

Пусть функция $f(x, y)$ подвергается дискретизации с помощью прямоугольной сетки с шагом выборки вдоль осей x и y соответственно Δx и Δy . Тогда дискретное представление функции $f(x, y)$ удобно записать с помощью гребенчатой функции

$$\text{comb}(x) = \sum_{n=-\infty}^{\infty} \delta(x-n) \quad (1.64)$$

в виде

$$f_s(x, y) = \text{comb}\left(\frac{x}{\Delta x}\right) \text{comb}\left(\frac{y}{\Delta y}\right) f(x, y). \quad (1.65)$$

Согласно теореме свертки (см. приложение 1) фурье-образ или спектр выборочной функции

$$F_s(\xi, \eta) = \mathcal{F}\left\{\text{comb}\left(\frac{x}{\Delta x}\right) \text{comb}\left(\frac{y}{\Delta y}\right)\right\} * F(\xi, \eta). \quad (1.66)$$

Используя свойства преобразования Фурье и δ -функции, выражение (1.66) приведем к виду

$$F_s(\xi, \eta) = \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} F[(\xi - n/\Delta x); (\eta - m/\Delta y)]. \quad (1.67)$$

Отсюда следует, что спектр выборочной функции $f_s(x, y)$ представляет собой сумму спектров функции $f(x, y)$, построенных возле каждой точки с координатами $[n/(\Delta x), m/(\Delta y)]$ в плоскости $\xi\eta$. Расстояния между соседними точками вдоль осей ξ и η равны $1/(\Delta x)$ и $1/(\Delta y)$. Смещенные друг относительно друга спектры, входящие в сумму (1.67), в общем случае могут перекрываться.

Допустим, что функция $f(x, y)$ имеет ограниченный спектр, определяемый областью D , которая располагается внутри прямоугольника

$$|\xi| \leq \xi_{\max}; \quad |\eta| \leq \eta_{\max} \quad (1.68)$$

так, что

$$F(\xi, \eta) = \begin{cases} \neq 0 & \text{внутри области;} \\ \equiv 0 & \text{вне этой области.} \end{cases}$$

Пусть Δx и Δy выбраны так, что выполняются следующие условия

$$\frac{1}{\Delta x} \geq 2\xi_{\max}; \quad \frac{1}{\Delta y} \geq 2\eta_{\max}. \quad (1.69)$$

В этом случае спектр $F_s(\xi, \eta)$ представляет собой совокупность смещенных друг относительно друга, но не перекрывающихся спектров $F[(\xi - n/\Delta x), (\eta - m/\Delta y)]$ (рис. 1.9).

Таким образом, если Δx и Δy достаточно малы, то расстояния $1/(\Delta x)$ и $1/(\Delta y)$ между различными спектральными областями будут достаточно большими для того, чтобы соседние спектры не перекрывались. Макси-

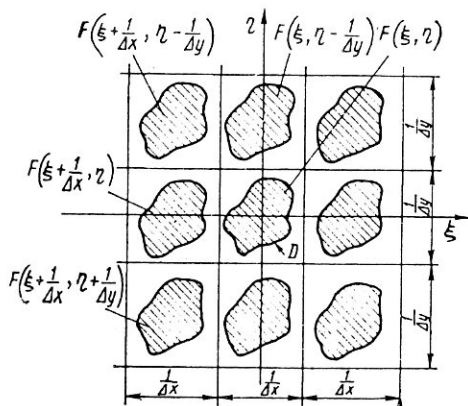


Рис. 1.9. Структура фурье-образа выборочных значений функции, восстанавливаемой с помощью теоремы отсчетов

моугольным отверстием со сторонами $2\xi_{\max} \times 2\eta_{\max}$; его передаточная функция

$$H(\xi, \eta) = \text{rect}\left(\frac{\xi}{2\xi_{\max}}\right) \text{rect}\left(\frac{\eta}{2\eta_{\max}}\right). \quad (1.71)$$

Следовательно, спектр функции $f(x, y)$ можно записать в виде

$$F(\xi, \eta) = F_s(\xi, \eta) H(\xi, \eta). \quad (1.72)$$

Обратное фурье-преобразование произведения в правой части данного соотношения определим с помощью операции свертки в координатной области. Учитывая, что $f_s(x, y) = \mathcal{F}^{-1}\{F_s(\xi, \eta)\}$ определяется выражением (1.65), получим

$$f(x, y) = \left[\text{comb}\left(\frac{x}{\Delta x}\right) \text{comb}\left(\frac{y}{\Delta y}\right) f(x, y) \right] * h(x, y), \quad (1.73)$$

где $h(x, y) = \mathcal{F}^{-1}\{H(\xi, \eta)\}$.

Приняв во внимание (1.71), найдем (см. приложение II):

$$\begin{aligned} h(x, y) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \text{rect}\left(\frac{\xi}{2\xi_{\max}}\right) \text{rect}\left(\frac{\eta}{2\eta_{\max}}\right) \exp[i2\pi(\xi x + \eta y)] d\xi d\eta = \\ &= 4\xi_{\max} \eta_{\max} \text{sinc}[2\xi_{\max}(x - n\Delta x)] \text{sinc}[2\eta_{\max}(y - m\Delta y)]. \end{aligned} \quad (1.74)$$

малые значения шагов дискретизации выборочной сетки, удовлетворяющие этому условию, определяются из равенств

$$\Delta x = \frac{1}{2\xi_{\max}}; \quad \Delta y = \frac{1}{2\eta_{\max}}. \quad (1.70)$$

Для восстановления спектра $F(\xi, \eta)$ достаточно пропустить фурье-образ (1.67) через фильтр, который прозрачен только для слагаемого с номером $n = 0, m = 0$. Такой фильтр представляет собой непрозрачный экран с пря-

Поскольку $\text{somb} \left(\frac{x}{\Delta x} \right) \text{somb} \left(\frac{y}{\Delta y} \right) f(x, y) = \Delta x \Delta y \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} f(n\Delta x, m\Delta y) \delta(x - n\Delta x, y - m\Delta y)$, то с учетом (1.74) свертку можно преобразовать к виду

$$f(x, y) = 4\xi_{\max} \eta_{\max} \Delta x \Delta y \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} f(n\Delta x, m\Delta x) \times \\ \times \text{sinc} [2\xi_{\max} (x - n\Delta x)] \text{sinc} [2\eta_{\max} (y - m\Delta y)].$$

Если вместо Δx и Δy подставить их максимально допустимые значения из (1.70), то последнее выражение можно представить в виде

$$f(x, y) = \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} f \left(\frac{n}{2\xi_{\max}}, \frac{m}{2\eta_{\max}} \right) \text{sinc} \left[2\xi_{\max} \left(x - \frac{n}{2\xi_{\max}} \right) \right] \times \\ \times \text{sinc} \left[2\eta_{\max} \left(y - \frac{m}{2\eta_{\max}} \right) \right]. \quad (1.75)$$

Это соотношение является математическим выражением теоремы Котельникова — Шеннона. Теорема может быть сформулирована так: если функция $f(x, y)$ имеет ограниченный спектр, определяемый областью $|\xi| \leq \xi_{\max}$; $|\eta| \leq \eta_{\max}$, то она может быть восстановлена путем интерполяции sinc-функциями по выборочным значениям (отсчетам), взятым с шагами $\Delta x = 1/(2\xi_{\max})$; $\Delta y = 1/(2\eta_{\max})$.

Таким образом, если двумерный оптический сигнал описывается непрерывной функцией с ограниченным спектром, что имеет место для большинства практических задач, то она может быть восстановлена по дискретным отсчетам, взятым с определенными шагами.

§ 1.7. Дискретное преобразование Фурье

Рассмотрим одномерное преобразование Фурье. Фурье-образ одномерной функции

$$F(\xi) = \int_{-\infty}^{\infty} f(x) \exp(-i2\pi\xi x) dx. \quad (1.76)$$

При анализе оптических сигналов с помощью ЭВМ, как правило, вместо непрерывных сигналов располагают набором их дискретных значений (отсчетов). Пусть функция $f(x)$ имеет ограниченный спектр в интервале $-\xi_{\max}$ и ξ_{\max} , задана отсчетами в последовательных точках, которые отстоят друг от друга на расстоянии $\Delta x = 0,5 \xi_{\max}$. Тогда, согласно теореме отсчетов (1.75), эта функция может быть представлена в виде следующего ряда:

$$f(x) = \sum_{m=-\infty}^{\infty} f \left(\frac{m}{2\xi_{\max}} \right) \text{sinc} \left[2\xi_{\max} \left(x - \frac{m}{2\xi_{\max}} \right) \right]. \quad (1.77)$$

Проанализируем получаемую форму фурье-образа $F(\xi)$, для чего выражение (1.77) подставим в (1.76). В результате получим

$$\begin{aligned}
 F(\xi) &= \int_{-\infty}^{\infty} \sum_{m=-\infty}^{\infty} f\left(\frac{m}{2\xi_{\max}}\right) \operatorname{sinc}\left[2\xi_{\max}\left(x - \frac{m}{2\xi_{\max}}\right)\right] \times \\
 &\times \exp(-i2\pi\xi x) dx = \sum_{m=-\infty}^{\infty} f\left(\frac{m}{2\xi_{\max}}\right) \int_{-\infty}^{\infty} \operatorname{sinc}\left[2\xi_{\max}\left(x - \frac{m}{2\xi_{\max}}\right)\right] \times \\
 &\times \exp(-i2\pi\xi x) dx = \frac{1}{2\xi_{\max}} \sum_{m=-\infty}^{\infty} f\left(\frac{m}{2\xi_{\max}}\right) \operatorname{rect}\left[\frac{\xi}{2\xi_{\max}}\right] \times \\
 &\times \exp\left(-i2\pi\xi \frac{m}{2\xi_{\max}}\right) = \frac{1}{2\xi_{\max}} \operatorname{rect}\left(\frac{\xi}{2\xi_{\max}}\right) \sum_{m=-\infty}^{\infty} f\left(\frac{m}{2\xi_{\max}}\right) \times \\
 &\times \exp\left(-i2\pi\xi \frac{m}{2\xi_{\max}}\right). \quad (1.78)
 \end{aligned}$$

Выражение (1.78) определяет фурье-образ исходной функции, заданной дискретными отсчетами. Действительно из (1.78) следует, что в интервале $-\xi_{\max} \leq \xi \leq \xi_{\max}$

$$F(\xi) = \frac{1}{2\xi_{\max}} \sum_{m=-\infty}^{\infty} f\left(\frac{m}{2\xi_{\max}}\right) \exp\left(-i2\pi\xi \frac{m}{2\xi_{\max}}\right); \quad (1.79)$$

для всех остальных значений ξ $F(\xi) = 0$.

Допустим, что функция $f(x)$ имеет не только ограниченный спектр, но и сама ограничена интервалом $-x_{\max} \leq x \leq x_{\max}$. Такое допущение неверно, так как функция, имеющая ограниченный спектр, не может одновременно быть ограниченной в пространстве. Однако при больших значениях произведения $x_{\max}\xi_{\max}$ погрешность, возникающая из-за этого допущения при вычислении фурье-образа функции, оказывается пренебрежимо малой. Таким образом, можно считать, что $f(x)$ ограничена в пространстве интервалом $[-x_{\max}, x_{\max}]$. Тогда из теоремы отсчетов следует, что спектр $F(\xi)$ функции $f(x)$ может быть восстановлен по отсчетам, взятым с шагом $\Delta\xi = 1/(2x_{\max})$, с помощью соотношения, аналогичного (1.77). Следовательно, достаточно вычислить значения фурье-образа в выборочных точках, что позволяет сократить объемы необходимой памяти и машинного времени, для расчета и хранения спектра сигнала. Подставив в (1.79) вместо аргумента ξ его дискретные значения в точках выборки $\xi_n = n\Delta\xi = n/(2x_{\max})$, получим выражение

$$F\left(\frac{n}{2x_{\max}}\right) = \frac{1}{2\xi_{\max}} \sum_{m=-M/2}^{M/2} f\left(\frac{m}{2\xi_{\max}}\right) \exp\left(i2\pi \frac{mn}{\xi_{\max} x_{\max}}\right), \quad (1.80)$$

называемое *дискретным преобразованием Фурье*, которое определяет значения отсчетов фурье-образа $F(\xi)$.

Число отсчетов функции $f(x)$ в интервале $-x_{\max} \leq x \leq x_{\max}$ $M = 2x_{\max}/(\Delta x) = 4x_{\max} \xi_{\max}$. Аналогично число дискретных отсчетов фурье-образа в интервале частот $-\xi_{\max} \leq \xi \leq \xi_{\max}$, $N = (2\xi_{\max})/(\Delta \xi) = 4x_{\max} \xi_{\max}$. Следовательно, $M = N = 4x_{\max} \xi_{\max}$, т. е. для исходной функции $f(x)$, как и для ее фурье-образа, требуются одинаковые объемы выборки.

Таким же образом можно найти двумерное дискретное преобразование:

$$F\left(\frac{n}{2x_{\max}}, \frac{k}{2y_{\max}}\right) = \frac{1}{4\xi_{\max} \eta_{\max}} \sum_{m=-M/2}^{M/2} \sum_{l=-N/2}^{N/2} f\left(\frac{m}{2\xi_{\max}}, \frac{l}{2\eta_{\max}}\right) \times \exp\left[-i2\pi\left(\frac{nm}{4\xi_{\max} \eta_{\max}} + \frac{kl}{4\xi_{\max} \eta_{\max}}\right)\right]. \quad (1.81)$$

Дадим математическое определение дискретного преобразования Фурье. Пусть функция $f(x)$ задана последовательностью значений f_0, f_1, \dots, f_{N-1} в N равноотстоящих друг от друга точках числовой прямой. Дискретным преобразованием Фурье (ДПФ) этой последовательности называют последовательность F_0, F_1, \dots, F_{N-1} , определяемую суммой вида

$$F_p = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} f(n) \exp\left(-i2\pi \frac{pn}{N}\right); \quad p=0, 1, 2, \dots, N-1. \quad (1.82)$$

ДПФ обладает свойством обратимости, т. е.

$$f_n = \frac{1}{\sqrt{N}} \sum_{p=0}^{N-1} F_p \exp\left(i2\pi \frac{pn}{N}\right); \quad n=0, 1, 2, \dots, N-1. \quad (1.83)$$

Выражения (1.82) и (1.83) представляют одномерные ДПФ: первое характеризует прямое ДПФ, второе — обратное ДПФ. Они отличаются знаком экспоненты.

Аналогичным образом определяются двумерные ДПФ. Если двумерная функция $f(x, y)$ задана матрицей выборочных значений $f(n, m)$ ($n = 0, 1, 2, \dots, N-1$; $m = 0, 1, 2, \dots, M-1$), то соответствующее ей прямое двумерное ДПФ имеет вид

$$F_{p,q} = \frac{1}{\sqrt{NM}} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} f(n, m) \exp\left[-i2\pi\left(\frac{pn}{N} + \frac{qm}{M}\right)\right]; \quad (1.84)$$

$p=0, 1, 2, \dots, N-1, q=0, 1, 2, \dots, M-1$.

Преобразование (1.84) также обратимо:

$$f_{n,m} = \frac{1}{\sqrt{NM}} \sum_{p=0}^{N-1} \sum_{q=0}^{M-1} F_{p,q} \exp\left[i2\pi\left(\frac{pn}{N} + \frac{qm}{M}\right)\right]; \quad (1.85)$$

$n=0, 1, 2, \dots, N-1; m=0, 1, 2, \dots, M-1$.

Для вычисления значений F_{pq} по формуле (1.84) даже при небольших значениях N и M потребуется большое количество машинного времени. Например, чтобы рассчитать ДПФ матрицы $N \times M = 1024 \times 1024$, необходимо около 6 ч машинного времени на ЭВМ, имеющей быстродействие около 10^6 оп/с. Поэтому ДПФ приобрело практическое значение только после изобретения так называемых быстрых алгоритмов (БПФ). Один из таких алгоритмов позволил сократить число операций в $M/\log_2 M$ раз ($M = N$). Для нашего случая ($M = N = 1024$) при использовании БПФ число операций сократится примерно в 100 раз, что приведет к снижению затрат машинного времени приблизительно до 4 мин. В настоящее время существует множество вариантов быстрых алгоритмов вычисления ДПФ [5].

ГЛАВА 2

ОПТИЧЕСКАЯ ГОЛОГРАФИЯ

§ 2.1. Физические принципы голографии

Основы голографии были заложены в 1948 г. английским физиком Д. Габором. Желая усовершенствовать электронный микроскоп, Д. Габор предложил регистрировать информацию не только об амплитудах, но и о фазах электронных волн путем наложения на одну из них когерентной волны. Однако отсутствие мощных источников когерентного света не позволило ему получить качественные голографические изображения. В 1971 г. за открытие метода голографии профессору Д. Габору была присуждена Нобелевская премия по физике. Второе рождение голография пережила в 1962 и 1963 гг., когда американские физики Э. Лейт и Ю. Упатниекс применили в качестве источника света лазер и разработали схему с наклонным опорным пучком. Советский ученый Ю. Н. Денисюк впервые предложил и осуществил запись голограмм в трехмерной среде [6].

Голография представляет собой способ записи и восстановления волнового поля, основанный на регистрации интерференционной картины (голограммы), которая образована волной (рис. 2.1, а), отраженной предметом, освещаемым источником света (объектная волна), и когерентной с ней волной, идущей непосредственно от источника света (опорная волна). Голограмма, освещенная опорной волной, создает такое же амплитудно-фазовое пространственное распределение волнового поля, как и при записи объектной волны (рис. 2.1, б).

Голографический метод применим ко всем волнам: электронным, рентгеновским, световым, акустическим и сейсмическим микроволнам, при условии, что они когерентны для создания требуемых интерференционных картин. Этот метод наиболее пригоден в оптическом диапазоне электромагнитного спектра. После создания лазеров оптическая голография стала быстро развиваться [2,6].

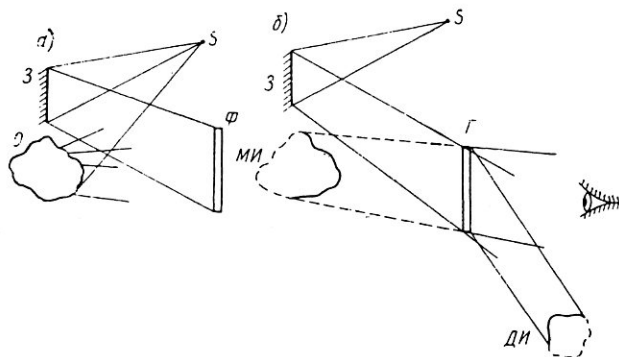


Рис. 2.1. Схемы получения голограммы (а) и восстановления волнового фронта (б):

S источник света; O — объект; Φ фотопластинка; $Г$ голограмма; $З$ — зеркало; $МИ$ — мнимое изображение; $ДИ$ — действительное изображение

Рассмотрение основ голографии начнем с волновых явлений интерференции и дифракции.

Интерференция. В голографии обычно имеют дело с интерференцией объектной и опорной волн. Простейший тип синусоидальных волн приведен на рис. 2.2. Расстояние между двумя гребнями волны называют *длиной волны* λ . Для световых волн λ характеризует один цвет. Волны A и B имеют одинаковые длины волн и амплитуды. Волна B опережает волну A по фазе на $\Delta\varphi$. Если волна B распространяется относительно волны A , то разность фаз этих волн изменяется во времени.

Для голографии интересны два случая взаимодействия волн: 1) волны находятся в фазе; 2) волны находятся в противофазе. При взаимодействии волн A и B получают новую волну, амплитуда которой является результатом сложения амплитуды волны A с амплитудой волны B в каждой точке. Если волны находятся в фазе, то при их наложении возникает усиливающая интерференция и результирующая волна имеет большую интенсивность, чем каждая из волн в отдельности (рис. 2.3, а). В том случае, когда волны находятся в противофазе, наблюдается ослабляющая интерференция; волны полностью гасят друг друга, амплитуда результирующей волны равна нулю (рис. 2.3, б). Процесс сложения нескольких волн, в результате которого образуется новая волна, называется *интерференцией*.

Рассмотрим интерференцию двух плоских волн, спроецированных на экран (рис. 2.4). Интерференция возникает в различных точках вдоль экрана и зависит от разности фаз волн, приходящих в данную точку. В одних точках гребни волны B совпадают с гребнями волны A , в других — гребень волны A встречается с впадиной волны B . Соответственно на экране наблюдаются области большой и малой интенсивности. Пространственное распределение интенсивности излучения, по-

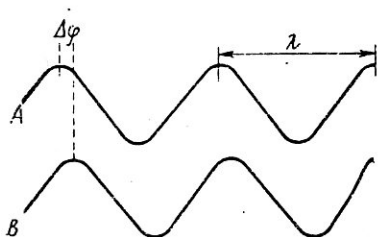


Рис. 2.2. Синусоидальные волны, сдвинутые по фазе

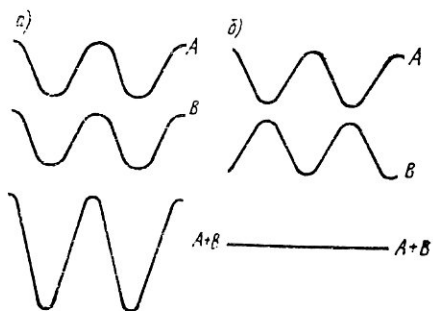


Рис. 2.3. Сложение двух синусоидальных волн

лучаемое при этом, называют *интерференционной картиной*. При интерференции двух волн $U_r = U_r \exp [i\varphi_r(x, y)]$; $U_0 = U_0 \exp [i\varphi_0(x, y)]$ результирующая интенсивность излучения

$$I = |U_r + U_0|^2 = |U_r|^2 + |U_0|^2 + U_r^* U_0 + U_r U_0^*$$

или

$$I = |U_r|^2 + |U_0|^2 + 2U_r U_0 \cos(\varphi_2 - \varphi_1). \quad (2.1)$$

Таким образом, интенсивность излучения в любой точке интерференционной картины, образованной двумя волнами, является суммой интенсивностей отдельных волн и интерференционной составляющей, содержащей информацию о разности фаз.

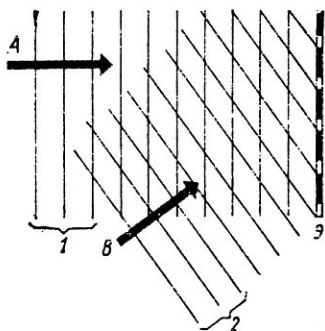


Рис. 2.4. Интерференция двух плоских волн на экране (вид сверху):

1 — гребни волны A; 2 — гребни волны B; Э — белый экран

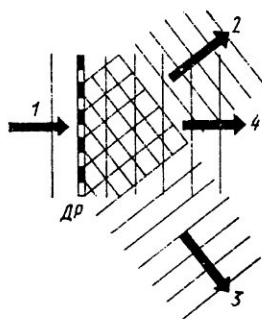


Рис. 2.5. Получение плоских волн методом дифракции:

1 — опорная волна; 2 — восстановленная плоская волна; 3 — сопряженная восстановленная плоская волна; 4 — часть плоской волны, прошедшая через решетку

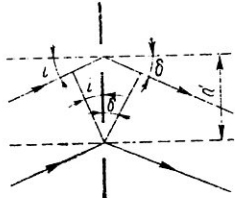


Рис. 2.6. К пояснению дифракции света в плоской дифракционной решетке

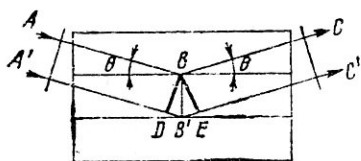


Рис. 2.7. К пояснению дифракции света в объемной дифракционной решетке

Дифракция. Дифракция — явление, заключающееся в искривлении световых лучей при встрече с препятствием, оптическое пропускание или отражение которого значительно изменяется на расстояниях, близких к длине волны света. На маленьких отверстиях в экранах дифрагируют все световые лучи, прошедшие через них, на больших — лучи, прошедшие вблизи краев отверстия. Световые лучи, проходящие через центральную часть большого отверстия, не испытывают дифракции.

Если на пути волнового фронта находится несколько небольших препятствий, он в результате дифракции изменяется таким образом, что световые лучи, распространяющиеся за препятствием, имеют качественно новый волновой фронт. Дифракцию можно рассматривать как механизм, посредством которого создают новый волновой фронт света.

Простейшее устройство, которое таким путем формирует новый волновой фронт, называют *дифракционной решеткой*. При падении лазерного пучка на дифракционную решетку ДР (рис. 2.5) часть его проходит через решетку прямо, а часть загибается; в результате формируются два новых пучка, выходящих из решетки под некоторым углом к исходному. Если исходный лазерный пучок имеет, например, плоский волновой фронт, то и два образовавшихся пучка также обладают плоскими волновыми фронтами. Так как ДР на рис. 2.5 аналогична картине на рис. 2.4, то ее можно рассматривать как простейший пример голограммы.

Голограмма сама по себе есть объект, вызывающий дифракцию. На рис. 2.6 представлена в разрезе плоская дифракционная решетка. Для плоской волны, падающей на решетку, условие синфазности дифрагированных пучков, ведущее к их взаимному усилению, является уравнением решетки:

$$d(\sin i + \sin \delta) = \lambda, \quad (2.2)$$

где d — постоянная решетки; i — угол падения; δ — угол дифракции. Как видно из рис. 2.6, при выполнении условия (2.2) под углом δ образуется главный максимум дифрагированной плоской волны. Кроме того, возможны отрицательные и более высокие порядки дифракции.

На рис. 2.7 показана (в разрезе) объемная дифракционная решетка, содержащая периодически расположенные рассеивающие плоскости, освещенные плоской волной. Здесь справедлив тот же самый принцип: интенсивность излучения максимальна в направлении, в котором происходит синфазное сложение световых волн, рассеянных последовательными плоскостями. Условие образования главного максимума дифрагированной плоской волны

$$2d \sin \theta = \lambda \quad (2.3)$$

называют *законом Брэгга*. Английский проф. У. Брэгг предположил, что дифракция рентгеновских лучей в кристалле обусловлена отражением падающей волны от кристаллических плоскостей. Максимум дифракции возникает, когда углы, образованные падающим и отраженным лучами с кристаллической плоскостью, равны (рис. 2.7), причем угол θ удовлетворяет условию (2.3).

Сравнение (2.2) и (2.3) обнаруживает, что последнее выражение накладывает более жесткие условия на наблюдение максимума дифракции. Для объемной решетки выбор угла падения определяет длину волны и угол дифракции. Для плоских решеток допускается произвольный выбор этих величин.

Образование голограмм. Для того чтобы получить голограмму, когерентный свет лазера необходимо разделить на две волны, одна из которых освещает объект, а другая является опорной (см. рис. 2.1, а). Опорная волна направляется таким образом, чтобы она пересекалась со световой волной, прошедшей через объект или отраженной от него. Если обе волны абсолютно когерентны, то интерференционная картина образуется во всем объеме, в котором они перекрываются. Фотопластинка, помещенная в область перекрытия, регистрирует интерференционную картину. После соответствующей химической обработки фотопластинки получают голограмму [2].

Восстановление волнового фронта. Если голограмма освещается опорной волной, то часть дифрагировавшего на ней света вновь воссоздает волновой фронт, который при регистрации шел от объекта. Восстановленная волна исходит из голограммы точно так же, как первоначальная объектная волна. Наблюдатель, видящий волну, идентичную объектной волне, воспринимает ее как бы исходящей от мнимого изображения предмета, расположенного там, где ранее находился предмет (см. рис. 2.1, б). В двухмерной голограмме одновременно восстанавливается сопряженная волна, образующая искаженное действительное изображение предмета.

Таким образом, голографический процесс включает в себя процесс получения голограммы и ее восстановления.

Что же дает этот трудоемкий процесс, кроме копии волнового фронта, которой располагали с самого начала? Ответы на этот вопрос будут получены при рассмотрении применения голографии в различных областях науки и производства. В большинстве случаев интерес представляет не восстановление волнового фронта, а некоторый объект

(возможно трехмерный), сквозь который волна проходит или от которого она отражается, распространяясь по направлению к плоскости записи. Возможность записи информации об объекте в виде голограммы, а не прямого изображения дает множество преимуществ, особенно в смысле характера и качества информации, которая может быть восстановлена.

Основное уравнение голографий. Идея голографической записи состоит в том, чтобы кроме волны, которая нас интересует (объектная волна), регистрировать дополнительно опорную волну, когерентную с объектной волной, таким образом, чтобы в результате их интерференции осуществлялось кодирование фазы φ . Введем следующие обозначения: $U_0 = U_0 \exp[i\varphi_1(x, y)]$ — объектная волна, $U_r = U_r \times \exp[i\varphi_2(x, y)]$ — опорная волна. Возьмем для записи пластинку с линейным откликом на интенсивность падающей волны. Она регистрирует распределение интенсивности, описываемое выражением $I = |U_0 + U_r|^2 = |U_0|^2 + |U_r|^2 + U_0 U_r^* + U_0^* U_r$. Ее пропускание после обычной фотохимической обработки будет пропорционально интенсивности I .

Осветим эту пластинку опорной волной. Волна, прошедшая сквозь пластинку,

$$U_r I = (|U_0|^2 + |U_r|^2) U_r + U_0 |U_r|^2 + U_0^* U_r^2. \quad (2.4)$$

Это выражение основное для голографии. В нем три слагаемых:

1) $(|U_0|^2 + |U_r|^2) U_r$ — опорная волна U_r , амплитуда которой модулирована коэффициентом $(|U_0|^2 + |U_r|^2)$;

2) $|U_r|^2 U_0 + U_r^2 U_0 \exp(i\varphi_0)$ — объектная волна U_0 или модулированная коэффициентом U_r^2 ;

3) $U_0^* U_r^2 + U_r^2 U_0 \exp(i(2\varphi_2 - \varphi_0))$ — волна, комплексно-сопряженная с объектной. Это слагаемое появляется из-за квадратичного детектирования. Оно несет информацию, близкую к информации об объекте, но отличается от объектной волны обратной фазой.

Таким образом, получение фазы и амплитуды объектной волны на этапе восстановления голограммы сопровождается восстановлением двух паразитных волн: опорной, которую ослабили, но не подавили совсем, и сопряженной.

§ 2.2. Основные типы голограмм

Двумерная и трехмерная голограммы. Важность третьего измерения, т. е. глубины голографической записи, установил еще на раннем этапе развития голографии Ю. Н. Денисюк. За большой вклад в становление и развитие голографии Ю. Н. Денисюк был удостоен Ленинской (1970) и Государственной (1982) премий СССР, а также премии им. Д. Габора (1983).

Голограмму можно рассматривать как тонкую дифракционную структуру лишь в том случае, когда ее оптическая толщина меньше длины волны. На практике, однако, на характеристики голограммы

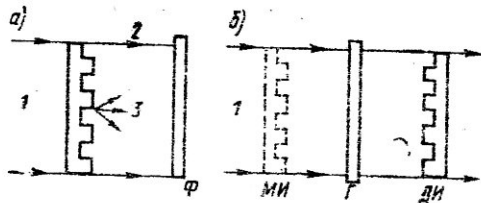


Рис. 2.8. Схемы получения (а) и восстановления (б) голограммы Габора:

1 — освещающий пучок; 2 — опорная волна; 3 — объектная волна; Ф — фотопластика; МИ — мнимое восстановленное изображение объекта; Г — голограмма; ДИ — действительное комплексно-сопряженное изображение объекта

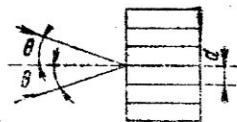


Рис. 2.9. Интерференционная картина, зарегистрированная в объеме светочувствительной эмульсии

влияет не отношение ее толщины к длине волны, а соотношение между толщиной и периодом самых тонких полос, записанных на голограмме. Если период тончайших полос больше толщины, то голограмма является двумерной, если же период полос меньше толщины, то — трехмерной [6].

К двумерным относятся голограммы Габора, Лейта и Упатниекса, имеющие период полос d дифракционной структуры больше ее толщины h ($d \gg h$).

Д. Габор, в распоряжении которого были только источники света с очень малой длиной когерентности, смешал объектную и опорную волны (рис. 2.8, а). Голограммы Габора имеют существенный недостаток: паразитные волны (опорная и сопряженная) восстанавливаются в том же направлении, что и объектная. В результате очень яркий однородный фон и наличие сопряженного расфокусированного изображения объекта мешают наблюдателю видеть объект (рис. 2.8, б). Устранить этот недостаток позволила схема записи голограмм Лейта, использующая наклонную опорную волну (см. рис. 2.1). При восстановлении волнового фронта наклонной опорной волной наблюдатель может видеть мнимое изображение объекта.

К трехмерным относятся голограммы Денисюка, имеющие период полос дифракционной структуры d меньше ее толщины h ($d \ll h$).

Рис. 2.9 поясняет образование полос в объеме светочувствительной эмульсии дифракционной решетки для плоских объектной и опорной волн. Если такую решетку осветить исходной опорной волной, то каждый луч до выхода из эмульсии последовательно рассеивается от большего числа периодически расположенных поверхностей максимальной плотности. Чтобы амплитуда результирующей дифрагированной волны была максимальной, волны, рассеянные последовательными слоями, должны быть синфазны. Для этого необходимо, чтобы выполнялось определенное соотношение между длиной волны λ , углом θ , который составляет освещающая голограмму волна с рассеивающими поверхностями, и расстоянием между этими поверхностями. Эту взаимосвязь устанавливает закон Брэгга (2.3), который можно записать в виде

$$2d \sin \theta = \lambda_0/n, \quad (2.5)$$

где λ_0 — длина волны в воздухе; n — средний показатель преломления светочувствительной среды; θ — угол, который освещающая и дифрагированная волны составляют с рассеивающими слоями. Закон Брэгга определяет угол падения, если длина волны и расстояние между слоями заданы. Если же угол падения и постоянная решетки выбираются независимо, то закон Брэгга определяет длину волны. Таким образом, трехмерные голограммы, свойства которых описываются законом Брэгга, являются селективными относительно освещающего их излучения.

На рис. 2.10, а приведена схема записи трехмерных голограмм Денисюка. При записи когерентное излучение W_S источника S , прошедшее через объем V , заполненный прозрачным светочувствительным материалом, падает на объект O . В результате интерференции отраженной от объекта волны W_0 и опорной волны W_S в пространстве перед объектом возникает стоячая волна, которую в некотором приближении можно представить в виде системы поверхностей пучностей d_1, d_2, d_3 и т. д., которым соответствует максимальная интенсивность поля. Воздействуя на светочувствительный материал, поле стоячей волны поразному засвечивает различные части заполненного им объема. В этом объеме после экспозиции и соответствующего проявления образуется трехмерная голограмма, подобная трехмерной дифракционной решетке. Зафиксированные голограммой поверхности пучностей стоячей волны d'_1, d'_2, d'_3 и т. д.

(рис. 2.10, б) можно рассматривать как систему зеркал, поверхности которых имеют сложную форму.

Трехмерную голограмму можно восстанавливать источником света, имеющим сплошной спектр излучения, так как голограмма является селективной по отношению к освещающему ее излучению [см. (2.5)]. При восстановлении точно воспроизводятся практически все параметры зарегистрированного на ней волнового поля объекта — амплитуда, фаза и спектральный состав. В частности, из сплошного спектра источника S' выбирается и отражается излучение той длины волны, которая совпадает с длиной волны излу-

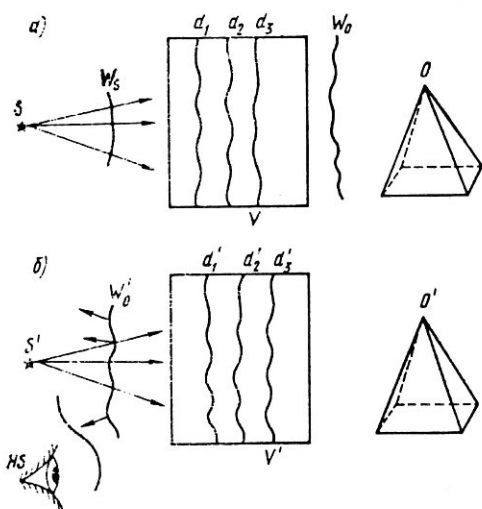


Рис. 2.10. Схемы получения (а) и восстановления (б) трехмерной голограммы:

S' — источник излучения со сплошным спектром; V' — трехмерная голограмма; W'_0 — волна излучения, восстановленного трехмерной голограммой; O' — восстановленное голограммой объемное изображение объекта

чения, с помощью которого была получена голограмма во время записи. Наблюдатель HS , воспринимающий восстановленную волну, не может отличить ее от первоначальной объектной и, следовательно, видит объемное изображение O' объекта в цвете, соответствующем длине волны излучения, освещавшего объект при записи (рис. 2.10, б).

Голографическое изображение обладает эффектами параллакса, благодаря чему удается увидеть объект в различных положениях (рис. 2.11).

Поскольку при освещении такой голограммы опорной волной восстановленная объектная волна распространяется ей навстречу, такие голограммы относят к голограммам отражательного типа.

Амплитудные и фазовые голограммы. Голограммная структура может быть зарегистрирована светочувствительным материалом путем: 1) изменения коэффициента пропускания или отражения света; 2) изменения коэффициента преломления или толщины (рельефа). Голограммы первого типа называют *амплитудными*, они осуществляют преимущественно амплитудную модуляцию восстанавливающей волны. Голограммы второго типа производят преимущественно фазовую модуляцию восстанавливающей волны и поэтому их называют *фазовыми*. Часто одновременно осуществляется фазовая и амплитудная модуляция. Например, обычная фотопластинка регистрирует голограммную структуру в виде изменения показателя преломления и рельефа. Такие голограммы называют *амплитудно-фазовыми*.

Голограммы Фраунгофера, Френеля и Фурье [6]. Структура голограммы зависит также от способа формирования объектной и опорной волн. Объект освещается когерентной световой волной; рассеянная им световая волна, несущая информацию о предмете, падает на фотопластинку, освещаемую опорной волной. В зависимости от взаимного

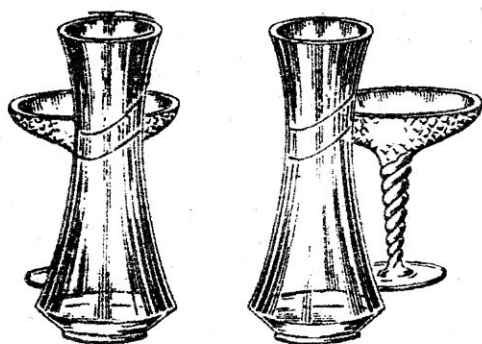


Рис. 2.11. Восстановленное изображение трехмерного объекта под разными углами зрения

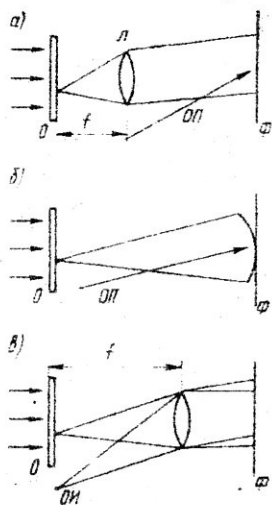


Рис. 2.12. Схемы получения голограмм различных типов

расположения объекта и пластинки, а также от наличия оптических элементов между ними связь между амплитудно-фазовыми распределениями объектной волны в плоскостях голограммы и объекта различна.

Когда объект O находится достаточно далеко от фотопластинки Φ либо в фокусе линзы L (рис. 2.12, а), то каждая его точка посылает на пластинку параллельную световую волну. При этом связь между амплитудно-фазовыми распределениями объектной волны в плоскости голограммы и в плоскости объекта характеризуется преобразованием Фурье. Комплексную амплитуду объектной волны на пластинке называют *фурье-образом* объекта. Такую голограмму называют *голограммой Фраунгофера*. Можно определить голограмму Фраунгофера как голограмму, полученную в голографическом поле образованной с помощью объектной волны, формируемой в области дифракции Фраунгофера.

Если комплексные амплитуды объектной и опорной волн являются фурье-образами объекта и опорного источника, то полученную голограмму называют *голограммой Фурье* (рис. 2.12, б).

Голограмма Френеля образуется в том случае, когда каждая точка объекта посылает на фотопластинку сферическую волну (рис. 2.12, в). При получении голограмм Френеля фотопластинку помещают в области дифракции ближнего поля (области дифракции Френеля) на произвольном расстоянии от источника опорной волны. Это наиболее простой способ регистрации, так как он позволяет получать голограмму и восстанавливать волновой фронт без использования линз или каких-либо других оптических устройств.

§ 2.3. Фурье-голограммы

Голограммы Фурье являются одним из наиболее распространенных видов голограмм и наиболее перспективными для применений в области вычислительной техники. Такой вид голограмм можно получить следующим образом (рис. 2.13). Плоский объект (транспарант) освещается когерентной коллимированной волной и фокусируется линзой L . В фокусной плоскости линзы устанавливается фотопластинка, на которую кроме объектной волны, прошедшей транспарант, направляется опорная волна. Фурье-голограмма образуется как результат интерференции фурье-образа транспаранта с опорной волной.

Если транспарант, описываемый функцией $t(x, y)$, установлен в передней фокальной плоскости линзы, то в ее задней фокальной плоскости фурье-образ транспаранта

$$T(\xi, \eta) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} t(x, y) \exp[-2\pi i(\xi x + \eta y)] dx dy. \quad (2.6)$$

Линза осуществляет частотный анализ функции транспаранта $t(x, y)$, т. е. если разложить эту функцию по гармоникам пространственных частот ξ, η , то каждая точка в фокальной плоскости линзы

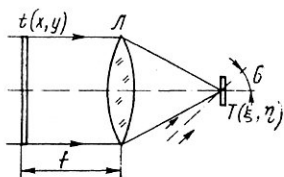


Рис. 2.13. Оптическая схема получения фурье-голограммы транспаранта

отвечает своей пространственной частоте с соответствующей фазой. Для отсчета фаз гармонических составляющих вводят вспомогательную волну, падающую под углом θ на фурье-плоскость. Угол θ определяет несущую частоту записи, модулируемую частотами транспаранта. Таким образом, фурье-голограмма регистрирует спектр пространственных частот входного сигнала $t(x, y)$.

На основании свойств фурье-преобразования можно объяснить достоинства фурье-голограмм. Поскольку информация о каждой точке транспаранта распределена по всей пространственно-частотной области, т. е. содержится в любой точке фурье-образа транспаранта, то потеря части спектра не приводит к потере всего образа, а лишь несущественно снижает разрешение и яркость его изображения при восстановлении. Этим объясняется высокая надежность и помехоустойчивость получения информации в виде фурье-голограмм.

Такое получение информации обладает инвариантностью к сдвигу. Если транспарант $t(x, y)$ во входной плоскости сдвинуть на (x_0, y_0) , то согласно теореме смещения

$$\mathcal{F}[t(x - x_0, y - y_0)] = T(\xi, \eta) \exp[-2\pi i(\xi x_0 + \eta y_0)], \quad (2.7)$$

т. е. сдвиг транспаранта на входе по координатам приводит к появлению постоянных фазовых множителей $\exp(-2\pi i \xi x_0)$ и $\exp(-2\pi i \eta y_0)$, которые не влияют на положение голограммы. Инвариантностью к сдвигу обладает и голограмма, ибо при восстановлении ее осуществляется обратное фурье-преобразование: $t'(x, y) = \mathcal{F}^{-1}\{T(\xi, \eta)\}$. Так как восстановление производят с помощью квадратичных детекторов (фотопластики, фотопленки и т. п.), голограмма формирует изображение, характеризующееся квадратом модуля амплитуды волны; поэтому фазовые члены исчезают. Это свойство фурье-преобразования является причиной нечувствительности восстановленного изображения к небольшим сдвигам голограммы.

Поскольку фурье-преобразующая линза собирает проходящую световую волну в своей фокальной плоскости на небольшой площади, то в фурье-голограмме информация регистрируется с максимальной плотностью хранения. Такие голограммы имеют меньшие aberrации, чем остальные типы голограмм, что является основным их достоинством, так как aberrации в голографии, так же как в оптике, являются обычным явлением, ограничивающим качество восстановленного изображения. Более того, можно получить плоские фурье-голограммы, свободные от aberrаций [6].

Фурье-голограммы, записанные специальным образом, могут также служить в качестве операционных фильтров в когерентных оптических системах обработки информации, в частности, в системах распознава-

ния и идентификации образов, основанных на принципах пространственно-частотной фильтрации (см. гл.6). Благодаря указанным свойствам фурье-голограммы находят широкое применение в системах голографической памяти и оптической обработки информации.

§ 2.4. Ассоциативные свойства голограмм

Зарегистрируем на фурье-голограмме два изображения, которые обозначим символами h и g . Для восстановления голограммы используем излучение, идущее от одного из изображений. Если излучение поступает от изображения h , то на выходе получают изображение g , и наоборот. Отметим, что одно из изображений, например h , может быть частью, фрагментом изображения g . Система осуществляет восстановление одного из изображений, если восстанавливаемое изображение смещается параллельно самому себе во входной плоскости, так как для фурье-голограмм имеет место сдвиговый инвариант (свойство ассоциативности).

На этом принципе может быть построен голографический корреляционный транслятор. Допустим, что изображением h служило слово *holos*, а изображением g — его перевод «целый»; голограмма записывалась в фурье-плоскости. Если голограмму освещать словом «целый» или частью его, то в выходной плоскости получим «*holos*».

Математически описать процесс ассоциативного опознавания изображений можно следующим образом. Если транспарант с изображениями h и g освещается плоской волной единичной амплитуды, то в фокальной плоскости объектива функции изображений $h(x, y)$ и $g(h, y)$ имеют комплексную амплитуду в виде суммы фурье-образов $H(\xi, \eta) + G(\xi, \eta)$. Фотопластинка при линейной регистрации имеет пропускающие

$$HH^* + GG^* + HG^* + H^*G. \quad (2.8)$$

Если осветить голограмму волной с комплексной амплитудой H , то четвертый член выражения H^*G характеризует изображение g , так как комплексная амплитуда волны в фурье-плоскости HH^*G ; в выходной плоскости волна описывается обратным фурье-преобразованием:

$$F^{-1}(HH^*G). \quad (2.9)$$

Используя свойства фурье-преобразования, выражение (2.9) запишем в виде

$$F^{-1}(HH^*G) = F^{-1}\{F[h^*(x_\xi, y_\xi)] * h(x_\xi, y_\xi)\} F[g(x_\xi, y_\xi)] \sim \\ \sim [h^*(x_\xi, y_\xi)] * h(x_\xi, y_\xi) \times g(x_\xi, y_\xi). \quad (2.10)$$

т. е. комплексная амплитуда выходного сигнала представляет собой свертку $g(x_\xi, y_\xi)$ с функцией автокорреляции изображения $h(x_\xi, y_\xi)$. Если автокорреляционная функция h представляет собой δ -функцию, то получают изображение объекта g в плоскости (x_ξ, y_ξ) .

При этом смещение изображения $n(x_1, y_1)$ во входной плоскости не нарушает процесса восстановления сигнала g , так как имеет место инвариантность сдвига.

Ассоциативные свойства голограмм наиболее полно выражены в трехмерных голограммах.

Рассмотренные ассоциативные свойства голограмм можно использовать при поиске информации по смысловым признакам в больших массивах данных голографического запоминающего устройства. Таким образом, по выбранному ключевому слову можно подобрать необходимую информацию.

§ 2.5. Синтез голограмм с помощью ЭВМ

В последнее время значительный интерес проявляется к цифровой голографии, т. е. к получению и восстановлению голограмм с помощью ЭВМ. Цифровые методы синтеза голограмм и восстановления изображений позволяют расширить возможности оптической голографии. Голограммы, синтезированные с помощью ЭВМ, находят широкое применение в таких областях, как оптическая обработка информации, распознавание образов, трехмерная индикация цифровых данных, проектирование физически еще несуществующих объектов, моделирование голографических процессов и т. п. Трудно переоценить полезность использования синтезированных голограмм для восстановления изображений в акустической и СВЧ-голографии. Машинный синтез голограмм свободен от ряда серьезных ограничений, которые имеют место при оптической записи и обусловлены частичной когерентностью света, несовершенством оптических элементов, влиянием вибраций и других возмущений. Часто синтез с помощью ЭВМ оказывается единственным способом получения голограммы с заданными свойствами. Главное же достоинство синтезированной голограммы заключается в том, что она является эффективным средством для преобразования цифровой информации в оптическую. Благодаря этому можно создать уникальные по производительности гибридные вычислительные системы, включающие в свой состав цифровой электронный и оптический процессоры и сочетающие гибкость и универсальность ЭВМ с огромной производительностью, присущей когерентным оптическим вычислительным машинам.

Следует отметить важную особенность синтезированных голограмм, заключающуюся в возможности получения оптических волновых фронтов от объектов, физически не существующих, а заданных лишь математическим описанием.

Синтез голограмм с применением ЭВМ осуществляют в четыре этапа. Сначала строят математическую модель голографируемого объекта и рассчитывают распределение комплексных амплитуд света, испущенного им, в плоскости регистрации голограммы. Затем полученное распределение комплексных амплитуд кодируют как действительную неотрицательную функцию, которую называют голографической. Далее

голографическая функция подвергается пространственной дискретизации путем выборки ее значений в узлах некоторой, обычно регулярной, сеточной области. Каждый отсчет состоит из амплитудного и фазового коэффициентов. Амплитуда и фаза каждого отсчета квантуются с использованием ограниченного числа уровней. Квантованные значения голографической функции отображаются выходным устройством ЭВМ (фототелеграф, графопостроитель, устройство отображения на основе ЭЛТ) в виде картины, представляющей увеличенное изображение получаемой голограммы. Оптическим уменьшением этой картины до размеров, обеспечивающих дифракцию света, используемого для восстановления изображения объекта, и фотографической записью ее на регистрирующую среду завершается процесс изготовления голограммы. Структурная схема синтеза голограммы показана на рис. 2.14.

Приведем алгоритм расчета голографической функции. Пусть исходный объект описан дискретной комплексной функцией $g(n_1, m_1)$; $n_1 = \overline{1, M}$; $m_1 = \overline{1, Q}$. Наибольший интерес представляет случай, когда голографическая функция является фурье-образом объекта. С помощью фурье-преобразования функции $g(n, m)$ получаем голографическую функцию

$$U_{1H}(x_H, y_H) = \sum_{n=1}^M \sum_{m=1}^Q g(n_1, m_1) \times \exp \left\{ 2\pi i \left[x_H \left(n_1 - \frac{M \cdot 1}{2} \right) + y_H \left(m_1 - \frac{Q \cdot 1}{2} \right) \right] \right\} \quad (2.11)$$

где x_H, y_H — координаты в плоскости голограммы. В дальнейшем мы часто будем обращаться к пространственно-частотным координатам $\xi = x_H / (\lambda f)$ и $\eta = y_H / (\lambda f)$.

Так как на ЭВМ можно рассчитать только конечное число отсчетов, голографическая функция $U_{1H}(x_H, y_H)$ также подвергается дискретизации, шаги которой должны удовлетворять теореме отсчетов Котельникова—Шеннона [см. (1.69)]: $\Delta\xi \leq \frac{1}{M\Delta x_1}$; $\Delta\eta \leq \frac{1}{Q\Delta y_1}$, где $\Delta x_1, \Delta y_1$ — шаги дискретизации в плоскости объекта. Обычно шаги дискретизации берут равными $\Delta v = \Delta\xi = \Delta\eta$. Отсчеты голографической функции в точках дискретизации рассчитывают с помощью алго-

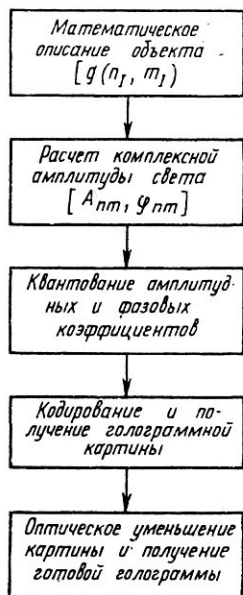
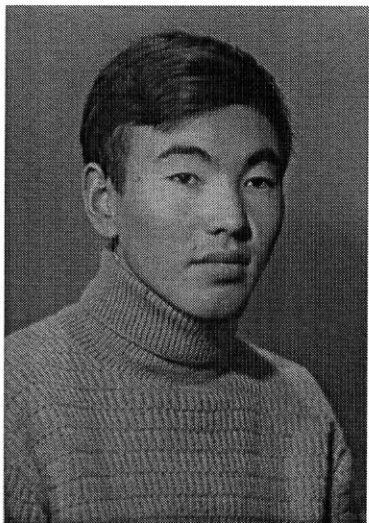


Рис. 2.14. Структурная схема получения синтезированной голограммы с помощью ЭВМ



*А. Акаев, студент 5-го курса кафедры
вычислительной техники ЛИТМО.
1967 год*



*Президент Кыргызской Республики А. А. Акаев на встрече с однокурсниками
в ноябре 2004 года во время пребывания в Alma Mater по случаю присвоения
ему звания почетного доктора Санкт-Петербургского государственного
университета информационных технологий, механики и оптики.
Слева – заведующий кафедрой вычислительной техники СПбГУ ИТМО,
профессор Т. И. Алиев*

ритма БПФ. В результате получают матрицу дискретных значений комплексных амплитуд голографической функции:

$$\left. \begin{aligned} U_{IH}(n\Delta v, m\Delta v) &= A(n\Delta v, m\Delta v) \exp [i\varphi(n\Delta v, m\Delta v)] = \\ &= A_{nm} \exp (i\varphi_{nm}), \quad A_{nm} = |U_{IH}| = \sqrt{a^2 + b^2}; \\ \varphi_{nm} &= \arg (U_{IH}) = \operatorname{arctg} (b/a); \quad a = \operatorname{Re} \{U_{IH}(n, m)\}; \\ b &= \operatorname{Im} \{U_{IH}(n, m)\}. \end{aligned} \right\} \quad (2.12)$$

Синтезированные голограммы делят на амплитудные, фазовые и поляризационные [7].

Впервые метод кодирования волнового фронта для синтеза амплитудных голограмм был предложен А. Ломаном и его сотрудниками. Принцип записи амплитудных голограмм методом Ломана состоит в том, что отсчет $U_{IH}(n\Delta v, m\Delta v)$ голографической функции объекта в узле $(n\Delta v, m\Delta v)$ отображается выходным устройством ЭВМ в виде прозрачной прямоугольной апертуры на непрозрачном фоне (рис. 2.15, а), причем ширина апертуры $S\Delta v$ постоянна для данной голограммы, высота ее $W_{nm}\Delta v$ с некоторым приближением прямо пропорциональна значению амплитуды отсчета A_{nm} , а ее смещение $P_{nm}\Delta v$ в направлении одной из координат относительно точки отсчета пропорционально фазе отсчета. Каждая апертура должна располагаться в пределах квадрата со стороной Δv и геометрическим центром в узле сеточной области (рис. 2.15, б). Это означает, что относительная ширина S , относительная высота апертуры W_{nm} и параметр P_{nm} , характеризующий ее положение, должны удовлетворять условиям

$$W_{nm} \leq 1; \quad |P_{nm}| + 0,5S \leq 1. \quad (2.13)$$

Область, заключенная внутри отдельного квадрата, является конечным элементом голограммы.

Процесс отображения голографической информации в картину осуществляется ЭВМ с помощью выходного устройства, например двухтонового графопостроителя. Полученная таким образом черно-белая картина представляет собой увеличенное изображение искомой голограммы в масштабе, согласованном с параметрами графопостроителя.

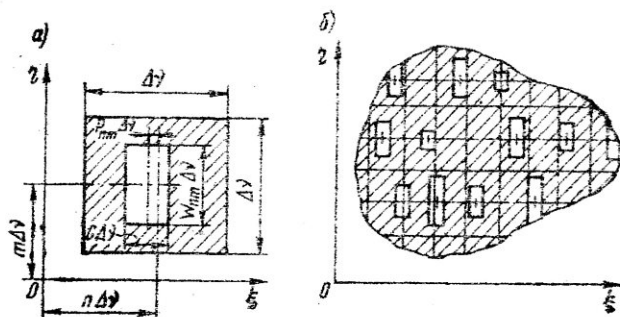


Рис. 2.15. Картина бинарной фурье-голограммы

теля. Амплитудное пропускание синтезированной голограммы является бинарной функцией, т. е. в любой точке голограммы принимает значение, равное нулю либо единице. Поэтому такие голограммы получили название бинарных фурье-голограмм.

Развитием метода Ломана является метод кодирования, заключающийся в следующем [8]. Каждый отсчет голографической функции разбивают на две компоненты следующим образом:

$$\left. \begin{aligned} U_{1H}(n, m) &= A_{nm} \exp(i\varphi_{nm}) = \\ &= \exp(i\varphi_{1nm}) + \exp(i\varphi_{2nm}). \\ \varphi_{1nm} &= \varphi_{nm} + \psi, \quad \varphi_{2nm} = \varphi_{nm} - \psi, \\ \psi &= \arccos(A_{nm}), \quad 0 \leq \psi \leq \frac{\pi}{2} \end{aligned} \right\} (2.14)$$

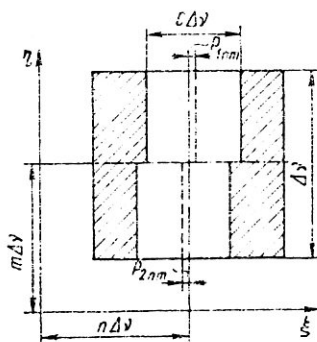


Рис. 2.16. Конечный элемент бинарной голограммы

Эти компоненты кодируют двумя апертурами, как показано на рис. 2.16. Смещения P_{1nm} , P_{2nm} пропорциональны фазам φ_{1nm} , φ_{2nm} отсчета, ширина апертур C равна половине конечного элемента, т. е. $\Delta\xi/2$. Преимущество таких голограмм заключается в том, что их дифракционная эффективность в 2 раза выше, чем у голограмм, полученных методом Ломана, без заметного снижения качества восстановленного изображения. Этот метод можно обобщить на случай, когда каждый отсчет разбивается на n ячеек. При этом дифракционная эффективность увеличивается незначительно, но ухудшается качество восстановленного изображения.

Другой оригинальный метод кодирования комплексной амплитуды был предложен В. Х. Ли [17]. Согласно этому методу каждый отсчет $U_{1H}(n, m)$ раскладывается на четыре компоненты следующим образом:

$$\left. \begin{aligned} U_{1H}(n, m) &= A_{nm} \exp(i\varphi_{nm}) = A_1(n, m) + iA_2(n, m) - \\ &- A_3(n, m) - iA_4(n, m), \end{aligned} \right\} (2.15)$$

где

$$\left. \begin{aligned} A_1(n, m) &= A_{nm} \cos(\varphi_{nm}) \cdot 1 \{\cos(\varphi_{nm})\}; \\ A_2(n, m) &= A_{nm} \sin(\varphi_{nm}) \cdot 1 \{\sin(\varphi_{nm})\}; \\ A_3(n, m) &= A_1(n, m) - A_{nm} \cos(\varphi_{nm}); \\ A_4(n, m) &= A_2(n, m) - A_{nm} \sin(\varphi_{nm}). \end{aligned} \right\} (2.16)$$

Функции $A_k(n, m)$ являются действительными и неотрицательными. Следовательно, каждая из них может быть записана на реальную регистрирующую среду, причем две компоненты из четырех всегда будут равны нулю. Компоненты A_{knm} ($k = \overline{1,4}$) отсчета голографической функции, взятого в узле с координатами $(n\Delta\xi, m\Delta\eta)$, отображаются в позициях с координатами $(\frac{4n-k+1}{4} \Delta\xi, m\Delta\eta)$ или $(n\Delta\xi, \frac{4m-k+1}{4} \Delta\eta)$, как показано на рис. 2.17, а. Величины A_{knm}

отображаются при этом в виде прямоугольных прозрачных апертур на непрозрачном фоне, площади которых пропорциональны A_{hnm} , либо в виде круговых апертур с переменными амплитудными пропусканиями, пропорциональными A_{hnm} (рис. 2.17, б). Первый способ отображения используют в том случае, когда в качестве выходного устройства ЭВМ применяют графопроектор или фототелеграфный аппарат, способный отображать двухтоновые изображения. Полученная голограмма является бинарной. Если голограммная структура воспроизводится на экране ЭЛТ, то предпочтительнее использовать второй способ отображения. При этом получают полутонную голограмму.

Рассмотрим получение фазовых голограмм с помощью ЭВМ. Встречаются случаи, когда комплексная амплитуда объектной волны $U(n, m) = A_{nm} \exp[i\varphi_{nm}]$ в плоскости регистрации практически постоянна по модулю:

$$|A(n, m) - \langle A(n, m) \rangle| \ll \langle A(n, m) \rangle, \quad (2.17)$$

где $\langle A(n, m) \rangle$ — среднее значение амплитуды в области регистрации. В таких случаях изображение объекта может быть восстановлено с использованием фазовой информации φ_{nm} . Как правило, это имеет место, когда голографируемый объект является диффузным или освещен диффузно рассеянным светом.

Таким образом, записав только фазовую информацию об объекте, можно восстановить его трехмерное изображение. Поскольку при регистрации фазово-информационных голограмм амплитудная информация об объекте отбрасывается, качество восстановленного изображения зависит от степени влияния потери амплитудной информации на процесс восстановления искомого изображения.

Вопросы синтеза фазовых голограмм только по фазовой информации, без использования опорной волны впервые были рассмотрены аме-

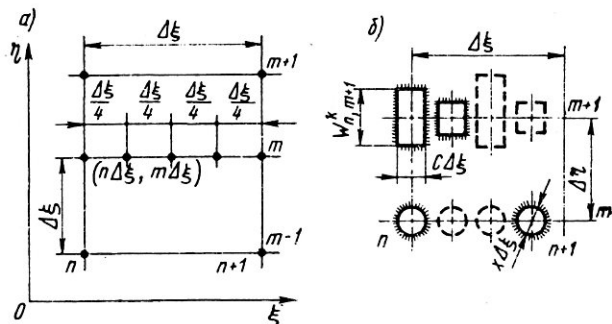


Рис. 2.17. К пояснению метода Ли кодирования волнового фронта:

а — выбор позиций для апертур; б — конечный элемент голограммы

риканскими учеными Л. Леземом, П. Хиршом и Дж. Жорданом [7]. Они же первыми синтезировали такие голограммы и назвали их *киноформами*. Такие голограммы в отличие от других типов при идеальном изготовлении восстанавливают только одно изображение — мнимое или действительное. Это означает, что весь световой поток, дифрагированный киноформом, концентрируется на одном изображении.

Процесс изготовления киноформа выглядит следующим образом. Допустим, что комплексная амплитуда объектной волны в плоскости регистрации описывается только фазовой информацией:

$$U(n, m) \cong \exp[i\varphi_{nm}]; \quad n(m) = \overline{1, N(Q)}. \quad (2.18)$$

На ЭВМ рассчитывают дискретные значения фазы φ_{nm} объектной волны. Полученные значения обрабатывают таким образом, чтобы их отклонения от начальной фазы лежали в интервале от 0 до 2π радиан по всей области выборки, т. е. из каждого значения фазы вычитают значения, кратные 2π радианам. В результате получают двумерный массив, состоящий из дискретных значений фазы $\widehat{\varphi}_{nm}$. Данный массив кодируют массивом значений яркости в многоградационной шкале, который в виде картины отображается на выходное устройство ЭВМ, например на экран ЭЛТ. Полученную картину фотографируют с необходимым уменьшением: конечный фотоснимок отбеливают в дубящем отбеливателе. При отбеливании градации фотографического почернения превращаются в соответствующее распределение значений оптической толщины. Поэтому киноформные элементы являются фазовыми. Такие фазовые элементы работают на пропускание световой волны. Наряду с этим изготавливают киноформы, которые работают на отражение волнового фронта. У них рельефные участки напыляют слоем алюминия или другими светоотражающими материалами. Киноформные элементы широко используют как плоские аналоги оптических элементов, таких, как линзы, объективы и т. п. В частности, на рис. 2.18 представлена амплитудная маска киноформного элемента, полученного с помощью ЭВМ и предназначенного для фокусирования излучения полупроводникового лазера.

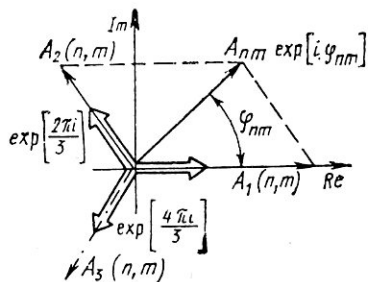


Рис. 2.18. Амплитудная маска киноформного элемента

§ 2.6. Перспективы создания трехмерного голографического дисплея

В настоящее время исследуются различные голографические устройства, предназначенные для воспроизведения трехмерных изображений. С точки зрения использования в вычислительной технике наи-

большой интерес вызывают устройства, основанные на применении синтезированных голограмм. В § 2.5 было подчеркнуто, что одним из свойств таких голограмм является то, что для их получения не требуется физической модели объекта, а достаточно лишь его математического описания. Благодаря этому голографические устройства воспроизведения изображений, основанные на применении синтезированных голограмм, идеальным образом подходят для осуществления трехмерной индикации цифровых данных, полученных в результате расчетов на ЭВМ. Трехмерный голографический дисплей существенно облегчит анализ сложных нестационарных полей (электромагнитных, тепловых, звуковых и др.), волновых процессов и других явлений. Трудно переоценить значение трехмерного индикатора в системе автоматизированного проектирования новых конструкций и других физических объектов. Трехмерный голографический дисплей может значительно повысить эффективность системы человек — ЭВМ.

Рассмотрим принцип действия голографического дисплея для ЭВМ. На первом этапе производится математическое описание объекта. Далее ЭВМ рассчитывает дифракционное поле в требуемой зоне. Полученные дискретные значения комплексной амплитуды дифракционного поля кодируются одним из методов кодирования волнового фронта и выводятся на экран воспроизводящего устройства. Экран представляет собой прозрачную подложку, на которую осаждается тонкий слой обратимого регистрирующего материала. В качестве регистрирующего материала могут быть использованы термопластические или фотохромные, а также электрооптические кристаллы. Воспроизведение голограммной структуры на экране ЭЛТ осуществляется сканирующим электронным пучком или лазерным лучом, управляемым от ЭВМ. Если разрешающая способность воспроизводящего устройства недостаточна, то синтезированная голограммная структура предварительно уменьшается с помощью оптической системы, а затем регистрируется на материале. Зарегистрированная голограмма освещается лазерным лучом. Дифрагированный голограммой световой пучок отфильтровывается с целью выделения той части, которая создает нужное изображение. Полученное изображение увеличивается оптической системой и воспроизводится для наблюдения.

Однако практическое создание трехмерных голографических дисплеев для ЭВМ — это дело будущего. Это обусловлено в основном двумя причинами. Во-первых, реальные трехмерные объекты описываются ансамблем точек, общее число которых достигает 10^5 и более. При этом для получения качественной голограммы требуется 10^6 и более отсчетов дифракционного поля, рассеянного объектом, поэтому быстродействия большинства современных ЭВМ недостаточно для того, чтобы выполнить расчет голографической функции в реальном масштабе времени. Во-вторых, к устройствам воспроизведения и регистрации синтезированных голограммных структур предъявляются высокие требования в отношении разрешающей способности (~ 1000 лин/мм), которые также на сегодня еще не могут быть реализованы.

КОМПОНЕНТЫ ОПТИЧЕСКИХ СИСТЕМ ХРАНЕНИЯ
И ОБРАБОТКИ ИНФОРМАЦИИ

§ 3.1. Лазеры

В настоящее время в оптических системах хранения и обработки информации в основном используют лазеры трех типов: газовые (атомарные и ионные), твердотельные и полупроводниковые. Наиболее совершенными среди них являются атомарные лазеры, например гелий-неоновые лазеры, работающие в режиме непрерывной генерации. Выходная мощность серийных гелий-неоновых лазеров, генерирующих красный свет с длиной волны 0,6328 мкм и имеющих приемлемые габариты, составляет 10—100 мВт. Достоинствами таких лазеров являются стабильность излучения и большой срок службы. Большую выходную мощность излучения могут обеспечить ионные лазеры, среди которых выделяются аргоновые благодаря удачному сочетанию параметров излучения. Выходная мощность таких лазеров составляет 1—10 Вт при длине волны излучения 0,488 и 0,5145 мкм. Однако аргоновые лазеры имеют относительно высокую стоимость и низкую эффективность преобразования электрической мощности в оптическую. Твердотельные лазеры наиболее приспособлены для работы в импульсном режиме. Широкое применение среди таких лазеров находят рубиновые, излучающие красный свет с длиной волны 0,6943 мкм, и лазеры на алюминивно-иттриевом гранате, активированном неодимом, которые дают излучение в инфракрасной области спектра с длиной волны 1,06 мкм. Рубиновые импульсные лазеры в режиме свободной генерации излучают импульсы с энергией порядка 1—10 Дж и длительностью, измеряемой долями миллисекунды. С помощью модулятора добротности длительность импульса может быть сокращена до десятков наносекунд. Имеются, например, одночастотные рубиновые лазеры, генерирующие импульсы длительностью 10—100 нс с энергией, достигающей нескольких джоулей. Выходная мощность излучения при этом составляет сотни мегаватт в импульсе. К недостаткам рубиновых лазеров следует отнести нестабильность излучения, малую частоту повторения импульсов (несколько импульсов в минуту), а также малый срок службы.

Перспективными твердотельными лазерами являются лазеры на алюминивно-иттриевом гранате (АИГ-лазеры), которые могут работать как в импульсном, так и в непрерывном режиме. Для работы в непрерывном режиме генерации требуется охлаждение. Выходная мощность излучения АИГ-лазера в импульсе достигает 1 ГВт при длительности импульса порядка 10 нс и энергии 10 Дж. Частота повторения импульсов у этого лазера несколько выше, чем у рубинового, и составляет примерно 1 имп/с. При работе в режиме непрерывной генерации выходная мощность АИГ-лазера достигает 50 Вт для основной попереч-

ной моды и 500 Вт для многомодового излучения. АИГ-лазеру присущи те же недостатки, которые были отмечены для рубинового. Заметим, что мощное инфракрасное излучение АИГ-лазеров может быть преобразовано с помощью удвоителя частоты в зеленый свет с длиной волны 0,53 мкм.

Полупроводниковые лазеры являются во многих отношениях наиболее перспективными. Они малогабаритны, обладают высокой эффективностью и низкой стоимостью, а также просты и удобны в обращении. Наиболее важным обстоятельством является возможность получения матрицы полупроводниковых лазеров методами интегральной технологии. Типичным представителем полупроводниковых лазеров является инжекционный лазер на арсениде галлия (GaAs), который генерирует излучение с длиной волны, лежащей в интервале от 0,82 до 0,9 мкм, в зависимости от температуры охлаждения. Так, например, при температуре жидкого азота он дает излучение с длиной волны 0,84 мкм и может работать как в импульсном, так и непрерывном режиме. При комнатной температуре этот лазер может работать только в импульсном режиме, излучая импульсы длительностью порядка 100 нс с частотой повторения порядка 1 кГц на длине волны 0,9 мкм. Средняя мощность излучения достигает 1 Вт.

В настоящее время уже изготавливаются линейки и небольшие матрицы из таких лазеров с достаточно однородными характеристиками по излучению. В последнее время были разработаны также лазерные экраны из полупроводникового материала с возбуждением излучения сканирующим электронным пучком. Наряду с указанными выше достоинствами, полупроводниковые лазеры имеют ряд серьезных недостатков, которые препятствуют широкому применению этих лазеров в устройствах памяти. Недостатками являются низкая пространственная когерентность, относительно широкий спектр излучения ($\sim 10 \text{ \AA}$) и значительная угловая расходимость излучения, доходящая до 20° . Однако полупроводниковые лазеры в настоящее время совершенствуются, так что отмеченные недостатки могут быть устранены в самое ближайшее время. Так, например, в лабораторных условиях уже была получена одномодовая генерация в инжекционном лазере на арсениде галлия с составным резонатором. Ширина спектра при этом составляла $0,3 \text{ \AA}$.

Характеристики некоторых лазеров приведены в табл. 3.1. При выборе лазера для конкретного устройства необходимо учитывать ряд требований к его характеристикам, например запись страницы цифровой информации объемом 10^4 бит на площади 1 мм^2 . Если запись голограммы производится на обычный фотоматериал для голографического ПЗУ, то необходим световой импульс длительностью 0,1 мс при мощности порядка 1–10 мВт с учетом потерь света в дефлекторе, оптических элементах и устройстве набора страниц. В этом случае пригодны гелий-неоновые лазеры. Для выделения импульсов нужной длительности из непрерывного излучения газовых лазеров используют специальные селекторы импульсов. Для записи голограммы на оперативные

Характеристика	Гелий-неоновый атомарный лазер	Аргонный ионный лазер	Рубиновый импульсный лазер	АИГ-лазер	Полупроводниковый лазер на GaAs
Длина волны излучения λ , мкм	0,6328	0,488—0,5145	0,6943	1,064—0,53	0,82—0,90
Выходная мощность в режиме непрерывной генерации одной моды P , Вт	0,002—0,05 (0,2)	1 (10)	—	50 (500)	0,01 (1)
Энергия импульса, $E_{\text{и}}$, Дж	—	—	1—5	10—100	10^{-4}
Длительность импульса $t_{\text{и}}$, нс	—	—	10—100	10—100	100
Максимальная мощность, мВт	—	—	$10-10^2$	$10-10^3$	$2 \cdot 10^{-3}$
Средняя мощность $P_{\text{ср}}$, Вт	—	—	1	10—100	0,01—1
Частота повторения импульсов, Гц	—	—	0,1	50	10^3-10^4
Диаметр выходного луча d , мм	2	1,4	8	4	—
Расходимость луча $\theta_{\text{р}}$, мрад	0,7	0,5	10	1—8	$10^{-2}-20$
КПД η , %	0,01	0,1	2	2	30

Примечание. В скобках приведены мощности для многомодового излучения.

среды необходимы значительно большие мощности. Так, запись голограммы страницы информации объемом 10^4 бит на термопластик-фотопроводник с достаточно высокой скоростью (порядка 1000 стр/с) может быть обеспечена только аргонным лазером. Наиболее жесткие требования к энергетическим характеристикам лазерного излучения выдвигаются при использовании магнитооптических материалов [7]. Для термомагнитной записи качественной голограммы на магнитную пленку из сплава MnVb требуется световой импульс длительностью 10—100 нс и энергией порядка 5—10 мДж. Такие импульсы могут быть получены от твердотельных лазеров с модулируемой добротностью (см. табл. 3.1). Однако существующие твердотельные лазеры, такие, как рубиновые или АИГ-лазеры, имеют очень низкую частоту повторения импульсов и не могут обеспечить приемлемой скорости ввода информации в ГЗУ.

Кроме того, для получения качественных голограмм с высокой плотностью хранения информации лазерное излучение должно обладать высокой степенью пространственно-временной когерентности. Когерентные свойства лазерного излучения определяются его модовой структурой. Временная когерентность (длина когерентности) характеризуется числом продольных мод, генерируемых одновременно. Если

лазер работает в режиме непрерывной генерации и излучает только одну продольную моду, то можно считать, что временная когерентность абсолютна. При этом длина когерентности практически не ограничена. С увеличением числа одновременно генерируемых продольных мод временная когерентность излучения лазера ухудшается. Для импульсного лазера даже в режиме генерации одной продольной моды длина когерентности составляет 1 м. Этого достаточно для использования в ГЗУ.

Пространственная когерентность лазерного излучения определяется его поперечной модовой структурой. В этом случае желательно, чтобы генерировалась одна поперечная мода. Наиболее высокими когерентными свойствами обладает основная поперечная мода — TEM_{00} , которая обеспечивает наибольшую равномерность освещения. Лазер, способный генерировать одну поперечную (основную) и одну продольную моду, называют *одночастотным*. Именно такой лазер, обладающий наивысшей временной и пространственной когерентностью, и необходим для использования в когерентных оптических системах хранения и обработки информации.

§ 3.2. Дефлекторы

Устройства отклонения лазерного пучка (дефлекторы) в основном применяют в ГЗУ. Дефлектор позволяет направить лазерный пучок в любую из M_x^* позиций на носителе информации. Благодаря этому стало возможным создание ГЗУ с двух- и трехкоординатной выборкой при произвольном доступе к любой голограмме, записанной на накопительной среде. Необходимо отметить, что основные характеристики ГЗУ — емкость и быстродействие — в значительной степени определяются соответствующими характеристиками дефлектора.

Основными характеристиками дефлектора являются разрешающая способность и быстродействие. Разрешающая способность дефлектора в одном из двух взаимно перпендикулярных направлений, например в направлении оси Ox (N_x), определяется отношением максимального угла отклонения светового пучка в плоскости xOz ($\Delta\theta$) к его угловой ширине ($\delta\theta$):

$$N_x = (\Delta\theta) / (\delta\theta). \quad (3.1)$$

Угловая ширина светового пучка зависит от дифракционного углового размытия пучка. Так как дифракционное размытие λ/d , где d — диаметр отклоняемого светового пучка, то (3.1) можно записать в виде

$$N_x = \frac{d}{\varepsilon \lambda} \Delta\theta, \quad (3.2)$$

где ε — параметр, зависящий от требуемой степени разрешения (критерия) двух соседних направлений с учетом закона распределения интенсивности света в поперечном сечении пучка. Обычно двухкоорди-