

МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ РОССИЙСКОЙ ФЕДЕРАЦИИ

ФЕДЕРАЛЬНОЕ АГЕНТСТВО ПО ОБРАЗОВАНИЮ

САНКТ - ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ  
ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ, МЕХАНИКИ И ОПТИКИ



ПОБЕДИТЕЛЬ КОНКУРСА ИННОВАЦИОННЫХ ОБРАЗОВАТЕЛЬНЫХ ПРОГРАММ ВУЗОВ

В.Т. Фисенко, Т.Ю. Фисенко

**КОМПЬЮТЕРНАЯ ОБРАБОТКА  
И РАСПОЗНАВАНИЕ ИЗОБРАЖЕНИЙ**  
**учебное пособие**



Санкт-Петербург

2008

В.Т. Фисенко, Т.Ю. Фисенко, Компьютерная обработка и распознавание изображений: учеб. пособие. - СПб: СПбГУ ИТМО, 2008. – 192 с.

В пособии приведены методы компьютерной обработки и распознавания изображений. Рассмотрены математические модели изображений, критерии качества изображений. Описываются основные алгоритмы цифровой обработки и распознавания изображений, в том числе основы яркостных преобразований, преобразования цветовых координатных пространств, пространственной и частотной фильтрации, морфологических операций, кодирования, сегментации и классификации, а также анализа изображений.

Предназначено для студентов, обучающихся по направлению подготовки 200600 – «Фотоника и оптоинформатика».

Рекомендовано к печати УМО по образованию в области приборостроения и оптотехники в качестве учебного пособия для студентов высших учебных заведений, обучающихся по направлению подготовки 200600 – «Фотоника и оптоинформатика».



В 2007 году СПбГУ ИТМО стал победителем конкурса инновационных образовательных программ вузов России на 2007–2008 годы. Реализация инновационной образовательной программы «Инновационная система подготовки специалистов нового поколения в области информационных и оптических технологий» позволит выйти на качественно новый уровень подготовки выпускников и удовлетворить возрастающий спрос на специалистов в информационной, оптической и других высокотехнологичных отраслях экономики.

©Санкт-Петербургский государственный университет информационных технологий, механики и оптики, 2008

©Фисенко В.Т., Фисенко Т.Ю., 2008

## СПИСОК ИСПОЛЬЗОВАННЫХ СОКРАЩЕНИЙ

АЦП -	аналого- цифровой преобразователь,
АФМ -	адаптивная медианная фильтрация,
ВО -	видеообъект,
ВП -	вейвлетное преобразование,
ВЧ -	высокочастотный,
ДИКМ -	дифференциальная импульсно-кодовая модуляция,
ДВП -	прямое дискретное вейвлетное преобразование,
ДКП -	прямое дискретное косинусное преобразование,
ДПФ -	прямое дискретное преобразование Фурье,
ЗС -	зрительная система,
ИК -	инфракрасный,
КИХ -	конечная импульсная характеристика,
МККР -	Международный консультативный комитет по радиосвязи ( <i>CCIR</i> )
МКО -	Международная комиссия по освещению ( <i>CIE</i> ),
МСП -	марковские случайные поля,
МСЭ -	Международный союз электросвязи,
МСЭ-Р -	комиссия по телевидению и радиовещанию МСЭ,
МФ -	медианный фильтр,
НЧ -	низкочастотный,
ОДВП -	обратное дискретное вейвлетное преобразование,
ОДКП -	обратное дискретное косинусное преобразование,
ОДПФ -	обратное дискретное преобразование Фурье,
ПЗС -	приборы с зарядовой связью,
ПК -	персональный компьютер,
ПСШ -	пиковое отношение сигнал/шум,
СКО -	среднеквадратическое отклонение,
ТВ -	телевидение,
ТВЧ -	телевидение высокой четкости,
ЦАП -	цифро - аналоговый преобразователь,
ЦТВ -	цифровое телевидение,
ЭДС -	электродвижущая сила,
ЭЛТ -	электронно-лучевая трубка,
ЭОП -	электронно-оптический преобразователь,
В кадр -	кадр, кодируемый с предсказанием в двух направлениях ( <i>Bidirectionally predictive coded picture</i> ),
<i>CCIR</i> -	МККР ( <i>International Radiocommunication Consultative Committee</i> ),
<i>CIE</i> -	МКО ( <i>Comission Internationale de l'Eclairage</i> ),
<i>FIFO</i> -	первым вошел, первым вышел,
<i>HDTV</i> -	телевидение высокой четкости,
<i>HIS</i> -	цветовое координатное пространство ( <i>hue</i> - тон, <i>saturation</i> - насыщенность, <i>intencity</i> - яркость),

<i>I</i> кадр -	независимо кодируемый кадр ( <i>Intra coded picture</i> ),
<i>IEC</i> -	Международный электротехнический комитет,
<i>ISO</i> -	Международная организация по стандартизации ( <i>International Organization for Standardization</i> ),
<i>ITU-R</i> -	Международный союз телекоммуникаций ( <i>International Telecommunication Union - Radio</i> )
<i>JND</i> -	порог "минимальной заметной разницы",
<i>JPEG</i> -	объединенная группа экспертов по фотографии ( <i>Joint Photographic Experts Group</i> ),
<i>LUT</i> -	просмотровая таблица ( <i>Look up Table</i> ),
<i>LZW</i> -	метод кодирования Лемпеля, Зива и Вельча ( <i>Lempel, Ziv, Welch</i> ),
<i>M-JPEG</i> -	стандарт внутрикадрового кодирования видеопоследовательностей ( <i>Motion JPEG</i> ),
<i>MPEG</i> -	Международная экспертная группа по обработке последовательностей изображений ( <i>Moving Picture Expert Group</i> ), набор инструментов для сжатия видео последовательностей,
<i>NTSC</i> -	аналоговая система телевизионного вещания (США, Канада, Япония и др.),
<i>P</i> -кадр -	кадр, кодируемый с предсказанием ( <i>Predictive coded picture</i> ),
<i>PDL</i> -	язык описания изображений,
<i>PAL</i> -	аналоговая система телевизионного вещания,
<i>RGB</i> -	красный ( <i>R</i> ), зеленый ( <i>G</i> ) и синий ( <i>B</i> ),
<i>RLE</i> -	кодирование длин серий ( <i>Run-Length Encoding</i> ),
<i>SECAM</i> -	аналоговая система телевизионного вещания (СНГ, Франция и др.),
<i>2D</i> -	двумерный,
<i>3D</i> -	трехмерный.

## **ПРЕДИСЛОВИЕ**

Цифровая обработка и распознавание изображений - одно из интенсивно развивающихся направлений исследования. Главная цель этого пособия - раскрыть предмет, не усложняя его сложными математическими преобразованиями. Имея большой опыт работы в области цифровой обработки изображений (более 34 лет), мы снабдили пособие большим числом иллюстраций, позволяющих продемонстрировать особенности реализации компьютерных методов обработки изображений и их анализа и синтеза. Считаем, что это позволит сделать изложение материала понятным и доступным как студентам, так и практикующим инженерам. Большое внимание уделено проблемам дискретизации и квантования сигналов изображений, поскольку эти вопросы важны при построении систем оптико - электронного преобразования сигналов. Все алгоритмы обработки описаны так, что они могут быть реализованы на компьютере. В пособии отсутствует описание конкретных пакетов программ, в которых можно выполнить эти алгоритмы. Выполнение может осуществляться, например, в системе *MatLab*, доступной студентам и специалистам-практикам. Это сделано, чтобы не загромождать представление особенностями, связанными с реализацией алгоритмов.

В первом разделе представлены предмет и задачи курса, а также описаны основные устройства формирования изображений.

Построение системы оптико - электронного преобразования базируется на модели сигнала; математические модели изображений описаны во втором разделе.

Наряду с широким применением датчиков черно-белого изображения, сегодня широко применяются цветные датчики изображения. Оценке цветовых характеристик изображений и неразрывно связанных с ними преобразованиям цветовых координатных пространств посвящен третий раздел.

Дискретизация и квантование сигналов позволяют представить изображение в цифровой форме в виде матрицы целых чисел. Процесс восстановления непрерывной двумерной функции по дискретным значениям яркости элементов изображения осуществляется интерполяционными методами. Этим важным проблемам преобразования сигналов посвящен четвертый раздел.

В компьютерных системах, когда получателем информации является человек, большое значение имеют методы улучшения изображений, позволяющие повысить заметность интересующих деталей на изображении. Кроме того, при предварительной обработке изображений, выполняемой в автоматических компьютерных системах, также важную роль играет предварительная обработка изображений, позволяющая сформировать пространство признаков объектов. Эти методы описаны в 5 разделе.

Изображения, полученные на выходе оптико - электронных преобразователей, искажены помехами. При анализе объектов на сложном фоне, фон тоже является помехой. Ослабление действия помех достигается фильтрацией. В зависимости от приложения фильтрация производится в пространственной или частотной области. Основам фильтрации посвящен 6 раздел.

Формированию и обработке бинарных изображений посвящен 7 раздел. При этом особое место отводится морфологии бинарных изображений, предназначеннной для представления и описания свойств формы и структуры объектов. Введены основные признаки объектов: геометрические признаки, яркостные характеристики, характеристики цветности, текстурные свойства и признаки движения.

Эффективность хранения, передачи и защиты цифровой информации в большой степени зависят от методов кодирования цифровых изображений. Методам кодирования и сжатия цифровых изображений посвящен 8 раздел.

Текстура является важной характеристикой изображения. Проблемы описания цветных текстур, их синтеза и анализа представлены в материале 9 раздела, посвященного анализу текстурных изображений.

Мы стремились к тому, чтобы материал пособия позволил читателям ознакомиться с основными современными методами и алгоритмами компьютерной обработки и распознавания изображений и помог ориентироваться в большом мире книг по цифровым методам обработки информации.

# **1 ПРЕДМЕТ И ЗАДАЧИ КОМПЬЮТЕРНОЙ ОБРАБОТКИ И РАСПОЗНАВАНИЯ ИЗОБРАЖЕНИЙ**

## **1.1 Определение компьютерной обработки изображений**

Компьютерная обработка и распознавание изображений представляет собой быстро развивающуюся самостоятельную дисциплину.

*Компьютерная обработка изображений предполагает обработку цифровых изображений с помощью компьютеров или специализированных устройств, построенных на цифровых сигнальных процессорах.*

При этом под 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 [1]. Термин «иконика» происходит от греческого «eikon», что означает изображение, образ. Сегодня под ним понимают «создание и обработку изображений с помощью ЭВМ, что совпадает с понятием компьютерной обработки изображений.

Области применения цифровой обработки в настоящее время значительно расширяются, вытесняя аналоговые методы обработки сигналов изображений. Методы цифровой обработки широко применяются в промышленности, искусстве, медицине, космосе. Они применяются при управлении процессами, автоматизации обнаружения и сопровождения объектов, распознавании образов и во многих других приложениях. Цифровая передача изображений с космических аппаратов, цифровые каналы передачи сигналов изображений требуют обеспечения передачи все больших потоков информации. Если при передаче цифрового сигнала цветного телевидения необходимо передавать потоки порядка 216 Мбит/с, то для передачи телевидения высокой четкости скорость передачи должна составлять порядка 1 Гбит/с. 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 медицинской технике широко применяются системы формирования изображения, его преобразования в цифровую форму, визуализация и документирование путем введения в компьютер изображений с помощью специализированных устройств захвата видео. Автоматический анализ в системах дистанционного наблюдения широко применяется при анализе местности, в лесном хозяйстве, например, для автоматического подсчета площади вырубок, в сельском хозяйстве для наблюдения за созреванием урожая, при разведке, в системах противопожарной безопасности. Контроль качества производимой продукции выполняется благодаря автоматическим методам анализа сцен.

Компьютерная обработка изображений применяется в задачах экспертизы живописи неразрушающими методами [2]. Для восстановления старых фильмов применяются методы автоматической компенсации дефектов видеоматериала, полученного после преобразования киноизображения в видео.

Сегодня трудно представить область деятельности, в которой можно обойтись без компьютерной обработки изображений. Интернет, сотовый телефон, видеокамера, фотоаппарат, сканер, принтер, так прочно вошедшие в наш быт, - немыслимы без компьютерной обработки изображений.

*При компьютерной обработке изображений решается широкий круг задач, таких как улучшение качества изображений; измерение параметров; спектральный анализ многомерных сигналов; распознавание изображений; сжатие изображений.*

Устройства формирования изображений получили широкое распространение и применение в самых различных областях науки, техники, промышленности, медицине, биологии и др. [3-6]. Они являются неотъемлемыми компонентами систем и устройств, применяемых в фотокинотехнике, телевидении, системах технического зрения: дневного, ночного и теплового видения, при дистанционном зондировании Земли. Назначение этих систем предполагает решение комплекса технических и научных задач, требующих синтеза и анализа методов обработки, бинаризации, классификации изображений. Развитие микроэлектроники, переход от аналоговой формы сигналов к цифровой позволяют расширить палитру и повысить сложность применяемых алгоритмов для решения поставленных задач [7,8]. Рассмотрим некоторые из устройств формирования изображений.

## 1.2 Устройства формирования изображений

Устройства формирования изображений позволяют создавать изображения, порождаемые электромагнитным излучением в спектре от гамма излучения до инфракрасного (ИК). Рассмотрим некоторые из устройств формирования, которые включают электровакуумные и твердотельные фоточувствительные приборы.

Электровакуумные фоточувствительные приборы имеют фотокатод, эmitтирующий электроны, и анод, на который подается положительный потенциал. Большинство передающих электровакуумных трубок и систем на их основе работают в видимом диапазоне. Отдельную группу оптико-электронных систем составляют приборы ночного видения, работающие в ближнем ИК диапазоне. Приемником и преобразователем ИК излучения в видимый диапазон в них является электронно - оптический преобразователь (ЭОП) [9]. В ЭОП анодом является люминисцентный экран, создающий видимое изображение при бомбардировке его электронами. Таким образом, он преобразует оптическое излучение в оптическое излучение, а не в электрический сигнал.

Передающие электронно - лучевые трубы (телевизионные) преобразуют оптическое излучение в электрический сигнал. В электровакуумных фотоприборах возможна регистрация предельно малых оптических сигналов на уровне единичных фотоэлектронов, высокое разрешение. Основные недостатки – это сложные вакуумные стеклянные и металлоконструкции, ограниченный срок службы катодов, мишеней и анодов, высокие напряжения питания; ослепление, ограничение спектральной чувствительности близким ИК диапазоном, большие массогабаритные характеристики.

Твердотельные фоточувствительные приборы обеспечивают преобразование оптических сигналов в электрические в твердом теле. Существует два класса твердотельных фоточувствительных приборов: квантовые и тепловые приемники излучения.

Квантовые фотоприемники основаны на изменении электрических свойств полупроводника при поглощении фотона. Фотон передает свою энергию электрону, переводя его на более высокий энергетический уровень. Этот процесс фотогенерации сопровождается формированием пары носителей электрон и дырка. К таким приборам относятся фотодиоды, фоторезисторы, приборы с зарядовой связью (ПЗС). ПЗС обеспечивают высокое разрешение (размеры элементов раstra менее 10 мкм), большие форматы (768x576 и больше), возможность управления экспонированием, высокую однородность чувствительности элементов, большее, чем в ЭЛТ, отношение сигнал/шум (С/Ш), малые габариты, вес и потребляемую мощность, имеют сравнительно низкую стоимость и большую номенклатуру изделий.

Тепловые твердотельные приемники излучения используют эффект изменения электрических свойств материала (емкости, сопротивления) при изменении его температуры вследствие нагрева при поглощении теплового излучения (при радиационном нагреве). Болометрическим эффектом называется изменение электрического сопротивления материала  $R$  при радиационном нагреве вследствие изменения температуры  $T$  этого материала. Болометрический эффект характеризуется температурным сопротивлением материала  $\gamma_\delta = \frac{1}{R} \frac{\partial R}{\partial T}$ , где  $R$  – сопротивление материала (чувствительного слоя болометра) при температуре  $T$ . Приемник, построенный на основе этого эффекта, называется болометром. К фотоприемным устройствам резисторного типа относятся микроболометрические матрицы. Термочувствительный слой элементов таких матриц изготавливается из пленок окислов ванадия, кремния и германия. Для устранения влияния температуры окружающей среды, приводящей к нестабильности параметров, микроболометр заключают в вакуумированный корпус с термоэлектрической системой стабилизации рабочей температуры. Оптический модулятор не нужен. Сегодня эти приемники уступают охлаждаемым фотонным матричным приемникам по чувствительности, размерам элементов и быстродействию [10].

Второй тип приборов – это пироэлектрические фотоприемники. Они используют тонкие пленки особых кристаллических диэлектриков, обладающих пироэлектрическим (ферроэлектрическим) эффектом. Чувствительный элемент в таких приемниках представляет собой ферроэлектрический конденсатор, при изменении температуры которого, изменяется его диэлектрическая постоянная, а, следовательно, и емкость. Изменение емкости при постоянном приложенном к конденсатору напряжении, приводит к изменению заряда, поступающего от конденсатора на схему считывания сигнала. Поскольку изменение заряда происходит только при изменении температуры, то для наблюдения за объектами с постоянной температурой, необходимо модулировать падающее излучение. Осуществляется это обтюратором, который с частотой кадров перекрывает падающий на приемник поток излучения. В качестве диэлектриков используются ниобат калия тантала (*KTN*), титанат бария-стронций (*BST*) и другие виды керамики. Достоинствами пироэлектрических приемников являются почти равномерная спектральная чувствительность в широком диапазоне (от 800 нм до 25 мкм), высокая времененная стабильность и низкая стоимость.

Третий тип тепловых приборов построен на использовании термопар (термоэлектрические матрицы). Два слоя разнотипных металлов/полупроводников образуют термопару. Один слой облучается радиацией и нагревается, второй – экранирован от облучения. Между ними возникает термо ЭДС  $U$ . Величина ЭДС пропорциональна производной от этой ЭДС по температуре  $\gamma_T = \partial U / \partial T$ . Термопара является генератором напряжения. При последовательном включении термоэлементов чувствительность возрастает пропорционально числу включенных элементов. Стабилизатор температуры не нужен. Термопары часто образуют пленками алюминия и поликристаллического кремния, расположенными друг под другом. Пленки разделены слоем  $SiO_2$  [11]. Термоэлементы имеют линейные рабочие характеристики, не требуют обтюрации (в отличие от пироэлектрических приемников), работают без термостабилизации, не требуют источников питания.

Независимо от типа матричного фотоприемника общей тенденцией остается увеличение формата. При этом снижение размеров чувствительных элементов подошло к теоретическому пределу. Совершенствование технологии производства приемников направлено, прежде всего, на уменьшение темнового тока с целью повышения пороговой чувствительности, повышение однородности чувствительности элементов, повышение чувствительности элементов, уменьшение перекрестных межэлементных помех.

Неохлаждаемые приемники ИК излучения все еще отстают от традиционных охлаждаемых по температурному разрешению. Однако их основные параметры (чувствительность, пространственное и температурное разрешение) в последние годы улучшаются, что в

сочетании с такими преимуществами, как отсутствие холодильника, низкая стоимость, формат кадра порядка 640x480 элементов, обусловливают все более широкое их применение в мобильных тепловизионных системах широкого применения.

## 2 МАТЕМАТИЧЕСКИЕ МОДЕЛИ ИЗОБРАЖЕНИЙ

### 2.1 Модели непрерывных изображений

Компьютерная обработка изображений возможна после преобразования сигнала изображения из непрерывной формы в цифровую форму. Эффективность обработки зависит от адекватности модели, описывающей изображение, необходимой для разработки алгоритмов обработки. При этом необходимо учитывать влияние передающей и приемной систем и канала связи на сигнал изображения. Модель изображения представляет систему функций, описывающих существенные характеристики изображения: функцию яркости, отражающую изменение яркости в плоскости изображения, пространственные спектры и спектральные интенсивности изображений, функции автокорреляции. Канал изображения содержит оптическую систему, оптико - электрический преобразователь, устройство аналого - цифрового преобразования (АЦП) и цифровой обработки сигналов изображения. В общем случае непрерывное изображение может быть представлено функцией пяти аргументов: трех пространственных координат, времени и длины волны электромагнитного излучения. Упрощения модели пространственно - временных сигналов в некотором диапазоне волн  $f(x, y, z, t, \lambda)$  приводят к моделям пространственно - временного сигнала  $f(x, y, z, t)$ , пространственного сигнала  $f(x, y, z)$ , временного сигнала  $f(t)$ . Здесь  $x, y, z$  - пространственные координаты,  $t$  - время,  $\lambda$  - длина волны электромагнитного излучения.

### 2.2 Пространственные спектры изображений

При обработке изображений широко используется анализ спектров изображений. Спектр изображения получают прямым двумерным преобразованием Фурье функции, описывающей изображение [12]:

$$F(\omega_x, \omega_y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \exp(-i(\omega_x x + \omega_y y)) dx dy, \quad (2.1)$$

где  $\omega_x, \omega_y$  - пространственные частоты;  $i = \sqrt{-1}$ , мнимая единица.

Функция  $\exp(-i(\omega_x x + \omega_y y))$  при фиксированных значениях пространственных частот описывает плоскую волну в плоскости изображения  $(x, y)$  (в соответствии с рисунком 2.1).

Формула (2.1) связывает вещественную функцию, описывающую яркость изображения  $f(x, y)$  с комплексной функцией частоты – спектром изображения  $F(\omega_x, \omega_y)$ :

$$F(\omega_x, \omega_y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \cos(\omega_x x + \omega_y y) dx dy +$$

$$+ i \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (-f(x, y)) \sin(\omega_x x + \omega_y y) dx dy = Re(\omega_x, \omega_y) + i Im(\omega_x, \omega_y), \quad (2.2)$$

где  $Re(\omega_x, \omega_y)$  - реальная часть спектра;  $Im(\omega_x, \omega_y)$  - мнимая часть спектра.

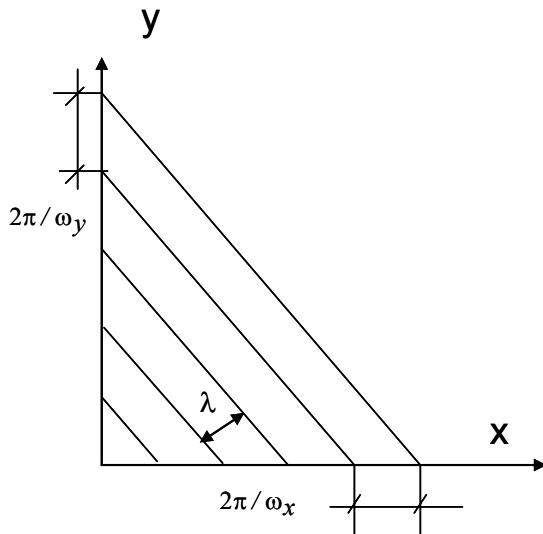


Рисунок 2.1 Определение пространственных частот изображения.

Амплитуда и фаза спектра определяются по формулам (2.3) и (2.4) соответственно:

$$F(\omega_x, \omega_y) = \sqrt{Re(\omega_x, \omega_y)^2 + Im(\omega_x, \omega_y)^2}, \quad (2.3)$$

$$\phi(\omega_x, \omega_y) = \arctg(Im(\omega_x, \omega_y) / Re(\omega_x, \omega_y)).$$

Из (2.3)

$$F(\omega_x, \omega_y) = F(\omega_x, \omega_y) \exp(i\phi(\omega_x, \omega_y)). \quad (2.4)$$

Обратное преобразование Фурье позволяет восстановить изображение по его спектру:

$$f(x, y) = (1/4\pi^2) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(\omega_x, \omega_y) \exp(i(\omega_x x + \omega_y y)) d\omega_x d\omega_y. \quad (2.5)$$

### 2.3 Спектральные интенсивности изображений

Спектральная интенсивность изображения характеризует распределение энергии по пространственным частотам. Она определяется как квадрат модуля спектра изображения:

$$S(\omega_x, \omega_y) = Re(\omega_x, \omega_y)^2 + Im(\omega_x, \omega_y)^2 = F^2(\omega_x, \omega_y). \quad (2.6)$$

Для ее названия используются термины спектральная плотность и энергетический спектр.

Энергия изображения определяется как интеграл энергетического спектра по пространственным частотам. В соответствии с теоремой Парсеваля энергия изображения может быть вычислена в соответствии с (2.7):

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f^2(x,y) dx dy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |F(\omega_x, \omega_y)|^2 d\omega_x d\omega_y. \quad (2.7)$$

## 2.4 Вероятностные модели изображений и функции автокорреляции

Вероятностные модели изображений широко используются для описания изображений. Изображение в этом случае рассматривается как случайная функция пространственных координат  $(x,y)$  и времени  $t$ . Случайный процесс называется *стационарным в широком смысле*, если он имеет постоянные значения математического ожидания и дисперсии, а его автокорреляционная функция зависит не от координат, а от их разностей (сдвига). Случайный процесс называется *стационарным в узком смысле*, если его  $n$ -мерная плотность распределения вероятностей инвариантна к сдвигу. В этом случае не зависят от времени и моменты более высокого порядка, в частности, асимметрия и эксцесс. Случайный процесс описывается плотностью вероятности распределения яркости в изображении по пространственным координатам для некоторого фиксированного момента времени  $t$   $p(x,y)$ .

В соответствии с определением математическое ожидание (среднее значение) стационарного процесса в широком смысле

$$Mf = \xi = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y) p(x,y) dx dy = \text{const.} \quad (2.8)$$

Дисперсия

$$Df = \sigma^2 = E(f(x,y) - \xi)^2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (f(x,y) - \xi)^2 p(x,y) dx dy = \text{const.} \quad (2.9)$$

Функция автокорреляции вычисляется в соответствии с (2.10):

$$R(\tau_x, \tau_y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y) f(x - \tau_x, y - \tau_y) dx dy, \quad (2.10)$$

где  $\tau_x, \tau_y$  задают сдвиги изображения по соответствующим осям координат.

Для действительной функции  $f$  автокорреляционная функция является действительной и четной.

Спектр двумерной автокорреляционной функции изображения (прямое преобразование Фурье автокорреляционной функции) равен энергетическому спектру изображения (спектральной плотности мощности) по определению:

$$S(\omega_x, \omega_y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} R(\tau_x, \tau_y) \exp(-i(\omega_x \tau_x + \omega_y \tau_y)) d\tau_x d\tau_y. \quad (2.11)$$

Стационарный случайный процесс называется *эргодическим*, если любая его вероятностная характеристика может быть получена из одной реализации путем усреднения по времени. При этом среднее по времени равно среднему по ансамблю реализаций. Свойство эргодичности используется при оценке вероятностных характеристик изображений.

## 2.5 Критерии качества изображений

Качество изображения может определяться статистическими, спектральными, яркостными характеристиками изображения. В большинстве практических применений качество рассматривается как мера близости двух изображений: реального и идеального или преобразованного и исходного. При таком подходе можно оценивать как субъективную степень похожести изображений, так и получать объективные оценки параметров сигналов изображения: моменты первого и второго порядка разностного сигнала сравниваемых изображений, такие параметры преобразования как отношение С/Ш, коэффициенты сжатия информации и другие.

Субъективные критерии - это критерии визуального восприятия, оцениваемые в процессе экспертизы некоторой группой наблюдателей (экспертов). Наибольшее распространение получил метод оценок, при котором наблюдатель оценивает качество изображения в баллах по определенной шкале, считая, что идеальное изображение имеет максимальный балл. Этот метод позволяет оценить такие характеристики изображения как правильность цветопередачи, координатные искажения, чистоту переходов и др. Основные шкалы оценок при использовании метода сравнения приведены в таблице 2.1 [13].

Для интерпретации полученных экспертных оценок разработаны методы их представления, например построение кумулятивных кривых распределения оценок как функции от искажений. Средняя оценка определяется по формуле

$$g_{cp} = (1/N) \sum_{i=1}^r i n_i,$$

где  $N$ -общее число оценок,  $n_i$ - число оценок равных  $i$  баллам,  $r$  - количество видов разных оценок.

Нормализованные оценки  $p$  выражают относительное качество в диапазоне  $[0,1]$ . При пятибалльной системе, когда  $g \in [1,5]$ :

$$p = (g - 1) / 4,$$

а средняя оценка вычисляется в соответствии с формулой:

$$p_{cp} = (n_5 + 0,75n_4 + 0,5n_3 + 0,25n_2) / N.$$

Таблица 2.1 Основные шкалы субъективных оценок качества изображения

Нормализованная	Пятибалльная шкала		Семибалльная шкала ухудшения
	Шкала качества	Шкала ухудшения	
1	5 (отлично)	5 (незаметно)	1 (незаметно)
0,875	-	-	2 (едва заметно)
0,75	4 (хорошо)	4 (заметно, но не мешает)	3 (заметно, но лишь слегка ухудшает)
0,625	-	-	4 (ухудшает, но не мешает)
0,5	3 (удовлетворительно)	3 (заметно, немного мешает)	5 (несколько мешает)
0,25	2 (плохо)	2 (мешает)	6 (определенко мешает)
0	1 (очень плохо)	1 (сильно мешает)	7 (крайне мешает)

Единицей ухудшения качества телевизионных (ТВ) изображений является имп (от *impairment* –ухудшение, повреждение). Эта единица введена Прессером, Аллнаттом и Льюисом в 1964 г. и используется МККР (Международным консультативным комитетом по радиосвязи (CCIR)). Ухудшение обратно пропорционально нормализованной оценке качества и изменяется от  $\infty$  до 0 при изменении  $p$  от 0 до 1 в соответствии с формулой:

$$I = 1/p - 1.$$

Достоинство методики оценки ухудшения состоит в том, что результирующая оценка ухудшения получается арифметическим суммированием оценок ухудшения, вызванных различными видами искажений сигналов изображения. Основываясь на психофизических свойствах наблюдателя, субъективные оценки позволяют характеризовать восприятие изображения. Интегральный критерий качества формируется по обобщенной формуле:

$$I_{\Sigma} = \sum_{i=1}^M I_i^v,$$

где  $M$ -число параметров, по которым оценивается качество изображения;  $v$  - показатель степени.

Значение показателя степени принимают равным 1, но могут быть использованы, например, такие значения как 0,78 или 2. В настоящее время применяются и другие оценки качества изображений. При разработке аппаратных средств специального назначения большое значение имеет оценка объективных характеристик качества преобразованного изображения.

Объективными критериями, используемыми при оценке качества изображений, являются критерии, позволяющие получить просто

вычисляемую характеристику изображения разностного сигнала. К таким критериям относится, прежде всего, среднеквадратический критерий. По нему мерой различия двух изображений  $f(x, y)$  и  $f_{\text{пр}}(x, y)$  является среднеквадратическое значение разностного сигнала двух изображений. Для непрерывных изображений, заданных при  $x \in [0, N]$  и  $y \in [0, M]$ , среднеквадратическое отклонение (СКО) вычисляется по формуле:

$$\sigma^2 = \int_0^M \int_0^N [f(x, y) - f_{\text{пр}}(x, y)]^2 dx dy. \quad (2.12)$$

В некоторых случаях используется критерий максимальной ошибки, который в отличие от (2.12), позволяет установить значение максимальной ошибки преобразования:

$$\epsilon_{\max} = \max_{(x, y)} |f(x, y) - f_{\text{пр}}(x, y)|. \quad (2.13)$$

Применяются и другие объективные критерии качества изображений [14].

Существует определенное разногласие в оценках качества, даваемых человеческим глазом (субъективных), и объективных, полученных в виде количественных показателей. Глаз является совершенным изобретением природы, с ним не могут соревноваться достаточно примитивные объективные оценки типа СКО, пикового отношения сигнал/шум (ПСШ) и др. Поэтому некоторые результаты, рассматриваемые с точки зрения объективных оценок как одинаковые, визуально могут восприниматься различно. Однако объективные критерии используются при компьютерной обработке изображений в системах с автоматическим принятием решений. Функционирование автоматических компьютерных систем полностью подчинено математическим критериям, и качество их работы оценивается только объективными показателями. Понятно, что и качество изображений, используемых в этих системах, также должно оцениваться только по объективным критериям.

### 3 ЦВЕТ КАК ВАЖНАЯ ХАРАКТЕРИСТИКА ВОСПРИЯТИЯ ОБЪЕКТА

Цветовые характеристики несут информацию об отражательных свойствах объекта. Различие отражательной способности объекта в разных участках спектрального диапазона обеспечивает возможность извлечения важной биологической информации об объекте. Ведь именно по цвету можно оценить, например, созрел ли плод, поражены ли заболеванием сельскохозяйственные культуры и многое другое.

Теория цветового зрения и сегодня находится в незавершенной стадии развития. Проблема состоит в том, что многочисленные модели описывают цветовое зрение, но не являются теорией цветового зрения [15], поскольку ни одной из них не удается строго ответить на все вопросы об установленных фактах, относящихся к психофизическому и

физиологическому аспектам цветового зрения. В книге [16] выделен специальный параграф “Некоторые нерешенные проблемы цветового зрения”. Начиная с опытов Ньютона и Максвелла, было предложено множество теорий, описывающих цветовое зрение человека. В классической трехцветной модели цветового зрения, разработанной Томасом Юнгом в 1802 году, предполагается, что существуют три компонента любого цветоощущения, которые являются аддитивными основными цветами: это красный (*R*), зеленый (*G*) и синий (*B*). На самом деле имеется бесконечное множество основных цветов, но чтобы получить максимальный диапазон смешанных цветов, следует пользоваться *RGB*. Единственное условие правильного выбора основных цветов состоит в том, что при смешении двух из них мы не должны получать третий цвет. Юнг постулировал, что поскольку трехкомпонентность цвета не имеет обоснования в теории света, то цвет является свойством самого глаза. Глаз анализирует каждый цвет в отдельности и передает сигналы о нем в мозг по трем типам нервных волокон: один тип передает сигнал о наличии *R*, второй - *G*, третий - *B*. На 50 лет теория Юнга была отвергнута и предана забвению. В 1852 году к ней одновременно обратились немецкий физик и физиолог Герман фон Гельмгольц и шотландский физик Джеймс Клерк Максвелл. Гельмгольц при попытке получить сине-зеленый цвет с длиной волны 500 нм смешением *BG*, заметил, что его нельзя получить путем аддитивного сложения трех основных цветов. Смесь получается белесая, менее насыщенная по сравнению со спектральным цветом. Через 10 лет Гельмгольц понял, что результаты опытов можно объяснить и на основе трех основных механизмов, исходя из предположения о том, что они обладают спектральной чувствительностью в широком, частично перекрывающемся диапазоне. При таком подходе, даже если раздражитель чистый в оптическом смысле, ответная реакция глаза таковой не является. Максвелл одним из первых признал теорию Юнга и занялся разработкой точных методов измерения цветов. Он использовал цветовой треугольник Юнга, поместив основные цвета *RGB* в вершины равностороннего треугольника. Результирующий цвет любой смеси *RGB* располагается в центре тяжести трех масс. Результирующая аддитивной смеси двух цветов находится в их центре тяжести и поэтому лежит на прямой, соединяющей эти цвета. Этот закон центра тяжести является свойством всех плоских цветовых диаграмм. Теория Юнга - Гельмгольца не соответствовала цветовым ощущениям. Человек в состоянии различать, по меньшей мере, четыре качественно разных цветовых ощущения: красного, желтого, зеленого и синего цветов, - если к ним добавить белый, то получится пять.

В 1870 году немецкий физиолог Эвальд Геринг сформулировал оппонентную теорию цветового зрения. Он опирался на существование пяти психологических ощущений и считал, что они действуют в противоположных парах. В парах: красный и зеленый, желтый и синий - цвета являются противоположными и не смешиваются.

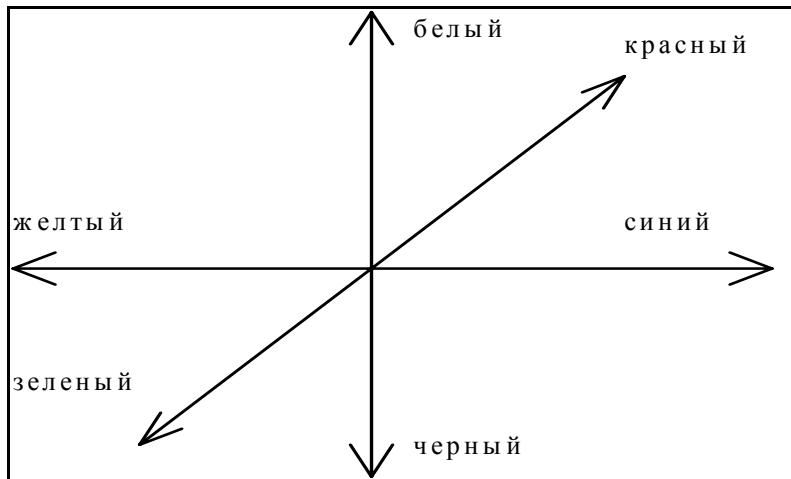


Рисунок 3.1 Оси противоположных цветов в соответствии с оппонентной теорией Геринга.

После Геринга был столетний перерыв в развитии теории цветового зрения. В 1953 году Томсон и Райт опубликовали кривые спектральной чувствительности к красному, синему и зеленому диапазонам спектра. В 1964 году две группы американских ученых (Маркс, Добелл, Мак - Никол в опытах на сетчатке серебряного карася, обезьяны и человека, и Браун и Уолд на сетчатке человека) обнаружили три типа колбочек, поглощающих свет в различных частях спектра. Согласно современным данным на рецептурном уровне свет регистрируется тремя различными типами колбочек (как постулировано в теории Юнга - Гельмгольца), и эти рецепторы обладают чувствительностью к  $R$ ,  $G$ ,  $B$  - частям спектра. Однако, поступающая от них информация, по - видимому, преобразуется в импульсные разряды и до передачи в мозг кодируется в сетчатке. Эта закодированная информация посыпается в виде сигнала о яркости из всех трех типов колбочек, а также в виде разностных сигналов каждого двух цветов. Подключается и второй яркостный сигнал, берущий начало, вероятно, от независимой палочковой системы. Мозг воспринимает закодированную информацию о яркости и разностные цветовые сигналы. Таков механизм цветного зрения в соответствии с зонной теорией Адамса [17].

Концепция построения систем цветного телевидения основана на принципе постоянной яркости и согласуется с зонной теорией Адамса.

В рамках трехкомпонентной теории цвета набор основных цветов можно выбрать многими способами, этим объясняется большое количество координатных систем, предложенных для количественного описания цвета. Описание этих координатных систем приводится как в фундаментальных исследованиях по цвету, так и в научно-технических статьях при описании различных алгоритмов обработки. Методы анализа цветных изображений зависят от цветового координатного пространства [18], выбор цветового координатного пространства определяет эффективность метода.

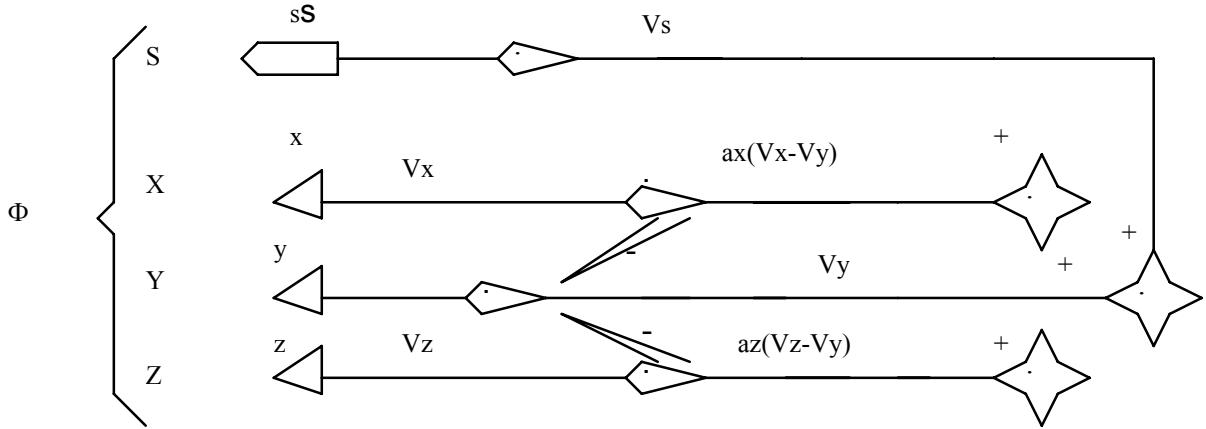


Рисунок 3.2 Механизм цветного зрения в соответствии с зонной теорией Адамса.

Рассмотрим некоторые цветовые модели.

### 3.1 Цветовая модель *RGB*

В цветовом координатном пространстве *RGB* любой цвет получается как сумма (смешение) красного, зеленого и синего цветов. Если представить это пространство в виде куба, то на главной диагонали куба, образованного из нормированных компонентов, будут расположены серые цвета (ахроматические). Наряду с тем, что накоплен большой объем информации о реакции и чувствительности глаза к трем стимулам *RGB*, это цветовое пространство является аппаратно ориентированным. Цветные электронно - лучевые трубы и жидкокристаллические дисплеи отображают цветные изображения, основываясь на аддитивной смеси этих трех компонентов.

### 3.2 Цветовая модель *YCrCb*

В этом пространстве компонент *Y* включает в себя только информацию о яркости пикселов, а компоненты *Cb* и *Cr* содержат только информацию о цвете и насыщенности. Поскольку органы зрения менее чувствительны к цвету предметов, чем к их яркости, такое пространство позволяет передать компонент яркости с большим разрешением, чем компоненты цветности.

Определение компонентного сигнала задается через сигналы основных цветов *R*, *G*, *B* из уравнения, рекомендованного стандартом федеральной комиссии связи (*FCC*) [19]:

$$Y=0,299 R+ 0,587G + 0,114B. \quad (3.1)$$

$$(R-Y)= R- 0,299 R- 0,587G - 0,114B = 0,701R- 0,587G- 0,114B.$$

$$(B-Y)= B- 0,299 R- 0,587G - 0,114B= - 0,299R- 0,587G + 0,886B.$$

Здесь *R*, *G*, *B* - исходные сигналы основных цветов, подвергнутых предварительной гамма - коррекции с целью обеспечения оптимального качества изображения на экране кинескопа. Если значения сигналов привести к единице (максимальный уровень сигнала - 1В), то получим значения для белого, черного и насыщенных основных и дополнительных цветов, представленные в таблице 3.1.

Таблица 3.1.

Цвет	<i>R</i>	<i>G</i>	<i>B</i>	<i>Y</i>	<i>Cr=R-Y</i>	<i>Cb=B-Y</i>
Белый	1,0	1,0	1,0	1,0	0	0
Черный	0	0	0	0	0	0
Красный	1,0	0	0	0,299	0,701	-0,299
Зеленый	0	1,0	0	0,587	-0,587	-0,587
Синий	0	0	1,0	0,114	-0,114	0,886
Желтый	1,0	1,0	0	0,886	0,114	-0,886
Голубой	0	1,0	1,0	0,701	-0,701	0,299
Пурпурный	1,0	0	1,0	0,413	0,587	0,587

Значения сигнала *Y* находятся в пределах от 0 до 1, значения цветоразностных сигналов изменяются от -0,701 до 0,701 для *Cr* и от -0,886 до 0,886 для *Cb*. Приведение диапазонов изменения цветоразностных сигналов к единице достигается введением нормирующих коэффициентов  
 $Kr=0,5/0,701=0,713$ ,  $Kb=0,5/0,886=0,564$ .

Сигнал яркости и нормированные цветоразностные сигналы связаны с сигналами основных цветов следующим матричным преобразованием:

$$\begin{bmatrix} Y \\ Cr \\ Cb \end{bmatrix} = \begin{bmatrix} 0,299 & 0,587 & 0,114 \\ 0,5 & -0,41869 & 0,08131 \\ -0,168874 & -0,33126 & 0,5 \end{bmatrix} \times \begin{bmatrix} R \\ G \\ B \end{bmatrix}. \quad (3.2)$$

Переход от цветового координатного пространства *RGB* к пространству *YCrCb* (3.1) соответствует рекомендациям Международного телекоммуникационного союза (*ITU - International Telecommunication Union*) *ITU-T* с идентификатором *BT.601*.

Определение компонентного сигнала задается через сигналы основных цветов *R*, *G*, *B*. При 8-ми разрядном представлении компонентов диапазон значений составляет для *Y* [0,255], а для *Cr* и *Cb* - [-128,127]. После перевода цветоразностных сигналов в диапазон [-128,127], получим матрицу для основных цветов, определяющую обратное преобразование в соответствии с уравнениями:

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1 & 1,402 & 0 \\ 1 & -0,714 & -0,344 \\ 1 & 0 & 1,772 \end{bmatrix} \times \begin{bmatrix} Y \\ Cr \\ Cb \end{bmatrix}. \quad (3.3)$$

Это пространство используется в ТВ системах *PAL* и *SECAM*, а также при кодировании неподвижных изображений и видеопоследовательностей. При формировании сигнала используется сокращение избыточности цветоразностных сигналов. Этот принцип основан на особенности человеческого зрения не различать или плохо различать цвета мелких деталей изображения. Экспериментально было установлено [20], что при расстоянии до экрана цветного телевизора  $L=4,5h$ , где  $h$  - высота экрана,

наблюдатель не ощущает мелкие синие детали как цветные при пространственной частоте этих деталей  $> 0,5 - 0,6$  МГц, а красные – при частоте  $> 1,3 - 1,5$  МГц.

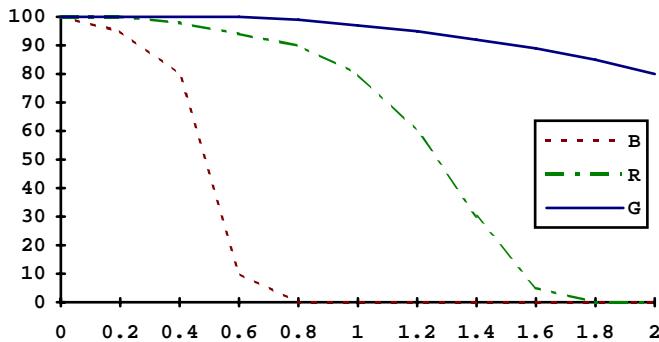


Рисунок 3.3 Графики зависимости видимой насыщенности от размеров деталей и их цветов.

На этой особенности человеческого зрения основано построение аналоговых и цифровых систем цветного телевидения, в которых частота дискретизации сигнала яркости в 2 раза превышает частоту дискретизации каждого из цветоразностных каналов. В форматах 4:2:2 и 4:2:0 вводится по одному цветоразностному отсчету на 2 отсчета яркости [21].

### 3.3 Цветовая модель $YIQ$

Эта модель используется в коммерческом цветном ТВ, тесно связана с цветной растровой графикой и представляет собой вариант кодирования цветов  $RGB$ , обеспечивающий совместимость с черно – белым телевидением. Это пространство используется в ТВ системе NTSC в США. Координата  $Y$  при этом совпадает с координатой  $Y$  в колориметрической системе МКО (Международная комиссия по освещению – *Comission Internationale de l'Eclairage, CIE*). Компонент  $I$  представляет тон, а компонент  $Q$  – насыщенность. Преобразование модели  $RGB$  в модель  $YIQ$  выполняется в соответствии с системой уравнений:

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0,299 & 0,587 & 0,114 \\ 0,596 & -0,275 & -0,321 \\ 0,212 & -0,523 & 0,311 \end{bmatrix} \times \begin{bmatrix} R \\ G \\ B \end{bmatrix}. \quad (3.4)$$

Соответственно обратное преобразование выполняется следующим образом:

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1 & 0,956 & 0,620 \\ 1 & -0,272 & -0,647 \\ 1 & -1,108 & 1,705 \end{bmatrix} \times \begin{bmatrix} Y \\ I \\ Q \end{bmatrix}. \quad (3.5)$$

### 3.4 Цветовая модель $L^*a^*b^*$ МКО 1976

Эта система координат обеспечивает относительно точное представление цветов в соответствии с системой цветов, разработанной в

1905 году художником Манселлом. Эта система может быть получена после преобразования системы  $RGB$  в цветовую координатную систему  $XYZ$  МКО 1931 в соответствии с уравнениями:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 2,769 & 1,7518 & 1,13 \\ 1 & 4,5907 & 0,0601 \\ 0 & 0,0565 & 5,5943 \end{bmatrix} \times \begin{bmatrix} R \\ G \\ B \end{bmatrix}.$$

Сумма коэффициентов при компонентах составляет 5,651. С учетом нормировки преобразование следует выполнять в соответствии с системой:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0,49 & 0,31 & 0,1999646 \\ 0,17695983 & 0,81242258 & 0,0106175 \\ 0 & 0,01008 & 0,989913 \end{bmatrix} \times \begin{bmatrix} R \\ G \\ B \end{bmatrix}. \quad (3.6)$$

Затем выполняется преобразование системы  $XYZ$  в систему  $L^*a^*b^*$  в соответствии с уравнениями (3.7). Координата  $L^*$  определяет яркость цвета,  $a^*$  - соотношение красного и зеленого цветов,  $b^*$  - соотношение синего и зеленого.

$$L^* = \begin{cases} 116(Y/Y_0)^{1/3} - 16, & Y/Y_0 > 0,008856 \\ 903,3Y/Y_0, & \text{иначе} \end{cases} \quad (3.7)$$

$$a^* = 250(f((X/X_0)^{1/3}) - f((Y/Y_0)^{1/3}))$$

$$b^* = 100(f((Y/Y_0)^{1/3}) - f((Z/Z_0)^{1/3})),$$

$$\text{где } f(t^{1/3}) = \begin{cases} t^{1/3}, & t > 0,008856 \\ 7,787t + 16/116, & \text{иначе} \end{cases}$$

$X_0, Y_0, Z_0$  - координаты опорного белого цвета в системе  $XYZ$ .

### 3.5 Цветовая модель $L^*H^\circ C^*$

Цилиндрические координаты этого пространства соответствуют как эмпирической системе Манселла, так и согласуются с физиологической моделью цветного зрения. Эти координаты известны как психометрическая яркость, тон и насыщенность и задаются по формуле:

$$\begin{aligned} L^* &= L^* \\ H^\circ &= \arctg(b^*/a^*) \\ C^* &= [(a^{*2} + b^{*2})]^{1/2}. \end{aligned} \quad (3.8)$$

При таком описании элементом является круговой цилиндрический сегмент, выделенный на рисунке 3.4. Поверхности элементарного объема формируются в соответствии с рисунком при заданных диапазонах изменения яркости и цветности. Горизонтальные срезы формируются при условии постоянной яркости, вертикальные срезы, проходящие через ахроматическую ось  $OL^*$ , получаются при постоянном тоне, а части

цилиндрических поверхностей, концентрических относительно оси  $OL^*$ , формируются при постоянной насыщенности.

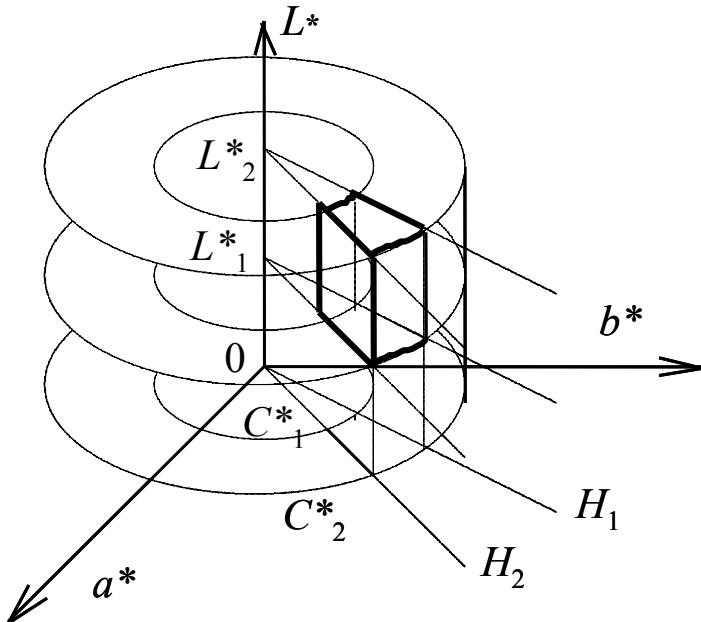


Рисунок 3.4 Элемент в цветовом координатном пространстве  $L^*a^*b^*$ .

### 3.6 Цветовая модель HSI

$$I = \frac{R + G + B}{3}. \quad (3.9)$$

$$S = 1 - \frac{3\min(R, G, B)}{R + G + B}. \quad (3.10)$$

$$H = \begin{cases} (G - B) / [3(R + G - 2B)], & \text{если } B = \min(R, G, B) \\ (B - R) / [3(G + B - 2R)] + 1/3, & \text{если } R = \min(R, G, B) . \\ (R - G) / [3(R + B - 2G)] + 2/3, & \text{если } G = \min(R, G, B) \end{cases} \quad (3.11)$$

Выполним обратное преобразование цветового координатного пространства  $HSI$  в пространство  $RGB$  и получим:

если  $H \leq 1/3$ , то

$$\begin{cases} B = (1 - S)I \\ G = 9SIH + B ; \\ R = 3I - (G + B) \end{cases} \quad (3.12)$$

если  $1/3 < H \leq 2/3$ , то

$$\begin{cases} R = (1 - S)I \\ B = 9SI(H - 1/3) + R ; \\ G = 3I - (R + B) \end{cases} \quad (3.13)$$

если  $H > 2/3$ , то

$$\begin{cases} G = (1 - S)I \\ R = 9SI(H - 2/3) + G . \\ B = 3I - (R + G) \end{cases} \quad (3.14)$$

### 3.7 Цветовая модель HLS

Эти цветовые координаты введены Тененбаумом (Стэнфордский исследовательский институт) и широко используются при анализе сцен [22].

Тон и насыщенность определяются через  $rgb$  координаты, определяемые как нормированные тристимульные значения:

$$r = \frac{R}{R+G+B}; \quad g = \frac{G}{R+G+B}; \quad b = \frac{B}{R+G+B}. \quad (3.15)$$

Локус  $r+g+b=1$  определяет треугольник Максвелла, изображенный на рисунках 3.5 и 3.6. На рисунке 3.5 приняты следующие обозначения:  $P$ -цветной элемент;  $W$ -серый,  $r=g=b=1/3$ ;  $P'$ -пересечение  $OP$  с плоскостью треугольника. Пересечение вектора  $OP$  с плоскостью треугольника Максвелла определяет тон и насыщенность в соответствии с выражениями 3.16:

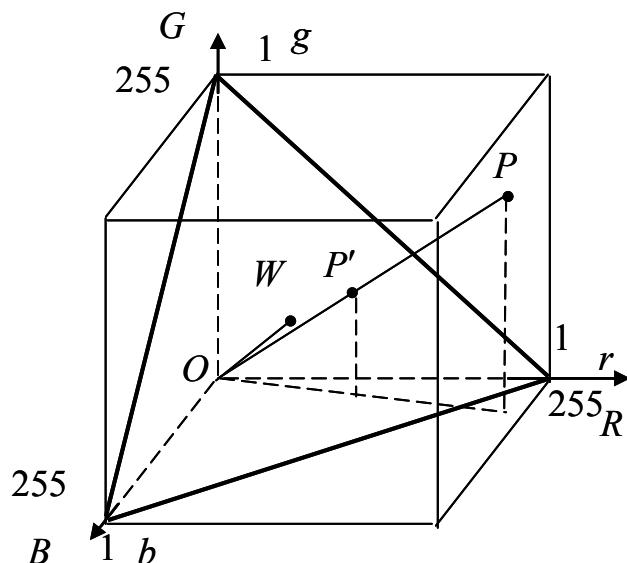


Рисунок 3.5 Цветовое координатное пространство  $RGB$ .

$$H=\phi \quad (0 \leq \phi \leq 2\pi); \quad (3.16 \text{ а})$$

$$S=WP'/WA \quad (0 \leq S \leq 1). \quad (3.16 \text{ б})$$

Яркость  $L$  пропорциональна длине вектора  $OP$  на рисунке 3.4 и определяется в соответствии с уравнением:

$$L=(R+G+B)/3. \quad (3.17)$$

Нейтральная точка, или точка серого,  $W$  представляет точку с равными компонентами  $R, G, B$ . Относительно этой точки определяются координаты  $H$  и  $S$  (в соответствии с рисунком 3.6).

Выполним прямое преобразование, чтобы затем получить формулы обратного преобразования из пространства  $HLS$  в пространство  $RGB$ .

Треугольник Максвелла задается тремя точками с координатами  $(1,0,0), (0,1,0)$  и  $(0,0,1)$  в координатной системе  $rgb$ . Уравнение плоскости, проходящей через эти точки, в соответствии с уравнением плоскости в отрезках, имеет вид:

$$\frac{r}{1} + \frac{g}{1} + \frac{b}{1} = 1, \text{ откуда } r+g+b=1. \quad (3.18)$$

Точка  $W$  является центром тяжести треугольника Максвелла и имеет координаты  $(1/3, 1/3, 1/3)$ . Угол между  $OW$  и плоскостью треугольника Максвелла составляет  $90^\circ$ :

$$\arcsin\left(\frac{\frac{1}{3} + \frac{1}{3} + \frac{1}{3}}{\sqrt{3}\sqrt{\frac{1}{9} + \frac{1}{9} + \frac{1}{9}}}\right) = \arcsin(1) = \frac{\pi}{2}, \quad (3.19)$$

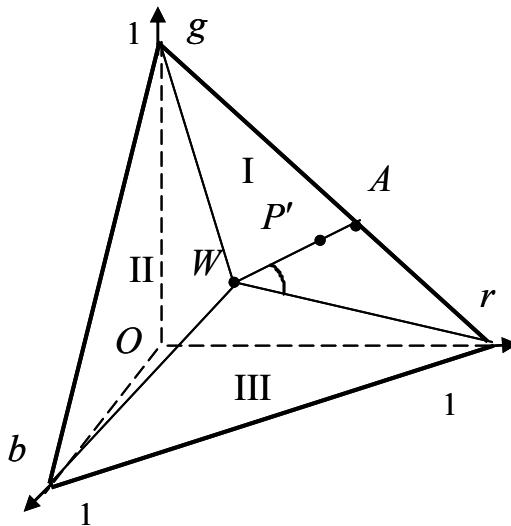


Рисунок 3.6 Цветовое координатное пространство  $HLS$ .

а модуль вектора  $OW$  определяется в соответствии с выражением:

$$|OW| = \sqrt{1/9 + 1/9 + 1/9} = \sqrt{3}/3. \quad (3.20)$$

Пусть точка  $P'$  имеет координаты  $(r_1, g_1, b_1)$ , тогда уравнение прямой  $OP$  можно записать в виде [23]:

$$\begin{aligned} \frac{r - r_O}{r_P - r_O} &= \frac{g - g_O}{g_P - g_O} = \frac{b - b_O}{b_P - b_O}, \\ r/r_1 &= g/g_1 = b/b_1. \end{aligned} \quad (3.21)$$

Отсюда направляющий вектор прямой  $OP$  имеет координаты  $(r_1, g_1, b_1)$ . Определим угол между прямой  $OP$  и плоскостью треугольника Максвелла:

$$\begin{cases} r + g + b - 1 = 0 \\ r/r_1 = g/g_1 = b/b_1 \end{cases}. \quad (3.22)$$

$$\sin\gamma_3 = \frac{r_1 + g_1 + b_1}{\sqrt{3}\sqrt{r_1^2 + g_1^2 + b_1^2}} = \frac{1}{\sqrt{3}\sqrt{r_1^2 + g_1^2 + b_1^2}}. \quad (3.23)$$

$$WP' = OW \operatorname{ctg} \gamma_3 = \frac{1}{\sqrt{3}} \frac{\sqrt{1 - \sin^2 \gamma_3}}{\sin \gamma_3}. \quad (3.24)$$

Уравнение прямой  $gr$ , где  $g(0,1,0)$ ,  $r(1,0,0)$  имеет вид:

$$\frac{r}{1} = \frac{g-1}{-1}, \text{ или } r=1-g. \quad (3.25)$$

Координаты точки  $P'$  определяются как координаты точки пересечения плоскости треугольника Максвелла и прямой  $OP$ :

$$\begin{cases} b+r+g-1=0 \\ r/r_1 = g/g_1 \\ r/r_1 = b/b_1 \end{cases}, \quad (3.26)$$

$$(b_1/g_1)g + (r_1/g_1)g + g - 1 = 0, \quad (3.27)$$

$$g = \frac{1}{b_1/g_1 + r_1/g_1 + 1} = \frac{g_1}{b_1 + g_1 + r_1} = g_1, \quad (3.28)$$

аналогично  $b=b_1$ ,  $r=r_1$ . То есть координаты точек  $P$  и  $P'$  совпадают.

Уравнение прямой  $WP'$  ( $W(1/3, 1/3, 1/3)$ ,  $P'(b_1, g_1, r_1)$ ) имеет вид:

$$\left(r - \frac{1}{3}\right) / \left(r_1 - \frac{1}{3}\right) = \left(g - \frac{1}{3}\right) / \left(g_1 - \frac{1}{3}\right) = \left(b - \frac{1}{3}\right) / \left(b_1 - \frac{1}{3}\right). \quad (3.29)$$

Рассмотрим 3 случая: первый, когда точка  $P'$  находится в секторе I, в треугольнике  $RWG$ ,  $\theta_0 = 0^\circ$ ; второй, когда точка  $P'$  находится в секторе II, в треугольнике  $GWB$ ,  $\theta_0 = 120^\circ$ ; третий, когда точка  $P'$  находится в секторе III, в треугольнике  $BWR$ ,  $\theta_0 = 240^\circ$ .

Рассмотрим сектор I. Координаты точки  $A$  определяются как координаты точки пересечения прямых  $WP'$  и  $GR$ . Прямая  $GR$  задается системой:

$$\begin{cases} b=0 \\ r=1-g \end{cases}. \quad (3.30)$$

Из (3.29) при  $b=0$  получим:

$$\frac{-\frac{1}{3}}{b_1 - \frac{1}{3}} = \frac{g - \frac{1}{3}}{g_1 - \frac{1}{3}}; g = \frac{b_1 - g_1}{3b_1 - 1} = \frac{b_1 - g_1}{2b_1 - r_1 - g_1}; r = 1 - g = \frac{b_1 - r_1}{2b_1 - r_1 - g_1}. \quad (3.31)$$

Из (3.31) координаты точки  $A$  задаются следующими значениями:

$$\begin{aligned} b_A &= 0 \\ r_A &= (b_1 - r_1) / (2b_1 - r_1 - g_1) \\ g_A &= (b_1 - g_1) / (2b_1 - r_1 - g_1) \end{aligned} \quad (3.32)$$

$$|WA| = \sqrt{\left(-\frac{1}{3}\right)^2 + \left(\frac{b_1 - r_1}{2b_1 - r_1 - g_1} - \frac{1}{3}\right)^2 + \left(\frac{b_1 - g_1}{2b_1 - r_1 - g_1} - \frac{1}{3}\right)^2} = \\ = \frac{\sqrt{(1-3b_1)^2 + (1-3r_1)^2 + (1-3g_1)^2}}{3|2b_1 - r_1 - g_1|}. \quad (3.33)$$

$$|WP'| = \sqrt{\left(r_1 - \frac{1}{3}\right)^2 + \left(g_1 - \frac{1}{3}\right)^2 + \left(b_1 - \frac{1}{3}\right)^2} = \\ = \frac{1}{3}\sqrt{(3r_1 - 1)^2 + (3g_1 - 1)^2 + (3b_1 - 1)^2}. \quad (3.34)$$

Насыщенность, задаваемая как отношение модулей  $WP'$  и  $WA$ , вычисляется делением (3.34) на (3.33):

$$S = \frac{|WP'|}{|WA|} = |2b_1 - r_1 - g_1| = |3b_1 - 1| = |1 - 3b_1|. \quad (3.35)$$

Для определения тона необходимо вычислить угол между прямой  $WR$  и прямой  $WP'$ :

$$\cos\phi = \frac{(r_{P'} - r_W)(r_R - r_W) + (g_{P'} - g_W)(g_R - g_W) + (b_{P'} - b_W)(b_R - b_W)}{|WR| \times |WP'|} \\ = \frac{\left(r_1 - \frac{1}{3}\right)\left(1 - \frac{1}{3}\right) + \left(g_1 - \frac{1}{3}\right)\left(-\frac{1}{3}\right) + \left(b_1 - \frac{1}{3}\right)\left(-\frac{1}{3}\right)}{\sqrt{\left(r_1 - \frac{1}{3}\right)^2 + \left(g_1 - \frac{1}{3}\right)^2 + \left(b_1 - \frac{1}{3}\right)^2} \times \sqrt{\left(\frac{2}{3}\right)^2 + \left(\frac{1}{3}\right)^2 + \left(\frac{1}{3}\right)^2}} = \\ = \frac{2r_1 - g_1 - b_1}{\sqrt{6\left[\left(r_1 - \frac{1}{3}\right)^2 + \left(g_1 - \frac{1}{3}\right)^2 + \left(b_1 - \frac{1}{3}\right)^2\right]}}. \quad (3.36)$$

Для сектора II  $\theta_0 = 120^\circ$ , производя вычисления, аналогичные выполненным для сектора I, с учетом того, что точка  $A$  определяется как точка пересечения прямых  $WP'$  и  $GB$ , а тон задается углом между прямыми  $WP'$  и  $GB$  плюс начальное смещение  $\theta_0 = 120^\circ$ , получим следующие выражения:

$$S = |2r_1 - g_1 - b_1|, \quad (3.37)$$

$$\cos\phi = \frac{2g_1 - r_1 - b_1}{\sqrt{6\left[\left(r_1 - \frac{1}{3}\right)^2 + \left(g_1 - \frac{1}{3}\right)^2 + \left(b_1 - \frac{1}{3}\right)^2\right]}}. \quad (3.38)$$

Для сектора III  $\theta_0 = 240^\circ$ , производя вычисления, аналогичные выполненным для сектора I, с учетом того, что точка  $A$  определяется как точка пересечения прямых  $WP'$  и  $BR$ , а тон задается углом между прямыми  $WP'$  и  $BR$  плюс начальное смещение  $\theta_0 = 240^\circ$ , получим следующие выражения:

$$S = |2g_1 - r_1 - b_1|, \quad (3.39)$$

$$\cos\phi = \frac{2b_1 - r_1 - g_1}{\sqrt{6\left[\left(r_1 - \frac{1}{3}\right)^2 + \left(g_1 - \frac{1}{3}\right)^2 + \left(b_1 - \frac{1}{3}\right)^2\right]}}. \quad (3.40)$$

Обобщая (3.35)÷(3.40), можно записать

$$H = \theta_0 + \arccos\left[N\sqrt{6\left(\left(r - \frac{1}{3}\right)^2 + \left(g - \frac{1}{3}\right)^2 + \left(b - \frac{1}{3}\right)^2\right)}\right], \quad (3.41)$$

где  $N=2r-g-b$ ,  $2g-b-r$ ,  $2b-r-g$  и  $\theta_0 = 0^\circ$ ,  $120^\circ$ ,  $240^\circ$  в секторах I, II, III соответственно.

$$S=1-3\min(r,g,b). \quad (3.42)$$

Обратное преобразование, так же как и прямое, будем выполнять для каждого сектора отдельно.

В секторе I при  $\theta_0 = 0^\circ$  исходные данные представлены системой (3.43).

$$\begin{cases} r + g + b = 1 \\ H = \arccos\left[(3r - 1)\sqrt{6\left(\left(r - \frac{1}{3}\right)^2 + \left(g - \frac{1}{3}\right)^2 + \left(b - \frac{1}{3}\right)^2\right)}\right] \\ S = |2b - r - g| = |3b - 1| = 1 - 3b \end{cases} \quad (3.43)$$

Обозначим

$$x = 3r - 1. \quad (3.44)$$

$$b = -\frac{S}{3} + \frac{1}{3},$$

$$g = 1 - r - b = \frac{2}{3} - r + \frac{S}{3}. \quad (3.45)$$

Отсюда  $3g - 1 = S - x$ ,  $3b - 1 = -S$ .

$$\cos\phi = x\sqrt{\frac{6}{9}(x^2 + (S - x)^2 + S^2)}. \quad (3.46)$$

Возведя правую и левую части уравнения (3.46) в квадрат, получим:

$$\frac{4}{3}\cos^2\phi\left(x^2 + S^2 - Sx\right) = x^2. \quad (3.47)$$

Корни уравнения (3.47):

$$x_{1,2} = 2S \frac{\cos^2 \phi \pm \sqrt{3} |\cos \phi| \sin \phi}{\cos^2 \phi - 3 \sin^2 \phi}. \quad (3.48)$$

При  $\phi \in \left[0, \frac{\pi}{2}\right]$  из (3.48) получим:

$$x_1 = S \frac{\cos \phi}{\cos \phi + \cos(120^\circ - \phi)} = S \frac{\cos \phi}{\cos(60^\circ - \phi)}, \quad (3.49 \text{ а})$$

$$x_2 = -S \frac{\cos \phi}{\cos(120^\circ - \phi)}. \quad (3.49 \text{ б})$$

Такую же пару значений  $x$  мы получим при раскрытии модуля в случае отрицательных значений косинуса при  $\phi \in \left[\frac{\pi}{2}, \frac{2\pi}{3}\right]$ , только  $x_1$  и  $x_2$  при этом поменяются местами. Поскольку область изменения функции, задаваемой уравнением (3.49 б), не удовлетворяет геометрическому смыслу задачи, то следует этот корень считать посторонним. Решением является корень  $x_1$ .

В соответствии с (3.49 а) из (3.44) выражение для  $r$  имеет вид:

$$r = \frac{1}{3} + \frac{S \cos \phi}{3 [\cos(60^\circ - \phi)]}. \quad (3.50)$$

В соответствии с (3.49 а) из (3.45) выражение для  $g$  имеет вид:

$$g = \frac{1+S}{3} - \frac{S \cos \phi}{3 [\cos(60^\circ - \phi)]}. \quad (3.51)$$

Выполнив аналогичные вычисления для секторов II, в котором  $\theta_0 = 120^\circ$ , и III, в котором  $\theta_0 = 240^\circ$ , и обобщив полученные решения, получим следующие уравнения для обратного преобразования из пространства HLS в пространство RGB:

$$\begin{cases} x_1 = (1-S)/3; \\ x_2 = S \cos \phi / [3 \cos(60^\circ - \phi)] + 1/3, \\ x_3 = 1 - (x_1 + x_2) \end{cases} \quad (3.52)$$

где  $\phi = H$ ,  $r = x_2$ ,  $g = x_3$ ,  $b = x_1$ , при  $H \leq 120^\circ$ ;

иначе если  $H \leq 240^\circ$ , то  $\phi = H - 120^\circ$ ,  $r = x_1$ ,  $g = x_2$ ,  $b = x_3$ ,

иначе если  $H > 240^\circ$ , то  $\phi = H - 240^\circ$ ,  $r = x_3$ ,  $g = x_1$ ,  $b = x_2$ .

$$R = 3Lr; G = 3Lg; B = 3Lb. \quad (3.53)$$

Преобразование пространства RGB в пространство HLS выполняется в соответствии с уравнениями (3.52), (3.53).

### 3.8 Цветовая модель $L^*u^*v^*$ МКО 1973

Эта модель равноконтрастного цветового пространства отличается от пространства  $L^*a^*b^*$  цветовыми координатами  $u^*v^*$  (яркости в этих пространствах совпадают). Если  $(a^*b^*)$  являются нелинейным

преобразованием  $(X, Y)$  МКО, то  $(u^*v^*)$  связаны с  $(X, Y)$  линейным преобразованием.

$$\begin{aligned} L^* &= 25(100Y/Y_0)^{1/3} - 16,1 \quad Y \leq 100 \\ u^* &= 13L^*(u' - u'_0) \\ v^* &= 13L^*(v' - v'_0) \end{aligned} \quad , \quad (3.54)$$

где

$$\begin{aligned} u' &= \frac{4X}{X + 15Y + 3Z}; \quad v' = \frac{9Y}{X + 15Y + 3Z}; \\ u'_0 &= \frac{4X_0}{X_0 + 15Y_0 + 3Z_0}; \quad v'_0 = \frac{9Y_0}{X_0 + 15Y_0 + 3Z_0}. \end{aligned} \quad (3.55)$$

### 3.9 Метрическое векторное цветовое пространство

На оппонентной теории цвета сегодня разработан новый подход к представлению цветов в линейных векторных пространствах [24,25]. В отличие от аффинного пространства, где отсутствуют определения угла и расстояния, полученное пространство является метрическим векторным цветовым пространством. Оно образуется откликами зрительной системы, модули векторов которых равны яркостным амплитудам. Реальные цвета образуют конус в метрическом векторном пространстве, положение векторов в котором зависит от соотношения яркостных амплитуд.

В цветовой фотометрии определяются основные величины для количественного описания цветов: амплитуда цвета, цветовая яркость, насыщенность и цветовой тон [26].

Цветовая амплитуда равна векторной сумме амплитуд основных откликов зрительной системы:  $\mathbf{A}_c = \mathbf{A}_r + \mathbf{A}_g + \mathbf{A}_b$  и может быть выражена через модули цветовых векторов в скалярном виде:

$$A_c = \sqrt{A_r^2 + A_g^2 + A_b^2 + 2A_r A_g \cos \alpha_{rg} + 2A_g A_b \cos \alpha_{gb} + 2A_b A_r \cos \alpha_{br}}, \quad (3.56)$$

где  $\alpha_{rg}$ ,  $\alpha_{gb}$ ,  $\alpha_{br}$  - углы между векторами откликов в метрическом векторном пространстве (в соответствии с рисунком 3.7). Цветовая яркость  $L_c = A_c^2$  может быть выражена через цветовые амплитуды и яркости  $L_r$ ,  $L_g$ ,  $L_b$  в соответствии с уравнением:

$$L_c = L_r + L_g + L_b + 2A_r A_g \cos \alpha_{rg} + 2A_g A_b \cos \alpha_{gb} + 2A_b A_r \cos \alpha_{br}. \quad (3.57)$$

Белому равнозернистому цвету в метрическом векторном цветовом пространстве соответствует вектор белого  $\mathbf{A}_W$ . Плоскость цветности перпендикулярна вектору равнозернистого белого  $\mathbf{A}_W$ .

Насыщенность цвета в метрическом векторном пространстве количественно определяется косинусом угла между вектором данного цвета и вектором равнозернистого белого цвета:

$$S = \cos \theta. \quad (3.58)$$

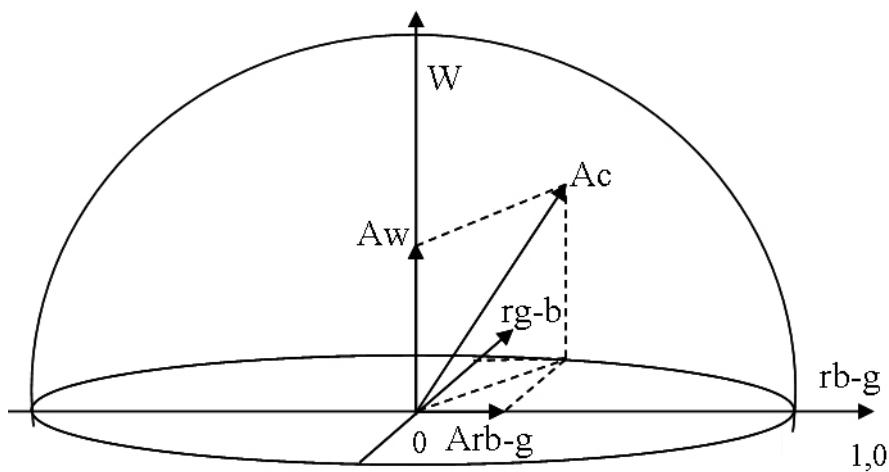


Рисунок 3.7 Представление цветов в метрическом векторном пространстве.

Цветовой тон определяется углом  $\varphi$  в плоскости цветности между составляющей вектора цвета в плоскости цветности  $A_{цс}$  и одним из векторов оппонентных сигналов цветности, например, вектором  $A_{rb-g}$ . Численно цветовой тон определяется выражением:

$$\varphi = (-1)^k \arccos(A_{rb-g}/A_c) + n\pi, \quad (3.59)$$

где постоянные  $n$  и  $k$  зависят от знака оппонентных сигналов  $rb-g$  и  $rg-b$ .

Ожидается, что построение ТВ систем на принципе постоянной цветовой яркости позволит приблизить воспроизведение цветных изображений к фотометрически точным условиям цветопередачи, повысить четкость на цветных деталях изображений, в том числе в насыщенных цветах, приблизив ее к четкости черно-белых изображений. Эта теория находится в развитии. Проблема состоит в определении основных величин. В колориметрии цвет численно определяется через функции относительного спектрального распределения реакций глаза, обусловленных работой колбочек, в неметрическом векторном пространстве. Предлагаемое метрическое векторное цветовое пространство позволит определить цветовую фотометрию, основанную на откликах цветовых зрительных каналов. Это пространство равномерно и в нем определены скалярные умножения. Основным законом цветовой фотометрии является аддитивность векторов. Предложена модель цветового зрения. Показано, что цветовые отклики могут быть описаны векторами в метрическом пространстве. В соответствии с предложенной моделью, основными величинами цветовой фотометрии являются амплитуды цветовых реакций и углы между векторами цветов. Гетерохромная яркость зависит от амплитуды хроматического отклика, а тон и насыщенность цвета зависят от углов.

#### 4 ПРЕДСТАВЛЕНИЕ ИЗОБРАЖЕНИЙ В ЦИФРОВОЙ ФОРМЕ

Для осуществления цифровой обработки изображений необходимо выполнить преобразование непрерывного (аналогового) сигнала

изображений в цифровой массив. Такое преобразование включает выполнение двух преобразований. Первое преобразование представляет замену реального непрерывного изображения набором отсчетов в дискретные моменты времени, такое преобразование называется дискретизацией. Второе - это преобразование непрерывного множества значений сигнала изображения в множество квантованных значений, такое преобразование называется квантованием.

#### 4.1 Дискретизация изображений

Пространственная дискретизация изображения предполагает формирование непрерывной функции в дискретных отсчетах пространственных координат. Пусть функция  $f_I(x, y)$  описывает исходное непрерывное изображение бесконечных размеров. В идеальной системе дискретизации изображения пространственные отсчеты исходного изображения получаются путем перемножения этой функции с пространственно дискретизирующей функцией, состоящей из бесконечного числа  $\delta$ -функций Дирака, заданных в узлах решетки с шагом  $(\Delta x, \Delta y)$  [12]:

$$s(x, y) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \delta(x - n\Delta x, y - m\Delta y), \quad (4.1)$$

где  $m, n \in \mathbb{Z}$ ,  $\mathbb{Z}$ - множество целых чисел.

На рисунке 4.1 представлен набор бесконечного числа дельта-функций в пространственной области (плоскости XY-координат) с шагом выборки  $\Delta x$ ,  $\Delta y$ .

Значения дельта-функций Дирака  $\delta(x, y)$  не равны нулю только при значениях аргументов, равных нулю:  $(x - n\Delta x) = 0$  и  $(y - m\Delta y) = 0$ . Произведения непрерывной функции и дельта функции имеют значения, равные:

$$f_I(x, y) \delta(x - n\Delta x, y - m\Delta y) = f_I(n\Delta x, m\Delta y) \delta(x - n\Delta x, y - m\Delta y), \quad (4.2)$$

а интеграл от этого произведения в бесконечных пределах равен значению функции в отсчете:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_I(n\Delta x, m\Delta y) \delta(x - n\Delta x, y - m\Delta y) dx dy = f_I(n\Delta x, m\Delta y).$$

Умножим непрерывное изображение на функцию пространственной дискретизации (4.1):

$$f(x, y) = f_I(x, y) s(x, y). \quad (4.3)$$

В соответствии с (4.2):

$$f(x, y) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} f_I(n\Delta x, m\Delta y) \delta(x - n\Delta x, y - m\Delta y). \quad (4.4)$$

Для того чтобы рассмотреть спектр дискретизованного изображения, обратимся к теореме о свертке.

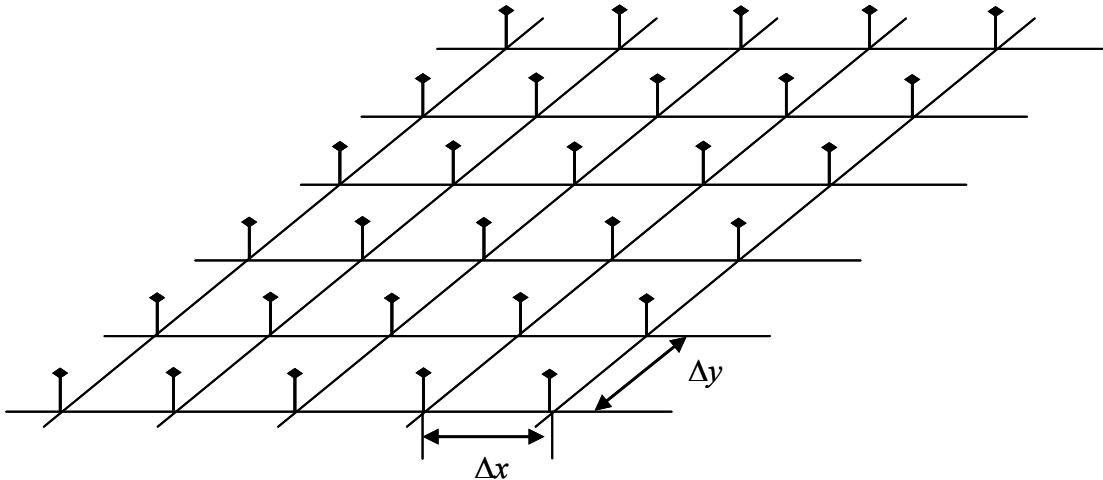


Рисунок 4.1 Функция пространственной дискретизации изображения.

Сверткой двух функций называется функция, определяемая соотношением:

$$f(x) = \int_{-\infty}^{\infty} f_1(\xi) f_2(x - \xi) d\xi. \quad (4.5)$$

Прямое преобразование Фурье от этой функции:

$$F(\omega) = \int_{-\infty}^{\infty} \left( \int_{-\infty}^{\infty} f_1(\xi) f_2(x - \xi) d\xi \right) \exp(-i\omega x) dx.$$

Изменим порядок интегрирования по теореме Фубини [27]:

$$F(\omega) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_1(\xi) f_2(x - \xi) \exp(-i\omega x) d\xi dx.$$

Выполним замену переменной  $x - \xi = \eta$ ; при этом  $dx = d\eta$ :

$$\begin{aligned} F(\omega) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_1(\xi) f_2(\eta) \exp(-i\omega(\xi + \eta)) d\xi d\eta = \\ &= \int_{-\infty}^{\infty} f_1(\xi) \exp(-i\omega\xi) d\xi \int_{-\infty}^{\infty} f_2(\eta) \exp(-i\omega\eta) d\eta = F_1(\omega) F_2(\omega). \end{aligned} \quad (4.6)$$

Из (4.6) следует, что спектр функции, полученной в результате свертки двух функций, равен произведению спектров этих функций. (Произведение спектров вычисляется их поэлементным умножением). И, обратно, обратное преобразование Фурье от произведения спектров двух функций равно свертке этих двух функций. И, наоборот, свертка спектров в частотной области приводит к умножению в пространственной области. Верно и обратное утверждение: умножение функций в пространственной области приводит к свертке их спектров в частотной области [28].

Для двумерного случая имеют место аналогичные соотношения по теореме о свертке:

$$f_1(x, y) * f_2(x, y) \Leftrightarrow F_1(\omega_x, \omega_y) F_2(\omega_x, \omega_y),$$

$$F_1(\omega_x, \omega_y) * F_2(\omega_x, \omega_y) \Leftrightarrow f_1(x, y)f_2(x, y). \quad (4.7)$$

В соответствии с теоремой о свертке, спектр дискретного изображения равен свертке спектров непрерывной функции изображения и спектра дискретизирующей функции:

$$F(\omega_x, \omega_y) = \frac{1}{4\pi^2} F_I(\omega_x, \omega_y) * S(\omega_x, \omega_y). \quad (4.8)$$

Преобразование Фурье от дельта-функции  $\delta(x - n\Delta x, y - m\Delta y)$  равно  $\delta(\omega_x - n\omega_{xs}, \omega_y - m\omega_{ys})$ . Преобразование Фурье от дискретизирующей функции (4.1)

$$S(\omega_x, \omega_y) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \exp(-i(n\Delta x\omega_x + m\Delta y\omega_y)). \quad (4.9)$$

По теореме Пуассона [27]:

$$\sum_{j=-\infty}^{\infty} \exp(-ij\Delta_x\omega_x) = (2\pi/\Delta x) \sum_{n=-\infty}^{\infty} \delta(\omega_x - 2\pi n/\Delta x). \quad (4.10)$$

В соответствии с (4.10) из (4.9) получим спектр дискретизирующей функции:

$$S(\omega_x, \omega_y) = (4\pi^2/(\Delta x\Delta y)) \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \delta(\omega_x - n\omega_{xs}, \omega_y - m\omega_{ys}), \quad (4.11)$$

где  $\omega_{xs} = 2\pi/\Delta x$ ,  $\omega_{ys} = 2\pi/\Delta y$ .

Предположим, что спектр исходного непрерывного изображения ограничен по ширине так, что

$$F_I(\omega_x, \omega_y) = 0 \text{ при } \begin{cases} |\omega_x| > \omega_{xc} \\ |\omega_y| > \omega_{yc} \end{cases}.$$

Вычисляя свертку согласно (4.8) найдем

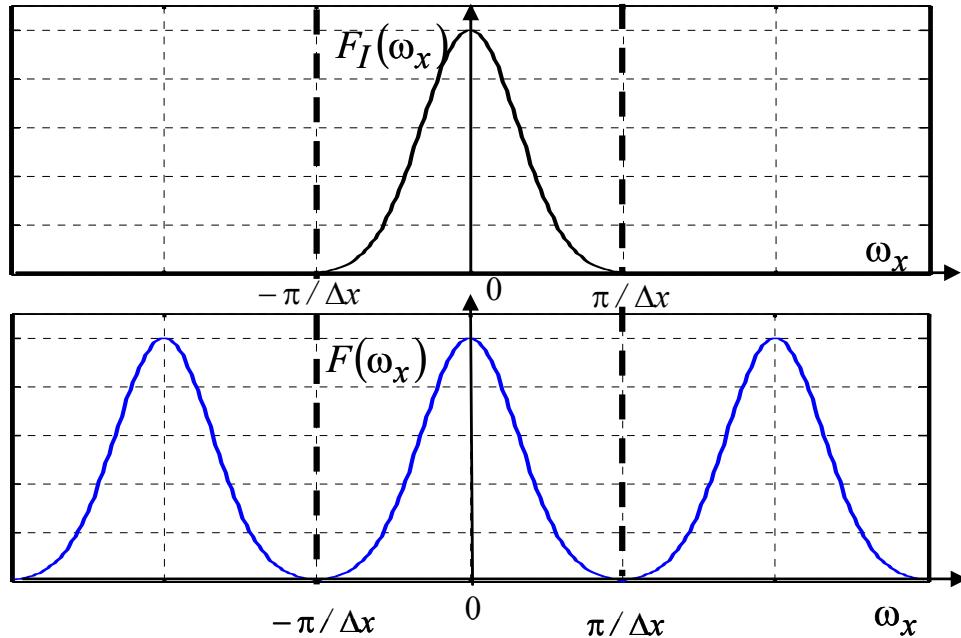
$$F(\omega_x, \omega_y) = \frac{1}{\Delta x \Delta y} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F_I(\omega_x - \alpha, \omega_y - \beta) \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \delta(\alpha - n\omega_{xs}, \beta - m\omega_{ys}) d\alpha d\beta.$$

Меняя порядок операций суммирования и интегрирования и учитывая основное свойство  $\delta$ -функции, получаем выражение для спектра дискретизованного изображения:

$$F(\omega_x, \omega_y) = \frac{1}{\Delta x \Delta y} \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} F_I(\omega_x - n\omega_{xs}, \omega_y - m\omega_{ys}). \quad (4.12)$$

Спектр дискретизированного изображения получается путем бесконечного повторения спектра исходного изображения со сдвигом на величины, кратные  $(2\pi/\Delta x, 2\pi/\Delta y)$ . Повторение спектра для сечения по строке показано на рисунке 4.2. Следует отметить, что при выборе  $\Delta x$  и  $\Delta y$  слишком большими, соседние спектры будут перекрываться друг с другом.

Обратная операция, позволяющая из цифрового массива получить непрерывное изображение, называется восстановлением непрерывных изображений. Из отсчетов функции  $f(x,y)$  можно получить непрерывное изображение путем линейной пространственной интерполяции или с помощью линейной пространственной фильтрации дискретизированного изображения. Пусть  $r(x,y)$  есть импульсный отклик интерполирующего фильтра, а  $R(\omega_x, \omega_y)$ - его частотная характеристика. Восстановленное изображение получается как свертка последовательности отсчетов с импульсным откликом восстанавливающего фильтра. Таким образом, восстановленное непрерывное изображение описывается соотношением  $f_R(x,y) = f(x,y) * r(x,y)$ . (4.13)



a)b)

Рисунок 4.2 а) Изменение спектра по строке для а) непрерывного сигнала, б) дискретизованного сигнала.

Подставляя  $f(x,y)$  из (4.4) и вычисляя свертку (4.13), получаем

$$f_R(x,y) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} f_I(n\Delta x, m\Delta y) r(x - n\Delta x, y - m\Delta y). \quad (4.14)$$

Отсюда видно, что импульсный отклик  $r(x,y)$  является двумерной функцией, интерполирующей отсчеты на всю плоскость.

Пространственно-частотный спектр изображения, восстановленного согласно равенству (4.14), есть произведение частотной характеристики восстанавливающего фильтра и спектра дискретизированного изображения, то есть

$$F_R(\omega_x, \omega_y) = F(\omega_x, \omega_y) R(\omega_x, \omega_y) = \frac{1}{\Delta x \Delta y} R(\omega_x, \omega_y) \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} F_I(\omega_x - n\omega_{xs}, \omega_y - m\omega_{ys}). \quad (4.15)$$

Из этого выражения видно, что спектры не должны перекрываться, а восстанавливающий фильтр  $R(\omega_x, \omega_y)$  должен пропускать без искажений основной спектр при  $n=0$  и  $m=0$  и полностью подавлять все побочные спектры при  $n, m \neq 0$ , чтобы спектр восстановленного непрерывного изображения совпадал со спектром исходного изображения. Только в этом случае исходное и восстановленное изображения будут одинаковыми. Для изображений с ограниченной шириной спектра первое условие выполняется, если интервал дискретизации выбран так, что прямоугольная область, ограниченная верхними граничными частотами спектра изображения  $(\omega_{xc}, \omega_{yc})$  лежит внутри прямоугольной области, определяемой половинами частот дискретизации  $\omega_{xs}/2, \omega_{ys}/2$  (в соответствии с рисунком 4.3). Следовательно, должны выполняться неравенства:

$$\omega_{xc} \leq \frac{\omega_{xs}}{2}, \quad \omega_{yc} \leq \frac{\omega_{ys}}{2} \text{ или } \Delta x \leq \frac{\pi}{\omega_{xc}}, \quad \Delta y \leq \frac{\pi}{\omega_{yc}}. \quad (4.16)$$

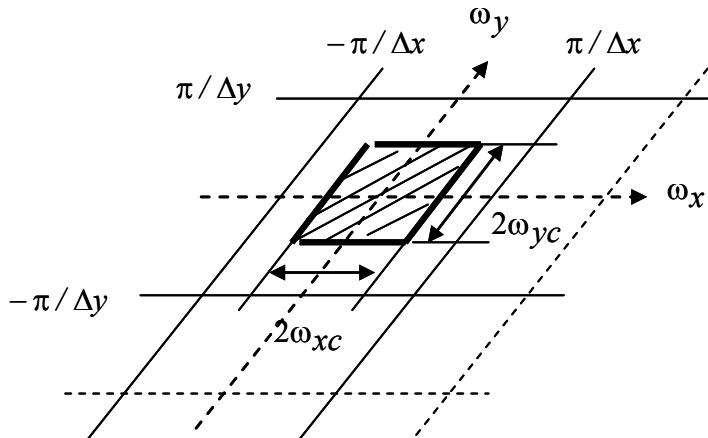


Рисунок 4.3 Выбор частоты дискретизации в соответствии с теоремой отсчетов.

Это означает, что шаг дискретизации не должен превышать половины периода пространственной гармоники, соответствующей самым мелким деталям изображения. Если  $\omega_{xc} = \frac{\omega_{xs}}{2}$  и  $\omega_{yc} = \frac{\omega_{ys}}{2}$ , то выборка осуществляется с частотой Котельникова, вдвое превышающей наивысшую частоту спектра исходного изображения. В тех случаях, когда пространственная частота дискретизации выбрана в соответствии с теоремой Котельникова, исходное изображение можно точно восстановить путем пространственной фильтрации отсчетов с помощью соответствующего фильтра. Так, например, фильтр, частотная характеристика которого имеет вид прямоугольного параллелепипеда (в соответствии с рисунком 4.4) и описывается выражением:

$$R(\omega_x, \omega_y) = \begin{cases} K, & \text{если } (\omega_x | \leq \omega_{xL}) \text{ и } (\omega_y | \leq \omega_{yL}) \\ 0, & \text{иначе} \end{cases}, \quad (4.17)$$

где  $K$ - масштабирующая постоянная,

удовлетворяет условию точного восстановления, если  $\omega_{xL} > \omega_{xc}$  и  $\omega_{yL} > \omega_{yc}$ .

Функция рассеяния точки, или импульсный отклик, данного восстанавливающего фильтра имеет вид [12]:

$$r(x,y) = \frac{K\omega_{xL}\omega_{yL}}{\pi^2} \frac{\sin(\omega_{xL}x)}{\omega_{xL}x} \frac{\sin(\omega_{yL}y)}{\omega_{yL}y}. \quad (4.18)$$

При использовании этого фильтра изображение восстанавливается с помощью бесконечной суммы функций вида  $\text{sinc}(x)$ .

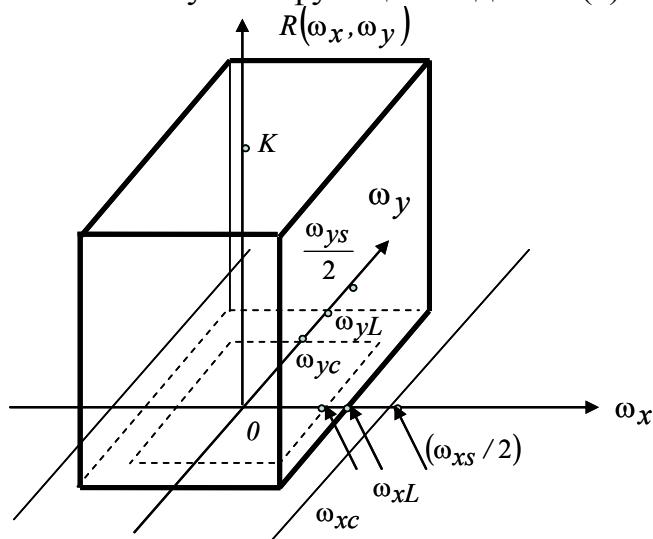


Рисунок 4.4 Частотная характеристика идеального восстанавливающего прямоугольного фильтра.

## 4.2 Квантование изображений

Для получения цифрового сигнала из непрерывного сигнала необходимо кроме дискретизации по времени произвести квантование по амплитуде [29]. Квантование состоит в том, что непрерывному по амплитуде сигналу ставится в соответствие конечное множество целочисленных значений сигнала, пропорциональных непрерывному значению.

Для этого динамический диапазон сигнала  $f = [f_{min}, f_{max}]$  разбивается на конечное число интервалов - интервалов квантования. Каждому интервалу ставится в соответствие одно значение, называемое уровнем квантования, кодируемое двоичным кодом. Все значения сигнала, попадающие в некоторый интервал, обозначаются одним числом, определенным для данного интервала.

Пусть  $f^q$  - амплитуда сигнала в дискретном представлении,  $f$  - значение сигнала, принадлежащее  $q$ -му интервалу квантования. При использовании  $L$ -разрядного кода число уровней квантования равно  $2^L$ . Ошибка квантования (шум квантования) может характеризоваться величиной:

$$\varepsilon = f - f^q.$$

Дисперсия ошибки  $\sigma_q^2$  отличия  $f$  от его квантованного представления  $f^q$ :

$$\sigma_q^2 = \int_{f_q}^{f_{q+1}} \varepsilon^2 p(f) df, \quad (4.19)$$

где  $f_q, f_{q+1}$  - границы  $q$ -го интервала квантования,  $p(f)$ - плотность вероятности распределения входного сигнала.

#### 4.2.1 Оптимальное квантование

Оптимальным квантованием будем считать такой выбор интервалов квантования и значений их представителей, при котором  $\sigma_q^2$  минимальна. Выбор дисперсии в качестве критерия обусловлен такими достоинствами этой меры, как универсальность; простота расчетов и построения алгоритмов; высокая коррелированность с субъективными показателями качества.

Пусть плотность вероятности значений исходного сигнала постоянна в пределах интервала квантования, тогда

$$\begin{aligned} \sigma_q^2 &= p_q \int_{f_q}^{f_{q+1}} (f - f^q)^2 df = p_q \frac{(f - f^q)^3}{3} \Big|_{f_q}^{f_{q+1}} \\ &= \frac{p_q}{3} \left[ (f_{q+1} - f^q)^3 - (f_q - f^q)^3 \right]. \end{aligned} \quad (4.20)$$

Оптимальное положение уровня квантования  $f^q$  в интервале  $[f_q, f_{q+1}]$  можно найти, решая задачу о минимуме ошибки как функции от  $f^q$ . Приравнивая нулю производную от  $\sigma_q^2$  по  $f^q$

$$\partial \sigma_q^2 / \partial f^q = 0,$$

получаем

$$f^q = (f_{q+1} + f_q) / 2. \quad (4.21)$$

Из (4.21) оптимальное значение уровня квантования соответствует середине интервала квантования, при этом максимальная ошибка квантования внутри интервала составляет не более половины интервала квантования.

Подставив выражения (4.21) в (4.20), получим

$$\sigma_q^2 = \frac{p_q}{12} (f_{q+1} - f_q)^3. \quad (4.22)$$

Дисперсия ошибки квантования

$$\sigma_Q^2 = \sum_{q=1}^{2^L} \sigma_q^2 = \frac{1}{12} \sum_{q=1}^{2^L} p_q (f_{q+1} - f_q)^3. \quad (4.23)$$

В общем случае оптимальное положение пороговых уровней и уровней квантования получают из точного уравнения ошибки квантования, полученного с учетом (4.19):

$$\sigma_Q^2 = \sum_{q=1}^{2^L} \int_{f_q}^{f_{q+1}} \left( f - f^q \right)^2 p(f) df. \quad (4.24)$$

Дифференцируя  $\sigma_Q^2$  по переменным  $f_q$  и  $f^q$  и приравнивая производные нулю, получим систему уравнений:

$$\begin{cases} \partial \sigma_Q^2 / \partial f_q = 0 \\ \partial \sigma_Q^2 / \partial f^q = 0 \end{cases}$$

После преобразований, она сводится к системе уравнений:

$$\begin{cases} f^q = 2f_q - f^{q-1} & a) \\ \int_{f_q}^{f_{q+1}} fp(f) df \\ f^q = \frac{f_q}{\int_{f_q}^{f_{q+1}} p(f) df} & b) \end{cases}, \quad (4.25)$$

где  $q=1..2^L$ .

Решая эти уравнения рекуррентным способом, для заданной плотности вероятностей находят оптимальные значения пороговых уровней и уровней квантования. Макс (*J. Max*) решил такую задачу для гауссовой плотности и составил таблицы размещения пороговых уровней в зависимости от числа уровней квантования. На рисунке 4.5 представлена амплитудная характеристика квантователя Макса [12] для трехразрядного представления сигнала.

Подставив (4.25б) в (4.24), получим, что дисперсия ошибки квантования для оптимального квантователя уменьшается до значения:

$$\sigma_Q^2 = M \left\{ f^2 \right\} - \sum_{q=1}^{2^L} (f^q)^2 \int_{f_q}^{f_{q+1}} p(f) df. \quad (4.26)$$

Для частного случая равномерной плотности распределения сигнала, при которой

$$p(f) = \frac{1}{f_{\max} - f_{\min}} = \text{const},$$

оптимальные значения уровней квантования из (4.25 б) определяются в соответствии с выражением:

$$f^q = \frac{(f_{q+1} + f_q)}{2}.$$

Оптимальные значения пороговых уровней в соответствии с (4.25 а)

$$f_q = \frac{(f^q + f^{q-1})}{2}.$$

Следовательно, при равномерной плотности вероятности сигнала изображения оптимальным является равномерное квантование, при котором интервал квантования:

$$\Delta f = (f_{\max} - f_{\min}) / 2^L, \quad (4.27)$$

а плотность распределения:

$$p(f) = 1 / (2^L \Delta f).$$

Дисперсия шума квантования в этом случае из (4.26):

$$\sigma_Q^2 = (\Delta f)^2 / 12. \quad (4.28)$$

Отношение сигнала к СКО шума квантования в этом случае составляет

$$\psi_{\text{кв}} = 20 \lg \frac{(f_{\max} - f_{\min})}{\sigma_Q} = 20 \lg \frac{(f_{\max} - f_{\min}) 2^L \sqrt{12}}{(f_{\max} - f_{\min})} \approx 10.8 + 6L, \text{ дБ.} \quad (4.29)$$

Из (4.29) следует, что увеличение числа разрядов квантования на 1 приводит к повышению отношения сигнал/шум примерно на 6 дБ.

Существующие устройства осуществляют обычно равномерное квантование сигналов. Используя такие устройства, оптимальное квантование можно выполнить, если перед равномерным квантованием сигнал подвергнуть нелинейному преобразованию (предыскажению), формирующему сигнал с равномерной плотностью вероятности.

#### 4.2.2 Квантование сигнала при наличии шумов

Рассмотрим воздействие аддитивного шума на процесс квантования при равномерной амплитудной характеристике квантователя. Входной сигнал представим в виде:

$$u(t) = f(t) + n(t),$$

где  $f(t)$ - входной полезный сигнал;  $n(t)$ - аддитивный шум.

Квантование считается безошибочным, если сигнал  $u(t)$  попадает в тот же интервал квантования, что и сигнал  $f(t)$ . Если же сигнал  $u(t)$  попадает в другие интервалы квантования, то возникают дополнительные ошибки квантования, вызванные шумом. Количественную оценку влияния шума на квантование дал Фридман [30]. Построенная им кривая представлена на рисунке 4.6 как кривая 0.

При построении этой кривой предполагается, что значения сигнала равновероятны в пределах диапазона квантования. Шум нормальный с СКО  $\sigma$ .

Кривая соответствует вероятности правильного присвоения двоичного числа, соответствующего незашумленному сигналу с ошибкой  $\pm 0$  в зависимости от  $\beta$ , равного отношению шага квантования  $\Delta f$  к  $\sigma$  шума. Фульц расширил рамки анализа, определив вероятность присвоения данного двоичного числа уровню, отличающемуся на величину  $\pm n$  от правильного уровня. Кривая 0 соответствует ошибке  $\pm 0$ , кривая 1 соответствует ошибке  $\pm 1$  и т.д.

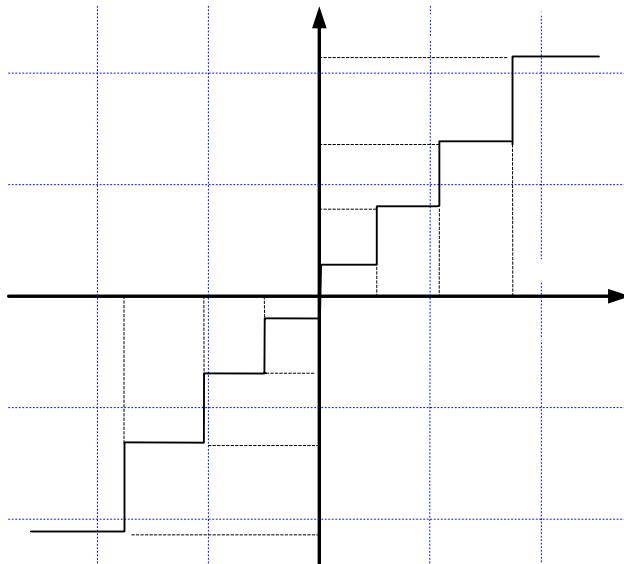


Рисунок 4.5 Амплитудная характеристика оптимального квантователя Макса

Формула интегрального распределения ошибок квантования, обусловленных шумом, имеет вид [31]:

$$P_n = 2(n+1)\Phi_u[(n+1)\beta] - 2n\Phi_u(n\beta) - 1 + \frac{1}{(\beta\sqrt{\pi/2})} \left[ \exp(-(n+1)^2\beta^2/2) - \exp(-n^2\beta^2/2) \right], \quad 1,0$$

где  $\beta = \Delta f / y$ ,

$$\Phi_u(U) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^U e^{-t^2/2} dt = \frac{1}{\sqrt{2\pi}} \int_0^U e^{-t^2/2} dt + 0,5 = \Phi(U) + 0,5, \quad (4.30)$$

-2,0	-1,0	0,0
------	------	-----

где  $\Phi(U)$  - интеграл вероятностей.

Определим понятие  $L$  полезных разрядов. Если имеется  $L$  полезных разрядов, то это означает, что в результате действия шума вероятность правильного формирования ( $L+1$ ) разряда составляет 0,5. То есть, если мы выбираем 8 полезных разрядов, то это означает, что 9-й разряд правильно не опознается, т.е. вероятность правильного прочтения этого разряда равна 0,5. Кривая 0 на рисунке 4.6 показывает, что разряду  $L+1$  (вероятность 0,5) соответствует значение  $\beta=1,47$ . Значит, при  $L$  полезных разрядах (( $L+1$ -й не нужен)  $\beta=1,47 \times 2=2,94$ , т.е. шаг квантования должен быть в 2,94 раза больше СКО шума.

-2,0

Качество работы системы можно оценивать также по уровню шума, вводимого квантователями (АЦП). Мощность шума ( $y_Q$ ), вносимого процессом квантования, определяется в соответствии с (4.28). Будем называть систему уравновешенной, если дисперсия шума квантования равна сумме дисперсий всех шумовых сигналов любого происхождения, генерируемых в системе до квантования, т.е.

$$y_{\text{до\_квантования}}^2 = \Delta f^2 / 12 \text{ или } \beta = \sqrt{12} = 3,464.$$

Таким образом, на основании двух критериев можно пользоваться общим правилом, согласно которому отношение шага квантования к СКО шума примерно равно 3.

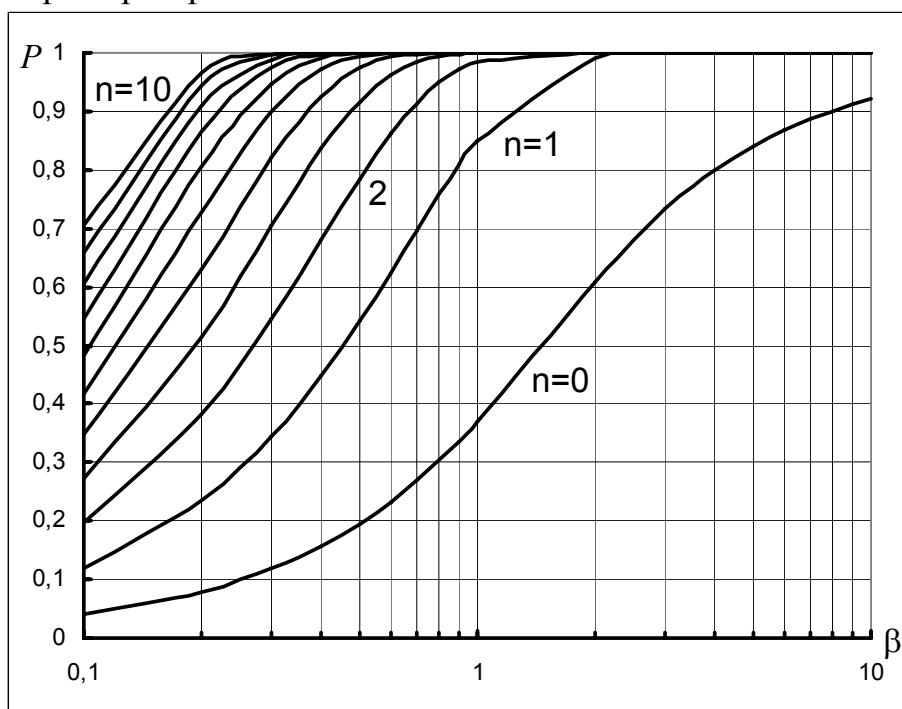


Рисунок 4.6 График зависимости вероятности квантования от отношения интервала квантования к СКО аддитивного шума  $\beta$ .

При  $n=0$  кривая соответствует вероятности безошибочного квантования при наличии шума. Ее можно использовать для выбора числа уровней квантования при заданной мощности шума и требуемой достоверности отсчета.

### 4.3 Погрешности дискретного представления изображений

Несоответствие дискретного представления сигнала изображения его физическому источнику, являющемуся непрерывной функцией пространственных координат и амплитуды сигнала, обусловлено ошибками дискретизации и квантования сигнала.

#### 4.3.1 Погрешность дискретизации

Погрешность дискретизации оценивается путем сравнения восстановленного по дискретным отсчетам непрерывного сигнала изображения с непрерывным исходным изображением. Ошибки

дискретизации обусловлены, прежде всего, шагом дискретизации и передаточной функцией восстанавливающего фильтра. В разделе 4.1 в соответствии с (4.16) задание шага дискретизации определяется шириной спектра непрерывного сигнала изображения. Шаг выборки может не соответствовать этому критерию вследствие, например, ограничений на размер массива данных. В этом случае идеальный восстанавливающий фильтр (4.18) приведет к восстановлению исходного изображения с некоторой помехой (возникновением муара, в иностранной литературе называется алиайзингом (*aliasing*)), вызванной наложением спектров [32]. Покажем это для случая одномерного сигнала. Пусть спектр непрерывной функции ограничен интервалом  $[-\pi/\Delta x, \pi/\Delta x]$ , как показано на рисунке 4.7а).

При дискретизации с шагом  $\Delta x$  происходит периодическое повторение спектра сигнала непрерывного изображения, как показано на рисунке 4.7 б). Полоса идеального низкочастотного (НЧ) фильтра соответствует спектральному диапазону исходного сигнала  $[-\pi/\Delta x, \pi/\Delta x]$  (рисунок 4.7 в). В этом случае после идеального фильтра спектр сигнала не изменится, как показано на рисунке 4.7 г). Следовательно, и обратное преобразование неискаженного спектра сигнала позволит восстановить его в точном соответствии с входным непрерывным сигналом. Однако, если полоса идеального НЧ фильтра меньше спектрального диапазона входного сигнала, то при дискретизации произойдет наложение высокочастотных частей спектра. Побочные спектры складываются с основным спектром. Формируется спектр сигнала изображения, отличный от исходного спектра. При использовании идеального восстанавливающего фильтра в полосе  $[-\pi/\Delta x, \pi/\Delta x]$  будет восстановлен сигнал, соответствующий искаженному спектру, представленному на рисунке 4.7 з). При этом уже невозможно разделить спектр помехи пространственной дискретизации и спектр исходного сигнала. Если нет возможности согласовать частоту дискретизации со спектральным диапазоном сигнала, то чтобы не допустить возникновения помехи пространственной дискретизации, перед дискретизацией необходимо ограничить спектр исходного непрерывного изображения, подвергнув его низкочастотной фильтрации, подавляющей частоты, большие  $\pi/\Delta x$ .

### 4.3.2 Погрешность квантования

В разделе 4.2 мы рассмотрели условия формирования оптимального квантователя, минимизирующего погрешности квантования. На рисунке 4.8 проиллюстрировано влияние ошибок квантования на восприятие изображения.

Возникновение ложных контуров особенно заметно на участках с плавным изменением яркости. В настоящее время чаще производится квантование изображений на 256 уровней (8 разрядов) по яркости или по каждому из *RGB* компонентов для цветных изображений. При

недостаточной разрядности иногда прибегают к наложению случайного шумового сигнала с небольшой дисперсией для уменьшения заметности ложных контуров.

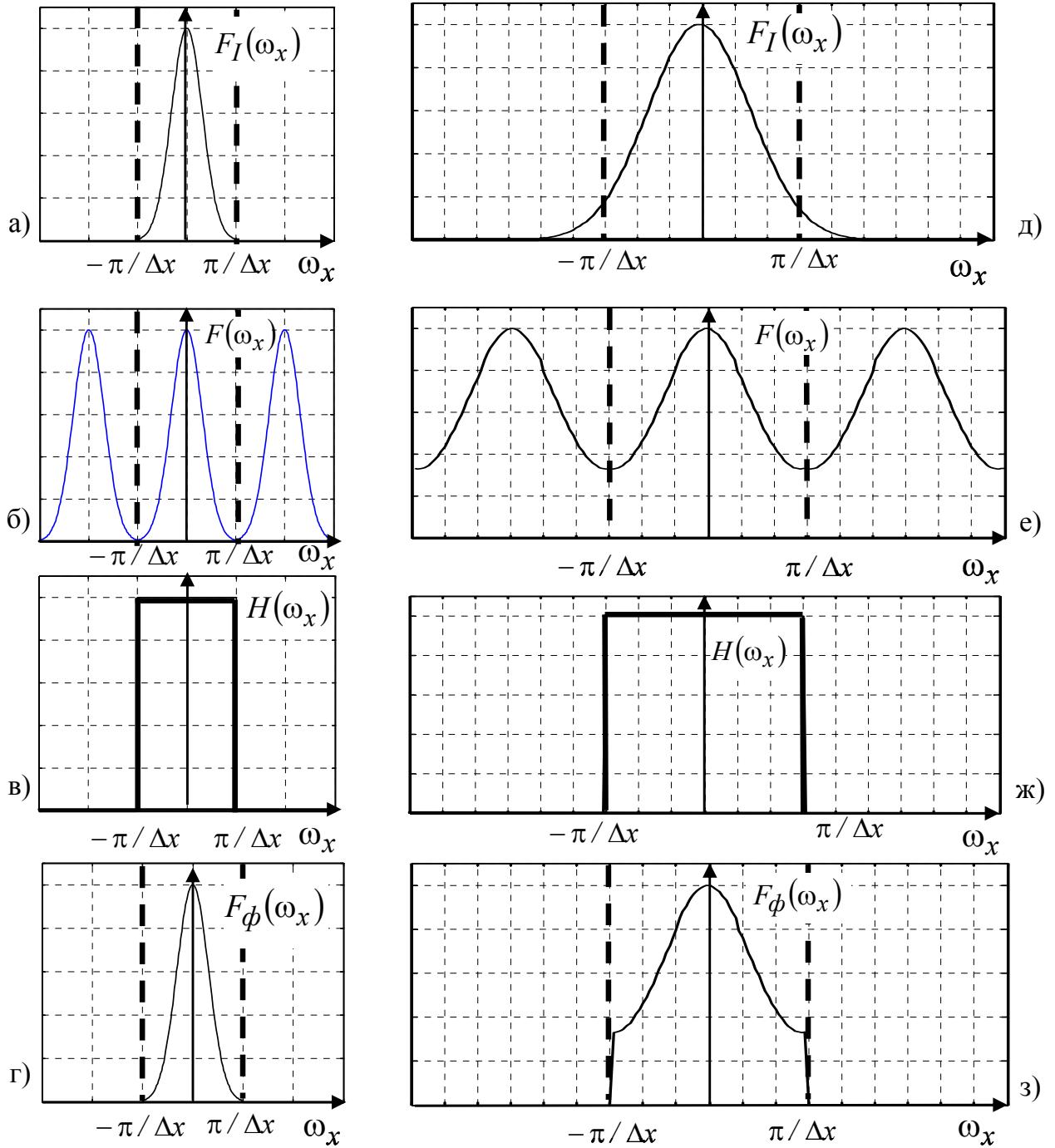


Рисунок 4.7 Идеальная НЧ фильтрация сигнала в полосе, равной полосе частотного спектра сигнала: а) спектр непрерывного сигнала; б) спектр сигнала после дискретизации; в) частотная характеристика идеального фильтра; г) спектр сигнала на выходе фильтра. Идеальная НЧ фильтрация сигнала в полосе, меньшей полосы частотного спектра сигнала: д) спектр непрерывного сигнала; е) спектр сигнала после дискретизации; ж) частотная характеристика идеального фильтра; з) спектр сигнала на выходе фильтра.

В этом случае сигналы одинаковой яркости, сложенные со случайными значениями, разрушают границы областей постоянной яркости. На рисунке 4.9 приведен пример такой обработки изображения.



а) б)

Рисунок 4.8 Равномерное квантование: а) число разрядов квантования  $L=8$ , б) число разрядов квантования  $L=4$ .



а) б)

Рисунок 4.9 а) Изображение получено равномерным квантованием на 16 уровней изображения «Лена»; б) изображение получено наложением нормального шума при СКО=5 на исходное изображение и равномерным квантованием на 16 уровней.

#### 4.4 Методы интерполяции изображений

Интерполяция предназначена для восстановления непрерывных значений амплитуды сигнала изображения по ее дискретным значениям. При масштабировании изображения необходимо заменить двумерный массив амплитуд сигнала, заданный на сетке отсчетов, некоторым другим массивом, положения отсчетов которого определяются коэффициентами масштабирования. Эта задача решается методами интерполяции. Широко применяется разделение интерполяционного преобразования на два независимых - сначала производится интерполяция по строкам, а затем по столбцам. Поэтому рассмотрим интерполяцию изображения по строкам, интерполяция по столбцам производится по аналогичным формулам в направлении столбцов изображения.

В разделе 4.1 рассмотрены условия точного восстановления сигнала при его дискретизации с частотой Котельникова, вдвое превышающей наивысшую частоту спектра исходного изображения. При этом исходный

непрерывный сигнал может быть точно восстановлен путем пространственной фильтрации отсчетов с помощью соответствующего фильтра. Частотная характеристика идеального восстанавливющего фильтра (4.17) приведена на рисунке 4.4. Функция рассеяния точки, или импульсный отклик, данного восстанавливющего фильтра определяется обратным преобразованием Фурье частотной характеристики фильтра (4.18). На рисунке 4.10 представлен график функции рассеяния точки вдоль оси абсцисс (координаты  $x$ ) для нулевого отсчета изображения. По оси абсцисс отложены номера отсчетов изображения  $i$ , соответствующие значениям  $i\Delta x$  относительно нулевого отсчета. Значение амплитуды сигнала изображения, учитывая (4.14 и 4.18), вычисляется по формуле:

$$f(x) = \sum_{i=-\infty}^{\infty} f[i\Delta x] r_x(x - i\Delta x) = \sum_{i=-\infty}^{\infty} f[i\Delta x] \frac{\sin(\pi(x - i\Delta x) / \Delta x)}{\pi(x - i\Delta x) / \Delta x}. \quad (4.31)$$

При использовании этого фильтра изображение восстанавливается с помощью бесконечной суммы произведений функции  $\text{sinc}(x)$  на соответствующие отсчеты сигнала изображения. Поскольку в каждом отсчете значения произведений равны нулю для всех отсчетов сигнала

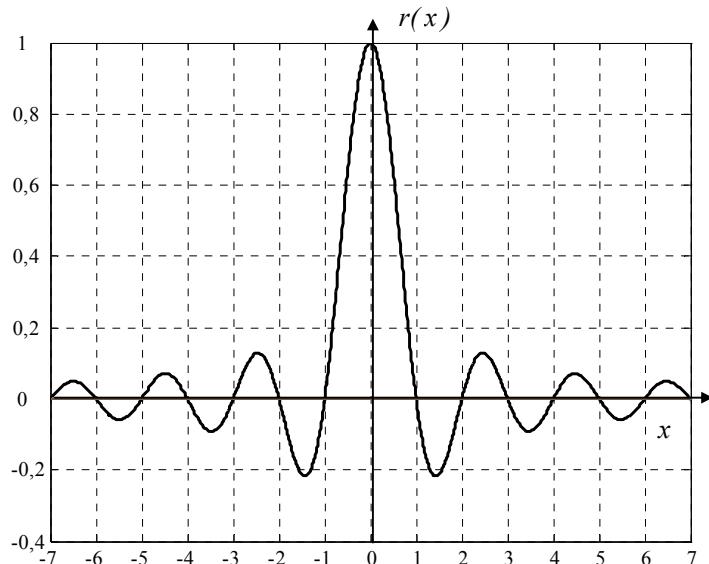


Рисунок 4.10 Импульсная характеристика идеального восстанавливющего фильтра по  $x$ -координате.

изображения, кроме отсчета с номером  $i$ , для которого  $x - i\Delta x = 0$ , а  $\text{sinc}(0)=1$ , то в положении отсчетов значения сигнала точно равны значениям сигнала исходного изображения. На интервале между отсчетами значение сигнала равно сумме взвешенных и сдвинутых  $\text{sinc}(x)$ . Функция  $\text{sinc}(x)$  сдвигается в каждое положение отсчета и масштабируется в соответствии со значением амплитуды сигнала изображения в этом отсчете.

Применение идеального восстанавливющего фильтра требует задания сигнала изображения на интервале от  $-\infty$  до  $\infty$ . Используется ограничение импульсной характеристики фильтра несколькими

интервалами  $\Delta x$  (до  $\pm 10$ ). В графических приложениях используется оконный метод, при котором значение функции  $\text{sinc}(x)$  умножается на некоторую оконную функцию. Главная задача при разработке такого фильтра - получить частотную характеристику фильтра наиболее близкую к частотной характеристике идеального НЧ фильтра. То есть фильтр должен пропускать сигнал с максимальным коэффициентом в полосе низких частот и максимально подавлять сигнал боковых полос с тем, чтобы уменьшить артефакты, вызванные наложением спектров. Одним из фильтров, удовлетворяющих этим требованиям, является фильтр Ланкцоса (Lanczos). Функции импульсной характеристики такого фильтра определяются в соответствии с формулами:

$$\begin{aligned} \text{Lanczos2}(x) &= \begin{cases} \frac{\sin(\pi x)}{\pi x} \frac{\sin(\pi x / 2)}{\pi x / 2}, & |x| < 2; \\ 0, & |x| \geq 2 \end{cases} \\ \text{Lanczos3}(x) &= \begin{cases} \frac{\sin(\pi x)}{\pi x} \frac{\sin(\pi x / 3)}{\pi x / 3}, & |x| < 3 \\ 0, & |x| \geq 3 \end{cases} \end{aligned}$$

Фильтры Ланкцоса используются в таких программах как *VirtualDub*, *IrfanView* и др. При заданном коэффициенте масштабирования могут применяться алгоритмы, позволяющие ускорить выполнение фильтрации за счет подбора коэффициентов аппроксимации оконной функции, кратных степени двойки. В этом случае операции плавающей арифметики заменяются операциями сдвигов и целочисленных умножений. Например, при коэффициенте масштабирования, равном  $1/2$ , применяются дециматоры Турковского (*Turkowski*) (коэффициенты равны  $-1/32$ ,  $9/32$ ,  $16/32$ ) или Габриэля (*Gabriel*) (коэффициенты равны  $-1/16$ ,  $5/16$ ,  $8/16$ ).

На практике применяются более простые методы интерполяции [33,34]. Наиболее распространенным видом интерполяции является полиномиальная интерполяция. При интерполяции полиномами нулевой степени значение  $f(x)$  определяется значением функции в ближайшем отсчете.

При линейной интерполяции значение функции интерполируется полиномом первой степени  $\varphi(x)$

$$\varphi_i(x) = f[i] + (f[i+1] - f[i]) / (x[i+1] - x[i])(x - x[i]). \quad (4.32)$$

Функция должна быть непрерывной:

$$\varphi_i(x[i+1]) = \varphi_{i+1}(x[i+1]), i=0,1,\dots,N-2. \quad (4.33)$$

Чтобы убедиться в этом, достаточно подставить в (4.32)  $x=x[i+1]$  для вычисления  $\varphi_i(x)$  и  $\varphi_{i+1}(x)$ . Если положить  $\Delta x = x[i+1] - x[i] = 1$ , то

$$\varphi_i(x) = (1-x)f[i] + xf[i+1], \quad i\Delta x \leq x \leq (i+1)\Delta x.$$

Использование полиномов низкой степени позволяет избежать многочисленных нестабильностей, возникающих при применении полиномов высоких степеней [35]. Но они не являются гладкими кривыми.

Полином третьей степени, называемый кубическим сплайном, описывает кривую наименшей степени, имеющую точку перегиба и имеющую возможность изгибаться [36]. Обозначим  $f[i] = y_i$ . Глобальный кубический сплайн должен удовлетворять следующим условиям:

$$\varphi(x_i) = y_i, i=0, \dots, N-1. \quad (4.34)$$

$$\varphi_i(x) = \begin{cases} a_i + b_i(x-x_i) + c_i(x-x_i)^2 + d_i(x-x_i)^3, & x \in [x_i; x_{i+1}] \\ 0, & x \notin [x_i; x_{i+1}] \end{cases}. \quad (4.35)$$

$$\varphi = \sum_{i=0}^{N-2} \varphi_i. \quad (4.36)$$

На каждом  $i$ -том отрезке  $[x_i, x_{i+1}]$  коэффициенты полинома  $a_i, b_i, c_i, d_i$  разные. Условия непрерывности функции  $\varphi(x)$ , ее первой и второй производных представим в виде:

$$\varphi_i(x_{i+1}) = \varphi_{i+1}(x_{i+1}), i=1, \dots, N-3, \quad (4.37)$$

$$\varphi'_i(x_{i+1}) = \varphi'_{i+1}(x_{i+1}), i=0, \dots, N-3, \quad (4.38)$$

$$\varphi''_i(x_{i+1}) = \varphi''_{i+1}(x_{i+1}), i=0, \dots, N-3. \quad (4.39)$$

Условие равенства нулю вторых производных функции на концах отрезка  $[x_0; x_{N-1}]$ , т.е.

$$\varphi''(x_0) = 0, \quad (4.40)$$

$$\varphi''(x_{N-1}) = 0. \quad (4.41)$$

Величины коэффициентов  $a_i, b_i, c_i, d_i$  находятся из решения системы, составленной из уравнений (4.34)-(4.41). Алгоритм масштабирования по строкам приведен в Пособии к лабораторным работам. Для масштабирования по столбцам необходимо выполнить все описанные процедуры над полученными данными в направлении по столбцам.

## 5 УЛУЧШЕНИЕ ИЗОБРАЖЕНИЙ

В компьютерных системах исходные изображения и результаты обработки данных в виде изображения отображаются на экране, при этом получателем информации является наблюдатель. Процедуру, обеспечивающую такое представление, называют визуализацией. Желательно при помощи обработки придать выводимому изображению такие качества, благодаря которым его восприятие человеком было бы по возможности комфортным [37]. Часто бывает полезным подчеркнуть, усилить какие-то свойства наблюданной картины с целью улучшения ее субъективного восприятия.

Последнее - субъективность восприятия - сильно усложняет применение формализованного подхода при 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 использует входные данные из некоторого множества точек исходного изображения, окружающих обрабатываемую точку. Однако имеется группа процедур, в которых осуществляется так называемая поэлементная обработка. В этом случае 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. Поэлементная обработка применяется как предварительная обработка, а также и как заключительный этап при анализе изображений.

Сущность поэлементной обработки изображений сводится к следующему. Пусть  $f(n, m)$  и  $g(n, m)$  - значения яркости исходного и получаемого после обработки изображений соответственно в точке кадра, имеющей декартовы координаты  $n$  (номер столбца) и  $m$  (номер строки). Поэлементная обработка означает, что существует однозначная функциональная зависимость между этими яркостями

$$g(n, m) = \phi(f(n, m)), \quad (5.1)$$

позволяющая по значению исходного сигнала определить значение выходного сигнала. То есть она характеризуется тем, что каждый элемент входного изображения математически преобразуется в новое значение элемента выходного изображения, независимо от значений других элементов входного изображения. Рассмотрим наиболее распространенные процедуры поэлементной обработки.

## 5.1 Яркостные преобразования изображений

### 5.1.1 Линейное контрастирование

Слабый контраст - наиболее распространенное свойство ТВ и ИК изображений, обусловленное условиями наблюдения, ограничением диапазона воспроизводимых яркостей и др. Задача контрастирования связана также и с улучшением 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. В качестве рабочего будем использовать диапазон значений сигнала  $[0,255]$ ; при этом значение 0 соответствует при визуализации уровню черного, а значение 255 - уровню белого. Предположим, что минимальная и максимальная яркости исходного изображения равны  $f_{\min}$  и  $f_{\max}$  соответственно. Если эти параметры или один из них существенно отличаются от граничных значений яркостного диапазона, то визуализированная картина выглядит

либо как темная, либо как ненасыщенная, неудобная, утомляющая при наблюдении. Пример такого неудачного представления приведен на рисунке 5.1а.

При линейном контрастировании используется линейное поэлементное преобразование вида:

$$g(n, m) = af(n, m) + b. \quad (5.2)$$

Параметры преобразования  $a$  и  $b$  определяются желаемыми значениями минимальной  $g_{\min}$  и максимальной  $g_{\max}$  выходной яркости. Решив систему уравнений:

$$\begin{cases} g_{\min} = af_{\min} + b \\ g_{\max} = af_{\max} + b \end{cases}$$

относительно параметров преобразования  $a$  и  $b$ , получим

$$\Rightarrow \begin{cases} a = (g_{\max} - g_{\min}) / (f_{\max} - f_{\min}) \\ b = (g_{\min} f_{\max} - g_{\max} f_{\min}) / (f_{\max} - f_{\min}) \end{cases}.$$

Можно привести (5.2) к виду:



а)



б)

Рисунок 5.1 а) Исходное изображение; б) автоматическое контрастирование (при 0,005 квантилях слева и справа).

$$g = \frac{f - f_{\min}}{f_{\max} - f_{\min}} (g_{\max} - g_{\min}) + g_{\min}. \quad (5.3)$$

Линейное контрастирование исходного изображения, результат которого представлен на рисунке 5.1, выполнено при  $g_{\min} = 0$  и  $g_{\max} = 255$ .

Сравнение двух изображений свидетельствует о значительно лучшем визуальном качестве обработанного изображения. Улучшение связано с

представлением изображения после контрастирования в полном динамическом диапазоне сигнала.

### 5.1.2 Выделение диапазона яркостей

Выделение диапазона яркостей производится преобразованием, амплитудная передаточная характеристика которого имеет вид, представленный на рисунке 5.2.

Преобразование, выполняемое в соответствии с характеристикой, представленной на рисунке 5.2, позволяет выделить те участки изображения, яркость которых соответствует заданному интервалу значений  $[f_{\text{л}} \dots f_{\text{п}}]$ . При этом остальные участки оказываются представленными некоторым «серым» фоном (имеют яркость, соответствующую уровню  $g_{\min}$ ). Такое преобразование позволяет повысить детальность наблюдаемой картины в выбранном диапазоне яркостей. Выходное изображение может использовать полный диапазон представления сигнала.

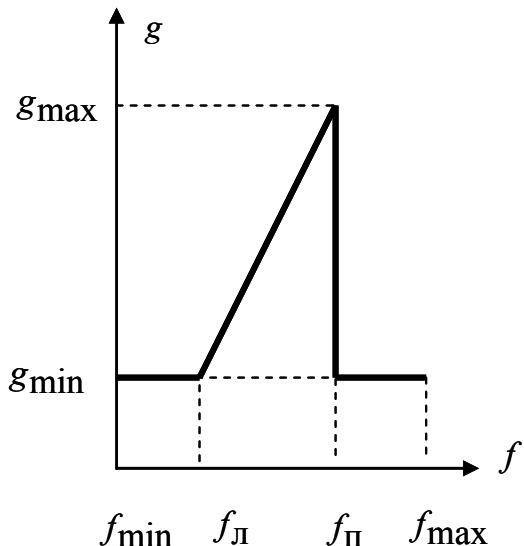


Рисунок 5.2 Амплитудная передаточная характеристика преобразования при выделении диапазона значений сигнала изображения.

### 5.1.3 Преобразование изображения в негатив

Поэлементное преобразование изображения, позволяющее сформировать негатив исходного изображения, выполняется в соответствии с уравнением:

$$g(n,m) = (2^L - 1) - f(m,n). \quad (5.4)$$

Такое преобразование позволяет повысить различимость белых и серых деталей на темном фоне. Эта процедура может быть также связана, например, с синтезом цветных текстурных изображений для формирования сигнала насыщенности при известном сигнале яркости. Такое применение основано на том, что чем выше яркость сигнала, тем меньше насыщенность вследствие более широкого диапазона спектра отраженного излучения. Иллюстрацией такого соответствия служит рисунок 5.3, на

котором представлены изображения негатива яркости и насыщенности, соответствующие одному и тому же изображению, полученному при натурной съемке.

#### 5.1.4 Степенные преобразования

К степенным преобразованиям яркости относятся преобразования вида:



а)б)

Рисунок 5.3 а) Изображение негатива яркостного сигнала изображения; б) изображение сигнала насыщенности этого же изображения.

$$g(n,m) = c(f(n,m) + f_0)^\gamma,$$

где  $c, f_0, \gamma$  - неотрицательные константы.

Одним из наиболее часто применяемых в обработке изображений является преобразование, называемое гамма-коррекцией. Пример применения гамма-коррекции со значениями параметров преобразования  $c = 1$ ,  $f_0 = 0$ ,  $\gamma = 0,5$  представлены на рисунке 5.4.



а)б)

Рисунок 5.4 Пример гамма-коррекции. а) исходное изображение, б) преобразованное.

### 5.1.5 Логарифмическое преобразование

В ряде случаев оказывается полезным нелинейное преобразование яркости. Одним из таких преобразований является логарифмическое преобразование яркости. Диапазон значений яркости автоматически вписывается в диапазон  $[0 \dots 2^{L-1}]$ . Строится гистограмма изображения и оценивается математическое ожидание - mean, минимальное, максимальное значения сигнала. Вычисляются: положительный диапазон  $PositiveRange = \max(2, f_{\max} - mean)$ ;  
отрицательный диапазон  
 $NegativeRange = \max(2, mean - f_{\min})$ ;  
и два коэффициента преобразования:  
 $PositiveAlpha = 2^{L-1} / \ln(PositiveRange)$ ;  
 $NegativeAlpha := 2^{L-1} / \ln(NegativeRange)$ .

Для всех отсчетов изображения вычисляется разность яркости и среднего значения сигнала:

$$Bufer = f(n,m) - mean.$$

На основании этого значения формируется выходное изображение:

$$g(n,m) = \begin{cases} mean + round(PositiveAlpha * \ln(Bufer)), & Bufer \geq 1 \\ mean - round(NegativeAlpha * \ln(\text{abs}(Bufer))), & Bufer \leq -1 \\ mean, & \text{иначе} \end{cases}$$

### 5.2 Преобразование гистограмм

При всех поэлементных преобразованиях происходит изменение закона распределения вероятностей, описывающего изображение. При линейном контрастировании сохраняется вид плотности вероятности, однако в общем случае, т.е. при произвольных значениях параметров линейного преобразования, изменяются параметры плотности вероятности преобразованного изображения.

Определение вероятностных характеристик изображений, прошедших нелинейную обработку, является прямой задачей анализа. При решении практических задач обработки изображений может быть поставлена обратная задача: по известному виду плотности вероятности  $p_f(f)$  и желаемому виду  $p_g(g)$  определить требуемое преобразование  $g = \varphi(f)$ , которому следует подвергнуть исходное изображение. В практике цифровой обработки изображений часто к полезному результату приводит преобразование изображения к равновероятному распределению. В этом случае

$$p_g(g) = \begin{cases} 1/(g_{\max} - g_{\min}), & g_{\min} \leq g \leq g_{\max} \\ 0, & \text{иначе} \end{cases}, \quad (5.5)$$

где  $g_{\min}$  и  $g_{\max}$  - минимальное и максимальное значения яркости преобразованного изображения. Определим характеристику

преобразователя, решающего данную задачу. Пусть  $f$  и  $g$  связаны функцией (5.1), а  $P_f(f)$  и  $P_g(g)$  - интегральные законы распределения входной и выходной яркостей. Учитывая (5.5), находим:

$$P_g(g) = \int_{g_{\min}}^g p_g(g) dg = \frac{g - g_{\min}}{g_{\max} - g_{\min}}.$$

Подставляя это выражение в условие вероятностной эквивалентности  $P_f(f) = P_g(g)$ ,

после простых преобразований получаем соотношение

$$g = (g_{\max} - g_{\min}) \cdot P_f(f) + g_{\min}, \quad (5.6)$$

представляющее собой характеристику (5.1) в решаемой задаче. Согласно (5.6) исходное изображение проходит нелинейное преобразование, характеристика которого  $P_f(f)$  определяется интегральным законом распределения исходного изображения. После этого результат приводится к заданному динамическому диапазону при помощи операции линейного контрастирования.

Таким образом, преобразование плотности вероятности предполагает знание интегрального распределения для исходного изображения. Как правило, достоверные сведения о нем отсутствуют. Аппроксимация аналитическими функциями, вследствие ошибок аппроксимации, может приводить к существенному отличию результатов от требуемых. Поэтому в практике обработки изображений преобразование распределений выполняют в два этапа.

На первом этапе измеряется гистограмма исходного изображения. Для цифрового изображения, шкала яркостей которого, например, принадлежит целочисленному диапазону  $[0, 255]$ , гистограмма представляет собой таблицу из 256 чисел. Каждое из них показывает количество точек в изображении (кадре), имеющих данную яркость. Разделив все числа этой таблицы на общий размер выборки, равный числу отсчетов в изображении, получают оценку распределения вероятностей яркости изображения. Обозначим эту оценку  $\hat{P}_f(f^q)$ ,  $0 \leq f^q \leq 255$ . Тогда оценка интегрального распределения получается по формуле:

$$\hat{P}_f(f^q) = \sum_{i=0}^q \hat{P}_f(f^i).$$

На втором этапе выполняется само нелинейное преобразование (5.6), обеспечивающее необходимые свойства выходного изображения. При этом вместо неизвестного истинного интегрального распределения используется его оценка, основанная на гистограмме. С учетом этого все методы поэлементного преобразования изображений, целью которых является видоизменение законов распределения, получили название гистограммных методов. В частности, преобразование, при котором

выходное изображение имеет равномерное распределение, называется *эквализацией (выравниванием) гистограммы*.

Отметим, что процедуры преобразования гистограмм могут применяться как к изображению в целом, так и к отдельным его фрагментам. Последнее может быть полезным при обработке нестационарных изображений, характеристики которых существенно различаются в различных областях. В этом случае лучшего эффекта можно добиться, применяя гистограммную обработку к отдельным участкам – областям интереса. Правда, при этом изменятся значения отсчетов и всех других областей. На рисунке 5.5 приведен пример эквализации, выполненной в соответствии с изложенной методикой.

Характерной чертой многих изображений, получаемых в реальных изображающих системах, является значительный удельный вес темных участков и сравнительно малое число участков с высокой яркостью. Эквализация гистограммы приводит к выравниванию интегральных площадей равномерно распределенных диапазонов яркостей. Сравнение исходного (рисунок 5.5 а) и обработанного (рисунок 5.5 б) изображений показывает, что перераспределение яркостей, происходящее при обработке, приводит к улучшению визуального восприятия.

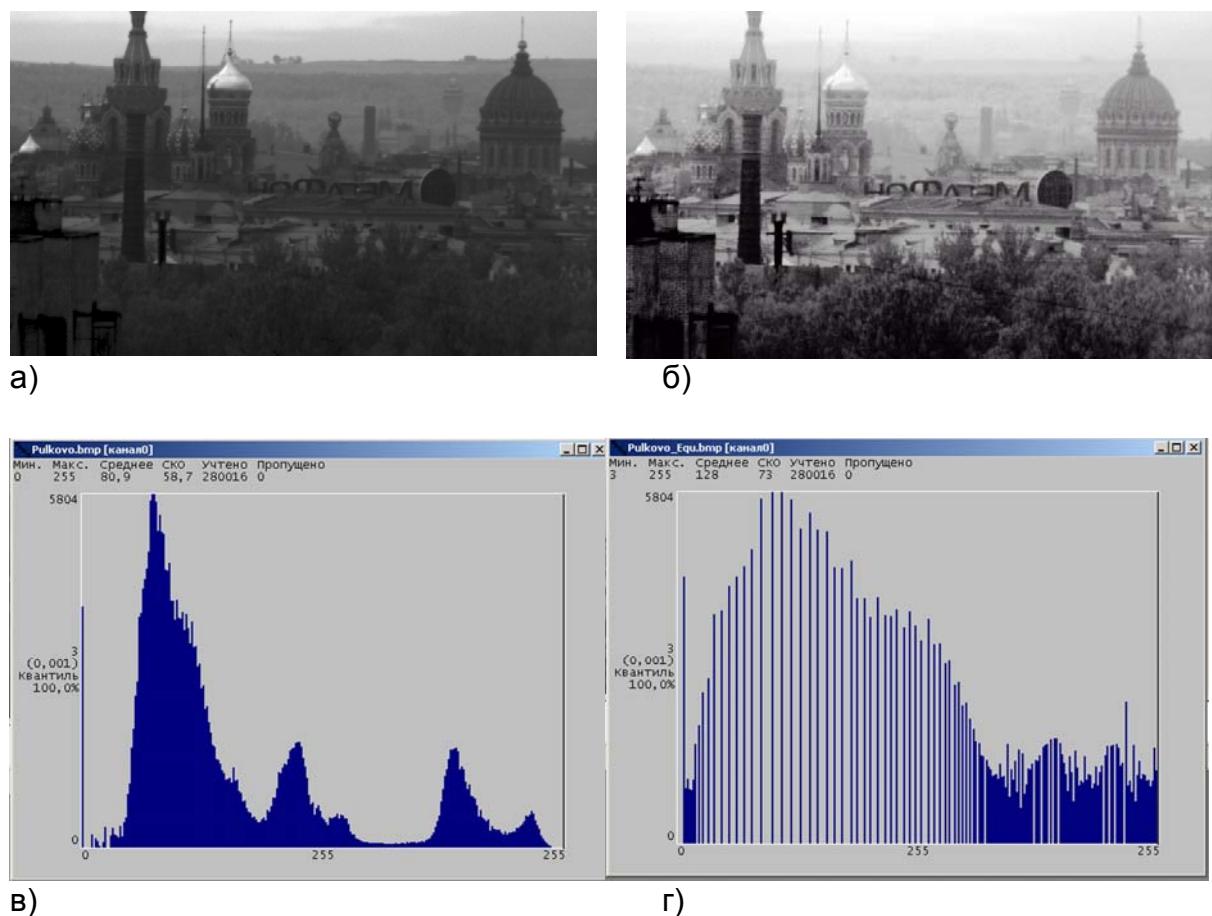


Рисунок 5.5 Пример эквализации гистограммы изображения. а) Исходное изображение и его гистограмма в); б) преобразованное изображение и его гистограмма г).

### 5.3 Пороговая обработка полутоновых изображений

Преобразование с амплитудными характеристиками  $g(f)$ , представленными на рисунке 5.6, превращает полутоновое изображение, содержащее все уровни яркости, в бинарное изображение, отсчеты которого имеют яркости  $g = 0$  или  $g = 1$ .

Такая операция называется бинаризацией или бинарным квантованием. Она применяется для выделения области объекта интереса, присутствующего на изображении. Иногда в качестве уровней белого и черного задают другие значения яркости, например, WhiteVM=170 и BlackVM=85 соответственно, с тем, чтобы визуализировать полученное

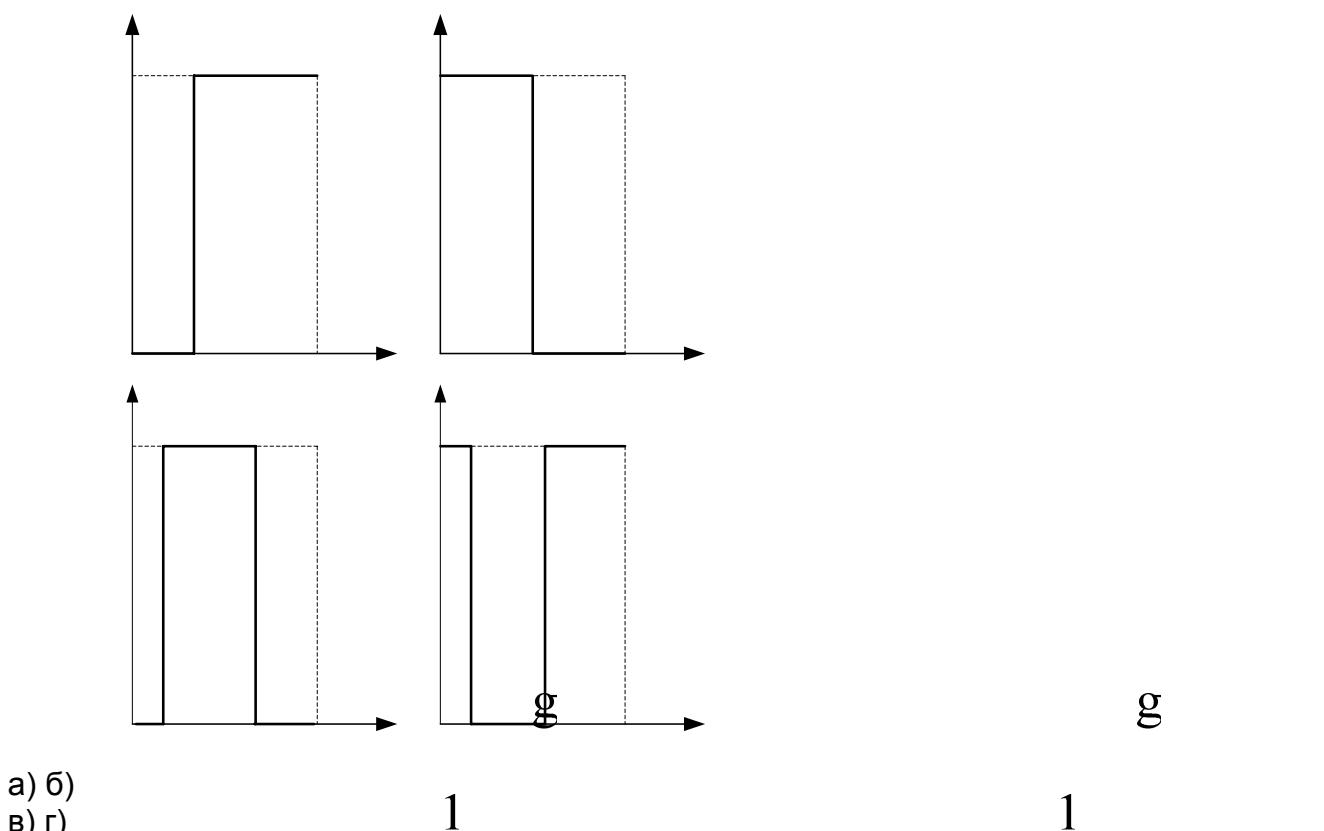
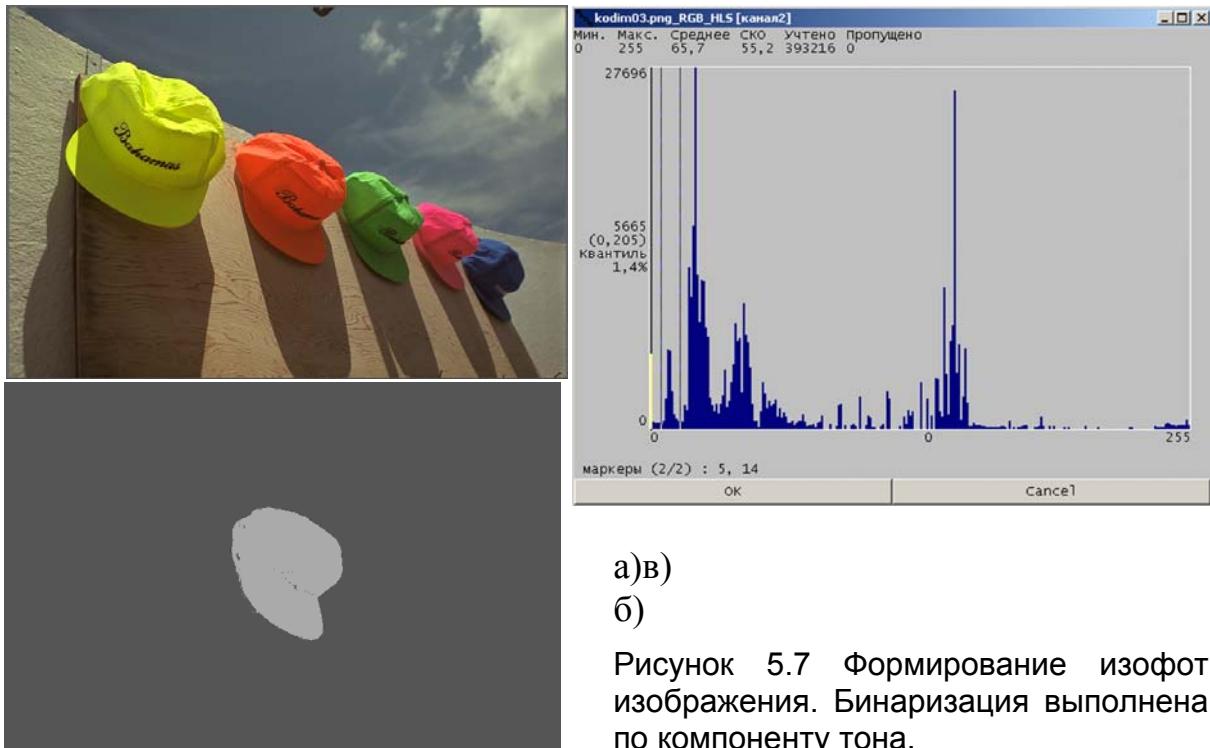


Рисунок 5.6 Амплитудные характеристики для бинаризации изображений.  
а), б) квантование по порогу прямое и инверсное; в) формирование изофот,  
г) формирование инверсных изофот.

бинарное изображение одновременно с полутоновым изображением. Основной проблемой при проведении такой обработки является определение порогов,  $f_L$ ,  $f_H$ , сравнение с которыми сигнала исходного изображения позволяет определить значение сигнала выходного изображения в каждой его точке. Когда при математическом описании изображения применяются модели случайных процессов и случайных полей, определение оптимального порога бинарного квантования представляет собой статистическую задачу.

Бинаризации подвергается обычно не исходное изображение, а некоторое производное, полученнное ~~после выделения~~ каких-либо свойств

объекта. Например, предварительно может выполняться высокочастотная (ВЧ) фильтрация яркостного компонента или преобразование цветовых координатных пространств и последующая обработка одного из компонентов сигнала изображения. На рисунке 5.7 представлено исходное изображение 5 кепок: желтой, красной, зеленой, малиновой и синей (слева направо) а), гистограмма тонового компонента изображения, полученная после выполнения цветового координатного преобразования  $RGB - HLS$  б) и бинарное изображение, полученное в соответствии с характеристикой формирования изофот, представленной на рисунке 5.6 в), выполненной по компоненту тона при порогах  $H_L=5$ ,  $H_P=14$ .



а)в)  
б)

Рисунок 5.7 Формирование изофот изображения. Бинаризация выполнена по компоненту тона.

#### 5.4 Применение табличного метода при поэлементных преобразованиях изображений

При визуализации изображений и при выполнении заранее известных преобразований  $\varphi(f)$  (например, гамма коррекции с заданным значением  $\gamma$ ) вместо вычисления функции преобразования

$$g(f)=\varphi(f) \quad (5.7)$$

в каждой точке изображения производится предварительное вычисление  $g(f)$  в соответствии с (5.7) для полного диапазона сигнала. При числе разрядов квантования равном  $L$ , число входных (и выходных) значений равно  $N=2^L$ . В результате таких вычислений формируется таблица 5.1. Такая таблица называется просмотровой (*LUT - Look up Table*).

Таблица 5.1 Вид просмотровой таблицы

<i>Входной сигнал, f</i>	0	1	2	...	$2^L - 2$	$2^L - 1$
<i>Выходной сигнал, g</i>	$g_1$	$g_2$	$g_3$	...	$g_{N-1}$	$g_N$

Применяется она следующим образом. Входным данным является значение сигнала изображения в текущем отсчете, определяющее номер столбца в первой строке таблицы, а выходным является значение сигнала  $g$  во второй строке этого столбца. Таким образом, входной сигнал является адресом ячейки, где хранится соответствующее ему значение выходного сигнала. При использовании  $LUT$  время преобразования не зависит от сложности функции  $\phi(f)$ , так как не связано со временем ее вычисления.

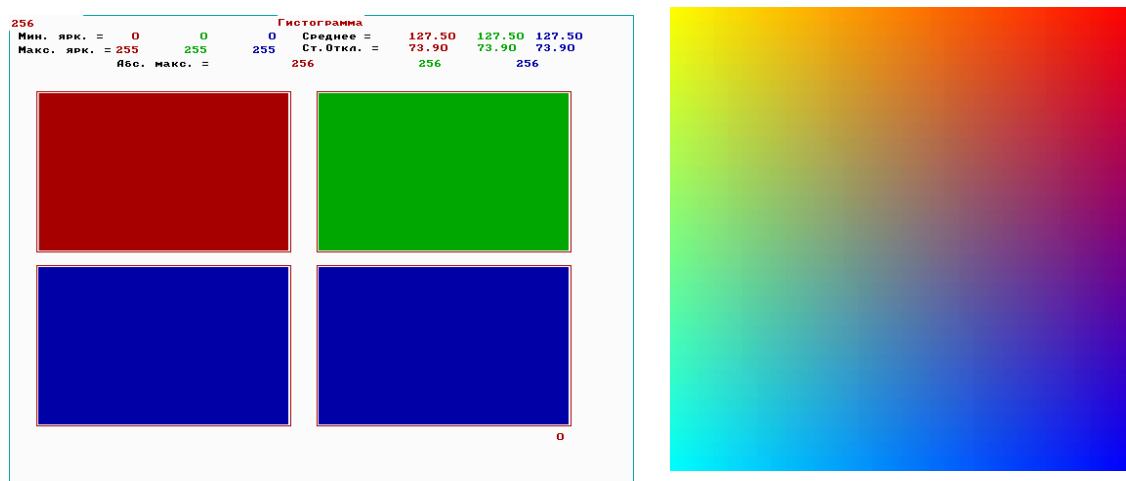
При использовании кодов большой длины для представления сигналов размер таблицы существенно увеличивается. Для ее сокращения весь диапазон входных значений разбивается на поддиапазоны (бины). В таблице каждому бину ставится в соответствие одно значение амплитуды отсчета. Множество значений входного сигнала превышает размер полученной таким образом таблицы. При попадании значения  $f$  между значениями, представленными в столбцах таблицы, применяют интерполяцию - приближенное определение недостающих значений функции  $g$  по имеющимся ее соседним значениям. Часто для этой цели используется *линейная интерполяция*, при которой на участке между заданными значениями функция аппроксимируется отрезком прямой.

Вместе с тем, если при квантовании исходного изображения использовано 8 разрядов, то размер полной таблицы составляет всего 256 адресов для черно-белого изображения и 768 – для цветного. В этом случае таблицы хранятся полностью. Формирование  $LUT$  для каждого из трех компонентов цветного изображения  $RGB$ , позволяет представить черно-белое изображение в псевдоцветах, отобразить тепловизионное изображение в ложных цветах, предъявить изображение тонового компонента после преобразования цветового координатного пространства  $RGB$  в пространство  $HLS$  в соответствующих цветах, выполнить гамма-коррекцию сигнала и многие другие функции. Табличное преобразование эффективно и обеспечивает максимальное быстродействие.

## 5.5 Особенности применения поэлементной обработки цветных изображений

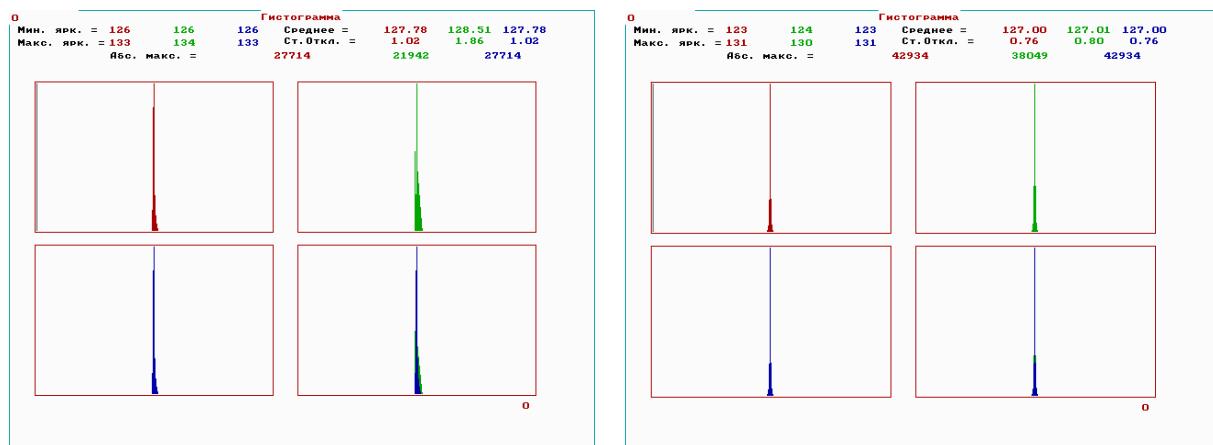
При улучшении цветных изображений часто необходимо сохранить характеристики цветности изображения. В этом случае обеспечивается независимая обработка по яркости и по цвету [38]. Выполняется преобразование  $RGB$  пространства в цветовое координатное пространство яркости и цветности, например, в пространство Стрикланда  $HLS$  или в пространство, в основе которого лежит цветовая система Оствальда [33]. В результате преобразования из пространства  $RGB$  в пространство  $HLS$  возможные значения компонентов изменяются в диапазоне от 0 до 255, а тоновая составляющая, принимающая значения в диапазоне от  $0^\circ$  для красного через оранжевый, желтый, зеленый, голубой, синий до  $360^\circ$  для пурпурного, может быть приведена к диапазону значений от 0 до 252.

Уровни тона с 253 по 255 соответствуют ахроматическому случаю (253-черный, 254-серый, 255-белый). Заметим, что нелинейные преобразования в пространствах Стрикланда и Оствальда приводят к совпадению тоновых компонентов. На рисунке 5.8 представлено тестовое изображение в пространстве *RGB* (а) и гистограмма распределения *RGB* компонентов тестового изображения (б). На рисунке 5.9 представлены гистограммы распределения ошибок прямого и обратного преобразования цветовых координатных пространств, связанные с восьмиразрядным представлением данных. Такие преобразования обеспечивают возможность использования алгоритмов обработки полутонаовых изображений без нарушения их цветовых характеристик, поскольку яркость практически не зависит от атрибутов цветности.



а) б)

Рисунок 5.8 Тестовое изображение. а) Изображение *RGB* компонентов; б) гистограмма распределения *RGB* компонентов тестового изображения.



а) б)

Рисунок 5.9 Гистограмма распределения ошибок преобразования цветового координатного пространства *RGB* в пространство *HLS*: а) для целочисленного преобразования Оствальда; б) для цветовой координатной системы Стрикланда.

Например, в пространстве *HLS* в соответствии с рисунком 3.5 в случае изменения яркостного компонента произойдет лишь изменение длины

вектора  $OP$  для каждого цветного элемента. Положение точки  $P'$  - точки пересечения вектора с плоскостью треугольника Максвелла - при этом останется на прежнем месте, и, соответственно, не изменятся значения тона и насыщенности.

Кроме того, можно использовать информацию ВЧ составляющей изображения, содержащуюся в компоненте насыщенности. Преобразование из аппаратно-ориентированного цветового координатного пространства в пространство, ориентированное на пользователя, каковым является, в частности, пространство  $HLS$ , позволяет выполнять преобразования над изображением, не нарушая баланса белого в изображении, поскольку операции над яркостным компонентом и компонентом насыщенности не приводят к его нарушению, в отличие от операций в пространстве  $RGB$ .

С целью повышения контраста изображения рекомендуется использовать линейное контрастирование яркостной составляющей и повышение насыщенности.

Применение алгоритма подчеркивания ВЧ яркостного компонента повышает четкость изображения. Рисунок 5.10 иллюстрирует представленную методику улучшения изображения на этапе его предварительной обработки.



Рисунок 5.10 Изображение фрагмента карты: а) исходное изображение в пространстве  $RGB$ , б) исходное изображение после повышения контраста, насыщенности и подчеркивания верхних частот.

При наличии неравномерности освещения можно использовать алгоритм компенсации неравномерности по яркостному компоненту, представляющий собой последовательность операций линейной фильтрации в скользящем окне  $128 \times 128$  элементов, вычитания полученной

оценки математического ожидания неравномерности из исходного сигнала и последующей операции линейного контрастирования.

Особый интерес представляет то обстоятельство, что в насыщенных областях изображения при низкой яркости изображения оценивается относительно высокий контраст мелких деталей изображения. То есть в областях изображения с низкими уровнями яркостного компонента, где становятся мало различимыми мелкие детали, сигнал компонента насыщенности имеет высокое значение, и, следовательно, улучшает условия наблюдаемости в темных областях. Для улучшения яркостного сигнала можно использовать информацию сигнала насыщенности в областях низких уровней яркости. При этом сигнал яркости может быть сложен со взвешенным сигналом контурного оператора негатива сигнала насыщенности.

Эффективным способом повышения цветового контраста в условиях низкой освещенности является накопление сигнала на матрице ПЗС. На рисунке 5.11 приведен пример, иллюстрирующий улучшение изображения при наблюдении за неподвижными объектами в условиях низкой освещенности.



а) б)  
в) г)

Рисунок 5.11 а) Исходное изображение без накопления, б) исходное изображение с накоплением 0,5 с; в) улучшенное изображение а; г) улучшенное изображение б).

## 6 ОСНОВЫ ФИЛЬТРАЦИИ ИЗОБРАЖЕНИЙ

Изображения, полученные на выходе оптико-электронных преобразователей, искажены помехами. Это затрудняет как визуальный анализ изображений человеком - оператором, так и их автоматическую компьютерную обработку. При обработке изображений помехами являются и некоторые области самого изображения. Например, при анализе объектов на сложном фоне, фон тоже представляет собой помеху. При цифровой обработке изображений необходимо устранять геометрические искажения изображений, подавлять шумы различной природы, производить апертурную коррекцию. Ослабление действия помех достигается фильтрацией.

Фильтрация изображений производится в пространственной и частотной областях. При пространственной фильтрации изображений преобразование выполняется непосредственно над значениями отсчетов изображения. Результатом фильтрации является оценка полезного сигнала изображения. Изображение представляет собой двумерную функцию пространственных координат, изменяющуюся медленнее, чем двумерная функция, описывающая помеху. Поэтому при оценке полезного сигнала в каждой точке кадра рассматривают окрестность этой точки (некоторое множество соседних с ней точек), используя общие характеристики сигнала в этой окрестности. В других случаях признаком полезного сигнала являются резкие перепады яркости. Однако, как правило, частота этих перепадов относительно невелика, так что на значительных промежутках между ними сигнал либо постоянен, либо изменяется медленно. И в этом случае свойства сигнала проявляются при наблюдении его не только в отдельной точке, но и при анализе ее окрестности. Заметим, что понятие окрестности является достаточно условным. На рисунке 6.1 представлена иерархия окрестностей отсчета, обозначенного «0».

9	8	7	6	7	8	9
8	5	4	3	4	5	8
7	4	2	1	2	4	7
6	3	1	0	1	3	6
7	4	2	1	2	4	7
8	5	4	3	4	5	8
9	8	7	6	7	8	9

Рисунок 6.1 Конфигурации окрестности элемента «0» в кадре изображения в иерархической последовательности.

«1» обозначена окрестность первого порядка, для которой расстояние между элементами равно 1. «2» обозначена окрестность второго порядка, к которой относятся диагональные элементы, расстояние от которых до центрального отсчета «0» равно  $\sqrt{2}$ . Окрестность третьего порядка представлена элементами, отстоящими от центрального элемента на расстояние, равное 2, и так далее.

В соответствии с рисунком 6.1 формируется иерархия конфигураций окрестности центрального отсчета рассматриваемого фрагмента (окна) кадра по возрастанию расстояний от него до отсчета окрестности.

Окрестность может быть образована лишь ближайшими соседями, но может содержать и достаточно много элементов кадра. При рассмотрении окрестности большого размера, иногда устанавливается различная степень влияния далеких и близких от центра окрестности точек на сигнал, формируемый на выходе фильтра в данной точке кадра. Таким образом, идеология фильтрации основывается на использовании как данных текущей точки, так и ее окрестности. В этом проявляется существенное отличие фильтрации от рассмотренных выше поэлементных процедур: фильтрация не может быть поэлементной процедурой обработки изображений.

Традиционная фильтрация в частотной области требует выполнения следующей последовательности преобразований [28]:

- двумерное дискретное преобразование изображения из пространственной области в частотную (например, посредством дискретного преобразования Фурье),
- преобразование дискретного спектра сигнала изображения,
- обратное двумерное дискретное преобразование, позволяющее восстановить полезный сигнал изображения в пространственной области.

Задача заключается в том, чтобы найти такую вычислительную процедуру, которая обеспечила бы получение наилучших результатов. Общепринято при решении этой задачи опираться на использование вероятностных моделей изображения и помехи, а также на применение статистических критериев оптимальности. Причины этого понятны - это случайные законы распределения полезного сигнала и помехи и стремление получить минимальное в среднем отличие результата обработки от результата обработки идеального сигнала. Многообразие методов и алгоритмов связано с большим разнообразием сюжетов, обусловливающих множество различных математических моделей, используемых для описания сигналов. Кроме того, применение различных критериев оптимальности также ведет к разнообразию методов фильтрации. Наконец, даже при совпадении моделей и критериев часто из-за математических трудностей не удается найти оптимальную процедуру. Сложность нахождения точных решений порождает различные варианты приближенных методов и процедур.

В практике цифровой обработки изображений широко используется масочная фильтрация. Ее линейная разновидность является одним из вариантов двумерной фильтрации с конечной импульсной характеристикой (КИХ) фильтра. В качестве маски используется множество весовых коэффициентов, заданных во всех точках окрестности  $S$ , обычно симметрично окружающих текущую точку кадра. Распространенным видом окрестности, часто применяемым на практике, является квадрат  $3 \times 3$  с текущим элементом в центре. Применяют различные маски, одним из эвристических вариантов является равномерная маска, все девять весовых коэффициентов которой равны  $1/9$ . Такой выбор коэффициентов отвечает условию сохранения средней яркости, вследствие чего выходной сигнал оказывается вписанным в диапазон входного сигнала.

Применение процедур фильтрации приводит к существенному снижению уровня шума в изображении.

## 6.1 Основы пространственной фильтрации изображений

### 6.1.1 Линейная пространственная фильтрация

Пространственная фильтрация изображения  $f(x, y)$ ,  $x \in [0, (N - 1)]$ ,  $y \in [0, (M - 1)]$  позволяет применять фильтры с КИХ [39]. Поскольку в изображении понятия прошлого и будущего времени становятся условными, мы можем использовать амплитуды отсчетов как в направлении увеличения индексов, так и в направлении уменьшения индексов. Пространственная фильтрация выполняется как операция двумерной свертки импульсной характеристики фильтра  $h(s, t)$  с изображением  $f(x, y)$ , где  $s$  – координата характеристики в горизонтальном направлении вдоль оси  $x$ ,  $s \in [-n/2, n/2]$ ,  $t$  – координата характеристики в вертикальном направлении вдоль оси  $y$ ,  $t \in [-m/2, m/2]$ :

$$g(x, y) = f(x, y) * h(x, y) = \sum_{t=-m/2}^{m/2} \sum_{s=-n/2}^{n/2} f(s, t)h(x-s, y-t) = \sum_{t=-m/2}^{m/2} \sum_{s=-n/2}^{n/2} f(x-s, y-t)h(s, t). \quad (6.1)$$

Прямоугольная область размером  $n \times m$ , на которой задана импульсная характеристика, называется маской или ядром фильтра. Рассмотрим, как соотносятся координаты изображения и импульсной характеристики фильтра на примере.

Пусть  $m = 3$ ,  $n = 3$ . Элементы импульсной характеристики фильтра и соответствующей области изображения представлены на рисунке 6.2. Начало координат фильтра устанавливается в центр импульсной характеристики, как показано на рисунке 6.2. Отсчеты импульсной характеристики отражаются относительно начала координат (что

равносильно повороту маски на  $180^\circ$ , и центр маски смещается в положение  $(x, y)$ .

	$s = -1$	$s = 0$	$s = 1$		$s = -1$	$s = 0$	$s = 1$
$t = -1$	$f(-1, -1)$	$f(0, -1)$	$f(1, -1)$		$h(1, 1)$	$h(0, 1)$	$h(-1, 1)$
$t = 0$	$f(-1, 0)$	$f(0, 0)$	$f(1, 0)$		$h(1, 0)$	$h(0, 0)$	$h(-1, 0)$
$t = 1$	$f(-1, 1)$	$f(0, 1)$	$f(1, 1)$		$h(1, -1)$	$h(0, -1)$	$h(-1, -1)$

Рисунок 6.2 Положение отсчетов импульсной характеристики при свертке с изображением  $f(x, y)$ .

Отклик фильтра  $g(x, y)$  вычисляется как сумма произведений отсчетов изображения на соответствующие отсчеты повернутой импульсной характеристики. Эта операция выполняется для каждого отсчета изображения.

Если импульсная характеристика фильтра симметрична, то есть  $h(s, t) = h(-s, -t)$ , то вместо свертки можно выполнять корреляцию:

$$g(x, y) = f(x, y) \otimes h(x, y) = \sum_{t=-m/2}^{m/2} \sum_{s=-n/2}^{n/2} f(x+s, y+t) h(s, t). \quad (6.2)$$

Эта операция не требует отражения импульсной характеристики, а представляет собой вычисление в скользящей по изображению маске фильтра произведений отсчетов изображения на соответствующие коэффициенты фильтра и их суммирование (в соответствии с рисунком 6.3).

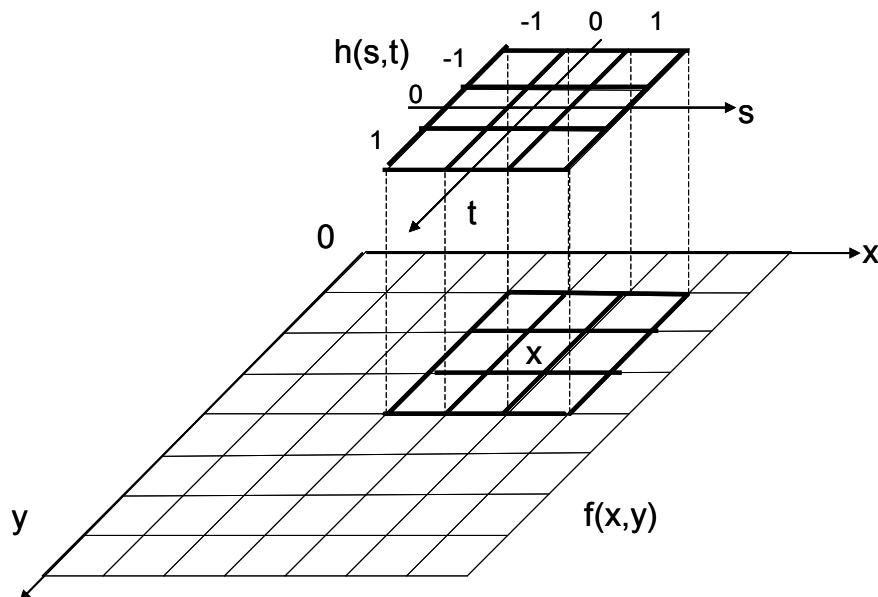


Рисунок 6.3 Корреляция изображения  $f(x, y)$  с маской  $h(s, t)$ .

Для обработки краевых эффектов необходимо увеличить изображение по строкам и столбцам на  $n/2$  отсчетов слева и справа и на  $m/2$  отсчетов сверху и снизу. При этом расширение может выполняться разными способами: дополнением нулями, повторением граничных элементов, периодическим повторением ( $x \bmod N$ ,  $y \bmod M$ ) или зеркальным отражением граничных элементов. Размер выходного изображения при этом сохраняется равным размеру входного изображения.

При построении линейных КИХ фильтров часто используют непараметрический подход.

*Линейные сглаживающие фильтры.* Для уменьшения шумов широко применяются НЧ фильтры, поскольку шум представляет собой ВЧ сигнал. В частности, для НЧ фильтрации применяется усреднение сигнала в маске, например, при  $n=m=3$ :

$$h_1 = \frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}; h_2 = \frac{1}{10} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix}; h_3 = \frac{1}{16} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}. \quad (6.3)$$

Нормировка необходима для того, чтобы привести значения отклика фильтра к диапазону входных данных. Нормирующий коэффициент определяется из условия равенства единице суммы всех коэффициентов КИХ в соответствии с уравнением:

$$k = 1 / \left( \sum_{t=-m/2}^{m/2} \sum_{s=-n/2}^{n/2} h(s, t) \right). \quad (6.4)$$

Линейная фильтрация широко применяется при подавлении шумов в изображении, для компенсации неравномерности чувствительности, создания эффектов размытия изображений.

Она также широко применяется в задачах выделения контуров на изображении, подчеркивания верхних пространственных частот [40,41]. В этом случае коэффициенты КИХ фильтра вычисляются на основе дифференцирования амплитуды сигнала, что эквивалентно дискретным разностям амплитуд отсчетов:

$$\begin{aligned} \partial f(x, y) / \partial x &\sim \Delta f_x(x, y) = f(x, y) - f(x - 1, y) \\ \partial f(x, y) / \partial y &\sim \Delta f_y(x, y) = f(x, y) - f(x, y - 1). \end{aligned} \quad (6.5)$$

Производные можно брать не только по горизонтали и вертикали, но и в произвольном направлении. Соответствующие импульсные характеристики фильтров имеют вид:

$$\begin{bmatrix} -1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 & 1 \end{bmatrix}, \quad (6.6)$$

$$\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}. \quad (6.7)$$

КИХ (6.7), формирующие производные в ортогональных направлениях с наклоном  $+45^\circ$  и  $-45^\circ$ , используются в операторе Робертса.

Доктор Джудит Превитт для обнаружения границ медицинских изображений применила оператор, маски которого получили ее имя [42]:

$$\begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix}. \quad (6.8)$$

Широко применяется в обработке изображений для выделения контуров оператор Собеля:

$$\begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}, \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}. \quad (6.9)$$

Для выделения контуров применяются также вторые производные амплитуды сигнала. Оператор Лапласа, имеющий вид:

$$\nabla^2 f(x, y) = \partial^2 f(x, y) / \partial x^2 + \partial^2 f(x, y) / \partial y^2,$$

можно применить в виде линейного КИХ фильтра с импульсной характеристикой:

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}. \quad (6.10)$$

Линейные КИХ фильтры просты в реализации.

### 6.1.2 Нелинейная пространственная фильтрация

В результате применения линейных сглаживающих фильтров происходит подавление шумов, но одновременно размываются границы между областями с разной амплитудой сигнала. Для уменьшения «смазы» границ разработаны различные нелинейные фильтры. Как и линейные КИХ фильтры, нелинейные фильтры работают в скользящем окне. Но, при линейной фильтрации вычисляется линейная комбинация отсчетов сигнала, а при нелинейной фильтрации выполняются нелинейные преобразования отсчетов сигнала в определяемой маской фильтра окрестности элементов.

#### 6.1.2.1 Сигма-фильтр

Сигма-фильтр предназначен для подавления шумов в изображении с сохранением контуров (резких границ областей) [43]. Центральный элемент маски замещается взвешенным средним значением, вычисленным только по тем амплитудам отсчетов, значения которых попадают в  $\pm k\sigma$  - область относительно яркости центрального элемента.  $\sigma$  выбирается либо как СКО подавляемого шума, либо как СКО в маске, либо устанавливается равной СКО, полученному по всему изображению:

$$g(x, y) = \sum_{s, t \in S} h(s, t) f(x - s, y - t), \quad (6.11)$$

где  $S$ -окрестность составляют те значения координат маски, в которых выполняется наложенное условие:

$$\mathbf{S} = \{(s, t) : |f(x - s, y - t) - f(x, y)| \leq k\sigma\}, \quad (6.12)$$

$h(s, t)$ -КИХ линейного сглаживающего фильтра, подобного (6.3).

При  $k = 2$  диапазон заменяемых значений составляет  $\pm 2\sigma$ , в случае нормального распределения шума вероятность попадания амплитуды за пределы диапазона равна 4,55%.

Фильтры, основанные на порядковых статистиках, также относятся к нелинейным фильтрам. Наиболее эффективными по совокупности воздействий: сглаживания шума на однородных участках изображения, сохранения скачков изменения яркости, минимального искажения формы границы, подавления импульсного шума, вычислительной эффективности является медианный фильтр.

### 6.1.2.2 Медианный фильтр

Медианный фильтр (МФ) (предложен Тьюки в 1974 г.) заменяет центральный элемент маски медианой упорядоченной (по невозрастанию или по неубыванию амплитуды) выборки, сформированной из всех амплитуд отсчетов, покрываемых маской фильтра. При применении МФ происходит последовательная обработка каждой точки кадра, в результате чего образуется последовательность оценок. При медианной фильтрации используется скользящее двумерное окно. В принципе, для каждого отсчета выполняется независимая оценка медианы в окне. В целях ускорения оценки целесообразно алгоритмически на каждом шаге использовать ранее выполненные вычисления. Размер окна устанавливается нечетным и равным  $m \times n$ . Отсчеты изображения, оказавшиеся в пределах окна, образуют рабочую выборку текущего отсчета. Если упорядочить последовательность  $\{f_i, i=[1, mn]\}$  по неубыванию, то ее медианой будет тот элемент выборки, который занимает центральное положение в этой упорядоченной последовательности. Этот элемент является  $(mn + 1)/2$  наибольшим и  $(mn + 1)/2$  наименьшим значением в выборке и определяет результат медианной фильтрации для текущей точки кадра. Введем формальное обозначение описанной процедуры в виде:

$$g_{\text{med}} = \text{med}(f_1, f_2, \dots, f_n). \quad (6.13)$$

Рассмотрим пример. Предположим, что упорядоченная последовательность  $\mathbf{Y}$  в окне размером  $3 \times 3$  имеет вид:  $\mathbf{Y} = \{99, 140, 97, 150, 255, 155, 158, 99, 175\}$ , где элемент 255 соответствует центру окна  $(x, y)$ . Большое значение яркости в этой точке кадра является результатом воздействия импульсной помехи. Упорядоченная по неубыванию выборка имеет вид:  $\{97, 99, 99, 140, 150, 155, 158, 175, 255\}$ , следовательно, в соответствии с процедурой (6.13), на выходе медианного фильтра получаем  $g_{\text{med}} = 150$ . Видим, что учет яркостей элементов окрестности при фильтрации в текущей точке привел к подавлению импульсной помехи. Если импульсная помеха не является точечной, а занимает некоторую область, то она также может быть подавлена, если

размер этой локальной области будет меньше, чем половина размера апертуры МФ. Поэтому для подавления импульсных помех, поражающих локальные участки изображения, следует увеличивать размеры апертуры МФ.

Из (6.13) следует, что действие МФ состоит в “игнорировании” как положительных, так и отрицательных выбросов значений входной выборки. Такой принцип подавления помехи может быть применен и для ослабления шума на изображении. Однако исследование подавления шума при помощи медианной фильтрации показывает, что ее эффективность при решении этой задачи ниже, чем у линейной фильтрации. Медианная фильтрация лучше сохраняет границы изображения, чем любая линейная фильтрация. Механизм этого явления очень прост и заключается в следующем. Предположим, что окно фильтра находится вблизи границы, разделяющей светлый и темный участки изображения, при этом его центр располагается в области темного участка. Тогда рабочая выборка будет содержать большее количество элементов с малыми значениями яркости, следовательно, и медиана будет находиться среди элементов с малыми значениями яркости. И наоборот, если центр окна смещен в область более высокой яркости, то и медиана будет находиться в области более высокой яркости. Это позволяет при применении МФ сохранить перепады яркости.

Медианные фильтры подавляют импульсные шумы. К таким шумам относится и шум типа «соль и перец», отсчеты которого имеют значения, соответствующие максимальному («соль») и минимальному («перец») уровням квантования в сигнале изображения. Резкие изменения амплитуды сохраняются медианным фильтром, а импульсная помеха, размер которой  $\leq mn / 2$ , таким фильтром подавляется. Однако при увеличении маски фильтра можно потерять информацию о малоразмерных областях изображения и произвести искажение границ областей, особенно в угловых положениях. На рисунке 6.4 приведен синтезированный пример, иллюстрирующий исчезновение линии толщиной 1 элемент на изображении, полученном после применения медианного фильтра размером  $3 \times 3$ .



а) б) в)

Рисунок 6.4 Применение медианного фильтра с окном  $3 \times 3$  элемента. а) исходное изображение; б) выходное изображение; в) изображение разности изображений а) и б).

Несмотря на подавление сигнала и в некоторых других областях (черные точки на разностном изображении рисунке 6.4 в) границы областей не размыты.

Цифровое цветное изображение, представленное в *RGB* цветовом координатном пространстве, является трехкомпонентным двумерным сигналом. Каждый элемент цифрового растрового цветного изображения размера  $M \times N$  может быть представлен вектором:

$$\mathbf{C}(x,y) = [R(x,y), G(x,y), B(x,y)],$$

где  $x$ - номер столбца,  $y$ - номер строки,  $x \in [0, N - 1]$ ,  $y \in [0, M - 1]$ .  $R(x,y)$ ,  $G(x,y)$ ,  $B(x,y)$ -  $R$ ,  $G$ ,  $B$ -компоненты элемента с координатами раstra  $(x,y)$ .

При создании векторных медианных фильтров вводят меру расстояния между векторами [44]. Функция расстояния между двумя векторами  $\mathbf{C}_i$  и  $\mathbf{C}_j$  в пространстве *RGB* в общем виде задается формулой:

$$d_{ij} = \|\mathbf{C}_i - \mathbf{C}_j\|_\gamma = \left( \sum_{n=1}^3 |C_i^n - C_j^n|^\gamma \right)^{1/\gamma},$$

где  $k$ - номер компонента вектора,  $\gamma$  - характеристика метрики.

В случае евклидова расстояния (метрика  $L^2$ ,  $\gamma=2$ )

$$d_{ij} = \|\mathbf{C}_i - \mathbf{C}_j\| = \sqrt{(R_i - R_j)^2 + (G_i - G_j)^2 + (B_i - B_j)^2}. \quad (6.14)$$

Множество входных векторов в скользящем окне фильтрации обозначим

$$\mathbf{W} = \{\mathbf{C}_0, \mathbf{C}_1, \dots, \mathbf{C}_{p-1}\},$$

где  $\mathbf{C}_i = (R_i, G_i, B_i)$ ,  $i \in [0, p - 1]$ .

Расстояние для  $j$ -го вектора оценивается как сумма расстояний от  $j$ -го вектора до каждого из векторов в окне:

$$D_j = \sum_{i=0}^{p-1} d_{ij}. \quad (6.15)$$

Последовательность расстояний  $\{D_0, D_1, \dots, D_{p-1}\}$  упорядочивают по возрастанию, и в качестве отклика фильтра используют вектор  $\mathbf{C}_i$ , соответствующий минимальному расстоянию  $D_i = \min_j \{D_j\}$ . Другие медианные векторные фильтры отличаются мерой расстояний и весовыми функциями, используемыми для взвешивания суммы в (6.15).

Другой подход основан на каскадной схеме многоканальной фильтрации, которая позволяет подавить импульсную помеху и в большей степени сохранить границы объектов, поскольку МФ выполняется в направлении границы и в минимально возможном окне [45].

Поскольку применение МФ приводит все же к подавлению ВЧ изображения, вызывая размывание краев и текстур, в настоящее время получают все большее развитие схемы адаптивной фильтрации [28,46],

которые позволяют изменить импульсную характеристику фильтра в зависимости от локального значения сигнала изображения.

Один из алгоритмов адаптивной медианной фильтрации (АМФ) выполняется следующим образом [28]. В окне фильтрации оцениваются минимальное значение сигнала  $f_{\min}$ , максимальное значение  $f_{\max}$  и медиана  $f_{\text{med}}$ . Фильтрации подвергается только тот центральный элемент окна  $f(x,y)$ , для которого выполняется условие 6.16 а): значение медианы больше минимального и меньше максимального значений в окне и не выполняется условие 6.16 б): значение сигнала в центре окна больше минимального и меньше максимального значений в окне.

- a)  $A1 = f_{\text{med}} - f_{\min}; A2 = f_{\text{med}} - f_{\max}; A1 > 0 \text{ и } A2 < 0,$  (6.16)
- b)  $B1 = f(x,y) - f_{\min}; B2 = f(x,y) - f_{\max}; B1 > 0 \text{ и } B2 < 0.$

Применение такого фильтра позволяет удалить биполярную импульсную помеху, обеспечить сглаживание шумов и уменьшить подавление ВЧ в изображении. На рисунке 6.5 приведены примеры подавления импульсной помехи с применением МФ и АМФ. Медианный фильтр относится к ранговым фильтрам.



Рисунок 6.5 Пример фильтрации изображения с импульсной помехой (1%): а) исходное изображение; б) МФ 3х3; в) АМФ.

Применяются и другие ранговые фильтры. Так, выбор минимального отсчета в маске (ранг равен 1), позволяет выполнить *эрозию*, а максимального (ранг равен  $m$ ) – *дилатацию*. Эти операции широко используются в морфологической обработке изображений.

### 6.1.3 Фильтры повышения верхних пространственных частот изображения

Поскольку приведенные КИХ линейных дифференциальных фильтров (уравнения (6.6)-(6.10)) производят положительные и отрицательные отклики, то выходной сигнал суммируется со значением равным  $2^{L-1}$ , где  $L$ - число разрядов при квантовании амплитуды (при  $L=8$  это значение равно 128). Это связано с тем, что поскольку в изображении велики корреляционные связи между близлежащими элементами, то большие значения разностей яркостей имеют малые вероятности, зато небольшие

изменения квантуются на большее число уровней, чем в случае приведения выходного динамического диапазона значений сигнала к диапазону его представления в системе. Значения выходного сигнала ограничиваются диапазоном представления сигнала  $[0, 2^L - 1]$ . В этом случае фильтр остается линейным. При формировании отклика в виде модуля градиента, что часто используется при формировании контуров изображения, фильтрация становится нелинейной.

Рассмотрим применение дифференциальных операторов для выделения контуров. При этом вместо (6.5) будем использовать модуль градиента:

$$g(x, y) = \sqrt{(f(x, y) - f(x - 1, y))^2 + (f(x, y) - f(x, y - 1))^2}.$$

Для повышения скорости вычисления в дифференциальных операторах вместо квадратов и квадратных корней используют абсолютные значения. В этом случае градиент оценивается в соответствии с выражением:

$$g(x, y) = |f(x, y) - f(x - 1, y)| + |f(x, y) - f(x, y - 1)|.$$

«Перекрестный градиент» Робертса вместо (6.7) вычисляется таким образом:

$$g_R(x, y) = \sqrt{(f(x, y) - f(x - 1, y - 1))^2 + (f(x - 1, y) - f(x, y - 1))^2}.$$

Упрощенное выражение для оценки имеет вид:

$$g_R(x, y) = |f(x, y) - f(x - 1, y - 1)| + |f(x - 1, y) - f(x, y - 1)|.$$

При этом градиент  $g(x, y)$  лучше выделяет контур в направлении  $+45^\circ$ , а оператор Робертса  $g_R(x, y)$  лучше выделяет вертикальные и горизонтальные границы. Розенфельд использовал оператор с предварительным сглаживанием:

$$g_{AR}(x, y) = |(f(x, y) + f(x - 1, y)) - (f(x, y - 1) + f(x - 1, y - 1))| + |(f(x - 1, y - 1) + f(x - 1, y)) - (f(x, y - 1) + f(x, y))|.$$

Использование попарных разностей значений яркости элементов, расположенных по обе стороны от элемента с координатами  $(x, y)$  позволяет сформировать оператор, хорошо выделяющий границы в направлениях  $\pm 45^\circ$ :

$$g_2(x, y) = |f(x - 1, y) - f(x + 1, y)| + |f(x, y - 1) - f(x, y + 1)|.$$

В [40] выполнено сравнение двумерных операторов для выделения границ объектов. В качестве модели были приняты: объект постоянной яркости, фон с постоянной яркостью фона, аддитивный нормальный шум с задаваемыми параметрами: математическим ожиданием и СКО. Получены оценки вероятности пропуска цели на выходе контурных операторов при выделении границ. При этом ложной тревогой называется превышение порога  $T$  шумом, а пропуском цели – пропуск истинного контурного оператора. Сравнение выполнено по критерию Неймана – Пирсона. Лучшим считается тот оператор, который обеспечивает минимальную вероятность пропуска цели при заданной вероятности ложной тревоги.

Показано, что вероятность пропуска контурного элемента существенно зависит от направления границы. Меньшую вероятность пропуска цели для вертикальных и горизонтальных границ обеспечивает оператор Робертса, а для границ под углами  $\pm 45^\circ$ - оператор  $g_2(x,y)$ . Показано, что замена квадратного корня из суммы квадратов на сумму абсолютных значений в операторе Робертса не ухудшает вероятностные характеристики сигналов на выходе контурного оператора. Предложен инвариантный к направлению границы оператор, полученный объединением операторов  $g_R(x,y)$  и  $g_2(x,y)$ :

$$g_3(x,y) = [g_R(x,y) \geq T] \vee [g_2(x,y) \geq T],$$

где  $T$  – значение порога.

## 6.2 Частотная фильтрация изображений

Частотная фильтрация выполняется в частотной области. Это означает, что при частотной фильтрации выполняются прямое и обратное пространственно-частотное преобразование.

### 6.2.1 Дискретное преобразование Фурье

Прямое двумерное дискретное преобразование Фурье (ДПФ) преобразует изображение, заданное в пространственной координатной системе  $(x,y)$ , в двумерное дискретное преобразование изображения, заданное в частотной координатной системе  $(u,v)$  [12]:

$$F(u,v) = \left(1 / \sqrt{NM}\right) \sum_{y=0}^{M-1} \sum_{x=0}^{N-1} f(x,y) \exp(-i2\pi(ux/N + vy/M)), \quad (6.17)$$

где  $i = \sqrt{-1}$ ,  $x \in [0, N-1]$ ,  $y \in [0, M-1]$ .

Обратное дискретное преобразование Фурье (ОДПФ) имеет вид:

$$f(x,y) = \left(1 / \sqrt{NM}\right) \sum_{v=0}^{M-1} \sum_{u=0}^{N-1} F(u,v) \exp(i2\pi(ux/N + vy/M)), \quad (6.18)$$

где  $i = \sqrt{-1}$ ,  $u \in [0, N-1]$ ,  $v \in [0, M-1]$ .

Из (6.17) и (6.18) видно, что ДПФ является комплексным преобразованием. Модуль этого преобразования представляет амплитуду спектра изображения и вычисляется как корень квадратный из суммы квадратов действительной и мнимой частей ДПФ. Фаза (угол сдвига фазы) определяется как арктангенс отношения мнимой части ДПФ к действительной. Энергетический спектр равен квадрату амплитуды спектра, или сумме квадратов мнимой и действительной частей спектра.

### 6.2.2 Теорема о свертке

В соответствии с теоремой о свертке, свертка двух функций в пространственной области может быть получена ОДПФ произведения их ДПФ [28], то есть

$$f(x,y) * h(x,y) \Leftrightarrow H(u,v)F(u,v). \quad (6.19)$$

Фильтрация в частотной области позволяет по ДПФ изображения подобрать частотную характеристику фильтра, обеспечивающую необходимое преобразование изображения. Рассмотрим частотные характеристики наиболее распространенных фильтров.

### 6.2.3 Низкочастотные фильтры

Частотная характеристика идеального НЧ фильтра имеет вид:

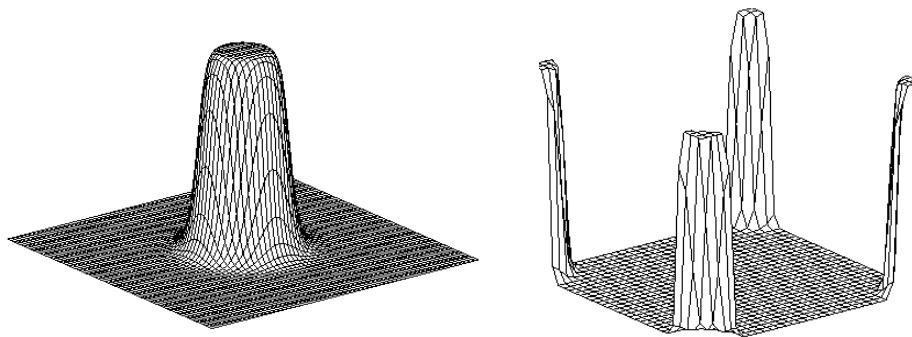
$$H_{\text{НЧ}}(u, v) = \begin{cases} 1 & , \text{при } r(u, v) \leq r_{\text{ср}}, \\ 0 & , \text{иначе} \end{cases}, \quad (6.20)$$

где  $r(u, v) = \sqrt{(u - u_0)^2 + (v - v_0)^2}$  - расстояние от центра маски фильтра до отсчета с координатами  $(u, v)$ ,  $(u_0, v_0)$  - координаты центра маски фильтра,  $r_{\text{ср}}$  - заданное неотрицательное число. При размере изображения  $N \times M$ ,  $u_0 = N/2$ ,  $v_0 = M/2$ .

Частотная характеристика НЧ фильтра Баттервортта (Butterworth) порядка  $k$ , подавляющего частоты, отстоящие на расстояние более  $r$  (по окружности) от начала координат, имеет вид [28]:

$$H(u, v) = 1 / (1 + (r(u, v) / r_{\text{ср}})^{2k}). \quad (6.21)$$

На рисунке 6.6 приведены графики частотной характеристики фильтра Баттервортта при  $r = 70$  и  $k = 10$ .



а) б)

Рисунок 6.6 График частотной характеристики фильтра Баттервортта при  $r=70$  и  $k=10$ : а) – центрированная характеристика; б) – нецентрированная.

Частотная характеристика гауссовского НЧ фильтра имеет вид:

$$H(u, v) = \exp(-r^2(u, v) / (2\sigma^2)), \quad (6.22)$$

где  $u$  - имеет смысл частоты среза ( $r_{\text{ср}}$ ).

Частотная фильтрация выполняется по алгоритму:

- выполнить двумерное ДПФ входного изображения  $f(x, y)$  (подвергаемого фильтрации) размером  $(N \times M)$ , получить  $F(u, v)$ ;
- вычислить передаточную характеристику фильтра в частотной области, например, в соответствии с одной из формул (6.20)-(6.22),

размер матрицы ( $N \times M$ ); выполнить децентрирование характеристики  $H(u, v)$ ;

- выполнить поточечное умножение  
 $S(u, v) = F(u, v)H(u, v), \forall (u \in [0, N - 1], v \in [0, M - 1]);$
- выполнить ОДПФ спектра  $S(u, v)$ .

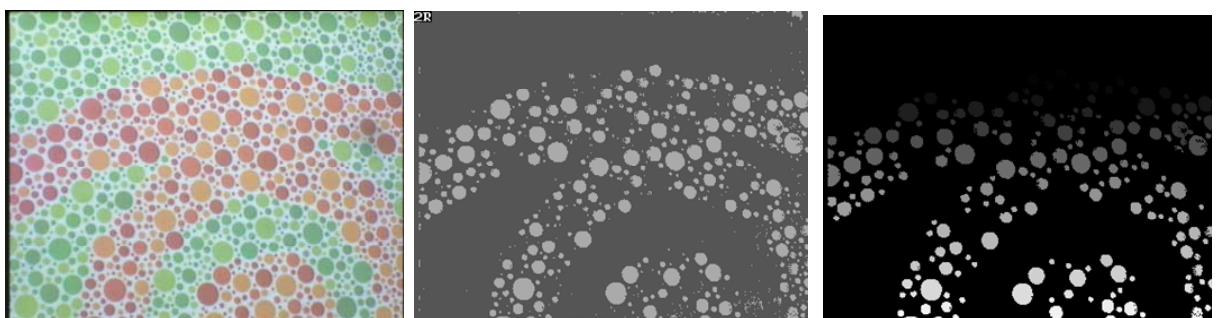
#### 6.2.4 Высокочастотные фильтры

По известной передаточной функции НЧ фильтра можно получить передаточную характеристику ВЧ фильтра в соответствии с уравнением:  
 $H_{ВЧ}(u, v) = 1 - H_{НЧ}(u, v).$  (6.23)

Фильтрация выполняется в соответствии с алгоритмом, приведенным в п. 6.2.3.

### 7 АНАЛИЗ БИНАРНЫХ ИЗОБРАЖЕНИЙ

Анализ бинарных изображений включает операции, применяемые в бинарном машинном зрении, например, анализ частиц, обработка документов, счет рыб, и при анализе полутоновых изображений. Он предполагает формирование и обработку бинарных изображений, обычно изображений, яркости элементов которых имеют значения 0 или 1, при этом 0 представляет фон, 1 представляет объект. На рисунке 7.1 представлено изображение клеток, окрашенных в разные цвета: желтый, оранжевый, светло-зеленый и темно-зеленый.



а) б) в)

Рисунок 7.1 Пример операций анализа цветного изображения. а) Исходное изображение. Каждая клетка является отдельным объектом. б) Бинарное изображение клеток оранжевого цвета после сегментации. в) Результат селекции связных компонентов.

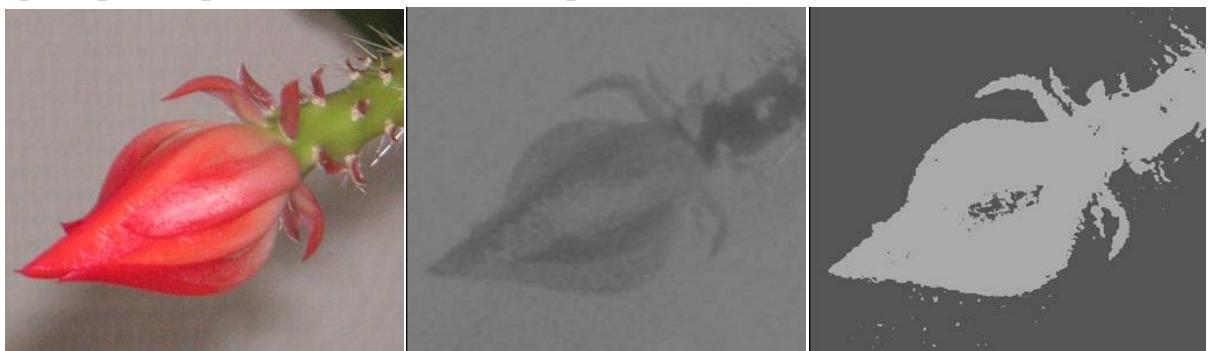
По признаку цветности из этого изображения выделяются объекты заданного (на рисунке 7.1 -оранжевого) цвета. В результате сегментации формируется бинарное изображение, на котором наряду с полезной информацией остаются области шума (рисунок 7.1 б). В результате селекции связных областей каждой связной области назначается определенный уровень яркости (номер области). После выполнения селекции и выделения областей в некотором заданном диапазоне площадей удается отфильтровать шум и сосчитать клетки заданного цвета, пронумеровав их, как это показано на рисунке 7.1 в).

## 7.1 Назначение основных операций бинарного анализа.

Первая операция бинарного анализа состоит в выделении объектов на фоне из множества других объектов. Только после этого выполняется объединение элементов изображения каждого объекта, как это показано на рисунке 7.1 б). После получения бинарного изображения объекта возможно вычисление признаков объекта.

## 7.2 Состав основных операций бинарного анализа

В соответствии с назначением операций бинарного анализа можно определить состав его основных операций. К ним относится, прежде всего, выделение признаков, по которым можно выделить объект на фоне. В связи с этим перед бинарными операциями выполняются полутоновые преобразования, позволяющие усилить признаки объекта. Следующая задача связана с определением порогов для выполнения бинаризации. Пороговая бинаризация полутоновых изображений в соответствии с амплитудными характеристиками, рассмотренными в п. 5.3, позволяет получить бинарное изображение объекта. На рисунке 7.2 представлен пример бинаризации цветного изображения.



а)б)в)

Рисунок 7.2 а) Изображение цветка; б) цветоразностный (*Cb*) компонент этого изображения; в) бинарное изображение цветка.

Поскольку отличительной характеристикой объекта в этом примере является красный цвет цветка и зеленый цвет его ножки, бинаризация может производиться по *Cb* компоненту. На рисунке 7.3 представлена бимодальная гистограмма значений *Cb* компонента исходного изображения. Одна мода соответствует цветку (значения  $\leq 115$ ), другая – фону.

Бинарное изображение неизбежно включает области, обусловленные помехами, поэтому разработаны операции, позволяющие исключить или уменьшить влияние помехи. Эти операции относятся к морфологическим операциям бинарных изображений. Они позволяют разделить соединенные объекты или соединить разорванные части объекта, заполнить дырки внутри областей объекта, исключить отдельные шумовые объекты. После выполнения таких операций по полученному бинарному изображению выполняется операция, называемая селекцией связных компонентов [47],

позволяющая представить каждый отдельный пространственно связанный объект номером.

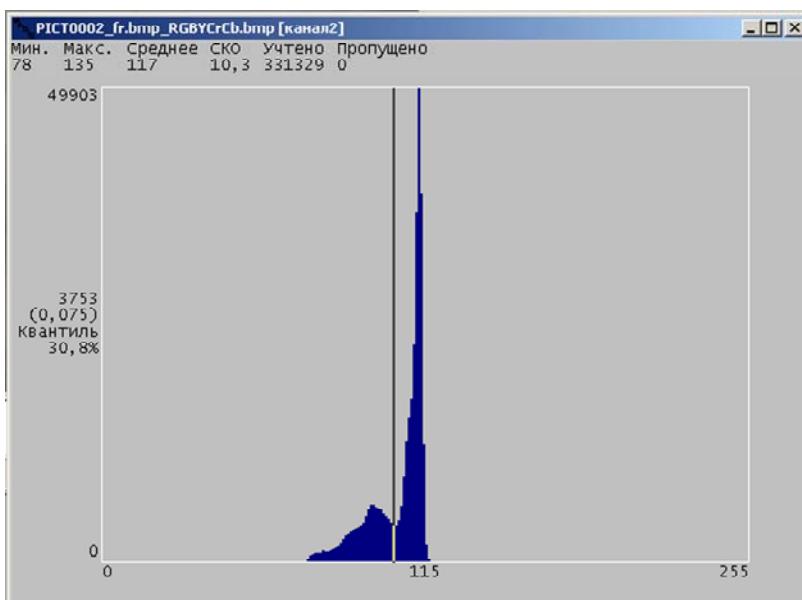


Рисунок 7.3 Гистограмма значений *Cb* компонента.

По бинарному изображению можно определить и проанализировать каждую связную область. Входным изображением для операции «селекция связных компонентов» является бинарное изображение. На выходе формируется изображение меток областей. Каждый отсчет изображения имеет значение номера связной области, которой он принадлежит, либо значение, равное нулю, если элемент принадлежит фону. Операция позволяет исключить из рассмотрения области, имеющие площади, находящиеся вне заданных пределов интересующих размеров объектов. В обработке изображений это приводит к формированию полутонаового изображения, каждый уровень яркости которого соответствует номеру связной области (объекта). По изображению связных компонентов и исходному полутональному изображению осуществляется оценка признаков изображения. В приведенном примере обнаружено 202 объекта. Объекты с площадью, меньшей 17 элементов, исключены как принадлежащие шуму. Наконец, последней операцией, необходимой для выполнения бинарного анализа, является оценка признаков объекта. На рисунке 7.4 представлен фрагмент яркостных и геометрических характеристик выделенных оранжевых клеток, изображенных на рисунке 7.1.в).

К геометрическим характеристикам относятся:  $X_{ц.т.}, Y_{ц.т.}$  - координаты центра тяжести объекта, Площ. – площадь объекта, Перим. – периметр объекта, Коэффиц. форм - коэффициент формы, оцениваемый как отношение квадрата периметра области объекта к ее площади,  $D_{min}$ ,  $D_{max}$ , Эксц., Угол – характеристики эллипса рассеяния: размер его малой и большой осей, эксцентриситет, оцениваемый как отношение малой оси эллипса к большой, и угол наклона большой оси. Яркостные характеристики представлены значением средней яркости по области и СКО яркости в этой области.

Геометрические признаки												
Метка	Хц.т.	Үц.т.	Плош.	Перим.	Коэф.форн	Дмин	Дмакс	Эксц.	Угол	Ср.ярк.	СКО	
4	280,00	62,00	63,00	27	11,57	8	10	0,85	66	192,62	8,06	
5	213,00	65,00	58,00	31	16,57	8	10	0,80	39	199,91	9,32	
6	255,00	66,00	73,00	30	12,33	9	11	0,81	39	195,47	9,28	
7	239,00	65,00	24,00	13	7,04	5	6	0,81	70	196,00	8,25	
8	269,00	69,00	48,00	21	9,19	7	8	0,90	70	173,08	9,49	
9	293,00	69,00	47,00	20	8,51	7	9	0,81	59	171,57	10,56	
10	322,00	72,00	97,00	29	8,67	10	12	0,84	49	169,06	10,00	
11	191,00	70,00	29,00	18	11,17	5	7	0,70	33	195,59	8,15	
12	101,00	77,00	113,00	33	9,64	12	12	0,92	22	175,01	7,99	
13	179,00	76,00	79,00	31	12,16	10	10	0,94	0	197,48	8,61	
14	342,00	75,00	20,00	12	7,20	4	6	0,77	43	154,40	11,97	
15	201,00	79,00	104,00	31	9,24	11	13	0,83	32	180,82	7,88	
16	253,00	77,00	50,00	20	8,00	7	9	0,83	42	173,04	9,25	

Рисунок 7.4 Оценка яркостных и геометрических признаков формы клеток.

### 7.3 Морфология бинарных изображений

Морфология предназначена для представления и описания свойств формы и структуры объектов, представленных областями изображения [28]. Математическая морфология определена в виде операций над множествами. Морфология бинарных изображений предназначена для обработки областей, полученных после выполнения квантования изображения на два уровня. Морфология бинарных изображений состоит из двух основных операций: дилатации (расширения), эрозии (уменьшения) и производных от них операций, замыкания (*closing*) и размыкания (*opening*).

Входными данными для операций морфологии являются бинарное изображение **B** и структурный элемент **S** (называемый также структурирующим элементом или примитивом), представляющий другое (меньшее) бинарное изображение. Оба множества являются подмножествами двумерного целочисленного пространства  $Z^2$  с элементами в виде двумерных векторов,  $(x, y)$  координаты которых указывают на белый (равный 1) элемент изображения. Форма области примитива определена единичными отсчетами в прямоугольной области, что позволяет формировать произвольную форму области. В морфологии определены следующие операции:

**Параллельный перенос.** Параллельный перенос  $\mathbf{B}_t$  множества пикселов **B** на вектор  $t = (x, y)$  задается в виде

$$\mathbf{B}_t = \{c | c = b + t, b \in \mathbf{B}\} \quad (7.1)$$

и предназначен для сдвига всех элементов множества **B** на заданное расстояние  $t$ , как показано на рисунке 7.5 при  $t = (x_0, y_0)$ .

**Центральное отражение множества B** представляет собой поворот множества **B** на  $180^\circ$  относительно начала координат (рисунок 7.6). Обозначается такое множество **B̄**.

$$\tilde{\mathbf{B}} = \{c | c = -b, b \in \mathbf{B}\}. \quad (7.2)$$

**Дилатация** множества  $\mathbf{B}$  по множеству  $\mathbf{S}$  задается в виде

$$\mathbf{B} \oplus \mathbf{S} = \{t | (\mathbf{S})_t \cap \mathbf{B} \neq \emptyset\}. \quad (7.3)$$

Дилатация выполняется следующим образом. Прежде всего, осуществляется центральное отражение примитива  $\mathbf{S}$ , формируется  $\tilde{\mathbf{S}}$ . Затем последовательно выполняются две операции:

- перенос примитива  $\tilde{\mathbf{S}}$  в точку  $t$ ;
- оценка выходного значения. Значение элемента выходного массива с координатами  $t$  равно 1, если единичные значения отсчетов  $(\tilde{\mathbf{S}})_t$  и  $\mathbf{B}$  совпадают хотя бы в одном элементе, равно 0 в противном случае.

Значение  $t$  последовательно изменяется в соответствии с координатами всех элементов массива  $\mathbf{B}$ , при этом центральное отражение примитива скользит по изображению  $\mathbf{B}$  подобно тому, как это происходит при свертке изображения с импульсной характеристикой фильтра. Таким образом формируется множество всех элементов  $t$ , для которых пересечение  $(\tilde{\mathbf{S}})_t$  и  $\mathbf{B}$  не является пустым множеством.

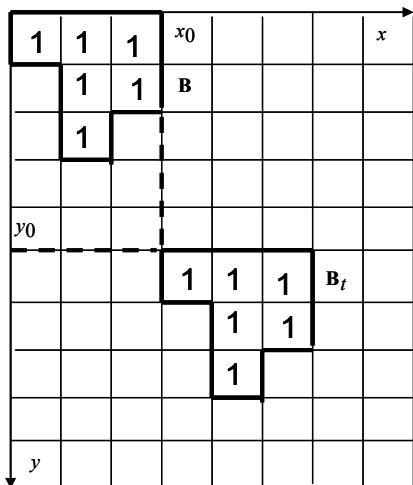


Рисунок 7.5 Параллельный перенос множества  $\mathbf{B}$  на вектор  $t = (x_0, y_0)$ .

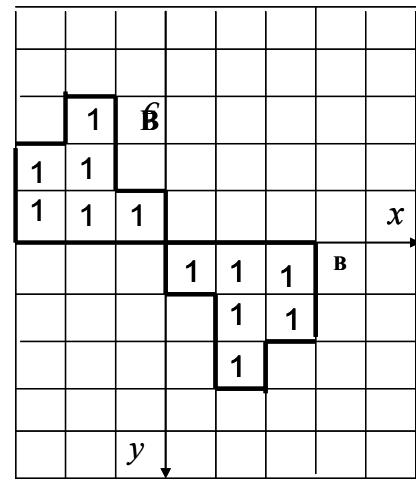


Рисунок 7.6 Центральное отражение множества  $\mathbf{B}$ .

Дилатация расширяет связанные единичные области бинарных изображений. Она используется для расширения областей, для заполнения внутренних «дыр» и устранения заливов вдоль границы области. Результат дилатации  $\mathbf{B} \oplus \mathbf{S}$  представлен на рисунке 7.7.

**Эрозия** бинарного изображения  $\mathbf{B}$  по структурному элементу  $\mathbf{S}$  обозначается  $\mathbf{B} \ominus \mathbf{S}$  и равна

$$\mathbf{B} \ominus \mathbf{S} = \{t | (\mathbf{S})_t \subseteq \mathbf{B}\}. \quad (7.4)$$

Структурный элемент скользит по всем элементам входного изображения. Если каждый единичный элемент  $\mathbf{S}$  совпадает с единичным элементом  $\mathbf{B}$ , то элементу выходного массива с координатами,

соответствующими положению центра  $S_t$ , присваивается значение 1. Таким образом формируется множество всех элементов  $t$ , для которых  $S_t$  полностью содержится в  $\mathbf{B}$ . Пример эрозии приведен на рисунке 7.8. Из него видно, что эрозия приводит к сжатию изображения области.

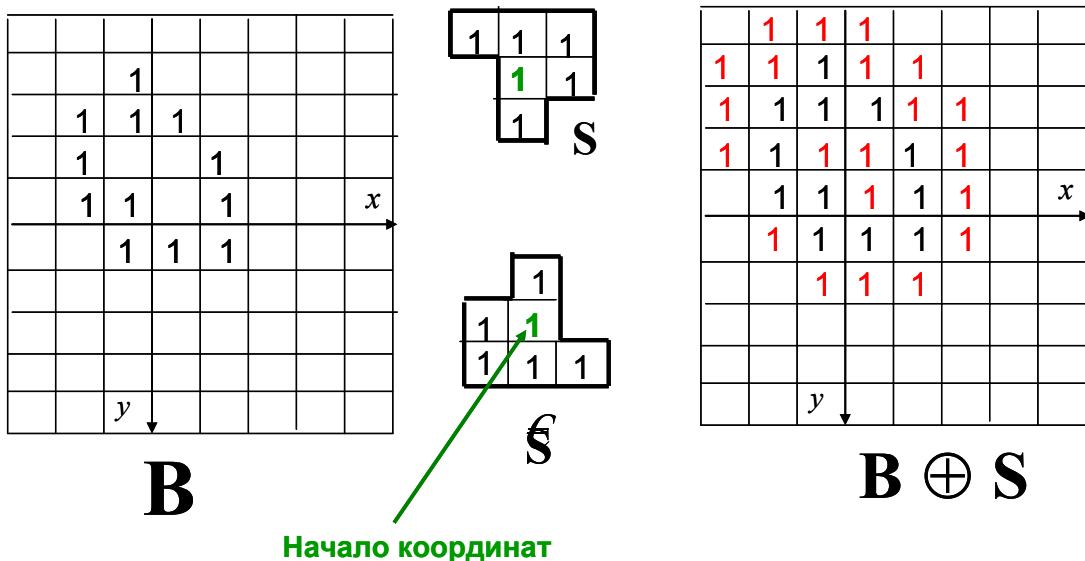


Рисунок 7.7 Дилатация множества  $\mathbf{B}$  по множеству  $\mathbf{S}$ .

Она может применяться для исключения областей, имеющих размер меньший, чем размер примитива, для уменьшения областей и удаления узких «мостов» между областями. Эрозия по примитиву, размер которого меньше размера объекта, который мы хотим сохранить, но больше размера подавляемого объекта, позволяет сохранить интересующие объекты, а последующая дилатация по этому же примитиву позволяет восстановить объекты в прежних размерах.

**Открытие** (размыкание) представляет собой последовательно выполненные две морфологические операции эрозию и дилатацию множества  $\mathbf{B}$  по примитиву  $\mathbf{S}$ :

$$\mathbf{B} \circ \mathbf{S} = (\mathbf{B} \ominus \mathbf{S}) \oplus \mathbf{S}. \quad (7.6)$$

Из рисунка 7.9 видно, что размыкание позволяет устраниТЬ узкие перешейки и выступы небольшой ширины, сглаживая контуры области, не вызывая общего уменьшения размера объекта.

**Закрытие** (замыкание) представляет собой последовательно выполненные две морфологические операции дилатацию и эрозию:

$$\mathbf{B} \bullet \mathbf{S} = (\mathbf{B} \oplus \mathbf{S}) \ominus \mathbf{S}. \quad (7.5)$$

В результате последовательного применения дилатации и эрозии, представленных на рисунках 7.7, 7.8 выполнено заполнение «дырки» внутри области (множество  $\mathbf{B}$ ) без изменения внешних границ этой области. Другой пример замыкания представлен на рисунке 7.10.

Изменение площади объектов при выполнении морфологических операций может быть упорядочено следующим образом:

$$\mathbf{B} \ominus \mathbf{S} \subseteq \mathbf{B} \circ \mathbf{S} \subseteq \mathbf{B} \subseteq \mathbf{B} \bullet \mathbf{S} \subseteq \mathbf{B} \oplus \mathbf{S}.$$

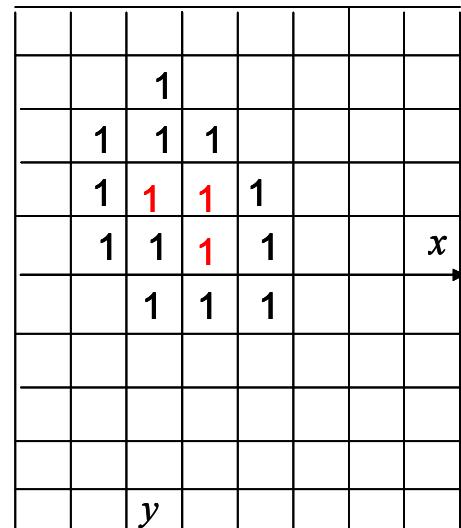
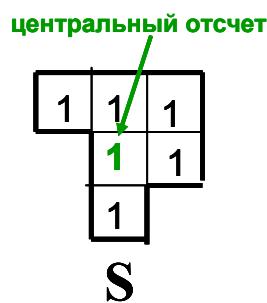
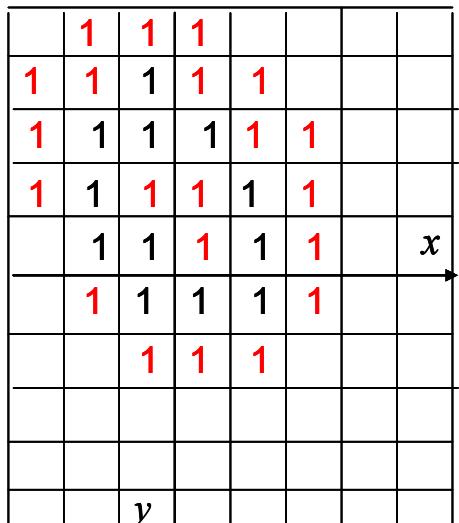


Рисунок 7.8 Эрозия множества **B** по примитиву **S**.

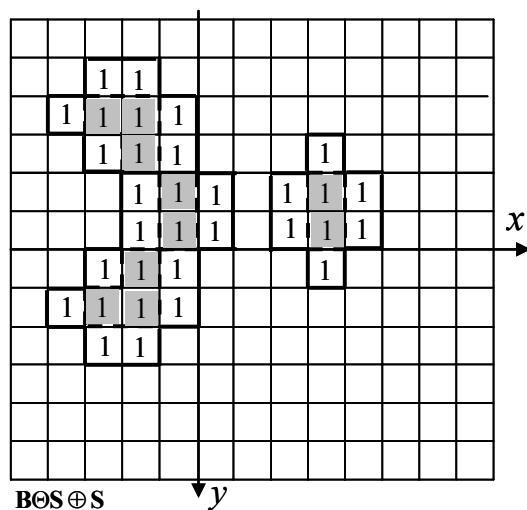
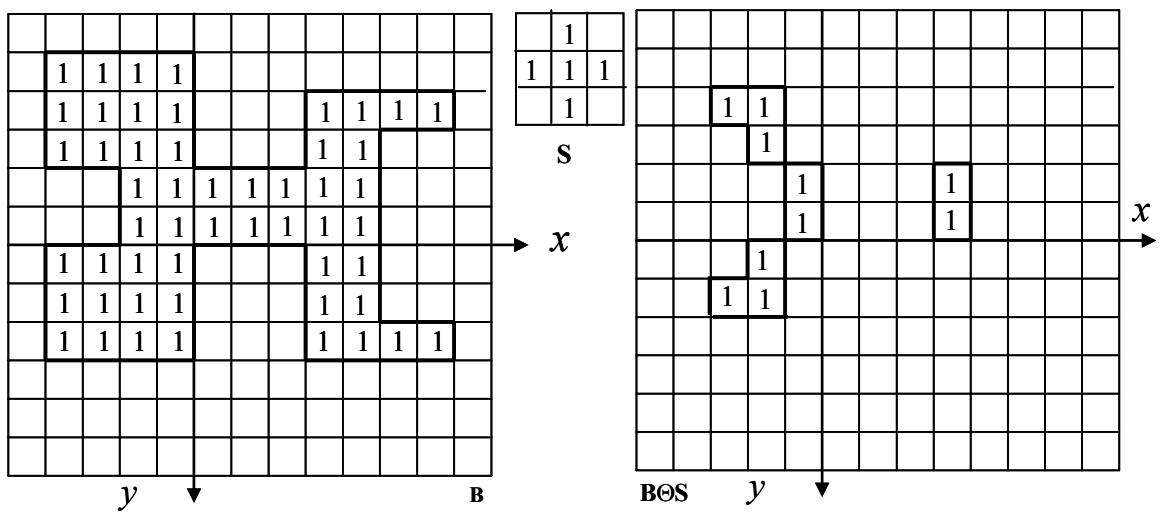
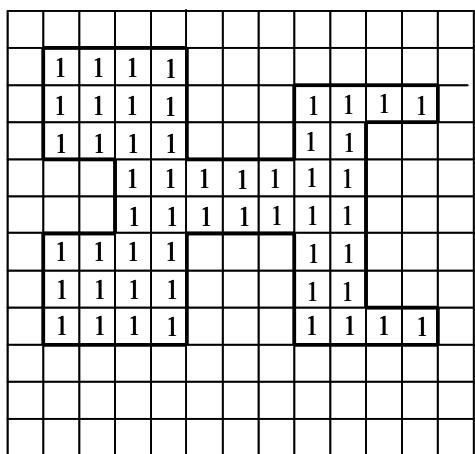
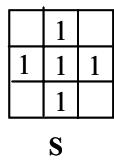


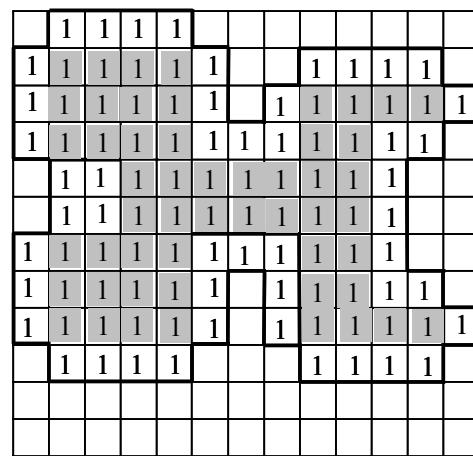
Рисунок 7.9 Размыкание множества **B** по примитиву **S**.



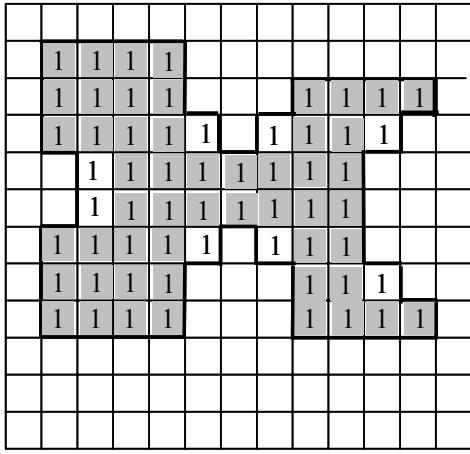
B



S



BOS



$$(B \oplus S)\Theta S$$

Рисунок 7.10 Замыкание множества **B** по примитиву **S**.

## 7.4 Признаки областей

После выполнения селекции связных областей по бинарному изображению, позволяющей промаркировать каждую область, формируются признаки этих областей. К ним относятся следующие признаки: геометрические признаки, характеристика формы; яркостные характеристики; цветовые характеристики; текстурные свойства; признаки движения.

К геометрическим характеристикам относятся такие характеристики, как площадь, центр тяжести, координаты описывающего прямоугольника и периметр [48]. Площадь равна количеству элементов изображения, принадлежащих области. Обозначим множество единичных отсчетов  $g(x,y)$  с координатами  $(x,y)$ , принадлежащих области  $A$ , заданной в плоскости изображения как:

$$g(x,y) = \begin{cases} 1, & (x,y) \in A \\ 0, & \text{иначе} \end{cases}$$

Пусть координаты левого верхнего и правого нижнего углов прямоугольника, описывающего область, равны  $(X_{\min}, Y_{\min})$  и  $(X_{\max}, Y_{\max})$  соответственно. Тогда площадь вычисляется в соответствии с формулой:

$$S = \sum_{y=Y_{\min}}^{Y_{\max}} \sum_{x=X_{\min}}^{X_{\max}} g(x, y). \quad (7.7)$$

Центр тяжести области задается координатами  $(X_c, Y_c)$ , определяемыми как среднее значение  $(x, y)$  координат, принадлежащих области в соответствии с уравнением:

$$\begin{aligned} X_c &= \frac{1}{S} \sum_{(x, y) \in A} x \\ Y_c &= \frac{1}{S} \sum_{(x, y) \in A} y \end{aligned} \quad (7.8)$$

Определение координат центра тяжести объекта позволяет нормализовать положение объекта, определив положение начала координат в плоскости изображения, относительно которого положение объекта является центральным.

Если число граничных отсчетов области равно  $N$ , то длина периметра  $P$  равна сумме расстояний между соседними граничными точками:

$$P = \sum_{i=1}^N r_i, \quad (7.9)$$

$$r_i = \sqrt{(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2}. \quad (7.10)$$

Отсчет является граничным, если хотя бы один отсчет из его окрестности не принадлежит области  $A$ .

В качестве характеристики формы используют оценку коэффициента формы, определяемую как отношение квадрата периметра к площади:

$$K = \frac{P^2}{S}. \quad (7.11)$$

Для оценки округлости области наряду с (7.11) используется коэффициент, определяемый в соответствии с выражением:

$$C = \frac{m_A}{\sigma_A}, \quad (7.12)$$

где  $m_A$  - среднее значение расстояний от центра тяжести области до граничных отсчетов;  $\sigma_A$  - СКО этих расстояний.

$$m_A = \frac{1}{N} \sum_{i=1}^N r_{ic}, \quad (7.13)$$

где  $r_{ic}$  - расстояние от  $i$ -того граничного отсчета до центра тяжести области, определяемое в соответствии с (7.10).

$$\sigma_A = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (r_{ic} - m_A)^2}. \quad (7.14)$$

Для нормализации ориентации объекта при анализе бинарных изображений используется статистический подход. Объект описывают некоторым эллипсом рассеяния [49]. В качестве ориентации выбирают

направление собственного вектора  $\mathbf{x}$  матрицы ковариации  $\mathbf{B}$  координат отсчетов ненулевой яркости, то есть принадлежащих области  $A$ . Собственный вектор должен соответствовать максимальному собственному значению  $\lambda$  матрицы ковариации. Пусть матрица ковариации имеет вид:

$$\mathbf{B} = \begin{pmatrix} B_{11} & B_{12} \\ B_{12} & B_{22} \end{pmatrix},$$

где  $B_{ij}$  - центральные моменты второго порядка:  $B_{11}$  - дисперсия  $x$ -координаты ненулевых отсчетов яркости,  $B_{22}$  - дисперсия  $y$ -координаты ненулевых отсчетов яркости,  $B_{12}$  - ковариация  $(x,y)$ -координат ненулевых отсчетов яркости.

Возможные собственные значения  $\lambda$  находятся из уравнения  $(\mathbf{B} - \lambda \mathbf{E}) \mathbf{x}_\lambda = 0$ , (7.15)

где  $\mathbf{E}$  - единичная матрица,  $\mathbf{x}_\lambda$  - собственный вектор, соответствующий числу  $\lambda$ .

Значения  $\lambda$  определяем из уравнения

$\mathbf{B} - \lambda \mathbf{E} = 0$ :

$$\begin{vmatrix} B_{11} & B_{12} \\ B_{12} & B_{22} \end{vmatrix} - \lambda \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 0$$

$$\begin{vmatrix} B_{11} - \lambda & B_{12} \\ B_{12} & B_{22} - \lambda \end{vmatrix} = 0$$

$$(B_{11} - \lambda)(B_{22} - \lambda) - B_{12}^2 = 0;$$

$$\lambda_{1,2} = \frac{B_{11} + B_{22}}{2} \pm \sqrt{\left(\frac{B_{11} + B_{22}}{2}\right)^2 - B_{11}B_{22} + B_{12}^2}. \quad (7.16)$$

Собственные векторы находятся из уравнения (7.15):

$$\begin{vmatrix} B_{11} - \lambda_1 & B_{12} \\ B_{12} & B_{22} - \lambda_1 \end{vmatrix} \times \begin{vmatrix} x_1 \\ x_2 \end{vmatrix} = 0;$$

$$(B_{11} - \lambda_1)x_1 + B_{12}x_2 = 0;$$

$$\begin{aligned} x_1 &= -B_{12}; \\ x_2 &= B_{11} - \lambda_1; \end{aligned} \quad (7.17)$$

$$\operatorname{tg} \varphi = \frac{x_2}{x_1}.$$

Размеры полуосей эллипса определяются следующим образом. Определяется соотношение собственных чисел матрицы ковариации (полуосей эллипса):

$$k = \sqrt{\frac{\operatorname{abs}(\lambda_2)}{\lambda_1}},$$

где  $\lambda_1$  - большее собственное значение,  $\lambda_2$  - меньшее собственное значение.

Площадь эллипса  $\text{Square} = \pi ab$ . Отношение малой и большой полуосей эллипса равно

$$a/b = k. \quad (7.18)$$

Тогда  $\text{Square} = \pi kb^2$ . Большая полуось эллипса равна:

$$b = \sqrt{\text{Square}/\pi k},$$

где  $\text{Square}=S$  - площадь бинарного изображения (количество отсчетов с ненулевой яркостью).

Малая полуось эллипса определяется в соответствии с уравнением (7.18):  $a=kb$ .

По бинарному изображению оцениваются параметры аппроксимирующего эллипса (в соответствии с рисунком 7.11): малая полуось  $a$  (на рисунке 7.4 обозначено  $D_{\min} = 2a$ ), большая полуось  $b$  (на рисунке 7.4  $D_{\max} = 2b$ ) и угол наклона большой оси эллипса  $\varphi$  (на рисунке 7.4 - «Угол») в соответствии с описанным статистическим подходом к нормализации бинарного изображения объекта.

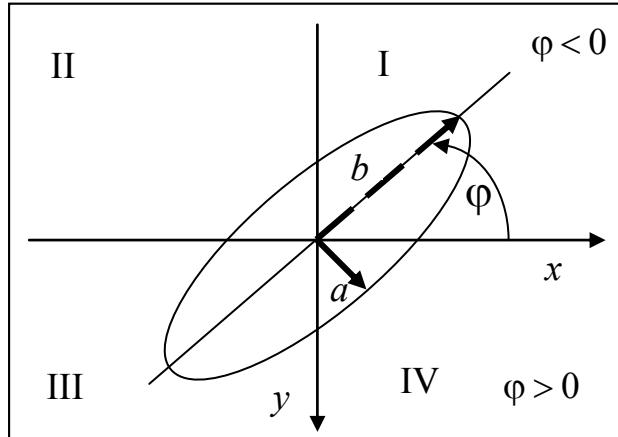


Рисунок 7.11 Определение ориентации по собственному вектору матрицы ковариации.

## 8 МЕТОДЫ КОДИРОВАНИЯ ЦИФРОВЫХ ИЗОБРАЖЕНИЙ

Рассмотрим методы кодирования цифровых изображений, позволяющие уменьшить объем данных, необходимый для представления изображений. Для передачи цифрового сигнала изображения необходимы каналы с высокой скоростью передачи. Повышенный интерес к улучшению алгоритмов сжатия вызван необходимостью передачи больших объемов информации по каналам связи и обеспечением эффективности хранения и защиты информации. Переход на цифровые технологии позволяет эффективно применять современные методы стеганографии.

Поскольку получателем информации является наблюдатель, то и методы кодирования основаны на учете особенностей зрительной системы человека. Данные психологии и физиологии зрения служат основой для постановки не только задач улучшения изображений, но и их кодирования

[50]. Методы кодирования строятся по принципу устранения избыточности информации. Остановимся на этих аспектах проблемы подробнее.

## 8.1 Особенности зрительной системы человека

Свойства зрительной системы (ЗС) человека можно разделить на 2 группы: низкоуровневые («физиологические») и высокоуровневые («психофизиологические»). Вплоть до середины 90-х годов прошлого века исследователи принимали во внимание, главным образом, низкоуровневые свойства зрения. В последние годы наметилась тенденция построения моделей с учетом и высокоуровневых характеристик ЗС. Выделим три наиболее важных низкоуровневых свойства, влияющих на заметность шума в изображении: чувствительность к изменению яркости изображения; частотная чувствительность; эффект маскирования. Отличием высокоуровневых свойств от низкоуровневых является то, что эти свойства проявляются «вторично». Мозг, обработавший первичную информацию от ЗС, выдает команды на ее «подстройку» под изображение. Основные свойства ЗС:

- Чувствительность к контрасту. Высококонтрастные участки изображения, перепады яркости обращают на себя значительное внимание.
- Чувствительность к размеру. Большие области изображения «заметнее» областей меньшего размера. Причем существует порог насыщения, когда дальнейшее увеличение размера не существенно.
- Чувствительность к форме. Длинные и тонкие объекты обращают на себя большее внимание, чем круглые и однородные.
- Чувствительность к цвету. Некоторые цвета, например красный, заметнее других. Этот эффект усиливается, если фон заднего плана отличен от цвета фигур на нем.
- Чувствительность к местоположению. Человек склонен в первую очередь рассматривать центр изображения. Люди обычно внимательнее к изображениям переднего плана, чем заднего. Если на изображении есть люди, в первую очередь человек обратит внимание на них. На фотографии человек обращает первоочередное внимание на лицо, глаза, руки, рот.
- Чувствительность к внешним раздражителям.

Движение глаз наблюдателя зависит от конкретной обстановки, от полученных им перед просмотром или во время него инструкций, дополнительной информации [15, 51]. ЗС может отнести к одному классу объекты, отличающиеся по ориентации, размеру, местоположению на сетчатке, имеющие разную освещенность и цвет. Это свойство инвариантности к различного рода преобразованиям дополняется системой анализа локальных признаков, обусловливающих многоканальность ЗС. Многочисленные клинические, психофизические и поведенческие исследования показали, что, например, форма объекта может опознаваться независимо от размера, цвета и местоположения объекта в поле зрения. Каждый компонент имеет следующие характеристики: пространственное

положение в изображении, частота изображения, ориентация сигнала (горизонтальное, вертикальное, диагональное направления).

– *Острота зрения.* Острота зрения определяется размером минимального элемента тестового объекта, от которого зависит опознание формы этого объекта. Острота зрения зависит от контраста изображения, яркости фона и времени предъявления теста. Ограничения разрешения ЗС зависят от оптических свойств глаза и сетчатки. Астигматизм приводит к зависимости разрешения от ориентации объекта. Сферические и хроматические aberrации влияют на фокусировку изображения. Дифракция света на границах зрачка приводит к размытию границ объектов. Для оценки остроты зрения пользуются критерием Рэлея, согласно которому две точки или линии воспринимаются как отдельные объекты, если дифракционный максимум одного изображения объекта совпадает с первым дифракционным минимумом другого изображения. Острота зрения определяется как минимальный угол разрешения. Типичными объектами для измерения разрешения являются две точки, две линии, решетки из параллельных полос, образцы типа шахматной доски. При оценке остроты с помощью колец Ландольта, нормальной остроте зрения соответствует различие разрыва, равного 1 угловой минуте. На сетчатке одна угловая минута превышает размер одной колбочки, величина которой в центре фовеолы составляет 0,4-0,6 угловой минуты, следовательно, она ограничена дифракцией, а не размером колбочки. На остроту зрения большое влияние оказывает неоднородность сетчатки глаза. Колбочки в центре более плотно упакованы, палочки отсутствуют. По мере удаления от центра к периферии расстояние между колбочками увеличивается, между ними располагается все больше и больше палочек. Однако на дальней периферии располагается слой колбочек, функциональное назначение которых не определено. Такая структура сетчатки обусловливает уменьшение остроты зрения от центра к периферии. Этим объясняется, что фовеола, геометрически связанная с центром поля зрения, обеспечивает распознавание мелких деталей изображения.

– *Контрастная маскировка.* Обнаружительная способность одного сигнала в присутствии другого сигнала.

– *Чувствительность к частоте.* Максимальная чувствительность ЗС наблюдается в области средних частот от 2 до 10 циклов/градус. Чувствительность в области ВЧ ограничивается оптикой глаза и дискретностью фоторецепторов сетчатки. В области НЧ снижение чувствительности вызвано особенностями нервной системы: размером рецептивных полей и торможением при действии стимула.

– *Чувствительность к яркости.* Порог обнаружительной способности шума на постоянном фоне.

Остановимся подробнее на некоторых особенностях ЗС человека.

### 8.1.1 Контрастная чувствительность

Порог "минимальной заметной разницы" (JND), или дифференциальный порог, определяется в соответствии с законом Вебера-Фехнера, формулируемым следующим образом: субъективная величина ощущения, измеренная в единицах минимальной заметной разницы, пропорциональна логарифму физической величины стимула:

$$\Delta I / I = \text{const.} \quad (8.1)$$

Закон утверждает: чтобы яркость одного светового пятна была чуть-чуть светлее яркости другого светового пятна, она должна быть больше на некоторую постоянную величину. Если яркость одного пятна равна  $I$ , а другого  $I + \Delta I$ , где  $\Delta I$  едва заметное различие, то отношение  $\Delta I / I$  будет постоянным. Отношение (8.1) называется пороговым контрастом, а  $\Delta I$  - дифференциальным порогом.  $\Delta I$  - это порог, внутри которого любые изменения соответствующего ощущения невидимы. Эксперимент по оценке контрастной чувствительности (также называемой чувствительностью к изменению яркости) строится следующим образом. Испытуемому предъявляют некоторую однотонную картинку (в соответствии с рисунком 8.1а). После того как глаз адаптировался к яркости  $I_0$ , постепенно изменяют яркость вокруг центрального пятна. Изменение яркости  $\Delta I$  продолжают до тех пор, пока оно не будет обнаружено. На рисунке 8.1 б) показана зависимость порогового контраста от яркости фона  $I$ . Для яркостей в среднем диапазоне контраст приблизительно постоянен, в то время как для малых и больших яркостей

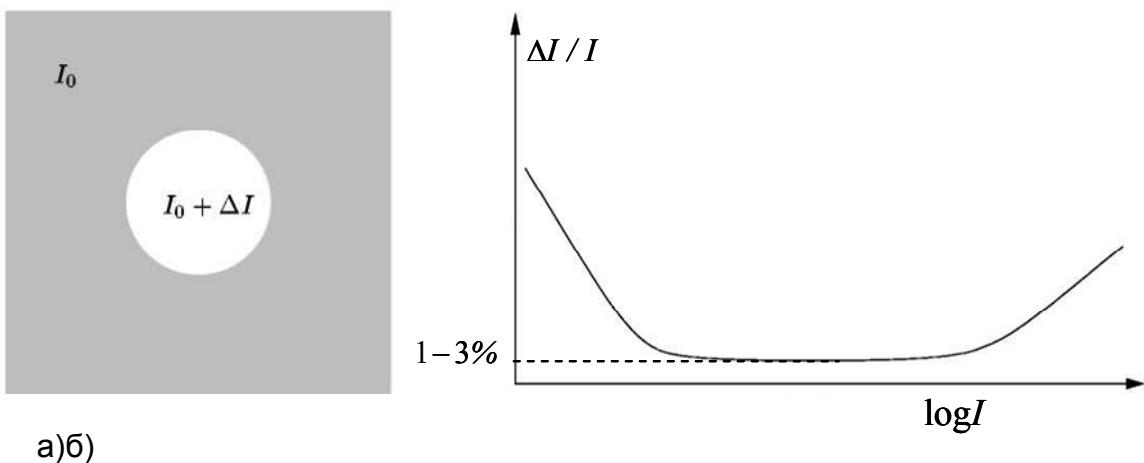


Рисунок 8.1 Чувствительность к контрасту и порог неразличимости  $\Delta I$ .

значение порога неразличимости возрастает. Приблизительно постоянное отношение в центре и называют отношением Вебера-Фехнера. Отношение Вебера-Фехнера, как установлено, составляет от 1 до 3%. Это означает, что изменение освещенности для средних значений яркости составляет приблизительно от 0,01 до 0,03 от яркости фона.

### 8.1.2 Пространственная чувствительность частоты

На чувствительность ЗС существенное влияние оказывает также пространственная частота ("форма"). Частотная чувствительность ЗС

проявляется в том, что человек гораздо более восприимчив к низкочастотному шуму, чем к высокочастотному.

Это связано с неравномерностью амплитудно-частотной характеристики ЗС, что приводит к различным явлениям, например сверхчувствительности в высококонтрастных областях и особенно на границах. Одно явление, названное эффектом полос Maxa, проиллюстрировано на рисунке 8.2.



Рисунок 8.2 Эффект полос Maxa.

Изображение представляет горизонтально изменяющуюся яркость с постоянным шагом. Хотя каждый шаг имеет однородную яркость, мы чувствуем неравномерную интенсивность в шаге, то есть более яркую слева и более темную справа у границ на каждом шаге.

### 8.1.3 Маскирование

Элементы ЗС разделяют поступающий видеосигнал на отдельные компоненты. Каждая составляющая возбуждает нервные окончания глаза через ряд подканалов. Выделяемые глазом компоненты имеют различные пространственные и частотные характеристики, а также различную ориентацию (горизонтальную, вертикальную, диагональную) [51].

В случае одновременного воздействия на глаз двух компонентов со сходными характеристиками возбуждаются одни и те же подканалы. Это приводит к эффекту маскирования, то есть увеличению порога обнаружения видеосигнала в присутствии другого сигнала, обладающего аналогичными характеристиками. Например, аддитивный шум намного более заметен в однородных областях изображения, чем в текстурированных.

### 8.1.4 Статистическая и визуальная избыточность изображений

Малоразмерные детали, имеющие малый контраст, вследствие недостаточной контрастной чувствительности зрения, не видны на изображении. Этот вид избыточности называется *психофизической избыточностью изображений*. Известно, что пороговый контраст зрения в сильной степени зависит от угловых размеров наблюдаемых объектов. Так, например, при наблюдении крупных объектов пороговый контраст составляет около 0,02, а при наблюдении точечных объектов он составляет

не более 0,1. Поэтому для представления малоразмерных объектов можно использовать меньшее число уровней квантования, чем для протяженных.

Внутрикадровая и межкадровая статистическая избыточность изображений используются при разработке алгоритмов кодирования изображений. Коэффициент сжатия определяется как отношение размера исходного изображения к размеру сжатого потока. Методы сжатия делятся на *методы сжатия без потерь информации* и *методы сжатия с потерями информации*. Метод сжатия изображения без потерь – это метод, при котором не производится никакой потери качества изображения по сравнению с исходным. Несжатое изображение математически идентично его оригиналу. Сжатие без потерь обычно обеспечивает меньшие коэффициенты сжатия, чем сжатие с потерями. Методы сжатия изображений «с потерями» - это методы, при которых жертвуют некоторым качеством изображения в обмен на уменьшение размера данных. Количество ухудшения зависит от используемого алгоритма сжатия и заданного пользователем коэффициента качества. Они основаны на устраниении избыточности изображений в соответствии со свойствами ЗС человека. При применении этих методов, декодированное изображение отличается от исходного изображения, то есть имеет место искажение изображения и соответственно, потеря информации.

Работы в области создания методов сжатия данных для кодирования неподвижных изображений были начаты в 50-х годах двадцатого столетия. Среди первых удачных разработок, сделанных в то время, следует назвать метод кодирования длин серий или *RLE* (*Run - Length Encoding*), являющийся методом сжатия без потерь информации, и метод дифференциальной импульсно- кодовой модуляции (ДИКМ), относящийся к методам сжатия с потерями информации. Метод *RLE* эффективнее при сжатии изображений с малым числом градаций, метод ДИКМ обеспечивает сжатие полутоновых изображений  $\approx$  в 2 раза. В конце 60-х годов был предложен и разработан метод кодирования полутоновых изображений на основе ортогональных преобразований (метод сжатия с потерями), который обеспечивал коэффициенты сжатия до 5. В дальнейшем этот метод был использован в формате архивации файлов изображений *JPEG*. Далее, в 80 - е годы были предложены методы второго поколения, и среди них кодирование на основе вейвлетных преобразований, анизотропное нестационарное кодирование с предсказанием, кодирование на основе выращивания областей, кодирование на основе разложения по направлениям. Особенностью этих методов является более глубокое использование свойств ЗС в целях устраниния из изображений психофизической избыточности. Эти методы при кодировании цветных изображений позволяли получать коэффициенты сжатия более 20. На границе 80-х и 90-х годов был разработан фрактальный метод кодирования, обеспечивающий коэффициенты сжатия от 50 до 2000 раз, в зависимости от типа кодируемого изображения. В области кодирования видео

последовательностей повышение эффективности кодирования связано с использованием не только внутrikадровой избыточности, но и межкадровой избыточности вследствие сильных корреляционных зависимостей между изображениями в смежных кадрах. По оценкам физиологов количество информации, воспринимаемое ЗС человека, не превышает 70 двоичных единиц в секунду. Это означает, что в течение часа зритель воспримет  $70 \times 60 \times 60 / 8 = 31,5$  КБайт. 1 час видеопоследовательности цветного изображения при схеме  $YCrCb$  4:2:0, формате кадра 720x576, при частоте 25 кадров/с составит:  $((720 \times 576 \times 2) \times 25) \times 60 \times 60 = 74\,649\,600\,000$  байт. Таким образом, только малая часть информации будет воспринята зрителем. Правда, каждый зритель воспримет какую-то свою малую часть информации, не совпадающую с той, которую запомнит другой зритель. Но, между тем, понятно, что этот поток информации является избыточным для ее получателя.

## 8.2 Кодирование длин серий

Метод кодирования длин серий *RLE* [52], или групповое кодирование, метод сжатия без потерь информации, является самым простым, понятным и быстрым методом. Он широко применяется при записи графических изображений в файлы [53]. Примерами могут служить кадры мультиплекса, выполненного в технике, подобной мультфильмам Диснея, графические изображения (чертежи, плакаты, и др.) и другие, содержащие большие области постоянной яркости или цвета. Серии повторяющихся значений отсчетов кодируются двумя байтами: длиной серии (числом повторяемых отсчетов) и значением яркости отсчета, за счет чего достигается сжатие данных. Первый байт называется счетчиком серии. Второй байт называется значением серии. Эта пара байтов формирует *RLE*- пакет. При изменении серии и в случае, когда размер серии превышает диапазон счетчика, формируется новый пакет. Например, пусть задана последовательность значений: {0, 0, 100, 100, 100, 200, 200, 200, 200, 200, 170, 170}, соответствующая ей кодовая последовательность 4 пакетов имеет вид: 2,0;3,100;6,200;2,170. 13 байтов последовательности мы заменили 8 байтами потока.

Есть несколько вариантов группового кодирования. Обычно сжатие выполняется вдоль строк изображения, при этом они представляют собой одномерный поток, а не двумерную таблицу данных. При этом растровое изображение кодируется слева направо и сверху вниз, начиная с левого верхнего угла. Альтернативные схемы кодирования *RLE* разрешают кодировать по столбцам, фрагментами 4x4 элемента и в зигзагообразном порядке (см. *JPEG*). Определим коэффициент сжатия. При формировании *RLE* по байтовой схеме, при которой пакет занимает 2 байта, при размере полутонового изображения равном  $M$  строк по  $N$  столбцов, при восьмиразрядном квантовании коэффициент сжатия равен:

$$k_{\text{сж}} = NM / (2N_{\text{гр}}),$$

где  $N_{\text{гр}}$  - общее число серий в изображении, размер которых меньше максимального значения счетчика серии.

Таким образом, если в изображении все соседние отсчеты отличаются друг от друга, то есть счетчики всех серий равны 1, то код окажется в два раза длиннее, чем объем памяти, занимаемый самим изображением. Но по мере уменьшения вероятности появления новых значений отсчетов, коэффициент сжатия увеличивается.

Этот метод не эффективен при кодировании полутоновых изображений, в которых практически отсутствуют повторяющиеся последовательности уровней яркости. Одна из схем кодирования *RLE* представлена на рисунке 8.3. В этой схеме для кодирования серии длиной в 1 отсчет используется пакет только в том случае, когда значение отсчета превышает максимальное значение счетчика группы, в противном случае в поток выводится только 1 байт, содержащий значение отсчета. В данном примере выбран размер серии, равный 127, максимальное значение счетчика задано равным  $128_{10}=80_{16}$ . Старший бит счетчика равен 1, он является флагом пакета. Остальные 7 разрядов позволяют записать длину серии (размер серии ограничен 127 отсчетами). Если серия одинаковых отсчетов превышает этот размер, то формируется новый пакет. Кроме того, на этой схеме кодирование выполняется по строкам. При этом серия завершается последним отсчетом текущей строки. Алгоритм, представленный на рисунке 8.3 выполняется для каждой строки изображения.

Схема алгоритма декодирования представлена на рисунке 8.4. Обнуляется счетчик строки  $i$ . Формируется маска *MaskValue*, позволяющая выделить счетчик серии. В нашем случае она равна \$7F. Считывается байт из потока. Если его значение не меньше *FlagRLE*, то этот байт содержит значение счетчика в пакете. Следовательно, производится выделение значения счетчика *RunCount*. Считывается следующий байт из потока, второй байт пакета *Curr*. В буфер строки выводится *RunCount* раз значение *Curr*. Если же значение байта *Curr* меньше *FlagRLE*, следовательно, это не пакет, а одиночный отсчет. Значение *RunCount* устанавливается в 1. В буфер строки выводится одно значение *Curr*. Счетчик длины строки  $i$  увеличивается на *RunCount*. Эта последовательность операций выполняется, пока значение счетчика остается меньшим длины строки изображения. Для каждой следующей строки снова выполняется алгоритм, представленный на рисунке 8.4.

Недостатком *RLE* кодирования является низкая помехоустойчивость метода. Изменение яркости вследствие помехи приводит к изменению яркости всей последовательности (штрихи вдоль строк) или к “раздергиванию” строк в случае изменения длины серии вследствие помехи. Достоинством метода является простота его реализации. Этот метод используется в графических форматах *BMP*, *PCX*, *TIFF* и *TARGA*, как дополнительный в формате *JPEG*.

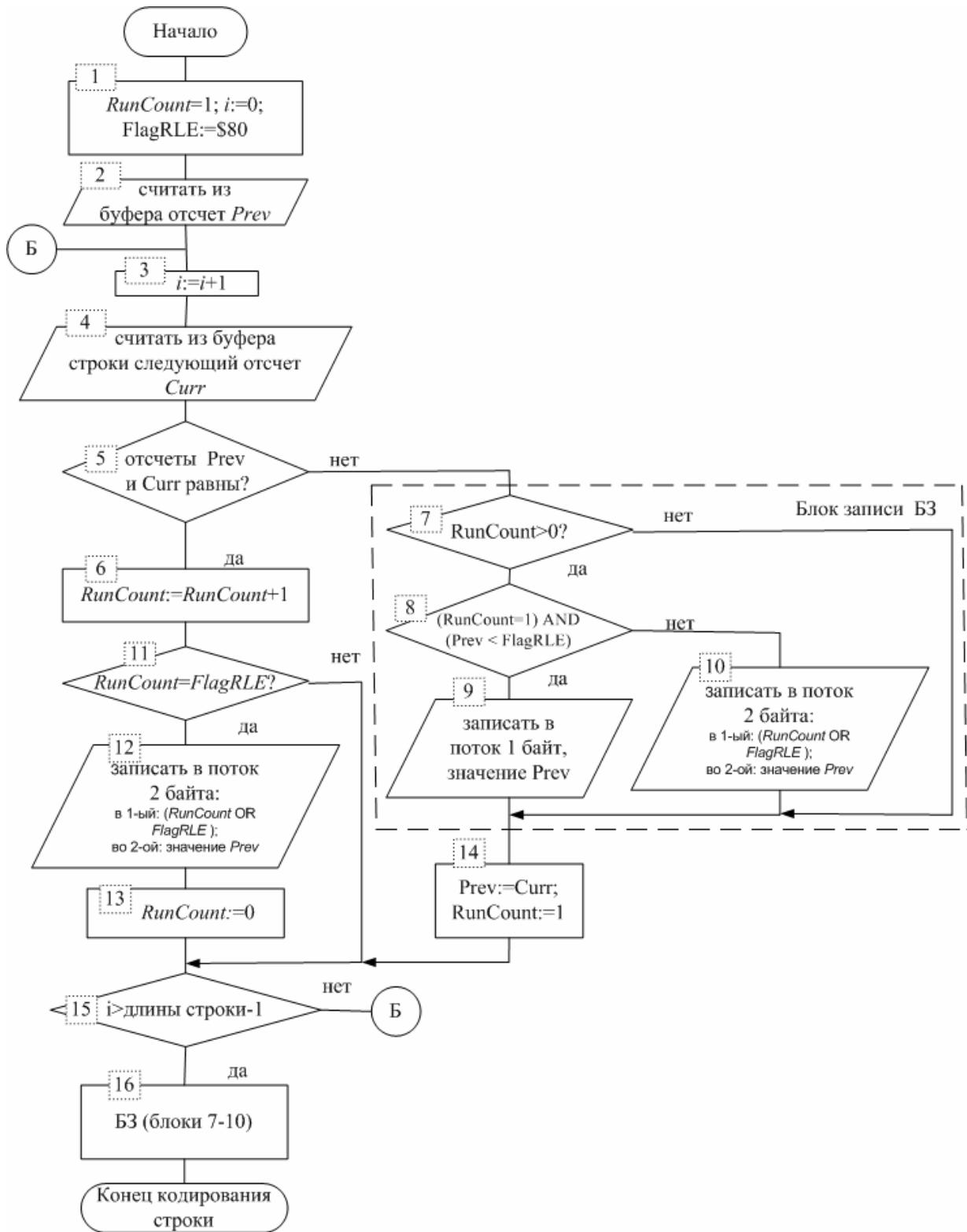


Рисунок 8.3 Блок-схема алгоритма кодирования длин серий строки изображения.

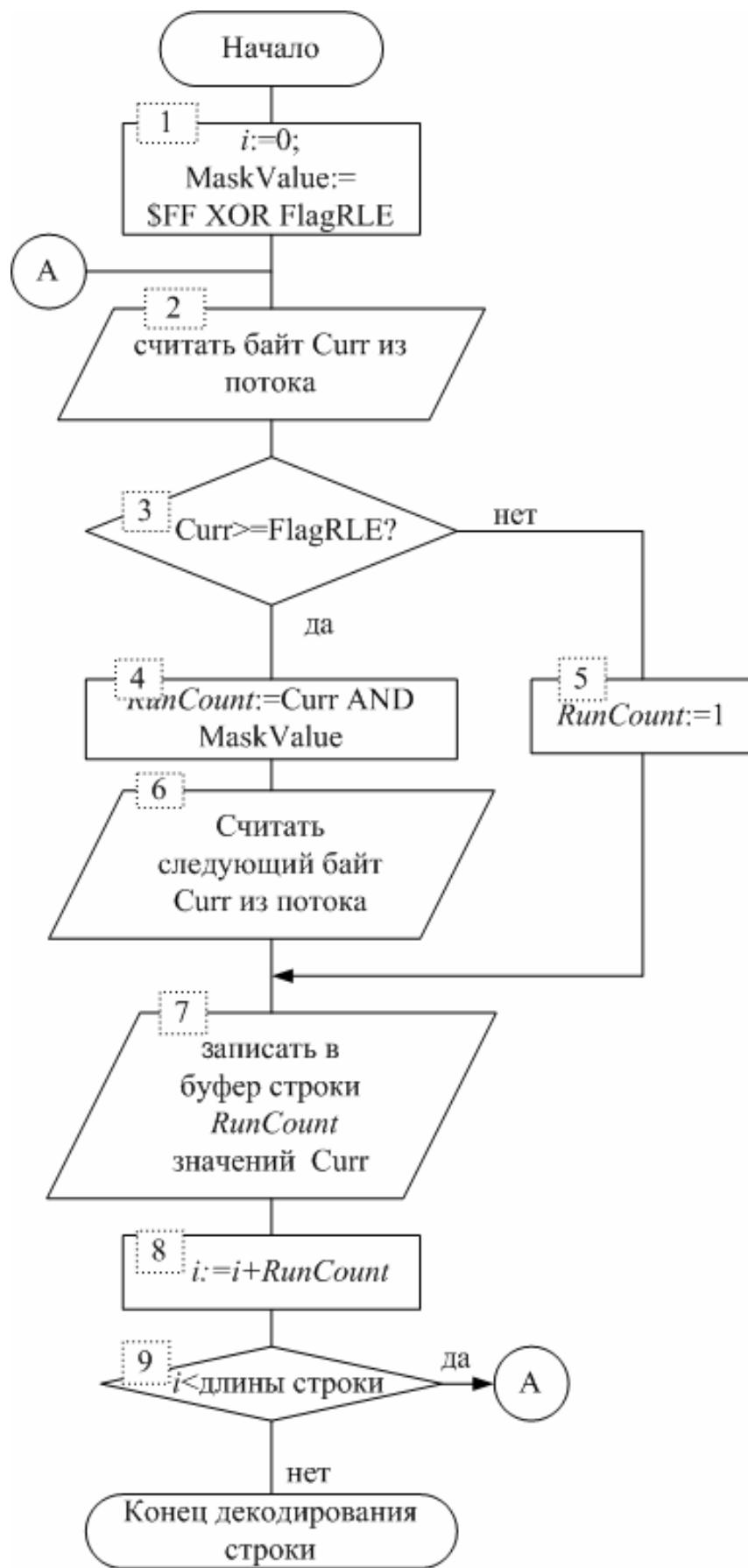


Рисунок 8.4 Блок-схема алгоритма декодирования строки изображения по методу RLE.

### 8.3 Кодирование по методу *LZW*

Кодирование по методу *LZW*- метод сжатия без потерь информации, названный по первым буквам фамилий разработчиков Абрахама Лемпела, Джекоба Зива и Терри Велча (*Lempel, Ziv и Welch*) [54]. Впервые алгоритм был опубликован в 1984 г. Он применяется для сжатия данных при записи на жесткий диск компьютера в формате *ARJ*, *ZIP*, *GZ*, применяется при записи файлов в форматах *TIFF* и *GIF*.

Особенностью метода является адаптивность и использование кодов переменной длины, с некоторым предопределенным числом разрядов.

#### 8.3.1 Кодирование

Начальная кодовая таблица инициализируется так, что каждому отсчету изображения из диапазона [0,255] ставятся в соответствие код, которым этот отсчет будет представлен в потоке, и два специальных кода: код очистки (256) и код конца потока (257), как показано в таблице 8.1.

Таблица 8.1 Таблица инициализации кодера/декодера *LZW*

№ по порядку	Десятичное значение строки байтов изображения или имя кода	Десятичное значение кода в потоке
0	0	0
...	...	...
254	254	254
255	255	255
256	Код очистки	256
257	Код конца записи	257

Пусть длина кода ограничена 12 разрядами. Тогда, если номера кодов превышают 12-разрядное значение, равное 4095, используют код очистки для новой инициализации таблицы. Код конца потока является признаком конца кодовой последовательности.

Алгоритм *LZW* можно записать формально:

- Инициализировать таблицу кодов (таблица 8.1).
- Записать в поток код очистки. Текущий буфер *CurBuf* пуст.
- Далее в цикле до конца файла изображения:
  - a) Прочитать очередной байт изображения в буфер (*Byte*).
  - b) Если в таблице кодов имеется код, соответствующий комбинации *CurBuf+ Byte*, то увеличить буфер на один байт и поместить в него *Byte*: *CurBuf := CurBuf + Byte*. Перейти в начало цикла а).
  - c) Иначе, (если код, соответствующий комбинации *CurBuf + Byte*, отсутствует) то:
    - получить из таблицы код *Code*, соответствующий содержимому буфера *CurBuf*,
    - вывести в выходной поток код *Code*,

- добавить в таблицу код, соответствующий  $CurBuf + Byte$ , присвоив ему значение, совпадающее со следующим порядковым номером,
  - переписать содержимое буфера  $Byte$  в  $CurBuf$ :  $CurBuf=Byte$ ,
  - перейти в начало цикла а).
- Поскольку по окончании файла изображения  $CurBuf$  не пуст, то необходимо:
    - получить из таблицы код  $Code$ , соответствующий содержимому буфера  $CurBuf$ ,
    - вывести в выходной поток код  $Code$ ,
    - вывести в выходной поток код конца записи.

Описанный алгоритм поясним на примере кодирования последовательности отсчетов: {7, 7, 7, 10, 10, 7, 7, 5, 5}. Кодирование начинается с инициализации таблицы кодов в соответствии с таблицей 8.1. В выходной поток записывается код очистки.  $CurBuf$  – пустой. Далее в цикле выполняются следующие действия:

Считываем нулевой байт изображения (7):  $Byte=7$ .  $CurBuf+Byte$  равен  $Byte$  и присутствует в таблице, поэтому  $CurBuf=7$ , переход к а).

Считываем первый байт изображения (7) в буфер  $Byte=7$ . Комбинация  $CurBuf+Byte$ , представляющая последовательность двух байтов (7,7), отсутствует в таблице. Поэтому получаем из таблицы код, соответствующий  $CurBuf$ ,  $Code = 7$  и записываем его в выходной поток. Добавляем в таблицу код, соответствующий ( $CurBuf+Byte$ ), присвоив ему значение, совпадающее со следующим порядковым номером  $Code = 258 \Leftrightarrow (7,7)$ . Переписываем содержимое буфера  $Byte$  в  $CurBuf$ :  $CurBuf = 7$  и переходим в начало цикла а).

Можно далее коротко записать процесс кодирования следующим образом:

Считываем (7). (7,7) есть в таблице.  $CurBuf = (7,7)$ .

Считываем (10). (7,7,10) нет в таблице. В поток - 258. В таблицу - 259  $\Leftrightarrow (7,7,10)$ .  $CurBuf = 10$ .

Считываем (10). (10,10) нет в таблице. В поток - 10. В таблицу - 260  $\Leftrightarrow (10,10)$ .  $CurBuf = 10$ .

Считываем (7). (10,7) нет в таблице. В поток - 10. В таблицу - 261  $\Leftrightarrow (10,7)$ .  $CurBuf = 7$ .

Считываем (7). (7,7) есть в таблице.  $CurBuf = (7,7)$ .

Считываем (5). (7,7,5) нет в таблице. В поток - 258. В таблицу - 262  $\Leftrightarrow (7,7,5)$ .  $CurBuf = 5$ .

Считываем (5). (5,5) нет в таблице. В поток - 5. В таблицу - 263  $\Leftrightarrow (5,5)$ .  $CurBuf = 5$ .

В поток-5. В поток-257.

Таким образом, выполняя перечисленные действия, мы получили результаты, приведенные в таблице 8.2.

Выходной поток представляет следующую последовательность кодов: {256, 7, 258, 10, 10, 258, 5, 5, 257}.

Таблица 8.2. Пример кодирования по алгоритму LZW

Чтение байта №	<i>Byte</i>	Код, выводимый в поток	Содержимое таблицы кодов	<i>CurBuf</i>
		256		пустой
0	7			7
1	7	7	258 ⇔ 7,7	7
2	7			7,7
3	10	258	259 ⇔ 7,7,10	10
4	10	10	260 ⇔ 10,10	10
5	7	10	261 ⇔ 10,7	7
6	7			7,7
7	5	258	262 ⇔ 7,7,5	5
8	5	5	263 ⇔ 5,5	5
		5		
		257		

### 8.3.2 Декодирование

При декодировании кодовая таблица не нужна, поскольку первоначально инициируется словарь в соответствии с таблицей 8.1. По мере поступления кодов комбинаций исходных отсчетов, в процессе декодирования составляется кодовая таблица, идентичная той, что была составлена кодером.

#### Алгоритм

- Прочитать код сжатых данных в буфер *Code*.

Далее в цикле до конца потока выполняются следующие операции:

- Если *Code* равен коду очистки, то
  - a) Инициализировать таблицу кодов.
  - b) Прочитать следующий код сжатых данных.
  - c) Найти последовательность байтов в кодовой таблице, соответствующую коду *Code* и записать ее в файл изображения.
  - d) Скопировать *Code* в буфер *OldCode*.
- Если *Code* равен коду конца записи, то завершить работу.
- Иначе:
  - Если *Code* находится в таблице, то
    - a) вывести соответствующую ему декодированную последовательность отсчетов в файл;
    - b) последовательность отсчетов, соответствующую *OldCode* + первый байт последовательности, соответствующей *Code*, добавить в кодовую таблицу;
    - c) скопировать *Code* в буфер, где хранится *OldCode*.
  - Иначе, если *Code* не находится в таблице, необходимо:

- a) сформировать последовательность, соответствующую  $OldCode +$  первый байт последовательности  $OldCode$  и вывести значения декодированного кода в файл изображения;
  - b) добавить в таблицу полученную последовательность байтов;
  - c) скопировать  $Code$  в буфер, где хранится  $OldCode$ .
- Прочитать код  $Code$  из входного потока. Перейти на начало цикла.

Для приведенного примера процесс декодирования выполняется в соответствии с таблицей 8.3.

Таблица 8.3. Пример декодирования по алгоритму  $LZW$

Чтение кода №	Код, вводимый из потока	Последовательность байтов, выводимых в файл изображения	Содержимое таблицы кодов	$OldCode$
0	256			
1	7	7	Инициализировать таблицу	7
2	258	7,7	$258 \Leftrightarrow 7,7$	258
3	10	10	$259 \Leftrightarrow 7,7,10$	10
4	10	10	$260 \Leftrightarrow 10,10$	10
5	258	7,7	$261 \Leftrightarrow 10,7$	258
6	5	5	$262 \Leftrightarrow 7,7,5$	5
7	5	5	$263 \Leftrightarrow 5,5$	5
8	257			5

В результате декодирования исходная последовательность байтов восстановлена без потери информации.

Метод сжатия  $LZW$  может быть применен не только для кодирования восьмиразрядных данных, но и для кодирования данных произвольной разрядности. В этом случае кодовые размерности объединяются в группы из восьми разрядов. Например, в четырехразрядных данных два отсчета объединяются в один, а при 16-разрядном представлении один отсчет разбивается на 2 группы. Коэффициент сжатия при использовании этого метода равен обычно 2-3.

#### 8.4 Метод кодирования Хаффмана

Метод кодирования Хаффмана [55] относится к группе методов сжатия данных без потерь информации. Этот метод используется для поддержки факсимальной связи и представления документов. Применяется также при записи графических изображений в файлы и является компонентом алгоритмов сжатия данных *JPEG* и *MPEG - 2*. Метод предложен Дэвидом Хаффманом в 1952 г. Особенностью метода является использование кодов переменной длины, при этом наиболее

вероятным символам присваиваются наиболее короткие кодовые слова, а менее вероятным – длинные. Благодаря такой стратегии, код Хаффмана дает минимальную среднюю длину кодовой последовательности, приближающуюся к энтропии источника сообщения. Рассмотрим построение кодовой таблицы на примере. Пусть кодируется 13 символов упорядоченных по невозрастанию вероятности их появления:  $\{A_1, A_2, \dots, A_{13}\}$ . Пусть вероятности появления  $A_i$  обозначаются  $p_i$  и составляют  $\{0,2; 0,18; 0,1; 0,1; 0,1; 0,06; 0,06; 0,04; 0,04; 0,04; 0,04; 0,03; 0,01\}$  соответственно. Алгоритм можно описать следующим образом.

Объединяются два символа с наименьшими вероятностями в узел кодового дерева, этому узлу приписывается вероятность, равная сумме вероятностей появления символов, составляющих узел. В примере минимальную вероятность имеют символы  $A_{12}$  и  $A_{13}$ , суммарная вероятность которых равна 0,04. Далее объединяются символы или узлы с минимальными вероятностями, образуя следующий узел, которому присваивается вероятность, равная сумме вероятностей составляющих.

Процесс повторяется до тех пор, пока не сойдется в один узел, расположенный в вершине. После этого левые и правые ветви дерева обозначаются «0» и «1» соответственно (или наоборот, «1» и «0»). Значение кодового слова формируется в виде последовательности кодов ветвей по пути от вершины кодового дерева к узлу. Кодовое дерево для приведенного примера представлено на рисунке 8.5. В результате кодирования по методу Хаффмана получены коды, приведенные в таблице 8.4. В таблице 8.4  $L_i$  - длина кодового слова  $i$ -го символа (в количестве разрядов). Код Хаффмана является префиксным кодом, то есть таким, что ни одно кодовое слово не является префиксом (началом) другого кодового слова. Префиксные коды декодируются однозначно, но обратное утверждение не выполняется. Например, код  $\{0, 01, 010, 0,0101, \dots\}$  не является префиксным, хотя декодируется однозначно.

Таблица 8.4 Таблица кодов для последовательности  $\{A_1, \dots, A_{13}\}$

Имя	Код	$L_i$	$p_i$	Имя	Код	$L_i$	$p_i$
A 1	00	2	0,2	A8	10101	5	0,04
A2	100	3	0,18	A9	10110	5	0,04
A3	110	3	0,1	A10	10111	5	0,04
A4	010	3	0,1	A11	11110	5	0,04
A5	011	3	0,1	A12	111110	6	0,03
A6	1110	4	0,06	A13	111111	6	0,01
A7	10100	5	0,06				

Средняя длина кода может быть получена в соответствии с формулой:

$$L_{cp} = \sum_{i=1}^N p_i L_i . \quad (8.2)$$

Для рассмотренной последовательности средняя длина кода составляет 3,42 разряда.

В приведенном примере таблица кодов построена в соответствии с частотой появления символов во входной последовательности.

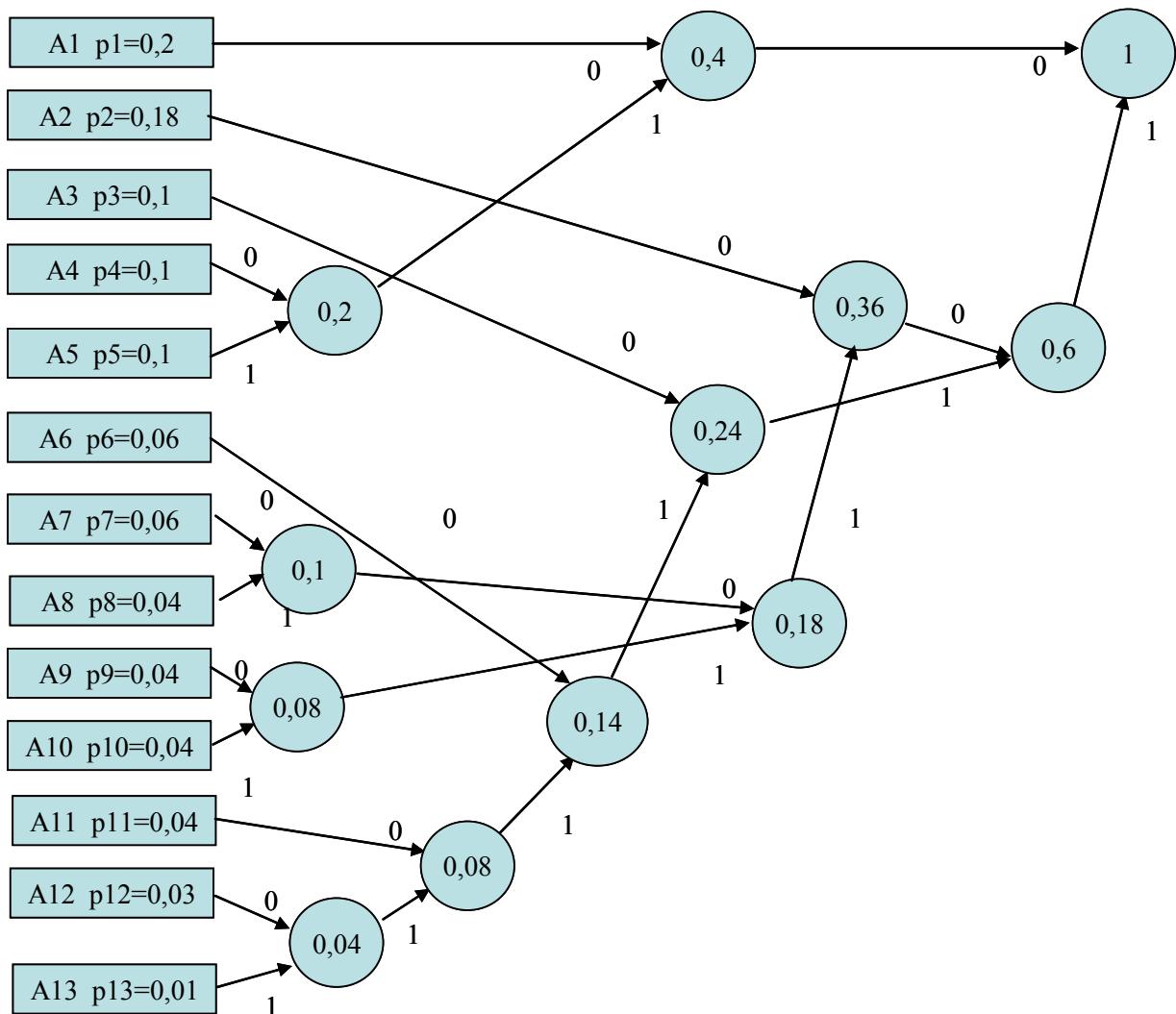


Рисунок 8.5 Формирование кодового дерева по методу Хаффмана.

Процедура кодирования в этом случае выполняется в два этапа. Вначале оцениваются частоты появления символов в последовательности, строится кодовая таблица, а затем, на втором этапе, производится собственно кодирование.

Для декодирования необходимо передать с потоком кодовую таблицу, содержащую алфавит (все символы от A1 до A13) и соответствующие им коды. На практике часто используют модифицированный алгоритм Хаффмана, при котором применяется заранее составленная на основании статистических данных кодовая таблица. Именно так выполняется кодирование в формате *JPEG*.

## 8.5 Принцип дифференциальной импульсно - кодовой модуляции

ДИКМ была предложена Катлером в 1952 г. Этот метод относится к методам кодирования с предсказанием [56]. Рассмотрим принцип действия на примере полутонового изображения. Значение каждого элемента

изображения предсказывается на основе значений предшествующих элементов. Оценка предсказания отсчета  $g(t_n)$  в момент  $t_n$  вычитается из фактического значения элемента  $f(t_n)$ . Полученное значение разности, является значением ошибки предсказания:

$$\varepsilon(t_n) = f(t_n) - g(t_n). \quad (8.3)$$

Оно квантуется и кодируется.

Восстановление сигнала осуществляется путем его предсказания в соответствии с процессом, выполненным при кодировании. Затем, к предсказанному значению прибавляется декодированное значение ошибки предсказания, то есть  $f(t_n) = g(t_n) + \varepsilon(t_n)$ .

Таким образом, при ДИКМ в поток записываются один опорный отсчет на некоторое количество отсчетов изображения, например, на строку, и значения ошибок предсказания для последующих отсчетов. Сокращение потока достигается квантованием и кодированием ошибки предсказания. Восстановленное значение сигнала отличается от исходного на величину шума квантования ошибки предсказания.

Разностный сигнал имеет больший динамический диапазон, чем исходный. Так, при 8-ми разрядном квантовании сигнала диапазон разностей составляет от -255 до 255. Однако плотность распределения вероятностей разностного сигнала имеет резкий пик слева и справа от нуля. На рисунке 8.6 представлены изображение, полученное дифференцированием по строке изображения «Лена» а), и гистограмма распределения яркостей этого изображения б).

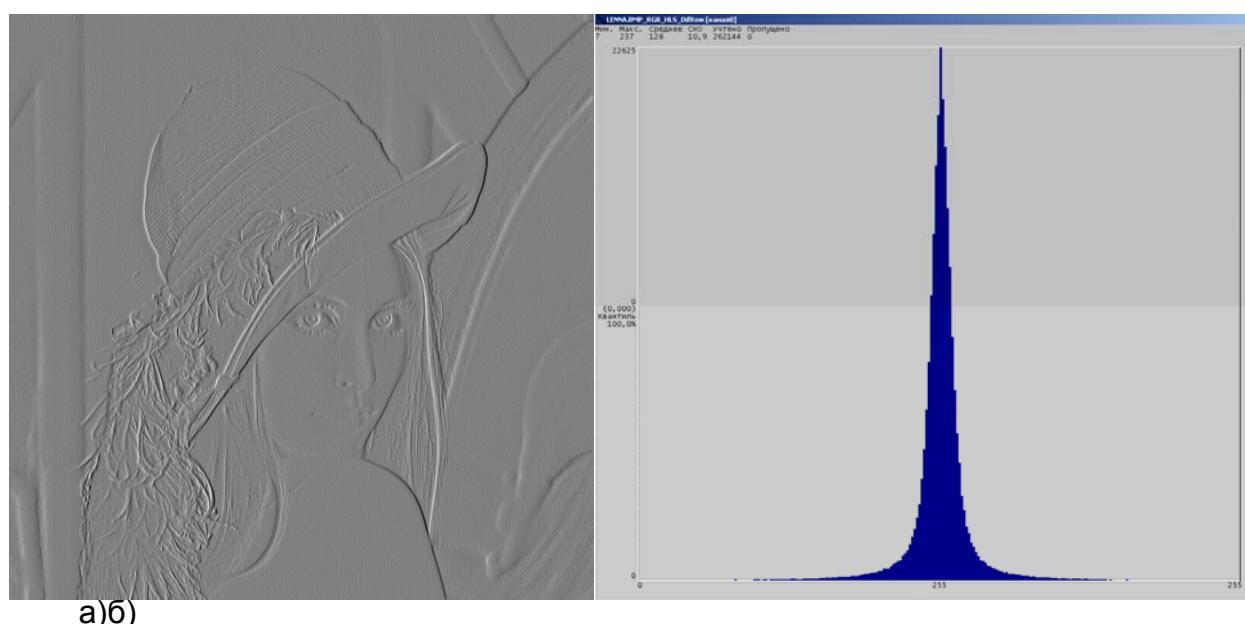


Рисунок 8.6 а) Изображение, полученное дифференцированием по строке изображения «Лена»; б) гистограмма распределения яркостей этого изображения.

Из рисунка 8.6 видно, что при формировании предсказания отсчета по предыдущему отсчету, происходит декорреляция изображения. Корреляционные связи между отсчетами остаются и в разностном сигнале,

о них свидетельствуют контуры на изображении (а). В 4.2.1 мы рассматривали решение задачи минимизации ошибки квантования. Было показано, что при неравномерной плотности распределения сигнала уровни квантования задаются неравномерно. Малые значения разностного сигнала, имеющие более высокую вероятность, квантуются точнее, чем менее вероятные большие значения, которые подвергают грубому квантованию.

Чаще всего при ДИКМ используется линейное предсказание. При этом предсказание значения отсчета вычисляется как линейная комбинация предыдущих отсчетов:

$$g(t_n) = \sum_{k=1}^N c_k f(t_k), \quad (8.4)$$

где  $k$  - номер отсчета, используемого при предсказании,  $N$  - число используемых для предсказания отсчетов;  $c_k$  - весовой коэффициент.

Весовые коэффициенты вычисляются по критерию минимума среднего квадрата ошибки предсказания. Ошибка предсказания в этом случае вычисляется в соответствии с (8.3):

$$\varepsilon(t_n) = f(t_n) - \sum_{k=1}^N c_k f(t_k). \quad (8.5)$$

Среднее значение квадрата ошибки предсказания вычисляется:

$$S = E(\varepsilon(t_n)^2) = E\left(\left(f(t_n) - \sum_{k=1}^N c_k f(t_k)\right)^2\right). \quad (8.6)$$

Минимум среднего значения квадрата ошибки достигается при тех значениях коэффициентов  $c_k$ , при которых его производная равна нулю:

$$\frac{\partial S}{\partial c_k} = 0. \quad (8.7)$$

Для определения коэффициентов  $c_k$  решается система уравнений, полученная дифференцированием  $S$  по  $c_k$  и приравниванием нулю этих частных производных. Например, при предсказании по одному предыдущему отсчету, уравнение (8.6) запишется в виде:

$$S = E((f(t_n) - c_1 f(t_{n-1}))^2) = E(f(t_n)^2 - 2c_1 f(t_n)f(t_{n-1}) + c_1^2 f^2(t_{n-1})).$$

Отсюда

$$\frac{\partial S}{\partial c_1} = -2E(f(t_n)f(t_{n-1})) + 2c_1 E(f^2(t_{n-1})). \quad (8.8)$$

Приравнивая нулю производную и решая относительно весового коэффициента, получаем:

$$c_1 = \frac{E(f(t_n)f(t_{n-1}))}{E(f^2(t_{n-1}))} = \rho(\tau), \quad (8.9)$$

где  $\rho(\tau)$  - коэффициент автокорреляции сигнала при смещении его реализации на интервал времени, равный интервалу между двумя соседними отсчетами  $\tau$ .

С учетом этого среднее значение квадрата ошибки предсказания можно записать в виде:

$$E(\varepsilon^2) = E(f^2)(1 - \rho(\tau))^2. \quad (8.10)$$

Из формулы (8.9) следует, что чем ближе значение коэффициента автокорреляции к единице, тем ближе к единице значение весового коэффициента  $c_1$ . В случае портретных изображений  $\rho(\tau)$  мало отличается от 1. Это позволяет в качестве предсказываемого значения сигнала  $g(t_n)$  использовать значение на предыдущем отсчете  $f(t_{n-1})$ . Такая замена, упрощая алгоритм, не приводит к заметному ухудшению качества сжатого изображения.

Из формулы (8.10) следует, что средний квадрат сигнала ошибки предсказания  $E(\varepsilon^2)$  много меньше среднего квадрата исходного сигнала при значениях коэффициента автокорреляции  $\rho(\tau)$  исходного сигнала, близких к 1. Следовательно, при квантовании сигнала ошибки предсказания, средний квадрат шума квантования будет гораздо меньше, чем значение, которое мы имели бы при квантовании самого сигнала.

При предсказании сигнала изображения по двум отсчетам осуществляется двумерное предсказание. В этом случае для предсказания сигнала  $f(x,y)$  используются значения сигналов  $f(x-1,y)$  и  $f(x,y-1)$ . Переход от одномерного предсказания к двумерному приводит к уменьшению  $E(\varepsilon^2)$ .

Структурная схема кодера ДИКМ представлена на рисунке 8.7. Схема кодера включена так, что ошибки квантования учитываются в предсказании. Это позволяет обеспечить устойчивое кодирование. В противном случае, если положение ключа будет отличаться от приведенного на рисунке, ошибки квантования  $n$ -го отсчета не будут учтены в предсказании ( $n+1$ ) отсчета.

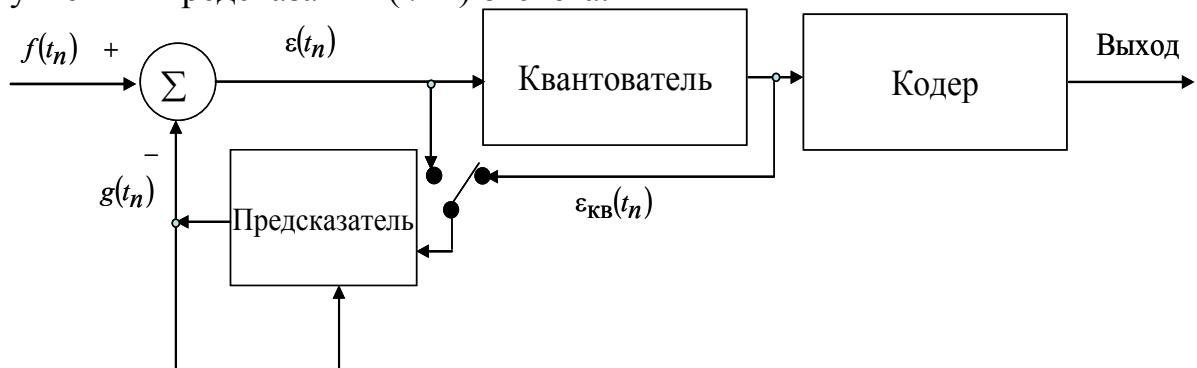


Рисунок 8.7 Структурная схема кодера ДИКМ.

В декодере эта ошибка квантования присутствует и приводит к ее распространению на последующие декодируемые отсчеты. Вследствие чего происходит накопление ошибок и увеличение их заметности на декодированных изображениях.

ДИКМ является методом сжатия данных с потерей информации [32], поскольку декодированное изображение отличается от исходного наличием шума квантования. В настоящее время ДИКМ используется как компонент сложных алгоритмов сжатия.

## 8.6 Кодирование с использованием ортогональных преобразований

Метод ДИКМ основан на декорреляции яркостей элементов изображения, квантовании и кодировании полученных ошибок квантования. Эти же основные операции применяются и при трансформационном кодировании. Однако, при ДИКМ декорреляция выполнялась в пространственной области, а в случае трансформационного кодирования декорреляция производится за счет преобразования изображения в матрицу коэффициентов преобразования (разложения). Размер матрицы коэффициентов преобразования равен размеру исходного изображения. Этот метод, как и ДИКМ, относится к методам кодирования с потерей информации. Квантованию и кодированию в этом случае подвергаются значения коэффициентов преобразования. Важным является то обстоятельство, что энергия коэффициентов преобразования распределена неравномерно между спектральными коэффициентами. При квантовании более информативные коэффициенты (НЧ) квантуются точнее, а менее информативные (ВЧ) квантуются грубее или отбрасываются. На этом этапе и осуществляется сжатие изображения. Дополнительного сжатия достигают применением кодирования. Обычно применяется энтропийное кодирование, например, рассмотренное кодирование по методу Хаффмана.

При декодировании изображения последовательно выполняются все операции кодирования в обратном порядке: декодируется поток, все коэффициенты преобразования домножаются на соответствующие весовые множители, получаемые как обратные коэффициентам квантования величины. А затем по восстановленным коэффициентам разложения путем обратного ортогонального преобразования восстанавливается само изображение.

Метод преобразования непрерывного сигнала в множество некоррелированных коэффициентов разработан Каруненом (*H. Karhunen*) и Лоэвом (*M. Loeve*). Хотеллинг (*H. Hotelling*) первым предложил метод преобразования дискретных сигналов в набор некоррелированных коэффициентов. В процессе преобразований изображения  $f(x,y)$ , имеющего сильные корреляционные связи между соседними отсчетами, происходит процесс декорреляции. Значения коэффициентов преобразования  $F(u,v)$  оказываются некоррелированными. Именно при преобразовании

Карунена-Лоэва достигается максимальная концентрация энергии. Однако применение получили те преобразования, вычислительная сложность которых меньше, хотя они и не позволяют полностью декоррелировать коэффициенты преобразования. К таким преобразованиям относится, например, дискретное преобразование Фурье.

Рассмотрим более подробно ортогональное преобразование изображения, представленного в виде массива (матрицы) чисел  $f(x,y)$ , размер которого  $N \times M$ , где  $N$  - число столбцов,  $M$  - число строк в изображении,  $x$  - номер столбца;  $y$  - номер строки. Спектральные коэффициенты  $F(u,v)$  вычисляются путем прямого ортогонального преобразования изображения следующим образом:

$$F(u,v) = \sum_{y=0}^{M-1} \sum_{x=0}^{N-1} f(x,y) a(x,y;u,v), \quad (8.11)$$

где  $a(x,y;u,v)$  - ядро прямого преобразования (базисные функции, по которым происходит разложение);  $u,v$  - индексы спектральных коэффициентов в матрице спектральных коэффициентов, соответствующие базисным функциям. Общее число базисных функций составляет  $NM$ , каждая определяется парой индексов  $(u,v)$ .

Как видно из (8.11), величина каждого из коэффициентов  $F(u,v)$  определяется всеми элементами изображения, поэтому не несет информации о локальных свойствах изображения в пространственной области.

Исходное изображение получается путем обратного ортогонального преобразования:

$$f(x,y) = \sum_{v=0}^{M-1} \sum_{u=0}^{N-1} F(u,v) b(x,y;u,v),$$

где  $b(x,y;u,v)$  - ядро обратного преобразования.

Если преобразование разделимо, то есть

$$\begin{aligned} a(x,y;u,v) &= a_x(x,u)a_y(y,v) \\ b(x,y;u,v) &= b_u(x,u)b_v(y,v), \end{aligned}$$

то оно может быть выполнено в два этапа, вначале по всем столбцам, а затем по всем строкам.

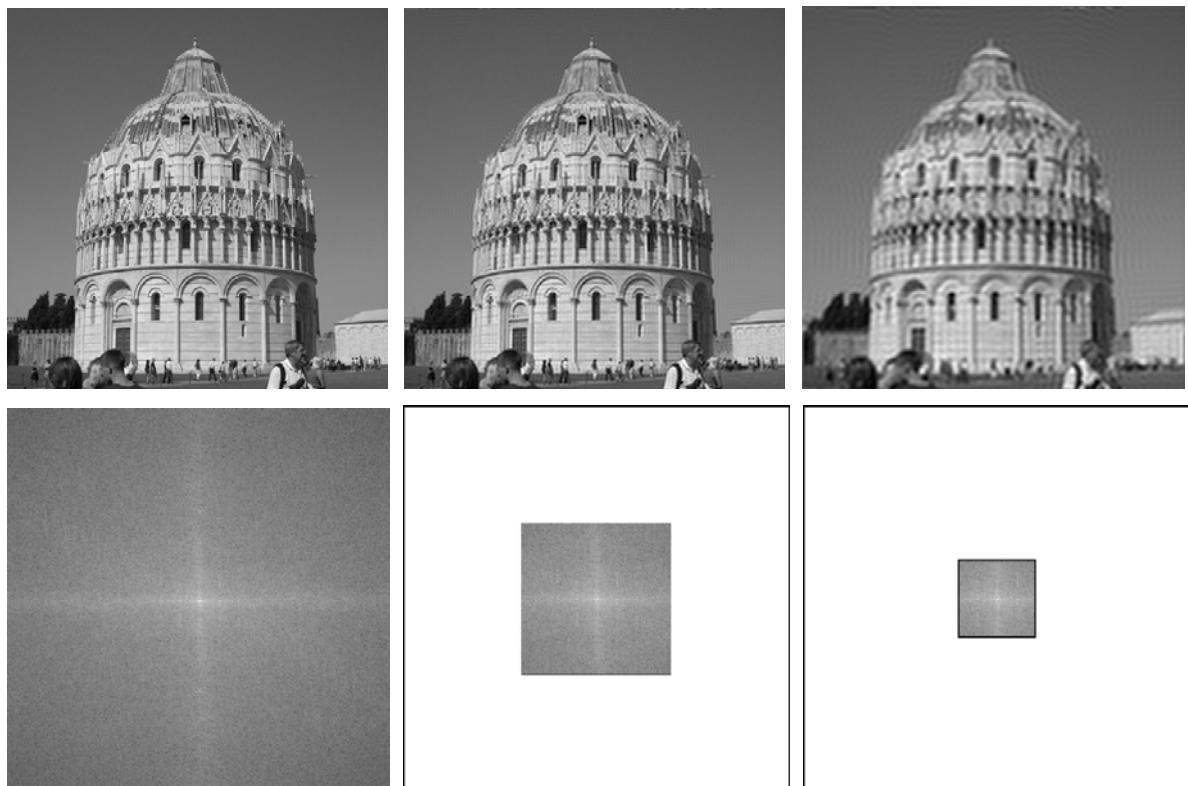
$$F(u,v) = \sum_{y=0}^{M-1} a_y(y,v) \sum_{x=0}^{N-1} f(x,y) a_x(x,u). \quad (8.12)$$

И соответственно

$$f(x,y) = \sum_{v=0}^{M-1} b_v(v,y) \sum_{u=0}^{N-1} F(u,v) b_u(x,u). \quad (8.13)$$

Ошибка преобразования определяется энергией отброшенных коэффициентов (суммой их квадратов). Чем больше коэффициентов мы отбрасываем, тем с большими ошибками будет восстановлено исходное изображение. Проиллюстрируем это на примере. В качестве исходного

используем изображение, представленное на рисунке 8.8 а), размер которого составляет 512x512 элементов. Выполним ортогональное преобразование Фурье. Спектр этого изображения представлен на рисунке 8.8 г). В области НЧ оставим по 200 коэффициентов в каждом направлении, по строкам и по столбцам, остальные коэффициенты обнулим, как показано на рисунке 8.8 д). Выполним обратное преобразование. Восстановленное изображение представлено на рисунке 8.8 б). А теперь оставим только по 100 коэффициентов в каждом направлении (в соответствии с рисунком 8.8 е). Восстановленное изображение представлено на рисунке 8.8 в). Из рисунка видно, что чем больше коэффициентов преобразования мы обнуляем, тем больше ошибки восстановления.



а) б) в)  
г) д) е)

Рисунок 8.8 Исходное изображение а) и два восстановленных изображения б) и в). Спектры соответствующих изображений г) - е).

При ортогональном преобразовании базисные функции ортогональны, то есть скалярные произведения всех пар базисных функций равны нулю. Как было рассмотрено в 6.2.1, в случае ДПФ базисные функции представляют собой комплексные экспоненты, а ортогональные преобразования имеют вид (6.17), (6.18). После выполнения прямого преобразования формируется матрица спектральных коэффициентов. Прямоугольная область размером  $NM$  имеет тот же размер, что и изображение. Основная доля энергии приходится на спектральные коэффициенты с малыми индексами ( $i, v$ ), то есть индексами низких пространственных частот. При сжатии спектральные коэффициенты,

соответствующие ВЧ, квантуются на малое число уровней, что позволяет для их представления использовать коды с малым числом двоичных единиц. Из (6.17) следует, что хотя значения отсчетов изображения являются действительными положительными числами, спектр Фурье - комплексный. Это одна из причин, по которой ДПФ не является лучшим преобразованием для сжатия изображений. Наиболее исследованным является широко применяемое в настоящее время дискретное косинусное преобразование (ДКП). Различаясь базисными функциями, эти преобразования различаются скоростью убывания спектральных коэффициентов с увеличением частот  $u, v$ .

Известны два метода отбора коэффициентов: зональный и пороговый. Первый заключается в том, что заранее, исходя из статистики изображений, в матрице спектральных коэффициентов выделяются зоны, и все спектральные коэффициенты, входящие в одну зону, квантуются на одно и то же число уровней, как показано в таблице 8.5.

Таблица 8.5 Распределение двоичных единиц кода между спектральными коэффициентами при зональном методе

	0	1	2	3	4	5	6	7
0	12	7	5	3	2	1	0	0
1	7	7	5	3	2	1	0	0
2	5	5	5	3	2	1	0	0
3	3	3	3	3	2	1	0	0
4	2	2	2	2	2	1	0	0
5	1	1	1	1	1	1	0	0
6	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	0

Второй заключается в том, что передаются спектральные коэффициенты, амплитуда которых превышает заранее установленный порог. Пороговый метод отбора требует кроме передачи значений спектральных коэффициентов записи адресов (индексов) этих коэффициентов.

### 8.6.1 Дискретное косинусное преобразование

Более эффективным, чем Фурье преобразование, является ДКП [12]. Преимущества ДКП связаны с тем, что в отличие от комплексного спектра на выходе ДПФ, при ДКП формируются вещественные спектральные коэффициенты преобразования. Кроме того, эти коэффициенты с ростом частоты убывают быстрее, чем коэффициенты ДПФ.

Рассмотрим двумерное ДКП изображения размером  $N$  столбцов и  $M$  строк. Двумерное ДКП вычисляется по формулам:

$$Z(u, v) = \frac{2}{\sqrt{NM}} \alpha(u)\alpha(v) \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f(n, m) \cos\left(\frac{\pi(2n+1)u}{2N}\right) \cos\left(\frac{\pi(2m+1)v}{2M}\right), \quad (8.14)$$

где  $u \in [0, N-1]$ ,  $v \in [0, M-1]$ , при  $k=(u,v)$   $\alpha(k) = \begin{cases} 1/\sqrt{2}, & k=0 \\ 1, & \text{иначе} \end{cases}$ .

Двумерное ДКП разделимо:

$$\begin{aligned} Z(u,v) &= \sqrt{\frac{2}{M}} \alpha(v) \left( \sum_{m=0}^{M-1} \left( \sqrt{\frac{2}{N}} \alpha(u) \sum_{n=0}^{N-1} f(n,m) \cos\left(\frac{\pi(2n+1)u}{2N}\right) \right) \cos\left(\frac{\pi(2m+1)v}{2M}\right) \right) \\ &= \sqrt{\frac{2}{M}} \alpha(v) \left( \sum_{m=0}^{M-1} Z(u,m) \cos\left(\frac{\pi(2m+1)v}{2M}\right) \right), \end{aligned} \quad (8.15)$$

где  $Z(u,m) = \sqrt{2/N} \alpha(u) \sum_{n=0}^{N-1} f(n,m) \cos\left[\frac{\pi(2n+1)u}{2N}\right]$  является одномерным

ДКП по строке  $m$ .

В соответствии с (8.15) двумерное преобразование ДКП можно выполнить как два последовательных одномерных преобразований ДКП. Сначала выполняется одномерное ДКП по строкам, а затем - по столбцам. В этом случае обратное ДКП (ОДКП) вычисляется в соответствии с уравнениями:

$$Z(u,m) = \sqrt{2/M} \sum_{v=0}^{M-1} \alpha(v) Z(u,v) \cos\left[\frac{\pi(2m+1)v}{2M}\right], \quad (8.16)$$

$$Z(n,m) = \sqrt{2/N} \sum_{u=0}^{N-1} \alpha(u) Z(u,m) \cos\left[\frac{\pi(2n+1)u}{2N}\right]. \quad (8.17)$$

С целью повышения быстродействия и сохранения локальных свойств изображения преобразование ДКП выполняется не над всем изображением, а разбивается на блоки, размер которых существенно меньше размера изображения.

При размере блока  $8 \times 8$  элементов, длина последовательности  $N=8$ . ДКП с точностью до множителей  $(\sqrt{2/N})\alpha(k)$  можно представить в матричном виде как произведение матрицы преобразования на вектор-столбец входных данных:

$$\begin{bmatrix} Z_0 \\ Z_1 \\ Z_2 \\ Z_3 \\ Z_4 \\ Z_5 \\ Z_6 \\ Z_7 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ \lambda & \gamma & \mu & v & -v & -\mu & -\gamma & -\lambda \\ \beta & \delta & -\delta & -\beta & -\beta & -\delta & \delta & \beta \\ \gamma & -v & -\lambda & -\mu & \mu & \lambda & v & -\gamma \\ \alpha & -\alpha & -\alpha & \alpha & \alpha & -\alpha & -\alpha & \alpha \\ \mu & -\lambda & v & \gamma & -\gamma & -v & \lambda & -\mu \\ \delta & -\beta & \beta & -\delta & -\delta & \beta & -\beta & \delta \\ v & -\mu & \gamma & -\lambda & \lambda & -\gamma & \mu & -v \end{bmatrix} \times \begin{bmatrix} f_0 \\ f_1 \\ f_2 \\ f_3 \\ f_4 \\ f_5 \\ f_6 \\ f_7 \end{bmatrix}, \quad (8.18)$$

где  $\alpha = \cos(\pi/4) \approx 0,707$ ;  $\beta = \cos(\pi/8) \approx 0,924$ ;  $\delta = \sin(\pi/8) \approx 0,383$ ;  
 $\lambda = \cos(\pi/16) \approx 0,981$ ;  $\gamma = \cos(3\pi/16) \approx 0,831$ ;  $\mu = \sin(3\pi/16) \approx 0,556$ ;  
 $\nu = \sin(\pi/16) \approx 0,195$ . На рисунке 8.9 приведены одномерные базисные функции ДКП для  $N=8$ .

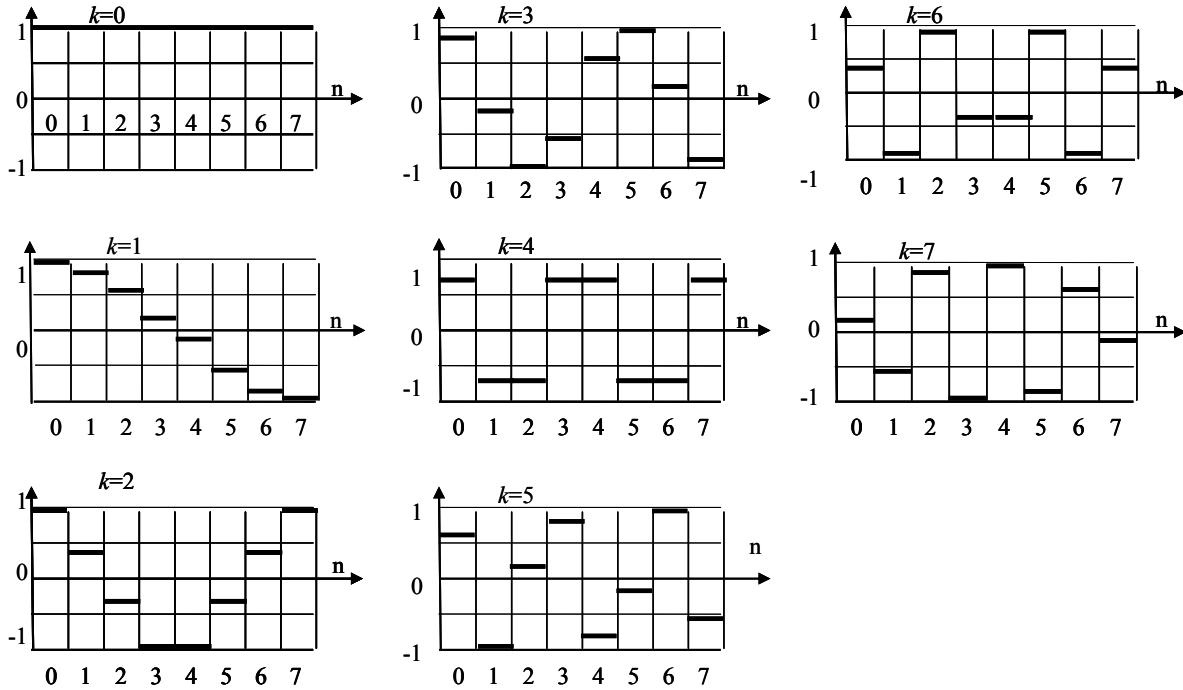


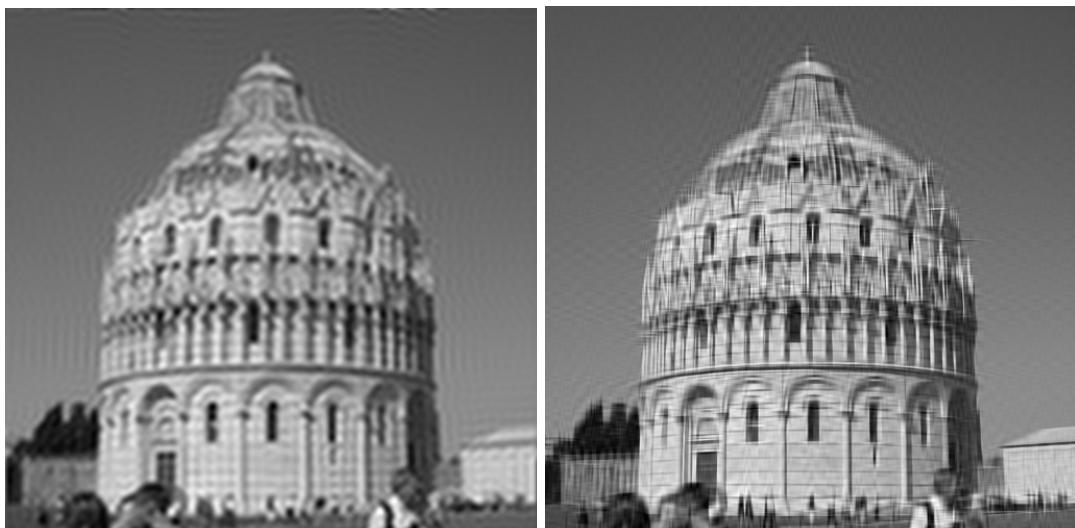
Рисунок 8.9 Базисные функции ДКП,  $n$ -порядковый номер отсчета входной последовательности,  $k$ -номер выходного значения коэффициента ДКП.

Значимые коэффициенты ДКП сосредоточены в области НЧ. Покажем на примере преимущество использования ДКП перед ДПФ при сжатии изображений. Возьмем изображение размером 512x512. Выполним преобразования. Оставим матрицу коэффициентов 100x100. Остальные коэффициенты обнулим. Выполним обратные преобразования.

Полученные восстановленные изображения приведены на рисунке 8.10. Видно, что ДКП позволяет получить более высокое качество восстановленного изображения, чем ДПФ.

## 8.7 Сжатие данных по методу JPEG

В 1992 ISO (*International Organization for Standardization* - Международная организация по стандартизации), ITU и международной электротехнической комиссией IEC (*International Electrotechnical Commission*- Международный электротехнический комитет) был установлен стандарт JPEG для сжатия неподвижных изображений [56]. В формате записи изображений JPEG применен метод сжатия данных с использованием ДКП, т.е. метод сжатия данных с потерями информации. Аббревиатура JPEG означает название организации, разработавшей этот стандарт, - *Joint Photographic Experts Group* (Объединенная группа экспертов по фотографии).



а) б)

Рисунок 8.10 Изображение, восстановленное по 3,8% коэффициентов: а) по алгоритму ОДПФ; б) по алгоритму ОДКП.

Метод предусматривает сжатие неподвижных изображений как полутоновых, так и цветных. При разработке метода предъявлялись следующие основные требования:

- достижение наилучшего соотношения коэффициента сжатия и точности восстановленного изображения по критерию визуального восприятия;
- параметризуемость кодера, то есть наличие параметра, управляющего коэффициентом сжатия;
- применимость к изображениям с различными характеристиками цветности и размерами;
- возможность эффективной программной и аппаратной реализации.

Схема кодера *JPEG* представлена на рисунке 8.11. Рассмотрим более общий случай сжатия цветных изображений, каждый элемент которых представлен 3 - мя байтами, по байту на каждый цветовой компонент (*R*, *G* и *B*).



Рисунок 8.11 Схема кодера *JPEG*.

Кодирование изображения начинается с того, что оно разбивается на отдельные блоки размером 16x16 отсчетов, которые затем кодируются (сжимаются) независимо друг от друга. Далее, в каждом блоке осуществляют переход от 3-х матриц спектральных коэффициентов для красного, зеленого и синего компонентов изображения, к трем матрицам, представляющим яркостный ( $Y$ ), и два цветоразностных компонента изображения ( $Cb$ ) и ( $Cr$ ). Поскольку острота зрения при наблюдении чисто хроматических изображений существенно ниже, чем в случае наблюдения изображений, имеющих яркостный контраст, переход к компонентам  $Cb$  и  $Cr$  выгоден, так как позволяет при их кодировании использовать меньшее количество отсчетов в блоке и за счет этого получить дополнительное сжатие [52]. Более того, хроматические компоненты квантуются грубее. Преобразование цветового координатного пространства  $RGB$  в пространство  $YCrCb$  в соответствии с (3.1) позволяет декоррелировать  $RGB$  компоненты сигнала. Затем матрица, имеющая размер 16x16 отсчетов, разбивается на 4 матрицы яркости  $Y$  размером 8x8 отсчетов каждая и две матрицы цветности размером 8x8 ( $Cb$ ) и ( $Cr$ ) в соответствии с тем, как показано на рисунке 8.12.

X	X	X	O	X	X	O	X
X	X	X	X	O	X	X	X
X	X	X	X	X	X	O	X
O	X	X	X	X	X	X	X
X	X	X	X	X	X	X	X
X	X	X	X	X	X	X	X
O	X	X	O	X	X	O	X
X	X	X	X	X	X	X	X
X	X	X	O	X	X	O	X
O	X	X	X	X	X	X	X

Рисунок 8.12 Положения отсчетов яркости и цветности при формировании макроблока. **X**- положение отсчета яркости, **O** –совпадающее положение отсчетов цветности  $Cb$  и  $Cr$ .

Каждая ячейка на рисунке 8.12 соответствует фрагменту изображения размером 2x2 отсчета. Положения отсчетов цветности двух компонентов совпадают и соответствуют положению центра этого фрагмента. На этом этапе кодирования, с одной стороны, в кодируемое изображение вносятся необратимые искажения за счет прореживания, т.е. происходит потеря информации, а с другой - происходит сжатие данных в два раза. Действительно, до прореживания полное количество отсчетов, которыми был представлен блок изображения, равнялось 3 блокам ( $Y$ ,  $Cr$  и  $Cb$ ) по 256 отсчетов, а после прореживания - только  $16 \times 16 + 2 \times 8 \times 8 = 384$ . Последовательность блоков в макроблоке

представлена на рисунке 8.13. Значения яркостного сигнала из диапазона  $[0,255]$  преобразуются к диапазону  $[-128,127]$  вычитанием уровня  $2^{L-1}=128$ , где  $L$  – число разрядов в представлении компонентов.

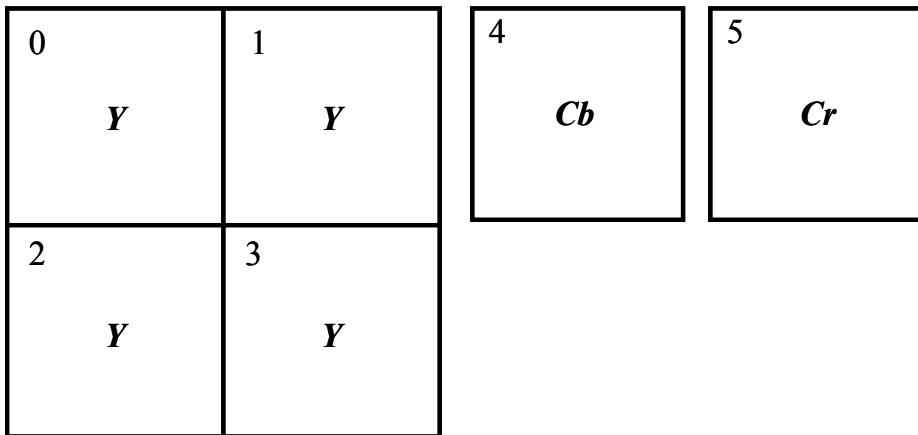


Рисунок 8.13 Расположение блоков в макроблоке при кодировании JPEG.

Далее, в соответствии с последовательностью блоков в макроблоке каждый блок подвергается следующей обработке: двумерное ДКП, квантование на соответствующую матрицу квантования, расположение квантованных спектральных коэффициентов в зигзагообразном порядке, кодирование квантованных спектральных коэффициентов. Двумерное ДКП выполняется в соответствии с (8.18). Квантование заключается в том, что матрица спектральных коэффициентов поэлементно делится на матрицу квантования, имеющую такую же размерность, т.е.  $8 \times 8$ . При этом значение проквантованного спектрального коэффициента определяется следующим образом:

$$A(u, v) = \text{round}(Z(u, v)/Q(u, v)) \text{ для всех } (u, v) \in [0, 7],$$

где  $Z(u, v)$  - исходное, не квантованное, значение спектрального коэффициента, а  $Q(u, v)$  - соответствующий ему по положению в матрице элемент матрицы квантования,  $\text{round}$  – операция округления результата до целого значения.

Матрица квантующих коэффициентов может формироваться в соответствии со спектром изображения. Обычно такая адаптация не выполняется. Часто задается коэффициент качества, определяющий степень сжатия. В зависимости от коэффициента качества *Quality*, матрица квантующих коэффициентов рассчитывается, например, в соответствии с алгоритмом:

$$Q(u, v) = 1 + \text{Quality} \times (1 + u + v), \quad \forall (u, v) \in [0, n-1],$$

где  $n=8$ ,  $\text{Quality} \in [1, 25]$ .

В стандарте *JPEG* для информации приводятся матрицы квантования, оптимизированные по критерию наилучшего визуального восприятия. Примером таких матриц являются следующие:

матрица квантования яркостного компонента:

$$Q_Y = \begin{bmatrix} 16 & 11 & 10 & 16 & 24 & 40 & 51 & 61 \\ 12 & 12 & 14 & 19 & 26 & 58 & 60 & 55 \\ 14 & 13 & 16 & 24 & 40 & 57 & 69 & 56 \\ 14 & 17 & 22 & 29 & 51 & 87 & 80 & 62 \\ 18 & 22 & 37 & 56 & 68 & 109 & 103 & 77 \\ 24 & 35 & 55 & 64 & 81 & 104 & 113 & 92 \\ 49 & 64 & 78 & 87 & 103 & 121 & 120 & 101 \\ 72 & 92 & 95 & 98 & 112 & 100 & 103 & 99 \end{bmatrix}$$

и компонентов цветности:

$$Q_C = \begin{bmatrix} 17 & 18 & 24 & 47 & 99 & 99 & 99 & 99 \\ 18 & 21 & 26 & 66 & 99 & 99 & 99 & 99 \\ 24 & 26 & 56 & 99 & 99 & 99 & 99 & 99 \\ 47 & 66 & 99 & 99 & 99 & 99 & 99 & 99 \\ 99 & 99 & 99 & 99 & 99 & 99 & 99 & 99 \\ 99 & 99 & 99 & 99 & 99 & 99 & 99 & 99 \\ 99 & 99 & 99 & 99 & 99 & 99 & 99 & 99 \\ 99 & 99 & 99 & 99 & 99 & 99 & 99 & 99 \end{bmatrix}.$$

Матрица квантования  $Q$  может быть построена по зональному принципу, в этом случае составляющие ее числа представляют собой величины равные  $2^{12-m}$ , где  $m$  — число уровней, на которое квантуется спектральный коэффициент, входящий в соответствующую зону. Эта процедура интересна тем, что деление обеспечивает приведение спектральных коэффициентов к значениям одного порядка, а округление обеспечивает собственно квантование по уровню. После выполнения операции квантования мы получаем матрицу квантованных спектральных коэффициентов  $A(u, v)$ , особенностью которой является наличие большого количества малых и нулевых спектральных коэффициентов, расположенных преимущественно в правом нижнем углу матрицы.

Следующий шаг алгоритма сжатия состоит в преобразовании полученной матрицы квантованных спектральных коэффициентов  $8 \times 8$  в последовательность из 64 элементов, в которой малые и нулевые спектральные коэффициенты должны быть по возможности сгруппированы. Эта цель достигается путем применения зигзагообразного сканирования, показанного на рисунке 8.14. Расположению квантованных спектральных коэффициентов ДКП в блоке (в соответствии с рисунком 8.14 а) соответствует зигзагообразное сканирование в соответствии с последовательностью номеров элементов, показанной на рисунке 8.14 б) в направлении повышения пространственных частот. Это обусловлено тем, что в области ВЧ коэффициенты квантования имеют большие значения, в

результате многие из спектральных коэффициентов становятся равными нулю.

$A_{00}$	$A_{01}$	$A_{02}$					$A_{07}$
$A_{10}$	$A_{11}$						$A_{17}$
$A_{70}$							$A_{77}$

а) б)

0	1	5	6	14	15	27	28
2	4	7	13	16	26	29	42
3	8	12	17	25	30	41	43
9	11	18	24	31	40	44	53
10	19	23	32	39	45	52	54
20	22	33	38	46	51	55	60
21	34	37	47	50	56	59	61
35	36	48	49	57	58	62	63

Рисунок 8.14 Зигзагообразное сканирование квантованных спектральных коэффициентов.

При таком сканировании нулевые значения располагаются друг за другом, что позволяет дополнительно сократить необходимый поток данных за счет эффективного применения кодирования длин серий *RLE* к последовательностям нулевых коэффициентов. На этапе квантования достигается основное сжатие данных благодаря тому, что спектральные коэффициенты с большими индексами, на которые приходится малая доля энергии изображения, квантуются на малое число уровней (или усекаются), и, следовательно, на их представление затрачивается мало двоичных единиц кода. На этом этапе происходит потеря информации, так как в изображение вносятся необратимые искажения (шум квантования). В соответствии со стандартом *JPEG* квантованный спектральный коэффициент  $A(0,0)$  называется *DC* коэффициентом, а остальные  $A(u,v)$  называются *AC* коэффициентами. *DC* коэффициент пропорционален среднему значению отсчетов изображения в блоке. Поскольку существует высокая корреляционная зависимость между *DC* коэффициентами в соседних блоках, кодируются не сами коэффициенты, а разности значений *DC* в текущем и предыдущем, уже закодированном, блоках (ДИКМ). Первоначальное значение *DC* устанавливается равным 0. Из ряда спектральных коэффициентов *AC* образуются пары чисел, одно из которых равно количеству предшествующих этому спектральному коэффициенту нулей, а другое - значению ненулевого спектрального коэффициента. Полученные пары сжимаются посредством кодирования длин серий. Наконец, применяется энтропийное кодирование без потери информации, обеспечивающее дополнительное сжатие информации. В стандарте *JPEG* представлены таблицы предопределенных кодов переменной длины (кодов Хаффмана) для разностей *DC* коэффициентов и значений *AC* коэффициентов.

Поскольку код Хаффмана является префиксным, то не требуется никаких разделителей между кодовыми словами. Можно использовать и таблицы, полученные по оценке вероятностей кодов данного изображения. Таблицы, так же как и таблицы квантования, хранятся в заголовке файла формата *JPEG* и являются необходимой информацией для декодера.

Если кодирование *JPEG* применяется для внутрикадрового кодирования видеопоследовательностей, так называемый *M-JPEG* (*Motion JPEG*) [52], то при построении кодера/декодера необходимо учитывать, что спектры изображений, подвергаемых кодированию, разные, поэтому задание коэффициента качества не обеспечивает точного значения коэффициента сжатия, что означает переменный размер выходного потока, полученного на выходе кодера. При необходимости обеспечения постоянной скорости передачи выходного потока, схема кодера должна содержать буфер памяти, в который данные поступают со скоростью кодирования, ачитываются с постоянной скоростью и схему управления коэффициентом сжатия, адаптивно изменяющимся в соответствии с размером выходного потока.

Алгоритм декодирования повторяет все операции кодирования в обратном порядке. Декодирование потока. Восстановление значений квантованных спектральных коэффициентов поэлементным умножением на значения соответствующих коэффициентов матрицы квантования  $Q(u,v)$ . ОДКП. Формирование блоков *YCrCb*. Преобразование в *RGB* пространство.

Достоинством описанного метода является высокая степень сжатия данных, которая для цветных изображений достигает 6-10. Большему коэффициенту сжатия соответствуют большие искажения в восстановленном изображении. При больших степенях сжатия на восстановленном изображении просматривается блочная структура. В настоящее время этот метод сжатия данных широко применяется практически во всех графических редакторах.

## 8.8 Кодирование на основе вейвлетных преобразований

В настоящее время все более широкое применение находит сжатие на базе вейвлетного (*Wavelet*) преобразования [57]. (Термин *Wavelet* переводят как маленькая или короткая волна). Этот метод обеспечивает более высокую степень сжатия данных, чем метод, применяемый в стандарте *JPEG*, благодаря тому, что в нем более полно учитываются свойства ЗС, что позволяет не передавать информацию о тех деталях в изображении, отсутствие которых менее заметно. Кодирование по вейвлетному алгоритму предполагает выполнение собственно вейвлетного преобразования для декорреляции информации в изображении, квантования полученных коэффициентов преобразования и кодирования кодами переменной длины полученных коэффициентов. Схема кодера/декодера представлена на рисунке 8.15. Сжатие данных при записи или передаче изображений на основе вейвлетного преобразования

относится к группе методов сжатия с потерей информации. Вейвлетное преобразование позволяет достичь оптимального компромисса между пространственным и частотным разрешением.

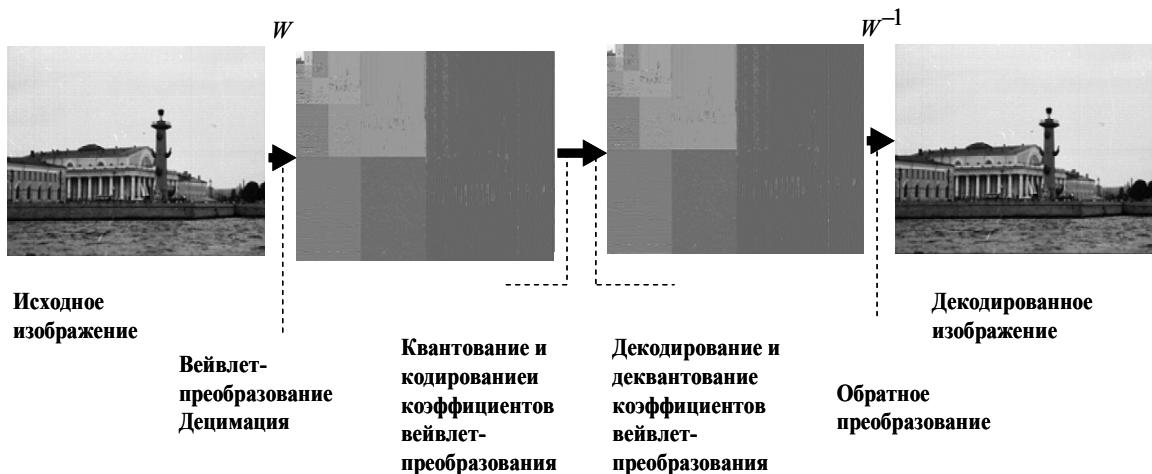


Рисунок 8.15 Условная схема процесса кодирования / декодирования при вейвлетном преобразовании.

В основу вейвлетного преобразования положена идея обеспечения высокого пространственного разрешения при низком частотном, и, наоборот, высокого разрешения по частоте при низком разрешении по пространству [58]. При вейвлетной схеме сжатия также используется разложение сигнала по набору базисных функций:

$$f(x) = \sum_i c_i \psi_i(x),$$

где  $\psi_i(x)$  - базисная функция,  $c_i$  - весовой коэффициент.

Невозможно достичь одновременно высокого разрешения по времени (по пространству  $\Delta x$ ) и по частоте  $\Delta f$ , поскольку их произведение ограничено  $\Delta x \Delta f \geq 1/2$ . Широкобазисные функции позволяют исследовать большие области и точно описать НЧ детали, а короткобазисные функции позволяют исследовать малоразмерные области (ВЧ детали). В связи с этим базисные функции формируются как множество  $\{\psi_i\}$  с конечными носителями разной ширины. В этом случае все базисные функции получаются из одного прототипа  $\psi$  (материнского вейвлета) путем его растяжения (или сжатия) и смещения по оси времени:

$$\psi(2^k t - j), k = \left[0, 2^k - 1\right].$$

Чтобы представить базис ортонормальным необходимо выполнить нормировку:

$$\psi_{kj}(t) = \sqrt{2^k} \psi(2^k t - j).$$

В этом случае вейвлетное преобразование можно представить в виде:

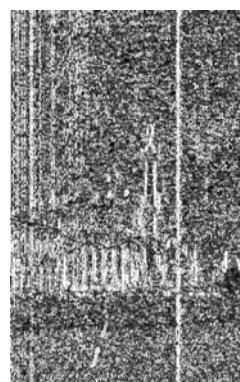
$$f(t) = \sum_k \sum_j c_{kj} \psi_{kj}(t).$$

Двумерное вейвлетное преобразование является разделимым преобразованием и сводится к двум независимо выполняемым одномерным преобразованиям. Одномерное вейвлетное преобразование - это совокупность процессов ВЧ и НЧ фильтрации и децимации. Преобразуемый сигнал подвергается фильтрации с помощью фильтров нижних и верхних частот, которые делят диапазон частот исходного сигнала на две субполосы. ВЧ и НЧ поддиапазоны (субполосы) сигнала, полученные при фильтрации, имеют в два раза более узкую полосу, чем исходный сигнал. Поэтому в соответствии с теоремой Котельникова выборку можно осуществлять с частотой, равной половине частоты дискретизации сигнала, что позволяет отбросить каждый второй отсчет в направлении фильтрации (по строкам или столбцам), то есть выполнить децимацию. Последующие процессы фильтрации производятся над НЧ и ВЧ коэффициентами преобразования. Границы объектов в изображении сигнала в значительной степени сосредоточены в ВЧ субполосе. НЧ субполоса делится снова на две части ВЧ и НЧ. Этот процесс продолжается, пока сигнал не будет полностью разделен или остановлен на заданной стадии разложения.

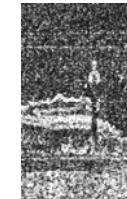
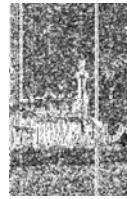
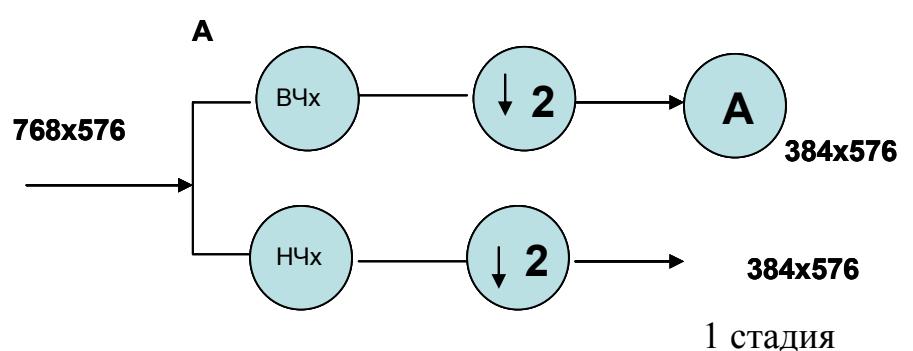
Из коэффициентов вейвлетного преобразования может быть восстановлен оригинальный сигнал. Процесс реконструкции выполняют с помощью обратного вейвлетного преобразования. При работе с изображениями применяют дискретные вейвлетные преобразования – прямое (ДВП) и обратное - ОДВП. Применяют различные схемы ВП, рассмотрим одну из них, представленную на рисунке 8.16. Эта схема является модифицированным представлением пирамидальной схемы Малла [27].

Пусть исходное изображение имеет размер 768x576 элементов. На первой стадии изображение подвергается ВЧ и НЧ фильтрации и децимации. При этом ВЧ составляющая образует блок А (обозначим  $H$  - соответствие ВЧ фильтру,  $L$  - НЧ фильтру), ширина которого (размер по строке) в два раза меньше, чем ширина изображения. Блок на выходе НЧ фильтра после децимации имеет размер, равный размеру блока А. Таким образом, общий размер матрицы коэффициентов преобразования равен размеру кодируемого изображения. Этот процесс представлен на рисунке 8.16 как стадия 1. Затем выполняются ВЧ и НЧ фильтрации с децимацией по строкам только над НЧ ( $L$ ) блоком 1 стадии. Новая пара блоков будет иметь размер 192x576 ( $LL$  и  $LH$  - блоки). Каждый из этих блоков подвергается ВЧ и НЧ фильтрации в направлении по столбцу. Из  $LH$  блока получаются блоки  $B$  и  $C$ . Из блока  $LL$  на выходе ВЧ фильтра формируется блок  $D$ , а на выходе НЧ фильтра формируется блок  $LL_1$ . Все эти блоки имеют размер 192x288. Далее для НЧ блока  $LL_1$  может повторяться несколько раз процедура, выполненная на стадии 2. Для цветного изображения выполняется преобразование цветового координатного пространства  $RGB$  в пространство  $YCrCb$ , поддискретизация компонентов

цветности по схеме 4:2:2 и вейвлет - преобразование компонентов  $Cr$  и  $Cb$  аналогично компоненту яркости.



**Исходное изображение**



**B**

**C**

**D**

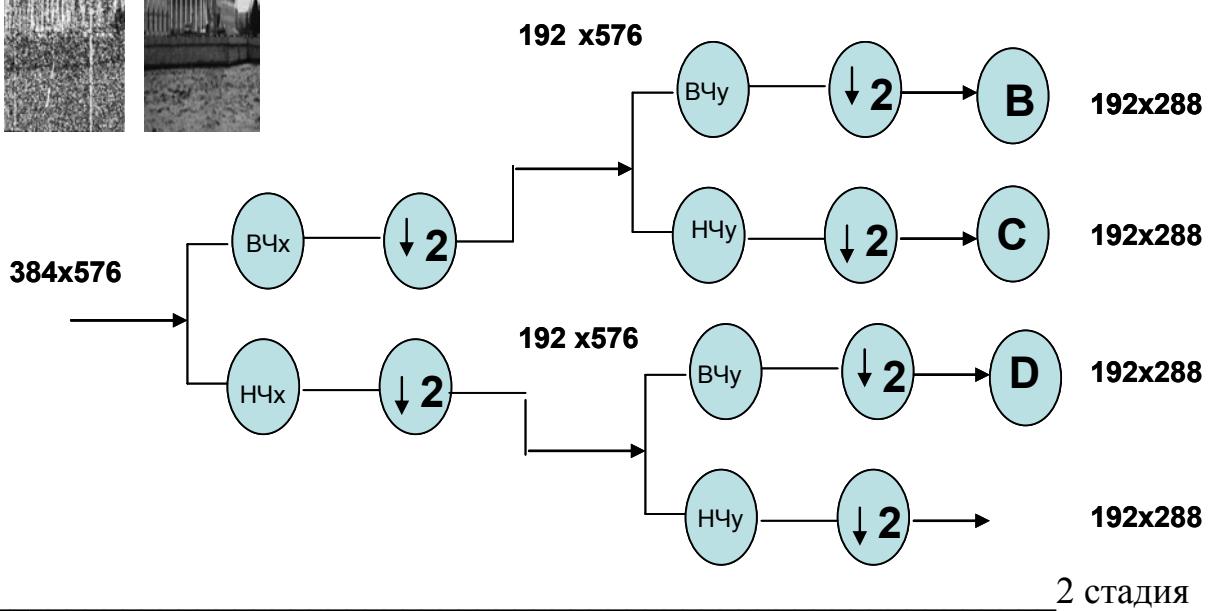


Рисунок 8.16 Схема вейвлет-кодирования изображения.

На рисунке 8.17 представлено пространственное расположение блоков коэффициентов вейвлетного преобразования в соответствии с пространственным расположением элементов изображения.

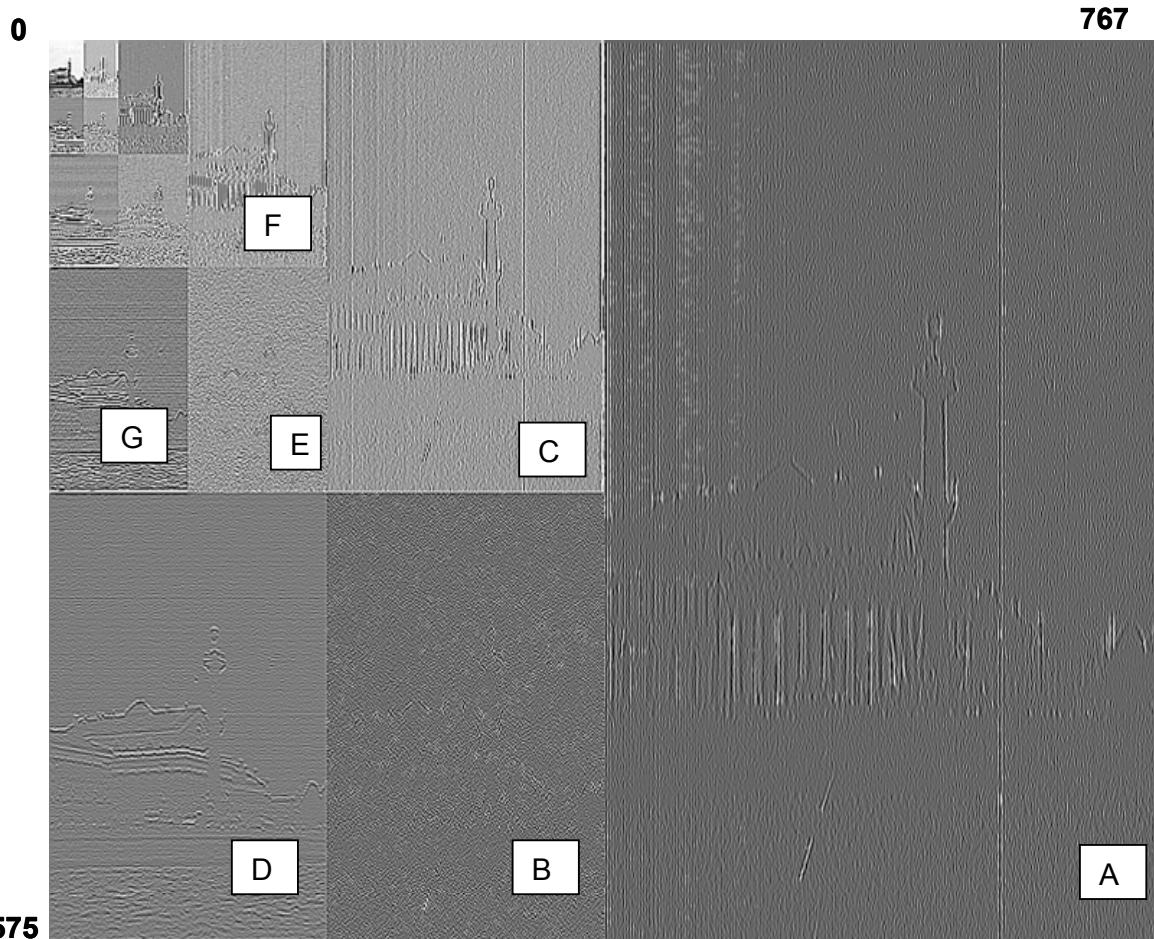


Рисунок 8.17 Пространственное расположение блоков коэффициентов в кадре изображения после вейвлетного преобразования.

Размер матрицы коэффициентов преобразования соответствует размерам кодируемого изображения. В 1997 г. фирмой *Analog Devices* выпущена первая микросхема *ADV601* [59], выполняющая вейвлет – преобразование в реальном времени в соответствии с представленной схемой, число стадий равно 5.

Выбор базиса вейвлетов для кодирования изображения является трудной задачей. Известен ряд критериев построения «хороших» вейвлетов, среди которых наиболее важными являются: гладкость, точность аппроксимации, величина области определения, частотная избирательность фильтра. Для кодирования/декодирования изображений удобно строить вейвлеты на основе ортогональных базисных функций.

Простейшим видом вейвлет-базиса для изображений является разделимый базис, получаемый сжатием и растяжением одномерных вейвлетов. Основополагающей в вейвлет-анализе является идея о выделении информации при различных уровнях детализации.

Доказано, что для ортогональных вейвлетов точное восстановление сигнала возможно при дополнительном использовании

аппроксимирующей функции. Рассмотрим подробнее ДВП и ОДВП с использованием вейвлета Хаара. Хаар в 1910 году описал полную ортонормальную систему базисных функций с локальной областью определения, то есть имеющих компактный носитель:

$$\psi(t) = \begin{cases} 1, & \text{если } 0 \leq t < 1/2 \\ -1, & \text{если } 1/2 \leq t < 1 \\ 0, & \text{иначе} \end{cases} \quad (8.21)$$

Функция (8.21) называется материнским вейвлетом. Масштабированный и сдвинутый вариант материнского вейвлета Хаара  $\psi_{k,j}(t)$  определяется следующим образом:

$$\psi_{k,j}(t) = \psi(2^k t - j), \quad j = 0, \dots, 2^k - 1. \quad (8.22)$$

Аппроксимирующая функция Хаара имеет вид:

$$\varphi(t) = \begin{cases} 1, & 0 \leq t < 1 \\ 0, & \text{иначе} \end{cases} \quad (8.23)$$

На рисунке 8.18 показано, как выглядят аппроксимирующая функция (а), материнский вейвлет  $\psi(t)$  Хаара (б), и масштабированный во времени вейвлет  $\psi(2t)$  и масштабированный и сдвинутый вейвлет  $\psi(2t - 1)$ .

Хаар не называл эту функцию вейвлетом. В области цифровой обработки и анализа сейсмических сигналов в работах А. Гроссмана и Ж. Морле было предложено посыпать вглубь Земли вместо импульсов одинаковой длины короткие волновые образования, полученные просто масштабированием одной функции. Именно ее потом и назвали вейвлетом.

На основании (8.21) можно построить фильтры Хаара: НЧ с элементами импульсной характеристики  $h_0=1/2$ ,  $h_1=1/2$  и ВЧ с импульсными характеристиками  $g_0=1/2$ ,  $g_1=-1/2$ . Прямым преобразованием Фурье можно получить передаточные характеристики этих фильтров:

$$H(\omega) = \sum_{k \in \mathbf{Z}} h_k \exp(-i\omega k) = 1/2 + 1/2 \exp(-i\omega) = \cos(\omega/2) \exp(-i\omega/2), \quad (8.24)$$

$$G(\omega) = \sum_{k \in \mathbf{Z}} g_k \exp(-i\omega k) = 1/2 - 1/2 \exp(-i\omega) = i \sin(\omega/2) \exp(-i\omega/2), \quad (8.25)$$

где  $\mathbf{Z}$ -любое целое число от  $-\infty$  до  $+\infty$ .

Из (8.24) реальная часть  $Re H(\omega) = \cos^2(\omega/2)$ , а мнимая часть  $Jm H(\omega) = -\cos(\omega/2) \sin(\omega/2)$ , откуда модуль:

$$|H(\omega)| = \cos(\omega/2). \quad (8.26 \text{ а})$$

Аналогично из (8.24)

$$|G(\omega)| = |\sin(\omega/2)|. \quad (8.26 \text{ б})$$

Рассмотрим входную последовательность  $\{f_n\}$ . Преобразование Фурье такой последовательности обозначим  $F(\omega)$ . Функция

$A(\omega) = H(\omega)F(\omega)$  представляет преобразование Фурье сигнала на выходе фильтра, а сам сигнал находится в виде свертки:

$$a_n = \sum_{k=0}^1 h_k f_{n-k} = \frac{1}{2} f_n + \frac{1}{2} f_{n-1}. \quad (8.27)$$

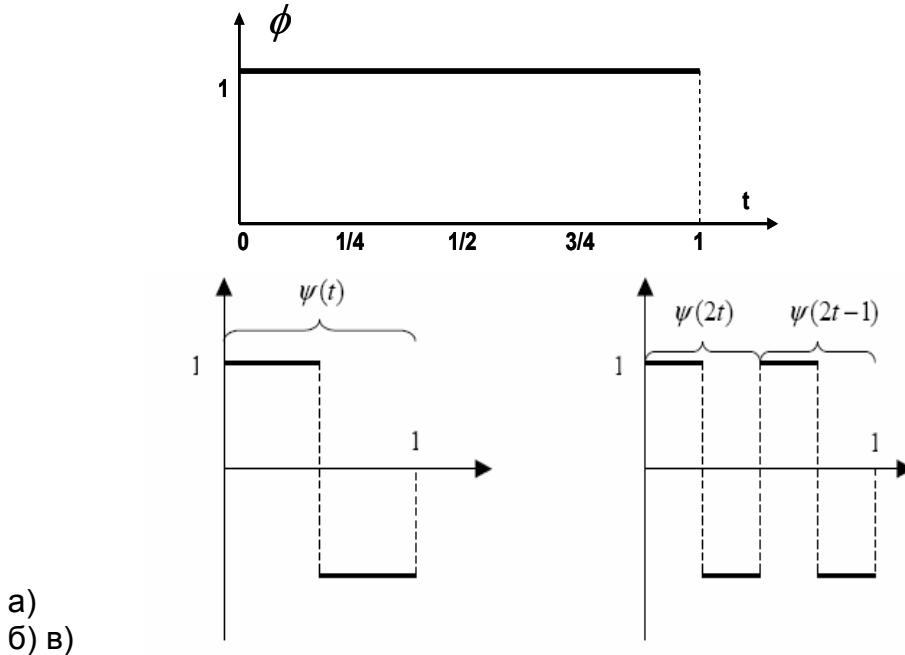


Рисунок 8.18 Вейвлет Хаара. а) Масштабирующая функция («отцовский вейвлет»), б) материнский вейвлет, в) производные вейвлеты Хаара.

Аналогично для высокочастотного фильтра

$$d_n = \sum_{k=0}^1 g_k f_{n-k} = \frac{1}{2} f_n - \frac{1}{2} f_{n-1}. \quad (8.28)$$

Таким образом, НЧ фильтр производит усреднение соседних отсчетов, а ВЧ фильтр формирует первую конечную разность. В результате исходную последовательность длиной  $N$  можно разложить на две последовательности длиной  $N/2$ : так называемые аппроксимирующую  $\{a_n\}$  и детализирующую  $\{d_n\}$  последовательности, или НЧ и ВЧ субполосы.

Из (8.27) и (8.28) сложением правых и левых частей получаем формулу обратного преобразования, позволяющую восстановить исходный сигнал:

$$f_n = a_n + d_n. \quad (8.29)$$

В общем случае задачу восстановления сигнала решают, строя квадратурно-зеркальные фильтры в соответствии с уравнениями (8.31), (8.32) [60]:

$$\tilde{H}(\omega) = \bar{H}(\omega), \quad (8.31)$$

$$\tilde{G}(\omega) = -\exp(-i\omega)H(\omega + \pi), \quad (8.32)$$

где черта означает комплексное сопряжение.

Для этих фильтров

$$G(\omega) = -\exp(i\omega)\bar{H}(\omega + \pi). \quad (8.33)$$

В случае вейвлета Хаара восстанавливающие фильтры получим в соответствии с (8.31), (8.32), используя (8.23) и (8.24) соответственно:

$$\tilde{H}(\omega) = 1/2 + 1/2\exp(i\omega),$$

$$\tilde{G}(\omega) = -\exp(-i\omega)H(\omega + \pi) = -1/2\exp(-i\omega) + 1/2\exp(-i\omega/2).$$

Отсюда получаем коэффициенты восстанавливающих НЧ и ВЧ фильтров соответственно:

$$\tilde{h}_{-1} = 1/2, \quad \tilde{h}_0 = 1/2; \quad (8.34)$$

$$\tilde{g}_1 = -1/2, \quad \tilde{g}_2 = 1/2. \quad (8.35)$$

Вейвлет Хаара обладает компактным носителем, но плохо локализован в частотной области. Ингрид Добеши показала в 1988 году, что компактный носитель возможен не только для вейвлетов Хаара. Добеши построила ортогональные вейвлеты с компактным носителем, локализованные и в частотной области, обладающие свойствами гладкости и исключения моментов. Для больших по площади и относительно гладких изображений они дают большое преимущество при сжатии. Для таких изображений не только уменьшается общая ошибка декодирования при заданном коэффициенте сжатия, но также существенно уменьшается блочность, имеющая место при сжатии с помощью вейвлетов Хаара.

Рассмотрим прямое и обратное дискретные вейвлетные преобразования на примере фильтра Добеши 4-го порядка, обозначаемого D4 ( $n=3$ ) [61]. Коэффициенты, определяющие вейвлеты Добеши, равны:

$$c_0 = \frac{1+\sqrt{3}}{4\sqrt{2}}, \quad c_1 = \frac{3+\sqrt{3}}{4\sqrt{2}}, \quad c_2 = \frac{3-\sqrt{3}}{4\sqrt{2}}, \quad c_3 = \frac{1-\sqrt{3}}{4\sqrt{2}}. \quad (8.36)$$

На основании этих коэффициентов НЧ фильтра можно получить коэффициенты ВЧ фильтра разложения и коэффициенты фильтров реконструкции в соответствии с уравнениями (8.31) - (8.33). В этом случае импульсная характеристика НЧ фильтра может быть представлена 4 значениями:  $h_2 = c_0 = 0,482962913145$ ,  $h_1 = c_1 = 0,836516303738$ ,  $h_0 = c_2 = 0,224143868042$ ,  $h_{-1} = c_3 = -0,129409522551$ .

В соответствии с (8.24) импульсная характеристика НЧ фильтра может быть представлена в виде:

$$H(\omega) = h_0 + h_1 \exp(-i\omega) + h_2 \exp(-2i\omega) + h_{-1} \exp(i\omega).$$

В соответствии с (8.33) импульсная характеристика ВЧ фильтра имеет вид:

$$G(\omega) = -h_0 \exp(-i\omega) + h_1 - h_2 \exp(-i\omega) + h_{-1} \exp(-2i\omega).$$

Отсюда коэффициенты импульсной характеристики ВЧ фильтра определяются соотношениями:  $g_2 = h_{-1} = c_3$ ,  $g_1 = -h_0 = -c_2$ ,

$$g_0 = h_1 = c_1, \quad g_{-1} = -h_2 = -c_0.$$

Строка изображения  $f(j)$ ,  $j = [0, N - 1]$  преобразуется в две подстроки.

Выполняется разложение сигнала и децимация, формируются НЧ и ВЧ компоненты по формуле:

$$a_j = \sum_{k=-1}^2 h_k f_{2j-k}, \quad (8.37)$$

$$d_j = \sum_{k=-1}^2 g_k f_{2j-k}. \quad (8.38)$$

Для вычисления коэффициентов преобразования используется циклическое повторение строки:

$$\dots a(N-1) [ a(0) \ a(1) \ a(2) \ \dots \ a(N-2) \ a(N-1) ] \ a(0) \ a(1) \ \dots$$

На рисунке 8.19 показано соотношение положения отсчетов сигнала и коэффициентов НЧ и ВЧ фильтров, соответствующее уравнениям (8.37)-(8.38).

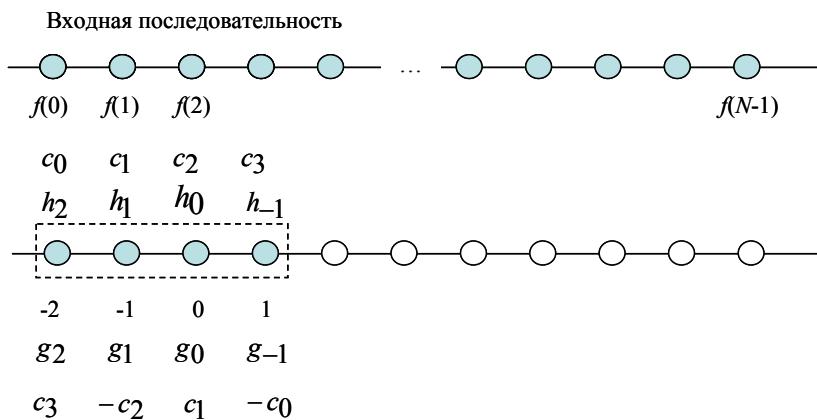


Рисунок 8.19 Положение коэффициентов фильтров и отсчетов сигнала входной последовательности.

Коэффициенты фильтра применяются ко всем последовательным отсчетам входной выборки в окне фильтра, но устанавливается окно через один отсчет, как показано на рисунке 8.20. Пунктиром обозначены положения окон фильтров при формировании последовательных отсчетов вейвлетного преобразования. Вейвлет – преобразование обратимо.

Значения НЧ и ВЧ коэффициентов реконструкции вычисляются в соответствии с (8.31), (8.32) по формулам:

$$\tilde{h}_{-n} = h_n; \quad (8.39)$$

$$\tilde{g}_{-n} = g_n, \quad (8.40)$$

откуда  $\tilde{h}_1 = c_3, \tilde{h}_0 = c_2, \tilde{h}_{-1} = c_1, \tilde{h}_{-2} = c_0; \tilde{g}_1 = -c_0, \tilde{g}_0 = c_1, \tilde{g}_{-1} = -c_2, \tilde{g}_{-2} = c_3$ .

Используя полученные коэффициенты ОДВП, можно точно восстановить исходное изображение. Декодирование выполняется путем применения ОДВП к прореженному массиву коэффициентов. При одномерном ОДВП при реконструкции отсчетов сигнала необходимо увеличить входную последовательность коэффициентов в два раза, вставив нули между отсчетами.

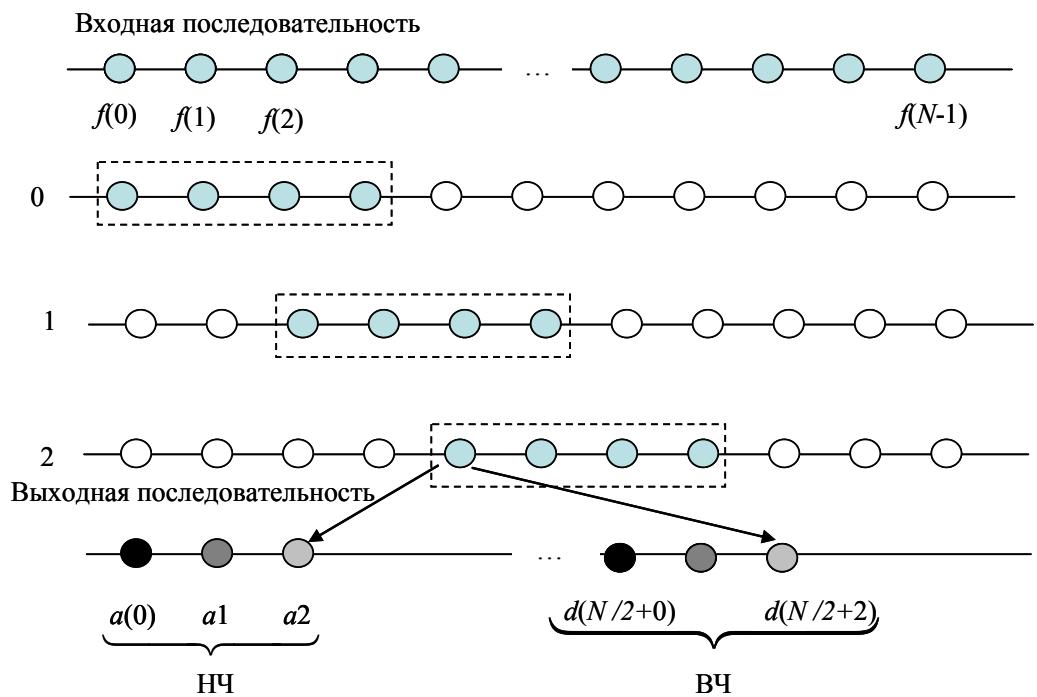


Рисунок 8.20 Схема выполнения одномерного ВП и децимации.

Выполнить свертку с соответствующими НЧ и ВЧ фильтрами и сложить полученные результаты. Схема реконструкции представлена на рисунке 8.21. Здесь черные кружки обозначают нулевые значения отсчетов коэффициентов преобразования. Значения двух последовательных отсчетов восстанавливаемой строки изображения вычисляются на основании информации о двух аппроксимирующих коэффициентах и двух детализирующих коэффициентах. Выполняется свертка с НЧ и ВЧ фильтрами восстановления и результатирующие значения складываются. Пунктирным контуром обозначены два последовательных положения окон фильтров для вычисления  $\tilde{f}(4)$  (левее) и  $\tilde{f}(5)$ . Восстановленные значения равны соответствующим значениям кодируемого сигнала:  $\tilde{f}(4)=f(4)$ ,  $\tilde{f}(5)=f(5)$ . Другое обозначение использовано только для того, чтобы показать, что это реконструированные значения. Вследствие нулевых значений коэффициентов восстановленные значения сигнала вычисляются по формулам:

$$\tilde{f}(4)=c_2 a(1)+c_0 a(2)+c_1 d(1)+c_3 d(2),$$

$$\tilde{f}(5)=c_3 a(1)+c_1 a(2)-c_0 d(1)-c_2 d(2).$$

6, 7 отсчеты восстанавливаются аналогично по парам коэффициентов  $a(2)$ ,  $a(3)$  и  $d(2)$ ,  $d(3)$ .

Краевые эффекты учитываются циклическим повторением строки. На примере фильтра  $D4$  мы рассмотрели схему одномерного ДВП, представленную на рисунке 8.22.

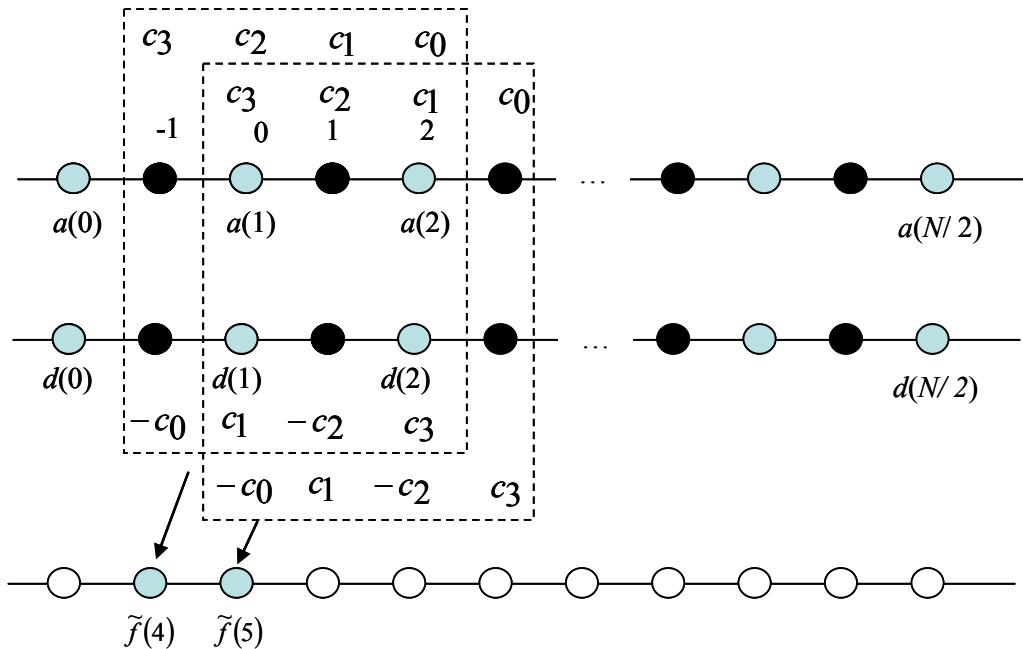


Рисунок 8.21 Схема выполнения одномерного ОДВП при использовании фильтра  $D4$ .

Рассмотренные вейвлеты Хаара и Добеши являются ортогональными. Они удовлетворяют следующим условиям:

$$\left. \begin{array}{l} (\phi_{k,j}, \phi_{k,l}) = 0 \\ (\psi_{k,j}, \psi_{k,l}) = 0 \end{array} \right\} \text{при } j \neq l. \quad (8.41)$$

$$(\phi_{k,j}, \psi_{k,l}) = 0 \text{ при всех } j, l. \quad (8.42)$$

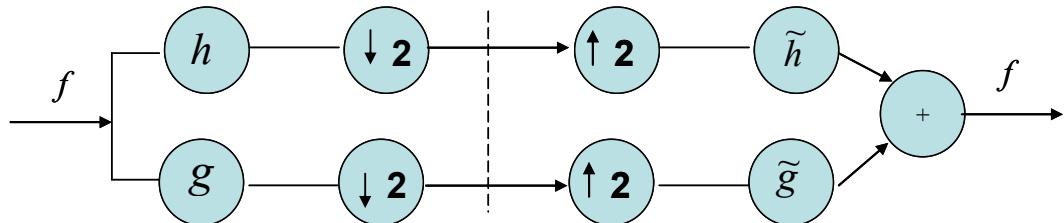


Рисунок 8.22 Схема прямого и обратного одномерного ВП.

При кодировании изображений широко используются биортогональные вейвлеты, позволяющие выполнить более эффективное сжатие. Биортогональность относится к концепции двойственности. Предположим, что  $\{u_1, u_2, \dots, u_n\}$ - множество неортогональных базисных функций. Мы можем представить функцию  $f$  в виде линейной комбинации этих базисных функций:

$$f(x) = \sum_{j=1}^n a_j u_j(x).$$

Отсутствие ортогональности усложняет определение коэффициентов  $a_j$ . Однако существует другой базис  $\tilde{u}_1, \tilde{u}_2, \dots, \tilde{u}_n$ , такой что  $a_j = (f, \tilde{u}_1)$ . Функции  $\tilde{u}_1$  обладают также свойством:

$$\left( u_k, \tilde{u}_j \right) = 0, \text{ если } j \neq k.$$

Базис  $\{\tilde{u}_1, \tilde{u}_2, \dots, \tilde{u}_n\}$  называется двойственным базисом, соответствующим  $\{u_1, u_2, \dots, u_n\}$ . Биортогональные вейвлетные системы состоят из четырех множеств функций: базиса масштабирующих функций  $\{\phi_{k,j}\}$ , двойственного к нему базиса  $\{\tilde{\phi}_{k,j}\}$ , базиса вейвлет-функций  $\{\psi_{k,j}\}$ , двойственного к нему базиса  $\{\tilde{\psi}_{k,j}\}$ . Условие биортогональности требует, чтобы эти множества функций удовлетворяли следующему свойству:

$$\left. \begin{array}{l} (\phi_{k,j}, \tilde{\psi}_{k,l}) = 0 \\ (\psi_{k,j}, \tilde{\phi}_{k,l}) = 0 \end{array} \right\} \text{при всех } j, k, l. \quad (8.43)$$

Кроме того, двойственность влечет за собой

$$\left. \begin{array}{l} (\phi_{k,j}, \tilde{\phi}_{k,l}) = 0 \\ (\psi_{k,j}, \tilde{\psi}_{k,l}) = 0 \end{array} \right\} \text{при } j \neq l. \quad (8.44)$$

Сжатие информации производится за счет квантования и кодирования коэффициентов преобразования. В качестве показателя качества используется дисперсия:

$$D_{\Sigma} = E \|\mathbf{X} - \mathbf{Y}\|^2, \quad (8.26)$$

где  $E$  - математическое ожидание,  $D_{\Sigma}$  - дисперсия, характеризующая суммарное искажение,  $\mathbf{X}$  и  $\mathbf{Y}$  - векторы исходного и реконструированного изображений, соответственно.

В качестве критерия оптимизации используется минимум дисперсии при заданных ограничениях на ресурс бит [57]:

$$D_{\Sigma} \rightarrow \min, \quad \text{при } R_{\Sigma} \leq R, \quad (8.27)$$

где  $R$  - ресурс разрядов, отводимых на кодирование.

От такой постановки условной задачи оптимизации осуществляется переход к безусловной оптимизации путем рассмотрения функционала Лагранжа:

$$(J(\lambda) = D_{\Sigma} + \lambda R_{\Sigma}) \rightarrow \min, \quad (8.28)$$

где  $\lambda$  - множитель Лагранжа. Критерием оптимизации является минимум функции Лагранжа:

$$(J(\lambda) = D_{\Sigma} + \lambda R_{\Sigma}) \rightarrow \min. \quad (8.29)$$

Известно, что задача (8.29) эквивалентна (8.27) для частного случая  $R_{\Sigma} = R$  [57].

Использование функционала Лагранжа в качестве критерия оптимизации является более общим, чем использование критериев только

скорости или только искажения. Эти критерии есть частные случаи, соответствующие крайним точкам ( $\lambda = \infty$ ) и ( $\lambda = 0$ ).

Таким образом, необходимо найти некоторые параметры алгоритма сжатия на основе вейвлет-преобразования, которые бы минимизировали комбинацию: 1) числа разрядов (или скорости  $R_{\Sigma}$ ), требуемых для представления сигнала; 2) искажения  $D_{\Sigma}$ , вычисляемого как дисперсия разности между оригинальным сигналом и его аппроксимацией с использованием  $R_{\Sigma}$  бит. Взаимообмен между скоростью и искажением представляется рабочей функцией скорость-искажение, которая определяет минимальное достижимое искажение при заданной скорости или минимально достижимую скорость при заданном искажении квантователя. Выбор каждого сочетания скорости и искажения определяет одну точку на кривой.

Решение этой задачи часто на практике заменяется интерполяцией коэффициентов квантования при заданных таблицах квантования для высокого и низкого качества реконструированного изображения, как это выполняется, например, при кодировании в рассмотренном методе, реализованном в *ADV601*. Главная задача при оценке коэффициентов квантования: обеспечить распределение искажений по полосам таким образом, чтобы большие искажения были допущены в ВЧ субполосах и меньшие в НЧ субполосе, в соответствии с особенностями ЗС человека. По этому критерию формируются опорные таблицы из 42 коэффициентов квантования для самого большого и самого низкого коэффициентов сжатия. Выполняется 5 стадий ДВП, в результате которого получается 14 субполос (блоков) для каждого компонента цветового пространства *YCrCb*, в соответствии с представленным на рисунке 8.17 для *Y* компонента. Здесь блокам от *A* до *N* соответствуют номера 0, 3, 6, 9, 12, 15, 18 и т.д. Соответствующим блокам *Cb* компонента – 1, 4, 7, 10, 13, и т.д. Соответствующим блокам *Cr* компонента – 2, 5, 8, 11, 14, и т.д. Затем вычисляются искажения для каждой субполосы, оцениваемые как функция суммы квадратов коэффициентов преобразования. Исходя из этой оценки и установленной скорости кодирования, путем интерполяции формируется 42 коэффициента квантования. Каждой субполосе, независимо от других субполос, назначается свой ресурс бит. Чем больше коэффициент квантования, тем больше искажения, но тем выше коэффициент сжатия. Типичные значения коэффициентов квантования представлены в таблице 8.6. Затем каждый коэффициент вейвлетного преобразования данной полосы квантуется. При этом ВЧ компоненты квантуются на меньшее число уровней, НЧ - на большее. Адаптация квантователей может производиться по критерию постоянного качества декодированного изображения либо постоянной скорости кодирования, необходимой при передаче сигнала по каналам связи.

Дальнейшее сжатие достигается применением энтропийного кодера и кодирования длин серий.

Таблица 8.6. Типичные значения коэффициентов квантования

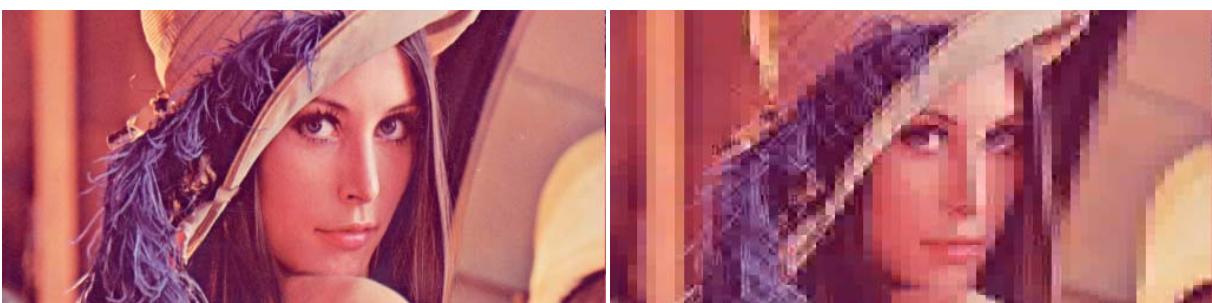
<i>Y</i>		<i>Cb</i>		<i>Cr</i>	
№ блока	Коэффициент квантования	№ блока	Коэффициент квантования	№ блока	Коэффициент квантования
0	164,09	1	198,16	2	198,16
3	35,83	4	43,27	5	43,27
6	29,86	7	36,06	8	36,06
9	29,86	10	36,06	11	36,06
12	12,10	13	14,61	14	14,61
15	10,09	16	12,18	17	12,18
18	10,09	19	12,18	20	12,18
21	3,0	22	3,63	23	3,63
24	2,5	25	3,02	26	3,02
27	2,5	28	3,02	29	3,02
30	0,89	31	1,08	32	1,08
33	0,74	34	0,898	35	0,898
36	0,74	37	0,898	38	0,898
39	0,496	40	0,6	41	0,6

Даже при адаптации коэффициентов квантования метод не позволяет сохранить постоянным коэффициент сжатия. Поскольку сокращение информации производится на стадии квантования и кодирования кодами переменной длины, то задание одних и тех же коэффициентов квантования приводит к формированию потоков разной длины. Длина выходного потока определяется информацией, содержащейся в изображении. Поддержание выходного потока постоянной скорости достигается за счет управляемой буферизации выходного потока. Причин успешного применения вейвлетов именно для изображений несколько:

- Вейвлет - преобразование хорошо аппроксимирует преобразование Карунена-Лоэва.
- В результате перераспределения дисперсий коэффициенты вейвлет - преобразования имеют существенно негауссовскую статистику и, таким образом, меньшую энтропию, чем гауссовский сигнал той же дисперсии.
- Дисперсии коэффициентов субполос вейвлет-преобразования распределены в широком диапазоне значений. Пусть дисперсии кодируются простым энтропийным кодером. Тогда стоимость кодирования всего изображения есть сумма стоимостей кодирования субполос. Различные энтропии субполос приведут к стоимости кодирования значительно меньшей, чем при непосредственном кодировании изображения.

- Коэффициенты вейвлетного преобразования имеют регулярные пространственно - частотные зависимости, которые с успехом используются в ряде алгоритмов кодирования.

При большом сжатии данных этот метод приводит к появлению артефактов на восстановленных изображениях в виде появления окантовок и посторонних узоров. Но, поскольку ДВП не требует формирования блоков изображения, артефакты сжатия на его основе более «естественны», т.е. выглядят менее чужеродными на типичных изображениях, чем, например, блочная структура при сжатии *JPEG*. Примеры артефактов, производимых при кодировании по вейвлетному алгоритму и алгоритму *JPEG* при коэффициенте сжатия 60 показаны на рисунке 8.23.



а) б)

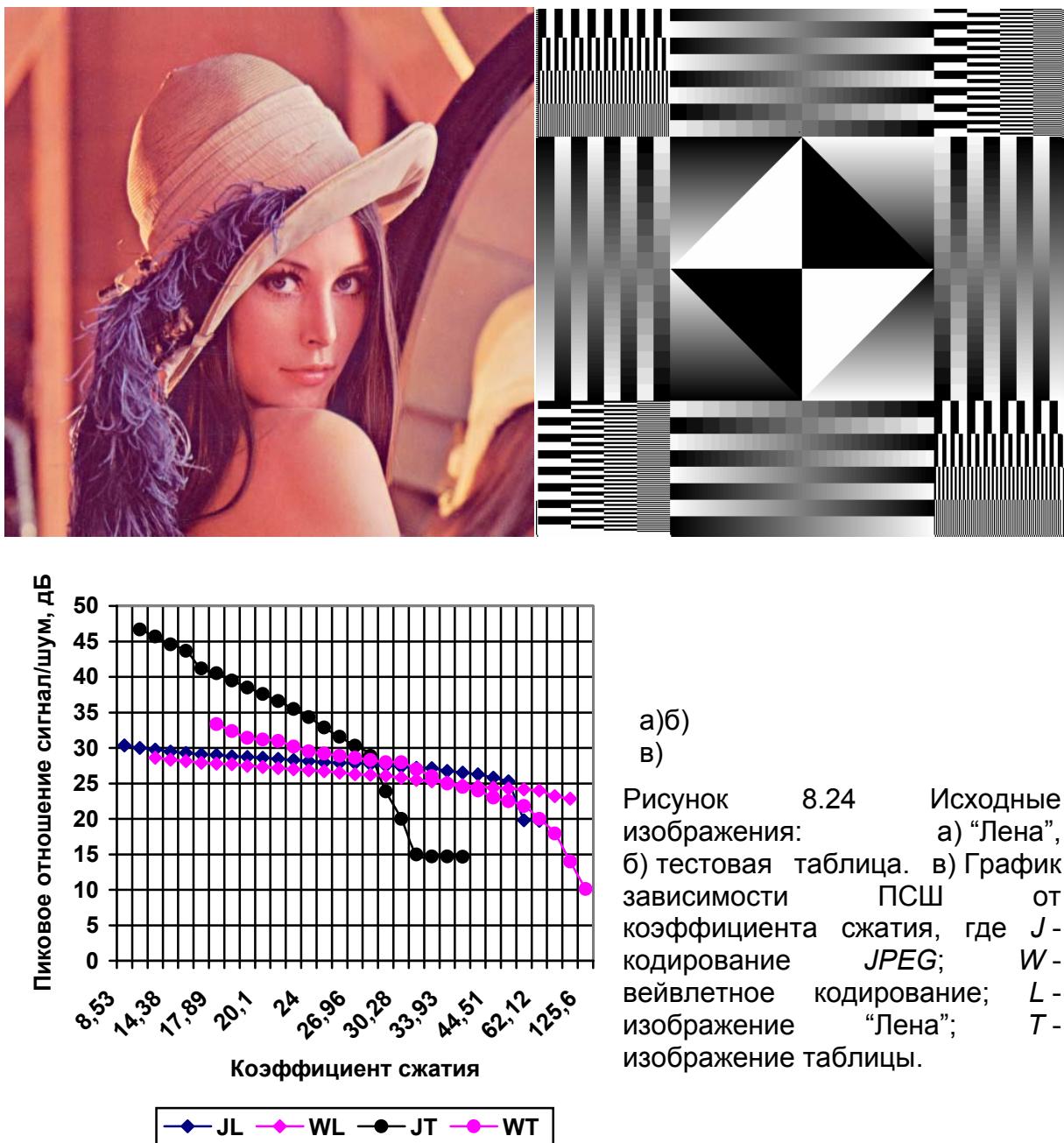
в)

Рисунок 8.23 Фрагменты изображения "Лена". а) Фрагмент исходного изображения; б) Фрагмент декодированного JPEG изображения; в) фрагмент декодированного изображения при использовании вейвлетного алгоритма.



Проведены исследования эффективности методов *JPEG* и вейвлетного кодирования на серии цветных и монохромных изображений. Оценка эффективности производилась по двум параметрам: коэффициенту сжатия и пиковому отношению сигнал/шум. Каждое изображение кодировалось по алгоритмам сжатия *JPEG* и вейвлетному. Производилось декодирование изображения и формировалось изображение разности исходного и декодированного изображений. По этому изображению оценивались коэффициент сжатия и ПСШ. Результаты исследования для двух тестовых изображений и сами эти изображения приведены на рисунке 8.24.

Проведенные исследования показали, что *JPEG* – кодеры обеспечивают более высокое качество изображения по сравнению с вейвлет-кодерами при коэффициентах сжатия до 15. Однако диапазон этих значений и сами оценки существенно изменяются в зависимости от типа изображения. Можно, однако, утверждать, что при больших коэффициентах сжатия вейвлет-кодер является более эффективным.



а), б)  
в)

Рисунок 8.24 Исходные изображения:  
а) "Лена",  
б) тестовая таблица. в) График зависимости ПСШ от коэффициента сжатия, где  $J$  - кодирование JPEG;  $W$  - вейвлетное кодирование;  $L$  - изображение "Лена";  $T$  - изображение таблицы.

## 8.9 Стандарт JPEG 2000

Разработан для кодирования неподвижных изображений. Является развитием стандарта JPEG. В качестве базисного метода использует ДВП. Обеспечивает большее сжатие, чем JPEG.

## 8.10 Стандарты сжатия видео последовательностей

Успех цифрового телевидения и DVD-видео обусловлен введением стандарта ISO/IEC 13818, известного как стандарт MPEG-2 по названию группы разработчиков. *MPEG* - *Moving Picture Experts Group* - эта организация, состоящая из нескольких сотен технических экспертов, занимающаяся разработкой и внедрением стандартов для сжатия,

обработки и представления движущихся изображений и звука, является рабочей группой *ISO* и *IEC*.

### 8.10.1 Стандарт *MPEG-1*

Стандарт разработан для просмотра и хранения видео на *CD* [14]. Принят в 1993 г, *ISO/IEC* 11172. Предназначен для сжатия видео и аудио информации для *CD*-проигрывателей. Предусмотрен для сжатия со скоростью 1,5 Мб/с. Типичный бытовой формат видео для *MPEG-1* в стандарте *PAL* составляет 352x288 пикселов, 25 кадров в секунду. Аудио часть - стереозвук с частотой дискретизации 44,1 кГц, сжатый в *MPEG-1 Layer II*. Качество видео ненамного превосходит *VHS*. В стандарте применена компенсация движения, ДКП и квантование.

*Временная обработка.* Временная (межкадровая) избыточность определяется тем, что два соседних кадра видеопоследовательности, как правило, мало отличаются друг от друга. Для устранения временной избыточности строится кадр-прогноз, который вычитается из текущего кадра. Простейший метод прогнозирования основан на предыдущем кадре, используемом как прогноз текущего кадра. Чем лучше прогноз, тем меньше энергия остаточного кадра, который кодируется и посыпается декодеру. Повышение точности предсказания при передаче видеопоследовательности достигается за счет оценки вектора движения и компенсации этого движения. Компенсация движения при формировании сигнала ошибки предсказания уменьшает его значение и, как следствие, увеличивает сжатие. Полностью сигнал ошибки предсказания не может быть сведен к нулю, так как при движении объекта обычно изменяется его ракурс и масштаб, однако он может быть значительно уменьшен.

Определено три типа изображений при кодировании. Внутрикадровое кодирование *I* - кадров (*Intra coded pictures, I - pictures*) выполняется без ссылок на другие изображения со средним коэффициентом сжатия. *I* - кадры обеспечивают возможность произвольного доступа к любому кадру, являясь своеобразными точками входа в поток данных для декодера и сжимаются независимо. Кадры, кодируемые с предсказанием (*Predictive coded pictures, P -pictures*) кодируются более эффективно, используя предсказание с компенсацией движения по предыдущему *I* - или *P* - кадру. В результате предсказания формируется кадр сигнала ошибки предсказания, который представляет собой разность между опорным и предсказанным кадрами с учетом векторов движения. Этот кадр подвергается кодированию посредством применения той же последовательности операций, которая применяется для кодирования *I* - кадров. Кадры, кодируемые с предсказанием в двух направлениях (*Bidirectionally predictive coded pictures, B - pictures*) обеспечивают наибольшую степень сжатия, но для компенсации движения требуют двух ссылок на предыдущий и последующий *I* или *P* кадры. *B* - кадры никогда не используются для предсказания. Кадры различных типов объединяются в группы. Каждая группа начинается с

кадра типа  $I$ , образующего опорный сигнал для предсказания при кодировании кадров типа  $P$  и  $B$ . Для того, чтобы получить высокую степень сжатия, группа должна быть достаточно большой. При воспроизведении изображений последовательность кадров может быть, например, такой:  $\{I, B, B, B, P, B, B, B, P\}$ . Взаимодействие кадров для приведенной последовательности показано на рисунке 8.25.

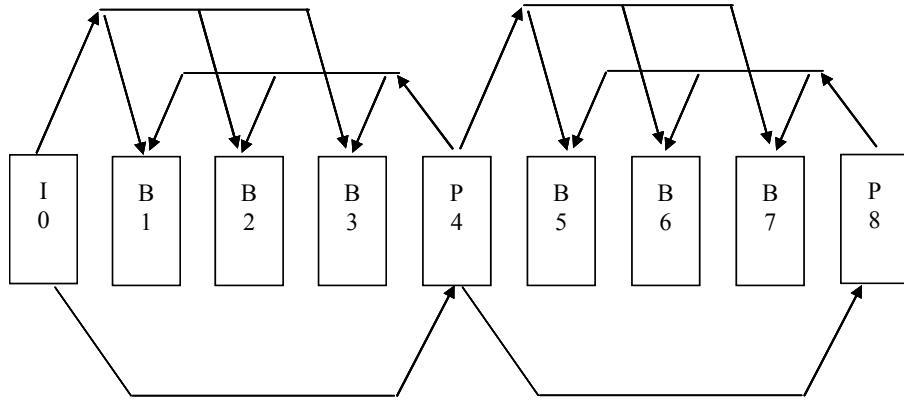


Рисунок 8.25 Схема формирования предсказания движения в группе в хронологической последовательности кадров.

Однако при кодировании и декодировании изображений порядок кадров различного типа в группе должен быть изменен на следующий:  $\{I(0), P(4), B(1), B(2), B(3), P(8), B(5), B(6), B(7)\}$ , поскольку кадры типа  $B$  создаются на основании кадров  $I$  и  $P$ , которые к этому времени уже должны иметься в кодере. Степень сжатия каждого из трех типов кадров различна: она меньше всего у кадров типа  $I$ , у кадров типа  $P$  она примерно в 3 раза больше, чем у кадров типа  $I$ , а у кадров типа  $B$  она примерно в 4,5 раза превышает степень сжатия кадров типа  $I$ .

Организация этих трех типов изображений в последовательность изменяется, в зависимости от приложения.

В стандарте предусмотрен еще один тип кадров –  $D$  (кодирование только  $DC$  коэффициентов,  $A(0,0)$ ). Эти кадры, обеспечивающие низкое качество изображения, используются только для быстрого поиска в режиме «быстрой прокрутки вперед».

Входное изображение видеопоследовательности должно быть оцифровано и представлено как яркость и два цветоразностных сигнала  $YCrCb$ . Далее выполняется предобработка и преобразование форматов для задания соответствующего разрешения. Для цветного телевидения широко используются два стандарта 525 строк разложения при кадровой частоте 29,97 Гц (*NTSC*) и 625 строк при кадровой частоте 25 Гц (*PAL*, *SECAM*). Размер кадра в стандартах *PAL*, *SECAM* составляет 720x576 элементов, а в стандарте *NTSC* – 640x480 элементов. Рекомендация МККР 601 (*CCIR 601*) (теперь *ITU BT-R 601*) определяет стандарты для цифрового кодирования цветных ТВ сигналов в компонентном режиме. В случае стандартов *PAL*, *SECAM*, в соответствии с рекомендацией для стандарта 4:2:2 тактовая частота для яркостного сигнала составляет 13,5 МГц, для цветоразностных

сигналов – 6,75 МГц. Число отсчетов яркостного компонента составляет 720 активных элементов в строке.

Для обеспечения эффективного кодирования при скорости данных от 1 до 1,5 Мб/с используют децимацию по строкам и столбцам, так называемый *SIF* (*source input format*) формат. Кодируемый кадр после соответствующей интерполяции имеет размер 360x288 для *Y* компонента и 180x144 для цветоразностных компонентов. Поскольку выделяется макроблок размером 16x16 элементов, чтобы получить целое число блоков в *SIF* формате игнорируются 4 первых и 4 последних элемента в каждой строке, в результате чего размер кодируемого кадра становится равным 352x288 для *Y* компонента и 176x144 – для цветоразностных компонентов. Изображения кодируются по полосам, состоящим из макроблоков, расположенных сверху вниз и слева направо. Макроблок соответствует макроблоку, рассмотренному в формате *JPEG* (рисунок 8.13), и состоит из 6 блоков.

*Оценка векторов движения.* Широкое распространение получил метод согласования блоков для оценки вектора движения. Метод состоит в следующем. Исходная последовательность кадров движущегося изображения разбивается на макроблоки размером 16x16 отсчетов, как это было сделано в стандарте *JPEG*. Выбор макроблока размером 16x16 элементов как единицы компенсации движения является результатом компромисса между повышением эффективности кодирования, обеспечиваемым компенсацией движения, и увеличением потока вследствие необходимости хранения информации о векторах движения. Текущий блок сравнивается со всеми блоками ссылочного (предыдущего, последующего или их комбинации) кадра в пределах некоторой области поиска. Критерием схожести является, например, величина СКО сигнала разности двух сравниваемых блоков. Тот блок, который меньше всего отличается от текущего, принимается за прогноз для целей предсказания, а смещение между прогнозом и текущим блоками определяет вектор движения. Остаточный блок формируется как разность текущего блока и блока прогноза. Остаточный блок кодируется и передается декодеру. Кроме того, декодер получает координаты вектора смещения текущего блока относительно прогноза. Декодер использует вектор движения для нахождения блока прогноза, декодирует остаточный блок и складывает его с прогнозом, реконструируя исходный блок. Центр области поиска на ссылочном кадре устанавливается в положение центра текущего блока.

Каждый макроблок может быть одного из введенных для кадров типа. Например, в *B*-кадрах могут быть макроблоки *I*, *B*, *P*. Вектор движения кодируется дифференциально относительно последнего закодированного вектора движения с использованием кодов переменной длины. Предсказанное значение сигнала вычитается из сигнала макроблока. Полученный сигнал ошибок разделяется на блоки (4 яркостных и 2 цветности).

*Сокращения пространственной избыточности* достигают ДКП блоков изображений для *I*-кадра и сигнала ошибок для блоков *B*, *P*-кадров. Дальнейшее сжатие достигается за счет квантования, зигзагообразного сканирования и кодирования длин серий для квантованных значений ДКП и энтропийного кодирования, как было рассмотрено для *JPEG*.

Если канал имеет фиксированную скорость, необходимо использовать буфер *FIFO* (первым вошел, - первым вышел) для согласования выхода кодера с каналом. Состояние этого буфера учитывается при управлении квантованием.

*Декодирование.* Этот процесс существенно проще, чем кодирование, так как не требует оценки векторов движения. При фиксированной скорости канала заполняется буфер *FIFO*. Декодер считывает и декодирует элементы данных потока в соответствии с определенным синтаксисом. Прежде всего, идентифицирует начало и тип изображения. Затем выполняется декодирование последовательности макроблоков. Тип макроблока и векторы движения, если они присутствуют, используются для реконструирования предсказания текущего макроблока по изображениям ссылок на предшествующие и последующие кадры, которые сохранены в декодере. Коэффициенты декодируются и деквантуются. Выполняется ОДКП каждого блока коэффициентов ДКП размером 8x8, результат складывается с сигналом предсказания. После декодирования всех макроблоков кадра, производится реконструкция кадра. *I* и *P* кадры заменяют прежние ссылочные кадры. Перед воспроизведением выполняется преобразование последовательности кадров в последовательность, соответствующую временной входной последовательности кадров.

### **8.10.2 Стандарт *MPEG-2***

Выпущен в 1995 г, предназначен для хранения и передачи изображения ТВ качества. Используется в компьютерном видео и в цифровом спутниковом телевидении. Основан на *MPEG-1*. По сравнению с *MPEG-1*, в аудио части добавлена поддержка многоканального звука. В видео части работает с изображениями полного цифрового формата в соответствии с МККР 601, позволяя осуществлять сжатие чересстрочного и прогрессивного видео. Обеспечивает более эффективное сжатие. В этом стандарте впервые были реализованы концепции профилей и уровней. Применяется во всем мире для широковещательной цифровой ТВ трансляции по кабельным сетям, спутниковым и наземным каналам.

### **8.10.3 Стандарт *MPEG-4***

Впервые был представлен в 1993 г. В 1999 году получил официальный статус стандарта *ISO/IEC* стандарт *MPEG-4 Visual* [14]. *MPEG-4* задумывался как способ передачи потоковых медиа - данных, в первую очередь видео, по каналам с низкой пропускной способностью. Стандарт

неожиданно завоевал популярность: применение более сложных алгоритмов компрессии позволило размещать полнометражные фильмы длительностью полтора - два часа с приемлемым качеством всего на одном компакт-диске. При одном и том же коэффициенте сжатия качество изображения фильма в *MPEG-4* может быть сравнимо или даже лучше, чем в случае применения *MPEG-2*. Однако применение новых алгоритмов сжатия повлекло за собой и существенное увеличение требований к вычислительным ресурсам, необходимым для качественного декодирования изображения из этого формата.

Стандарт *MPEG-4* поддерживает не только кадры и полукадры, как предыдущие форматы *MPEG*, а еще и оперирует с видеообъектами (областями произвольной формы с движущимися предметами), неподвижными текстурами, гибридными синтетическими и натуральными видеообъектами (ВО), сеточными 2D и 3D объектами. Фактически данный формат задает правила организации среды, причем среды объектно ориентированной. *MPEG-4* определяет видеообъект как «пластичную сущность, доступную пользователю, к которой он может обращаться, наблюдать и манипулировать (вырезать и вставлять)». Каждый объект кодируется отдельно. Простой профиль объединяет инструменты для обращения с прямоугольными ВО.

Алгоритм компрессии видео в *MPEG-4* работает по той же схеме, что и в предыдущих стандартах. При кодировании исходного изображения кодек ищет и сохраняет ключевые кадры, на которых происходит смена сюжета. А вместо сохранения промежуточных кадров прогнозирует и сохраняет лишь информацию об изменениях в текущем кадре по отношению к предыдущему. Полученная таким образом информация сжимается по алгоритмам компрессии, аналогичным тем, что применяются в *MPEG-1*, *MPEG-2*. Для кодирования текстур применяется не ДКП, а ДВП.

Естественно, алгоритмы поиска и обработки подобных объектов требуют гораздо больше вычислительных ресурсов, нежели в случае *MPEG-1* или *MPEG-2*. Но с учетом быстродействия современных компьютеров последнее обстоятельство нельзя рассматривать как крупное препятствие на пути широкого распространения формата *MPEG-4*. Для сравнения - во времена выхода *MPEG-2*, в 1995 году, частота процессора новых ПК составляла около 100 МГц. Сегодня эта цифра возросла на порядок.

## **9 АНАЛИЗ ТЕКСТУРНЫХ ИЗОБРАЖЕНИЙ**

### **9.1 Описание текстур**

Цвет и текстура являются важными характеристиками изображения. Проблема анализа цветных текстур включает в себя такие аспекты, как описание цветных текстур, их классификация, то есть формирование кластеров (при этом под кластером понимают обычно группу объектов, образующих в пространстве описания компактную в некотором смысле

область), и сегментация, то есть разбиение изображения на области, которые являются однородными относительно одной или нескольких характеристик, или принадлежат некоторому кластеру. Харалик [62] выделяет два подхода к описанию текстур: статистический и структурный. Сегодня можно сказать, что развит и получил распространение третий подход к описанию текстур - фрактальный [63]. Остановимся более подробно на этих подходах.

### 9.1.1 Статистический подход к описанию текстур

В первых работах по анализу текстур использовались такие признаки, как средняя оптическая плотность снимка и средняя пространственная частота [64].

Авторы работы [65] в качестве признаков при автоматической классификации аэрофотоизображений используют среднее значение оптической плотности и ее дисперсию. В работе [66] для автоматического анализа формы облачности применяются собственные векторы ковариационной матрицы яркостей элементов скользящего окна (прямоугольного фрагмента изображения) размером  $6 \times 6$  элементов. Распределение яркостей пар соседних элементов и двумерная автокорреляционная функция в качестве описания изображений облачных полей исследуются в работе [67]. В более поздних работах Розенфельд описывает текстуру плотностью перепадов яркости, то есть количеством перепадов яркости на единицу площади [68].

Харалик [69] для различия текстур использует ковариационные матрицы, характеризующие статистики второго порядка и описывающие пространственные связи пар яркостей элементов в цифровом изображении текстуры. Этот подход основан на предположении Юлеша [70] о том, что человеческий глаз использует статистики не выше второго порядка для распознавания текстур. В 1978-1980 годах Юлеш и Гагалович [71, 72] посредством моделирования построили контрпримеры, состоящие однако из искусственных текстур, не встречающихся в природе. В работе [73] элементы ковариационной матрицы  $\mathbf{B}(i,j)$  являются счетчиками числа переходов яркости  $i$  в яркость  $j$  для пар элементов изображения, отстоящих на заданном расстоянии  $d$ . Эта характеристика не инвариантна к повороту, поэтому необходимо считать число переходов для каждого анализируемого направления.

Чаще используется не вычисление ковариационных матриц как таковых, а оценка основанных на них признаков, характеризующих текстуру. В работе [69] Харалик предлагает использовать 14 признаков текстур, основанных на статистике и теории информации. В работе [73] на основе анализа этих признаков, авторами предложено использовать только 6 из них, представляющих оценку изображения по его информационным и статистическим характеристикам, таким как характеристика однородности, контраст, энтропия, количество информации.

Для анализа текстур авторами работ [74 – 76] на основании теории статистического кодирования и данных физиологии зрения предлагается использовать распределение отрезков контуров и интервалов между контурными элементами. Эти характеристики можно измерять за время одного ТВ кадра посредством простых схем нелинейных фильтров, регистров сдвига и фильтров с обратной связью. При построении ТВ автомата, работающего в реальном времени [77], для классификации неоднородных текстур использовалась оценка максиминных расстояний между контурами.

Рассмотрим работу этого автомата подробнее. При создании автоматической системы для анализа неоднородных текстурных изображений выполняется следующая процедура. Анализируемое изображение разбивается на фрагменты. Для каждого фрагмента оцениваются характеристики и строятся гистограммы их распределения. Отдельные информативные отсчеты полученных распределений выбираются в качестве признаков для распознавания. Параметры классификации такие, как размер фрагмента, количество уровней квантования гистограммы распределения признаков и количество информативных признаков, остаются неопределенными.

Для нахождения параметров используется метод «стабильности» моды. Для некоторого признака строится зависимость количества мод гистограммы от определяемого параметра. Наибольшая область этого параметра, в которой количество мод остается постоянным, называется зоной «стабильности» моды, а середину этой области используют в качестве значения неизвестного параметра. Если распределение какого-то признака не имеет мод или имеет только 1 моду, то признак считается неинформативным и отбрасывается.

Алгоритм предполагает выполнение следующих операций. Прежде всего, формируется контурное изображение путем сравнения лапласиана с

$$\text{порогом } g(x,y) = \begin{cases} 1, & \text{если } G \geq T \\ 0, & \text{иначе} \end{cases},$$

где  $G = 4f(x,y) - f(x-1,y) - f(x+1,y) - f(x,y-1) - f(x,y+1)$ ,  $T$  – порог, пропорциональный среднему значению яркости элементов в окрестности.  
 $T = \alpha(f(x,y) + f(x-1,y) + f(x+1,y) + f(x,y-1) + f(x,y+1))/5$ ,

где  $\alpha$  – коэффициент пропорциональности.

Подсчитывается число контурных элементов во фрагменте, получается достаточно эффективная характеристика для классификации некоторых изображений.

Оценивается распределение интервалов между контурами (внутри каждой строки при построчном сканировании изображения) по формуле:

$$g^*(x,y) = \begin{cases} g^*(x,y-1) + 1, & \text{если } g(x,y) = 0 \\ 0, & \text{если } g(x,y) = 1 \end{cases}$$

и

$$c_k(i,j) = \begin{cases} 0, & \text{если } g(x,y)=0 \\ g^*(x,y-1)+1, & \text{если } g(x,y)=1 \end{cases}$$

где \* обозначены уже обработанные отсчеты.

Таким образом, в матрице  $\|c_k(x,y)\|$  будут записаны расстояния между контурными элементами вдоль строки.

Производится оценка расстояний между элементами контурного (бинарного) изображения. Пусть  $P$  и  $Q$  – две точки бинарного изображения, а  $d(P,Q)$  – такое наименьшее положительное целое, что существует последовательность отдельных точек  $P = \{P_1, P_2, \dots, P_n\} = Q$ , причем  $P_k$  является соседом  $P_{k-1}$  ( $1 \leq k \leq n$ ).  $d(P,Q)$  называется расстоянием от точки  $P$  до точки  $Q$ , то есть представляет собой минимальное количество «шагов», за которое можно попасть из точки  $P$  в точку  $Q$  по соседним точкам. Очевидно, это расстояние зависит от определения соседства. В этом методе соседство определяется по окрестности 1-го порядка (в соответствии с рисунком 6.1), то есть по отношению к элементу с координатами  $(x,y)$ , соседними считаются 4 элемента с координатами:  $(x-1,y)$ ,  $(x+1,y)$ ,  $(x,y-1)$ ,  $(x,y+1)$ .

Наконец, строится матрица максиминных расстояний. Для этого матрица контурного изображения  $\|g(x,y)\|$  обрабатывается по алгоритму слева направо и сверху вниз, и строятся матрицы:

$$b^1(x,y) = g^*(x,y) = \begin{cases} \min(g^*(x-1,y); g^*(x,y-1)) + 1, & \text{если } g(x,y)=0 \\ 0, & \text{если } g(x,y)=1 \end{cases}$$

и матрица:

$$c^1(x,y) = \begin{cases} 0, & \text{если } g(x,y)=0 \\ \min(g^*(x-1,y), g^*(x,y-1)) + 1, & \text{если } g(x,y)=1 \end{cases}$$

Затем выполняется сканирование матрицы  $\|g(x,y)\|$  справа налево и сверху вниз, формируя матрицы

$$b^2(x,y) = g^*(x,y) = \begin{cases} 0, & \text{если } g(x,y)=1 \\ \max[\min(g^*(x-1,y); g^*(x,y+1)) + 1, b^1(x,y)], & \text{если } g(x,y)=0 \end{cases}$$

$$c^2(x,y) = \begin{cases} 0, & \text{если } g(x,y)=0 \\ \max[\min(g^*(x-1,y), g^*(x,y+1)) + 1, c^1(x,y)], & \text{если } g(x,y)=1 \end{cases}$$

Выполняется сканирование матрицы  $\|g(x,y)\|$  слева направо и снизу вверх, затем справа налево и снизу вверх. В матрице  $\|c(x,y)\|$  получаются кратчайшие расстояния между элементами контурного изображения.

Для классификации выбран алгоритм самообучения, работающий по критерию минимальной ошибки классификации [78]. Исследования показали, что такие методы эффективно работают в задачах анализа

изображений, когда не предъявляется строгих требований к форме границы между статистически однородными областями. Использование оценки максиминных расстояний эффективнее оценки числа контурных элементов и оценки распределения интервалов между контурами.

Для анализа анизотропных текстур предложены алгоритмы, учитывающие зависимости распределения интервалов между контурами и отрезков контуров вдоль направлений сканирования текстуры [49, 79].

Работы [62, 64-81] построены на статистическом подходе к описанию текстур. При статистическом подходе к описанию текстур изображения трактуются как реализации некоторого случайного процесса. В работе [82] введено понятие типового изображения, порожденного эргодическим источником. *Под текстурами понимают такие типы изображений, класс которых обладает свойством эргодичности, то есть любое изображение данной текстуры является типовым и полностью характеризует все другие изображения, относящиеся к этому классу.*

### 9.1.2 Структурный подход к описанию текстур

В работе [62] Харалик выделяет другой подход - структурный, или синтаксический. Этот подход развился после возникновения теории формальных языков в середине пятидесятых годов прошлого века. Он основан на том, что текстура составлена из регулярно или почти регулярно повторяющихся непроизводных элементов. Поэтому описание текстуры, как считают сторонники такого подхода, должно состоять из описания непроизводных элементов и правил их размещения [83]. Синтаксическому распознаванию образов посвящена и глава в книге [84]. Выделяют три направления в лингвистическом распознавании:

- Синтаксически ориентированное распознавание, когда устанавливается синтаксис грамматики и разбор грамматики осуществляется в двух направлениях - снизу вверх и сверху вниз.
- Распознавание образов, представленных графами. В рамках разработки этого направления был создан язык описания изображений *PDL*, предложенный Шоу [85]. Непроизводным элементом в *PDL* служит любая *n*-мерная структура с двумя выделенными точками - хвостовой и головной, при этом любая структура рассматривается как ориентированный отрезок прямой, заданный хвостовой и головной точками. Устанавливается допуск на соединение структур примыканием только в головной или хвостовой точках. Таким образом формируются графы структур, а для их обработки используют грамматики цепочек. Этот же подход использован Ледли при анализе хромосом [86].
- Распознавание древовидных структур, при котором модифицируется определение грамматики, дополнительно определяется функция ранжирования. В связи с этим соответственно меняются и правила подстановки.

Стохастические грамматики и языки вводятся для учета случайного характера текстур. Такой подход представлен в работе [87]. Однако авторы указывают, что их описание чувствительно к шумовым искажениям и не дает хороших результатов при анализе текстур с неявно выраженнымими периодическими свойствами, к которым принадлежит большинство реальных изображений.

Эрих и Фойт [88] представляют яркости элементов вдоль строки также в виде дерева соотношений, которое описывает рекурсивное разбиение функции яркости в точке наименьшего из относительных минимумов. Точки относительных минимумов вновь построенных ветвей дерева и значения функции яркости слева и справа от минимума используются для дальнейшего разбиения на следующем шаге рекурсии.

Основные проблемы в реализации стохастических грамматик связаны с решением задачи вывода грамматик и получения вероятностей правил подстановки с помощью обучения, и, до настоящего времени, они имеют ограниченную область применения.

### **9.1.3   Фрактальный подход к описанию текстур**

Описание широкого класса процессов и явлений, таких как процессы ограниченной диффузной агрегации, образование вязких пальцев в пористых средах, турбулентность, процессы диффузии, называемые протеканием, или перколяционными процессами, а также описание объектов природы, таких как облака, земная поверхность и многие другие, в терминах фрактальной геометрии определило новое направление в исследованиях - анализ фракталов. При таком подходе авторы не называют объект текстурой, а называют его фракталом. Поскольку анализ фракталов, по существу, дает характеристику текстуры, и на сегодняшний день нет устоявшегося определения ни текстуры, ни фрактала, то представляется возможным говорить о фрактальных свойствах текстур. Бенуа Б. Мандельброт в книге [89], первый назвавший объекты фракталами, изложил как элементарные понятия фрактальной геометрии, так и новые идеи в этой области, издав общепризнанный стандартный справочник по фракталам. В евклидовой геометрии введено понятие топологической размерности. Так, размерность кривой - 1, размерность плоскости - 2, поверхности - 3. Таким образом, топологическая размерность имеет целочисленное значение. Во фрактальной геометрии размерность кривой может иметь значение в интервале [1,2] в зависимости от сложности кривой, размерность поверхности заключена в интервале [2,3]. Концепция дробной размерности была впервые сформулирована математиками Хаусдорфом и Безиковичем. Мандельброт назвал такую размерность фрактальной и ввел такое определение фрактала [89]: “*Фракталом называется множество, размерность Хаусдорфа-Безиковича которого строго больше его топологической размерности*”.

Распространение фрактального описания объясняется тем, что большинство пространственных систем в природе является нерегулярным

и фрагментарным, форма этих систем плохо поддается описанию аппаратом евклидовой геометрии. Например, береговая линия острова не прямая и не круглая, и никакая другая классическая кривая не может служить для описания и объяснения ее формы без чрезмерной искусственности и усложнения.

Определение фрактала претерпевает изменения. В 1987 году в частном сообщении [63] Мандельброт сузил свое определение: “*Фракталом называется структура, состоящая из частей, которые в каком-то смысле подобны целому*”.

Существенным отличительным признаком второго определения является то, что при определении фрактала используется свойство самоподобия фрактала. Многие кривые и поверхности статистически самоподобны, то есть каждая часть может считаться изображением целого в уменьшенном виде.

Размерность фрактала  $D$  определяется как

$$D = \frac{\log N}{\log(1/r)}, \quad (9.1)$$

где  $1/r$  есть отношение подобия,  $N$  - число шагов, необходимое для того, чтобы покрыть кривую.

На рисунке 9.1 проиллюстрировано соотношение между числом шагов и отношением подобия на примере четырех ломаных прямых.

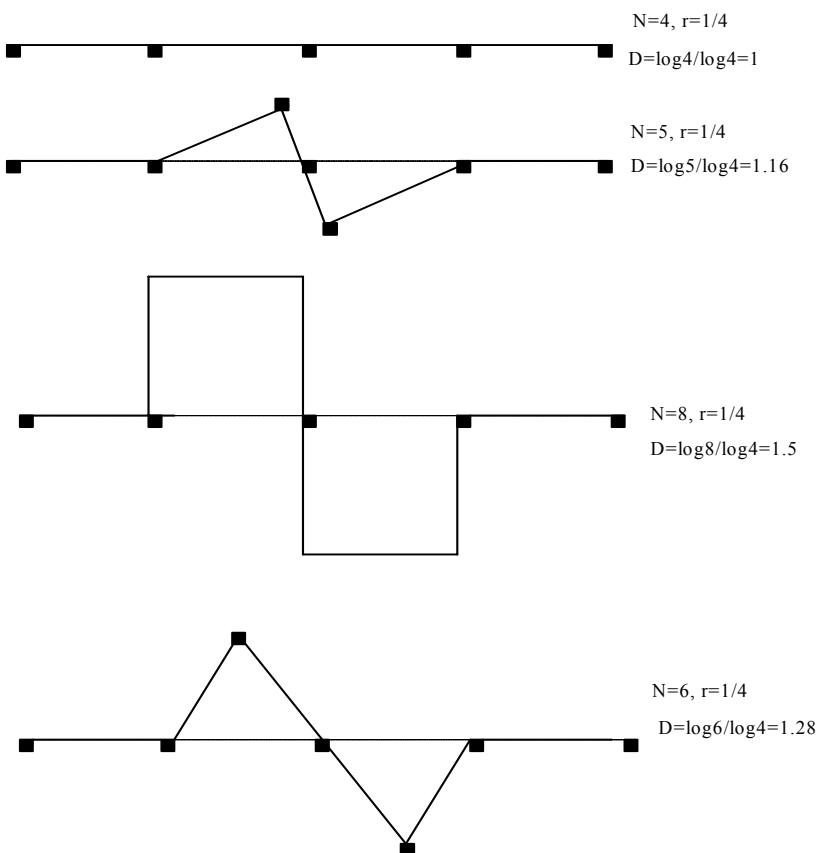


Рисунок 9.1 Определение размера фрактала на примере четырех ломанных линий.

Практически размер фрактала для кривой оценивается путем измерения длин кривой при различных размерах шага. Размерность фрактала  $D$  может быть оценена с помощью следующего уравнения регрессии:

$$\log L = C + B \log G, \quad (9.2)$$

$$D = 1 - B, \quad (9.3)$$

где  $L$  – длина кривой,  $B$  – наклон регрессии,  $G$  – величина шага,  $C$  – константа.

Рассмотрим подробнее реализацию фрактального подхода к анализу облаков. В основу этого метода положено выведенное Мандельбротом соотношение между периметром и площадью объекта [90]. Для окружностей, квадратов, равносторонних треугольников и других многоугольников отношение периметра к квадратному корню из ограничивающей ими площади не зависит от размера фигуры и является величиной постоянной для данного семейства. Аналогично для семейства подобных островов отношение длины нефрактальной береговой линии любого острова к квадратному корню из его площади не зависит от размера площади. Однако, если береговая линия фрактальна, то ее длина  $L(\delta)$  зависит от длины эталона  $\delta$  и стремится к бесконечности при стремлении эталона к нулю.

При этом площадь острова  $A(\delta)$ , определяемая количеством располагаемых на ней квадратов  $\delta^2$ , остается конечной. Таким образом, отношение периметра к квадратному корню из площади расходится. Мандельброт получил для случая фрактальной береговой линии следующее соотношение между периметром и площадью:

$$L(\delta) = C \delta^{(1-D)} [A(\delta)]^{D/2}. \quad (9.4)$$

Это соотношение выполняется для любого эталона длины  $\delta$ , достаточно малого, чтобы удовлетворительно обмерять самый малый из островов.

Соотношение (9.4) применимо при исследовании геометрии облаков и зон дождя, размеры которых заключены в широких пределах от 1 до  $1,2 \times 10^6$  км<sup>2</sup>. Выяснилось, что периметр облака связан с его площадью соотношением (9.4) с фрактальным размером  $D = 1,35 \pm 0,05$  [90]. При этом эти оценки оказались справедливы как для кучевых, так и для перистых облаков. В работе Ф. Риса и А. Вальдфогеля [91], посвященной анализу фрактальной размерности облаков с мощными конвективными токами, было установлено соотношение между периметром и площадью для последовательности моментов времени (с интервалом в 1 минуту) в плоскости сечения для постоянного коэффициента отражения. Основные выводы могут быть следующими: для облаков, периметр которых больше 8 км, размер фрактала примерно совпадает с размером менее мощных облаков и составляет  $1,36 \pm 0,1$ ; для облаков периметра от 3 км до 8 км

$D=1,0 \pm 0,1$  и, наконец, облака с периметром менее 3 км не являются фракталами.

Исследования фрактальных поверхностей от молекулярных поверхностей белков, обшивки супертанкеров, поверхностей суставов до взлетных полос аэродромов проводились различными авторами. При этом использовались различные методы оценки размера фрактала, основные из которых более подробно будут рассмотрены в разделе 9.6.

Таким образом, важной задачей анализа текстур является выделение признаков. Можно отметить три основных подхода к описанию текстур, на основании которых формируются признаки текстур. Перспективным представляется использование набора признаков, комбинируемых из признаков, выделяемых при различных подходах к описанию текстур.

## 9.2 Оценка наличия текстуры в изображении

При применении методов текстурного анализа мы предполагаем, что входное изображение имеет текстурные характеристики. Поскольку в настоящее время нет единого определения текстуры, то каждый метод текстурного анализа предполагает описание текстуры некоторым набором признаков, извлекаемых из изображения. Поэтому при проведении текстурного анализа, прежде всего, необходимо выделить на изображении текстурные области. В [92] авторы приводят следующую схему оценки текстуры в изображении:

Первые три класса объектов ( $A, B, C$ ), по мнению авторов, текстурами не являются, к ним в соответствии с рисунком 9.2 относятся: области постоянной яркости, области белого нормального шума и объекты, полностью описываемые их формой. *Белым шумом* называется стационарный случайный процесс  $n(t)$ , у которого спектральная плотность мощности не зависит от частоты и имеет постоянное значение, равное дисперсии значений  $n(t)$ . Другими словами, все спектральные составляющие белого шума имеют одинаковую мощность. По существу, это идеализированный случайный процесс с бесконечной энергией. Считают, что спектральная плотность мощности шумов слабо изменяется в диапазоне спектра сигнала, который существенно уже спектра шума. Понятие "белый шум" определяет только равномерный энергетический спектр шума, а законы распределения амплитуды могут быть любыми. Амплитуда отсчетов белого нормального шума распределена по нормальному закону.

Последние три класса изображений ( $D, E, F$ ) содержат текстуры и, следовательно, подлежат текстурному анализу. Предварительная грубая сегментация изображения на текстурные и нетекстурные области производится по методу оценки зернистости текстуры [79]. Зернистость текстуры оценивается плотностью распределения локальных экстремумов яркости изображения по столбцам и по строкам. Размер фрагмента, в котором производится оценка зернистости, зависит от ожидаемого диапазона зернистости текстур на изображении.

Чем больше зернистость, тем больший размер фрагмента должен задаваться при анализе. В работе [92] авторы предлагают размер квадратного фрагмента оценивать по формуле:

$$s = \sqrt{\frac{2n}{p_{\min} + p_{\max}}}, \quad (9.5)$$

где:  $s$  - размер стороны фрагмента;  $p_{\min}, p_{\max}$  - минимальная и максимальная плотности локальных экстремумов изображений, которые могут быть отнесены к классу текстур, в [92] авторы устанавливают  $p_{\min}=0,04$ ,  $p_{\max}=0,16$ ;  $n$  - некоторое, задаваемое пользователем, число локальных экстремумов во фрагменте, например, 32.

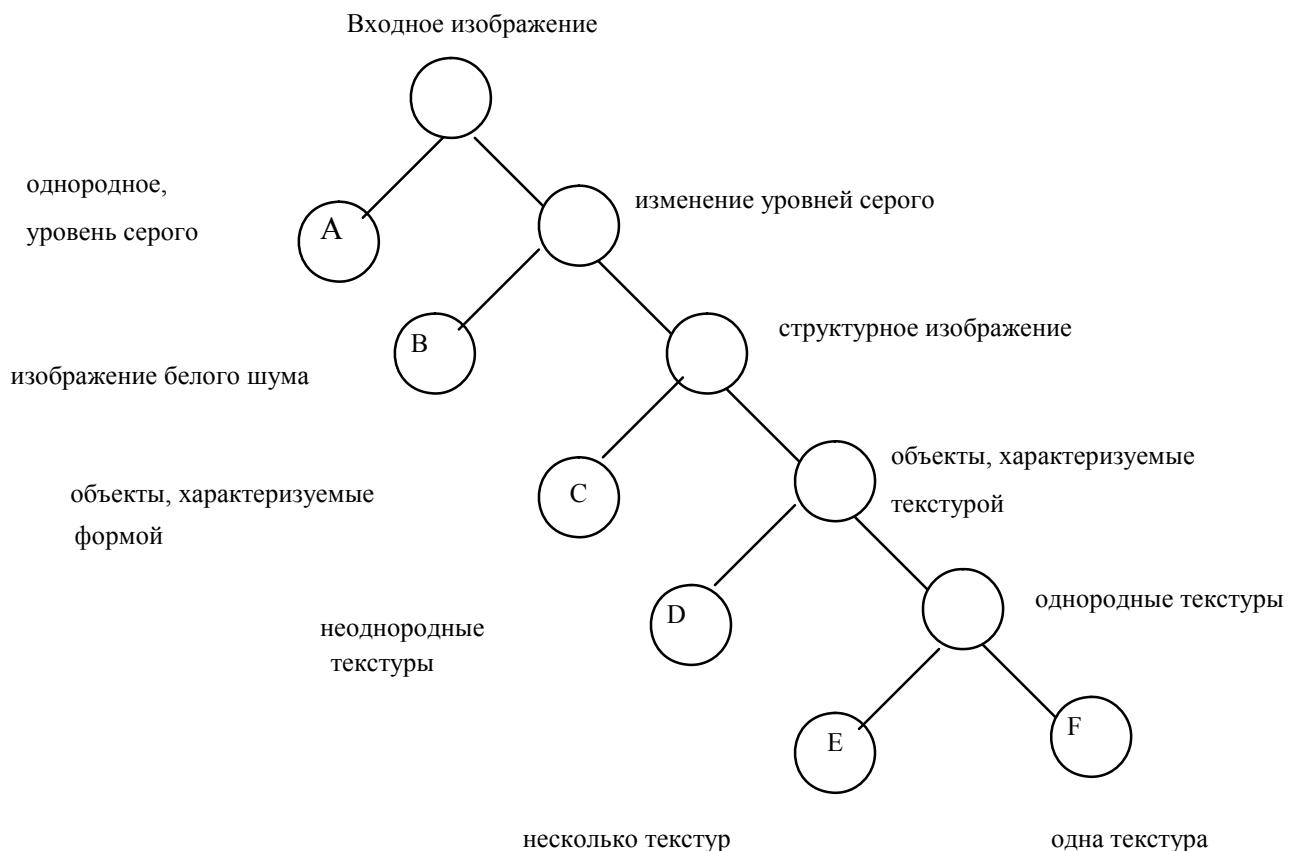


Рисунок 9.2 Граф выделения на изображении текстурных областей.

Поскольку быстродействие этого алгоритма достаточно велико, авторам представляется целесообразным использование его как первого шага при анализе текстур с тем, чтобы выделить те области на изображении, которые подлежат более подробному текстурному анализу.

### 9.3 Методы сегментации цветных текстур в цветовых координатных пространствах

Мы говорили о том, что цвет является важной характеристикой объектов. Поскольку форма и распределение кластеров при сегментации цветных текстур зависят от цветового координатного пространства, авторы

проводят анализ цветных текстур в различных цветовых координатных пространствах. Выбор цветового координатного пространства также важен, как и собственно метод сегментации. Вследствие этого в аналитическом обзоре статей будем указывать цветовое координатное пространство, выбранное авторами.

В работе [93] авторы используют для сегментации цветных изображений выбор порогов квантования по многомерным гистограммам распределения компонентов сигнала. Анализ проводится по девяти компонентам, полученным для трех цветовых координатных пространств  $RGB$ ,  $YIQ$  и  $HSI$ . Исследуется изображение комнаты с обстановкой для отдыха. При этом авторами, прежде всего, выполняется сегментация текстурной части изображения. А затем для областей изображения, в которых отсутствует текстура, проводится сегментация по цвету на основании многомерного порогового ограничения. В противном случае при таком подходе области с однородной текстурой распались бы на множество отдельных областей. Модельный эксперимент показал, что метод сегментации Олэндера оказывается весьма эффективным. При этом более информативным авторами признан компонент тона.

В работе [94] предложен метод кластеризации для сегментации цветных изображений в пространстве  $L^*a^*b^*$  МКО 1976. При таком описании элементом кластера является круговой цилиндрический сегмент, выделенный на рисунке 3.4. Оценка кластеров производится без предварительных оценок их формы и без предположений о законах распределения кластеров. Определение кластеров производится путем оценки одномерных распределений в хорошо определенном пространстве решений для вычисленных наборов постоянных значений яркости и насыщенности. Процедура определения кластеров состоит в следующем. По изображению строятся одномерные гистограммы распределений цветовых координат. Из этих распределений выбирается то, глубина мод которого больше. По этому распределению выбираются два кластера. Затем уже для каждого из оцененных кластеров строятся одномерные гистограммы распределений по каждой из оставшихся координат. Использование одномерных гистограмм для определения трехмерных кластеров уменьшает количество необходимых вычислений. Для разделения кластеров используется метод линейного дискриминанта Фишера. Такая процедура обеспечивает успешное разделение по одномерному распределению, что выгодно отличает этот метод от приведенного в работе [93], однако его использование затруднено при анализе сложных текстур.

В работе [95] предлагается способ классификации цветных изображений в реальном времени. Авторы используют модель нелинейного преобразования цветового координатного пространства  $RGB$  в пространство  $HSI$  (3.9)-(3.11). На модуль классификации подается три  $RGB$  сигнала от цветной ПЗС камеры, преобразуемые в 6-ти разрядные коды. 18-ти разрядная  $LUT$  осуществляет преобразование  $RGB$  координат в

координаты  $HSI$ , а эти компоненты используются для адресации второй  $LUT$  типа ОЗУ  $256K \times 8$ , которая запрограммирована как классификатор реального времени. Обучение осуществляется для нескольких типичных культур. Каждый байт ОЗУ дает код идентификации класса и меру значимости соответствия, пропорциональную расстоянию от исследуемого вектора  $HSI$  до центра кластера. Во время автоматической классификации по цвету просмотровая таблица устанавливается в режим чтения. Результат классификации для каждого элемента изображения передается на процессор управления роботом вместе с данными формы, полученными с контурного модуля. Успех работы такого классификатора во многом зависит от получения хороших слаженных кластеров во время обучения, что ограничивает возможности применения таких систем.

В работе [96] авторы используют рекомендации МКО 1973 по применению равноконтрастного цветового пространства  $L^* u^* v^*$ . При сегментации цветного изображения авторы используют метод обучения без учителя, доказывая, что он сходится к методу кластеризации по критерию минимума суммы квадратов ошибки. При таком подходе изменяется форма кластера. Если при сегментации по методу квантования многомерных гистограмм распределения кластеры имеют прямоугольную форму, то при кластеризации по методу минимума квадратов ошибок, когда все пространство разбивается на области, принадлежащие ближайшему весовому вектору, мы имеем более гибкую форму кластера вида мозаик Вороного (рисунок 9.3). Векторное квантование предполагает разбиение множества входных векторов на некоторое количество весовых векторов  $\mathbf{W}_i$ . В [96] доказывается, что задача кластеризации, то есть задача разбиения входного множества векторов на отдельные области, характеризуемые некоторым центром кластера, по критерию минимума суммы квадратов ошибок сходится к задаче векторного квантования.

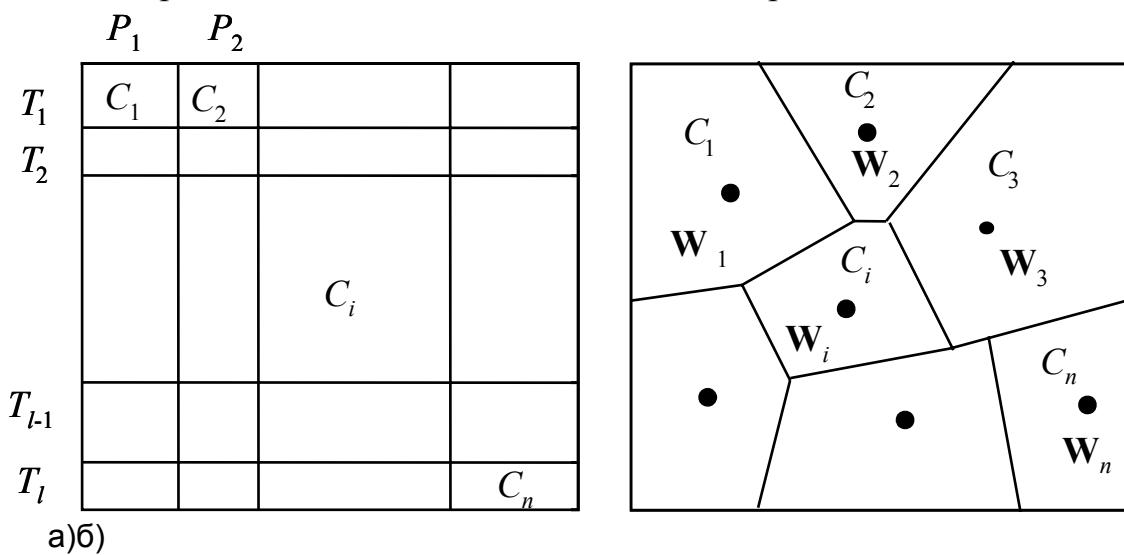


Рисунок 9.3 Форма кластеров при кластеризации двумерных данных а) по методу квантования многомерных гистограмм, б) по методу векторного квантования.

Пусть  $\mathbf{X}_j$  ( $j=1,\dots,m$ ) - конечное множество входных векторов ( $\mathbf{X}$ ),  $C_i$  ( $i=1,\dots,n$ ) - конечное множество кластеров ( $\mathbf{C}$ ),  $\mathbf{W}_i$  ( $i=1,\dots,n$ ) - конечное множество весовых векторов ( $\mathbf{W}$ ). Определим функцию разбиения  $C_f(W)$  таким образом:

$$\mathbf{X}_j \in C_i, \text{ если для любых } k \neq i \quad \|\mathbf{X}_j - \mathbf{W}_i\|^2 < \|\mathbf{X}_j - \mathbf{W}_k\|^2. \quad (9.6)$$

Определим функцию весового вектора  $\mathbf{W}_f(\mathbf{C})$  как центр масс кластера  $C_i$

$$\mathbf{M}_i \equiv \frac{\sum_{\mathbf{X}_j \in C_i} p_j \mathbf{X}_j}{\sum_{\mathbf{X}_j \in C_i} p_j}, \quad (9.7)$$

$$\mathbf{W}_f(\mathbf{C}) = \{\mathbf{M}_i / i=1,\dots,n\}, \quad (9.8)$$

где  $p_j$  - вероятность входного вектора  $X_j$ ,  $p_j = 1/m$ , при условии равновероятных входных векторов;  $n$  - число кластеров.

Сумма квадратов отклонений от центра масс множества входных векторов может быть представлена в виде:

$$S_o = \sum_{i=1}^n \sum_{j=1}^m p_j \|\mathbf{X}_j - \mathbf{M}_0\|^2. \quad (9.9)$$

Сумма квадратов межкластерных отклонений задается следующим уравнением:

$$S_M = \sum_{i=1}^n \sum_{\mathbf{X}_j \in C_i} p_j \|\mathbf{M}_i - \mathbf{M}_0\|^2, \quad (9.10)$$

где  $\mathbf{M}_0$  - центр масс множества входных векторов:

$$\mathbf{M}_0 = \sum_{j=1}^m p_j \mathbf{X}_j / \sum_{j=1}^m p_j. \quad (9.11)$$

При использовании критерия минимума суммы квадратов ошибок показатель качества для кластеризации имеет вид:

$$S_W(\mathbf{C}) = \sum_{i=1}^n \sum_{\mathbf{X}_j \in C_i} p_j \|\mathbf{X}_j - \mathbf{M}_i\|^2. \quad (9.12)$$

$$S_o = S_M + S_W. \quad (9.13)$$

Так как  $S_o$  не зависит от кластеризации, то минимизация суммы квадратов ошибок  $S_W$  приводит к достижению максимума суммы квадратов межкластерных отклонений, обеспечивая тем самым наилучшую разделимость кластеров. Из этой посылки исходят при использовании критерия минимума суммы квадратов ошибок.

Показатель качества для векторного квантования определяется следующим уравнением:

$$E(\mathbf{W}) = \sum_{i=1}^n \sum_{X_j \in C_i} p_j \|\mathbf{X}_j - \mathbf{W}_i\|^2. \quad (9.14)$$

В [96] доказано, что если разбиение на кластеры проводить по алгоритму мозаик Вороного таким образом, что векторы центров масс кластеров совпадают с векторами квантования  $\mathbf{W}_i = \mathbf{m}_i$ , то задача векторного квантования сходится к задаче кластеризации по критерию минимума суммы квадратов ошибок.

Алгоритм кластеризации выполняется следующим образом.

Задаются параметры кластеризации: размер кластера  $\theta$  и количество итераций  $K$ .

1. Инициализация. Один весовой вектор устанавливается равным  $\mathbf{m}_0$ , соответствующий ему счетчик количества векторов, вошедших в кластер,  $N_1$ , обнуляется. Счетчик итераций  $K$  обнуляется.

2. Состязательное обучение.

2.1. Из множества  $\mathbf{X}$  выбирается случайным образом входной вектор  $\mathbf{X}_j$ . Определяется вектор квантования  $\mathbf{W}_W = \mathbf{W}_i$ , имеющий  $\min$  квадрат евклидова расстояния  $\|\mathbf{X}_j - \mathbf{W}_i\|^2$ , если таких векторов оказывается несколько, то случайным образом выбирается один из них.

2.2. Весовой вектор  $\mathbf{W}_W$  заменяется вектором  $\Delta\mathbf{W}_W = \gamma(\mathbf{X}_j - \mathbf{W}_W)$ .

Счетчик количества векторов, вошедших в кластер  $\mathbf{W}_W$ ,  $N_w$  увеличивается на 1.

2.3. Если счетчик количества векторов, вошедших в кластер,  $N$  равен  $\theta$  и количество весовых векторов меньше  $n$ , то генерируется новый весовой вектор, равный  $\mathbf{W}_W$ , и счетчики векторов, вошедших в кластер  $\mathbf{W}_W$  и кластер  $\Delta\mathbf{W}_W$ , обнуляются.

2.4.  $N_i$  увеличивается на 1. Если  $N_i = I$ , то процесс обучения останавливается, в противном случае процесс обучения продолжается с пункта 2.1.

3. Определение кластеров по формуле (9.6).

Увеличение параметра  $\theta$  обеспечивает большую точность кластеризации, но требует соответственно большего допустимого числа итераций  $I$ . Авторы показывают, что  $I \geq (2n-3)\theta$ , и рекомендуют задавать  $I = (2n-3)\theta(n+7)$ ,  $\theta = 400\sqrt{n}$ . Коэффициент  $\gamma \in [0,1]$  и выбирается авторами равным 0,015. Авторы подчеркивают важное значение интенсивности при сегментации цветных изображений, хотя оценка тона представляется им более информативной характеристикой, чем интенсивность.

Таким образом, можно сделать следующие выводы по анализу цвета текстур:

- При наличии текстурных и нетекстурных областей в изображении необходимо выделить текстурные области и провести сегментацию нетекстурных областей.
- Эффективность выбора признаков зависит от выбора цветового координатного пространства.
- Для многих применений возможно проводить анализ текстур только по их цветовым характеристикам.

## 9.4 Синтез цветных пространственных текстур

Задача синтеза текстур возникает при формировании образцов цветных текстур для проверки эффективности формирования системы признаков и алгоритмов кластеризации при выполнении сегментации текстур. Единые методы синтеза текстур позволяют выполнить сопоставление методов анализа цветных текстур, разработанных в разных учреждениях.

Выбор признаков при анализе текстур определяет во многом успех кластеризации при разделении изображения на отдельные области. Понятно, что выбор признаков цветных текстур опирается на то описание текстур, на основании которого производится их классификация. Как уже обсуждалось в разделе 9.1, цветная текстура описывается выбранным цветовым координатным пространством, а яркость ее описывается на основании подходов, структурного, статистического или фрактального. Поскольку предпочтительным представляется использование статистического и фрактального подходов к описанию текстур, то и синтез текстур должен базироваться на этих описаниях.

### 9.4.1 Статистический метод синтеза цветных текстур

При статистическом описании текстур для синтеза яркостного компонента текстур будем использовать статистическую модель синтеза монохромных текстур, предполагая, что наиболее строгими моделями являются стохастические коррелированные поля [97]. Прэтт использовал изображения двумерных полей для изучения различия текстур зрительным анализатором человека. При таком подходе текстура генерируется как двумерный массив случайных чисел с заданной совместной плотностью вероятности.

Процедура синтеза стохастического поля с заданными статистическими свойствами строится на основании одномерной модели. Пусть элементу изображения  $f_0$  предшествует  $n$  элементов, составляющих множество  $\{f_1, f_2, \dots, f_n\}$ . Тогда условная плотность вероятности того, что данный элемент будет иметь некоторое значение яркости  $f_0$  при условии, что предшествующее множество было  $\{f_1, f_2, \dots, f_n\}$ , выражается в соответствии с уравнением:

$$p(f_0 / f_1, f_2, \dots, f_n) = \frac{p(f_0, f_1, f_2, \dots, f_n)}{p(f_1, f_2, \dots, f_n)}. \quad (9.15)$$

Если закон плотности вероятности является нормальным, то уравнение (9.15) может быть представлено в виде:

$$p(f_0 / f_1, f_2, \dots, f_n) = \frac{\frac{1}{\sqrt{(2p)^{n+1} |\mathbf{B}_{n+1}|}} \times e^{-\frac{(\mathbf{h}_{n+1} - \mathbf{z}_{n+1})^T \mathbf{B}_{n+1}^{-1} (\mathbf{h}_{n+1} - \mathbf{z}_{n+1})}{2}}}{\frac{1}{\sqrt{(2p)^n |\mathbf{B}_n|}} \times e^{-\frac{(\mathbf{h}_n - \mathbf{z}_n)^T \mathbf{B}_n^{-1} (\mathbf{h}_n - \mathbf{z}_n)}{2}}}, \quad (9.16)$$

где  $\mathbf{h}_n = \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ \vdots \\ f_n \end{bmatrix}$ ,  $\mathbf{h}_{n+1} = \begin{bmatrix} f_0 \\ f_1 \\ \vdots \\ \vdots \\ f_n \end{bmatrix}$ ,  $\mathbf{B}$  - ковариационная матрица,

$\mathbf{z}$  - математическое ожидание.

Для стационарного процесса, когда математическое ожидание и дисперсия яркости постоянны по кадру, условная плотность вероятности определяется ковариационной матрицей в соответствии с (9.17):

$$\mathbf{B}_{n+1} = y^2 \begin{bmatrix} 1 & \alpha & \beta & \gamma & \dots & \dots & \dots \\ \alpha & \ddots & & & & & \\ \beta & & \ddots & & & & \\ \gamma & & & \ddots & & & \\ \vdots & & & & \ddots & & \\ \vdots & & & & & \ddots & \\ \vdots & & & & & & \ddots \end{bmatrix}, \quad (9.17)$$

где  $\alpha$  - коэффициент корреляции между соседними элементами,  $\beta$  - коэффициент корреляции между элементами, расположенными через элемент,  $\gamma$  - коэффициент корреляции между элементами, расположенными через два элемента друг от друга и т.д.

В этом случае процесс формирования элемента  $f_0$  описывается следующим уравнением:

$$f_0 = a_w W_0 + a_0 + \sum_{j=1}^n a_j f_j, \quad (9.18)$$

где:  $W_0$  - случайное число, распределенное по нормальному закону с математическим ожиданием, равным 0, и СКО равным 1;  $a_j$  - весовые коэффициенты, которые получают из уравнения

$$b_j a_j = d_j, \quad (9.19)$$

$$\text{где } d_j = E\{(f_0 - \eta)(f_j - \eta)\} \text{ для } 1 \leq j \leq n. \quad (9.20)$$

Смещение оценивается по формуле:

$$a_0 = \eta(1 - \sum_{j=1}^n a_j), \quad (9.21)$$

а весовой коэффициент шума  $a_W = \sigma [1 - a_1\alpha - a_2\beta - a_3\gamma - \dots]^{1/2}$ , где  $\sigma$  - СКО.

Изображение задается матрицей с  $j$  столбцами и  $i$  строками  $\{f(s), s \in \Omega\}$ ,  $\Omega = \{(i,j), 0 \leq i, j \leq M-1\}$ . Одномерный процесс (9.18) используется на первом этапе для генерации всех элементов яркостного компонента изображения. Затем, над элементами всех строк, начиная со второй, производится аналогичная операция для обеспечения заданной корреляции по столбцу в соответствии с уравнением (9.20):

$$g_0 = b_W \left( \frac{f_0 - \eta}{y} \right) + b_0 + \sum_{j=1}^n b_j g_j \quad (9.20)$$

Поскольку человек способен различать текстуры с плотностями распределения не выше второго порядка, то представляется целесообразным синтезировать текстуры с заданной плотностью распределения второго порядка. В этом случае уравнения (9.18)÷(9.20) будут иметь следующий вид:

$$a_1 = \frac{E[(f_0 - \eta)(f_1 - \eta)]}{y^2} = \alpha, \quad (9.21)$$

где  $\alpha$  - коэффициент корреляции по строке,  $\eta$  - средняя яркость,  $\sigma^2$  - дисперсия яркости,  $f_1$  - яркость элемента слева.

Смещение равно  $a_0 = \eta(1 - a_1) = \eta(1 - \alpha)$ , а шумовой коэффициент  $a_W = \sigma \sqrt{1 - a_1 \alpha} = \sigma \sqrt{1 - \alpha^2}$ .

Тогда (9.21) можно представить в виде:

$$f_0 = W_0 y \sqrt{1 - \alpha^2} + z(1 - \alpha) + \alpha f_1. \quad (9.22)$$

Соответственно (9.20) для плотности распределения второго порядка может быть представлено в виде:

$$g_0 = (f_0 - \eta) \sqrt{1 - \beta^2} + \eta(1 - \beta) + \beta g_1. \quad (9.23)$$

Изменением параметров распределений, а именно коэффициентов корреляции по строке и по столбцу, можно создавать разные типы текстур, а задавая математическое ожидание и СКО, обеспечивать разную энергетику процессов.

При задании более сложных локальных пространственных зависимостей между яркостями соседних элементов используют гауссово-марковские случайные поля (МСП) в качестве моделей текстур [97]. Два элемента изображения  $f(x,y)$  и  $f(x',y')$  являются ближайшими соседями, если  $x = x'$  и  $y = y' \pm 1$  или если  $y = y'$  и  $x = x' \pm 1$ . Условная плотность вероятности элемента  $f(x,y)$  при таком подходе для МСП 1-го порядка определяется следующим выражением:

$$p(f(x,y) / f(x-1,y), f(x+1,y), f(x,y-1), f(x,y+1)). \quad (9.24)$$

В соответствии с этим представлением определяется иерархия МСП более высокого порядка. Для МСП 1-го порядка (9.24) соседи определяются как элементы, расстояние до которых равно 1,  $N=\{(0,1), (0,-1), (-1,0), (1,0)\}$ , всего их 4, и они отмечены на рисунке 6.1 цифрой 1. Конфигурация соседей для МСП 2-го порядка определяется элементами, расстояние до которых  $\leq \sqrt{2}$ ,  $N=\{(0,1), (0,-1), (-1,0), (1,0), (-1,1), (1,-1), (1,1), (-1,-1)\}$ , всего таких элементов 8, и они отмечены на рисунке 6.1 цифрами 1 и 2, и так далее.

При условии, что яркость элемента  $f(s)$  имеет нормальное распределение, она задается линейной комбинацией яркостей соседних элементов  $f(s+r)$ ,  $r \in N$  плюс нормальный шум  $n(s)$  с нулевым математическим ожиданием и дисперсией  $\sigma^2$ :

$$g(s)=\sum_{r \in N} \theta_r f(s+r) + n(s). \quad (9.25)$$

В (9.25)  $\theta_r$ ,  $r \in N$  и  $\sigma$  являются параметрами модели, при этом рассматривается случай, когда математическое ожидание яркостного компонента равно 0,  $E(g(s))=0$ . Достаточным условием стационарности яркостного процесса является выполнение условия:

$$1 - \sum_{(i,j) \in N} \theta_{i,j} f_1^i f_2^j \neq 0 \text{ для всех } f_1, f_2 \text{ таких что } |f_1|=1, |f_2|=1; \quad (9.26)$$

$N$  не должно быть симметричным. Если  $N$  симметрично, необходимо выполнить условие противоположной симметрии:

$$\theta_{i,j} = \theta_{-i,-j}. \quad (9.27)$$

В случае симметричного выбора соседей в силу выполнения (9.27)  $N$  характеризуется полностью одной из симметричных половин набора соседей, например, для МСП 2-го порядка  $N_S=\{(-1,-1), (-1,0), (-1,1), (0,-1)\}$ , то есть, если  $r \in N_S$ , то  $-r \notin N_S$  и  $N=(r: r \in N_S) \cup (-r: r \in N_S)$ . Уравнение (9.25) в этом случае можно представить в виде:

$$g(s)=\sum_{r \in N_S} \theta_r (f(s+r) + f(s-r)) + n(s). \quad (9.28)$$

Корреляционные свойства гауссовского шума при этом описываются следующим соотношением:

$$E[n(s)n(r)] = \begin{cases} -\theta_{s-r}\sigma^2, & (s-r) \in N \\ \sigma^2, & s=r \\ 0, & (s-r) \notin N, s \neq r \end{cases}. \quad (9.29)$$

Возможные значения параметров  $\theta_r$  ограничиваются требованием положительной определенности матрицы ковариаций:

$$\mathbf{M}_S = (1 - 2\mathbf{i}^T \mathbf{u}_S) > 0, \quad s \in \Omega, \quad (9.30)$$

где  $\mathbf{i}=[\theta_r, r \in N_S]$ , упорядоченные в виде вектора-столбца,

$\mathbf{u}_S = [\cos \frac{2\pi}{M} s^T r, r \in N_S]$ , также представленные в виде вектора-столбца.

На основании такого описания вектор признаков текстуры включает в себя оценки параметров  $\mathbf{i}$ , дисперсии и математического ожидания. При решении этой задачи по методу наименьших квадратов

$$\mathbf{i}^* = \left[ \sum_{\Omega} \mathbf{q}_s \mathbf{q}_s^T \right]^{-1} \sum_{\Omega} (\mathbf{q}_s \mathbf{y}_s), \quad (9.31)$$

где  $\mathbf{q}_s$  - вектор-столбец, составленный из  $f(s+r)$ ,  $r \in N_S$ .

$$\sigma^{*2} = \frac{1}{M^2} \sum_{\Omega} (\mathbf{g}_s - \mathbf{i}^{*T} \mathbf{q}_s)^2, \quad s \in \Omega. \quad (9.32)$$

Таким образом, уравнение (9.25) задает описание яркостного компонента цветной текстуры. Для синтеза цветной текстуры необходимо задать изменение компонентов тона и насыщенности. Поскольку в реальных изображениях компонент насыщенности оказывается коррелированным с негативом яркости, то в соответствии с физическими представлениями будем синтезировать его таким образом:

$$\text{sat}(s) = \alpha \bar{g}(s). \quad (9.33)$$

Компонент тона является характеристикой объекта и должен задаваться пользователем. При этом компонент тона будем задавать в виде случайного нормального распределения с заданными математическим ожиданием и дисперсией. На рисунке 9.4 представлены примеры синтеза текстур по этому методу.

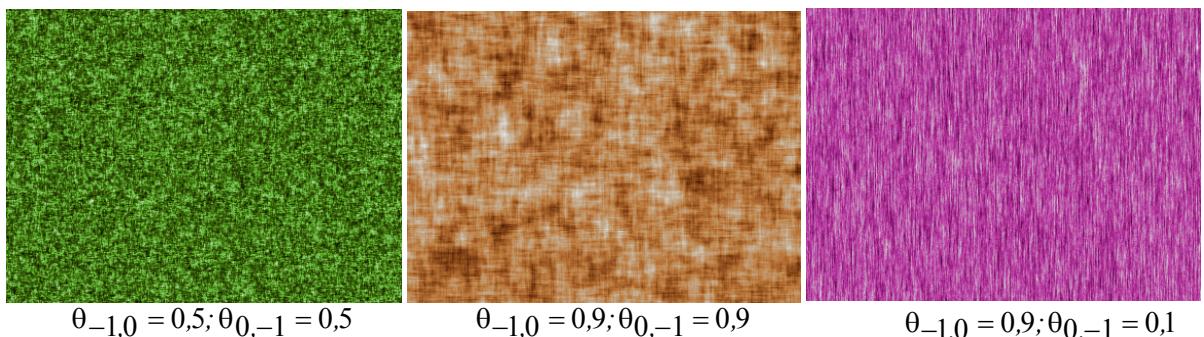


Рисунок 9.4 Примеры синтезированных текстур по статистическому методу.

#### 9.4.2 Синтез цветных фрактальных текстур.

Синтез яркостного компонента цветной фрактальной текстуры будем осуществлять на основании алгоритма последовательных случайных сложений, предложенного Фоссом [63]. Алгоритм построен в соответствии с законом обобщенного броуновского движения, для которого дисперсия разностей яркостей отсчетов, отстоящих на заданном расстоянии  $t$  друг от друга, подчиняется соотношению:

$$\sigma^2 = |t|^{2H} \sigma_0^2,$$

где  $\sigma_0^2$  - начальная дисперсия случайных сложений,  $H$  - показатель Херста.

Синтез яркостного компонента изображения производится на сетке  $s = (i,j)$ ,  $0 \leq i,j \leq M$  при использовании датчика случайных чисел, распределенных по нормальному закону с нулевым математическим ожиданием и дисперсией, изменяющейся на каждом шаге в соответствии с формулой:

$$\sigma_n^2 = r^{2nH}, \quad (9.34)$$

где  $r = 1/\sqrt{2}$  - масштабный коэффициент изменения расстояния между старыми и новыми отсчетами;  $n$  - номер шага. Фрактальная размерность при этом  $D=3-H$ .

Алгоритм синтеза можно описать следующим образом.

Шаг 0. Исходные значения яркостного компонента в точках с координатами  $(0,0)$ ,  $(0,M)$ ,  $(M,0)$ ,  $(M,M)$  задаются равными нулю. Значение яркостного компонента в точке с координатами  $(M/2, M/2)$  выбирается как случайное число (распределенное поциальному закону с математическим ожиданием равным нулю и дисперсией, равной 1).

Шаг 1.

1а) Вычисляются значения яркостного компонента на границах области в точках с координатами  $(0, M/2)$ ,  $(M/2, 0)$ ,  $(M/2, M)$ ,  $(M, M/2)$  как среднее арифметическое значений яркостного компонента в двух ближайших узлах, например  $f(0, M/2) = [f(0,0) + f(0, M)]/2$ .

В результате вычислений четыре области размером  $(M/2) \times (M/2)$  оказываются определены четырьмя узловыми точками.

1б) Для каждой области вычисляются значения яркостного компонента в центре области, как среднее арифметическое яркостей в узлах области, например:

$$f\left(\frac{M}{4}, \frac{M}{4}\right) = \frac{1}{4} \left( f(0,0) + f\left(\frac{M}{2}, 0\right) + f\left(\frac{M}{2}, \frac{M}{2}\right) + f\left(0, \frac{M}{2}\right) \right).$$

1в) К каждому из вычисленных значений яркости прибавляется случайное число  $(0, \sigma_1^2)$ .

Шаг 2.

2а) Число областей, заданных своими узловыми точками, увеличивается в 4 раза. Для каждой области вычисляются значения яркостного компонента на границах области как среднее арифметическое значений яркостного компонента в двух ближайших узлах.

2б) Для каждой области вычисляются значения яркостного компонента в центре области, как среднее арифметическое яркостей в узлах области.

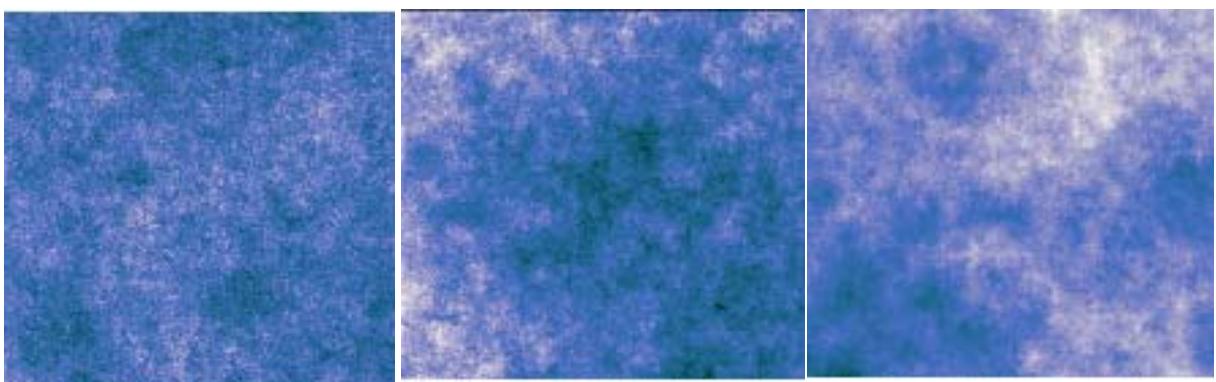
2в) К каждому из вычисленных значений яркости прибавляется случайное число  $(0, \sigma_2^2)$ .

Шаги  $[3,m]$ . (Здесь  $M=2^m$ ) Выполняются также как и шаг 2, число определенных областей при этом увеличивается в 4 раза. После

выполнения шага  $m$  все отсчеты на сетке изображения оказываются заданными.

Синтез цветного фрактала будем осуществлять на основе подхода, предложенного при синтезе цветных гауссово-марковских полей. Компонент насыщенности будем задавать как некоторое взвешенное значение негатива яркостного компонента, компонент тона задаем как реализацию случайного нормального процесса с математическим ожиданием и СКО, определяющими тоновые характеристики синтезируемого фрактала.

На рисунке 9.5 представлены примеры текстур, синтезированных в соответствии с разработанным алгоритмом. Значение яркостного компонента приведено к диапазону АЦП, то есть значениям из интервала [0,255]. Значение насыщенности получено как негатив яркостного компонента.



а) б) в)

Рисунок 9.5 Пример синтезированных текстур при  $H$  равном 0,1; 0,5 и 0,8 соответственно.

### 9.5 Сегментация цветных текстурных изображений

Под сегментацией цветных текстур мы понимаем процесс пространственного разбиения изображения на области, однородные относительно некоторого набора характеристик, то есть принадлежащие одному кластеру.

Процессу сегментации предшествует процесс формирования пространства признаков, основанный на исследовании текстур и, в свою очередь, опирающийся на описание текстур. Текстуры, имеющие различные цветовые характеристики, могут быть успешно разделены на основании цветовых признаков.

При сегментации цветных текстур необходимо учитывать как спектральные характеристики объекта, так и их пространственные характеристики. При этом наиболее информативным из цветовых признаков представляется компонент тона. Однако разбиение изображения только по тоновому компоненту является недостаточным, поскольку объекты, имеющие один тон, но разную яркость, наблюдатель определяет как разные. Например, объекты, окрашенные в желтый или коричневый

цвета, для наблюдателя различны. Это же можно отнести и к объектам, обладающим разной насыщенностью. Сегментация по цвету на основании многомерного порогового ограничения сопряжена, прежде всего, с тем, что в сложных изображениях, содержащих множество небольших областей, имеющих различные цветовые признаки, нет выраженных провалов гистограмм, что в свою очередь затрудняет определение порогов и в конечном итоге приводит к ошибкам сегментации. Исходя из этого, представляется целесообразным первоначально выполнить сегментацию изображения по тоновому компоненту, затем применить алгоритм кластеризации по методу  $K$ -внутригрупповых средних [98] по критерию минимальной удаленности элемента изображения от центра кластера, а затем выполнить сегментацию кластеров по гистограммам  $H$ ,  $R$ ,  $G$  или  $B$  компонентов [99]. Причем, с целью уменьшения ошибок при выборе порогов, гистограмму распределений компонентов следует оценивать с учетом пространственных характеристик кластера [100]. С этой целью производится предварительная селекция связных компонентов и в выборку включаются только те из них, которые превышают некоторый заданный размер области. Затем для доопределенного множества кластеров производится кластеризация элементов изображения по методу  $K$ -внутригрупповых средних.

Предложенный алгоритм обладает следующими характеристиками:

- 1) Сочетание сегментации и кластеризации позволяет учитывать как пространственные, так и цветовые характеристики изображения.
- 2) Количество кластеров не является предопределенным, а вычисляется в процессе обработки в соответствии с информацией, содержащейся в обрабатываемом изображении.
- 3) Определение границы сегмента производится с точностью до элемента растра в отличие от фрагментарных методов, при использовании которых точность определения границы зависит от размера фрагмента.

Поскольку метод  $K$ -внутригрупповых средних является локально сходимым и не производит новых кластеров, успех его применения зависит от первоначального разбиения пространства признаков, в нашем случае это значения  $R$ ,  $G$ ,  $B$  компонентов сигнала. Чтобы обеспечить лучшие условия сходимости алгоритма, мы выполняем первоначальное разбиение по методу порогового ограничения, причем используем гистограмму распределения тонового компонента для сокращения времени выполнения разбиения.

Пусть  $C_j$  есть кластер с количеством элементов  $m_j \geq 2$ . Пусть  $C_k$  является непустым подмножеством  $C_j$ , то есть  $C_k \neq 0$ ,  $C_k \neq C_j$ ,  $C_k \in C_j$ . Пусть  $C_p = C_j - C_k$  является также непустым множеством разности  $C_j$  и  $C_k$ .

Центр масс кластера задается как

$$\bar{\mathbf{x}}_q = \frac{1}{m_q} \sum_{i \in C_q} \mathbf{x}_i, (q=j,k,p). \quad (9.35)$$

Сумма квадратов расстояний элементов от центра кластера определяется по формуле

$$e_q = \sum_{i \in C_q} \|\mathbf{x}_i - \bar{\mathbf{x}}_q\|^2, (q=i,k,p). \quad (9.36)$$

Возводя в квадрат правую часть уравнения (9.36), получим:

$$e_q = \sum_{i \in C_q} \|\mathbf{x}_i\|^2 - m_q \|\bar{\mathbf{x}}_q\|^2. \quad (9.37)$$

Центр масс кластера  $C_p$  определяется:

$$\bar{\mathbf{x}}_p = \frac{m_j \bar{\mathbf{x}}_j - m_k \bar{\mathbf{x}}_k}{m_j - m_k}. \quad (9.38)$$

Сумма квадратов расстояний элементов от центра кластера  $p$  может быть определена в соответствии с (9.37):

$$\begin{aligned} e_p &= \sum_{i \in C_p} \|\mathbf{x}_i\|^2 - m_p \|\bar{\mathbf{x}}_p\|^2 = \sum_{i \in C_j} \|\mathbf{x}_i\|^2 - \sum_{i \in C_k} \|\mathbf{x}_i\|^2 - \frac{1}{m_p} \|m_j \bar{\mathbf{x}}_j - m_k \bar{\mathbf{x}}_k\|^2 = \\ &e_j - e_k - \frac{m_j m_k}{m_j - m_k} \|\bar{\mathbf{x}}_j - \bar{\mathbf{x}}_k\|^2. \end{aligned} \quad (9.39)$$

В нашем случае, когда решение необходимо принимать для каждого элемента растра, оценка производится в соответствии с формулами (9.38) и (9.39) для числа векторов в кластере  $C_k$   $m_k = 1$ :

$$\bar{\mathbf{x}}_p = \frac{m_j \bar{\mathbf{x}}_j - \mathbf{x}_k}{m_j - 1}, \quad (9.40)$$

$$e_p = e_j - \frac{m_j}{m_j - 1} \|\bar{\mathbf{x}}_j - \mathbf{x}_k\|^2. \quad (9.41)$$

Для кластера  $C_p$ , образованного объединением двух кластеров  $C_j$  и  $C_k$ ,  $C_p = C_j \cup C_k$  при условии, что  $C_j \cap C_k = 0$  соответствующие формулы задаются следующими уравнениями:

$$\bar{\mathbf{x}}_p = \frac{m_j \bar{\mathbf{x}}_j + m_k \bar{\mathbf{x}}_k}{m_j + m_k}, \quad (9.42)$$

$$e_p = e_j + e_k + \frac{m_j m_k}{m_j + m_k} \|\bar{\mathbf{x}}_j - \bar{\mathbf{x}}_k\|^2. \quad (9.43)$$

Для нашего случая, когда кластер  $C_k$  состоит из одного вектора, эти уравнения принимают вид:

$$\bar{\mathbf{x}}_p = \frac{m_j \bar{\mathbf{x}}_j + \mathbf{x}_k}{m_j + 1} , \quad (9.44)$$

$$e_p = e_j + \frac{m_j}{m_j + 1} \|\bar{\mathbf{x}}_j - \mathbf{x}_k\|^2 . \quad (9.45)$$

Алгоритм  $K$ -внутригрупповых средних строится следующим образом. Для определенного по тоновому компоненту первоначального разбиения вычисляются оценки центров масс кластеров и суммы квадратов отклонений векторов, принадлежащих кластеру, от центров масс этих кластеров в соответствии с уравнениями (9.35), (9.36) и определяется сумма квадратов отклонений от центров масс кластеров по всем кластерам. Затем для каждого вектора  $\mathbf{x}_i$ , принадлежащего кластеру  $C_r$ , отыскивается тот кластер  $j \neq r$ , для которого выполняется условие:

$$\frac{m_r}{m_r - 1} \|\bar{\mathbf{x}}_r - \mathbf{x}_i\|^2 > \frac{m_j}{m_j - 1} \|\bar{\mathbf{x}}_j - \mathbf{x}_i\|^2 . \quad (9.46)$$

Если таким кластером оказывается кластер  $C_v$ , то сумма квадратов отклонений векторов от центров масс их кластеров уменьшается:

$$e_r - \frac{m_r}{m_r - 1} \|\bar{\mathbf{x}}_r - \mathbf{x}_k\|^2 + e_v + \frac{m_v}{m_v - 1} \|\bar{\mathbf{x}}_v - \mathbf{x}_k\|^2 . \quad (9.47)$$

Для кластера  $C_v$  вычисляется новое значение центра масс и суммы квадратов отклонений по формулам (9.44), (9.45), (9.40), (9.41) для кластеров  $C_v$  и  $C_r$  соответственно.

Такая перестановка приводит к уменьшению общей суммы квадратов отклонений векторов от центров масс кластеров, которым они принадлежат. Классический алгоритм  $K$ -внутригрупповых средних предполагает выполнение стольких итераций этого процесса, сколько потребуется для того, чтобы при двух последовательных итерациях сумма квадратов отклонений не изменилась.

Представим подробнее схему выполнения алгоритма. Размерность вектора  $\mathbf{x}$   $L=3$  (вектор задается своими  $RGB$  компонентами). Первоначальное разбиение выполняется по тоновому компоненту, и результат сегментации записывается в виде уровней отсчета изображения. Значение отсчета равно номеру кластера, сформированного после выполнения порогового ограничения.  $N$  - количество кластеров, полученных после сегментации, является параметром алгоритма.

Затем производится оценка центров кластеров  $S[j, k] \forall j \in [1, N]$ ,  $k \in [1, L]$  и суммы квадратов отклонений всех векторов кластера от центра кластера  $e[r] \forall r \in [1, N]$ , причем отклонение для каждого вектора определяется в пространстве  $RGB$ , то есть

$$z[r, j] = \sum_{k=1}^L (S[r, k] - x[k, j])^2 , \quad (9.48)$$

где  $k$  - номер компонента вектора,  $r$  - номер кластера,  $j$ - номер элемента в кластере;  $x[k, j]$  - значение  $k$ -го компонента  $j$ -го элемента изображения, принадлежащего кластеру  $r$ .

Вычисляется сумма квадратов отклонений от центра кластера по всем векторам, составляющим кластер:

$$e[r] = \sum_j z[r, j]. \quad (9.49)$$

Вычисляется сумма квадратов отклонений от центров кластеров по всем кластерам, составляющим изображение:

$$D = \sum_{r=1}^N e[r]. \quad (9.50)$$

Выполняется перераспределение векторов между кластерами таким образом, чтобы минимизировать  $D$ . Формируются новые оценки центров кластеров и суммы квадратов отклонений векторов, входящих в кластер, от центра кластера в соответствии с формулами (9.40), (9.41), 9.44, 9.45.

На рисунке 9.6 представлен график зависимости нормированного к максимальному значению значения суммы квадратов отклонений векторов от центров кластеров  $D/D_{\max}$  от числа итераций  $ntrace$ .

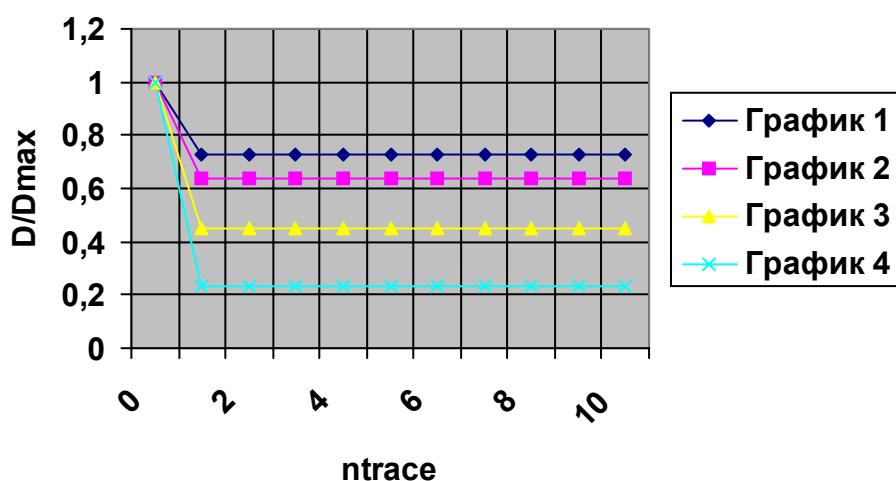
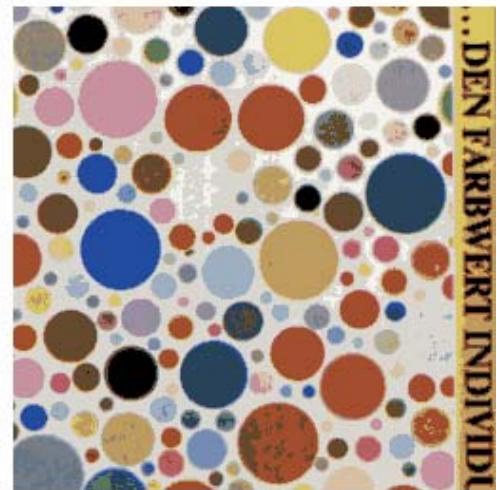
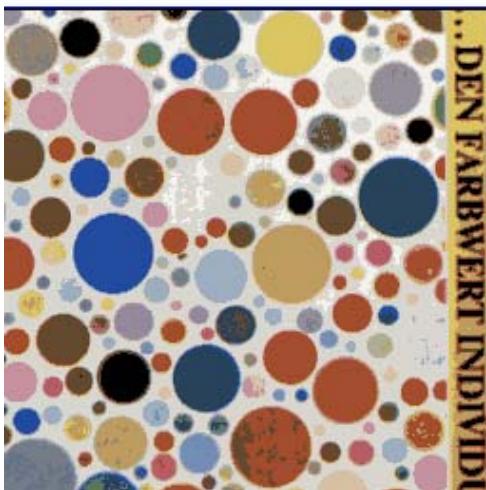
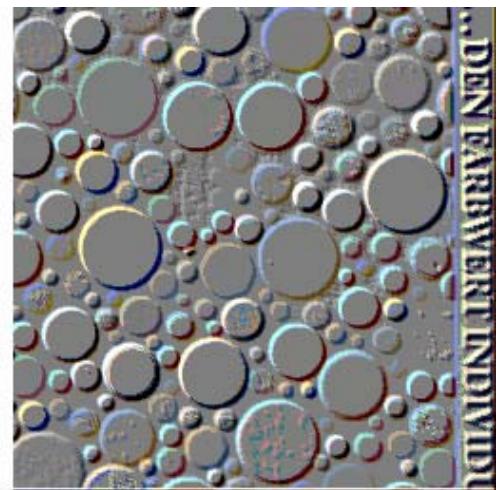
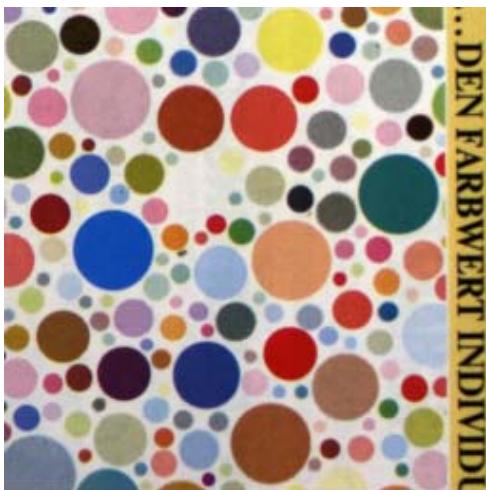


Рисунок 9.6 График зависимости нормированной величины внутрикластерных ошибок  $D/D_{\max}$  от числа итераций  $ntrace$ .

Исследования проведены по 50 различным цветным изображениям. Коэффициент уменьшения  $D$  от итерации к итерации изменяется, но характер зависимости соответствует представленному на рисунке. Из графика видно, что увеличение числа итераций не приводит к существенному уменьшению  $D$ . Метод  $K$ -внутригрупповых средних сходится локально. Эффективность кластеризации зависит от первоначального разбиения. На основании полученных результатов исследования ограничим число итераций:

$$ntrace=1. \quad (9.51)$$

На рисунке 9.7 приведен пример изображений, в которых кластеры представлены центрами масс.



а) б)  
в) г)

Рисунок 9.7 Пример классификации изображений. Кластеры представлены центрами масс. а) Исходное изображение; в) изображение получено после кластеризации по методу  $K$ -внутригрупповых средних при числе итераций  $ntrace=1$ ; г) изображение получено после кластеризации по методу  $K$  – внутригрупповых средних при числе итераций  $ntrace=10$ ; б) изображение, сформированное как разность изображений в) и г).

Изображения получены после кластеризации по методу  $K$ -внутригрупповых средних при числе итераций  $ntrace=1$  (рисунок 9.7 в) и  $ntrace=10$  (рисунок 9.7 г). На рисунке 9.7 б) показано изображение, сформированное как разность изображений в) и г), по которому видно, что увеличение числа итераций приводит к уточнению разбиения прежде всего на границах кластеров, но не уменьшает ошибок разбиения, вызванных первоначальным разбиением. Например, изображения красно-коричневых кругов (рисунок 9.7 а) отнесены к одному кластеру. Увеличение числа итераций не приводит к их разделению, хотя глаз хорошо различает эти круги. Для уменьшения ошибок кластеризации выполним дополнительное разбиение кластеров посредством сегментации по гистограммам  $H$ ,  $R$ ,  $G$ ,  $B$  компонентов, полученным для каждого кластера, по методу порогового

ограничения. Сформируем кластеры и применим алгоритм  $K$  – внутригрупповых средних для доопределенного множества кластеров. На рисунке 9.8 г) видно, что вследствие доопределения кластеров, выполнено разделение кругов красно-коричневого цвета на кластеры, соответствующее зрительному восприятию.

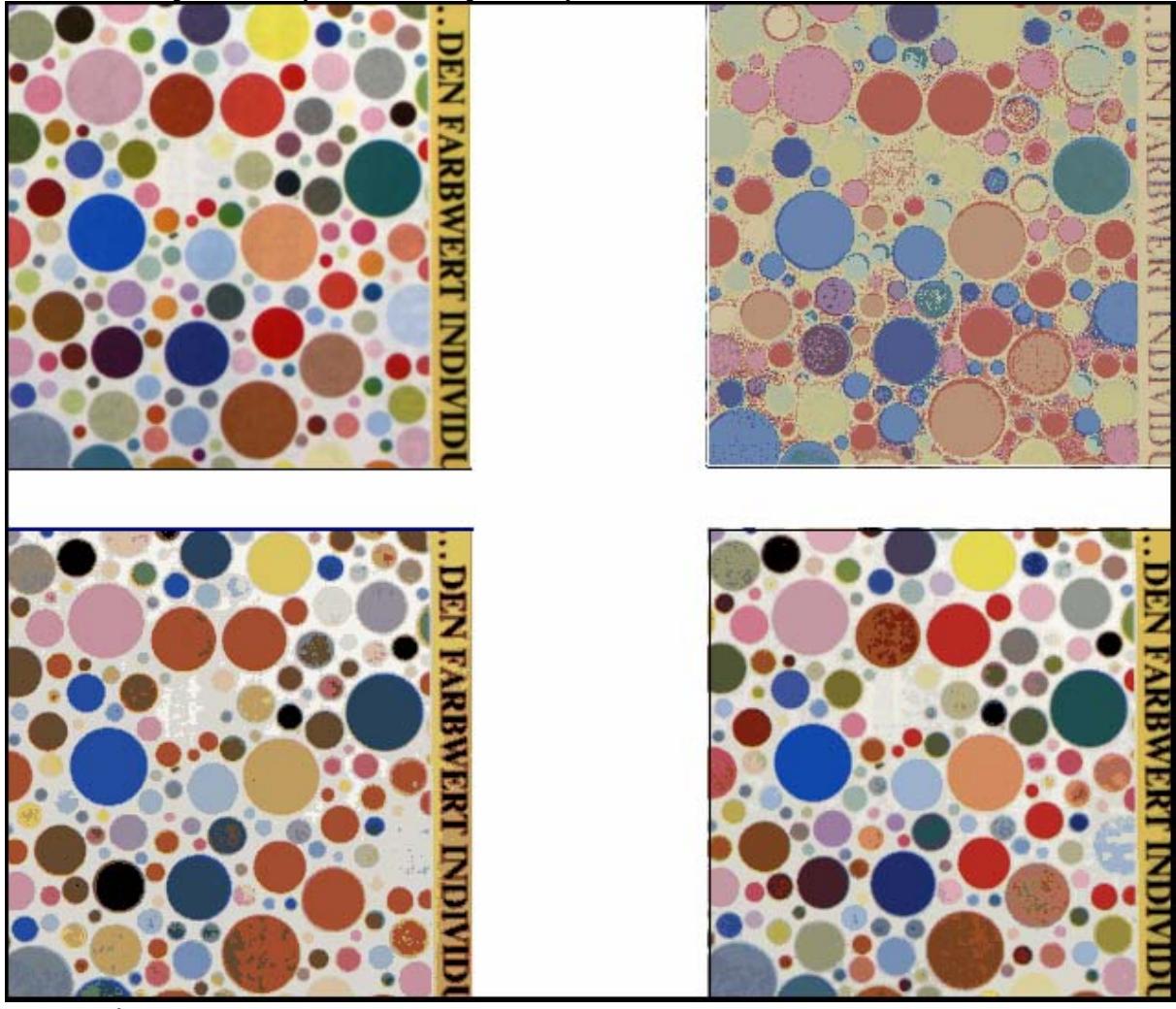


Рисунок 9.8 Пример классификации: а) исходное изображение; б) изображение, представленное центрами масс кластеров после первоначального разбиения по тоновому компоненту; в) изображение, представленное центрами масс кластеров, при кластеризации по методу  $K$ -внутригрупповых средних после выполнения 1 итерации; г) изображение, представленное центрами масс кластеров, после доопределения кластеров и кластеризации по методу  $K$ -внутригрупповых средних.

На рисунке 9.9 представлены маски красно-коричневого кластера, полученные на различных шагах алгоритма. Для повышения эффективности алгоритма на этапе оценки гистограмм распределения компонентов сигнала, наряду с цветовыми характеристиками, используется пространственная характеристика изображения. А именно, для уменьшения влияния ошибок первоначального разбиения по гистограмме тонового компонента, для каждого кластера выполняется

селекция связных компонентов, исключаются из рассмотрения все связные области, имеющие некоторый заданный размер, оценка гистограммы производится только для связных областей кластера, превышающих этот заданный размер.

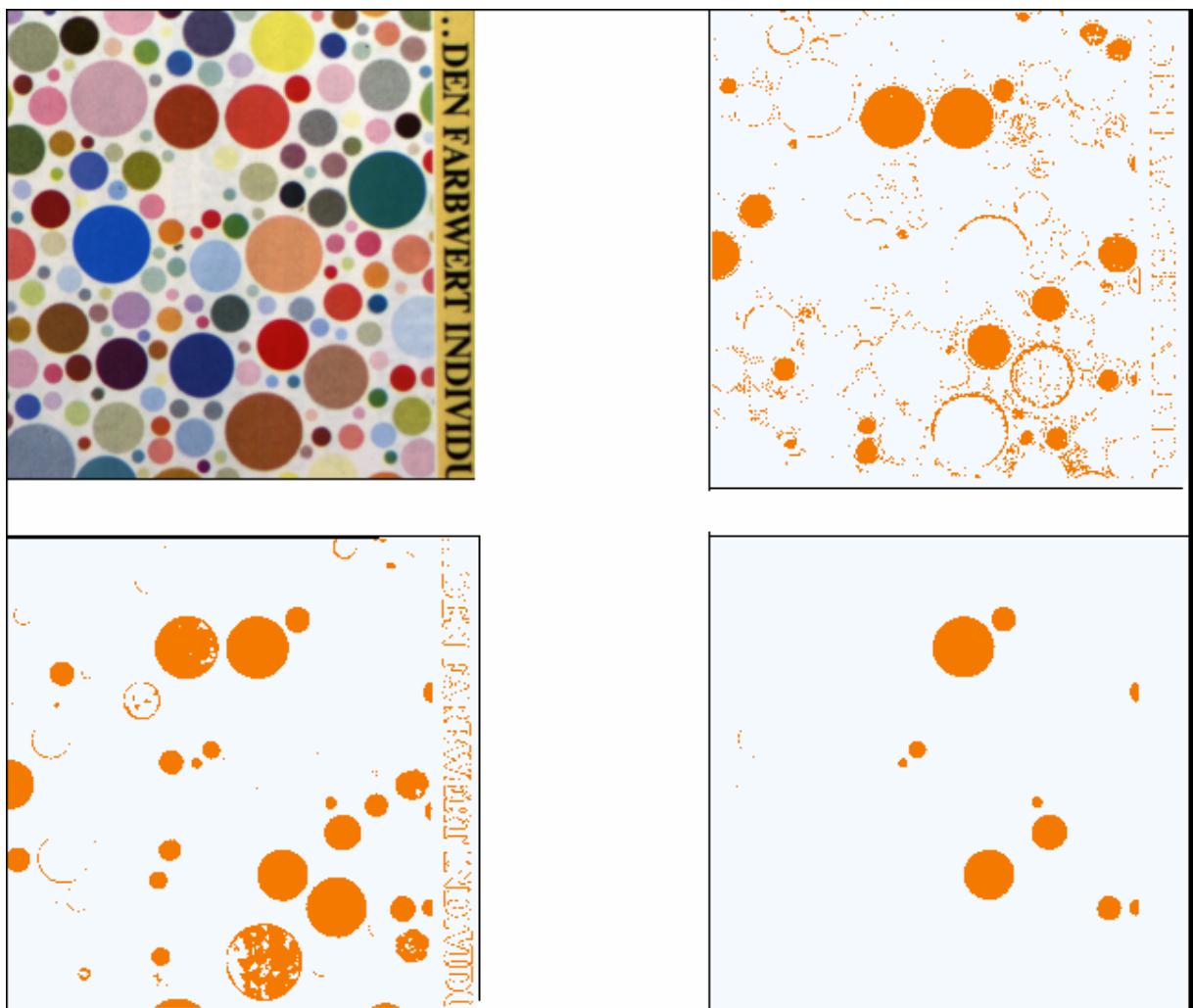


Рисунок 9.9 Маски красно-коричневого кластера, полученные на различных шагах.

Для сокращения избыточности кластеризации используется метод иерархического слияния кластеров [99]. В качестве меры межкластерных расстояний используется мера Махаланобиса. При работе алгоритма  $K$ -внутригрупповых средних мы использовали меру удаленности векторов от центра кластеров в пространстве  $RGB$  для вычисления в соответствии с формулой (9.36) .

Аналогично можно определить и расстояние между кластерами:

$$d[i, j] = \sqrt{(R[i] - R[j])^2 + (G[i] - G[j])^2 + (B[i] - B[j])^2}, \quad (9.52)$$

где  $R[i]$ ,  $G[i]$ ,  $B[i]$  -  $RGB$  координаты центра масс кластера  $i$ ;  $R[j]$ ,  $G[j]$ ,  $B[j]$  -  $RGB$  координаты центра масс кластера  $j$ .

При объединении кластеров будем учитывать также дисперсию плотности распределения  $RGB$  компонентов кластера. Это можно сделать

при использовании меры Махalanобиса [84], описываемой следующим уравнением:

$$md[i, j] = \frac{d[i, j]}{\sqrt{\sigma_i^2 + \sigma_j^2}}, \quad (9.53)$$

где  $d[i, j]$  определяется в соответствии с формулой (9.52.);  $\sigma_i^2$ ,  $\sigma_j^2$  - дисперсии плотностей распределения  $RGB$  компонентов кластеров  $i, j$  соответственно.

Рассмотренный метод автоматической классификации цветных текстурных изображений является синтезом метода квантования гистограмм и метода кластеризации по  $K$ -внутригрупповым средним. Такой синтез методов позволяет, не делая предположений о законах распределения кластеров, на основании информации, содержащейся в изображении, получить более гибкую форму кластера. При автоматической сегментации цветных текстурных изображений на первом шаге выполняется сегментация по гистограмме распределения тонового компонента для сокращения времени выполнения разбиения. На втором шаге выполняется алгоритм кластеризации по методу  $K$ -внутригрупповых средних по критерию минимальной удаленности элемента изображения от центра кластера в пространстве  $RGB$ . На третьем шаге для каждого кластера производится селекция связных компонентов с целью уменьшения ошибок при выборе порогов по гистограммам распределений компонентов с учетом пространственных характеристик кластера. В выборку включаются только те связные области кластера, которые превышают некоторый заданный размер области. Вычисляются гистограммы распределения компонентов  $R, G, B$ .

Определяются пороги квантования и производится дополнительное разбиение кластеров. На четвертом шаге для доопределенного множества кластеров производится кластеризация элементов изображения по методу  $K$ -внутригрупповых средних. На пятом шаге выполняется алгоритм иерархического объединения кластеров по критерию минимума меры Махalanобиса. На шестом шаге выполняется объединение кластеров. Предложенный алгоритм обладает следующими преимуществами. Он учитывает как пространственные, так и цветовые характеристики изображения. Количество кластеров не является предопределенным, а вычисляется в процессе обработки в соответствии с информацией, содержащейся в обрабатываемом изображении. Определение границы сегмента производится с точностью до элемента растра в отличие от фрагментарных методов, при использовании которых точность определения границы зависит от размера фрагмента. Алгоритм обеспечивает сокращение пространства признаков с 16 миллионов до нескольких десятков кластеров.

## **9.6 Фрактальный анализ сложных текстурных изображений**

### **9.6.1 Оценка фрактальности признаков цветных текстур**

В то время как объекты, построенные человеком, такие как промышленные и жилые здания, могут быть эффективно описаны набором простых геометрических примитивов: кубов, сфер, цилиндров, конусов, цветные текстуры природного происхождения, являясь нерегулярными и фрагментарными, плохо поддаются такому описанию. При включении в систему признаков геометрических признаков используются некоторые аппроксимирующие оценки в виде равновеликого эллипса рассеяния, размеров его большой и малой полуосей и тому подобное. В связи с этим, для анализа таких текстур оказывается естественным представление их фракталом с некоторым размером фрактала  $D$ . В настоящее время нет еще окончательного определения фрактала. Ключевая концепция фракталов заключается в использовании самоподобия в определении размера фрактала.

В настоящее время фракталы нашли свое применение при анализе текстур ландшафтов, полученных при аэрокосмической съемке, при анализе поверхностей порошков и других пористых сред, при анализе поверхности облаков и так далее.

Однако размер фрактала цветной текстуры во многом зависит от выбора метода оценки. Так, при использовании различных методов оценки размера фрактала, мы получим соответственно и разные его размеры. Сопоставление текстур, таким образом, возможно при использовании одного и того же метода (группы методов).

Более того, не всякие текстуры хорошо различимы по размеру фрактала. В связи с этим прежде, чем включать в систему признаков размер фрактала, необходимо оценить фрактальность текстуры. Оценка фрактальности текстуры производится на основе выбранного метода оценки размера фрактала. Поскольку размер фрактала вычисляется через оценку выборочной регрессии, то естественно оценивать фрактальность текстуры по коэффициенту корреляции между логарифмом случайной величины и логарифмом заданной функции шага. При этом принятие решения о фрактальности текстуры можно строить следующим образом:

- 1) построить зависимость коэффициента корреляции от шага; значение шага, при котором функция имеет максимум, является максимальным шагом в диапазоне задаваемых шагов при оценке размера фрактала;
- 2) не учитывать оценку размера фрактала при низком коэффициенте корреляции в тех методах, где используется оценка фрактала как среднее значение в серии экспериментов;
- 3) не включать размер фрактала в систему признаков для сегментации текстур при значениях коэффициента корреляции  $< 0,7$ .

Оценка фрактальности текстуры является важной характеристикой при сегментации по размеру фрактала.

## 9.6.2 Возможности и ограничения применения алгоритма оценки размера фрактала по длине контура при анализе сложных текстурных изображений

Алгоритм оценки размера фрактала текстуры по длине контура [101] состоит в развитии алгоритма оценки размера фрактала линии для оценки размера фрактала поверхности. Для оценки фрактала текстуры производится разбиение динамического диапазона яркостей изображения на равные интервалы. Для полученного набора пороговых уровней строится бинарное изображение. При этом отсчетам, яркость которых меньше порога, приписывается значение 0, а отсчетам, яркость которых выше или равна порогу, приписывается значение 1. Таким образом, исходное изображение представляется набором бинарных изображений. Для каждого из таких изображений производится оценка размера фрактала контуров единичных областей. А в качестве оценки размера фрактала исходного изображения используется среднее значение полученных фракталов для бинарных изображений. При этом предлагается оценивать размер фрактала бинарных изображений только по строкам, только по столбцам, а также совместно по строкам и столбцам, что имеет особое значение при распознавании анизотропных текстур.

Процедура оценки размера фрактала контура строится следующим образом. Производится оценка длин контуров единичной области для серии размеров элемента разложения (шагов)  $S_i$ . Увеличение шага эквивалентно интерпретации анализируемого изображения с меньшим разрешением, чем разрешение, с которым изображение получено. Длина контура  $L$  аппроксимируется числом переходов уровней яркости бинарного изображения из 0 в 1 и из 1 в 0 для каждого значения шага. По полученным значениям оценивается регрессия логарифма длины контура на логарифм шага [102] в виде функции

$$y = \theta_0 + \theta_1 x + \theta_2 x^2 + \dots + \theta_{m-1} x^{m-1}, \quad (9.54)$$

где

$$\mathbf{x} = \begin{bmatrix} 1 & x_1 & x_1^2 & \dots & x_1^{m-1} \\ 1 & x_2 & x_2^2 & \dots & x_2^{m-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \dots & x_n^{m-1} \end{bmatrix}; \mathbf{\theta} = \begin{bmatrix} \theta_0 \\ \theta_1 \\ \vdots \\ \theta_{m-1} \end{bmatrix}, \quad (9.55)$$

где  $n$ -число шагов,  $y_i = \log L_i$ ,  $x_i = \log S_i$ .

По методу наименьших квадратов оценка  $\hat{\mathbf{\theta}}$  является решением системы нормальных уравнений:

$$\begin{bmatrix} n & \sum x_i^1 & \sum x_i^2 & \dots & \sum x_i^{m-1} \\ \sum x_i & \sum x_i^2 & \sum x_i^3 & \dots & \sum x_i^m \\ \vdots & \vdots & \vdots & \dots & \sum x_i^{2(m-1)} \end{bmatrix} \times \begin{bmatrix} \theta_0 \\ \theta_1 \\ \vdots \\ \theta_{m-1} \end{bmatrix} = \begin{bmatrix} \sum y_i \\ \sum x_i y_i \\ \vdots \\ \sum x_i^{m-1} y_i \end{bmatrix}. \quad (9.56)$$

При вычислении размера фрактала используется линейная выборочная регрессия:

$$\hat{\theta} = \begin{bmatrix} \hat{\theta}_0 \\ \hat{\theta}_1 \end{bmatrix} = \begin{bmatrix} \frac{\sum x_i^2 \sum y_k - \sum x_i \sum x_k y_k}{n \sum x_i^2 - (\sum x_i)^2} \\ \frac{n \sum x_i y_i - \sum x_i \sum y_k}{n \sum x_i^2 - (\sum x_i)^2} \end{bmatrix}. \quad (9.57)$$

Размер фрактала оценивается по формуле

$$D = 2 - \hat{\theta}_1. \quad (9.58)$$

Качество “наилучшего” линейного приближения оценивается значением коэффициента корреляции логарифма длины контура и логарифма шага [103]:

$$r_{\log L, \log S} = \frac{\text{cov}(\log L, \log S)}{\sqrt{D_{\log L} \times D_{\log S}}} = \frac{E[(\log L - M_{\log L})(\log S - M_{\log S})]}{\sqrt{D_{\log L} \times D_{\log S}}}, \quad (9.59)$$

где  $D$  - дисперсия,  $M$  - математическое ожидание соответствующих случайных величин  $\log L$  и  $\log S$ .

При низком коэффициенте корреляции полученное значение размера фрактала исключается из процедуры усреднения.

Измерение размера фрактала по методу оценки длины контура при сканировании по строкам и по столбцам раздельно позволяет оценивать анизотропные свойства текстуры, в то время как комбинированный метод, при котором осуществляется подсчет краев как вдоль строк, так и вдоль столбцов, пригоден для анализа изотропных текстур.

В качестве модели для исследования оценки размера фрактала по длине контура используем синтезированные в соответствии с алгоритмом, представленным в разделе 9.4.2, фракталы с показателем Херста от 0,1 до 0,9.

Моделирование выполнено на серии из 50 реализаций фракталов. Поскольку распределение оценки размера фрактала имеет большую дисперсию, произведена низкочастотная фильтрация оценки размера фрактала. Значения  $m$  и  $\sigma$ , соответствующие значениям математического ожидания и СКО размеров фракталов, приведенных к диапазону значений [2,3] и представленных уровнями [0,255] приведены в таблице 9.1.

Таблица 9.1. Оценка размера фрактала по длине контура (по строкам и по столбцам) после низкочастотной фильтрации

N	Фрагмент 32x32		Фрагмент 16x16		Фрагмент 8x8	
	$m$	$\sigma$	$m$	$\sigma$	$m$	$\sigma$
0,1	192,19	5,28	197,89	3,93	201,87	6,0
0,2	174,97	5,4	190,41	4,1	195,79	5,6
0,3	158,85	9,0	183,58	4,5	190,78	4,9
0,4	168,93	8,1	179,14	4,2	186,87	3,8
0,5	158,55	7,46	174,04	5,14	178,9	3,64
0,6	154,39	9,42	167,04	5,91	175,64	5,16
0,7	146,01	8,08	159,59	5,73	168,67	4,82
0,8	137,23	8,45	150,25	6,17	163,6	4,83
0,9	137,57	8,14	148,45	6,69	164,17	5,6

На основании этих данных построены матрицы расстояний между этими 9 фракталами, вычисленные как мера Фишера. В качестве примера в таблице 9.2 приведены значения межфрактальных расстояний для окна 16x16.

Таблица 9.2 Матрица межфрактальных расстояний, определенных как мера Фишера при оценке размера фрактала по длине контура (по строкам и столбцам, окно 16x16)

$H$	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9
0,1	0	0,93	1,7	2,31	2,63	3,14	3,96	4,66	4,72
0,2		0	0,79	1,36	1,77	2,33	3,14	3,89	3,91
0,3			0	0,51	0,99	1,59	2,35	3,12	3,14
0,4				0	0,55	1,2	1,97	2,79	2,82
0,5					0	0,63	1,33	2,1	2,16
0,6						0	0,64	1,39	1,48
0,7							0	0,78	0,9
0,8								0	0,14
0,9									0

На основании проведенного исследования можно сделать следующие выводы:

- Метод оценки размера фрактала по длине контура можно использовать при условии низкочастотной фильтрации оценки размера фрактала.

- Исключение некоторых изаритмов из рассмотрения при оценке размера фрактала может приводить к неразличимости фракталов, имеющих разный размер.
  - Для изотропных текстур большей эффективностью обладает алгоритм оценки размера фрактала по строкам и по столбцам.
  - В задачах сегментации, когда необходимо различить объекты, а не собственно оценить размер фрактала, представляется целесообразным не вычислять размер фрактала, поскольку это лишь приводит к дополнительным ошибкам, связанным с оценкой тангенса угла наклона линейной выборочной регрессии, а использовать изменение характера зависимостей, по которым фрактал оценивается.

### **9.6.3 Сегментация текстурных изображений по методу треугольной пирамиды**

Метод треугольной пирамиды [104] устанавливает соотношение между площадью поверхности, создаваемой яркостью изображения, и пространственным разрешением двумерных единиц, используемых для измерения этой площади. Изображение рассматривается на квадратной сетке и измерения производятся для серии размеров этой сетки. Треугольная пирамида строится как показано на рисунке 9.10.

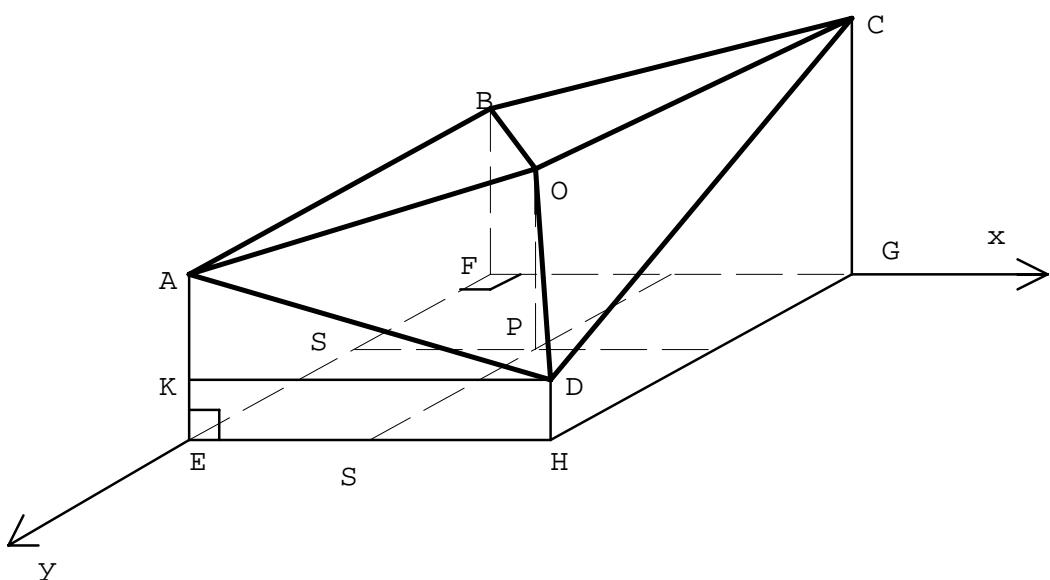


Рисунок 9.10 Построение пирамиды при оценке размера фрактала по методу пирамиды.

На плоскости раstra на расстоянии заданного шага  $S$  ( $FG = EH = FE = GH$ ) по строке и столбцу восстанавливаются перпендикулярно к раstrу 4 ребра, длины которых равны яркости соответствующих отсчетов изображения. Соединение вершин 4 ребер задает основание треугольной пирамиды  $ABCD$ . Вершина пирамиды  $O$  строится как вершина перпендикуляра к плоскости раstra, восстановленного из центра квадратной площадки размером  $S \times S$ , и равного среднему значению четырех опорных отсчетов яркостей:

$$OP = \frac{AE + DH + BF + CG}{4}. \quad (9.60)$$

Вычисляется площадь боковой поверхности полученной треугольной пирамиды  $OABCD$ .

Для вычисления площади боковой поверхности треугольной пирамиды необходимо определить 4 площади треугольников:

$$S_{OABCD} = S_{OAD} + S_{AOB} + S_{BOC} + S_{COD}. \quad (9.61)$$

Площадь каждого из треугольников определяется аналогично площади  $\Delta OAD$ :

$$S_{OAD} = \sqrt{p(p - AO)(p - AD)(p - OD)}, \quad (9.62)$$

где  $p = \frac{AO + AD + DO}{2}$  есть полупериметр  $\Delta OAD$ ,

$$AO = \sqrt{S^2/2 + (OP - AE)^2}, \quad AD = \sqrt{S^2 + (AE - DH)^2},$$

$OD = \sqrt{S^2/2 + (OP - DH)^2}$ ,  $AE$  и  $DH$  - яркости в соответствующих отсчетах.

Для каждого шага  $S_i$  на растре формируются пирамиды и вычисляется суммарная площадь боковых поверхностей этих пирамид. Так, при  $S_i = 1$  в вычислениях используются все отсчеты яркостного сигнала, при  $S_i = 2$  - в четыре раза меньше, при шаге  $S_i = 4$  - в 16 раз меньше отсчетов участвует в вычислениях и так далее. Сканирование осуществляется сверху вниз, слева направо. Затем строится выборочная регрессия логарифма суммарной площади боковых поверхностей пирамид на логарифм площади элемента раstra  $S_i^2$  в соответствии с уравнением (9.57).

Размер фрактала вычисляется по формуле (9.58), а качество оценки производится по коэффициенту корреляции в соответствии с (9.59).

Выполнены исследования оценки размера фрактала по алгоритму пирамиды после низкочастотной фильтрации для 9 различных фракталов и 3 размеров фрагментов, по которым производилась оценка. В таблице 9.3 приведена матрица межфрактальных расстояний для 9 различных фракталов и окна  $17 \times 17$ .

На основании проведенных исследований можно сделать следующие выводы:

- Метод пирамиды можно использовать для оценки размера фрактала при условии низкочастотной фильтрации оценки размера фрактала.
- Метод пирамиды имеет большую эффективность, чем алгоритм оценки размера фрактала по длине контура.
- В задачах сегментации, когда необходимо различить объекты, а не собственно оценить размер фрактала, представляется целесообразным не вычислять размер фрактала, поскольку это лишь приводит к

дополнительным ошибкам, связанным с оценкой тангенса угла наклона линейной выборочной регрессии, а использовать изменение характера зависимостей, по которым фрактал оценивается, если это возможно.

Таблица 9.3. Матрица межфрактальных расстояний при оценке размера фрактала по площади пирамиды (окно17x17)

$H$	$m$	$\sigma$	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9
0,1	75,2	3,2	0	0,4	0,91	1,65	2,08	2,49	3,71	4,49	5,17
0,2	72,8	2,8		0	0,57	1,32	1,79	2,24	3,53	4,37	5,1
0,3	69,5	3,0			0	0,68	1,15	1,6	2,78	3,58	4,26
0,4	65,8	2,5				0	0,53	1,05	2,29	3,16	3,91
0,5	63,1	2,8					0	0,53	1,68	2,52	3,24
0,6	60,2	2,3						0	1,07	1,86	2,54
0,7	54,8	2,2							0	0,86	1,59
0,8	51,0	2,1								0	0,73
0,9	47,8	2,1									0

#### 9.6.4 Оценка размера фрактала по модулю разности яркостей отсчетов

В основе этого метода лежит концепция статистического самоподобия цветных текстур природного происхождения, основывающаяся на том факте, что фракталы природного происхождения статистически инвариантны в широком диапазоне масштабов и каждый из компонентов статистически подобен другим компонентам. Математической моделью таких фракталов является фрактальная (обобщенная) броуновская функция [105].

Фрактальная броуновская функция  $f(x)$  является вещественной случайной функцией, такой, что для всех  $x$  и  $\Delta x$

$$P\left\{\frac{f(x + \Delta x) - f(x)}{\|\Delta x\|^H} < t\right\} = P(t), \quad (9.63)$$

где  $x$  представляет точку в  $n$ -мерном евклидовом пространстве  $R^E$  и  $P(t)$  является функцией распределения случайной величины  $t$ .

Обобщение броуновской функции состоит в том, что вместо  $1/2$  вводится действительный параметр  $H$ , некоторая постоянная, диапазон изменения которой  $[0,1]$ . Размер фрактала задается соотношением  $D=n+1-H$ . Для  $3D$  поверхностей ( $n=2$ ), размер фрактала определяется выражением

$$D=3-H. \quad (9.64)$$

Рассмотрим основные свойства фрактальной броуновской функции.  $P(t)$  описывает нормальное гауссовское распределение  $N(0, \sigma^2)$  с математическим ожиданием, равным 0, и дисперсией  $\sigma^2$ , то есть  $P(t)$  задается следующим уравнением:

$$P(t) = \frac{1}{\sqrt{2\pi}y} \int_{-\infty}^t \exp\left(-\frac{s^2}{2y^2}\right) ds. \quad (9.65)$$

При задании броуновской функции в соответствии с уравнением (9.63) в [110] установлено следующее соотношение:

$$E[f(x + \Delta x) - f(x)] \times \|\Delta x\|^{-H} = C, \quad (9.66)$$

где  $E[f(x + \Delta x) - f(x)]$  - математическое ожидание разностей значений функции, находящихся на расстоянии  $\|\Delta x\|$  друг от друга.

Постоянная  $C$  равна математическому ожиданию случайной величины  $|t|$  и является средним абсолютным отклонением. Из (9.65) соотношение между средним абсолютным отклонением  $C$  и СКО  $\sigma$  выражается следующей зависимостью:

$$\begin{aligned} C &= 2 \int_0^\infty s \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{s^2}{2\sigma^2}\right) ds = \frac{2}{\sqrt{2\pi}} \sigma \int_0^\infty \exp\left(-\frac{s^2}{2\sigma^2}\right) d\left(\frac{s^2}{2\sigma^2}\right) = \\ &= \sigma \sqrt{\frac{2}{\pi}} \left( -\exp\left(-\frac{s^2}{2\sigma^2}\right) \right) \Big|_0^\infty = \sigma \sqrt{\frac{2}{\pi}}. \end{aligned} \quad (9.67)$$

Логарифмируя уравнение (9.66), получим:

$$\log E[f(x + \Delta x) - f(x)] - H \log \|\Delta x\| = \log C. \quad (9.68)$$

Поскольку  $H$  и  $C$  являются постоянными, из уравнения (9.68) следует, что логарифм математического ожидания модуля разностей случайной величины  $f(x)$ , отстоящих на расстоянии  $\|\Delta x\|$ , линейно зависит от расстояния. Причем  $H$  определяет тангенс угла наклона этой прямой.  $E[f(x + \Delta x) - f(x)]$  является статистикой второго порядка, используемой в текстурном анализе [69]. Полученное соотношение (9.62) указывает на то, что в качестве признака текстуры может быть использован параметр  $H$ .

Таким образом, можно рассматривать изображение как двумерную функцию яркости  $f(x,y)$ , которая определена для  $(x,y) \in \mathbb{R}^2$ . Функция  $z=f(x,y)$  формирует трехмерную поверхность. Для оценки фрактальных признаков этой поверхности будем использовать аппроксимацию этой поверхности фрактальной броуновской функцией.

Размер фрактала  $D$  является существенным признаком при использовании фрактальной броуновской функции для описания

природных поверхностей. В этом случае размер фрактала оценивается в соответствии с (9.64). Таким образом, чтобы оценить размер фрактала, необходимо вычислить параметр  $H$ . На основании (9.68), чтобы вычислить  $H$ , необходимо получить оценку линейной выборочной регрессии логарифма СКО всех разностей пар элементов, отстоящих на определенном расстоянии друг от друга, на логарифм этого расстояния. Расстояния объединяются в кластеры, и СКО вычисляются для каждого кластера. Размер фрактала представляет характеристику сложности поверхности. При одной и той же функции распределения  $P(t)$  малые значения размера фрактала  $D$  описывают гладкие поверхности, а большие - более сложные изрезанные поверхности.

Функция распределения  $P(t)$  из уравнения (9.63) также характеризует вид поверхности. В случае фрактальной броуновской модели предполагается гауссовское распределение с нулевым математическим ожиданием и дисперсией, равной  $\sigma^2$ , таким образом, распределение характеризуется только одним параметром  $\sigma^2$ .

Хотя размер фрактала фрактальной броуновской функции инвариантен при изменении масштаба, реальные природные поверхности не обладают одним и тем же размером фрактала во всем возможном диапазоне масштабов. Реальные природные поверхности имеют один и тот же размер фрактала для большего или меньшего диапазона шкал. В связи с этим следует производить оценку диапазона расстояний  $[\|\Delta x\|_{\min}, \|\Delta x\|_{\max}]$ , в

котором поверхность может быть описана фрактальной броуновской функцией. При этом качество "наилучшего" линейного приближения оценивается по формуле, аналогичной формуле (9.59). Максимальное расстояние, на котором можно оценивать дисперсию разностей отсчетов яркости, определяется как точка максимума зависимости коэффициента корреляции от расстояния.

В качестве модели для исследования оценки размера фрактала по методу броуновской функции используем синтезированные фракталы с показателем Херста от 0,1 до 0,9. Поскольку распределение оценки размера фрактала имеет большую дисперсию, произведена низкочастотная фильтрация оценки размера фрактала. Моделирование выполнено на серии из 50 реализаций. Исследовано 9 различных фракталов при 3 размерах фрагментов, по которым производилась оценка фракталов.

В таблице 9.4 приведены данные для оценки размера фрактала по методу броуновской функции для размера фрагмента 16x16.

На основании выполненных исследований можно сделать следующие выводы:

- Эффективность метода броуновской функции может быть повышена за счет низкочастотной фильтрации оценки фрактала.
- Оценку размера фрактала можно также производить, учитывая изменения яркости только по строкам или только по столбцам, что

позволит повысить эффективность сегментации анизотропных текстур.

- Метод броуновской функции обладает наибольшей эффективностью по сравнению с другими рассмотренными методами.

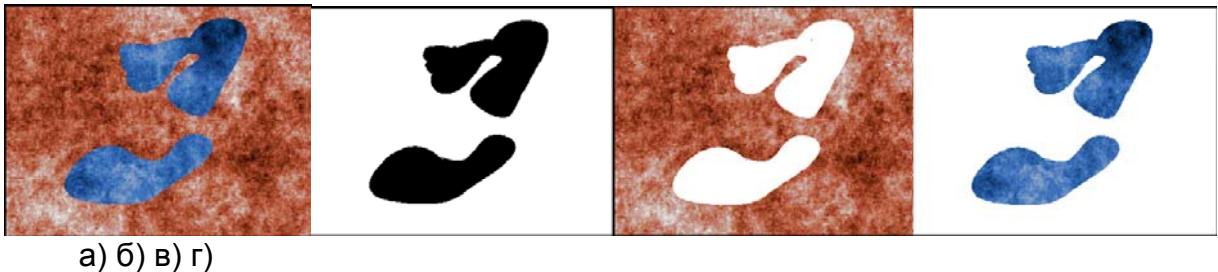
Таблица 9.4 Матрица межфрактальных расстояний при оценке размера фрактала по модулю разности яркостей (окно 16x16)

$H$	0,1	0,2	0,3	0,4	0,5	0,6	0,7	0,8	0,9
0,1	0	1,86	2,08	2,92	3,83	4,51	4,61	5,81	6,52
0,2		0	0,19	1,01	1,7	2,32	2,49	3,65	4,26
0,3			0	0,83	1,51	2,13	2,3	3,49	4,09
0,4				0	0,59	1,2	1,4	2,58	3,16
0,5					0	0,67	0,92	2,25	2,89
0,6						0	0,28	1,65	2,28
0,7							0	1,32	1,91
0,8								0	0,53
0,9									0

Осуществляется преобразование цветового координатного пространства  $RGB$  в пространство  $HLS$ . Оценка размера фрактала производится по яркостному компоненту изображения. Хотя размер фрактала инвариантен к масштабу, реальные природные поверхности не обладают одним и тем же размером фрактала во всем возможном диапазоне масштабов. Они имеют один и тот же размер фрактала для большего или меньшего диапазона шкал. Исследования показали, что для большинства приложений приходится ограничивать диапазон изменения шага [1,3].

Оценка межклusterных расстояний для фрактальных методов представлена в таблице 9.2-9.4. На рисунке 9.11 приведен пример наложения по маске (рисунок 9.11 б) двух синтезированных фрактальных текстур, имеющих разные цветовые характеристики. Изображения получены в соответствии с алгоритмом синтеза цветных фракталов, рассмотренным в разделе 9.4. При отличии по цвету текстуры могут быть успешно разделены на основании простого алгоритма: преобразование пространства  $RGB$  в пространство  $HLS$ , пороговое ограничение по гистограмме тонового компонента. На рисунке 9.11 в) и г) приведены текстуры, полученные в результате сегментации по компоненту тон. В данном случае мы получили точные маски областей текстур коричневого и синего цветов, вероятность ошибки равна 0. Таким образом, использование

цветовых характеристик текстуры в качестве дополнительного признака, обеспечивает эффективную классификацию.



а) б) в) г)

Рисунок 9.11 Пример сегментации фрактальных текстур, имеющих разные характеристики цветности. а) исходное изображение, б) изображение масок аппликативного наложения текстур, в) г) результат сегментации по тону.

Однако когда текстуры имеют одинаковые цветовые характеристики, необходимо включать в пространство признаков текстурные характеристики для различения текстур. Примеры фрактального анализа приведены на рисунках 9.12-9.14.



а) б) в)

Рисунок 9.12 Пример сегментации фрактальных текстур, имеющих одинаковые характеристики цветности. а) исходное изображение, б) результат сегментации при размере окна 16x16, в) результат сегментации при размере окна 8x8.

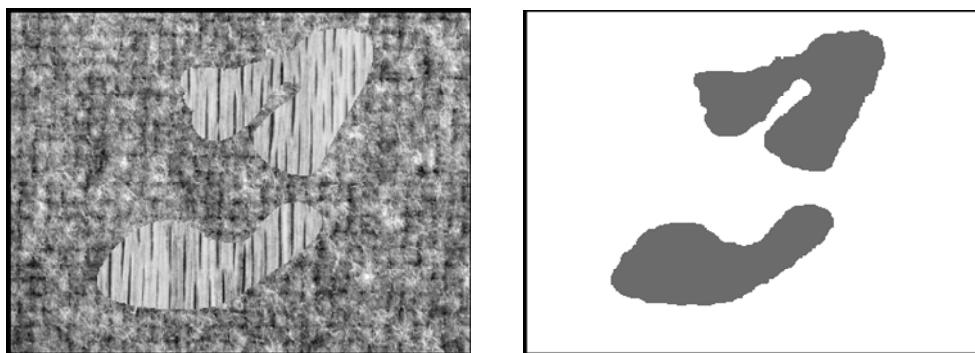
Анализ полученных результатов позволяет сделать вывод о большей эффективности метода броуновской функции по сравнению с методами оценки размера фрактала по длине контура и по методу пирамиды. На рисунке 9.12 представлены результаты сегментации фрактальных текстур по оценке размера фрактала по модулю разности яркостей отсчетов изображения для двух размеров окна 16x16 (рисунок 9.12 б) и 8x8 (рисунок 9.12 в).

На рисунке 9.13 а) показаны реальные текстуры из альбома Бродатца [106] и результат их сегментации (б). На рисунке 9.14 представлены результаты сегментации текстур, полученных с помощью гауссово - марковских полей, по оценке размера фрактала по модулю разности яркостей отсчетов изображения для размера окна 8x8 (рисунок 9.13 б). На основании выше изложенного можно сделать следующие выводы.

- 1) Представлены методы сегментации цветных текстур на основании объединения цветовых и фрактальных характеристик текстур при описании текстур в цветовом координатном пространстве *HLS*.

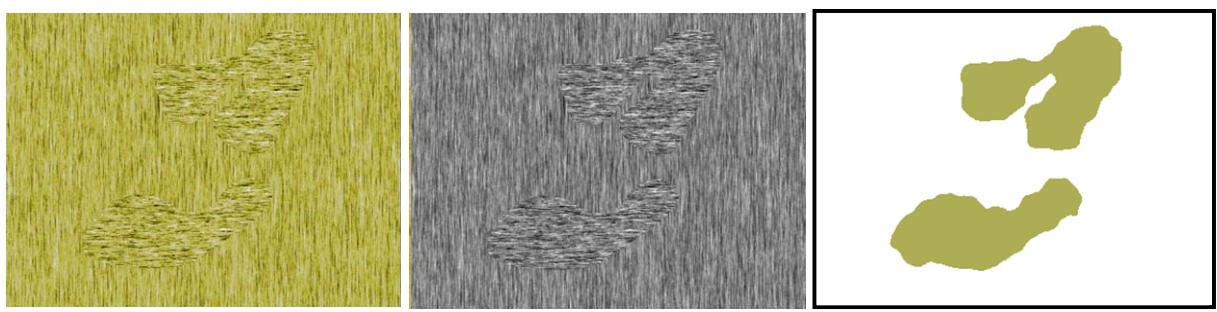
Исследования выполнялись на синтезированных цветных текстурах и на природных текстурах.

- 2) Были исследованы три метода сегментации текстур, построенных на оценке размера фрактала. Метод оценки размера фрактала по модулю разности яркостей отсчетов более эффективен, чем другие представленные методы. Оценку размера фрактала можно производить, используя изменения модуля разности яркостей элементов изображения в направлениях только по строкам, только по столбцам или только в диагональных направлениях для улучшения сегментации анизотропных текстур. Эффективность сегментации повышается благодаря низкочастотной фильтрации оценки размера фрактала.



а) б)

Рисунок 9.13 Пример сегментации текстур из альбома Бродатца а) исходное изображение, б) результат сегментации при размере окна 8x8.



а) б) в)

Рисунок 9.14 Пример сегментации текстур а) исходное изображение; б) исходное изображение яркостного компонента; в) результат сегментации при размере окна 8x8.

- 3) Использование признаков цветности и фрактальных характеристик позволяет уменьшить ошибки сегментации.  
4) Оценка размера фрактала вызывает дополнительные ошибки при вычислении тангенса угла наклона линии выборочной регрессии. Для сегментации можно использовать характеристики, по которым вычисляется размер фрактала, поскольку задача состоит в различении текстур, а не в оценке размера фрактала этих текстур.

## **ЗАКЛЮЧЕНИЕ**

В последние годы, когда цифровые системы все более активно заменяют аналоговые системы обработки изображений, очень важно владеть современными компьютерными методами описания и обработки изображений. Мы стремились дать понятное теоретическое обоснование и основы практического применения наиболее востребованных методов обработки изображений в компактной форме. Использование в пособии единой системы обозначений функций и переменных, позволяет упростить восприятие теоретических основ различных разделов, облегчает анализ многочисленных монографий, посвященных подробному описанию отдельных проблем и тем, поднимаемых в пособии. Большое количество примеров позволяет проиллюстрировать основные особенности применения алгоритмов и эффективность их использования. Надеемся, что пособие будет полезным студентам, аспирантам и специалистам при изучении компьютерных методов обработки изображений.

## БИБЛИОГРАФИЧЕСКИЙ СПИСОК

- 1 Мирошников М.М., Лисовский В.А., Филиппов Е.В. и др. Иконика в физиологии и медицине // Под ред. Уголова А.М. АН СССР. Отд-ние физиологии. Л.: Наука, 1987. 392 с.
- 2 Фисенко Т.Ю. и др. Аппаратно-программные средства МАГИСОФТ в задачах экспертизы живописи // 1 научная конференция "Экспертиза произведений изобразительного искусства". Тезисы докладов. М., 1995. С.18-20.
- 3 Мирошников М.М. Теоретические основы оптикоэлектронных приборов. Л.: Машиностроение, 1983. 696 с.
- 4 Дрязгов В.Д., Колобков В.С., Петров В.С., Фисенко В.Т. Способ получения стереоскопического цветного телевизионного изображения. Патент РФ на изобретение № 2189121, 2002 г.
- 5 Иванов В.А., Кованько В.В., Колобков В.С., Фисенко В.Т. Способ выполнения подводных подъемно-транспортных операций. Патент РФ на изобретение № 2203829, 2003 г.
- 6 Белоусов Ю.И., Можейко В.И., Силаев Н.А., Регинский В.Д., Фисенко В.Т. Некоторые аспекты построения корабельных оптико-электронных систем //Морская радиоэлектроника, вып. 2 (12), июнь 2005. С. 23-26.
- 7 Можейко В.И., Фисенко В.Т., Фисенко Т.Ю. Адаптивный метод автоматического сопровождения объектов по последовательности цифровых сигналов телевизионных изображений // Материалы 4-й международной конференции "Телевидение: передача и обработка изображений", С.-Пб., 2005 г. С. 91-92.
- 8 Можейко В.И., Фисенко В.Т., Фисенко Т.Ю. Автоматическое сопровождение объектов в компьютерных системах обработки изображений // Оптический журнал, том 74, номер 11, 2007. С. 39-46.
- 9 Апреликов В.В., Сущев Г.А., Фисенко В.Т. Основные тенденции в построении высокочувствительных ТВ камер // Научно-техническая конференция С-Пб Гос. Универс. Телекоммуникаций. С.-Пб., 1997. С. 81.
- 10 Тарасов В.В., Якушенков Ю.Г. Инфракрасные системы «смотрящего» типа. М.: Логос, 2004. 444 с.
- 11 Иванов В.П., Курт В.И., Овсянников В.А., Филиппов В.Л. Моделирование и оценка современных тепловизионных приборов. Казань, ФНПЦ НПО ГИПО, 2006. 594 с.
- 12 Прэтт У. Цифровая обработка изображений. М.: Мир, 1982. Т. 1, 2. 791 с.
- 13 Певзнер Б.М. Качество цветных телевизионных изображений. М.: Радио и связь, 1988. 2-е изд., доп. и перераб. 224 с.
- 14 Ричардсон Я. Видеокодирование. H.264 и MPEG-4 – стандарты нового поколения. М.: Техносфера, 2005. 368 с.

- 15 Марр Д. Зрение. Информационный подход к изучению представления и обработки зрительных образов. М., Радио и связь, 1987.
- 16 Джадд Д., Вышецки Г. Цвет в науке и технике. М., Мир, 1978. 592 с.
- 17 Lang H. 40 Jahre Farbfernsehen nach dem Prinzip der konstanten Luminans. // Ferseh- und Kino- Technik, 1995. 49 jahrgang, № 1-2. S. 35-40.
- 18 Быков Р.Е., Гуревич С.Б. Анализ и обработка цветных и объемных изображений. М.: Радио и связь, 1984.
- 19 Быков Р.Е. Основы телевидения и видеотехники: Учеб. для вузов. М.: Горячая линия - телеком, 2006. 399 с.
- 20 Цуккерман И.И. Теоретические основы цифрового телевидения. Учебное пособие. М.: МПСС. Институт повышения квалификации руководящих работников и специалистов, 1988. 69 с.
- 21 Полосин Л.Л. Цифровые системы вещательного телевидения. Учебное пособие // Балт. Гос.техн.ун-т. СПб., 2004. 120 с.
- 22 Strikland R. N., Kim Cheol-Sung, McDonnell William F. Luminance, hue, and saturation processing of digital color images. // SPIE. 1986. V. 697. Applications of Digital Image Processing IX . P. 286
- 23 Корн, Г., Корн Т. Справочник по математике для научных работников и инженеров. М.: Наука. Главная редакция физико-математической литературы, 1974. 832 с.
- 24 Полосин Л.Л. Принцип постоянной цветовой яркости в телевидении. // Телевизионная техника и связь. С-Пб, 1995. С.28-37.
- 25 Полосин Л.Л. Новый подход к определению метрического цветового пространства //Техника средств связи, сер. “Техника телевидения”, вып. 2, Л., 1991. С. 3-12.
- 26 Polosin L.L. Basic quantities of colour photometry//ECVP 2006. 29<sup>th</sup> European Conference on Visual Perception. St Peterburg, 2006, Abstracts, p.186.
- 27 Малла С. Вэйвлеты в обработке сигналов: Пер. с англ. М.: Мир, 2005. 671 с.
- 28 Гонсалес Р., Вудс Р. Цифровая обработка изображений: Пер. с англ./ Под ред. Чочиа П.А. М.: Техносфера, 2005. 1072 с.
- 29 Красильников Н.Н. Цифровая обработка изображений. М.:Вузовская книга, 2001. 320 с.
- 30 Горьян И.С., Межов Ф.Д., Фисенко В.Т. Введение в цифровую обработку изображений. Учебное пособие. СПб.: ЭИС им. М.Бонч-Бруевича, 1992. 60 с.
- 31 Горелик С.Л., Кац Б.М., Киврин В.И. Телевизионные измерительные системы. М.: Связь, 1980. 168 с.
- 32 Цуккерман И.И., Кац Б.М.,Лебедев Д.С. и др. Цифровое кодирование телевизионных изображений.// Под ред. Цуккермана И.И.М.: Радио и связь, 1981. 240 с.
- 33 Фоли Дж., вэн Дэм А. Основы интерактивной машинной графики: В 2-х книгах: Пер. с англ. М.: Мир, 1985. 368 с.

- 34 Фисенко Т.Ю. и др. Интерполяционные алгоритмы определения положения центра изображения объекта с помощью ПЗС //Техника средств связи, сер. “Техника телевидения”, вып. 4. Л., 1986. С. 37-47.
- 35 Fissenko T.Y., Fisssenko V.T., Rogatchev K.S., Sushchev G.A. An Interpolation Method for Color Image Reconstruction in Deep UnderWater Observation // Proceedings of D.S. Rozhdestvensky Optical Society. Proceedings of the International Conference “Current Problems in Optics of Natural Waters” (ONW’2001). St. Peterburg. Russia, 2001. September 25-29. P. 113-118.
- 36 Кац Б.М. Фисенко Т.Ю. Синтез промежуточных фаз по контурному мультиплексу //Методы и средства обработки сложной графической информации. Тезисы докладов IV Всесоюзной конференции. Нижний Новгород, 1991. С.34.
- 37 Фисенко В.Т., Фисенко Т.Ю. и др. Особенности построения ТВ камеры для наблюдения сцен под водой // 6 Международная конференция. Труды конференции. Прикладные технологии гидроакустики и гидрофизики. С.-Пб., Россия, 28-31 мая 2002. С. 200 - 204.
- 38 Фисенко Т.Ю. Улучшение цифровых цветных изображений микрообъектов // XLIX научная сессия, посвященная дню радио. Тезисы докладов. М., 1994. Часть 2. С. 86-87.
- 39 Тимофеев Б.С. Цифровое телевидение. Учебное пособие//СПбГУАП. С.Пб, 1998. 49 с.
- 40 Фисенко В.Т., Валышков В.И. Сравнение двумерных операторов для выделения контуров //Техника средств связи, сер. “Техника телевидения”, вып. 6. Л., 1979. С. 47-54.
- 41 Аристархов Ю.Н., Валышков В.И., Горьян И.С., Фисенко В.Т. Определение эффективности работы контурных операторов //Техника средств связи, сер. “Техника телевидения”, вып.1, Л., 1985. с.
- 42 Шапиро Л. Компьютерное зрение/ Шапиро Л., Стокман Дж.: Пер. с англ. М.: БИНОМ. Лаборатория знаний, 2006. 752 с.
- 43 Виноградов А.Ю., Горьян И.С., Фисенко В.Т., Шостацкий Н.Н. Методические указания к лабораторным работам по цифровым преобразованиям в телевидении. ЛЭИС. Л., 1988. 36 с.
- 44 Lukac R., Smolka B: Application of the Adaptive Center-Weighted Vector Median Framework for the Enhancement of cDNA Microarray Images // Int. J. Appl. Math. Comput. Sci., 2003. V. 13, N3. P 369-383.
- 45 Можейко В.И., Фисенко В.Т., Фисенко Т.Ю. Улучшение цифровых телевизионных цветных изображений по методу многоканальной медианной фильтрации // Материалы международной конференции «Телевидение: передача и обработка изображений», С.-Пб., Россия, 2000. С. 22-24.
- 46 Методы компьютерной обработки изображений// Под редакцией Сойфера В.А., М.: Физматлит, 2001. 784 с.

- 47 Павлидис Т. Алгоритмы машинной графики и обработки изображений: Пер. с англ. М.: Радио и связь, 1986. 400 с.
- 48 Фисенко Т.Ю. и др. Исследование устойчивых радиационных дефектов в ПЗС-матрице, облучаемой высокоэнергетическими частицами // Вопросы атомной науки и техники. Серия: Физика радиационного воздействия на радиоэлектронную аппаратуру. Научно-технический сборник, М., 2000. Вып. 3-4. С.8-11.
- 49 Авдеев В.П., Горьян И.С., Фисенко В.Т. Сокращенное описание некоторых текстурных характеристик изображений //Техника средств связи, сер. “Техника телевидения”, вып. 4. Л., 1980. С.33
- 50 Лебедев Д.С., Цуккерман И.И. Телевидение и теория информации. М.-Л.: Энергия, 1965.
- 51 Бондаренко В.М., Данилова М.В., Красильников Н.Н., Леушина Л.И., Невская А.А., Шелепин Ю.Е. Пространственное зрение. СПб.: Наука, 1999. 218 с.
- 52 Брайс Р. Справочник по цифровому телевидению: Пер. с англ. Жуковский: Эра, 2001. 225 с.
- 53 Климов А.С. Форматы графических файлов. К.: НИПФ «ДиаСофт Лтд», 1995. 480 с.
- 54 Ватолин Д., Ратушняк А., Смирнов М., Юкин В. Методы сжатия данных. Устройство архиваторов, сжатие изображений и видео. М.: ДИАЛОГ-МИФИ, 2002. 384 с.
- 55 Huffman D.A. A Method for the Construction of Minimum-Redundancy Codes // Proceedings of the I.R.E. 1952. P. 1098-1101.
- 56 Дворкович А.В. и др. Цифровая обработка телевизионных и компьютерных изображений / Под редакцией Зубарева Ю.Б. и Дворковича В.П. М.: Международный центр научной и технической информации, 1997. 212 с.
- 57 Воробьев В.И., Грибунин В.Г. Теория и практика вейвлет-преобразования. СПб.: ВУС, 1999. 208 с.
- 58 Чуи К. Введение в вэйвлеты: Пер. с англ. М.: Мир, 2001. 412 с.
- 59 ADV601
- 60 Смоленцев Н.К. Основы теории вейвлетов. Вейвлеты в MATLAB. М.: ДМК Пресс, 2005. 304 с.
- 61 Добеши И. Десять лекций по вейвлетам. Пер. с англ. – Ижевск: НИЦ «Регулярная и хаотическая динамика», 2001. 464 с.
- 62 Харалик Р.М. Статистический и структурный подходы к описанию текстур // ТИИЭР, 1979. Т. 67, № 5. С. 98.
- 63 Feder E. Фракталы. Учебное издание. М., Мир, 1991. 259 с.
- 64 Rosenfeld A. Automatic recognition of basic terrain types on aerial photographs // Photogrammetric Eng., 1962. V. 28. P. 115.
- 65 Горьян И.С., Януш Д.А. К вопросу об автоматическом дешифрировании аэроснимков // В сб.: Исследование оптических свойств природных объектов и их аэрофотографического изображения. Л: Наука, 1970.

- 66 Латышева В.И. и др. Использование статистических характеристик телевизионного сигнала для определения количества и форм облаков на снимках, получаемых с метеорологических спутников // Труды ГМЦ, 1966, вып. II.
- 67 Горьян И.С., Кац Б.М., Цуккерман И.И. Исследование статистики изображений с помощью цифровых вычислительных машин // Вопросы радиоэлектроники, сер «Техника телевидения», 1971. Вып. 4. С. 3.
- 68 Rosenfeld A., Thurston M. Edge and curve detection for visual scene analysis // IEEE Trans., 1971. V. C-20, № 5. P. 562.
- 69 Haralick R.M. and oth. Textural features for image classification // IEEE Trans., 1973. V. SMC-3. P. 610.
- 70 Julesz. B. Experiments in the visual perception of texture // Sci. Amer. 232, 1975. № 4. P.2.
- 71 Caelli T. and Julesz B. On perceptual analysis underlying visual texture discrimination // Parts 1,2, Biol. Cybern. 1978. 28. P.167. 1978. 29. P.201.
- 72 Gagalowicz A. Visual discrimination of stochastic texture fields based on their second order statistics // Proceedings 5-th Int'l. Conf. on Pattern Recognt., Florida Beach, 1980. V. 1. P.768.
- 73 Gotlieb C. and Kreyszig H. Texture descriptors based on Co-occurrence Matrices //Comput. Vision Graphics Image Processing, 1990. V. 51, №1. P. 70.
- 74 Горьян И.С., Цуккерман И.И. Распознавание изображений по их статистическим свойствам //Техника кино и телевидения, 1966. № 2. С. 57.
- 75 Горьян И.С., Фисенко В.Т., Януш Д.А. Анализ неоднородных текстурных изображений // 1 Всесоюзная конференция «Статистические свойства микроструктур». Тезисы докладов. М. 1978. С. 53-55.
- 76 Hayes K.C. Jr., Shah A.N., Rosenfeld A. Texture coarseness: further experiments // IEEE Trans. Systems, Man, and Cybernetics, 1974. SMC-4,5. P.467-472.
- 77 Горьян И.С., Фисенко В.Т. Классификация изображений самообучающимся автоматом //Техника средств связи, сер. “Техника телевидения”. Вып.4. Л., 1978. С. 72-78.
- 78 Горьян И.С., Зеленцов В.Н., Фисенко В.Т. Рекурсивный логический фильтр для обнаружения протяженных связных объектов //Техника средств связи, сер «Техника телевидения», 1987. Вып. 1. С. 48.
- 79 Weszka J. and oth. A comparative study of texture measures for terrain classification //IEEE Trans., 1976. V. SMC-26, № 4. P. 269.
- 80 Авдеев В.П., Горьян И.С., Фисенко В.Т. Статистические свойства изображений и адаптивные методы выделения признаков объектов // Тезисы докладов Всесоюзного симпозиума «Проблемы цифрового кодирования и преобразования изображений». Тбилиси, 1980.

- 81 Yhann S.R., Young T.Y. Boundary Localization in Texture Segmentation// IEEE –T-IP, V.4, №6, 1995. P. 849-856.
- 82 Горьян И.С., Зеленцов В.Н., Фисенко В.Т. Выделение признаков для систем структурного анализа рисунков в реальном масштабе времени // Третья Всесоюзная конференция «Математические методы распознавания образов». Тезисы докладов. Львов, 1987. С. 147-148.
- 83 Фу К. Структурные методы в распознавании образов. М.: Мир, 1977.
- 84 Ту Дж., Гонсалес Р. Принципы распознавания образов. М.: Мир, 1978. 412 с.
- 85 Shaw A.C. Parsing of graph-representable pictures // J. ACM, 17, 3. P. 453.
- 86 Ledley R.S. High-speed automatic analysis of biomedical pictures // Science, 146, 3641, 1964. P. 216.
- 87 Lu S.Y., Fu K.S. A syntactic approach to texture analysis // Comput. Graph. Image Proc., 1978. V. 7, № 3. P. 303.
- 88 Ehrich R.W., Foith J.P. A view of texture topology and texture description // Comput. Graph. Image Proc., 1978. V. 8, № 2. P. 174.
- 89 Mandelbrot B. The fractal geometry of Nature. W.H. Freeman and Co., NY., 1982.
- 90 Lovejoy S. Area-perimeter relation for rain and cloud areas // Science, 216, 1982. P. 185.
- 91 Под редакцией Пьетронеро Л. и Тозитти Э. Фракталы в физике. М.: Мир, 1988. 644 с.
- 92 Kalle Karu, Anil K. Jain, Rund M.Bolle. Is there any texture in the image? // Pattern Recognition, 1996. V. 29, № 9, Sept., P. 1437-1446.
- 93 Ohlander R., Price K. and Reoldy D.R. Picture segmentation using a recursive region splitting method.//Comput. Graphics Image Process.- 1978.-Vol.8.-P.313.
- 94 Mehment Celenk. A color clustering technique for image segmentation.// Computer Vision, Graphics, and Image Processing, 1990. V. 52. № 2. P. 145.
- 95 Masson R.C., Bottcher P. and Leisinger V. Real-time grey-level and color image pre-processing for a vision guided biotechnology robot. // Technol. Proc. 7-th Int. Conf. Rob. Vision and Sens Contr. Zurich
- 96 Uchiyama Toshio, Arbib Michael.A. Color image segmentation using competitive learning // IEEE T-PAMI, 1994. V. 16. № 12. P.1197-1206.
- 97 Chellappa R., Chatterjee S., Bagdazian R. Texture synthesis and compression using Gaussian-Markov random field models // IEEE Trans. Systems, Man, and Cybernetics, 1985. V. 15, № 2. P. 298-303.
- 98 Spath H. Cluster analysis Algorithms for data reduction and classification of objects. Chichester, England, 1980. - P.
- 99 Фисенко В.Т., Фисенко Т.Ю. Метод автоматического анализа цветных изображений // Оптический журнал, том 70, номер 9, сентябрь 2003. С.18-23.

- 100 Горьян И.С., Фисенко В.Т. Автоматический выбор порогов квантования при сегментации изображений // Четвертая Всесоюзная конференция «Математические методы распознавания образов». Тезисы докладов. Рига, 1989. С. 62-64.
- 101 Lam N.S-N. Description and Measurement of Landsat TM Images Using Fractals // Photogrammetric Engineering and Remote Sensing, 1990. № 2, P. 187-195.
- 102 Мак-Кракен Д., Дорн У. Численные методы и программирование на Фортране. М.: Мир, 1977. 584 с.
- 103 Королюк В.С., Портенко Н.И., Скороход Ф.И., Турбин А.Ф. Справочник по теории вероятностей и математической статистике. М.: Наука. Главная редакция физико-математической литературы, 1985. 640 с.
- 104 Jaggi S., Quattrochi D., Lam N. Multiresolution processing for fractal analysis of airborne remotely sensed data // Int. Conf. “AeroSpace 92”, USA, Orlando, April 1992.
- 105 Yokoya N., Yamamoto K., Funakubo N. Fractal-Based Analysis and Interpolation of 3D Natural Surface Shapes and Their Application to Terrain Modeling // Comput. Vision Graphics Image Process., 1989. V. 46, P. 284-302
- 106 Brodatz P. A Photographic Album for Artists and Designers. New York, Dover, 1966.

## **ПЕРЕЧЕНЬ ИЛЛЮСТРАЦИЙ**

- 1 Рисунок 2.1 Определение пространственных частот изображения.12
- 2 Рисунок 3.1 Оси противоположных цветов в соответствии с оппонентной теорией Геринга.18
- 3 Рисунок 3.2 Механизм цветного зрения в соответствии с зонной теорией Адамса.19
- 4 Рисунок 3.3 Графики зависимости видимой насыщенности от размеров деталей и их цветов.21
- 5 Рисунок 3.4 Элемент в цветовом координатном пространстве  $L^*a^*b^*$ .  
23
- 6 Рисунок 3.5 Цветовое координатное пространство  $RGB$ .24
- 7 Рисунок 3.6 Цветовое координатное пространство  $HLS$ .25
- 8 Рисунок 3.7 Представление цветов в метрическом векторном пространстве.31
- 9 Рисунок 4.1 Функция пространственной дискретизации изображения.  
33
- 10 Рисунок 4.2 а) Изменение спектра по строке для а) непрерывного сигнала, б) дискретизованного сигнала.35
- 11 Рисунок 4.3 Выбор частоты дискретизации в соответствии с теоремой отсчетов.36
- 12 Рисунок 4.4 Частотная характеристика идеального восстанавливающего прямоугольного фильтра.37
- 13 Рисунок 4.5 Амплитудная характеристика оптимального квантователя Макса.41
- 14 Рисунок 4.6 График зависимости вероятности квантования от отношения интервала квантования к СКО аддитивного шума  $\beta$ .42
- 15 Рисунок 4.7 Идеальная НЧ фильтрация сигнала в полосе, равной полосе частотного спектра сигнала: а) спектр непрерывного сигнала; б) спектр сигнала после дискретизации; в) частотная характеристика идеального фильтра; г) спектр сигнала на выходе фильтра. Идеальная НЧ фильтрация сигнала в полосе, меньшей полосы частотного спектра сигнала: д) спектр непрерывного сигнала; е) спектр сигнала после дискретизации; ж) частотная характеристика идеального фильтра; з) спектр сигнала на выходе фильтра.44
- 16 Рисунок 4.8 Равномерное квантование: а) число разрядов квантования  $L=8$ , б) число разрядов квантования  $L=4$ .45

- 17 Рисунок 4.9 а) Изображение получено равномерным квантованием на 16 уровней изображения «Лена»; б) изображение получено наложением нормального шума при СКО=5 на исходное изображение и равномерным квантованием на 16 уровней.45
- 18 Рисунок 4.10 Импульсная характеристика идеального восстанавливающего фильтра по  $x$ -координате.46
- 19 Рисунок 5.1 а) Исходное изображение; б) автоматическое контрастирование (при квантилях слева и справа = 0,005).50
- 20 Рисунок 5.2 Амплитудная передаточная характеристика преобразования при выделении диапазона значений сигнала изображения.51
- 21 Рисунок 5.3 а) Изображение негатива яркостного сигнала изображения; б) изображение сигнала насыщенности этого же изображения.52
- 22 Рисунок 5.4 Пример гамма-коррекции. а) исходное изображение, б) преобразованное.52
- 23 Рисунок 5.5 Пример эквалайзации гистограммы изображения. а) Исходное изображение и его гистограмма в); б) преобразованное изображение и его гистограмма г).55
- 24 Рисунок 5.6 Амплитудные характеристики для бинаризации изображений. а), б) квантование по порогу прямое и инверсное; в) формирование изофот, г) формирование инверсных изофот.56
- 25 Рисунок 5.7 Формирование изофот изображения. Бинаризация выполнена по компоненту тона.57
- 26 Рисунок 5.8 Тестовое изображение. а) Изображение RGB компонентов; б) гистограмма распределения RGB компонентов тестового изображения.59
- 27 Рисунок 5.9 Гистограмма распределения ошибок преобразования цветового координатного пространства  $RGB$  в пространство  $HLS$ : а) для целочисленного преобразования Оствальда; б) для цветовой координатной системы Стрикланда.59
- 28 Рисунок 5.10 Изображение фрагмента карты: а) исходное изображение в пространстве  $RGB$ , б) исходное изображение после повышения контраста, насыщенности и подчеркивания верхних частот.60
- 29 Рисунок 5.11 а) Исходное изображение без накопления, б) исходное изображение с накоплением 0,5 с; в) улучшенное изображение а; г) улучшенное изображение б).61
- 30 Рисунок 6.1 Конфигурации окрестности элемента «0» в кадре изображения в иерархической последовательности.62

- 31 Рисунок 6.2 Положение отсчетов импульсной характеристики при свертке с изображением  $f(x,y)$ .65
- 32 Рисунок 6.3 Корреляция изображения  $f(x,y)$  с маской  $h(s,t)$ .65
- 33 Рисунок 6.4 Применение медианного фильтра с окном  $3 \times 3$  элемента.  
а) исходное изображение; б) выходное изображение; в) изображение разности изображений а) и б).69
- 34 Рисунок 6.5 Пример фильтрации изображения с импульсной помехой (1%): а) исходное изображение; б) МФ  $3 \times 3$ ; в) АМФ.71
- 35 Рисунок 6.6 График частотной характеристики фильтра Баттервортса при  $r=70$  и  $k=10$ : а) – центрированная характеристика; б) – нецентрированная.74
- 36 Рисунок 7.1 Пример операций анализа цветного изображения. а) Исходное изображение. Каждая клетка является отдельным объектом.  
б) Бинарное изображение клеток оранжевого цвета после сегментации. в) Результат селекции связных компонентов.75
- 37 Рисунок 7.2 а) Изображение цветка; б) цветоразностный ( $Cb$ ) компонент этого изображения; в) бинарное изображение цветка.76
- 38 Рисунок 7.3 Гистограмма значений  $Cb$  компонента.77
- 39 Рисунок 7.4 Оценка яркостных и геометрических признаков формы клеток.78
- 40 Рисунок 7.5 Параллельный перенос множества  $\mathbf{B}$  на вектор  $t = (x_0, y_0)$ .79
- 41 Рисунок 7.6 Центральное отражение множества  $\mathbf{B}$ .79
- 42 Рисунок 7.7 Дилатация множества  $\mathbf{B}$  по множеству  $\mathbf{S}$ .80
- 43 Рисунок 7.8 Эрозия множества  $\mathbf{B}$  по примитиву  $\mathbf{S}$ .81
- 44 Рисунок 7.9 Размыкание множества  $\mathbf{B}$  по примитиву  $\mathbf{S}$ .81
- 45 Рисунок 7.10 Замыкание множества  $\mathbf{B}$  по примитиву  $\mathbf{S}$ .82
- 46 Рисунок 7.11 Определение ориентации по собственному вектору матрицы ковариации.85
- 47 Рисунок 8.1 Чувствительность к контрасту и порог неразличимости  $\Delta I$ .88
- 48 Рисунок 8.2 Эффект полос Maxa.89
- 49 Рисунок 8.3 Блок-схема алгоритма кодирования длин серий строки изображения.93
- 50 Рисунок 8.4 Блок-схема алгоритма декодирования строки изображения по методу *RLE*.94
- 51 Рисунок 8.5 Формирование кодового дерева по методу Хаффмана.100

- 52 Рисунок 8.6 а) Изображение, полученное дифференцированием по строке изображения «Лена»; б) гистограмма распределения яркостей этого изображения.101
- 53 Рисунок 8.7 Структурная схема кодера ДИКМ.103
- 54 Рисунок 8.8 Исходное изображение а) и два восстановленных изображения б) и в). Спектры соответствующих изображений г) - е). 106
- 55 Рисунок 8.9 Базисные функции ДКП,  $n$ -порядковый номер отсчета входной последовательности,  $k$ -номер выходного значения коэффициента ДКП.109
- 56 Рисунок 8.10 Изображение, восстановленное по 3,8% коэффициентов: а) по алгоритму ОДПФ; б) по алгоритму ОДКП.110
- 57 Рисунок 8.11 Схема кодера JPEG.110
- 58 Рисунок 8.12 Положения отсчетов яркости и цветности при формировании макроблока. **X**- положение отсчета яркости, **O** – совпадающее положение отсчетов цветности  $Cb$  и  $Cr$ .111
- 59 Рисунок 8.13 Расположение блоков в макроблоке при кодировании JPEG.112
- 60 Рисунок 8.14 Зигзагообразное сканирование квантованных спектральных коэффициентов.114
- 61 Рисунок 8.15 Условная схема процесса кодирования / декодирования при вейвлетном преобразовании.116
- 62 Рисунок 8.16 Схема вейвлет-кодирования изображения.118
- 63 Рисунок 8.17 Пространственное расположение блоков коэффициентов в кадре изображения после вейвлетного преобразования.119
- 64 Рисунок 8.18 Вейвлет Хаара. а) Масштабирующая функция («отцовский вейвлет»), б) материнский вейвлет, в) производные вейвлеты Хаара.121
- 65 Рисунок 8.19 Положение коэффициентов фильтров и отсчетов сигнала входной последовательности.123
- 66 Рисунок 8.20 Схема выполнения одномерного ВП и децимации.124
- 67 Рисунок 8.21 Схема выполнения одномерного ОДВП при использовании фильтра  $D4$ .125
- 68 Рисунок 8.22 Схема прямого и обратного одномерного ВП.125
- 69 Рисунок 8.23 Фрагменты изображения "Лена". а) Фрагмент исходного изображения; б) Фрагмент декодированного JPEG изображения; в) фрагмент декодированного изображения при использовании вейвлетного алгоритма.129

- 70 Рисунок 8.24 Исходные изображения: а) “Лена”, б) тестовая таблица. в) График зависимости ПСШ от коэффициента сжатия, где  $J$  - кодирование JPEG;  $W$  - вейвлетное кодирование;  $L$  - изображение “Лена”;  $T$  - изображение таблицы.130
- 71 Рисунок 8.25 Схема формирования предсказания движения в группе в хронологической последовательности кадров.132
- 72 Рисунок 9.1 Определение размера фрактала на примере четырех ломаных линий.141
- 73 Рисунок 9.2 Граф выделения на изображении текстурных областей. 144
- 74 Рисунок 9.3 Форма кластеров при кластеризации двумерных данных а) по методу квантования многомерных гистограмм, б) по методу векторного квантования.146
- 75 Рисунок 9.4 Примеры синтезированных текстур по статистическому методу.153
- 76 Рисунок 9.5 Пример синтезированных текстур при  $H$  равном 0,1; 0,5 и 0,8 соответственно.155
- 77 Рисунок 9.6 График зависимости нормированной величины внутрикластерных ошибок  $D/D_{max}$  от числа итераций  $ntrace$ .159
- 78 Рисунок 9.7 Пример классификации изображений. Кластеры представлены центрами масс. а) Исходное изображение; в) изображение получено после кластеризации по методу  $K$ -внутригрупповых средних при числе итераций  $ntrace=1$ ; г) изображение получено после кластеризации по методу  $K$  – внутригрупповых средних при числе итераций  $ntrace=10$ ; б) изображение, сформированное как разность изображений в) и г).160
- 79 Рисунок 9.8 Пример классификации: а) исходное изображение; б) изображение, представленное центрами масс кластеров после первоначального разбиения по тоновому компоненту; в) изображение, представленное центрами масс кластеров, при кластеризации по методу  $K$ -внутригрупповых средних после выполнения 1 итерации; г) изображение, представленное центрами масс кластеров, после доопределения кластеров и кластеризации по методу  $K$  – внутригрупповых средних.161
- 80 Рисунок 9.9 Маски красно-коричневого кластера, полученные на различных шагах.162
- 81 Рисунок 9.10 Построение пирамиды при оценке размера фрактала по методу пирамиды.168

- 82 Рисунок 9.11 Пример сегментации фрактальных текстур, имеющих разные характеристики цветности. а) исходное изображение, б) изображение масок аппликативного наложения текстур, в) г) результат сегментации по тону.<sup>174</sup>
- 83 Рисунок 9.12 Пример сегментации фрактальных текстур, имеющих одинаковые характеристики цветности. а) исходное изображение, б) результат сегментации при размере окна 16x16, в) результат сегментации при размере окна 8x8.<sup>174</sup>
- 84 Рисунок 9.13 Пример сегментации текстур из альбома Бродатца а) исходное изображение, б) результат сегментации при размере окна 8x8.<sup>175</sup>
- 85 Рисунок 9.14 Пример сегментации текстур а) исходное изображение; б) исходное изображение яркостного компонента; в) результат сегментации при размере окна 8x8.<sup>175</sup>

## ОГЛАВЛЕНИЕ

<b>СПИСОК ИСПОЛЬЗОВАННЫХ СОКРАЩЕНИЙ .....</b>	<b>3</b>
<b>ПРЕДИСЛОВИЕ .....</b>	<b>5</b>
<b>1 ПРЕДМЕТ И ЗАДАЧИ КОМПЬЮТЕРНОЙ ОБРАБОТКИ И РАСПОЗНАВАНИЯ ИЗОБРАЖЕНИЙ .....</b>	<b>7</b>
1.1     Определение компьютерной обработки изображений	7
1.2     Устройства формирования изображений	8
<b>2 МАТЕМАТИЧЕСКИЕ МОДЕЛИ ИЗОБРАЖЕНИЙ .....</b>	<b>11</b>
2.1     Модели непрерывных изображений	11
2.2     Пространственные спектры изображений	11
2.3     Спектральные интенсивности изображений	12
2.4     Вероятностные модели изображений и функции автокорреляции	13
2.5     Критерии качества изображений	14
<b>3 ЦВЕТ КАК ВАЖНАЯ ХАРАКТЕРИСТИКА ВОСПРИЯТИЯ ОБЪЕКТА .....</b>	<b>16</b>
3.1     Цветовая модель <i>RGB</i>	19
3.2     Цветовая модель <i>YCrCb</i>	19
3.3     Цветовая модель <i>YIQ</i>	21
3.4     Цветовая модель <i>L*a*b*</i> МКО 1976	21
3.5     Цветовая модель <i>L*H°C*</i>	22
3.6     Цветовая модель <i>HSI</i>	23
3.7     Цветовая модель <i>HLS</i>	24
3.8     Цветовая модель <i>L*u*v*</i> МКО 1973	29
3.9     Метрическое векторное цветовое пространство	30
<b>4 ПРЕДСТАВЛЕНИЕ ИЗОБРАЖЕНИЙ В ЦИФРОВОЙ ФОРМЕ</b>	<b>31</b>
4.1     Дискретизация изображений	32
4.2     Квантование изображений	37
4.2.1     Оптимальное квантование .....	38
4.2.2     Квантование сигнала при наличии шумов .....	40
4.3     Погрешности дискретного представления изображений	42
4.3.1     Погрешность дискретизации .....	42
4.3.2     Погрешность квантования.....	43
4.4     Методы интерполяции изображений	45
<b>5 УЛУЧШЕНИЕ ИЗОБРАЖЕНИЙ .....</b>	<b>48</b>
5.1     Яркостные преобразования изображений	49
5.1.1     Линейное контрастирование .....	49
5.1.2     Выделение диапазона яркостей .....	51
5.1.3     Преобразование изображения в негатив.....	51
5.1.4     Степенные преобразования.....	52
5.1.5     Логарифмическое преобразование .....	53
5.2     Преобразование гистограмм	53
5.3     Пороговая обработка полутоновых изображений	56

<b>5.4</b>	<b>Применение табличного метода при поэлементных преобразованиях изображений</b>	<b>57</b>
<b>5.5</b>	<b>Особенности применения поэлементной обработки цветных изображений</b>	<b>58</b>
<b>6</b>	<b>ОСНОВЫ ФИЛЬТРАЦИИ ИЗОБРАЖЕНИЙ.....</b>	<b>62</b>
<b>6.1</b>	<b>Основы пространственной фильтрации изображений</b>	<b>64</b>
6.1.1	Линейная пространственная фильтрация .....	64
6.1.2	Нелинейная пространственная фильтрация .....	67
6.1.3	Фильтры повышения верхних пространственных частот изображения.....	71
<b>6.2</b>	<b>Частотная фильтрация изображений</b>	<b>73</b>
6.2.1	Дискретное преобразование Фурье .....	73
6.2.2	Теорема о свертке.....	73
6.2.3	Низкочастотные фильтры .....	74
6.2.4	Высокочастотные фильтры .....	75
<b>7</b>	<b>АНАЛИЗ БИНАРНЫХ ИЗОБРАЖЕНИЙ.....</b>	<b>75</b>
<b>7.1</b>	<b>Назначение основных операций бинарного анализа.</b>	<b>76</b>
<b>7.2</b>	<b>Состав основных операций бинарного анализа</b>	<b>76</b>
<b>7.3</b>	<b>Морфология бинарных изображений</b>	<b>78</b>
<b>7.4</b>	<b>Признаки областей</b>	<b>82</b>
<b>8</b>	<b>МЕТОДЫ КОДИРОВАНИЯ ЦИФРОВЫХ ИЗОБРАЖЕНИЙ .</b>	<b>85</b>
<b>8.1</b>	<b>Особенности зрительной системы человека</b>	<b>86</b>
8.1.1	Контрастная чувствительность.....	88
8.1.2	Пространственная чувствительность частоты .....	88
8.1.3	Маскирование .....	89
8.1.4	Статистическая и визуальная избыточность изображений ..	89
<b>8.2</b>	<b>Кодирование длин серий</b>	<b>91</b>
<b>8.3</b>	<b>Кодирование по методу <i>LZW</i></b>	<b>95</b>
8.3.1	Кодирование .....	95
8.3.2	Декодирование .....	97
<b>8.4</b>	<b>Метод кодирования Хаффмана</b>	<b>98</b>
<b>8.5</b>	<b>Принцип дифференциальной импульсно - кодовой модуляции</b>	<b>100</b>
<b>8.6</b>	<b>Кодирование с использованием ортогональных преобразований</b>	<b>104</b>
8.6.1	Дискретное косинусное преобразование.....	107
<b>8.7</b>	<b>Сжатие данных по методу <i>JPEG</i></b>	<b>109</b>
<b>8.8</b>	<b>Кодирование на основе вейвлетных преобразований</b>	<b>115</b>
<b>8.9</b>	<b>Стандарт <i>JPEG 2000</i></b>	<b>130</b>
<b>8.10</b>	<b>Стандарты сжатия видео последовательностей</b>	<b>130</b>
8.10.1	Стандарт <i>MPEG-1</i> .....	131
8.10.2	Стандарт <i>MPEG-2</i> .....	134
8.10.3	Стандарт <i>MPEG-4</i> .....	134
<b>9</b>	<b>АНАЛИЗ ТЕКСТУРНЫХ ИЗОБРАЖЕНИЙ.....</b>	<b>135</b>
<b>9.1</b>	<b>Описание текстур</b>	<b>135</b>

9.1.1	Статистический подход к описанию текстур.....	136
9.1.2	Структурный подход к описанию текстур .....	139
9.1.3	Фрактальный подход к описанию текстур .....	140
<b>9.2</b>	<b>Оценка наличия текстуры в изображении</b>	<b>143</b>
<b>9.3</b>	<b>Методы сегментации цветных текстур в цветовых координатных пространствах</b>	<b>144</b>
<b>9.4</b>	<b>Синтез цветных пространственных текстур</b>	<b>149</b>
9.4.1	Статистический метод синтеза цветных текстур.....	149
9.4.2	Синтез цветных фрактальных текстур.....	153
<b>9.5</b>	<b>Сегментация цветных текстурных изображений</b>	<b>155</b>
<b>9.6</b>	<b>Фрактальный анализ сложных текстурных изображений</b>	<b>164</b>
9.6.1	Оценка фрактальности признаков цветных текстур .....	164
9.6.2	Возможности и ограничения применения алгоритма оценки размера фрактала по длине контура при анализе сложных текстурных изображений .....	165
9.6.3	Сегментация текстурных изображений по методу треугольной пирамиды .....	168
9.6.4	Оценка размера фрактала по модулю разности яркостей отсчетов .....	170
<b>ЗАКЛЮЧЕНИЕ</b>		<b>176</b>
<b>БИБЛИОГРАФИЧЕСКИЙ СПИСОК</b>		<b>177</b>
<b>ПЕРЕЧЕНЬ ИЛЛЮСТРАЦИЙ</b>		<b>184</b>
<b>ОГЛАВЛЕНИЕ</b>		<b>190</b>
<b>КАФЕДРА КОМПЬЮТЕРНОЙ ФОТОНИКИ</b>		<b>193</b>

В 2007 году СПбГУ ИТМО стал победителем конкурса инновационных образовательных программ вузов России на 2007–2008 годы. Реализация инновационной образовательной программы «Инновационная система подготовки специалистов нового поколения в области информационных и оптических технологий» позволит выйти на качественно новый уровень подготовки выпускников и удовлетворить возрастающий спрос на специалистов в информационной, оптической и других высокотехнологичных отраслях экономики.

---

## КАФЕДРА КОМПЬЮТЕРНОЙ ФОТОНИКИ

Достижения в оптической науке, технике и технологиях за последние годы способствовали появлению нового направления - фотоники. Этот термин охватывает область науки и техники, связанную с использованием светового излучения (или потока фотонов) в оптических элементах, устройствах и системах.

На рубеже ХХ - ХХI веков электронные информационные технологии достигли фундаментальных и технических пределов производительности при продолжающемся росте потребительского спроса на скорость и объем обрабатываемой и передаваемой информации. Решение данной проблемы потребовало разработки нового поколения информационно - телекоммуникационных систем, основанных на технологиях фотоники. В фотонике появилось новое динамично развивающееся направление, определяющее прогресс мировой науки и техники, - «оптоинформатика». Под «оптоинформатикой» понимают область науки и техники, связанную с исследованием, разработкой, созданием и эксплуатацией новых материалов, технологий, приборов и устройств, направленных на передачу, прием, обработку, хранение и отображение информации.

Изучение фотоники основывается на знании принципов формирования, преобразования, анализа изображений, теории построения информационных систем. Интеграция фотоники и компьютерных технологий позволяет создавать методы, которые возможно реализовать исключительно средствами компьютерной фотоники, обеспечивая развитие технологий качественно нового уровня.

По многим направлениям фотоники и оптоинформатики Россия находится на уровне промышленно - развитых стран (интегральная оптика, системы приема, обработки и отображения информации и др.), а, по некоторым - даже опережает. Приоритетными направлениями

являются: волоконная оптика (работы академика Дианова Е.М. - ИОФ РАН), голография (академик Денисюк Ю.Н. - ГОИ им. СИ. Вавилова), полупроводниковые лазеры (академик Алферов Ж.И. - ФТИ РАН им. А.Ф. Иоффе), полифункциональные оптические материалы (академик Петровский Г.Т. - ГОИ им. С.И. Вавилова) и др.

Ввиду большого научного и практического значения направления "Фотоника и оптоинформатика", а также спроса на него на потребительском рынке, в 2002 г. в СПбГУ ИТМО был организован факультет «Фотоники и оптоинформатики» под руководством доктора физ.-мат. наук, профессора С.А. Козлова. По инициативе профессорско-преподавательского состава, начиная с 2005 года, на факультете стала работать выпускающая кафедра «Компьютерной фотоники», которую возглавил доктор технических наук, профессор И.П. Гуров.

История кафедры началась в 1946 году. На всех этапах развития результаты научных исследований, проводимых сотрудниками кафедры, неизменно использовались в учебном процессе. Совершенствовались направления подготовки студентов, изменялось название кафедры, но всегда кафедра гордилась своими выпускниками.

Выпускники кафедры занимают видное место в оптической науке: академик РАН Ю.Н. Денисюк, изобретатель трехмерной голографии; член-корр. РАН, профессор Н.Г. Бахшиев, известный специалист в области спектроскопии межмолекулярных взаимодействий; Заслуженный деятель науки РФ, профессор Г.Н. Дульнев, крупный ученый в области теплофизики, долгие годы бывший ректором ЛИТМО; профессор И.М. Нагибина, исследования которой в области физической оптики получили широкое признание.

Одной из важнейших задач кафедры является организация учебного процесса и подготовка профессионалов в области компьютерной фотоники. Направление работы кафедры определяется развитием информационных технологий и компьютерных систем в области формирования, синтеза, обработки и анализа изображений на основе интеграции эффективных компьютерных систем с системами фотоники.

Проводимые исследования в области компьютерной обработки когерентных и некогерентных изображений обеспечивают решение научно-технических задач оптической томографии, цифровой голографии, синтеза, анализа, распознавания и классификации изображений.

Научным консультантом работ кафедры в области компьютерной обработки изображений – иконики – является член-корреспондент РАН М.М. Мирошников.

Кафедра проводит работы в рамках международных научных проектов в сотрудничестве с ведущими зарубежными университетами, институтами и исследовательскими лабораториями Италии, Финляндии, Франции, Германии, Великобритании, Японии, США и других стран в области оптической когерентной томографии для биомедицинских исследований, цифровой голографии для исследования микро- и

nanostructures, three-dimensional photographs of micro- and macroscopic objects, hyperspectral processing of images.

In 2007, St. Petersburg GUITMO became the winner of the Russian competition for the development of an innovative educational program, there is also merit of the teaching staff of the department of Computer Photonics. Participation in the implementation of the National Project «Education» allowed in 2007-2008 academic year to develop innovative programs for the training of bachelors and masters, to equip the teaching laboratories with modern equipment, to prepare and publish new textbooks.