

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ
Федеральное государственное автономное образовательное учреждение
высшего профессионального образования
САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
АЭРОКОСМИЧЕСКОГО ПРИБОРОСТРОЕНИЯ

М. Б. Рыжиков

**ФОРМИРОВАНИЕ И ОБРАБОТКА
ИЗОБРАЖЕНИЙ
В ЛАЗЕРНЫХ СИСТЕМАХ ВИДЕНИЯ**

Учебное пособие



Санкт-Петербург
2013

УДК 004.932
ББК 32.86-5
Р93

Рецензенты:

доктор технических наук, профессор *А. Ф. Крячко*;
доктор технических наук, профессор *Ю. В. Юханов*

Утверждено

редакционно-издательским советом университета
в качестве учебного пособия

Рыжиков М. Б.

Р93 Формирование и обработка изображений в лазерных системах
видения: учеб. пособие. – СПб.: ГУАП, 2013. – 210 с.: ил.
ISBN 978-5-8088-0878-2

В пособии изложены основы теории получения изображений объектов, находящихся на больших расстояниях или в условиях недостаточной освещенности за счет применения систем с использованием лазерного подсвета – лазерных систем видения. Рассмотрены вопросы технической реализации оптического и электронного тракта для систем видения, предназначенных для приема и регистрации изображений, представлены основные характеристики систем видения, предназначенные для оценки предельной дальности их действия. Дан математический аппарат линейных изображающих систем, используемый для описания процессов передачи, приема и обработки изображения в лазерных системах видения. Особое внимание уделяется вопросам цифровой обработки изображения, получаемого с фотоприемной матрицы в электронном тракте, а также решению задач автоматического распознавания образов.

Учебное пособие предназначено для студентов технических вузов, обучающихся по направлению «Оптотехника», а также может быть полезным разработчикам оптико-электронных систем различного назначения.

УДК 004.932
ББК 32.86-5

ISBN 978-5-8088-0878-2

© Санкт-Петербургский государственный
университет аэрокосмического
приборостроения (ГУАП), 2013
© М. Б. Рыжиков, 2013

СПИСОК СОКРАЩЕНИЙ

- АЦП – аналого-цифровой преобразователь
- ИК – инфракрасный
- ЛСВ – лазерная система видения
- МДП – металл – диэлектрик – полупроводник
- МКП – микроканальная пластина
- МПФ – модуляционная передаточная функция
- МРК – минимальный разрешаемый контраст
- НЧ – низкочастотный
- ОПФ – оптическая передаточная функция
- ПЗС – прибор с зарядовой связью
- ПНВ – прибор ночного видения
- ТВ – телевизионный
- УФ – ультрафиолетовый
- ФПЗС – фотоприемная матрица на основе приборов с зарядовой связью
- ФПУ – фотоприемное устройство
- ФР – функция рассеяния
- ФЧХ – фазо-частотная характеристика
- ЭОП – электронно-оптический преобразователь

ПРЕДИСЛОВИЕ

Данное учебное пособие предназначено для изучения курса «Лазерные системы видения» студентами специальности 200202 «Лазерные системы в ракетной технике и космонавтике» по направлению 200200 «Оптотехника».

Цель данного учебного пособия – ознакомить слушателей курса с физическими принципами, заложенными в основу работы лазерных систем видения (ЛСВ), дать представление об областях их применения, о технической реализации, о математических моделях, используемых при описании процессов получения, приема и обработки изображений и расчете основных технических характеристик системы, которые определяются данными процессами.

Учебное пособие базируется на классических подходах к анализу систем видения, т. е. на методах описания линейных систем, на теории передачи и обработки оптических сигналов в областях пространственных и временных частот. Рассматривается не только процесс формирования изображения в оптико-электронном тракте системы, но и процесс переноса отраженного от объекта наблюдения сигнала в различных средах.

Автором предлагается следующая последовательность изложения приведенного в учебном пособии материала.

Раздел 1 включает в себя описание тех физических процессов и характеризующих их основных параметров, которые необходимы для понимания принципов работы и различий в технической реализации систем видения. Рассматриваются оптические, электрические и обобщенные системные параметры систем, предназначенных для получения изображений удаленных объектов, а также причины, приводящие к ухудшению восприятия изображения в реальных средах распространения. Приводятся сведения о классификации, об общей структуре, об особенностях реализации систем видения, использующих лазерный подсвет, а также об областях их применения.

Раздел 2 знакомит читателя с математическим аппаратом, применяемым для описания процесса передачи и формирования изображения в области пространственных частот различными блоками ЛСВ. Вводятся понятия оптической передаточной функции (ОПФ), модуляционной передаточной функции (МПФ) и фазо-частотной характеристики (ФЧХ) в области пространственных частот. Анализируется связь, существующая между распределением яркости на объекте и его изображением посредством расчета амплитудной и фазовой характеристик ОПФ. Приводятся основные

соотношения, позволяющие рассчитать МПФ всех основных звеньев ЛСВ, включая среду распространения. Дается описание работы фотоприемных матриц на основе приборов с зарядовой связью, анализ источников шума в приемном тракте. Представлен пример расчета предельной дальности обнаружения и распознавания ЛСВ в подводной среде, которая может служить для поиска затонувших объектов или для контроля состояния подводных сооружений, таких как нефтяные или газовые трубопроводы.

Раздел 3 посвящен алгоритмам обработки изображений, формирующихся первично в фотоприемных матрицах и реализуемых в цифровых процессорах, позволяющих улучшить их визуализацию. Рассматриваются способы повышения контрастности, резкости, выделения границ объектов на изображениях. Приводятся методики обработки как плоскости изображения, так и в области пространственных частот; результаты, демонстрирующие возможности цифровой обработки изображений.

Раздел 4 посвящен описанию алгоритмов автоматического распознавания образов на изображениях, показывает перспективы применения указанных алгоритмов в лазерных системах видения. Рассматриваются основы теории нейронных сетей, которые можно использовать для идентификации объектов изображения. Возможности практического моделирования показываются на примере актуальной в системах обработки изображений проблеме автоматического распознавания.

Учебное пособие не затрагивает всего круга вопросов, необходимых для проектирования лазерных систем видения, поскольку не содержит подробного расчета энергетических характеристик оптико-электронных приборов, входящих в состав лазерных систем видения. Особое внимание уделяется техническим характеристикам и объяснению физических процессов, которые необходимы для решения главной задачи применения лазерных систем видения – получения качественного изображения объекта, находящегося на большом расстоянии в рассеивающей среде или в условиях малой освещенности, когда возможности пассивных оптико-электронных систем видения ограничены. Затрагиваются вопросы распознавания и идентификации объектов по изображению.

1. НАЗНАЧЕНИЕ, ФИЗИЧЕСКИЕ ОСНОВЫ РАБОТЫ И ТЕХНИЧЕСКИЕ ХАРАКТЕРИСТИКИ ЛАЗЕРНЫХ СИСТЕМ ВИДЕНИЯ

1.1. Принципы действия и устройство оптической части оптико-электронных систем видения

Под *видением* в широком смысле понимают восприятие наблюдателем удаленных объектов. Если речь идет о возможности непосредственного зрительного восприятия объектов, то пользуются понятием «*видимость*». Для технических средств и систем, предназначенных для получения изображения удаленных объектов, используется термин «видение».

Под *удаленным объектом* следует понимать не объект, находящийся на большом расстоянии от наблюдателя, а объект таких размеров, формы и контраста по отношению к фону, которые не позволяют увидеть его на необходимом для обнаружения и идентификации расстоянии. Так, для человека с пониженной остротой зрения шрифт малого размера можно считать удаленным объектом.

Прежде чем перейти к непосредственному изучению лазерных систем видения, необходимо ознакомиться с известными системами видения, историей их развития и физическими законами, которые были использованы в процессе их создания и совершенствования.

Для понимания принципов действия оптико-электронных систем видения вначале рассмотрим техническую реализацию и процессы в оптической части, а в следующем подразделе в электронной части.

Следует сразу отметить, что при использовании технических средств для реализации процесса видения редко можно обойтись без участия оператора, осуществляющего анализ и принятие решений. Таким образом, конечным звеном многих систем видения или передачи изображений является органическая система зрительного восприятия человека [1]. Данная органическая система может быть представлена в виде оптико-электронного прибора, поскольку глаз человека является оптической системой, проецирующей изображение на чувствительный приемник – сетчатку, а дальнейший процесс восприятия изображения требует передачи последовательности электрических импульсов в центральную нервную систему, т. е. может быть смоделирован как электронный тракт для приема сигнала. Головной мозг оператора включается в процесс обработки как последнее звено цепи, анализируя полученное изображение. Более того при проектировании технических систем, связанных с

формированием, передачей, регистрацией изображений, как будет показано далее, используют для расчета световые величины, которые вводятся на основе тех данных, которые были получены в процессе исследования возможностей и особенностей восприятия человеком света.

Именно необходимость повышения возможности наблюдения сильно удаленных объектов в условиях необходимости или в условиях получения зрительной информации в рассеивающих средах: в воде или в атмосфере при наличии тумана, дождя или снега, привела к созданию первых технических систем видения [2].

Самые первые технические системы для улучшения эффективности процесса видения были оптическими. Никто не знает имена тех людей, которые изобрели первую пару очков для коррекции остроты зрения и подзорную трубу для увеличения дальности видимости наблюдаемых объектов. Известно, что в 1266 году Роджер Бэкон увеличил буквы в книге, прикладывая к ним увеличительное стекло, а портрет кардинала Югона, датированный 1352 годом, изображает человека в очках, которые состояли из двух обрамленных линз и двух дужек, скрепленных вместе и помещенных возле глаз. В 1629 году английский король Чарльз I утвердил устав Гильдии мастеров по производству очков, а в 1784 году Бенджамин Франклин изобрел очки с двойным фокусом. В 1609 году Галилео Галилей сконструировал подзорную трубу для наблюдения космических объектов, создав первый телескоп. Дальнейшие конструкции телескопов, сделанных Галилеем, позволили ему открыть горы на Луне, спутники у Юпитера и даже пятна на Солнце. Именно создание телескопа Галилеем и его предназначение позволяют считать его первой технической системой видения [3].

Следует отметить, что как тогда, так и сейчас, любая система видения содержит в своем составе оптическую приемную систему, представляющую собой конструкцию из расположенных определенным образом линз. Линзы, используемые для приема изображения, бывают двух типов: собирающие и рассеивающие. Собирающая линза в середине толще, чем у краев, рассеивающая линза, наоборот, в средней части тоньше (рис. 1).

Если на линзу направить пучок лучей, параллельных главной оптической оси, то после прохождения через линзу лучи (или их продолжения) соберутся в одной точке F , которая называется главным фокусом линзы. У тонкой линзы имеются два главных фокуса, расположенных симметрично на главной оптической оси относительно линзы. У собирающих линз фокусы действительные,

у рассеивающих – мнимые. Пучки лучей, параллельных одной из побочных оптических осей, после прохождения через линзу также фокусируются в точку F' , которая расположена при пересечении побочной оси с **фокальной плоскостью Φ** , т. е. плоскостью, перпендикулярной главной оптической оси и проходящей через главный фокус (рис. 2). Расстояние между оптическим центром линзы O и главным фокусом F называется **фокусным расстоянием**. Оно обозначается той же буквой F .

Положение изображения и его характер можно определить с помощью геометрических построений. Для этого используют свойства некоторых стандартных лучей, ход которых известен. Стандартные лучи – это лучи, проходящие через оптический центр или один из фокусов линзы, а также лучи, параллельные главной или одной из побочных оптических осей. Примеры таких построений для собирающей и рассеивающей линз представлены на рис. 3 и 4, соответственно.

Положение изображения и его характер (действительное или мнимое) можно также рассчитать с помощью **формулы тонкой линзы**. Если расстояние от предмета до линзы обозначить через d ,

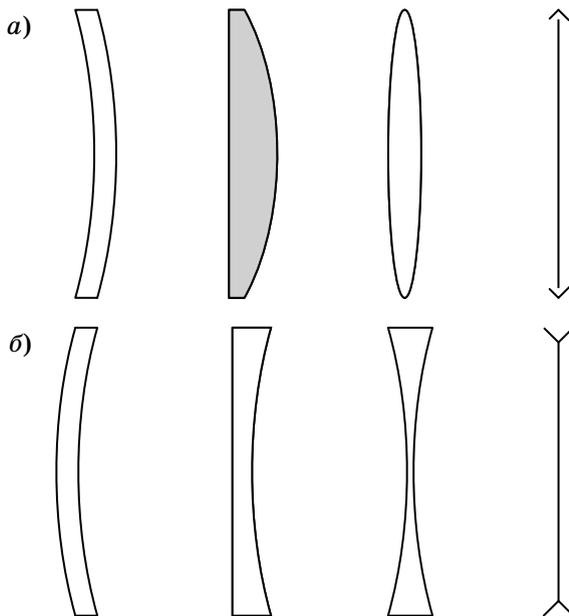


Рис. 1. Собирающие (а) и рассеивающие (б) линзы

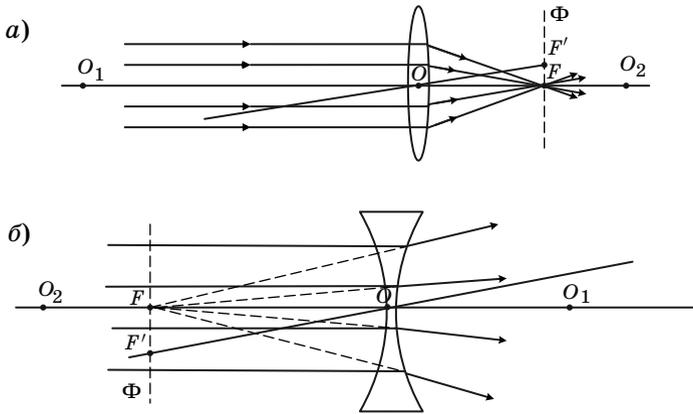


Рис. 2. Преломление параллельного пучка лучей в собирающей (а) и рассеивающей (б) линзах. Точки O_1 и O_2 – центры сферических поверхностей; O_1O_2 – главная оптическая ось; O – оптический центр; F – главный фокус; F' – побочный фокус; OF' – побочная оптическая ось; Φ – фокальная плоскость

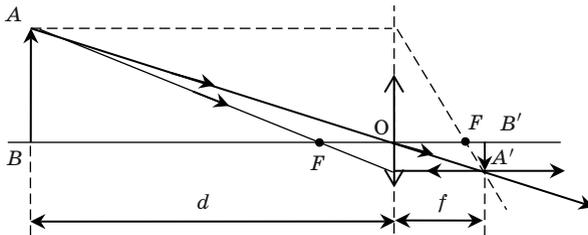


Рис. 3. Построение изображения в собирающей линзе

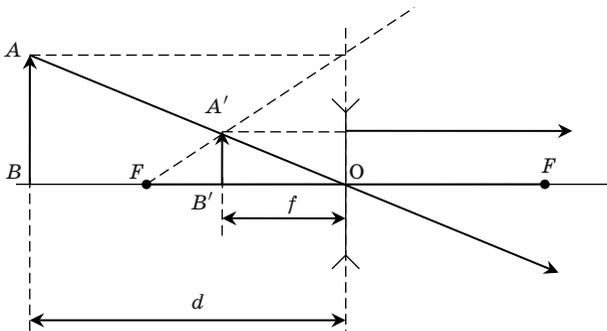


Рис. 4. Построение изображения в рассеивающей линзе

а расстояние от линзы до изображения через f , то формулу тонкой линзы можно записать в виде [3]:

$$\frac{1}{d} + \frac{1}{f} = \frac{1}{F} = D.$$

Величину D , обратную фокусному расстоянию, называют **оптической силой** линзы. Единицей измерения оптической силы является **диоптрия** (дптр). Диоптрия – оптическая сила линзы с фокусным расстоянием 1 м, таким образом, 1 дптр = 1 м⁻¹. Оптическая сила D линзы зависит как от радиусов кривизны R_1 и R_2 ее сферических поверхностей, так и от показателя преломления n материала, из которого изготовлена линза [3]:

$$D = (n - 1) \left(\frac{1}{R_1} + \frac{1}{R_2} \right).$$

Радиус кривизны выпуклой поверхности считается положительным, вогнутой – отрицательным. Эта формула используется при изготовлении линз с заданной оптической силой. В зависимости от положения предмета по отношению к линзе изменяются линейные размеры изображения. Линейным увеличением линзы Γ называют отношение линейных размеров изображения h' и предмета h . Величине h' принято приписывать знаки плюс или минус в зависимости от того, является изображение прямым или перевернутым. Величина h всегда считается положительной.

Поэтому для прямых изображений $\Gamma > 0$, а для перевернутых $\Gamma < 0$. Для расчета линейного увеличения тонкой линзы используют формулу:

$$\Gamma = \frac{h'}{h} = -\frac{f}{d}.$$

Во многих оптических приборах свет последовательно проходит через две или несколько линз. Изображение предмета, даваемое первой линзой, служит предметом (действительным или мнимым) для второй линзы, которая строит второе изображение предмета.

Это второе изображение также может быть действительным или мнимым. Расчет оптической системы из двух тонких линз сводится к двукратному применению формулы линзы, при этом расстояние d_2 от первого изображения до второй линзы следует положить равным величине $l - f_1$, где l – расстояние между линзами. Рассчитанная по формуле линзы величина f_2 определяет положение вто-

рого изображения и его характер ($f_2 > 0$ – действительное изображение, $f_2 < 0$ – мнимое). Общее линейное увеличение Γ системы из двух линз равно произведению линейных увеличений обеих линз: $\Gamma = \Gamma_1 \cdot \Gamma_2$. Если предмет или его изображение находятся в бесконечности, то линейное увеличение утрачивает смысл и пользуются **угловым увеличением** γ . Если удаленный предмет виден невооруженным глазом под углом ψ , а при наблюдении через телескоп под углом ζ , то угловым увеличением называют отношение

$$\gamma = \frac{\zeta}{\psi}.$$

Ход лучей в оптическом приемном тракте простейшей системы видения – телескопа Галилея представлен на рис. 5. Оптическая часть телескопа состоит из двояковыпуклого объектива (длиннофокусная собирающая линза) и двояковогнутого стекла – окуляра (короткофокусная рассеивающая линза). Назначение окуляра – превратить пучок сходящихся лучей в пучок лучей параллельных, которые затем входят в зрачок глаза наблюдателя.

Увеличение такой оптической системы может быть также вычислено как отношение фокусного расстояния объектива к фокусному расстоянию окуляра

$$\Gamma = \frac{F_1}{F_2}.$$

При работе фотоприемных систем чаще используют обобщенное увеличение, которое определяется как отношение линейной величины изображения к угловому размеру объекта.

В реальных оптических системах при формировании изображения объекта увеличение для различных его точек не является идентичным, при этом говорят, что система обладает дисторсией. Дис-

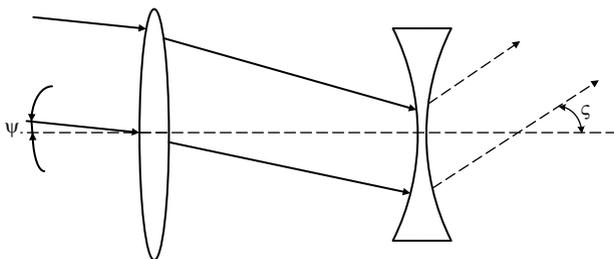


Рис. 5. Ход лучей в оптической системе телескопа Галилея

торсия проявляется в том, что на полученном изображении исходные прямые линии объекта искривляются. Вместо изображения квадрата можно получить фигуру как с вогнутыми внутрь квадрата линиями, так и с выпуклыми. Такие искажения называются геометрическими, причем чаще соответственно говорят о «подушкообразных» или «бочкообразных» искажениях оптической системы.

Допустимая относительная дисторсия, при которой искаженное изображение объекта воспринимается глазом как правильное, составляет 5–10%. Разумеется, при работе с измерительными оптическими приборами на практике часто важно иметь большую точность.

Как уже говорилось, глаз оператора является составной частью практически всех систем видения и представляет собой сложную оптическую систему [4]. Схематическое устройство глаза представлено на рис. 6.

Глаз имеет почти шарообразную форму и диаметр около 2,5 см. Снаружи он покрыт защитной оболочкой 1 белого цвета – склерой. Передняя прозрачная часть 2 склеры называется роговицей. На некотором расстоянии от нее расположена радужная оболочка 3, окрашенная пигментом. Отверстие в радужной оболочке представляет собой зрачок. В зависимости от интенсивности падающего света зрачок рефлекторно изменяет свой диаметр приблизительно от 2 до 8 мм, изменяя величину светового потока в зависимости от освещенности. Между роговицей и радужной оболочкой находится прозрачная жидкость. За зрачком находится хрусталик 4 – эластичное

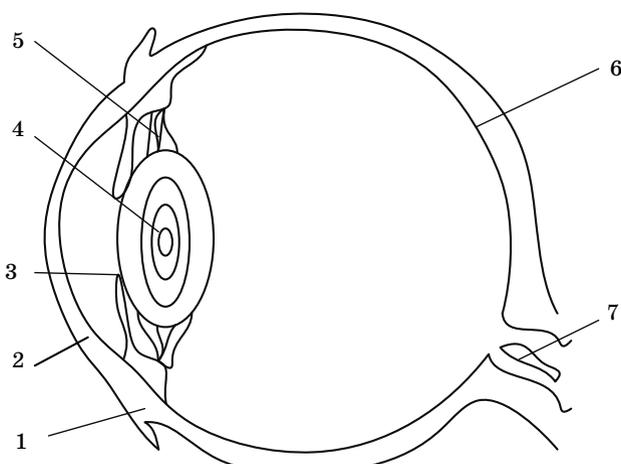


Рис. 6. Схематическое изображение глаза человека

линзоподобное тело. Особая мышца 5 может изменять в некоторых пределах форму хрусталика, изменяя тем самым его оптическую силу. Остальная часть глаза заполнена стекловидным телом. Задняя часть глаза – глазное дно, оно покрыто сетчатой оболочкой 6, представляющей собой сложное разветвление зрительного нерва 7 (слепое пятно) с нервными окончаниями – палочками и колбочками, которые являются светочувствительными элементами.

Лучи света от предмета, преломляясь на границе воздух–роговица, проходят далее через хрусталик (линзу с изменяющейся оптической силой) и создают изображение на сетчатке.

Роговица, прозрачная жидкость, хрусталик и стекловидное тело образуют оптическую систему, оптический центр которой расположен на расстоянии около 5 мм от роговицы. При расслабленной глазной мышце оптическая сила глаза приблизительно равна 59 дптр, при максимальном напряжении мышцы – 70 дптр.

Основная особенность глаза как оптического инструмента состоит в способности рефлекторно изменять оптическую силу глазной оптики в зависимости от положения предмета. Такое приспособление глаза к изменению положения наблюдаемого предмета называется *аккомодацией*.

Для невооруженного глаза наименьший угол зрения приблизительно равен 1'. Этот угол определяется мозаичным строением сетчатки, а также волновыми свойствами света. При анализе работы приборов для визуальных наблюдений удобнее всего полагать, что глаз наблюдателя аккомодирован на бесконечность. Это означает, что лучи от каждой точки предмета, пройдя через прибор, попадают в глаз в виде параллельного пучка.

1.2. Принципы действия и устройство электронной части оптико-электронных систем видения

Следующим значительным этапом развития систем видения следует считать добавление к оптической системе, которая решала задачу улучшения дальности видения, технических средств регистрации изображения с возможностью его последующего воспроизведения. Вначале изображение регистрировалось на фотопленке за счет химических процессов, реализующихся под воздействием света, а затем для этого стали использовать матрицы фоточувствительных элементов. Таким образом, к оптической части систем видения добавился электронный тракт, основное назначение которого зарегистрировать и произвести обработку изображения.

Современный **фотоаппарат** представляет собой электронно-оптическую систему видения, работающую эффективно в условиях хорошей освещенности. Рассмотрим подробнее его устройство и основные принципы его функционирования.

Конструктивно фотоаппарат представляет собой светонепроницаемую камеру, которая содержит объектив, затвор, фотоприемную матрицу, процессор обработки изображения, перезаписываемую карту памяти для хранения информации и дисплея, на котором визуализируется полученный в процессе съемки результат. Процессор обработки служит для автоматического определения параметров съемки и для последующей обработки изображения, например, фильтрации шумов. Изображение фотографируемых предметов создается на приёмной матрице системой линз, которая называется **объективом**.

В составе любого объектива фотоаппарата присутствуют тонкие собирающие линзы. Тонкие линзы обладают рядом недостатков, не позволяющих получать высококачественные изображения. Искажения, возникающие при формировании изображения, называются **абберациями**. Главные из них – геометрическая и хроматическая **абберации**. **Геометрическая абберация** проявляется в том, что лучи, далекие от оптической оси, пересекают ее не в фокусе, поэтому формула тонкой линзы справедлива только для лучей, близких к оптической оси. Таким образом, изображение удаленного точечного источника, расположенного не на оптической оси оказывается размытым.

Хроматическая абберация возникает вследствие того, что показатель преломления материала линзы зависит от длины волны света λ . Это свойство прозрачных сред называется **дисперсией**. Фокусное расстояние линзы оказывается различным для света с разными длинами волн, что приводит к размытию изображения при использовании немонахроматического света.

В современных оптических приборах применяются не только тонкие линзы, а сложные многолинзовые системы, в которых удается приближенно устранить абберации.

Особенностью работы фотоаппарата является то, что на плоской приемной матрице должны получаться достаточно резкими изображения предметов, находящихся на разных расстояниях. Реально в плоскости матрицы получаются резкими только изображения предметов, находящихся на определенном расстоянии (точка А на рис. 7). Наведение на резкость достигается перемещением объектива относительно плоскости матрицы. Изображения точек, не лежащих

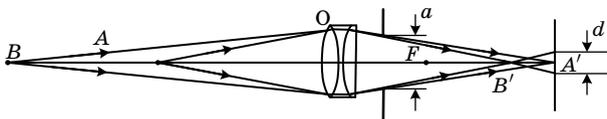


Рис. 7. Ход лучей в оптической системе фотоаппарата

в плоскости резкого наведения (точка B на рис. 7), получают размытыми в виде **кружков рассеяния**. Размер d этих кружков может быть уменьшен путем **диафрагмирования** объектива, т. е. уменьшения **относительного отверстия** a / F (рис. 7). Это приводит к увеличению **глубины резкости**, т. е. к увеличению зоны пространства вокруг точки A , которая слабо размывается, т. е. имеет допустимый диаметр кружка рассеяния.

Для достижения данного эффекта в оптическую систему фотоаппарата включают пластинку с отверстием, диаметр которой можно изменять. Она называется **диафрагмой**. Уменьшение отверстия называют «закрытием», а увеличение – «открытием» диафрагмы. Измеряют степень диафрагмирования безразмерным **числом диафрагмы**, равным отношению фокусного расстояния к диаметру отверстия F / a . Стандартный ряд значений чисел диафрагмы, используемых в фотоаппаратуре: 1,4; 2; 2,8; 4; 5,6; 8; 11,16; 32.

Электрические процессы фиксации изображения матрицей требуют определённого количества световой энергии, приходящейся на единицу площади, т. е. **освещённости** матрицы. Чем меньше допустимое количество этой энергии, тем выше так называемая **чувствительность матрицы**. При заданной чувствительности на поверхности матрицы необходимо накопить такую энергию, которая равна произведению освещённости на время действия, называемое **выдержкой**.

Для дальнейшего понимания энергетических процессов, происходящих как в фотоаппарате, так и в любой технической системе, использующейся для передачи или восприятия изображений необходимо ввести систему световых величин.

Здесь следует отметить, что освещённость изначально связана с **спектральной мощностью излучения** источника, которая рассчитывается по формуле

$$\Phi_{\lambda} = \frac{dQ}{dt d\lambda}$$

и представляет собой энергию, испускаемую с поверхности источника за 1 секунду на длине волны λ в пределах малой ширины спек-

тра излучения $d\lambda$. Строго говоря, источник излучения может излучать в определенном диапазоне длин волн, и сама **мощность излучения** источника будет представлять собой следующий интеграл:

$$\Phi = \int_{\lambda} \Phi_{\lambda} d\lambda.$$

Мощность излучения связана со **световым потоком**, представляющим собой ту часть мощности излучения источника, которая воспринимается человеческим глазом. Для вычисления величины светового потока необходимо проинтегрировать в диапазоне от $\lambda_1 = 380$ нм до $\lambda_2 = 800$ нм спектральную мощность излучения Φ_{λ} , помноженную на кривую спектральной чувствительности глаза $V(\lambda)$ для дневного зрения; результат следует умножить на фотометрический эквивалент излучения $K_m = 683$ лм/Вт:

$$\Phi_v = K_m \int_{\lambda_1}^{\lambda_2} \Phi_{\lambda} V_{\lambda} d\lambda .$$

Как следует из представленной формулы, световой поток измеряется в люменах (лм). На рис. 8 представлены кривые спектральной чувствительности человеческого глаза (*a* – для дневного, *b* – для сумеречного зрения), из которых следует, что строго говоря, диапазон спектра эффективного восприятия нормального глаза несколько уже, чем используемый для вычисления светового потока [5].

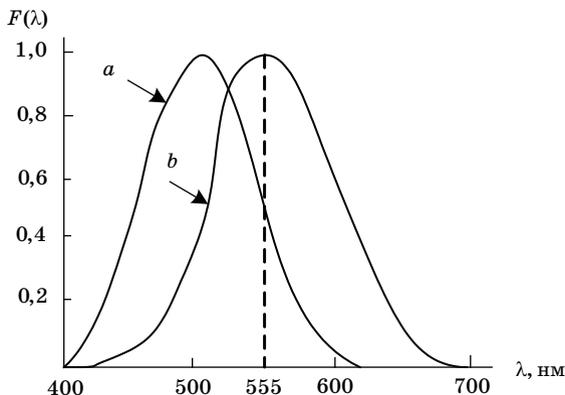


Рис. 8. Кривые спектральной чувствительности человеческого глаза

Особенность нашего зрения такова, что при равной мощности излучения всех длин волн видимого спектра мы лучше всего воспринимаем желто-зеленый цвет, т. е. излучение с длиной волны, равной 555 нм. Поэтому чувствительность глаза на этой длине волны принимается за единицу, а для остальных длин волн светового излучения она будет меньше единицы (при одинаковой мощности излучения). Максимумы на кривых а и б, равные единице, относительны. Дело в том, что палочковый аппарат ночного зрения человека намного чувствительнее, и для восприятия предельно малого светового сигнала (например, едва видимой точки на темном фоне) палочкам необходима примерно в пятьсот раз меньшая мощность, чем колбочкам. При этом палочки, действующие при периферическом (боковом) зрении, не позволяют определить цвета точки, в то время как колбочки, фиксирующие точку при прямом зрении, дают возможность увидеть и ее цвет.

Следует отметить, что оптическое излучение соответствует более широкому диапазону длин волн электромагнитного излучения от 1 нм до 1 мм и состоит из трех областей: ультрафиолетовой (УФ), видимой и инфракрасной (ИК).

Ультрафиолетовая область оптического излучения лежит в пределах 1...380 нм. Международная комиссия по освещению предложила следующее деление УФ-излучений с длинами волн от 100 нм до 400 нм: УФ-А – 315...400 нм; УФ-В – 280...315 нм; УФ-С – 100...280 нм.

Видимое излучение (свет), попадая на сетчатую оболочку глаза, в результате осознанного превращения энергии внешнего раздражителя вызывает зрительное ощущение. Диапазон длин волн монохроматических составляющих данного излучения соответствует 380...800 нм.

С изменением длины волны монохроматического излучения меняется и его цветовое восприятие глазом.

При разложении призмой видимого (белого) света в непрерывный спектр в последнем цвета плавно переходят один в другой так, что точно определить границы каждого цвета и связать их с определенной длиной волны трудно. Но приблизительно они выглядят так:

- фиолетовый — 380...440 нм;
- синий — 440...480 нм;
- голубой — 480...510 нм;
- зеленый — 510...550 нм;
- желто-зеленый — 550...575 нм;

- желтый — 575...585 нм;
- оранжевый — 585...620 нм;
- красный — 620...780 нм.

Длины волн монохроматических составляющих инфракрасного излучения больше длин волн видимого излучения (но не более 1 мм). МКО предложила следующее деление области ИК-излучений:

а) ближний диапазон, разделяющийся на две области: ИК-А – 780-1400 нм и ИК-В – 1400-3000 нм;

б) дальний (тепловой) ИК-С – 3000 нм (3 мкм)...1000 мкм.

Следующей светотехнической величиной используемой в процессе проектирования светочувствительных приборов является **сила света** – это световой поток, приходящийся на единицу телесного угла, в пределах которого он распространяется, т. е.:

$$I(\theta, \varphi) = \frac{d\Phi_v}{d\Omega},$$

где θ , φ – углы, задающие направление излучения; Ω – телесный угол.

Значение фотометрического эквивалента излучения K_m однозначно задаётся определением основной фотометрической величины – **канделы**, а именно, одна кандела – это сила света, излучаемая в каком-либо направлении источником монохроматического излучения с частотой $540 \cdot 10^{12}$ Гц, имеющим в этом направлении мощность излучения 1/683 Ватт/стерадиан. Излучение с частотой $540 \cdot 10^{12}$ Гц соответствует в воздухе длине волны 555 нм, т. е. максимуму кривой спектральной чувствительности адаптированного глаза для дневного зрения. Поэтому коэффициент K_m находится из тождества [5]

$$1 \text{ кд} = K_m \cdot V_\lambda(555) \cdot 1/683 \text{ Вт/ср},$$

откуда $K_m = 683 \text{ (кд·ср)/Вт} = 683 \text{ лм/Вт}$.

Как уже говорилось, освещенность представляет собой поверхностную плотность светового потока, падающего на поверхность (рис. 9). При равномерном распределении светового потока в пределах освещаемой поверхности S значение освещенности можно определить как:

$$E = \Phi_v / S.$$

Единицей освещенности является люкс (лк, от лат. lux — свет); т. е., 1 лк = 1 лм/ 1 м².

Освещенность и сила света точечного источника света при нормальном падении лучей (поверхность перпендикулярна лучам) связаны следующим соотношением:

$$E = I / r^2,$$

где r – расстояние от источника света до освещаемой поверхности. Следует помнить, что освещенность будет оставаться постоянной вдоль пучка лучей только тогда, когда они параллельны.

Если лучи от источника падают на поверхность под углом φ к нормали n (рис. 5), то тот же световой поток распределяется по площади, в $1/\cos\varphi$ раз большей, чем S (по площади $S/\cos\varphi$), и тогда:

$$E = I \cos\varphi / r^2.$$

Яркость поверхности изотропных излучателей для заданного направления – это отношение силы света, излучаемого в данном направлении, к площади проекции светящейся поверхности на плоскость, перпендикулярную к этому направлению (рис. 10):

$$L = I / S \cos\varphi.$$

При равномерном освещении диффузно отражающей поверхности уравнение, связывающее яркость этой поверхности с ее освещенностью, будет иметь вид:

$$L = \rho E / \pi.$$

где ρ – коэффициент отражения поверхности. Единицей яркости является кандела на квадратный метр ($\text{кд}/\text{м}^2$). $1 \text{ кд} / \text{м}^2$ – это яр-

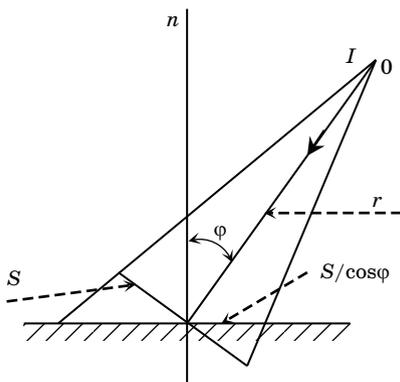


Рис. 9. К определению освещенности поверхности

кость такой плоской поверхности, которая в перпендикулярном направлении излучает силу света в $1 \text{ кд с } 1 \text{ м}^2$ поверхности.

Переход от дневного к сумеречному зрению у человека начинается тогда, когда яркость наблюдаемых им объектов становится меньше пороговой величины около 100 кд/м^2 .

Также при расчете и проектировании светочувствительных приборов часто оперируют энергетическими характеристиками. Среди них можно выделить следующие: облученность в точке поверхности, энергетическая светимость точки поверхности.

Под *облученностью в точке поверхности* понимают величину

$$E_s = d\Phi/dS,$$

измеряемую в Вт/м^2 , где $d\Phi$ – поток излучения приходящий на малый участок поверхности dS . *Энергетическая светимость в точке поверхности* представляет собой ту часть мощности $d\Phi_{\text{и}}$, которая испускается с малого участка поверхности dS , вокруг точки поверхности и рассчитывается по формуле

$$E_s = d\Phi_{\text{и}}/dS.$$

Для контроля над процессом регистрации изображений на матрице фотоаппарата – *экспозицией*, требуется управлять ее освещённостью.

Освещённость матрицы контролируется диафрагмой, поскольку освещённость обратно пропорциональна квадрату числа диафрагмы. Таким образом, диафрагма 2 даст на единицу площади матрицы вчетверо больше света, чем диафрагма 4. Именно поэтому диафрагму маркируют по степеням корня из двойки, так как каждое последующее увеличение ее диаметра, увеличивает освещённость матрицы вдвое. Специальный затвор позволяет открывать объектив на время экспозиции.

После накопления изображения на матрице, оно считывается. Как будет показано далее, именно датчик, предназначенный для регистрации изображения, способен внести значительные искажения изображения за счет собственных шумов. По этой причине

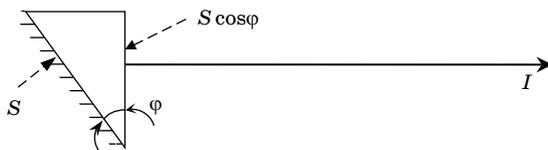


Рис. 10. К определению освещенности поверхности

перед выводом изображения на экран дисплея оператора его подвергают процессам медианной фильтрации, программному увеличению резкости в процессоре обработки изображения. Источники шумов, их влияние на характеристики лазерных систем видения будут рассмотрены в разд. 3.

1.3. Проблема видения в рассеивающих средах

Видимость в атмосфере или под водой – это возможность различать зрением удаленные объекты, отделенные от наблюдателя слоем воздуха той или иной степени мутности или толщей водной среды со своими оптическими характеристиками. Если же для наблюдения используются технические средства, формирующие изображение объекта, которое затем анализируется оператором, то говорят о видении, а изображающие средства, включая оператора, называют **системами видения**. Системы видения могут работать не только в видимом, но и в других, недоступных для зрительного восприятия участках спектрального диапазона, в соответствии с которыми они получили название систем радиовидения, тепловидения, телевидения и др.

Все системы видения можно подразделить на две большие группы: пассивные и активные. В пассивных системах видения изображение формируется при естественном освещении или за счет собственного (теплого) излучения объектов. К этим системам относятся, например, тепловизионные системы, системы низкоуровневого телевидения, большинство приборов ночного видения и др.

Активные системы видения используют для работы искусственную подсветку, создаваемую специальными устройствами различных типов – прожекторами, лампами-фарами, лазерными источниками. Системы видения с лазерным подсветом получили название **лазерных систем видения**.

Для понимания физического смысла технических характеристик, используемых при разработке ЛСВ, необходимо познакомиться с понятиями пространственной частоты и угловой пространственной частоты.

Процесс распространения лазерного света может быть описан достаточно строго в рамках элементарной теории идеально монохроматических волн. В большинстве случаев можно полагать, что все рассматриваемые световые волны линейно поляризованы в одном и том же направлении, и это позволяет описать световую волну в скалярном приближении.

Распространение монохроматической плоской волны можно описать функцией изменения напряженности поля во времени и пространстве следующим образом:

$$E = A \cos(\omega t - \mathbf{r}\mathbf{k}),$$

где A – амплитуда колебания; ω – круговая частота; t – время; \mathbf{r} – радиус вектор точки наблюдения; \mathbf{k} – волновой вектор. Колебания происходят синфазно во всех точках плоскости волнового фронта, т. е. $\mathbf{kr} = \text{const}$.

Плоская монохроматическая волна в комплексном представлении записывается в виде

$$E = A \exp(i\mathbf{kr} - i\omega t) = A \exp(i\mathbf{kr}) \exp(-i\omega t),$$

где i – мнимая единица. Выражение $A \exp(i\mathbf{kr})$ представляет собой комплексную амплитуду, определяющую собой как амплитуду, так и фазу волны. Поскольку множитель $\exp(i\omega t)$ представляет собой временную функцию, которая не подвергается изменениям при линейных операциях, из дальнейшего рассмотрения его можно исключить.

В общем случае монохроматическая волна имеет комплексную амплитуду, которую можно записать в виде функции

$$f(\mathbf{r}) = A(\mathbf{r}) \exp(\varphi(\mathbf{r})).$$

Эта функция содержит информацию как об амплитуде $A(\mathbf{r})$, так и о фазе колебаний $\varphi(\mathbf{r})$ в любой точке пространства.

Запишем выражение для комплексной амплитуды плоской монохроматической волны, выразив скалярное произведение \mathbf{kr} через проекции волнового вектора (k_x, k_y, k_z). Для простоты ограничимся рассмотрением двумерного случая, когда вектор \mathbf{k} лежит в плоскости xOz , (рис. 11).

В этом случае

$$f(x, z) = A \exp(i(k_x x + k_z z)),$$

при этом $k = \sqrt{k_x^2 + k_z^2}$, и если обозначить $u = k_x = k \sin \theta$ и выразить k_z через k , то получим $k_z = \sqrt{k^2 - k_x^2}$, и можно записать

$$f(x, z) = f(x, 0) \exp\left(iz \sqrt{k^2 - k^2 (\sin \theta)^2}\right).$$

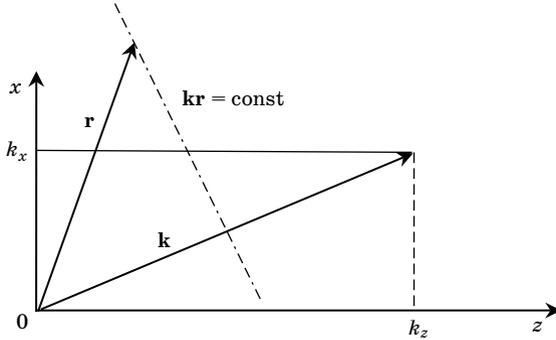


Рис. 11. Распространение плоской волны

Выражение $f(x, 0)$ – комплексная амплитуда в плоскости $z = 0$. Для расчета комплексной амплитуды в произвольной плоскости z необходимо помножить ее на функцию $\exp\left(iz\sqrt{k^2 - k^2(\sin\theta)^2}\right)$, которая носит название функции передачи свободного пространства. В плоскости $z = 0$ амплитуда плоской волны представляет собой гармоническую функцию $A\exp(iux)$, изменяющуюся вдоль оси x с **пространственной частотой** $\omega_x = k\sin\theta/(2\pi)$ и **пространственным периодом** $d_x = 1/\omega_x = 2\pi/u$.

Данный пример говорит о том, что при описании изменяющихся в пространстве процессов можно использовать понятие пространственной частоты, имеющей размерность $1/\text{м}$. Для понимания терминов пространственная частота и угловая пространственная частота можно использовать два следующих простых случая, проиллюстрированных рис. 12, а и 12, б, соответственно.

Представим, что на рис. 12, а дано изображение черных и белых прямых полос, соединенных между собой встык и имеющих размер вдоль оси x , равный 0,5 метра. Тогда период следования черных и белых полос составит 1 метр, а частота встреч черных (или белых) полос на 1м будет равна единице. Таким образом, пространственная частота для черных полос на рис. 12, а вдоль оси x составит $\nu_x = 1 \text{ м}^{-1}$.

На рис. 12, б представлены объекты, которые расположены относительно наблюдателя через одинаковое угловое расстояние α_x , которое можно считать угловым пространственным периодом. Тогда угловая пространственная частота определяется по формуле $\nu'_x = 1/\alpha_x = 1 \text{ град.}^{-1}$.

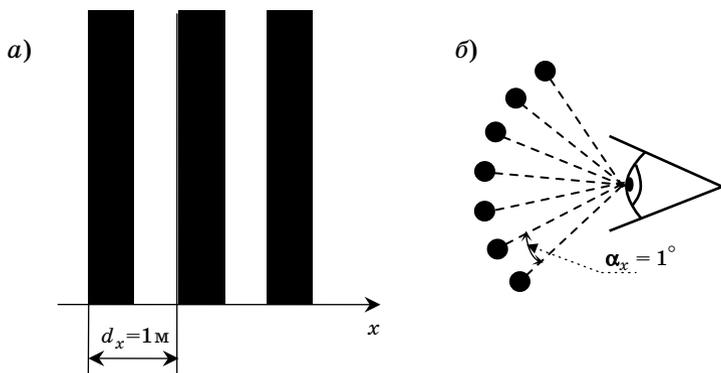


Рис. 12. К определению линейной (а) и угловой (б) пространственных частот

Отметим, что при повышении линейной или угловой пространственной частоты можно достичь такого предела, когда различение наблюдаемых объектов становится невозможным. В пределе последовательность черно-белых полос, показанных на рис. 12, а превращается в серое поле, а последовательность черных точек на рис. 12, б наблюдается как сплошная черная дуга окружности. Отметим, что способность глаза различать минимальные различия яркости смежных пространственных областей изображения характеризуется **контрастной чувствительностью**.

Чем выше яркостный контраст наблюдаемого на некотором фоне объекта, тем лучше его различает глаз человека. Помимо яркостного контраста существует также цветовой контраст, поскольку человек воспринимает целый спектр отраженных от объекта электромагнитных волн оптического диапазона и различает условия наблюдения, например, последовательности из синих и желтых полос или красных и зеленых. Для лазерных систем видения, работающих в монохромном режиме, важен именно контраст наблюдаемого изображения по яркости. Примем значение яркости для белой полосы на рис. 12, а равной, например, единице, а для черной – равной нулю. Рассмотрим процесс видения белых полос (объект) на черном поле (фон). Истинный (т. е. измеренный в плоскости объекта) **модуляционный контраст объекта** на данном фоне (при $L_{об} > L_{ф}$) определяется по формуле

$$k_{об} = \frac{L_{об} - L_{ф}}{L_{об} + L_{ф}} \quad (1.1)$$

и при заданных нами условиях равен единице. Если яркость фона увеличивается до значения 0,1, то контраст падает до значения 0,818.

Рассмотрим теперь, каким образом происходит процесс видения объектов на удаленном расстоянии в рассеивающих средах. Пусть объект наблюдается в приземном слое атмосферы на некотором фоне и регистрируется за счет различия яркостей объекта $L_{об}$ и фона $L_{ф}$. Модуляционный контраст может изменяться от 0 до 1 (см. ф-лу 1.1), причем термин «модуляционный» употребляется в пространственном смысле. Если пространственная модуляция яркости объекта отсутствует, то его изображение представляется как равномерно освещенное поле, которое не несет никакой информации.

Если объект из черно-белых полос обладает слишком высокой пространственной частотой, то человек воспринимает его изображение как серое, т. е. сплошное с нулевым контрастом. Это означает, что при исходном заданном единичном контрасте объекта воспринимаемый контраст изображения зависит от пространственной частоты.

Введение понятия модуляционного контраста для описания пространственных соотношений яркостей и фонов чрезвычайно удобно, поскольку позволяет использовать хорошо развитый математический аппарат теории линейной пространственно-частотной фильтрации для объективной оценки качества работы любой изображающей системы.

Широко применяемая при анализе качества изображения *модуляционная передаточная функция* (МПФ) является отношением модуляционных контрастов изображения $k_{из}$ и объекта $k_{об}$ в виде функции пространственной частоты ν [1]

$$T(\nu) = \frac{k_{из}(\nu)}{k_{об}}. \quad (1.2)$$

Способность обнаружения и опознавания удаленных объектов зависит от способности системы наблюдения пропускать спектр пространственных частот.

По мере удаления от объекта истинный контраст начинает уменьшаться из-за изменения яркостей $L_{об}$ и $L_{ф}$. Эти изменения обусловлены двумя причинами:

– во-первых, яркости $L_{об}$ и $L_{ф}$ уменьшаются вследствие ослабления излучения при прохождении им замутненного слоя z от объекта до ЛСВ;

– во-вторых, при работе ЛСВ в дневное время между объектом и аппаратурой наблюдения появляется слой дымки, яркость которой L_d увеличивается с расстоянием.

Наличие в атмосфере большого количества взвешенных частиц (аэрозоль, минеральная и органическая пыль, дымовые частицы и др.) приводит к интенсивному рассеянию света и возникновению своеобразной световой пелены. Этот рассеянный свет и создает атмосферную дымку, которая приводит к уменьшению видимого контраста (т. е. контраста, наблюдаемого на расстоянии z от объекта). Следовательно, результирующие наблюдаемые яркости объекта и фона с расстояния z будут иметь такой вид

$$L_{об.н} = L_{об} \exp(-\bar{\varepsilon}z) + L_d(z);$$

$$L_{ф.н} = L_{ф} \exp(-\bar{\varepsilon}z) + L_d(z),$$

где $\bar{\varepsilon}$ – среднее значение показателя ослабления светового потока в пределах слоя z .

Тогда выражение для контраста с помощью (1.2), наблюдаемого на расстоянии z от плоскости объекта, преобразуется к виду

$$k_{об.н} = \frac{k_{об}}{1 + \frac{2L_d(z)\exp(\bar{\varepsilon}z)}{L_{об} + L_{ф}}}.$$

Когда при зрительном восприятии наблюдаемый контраст $k_{об.н}$ становится равным пороговому контрасту $k_{пор}$, воспринимаемому глазом, то расстояние z , на котором это происходит, называют дальностью видимости объекта S . И тогда пороговый контраст глаза можно определить по соотношению

$$k_{об.н} = \frac{k_{об}}{1 + \frac{2L_d(S)\exp(\bar{\varepsilon}S)}{L_{об} + L_{ф}}},$$

где $\bar{\varepsilon}$ – среднее значение показателя ослабления светового потока на расстоянии, равном дальности видимости; $L_d(S)$ – яркость дымки в атмосферном слое протяженностью S . Тогда дальность видимости можно вычислить по формуле

$$S = \frac{1}{\bar{\varepsilon}(S)} \left[\ln \left(\frac{k_{об}}{k_{пор}} - 1 \right) + \ln \left(\frac{L_{об} + L_{ф}}{2L_d(S)} \right) \right].$$

Итак, дальность видимости несамосветящегося объекта сложным образом зависит от показателя ослабления, яркостей объекта и фона при данном освещении и яркости дымки.

В общем случае яркость слоя дымки по мере увеличения протяженности этого слоя сначала возрастает, но затем рост яркости с увеличением расстояния z замедляется, и при $z \rightarrow \infty$ величина $L_d(z)$ стремится к некоторому предельному значению L_∞ , которое обычно отождествляется с яркостью неба у горизонта. Зависимость $L_d(z)$ называется световоздушным уравнением, которое с некоторыми допущениями может быть представлено в виде

$$L_d(z) = L_\infty \left[1 - \exp^{-\tau(z)} \right],$$

где оптическая толщина слоя $\tau(z)$ равна

$$\tau(z) = \bar{\varepsilon}z.$$

При сильной дымке ее яркость в горизонтальном направлении достигает уровня насыщения уже на расстояниях нескольких километров, а в туманах – на расстояниях сотен метров. С помощью формулы (1.3) можно оценить дальность видимости несамосветящегося объекта при условии выполнения световоздушного уравнения. Для этого вначале определяют величину [1]:

$$k_{\text{пор}} = \frac{k_{\text{об}}}{1 + \frac{2L_\infty(1 - \exp(-\bar{\varepsilon}S))\exp(\bar{\varepsilon}S)}{L_{\text{об}} + L_\Phi}},$$

а затем дальность видимости

$$S = \frac{1}{\bar{\varepsilon}} \ln \left[\left(\frac{k_{\text{об}}}{k_{\text{пор}}} - 1 \right) \left(\frac{L_{\text{об}} + L_\Phi}{2L_\infty} \right) + 1 \right].$$

Проблема видимости в атмосфере связана с вопросами обнаружения и распознавания удаленных объектов и, в первую очередь, с определением расстояния, на котором объект обнаруживается на пределе восприятия. В этом случае под *пороговым контрастом* $k_{\text{пор}}$ понимают пороговый контраст обнаружения (порог обнаружения). Согласно [6] при определении оптических характеристик атмосферы используется значение $k_{\text{пор}} = 0,02$. Поскольку атмосфера является фильтром пространственных частот, то $k_{\text{об.н}}$ также зависит от пропускаемой пространственной частоты. Зависимость

$k_{\text{пор}}$ от пространственной частоты обусловлена специфическими процессами пространственно-частотной фильтрации, присущими зрительному анализатору.

Согласно современным представлениям физиологии зрительного восприятия, рецептивные поля зрительного анализатора являются узкополосными пространственно-частотными фильтрами, каждый из которых настроен на определенную среднюю пространственную частоту. При восприятии изображения в высших отделах зрительного анализатора формируется оптимальный для решения данной задачи фильтр путем выбора соответствующих резонансных фильтров из имеющегося их набора. Огибающая этого набора фильтров как раз и характеризует зависимость порогового модуляционного контраста от пространственной частоты.

Дальность видимости в воде ограничивается практически полным ослаблением в поверхностном слое электромагнитного излучения оптического диапазона. Только в узком спектральном диапазоне, соответствующем сине-зеленой области видимого спектра, излучение проникает до значительных глубин.

Ослабление светового потока при его распространении в природных водах примерно на две трети вызвано рассеянием и на одну треть – поглощением. Ослабление приводит к уменьшению освещенности объектов по мере их удаления от источников подсвета и к ослаблению их яркости водой.

Другим важным фактором, ограничивающим дальность видимости в воде, является образование световой дымки, вызванное обратным рассеянием воды и приводящее к снижению контраста изображения.

Дальность видимости подводных объектов обычно оценивают по значениям предельной безразмерной оптической глубины обнаружения $\tau_{\text{пр}}$:

$$\tau_{\text{пр}} = \int_0^{z_0} \varepsilon(z) dz = \bar{\varepsilon} z_0,$$

которая в общем случае определяется не только оптическими свойствами воды, но и направлением визирования.

Вследствие того, что основным фактором, определяющим дальность видимости в воде, является затухание света, дальность слабо зависит от изменения интенсивности подсвета. Действительно, для увеличения предельной оптической глубины $\tau_{\text{пр}}$ всего на две единицы нужно увеличить мощность источника излучения в 10 раз [7].

Изложенные основные положения теории видимости в принципе могут быть привлечены и для анализа характеристик систем видения. Отличие состоит лишь в том, что при использовании систем видения оператор наблюдает не сам объект, а его изображение на экране дисплея, причем характеристики изображения (яркость, контраст и др.) могут регулироваться электронным способом независимо от характеристик объекта. Приемное устройство любой системы видения регистрирует дифференциальный сигнал – разность мощностей (энергий) оптического излучения, отраженного объектом и фоном, а также мощность помехи (внешнее фоновое излучение, помеха обратного рассеяния и др.). Сигнал на выходе фотодетектора ΔI должен не только превышать некоторое минимальное значение, соответствующее порогу чувствительности фотоприемной системы, но и превосходить среднеквадратическое значение системного шума $\sigma_{\text{сист}}$.

Этот флуктуационный шум обусловлен различными причинами, в том числе флуктуациями числа фотонов регистрируемого оптического излучения, обусловленными квантовой природой света, флуктуациями оптических характеристик канала распространения излучения, внутренним шумом фотоэлектронного тракта и другими видами шумов, которые будут рассмотрены далее. Обе величины ΔI и $\sigma_{\text{сист}}$ определяют важнейшие характеристики изображения – контраст и отношение сигнал/шум, поэтому эффективность системы видения в общем случае может ограничиваться любой из них.

Каждой изображающей системе присущ некоторый минимальный контраст, определяемый на выходе системы (например, на экране дисплея) в соответствии с требованиями наблюдателя или другого анализирующего устройства. Данный пороговый контраст зависит от пространственной частоты ν в плоскости анализа, что наиболее отчетливо проявляется при наблюдении с помощью изображающей системы штриховой миры в виде последовательности белых и черных прямоугольных полос различной пространственной частоты. При определенной высокой пространственной частоте полосы перестают различаться. Эта максимальная частота и характеризует предельное пространственное разрешение.

Иная ситуация возникает при ограничении пространственной разрешающей способности уровнем шума. Помехи, воздействуя на сигнал изображения, вызывают флуктуации яркости в изображении, которые искажают его, а при высоком уровне помех делают неразличимым. Особенно отчетливо флуктуационный шум прояв-

ляется на дальностях, близких к предельной, когда регистрируются слабые сигналы. В этом случае изображение объекта формируется на экране дисплея на фоне случайного шумового поля яркости, проявляющегося в виде «снега», и именно это может ограничить пространственное разрешение.

1.4. Назначение и технические характеристики ЛСВ

Появление лазеров предопределило создание таких технических систем, в основу реализации которых были положены современные возможности по созданию и применению лазерного излучения. Данные системы предоставляют сегодня возможность использовать лазеры в различных системах, и в частности, в лазерных системах видения, предназначенных для получения изображения там, где дальность действия пассивных оптических систем видения и их разрешающая способность невысока.

Области применения лазерных систем видения:

- системы ночного видения;
- системы охраны наземных объектов;
- системы подводного видения;
- системы видения для проведения работ в космосе, например, стыковки;
- системы дефектопической проверки результатов работы различных технологических процессов, например, сварочных швов.

Большинство систем видения решает задачу формирования изображений удаленных объектов с целью их обнаружения или распознавания в условиях недостаточной естественной освещенности или ночью. Они позволяют осуществлять наблюдение объектов при крайне низких уровнях освещенности и наличии помех различного рода, т. е. в условиях, когда невозможно получить удовлетворительное изображение с помощью прибора ночного видения (ПНВ) или камеры низкоуровневого телевидения.

Применение в качестве внешней подсветки традиционных ИК-прожекторов неэффективно уже на дальностях, превышающих 700–800 метров из-за слишком высоких уровней требуемой мощности [1]. Кроме того, возникающая в канале распространения помеха обратного рассеяния создает значительные сложности при приеме слабого отраженного сигнала.

Системы видения, применяемые в ночное время на судах и кораблях, могут помочь навигаторам избежать опасных столкновений судна с посторонними объектами на поверхности воды, которые мо-

гут встретиться по курсу судна и стать причиной его повреждения или даже гибели.

С их помощью могут быть обнаружены и идентифицированы объекты небольших размеров, которые иногда невозможно увидеть на экране радиолокатора, такие как небольшие камни, торчащие над водой, небольшие бревна и плавающие льдины, суда малых размеров, изготовленные из стекловолокна или других материалов, а также спасательные плоты или люди, находящиеся в воде.

Системы охраны объектов способны получать изображения автомобилей в условиях плохой видимости на больших расстояниях, что позволяет оператору распознать тип, а, следовательно, и потенциальную опасность.

Системы подводного видения являются эффективными средствами контроля морского дна в прибрежных водах и способны производить дистанционный поиск затонувших малогабаритных объектов, способных своим присутствием нарушить экологический баланс в промысловых зонах. Также такие системы позволяют производить контроль строящихся и эксплуатирующихся объектов, например, подводных нефтяных и газовых трубопроводов, осуществлять поиск и обнаружение дефектов, которые могут приводить к утечкам газа и нефти.

Ощутимого прогресса в развитии активных систем видения удалось достигнуть в 60-х годах в связи с созданием лазеров, излучение которых обладает высокими спектральной интенсивностью, направленностью и монохроматичностью. Особый интерес проявляется к ЛСВ, работающим в режиме стробирования по дальности, когда объект наблюдения зондируется короткими (10...50 нс) мощными световыми импульсами, длительность которых значительно меньше времени распространения импульса до объекта и обратно. Приемная система снабжена быстродействующим затвором, открывающимся в такт с посылкой импульсов подсвета на короткое время, через определенную временную задержку, которая соответствует дальности локации. При этом оператор видит только сам объект и участок пространства, непосредственно его окружающий.

Широкое распространение ЛСВ получили в системах авиационной разведки наземных объектов, подвижных системах переднего обзора для получения изображения объектов ночью и в плохих погодных условиях, в устройствах высокоскоростной фотографии и аэрофотосъемки, а также в подводном видении.

Для ЛСВ важное значение имеет качество воспроизводимого изображения, которое определяется субъективным впечатлением,

изменяющимся от крайне слабого до великолепного. Это воспринимаемая метрика, зависящая не только от оптимального функционирования системы «глаз–мозг», но и от входных параметров других сенсорных систем, эмоций, опыта и т. д. Воспринимаемое качество одного и того же изображения существенно различается между наблюдателями, а также отличается по времени для одного наблюдателя. Поэтому качество изображения не может оцениваться в абсолютных величинах, а визуальные психофизиологические исследования не позволяют количественно измерить все показатели изображающих систем. Существует множество формул для оценки качества изображения, каждая из которых получена для конкретных условий наблюдения эмпирическим путем. Все они в той или иной мере используют две основные характеристики – пространственную и энергетическую разрешающую способность.

Пространственная разрешающая способность ЛСВ характеризует возможность системы различать минимальные пространственные размеры объектов и их элементов. В общем случае она ограничивается дифракцией и аберрациями оптической приемопередающей системы, размерами фотодетектора (пиксела) и шагом матричной фотоприемной структуры, шириной полосы видеочастот электронного тракта, разрешением телевизионного монитора. Каждая из подсистем ЛСВ обычно характеризуется собственной метрикой разрешающей способности. Так, разрешение оптических систем оценивается критерием Рэля или кружком размытия – изображением точечного источника, формируемого оптической системой. **Разрешающая способность по Рэлю** – это способность оптической системы изображать отдельно два близко расположенных точечных предмета. Иногда вместо понятия «разрешающая способность» используют понятие предела разрешения, то есть минимального расстояния, при котором два близко расположенных точечных предмета будут изображаться как отдельные. Критерий Рэля заключается в том, что при провале в распределении интенсивности изображения двух близких точек в 20% эти точки будут восприниматься как отдельные. Для этого необходимо, чтобы центральный максимум в изображении одной точки приходился бы на первый минимум в изображении другой (рис. 13).

Иллюстрацией различения двух ярких точек в оптических системах может служить случай слияния изображения двух звезд в телескопе тогда, когда угловое расстояние между ними меньше разрешающей способности телескопа. Таким образом, разрешение по Рэлю характеризует качество изображения астрономических

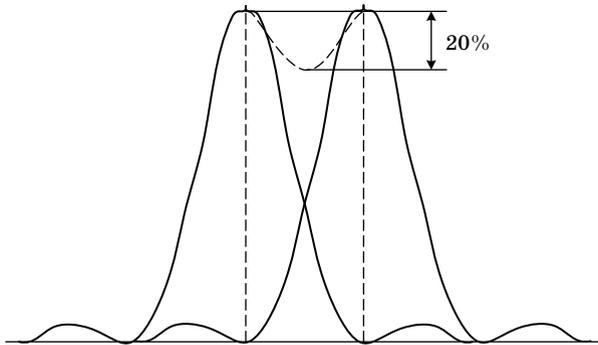


Рис. 13. Минимальная разрешающая способность по Релею

телескопов, спектральных приборов и других оптических систем, для которых предметами являются точки или линии.

Разрешающая способность фотоприемного матричного устройства определяется угловым размером чувствительного элемента (пиксела), телевизионного монитора – числом телевизионных линий, а электронного тракта – частотой Найквиста.

Современный подход к оценке разрешающей способности ЛСВ основан на теории линейной пространственной фильтрации, в рамках которой процесс видения рассматривается как пространственная фильтрация наблюдаемого объекта многозвенным двумерным фильтром. При этом модуль передаточной функции такого фильтра – модуляционная передаточная функция – характеризует зависимость контраста мелких деталей изображения от их размеров. Предельное разрешение определяется пространственной частотой, на которой МПФ падает до определенного уровня, например, до уровня 0,1 [1].

Энергетическая разрешающая способность (контрастная чувствительность) характеризует возможность различения сигналов от крупноразмерных объектов на окружающем фоне за счет существующего контраста. Она зависит от светоэнергетических характеристик оптической системы, чувствительности фотодетектора и системного шума и не зависит от пространственной разрешающей способности.

Чувствительность определяется тем минимальным уровнем сигнала, который может быть зарегистрирован приемной системой. Обычно это такой входной сигнал, который формирует на выходе фотодетектора отношение сигнал/шум, равное 1. Вместе с тем, качество изображения, воспроизводимое ЛСВ и наблюдаемое опера-

тором на экране монитора, определяется воспринимаемым отношением сигнал/шум $(S/N)_{\text{восп}}$, которое может превышать $(S/N)_{\text{вид}}$ на порядок за счет интегрирующих свойств зрительного анализатора. Поэтому в современных ЛСВ формируется удовлетворительное по качеству изображение на экране монитора даже при значениях $(S/N)_{\text{вид}}$, существенно меньших 1, например 0,1.

Предельная дальность видения – максимальное расстояние, на котором система видения реализует требуемый уровень восприятия с заданной вероятностью (например, дальность обнаружения конкретных объектов с вероятностью 0,9). Предельная дальность видения зависит одновременно от пространственного разрешения и чувствительности.

Если эффективность системы ограничена пространственным разрешением, то дальность видения z определяется только размером объекта r_0 и предельным угловым разрешением α' , т. е. $z = r_0/\alpha'$.

В изображающих системах предельное разрешение ограничивается частотой пространственной дискретизации (или частотой Найквиста). Важно отметить, что при такой предельной оценке качества ЛСВ предполагается наличие высококонтрастного изображения, т. е. достаточно высокое отношение сигнал/шум на выходе приемной системы.

В другом предельном случае максимальная дальность видения ограничивается чувствительностью ЛСВ или ее энергетической разрешающей способностью. Эта ситуация возникает при наблюдении крупноразмерных малоконтрастных объектов. Для таких систем определяющей характеристикой является отношение сигнал/шум, вычисляемое на нулевой пространственной частоте

$$\frac{S}{N} = \frac{\exp(-\bar{\varepsilon}z)\Delta I}{\sqrt{\sigma_{\text{сист}}^2}},$$

где ΔI – дифференциальный сигнал (например, разность средних значений сигнального и фонового фототоков па выходе фотодетектора); $\sigma_{\text{сист}}^2$ – дисперсия системного шума.

Следует отметить, что формула применима только в тех условиях, когда угловой размер наблюдаемого объекта существенно превышает угловую пространственную разрешающую способность ЛСВ. Но пространственная разрешающая способность однозначно не определяет эффективность действия или качество изображения, так как она не учитывает чувствительность системы, так как

крупноразмерные низкоконтрастные объекты не обнаруживаются на дальностях, допустимых по разрешению, если уровень сигнала ниже уровня шума фотоприемного устройства.

Минимальный разрешаемый контраст. Для корректного анализа ЛСВ необходима характеристика, учитывающая одновременно и пространственное разрешение и чувствительность. Такой характеристикой может служить минимальный разрешаемый контраст (МРК) объекта, определяемый как минимальный контраст наблюдаемого объекта k_{\min} , при котором воспринимаемое на экране монитора оператором отношение сигнал/шум превышает пороговое значение.

Реальный контраст объекта и контраст его изображения на экране не имеют однозначной связи, поскольку регулировкой яркости и контрастности ТВ-приемника можно изменять контраст изображения, в то же время как реальный контраст объекта остается постоянным, не зависящим от рабочих характеристик ЛСВ. При увеличении контраста изображения на экране монитора электронным способом одновременно увеличивается контрастность шумовой картины, в результате чего восприятие объекта не улучшается. Психофизические исследования показывают, что человеческий глаз в силу природных особенностей и присущих ему свойств пространственного и временного фильтра наиболее чувствителен не к воспринимаемому контрасту, а к воспринимаемому отношению сигнал/шум $(S/N)_{\text{восп}}$, определяемому выражением:

$$\left(\frac{S}{N}\right)_{\text{восп}} = \frac{\Delta I}{\sqrt{\sigma_{\text{сист}}^2}} T_{\text{сист}}(v_x, v_y) T_{\text{гл}}(v_x, v_y),$$

где $T_{\text{сист}}(v_x, v_y)$ – системная МПФ; $T_{\text{гл}}(v_x, v_y)$ – МПФ зрительного анализатора.

Связь между воспринимаемым отношением сигнал/шум и МРК по пороговому значению отношения сигнал шум $(S/N)_{\text{восп.пор}}$

$$\text{МРК} = (S/N)_{\text{восп.пор}} \frac{\sqrt{\sigma_{\text{сист}}^2}}{T_{\text{сист}}(v_x, v_y) T_{\text{гл}}(v_x, v_y) (L_{\text{об}} + L_{\text{ф}})}.$$

Минимальный разрешаемый контраст различается для каждой конкретной ЛСВ, так как зависит от системной МПФ и системного шума. Характерно, что МРК не учитывает энергетическое уменьшение контраста объекта при его наблюдении в поглощающих и рассеивающих средах, например в атмосфере. Поэтому предельная

дальность действия ЛСВ должна определяться как дальность действия, при которой для заданного уровня восприятия реализуется наблюдаемый МРК.

Системы видения, ограниченные разрешением и чувствительностью, могут изменять характер ограничения в зависимости от пространственной частоты.

1.5. Обобщенная структурная схема ЛСВ

Специфика построения ЛСВ обусловлена главным образом уникальными свойствами лазерного излучения, которые проявляются в виде различных эффектов при распространении излучения в рассеивающей и случайно-неоднородной среде, а также в процессе формирования изображения приемным каналом.

В качестве обобщенной структуры ЛСВ может быть использована схема, приведенная на рис. 14 [1].

Передающий канал формирует зондирующее лазерное излучение. Основными его элементами являются лазер 1, формирующая оптическая система 3 для изменения угловой расходимости лазерного излучения, устройство управления лазерным излучением 2 для модуляции и сканирования лазерного пучка. Сформированное передатчиком лазерное излучение проходит через среду распространения 4 (атмосфера, вода) и подсвечивает объект 5, имеющий некоторое пространственное распределение коэффициента отражения по интенсивности – переизлучающего отраженный сигнал определенного спектрального состава.

Отраженное излучение оказывается промодулированным по интенсивности в соответствии с двухмерным распределением коэффициента отражения и после обратного прохода через среду 4 регистрируется приемным каналом, содержащим приемную оптическую систему 6, усилитель – преобразователь 7 (например, электронно-оптический преобразователь – ЭОП) и фотоприемное устройство 8 (ФПУ). Помимо фотоприемника ФПУ может содержать сканирующее устройство, перемещающее в пространстве синхронно с лазерным пучком угловую приемную диаграмму ФПУ, осуществляя тем самым эффективную пространственную селекцию.

Центральное устройство управления содержит блок обработки данных 9 и контроллер управления 10. Блок обработки данных предназначен для приема, цифрового преобразования и обработки входного сигнала с ФПУ, а также для формирования выходного аналогового сигнала на входе монитора. Контроллер управления



Рис. 14. Обобщенная структура лазерной системы видения

осуществляет формирование синхронизирующих и управляющих импульсов, индикацию режимов работы ЛСВ, выполнение команд оператора.

Наведение ЛСВ на объект в условиях достаточной естественной освещенности проводится оператором с помощью визирного устройства 11, установленного на подвижной платформе 12 таким образом, что его оптическая ось соосна с оптической осью приемного канала. В этом режиме оператор 13 поворачивает платформу по азимуту и углу места с помощью рукояток ручного привода. При работе ночью или в автоматизированном режиме оператор наблюдает участок освещаемой передатчиком местности на экране ТВ-монитора 14, а поворот платформы осуществляется электромеханическими приводами 15 по командам контроллера управления.

Замыкающим звеном в ЛСВ является оператор. Поэтому все тактико-технические характеристики должны быть ориентированы на оператора как на звено, принимающее решение о наличии изображения и его качестве. Именно зрительный аппарат человека-оператора является исходным пунктом при проведении светонергетического расчета ЛСВ.

Более того, помимо тех технических параметров, которые следует учитывать при разработке лазерных систем видения, следует помнить о том, что конечное решение о наличии цели, ее ракурсе с последующей ее идентификацией производит непосредственно человек по анализу изображения, которое формирует система видения. Как следствие следует говорить о необходимости при проектировании лазерных систем видения учитывать особенности зрительного восприятия человеком изображения объекта зондирования.

Зрительное восприятие изображения объекта начинается с выделения его общих структурных особенностей. В первую очередь воспринимается отношение предметов и пространства, т. е. объекта и фона, которое зависит от контраста изображения, способности передавать мелкие детали изображения, его «текстуру». В процессе зрительного восприятия осваиваются отношения между входящими в изображение предметами, затем между более мелкими деталями предметов для их идентификации. Только после этого создается четкое представление о целом и производится окончательное принятие решения оператором. Поэтому вопросы о том, как получить контрастное, детализированное изображение и как его обработать для того, чтобы предоставить оператору, имеют столь важное значение при разработке лазерных систем видения и рассматриваются подробно в следующих разделах учебного пособия.

1.6. Устройство передающего канала лазерных систем видения

Рассмотрим реализацию передающего канала ЛСВ на примере реализации систем активного подсвета для активных приборов ночного видения. Раньше для реализации подсвета объектов в системах ночного видения использовались прожекторы инфракрасного (ИК) излучения на лампах накаливания и газоразрядных лампах. Согласно [8] унифицированный малогабаритный прожектор на основе эффективного импульсного лазерного полупроводникового излучателя, позволяет не только увеличить дальность видимости в активном режиме работы, но и повысить помехоустойчивость

и эффективность работы комплекса в целом в условиях эксплуатации. Преимуществом применения лазерного полупроводникового излучателя является и возможность реализации режима активно-импульсного режима работы приборов ночного видения [8].

Пржектор ПЛ-1 представляет собой основанное на базе импульсного лазерного полупроводникового излучателя современное поколение источников ИК-подсвета. Он содержит единый блок, объединяющий все функциональные звенья электронного передающего тракта излучателя, блок питания и формирующую оптическую систему, а также систему обогрева защитного стекла [9]. Пржектор ПЛ-1 формирует удобное для оператора пятно излучения прямоугольной формы с однородным распределением энергетической яркости излучения, в то время как лампы прожекторы формируют «колоколообразное» распределение (рис. 15).

Пржектор ПЛ-1 характеризуется меньшим энергопотреблением и повышенным ресурсом работы по сравнению с существующими аналогами, имеет меньшую массу и габариты. В частности, пржектор ПЛ-1 имеет массу 7,0 кг, габариты 246×174×177 мм, энергопотребление 50 Вт, в то время как ламповый пржектор Л4 имеет аналогичные параметры соответственно 20,5 кг, Ø300×280 мм, 400 Вт [9]. При этом исключаются такие дефекты ламповых пржекторов, как взрыв лампы, незажигание, нестабильность яркости, разрушение отражателя и светофильтра при взрыве лампы и др. Основные параметры лазерных пржекторов приведены в табл. 1 [9].

Оптическая схема лазерного ИК пржектора ПЛ-1 представлена на рис. 15 [9]. Полупроводниковый лазер 1 создает поток излучения. Фильтр 2 предназначен для коррекции спектра излучения источника 1 с целью уменьшения демаскировки работающего пржектора. Дальность видения излучения ИК-пржектора невооруженным глазом не превышает 200 м.

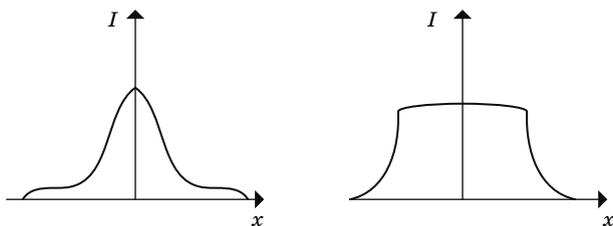


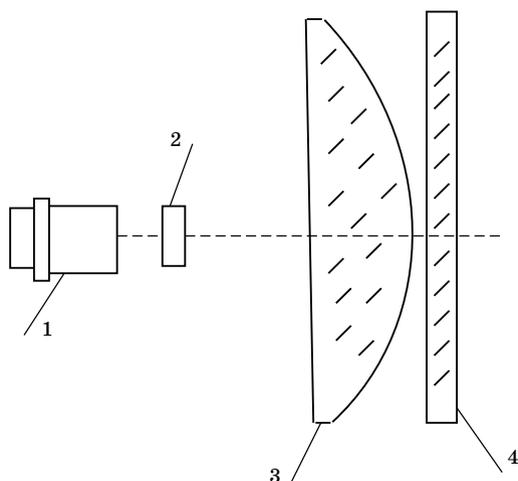
Рис. 15. Распределения интенсивности ИК подсвета, формируемые посредством лампового пржектора (слева) и полупроводникового лазерного излучателя ПЛ-1 (справа)

Таблица 1

Технические характеристики прожектора ПЛ-1		
Выходная мощность излучения, не менее, Вт	0,15	
Длина волны излучения, мкм	0,85	
Угловая расходимость излучения на уровне 0,25 от максимума энергии, град:		
– по вертикали	0,75°	
– по горизонтали	1,5°	
Время непрерывной работы с сохранением характеристик, не менее, час	6	
Потребляемая мощность, не более, Вт:		
– без системы обогрева защитного стекла	20	
– с системой обогрева защитного стекла	50	
Напряжение питания, В	27	
Масса, не более, кг	7	
Технические характеристики прожекторов ОУ-6 и ОУ-6-01		
Выходная мощность излучения, не менее, Вт	0,15	
Длина волны излучения, мкм	0,85	
Угловая расходимость излучения на уровне 0,25 от максимума энергии, град:		
– по вертикали	0,75°	
– по горизонтали	1,5°	
Углы слежения прожектора за визирной осью ПНВ:	ОУ-6	ОУ-6-01
– в вертикальной плоскости	От минус 6° до +30°	От минус 8° до +32°
– в горизонтальной плоскости	От минус 5° до +5°	От минус 5° до +5°
Время непрерывной работы с сохранением характеристик, не менее, час	6	
Потребляемая мощность, не более, Вт:	100	
Напряжение питания, В	27	
Масса, не более, кг	15	

Линза 3 служит для формирования пятна подсвета требуемых размеров. Она имеет фокусное расстояние 114,5 мм при относительном отверстии 1:1,4. Особо следует рассмотреть вопрос об оптике формирования излучения лазерного прожектора [9]. В простейшем случае в качестве такой оптики может быть использована полусферическая линза 3, рассчитанная на минимум сферической аберрации (рис. 16).

Под сферической аберрацией подразумевают эффект размытия фокуса. Световые лучи, проходящие сквозь линзу вблизи оптиче-



*Рис. 16. Оптическая схема ИК-проектора ПЛ-1:
1 – импульсный лазерный полупроводниковый излучатель;
2 – фильтр; 3 – линза; 4 – защитное стекло*

ской оси, фокусируется в одной точке. Световые лучи, проходящие сквозь краевые зоны линзы, при этом могут фокусироваться в зависимости от исполнения линзы как до точки фокуса осевых лучей (ближе к линзе), так и за ней.

Изображение точки, сформированное линзой со сферической aberrацией – обычно яркая точка, окруженная ореолом света. Эффект такой aberrации на изображении в целом проявляется в смягчении контраста и размытии деталей.

Такая оптика проста, но из-за значительных aberrаций излучение сильно рассеивается вне заданного угла подсвета, что приводит к энергетическим потерям, достигающим до 25–30 %. Для снижения aberrаций и повышения коэффициента полезного действия оптики могут быть рекомендованы линзы с асферической поверхностью (параболические, гиперболические) [9]. Данные линзы уменьшают aberrации и дают в итоге подсвет объекта с меньшими энергетическими потерями. Для работы на различных дальностях целесообразно соответственно изменять и угол подсвета, который можно функционально связать с соответствующим изменением временной задержки в приемном тракте лазерной системы видения. Для этого может быть использован вместо линзы вариообъектив с плавно изменяемым фокусным расстоянием.

Конструкция рассматриваемого прожектора ПЛ-1 показана на рис. 17 [9].

Его основными узлами являются корпус 1, задняя 2 и защитная 3 крышки. В корпусе прожектора установлены полупроводниковый источник лазерного подсвета 4, светофильтр 5 в оправе, линза 6 в оправе, блок питания источника подсвета 7, влагопоглотитель 10 и терморегулятор 12. Передняя часть корпуса закрыта защитным стеклом 8 в оправе. Для предохранения стекла 8 от запотевания и заиндевления используется терморегулятор в сочетании с нагревательным элементом – токопроводящим покрытием, нанесенным на внутреннюю поверхность стекла 8. В крышке 2 установлена катушка индуктивности (дрессель) 9, являющаяся фильтром низких частот. Она предназначена для подавления помех по цепи питания при работе блока питания источника лазерного подсвета. Защитная крышка 3 в закрытом и открытом состоянии крепится на корпусе винтом 11.

Дальнейшим развитием схемы прожектора ПЛ-1 следует считать лазерные прожекторы ОУ-6 и ОУ-6-01 (табл. 1). Их важным достоинством является возможность слежения за оптической осью лазерной системы видения по горизонту и по вертикали. Это достигается с помощью электромеханического привода. При этом может быть реализовано автоматическое удержание перемещающегося по фронту объекта наблюдения в пределах пятна подсвета за счет формирования вспомогательных пятен подсвета, расположенных по периметру основного пятна. Для этого по периметру основного

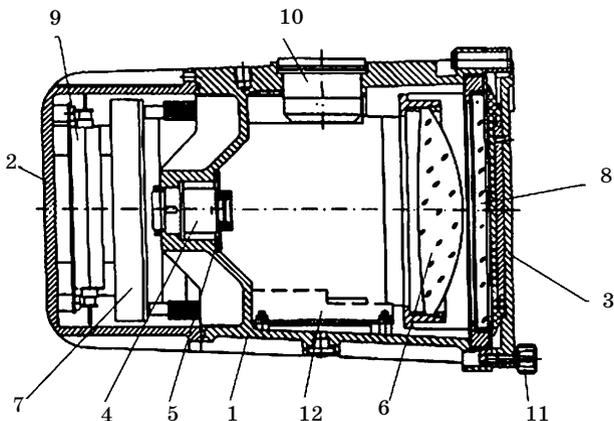


Рис. 17. Конструкция лазерного ИК-прожектора

лазерного источника подсвета следует расположить четыре вспомогательных. Их рабочая частота должна быть пренебрежимо малой по сравнению с рабочей частотой основного источника подсвета, чтобы оператор не видел через ПНВ эти пятна подсвета. Пусть частота вспомогательных источников подсвета, находящихся слева и справа от основного, составляет соответственно 3 Гц и 5 Гц, а находящихся сверху и снизу – соответственно 7 Гц и 9 Гц. При этом на выходе ПНВ установлен фотоприемник для регистрации сигналов с этими частотами. Фотоприемник подключен к компаратору, который через регистр управления связан с приводом перемещения оси прожектора по горизонту и по вертикали. Предположим, что объект наблюдения сместился вправо от пятна подсвета, формируемого основным источником подсвета. Тогда объект попадет в пятно подсвета, формируемого вспомогательным источником с рабочей частотой 5 Гц. Фотоприемник воспримет этот сигнал с экрана электронно-оптического преобразователя и преобразует его в электрический сигнал, который поступит в компаратор. Последний выработает разностный сигнал по отношению к частоте основного источника подсвета и будет через регистр управлять работой привода в направлении устранения сигнала с частотой 5 Гц. Это произойдет тогда, когда объект снова окажется в пределах основного пятна подсвета. Аналогичным образом осуществляется управление приводом при попадании объекта в другие вспомогательные пятна подсвета. Таким образом, привод осуществляет слежение за объектом наблюдения, обеспечивая его постоянное положение в пределах основного пятна подсвета.

2. ОПИСАНИЕ ПРОЦЕССА ПЕРЕДАЧИ И ПРИЕМА ИЗОБРАЖЕНИЯ В ЛАЗЕРНЫХ СИСТЕМАХ ВИДЕНИЯ

2.1. Понятие оптической передаточной функции

Любой удаленный объект в определенном пространственном сечении можно охарактеризовать *двумерным распределением яркости в плоскости объекта* $L(x_1, y_1)$, а изображение этого объекта – *распределением освещенности в плоскости анализа* $E(x_2, y_2)$. Поведение изображающей системы описывается математическим оператором P , который показывает, как нужно воздействовать на функцию $L(x_1, y_1)$, чтобы на выходе системы получить функцию $E(x_2, y_2)$. Такое воздействие изображающей системы при выполнении определенных требований (линейность, временная и пространственная инвариантность), можно описать с помощью передаточных функций. Эти функции должны быть достаточно универсальны, чтобы характеризовать не только подсистемы различной физической природы, входящие в состав лазерной системы видения, но и оценивать степень влияния на качество изображения внешних факторов (оптических характеристик канала распространения, движения объекта и т. п.).

Двумерное распределение яркости на объекте однозначно связано со спектром яркостного сигнала (в области пространственных частот) посредством двумерного преобразования Фурье [10]

$$\tilde{L}(v_x, v_y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} L(x, y) \exp(-i2\pi(v_x x + v_y y)) dx dy,$$

где v_x, v_y – пространственные частоты.

Необходимой предпосылкой использования передаточных функций является линейность изображающих систем. Пусть $P\{\}$ – линейный оператор, преобразующий функцию $f(x)$ в другую функцию $g(\xi)$

$$P\{f(x)\} = g(\xi). \quad (2.1)$$

Обозначим функциями $g_1(\xi)$ и $g_2(\xi)$ реакции (отклики) системы на соответствующие входные воздействия $f_1(x)$ и $f_2(x)$. Для линейной системы реакция на сумму входных воздействий равна сумме реакций на отдельные входные воздействия, т.е.

$$P\{a_1 f_1(x) + a_2 f_2(x)\} = a_1 g_1(\xi) + a_2 g_2(\xi),$$

где a_1 и a_2 – произвольные (в общем случае комплексные) постоянные.

Применительно к изображающей системе это означает, что освещенность изображения, получаемая от суммарного воздействия нескольких яркостных элементов объекта, равна сумме освещенностей, которые создаются каждым яркостным элементом в отдельности. Для линейной системы ее отклик на произвольный входной сигнал можно выразить через отклик на определенные элементарные входные воздействия, на которые следует разложить входной сигнал. Такое разложение можно получить, используя фильтрующее свойство δ -функции, которое позволяет представить входной сигнал $L(x_0, y_0)$, характеризующий двумерное распределение яркости в плоскости объекта

$$L(x_0, y_0) = \int \int_{-\infty}^{\infty} L(\xi, \eta) \delta(x_0 - \xi) \delta(y_0 - \eta) d\xi d\eta. \quad (2.2)$$

Смысл выражения (2.2) состоит в том, что любую функцию можно представить как бесконечную сумму взвешенных и смещенных δ -функций. С оптико-физической точки зрения формулу (2.2) можно рассматривать как непрерывную двумерную сумму точечных источников $\delta(x_0 - \xi)\delta(y_0 - \eta)$ с амплитудой $L(\xi, \eta)$, локализованных в фиксированных точках (x_0, y_0) плоскости объекта (ξ, η) .

Реакцию системы на входной сигнал $L(x_0, y_0)$, т. е. распределение освещенности в плоскости изображения $(x_{\text{и}}, y_{\text{и}})$, получим, подставив (2.2) в (2.1) с учетом свойства линейности:

$$E(x_{\text{и}}, y_{\text{и}}) = \int \int_{-\infty}^{\infty} L(\xi, \eta) P\{\delta(x_0 - \xi) \delta(y_0 - \eta)\} d\xi d\eta. \quad (2.3)$$

Обозначим реакцию (отклик) системы на входной типовой δ -сигнал, т. е. на воздействие точечного источника излучения в виде

$$P\{\delta(x_0 - \xi, y_0 - \eta)\} = H(x_{\text{и}}, y_{\text{и}}; x_0, y_0). \quad (2.4)$$

Функция $H(x_{\text{и}}, y_{\text{и}}; x_0, y_0)$ характеризует дифракционно-абerrационное изображение точечного источника, локализованного в точке (x_0, y_0) предметной плоскости (ξ, η) , и называется *функцией рассеяния* (ФР) или *функцией рассеяния точки* (ФРТ).

Функция рассеяния точки, которая лежит на оси идеальной оптической системы, представляет собой симметричную относительно данной оси функцию (рис. 18, а) и состоит из центрального максимума (на изображении от точки – диска, рис. 18, б) и вторичных максимумов (на изображении – кольца, рис. 18, в) [4].

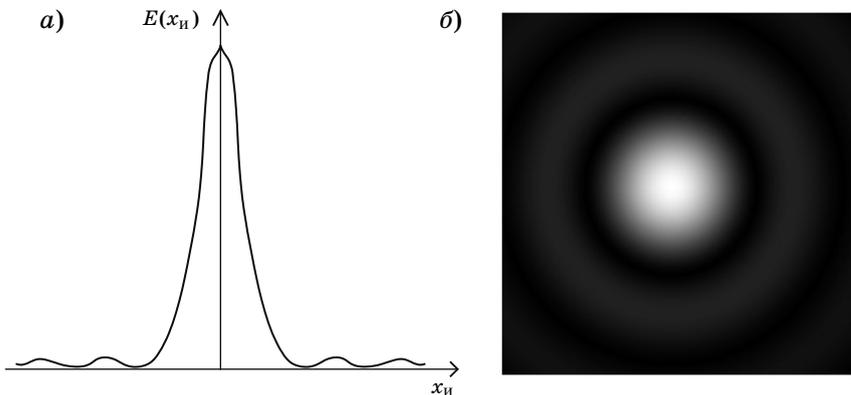


Рис. 18. Вид функции рассеяния точки в сечении (а) и визуализация ее изображения (б)

Если изображающая система имеет линейное увеличение β , то ее ФР обозначается $H_{\beta}(x_{и}, y_{и}; x_0, y_0)$. Выражение (2.3) называется интегралом суперпозиции, так как выходной сигнал $E(x_{и}, y_{и})$ представляется в виде бесконечной суммы функций рассеяния, характеризующих реакцию системы на воздействия точечных источников, которые расположены в точках (x_0, y_0) предметной плоскости. В этой связи целесообразно переменные интегрирования в (2.3) обозначать теперь через x_0 и y_0

$$E(x_{и}, y_{и}) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} L(x_0, y_0) H(x_{и}, y_{и}; x_0, y_0) dx_0 dy_0. \quad (2.5)$$

Геометрическая интерпретация интеграла суперпозиции представлена на рис. 19. Распределение освещенности в плоскости изображения находят при разложении одномерного распределения яркости объекта на взвешенные δ -функции, заменой каждой из них взвешенной функцией рассеяния и последующим суммированием.

Особый интерес представляют **пространственно-инвариантные изображающие системы**, у которых изображение точечного источника при перемещении его в предметной плоскости изменяет только свое пространственное положение, но не форму, т.е.

$$H(x_{и}, y_{и}; x_0, y_0) = H(x_{и} - x_0; y_{и} - y_0),$$

причем разность координат рассматривается в плоскости изображения $(x_{и}, y_{и})$.

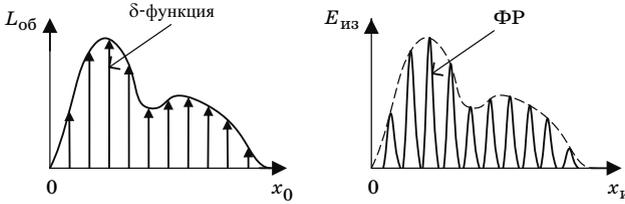


Рис. 19. Представление яркостного поля объекта и его изображения в виде последовательности δ -функций и в виде суперпозиции отдельных ФР, соответственно

Для пространственно-инвариантной изображающей системы интеграл суперпозиции (2.5) преобразуется в интеграл свертки

$$E(x_{из}, y_{из}) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} L(x_0, y_0) H(x_{из} - x_0; y_{из} - y_0) dx_0 dy_0. \quad (2.6)$$

Выражение (2.6) обычно записывается в символической форме:

$$E(x_{из}, y_{из}) = L \otimes H(x_{из}, y_{из}). \quad (2.7)$$

Выполнив преобразование Фурье обеих частей соотношения (2.7) и применив теорему свертки, получим выражение, связывающее спектры входного $\tilde{L}(v_x, v_y)$ и выходного $\tilde{E}(v_x, v_y)$ сигналов

$$\tilde{E}(v_x, v_y) = \tilde{H}(v_x, v_y) \tilde{L}(v_x, v_y). \quad (2.8)$$

Фурье-образ нормированной ФР при некогерентном освещении называют *оптической передаточной функцией* (ОПФ) изображающей системы:

$$\tilde{H}(x_{из}, y_{из}) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} H(x_{из}, y_{из}) \exp[-i2\pi(v_x x_{из} + v_y y_{из})] dx_{из} dy_{из}. \quad (2.9)$$

В общем случае ОПФ представляет собой комплексную функцию

$$\tilde{H}(v_x, v_y) = \tilde{T}(v_x, v_y) \exp(i\psi(v_x, v_y)), \quad \tilde{T}(v_x, v_y) = |\tilde{H}(v_x, v_y)|,$$

$$\psi(v_x, v_y) = \arg[\tilde{H}(v_x, v_y)].$$

Модуль ОПФ $\tilde{T}(v_x, v_y)$ называется *модуляционной передаточной функцией* (МПФ) или функцией передачи модуляции. Аргу-

мент ОПФ $\psi(v_x, v_y)$ называется *фазо-частотной характеристикой* (ФЧХ), или функцией передачи фазы.

Модуляционная передаточная функция играет ключевую роль при расчете, проектировании и оптимизации ЛСВ, при этом ЛСВ представляется в виде составной (многоступенчатой) изображающей системы, и системная МПФ всей системы (системная МПФ) определяется в виде произведения отдельных подсистем.

Системная МПФ показывает, насколько качественно, без искажений система воспроизводит изображение объекта. Наивысшая пространственная частота v_{\max} , воспроизводимая лазерной системой видения без искажения для оператора, называется максимальной частотой пропускания, или частотой среза $v_{\text{ср}}$.

Так, при передаче изображения считается, что *частота среза* – это такая частота, на которой значение МПФ достигает уровня 0,1 [1]. Указанный уровень соответствует пороговому контрасту для оператора, поскольку при значениях МПФ ниже данного уровня он уже не может выделить объект из окружающего фона и обнаружить его изображение на экране монитора.

Представление ЛСВ в виде составной многоступенчатой системы позволяет определить МПФ всей системы в виде произведения отдельных подсистем

$$T_{\text{сист}}(v_x, v_y) = T_{\text{ср}}(v_x, v_y) T_{\text{опт}}(v_x, v_y) T_{\text{фд}}(v_x, v_y) T_{\text{ов}}(v_x, v_y) T_{\text{в}}(v_x, v_y) T_{\text{гл}}(v_x, v_y),$$

где $T_{\text{опт}}(v_x, v_y)$ – МПФ приемопередающей оптической системы; $T_{\text{фд}}(v_x, v_y)$ – МПФ фотодетектора; $T_{\text{ов}}(v_x, v_y)$ – МПФ электронного тракта для формирования и обработки видеосигнала; $T_{\text{в}}(v_x, v_y)$ – МПФ видеомонитора; $T_{\text{гл}}(v_x, v_y)$ – МПФ зрительного анализатора оператора.

2.2. Передача модуляционного контраста изображающей системой

По определению МПФ – это отношение глубин модуляции выходного и входного сигналов синусоидальной формы, нормированное к единице на нулевой пространственной частоте. Для понимания смысла модуля ОПФ и аргумента комплексной ОПФ рассмотрим объект с косинусоидальным распределением коэффициента отражения относительно среднего значения $\bar{\rho}$:

$$\rho(x) = \bar{\rho} + \Delta\rho \cos(2\pi\nu_x t),$$

и определим, как изображается отдельная частотная составляющая изображающая система с известной функцией рассеяния. Распределение коэффициента отражения в изображении этого объекта согласно [1] будет иметь вид

$$\rho'(x) = \bar{\rho} + \Delta\rho T(\nu_x) \cos(2\pi\nu_x x - \psi(\nu_x)). \quad (2.10)$$

Из (2.10) следует, что изображение отличается от объекта двумя особенностями. Во-первых, глубина модуляции, равная отношению амплитуды переменной составляющей к среднему значению (к постоянной составляющей), оказывается меньше, чем на объекте, поскольку

$$m = \frac{\Delta\rho}{\bar{\rho}} \text{ (глубина модуляции объекта),}$$

$$m' = \frac{\Delta\rho T(\nu)}{\bar{\rho}} \text{ (глубина модуляции в изображении объекта).}$$

Контраст объекта для косинусоидального распределения яркости

$$k_{об} = \frac{(\bar{L} + \Delta L) - (\bar{L} - \Delta L)}{(\bar{L} + \Delta L) + (\bar{L} - \Delta L)} = \frac{\Delta L}{\bar{L}} = \frac{\Delta\rho}{\bar{\rho}} = m.$$

Контраст изображения

$$k_{из} = \frac{\Delta L T(\nu)}{\bar{L}} = \frac{\Delta\rho T(\nu)}{\bar{\rho}} = m',$$

а МПФ по определению $T(\nu_x) = \frac{m'}{m} = \frac{k_{из}}{k_{об}}$.

Таким образом, значение модуля ОПФ для каждой пространственной частоты, равное отношению глубин модуляции, называется коэффициентом передачи модуляции. Их совокупность составляет функцию передачи модуляции или МПФ. Вторая особенность состоит в сдвиге косинусоиды в угловой мере, проявляющемся при абберации оптических систем, дефектах юстировки. В системах, где данные негативные явления сведены к минимуму, можно считать, что значение ФЧХ равно 0 или π .

Однако определение контраста подобным образом приводит к тому, что его значение не совпадает со значением воспринимаемо-

го контраста. Наиболее отчетливо эти расхождения проявляются в электронных изображающих системах, где изображение формируется на экране монитора и анализируется оператором.

Пусть изображение содержит $m \times n$ элементов (пикселей), а распределение яркости описывается двумерной дискретной функцией $L(x, y) = L_{mn}$, тогда *средний уровень яркости в изображении*

$$\bar{L} = \frac{1}{mn} \sum_m \sum_n L_{mn}.$$

Воспринимаемый контраст в изображении характеризуется отклонением яркостного уровня его элементов от среднего

$$k_{\text{воспр}} = \frac{\sqrt{\frac{1}{mn} \sum_m \sum_n (L_{mn} - \bar{L})^2}}{\bar{L}}.$$

Пусть диапазон яркости квантуется по уровням от 0 (черный) до 255 (белый). Если один из элементов изображения будет иметь уровень 255, другой – 0, а все остальные – 128, то модуляционный контраст по (1.1) равен 1, а наблюдатель видит равномерный фон нулевого контраста с двумя шумовыми пикселями.

Определим, каким образом вид МПФ лазерной системы видения искажает исходное двумерное распределение яркости объекта в процессе получения изображения при заданной МПФ. Рассмотрим для примера случай наблюдения объекта прямоугольной формы, который описывается равномерным двумерным распределением яркости, равной единице на поверхности объекта, на фоне с нулевой яркостью (рис. 20, а).

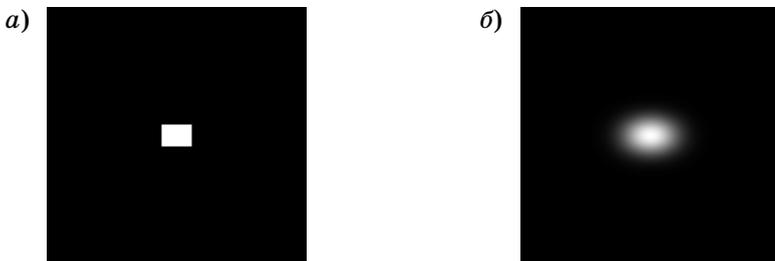


Рис. 20. Исходное изображение (а) и изображение (б) полученное на выходе ЛСВ

$$L(x, y) = \begin{cases} 1 & |x| \leq a/2; |y| \leq b/2 \\ 0 & |x| > a/2; |y| > b/2 \end{cases}$$

где a – длина стороны прямоугольного объекта вдоль оси x , а b – вдоль оси y . Пространственно-частотный спектр в этом случае равен

$$\begin{aligned} \tilde{L}(v_x, v_y) &= \int_{-a/2}^{a/2} \int_{-b/2}^{b/2} \exp(-i2\pi(v_x x + v_y y)) dx dy = \\ &= \int_{-a/2}^{a/2} \exp(-i2\pi v_x x) dx \int_{-b/2}^{b/2} \exp(-i2\pi v_y y) dy. \end{aligned}$$

С учетом того, что каждый интеграл преобразуется к виду

$$\int_{-a/2}^{a/2} \exp(-i2\pi v_x x) dx = a \frac{\sin\left(2\pi v_x \frac{a}{2}\right)}{2\pi v_x \frac{a}{2}},$$

получаем выражение для вычисления модуля пространственно-частотного спектра

$$\begin{aligned} |\tilde{L}(v_x, v_y)| &= ab \left| \frac{\sin\left(2\pi v_x \frac{a}{2}\right) \sin\left(2\pi v_y \frac{b}{2}\right)}{2\pi v_x \frac{a}{2} \cdot 2\pi v_y \frac{b}{2}} \right| = \\ &= ab \left| \operatorname{sinc}\left(2\pi v_x \frac{a}{2}\right) \operatorname{sinc}\left(2\pi v_y \frac{b}{2}\right) \right|. \end{aligned}$$

Предположим, что данный спектр проходит через МПФ ЛСВ, которая имеет форму фильтра Гаусса [11].

$$|\tilde{L}_\Phi(v_x, v_y)| = \exp\left[-(2\pi v_x)^2 a/2\right] \exp\left[-(2\pi v_y)^2 b/2\right].$$

Произведение полученного спектра сигнала на спектр фильтра Гаусса в частотной области приведет к формированию изображения (рис. 20, б), которое будет представлять свертку исходного распределения яркости со следующей функцией

$$S(x, y) = \exp\left(-x^2 / a^2\right) \exp\left(-y^2 / b^2\right).$$

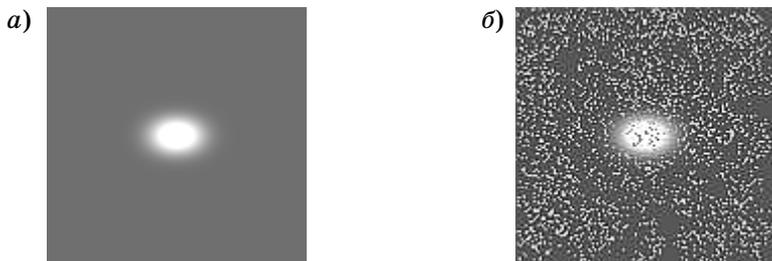


Рис. 21. Изображение (а), полученное на выходе ЛСВ при наличии солнечной засветки, изображение (б) – при добавлении шума в приемном тракте ЛСВ

Вследствие наличия помехи обратного рассеяния от солнечной или лунной засветки наблюдаемый контраст фона и объекта меняется, что приводит к дополнительному искажению изображения, полученного на рис 20, б. На рис. 21, а представлено влияние наличия возможной сторонней засветки изображения.

Чтобы снизить уровень засветки на изображении, применяют импульсное лазерное излучение малой длительности менее 10 нс, а приемник излучения не открывают до тех пор, пока не придет отраженный сигнал от просматриваемого участка дальности.

Кроме того, на регистрируемое изображение накладывается шум, возникающий в электронном тракте ЛСВ. Влияние шума дополнительно искажает получаемое при помощи ЛСВ изображение (рис. 21, б). Причины возникновения шума и способы борьбы с ним рассматриваются в следующей главе.

Шум на экране индикатора может существенно снизить вероятность обнаружения и идентификации объекта в том случае, когда размеры объекта на изображении или детали изображения, по которым производится его идентификация, обладают малыми размерами и являются малоконтрастными по отношению к фону.

2.3. Модуляционная передаточная функция атмосферы и морской воды

Подробное описание процесса распространения лазерного излучения в атмосфере или под морской поверхностью можно найти в [1,7]. Остановимся на основных соотношениях, позволяющих рассчитать только МПФ указанных сред при различных внешних условиях.

Атмосфера Земли представляет собой среду, состоящую из смеси газов и взвешенных в воздухе твердых и жидких частиц, называемых аэрозолями. Общее ослабление лазерного излучения в атмосфере обусловлено молекулярным и аэрозольным поглощением и рассеянием.

Наиболее сильно излучение поглощается парами воды, углекислым газом и озоном. Молекулярное поглощение имеет ярко выраженный селективный характер и проявляется в виде полос поглощения, разделенных окнами пропускания, где поглощение либо вообще отсутствует, либо является весьма малым. Такие участки спектра называют «окнами прозрачности» атмосферы, которые для лазерного излучения лежат в диапазонах 0,4...0,75; 0,95...1,06; 1,2...1,3; 1,5...1,8; 2,1...2,4; 3,3...4,0; 8...9,3; 10...12,5 мкм [12].

С достаточной для практики точностью процесс распространения лазерных пучков в аэрозольной атмосфере можно описывать, располагая всего тремя оптическими характеристиками: *показателем ослабления ϵ* , *показателем рассеяния σ* и *индикатрисой рассеяния $x(\gamma)$* , где γ – угол, отсчитываемый от направления распространения падающей волны. Для оценки помехи обратного рассеяния наиболее удобной характеристикой является лидарное отношение, которое определяет в относительных единицах долю мощности излучения рассеянного единицей объема среды в направлении «назад». Для индикатрисы рассеяния, нормированной в телесном угле 4π условием:

$$\int x(\gamma) \frac{d\omega}{4\pi} = 1, \text{ лидарное отношение } b = \frac{x(\pi)}{4\pi}.$$

В приземном слое атмосферы выделяют несколько качественно различных типов оптического состояния атмосферы: дымка, для которой типична метеорологическая дальность видимости $S_M > 3$ км, туманная дымка с характерными значениями $1 \text{ км} < S_M < 3$ км, туманы $S_M < 1$ км. Состояния дымки и туманной дымки, в которых приземный слой находится приблизительно 90% времени, являются наиболее естественной ситуацией для применения лазерных устройств. В континентальной зоне умеренных широт для дымки и туманной дымки в диапазоне длин волн 0,55... 10 мкм расчет аэрозольного показателя ослабления (в «окнах прозрачности» атмосферы) проводится по формуле [1]:

$$\epsilon(\lambda) = \epsilon(0,55) \left[n_0 + n_1 \lambda^{-n_2} \right]; \quad \epsilon(0,55) = 3,91 / S_M,$$

где λ подставляется в мкм, S_M – в км.

Эмпирические коэффициенты n_0 , n_1 , n_2 для различных типов дымок приведены в табл. 2.

Таблица 2

Период	Тип погоды	n_0	n_1	n_2
Летний	Дымка после сильного дождя	0	0,4	1,88
	Устойчивая дымка $S_M > 4$ км	0,06	0,36	1,88
Весенне-осенний	Дымка	0,1	0,45	1,3
	Туманная дымка	0,01	0,8	0,5
	Дымка с моросью	0,3	0,5	0,6
Зимний	«Ледяная» дымка	0,248	0,447	1,24
	Зимняя дымка	0	0,58	1,24
	Дымка со снегом	0,77	0,145	1,24

Для определения показателя поглощения и рассеяния используются следующие соотношения [1]:

$$\sigma(\lambda) = \frac{\varepsilon(\lambda)}{1 + \beta(\lambda)}; \quad k_{II}(\lambda) = \frac{\varepsilon(\lambda)\beta(\lambda)}{1 + \beta(\lambda)}, \quad \text{где } \beta(\lambda) = 0,2 \left(\frac{\lambda}{0,55 \cdot 10^{-6}} \right)^{0,8}.$$

Индикатрисса рассеяния определяется по формуле

$$\chi(\gamma) = \frac{2}{\mu^2} \exp\left(-\frac{\gamma}{\mu}\right),$$

и в условиях атмосферной дымки для наиболее вероятного значения параметра $\mu = 0,3$ в обратном направлении ($\gamma = \pi$) определяется по формуле

$$x_\pi = 0,33\sigma^{-0,31},$$

где σ определена в км^{-1} . При наличии осадков имеют место несколько другие соотношения, поскольку в отличие от туманов, частицы в осадках имеют более крупные размеры (100 мкм и более) и более вытянутую по направлению изучения индикатриссу рассеяния. При оценках эффектов рассеяния параметр индикатриссы рассеяния может быть принят $\mu = 0,06$, а показатель ослабления $\varepsilon = \sigma$ (вне полос поглощения воды). Зависимость показателя ослабления в зависимости от интенсивности дождя имеет вид

$$\varepsilon = 0,21J^{0,74}.$$

Модуляционная передаточная функция в условиях атмосферных дымок и осадков определяется с учетом указанных параметров по формуле

$$T_{\text{атм}}(\nu) = \exp \left(-\Lambda \tau + \frac{\Lambda \tau}{\sqrt{(2\pi\mu z \nu)^2 + 1}} \right),$$

где $\Lambda = \sigma(\lambda)/\varepsilon(\lambda)$ – вероятность выживания фотона, оптическая толщина слоя среды $\tau = \varepsilon(\lambda)z$ (z – длина трассы в км).

Лазерное излучение, распространяющееся в морской среде, ослабевает в результате рассеяния и поглощения. Рассеяние обусловлено взвешенными в воде частицами и зависит от их состава. Показатель рассеяния (σ) слабо зависит от длины волны.

Поглощение в море зависит от характеристик воды и содержания в ней органических веществ (фитопланктона) [7]. Область наименьшего ослабления в воде соответствует минимуму показателя поглощения $k_{\text{п. min}}$. В различных районах мирового океана окна прозрачности немного отличаются, но всегда лежат в пределах спектрального диапазона $\lambda = 470\text{-}570$ нм.

Показатель ослабления $\varepsilon = \sigma + k_{\text{п}}$ измерить проще, чем показатели поглощения и рассеяния по отдельности. Наиболее легко измеряемая гидрооптическая характеристика – **прозрачность воды по белому диску**, которая представляет собой глубину видимости белого стандартного 30 сантиметрового диска $z_{\text{б}}$, результаты измерения которой имеются для многих океанов или прибрежных вод. Значение $z_{\text{б}} < 10$ метров соответствует водоемам с плохой видимостью, со значением $z_{\text{б}} > 40$ – прозрачным водоемам с наилучшей видимостью. Например, значение $z_{\text{б}}$ для Черного моря колеблется в диапазоне от 5 до 28 метров, для Финского залива от 6 до 12 метров, а в Бенгальском заливе составляет от 40 до 45 метров [1].

Показатель ослабления (ε) и вероятность выживания фотона в воде (Λ) и показатель рассеяния (σ) связаны с прозрачностью воды по белому диску соотношениями

$$\varepsilon = \frac{5}{z_{\text{б}}}; \quad \Lambda = 0,955 - 0,035 / \varepsilon; \quad \sigma = \Lambda \varepsilon.$$

МПФ рассеивающей морской среды можно определить по соотношению

$$T(\nu') = \exp(-\Lambda \varepsilon z),$$

где ν' – угловая пространственная частота, а z – дальность до объекта, которые связаны с пространственной частотой соотношением $\nu = \nu'/z$.

2.4. Модуляционная передаточная функция приемопередающей оптической системы и фотодетектора

Общая МПФ оптического тракта ЛСВ зависит от соотношения диаграмм направленности передающего и приемного каналов. Модуляционная передаточная функция оптической приемопередающей системы однозначно определяется МПФ звена с узкой диаграммой направленности. Лазерная система видения чаще всего выполняется с импульсной подсветкой (широкая диаграмма направленности источника) и фотоприемным устройством на основе ПЗС – матрицы (ПЗС – прибор с зарядовой связью) с узкой диаграммой приемного объектива.

Модуляционная передаточная функция объектива – системы с круглым входным зрачком (идентичная как для горизонтальной и вертикальной плоскостей изображения) имеет вид [1]:

$$T_{\text{опт}}(\delta) = \frac{2}{\pi} \left(\arccos \delta - \delta \sqrt{1 - \delta^2} \right) \text{ при условии } \delta \leq 1;$$
$$T_{\text{опт}}(\delta) = 0 \text{ при условии } \delta > 1,$$

где δ – относительная пространственная частота, равная отношению пространственной частоты ν к её предельному значению $\nu_{\text{ср}}$, еще пропускаемому оптической системой.

Эта максимальная пространственная частота зависит от диаметра объектива D и фокусного расстояния оптической системы $f_{\text{об}}$: $\nu_{\text{ср}} = D/(\lambda f_{\text{об}})$. Для угловых пространственных частот МПФ такой оптической системы вычисляется по той же формуле, а частота среза в этом случае равна $\nu'_{\text{ср}} = D/\lambda$.

Поиск габаритных параметров объектива можно осуществить, используя для общей методики расчета диаметра объектива, которая подробно рассмотрена в [13], параметры, определяющие качество работы лазерной системы видения.

Вначале вычисляем минимальный угловой размер, который должен разрешить объектив приемной оптической системы

$$\Delta\alpha_{\text{min}} = H_{\text{кр.об}}/(Nz_{\text{max}}),$$

где z_{max} – предельная дальность видения ЛСВ; $H_{\text{кр.об}}$ – критический (минимальный) размер объекта, подлежащего обнаружению на заданном расстоянии; $N \geq 1$ – положительный множитель, соответствующий числу эквивалентной штриховой миры, который выбирается, исходя из критериев Джонсона, которые рассмотрены подробно в завершении данного раздела учебного пособия. Отметим предварительно, что процесс принятия решения оператором

ТВ-системы включает в себя четыре взаимосвязанных этапа: обнаружение, классификацию (и определение ориентации), различение и идентификацию (опознавание). Под **обнаружением** подразумевается выделение объекта на фоне и отнесение его к классам объектов, представляющим потенциальный интерес. **Классификация** означает отнесение обнаруженного объекта к одному из широких классов (человек, транспортное средство) и зачастую его ракурс по отношению к наблюдателю (определение ориентации). **Различение** означает отнесение наблюдаемого объекта к более узкому подклассу (грузовой автомобиль, легковой автомобиль). Наконец, на стадии **идентификации** можно установить тип объекта (марка автомобиля). Таким образом, Джонсон ввел для данных этапов четыре соответствующих значения множителя N .

После вычисления минимального углового размера, определяем предел порогового углового разрешения объектива как

$$\Delta\alpha_{\text{пор}} = 2,44\lambda k_{\text{аб}}/D,$$

где λ – граничная (наибольшая) длина волны в рабочем спектральном диапазоне; $k_{\text{аб}}$ – коэффициент, учитывающий наличие aberrаций в объективе ($k_{\text{аб}} > 1$). Малым aberrациям соответствуют коэффициенты 1–1,25, для расчета средних или больших aberrаций дешевых объективов можно выбрать значения 3 и 5 соответственно [14].

Учитывая, что требуется выполнять неравенство $\Delta\alpha_{\text{мин}} \geq \Delta\alpha_{\text{пор}}$, получаем для вычисления диаметра объектива приемной оптической системы ЛСВ следующее выражение:

$$D \geq 2,44 \lambda k_{\text{аб}} z_{\text{max}} N / H_{\text{кр.об.}}$$

Например, для различения объекта ($N = 3$) на границе зоны обзора по дальности, т. е. на максимальной дальности, равной $z_{\text{max}} = 2$ км с $H_{\text{кр.об.}} = 0,5$ м, при использовании неодимового лазера с $\lambda = 1,06$ мкм [3] и объектива со средними хроматическими aberrациями ($k_{\text{аб}} = 3$) получаем минимальный диаметр объектива, равный $D_{\text{мин}} = 10$ см.

Одним из наиболее важных элементов приемного канала ЛСВ с импульсным подсветом является матричный фоточувствительный детектор оптического излучения на основе приборов с зарядовой связью – фотоприемная ПЗС матрица (ФПЗС-матрица). Входным оптическим сигналом для ФПЗС – матрицы является непрерывное распределение освещенности на поверхности её чувствительных элементов $E_{\text{из}}(x, y, z)$, сформированное предшествующими компонентами оптической системы, а выходным – видеосигнал $u(t)$.

Существует взаимосвязь между размерами ФПЗС-матрицы ($m_x \times m_y$), критериями Джонсона (N) и фокусным расстоянием объектива f_0 . Определим в качестве примера фокусное расстояние объектива для рассмотренного случая различения объекта.

Отметим, что ФПЗС-матрицы различаются по формату, представляющему собой округленное значение длины диагонали матрицы, выраженное в дюймах. Так матрица формата 1/3" дюйма имеет размеры: $m_x \times m_y = 4,8 \times 3,6$ мм, а матрицы форматов 1/2" – $m_x \times m_y = 6,4 \times 4,8$ мм, 1/4" – $m_x \times m_y = 3,6 \times 2,7$ мм.

Ширина (W_x) и высота (H_y) наблюдаемого участка пространства на дальней границе зоны обзора ЛСВ, т. е. в плоскости объекта (удаленного на расстояние z_{\max}) связаны с размерами матрицы соотношениями и фокусным расстоянием объектива:

$$W_x = m_x z_{\max} / f_0; H_y = m_y z_{\max} / f_0.$$

Тогда при заданном критическом размере объекта и известном количестве линий, для его отображения на экране дисплея оператора, равном N , получаем следующую пропорцию:

$$W_x / H_{\text{кр.об}} = N_{\text{дисп}} / N,$$

где $N_{\text{дисп}}$ – общее число телевизионных столбцов изображения на высоту или строк на ширину экрана, в зависимости от того, минимальный критический размер объекта задан по высоте или по горизонтальной протяженности. Если заранее неизвестен наиболее вероятный ракурс обнаруживаемого объекта, то в качестве величины $N_{\text{дисп}}$ можно использовать число строк на экране дисплея.

Исходя из представленных соотношений, фокусное расстояние может быть найдено из известных параметров телевизионной системы для отображения объекта по следующему соотношению:

$$f_0 = m_x z_{\max} N / (N_{\text{дисп}} H_{\text{кр.об}}).$$

Подставляя использовавшиеся в предыдущем примере значения, а также то, что для отображения информации в системах охранного телевидения принято использовать от 380 до 420 линий, получаем при $N_{\text{дисп}} = 320$ и матрице формата 1/3" дюйма, что $f_{0\max} = 180$ мм. Для идентификации объекта ($N = 7$) на данном расстоянии требуется использовать объектив с фокусным расстоянием $f_0 = 420$ мм.

Следует отметить, что необходимость применять в ЛСВ оптических систем с длиннофокусными объективами (120 мм и более),

предназначенными для наблюдения удаленных или протяженных объектов приводит к потере разрешающей способности. Эти потери обусловлены действием нескольких факторов. Во-первых, кружок рассеяния объектива возрастает с увеличением фокусного расстояния. Во-вторых, при наблюдении на больших расстояниях, оказывает заметное влияние турбулентность воздуха, особенно если недалеко от камеры вдоль оси ее зрения имеются открытые окна теплых помещений, трубы отопительной системы, работающие моторы механизмов или другие теплые объекты. В результате возникновения сильных потоков воздуха, происходит размывание и дрожание мелких деталей изображения, что приводит к потере разрешающей способности. Кроме того, при наблюдении на дальних дистанциях, даже незначительные атмосферные осадки и туман вызывают заметное светорассеяние и потерю четкости и контраста изображения. Еще одной причиной ухудшения разрешающей способности становится естественное загрязнение стекол объективов и иллюминаторов наружных телекамер в процессе работы. Одновременно с потерей разрешающей способности, в этом случае возможно появление пятен и полос на изображении.

Рассмотрим далее, каким образом формируется первичное «электронное изображение» объекта. Преобразование $E_{из}(x, y, z) \rightarrow u(t)$ осуществляется путем коммутации сигналов отдельных пикселей ФПЗС-матрицы с установленными траекторией считывания изображения и скоростью. Как дискретный датчик ФПЗС осуществляет двумерную пространственную и одномерную временную выборку отсчетов. Из-за того, что непрерывное распределение освещенности преобразуется в дискретный сигнал распределения зарядов, изображающая система с выборкой оказывается *пространственно неинвариантной*, т. е. изображение зависит от его положения относительно двумерной решетки выборки (матрицы), как показано на рис. 22.

Схематическое изображение раstra ФПЗС-матрицы в виде прямоугольных элементов (пикселей) размерами $a_x \times a_y$, с расстоянием между их центрами d_x и d_y показано на рис. 23. Количество пикселей по горизонтали и вертикали $m \times n$ зависит от формата матрицы одинаковых элементов.

Процесс формирования выходного сигнала ФПЗС-матрицы можно представить в виде следующей цепочки преобразований. Вначале происходит преобразование непрерывного распределения освещенности в дискретные заряды отсчетов, пространственно-частотная фильтрация и по завершении экспозиции изображения на

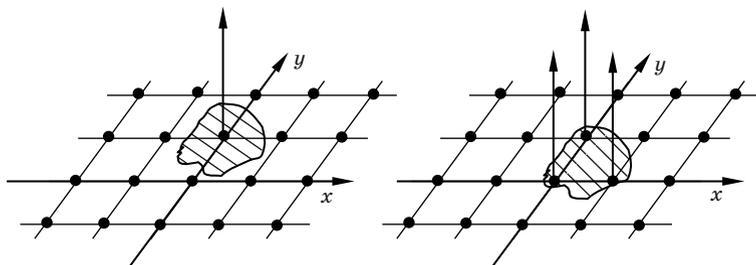


Рис. 22. Выборка двумерного сигнала

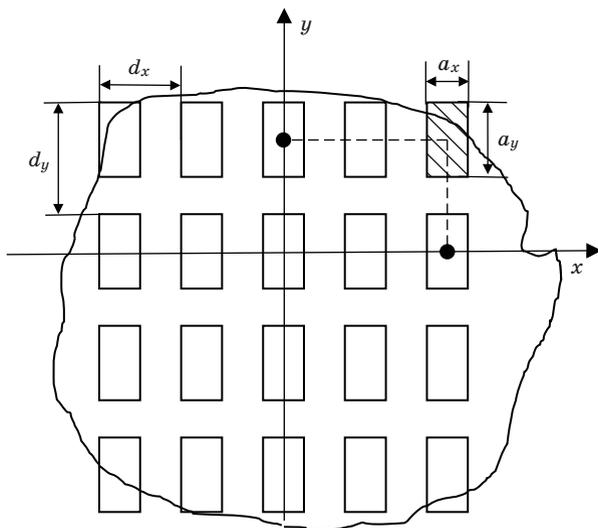


Рис. 23. Часть ФПЗС-матрицы с распределением освещенности от объекта

матрице реализуется считывание зарядов – **фотодетектирование**. При переносе зарядовых пакетов к выходному регистру и их преобразовании в отсчеты напряжения формируется дискретный видеосигнал, содержащий нелинейные искажения, обусловленные выборкой. После прохождения через активный НЧ-фильтр дискретные сигналы напряжения преобразуются в непрерывный аналоговый видеосигнал.

Для восстановления неискаженного исходного изображения по совокупности передаваемых отсчетов видеосигнала, формируемых при считывании зарядовых пакетов и следующих с интервалами

дискретизации в продольном d_x и в поперечном d_y направлениях, в соответствии с теоремой отсчетов необходимо выполнить два условия. Ширина спектра передаваемых частот должна быть меньше, чем частота дискретизации, т. е.:

$$v_x \leq \frac{1}{2} \frac{1}{d_x} = \frac{1}{2} v_{sx}; \quad v_y \leq \frac{1}{2} \frac{1}{d_y} = \frac{1}{2} v_{sy},$$

где v_{sx}, v_{sy} – частоты дискретизации по двум осям.

Во-вторых выходной сигнал для ограничения влияния искажений изображения, возникающего вследствие инвариантности его считывания и дрожания относительно матрицы во время экспозиции, пропустить полученный результат через двумерный фильтр низких частот с прямоугольной частотной характеристикой и частотами среза $0,5v_{sx}; 0,5v_{sy}$.

В случае невыполнения данных преобразований пространственные частоты в спектре изображения превышают частоты среза, вследствие чего возникает наложение компонентов частот побочных спектров дискретизации на компоненты основного спектра. Это приводит к тому что на изображении появляются посторонние узоры (муар). При указанных ограничения передаточную функцию фотодетектора определяют как

$$T_{\text{фд}}(v_x, v_y) = \text{sinc}(a_x v_x) \text{sinc}(a_y v_y),$$

где функция $\text{sinc}(x)$ является функцией $\sin(x)/x$.

2.5. Модуляционная передаточная функция электронно-оптического преобразователя и репродукционного объектива

В условиях крайне низкой освещенности (например, под водой или в космосе) ЛСВ содержат в приемном канале усилитель яркости изображения на основе электронно-оптического преобразователя (ЭОП), расположенного до ФПЗС-матричного фотоприемника и сопряженного с ней при помощи репродукционного объектива.

Электронно-оптический преобразователь представляет собой электровакуумную колбу, внутри которой размещены фотокатод, люминесцентный экран, микроканальная пластина и реализует функцию усиления яркости изображения. Характеристики и параметры ЭОП зависят от используемых в ЭОП фотокатода и люминесцирующего вещества. Преобразование изображения в

указанном приемном тракте происходит следующим образом: под воздействием внешнего излучения фотокатод электронно-оптического преобразователя генерирует фотоэлектроны. Таким образом, возникает *электронное изображение*. Эмитированные электроны проходят через микроканальную пластину (МКП), в которой число электронов увеличивается в сотни раз. Микроканальная пластина представляет собой набор нескольких миллионов стеклянных трубочек, сложенных в единый пакет в виде диска. Каждая трубочка имеет диаметр 10–20 мкм, длину 1–3 мм [15]. Внутренняя поверхность трубочки покрыта тонким (толщиной в несколько долей нанометра) слоем высокоомного полупроводника, который является эмиттером вторичных электронов, причем один первичный электрон выбивает при столкновении до 3–5 вторичных электронов. Таким образом, достигается усиление электронного изображения. Вторичные электроны создают один элемент светящегося изображения. Чтобы электроны не пролетали МКП без соударения со стенками каналов, каналы стараются располагать под небольшим углом к оси ЭОП. Это, однако, снижает качество изображения, так как заметен разброс электронов по скоростям и поэтому ощущается различие в фокусировке на экране быстрых и медленных электронов. Современные МКП позволяют получать изображения в самых сложных условиях недостаточной освещенности ФПУ, поскольку обладают коэффициентом усиления в несколько сотен тысяч раз.

Для преобразования энергии электронов в световую энергию служат люминесцирующие экраны. Электронный луч, попадая на такой экран, отдает целиком или частично свою энергию зернам люминофора, которые возбуждаются. Такой вид люминесценции называется катодолюминесценцией. Люминесценция, независимо от способа возбуждения, делится на флюоресценцию и фосфоресценцию. Флюоресценция исчезает немедленно после снятия возбуждения, в то время как фосфоресценция продолжает существовать и после прекращения возбуждения. Если изображение на экране рассматривается глазом человека, то спектральную характеристику люминофора стремятся совместить с кривой видимости глаза. Если в качестве ФПУ используется фотопленка, то удобнее использовать свечение люминофора в синей области, где квантовый выход фотоматериалов имеет большие значения. Свечение люминофора может вызываться не только катодолюминесценцией, но и непосредственно оптическим облучением. Люминофор следует возбудить коротковолновым излучением, а после того, как видимое излучение погаснет, его нужно возбудить снова, но уже

длинноволновым световым потоком. При этом будет происходить свечение люминофора в видимом диапазоне спектра. Лучший для этого люминофор – сульфид стронция, активированный самарием и церием, который дает зеленое стимулированное излучение. Этот люминофор может возбуждаться частицами небольшой порции радия. Другие люминофоры на основе сульфида цинка чувствительны до 1,3 мкм и выше [16]. Охлаждением до температуры жидкого азота спектральную характеристику удастся продлить до 2 мкм.

Модуляционная передаточная функция ЭОП имеет вид

$$T_{\text{эоп}}(v_x, v_y) = \exp \left[- \frac{\pi^2 \sigma_{\text{эоп}}^2 (v_x^2 + v_y^2)}{8} \right],$$

где $\sigma_{\text{эоп}}$ – диаметр пятна рассеяния на люминофоре ЭОП, который для преобразователей второго и третьего поколений можно считать равным $\sigma_{\text{эоп}} = 1,5 \cdot 10^{-2}$ мм.

Для того чтобы спроецировать изображение, полученное на поверхности люминофора, на фотоприемную матрицу используется **репродукционный объектив**, МПФ которого рассчитывается по соотношению

$$T_{\text{ро}}(v_x, v_y) = \exp \left[- \frac{\pi^2 (\lambda f_{\text{ро}})^2 (v_x^2 + v_y^2)}{D_{\text{ро}}^2} \right],$$

где $f_{\text{ро}}$ – фокусное расстояние репродукционного объектива; $D_{\text{ро}}$ – диаметр объектива; λ – длина волны свечения люминофора.

2.6. Модуляционная передаточная функция канала формирования видеосигнала

В общем случае сформированный матричным фотоприемником видеосигнал не может быть подан непосредственно на монитор, поскольку обладает рядом специфических особенностей. Главная из них состоит в импульсном характере экспонирования матрицы полученным сигналом, причем частота следования импульсов под света обычно меньше частоты кадров фотоприемного устройства. В результате видеосигнал содержит кадры, в которых отсутствует информация об объекте, поэтому изображение на экране дисплея начинает мелькать. Другая особенность выходного видеосигнала состоит в том, что для улучшения качества изображения при низ-

ком уровне освещенности требуется его дополнительная обработка. В общем виде процесс преобразования сигнала после фотоприемной матрицы представляется следующим образом:

$$L_{\text{ЭК}}(x, y, t) = P_{\text{ОВ}}(u_{\text{ФД}}(t)),$$

где $L_{\text{ЭК}}(x, y, t)$ – распределение яркости на экране дисплея; $P_{\text{ОВ}}$ – линейный преобразующий оператор; $u_{\text{ФД}}$ – выходное напряжение ФПЗС-матрицы. Рассмотрим поэтапное преобразование видеосигнала. Входной аналоговый модуль осуществляет временную фильтрацию выходного напряжения и по существу является низкочастотным фильтром. Полоса пропускания такого фильтра располагается на шкале частот от $f = 0$ до некоторой граничной частоты $f_{\text{ср}}$, а полоса непропускания (задержания) простирается от этой частоты до бесконечно больших частот. Передаточная функция такого идеального фильтра имеет вид

$$T_{\text{ид.ф}} = \text{rect}\left(\frac{f}{f_{\text{ср}}}\right),$$

где f – электрическая видеочастота, Гц; $f_{\text{ср}}$ – максимальная частота пропускания видеотракта (частота среза).

Связь видеочастоты с линейной v и угловой v' пространственными частотами выражается зависимостями

$$f = \frac{d_x m}{t_{\text{СТ}}} v_x; \quad f = \frac{\alpha_x m}{t_{\text{СТ}}} v'_x,$$

где d_x , α_x – линейный и угловой размеры шага между пикселями матричного фотоприемника; m – число элементов (пикселей) в строке; $t_{\text{СТ}}$ – активное время строки. Если максимальная пространственная частота ограничена частотой среза, равной

$$f_{\text{ср}} = \frac{m}{2t_{\text{СТ}}},$$

то идеальный НЧ-фильтр пропускает ограниченный спектр полезного сигнала без искажений и подавляет шум вне полосы пропускания, максимизируя тем самым отношение S/N .

Передаточная функция реального низкочастотного НЧ-фильтра существенно отличается от идеального и обычно аппроксимируется полиномом. Наибольшее распространение получили НЧ-фильтры Баттерворта, обеспечивающие наиболее плоскую характеристику в

полосе пропускания. Передаточная функция такого фильтра описывается формулой

$$T(f) = \frac{1}{\sqrt{1 + \left(\frac{f}{f_{3дБ}}\right)^{2n}}},$$

где $f_{3дБ}$ – частота, на которой амплитуда выходного сигнала уменьшается на 3 дБ; n – степень полинома или порядок фильтра. Чем выше порядок фильтра n , тем более плоским можно сделать участок МПФ в полосе пропускания и увеличить крутизну спада от полосы пропускания к полосе подавления. Когда видеочастота $f_{3дБ}$ соответствует частоте Найквиста, теорема выборки удовлетворяется, однако прямоугольные полосы тест-объекта будут воспроизводиться как синусоидальные, поскольку для воспроизведения четких границ требуется передача более высоких частот. Поэтому рекомендуется выбирать частоту $f_{3дБ}$ существенно более высокой, т. е. использовать более широкополосные фильтры.

Вторым элементом канала формирования видеосигнала является цифровой модуль, который преобразует аналоговый сигнал в цифровой двоичный код, после чего осуществляет его запоминание и обратное преобразование в аналоговый сигнал. После цифро-аналогового преобразователя стоит НЧ-фильтр, позволяющий отфильтровать побочные спектральные составляющие. Результирующая передаточная функция тракта обработки видеосигнала, выраженная в угловых пространственных частотах представляется в виде [1]:

$$T_{эл}(v'_x) = \text{sinc}(v'_x a_x) \frac{1}{1 + \left(\frac{v'_x}{v'_{x.сп}}\right)^{2n}}.$$

2.7. Модуляционная передаточная функция видеомонитора и зрительного анализатора

При работе любой изображающей системы качество формируемого изображения – является функцией не только первичной и вторичной обработки сигналов, но и зависит от выбора монитора. Для ЛСВ обычно применяются монохромные мониторы с электронно-лучевой трубкой. Основной их характеристикой является *разрешающая способность монитора*, которая зависит от размеров

зерна жидкокристаллической матрицы и представляет собой число пикселей, приходящееся на единицу длины экрана. Для мониторов с ЭЛТ эта зависимость количества телевизионных линий (ТВ-линий) $N_{ТВ}$, которые приходятся на высоту экрана H_M и числа эквивалентных линий $J_{ТВ}$, которые приходятся на длину экрана L_M , которые определяются размерами перемещающегося вдоль строки электронного пятна. Если профиль пятна радиально-симметричный и аппроксимируется гауссовой кривой, то интенсивность светового пятна имеет распределение [1]

$$I(r) = I_0 \exp\left(-\frac{r^2}{2\sigma_{\Pi}^2}\right),$$

где I_0 – интенсивность в центре пятна; σ_{Π} – среднеквадратическое отклонение. В этом случае разрешающая способность монитора зависит только от числа ТВ-линий на высоту экрана, а диаметр пятна d_{Π} , определенный по уровню $0,5I_0$, равен

$$d_{\Pi} = 2,35\sigma_{\Pi},$$

а передаточная функция монитора, выраженная через линейную пространственную частоту монитора ν_M , имеет вид

$$T_M(\nu_M) = \exp\left(-2\pi^2\sigma_{\Pi}^2\nu_M^2\right).$$

Пространственная частота ν_M , определяется числом ТВ-линий на экране, приходящимся на единицу длины (например, 1/мм). Она связана с линейными пространственными частотами в пространстве изображений ν_x и ν_y соотношениями

$$\nu_x = \nu_M\gamma_x; \quad \nu_y = \nu_M\gamma_y,$$

где $\gamma_x = \frac{L_M}{md_x}$, $\gamma_y = \frac{H_M}{md_y}$ – электронные увеличения, равные отношению размеров экрана к соответствующим размерам ФПЗС-матрицы.

В том случае, если диаметр пятна d_{Π} неизвестен, значение σ_{Π} можно оценивать параметром $N_{ТВ}$, характеризующим число ТВ-линий, укладываемых на высоте экрана H_M . В соответствии с критерием разрешения [1], две соседние линии разделены расстоянием $1,18\sigma_{\Pi}$. Поскольку на одном пространственном периоде укладываются две ТВ-линии, то линейный размер периода по высоте

равен $2,36\sigma_{\text{п}}$. Если масштаб изображения составляет $4/3$, то передаточная функция монитора определяется соотношениями

$$T_{\text{М}}(v_x) = \exp\left(-2\pi^2 \frac{3md_x}{2,35 \cdot 4N_{\text{ТВ}}} v_x\right)^2,$$

$$T_{\text{М}}(v_y) = \exp\left(-2\pi^2 \frac{nd_y}{2,35N_{\text{ТВ}}} v_x\right)^2.$$

Из представленных соотношений видно, что с увеличением диаметра пятна полоса пропускания дисплея уменьшается, что может вызвать снижение эффективности передачи изображения объекта.

Если обозначить расстояние от монитора до оператора через величину $R_{\text{М}}$, то угловые пространственные частоты для монитора и воспринимаемые наблюдателем, связаны с пространственными частотами соотношениями:

$$v'_x = v_x R_{\text{М}}; \quad v'_y = v_y R_{\text{М}}.$$

В процессе выбора параметров монитора для ЛСВ подбирают параметры монитора так, чтобы частоты, представляющие наибольший интерес, приходились бы на максимум МПФ глаза оператора.

В общем случае на МПФ глаза сильно влияют самые разнообразные характеристики органа зрения, и главным образом, дифракция на входном зрачке, аберрации хрусталика, конечные размеры фоточувствительных рецепторов и др. Кроме того, работа зрительного анализатора зависит от внешних условий наблюдения, например, от средней яркости экрана монитора, продолжительности наблюдения и т.д.

Однако МПФ глаза не учитывает пространственный шум, фоновую освещенность, положение оператора относительно монитора и времени экспонирования. Тем не менее, каждый из этих параметров влияет на качество восприятия изображения, поэтому эмпирические зависимости, характеризующие МПФ глаза, являются лишь аппроксимацией реальной МПФ.

Модуляционную передаточную функцию глаза можно измерить пороговым методом по тест-объекту с синусоидальным распределением яркости, определяя зависимость контраста, при котором с заданной вероятностью разрешается тест-объект, от пространственной частоты. Однако непосредственно измерить амплитуду или контраст воспринятого изображения синусоидальной картины в человеческом глазу не представляется возможным, поэтому

обычно измеряют функцию контрастной чувствительности (ФКЧ). Контраст пространственной решетки, воспроизводимой на экране ТВ-монитора, без изменения среднего значения яркости, уменьшается до порогового значения, т. е. до тех пор, пока наблюдатель не увидит однородную картину. Измерения повторяются на различных пространственных частотах, в результате чего определяется кривая зависимости порогового уровня по контрасту (в относительных единицах) для глаза от угловой пространственной частоты (модуляционная кривая глаза), вид которой представлен на рис. 24.

Величина, обратная пороговому уровню контраста, является пороговой чувствительностью, а ее зависимость от пространственной частоты есть *функция контрастной чувствительности*. Модуляционная передаточная функция глаза и функция контрастной чувствительности должны находиться в прямой пропорциональности, но, согласно [1], проводимые эксперименты показали, что это не совсем так. Модуляционная передаточная функция глаза имеет максимум в области угловых пространственных частот $\nu' = 6 \text{ град}^{-1}$. Эти значения согласуются с концепцией остроты человеческого зрения. В то же время пороговый контраст зрительного анализатора находится в области угловых пространственных частот $\nu' = 2 \text{ град}^{-1}$, т. е. максимальная чувствительность глаза при наблюдении малоконтрастных объектов сдвигается в сторону более низких пространственных частот. Так, при малом контрасте легче всего обнаруживаются крупноразмерные объекты (угловой размер более $0,25^\circ$), а при большом – легче обнаруживаются объекты с угловым размером $0,083^\circ$.

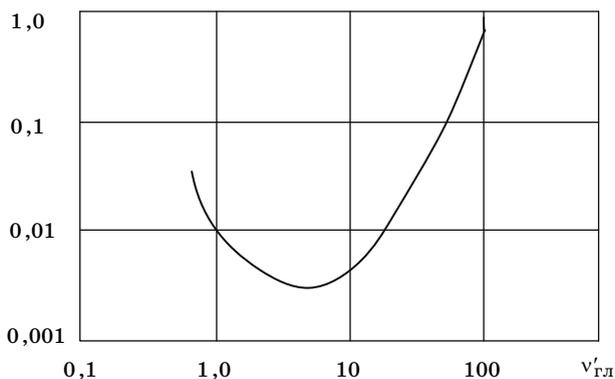


Рис. 24. Модуляционная кривая глаза

Таким образом, МПФ зрительного анализатора определяется соотношением

$$T_{\text{ГЛ}}(v'_{\text{ГЛ}}) = k_{\text{пор}}(v'_{\text{ГЛ}}) \Phi \text{КЧ}(v'_{\text{ГЛ}})$$

и, следовательно, вблизи порога зрительного восприятия МПФ пропорциональна $\Phi \text{КЧ}$.

Современная методология оценки качества изображающих систем, предназначенных для обнаружения и опознавания объектов рекомендует аппроксимацию Корнфельда-Лоусона [1]:

$$T_{\text{ГЛ}}(v'_{\text{ГЛ}}) = \exp\left(-B \frac{v'_{\text{ГЛ}}}{17,45}\right),$$

где $B = 1,444 - 0,344 \lg(B_{\text{ЭК}}) + 0,0395 \lg^2(B_{\text{ЭК}})$; $L_{\text{ЭК}}$ – яркость экрана.

Пространственная частота $v'_{\text{ГЛ}}$, регистрируемая зрительным анализатором, зависит от размеров экрана монитора и расстояния до наблюдателя и связана с угловой частотой в пространстве изображения соотношениями:

$$v'_x = \frac{L_{\text{М}}}{R_{\text{М}}} \frac{f'}{m d_x} v'_{x.\text{ГЛ}} = \Gamma_x v'_{x.\text{ГЛ}}; \quad v'_y = \frac{H_{\text{М}}}{R_{\text{М}}} \frac{f'}{n d_y} v'_{y.\text{ГЛ}} = \Gamma_y v'_{y.\text{ГЛ}}.$$

При нормальной яркости монитора (85 кд/м^2) и для небольших углов наблюдения, МПФ можно рассчитать по соотношениям:

$$T_{\text{ГЛ}}(v'_x) = \exp\left(-1,1 \frac{v'_x}{17,45 \Gamma_x}\right); \quad T_{\text{ГЛ}}(v'_y) = \exp\left(-1,1 \frac{v'_y}{17,45 \Gamma_y}\right).$$

Таким образом, вычисляя последовательно МПФ всех звеньев ЛСВ, можно определить результирующую МПФ всей системы в целом. В том случае, когда кривая, которая соответствует произведению всех звеньев ЛСВ, кроме МПФ среды приближается к единице, а МПФ среды резко падает, можно снижать требования к техническим характеристикам ЛСВ, которые отвечают за спад передаточной функции. Следовательно, только определив заданные условия окружающей среды и проанализировав ее влияние на передаточную характеристику в области пространственных частот, целесообразно определять структуру и элементную базу ЛСВ в целом. При уменьшении технических требований к узлам ЛСВ, не следует изменять тех параметров, которые отвечают за расчет энергетических характеристик и способствуют достижению требуемого отношения сигнал – помеха. Если для улучшения качества воспринимаемого

оператором изображения применяются алгоритмы цифровой обработки изображения, полученного после фотодетектирования, то для оценки их влияния на прием изображения требуется вводить соответствующую МПФ в области пространственных частот.

2.8. Пример практической реализации ЛСВ

Рассмотрим пример практической реализации ЛСВ [1], которая содержит лазерный канал подсвета и функционирует в импульсном режиме, обеспечивая одномоментную засветку всего заданного поля обзора. Приемный канал содержит фотоприемную матрицу для формирования изображения и обладает узкой диаграммой направленности приемного устройства.

Двухчастотная лазерная система видения с импульсным подсветом осуществляет подсвет объектов лазерным излучением в видимом и ближнем ИК-диапазонах спектра и предназначена для наблюдения объектов в условиях низкой естественной освещенности или ночью на дальностях до 4 км. Структурная схема ЛСВ приведена на рис. 25 и включает в качестве основных блоков:

- передающий канал, содержащий многофункциональный лазерный излучатель и формирующую оптическую систему;
- приемный канал, представляющий собой телевизионный визир на основе приемного объектива-трансфокатора и фокальной ФПЗС-матрицы с прогрессивной разверткой;
- электронный блок управления, содержащий канал формирования и обработки видеосигнала, канал синхронизации и вычислительный канал;
- опорно-поворотное устройство для управления пространственным положением оптических осей приемопередающей системы.

Передающий канал обеспечивает формирование пространственно-энергетических характеристик лазерного излучения подсвета на двух рабочих частотах. Существенной особенностью лазера является его многофункциональность, т. е. возможность работы в различных режимах генерации. Переключение режимов генерации осуществляется по командам блока управления внутрирезонаторным модулятором добротности (затвором).

Оптическая схема лазера, формирующего излучение с основной длиной волны $\lambda = 1,06$ мкм (рис. 26), включает задающий генератор и усилитель. На рис. 26 введены следующие обозначения: 1 – зеркало резонатора; 2 – оптический затвор; 3 – активный элемент задающего генератора; 4 – плоскопараллельная пластина (поляри-

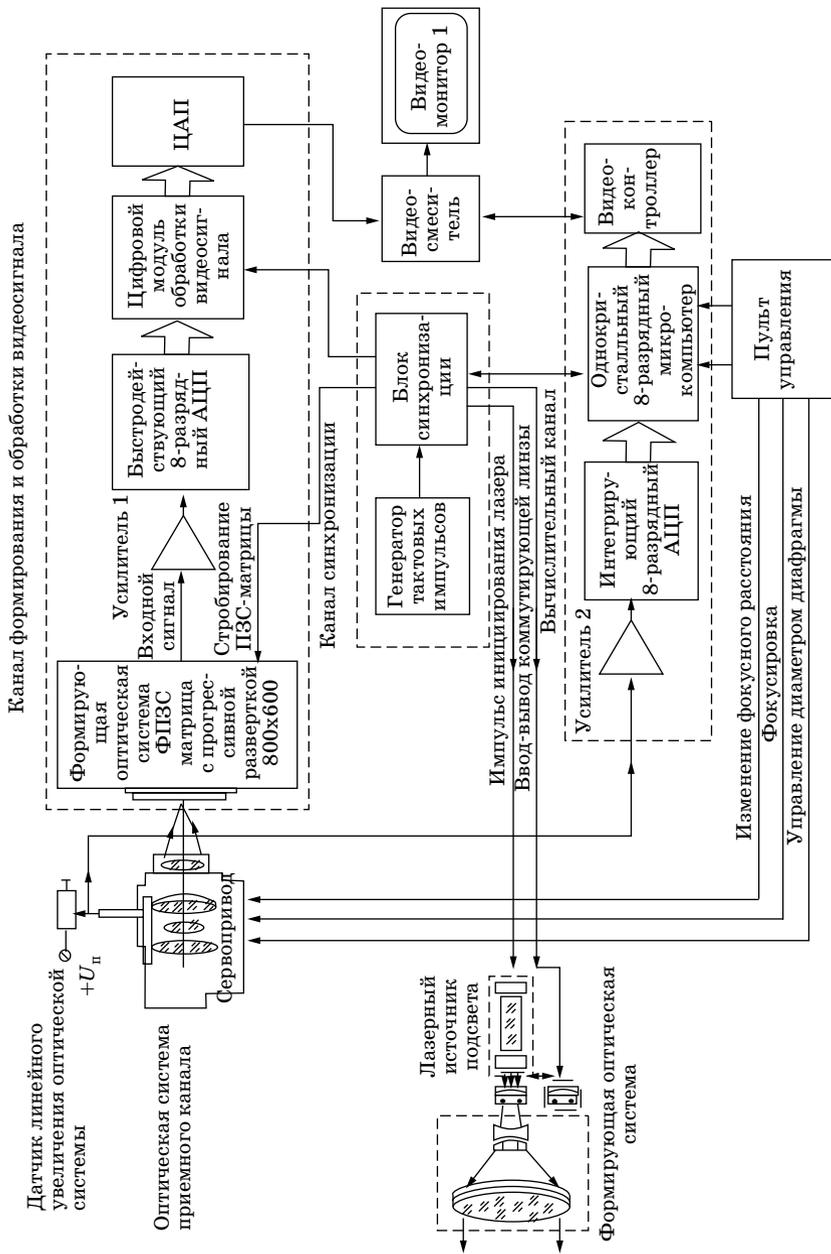


Рис. 25. Двухчастотная ЛСВ

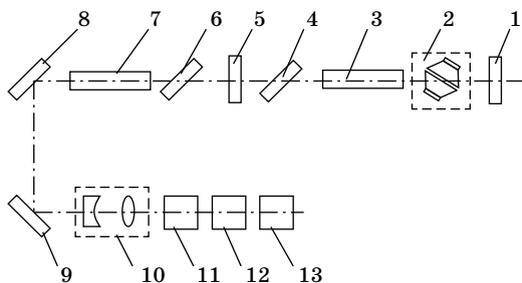


Рис. 26. Источник излучения системы подсвета

затвор); 5 – выходное зеркало резонатора; 6 – поляризатор; 7 – усилитель, 8, 9 – поворотные зеркала; 10 – телескоп, 11, 13 – удвоитель частоты; 12 – кварцевый фазовращатель.

Специально разработанный оптический затвор на основе нарушения полного внутреннего отражения позволяет реализовать два режима генерации: моноимпульсный и периодической модуляции добротности (режим формирования пачки импульсов). Для формирования поляризованного излучения в резонаторе расположены плоскопараллельные пластины, установленные под углом Брюстера. Удвоитель частоты выполнен на нелинейном кристалле и предназначен для преобразования инфракрасного излучения с основной длиной волны $\lambda_1 = 1,062$ мкм в видимое излучение с длиной волны $\lambda_2 = 0,53$ мкм. Для уменьшения влияния деполаризации излучения накачки на коэффициент полезного действия преобразования в оптическую схему введены последовательно расположенный второй кристалл и фазовращатель между ними.

Передающая оптическая система формирует три различных диаграммы направленности:

широкую ($2\theta = 1,5^\circ$), используемую в режиме поиска и обнаружения;

узкую в режиме пачки импульсов ($2\theta \leq 10'$);

узкую в режиме моноимпульса ($2\theta = 1,5'$).

Две последние диаграммы используются в процессе работы ЛСВ при различных уровнях восприятия (различение, идентификация).

Оптическая система (см. рис. 26) состоит из телескопа с фиксированным видимым увеличением, который формирует обе узкие диаграммы направленности и вбрасываемый оптический компонент (линза-коммутатор). Установка линзы-коммутатора перед телескопом обеспечивает ширину диаграммы направленности излучения подсвета $2\theta = 1,5^\circ$.

Ввод и вывод линзы-коммутатора осуществляется электродвигателем с редуктором, а ее фиксация в рабочих положениях обеспечивается за счет момента инерции ротора двигателя. Управление вбрасываемым компонентом осуществляется командами, вырабатываемыми в блоке управления.

Приемный канал предназначен для формирования изображений заданного сектора обзора в пассивном и активном режимах работы и точного наведения приемной системы на объект с целью определения его координат. Он выполнен на основе ПЗС-камеры. Сигналы с выхода ПЗС-камеры преобразуются в телевизионный сигнал. На экране видеомонитора вместе с изображением формируются координатная визирная сетка с перекрестием и служебные символы, информирующие оператора о текущем состоянии всех подсистем аппаратуры ЛСВ.

Электронный блок управления выполняет две основные функции:

- формирование и обработку видеосигнала;
- управление работой всех блоков и модулей ЛСВ.

Канал формирования и обработки видеосигнала (ФОВ) (см. рис. 26) предназначен для формирования видеосигнала ФПЗС-матрицей. Сформированный видеосигнал не может быть непосредственно подан на видеомонитор из-за импульсного характера экспонирования матрицы локационным сигналом. Поскольку частота импульсов экспонирования меньше частоты формирования видеокадров ФПЗС-матрицей, то часть видеокадров не содержит локационный сигнал. Это приводит к мельканию изображения на экране.

Для преобразования нестандартного видеосигнала с выхода ФПЗС-матрицы в видеосигнал служит цифровой канал обработки. Это преобразование осуществляется с частотой подсвета 5 Гц, а считывание – с частотой 50 Гц. Кроме того, цифровой модуль канала ФОВ реализует электронное увеличение в два или в четыре раза за счет фрагментации локального участка матрицы и последующего вывода изображения на весь экран.

Канал синхронизации осуществляет временное согласование работы лазерного излучателя и ФПЗС-матрицы, а также выполняет функции контроллера управления.

Лазерные импульсы подсвета, отраженные объектом и подстилающей поверхностью, должны попасть в определенный временной интервал накопления рабочего цикла ФПЗС-матрицы. Для синхронизации используется высокостабильный тактовый генератор, импульсы которого одновременно инициируют генерацию лазерного излучения и начало периода накопления ФПЗС-матрицы.

Эти же импульсы запускают схему стробирования, которая забирает матрицу на время эффективного воздействия на приемный канал помехи обратного рассеяния. Кроме того, канал синхронизации осуществляет управление вводом-выводом линзы-коммутатора в передающем канале для изменения угловой расходимости лазерного излучения.

Для формирования динамического изображения визирной сетки и служебных символов на экране монитора, существенно влияющих на точность наведения, блок управления содержит вычислительный канал на основе однокристалльного микропроцессора. Положение визирных меток на экране видеомонитора рассчитывается в соответствии с введенной программой, учитывающей нелинейный характер изменения увеличения при перемещении компонентов приемного объектива. На основе сигнала с датчика линейного увеличения микропроцессор строит рамку визира на экране.

С помощью пульта управления осуществляются дистанционное управление линейным увеличением приемного объектива и его фокусировкой, регулировку апертурой диафрагмы в соответствии с уровнем яркости фона, дискретное изменение электронного увеличения и частоты следования импульсов подсвета, а также переключение режимов работы.

В процессе работы оператор с помощью приводов опорно-поворотного устройства направляет приемопередающую систему в заданную пространственную область обзора и включает канал подсвета. Осуществляется режим поиска и обнаружения объекта, при котором формируется широкая ($2\theta = 1,5^\circ$) диаграмма направленности лазерного излучения. При этом лазерный излучатель генерирует пачку импульсов длительностью 200–300 мкс и суммарной энергией в пачке не менее 5 Дж.

В случае обнаружения объекта его изображение выводится механизмами опорно-поворотного устройства в центральную зону визирной рамки, после чего линза-коммутатор удаляется из оптической системы передающего канала и диаграмма направленности лазерного излучения «схлопывается» с $1,5^\circ$ до $8'$. В результате освещенность в плоскости наблюдаемого объекта возрастает более чем в 200 раз, что приводит к существенному увеличению отношения сигнал-шум, необходимому для достоверной идентификации объекта.

Изображение легкового автомобиля, полученного в лазерной системе видения в степном холмистом регионе на дальности 2,7 км, приведено на рис. 27 [1]. Изображение построено при узкой диа-

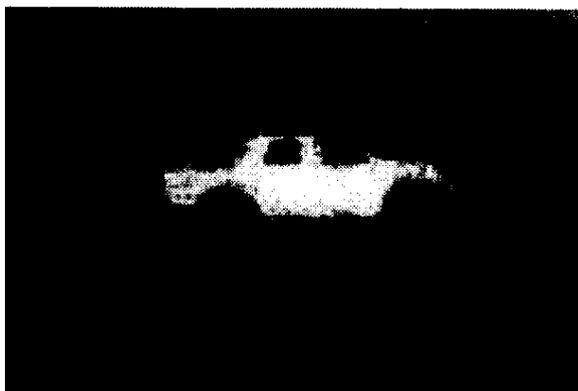


Рис. 27. Изображение легкового автомобиля, полученного при помощи ЛСВ

грамме направленности и суммарной энергии в пачках на обеих длинах волн, не превышающей 4,8 Дж.

Дальность обнаружения указанного объекта ограничивалась длиной трассы $z = 4,2$ км, на этой предельной дальности вероятность превышала 80%.

2.9. Фотодетектирование в приемных матрицах

При зондировании области пространства лазерным лучом формируется распределение освещенности на ФПЗС-матрице, создавая первичное электронное изображение. Следует отметить, что процесс формирования изображения в электронном тракте очень важен, так как именно по полученной в итоге на экране монитора картинке оператор будет принимать решение о наличии или отсутствии объекта, производить его идентификацию и принимать дальнейшие решения. Рассмотрим, как происходит формирование первичного электронного изображения в электронном тракте и на ФПЗС-матрице.

Отметим, что в общем случае электронный тракт выполняет три основные функции:

- а) фотодетектирование и цифровая обработка изображения для улучшения его качества;
- б) формирование видеосигнала, несущего информацию об объекте наблюдения;
- в) реализацию алгоритма управления всеми каналами и подсистемами ЛСВ, например, с целью синхронизации работы фотопри-

емника в приемном канале и излучающего импульсного лазера в передающем канале.

Рассмотрим устройство ФПЗС-матриц и процесс фотодетектирования, осуществляемый с их помощью. Основой фоточувствительного элемента ФПЗС матрицы является МДП-конденсатор, в котором в отличие от обычного конденсатора одна из обкладок представляет собой полупроводниковый материал с примесью проводимостью (обычно р-типа). Отсюда сокращение МДП: металл – диэлектрик – полупроводник (рис. 28) [17].

Если подать положительное напряжение на металлическую обкладку (электрод) этого конденсатора, то под ней сформируется электрическое поле, распространяющееся через диэлектрик вглубь полупроводника. При этом в полупроводнике произойдет пространственное разделение зарядов: положительно заряженные дырки (основные носители для полупроводника р-типа) будут отталкиваться от границы раздела полупроводник-диэлектрик (т. е. от поверхности полупроводника, граничащего через диэлектрик с электродом), а не основные (электроны) – скапливаться у поверхности полупроводника. Необходимо отметить, что минимальным напряжением на электроде является напряжение, при котором на-

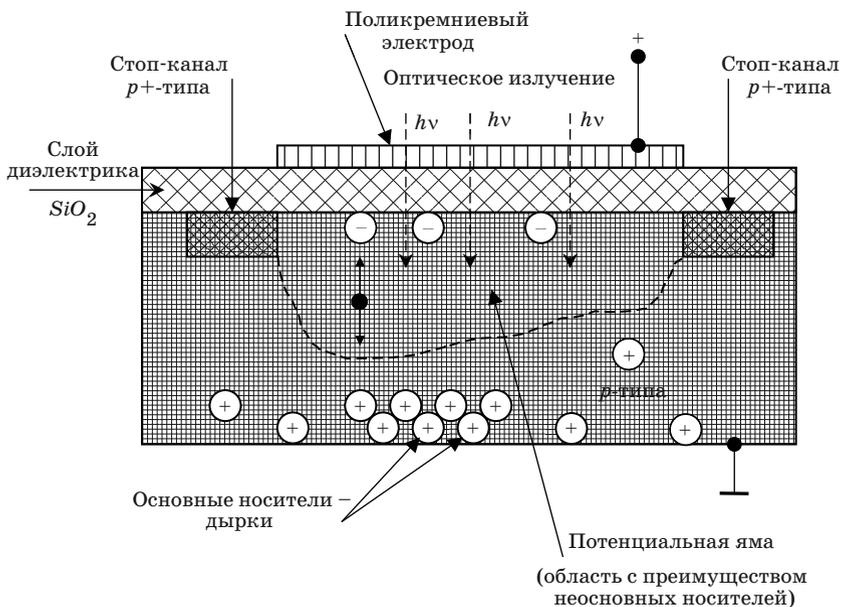


Рис. 28. К пояснению принципа действия ФПЗС-матрицы

ступает инверсия приповерхностной зоны полупроводника, т. е. состояние, при котором отсутствуют основные носители – дырки, а рассматриваемая зона становится отрицательно заряженной. Таким образом, потенциальная яма заполняется электронами, а ее глубина зависит от геометрических параметров МДП-конденсатора и величины напряжения на электроде.

Поступление неосновных носителей (электронов) в потенциальную яму обусловлено тремя причинами:

а) тепловой генерацией электронов;

б) воздействием электрического поля;

с) генерацией фотоэлектронов в результате поглощения квантов оптического излучения.

Тепловая генерация (термогенерация) электронов является результатом теплового движения частиц, при котором возможна ионизация нейтрального атома, т. е. образования пары электрон – дырка. Под действием внешнего электрического поля, создаваемого электродом, дырка отталкивается в глубь полупроводника, а электрон остается в потенциальной яме. Процесс термогенерации носит случайный характер и приводит к появлению теплового шума.

Генерация фотоэлектронов обусловлена поглощением фотонов внешнего излучения, чья энергия $h\nu$ ($h = 6 \cdot 10^{-34}$ Дж·с – постоянная Планка; ν – частота излучения) превышает ширину запрещенной зоны используемого полупроводникового материала. При облучении ячейки ФПЗС-матрицы фотоны, поглощенные нейтральными атомами в обедненной области, приводят к их возбуждению. В результате образуется пара: фотоэлектрон и ионизированный атом (дырка), которые разделяются электрическим полем. Фотоэлектрон остается в потенциальной яме, а дырка перемещается вглубь полупроводника. Если продолжать удерживать положительное напряжение на электроде, будет происходить накопление фотоэлектронов в потенциальной яме, а МДП-конденсатор начнет заряжаться. Для предотвращения растекания накапливаемых фотоэлектронов в соседние области за счет диффузии, область потенциальной ямы ограничивают стоп-каналами.

Для того чтобы фотоны могли проникнуть в обедненную область, электрод изготавливают из поликремния с металлическим напылением таким образом, чтобы он был прозрачен для входного излучения. В качестве диэлектрика используется окисел кремния SiO_2 толщиной 0,1-0,01 мкм, поэтому МДП-конденсатор часто называют МОП-конденсатором (металл-окисел-полупроводник). В последнее время получил распространение способ облучения

ФПЗС-ячейки со стороны полупроводника. При этом для уменьшения потерь входного потока полупроводниковую подложку делают тонкой (200–300 мкм) [17]. Такие ФПЗС-ячейки являются более чувствительными к оптическому излучению, особенно в коротковолновой области спектра, поскольку в этом случае коэффициенты пропускания электрода и окисла значения не имеют.

Простейший фотоприемник на основе ФПЗС представляет собой полупроводниковую подложку с последовательно установленными на ней МДП-конденсаторами, играющими роль фоточувствительных ячеек. Для того чтобы считать накопленный заряд фотоэлектронов с каждого МДП-конденсатора, используется механизм переноса заряда из одной фоточувствительной ячейки в другую с использованием дополнительных ячеек – регистров переноса.

Каждый элемент ФПЗС (ячейка, пиксел) представляет собой фоточувствительную зону, окруженную нечувствительной к свету областью, куда входят элементы регистров переноса, каналы стоп-диффузии и т.д. Соотношение между ними характеризуется *фактором заполнения*, представляющим собой отношение площадей фоточувствительной зоны к общей площади элемента. С целью максимизации накопленного заряда необходимо, чтобы все входное излучение попадало на фоточувствительную область. Одним из вариантов решения этой проблемы является применение микролинз (рис. 29). У современных ФПЗС приемников фактор заполнения достигает 100 %.

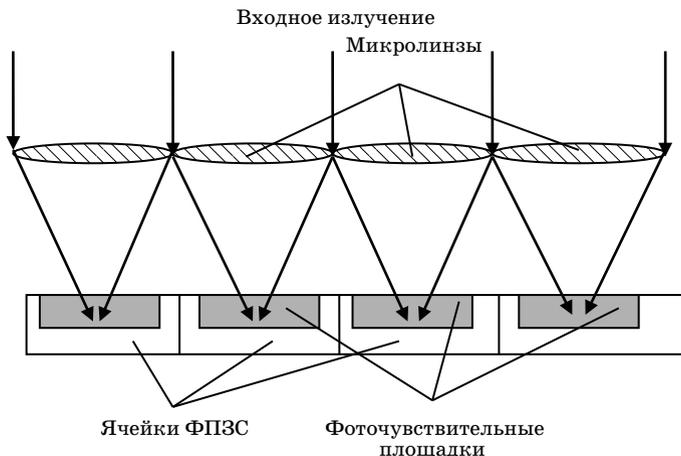


Рис. 29. ФПЗС-матрица с микролинзами

Процесс накопления фотоэлектронов занимает определенное время, называемое периодом накопления. После этого происходит перенос накопленных зарядовых пакетов к выходному устройству, преобразующему заряд каждой ячейки в напряжение. Это напряжение, сложенное с кадровыми и строчными синхросигналами, представляет собой электрический композитный видеосигнал. В свою очередь, видеосигнал несет информацию об изображении, принятом ФПЗС-приемником.

2.10. Воспринимаемое оператором отношение сигнал/шум

Основным назначением ЛСВ является обнаружение и распознавание объектов на предельных дальностях, когда уровень полезного сигнала соизмерим с уровнем шумов. Оператор наблюдает на экране дисплея зашумленное, «заснеженное» изображение вследствие случайных временных и пространственных флуктуаций освещенности. Даже высококонтрастные детали изображения могут оказаться неразличимыми при большом уровне шума. Возможность регулировки яркости и контраста изображения, воспроизводимого на экране монитора, приводит к тому, что реальный контраст на входе ЛСВ существенно отличается от контраста наблюдаемого изображения. Поэтому для ЛСВ, разрешающая способность которых ограничивается не контрастной чувствительностью, а уровнем шумов, т. е. отношением сигнал/шум (S/N), целесообразно пороговый уровень определять не через воспринимаемый контраст, а через воспринимаемое отношение сигнал/шум $(S/N)_{\text{восп}}$. Установив связь между контрастом тест-объекта в виде прямоугольных мир и величиной $(S/N)_{\text{восп}}$ можно найти такой минимально разрешаемый контраст (МРК) черно-белых прямоугольных полос миры, при котором $(S/N)_{\text{восп}}$ равно пороговому значению $(S/N)_{\text{восп.пор}}$.

Понятие МРК применимо только для обнаружения объектов, имеющих форму чередующихся черно-белых полос различной пространственной частоты. Существуют критерии (например, критерий Джонсона), которые устанавливают эквивалентность этих эталонных мир и реальных объектов с позиции вероятности их восприятия. Критерий Джонсона подробнее будет рассмотрен далее. В результате понятие МРК распространяется на любые реальные объекты, а связующим звеном является пространственная частота эквивалентной миры. Понятие МРК применимо исключительно к монохромным оптико-электронным изображающим системам, причем предполагается, что контраст изображения на экране дис-

плея достаточно высок и не ограничивает разрешение и дальность действия ЛСВ.

Значение воспринимаемого оператором отношения сигнал/шум, пересчитанное к выходу фотоприемного устройства, определяется выражением

$$\left(\frac{S}{N}\right)_{\text{восп}} = \frac{\Delta n_c T_{\text{лсв}}(v_x, v_y)}{n_{\text{сист}} M(v_x, v_y) N(f)} = \left(\frac{S}{N}\right)_{\text{сист}} \frac{T_{\text{гл}}(v_x, v_y) T_{\text{м}}(v_x, v_y)}{M(v_x, v_y) N(f)},$$

где Δn_c – разность средних чисел фотоэлектронов, формируемых от объекта и фона элементарными ячейками (пикселями) фотоприемного устройства под воздействием излучения за определенный временной интервал; $n_{\text{сист}}$ – среднее квадратическое значение числа шумовых фотоэлектронов, приведенное к одной ячейке; $T_{\text{сист}}(v_x, v_y)$ – МПФ изображающей системы на выходе электронного тракта; $T_{\text{гл}}(v_x, v_y)$ $T_{\text{м}}(v_x, v_y)$ – МПФ зрительного анализатора и дисплея; $(S/N)_{\text{сист}}$ – аппаратное отношение сигнал/шум системы на выходе электронного блока; $M(v_x, v_y)$, $N(f)$ – функции, описывающие пространственную и временную интегрирующую способность зрительного анализатора.

Принципиальная разница между аппаратным и воспринимаемым отношениями сигнал/шум состоит в том, что аппаратное отношение сигнал шум определяется в точке, т. е. для элемента изображения, площадь которого $S_{\text{кор}}$ (область корреляции) равна размеру одного пикселя. С увеличением размера это отношение существенно не изменяется, хотя глазу становится легче различать протяженные объекты. Из этого следует, что зрительный аппарат воспринимает отношение сигнал/шум, характерное не для элемента изображения, а для всего изображения в целом.

В результате экспериментальных исследований установлено, что эффект пространственного интегрирования приводит к улучшению отношения сигнал/шум. Эффект временного интегрирования связан с *инерцией зрения*, т. е. способностью зрительного анализатора сохранять или накапливать сигнал в течение определенного промежутка времени. Многочисленные исследования подтвердили способность глаза воспринимать не мгновенное значение случайного шума, а его среднее квадратическое значение, усредненное за конечный период. Это означает, что в системах с кадровой разверткой, при условии некоррелированности шума в пространстве от кадра к кадру, изображение, воспринимаемое в результате суммирования нескольких кадров, имеет меньший уровень шумов, чем изображение одного кадра. При этом отношение сигнал/шум

увеличивается пропорционально величине $\sqrt{f_k t_{\text{ГЛ}}}$, где f_k – частота кадров, а $t_{\text{ГЛ}}$ – время инерции глаза.

Таким образом, без учета селективных свойств глаза к воздействию различных видов шума при пространственно-временном интегрировании воспринимаемое отношение сигнал/шум связано с аппаратным формулой [1]:

$$\left(\frac{S}{N}\right)_{\text{восп}} = \left(\frac{S}{N}\right)_{\text{сист}} \frac{T_{\text{ГЛ}}(v_x, v_y) T_{\text{М}}(v_x, v_y)}{\left(\frac{1}{\sqrt{f_k t_{\text{ГЛ}}}}\right) \sqrt{\frac{1}{q_{\text{М}} f_{\text{об}}} v'_x} \sqrt{\frac{1}{q_{\text{М}} f_{\text{об}}} v'_y}},$$

где $q_{\text{М}}$ принимают равным 5 для систем видимого и ближнего инфракрасного диапазонов, а значения функций $M(v_x, v_y)$, $N(f)$ определяются соответственно по формулам:

$$M(v_x, v_y) = \sqrt{\frac{1}{q_{\text{М}} f_{\text{об}}} v'_x} \sqrt{\frac{1}{q_{\text{М}} f_{\text{об}}} v'_y}; \quad N(f) = \frac{1}{\sqrt{f_k t_{\text{ГЛ}}}}.$$

Эффекты пространственного и временного интегрирования, проявляющиеся в процессе восприятия изображений зрительным анализатором, приводят к существенному повышению воспринимаемого отношения $(S/N)_{\text{восп}}$ относительно аппаратного значения. В рационально спроектированных ЛСВ оператор уверенно обнаруживает изображения объектов при $(S/N)_{\text{сист}}$, не превышающем 0,05.

2.11. Шумы в электронном тракте ЛСВ

Величина $n_{\text{сист}}$ характеризует суммарный шум системы, приведенный к накопительной ячейке ФПЗС-матрицы. Суммарный шум проявляется в виде случайных флуктуаций фотоэлектронов как от ячейки к ячейке фоточувствительной матрицы, так и от кадра к кадру для некоторой фиксированной ячейки. Шум, зависящий от времени, т. е. изменяющийся от кадра к кадру, проявляется на экране в виде мерцания отдельных линий и точек. Шум, зависящий только от пространственных координат, проявляется на экране как случайное, но не изменяющееся со временем распределение яркости. Принимая все компоненты шума независимыми друг от друга, получаем

$$\left\langle n_{\text{сист}}^2 \right\rangle^{\frac{1}{2}} = \sqrt{\left\langle n_1^2 \right\rangle + \left\langle n_2^2 \right\rangle + \dots + \left\langle n_n^2 \right\rangle},$$

где $\left\langle n_i^2 \right\rangle$ – дисперсия i -го шумового компонента.

Фотоприемный канал в силу своей сложной внутренней структуры содержит наибольшее количество шумовых источников. Рассмотрим основные виды шумов ФПЗС матрицы и дальнейшего электронного тракта.

Дробовый шум входного излучения присутствует всегда и обусловлен флуктуациями фотонов входного оптического излучения относительно среднего значения за интервал наблюдения, поэтому он часто называется флуктуационным шумом. При работе системы видения в активном режиме со стробированием по дальности входное оптическое излучение содержит две составляющие: сигнальную (лазерное, а также солнечное или лунное излучения, отраженные объектом) и фоновую (излучения, отраженные поверхностью, на которой наблюдается объект).

Контраст в изображении возникает при регистрации обеих составляющих раздельно соседними элементарными фоточувствительными ячейками ФПЗС-матрицы. Таким образом, предполагается, что некоторый фиксированный пиксель матричного фотоприемника накапливает за определенный временной интервал среднее число $\bar{n}_{об}$ сигнальных фотоэлектронов, а соседний пиксель в среднем регистрирует $\bar{n}_{ф}$ фоновых фотоэлектронов. Разность между этими значениями $\Delta n = \bar{n}_{об} - \bar{n}_{ф}$ и является в ЛСВ полезным сигналом, несущим информацию об объекте наблюдения.

Реальные числа использованных фотонов $n_{об}$ и $n_{ф}$ флуктуируют относительно соответствующих средних значений из-за квантовой природы света и обусловленных ею дискретностью и неопределенностью распределения фотонов. Эти флуктуации проявляются в фоторегистрирующих устройствах в виде дробового шума. Если регистрируемая интенсивность входного оптического излучения невелика, что соответствует работе ЛСВ на предельных дальностях, то распределение фотоэлектронов подчиняется статистике Пуассона со среднеквадратическим значением

$$\left\langle n_{\text{фот}}^2 \right\rangle^{\frac{1}{2}} = \sqrt{\frac{\bar{n}_{об} + \bar{n}_{ф}}{2}}.$$

Наличие помехи обратного рассеяния излучения подсвета и внешняя фоновая засветка, обусловленная рассеянием солнечного излучения толщей атмосферы, вызывает не только снижение контраста, но и уменьшение отношения сигнал/шум за счет увеличения уровня дробового шума.

Даже в отсутствие входного излучения в накопительной ячейке ФПЗС-матрицы образуется заряд, формируемый термогенерацией.

В процессе генерации заряда накопительная ячейка, представляющая собой конденсатор, заряжается темновым током, плотность которого J_T характеризует для данной матрицы скорость термогенерации. Для кремниевых ПЗС в зависимости от производителя плотность темнового тока лежит в пределах $0,1 \dots 10$ пА/см² и может быть сильно уменьшена за счет охлаждения. Среднее число (математическое ожидание) термогенерированных электронов в ячейке ПЗС определяется выражением

$$\bar{n}_T = \frac{J_T S_{\text{э}} t_{\text{н}}}{e},$$

где e – заряд электрона; $S_{\text{э}}$ – площадь одного элемента; $t_{\text{н}}$ – время накопления.

Плотность темнового тока сильно зависит от температуры

$$J_T = k \exp\left(-\frac{E}{ckT}\right),$$

где k – постоянная Больцмана; E – энергетическая ширина запрещенной зоны данного полупроводника; T – температура; c – коэффициент, зависящий от типа полупроводника (от 1 до 2).

Электроны, эмиссия которых обусловлена термогенерацией, также вносят свой вклад в дробовый шум матрицы. В соответствии со статистикой Пуассона дисперсия дробового шума темнового тока

$$\langle n_T^2 \rangle = \bar{n}_T.$$

Таким образом, суммарный дробовый шум матрицы состоит из трех компонентов: дробового шума сигнальных фотоэлектронов, дробового шума фоновых фотоэлектронов и дробового шума темнового тока.

$$\langle n_{\text{др}}^2 \rangle^{\frac{1}{2}} = \sqrt{\bar{n}_T + \frac{\bar{n}_{\text{об}} + \bar{n}_{\text{ф}}}{2}}.$$

Шум восстановления связан с особенностями работы узла преобразования накопленного заряда в напряжение, содержащего в своем составе резистор R и конденсатор C . На резисторе происходит термогенерация дополнительных электронов, создающих шумовой ток. Дисперсия этого тока в соответствии с формулой Джонсона [1]:

$$\langle i_{ш}^2 \rangle = \frac{4kT}{R} \Delta f ; \Delta f = \frac{RC}{4},$$

а среднеквадратическое значение эквивалентных электронов:

$$\langle n_{вс}^2 \rangle^{\frac{1}{2}} = \frac{\sqrt{kTC}}{e}.$$

Его можно уменьшить, применив охлаждение матрицы.

Шум квантования. Причина возникновения – наличие в приемном тракте ЛСВ аналого-цифрового преобразователя (АЦП) и соответствие разным значениям напряжения одного уровня АЦП.

Среднеквадратическое значение напряжения ошибки:

$$V_{кв} = \frac{U_{\max}}{2^N \sqrt{12}},$$

где N – разрядность АЦП; U_{\max} – максимальное напряжение. Если вход АЦП согласован с выходом фотоприемника, то среднеквадратическое значение шумовых электронов:

$$\langle n_{кв}^2 \rangle = \frac{1}{2^N} \frac{N_{\max}}{\sqrt{12}},$$

N_{\max} – максимальный заряд, который может накопить ячейка ПЗС.

Идеальной матрице присущ только дробовой шум, причем та его часть, которая определяется входным излучением.

Даже в его отсутствие в приемнике существуют шумы, которые определяют шумовой порог:

$$\langle n_{\text{пор}}^2 \rangle = \langle n_{\text{т}}^2 \rangle + \langle n_{\text{вс}}^2 \rangle + \langle n_{\text{кв}}^2 \rangle.$$

Если в приемном канале ЛСВ используется электронно-оптический преобразователь ЭОП с коэффициентом преобразования $K_{\text{эоп}}$, то дисперсия суммарного шума:

$$n_{\Sigma}^2 = \frac{\bar{n}_{\text{об}} + \bar{n}_{\text{ф}}}{2} \left(1 + K_{\text{эоп}} + \frac{\langle n_{\text{пор}}^2 \rangle}{\langle n_{\text{фот}}^2 \rangle} \right).$$

В общем случае ЛСВ является многофункциональной системой, способной работать в различных условиях: от полной темноты и слабого входного сигнала до мощного сигнала и сильного помехового фона. В связи с этим в различных ситуациях применения ЛСВ от-

дельные шумовые компоненты по-разному оказывают влияние на качество получаемого изображения. Согласно [1] при слабом входном излучении суммарный шум ПЗС определяется шумовым порогом порядка 200–500 шумовых электронов на ячейку. Эта ситуация характерна для случая, когда ЛСВ с импульсным подсветом работает в полной темноте (фон практически отсутствует) и на предельной дальности от объекта. Интерес представляет ситуация, когда система видения работает в режиме обнаружения слабого сигнала на предельных дальностях при наличии подсветки либо в пассивном режиме при недостаточной естественной освещенности. В этом случае доминирующую роль играют дробовые шумы темнового тока ФПЗС.

2.12. Определение предельной дальности обнаружения и распознавания ЛСВ

Предельная дальность видения полностью определяет эффективность использования ЛСВ в заданных условиях наблюдения. Определение дальности видения при использовании ЛСВ связано с получением математического выражения, которое описывает процесс восприятия и интерпретации изображения зрительным анализатором. Выделено несколько уровней восприятия, нижний из которых (обнаружение объекта) соответствует выделению какого-то размытого пятна на фоне помех, а высший уровень соответствует точной идентификации объекта и определению его специфических особенностей. Между этими уровнями находится ряд различных уровней восприятия. В качестве параметра, характеризующего возможный уровень восприятия при работе с оптико-электронной системой видения предложено использовать разрешение штриховых мир, эквивалентных объекту. Под одним периодом мира понимается одна пара из черной и белой полосы. При этом предполагается, что объект характеризуется неким минимальным размером $H_{кр}$, существенным для его восприятия.

Под эквивалентной штриховой мирой понимают миру прямоугольной формы, ширина которой равна минимальному или критическому (при двумерном восприятии) размеру объекта, а длина соответствует размеру объекта в направлении, перпендикулярном критическому.

В соответствии с таким подходом, впервые развитым Джонсоном, разрешение эквивалентной мира связывалось с качеством (уровнем) видения, которое определялось максимальной разрешаемой пространственной частотой мира, имеющей тот же контраст и на-

блюдаемой при тех же условиях, что и объект. На примере военных транспортных средств Джонсон провел сравнение способности наблюдателя разрешать изображение миры с его способностью воспринимать объект с различным качеством видения. В результате каждому уровню видения он поставил в соответствие число некоторых разрешаемых периодов эквивалентной миры N , укладывающихся в пределах критического размера объекта. Идентификация объекта с заданной вероятностью требует, чтобы на его критическом размере уложилось определенное количество периодов эквивалентной миры.

Количество периодов эквивалентной миры для различных уровней восприятия представлено в табл. 3 [1].

Таблица 3

Уровень видения	Число периодов эквивалентной миры, N
Обнаружение	1
Определение ориентации	1,5
Различение	3
Идентификация (опознавание)	7

Джонсону удалось также для определения вероятности распознавания объекта на телевизионном растре ввести следующий критерий:

$$P_{\text{расп}} = \left(1 - e^{-0,15(N-1)^2}\right).$$

Данный критерий позволяет по заданной вероятности распознавания объекта определить требуемое число периодов эквивалентной миры, используя следующее соотношение:

$$N = 1 + \left(-\frac{\ln(1 - P_{\text{расп}})}{0,15} \right)^{0,5}.$$

Число периодов эквивалентной миры N , укладывающихся на критическом размере наблюдаемого объекта, связано с угловой пространственной частотой соотношением

$$v' = \frac{Nz}{H_{\text{кр}}},$$

где z – дальность до объекта.

Таким образом, при заданном критическом размере можно результирующую МПФ ЛСВ представить в виде функции от расстояния до объекта. При этом, если график функции МПФ ЛСВ от

угловой пространственной частоты по x спадает быстрее, чем по y , то соответствующая функция и используется для определения предельной дальности действия системы.

Знание шумов ФПЗС-матрицы и воспринимаемого отношения сигнал/шум для оператора позволяет определить минимальный разрешаемый контраст ЛСВ. Согласно [1] его можно вычислить по формуле

$$k_{\min}(v_x, v_y) = \frac{\pi^2}{16} \frac{(S/N)_{\text{воспр.пор}} \sqrt{1 + \frac{\langle n_{\text{пор}}^2 \rangle}{\langle n_{\text{др}}^2 \rangle}}}{T_{\text{лсв}}(v_x, v_y) \sqrt{\frac{\bar{n}_{\text{об}} + \bar{n}_{\text{ф}}}{2}}} M(v_x, v_y) N(f),$$

где величина $(S/N)_{\text{воспр.пор}}$ принимается равной 2,5.

Формально следует, что при стремлении МПФ к нулю значение МРК неограниченно возрастает, но в соответствии с определением контраста, его значение не может превышать 1.

Это кажущееся противоречие легко устраняется, если принять во внимание, что произведение $k_{\min} T_{\text{лсв}}$ представляет собой воспринимаемый контраст изображения на пороге обнаружения объекта, который, так же как МРК, изменяется в пределах от 0 до 1. При фиксированной интенсивности излучения подсвета увеличение пространственных частот приводит к одновременному снижению уровня системной МПФ и росту МРК. Максимальные значения пространственных частот, при которых значение МРК достигает 1 (без учета влияния канала распространения излучения), являются частотами среза ЛСВ. На практике часто задают минимально допустимый уровень МПФ системы $T_{\text{лсв}}$ и по данному уровню (обычно 0,1) находят значения частот среза и максимально допустимого МРК. Если в этом случае расчетное значение МРК превысит 1, то это означает, что контраст изображения недостаточен для порогового обнаружения и, следовательно, необходимо повысить уровень $T_{\text{лсв}}$ до требуемого значения.

Зависимость наблюдаемого контраста объекта от дальности действия выражается соотношением

$$k_{\text{н}}(z) = \frac{k_{\text{об}} T_{\text{лсв}}(z)}{1 + 2n_{\text{пор}} / (n_{\text{об}} + n_{\text{ф}})},$$

где $n_{\text{об}}$, $n_{\text{ф}}$, $n_{\text{пор}}$ – числа фотоэлектронов, регистрируемых одним пикселем ФПЗС матрицы от объекта, фона и помехи обратного рассеяния, рассчитываемые по соотношениям, приведенным в [1].

Для определения предельной дальности работы ЛСВ необходимо наложить график наблюдаемого контраста объекта от дальности действия ЛСВ при обнаружении и распознавании на график зависимостей минимального контраста от дальности действия ЛСВ. Предельная дальность обнаружения $z_{\text{обн}}$ (или распознавания) будет определяться абсциссой точки пересечения кривых $k_{\text{н}}(z)$ и $k_{\text{min}}(z)$.

3. ФОРМИРОВАНИЕ И ОБРАБОТКА ИЗОБРАЖЕНИЙ В ЭЛЕКТРОННОМ ТРАКТЕ ЛАЗЕРНЫХ СИСТЕМ ВИДЕНИЯ

3.1. Фильтрация, дискретизация и квантование изображений в электронном тракте

При разработке лазерных систем видения нельзя не учитывать тот фактор, что особенности восприятия человеком изображений далеко не всегда могут быть описаны строго математически в силу того, что сам процесс восприятия достаточно сложен.

Для того чтобы сформировать правильное, удобное для анализа изображение, крайне важно определить, какие детали изображения мозг воспринимает особенно четко, а какие являются второстепенными для анализа. Следует сразу сказать, что однозначный ответ вряд ли возможен: известно, что восприятие изображения в сильной степени зависит от развития индивидуума, от задачи, которую ставит перед собой наблюдатель. Хорошо известно, что человек, собирающий грибы, быстро настраивается на поиск круглой шляпки среди овальных листьев. Точно так же направленно обучается оператор, ведущий по экрану телевизора отбор бракованных деталей конвейера. При этом, чем меньше лишних деталей на изображении, шумов, чем четче выделены границы объектов, тем проще работать оператору лазерной системы видения.

В результат считывания, квантования сигналов с каждого пикселя ФПЗС-матрицы мы получаем двумерное представление распределения, которое в дальнейшем будем называть *изображением* или *кадром ФПЗС-матрицы*.

Напомним, что формирование изображения любого объекта производится в результате отражения от него энергии источников светового излучения и последующей регистрации его светочувствительным элементом, обладающим свойством спектральной избирательности, т. е. осуществляющим фильтрующее действие. Обозначим кривую спектральной чувствительности ФПЗС-матрицы функцией $s(\lambda)$. Распределение внешней освещенности на ФПЗС-матрице $E(x_1, x_2, \lambda)$, строго говоря, представляет собой большой набор спектральных составляющих отраженной или выделяемой самим объектом энергии, которые доходят до матрицы. Фоточувствительный приемник обладает спектральной избирательностью, может регистрировать только энергию только в конечном диапазоне от минимальной и максимальной, а также имеет определенные размеры, т. е. в результате своей работы ограничивает внешний по-

ток отраженной энергии по диапазону значений, по спектру и по пространству.

Регистрируемое изображение является результатом усреднения ограниченной в диапазоне значений функции $E_{\text{орг}}(x_1, x_2, \lambda)$ по диапазону длин волн от λ_{min} до λ_{max} с весовой функцией $s(\lambda)$ и может быть определено по соотношению

$$x_{i,j} = \int_{\lambda_{\text{min}}}^{\lambda_{\text{max}}} E_{\text{орг}}(i = x_1, j = x_2) s(\lambda) d\lambda,$$

где i, j – координаты пикселей на матрице, которые соответствуют пространственным координатам x_1, x_2 .

Если для формирования изображения используется лазерная система видения и при этом ФПЗС-матрица имеет ярко выраженный максимум спектральной чувствительности, сопряженный со значением длины волны источника подсвета (например, в инфракрасном диапазоне), то на исходном изображении в качестве существенной паразитной засветки может выступать только та часть солнечной энергии, которая соответствует данной области спектра. Тогда для заданной длины волны подсвета распределение полученной в итоге яркости на изображении переписывается в виде

$$x_{i,j} = s E_{\text{орг}}(i = x_1, j = x_2),$$

где s – величина постоянного множителя, соответствующего спектральной чувствительности для рабочей частоты ЛСВ.

Наличие шумов, формируемых ФПЗС-матрицей, приводит к тому, что изображение искажается. Шум на изображении проявляется в виде случайным образом расположенных элементов с размерами, которые определяются размером отображаемого на экране пикселя ФПЗС-матрицы. Неоднородности, возникающие при изготовлении МДП структуры матрицы, которые приводят к утечке накопленного заряда, а также дефектные (неработающие) ячейки проявляются в виде темных точек, а шумы проявляются в виде светлых точек. Таким образом, исходное изображение «Фотоаппарат» (рис. 30, *а*) превращается в «зашумленное» (рис. 30, *б*) с большим количеством светлых и темных точек. Данный «снег» на экране мешает восприятию мелких деталей на изображении «Фотоаппарат» и даже, если оператору удастся обнаружить изображение – определить наличие корпуса фотоаппарата на указанном примере, то он не может идентифицировать его марку – прочесть его название.

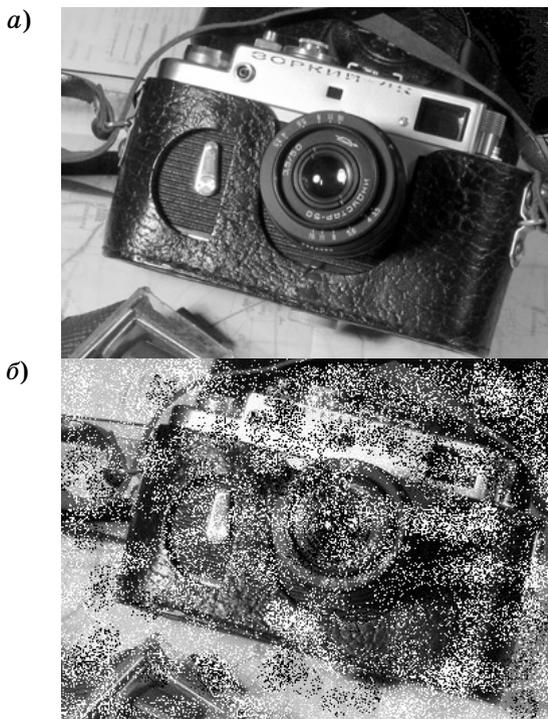


Рис. 30. Исходное (а) и искаженное (б) наличием шумов изображение

Таким образом, серьезно ухудшаются характеристики возможного обнаружения и распознавания. То же самое происходит и в лазерных системах видения. Например, при реализации наблюдения за дорогой, ведущей к охраняемой территории при помощи лазерной системы видения можно обнаружить наличие движущегося автомобиля, но затруднительно распознать его тип в силу зашумленности мелких деталей на изображении.

Для улучшения такого изображения обычно применяют цифровые методы обработки изображений с помощью сигнальных процессоров, входящих в состав аппаратуры обработки сигнала в приемном тракте ЛСВ. Теория и практика цифровых методов обработки изображений достаточно сильно развита. Связано это с большим накопленным опытом в процессе практического использования данной области. При этом под **цифровой обработкой изображений** понимается не только улучшение зрительного восприятия изображений, но и классификация объектов, выполняемая при анализе

изображений. В 60-е годы прошлого века получила развитие особая наука об изображениях – «*иконика*», которая посвящена исследованиям общих свойств изображений, целей и задач их преобразования, обработки и воспроизведения, распознавания графических образов. Термин «иконика» происходит от греческого «*eikon*», что означает изображение, образ. Сегодня под ним понимают создание и обработку изображений с помощью ЭВМ, пользуясь зачастую понятием компьютерной обработки изображений [10]. Методы цифровой обработки широко применяются в промышленности, искусстве, медицине, космосе. Формирование изображений, улучшение качества и автоматизация обработки медицинских изображений, включая изображения, создаваемые электронными микроскопами, рентгеновскими аппаратами, томографами и т.д., являются предметом исследования и разработки. Сегодня в технике широко применяются системы формирования изображений, его преобразования в цифровую форму, визуализация и документирование путем введения в компьютер изображений с помощью специализированных устройств захвата видео. Автоматический анализ в лазерных системах дистанционного наблюдения широко применяется при анализе местности в целях обеспечения безопасности объектов, при реализации контроля качества производимой продукции выполняется благодаря автоматическим методам анализа сцен, при проведении поиска определенных затонувших объектов или контроле при реализации подводных строительных работ. Рассмотрим на примере фильтрации шумов изображений, как с помощью цифровых методов обработки можно достичь улучшения визуализации анализируемых оператором сцен.

Для того чтобы очистить цифровое изображение от шумов применяют нелинейную пространственную фильтрацию. Чаще всего для устранения шумов используется медианный фильтр. *Медианный фильтр* представляет собой скользящее двумерное окно заданных размеров («маска») с нечетным количеством элементов в окне, которое заменяет центральный элемент «маски» медианой упорядоченных по возрастанию (убыванию) значений яркости. На рис. 31 представлен результат медианной фильтрации изображения, представленного на рис. 30, б. Медианная фильтрация произведена скользящим окном размером 3×3 элементов изображения.

Следует отметить, что после проведения медианной фильтрации шумы значительно уменьшились, но при этом само изображения утратило четкость. Искажения изображения могут также возникнуть и при его дискретизации.



*Рис. 31. Результат применения
медианной фильтрации*

Погрешность дискретизации изображения определяется посредством сравнения восстановленного сигнала подвергнутого дискретизации и исходного. Дискретизация исходного изображения в виде распределения освещенности происходит непосредственно при получении цифрового изображения на ФПЗС-матрице.

В том случае, когда спектр сигнала узкий, и выполняется теорема Котельникова, то лучше всего, когда изображение поступает с матрицы, которая обладает большими светочувствительными элементами, поскольку чувствительность элемента матрицы пропорциональна ее площади.

Современные матрицы высокой чувствительности имеют разрешение 796×582 , размещаются в корпусе и выполняют функцию обработки и формирования видеосигнала. Более того, изображения такого размера удобно визуализировать непосредственно на экране монитора или дисплея.

Однако, если необходимо принять изображение, которое обладает широким эффективным спектром, то использование матрицы, частота дискретизации которой не удовлетворяет теореме Котельникова, способно приводить к искажениям изображения, называемых «муаром». Рассмотрим пример, поясняющий возникновение искажений изображения. Пусть спектр функции $F_1(\omega_x)$ ограничен интервалом $[-\pi/\Delta x; \pi/\Delta x]$ (рис. 32, а).

Дискретизация с шагом Δx приводит к периодическому размножению спектров $F(\omega_x)$ (рис. 32, б). Полоса идеального низкочастотного фильтра $H(\omega_x)$ совпадает с полосой сигнала (рис. 32, в), а ре-

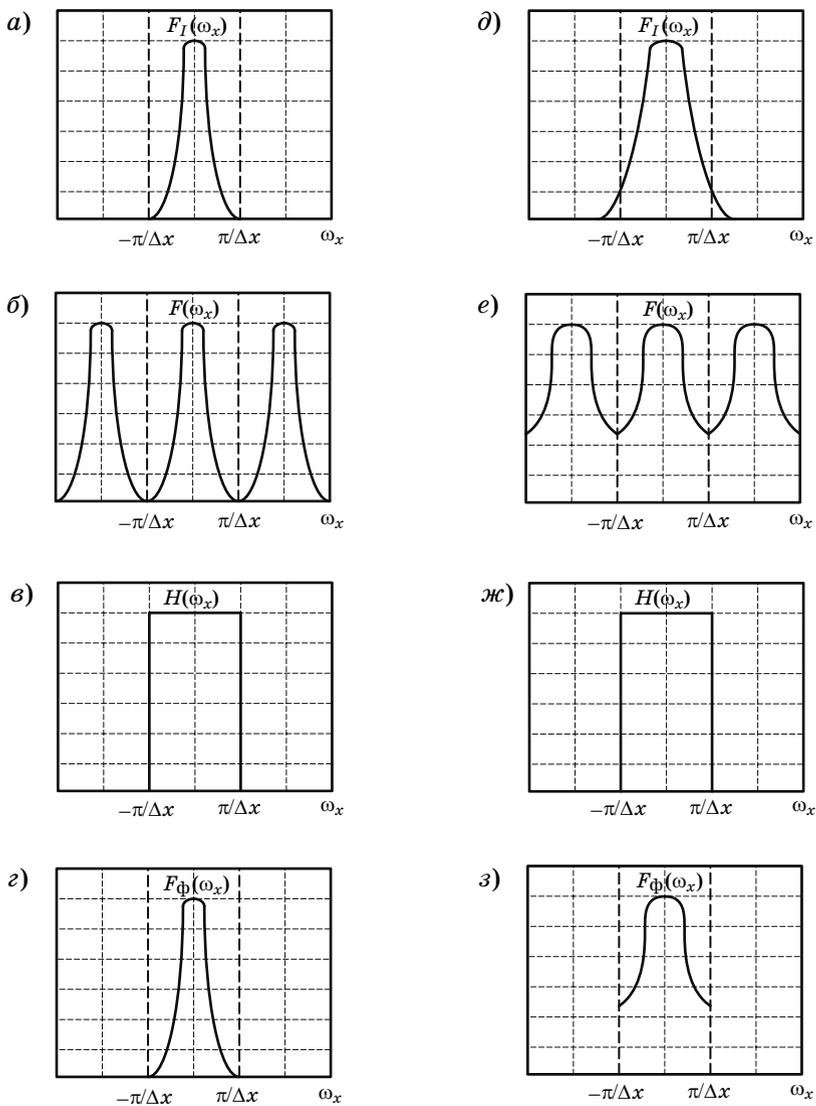


Рис. 32. Результат применения полосовой низкочастотной фильтрации в случае согласования спектра сигнала и спектра фильтра (а-г), и в случае, когда спектр сигнала шире спектра полосового фильтра (д-з)

зультат фильтрации спектра $F_{\Phi}(\omega_x)$ (рис. 32, *г*) говорит о том, что спектр сигнала $F_I(\omega_x)$ не изменился.

А вот если полоса того же низкочастотного фильтра меньше, чем ширина спектра изображения, то ситуация изменяется так, как это показано на рис. 32, *д, е, ж, з*, соответственно.

При такой дискретизации появляется наложение высокочастотных частей спектра, что приведет к тому, что полученный в итоге спектр изображения (рис. 32, *з*) будет отличаться от исходного.

Приведем пример, того, как неправильная дискретизация может сказаться на изображении. На рис. 33, *а*, представлено изображение «Отпечаток пальца», восстановленное при выполнении правильной дискретизации, а на рис. 33, *б* изображение, восстановленное для случая, когда частота дискретизации в три раза меньше требуемой [18].

Кроме того, полученное после ФПЗС-матрицы изображение в виде значений напряженности, пропорциональных внешней освещенности фоточувствительных элементов матрицы, подвергается операции квантования в аналогово-цифровом преобразователе для получения оцифрованного изображения, порождая шум квантования. Следует отметить, что качество изображения зависит от разрядности АЦП, осуществляющего квантование. Например, на рис. 34, *а* и 34, *б* представлены результаты квантования изображения «Рыбачий бастион», осуществляемого восьми и четырехрядным АЦП. Как видно при недостаточной разрядной сетке АЦП на изображении появляются резкие контуры, которые бывают иногда и «ложными» [19]. В [19] также показано, что устранить лож-

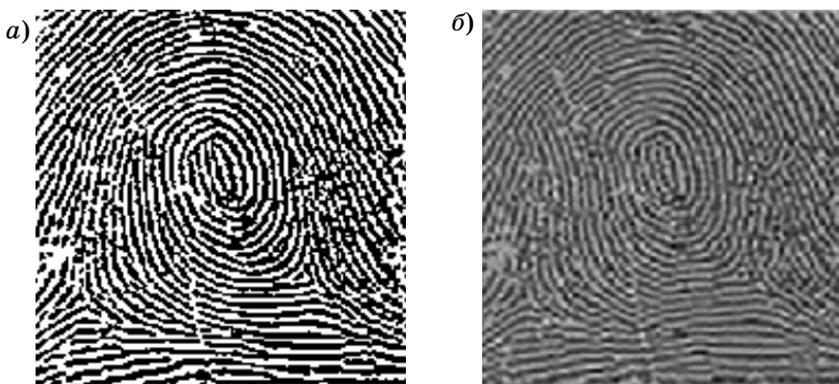


Рис. 33. Пример правильной (*а*) и неправильной (*б*) дискретизации изображения

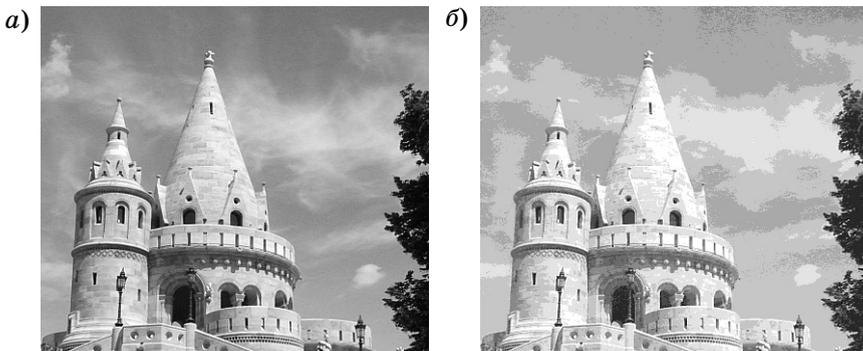


Рис. 34. Пример шестнадцатиразрядного (а) и восьмиразрядного (б) квантования изображения

ные контуры можно добавлением в изображение шума с равномерной плотностью вероятности и столь малым значением среднеквадратического отклонения, которое не вызывало эффект появления заметного влияния шума на изображение.

Для улучшения качества воспроизводимого на дисплее оператора изображения можно использовать различные алгоритмы цифровой обработки, позволяющие повысить его контрастность, выявить контуры или устранить искажения, вызванные неравномерностью внешней освещенности на ФПЗС-матрице ЛСВ.

3.2. Контрастирование изображений. Бинаризация

В большом числе систем видения применяется представление результатов обработки данных в виде изображения, выводимого на экран для использования наблюдателем. Процедуру, обеспечивающую такое представление результатов обработки, называют визуализацией. Желательно при помощи обработки придать выводимому изображению такие качества, благодаря которым его восприятие человеком было бы по возможности комфортным. В лазерных системах видения, предназначенных для работы в темноте или плохих метеоусловиях, обычно используют лазерный подсвет ближнего инфракрасного диапазона. Отражательные характеристики различных материалов в ближней ИК-области и в видимом диапазоне весьма близки, поэтому общий характер изображения практически повторяет изображение видимого диапазона. Это приводит к тому, что результаты применения методов, предназначенных для улучшения качества цифровых монохромных фотографических сним-

ков, можно использовать и для демонстрации их эффективности при использовании аналогичной обработки изображений в лазерных системах видения, что и используется в дальнейшем при изложении материала.

Таким образом, для улучшения качества воспроизводимого на дисплее оператора изображения можно использовать различные процедуры их обработки, которые часто используются в фотографической практике. Часто бывает полезным подчеркнуть, усилить какие-то черты, особенности, нюансы наблюдаемой картины с целью улучшения ее субъективного восприятия. Последнее – субъективность восприятия – сильно усложняет применение формализованного подхода в достижении данных целей. Поэтому при обработке изображений для визуализации получили распространение методы, в которых часто отсутствуют строгие математические критерии оптимальности. Их заменяют качественные представления о целесообразности той или иной обработки, опирающиеся на субъективные оценки результатов. Подавляющее большинство процедур обработки для получения результата в каждой точке кадра привлекает входные данные из некоторого множества точек исходного изображения, окружающих обрабатываемую точку. Однако имеется группа процедур, где осуществляется так называемая поэлементная обработка. Здесь результат обработки в любой точке кадра зависит только от значения входного изображения в этой же точке. Очевидным достоинством таких процедур является их предельная простота. Вместе с тем, многие из них приводят к очевидному субъективному улучшению визуального качества. Этим определяется внимание, которое уделяется поэлементным процедурам. Не преувеличивая их роли, отметим, что очень часто поэлементная обработка применяется как заключительный этап при решении более сложной задачи обработки изображения. Сущность поэлементной обработки изображений сводится к следующему. Пусть $x(i,j) = x_{i,j}$, $y(i,j) = y_{i,j}$ – значения яркости исходного и получаемого после обработки изображений соответственно в точке кадра, имеющей декартовы координаты i (номер строки) и j (номер столбца). Поэлементная обработка означает, что существует функциональная однозначная зависимость между этими яркостями $y_{i,j} = f_{i,j}(x_{i,j})$, позволяющая по значению исходного сигнала определить значение выходного продукта. В общем случае, как это учтено в данном выражении, вид или параметры функции $f_{i,j}(\cdot)$, описывающей обработку, зависят от текущих координат. При этом обработка является неоднородной. Однако в большинстве прак-

тически применяемых процедур используется *однородная поэлементная обработка*. В этом случае индексы i и j в описании функции, осуществляющей обработку могут отсутствовать. При этом зависимость между яркостями исходного и обработанного изображений описывается функцией: $y = f(x)$, одинаковой для всех точек кадра.

Для выделения изображения объектов на определенном фоне необходимо наличие контраста объекта и фона по коэффициенту отражения для падающих электромагнитных волн. Наличие такого контраста можно подчеркнуть при визуализации, сделав наблюдаемый на дисплее оператор контраст выше исходного. Простейшим процессом обработки изображения с целью повышения наблюдаемого контраста является линейное контрастирование.

Выполнение операции *контрастирования* необходимо для улучшения согласования динамического диапазона изображения и экрана, на котором выполняется визуализация. Если для цифрового представления каждого отсчета изображения отводится 1 байт (8 бит) запоминающего устройства, то входной или выходной сигналы могут принимать одно из 256 значений. Обычно в качестве рабочего используется диапазон 0–255; при этом значение 0 соответствует при визуализации уровню черного, а значение 255 – уровню белого. Предположим, что минимальная и максимальная яркости исходного изображения равны x_{\min} и x_{\max} соответственно. Если эти параметры или один из них существенно отличаются от граничных значений яркостного диапазона, то визуализированная картина выглядит как ненасыщенная, неудобная, утомляющая при наблюдении. Пример такого неудачного представления приведен на рис. 35, а, где диапазон яркостей имеет границы $x_{\min} = 104$, $x_{\max} = 254$. При линейном контрастировании используется линейное поэлементное преобразование вида: $y = ax + b$, параметры которого и определяются желаемыми значениями минимальной и максимальной выходной яркости. Решив систему уравнений:

$$\begin{cases} y_{\min} = ax_{\min} + b \\ y_{\max} = ax_{\max} + b \end{cases}$$

относительно параметров преобразования a и b , нетрудно привести поэлементное преобразование к виду:

$$y = \frac{x - x_{\min}}{x_{\max} - x_{\min}}(y_{\max} - y_{\min}) + y_{\min}.$$



Рис. 35. Пример результата применения операции линейного контрастирования (а – до контрастирования, б – после контрастирования)

Пример реализации линейного контрастирования изображения «Башня», представленного на рис. 35, а, приведен на рис. 35, б при следующих параметрах $y_{\min} = 12$ и $y_{\max} = 255$. Сравнение двух изображений на представленном рисунке свидетельствует о значительно лучшем визуальном качестве обработанного изображения. Улучшение связано с использованием после контрастирования большего динамического диапазона, который значительно меньше у исходного изображения.

Так, например, нелинейное преобразование яркости может способствовать формированию сильного контраста в области малых яркостей изображения и слабого – при высоких значениях яркости элементов изображения. Одной из часто встречающихся операций обработки изображений по расширению входного динамического диапазона для визуализации является *гамма-коррекция*, которая осуществляется по следующей формуле:

$$y_{i,j} = f(x_{i,j}) = c(x_{i,j} + f_0)^\gamma,$$

где c , f_0 , γ представляют собой неотрицательные числовые параметры.

Так, на рис. 36, а представлено исходное изображение «Дом Батло», а на рис. 36, б – визуализация «Дом Батло» после применения гамма-коррекции с параметрами $c = 1$, $f_0 = 0$, $\gamma = 1,25$. Следует от-



Рис. 36. Пример результата применения операции гамма-коррекции для расширения динамического диапазона (а – исходное изображение; б – после коррекции)

метить, что некорректное применение параметров гамма-коррекции может привести к ухудшению исходного изображения. Так, на рис. 37 представлено изображение «Дом Батло», подвергнутое гамма-коррекции с недопустимо высоким параметром $\gamma = 1,5$, при котором расчетная яркость некоторых элементов изображения превышает значение 255.

Данный пример показывает, что использование нелинейных преобразований для обработки изображений требует четкого определения допустимых диапазонов значений для реализации процедур обработки изображений.

Например, выбор максимально допустимого параметра γ для процедуры гамма-коррекции при заданном параметре c может быть осуществлен следующим образом:

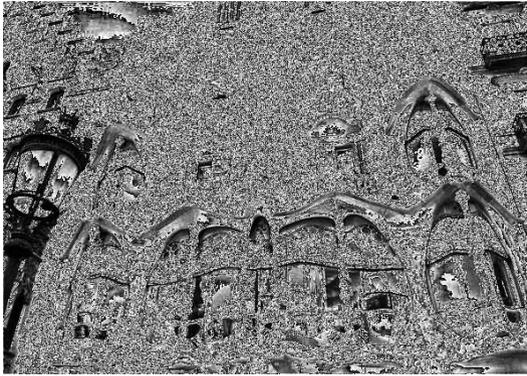


Рис. 37. Изображение «Дом Батло» после применения операции гамма-коррекции с неправильно выбранными параметрами

а) вначале оценивается максимальная и минимальная яркости входного изображения $x_{\max} = \max(x_{i,j})$ и $x_{\min} = \min(x_{i,j})$ и определяется их разность $\Delta x = x_{\max} - x_{\min}$;

б) значение f_0 принимается равным $f_0 = -x_{\min}$;

в) задается максимально допустимое значение яркости элемента изображения после применения гамма-коррекции y_{\max} ;

г) для заданного c вычисляется максимально допустимое значение по формуле $\gamma = \log_{\Delta x}(y_{\max}/c)$.

Например, в исходном изображении «Дом Батло» параметры: $x_{\max} = 83$, $x_{\min} = 0$, а выходная максимальная яркость $y_{\max} = 255$. При выбранном параметре $c = 1$ получаем, что максимально допустимое значение $\gamma = 1,254$.

Такой параметр как динамический контраст может сильно варьироваться в зависимости от предустановленных в специальном программном обеспечении параметров функции, осуществляющей преобразование контраста, и потому не может служить критерием качества визуализации контраста, который может быть непосредственно отображен на экране монитора или телевизора. Поэтому для оценки реального динамического диапазона яркостей, который может быть воспроизведен на экране монитора следует использовать только значение контраста (часто называемого статическим). На практике для визуализации изображений в ЛСВ следует использовать полный динамический диапазон для изображения, выводимого на дисплей оператора. В современных мониторах, дисплеях и индикаторах принято 8-битное отображение, т. е. значения яркости любого пикселя может изменяться в диапазоне значений от 0 до 255.

Для выделения объекта на фоне можно также пользоваться определенной пороговой процедурой, которая представляет визуализируемое изображение всего в двух градациях по яркости: нулевой и максимальной, которая называется бинаризацией. Данная поэлементная процедура выполняется следующим образом: если яркость элемента изображения превышает заданное значение, например равное 128, то ей присваивается максимальное значение – 255. Если не превышает, то присваивается значение, равное нулю.

На рис. 38 приведены результаты бинаризации изображения «Башня», которая была произведена без применения предварительной процедуры линейного контрастирования (рис. 38, а) и бинаризации того же изображения после контрастирования (рис. 38, б). Полученные результаты показывают, что процедуру бинаризации изображения необходимо производить после контрастирования, т. е. после расширения динамического диапазона визуализируемого изображения.

Необходимо отметить, что получение высокого контраста итогового изображения в лазерных системах видения с уменьшением шумов возможно при наличии малого обратного рассеяния и сла-

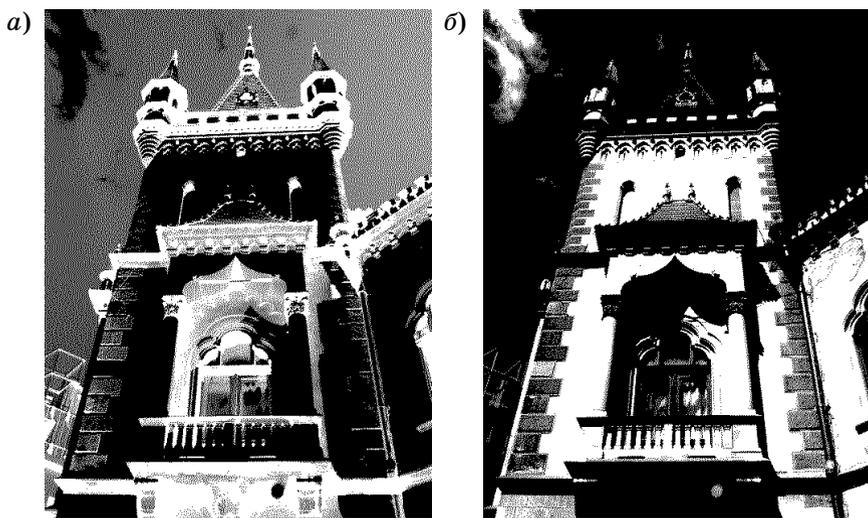


Рис. 38. Пример результата применения операции бинаризации (а – без предварительного контрастирования; б – при наличии предварительного контрастирования)

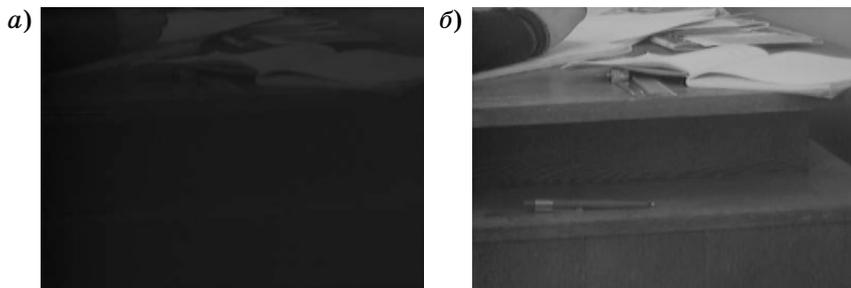


Рис. 39. Рисунок демонстрирующий улучшение изображения при росте времени накопления на ФПЗС-матрице (а – время накопления 0,01 с; б – время накопления 0,5 с)

бой сторонней засветки только посредством накопления сигнала на ФПЗС-матрице.

Так, на рис. 39, а и 39, б представлено изображение, полученное посредством накопления в низких условиях освещенности, заимствованное из [10].

Как видно из представленного изображения рост времени накопления сигнала позволяет улучшить контраст неподвижного изображения. Однако для движущегося изображения рост времени накопления приводит к таким негативным явлениям как «размытие» границ объекта и фона, утрачивание мелких его деталей, т. е. ухудшению условий распознавания.

3.3. Компенсация неравномерной засветки ФПЗС-матрицы посредством применения алгоритма выравнивания освещенности изображения

Помимо отраженного от объекта и фона сигнала лазерного подсвета в формировании внешнего распределения освещенности в плоскости ФПЗС-матрицы участвуют и другие источники излучения: чаще всего солнечная или лунная засветка. При работе лазерных систем видения данное освещение, как было показано ранее, порождает помеху обратного рассеяния и более того в процессе накопления полезного сигнала может непосредственно быть зафиксировано на матрице. Для того чтобы избежать эффекта прямого попадания стороннего света на матрицу применяют защитные бленды.

Бленда представляет собой тонкостенную полую насадку конической, пирамидальной или цилиндрической формы или более сложную лепестковую конструкцию для предотвращения затемне-

ния изображения по краям кадра, которое называется *виньетирование*. Последнее характеризует как ослабление проходящего под углом по отношению к оптической оси потока лучей в оптической системе, так и затемнение части изображения из-за различных преград на пути света, таких как затемнение блендой.

Практически, из-за разницы в угле зрения, диаметре передней линзы объектива, для каждого из них необходимо проектировать или подбирать бленду индивидуальной формы.

Рассмотрим идеализированный случай, при котором центр симметрии объекта (по горизонтали и вертикали) с коэффициентом отражения, равным единице, расположен на оптической оси лазерной системы видения. Предположим, что при этом объект имеет такие геометрические размеры, которые позволяют ему занимать большую часть кадра. При этом фон в виде квадрата имеет малый коэффициент отражения, который равен 0,1.

Идеальное оцифрованное монохромное изображение данной картины показано на рис. 40, а. Пусть источник лазерного подсвета, направленный в рассматриваемом случае на центр симметрии зондируемого объекта, имеет диаграмму направленности гауссовой формы. Тогда объект в этом случае подсвечивается неравномерно, что совместно с виньетированием приводит к получению изображения представленного на рис. 40, б.

Изображение, показанное на рис. 40, б, можно привести к изображению на рис. 40, а следующим образом – добавить яркости всему изображению, а потом произвести бинаризацию, подбирая порог так, чтобы получить исходное изображение.

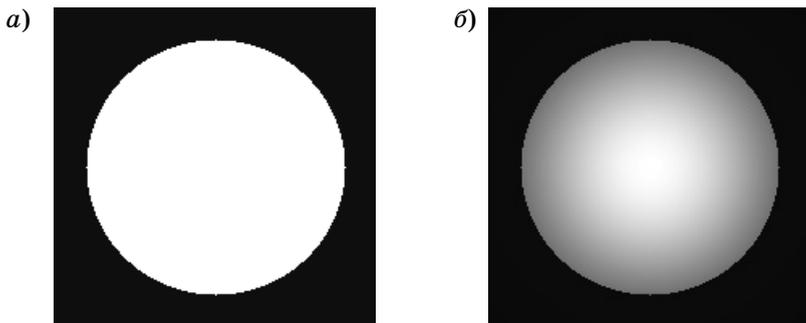


Рис. 40. Влияние неравномерности лазерного подсвета и виньетирования на получение изображения

Иную ситуацию будем иметь, когда малоразмерный объект наблюдается на неравномерном фоне, состоящем из множества объектов другой природы с близкими к объекту коэффициентами отражения. Если при этом сам объект расположен не по максимуму диаграммы направленности лазерного подсвета, то необходимо решить задачу выделения слабоконтрастного объекта из кадра с неравномерным освещением.

Чтобы улучшить вид изображения в таких случаях, применяется *метод выравнивания освещенности*.

Освещенность меняется в пространстве достаточно медленно и ее можно считать низкочастотным сигналом. Само же изображение можно считать в среднем более высокочастотным сигналом. Если бы в процессе получения изображения эти сигналы складывались, то их можно было бы разделять с помощью обычных фильтров. Однако на ФПЗС-матрице получается произведение той «исходной картины», которую мы хотим видеть, и «карты» освещенности [20]. И поскольку эти сигналы не складываются, а перемножаются, то избавиться от неравномерностей освещенности простой фильтрацией не удастся.

Для решения таких задач применяется гомоморфная обработка. Идея обработки заключается в сведении нелинейной задачи к линейной. Для этого производится нелинейная обработка исходного сигнала для представления его в форме сложения двух сигналов. Для ее реализации необходимо взять логарифм от произведения изображений «исходной картины» и «карты» освещенности, который будет равен сумме логарифмов сомножителей. При этом задача разделения произведения сигналов сводится к задаче разделения суммы низкочастотного и высокочастотного сигналов и решается с помощью высокочастотного фильтра, который удалит из суммы сигналов низкие частоты. Останется взять от полученного сигнала экспоненту, чтобы вернуться к исходному масштабу амплитуд.

Высокочастотный фильтр можно реализовать следующим образом. Сначала к изображению применяется операция размытия, т. е. осуществляется операция низкочастотной фильтрации, а потом из исходного изображения вычитается «размытое». Для размытия изображения можно применить двумерный гауссовский фильтр, так как это было показано в разд. 2.2.

Эквивалентного эффекта можно достичь, отфильтровав одномерным гауссовым фильтром сначала строки изображения, а затем столбцы полученного изображения. Полученный от выравнивания освещенности эффект может оказаться слишком сильным (тем-

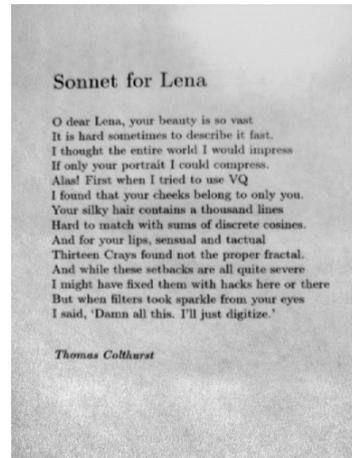
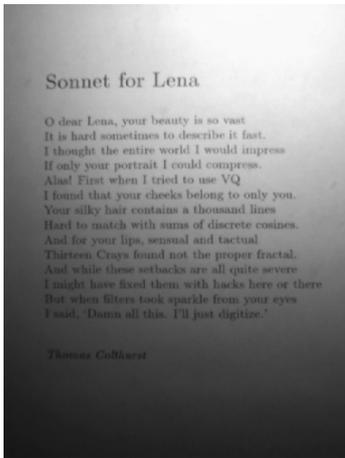


Рис. 41. Выравнивание освещенности изображения

ные области станут по яркости такими же, как и светлые). Чтобы уменьшить эффект и добиться лучшей визуализации изображения можно смешать обработанное изображение с исходным в определенной пропорции.

Результат применения алгоритма для компенсации неравномерности освещенности изображения показан на рис. 41. Как следует из представленных изображений, использование алгоритма выравнивания освещенности позволяет повысить контраст тех объектов (в данном случае – букв), которые изначально располагаются в затененной части.

Тем не менее, при применении лазерных систем видения необходимо задумываться о том, каким образом будет производиться обзор заданной области пространства и при каких условиях, чтобы избегать по мере возможности наличия сторонней солнечной заливки. Также следует стремиться к подсвету объектов максимум диаграммы направленности источника лазерного подсвета и делать данный подсвет как можно более равномерным.

3.4. Алгоритмы выделения границ объектов на изображениях

Исследованиями психологов установлено, что с точки зрения распознавания и анализа объектов на изображении наиболее информативными являются не значения яркости объектов, а характеристики их границ – контуров. Таким образом, основная инфор-

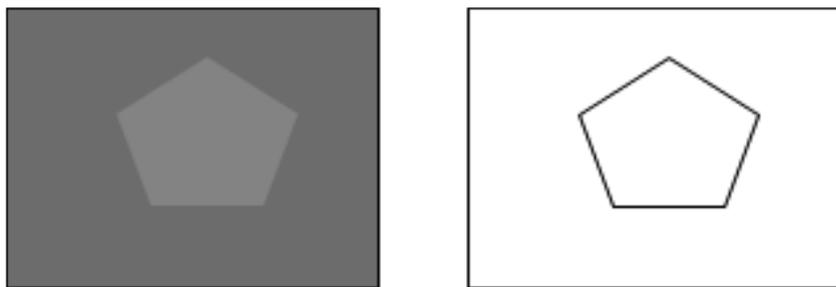


Рис. 42. Переход от слабоконтрастного изображения к контурному

мация заложена в очертаниях областей. Особенно хорошо это проявляется при малых контрастах объекта и фона. Переход от контрастно-яркостного к контурному иллюстрирует рис. 42.

Поэтому в практических задачах видения, предназначенных для обнаружения слабоконтрастных по отношению к фону объектов и их распознавания, можно предоставить оператору не только яркостное изображение, но и контурное, используя процедуры, которые позволяют выделить из полученного зарегистрированного изображения границы объекта для его визуальной идентификации [21].

С точки зрения распознавания и анализа объектов на изображении наиболее информативными являются не значения яркостей объектов, а характеристики их границ – контуров. Другими словами, основная информация заключена не в яркости отдельных областей, а в их очертаниях. Задача *выделения контуров* состоит в построении изображения именно границ объектов и очертаний однородных областей.

Как правило, граница предмета отображается перепадом яркости между двумя сравнительно однотонными областями. Кроме этого перепады яркости могут быть вызваны также текстурой объекта, тенями, бликами, перепадами освещенности и т.п.

Будем называть контуром изображения совокупность его пикселей, в окрестности которых наблюдается скачкообразное изменение функции яркости. Так как при цифровой обработке изображение представлено как функция целочисленных аргументов, то контуры представляются линиями шириной, как минимум, в один пиксель. Если исходное изображение, кроме областей постоянной яркости, содержит участки с плавно меняющейся яркостью, то непрерывность контурных линий не гарантируется, что иллюстрируется рис. 43, заимствованным из [22]. С другой стороны, если на «кусочно-постоянном» изображении присутствует шум, то могут

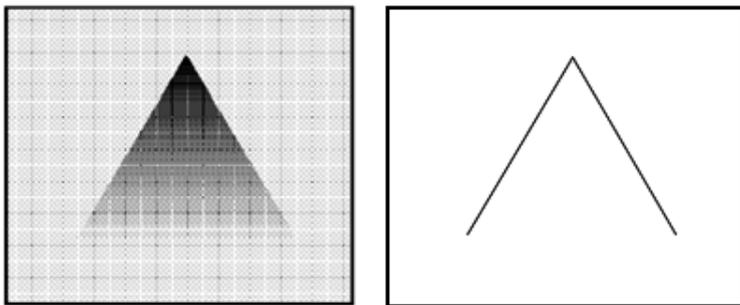


Рис. 43. Иллюстрация потери «замкнутости» объекта при реализации формирования контура

быть обнаружены «лишние» контуры в точках, которые не являются границами областей.

При разработке алгоритмов выделения контуров нужно учитывать указанные особенности поведения контурных линий. Специальная дополнительная обработка выделенных контуров позволяет устранять разрывы и подавлять ложные контурные линии.

Процедура построения бинарного изображения границ объектов обычно складывается из двух последовательных операций: выделения контуров и их пороговой обработки.

Одним из самых простых способов выделения границ на изображении является градиентный метод [21]. Для одномерной непрерывной функции изменения яркости $F(x)$ выделение границ вдоль оси Ox данным методом иллюстрируется рис. 44.

В задаче обработки изображения требуется выделить контуры, направление которых произвольно на плоскости xOy . Для данной цели можно использовать *модуль градиента функции яркости*

$$|f(x, y)| = \sqrt{\left(\frac{\partial f(x, y)}{\partial x}\right)^2 + \left(\frac{\partial f(x, y)}{\partial y}\right)^2},$$

который пропорционален максимальной скорости изменения функции яркости в точке. Модуль градиента в отличие от частной производной принимает только неотрицательные значения, поэтому контуры представляются точками с повышенным уровнем яркости. Для цифровых изображений расчет градиента производится по формуле

$$y_{i,j} = |f_{i,j}| = \sqrt{(x_{i,j} - x_{i-1,j})^2 + (x_{i,j} - x_{i,j-1})^2}.$$

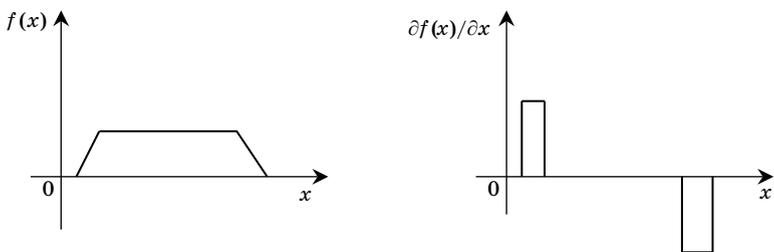


Рис. 44. Градиентный метод выделения границ по изменению яркости

Таким образом, операция выделения контуров заключается в выполнении нелинейной локальной обработки изображений скользящим «окном» 2×2 без одной точки, который называется *апертурой*.

Практический пример выделения границ изображения «Бурундук» приведен на рис. 45. Исходное изображение (а) является одно-тонным. На изображении (б) представлен результат вычисления вектора градиента яркости. Отфильтровав пиксели с длиной градиента, большей определенного порога, мы получим изображение границ (в).

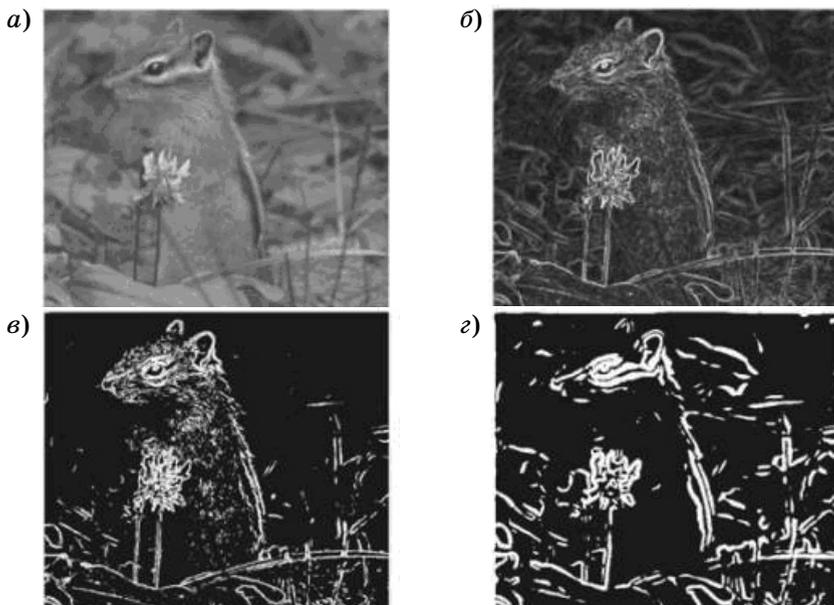


Рис. 45. Выделение границ объектов на малоконтрастном изображении

Недостаток алгоритма – пропуск границ с малыми перепадами яркости и включение в число границ деталей изображения с большими изменениями яркости, в качестве которых могут выступать и шумы изображения. Это происходит вследствие того, что не учитывается, что граничные точки соответствуют не просто перепадам яркости, а перепадам яркости между относительно монотонными областями.

Для снижения влияния данного недостатка изображение сначала подвергают сглаживающей гауссовой фильтрации. При этом мелкие несущественные детали размываются быстрее перепадов между областями. Затем производят операцию выделения контуров, результат которой представлен на рис. 45, *г* и показывает, что при такой обработке пропадают мелкие детали изображения, а четко выраженные границы расплываются в жирные линии.

Существуют и другие распространенные алгоритмы выделения контуров изображений. Рассмотрим принципы работы и результаты применения самых распространенных из них. Так, часто для простого вычисления дискретного градиента осуществляют **преобразование Робертса** [21]:

$$y_{i,j} = \sqrt{(x')^2 + (x'')^2}, \quad (3.1)$$

где $x' = x_{i,j} - x_{i+1,j+1}$; $x'' = x_{i,j+1} - x_{i+1,j}$.

Реализацию данной процедуры можно пояснить следующим наглядным примером. Процедуру преобразования изображения можно представить в виде обработки каждого пикселя двумя двумерными апертурами, общий вид которых приведен на рис. 46, *а*. Первая двумерная апертура формирует разность x' , а вторая – x'' .

а)

$A \cdot c1$	$C \cdot c2$
$B \cdot c3$	$D \cdot c4$

б)

$A \cdot c1$	$B \cdot c2$	$C \cdot c3$
$D \cdot c4$	$E \cdot c5$	$F \cdot c6$
$G \cdot c7$	$H \cdot c8$	$I \cdot c9$

Рис. 46. Геометрическая интерпретация алгоритмов выделения границ изображения при применении апертур различного вида

Первая апертура применяется со следующими значениями сомножителей: $c_1 = 1, c_2 = 0, c_3 = 0, c_4 = -1$, а вторая: $c_1 = 0, c_2 = -1, c_3 = 1, c_4 = 0$ (рис. 46, б).

Тогда правило обработки пикселя изображения таким «скользящим» двумерным окном можно сформулировать так

$$A' = \sqrt{(A - D)^2 + (B - C)^2},$$

где A' – значение новой яркости в обрабатываемом элементе изображения с текущей яркостью A ; B, C, D – значения яркостей в смежных элементах изображения.

Следующим распространенным методом для применения выделения границ изображения является *метод Собеля*, использующий апертуру размером 3×3 , для получения значений x' и x'' , подставляемых в формулу (3.1).

Процедуру преобразования изображения можно представить в виде обработки каждого пикселя двумя двумерными апертурами, приведенными на рис. 46, б, которые формируют разности x' и x'' . Для получения разности x' используется апертура со значениями: $c_1 = -1, c_2 = 0, c_3 = 1, c_4 = -2, c_5 = 0, c_6 = 2, c_7 = -1, c_8 = 0, c_9 = 1$. Для получения разности x'' используется апертура: $c_1 = -1, c_2 = -2, c_3 = -1, c_4 = 0, c_5 = 0, c_6 = 0, c_7 = 1, c_8 = 2, c_9 = 1$.

Для решения задачи выделения перепадов яркости иногда используют дифференциальные операторы более высокого порядка, например *оператор Лапласа*

$$\nabla^2 f(x, y) = \frac{\partial^2 f(x, y)}{\partial x^2} + \frac{\partial^2 f(x, y)}{\partial y^2}. \quad (3.2)$$

Для случая обработки изображения данную операцию осуществляют в следующем порядке. Вначале представляют оператор Лапласа для дискретной обработки как сумму следующих компонент:

$$\frac{\partial^2 f(x, y)}{\partial x^2} \rightarrow x_{i+1, j} - 2x_{i, j} + x_{i-1, j}; \quad \frac{\partial^2 f(x, y)}{\partial y^2} \rightarrow x_{i, j+1} - 2x_{i, j} + x_{i, j-1}.$$

Тогда процедуру преобразования изображения можно представить в виде обработки каждого пикселя двумерной апертурой, приведенной на рис. 46, б, со следующим набором коэффициентов: $c_1 = 0, c_2 = 1, c_3 = 0, c_4 = 1, c_5 = -4, c_6 = 1, c_7 = 0, c_8 = 1, c_9 = 0$.

Согласно формуле (3.2) результат может принимать как положительные, так и отрицательные значения. Поскольку значение

яркости изображения является величиной положительной, то от полученного значения следует брать модуль. Таким образом, формула для вычисления новой яркости в точке изображения

$$y_{i,j} = |x_{i+1,j} + x_{i-1,j} + x_{i,j+1} + x_{i,j-1} - 4x_{i,j}|.$$

Использование одной апертуры для формирования данного изображения показывает, что в отличие от применения оператора Робертса или оператора Собеля, реализация преобразования изображения с помощью оператора Лапласа оказывается нечувствительной к ориентации границ яркости контуров на изображении, что можно считать его достоинством. Но следует учесть то, что данный метод выделения контуров не является градиентным, в отличие от представленных ранее. Различие заключается в том, что градиентные методы выделяют участки функции изменения яркости с разным наклоном, а не участки ее перегибов. Если граница размыта, то после обработки лапласианом она раздваивается, таким образом, следует проводить дополнительную обработку для поиска соответствующих раздвоенных границ и их сливанию, что негативно сказывается на вычислительной сложности алгоритма. Кроме этого, практика его применения показывает, что он также чувствителен к наличию шумов на изображении.

Таким образом, перед применением операторов Робертса, Собеля или Лапласа также следует производить сглаживание яркости в заданном окне, но учесть то, что при указанной обработке пропадают мелкие детали изображения.

В качестве такого оператора со сглаживанием можно использовать *оператор Превита*, представляющий собой двумерную апертуру размером 3×3 со следующим набором коэффициентов $c_1 = 2/3$, $c_2 = -1/3$, $c_3 = 2/3$, $c_4 = -1/3$, $c_5 = -4/3$, $c_6 = -1/3$, $c_7 = 2/3$, $c_8 = -1/3$, $c_9 = 2/3$. Расчет яркости точки также должен быть произведен с использованием модуля от полученной величины.

Результаты сравнения работы различных алгоритмов для выделения контуров изображения представлены на рис. 47. На рис. 47, *a* – приведено исходное изображение, на рис. 47, *б* – результат обработки изображения градиентным методом, на рис. 47, *в-д* – результат обработки изображения операторами Робертса, Собеля, Лапласа и Превита.

Покажем, что иногда вместо выделения контуров для визуализации малоcontrastного объекта на экране можно использовать и другую последовательность операций, например для представленного изображения: линейное контрастирование, бинаризацию. На

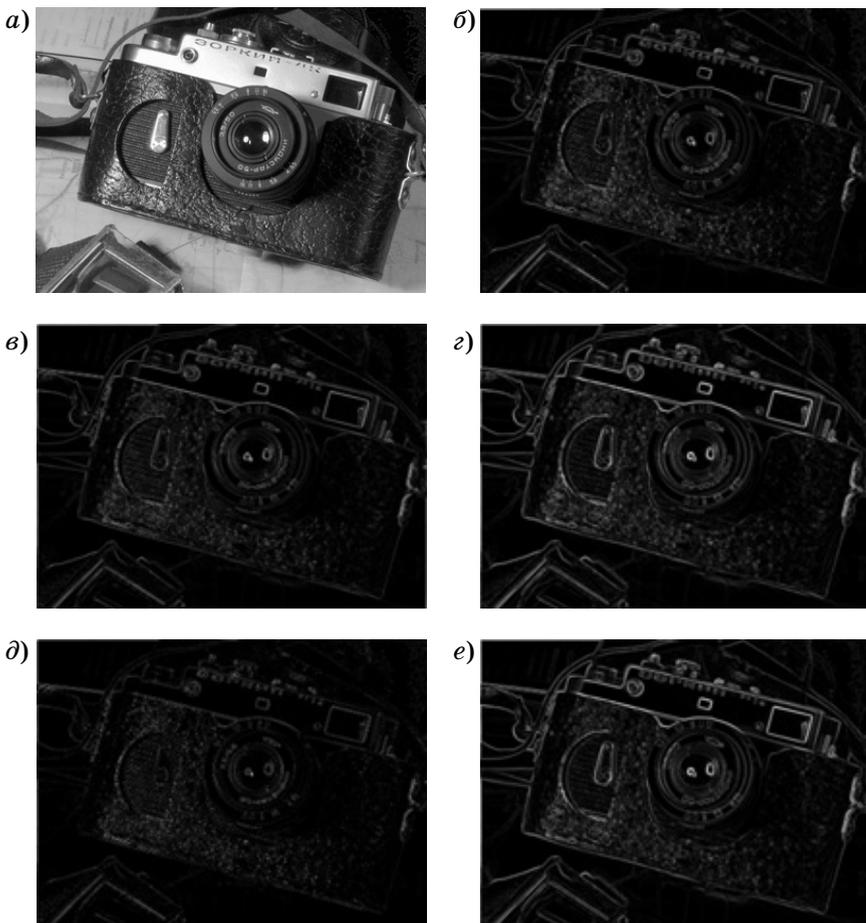


Рис. 47. Результат применение различных алгоритмов для выделения краев изображения

рис. 48, *а* приведено зашумленное изображение «Бурундук» после линейного контрастирования, на рис. 48, *б* – изображение «Бурундук» после бинаризации.

Как следует из рисунка 48, *б*, объект для оператора можно выделить и посредством простой бинарной обработки. Однако при этом следует отметить, что как при представленных ранее методах обработки изображения для поиска контуров, так и в процедуре обработки изображения посредством бинаризации результат преобразования зависит от величины порога, которая используется для

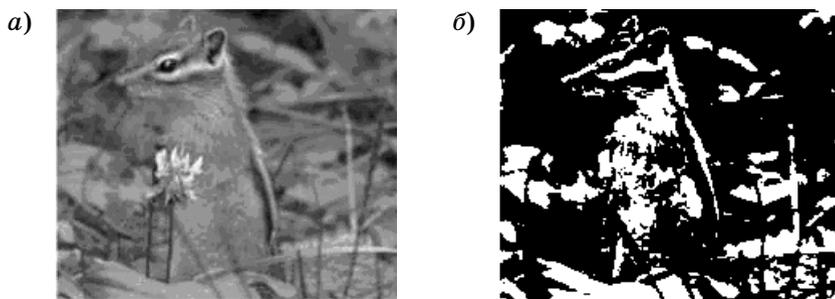


Рис. 48. Выделение объекта на малококонтрастном изображении посредством бинаризации

выделения объекта. Причем в последнем случае величина выбранного порога может оказать существенное влияние на визуальное восприятие изображения.

3.5. Преобразования гистограмм, реализация процедуры эквализации для повышения средней яркости на изображении

Вероятностные модели изображений широко используются для описания изображений. Введем для начала те понятия и определения, которые необходимы для понимания и расчета параметров, описывающих случайные процессы.

Под *случайным процессом (сигналом)* будем понимать такие изменения параметров физической системы во времени и пространстве, которые заранее в точности предсказать невозможно. Если взять изображения различных участков земной поверхности, сделанные в процессе аэрофотосъемки, то они при всей своей общей схожести – различны (рис. 49, а и 49, б) [21].

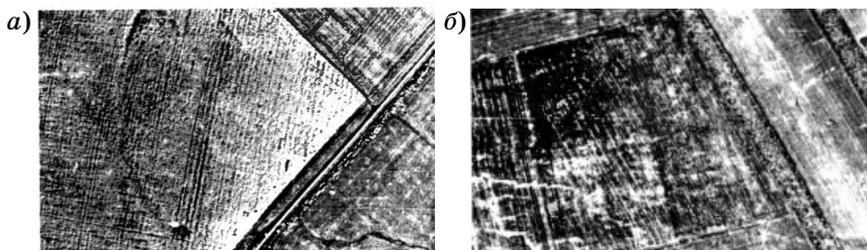


Рис. 49. Изображения различных участков земной поверхности

Учтем, что таких снимков различных участков земной поверхности может быть получено довольно много. Можно считать, что каждый снимок отдельного участка представляет собой случайную функцию пространственных координат и времени. Таким образом, все множество таких снимков представляет собой *генеральную совокупность* случайных реализаций, которую можно обрабатывать в соответствии с известными из теории математической статистики методами. *Выборочной совокупностью*, или просто *выборкой*, для рассматриваемого нами случая называют совокупность случайно отобранных снимков.

После того, как снимки отобраны, выполняется их обработка, при этом необходимо обеспечить наглядность представления данных, позволяющую получить какие-то первоначальные представления об их закономерности. Эта наглядность достигается путем построения таблиц или графических зависимостей.

Рассмотрим в качестве примера процедуру представления результатов при исследовании одной *случайной реализации*, т. е. одного полученного изображения. Если расположить все значения яркости пикселей изображения в порядке возрастания, то такая процедура будет называться *ранжированием*.

Ранжированный ряд значений яркости называется *вариационным рядом*. Количество значений яркости в данном ряду, равное числу пикселей ($V = MN$, где N – общее число строк изображения; M – общее число столбцов изображения) называется *объемом выборки*. Величина (ν_x), показывающая сколько раз встречается то или иное значение яркости, называется *частотой*. *Относительной частотой* или *статистической вероятностью* появления того или иного значения яркости будем называть величину

$$\omega_x = \nu_x/V.$$

Поскольку на изображениях, получаемых с современных ФПЗС-матриц, встречается в зависимости от разрядности АЦП от 2^8 до 2^{12} значений яркости (от 256 до 4096), то для такого большого числа значений, следует рассматривать не дискретные значения яркости, а ее значения в пределах определенного интервала.

В качестве частоты тогда принимается число случаев, в которых признак принял значения, входящие в некоторый интервал. Такую величину называют *интервальной частотой* и обозначают $\nu_{\Delta x}$ (соответственно рассматривается также и *интервальная относительная частота* – $\omega_{\Delta x}$). Полученный таким образом ряд называют *интервальным вариационным рядом*.

Для расчетов тогда необходимо определиться с величиной интервала Δx . Для вычисления Δx можно использовать формулу Стэрджесса [22]

$$\Delta x = (x_{\max} - x_{\min}) / (1 + 3.32211 \lg V),$$

где величина $(x_{\max} - x_{\min})$, называется вариационным размахом и является мерой разброса данных. Когда Δx выбрана, строят шкалу интервалов. За максимальную границу первого интервала принимают обычно величину

$$x_0 = x_{\min} + \Delta x / 2.$$

Следующая граница каждого последующего интервала определяется добавлением к верхней границе предыдущего значения интервала Δx

$$x_k = x_{k-1} + \Delta x, k = 1, 2, \dots$$

до тех пор, пока начало очередного интервала не окажется больше x_{\max} . Затем все значения яркости распределяются между соответствующими интервалами, и строится интервальный вариационный ряд.

Для обработки случайных данных удобно использовать их графическое представление: полигон или гистограмма, позволяющие выявить преобладающие значения яркости и характер распределения частот и относительных частот.

Полигон служит обычно для представления дискретного вариационного ряда. В системе координат (x, ν_x) или (x, ω_x) строятся точки, соответствующие значениям частот или относительных частот ряда, а затем эти точки соединяются прямыми линиями. На рис. 50 показана часть полигона относительных частот яркости, в диапазоне значений яркости от нуля до восьми для изображения, показанного на рис. 49, а).

Гистограмма — это диаграмма для наглядного отображения интервального вариационного ряда. Частота или относительная частота на указанной диаграмме отображаются не точкой, а прямой, параллельной оси абсцисс (Ox) на всем интервале. Это объясняется тем, что частота относится не к дискретному значению яркости x , а к интервалу яркостей Δx . Пример гистограммы распределения относительных частот яркости изображения, представленного на рис. 49, а) можно увидеть на рис. 51.

К числовым характеристикам, описывающим выборку, традиционно относят следующие величины: среднее арифметическое, медиана, мода, выборочная дисперсия, стандартное отклонение.

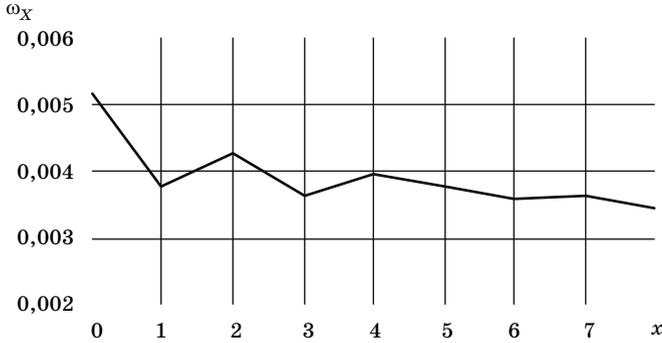


Рис. 50. Визуализация части полигона относительных частот для минимальных значений яркости изображения на рис. 49, а)

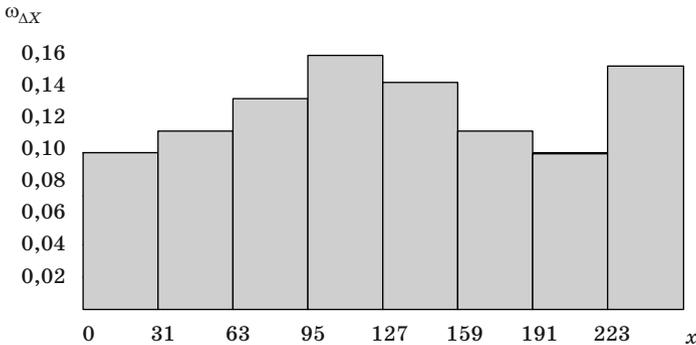


Рис. 51. Гистограмма относительных частот для всех значений яркости изображения на рис. 49, а)

Среднее арифметическое определяется по соотношению

$$\bar{x} = \frac{1}{V} \sum_{k=0}^{V-1} (x_k \cup_{xk}) = \sum_{k=0}^{V-1} (x_k \omega_{xk}).$$

Медиана в рассматриваемом случае представляет собой такое значение яркости, пиксели со значениями яркости меньше которой составляют практически половину от всего объема выборки.

Мода – это значение яркости, которая наиболее часто встречается на изображении, или для интервальной оценки это диапазон

значений яркостей, которые имеют максимальную относительную частоту. Из рис. 51 следует, что с учетом выбранного интервала ($\Delta x = 32$), данный диапазон составляет от 95 до 127.

Следующая часто используемая на практике величина – *выборочная дисперсия*, характеризующая меру разброса значений яркости от среднего арифметического

$$D_x = \frac{1}{V} \sum_{k=0}^{V-1} (x_k - \bar{x})^2.$$

Поскольку среднее арифметическое и выборочная дисперсия являются оценками случайного процесса, обладающими разной размерностью, то на практике чаще используют *выборочное стандартное отклонение*

$$\sigma_x = \sqrt{D_x}.$$

Вероятность того, что случайное значение яркости x_k (принимает значение, меньшее x , называется *функцией распределения* или *интегральным законом распределения* случайной величины x и обозначается $F(x)$:

$$F(x) = P\{x_k < x\}.$$

Функция распределения является универсальным видом *закона распределения вероятности*, пригодным для любой случайной величины. Она является монотонно возрастающей, диапазон изменения ее значений от нуля до единицы.

Чаще используют для анализа случайных процессов плотность распределения, определяемую соотношением $p(x) = F'(x)$. Анализируя любое изображение с ФПЗС-матрицы получаем множество результатов измерений. Найдем по полученным данным вероятность того, что величина x окажется в определенном интервале $[x_{k-1}, x_k]$. Эта вероятность будет прямо пропорциональна ширине интервала $\Delta x = x_k - x_{k-1}$. Коэффициент пропорциональности может зависеть от x . Таким образом, со случайной величиной x связана некоторая функция $f(x)$, которая по определению и является функцией распределения плотности вероятности $p(x)$, поскольку величина $f(x) dx$ равна вероятности события, состоящего в том, что случайная величина заключена в интервале от x до $x+dx$ (здесь dx – интервал *бесконечно малой ширины*).

Учитывая, что высота каждого столбика гистограммы представляет собой статистическую вероятность, получаем, что она пред-

ставляет собой визуализацию приближенного представления закона распределения вероятностей для случайной величины в виде ступенчатой функции, аппроксимация (выравнивание) которой до некоторой кривой $r(x)$ даст в первом приближении плотность распределения $p(x)$.

В качестве числового параметра, при помощи которого нередко характеризуют плотность распределения вероятности, может выступать *математическое ожидание*. Так, если имеется множество значений случайной яркости — x_k , а вероятность ее формирования на изображении составляет величину P_k , то математическое ожидание находится из соотношения

$$m_x = \sum_{k=1}^{V-1} x_k P_k.$$

В случае равномерного закона распределения, когда вероятность появления значения яркости составляет $1/V$, значение математического ожидания и арифметического среднего совпадают.

Покажем теперь, что всех поэлементных преобразованиях (кроме $y_{i,j} = x_{i,j}$) происходит изменение закона распределения вероятностей, описывающего изображение. Рассмотрим механизм этого изменения на примере произвольного преобразования яркости, описываемого монотонной нелинейной функцией $y = f(x)$ (рис. 52).

У представленной на графике (рис. 52) функции монотонной имеется обратная, которую можно обозначить как $x = q(y)$. Пусть случайная величина x обладает плотностью вероятности $p(x)$. Если задать произвольный малый интервал dx , то ему будет соответствовать интервал Δy преобразованной случайной величины. Попада-

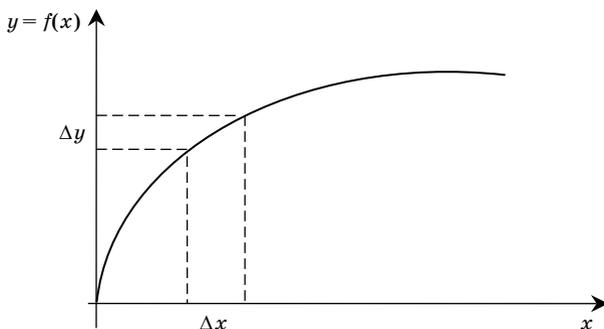


Рис. 52. Нелинейная функция преобразования яркости

ние величины x в интервал dx влечет за собой попадание величины y в интервал dy . Таким образом, должно выполняться следующее вероятностное равенство:

$$p(x)dx = p^*(y)dy,$$

откуда следует, что

$$p^*(y) = p(x)(dx/dy) = p(x)[d(q(y))/dy].$$

Полученное выражение позволяет вычислить плотность вероятности результата преобразования, откуда видно, что плотность вероятности преобразованной величины $p^*(y)$ отлична от исходной $p(x)$. При этом на полученный результат сильно влияет выбор вида функции преобразования $f(x)$, поскольку при расчете новой плотности вероятности участвует производная от обратной ей функции.

При линейном преобразовании изображения, например линейном контрастировании, сохранится вид плотности вероятности, но изменятся параметры, ее характеризующие. Пусть гистограмма исходного изображения может быть описана нормальной плотностью распределения вероятности вида

$$p(x) = \frac{1}{\sigma_x \sqrt{2\pi}} e^{-\frac{(x-m_x)^2}{2\sigma_x^2}},$$

тогда при линейном контрастировании изображения случайная величина y , будет иметь тот же закон распределения, но с другим средним арифметическим и дисперсией. При преобразовании данного вида $y_k = hx_k + b$ математическое ожидание величины y равно $m_y = hm_x + b$, а дисперсия — $D_y = h^2 D_x$.

Определение вероятностных характеристик изображений, прошедших нелинейную обработку, является прямой задачей анализа. Для решения практических задач обработки изображений актуально найти решение обратной задачи. Это означает, что при заданном или определенном виде плотности вероятности $p(x)$ требуется сформировать изображение с требуемым видом плотности вероятности $p^*(y)$. Для осуществления данной операции необходимо отыскать функцию преобразования, определить требуемое преобразование $y = f(x)$, которому следует подвергнуть исходное изображение. В практике цифровой обработки изображений часто к полезному результату приводит преобразование изображения для получения равномерной плотности распределения вероятности [18], которая имеет следующий вид:

$$p^*(y) = \begin{cases} \frac{1}{y_{\max} - y_{\min}}; & y \in [y_{\min}; y_{\max}] \\ 0; & y \in [-\infty; y_{\min}) \cup (y_{\max}; \infty] \end{cases}$$

Учтем, что при преобразовании изображений, как было показано ранее, выполняется условие: $p(x)dx = p^*(y)dy$ или $F(x) = F(y)$. Найдем, учитывая знание требуемой плотности вероятности $p^*(y)$, функцию распределения

$$F(y) = \int_{-\infty}^y p^*(y)dy = \int_{y_{\min}}^y \frac{1}{y_{\max} - y_{\min}} dy = \frac{y - y_{\min}}{y_{\max} - y_{\min}}$$

Учитывая, что $F(x) = F(y)$, то значение яркости для каждого пикселя нового изображения определяется по соотношению

$$y_{i,j} = (y_{\max} - y_{\min})F(x_{i,j}) + y_{\min}. \quad (3.3)$$

Таким образом, исходное изображение проходит нелинейное преобразование, которое определяется через функцию распределения исходного изображения.

Преобразование плотности вероятности требует определения функции распределения для исходного изображения. Поскольку априорные сведения о ней отсутствуют, то для получения нового изображения с требуемым видом плотности вероятности необходимо выполнить последовательно две операции. Вначале так, как это было описано ранее, формируется гистограмма исходного изображения, представляющая собой оценку яркостной плотности вероятности $p_{\text{оц}}(x_k)$. Тогда оценка функции распределения определяется по соотношению

$$F_{\text{оц}}(x_k) = \sum_{i=0}^k p_{\text{оц}}(x_k).$$

После чего выполняется соотношение (3.3), в котором используется оценка функции распределения, полученная через определение гистограммы. Поэтому методы преобразования яркостной плотности вероятности зачастую называют **гистограммными методами**. На последнем этапе, чтобы получить более высокий контраст визуализируемого изображения и привести изображения к заданному динамическому диапазону, используют операцию линейного контрастирования. Применение линейного контрастиро-

вания, как было показано ранее, не меняет вида плотности вероятности, а приводит только к ее новым параметрам.

Результат *эквализации* изображения, т. е. результат приведения исходного изображения к равномерной плотности распределения вероятности представлен на рис. 53.

Изображения, которые получаются в результате работы ЛСВ, могут иметь более значительный удельный вес темных участков и меньшее число участков с высокой яркостью в силу того, что инфракрасный подсвет формируется с узкой диаграммой направленности. Таким образом, в гистограмме исходного изображения математическое ожидание смещено в область значений с меньшей яркостью. Эквализация гистограммы приводит к выравниванию интегральных площадей равномерно распределенных диапазонов яркостей. Сравнение исходного (рис. 53, *а*) и обработанного (рис. 53, *б*) изображений, заимствованных из [10], показывает улучшение визуального восприятия.

а)



б)



Рис. 53. Пример эквализации изображения

Использование приведенных соотношений, строго говоря, справедливых для изображений с непрерывным распределением яркости, является не вполне корректным для цифровых изображений. Поэтому часто в результате обработки не удается получить идеальное распределение вероятностей выходного изображения и для улучшения качества изображения лучше проводить контроль гистограммы.

Под *контролем гистограммы* следует понимать отсутствие резких одиночных выбросов, в частности нулевых значений в полученной гистограмме. Простейшей процедурой контроля гистограммы может служить ее медианная фильтрация, которая была подробно рассмотрена ранее.

Более сложной операцией, которую можно реализовать для контроля гистограммы, является получение непрерывной функции $p^*(x)$ из рассчитанных экспериментальных отдельных гистограммных значений $p_{\text{оц}}(x_k)$, которая будет использована в дальнейшем для получения соответствующей непрерывной функции распределения.

Поскольку с ФПЗС-матрицы считывается от 256 до 4096 значений яркости, то зачастую в условиях реальной работы при недостаточной освещенности и при наличии шумов на матрице, в силу большого числа шумовых колебаний, обладающих приблизительно одинаковым значением яркости, может быть испорчена исходная интервальная оценка относительной вероятности.

В этом случае способы локальной интерполяции гистограммы, например, с помощью сплайнов, не имеют смысла из-за локальных выбросов интерполирующей функции, в этом случае гистограмму необходимо сглаживать в среднем, чаще всего многочленом, коэффициенты которого находят с помощью минимизации отклонения сглаживающей функции от заданных точек (рис. 54). Сделать это можно, используя метод наименьших квадратов, который получил широкое распространение при регрессионном анализе данных.

Рассмотрим необходимые для его применения математические соотношения для поиска нелинейной функции, обладающей минимальным среднеквадратическим отклонением от заданных случайным образом точек, расположенных вокруг нее. Пусть имеется ряд из N значений яркостей (x_k) исходного изображения, соответствующих центрам интервалов, для которых произведен расчет гистограммных значений, а также значения интервальных относительных частот, обозначенных для краткости изложения, как

$$u_k = (\omega_{\Delta x})_k.$$

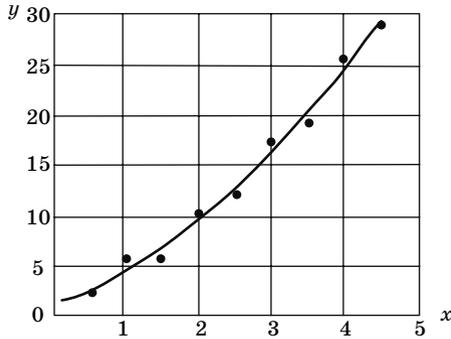


Рис. 54. К пояснению метода наименьших квадратов

Искомую функцию $p(x)$, которая служит для преобразования из значений яркости в интервальные относительные частоты, представим как

$$p(x) = a_0 f_0(x) + a_1 f_1(x) + a_2 f_2(x) + a_3 f_3(x) + \dots + a_m f_m(x),$$

где $f_0(x)$ – базисные функции, которые будут использованы для задания кривой; $a_0, a_1, a_2, \dots, a_m$ – коэффициенты, подлежащие определению. В частности, если в качестве базисных функций использовать степенные, то решение задачи будет сводиться к поиску полинома степени m ($m \ll N$). При этом подбор коэффициентов для искомой функции должен быть сделан таким образом, чтобы гистограммные значения относительных частот были бы близки к значениям частот, рассчитанным по этой функции, в смысле достижения минимального среднеквадратического отклонения. Для этого вначале задается следующая вспомогательная функция:

$$\Delta f = \sum_{k=0}^{N-1} (u_k - p(x_k))^2, \quad (3.4)$$

которая представляет собой разность квадратов между значениями функции и гистограммных значений, а затем производится поиск ее минимума.

Необходимым условием минимума функции многих переменных является равенство нулю ее частных производных первого порядка по независимым переменным. В (3.4) такими независимыми переменными являются коэффициенты полинома, которые до их определения являются не постоянными, а варьируемыми переменными. Данное условие представимо в виде следующей системы:

$$a = \frac{k_y k_{x2}^2 + k_x^2 k_{yx2} - k_{x2} (k_{yx2} + k_x k_{yx}) + k_{x3} k_{yx} - k_x k_y k_{x3}}{k_{x4} k_x^2 - 2k_x k_{x2} k_{x3} + k_{x2}^3 - k_{x4} k_{x2} + k_{x3}^3};$$

$$b = \frac{k_{yx} k_{x2}^2 + k_{x3} (k_{yx2} - k_y k_{x2}) - k_{x4} k_{yx} + k_x k_y k_{x4} - k_x k_{x2} k_{yx2}}{k_{x4} k_x^2 - 2k_x k_{x2} k_{x3} + k_{x2}^3 - k_{x4} k_{x2} + k_{x3}^3};$$

$$c = \frac{k_y k_{x3}^2 + k_{x2}^2 k_{yx2} + k_x (k_{x4} k_{yx} - k_{x3} k_{yx2}) - k_y k_{x2} k_{x4} - k_{x2} k_{x3} k_{yx}}{k_{x4} k_x^2 - 2k_x k_{x2} k_{x3} + k_{x2}^3 - k_{x4} k_{x2} + k_{x3}^3};$$

$$k_x = \frac{1}{N} \sum_{k=0}^{N-1} x_k; \quad k_{x2} = \frac{1}{N} \sum_{k=0}^{N-1} x_k^2; \quad k_{x3} = \frac{1}{N} \sum_{k=0}^{N-1} x_k^3; \quad k_{x4} = \frac{1}{N} \sum_{k=0}^{N-1} x_k^4;$$

$$k_y = \frac{1}{N} \sum_{k=0}^{N-1} y_k; \quad k_{yx} = \frac{1}{N} \sum_{k=0}^{N-1} (y_k x_k); \quad k_{yx2} = \frac{1}{N} \sum_{k=0}^{N-1} (y_k x_k^2).$$

На рис. 54 приведен пример того, как параболическая функция, построенная в соответствии с методом наименьших квадратов, приближается к экспериментально полученным значениям, отмеченным точками.

Как следует из рис. 54, полученная в соответствии с методом наименьших квадратов функция может сгладить отдельные выбросы гистограммных значений, что может способствовать улучшению визуализации после гистограммного преобразования изображения.

Существуют и другие применяемые в практике улучшения визуализации изображения примеры приведения исходной плотности вероятности к заданному виду. Так иногда осуществляется гиперболизация распределения

$$p^*(y) = \begin{cases} \frac{1}{\ln(y_{\max}) - \ln(y_{\min})}; & y \in [y_{\min}; y_{\max}] \\ 0; & y \in [-\infty; y_{\min}) \cup (y_{\max}; \infty] \end{cases}$$

Если учесть, что при прохождении света через глаз входная яркость логарифмируется его сетчаткой, то итоговая плотность вероятности оказывается равномерной с учетом физиологических свойств зрения.

3.6. Расширение динамического диапазона при регистрации изображения

Как было показано ранее, одним из самых важных этапов при проектировании лазерных систем видения является получение изображения с высоким контрастом объектов, расположенных на нем по отношению к фону. Первичный контраст изображения получается непосредственно на кадре ФПЗС-матрицы, обладающей определенным динамическим диапазоном. Понятие динамического диапазона обозначает максимально возможное количество различных градаций яркости в кадре. Наилучшее качество изображения достигается тогда, когда в поле зрения все необходимые для съемки объекты освещены примерно одинаково и обладают приблизительно равным коэффициентом отражения. В этом случае имеющийся динамический диапазон камеры используется наиболее эффективно, все объекты на изображении видны и имеют достаточную контрастность.

В области пространства разрешаемого ЛСВ одновременно существует область, освещенная по максимуму диаграммы направленности, а также окружающее пространство, в котором интенсивность лазерного излучения падает при приближении к границе диаграммы направленности (рис. 40, б). Следовательно, на будущем изображении объекта будут более светлая и более темные области. Кроме того, области объекта и фона с различными коэффициентами отражения также формируют различной интенсивности сигналы, регистрируемые на ФПЗС-матрице. Для выделения областей объекта с близкими коэффициентами отражения, а также при нахождении объекта в области быстрого изменения функции направленности излучения, градаций яркости, передаваемых ФПЗС-матрицей, может оказаться недостаточно, чтобы получить детализированное контрастное изображение объекта. Если экспозиция при съемке объекта выбрана так, чтобы получить изображение освещенных ярко участков объекта, то все остальные его части, находящиеся в условиях плохой освещенности или обладающие малым коэффициентом отражения, на изображении становятся черными и неразличимыми. В том случае, когда для получения изображения в области с плохой освещенностью используется более длительное время накопления полезного сигнала, оказывается засвеченной светлая часть и соответственно падает контраст деталей, расположенных в области с хорошей освещенностью. Приходится выбирать что-то одно, а другую часть кадра безвозвратно терять,

либо снижать общую контрастность изображения, делая неразличимыми многие детали.

Самый простой способ решить проблему с засветкой или с затемнением части кадра – это вырезать такие участки в кадре и запрашивать их от камеры отдельным потоком со своими настройками выдержки. Либо запрашивать целые кадры с разной выдержкой двумя отдельными потоками. Однако этот способ неэффективен, вместо одного изображения оператору приходится контролировать два изображения. Эффективным решением данной проблемы является технология *HDRi* (*High Dynamic Range imaging*), или «высокий динамический диапазон». Основной принцип – это получение двух кадров с разной выдержкой, один настроен на светлый участок сцены, а другой – на темный участок сцены, а потом слияние этих кадров в один, в котором присутствуют изображения и светлых, и темных участков.

Рассмотрим основы получения высококонтрастного изображения по технологии *HDRi*, описание которого представлено в [24]. В первом приближении процесс получения изображения можно представить в виде блок-схемы, представленной на рис. 55. На схеме приняты следующие обозначения: Ус – усилитель; МК – микроконтроллер.

Входом $S_{ВХ}$ является освещенность пикселя на сенсоре, выходом $S_{ВЫХ}$ является оцифрованное значение яркости пикселя, определяемое разрядностью применяемого АЦП. Используемый в схеме микроконтроллер управляет коэффициентом усиления сигнала с выхода ФПЗС-матрицы и выдержкой матрицы. Зависимость выхода $S_{ВЫХ}$ от входа $S_{ВХ}$ можно приближенно описать линейной функцией, прямой, ограниченной сверху и снизу (рис. 48, а) – сплошная линия). При слишком большом значении входного сигнала все элементы схемы уходят в режим насыщения («засветка»), поэтому

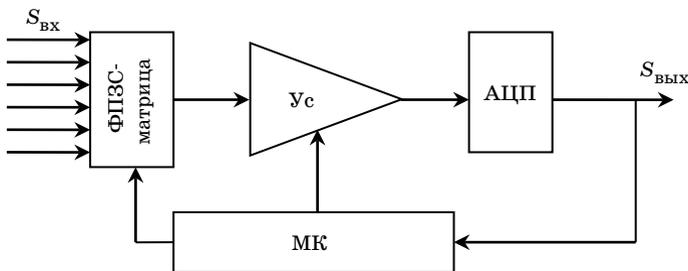


Рис. 55. Получение яркости пикселя ФПЗС-матрицы с различной экспозицией (выдержкой)

прямая ограничена сверху. При слишком малом значении сигнала все элементы схемы уходят в режим отсечки («затемнение»), поэтому прямая ограничена снизу.

Изменение времени выдержки приводит к различному накоплению на пикселе ФПЗС-матрицы и с учетом применения разного коэффициента усиления позволяет менять наклон прямой. При более короткой длительности накопления сигнала и меньшем коэффициенте усиления получаем другую зависимость $S_{\text{ВЫХ}}$ от входа $S_{\text{ВХ}}$ (штрихпунктирная линия на рис. 56, а). Будем называть для краткости их двумя кривыми, полученными при различной экспозиции.

Далее следует определить разницу между коэффициентами наклона двух прямых с разной экспозицией (рис. 56, б). Далее, для того чтобы получить изображение с высоким динамическим диапазоном, достаточно вместо засвеченных пикселей из первого изображения брать пиксели из второго изображения, яркость для которых будет помножена на значение величины R . В результате из двух изображений с разной выдержкой и малым динамическим диапазоном D_0 получим одно с увеличенным динамическим диапазоном – D_1 (рис. 56, в). При этом получим увеличенную разрядность пикселя.

Недостатком такого пути расширения динамического диапазона является то, что вначале требуется время для экспозиции кадра с первой выдержкой, а затем время для экспозиции кадра со второй выдержкой. Если в это время в зоне обзора лазерной системы видения имеется движущийся объект, то это приводит к существенному смазыванию его изображения. Поэтому технология *HDRi* дает хорошие результаты в случае наличия в зоне видимости объектов, которые обладают малой скоростью перемещения по кадру. Если производится автоматический поиск объектов на заданном фоне по

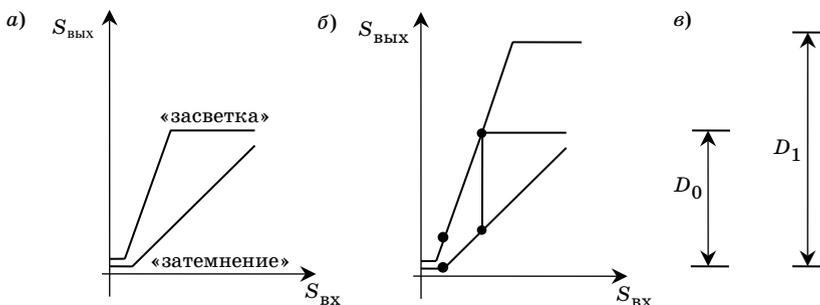


Рис. 56. Принцип расширения динамического диапазона изображения

контрасту, то дальнейшее преобразование изображения не требуется, но для непосредственной визуализации изображения с высоким динамическим диапазоном следует привести его к такому виду, который нормально воспринимается человеком-оператором. Строго говоря, человеческий глаз различает до 600 градаций яркости [25]. Это означает, что для адекватного представления яркостного изображения, содержащего все различаемые человеком градации яркости, требуется отображать изображения на экране, воспроизводящем 10-битное представление цвета, т. е. 1024 градации яркости. На сегодняшний день большинство индикаторов отображают только 256 градаций яркости, т. е. работают с 8-битным представлением яркости. Поскольку изображение с высоким динамическим диапазоном может содержать намного больше градаций яркости, чем 256, то его нельзя просмотреть на индикаторе потери информации. Для рассмотрения всех деталей изображения в областях с разным уровнем яркости требуется просматривать визуализацию кадра с последовательным изменением настройки яркости на экране инди-

а)



б)



Рис. 57. Исходные изображения с разной экспозицией



Рис. 58. Визуализация изображения, выполненного по технологии HDRi

катора. Поэтому для визуализации изображения, полученного по технологии *HDRi* на экране, отображающем только 256 градаций яркости, необходимо осуществлять сжатие полученного динамического диапазона. Данную операцию можно совершить, используя нелинейное преобразование яркости: гамма-коррекцию, логарифмическое преобразование.

Приведем пример реализации получения изображения с высоким динамическим диапазоном с учетом преобразования его для визуализации. На рис. 57, *a* и 57, *б* представлены два изображения с разной экспозицией. Изображение на рис. 58 – результат визуализации применения технологии *HDRi* с преобразованием диапазона яркости для визуализации в 8-битном представлении.

Как видно из представленного изображения можно получить высококонтрастные объекты, как в области с хорошей освещенностью, такие как небо, так и в области с малой освещенностью – берег со скамейкой. Это означает, что использование подобной технологии в лазерных системах видения может улучшить визуализацию изображений. Однако для того, чтобы улучшить результат, необходимо при получении двух кадров с разной экспозицией применять алгоритм компенсации неравномерной засветки ФПЗС-матрицы.

3.7. Увеличение резкости изображений.

Препарирование изображений

Для объектов, у которых коэффициент отражения для заданной длины волны подсвета ЛСВ близок к коэффициенту отражения фона, задача их обнаружения и идентификации может быть сильно осложнена в силу того, что в итоге осуществить операцию

выделения контуров на изображениях становится сложнее. Области с близкими коэффициентами отражения визуализируются для указанного случая в виде участков изображения с похожей яркостью, что затрудняет работу алгоритмов выделения границ на изображении, которые базируются на использовании операторов Робертса, Собеля, Лапласа или Превита. В том случае, если при этом предварительно осуществляется медианная фильтрация для устранения наблюдаемых на изображении шумов ФПЗС-матрицы, то в результате ее применения изображение также становится более размытым. Повысить визуальную четкость изображения для рассматриваемого случая можно посредством использования методов их обработки «скользящей» апертурой с соответствующим образом подобранными коэффициентами.

К ухудшению наблюдаемости тонкой структуры изображения, т. е. воспроизведению четких мелких деталей, границ между объектом и фоном приводит и ослабление верхних пространственных частот, поскольку МПФ оптико-электронной части ЛСВ (как и любой другой оптической системы) представляет собой монотонно убывающую функцию пространственной частоты.

Таким образом, *увеличение резкости изображения* – операция, которая реализует повышение коэффициента усиления для верхних пространственных частот [21]. При этом не произойдет обратного преобразования к оригиналу, поскольку часть верхних пространственных частот объекта теряется уже при прохождении канала распространения до матричного фотоприемника. Кроме того, объекты, границы которых представлены завышенными или даже максимальными значениями яркости, позволяют оператору гораздо лучше идентифицировать объект [21].

Таким образом, становится очевидной и разница между понятиями контрастность и резкость изображения. Контрастность изображения является характеристикой, которая определяется диапазоном яркости изображения. Чем он больше, тем выше контрастность. Понятие резкость, относится к локальным характеристикам, описывающим скорость изменения функции яркости по некоторому направлению, т. е. фактически ее производную. Естественно, может быть определено и среднее значение резкости по всему изображению. Увеличение контрастности изображения путем расширения его диапазона яркости всегда приводит к увеличению резкости. Однако изменение резкости может происходить и без изменения контрастности, например, путем подчеркивания контуров.

Рассмотрим, каким образом в соответствии с [21] можно сделать простую обработку изображения для увеличения видимости границ объекта и фона, а также мелких деталей изображения. Пусть x_i – значения яркости в произвольной i -й строке исходного изображения, для которой поставлена задача увеличения ее резкости.

Кривая 1 (рис. 59) представляет собой строку изображения с расфокусированной границей объекта. Обработку данной строки для выделения более четкой границы между объектом и фоном производят в несколько этапов. Вначале осуществляется низкочастотная фильтрация, представляющая собой дополнительное сглаживание значений яркости в строке, в результате которой получается еще более размытое изображение (кривая 2, рис. 59). Затем формируется разностный сигнал (кривая 3, рис. 59), как результат вычитания из исходной яркости пикселей яркости пикселей сглаженного изображения. После чего яркости пикселей для разностного сигнала складываются соответственно с яркостями пикселей исходного изображения с неким поправочным коэффициентом, сформированным так, чтобы низкочастотные компоненты не изменились, а получившийся стал более высокочастотным (кривая 4, рис. 59). Таким образом, будут подчеркнуты локальные особенности – границы, мелкие детали.

Согласно [21] данную процедуру можно осуществить, используя апертуру 3×3 (рис. 50, б). Коэффициенты скользящего окна для выделения высокочастотных деталей при усреднении по 5 точкам следующие: $c_1 = 0$, $c_2 = -1$, $c_3 = 0$, $c_4 = -1$, $c_5 = 5$, $c_6 = -1$, $c_7 = 0$, $c_8 = -1$, $c_9 = 0$. При усреднении по 9 точкам коэффициенты для апертуры следующие $c_1 = c_2 = c_3 = c_4 = c_6 = c_7 = c_8 = c_9 = -1$, $c_5 = 9$. Поскольку результат обработки подобными фильтрами мо-

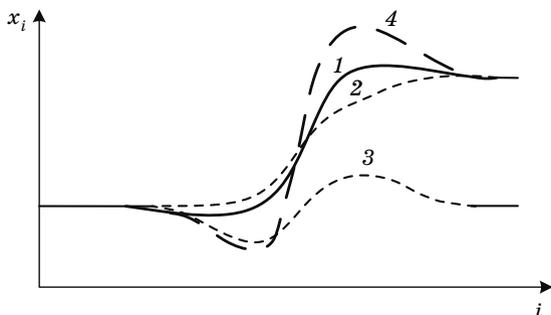


Рис. 59. К объяснению алгоритма увеличения четкости в i -строке изображения

жет дать как положительные, так и отрицательные значения, то для формирования итогового изображения следует применить вначале операцию линейного контрастирования или взять модуль от итоговых значений яркости пикселей.

Исходное изображение представлено на рис. 60, *а*, а на рис. 60, *б* – обработанное при помощи первой апертуры с последующим линейным контрастированием.

Как видно из представленного рисунка, повышение резкости изображения сопровождается увеличением шума на изображении. Так, при применении первой апертуры дисперсия шума согласно [21] увеличится в 29 раз. Кроме этого, на изображении появляются дополнительные шумы. Поэтому для выделения малоконтрастных объектов в данном примере предпочтительнее использовать процедуры выделения границ на изображении или его бинаризацию (см. рис. 48, *б*).

Основная проблема, которая возникает при проведении процедуры бинаризации, заключается в выборе порога. При определенной разнице между коэффициентами отражения между объектом и фоном за счет правильного выбора порога на итоговом изображении можно убрать фон. Особенно хорошо данная процедура работает в случаях обработки изображений, объект и фон на которых дают разный уровень освещенности на ФПЗС-матрице. В качестве простого примера контрастного изображения можно привести снимок машинописного текста.

В указанном случае плотность распределения вероятности и, следовательно, полученная экспериментальным образом идеальная гистограмма будут представлять собой два отдельно стоящих прямоугольника (рис. 61, *а*).

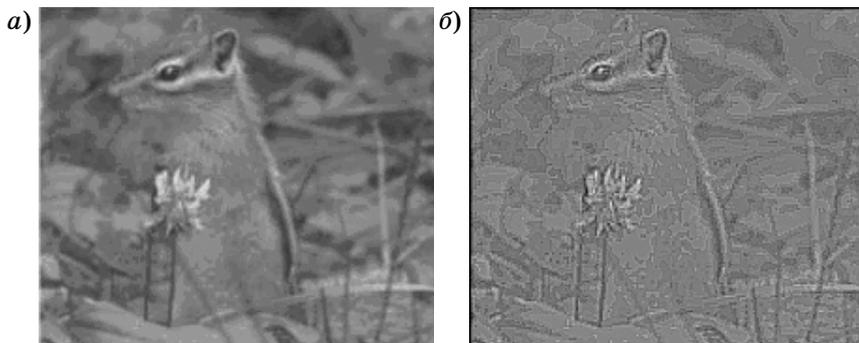


Рис. 60. К увеличению резкости изображения

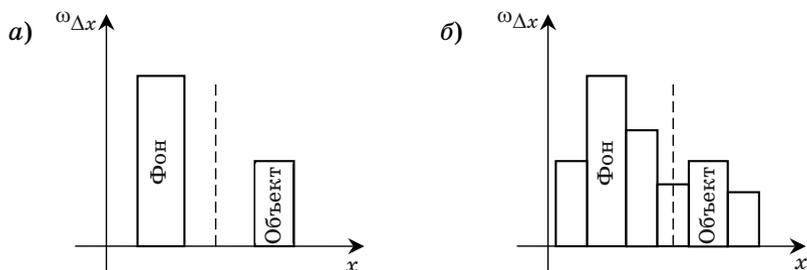


Рис. 61. Гистограммы, иллюстрирующие выбор порога при бинаризации

В указанном случае процедура установления порога очень проста – достаточно принять любое значение яркости в интервале между фоном и объектом в качестве порога (штриховая линия на рис. 61, а). В реальном случае в ЛСВ всегда имеется помеха обратного рассеяния, шумы ФПЗС-матрицы, неравномерность освещенности зоны обзора. Все это приводит к тому, что итоговая гистограмма изображения отличается от реальной наличием множества других яркостей в изображении (рис. 61, б). В такой ситуации в качестве порога можно принять любое значение яркости, которое лежит в интервале гистограммы, соответствующем минимальному значению относительной интервальной частоты, располагающейся между объектом и фоном. Лучше всего найти для гистограммы аппроксимирующую функцию, рассчитанную, например, по методу наименьших квадратов, и после найти соответствующее ее минимуму значение яркости, которое и будет являться оценкой порога.

Другим способом выделения изображения при заранее известной разнице между коэффициентами отражения от объекта и фона может служить *препарирование* изображения, представляющее собой различную обработку для пикселей изображения в зависимости от их яркости. Простейшей операцией препарирования является обработка изображения с двумя порогами. Данная пороговая обработка заключается в том, что обнуляются все значения яркости пикселей, лежащих ниже первого порога, также обнуляются все значения яркости выше второго порога.

Пример простейшей операции препарирования по выделению малококонтрастного изображения приведен на рис. 62 (а – исходное изображение; б – препарированное изображение).

Препарирование производится обычно в диалоговом окне, параметры и способы препарирования конкретной реализации изо-

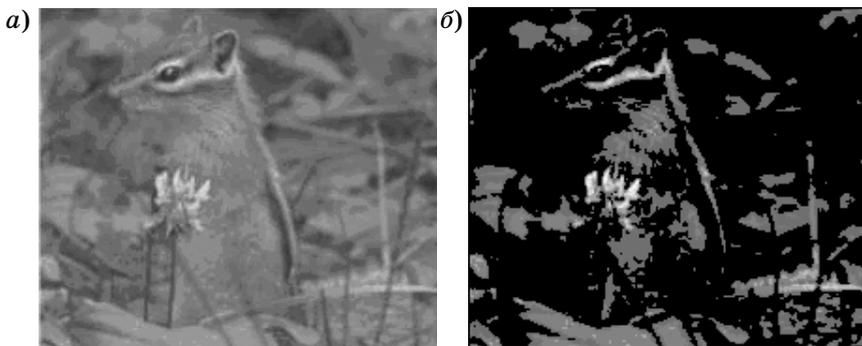


Рис. 62. Иллюстрация операции приготовления изображения

бражения при этом должны быть выбраны оператором, исходя из визуального улучшения выделения объекта на фоне. Однако данная операция требует значительного времени для анализа, что затрудняет решение задач по идентификации объектов оператором в реальном масштабе времени.

3.8. Фурье-образ изображения в частотной области

До этого момента мы рассматривали только пространственные методы обработки изображений, которые работают непосредственно с яркостными отсчетами изображения. Однако существует возможность также использовать способы вторичной обработки изображений, которая может быть реализована в области пространственных частот для улучшения качества визуализации или преобразования их к виду, удобному для дальнейшей обработки в пространственной области. Изображение описывается некоторой матрицей значений яркости дискретных отсчетов изображения в пространственной области, а его представление в области пространственных частот является матрицей, полученной путем разложения исходной матрицы изображения по выбранному базису. Как было показано ранее, дискретизация изображения должна быть реализована в соответствии с теоремой Котельникова, а полученный после дискретизации периодический спектр отфильтрован с использованием фильтра с прямоугольной амплитудно-частотной характеристикой. Результат такого преобразования будем далее для краткости называть спектром оцифрованного изображения. Рассмотрим для примера использование для обработки изображений разложение их в ряд отдельных гармонических составляющих, по-

лученных с использованием преобразования Фурье. Это будет дискретное представление спектра оцифрованного изображения, которое для краткости далее будем называть просто спектром.

Так, прямое двумерное дискретное преобразование Фурье (ДПФ) предназначено для получения из изображения, заданного в пространстве по координатам (i, j) , двумерного спектра в частотных координатах (m, n) [18]:

$$F(m, n) = \frac{1}{\sqrt{MN}} \sum_{j=0}^{N-1} \sum_{i=0}^{M-1} x_{i,j} e^{-2\pi i \left(m \frac{i}{M} + n \frac{j}{N} \right)},$$

где $x_{i,j}$ – значения отдельных элементов X – матрицы яркостных значений исходного изображения; M – число строк матрицы яркостных значений; N – число её столбцов; $i^* = \sqrt{-1}$.

Обратное дискретное преобразование Фурье (ОДПФ), которое служит для получения изображения по заданному спектру, имеет вид

$$x_{i,j} = \frac{1}{\sqrt{MN}} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} F(m, n) e^{2\pi i \left(i \frac{m}{M} + j \frac{n}{N} \right)}.$$

Дискретное преобразование Фурье является комплексным преобразованием. Результатом выполнения операции ДПФ являются значения амплитуд $Am(m, n)$ и фаз спектра изображения $\phi(m, n)$. Расчет амплитуды составляющей спектра производится по формуле

$$Am(m, n) = \sqrt{\text{Re}^2(F(m, n)) + \text{Im}^2(F(m, n))},$$

где Re – функция вещественной части; Im – функция мнимой части комплексного числа.

Если ввести следующее обозначение

$$\text{psi}^* = \arctg[\text{Im}(m, n)/\text{Re}(m, n)],$$

то фаза составляющей спектра определяется как

$$\phi(m, n) = \text{psi}^* \text{ при } \text{Im}(m, n) \geq 0 \text{ и } \text{Re}(m, n) > 0;$$

$$\phi(m, n) = \pi/2 - \text{psi}^* \text{ при } \text{Im}(m, n) \geq 0 \text{ и } \text{Re}(m, n) < 0;$$

$$\phi(m, n) = \pi + \text{psi}^* \text{ при } \text{Im}(m, n) \leq 0 \text{ и } \text{Re}(m, n) < 0;$$

$$\phi(m, n) = 2\pi + \text{psi}^* \text{ при } \text{Im}(m, n) \leq 0 \text{ и } \text{Re}(m, n) > 0;$$

$$\phi(m, n) = \pi/2 \text{ при } \text{Im}(m, n) > 0 \text{ и } \text{Re}(m, n) = 0;$$

$$\phi(m, n) = 3\pi/2 \text{ при } \text{Im}(m, n) < 0 \text{ и } \text{Re}(m, n) = 0.$$

Рассмотрим далее, каким образом можно реализовать визуализацию амплитудного спектра действительно изображения, которую часто принято называть в оптике *фурье-образом изображения*. Для этого сформируем изначально два изображения с различными законами изменения яркости. Для первого из них яркость изменяется по гармоническому закону с заданным периодом $T1$ только вдоль оси i

$$x_{i,j} = 0.5x_{\max}(\cos(2\pi i/T1)+ 1),$$

где x_{\max} – максимальная яркость в изображении. Заданное распределение яркости представлено на рис. 63, *а*. Второе изображение характеризуется как гармоническим изменением яркости с периодом $T1$ вдоль оси i , так и изменением яркости с периодом $T2 = T1$ вдоль оси j , которое описывается соотношением

$$x_{i,j} = 0.5x_{\max}(\cos(2\pi i/T1)\cos(2\pi j/T2)+1).$$

Если взять ДПФ от рассматриваемых изображений, а потом вычислить соответственно амплитудные спектры, то получится, что наибольшее значение амплитуды приходится на нулевую спектральную составляющую. Линейное контрастирование амплитудного спектра для получения его изображения, состоящего из 256 градаций яркости, приводит к тому, что ненулевые частотные составляющие, которые соответствуют резким изменениям яркости изображения: границам объектов, мелким деталям – становятся слабо различимыми. Для улучшения их визуальной наблюдаемости можно сделать предварительную обработку, заключающуюся в использовании нелинейного преобразования над значениями ам-

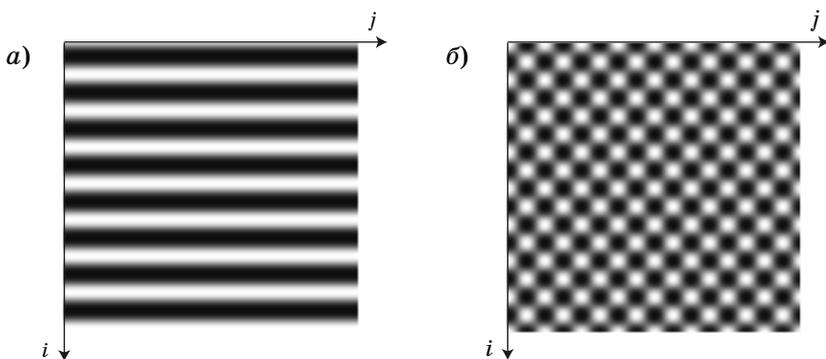


Рис. 63. Изображения с гармоническим изменением яркости по одной (а) и двум (б) координатным осям

плитуд такого рода, которое уменьшает значения больших амплитуд сильнее остальных. В качестве такого преобразования можно использовать логарифмическое. Другим способом визуализации ненулевых частотных составляющих является линейное контрастирование, выполняемое для всех частот кроме нулевой, с последующим присваиванием ей наибольшей градации яркости. Для представленных далее амплитудных спектров изображений использовался именно данный способ визуализации.

На рис. 64, *a* и 64, *б* представлены полученные таким образом визуализации для изображений 63, *a* и 63, *б*, соответственно. Рассмотрим вначале рис. 64, *a*. Отметим, что исходное изображение было квадратным с размерами $M = N = 32$. Периоды $T1 = 4$. Полученное изображение амплитудного спектра имеет следующие особенности: первая яркая точка с координатами $(0;0)$ соответствует постоянной составляющей, вторая точка $(8;0)$ соответствует частоте изменения яркости, ее координата по оси m может быть рассчитана по формуле

$$mf = M/T1.$$

Если mf не будет целым числом, то вместо одной точки в нулевом столбце будет спектр точек с ненулевой амплитудой. Вторая координата рассматриваемой точки по оси $n = 0$, поскольку вдоль указанной оси яркость постоянна. Третья точка, координаты которой равны $(24;0)$, является «зеркальной» к точке $(8;0)$ относительно центра оси m , т. е. ее координата по этой оси определяется как $M-mf$.

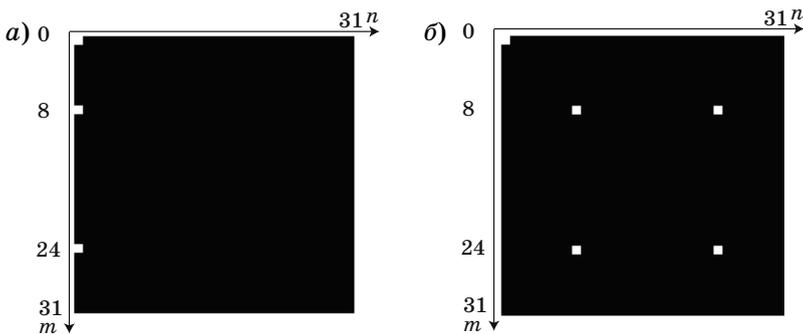


Рис. 64. Визуализация амплитудных спектров изображений, расположенных на рис. 63

Это является следствием симметричности амплитудного спектра аналогового изображения относительно нулевой частоты и того, что дискретизация изображения приводит к периодизации его спектра. Поскольку дискретизация изображения производится по каждой пространственной оси, то и периодизация спектра происходит по двум частотным осям. Иллюстрацией данного эффекта является рис. 64, б.

В когерентной оптике преобразование Фурье имеет реальную физическую интерпретацию. Оно описывает дифракцию Фраунгофера при прохождении когерентного пучка через оптическую систему с достаточно малой угловой апертурой. Действительно, любая дифракционная оптическая система с помощью когерентных волн изображения объекта, определяемого законами геометрической оптики, ставит ему в соответствие двумерный фурье-образ на плоскости, определяемый законами дифракции.

Пусть исходная форма сигнала в пространстве изображений – круг с равномерным распределением яркости. Физический аналог данного явления – дифракция Фраунгофера на круглом отверстии. Изображение, его фурье-образ и модуль ДПФ представлены на рис. 65 а, б, в, соответственно. Для комфортного зрительного восприятия фурье-образа и получения изображения, которое демонстрировало бы результат дифракции на круглом отверстии обычно применяют следующую процедуру. Вначале производят возведение модуля его ДПФ в степень, меньшую единицы. Потом начало координат смещают по обеим осям на половину интервала и формируют симметричное дополнение вокруг центра, применяя при этом процедуру линейного контрастирования изображения. Реальный же ДПФ после линейного контрастирования визуально предстанет в виде яркой белой точки в начале координат. Это видно, исходя из анализа модуля ДПФ, поскольку при большом радиусе круга яркости идет быстрое спадание до уровня боковых лепестков. Так, максимальный боковой лепесток меньше основного на 13,5 дБ, что означает, что его яркость после линейного контрастирования составит лишь 12 единиц от максимума в 255.

На рис. 66 представлены изображения эллипса, квадрата и прямоугольника, и их фурье-образы, заимствованные из [26].

Как следует из представленных рисунков, фурье-образы изображений соответствуют повернутым на 90 градусов исходным изображениям. Тем не менее, следует отметить, что визуализация фурье-образов простейших изображений в пространстве может что-то сказать об исходном изображении, а вот визуализация фурье-об-

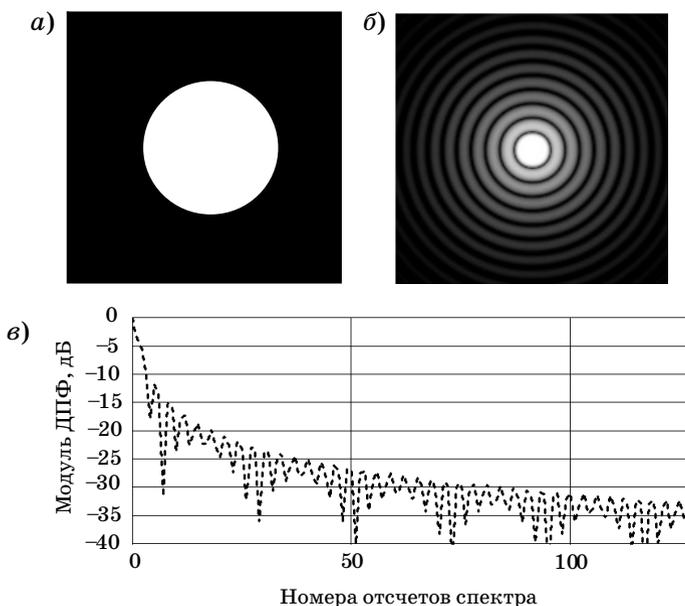


Рис. 65. Изображение, его фурье-образ и модуль ДПФ

разов реальных изображений только при определенных условиях и соответствующем выборе параметров визуализации.

Рассмотрим для примера, как отличаются визуализации логарифмов амплитудных спектров реальных изображений. На рис. 67, а и 67, б, соответственно представлены четкое и смазанное изображения архитектурного сооружения, украшенного мозаикой, т. е. имеющего множество мелких геометрических деталей.

На рис 68, а и 68, б представлены визуализации их амплитудных спектров. Сравнивая их между собой можно говорить только о том, что более высокие пространственные гармоники «размытого» изображения обладают гораздо меньшими значениями амплитуд. Это наглядно демонстрирует тот факт, что фильтрация верхних пространственных частот ведет к снижению четкости восприятия мелких деталей изображения. Общая яркость изображения также уменьшается при фильтрации гармоник в спектре.

Сравнительный анализ спектров изображений наблюдаемых объектов и структуры фона позволяет проводить операцию фильтрации в частотной области для удаления шумов, муара с потерей детализации изображения.

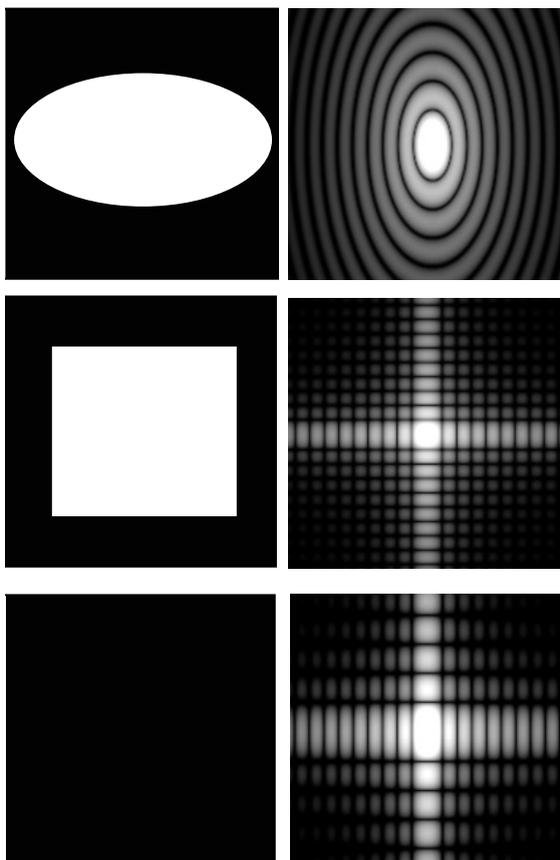


Рис. 66. Изображения простейших фигур и их фурье-образы

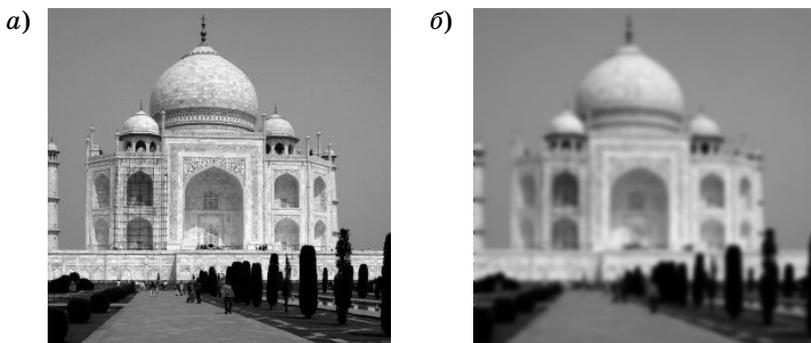


Рис. 67. Детализированное (а) и «размытое» (б) изображения

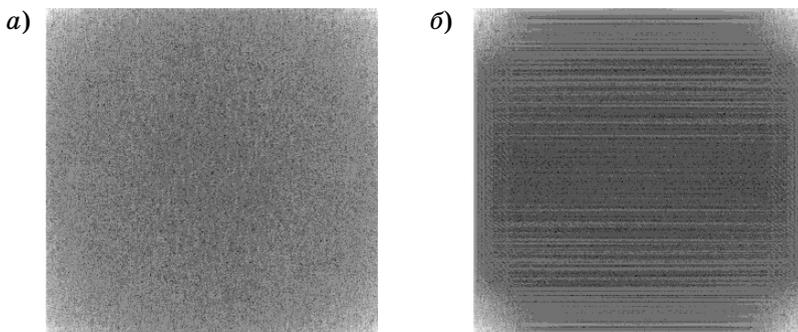


Рис. 68. Амплитудные спектры детализированного (а) и «размытого» (б) изображения

3.9. Перспективы применения автоматического распознавания образов в ЛСВ

На сегодняшний день развиваются многочисленные алгоритмы и прикладные программы, которые служат для автоматического обнаружения и распознавания объектов, например, человеческих лиц на изображениях. Тем не менее, решение данных задач в реальном масштабе времени представляет собой сложную задачу, требующую применения больших вычислительных возможностей, а главное до сих пор согласно [27] уступает по эффективности принятия решений обученному оператору.

Данную ситуацию можно прояснить на некоторых примерах по возможности распознавания определенных изображений человеком в условиях априорной информации о том, какого рода объекты он должен там обнаружить. Как пример можно привести монохромное изображение «Пять лошадей на фоне гор» (рис. 69) с автоматическим распознаванием образов, с которым сложно справиться компьютеру в силу изменения ракурса, размеров и возможных контуров искомых объектов. В то же время, как показывают исследования в отсутствие априорной информации о располагающихся на данном изображении объектах, человеку также сложно их распознать [28].

Для эффективного решения задач обнаружения и распознавания необходимо разрабатывать ЛСВ так, чтобы обеспечить требуемую итоговую разрешающую способность и достаточный контраст наблюдаемого изображения на экране монитора, которые и определяют предельную дальность обнаружения и распознавания ЛСВ.



Рис. 69. Иллюстрация возможностей распознавания образов оператором

Как показывает практика, данные параметры серьезно зависят от того, каким образом представляется изображение на экране монитора и от степени подготовки оператора. Для формирования слабо искаженных изображений в ЛСВ необходимо избегать aberrаций в оптическом тракте, использовать матрицы, которые, обладая минимальными шумами, обеспечивают требуемую частоту пространственной дискретизации. Итоговая МПФ ЛСВ должна быть по максимуму прямоугольной и близкой к единице в пределах выделяемого спектра пространственных частот объекта. При этом для улучшения качества цифрового изображения рекомендуется использовать алгоритмы цифровой обработки, направленные на увеличение контраста и минимизацию шумов при условии сохранения имеющихся резких границ по яркости. Тем не менее, число задач по распознаванию образов, которые успешно решаются на сегодняшний день, постоянно растет, появляются новые методики и соответствующие алгоритмы.

Их использование в лазерных системах видения может позволить улучшить эффективность решения задач в полуавтоматическом режиме работы или обойтись без участия оператора в автоматическом режиме работы системы. Основы теории распознавания образов, а также основные технические проблемы, возникающие при решении указанного круга задач, рассмотрены в следующих разделах.

Сегодня при практическом применении ЛСВ для обнаружения и распознавания объектов в качестве решающего устройства нередко выступает человек-оператор. Его скорость принятия решений по распознаванию образов при соответствующей подготовке превосходит возможности автоматических средств, выполняющих данные

операции. Даже сложные программные пакеты, предназначенные для формирования изображений в формате растровой и векторной графики (*Adobe Illustrator, PhotoShop, CorelDraw, CorelPhotoPaint, 3D Studio Max*), не могут справиться с задачей преобразования графика, иллюстрирующего квадратичную зависимость на фоне прямоугольной сетки, из растрового формата в векторный. Это выражается в избыточном выделении элементов на полученном растровом изображении, в то время как человек производит данную операцию, безошибочно отделяя прямые линии от искомой кривой за короткий промежуток времени. Тем не менее, следует учитывать, что значительного прогресса достигла обработка, направленная на преобразование рукописных или растровых печатных текстов в набор отдельных символов. Так программный пакет *FineReader* позволяет производить автоматическую сегментацию страниц документа, т. е. выделение блоков текста, таблиц, картинок и графиков, подрисующих подписей с надежным распознаванием текста в соответствующем блоке.

Однако в других смежных областях, напрямую связанных с задачами распознавания текстов, например в автоматической идентификации номеров автомобилей, возникают специфические сложности, связанные с особенностями формирования сигналов на аппаратном уровне и необходимости идентификации формы номера и расположения на нем подблестей с символами, несущими информацию о типе и государственной принадлежности транспортного средства [29]. В качестве примера на рис. 70 представлены изображения различных номеров транспортных средств, подлежащих идентификации (*a* – номер автомобиля РФ; *б* – номер мотоцикла РФ; *в* – номер автомобиля Украины), которые демонстрируют разное расположение символики. Рассмотрим существующие трудно-

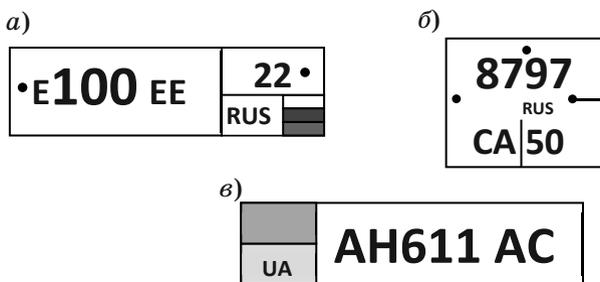


Рис. 70. Номера различных типов транспортных средств, подвергающихся распознаванию

сти по автоматическому распознаванию опознавательных знаков транспортных средств подробнее на примере накопленного опыта работы системы «ТелеВизард-Авто» [29]. Для этого вначале приведем основные задачи, которые реализуются при распознавании транспортного номера:

- приведение исходного изображения к виду, который не зависит от условий получения изображения: степень освещенности ФПЗС-матрицы, неравномерность распределения яркости источников подсвета, размытие вследствие движения объекта съемки при длительной экспозиции в условиях малой освещенности;

- выделение на исходном полученном изображении областей-кандидатов, которые могут содержать снимок пластины с номером транспортного средства;

- отбор изображения пластины по формальным геометрическим признакам: таким как соотношение сторон;

- приведение графического изображения пластины с номером к стандартному, удобному для дальнейшего анализа размеру с коррекцией качества;

- предварительное определение типа номерной пластины (вида и принадлежности номерной пластины) с привязкой к действующим стандартам;

- извлечение отдельных символов и их распознавание с проведением анализа символов по ключевым характеристикам, независимым от геометрических искажений и разрывов;

- уточнение результатов распознавания на основе информации о типе номера, а также в сравнении с результатами обработки предыдущих кадров.

Эффективность работы алгоритмов по распознаванию изображений зависит от условий съемки и параметров системы регистрации изображения. Для распознавания номерных знаков необходимо получить изображение:

- с высоким контрастом номерных знаков на пластине;

- с высоким пространственным разрешением изображения пластины;

- в условиях правильной освещенности для получения высокого контраста;

- при грамотно выбранной взаимной ориентации объекта съемки и системы регистрации изображения.

В реальных условиях эксплуатации оптико-электронных систем, предназначенных для регистрации изображений номерных табличек автомобилей, практически всегда возникают искажения,

которые способны приводить к ошибкам работы алгоритмов распознавания. Типичные проблемные изображения приведены на рис. 71 (*a* – низкое разрешение изображения номера; *б* – смазанное изображение; *в* – низкий контраст в плохую погоду, равномерно зашумленное изображение; *г* – передержка, засвеченное изображение с низким контрастом; *д* – изображение с неравномерной освещенностью; *е* – изображение деформированного номера).

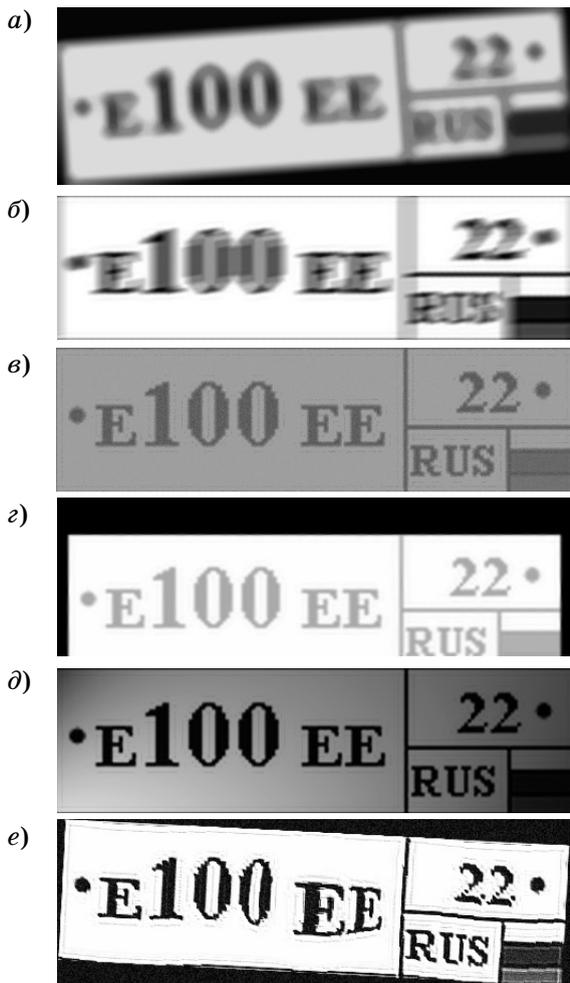


Рис. 71. Изображения номерных знаков с искажениями, ухудшающими эффективность распознавания символов

Допустим, что в системе автоматической идентификации и регистрации номерных знаков автомобилей задана вероятность правильного распознавания номера на статическом изображении, равная 95%. Найдем требуемую вероятность идентификации одного символа алгоритмами при наличии исходного изображения номера с ФПЗС-матрицы. Вначале запускается алгоритм, локализирующий на изображении область, содержащую номер транспортного средства. Требуемая вероятность правильного нахождения области P_{find} должна удовлетворять условию: $P_{\text{find}} \geq 98,5\%$. Далее работает алгоритм коррекции изображения, улучшающий контраст и неравномерность засветки. Требуемая вероятность получения необходимого итогового контраста на изображении $P_{\text{contr}} \geq 99,7\%$. Далее запускается алгоритм определения области, содержащей символы номера. Вероятность правильного определения местоположения всех символов $P_{\text{area}} \geq 99\%$. Тогда вероятность правильной идентификации символов номера алгоритмами распознавания должна составлять $P_{\Sigma} \geq 99,7\%$.

Учитывая то, что на основных российских государственных регистрационных знаках число символов составляет девять (код региона включает три символа), то вероятность распознавания отдельного символа: $P_1 = \sqrt[9]{P_{\Sigma}}$. Таким образом, значение вероятности одного символа также должно удовлетворять условию $P_1 \geq 99,975\%$. Это означает, что из 10000 изображений отдельных символов, поданных на вход процедуры распознавания, допускается возможность не распознавания или распознавания с ошибкой только 25 из них.

В то же самое время, практика применения различных российских видеосистем для автоматического распознавания номерных знаков транспортных средств показывает, что в реальных условиях достигаются следующие вероятности правильного распознавания: порядка 80% днем и 70% ночью [30]. Очевидно, что представленных значений недостаточно для желаемого уровня обеспечения безопасности дорожного движения. Проблема усугубляется тем, что, как было показано в предыдущих разделах, на процесс получения изображения с высоким контрастом и достаточным пространственным разрешением серьезное влияние оказывают условия распространения энергии источника подсвета: наличие дымки, тумана, дождя, паразитной солнечной засветки, помехи обратного рассеяния. Повысить эффективность распознавания номеров за счет улучшения условий съёмки можно, если использовать лазерные системы видения в инфракрасном диапазоне или в сине-зеленой части спектра, гарантирующей слабое затухание отраженного

сигнала при наличии большой влажности. Помимо обеспечения работы в сложных погодных условиях, подобная система, обладая высоким пространственным разрешением, позволит увеличить и вероятность правильного распознавания символов.

Аналогичный принцип работы лазерных систем видения предлагается к использованию в системах видения для обеспечения безопасности на объекте. Структурная схема такой лазерной системы мониторинга объектов, разработанной в США (*Laser-Based Item Monitoring System, LBIMS*), представлена на рис. 72 [31].

Ретроотражательные метки наносятся на защищаемые объекты. Термин «ретроотражательные» означает, что они выполнены таким образом, чтобы максимум их индикатрисы рассеяния совпадал с направлением прихода падающего излучения. Такая ретроотражательная метка может представлять собой набор небольших угловковых отражателей или полусферических металлических поверхностей, формирующих определенный символ, который можно считать и распознать системой.

Лазер и фотоприемник подключаются к контроллеру. Он задает программу обхода сканируемых объектов, реализуемую посредством отклонения системой зеркал, сравнивает данные о расположении меток с информацией, получаемой в ходе сканирования. Для обработки данных и распознавания геометрической конфигурации ретроотражательных меток к контроллеру может подключаться компьютер.

Представленные примеры показывают, что дальнейшее развитие лазерных систем видения сложно представить без применения в них функций автоматической идентификации объектов или полуавтоматического распознавания в помощь оператору.

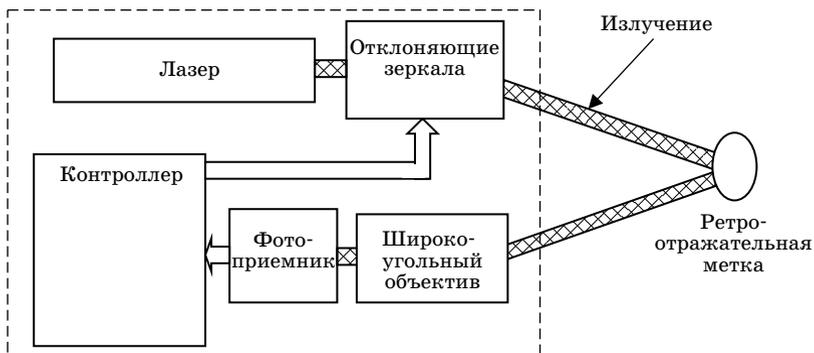


Рис. 72. Лазерная система мониторинга охраняемых объектов

4. МЕТОДЫ АВТОМАТИЧЕСКОГО РАСПОЗНАВАНИЯ ИЗОБРАЖЕНИЙ

4.1. Корреляционное сопоставление

Корреляционное сличение является классическим средством поиска эквивалентов эталона $w(x,y)$ размерами $J \times K$ на изображении $f(x,y)$ размерами $M \times N$. При этом предполагается, что $J \leq M$, $K \leq N$. В самой простой форме корреляция между изображениями $f(x,y)$ и $w(x,y)$ задается выражением [32]

$$c(x,y) = \sum_s \sum_t f(s,t)w(x+s,y+t)$$

для $x = 0, 1 \dots M-1$; $y = 0, 1 \dots N-1$; т. е. корреляция вычисляется в процессе скользящего поиска, когда каждый очередной фрагмент получается из предыдущего простым сдвигом на один шаг. В каждой точке (x_0, y_0) изображения $f(x,y)$ вычисляется одно значение $c(x_0, y_0)$.

Процедуру корреляции иллюстрирует рис. 73. Следует отметить, что для значений x и y вблизи краев изображения f точность теряется. Величина ошибки корреляции пропорциональна той доли площади эталона w , которая выходит за границы изображения f .

Недостатком корреляционной функции $c(x, y)$ является ее чувствительность к изменениям амплитуд f и w . Чтобы преодолеть это

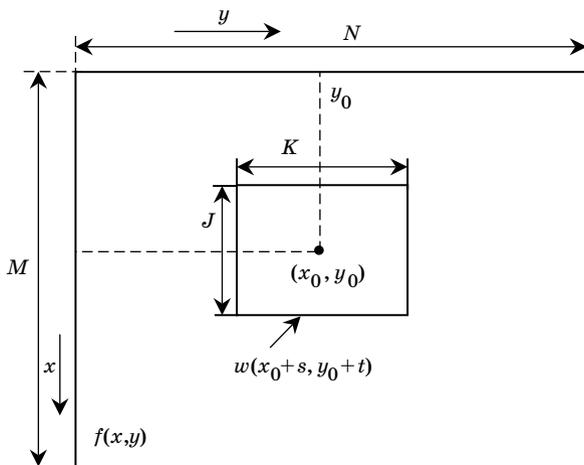


Рис. 73. Получение значения корреляции f и w в точке (x_0, y_0)

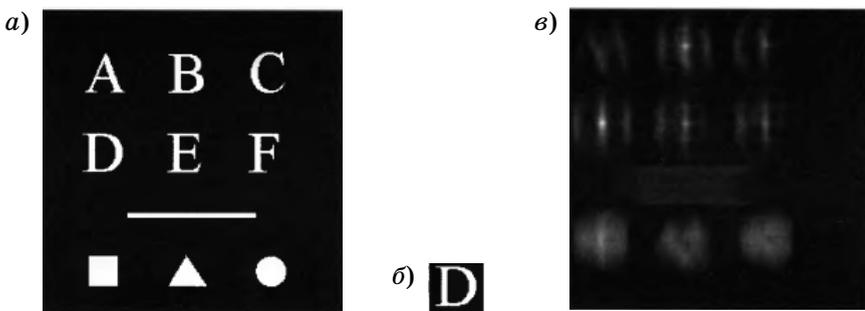


Рис. 74. Изображение (а), эталон (б), и коэффициент корреляции между ними (в)

затруднение, часто используется подход, при котором сопоставление осуществляется с помощью коэффициента корреляции, определяемого выражением

$$\gamma(x,y) = \frac{\sum_s \sum_t [f(s,t) - \bar{f}][w(x+s,y+t) - \bar{w}]}{\sqrt{\sum_s \sum_t [f(s,t) - \bar{f}]^2 \sum_s \sum_t [w(x+s,y+t) - \bar{w}]^2}},$$

где $x = 0, 1 \dots M-1$; $y = 0, 1 \dots N-1$, \bar{w} – среднее значение пикселей в эталоне w ; \bar{f} – среднее значение элементов области изображения f , совпадающей с текущим положением w . Коэффициент корреляции $\gamma(x,y)$ изменяется в диапазоне от минус 1 до 1 и не зависит от изменения амплитуд f и w .

Иллюстрирует описанный принцип рис. 74 [33]: на рис. 74, а приведено изображение $f(x,y)$, а на рис. 74, б – эталон $w(x,y)$. Коэффициент корреляции $\gamma(x,y)$ показан как изображение на рис. 74, в. Наибольшее значение коэффициента $\gamma(x,y)$ достигается в точке наилучшего совпадения f и w .

Серьезным недостатком корреляционной меры сходства является ее чувствительность к геометрическим искажениям видимых размеров сопряженных фрагментов при изменении ракурса съемки, в связи с чем корреляция редко используется в тех случаях, когда возможен произвольный поворот распознаваемого объекта.

4.2. Кросс-спектральная мера сходства

Сопоставление изображений может осуществляться не только в пространственной, но и в частотной области. Зачастую такой под-

ход оказывается более рациональным при сопоставимых размерах текущего изображения $f(x, y)$ и эталона $w(x, y)$.

Для единообразия запишем матрицу эталонного фрагмента w в левом верхнем углу нулевой матрицы порядка $L \times L$. Это позволит перейти от центрированных переменных $\bar{f}(x, y) = f(x, y) - \bar{f}$ и $\bar{w}(x, y) = w(x, y) - \bar{w}$ к спектральным компонентам (Фурье-образам) изображений \bar{F} , \bar{W} :

$$S_{p,q}(\bar{F}) = |S_{p,q}(\bar{F})| \exp\{j\varphi_{p,q}(\bar{F})\}; \quad p, q = 0, 1 \dots L-1;$$

$$|S_{p,q}(\bar{F})| = \sqrt{\operatorname{Re}^2 S_{p,q}(\bar{F}) + \operatorname{Im}^2 S_{p,q}(\bar{F})}, \quad \varphi_{p,q}(\bar{F}) = \arctg\left(\frac{\operatorname{Im} S_{p,q}(\bar{F})}{\operatorname{Re} S_{p,q}(\bar{F})}\right).$$

Поскольку при смещении эталона $\bar{w}(x, y)$ в области поиска меняется только его фазовый спектр

$$S_{p,q}(\bar{W}(x+k, y+l)) = S_{p,q}(\bar{W}) \exp\left\{j\frac{2\pi}{L}(pk+ql)\right\},$$

то будем минимизировать функционал [34]

$$G_{k,l} = \sum_{p,q} h_{p,q} \left(S_{p,q}(\bar{F}) - S_{p,q}(\bar{W}) \exp\left\{j\frac{2\pi}{L}(pk+ql)\right\} \right)^* \times \\ \times \left(S_{p,q}(\bar{F}) - S_{p,q}(\bar{W}) \exp\left\{j\frac{2\pi}{L}(pk+ql)\right\} \right)$$

по параметрам сдвига k , l , где звездочкой обозначена операция комплексного сопряжения, а функция h_{pq} осуществляет «взвешивание» разностной меры спектральных компонентов. Поскольку «энергия» изображений $|S_{p,q}(\bar{F})|^2$, $|S_{p,q}(\bar{W})|^2$ не зависит от параметров k и l , то положение минимума функционала $G_{k,l}$ соответствует точке экстремума перекрестного члена

$$C_{k,l}(\bar{F}, \bar{W}) = \sum_{p,q} h_{p,q} S_{p,q}^* S_{p,q}(\bar{W}) S_{p,q}(\bar{F}) \exp\left\{j\frac{2\pi}{L}(pk+ql)\right\}.$$

При $h_{p,q} \equiv 1$ данное выражение соответствует классическому корреляционному алгоритму, реализуемому в спектральной области.

4.3. Фурье-дескрипторы и инвариантные моменты Ху

В распознавании изображений существует проблема автоматического отождествления фрагментов независимо от их размеров, положения и взаимной ориентации. Для решения данной задачи пространство признаков, используемых при описании изображения, должно обладать свойствами инвариантности к таким видам аффинных преобразований, как масштабирование, сдвиг и поворот. Наиболее распространенными пространствами признаков среди удовлетворяющих данным условиям являются Фурье-дескрипторы и инвариантные моменты Ху, широко используемые системами распознавания, индексации и поиска изображений.

Фурье-дескрипторы являются пограничными признаками и базируются на учете элементов изображения, расположенных вдоль его контура. При расчете инвариантных моментов Ху, напротив, учитываются значения всех пикселей, входящих в состав фрагмента изображения. Как Фурье-дескрипторы, так и моменты Ху, не зависят от аффинных преобразований исходных фрагментов, включая масштабирование, сдвиг и поворот.

Фурье-дескрипторы. Дано изображение замкнутой формы в двухмерной декартовой системе координат. Опишем границу изображения s равномерно распределенными вдоль нее K точками и повернем против часовой стрелки, как показано на рис. 75. Координаты каждой точки $(x_0, y_0), (x_1, y_1) \dots (x_{K-1}, y_{K-1})$ выразим в формате $x(k) = x_k, y(k) = y_k$. В данных условиях контур может быть представлен в виде последовательности комплексных чисел:

$$s(k) = x_k + jy_k, k = 0, 1 \dots K - 1,$$

что эквивалентно отождествлению оси x изображения с осью вещественных, а оси y – с осью мнимых чисел на комплексной плоскости

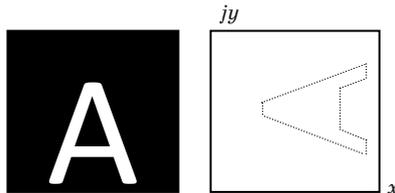


Рис. 75. Граница изображения, представленная на комплексной плоскости

сти. Коэффициенты дискретного преобразования Фурье (ДПФ) комплексной последовательности $s(k)$ определяются выражением [35]:

$$z(u) = \frac{1}{K} \sum_{k=0}^{K-1} s(k) \exp(-j2\pi uk / K), u = 0, 1 \dots K - 1.$$

Комплексные коэффициенты $z(u)$ называются Фурье-дескрипторами контура. Они не являются инвариантными к операциям масштабирования, сдвига и поворота изображения, однако изменения этих параметров связаны с простыми операциями над Фурье-дескрипторами. Чтобы избавиться от компонент, зависящих от аффинных преобразований, используют дескрипторы вида

$$c(u-2) = \left| \frac{z(u)}{z(1)} \right|, u = 2, 3 \dots K - 1.$$

Инвариантные моменты Ху. Множество инвариантов, основанных на нелинейной комбинации регулярных моментов, были впервые введены Ху в 1961 г. Регулярные моменты $(p+q)$ -го порядка непрерывной функции $f(x,y)$ задаются выражениями следующего вида:

$$m_{p,q} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x^p y^q f(x,y) dx dy \quad p, q = 0, 1, 2 \dots$$

Центральные моменты функции $f(x,y)$ определяются:

$$\mu_{p,q} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \bar{x})^p (y - \bar{y})^q f(x,y) dx dy \quad p, q = 0, 1, 2 \dots$$

где $\bar{x} = m_{1,0} / m_{0,0}$, $\bar{y} = m_{0,1} / m_{0,0}$ – координаты центра тяжести.

Центральные моменты рассчитываются относительно центра тяжести функции яркости $f(x,y)$ и являются инвариантными относительно сдвига изображения. В случае дискретного бинарного изображения размерностью $N \times M$ формула для вычисления центральных моментов имеет вид

$$\mu_{p,q} = \sum_{x=1}^N \sum_{y=1}^M x^p y^q A(x,y) \quad p, q = 0, 1, 2 \dots$$

Функция яркости $A(x,y)$ принимает в каждой точке изображения нулевое либо единичное значение.

Нормированные центральные моменты, инвариантные относительно масштабирования, выражаются:

$$\eta_{p,q} = \frac{\mu_{p,q}}{\mu_{0,0}^\gamma}, \quad \gamma = (p + q + 2) / 2, \quad p + q = 2, 3 \dots$$

Семь моментных инвариантов Ху на основе нормированных центральных моментов определяются следующими выражениями [36]:

$$\left\{ \begin{array}{l} \varphi_1 = \eta_{2,0} + \eta_{0,2}, \\ \varphi_2 = (\eta_{2,0} - \eta_{0,2})^2 + 4\eta_{1,1}^2, \\ \varphi_3 = (\eta_{3,0} - 3\eta_{1,2})^2 + (3\eta_{2,1} - \eta_{0,3})^2, \\ \varphi_4 = (\eta_{3,0} + \eta_{1,2})^2 + (\eta_{2,1} + \eta_{0,3})^2, \\ \varphi_5 = (\eta_{3,0} - 3\eta_{1,2})(\eta_{3,0} + \eta_{1,2}) \left[(\eta_{3,0} + \eta_{1,2})^2 - 3(\eta_{2,1} + \eta_{0,3})^2 \right] + \\ + (3\eta_{2,1} - \eta_{0,3})(\eta_{2,1} + \eta_{0,3}) \left[3(\eta_{3,0} + \eta_{1,2})^2 - (\eta_{2,1} + \eta_{0,3})^2 \right] \\ \varphi_6 = (\eta_{2,0} - \eta_{0,2}) \left[(\eta_{3,0} + \eta_{1,2})^2 - (\eta_{2,1} + \eta_{0,3})^2 \right] + \\ + 4\eta_{1,1}(\eta_{3,0} + \eta_{1,2})(\eta_{2,1} + \eta_{0,3}), \\ \varphi_7 = (3\eta_{2,1} - \eta_{0,3})(\eta_{3,0} + \eta_{1,2}) \left[(\eta_{3,0} + \eta_{1,2})^2 - 3(\eta_{2,1} + \eta_{0,3})^2 \right] + \\ + (3\eta_{2,1} - \eta_{0,3})(\eta_{2,1} + \eta_{0,3}) \left[3(\eta_{3,0} + \eta_{1,2})^2 - (\eta_{2,1} + \eta_{0,3})^2 \right]. \end{array} \right.$$

Данные моменты удовлетворяют требованиям инвариантности к сдвигу, повороту и масштабированию изображения. Существуют инварианты Ху более высокого порядка, однако их расчет требует больших вычислительных затрат, в результате чего на практике используют первые семь моментных инвариантов.

Пространственное разрешение. В соответствии с определением аффинного преобразования, каждому входному пикселю должен соответствовать пиксель на выходе. Общее количество элементов после преобразования не должно изменяться. Однако процесс визуализации цифровых изображений после аффинных преобразований связан с необходимостью согласования пространственного разрешения.

Размер цифрового изображения пропорционален его пространственному разрешению. Изображение, состоящее из 640 пикселей в длину и 480 пикселей в высоту, обладает разрешением 640×480

и содержит 307200 элементов. Чем больше пространственное разрешение изображения, тем большее количество пикселей оно содержит и большей четкостью обладает. Для уменьшения масштаба из исходного изображения удаляются строки и столбцы. Процесс масштабирования изображения, сопровождающийся изменением его пространственного разрешения иллюстрирует рис. 76.

Результаты приведения уменьшенных изображений с рис. 76 к исходному размеру 256×256 показаны на рис. 77. Как видно из рис. 77, восстановленные изображения искажены.

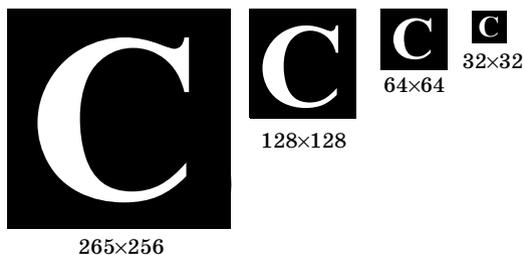


Рис. 76. Изображения с различным пространственным разрешением

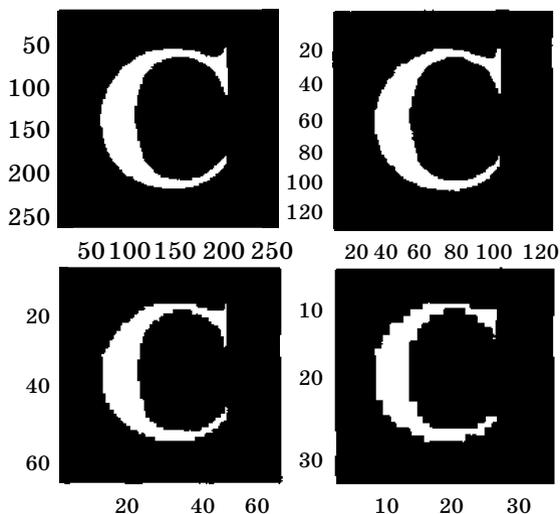


Рис. 77. Приведение уменьшенных изображений к исходному размеру 256×256

Экспериментальная оценка. В [35] проводилось экспериментальное исследование эффективности использования Фурье-дескрипторов и инвариантных моментов Ху для распознавания растровых изображений букв латинского алфавита в среде Matlab. Изображения подвергались аффинным преобразованиям и обладали пространственным разрешением от 512×512 до 16×16 . Структура использованной математической модели представлена на рис. 78. Такая модель позволила получить образы изображений, собрать их в библиотеке и произвести распознавание.

При исследовании Фурье-дескрипторов составлялся вектор признаков, содержащий первые десять дескрипторов. Именно они являются наиболее значимыми ввиду близости к нулю остальных элементов. В случае инвариантных моментов Ху вектор признаков содержал все семь инвариантов. Фурье-дескрипторы рассчитывались для изображений всех букв алфавита от «А» до «Z» с пространственным разрешением от 512×512 до 32×32 . Для инвариантов Ху диапазон разрешений составил от 512×512 до 16×16 .

Зависимости Фурье-дескрипторов изображений букв «А» и «Z» от пространственного разрешения показана на рис. 79. Из рис. 79 видно, что пределом для Фурье-дескрипторов является разрешение 64×64 , после которого наблюдаются их очевидные изменения. При разрешении 32×32 количество успешных распознаваний сократилось до 53,8%.

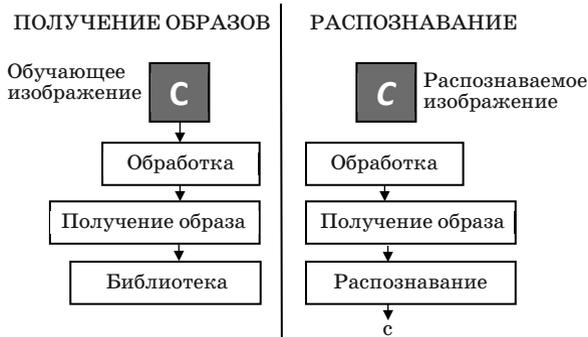


Рис. 78. Структура математической модели распознавания изображений

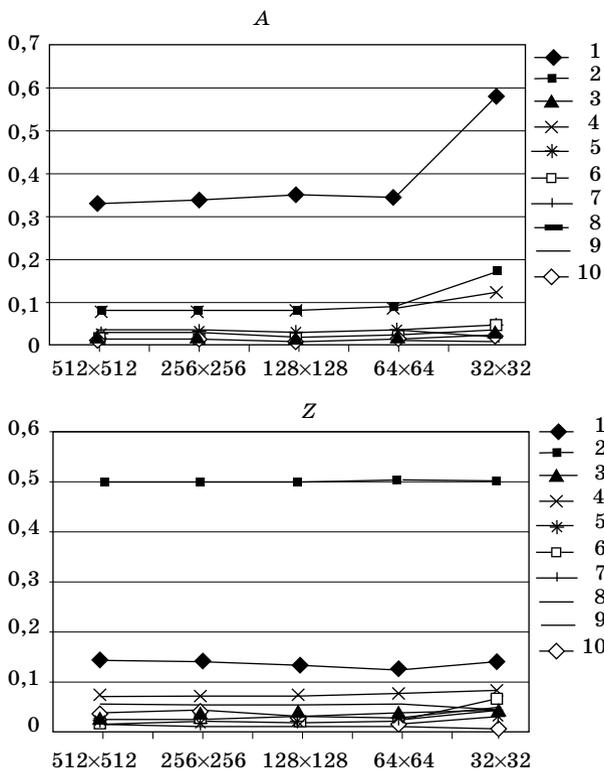


Рис. 79. Зависимость Фурье-дескрипторов изображений букв «А» и «Z» от пространственного разрешения

Зависимость от пространственного разрешения инвариантных моментов X_u изображения буквы «С» показана на рис. 80. В данном случае пороговым является разрешение 128×128 . При разрешении 32×32 количество успешных распознаваний упало со 100% до 84,6%.

Таким образом, Фурье-дескрипторы и моменты X_u обладают свойством инвариантности относительно таких видов преобразований изображения, как масштабирование, сдвиг и поворот. Однако для обоих видов описания изображений существует порог пространственного разрешения. Проведенные экспериментальные исследования показали, что при использовании вектора признаков из первых десяти Фурье-дескрипторов для обеспечения 100% успешного распознавания пространственное разрешение изображения не

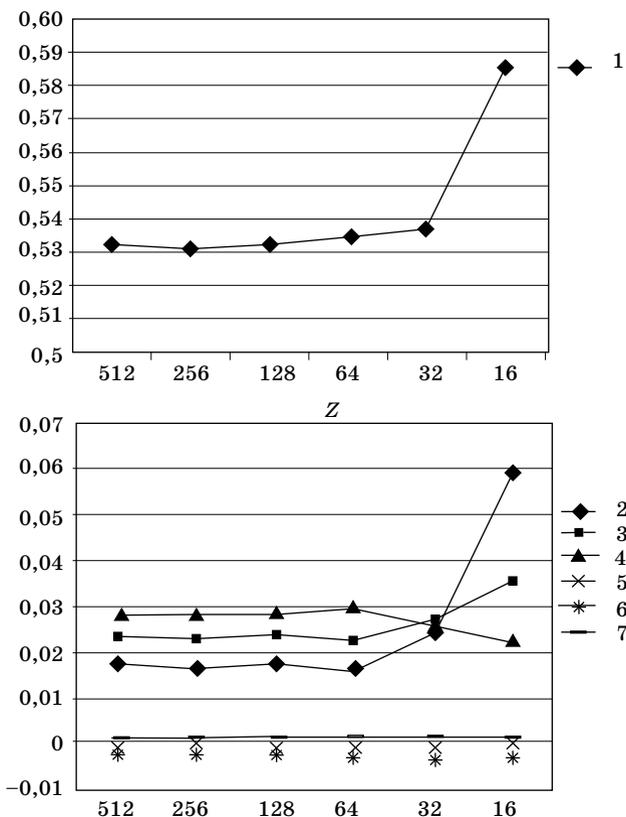


Рис. 80. Зависимость инвариантов X_u изображения буквы «С» от пространственного разрешения

должно быть меньше, чем 64×64 . Для семи инвариантных моментов X_u минимальное разрешение составило 128×128 .

4.4. Моменты Лежандра, Зернике и псевдо-Зернике

Моменты Лежандра порядка $(m+n)$ определяются следующим образом:

$$\lambda_{m,n} = \frac{(2m+1)(2n+1)}{4} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} P_m(x)P_n(y)f(x,y)dx dy,$$

где $m, n = 0, 1, 2, \dots$. Полиномы Лежандра $P_m(x)$ являются ортогональным базисным множеством функций на интервале $[-1, 1]$:

$$\int_{-\infty}^{+\infty} P_m(x)P_n(y) = \frac{2}{2m+1} \delta_{m,n},$$

где

$$\delta_{m,n} = \begin{cases} 1, & m = n \\ 0, & m \neq n \end{cases}.$$

Полином Лежандра n -го порядка равен

$$P_n(x) = \sum_{j=0}^n a_{n,j} x^j = \frac{1}{2^n n!} \frac{d^n}{dx^n} (x^2 - 1)^n.$$

По принципу ортогональности функция изображения $f(x,y)$ может быть записана как бесконечное разложение в ряд полиномов Лежандра внутри квадрата изображения $[-1 \leq x, y \leq 1]$:

$$f(x,y) = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \lambda_{m,n} P_m(x) P_n(y),$$

где моменты Лежандра $\lambda_{m,n}$ вычисляются по точкам квадрата заданного изображения. В случае, если моменты Лежандра имеют порядок $\leq N$, функция $f(x,y)$ может быть аппроксимирована непрерывной функцией, представляющей собой усеченный ряд:

$$f(x,y) = \sum_{m=0}^M \sum_{n=0}^m \lambda_{m-n,n} P_{m-n}(x) P_n(y).$$

Формула, связывающая моменты Лежандра и регулярные моменты, имеет вид:

$$\lambda_{m,n} = \frac{(2m+1)(2n+1)}{4} \sum_{j=0}^m \sum_{k=0}^n a_{m,j} a_{n,k} m_{j,k}.$$

Таким образом, моменты Лежандра зависят только от регулярных моментов того же или низшего порядка и наоборот.

Моменты Зернике инвариантны к повороту изображения. Комплексные моменты Зернике порядка n с m повторениями для функции изображения $f(x,y)$, заданной внутри круга $x^2 + y^2 \leq 1$, определяются следующим образом:

$$A_{m,n} = \frac{n+1}{\pi} \int_0^{2\pi} \int_0^{\infty} [V_{n,m}(\rho, \theta)]^* f(\rho \cos \theta, \rho \sin \theta) \rho d\rho d\theta$$

или

$$A_{m,n} = \frac{n+1}{\pi} \iint_{x^2+y^2 \leq 1} [V_{n,m}(\rho, \theta)]^* f(x,y) dx dy,$$

где n – положительное целое число или нуль; m – положительное или отрицательное целое число, удовлетворяющее условиям $n - |m|$ – четно, $|m| \leq n$; ρ – длина вектора от начала координат к точке (x,y) ; θ – угол между вектором ρ и осью x в направлении по часовой стрелке; $x = \rho \cos \theta$, $y = \rho \sin \theta$; символ $*$ означает «комплексно-сопряженный».

Для дискретного изображения интегралы заменяются на сумму:

$$A_{m,n} = \frac{n+1}{\pi} \sum_x \sum_y [V_{n,m}(\rho, \theta)]^* f(x,y), \quad x^2 + y^2 \leq 1.$$

Полиномы Зернике представляют собой множество комплекснозначных функций, ортогональных на внутренней области круга $x^2+y^2 \leq 1$. Обозначим множество этих полиномов $V_{n,m}(x,y)$:

$$V_{n,m}(x,y) = V_{n,m}(\rho \cos \theta, \rho \sin \theta) = R_{n,m}(\rho) \exp\{im\theta\},$$

где $i = \sqrt{-1}$.

Полиномы Зернике ортогональны и удовлетворяют условию

$$\iint_{x^2+y^2 \leq 1} [V_{n,m}(\rho, \theta)]^* V_{p,q}(\rho, \theta) \rho dx dy = \frac{\pi}{n+1} \delta_{p,n} \delta_{q,m},$$

где

$$\delta_{a,b} = \begin{cases} 1, & a = b \\ 0, & a \neq b. \end{cases}$$

Вещественнозначные радиальные полиномы $R_{n,m}(\rho)$ удовлетворяют соотношениям

$$\int_0^1 R_{n,m}(\rho) R_{p,m}(\rho) \rho d\rho = \frac{\pi}{2(n+1)} \delta_{p,n}$$

и определяются как

$$R_{n,m}(\rho) = \sum_{s=0}^{(n-|m|)/2} \frac{(-1)^s [(n-s)!] \rho^{n-2s}}{s! \left(\frac{n+|m|}{2} - s\right)! \left(\frac{n-|m|}{2} - s\right)!} = \sum_{k=|m|}^n \frac{B_{n,|m|,k} \rho^k}{(n-k) \text{—четно}}.$$

Функцию $f(x, y)$ можно разложить на основе полиномов Зернике внутри круговой области $x^2 + y^2 \leq 1$ следующим образом:

$$f(x, y) = \sum_n \sum_{\substack{m \\ (n-|m|)\text{-четно} \\ |m| \leq n}} A_{n,m} V_{n,m}(x, y),$$

где моменты Зернике $A_{n,m}$ вычисляются для данной круговой области $x^2 + y^2 \leq 1$.

Моменты Зернике и регулярные моменты связаны следующим образом:

$$A_{n,m} = \frac{n+1}{\pi} \sum_{k=|m|}^n \sum_{j=0}^q \sum_{l=0}^{|m|} w^l \binom{q}{j} \binom{|m|}{l} B_{n,|m|,k} m_{k-2j-l, 2j+l},$$

(n-k)-четно

где

$$w = \begin{cases} -i, & m > 0 \\ +i, & m \leq 0 \end{cases}, \quad q = \frac{1}{2}(k - |m|) \text{ и } i = \sqrt{-1}.$$

При вычислении моментов Зернике для заданного изображения центр изображения берется за начало координат, и используются только те точки, координаты которых попадают внутрь круга $x^2 + y^2 \leq 1$.

Рассмотрим поворот изображения на угол φ . Соотношение между моментом Зернике повернутого изображения $A'_{n,m}$ и моментом неповернутого изображения $A_{n,m}$ следующее:

$$A'_{n,m} = A_{n,m} \exp\{-im\varphi\}.$$

Это соотношение показывает, что моменты Зернике обладают простыми свойствами преобразования поворотов: при повороте каждый момент просто подвергается сдвигу по фазе, следовательно, абсолютные значения моментов Зернике для функции повернутого изображения остаются идентичными значениям до поворота. В связи с этим значение момента $|A_{n,m}|$ может рассматриваться как признак функции изображения, инвариантный к повороту. Заметим, что поскольку $A_{n,-m} = A_{n,m}^*$, то $|A_{n,-m}| = |A_{n,m}|$, поэтому можно сосредоточиться на вычислениях $|A_{n,m}|$ при $m \geq 0$. Список инвариантных к повороту моментов Зернике порядка от 0 до 12 и соответствующее каждому порядку число признаков (моментов) приведен в табл. 4.1 [37].

Таблица 4

Порядок	Моменты	Число моментов
0	$A_{0,0}$	1
1	$A_{1,1}$	1
2	$A_{2,0}; A_{2,2}$	2
3	$A_{3,1}; A_{3,3}$	2
4	$A_{4,0}; A_{4,2}; A_{4,4}$	3
5	$A_{5,1}; A_{5,3}; A_{5,5}$	3
6	$A_{6,0}; A_{6,2}; A_{6,4}; A_{6,6}$	4
7	$A_{7,1}; A_{7,3}; A_{7,5}; A_{7,7}$	4
8	$A_{8,0}; A_{8,2}; A_{8,4}; A_{8,6}; A_{8,8}$	5
9	$A_{9,1}; A_{9,3}; A_{9,5}; A_{9,7}; A_{9,9}$	5
10	$A_{10,0}; A_{10,2}; A_{10,4}; A_{10,6}; A_{10,8}; A_{10,10}$	6
11	$A_{11,1}; A_{11,3}; A_{11,5}; A_{11,7}; A_{11,9}; A_{11,11}$	6
12	$A_{12,0}; A_{12,2}; A_{12,4}; A_{12,6}; A_{12,8}; A_{12,10}; A_{12,12}$	7

Предложенные признаки Зернике инвариантны только к повороту. Чтобы получить инвариантность к масштабу и сдвигу, надо сначала нормализовать изображение с помощью регулярных моментов $m_{p,q}$. Инвариантность к сдвигу достигается преобразованием исходного изображения в новое, у которого моменты первого порядка $m_{1,0}$ и $m_{0,1}$ равны нулю. Для этого выполняется преобразование исходного изображения $f(x,y)$ в изображение $f(x+\bar{x}, y+\bar{y})$, где \bar{x} и \bar{y} – координаты центра исходного изображения: $\bar{x} = m_{1,0} / m_{0,0}$, $\bar{y} = m_{0,1} / m_{0,0}$.

Инвариантность к масштабу достигается увеличением или уменьшением каждого объекта таким образом, чтобы его момент нулевого порядка $m_{0,0}$ стал равным некоторому предопределенному значению β . Заметим, что в случае бинарного изображения $m_{0,0}$ равен общему числу точек объекта на изображении. Обозначим $f(x/a, y/a)$ – масштабированный вариант функции изображения $f(x,y)$. Тогда момент $m_{p,q}$ функции $f(x,y)$ и момент $m'_{p,q}$ функции $f(x/a, y/a)$ связаны следующим соотношением:

$$m'_{p,q} = \int \int_{x \ y} x^p y^q f(x/a, y/a) dx dy = \int \int_{x \ y} a^p x^p a^q y^q f(x,y) a^2 dx dy = \\ = \int \int_{x \ y} a^{p+q+2} x^p y^q f(x,y) dx dy = a^{p+q+2} \int \int_{x \ y} x^p y^q f(x,y) dx dy = a^{p+q+2} m_{p,q}.$$

Поскольку нужно получить $m'_{0,0} = \beta$, возьмем $a = \sqrt{\beta / m_{0,0}}$. Подставив $a = \sqrt{\beta / m_{0,0}}$ в $m'_{0,0}$, получим $m'_{0,0} = a^2 m_{0,0} = \beta$. Таким образом, инвариантность к масштабу достигается путем преобразования функции исходного изображения $f(x, y)$ в новую функцию $f(x/a, y/a)$, где $a = \sqrt{\beta / m_{0,0}}$.

Таким образом, функцию изображения $f(x, y)$ можно нормализовать по масштабу и сдвигу, если преобразовать ее в $g(x, y)$:

$$g(x, y) = \bar{f} \left(\bar{x} + \frac{x}{a}, \bar{y} + \frac{y}{a} \right),$$

где (\bar{x}, \bar{y}) – координаты центра $f(x, y)$; $a = \sqrt{\beta / m_{0,0}}$; β – предопределенное значение для числа точек на изображении.

Процесс нормализации по масштабу и сдвигу затрагивает два момента Зернике: $|A_{0,0}|$ и $|A_{1,1}|$. Значение $|A_{0,0}|$ стремится быть одинаковым для всех изображений, а значение $|A_{1,1}|$ равно нулю. Это видно из следующего соотношения:

$$\begin{aligned} A_{0,0} &= \frac{1}{\pi} \iint_{x^2+y^2 \leq 1} R_{0,0}(\rho) g(x, y) dx dy = \\ &= \frac{1}{\pi} \iint_{x^2+y^2 \leq 1} g(x, y) dx dy = \frac{1}{\pi} m_{0,0}. \end{aligned}$$

Поскольку $m_{0,0} = \beta$, то $|A_{0,0}| = \beta / \pi$ для всех нормализованных изображений. Поэтому $|A_{0,0}|$ не учитывается как признак, используемый при распознавании.

Что касается $|A_{1,1}|$, то:

$$\begin{aligned} A_{1,1} &= \frac{2}{\pi} \iint_{x^2+y^2 \leq 1} R_{1,1}(\rho) g(x, y) \exp\{i\theta\} dx dy ; \\ \operatorname{Re}[A_{1,1}] &= \frac{2}{\pi} \iint_{x^2+y^2 \leq 1} R_{1,1}(\rho) g(x, y) \cos \theta dx dy = \\ &= \frac{2}{\pi} \iint_{x^2+y^2 \leq 1} g(x, y) \rho \cos \theta dx dy = \\ &= \frac{2}{\pi} \iint_{x^2+y^2 \leq 1} g(x, y) x \cos \theta dx dy = \frac{2}{\pi} m_{1,0} \end{aligned}$$

и

$$\begin{aligned} \operatorname{Im}[A_{1,1}] &= \frac{2}{\pi} \iint_{x^2+y^2 \leq 1} R_{1,1}(\rho) g(x,y) \sin \theta dx dy = \\ &= \frac{2}{\pi} \iint_{x^2+y^2 \leq 1} g(x,y) \rho \sin \theta dx dy = \\ &= \frac{2}{\pi} \iint_{x^2+y^2 \leq 1} g(x,y) y \sin \theta dx dy = \frac{2}{\pi} m_{0,1}. \end{aligned}$$

Поскольку для всех нормализованных изображений $m_{1,0} = m_{0,1} = 0$, то для них $|A_{1,1}| = 0$, и поэтому $|A_{1,1}|$ также не используется как признак.

Моменты псевдо-Зернике. Полиномы Зернике, будучи инвариантными к повороту изображения вокруг начала координат, являются полиномами от x и y . Полиномы псевдо-Зернике представляют собой сходное ортогональное множество полиномов от x , y и r . Они обладают аналогичными полиномам Зернике свойствами, но отличаются определением вещественнозначных радиальных полиномов:

$$\begin{aligned} R_{n,m}(\rho) &= \sum_{s=0}^{(n-|m|)/2} (-1)^s \frac{(2n+1-s)!}{s! (n-|m|-s)! (n+|m|+1-s)!} \rho^{n-s} = \\ &= \sum_{k=|m|}^n S_{n,|m|,k} \rho^k, \end{aligned}$$

где $n = 0, 1, 2, \dots, \infty$, и m принимает положительные и отрицательные целые значения, удовлетворяющие только одному условию $|m| \leq n$. Простым подсчетом получим, что множество полиномов псевдо-Зернике состоит из $(n+1)^2$ линейно независимых полиномов порядка $\leq n$, тогда как множество полиномов Зернике состоит только из $1/2(n+1)(n+2)$ линейно независимых полиномов порядка $\leq n$ из-за дополнительного условия, что $n - |m|$ — четное число.

Моменты Зернике становятся моментами псевдо-Зернике, если использовать полиномы $R_{n,m}(\rho)$, исключая условие, что $n - |m|$ должно быть четным числом. Поскольку полиномы псевдо-Зернике также представляют собой полный набор функций, ортогональных на круговой области, то для разложения в ряд функции $f(x,y)$ также исключается условие, что $n - |m|$ — четное число, и $A_{n,m}$ и $V_{n,m}(x,y)$ становятся соответственно моментами и полиномами псевдо-Зернике.

В [38] проведено экспериментальное исследование применения моментов Ху, Лежандра и Зернике для распознавания и классификации изображений из базы данных ETN-80 при 8 различных классах (рис. 81 и рис. 82).

Объем обучающей и текстовой выборки для каждого класса составил, соответственно, 5 и 160 изображений. В качестве признаков использовалось 7 моментов Ху, 5 моментов Зернике и 16 моментов Лежандра. Результаты эксперимента приведены в табл. 5.

Таблица 5

Объект	Доля успешных распознаваний при использовании моментов:		
	Ху	Зернике	Лежандра
Яблоко	85%	75%	85%
Автомобиль	65%	65%	65%
Корова	65%	65%	65%
Чашка	65%	65%	72%
Собака	65%	75%	75%
Лошадь	85%	75%	85%
Персик	75%	65%	78%
Томат	65%	65%	65%
Среднее	71%	69%	74%

Как видно из табл. 5, наилучшие результаты были получены при использовании моментов Лежандра.

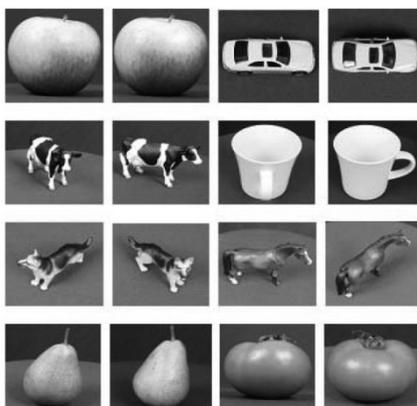


Рис. 81. Образцы изображений из базы данных ETN-80

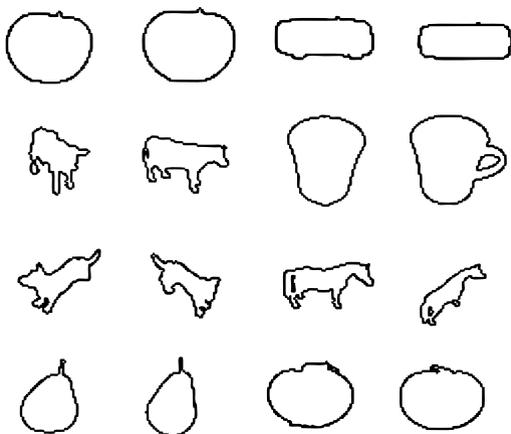


Рис. 82. Образцы контуров изображений
с рис. 75

В свою очередь, моменты Зернике и псевдо-Зернике показали наилучшее качество по сравнению с другими в плане чувствительности к зашумлению изображения, объему информации и способности воспроизводить точное изображение [39]. В частности, алгоритм воспроизведения изображений посредством коэффициентов моментов Зернике, предварительно соответствующим образом нормализованных с помощью регулярных моментов малого порядка, для задачи инвариантного распознавания символов дает наилучшие результаты по сравнению с другими методами, основанными на моментах.

4.5. Поворотные моменты, комплексные моменты и моментные инварианты для классификации неравномерно масштабированных изображений

Поворотные моменты порядка n с m повторениями определяют следующим образом:

$$D_{n,m} = \int_0^{2\pi} \int_0^{\infty} \rho^n \exp\{-im\theta\} f(\rho \cos \theta, \rho \sin \theta) \rho d\rho d\theta,$$

где $n = 0, 1, 2, \dots, \infty$, и m принимает любые положительные и отрицательные целые значения. Моменты Зернике и поворотные моменты связаны следующим образом:

$$A_{n,m} = \frac{n+1}{\pi} \sum_{\substack{k=|m| \\ (n-k)\text{—четно}}}^n B_{n,|m|,k} D_{k,m},$$

откуда следует, что поворотные моменты также можно получить из регулярных моментов

$$D_{n,m} = \sum_{j=0}^q \sum_{l=0}^{|m|} w^l, \binom{q}{j}, \binom{|m|}{l}, m_{n-2j-l, 2j+l},$$

где

$$w = \begin{cases} -i, & m > 0 \\ +i, & m \leq 0 \end{cases} \quad q = \frac{1}{2}(n - |m|).$$

Комплексные моменты. Понятие комплексных моментов было введено как простой путь получения моментных инвариантов. Комплексные моменты порядка $(p+q)$ определяются следующим образом:

$$C_{p,q} = \int_{-\infty-\infty}^{+\infty+\infty} \int_{-\infty-\infty}^{+\infty+\infty} (x+iy)^p (x-iy)^q f(x,y) dx dy,$$

где $p, q = 0, 1, 2, \dots, \infty$. Комплексные моменты порядка $(p+q)$ представляют собой линейную комбинацию с комплексными коэффициентами регулярных моментов $m_{r,s}$, удовлетворяющих условию $r+s = p+q$:

$$C_{p,q} = \sum_{r=0}^p \sum_{s=0}^q \binom{p}{r} \binom{q}{s} i^{p+q-(r+s)} (-1)^{q-s} m_{r+s, p+q-(r+s)}.$$

В полярной системе координат комплексные моменты порядка $(p+q)$ можно записать следующим образом:

$$C_{p,q} = \int_0^{2\pi} \int_0^{\infty} \rho^{p+q} \exp\{i(p-q)\theta\} f(\rho \cos \theta, \rho \sin \theta) \rho d\rho d\theta,$$

таким образом, они связаны с поворотными моментами соотношением

$$D_{n,m} = C_{(n-m)/2, (n+m)/2}$$

откуда следует, что моменты Зернике и комплексные моменты также связаны:

$$A_{n,m} = \frac{n+1}{\pi} \sum_{k=|m|}^n B_{n,|m|,k} C_{(k-m)/2, (k+m)/2} \cdot$$

(n-k)–четно

Моментные инварианты для классификации неравномерно масштабированных изображений. Далее представлен метод распознавания неравномерно горизонтально и вертикально масштабированных изображений, а также показано, что традиционные моментные инварианты не остаются инвариантными к изображению, масштабированному неравномерно в x и y направлениях. Описывается метод формирования моментных инвариантов, не изменяющихся при неравномерном масштабировании.

Центральные моменты, инвариантные к масштабу:

$$\eta_{p,q} = \frac{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \bar{x})^p (y - \bar{y})^q f(x,y) dx dy}{\left[\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y) dx dy \right]^{\frac{p+q+2}{2}}}.$$

Обозначим $g(x,y)$ неравномерно масштабированное изображение. Оно связано с исходным изображением следующим образом:

$$g(x,y) = f(ax, by).$$

Оценим нормализованные моменты изображения $g(x,y)$:

$$\gamma_{p,q} = \frac{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \bar{x}_1)^p (y - \bar{y}_1)^q g(x,y) dx dy}{\left[\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x,y) dx dy \right]^{\frac{p+q+2}{2}}},$$

где \bar{x}_1 и \bar{y}_1 – координаты центра изображения $g(x,y)$, которые равны

$$\bar{x}_1 = \frac{\bar{x}}{a}, \quad \bar{y}_1 = \frac{\bar{y}}{b}.$$

Определив $X = ax$ и $Y = by$, получим:

$$\begin{aligned} \gamma_{p,q} &= \frac{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left(x - \bar{x}\right)^p \left(y - \bar{y}\right)^q f(ax, by) dx dy}{\left[\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(ax, by) dx dy \right]^{\frac{p+q+2}{2}}} = \\ &= \frac{1}{a^{p+1} b^{q+1}} \frac{\int \int (X - \bar{x})^p (Y - \bar{y})^q f(X, Y) dX dY}{\left[\int \int f(X, Y) dX dY \right]^{\frac{p+q+2}{2}}} = \left(\frac{b}{a}\right)^{\frac{p-q}{2}} \eta_{p,q}. \\ &= \frac{1}{(ab)^{\frac{p+q+2}{2}}} \left[\int \int f(X, Y) dX dY \right]^{\frac{p+q+2}{2}} \end{aligned}$$

Из последнего уравнения понятно, что инвариантность к масштабу, полученная в предыдущих подразделах, возможна только при $a = b$. Когда $a \neq b$, изображение не может быть классифицировано с использованием указанных инвариантных моментов. Чтобы классифицировать изображение при $a \neq b$, возможны следующие инварианты:

$$R_{p,q} = \gamma_{p,q} \gamma_{q,p} \left(\frac{b}{a}\right)^{\frac{p-q}{2}} \eta_{p,q} \left(\frac{b}{a}\right)^{\frac{q-p}{2}} \eta_{q,p} = \eta_{p,q} \eta_{q,p},$$

$R_{p,q}$ не зависит от значений a и b , а, значит, будет правильно классифицировать неравномерно масштабированные изображения.

4.6. Нейронные сети высших порядков

Классификация и распознавание изображений, инвариантных к сдвигу, повороту и масштабу, обычно выполняются в два этапа:

- из классифицируемых образцов извлекается подходящий набор признаков с нужными свойствами инвариантности;
- выделенные признаки предъявляются классификатору, который разделяет пространство признаков на решающие области, соответствующие каждому классу изображений.

Далее исследуются возможности реализации инвариантности на нейронных сетях высших порядков при решении задач распознавания изображений. Эти сети обрабатывают произведения значений пикселей на изображении. Сети высших порядков можно настроить таким образом, чтобы кодировать в значениях их синаптических весов свойство инвариантности к двумерному преоб-

разованию координат. Таким образом, они могут рассматриваться как эквивалент обычным сетям первого порядка, обрабатывающим инвариантные признаки, равные соответствующим суммам произведений значений пикселей. С этой точки зрения, определение данных сумм представляет собой стадию предобработки, и нейронная сеть берет на себя традиционную роль классификатора признаков.

Следует заметить, что нейронные сети высших порядков инвариантны к аффинным преобразованиям только для непрерывных изображений. Точные преобразования, соответствующие произвольному сдвигу, повороту и масштабу, невыполнимы на квадратной решетке дискретного изображения. Например, под воздействием преобразования поворота или масштабирования точка может быть преобразована сама в себя, что делает это преобразование необратимым.

Далее приведено описание структур нейронных сетей третьего порядка для распознавания образов: инвариантного к сдвигу, повороту и масштабу изображения [40,41].

Сети высших порядков для инвариантного распознавания образов. Рассмотрим образ, представленный на квадратной растровой решетке, состоящей из N точек. Цвет точки обозначим s_i , где $i = 1, 2, \dots, N$ – номер точки. Рассмотрим однослойную нейронную сеть, обрабатывающую произведение вида $s_{i_1} \cdot s_{i_2} \dots s_{i_p}$, где p – порядок сети ($p \leq N$), и i_m – индекс номера точки ($i_m = 1, 2, \dots, N$, $m = 1, 2, \dots, p$). Тогда выход j -го нейрона (рис. 83) равен

$$Q_j(s) = f \left(\sum_{i_1} \sum_{i_2} \dots \sum_{i_p} w_{i_1, i_2, \dots, i_p, j} \cdot s_{i_1} \cdot s_{i_2} \dots s_{i_p} \right),$$

где f – функция активации j -го нейрона.

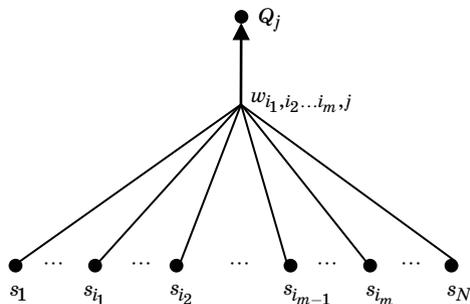


Рис. 83. Связность в сети высшего порядка

Расположение i -й точки определяется радиус-вектором $r(i)$, начало которого находится в центре экрана. Необходимо добиться, чтобы выход j -го нейрона Q_j был инвариантен к группе преобразований G координат пиксельной решетки. Под воздействием преобразования R , принадлежащего группе преобразований G , i -я точка изменит свой цвет s_i на цвет s'_i . Другими словами, новый цвет s'_i i -й точки преобразованного изображения является цветом $s_{i'}$ другой точки с номером i' непреобразованного изображения:

$$s'_i = s_{i'} \text{ при } r(i') = Rr(i).$$

Под воздействием данного преобразования R выход сети примет вид

$$\begin{aligned} Q'_j(s) &= f \left(\sum_{i_1} \sum_{i_2} \dots \sum_{i_p} w_{i_1, i_2 \dots i_p, j} \cdot s'_{i_1} \cdot s'_{i_2} \dots s'_{i_p} \right) = \\ &= f \left(\sum_{i_1} \sum_{i_2} \dots \sum_{i_p} w_{i_1, i_2 \dots i_p, j} \cdot s_{i'_1} \cdot s_{i'_2} \dots s_{i'_p} \right), \end{aligned}$$

где $r(i'_m) = Rr(i_m)$, $m = 1, 2 \dots p$. Данное уравнение можно переписать следующим образом:

$$Q'_j(s) = f \left(\sum_{i_1} \sum_{i_2} \dots \sum_{i_p} w_{k_1, k_2 \dots k_p, j} \cdot s_{i_1} \cdot s_{i_2} \dots s_{i_p} \right),$$

где $r(k_m) = R^{-1}r(i_m)$, $m = 1, 2 \dots p$, R^{-1} – преобразование, обратное R . Поскольку R^{-1} является элементом G , значит Q_j будет инвариантным к воздействию всех элементов из G , если веса j -го нейрона выбраны таким образом, что

$$w_{i_1, i_2 \dots i_p, j} = w_{k_1, k_2 \dots k_p, j},$$

и при этом в G существует преобразование, связывающее $r(i_m)$ и $r(k_m)$ для всех $m = 1, 2 \dots p$. Таким образом, нужная инвариантность оказывается встроенной в архитектуру сети при наложении соответствующих ограничений на синаптические веса.

Для того чтобы гарантировать инвариантность к преобразованию, надо устанавливать $w_{i_1, i_2 \dots i_p, j} = w_{k_1, k_2 \dots k_p, j}$ всякий раз, когда координаты любой пары точек (i_a, i_b) ($a, b = 1 \dots p$) могут быть получены из координат соответствующей пары точек (k_a, k_b) через преобразование G .

Например, чтобы получить сеть второго порядка, выход которой инвариантен к сдвигу, надо установить равными все веса $w_{i_1, i_2, j}$ и $w_{k_1, k_2, j}$ для всех линейных сегментов (i_1, i_2) и (k_1, k_2) , которые могут быть преобразованы один в другой через этот сдвиг, т. е. веса, для которых сегменты равны по длине и параллельны один другому. Примеры распознавания изображений на нейронных сетях высших порядков, инвариантных к различным группам аффинных преобразований показаны на рис. 84.

Заметим, что эта процедура встраивания в архитектуру сети инвариантностей к группе преобразований присваивает p -кортежам точек на плоскости эквивалентные классы C_h , $h = 1, 2, \dots$. Поэтому мы можем заменить обозначение весов $w_{i_1, i_2, \dots, i_p, j}$ на обозначение $w_{h, j}$, где h – это общий индекс, соответствующий классу C_h . Выход узла первого слоя сети высшего порядка теперь можно записать как

$$Q_j(s) = f \left(\sum_h w_{h, j} \sum_{(i_1 \dots i_p) \in C_h} s_{i_1} \cdot s_{i_2} \dots s_{i_p} \right),$$

что эквивалентно однослойной нейронной сети первого порядка (персептрон) с весами $w_{h, j}$ и действительными входами:

$$I_h = \sum_{(i_1 \dots i_p) \in C_h} s_{i_1} \cdot s_{i_2} \dots s_{i_p},$$

представляющими инвариантные признаки изображений, которые сеть должна классифицировать.

Во всех случаях треугольник T представляет изображение, которое сеть должна распознать.

Значения смещений нейронов (т. е. веса от единичного фиксированного входа, используемые в сетях первого порядка) не влияют ни на инвариантность свойств сети, ни на последующие слои весов. Будем называть эти многослойные сети сетями высшего порядка, хотя произведения вида $s_{i_1} \cdot s_{i_2} \dots s_{i_p}$ получаются только для узлов первого скрытого слоя. Таким образом, сеть имеет такую же архитектуру, как и многослойные сети первого порядка (персептроны), а задачи дополнительного программирования ограничиваются определением классов эквивалентности и вычислением действительных входов I_h .

Очевидно, что точные преобразования, соответствующие сдвигу на произвольные векторы, повороту на произвольные углы и масштабированию на произвольные множители, невозможно реализовать

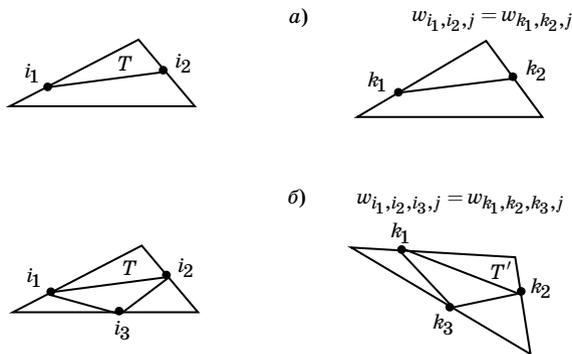


Рис. 84. Правила построения сетей высших порядков для инвариантного распознавания:

а – инвариантное к сдвигу и масштабу распознавание посредством сети второго порядка: для параллельных отрезков i_1, i_2 и k_1, k_2 веса $w_{i_1, i_2, j}$ и $w_{k_1, k_2, j}$ установлены равными друг другу; б – инвариантное к сдвигу, повороту и масштабу распознавание посредством сети третьего порядка: для подобных треугольников i_1, i_2, i_3 и k_1, k_2, k_3 веса $w_{i_1, i_2, i_3, j}$ и $w_{k_1, k_2, k_3, j}$ установлены равными друг другу

вать на квадратной решетке, потому что при этих преобразованиях одни точки решетки не обязательно трансформируются в другие точки решетки. Например, соседние точки решетки могут быть преобразованы в те же точки под воздействием преобразования масштабирования решетки, которое поэтому является необратимым. Тем не менее, мы можем использовать некоторые из главных моментов приведенного рассуждения для построения сетей, гарантирующих приближенную инвариантность к сдвигу, масштабу и повороту.

При рассмотрении изображения, представленного бинарными входами (0 – для пустых мест на экране; 1 – для «активных» точек изображения), нелинейное масштабирование на коэффициент ρ приводит к масштабированию числа активных точек приблизительно на ρ^2 и к масштабированию действительных входов сети I_h на ρ^4 . Мы можем компенсировать этот эффект через нормализацию действительного входного вектора ($I_1, I_2 \dots$) на стандартную евклидову длину.

Рассмотрим реализацию инвариантности одновременно к сдвигу, масштабу и повороту изображения. Заметим, что два любых линейных сегмента на плоскости могут быть преобразованы один в другой через комбинацию этих преобразований. Значит, если ис-

пользуется сеть второго порядка, то все ее веса будут равными друг другу, т. е. будет невозможно классифицировать различные изображения. Однако сеть третьего порядка генерирует выход, приблизительно инвариантный ко всем трем преобразованиям, если два веса $w_{i_1, i_2, i_3, j}$ и $w_{k_1, k_2, k_3, j}$ установлены равными всякий раз, когда треугольники с вершинами в i_1, i_2, i_3 и k_1, k_2, k_3 соответственно могут быть преобразованы один в другой через комбинацию трех произведений. Это означает, что два треугольника должны быть подобны друг другу, иметь равные углы, встречающиеся в одном и том же порядке, когда периметры треугольников пересекаются по часовой стрелке (рис. 84, б). К тому же, все изображения должны быть нормализованы для компенсации воздействия на действительные выходы I_n , вызванного преобразованием масштабирования.

Использование нейронных сетей третьего порядка для выделения признаков изображения. Для выделения признаков входного двумерного изображения введем в структуру нейронной сети понятие сферичности.

Сферичность треугольного преобразования, отображающего один треугольник в другой, является мерой подобия двух треугольников. Как показано на рис. 85, при преобразовании треугольника вписанный круг одного треугольника преобразуется во вписанный эллипс другого треугольника. Сферичность определяется как отношение геометрической средней к арифметической средней длин главных осей вписанного эллипса, т. е.

$$Y_g = \frac{\sqrt{d_1 d_2}}{(d_1 + d_2)/2} = \frac{2\sqrt{d_1 d_2}}{d_1 + d_2}.$$

Сферичности двух подобных треугольников одинаковы. Заметим, что значения сферичности изменяются в диапазоне между 0 и 1, что понятно из последнего уравнения. Параметр сферичности инвариантен к положению, сдвигу и масштабу изображения.

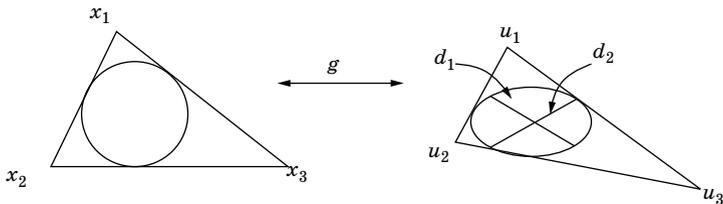


Рис. 85. Вписанный круг одного треугольника преобразуется во вписанный эллипс другого треугольника

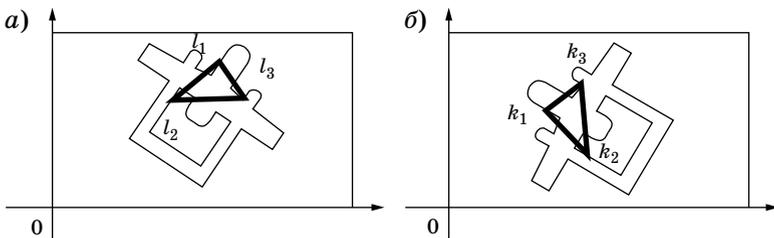


Рис. 86. Подобные треугольники одного и того же преобразованного изображения

Как видно на рис. 86, а и 86, б инвариантность к сдвигу, повороту и масштабу достигается посредством установки всех весов третьего порядка равными для множества входов j , k и l , образующих подобные треугольники.

Примем левый верхний угол изображения за начало координат. Возьмем три любые точки, не лежащие на одной прямой, за вершины эталонного треугольника. Для любых трех точек изображения существует свое отношение сферичности к эталонному треугольнику. Разделим диапазон сферичности $[0, 1]$ на l подынтервалов. Тогда можно объединить треугольники, сферичность которых попадает в один подынтервал, в один класс. Таким образом, подынтервал сферичности будет включать множество треугольников, подобных до определенной степени эталонному треугольнику.

Нейронная сеть третьего порядка формирует из исходного изображения вектор признаков $F = [f_1, f_2 \dots f_l]$, где l – число классов (рис. 87) [40]. Этот вектор признаков затем подается на вход многослойного персептрона, выполняющего классификацию изображений. Элемент вектора признаков f_m равен

$$F_m = W_m t_m,$$

где W_m – вес m -го признака; t_m – выход m -й пластины (рис. 87), равный числу треугольников m -го класса, содержащихся во входном изображении:

$$t_m = \sum_{(i_1, i_2, i_3) \in C_m} x_{i_1} x_{i_2} x_{i_3}.$$

Получившийся таким образом вектор признаков F подается на вход нейронной сети первого порядка (многослойной нейронной сети с полными последовательными связями), выступающей в качестве классификатора.

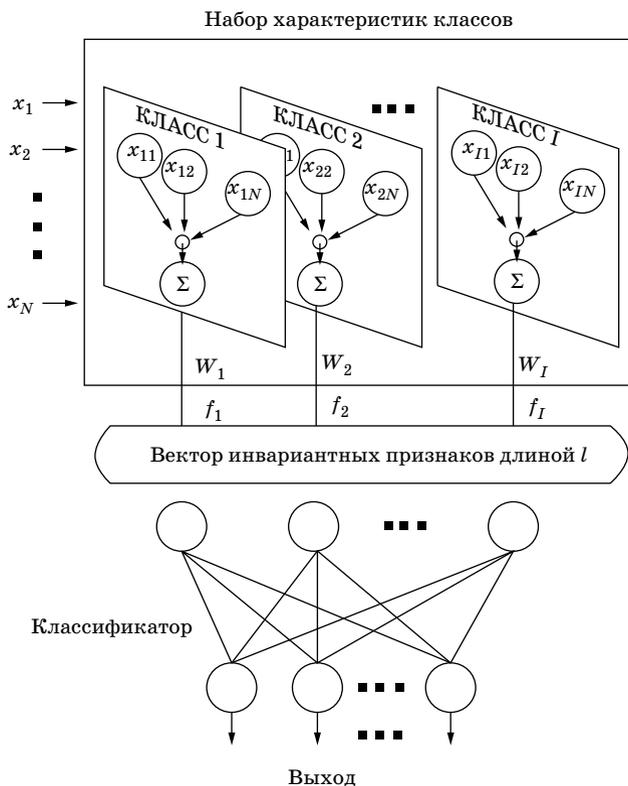


Рис. 87. Реализация инвариантной системы

Использование адаптивного усреднения в нейронной сети второго порядка для реализации инвариантности к циклическим сдвигам и поворотам на 90° . Рассмотрим архитектуру нейронной сети для инвариантности к циклическим сдвигам одномерного входного вектора из $N = 2^n$ элементов.

Нейронная сеть имеет архитектуру прямого распространения сигнала и своей структурой, подобной двоичному дереву, похожа на граф. Размер сети определяется числом входных элементов. Слои сети пронумерованы от n для входного слоя и до 0 для выходного слоя. Число нейронов в i -м слое ($0 \leq i \leq n$) равно 2^i . Каждый нейрон получает входной сигнал только от двух нейронов предшествующего слоя, т. е. k -й нейрон ($0 \leq k \leq 2^i$) i -го слоя получает входной сигнал от k -го и $(k+2^i)$ -го нейронов слоя $(i+1)$. Таким образом, сеть имеет малосвязную архитектуру, что обеспечивает быстрый проход сети.

Нейроны одного слоя совместно используют все веса и функции активации. Поэтому действие i -го слоя можно полностью описать с помощью функции активации f_i . На рис. 88 показана архитектура сети для одномерного входного образа длиной 8 [40].

Кроме того, функции узла симметричны относительно двух входных аргументов x и y , т. е. $f(x, y) = f(y, x)$.

Обозначим выходной вектор i -го слоя $O^{(i)}$, $O^{(n)} = [s_0 \dots s_{N-1}]$ – входной вектор сети и $O^{(0)}$ – выход сети. Тогда выход k -го нейрона i -го слоя сети равен

$$O^{(i)}[k] = f_i(O^{(i+1)}[k], O^{(i+1)}[k + 2^i]), \quad 0 \leq k \leq 2^i.$$

Такая сеть дает выход, инвариантный к циклическим сдвигам и отражениям входных элементов.

Формальное выражение выхода нейрона i -го слоя O_i имеет следующий вид:

$$O_i = f_i \left(\sum_{j=1}^2 w_{i,j} s_j + w_{i,0} \right),$$

где s_j – значение j -го, связанного с данным, нейрона предыдущего слоя; $w_{i,j}$ – вес связи от j -го, связанного с данным, нейрона предыдущего слоя к данному нейрону i -го слоя; $w_{i,0}$ – пороговый элемент; f_i – нелинейная передаточная функция активации i -го слоя.

Выбор весов должен удовлетворять требованию симметричности уравнения для $O^{(i)}[k]$. Одинаковые значения обоих выходных весов

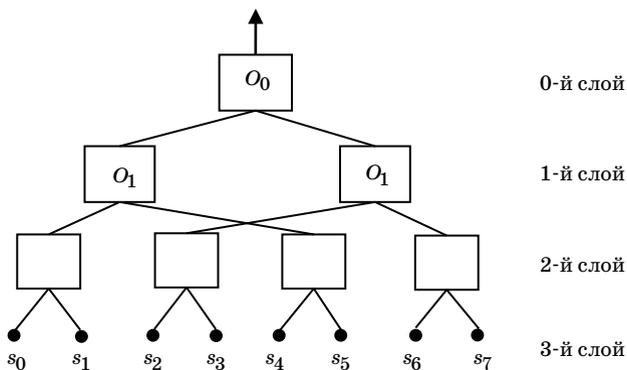


Рис. 88. Архитектура нейронной сети, инвариантной к циклическим сдвигам одномерного входного образа длиной 8

одного нейрона $w_{i,1} = w_{i,2} = w_i$ не являются подходящим решением, поскольку образы с равными суммами s_1 и s_2 , но с разными значениями не будут различаться. Поэтому необходимы различные значения для $w_{i,1}$ и $w_{i,2}$, нарушающие симметричность функции f_i . Чтобы решить эту проблему, передаточная функция активации f_i заменяется парой аналогичных функций с переставленными весами от входов. Также вводится дополнительный слой, в котором выход каждой передаточной функции f_i взвешивается на фиксированный вес $1/2$, затем обрабатывается второй передаточной функцией u_i . Тогда выходное выражение нейрона i -го слоя $Q_{(i)}$ примет следующий вид:

$$O_i = u_i \left(\frac{1}{2} f_i(w_{i,1}s_1 + w_{i,2}s_2 + w_{i,0}) + \frac{1}{2} f_i(w_{i,2}s_1 + w_{i,1}s_2 + w_{i,0}) \right).$$

Данная функция симметрична и различает два входа s_1 и s_2 .

Нами была описана архитектура сети, инвариантной к циклическим сдвигам одномерного входного образа. Та же сеть может

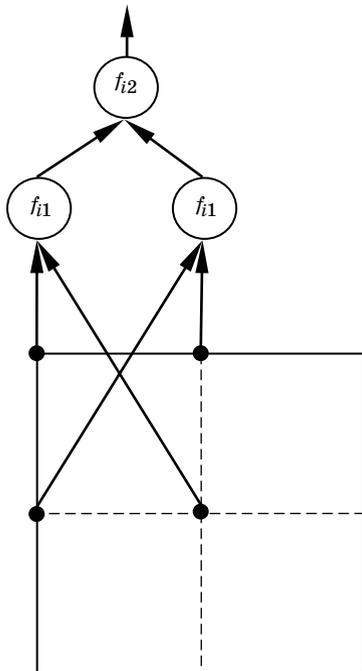


Рис. 89. Связь четырех входных элементов

использоваться для реализации инвариантности к циклическим сдвигам двумерного входного образа. Для этого входные элементы высшего порядка выстраиваются в один столбец или в один ряд. Сеть сначала вычисляет инвариантности к циклическим сдвигам вправо-влево, а затем к сдвигам вверх-вниз, или наоборот. Другой возможностью выполнять адаптивное усреднение на двумерных образах является одномерная сеть, которую можно перестроить в двумерную с четырьмя входными элементами, соединенными в один. Если это соединение выполняется так, как показано на рис. 89, то добавок к инвариантности к циклическим сдвигам достигается инвариантность к повороту на углы, кратные 90° .

4.7. Комбинированные нейронные сети

В последнее время широкое распространение получили комбинированные нейронные сети, в которых многослойный персептрон объединен с нейронной сетью для предварительной обработки изображения. Этот метод аналогичен методу реализации инвариантности (подразд. 4.3–4.5), с той разницей, что предварительная обработка изображения (т. е. выделение признаков изображения) выполняется не с использованием моментов различных порядков, а с помощью создания так называемой инвариантной нейронной сети, выходом которой являются инвариантные к некоторому аффинному преобразованию признаки. Эти признаки подаются на вход второй нейронной сети, выполняющей их классификацию. Как правило, в качестве классифицирующей нейронной сети используется многослойный персептрон или, другими словами, многослойная нейронная сеть с полными последовательными связями.

Рассмотрим архитектуры комбинированных нейронных сетей, реализующих инвариантность к сдвигу и повороту.

Инвариантность к сдвигу. Комбинированная нейронная сеть, реализующая инвариантность к сдвигу вправо-влево и вверх-вниз показана на рис. 90. Инвариантная сеть состоит из множества пластин. Пластина инвариантной к сдвигу сети представлена на рис. 91 [39].

Пластина состоит из множества нейронов, называемых ADALINE, и одного мажоритарного нейрона (на рис. 91 он обозначен MAJ). Каждый нейрон ADALINE связан со всеми точками входного изображения посредством весов. Входы нейронов ADALINE подаются на вход мажоритарного нейрона. Выходом пластины является выходное значение этого мажоритарного нейрона.

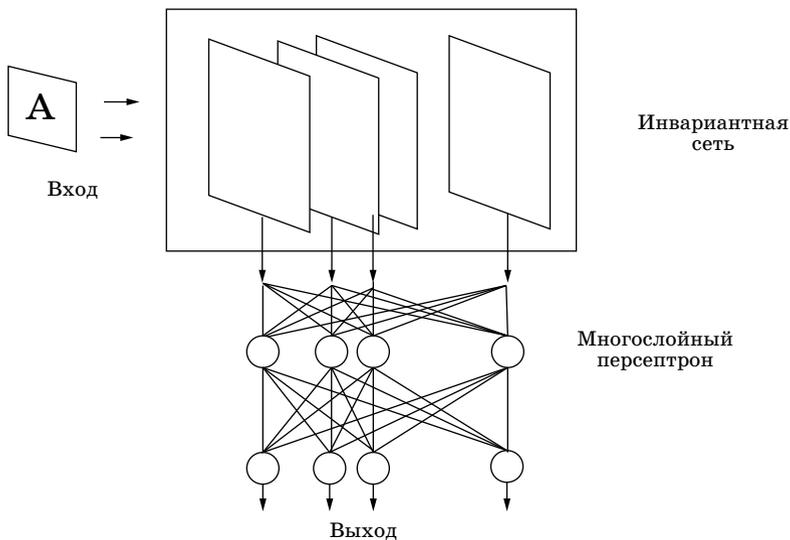


Рис. 90. Инварантная к сдвигу сеть + многослойный перцептрон (MLP)

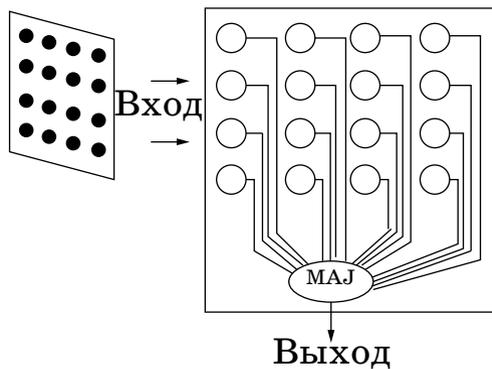


Рис. 91. Пластина инварантной к сдвигу сети

Таким образом, число выходов инварантной к сдвигу сети равно числу пластин, содержащихся в ней. Вектор выходных значений инварантной к сдвигу сети подается на вход многослойного перцептрона (MLP), распознающего изображение.

Обозначим через \mathbf{W} квадратную матрицу весов, а $\mathbf{T}_{D1}(\mathbf{W})$ матрицу с теми же элементами, что и \mathbf{W} , но сдвинутую вниз на один ряд, так что последний ряд матрицы \mathbf{W} становится первым рядом матри-

цы $T_{D1}(W)$. Обозначим через $T_{R1}(W)$ матрицу с теми же элементами, что и W , но сдвинутую вправо на один столбец, так что последний столбец матрицы W становится первым столбцом матрицы $T_{R1}(W)$.

Итак, $T_{Di+1}(W) = T_{D1}(W)T_{Di}(W)$ и $T_{Ri+1}(W) = T_{R1}(W)T_{Ri}(W)$. Тогда получим следующую матрицу весов для массива узлов ADALINE первой пластины:

$$\begin{pmatrix} (W_1) & T_{R1}(W_1) & T_{R2}(W_1) & T_{R3}(W_1) \\ T_{D1}(W_1) & T_{R1}T_{D1}(W_1) & T_{R2}T_{D1}(W_1) & T_{R3}T_{D1}(W_1) \\ T_{D2}(W_1) & T_{R1}T_{D2}(W_1) & T_{R2}T_{D2}(W_1) & T_{R3}T_{D2}(W_1) \\ T_{D3}(W_1) & T_{R1}T_{D3}(W_1) & T_{R2}T_{D3}(W_1) & T_{R3}T_{D3}(W_1) \end{pmatrix}.$$

Все выходы нейронов ADALINE взвешиваются на одно и то же значение и тракуются симметрично мажоритарным элементом пластины, поэтому выход пластины является инвариантным к сдвигам вправо-влево и вверх-вниз.

В [39] предлагается использовать случайные значения для получения матрицы весов первой пластины W_1 . Веса для остальных пластин должны сохранять ту же симметрию относительно сдвига, что и первая пластина. Матрица весов для i -й пластины W_i также составляется из случайных значений, но для всех пластин они должны быть разными.

Использование случайных значений для матриц весов W_i не гарантирует однозначного соответствия для преобразовывающей сети. Кроме того, сеть со случайными значениями весов очень чувствительна к шуму. Поэтому в [41] предлагается алгоритм обратного распространения ошибки (*back propagation*) для одновременного обучения преобразовывающей сети и MLP (многослойного перцептрона), что позволяет реализовать инвариантность к сдвигу и устойчивость к шуму.

Чтобы использовать алгоритм обратного распространения, надо внести некоторые изменения в преобразовывающую сеть. Для нейронов пластин должна использоваться дифференцируемая, а не знаковая (*sgn*) функция. Подходящей дифференцируемой функцией является сигмоидальная функция активации с порогом. Мажоритарный нейрон также должен иметь сигмоидальную функцию активации (без порога), а все его входы должны иметь одинаковые веса. Такие условия налагают симметрию на входные значения, гарантирующую инвариантность к сдвигу.

Использование дифференцируемой функции для преобразовывающей сети означает, что алгоритм обратного распространения

ошибки можно применить для всей сети в целом (предобрабатывающей сети + MLP). Остается только учесть тот факт, что в обрабатывающей сети есть одинаковые веса. Рассмотрим, как изменятся уравнения для таких весов.

Обозначим через M число узлов в пластине, которое равно числу входов обрабатывающей сети. Имеем $M+1$ различных весов (число входов плюс порог) и M копий каждого веса, каждая копия находится в разных узлах.

Обозначим через w_i^k i -й вес k -й пластины при $0 \leq i \leq M$. Порог реализуется как дополнительный фиксированный вход, установленный в единицу, и он будет включен в сумму с нулевым индексом. Изменение весов выглядит следующим образом:

$$w_i^k(n+1) = w_i^k(n) - \mu \left. \frac{\partial E_p}{\partial w_i^k} \right|_{w^k(n)},$$

где выражение для градиента

$$\frac{\partial E_p}{\partial w_i^k} = \frac{\partial}{\partial w_i^k} \left(\frac{1}{2} \sum_{r=1}^N (d_r^p - o_r)^2 \right),$$

где N – число выходов MLP; d_r^p – желаемый выход r -го узла p -го образа; o_r – действительный выход r -го узла выходного слоя MLP.

Обозначим через $w_{l,j}^k$ вес связи от x_l -го входа к j -му узлу k -й пластины. Чтобы установить взаимосвязь между w_i^k и $w_{l,j}^k$, положим, без потери общности, что $w_{i,1}^k$ является копией веса w_i^k . Значит, $w_{l,j}^k$ является копией веса w_i^k , если $l = (i+j-1) \bmod(M)$.

Получим

$$\frac{\partial E_p}{\partial w_i^k} = \sum_{j=1}^M \frac{\partial E_p}{\partial w_{(i+j-1) \bmod(M),j}^k}.$$

Чтобы вычислить каждое слагаемое, рассмотрим функцию, реализуемую каждой пластиной:

$$O^k = \text{sgn}(y^k) = \text{sgn} \left(\frac{1}{K} \sum_{j=1}^M \text{sgn} \left(\sum_{l=1}^M w_{l,j}^k x_l + \theta_k \right) \right),$$

где O^k – выход k -й пластины; y^k – промежуточная сумма выходного узла k -й пластины; K – вес выходного узла, на который умножаются выходные значения предыдущих узлов пластины.

Применяя цепное правило, получим

$$\frac{\partial E_p}{\partial w_{l,j}^k} = \frac{\partial E_p}{\partial y^k} \frac{\partial y^k}{\partial w_{l,j}^k}.$$

Поскольку выход сети каждой пластины связан с входным слоем MLP, мы можем вычислить $\frac{\partial E_p}{\partial y^k}$ так, как будто выходной узел пластины является промежуточным узлом MLP:

$$-\frac{\partial E_p}{\partial y^k} = \delta^k = O^k \left(1 - O^k\right) \sum_{p=1}^{N_1} \delta_p^1 \cdot w_{k,p}^1,$$

где δ^k – значение ошибки для k -й пластины; O^k – выход k -й пластины; N_1 – число узлов первого слоя MLP (скрытый слой); δ_p^1 – значение ошибки для p -го узла 1-го слоя MLP; $w_{k,p}^1$ – вес связи от k -й пластины к p -му узлу 1-го слоя MLP.

Оставшаяся часть выражения вычисляется следующим образом:

$$\frac{\partial y^k}{\partial w_{l,j}^k} = \frac{\partial \left(\frac{1}{K} \sum_{j=1}^M \operatorname{sgn} \left(\sum_{l+1}^M w_{l,j}^k x_l \right) \right)}{\partial w_{l,j}^k} = \frac{1}{K} o_j^k \left(1 - o_j^k\right) x_l,$$

где o_j^k – выход j -го узла k -й пластины.

Подставив два последних уравнения в $\frac{\partial E_p}{\partial w_{l,j}^k}$, получим

$$-\frac{\partial E_p}{\partial w_{l,j}^k} = \delta^k \frac{1}{K} o_j^k \left(1 - o_j^k\right) x_l,$$

где

$$\delta^k = O^k \left(1 - O^k\right) \sum_{p=1}^{N_1} \delta_p^1 w_{k,p}^1.$$

Конечное выражение, соответствующее общему изменению весов:

$$w_i^k(n+1) = w_i^k(n) - \mu \delta^k \sum_{j=1}^M \frac{1}{K} o_j^k \left(1 - o_j^k\right) x_{(i+j-1) \bmod(M)}.$$

Используя данное выражение можно применить алгоритм обратного распространения ошибки одновременно к преобразовывающей сети и MLP.

Инвариантность к повороту. Аналогично инвариантности к сдвигу реализуется инвариантность к повороту с помощью комбинированной нейронной сети, в которой многослойный персептрон объединен с нейронной сетью для предварительной обработки изображения. Такая комбинированная нейронная сеть показана на рис. 92. Сеть для предварительной обработки состоит из множества пластин (рис. 93) [40].

Предобработчик (см. рис. 92) состоит из множества пластин, в которых элементы, обозначенные через M , являются сигмоидальными нейронами, а элементы, обозначенные через M , являются мажоритарными нейронами [40]. Каждая пластина имеет один выход, который нечувствителен к повороту и является входным сигналом

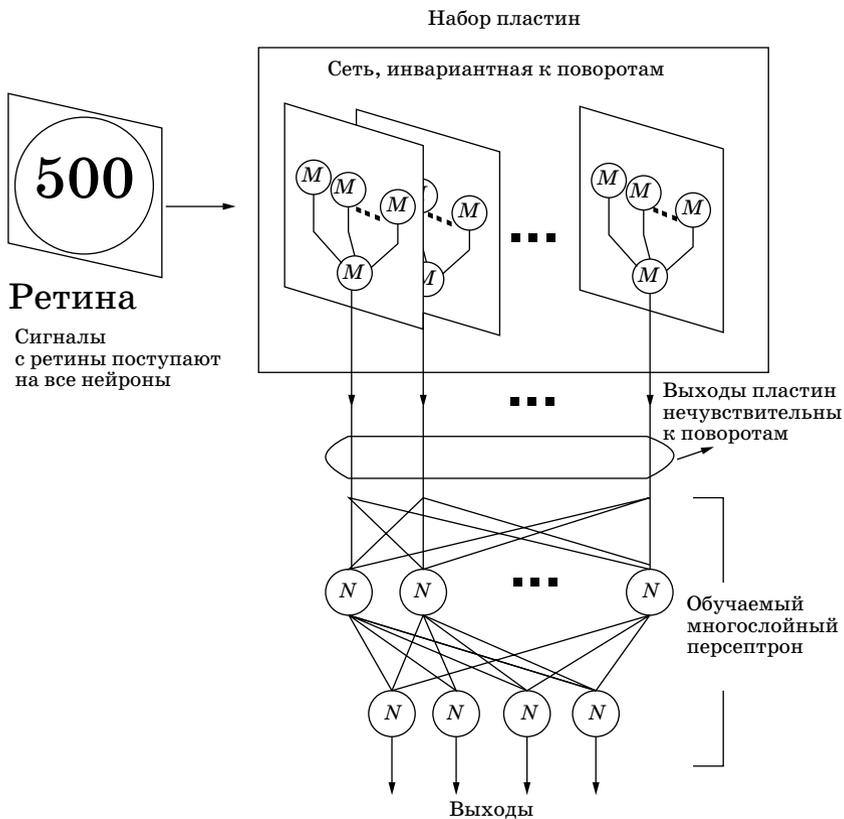


Рис. 92. Инвариантная к повороту сеть + многослойный персептрон (MLP)

для многослойного персептрона. Поэтому число входных нейронов многослойного персептрона равно числу пластин предобработчика.

На рис. 93 показана одна пластина. Возникает проблема, как определить веса пластины так, чтобы выходы предобработчика были инвариантны к повороту. Расположение весов по типу квадратной матрицы, показанной на рис. 94, б, не является подходящим решением для инвариантности к повороту. Поэтому используется круговая матрица весов, показанная на рис. 94, а [40].

Пусть соответствующие точки на ретине также расположены в круговом образе. Предположим, что пластины предобработчика инвариантны к повороту на каждые 30° и каждые из 12 нейронов на пластине имеют круговую матрицу весов. Если окружность и радиус наиболее удаленного круга на рис. 94, а разделить на N_1 и N_2 элементов, соответственно, то число весов к каждому нейрону составит $NN = (N_2 - 1)N_1 + 1$. Для случая, когда $N_1 = 12$ и $N_2 = 4$, образ на ретине должен иметь такую же круговую матрицу из 37 точек.

Обозначим матрицу весов крайнего левого нейрона на рис. 93 через (W_1) ; эта матрица состоит из NN элементов (рис. 94, в). Пусть (W_1) при повороте на 30° преобразуется в матрицу весов $R_{30}(W_1)$ (рис. 94, г). Оператор R_{30} представляет собой поворот на 30° . Тогда все веса, кроме фиксированного центрального веса, повернуты. Веса следующего нейрона на пластине равны $R_{30}(W_1)$, что соответствует весам крайнего левого нейрона, повернутым на 30° . Анало-

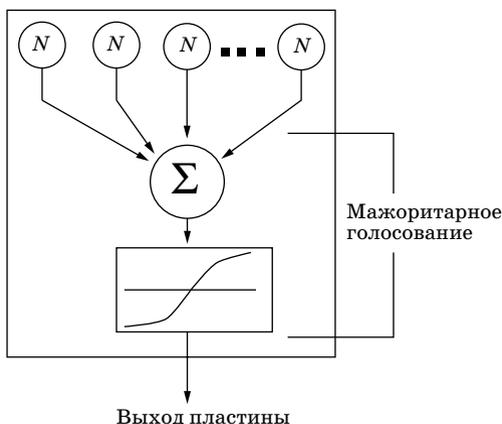


Рис. 93. Пластина инвариантной к повороту сети

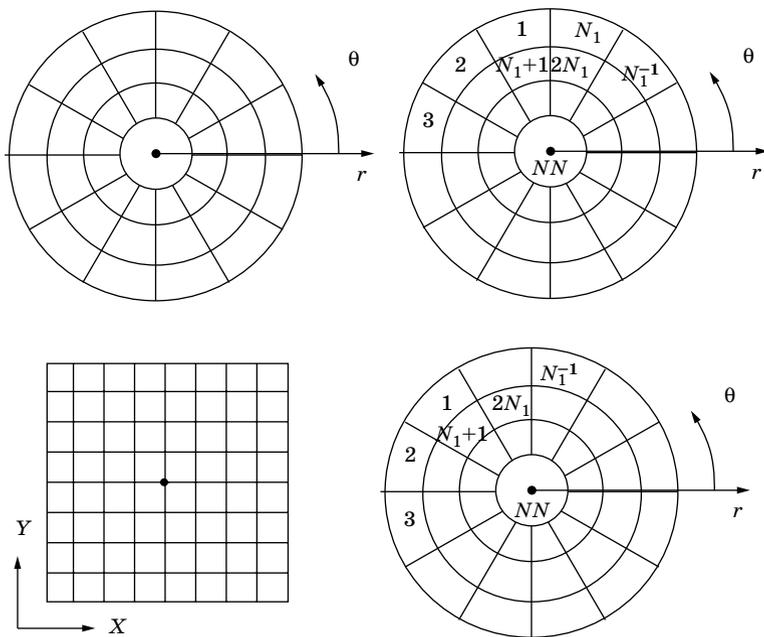


Рис. 94. Структура матрицы весов для нейронов пластины:
 а – круговой массив; б – квадратный массив; в – расположение весов (W_1)
 (номера означают порядок весов); г – расположение весов $R_{30}(W_1)$

гично получим $R_{60}(W_1), R_{90}(W_1) \dots R_{330}(W_1)$. В результате получим 12 наборов весов для нейронов одной пластины:

$$W_1, R_{30}(W_1), R_{60}(W_1), R_{90}(W_1) \dots R_{330}(W_1).$$

Помещение образа на ретину вызывает немедленную реакцию от выходного мажоритарного логического элемента (см. рис. 93). Понятно, что эта реакция не изменится при повороте образа на ретине. Поворот образа на 30° вызывает перестановку ролей нейронов в формировании реакций, но поскольку все они взвешены одинаково и суммируются выходным мажоритарным логическим элементом, то выходная реакция не меняется при повороте на 30° . Если необходима инвариантность к повороту на 15° , то пластине на рис. 86 потребуется 24 нейрона. Инвариантность к повороту на малые угловые приращения достигается при увеличении числа нейронов в пластине.

Для весов второй пластины потребуется такая же поворотная симметрия, но другой случайно выбранный набор весов (W_2). По-

этому функция отображения, получаемая с помощью второй пластины, будет отличаться от функции отображения первой пластины. Веса $\mathbf{W}_1, \mathbf{W}_2 \dots \mathbf{W}_L$ (L – число пластин) выбираются случайным образом. Единственным требованием является однозначное соответствие входного образа выходу преобразовывающей сети.

Пластину преобразовывающей сети можно обучить таким образом, чтобы ее выход для одного образа отличался от выхода для другого образа. Это обучение пластины позволяет достичь лучшей точности распознавания, хотя на это уходит много времени.

4.8. Байесовский метод распознавания

В процессе регистрации объекта и измерения его характерных признаков получают множество чисел, которые составляют вектор наблюдения. Будем считать, что этот вектор наблюдений x представляет собой случайный вектор с условной плотностью вероятности, зависящей от принадлежности этого вектора определенному классу. При распознавании объектов задачу формально сводят к проверке многих гипотез H_1, H_2, \dots, H_k , где H_i – гипотеза, предполагающая принадлежность объекта классу C_i . Здесь принято считать, что априорные распределения вероятностей этих гипотез заданы, т. е. известно, с какой вероятностью $P(H_i)$ объект может принадлежать классу C_i (или как часто появляется объект данного

класса). Причем $\sum_{i=1}^k P(H_i) = 1$, поскольку объект должен принадлежать какому-либо классу.

Процесс принятия решений в распознавании объектов можно рассматривать как игру статистического характера, которую классификационный механизм системы распознавания ведет с природой. При каждой реализации игры природа выбирает стратегию (в виде состояний природы, соответствующих образам или классам объектов), обозначаемую через $\{C_i, P(H_i)\}$. Стратегии игры, применяемые алгоритмом классификации, представляют собой решения, относящиеся к состояниям природы. Каждой паре действий, предпринятой игроками «природа – классификатор» ставится в соответствие некоторая функция потерь (или выигрыша). Считается, что число решений соответствует числу состояний природы (числу классов).

При каждой реализации игры природа выбирает стратегию (класс) C_i в соответствии с вероятностью $P(H_i)$. В результате хода игры, реализованного природой, появляется выборочный образ (объ-

ект) x . Классификатору неизвестно, какой именно класс предпочла природа. Вся информация, имеющаяся в его распоряжении, ограничивается самим вектором признаков объекта. Задача классифицирующего механизма – определить, опираясь на эту информацию, к какому классу принадлежит объект x . Ход игры классификатора, следовательно, представляет собой некоторое решение, определяющее класс C_j , который, «по мнению» классификатора, выбрала природа.

Игры рассматриваемого типа часто называют статистическими. Здесь природа не является «разумным противником», который способен сознательно выбирать свои стратегии таким образом, чтобы добиться максимизации потерь классификатора. Кроме того, у классификатора существует возможность «подсматривать» за игрой природы: он может осуществлять эксперименты и регистрировать обучающее множество объектов, которое затем используется при построении стратегии своей игры.

Пусть при реализации игры между природой и классификатором природа выбирает класс C_i (стратегию игры) и предъявляет объект x . Вероятность принадлежности объекта x классу C_i обозначим как $P(H_i|x)$. Если классификатор принимает решение о том, что объект x принадлежит классу C_j , когда на самом деле он принадлежит классу C_i , то классификатор несет потери, равные $L_{i,j}$. Поскольку объект x может принадлежать любому из k рассматриваемых классов, то математическое ожидание потерь, связанных с отнесением наблюдаемого объекта к классу C_j , определяется следующим выражением:

$$\pi_j(x) = \sum_{i=1}^k L_{i,j} p(H_i | x).$$

В теории статистических решений эту величину часто называют условным средним риском или условными средними потерями.

При распознавании каждого объекта, предъявляемого природой, классификатор может отнести его к одному из k возможных образов. Если для каждого объекта x вычисляются значения условных средних потерь $\pi_1(x)$, $\pi_2(x)$... $\pi_k(x)$ и классификатор причисляет объект к классу, которому соответствуют наименьшие условные потери, то очевидно, что и математическое ожидание полных потерь на множестве всех решений также будет минимизировано. Классификатор, минимизирующий математическое ожидание общих потерь, называется байесовским [33]. Со статистической точки зрения байесовский классификатор соответствует оптимальному качеству классификации.

Пусть $P(H_i | x)$ есть плотность распределения элементов вектора x при условии, что он принадлежит классу C_i . Хорошо известно, что вероятность принадлежности x классу C_i определяется формулой Байеса:

$$p_i = p(H_i | x) = \frac{P(H_i)p(x | H_i)}{p(x)},$$

так как безусловная плотность распределения

$$p(x) = \sum_{i=1}^k P(H_i)p(x | H_i).$$

Поскольку выражение $1/p(x)$ входит во все формулы вычисления средних потерь

$$\pi_j(x) = \frac{1}{p(x)} \sum_{i=1}^k L_{i,j} P(H_i) p(x | H_i), \quad j = 1, 2, \dots, k$$

в качестве общего множителя, его можно устранить из данного соотношения. В таком случае выражение для средних потерь сводится к следующему:

$$\pi_j(x) = \sum_{i=1}^k L_{i,j} P(H_i) p(x | H_i), \quad j = 1, 2, \dots, k.$$

При $k = 2$ и выборе классификатора стратегии (гипотезы) H_1 средние его потери для предъявленного природой объекта x равны

$$\pi_1(x) = L_{1,1} P(H_1) p(x | H_1) + L_{2,1} P(H_2) p(x | H_2),$$

а при выборе стратегии (гипотезы) H_2 :

$$\pi_2(x) = L_{1,2} P(H_1) p(x | H_1) + L_{2,2} P(H_2) p(x | H_2).$$

Как мы знаем, байесовский классификатор обеспечивает отнесение объекта x к классу с наименьшим значением средних потерь $\pi(x)$. Поэтому объект x причисляется к классу C_1 , если выполняется условие $\pi_1(x) < \pi_2(x)$; это должно означать, что

$$\begin{aligned} & L_{1,1} P(H_1) p(x | H_1) + L_{2,1} P(H_2) p(x | H_2) < \\ & < L_{1,2} P(H_1) p(x | H_1) + L_{2,2} P(H_2) p(x | H_2) \end{aligned}$$

или

$$(L_{2,1} - L_{2,2}) p(x | H_2) P(H_2) < (L_{1,2} - L_{1,1}) L_{1,1} p(x | H_1) P(H_1).$$

Принято считать, что потери от ошибочно принятого решения выше «потерь» при правильном выборе. Этому соответствуют неравенства $L_{i,j} > L_{i,i}$. Тогда байесовское решающее правило принимает следующий вид:

$$x \in C_1, \text{ если } \frac{P(H_1)p(x|H_1)}{P(H_2)p(x|H_2)} > \frac{L_{2,1} - L_{2,2}}{L_{1,2} - L_{1,1}}$$

или

$$\frac{p(x|H_1)}{p(x|H_2)} > \frac{P(H_2)(L_{2,1} - L_{2,2})}{P(H_1)(L_{1,2} - L_{1,1})}.$$

Величину $\frac{p(x|H_1)}{p(x|H_2)}$ называют отношением правдоподобия и обозначают через $\Lambda(x)$. Поскольку $\Lambda(x)$ представляет собой отношение двух функций случайной величины, то и само является случайной величиной. Величина

$$\eta = \frac{P(H_2)(L_{2,1} - L_{2,2})}{P(H_1)(L_{1,2} - L_{1,1})}$$

является пороговым значением критерия отношения правдоподобия, к которому в итоге свелось байесовское решающее правило:

$$x \in C_1, \text{ если } \Lambda(x) > \eta.$$

Отсюда видно, что вся процедура принятия решения сводится к вычислению отношения правдоподобия (зависящего лишь от вектора признаков и параметров распределений классов), и распределение априорных вероятностей или величины потерь на данное соотношение $\Lambda(x)$ влияния не оказывает. Указанная инвариантность процедуры обработки информации имеет большое практическое значение. Часто величины потерь и априорные вероятности являются квалифицированными предположениями на основе предыдущего опыта (интуиции). Последнее неравенство позволяет построить решающее правило, рассматривая η как переменный порог, учитывающий изменения в оценках априорных вероятностей и потерь в процессе накопления опыта.

Хорошо известно, что проведение статистического анализа и классификации многомерных наблюдений (признаков природных объектов) зачастую невозможно ограничить применением некоторых стандартных методов. Необходим детальный анализ структу-

ры наблюдаемой совокупности данных, чтобы путем углубленного исследования представленного числового материала выявить скрытые в нем закономерности, его вероятностную и геометрическую природу. Такой предмодельный (разведочный) анализ данных может оказать решающую помощь в компактном и понятном описании структуры наблюдений. Отталкиваясь от него, можно «осознанно» поставить вопрос о направлении более детального исследования данных с помощью того или иного метода, а также, возможно, сделать некоторые заключения о причинности модели данных. Если мы хотим использовать двумерное отображение входных данных для понимания внутренней структуры полученной информации с целью последующей классификации, то должны выбрать преобразование (многомерных) данных, сохраняющее разделимость классов. Мы видели, что в случае двух классов отношение правдоподобия несет полную информацию о разделимости классов в байесовском смысле. Следовательно, две плотности вероятностей или монотонные функции от них (отрицательные логарифмы, например) являются подходящей парой переменных. Отображение с использованием этих переменных показано на рис. 95. В этом пространстве байесовская граница представляет прямую, проходящую под углом 45° , независимо от вида распределений.

Отображение на рис. 95 не приводит к потере информации, необходимой для классификации. Единственная сложность здесь – это сложность вычисления функции $-\ln p(x|H_i)$. Если плотности вероятностей задаются набором параметров, решение этой задачи связано с оцениванием параметров. Например, если известно, что плотности вероятностей нормальны, функция $-\ln p(x|H_i)$ примет вид

$$-\ln p(x|H_i) = \frac{1}{2} \left(x - \hat{\mu}_i^T \right) \hat{\Sigma}_i^{-1} \left(x - \hat{\mu}_i \right) + \frac{1}{2} \ln \left\{ (2\pi)^m \left| \hat{\Sigma}_i \right| \right\},$$

где $\hat{\mu}_i$ и $\hat{\Sigma}_i$ – оценки математического ожидания и ковариационной матрицы.

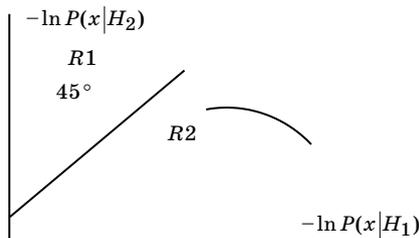


Рис. 95. Двумерное отображение данных

Пример 1. Пусть при гипотезе H_1 наблюдаемый фрагмент изображения соответствует постоянному «фону» с яркостью $b_1 > 0$, а по гипотезе H_2 фрагмент изображения соответствует «объекту» с постоянной яркостью $b_2 > b_1$ (рис. 96) [34].

Наблюдаемый яркостный сигнал подвержен шумовым искажениям. Будем считать, что фрагмент однородный по яркости и содержит N отсчетов. Результаты наблюдений представляют ряд из N независимых гауссовых величин x_1, x_2, \dots, x_N с известным средним значением: либо b_1 при гипотезе H_1 , либо b_2 при гипотезе H_2 .

Вследствие статистической независимости нетрудно записать совместные плотности вероятности величин $x_j, x_j = 1, 2, \dots, N$ для каждой из гипотез:

$$P(x | H_i) = \prod_{j=1}^N \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{(x_j - b_i)^2}{2\sigma^2}\right),$$

где σ^2 – известная дисперсия шума. Критерий отношения правдоподобия в этом случае имеет простой вид:

$$\Lambda(x) = \frac{\prod_{j=1}^N \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{(x_j - b_1)^2}{2\sigma^2}\right)}{\prod_{j=1}^N \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{(x_j - b_2)^2}{2\sigma^2}\right)}.$$



Рис. 96. Изображение сцены в инфракрасном диапазоне: рамками отмечены фоновый фрагмент (вверху) и фрагмент, соответствующий изображению объекта (яркая область повышенной теплоотдачи работающего двигателя)

После приведения подобных членов и взятия логарифма получим

$$\ln \Lambda(x) = \frac{b_2 - b_1}{\sigma^2} \sum_{j=1}^N x_j + \frac{N(b_2 - b_1)^2}{2\sigma^2}$$

и критерий Байеса запишется в виде $x \in C_1$, если

$$\frac{b_2 - b_1}{\sigma^2} \sum_{j=1}^N x_j + \frac{N(b_2 - b_1)^2}{2\sigma^2} > \ln \eta$$

или в эквивалентной форме

$$\sum_{j=1}^N x_j < \frac{\sigma^2}{b_2 - b_1} \ln \eta + \frac{N(b_2 + b_1)}{2}.$$

Нетрудно видеть, что процедура классификации сводится просто к суммированию результатов наблюдения яркости на распознаваемом фрагменте и сравнению суммы с порогом

$$\gamma = \frac{\sigma^2}{b_2 - b_1} \ln \eta + \frac{N(b_2 + b_1)}{2}.$$

Поскольку реализация байесовского классификатора предполагает знание плотности распределения для каждого класса, то становится совершенно очевидным, что оценка плотностей – основная проблема такой схемы классификации. С точки зрения статистического анализа при выборе модели распределений вполне обоснованным представляется принцип: среди множества моделей следует использовать модель, которая позволяет делать максимально надежные выводы о лежащей в основе структуры данных статистики (функции данных). Общим выражением этого положения является принцип максимальной энтропии, который гласит: если мы делаем выводы по неполной информации, то должны опираться на такое распределение вероятностей, которое имеет максимальную энтропию, допускаемую нашей априорной информацией. При этом признается, что модельное распределение с более высокой энтропией в некотором смысле «предпочтительнее» модели распределения с малой энтропией (позволяет извлечь больше информации из наблюдаемых данных).

Пример 2. Рассмотрим модель двух многомерных нормальных совокупностей с равными ковариационными матрицами Σ : $N(\mu_1, \Sigma)$ и $N(\mu_2, \Sigma)$, где $\bar{\mu}_1$ и $\bar{\mu}_2$ – вектора средних значений классов. Соответствующие (предполагающиеся заданными) ковариационные матрицы имеют вид

$$\Sigma = \left\{ \sigma_{i,j} \right\}_{i,j=1}^m,$$

где $\sigma_{i,j}$ – ковариация i -й и j -й компонент вектора признаков \mathbf{x} ; m – оценка размерности пространства.

Поскольку в случае нормального распределения имеем

$$p(\mathbf{x} | H_i) = (2\pi)^{-m/2} |\Sigma|^{-1/2} \exp \left\{ -\frac{1}{2} (\mathbf{x} - \bar{\mu}_i)^T \Sigma^{-1} (\mathbf{x} - \bar{\mu}_i) \right\}$$

($\bar{\mu}_i$ – известный вектор математического ожидания \mathbf{x} при гипотезе H_i), то отношение двух плотностей для байесовского правила определяется выражением

$$\Lambda(\mathbf{x}) = \frac{p(\mathbf{x} | H_1)}{p(\mathbf{x} | H_2)} = \exp \left\{ -\frac{1}{2} \left[(\mathbf{x} - \bar{\mu}_1)^T \Sigma^{-1} (\mathbf{x} - \bar{\mu}_1) - (\mathbf{x} - \bar{\mu}_2)^T \Sigma^{-1} (\mathbf{x} - \bar{\mu}_2) \right] \right\}.$$

Область R_1 , при попадании в которую наблюдение классифицируется как принадлежащее первому классу, является множество векторов \mathbf{x} , для которых величина $\Lambda(\mathbf{x})$ больше некоторой константы η , выбираемой подходящим образом. В частности, при выборе двоичной функции потерь ($L_{1,1} = L_{2,2} = 0$ – правильное решение, $L_{1,2} = L_{2,1} = 1$ – ошибка) условие, определяющее принадлежность образа \mathbf{x} классу C_1 принимает вид

$$\ln \Lambda(\mathbf{x}) = \ln \left(\frac{p(\mathbf{x} | H_1)}{p(\mathbf{x} | H_2)} \right) > \ln \eta,$$

где

$$\ln \eta = \ln \left(\frac{P(H_2)}{P(H_1)} \right),$$

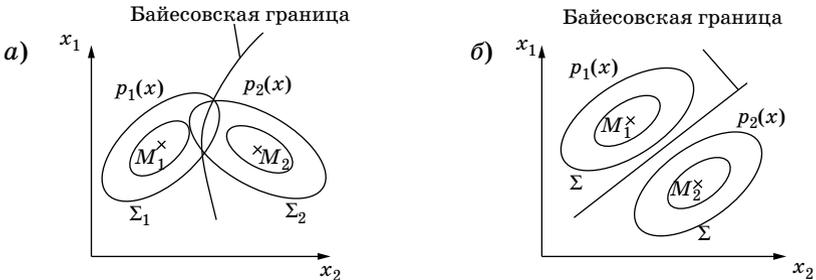


Рис. 97. Решающие границы для нормальных распределений:
а – $\Sigma_1 \neq \Sigma_2$, б – $\Sigma_1 = \Sigma_2$

если $\mathbf{x} \in C_1$ и известны априорные вероятности $\{P(H_1), P(H_2)\}$. Группируя соответствующие члены представления $\ln \Lambda$, получаем

$$\ln \Lambda = \mathbf{x}^t \Sigma^{-1} (\bar{\mu}_1 - \bar{\mu}_2) > \frac{1}{2} (\bar{\mu}_1 + \bar{\mu}_2)^t \Sigma^{-1} (\bar{\mu}_1 - \bar{\mu}_2) - \ln \left(\frac{P(H_1)}{P(H_2)} \right) = \eta_1.$$

Рассмотрим случай, когда два многомерных нормальных распределения имеют разные ковариационные матрицы. Решающее правило приобретает вид $\mathbf{x} \in C_1$ если

$$d(\mathbf{x}) = (\mathbf{x} - \bar{\mu}_1^T)^t \Sigma_1^{-1} (\mathbf{x} - \bar{\mu}_1) - (\mathbf{x} - \bar{\mu}_2^T)^t \Sigma_2^{-1} (\mathbf{x} - \bar{\mu}_2) + \ln \frac{|\Sigma_1|}{|\Sigma_2|} \leq 2 \ln \frac{P(H_1)}{P(H_2)}.$$

Видно, что решающая граница является квадратичной формой относительно вектора признаков (рис. 97 а, б). Для таких ситуаций рекомендован простой подход к синтезу линейной разделяющей функции. Он состоит в замене каждой из ковариационных матриц их средним значением, т. е. $P(H_1)\Sigma_1 + P(H_2)\Sigma_2$.

4.9. Распознавание на основе вейвлет-преобразования

Вейвлеты являются математическим инструментом для иерархического представления функций и появились в восьмидесятых годах как альтернатива оконным преобразованиям Фурье для анализа сигналов. С позиций точного представления произвольных сигналов и функций преобразование Фурье имеет ряд недостатков, которые привели к появлению оконного преобразования Фурье и стимулировали развитие вейвлет-преобразования [35]:

- ограниченная информативность анализа нестационарных сигналов и практически полное отсутствие возможностей анализа их особенностей (сингулярностей), так как в частотной области происходит «размазывание» особенностей сигналов (разрывов, ступенек, пиков и т.п.) по всему частотному диапазону спектра;
- появление эффекта Гиббса на скачках функций, при усечении сигналов и при вырезке отрезков сигналов для локального детального анализа;
- гармонический характер базисных функций, определенных в интервале от $-\infty$ до $+\infty$.

Термин «вейвлет» (*wavelet*) в переводе с английского означает «маленькая (короткая) волна». Вейвлеты – это обобщенное назва-

ние семейств математических функций определенной формы, которые локальны во времени и по частоте, и в которых все функции получаются из одной базовой (порождающей) посредством ее сдвигов и растяжений по оси времени. Вейвлет-преобразования рассматривают анализируемые временные функции в терминах колебаний, локализованных по времени и частоте.

В компьютерной графике вейвлет-преобразование – это свертка исходного изображения с некоторой функцией Ψ – вейвлетом. Двумерный вейвлет Габора $\Psi_k(x)$ – это [39]:

$$\Psi_k(x) = \left(\frac{\|k\|^2}{\sigma^2} \right) \exp \left(-\frac{\|k\|^2 \|x\|^2}{2\sigma^2} \right) \left(\exp(jkx) - \exp \left(\frac{-\sigma^2}{2} \right) \right),$$

где $k = (\omega \cos \theta, \omega \sin \theta)$, j – мнимая единица, $x \in R^2$.

Обычно рассматривают вещественную (четную) компоненту $R_k(x) = \text{Re}(\Psi_k(x))$ (рис. 98, а) и мнимую $S_k(x) = \text{Im}(\Psi_k(x))$ (рис. 98, б) вейвлета Габора.

Вектор k определяет частоту и направление функций $R_k(x)$ и $S_k(x)$, ω отвечает частоте, а θ – повороту. Функции $R_k(x)$ и $S_k(x)$ при разных значениях ω и θ показаны на рис. 99.

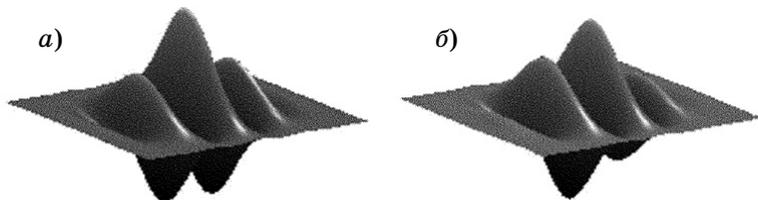


Рис. 98. Вещественная (а) и мнимая (б) компоненты вейвлета Габора

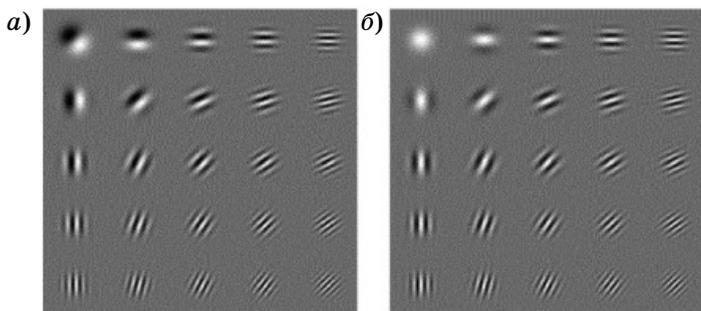


Рис. 99. Вещественная (а) и мнимая (б) компоненты вейвлета Габора при разных значениях ω и θ

Вейвлет-преобразование $T_k(x,y)$ изображения $I(x,y)$ представляется в виде:

$$T_k(x,y) = \iint I(x,y) \psi_k(x-u, y-v) du dv.$$

Рассмотрим функцию $R_k(x)$ при различных значениях ω и θ (рис. 100) и изображение $I(x,y)$ (рис. 101). Для каждого значения ω и θ вычислим $T_k(x,y)$ (рис. 102).

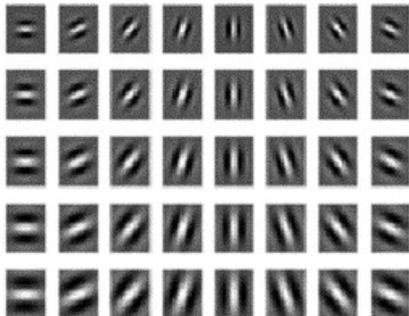


Рис. 100. Функция $R_k(x)$ при различных значениях ω и θ



Рис. 101. Изображение лица $I(x,y)$

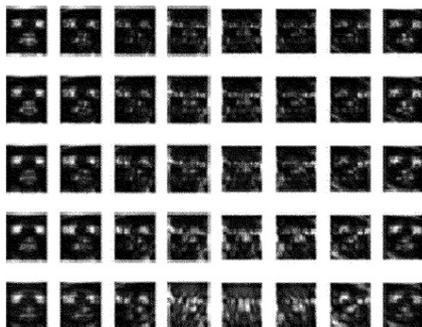


Рис. 102. Функции $T_k(x,y)$ при различных значениях ω и θ

С помощью полученного семейства изображений $T_k(x, y)$ можно найти на лице характерные точки. Точку (x_0, y_0) назовем характерной, если

$$T_k(x_0, y_0) = \max_{(x, y) \in P} (T_k(x, y))$$

и

$$T_k(x_0, y_0) > \frac{1}{WH} \sum_{x=1}^W \sum_{y=1}^H T_k(x_0, y_0), \quad k = 1 \dots 40,$$

где P – прямоугольное окно с центром в точке (x_0, y_0) ; W – ширина изображения; H – высота изображения. Первое уравнение говорит о том, что значение $T_k(x_0, y_0)$ максимально в P – окрестности (x_0, y_0) . Второе – о том, что это максимум не локальный. Здесь важен размер окна P . Если рассмотреть полученные характерные точки на исходном изображении $I(x, y)$ (размеры окна P 9×9 пикселей), то они сконцентрируются около глаз, носа, рта и т. д. (рис. 103) [35].

Для каждой характерной точки (x_k, y_k) каждого изображения i определим характерные вектора (вектора черт лица) $v_{i,k} = (x_k, y_k, T_{i,j}(k))$, $j = 1, 2 \dots 40$). Такой вектор состоит из координат особой точки и 40 вейвлет-коэффициентов, отвечающих этой точке. С помощью характерных векторов $v_{i,k}$ и $v_{j,r}$ определим степень схожести изображений i и j :

$$S_{k,r}(i, j) = \frac{\sum_{m=3}^{42} |v_{i,k}(m)| |v_{j,r}(m)|}{\sqrt{\sum_{m=3}^{42} |v_{i,k}(m)|^2 \sum_{m=3}^{42} |v_{j,r}(m)|^2}},$$

где $v_{i,k}(m)$ – m -я компонента вектора $v_{i,k}$.

Пусть даны база данных лиц и лицо X , которое требуется распознать. Алгоритм распознавания следующий: для каждого харак-



Рис. 103. Характерные точки изображения

терного вектора $v_{x,k}$ и лица i из базы данных составим множество N_i^k ($k = 1, 2 \dots 40$), состоящее из векторов $v_{i,r}$ со следующими свойствами:

$$- \sqrt{(x_k - x_r)^2 + (y_k - y_r)^2} < C_1,$$

где C_1 – радиус окружности, примерно содержащий область глаз, носа и т. д.;

$$- S_{k,r}(i, X) > C_2,$$

где C_2 – стандартное отклонение функции $S_{k,r}(i, X)$ по всем i из базы данных.

$$\text{Во всех непустых множествах } S_i^k = \max_{r \in N_i^k} (S_{k,r}(i, X)).$$

Тогда степень схожести изображений i и X можно характеризовать величиной $S_i = E(S_i^k)$ – средним значением S_i^k по всем k .

В этом алгоритме следует отметить следующие моменты:

– Выбор константы C_1 следует из условий съемки изображений, размера изображения в пикселях и т.д. C_1 подбирается эмпирически.

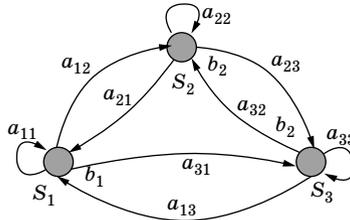
– Для двух разных изображений, даже одного человека, количество характерных точек может быть различным.

Кроме того, на практике получившиеся изображения требуют предварительной обработки. Основная проблема предварительной обработки состоит в определении местонахождения лица на изображении. Далее требуется изменить размер изображения лица до эталонного, после чего провести выравнивание гистограммы освещенности. При этом требуется не потерять качества изображения.

4.10. Скрытые марковские модели

Марковские модели являются мощным средством моделирования различных процессов и распознавания образов. По своей природе марковские модели позволяют учитывать непосредственно пространственно-временные характеристики сигналов, и поэтому получили широкое применение в распознавании речи, а в последнее время – изображений (в частности изображений лиц).

Каждая модель $\lambda = (A, B, \pi)$ (рис. 104) представляет собой набор N состояний $S = \{S_1, S_2, \dots, S_N\}$ (на рис. 104 – вершины графа), между которыми возможны переходы (на рис. 104 – дуги). В каждый момент времени система находится в строго определенном состоянии.



$O = \{Y, G, R, G, R, G, B, R, R, Y, B\}$
 $S = \{2, 1, 1, 3, 2, 2, 2, 3, 3, 1\}$

Рис. 104. Марковская модель, пример последовательности наблюдений O и последовательности состояний S

В наиболее распространённых марковских моделях первого порядка полагаются, что следующее состояние зависит только от текущего состояния.

При переходе в каждое состояние генерируется наблюдаемый символ, который соответствует физическому сигналу с выхода моделируемой системы. Набор символов для каждого состояния $\mathbf{V} = \{v_1, v_2, \dots, v_M\}$, количество символов M . Выход, генерируемый моделью, может быть также непрерывным. Существуют также модели, в которых набор символов для всех состояний одинаков. Символ в состоянии $q_t = S_j$ в момент времени t генерируется с вероятностью $b_j(k) = P[v_{k,t} | q_t = S_j]$. Набор всех таких вероятностей составляет матрицу $\mathbf{B} = \{b_j(k)\}$.

Матрица $\mathbf{A} = \{a_{ij}\}$ определяет вероятность перехода из одного состояния в другое состояние: $a_{ij} = P[q_{t+1} = S_j | q_t = S_i]$, $1 \leq i, j \leq N$. Считается, что A не зависит от времени. Если из каждого состояния можно достичь любого другого за один переход, то все $a_{ij} > 0$, и модель называется эргодической.

Также модель имеет вероятность начальных состояний $\pi = \pi_i$, где $\pi_i = P[q_1 = S_i]$.

Обычно в реальных процессах последовательность состояний является скрытой от наблюдения и остаётся неизвестной, а известен только выход системы, последовательность наблюдаемых символов $O = \{O_1, O_2, \dots, O_T\}$, где каждое наблюдение O_t – символ из \mathbf{V} , и T – число наблюдений в последовательности. Поэтому такие модели называют скрытыми марковскими моделями (СММ, по-английски – Hidden Markov Models, НММ).

Модель $\lambda = (\mathbf{A}, \mathbf{B}, \pi)$ с настроенными параметрами может быть использована для генерирования последовательности наблюдений. Для этого случайно, в соответствии с начальными вероятностями π выбирается начальное состояние, затем на каждом шаге вероятность \mathbf{B} используется для генерации наблюдаемого символа, а вероятность \mathbf{A} – для выбора следующего состояния. Вероятность \mathbf{P} генерирования моделью λ последовательности состояний

$$O: P(O|Q, \lambda) = \prod_{t=1}^T b_{q_t}(O_t), \text{ где } Q = q_1, q_2 \dots q_T \text{ – последовательность}$$

состояний. Предполагается, что наблюдения статистически независимы.

В распознавании образов скрытые марковские модели применяются следующим образом. Каждому классу i соответствует своя модель λ_i . Распознаваемый образ (речевой сигнал, изображение и т.д.) представляется в виде последовательности наблюдений O . Затем для каждой модели λ_i вычисляется вероятность того, что эта последовательность могла быть сгенерирована именно этой моделью. Модель λ_i , получившая наибольшую вероятность, считается наиболее подходящей, и образ относят к классу j .

В связи с этим появляются несколько вопросов, называемых тремя основными задачами скрытых марковских моделей:

– Имея последовательность наблюдений и настроенную модель, как оценить вероятность генерации этой моделью данной последовательности наблюдений? Эта задача называется задачей распознавания.

– Имея последовательность наблюдений и настроенную модель, как подобрать последовательность состояний, чтобы она была оптимальной в соответствии с некоторым критерием? Другими словами, это задача объяснения. Она нужна для последующей коррекции параметров модели.

– Каким образом корректировать параметры модели, для того чтобы максимизировать вероятность генерирования моделью последовательности наблюдаемых симсостояний; т. е. как сделать так, чтобы модель больше соответствовала своему классу, одним из образов которого является данная последовательность наблюдений (или несколько различных последовательностей)? Это задача обучения.

Первая задача имеет точное аналитическое решение, называемое процедурой прямого-обратного прохода. Последующие две задачи не имеют точного аналитического решения. Для решения второй задачи используется алгоритм Витерби, для третьей – ал-

горитм Баума-Вельча. Оба этих метода являются разновидностями градиентного спуска и решаются оптимизационными методами.

Для того чтобы сократить вычисления, в распознавании речи используются линейные модели (рис. 105). В таких моделях каждое состояние имеет только одно последующее, также переход возможен обратно в то же состояние. Такие модели учитывают временные характеристики речевого сигнала: определённый порядок следования участков сигнала, их взаимное расположение, возможность локальных растяжений или сжатий. Это позволяет их применять и в распознавании изображений.

Одна из первых работ, применяющая СММ для распознавания изображений лиц – это диссертация Фердинанда Самарии, которой предшествовали работы по распознаванию изображений других видов скрытыми марковскими моделями. В этой работе распознавание осуществлялось как простейшими одномерными линейными СММ, так и псевдодвумерными. Введение второго измерения позволило повысить точность распознавания с 85% до 95%.

Суть двумерных марковских моделей заключается в том, что в отличие от одномерных линейных СММ, они позволяют моделировать искажения изображения и взаимное расположение участков не отдельно по горизонтали или вертикали, а в обоих направлениях одновременно. Для уменьшения вычислительной сложности применяются псевдодвумерные СММ (Pseudo-2D Hidden Markov Models, P2D-НММ). Такая модель состоит из нескольких линейных вертикальных моделей нижнего уровня и одной линейной горизонтальной модели верхнего уровня, на вход которой поступают выходы моделей нижнего уровня (рис. 106). Каждое состояние модели верхнего уровня включает в себя последовательность состояний соответствующей модели нижнего уровня. Модели нижнего уровня не связаны между собой. Изначально модели верхнего уровня были вертикальными. Затем модели верхнего уровня были сделаны горизонтальными (как это и показано на рис. 106), чтобы вертикальные модели нижнего уровня могли учесть, что глаза могут находиться на разной высоте. Таким образом, псевдодвумерная модель позволяет учесть локальные деформации и взаимное расположение

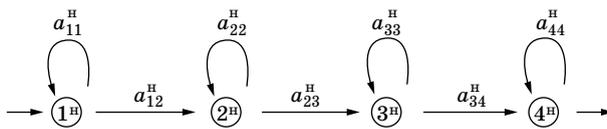


Рис. 105. Линейная Марковская модель

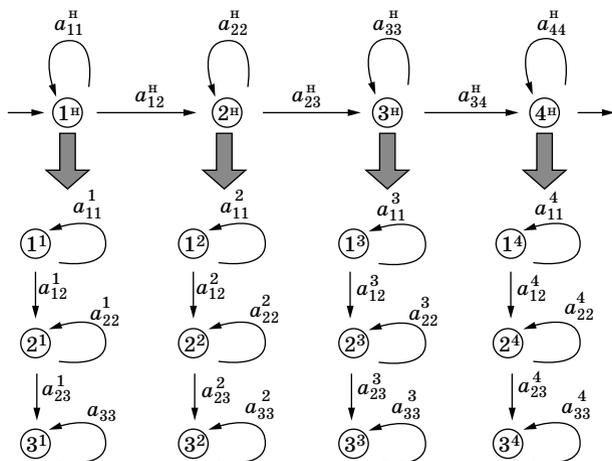


Рис. 106. Псевдодвумерная скрытая марковская модель

участков изображений. Но в отличие от оптических потоков и других методов сопоставления деформациями, псевдодвумерная модель учитывает характер деформаций, а то, какими именно могут быть возможные деформации, псевдодвумерные СММ усваивают в процессе обучения. Другими словами, участок, соответствующий глазу, никогда не будет сопоставлен например участку на месте рта, как это может быть в оптическом потоке.

Наблюдениям, подаваемым на вход СММ, являлись квадратные участки изображений (рис. 107). Было обнаружено, что участки, извлекаемые с 75% перекрытием друг друга, давали наилучшую точность распознавания.

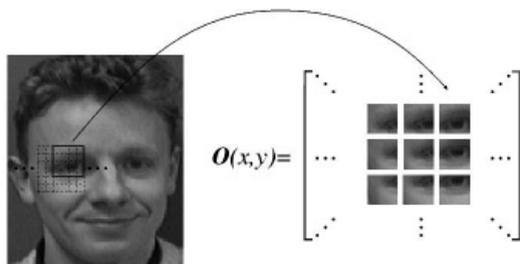


Рис. 107. Извлечение участков-образцов наблюдения

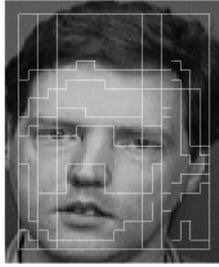


Рис. 108. Сегментация изображения. Линии отмечают области, соответствующие одинаковым состояниям

Одним из полезных свойств СММ является способность сегментировать распознаваемое изображение. Результат работы алгоритма Витерби, разбившего изображение на последовательность состояний, показан на рис. 108.

Впоследствии использовались дальнейшие улучшения способов начального представления изображения и алгоритмов тренировки.

Для каждого квадратного участка изображения 16×16 вычислялось двумерное дискретное косинусное преобразование, и этот участок представлялся в виде набора первых 15 коэффициентов. Это позволило повысить точность распознавания на 2%. Кроме того, такое представление позволяет более точно, чем при масштабировании, представлять изображение, используя меньший объём информации.

Для увеличения тренировочного набора использовались также зеркально отражённые по вертикали изображения. Это позволило учесть более широкий диапазон ракурсов.

Полезное свойство распознавания по коэффициентам дискретного косинусного преобразования заключается в том, что оно позволяет работать непосредственно со сжатыми изображениями, такими как JPEG и MPEG, которые на сегодняшний день являются распространёнными форматами хранения изображений и видео.

Как можно видеть, правильный выбор метода классификации (моделирования), начального представления изображения и учёт его особенностей в сочетании с применением априорных знаний о предметной области, дали успешный результат.

Недостатком СММ является то, что СММ не обладает различающей способностью. Таким образом, алгоритм обучения только максимизирует отклик каждой модели на свои классы, но не минимизирует отклик на другие классы, и не выделяются ключевые

признаки, отличающие один класс от другого. Например, для определения того, содержится ли лицо в обучающей выборке, используется алгоритм ранжирования вероятностей, заключающийся в следующем. На обучающем наборе каждая модель реагирует на изображения-примеры с некоторой вероятностью. Отсортированные таким образом модели образуют исходное ранжирование. Для неизвестного изображения модели также ранжируются по вероятностям отклика на неизвестное изображение. Большая величина отклонения полученного ранжирования от исходного сигнализирует о том, что изображение принадлежит лицу неизвестного класса.

Таким образом, похожие классы могут оказаться слабо различимыми, и при увеличении объёма базы или использования в более широких условиях СММ может оказаться ненадёжными. Много-слойные нейронные сети лишены такого недостатка.

Библиографический список

1. *Карасик В. Е., Орлов В. М.* Лазерные системы видения: учеб. пособие. М.: Изд-во МГТУ им. Баумана, 2001. 352 с.
2. *Борн М., Вольф Э.* Основы оптики: пер. с англ. / под ред. Г. П. Мотулевич. М.: Наука, 1970. 856 с.
3. *Ахманов С. А., Никитин С. Ю.* Физическая оптика: учебник. М.: Изд-во МГУ, 2004. 656 с.
4. *Иванова Т. В.* Введение в прикладную и компьютерную оптику: конспект лекций / СПб ГИТМО. СПб., 2002. 92 с.
5. *Прокопенко В. Т., Трофимов В. А., Шарок Л. П.* Психология зрительного восприятия: учеб. пособие / СПбГУИТМО. СПб., 2006. 73 с.
6. *Грегори Р. Л.* Глаз и мозг. М.: Прогресс, 1979. 269 с.
7. *Долин Л. С., Левин И. М.* Справочник по теории подводного видения. Л.: Гидрометеоздат, 1991. 229 с.
8. *Волков В. Г.* Приборы ночного видения для бронемашин // Специальная техника. 2004. № 5. С. 2–13.
9. *Гейхман И. Л., Волков В. Г.* Основы улучшения видимости в сложных условиях. М.: Недра. 1999. 286 с.
10. *Фисенко В. Т., Фисенко Т. Ю.* Компьютерная обработка и распознавание изображений: учеб. пособие / СПбГУИТМО. СПб., 2008. 192 с.
11. *Васильев В. П., Муру Э. Л., Смольский С. М.* Основы теории и расчета цифровых фильтров. М.: Академия, 2007. 272 с.
12. *Бисярин В. П., Соколов А. П., Сухонин Е. В.* Ослабление лазерного излучения в гидрометеорах. М.: Наука, 1977. 176 с.
13. *Мишура Т. П., Платонов О. Ю.* Проектирование лазерных систем: учеб. пособие / ГУАП. СПб., 2006. 98 с.
14. *Бакуменко В. Л., Свиридов А. Н., Таубкин И. И.* Анализ предельных возможностей идеальных тепловизоров при наблюдении «точечных» излучателей // Прикладная физика. 2002. № 1. С. 63–70.
15. *Коцавцев Н. Ф., Волков В. Г.* Приборы ночного видения // Вопросы оборонной техники. Сер. 11. 1993. Вып. 3(138). С. 16–23.
16. *Рохлин Г. Н.* Разрядные источники света. М.: Энергоатомиздат, 1991. 720 с.
17. *Бокшанский В. Б., Карасик В. Е.* Расчет характеристик фоточувствительных приборов с зарядовой связью: учеб. пособие. М.: Изд-во МГТУ им. Баумана, 2001. 43 с.
18. Прэтт У. Цифровая обработка изображений: в 2-х кн. Кн. 1: пер. с англ. М.: Мир, 1982. 312 с.
19. *Гонсалес Р, Вудс Р.* Цифровая обработка изображений. М.: Техносфера, 2005. 1070 с.

20. *Лукин А.* Введение в цифровую обработку сигналов (Математические основы) М.: Изд-во МГУ, 2002. 44 с.
21. *Сойфер В. А.* Компьютерная обработка изображений: Ч. 2. Методы и алгоритмы // Соросовский образовательный журнал. 1996. № 3. С. 110–121.
22. *Гмурман В. Е.* Теория вероятностей и математическая статистика. М.: Высшая школа, 2003. 479 с.
23. *Мудров В. И., Кушко В. Л.* Методы обработки измерений. Квазиправдоподобные оценки. М.: Радио и связь, 1983. 304 с.
24. *Анисимов И.* Технология HDR1 в видеонаблюдении // Алгоритм безопасности. № 3. 2010. С. 40–41.
25. *Передерий В. А.* Глазные болезни. Полный справочник. М.: Эксмо, 2008. 704 с.
26. *Магурин В. Г., Тарлыков В. А.* Когерентная оптика: учеб. пособие по курсу «Когерентная и нелинейная оптика» / СПбГУ ИТМО. СПб., 2006. 122 с.
27. *Авласенок А. В., Алексеев Е. Г., Литвинов С. П.* Экспериментальные исследования возможности распознавания воздушных целей человеком-оператором по инфракрасным изображениям, формируемым пассивными ОЭС // Современные проблемы проектирования, производства и эксплуатации радиотехнических систем: сб. науч. тр. Ульяновск. 2008. С. 35–42.
28. *Журавлёв Ю. И.* Распознавание образов и анализ изображений // Ю. И. Журавлёв, И. Б. Гуревич. Искусственный интеллект: Модели и методы. Т. 1. М.: Радио и связь, 2000. 310 с.
29. *Свирин И., Ханин А.* Некоторые аспекты автоматического распознавания автомобильных номеров // Алгоритм безопасности. 2010. №3. С. 27–29.
30. Тестирование систем распознавания автомобильных номеров// <http://www.proscvtv.ru/jurnal/jurnal.html>. Электронная версия журнала PROSystem CCTV. 2009. № 6.
31. Лазер как альтернатива видеонаблюдению // http://www.secnews.ru/articles/10148.htm?sphrase_id=33946. Электронная версия газеты Security News, 21.09.2007.
32. *Андреев А. Л.* Автоматизированные телевизионные системы наблюдения. Ч. II. Арифметико-логические основы и алгоритмы: учеб. пособие / СПбГУИТМО. СПб., 2005. 88 с.
33. *М. Н. Yang, D. J. Kriegman, N. Ahuja.* Detecting faces in images // IEEE Trans. Pattern Analysis and Machine Intelligence. 2002. Vol. 24. No. 1. Pp. 34–58.

34. *Грузман И. С., Киричук В. С.* и др. Цифровая обработка изображений в информационных системах: учеб. пособие / Новосибирск: Изд-во НГТУ, 2002. 352 с.

35. *Qing Chen, Emil Petriu, Xiaoli Yang.* A Comparative Study of Fourier Descriptors and Hu's Seven Moment Invariants for Image Recognition // CCECE 2004. Pp. 0103–0106.

36. *Столниц Э., ДеРоуз Т., Салезин Д.* Вейвлеты в компьютерной графике. М.: Регулярная и хаотическая динамика, 2002. 272 с.

37. *Волченков М. П., Самоненко И. Ю.* Об автоматическом распознавании лиц // Интеллектуальные системы. 2005. Т. 9. Вып. 1–4. С. 135–156.

38. *Thawar Arif, Zyad Shaaban, Lala Krekor, Sami Baba.* Object classification via geometrical, zernike and legendre moments // Journal of Theoretical and Applied Information Technology, 2009. Vol. 7. No. 1. Pp. 31–37.

39. *Томашевич Н. С., Томашевич Д. С., Галушкин А. И.* Методы реализации инвариантности к аффинным преобразованиям при распознавании двумерных изображений // Информационные технологии. 2001. Прил. к № 1. С. 1–18.

40. *Cruz V., Criclobal G., Michaux T. and Barquin S.* Invariant image recognition using a multi-network neural model // Electronic Neurocomputers, Proc. Int. Joint Conf. Neural Networks. 1989. Vol. 2. Pp. 17–21.

41. *Pedro J. Zufiria, Javier Munoz.* Extended backpropagation for invariant pattern recognition neural networks // IJCNN'93 – Nagoya. Japan. 1993. October. Vol. 3. Pp. 2097–2100.

Учебное издание

Рыжиков Максим Борисович

**ФОРМИРОВАНИЕ И ОБРАБОТКА
ИЗОБРАЖЕНИЙ
В ЛАЗЕРНЫХ СИСТЕМАХ ВИДЕНИЯ**

Учебное пособие

Редактор *В. П. Зуева*
Верстальщик *С. Б. Мацанура*

Сдано в набор 02.09.13. Подписано к печати 25.12.13.
Формат 60×84 1/16. Бумага офсетная. Усл. печ. л. 12,0.
Уч.-изд. л. 12,9. Тираж 100 экз. Заказ № 670.

Редакционно-издательский центр ГУАП
190000, Санкт-Петербург, Б. Морская ул., 67