi_3(z) \} = \{ \varphi_1(y) \varphi_2(y) \} \varphi_3(z). \quad (3.2.8б)$$



На основе прямого произведения можно определить функционал свертки обобщенных функций:

$$\begin{aligned}
 (\varphi_1(x) \otimes \varphi_2(y), f(x)) &= \int \left\{ \int \varphi_1(\xi) \varphi_2(x - \xi) d\xi \right\} f(x) dx = \\
 &= \iint \varphi_1(\xi) \varphi_2(\eta) f(\xi + \eta) d\xi d\eta = \\
 &= (\varphi_1(x) \times \varphi_2(y), f(x + y)). \tag{3.2.8}
 \end{aligned}$$

Следует отметить, что  $f(x + y)$  – уже не финитная функция (финитная – отличная от нуля только на конечном интервале) в пространстве переменных  $x$  и  $y$ , поэтому это выражение в общем случае не имеет смысла. Но (3.2.8) имеет смысл, если один из функционалов  $\varphi_1(x)$  или  $\varphi_2(y)$  имеет ограниченный носитель и если носители обоих функционалов ограничены с одной и той же стороны (например,  $\varphi_1(x) = 0$  при  $x < a$ ,  $\varphi_2(x) = 0$  при  $y < b$ ).

Пример. Для любой обобщенной функции  $\varphi(x)$

$$\varphi(x) \otimes \delta(x) = \varphi(x). \tag{3.2.9}$$

Действительно,

$$\begin{aligned}
 (\varphi(x) \otimes \delta(y), f(x)) &= (\varphi(x) \delta(y), f(x + y)) = \\
 &= (\varphi(x), (\delta(y), f(x + y))) = (\varphi(x), f(x)).
 \end{aligned}$$

### 3.3. СВОЙСТВА НЕКОТОРЫХ ОБОБЩЕННЫХ ФУНКЦИЙ, НЕОБХОДИМЫХ ДЛЯ ДИСКРЕТИЗАЦИИ

Рассмотрим некоторые обобщенные функции, которые нам понадобятся при дискретизации изображений.

*Дельта-функция* – сингулярная обобщенная функция. Эта функция впервые была введена английским физиком Дираком в 1926 г.



Поль Дирак родился 8 августа 1902 г. в английском городе Бристоле. В 1933 г. Дирак и Шредингер получили Нобелевскую премию по физике «за открытие новых продуктивных форм атомной теории».

«Казалось, что он человек из будущего. Потому что он говорил только утверждениями. Ум его был так устроен, что если он ничего не мог сказать утвердительно, он вообще не говорил. Если ему сказать: “Я не понял того-то”, он делал удивленное лицо: “Как вам объяснить?” и опять повторял, несколько медленнее, те же изречения и в том же порядке, что и раньше. Он был на вершине своей славы в тридцатом году, и ему Макс Борн говорил: “Слушайте, Дирак, вам 28 лет, сколько вы еще создадите ценного!” А он с полным убеждением вещает: “Nothing more, никогда больше ничего я не сделаю!” Я помню, мы тогда все рассмеялись, а Дирак удивленно на нас посмотрел: “Почему вы смеетесь?” Ну и мы разговор замяли. Но ведь он же действительно с тех пор ничего не сделал! Так что он абсолютно точно знал, что всё, что мог сделать, он сделал». (Воспоминания Ю.Б. Румера)

Дельта-функцию (функция Дирака) одной вещественной переменной можно определить как функцию  $\delta(x)$ , удовлетворяющую следующим условиям:

$$\delta(x) = \left. \begin{array}{l} \infty \text{ при } x = 0 \\ 0 \text{ в остальных случаях} \end{array} \right\}, \quad (3.3.1a)$$

$$\int_{-\infty}^{\infty} \delta(x) dx = 1. \quad (3.3.1b)$$



Дельта-функция на всей оси времени равна нулю. И только в точке  $x = 0$  она устремляется в бесконечность.

Действие дельта-функции на функцию  $f(x)$  можно определить как

$$(f, \delta) = \int_a^b f(x) \delta(x) dx = f(0), \quad a < 0 < b, \quad (3.3.2)$$

где  $f(x)$  – произвольная непрерывная функция  $x$ .

Определим преобразование Фурье от дельта-функции  $\delta(x)$ . Пользуясь свойством (3.2.6), получим

$$\begin{aligned} (F(\omega), \mathfrak{T}(\delta(x))) &= 2\pi (f(x), \delta(x)) = \\ &= \left( \int_{-\infty}^{\infty} F(\omega) e^{i\omega x} d\omega, \delta(x) \right) = \\ &= \int_{-\infty}^{\infty} F(\omega) e^{i\omega 0} d\omega = \int_{-\infty}^{\infty} F(\omega) d\omega = (F(\omega), 1(\omega)). \end{aligned} \quad (3.3.3)$$

Таким образом, преобразование Фурье от дельта-функции  $\delta(x)$  равно  $1(\omega)$ .

На рис. 3.1 показаны графики дельта-функции и ее фурье-образа.

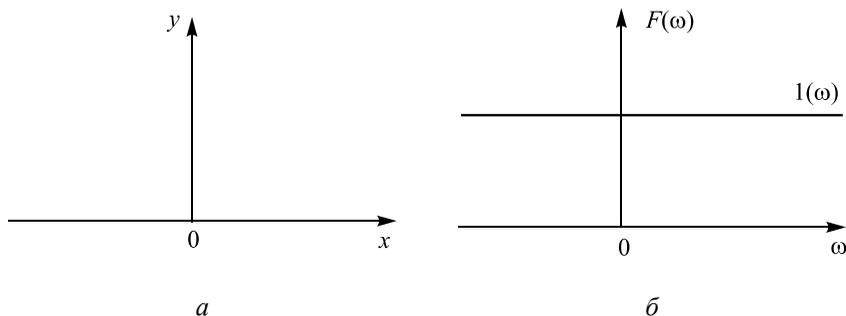


Рис. 3.1. Графическое представление дельта-функции:

$a$  – вид дельта-функции;  $b$  – график фурье-образа дельта-функции



**Двумерная дельта-функция.** Двумерная дельта-функция Дирака есть сингулярный оператор, обладающий следующими свойствами:

$$\delta(x, y) = \left\{ \begin{array}{l} \infty, x = 0, y = 0 \\ 0 \text{ в остальных случаях} \end{array} \right\}. \quad (3.3.4)$$

Пусть функция  $f(x, y)$  описывает исходное непрерывное изображение бесконечных размеров, представляя значение интенсивности или какого-либо другого параметра реального волнового поля, тогда

$$(f(x, y), \delta(x, y)) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \delta(x, y) dx dy = f(0, 0). \quad (3.3.5)$$

**Смещенная дельта-функция** определяется следующим образом:

$$\delta(x - x_0) = \left\{ \begin{array}{l} \infty, x = x_0 \\ 0 \text{ в остальных случаях} \end{array} \right\}. \quad (3.3.6)$$

Действие смещенной функции  $\delta(x - x_0)$  на  $f(x)$  можно описать как

$$(f(x), \delta(x - x_0)) = \int_a^b f(x) \delta(x - x_0) dx = f(x_0), \quad a < x_0 < b. \quad (3.3.7)$$

Область интегрирования должна включать точку  $x_0$ , иначе интеграл (3.3.5) обращается в ноль.

Преобразование Фурье от  $\delta(x - x_0)$  (используем свойство преобразования Фурье – сдвиг функции на  $x_0$  приводит к умножению ее фурье-образа на  $e^{-i\omega x_0}$ ):

$$\mathfrak{Z}(\delta(x - x_0)) = \mathfrak{Z}(\delta(x)) e^{-i\omega x_0} = e^{-i\omega x_0}. \quad (3.3.8)$$

**Двумерная смещенная дельта-функция.** Двумерная дельта-функция Дирака есть сингулярный оператор, обладающий следующими свойствами:

$$\delta(x - x_0, y - y_0) = \left\{ \begin{array}{l} \infty \text{ в точке } (x_0, y_0) \\ 0 \text{ в остальных случаях} \end{array} \right\}. \quad (3.3.9)$$



Двумерная дельта-функция представляется функционалом

$$\begin{aligned} & (f(x, y), \delta(x - x_0, x - y_0)) = \\ & = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \delta(x - x_0, x - y_0) dx dy = f(x_0, y_0). \end{aligned} \quad (3.3.10)$$

**Гребенка Дирака.** С помощью дельта-функций можно определить так называемую *гребенку Дирака* – набор смещенных дельта-функций Дирака (рис. 3.2).

Гребенка Дирака определяется как

$$\text{comb}_{\Delta x}(x) = \left\{ \begin{array}{l} \infty \text{ в точках } (n\Delta x), n = (-\infty, \infty) \\ 0 \text{ в остальных случаях} \end{array} \right\}. \quad (3.10.11)$$

Как видно, гребенка Дирака – это прежде всего периодическая функция, где  $\Delta x$  – период. Кроме этого, гребенка Дирака определена на всем интервале от  $-\infty$  до  $\infty$ . Она, так же как и дельта-функция, на всей оси  $x$  равна нулю, кроме точек  $n\Delta x$ , в которых она устремляется в бесконечность.

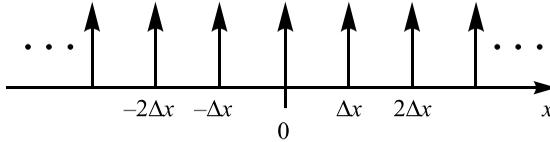


Рис. 3.2. Вид гребенки Дирака

Действие гребенки Дирака на функцию  $f(x)$  можно описать как

$$(f(x), \text{comb}_{\Delta x}(x)) = \mathcal{S}_{n=-\infty}^{\infty} f(n\Delta x), \quad (3.3.12)$$

где символ  $\mathcal{S}_{n=-\infty}^{\infty}$  означает набор (set) значений функций  $f(x)$ , заданных в точках  $n\Delta x$ . В литературе вместо использования этого символа

применяют обозначение суммы  $\left( \sum_{n=-\infty}^{\infty} \right)$  или двойной суммы для

двумерных функций [5]. Однако это обозначение в дальнейшем пони-



мается как обычная сумма значений, что в большинстве случаев приводит к неправильной интерпретации выражений.

Выражение (3.3.12) означает, что значения  $f(x)$  задаются только в точках отсчетов. Таким образом, описывается дискретизация функции  $f(x)$  – перевод непрерывной функции в дискретный вид  $f(n\Delta x)$ , т. е. некоторый набор дискретных значений.

Найдем фурье-образ гребенки Дирака:

$$F_{\text{comb}_{\Delta x}}(\omega) = \sum_{n=-\infty}^{\infty} \mathfrak{T}(\delta(x-n\Delta x)) = \sum_{n=-\infty}^{\infty} e^{-i\omega n\Delta x}. \quad (3.3.13)$$

В каждой точке  $\omega$  произойдет суммирование от всех  $e^{-i\omega n\Delta x}$ , поэтому можно записать

$$F_{\text{comb}_{\Delta x}}(\omega) = \sum_{n=-\infty}^{\infty} e^{-i\omega n\Delta x} = \frac{2\pi}{\Delta x} \sum_{n=-\infty}^{\infty} \delta\left(\omega - n\frac{2\pi}{\Delta x}\right). \quad (3.3.14)$$

Следовательно, фурье-образ гребенки Дирака также является гребенкой Дирака.

**Двумерная гребенка Дирака.** Эта функция представляет бесконечное число дельта-функций, заданных в узлах решетки с шагом  $(\Delta x, \Delta y)$  (рис. 3.3).

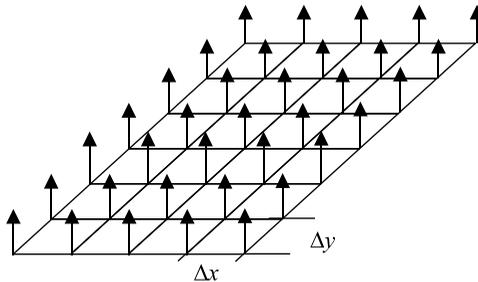


Рис. 3.3. Набор двумерных дельта-функций, осуществляющих дискретизацию изображений

Действие двумерной гребенки Дирака на функцию задается как

$$(f(x, y), \text{comb}_{\Delta x, \Delta y}(x, y)) = \sum_{j_1=-\infty}^{\infty} \sum_{j_2=-\infty}^{\infty} f(j_1\Delta x, j_2\Delta y). \quad (3.3.15)$$



Поскольку с помощью этой функции осуществляется идеальная дискретизация изображений, то ее иногда называют дискретизирующей функцией.

Преобразование Фурье от  $\text{comb}_{\Delta x, \Delta y}(x, y)$  является бесконечным набором дельта-функций в плоскости частот с шагом  $\omega_{xs} = \frac{2\pi}{\Delta x}$ ,

$$\omega_{ys} = \frac{2\pi}{\Delta y}, \text{ т. е.}$$

$$\begin{aligned} F_{\text{comb}_{\Delta x, \Delta y}}(\omega_x, \omega_y) &= \\ &= \frac{4\pi^2}{\Delta x \Delta y} \sum_{j_1=-\infty}^{\infty} \sum_{j_2=-\infty}^{\infty} \delta(\omega_x - j_1 \omega_{xs}, \omega_y - j_2 \omega_{ys}). \end{aligned} \quad (3.3.16)$$

**Ограниченная гребенка Дирака.** При дискретизации важен вид гребенки Дирака с ограниченным числом смещенных дельта-функций

$$\sum_{n=-N}^N \delta(x - nT).$$

Учитывая, что  $\mathfrak{F}(\delta(x - x_0)) = e^{-i\omega x_0}$ , получим, что преобразование Фурье от этой функции равно

$$\begin{aligned} \mathfrak{F}\left(\sum_{n=-N}^N \delta(x - nT)\right) &= \sum_{n=-N}^N e^{-i\omega nT} = \\ &= \frac{e^{i\omega NT} - e^{-i\omega(N+1)T}}{1 - e^{-i\omega T}} = \frac{\sin\left[\left(N + \frac{1}{2}\right)\omega T\right]}{\sin\left(\frac{\omega T}{2}\right)}. \end{aligned} \quad (3.3.17)$$

На рис. 3.4 показан вид этой функции.

Функция

$$D_N(t) = \frac{1}{2\pi} \frac{\sin\left[\left(N + \frac{1}{2}\right)t\right]}{\sin\left(\frac{t}{2}\right)} \quad (3.3.18)$$



называется ядром Дирихле. Эта функция обладает следующим свойством:

$$\int_{-\pi}^{\pi} D_N(t) dt = 1. \quad (3.3.19)$$

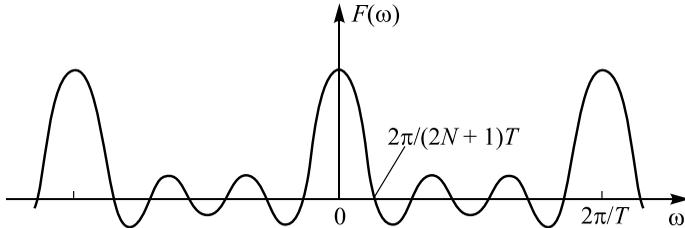


Рис. 3.4. График фурье-образа ограниченной гребенки Дирака

### Двумерная ограниченная гребенка Дирака

$$D(x, y) = \sum_{j_1=-N_1}^{N_1} \sum_{j_2=-N_2}^{N_2} \delta(x - j_1 \Delta x, y - j_2 \Delta y). \quad (3.3.20)$$

Преобразование Фурье от  $D(x, y)$

$$\mathfrak{F}(D(x, y)) = \frac{\sin \left[ \left( N_1 + \frac{1}{2} \right) \omega \Delta x \right]}{\sin \left( \frac{\omega \Delta x}{2} \right)} \frac{\sin \left[ \left( N_2 + \frac{1}{2} \right) \omega \Delta y \right]}{\sin \left( \frac{\omega \Delta y}{2} \right)}. \quad (3.3.21)$$

**Прямоугольный импульс.** Прямоугольный импульс определяется так:

$$rect_{\tau}(x) = \left. \begin{array}{l} 1, \quad -\frac{\tau}{2} < x < \frac{\tau}{2} \\ \frac{1}{2}, \quad |x| = \frac{\tau}{2} \\ 0 \text{ в остальных случаях} \end{array} \right\}. \quad (3.3.22)$$



Прямоугольный импульс с произвольной амплитудой  $A_0$

$$P(x) = A_0 \text{rect}_\tau(x) = \left. \begin{array}{l} A_0, \quad -\frac{\tau}{2} < x < \frac{\tau}{2} \\ \frac{A_0}{2}, \quad |x| = \frac{\tau}{2} \\ 0 \text{ в остальных случаях} \end{array} \right\}. \quad (3.3.23)$$

Преобразование Фурье от прямоугольного импульса равно

$$\begin{aligned} \mathfrak{F}(P(x)) &= \int_{-\tau/2}^{\tau/2} A_0 e^{-i\omega t} d\omega = \frac{A_0}{-i\omega} e^{-i\omega t} \Big|_{-\tau/2}^{\tau/2} = \frac{A_0 \left( e^{-i\omega \frac{\tau}{2}} - e^{i\omega \frac{\tau}{2}} \right)}{-i\omega} = \\ &= \frac{-A_0 2i \sin\left(\omega \frac{\tau}{2}\right)}{-i\omega} = \frac{A_0 2 \sin\left(\omega \frac{\tau}{2}\right)}{\omega \frac{\tau}{2}} \frac{\tau}{2} = A_0 \tau \text{sin} c\left(\omega \frac{\tau}{2}\right), \end{aligned} \quad (3.3.24)$$

где функция  $\text{sin} c(x) = \frac{\sin(x)}{x}$  определяется как

$$\text{sin} c(x) = \left\{ \begin{array}{ll} \frac{\sin(x)}{x}, & x \neq 0 \\ 1, & x = 0 \end{array} \right\}. \quad (3.3.25)$$

Ее график показан на рис. 3.5.

Вид прямоугольного импульса и его фурье-преобразование (3.3.24) показаны на рис. 3.6.

Если определить действие прямоугольного импульса на функцию как

$$S_\tau(0) = (f(x), \text{rect}_\tau(x)) = \int_{-\infty}^{\infty} f(x) \text{rect}_\tau(x) dx = \int_{-\tau/2}^{\tau/2} f(x) dx, \quad (3.3.26)$$



то этот интеграл определит площадь функции  $f(x)$  при изменении  $x$  в диапазоне  $\left[-\frac{\tau}{2}, \frac{\tau}{2}\right]$  в точке с центром 0.

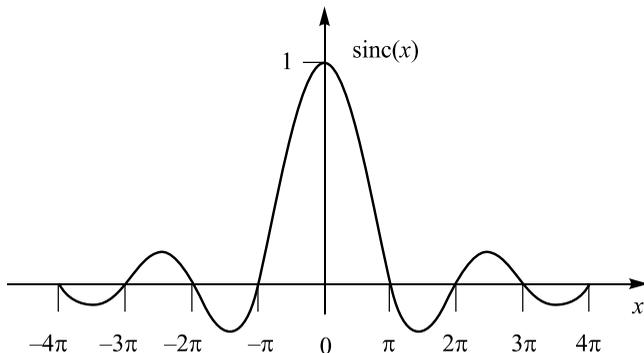


Рис. 3.5. График функции  $\text{sinc}(x)$

Соответственно, смещенный импульс можно определить как

$$S_{\tau}(\Delta x) = (f(x), \text{rect}_{\tau}(x - \Delta x)) = \int_{-\infty}^{\infty} f(x) \text{rect}_{\tau}(x - \Delta x) dx = \int_{\Delta x - \tau/2}^{\Delta x + \tau/2} f(x) dx. \quad (3.3.27)$$

**Бесконечная последовательность прямоугольных импульсов.** Определим спектр периодической последовательности прямоугольных импульсов длительностью  $\tau$  и амплитудой  $A_0$ , следующих с частотой  $\Delta x$  симметричной относительно нуля (рис. 3.7).

Периодическую последовательность можно представить в виде набора:

$$A_0 \sum_{n=-\infty}^{\infty} \text{rect}_{\tau}(x - k\Delta x). \quad (3.3.28)$$

Определим действие суммы прямоугольных импульсов на функцию

$$(f(x), \sum_{n=-\infty}^{\infty} \text{rect}_{\tau}(x - k\Delta x)) = \sum_{n=-\infty}^{\infty} S_{\tau}(k\Delta x). \quad (3.3.29)$$

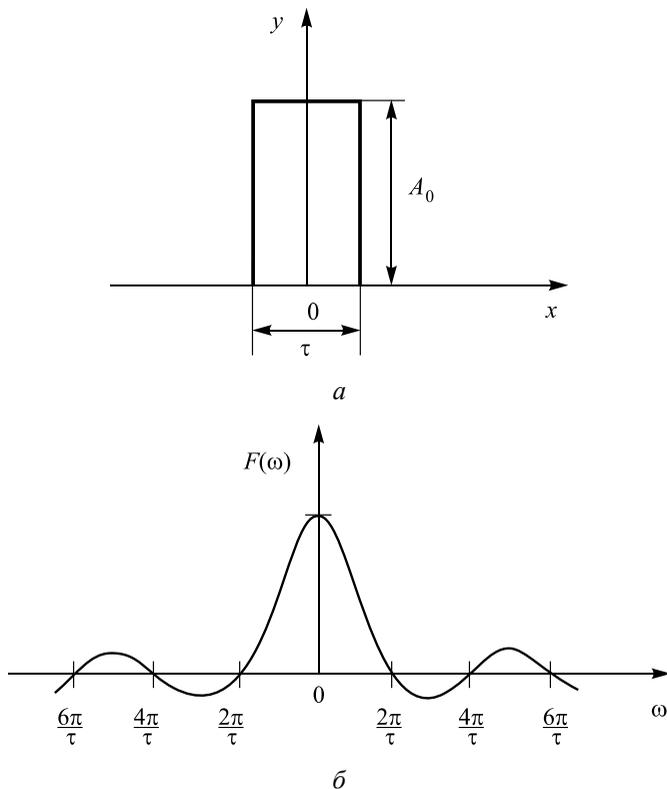


Рис. 3.6. Графическое представление прямоугольного импульса:

а – вид прямоугольного импульса; б – график фурье-образа прямоугольного импульса

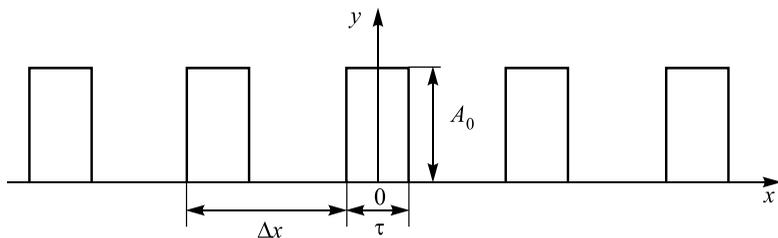


Рис. 3.7. Вид периодической последовательности прямоугольных импульсов



То есть эта функция определяет площадь функции, ограниченной прямоугольными импульсами, и задает их серией значений в точках, соответствующих гребенке Дирака.

Бесконечную последовательность прямоугольных импульсов можно записать в виде ряда Фурье:

$$P(x) = \frac{A_0\tau}{\Delta x} \left[ 1 + 2 \sum_{k=1}^{\infty} \frac{\sin\left(k\omega\frac{\tau}{2}\right)}{\left(k\omega\frac{\tau}{2}\right)} \cos(\omega x) \right]. \quad (3.3.30)$$

**Ограниченная последовательность прямоугольных импульсов.**

Ограниченную периодическую последовательность можно представить в виде набора:

$$A_0 \sum_{n=-N}^N \text{rect}_{\tau}(x - k\Delta x).$$

Определим действие суммы прямоугольных импульсов на функцию

$$(f(x), \sum_{n=-N}^N \text{rect}_{\tau}(x - k\Delta x)) = \sum_{n=-N}^N S_{\tau}(k\Delta x).$$

Преобразование Фурье от этого ряда

$$\begin{aligned} \mathfrak{F}\left(A_0 \sum_{n=-N}^N \text{rect}_{\tau}(x - k\Delta x)\right) &= \sum_{k=-N}^N \mathfrak{F}(\text{rect}_{\tau}(x - k\Delta x)) = \\ &= A_0\tau \sum_{k=-N}^N \text{sinc}\left(\omega\frac{\tau}{2}\right) e^{-ik\Delta x\omega}. \end{aligned} \quad (3.3.31)$$

В общем случае для последовательности из  $N$  импульсов длительностью  $\tau$  каждый, следующих с пространственной периодичностью  $\Delta x$ , спектр имеет вид

$$F(\omega) = A_0\tau \text{sinc}\left(\frac{\omega\tau}{2}\right) \frac{\sin\left[\left(N + \frac{1}{2}\right)\omega\Delta x\right]}{\sin\left(\frac{\omega\Delta x}{2}\right)}. \quad (3.3.32)$$



В случае пяти прямоугольных импульсов получим следующий спектр (рис. 3.8).

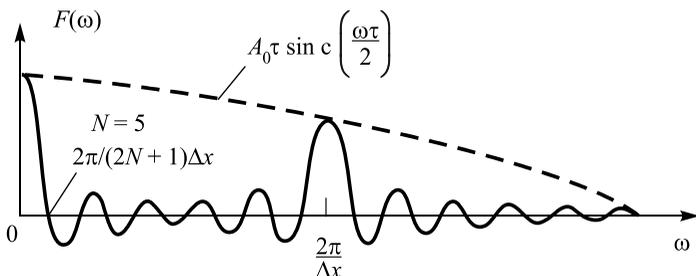


Рис. 3.8. Спектр пяти ( $N=5$ ) прямоугольных импульсов с  $\tau = \Delta x / 2$

**Двумерная последовательность ограниченного числа прямоугольных импульсов.** С помощью этой функции осуществляется дискретизации с усреднением по прямоугольной апертуре:

$$P_{\tau_x, \tau_y}(x, y) = \sum_{j_1=-N_1}^{N_1} \sum_{j_2=-N_2}^{N_2} \text{rect}_{\tau_x, \tau_y}(x - j_1 \Delta x, y - j_2 \Delta y). \quad (3.3.33)$$

Ее фурье-образ равен

$$F(\omega_x, \omega_y) = \mathfrak{Z}(P_{\tau_x, \tau_y}(x, y)) = \tau_x \tau_y \sin c\left(\frac{\omega_x \tau_x}{2}\right) \sin c\left(\frac{\omega_y \tau_y}{2}\right) \times \frac{\sin\left[\left(N_1 + \frac{1}{2}\right)\omega_x \Delta x\right] \sin\left[\left(N_2 + \frac{1}{2}\right)\omega_y \Delta y\right]}{\sin\left(\frac{\omega_x \Delta x}{2}\right) \sin\left(\frac{\omega_y \Delta y}{2}\right)}. \quad (3.3.34)$$

### 3.4. ОПРЕДЕЛЕНИЕ СВЕРТКИ И ПРОИЗВЕДЕНИЯ ОБОБЩЕННОЙ ФУНКЦИИ С ОБЫЧНОЙ ФУНКЦИЕЙ

Запись скалярного произведения «хорошей» функции и обобщенной в виде интеграла является чисто условной, так как значение интеграла зависит от последовательного приближения к пределу, и поэтому



интеграл, строго говоря, не имеет смысла. Перемножить две произвольные обобщенные функции в общем случае нельзя. Чаще всего определяется произведение обобщенных функций  $\varphi(x)$  на обычные функции  $f(x)$ .

Свертка двух функций равна  $f(x) \otimes \varphi(x) = \int_{-\infty}^{\infty} f(\xi) \varphi(x - \xi) d\xi$ , следовательно, свертка с дельта-функцией равна

$$f(x) \otimes \delta(x) = \int_{-\infty}^{\infty} f(\xi) \delta(x - \xi) d\xi = f(x). \quad (3.4.1)$$

Таким образом, операция свертки  $f(x)$  с дельта-функцией  $\delta(x)$  возвращает исходную функцию.

Свертка со смещенной дельта-функцией сдвигает функцию на определенное значение:

$$f(x) \otimes \delta(x - x_0) = \int_{-\infty}^{\infty} f(\xi) \delta(x - x_0 - \xi) d\xi = f(x - x_0). \quad (3.4.2)$$

Свертка функции  $f(x)$  с решеткой Дирака  $\text{comb}_T(x)$  записывается как

$$\begin{aligned} f(x) \otimes \text{comb}_{\Delta x}(x) &= f(x) \otimes \sum_{n=-\infty}^{\infty} \delta(x - n\Delta x) = \\ &= \sum_{n=-\infty}^{\infty} f(x) \otimes \delta(x - n\Delta x) = \sum_{n=-\infty}^{\infty} f(x - n\Delta x). \end{aligned} \quad (3.4.3)$$

Таким образом, свертка  $f(x)$  с  $\text{comb}_T(x)$  приводит к бесконечному повторению реплик исходной функции (рис. 3.9).

Рассмотрим преобразование Фурье от периодической функции  $f(x) = \sum_{n=-\infty}^{\infty} g(x - nT)$ , где  $g(x)$  – функция одного периода, повторяющаяся с периодом  $T$ .

Из математического анализа известно, что спектр любой периодической функции будет дискретным. Периодическую функцию можно



разложить в ряд Фурье, и этот ряд будет сходиться, если функция удовлетворяет следующим условиям [7].

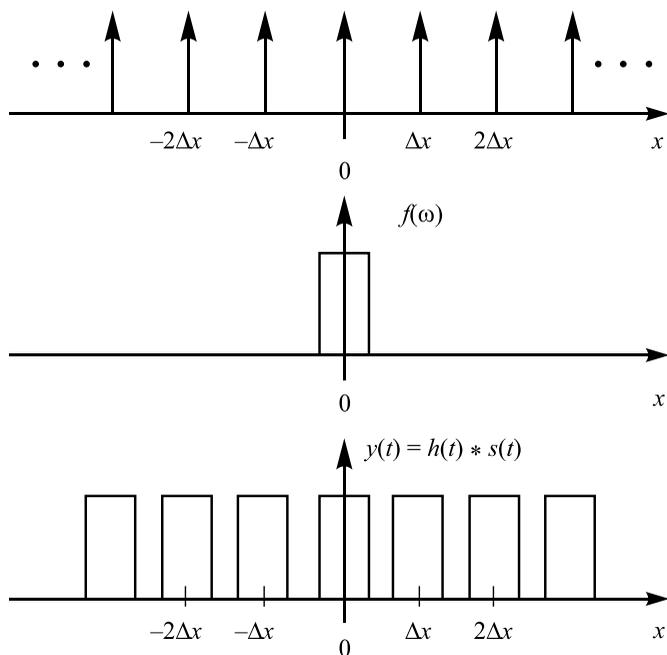


Рис. 3.9. Свертка  $f(x)$  с решеткой Дирака приводит к бесконечному повторению реплик исходной функции

**Теорема.** Пусть функция  $f(x)$  и ее производная  $f'(x)$  – непрерывные функции на отрезке  $[-l, l]$  или же имеют на нем конечное число точек разрыва первого рода. Тогда ряд Фурье функции  $f(x)$  сходится на всей числовой прямой, причем в каждой точке  $x \in (-l, l)$ , в которой  $f(x)$  непрерывна, сумма ряда равна  $f(x)$ , а в каждой точке  $x_0$  разрыва функции сумма ряда равна  $\frac{f(x_0-) + f(x_0+)}{2}$ , где  $f(x_0-) = \lim_{x \rightarrow x_0-} f(x)$  и  $f(x_0+) = \lim_{x \rightarrow x_0+} f(x)$ . На концах отрезка  $[-l, l]$  сумма ряда равна  $\frac{f(-l) + f(l)}{2}$ .



Разложение функции в ряд Фурье запишется как

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} \left( a_n \cos \frac{n\pi x}{l} + b_n \sin \frac{n\pi x}{l} \right), \quad (3.4.4)$$

где

$$a_0 = \frac{1}{l} \int_{-l}^l f(x) dx; \quad (3.4.5)$$

$$a_n = \frac{1}{l} \int_{-l}^l f(x) \cos \frac{n\pi x}{l} dx; \quad (3.4.6)$$

$$b_0 = \frac{1}{l} \int_{-l}^l f(x) \sin \frac{n\pi x}{l} dx. \quad (3.4.7)$$

Таким образом, спектр периодической функции с периодом  $\Delta x$  ( $l = \Delta x / 2$ ) существует только в отдельных точках с дискретным шагом  $\frac{2\pi}{\Delta x}$  (рис. 3.10).

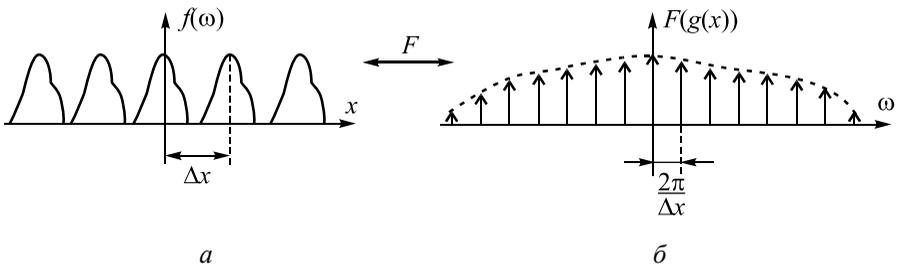


Рис. 3.10. Периодическая функция (а) и ее спектр (б)

Периодическую функцию можно описать как свертку функции  $\text{comb}_{\Delta x}(x)$  с функцией одного периода  $g(x)$ :

$$f(x) = g(x) \otimes \text{comb}_T(x). \quad (3.4.8)$$



Тогда, согласно свойствам преобразования Фурье, спектр свертки определяется как произведение фурье-образов  $\text{comb}_T(x)$  и  $g(x)$ :

$$\begin{aligned} \mathfrak{F}(f(x)) &= \mathfrak{F}(g(x))\mathfrak{F}(\text{comb}_{\Delta x}(x)) = \mathfrak{F}(g(x)) \frac{2\pi}{\Delta x} \sum_{n=-\infty}^{\infty} \delta\left(\omega - n \frac{2\pi}{\Delta x}\right) = \\ &= \frac{2\pi}{\Delta x} \mathfrak{F}(g(x)) \sum_{n=-\infty}^{\infty} \delta\left(\omega - n \frac{2\pi}{\Delta x}\right) = \frac{2\pi}{\Delta x} \mathfrak{F}(g(x)) \text{comb}_{\frac{2\pi}{\Delta x}}(x). \end{aligned} \quad (3.4.9)$$

Таким образом определяется произведение обобщенной функции  $\text{comb}_T(x)$  на обычную функцию. Произведение обычной функции и гребенки Дирака определяет дискретизацию функции, т. е. задание ее в виде отсчетов.

В силу обратимости функций и их спектров можно сделать вывод, что и спектр дискретного сигнала будет периодически повторяющейся функцией.

### 3.5. ИДЕАЛЬНАЯ ДИСКРЕТИЗАЦИЯ

**Одномерная дискретизация.** Возьмем произвольную функцию  $f(x)$ .

Дискретное значение функции можно определить как воздействие на эту функцию  $f(x)$  гребенки Дирака:

$$\begin{aligned} \sum_{n=-\infty}^{\infty} f(n\Delta x) &= (f(x), \text{comb}_{\Delta x}(x)) = \\ &= \int_{-\infty}^{\infty} f(x) \sum_{n=-\infty}^{\infty} \delta(x - n\Delta x) dx = \sum_{n=-\infty}^{\infty} \int_{-\infty}^{\infty} f(x) \delta(x - n\Delta x) dx. \end{aligned} \quad (3.5.1)$$

Выражение (3.5.1) определяет функцию  $\text{comb}_{\Delta x}(x)$  в виде скалярного произведения.

Произведение  $f(x) \text{comb}_{\Delta x}(x)$  определяет задание ее в виде отсчетов:

$$f_p(x) = f(x) \text{comb}_{\Delta x}(x). \quad (3.5.2)$$



Спектр дискретного изображения соответственно равен

$$F_p(\omega) = \frac{2\pi}{\Delta x} F_I(\omega) \otimes \text{comb}_{\frac{2\pi}{\Delta x}}(\omega). \quad (3.5.3)$$

Из выражения (3.5.3) видно, что дискретизация изображения приводит к периодическому повторению ее спектра с периодом  $\frac{2\pi}{\Delta x}$  (рис. 3.11).

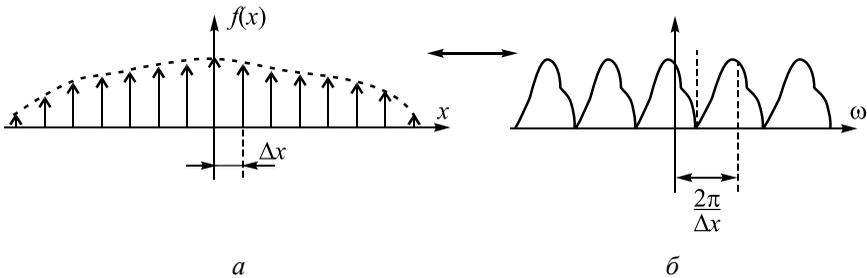


Рис. 3.11. Спектр дискретной функции:

*a* – дискретная функция; *б* – ее спектр

**Двумерная идеальная дискретизация.** Для двумерного случая операция дискретизации осуществляется аналогично. В качестве дискретизирующей функции берется двумерная решетка Дирака.

Дискретное двумерное изображение можно представить как

$$(f(x, y), \text{comb}_{\Delta x, \Delta y}(x, y)) = \sum_{j_1=-\infty}^{\infty} \sum_{j_2=-\infty}^{\infty} f(j_1 \Delta x, j_2 \Delta y). \quad (3.5.4)$$

Выражение (3.5.4) означает, что значения  $f(x, y)$  задаются только в точках отсчетов.

Спектр дискретного изображения можно представить в виде

$$F_p(\omega_x, \omega_y) = \frac{4\pi^2}{\Delta x \Delta y} \sum_{j_1=-\infty}^{\infty} \sum_{j_2=-\infty}^{\infty} \mathfrak{F}_I \left( \omega_x - j_1 \frac{2\pi}{\Delta x}, \omega_y - j_2 \frac{2\pi}{\Delta y} \right), \quad (3.5.5)$$

где  $\mathfrak{F}_I(\omega_x, \omega_y)$  – спектр исходного изображения  $f(x, y)$ ,  $\mathfrak{F}_P(\omega_x, \omega_y)$  – спектр дискретизирующей функции.



Спектр дискретного изображения получается путем *бесконечного повторения спектра исходного изображения* со сдвигом на величины,

кратные  $\omega_{xs} = \frac{2\pi}{\Delta x}$ ,  $\omega_{ys} = \frac{2\pi}{\Delta y}$  (рис. 3.12).

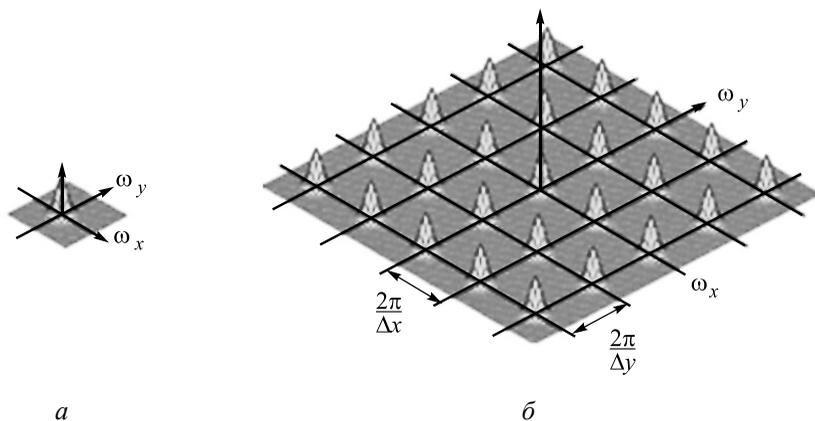


Рис. 3.12. Спектр дискретизованного изображения:

*a* – спектр исходного изображения; *б* – спектр дискретизованного изображения

Если спектр ограничен, то необходимо, чтобы шаг дискретизации ( $\Delta x, \Delta y$ ) выбирался таким, чтобы перекрытие спектров было минимальным.

### 3.6. НЕОБХОДИМЫЕ УСЛОВИЯ ДИСКРЕТИЗАЦИИ ФУНКЦИЙ (ТЕОРЕМА КОТЕЛЬНИКОВА)

Пусть функция  $f(x)$  имеет ограниченный спектр  $F_I(\omega)$  (рис. 3.13).

Непрерывный аналоговый сигнал представляется последовательностью его значений (отсчетов). Эти отсчеты берутся в точках, отделенных друг от друга интервалом  $\Delta x$ , который называется *интервалом дискретизации*. Тогда спектр дискретного сигнала будет таким, как показано на рис. 3.14.

Величину, обратную интервалу между отсчетами, называют *частотой дискретизации*. *Пространственная частота* – отношение периода сигнала к интервалу дискретизации.

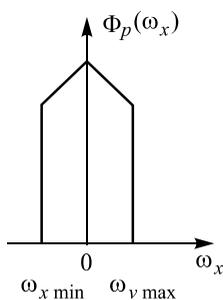


Рис. 3.13. Спектр функции  $f(x)$

Интуитивно понятно, что чем меньше интервал дискретизации и, соответственно, выше частота дискретизации, тем меньше различия между исходным сигналом и его дискретизированной копией. Однако если спектр сигнала ограничен, нет необходимости увеличивать число интервалов дискретизации до бесконечности.

В 1936 г. В.А. Котельников доказал теорему о числе интервалов дискретизации, при которых всегда можно точно восстановить непрерывный сигнал.

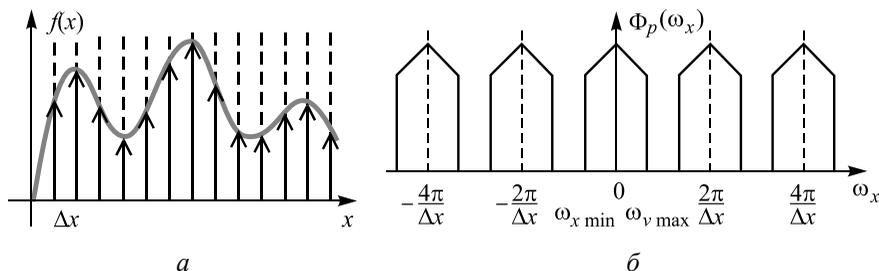


Рис. 3.14. Спектр дискретного сигнала:

$a$  – функция  $f(x)$  и ее отсчеты;  $b$  – спектр дискретного сигнала

В работе [7] В.А. Котельников доказывает, что непрерывный сигнал может быть представлен в виде суммы интерполяционного ряда:

$$\begin{aligned}
 f(x) &= \sum_{n=-\infty}^{\infty} f(n\Delta x) \frac{\sin \left[ \frac{\pi}{\Delta x} (x - n\Delta x) \right]}{\frac{\pi}{\Delta x} (x - n\Delta x)} = \\
 &= \sum_{n=-\infty}^{\infty} f(n\Delta x) \operatorname{sinc} \left[ \frac{\pi}{\Delta x} (x - n\Delta x) \right]. \quad (3.6.1)
 \end{aligned}$$

Интервал дискретизации должен быть таким, чтобы  $\Delta x \leq \frac{\pi}{\omega_{xc}}$ .



*Теорема была доказана В.А. Котельниковым в 1933 г. в работе «О пропускной способности эфира и проволоки в электросвязи» [5]. Работа была заявлена как доклад на намечавшийся I Всесоюзный съезд по вопросам технической реконструкции дела связи и развития слаботочной промышленности. Однако съезд не состоялся, но материалы к нему были опубликованы. Понимая значение теоремы отсчетов, В.А. Котельников в 1936 г. попытался опубликовать статью в более широко читаемом специалистами журнале «Электричество». Получив отказ, он не стал повторять попытки. Поэтому эта работа была неизвестна научной общественности.*

*Независимо от него эту теорему в 1949 г. (через 16 лет) доказал К. Шеннон. Исторические изыскания показывают, что теорема отсчетов как в части утверждения возможности реконструкции аналогового сигнала по дискретным отсчетам, так и в части способа реконструкции рассматривалась многими учеными и ранее. Как один из частных математических результатов теории интерполяции функции эта теорема была открыта еще в начале XX в. английскими математиками Е.Т. Уиткером и Дж.М. Уиткером.*

По теореме Котельникова непрерывная функция  $f(x)$  может быть всегда безошибочно восстановлена по дискретным значениям, если спектр сигнала ограничен частотой Найквиста –  $\omega_{xc}$ .

*Теорема Котельникова формулируется следующим образом.*

Пусть  $\omega_{xc} \leq \left| \frac{\omega_{xs}}{2} \right|$ ,  $\omega_{yc} \leq \left| \frac{\omega_{ys}}{2} \right|$  максимальные значения  $\omega$ , при ко-

тором Фурье-образ сигнала не равен нулю.



Тогда сигнал  $F(x, y)$  поддается точному восстановлению по выборочным отсчетам, если расстояние между соседними отсчетами меньше или равно  $\Delta x \leq \frac{\pi}{\omega_{xc}}$ ,  $\Delta y \leq \frac{\pi}{\omega_{yc}}$ .

Частоты  $\omega_{xc} = \frac{\pi}{\Delta x}$ ,  $\omega_{yc} = \frac{\pi}{\Delta y}$  – называются *частотами среза* или *частотами Найквиста*.

Это означает, что шаг дискретизации не должен превышать половины периода пространственной гармоники, соответствующей самым мелким деталям изображения. В этом случае непрерывное изображение можно полностью восстановить по дискретному изображению. Если это условие не выполняется, то дискретизация может сопровождаться необратимыми искажениями.

В 1977 г. теорему было предложено называть WKS-теоремой или УКШ-теоремой в честь Уиттекера–Котельникова–Шеннона (Whittaker–Kotelnikov–Shannon).

Если  $\Delta x$  и  $\Delta y$  выбраны слишком большими по сравнению с шириной спектра  $\mathfrak{F}_f(\omega)$ , то соседние спектры будут перекрываться (рис. 3.15).

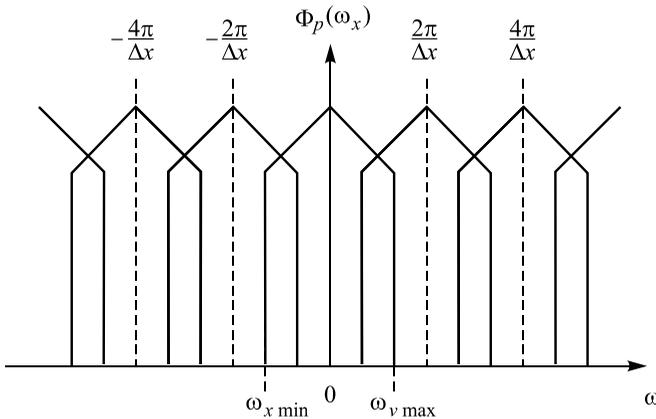


Рис. 3.15. Повторение спектров в результате дискретизации



На рис. 3.15 спектр сигнала такой, что  $\Phi_I(\omega) \neq 0$  при  $\omega > \left| \frac{\pi}{\Delta x} \right|$ .

В этом случае восстановить сигнал без искажений из-за перекрытия спектров не удается.

При дискретизации изображений с недостаточной частотой происходит наложение спектров. Это приводит к появлению в восстановленном изображении *ложных низкочастотных гармоник*. Высокочастотный сигнал после дискретизации со слишком малой частотой приобретает вид ложного низкочастотного рисунка.

На рис. 3.16 показаны сигнал и точки дискретизации, взятые с шагом в два раза больше, чем период гармоники. В этом случае вместо синусоидального сигнала после дискретизации мы можем ошибочно восстановить ложный низкочастотный сигнал.

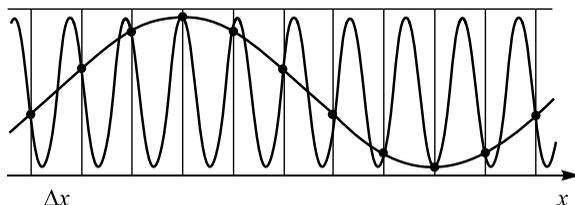


Рис. 3.16. Образование ложного низкочастотного сигнала вследствие недостаточной частоты дискретизации.

Этот же эффект показан на рис. 3.17 для синусоидальных полос с изменяемой частотой.

На рис. 3.17 приведены изображения исходного аналогового и дискретного сигналов. Под изображением аналогового сигнала (верхний рисунок) указаны точки, в которых берутся отсчеты. Сверху приведены значения, показывающие, сколько точек дискретизации приходится на один период. Значения пространственной частоты сначала увеличиваются, а затем уменьшаются. Как видно, низкочастотные компоненты, для которых на один период приходится более двух дискретных точек, восстанавливаются без искажений. Высокочастотные компоненты, для которых это условие не выполняется, искажаются. При этом возникают ложные низкочастотные полосы.

Однако важно отметить, что в теореме Котельникова не говорится и не подразумевается, что сигнал не может быть восстановлен в случае нарушения условия Найквиста.

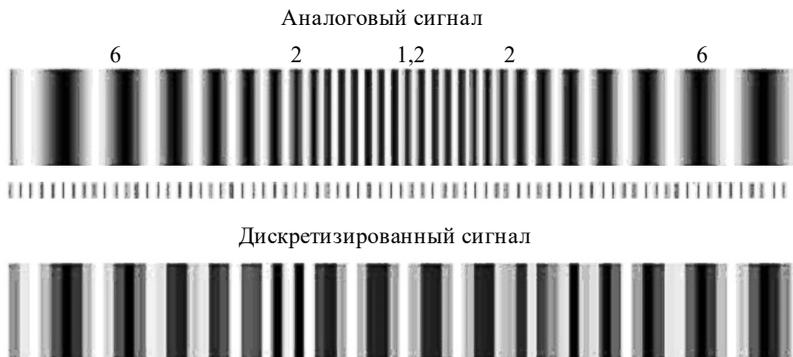


Рис. 3.17. Аналого-цифровое преобразование. Искажение дискретизации вследствие недостаточной частоты дискретизации

Можно ли восстановить сигнал, если условие Найквиста нарушается? Возможность перекрытия этого ограничения для сигналов определенного вида зависит от априорной информации о характеристиках входного сигнала. В теореме Котельникова говорится о числе отсчетов, при которых *всегда* можно восстановить сигнал, но при некоторых видах сигнала и наличии априорной информации его можно восстановить и при недостаточной дискретизации. Не существует фундаментальных причин, которые препятствуют определению частот, больших частоты Найквиста.

Пример спектра, при котором можно восстановить сигнал, показан на рис. 3.18, б.

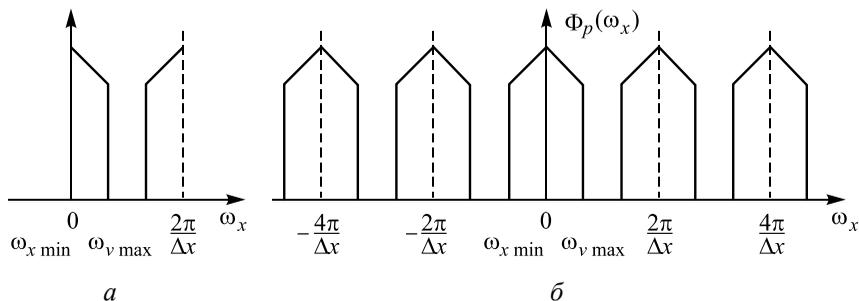


Рис. 3.18. Восстановление спектра сигнала при нарушении условий Найквиста:

$\bar{a}$  – спектр сигнала после дискретизации;  $\bar{b}$  – восстановленный спектр с использованием априорной информации



Входной сигнал должен содержать пространственные частоты в некотором диапазоне от нуля до частоты Найквиста и частоты выше частоты Найквиста. Эти области частот должны быть известны (рис. 3.18, *a*).

После дискретизации пространственные частоты выше частоты Найквиста будут накладываться. На рис. 3.19, *a* показан результат низкочастотной фильтрации, после которой остаются частоты только в диапазоне, ограниченном разрешением фотоприемника.

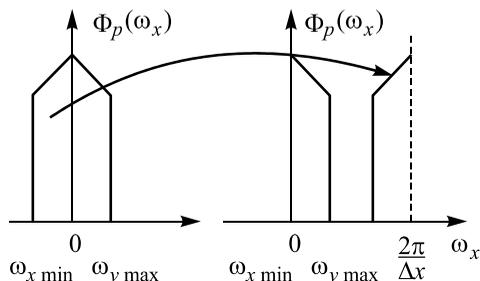


Рис. 3.19. Восстановление спектра переносом части спектра

Однако если имеется априорная информация о том, что область около нуля не содержит частот, ошибочно отображаемые частоты можно вернуть на свое место (рис. 3.18, *б*). Такая процедура будет успешной для сигналов, у которых нет перекрытия от других реплик в области разрешения фотоприемника.

### 3.7. ВОССТАНОВЛЕНИЕ НЕПРЕРЫВНЫХ СИГНАЛОВ ПО ДИСКРЕТНЫМ ОТЧЕТАМ

По теореме Котельникова непрерывная функция  $f(x)$  может быть всегда безошибочно восстановлена по дискретным значениям, если спектр сигнала ограничен частотой Найквиста  $-\omega_{xc}$ . Процесс восстановления становится понятным, если рассмотреть спектр дискретного сигнала (рис. 3.20, верхний график).

Из рисунка видно, что спектр дискретного сигнала представляет периодически повторяющийся спектр исходного сигнала. Для того чтобы восстановить исходный сигнал, надо выделить одиночный



спектр. Это можно сделать умножением периодического спектра на прямоугольный импульс  $rect_{\frac{2\pi}{\Delta x}}(x)$  (рис. 3.20, средний график).

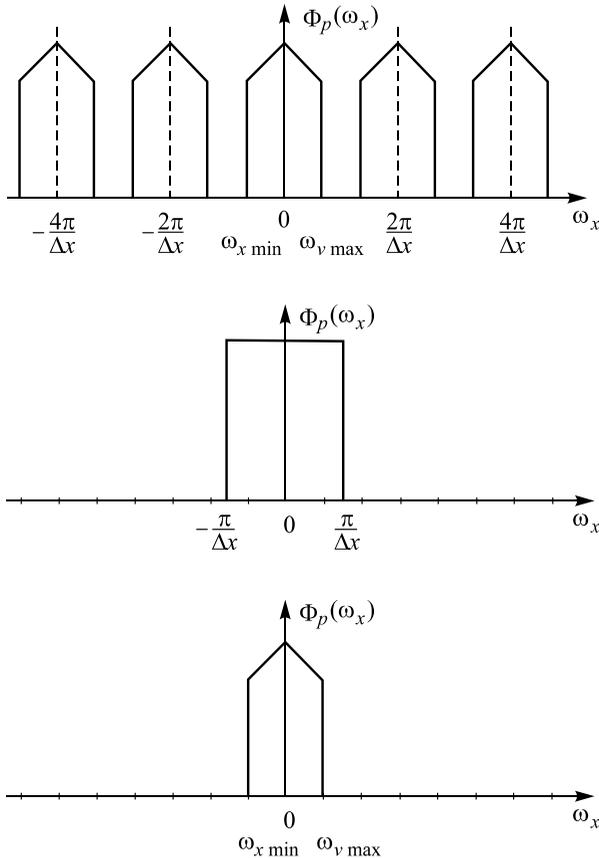


Рис. 3.20. Выделение спектра сигнала из спектра дискретного сигнала

После этого для восстановления исходного непрерывного сигнала необходимо сделать обратное преобразование Фурье.

Умножение спектров приведет в свертке их фурье-образов:

$$\mathfrak{F}^{-1}\left(\Phi(\omega)rect_{\frac{2\pi}{\Delta x}}(\omega)\right) = \mathfrak{F}^{-1}(\Phi(\omega)) \otimes \mathfrak{F}^{-1}\left(rect_{\frac{2\pi}{\Delta x}}(\omega)\right). \quad (3.7.1)$$



Учитывая, что

$$\Phi(\omega) = \frac{2\pi}{\Delta x} \sum_{k=-\infty}^{\infty} F_I \left( \omega - k \frac{2\pi}{\Delta x} \right), \quad (3.7.2)$$

$$\mathfrak{F}^{-1}(\Phi(\omega)) = \frac{1}{\Delta x} \sum_{i=-\infty}^{\infty} f(i\Delta x), \quad (3.7.3)$$

$$\mathfrak{F}^{-1} \left( \text{rect} \frac{2\pi}{\Delta x}(\omega) \right) = \Delta x \sin c \left( x \frac{\pi}{\Delta x} \right), \quad (3.7.4)$$

по теореме о свертке спектров получим

$$\begin{aligned} \mathfrak{F}^{-1} \left( \Phi(\omega) \text{rect} \frac{2\pi}{\Delta x}(\omega) \right) &= \mathfrak{F}^{-1}(\Phi(\omega)) \otimes \mathfrak{F}^{-1} \left( \text{rect} \frac{2\pi}{\Delta x}(\omega) \right) = \\ &= \sum_{n=-\infty}^{\infty} f(n\Delta x) \sin c \left[ \frac{\pi}{\Delta x}(x - n\Delta x) \right]. \end{aligned} \quad (3.7.5)$$

Таким образом, по дискретным значениям можно точно восстановить непрерывную функцию  $f(x)$  интерполяцией дискретных значений  $f(n\Delta x)$  функцией  $\sin c$ .

Идеальная дискретизация – это теоретическая абстракция, поэтому необходимо рассмотреть дискретизацию, которая имеет место при оцифровке голограмм, в реальных системах.

### 3.8. ДИСКРЕТИЗАЦИЯ СИГНАЛОВ В РЕАЛЬНЫХ СИСТЕМАХ

В реальных системах дискретизации изображений, в отличие от идеального случая, выполняются следующие условия:

- дискретизирующая решетка имеет конечные размеры;
- ширина дискретизирующих импульсов заметно отличается от нуля.

Последовательность точек, в которых берутся отсчеты, называется *растром*. Различают регулярные растры и стохастические. Основной признак регулярных растров – это периодическая структура. Как правило, применяются растры, точки которых находятся в узлах квадратов.



ной сетки (рис. 3.21). Практически эта операция осуществляется измерением сигнала с помощью датчика с некоторой конечной апертурой (площадкой по которой происходит усреднение). Числовые значения отсчетов изображения получаются путем пространственного интегрирования по некоторой конечной площадке (апертуре). Набор значений  $\{ f_p(r\Delta x, k\Delta y) \}$  представляет дискретное преобразование сигнала.

Используются апертуры различных форм (например, эллиптические, ромбовидные, гексагональные), но наиболее часто используются апертуры прямоугольной формы. Пусть матрица фотоприемника представляет собой прямоугольный массив прямоугольных датчиков (рис. 3.21, б). Каждый элемент фотоприемника имеет размер  $a_{\tau_x}, a_{\tau_y}$ .

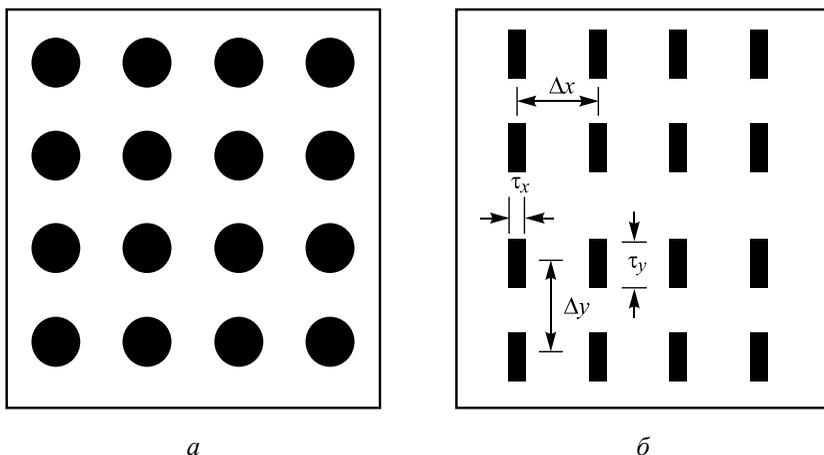


Рис. 3.21. Примеры растров:

$a$  – регулярный растр с круглой апертурой;  $b$  – с прямоугольной апертурой

В этом случае пространственную дискретизацию изображения можно записать с помощью следующего определения для обобщенной функции  $rect$  (для упрощения рассмотрим одномерный случай):

$$S_{\tau}(0) = (f(x), rect_{\tau}(x)) = \int_{-\infty}^{\infty} f(x) rect_{\tau}(x) dx = \int_{-\tau/2}^{\tau/2} f(x) dx. \quad (3.8.1)$$



Для смещенного импульса

$$(f(x), \text{rect}_\tau(x - \Delta x)) = \int_{-\infty}^{\infty} f(x) \text{rect}_\tau(x - \Delta x) dx = \int_{\Delta x - \tau/2}^{\Delta x + \tau/2} f(x) dx. \quad (3.8.2)$$

В результате дискретизации произойдет усреднение значений интенсивности на прямоугольной апертуре.

Определим действие бесконечного набора прямоугольных импульсов  $P(x)$  на функцию:

$$(f(x), P(x)) = \left( f(x), \sum_{k=-\infty}^{\infty} \text{rect}_\tau(x - k\Delta x) \right) = \sum_{k=-\infty}^{\infty} \int_{k\Delta x - \tau/2}^{k\Delta x + \tau/2} f(x) dx. \quad (3.8.3)$$

Бесконечный набор прямоугольных функций – это периодическая функция. Ее можно описать как свертку прямоугольной функции с решеткой Дирака:

$$P(x) = \text{rect}_\tau(x) \otimes \text{comb}_{\Delta x}(x). \quad (3.8.4)$$

Свертка  $\text{rect}_\tau(x)$  с функцией  $f(x)$ :

$$f(x) \otimes \text{rect}_\tau(x) = \int_{-\infty}^{\infty} f(\xi) \text{rect}_\tau(x - \xi) d\xi. \quad (3.8.5)$$

Дискретизация с помощью бесконечного набора прямоугольных апертур опишется как

$$f_P(x) = \sum_{k=-\infty}^{\infty} (f(x) \otimes P(x)) \text{comb}_{\Delta x}(x). \quad (3.8.6)$$

Дискретизацию с помощью конечного набора прямоугольных апертур можно задать, ограничивая эту функцию с помощью конечного набора прямоугольных импульсов:

$$f_P(x) = \sum_{k=-\infty}^{\infty} S_\tau(k\Delta x) = \sum_{k=-\infty}^{\infty} (f(x) \otimes P_N(x)) \text{comb}_{\Delta x}(x). \quad (3.8.7)$$



В результате дискретизации произойдет усреднение значений интенсивности на прямоугольных апертурах, и результаты усреднения на каждой апертуре будут задаваться в массиве чисел  $(r, k)$ .

Пусть  $F_I(\omega)$  – спектр исходной функции,  $\mathfrak{Z}(P_N(x)) = F_{rect_\tau}(\omega)$  – спектр ограниченного набора прямоугольных импульсов

$$F_{rect}(\omega) = \mathfrak{Z}\left(\sum_{k=-N}^N \text{rect}_\tau(x - k\Delta x)\right) = \tau \sin c\left(\frac{\omega\tau}{2}\right) \frac{\sin\left[\left(N + \frac{1}{2}\right)\omega\Delta x\right]}{\sin\left(\frac{\omega\Delta x}{2}\right)}, \quad (3.8.8)$$

тогда преобразование Фурье от дискретизированного сигнала можно записать как

$$\begin{aligned} F_p(\omega) &= (F_I(\omega)F_{rect}(\omega)) \otimes \mathfrak{Z}(\text{comb}_{\Delta x}(x)) = \\ &= \frac{2\pi}{\Delta x} \left[ (F_I(\omega)F_{rect}(\omega)) \right] \otimes \text{comb}_{\frac{2\pi}{\Delta x}}(\omega). \end{aligned} \quad (3.8.9)$$

То есть спектр исходного изображения умножается на спектр ограниченного набора прямоугольных импульсов и периодически повторяется с периодом  $\frac{2\pi}{\Delta x}$ .

Полнокадровые модели фотокамер снабжаются сенсором, размер которого соответствует размеру кадра 35 мм пленки, т. е.  $36 \times 24$  мм. У неполнокадровых моделей размеры матрицы незначительно отличаются: например, у Nikon  $\sim 24 \times 16$  мм, у Canon APS-C  $\sim 22,3 \times 14,9$  мм.

Состав спектра периодической последовательности прямоугольных импульсов зависит от величины отношения периода последовательности к длительности импульса, называемого скважностью импульсов, и

числа этих импульсов. Скважность импульсов равна  $S = \frac{\Delta x}{\tau} = \frac{1}{D}$ , где

$D$  называется коэффициентом заполнения. Малая величина коэффициента заполнения достигается, когда элементы детектора малы и широко разделены. Традиционные датчики, используемые для ввода изображений, разработаны для улавливания максимально возможного количества освещения и имеют величину  $D$  между 5 и 1.



В настоящее время наибольший коэффициент заполнения (близкий к 100 %) имеют ССД матрицы, у CMOS матриц этот коэффициент составляет 80...90 %. На рис. 3.22 показан фрагмент CMOS матрицы, полученный с помощью электронного микроскопа.

Рассмотрим, как сказывается увеличение числа импульсов на спектральную характеристику. Для одномерной последовательности прямоугольных импульсов

$$F(\omega) = A_0 \tau \operatorname{sinc} \left( \frac{\omega \tau}{2} \right) \frac{\sin \left[ \left( N + \frac{1}{2} \right) \omega \Delta x \right]}{\sin \left( \frac{\omega \Delta x}{2} \right)}.$$

Пусть  $L = 24$  мм. Теперь будем увеличивать число прямоугольных импульсов при коэффициенте заполнения 0,8.

На рис. 3.23 показана амплитудно-частотная характеристика от 16 прямоугольных импульсов с единичной амплитудой  $\Delta x = 1,75$  мкм,  $\tau = 1,4$  мкм. Шаг между отдельными (большими пиками на рисунке) равен

$$\omega = \frac{2\pi}{\Delta x}.$$

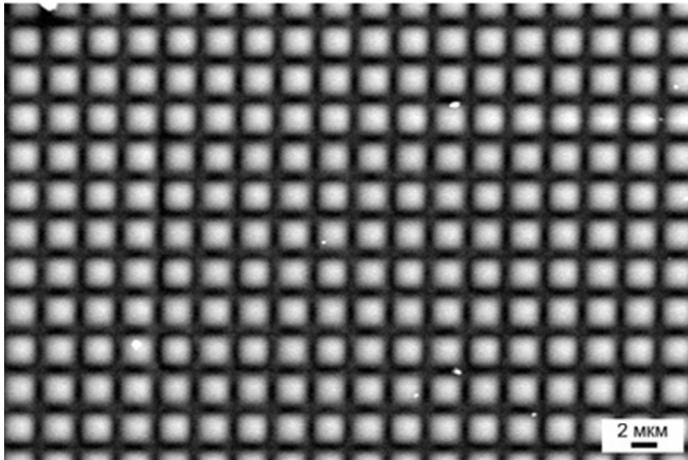


Рис. 3.22. Пример регулярного раstra CMOS матрицы (фотография матрицы сделана с помощью электронного микроскопа)

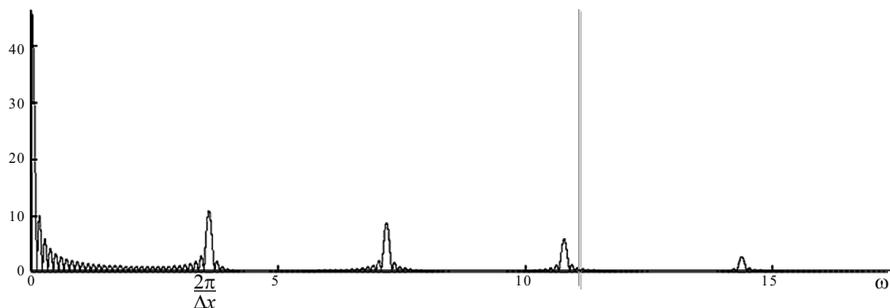


Рис. 3.23. Амплитудно-частотный спектр последовательности прямоугольных импульсов ( $N = 16$ ). Число небольших пиков от 0 до  $\frac{2\pi}{\Delta x}$  равно 32

Число небольших пиков в пределах от 0 до  $\frac{2\pi}{\Delta x}$  составит  $2N$ .

Множитель равен 2, поскольку амплитудно-частотная характеристика и отрицательные значения на рис. 3.23 отображаются в положительные значения.

На рис. 3.24 показана амплитудно-частотная характеристика функции  $\text{sinc}\left(\frac{\omega\tau}{2}\right)$  от 16 прямоугольных импульсов, которая является модулирующей функцией в (2.4.36).

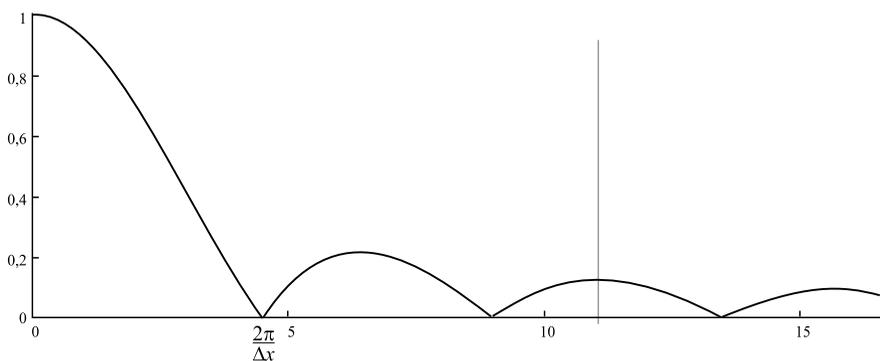


Рис. 3.24. Амплитудно-частотный спектр функции  $\text{sinc}\left(\frac{\omega\tau}{2}\right)$ ,  $\tau = 1,4$  мм



На рис. 3.25 показана амплитудно-частотная характеристика от 32 прямоугольных импульсов с единичной амплитудой  $\Delta x = 0,875$  мм,  $\tau = 0,7$  мм; на рис. 3.26 – амплитудно-частотная характеристика функции  $\sin c\left(\frac{\omega\tau}{2}\right)$  от 32 прямоугольных импульсов.

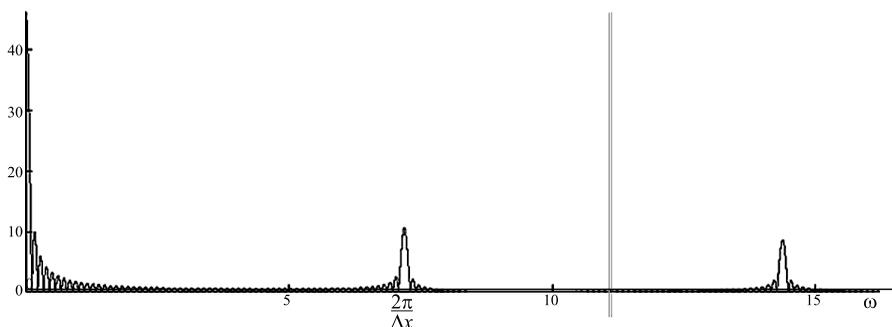


Рис. 3.25. Амплитудно-частотный спектр последовательности прямоугольных импульсов ( $N = 32$ )

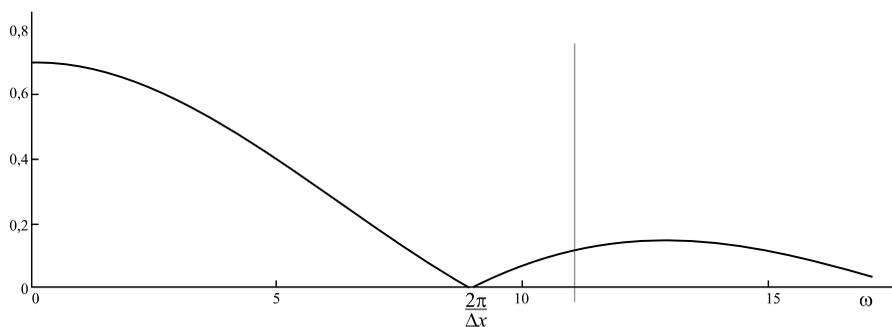


Рис. 3.26. Амплитудно-частотный спектр функции

$$\sin c\left(\frac{\omega\tau}{2}\right), \tau = 0,7 \text{ мм}$$

На рис. 3.27 показана амплитудно-частотная характеристика от 1024 прямоугольных импульсов с единичной амплитудой,  $\Delta x = 0,027$  мм,  $\tau = 0,022$  мм.



На рис. 3.27 пик  $\omega = \frac{2\pi}{\Delta x} = 229,785$  не показан, но его амплитуда такая же, как и у второго пика на предыдущих рисунках. В показанном на рис. 3.27 диапазоне частот функции  $\sin c\left(\frac{\omega\tau}{2}\right) \approx 1$ .

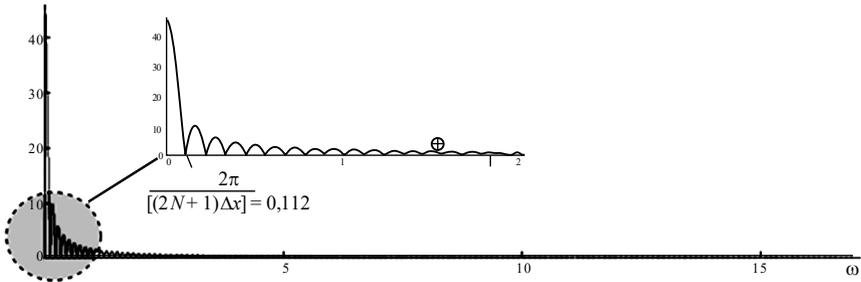


Рис. 3.27. Амплитудно-частотный спектр последовательности прямоугольных импульсов ( $N = 1024$ )

Для двумерного случая

$$F_{rect}(\omega_x, \omega_y) = \tau_x \tau_y \sin c\left(\frac{\omega_x \tau_x}{2}\right) \sin c\left(\frac{\omega_y \tau_y}{2}\right) \times$$

$$\frac{\sin\left[\left(N_1 + \frac{1}{2}\right)\omega_x \Delta x\right] \sin\left[\left(N_2 + \frac{1}{2}\right)\omega_y \Delta y\right]}{\sin\left(\frac{\omega_x \Delta x}{2}\right) \sin\left(\frac{\omega_y \Delta y}{2}\right)}. \quad (3.8.10)$$

Двумерный спектр массива прямоугольных импульсов показан на рис. 3.27.

Если убрать пики по первому столбцу и первой строке, можно более детально увидеть вид остальных пиков (рис. 3.29).

Из рис. 3.28 и 3.29 видно, что при увеличении  $N$  спектр функции  $F_{rect}(\omega)$  приближается к дискретному виду.

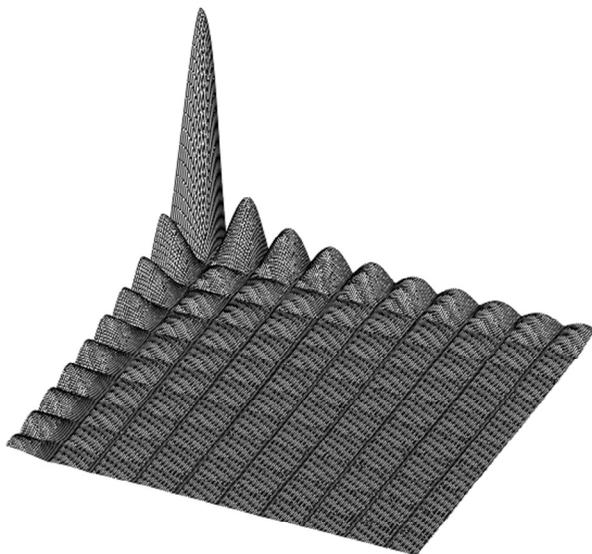


Рис. 3.28. Спектр дискретизирующей функции (показаны только первые десять пиков)

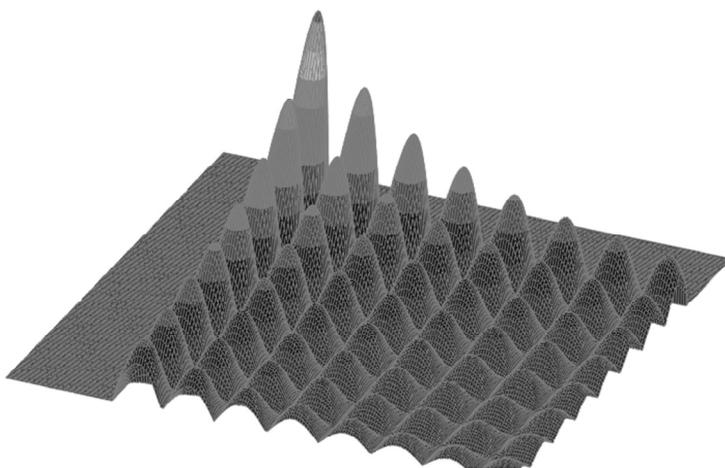


Рис. 3.29. Спектр дискретизирующей функции без первых пиков



Использование массива апертур с конечными размерами не всегда является недостатком. Рассмотрим случай, когда спектр изображения очень широкий и изображение оцифровывается с недостаточной частотой. В этом случае по полученной цифровой выборке нельзя будет верно восстановить исходный сигнал. Восстановленный сигнал будет содержать искажения. Этот эффект называется наложением спектров или алиасингом (aliasing). При оцифровке изображения алиасинг может привести к дефектам в изображении, таким как «блочные», границы или муар.

Для устранения таких искажений необходимо использовать более высокую частоту дискретизации или же искусственно ограничить спектр сигнала перед оцифровкой.

Использование апертуры фактически осуществляет низкочастотную фильтрацию исходного изображения. Высокочастотные компоненты сигнала резко уменьшаются. Это приводит к сужению спектра и, следовательно, уменьшаются ошибки, вызванные наложением спектров.

### БИБЛИОГРАФИЧЕСКИЙ СПИСОК К ГЛАВЕ 3

1. *Schwartz L.* Theorie des distributions : in 2 t. / L. Schwartz. – Paris : Hermann, 1950–1951. – 2 t.
2. *Гельфанд И. М.* Обобщенные функции и действия над ними / И. М. Гельфанд, Г. Е. Шилор. – Москва : Физматлит, 1959. – 470 с.
3. *Сороко Л. М.* Основы голографии и когерентной оптики / Л. М. Сороко. – Москва : Наука, 1971. – 616 с.
4. *Владимиров В. С.* Обобщенные функции в математической физике / В. С. Владимиров. – Москва : Наука, 1979. – 320 с.
5. Использование обобщенных функций для дискретизации изображений [Электронный ресурс] / В. И. Гужов, И. О. Марченко, Д. С. Хайдуков, С. П. Ильиных // Инженерный вестник Дона. – 2017. – № 2. – Режим доступа: <http://ivdon.ru/ru/magazine/archive/n4y2017/4127> (дата обращения: 13.06.2018).
6. *Шипачев В. С.* Курс высшей математики : учебник / В. С. Шипачев ; под ред. А. Н. Тихонова. – Москва : Проспект, 2004. – 600 с.
7. *Котельников В. А.* О пропускной способности эфира и проволоки в электросвязи / В. А. Котельников // Материалы к I Всесоюзному съезду по вопросам технической реконструкции дела связи и развития слаботочной промышленности / Всесоюзный энергетический комитет. – [Б. м. : б. и.], 1933. – Репринт статьи в журнале УФН. – 2006. – Т. 176, № 7. – С. 762–770.

---

---

## ГЛАВА 4

# КОМПЬЮТЕРНОЕ МОДЕЛИРОВАНИЕ ГОЛОГРАФИЧЕСКОГО ПРОЦЕССА

---

---

Математические выражения для описания отдельных этапов голографического процесса показаны в главе 2. В этой главе рассмотрим особенности компьютерной реализации алгоритмов.

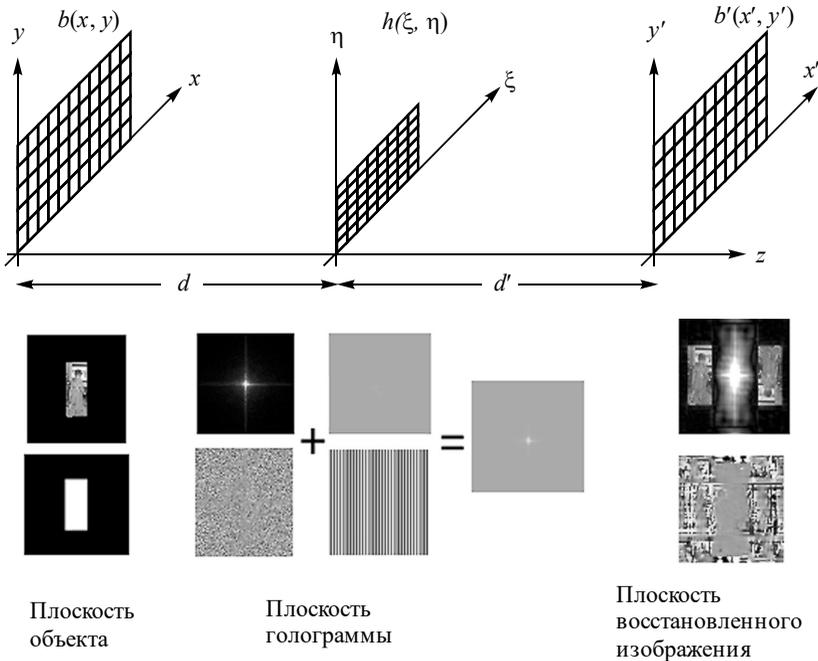


Рис. 4.1. Этапы моделирования голографического процесса



На рис. 4.1 показаны основные этапы, которые необходимо выполнить для исследования процесса получения и восстановления голограмм:

- задание объекта;
- преобразование свободного пространства от  $(x, y)$  до  $(\xi, \eta)$ ;
- получение голограммы;
- восстановление волнового фронта из голограммы сразу же за плоскостью  $(\xi, \eta)$ ;
- преобразование свободного пространства от  $(\xi, \eta)$  до  $(x', y')$ .

Особенность компьютерного моделирования – представление математических выражений в дискретном виде. Наиболее сложным является представление непрерывных преобразований Фурье и Френеля, с помощью которых описывается преобразование свободного пространства, в дискретном виде.

#### 4.1. ПРЕДСТАВЛЕНИЕ ПРЕОБРАЗОВАНИЯ ФУРЬЕ В ДИСКРЕТНОЙ ФОРМЕ

Наиболее простым описанием  $b_1(x, y)$  в дискретном виде является представление ее в виде матрицы отсчетов  $b_{1d}(k, l)$ , взятых на прямоугольном растре с некоторым шагом по координатам  $\Delta x$  и  $\Delta y$  [1].

Представим  $b_1(x, y)$  в дискретной форме:

$$b_{1d}(k, l) = \sum_{k=0}^{N_x-1} \sum_{l=0}^{N_y-1} b_1(k\Delta x, l\Delta y). \quad (4.1.1)$$

Если  $b_1(x, y)$  отлично от нуля в области

$$[-X_{\max}, X_{\max}; -Y_{\max}, Y_{\max}],$$

то

$$N_x = \text{int}(2X_{\max} / \Delta x), \quad N_y = \text{int}(2Y_{\max} / \Delta y), \quad (4.1.2)$$

где  $\text{int}(\cdot)$  – целая часть аргумента. По теореме Котельникова точная интерполяция  $b_1(x, y)$  по  $b_{1d}(k, l)$  возможна, если

$$\Delta x = 1/2v_{x_{\max}}, \quad \Delta y = 1/2v_{y_{\max}}, \quad (4.1.3)$$



где  $[-v_{x\max}, v_{x\max}; -v_{y\max}, v_{y\max}]$  – область, за пределами которой пространственный спектр  $b_1(x, y)$  можно считать равным нулю. Интерполяция может быть описана как свертка  $b_{1d}(k, l)$  с некоторой интерполирующей функцией  $W(x, y)$ .

Запишем выражение для фурье-преобразования:

$$\mathfrak{F}(v_x, v_y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} b_1(x, y) \exp\{-i2\pi[xv_x + yv_y]\} dx dy. \quad (4.1.4)$$

Подставив (4.1.1) в (4.1.4), получим

$$F_b(v_x, v_y) = \sum_{k=0}^{N_x-1} \sum_{l=0}^{N_y-1} b_{1d}(k, l) \exp\left[-i2\pi\left(\frac{kv_x}{2v_{x\max}} + \frac{lv_y}{2v_{y\max}}\right)\right]. \quad (4.1.5)$$

В этом выражении  $b_{1d}(k, l)$  – дискретная матрица отсчетов, взятая с шагом  $\Delta x$  и  $\Delta y$ , но  $v_x$  и  $v_y$  – все еще непрерывные величины. Проведем дальнейшую дискретизацию выражения (4.1.5) по  $v_x$  и по  $v_y$ .

Из ограниченности области задания  $b_{1d}(k, l)$  вытекает, что по теореме отсчетов непрерывную функцию  $F_b(v_x, v_y)$  можно восстановить линейной интерполяцией дискретной функции

$$F(r, s) = \sum_{r=-\infty}^{\infty} \sum_{s=-\infty}^{\infty} F_b(r\Delta v_x, s\Delta v_y), \quad (4.1.6)$$

где  $F_b(r, s)$  – отсчеты  $F_b(v_x, v_y)$ , взятые на прямоугольном растре с шагом  $\Delta v_x$  и  $\Delta v_y$  по координатам  $v_x$  и  $v_y$ .

Из (4.1.6) и (4.1.5) получим

$$F(r, s) = \sum_{k=0}^{N_x-1} \sum_{l=0}^{N_y-1} b_{1d}(k, l) \exp\left[-i2\pi\left(kr \frac{\Delta v_x}{2v_{x\max}} + ls \frac{\Delta v_y}{2v_{y\max}}\right)\right]. \quad (4.1.7)$$

В этом выражении  $F(r, s)$  – дискретная матрица отсчетов, взятая с шагом  $\Delta v_x$  и  $\Delta v_y$ .



Для  $\Delta v_x$ ,  $\Delta v_y$ ,  $\Delta v_{x\max}$ ,  $\Delta v_{y\max}$ ,  $N_x$ ,  $N_y$  выполняются соотношения

$$M_x = \frac{2v_{x\max}}{\Delta v_x}, \quad N_x = \frac{2X_{\max}}{\Delta x},$$

$$M_y = \frac{2v_{y\max}}{\Delta v_y}, \quad N_y = \frac{2Y_{\max}}{\Delta y}. \quad (4.1.8)$$

Обычно для расчетов выбираются значения  $M_x = N_x$  и  $M_y = N_y$ .

Подставив (4.1.8) в (4.1.7), получим окончательную формулу для вычисления элементов матрицы  $F(r, s)$  по матрице чисел  $b_1(k, l)$ :

$$\mathfrak{F}(v_x, v_y) = F(r, s) = \sum_{k=0}^{N_x-1} \sum_{l=0}^{N_y-1} b_1(k, l) \exp \left[ -i2\pi \left( \frac{kr}{N_x} + \frac{ls}{N_y} \right) \right]. \quad (4.1.9)$$

Это соотношение описывает *двумерное дискретное преобразование Фурье*.

Без учета фазовых множителей, стоящих перед интегралом  $\Gamma(\xi, \eta) = \mathfrak{F}(v_x, v_y)$ , значения координат  $(\xi, \eta)$  (рис. 4.1) и пространственные частоты связаны соотношениями (рис. 4.2).

$$v_x = \frac{\xi}{\lambda d}, \quad v_y = \frac{\eta}{\lambda d}. \quad (4.1.10)$$

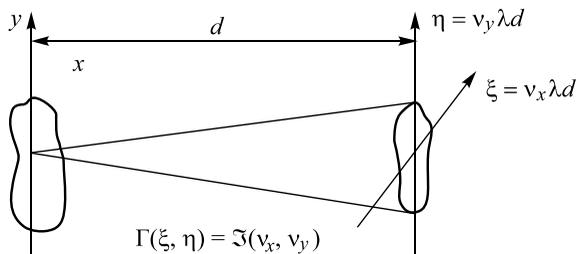


Рис. 4.2. Пространственные частоты в области наблюдения

Дискретное преобразование Фурье (ДПФ) в англоязычной литературе обозначается как Discrete Fourier Transform (DFT).



### 4.1.1. СВОЙСТВА ДИСКРЕТНОГО ПРЕОБРАЗОВАНИЯ ФУРЬЕ

Фурье-спектр дискретного объекта  $f(k, l)$  может быть вычислен в виде конечной суммы:

$$F(\omega_x, \omega_y) = \sum_{k=0}^{N_x-1} \sum_{l=0}^{N_y-1} f(k, l) \exp \left[ -i2\pi \left( \frac{k\omega_x}{N_x} + \frac{l\omega_y}{N_y} \right) \right]. \quad (4.1.11)$$

Обратное преобразование определяется с помощью следующего выражения:

$$f(k, l) = \frac{1}{N_x N_y} \sum_{\omega_x=0}^{N_x-1} \sum_{\omega_y=0}^{N_y-1} F(\omega_x, \omega_y) \exp \left[ i2\pi \left( \frac{k\omega_x}{N_x} + \frac{l\omega_y}{N_y} \right) \right]. \quad (4.1.12)$$

Встречаются и другие определения дискретного преобразования Фурье. Иногда знаки экспонент заменяются на противоположные:

$$F(\omega_x, \omega_y) = \frac{1}{\sqrt{N_x N_y}} \sum_{k=0}^{N_x-1} \sum_{l=0}^{N_y-1} f(k, l) \exp \left[ i2\pi \left( \frac{k\omega_x}{N_x} + \frac{l\omega_y}{N_y} \right) \right]; \quad (4.1.13)$$

$$f(k, l) = \frac{1}{\sqrt{N_x N_y}} \sum_{\omega_x=0}^{N_x-1} \sum_{\omega_y=0}^{N_y-1} F(\omega_x, \omega_y) \exp \left[ -i2\pi \left( \frac{k\omega_x}{N_x} + \frac{l\omega_y}{N_y} \right) \right]. \quad (4.1.14)$$

Как правило, в математической литературе прямое и обратное (с заменой  $i$  на  $-i$  в показателе экспоненты) преобразования умножаются на множитель  $1/\sqrt{2\pi}$  (в дискретном преобразовании  $1/\sqrt{N_x N_y}$ ), в то время как физики и инженеры чаще всего при прямом преобразовании множитель не ставят, а при обратном ставят  $1/2\pi$  (в дискретном преобразовании  $1/N_x N_y$ ). Особого смысла в этом нет, но надо быть внимательным при сравнении своих результатов с другими результатами, встречающимися в литературе.

Двумерное преобразование можно выполнить в виде последовательных одномерных преобразований по строкам и по столбцам. Аналогичные соотношения и теоремы справедливы и для двумерного случая с той разницей, что последовательности следует рассматривать как периодические по двум индексам.



Рассмотрим свойства одномерного дискретного преобразования Фурье. Пусть выражения для них задаются в следующем виде:

$$a_r = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} \alpha_k \exp \left[ i \frac{2\pi kr}{N} \right] - \text{прямое ДПФ}; \quad (4.1.15)$$

$$a_k = \frac{1}{\sqrt{N}} \sum_{r=0}^{N-1} \alpha_k \exp \left[ -i \frac{2\pi kr}{N} \right] - \text{обратное ДПФ}. \quad (4.1.16)$$

Дискретное преобразование Фурье обладает следующими свойствами.

1. Теорема о комплексно-сопряженной симметрии:

$$\begin{aligned} \{a_{N-k}\} &\leftarrow \text{ДПФ} \rightarrow \{\pm \alpha_{N-r}\}, \\ \{a_k^*\} &\leftarrow \text{ДПФ} \rightarrow \{\alpha_{N-r}^*\}, \\ \{a_k = \pm a_k^*\} &\leftarrow \text{ДПФ} \rightarrow \{\alpha_r = \pm \alpha_{N-r}^*\}. \end{aligned} \quad (4.1.17)$$

Понятия симметрии, четности и нечетности, в отличие от непрерывных сигналов, определены не относительно точки с номером 0, а относительно точки с номером  $N/2$ .

2. Спектральная составляющая в центре координат частотной плоскости равна увеличенному в  $\sqrt{N}$  раз среднему (по исходной плоскости) значению яркости изображения:

$$a_0 = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} \alpha_k. \quad (4.1.18)$$

В результате центральный пик становится намного больше, чем все остальные гармоники. Для визуализации высокочастотных гармоник приходится обрезать значения, большие определенного порядка (рис. 4.3).

3. Цикличность:

$$a_r = \alpha_{(r) \bmod N}. \quad (4.1.19)$$

Спектр Фурье является, по существу, представлением исходного поля в виде ряда Фурье. Для того чтобы это представление было верным (ряд должен сходиться), предполагается, что исходное изображе-



ние также обладает периодичностью. Поэтому последовательность коэффициентов ДПФ также является периодической функцией.

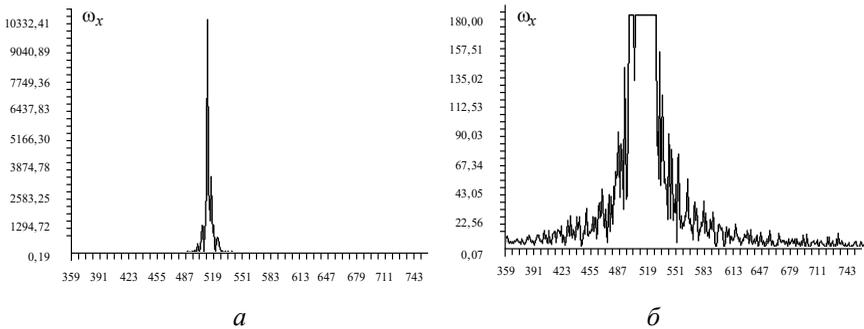


Рис. 4.3. Амплитуда спектра:

*a* – действительный график амплитуды спектра по центральной строке;  
*б* – график амплитуды спектра после обрезания

#### 4. Теорема сдвига:

$$\{a_{k+k_0 \bmod N}\} \leftarrow \text{ДПФ} \rightarrow \alpha_r \exp(-i2\pi k_0 r / N). \quad (4.1.20)$$

Если сигнал линейно сдвинуть по координате, то его фурье-преобразование умножается на комплексную экспоненту. При умножении на экспоненту коэффициенты преобразования Фурье сдвигаются.

#### 5. Теорема о циклической свертке:

$$\left\{ \sum_{k=0}^{N-1} a_n b_{(k-n) \bmod N} \right\} \leftarrow \text{ДПФ} \rightarrow \sqrt{N} \cdot \alpha_r \beta_r. \quad (4.1.21)$$

В отличие от интегрального преобразования Фурье, ДПФ от произведения дискретных спектров сигналов дает не арифметическую, а циклическую свертку, т. е. свертку сигналов, периодически продолженных за пределы заданного интервала.

У ДПФ есть еще одно замечательно свойство – свойство двойственности, которое заключается в том, что все свойства ДПФ справедливы как для сигнала, так и для спектра. Например, спектр циклической свертки сигналов есть произведение спектров. В то же время это можно сформулировать и в обратную сторону: спектр произведения сигналов есть циклическая свертка спектров этих сигналов. Аналогично сдвиг во времени приводит к умножению спектра на комплекс-



ную экспоненту, в то время как умножение сигнала на комплексную экспоненту приводит к циклическому сдвигу спектра.

В дифракционной картине, полученной с помощью оптической системы, гармоника с нулевой частотой оказывается в центре изображения (рис. 4.4). В результате дискретного преобразования нулевые частоты располагаются в начале координат. Для того чтобы начало координат оказалось в центре, необходимо переставить координаты точек.

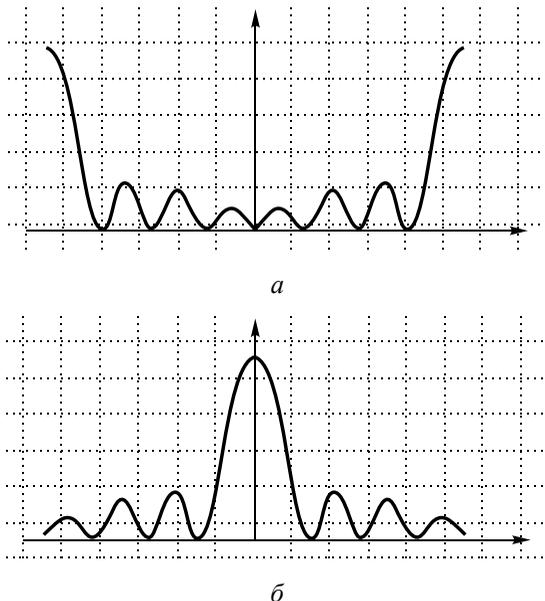


Рис. 4.4. Циклическое смещение дискретного преобразования на  $N/2$  :

$a$  – непрерывное преобразование Фурье;  
 $b$  – дискретное преобразование Фурье

Это можно сделать с помощью теоремы сдвига. Для того чтобы сдвинуть спектр на  $N/2$ , достаточно сигнал умножить на

$$a_k \exp\left(-i \frac{2\pi N}{N} \frac{k}{2}\right) = (-1)^k a_k.$$

Таким образом, достаточно каждый второй отсчет его спектра умножить на минус единицу.



Однако, если мы будем делать преобразование Френеля, такая операция может привести к фазовым искажениям. В этом случае проще сделать простую перестановку координат, как показано на рис. 4.5.

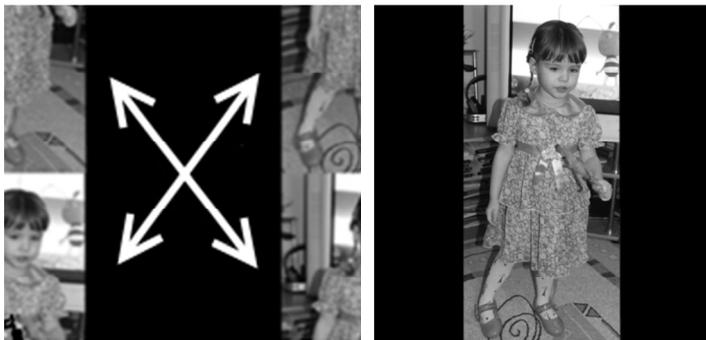


Рис. 4.5. Перестановка точек для устранения циклического смещения

Такую перестановку необходимо выполнить для переноса частот в центр изображения после преобразования Френеля или Фурье (рис. 4.6).



*а б в*

Рис. 4.6. Двумерное преобразование Фурье:

*а* – амплитуда объекта, фаза равна нулю; *б* – амплитуда фурье-спектра после преобразования; *в* – амплитуда спектра после перестановки координат

В этом разделе рассмотрены основные свойства ДПФ: линейность, свойства временного и частотного сдвигов, спектр свертки и произведения сигналов.

Если непосредственно производить вычисления по формулам (4.1.15) и (4.1.16), то для нахождения всех  $N$  коэффициентов необходимо опреде-



лать примерно  $N^2$  комплексных операций ( $N$  умножений и  $N(N - 1)$  комплексных сложений). Даже при небольших значениях  $N$  это число может оказаться очень большим.

Практическое значение при числовой обработке сигналов ДПФ приобрело после изобретения так называемых быстрых алгоритмов преобразования Фурье (БПФ).

#### 4.1.2. БЫСТРОЕ ПРЕОБРАЗОВАНИЕ ФУРЬЕ

Алгоритм быстрого преобразования Фурье предложили Дж. Кули и Дж. Таки в 1965 г. [3].

Возможность сокращения числа операций при вычислении ДПФ станет очевидной, если рассмотреть двумерное ДПФ. Действительно, для двумерного массива размерностью  $NM$ , благодаря тому что двумерное ДПФ разбивается на два одномерных, количество операций равно приблизительно  $NM^2$  для ДПФ по строкам и  $MN^2$  по столбцам, т. е.  $NM(M + N)$ , а не  $N^2M^2$  операций, как это было бы для одномерного массива той же длины. Для  $n$ -мерного массива размерностью  $N_1N_2 \dots N_n$  количество требуемых операций будет порядка  $N_1N_2 \dots N_n \times (N_1 + N_2 + \dots + N_n)$ , т. е. в  $(N_1N_2 \dots N_n) / (N_1 + N_2 + \dots + N_n)$  раз меньше, чем для одномерного массива такой же длины. То есть если представить одномерное ДПФ в виде многомерного, то число операций резко сократится. Такое представление возможно, если размерность массива определяется составным числом:

$$N = N_1N_2 \dots N_n. \quad (4.1.22)$$

Разберем эту возможность на примере двух сомножителей  $N = N_1N_2$ . В этом случае индексы  $k$  и  $s$  могут быть представлены в системе счисления с основаниями  $N_1$  и  $N_2$ :

$$k = k_2N_1 + k_1, \quad s = s_1N_2 + s_2, \quad (4.1.23)$$

где  $k_1, s_1 = 0, 1, \dots, N_1 - 1$ ;  $k_2, s_2 = 0, 1, \dots, N_2 - 1$ .



Подставив (4.1.22) в выражение для одномерного ДПФ (4.1.15), получим

$$\begin{aligned}
 a(s_1, s_2) &= \frac{1}{\sqrt{N_1 N_2}} \sum_{k_2=0}^{N_2-1} \sum_{k_1=0}^{N_1-1} a(k_2, k_1) \exp \left[ i2\pi \frac{(k_2 N_1 + k_1)(s_1 N_2 + s_2)}{N_1 N_2} \right] = \\
 &= \frac{1}{\sqrt{N_1 N_2}} \sum_{k_1=0}^{N_1-1} \exp \left[ i2\pi \left( \frac{k_1 s_1}{N_1} + \frac{k_1 s_2}{N_1 N_2} \right) \right] \sum_{k_2=0}^{N_2-1} a(k_2, k_1) \exp \left[ i2\pi \frac{k_2 s_2}{N_2} \right] = \\
 &= \text{ДПФ}_{k_1} \left\{ \exp \left[ i2\pi \frac{k_1 s_2}{N} \right] \text{ДПФ}_{k_2} (a(k_1, k_2)) \right\}. \quad (4.1.24)
 \end{aligned}$$

Таким образом, исходное ДПФ оказалось сведенным к двум ДПФ, производимым над уменьшенными массивами.

Простые для реализации алгоритмы получаются тогда, когда  $N = 2^n$ . В этом случае количество требуемых операций уменьшается от  $N^2$  до  $N \log_2 N = Nn$ , т. е. при больших  $N$  число операций уменьшается в десятки и сотни раз.

Заметим, что при такой организации вычислений коэффициенты получаются в так называемом инверсном порядке, определенном инверсией разрядов в позиционно-численном представлении их номеров по сложному основанию.

Рассмотрим  $N = 2^3$  (табл. 4.1).

Т а б л и ц а 4.1

#### Определение номера инверсии

0	000	–	000	0
1	001	–	100	4
2	010	–	010	2
3	011	–	110	6
4	100	–	001	1
5	101	–	101	5
6	110	–	011	3
7	111	–	111	7



Во втором столбце этой таблицы приведены двоичные представления номеров первого столбца. Число двоичных разрядов равно степени числа размера массива  $N = 2^3$  (степень равна трем, число разрядов тоже равно трем). В четвертом столбце представлены инвертированные относительно центра двоичные разряды, в пятом – соответствующие им числа в десятичном представлении.

Вычислительный процесс БПФ наглядно характеризуется графом. Для  $N = 16$   $m = 4$  ( $2^4 = 16$ ). Перевод в коэффициенты Фурье осуществляется за четыре итерации (рис. 4.7).

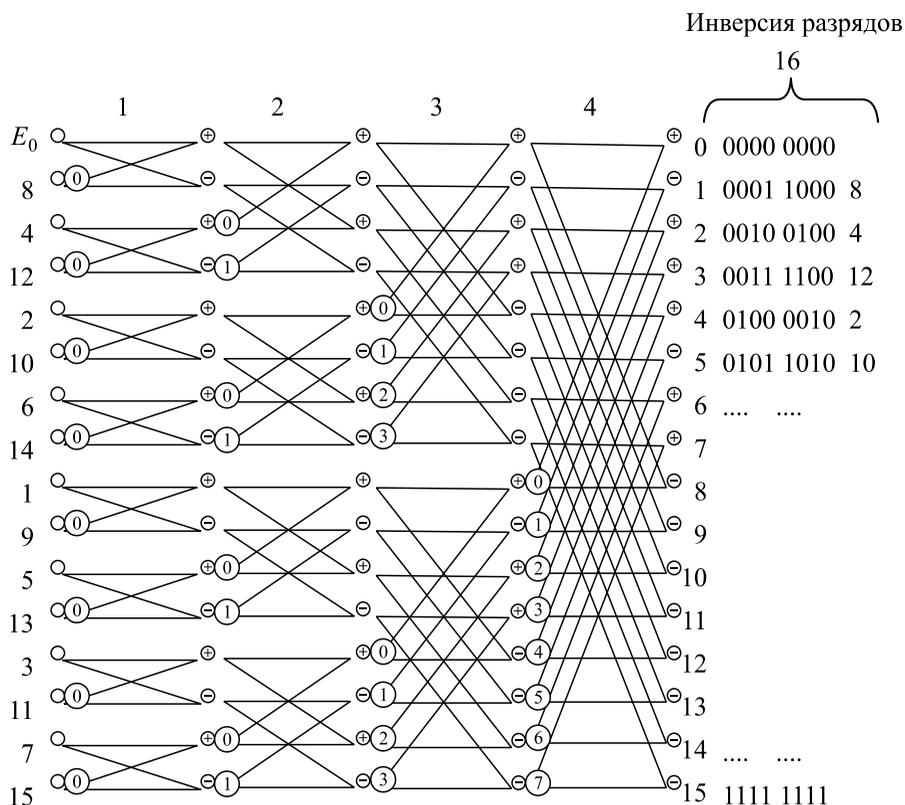


Рис. 4.7. Схема вычислительного алгоритма быстрого преобразования Фурье для  $N = 2^4 = 16$



На каждой итерации предыдущий массив переводится в новый массив путем сложения и вычитания. Знаки в кружках означают операции, которые нужно проделать с числом исходного массива, чтобы получить число, соответствующее данному кружку. Парные кружки означают, что данное число нужно умножить на  $\exp\left(i2\pi\frac{p}{2^k}\right)$ , причем  $k$  соответствует номеру итерации, а значение  $p$  записано во втором кружочке.

Существует множество программных реализаций этой схемы. Ниже показан наиболее простой текст программы на языке C++. Исходный алгоритм приведен в [2] на языке Фортран.

```
void fft(double_complex a[], int m, int kk)
{ double_complex u,w,t;
  inti,j,ip,k,l;
  if (kk> 0) kk=1; else kk=-1;
  int n=(int)pow(2.,m); int n1=n>>1;
```

```
for (i=0, j=0, k=n1; i<n-1; i++, j=j+k)
{ if (i<j) { t=a[j]; a[j]=a[i]; a[i]=t; }
  k=n1;
while(k<=j){ j=j-k; k=k>>1; }
}
```

```
double pi=3.141592653589793;
```

```
for (l=1; l<=m; l++)
{ int l1=(int)pow(2.,l), l11=l1>>1;
  u=double_complex(1.,0.);
  w=double_complex(cos(pi/l11),kk*sin(pi/l11));
for (j=1; j<=l11; j++)
{ for (i=j-1; i<n; i=i+l1)
{ ip=i+l11; t=a[ip]*u; a[ip]=a[i]-t; a[i]=a[i]+t; }
  u=u*w;
}
}
```

```
for (i=0; i<n; i++) a[i]=a[i]/sqrt((double)n);
}
```



В подпрограмме три параметра:

- 1) комплексный массив входных значений  $\mathbf{a} [ ]$ ;
- 2) степень, соответствующая размеру массива  $\mathbf{m}$  (например, для 256 значений степень равна 8, поскольку  $256 = 2^8$ );
- 3)  $\mathbf{k}\mathbf{k} = 0$  для быстрого преобразования,  $\mathbf{k}\mathbf{k} = 1$  для обратного преобразования.

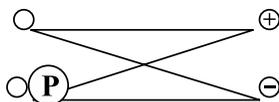
Результаты преобразования записываются в комплексный входной массив.

Программу можно усовершенствовать, перейдя к целочисленной арифметике, и хранить все вычисляемые комплексные значения  $\sin( )$  и  $\cos( )$  в промежуточных массивах. При этом скорость выполнения значительно возрастет.

Подпрограмма состоит из двух частей.

1. Первый выделенный прямоугольник. Выполняется инверсное перемешивание. Эту часть алгоритма можно написать в машинных кодах для ускорения выполнения.

2. Второй прямоугольник. Операторы, которые выполняют основную вычислительную процедуру на каждой итерации над парой комплексных чисел.



На каждой итерации предыдущий массив переводится в новый массив путем сложения и вычитания. Знаки в кружках означают операции, которые нужно проделать с числом исходного массива, чтобы получить число, соответствующее данному кружку. Парные кружки означают, что данное число нужно умножить на  $\exp\left(i2\pi\frac{p}{2^k}\right)$ , причем  $k$  соответствует

номеру итерации, а  $p$  записано во втором кружке на рис. 4.7. Эта операция получила название «бабочка» из-за внешнего вида графа. Основное время вычисления алгоритма зависит от выполнения этой элементарной операции. Для вычисления в реальном времени создаются специальные спецпроцессоры, в которых «бабочка» выполняется с максимальной скоростью. Параллельные одновременные вычисления всех таких опе-



раций в одной итерации позволяют создать эффективные быстродействующие специализированные устройства.

Можно сократить время вычислений, поскольку вычисления порядка перемешивания (первый прямоугольник) и вычисление экспонент для операции «бабочка» (во втором прямоугольнике) одни и те же для строк одинакового размера. Поэтому если вычисления производятся над некоторым массивом со строками одинакового размера, можно вынести эти операции за пределы итераций. Существуют сотни реализаций этого алгоритма, которые несложно найти в сети Интернет. Программист высокого уровня может значительно сократить время выполнения данной операции. Однако общий принцип остается таким же.

Если необходимо выполнить быстрое преобразование Фурье над массивом чисел, несложно показать, что для этого следует провести одномерные преобразования над каждой строкой, а затем одномерные преобразования над столбцами чисел, которые определились в результате действий над строками. В этом случае необходимо, чтобы размер строк  $N = 2^{m_1}$  и размер столбцов  $M = 2^{m_2}$  равнялись степени два.

Изобретение быстрого преобразования Фурье приписывается Кули (Cooley) и Дж. Таки [3] в 1965 г. На самом деле БПФ неоднократно изобреталось до этого, но важность его в полной мере не осознавалась до появления современных компьютеров. Некоторые исследователи приписывают открытие БПФ К. Рунге (K. Runge) и Д. Кёнигу (D. König) в 1924 г. Наконец, открытие этого метода приписывается еще и И.К. Гауссу (J.K. Gauss) в 1805 г.

Следствием популярности БПФ алгоритма Кули–Тьюки явилось широко распространенное мнение, что дискретное преобразование Фурье практично применять лишь при длине блока, равной степени два. Существуют хорошие БПФ алгоритмы практически для всех произвольных длин блока.

Алгоритм быстрого преобразования Фурье, основанный на разбиении числа точек на последовательности, кратные двум, не является наиболее быстрым. В 1969 г. значительный шаг в ускорении вычислений дискретного преобразования сделал Ш. Виноград [4], который разработал алгоритм, снижающий число умножений примерно в пять раз. Такое сокращение числа умножений сопровождается небольшим увеличением числа сложений. В большинстве случаев общий рост производительности не превышает 20 %. Алгоритм достаточно эффективен.



Алгоритм Винограда основан на разбиении  $N$  на число сомножителей, которые являются простыми числами:

$$N = \prod_{k=1}^{\chi} N_k, \quad (4.1.25)$$

где  $N_k$  – взаимно простые числа. К сожалению, этот алгоритм не получил широкого распространения из-за сложности построения алгоритмов вычислительных схем при компьютерной реализации.

### 4.1.3. АЛГОРИТМ БЛЮСТЕЙНА (BLUESTEIN) ДЛЯ ВЫЧИСЛЕНИЯ ДИСКРЕТНОГО ПРЕОБРАЗОВАНИЯ ФУРЬЕ НАД ПРОИЗВОЛЬНЫМ КОЛИЧЕСТВОМ ТОЧЕК

Иногда выгоднее вычислять свертку через преобразование Фурье, хотя и на длине массивов, отличных от исходных [13]. Два различных способа перехода от преобразований Фурье к свертке дают чирп-алгоритм Блюстейна и алгоритм Рейдера для простых чисел.

Алгоритм Блюстейна наиболее часто встречается в программных пакетах, осуществляющих преобразование Фурье. Достоинство алгоритма заключается в том, что на число точек в исходной последовательности не накладываются никакие ограничения. Рассмотрим сам алгоритм и его программную реализацию.

Дискретное преобразование Фурье задается следующей формулой:

$$X_k = \sum_{n=0}^{N-1} x_n e^{-i \frac{2\pi}{N} nk}, \quad k = 0, \dots, N-1. \quad (4.1.26)$$

Заменим  $nk$  в экспоненте на выражение

$$nk = \frac{-(k-n)^2}{2} + \frac{n^2}{2} + \frac{k^2}{2}. \quad (4.1.27)$$

Тогда выражение для дискретного преобразования примет вид

$$X_k = \sum_{n=0}^{N-1} x_n e^{-i \frac{2\pi}{N} \left( \frac{-(k-n)^2 + n^2 + k^2}{2} \right)} = e^{-i \frac{\pi k^2}{N}} \sum_{n=0}^{N-1} \left( x_n e^{-i \frac{\pi}{N} n^2} \right) e^{i \frac{\pi}{N} ((k-n)^2)},$$

$$k = 0, \dots, N-1. \quad (4.1.28)$$



Проведем замену переменных

$$a_n = x_n e^{-i\frac{\pi}{N}n^2}, \quad b_n = e^{i\frac{\pi}{N}n^2}. \quad (4.1.29)$$

Заметим, что последовательность  $a_n$  состоит из исходной последовательности данных  $x_n$ , умноженной на сигнал  $e^{-i\frac{\pi}{N}n^2}$ . Такие сигналы называются чирп-сигналами по очень простой причине: звук от этого сигнала подобен чириканью (chirp) птицы, если его прослушать через громкоговоритель.

По этой же причине алгоритм Блустейна также называют чирп-преобразованием. В русскоязычной литературе встречается следующее название – алгоритм  $z$ -преобразования с ЛЧМ (линейно-частотной модуляцией) фильтрацией [2].

Выражение (4.1.28) есть свертка последовательностей  $a_n$  и  $b_n$ , умноженная на  $N$  фазовых коэффициентов  $b_k^*$ :

$$X_k = b_k^* \sum_{n=0}^{N-1} a_n b_{k-n}, \quad k = 0, \dots, N-1. \quad (4.1.30)$$

Таким образом, от прямого вычисления дискретного преобразования Фурье мы перешли к свертке двух последовательностей. Для нахождения свертки достаточно выполнить произведение образов функций в частотной области и затем взять обратное преобразование. Однако с помощью дискретного преобразования Фурье мы можем получить только циклическую свертку.

Из теоремы о циклической свертке (4.1.21)

$$\left\{ \sum_{k=0}^{N-1} a_n b_{(k-n) \bmod N} \right\} \leftarrow \text{ДПФ} \rightarrow \sqrt{N} \alpha_r \beta_r$$

видно, что в результате такой операции возникают перекрытия на краях. Нам необходимо определить линейную свертку с помощью циклической свертки.

Линейной сверткой называют последовательность, определяемую соотношением  $y_n = \sum_{k=0}^{N-1} a_n b_{(k-n)}$ , где  $a_n$  и  $b_{(k-n)}$  равны нулю вне со-



ответствующих интервалов. На рис. 4.8 приведены примеры последовательностей  $a_n$ ,  $b_n$  и  $y_n$ . Из рисунка видно, что последовательность  $y_n$  является конечной и имеет длину  $N_1 + N_2 - 1$  отсчетов [2].

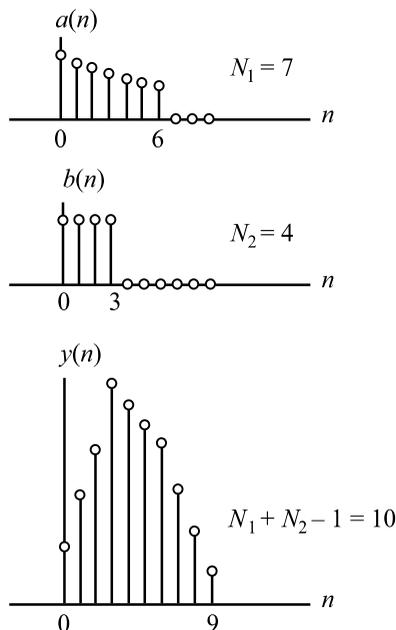


Рис. 4.8. Линейная свертка

$$y_n = \sum_{k=0}^{N-1} a_n b_{(k-n)}$$

Свертка периодических последовательностей периодична и имеет тот же период, что и сами последовательности. Поскольку период свертки (рис. 4.8) равен  $N_1 + N_2 - 1$  отсчетам, то для получения такого периода при круговой свертке необходимо, чтобы  $a_n$  и  $b_n$  содержали по  $N_1 + N_2 - 1$  отсчетов, что достигается дополнением каждой из двух последовательностей соответствующим числом нулевых отсчетов. После этого можно найти  $N_1 + N_2 - 1$  точечные ДПФ дополненных последовательностей, перемножить их и выполнить обратное ДПФ произведения.



На рис. 4.9 изображены эквивалентные периодические последовательности, используемые при вычислении круговой свертки. Ясно, что дополнение исходных последовательностей конечной длины  $a_n$  и  $b_n$  нулевыми отсчетами доводит период до нужной величины и позволяет устранить круговые наложения, характерные для циклической свертки.

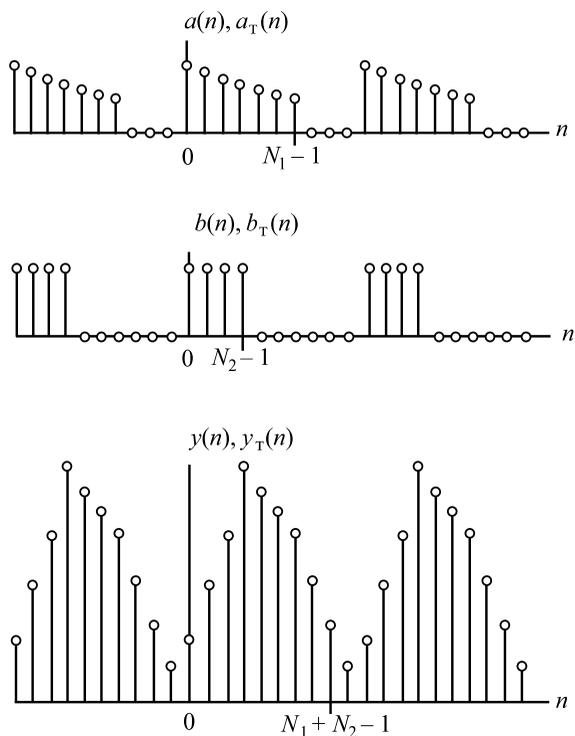


Рис. 4.9. Вычисление линейной свертки с помощью круговой свертки

В результате каждый период последовательности  $y_T(n)$  (рис. 4.9) совпадает с  $y_n$  (рис. 4.8). Для практических приложений важно отметить, что в рассмотренном выше примере размер ДПФ не обязательно ограничивать величиной  $N_1 + N_2 - 1$ , ДПФ можно выполнять по любому числу отсчетов, удовлетворяющему условию  $L \geq N_1 + N_2 - 1$ . Если это условие удовлетворяется, то последовательности  $a_n$  и  $b_n$  до-



полняются нулевыми отсчетами до размера  $L$ . В результате периодическая последовательность  $y_T(n)$  будет иметь в конце периода нули. Ясно, что эти отличия никак не искажают желаемого результата.

Рассмотренный метод вычисления свертки двух конечных последовательностей с применением алгоритма ДПФ называется быстрой сверткой в противоположность методу прямого вычисления суммы

$$y_n = \sum_{k=0}^{N-1} a_n b_{(k-n)},$$

называемой прямой или медленной сверткой. Термин

«быстрая» применяется потому, что ДПФ можно вычислить быстро и эффективно, используя один из алгоритмов БПФ. Для этого число отсчетов  $N_1 + N_2 - 1$  нужно дополнить до ближайшей степени числа два.

Для вычисления свертки мы используем алгоритм быстрого преобразования Фурье с основанием два. Для этого увеличиваем размер массивов до ближайшей степени числа два. Этот размер должен быть больше, чем  $2N - 1$ . Можно показать, что этот размер будет не больше, чем  $4N - 3$ .

Однако в этом случае размер массива может быть достаточно большим. Например, если вместо  $5184 \times 3456$  взять  $10361 \times 6911$  ( $2N - 1$ ) и дополнить размер до ближайшей степени числа два, получим размер массива  $16384 \times 8192$ .

Алгоритм Блустейна состоит из двух шагов.

1. Производим БПФ над последовательностями  $a_n$  и  $b_n$ , дополненными нулями до размера массивов к ближайшей степени числа два,

большого чем  $2N - 1$ . Поскольку  $b_n = e^{i\frac{\pi}{N}n^2}$  не зависит от исходных данных, для массивов можно сделать только одно вычисление БПФ для  $b_n$  в начале цикла;

2) затем делаем обратное преобразование и умножаем на фазовые коэффициенты  $b_k^*$ .

Поскольку используется быстрое преобразование Фурье, количество операций пропорционально  $2N_1 \log_2 N_1$ , где  $N_1$  – ближайшее число степени два от  $2N - 1$ .

Конечно, число вычислений намного больше, чем для БПФ. Однако алгоритм очень простой и легко реализуется на графических ускорителях.



#### 4.1.4. БЫСТРОЕ ПРЕОБРАЗОВАНИЕ ФУРЬЕ ДЛЯ ЧЕТНОГО КОЛИЧЕСТВА ТОЧЕК

Мы рассмотрели случай, когда число элементов преобразования равно степени два. К сожалению, на данный момент нет столь же эффективного алгоритма вычисления дискретного преобразования Фурье для произвольного  $N$ . Однако существует алгоритм, который лучше прямого вычисления преобразования Фурье. В этом алгоритме требуется, чтобы число  $N$  было четным [10]. Допустим, что  $N = M2^T = ML$ .

Рассмотрим сам алгоритм. Он состоит из двух этапов.

**Этап 1.** Выбираем  $L$  элементов с шагом  $M$ . Так как  $L$  – степень два, мы можем применить к полученной последовательности обычный алгоритм БПФ:

$$Y_{h+gM} = \sum_{l=0}^{L-1} x_{h+lM} e^{-\frac{2\pi i g l L}{L}}, \quad (4.1.31)$$

где  $h$  проходит значения  $0, 1, \dots, M-1$ , а  $g$  – значения  $0, 1, \dots, L-1$ .

Результаты преобразования записываем в те же ячейки, откуда они были взяты.

**Шаг 0:** выбираем элементы с номерами  $0, M, 2M, \dots, (L-1)M$ . Выполняем БПФ. Результаты преобразования записываем в те же ячейки.

**Шаг 1:** выбираем элементы с номерами  $1, 1+M, 1+2M, \dots, 1+(L-1)M$ . Выполняем БПФ. Результаты преобразования записываем в те же ячейки.

**Шаг 2:** выбираем элементы с номерами  $2, 2+M, 2+2M, \dots, 2+(L-1)M$ . Выполняем БПФ. Результаты преобразования записываем в те же ячейки.

...

**Шаг  $M-1$ :** выбираем элементы с номерами  $M-1, M-1+M, M-1+2M, \dots, M-1+(L-1)M$ . Выполняем БПФ. Результаты преобразования записываем в те же ячейки.

**Этап 2.** На втором этапе вводим новый массив  $Y$  размером в  $N$  элементов. Новые элементы  $X$  определяем по элементам  $Y$ , вычисленным на первом шаге, по формуле

$$X_{r+sL} = \sum_{m=0}^{M-1} Y_{m+rM} e^{-\frac{2\pi i m (r+sL)}{N}}. \quad (4.1.32)$$



В двойном цикле величина  $s$  проходит значения  $0, 1, \dots, M-1$ , а величина  $r$  – значения  $0, 1, \dots, L-1$ . Общее число итераций, таким образом, равно  $N = ML$ . Каждая итерация требует суммирования  $M$  элементов.

Несложно проверить, что формула (4.1.32) действительно определяет дискретное преобразование Фурье.

Сложность этого алгоритма равна  $N^2 / L + N \log_2 L$ . Видно, что алгоритм тем эффективнее, чем больше  $L$ , т. е. чем больше число элементов  $N$  «похоже» на степень два. В худшем случае, когда  $L = 2$ , алгоритм лишь немногим лучше прямого вычисления, которое имеет сложность  $N^2$ .

Пусть необходимо проанализировать изображение, полученное с матрицы Canon, размером  $5184 \times 3456$ . Размер по оси  $x$  – 5184. Если использовать стандартный алгоритм БПФ, придется отрезать часть сигнала до  $4096 = 2^{12}$ , т. е. выбросить 1088 точек или дополнить исходный сигнал до  $8192 = 2^{13}$  точек. При этом возникает проблема: как интерполировать сигнал на дополнительные точки, чтобы не появились лишние гармоники (шумы). При использовании нового алгоритма необходимо разбить  $N = 5184 = 81 \cdot 64 = 81 \cdot 2^6$  ( $M = 81$ ,  $L = 64$ ). Если выполнять обычное БПФ, то придется дополнить число точек по строкам до  $8192 = 2^{13}$ . Сложность алгоритма БПФ по основанию два равна  $N \log_2 N = 8192 \cdot 13 = 106\,496$ . Сложность нового алгоритма

$$N^2 / L + N \log_2 L = 5184^2 / 64 + 5184 \cdot 6 = 419\,904 + 31\,104 = 451\,008.$$

Таким образом, сложность увеличивается в 4,3 раза.

По столбцам  $N = 3456 = 27 \cdot 128 = 81 \cdot 2^7$  ( $M = 27$ ,  $L = 128$ ). Для вычисления по обычному алгоритму БПФ придется дополнить число точек по строкам до  $4096 = 2^{12}$ . Сложность алгоритма БПФ по основанию два равна  $N \log_2 N = 4096 \cdot 12 = 49\,152$ . Сложность нового алгоритма

$$N^2 / L + N \log_2 L = 3456^2 / 128 + 3456 \cdot 7 = 93\,312 + 24\,192 = 117\,504.$$

Таким образом, сложность увеличивается всего в 2,4 раза.



Основная часть увеличения времени выполнения зависит от члена в выражении сложности  $N^2/L$ . Видно, что с ростом  $L$  быстродействие увеличивается. Если подобрать число точек с большим множителем  $L$ , можно значительно увеличить быстродействие. Например, если вместо  $5184 \times 3456$  взять 5120 точек, то  $M = 5$ ,  $L = 1024$ , время выполнения сократится в 2,5 раза.

Время выполнения на компьютере с процессором Pentium 2.6 ГЦ показано в табл. 4.2.

Таблица 4.2

#### Быстродействие алгоритма при разном числе точек

БПФ по основанию два	$8192 \times 4092$	28 с 514 мс
БПФ с четным $N$	$5184 \times 3456$	1 мин 8 с 922 мс
БПФ с четным $N$	$5120 \times 3456$	27 с 860 мс

Таким образом, этот алгоритм зависит от правильного подбора размеров исходных массивов. В некоторых случаях незначительное изменение размеров массива приводит к значительному сокращению времени вычислений.

Видно, что для размеров массивов, показанных в табл. 4.2, скорость вычисления БПФ с четным  $N$  примерно такая же, как и для БПФ с основанием два над расширенным массивом. Однако искажений гармоник, вызванных добавлением нулей при расширении массива, не произойдет.

Для стандартных размеров матрицы Canon можно предложить следующие ограничения (табл. 4.3).

Таблица 4.3

#### Оптимальные размеры точек

Число точек	$(M*L)$	Новое число точек	$(M*L)$
$5184 \times 3456$	$(81*64) \times (27*128)$	$5120 \times 3328$	$(5*1024) \times (13*256)$
$4752 \times 3168$	$(297*17) \times (99*32)$	$4736 \times 3072$	$(37*128) \times (3*1024)$

Такая процедура может значительно сократить время вычислений, а число потерянных точек может быть нечувствительным к результатам эксперимента в тех случаях, когда значимая часть изображения занимает не весь кадр.



## 4.1.5. МАТРИЧНОЕ ДИСКРЕТНОЕ ПРЕОБРАЗОВАНИЕ ФУРЬЕ

Дискретное преобразование Фурье может быть представлено в матричном виде. ДПФ является линейным преобразованием, которое переводит вектор отсчетов  $\vec{x}$  в вектор спектральных отсчетов той же длины. Таким образом, преобразование может быть реализовано как умножение симметричной квадратной матрицы на вектор:

$$\vec{X} = A\vec{x}. \quad (4.1.33)$$

Матрица  $A$  имеет вид

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 & \dots & 1 \\ 1 & e^{-\frac{2\pi i}{N}} & e^{-\frac{4\pi i}{N}} & e^{-\frac{6\pi i}{N}} & \dots & e^{-\frac{2\pi i}{N}(N-1)} \\ 1 & e^{-\frac{4\pi i}{N}} & e^{-\frac{8\pi i}{N}} & e^{-\frac{12\pi i}{N}} & \dots & e^{-\frac{2\pi i}{N}2(N-1)} \\ 1 & e^{-\frac{6\pi i}{N}} & e^{-\frac{12\pi i}{N}} & e^{-\frac{18\pi i}{N}} & \dots & e^{-\frac{2\pi i}{N}3(N-1)} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & e^{-\frac{2\pi i}{N}(N-1)} & e^{-\frac{2\pi i}{N}2(N-1)} & e^{-\frac{2\pi i}{N}3(N-1)} & \dots & e^{-\frac{2\pi i}{N}(N-1)^2} \end{pmatrix}. \quad (4.1.34)$$

Элементы матрицы задаются следующей формулой:

$$A(m, n) = \exp\left(-2\pi i \frac{(m-1)(n-1)}{N}\right). \quad (4.1.35)$$

Обратное преобразование задается обратной матрицей  $A^{-1}$ .



## 4.2. ПРЕДСТАВЛЕНИЕ ПРЕОБРАЗОВАНИЯ ФРЕНЕЛЯ В ДИСКРЕТНОЙ ФОРМЕ

Теперь рассмотрим дискретное преобразование Френеля [5].

$$\begin{aligned} \Gamma(\xi, \eta) = & \frac{1}{i\lambda d} \exp\left[i\frac{2\pi d}{\lambda}\right] \exp\left[i\frac{\pi(\xi^2 + \eta^2)}{\lambda d}\right] \times \\ & \times \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} b_1(x, y) \exp\left[i\frac{\pi(x^2 + y^2)}{\lambda d}\right] \times \\ & \times \exp\left[-i\frac{2\pi(x\xi + y\eta)}{\lambda d}\right] dx dy. \end{aligned} \quad (4.2.1)$$

Это соотношение получено в разделе 2.2.

Пусть объект, описываемый функцией  $b_1(x, y) \exp\left[i\frac{\pi(x^2 + y^2)}{\lambda d}\right]$ , имеет ограниченные размеры  $[-X_{\max}, X_{\max}; -Y_{\max}, Y_{\max}]$  и

$$\Delta x = \frac{2X_{\max}}{N_x}, \quad \Delta y = \frac{2Y_{\max}}{N_y}. \quad (4.2.2)$$

Введем пространственные частоты

$$v_x = \frac{\xi}{\lambda d}, \quad v_y = \frac{\eta}{\lambda d}. \quad (4.2.3)$$

По теореме отсчетов функцию

$$\Gamma(\xi, \eta) \exp\left[-i\frac{2\pi d}{\lambda}\right] \exp\left[-i\frac{\pi(\xi^2 + \eta^2)}{\lambda d}\right]$$

можно восстановить интерполяцией своих отсчетов, если

$$\Delta x = 1/2v_{x\max}, \quad \Delta y = 1/2v_{y\max}, \quad (4.2.4)$$



где  $[-v_{x\max}, v_{x\max}; -v_{y\max}, v_{y\max}]$  – область, за пределами которой пространственный спектр  $b_1(x, y) \exp\left[i \frac{\pi(x^2 + y^2)}{\lambda d}\right]$  можно считать равным нулю.

Это значит, что и функция  $\Gamma(\xi, \eta)$  может быть восстановлена по ее отсчетам:

$$\begin{aligned} \Gamma(r\Delta\xi, s\Delta\eta) \exp\left[-i \frac{2\pi d}{\lambda}\right] \exp\left[-i \frac{\pi[(r\Delta\xi)^2 + (s\Delta\eta)^2]}{\lambda d}\right] &= \\ = \frac{1}{i\lambda d} \int_{-X_{\max}}^{X_{\max}} \int_{-Y_{\max}}^{Y_{\max}} b_1(x, y) \exp\left[i \frac{\pi(x^2 + y^2)}{\lambda d}\right] \times \\ \times \exp\left[-i \frac{2\pi(xr\Delta\xi + ys\Delta\eta)}{\lambda d}\right] dx dy. \end{aligned} \quad (4.2.5)$$

Предположим, что объект описывается достаточно гладкой функцией  $b_1(x, y)$ . В этом случае  $b_1(x, y)$  может быть восстановлена по ее отсчетам путем интерполяции их некоторой функцией  $\varphi(x, y)$ :

$$\begin{aligned} b_1(x, y) \exp\left[i \frac{\pi(x^2 + y^2)}{\lambda d}\right] &= \sum_{k=0}^{N_x-1} \sum_{l=0}^{N_y-1} b_1(k\Delta x, l\Delta y) \times \\ \times \exp\left[i \frac{\pi(k^2\Delta x^2 + l^2\Delta y^2)}{\lambda d}\right] \varphi(x - k\Delta x, y - l\Delta y), \end{aligned} \quad (4.2.6)$$

где  $\Delta x$  и  $\Delta y$  – интервалы дискретизации по координатам  $x$  и  $y$ .

Выражение (4.2.5) будет выполняться, если для  $(x, y)$  внутри прямоугольника размером  $(\Delta x, \Delta y)$  с центром в точке  $(k\Delta x, l\Delta y)$

$$x^2 - (k\Delta x)^2 + y^2 - (l\Delta y)^2 \ll \lambda d. \quad (4.2.7)$$



Перепишем выражение (4.2.5), используя (4.2.6):

$$\begin{aligned}
 & \Gamma(r\Delta\xi, s\Delta\eta) \exp\left[-i\frac{2\pi d}{\lambda}\right] \exp\left[-i\frac{\pi\left[(r\Delta\xi)^2 + (s\Delta\eta)^2\right]}{\lambda d}\right] = \\
 & = \frac{1}{i\lambda d} \sum_{k=0}^{N_x-1} \sum_{l=0}^{N_y-1} b_1(k\Delta x, l\Delta y) \exp\left\{i\frac{\pi\left[k^2\Delta x^2 + l^2\Delta x^2\right]}{\lambda d}\right\} \times \\
 & \times \int_{-X_{\max}}^{X_{\max}} \int_{-Y_{\max}}^{Y_{\max}} \varphi(x - k\Delta x, y - l\Delta y) \exp\left\{-i\frac{2\pi\left[xr\Delta\xi + ys\Delta\eta\right]}{\lambda d}\right\} dx dy. \quad (4.2.8)
 \end{aligned}$$

Преобразуем интеграл следующим образом:

$$\begin{aligned}
 & \int_{-X_{\max}}^{X_{\max}} \int_{-Y_{\max}}^{Y_{\max}} \varphi(x - k\Delta x, y - l\Delta y) \exp\left\{-i\frac{2\pi\left[xr\Delta\xi + ys\Delta\eta\right]}{\lambda d}\right\} dx dy = \\
 & = \left\{ \int_{-X_{\max}-k\Delta x}^{X_{\max}-k\Delta x} \int_{-Y_{\max}-l\Delta y}^{Y_{\max}-l\Delta y} \varphi(x, y) \exp\left[-i\frac{2\pi\left(xr\Delta\xi + ys\Delta\eta\right)}{\lambda d}\right] dx dy \right\} \times \\
 & \times \exp\left[-i\frac{2\pi\left(kr\Delta\xi\Delta x + ls\Delta\eta\Delta y\right)}{\lambda d}\right]. \quad (4.2.9)
 \end{aligned}$$

Протяженность интерполирующей функции  $\varphi(x, y)$  по  $x$  и  $y$  обычно намного меньше, чем размер объекта:  $[-X_{\max}, X_{\max}; -Y_{\max}, Y_{\max}]$ , поэтому в (4.2.9) с точностью до краевых эффектов можно использовать бесконечные пределы:

$$\begin{aligned}
 & \int_{-X_{\max}-k\Delta x}^{X_{\max}-k\Delta x} \int_{-Y_{\max}-l\Delta y}^{Y_{\max}-l\Delta y} \varphi(x, y) \exp\left[-i\frac{2\pi\left(xr\Delta\xi + ys\Delta\eta\right)}{\lambda d}\right] dx dy \approx \\
 & \approx \int_{-\infty\Delta x}^{\infty} \int_{-\infty}^{\infty} \varphi(x, y) \exp\left[-i\frac{2\pi\left(xr\Delta\xi + ys\Delta\eta\right)}{\lambda d}\right] dx dy = \Phi(r\Delta\xi, s\Delta\eta). \quad (4.2.10)
 \end{aligned}$$



Подставив (4.2.10) в (4.2.8), получим

$$\begin{aligned} \Gamma(r\Delta\xi, s\Delta\eta) &= \frac{1}{i\lambda d} \exp\left[i\frac{2\pi d}{\lambda}\right] \exp\left[i\frac{\pi\left[(r\Delta\xi)^2 + (s\Delta\eta)^2\right]}{\lambda d}\right] \times \\ &\times \left\{ \sum_{k=0}^{N_x-1} \sum_{l=0}^{N_y-1} b_1(k, l) \exp\left[-i\frac{\pi\left[(k\Delta x)^2 + (l\Delta y)^2\right]}{\lambda d}\right] \right\} \times \\ &\times \exp\left[-i\frac{2\pi(kr\Delta\xi\Delta x + ls\Delta\eta\Delta y)}{\lambda d}\right] \Phi(r\Delta\xi, s\Delta\eta) \}. \end{aligned} \quad (4.2.11)$$

Сумма в правой части может рассматриваться как дискретное представление интеграла Френеля:

$$\begin{aligned} \Gamma(r, s) &= \frac{1}{i\lambda d} \exp\left[i\frac{2\pi d}{\lambda}\right] \exp\left[i\frac{\pi\left[(r\Delta\xi)^2 + (s\Delta\eta)^2\right]}{\lambda d}\right] \times \\ &\times \sum_k^{N_x-1} \sum_l^{N_y-1} b_1(k, l) \exp\left[i\frac{\pi(k^2\Delta x^2 + l^2\Delta y^2)}{\lambda d}\right] \times \\ &\times \exp\left[-i\frac{2\pi(kr\Delta\xi\Delta x + ls\Delta\eta\Delta y)}{\lambda d}\right]. \end{aligned} \quad (4.2.12)$$

Интерполяция при восстановлении объекта из голограммы производится маскированием голограммы функцией  $\Phi(r\Delta\xi, s\Delta\eta)$ .

Эффективная реализация этого алгоритма может быть осуществлена с помощью цифрового преобразования Фурье. Для того чтобы преобразовать это выражение к виду преобразования Фурье, допустим, что  $\Delta x$  и  $\Delta\xi$  такие, что

$$\Delta\xi = \frac{\lambda d}{2X_{\max}} = \frac{\lambda d}{N_x \Delta x}. \quad (4.2.13)$$



Тогда выражение (4.2.12) можно переписать в виде

$$\begin{aligned} \Gamma(r, s) = & \frac{1}{i\lambda d} \exp\left[i\frac{2\pi d}{\lambda}\right] \exp\left[i\frac{\pi\left[(r\Delta\xi)^2 + (s\Delta\eta)^2\right]}{\lambda d}\right] \times \\ & \times \sum_k^{N_x-1} \sum_l^{N_y-1} b_1(k, l) \exp\left[i\frac{\pi(k^2\Delta x^2 + l^2\Delta y^2)}{\lambda d}\right] \times \\ & \times \exp\left[-i2\pi\left(\frac{kr}{N_x} + \frac{ls}{N_y}\right)\right]. \end{aligned} \quad (4.2.14)$$

Второй множитель представляет собой преобразование Фурье над аргументом  $b_1(k, l) \exp\left[i\frac{\pi(k^2\Delta\xi^2 + l^2\Delta\eta^2)}{\lambda d}\right]$ .

Выражение (4.2.14) точно выполняется только при условии  $\Delta\xi = \frac{\lambda d}{N_x \Delta x}$  (4.2.13). Это условие фактически определяет условие равенства размеров объекта, размер области регистрации и  $\Delta x = \Delta\xi$ . Расстояние, на котором размеры пикселей совпадают, можно определить следующим образом:

$$d_{eq} = \frac{N_x \Delta x^2}{\lambda}. \quad (4.2.15)$$

Если эти условия не выполняются, возникают фазовые ошибки.

Выставлять точные значения для расстояния от объекта до регистратора сложно. Возьмем другое расстояние:  $d = \gamma_x d_{eq}$ .

Отношение размеров объекта и области регистрации преобразования Френеля связаны с расстоянием  $d$ . Следует отметить, что размеры пикселей  $\Delta x$  и  $\Delta\xi$  могут отличаться. Введем параметр  $\gamma$ , характеризующий изменение размера апертуры регистратора в области преобразования Френеля [11]:

$$\gamma_x = \frac{\Delta\xi}{\Delta x} = \frac{\lambda d}{N_x \Delta\xi^2}. \quad (4.2.16)$$



Для упрощения считаем, что  $\gamma_x = \gamma_y = \gamma$ . На рис. 4.10 показано соответствие размеров пикселей в предметной области и области регистрации.

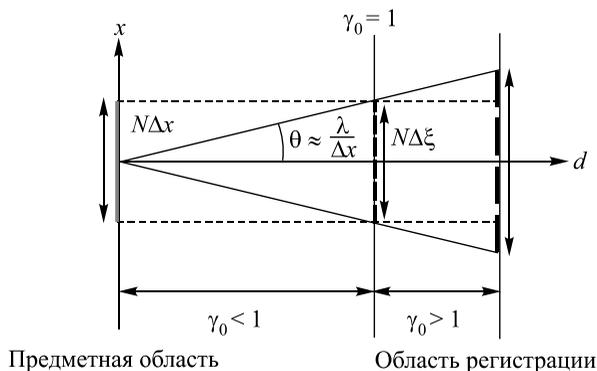


Рис. 4.10. Соответствие размеров пикселей в предметной области и области регистрации

Получим значения всех экспонент выражения (4.2.12).

$$\text{Поскольку } \Delta x = \frac{1}{\gamma} \Delta \xi,$$

$$\exp \left[ i \frac{\pi}{\lambda d} \left[ (k \Delta x)^2 + (l \Delta y)^2 \right] \right] = \exp \left[ i \frac{\pi}{\lambda d} \left[ \left( \frac{k}{\gamma} \Delta \xi \right)^2 + \left( \frac{l}{\gamma} \Delta \eta \right)^2 \right] \right]. \quad (4.2.17)$$

$$\text{Учитывая, что } \Delta \xi_{eq} = \frac{\lambda d_{eq}}{N_x \Delta \xi},$$

$$\exp \left[ -i \frac{2\pi(kr \Delta \xi \Delta x)}{\lambda d} \right] = \exp \left[ -i \frac{2\pi(kr \Delta \xi \Delta x)}{\gamma \lambda d_{eq}} \right] = \exp \left[ -i \frac{2\pi k}{N} \frac{r}{\gamma} \right]; \quad (4.2.18)$$

$$\exp \left[ -i \frac{2\pi(kr \Delta \xi \Delta x + ls \Delta \eta \Delta y)}{\lambda d} \right] = \exp \left[ -i \frac{2\pi}{N} \left( \frac{k}{\gamma} r + \frac{l}{\gamma} s \right) \right]. \quad (4.2.19)$$



Подставим значения экспонент в (4.2.14):

$$\begin{aligned} \Gamma(r, s) = & \frac{1}{i\lambda d} \exp\left[i \frac{2\pi d}{\lambda}\right] \exp\left[i \frac{\pi\left[(r\Delta\xi)^2 + (s\Delta\eta)^2\right]}{\lambda d}\right] \times \\ & \times \sum_k^{N_x-1} \sum_l^{N_y-1} b_1(k, l) \exp\left[i \frac{\pi}{\lambda d} \left[\left(\frac{k}{\gamma} \Delta\xi\right)^2 + \left(\frac{l}{\gamma} \Delta\eta\right)^2\right]\right] \times \\ & \exp\left[-i \frac{2\pi}{N} \left(\frac{k}{\gamma} r + \frac{l}{\gamma} s\right)\right]. \end{aligned} \quad (4.2.20)$$

Если сделать замену переменных  $k^I = \frac{k}{\gamma}$ , то выражение (2.8.14)

примет вид

$$\begin{aligned} \Gamma(r, s) = & \frac{1}{i\lambda d} \exp\left[i \frac{2\pi d}{\lambda}\right] \exp\left[i \frac{\pi\left[(r\Delta\xi)^2 + (s\Delta\eta)^2\right]}{\lambda d}\right] \times \\ & \times \sum_k^{N_x-1} \sum_l^{N_y-1} b_1(k, l) \exp\left[i \frac{\pi}{\lambda d} \left[(k\Delta\xi)^2 + (l\Delta\eta)^2\right]\right] \times \\ & \exp\left[-i \frac{2\pi}{N} (kr + ls)\right]. \end{aligned} \quad (4.2.21)$$

Выражение (4.3.21) можно использовать для любого расстояния  $d$  [12].

#### 4.2.1. ОСОБЕННОСТИ РАСЧЕТА ПРЕОБРАЗОВАНИЯ ФРЕНЕЛЯ ПРИ МОДЕЛИРОВАНИИ ГОЛОГРАММ

Для вычислений удобно дискретное преобразование Френеля определять с помощью дискретного преобразования Фурье, поскольку для преобразования Фурье существуют хорошо изученные методы сокращения вычислений.



Выражение для преобразования Френеля можно выразить через дискретное преобразование Фурье:

$$\Gamma(r, s) = \frac{1}{i\lambda d} \exp\left[i \frac{2\pi d}{\lambda}\right] \exp\left[i \frac{\pi\left[(r\Delta\xi)^2 + (s\Delta\eta)^2\right]}{\lambda d}\right] \times \\ \times \Im\left(b_1(k, l) \exp\left[i \frac{\pi(k^2\Delta\xi^2 + l^2\Delta\eta^2)}{\lambda d}\right]\right). \quad (4.2.22)$$

Рассмотрим первый множитель выражения (4.2.22).

$$R(r, s) = \frac{1}{i\lambda d} \exp\left[i \frac{2\pi d}{\lambda}\right] \exp\left[i \frac{\pi\left[(r\Delta\xi)^2 + (s\Delta\eta)^2\right]}{\lambda d}\right]. \quad (4.2.23)$$

Следует учесть две особенности.

1. Множитель  $-\frac{i}{\lambda d} \exp\left[i \frac{2\pi d}{\lambda}\right]$  – это большая постоянная добавка к фазе, которая при вычислении дискретного преобразования Френеля может быть отброшена, поскольку она не влияет на вид фазового фронта. Таким образом, множитель выражения (4.2.22) можно представить в виде

$$R(r, s) = -\exp\left[i \frac{\pi\left[(r\Delta\xi)^2 + (s\Delta\eta)^2\right]}{\lambda d}\right]. \quad (4.2.24)$$

2. Выражение (4.2.24) описывает сферический фазовый фронт (рис. 4.11), симметричный относительно центра изображения.

Точка с координатами  $(r, s) = (0, 0)$  находится в центре изображения, поэтому фазовый фронт (4.2.24) можно представить как

$$R(r, s) = -\exp\left\{i \frac{\pi}{\lambda d} \left[ \left( \Delta\xi \left( r - \frac{N_x}{2} \right) \right)^2 + \left( \Delta\eta \left( s - \frac{N_y}{2} \right) \right)^2 \right] \right\}. \quad (4.2.25)$$

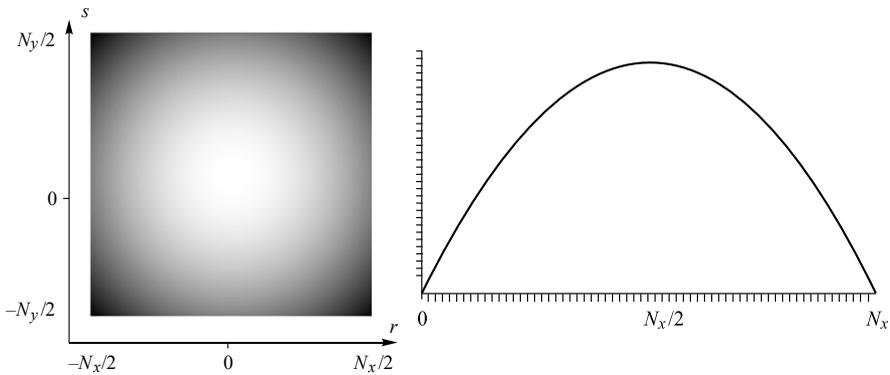


Рис. 4.11. Сферический фазовый фронт (справа график по центральной строке)

То же самое и у экспоненты, на которую умножается дискретный образ объекта в преобразовании Фурье:

$$R_1(k, l) = \exp \left\{ i \frac{\pi}{\lambda d} \left[ \left( \Delta \xi \left( k - \frac{N_x}{2} \right) \right)^2 + \left( \Delta \eta \left( l - \frac{N_y}{2} \right) \right)^2 \right] \right\}. \quad (4.2.26)$$

Таким образом, дискретное преобразование Френеля можно записать в виде

$$\Gamma(r, s) = -\exp \left\{ i \frac{\pi}{\lambda d} \left[ \left( \Delta \xi \left( r - \frac{N_x}{2} \right) \right)^2 + \left( \Delta \eta \left( s - \frac{N_y}{2} \right) \right)^2 \right] \right\} \times \\ \times \mathfrak{F} \left\{ b_1(k, l) \exp \left\{ i \frac{\pi}{\lambda d} \left[ \left( \Delta \xi \left( k - \frac{N_x}{2} \right) \right)^2 + \left( \Delta \eta \left( l - \frac{N_y}{2} \right) \right)^2 \right] \right\} \right\}. \quad (4.2.27)$$

Кривизна сферических фронтов  $R(r, s)$  и  $R_1(k, l)$  уменьшается с расстоянием  $d$ . Пусть  $\lambda = 0,5$  мкм, размер объекта – 1 см,  $N_x \times N_y =$



$= 1024 \times 1024$ , тогда при изменении расстояния кривизна фронта меняется следующим образом:

$d = 10$ см	$0 \dots 996\pi$ ,
$d = 1$ м	$0 \dots 99\pi$ ,
$d = 10$ м	$0 \dots 9,9\pi$ ,
$d = 100$ м	$0 \dots 0,996\pi$ ,
$d = 1$ км	$0 \dots 0,099\pi$ ,
$d = 6$ км	$0 \dots 0,017\pi$ .

На больших расстояниях сферический фронт становится все более плоским и постепенно преобразование Френеля переходит в преобразование Фурье.

Для расчета волнового фронта в плоскости голограммы (рис. 4.12) исходными данными являются размер апертуры фоторегистрирующей матрицы  $[-\xi_{\max}, \xi_{\max}; -\eta_{\max}, \eta_{\max}]$ , длина волны  $\lambda$ , расстояние до объекта  $d$ .

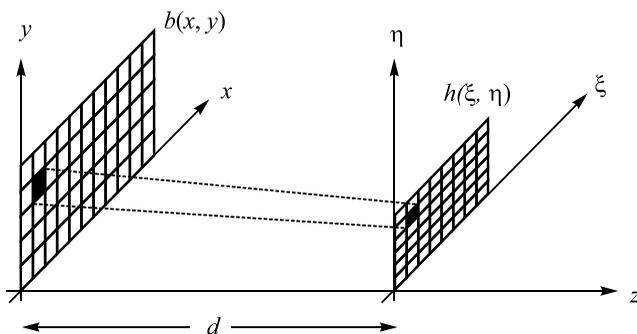


Рис. 4.12. Соответствие размеров пикселей в предметной области и области голограммы

Параметр  $\gamma = \frac{\Delta\xi}{\Delta x} = \frac{\lambda d}{N_x \Delta x^2}$   $d_{eq} = \frac{N_x \Delta \xi^2}{\lambda}$  – расстояние, на котором

$\Delta x = \Delta \xi$ . На этом расстоянии размер объекта  $2X_{\max} = 2\xi_{\max}$ .

Для других расстояний  $2X_{\max} = \frac{1}{\gamma} 2\xi_{\max}$ .



Выражение для расчета преобразования Френеля (4.2.26).

Пусть  $\lambda = 0,5$  мкм,  $N_x \times N_y = 1024 \times 1024$ , тогда при следующих размерах апертуры в области регистрации расстояние  $d_{eq}$ , на котором размеры  $\Delta x = \Delta \xi$ , будет следующим:

$$\begin{aligned} 2X_{\max} = 2\xi_{\max} = 1 \text{ см} & \quad d_{eq} = 195,3 \text{ мм}, \\ 2X_{\max} = 2\xi_{\max} = 2 \text{ см} & \quad d_{eq} = 781,25 \text{ мм}, \\ 2X_{\max} = 2\xi_{\max} = 3 \text{ см} & \quad d_{eq} = 1757,81 \text{ мм}. \end{aligned}$$

Если взять расстояние  $d = 400$  мм, то при тех же параметрах можно определить возможный размер объекта в зависимости от апертуры:

$$\begin{aligned} 2\xi_{\max} = 1 \text{ см} & \quad 2X_{\max} = 4,88 \text{ мм}, \\ 2\xi_{\max} = 2 \text{ см} & \quad 2X_{\max} = 39,06 \text{ мм}. \end{aligned}$$

Алгоритм преобразования Френеля обеспечивает простое масштабирование восстановленного изображения, однако это накладывает ряд ограничений на конструкцию измерительной системы. В частности, верхняя и нижняя границы допустимого расстояния записи голограммы становятся значимым фактором.

Необходимо определить, при каких расстояниях мы можем использовать дискретное преобразование Френеля.

#### 4.2.2. ЗОНЫ ДЕЙСТВИЯ ДИСКРЕТНЫХ ПРЕОБРАЗОВАНИЙ ФРЕНЕЛЯ И ФУРЬЕ

В ряде случаев можно использовать дискретные преобразования Френеля и Фурье, когда расстояние до объекта сравнимо с размерами объекта и голограммы, при этом объект должен находиться в ближней зоне дифракции. В этом случае условия приближений Фурье или Френеля не выполняются.

Однако для дискретных преобразований Фурье и Френеля границы применимости существенно уменьшаются.

**Границы применимости дискретных преобразований Фурье и Френеля.** Для преобразования Фраунгофера расстояние, на котором преобра-



зование можно применять, определяется выражением  $d \gg \frac{\pi}{\lambda}(x^2 + y^2)_{\max}$ .

Для преобразования Френеля

$$d^3 \gg \frac{\pi}{4\lambda} [(x-\xi)^2 + (x-\eta)^2]_{\max}^2.$$

Если размер объекта 20 см, то допустимое расстояние, на котором выполняется преобразование Фраунгофера, получается достаточно большим:

$$x_{\max} = 10 \cdot 10^{-3}, \quad y_{\max} = 10 \cdot 10^{-3} \text{ м},$$

$$x_{\max}^2 + y_{\max}^2 = 200 \cdot 10^{-6} \text{ м}^2, \quad \frac{\pi}{\lambda_1} \approx 8,6 \cdot 10^6 \text{ м}^{-1},$$

откуда  $z \gg 1720$  м при  $\lambda_1 = 365$  нм.

Если размер объекта 5 см, то допустимое расстояние, на котором может выполняться преобразование Фраунгофера, все еще получается достаточно большим:

$$x_{\max} = 2,5 \cdot 10^{-3} \text{ м}, \quad y_{\max} = 2,5 \cdot 10^{-3} \text{ м},$$

$$x_{\max}^2 + y_{\max}^2 = 12,5 \cdot 10^{-6} \text{ м}^2, \quad \frac{\pi}{\lambda_1} \approx 8,6 \cdot 10^6 \text{ м}^{-1},$$

откуда  $z \gg 107,5$  м при  $\lambda_1 = 365$  нм.

Однако если размер объекта уменьшается, то допустимое расстояние тоже резко уменьшается.

Рассмотрим рис. 4.13. На этом рисунке показано отверстие размером 50 мкм, которое является источником сферических волн.

Для преобразования Фраунгофера при диаметре отверстия 50 мкм используются значения:

$$x_{\max} = 25 \text{ мкм}, \quad y_{\max} = 25 \text{ мкм},$$

$$x_{\max}^2 + y_{\max}^2 = 1250 \text{ мкм}^2, \quad \frac{k}{2} = \frac{\pi}{\lambda_1} \approx 8,6 \text{ мкм}^{-1},$$

откуда  $d \gg 10,75$  мм при  $\lambda_1 = 365$  нм.

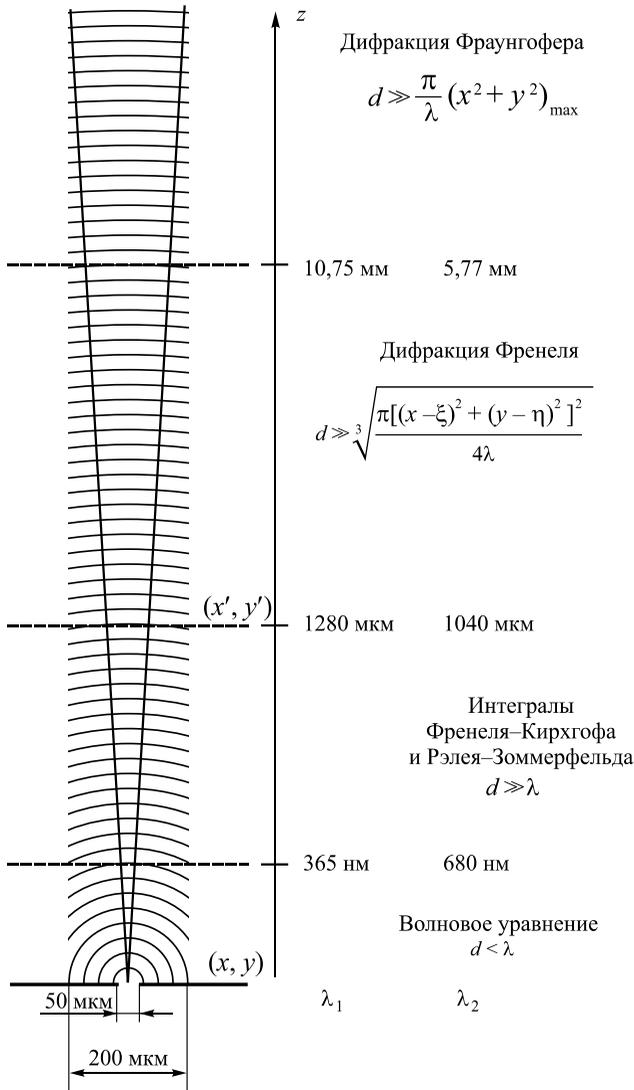


Рис. 4.13. Зоны дифракций Френеля и Фраунгофера:

на рисунке показаны зоны, в которых действуют модели дифракции Френеля и Фраунгофера при размере отверстия 50 мкм [6]. Рассматривается распространение волн с длиной  $\lambda_1 = 365$  нм и  $\lambda_2 = 680$  нм (границы видимого диапазона)



Для оценки границ применимости приближения Френеля принято, что анализ дифракционных картин ведется в области, размер диаметра которой больше, чем у объекта рассеивания, в четыре раза. Для преобразования Френеля

$$\xi_{\max} = 200 \text{ мкм}, \quad (x - \xi)_{\max} = 125 \text{ мкм} \text{ и } (x - \eta)_{\max} = 125 \text{ мкм},$$

$$\sqrt[3]{\left[ (x_{\max}^2 + y_{\max}^2)^2 \right]} = 992,1256 \text{ мкм}^2, \quad \sqrt[3]{\frac{\pi}{4\lambda_1}} \approx 1,29 \text{ мкм}^{-1},$$

откуда  $d \gg 1,280 \text{ мм}$ .

Для объектов, которые разбиваются на отдельные дискретные значения, можно провести анализ для каждого из них.

Пусть размер объекта  $2X_{\max} = 5 \cdot 10^{-3} \text{ м}$  (5 см), число элементов  $N_x = 1000$ , тогда  $\Delta x = 2X_{\max} / N_x = 5 \cdot 10^{-3} / 5000 = 5 \cdot 10^{-6} \text{ м}$  (5 мкм).

Другое условие, которое было получено при выводе дискретного преобразования Френеля (2.7.8)  $x^2 - (k\Delta x)^2 + y^2 - (l\Delta y)^2 \ll \lambda d$ , для  $(x, y)$  внутри прямоугольника размером  $(\Delta x, \Delta y)$  с центром в точке  $(k\Delta x, l\Delta y)$  является менее жестким.

#### **Необходимое разрешение фотоприемника для регистрации.**

Определим необходимое разрешение фотоприемника для регистрации преобразования Фраунгофера.

Рассмотрим, какая связь между отсчетами и физическим размерами объекта и распределением частот в области регистрации (рис. 4.14).

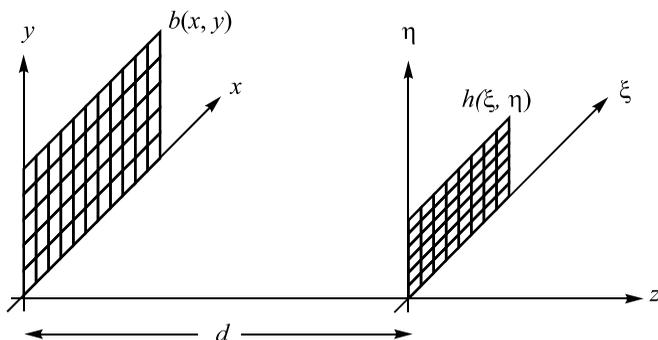


Рис. 4.14. Соответствие размеров пикселей в предметной области и области голограммы



Размеры пикселя в плоскости преобразований Фурье будут зависеть от расстояния восстановления и длины волны. Рассмотрим эти значения для конкретных величин.

Дискретное преобразование Френеля использует сумму волн, каждая из которых рассеивается на площади размером  $\Delta x$ . Если  $\Delta x$  достаточно мало, то возможно выбрать расстояние  $d$ , на котором погрешность определения фазы будет малой.

Приближение Фраунгофера справедливо, если фазовая ошибка, вызванная неучтенными членами разложения, удовлетворяет следующему выражению:

$$\frac{\pi}{\lambda d} (x^2 + y^2 + \xi^2 + \eta^2) \ll \frac{\pi}{N}. \quad (4.2.28)$$

Аналогично условие для выполнения приближения Френеля:

$$\frac{\pi}{\lambda d^3} \left( \frac{(x-\xi)^4}{4} + \frac{(y-\eta)^4}{4} \right) \ll \frac{\pi}{N}. \quad (4.2.29)$$

Допустим, что необходимо, чтобы фазовая ошибка была меньше, чем  $\pi/1000$ . Определим расстояние  $d$ , на котором будет выполняться это условие.

Для приближения Фраунгофера

$$d \gg \frac{N(\Delta x^2 + \Delta y^2 + \Delta \xi^2 + \Delta \eta^2)}{\lambda}. \quad (4.2.30)$$

Пусть максимальный размер  $N = 1000$ ,  $\Delta x = 4$  мкм,  $\Delta \xi = 5$  мкм,  $\lambda = 0,5$  мкм, тогда  $d \gg 164$  мм.

Проведем компьютерное моделирование.

Рассмотрим рассеяние сферической волны на некотором расстоянии:

$$Phase(i, j) =$$

$$= -i \frac{2\pi}{\lambda} \sqrt{\left( \frac{2\xi_{\max}}{N_x} \left( i - \frac{N_x}{2} \right) \right)^2 + \left( \frac{2\eta_{\max}}{N_y} \left( j - \frac{N_y}{2} \right) \right)^2} + d^2. \quad (4.2.31)$$



Пусть  $\Delta\xi = 5$  мкм (это характерный размер пикселя цифровой фотоматрицы),  $\lambda = 0,5$  мкм,  $N_x \times N_y = 1024 \times 1024$ , тогда на некотором расстоянии  $d$  кривизна профиля (рис. 4.15) будет равна:

$$\begin{aligned} d = 5 \text{ см} & \quad 0 - 0,0005\pi, \\ d = 10 \text{ см} & \quad 0 - 0,0002\pi, \\ d = 20 \text{ см} & \quad 0 - 0,00012\pi. \end{aligned}$$

Таким образом, на расстоянии 20 см фазовая ошибка меньше, чем  $\pi/1000$ .

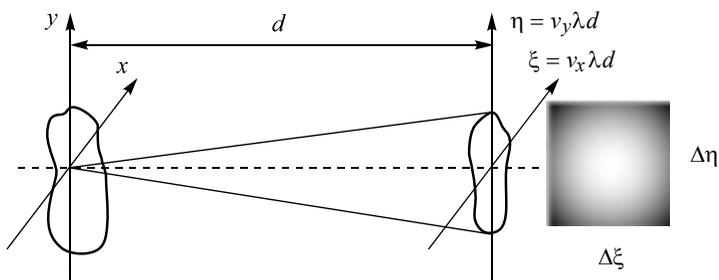


Рис. 4.15. Кривизна фронта при рассеивании сферического фронта на одном пикселе

Возьмем парааксиальное приближение

$$\begin{aligned} & \text{Phase}(i, j) = \\ & = -i \frac{2\pi}{\lambda} \left[ d + \frac{1}{2d} \left( \frac{2\xi_{\max}}{N_x} \left( i - \frac{N_x}{2} \right) \right)^2 + \frac{1}{2d} \left( \frac{2\eta_{\max}}{N_y} \left( j - \frac{N_y}{2} \right) \right)^2 \right]. \end{aligned} \quad (4.2.32)$$

Эти результаты показывают, что можно использовать дискретные преобразования Френеля и Фурье для определения амплитуды и фазы волнового фронта в случае, когда расстояние до объекта сравнимо с размерами объекта и голограммы.



### 4.3. КОМПЬЮТЕРНОЕ МОДЕЛИРОВАНИЕ ПЛОСКИХ И СФЕРИЧЕСКИХ ВОЛНОВЫХ ПОЛЕЙ

**Плоская волна.** Комплексная пространственная амплитуда плоской волны может быть представлена как

$$\vec{A}_r(x, y) = a_r \exp\left(i \frac{2\pi}{\lambda} (\sin \theta_x x + \sin \theta_y y)\right). \quad (4.3.1)$$

Если плоская волна падает перпендикулярно плоскости наблюдения, то

$$\vec{A}_r(x, y) = a_r + i0 = a_r. \quad (4.3.2)$$

Для компьютерного моделирования необходимо знать физический размер поля  $X_{\max} Y_{\max}$  (например, в мкм) и число точек  $N_x N_y$ . Тогда массив распределений комплексной амплитуды в плоскости  $(x, y)$  можно записать как

$$U_r(i, j) = a_r \exp\left(i \frac{2\pi}{\lambda} \left(\sin \theta_x \frac{X_{\max}}{N_x} i + \sin \theta_y \frac{Y_{\max}}{N_y} j\right)\right), \quad (4.3.3)$$

где  $i = 0 \dots N_x - 1$ ,  $j = 0 \dots N_y - 1$ .

Пример интерференционных полос, полученных в результате интерференции (сложения комплексных амплитуд) плоских волн, показан на рис. 4.16. Одна из волн перпендикулярна плоскости наблюдения, а другая имеет наклон  $\theta_x$  с осью  $X$ . На рисунке показана реальная часть комплексного сигнала.

Из рис. 4.16 видно, что с увеличением угла  $\theta_x$  число полос растет. Расстояние между пиками интерференционных полос можно найти из выражения

$$\Delta x = \frac{\lambda}{2 \sin(\theta_x / 2)}. \quad (4.3.4)$$

В этом случае число точек, которые приходятся на одну полосу, можно определить как

$$i_n = \frac{\Delta x N_x}{X_{\max}} = \frac{\lambda N_x}{2 X_{\max} \sin(\theta_x / 2)}. \quad (4.3.5)$$

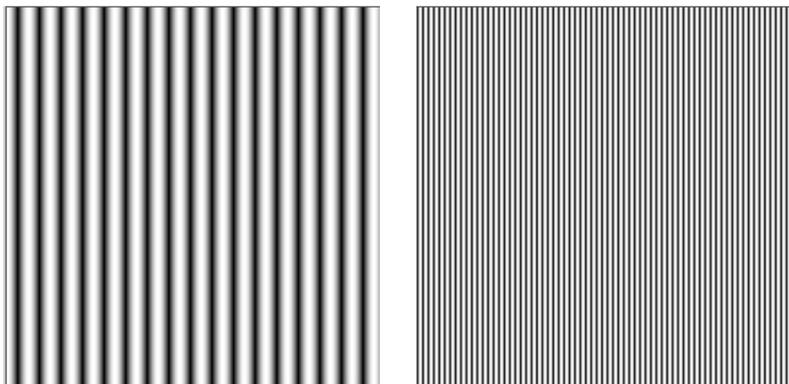


Рис. 4.16. Интерференция двух плоских волн

(слева угол между интерферирующими пучками  $\theta_x = 0,05^\circ$ ,  $\theta_y = 0^\circ$ ;  
справа  $\theta_x = 0,2^\circ$ ,  $\lambda = 0,5$  мкм,  $X_{\max} = 10$  мм,  $N_x N_y = 512 \times 512$  точек)

Для приведенных на рис. 4.16 примеров при угле  $\theta_x = 0,05^\circ$  число точек на одну полосу равно 29, а при  $\theta_x = 0,2^\circ$  равно 7. Для разрешения полос по теореме Найквиста необходимо, чтобы было больше двух точек на полосу, т. е. три и более. При приведенных на рис. 4.16 условиях угол не может превышать  $\theta_x = 0,5^\circ$ . При больших значениях угла необходимо увеличивать разрешение. Например, при  $N_x = 1024$  максимальный угол составит  $1,0^\circ$ .

При увеличении числа точек разрешение увеличивается. Для того чтобы можно было зафиксировать интерференцию плоских пучков при угле  $\theta_x = 30^\circ$  необходимо, чтобы число точек было больше, чем  $30\,000 \times 30\,000$ .

**Сферическая волна.** Сферическую опорную волну, распространяющуюся из центральной точки объектной плоскости на расстоянии  $d$  от плоскости наблюдения, можно записать как

$$\bar{A}(x, y) = \frac{1}{\sqrt{x^2 + y^2 + d^2}} \exp\left(-i \frac{2\pi}{\lambda} \sqrt{x^2 + y^2 + d^2}\right). \quad (4.3.6)$$



Массив распределений комплексной амплитуды в плоскости  $(x, y)$  можно записать как

$$U_r(i, j) = Amp(i, j) \exp(Pase(x, y)). \quad (4.3.7)$$

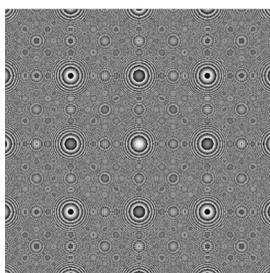
Центр сферы будет находиться в точке с координатами  $x_0 = 0, y_0 = 0$ . Если нужно, чтобы центр находился в центре изображения, необходимо центр сферы перенести в центр изображения:

$$Amp(i, j) = \left( \sqrt{\left( \frac{X_{\max}}{N_x} \left( i - \frac{N_x}{2} \right) \right)^2 + \left( \frac{Y_{\max}}{N_y} \left( j - \frac{N_y}{2} \right) \right)^2 + d^2} \right)^{-1}; \quad (4.3.8)$$

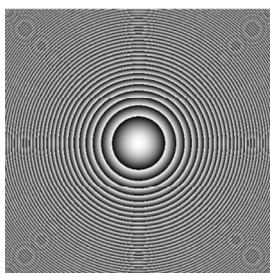
$$Phase(i, j) = -i \frac{2\pi}{\lambda} \sqrt{\left( \frac{X_{\max}}{N_x} \left( i - \frac{N_x}{2} \right) \right)^2 + \left( \frac{Y_{\max}}{N_y} \left( j - \frac{N_y}{2} \right) \right)^2 + d^2}. \quad (4.3.9)$$

Реальная часть сферической волны показана на рис. 4.17. Сферическая волна формировалась при следующих условиях:  $\lambda = 0,5$  мкм,  $X_{\max} = 10$  мм,  $N_x N_y = 512 \times 512$  точек. Менялось лишь расстояние до объекта  $d$ .

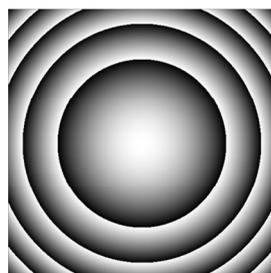
На рис. 4.17, *a* и *б* кривизна волнового фронта велика и видны артефакты, возникающие в результате недостаточной дискретизации.



*a* –  $d = 10$  см



*б* –  $d = 1$  м



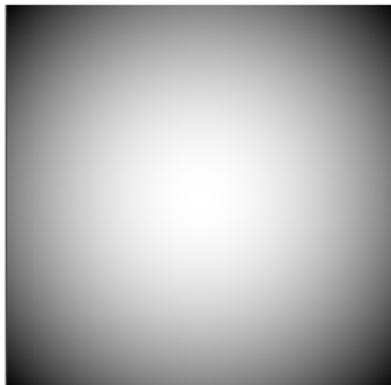
*в* –  $d = 10$  м

Рис. 4.17. Распространение сферической волны (показаны фазовые значения комплексной амплитуды)



При больших расстояниях и при заданных размерах объекта кривизна волнового фронта становится малой. Тот же эффект возникает при уменьшении размера поля. В этом случае сферический волновой фронт можно аппроксимировать плоским фронтом. Рассмотрим, при каких размерах поля это можно сделать.

Определим отклонение фазы от плоскости при разных размерах объекта на расстоянии  $d = 10$  см (рис. 4.18).



$$X_{\max} = 100 \text{ мкм} - \Delta\varphi = 0,3 \text{ rad},$$

$$X_{\max} = 20 \text{ мкм} - \Delta\varphi = 0,01 \text{ rad},$$

$$X_{\max} = 10 \text{ мкм} - \Delta\varphi = 0,003 \text{ rad},$$

$$X_{\max} = 5 \text{ мкм} - \Delta\varphi = 0,000785 \text{ rad}.$$

Рис. 4.18. Фаза сферической волны при указанных размерах поля

Допустим, что отклонение фазы сферической волны от плоскости не должно превышать  $\frac{2\pi}{1000} \approx 0,006 \text{ rad}$ . В этом случае при размерах поля менее 10 мкм это условие будет выполняться. Сферический фронт можно представить в виде набора плоских волн.

#### 4.4. КОМПЬЮТЕРНОЕ МОДЕЛИРОВАНИЕ ГОЛОГРАФИЧЕСКОГО ПРОЦЕССА В ОБЛАСТИ ФРАУНГОФЕРА

На рис. 4.19 показаны координатные плоскости, на которых нам необходимо исследовать процесс получения и восстановления голограмм.

Рассмотрим моделирование голографического восстановления амплитудных объектов [7].

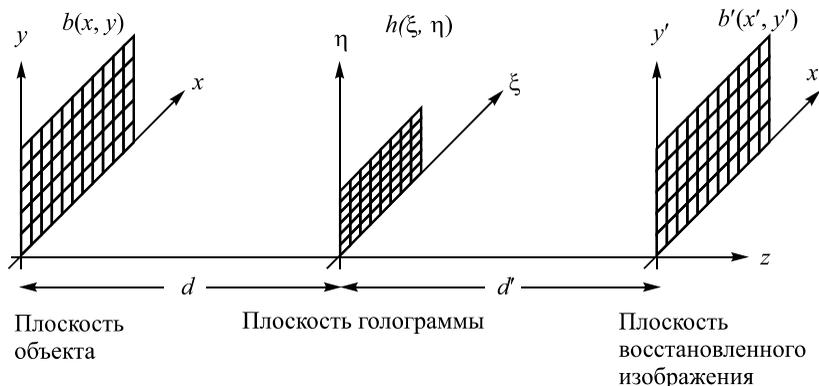


Рис. 4.19. Координатная система при получении голограмм и голографическом восстановлении

Для моделирования необходимо задать амплитуду и фазу волнового фронта. В плоскости объекта, зададим комплексный массив размером  $1024 \times 1024$  числа (рис. 4.20, касательной к. 4.20).



Рис. 4.20. Распределение комплексных амплитуд в плоскости, касательной к объекту:

$a$  – амплитуда;  $b$  – фаза

Амплитуда – произвольное распределение интенсивности. Диапазон от 0 до 255. Фаза – постоянное значение в диапазоне от  $-\pi$  до  $\pi$ . Такой комплексный массив моделирует амплитудный объект с зеркальным отражением.



Плоскость  $(\eta, \xi)$  – плоскость формирования голограммы. При реальном физическом процессе нулевые частоты должны быть в центре массива. При цифровом преобразовании Фурье происходит циклический сдвиг. Поэтому перед преобразованием необходимо сделать так, чтобы точки в центре массива располагались по краям изображения (рис. 4.21). В этом случае при преобразовании Фурье или Френеля нулевые частоты будут находиться в центре.

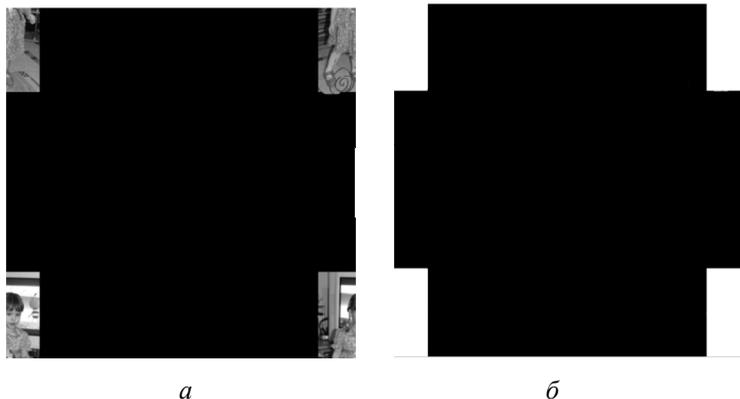


Рис. 4.21. Распределение комплексных амплитуд в плоскости, касательной к объекту, после циклического сдвига:

$a$  – амплитуда;  $b$  – фаза

Для пересчета в плоскость голограммы в зависимости от расстояния  $d$  необходимо сделать преобразование Френеля или Фурье (рис. 4.22). Размер объекта 7 мм, длина волны 0,5 мкм.

Для формирования голограммы необходимо добавить опорный волной фронт:

$$U_r(i, j) = a_r \exp \left( i \left[ \frac{2\pi}{\lambda} \left( \sin \theta_x \frac{X_{\max}}{N_x} i + \sin \theta_y \frac{Y_{\max}}{N_y} j \right) \right] \right), \quad (4.4.1)$$

где  $i = 0 \dots N_x - 1$ ,  $j = 0 \dots N_y - 1$ . В нашем случае  $N_x \times N_y = 1024 \times 1024$ ,  $X_{\max} Y_{\max} = 10 \text{ мм} \times 10 \text{ мм}$ ,  $a_r$  – постоянная по всему полю величина.

Амплитуда суммы двух волновых фронтов (фронта, рассеянного от объекта, и опорного пучка) пропорциональна интенсивности и может



рассматриваться как голограмма. Для обеспечения хорошего контраста при восстановлении изображений амплитуды опорного и объектного пучков должны быть выровнены.

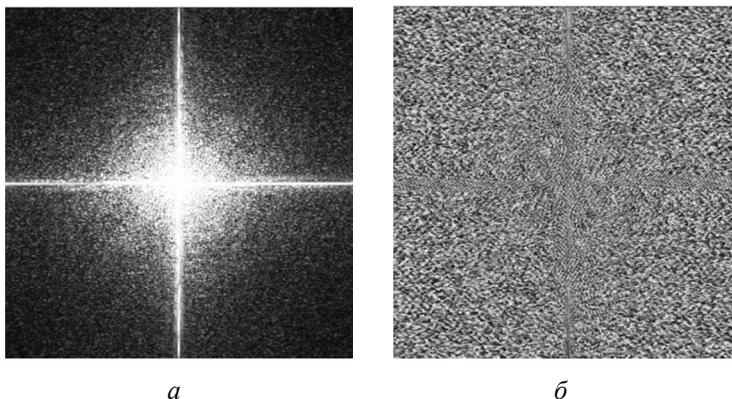


Рис. 4.22. Распределение амплитуд комплексного массива после преобразования Фурье и циклического сдвига:

$a$  – амплитуда;  $b$  – фаза

На следующем рисунке (рис. 4.23) приведены распределения фаз опорного пучка и голограмма в области Фраунгофера.

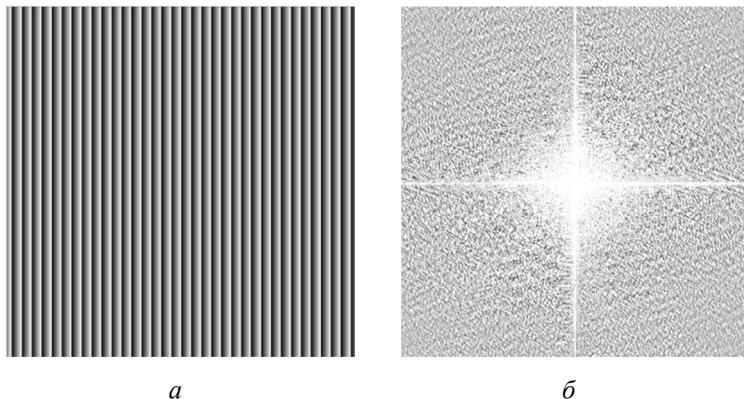


Рис. 4.23. Вид фразы опорного фронта и голограммы:

$a$  – распределение фаз опорного фронта ( $\theta_x = 0,3^0$ );  $b$  – голограмма



Для восстановления голограмму освещают световой волной, аналогичной той, что была использована при записи. Если использовать плоскую волну, которая падает перпендикулярно плоскости наблюдения, то

$$U_r(i, j) = a_r + i0 = a_r. \quad (4.4.2)$$

В нашем случае это умножение голограммы на какую-то константу.

Плоскость  $(x', y')$  – плоскость наблюдения. Для получения изображения объекта производится преобразование Френеля или Фурье над голограммой в зависимости от расстояния  $d'$ . В плоскости изображения мы можем определить как амплитуду, так и фазу исходного объекта.

При умножении голограммы на волну, перпендикулярную плоскости голограммы, нулевой пучок будет в центре, а мнимое и действительное изображения будут разъезжаться в зависимости от угла с опорной волной (рис. 4.24). На рисунках действительное изображение справа перевернутое.

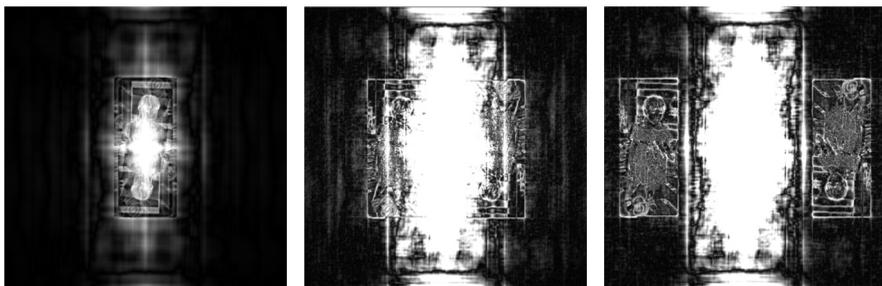


Рис. 4.24. Массив амплитуд в плоскости восстановления изображения в области Фраунгофера (слева направо: опорный фронт под углами  $\theta_x$  с нормалью к регистрирующей плоскости  $0^\circ, 0,7^\circ$  и  $1,4^\circ$ )

При умножении голограммы на волну, сопряженную с опорной при восстановлении, в центре будет действительное изображение. На рис. 4.25 показаны восстановленные в области Фраунгофера амплитуда и фаза.

Из приведенных рисунков видно следующее:

- 1) амплитуда объекта искажается из-за влияния центрального пятна;
- 2) информацию о фазе восстановить практически не удастся.

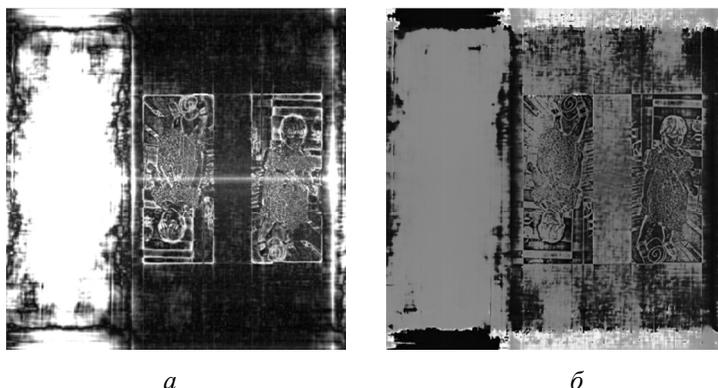


Рис. 4.25. Амплитуда и фаза восстановленного изображения:

*a* – массив амплитуд и фаз; *б* – массив в плоскости восстановления изображения (опорный фронт под углом  $\theta_x = 1,4^\circ$ )

Для повышения качества голограммы необходимо ослабить или полностью устранить влияние центрального пучка.

Рассмотрим основные способы устранения влияния центрального пучка.

#### 4.5. УСТРАНЕНИЕ НУЛЕВОГО ПОРЯДКА ДИФРАКЦИИ ПРИ ДИФФУЗНОМ РАССЕЙВАНИИ ПОВЕРХНОСТИ ОБЪЕКТОВ

Влияние центрального пятна можно сильно ослабить, если объект имеет случайную фазу. Пусть объект имеет вид, представленный на рис. 4.26. Фаза – случайное распределение в диапазоне от  $-\pi$  до  $\pi$ .

Этот случай соответствует диффузному распределению поверхности объекта. Для повышения качества голографии объект покрывают специальными красителями, рассеивающими световую волну в разных направлениях. При диффузном отражении от объекта центральный пучок размывается по всему полю. Амплитуда восстановленного объекта имеет диффузную структуру, но качество восстановления лучше, чем для зеркальных объектов.

В качестве опорного фронта можно выбрать плоскую волну под углом  $\theta_x = 1,4^\circ$  с нормалью к регистрирующей плоскости. В плоскости изображения в зоне Фраунгофера получим следующее распределение (рис. 4.27).

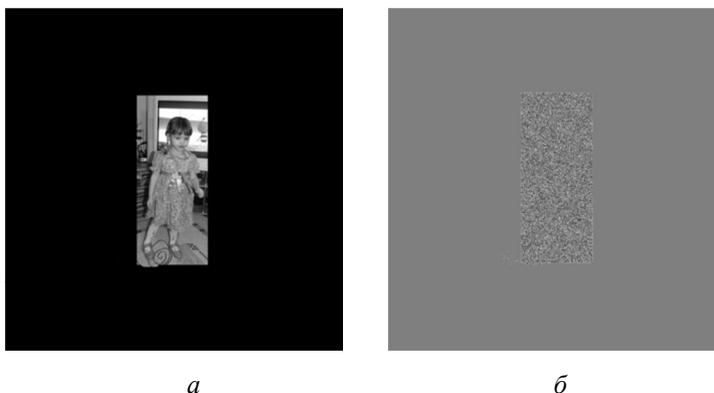


Рис. 4.26. Распределение комплексной амплитуды в плоскости, касательной к объекту:

$a$  – амплитуда;  $b$  – фаза

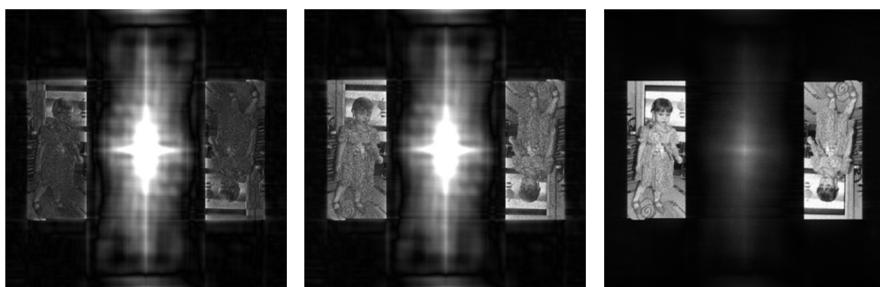


Рис. 4.27. Распределение амплитуд комплексного массива после преобразования (слева направо: исходная фаза, случайные значения с распределением 20, 40 и 80 % от  $-\pi$  до  $\pi$  соответственно)

Видно, что качество восстановленного изображения значительно ухудшается.

То же самое происходит, если мы получаем голограмму в области Френеля (рис. 4.28). Из рисунка видно, что мнимое изображение расплывается и только в области Фраунгофера становится резким.

Рассмотрим, как происходит восстановление фазовых объектов при небольших углах сдвига. Зададим объект с комплексной амплитудой, показанной на рис. 4.29.

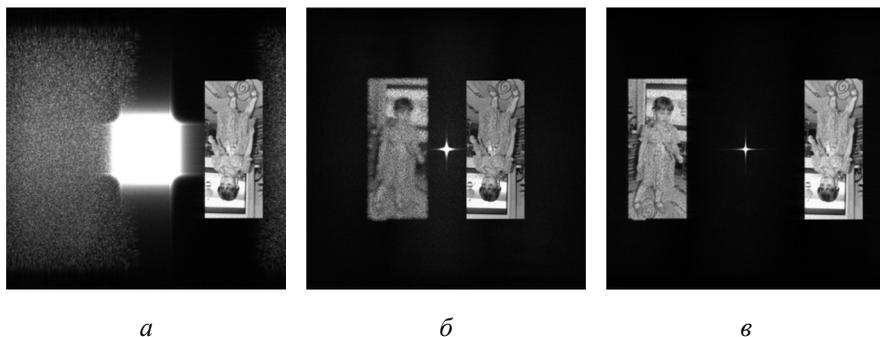


Рис. 4.28. Массив амплитуд в плоскости восстановления изображения:

$$a - d = 60 \text{ см}; \quad б - d = 10 \text{ м}; \quad в - d = 100 \text{ м}$$

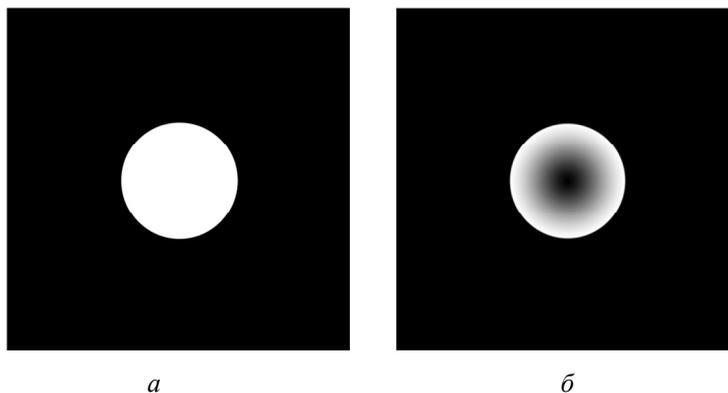


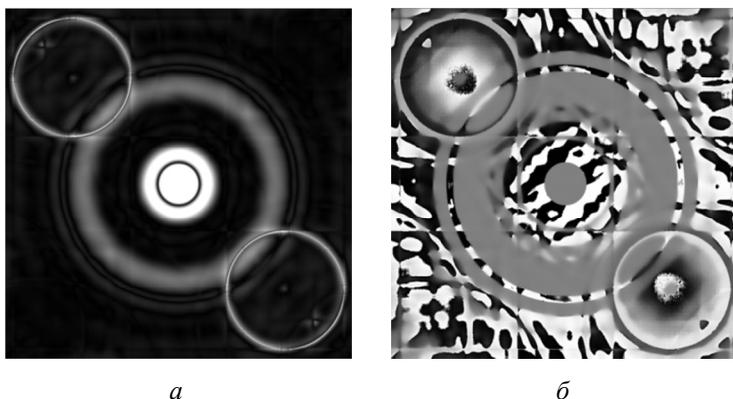
Рис. 4.29. Распределение комплексных амплитуд в плоскости, касательной к объекту:

$a$  – амплитуда;  $б$  – фаза. Фаза объекта распределена в диапазоне от  $-\pi$  до  $\pi$

Для формирования голограммы добавим опорный волновой фронт под углами  $\theta_x = 0,9^\circ$ ,  $\theta_y = 0,9^\circ$  с нормалью к регистрирующей плоскости. В результате получим изображение, показанное на рис. 4.30.

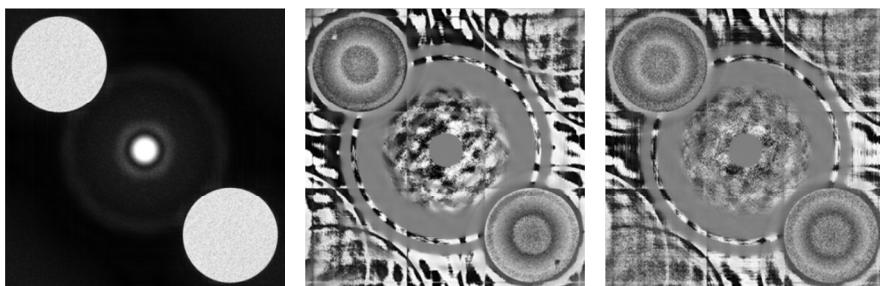
В правом нижнем углу на рис. 4.30 действительное изображение, в левом верхнем – мнимое.

Пусть объект имеет диффузную поверхность. Тогда при восстановлении мы получим изображение, представленное на рис. 4.31.



*Рис. 4.30.* Распределение амплитуд комплексного массива после восстановления из голограммы:

*a* – амплитуда; *б* – фаза



В классической голографии для того, чтобы придать объекту случайный характер отражения от поверхности при голографических исследованиях, объект покрывается специальной диффузно-рассеивающей краской. Это приводит к значительному улучшению качества восстановленного изображения.

К сожалению, и в этом случае фазовая информация сильно искажается при восстановлении.



#### 4.6. МЕТОДЫ ПОДАВЛЕНИЯ НУЛЕВОГО ПОРЯДКА ДИФРАКЦИИ НА ОСНОВЕ ЦИФРОВОЙ ФИЛЬТРАЦИИ ГОЛОГРАММЫ

При малых углах волновые фронты, находящиеся в разных дифракционных порядках, частично или полностью перекрываются.

Если рассматривать яркость частей изображения в разных дифракционных порядках, то яркость постоянной составляющей будет на порядок выше остальных. Следовательно, при цифровом восстановлении в процессе нормализации значений большая часть диапазона будет приходиться на постоянную составляющую, что не обеспечит необходимого контраста для выделения остальных составляющих.

Интенсивность постоянной составляющей в спектральном диапазоне согласно [8, 9] определяется следующим соотношением:

$$Z = \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} U_h(k\Delta x, l\Delta y), \quad (4.6.1)$$

где  $N$  – размер голограммы в пикселях. Для того чтобы удалить постоянную составляющую, можно отфильтровать голограмму фильтром частот, эквивалентным средней интенсивности голограммы:

$$U'_h(k\Delta x, l\Delta y) = U_h(k\Delta x, l\Delta y) - \frac{1}{N^2} \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} U_h(k\Delta x, l\Delta y). \quad (4.6.2)$$

Единственным достоинством этого метода является его вычислительная простота. Попробуем его улучшить. Фильтр (4.6.2) вычисляет среднюю интенсивность всей голограммы и вычитает ее из каждой точки голограммы перед восстановлением. Основная сложность применения данного метода заключается в том, что интенсивность постоянной составляющей может на порядки превышать интенсивности остальных частей восстановленного изображения. Это значит, что при вычитании могут образовываться отрицательные значения интенсивностей в областях низких частот. Поэтому после фильтрации неизбежно появляются низкочастотные остатки.

Необходим алгоритм, который вычисляет величину фильтра адаптивно, в зависимости от места в голограмме, где он приложен. В основе нового метода лежит фрагментирование голограммы и взвешивание полученных фрагментов.



Возьмем на голограмме область величиной  $M \times M$  пикселей с центром в пикселе, для которого вычисляется фильтр. Взвесим каждый пиксель внутри области коэффициентом  $W(k, l)$ , где  $(k, l)$  – координаты пикселя внутри области. Характер функции взвешивания определим экспериментально, исходя из параметров измерительной системы, величин объектов исследования и размеров голограммы. На рис. 4.32 представлен пример подобного адаптивного фильтра.

	0,5	0,5	0,5	0,5	0,5		
	0,5	1,0	1,0	1,0	0,5		
	0,5	1,0	1,0	1,0	0,5		
	0,5	1,0	1,0	1,0	0,5		
	0,5	0,5	0,5	0,5	0,5		

Рис. 4.32. Пример адаптивного фильтра

Закрашенный пиксель в центре обозначает пиксель, для которого вычисляется фильтр, числа – веса каждого из пикселей. За пределами рамки фильтра считаем, что веса нулевые. Если размер фильтра  $M$  совпадает с размером голограммы  $N$ , а взвешивание осуществляется равномерно ( $W_{k,l} = 1$ ), то фильтр вырождается в представленный ранее частотный фильтр (4.6.2).

Представим описанный фильтр в математическом виде. Для этого введем величину  $w(k, l)$ , обозначающую нормализованный в области фильтра вес пикселя. Тогда адаптивный частотный фильтр для голограммы может быть вычислен следующим образом [9]:

$$U'_h(m, n) = U_h(m, n) - \frac{\sum_{k=m-\frac{M}{2}}^{m+\frac{M}{2}} \sum_{l=n-\frac{M}{2}}^{n+\frac{M}{2}} U_h(k, l) w(k, l)}{M^2}, \quad (4.6.3)$$

где  $(m, n)$  – координаты пикселя в голограмме.



Выбор функции взвешивания осуществляется экспериментально. При восстановлении искусственных голограмм, рассматриваемых в этом разделе, используется фильтр  $3 \times 3$  пикселя с равномерным взвешиванием:

$$U'_h(m, n) = U_h(m, n) - \frac{1}{9} [U_h(m-1, n-1) + U_h(m-1, n) + U_h(m-1, n+1) + U_h(m, n-1) + U_h(m, n) + U_h(m, n+1) + U_h(m+1, n-1) + U_h(m+1, n) + U_h(m+1, n+1)]. \quad (4.6.4)$$

Хорошо зарекомендовал себя также фильтр с весовыми коэффициентами:

$$U'_h(m, n) = U_h(m, n) - \frac{1}{16} [U_h(m-1, n-1) + 2U_h(m-1, n) + U_h(m-1, n+1) + 2U_h(m, n-1) + 4U_h(m, n) + 2U_h(m, n+1) + U_h(m+1, n-1) + 2U_h(m+1, n) + U_h(m+1, n+1)]. \quad (4.6.5)$$

Результаты подавления постоянной составляющей при помощи адаптивного фильтра (4.6.5) приведены на рис. 4.33. Угол между интерферирующими пучками  $\theta_x = 0,7^\circ$ .

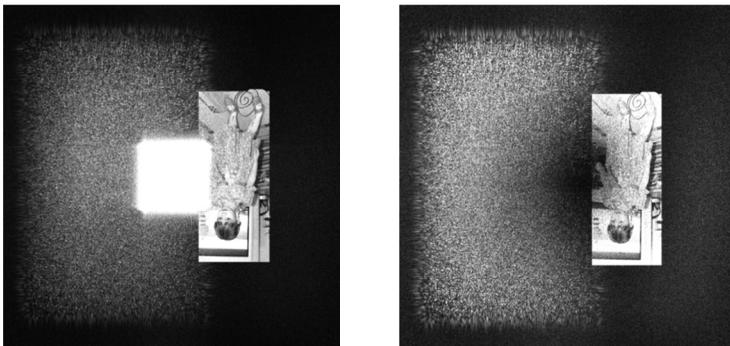


Рис. 4.33. Подавление постоянной составляющей при помощи адаптивного частотного фильтра при восстановлении голограммы



При правильном подборе параметров адаптивного фильтра постоянная составляющая подавляется достаточно хорошо. Методы, основанные на фильтрации, дают хорошие результаты при восстановлении амплитуды, но при этом фазовая информация искажается достаточно сильно.

#### 4.7. МЕТОДЫ ПОДАВЛЕНИЯ НУЛЕВОГО ПОРЯДКА ДИФРАКЦИИ НА ОСНОВЕ ОБРАЗОВАНИЯ РЕПЛИК СИГНАЛА

Поскольку для восстановления действительного и мнимого изображений необходимо произвести преобразование Фурье, то можно воспользоваться алгоритмом образования реплик спектра [14].

Рассмотрим, что произойдет со спектром объекта, если его значения разрядить нулями. Пусть объект представляет собой синусоидальную решетку (рис. 4.34). Тогда в частотной области амплитуда спектра сигнала будет представлять собой три пика – центральный пик на 256-м отсчете и два пика, соответствующих числу полос. Чем больше частота, тем больше пики будут разъезжаться от центрального пика.

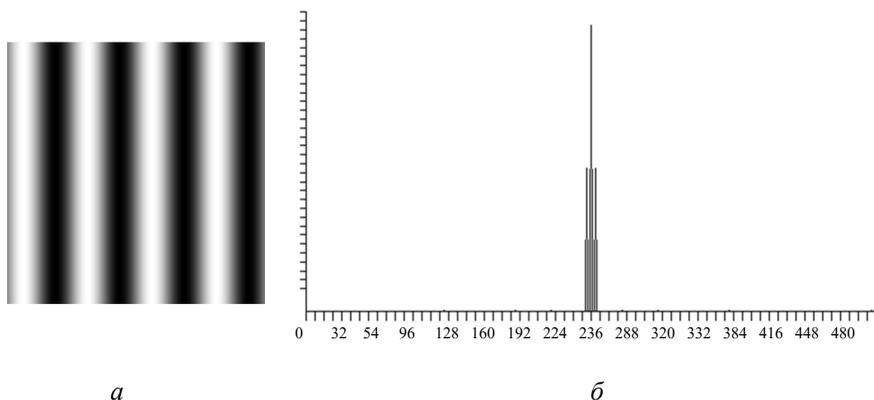


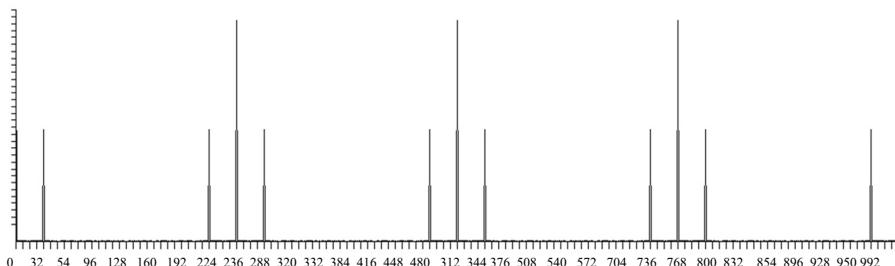
Рис. 4.34. Спектр синусоидального сигнала:

*a* – синусоидальные полосы (размер объекта  $512 \times 512$  пикселей, четыре полосы, диапазон яркости от 0 до 256); *б* – спектр по строке

Если изображение увеличить, добавляя нули после каждой точки, то спектр будет повторяться с некоторым периодом. На рис. 4.35 показан спектр сигнала с добавлением нуля после каждой точки. Из рисунка

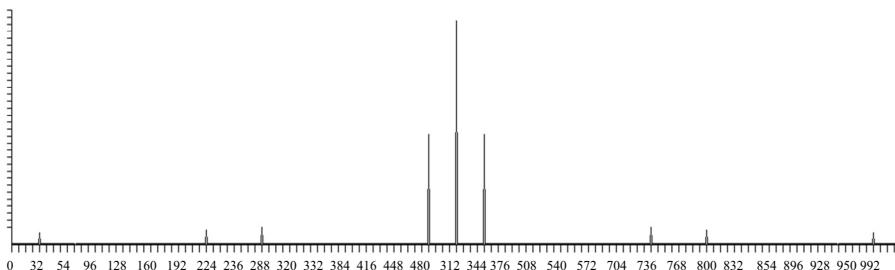


ка видно, что в этом случае образуются реплики спектра. Период повторения реплик соответствует размеру исходного объекта.



*Рис. 4.35.* Спектр сигнала с частотой 32 полосы, увеличенный тремя нулями. Размер исходного объекта  $256 \times 256$  пикселей, результирующего –  $1024 \times 256$  пикселей

Если вместо добавлений между исходными значениями нулей мы просто повторим исходные значения (рис. 4.36), то можно подавить влияние центрального пика в репликах спектра. Из рисунка видно, что спектр повторяется, но центральный пик отсутствует в репликах спектра.



*Рис. 4.36.* Спектр сигналов с частотой 32 полосы, увеличенный тремя повторениями исходной величины. Размер исходного объекта  $256 \times 256$  пикселей, результирующего –  $1024 \times 256$  пикселей

На рис. 4.37 показана реальная голограмма, оцифрованная с размером  $1728 \times 1252$  пикселя.

Видно, что в этом случае нулевой пучок перекрывает действительное и мнимое изображения.

На рис. 4.38 показана голограмма размером  $6912 \times 1252$  пикселя, полученная простым повторением значений по  $x$  координате четыре раза.

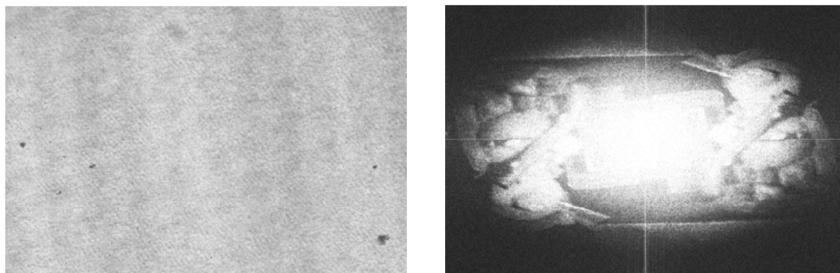
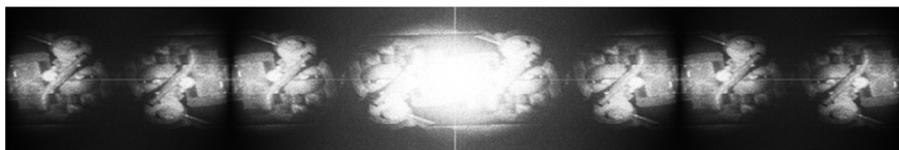


Рис. 4.37. Восстановление изображения из голограммы:

$a$  – голограмма;  $b$  – восстановленное из нее действительное и мнимое изображения



$a$



$b$

Рис. 4.38. Голограмма, полученная четырехкратным повторением значений ( $a$ ), и восстановленные из нее действительные и мнимые изображения ( $b$ )

Видно, что в репликах спектра влияние центрального пучка ослаблено. Взяв вторую или третью реплики, можно получить неискаженное нулевым пучком действительное и мнимое изображения.

В отличие от использования частотных фильтров, искажения спектра уменьшаются. Однако увеличение размера голограммы приводит к увеличению времени расшифровки.

В этой главе проведено компьютерное моделирование получения классических тонких голограмм, т. е. получение картин интенсивностей от интерференции объектного и опорного пучков и восстановление из них амплитудных и фазовых значений волновых полей. Из-за недостаточного разрешения матричных устройств фотодетекторов это



возможно только при небольших углах между интерферирующими пучками. Однако при этом влияние центрального пучка искажает фазовую информацию, поэтому возможна только качественная оценка параметров объекта. Получить точную количественную информацию при восстановлении голограмм при малых углах возможно только при использовании методов расшифровки на основе получения математических голограмм.

#### БИБЛИОГРАФИЧЕСКИЙ СПИСОК К ГЛАВЕ 4

1. *Ярославский Л. П.* Цифровая голография / Л. П. Ярославский, Н. С. Мерзляков. – Москва: Наука, 1982. – 219 с.
2. *Рабинер Л.* Теория и применение цифровой обработки сигналов / Л. Рабинер, Б. Гоулд. – Москва: Мир, 1978. – 848 с.
3. *Cooley J. M.* An algorithm for the machine calculation of complex Fourier series / J. M. Cooley, J. W. Takey // *Mathematics of Computation*. – 1965. – Vol. 19. – P. 297–301.
4. Быстрые алгоритмы в цифровой обработке изображений: пер с англ. / Т. С. Хуанг, Дж.-О. Эклунд, Г. Дж. Нуссбаумер [и др.]; под ред. Т. С. Хуанга. – Москва: Радио и связь, 1984. – 224 с.
5. *Ярославский Л. П.* Цифровая обработка сигналов в оптике и голографии: введение в цифровую оптику / Л. П. Ярославский. – Москва : Радио и связь, 1987. – 296 с.
6. *Домненко В. М.* Моделирование формирования оптического изображения: учебное пособие / В. М. Домненко, М. В. Бурсов, Т. В. Иванова. – Санкт-Петербург: НИУ ИТМО, 2011. – 141 с.
7. *Гужов В. И.* Цифровое моделирование оптических систем для фазометрии / В. И. Гужов, В. Ф. Ким, В. А. Родионов // *Системы автоматизации обработки оптической информации*. – Новосибирск, 1984. – С. 112–120.
8. *Kim M. K.* Principles and techniques of digital holographic microscopy / M. K. Kim // *SPIE Reviews*. – 2010. – Vol. 1. – P. 018005-1–018005-50.
9. *Кузнецов Р. А.* Разработка системы неразрушающего контроля на основе методов цифровой голографической интерферометрии: дис. ... канд. техн. наук / Р. А. Кузнецов. – Новосибирск, 2013. – 117 с.
10. *Войнаровский М.* Программирование. Быстрое преобразование Фурье [Электронный ресурс] / М. Войнаровский. – 2002. – Режим доступа: <http://bourabai.ru/graphics/fft.htm> (дата обращения: 13.06.2018).



11. *Verrier N.* Off-axis digital hologram reconstruction: some practical considerations / N. Verrier, M. Atlan // *Applied Optics*. – 2011. – Vol. 50, N 34. – P. H136–H146.

12. *Гужов В. И.* Представление преобразования Френеля в дискретном виде / В. И. Гужов, Р. Б. Несин, В. А. Емельянов // *Автоматика и программная инженерия*. – 2016. – № 1 (15). – С. 91–96.

13. *Блейхут Р.* Быстрые алгоритмы цифровой обработки сигналов / Р. Блейхут. – Москва: Мир, 1989. – 448 с.

14. *Гужов В. И.* Новый способ устранения влияния нулевого пучка при цифровом восстановлении изображений из голограмм / В. И. Гужов, Е. Н. Денежкин, Н. С. Зарубин // *Автоматика и программная инженерия*. – 2018. – № 1 (2). – С. 80–88.

---

---

## ГЛАВА 5

# ВОССТАНОВЛЕНИЕ АМПЛИТУДЫ И ФАЗЫ ОБЪЕКТА ПО МАТЕМАТИЧЕСКОЙ ГОЛОГРАММЕ

---

---

Классические методы голографии основаны на том, что мы можем регистрировать только интенсивность голограммы.

В отличие от восстановления классических голограмм, при использовании методов компьютерной голографии можно получить математическую голограмму, которая состоит из амплитуды и фазы объектного поля:

$$G(x, y) = a_p(x, y) \exp(\varphi_p(x, y)), \quad (5.1)$$

где  $a_p(x, y)$  – амплитуда поля;  $\varphi_p(x, y)$  – фаза поля, рассеянного от объекта в плоскости голограммы  $(\eta, \xi)$ . Если удастся найти математическую голограмму  $G(x, y)$ , то в плоскости изображения можно восстановить комплексную амплитуду поля, рассеянного от объекта.

Математическая голограмма, в отличие от обычной голограммы, это не картина интенсивностей, а комплексная амплитуда.

Найти значения амплитуды и фазы в (5.1) можно по нескольким голограммам с помощью методов пошагового фазового сдвига (Phase Sampling, Phase Shifting Interferometry – PSI) [1–4].

### 5.1. ОПРЕДЕЛЕНИЕ РАЗНОСТИ ФАЗ МЕТОДОМ ПОШАГОВОГО ФАЗОВОГО СДВИГА (PSI)

Метод пошагового фазового сдвига основан на регистрации нескольких интерференционных картин при управляемом изменении фазы опорной волны на известную величину.



Фазовый сдвиг между интерферирующими пучками может быть реализован различными способами. На рис. 5.1 показана принципиальная схема оптической установки, в которой фазовый сдвиг задается перемещением зеркала, закрепленного на пьезокерамике. Пучок света от когерентного источника излучения, попадая на делительный куб, разделяется на опорный и объектный пучки. В опорном плече находится зеркало, предназначенное для внесения эталонных фазовых сдвигов. Конфигурация объектного плеча зависит от задачи измерения и от формы поверхности тестируемого объекта. В выходной плоскости помещен массив детекторов для регистрации интенсивности в каждой точке поля. После каждого фазового сдвига информация о поле интенсивностей  $I(x, y)$  вводится в компьютер.

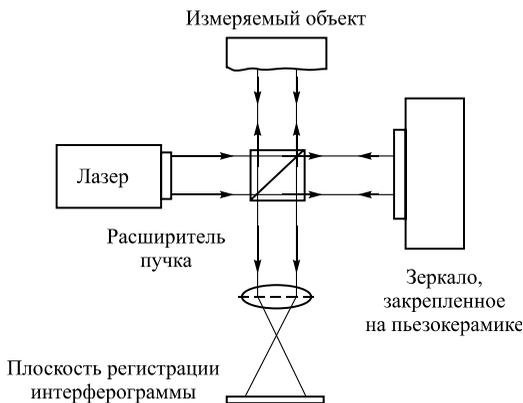


Рис. 5.1. Схема интерферометра Тваймана–Грина с зеркалом, закрепленным на пьезо-керамике, в опорном плече

В зависимости от числа фазовых сдвигов существуют различные алгоритмы расшифровки. В [5] показано, как получить вычислительные процедуры, которые с помощью комбинаций интенсивности интерференционных картин получают распределение фаз для произвольных углов сдвига. В [6, 7] получена обобщенная схема алгоритма для различного числа сдвигов.

Результат интерференции двух волновых фронтов можно записать в виде (2.4.13). Это выражение имеет вид

$$I(x, y) = I_0(x, y) (1 + V(x, y) \cos(\varphi_p(x, y) - \varphi_r(x, y))). \quad (5.1.1)$$



В этом выражении три неизвестных:  $I_0(x, y)$  – средняя интенсивность,  $V(x, y)$  – контраст,  $\varphi(x, y) = \varphi_p(x, y) - \varphi_r(x, y)$  – фазовая разность между интерферирующими фронтами. Нам необходимо найти разность фаз  $\varphi(x, y)$ .

Для идеальных картин, в которых средняя интенсивность и контраст постоянны для всех точек поля, это несложно сделать, взяв  $\arccos()$  от найденных значений интенсивности. Однако в общем случае этого сделать нельзя, так как  $I_0(x, y)$  и  $V(x, y)$  меняются в зависимости от  $(x, y)$ . Взаимная независимость достигается, например, если в каждое из этих уравнений добавить постоянный фазовый сдвиг  $\delta_i$ . В таком случае можно определить значения разности фаз  $\varphi(x, y)$  из системы уравнений

$$I_i(x, y) = I_0(x, y)[1 + V(x, y)\cos(\varphi(x, y) + \delta_i)], \quad (5.1.2)$$

здесь  $i = 0, 1, \dots, m-1$ , где  $m$  – число фазовых сдвигов, большее трех или равное трем.

Для нахождения разности фаз необходимо найти решение системы трансцендентных уравнений (5.1.2). Для однозначного определения разности фаз необходимо не менее трех взаимно независимых уравнений.

Рассмотрим метод решения системы (5.1.2) в общем виде.

Представим выражение (5.1.2) в векторной форме (учитывая, что  $\cos(\alpha + \beta) = \cos\alpha\cos\beta - \sin\alpha\sin\beta$ ):

$$\vec{I} = I_0\vec{R} + (I_0V\cos\varphi)\vec{C} - (I_0V\sin\varphi)\vec{S}, \quad (5.1.3)$$

где  $\vec{I} = (I_0, \dots, I_{m-1})^T$  – набор измеренных интенсивностей с различными фазовыми сдвигами  $\delta_i$  в каждой точке интерферограммы  $I_i(x, y)$ . Пусть  $\vec{C} = (\cos\delta_0, \dots, \cos\delta_{m-1})^T$ ,  $\vec{S} = (\sin\delta_0, \dots, \sin\delta_{m-1})^T$  – набор синусов и косинусов от известных фазовых сдвигов,  $\vec{R} = (1, \dots, 1)^T$  – единичный вектор. Размерность векторов определяется числом фазовых сдвигов  $m$ .

Для выделения  $\sin\varphi$  и  $\cos\varphi$  воспользуемся свойством скалярного произведения ортогональных векторов  $(aa^\perp) = 0$ . Умножая правую и



левую часть выражения (5.1.2) на вектор  $\vec{C}^\perp$  (вектор, ортогональный вектору  $\vec{C}$ ) и вектор  $\vec{S}^\perp$  (вектор, ортогональный вектору  $\vec{S}$ ), получим

$$\vec{I}\vec{C}^\perp = I_0\vec{R}\vec{C}^\perp + (I_0V \cos \varphi) \vec{C}\vec{C}^\perp - (I_0V \sin \varphi) \vec{S}\vec{C}^\perp; \quad (5.1.4)$$

$$\vec{I}\vec{S}^\perp = I_0\vec{R}\vec{S}^\perp + (I_0V \cos \delta\varphi) \vec{C}\vec{S}^\perp - (I_0V \sin \delta\varphi) \vec{S}\vec{S}^\perp. \quad (5.1.5)$$

В силу свойства скалярного произведения ортогональных векторов получим

$$\vec{I}\vec{C}^\perp = I_0\vec{R}\vec{C}^\perp - (I_0V \sin \varphi) \vec{S}\vec{C}^\perp; \quad (5.1.6)$$

$$\vec{I}\vec{S}^\perp = I_0\vec{R}\vec{S}^\perp + (I_0V \cos \varphi) \vec{C}\vec{S}^\perp. \quad (5.1.7)$$

Ортогональный вектор  $\vec{a}^\perp$  можно найти с помощью матричного уравнения

$$\vec{a}^\perp = M\vec{a}. \quad (5.1.8)$$

Матрица преобразования  $M$  должна удовлетворять следующему требованию:

$$M[1 \ 1 \ \dots \ 1 \ 1]^T = 0, \quad (5.1.9)$$

Учитывая (5.1.9), можно увидеть, что скалярные произведения векторов  $\vec{R}\vec{C}^\perp = 0$  и  $\vec{R}\vec{S}^\perp = 0$ , поэтому выражения (5.1.6) и (5.1.7) можно записать в виде

$$I_0V \sin \varphi = -\frac{\vec{I}\vec{C}^\perp}{\vec{S}\vec{C}^\perp}; \quad (5.1.10)$$

$$I_0V \cos \varphi = \frac{\vec{I}\vec{S}^\perp}{\vec{C}\vec{S}^\perp}. \quad (5.1.11)$$

Учитывая свойство скалярного произведения  $(ab^\perp) = -(a^\perp b)$ , получим  $(\vec{C}\vec{S}^\perp) = -(\vec{S}\vec{C}^\perp)$ . Тогда алгоритм определения разности фаз в векторной форме примет следующий вид:



$$\frac{\sin \varphi}{\cos \varphi} = \frac{\vec{I}\vec{C}^\perp}{\vec{I}\vec{S}^\perp}, \quad (5.1.12)$$

$$\varphi = \arctan \left( \frac{\vec{I}\vec{C}^\perp}{\vec{I}\vec{S}^\perp} \right). \quad (5.1.13)$$

Приведем выражение (5.1.13) к более часто используемому виду, умножая числитель и знаменатель (5.1.10) и (5.1.11) на левую и правую части уравнения  $(\vec{C}\vec{S}^\perp) = -(\vec{C}^\perp\vec{S})$ :

$$I_0 V \sin \varphi = -\frac{\vec{I}\vec{C}^\perp}{\vec{S}\vec{C}^\perp} = -\frac{\vec{I}^\perp\vec{C}}{\vec{S}^\perp\vec{C}}; \quad (5.1.14)$$

$$I_0 V \cos \varphi = \frac{\vec{I}\vec{S}^\perp}{\vec{C}\vec{S}^\perp}, \quad (5.1.15)$$

откуда

$$\varphi = \arctan \left( \frac{\vec{I}^\perp\vec{C}}{\vec{I}^\perp\vec{S}} \right). \quad (5.1.16)$$

Выражение (5.1.16) проще вычислять, чем (5.1.13), так как в этом случае требуется вычислить лишь один ортогональный вектор  $\vec{I}^\perp$ , который можно найти с помощью матричного уравнения

$$\vec{I}^\perp = M\vec{I}. \quad (5.1.17)$$

При  $m$  фазовых сдвигов матрица  $M$  будет иметь вид

$$M = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & -1 \\ -1 & 0 & 1 & \cdots & 0 & 0 \\ 0 & -1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 1 & 0 & 0 & \cdots & -1 & 0 \end{bmatrix}_{m \times m}. \quad (5.1.18)$$



В приведенных формулах расшифровки разность фаз определяется с помощью функции  $\arctan()$ . Область определения функции  $\arctan()$  определена от  $-\pi/2$  до  $\pi/2$  (рис. 5.2).

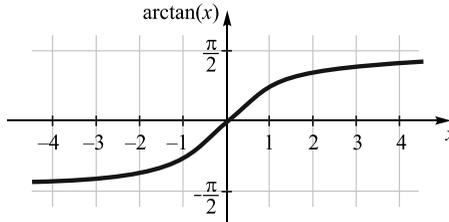


Рис. 5.2. График функции  $y = \arctan(x)$

Анализируя знаки числителя и знаменателя выражения (5.1.1), можно расширить область определения. Видно, что знак числителя совпадает со знаком синуса, а знак знаменателя – со знаком косинуса. Это относится ко всем разрешающим выражениям этого раздела.

В современных языках программирования вводится специальная функция двух аргументов  $\text{atan2}(y, x)$ . Эта функция имеет область допустимых значений  $(-\pi, \pi]$ , которая добавлением  $2\pi$  к отрицательным значениям может быть отображена на область  $[0, 2\pi)$ . Эта функция определяется как

$$\text{atan2}(y, x) = \begin{cases} \arctan\left(\frac{y}{x}\right) & \text{if } x > 0 \\ \arctan\left(\frac{y}{x}\right) + \pi & \text{if } x < 0 \text{ and } y \geq 0 \\ \arctan\left(\frac{y}{x}\right) - \pi & \text{if } x < 0 \text{ and } y < 0 \\ \frac{\pi}{2} & \text{if } x = 0 \text{ and } y > 0 \\ -\frac{\pi}{2} & \text{if } x = 0 \text{ and } y < 0 \\ \text{не определено} & \text{if } x = 0 \text{ and } y = 0 \end{cases} \quad (5.1.19)$$



В литературе встречаются различные формы записи разрешающих уравнений для нахождения фазовых значений:

$$\varphi = \arctan \left( \frac{\vec{I}^\perp \vec{C}}{\vec{I}^\perp \vec{S}} \right) = \text{atan } 2(x, y) \quad \text{или} \quad \varphi = \arctan \left( \frac{\vec{I}^\perp \vec{S}}{\vec{I}^\perp \vec{C}} \right) = \text{atan } 2(y, x),$$

т. е. числитель и знаменатель переставлены местами.

Посмотрим, к чему приводит использование различных видов уравнения. В зависимости от числа фазовых сдвигов значения  $x = \vec{I}^\perp \vec{C}$  и  $y = \vec{I}^\perp \vec{S}$  меняются от  $-X_{\max}$  до  $X_{\max}$ . Если отложить по оси  $x$  и по оси  $y$  значения от  $-X_{\max}$  до  $X_{\max}$ , то величина  $\text{atan } 2(x, y)$  будет меняться от  $(-\pi, \pi]$  (рис. 5.3).

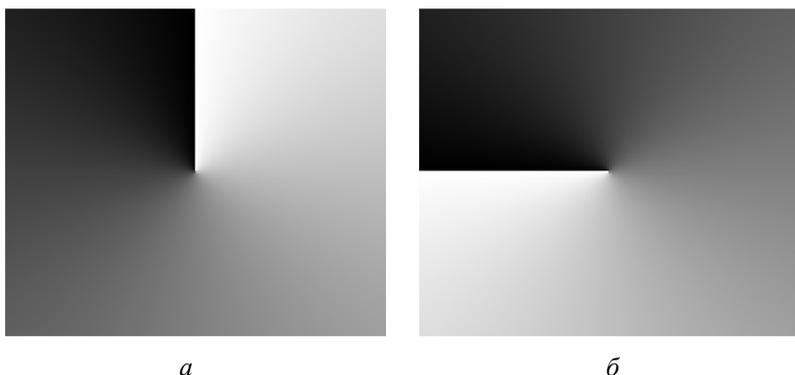


Рис. 5.3. Функция  $\text{atan } 2(x, y)$  (а) и  $\text{atan } 2(y, x)$  (б)

Из рисунка видно, что

$$\text{atan } 2(x, y) = -\text{atan } 2(y, x) - \frac{3\pi}{2}, \quad \text{если} \quad \left( -\text{atan } 2(y, x) - \frac{3\pi}{2} \right) > -\pi, \quad (5.1.20)$$

и

$$\begin{aligned} \text{atan } 2(x, y) &= -\text{atan } 2(y, x) - \frac{3\pi}{2} + 2\pi, \\ &\text{если} \quad \left( -\text{atan } 2(y, x) - \frac{3\pi}{2} \right) \leq -\pi. \end{aligned} \quad (5.1.21)$$



Это приведет к следующему (рис. 5.3). Если фаза – линейно возрастающая функция (рис. 5.4, а), то использование функции  $\text{atan} 2(y, x)$  вместо  $\text{atan} 2(x, y)$  приведет к изменению знака (функция линейно убывающая) и сдвигу на  $\frac{3\pi}{2}$  (рис. 5.4, б).

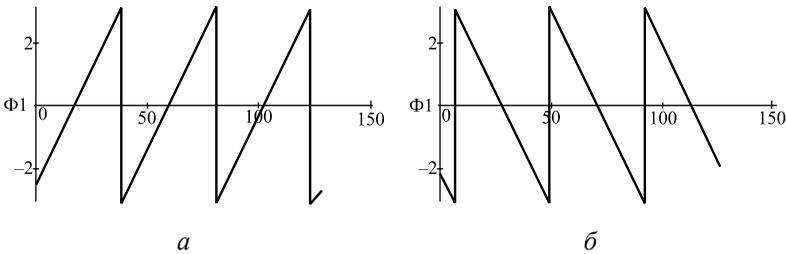


Рис. 5.4. Искажения при изменении числителя и знаменателя в формуле расшифровки:

а –  $\text{atan} 2(x, y)$  ; б –  $\text{atan} 2(y, x)$

Таким образом, выражение (5.1.16) можно записать в виде

$$\varphi = \arctan \left( \frac{\vec{I}^\perp \vec{C}}{\vec{I}^\perp \vec{S}} \right) = -\arctan \left( \frac{\vec{I}^\perp \vec{S}}{\vec{I}^\perp \vec{C}} \right) + \frac{3\pi}{2}. \quad (5.1.22)$$

При трех фазовых сдвигах

$$\vec{I}^\perp = \begin{bmatrix} 0 & 1 & -1 \\ -1 & 0 & 1 \\ 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} I_1 \\ I_2 \\ I_3 \end{bmatrix} = \begin{bmatrix} I_2 - I_3 \\ -I_1 + I_3 \\ I_1 - I_2 \end{bmatrix}.$$

Перемножая полученный вектор на векторы  $\vec{C}$  и  $\vec{S}$  соответственно, получим

$$\varphi(x, y) = \arctan \times \frac{(I_2 - I_3) \cos \delta_1 + (I_3 - I_1) \cos \delta_2 + (I_1 - I_2) \cos \delta_3}{(I_2 - I_3) \sin \delta_1 + (I_3 - I_1) \sin \delta_2 + (I_1 - I_2) \sin \delta_3}. \quad (5.1.23)$$



В литературе часто формулы расшифровки записываются с переменной мест числителя и знаменателя:

$$\varphi(x, y) = \frac{(I_2 - I_3) \sin \delta_1 + (I_3 - I_1) \sin \delta_2 + (I_1 - I_2) \sin \delta_3}{(I_2 - I_3) \cos \delta_1 + (I_3 - I_1) \cos \delta_2 + (I_1 - I_2) \cos \delta_3}. \quad (5.1.24)$$

При использовании различных формул расшифровки необходимо учитывать эти различия между двумя вариантами (5.1.22).

Если подставить в (5.1.24) значения конкретных фазовых сдвигов, получим более простые выражения:

$\delta_1 = 45^\circ, \delta_2 = 135^\circ, \delta_3 = 225^\circ$ $\left( \frac{\pi}{4}, \frac{3\pi}{4}, \frac{5\pi}{4} \right)$	$\varphi(x, y) = \arctan \frac{(I_3 - I_2)}{(I_1 - I_2)}$ <p>Это выражение использовано в [12]</p>
$\delta_1 = 0, \delta_2 = 120^\circ, \delta_3 = 240^\circ$ $\left( 0, \frac{2\pi}{3}, \frac{4\pi}{3} \right)$	$\varphi(x, y) = \arctan \sqrt{3} \frac{(I_3 - I_2)}{2I_1 - I_2 - I_3}$ <p>Это выражение описано в [10, 11].</p>
$\delta_1 = 0^\circ, \delta_2 = 90^\circ, \delta_3 = 180^\circ$ $\left( 0, \frac{\pi}{2}, \pi \right)$	$\varphi(x, y) = \arctan \frac{I_1 - 2I_2 + I_3}{I_1 - I_3}$
$\delta_1 = 0^\circ, \delta_2 = 60^\circ, \delta_3 = 120^\circ$ $\left( 0, \frac{\pi}{3}, \frac{2\pi}{3} \right)$	$\varphi(x, y) = \arctan \frac{(2I_1 - 3I_2 + I_3)}{\sqrt{3}(I_2 - I_3)}$

При четырех фазовых сдвигах

$$\vec{I}^\perp = \begin{bmatrix} 0 & 1 & 0 & -1 \\ -1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \\ 1 & 0 & -1 & 0 \end{bmatrix} \begin{bmatrix} I_1 \\ I_2 \\ I_3 \\ I_4 \end{bmatrix} = \begin{bmatrix} I_2 - I_4 \\ -I_1 + I_3 \\ -I_2 + I_4 \\ I_1 - I_3 \end{bmatrix}. \quad (5.1.25)$$



Выражение для определения разности фаз примет вид

$$\varphi(x, y) = \arctan \times \frac{(I_2 - I_4) \sin \delta_1 + (I_3 - I_1) \sin \delta_2 + (I_4 - I_2) \sin \delta_3 + (I_1 - I_3) \sin \delta_4}{(I_2 - I_4) \cos \delta_1 + (I_3 - I_1) \cos \delta_2 + (I_4 - I_2) \cos \delta_3 + (I_1 - I_3) \cos \delta_4}. \quad (5.1.26)$$

При четырех фазовых сдвигах

$$\delta_1 = 0^\circ, \delta_2 = 90^\circ, \delta_3 = 180^\circ, \delta_4 = 270^\circ \left( \delta_1 = 0, \delta_2 = \frac{\pi}{2}, \delta_3 = \pi, \delta_4 = \frac{3\pi}{2} \right)$$

получается простое выражение, которое наиболее часто используется при компьютерной обработке интерференционных картин [13]:

$$\varphi(x, y) = \arctan \frac{I_4 - I_2}{I_1 - I_3}. \quad (5.1.27)$$

Помимо приведенных существует достаточно много алгоритмов с различным числом фазовых сдвигов и их различными значениями.

При произвольных значениях фазовых сдвигов проще выполнять вычисления по формуле (5.1.16).

Рассмотрим код на языке C#. Пусть `n_sdv` – число фазовых сдвигов.

Процедура `Vector_orto(double[] vect)` возвращает вектор, ортогональный входному вектору.

```
public static double[] Vector_orto(double[] vect)
{
    int n = vect.Length;
    double[] v_sdv = new double[n];

    v_sdv[0] = vect[1] - vect[n - 1];
    v_sdv[n - 1] = vect[0] - vect[n - 2];
    for (int ii = 1; ii < n - 1; ii++) { v_sdv[ii] = vect[ii + 1] - vect[ii - 1]; }

    return v_sdv;
}
```

Следующий код задает значения синуса и косинуса для `fz[k]` – массива фазовых сдвигов, `k` меняется от 0 до `n_sdv-1`.



```

for (inti = 0; i<n_sdv; i++)
{
k_sin[i] = Math.Sin(fz[i]);
k_cos[i] = Math.Cos(fz[i]);
}

```

Следующий участок кода определяет разность фаз  $\varphi = \arctan\left(\frac{\vec{I}^\perp \vec{C}}{\vec{I}^\perp \vec{S}}\right)$ .

Входной массив: `zArrayPicture[k].array[i,j]` – оцифрованные значения  $k$  интерференционных картин. Выходной массив: `Faza.array[i,j]`. Размеры каждого массива по  $x$  от  $0$  до  $w1-1$ , по  $y$  от  $0$  до  $h1-1$ .

```

for (inti = 0; i < w1; i++)
{
for (int j = 0; j < h1; j++)
{
for (intk = 0; k<n_sdv; k++){i_sdv[k] =
= zArrayPicture[k].array[i, j]; }
double[] v_sdv = Vector_orto(i_sdv);
double fz1 = Vector_Mul(v_sdv, k_sin);
double fz2 = Vector_Mul(v_sdv, k_cos);
faza.array[i, j] = Math.Atan2(fz2, fz1);
}
}

```

`Vector_Mul(v_sdv, k_cos)` – процедура для определения произведения двух векторов

```

publicstaticdoubleVector_Mul(double[] vect1, double[] vect2)
{
int n = vect1.Length;
double s = 0;

for (int ii = 0; ii < n; ii++) { s += vect1[ii] * vect2[ii];
}

return s;
}

```



Зная разность фаз  $\varphi(x, y)$  и фазу опорной волны  $\varphi_r(x, y)$ , можно определить и исходное фазовое распределение волнового фронта  $\varphi_p(x, y)$ , рассеянного объектом:

$$\varphi_p(x, y) = \varphi(x, y) - \varphi_r(x, y). \quad (5.1.28)$$

Для формирования математической голограммы (5.1) необходимо также определить амплитуду волнового поля  $a_p(x, y)$ , отраженного от объекта, в плоскости голограммы. Значения амплитуд объектного и опорного пучков можно определить, физически перекрыв с помощью заслонки пучки в оптической схеме.

Но если у нас уже есть набор зарегистрированных голограмм, используемых для определения фазовых значений, то можно амплитуду объектного пучка получить методом пошагового сдвига, используя тот же набор голограмм.

## 5.2. ОПРЕДЕЛЕНИЕ АМПЛИТУДЫ ОБЪЕКТНОГО ПУЧКА МЕТОДОМ ПОШАГОВОГО ФАЗОВОГО СДВИГА

В отличие от классической интерферометрии, где основной задачей расшифровки является получение разности фаз интерферирующих пучков, для формирования математической голограммы нужно кроме фазы знать и амплитуду исходной волны [8]. Наиболее просто это можно сделать, если в качестве опорного пучка использовать плоскую волну с постоянной амплитудой.

Интенсивность интерференционной картины можно представить в виде (2.4.12):

$$I(x, y) = A_p^2(x, y) + A_r^2(x, y) + 2A_p(x, y)A_r(x, y)\cos(\varphi_p(x, y) - \varphi_r(x, y)), \quad (5.2.1)$$

где  $A_p(x, y)$  и  $A_r(x, y)$  – амплитуды объектной и опорной волны.

Обозначим  $A(x, y) = A_p^2(x, y) + A_r^2(x, y)$  – средняя яркость,  $B(x, y) = 2A_p(x, y)A_r(x, y)$  – амплитуда, а  $\varphi(x, y) = \varphi_p(x, y) - \varphi_r(x, y)$  – раз-



ность фаз, тогда интерференционные картины, полученные с фазовым сдвигом  $\delta_i$ , могут быть записаны как

$$I_i(x, y) = A(x, y) + B(x, y) \cos(\varphi(x, y) + \delta_i). \quad (5.2.2)$$

Выражения (5.1.10) и (5.1.11) можно переписать в виде

$$B \sin \varphi = -\frac{\vec{I}\vec{C}^\perp}{\vec{S}\vec{C}^\perp}; \quad (5.2.3)$$

$$B \cos \varphi = \frac{\vec{I}\vec{S}^\perp}{\vec{C}\vec{S}^\perp}. \quad (5.2.4)$$

$B(x, y)$  можно получить из выражений (5.2.3) и (5.2.4):

$$(B \sin \varphi)^2 + (B \cos \varphi)^2 = \left( \frac{\vec{I}\vec{S}^\perp}{\vec{C}\vec{S}^\perp} \right)^2 + \left( \frac{\vec{I}\vec{C}^\perp}{\vec{S}\vec{C}^\perp} \right)^2, \quad (5.2.5)$$

откуда следует, что

$$B = \frac{1}{|\vec{C}\vec{S}^\perp|} \sqrt{(\vec{I}\vec{S}^\perp)^2 + (\vec{I}\vec{C}^\perp)^2}$$

или

$$B = \frac{1}{|\vec{C}\vec{S}^\perp|} \sqrt{(\vec{I}^\perp\vec{S})^2 + (\vec{I}^\perp\vec{C})^2}. \quad (5.2.6)$$

Для четырех сдвигов:

$$(\vec{I}^\perp\vec{S}) = (I_2 - I_4) \sin \delta_1 + (I_3 - I_1) \sin \delta_2 + (I_4 - I_2) \sin \delta_3 + (I_1 - I_3) \sin \delta_4;$$

$$(\vec{I}^\perp\vec{C}) = (I_2 - I_4) \cos \delta_1 + (I_3 - I_1) \cos \delta_2 + (I_4 - I_2) \cos \delta_3 + (I_1 - I_3) \cos \delta_4;$$

$$\begin{aligned} \vec{C}\vec{S}^\perp &= (\cos \delta_2 - \cos \delta_4) \sin \delta_1 + (\cos \delta_3 - \cos \delta_1) \sin \delta_2 + \\ &+ (\cos \delta_4 - \cos \delta_2) \sin \delta_3 + (\cos \delta_1 - \cos \delta_3) \sin \delta_4. \end{aligned}$$



Обычно в качестве опорного пучка используется пучок, амплитудное распределение которого известно. Это может быть плоский или сферический волновой фронт. Зная  $B(x, y) = 2A_p(x, y)A_r(x, y)$  и значение амплитуды опорного пучка  $A_r(x, y)$ , можно определить и  $A_p(x, y)$ .

Код на языке C# для нахождения амплитуды и фазы математической голограммы по набору интерференционных картин может быть следующим:

```
for (inti = 0; i < w1; i++)
{
for (int j = 0; j < h1; j++)
{
for (int k = 0; k < n_sdv; k++){i_sdv[k] =
                        = zArrayPicture[k].array[i, j]; }

double[] v_sdv = Vector_orto(i_sdv);
double fz1 = Vector_Mul(v_sdv, k_sin);
double fz2 = Vector_Mul(v_sdv, k_cos);
double a = Math.Atan2(fz2, fz1);

double[] cos_orto = Vector_orto(k_cos);
double znmt = Vector_Mul(cos_orto, k_sin);
double am = Math.Sqrt(fz1 * fz1 + fz2 * fz2) /
Math.Abs(znmt);
am = am / (2 * amplit);
faza.array[i, j] = Complex.FromPolarCoordinates(am, a);
}
}
```

`faza.array[i, j]` – исходный комплексный массив.

В большинстве случаев амплитуда опорного поля неизвестна. Покажем, как можно найти отдельно и амплитуду опорного, и амплитуду объектного пучка, вводя переменный светофильтр в плечо опорного пучка [9].



При изменении интенсивности опорного пучка  $A_r(x, y)$  уравнение (5.2.1) можно представить как

$$I(n)_i = A_p^2 + (nA_r)^2 + 2nA_pA_r \cos \varphi_i, \quad (5.2.7)$$

где  $n$  – коэффициент изменения интенсивности опорного пучка. Введем обозначения:

$$B(n) = 2nA_pA_r = nB; \quad (5.2.8)$$

$$A(n) = A_p^2 + (nA_r)^2 = A_p^2 + n^2A_r^2. \quad (5.2.9)$$

Решая систему из двух уравнений (5.2.8) и (5.2.9), получим

$$A_r = + \sqrt{\frac{A - A(n)}{1 - n^2}}; \quad (5.2.10a)$$

$$A_p = + \sqrt{\frac{A(n) - n^2A}{1 - n^2}}. \quad (5.2.10b)$$

Представим систему уравнений (5.2.2) в векторном виде:

$$\vec{I} = \vec{A} + (B \cos \varphi) \vec{C} - (B \sin \varphi) \vec{S}. \quad (5.2.11)$$

Для трех фазовых сдвигов

$$\vec{S} = [\sin \delta_0, \sin \delta_1, \sin \delta_2]^T, \quad \vec{C} = [\cos \delta_0, \cos \delta_1, \cos \delta_2]^T \text{ и}$$

$$\vec{A} = (A_p^2 + A_r^2)[1, 1, 1]^T, \quad \vec{I} = [I_0, I_1, I_2]^T.$$

Анализ выражения (5.2.11) показывает, что для нахождения постоянной составляющей  $\vec{A}$  достаточно умножить вектор  $\vec{I}$  на вектор, ортогональный векторам  $\vec{C}$  и  $\vec{S}$ .

При трех фазовых сдвигах такому условию удовлетворяет вектор

$$\vec{U} = \vec{S} \times \vec{C}, \quad (5.2.12)$$

здесь символ  $\times$  обозначает векторное произведение векторов.



Матрицы оператора векторного произведения существуют только в трех- и семимерном пространствах [10]. В нашем случае это соответствует внесению трех или семи фазовых сдвигов. В трехмерном пространстве

$$\vec{S} \vec{C} = \begin{bmatrix} 0 & S_2 & -S_1 \\ -S_2 & 0 & S_0 \\ S_1 & -S_0 & 0 \end{bmatrix} \begin{bmatrix} C_0 \\ C_1 \\ C_2 \end{bmatrix}. \quad (5.2.13)$$

В семимерном пространстве

$$\vec{S} \vec{C} = \begin{bmatrix} 0 & S_2 & -S_1 & S_4 & -S_3 & -S_6 & S_5 \\ -S_2 & 0 & S_0 & S_5 & S_6 & -S_3 & -S_4 \\ S_1 & -S_0 & 0 & S_6 & -S_5 & S_4 & -S_3 \\ -S_4 & -S_5 & -S_6 & 0 & S_0 & S_1 & S_2 \\ S_3 & -S_6 & S_5 & -S_0 & 0 & -S_2 & S_1 \\ S_6 & S_3 & -S_4 & -S_1 & S_2 & 0 & -S_0 \\ -S_5 & S_4 & S_3 & -S_2 & -S_1 & S_0 & 0 \end{bmatrix} \begin{bmatrix} C_0 \\ C_1 \\ C_2 \\ C_3 \\ C_4 \\ C_5 \\ C_6 \end{bmatrix}. \quad (5.2.14)$$

При другом числе сдвигов можно построить вектор, удовлетворяющий условиям  $(\vec{U} \vec{C}) = 0$  и  $(\vec{U} \vec{S}) = 0$ . Для этого в пространстве размерностью  $N$  будем рассматривать некоторую гиперплоскость  $P$ , образуемую набором  $N$  векторов  $V$ :

$$P = V_0 e_0 + V_1 e_1 + \dots + V_{N-1} e_{N-1} + V_N e_N + e_{N+1} = 0, \quad (5.2.15)$$

где  $e_i$  – коэффициенты уравнения гиперплоскости  $P$ . Для ее построения необходимо доопределить имеющееся пространство векторов  $\{\vec{S}, \vec{C}\} \in R^2$  до пространства  $P \in R^N$  набором произвольных векторов, не совпадающих с вектором  $\vec{E} = \{[1 \ 1 \dots \ 1]_{1 \times N}\}^T$ , который параллелен вектору  $\vec{A}$ .

Коэффициенты  $e_i$  можно найти путем разложения определителя, строки которого составлены из векторов, принадлежащих данной гиперплоскости:



$$\begin{vmatrix}
 e_0 & e_1 & \cdots & e_N & e_{N+1} \\
 S_0 & S_1 & \cdots & S_N & 1 \\
 C_0 & C_1 & \cdots & C_N & 1 \\
 V^1_0 & V^1_1 & \cdots & V^1_N & 1 \\
 V^2_0 & V^2_1 & \cdots & V^2_N & 1 \\
 \cdots & \cdots & \cdots & \cdots & \cdots \\
 V^{N-2}_0 & V^{N-2}_1 & \cdots & V^{N-2}_N & 1 \\
 V^{N-1}_0 & V^{N-1}_1 & \cdots & V^{N-1}_N & 1
 \end{vmatrix} = 0. \quad (5.2.16)$$

Коэффициент  $e_{N+1}$  равен нулю, поскольку гиперплоскость натянута на радиус-векторы  $i$ , следовательно, проходит через начало системы координат, поэтому определитель (5.2.16) можно упростить и соответственно снизить его размерность на единицу:

$$\begin{vmatrix}
 e_0 & e_1 & \cdots & e_N \\
 S_0 & S_1 & \cdots & S_N \\
 C_0 & C_1 & \cdots & C_N \\
 V^1_0 & V^1_1 & \cdots & V^1_N \\
 V^2_0 & V^2_1 & \cdots & V^2_N \\
 \cdots & \cdots & \cdots & \cdots \\
 V^{N-1}_0 & V^{N-1}_1 & \cdots & V^{N-1}_N
 \end{vmatrix} = 0. \quad (5.2.17)$$

Коэффициенты уравнения общего вида  $e_i$  гиперплоскости  $P$  соответственно образуют координаты вектора нормали к ней. Например, при четырех сдвигах уравнение четырехмерной плоскости получаем, раскладывая определитель:

$$\begin{vmatrix}
 e_0 & e_1 & e_2 & e_3 \\
 \cos(\delta_0) & \cos(\delta_1) & \cos(\delta_2) & \cos(\delta_3) \\
 \sin(\delta_0) & \sin(\delta_1) & \sin(\delta_2) & \sin(\delta_3) \\
 1 & 1 & 1 & -1
 \end{vmatrix}. \quad (5.2.18)$$



Тогда коэффициенты уравнения плоскости будут равны

$$e_0 = \begin{vmatrix} \cos(\delta_1) & \cos(\delta_2) & \cos(\delta_3) \\ \sin(\delta_1) & \sin(\delta_2) & \sin(\delta_3) \\ 1 & 1 & -1 \end{vmatrix} =$$

$$= \sin(\delta_1 - \delta_2) + \sin(\delta_1 - \delta_3) - \sin(\delta_2 - \delta_3); \quad (5.2.19a)$$

$$e_1 = - \begin{vmatrix} \cos(\delta_0) & \cos(\delta_2) & \cos(\delta_3) \\ \sin(\delta_0) & \sin(\delta_2) & \sin(\delta_3) \\ 1 & 1 & -1 \end{vmatrix} =$$

$$= -\sin(\delta_0 - \delta_2) - \sin(\delta_0 - \delta_3) + \sin(\delta_2 - \delta_3); \quad (5.2.19б)$$

$$e_2 = \begin{vmatrix} \cos(\delta_0) & \cos(\delta_1) & \cos(\delta_3) \\ \sin(\delta_0) & \sin(\delta_1) & \sin(\delta_3) \\ 1 & 1 & -1 \end{vmatrix} =$$

$$= \sin(\delta_0 - \delta_1) + \sin(\delta_0 - \delta_3) - \sin(\delta_1 - \delta_3); \quad (5.2.19в)$$

$$e_3 = - \begin{vmatrix} \cos(\delta_0) & \cos(\delta_1) & \cos(\delta_2) \\ \sin(\delta_0) & \sin(\delta_1) & \sin(\delta_2) \\ 1 & 1 & 1 \end{vmatrix} =$$

$$= \sin(\delta_0 - \delta_1) + \sin(\delta_0 - \delta_2) - \sin(\delta_1 - \delta_2), \quad (5.2.19г)$$

и тогда  $\vec{U} = [e_0 \ e_1 \ e_2 \ e_3]^T$ .

Умножая полученный вектор  $\vec{U}$  на вектор  $\vec{I}$ , получим

$$\vec{I}\vec{U} = \vec{A}\vec{U} + (B \cos \varphi) \vec{C}\vec{U} - (B \sin \varphi) \vec{S}\vec{U} = \vec{A}\vec{U}. \quad (5.2.20)$$

Тогда средняя яркость  $A$  может быть рассчитана с помощью следующего выражения:

$$A = \frac{\vec{U}\vec{I}}{\vec{U}\vec{E}}, \quad (5.2.21)$$

где вектор  $\vec{E} = [1 \ 1 \ 1 \ 1]^T$ .



Подставляя найденное значение  $A$  в выражения (5.2.10а) и (5.2.10б), окончательно получим выражения для определения амплитуды опорного и объектного пучков:

$$A_r = \left\{ \frac{1}{1-n^2} \frac{\vec{U}(\vec{I} - \vec{I}(n))}{\vec{U} \vec{E}} \right\}^{\frac{1}{2}}; \quad (5.2.22a)$$

$$A_p = \left\{ \frac{1}{1-n^2} \frac{\vec{U}(\vec{I}(n) - n^2 \vec{I})}{\vec{U} \vec{E}} \right\}^{\frac{1}{2}}. \quad (5.2.26)$$

Для предотвращения усиления ошибок вектор желательно нормировать по длине:  $\vec{U} = \frac{\vec{U}}{|\vec{U}|}$ .

Измерения амплитуды опорного пучка прямо зависят от истинных значений коэффициента пропускания  $n$  нейтрального фильтра (или иных устройств, которые используются для изменения амплитуды). Эти значения нужно измерить.

Из (5.2.8)

$$n = B(n) / B. \quad (5.2.23)$$

Значение для определения  $B$  мы уже получили (см. 5.2.6), таким же образом можно получить и  $B(n)$ :

$$B = \left\{ \left( \frac{\vec{I} \vec{S}^\perp}{\vec{S} \vec{C}^\perp} \right)^2 + \left( \frac{\vec{I} \vec{C}^\perp}{\vec{S} \vec{C}^\perp} \right)^2 \right\}^{\frac{1}{2}}; \quad (5.2.24a)$$

$$B(n) = \left\{ \left( \frac{\vec{I}(n) \vec{S}^\perp}{\vec{S} \vec{C}^\perp} \right)^2 + \left( \frac{\vec{I}(n) \vec{C}^\perp}{\vec{S} \vec{C}^\perp} \right)^2 \right\}^{\frac{1}{2}}. \quad (5.2.24б)$$



Учитывая, что для трех сдвигов

$$\vec{I}^\perp = \begin{bmatrix} 0 & 1 & -1 \\ -1 & 0 & 1 \\ 1 & -1 & 0 \end{bmatrix} \vec{I}; \quad (5.2.25a)$$

$$\vec{I}^\perp(n) = \begin{bmatrix} 0 & 1 & -1 \\ -1 & 0 & 1 \\ 1 & -1 & 0 \end{bmatrix} \vec{I}(n), \quad (5.2.25b)$$

из выражений (6.2.24a) и (6.2.24b) получим

$$\vec{I}^\perp = [I_2 - I_1, I_0 - I_2, I_1 - I_0]^T = \begin{bmatrix} B(\cos(\varphi + \delta_2) - (\varphi + \delta_1)) \\ B(\cos(\varphi + \delta_0) - (\varphi + \delta_2)) \\ B(\cos(\varphi + \delta_1) - (\varphi + \delta_0)) \end{bmatrix}, \quad (5.2.26)$$

$$\begin{aligned} \vec{I}^\perp(n) &= [I_2(n) - I_1(n), I_0(n) - I_2(n), I_1(n) - I_0(n)]^T = \\ &= \begin{bmatrix} Bn(\cos(\varphi + \delta_2) - (\varphi + \delta_1)) \\ Bn(\cos(\varphi + \delta_0) - (\varphi + \delta_2)) \\ Bn(\cos(\varphi + \delta_1) - (\varphi + \delta_0)) \end{bmatrix}. \end{aligned} \quad (5.2.27)$$

Отсюда покомпонентное сложение элементов векторов дает

$$\frac{\sum_i^3 \vec{I}^\perp(n)}{\sum_i^3 \vec{I}^\perp} = \frac{3Bn\tilde{I}^\perp}{B\tilde{I}^\perp} = 3n, \quad (5.2.28)$$

где  $\tilde{I}^\perp$  – сумма компонент вектора  $\vec{I}^\perp$ , и тогда соответственно получим

$$n = \frac{1}{3} \frac{\sum_i^3 \vec{I}^\perp(n)}{\sum_i^3 \vec{I}^\perp}, \quad (5.2.29)$$



Таким образом, получено простое выражение для определения коэффициента пропускания.

Учет влияния амплитудного распределения может существенно увеличить качество восстановленного изображения из цифровых голограмм.

В этом разделе показано, каким образом можно определить амплитуду  $A_p(x, y)$  и фазу  $\varphi_p(x, y)$  исходного волнового фронта в плоскости голограммы по значениям интерференционных картин, полученных с пошаговым фазовым сдвигом. Определив комплексное распределение  $G(x, y) = a_p(x, y) \exp(i\varphi_p(x, y))$  для восстановления поля в плоскости изображения, необходимо выполнить преобразование Френеля в зависимости от расстояния между объектом и плоскостью регистрации.

### 5.3. ВОССТАНОВЛЕНИЕ КОМПЛЕКСНОЙ АМПЛИТУДЫ ВОЛНОВОГО ФРОНТА, ОТРАЖЕННОГО ОТ ОБЪЕКТА ПО МАТЕМАТИЧЕСКОЙ ГОЛОГРАММЕ

1. Зададим комплексную амплитуду поля в плоскости, касательной к объекту ( $x, y$ ) (рис. 5.5).

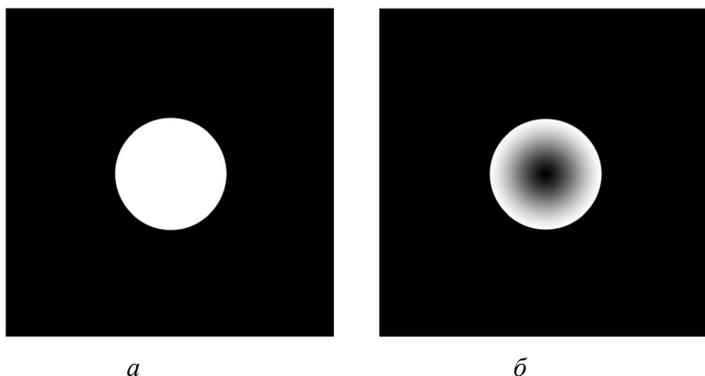


Рис. 5.5. Распределение комплексных амплитуд в плоскости, касательной к объекту:

$a$  – амплитуда;  $b$  – фаза. Фаза объекта распределена в диапазоне от  $-\pi$  до  $\pi$

2. Определим волновое поле на некотором расстоянии от объекта (рис. 5.6).

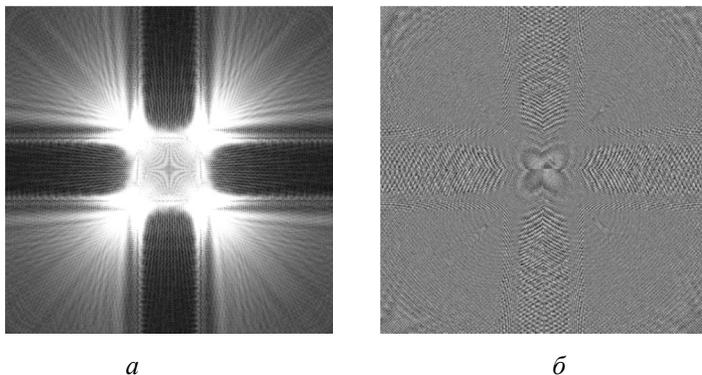


Рис. 5.6. Распределение комплексных амплитуд в плоскости голограммы ( $\eta, \xi$ ):

$a$  – амплитуда;  $b$  – фаза. Размер объекта 5 см, длина волны 500 нм, расстояние до плоскости голограммы 20 см

3. Регистрируем картины интенсивности на матрице фотоприемника при интерференции волнового поля, отраженного от объекта, и опорных полей с известным фазовым сдвигом ( $m$  – число картин интенсивности). На рис. 5.7 показаны четыре голограммы с фазовым сдвигом  $0^\circ, 90^\circ, 180^\circ, 270^\circ$ .

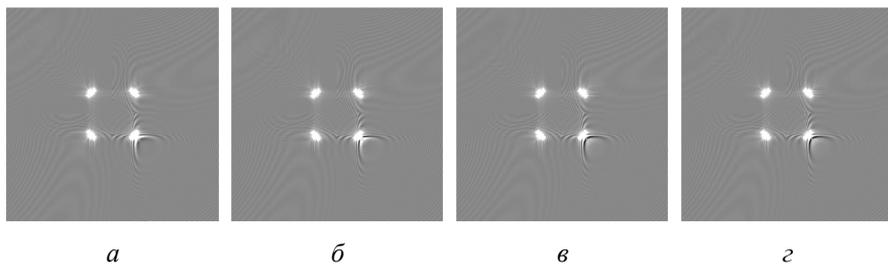


Рис. 5.7. Четыре голограммы с фазовым сдвигом:

$a - 0^\circ$ ;  $b - 90^\circ$ ;  $v - 180^\circ$ ;  $z - 270^\circ$

4. По зарегистрированным картинам интенсивности восстанавливаем фазовую разность волновых полей  $\varphi(x, y) = \arctan \left( \frac{\vec{I}^\perp \vec{S}}{\vec{I}^\perp \vec{C}} \right)$



и  $B(x, y) = \frac{1}{|\vec{C}\vec{S}^\perp|} \sqrt{(\vec{I}^\perp \vec{S})^2 + (\vec{I}^\perp \vec{C})^2}$ . Поскольку  $B(x, y) = 2A_p(x, y) \times A_r(x, y)$ , можно определить амплитуду поля, отраженного от объекта в плоскости голограммы  $A_p(x, y) = B(x, y) / (2A_r(x, y))$ .

5. Формируем математическую голограмму. Формируем волновой фронт  $A_p(x, y) \exp(i\varphi(x, y))$ . Умножаем на волну, соответствующую опорному пучку  $A_r(x, y) \exp(i\varphi_r(x, y))$ . В результате получим  $A_r(x, y) \times A_p(x, y) \exp(i\varphi_p(x, y))$ , поскольку  $\varphi(x, y) = \varphi_p(x, y) - \varphi_r(x, y)$ . Зная распределение амплитуды опорного пучка, можно определить и искомое выражение для математической голограммы  $A_p(x, y) \exp(i\varphi_p(x, y))$ .

6. В зависимости от расстояния до плоскости изображения осуществляем над математической голограммой преобразование Френеля или преобразование Фурье для получения распределения волнового фронта в области наблюдения (рис. 5.8).

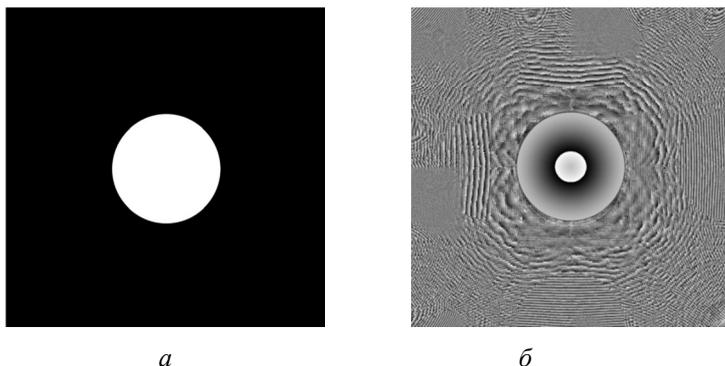


Рис. 5.8. Распределение амплитуды и фазы после восстановления из голограммы в плоскости наблюдения:

$a$  – амплитуда;  $b$  – фаза

7. Устраняем фазовую неоднозначность. Фазовая неоднозначность может быть устранена одним из способов, рассмотренных в [11–16].

Заметим, что шаги 1–3 получаются в результате работы голографического интерферометра с пошаговым фазовым сдвигом, шаги 4–6 реализуются в цифровом виде на компьютере.



Использование этого алгоритма устраняет все сложности, связанные с устранением влияния центрального пучка, однако требует записи нескольких интерференционных картин и специального оборудования в составе измерительной установки для внесения точных значений фазового сдвига.

Однако при небольших углах между интерферирующими пучками возможны ошибки при определении точного значения углов. Эта ошибка приводит к искажению фазовых значений. Приведем пример: зададим комплексную амплитуду поля в плоскости  $(x, y)$ , касательной к объекту (рис. 5.9).

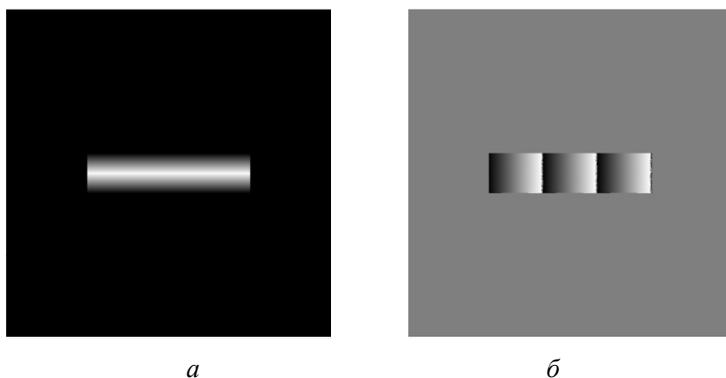


Рис. 5.9. Распределение комплексных амплитуд в плоскости, касательной к объекту:

*a* – амплитуда; *б* – фаза. Фаза объекта распределена в диапазоне от  $-\pi$  до  $\pi$

Пусть угол между интерферирующими пучками  $0,7^\circ$ . Однако при восстановлении умножим математическую голограмму на опорный фронт под углом  $0,8^\circ$ . В результате возникнут фазовые ошибки (рис. 5.10).

Использование небольших углов между опорным и объектным пучками значительно снижает требования к разрешению регистрирующих сред, что позволяет использовать цифровые матрицы при регистрации голограмм. Метод хорошо работает даже при центральном опорном пучке (голограммы Габора).

В большинстве случаев достаточно информации о пространственном распределении амплитуды сигнала. Для прямого определения фазового распределения необходимо точное знание об угле между ин-



терферирующими полосами. Значение этого угла может быть найдено только в результате предварительной калибровки.

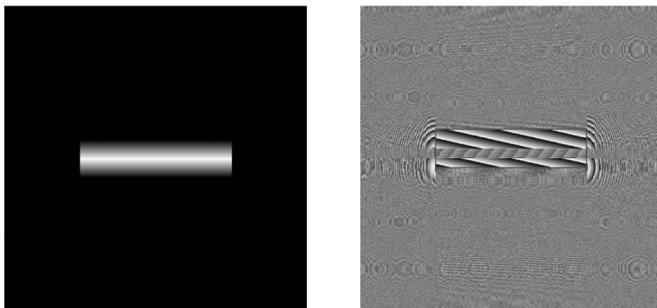


Рис. 5.10. Распределение амплитуды и фазы после восстановления из математической голограммы в плоскости наблюдения при неправильном определении угла сдвига

В качестве регистратора используются матрицы фотоприемников, которые имеют конечное число уровней квантования. Рассмотрим, как количество этих уровней сказывается на качестве восстановления.

#### 5.4. ВЛИЯНИЕ ЧИСЛА УРОВНЕЙ КВАНТОВАНИЯ НА КАЧЕСТВО ВОССТАНОВЛЕНИЯ КОМПЛЕКСНОЙ АМПЛИТУДЫ ОБЪЕКТНОЙ ВОЛНЫ

Квантование – это замена величины отсчета сигнала ближайшим значением из набора фиксированных величин – уровней квантования. Уровни квантования делят весь диапазон возможного изменения значений сигнала на конечное число интервалов – шагов квантования. Расположение уровней квантования обусловлено шкалой квантования. Используются как равномерные, так и неравномерные шкалы.

Искажения сигнала, возникающие в процессе квантования, называют *шумом квантования*. При инструментальной оценке шума вычисляют разность между исходным сигналом и его квантованной копией, а в качестве объективных показателей шума принимают, например, среднеквадратичное значение этой разности:

$$D_{\Sigma} = \sum_{k=1}^n D[\Delta x_k] = \frac{1}{12} q^2,$$

где  $\Delta x$  – погрешность квантования;  $q$  – величина интервала квантования.



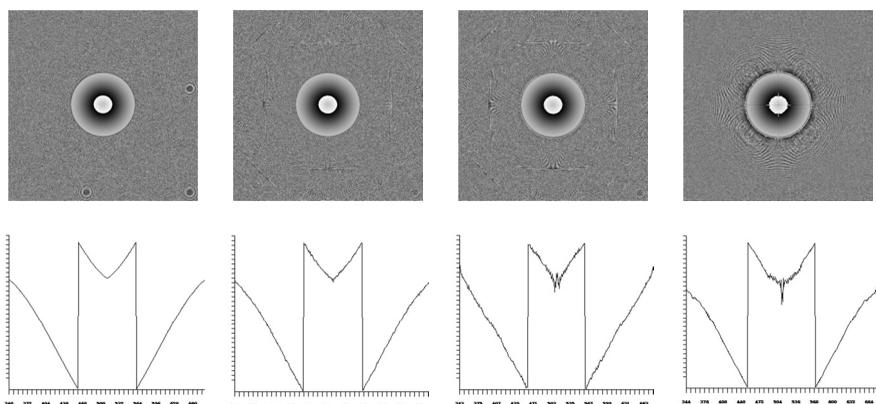


Рис. 5.12. Восстановленное распределение поля фаз в плоскости изображения после квантования слева направо 256, 64, 32, 16 уровней. Внизу графики вдоль окружности

Видно, что при уровнях более 64 качество восстановления достаточно для количественной расшифровки. Однако для снижения погрешности измерений необходимо не менее 256 уровней квантования.

Современные фотокамеры, используемые для регистрации, позволяют оцифровывать значения сигналов с 12 или 14 двоичными разрядами. Этого достаточно для регистрации и расшифровки цифровых голограмм с высокой точностью.

## БИБЛИОГРАФИЧЕСКИЙ СПИСОК К ГЛАВЕ 5

1. Гуров И. П. Компьютерная обработка интерференционных сигналов на основе алгоритма управляемого фазового сдвига / И. П. Гуров // Оптический журнал. – 1998. – № 10. – С. 38–42.
2. Современные методы цифровой голографии / С. А. Балтийский, И. П. Гуров, С. Де Никола [и др.] // Проблемы когерентной и нелинейной оптики / под ред. И. П. Гурова и С. А. Козлова. – Санкт-Петербург: СПбГУ ИТМО, 2004. – С. 91–117.
3. Liu J.P. Two-step-only quadrature phase-shifting digital holography / J. P. Liu, T. C. Poon // Optics Letters. – 2009. – Vol. 34. – P. 250–252.
4. Strain measurement using phase-shifting digital holography with two cameras / M. Fujigaki, R. Nishitani, Y. Morimoto, M. Fujigaki, R. Nishitani, Y. Morimoto // EPJ Web of Conferences. – 2010. – Vol. 6. – P. 30001. – doi: 10.1051/epjconf/20100630001.



5. *Гужов В. И.* Оптические измерения. Компьютерная интерферометрия / В. И. Гужов, С. П. Ильиных. – 2-е изд. – Москва: Юрайт, 2018. – 258 с. – ISBN 978-5-534-06855-9.

6. Универсальный алгоритм расшифровки / В. И. Гужов, С. П. Ильиных, Д. С. Хайдуков, А. Р. Вагизов // Научный вестник НГТУ. – 2010. – № 4 (41). – С. 51–58.

7. Generic algorithm of phase reconstruction in phase-shifting interferometry / V. Guzhov, S. Ilinykh, R. Kuznetsov, D. Haydukov // Optical Engineering. – 2013. – Vol. 52 (3). – P. 030501-1–030501-2.

8. Determination of complex amplitudes of the reference and subject the wave fields on interference patterns / V. Guzhov, S. Ilinykh, I. Marchenko, V. Emelyanov // The 11<sup>th</sup> International Forum on Strategic Technologies: proceedings of IFOST 2016, Novosibirsk, June 1–3, 2016. – Novosibirsk, 2016. – Vol. 1. – P. 531–534.

9. *Гужов В. И.* Определение интенсивности опорного и объектного пучков при использовании метода пошагового фазового сдвига / В. И. Гужов, С. П. Ильиных // Автоматика и программная инженерия. – 2017. – № 4 (22). – С. 68–73.

10. *Massey W. S.* Cross products of vectors in higher dimensional euclidean spaces / W. S. Massey // The American Mathematical Monthly. – 1983. – Vol. 90, N 10. – P. 697–701.

11. *Гужов В. И.* Методы измерения 3D профиля объектов. Фазовые методы: учебное пособие / В. И. Гужов. – Новосибирск: Изд-во НГТУ, 2016. – 83 с.

12. *Gushov V. I.* Automatic processing of fringe patterns in integer interferometers / V. I. Gushov, Yu. N. Solodkin // Optics and Lasers in Engineering. – 1991. – Vol. 14, iss. 4–5. – P. 311–324.

13. *Гужов В. И.* Анализ точности определения полной разности фаз в целочисленных интерферометрах / В. И. Гужов, Ю. Н. Солодкин // Автометрия. – 1992. – № 6. – С. 24–30.

14. *Gushov V. I.* The accuracy of determining the total phase difference in integer interferometers / V. I. Gushov, Yu. N. Solodkin // Optoelectronics, Instrumentation and Data Processing. – 1992. – N 6. – P. 23–28.

15. Решение проблемы фазовой неоднозначности методом целочисленной интерферометрии / В. И. Гужов, С. П. Ильиных, Р. А. Кузнецов, А. Р. Вагизов // Автометрия. – 2013. – Т. 49, № 2. – С. 85–91.

16. Solution of the problem of phase ambiguity by integer interferometry / V. I. Guzhov, S. P. Ilinykh, R. A. Kuznetsov, A. R. Vagizov // Optoelectronics, Instrumentation and Data Processing. – 2013. – Vol. 49, iss. 2. – P. 178–183.

---

---

## ГЛАВА 6

# ИСПОЛЬЗОВАНИЕ МЕТОДОВ КОМПЬЮТЕРНОЙ ГОЛОГРАФИИ ДЛЯ ЭКСПЕРИМЕНТАЛЬНЫХ ИССЛЕДОВАНИЙ

---

---

Компьютерная голография включает следующие этапы:

- получение голограммы с помощью оптической установки;
- регистрация голограмм;
- восстановление полезной информации из зарегистрированной голограммы с помощью компьютерных методов.

Для получения голограммы необходима оптическая установка, которая включает механические подставки, оптические элементы и лазер с достаточной глубиной когерентности [1].

Новыми элементами являются устройства ввода оптической информации для регистрации голограмм, компьютер и устройства для внесения управляемого фазового сдвига при использовании методов PSI.

В этой главе рассмотрены основные элементы голографической установки и приведены два примера: расшифровка голограмм, зафиксированных на фотопластинке, и получение и расшифровка цифровых голограмм методом пошагового фазового сдвига.

В качестве устройств регистрации в современной фотоаппаратуре применяются два вида фотоматриц: CMOS (от англ. *complementary metal-oxide-semiconductor*) или КМОП (комплементарная структура металл–оксид–полупроводник) и CCD (от англ. *charge-coupled device*) или ПЗС (прибор с зарядовой связью).

Достоинством КМОП (CMOS) технологии является низкое энергопотребление в статическом состоянии и дешевизна производства. В большинстве современных фотокамер используется именно эта технология.



Профессиональные устройства и бытовая фотоаппаратура основаны на одних и тех же принципах, часто в них используются одни и те же фотоприемные матрицы. Однако по стоимости они значительно различаются. В следующем разделе рассмотрим возможность использования коммерческих фотокамер для ввода голограмм.

### **6.1. ИСПОЛЬЗОВАНИЕ КОММЕРЧЕСКИХ ФОТОКАМЕР ДЛЯ ВВОДА ГОЛОГРАММ**

Наиболее распространенными устройствами, которые допускают проекцию изображения непосредственно на матрицу, являются фотокамеры Canon и Nikon со снимаемыми объективами.

Полнокадровые модели – как Canon, так и Nikon – снабжаются сенсором (CMOS-матрицей), размер которого соответствует размеру кадра 35 мм пленки, т. е.  $36 \times 24$  мм. У неполнокадровых моделей размеры матрицы незначительно отличаются: у Nikon  $\sim 24 \times 16$  мм, у Canon APS-C  $\sim 22,3 \times 14,9$  мм. Кроме размера важным параметром является число возможных элементов детектора.

Увеличить количество фотодатчиков можно двумя способами: путем увеличения матрицы или же уменьшения площади самих датчиков. В 2016 г. фирма Canon разработала матрицу с разрешением 250 мегапикселей. При этом сенсор имеет размеры  $29,2 \times 20,2$  мм. Разрешение составит  $19\,580 \times 12\,600$  пикселей. Также Canon разработала и самую большую на сегодняшний день CMOS-матрицу. Ее размер  $202 \times 205$  мм.

Чем большее количество света попадает на светочувствительные элемент матрицы, тем больший электрический заряд на нем формируется.

Все элементы, или ячейки матрицы, одинаковы с точки зрения спектральной чувствительности. Иначе говоря, они различают только уровень яркости падающего на них света, но не цвет. Для регистрации цвета используется нанесенный на матрицу набор светофильтров (рис. 6.1).

При расположении пикселей в виде прямоугольного раstra цвета приходится разбивать не на триады, как в основной цветовой схеме RGB (красный, зеленый, синий), а на тетрады. В большинстве случаев приоритет отдается зеленому цвету, и наиболее часто используемым элементом в схеме расположения светофильтров выступает зеленый как самый распространенный.

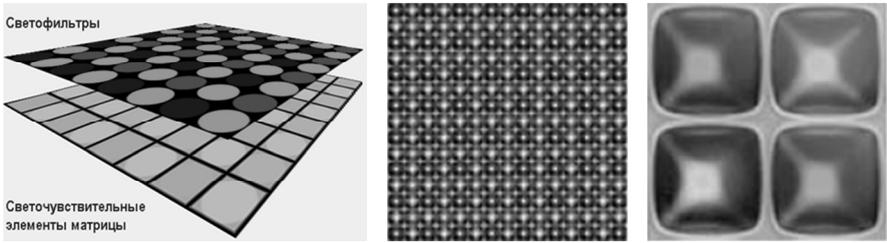


Рис. 6.1. Нанесение светофильтров на матрицу фотоприемников

Наиболее распространены три схемы фильтрации:

- Bayer Mosaic Filter,
- Complementary-Primary Mosaic Filter;
- Hybrid Complementary-Primary Mosaic Filter.

Во всех этих схемах для считывания информации как о яркости, так и о цвете используются элементы с минимальными элементами кластера  $2 \times 2$  (рис. 6.2).

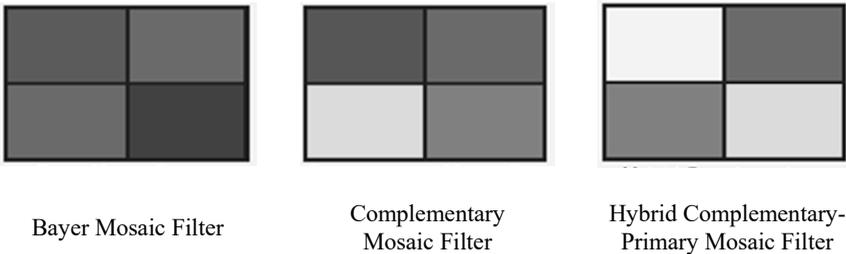


Рис. 6.2. Схемы фильтрации

Фильтр Байера состоит из 25 % красных элементов, 25 % синих и 50 % зеленых элементов. Он назван в честь его создателя, доктора Брайса Э. Байера (Brucе Bayer), сотрудника компании Kodak, запатентовавшего предложенный им фильтр в 1976 г. [2]. Поскольку зеленого цвета в каждой ячейке в два раза больше, чем фотодиодов других цветов, разрешающая способность такой структуры максимальна в зеленой области спектра, что соответствует особенностям человеческого зрения.

Для отличия от других разновидностей его называют GRGB, RGBG или (если надо подчеркнуть диагональное расположение красного и синего пикселей) RGGB.



Вследствие использования фильтров каждый фотоприемник воспринимает лишь  $1/3$  цветовой информации участка изображения, а  $2/3$  отсекается фильтром. Для получения остальных цветовых компонент используются значения из соседних ячеек. Недостающие компоненты цвета рассчитываются процессором камеры на основании данных из соседних ячеек в результате интерполяции. Таким образом, в формировании конечного значения цветного пикселя участвуют девять (или более) элементов матрицы.

Фильтр Байера исторически самый первый из массивов цветных фильтров. Однако искажения при цветопередаче и потери  $2/3$  светового потока привели к необходимости поиска других моделей световой фильтрации.

Иногда вместо GRGB-схемы используется CYGM-схема расположения фильтров (голубой, желтый, зеленый, пурпурный). Данный фильтр интересен тем, что пропускает приблизительно  $2/3$  падающего света, задерживая  $1/3$ . Тем самым достигается рост общей светочувствительности матрицы. Основная сложность при использовании CYGM-схемы – повышенная сложность при расчете цвета. Дело в том, что диапазон цветов, генерируемых аддитивными субтрактивным методами, разный: с помощью модели RGB оттенков генерируется больше, а погрешностей цветопередачи меньше. Поэтому такие альтернативные схемы большого распространения не получили, и сегодня в большинстве цифровых камер по-прежнему применяются классические байеровские фильтры.

Регулярная структура размещения элементов в некоторых случаях приводит к появлению муара. Возможность появления данного искажения зависит от сложности алгоритма, ответственного за расчет цвета. Для таких расчетов требуются высокопроизводительные микропроцессоры. Для снижения вероятности появления муара из-за повторяющихся участков иногда вносится некоторая псевдослучайность в размещении «синих» и «красных» элементов и используются модифицированные байеровские схемы с опорными группами, состоящими не из четырех элементов (RGBG), а из 12 либо 24.

Современные модели цифровых фотоаппаратов позволяют записывать изображения в RAW-формате – в виде сигналов яркости на каждом элементе, т. е. в черно-белом виде, и в файл записываются данные, полученные напрямую с матрицы. Производители цифровых фотоап-



паратов и RAW-конвертеров используют собственные адаптивные алгоритмы выделения интенсивности и цветовой информации, как правило, объявляемые производителем ноу-хау. Однако из изображения в формате RAW можно выделить большее число квантов на один пиксель. Если использовать изображения, полученные в стандартном JPEG-формате, на каждый цвет приходится по восемь разрядов. В форматах RAW для полупрофессиональных фотоаппаратов используется как минимум 12 разрядов.

Необходимость восстановления информации о цвете за счет соседних пикселей приводит к потере пространственного разрешения. Некоторые разработчики CMOS-матриц считают разрешение системы (в линиях) равным количеству считываемых с матрицы пикселей, разделенному на 1,5. Поскольку при оценке разрешающей способности объектива принято измерение в парах черной и белой линий мира Фурко на миллиметр (определяющих не одиночный пик, а пространственную частоту), то коэффициент пересчета разрешения матрицы в пары линий требует поправочного коэффициента 3,0. Поэтому при оцифровке изображений необходимо аккуратно выбирать частоту дискретизации, которая должна соответствовать заданным параметрам фото-матрицы.

Для увеличения пространственного разрешения используются технологии пиксельного сдвига сенсора. При помощи прецизионных пьезопроводов матрица передвигается в процессе экспозиции на один пиксель. Изображение экспонируется несколько раз с разным фильтром. Однако подобная технология очень дорогая, а кроме того, увеличивает время экспозиции. Поэтому применяется она преимущественно в профессиональной аппаратуре [3].

Кроме светофильтров, современные фотоаппараты обычно снабжаются и инфракрасными фильтрами (cut-off-filter), чтобы устранить влияние паразитного излучения. Устанавливается фильтр между матрицей камеры и объективом. Наличие этих фильтров необходимо учитывать при использовании инфракрасных лазеров.



## 6.2. УСТРОЙСТВА ДЛЯ ВНЕСЕНИЯ УПРАВЛЯЕМОГО ФАЗОВОГО СДВИГА

Для определения фазовых распределений используется метод пошагового фазового сдвига (PSI), описанный в главе 5. Для внесения фазовых сдвигов в схему интерферометра необходимо внести устройство для задания управляемых фазовых сдвигов.

Разность фаз между опорным и объектным пучками интерферометра определяется разностью длин оптических путей, которые проходят эти пучки. Таким образом, существует возможность управления разностью фаз за счет изменения либо разности хода лучей при постоянном значении показателя преломления либо за счет изменения скорости распространения световой волны, т. е. изменения показателя преломления. Поэтому всю совокупность известных модуляторов можно разделить на два широких класса – «отражающие» и «преломляющие».

Рассмотрим основные технические конструкции фазовых модуляторов (рис 6.3).

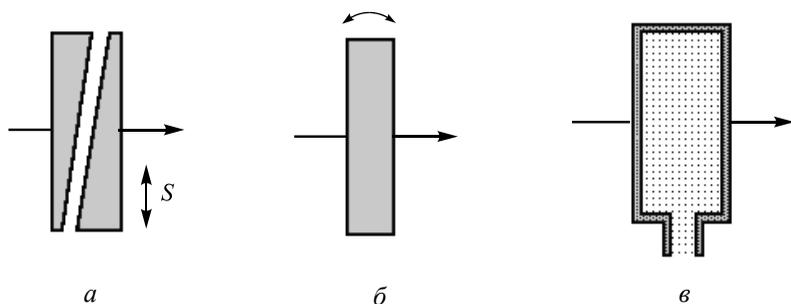


Рис. 6.3. Способы внесения сдвига фаз:

*a* – пространственный сдвиг стеклянного клина; *б* – поворот плоскопараллельной пластины; *в* – использование барокамеры

На рис. 6.3, *a* показан стеклянный клин, расположенный перпендикулярно падающему свету. Фазовый сдвиг происходит за счет пространственного сдвига клина:

$$\varphi = \frac{2\pi}{\lambda}(n-1)stg\gamma, \quad (6.2.1)$$

где  $\gamma$  – угол клина, а  $n$  – коэффициент преломления стекла клина.



Фазовый сдвиг может быть получен также наклоном на угол  $\alpha$  плоско-параллельной пластины толщины  $t$  (рис. 6.3, б):

$$\varphi = \frac{2\pi}{\lambda} t (\sqrt{n^2 - \sin^2 \alpha} - \cos \alpha). \quad (6.2.2)$$

Еще одним устройством фазового сдвига является барокамера (рис. 6.3, в), устанавливаемая в одном из плеч интерферометра. Этот технический прием хорошо известен из конструкции интерферометра Фабри–Перо.

В интерферометрии для получения фазового сдвига часто используют дифракционные решетки. Известно, что свет, дифрагированный при перемещении решетки, сдвигает фазу. Рассмотрим синусоидальную решетку пространственной частоты  $\theta$ , перемещающуюся со скоростью  $V$  (рис. 6.4, а). Пучки, сдвинутые на  $\pm\theta$ , имеют фазовые сдвиги, зависящие от времени:

$$\varphi = \pm 2\pi\theta Vt. \quad (6.2.3)$$

Это позволяет использовать первые порядки дифракции как опорные волны в интерференционном эксперименте.

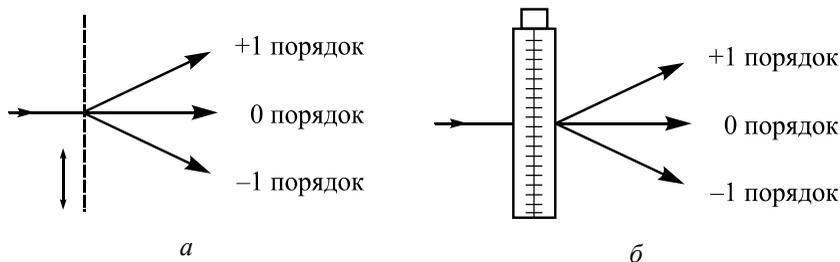


Рис. 6.4. Способы внесения сдвига фаз спомощью дифракционных решеток:

а – дифракционная решетка; б – акустооптический модулятор

Решетки бывают амплитудные и фазовые. Используются как вращающиеся радиальные решетки, так и акустооптические модуляторы (рис. 6.4, б), где бегущие в оптической среде волны используются в качестве ячеек Брэгга. В них создается поперечная акустическая волна, в результате чего ячейка начинает функционировать как толстая дифракционная решетка.



Из класса преломляющих модуляторов наиболее широкое применение получили электрооптические. Их действие основано на электрооптическом эффекте (эффекте Погкельса), характеризующемся возникновением сдвига фаз между компонентами обыкновенного и необыкновенного лучей световой волны, зависящего от величины приложенного электрического поля [4]. В качестве модулируемой среды в основном используются электрооптические кристаллы типа метаниобата лития или арсенида галлия (рис. 6.5).

Если поместить такой кристалл (называемый ячейкой Погкельса) между двумя поляризационными фильтрами (линейными поляризатором и анализатором) с углом  $90^\circ$  между их плоскостями поляризации, то получится модулятор. Поляризованный полуволновой пластиной пучок проходит через модулятор. На выходе модулятора он содержит две компоненты, одна из которых поляризована в вертикальной плоскости, другая – в горизонтальной. Относительные амплитуды этих компонент определяются ориентацией полуволновой пластины, установленной на выходе модулятора. Компоненты разделяются поляризационным делителем, и вторая полуволновая пластина делает плоскость поляризации горизонтально поляризованного пучка вертикальной. Фазовый сдвиг между этими пучками задается приложенным к электрооптическому модулятору напряжением.

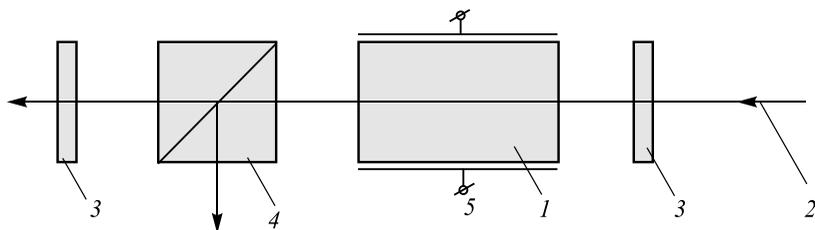


Рис. 6.5. Электрооптический модулятор:

- 1 – модулятор; 2 – модулируемое освещение; 3 – полуволновые пластины;  
4 – поляризационный делитель; 5 – подаваемое напряжение

Жидкие кристаллы, в которых тоже имеет место электрооптический эффект, мало используются из-за невысокой оптической неоднородности, приводящей к амплитудно-фазовым искажениям. Важным их достоинством является низкое (порядка десятков вольт) управляющее напряжение. Фазовые модуляторы, построенные на основе элек-



трооптических кристаллов, характеризуются высоким быстродействием, линейностью фазового сдвига, точностью, малой потребляемой мощностью. Однако они требуют применения дополнительных оптических элементов, таких как полуволновые пластины, поляризационные делители, имеющие высокую стоимость и трудоемкость изготовления. Важным недостатком электрооптических модуляторов является низкая предельно допустимая мощность модулируемого излучения.

В поляризационных интерферометрах фазовый сдвиг можно получить вращением полуволновой пластины или вращением поляризаторов (рис. 6.6). Устройство состоит из полуволновой пластины, смонтированной на валу шагового двигателя и четвертьволновой пластины, закрепленной стационарно так, что ось  $F$  образует с осью  $x$  угол в  $45^\circ$ .

Когда горизонтально поляризованная волна проходит через полуволновую пластину, ориентированную так, что ее ось образует с осью  $x$  угол  $\theta$ , линейная поляризация поворачивается на угол  $2\theta$ . Эта волна может быть представлена как линейная комбинация двух волн с круговой поляризацией и разностью фаз  $4\theta$ .

После того как волны проходят через полуволновую пластину, они из круговой поляризации преобразуются в горизонтальную и вертикальную, сохраняя фазовые отличия между собой.

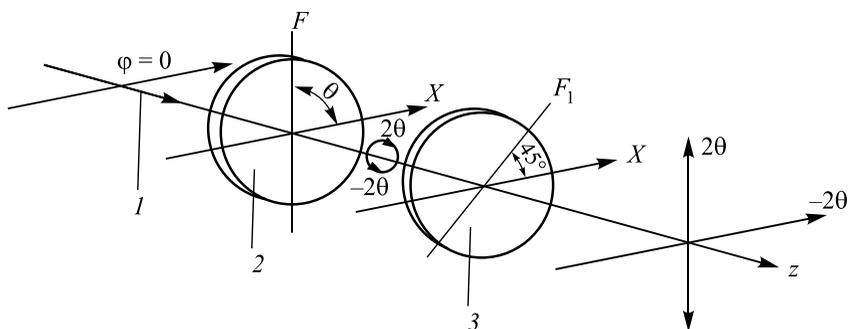


Рис. 6.6. Электромеханический модулятор:

1 – модулируемое освещение; 2 – полуволновая пластина;  
3 – четвертьволновая пластина

Итак, изменяя  $\theta$ , относительная разность между двумя ортогонально поляризованными волнами будет  $4\theta$ . Далее две волны разделяются поляризационным делителем, как и в предыдущем случае.



Несмотря на хорошую точность сдвига фаз, такие устройства характеризуются низким быстродействием, наличием механических узлов и дорогостоящей поляризационной оптики.

Лазеры с перестраиваемой длиной волны также могут использоваться для измерения фазы:

$$\varphi = \frac{2\pi}{\lambda^2} L \Delta\lambda, \quad (6.2.4)$$

где фазовое изменение длины волны  $\Delta\lambda$  достигается изменением управляющего тока.

Наиболее просто фазовый сдвиг осуществляется перемещением зеркала. Фазовый сдвиг происходит за счет поступательного движения зеркала на расстояние  $s$ :

$$\varphi = \frac{2\pi}{\lambda} s 2 \cos \alpha, \quad (6.2.5)$$

где  $\alpha$  – угол падения;  $\lambda$  – длина волны.

Перемещение зеркала может осуществляться электродвигателем с помощью микрометрического винта. Однако из-за высокой сложности изготовления прецизионного механического двигателя и низкой точности внесения фазового сдвига подобные приборы не нашли широкого применения.

Более совершенными являются магнитострикционные модуляторы. Принцип их действия основан на деформации стержня из магнитного материала, помещенного в магнитное поле. Модулятор представляет собой стержень с приклеенным зеркалом и катушку, в которую подается управляющий ток. Хотя точность такого модулятора выше, необходимость создания сильных полей и токов затрудняет его применение.

Широкое использование при создании фазовых модуляторов нашел пьезоэлектрический эффект. Это объясняется малостью габаритов таких устройств, низкой стоимостью, высокой точностью внесения фазового сдвига, относительно низкими управляющими напряжениями (порядка сотен вольт), низкой потребляемой мощностью, большой механической прочностью, надежностью, высоким быстродействием. Для этих целей используется пьезокерамика, создаваемая из материалов на основе твердых растворов цикроната-титана свинца. На рис. 6.7



показана конструкция модулятора на основе пьезоэлектрического преобразователя.

При подаче на электроды управляющего напряжения возникает пьезоэффект, который приводит к сжатию или растяжению пьезокерамики и закрепленного на ней отражающего покрытия. Зависимость фазового сдвига от величины приложенного напряжения выражается аналитической зависимостью

$$\varphi = \frac{\pi V_k}{V_0}, \quad (6.2.6)$$

где  $V_0$  – полуволновое напряжение (т. е. напряжение, при котором произойдет перемещение на половину длины волны);  $V_k$  – приложенное напряжение.

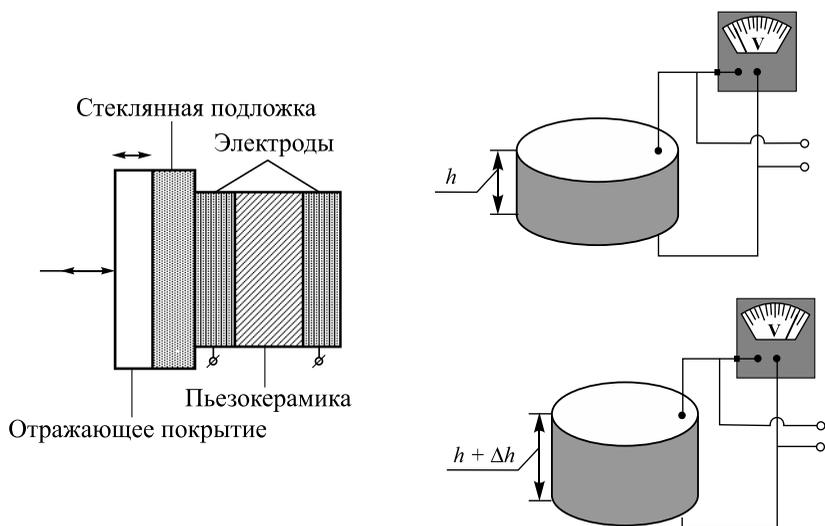


Рис. 6.7. Сдвиг фаз перемещением зеркала, закрепленного на пьезоKERAMИКЕ

Помимо относительной простоты внедрения пьезокерамики в интерферометрические схемы, простоты сопряжения с компьютером, точности позиционирования, удобства калибровки и управления сдвигом, данный метод обладает рядом недостатков. Прежде всего это высокая чувствительность пьезокерамики к изменению температуры.



При работе пьезокерамика существенно нагревается. На точность внесения фазового сдвига также оказывают влияние нелинейность перемещения пьезокерамики и гистерезис. Для проведения измерений необходимо выбирать наиболее линейный участок петли гистерезиса в зависимости от подаваемого напряжения. Поэтому при проведении высокоточных измерений необходимо производить калибровку каждый раз перед сбором данных.

На рис. 6.8 показана структурная схема устройства сдвига фаз на основе пьезокерамики ЦТС-19. Для управления от компьютера использовалось устройство сопряжения, которое управляет двенадцатиразрядным цифроаналоговым преобразователем (ЦАП). Полуволновое напряжение модуляторов 170 В, рабочая апертура 30 мм.

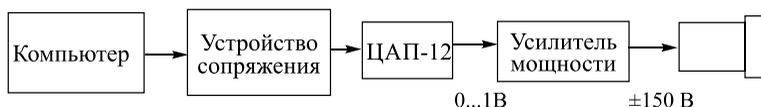


Рис. 6.8. Структурная схема устройства управления пьезокерамикой

На рис. 6.9 показано устройство сдвига фаз, смонтированное на механической подставке. Оно состоит из зеркала, закрепленного на пьезокерамике. На пьезокерамику подается напряжение, значение которого задается компьютером.

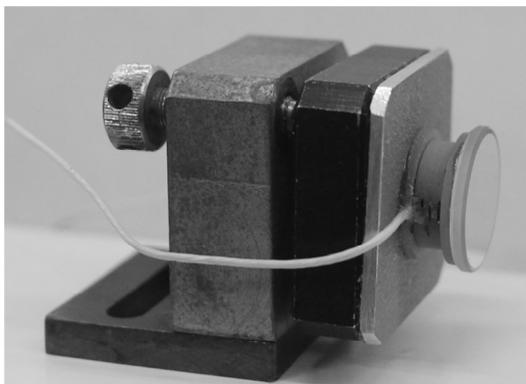


Рис. 6.9. Устройство внесения фазового сдвига на основе пьезокерамики

В зависимости от типа и размера керамики для изменения разности хода на одну длину волны может потребоваться напряжение от несколь-



ких вольт до 300 и более вольт. Поэтому для согласования устройства управления с полуволновым напряжением пьезокерамики используются усилители мощности. Выходное напряжение усилителей  $\pm 150$  В. Линейность передаточной характеристики обеспечивалась на уровне 0,1 %. Суммарное время установки устройства позиционирования  $\sim 1$  мс.

### **6.3. РАСШИФРОВКА АНАЛОГОВЫХ ГОЛОГРАММ, ЗАФИКСИРОВАННЫХ НА ФОТОПЛАСТИНКЕ**

В этом разделе рассмотрен процесс компьютерного восстановления волнового фронта с фотопластинки, на которой зарегистрирована голограмма объекта.

Стандартные схемы получения голограмм не предусматривали включение устройств внесения фазового сдвига. Существует много систем, которые не планируется модифицировать, но количественную компьютерную обработку требуется проводить.

При записи голограмм используются материалы с большим разрешением. Для компьютерной расшифровки голограммы необходимо оцифровать ее с достаточным разрешением. Поэтому для непосредственного ввода голограмм с фотопластинки использовался оптический микроскоп.

Размер поля микроскопа при использовании микрообъективов составляет около одного миллиметра, в то время как голограммы часто имеют пространственный размер более 5 см. Поэтому для ввода всей голограмм необходимо сканировать поле по частям и сшивать отдельные изображения. Для автоматизации процесса ввода проведена модификация существующего оптического микроскопа «Ломо Метам Р-1».

#### **6.3.1. МОДИФИКАЦИЯ ОПТИЧЕСКОГО МИКРОСКОПА ДЛЯ ВВОДА ГОЛОГРАММ**

Для ввода голограмм в компьютер использовался модифицированный микроскоп «Ломо Метам Р-1» [5]. Модификация микроскопа состояла в установке фотокамеры в качестве окуляра и автоматизированного предметного столика для перемещения объекта [6]. Наблюдение может проводиться в отраженном и проходящем свете.

Для регистрации изображений использовалась серийная цифровая камера Canon 650D. В качестве датчика в ней используется CMOS-



матрица с разрешением 18 млн пикселей ( $5184 \times 3456$  пикселей). Размер матрицы  $22,3 \times 14,9$  мм (рис. 6.10). Физический размер пикселя в матрице составляет 4,3 мкм.

Изображение напрямую проецируется на матрицу фотоприемников без использования объектива. Фотокамера подключается через интерфейсный кабель, с помощью которого цифровой кадр передается в компьютер.



Рис. 6.10. Цифровая фотокамера Canon 650D

Предметный столик в микроскопе был заменен на моторизованный двухосный линейный транслятор. Движение по осям осуществляется с помощью двух шаговых двигателей [7] (рис. 6.11). Величина перемещения на один шаг составляет 2,5 мкм. Возможно перемещение на  $1/8$  и  $1/16$  шага. Максимальная скорость перемещения – 10 мм/с.

Управление линейным транслятором осуществляется с помощью контроллера шагового двигателя OSM-42RA фирмы «Онитекс» по протоколу Modbus [8, 9].

На рис. 6.12 показан общий вид интерференционного микроскопа.

Общее изображение сшивалось из отдельных кадров при перемещении столика. Сшивка кадров позволяет увеличить поле зрения до  $10 \times 10$  см.

Перемещение столика зависит от механической нестабильности установки, поэтому возможны ошибки при сшивке кадров. Точное позиционирование должно быть обеспечено сшиванием кадров с совпадением до пикселя. Такая точность достигалась с помощью выделения особых точек. Изображения с большим увеличением снимают по ча-



стям с достаточным разрешением, а затем программа распознает особые точки на краях изображений и пытается скомпоновать их в одно целое изображение.

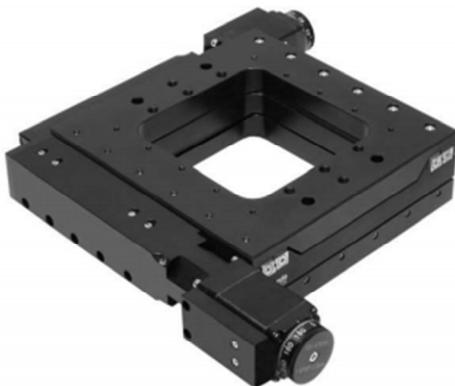
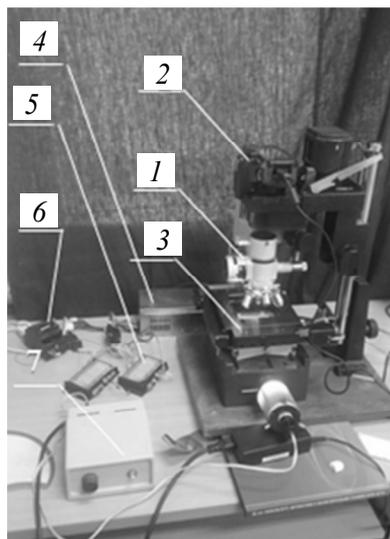


Рис. 6.11. Моторизованный двухосный линейный транслятор

Рис. 6.12. Конструкция автоматизированного оптического микроскопа для ввода голографических изображений:

1 – микроскоп; 2 – фотоаппарат; 3 – автоматизированный предметный столик; 4 – блок питания; 5 – контроллер шагового двигателя; 6 – преобразователь интерфейса RS-485/RS-232



Использование программной сшивки различных кадров позволило отказаться от использования аппаратных измерительных средств контроля при перемещении предметного столика в случае сбоя шаговых



двигателей. Склеивание изображений позволяет расширить поле зрения микроскопа и получить одно изображение всей области в высоком разрешении.

Поле зрения одного кадра зависит от используемого микроскопа и матрицы фотоприемника. Для определения чувствительности вводилось изображение линейки с миллиметровыми штрихами (рис. 6.13).

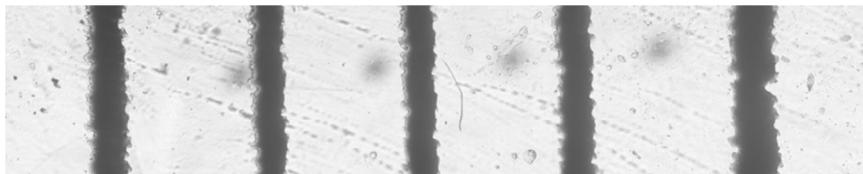


Рис. 6.13. Изображение линейки после сшивания нескольких кадров (размер 14 538 точек по  $x$ -координате, физический размер 5,3 мм)

Разрешающая способность оптического микроскопа определяется оптическими свойствами объектива. Э. Аббе показал, что в результате дифракции существует теоретический предел разрешающей способности (дифракционный предел), определяемый длиной волны  $\lambda$  используемого света. Классический предел разрешения идеальной оптической системы определяется критерием Рэля (J.W. Rayleigh, 1879) и равен

$$R = \frac{\lambda}{2NA}, \quad (6.3.1)$$

где  $NA$  – числовая апертура.

В оптическом микроскопе пространственное разрешение  $R$  по осям  $x, y$  определяется как

$$R = 0,61 \frac{\lambda}{NA^{obj}}, \quad (6.3.2)$$

где  $NA^{obj}$  – числовая апертура микрообъектива. Числовая апертура для микрообъективов вычисляется по следующей формуле:

$$NA^{obj} = n \sin(\mu), \quad (6.3.3)$$

где  $n$  – индекс преломления среды между передней линзой объектива ( $n = 1$  для воздуха,  $n = 1,51$  для масла);  $\mu$  – половина угловой апертуры (рис. 6.14).



Чем больше  $\mu$ , тем больше числовая апертура. Каким бы ни был угол  $\mu$ , величина  $\sin(\mu)$  не может быть больше единицы. Поэтому при работе в воздухе теоретически максимальное значение числовой апертуры  $NA^{obj} = 1$ . На практике  $NA^{obj}$  не превышает 0,95. Для достижения более высоких значений  $NA^{obj}$  необходима иммерсионная среда между передней линзой и образцом. Для этой цели используется масло или вода.

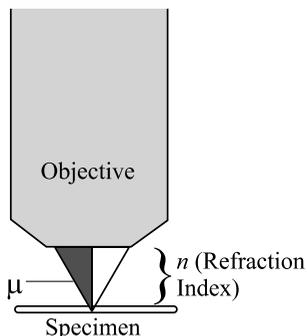


Рис. 6.14. Определение числовой апертуры



Рис. 6.15. Маркировка микрообъективов

На объективе обязательно указывается увеличение и числовая апертура (рис. 6.15).

Размер числовой апертуры зависит от конструкции микрообъектива. При одном и том же увеличении числовая апертура современных микрообъективов больше.

Таким образом, зная числовую апертуру, можно определить разрешение системы. Например, для объективов с увеличением в восемь раз (8X) числовая апертура  $NA^{obj} = 0,2$  и разрешение составит  $R \approx 3\lambda$ , для объектива с увеличением в десять раз (10X) числовая апертура  $NA^{obj} = 0,3$  и  $R \approx 2\lambda$ . При длине волны  $\lambda \approx 0,5$  мкм для микрообъектива 8X разрешение  $R \approx 1,5$  мкм, для 10X –  $R \approx 1$  мкм.

В модифицированном микроскопе при использовании микрообъективов с различным увеличением на один пиксель приходится 8X – 365 нм,



10X – 304 нм, 20X – 183 нм, 40X – 103 нм. Таким образом, сшивка кадров обеспечивает необходимое разрешение.

Расстояние между полосами при интерференции плоских пучков определяется выражением

$$\Delta x = \frac{\lambda}{2 \sin(\alpha / 2)}.$$

При угле  $5^\circ$  между интерферирующими пучками  $\Delta x = 5,7$  мкм ( $\lambda = 0,5$  мкм). По теореме Котельникова необходимо как минимум две точки на полосу, поэтому разрешение системы должно быть менее 2,8 мкм.

При угле  $10^\circ$  необходимое пространственное разрешение 1,4 мкм, при  $15^\circ$  – 1 мкм, при  $20^\circ$  – менее 0,7 мкм.

Для нашей системы при увеличении 8X на одну точку приходится 365 нм, а разрешение  $\sim 3\lambda$ . Это значит, что увеличение системы избыточно. Поэтому для сокращения вычислений можно результирующий массив цифровых значений усреднить по окрестности  $4 \times 4$  точки. Это не приведет к ухудшению результатов при восстановлении изображений из голограмм, поскольку разрешение системы больше, но позволит значительно сократить время вычислений.

### 6.3.2. РАСШИФРОВКА АНАЛОГОВЫХ ГОЛОГРАММ

Голограммы получались способом, предложенным Э. Лейтом и Ю. Упатниексом.

Схема записи голограммы показана на рис. 6.16. Пучок света от лазера с помощью делителя пучка (3) и зеркала (4) проходят разные пути. Объектный и опорный пучки расширяются с помощью расширителей пучка (5) и (6). Объектный пучок освещает объект (9) и интерферирует с опорным пучком на фотопластинке (8).

Угол между опорным и объектным пучком составлял от  $5$  до  $20^\circ$ .

На рис. 6.17 показана фотография голографической установки [10].

Голограммы фиксировались на голографической пластинке.

На рис. 6.18, *a* показана керамическая фигурка, которая использовалась в качестве объекта, а на рис. 6.18, *б* – результат восстановления изображения из голограммы с помощью лазерного пучка.

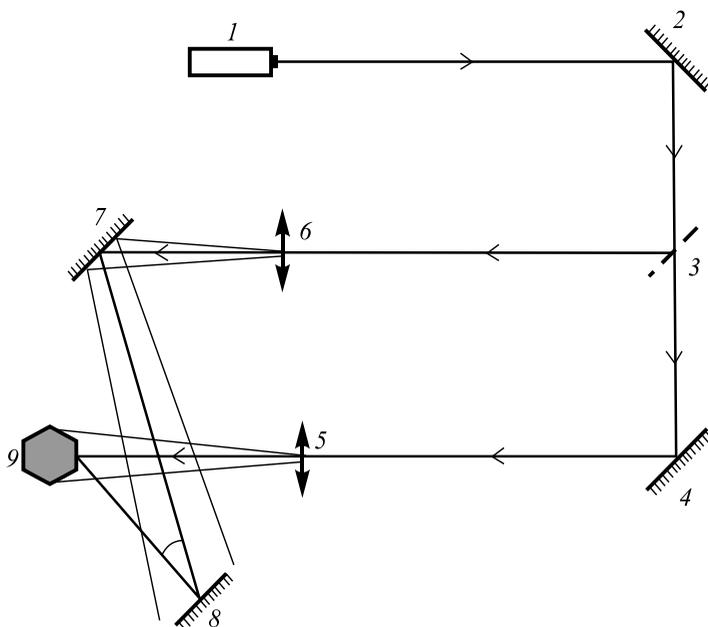


Рис. 6.16. Схема записи голограммы:

1 – лазер; 2, 4, 7 – отражающие зеркала; 3 – делитель светового пучка; 5, 6 – расширители пучка; 8 – фотопластинка; 9 – объект

Наличие моторизованного столика позволяет значительно расширить поле зрения микроскопа при оцифровке голограмм. Посмотрим, как зависит качество восстановленного изображения от увеличения размера оцифрованной области при одинаковом разрешении.

Для ввода использовался микрообъектив с восьмикратным увеличением. Разрешение системы составит около 1,5 мкм. На рис. 6.19 показан результат сшивки 25 кадров (пяти по оси  $X$  и пяти по оси  $Y$ ). В одном кадре содержится  $5184 \times 3456$  пикселей. Физический размер оцифрованной голограммы равен  $9,4 \text{ мм} \times 4,3 \text{ мм}$ .

Поскольку голограмма записана в области Фраунгофера, для восстановления использовалось преобразование Фурье. Результат восстановления при увеличении площади голограммы показан на рис. 6.20.

На рис. 6.21 показаны увеличенные изображения объекта. Видно, что при увеличении площади голограммы размер спеклов уменьшается и увеличивается качество восстановленного изображения. Такой же



эффект получается при восстановлении изображения с физической голограммы при увеличении апертуры восстанавливающего пучка.

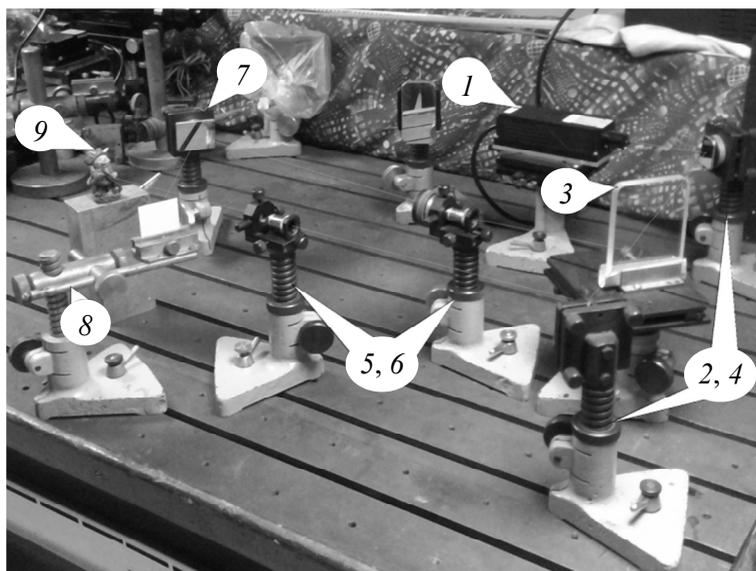


Рис. 6.17. Голографическая установка:

1 – лазер; 2, 4, 7 – отражающие зеркала; 3 – делитель светового пучка; 5, 6 – расширители пучка; 8 – фотопластинка; 9 – объект



а



б

Рис. 6.18. Объект для записи голограммы (а); действительное и мнимое изображение, восстановленное из голограммы лазерным пучком (б)



Рис. 6.19. Оцифрованная голограмма размером  $25\,861 \times 11\,690$  точек

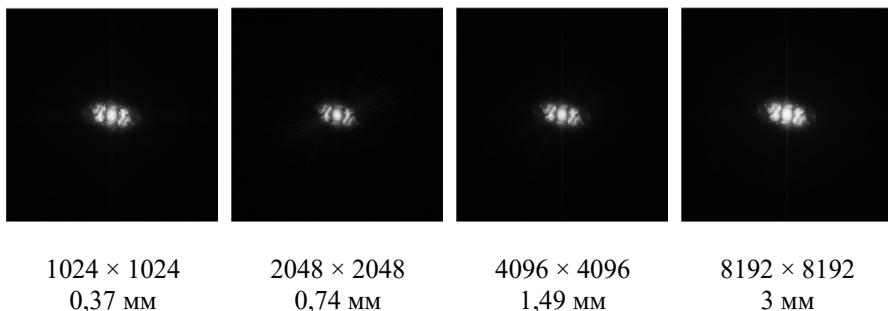


Рис. 6.20. Восстановленные на компьютере действительное и мнимое изображения объекта при различном размере участка голограммы, из которого производилось восстановление, но при одном разрешении

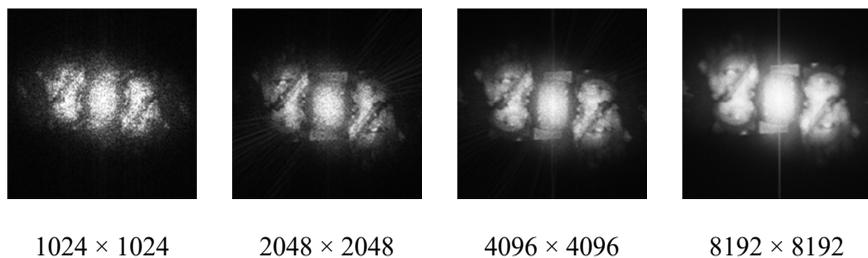


Рис. 6.21. Сравнение качества действительного и мнимого изображения объекта при увеличении анализируемого участка голограммы



При небольших углах сдвига бóльшая часть кадра неинформативна (рис. 6.22). Можно снизить время вычислений, используя следующий прием. Выполняем усреднение по четырем соседним точкам и уменьшаем размер массива в два раза по  $x$  и  $y$ :

$$U'_h(m/2, n/2) = \frac{1}{4} [U_h(m, n) + U_h(m, n+1) + U_h(m+1, n) + U_h(m+1, n+1)]. \quad (6.3.4)$$

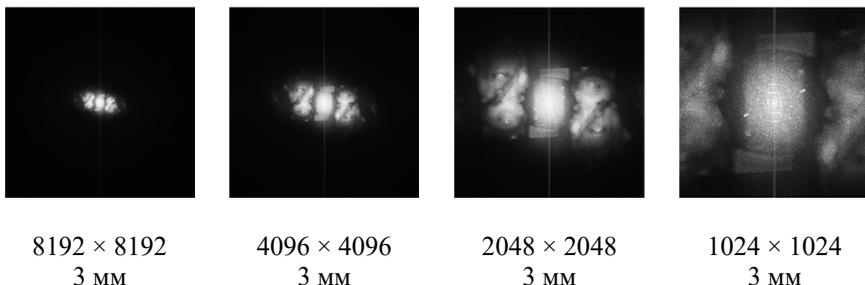


Рис. 6.22. Восстановленные на компьютере действительное и мнимое изображения объекта при снижении разрешения. Физический размер участка голограммы одинаков

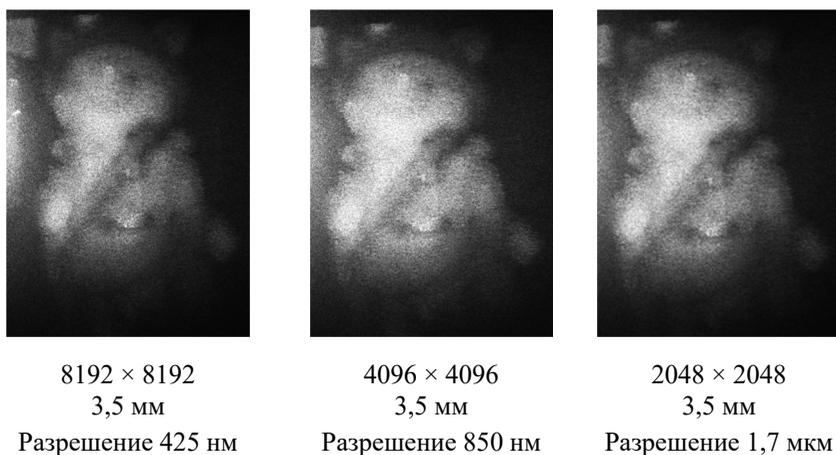


Рис. 6.23. Увеличенные участки изображений, показанных на рис. 6.22



Таким образом, мы снижаем разрешение, с которым оцифровывается голограмма при одном и том же физическом размере участка. Затем выполняем преобразование Фурье (рис. 6.22).

На рис. 6.23 показаны увеличенные части восстановленного изображения.

В разделе 6.3.1 показано, что число точек, которое фиксируется на фотоматрице камеры, является избыточным. Реальное разрешение определяется только типом используемого микрообъектива. Из рис. 6.23 видно, что качество восстановления при числе пикселей  $8192 \times 8192$  и  $2049 \times 2048$  примерно одинаковое. Поэтому без потери качества размер массива можно уменьшить в 16 раз (усредняя точки по окрестности  $4 \times 4$  пикселя).

### 6.3.3. РАСШИФРОВКА ГОЛОГРАММ ПРИ РАЗНЫХ УГЛАХ МЕЖДУ ИНТЕРФЕРИРУЮЩИМИ ПОЛОСАМИ

При увеличении угла между объектным и опорным пучками влияние центрального пучка можно снизить, поэтому качество восстановленного изображения увеличивается. Так как при этом размер полос на голограмме уменьшается, то требуется большее разрешение фотоматериалов, которые используются для регистрации голограмм.

На оптической установке (рис. 6.16, 6.17) были получены голограммы с разными углами между интерферирующими пучками – 5, 10, 15, 20°.

Для оцифровки и ввода в компьютер использовался микрообъектив с восьмикратным увеличением (8X). Размер оцифрованного кадра  $5184 \times 3456$ . Для восстановления изображения из голограммы использовался алгоритм БПФ для произвольного количества точек.

На рис. 6.24 показаны восстановленные изображения из разных голограмм, полученных при различных углах.

Ниже показаны увеличенные участки восстановленных изображений (рис. 6.25).

Видно, что хотя пучки, которые формируют действительное и мнимое изображения, расходятся, но при оцифровке голограмм с углами 15 и 20° выбранного при оцифровке разрешения уже не хватает.

Для очистки перекрытия спектра с центральным пучком можно использовать адаптивный частотный фильтр, описанный в разделе 4.6 (рис. 6.26).





Для расшифровки голограмм, полученных при больших углах сдвига, необходимо увеличивать разрешение при оцифровке. Для этого нужно использовать микрообъектив, у которого пространственное разрешение больше.

На рис. 6.27 показаны увеличенные части восстановленного изображения из голограммы, полученной при  $15^\circ$  между опорным и объектным пучками. При оцифровке использовались микрообъективы с восьмикратным и десятикратным увеличением.

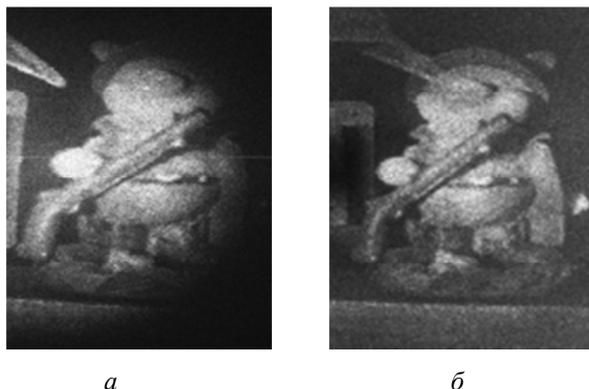


Рис. 6.27. Восстановленные на компьютере увеличенные части изображения при оцифровке с использованием микрообъективов:

*a* – с восьмикратным увеличением; *б* – с десятикратным увеличением

Видно, что правая часть изображения содержит элементы (прорисована левая рука), которые отсутствуют на изображении на левой части рисунка. Эти элементы соответствуют области высоких пространственных частот в голограмме, которые не разрешаются объективом с низким увеличением.

В следующем разделе рассмотрена оптическая схемы для непосредственного ввода голограмм без использования промежуточных фотоносителей.



#### 6.4. РЕГИСТРАЦИЯ И ВОССТАНОВЛЕНИЕ ЦИФРОВЫХ ГОЛОГРАММ

Для использования метода расшифровки, основанного на получении математической голограммы, необходима модификация оптической установки. В схему необходимо добавить устройство для внесения фазового сдвига. В качестве такого устройства использовалось зеркало, закрепленное на пьезокерамике [11, 12].

При оцифровке голограмм, регистрируемых непосредственно с помощью одной матрицы фотоприемников (без использования оптического увеличения с помощью микрообъективов), необходимо согласовывать угол между интерферирующими пучками с размером единичного датчика. В цифровых камерах размер единичного датчика меняется от 1 до 8 мкм. Поэтому угол между интерферирующими пучками не должен превышать  $1...3^\circ$ .

Нами использовалась простая оптическая схема, показанная на рис. 6.28. Для образования интерференционной картины использовался интерферометр Тваймана–Грина. Пучок света от лазера расширяется (3) и попадает на разделительный кубик. Часть пучка попадает на зеркало, закрепленное на пьезокерамике (5). При отражении от этого зеркала формируется объектный пучок. Другая часть пучка попадает на объект. При отражении от объекта (2) формируется опорный пучок. Для выравнивания интенсивностей опорного и объектного пучков при формировании голограммы используется световой фильтр (4).

Интерференция опорного и объектного пучков и формирование голограмм происходит на матрице фотокамеры (9).

На рис. 6.29 показана фотография голографической установки.

При формировании цифровой голограммы необходимо, во-первых, обезопасить установку от внешних вибраций, поскольку это препятствует формированию устойчивой интерференции. Для этого используется antivибрационный стол. Во-вторых, в зоне интерференции влияние постороннего освещения должно быть минимальным, камера должна фиксировать только интерференцию монохроматического когерентного света лазера. Для этого эксперимент должен проводиться без освещения естественными или искусственными источниками света. В-третьих, интенсивность света, отраженного от объекта, должна



примерно совпадать с интенсивностью опорного пучка. В противном случае произойдет перекрытие по яркости одного пучка другим, что не позволит корректно сформировать голограмму. В-четвертых, камера должна фиксировать только интерференционную картину. Для этого над оптическим кубом размещают специальный щит, препятствующий попаданию прямых лучей света от объекта на матрицу цифровой камеры.

В качестве объекта использовался юбилейный серебряный значок Новосибирского государственного технического университета (рис. 6.30).

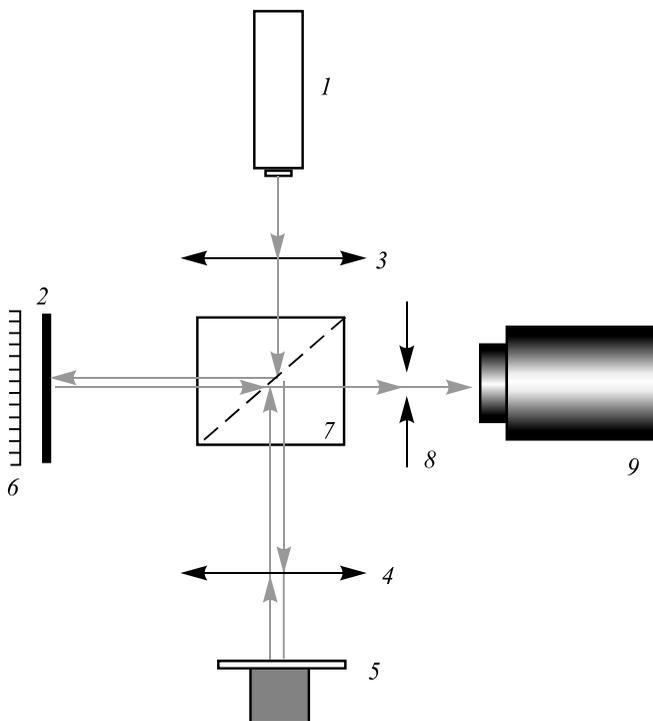


Рис. 6.28. Схема записи цифровой голограммы:

1 – лазер; 2 – объект; 3 – расширитель пучка; 4 – свето-  
фильтр для выравнивания уровня интенсивности; 5 – опор-  
ное зеркало, закрепленное на пьезокерамике; 6 – зеркало для  
юстировки установки; 7 – делитель светового пучка;  
8 – диафрагма; 9 – камера

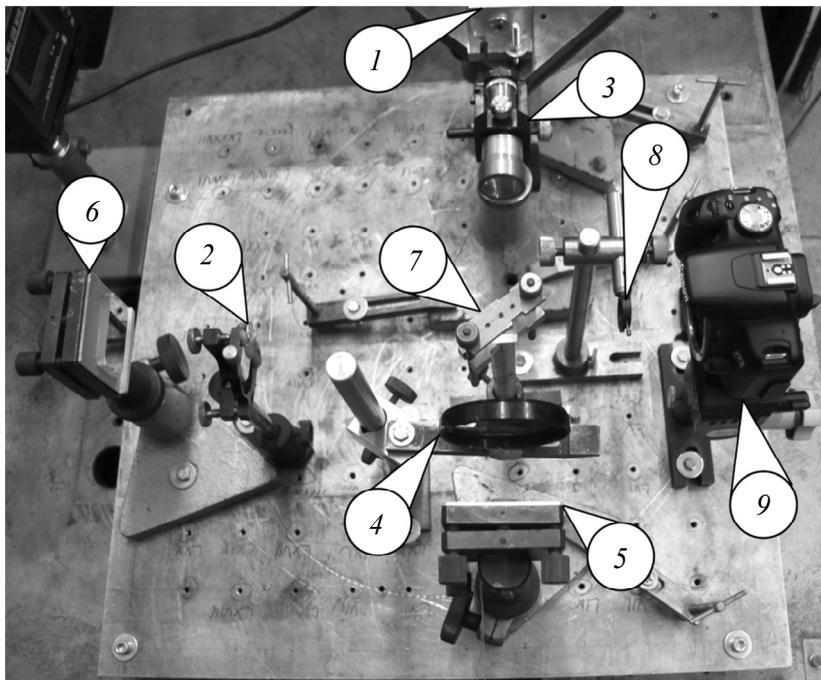


Рис. 6.29. Голографическая установка:

1 – лазер; 2 – объект; 3 – расширитель пучка; 4 – светофильтр для выравнивания уровня интенсивности; 5 – опорное зеркало, закрепленное на пьезокерамике; 6 – зеркало для юстировки установки; 7 – делитель светового пучка; 8 – диафрагма; 9 – камера



Рис. 6.30. Объект для записи голограммы



На рис. 6.31 показаны результаты интерференции между опорным и объектным пучками при изменении фазового угла сдвига:

$$\delta_1 = 0^\circ, \delta_2 = 90^\circ, \delta_3 = 180^\circ, \delta_4 = 270^\circ .$$

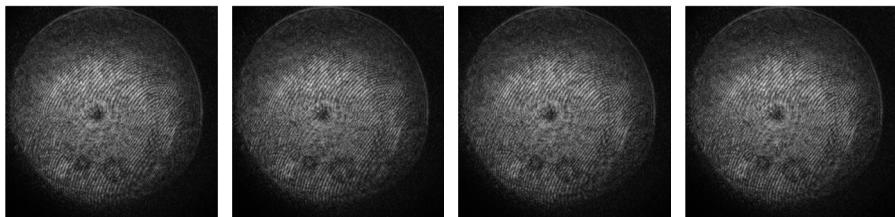


Рис. 6.31. Интерференционные картины при изменении фазового угла сдвига:  $\delta_1 = 0^\circ, \delta_2 = 90^\circ, \delta_3 = 180^\circ, \delta_4 = 270^\circ$

По этим картинам определялись фазовое распределение и амплитуда по полю с помощью выражений, приведенных в разделах 5.1 и 5.2. Затем формировалась математическая голограмма  $G(x, y)$  с помощью выражения (5.1). На рис. 6.32 показаны амплитуда  $a_p(x, y)$  и фаза  $\varphi_p(x, y)$  математической голограммы  $G(x, y)$ .

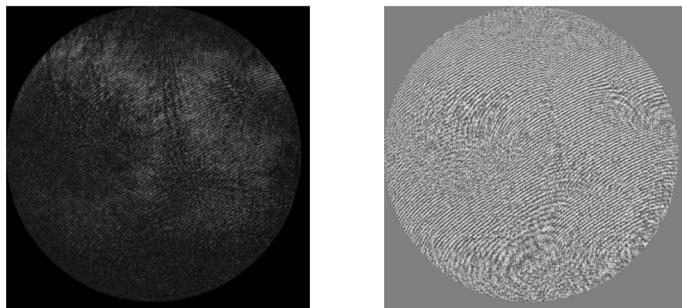


Рис. 6.32. Амплитуда и фаза математической голограммы

Изображение восстанавливалось по математической голограмме с помощью преобразования Френеля. Размер объекта 7 мм, расстояние до объекта 135 мм. Результат восстановления показан на рис. 6.33.

Шумы на рис. 6.33 вызваны отклонением опорного пучка от плоского. Повышением качества оптических элементов этот фактор можно устранить.



*Рис. 6.33.* Результат восстановления действительного изображения из математической голограммы

К сожалению, для диффузных объектов информацию о фазе определить по одному состоянию объекта не удастся. Это связано с фазовыми переходами через  $2\pi$  в соседних точках. Для выделения фазовой информации необходимо использовать методы голографической интерферометрии.

В этой главе показаны примеры, иллюстрирующие возможности цифровой обработки голограмм. Для реальных измерительных систем оптические схемы должны быть усовершенствованы для получения необходимой информации об объекте и увеличении качества восстановленных изображений.

### БИБЛИОГРАФИЧЕСКИЙ СПИСОК К ГЛАВЕ 6

1. Голография. Методы и аппаратура / В. М. Гинзбург, Б. М. Степанов, Е. А. Антонов [и др.]. – Москва: Советское радио, 1974. – 376 с.
2. Patent US 3971065. Color Imaging Array / В. Е. Bayer. – Appl. date 05.03.1975; publ. date 20.07.1976.
3. Видеокамеры сверхвысокого разрешения с технологией pixel shift [Электронный ресурс]. – Режим доступа: <https://www.cameraiq.ru/catalog/series/368-VN-videokamery-vysokogo-razresheniia-pixel-shift> (дата обращения: 14.06.2018).
4. Романова Г. Э. Конспект лекций по курсу «Компьютерные методы контроля оптики» / Г. Э. Романова, М. А. Парпин, Д. А. Серегин. – Санкт-Петербург: НИУ ИТМО, 2012. – 188 с.



5. Микроскоп стереоскопический МБС-10. Руководство по эксплуатации АЦ3.850.005 РЭ [Электронный ресурс]. – Режим доступа: [http://www.mbs10.ru/pdf/MBS-10\\_manual.pdf](http://www.mbs10.ru/pdf/MBS-10_manual.pdf) (дата обращения: 14.06.2018).

6. Модификация оптических микроскопов / В. И. Гужов, Д. В. Ильтимиров, Д. С. Хайдуков, О. В. Чернов, С. Л. Полубинский // Автоматика и программная инженерия. – 2016. – № 2 (16). – С. 71–76.

7. 8MTF Моторизованный двухосный линейный транслятор [Электронный ресурс] // Vicon Standa : web-сайт. – Режим доступа: [http://vicon-se.ru/catalog/motorizovannye\\_pozicionery\\_i\\_kontrollery\\_nanopozicionery/linejnye\\_translyatory/motorizovannyj\\_dvuhosnyj\\_linejnyj\\_translyator](http://vicon-se.ru/catalog/motorizovannye_pozicionery_i_kontrollery_nanopozicionery/linejnye_translyatory/motorizovannyj_dvuhosnyj_linejnyj_translyator) (дата обращения: 14.06.2018).

8. Modbus over serial line specification and implementation guide. V1.02 [Electronic resource]. – Hopkinton, MA, 2006. – 44 p. – Available at: [http://www.modbus.org/docs/Modbus\\_over\\_serial\\_line\\_V1\\_02.pdf](http://www.modbus.org/docs/Modbus_over_serial_line_V1_02.pdf) (accessed: 14.06.2018). – 44 с.

9. Контроллеры шагового двигателя OSM17R, OSM-42RA [Электронный ресурс]: полное описание и руководство по эксплуатации: версия 25-0413 / Компания Онитекс. – Санкт-Петербург, 2013. – 44 с. Режим доступа: [http://onitex.ru/files/Documentation/OSM/datasheet\\_OSM17R\\_OSM42R.pdf](http://onitex.ru/files/Documentation/OSM/datasheet_OSM17R_OSM42R.pdf) (дата обращения: 14.06.2018).

10. Восстановление изображений из реальных голограмм, зафиксированных на фотопластинках / В. И. Гужов, Е. Н. Денежкин, О. В. Чернов, Н. С. Зарубин // Автоматика и программная инженерия. – 2017. – № 1 (19). – С. 76–80.

11. Восстановление фазовой информации на основе методов пошагового фазового сдвига при малых углах между интерферирующими пучками / В. И. Гужов, С. П. Ильиных, С. В. Хайбулин // Автометрия. – 2017. – Т. 53, № 3. – С. 101–106.

12. *Guzhov V.I.* Phase information recovery based on the methods of phase shifting interferometry with small angles between interfering beams / V. I. Guzhov, S. P. Il'Inykh, S. V. Khaibullin // *Optoelectronics, Instrumentation and Data Processing*. – 2017. – Vol. 53, iss. 3. – P. 288–293. – doi: 10.3103/S875669901703013X.

## ЗАКЛЮЧЕНИЕ

Голографические измерительные системы имеют широкий диапазон применений. Такие системы могут использоваться при измерении смещений и деформаций, определении механических и термических свойств материалов, таких как модуль Юнга, коэффициент Пуассона, коэффициент термального расширения.

Однако большие надежды, которые возлагались на голографические системы в конце XX в., в полной мере не оправдались. Необходимость хранения промежуточных голограмм на фотографических пластинах сдерживало использование измерительных голографических систем в условиях реального производства. Попытки использования для регистрации голограмм других сред не дали существенных практических результатов.

Качество цифровых средств регистрации и возможности вычислительной техники увеличиваются, а их стоимость уменьшается с каждым годом, поэтому открываются новые перспективы в разработке голографических измерительных систем. Прослеживая тенденцию развития когерентно-оптических систем измерения, контроля и диагностики на ближайшее время, можно заключить, что они состоят на 10 % из оптики, на 10 % из механики, на 20 % из электроники и на 60 % из алгоритмического и математического обеспечения. Исследование новых алгоритмических подходов, методов расшифровки и математических методов обработки и получения голографических изображений приобретает в настоящее время, особенно важное значение.

Использование небольших углов сдвига между интерферирующими пучками значительно снижает требования к разрешению регистрирующих сред, что позволяет использовать матрицы фотоприемников при регистрации голограмм.

В этой книге приведены новые алгоритмы восстановления фазовой информации из цифровых голограмм, основанные на сдвиге опорного



пучка на заданные значения. Использование этих алгоритмов позволяет эффективно устранять влияние центрального пучка и эффективно выделять не только информацию об амплитуде, но и о фазе исследуемого волнового поля, отраженного от объекта.

Рассмотрен новый алгоритм для численного расчета дискретного преобразования Френеля, свободный от фазовых искажений при любых расстояниях. Исследована область возможного применения этого преобразования. Показано, что в области оптических частот условия, необходимые для выполнения дифракции Фраунгофера, могут быть менее жесткими, чем классические ограничения. Можно использовать дискретные преобразования Френеля и Фурье для определения амплитуды и фазы волнового фронта даже в тех случаях, когда расстояние до объекта сравнимо с размерами объекта и голограммы.

Полученные результаты позволяют надеяться на эффективное использование метода цифровой голографии в экспериментальных голографических измерительных системах для изучения объектов с диффузной поверхностью и в системах для анализа напряженно-деформированного состояния объектов методами цифровой голографической интерферометрии, которые будут рассмотрены в нашей следующей публикации.

Автор благодарен многим людям за их помощь и поддержку при написании книги. Методы и алгоритмы, рассмотренные в этом издании, не были бы реализованы без участия многих сотрудников лаборатории голографии, которая была создана в НЭТИ (НГТУ) в 70-е гг. прошлого века. Автор обязан руководителям лаборатории А.Г. Козачку и А.В. Логинову, которые оказали поддержку в начальный период работы и стимулировали интерес к голографии. В разработку оптических схем голографических установок большой вклад внесли сотрудники лаборатории Е.Н. Денежкин, В.Г. Нечаев и О.В. Чернов. Программные средства расшифровки цифровых голограмм были разработаны при участии Д.С. Хайдукова, Р.А. Кузнецова и Н.С. Зарубина. Необходимо отдельно отметить С.П. Ильиных, который оказывал постоянную помощь на протяжении многих лет, а также большую работу профессора Ю.В. Чугуя при рецензировании рукописи, полезные замечания которого позволили исправить значительное число ошибок.

## ОГЛАВЛЕНИЕ

Введение .....	7
Библиографический список к введению .....	12
<b>Глава 1. Основы голографии</b> .....	<b>15</b>
Библиографический список к главе 1 .....	34
<b>Глава 2. Математическая модель голографического процесса</b> .....	<b>35</b>
2.1. Основные этапы голографического процесса .....	35
2.2. Виды электромагнитных волн .....	37
2.3. Математическая модель распространения света в пространстве .....	57
2.4. Интерференция волн .....	73
2.5. Регистрация голограмм .....	80
2.6. Восстановление волнового фронта из голограммы .....	83
2.7. Необходимое пространственное разрешение при регистрации голограммы .....	87
Библиографический список к главе 2 .....	92
<b>Глава 3. Дискретизация голограмм</b> .....	<b>95</b>
3.1. Основные свойства преобразования Фурье .....	96
3.2. Основные свойства обобщенных функций .....	100
3.3. Свойства некоторых обобщенных функций, необходимых для дискретизации .....	107
3.4. Определение свертки и произведения обобщенной функции с обычной функцией .....	119
3.5. Идеальная дискретизация .....	123
3.6. Необходимые условия дискретизации функций (теорема Котельникова) .....	125
3.7. Восстановление непрерывных сигналов по дискретным отчетам .....	131
3.8. Дискретизация сигналов в реальных системах .....	133
Библиографический список к главе 3 .....	142



<b>Глава 4. Компьютерное моделирование голографического процесса</b> ....	143
4.1. Представление преобразования Фурье в дискретной форме .....	144
4.1.1. Свойства дискретного преобразования Фурье .....	147
4.1.2. Быстрое преобразование Фурье .....	152
4.1.3. Алгоритм Блустейна (Bluestein) для вычисления дискретного преобразования Фурье над произвольным количеством точек .....	158
4.1.4. Быстрое преобразование Фурье для четного количества точек .....	163
4.1.5. Матричное дискретное преобразование Фурье .....	166
4.2. Представление преобразования Френеля в дискретной форме .....	167
4.2.1. Особенности расчета преобразования Френеля при моделировании голограмм .....	173
4.2.2. Зоны действия дискретных преобразований Френеля и Фурье .....	177
4.3. Компьютерное моделирование плоских и сферических волновых полей .....	183
4.4. Компьютерное моделирование голографического процесса в области Фраунгофера .....	186
4.5. Устранение нулевого порядка дифракции при диффузном рассеивании поверхности объектов .....	191
4.6. Методы подавления нулевого порядка дифракции на основе цифровой фильтрации голограммы .....	195
4.7. Методы подавления нулевого порядка дифракции на основе образования реплик сигнала .....	198
Библиографический список к главе 4 .....	201
<b>Глава 5. Восстановление амплитуды и фазы объекта по математической голограмме</b> .....	203
5.1. Определение разности фаз методом пошагового фазового сдвига (PSI) .....	203
5.2. Определение амплитуды объектного пучка методом пошагового фазового сдвига .....	214
5.3. Восстановление комплексной амплитуды волнового фронта, отраженного от объекта по математической голограмме .....	223
5.4. Влияние числа уровней квантования на качество восстановления комплексной амплитуды объектной волны .....	227
Библиографический список к главе 5 .....	229



---

<b>Глава 6. Использование методов компьютерной голографии для экспериментальных исследований</b> .....	231
6.1. Использование коммерческих фотокамер для ввода голограмм.....	232
6.2. Устройства для внесения управляемого фазового сдвига.....	236
6.3. Расшифровка аналоговых голограмм, зафиксированных на фотопластинке.....	243
6.3.1. Модификация оптического микроскопа для ввода голограмм.....	243
6.3.2. Расшифровка аналоговых голограмм.....	248
6.3.3. Расшифровка голограмм при разных углах между интерферирующими полосами.....	253
6.4. Регистрация и восстановление цифровых голограмм.....	256
Библиографический список к главе 6.....	260
Заключение.....	262

## CONTENTS

Introduction.....	7
References to the Introduction .....	12
<b>Chapter 1. Holography Principles</b> .....	15
References to Chapter 1 .....	34
<b>Chapter 2. A mathematical model of the holographic process</b> .....	35
2.1. The main stages of the holographic process .....	35
2.2. Types of electromagnetic waves .....	37
2.3. A mathematical model of light propagation in space .....	57
2.4. Wave interference .....	73
2.5. Hologram recording .....	80
2.6. Wave front reconstruction from holograms .....	83
2.7. The necessary spatial resolution in hologram recording .....	87
References to Chapter 2 .....	92
<b>Chapter 3. Hologram Sampling</b> .....	95
3.1. The main properties of the Fourier transform .....	96
3.2. The main properties of generalized functions .....	100
3.3. Properties of some generalized functions necessary in sampling .....	107
3.4. Determination of the convolution and product of generalized and ordinary functions .....	119
3.5. Ideal sampling .....	123
3.6. Conditions necessary for function sampling ( Sampling theorem) .....	125
3.7. Restoration of continuous signals by samples .....	131
3.8. Signal sampling in real systems .....	133
References to Chapter 3 .....	142



<b>Chapter 4. Computer Modeling of the Holographic Process</b> .....	143
4.1. Representation of the discrete Fourier transform .....	144
4.1.1. Properties of the discrete Fourier transform .....	147
4.1.2. Fast Fourier transform (FFT) .....	152
4.1.3. The Bluestein algorithm to compute the discrete Fourier transform for an arbitrary number of points .....	158
4.1.4. FFT for an even number of points .....	163
4.1.5. The matrix discrete Fourier transform .....	166
4.2. Representation of the discrete Fresnel transform .....	167
4.2.1. Peculiarities of the Fresnel transform computation in modeling holograms .....	173
4.2.2. The effective area of the Fresnel and Fourier transforms .....	177
4.3. Computer modeling of planar and spherical wave fields .....	183
4.4. Computer modeling of the holographic process in the Fraunhofer area .....	186
4.5. Elimination of the diffraction zero order in diffuse scattering of the objects surface .....	191
4.6. Methods of the diffraction zero order suppression based on the hologram digital filtration .....	195
4.7. Methods of the diffraction zero order suppression based on the formation of signal replicas .....	198
References to Chapter 4 .....	201
<b>Chapter 5. Recovery of an Object's Amplitude and Phase by a Mathematical Hologram</b> .....	203
5.1. Determination of the phase difference by the PSI method .....	203
5.2. Determination of an object's beam amplitude by the PSI method .....	214
5.3. Recovery of a complex amplitude of the wave front reflected from an object by a mathematical hologram .....	223
5.4. An effect of the number of quantizing levels on the recovery quality of a complex amplitude of an object wave .....	227
References to Chapter 5 .....	229
<b>Chapter 6. Application of Computer Holography Methods in Experimental investigations</b> .....	231
6.1. The use of commercial cameras to enter holograms .....	232
6.2. Devices for entering a controlled phase shift .....	236




---

6.3. Interpretation of analogue holograms recorded on a photo plate .....	243
6.3.1. Modification of an optical microscope for entering holograms .....	243
6.3.2. Interpretation of analogue holograms .....	248
6.3.3. Interpretation of holograms at different angles between interfering bands .....	253
6.4. Recording and reconstruction of digital holograms .....	256
References to Chapter 6 .....	260
Conclusions .....	262

НАУЧНОЕ ИЗДАНИЕ

**Гужов Владимир Иванович**

**КОМПЬЮТЕРНАЯ ГОЛОГРАФИЯ**

**Монография**

Редактор *Е.Н. Николаева*  
Выпускающий редактор *И.П. Брованова*  
Художественный редактор *А.В. Ладыжская*  
Корректор *Л.Н. Кишит*  
Компьютерная верстка *Л.А. Веселовская*

Подписано в печать 20.11.2018  
Формат 60 × 90 1/16. Бумага офсетная  
Уч.-изд. л. 17,0. Печ. л. 17,0.  
Тираж 3000 экз. (1-й з-д – 1–70 экз.)  
Изд. № 118. Заказ № 1563.

Налоговая льгота – Общероссийский классификатор продукции  
Издание соответствует коду 95 3000 ОК 005-93 (ОКП)

Издательство Новосибирского государственного  
технического университета  
630073, г. Новосибирск, пр. К. Маркса, 20  
Тел. (383) 346-31-87  
E-mail: office@publish.nstu.ru

Отпечатано в типографии  
Новосибирского государственного технического университета  
630073, г. Новосибирск, пр. К. Маркса, 20