

Министерство общего и профессионального образования
Российской Федерации

ОРЕНБУРГСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ

Кафедра радиофизики и электроники

Ю.Д.Лантух

ОПТИЧЕСКИЕ МЕТОДЫ В ИНФОРМАТИКЕ

учебное пособие для студентов естественнонаучных и технических специальностей

Оренбург 1999

Предисловие

В настоящее время перспективы развития информационных технологий немыслимы без активного внедрения достижений современной оптики в процессы передачи, обработки и отражения информации. Более того, именно с оптической электроникой связывается магистральное направление развития технических средств информатики. В пользу такого вывода говорит тот факт, что в каждой из трех основных составляющих информационных технологий уже сейчас имеются существенные достижения, обусловленные прогрессом оптоэлектроники. Это, во-первых, создание лазерных компакт-дисков – наиболее емких машиночитаемых носителей информации. Во – вторых, это внедрение волоконных и планарных световодов – оптимальных проводников информационных сигналов. И, в - третьих, появление оптических процессоров – наиболее быстродействующих систем для обработки информации.

В Оренбургском государственном университете предмет "Оптические методы в информатике" предлагается студентам некоторых специальностей Института энергетики и информатики в качестве дисциплины по выбору. Этим и обусловлена направленность данного издания. Его цель состоит в ознакомлении студентов с основами оптической электроники и оптических информационных технологий в рамках относительно небольшого (34 часа) лекционного курса.

Изложение основ и приложений оптических методов в информатике является предметом многих оригинальных и популярных статей, учебников и монографий. При написании настоящего учебного пособия автором взяты за основу некоторые из них. Это книги Ю.Р. Носова /3 /, А.А. Акаева и С.А. Майорова /4/, А.Л. Микаэляна /11/. Кроме списка использованных источников в пособии приведен список литературы, рекомендуемой для более подробного изучения данной темы.

Настоящее пособие состоит из шести разделов.

В разделе "Введение" рассмотрены общие проблемы традиционной электронной вычислительной техники, а также особенности систем оптической обработки информации, позволяющие вывести технические средства информатики на качественно новый уровень.

В разделе "Элементная база оптической вычислительной техники" приводятся сведения об основных компонентах и устройствах (таких как светодиоды, дефлекторы, управляемые транспаранты и т.д.), используемых в оптоэлектронике и оптической вычислительной технике.

Раздел "Интегрально-оптические элементы" посвящен изложению основ функционирования пассивных и активных элементов интегрально-оптических схем.

В четвертом разделе "Лазеры" в краткой форме излагаются принципы работы лазеров, являющихся основными источниками изучения в оптической электронике, а также рассмотрены некоторые типы лазерных излучателей.

Раздел "Основы голографии" содержит основные сведения о голографическом принципе записи и восстановлении изображений. Здесь же приведены классификация и свойства голограмм, в той или иной мере используемые в оптических системах обработки информации.

В шестом разделе "Оптические запоминающие устройства" рассмотрены регистрирующие среды, голографические системы записи информации и оптические дисковые системы.

Настоящее пособие будет полезным при изучении таких дисциплин как "Информатика", "Электронные вычислительные машины", "Оптоэлектроника", "Оптика" и др., а также всем кто интересуется проблемами современной электроники и компьютерной техники.

1 Введение

1.1 Информация и информационные технологии

В наше время общепризнанным стал тезис о том, что без накопления и умелого использования информации немислим научно-технический прогресс. Информация служит определяющим фактором развития экономической, технической и научной сфер человеческой деятельности. В современном сложном и многоликом мире ни одну крупную проблему невозможно решить без переработки значительных объемов информации и налаженных коммуникационных процессов /1/.

Информация позволяет человеку ориентироваться в окружающей обстановке, стимулирует его деятельность, определяет его поведение и обеспечивает его жизнедеятельность.

Принципиальные изменения в отношении человека к информации произошли в связи с появившейся возможностью перейти от ручных способов сбора и обработки информации к автоматизированным. Компьютер, открывший перед человеком почти неограниченные возможности в вопросах оперирования информацией, позволил превратить ее в ресурс, от состояния которого зависит развитие экономики любой страны.

Информационные ресурсы, являющиеся продуктом интеллектуальной деятельности наиболее квалифицированной и творчески активной части трудоспособного населения, в последние годы растут особенно быстро. Этот рост в последней четверти XX в. достиг таких рекордных величин, что многие специалисты начали говорить об «информационном взрыве» и «информационной революции». Действительно, с начала XX в. информационный поток увеличился примерно в 30 раз. Ежегодно в мире публикуется около 100 тысяч журналов на 60 языках, 5 млн. научных статей, книг, брошюр, 250 тысяч диссертаций и отчетов.

Вполне очевидно, что в создавшейся ситуации необходимо направить достижения науки и техники, связанные с переработкой и передачей информации, на совершенствование информационного обслуживания общества. Необходимо применение принципиально новых методов и средств обработки, хранения и передачи информации, оперирующих с большими ее объемами в реальном времени. В противном случае значительная доля генерируемой информации оседает в некоторой окрестности ее источника и остается незамеченной.

Отраслью науки, изучающей информационные процессы и пути их автоматизации, является информатика.

У нас в стране информатика своим происхождением обязана информационной деятельности, связанной со сбором, хранением, переработкой и распространением научной информации. По мере разработки и внедрения в практику средств автоматизации информационной деятельности, создания информационно-поисковых систем информационная деятельность по своему содержанию стала приближаться к разделу кибернетики, изучающему методы и средства управления информационными процессами.

В документах ЮНЕСКО термин "информатика" определяется в широком смысле, охватывая собственно информацию, её сбор, анализ и обработку, а также средства для обработки информации, включая микропроцессоры как таковые или в сочетании с другими электронными системами. Информатика названа крупным научным направлением, заслуживающим активного развития в интересах всего человечества. Она способна (при соответствующем освоении её методов и средств) помочь человеку полнее использовать информационные ресурсы в интересах научно-технического прогресса и социального развития. Уже сейчас информатика имеет весьма широкий спектр приложений, включая материальное производство, государственное планирование и управление, науку, оборону, административную сферу, охрану окружающей среды. Информатика развивается исключительно высокими темпами, несмотря на отсутствие пока общепринятого мнения относительно её предмета и связей с другими научными направлениями. По-видимому, ещё потребуются время, пока сложится единое мнение об информатике как научно-технической дисциплине.

Методы и средства информатики материализуются и доходят до потребителя в виде так называемых новых *информационных технологий*, под которыми подразумеваются современные виды информационного обслуживания, организованные на базе средств вычислительной техники и средств связи. Иными словами, информационные технологии это совокупность методов и средств для получения, хранения, переработки и отображения информации.

Человек постоянно пользовался информационными технологиями, однако в последнее время степень их совершенства пока отстаёт от его потребностей.

На всех этапах развития общества информационные технологии обеспечивали информационный обмен между людьми, коллективами, институтами, отражали соответствующий уровень и возможности систем регистрации, хранения, обработки и передачи информации и, по существу, являлись синтезом

методов оперирования человека с информацией в интересах той или иной сферы его деятельности.

Техника оперирования человека с информацией прошла долгий и интересный путь развития. Письменность, книгопечатание, телеграф, использование радиоволн и магнитных носителей – все это этапы совершенствования информационных технологий.

Основу современных информационных технологий составляют три технических достижения:

1) появление новых сред накопления информации на машиночитаемых носителях (например, магнитные ленты, микроформы, магнитные диски, сдвиговые регистры на полупроводниковых приборах, а также появившиеся в последнее время лазерные компакт-диски);

2) развитие средств связи, обеспечивающих доставку информации практически в любую точку земного шара без существенных ограничений во времени и расстоянии, широкий охват населения средствами связи (радиовещание, телевидение, сети передачи данных, спутниковая связь, телефонная сеть, оптоволоконные сети);

3) возможность автоматизированной обработки информации с помощью компьютера по заданным алгоритмам (сортировка, классификация, представление в нужной форме).

Одним из примеров информационной технологии может служить система электронной почты, для формирования которой необходимы компьютер, устройства долговременной памяти, система факсимильной передачи изображений и сеть передачи данных. Электронная почта является новым видом информационного обслуживания, способным заменить в будущем традиционную почту.

В целом современные информационные технологии направлены на увеличение степени автоматизации всех информационных процессов, что является основой ускорения темпов научно-технического прогресса.

В настоящее время уровень развития информационных технологий пока еще отстает от потребностей общества.

Наиболее перспективным путем развития информационной техники является применение достижений современной оптики в процессах регистрации, обработки и хранения информации.

Целью настоящего пособия является формирование у студентов общих представлений об оптических методах, используемых в информатике в настоящее время. Будут рассмотрены причины, побудившие к внедрению оптических методов в информационные технологии, их преимущества, недостатки, а также перспективы, появляющиеся при использовании методов современной оптики в информатике.

1.2 Проблемы традиционной электронной вычислительной техники

Как уже отмечалось, информатика возникла как дисциплина, изучающая информационные процессы и пути их автоматизации. Вся история цивилизации это, по сути, история развития средств и методов автоматизации информационных процессов. В настоящее время, благодаря изобретению компьютера, процессы обработки информации кардинально преобразились. Сейчас вполне привычным стало употребление термина "компьютеризация", обозначающего именно процессы автоматизации на основе компьютерных технологий. Компьютер и микроэлектроника открыли поистине безграничные возможности в оперировании с информацией. Все средства информатики (вычислительная техника, управляющие электронные комплексы, информационные системы) достигли современного уровня и продолжают непрерывно совершенствоваться благодаря стремительному прогрессу микроэлектроники. Основа микроэлектроники – интегральные схемы (ИС) год от года совершенствуются и усложняются. На одном кремниевом кристалле площадью в несколько квадратных миллиметров размещается до миллиона транзисторов. Разрешающая способность "инструментария" в производстве интегральных схем составляет ~ 1 мкм; на очереди появление новой отрасли электроники – наноэлектроники. Одновременно с уменьшением размеров транзисторов растет скорость выполнения ими элементарных операций над поступающими электрическими импульсами. Значение времени переключения транзисторов в традиционных кремниевых ИС составляет менее 1 наносекунды.

Однако сверхминиатюризация, оставаясь важным направлением микроэлектроники, не является панацеей при решении многих проблем, стоящих перед современной электронной техникой /2/.

Вскроем корпус обычного карманного калькулятора. Что мы увидим? Лишь одна - две ИС выполняют всю вычислительную работу и хранят всю информацию. Объем их не превышает сотых долей объема калькулятора. Остальное - это индикатор, клавиатура, батарейка, плата, соединительные провода и сам корпус. Можно ли здесь обойтись без чего - либо? Разумеется, нет. Налицо парадоксальная ситуация, типичная для всего компьютерного парка. Сколько бы мы не совершенствовали микроэлектронную "начинку" калькулятора, его важнейшие характеристики, такие, как масса и габариты, надежность, долговечность, стоимость практически не изменяются. Многие проблемы развития информатики лежат вне "сферы действия" микроэлектроники и "магнитного" принципа записи данных, составляющих основу современной компьютерной базы.

Какие же это проблемы? Перечислим некоторые из проблем, стоящих перед традиционной электронной вычислительной техникой.

Проблема вывода (отображения) информации.

После завершения работы компьютера по заданной программе полученный результат необходимо сделать доступным для пользователя. Вывод информации может осуществляться в виде твердой копии, т. е. на бумагу или на экран дисплея. Учитывая то обстоятельство, что 80-90% всей информации опе-

ратор усваивает в зрительной форме, можно сделать вывод о важности своевременного и качественного отображения результатов. Разработчики средств отображения информации предлагают дисплеи, основанные на различных принципах. На очереди появление трехмерных - голографических дисплеев. Поиск эффективных способов отображения информации является важной задачей и во многом определяется развитием оптоэлектроники.

Проблема ввода.

Чем мощнее, производительнее компьютер, тем больший объем информации следует вводить для обработки. И часто ввод информации становится "узким местом" в работе компьютера. Одна из причин такого положения состоит в том, что очень часто исходная информация, подлежащая обработке, представляет собой некое изображение, то есть двухмерный массив оптической информации. В традиционной электронике весь этот массив необходимо перевести в цифровую форму и последовательно вводить в компьютер. Этот процесс занимает время, появляется вероятность возникновения ошибок, сбоев и т. д. Оптические вычислительные машины позволяют осуществлять ввод информации без аналого-цифрового преобразования и производить многие операции над такими массивами и практически мгновенно.

Проблема плотности записи информации.

Плотность записанной информации, т. е. объем информации, приходящийся на единицу площади носителя – одна из главных характеристик запоминающего устройства (ЗУ). Согласно теории поток оптического излучения принципиально может быть сфокусирован до пятна с размером около $\lambda/2$ (λ - длина волны света), – таков же и минимальный шаг дискретности оптических воздействий. Это значит, что максимальная плотность записи оптической информации может достигать $4/\lambda^2$, то есть 10^9 бит/см². Напомним, для сравнения, что предельно достижимая плотность записи на магнитных носителях составляет всего 10^5 бит/см².

Наличие механического контакта между поверхностью носителя и считывающей (записывающей) головкой, а также "старения" магнитной записи (необходимо обновлять запись спустя некоторое время) также относятся к недостаткам магнитной памяти. Оптические ЗУ свободны от указанных недостатков.

Проблема межэлементных соединений.

Иначе эту проблему можно назвать проблемой электрических проводов и кабелей. Проводные связи в значительной мере определяют массовые и габаритные показатели компьютерных систем, сложность монтажа и ремонта, а в конечном счете их стоимость. Но дело не только в этом. При прохождении электрического сигнала по металлическому проводу возникает, хотя и слабое, излучение электромагнитных волн в окружающее пространство и современная сверхчувствительная аппаратура позволяет узнать "чем занимается" компьютер. Таким же путем можно вмешаться в работу машины. Возникла весьма актуальная проблема так называемого "несанкционированного доступа" к данным, программам и процессам в компьютерах.

Использование в качестве носителя информации электрически нейтральных фотонов обуславливает бесконтактность оптической связи. Отсюда следуют: идеальная электрическая развязка входа и выхода, однонаправленность потока информации, помехозащищенность оптических каналов связи, скрытность передачи информации по оптическому каналу связи. Объединение компьютеров в сети посредством волоконных кабелей существенно увеличивает скорость передачи данных и улучшает эксплуатационные характеристики.

Проблема пропускной способности каналов передачи данных.

Частота оптических колебаний на 3-5 порядков выше, чем в обыкновенном радиотехническом диапазоне, - это значит, что во столько же раз возрастает и пропускная способность оптического канала передачи информации.

Суммируя вышесказанное, можно отметить следующее.

Интерес к использованию оптических методов в информатике связан с фундаментальными свойствами фотонов при использовании их в качестве переносчиков информации. Это, прежде всего, отсутствие взаимодействия между фотонами при их распространении, а также пространственный масштаб 10^{-6} м, определяющий размеры области взаимодействия с веществом. В этой связи можно выделить следующие основные преимущества оптических методов:

- Высокая плотность записи информации (10^5 бит/мм² и более);
- Отсутствие непосредственного контакта между носителем и системой записи/считывания;
- Применимость голографических методов записи и обработки информации;
- Разнообразие способов практической реализации, основанное на множестве механизмов взаимодействия света с регистрирующими средами;
- Увеличение на 3-4 порядка быстродействия и помехозащищенности информационных систем.

К недостаткам оптоэлектронных средств обработки и передачи информации могут быть отнесены /3/:

1. Неудовлетворительная энергетика.

Коэффициент полезного действия преобразований вида $E \rightarrow L$ (электричество-свет) и $L \rightarrow E$ в лучших современных оптоэлектронных приборах (лазерах, светодиодах, р-і-n-фотодиодах), как правило не превышает 10–20%. Поэтому, если в устройстве эти преобразования осуществляются лишь дважды (на входе и на выходе), как, например, в оптопарах или волоконно-оптических линиях связи (ВОЛС), то общий КПД падает до единиц процентов. Введение каждого дополнительного акта преобразования информационных сигналов из одной формы в другую ведет к уменьшению КПД еще на порядок и более. Низкое значение КПД вызывает рост энергопотребления, что недопустимо из-за ограниченных возможностей источников питания, затрудняет миниатюризацию, теплоотвод, снижает эффективность и надежность большинства оптоэлектронных устройств.

2. Гибридность.

Составляющие оптоэлектронное устройство отдельные элементы и приборы, как правило, изготавливаются из различных материалов. Оптопара состоит из арсенида галлия (излучатель), полимерного оптического клея, кремния (фотоприемник). В ВОЛС к этим материалам добавляется кварц (световод).

Наличие разнородных материалов обуславливает: низкий общий КПД устройства из-за поглощения излучения в пассивных областях структур, отражения и рассеяния на оптических границах; снижение надежности из-за различия температурных коэффициентов расширения, разъюстировки при механических воздействиях, сложности общей герметизации устройства, технологическую сложность и высокую стоимость.

Заметим, что в традиционной микроэлектронике эти недостатки предопределили доминирование монокристаллических интегральных микросхем над гибридными.

3. Деградация.

В данном случае это понятие используется в широком смысле как снижение эффективности оптоэлектронных приборов при воздействии температуры, радиации, долговременной работе. Принципиальная особенность процессов распространения излучения в веществе и оптоэлектронных преобразований (обусловленная малостью длины волны света) состоит в исключительно высокой их чувствительности к нарушениям оптической однородности материалов, даже к субмикронным включениям. К появлению таких дефектов ведут вышеуказанные воздействия. Практически у всех видов излучателей имеет место снижение мощности излучения при возрастании температуры; у фотоприемников происходит увеличение темновых токов и шумов. Так же проявляется воздействие проникающей радиации, с той лишь разницей, что возникающие нарушения необратимы. Также неизбежно помутнение оптических сред и ухудшение светопропускания на границах разнородных материалов.

Сопоставление перечисленных достоинств и недостатков, значимость первых и возможность преодоления (хотя бы частично) вторых, позволяет сделать общий оптимистический вывод об огромных возможностях оптических методов в информатике.

1.3 Двумерный оптический сигнал и его информационная структура

Оптическим сигналом называют световую волну, несущую определенную информацию [4]. Особенностью световой волны по сравнению с радиоволной является то, что вследствие малости длины волны λ можно практически реализовать прием, передачу и обработку сигналов, промодулированных не только по времени, но и по пространственным координатам. Это позволяет значительно увеличить объем вносимой в оптический сигнал информации.

Таким образом, оптический сигнал в общем случае является функцией четырех переменных: трех пространственных координат x, y, z и времени t . Рассмотрим его математическое описание.

Электромагнитная волна представляет собой изменение во времени в каждой точке пространства электрического и магнитного полей, которые связаны между собой по закону индукции. Изменение магнитного поля создает переменное электрическое поле, которое в свою очередь порождает переменное магнитное поле. Электромагнитная волна распространяется в пространстве от одной точки к другой. Она характеризуется взаимно перпендикулярными векторами напряженностей электрического \vec{E} и магнитного \vec{H} полей, которые изменяются во времени по одному и тому же гармоническому закону:

$$\vec{U}(x, y, z, t) = \vec{P}(x, y, z) U(x, y, z, t) = \vec{P}(x, y, z) A(x, y, z) \cos(2\pi\nu t + \varphi(x, y, z)). \quad (1.1)$$

Световую волну можно представить с помощью электрического, либо магнитного поля. В оптике чаще всего для этой цели используют электрическое поле, поскольку оно играет более важную роль. Например, в оптической голографии в результате действия электрического поля можно получать голограммы. Поэтому в дальнейшем будем считать, что (1.1) описывает электрическое поле световой волны. В этом случае $\vec{P}(x, y, z)$ – единичный вектор, определяющий в пространстве прямую, вдоль которой осуществляется колебание электрического поля в точке пространства с координатами (x, y, z) и характеризующий плоскость поляризации в данной точке. Функция $U(x, y, z, t)$ – скалярная функция координат пространства и времени, численно равная мгновенному значению модуля вектора напряженности электрического поля $E(x, y, z, t)$; $A(x, y, z)$ – амплитуда колебания напряженности электрического поля в точке (x, y, z) , ν – частота колебаний, $\varphi(x, y, z)$ – фаза световой волны в точке с координатами (x, y, z) . Начальную фазу можно принять равной нулю в любой произвольной точке пространства. Тогда функция координат $\varphi(x, y, z)$ будет характеризовать разность фаз напряженности в этой точке и точке (x, y, z) . Кроме того параметры $\nu = const$; $\vec{P}(x, y, z)$; $A(x, y, z)$; $\varphi(x, y, z)$ не зависят от времени, так как рассматриваются только когерентные волны, а модуляция осуществляется по пространственным координатам.

Из (1.1) следует скалярная форма записи уравнения световой волны:

$$U(x, y, z, t) = E(x, y, z, t) = A(x, y, z) \cos(2\pi\nu t + \varphi(x, y, z)). \quad (1.2)$$

Обычно используют комплексную форму записи, которая является наиболее удобной для выполнения математических операций и преобразований; например, $y = a \cos \omega t$ записывают в виде $y = a \exp(i\omega t)$. Согласно формуле Эйлера $\exp(\pm i\omega t) = \cos \omega t \pm i \sin \omega t$, так что действительная функция y может быть получена из комплексной: $y = a \cos \omega t = \operatorname{Re} y = \operatorname{Re}[a \exp(i\omega t)]$, где символ Re обозначает действительную часть комплексной функции. Тогда:

$$\begin{aligned}
 U(x,y,z,t) &= A(x,y,z) \exp \{i[2\pi \nu t + \varphi(x,y,z)]\} = \\
 &= a(x,y,z) \exp(2\pi \nu t)
 \end{aligned}
 \tag{1.3}$$

Величину

$$U(x,y,z) = a(x,y,z) = A(x,y,z) \exp [i \varphi(x,y,z)]
 \tag{1.4}$$

называют *комплексной амплитудой* световой волны. Она описывает пространственное распределение амплитуд $A(x,y,z)$ и фаз $\varphi(x,y,z)$ световой волны и является важной характеристикой, монохроматической волны.

Временной множитель $\exp(i 2\pi \nu t)$, являющийся гармонической функцией времени, обычно опускают. Он может быть введен на любом этапе преобразований. Поэтому в дальнейшем оптический сигнал будем представлять в виде (1.4). Таким образом, основными характеристиками световой волны являются амплитуда $A(x,y,z)$, фаза $\varphi(x,y,z)$ и поляризация, определяемая единичным вектором $\vec{P}(x,y,z)$. В оптических системах хранения и обработки информации, как правило, используют двумерный оптический сигнал, который описывается распределением комплексной амплитуды, фазы или поляризации световой волны по точкам пространства, летящим в плоскости, перпендикулярной направлению распространения волны. Если в этой плоскости ввести координаты x,y , то информация, содержащаяся в двумерном сигнале, будет определяться комплексной амплитудой

$$U(x,y) = a(x,y) = A(x,y) e^{i\varphi(x,y)}
 \tag{1.5}$$

и поляризацией $\vec{P}(x,y)$. Итак, информация в световую волну может быть введена путем модуляции амплитуды, фазы и поляризации по двум пространственным координатам x и y .

1.4 Обобщенная схема функционирования оптической вычислительной машины

Понятие *оптической вычислительной техники* [3] охватывает комплекс алгоритмических и технических средств, предназначенных для выполнения функций электронной вычислительной техники, но оперирующих не с электрическими, а с оптическими информационными сигналами. К числу основных относятся функции процессора – математические и логические преобразования, запоминающего устройства – запись, хранение, стирание, считывание информации, устройств ввода-вывода, интерфейса и канала передачи данных.

Обычно это определение сужают и распространяют понятие оптической вычислительной техники лишь на оптические процессоры и запоминающие устройства (а иногда только на процессоры).

Принципиальная схема системы, осуществляющей оптическую обработку информации представлена на рисунке 1.1, где цифрами обозначены:

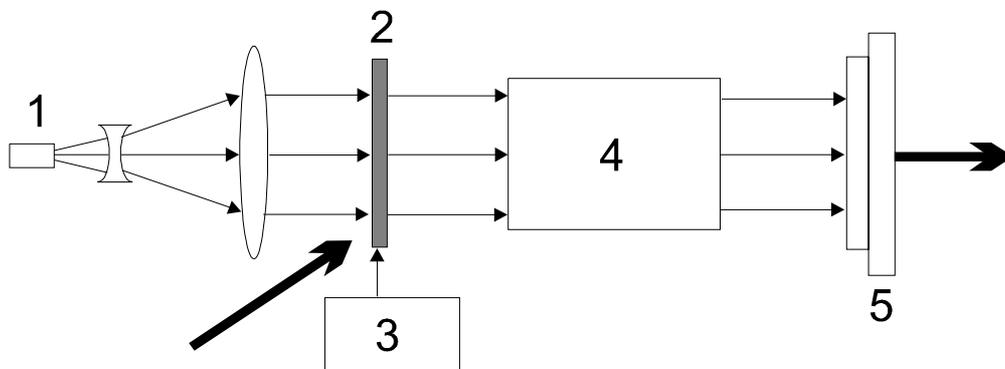


Рисунок 1.1

1. Лазер
2. Пространственный модулятор
3. Устройство управления пространственным модулятором
4. Оптический процессор
5. Преобразователь выходных оптических сигналов, запоминающее устройство.

Лазер – это источник когерентного оптического излучения. Пространственный модулятор – электрически или оптически управляемый транспарант, создает пространственно модулированный лазерный пучок, поступающий в процессор. Устройство управления пространственным модулятором осуществляет подачу и смену информации, подлежащей обработке. Оптический процессор – оптическое решающее устройство, в котором осуществляется заданное преобразование входного пространственного сигнала. Это устройство представляет собой набор различных элементов (линз, диафрагм, специальных фильтров, дифракционных решеток, голограмм, зеркал и т.д.), расположенных в пространстве определенным образом. Преобразователь выходных оптических сигналов фиксирует результат обработки информации, например, в виде фотографического изображения или голограммы. В этом блоке осуществляется преобразование выходящих оптических сигналов в электрические (фотоумножитель, видикон и т.д.). Рассматриваемый преобразователь может выполнить роль оптического запоминающего устройства или может быть связан с запоминающим устройством.

Оптический процессор является основным компонентом устройств оптической вычислительной техники.

Принцип действия оптического процессора основан на следующих характерных особенностях, присущих оптическому сигналу как переносчику информации и оптической вычислительной технике в целом:

- Параллельность обработки информации;
- Двумерность оптических систем;
- Возможность (и удобство) обработки аналоговой информации;
- Оперирование с информацией в оптической форме;
- Простота выполнения комплексного умножения, преобразования Фурье и других операций;
- Высокая предельная скорость переключения и отсутствие реактивностей при работе с фотонами.

Рассмотрим в качестве примера работу оптического процессора по выполнению операции умножения функций (рисунок 1.2).

Если световое поле с комплексной амплитудой $E(x,y)$ проходит через среду с коэффициентом амплитудного пропускания $\Phi(x,y)$, то напряженность поля в пространстве за средой изменяется по закону $E(x,y)\Phi(x,y)$.

Пусть напряженность электрического поля световой волны лазерного пучка на входе в процессор постоянна: $E_0(x,y) = E = const$.

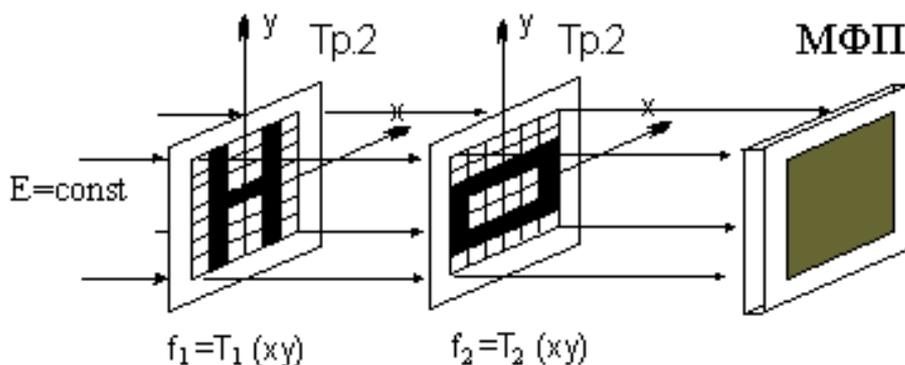


Рисунок 1.2

Далее такой пучок проходит через транспарант (диапозитив) $Tr1$, прозрачность которого $\Phi_1(x,y)$ определенным образом меняется в зависимости от координат x и y . Это означает, что в пучок введена информация.

Напряженность электрического поля световой волны в пучке, прошедшем транспарант $Tr1$, прозрачность которого неодинакова в плоскости xy , равна $E\Phi_1(x,y)$. Сигнал за транспарантом промодулирован и распределение его интенсивности по сечению соответствует функции $f_1(x,y)$.

Если на пути пучка поставить второй транспарант $Tr2$, пропускание которого есть функция $f_2 = \Phi_2(x,y)$, то после него распределение поля в пучке $E\Phi_1(x,y)\Phi_2(x,y)$ будет соответствовать произведению функций f_1 и f_2 .

Оптический сигнал, являющийся произведением функций f_1 и f_2 , может быть преобразован в последовательность электрических сигналов с помощью, например, матрицы фотоприемников (МФП). Опрашивая по порядку фотопри-

емники одной строки, затем другой, третьей, и т.д., мы получим последовательность электрических сигналов, каждый из которых пропорционален произведению функций f_1 и f_2 при определенных координатах x и y .

1.5 Особенности систем оптической обработки информации

Одним из главных принципов, на которых базируется работа оптических вычислительных машин, является *принцип параллельности* обработки оптической информации. Это означает, что операции над оптическим сигналом (умножение, деление и т.д.) могут осуществляться параллельно, т. е. одновременно по многим каналам, что придает вычислительным устройствам простоту и компактность. Именно такой принцип был продемонстрирован в рассмотренном нами примере работы оптического процессора. Возможность параллельной обработки информации базируется на двумерном характере оптического сигнала, представляющего собой распределение поля в пространстве двух измерений и связанным с этим свойством многоканальности, заложенным в оптическом сигнале.

Важным следствием принципа параллельности обработки информации является возможность создания *ассоциативного запоминающего устройства*. Если при обычной организации запоминающего устройства обращение к памяти происходит только адресным способом, то есть поиск информации происходит путем последовательного анализа содержимого всех ячеек памяти, то в ассоциативном запоминающем устройстве обращение осуществляется непосредственно ко всем хранящимся в памяти данным по некоторому формальному признаку. В случае оптической вычислительной машины, для которой характерен параллельный доступ ко всей информации, хранящейся в ЗУ с ассоциативной организацией, становится возможным извлечение полезной информации за время одного периода обращения, то есть существенно уменьшается время выборки. Примером устройств, осуществляющих ассоциативный принцип, служат голографические запоминающие устройства и оптический коррелятор.

Оперирование с информацией в оптической форме. Когда входные и выходные сигналы световые (например, в ретрансляторах волоконных оптических линий связи), использование для их обработки оптического процессора вместо электронного избавляет от необходимости двойного оптоэлектронного преобразования и связанных с этим энергетических потерь и искажений.

Аналоговая и цифровая обработка информации. Оптические методы позволяют производить как аналоговую, так и цифровую обработку информации.

В современных цифровых вычислительных машинах микропроцессоры производят операции с использованием двоичной системы счисления, в основе которой только две цифры "0" и "1". Для представления литер русского и латинского алфавита, а также других печатных знаков требуется семь, восемь и более двоичных цифр. Двоичный код маломощен, зато используемые в нем цифры легко распознать, так как в памяти ЭВМ они представлены как отсутствие или наличие заряда, напряжения, тока, намагниченности и проч., а при считывании или

передаче информации от устройства к устройству – как отсутствие или наличие импульса тока или напряжения. Это, в свою очередь, позволяет строить электронные микропроцессоры на базе электронных переключателей, которые могут находиться только в двух состояниях: "выключен" или "включен".

В силу этого подавляющее большинство современных вычислительных машин не приспособлены для непосредственной обработки информации находящейся в аналоговой форме: в виде картинка, графика, изображения. Если аналоговая информация все же должна быть обработана на ЭВМ, предварительно эту информацию надо перевести в двоичный код.

Именно для обработки аналоговой информации наиболее перспективны оптические и оптоэлектронные системы. Уже сейчас созданы высокопроизводительные когерентные оптические машины аналогового типа. Они успешно используются для решения таких сложных и трудоемких задач, как обработка изображений и распознавание образов. Характерной чертой оптических аналоговых вычислительных машин является то, что все элементы на входе преобразуются в результирующий сигнал на выходе одновременно (скорость света составляет $3 \cdot 10^8$ м/сек). Благодаря этому достигается огромная производительность – более 10^{12} оп/сек.

Однако такие вычислительные машины (также как и электронные аналоговые вычислительные машины) имеют ограниченную точность вычислений – порядка 1%; другой их недостаток – это отсутствие гибкости присущей электронной технике, то есть преобразовать оптический процессор под решение другой задачи трудно. Оптические аналоговые вычислительные машины - специализированные устройства.

Цифровая оптическая обработка информации предполагает замену элементной базы, логических элементов и функциональных узлов ЭВМ на оптические аналоги. Для создания устройств оптических вычислительных машин, осуществляющих последовательно-параллельную обработку оптической информации, требуется наличие пороговых элементов, обладающими двумя устойчивыми состояниями; набора логических элементов, реализующих функционально полную систему логических функций; запоминающих устройств, позволяющих записывать, хранить и считывать информацию и ряда других компонент.

2 Элементная база оптической вычислительной техники

2.1 Оптоэлектроника и оптическая вычислительная техника

Традиционная электроника имеет в своем распоряжении такие элементы как резисторы, конденсаторы, диоды, транзисторы и другие, но в то же время ИС, жидкокристаллические индикаторные панели, модули памяти... - всё это тоже "элементная база". Понятие "элементная база" по сути, переродилось и означает теперь набор компонентов с очень высокой степенью интегрированности многих отдельных деталей, собранных в одном корпусе. Для оптоэлектроники и оптической вычислительной техники эта тенденция проявляется ещё более явно. Лазеры, фотоприёмные устройства, модуляторы, ячейки памяти - это достаточно сложные устройства, которые с трудом можно отнести к исходным "кирпичикам", из которых строится здание любого узла или прибора системы оптической обработки информации. Помня об этом, мы в понятие "элементная база" будем вкладывать двоякий смысл: во-первых, это действительно элементная база, например, светодиод или оптический аналог транзистора - трансфазор, во-вторых, - более сложные устройства, как по компоновке, так и по назначению: лазер, модулятор, фотоприемная матрица. А в общем все они составляют компоненты оптических систем обработки информации.

Мы очень часто употребляем термин - "оптоэлектроника". Оптоэлектроника и оптическая вычислительная техника переплетены между собой. Особенно часто понятие "оптоэлектроника" употребляется при перечислении и рассмотрении конкретных приборов, компонентов оптической вычислительной техники (т.е. той самой "элементной базы"). Почему так ? Для этого надо ответить на вопрос: что такое оптоэлектроника ?

Оптоэлектроника /3/- раздел электроники, связанный, главным образом, с изучением эффектов взаимодействия между электромагнитными волнами оптического диапазона и электронами вещества (преимущественно твердых тел) и охватывающий проблемы создания оптоэлектронных приборов (в основном методами микроэлектронной технологии), в которых эти эффекты используются для генерации, обработки, хранения и отображения информации. Т.е. функциональное назначение оптоэлектроники состоит в решении задач информатики! Оптоэлектроника даёт средства для решения задач информатики, поэтому рассмотрение этих средств и есть, по существу, изучение элементной базы оптоэлектроники.

О каких же средствах идет речь? На этот вопрос не трудно ответить. Достаточно представить себе всю цепочку "информационного процесса": генерация, передача, обработка, хранение и отображение информации. На каждом этапе есть свой элемент или устройство. Это излучатели (генерация информации), световоды и другие элементы волоконной оптики (передача), элементы оптических процессоров и запоминающие устройства (обработка), регистри-

рующие среды и устройства памяти (хранение). Таким образом, знакомство с элементной базой оптоэлектроники и оптической вычислительной техники по сути сводится к решению следующих задач: во-первых, это ознакомление с принципом действия устройств, соответствующих каждому шагу цепочки и краткая характеристика самих этих устройств, во-вторых, (и это главное) нужно понять смысл выполняемых данным устройством операций и преобразований. Так, при изучении работы оптического процессора необходимо уяснить как смысл выполняемой этим устройством операции (например, "оптического сложения"), так и принцип работы управляемого транспаранта, - элемента, при взаимодействии световых волн с которым, осуществляется данная операция.

2.2 Нелазерные излучатели

Излучатель (источник света) - прибор, преобразующий электрическую энергию возбуждения в энергию оптического излучения заданного спектрального состава и пространственного распределения.

Излучатели представляют собой основу практически любой оптоэлектронной системы, в значительной степени определяя её функциональные возможности, эксплуатационные и стоимостные характеристики. Выбор излучателей для той или иной цели, как правило, ограничен и практически predetermined, поэтому под излучатель "подстраиваются" остальные элементы системы – фотоприёмники и пассивные оптические элементы.

2.2.1 Требования к излучателям

Требования к излучателям таковы:

1. Высокая эффективность преобразования энергии из электрической формы в оптическую.
2. Узкая спектральная полоса излучения.
3. Направленность излучения.
4. Быстродействие (т.е. быстрое возникновение и гашение излучения).
5. Совместимость с интегральными микросхемами.
6. Высокая технологичность и низкая стоимость.
7. Высокие эксплуатационные характеристики, долговечность.
8. Миниатюрность, твердотельность.
9. Когерентность генерируемого излучения.

О последнем требовании следует сказать особо. Многие разделы оптоэлектроники, например, голографическая запись и обработка информации, могут существовать лишь при использовании источников когерентного излучения, причём их характеристики тем лучше, чем выше степень когерентности излучения. В других разделах требуются лишь такие косвенные проявления когерентности как узкий спектр и направленность излучения. А в ряде случаев – в индикаторной технике, оптронах - когерентность не нужна. Т.е. последнее требование может быть важнейшим - и тогда определяющим видом оптоэлектронного излучателя является лазер, а может и вообще не приниматься во внимание - в этом случае в качестве источника света в большинстве случаев используются светоизлучающие диоды - излучатели, основанные на инжекционной спон-

танной электролюминесценции полупроводниковых диодов. Упомянув о лазерах, добавим, что значительную их часть в оптоэлектронике составляют полупроводниковые лазеры (ППЛ) или лазерные диоды – приборы, которые отличаются от светоизлучающих диодов лишь наличием оптического резонатора. Поэтому, говоря о светоизлучающих диодах, их можно (и нужно) сравнивать с ППЛ.

Излучатели, основанные на других принципах, нашли ограниченное применение. Миниатюрные лампочки накаливания, у которых в качестве тела излучения используется вольфрамовая нить, вмонтированная в стеклянный баллон и нагреваемая электрическим током до 2100 ...2300 К, характеризуются очень широким спектром, охватывающим часть ультрафиолетового, весь видимый и большую часть инфракрасного диапазона, и значительной инерционностью (10^{-1} ... 10^{-2} с). Правда, эти простейшие излучатели очень дешевы, обладают высокой температурной и радиационной стойкостью, а также интегральной яркостью, но все это обуславливает их использование лишь в нескольких устаревших разновидностях оптопар.

Неоновые лампочки с люминесценцией газового разряда применяются в качестве простейших индикаторов "дооптоэлектронной эры"; лишь конструктивно-технологические принципы интегральной электроники (плоская панельная мозаичная структура) открыли перед этим видом люминесценции перспективы в одном из разделов оптоэлектроники (отображающие плазменные панели).

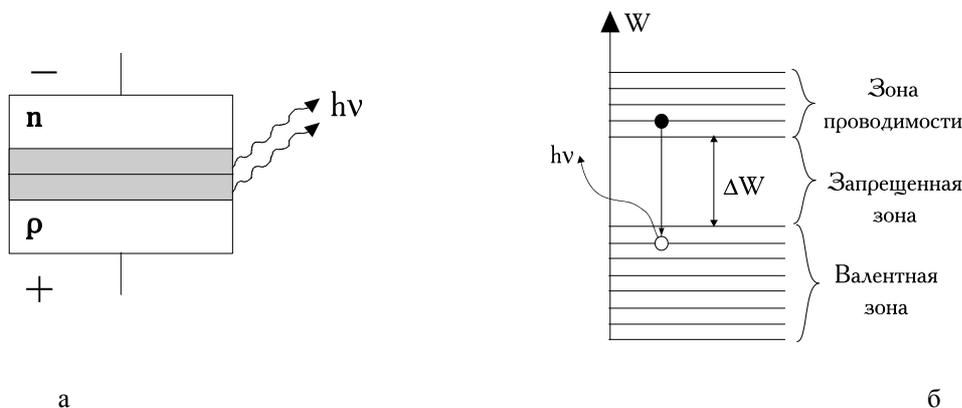
Ограниченное применение в оптоэлектронике находят также дискретные электролюминесцентные ячейки, использующие свечение порошковых или пленочных люминофоров. При сопоставлении со светодиодами они проигрывают по всем показателям — более высокие напряжения и мощности возбуждения, низкая светоотдача, широкий спектр излучения, низкое быстродействие, неудовлетворительные характеристики. Лишь в индикаторной технике, где требуются большие светящиеся площади, этот механизм излучения получил реализацию.

Сказанное относится и к другим видам излучателей — ни один из них не может хотя бы поколебать то доминирующее положение, которое занимают в оптоэлектронике лазеры и светодиоды.

2.2.2 Светоизлучающие диоды

Механизм возникновения люминесценции в светоизлучающих диодах /5/ (СИД) обусловлен инжекцией носителей заряда (электроны и дырки) в область *p-n* перехода при включении диода в прямом направлении (рисунок 2.1 а). Попав в область *p-n* перехода, эти носители становятся неосновными и рекомбинируют с основными. Процесс рекомбинации сопровождается переходом электронов с более высоких энергетических уровней зоны проводимости на более низкие уровни валентной зоны. Выделение энергии при этом происходит в форме излучения квантов света $h\nu$ (рисунок 2.1 б).

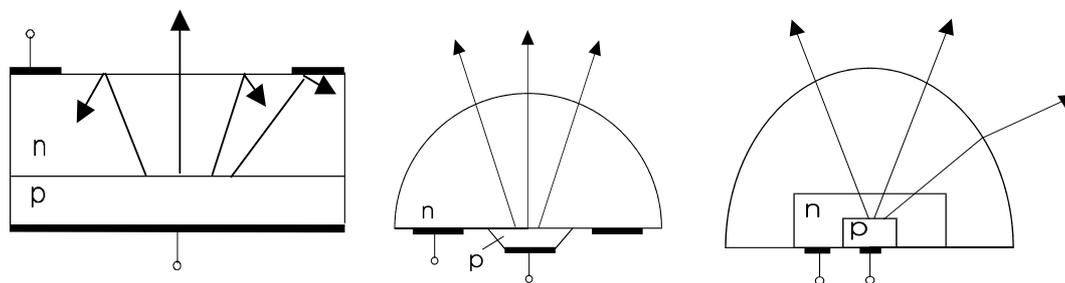
Энергия выделяющихся фотонов почти равна ширине запрещенной зоны ΔW , т.е. $h\nu = hc/\lambda = \Delta W$. Подставив значение постоянной Планка и скорости света, определим ширину запрещенной зоны ΔW (в эВ), необходимую для получения излучения с длиной волны λ (в микрометрах) $\Delta W = 1.23/\lambda$.



а
Рисунок 2.1

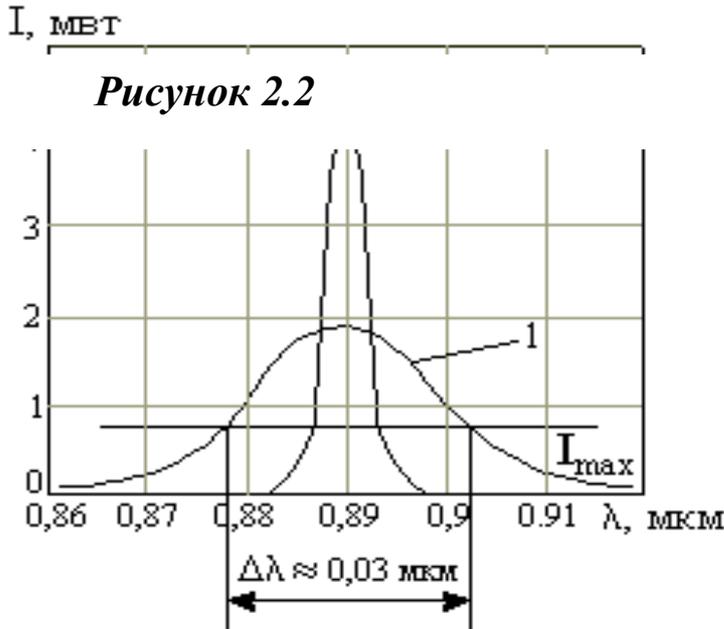
Отсюда следует, что для получения видимого (0.38 - 0.75 мкм) или ближнего инфракрасного излучения (0.7 - 3 мкм) необходимы полупроводники с $\Delta W > 1.7$ эВ для видимой области и $\Delta W > 0.4$ эВ для ближней инфракрасной области. Материалы для изготовления СИД: фосфид галлия GaP, карбид кремния SiC, арсенид галлия GaAs, арсенид индия InAs, а также некоторые тройные соединения: арсенид-галлия-алюминия GaAlAs, арсенид-фосфид галлия GaAsP и т.д.

Конструкция светодиодов (рисунок 2.2) помимо технологичности определяется необходимостью вывода наружу возможно большего светового потока. Однако значительная часть потока излучения всё же теряется из-за поглощения в самом полупроводнике и полного внутреннего отражения на границе кристалл-воздух (показатель преломления GaAs равен 3.3). Для уменьшения потерь, одну из областей делают в виде полусферы или снабжают светодиод плоской полусферической линзой.



Основные параметры СИД:

1. Мощность излучения I , (единицы милливатт);
2. Рабочее напряжение U_p , (единицы вольт);
3. Спектральный состав излучения $\Delta\lambda$, (30 - 40 нм, что на порядок больше чем у ППЛ – рисунок 2.3, где кривая 1 представляет спектральное распределение мощности излучения СИД, а кривая 2 - ППЛ);



4. Быстродействие (десят-

ки наносекунд – микро-секунды).

5. Долговечность ($\sim 10^5$ часов);

6. Коэффициент преобразования $E - L$ (до 10%);

7. Диаграмма направленности излучения (фронтальный (поверхностный) СИД - 180° , торцевой - $10-20^\circ$).

Основными областями

применения СИД являются оптопары, ВОЛС, устройства индикации.

2.3 Дефлекторы

2.3.1 Общая характеристика

Как можно отклонить световой (лазерный) луч? Это можно сделать с помощью зеркал или призм. Однако на практике часто требуется не просто отклонить световой луч, но сделать это очень быстро! Более того, часто требуется быстро и определённым образом менять направление луча (сканировать луч). Т.е. требуется управлять лучом по некоторой программе. Здесь механическое отклонение зеркал или призм не годится, так как оно совершается относительно медленно. В подобных случаях требуются немеханические методы управления лазерным лучом в пространстве.

Устройства, обеспечивающие отклонение или сканирование светового луча, называют *оптическими дефлекторами* (лат. deflecto - отклонять). Различают дефлекторы непрерывного и дискретного отклонения. Первые используются для непрерывного сканирования луча в пространстве, а вторые - для дискретного. Применяя последние, можно последовательно ориентировать луч в различных определённых направлениях или дискретно менять положение луча

в пространстве, не меняя его направления.

На практике широко используют дефлекторы двух типов: электрооптические и акустооптические. Основными характеристиками дефлектора являются разрешающая способность и быстродействие. Разрешающая способность дефлектора в одном из двух взаимно перпендикулярных направлениях, (например, $Ox - N_x$) определяется отношением максимального угла отклонения светового пучка $\Delta\theta$ к его угловой ширине $\delta\theta$, (рисунок 2.4).

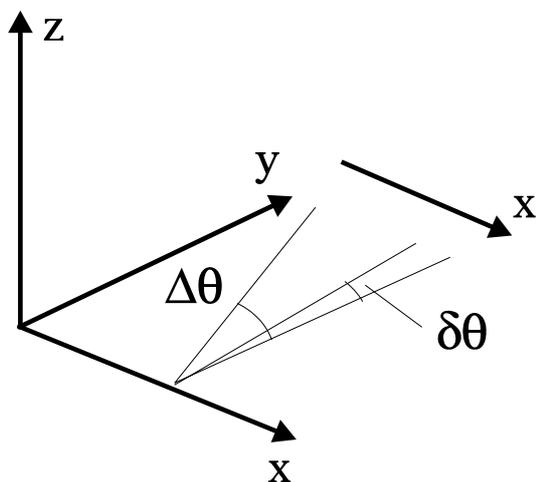


Рисунок 2.4

Например, если $\Delta\theta = 1$ рад, а $\delta\theta = 1$ мрад, то $N_x = \frac{\Delta\theta}{\delta\theta} = 10^3$.

Быстродействие дефлектора определяется временем t_n , необходимым для перемещения светового луча из одной позиции в другую. Время t_n это время произвольного доступа к голограммам на носителе или соответствующим страницам. Часто под быстродействием дефлектора понимают скорость переключения пучка v_n , которая измеряется числом переключений в секунду. Очевидно $v_n = 1/t_n$. Пусть $t_n = 0.01$ с, тогда $v_n = 100$ переключений в секунду. Типичные значения t_n - 0.1 - 10 мкс.

2.3.2 Электрооптические дефлекторы непрерывного отклонения

Действие электрооптических дефлекторов непрерывного отклонения (ЭОД НО) основано на электрооптическом эффекте - изменении показателя преломления n среды под действием электрического поля. Простейший такой дефлектор представляет собой обычную трёхгранную призму с отклоняющим углом Φ из электрооптического материала, на торцах которой находятся металлические электроды (рисунок 2.5). Значение n зависит от напряженности электрического поля. Изменяя напряжение на электродах, можно плавно менять n , а, следовательно, угол отклонения пучка θ :

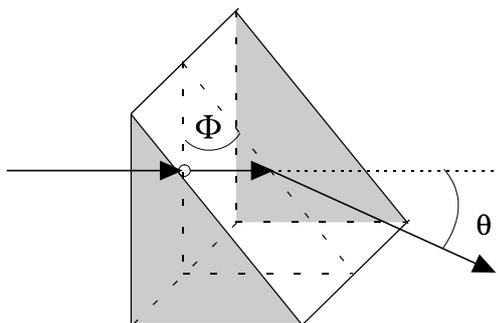


Рисунок 2.5

$$\theta = \arcsin(n \sin \Phi) - \Phi . \quad (2.1)$$

Поместив на пути луча два последовательно расположенных призмных дефлектора, ориентированных под углом в 90° друг относительно друга, можно осуществить двумерное сканирование луча. Быстродействие и разрешающая способность ЭОД НО: 0.1 - 1 мкс и 100 x 100 позиций.

2.3.3 Электрооптические дефлекторы дискретного действия

Электрооптический дефлектор дискретного действия (ЭОД ДД) способен менять положение светового пучка, сохраняя его направление. Принцип действия ЭОД ДД поясняется схемой, изображенной на рисунке 2.6. Схема состоит из 2-х каскадов. 1 и 2 - ячейки Поккельса, которые изменяют поляризацию проходящего через них излучения на 90° ; они играют роль "переключателей поляризации". K_1 и K_2 - двулучепреломляющие кристаллы, которые по разному преломляют лучи с исходной и ортогональной поляризацией ("обыкновенный" и "необыкновенный" лучи).

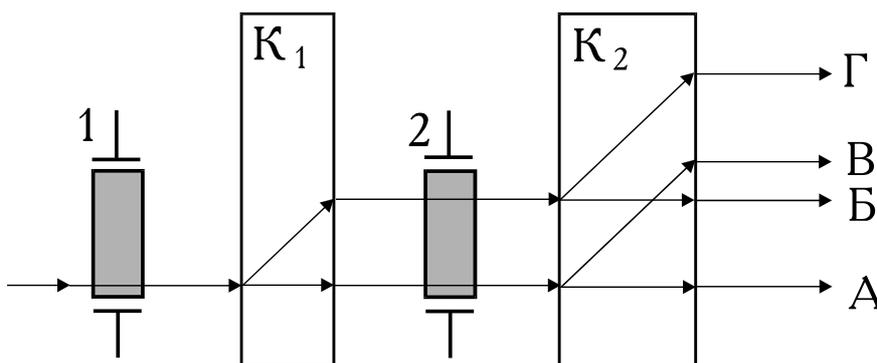


Рисунок 2.6

Работа ЭОД ДД осуществляется следующим образом:

1. Пусть обе ячейки выключены. Световой луч проходит оба кристалла не отклоняясь ("обыкновенный" луч) и выходит из дефлектора в точку А.
2. Пусть обе ячейки включены. Теперь при прохождении ячейки 1 плоскость поляризации пучка поворачивается на 90° . Для кристалла K_1 он становится "необыкновенным" лучом и отклоняется в кристалле. После ещё одного поворота плоскости поляризации в ячейке 2, луч снова становится "обыкновенным". Из дефлектора этот луч попадет в точку Б.
3. Легко сообразить, что если выключить ячейку 1, но включить 2, то луч преломится в K_2 и попадет в точку В.
4. Если ячейка 1 включена, а 2 выключена, то луч будет необыкновенным для K_1 и K_2 и в итоге попадет в точку Г.

При наличии в схеме N каскадов, число возможных положений луча на выходе дефлектора равно 2 в степени N . В современных ЭОД ДД уверенно реализуются 1024 положения, что соответствует 10 каскадам.

2.3.4 Акустооптические deflectоры непрерывного отклонения

Отклонение светового луча в акустооптических deflectорах непрерывного отклонения (АОД НО) можно рассматривать как результат дифракции света на искусственно созданной в среде дифракционной решетке [4]. Решетка создается в специальных кристаллах (молибдат свинца, парателлуриит и др.) ультразвуковыми (акустическими) волнами. Отклоняющая ячейка АОД состоит из следующих основных элементов (рисунок 2.7). 1 - светозвукопровод. 2 - электроакустический преобразователь, который возбуждает в среде акустические волны.

Рассмотрим действие АОД НО с помощью рисунка 2.7.

Пусть преобразователь возбуждает в среде плоскую акустическую волну, распространяющуюся в кристалле в направлении Ox .

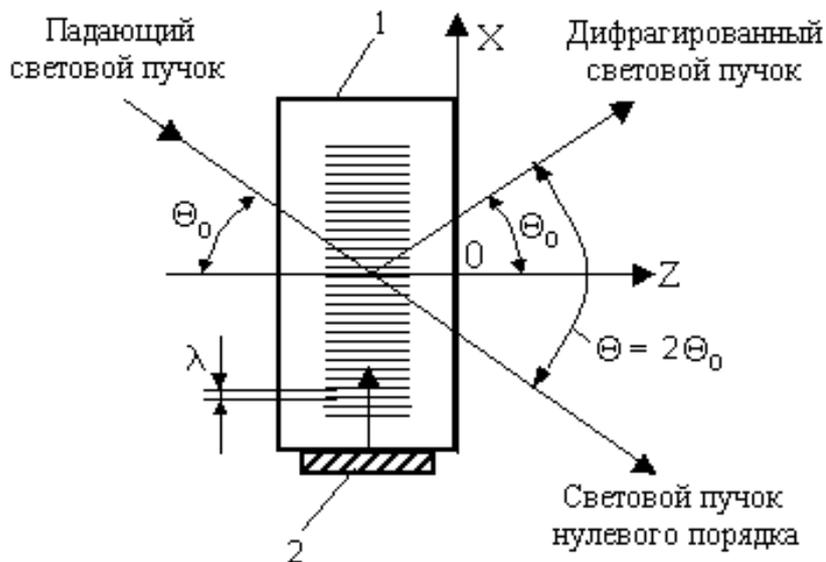


Рисунок 2.7

Акустическая волна вызывает синусоидальные изменения показателя преломления среды. Эти изменения приводят к образованию фазовой объёмной дифракционной решетки, период которой Λ равен длине акустической волны ($v_a T_a$). Угол дифракции падающего светового пучка определяется выражением:

$$\Theta = \arcsin\left(\frac{\lambda}{2\Lambda}\right) = \frac{\lambda}{2\Lambda} = \frac{1}{2v_a} v_a, \quad (2.2)$$

где λ - длина световой волны;
 v_a - скорость звуковой волны;

ν_a - частота ультразвука.

Изменяя частоту ультразвука ν_a , мы можем плавно менять Λ , а следовательно угол отклонения лазерного пучка дефлектором Θ . Быстродействие дефлектора можно оценить по времени прохождения акустической волной расстояния, соответствующего сечению светового пучка $\tau = d/\nu_a$. Если d (диаметр лазерного пучка) = 2 мм, $\nu_a = 2000$ м/с, $\tau = 1$ мкс. Разрешающая способность $N = 100 \times 100$ позиций.

2.4 Управляемые транспаранты

2.4.1 Определения

Управляемый оптический транспарант представляет собой двухкоординатную матрицу элементарных ячеек, оптические свойства которых могут изменяться независимо друг от друга и обратимо во времени (реверсивно) под влиянием внешнего воздействия [3]. Транспарант служит для управляемого изменения характеристик падающего на него потока излучения. Управляющее воздействие может быть в форме электрического, электронно-лучевого, оптического возбуждения. Наиболее широкое распространение получили электрически (рисунок 2.8) и оптически (рисунок 2.9) управляемые транспаранты (ЭУТ и ОУТ). Смысл электрического управления очевиден - это преобразование информации из одной формы в другую. При оптическом управлении форма информации остается той же самой, но может иметь место изменения интенсивности, спектрального состава и, что самое главное, переход от некогерентного излучения к когерентному. Из-

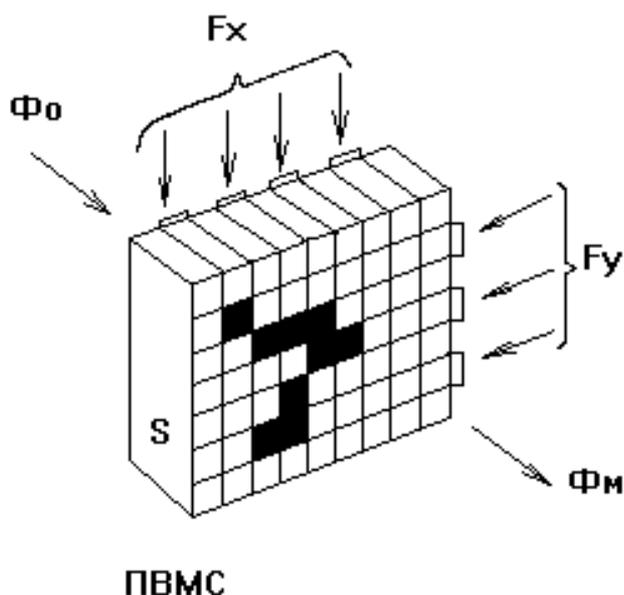


Рисунок 2.8

менение оптических свойств по координатам происходит как дискретно (типично для ЭУТ) - элементарные ячейки конструктивно обозначены, выделены,

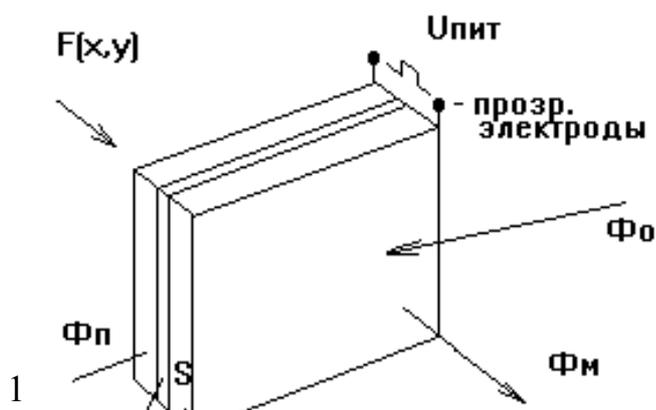


Рисунок 2.9

так и непрерывно (типично для ОУТ) - об элементарных ячейках можно говорить условно, имея в виду конечность разрешающей способности транспаранта. Изменение оптических свойств ячеек также может осуществляться скачкообразно или непрерывно, соответственно транспарант

может оперировать с аналоговой и цифровой информацией. Управление ячейками может осуществляться по мозаичному (индивидуально каждая ячейка) или по матричному (двухкоординатная выборка) принципу. В последнем случае говорят: "по принципу совпадений". В зависимости от устройства транспаранты работают на пропускание, либо на отражение. Модулируемыми параметрами светового пучка могут быть его амплитуда $A(x, y)$, фаза или поляризация $P(x, y)$ светового потока.

Транспаранты называют также пространственно-временными модуляторами света (ПВМС), подчеркивая этим присущую оптическому излучению возможность двойной – временной и пространственной модуляции. Применительно к ЭУТ говорят о "формирователях страниц информации" или "устройствах набора страниц". Разработки ПВМС во многом опираются на достижения в создании "обычных" дискретных модуляторов – между этими изделиями имеется функциональное сходство. Следует, однако, подчеркнуть, что ни сами дискретные модуляторы, ни их матричные наборы не в состоянии хоть в какой-то мере заменить ПВМС – интегрированные оптоэлектронные устройства. В разработках ПВМС много пересечений с индикаторными приборами и с реверсными запоминающими средами.

2.4.2 Принцип действия управляемых транспарантов

В работе транспаранта можно выделить 3 фазы: А - запись; Б - считывание; В - стирание. На стадии А реализуется ввод информационного рельефа в активную среду; Б - перенос этого рельефа в информационный поток излучения. В фазе В происходит возвращение транспаранта в исходное однородное состояние.

Если сопоставить объектам нашего рассмотрения некоторые математические функции в виде:

- | | | |
|-----------------------|-----------------|------------------|
| 1) Транспарант | $S_0(x, y)$ | $S_1(x, y)$ |
| 2) Управление | $F_{зан}(x, y)$ | $F_{см}(x, y)$ |
| 3) Модулируемый поток | $\Phi_0(x, y)$ | $\Phi_M(x, y)$, |

то схему работы транспаранта можно представить следующим образом (рисунки 2.8 и 2.9):

- А. $S_0 \xrightarrow{F_{зан}} S_1$
 Б. $\Phi_0 \xrightarrow{S_1} \Phi_M$
 В. $S_1 \xrightarrow{F_{см}} S_0$.

Практически, воздействие $F_{зан}$ приводит к возникновению пространственной неоднородности амплитудной, фазовой, поляризационной характеристик активной среды (каждой в отдельности или всех вместе), т.е.

$$S_1(x, y) = f \{A(x, y), \varphi(x, y), P(x, y)\}, \quad (2.3)$$

так, что при падении плоской немодулированной волны имеем:

$$\Phi_0 = A_0 \exp(-i\omega t) \xrightarrow{S_1} \Phi_M = \overset{P}{P} A_0 A_1 \exp[-i(\omega t + \varphi(x, y))]. \quad (2.4)$$

Перечислим функции, выполняемые управляемыми транспарантами:

1) Основной является модулирующая функция. Применение транспаранта позволяет наиболее плотно "загрузить" световой поток информацией; неиспользованной остается лишь "спектральная координата", т.е. модуляция по длине волны.

2) Усиление светового сигнала. Слабый световой поток $F_{зан}(x, y)$ может управлять высокоинтенсивным лучом $\Phi_0(x, y)$. Коэффициент усиления интенсивности света может достигать от 100 до 1000.

3) Кодировка и декодировка светового потока $\Phi_0(x, y)$ и $\Phi_M(x, y)$. Эти процессы осуществляются в форме пространственной фильтрации оптических сигналов.

4) Перевод исходного информационного массива $F_{зан}(x, y)$ на когерентную несущую Φ_0 для дальнейшей её обработки методами когерентных преобразований.

Фактически транспарант может быть отнесен к простейшим оптическим процессорам, в большей степени это относится к набору транспарантов.

2.4.3 Основные параметры и характеристики

Основные параметры и характеристики ПВМС рассмотрим на примере амплитудного транспаранта, оперирующего с цифровой информацией.

1) Геометрические характеристики:

а) размеры рабочего поля A_x и A_y .

б) разрешающая способность f (мм^{-1}) - число воспроизводимых линий на единицу длины. Если размеры ячейки 100×100 мкм, то $f = 10 \text{ мм}^{-1}$. Разрешающая способность ЭУТ - от 10 до 100 мм^{-1} , ОУТ - до 3000 мм^{-1} . Для записи 1-го бита информации оптическим способом достаточно 1 мкм^2 площади, однако на практике определяющей является разрешающая способность оптической системы.

в) информационная емкость транспаранта $N = Sf^2$, где S - площадь рабочего поля, $S = A_x A_y$. Пусть $f = 100 \text{ мм}^{-1}$, $S = 1000 \text{ мм}^2$, тогда $N = 10^7$ бит. Фактически N есть число ячеек, размещающихся на рабочей апертуре: $N = S/s$, s - площадь ячейки.

2) Эффективность преобразования информационного (светового) потока:

$$\eta = \frac{I_{вых}}{I_{вх}}, \quad (2.5)$$

где $I_{вх}$ и $I_{вых}$ - интенсивности светового потока на входе и выходе.

3) Спектральный диапазон:

а) Диапазон рабочих длин волн $\Delta\lambda_{раб}$ для модулируемого пучка Φ ;

б) Диапазоны управляющих световых воздействий $\Delta\lambda_{зан}$ и $\Delta\lambda_{стипр}$ (для ОУТ).

4) Контраст или глубина модуляции:

$$K = \frac{I^1 - I^0}{I^1 + I^0}, \quad (2.6)$$

где - I^1 и I^0 - интенсивности выходящего потока в открытой ячейке (логическая "1") и в закрытой ячейке (логический "0").

5) Быстродействие: времена переключения состояния ячейки τ_{0-1} и τ_{1-0} из открытого в закрытое состояние (и наоборот) или частота f .

6) Эффективность памяти: время $t_{хр}$ хранения информации без разрушения после окончания записывающего воздействия $F_{зан}$.

7) Нелинейность передаточной характеристики $k = f(F_{зан})$.

Для обработки информации в цифровом виде нелинейность должна быть велика. Для аналоговой обработки, наоборот, необходима линейность передаточной характеристики, лишь при этом условии может быть воспроизведено наибольшее число градаций яркости $\Phi_0(x, y)$.

2.4.4 Конструкции управляемых транспарантов

В качестве примера ЭУТ рассмотрим ПВМС на жидких кристаллах (рисунок 2.8). Жидкие кристаллы (ЖК) обладают электрооптическими свойствами и являются почти совершенным материалом для создания управляемых транспарантов УТ, легко управляемых электрическим полем. Основу такого ЭУТ составляет тонкий слой ЖК (до 10 мкм), который помещают между двумя прозрачными проводящими слоями разделенными на "дорожки". Управляющие напряжения - от 10 до 50 В. В исходном состоянии жидкокристаллическая ячейка совершенно прозрачна. При подаче возбуждающего электрического импульса наблюдаются эффекты двулучепреломления или динамического рассеяния света, т.е. молекулы ЖК ориентируются таким образом, что становятся идеальными (матовыми) рассеивателями. Происходит ослабление Φ_0 в сотни раз. Потребление мощности при этом - 1 мкВт/см². Тип модуляции: амплитудная, как проходящего так и отраженного пучка. Для ЖК ЭУТ характерны: пороговый характер срабатывания, время срабатывания – миллисекунды, емкость 32 x 32 бит, $\eta \sim 40\%$, $K \sim 0.96$.

Конструкцию ОУТ рассмотрим на примере устройства ФОТОТИТУС - аналогового отражательного транспаранта (рисунок 2.10). В качестве активной среды использован электрооптический кристалл DKDP (2). Слой фотополупроводника (ФП) под воздействием света $F_{зан}(x, y)$ становится проводящим. Диэлектриче-

ское зеркало (ДЗ) (4) отражает $F(x, y)$ и $\Phi(x, y)$. Прозрачные электроды (7) служат для подачи управляющих напряжений. $U_{раб} \sim 80$ В.

Изображение предмета (5) с помощью оптической системы (6) проецируется на транспарант со стороны ФП. Управляющий световой поток $F(x, y)$ с записанной информацией экспонирует ФП (3). Освещенные участки ФП становятся проводящими.

Электрическое поле прикладывается к активному кристаллу. Под воздействием поля некоторые области в DKDP становятся двулучепреломляющими - это значит, что они способны управлять поляризацией управляемого светового потока Φ_0 . В отражательном варианте УТ световой поток Φ_0 отражается от ДЗ и еще раз проходит через DKDP тем самым в Φ_0 вносится информация. Параметры системы подбираются таким образом, чтобы на выходе $\Phi_M(x, y)$ имел ортогональную поляризацию. Такой свет можно отделить от Φ_0 , используя поляризационные светоделители (1).

Рассмотренный ОУТ обладает следующими характеристиками. Площадь рабочего поля - $A_x \times A_y \sim 3 \times 4$ см. Разрешающая способность 150 мм^{-1} . Коэффициент контрастности $K = 0,99$. Чувствительность - от 2 до 10 мкДж/см^2 .

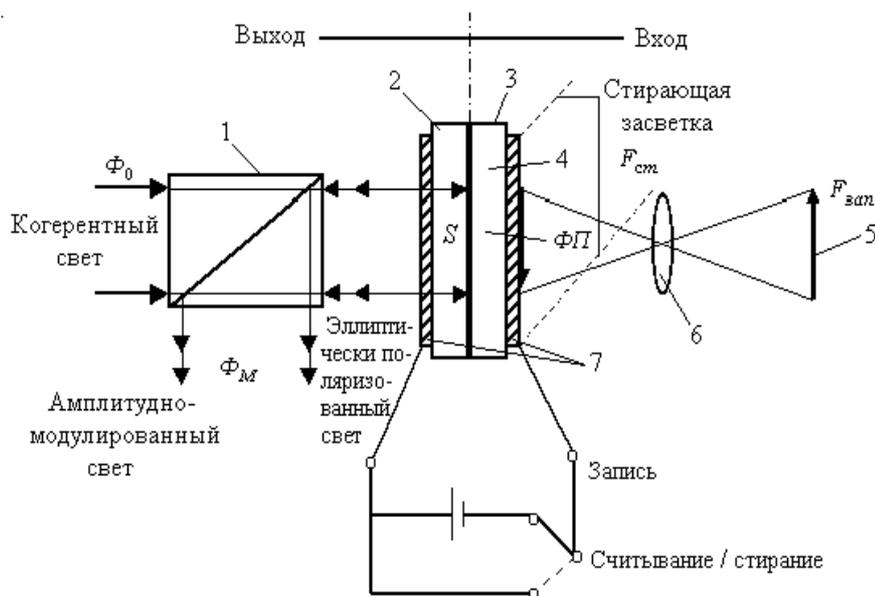


Рисунок 2.10

В качестве физических принципов, используемых в конкретных разработках ПВМС, можно отметить следующие:

1. Изменение показателя преломления рабочего вещества благодаря электрооптическим эффектам Поккельса и Керра.
2. Наведение искусственного двулучепреломления. Вращение плоскости поляризации происходит вследствие электрооптического или магнитооптического эффектов Поккельса, Фарадея, Керра.

3. Изменение поглощения или рассеяния света вследствие фото- или электрохромных эффектов, эффектов гидродинамического рассеяния света.

4. Изменение фотоотражающих свойств. Эффект достигается путём деформации микрорельефа поверхности вещества, термооптических эффектов, пьезоэффектов.

Наиболее часто используются следующие материалы: KDP, DKDP, (KN_2PO_4 , KD_2PO_4), электрооптическая керамика, жидкие кристаллы, фотополупроводники, фотохромные соединения, фототермопластические носители и т.д.

2.5 Трансфазор

Первоначально для создания оптических переключаемых элементов /4/ были предложены инжекционные полупроводниковые лазеры (ПЛ) и оптроны (сочетание СИД и фототранзистора).

Однако разработка интегральной технологии получения ПЛ и объединения их в логические схемы еще находится на начальной стадии и встречает большие трудности. Оптроны как элементная база для построения вычислительных устройств также не могут конкурировать с транзисторами. Потери, связанные с двойным преобразованием энергии обуславливают низкую энергетическую добротность таких устройств. Необходимость спектрального согласования СИД и фотоприемников создают дополнительные проблемы на пути реализации оптических ключевых элементов на базе оптронов.

Сотрудники Эдинбургского университета во главе с проф. Д. Смитом создали оптический аналог электронного транзистора и назвали его трансфазором. Трансфазор переключается при небольшом изменении интенсивности падающего на него лазерного луча. Время переключения экспериментального образца трансфазора составило несколько пикосекунд ($1 \text{ пс} = 10^{-12} \text{ с}$), что на два-три порядка лучше, чем у электронного транзистора.

В основе работы трансфазора лежит свойство некоторых кристаллов изменять показатель преломления при увеличении интенсивности падающего на них света. При создании конструкции трансфазора была заимствована идея, заложенная в интерферометре Фабри — Перо.

Простейший интерферометр Фабри — Перо (рисунок 2.11) состоит из двух плоских зеркал, расположенных параллельно друг другу и разделенных некоторым пространством, именуемым резонатором, которое заполняется веществом, пропускающим свет только определенной длины волны. Расстояние между зеркалами может быть изменено. Зеркала обладают высокой отражательной способностью. Коэффициент отражения R зеркал

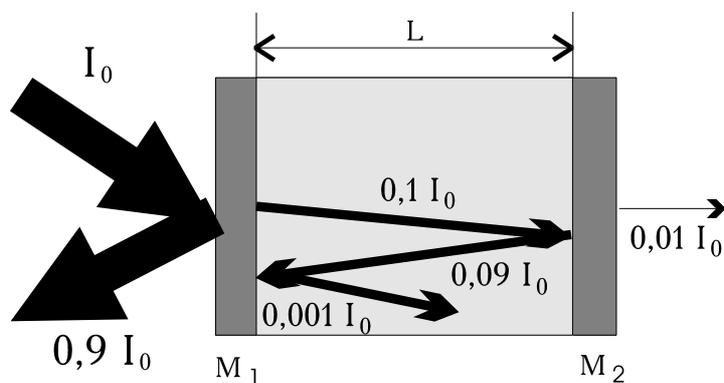


Рисунок 2.11

таков, что при освещении любого из них 90 — 95 % падающего света отражается обратно и только 5 — 10 % пропускается независимо от длины волны. Допустим, что $R = 0,9$. Тогда, если на зеркало M_1 , образующее переднюю стенку интерферометра, падает пучок света интенсивностью I_0 , то пучок, прошедший внутрь резонатора, будет иметь интенсивность, равную $0,1 I_0$. Поскольку зеркало M_2 , обладает теми же свойствами, что и M_1 , выходной пучок света, прошедший через интерферометр, будет иметь интенсивность, равную $0,01 I_0$, т. е. в 100 раз меньшую интенсивности входного пучка. Пучок света, отраженный от задней стенки интерферометра и имеющий интенсивность $0,09 I_0$, многократно отражаясь от зеркал резонатора, будет становиться все слабее и слабее, пока весь свет не выйдет из интерферометра. Каждый проход пучка вносит определенный вклад в выходной пучок, увеличивая его энергию, но несущественно.

Однако прямой и обратный пучки света в резонаторе нельзя разделить так, как это сделано в приведенном схематическом описании. Входной пучок падает на интерферометр перпендикулярно поверхности его зеркал, поэтому все пучки в резонаторе распространяются по одному и тому же пути, взаимодействуя между собой. Результат взаимодействия зависит от соотношения фаз прямого (φ_n) и обратного (φ_0) пучков: если разность фаз кратна $(\varphi_n - \varphi_0) = 2k\pi$, то происходит конструктивная интерференция (пучки в резонаторе усиливают друг друга); если же разность фаз составляет $(2k + 1)\pi$, то происходит деструктивная интерференция (пучки ослабляют друг друга). Может иметь место и любая промежуточная картина. Полная конструктивная интерференция возникает, когда оптическая длина резонатора L равна целому числу полуволин падающего света:

$$L = k\lambda / 2, \quad (2.7)$$

где λ - длина волны падающего света,
 k - целое число.

В этом случае световой пучок, входящий в резонатор, и пучки, образованные вследствие многократного отражения от зеркал внутри резонатора, а также пучок, выходящий из интерферометра, совпадают по фазе. Если $L = (k + 0,5) \lambda/2$, то наблюдается полная деструктивная интерференция.

Интерференция света внутри резонатора существенно влияет на пропускание интерферометра Фабри — Перо:

$$J = I_{\text{вых}} / I_{\text{вх}}, \quad (2.8)$$

где $I_{\text{вх}}$, $I_{\text{вых}}$ - интенсивности входного и выходного световых пучков.

При полной деструктивной интерференции интенсивность выходного светового пучка минимальна:

$$\min I_{\text{вых}} = \left(\frac{1 - R}{1 + R} \right)^2 I_{\text{вх}}. \quad (2.9)$$

В этом случае пропускание интерферометра незначительно. Так, если $R = 0,9$, то интенсивность выходящего из интерферометра пучка $I_{\text{вых}}$ равна $0,0028I_{\text{вх}}$. При полной конструктивной интерференции многочисленных прямых и обратных пучки света усиливают друг друга, складываясь по амплитуде, в результате чего в резонаторе создается поле, интенсивность которого во много раз превышает интенсивность падающего на интерферометр пучка. Интенсивность светового пучка на выходе интерферометра равна интенсивности входного пучка, т. е. максимальное значение $I_{\text{вых}} = I_{\text{вх}}$. Таким образом, изменяя оптическую длину резонатора L , можно менять пропускание интерферометра от незначительного

до единицы: $\left(\frac{1-R}{1+R}\right)^2 \leq J \leq 1$. Зависимость пропускания J интерферометра от длины резонатора L описывается функцией Эйри (рисунок 2.12).

Функция Эйри показывает, что при постепенном приближении L к значению, определяемому соотношением (2.7), пропускание будет меняться медленно до определенного порогового значения, после чего произойдет резкий скачок. Это означает, что интерферометр Фабри — Перо способен резко переходить из одного состояния в другое и может служить основой для создания оптического переключательного элемента. Причем чем больше коэффициент отражения, тем лучше различимы эти состояния (рисунок 2.12).

Весь вопрос заключается в том, как можно быстро изменять оптическую длину резонатора. Это можно сделать двумя путями: изменить расстояние между

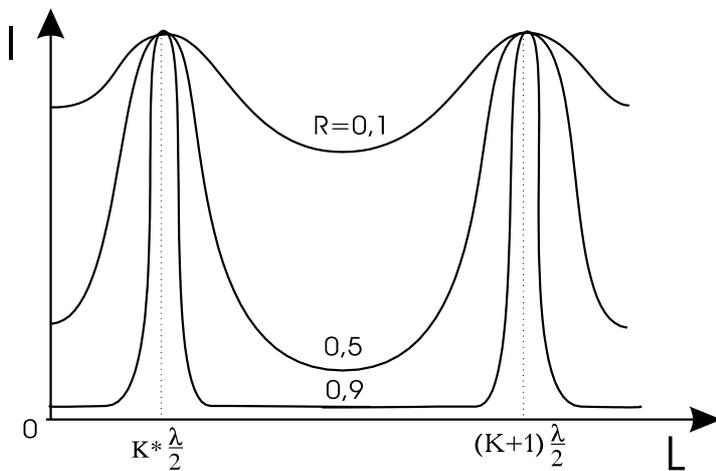


Рисунок 2.12

зеркалами l либо показатель преломления n вещества в резонаторе. Поскольку оптическая длина резонатора $L = ln$ то пригодны оба способа, однако управление изменением длины резонатора не позволит создать быстродействующий прибор.

Долгое время считалось, что показатель преломления вещества не зависит от интенсивности света. С появлением лазеров, способных генерировать мощное когерентное излучение, было показано, что при взаимодействии лазерного излучения с некоторыми веществами показатель преломления последних изменяется в зависимости от интенсивности падающего излучения, т. е. такие вещества имеют нелинейный показатель преломления. Если в резонатор интерферометра Фабри — Перо поместить вещество, обладающее таким свойством, то можно управлять его показателем преломления путем изменения интенсивности входного светового пучка. Данный способ управления пропусканием интерферометра оказался эффективным и был принят за основу при создании трансфазора.

Весь вопрос заключается в том, как можно быстро изменять оптическую длину резонатора. Это можно сделать двумя путями: изменить расстояние между зеркалами l либо показатель преломления n вещества в резонаторе. Поскольку оптическая длина резонатора $L = ln$ то пригодны оба способа, однако управление изменением длины резонатора не позволит создать быстродействующий прибор.

Рассмотрим более подробно, как меняется пропускание интерферометра Фабри — Перо с нелинейным показателем преломления. Допустим, что интенсивность пучка падающего лазерного излучения можно плавно менять в определенном интервале. Пусть интенсивность такова, что оптическая длина L , обусловленная естественным показателем преломления вещества n_0 , имеет значение, лежащее между пиками функции Эйри (рисунок 2.12), а пропускание интерферометра мало. Если медленно увеличивать $I_{\text{вх}}$, то n и L постепенно изменятся, а пропускание J медленно возрастает. При определенном значении интенсивности света в резонаторе, называемом пороговым, между ним и приращением показателя преломления Δn образуется положительная обратная связь — они начинают увеличивать друг друга. Вследствие изменения $n = n_0 + \Delta n$ прямые и обратные световые пучки смещаются по фазе, увеличивая взаимное усиление, интенсивность света в резонаторе резко возрастает, еще более изменяя показатель преломления. На графике функции Эйри это соответствует переходу на крутой участок, где J резко возрастает. Значение L удовлетворяет (2.7), а $J = 1$, что соответствует пику кривой Эйри. Дальнейшее увеличение $I_{\text{вх}}$ сопровождается линейным увеличением $I_{\text{вых}}$.

Если постепенно уменьшать $I_{\text{вх}}$, то пропускание интерферометра некоторое время остается неизменным, пока интенсивность света в резонаторе достаточно велика, чтобы значения n и L удерживались на уровне, соответствующем максимуму пропускания. Наступает момент, когда пропускание начинает плавно уменьшаться. При некотором критическом значении $I_{\text{вх}}$ показатель преломления и интенсивность света в резонаторе начинают резко ослаблять друг друга. Пропускание интерферометра падает до минимума.

Интерферометр Фабри — Перо, в резонатор которого помещено вещество с нелинейным показателем преломления, является оптически бистабильным прибором, так как имеет два устойчивых состояния, в которых интенсивность выходного светового пучка изменяется незначительно при значительном изменении пропускания. Очевидно, что в качестве двоичной или логической единицы можно принять одно из двух устойчивых состояний интерферометра, например, состояние, соответствующее высокому уровню пропускания ($J_{\text{max}} = 1$), а другое — низкому уровню J_{min} . приписать нулю или наоборот. Любое из этих состояний интерферометра может поддерживаться сколь угодно долго с помощью пучка лазерного излучения промежуточной интенсивности.

В переключательных элементах, применяемых при построении операционных устройств, используются устройства, кривая пропускания которых приведена на рисунке 2.13. Ее форма напоминает характеристику тока коллектора электронного транзистора. Кривая пропускания такого вида принимается за основу в оптическом аналоге транзистора. Оптический прибор с такой характеристикой, созданный на основе оптического интерферометра Фабри — Перо, был назван трансфазором, поскольку его работа основана на управлении фазой световой волны, многократно проходящей вещество резонатора.

В первых экспериментальных образцах трансфазора в качестве вещества резонатора были использованы кристаллы антимонида индия InSb, обладающе-

го сильно выраженной нелинейной характеристикой. Из этого кристалла вырезались прямоугольные образцы, которые имели грани размером несколько миллиметров. Передние и задние грани кристаллов покрывались тончайшим слоем отражающего вещества, которые служили зеркалами. Исследования показали, что применение покрытия даже необязательно: полированная поверхность кристалла сама может служить зеркалом. Как видим, устройство трансфазора чрезвычайно просто и технологично.

Кристалл InSb прозрачен только для некоторой области инфракрасной части спектра и лучше всего проявляет свою оптическую нелинейность при температуре 77 К и длине волны излучения, равной приблизительно 5 мкм. Такое излучение может генерироваться химическим лазером на оксиде углерода, излучение которого можно регулировать в узком диапазоне длин волн вокруг $\lambda = 5$ мкм. В последнее время найдены вещества, например кристаллы GaSe, которые проявляют оптическую бистабильность в видимом диапазоне спектра (0,4 - 0,7 мкм), что значительно облегчает выбор лазера и открывает возможность дальнейшего уменьшения размеров трансфазора. Из кристаллов GaSe были изготовлены резонаторы оптической длиной $L = 10$ мкм.

На вход трансфазора подводят два хорошо отъюстированных лазерных пучка, например, с помощью световолокон, которые затем фокусируются в одной точке на его передней грани. Один из пучков имеет относительно большую и неизменную интенсивность I_0 (рисунок 2.13), второй имеет значительно меньшую интенсивность I_y и может модулироваться по интенсивности. Этот пучок является управляющим. Интенсивность постоянного пучка подбирается так, чтобы пропускание трансфазора было близко к порогу усиления. Интенсивность же управляющего пучка такова, что когда он добавляется к постоянному, под действием интенсивности I_m результирующего входного пучка трансфазор переключается в состояние с максимальным пропусканием.

Так как характеристика трансфазора крутая, даже незначительная модуляция управляющего пучка сильно увеличивает его пропускание. Трансфазор работает аналогично электронному транзистору. Действительно, постоянный

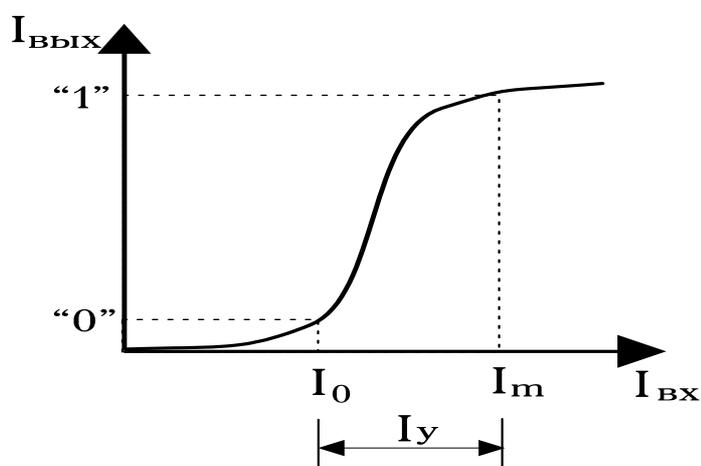


Рисунок 2.13

пучок аналогичен постоянному току смещения, который протекает от эмиттера к коллектору транзистора, а управляющий пучок — меньшему току, протекающему от базы к коллектору. Как небольшое изменение тока базы вызывает в транзисторе большое увеличение тока коллектора, так и незначительное изменение управляющего пучка вызывает в трансфазоре увеличение пропускания. Подобно транзистору, трансфазор может переключаться в одно из двух четко различимых

состояний, только намного быстрее (за время, измеряемое пикосекундами). Быстродействие трансфазора ограничивается временем установления поля внутри резонатора $t' = 2Ln / c$. Для оптической длины резонатора $L = 10$ мкм t' равно 10^{-13} с (0,1 пс). Частота переключения трансфазора зависит от частоты модуляции управляющего сигнала, составляющей порядка 1 ГГц. Трансфазор может быть таким же миниатюрным, как и электронный транзистор. Его поперечные размеры определяются сечением входного лазерного пучка, которое ограничено длиной волны и возможностью фокусирующей системы. Получить пучок диаметром сечения порядка 10λ не представляет трудности. Длина резонатора, как отмечалось, составляет порядка 10 мкм. Для поддержания бистабильного состояния в трансфазоре требуется мощность порядка 10 мВт, а энергия переключения порядка 10^{-15} Дж. Таким образом, энергетическая добротность трансфазора практически может достигать значения 10^{-14} , что на 2-3 порядка лучше, чем у электронного транзистора.

Трансфазор относится к потенциальной системе элементов, поскольку он реализует потенциальный способ представления цифровой информации. Двоичные переменные "1" и "0" кодируются соответственно высоким и низким уровнем пропускания (или, что, то же самое, интенсивностью выходного светового сигнала).

Рассмотрим принципы построения логических элементов на основе трансфазора при реализации простейшей функционально полной системы логических элементов: *И*, *ИЛИ* и *НЕ*. Один и тот же трансфазор может служить как элементом *И*, так и элементом *ИЛИ* в зависимости от подведенных к нему световых сигналов. Если интенсивности входных световых пучков подобраны так, что $I_1 = I_2 = I_m/2$ (если $\varphi_1 - \varphi_2 = k\pi$, то достаточно $I_1 = I_2 = I_m/4$, поскольку пучки взаимно когерентны и интерферируют на входе трансфазора), то образуется элемент *И*, так как трансфазор переключается только при одновременном появлении обоих сигналов (рисунок 2.14 а). Если $I_1 = I_2 = I_m$, то образуется логический элемент *ИЛИ*, поскольку любой из входных световых сигналов способен переключить трансфазор (рисунок 2.14 б). Если в качестве выходного сигнала использован отраженный пучок, то трансфазор работает как элемент *НЕ*.

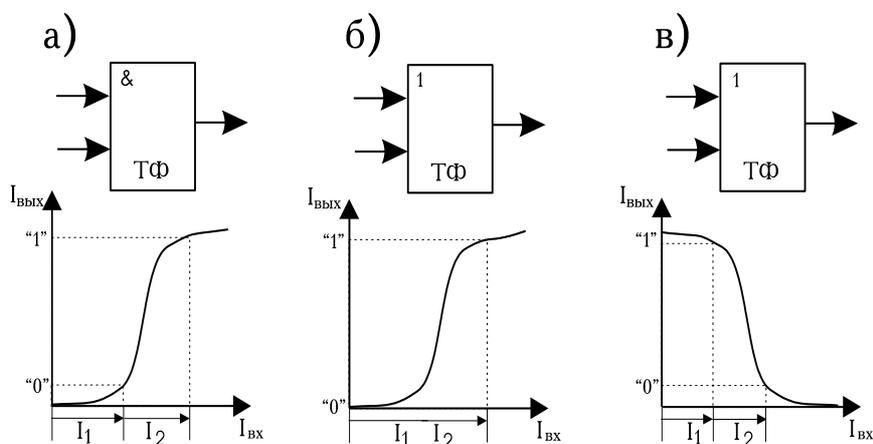


Рисунок 2.14

Действительно, отраженный пучок является инверсией прошедшего пучка, поэтому повышение интенсивности входного пучка до I_m уменьшает выходной сигнал до минимума наоборот. Характеристика и условное обозначение элемента *НЕ* представлены на рисунке 2.14 в.

На этих рисунках ТФ – логический элемент на базе трансфазора.

Из этих логических элементов можно строить любые логические устройства и функциональные узлы вычислительных машин. При этом межэлементные соединения могут быть выполнены с помощью оптических волноводов. Успехи интегральной оптики позволяют надеяться, что интегральная технология получения оптических логических устройств и функциональных узлов на базе трансфазора сможет конкурировать с электронной технологией. Еще не решена проблема межэлементных соединений и подвода световых пучков к кристаллу, который является однородным для всей логической схемы и может быть выполнен в виде тонкой пленки. Так как параллельные пучки света в кристалле практически не взаимодействуют друг с другом, различные участки кристалла, примыкающие друг к другу, могут играть роль резонатора для различных логических элементов. Благодаря этому можно создать огромное число параллельных каналов обработки информации. Преобразование информации в канале осуществляется последовательно от каскада к каскаду, каждый из которых представляет собой пластинку кристалла, связанного с помощью световолокон с соседями, а также источниками когерентного излучения. Такие оптические вычислительные устройства будут представлять собой по существу объемные схемы.

На базе трансфазорных логических элементов может быть построена цифровая оптическая вычислительная машина с архитектурой, подобной архитектуре современных ЭВМ. Преимуществом такой машины было бы сравнительно высокое быстродействие ($\sim 10^9$ оп/с), превышающее быстродействие аналогичных ЭВМ на 2 — 3 порядка. Параллельная организация позволит повысить производительность оптикоэлектронных вычислительных машин еще на 2 — 3 порядка, до 10^{12} оп/с.

Интересно отметить, что трансфазор может иметь более двух устойчивых состояний. Некоторые кристаллы могут выполнять несколько последовательных переключений, сопровождающихся ступенчатым увеличением интенсивности выходного сигнала при ступенчатом росте интенсивности входного сигнала. Следовательно, на базе трансфазора может быть создан многоуровневый логический элемент, что откроет новые перспективы в разработке многозначной машинной логики.

На пути промышленной реализации трансфазора и ИС на его основе возникает ряд трудностей:

1) Необходимость работы при низких температурах. Создание элементов, работающих при комнатной температуре, значительно упростило бы конструкцию и эксплуатацию оптических вычислительных устройств на трансфазорах. Для отдельных материалов и длины волны падающего излучения уже удалось получить оптическое переключение при комнатной температуре, что было достигнуто при больших плотностях мощности излучения, приводящих к преждевременному разрушению резонатора.

2) Стремление к уменьшению мощности, требуемой для переключения элемента, противоречит условию повышения быстродействия, которое зависит

от энергии входного излучения. Поэтому снижения мощности падающего излучения можно добиться лишь путем выбора материала с сильным нелинейным оптическим эффектом. Вопрос создания материалов с необходимыми свойствами является ключевым в проблеме реализации логических схем и вычислительных устройств на основе трансфазора.

2.6 Фотоприемники

2.6.1 Общая характеристика фотоприемников

Фотоприемник – прибор, в котором под действием оптического излучения происходят изменения, позволяющие обнаружить это излучение и измерить его характеристики. Фотоприемник – прибор, в котором имеет место преобразование энергии вида $L \rightarrow E$.

Известно, что по принципу действия все фотоприемники подразделяются на две группы: тепловые, интегрирующие результаты воздействия излучения за длительное время и фотоэлектрические (или фотонные), использующие внешний или внутренний фотоэффект. На основе внешнего фотоэффекта действуют вакуумные фотоэлектрические приборы, на основе внутреннего – твердотельные, главным образом полупроводниковые. Именно эти (последние) приборы удовлетворяют общим концепциям оптоэлектроники и являются пригодными для решения задач информатики. Напомним, что основными концепциями оптоэлектроники являются следующие положения: 1) физическая основа оптоэлектроники – это процессы преобразования энергии $E \rightarrow L$, $L \rightarrow E$, $L \rightarrow L$; 2) техническая основа – микроэлектроника; 3) функциональное назначение – решение задач информатики.

Оптическая информация в оптоэлектронных устройствах практически сводится к двум видам /3/:

- дискретные (во времени, в пространстве и в данной спектральной области) оптические сигналы;
- световые образы, картины. В соответствии с этим и приемники излучения делятся на 2 группы: дискретные и многоэлементные.

К фотоприемникам дискретных сигналов предъявляются следующие требования:

1. Высокая чувствительность на заданной длине волны. Наиболее интересными представляются спектральные области, занятые важнейшими направлениями оптоэлектроники и линии (длины волн) внутри них, генерируемые основными типами лазеров и светодиодов (0,49; 0,51; 0,63; 0,647; 0,69 мкм) – видимая область; (0,8 – 1,55 мкм) близкая инфракрасная область и (3..5 мкм); (8..14 мкм) – "окна прозрачности атмосферы" – специфический раздел оптоэлектроники.

2. Высокий КПД энергетического преобразования. Лишь при этом условии могут быть созданы подлинно микроэлектронные устройства. В противном случае возникает проблема теплоотвода.

3. Высокое быстродействие, обязательное практически для всех устройств генерации, обработки и передачи информации. Современный уровень инерционности характеризуется временами переключения $\tau \approx 10^{-7} - 10^{-10}$ с, для решения перспективных задач информатики требуется продвижение в область $10^{-10} - 10^{-13}$ с. Режим высокоскоростного переключения автоматически ведет к использованию достаточно низких нагрузочных резисторов (в противном случае перезарядка паразитных емкостей оказывается слишком инерционной: $\tau = RC$), а это не позволяет работать с минимальным отношением сигнал/шум. Иными словами, быстродействующие фотоприемники должны использоваться вдали от порога чувствительности, поэтому такие характеристики как обнаружительная способность, уровень шумов оказываются не определяющими.

4. Из требования малой емкости однозначно вытекает и необходимость малой площади фоточувствительной поверхности прибора. Это обстоятельство практически редко является ограничивающим, т.к. дискретные оптические сигналы обычно поступают из малых апертур: лазерный пучок, торец волоконного световода и т. д.

Существенно по иному формируются требования к приемникам световых образов:

1. Многоэлементность. Лишь прибор с большим числом элементарных фоточувствительных ячеек способен воспринимать весь образ сразу; при этом пространственная разрешающая способность восприятия тем выше, чем больше число ячеек.

2. Совместимость фоточувствительной части с электронной схемой считывания (сканирования) информации. Оптимальным является вариант когда обе части устройства конструктивно объединены в виде интегральной схемы (в этом случае говорят о самосканировании или встроенном сканировании).

3. Широкий спектр, например, видимый диапазон.

4. Режим фотонного накопления, в котором могут использоваться рассматриваемые фотоприемники, так как смена воспринимаемых образов достаточно инерционна (время смены кадров в TV-стандарте - $1/25$ с). Это, однако, не означает отсутствия требования к быстродействию каждой элементарной ячейки.

5. Широкий рабочий динамический диапазон, т.к. в любой картине практически соседствуют яркие и бледные области.

6. Минимальный уровень шумов (при сохранении высокой фоточувствительности).

Итак, кратко две основных группы фотоприемников в оптоэлектронике (информатике) характеризуются как:

1) Дискретные (одноэлементные) фотоприемники с малой апертурой (малой площадью активной поверхности), высокой монохроматической чувствительностью в заданной области спектра (в основном в области длин волн $0.4 - 1.55$ мкм) с высоким быстродействием, предназначены для приема коротких оптических импульсов.

2) Многоэлементные фотоприемники с самосканированием, высокой разрешающей способностью, высокой фоточувствительностью и низким уровнем шумов в широкой спектральной области (главным образом, в видимом диапазоне), предназначенные для восприятия световых образов.

Лучшими среди фотоприемников первой группы являются фотодиоды (особенно кремниевые с *p-i-n* структурой и лавинные), среди второй – фоточувствительные приборы с зарядовой связью.

2.6.2 Фотодиоды

***P-n* фотодиоды** (рисунок 2.15). Фотодиоды /5/ представляют собой монокристаллическую полупроводниковую структуру, содержащую две области с различными типами фотопроводимости (*n* и *p* типа). На границе этих областей образуется область объемного пространственного заряда (положительный потенциал в полупроводнике *n*-типа и отрицательный в полупроводнике *p*-типа), называемый *p-n* переходом или барьером, который обладает выпрямляющими свойствами. При освещении этих областей образуются пары носителей заряда - дырки и электроны. Направление тока этих носителей совпадает с направлением обратного тока перехода. Фотодиоды используются в двух режимах: фотодиодном и фотогальваническом. Во втором случае на выходах прибора образуются заряды, разделяемые полем *p-n* перехода, и фотодиод становится источником ЭДС (режим фотогенератора). В первом режиме фотодиод включают по схеме с обратным смещением, и фототок является функцией освещенности (грубо говоря, падает сопротивление перехода). В этом режиме фотодиод обладает рядом преимуществ по сравнению с режимом фотогенератора: пониженная инерционность, большой динамический диапазон линейности зависимости фототока от освещенности.

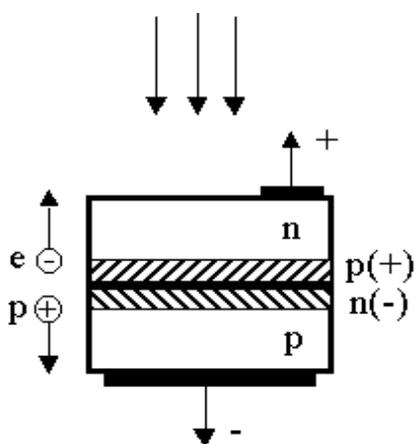


Рисунок 2.15

В *p-n* переходе, на который подается обратное смещение, существует обедненная зона, в которой отсутствуют свободные носители заряда. Фотоны с достаточно большой энергией, попадают в эту зону, выбивают электроны из валентной зоны. Последние переходят в зону проводимости, в результате образуется пара свободных носителей заряда, эти носители образуют фототок. Если электрическое поле в обедненной зоне достаточно сильное, то ускоренные им носители могут приобретать энергию, равную энергии ионизации и выбить другие носители заряда, которые в свою очередь, также ионизируют атомы и т.д. Концентрация носителей будет лавинообразно нарастать, что характерно для лавинных фотодиодов.

Тип фотодиода зависит от материала, конфигурации и структуры *p-n* перехода, а также от приложенного напряжения. Материал определяет ко-

эффицент поглощения света k и квантовый выход. Квантовый выход $\eta_{фд}$ измеряется числом пар (e-p), образованных в результате поглощения одного фотона, т.е. по существу является КПД фотоприемника. Основной характеристикой фотодиода является чувствительность $S_{фд}$, измеряемая полным КПД преобразования световой мощности в электрический ток. $S_{фд}$ прямо пропорциональна $\eta_{фд}$; она также зависит от длины волны λ падающего излучения.

В качестве материала для фотодиодов наиболее широкое применение находят три полупроводника: кремний, германий и арсенид галлия. Германиевые фотодиоды работают при $\lambda < 1,6$ мкм, кремниевые при $\lambda < 1,1$ мкм, а фотодиоды из GaAs при $\lambda < 0,9$ мкм. Наиболее распространены кремниевые фотодиоды, которые достаточно хорошо отработаны в производстве, обладают удовлетворительными характеристиками, высокой надежностью и довольно дешевы.

p-i-n фотодиоды. Фотодиоды на основе обычных p-n переходов с размерами фоточувствительной площадки в несколько квадратных миллиметров обладают недостаточно высоким быстродействием (постоянная времени τ не более $10^{-5} - 10^{-6}$ с), что ограничивает их применение во многих областях информатики. Постоянную времени можно существенно уменьшить путем создания p-i-n – структуры (рисунок 2.16). Последняя создается за счет образования сильнолегированных областей (p и n) с двух сторон пластинки из высокоомного кремния (сопротивление в $10^6 - 10^9$ раз больше сопротивлений n- и p- областей). При приложении к прибору напряжения обратного смещения, в i- области создается высокая напряженность электрического поля. Область пространственного заряда растягивается на всю ширину i- области и в ней происходит увеличение скорости дрейфа носителей

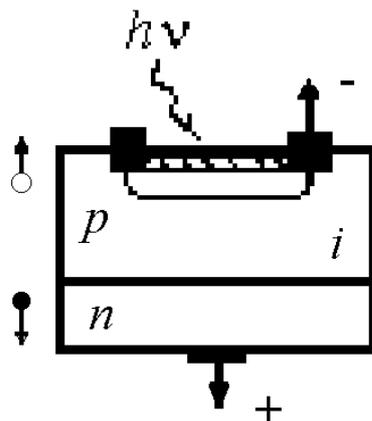


Рисунок 2.16

заряда, образовавшихся под действием света до 10^7 см/с. Напряженность электрического поля при этом достигает $E \sim 10^3$ В/см. Эту область называют пространством дрейфа.

За пределами обедненного слоя движение носителей носит диффузионный характер с относительно низкой скоростью - примерно 10^4 см/с. Это обстоятельство ухудшает быстродействие. Для его повышения необходимо сконцентрировать поглощение излучения в обедненном слое. С этой целью слой p^+ делают очень тонким, а толщину слоя i - больше длины поглощения излучения ($1/k$). Длина поглощения для кремния на длине волны 0.8 мкм равна примерно 10...20 мкм и рабочее напряжение, при котором обедненный слой имеет требуемую ширину, не превышает 10-20 В.

Лавинные фотодиоды. Лавинный фотодиод представляет собой фотодиод, предназначенный для использования в режиме лавинного умножения фототока. Практически на фотодиод подают обратное напряжение, близкое к напря-

жению лавинного пробоя: генерируемые в базе фотоносители диффундируя или дрейфуя, достигают области сильного поля (10^5 В/см), в которой происходит их лавинное размножение. Коэффициент усиления может достигать 1000. При напряжении 100-150 В, быстродействие лавинных фотодиодов оказывается равным $\sim 0,3$ нс. Иными словами, в лавинных фотодиодах имеется внутреннее усиление, а поскольку лавина развивается очень быстро, это усиление не сопровождается снижением быстродействия.

Быстродействие ограничено временем пробега фотовозбужденных носителей и постоянной времени $\tau = RC$, где R - сопротивление нагрузки (R_n), C - емкость диода. Время пробега носителей при скорости 10^7 см/с и ширине обедненного слоя 100 мкм - примерно 1 нс. При меньшей ширине обедненного слоя может быть достигнута граничная частота до нескольких гигагерц. Емкость диода - 1..2 пФ (сумма паразитной емкости корпуса и емкости перехода). Если сопротивление нагрузки принять равным 50 Ом, то постоянная времени $\tau = 0,05...0,1$ нс.

Возникает, однако, вопрос, для чего необходимо внутреннее усиление, если лавинные фотодиоды также как и $p-i-n$ - фотодиод, подключаются к электронной схеме, а ее усиление может быть сколько угодно большим. Почему недостаточно практически идеального $p-i-n$ фотодиода, квантовая эффективность которого близка к 100%, а шумы очень малы? Дело заключается в том, что шумы входных каскадов усилителей фотосигналов обычно на несколько порядков превосходят шумы качественного $p-i-n$ фотодиода. Поэтому даже при оптимальном сопряжении с электронной схемой реализовать все потенциальные возможности фотодиода не удастся, его очень низкий порог чувствительности остается "вещью в себе".

Введение внутреннего усиления (воздействующего и на полезный сигнал и на шум) приводит к возрастанию отношения сигнал/шум при условии, что усиленный шум фотоприемника остается меньше шумов входного каскада усилителя. Иными словами, лавинные фотодиоды имеют преимущество перед $p-i-n$ фотодиодами в условиях приема слабых минимально различимых сигналов.

2.6.3 Фоточувствительные приборы с зарядовой связью

Фоточувствительный прибор с зарядовой связью (ФПЗС) представляет собой специфическую фоточувствительную МДП-микросхему с регулярной системой электродов, расположенных на поверхности диэлектрика столь близко друг от друга, что существенным становится их взаимовлияние, т. е. перекрытие электрических полей соседних электродов внутри кристалла полупроводника /3/.

Электроды располагаются в виде линейки (строки) или матрицы, типичные размеры электрода: длина $Z = 40$ мкм; ширина $L = 5$ мкм; зазоры между ними $l < 1...2$ мкм; число электродов обычно $5 \cdot 10^2 ... 2 \cdot 10^3$ в линейном и $10^4 ... 10^6$ в матричном ФПЗС. Электроды изготавливаются из алюминия или

(для увеличения прозрачности) из поликремния. Строка от строки матричного ФПЗС отделяется узкими областями стоп-канальной диффузии.

Функционально ФПЗС — это прибор, воспринимающий изображение, осуществляющий его разложение на элементарные фрагменты, поэлементное электронное считывание (сканирование) и формирование на выходе видеосигнала, адекватного изображению. Принцип действия ФПЗС рассмотрим на примере классической трехтактной (трехфазной) схемы управления (рисунок 2.17). В этом случае элементарная ячейка ФПЗС включает три соседних электрода 1—3 одной строки. В течение первой фазы (рисунок 2.17 а) к электроду 2 прикладывается положительное напряжение хранения (обычно $U_{xp} = 10...20$ В).

Благодаря возникающему электрическому полю дырки (основные носители) оттесняются в глубь полупроводника и у поверхности образуется обедненный слой (ОС) глубиной $x_{o,c} = 0,5...2$ мкм, представляющий собой на энергетической диаграмме потенциальную яму для электронов. Освещение ФПЗС порождает в объеме полупроводника электронно-дырочные пары, при этом электроны втягиваются в потенциальную яму и локализуются в тонком (около 10 нм) приповерхностном слое. Перемещение электронов осуществляется путем дрейфа (внутри ОС) или диффузии (вне ОС). Накопление электронов ведет к образованию зарядового пакета, который определяется локальной интенсивностью и временем засветки (типичная объемная плотность электронов в пакете около 10^{18} см⁻³). Зарядовый пакет может относительно долго (1 ... 100 мс) сохраняться, однако постепенно термогенерация электронов объемными и поверхностными ловушками приводит к искажению хранимой информации.

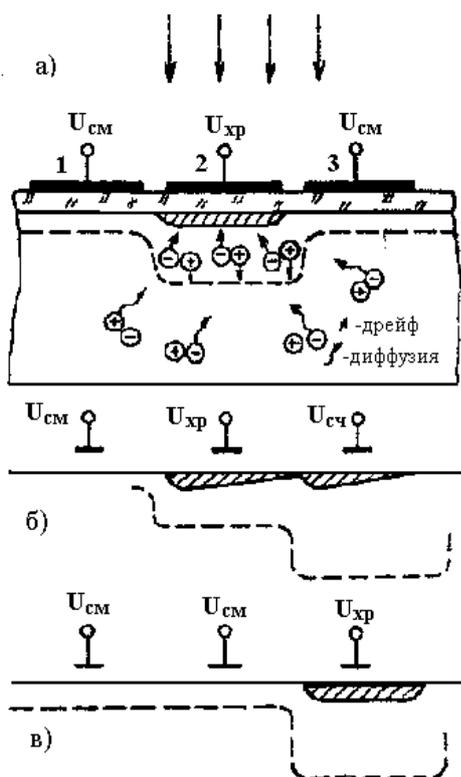


Рисунок 2.17

Во время второй фазы к электроду 3 прикладывается напряжение считывания $U_{sc} > U_{xp}$; вследствие близости электродов 2 и 3 барьер между ними исчезает и зарядовый пакет перетекает в более глубокую потенциальную яму (рисунок 2.17 б). Здесь также имеют место дрейф (между электродами и вблизи их краев) и диффузия электронов (под электродами). Понятно, что и на этой фазе происходит некоторая потеря информации: часть электронов зарядового пакета рекомбинирует при взаимодействии с поверхностными ловушками, а часть пропадает вследствие неполного перетекания зарядов.

Во время третьей фазы напряжение на электроде 3 снижается до U_{xp} , а с электрода 2 потенциал снимается (рис. 2.17 в). Заметим, что на электродах, к которым не приложены напряжения U_{xp} или U_{sc} , все время поддерживается

небольшое напряжение смещения $U_{см} < U_{xp}$, обеспечивающее обеднение всей поверхности полупроводника и ослабление поверхностной рекомбинации. Итак, зарядовый пакет перенесен на один шаг вправо и подготовлен к новому перемещению. Электрод 1 в этом процессе играет роль буфера, иначе слева от электрода 2 оказался бы 3-й электрод предыдущей ячейки и во втором такте зарядовый пакет равновероятно мог бы перетекать как вправо, так и влево.

Таким образом, в ФПЗС пространственное распределение интенсивности излучения преобразуется в рельеф электрических зарядов, локализуемых в приповерхностной области. Зарядовые пакеты перемещаются от элемента к элементу, выводятся наружу и дают последовательность видеоимпульсов, адекватную полю излучения, т. е. осуществляется стандартный телевизионный алгоритм восприятия образа.

2.6.4 Фотоприемные матрицы

Матрица фотоприемников (МФП) служат для преобразования информации, поступающей от оптического процессора, в электрические сигналы [4]. При использовании фотоприемных матриц в ЗУ число ее элементов (фотоприемников) соответствует объему страницы, причем каждый элемент функционирует как пороговый детектор (дискретного типа), указывающий наличие или отсутствие светового сигнала в соответствующей позиции.

Различают два режима работы фотоприемных элементов:

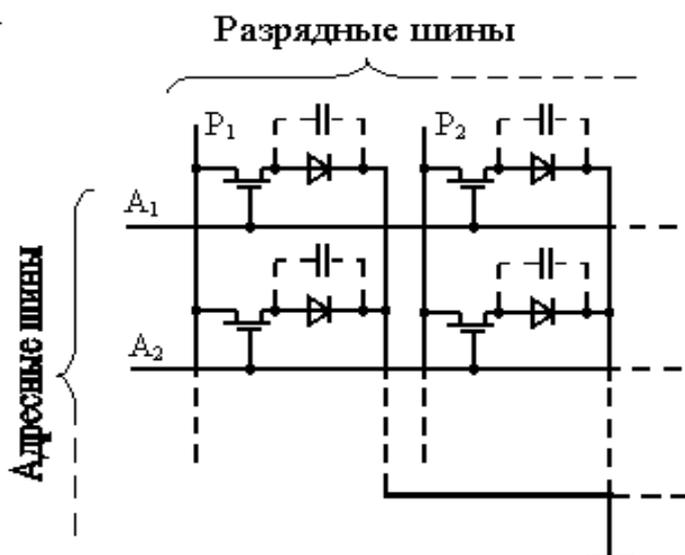
1. Режим непосредственного отсчета. В этом случае выходной электрический сигнал фотоприемника в каждый момент времени пропорционален интенсивности падающего на него оптического сигнала.

2. Режим работы с накоплением заряда. В этом случае выходной электрический сигнал пропорционален полному световому потоку, падающему за время накопления. Т.к. мощность оптического сигнала, поступающего на вход отдельно элемента ФПМ очень мала, то второй режим работы предпочтительней. При разработке фотоматрицы в последнее время наблюдается тенденция объединения фотоприемников с элементами транзисторной памяти. При этом к вы-

ходным сигналам ФП предъявляется единственное требование - устанавливать триггерную схему обработки принимаемой информации в нужное состояние.

В качестве примера ФП, элементы которой работают в режиме накопления заряда, рассмотрим следующую схему (рисунок 2.18).

Каждый ФД матрицы ($p-i-n$ типа) включен последова-



тельно с ключевым МОП-транзистором.

1) В начале цикла работы МФП на затворы МОП транзисторов подаются коммутирующие импульсы и через открытые транзисторы, производится заряд емкостей $p-n$ переходов фотодиодов до максимального напряжения источника питания.

2) Матрица освещается оптическим изображением страницы. При этом под действием падающего светового потока происходит разряд емкостей $p-n$ переходов диодов и напряжение на них падает на значение, пропорциональное мощности светового потока и длительности освещения (экспозиции).

3) Для считывания требуемого слова на соответствующую адресную шину подается коммутирующий импульс, который отпирает ключевые транзисторы, соединенные с выбранной шиной. Тогда через открытые транзисторы, фотодиоды и входные цепи усилителей считывания потекут токи дозаряда. Ток, протекающий через отдельный фотодиод, зависит от потери заряда его емкостью за время экспозиции и пропорционален числу фотонов света, попавших на фотодиод за время накопления.

4) Выделение полезного сигнала из помех производится усилителями и дальнейшей схемой. Т.е. требуемое слово оказывается считанным.

Основным достоинством фотоприемной матрицы с накоплением заряда является очень высокая чувствительность $\sim 10^{-14}$ Дж/бит. Кроме того, она обладает свойствами памяти с ограниченным временем хранения информации и однократным считыванием.

3 Интегрально-оптические элементы

3.1 Общие положения

Интегральная оптика - раздел оптоэлектроники, изучающий и использующий на практике процессы генерации, распространения и преобразования оптического излучения в тонкопленочных волноводах и других пленочных оптоэлектронных элементах, совместно образующих миниатюрные твердотельные устройства - *оптические интегральные схемы* [3].

Создание интегрально-оптических систем, содержащих тонкопленочные генераторы излучения, тонкопленочные волноводы, модуляторы, дефлекторы, направленные ответвители, фотодетекторы и другие элементы, выполненные на базе одного технологического процесса и на общей подложке, открывает широкие перспективы микроминиатюризации оптико-электронных приборов и устройств, расширяет их функциональные возможности, открывает принципиально новые пути создания ЭВМ, оптических систем связи и систем оптической обработки информации.

Непосредственным толчком к зарождению интегральной оптики послужили проблемы, выдвинутые развитием оптико-электронных приборов. Любой оптико-электронный прибор состоит из ряда активных и пассивных элементов: источников света, фотоприемников, модуляторов, фильтров, линз, зеркал, призм и др., расположенных на общем основании. Вполне естественен вопрос, а

нельзя ли все эти элементы объединить в одно целое, используя общую подложку и единый технологический процесс подобно тому, как объединяются элементы в интегральной электронной микросхеме? Для такого объединения оптических элементов в одну интегрально-оптическую микросхему нужен элемент, который бы обеспечил направленное распространение светового излучения без потерь и искажений и который можно было бы изготовить методами интегральной технологии. Таким элементом является *плоский диэлектрический волновод* - основной элемент интегральной оптики. Волноводные структуры могут быть сформированы на поверхности или в объеме диэлектрической подложки технологическими методами, близкими к методам создания электронно-интегральных микросхем. Размещая на той же подложке источник и приемник излучения и элементы, служащие для преобразования оптического излучения: модуляторы, дефлекторы, смесители и другие, можно создавать малогабаритные интегрально-оптические устройства с повышенной надежностью, малой потребляемой мощностью и низкими управляющими напряжениями 1-10В.

Имеется два основных типа оптико-интегральных схем. К первому типу относятся монокристаллические схемы, в которых все элементы выполнены из одного материала. Так как большинство схем требуют источников света и приемников, которые могут быть изготовлены из активных материалов, то монокристаллические схемы изготавливаются из таких материалов как полупроводники. Ко второму типу относятся гибридные схемы, в которых два или несколько видов материалов объединяются для оптимизации характеристик устройства. Гибридный подход удобен широким спектром выполняемых схемой операций, но обладает тем недостатком, что из-за тепловых деформаций и вибраций возможно рассогласование элементов и даже отказ схемы.

Реально оптическая интегральная "схема" (рисунок 3.1) изготавливается на

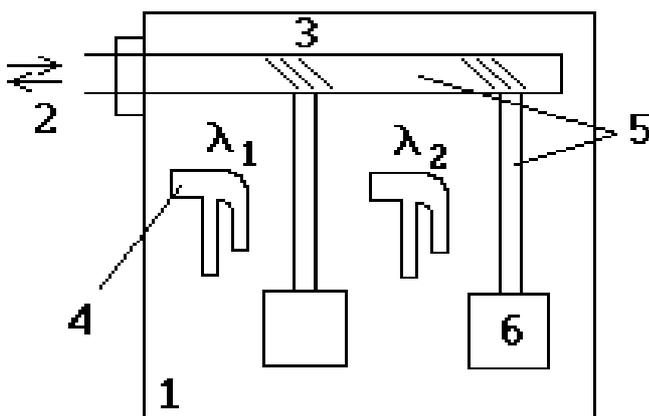


Рисунок 3.1

плоской подложке 1, площадью в несколько квадратных сантиметров. Ее тонкопленочные элементы располагаются на поверхности подложки (или непосредственно над поверхностью). Плоские волноводы 5 составляют основу любой оптической интегральной схемы, а функционально они играют роль проводников и являются базой для создания других: активных и пассивных элементов. В определенном смысле можно считать, что вся схема состоит из волноводов. На рисунке: 3 - спектральные фильтры (отражатели), 2 - устройства ввода-вывода, 4 - встроенные лазеры, 6 - приемники излучения.

Количественным критерием, по которому пленочный элемент относится к сфере интегральной оптики, является соизмеримость толщины пленки и длины

световой волны: лишь при этом возникают качественные отличия от традиционной ("объемной") оптики.

Физическую основу интегральной оптики составляют волновые процессы в плоских оптических волноводах, обычно диэлектрических, иногда полупроводниковых; в этой связи правомерен термин "волноводная оптика", который нередко используется в публикациях расчетно-теоретического характера.

Конструктивно-технологические особенности нового направления связаны с расположением пленочных элементов в одной плоскости (в плане) – отсюда происходит термин "планарная оптика", также получивший некоторое распространение. При его употреблении обычно имеется в виду не только указанный геометрический фактор, но и то, что при изготовлении интегрально-оптических элементов широко используются процессы планарной технологии микроэлектроники: эпитаксия, диффузия и другие.

Приборную сущность нового направления определяет интеграция функционально различных оптических и оптоэлектронных элементов в единое устройство, в схему; отсюда и термин "интегральная оптика", получивший наиболее широкое распространение.

Оптические волноводы. Оптические волноводы представляют собой тонкие диэлектрические слои с малыми потерями на рабочих длинах волн.

Размеры слоев сравнимы с длиной волны света λ , поэтому точность их изготовления должна быть очень высокой. Толщина волноводных слоев 0,3 - 3 мкм, ширина 1 - 5 мкм.

Планарный волновод представляет собой пленку с показателем преломления n_2 , заключенную между двумя слоями из материалов с меньшими значениями n : n_1 и n_3 . В зависимости от характера изменения n планарные волноводы бывают ступенчатыми и градиентными - с плавным изменением показателя преломления.

На рисунке 3.2 представлены различные типы волноводных каналов.

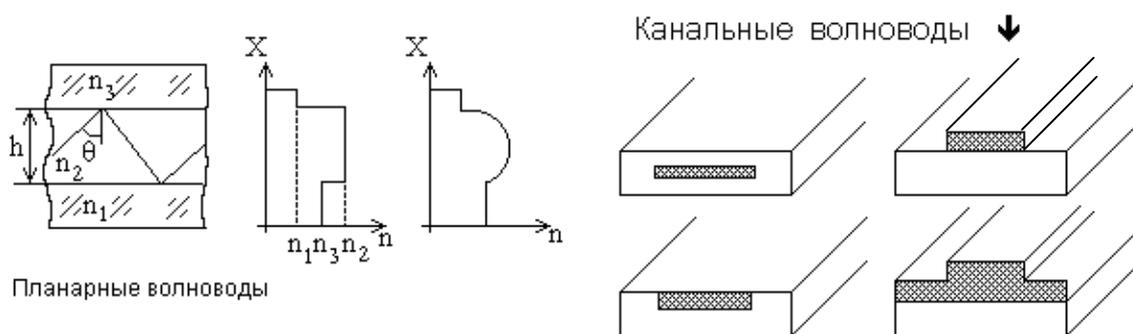


Рисунок 3.2

Методы изготовления. Ионный обмен: суть этого метода заключается в том, что при погружении стеклянной подложки в расплав солей, содержащих ионы тяжелых металлов, происходит частичное замещение ионов стекла более тяжелыми ионами расплава. В результате показатель преломления поверхност-

ного слоя стекла возрастает. Осуществляя ионообменные реакции в электромагнитном поле, можно регулировать как последовательность внедрения различных ионов в стекло, так и глубину их проникновения, т.е. создавать на поверхности стекла слои заданной толщины с различными показателем преломления. Используя специальные маски, через окна в которых происходит ионный обмен, можно получать волновые структуры разной конфигурации.

В кварцевом стекле волноводы обычно создаются методами ионной имплантации - облучением ионами, разогнанными электрическим полем до больших энергий.

Наряду со стеклом в интегральной оптике находят применение структуры из различных оксидов и нитридов, получаемые другими методами тонкопленочной технологии: вакуумным распылением, химическим осаждением, термическим окислением и т.д. Используются также и полимерные пленки, осаждаемые из растворов или получаемые другими методами.

Материалы. Из материалов, для изготовления волноводов и других элементов интегральной оптики, кроме стекол, активно применяются электрооптические материалы: кристаллы ниобата лития (LiNbO_3) или танталата лития.

Из полупроводниковых материалов для интегральной оптики наиболее перспективны соединения типа $A_3 B_4$ (соединения 3 и 4 групп таблицы Менделеева) GaAs. Из них создаются как пассивные (волноводы), так и активные элементы: инжекционные тонкопленочные лазеры, фотоприемники, различные управляющие и другие устройства.

3.2 Ввод-вывод излучения из волновода

При практическом использовании диэлектрических пленочных волноводов излучение источника света (обычно лазера) необходимо ввести в волновод или наоборот вывести из него. Для решения этой задачи разработаны различные устройства ввода (вывода). Связь источника излучения, например лазера, с волноводом может быть осуществлена с помощью линзы (рисунок 3.3).

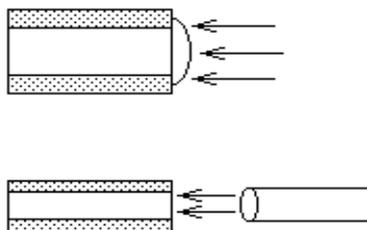
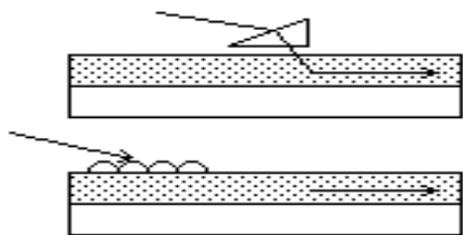


Рисунок 3.3

При линзовой системе связи нужна весьма тонкая фокусировка излучения на торец волновода, так как даже незначительное смещение луча относительно торца резко уменьшает эффективность ввода (долю вводимого излучения). Тре-

бования к фокусировке несколько ниже, если линза расположена вплотную к волноводу или к источнику. Связь с помощью линз называется поперечной.

В интегральной оптике широко применяется и продольная связь (рисунок 3.4). Такую связь можно осуществить с помощью призмы, помещенной над пленкой на расстоянии менее длины волны света. При постоянной толщине воздушного зазора (используя эффект полного внутреннего отражения) в пленку можно ввести до 80 % лазерного излучения.



ния.

Рисунок 3.4

Однако получение такого зазора достаточно сложно. Для ввода-вывода света используют также периодические решеточные структуры. На пленочном волноводе создаются решетки из линий, перпендикулярных к направлению волны. Связь основана на дифракции света на решетке и на существовании определенных мод в волноводе.

Проблема ввода-вывода – одна из ключевых для интегральной оптики; поиски оптимального ее решения продолжаются.

3.3 Оптическое туннелирование и волноводная дисперсия

Особенность интегрально-оптических устройств заключается в разнообразных проявлениях эффекта оптического туннелирования, суть которого состоит в том, что между двумя близко расположенными (но не соприкасающимися) волноводами может происходить когерентная перекачка энергии. На основе этого эффекта созданы специфические элементы интегральной оптики - *направленные ответвители* - устройства, обеспечивающие обмен мощностью между волнами распространяющихся в параллельных волноводах. На рисунке 3.5 (а) представлена схема такого устройства.

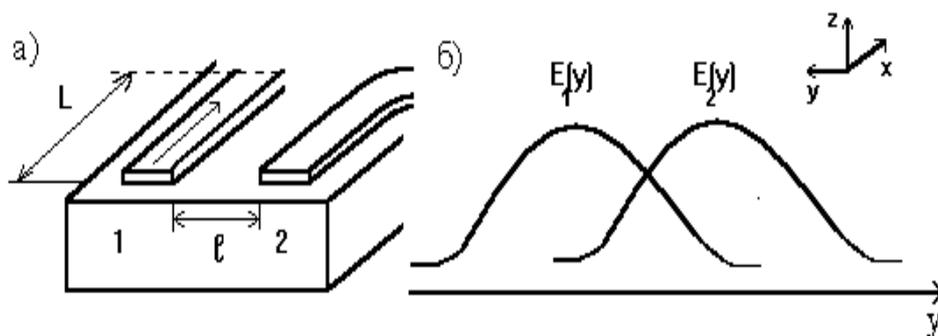


Рисунок 3.5

Здесь используется эффект связи между волноводами 1 и 2, расположенными близко друг к другу на расстоянии l ($l \leq \frac{\lambda_p}{4}$). В зависимости от параметров волноводов и характеристик волн, можно добиться вполне определенной

доли ответвляемой в соседний волновод интенсивности. На рисунке 3.5 (б) представлено распределение амплитуд электрического поля излучения, распространяющегося по соседним волноводам.

В интегрально оптическом исполнении могут быть и пассивные оптические элементы типа линз и призм. При создании этих элементов локальное изменение показателя преломления, а следовательно и хода световых лучей достигается локальным изменением толщины элемента. Известно, что при распространении света в объеме какого-либо слабо поглощающего вещества скорость света $v = \frac{c}{n}$ зависит от его показателя преломления n . В тонкопленочных же

волноводах $v = \frac{c}{n^*}$, где n^* - фазовый коэффициент замедления волны. n^* зависит от материала волновода и граничащих с ним сред, а также от типа распространяющегося по волноводу колебания (моды). При переходе из тонкопленочного волновода с коэффициентом n_1^* в волновод с коэффициентом n_2^* изменяется направление распространения света, т.е. свет преломляется. Закон преломления света аналогичен закону Снеллиуса:

$$\frac{\sin(i_1)}{\sin(i_2)} = \frac{n_2^*}{n_1^*} = n^* = f(d, n).$$

Коэффициент замедления n^* изменяется в зависимости от толщины пленки (явление "волноводной дисперсии" - чем толще пленка, тем больше n^*), и от ее показателя преломления n . Поэтому, если волновод выполнен из одного материала, но имеет разную толщину, то при переходе света из области с толщиной d_1 в область с толщиной d_2 , направление распространения света изменяется (рисунок 3.6 а).

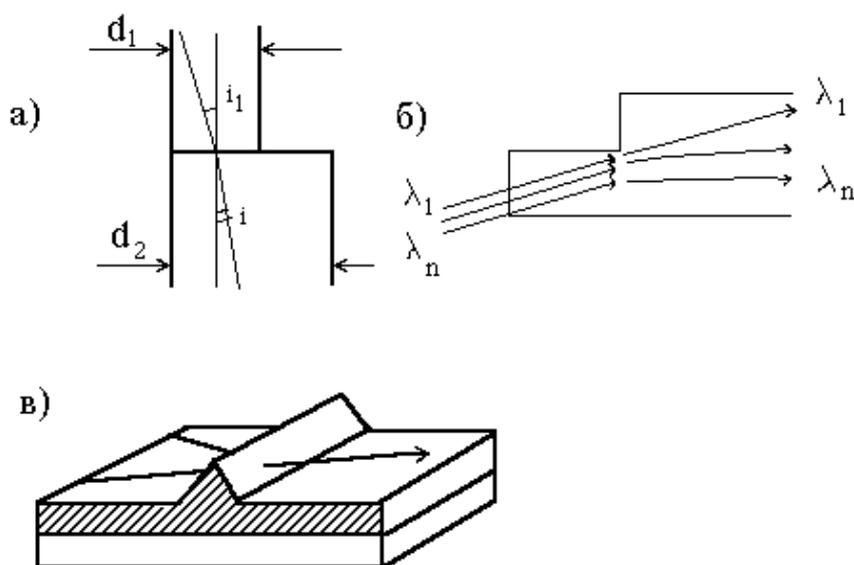


Рисунок 3.6

Так будет работать и волноводная призма (рисунок 3.6 в). По этой же причине (как и при падении белого света на границу раздела двух сред, где в преломленном пучке происходит пространственное разделение колебаний с различными длинами волн (разложение в спектр)), при распространении белого света в ступенчатом волноводе с разным коэффициентом n^* происходит пространственное разделение цветов (рисунок 3.6 б).

3.4 Интегрально-оптические линзы

Для фокусировки излучения в интегральной оптике используются системы с плавно меняющимися параметрами - тонкопленочные линзы. Первые такие линзы создавались путем с изменением толщины волновода на участке, имитирующем по форме обычную объемную линзу (рисунок 3.7 а).

Линза Люнеберга (рисунок 3.7 б) представляет собой подложку с показателем преломления n_4 , на которую нанесены два волноводных слоя с показателями преломления n_2 и n_3 . Верхний волноводный слой имеет круговую сим-

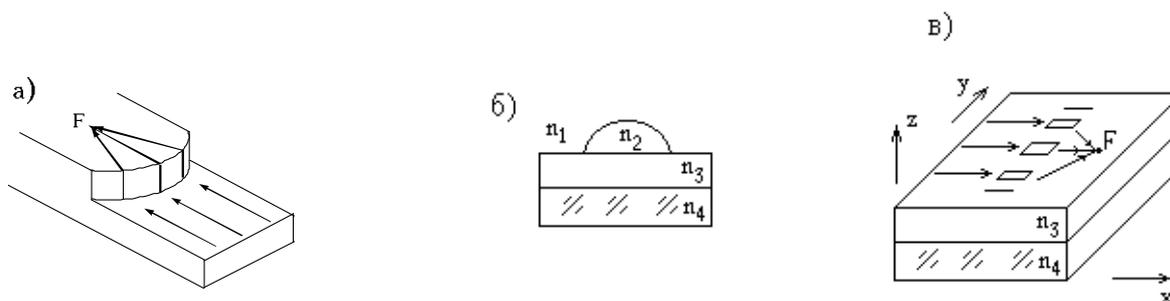


Рисунок 3.7

метрию и переменную толщину за счет чего меняется коэффициент замедления.

Линза Френеля (рисунок 3.7 в) состоит из подложки с показателем преломления n_4 , волновода с показателем преломления n_3 и отрезков пленки с показателем преломления n_2 . Фокусирующие свойства такой линзы определяются процессами дифракции на периодической структуре с переменным периодом.

3.5 Активные элементы интегрально-оптических схем

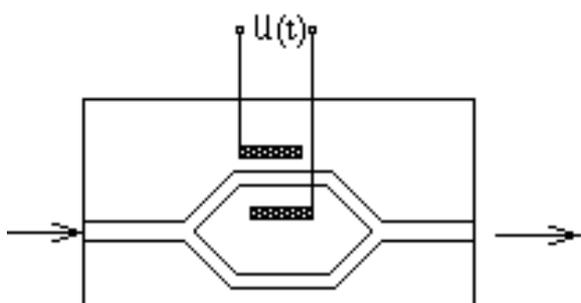


Рисунок 3.8

Интерференционно-волноводный модулятор (рисунок 3.8). В интерференционно-волноводном модуляторе входной оптический поток разводится по двум идентичным каналам, разделенным

в пространстве. При подаче на электроды управляющего сигнала $U(t)$ изменяется показатель преломления (волноводы изготовлены из электрооптического материала, обычно ниобата лития) вследствие чего возникает фазовый сдвиг между волнами в 2-х каналах. На выходе в результате интерференции этих волн фазовая модуляция преобразуется в модуляцию интенсивности суммарного потока. Используемая здесь структура представляет собой интерферометр Маха-Цандера – элемент очень широко распространенный в интегральной оптике. В общем случае изменение оптической длины в одном из плеч интерферометра может осуществляться не только при воздействии электрического поля, но и света, температуры, давления....

Оптический переключатель 4x4 (рисунок 3.9) по конструкции подобен ответвителю, но изготавливается из электрооптического материала и снабжен металлическими электродами; при подаче сигналов на электроды изменяется коэффициент связи, и следовательно, доля излучения, переходящая в другой канал. Основные характеристики переключателя: время переключения $T=10^{-12}$ с, управляющие напряжения 0,1-10 В, полоса частот до 10^{10} Гц.

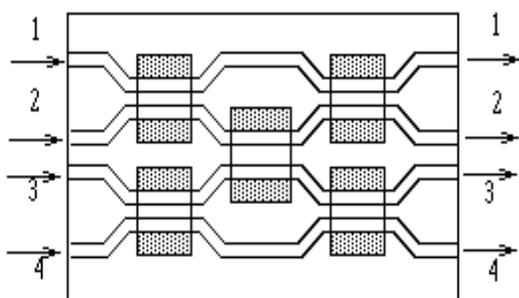


Рисунок 3.9

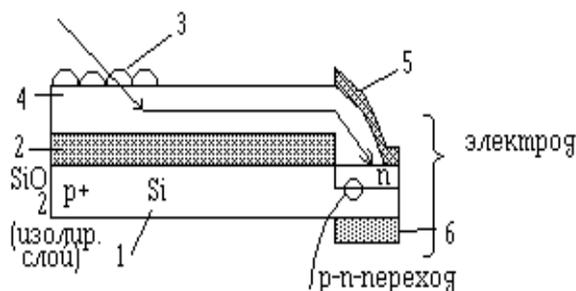
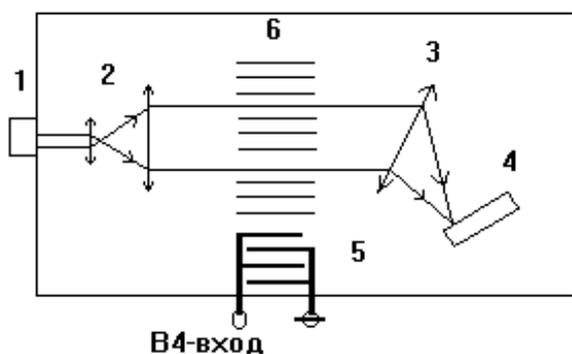


Рисунок 3.10

Интегрально-оптический фотодиод. В этой схеме кремниевый фотодиод p^+n , объединен с волноводом (рисунок 3.10). Свет вводится в стеклянный тонкопленочный волновод 4 с помощью решетчатого элемента связи 3 и попадает в кремниевую подложку в том месте, где обрывается изолирующий слой 2 диоксида кремния (SiO_2). В области волноводного соединения сформирован p - n переход, играющий роль приемника света. Выходной сигнал снимается с электродов 5 и 6.

Интегрально-оптический анализатор спектра. (рисунок 3.11). В интегрально-оптическом спектроанализаторе высокочастотный радиосигнал возбуждает в планарном волноводе по-

верхностную акустическую волну (ПАВ), на которой дифрагирует распространяющийся по этому волноводу лазерный луч. При этом каждой частоте радиосигнала соответствуют



В4-вход

определенная длина акустической волны и угол отклонения дифрагируемого луча. Отклоненный луч попадает на определенный элемент в линейке фотоприемников. Особенность этого устройства, как и других интегрально-оптических устройств, - очень широкая полоса частот анализируемых сигналов (выше 1 ГГц). Цифрами на рисунке обозначены:

1. лазер;
2. линзовый расширитель луча;
3. фокусирующая линза;
4. линейка фотоприемников;
5. возбудитель ПАВ;
6. ПАВ.

Рисунок 3.11

4 Лазеры

4.1. Принцип работы лазера

4.1.1 Энергетические уровни атомов и молекул

Энергия атома или молекулы может принимать лишь некоторые определённые (дискретные) значения /6/. Этим значениям соответствуют энергетические уровни. Переход атома (молекулы) с одного энергетического уровня на другой совершается скачком; это есть так называемый квантовый переход. Квантовые переходы атома или молекулы обусловлены различными причинами. В частности, они могут происходить при взаимодействии атомов и молекул с оптическим излучением.

Атомы каждого химического элемента характеризуются своей системой энергетических уровней. Характерные расстояния между уровнями энергии атома составляют примерно $1 \dots 5 \text{ эВ}$. Система энергетических уровней атома отражает особенности движения электронов в данном атоме.

Структура энергетических уровней молекулы отражает три типа движений: электронные движения, колебания атомов в молекуле, вращение молекулы. Поэтому система уровней у молекулы сложнее, чем у атома. Наряду с уровнями, отвечающими электронным движениям, существуют колебательные (расстояние между уровнями порядка $0,1 \text{ эВ}$) и вращательные уровни (расстояние между уровнями порядка $0,01 \text{ эВ}$ и меньше).

4.1.2 Поглощение света

Для простоты будем рассматривать всего два энергетических уровня некоего атома. Обозначим энергию нижнего уровня через E_1 , а энергию верхнего - через E_2 . Пусть атом находится на нижнем уровне и пусть вблизи атома пролетает фотон с энергией $\varepsilon_{12} = E_2 - E_1$. Атом может поглотить фотон и совершить

квантовый переход с уровня E_1 на уровень E_2 (рисунок 4.1 а). Это есть элементарный акт поглощения света веществом.

Обозначим через ω_n отнесенную к единице времени вероятность поглощения света, связанную с переходами $E_1 \rightarrow E_2$ атомов вещества. Указанная вероятность пропорциональна числу фотонов. Она может быть записана в виде

$$\omega_n = BN\varepsilon_{12}, \quad (4.1)$$

где N - число фотонов в единице -объема, имеющих энергию ε_{12} ;

B - коэффициент, характеризующий рассматриваемый переход $E_1 \rightarrow E_2$.

4.1.3 Вынужденное испускание света

Если атом находится на верхнем энергетическом уровне, то пролетающий фотон с энергией $\varepsilon_{12} = E_2 - E_1$ может сыграть роль своеобразного "спускового крючка": он может вызвать (инициировать) переход $E_2 \rightarrow E_1$, т.е. возвращение атома на нижний уровень. Переход

$E_2 \rightarrow E_1$ сопровождается испусканием фотона (рисунок 4.1, б). Новый фотон, как и исходный фотон, имеет энергию $\varepsilon_{12} = E_2 - E_1$. Более того, он имеет такое же направление импульса и такую же поляризацию. Иными словами, вторичный фотон (фотон, испущенный атомом в процессе перехода $E_2 \rightarrow E_1$) оказывается в том же самом состоянии, в каком находится первичный фотон (фотон, вызвавший рассматриваемый переход). В этом как раз и проявляется бозонный характер статистики фотонов - тенденция фотонов накапливаться в одном и том же состоянии.

Рассмотренный процесс называют вынужденным (индуцированным) испусканием света. Чем больше имеется первичных фотонов, тем выше вероятность того, что атом, находящийся на уровне E_2 , совершит переход на уровень E_1 . Здесь проявляется определенное сходство между вынужденным испусканием света и поглощением света:

вероятности обоих процессов пропорциональны числу первичных фотонов. Ве-

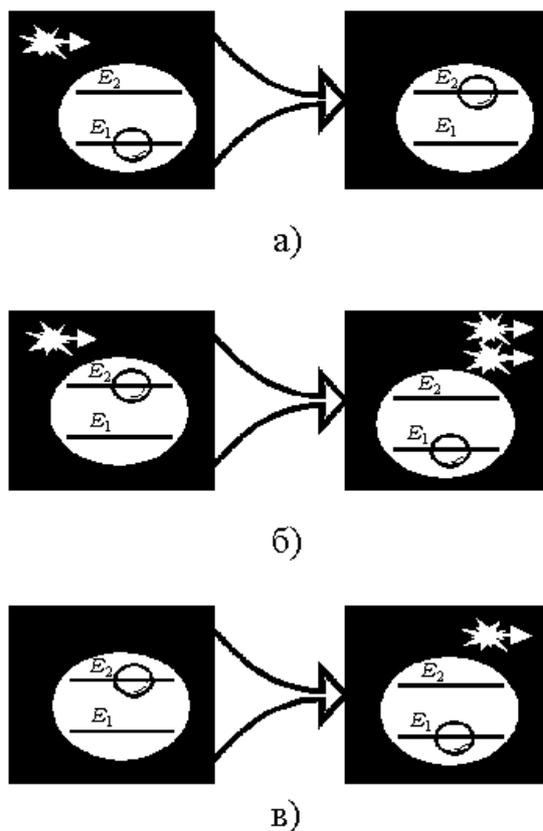


Рисунок 4.1

роятность $\omega_{u,v}$ вынужденного испускания света, отнесенная к единице времени, есть

$$\omega_{u,v} = BN\varepsilon_{12}. \quad (4.2)$$

Выражение (4.2) совпадает с (4.1).

Итак, если атом находится на уровне E_1 , то фотон с энергией $\varepsilon_{12} = E_2 - E_1$ вызывает переход $E_1 \rightarrow E_2$ (и при этом уничтожается); если же атом находится на уровне E_2 , то рассматриваемый фотон с такой же вероятностью вызывает переход $E_2 \rightarrow E_1$ (и при этом рождается еще один фотон).

Предположим, что имеется много атомов на уровне E_2 . Пролетая мимо них, первичный фотон может инициировать переход $E_2 \rightarrow E_1$ во многих атомах, т.е. может инициировать рождение не одного, а целой лавины вторичных фотонов. Все эти фотоны будут рождаться в том же самом состоянии, в каком находится первичный фотон. Таким образом, один фотон может инициировать рождение целого волнового цуга, причем длина этого цуга (определяемая количеством родившихся вторичных фотонов) может быть, в принципе, сколь угодно большой.

4.1.4 Спонтанное испускание света

Находящийся на уровне E_2 атом может перейти на уровень E_1 не только под действием первичного фотона, но и самопроизвольно или, как обычно говорят, спонтанно (рисунок 4.1, в). Родившийся при спонтанном переходе $E_2 \rightarrow E_1$ фотон имеет энергию $\varepsilon_{12} = E_2 - E_1$; остальные же его характеристики (направление импульса, поляризация) оказываются произвольными. Вероятность спонтанного испускания фотона определяется только свойствами перехода $E_2 \rightarrow E_1$; ее записывают в виде

$$\omega_{u,c} = A. \quad (4.3)$$

Коэффициенты A и B , входящие в вероятности процессов испускания и поглощения света, называют коэффициентами Эйнштейна. Важно подчеркнуть, что эти коэффициенты не зависят от каких-либо внешних условий (температуры, давления, интенсивности оптического излучения и т. п.).

Итак, возможны два типа процессов испускания света атомами вещества - вынужденное испускание и спонтанное испускание. Первый процесс можно рассматривать как управляемый: он вызывается первичным фотоном, который не только инициирует переход $E_2 \rightarrow E_1$, но и определяет характеристики родившегося при этом переходе фотона. Второй процесс происходит самопроизвольно; он имеет ярко выраженный случайный характер.

4.1.5 Активная среда и активные центры

"Сердцем" лазера является активный элемент - некоторая среда (твердая, жидкая или газообразная), содержащая специально подобранные атомы, ионы или молекулы, в результате высвечивания которых и возникает лазерное излучение. Эти атомы (ионы, молекулы) называют активными центрами. Они составляют относительно малую долю от всех атомов (молекул), образующих активную среду. Их количество в 1 см^3 активной среды составляет примерно 10^{15} К 10^{17} в газообразных средах и 10^{19} К 10^{20} в твердых и жидких средах.

Будем рассматривать только те уровни энергии активного центра, которые важны с точки зрения возникновения генерации в лазере. Прежде всего выделим два уровня, на переходе между которыми рождается лазерное излучение. Этот переход называют рабочим переходом, а соответствующие энергетические уровни рабочими уровнями. Различают нижний и верхний рабочие уровни. Пусть $\varepsilon_{12} = E_2 - E_1$ - разность энергий рабочих уровней; $\nu_{12} = \varepsilon_{12} / h$ есть частота рабочего перехода (частота излучения лазера). Обозначим через n_1 и n_2 число активных центров в единице объема активной среды, находящихся на нижнем и верхнем рабочих уровнях соответственно; эти величины называют заселенностями рабочих уровней.

4.1.6 Инвертированная активная среда

Как уже отмечалось, фотон с энергией $\varepsilon_{12} = E_2 - E_1$ может с равной вероятностью инициировать как переход $E_1 \rightarrow E_2$, так переход $E_2 \rightarrow E_1$. Все дело в том, на каком энергетическом уровне находится взаимодействующий с фотоном активный центр. Если на нижнем рабочем уровне (E_1) находится больше активных центров, чем на верхнем уровне (E_2), то будут преобладать процессы поглощения света. Если же, наоборот, на уровне E_1 находится меньше активных центров, чем на уровне E_2 , то будут преобладать процессы вынужденного испускания света.

В обычных условиях и, в частности, при термодинамическом равновесии, заселенности энергетических уровней уменьшаются по мере увеличения энергии уровней, т.е. $n_2 < n_1$. Поэтому обычно процессы поглощения света преобладают над процессами вынужденного испускания света.

Нам важно, чтобы, наоборот, преобладали процессы вынужденного испускания света. Следовательно, необходимо позаботиться о том, чтобы заселенность верхнего рабочего уровня оказалась выше заселенности нижнего уровня, т.е. чтобы выполнялось условие

$$n_2 > n_1, \quad (4.4)$$

называемое условием инверсии (обращения) рабочих уровней. Активную среду, для активных центров которой выполнено условие инверсии называют инвертированной активной средой.

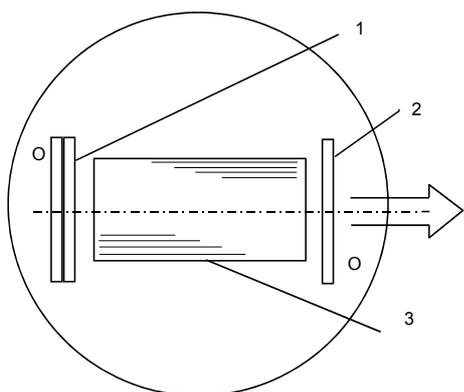
Предположим, что приготовлена инвертированная активная среда. Пусть в эту среду проходит направленный световой пучок, имеющий частоту ν_{12} . Проходя через инвертированную активную среду, световой пучок может усиливаться вследствие преобладания процессов вынужденного испускания над процессами поглощения. Существенно, что усиливающий пучок фотоны рождаются в том же состоянии, в каком находятся фотоны исходного светового пучка, играющие роль первичных фотонов. Обозначим через W световую энергию, генерируемую на частоте ν_{12} в единице объема активной среды в единицу времени (это есть плотность световой мощности). Нетрудно показать, что

$$W = (n_2 - n_1)BNh^2\nu_{12}^2. \quad (4.5)$$

Действительно, за единицу времени в единице объема происходят $n_2\omega_{u.v}$ вынужденных переходов $E_2 \rightarrow E_1$ и $n_1\omega_n$ обратных переходов. При каждом переходе рождается (либо, напротив, уничтожается) фотон, имеющий энергию $h\nu_{12}$. Таким образом, $W = (n_2\omega_{u.v} - n_1\omega_n)h\nu_{12}$

4.1.7 Оптический резонатор

Выделение одних фотонных состояний и подавление остальных состояний осуществляется на практике при помощи оптического резонатора - принципиально важного элемента лазера. В простейшем случае оптический резонатор представляет собой пару зеркал на общей оптической оси, которая фиксирует в пространстве направление лазерного луча. Между зеркалами оптического резонатора помещается активный элемент. Твердотельные активные элементы имеют чаще всего форму цилиндрического тела, ось которого совпадает с оптической осью резонатора; длина цилиндра примерно на порядок больше его диаметра. По крайней мере, одно из зеркал оптического резонатора обладает некоторой прозрачностью. Через это зеркало (его называют выходным зеркалом) из резонатора лазера выходит излучение.



Простейший оптический резонатор (вместе с активным элементом) показан на рисунке 4.2. Спонтанные фотоны, случайно родившиеся в направлении оси OO или достаточно близком к нему, будут проходить внутри активного элемента (3) относительно большой путь, который, к тому же, существенно увеличивается из-за многократных отражений излучения от зеркал (1,2) резонатора. Взаимодействуя с возбужденными активными центрами, эти фотоны иници-

Рисунок 4.2

руют в конечном счете мощную лавину вынужденно испущенных фотонов, которые случайно родились в других направлениях, то они (и соответствующие лавины вторичных фотонов) пройдут в активном элементе относительно короткий путь и быстро "выйдут из игры".

Итак, оптический резонатор обеспечивает избирательность для фотонных состояний прежде всего с точки зрения направления движения фотонов. Он выделяет в пространстве определенное направление, вдоль которого и реализуется лазерная генерация.

4.2 Функциональная схема работы и классификация лазеров

4.2.1 Обобщенная схема лазерного генератора

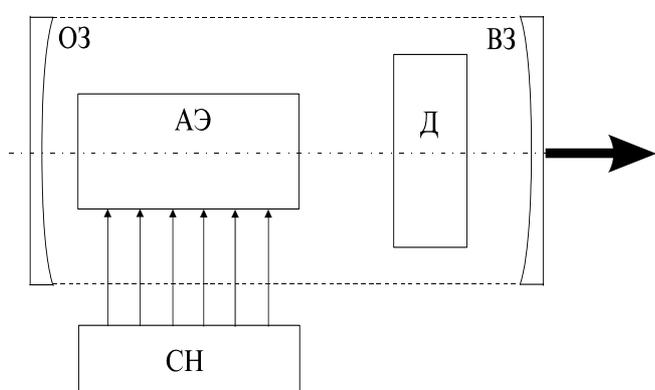


Рисунок 4.3

Любой лазер, по существу, состоит из четырёх компонент (рисунок 4.3). АЭ - активный элемент или активная среда, в которой происходит генерация и усиление световых волн. СН - система накачки, селективно накачивающая энергию в активную среду, чтобы заселить выбранные уровни и достичь инверсной населённости. Оптический резонатор, состоящий из двух параллельных друг другу и перпендикулярных оси лазера зеркал. ОЗ – полностью отражающее

зеркало резонатора, ВЗ – выходное зеркало резонатора. Оптический резонатор накапливает индуцированное излучение, сконцентрированное в нескольких модах. Оптический резонатор также формирует лазерный пучок. Д - дополнительные элементы, вводимые внутрь резонатора и предназначенные для управления характеристиками лазерного излучения.

4.2.2 Классификация лазеров

Существует несколько типов классификации, лазеров, основными из которых являются три:

а) По спектральному диапазону генерируемого излучения: УФ, ИК-лазеры и лазеры видимого диапазона.

б) По типу применяемой накачки. Накачка – это процесс создания инверсной населённости. Основными видами накачки являются оптическая и электрическая – чаще всего с использованием самостоятельного электрического разряда в активной среде. Рассматривая этот способ классификации, следует иметь в виду, что один и тот же вид накачки может допускать две разновидности накачки – импульсную непрерывную. Так для оптической накачки могут

быть использованы как газоразрядные импульсные лампы, так и лампы непрерывного горения. В любом случае для поддержания процесса лазерной генерации, мы подводим к активной среде энергию извне. Причём количество подводимой энергии гораздо больше, чем количество энергии излучаемой. КПД лазеров составляет единицы процентов.

в) Наиболее распространённый вид классификации - по типу активной среды, используемой в лазере. Существуют газовые (атомарные, ионные, молекулярные), твердотельные, жидкостные, полупроводниковые.

4.3 Отдельные типы лазеров

4.3.1 Гелий-неоновый лазер

Газовые лазеры - наиболее распространенный тип лазеров. А среди газовых лазеров наиболее распространен He-Ne лазер. Этому способствуют несколько причин:

- He-Ne лазеры конструктивно закончены, давно выпускаются промышленностью, они дешевле, других типов лазеров, доступнее и проще.

- He-Ne лазеры обладают рекордной долговечностью при сохранении выходных параметров.

Лазер на смеси гелия и неона - пример газоразрядного лазера на атомных переходах. Генерация происходит на переходах между уровнями атомов неона (они являются активными центрами). Кроме неона, в состав активной среды входит буферный газ - гелий. Как будет показано, он играет важную роль в возбуждении атомов неона. Давление газовой смеси порядка 1 мм рт. ст. ; парциальное давление гелия в 5-10 раз больше давления неона. В гелий - неоновом лазере используется стационарный тлеющий разряд, возбуждаемый постоянным током.

На рисунке 4.4 показана газоразрядная трубка лазера, ориентированная вдоль оси OO резонатора (2). На рисунке хорошо видны катод K и анод A . Когда разность потенциалов между ними (рабочее напряжение) достигает примерно 10^3 В , в рабочем капилляре газоразрядной трубки (1), имеющей диаметр в несколько миллиметров, зажигается тлеющий разряд. (3) - газопоглотитель с подогревом для поглощения примесей посторонних газов. Плоскости выходных окон газоразрядной трубки ориентированы не перпендикулярно к оси резонатора, а наклонно - так, чтобы перпенди-

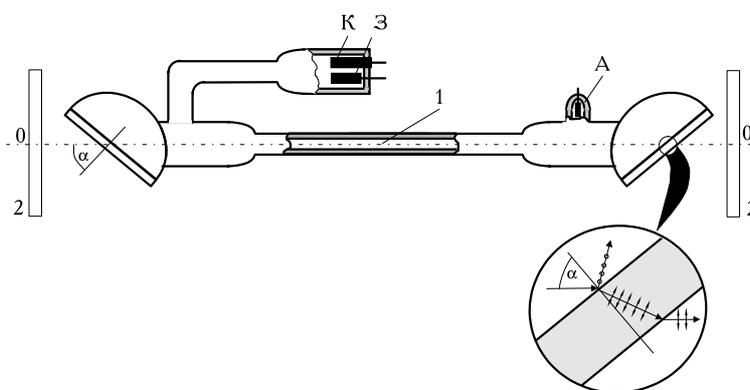


Рисунок 4.4

куляр к плоскости окна составлял с осью резонатора так называемый угол Брюстера, соответствующий показателю преломления вещества, из которого изготовлено выходное окно трубки. Данная ситуация выделена на рисунке 4.4; короткие стрелки показывают, что колебания электрического вектора происходят в плоскости рисунка (в плоскости падения), а кружочки указывают на колебания в плоскости, перпендикулярной к плоскости рисунка. Заметим, что при угле падения, равном α - углу Брюстера, отраженный и преломленный лучи составляют друг с другом прямой угол. Используя окна трубки, срезанные под углом Брюстера, достигаем сразу две цели. Во-первых, получаем поляризованное лазерное излучение. Во-вторых, исключаем потери на отражение от поверхностей выходных окон газоразрядной трубки.

На рисунке 4.5 показаны основные переходы, происходящие в активной среде гелий - неонового лазера. В левой половине рисунка представлены переходы в атомах гелия, а в правой - в атомах неона. Роль верхних рабочих уровней играют термы $3s$ и $2s$, а роль нижних рабочих уровней - термы $3p$ и $2p$. Основные рабочие переходы: $3s \rightarrow 3p$ (3,39 мкм), $2s \rightarrow 2p$ (1,15 мкм), $3s \rightarrow 2p$ (0,63 мкм; красная линия).

Инверсия в гелий - неоновом лазере возникает за счет того, что скорость заселения верхних рабочих уровней значительно выше скорости заселения нижних. Образующиеся в газоразрядной плазме свободные электроны сталкиваются с атомами гелия и неона и возбуждают их, передавая им часть своей кинетической энергии. Неоптические переходы, отвечающие процессам электронного возбуждения, показаны на рисунке 4.5 штриховыми стрелками. Электроны возбуждают уровни $1S_0$ и $3S_1$ в гелии, а также различные энергетические полосы неона (соответствующие как верхним, так и нижним рабочим уровням).

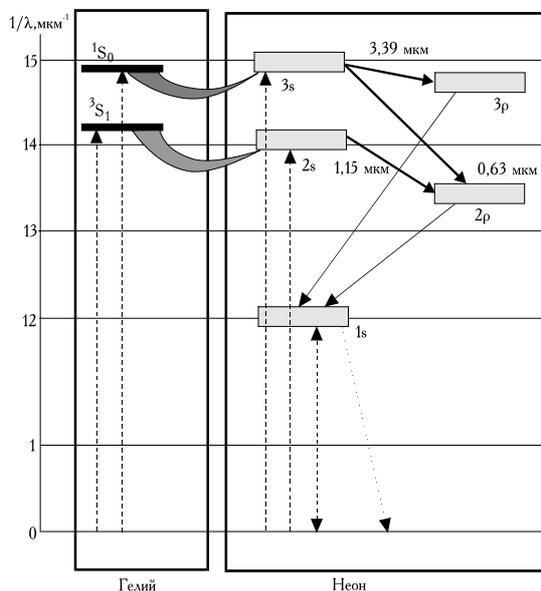


Рисунок 4.5

Возбужденные атомы гелия, сталкиваясь с возбужденными атомами неона, резонансно передают им свою энергию возбуждения, что условно показано на рисунке полукруглыми стрелками. Резонансная передача энергии от гелия к неону и является тем решающим фактором, благодаря которому скорость заселения верхних рабочих уровней неона оказывается выше скорости заселения нижних рабочих уровней. Процессу резонансной передачи энергии от гелия к неону благоприятствуют три обстоятельства: близость соответствующих уровней энергии атомов гелия и неона; метастабильность возбужденных уровней атомов гелия (они относительно медленнее очищаются процессами спон-

тупного излучения); метастабильность возбужденных уровней атомов неона (они относительно медленнее очищаются процессами спон-

танного испускания); более высокое давление гелия в газовой смеси, вследствие чего уменьшается вероятность обратной передачи энергии (от неона к гелию).

4.3.2 Твердотельные лазеры

Твердотельными называют лазеры, активная среда которых представляет собой диэлектрический кристалл или стекло, в которые введены (в виде примеси) специальные ионы, играющие роль активных центров. В твердотельных лазерах применяется только оптическая накачка.

Оптическая накачка. При оптической накачке происходит возбуждение активных центров за счет излучения специального источника света (источника излучения накачки). В качестве таких источников используются газоразрядные импульсные лампы, лампы непрерывного горения (газоразрядные и накаливания), искровые разрядники, пламя и так далее. Источником накачивающего излучения может служить также вспомогательный лазер.

Чтобы оптическая накачка была эффективной, необходимо выполнение двух условий. Во-первых, уровень возбуждения активного центра должен иметь достаточно большую ширину. Иначе говоря, линии поглощения накачивающего излучения должен соответствовать относительно широкий участок, а шкале частот (шкале длин волн). Во-вторых, необходимо, чтобы максимум линии поглощения активного центра попадал в область спектра частот накачивающего излучения, характеризующуюся высокой интенсивностью. Только при выполнении этих условий оптическая накачка может стать достаточно эффективной; в противном случае доля полезно расходуемой мощности накачивающего излучения будет незначительной.

Твердотельный лазер. В твердотельном активном элементе различают матрицу (основу) и введенный в матрицу в виде примеси активатор (активные центры). Используются как кристаллические, так и аморфные (стеклянные) матрицы. Наиболее широко используются так называемые оксидные кристаллы с упорядоченной структурой. К ним относятся, в частности, кристалл окиси алюминия (Al_2O_3), активированный ионами хрома (Cr^{3+}), а также кристалл иттрий-алюминиевого граната ($Y_3Al_5O_{12}$), активированный ионами неодима (Nd^{3+}). Первый кристалл есть активный элемент лазера на рубине, а второй - на гранате с неодимом. Указанные лазеры являются, пожалуй, наиболее широко применяемыми твердотельными лазерами.

Лазерная генерация была осуществлена на рубине. Благодаря большому оптическому качеству рубина до сих пор широко используются на практике лазеры на рубине, содержащие в качестве примеси ионы хрома. В лазерах обычно используют розовый рубин с содержанием хрома 0,05% ($1,6 \cdot 10^{19}$ ионов хрома в 1 см^3). На рисунке 4.6 представлена схема уровней активного центра (иона хрома) в рубине. По вертикальной оси отложены значе-

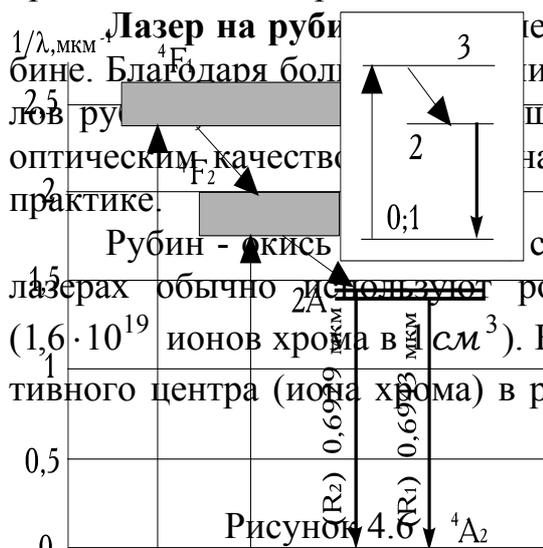


Рисунок 4.6

ния $1/\lambda$ в мкм^{-1} . Переход от $1/\lambda$ к энергии E осуществляется по формуле $E = hc/\lambda$. В процессе оптической накачки возбуждаются состояния (уровни), обозначаемые как 4F_1 и 4F_2 . Каждое из этих состояний изображено на рисунке в виде некоторой энергетической полосы, ширина которой отвечает "размытию" соответствующего уровня. Передавая часть энергии тепловым колебаниям кристаллической решетки рубина, возбужденные ионы хрома относительно быстро переходят из состояний 4F_1 и 4F_2 на два близко расположенных метастабильных уровня, условно обозначаемых $\bar{2}\bar{A}$ и \bar{E} . Эти уровни играют роль верхних рабочих уровней. Нижний рабочий уровень обозначен 4A_2 , он является также основным уровнем. Таким образом, лазер на рубине может быть описан трехуровневой схемой.

На рисунке 4.7, а представлен спектр поглощения иона хрома в рубине. Спектр имеет два максимума - для состояний 4F_1 и 4F_2 . Эти максимумы отвечают длинам волн $\lambda' = 0,41 \text{ мкм}$ (синяя полоса поглощения) и $\lambda'' = 0,55 \text{ мкм}$ (желто - зеленая полоса поглощения). Ширина каждой полосы поглощения составляет примерно $0,1 \text{ мкм}$. Для накачки рубинового лазера применяют газоразрядную ксеноновую лампу. На рисунке 4.7, б показан спектр излучения, спонтанно испускаемого ионами хрома на рабочем переходе (спектр люминесценции хрома в рубине). В спектре имеются две линии. Максимум одной из них соответствует длине волны $0,6943 \text{ мкм}$ (так называемая R_1 - линия); максимум другой - длине волны $0,6929 \text{ мкм}$ (R_2 - линия). При лазерной генерации реализуется практически только R_1 - линия.

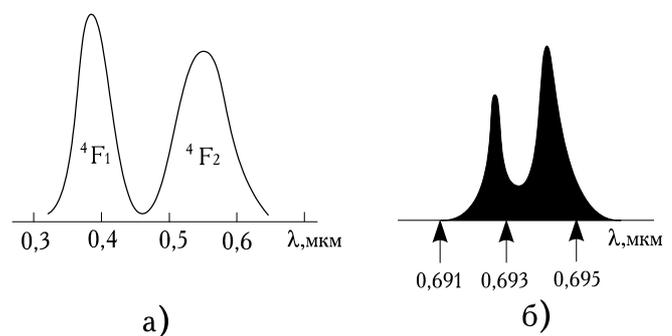


Рисунок 4.7

Лазер на гранате с неодимом. Лазер на иттрий - алюминиевом гранате, активированном ионами неодима, является в настоящее время наиболее широко применяемым твердотельным лазером. Он имеет сравнительно низкий порог возбуждения и высокую теплопроводность, что позволяет реализовать генерацию при большой частоте следования световых импульсов, а также генерацию в непрерывном режиме. КПД лазера сравнительно высок; он достигает нескольких процентов.

При упрощенном рассмотрении лазера на гранате с неодимом можно пользоваться четырехуровневой схемой.

Для накачки рассматриваемого лазера применяют криптоновые лампы. Иногда в кристаллическую решетку граната (содержащую в виде примеси ионы неодима) дополнительно вводят еще одну примесь - ионы хрома, что позволяет применять для накачки ксеноновые лампы. Ионы хрома в гранате имеют две довольно широкие полосы поглощения (при $\lambda = 0,43$ и $0,59 \text{ мкм}$) хорошо со-

гласующиеся со спектром излучения ксеноновых ламп. Возбужденные ионы хрома передают энергию возбуждения активным центрам - ионам неодима.

Достаточно широко применяются лазеры на стеклах, активированных редкоземельными элементами, в частности неодимом. Лазер на стекле с неодимом, как и лазер на гранате с неодимом, работает по четырехуровневой схеме. Генерируемое излучение имеет $\lambda = 1,06 \text{ мкм}$.

4.3.3 Полупроводниковые инжекционные лазеры

Излучение квантов света в полупроводниках происходит благодаря переходам электронов из зоны проводимости (ЗП) в валентную зону (ВЗ). При этом высвобождается энергия $E = h\nu$, приблизительно равная ширине запрещенной зоны. Этот процесс эквивалентен рекомбинации электронов и дырок. Высвечивание может быть спонтанным и вынужденным. Создать лазер на основе беспримесного полупроводника можно накачивая его оптически, либо электронным пучком. Однако создать инверсную населённость, т.е. преобладание электронов на нижнем крае ЗП над их концентрацией в верхнем крае ВЗ трудно.

Гораздо легче создать инверсную населённость в примесных полупроводниках, которые, благодаря разнообразию вводимых примесей, способны демонстрировать широкий спектр выходных характеристик. Уже при комнатных температурах большое число электронов с уровня донорной примеси (УДП) переходит в ЗП, а одновременно электроны переходят из ВЗ на уровень акцепторной примеси (УАП), освобождая в ВЗ множество дырок (рисунок 4.8).

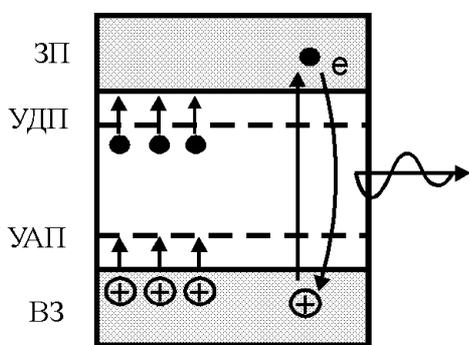


Рисунок 4.8

Для создания инверсии заселенностей в полупроводниковой активной среде можно привести в контакт друг с другом два вырожденных полупроводника разного типа (n-полупроводник и p-полупроводник). Такой контакт называют *p-n переходом*. Приложим к p-n переходу электрическое напряжение так, как показано на рисунке 4.9. Под действием поля электроны проводимости из n-полупроводника будут инжектироваться в область p-n перехода; одновременно в эту же область будут инжектироваться дырки из p-полупроводника. В области p-n перехода будут

происходить переходы электронов из зоны проводимости в валентную зону (рекомбинация электронов проводимости и дырок), на которых при наличии оптического резонатора может быть реализована лазерная генерация. Таков принцип действия полупроводниковых лазеров, называемых *инжекционными лазерами*. Наиболее широко применяется инжекционный лазер на арсениде галлия GaAs. Вырожденный n-полупроводник получается при внесении в арсенид галлия примеси теллура; концентрация

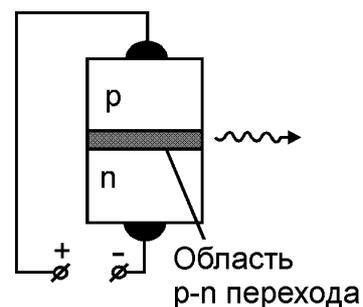
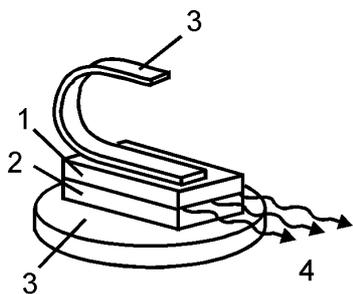


Рисунок 4.9

примеси $3 \dots 5 \cdot 10^{18} \text{ см}^{-3}$. Вырожденный p-полупроводник получается при внесении в GaAs примеси цинка; концентрация примеси 10^{19} см^{-3} . Генерация осуществляется на длинах волн от 0,82 мкм до 0,9 мкм (инфракрасное излучение).



Рисунок

Инжекционный лазер на GaAs показан на рисунке 4.10. Видно, что зеркалами оптического резонатора служат отполированные грани кристалла, ориентированные перпендикулярно к плоскости, по которой контактируют друг с другом p-полупроводник (2) и n-полупроводник (1). (3) – металлические контакты, (4) – излучение лазера. Инжекционный лазер отличается своей миниатюрностью; линейные размеры граней полупроводникового кристалла составляют всего 1 мм. Толщина

излучающей области (области перехода) равна всего лишь 2 мкм. Мощность генерируемого излучения такого лазера порядка 10 мВт достигается в непрерывном режиме; в импульсном режиме она достигает 100 Вт. Наибольшие мощности реализуются при охлаждении до 4,2 К. Разработаны образцы полупроводниковых лазеров, которые могут генерировать в сине-зеленой области спектра. Инжекционные лазеры характеризуются высоким КПД, который может достигать 50...60 %.

5 Основы голографии

Голография это интерференционный метод регистрации световых волн, дифрагировавших на объекте, который освещен когерентным светом [7]. Иначе говоря, голографический способ записи и восстановления волнового поля основан на регистрации интерференционной картины, которая образована волной, отраженной предметом (*предметная волна*) и когерентной с ней волной, идущей непосредственно от источника света (*опорная волна*). Зарегистрированная интерференционная картина называется *голограммой*.

5.1 История развития голографии

Основы голографии были заложены в 1948 году английским физиком Деннисом Габором. Само слово голография происходит от греческого *holos*, что означает "весь", "целый". Этим изобретатель хотел подчеркнуть, что в голографии регистрируется полная информация о волне, как амплитудная, так и фазовая.

Габор поставил перед собой задачу усовершенствования принципов электронной микроскопии. Он предложил регистрировать, кроме амплитудной, также и фазовую информацию об электронных волнах путем наложения когерентной опорной волны. Из-за отсутствия когерентных электронных пучков он провел оптические модельные опыты, положившие начало голографии. Однако трудности, связанные с отсутствием мощных источников когерентного света,

были столь существенными, что голография оставалась лишь "оптическим парадоксом".

Второе рождение голография пережила в 1962-1963 годах, когда американские ученые Э.Лейт и Ю.Упатниекс применили для голографии изобретенные к тому времени лазеры. Ими же была предложена схема получения голограмм с наклонным опорным пучком, а советский физик Ю.Н.Денисюк получил первые голограммы с записью в трехмерной среде, объединив, таким образом, идею голографии Габора с методом цветной фотографии Липпмана.

Последующее развитие голографии шло весьма бурными темпами. К 1965-1966 годам были заложены все ее теоретические и экспериментальные основы, а в последующие годы развитие голографии шло по пути совершенствования применений.

5.2 Основные схемы записи голограмм

Голография является многоступенчатым процессом: сначала регистрируется голограмма объекта, затем восстанавливается его изображение, а в научно-технических приложениях восстановленное изображение обрабатывается.

Рассмотрим некоторые основные схемы, использующиеся для голографической записи. Такое рассмотрение, во-первых, отражает исторический ход развития голографии, во-вторых, помогает понять физику процессов, происходящих при записи голограмм, а, в третьих, будет использовано при рассмотрении классификации голограмм.

По геометрии падения волн на регистрирующую среду голографические схемы можно разделить на однолучевые (схема Габора), двухлучевые (схема Лейта и Упатниекса) и схемы во встречных пучках (схема Денисюка) /8/.

В однолучевой схеме опорной волны как таковой нет (рисунок 5.1). Она формируется из волны, не претерпевшей рассеяния при прохождении через объект. Поэтому класс объектов, которые можно регистрировать с помощью однолучевой схемы, ограничен.

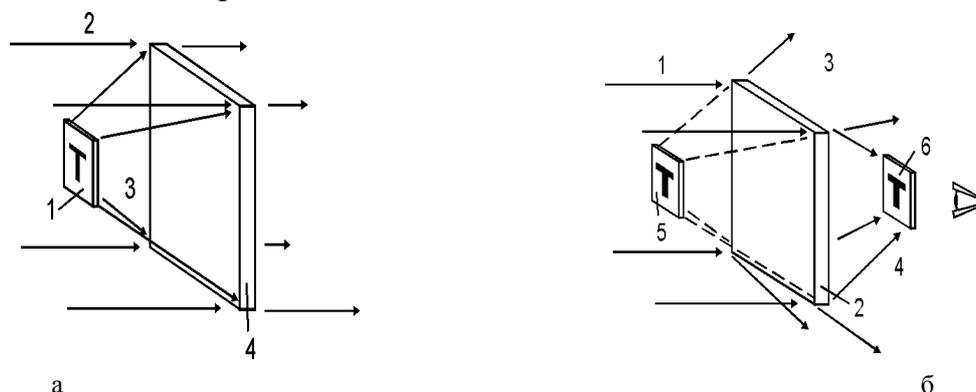


Рисунок 5.1

На рисунке 5.1 а представлена схема записи однолучевой голограммы. Предмет-транспарант (1), освещающий пучок (2), предметная волна (3) и фотопластинка (4) находятся на одной оси. При восстановлении (рисунок 5.1 б) та-

кой голограммы (2) исходной волной (1) образуются пучки (3) и (4), которые формируют действительное (6) и мнимое (5) изображения предмета. Эти изображения также расположены на одной оси, что создает трудности для наблюдения.

В двухлучевой схеме и в схеме во встречных пучках сигнальные и опорные лучи разделены в пространстве и падают на регистратор под разными углами. Эти схемы отличаются друг от друга тем, что в двухлучевой схеме интерферирующие волны падают на регистрирующую среду с одной стороны, а в схеме во встречных пучках – с противоположных сторон.

На рисунках 5.2 а и 5.2 б изображена геометрия записи и восстановления голограмм по схеме Лейта и Упатниекса. Опорный пучок (1) и отраженный от объекта (2) предметный пучок (3) интерферируют на регистрирующем материале (4). При восстановлении опорным пучком голограммы (6), записанной по такой двухлучевой схеме, образуются два изображения: мнимое (2) и действительное (3). Мнимое изображение, расположенное за голограммой, является точной копией объекта и позволяет производить исследования с помощью традиционных оптических методов, например, фотографирования. Действительное изображение, расположенное перед голограммой, является псевдоскопической копией объекта.

В двухлучевой схеме, в отличие от однолучевой, волны (3) и (4), образующие мнимое и действительное изображения, и волна нулевого порядка дифракции, возникающие при восстановлении голограммы, разделены в пространстве и не вносят взаимных помех.

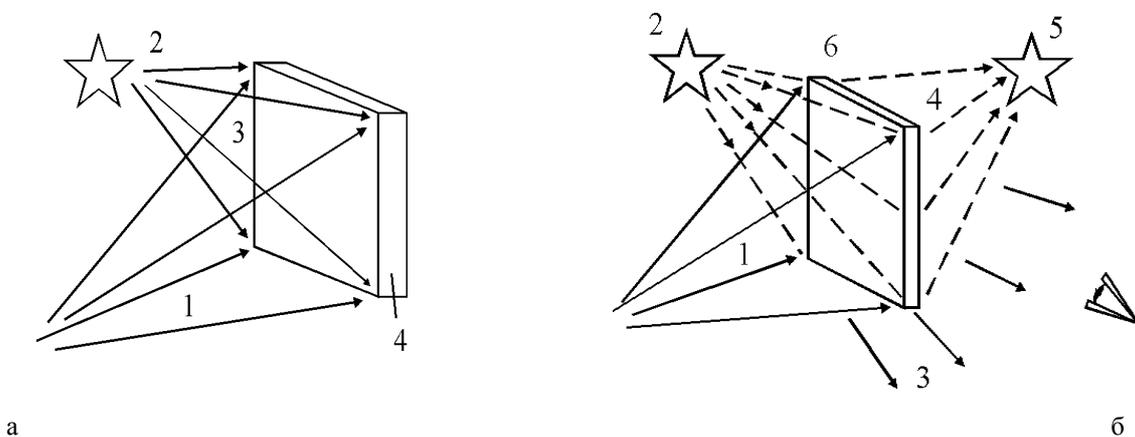


Рисунок 5.2

Голограммы, полученные по схеме во встречных пучках, обладают рядом особенностей. Избирательность голограммы по отношению к длине волны восстанавливающего излучения позволяет восстанавливать изображение от источника, имеющего сплошной спектр излучения, например Солнца. Кроме того, голограмма, полученная по схеме во встречных пучках, при восстановлении мнимого изображения не создает дополнительно волны нулевого порядка и действительного изображения объекта.

На рисунках 5.3 а и 5.3 б представлены схемы записи и восстановления голограмм Денисюка (во встречных пучках). Для записи такой голограммы (рисунок 5.3 а) излучение лазера (1) проходит через расширитель пучка (2), формируя опорную волну (6). Отразившись от различных точек предмета (4), волны (5) образуют предметный пучок, который интерферирует с опорным на регистрирующем материале (3) под большими углами. Записанная голограмма обладает свойствами объемной дифракционной решетки. При восстановлении (рисунок 5.3 б) световая волна (2) от источника сплошного спектра (1) дифрагирует на голограмме (3). Наблюдатель видит изображение предмета, записанного на голограмме, т.к. восстановленная волна (4) сохраняет всю информацию о предмете.

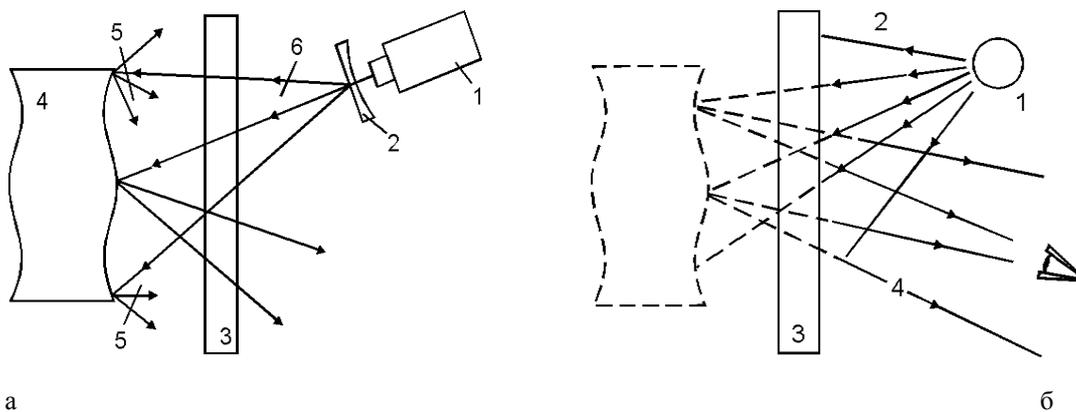


Рисунок 5.3

5.3 Основные уравнения

Поскольку голограмма является интерференционной картиной, образованной предметной и опорной волнами, то для анализа результата восстановления голограммы можно воспользоваться уравнениями, описывающими взаимодействие этих волн /9/.

Обратимся к рисунку 5.4, где изображена принципиальная схема получения голограммы. На фотопластинку падают световые волны от предмета и от опорного источника. Мы будем считать, что комплексные векторы амплитуд этих волн направлены перпендикулярно к плоскости чертежа, и потому ограничимся скалярным рассмотрением. Комплексные амплитуды предметной и опорной волн в плоскости голограммы (x, y) можно записать в следующем виде:

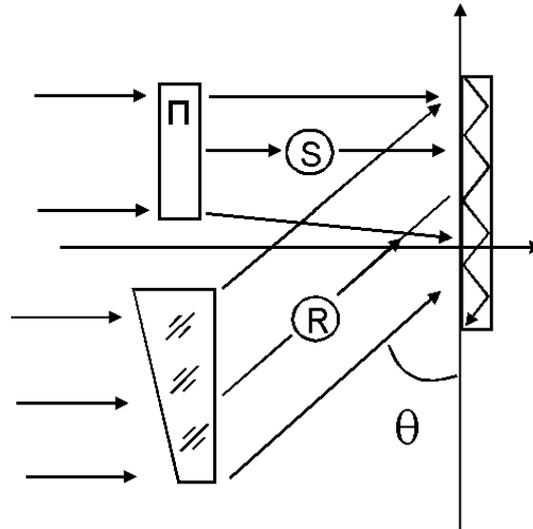


Рисунок 5.4

Соответствующее объекту распределение света

$$S(xy) = S_0(xy) e^{i[\omega t - kz + \varphi_0(xy)]}, \quad (5.1)$$

где: x, y - координаты в плоскости фотопластинки,

z - направление распространения света,

ω - циклическая частота,

S_0 - амплитудная функция,

φ_0 - начальная фаза.

Мгновенное значение напряженности электрического поля световой волны опорного пучка:

$$R(xy) = R_0 e^{i[\omega t - kx \cos \theta - kz \sin \theta]}. \quad (5.2)$$

Введем приближения: 1) поскольку процесс записи предполагает усреднение светового поля по многим циклам, $e^{i\omega t}$ - постоянный коэффициент, которым можно пренебречь; 2) фотопластинка относительно тонкая, то есть изменение координаты внутри нее несущественно $z = 0$; 3) $k \cos \theta = \text{const} = \alpha$. Тогда система уравнений для опорной и предметной волн запишется в виде:

$$S(xy) = S_0(xy) e^{i\varphi_0(xy)}, \quad (5.3)$$

$$R(xy) = R_0 e^{-i\alpha x}. \quad (5.4)$$

Если эти волны когерентны, то они образуют на фотопластинке - голограмме интерференционную картину, распределение освещенности в которой равно:

$$I(xy) = |(S + R)|^2 = S_0^2 + R_0^2 + S R^* + S^* R, \quad (5.5)$$

$$I(xy) = S_0^2 + R_0^2 + 2S_0R_0 \cos(\varphi_0 + \alpha x). \quad (5.6)$$

Итак на фотопластинке должны быть зарегистрированы сумма квадратов амплитудных функций (относительно медленно меняющийся фон) плюс член, осциллирующий по закону косинуса, который характеризует интерференционную структуру с высокой пространственной частотой. Запись содержит информацию как об амплитудной функции предмета $S_0(x, y)$, так и о фазовой функции $\varphi_0(x, y)$.

Мы здесь ограничимся случаем чисто амплитудной записи этого распределения, когда фотослой реагирует на освещенность только своим пропусканием.

Фотографические свойства фотопластинки обычно описывают так называемой характеристической кривой, представляющей собой зависимость от логарифма экспозиции ($H = I \cdot t$, где t - время экспозиции) оптической плотности почернения фотопластинки, то есть величины $D = -\lg T$, где T - коэффициент пропускания проявленного фотослоя по интенсивности, связанный с амплитудным пропусканием слоя как $T = T_a^2$. Проявление голограммы ведут, выбирая линейный участок характеристической кривой $D = A \lg H$, где $A = \text{const}$. Наклон характеристической кривой определяет коэффициент контрастности фотоматериала γ .

$$-2 \lg T_a = A \lg H \quad \text{и} \quad A = -2 = \gamma. \quad (5.7)$$

Таким образом, амплитудное пропускание правильно проявленной голограммы должно быть линейной функцией от распределения интенсивности светового поля на этапе записи

$$T_a(xy) = AI(xy). \quad (5.8)$$

Следующим этапом является восстановление голограммы. Осветив голограмму опорной волной $R(x, y)$, мы имеем непосредственно за ее плоскостью следующее распределение комплексных амплитуд (рисунок 5.5):

$$W = RT_a(xy) = RAI(xy) = RA(S_0^2 + R_0^2 + SR^* + S^*R). \quad (5.9)$$

q

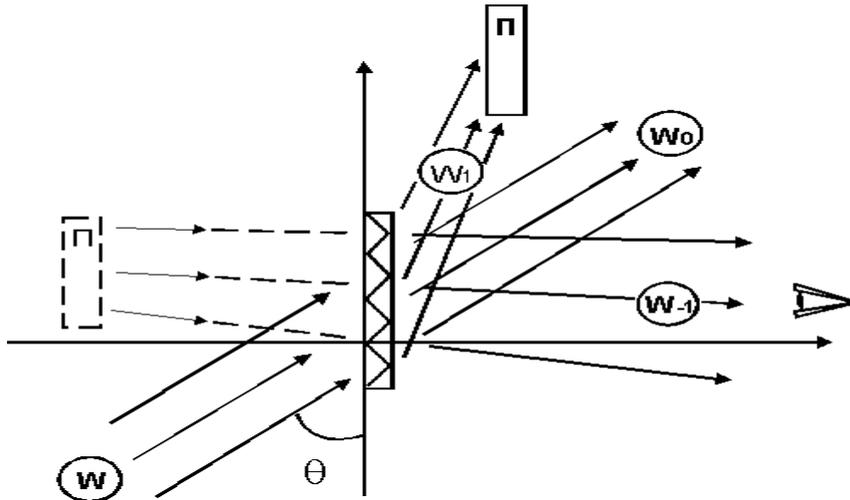


Рисунок 5.5

Первый член в правой части этого равенства с точностью до множителя равен комплексной амплитуде опорной волны и соответствует волне нулевого порядка

$$W_1(xy) = RA(S_0^2 + R_0^2) = A_1 R. \quad (5.10)$$

Второй член только действительным множителем $AR_0^2 = A_2$ отличается от комплексной амплитуды предметной волны. Этот член описывает восстановленную голограммой предметную волну (волну минус 1-го порядка). Эта волна образует мнимое трехмерное изображение предмета, расположенное там же, где находится предмет во время записи голограммы.

$$W_2(xy) = RASR^* = AR_0^2 S(xy) = A_2 S_0(xy) e^{i\varphi_0(xy)}. \quad (5.11)$$

Третий член уравнения отличается от волны, сопряженной предметной волне, комплексным множителем. Он описывает волну плюс 1-го порядка, образующую искаженное действительное изображение объекта. Это изображение называют псевдоскопическим. Для его наблюдения нужно установить экран в направлении, составляющем угол θ_1 с оптической осью, который определяется выражением $2\alpha = k \cos \theta_1$. Чтобы получить неискаженное действительное изображение, голограмму надо осветить волной, сопряженной опорной волне.

$$W_3(xy) = RAS^* R = AR_0^2 e^{-2i\alpha x} S_0 e^{-i\varphi_0} = AR_0^2 S_0^2 e^{-i(\varphi_0 + 2\alpha x)}. \quad (5.12)$$

Из рассмотренной схемы видно, что углы, под которыми распространяются волны нулевого ± 1 -го порядков, определяются углами падения на голограмму предметной и опорной волны. В схеме Габора опорный источник и

предмет располагались на оси голограммы (осевая схема). При этом все три волны на этапе восстановления распространялись за голограммой в одном и том же направлении, создавая взаимные помехи. В схеме Лейта и Упатниекса такие помехи были устранены использованием наклонной опорной волны (внеосевая схема).

5.4 Классификация голограмм

Голограммы можно классифицировать по способу формирования предметной и опорной волн, по способу записи интерференционной структуры и т.д. Рассмотрим некоторые типы классификации.

5.4.1 По способу формирования предметной и опорной волн

Предметную волну формируют следующим образом. Предмет освещается пучком когерентного света. В качестве предмета будем рассматривать двумерный объект переменной прозрачности – транспарант. Рассеянная предметом волна, несущая информацию о предмете, попадает на голограмму. При этом в зависимости от взаимного положения предмета и голограммы, а также наличия оптических элементов между ними, связь между амплитудно-фазовым распределением в плоскости голограммы и соответствующим распределением непосредственно за объектом может описываться преобразованиями Френеля или Фурье:

$$a(x, y) = \iint f(x_1, y_1) \frac{1}{\lambda d} e^{\frac{ip}{\lambda d} [(x_1 - x)^2 + (y_1 - y)^2]} dx_1 dy_1, \quad (5.13)$$

$$a(x, y) = \iint f(x_1, y_1) e^{-\frac{2\pi i}{\lambda f} (x_1 x + y_1 y)} dx_1 dy_1, \quad (5.14)$$

где x, y и x_1, y_1 – соответственно координаты в плоскости голограммы и предмета,

d – расстояние от предмета до голограммы,

f – фокусное расстояние линзы, с помощью которой осуществляется преобразование Фурье.

Если предмет находится в плоскости голограммы или сфокусирован на неё, то амплитудно-фазовое распределение на голограмме будет тем же, что и в плоскости объекта (рисунок 5.6 а). Соответствующие голограммы носят название *голограмм сфокусированного изображения*.

Когда голограмма находится на бесконечном расстоянии от предмета, то есть в области дифракции Фраунгофера, она носит название *голограммы Фраунгофера*. В этом случае каждая точка предмета посылает на голограмму параллельный световой пучок, а связь между амплитудно-фазовым распределением предметной волны в плоскости голограммы и в плоскости предмета даётся пре-

образованием Фурье. Для получения такой голограммы предмет должен располагаться достаточно далеко от пластинки, либо в фокусе линзы (рисунок 5.6 б).

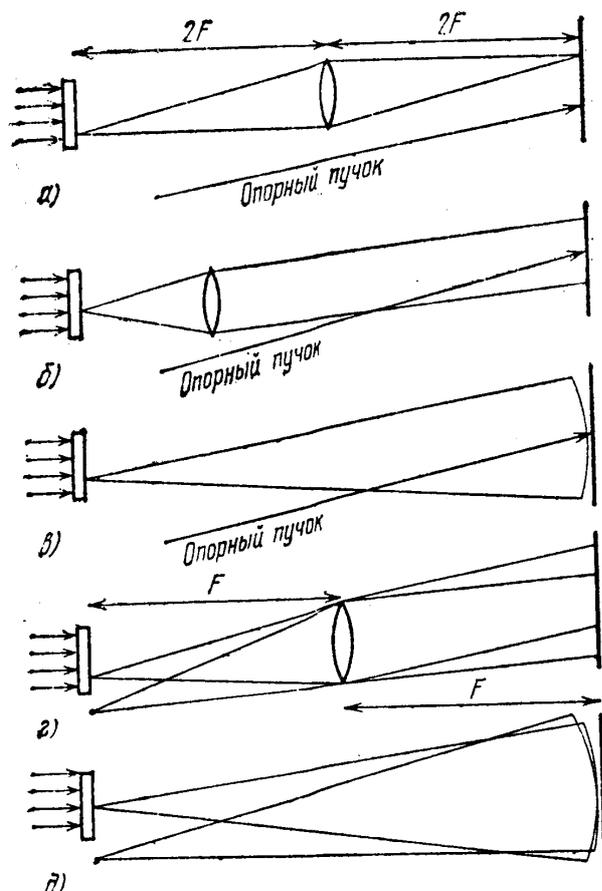


Рисунок 5.6

образом распределения амплитуд соответственно предмета и опорного источника. Такую голограмму называют *фурье-голограммой*. Для получения фурье-голограммы и предмет, и опорный источник обычно располагают в фокальной плоскости линзы (рисунок 5.6 г). Другая схема получения фурье-голограммы изображена на рисунке 5.6 д. В этом случае и предмет и опорный источник расположены на конечном расстоянии от фотопластинки, однако, поскольку они находятся от нее на одинаковом расстоянии, фронт опорной волны и фронты элементарных волн, рассеянных отдельными точками объекта, имеют одинаковую кривизну. Поэтому структура и свойства такой "*безлинзовой фурье-голограммы*" практически такие же, как и у голограммы, полученной по схеме 5.6 г.

Наиболее общим типом голограмм являются голограммы Френеля (рисунок 5.6 в). Они образуются в том случае, когда регистрирующий материал располагается в области дифракции ближнего поля (область дифракции Френеля). По мере увеличения расстояния между объектом и голограммой они переходят в голограммы Фраунгофера, а если это расстояние стремится к нулю, получаем голограммы сфокусированных изображений.

Поскольку на голограмме записывается интерференционная картина, образованная предметной и опорной волнами, для классификации голограмм существенна и форма волнового фронта последней.

Если и предмет, и точечный источник опорной волны расположены на бесконечности, то распределение амплитуд каждой из волн в плоскости голограммы совпадёт с Фурье-

5.4.2 По способу записи интерференционной структуры

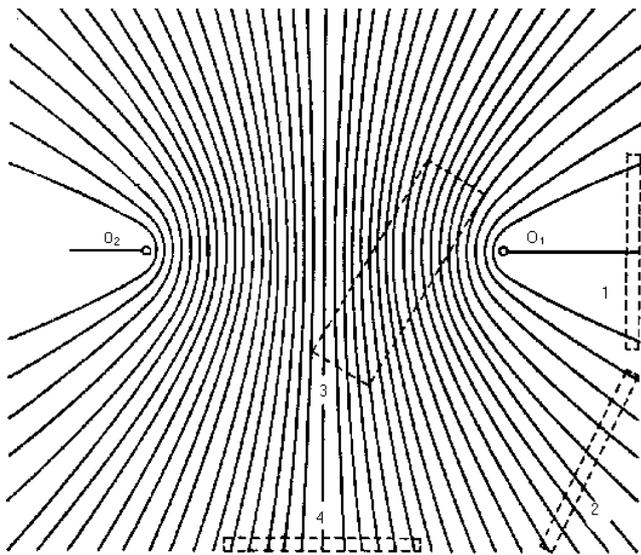


Рисунок 5.7

Структура объёмной интерференционной картины, образующейся при суперпозиции двух сферических волн, представлена на рисунке 5.7. Здесь изображена система гиперboloидов вращения, соответствующая совокупности поверхностей интерференционных минимумов и максимумов от точечного опорного источника (O_1) и точечного предмета (O_2) в сечении плоскостью чертежа.

структуры, определяется углом θ , под которым сходятся в данной точке световые лучи, исходящие из опорного источника и предмета

Пространственная частота ν

$$\nu = \frac{1}{\Lambda} = \frac{2 \sin (u / 2)}{\lambda}, \quad (5.15)$$

где Λ - период интерференционной структуры;
 λ - длина волны света.

В осевой схеме Габора опорный источник и предмет расположены на оси голограммы. В этом случае угол θ близок к нулю и пространственная частота интерференционной структуры минимальна. Осевые голограммы иногда называют также однолучевыми, поскольку в схеме Габора используется один пучок света, рассеянный предметом, часть которого образует предметную волну, а часть исходного пучка, прошедшая через объект без искажения является опорной волной. Расположение фотопластинки при получении осевой голограммы обозначено на рисунке 5.7 цифрой 1.

Схеме Лейта и Упатниекса с наклонным опорным пучком соответствует положение 2 на рисунке 5.7. Обычно в этой схеме когерентный опорный пучок формируется отдельно. Поэтому голограмму, полученную в схеме с наклонным опорным пучком, иногда называют двухлучевой. Пространственная частота интерференционной структуры для голограммы с наклонным опорным пучком выше, чем для однолучевых голограмм. Поэтому для регистрации таких голограмм требуются светочувствительные материалы с высоким пространственным разрешением.

Положению 3 на рисунке 5.7 соответствует голограмма, полученная во встречных пучках. В этом случае опорный и предметный пучки падают на светочувствительный слой с разных сторон, а угол θ между ними близок к 180° . Пространственная частота структуры при этом максимальна и близка к $2/\lambda$. При записи голограмм во встречных пучках интерференционные максимумы располагаются вдоль поверхности материала в его толще. Эта схема была впервые

предложена Ю. Н. Денисюком. Поскольку при освещении такой голограммы опорным пучком восстановленная предметная волна распространяется навстречу освещающему пучку, такие голограммы называют также отражательными.

На рисунке 5.7 изображено также положение фотопластинки (4) при записи "безлинзовой фурье-голограммы".

5.4.3 Двумерные и трехмерные голограммы

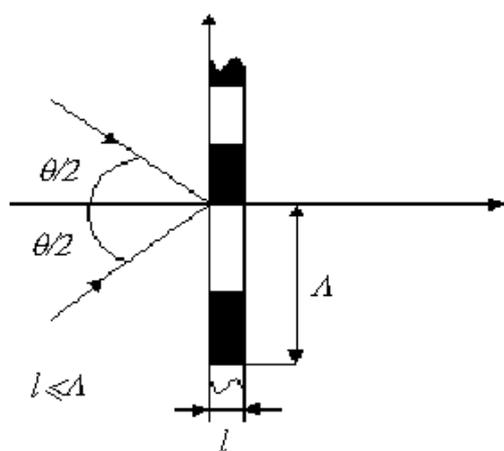


Рисунок 5.8

Если запись голографической структуры происходит не в толщине регистрирующего слоя, а на его поверхности или если толщина слоя l достаточно мала по сравнению с расстоянием между соседними элементами структуры Λ , то такие голограммы называют двумерными или плоскими.

На рисунке 5.8 представлена схема записи и структура такой голограммы. Например, если излучением рубинового лазера ($\lambda \approx 0,7$ мкм) производится запись голографической решетки на фотоматериале с эмульсией толщиной $l=2$ мкм, то период Λ

структуры определяется выражением

$$\Lambda = \frac{l}{2 \sin (\theta / 2)} . \quad (5.16)$$

Период решетки равен 2 мкм, если угол θ между пучками равен 10° . В этом случае записана плоская голограмма-решетка.

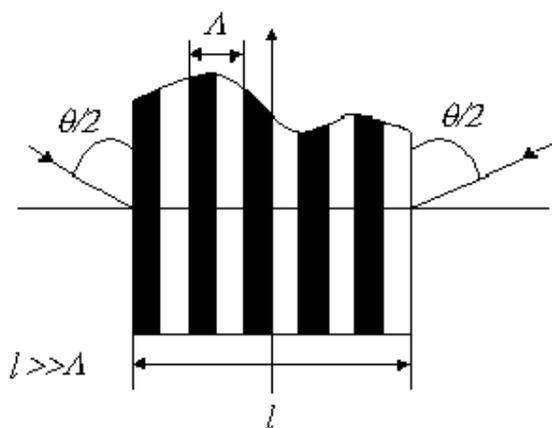


Рисунок 5.9

Если толщина светочувствительного слоя много больше расстояния между соседними поверхностями интерференционных максимумов, то голограмму можно рассматривать как объёмную или трёхмерную. Объёмные свойства голограмм особенно чётко проявляются в схеме со встречными пучками. На рисунке 5.9 изображена схема записи объёмной голограммы. Если углы, под которыми сходятся опорная и предметная волна составляют 80° с плоскостью голограммы, то период решетки записанный излучением рубинового лазера, составит 0,35 мкм. Фотопластинки, предна-

значенные для записи голограмм во встречных пучках, политы эмульсией толщиной около 15 мкм. Таким образом, в данном примере записывается объёмная голограмма. Критерий перехода от двумерных голограмм к трёхмерным определяется выражением

$$l \geq 1,6 \frac{\Lambda^2}{\lambda_n}, \quad (5.17)$$

где λ_n - длина волны света в регистрирующей среде.

5.4.4 Другие виды классификации

Интерференционная структура может быть зарегистрирована светочувствительным материалом одним из следующих способов:

– в виде вариаций коэффициента пропускания или отражения. Такие голограммы модулируют при восстановлении волнового фронта амплитуду освещающей волны и поэтому носят название амплитудных;

– в виде вариаций толщины или коэффициента преломления. Такие голограммы модулируют фазу освещающей волны и поэтому носят название фазовых голограмм.

Во многих случаях одновременно осуществляется как фазовая, так и амплитудная модуляции. Например, обычная фотопластинка регистрирует интерференционную структуру в виде вариаций почернения, показателя преломления и рельефа. После отбеливания такой голограммы остаётся только фазовая модуляция.

Обычно зарегистрированная на голограмме структура сохраняется долго и процесс записи отдалён во времени от процесса восстановления волнового фронта. В этом случае голограмму называют стационарной. Однако существуют среды (некоторые красители, кристаллы, пары металлов), которые почти безинерционно реагируют своими фазовыми и амплитудными характеристиками на воздействие освещённости. В этом случае голограмма существует только во время воздействия на среду предметной и опорной волн, а восстановление волнового фронта производится одновременно с записью, в результате взаимодействия опорной и предметной волн с образованной ими же структурной голограммы. Такие голограммы называются динамическими.

5.5 Основные свойства голограмм

1. Основное свойство голограммы, отличающее её от обычной фотографии, заключается в том, что на фотографии регистрируется лишь распределение освещённости в падающей на неё световой волне, в то время как на голограмме регистрируется, кроме того, распределение фазы предметной волны относительно фазы опорной волны. Информация об амплитуде предметной волны записана на голограмме в виде контраста интерференционной структуры, а информация о фазе - в виде формы и частоты интерференционных полос. В ре-

зультате голограмма при освещении её опорной волной восстанавливает копию предметной волны со всеми амплитудными и фазовыми подробностями.

2. Обычно амплитудные голограммы регистрируют на негативном фотоматериале. При этом свойства голограммы остаются такими же, как и у позитивной голограммы – светлым местам объекта соответствуют светлые места восстановленного изображения, а тёмным – тёмные. Это свойство голограммы легко понять, принимая во внимание, что информация об амплитуде предметной волны заключена в контрасте интерференционной структуры, распределение которого на голограмме не меняется при замене позитивного процесса на негативный. При такой замене лишь сдвигается на π фаза восстановленной предметной волны, что незаметно при визуальном наблюдении, но проявляется в некоторых экспериментах по голографической интерферометрии.

3. В тех случаях, когда при записи голограмм свет от каждой точки объекта освещает всю поверхность голограммы, каждый малый участок последней способен восстановить всё изображение объекта. Конечно, меньший участок голограммы восстановит соответственно меньший участок волнового фронта, несущего информацию об объекте. Если этот участок будет очень мал, то качество восстановленного изображения ухудшится.

В случае голограмм сфокусированного изображения каждая точка объекта посылает свет на соответствующий ей малый участок голограммы. Поэтому фрагмент такой голограммы восстанавливает лишь соответствующий ему участок объекта.

4. Полный интервал яркостей, передаваемый фотографической пластинкой, как правило, не превышает одного-двух порядков. Между тем реальные объекты часто имеют гораздо большие перепады яркостей. Голограмма, обладая фокусирующими свойствами, использует для построения наиболее ярких участков изображения свет, передающий на всю её поверхность, и способна передать градации яркости до пяти-шести порядков.

5. Свойство "некогерентного" сложения или "памяти". На одной и той же фотопластинке можно независимо записать несколько голограмм. Это можно сделать изменяя угол падения опорного пучка. При восстановлении под разными углами, голограммы можно поочерёдно считывать. На применении этого свойства основан принцип действия некоторых типов голографических запоминающих устройств.

6. Свойство взаимности опорного и предметного пучков. Обычно при восстановлении голограммы стремятся получить изображение предмета (действительное или мнимое). Для этого голограмму освещают опорной волной, а голограмма восстанавливает предметную. Но если освещать голограмму предметной волной на этапе восстановления, то результатом такой операции будет воссоздание опорной волны. Это свойство оказывается полезным при разработке голографических систем опознавания образов.

7. Особыми свойствами обладают объёмные голограммы. Такие голограммы представляют собой трёхмерную структуру, в которой поверхности узлов и пучностей зарегистрированы в виде вариаций показателя преломления

или коэффициента отражения среды (рисунок 5.10). Такая структура при освещении её опорной волной действует подобно объемной дифракционной решётке.



Рисунок 5.10

Свет, зеркально отражённый от слоёв, будет восстанавливать волну от предмета. Действительно, поверхности узлов и пучностей направлены по биссектрисе угла, который составляют предметный и опорный пучки, что и обеспечивает это свойство трёхмерных голограмм.

Пучки, отражённые от разных слоёв, будут усиливать друг друга только в том случае, если они синфазны, то есть разность хода между ними должна равняться λ . Указанное условие Липпмана - Брэгга будет автоматически выполняться лишь для

той длины волны, в которой регистрировалась голограмма. Это приведёт к избирательности голограммы по отношению к длине волны источника, в свете которого происходит восстановление волнового фронта. Поэтому появляется возможность восстанавливать изображение с помощью источника сплошного спектра (лампа накаливания, Солнце). Если голограмма экспонировалась в свете нескольких спектральных линий (например, синей, зелёной и красной), то каждая длина волны образует свою трёхмерную структуру. Соответствующие длины волн будут выделяться из сплошного спектра при освещении голограммы, что приведёт к восстановлению не только структуры, но и спектрального состава световой волны, то есть к получению цветного изображения.

Яркость голограмм принято оценивать величиной дифракционной эффективности η , которая определяется как отношение интенсивности света, формирующего изображение I_1 при восстановлении голограммы, к интенсивности восстанавливающего пучка I , т.е.

$$\eta = I_1 / I. \quad (5.18)$$

Максимально достижимые эффективности голограмм разных типов, выраженные в процентах, представлены в таблице 5.1.

Таблица 5.1

Тип голограмм	Пропускающие		Отражательные	
	Амплитудные	Фазовые	амплитудные	фазовые

<i>Двумерные</i>	6,25	33,9	6,25	100
<i>Трёхмерные</i>	3,7	100	7,2	100

5.6 Пространственная фильтрация и опознавание образов

В различных областях науки и техники нередко приходится прибегать к выделению определенного сигнала из совокупности сигналов, более или менее отличающихся от него /10/. Такую задачу решают, например, радист, выделяющий излучение одной определенной радиостанции из заполняющих эфир излучений тысяч станций; спектроаналитик, выискивающий в сложном частотном спектре спектральных линий анализируемого образца линии, принадлежащие определенному элементу или криминалист, сопоставляющий отпечатки пальцев на месте преступления, с имеющимися в картотеке.

Существует общий способ оптимального решения таких задач. Рассмотрим его сначала применительно к задачам, которые решает радист или спектроаналитик. Объединяет эти задачи то, что электромагнитные излучения различных радиостанций, так же как и оптические излучения различных элементов, отличаются спектром своих частот. В каждом радиоприемнике есть элемент, называемый частотным фильтром, который пропускает лишь излучения определенных частот.

В спектральном анализе также иногда используются эти фильтры, но так как изготовить их узкополосными достаточно трудно, то фильтрацию часто осуществляют следующим образом. Сначала исследуемое излучение разлагают в спектр, т. е. направляют световые волны разных частот под разными углами и, следовательно, в разные точки фокальной плоскости. Эта операция обычно производится с помощью спектрографа с призмой или дифракционной решеткой. Затем полученный таким образом спектр сопоставляют со спектрами известных элементов и выискивают в них совпадающие спектральные линии. Для автоматизации этого процесса в фокальной плоскости спектрографа можно установить выходные щели в тех местах, где расположены линии различных элементов, и следить за сигналами фотоумножителей, установленных за этими щелями.

Такой способ обладает тем недостатком, что выделяются не все спектральные линии каждого элемента, а лишь одна. При этом велика вероятность помех: в полосу спектра, вырезаемую щелью, может попасть линия другого элемента — сигнал, зарегистрированный фотоумножителем, окажется ложным.

Предложен и более совершенный способ фильтрации. В фокальную плоскость спектрографа помещают позитив спектра разыскиваемого элемента, полученный на том же спектрографе. За пластинкой устанавливают свето-сборник и фотоумножитель, сигнал которого будет теперь определяться степенью корреляции всего анализируемого спектра (а не одной линии) со спектром данного элемента. Эта система представляет собой идеально согласованный фильтр. Влияние помех и случайных совпадений на сигнал в этом случае минимально.

Мы так подробно описали этот процесс потому, что операция фильтрации изображений осуществляется с помощью голограмм совершенно аналогично, с той лишь разницей, что изображение разлагается в спектр пространственных (а не временных) частот, причем разложение (и фильтрация) осуществляется одновременно в двух координатах.

Следует заметить, что основы излагаемых ниже идей заложены немецким оптиком Эрнстом Аббе около ста лет тому назад.

Каждое двумерное изображение может быть разложено в двумерный же спектр пространственных частот. Эта операция соответствует представлению изображения в виде набора синусоидальных дифракционных решеток разных периодов и ориентаций, аналогично тому, как в радиотехнике или спектроскопии при разложении сигнала в спектр его представляют в виде набора синусоидальных колебаний разных частот.

Операция разложения изображения транспаранта в спектр по пространственным частотам осуществляется обычно с помощью линзы (левая часть рисунка 5.11).

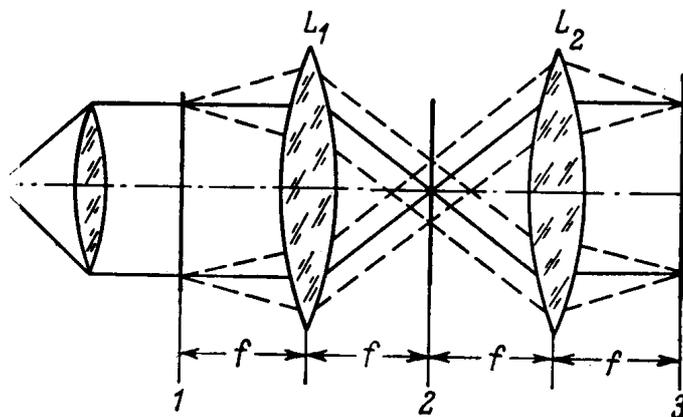


Рисунок 5.11

На этом рисунке двумерный объект – транспарант помещен в плоскость 1 и просвечивается плоскопараллельным пучком света. Каждая из синусоидальных решеток, на которые можно разложить изображение, действует независимо. Решетка большей пространственной частоты отклоняет лучи первых порядков на большие углы. Эти лучи фокусируются линзой L_1 в точку, удаленную от центра плоскости 2. Решетки с меньшим периодом создают на плоскости 2 освещенные точки, менее удаленные от центра. Примеры спектров, полученных таким образом, даны на рисунке 5.12, где изображены пространственные спектры букв (слева — буква, справа — ее спектр). Не останавливаясь на математической стороне вопроса, упомянем только, что операция, производимая с изображением в такой схеме, называется преобразованием Фурье.

Отметим важные свойства установки на рисунке 5.11. При вращении транспаранта вокруг оптической оси будет вращаться и спектр. Спектр также меняется при изменении масштабов транспаранта: расширяется при его уменьшении и сужается при его увеличении. Поступательное движение транспаранта

в плоскости 1 на спектре не скажется. Пучок нулевого порядка создает в центре плоскости 2 яркую точку, соответствующую постоянному члену в разложении изображения в ряд Фурье.

Теперь рассмотрим весь рисунок 5.11 (включая и его правую часть). Нетрудно видеть, что софокусные линзы L_1 и L_2 построят в плоскости 3 перевернутое изображение транспаранта.

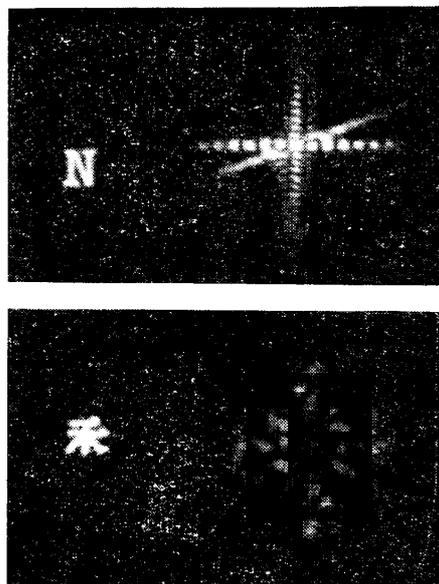


Рисунок 5.12

Помещая в плоскость 2 различные фильтры или маски, мы можем пропускать для образования изображения те или иные части пространственного спектра объекта. Так осуществляют, например, исправление изображения, ослабляя или выявляя высокие и низкие пространственные частоты. Можно из всего изображения выделить только определенные детали, например, из страницы текста только буквы N. Для этого нужно поместить в плоскость 1 (рисунок 5.11) транспарант анализируемой страницы текста; в плоскости 3 мы увидим его изображение. Если теперь в плоскость 2 ввести фильтр пространственных частот буквы N (рисунок 5.12), то с изображения страницы исчезнут все детали, кроме этой буквы.

Приготовленный специальным образом фильтр, помещенный в эту плоскость, может также улучшить качество изображения, исключив искажения, внесенные ранее аберрациями оптической системы, неоднородностями атмосферы или неблагоприятными условиями съемки (смазывание изображения вследствие перемещения объекта во время экспозиции).

Изложенный способ пространственной фильтрации изображения обладает существенным недостатком: фильтр содержит не всю информацию об объекте, по которому он изготовлен, и фазовая информация теряется при записи. Поэтому световой сигнал на выходе системы содержит паразитные компоненты, накладывающиеся на опознаваемые изображения и затрудняющие интерпретацию результатов.

В голографическом способе получения согласованного пространственно-го фильтра фазовая информация об объекте сохраняется, и шумы резко снижаются. Схема получения голографического согласованного фильтра пространственных частот представлена в левой части рисунка 5.13. В плоскости 2 по-прежнему образуется фурье-образ транспаранта, помещенного в плоскость 1, но в результате интерференции с когерентным фоном, создаваемым клином, в плоскости 2 образуется голографическая дифракционная решетка, так называемая фурье-голограмма.

Теперь уже не надо делать позитивную копию с голограммы-фильтра — мы знаем, что это никаких свойств голограммы не изменит.

Поместив в плоскость 1 объект, а в плоскость 2 голографический фильтр какой-либо его части, мы по-прежнему увидим изображение в середине плоскости 3 за счет нулевого порядка — фильтр практически его не исказит, а только несколько ослабит. В изображении первого порядка мы увидим яркие

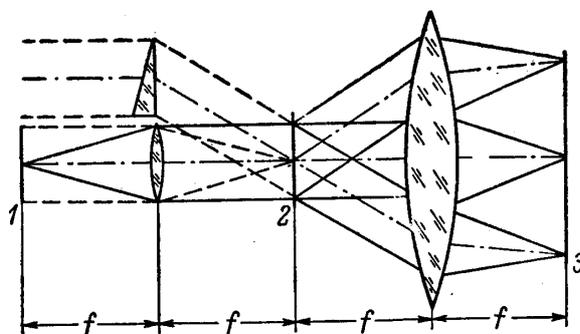


Рисунок 5.13

точки опознавания, координаты которых соответствуют распределению по объекту тех его деталей, с которых снят голографический фильтр.

Изложенный здесь метод опознавания образов тем надежней, чем сложнее объект, который надо опознать. Удовлетворительные результаты получены, например, при опознавании отпечатков пальцев, причем даже при сохранении незначительной части отпечатка яркость точки опознавания сохраняется достаточно высокой. На базе устройства, схема которого представлена на рисунке 5.13, уже работают автоматические чтецы, осуществлено выделение предметов определенной формы или направления на аэрофотоснимках, разработаны приборы для обработки геофизических данных и т. д.

Фильтрующие свойства голограммы легко понять, если вспомнить об отмечавшейся ранее обратимости опорной и предметной волн. Если голограмму, полученную при экспонировании пластинки светом от предметов A и B , осветить затем волной от A , то восстановится волна от предмета B . Наоборот, если ту же голограмму осветить предметом B , восстановится волна, рассеянная A . Если A — точечный источник, то его изображение голограмма создаст лишь в случае, когда освещающий предмет — B .

Голограмма, таким образом, осуществляет операцию узнавания "своего" предмета. Тот же принцип может быть применен для воссоздания целого изображения по его части.

6 Оптические запоминающие устройства

6.1 Регистрирующие среды

Регистрирующая среда (носитель информации) является основным компонентом запоминающего устройства (ЗУ) и служит для регистрации и хранения данных, представленных в побитовой или голографической форме [4]. Рассмотрим некоторые типы регистрирующих материалов, применяемых для голографической записи информации. Существующие регистрирующие материалы и устройства чувствительны только к интенсивности света и изменяют свои оптические свойства в соответствии с ее пространственным распределением. В различных материалах это происходит по-разному: в одних изменяется амплитудное пропускание, в других — показатель преломления или рельеф поверхности. В первом случае образуется амплитудная голограмма, а во втором — фазовая. В некоторых материалах (например, халькогенидные стеклообразные полупроводники) наблюдается одновременное изменение, как амплитудного пропускания, так и показателя преломления, что приводит к образованию амплитудно-фазовых голограмм. Магнитные голограммы, получаемые термомагнитным способом на тонких магнитооптических пленках, представляют собой третий тип голограмм — поляризационный.

Важнейшей особенностью регистрирующего материала является свойство обратимости, благодаря которому можно обновлять информацию на носителе, т. е. стирать записанную голограмму и записывать новую. К необратимым материалам относится обычный фотоматериал. Для построения оперативных ЗУ необходимы обратимые регистрирующие материалы. К числу обратимых регистрирующих материалов относятся магнитооптические пленки, термопластические и фотохромные материалы, электрооптические кристаллы, халькогенидные стеклообразные полупроводники и др. Особый интерес представляют материалы с объемными изменениями оптических свойств показателя преломления, которые позволяют осуществить трехмерное хранение информации в виде матрицы наложенных объемных голограмм с высокой дифракционной эффективностью. К таким материалам относятся электрооптические кристаллы и, в частности, хорошо известный ниобат лития.

Характеристики, которыми должен обладать регистрирующий материал, пригодный для использования в голографических запоминающих устройствах (ГЗУ) оперативного типа следующие:

Разрешающая способность ν_m , лин/мм	≥ 1000
Дифракционная эффективность η_n , %	$\geq 0,1$
Энергия записи E_s , мкДж/мм ²	$\leq 0,5$

Время записи τ_w , мс	$\leq 0,1$
Время стирания τ_e , мс	$\leq 0,1$
Допустимое число циклов перезаписи M ,	10^8

Было исследовано большое число регистрирующих материалов. Однако материал, удовлетворяющий одновременно всем перечисленным требованиям, пока еще не найден. Но ряд материалов полностью удовлетворяет большей части требований и хорошо согласуется с остальными. Характеристики регистрирующих материалов, пригодных для использования в оптических устройствах хранения и обработки информации, представлены в таблице 6.1.

Галогенидосеребряные фотоматериалы нашли широкое применение в голографии благодаря высокой светочувствительности и разрешающей способности, а также доступности. Так как такие фотоматериалы не допускают перезаписи, они не могут быть использованы в ГЗУ оперативного типа (ГОЗУ); их применяют в качестве носителей информации для ГЗУ постоянного типа (ГПЗУ).

Термопластические материалы относятся к числу регистрирующих сред, в которых сочетаются постоянство хранения, возможность стирания и повторной записи. Термопластик сам по себе нечувствителен к свету, поэтому его соединяют с фотопроводниковым материалом в одну пленочную структуру. Готовый носитель представляет собой многослойную структуру и состоит из стеклянной подложки с токопроводящим покрытием (обычно из оксида индия InO) с нанесенными на нее слоями фотопроводника и термопластика. Токопроводящий слой служит для нанесения электрического заряда на поверхность термопластика с помощью установки коронного разряда, а также используется в качестве тепловыделяющего элемента при тепловом проявлении зарегистрированной фазовой голограммы.

Для существующих способов записи голограмм на фототермопластике характерны два основных момента: 1) создание электрического поля в слое термопластика, соответствующего распределению интенсивности света при экспонировании; 2) тепловое проявление, при котором под действием кратковременного теплового импульса температура термопластика повышается до температуры размягчения, в результате чего электростатические силы деформируют поверхность термопластика в соответствии с распределением электрического поля на ней, образуя двумерную фазовую голограмму. Фиксация полученной голограммы осуществляется путем быстрого охлаждения («замораживания») до комнатной температуры. Для стирания записанной голограммы достаточно вновь нагреть термопластик до температуры размягчения. В отличие от других обратимых материалов для получения голограмм на фототермопластике требуется небольшая экспозиция порядка 10 мкДж/мм, что обусловлено высокой чувствительностью фотопроводящего слоя. Дифракционная эффективность фазовых голограмм, записанных на, составляет 10-30 %. Несмотря на сложность механизма записи, фото термопластик на сегодняшний день является лучшим обратимым материалом.

Основными недостатками фототермопластика как носителя информации для ГОЗУ являются малая скорость записи – стирание и небольшая цикличность. Было достигнуто время цикла запись – стирание порядка 100 мс и до

Таблица 6.1

Вид регистрирующей среды	Спектральная чувствительность $\Delta\lambda, \text{мкм}$	Энергия записи $E_s, \text{мДж/мм}^2$	Разрешающая способность $\nu_m * 10^3, \text{лин/мм}$	Тип голограммы	Дифракционная эффективность $\eta_{\text{max}}, \%$	Время записи τ_W	Время стирания τ_E	Допустимое число циклов		Продолжительность хранения
								Записи, стирания	Считывания	
Галогенидосеребряные эмульсии	0,4 – 0,7	$10^{-5} - 10^{-4}$	1 – 10	<i>Амплитудная и фазовая (отбеленные)</i>	20 – 50	Минуты	---	---	Не ограничено	Не ограничено
Термопластические материалы	0,488-0,633	$10^{-3} - 10^{-2}$	0,6 - 2	Фазовая двумерная	10-30	10 мс	100 мс	$<10^3$	то же	то же
Фотохромные материалы	Активная 0,3 – 0,45 Отбелен. 0,55 – 0,7	0,1 – 1,0	1,5 - 10	Амплитудная трехмерная	0,4 – 3	10 мс	Такое же, как при записи	$<10^5$	Имеет место обесцвечивание	От минут до месяца
Электрооптические кристаллы	0,488 – 0,633	0,02 – 0,1	1,5	Фазовая трехмерная	60 – 80	30 нс. - мин	То же	Не ограничено	Считывание при другой λ	От недель до месяцев
Магнитооптические пленки (MnBi)	Неселективные	0,1 – 1,0	1,5 – 2,0	Поляризационная	0,01	10–100 нс.	0,1 мс	То же	Не ограничено	Не ограничено

Халько- генидные стеклооб- разные полупро- водники	0,3 – 0,75	10 - 100	3 - 10	Амплитуд- но-фазовая	20 – 80	1-10 мин	Не- сколь- ко минут	Неиз- вестно	<i>же</i> <i>To</i>	То же
---	------------	----------	--------	-------------------------	---------	-------------	------------------------------	-----------------	----------------------------	-------

1000 циклов записи – стирания без изменения дифракционной эффективности. Скорость записи—стирания ограничивается скоростью отвода теплоты от подложки после воздействия каждого проявляющего и стирающего теплового импульса. Поэтому для увеличения быстродействия необходимо создать устройства для эффективного быстрого отвода тепла от подложки и уменьшить длительность тепловых проявляющих и стирающих импульсов. Увеличение цикличности возможно при уменьшении вредного воздействия коронного разряда на термопластический носитель. Недостатком термопластических материалов является также то, что они реагируют только на определенную полосу пространственных частот. В связи с этим существует соотношение между дифракционной эффективностью и разрешающей способностью, которое зависит от угла между опорным и объектным пучками. Максимальный отклик приблизительно соответствует пространственной частоте $(2/h)$, где h – толщина термопластического слоя. В обе стороны от центральной пространственной частоты отклик, быстро снижаясь, падает до нуля. Эффективная ширина полосы пространственных частот – 50% от центральной частоты. Если, например, толщина термопластического слоя равна 0,5 мкм, то центральная пространственная частота составляет 1000 лин/мм, что соответствует углу 30° между опорным и объектным световыми пучками при использовании аргонового лазера в качестве источника света. При этом ширина полосы пространственных частот 500 лин/мм, поэтому эффективно могут быть зарегистрированы детали интерференционной картины с пространственной частотой 750 – 1250 лин/мм.

Электрооптические кристаллы типа ниобата лития ($LiNbO_3$), ниобата стронция и бария $[(Sr, Ba)Nb_2O_6, SbN$ или $Ba_{0,75}Sr_{0,25}Nb_2O_6]$ позволяют получить высокоэффективные трехмерные фазовые голограммы. Процесс записи основан на оптическом возбуждении и последующем дрейфе или диффузии фотоэлектронов, которые захватываются новыми узлами кристаллической решетки. Вследствие такого перераспределения зарядов в кристалле возникает электрическое поле, которое изменяет показатель преломления вследствие электрооптического эффекта. Таким образом, создается трехмерная фазовая голограмма, дифракционная эффективность которой может достигать теоретического предела – 100%. Голограмма фиксируется воздействием электрического поля или теплоты. Стирание может быть осуществлено оптическим путем либо нагреванием кристалла с одновременным облучением.

Наиболее исследованным электрооптическим кристаллом является ниобат лития. Он, как и другие электрооптические кристаллы, обладает высокой дифракционной эффективностью, требуемой для носителя ГЗУ разрешающей способностью, неограниченной цикличностью и незначительным светорассеянием. С помощью этого кристалла были получены голограммы с дифракционной эффективностью 60—80 %. Ниобат лития в чистом виде имеет очень низкую светочувствительность. Однако введение примеси позволяет ее значительно повысить. Например, введение добавок железа дает возможность повысить светочувствительность ниобата лития в сотни раз.

Ниобат лития допускает высокую скорость записи голограмм. Важным его достоинством является возможность записи большого числа голограмм с необходимой дифракционной эффективностью на одном и том же участке кристалла путем наложения их друг на друга. Число наложенных голограмм ограничивается диапазоном изменения показателя преломления. Наибольшая наблюдавшаяся модуляция показателя преломления ниобата лития составляет $4 \cdot 10^{-4}$. Но даже при таком небольшом динамическом диапазоне показателя преломления можно записать сотни голограмм путем наложения их друг на друга. Благодаря высокой угловой селективности трехмерных голограмм легко осуществляется селективное восстановление наложенных голограмм. Возможно также их селективное стирание.

К недостаткам кристаллов ниобата лития относятся малая светочувствительность, нестабильность и низкая термостойкость записанных голограмм. Нестабильность особенно проявляется при наложении их друг на друга. Так как при этом не производится фиксирования записанных голограмм, при наложении следующих происходит частичное разрушение (стирание) предыдущих. Несмотря на указанные недостатки, ниобат лития по своим физическим свойствам наиболее близок к идеальному регистрирующему материалу для ГОЗУ с трехкоординатной выборкой.

6.2 Голографические системы записи информации

6.2.1 Запись двумерных массивов информации в виде голограмм

Рассмотрим метод записи информации в виде двумерных голограмм /11/. Исторически развитие голографических систем памяти началось с исследования именно этого метода (а не одномерных или трехмерных голограмм), когда ввод информации осуществляется постранично, т.е. большими массивами сигналов.

Однако голографические ЗУ, использующие этот принцип, не получили достаточного развития. Это, прежде всего, связано с недостаточным развитием элементной базы таких ЗУ, которая оказывается более сложной (импульсные лазеры, управляемые транспаранты и другое), чем в случае дисковых ОЗУ с построчной записью в виде одномерных голограмм. Другая причина, быть может более важная, - это неподготовленность методов обработки двумерных массивов цифровых сигналов, что не позволяет в полной мере достоинства этого класса голографических ЗУ (высокий темп считывания - до 10^{10} бит/с и др.).

Последние успехи в развитии оптических методов обработки цифровых сигналов (в частности, созданием многоэлементной матрицы оптических логических элементов), а также возросший интерес к проблеме оптических вычислительных машин позволяют думать, что указанные препятствия будут устранены и голографические ЗУ с записью информации массивами найдут широкое применение в оптической вычислительной технике.

Исходная информация, записываемая в виде двумерной голограммы на светочувствительную поверхность носителя, может быть представлена как в аналоговой, так и в цифровой форме. В первом случае это могут быть, тексты, формулы, чертежи, карты или какие-либо другие изображения, которые записываются на промежуточный носитель, например рулонную фотопленку (рисунок 6.1). Во втором, первичная информация представляет собой многоразрядные числа в двоичном коде с числом разрядов порядка десятков или сотен. Эти числа, выражаемые определённой последовательностью электрических импульсов, предварительно также записываются на промежуточный носитель. Каждое многоразрядное число после этого будет представлено в виде набора светлых и темных пятнышек. Заполненный такими числами кадр принято называть транспарантом.

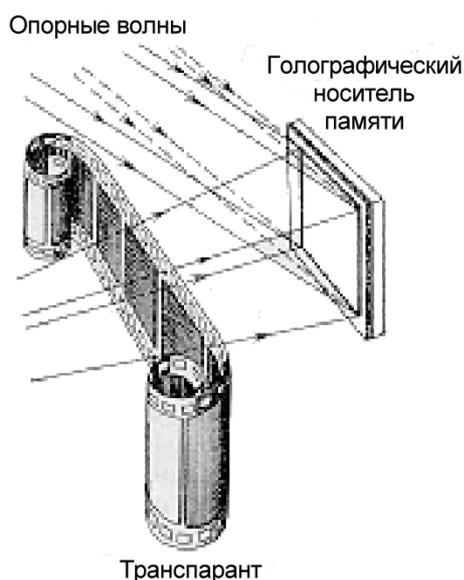


Рисунок 6.1

изменять или записывать их на отдельные небольшие участки (рисунок 6.1).

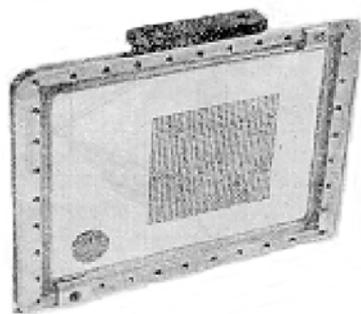


Рисунок 6.2

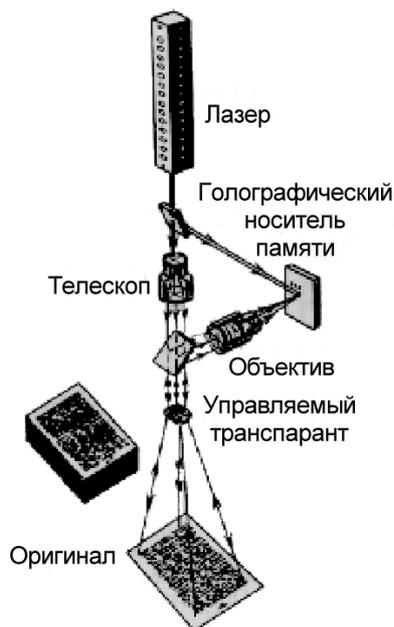
В каждом транспаранте содержится обычно 10^3 - 10^5 пятнышек, т.е. единиц, или бит, информации. По мере заполнения одного кадра пленка сдвигается и записывается следующий кадр. Т.е., первичная информация преобразуется к виду удобному для голографирования. Если, например, объем памяти голографического носителя должен составлять 10^8 бит, а в каждый транспарант записывается 10^5 бит информации, то потребуется промежуточный носитель, содержащий 1000 кадров. Далее вся информация вводится в основной носитель памяти, в простейшем случае - в фотопластинку с высокоразрешающей эмульсией, на которую последовательно записываются голограммы всех транспарантов. При этом голограммы можно накладывать друг на друга, наклон опорной волны при этом следует каждый раз

В методе наложения голограмм надежность записи выше, меньше влияют помехи типа царапин, пылинок и др., используется вся площадь фотопластинки; во втором методе меньше проявляются шумы голограмм и достигается более высокая плотность записи. Оптимальным является сочетание этих способов. В этом случае голографический носитель информации содержит растр микроголограмм, в каждой из которых можно записать под разными углами несколько транспарантов.

На рисунке 6.2 представлен голографический носитель - фотопластинка, на которой запи-

сано 10^7 бит, 32×32 микроголограмм с шагом 1,5 мм (всего $\sim 5 \times 5$ см²). Обычно фотопластинка отбеливается, при этом голограммы становятся фазовыми, это приводит к уменьшению потерь света при считывании информации.

Описанный выше процесс ввода информации в голографический носитель неудобен и требует большого времени. Поэтому на практике применяются быстродействующие системы записи, основанные на использовании управляемых транспарантов. В случае записи изображения в виде УТ можно использовать, например, устройство типа МДП-ЖК, представляющее собой многослойную структуру на основе жидких кристаллов, управляемую внешним электрическим полем.



На рисунке 6.3 показана схема записи микроголограмм в голографический носитель. Исходный документ освещается и проецируется на поверхность управляемого транспаранта, запоминающего эту информацию на время микроголограммы. Затем с помощью внешнего высокочастотного электрического поля информация стирается, и повторяется процесс записи следующего документа. Таким способом можно быстро и без промежуточных операций вводить самую разнообразную информацию в голографический носитель памяти.

С помощью управляемого транспаранта можно также записывать информацию, передаваемую по каналу связи. Для этого используется приемопередающая телевизионная; исходная информация воспринимается передающей телевизионной камерой, а изображение с экрана монитора проецируется на поверхность транспаранта.

Рисунок 6.3

6.2.2 Принцип параллельной записи сигналов на движущийся носитель в виде голограмм

Параллельная запись большого числа сигналов одновременно по многим каналам заложена в самом принципе голографической записи и радикально решает проблему повышения быстродействия оптических ЗУ. Переход к многоканальности позволяет легко увеличить скорость записи и считывания сигналов (в десятки и сотни раз) и в случае необходимости заметно понизить скорость перемещения носителя. (Если скорость записи побитовым способом на вращающийся оптический диск увеличивать за счет частоты вращения, то это неизбежно приведет к увеличению числа ошибок, возникновению технических трудностей и т.д.). Принцип создания многоканальных ЗУ основан на возможности одновременной записи большого числа голограмм на одну и ту же по-

верхность фоточувствительного материала. На рисунке 6.4 в самом общем виде показана схема параллельной записи сигналов по многим каналам.

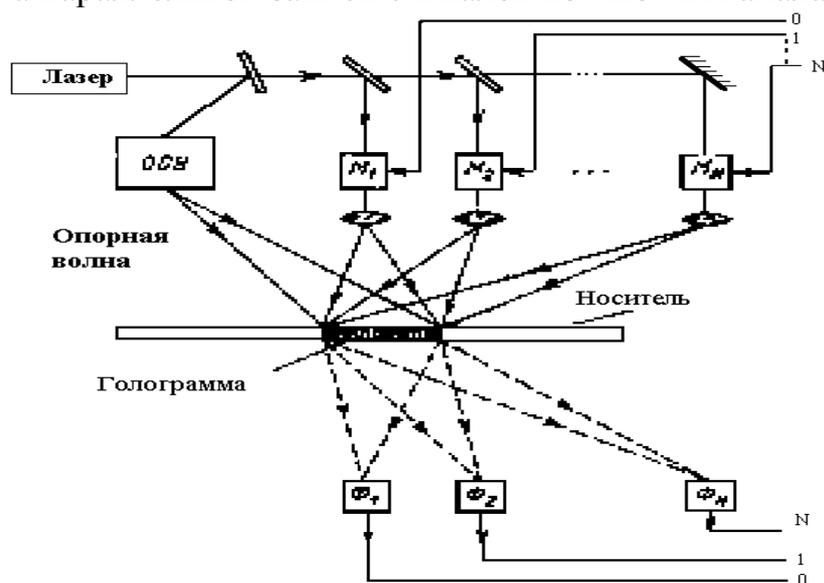


Рисунок 6.4

Луч лазера делится на два луча: опорный и сигнальный. Сигнальный луч, в свою очередь, делится на N каналов в соответствии с числом одновременно записываемых двоичных сигналов. В каждом из этих каналов имеется модулятор M , работающий в режиме выключателя. При наличии управляющего напряжения он пропускает луч лазера, а при отсутствии напряжения будет непрозрачным. Таким образом, на выходе модулятора возникает комбинация M лучей, которые вместе с опорным лучом записываются в виде голограммы на один и тот же участок носителя. Следующая комбинация N лучей записывается на соседний участок фоточувствительного слоя, который смещается на соответствующее расстояние.

При считывании информации опорный луч, проходя через оптоэлектронную систему управления (ОСУ), просвечивает микроголограмму, в результате чего формируется изображение в виде системы ярких точек, число и взаимное расположение которых определяются комбинацией включенных модуляторов при записи. Это изображение проецируется на систему фотоприемников Φ , на выходе которых восстанавливаются записанные электрические сигналы. Число каналов при голографической записи может достигать сотен и, в специальных случаях, тысяч. В соответствии с этим должно быть выбрано и число элементов модуляторов и фотоприемников, выполняемых в виде интегральных матриц или линеек.

Таким образом, в голографических ЗУ, благодаря многоканальности, реализуется высокая скорость записи и считывания сигналов, что особенно важно для супер-ЭВМ.

В рассмотренном методе записи обеспечивается также высокая надежность и сохраняемость записанной информации. Это связано с тем, что каждый

бит занимает большую площадь и повышение плотности записи не сопровождается ее уменьшением (что имеет место при побитовой записи сфокусированным лазерным лучом). Кроме того, поскольку каждый участок микроголограммы содержит информацию обо всех битах, осуществляется своеобразное резервирование вышедших из строя участков носителя информации. Все это происходит автоматически и не требует каких-либо переключений.

Большим достоинством дисковых голографических ЗУ является то, что ширина спиральных дорожек в десятки и сотни раз больше, чем в существующих оптических дисках с одноканальной записью. Благодаря этому упрощаются требования к электромеханическим системам слежения и управления, а также к быстродействию этих систем. Последнее связано с тем, что при многоканальной записи скорость вращения диска может быть сделана небольшой.

Плотность записи при многоканальной записи сигналов в виде микроголограмм определяется формулой:

$$\rho = \frac{N}{S} = \frac{N_r \cdot N_\varphi}{d_r \cdot d_\varphi},$$

где $N = N_r, N_\varphi$ - общее число каналов, равное числу бит и совпадающее с числом элементов модулятора;

S - площадь голограммы, равная примерно $d_r \cdot d_\varphi$;

d_r - размер голограммы в радиальном направлении,

d_φ - размер в направлении вращения диска.

Однако, при голографическом методе записи имеется возможность увеличения плотности записи путем применения толстых (по сравнению с длиной волны) носителей информации. В этом случае, записав информацию на диск по описанной схеме, следует провести второй цикл записи, изменив с помощью оптоэлектронной системы управления направление опорной волны, и аналогичным способом повторить несколько циклов записи. Естественно, что при считывании информации необходимо после каждого цикла изменять направление опорной волны. Как показывают расчеты и эксперименты, при использовании толстых фоточувствительных носителей можно за счет наложения голограмм увеличить плотность записи почти на порядок и достичь емкости на одном диске более 10^{11} бит, что открывает возможность создания цифрового видеодиска.

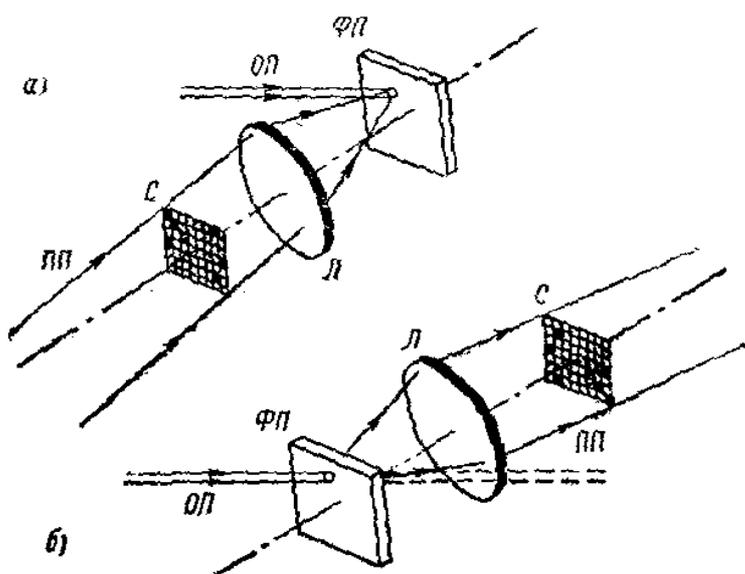
Таким образом, голографический метод записи электрических сигналов открывает возможность создания дисковых ЗУ, которые по основным характеристикам (информационная емкость, скорость считывания и др.) существенно превосходят одноканальные системы с побитовой записью информации.

6.2.3 Голографический способ записи, хранения и считывания цифровой информации

Цифровая информация, предназначенная для голографической записи, разбивается на страницы, и каждая страница записывается в виде отдельной голограммы. Массив данных записывается и хранится в виде матрицы голограмм на специальном регистрирующем материале, называемом носителем информации.

Страница данных, представленная в двоичном коде, подготавливается в виде транспаранта, прозрачные отверстия которого соответствуют двоичным единицам. Такой транспарант называют *входной страницей*. Отверстия круглой или прямоугольной формы пробиваются в узлах регулярной двумерной сетки, наложенной на непрозрачный материал транспаранта. Непрозрачные узлы соответствуют двоичным нулям. Узлы сетки называют информационными точками. Каждая информационная точка несет бит информации. Описанный способ физического представления страницы наиболее удобен как для ее формирования при записи, так и для считывания информации с ее изображения, восстановленного голограммой. В реальных запоминающих устройствах входная страница формируется *устройством набора страниц* (УНС), которое представляет собой пространственный матричный модулятор световой волны с электронной схемой управления. УНС осуществляет пространственную модуляцию проходящей или отраженной световой волны по амплитуде, фазе, поляризации или по совокупности этих параметров. В дальнейшем под входной страницей подразумевается амплитудный транспарант, работающий на пропускание.

Голограмма входной страницы, как правило, записывается по схеме Фурье-голографии (рисунок 6.5 а). Такая запись обладает рядом важных достоинств, которые обусловлены двумя основными свойствами преобразования Фурье:



1) Фурье-образ точечного источника света представляет собой равномерное распределение амплитуд света по всей частотной плоскости; 2) смещение точечного источника в координатной области вызывает только линейный фазовый сдвиг в частотной области. Согласно первому свойству Фурье-голограмма обладает большой избыточностью регистрации и хранения, так как информация о каждой точке вход-

Рисунок 6.5 Голографическая запись (а) и схема восстановления (б) страницы двоичных данных

ной страницы распределяется по всей площади записи. Избыточная регистрация является одним из важнейших достоинств такой голограммы, обеспечивающим высокую надежность хранения и помехозащищенность против локальных дефектов регистрирующей среды, таких, как неоднородность, пыль, царапины и т. п. Поскольку входная страница представляет собой совокупность пространственно разнесенных идентичных световых источников, ее фурье-образ складывается из фурье-образов отдельных информационных точек, которые, согласно второму из приведенных свойств преобразования Фурье, отличаются только фазовыми множителями. Это означает, что фурье-образы всех информационных точек входной страницы попадают на один и тот же участок носителя, расположенного в фурье-плоскости. Следовательно, для записи фурье-голограммы страницы достаточно площади, требуемой для записи только одного бита с необходимой избыточностью. Благодаря этому достигается высокая плотность записи, помехоустойчивость, которая в идеальном случае соответствует помехоустойчивости одного бита на всей площади голограммы. Фурье-голограммы обладают наибольшей информационной емкостью по сравнению с другими типами голограмм.

На рисунке 6.5 б показана схема восстановления страницы с двоичными данными. За линзой формируется восстановленное изображение страницы, которое представляет собой совокупность светлых и темных пятен, соответствующих единицам и нулям данных.

Длительность хранения голограммы без потери голографической информации зависит от свойств регистрирующей среды. Например, в фотографических средах реализуется постоянная запись с неограниченным сроком хранения. Однако они не допускают перезапись, что необходимо для оперативной памяти. Голограммы, записанные на фотохромных материалах, имеют малую продолжительность хранения, так как легко теряют информацию в результате обесцвечивания восстанавливающим световым пучком, а также вследствие тепловой релаксации. К числу регистрирующих материалов, в которых сочетается постоянство записи, возможность стирания и перезаписи, относятся фототермопластические материалы и магнитооптические пленки.

6.3 Оптические диски: принципы и применения

6.3.1 Принцип оптической побитовой записи

Рассмотрим, как осуществляется побитовая запись сигнала на движущийся носитель, широко используемая в современных оптических системах памяти. Ради конкретности в качестве движущегося носителя будем рассматривать диск со светочувствительным покрытием, хотя принцип построения и технические решения являются общими и для других видов носителей (лента, барабан, пластина и другие). Принцип записи поясняет рисунок 6.6.

Излучение лазера с помощью оптической системы фокусируется на поверхность диска в виде небольшого пятна. На поверхности диска нанесено спе-

циальное покрытие в виде тонкой металлической пленки, которая под воздействием лазерного луча меняет свои свойства. В пленке может образоваться углубление, выступ, отверстие или какой-либо другой след (называемый питом), который может быть легко зафиксирован при считывании информации лазером малой мощности.

При вращении диска на спиральных (или кольцевых) дорожках в определенной последовательности образуютсяпиты, соответствующие записываемым информационным сигналам, которые в виде управляющих электрических импульсов поступают на оптический модулятор. Если производить запись цифро-

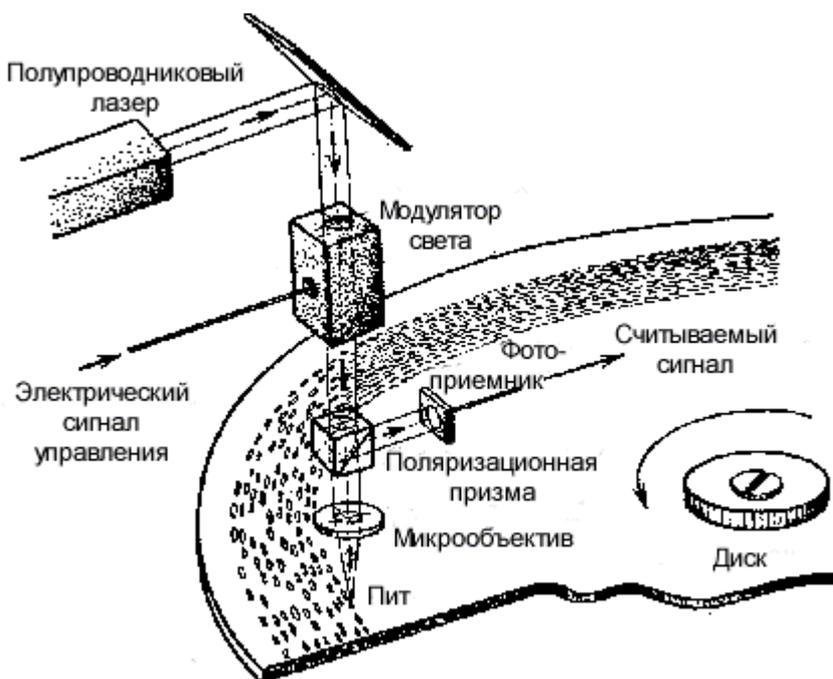


Рисунок 6.6

вых сигналов в двоичном коде, то единицы, т. е. информационные биты, соответствуют открытому состоянию модулятора; при этом лазерное излучение попадает на диск и оставляет след.

Плотность записи ρ определяется диаметром сфокусированного пятна, который примерно равен

$$d \approx 2\lambda F/D, \quad (6.1)$$

где F — фокусное расстояние объектива;
 λ — длина волны;
 D — диаметр выходной линзы.

При $F/D = 1$ размер сфокусированного пятна $d = 2\lambda$ и плотность записи (для случая, когда пятна расположены вплотную друг к другу) определяется формулой

$$c = 1/d^2 = 1/4\lambda^2 . \quad (6.2)$$

Например, на волне полупроводникового лазера, т. е. при $\lambda \approx 0,8$ мкм, плотность записи будет примерно равна $c \approx 3,7 \cdot 10^7$ бит/см², а ширина каждой дорожки имеет порядок микрона. (Следует отметить, что фактический размер записанного пятна может быть значительно меньше, чем предельный размер, определяемый формулой (6.1), — это зависит от физики взаимодействия лазерного излучения с фоточувствительным материалом, на котором записывается информация).

При считывании информации фиксируется изменение коэффициента отражения. При этом сфокусированный лазерный луч не должен перекрывать два соседних бита, иначе появятся помехи. Поэтому если для считывания информации применяется тот же лазер, что и для записи (как показано на рисунке 6.6), то информационные биты (т. е. отверстия) надо располагать с шагом $d \approx 2\lambda$, т.е. порядка 1,6 мкм.

При считывании информации с оптического диска необходимо обеспечить слежение луча лазера за спиральной дорожкой (с точностью порядка 0,1 - 0,2 мкм). Кроме того, в процессе записи и считывания надо скомпенсировать биения диска, чтобы расстояние между выходной линзой и плоскостью диска сохранялось неизменным (иначе размер пятна будет увеличиваться вследствие расфокусировки, что приведет к уменьшению плотности). Требуемая точность слежения за фокусом зависит от глубины резкости линзы. Для применяемых обычно короткофокусных объективов ($F/D \approx 1 - 1,5$) достаточно следить за фокусным расстоянием с точностью 1—2 мкм. Быстродействие систем слежения определяется частотой вращения диска, которая лежит в пределах 500 -3000 оборотов в минуту.

Для того чтобы при считывании свести к минимуму вероятность ошибки, вводится избыточное кодирование информации при записи; на это обычно тратится 20 - 30% информационной емкости диска.

Метод записи информации путем выжигания пятен в тонких (порядка десятых долей микрометра) пленках оказался чрезвычайно удобным при создании оптических дисков. В этом случае достигается разрешение, ограниченное лишь дифракцией, а главное, не требуется химической обработки (или каких-либо других операций) для проявления и фиксации записанной информации. Это позволяет осуществлять контрольное считывание в процессе записи и при обнаружении ошибок проводить повторную запись на соседние участки носителя. Недостатком этого метода является низкая чувствительность указанных носителей (порядка $10^{-1} - 10^{-2}$ Дж/см²), что ограничивает скорость записи информации. Скорость записи информации определяется формулой

$$v = \frac{P}{\sigma/\rho}, \quad (6.3)$$

где P – мощность излучения лазера в сфокусированном пятне в ваттах;

u -плотность энергии записи [Дж/см²];

ρ -плотность записи [бит/см²];

σ/ρ - энергия, затрачиваемая на запись одного бита [Дж/бит].

Подставляя в (6.3) $P = 10$ мВт (случай полупроводникового лазера), $\sigma = 10^{-1}$ Дж/см² и $\rho = 10^7$ бит/см², получим $\nu = 10^6$ бит/с.

В настоящее время существует большое разнообразие оптических запоминающих устройств (ЗУ) с однократной записью информации, которые различаются между собой главным образом объемом памяти и используют диски различных размеров. Например, для диска диаметром 300 мм достигается емкость порядка 1 Гбайт при постоянной частоте вращения порядка 1000 оборотов в минуту. При постоянной линейной скорости емкость может быть увеличена примерно до 1,8 Гбайт. Запись производится полупроводниковым лазером, причем для записи требуется мощность (на поверхности диска) порядка 5 - 8 мВт. При считывании сигналов мощность снижается до 1 мВт. Шаг между записанными битами составляет примерно 1,6 мкм, а размер отверстий в диске - 0,4 мкм по радиусу и около 1 мкм вдоль дорожки.

То обстоятельство, что размер пятна, может быть заметно меньше, чем размер сфокусированного, пятна, определяемый дифракционной формулой (1), открывает возможность увеличения информационной емкости оптического диска. Для этого необходимо ввести в полупроводниковый лазер нелинейный элемент с тем, чтобы он генерировал одновременно на двух длинах волн (на основной частоте и на гармонике). Поскольку мощность при считывании должна быть в 5—10 раз меньше, чем при записи (иначе записанная информация будет разрушаться при считывании), для считывания сигнала мощность излучения лазера на второй гармонике окажется достаточной. Понятно, что такой способ увеличения емкости диска с выжиганием пиков потребует увеличения скорости вращения диска, применения более сложной оптики и т.д.

В оптических дисковых ЗУ можно применять двух лучевые лазеры. При этом луч считывания всего на несколько миллиметров отстает от луча записи, что позволяет сразу выявить ошибку и исправить ее, не дожидаясь следующего оборота диска.

Развитие техники полупроводниковых лазеров и устройств интегральной оптики, существенный прогресс в технологии изготовления и тиражирования оптических дисков создали условия широкого внедрения оптических ЗУ в вычислительную технику. Этому способствуют также высокие темпы развития и производства персональных ЭВМ, для которых требуется высокоинформативная внешняя память, недорогая и простая в эксплуатации. Оптические диски, применяемые в персональных ЭВМ, имеют диаметр 120 или 133 мм и обладают информационной емкостью примерно 600 Мбайт, что во много раз больше емкости магнитных дисков, применяемых в настоящее время.

Следует также сказать о широком применении оптических дисков в бытовой технике. На основе оптических компакт-дисков созданы различные вариан-

ты цифровых проигрывателей звука. При этом достигается высокое качество звучания (отношение сигнал/шум порядка 96 дБ) при длительности воспроизведения более одного часа. Частота вращения диска составляет 500 - 800 оборотов в минуту. В проигрывателе используется высокоточное устройство слежения за дорожкой (погрешность порядка $\pm 0,1$ мкм). Довольно жесткие требования предъявляются к качеству материалов, как диска-оригинала, так и массовых копий. В проигрывателе используется помехоустойчивое кодирование информации (на основе кода Рида — Соломона) для повышения надежности считывания.

Главное преимущество оптических проигрывателей перед обычными состоит в том, что они не портятся в процессе эксплуатации, поскольку считывание информации с оптических дисков производится лазерным излучением небольшой мощности, в то время как в обычных дисках вследствие механического контакта с иглой качество звучания ухудшается с увеличением числа проигрываний. Цифровые оптические проигрыватели быстро завоевывают широкий рынок, постепенно вытесняя проигрыватели с обычными грампластинками.

Определенное распространение получили также видеопроигрыватели, использующие оптические диски. В этом случае записывается не цифровой, а аналоговый телевизионный сигнал с широтно-импульсной модуляцией и лазерный луч выжигает на дорожках вращающегося диска отверстия разной длины, которые при считывании полупроводниковым лазером отображают записанную информацию.

Дальнейший шаг в развитии дисковых ЗУ с побитовой записью связан с использованием носителей, допускающих перезапись информации. Среди материалов, пригодных для создания таких дисков, в последние годы особенно успешные исследования проводятся по магнитным структурам. Эти исследования были начаты более 20 лет назад, однако лишь в последние годы удалось создать, магнитные структуры, пригодные для оптических ЗУ.

Принцип записи основан на термомагнитных эффектах и осуществляется сфокусированным лазерным лучом, нагревающим участок среды до температуры порядка температуры Кюри. При одновременном воздействии лазерного излучения и внешнего магнитного поля происходит перемагничивание доменов локально нагретых участков пленки. Подъем температуры ΔT , необходимый для записи, определяется как

$$\Delta T = T_C + T_0$$

где T_C - температура Кюри;

T_0 - температура окружающей среды.

При считывании информации фиксируется поворот плоскости поляризации лазерного излучения за счет магнитооптических эффектов Фарадея или Керра. Поэтому такие ЗУ часто называют магнитооптическими. В качестве светочувствительных материалов используются различные ферромагнитные пленки с добавками редкоземельных элементов.

По информационным характеристикам магнитооптические ЗУ примерно такие же, как и ЗУ с однократной записью. Чувствительность магнитооптических дисков также невысокая (порядка 10^{-2} Дж/см²), что ограничивает темп записи.

Создание магнитооптических дисков, допускающих перезапись информации, заметно расширяет область применения оптических ЗУ. Прежде всего, они способны составить серьезную конкуренцию магнитным дискам, широко применяемым в качестве носителей внешней памяти в существующих ЭВМ. Обладая большой информационной емкостью, магнитооптические диски открывают также возможность создания нового класса аппаратуры — дисковых магнитофонов для цифровой записи и воспроизведения звуковых программ с высоким качеством звучания. Такие оптические магнитофоны будут более долговечными, чем применяемые в настоящее время на магнитных лентах, которые портятся в процессе эксплуатации из-за наличия контакта между магнитной головкой и лентой. Кроме того, здесь открывается возможность создания единого устройства, которое может использоваться и как магнитофон, и как проигрыватель.

6.3.2 Как устроен CD-ROM

Оптические диски с информацией, записанной побитовым способом, принято называть компакт-дисками (CD). Если информация на диске уже записана изготовителем и может только считываться пользователем, то такой диск называют CD-ROM (compact disc – read only memory).

Все компакт-диски имеют один и тот же физический формат изготовления /12/. Они представляют собой диск (рисунок 6.7) диаметром 120 мм, толщиной в 1,2 мм и центральным отверстием диаметром 15 мм. Центральная область вокруг отверстия шириной 6 мм называется зоной крепления (clamping area). За ней непосредственно следует заголовочная область (lead-in area), содержащая оглавление диска (TOC – Table of Content). Далее расположена область шири-

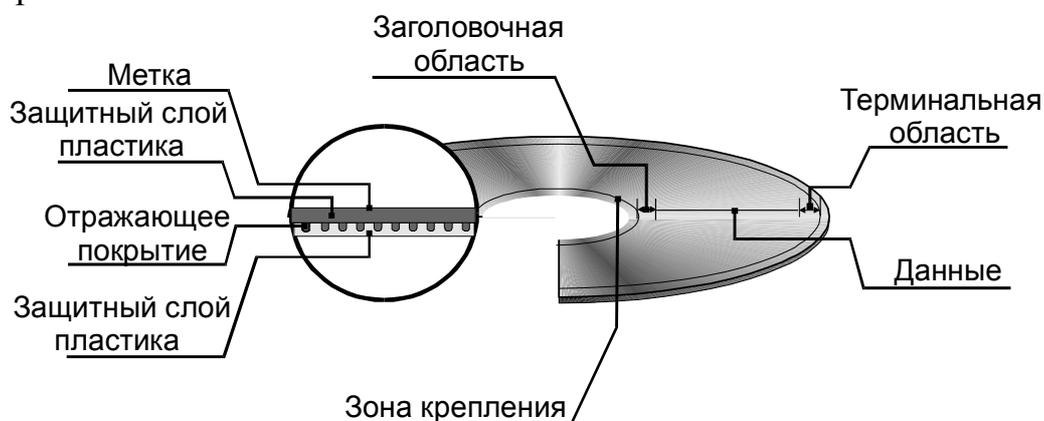


Рисунок 6.7

ной 33 мм, предназначенная для хранения данных и физически представляющая собой единый трек, закрученный в виде спирали. Завершающей является терминальная (lead out) область шириной 1мм. Внешний обод диска шириной 3 мм предназначен для того, чтобы его было удобно брать в руки.

Область хранения данных логически может содержать от 1 до 99 треков, в каждом из которых записывается текстовая, видео- или аудиоинформация. Однако, разнородная информация не может быть смешана на одном и том же треке.

Цифровые данные хранятся на CD-ROM в виде чередующихся между собой по ходу спирали ямок, нанесенных на поверхность полиуглеродного пластика, и ровных областей. Поверх этого пластикового слоя для лучшего отражения напыляется тонкий слой алюминия (или золота), который, в свою очередь, также покрывается защитным слоем пластика. Луч лазера, считывающий информацию, воспринимает ямки как рассеивающие центры, что в цифровой электронике является аналогом состояния "выключено". Когда же лазерный луч сканирует гладкую поверхность, свет отражается строго в обратном направлении, и это представляет собой аналог состояния "включено". Ямки и "гладкие" области компакт-диска не являются в чистом виде нулями и единицами двоичной информации, а представляют собой так называемые "канальные биты" (channel bits). Четырнадцать таких канальных битов определяют один информационный байт. Такая схема обеспечивает избыточность данных, благодаря чему информация остается при каждом считывании одинаковой, даже если на поверхности диска появились царапины. Это и снискало компакт-диску репутацию очень надежного устройства.

6.3.3 Оптические диски с многократной записью

Материалы. Оптические диски с возможностью перезаписи в определенном смысле ничем не отличаются от других видов сменных носителей массовой памяти: они должны представлять собой надежное средство долговременного хранения информации. Чтобы достичь этой цели, изготовители оптических дисков сосредоточили свое внимание на шести проблемах /13/:

- обеспечение высокой ударопрочности;
- применение материалов и технологических процессов, обеспечивающих простое, но при этом точное формование дисков;
- создание дисков с однородной структурой;
- обеспечение минимальности деформации;
- обеспечение высокой устойчивости к внешним воздействующим факторам;
- использование недорогих материалов.

Для изготовления оптических дисков большинство фирм используют один из трех видов материалов: поликарбонат, стекло или полиметилметакрилат (PMMA). Каждый из этих материалов имеет свои достоинства.

Поликарбонат является наиболее распространенным материалом основы-подложки для оптического носителя, поскольку он отличается высокой прочно-

стью, малой удельной плотностью, долговечностью, химической стойкостью и стойкостью к воздействию тепла и влаги. Благодаря физической и химической стабильности структуры поликарбоната он характеризуется пренебрежимо малым ухудшением параметров со временем и практически не проявляет признаков кристаллизации или атомной диффузии.

К числу главных достоинств стекла относится то, что оно практически не поглощает влаги и устойчиво к воздействию тепла. К сожалению, стекло имеет большую удельную плотность, боится ударов, дорого в производстве и исключительно хрупко. Кроме того, в отличие от поликарбонатного носителя, диски со стеклянной основой необходимо изготавливать с воздушной полостью, из-за которой записывающий слой может оказаться подверженным загрязнению.

PMMA – это сополимер метилметакрилата и акрилата, самый прозрачный из всех пластиков, при этом стойкий к царапинам. PMMA имеет относительно низкую температуру плавления и малую вязкость в состоянии текучести, что позволяет достичь необходимой точности при формовании этого компаунда, однако он активно поглощает влагу и поэтому подвержен физическим деформациям. Последнее обстоятельство означает, что PMMA не может считаться идеальным материалом при изготовлении дисков, используемых для долговременного архивного хранения информации.

Запись информации. Существует шесть основных способов записи информации на оптический диск: путем преобразования состава сплава записывающего слоя поверхности, путем образования темных пятен-впадин, или кратеров (pit), путем образования "пузырьков", с использованием слоя органического полимерного красителя, путем изменения фазы вещества, и магнитооптический способ. Изготовители оптических дисков обычно выбирают для производства один или несколько из этих технологических процессов.

Запись путем преобразования состава сплава – это технология однократной записи, которая обеспечивает занесение данных на диск для постоянного хранения путем преобразования состава и структуры записывающего слоя. Например, один из поставщиков оптических дисков использует поликарбонатную основу и записывающий слой, состоящий из трех слоев двухэлементных сплавов: сурьма-селен Sb_2Se_3 , висмут-теллур Bi_2Te_3 и сурьма-селен Sb_2Se_3 . Когда на такой записывающий слой направляется луч лазера высокой мощности, на участке, подвергнутом воздействию, образуется слой одного четырехэлементного сплава сурьма-селен-висмут-теллур $SbSeBiTe$. Вследствие этого отражательная способность обработанного участка увеличивается.

Для чтения данных используется слабый лазерный луч, который сканирует записывающий слой. Этот луч отражается на фотодетектор, преобразующий колебания яркости света в двоичные данные. Участки диска, обработанные в процессе записи информации, имеют на 300 - 400% большую отражательную способность, чем необработанные участки, что и позволяет накопителю считывать записанные данные. Важнейшим преимуществом подобного метода записи является то, что он не требует наличия воздушного кармана между записываю-

щим слоем и основой, а это снижает риск загрязнения, повышает надежность хранения данных и долговечность оптического диска.

Запись путем образования темных впадин – это разрушающий способ однократной записи, который чаще всего используется для оптических дисков со стеклянной основой. В процессе образования впадин на записывающий слой диска воздействует лазерный луч, вызывающий повышение температуры обрабатываемого участка. Тепловой удар, в свою очередь, вызывает испарение материала, так что образуется впадина, или кратер (pit), т.е. физически меняется поверхность носителя. Записывающий слой – это, как правило, тонкая пленка материала типа теллура, нанесенная с подслоем воздуха. При выполнении операций чтения оптический накопитель направляет на диск лазерный луч, причем этот луч отражается по-разному от впадин поверхности. Одним из недостатков записи путем формирования впадин является то, что воздушный «сэндвич» может привести к загрязнению записывающего слоя и к изменению данных, хранящихся на диске.

Запись путем образования «пузырьков» – также способ однократной записи, меняющий поверхность диска. Однако здесь вместо образования впадин под воздействием лазерного луча из-за локального нагрева создаются «пузырьки-шишки»; эти пузырьки, как и впадины в предыдущем способе, при чтении отражают лазерный луч не так, как необработанные участки. Как правило, для образования пузырьков используются два слоя: слой материала, поглощающего тепло, поверх подслоя диэлектрика. Верхний слой накапливает тепло при воздействии лазерного луча, и когда температура повысится, в нижнем диэлектрическом слое начнется испарение. Из-за локального испарения в слое, поглощающем тепло, образуются пузырьки. Хотя такой процесс предназначается для однократной записи данных на диск, он позволяет стереть данные путем уничтожения пузырьков воздействием тепла или давления.

Запись с использованием слоя органического полимерного красителя предназначается преимущественно для оптических дисков с возможностью перезаписи, но может также применяться для дисков с однократной записью. Органический краситель играет роль записывающего слоя. При воздействии на этот записывающий слой лазерного луча создаются пузырьки, которые отражают свет иначе, чем гладкие, необработанные участки диска. Чтобы записать на подобный диск новые данные, краситель подвергается повторному нагреву и преобразованию. Теоретически технология записи с использованием полимерного органического красителя дешевле, чем другие способы однократной или многократной записи, однако для чтения и записи она требует двух лазеров, работающих на различных длинах волн; поэтому накопители с полимерным красителем дороги в изготовлении и сложны в техническом обслуживании. Кроме того, носитель с полимерным красителем может выдержать всего 10 тысяч циклов записи; в то же время магнитооптический носитель компании Sony способен выдержать более 10 миллионов циклов записи.

Технология *записи путем изменения фазы вещества* применяется в оптических дисках как с однократной, так и многократной записью. Когда на запи-

сыывающий слой диска воздействует лазерный луч, на обратном участке происходит переход материала из неструктурированного аморфного состояния в структурированное кристаллическое состояние, аналогичный переходу воды в лед при замерзании. В качестве записывающего слоя обычно используется тонкая пленка аморфного материала типа окисла теллура. Поскольку температурный диапазон, при котором обеспечивается переход, узок, а измененное состояние зачастую нестабильно, подобный метод требует использования дорогостоящего мощного лазера, способного четко контролировать мощность своего луча. Кроме того, как и в случае записи с полимерным красителем, технология записи путем изменения фазы вещества ограничивает допустимое число перезаписей информации на диске.

На сегодняшний день в области НОД с возможностью перезаписи лидирует *магнитооптическая (МО) технология*, представляющая собой сочетание магнитной и лазерной технологии для чтения, записи, стирания и перезаписи данных. Магнитооптическая запись очень похожа на традиционный способ магнитной записи. Цифровая информация записывается на магнитный слой диска как последовательность битовых элементов с различными направлениями магнитного потока. Каждый битовый элемент имеет полярность - северный полюс магнита может быть направлен вниз (представляя цифру "0") или вверх (представляя цифру "1"). Изменения направления магнитного потока битового элемента – это оптический процесс, требующий участия лазера и смещающего магнита (рисунок 6.8).

Магниты, применяемые обычно в МО – накопителях, являются не настолько мощными, чтобы изменить ориентацию битового элемента, тем более что применение сильных магнитов влияло бы на соседние битовые элементы диска. Для изменения полярности конкретного элемента без разрушения других данных в МО–накопителе лазер нагревает поверхность над конкретным битовым элементом (в течение приблизительно 800 нс) до точки Кюри магнитной пленки (150 градусов Цельсия). При этой температуре магнитные свойства пленки изменяются таким образом, что магнит накопителя может переключить направление битового элемента.

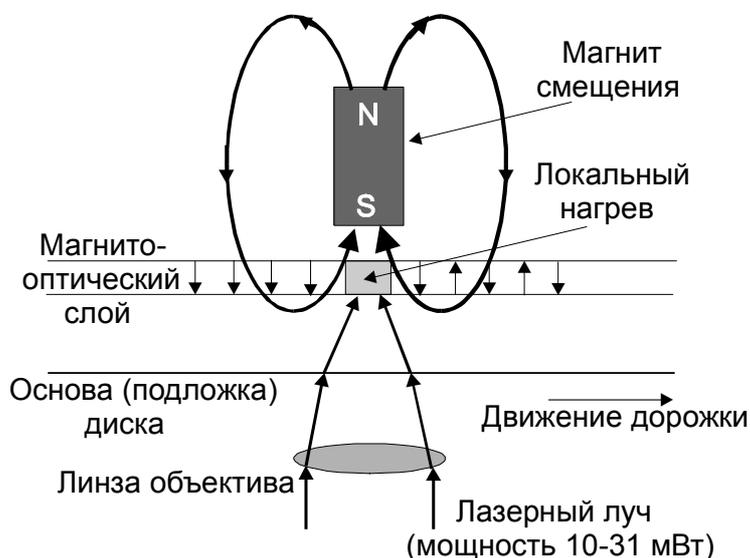


Рисунок 6.8. Принцип магнитооптической записи.

Можно считать, что все битовые элементы чистого (без записи) МО–диска имеют ориентацию северным полюсом вниз. Когда диск вращается в накопителе, лазерный луч включается и выключается в импульсном режиме. Если лазер включится, когда магнит пытается изменить полярность, чтобы северный полюс был направлен вверх, записывается цифра "1". Цифра "0" записывается в случае, когда магнит стремится перевести северный полюс вниз и на соответствующий битовый элемент воздействует луч лазера. Если лазер выключен, информация остается неизменной.

Для изменения информации, записанной на диск, накопитель должен выполнить две операции. Первая заключается в стирании существующих данных. Эта процедура выполняется путем реверсирования полярности магнита и включения лазера в течение одного оборота диска. Все перезаписываемые битовые элементы переходят в состояние "0". При втором проходе, т.е. при перезаписи, полярность магнита меняется избирательно. Снова на один оборот диска включается лазер, и в соответствующих битовых элементах записывается "1".

Считывание информации с МО–диска осуществляется с использованием эффекта Керра (рисунок 2): лазерный луч попадает на битовый элемент и в зависимости от поляризации этого элемента отраженный луч вращается либо по часовой стрелке (если записана цифра "1"), либо против часовой стрелки (если записана цифра "0").

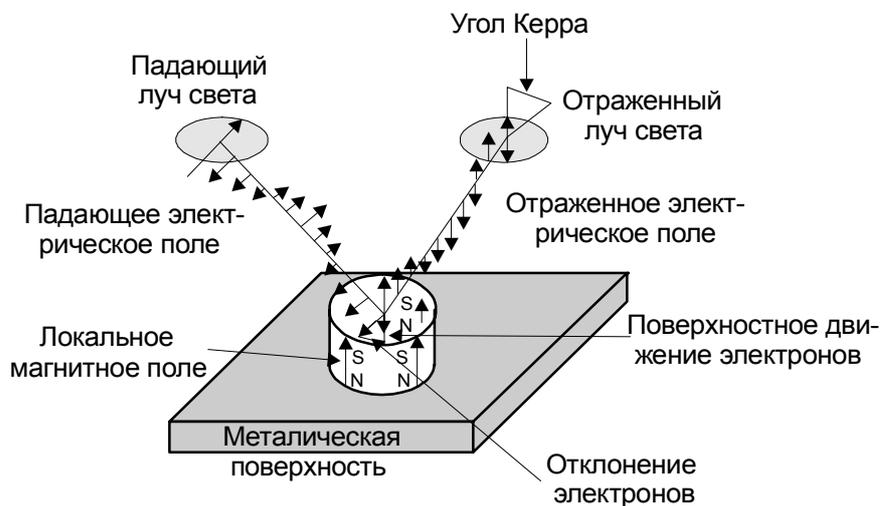


Рисунок 6.9. Использование эффекта Керра для считывания данных, записанных магнитооптическим способом.

Заканчивая раздел об оптических дисках с побитовой записью информации, следует отметить, что, несмотря на их широкое внедрение в различные системы вычислительной техники и информатики, дальнейшее развитие техники таких дисков имеет довольно ограниченные возможности. Здесь, прежде

всего, имеется в виду то, что увеличение плотности записи связано с уменьшением надежности хранения информации, даже при очень высоких требованиях к однородности носителя и к системам управления.

Второе, что следует отметить, это отсутствие многоканальности, что ограничивает темп считывания (и записи) даже при большой скорости вращения диска величинами порядка единиц (в лучшем случае десятка) мегабит в секунду. Правда, многоканальность можно ввести, используя решетку лазеров, однако в этом случае возникают технические сложности, связанные главным образом с реализацией фокусирующей оптики и систем слежения. Поэтому в настоящее время высокие скорости записи и считывания реализуются путем применения нескольких дисков, управляемых по определенному закону.

7 Литература, рекомендуемая для изучения темы

1. Исихара С. Оптические компьютеры. - М.: Наука, 1992 г.
2. Гонда С., Сэко Д. Оптоэлектроника в вопросах и ответах. - Л.: Энергоатомиздат., 1989 г.
3. Морозов В.Н. Оптоэлектронные матричные процессоры. - М.: Радио и связь, 1986 г.
4. Основы оптоэлектроники: Пер. с япон. М.: Мир, 1988 г.
5. Свечников Г.С. Интегральная оптика. – Киев: Наукова думка, 1988 г.
6. Микаэлян А.Л. Голография. - М.: Знание, 1968 г.
7. Михайлов В.И. и др. Запоминающие устройства на оптических дисках.- М.: Радио и связь, 1991г.
8. Мнейн М.Г. Физика машинной памяти. – М.: Высш.шк., 1990 г.
9. Волоконно-оптические датчики:/ Под ред. Т. Оокоси, Пер. с япон. - Л.: Энергоатомиздат., 1991 г.
10. Оокоси Т. Оптоэлектроника и оптическая связь. – М.: Мир, 1988 г.
11. Семенов А.С. и др. Интегральная оптика – М.: Радио и связь, 1990 г.
12. Аблеков В.К. и др. Оптическая и оптоэлектронная обработка информации. – М.: Машиностроение, 1976 г.
13. Пространственные модуляторы света /А.А. Васильев, Д.Касасент и др. – М.: Радио и связь, 1987 г.
14. Сиаккоу М. Физические основы записи информации.- М.:Связь, 1980 г.
15. Суэмацу Я. и др. Основы оптоэлектроники. – М: Мир, 1988 г.
16. Гауэр Дж. Оптические системы связи. – М.: Радио и связь, 1988 г.
17. Гончаренко А.М., Редько В.П. Введение в интегральную оптику. – Минск: Наука и техника, 1975 г.
18. Фотоника: Пер. с англ. и фр./ Под ред. М.Балкански, П. Лалемана.- М.: Мир, 1978 г.
19. Кохонен Т. Ассоциативная память. - М.: Мир, 1980 г.
20. Гудмен Дж. Введение в фурье-оптику. - М.: Мир, 1970 г.
21. Папулис А. Теория систем и преобразований в оптике. – М.: Мир, 1971 г.
22. Престон К. Когерентные оптические вычислительные машины. – М.: Мир, 1974 г.

23. Юу Ф.Т.С. Введение в теорию дифракции, обработку информации и голографию. – М.: Сов. радио, 1979 г.
24. Когельник Г. Введение в интегральную оптику // Успехи физ. наук. – 1977. –Т.121, вып.4. – с. 659-726.
25. Эйбрэхэм А., и др. Оптический компьютер// В мире науки: М.: Мир, 1983 г. №4.

Список использованных источников

1. Свириденко С. С. Современные информационные технологии. - М.: Радио и связь, 1989. - 304 с.
2. Носов Ю. Р. Достижения и перспективы оптоэлектроники // Оптоэлектронная обработка информации. - М.: Знание, 1986. С. 3-33.
3. Носов Ю. Р. Оптоэлектроника. - М.: Радио и связь, 1989. – 360 с.
4. Акаев А. А., Майоров С. А. Оптические методы обработки информации. - М.: Высшая школа, 1988. – 237 с.
5. Василевский А. М. и др. Оптическая электроника. - Л.: Энергоатомиздат, 1990. – 167 с.
6. Тарасов Л. В. Лазеры и их применение. - М.: Радио и связь, 1983. – 152 с.
7. Кольер Р. и др. Оптическая голография. - М.: Мир, 1973. – 686 с.
8. Оптическая голография. Практические применения. Под ред. Гинзбург В.М., Степанова Б.М., М.: Сов. радио, 1973. – 304 с.
9. Островский Ю. И. и др. Голографическая интерферометрия. - М.: Наука, 1977. – 339 с.
10. Островский Ю. И. Голография и ее применение. - Л.: Наука, 1973. – 179 с.
11. Микаэлян А.Л. Оптические методы и информатике. - М.: Наука, 1990. – 232с.
12. Костяков С. CD – ROM как устройство хранения информации // CompUnity. – 1995. № 1. - С. 54 – 61.
13. Уилсон Р. Оптические диски с многократной записью // Мир ПК. – 1992. № 1. - С. 13 –19.