Министерство образования и науки Российской Федерации
Южный федеральный университет

**А.Н. Каркищенко**

# Математические методы обработки изображений

**Курс лекций**

Таганрог, 2017

# ОГЛАВЛЕНИЕ

ПРЕДИСЛОВИЕ

**РАЗДЕЛ 1. ОСНОВНЫЕ ПОНЯТИЯ И ЗАДАЧИ ОБРАБОТКИ ИЗОБРАЖЕНИЙ**

Глава 1. Цифровое изображение: базовые понятия и основные задачи обработки 4

1.1. Понятие цифрового изображения. 4

1.2. Типы изображений и способы их описания 5

1.3. Основные стадии и задачи обработки изображений 10

Глава 2. Регистрация и получение цифровых изображений 12

2.1. Регистрация изображений с помощью сенсоров и формирование цифрового изображения 12

2.2. Дискретизация и квантование аналогового изображения 15

2.3. Пространственное и яркостное разрешение 18

2.4. Основные понятия, описывающие цифровые изображения (пиксель, смежность, связность, области, границы) 20

2.5. Расстояния на изображениях 23

2.6. Операции и преобразования на изображениях 24

2.7. Геометрические преобразования изображений 26

2.8. Двумерные линейные преобразования изображений 28

2.9 Вероятностные методы описания изображений 29

**РАЗДЕЛ 2. ПРОСТРАНСТВЕННЫЕ МЕТОДЫ УЛУЧШЕНИЯ ИЗОБРАЖЕНИЙ**

Глава 3. Пространственные преобразования изображений 31

3.1. Общие принципы пространственных преобразований 31

3.2. Основные градационные преобразования 32

Глава 4. Гистограммные преобразования 37

4.1. Основные понятия и определения 37

4.2. Эквализация гистограмм 38

4.3. Моделирование гистограммного распределения 42

Глава 5. Пространственная фильтрация изображений 44

5.1. Принципы пространственной фильтрации 44

5.2. Корреляция и свертка 46

5.3. Формирование фильтров 47

Глава 6. Сглаживающие пространственные фильтры 48

6.1. Линейные сглаживающие фильтры 48

6.2. Пример нелинейного фильтра. Медианный фильтр 51

Глава 7. Дифференциальные фильтры 51
7.1. Основные принципы построения дифференциальных фильтров 51
7.2. Повышение резкости изображений на основе лапласиана 53
7.3. Повышение резкости: операторы Робертса и Собела 55

**РАЗДЕЛ 3. ЧАСТОТНЫЕ МЕТОДЫ УЛУЧШЕНИЯ ИЗОБРАЖЕНИЙ**

Глава 8. Введение в Фурье-анализ функций одной переменной 58
8.1. Предварительные сведения 58
8.2. Свертка 60
8.3. Преобразование Фурье дискретных функций 61
8.4. Дискретное преобразование Фурье функций одной переменной 64

Глава 9. Введение в Фурье-анализ функций двух переменных 67
9.1. Двумерные непрерывные преобразования Фурье 67
9.2. Двумерная дискретизация и двумерное дискретное преобразование Фурье 68
9.3. Фурье-спектр и Фурье-фаза 71
9.4. Основы частотной фильтрации 73

Глава 10. Низкочастотная фильтрация изображений 76
10.1. Идеальный фильтр низких частот 76
10.2. Фильтры низких частот Баттерворта 78
10.3. Гауссовы фильтры низких частот 80

Глава 11. Высокочастотная фильтрация изображений 82
11.1. Идеальный фильтр высоких частот 82
11.2. Фильтры высоких частот Баттерворта 83
11.3. Гауссовы фильтры высоких частот 84

**РАЗДЕЛ 4. МОРФОЛОГИЧЕСКАЯ ОБРАБОТКА ИЗОБРАЖЕНИЙ**

Глава 12. Базовые понятия и операции морфологической обработки изображений 86
12.1. Основные определения 86
12.2. Эрозия и дилатация 87
12.3. Размыкание и замыкание 91

Глава 13. Основные морфологические алгоритмы 94
13.1. Выделение границ 94
13.2. Заполнение дырок 95
13.3. Выделение связных компонент 96
13.4. Построение скелета 97

**РАЗДЕЛ 5. СЕГМЕНТАЦИЯ ИЗОБРАЖЕНИЙ**

Глава 14. Сегментация: основные понятия и базовые операции 100
14.1. Основные понятия и определения 100
14.2. Обнаружение точек, линий и перепадов яркостей 101
14.3. Простые методы обнаружения контурных перепадов 106

Глава 15. Обнаружение линий и контуров 110
15.1. Детектор контуров Кэнни 110
15.2. Преобразование Хафа 114
Глава 16. Структурные методы сегментации 118
16.1. Выращивание областей 118
16.2. Разделение и слияние областей 120

БИБЛИОГРФИЧЕСКИЙ СПИСОК 125

Раздел 1

# ОСНОВНЫЕ ПОНЯТИЯ И ЗАДАЧИ ОБРАБОТКИ ИЗОБРАЖЕНИЙ

## Глава 1. Цифровое изображение: базовые понятия и основные задачи обработки

### 1.1. Понятие цифрового изображения

Цифровая обработка изображений, как направление исследований, возникла в связи с необходимостью повышать качество изображений для улучшения его визуального восприятия, а также для эффективного хранения, передачи и представления в системах машинного зрения.

Методы обработки изображений, как правило, вычислительно затратные, поэтому реальное использование этих методов оказалось возможным лишь с появлением компьютеров, обладающих высокой производительностью и большим объемом *оперативной* и *внешней* памяти.

Под изображением понимается функция двух переменных $f(x,y)$, где $x$ и $y$ – координаты на плоскости.

| | |
|---|---|
| **Определение** | Значение функции в любой точке, задаваемой парой координат $(x,y)$, называется *яркостью*, *интенсивностью* или *уровнем серого* изображения в этой точке. Если величины $x$, $y$ и $f(x,y)$ принимают конечное число дискретных значений, то такое изображение называется *цифровым изображением*. |

| | |
|---|---|
| **Определение** | Элементы с координатами $(x,y)$ называются пикселями (от англ. *picture element*). |

Обработка изображений – широкое понятие. Между тем, оно не покрывает всех задач работы с цифровыми изображениями. Не существует общепринятой точки зрения, где заканчивается обработка изображений, и начинаются другие смежные области. Однако можно выделить несколько основных направлений исследований по мере возрастания уровня:

Обработка изображений

Распознавание объектов на изображениях

Анализ сцен

Анализ изображений

Машинное зрение

На всех уровнях работы с изображениями – от обработки изображений до машинного зрения можно различать процессы низкого, среднего и высокого уровня.

Процессы *низкого* уровня касаются только примитивных операций – уменьшение шума, повышение контраста, улучшение резкости изображений.

На *среднем* уровне решаются такие задачи, как сегментация (разделение изображения на области или выделение на нем объектов), описание объектов и сжатие их, а также распознавание отдельных объектов.

*Высокоуровневая* обработка включает в себя «понимание» распознанных объектов и изображенных сцен. Этот уровень близок к тому, что понимается под зрением.

**Пример.** Автоматическое распознавание печатного или рукописного текста.

В сферу цифровой обработки изображений входят процессы получения изображения области, содержащей текст, предварительной обработки полученного изображения, выделения (сегментации) отдельных символов текста, описания символов в подходящей для компьютерной обработки форме и, наконец, распознавания этих символов.

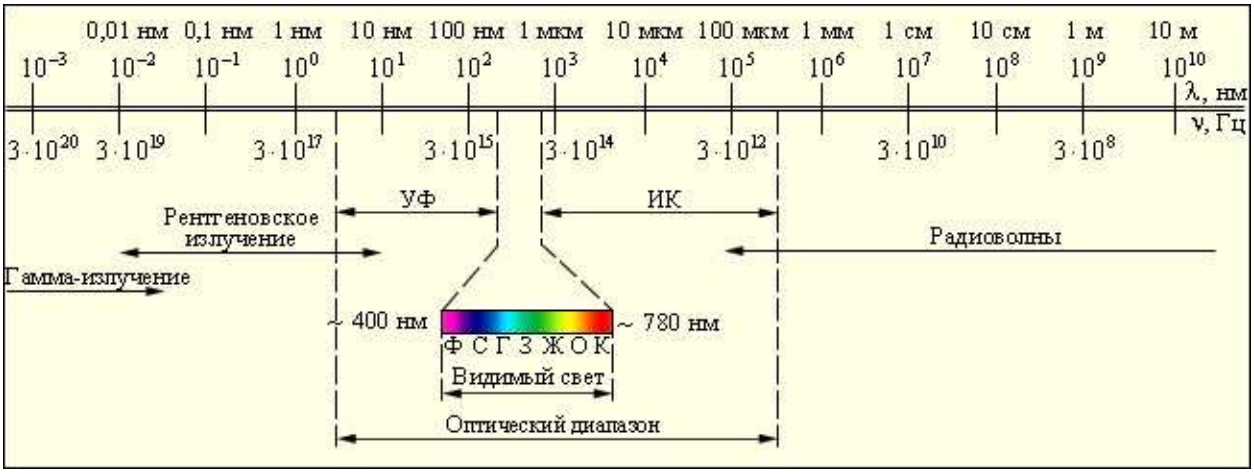
Однако осмысление содержимого текста уже относится к сфере анализа изображений или даже машинного зрения, в зависимости от уровня сложности, который подразумевается под этим «осмыслением».

## 1.2. Типы изображений и способы их описания

Люди способны воспринимать электромагнитное излучение лишь в видимом диапазоне. Однако визуализации поддаются излучения в значительно более широком спектре. Поэтому машинная обработка изображений охватывает практически весь электромагнитный спектр от гамма-излучения до радиоволн. Более того, могут визуализироваться и обрабатываться изображе-

ния, которые получаются не из электромагнитных источников, например, акустическими, гидроакустическими, ультразвуковыми волнами.

Ниже приведена логарифмическая шкала частот (и длин волн) электромагнитных излучений, изображения в которых обрабатываются цифровыми методами. Видимая часть электромагнитного спектра лежит в пределах примерно от 400 до 700 нанометров. Слева и справа на этой шкале – зоны электромагнитных частот, не видимые глазу.




С физической точки зрения большинство изображений представляют собой двумерное распределение интенсивности *электромагнитного* излучения, отраженное регистрируемым объектом, или пошедшее сквозь него. Кроме электромагнитного излучения источниками изображений могут быть также *акустические*, *гидроакустические* и *ультразвуковые* изображения. В настоящее время интенсивно развиваются также методы обработки и восстановления изображений, полученных на основе так называемых полей дальностей, формируемых, как правило, различными типами *лазерных* сканеров. И это далеко не полный перечень источников получаемых изображений.

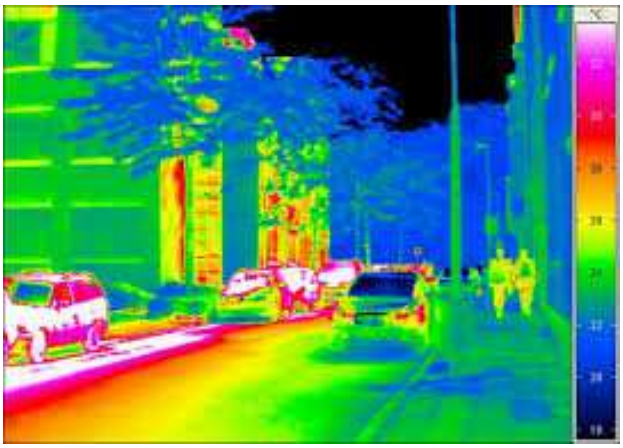
Изображения, полученные в различных диапазонах электромагнитных волн, а также с помощью волн другой природы, имеют свои специфические особенности, которые определяют области их применения. Рассмотрим примеры изображений различной физической природы.

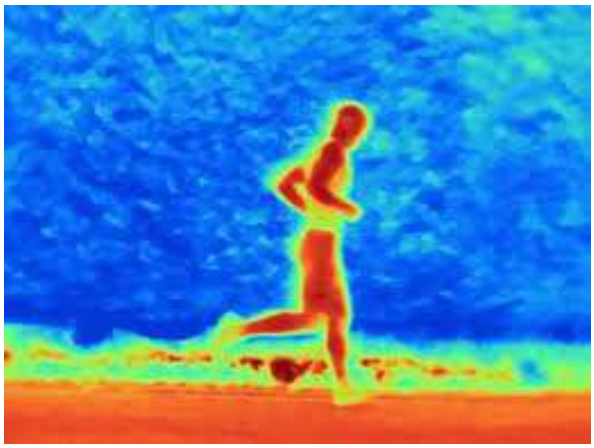
**Рентгеновские изображения**

Рентгеновское излучение имеет достаточно большую мощность, чтобы проникать внутрь физических тел, что позволяет сделать видимое изображение. Такие изображения используются в медицине, системах технического контроля и системах безопасности. Примеры рентгеновских изображений:

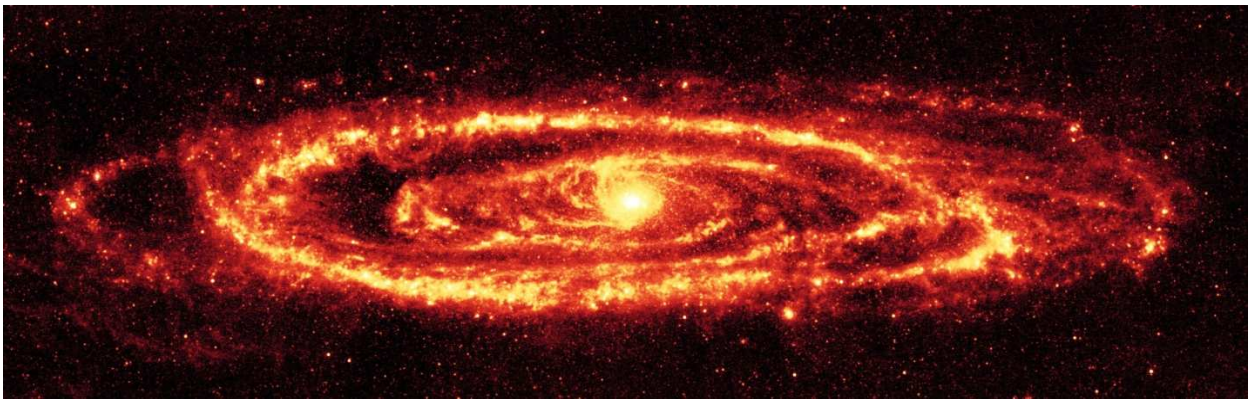
**Изображения в инфракрасном диапазоне**

Большое количество инфракрасного излучения исходит от теплых предметов. Холодные объекты на таких изображениях выглядят более темными, а теплые – более яркими. Используются при создании приборов ночного видения, в системах экологического и промышленного мониторинга, в системах наведения на военную технику, например, на двигатели, «факелы» ракет и т.п.  Примеры изображений в инфракрасном диапазоне:



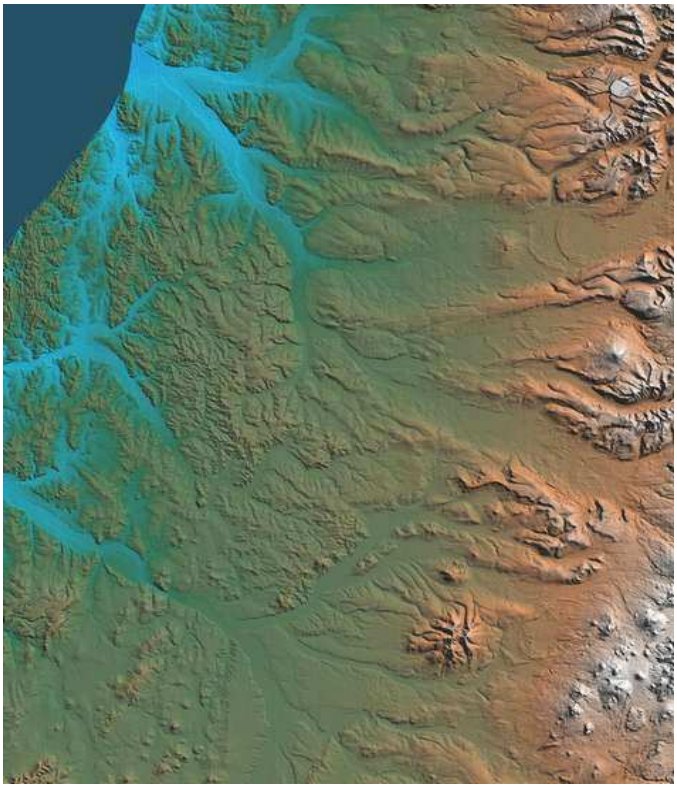


А вот туманность Андромеды, снятая в инфракрасном диапазоне:

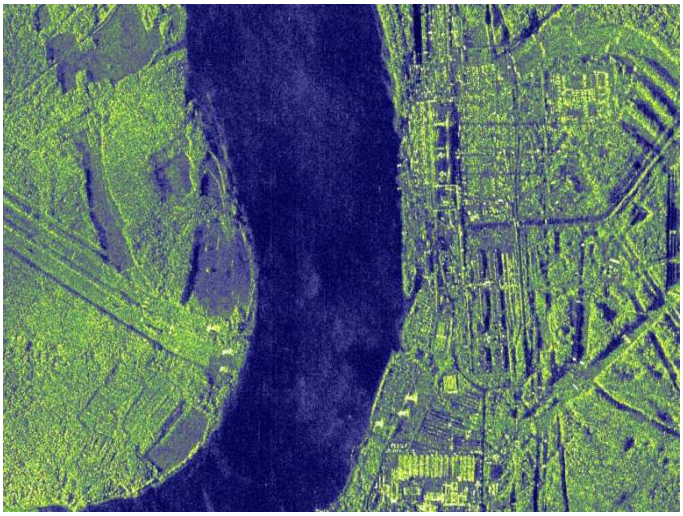


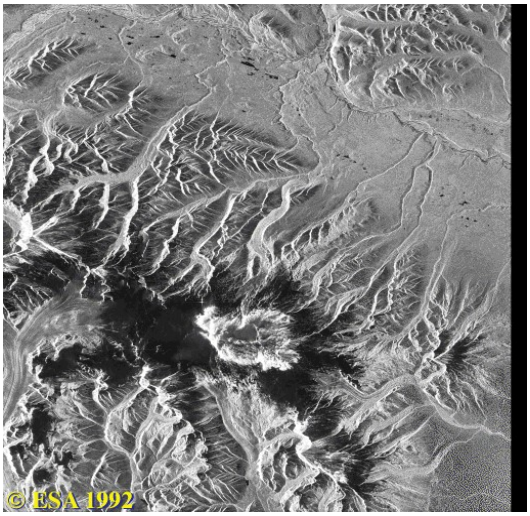
## Радиолокационные изображения

Для формирования изображений используется генератор микроволнового излучения, который сканирует поверхность и фиксирует отраженные сигналы. Получение этих изображений не зависят от погодных условий, времени суток и наличия, например, облаков. Ниже приведены примеры изображений земной поверхности, полученных в диапазоне радиоволн с помощью так называемых радаров с синтезированной апертурой:



И еще …





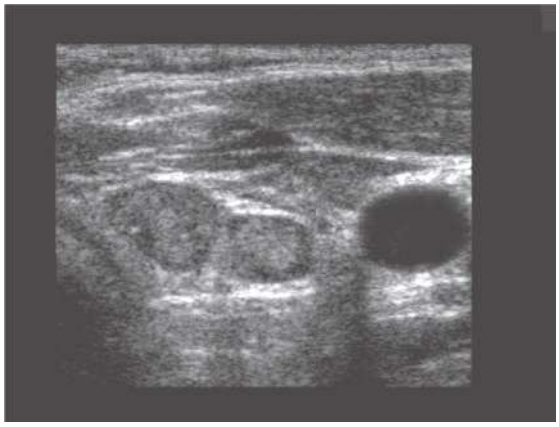
## Лазерное сканирование

Лазерные сканеры – это дальнометрические системы, которые работают на основе испускаемых электромагнитных импульсов, отражение которых фиксируется приемником, На основе времени прохождения импульса вычисляется расстояние до объекта.

Ниже приведен пример изображения, полученного посредством лазерного сканирования. Диапазон, в котором работают лазерные сканеры, покрывает ультрафиолетовую, видимую и инфракрасную части электромагнитного спектра.

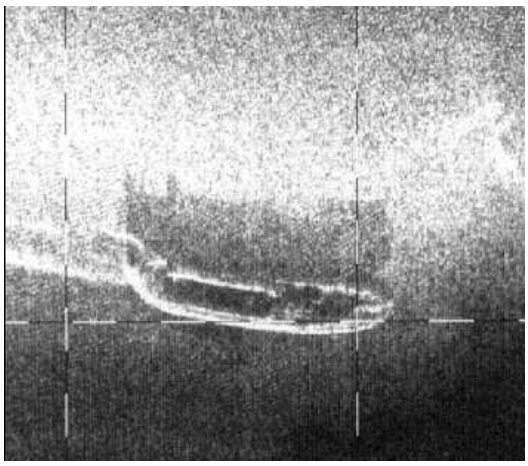
## Ультразвуковые изображения

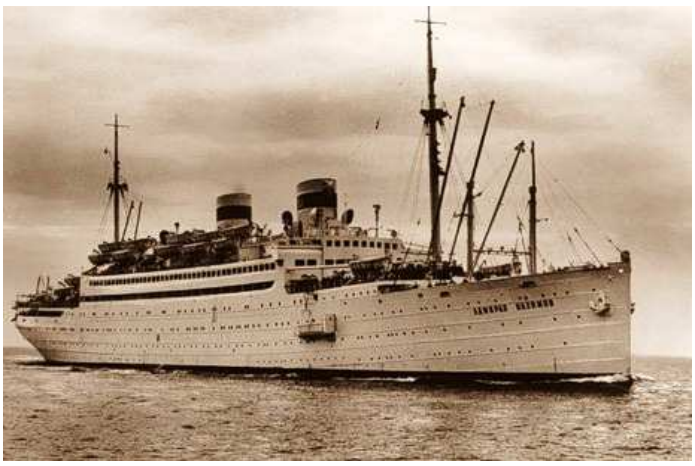
Ультразвуковые волны многократно отражаются на границах сред разной плотности и фиксируются приемником волн. Применяются в медицине и промышленности для неразрушающего контроля.



**Гидроакустические изображения**

Здесь также не электромагнитные волны формируют изображение, а акустические (т.е. звуковые) волны специальной частоты. Изображение формируется за счет отражения акустических волн от препятствий. Ниже приведен пример изображения, полученного с помощью гидролокационной станции:





На этом снимке изображен теплоход «Адмирал Нахимов», затонувший 31 августа 1986 года в бухте Новороссийска.

Многие современные системы технического зрения основаны на получении и анализе не одного типа изображений, а сразу нескольких, полученных в разных диапазонах. В этом случае говорят о *комплексировании* изображений, *мультиспектральной съемке*, *многозональных изображениях* и т.п.

Перечисленные возможности делают автоматическую обработку изображений еще более важным инструментом, как теоретического исследования природы, так и создания совершенно новых приборов прикладного, прагматического назначения.

## 1.3. Основные стадии и задачи обработки изображений

Задачи, связанные с обработкой изображений, можно сгруппировать в ряд последовательных стадий или этапов.

*Регистрация изображений* – это в большей степени технический этап, связанный с получением цифрового изображения в виде графического файла. На этом этапе могут выполняться лишь саамы простые преобразования, такие, например, как изменения масштаба.

*Улучшение изображений* – это процесс различных преобразований изображения, в результате которого оно становится более удобным для восприятия. Оценка результата таких улучшений осуществляется субъективно и зависит от характера изображения, физической природы его получения, а также области применения. Соответственно, методы улучшения изображений различны и зависят от цели улучшения. Единой теории улучшения изображений не существует, а критерии качества методов субъективны.

*Восстановление изображений* – это группа методов, также связанная с повышением качества изображений. Однако эти методы основаны на достаточно строгих математических методах исправления искажений или разрушений изображения. В этом смысле они объективны.

*Обработка цветных изображений* – этап, на котором применяются методы работы с цветными изображениями. В настоящее время цвет играет очень важную роль в выразительных свойствах изображения практически во всех приложениях.

*Сжатие изображений*. Цифровые изображения, особенно в радиоволновом диапазоне, требуют для хранения очень большого объема компьютерной памяти. Поэтому сжатие этой информации с целью уменьшения требуемой памяти – один из этапов цифровой обработки. Особенно это важно для передачи изображений, поскольку пропускная способность современных каналов связи не очень высока.

*Морфологическая обработка* – это этап, на котором из изображения извлекаются или изменяются некоторые компоненты изображения, которые могут оказаться полезными для последующего описания формы представленных на изображении объектов.

*Сегментация* – процесс разделения изображения на отдельные части, предполагающий выделение отдельных объектов. При сегментации требуется обнаружение границ, углов и других внешних признаков объекта, а также выявление текстурных характеристик.

*Представление и описание*. Этап, который следует за этапом сегментации. Основная задача на этом этапе – перейти от сегментированных изображений к описанию объектов с помощью некоторых признаков, которые могли бы стать основой для компактного описания объектов и их дальнейшего распознавания.

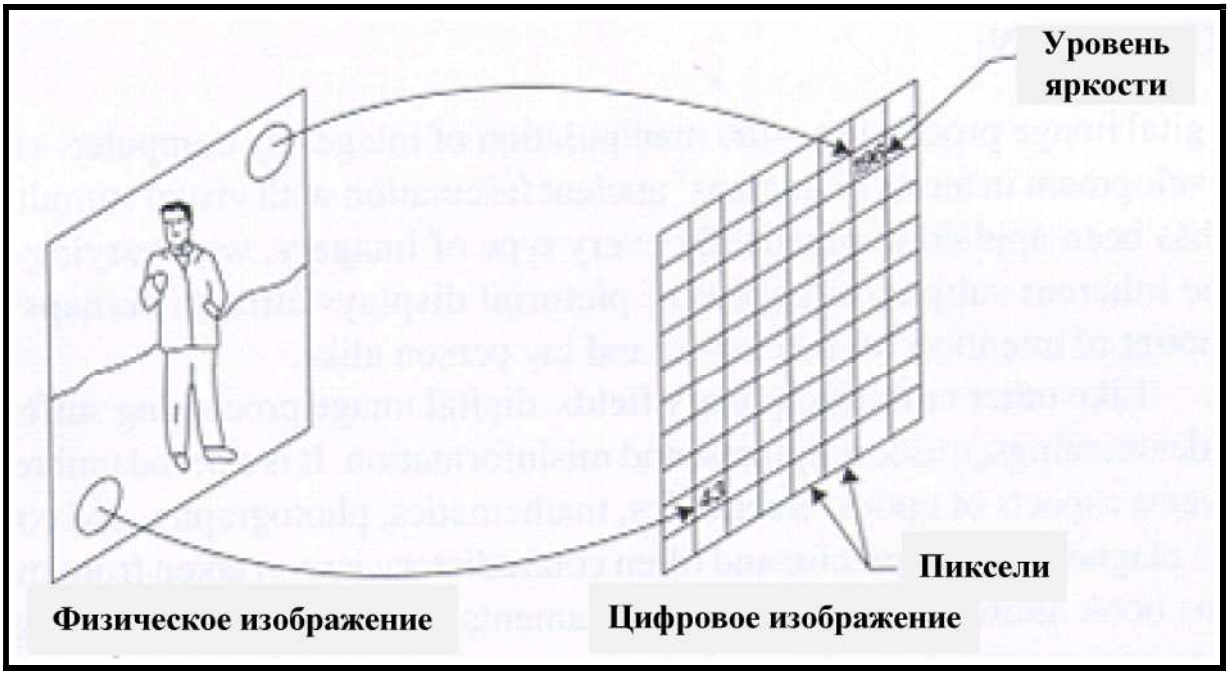
*Распознавание* – завершающий этап цифровой обработки. На этом этапе осуществляется распознавание объекта. В общем случае это означает от-

несение объекта на основе его описания к одному из классов объектов, имеющихся в базе данных (или знаний) системы. Для такого отнесения используются признаки, полученные на предыдущем этапе. Алгоритмы распознавания, как правило, строятся на основе достаточно тонких математических методов.

# Глава 2. Регистрация и получение цифровых изображений

## 2.1. Регистрация изображений с помощью сенсоров и формирование цифрового изображения

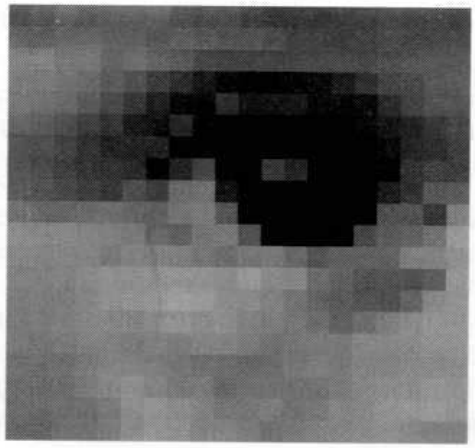
Большинство устройств ввода данных для изображений (в том числе цифровых камер) используют так называемую *матрицу сенсоров*, которая объединяет большое количество светочувствительных элементов (4000x4000 и более). Двумерность матрицы сенсоров позволяет сформировать сразу все изображение, если сфокусировать на поверхности матрицы отвечающий этому изображению пространственный поток лучистой энергии.




Отклик каждого сенсора пропорционален интегралу световой энергии, попадающей на его поверхность за время экспозиции. Благодаря этому снижается уровень шума на изображении. Чем дольше экспозиция, тем меньше шум.

Как показано на рисунке, энергия, излучаемая источником освещения, отражается от объекта сцены (энергия также может и проникать сквозь объекты сцены). Поступающая энергия фокусируется на плоскости изображения. Если для освещения используется источник видимого света, то на входе системы формирования изображения имеется оптическая линза, которая проецирует наблюдаемую сцену на плоскость изображения.

Чувствительная матрица, которая совмещена с этой плоскостью, генерирует набор выходных сигналов, каждый из которых пропорционален интегралу световой энергии, принятой соответствующим сенсором. Тот факт, что регистрация двумерного сигнала осуществляется дискретно расположенными в пространстве сенсорами, обеспечивает пространственную дискретизацию сигнала. В следующем блоке системы осуществляется квантование сигналов. На выходе получается цифровое изображение.



| | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 117 | 125 | 133 | 127 | 130 | 130 | 133 | 121 | 116 | 115 | 100 | 91 | 93 | 94 | 99 | 103 | 112 | 105 | 109 | 106 |
| 134 | 133 | 138 | 138 | 132 | 134 | 130 | 133 | 128 | 123 | 121 | 113 | 106 | 102 | 99 | 106 | 113 | 109 | 109 | 113 |
| 146 | 147 | 138 | 140 | 125 | 134 | 124 | 115 | 102 | 96 | 93 | 94 | 99 | 96 | 99 | 100 | 103 | 110 | 109 | 110 |
| 144 | 141 | 136 | 130 | 120 | 108 | 88 | 74 | 53 | 37 | 31 | 37 | 35 | 39 | 53 | 79 | 93 | 100 | 109 | 116 |
| 139 | 136 | 129 | 119 | 102 | 85 | 58 | 31 | 41 | 77 | 51 | 53 | 53 | 33 | 37 | 41 | 69 | 94 | 105 | 108 |
| 132 | 127 | 117 | 102 | 87 | 57 | 49 | 77 | 42 | 28 | 17 | 15 | 13 | 13 | 17 | 41 | 53 | 69 | 88 | 100 |
| 124 | 120 | 108 | 94 | 72 | 74 | 72 | 31 | 35 | 31 | 15 | 13 | 15 | 11 | 15 | 13 | 46 | 75 | 83 | 96 |
| 125 | 115 | 102 | 93 | 88 | 82 | 42 | 79 | 113 | 41 | 19 | 100 | 82 | 11 | 11 | 17 | 31 | 91 | 99 | 100 |
| 124 | 116 | 109 | 99 | 91 | 113 | 99 | 140 | 144 | 57 | 20 | 20 | 15 | 11 | 15 | 17 | 63 | 87 | 119 | 124 |
| 136 | 133 | 133 | 135 | 138 | 133 | 132 | 144 | 150 | 120 | 24 | 17 | 15 | 15 | 17 | 20 | 115 | 113 | 88 | 150 |
| 158 | 157 | 157 | 154 | 149 | 145 | 133 | 127 | 146 | 150 | 116 | 35 | 20 | 19 | 28 | 105 | 124 | 128 | 141 | 171 |
| 155 | 154 | 156 | 155 | 146 | 155 | 154 | 154 | 147 | 139 | 148 | 150 | 138 | 120 | 128 | 129 | 130 | 151 | 156 | 165 |
| 150 | 151 | 154 | 162 | 166 | 167 | 169 | 174 | 172 | 167 | 177 | 166 | 164 | 140 | 134 | 120 | 121 | 120 | 127 | 172 |
| 145 | 149 | 151 | 157 | 165 | 169 | 173 | 179 | 176 | 166 | 166 | 157 | 145 | 136 | 129 | 124 | 120 | 136 | 163 | 168 |
| 144 | 148 | 153 | 160 | 159 | 158 | 165 | 172 | 165 | 169 | 157 | 151 | 149 | 141 | 130 | 140 | 151 | 162 | 169 | 167 |
| 144 | 141 | 147 | 155 | 154 | 149 | 156 | 151 | 157 | 157 | 151 | 144 | 147 | 147 | 149 | 159 | 158 | 159 | 166 | 165 |
| 139 | 140 | 140 | 150 | 153 | 151 | 150 | 146 | 140 | 139 | 138 | 140 | 145 | 151 | 149 | 156 | 156 | 162 | 162 | 161 |
| 136 | 134 | 138 | 146 | 156 | 164 | 153 | 146 | 145 | 136 | 139 | 139 | 140 | 141 | 149 | 157 | 159 | 161 | 169 | 166 |
| 136 | 133 | 136 | 135 | 144 | 159 | 168 | 159 | 151 | 142 | 141 | 145 | 139 | 146 | 153 | 156 | 164 | 167 | 172 | 168 |
| 133 | 129 | 140 | 142 | 146 | 159 | 167 | 165 | 154 | 151 | 146 | 141 | 147 | 154 | 156 | 160 | 161 | 157 | 153 | 154 |

Значение функции $f(x,y)$ в точке с пространственными координатами $(x,y)$ является положительной скалярной величиной, физический смысл которой определяется источником изображения. Эту функцию можно характеризовать двумя компонентами: величиной светового потока $i(x,y)$, который

падает на наблюдаемую сцену от источника, и относительной долей $r(x,y)$ светового потока, отраженного от объектов этой сцены.

| | |
|---|---|
| **Определение** | Компоненты $i(x,y)$ и $r(x,y)$ называются соответственно *освещенностью* и *коэффициентом отражения*. |

Произведение этих функций дает функцию изображения:

$$f(x,y) = i(x,y)r(x,y),$$

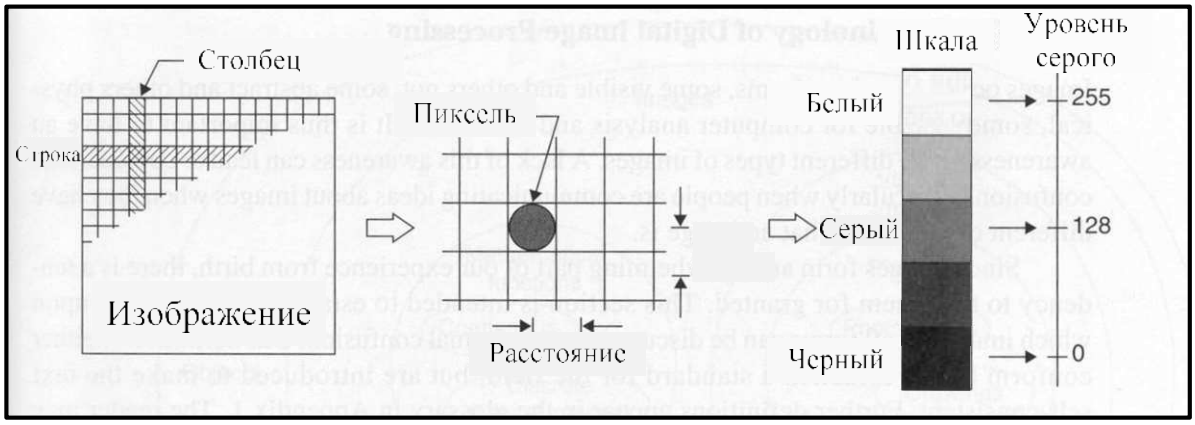
причем $0 < i(x,y) < \infty$, $0 \le r(x,y) \le 1$. Значение $r(x,y) = 0$ соответствует полному поглощению, а $r(x,y) = 1$ – полному отражению. Значение функции $i(x,y)$ зависит только от источника освещения, а $r(x,y)$ определяется свойствами объектов на сцене.

Приведенные выражения применимы и к изображениям, сформированным в проходящем освещении (например, при рентгеноскопии). В этом случае функцию $r(x,y)$ называют *коэффициентом пропускания* вместо коэффициента отражения.

Теоретически к границам этого интервала предъявляются только требования, чтобы $F_{\min}$ было положительно, а $F_{\max}$ – конечно.

| | |
|---|---|
| **Определение** | Интервал $[F_{\min}, F_{\max}]$ называется *диапазоном яркостей*. |

На практике его сдвигают по числовой оси, получая отрезок [0, $L$ – 1], границы которого принимаются за уровень черного $f(x,y)$ = 0 и уровень белого $f(x,y) = L - 1$. Все промежуточные значения на этом отрезке соответствуют некоторым оттенкам серого при изменении от черного до белого.

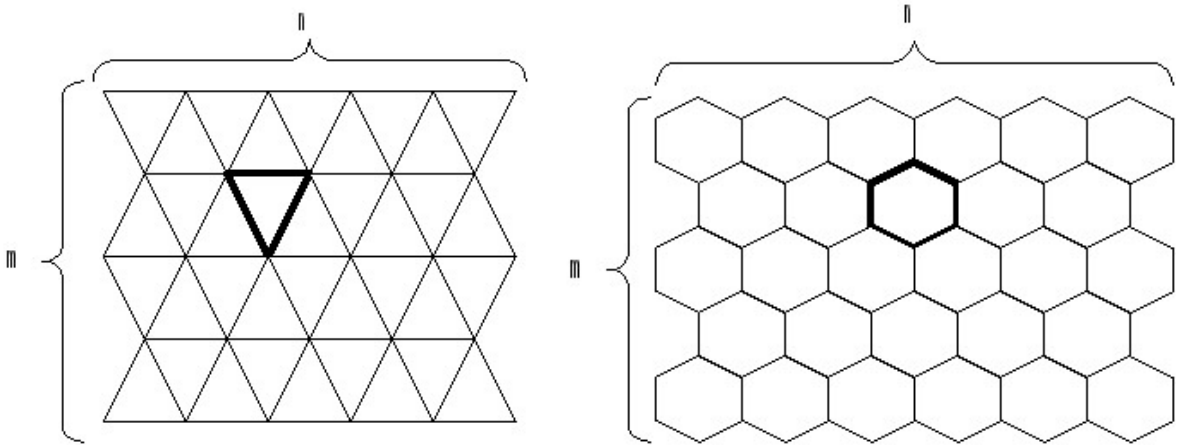


Отметим в заключение, что в качестве элемента сенсорной матрицы можно использовать квадраты, но фигуры другой формы. Эти фигуры должны удовлетворять следующим требованиям:

– все фигуры должны быть одинаковые;

– должны полностью покрывать плоскость без перекрытия и дырок.

Этим условиям удовлетворяют только правильный треугольник и правильный шестиугольник – гексагон.



Во многих случаях гексагональная матрица оказывается предпочтительнее, чем прямоугольная. Например, при отображении прямых линий (и не только прямых), а также при осуществлении операций над изображениями, например, поворотов. Однако технологически изготовление таких матриц существенно сложнее, а математический аппарат, реализующий такие преобразования также оказывается более сложным и ресурсно более затратным.

## 2.2. Дискретизация и квантование аналогового изображения

Большинство сенсоров выдают на выходе аналоговый сигнал в форме, который непрерывно меняется по величине. Форма и амплитуда сигнала связаны с регистрируемым физическим явлением. Для того чтобы получить цифровое изображение, необходимо преобразовать непрерывно сигнал в цифровую форму. Эта операция включает в себя два процесса — *дискретизацию* и *квантование*.

| | |
|---|---|
| **Определение** | Представление координат в виде конечного множества точек (отсчетов) называется *дискретизацией*, а представление значений функции яркости конечным множеством значений называется *квантованием*. |

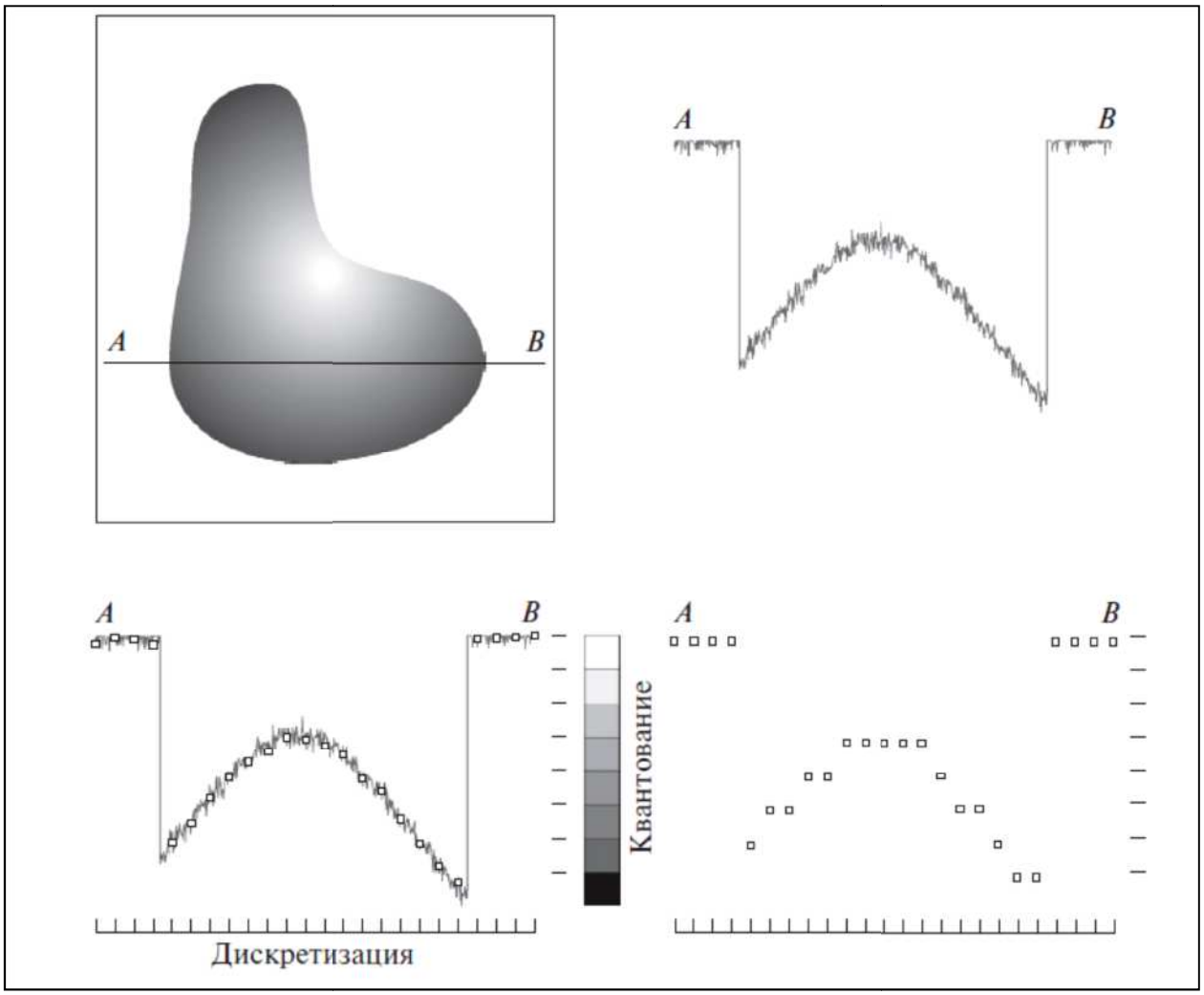
Процессы дискретизации и квантования иллюстрируются на приведенном ниже рисунке. Слева вверху показано исходное непрерывное изображение, которое должно быть преобразовано в цифровую форму. Изображенная на этом рисунке одномерная функция представляет собой график изменения значений яркости непрерывного изображения вдоль отрезка *AB*. Случайные отклонения на графике вызваны наличием шумов в изображении.

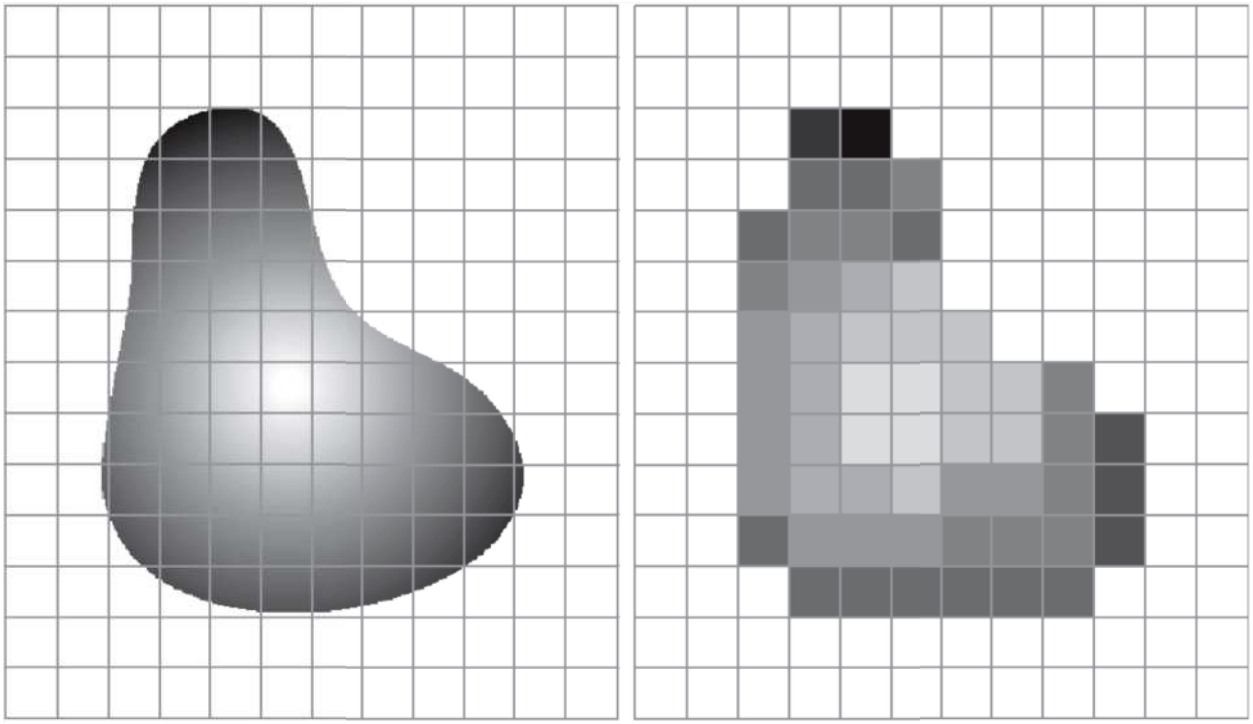
Чтобы дискретизировать эту функцию, разобьем отрезок *AB* на равные интервалы. Значения в выбранных точках отсчета представлены на графике функции квадратиками. Построенный набор значений в точках дискретизации описывает функцию в виде совокупности дискретных отсчетов, однако сами значения пока еще охватывают весь непрерывный диапазон яркостей.

Чтобы построить цифровую функцию, диапазон яркостей также необходимо преобразовать в дискретные величины (*проквантовать*). Справа изображена шкала яркостей, разбитая на восемь дискретных интервалов от черного до белого. Засечки указывают конкретное значение яркости, присвоенное каждому интервалу. Квантование непрерывных значений яркости в точках дискретизации осуществляется простым сопоставлением каждому отсчету одного из ближайших восьми отмеченных значений.

Выполняя такую процедуру построчно, мы получим двумерное цифровое изображение.

На следующем рисунке показано непрерывное изображение, спроецированное на плоскость сенсорной матрицы, и это же изображение после дискретизации и квантования. Качество получаемого представления в большой степени зависит от числа отсчетов в разбиении и числа уровней квантования.


A
B
A
B
A
B
Квантование
A
B
Дискретизация



Как теперь представить цифровое изображение в компьютере для дальнейших операций с ним? Предположим, что сенсорная матрица имеет $M$ строк и $N$ столбцов. Тогда значения, принимаемые квантованной функцией яркости $f(x,y)$, можно записать в виде следующей матрицы:

$$f(x,y)=\begin{bmatrix} f(0,0) & f(0,1) & \cdots & f(0,N-1) \\ f(1,0) & f(1,1) & \cdots & f(1,N-1) \\ \vdots & \vdots & & \vdots \\ f(M-1,0) & f(M-1,1) & \cdots & f(M-1,N-1) \end{bmatrix}.$$

При обработке изображений элементы этой матрицы принято называть элементами изображения или пикселями. Понятно, что для записи изображения его можно представлять и в виде матрицы, записанной в привычных обозначениях:

$$A=\begin{bmatrix} a_{00} & a_{01} & \cdots & a_{0,N-1} \\ a_{10} & a_{11} & \cdots & a_{1,N-1} \\ \vdots & \vdots & & \vdots \\ a_{M-1,0} & a_{M-1,1} & \cdots & a_{M-1,N-1} \end{bmatrix}.$$

Значения $M$ и $N$ могут быть любыми целыми положительными числами. Однако число градаций яркости по технологическим причинам обычно является целочисленной степенью двойки.

## 2.3. Пространственное и яркостное разрешение

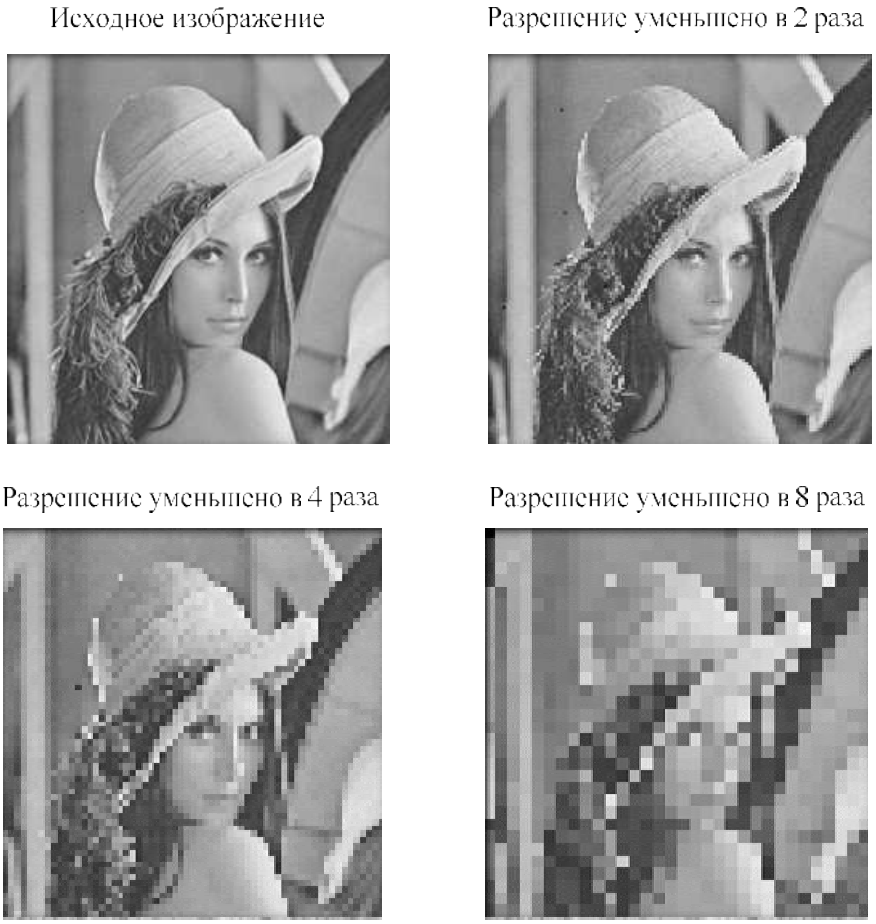
В качественном смысле пространственное разрешение – это размер мельчайших различимых деталей на изображении. Количественно же *пространственное разрешение* определяется, как правило, двумя способами.

Число пар линий на единицу длины. Пусть имеется чертеж, состоящий из чередующихся черных и белых вертикальных линий, каждая шириной $w$. Ширина *пары линий* составляет $2w$, и на единице длины поместится $1/2w$ таких пар. Например, если ширина линии равна 0,1 мм, то на одном миллиметре поместится пять пар линий.

Разрешение определение указанием максимального числа *различимых* пар линий на единицу длины; например 100 пар линий на миллиметр.

Число точек (пикселей) на единицу длины. Это количественная мера разрешения изображения, которая широко применяется в полиграфии. Ее выражают в *точках на дюйм* (dpi – *dots per inch*).

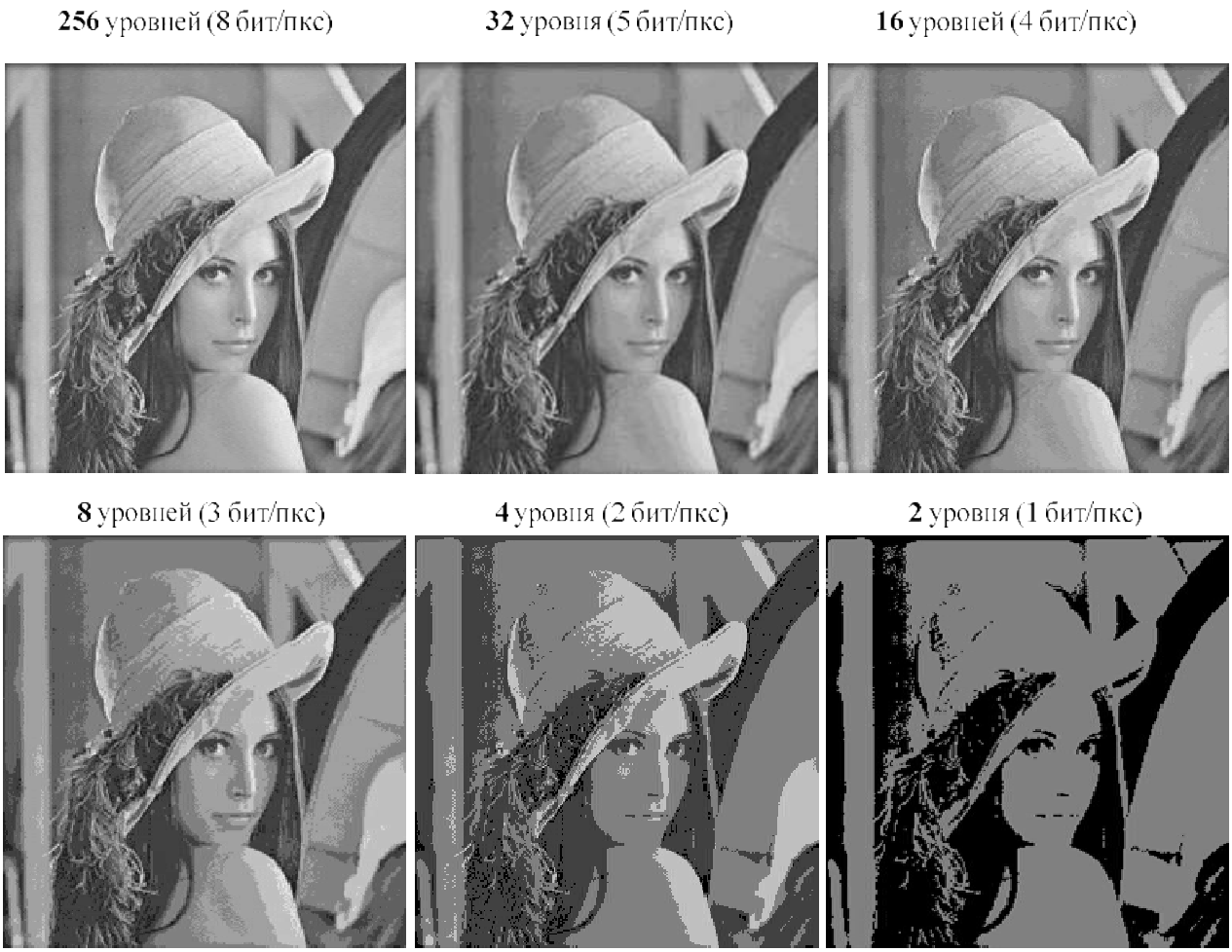
Важно обратить внимание, что мера пространственного разрешения должна определяться по отношению к пространственным единицам измерения. Размеры изображения сами по себе не дают полной информации. Например, если изображение имеет разрешение 1024x1024 пикселя, это мало что говорит о качестве этого изображения, поскольку не указаны размеры, занимаемые этим изображением. Например, квадрат с шириной 128 пикселей при отображении с разрешением 72 dpi имеет физическую ширину 45 мм (1 дюйм = 2.54 см). На рисунке ниже приведены примеры одного и того же изображения, но с разным пространственным разрешением.




*Яркостным* (или *полутоновым*) *разрешением* аналогично называется мельчайшее различимое изменение яркости. В отличие от дискретизации при выборе числа градаций яркости приходится учитывать, что число градаций обычно равно степени 2. В подавляющем большинстве случаев используется 8-битовое представление (256 градаций яркости).

В отличие от пространственного разрешения, которое обязательно дается в отношении к единице длины, *яркостным разрешением* принято называть число бит, используемых при квантовании величины яркости. На сле-

дующем рисунке представлены примеры того же изображения, но с разным яркостным разрешением.




## 2.4. Основные понятия, описывающие цифровые изображения (пиксель, смежность, связность, области, границы)

Соседство и шаблоны пикселей

Рассмотрим произвольный пиксель $p(x,y)$, не лежащий на краю изображения. Такой пиксель будем называть *внутренним* пикселем изображения.

| | |
|---|---|
| **Определение** | Пиксели, соседние с внутренним пикселем $p(x,y)$ по горизонтали и вертикали, т.е. пиксели с координатами $(x-1,y)$, $(x+1,y)$, $(x,y-1)$ и $(x,y+1)$ называются *четверкой соседей* пикселя $p(x,y)$ и обозначаются $N_4(p)$.<br>Указанную окрестность соседних пикселей называют также *четырехточечным шаблоном*. |

| | |
|---|---|
| **Определение** | Пиксели, соседние с внутренним пикселем $p(x,y)$ по диагоналям, т.е. пиксели с координатами $(x-1,y-1)$, $(x-1,y+1)$, $(x+1,y-1)$ и $(x+1,y+1)$ называются *диагональными соседями* пикселя $p(x,y)$ и обозначаются $N_D(p)$. |

| | |
|---|---|
| **Определение** | Пиксели, соседние с внутренним пикселем $p(x,y)$ как по горизонтали и вертикали, так и по диагоналям, называются *восьмеркой соседей* пикселя $p(x,y)$ и обозначаются $N_8(p)$.<br>Данную окрестность пикселя $p(x,y)$ называют *восьмиточечным шаблоном*. |

Множества $N_4(p)$, $N_D(p)$ и $N_8(p)$ называют также окрестностями пикселя $p$. Иногда по аналогии определяется понятие 15-точечного шаблона. Но оно используется значительно реже.

Смежность, связность, области и границы

Помимо соседства на множестве пикселей определяется важное понятие смежности. Если в понятии соседства фигурировала только геометрическая близость пикселей, то смежность подразумевает еще и близость функции яркости, задаваемой на этих пикселях.

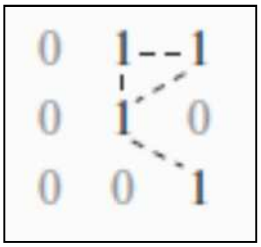
Пусть $V$ – некоторое подмножество возможных значений яркости пикселей.

| | |
|---|---|
| **Определение** | Два пикселя $p$ и $q$ называются *4-смежными*, если $q$ входит во множество $N_4(p)$ и значения яркости на этих пикселях принадлежат множеству $V$. |

| | |
|---|---|
| **Определение** | Два пикселя $p$ и $q$ называются *8-смежными*, если $q$ входит во множество $N_8(p)$ и значения яркости на этих пикселях принадлежат множеству $V$. |

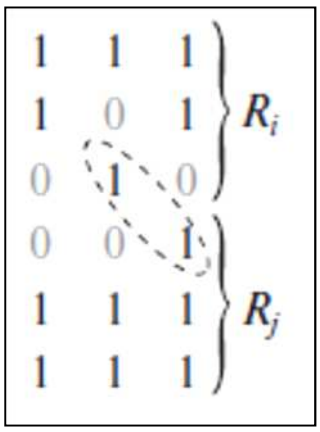
| | |
|---|---|
| **Определение** | *Путем* или *кривой* от пикселя $p(x,y)$ до пикселя $q(s,t)$ называется неповторяющаяся последовательность пикселей с координатами<br>$$(x_0,y_0),(x_1,y_1),\ldots,(x_n,y_n)$$ |

| | |
|---|---|
| | такая, что $(x_0, y_0) = (x, y)$, $(x_n, y_n) = (s, t)$, и пиксели $(x_{t-1}, y_{t-1})$ и $(x_t, y_t)$ будут смежными для всех $t = 1, 2, ..., n$. Число $n$ называется *длиной* пути. Если $(x_0, y_0) = (x_n, y_n)$, то путь называется *замкнутым*. |




В зависимости от принятого типа смежности можно говорить о 4-пути или 8-пути. На рисунке показаны два 8-пути между верхним и нижним правыми элементами.

Пусть $S$ – подмножество элементов изображения. Два элемента $p$ и $q$ из $S$ называются *связными* в $S$, если между ними существует путь, целиком состоящий из элементов подмножества $S$. Для любого пикселя $p$ из $S$ множество всех пикселей, связных с ним в $S$, называется *связной компонентой* (или *компонентой связности*) $S$. Если множество $S$ содержит только одну компоненту связности, оно называется *связным* множеством.




Подмножество $S$ элементов изображения будем называть *областью*, если $S$ – связное множество. Две области $S_1$ и $S_2$ называются *смежными*, если их объединение является связным множеством. Области, не являющиеся смежными, называются *раздельными* (или *несмежными*). При этом необходимо указать используемый тип смежности, т.е. либо 4-смежность, либо 8-смежность. Например, две области единичных пикселей на следующем рисунке являются смежными только в смысле 8-смежности, поскольку в соответствии с данным выше определением, не существует 4-пути между этими областями, и потому их объединение не является связным множеством.

Множество связных областей иногда называют *объектами* изображения, а множество пикселей, не принадлежащих объектам – *фоном*.

*Границей* области $S$ (или *краем*) называется множество точек этой области, которые являются смежными к точкам фона, т.е. граница области — это множество пикселей этой области, у которых один или более соседей принадлежат фону. Например, точка, обведенная кружком на рисунке, не является частью границы области единичных пикселей, если при рассмотрении смежности интересующей области и фона используется 4-связность. Во из-

| | | | | |
|---|---|---|---|---|
| 0 | 0 | 0 | 0 | 0 |
| 0 | 1 | 1 | 0 | 0 |
| 0 | 1 | 1 | 0 | 0 |
| 0 | 1 | 1 | 1 | 0 |
| 0 | 1 | 1 | 1 | 0 |
| 0 | 0 | 0 | 0 | 0 |

бежание подобных ситуаций, как правило, смежность точек области и фона определяют в терминах 8-связности.

Предыдущее определение дает определение *внутренней границы* области, в отличие от *внешней границы*, определяемой аналогичным образом для фона. Иногда важно различать внутренние и внешние границы. Например, внутренняя граница области пикселей с единичным значением на рисунке справа совпадает с самой областью. Такая граница не удовлетворяет данному выше определению замкнутого пути. С другой стороны, внешняя граница этой области образует замкнутый путь вокруг нее.

| | | |
|---|---|---|
| 0 | 0 | 0 |
| 0 | 1 | 0 |
| 0 | 1 | 0 |
| 0 | 1 | 0 |
| 0 | 1 | 0 |
| 0 | 0 | 0 |

В заключение надо сказать о границе всего изображения. Если $S$ – все изображение, то его граница состоит из элементов первого и последнего столбцов, а также первой и последней строки. Поскольку у этих элементов нет соседей за пределами изображения, то под областью понимают все изображение, а элементы на границе области, которые совпадают с краем изображения, всегда включаются в состав границы этой области.

## 2.5. Расстояние на изображениях

При разработке методов и построении алгоритмов обработки изображений часто приходится иметь дело с расстояниями между пикселями. Расстояние можно вводить многими способами, но в любом случае оно должно удовлетворять следующему определению.

Пусть $p$, $q$ и $z$ – произвольные пиксели изображения. Функция $d$ называется функцией расстояния, просто расстоянием или метрикой, если для любых $p$, $q$ и $z$ выполняются условия:

1. $d(p,q) \geq 0$, причем $d(p,q) = 0$ в том и только том случае, когда $p = q$.
2. $d(p,q) = d(q,p)$;
3. $d(p,z) \leq d(q,p) + d(q,z)$.

Пусть теперь пиксели заданы своими координатами: $p(x,y)$, $q(s,t)$ и $z(v,w)$. Наиболее используемыми являются следующие функции расстояния.

*Эвклидова метрика*: $d_2(p,q)=\sqrt{(x-s)^2+(y-t)^2}$. В случае этой метрики пиксели, которые находятся на одинаковом расстоянии от центрального пикселя, образуют окружности.

*Метрика Манхэттена*: $d_1(p,q)=|x-s|+|y-t|$. Если принять расстояние между парой соседних пикселей на единицу, то распределение расстояний от центрального пикселя до ближайших к нему пикселей выглядит так, как показано на рисунке внизу слева.

*Бесконечная метрика*: $d_\infty(p,q)=\max\left(|x-s|,|y-t|\right)$. Распределение расстояний от центрального пикселя выглядит в случае такой метрики как показано на рисунке справа.

| | | | | |
|---|---|---|---|---|
| | | 2 | | |
| | 2 | 1 | 2 | |
| 2 | 1 | 0 | 1 | 2 |
| | 2 | 1 | 2 | |
| | | 2 | | |

| | | | | |
|---|---|---|---|---|
| 2 | 2 | 2 | 2 | 2 |
| 2 | 1 | 1 | 1 | 2 |
| 2 | 1 | 0 | 1 | 2 |
| 2 | 1 | 1 | 1 | 2 |
| 2 | 2 | 2 | 2 | 2 |

## 2.6. Операции и преобразования на изображениях

Над изображениями выполняются в основном так называемые *поэлементные* и *попиксельные* операции над соответствующими элементами одного или большего числа изображений. При этом изображения рассматриваются как матрицы, о чем говорилось выше.

Можно говорить об *унарных* поэлементных операциях. Например, возведение каждого элемента изображения (т.е. яркости соответствующего пикселя) в степень или логарифмирование каждого элемента. Другим примером является умножение изображения на число: если $\mathrm{I}=[a_{ij}]$ и $\alpha$ – действительное число, то $\alpha\mathrm{I}=[\alpha a_{ij}]$.

В случае *бинарных* операций подразумеваются бинарные поэлементные операции над матрицами одного и того же размера. Пусть имеется два изображения $\mathrm{I}_1=[a_{ij}]$ и $\mathrm{I}_2=[b_{ij}]$ одного размера. Тогда можно определить следующие операции.

*Суммой* изображений: $\mathrm{I}_3=\mathrm{I}_1+\mathrm{I}_2=[a_{ij}+b_{ij}]$.

*Разность* изображений: $\mathrm{I}_3=\mathrm{I}_1-\mathrm{I}_2=[a_{ij}-b_{ij}]$.

*Произведение* изображений: $\mathrm{I}_3 = \mathrm{I}_1 \cdot \mathrm{I}_2 = [a_{ij} b_{ij}]$.

*Частное* изображений: $\mathrm{I}_3 = \mathrm{I}_1 / \mathrm{I}_2 = [a_{ij} / b_{ij}]$. Эта операция может быть применена, если в изображении $\mathrm{I}_2$ нет нулевых элементов.

Понятно, что результирующие изображения будут того же размера. Данные операции могут быть естественным образом распространены на любое количество матриц. Эти операции иногда называют *арифметическими операциями* над изображениями.

Сложение изображений, в частности усреднение, используется для уменьшения уровня шумов. Вычитание – для фиксации, например, изменений в наблюдаемой картине. Если камера в одном и том же положении делает ряд снимков, то по разности изображений можно наблюдать изменения, происходящие на наблюдаемой сцене.

В случае поэлементных операций над изображениями можно определить и множество других операций, которые в принципе можно осуществить над парой (или большим количеством) чисел. Например, можно определить полезные операции вычисления максимума или минимума изображений:

$$\max\left(\mathrm{I}_1, \mathrm{I}_2\right) = \left[\max\left(a_{ij}, b_{ij}\right)\right],$$
$$\min\left(\mathrm{I}_1, \mathrm{I}_2\right) = \left[\min\left(a_{ij}, b_{ij}\right)\right].$$

Попиксельные операции, с одной стороны, очень просты, а с другой стороны оказываются весьма полезными при работе с изображениями. Существуют и более сложные операции над изображениями, некоторые из которых будут определены в последующих разделах курса.

Следует заметить, что в результате выполнения арифметических и других операций значения получившихся яркостей могут выйти за допустимый диапазон. В этом случае осуществляется нормировка новых значений. Это можно сделать многими способами. Наиболее часто применяемый способ нормировки состоит в следующем.

Изображение $f(x,y)$ сначала преобразуется по правилу:

$$f_1(x,y) = f(x,y) - \min_{(x,y)} f(x,y).$$

Затем $f_2(x,y) = (L-1)\dfrac{f_1(x,y)}{\max\limits_{(x,y)} f_1(x,y)}$. Очевидно, значения яркости изображения $f_2(x,y)$ будут лежать в исходном диапазоне $[0, L-1]$.

Важной характеристикой операции над изображениями является ее линейность. Оператор $A$, действующий на изображения, называется *линейным*, если он удовлетворяет следующим условиям:

1. $A(I_1 + I_2) = AI_1 + AI_2$ для любых изображений $I_1$ и $I_2$ одинакового размера (*аддитивность*);

2. $A(\alpha I) = \alpha AI$ Для любого изображения $I$ и любого действительного числа $\alpha$ (*однородность*).

Это определение эквивалентно условию: $A(\alpha_1 I_1 + \alpha_2 I_2) = \alpha_1 AI_1 + \alpha_2 AI_2$ для любых действительных $\alpha_1$ и $\alpha_2$.

Приведем в качестве примера еще одно важное преобразование изображений. Пусть $\varphi$ – строго монотонно возрастающее преобразование, определенное на множестве $[0, L-1]$ значений яркостей, принимаемых пикселями, такое, что $\varphi(0) = 0$ и $\varphi(L-1) = L-1$, т.е. для любых яркостей $a_1$ и $a_2$ таких, что $a_1 < a_2$, выполняется $\varphi(a_1) < \varphi(a_2)$. Тогда *строго монотонным преобразованием* изображения $\mathrm{I} = [a_{ij}]$ называется преобразование $\varphi\mathrm{I} = \left[\varphi(a_{ij})\right]$.

Нетрудно видеть, что при таком преобразовании во вновь полученном изображении $J = \varphi\mathrm{I}$ для любой пары пикселей сохраняются те же отношения яркостей (больше или меньше), которые были и в исходном изображении. Еще одной важной особенностью таких преобразований является то, что они в математическом смысле образуют группу относительно операции обычной композиции функций. Для этого достаточно вспомнить, что преобразование, обратное к строго монотонному, также будет строго монотонным, и принять во внимание, что единичным элементом в этой группе будет тождественное преобразование. Но последнее справедливо только в том случае, если рассматриваются непрерывные (а не квантованные!) значения функции яркости.

В последующем мы познакомимся с большим количеством различных преобразований цифровых изображений, ориентированных на решение конкретных задач обработки.

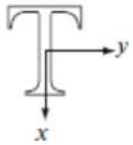
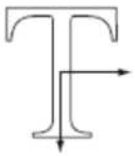
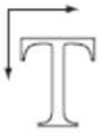
## 2.7. Геометрические преобразования изображений

Геометрические преобразования предполагают выполнения двух последовательных этапов: (1) собственно преобразование изображения и (2) интерполяция значений функций яркости в пикселях преобразованного изо-

бражения. Преобразование изображения сводится к преобразованию координат пикселей.

Наиболее широкий класс преобразований представляют аффинные преобразования. В общем случае они могут быть записаны следующим образом:

$$\begin{pmatrix} x & y & 1 \end{pmatrix} = \begin{pmatrix} v & w & 1 \end{pmatrix} \begin{pmatrix} t_{11} & t_{12} & 0 \\ t_{21} & t_{22} & 0 \\ t_{31} & t_{32} & 1 \end{pmatrix}.$$

| Название преобразования | Аффинная матрица T | Преобразование координат | Пример |
|---|---|---|---|
| Тождественное преобразование | $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ | $x = v$<br>$y = w$ |  |
| Изменение масштаба | $\begin{bmatrix} c_x & 0 & 0 \\ 0 & c_y & 0 \\ 0 & 0 & 1 \end{bmatrix}$ | $x = c_x v$<br>$y = c_y w$ |  |
| Поворот | $\begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$ | $x = v\cos\theta - w\sin\theta$<br>$y = v\sin\theta + w\cos\theta$ | |
| Параллельный перенос (сдвиг) | $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ t_x & t_y & 1 \end{bmatrix}$ | $x = v + t_x$<br>$y = w + t_y$ |  |
| Вертикальный скос | $\begin{bmatrix} 1 & 0 & 0 \\ s_v & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ | $x = v + s_v w$<br>$y = w$ | |
| Горизонтальный скос | $\begin{bmatrix} 1 & s_h & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ | $x = v$<br>$y = s_h v + w$ | |

С помощью приведенных преобразований можно осуществлять изменение масштаба изображения, повороты, сдвиги и скос (пропорциональный сдвиг). Каждое из этих преобразований зависит от значений элементов матрицы преобразования. Если комбинировать указанные преобразования, то можно строить самые разнообразные преобразования изображения.

После указанных преобразований необходимо вычислить значения новые яркости пикселей на новом изображении. Данная задача решается с по-

мощью интерполяционных методов – интерполяцию по ближайшему соседу, билинейной интерполяции или бикубической интерполяции.

Опишем еще одно полезное представление изображений, которое позволяет более строго в математическом смысле интерпретировать преобразование изображений. Представим изображение размера $M \times N$ в виде $MN$-мерного вектора $\mathbf{f}$, считая, что первые $N$ компонент вектора образованы первой строкой изображения, вторые $N$ компонент – второй и т.д. Тогда преобразование изображения можно рассматривать в виде $\mathbf{g} = \mathbf{Tf} + \mathbf{n}$, где $\mathbf{n}$ – вектор, соответствующий шуму. Такие представления оказываются очень удобными для анализа шума и восстановления зашумленных изображений.

## 2.8. Двумерные линейные преобразования изображений

Пусть $f(x,y)$ - исходное изображение. Двумерным линейным преобразованием называется преобразование вида:

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

где $r(x,y,u,v)$ – числа, которые называются *ядром прямого преобразования*. Линейность этого преобразования очевидна. Здесь переменные $u$ и $v$ принимают такие же значения, что и $x$ и $y$ соответственно. Таким образом, для вычисления данного преобразования нужно вычислить значения $T(u,v)$ для всех $u = 1,2,...,M-1$ и $v = 1,2,...,N-1$. $T(u,v)$ называется *прямым преобразованием* изображения $f(x,y)$, а $u$ и $v$ – переменными преобразования.

Особую роль играют так называемые обратимые преобразования, а именно такие, которые позволяют по известным значениям $T(u,v)$ восстановить значения исходной функции. В этом случае можно говорить об обратном преобразовании, которое имеет точно такую же форму, как и прямое:

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

где $s(x,y,u,v)$ – *ядро обратного преобразования*.

| | |
|---|---|
| **Определение** | Ядро преобразования называется *разделимым* или *сепарабельным*, если оно представимо в виде: $$r(x,y,u,v) = r_1(x,y) r_2(u,v).$$ |

| | |
|---|---|
| **Определение** | Ядро называется *симметричным*, если $r_1(x,y)$ и $r_2(u,v)$ представляются одной и той же функцией, т.е. $r_2(u,v)=r_1(u,v)$. |

Двумерные линейные преобразования играют исключительно важную роль в частотных методах обработки изображений. При задании конкретных ядер прямых преобразований получаются многие полезные и широко используемые преобразования изображений. С некоторыми из них мы познакомимся далее.

## 2.9. Вероятностные методы описания изображений

В обработке изображений широко используются вероятностные и статистические методы. Рассмотрим простейшие вероятностные характеристики цифровых изображений.

Предположим, как и раньше, что изображение имеет размеры $M \times N$. Пусть все возможные значения яркостей задаются множеством $\{0,1,2,...,L-1\}$. Обозначим через $n_k$ количество пикселей изображения, имеющих яркость $z_k$. Тогда вероятность появления пикселя с яркостью $z_k$ может быть оценена частотой его появления $p_k = \dfrac{n_k}{MN}$, $k = 0,1,2,...,L-1$. Очевидно, что $\sum_{k=0}^{L-1} p_k = 1$. Таким образом, мы рассматриваем случайную величину $z$ (имеющую смысл яркости) с указанным распределением вероятностей.

Зная распределение вероятностей яркостей пикселей, можно вычислить обычные числовых характеристики. Например, *математическое ожидание* равно $M[z] = \sum_{k=0}^{L-1} z_k p_k$ и имеет смысл среднего значения яркостей по всему изображению.

Аналогично можно вычислить *дисперсию*: $D[\mathrm{z}] = \sum_{k=0}^{L-1} \left(z_k - M[z]\right)^2 p_k$. Дисперсия служит мерой разброса яркостей относительно среднего значения. Однако практически более удобной мерой разброса является *среднеквадратическое* (или *стандартное*) отклонение: $\sigma[\mathrm{z}] = \left[\sum_{k=0}^{L-1} \left(z_k - M[z]\right)^2 p_k\right]^{\frac{1}{2}}$. Это

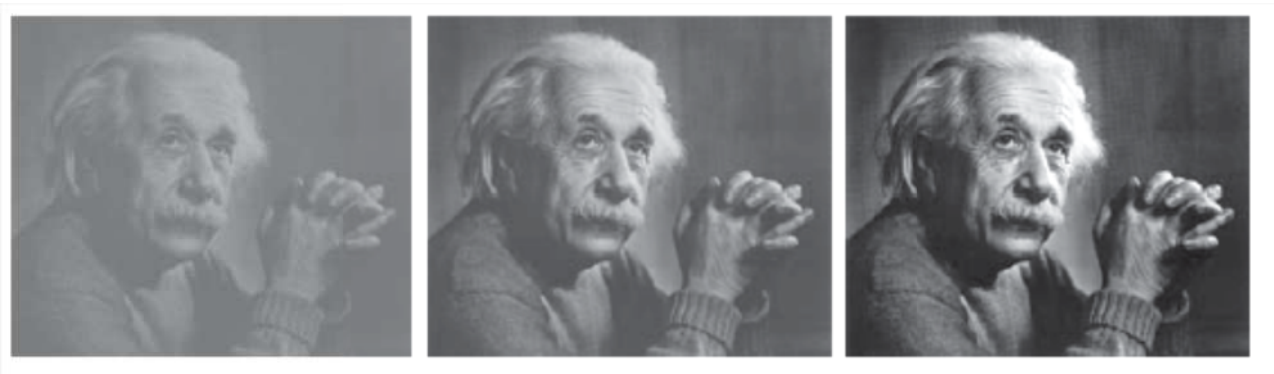
объясняется тем, что среднеквадратическое отклонение имеет ту же размерность, что и случайная величина $z$ и поэтому значительно легче поддается интерпретации.

Можно ввести также и числовые характеристики более высоких порядков. Например, начальные и центральные моменты $m$ -го порядка:

$$\alpha[\mathrm{z}] = \sum_{k=0}^{L-1} z_k^m p_k \,, \qquad D[\mathrm{z}] = \sum_{k=0}^{L-1} \left(z_k - M[z]\right)^m p_k \,, \qquad m = 3, 4, \ldots$$

Такие моменты для $m \geq 3$ значительно труднее поддаются интерпретации.

Стандартное отклонение, характеризуя меру разброса яркостей пикселей, является удобной мерой контрастности изображения. На следующем рисунке показаны три изображения, отличающиеся величиной стандартного отклонения.



На этом рисунке стандартные отклонения имеют значения соответственно 14.3, 31.6 и 49.2 уровней яркости.

Вероятностно-статистические методы чрезвычайно широко используются при обработке изображений.

## РАЗДЕЛ 2

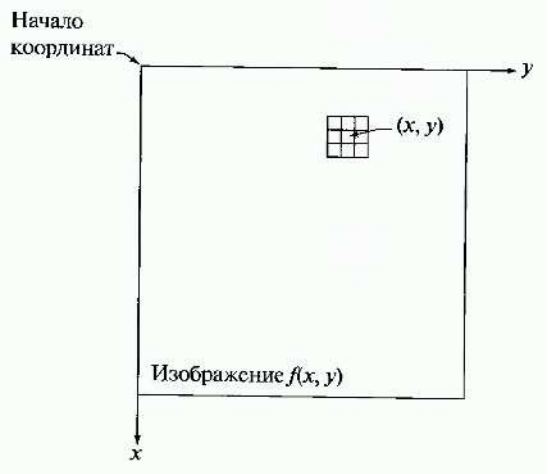
# ПРОСТРАНСТВЕННЫЕ МЕТОДЫ УЛУЧШЕНИЯ ИЗОБРАЖЕНИЙ

## ГЛАВА 3. ПРОСТРАНСТВЕННЫЕ ПРЕОБРАЗОВАНИЯ ИЗОБРАЖЕНИЙ

### 3.1. Общие принципы пространственных преобразований

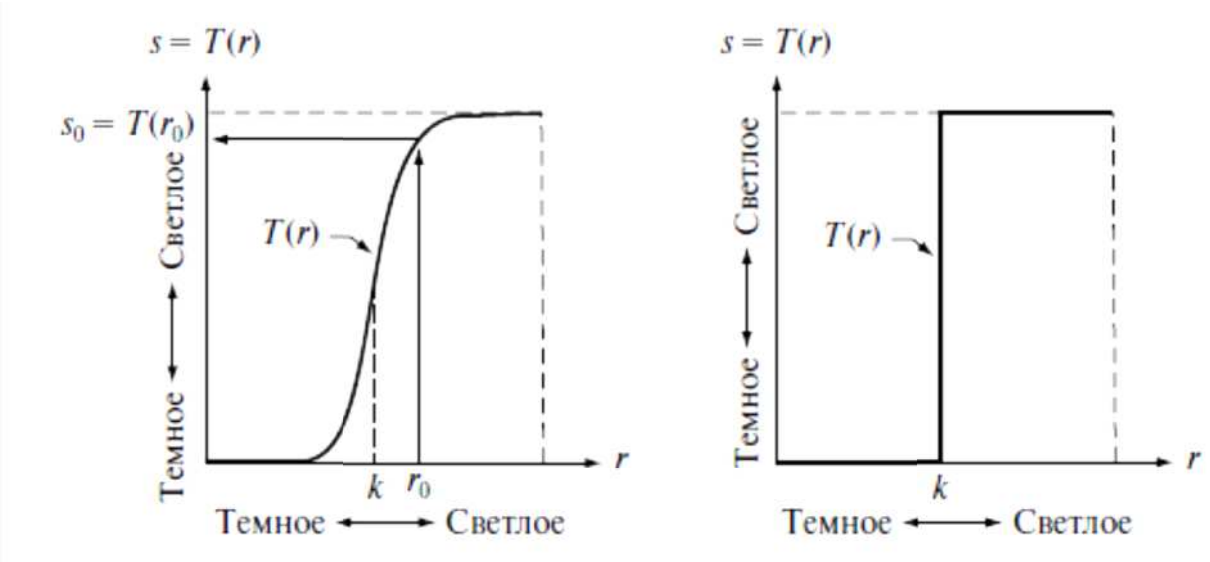
Пространственные методы обработки изображений подразумевают непосредственное преобразование изображений, как правило, с помощью изменения значения каждого пикселя (или какой-то совокупности пикселей) на основе яркостей пикселей из некоторой окрестности. Эти методы используются для улучшения качества изображений, но строгих критериев такого улучшения не существует. Оценки в этом случае носят, как правило, субъективный характер.

В общем случае процедуру пространственного преобразования можно записать следующим образом. Пусть $f$ – исходное изображение, $T$ – некоторый оператор, действие которого определено в некоторой окрестности рассматриваемой точки изображения, а $g$ – изображение, полученное в результате обработки этим оператором. Тогда это можно записать так: $g = T(f)$. Эта запись подразумевает, что определена окрестность, в которой действует оператор $T$; эта окрестность последовательно перемещается по изображению, и каждый раз оператор применяется к пикселям этой окрестности, вычисляя значение каждого пикселя изображения.



Обычно окрестность представляет собой квадрат 3x3 или 5x5. Иногда такая окрестность не квадратная, а прямоугольная. В отдельных случаях окрестность может иметь даже форму круга (точнее, форму совокупности пикселей, покрывающих круг). В тех случаях, когда окрестность выходит за пределы изображения, недостающие пиксели либо игнорируются, либо доопределяются некоторым образом. Такое преобразование называют *пространственной фильтрацией*.

Наиболее простой случай, когда окрестность имеет размер 1x1. Очевидно, в этом случае новое значение яркости пикселя $s$ зависит только от его значения $r$ в исходном изображении и определяется введенным преобразованием $s = T(r)$. На следующем рисунке приведены два примера такого преобразования. На нем представлена функция повышения контраста и пороговая функция.




Такие методы по понятным причинам называются методами *поэлементной обработки* или методами *градационных преобразований*.

## 3.2. Основные градационные преобразования

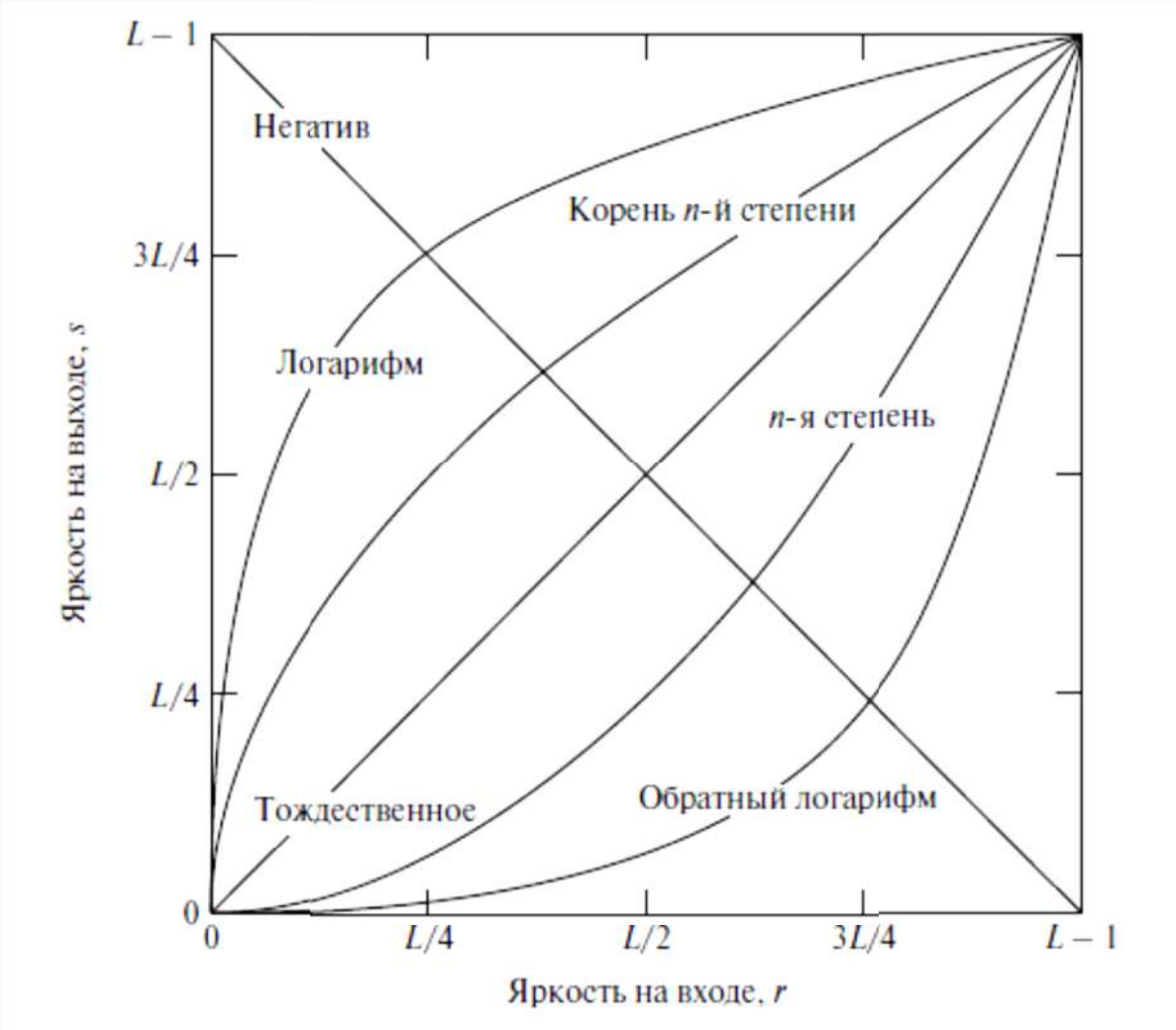
Рассмотрим основные яркостные преобразования, смысл которых иллюстрируется на приведенном ниже рисунке.

**Негативное изображение**

Преобразование изображения в *негативное* осуществляется простейшим преобразованием:

$$s = L - 1 - r.$$

Во многих случаях негативное изображение позволяет лучше рассмотреть изображение, особенно в тех случаях, когда на исходном изображении преобладают либо темные, либо светлые тона.




**Логарифмическое преобразование**

*Логарифмическое преобразование* осуществляется по формуле:

$$s = c\lg(1+r).$$

В этой формуле $c$ – это нормирующая константа. Данное преобразование «растягивает» малые значения яркостей в более широкий диапазон, и наоборот, «сжимает» диапазон ярких пикселей.

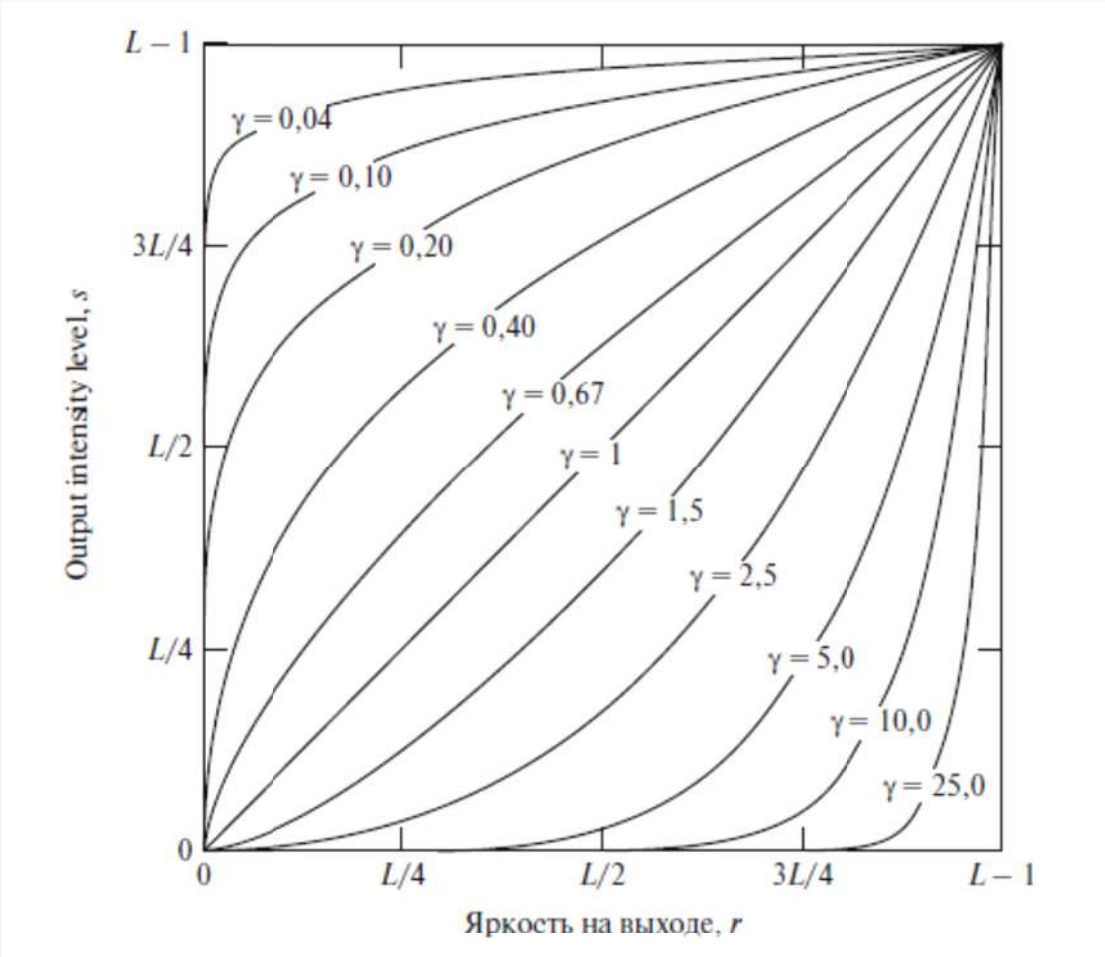
Обратное *логарифмическое преобразование* или показательное *преобразование* действует в точности противоположным образом. Растягивает диапазон ярких пикселей и сжимает диапазон темных. Эти преобразования позволяют более детально рассмотреть либо темные части изображений, либо, наоборот, светлые.

**Степенное преобразование**

Во многих случаях применение логарифмического преобразования ограничено, так как оно недостаточно параметризовано. Очень широкий класс преобразований представляют *степенные преобразования*. В общем виде оно задается следующим образом:

$$s = cr^{\gamma}.$$

Здесь $c$ и $\gamma$ – неотрицательные константы, параметры степенного преобразования. Степенное преобразование действует так же, как и логарифмическое, но благодаря параметру $\gamma$ возникает целое семейство кривых, отличающихся характером роста. На рисунке ниже представлены степенные кривые для разных значений этого параметра. Качественный характер этих кривых разный в зависимости от того является ли $\gamma$ большим единицы или меньшим единицы.



Следует отметить, что в качестве обозначения параметр $\gamma$ выбран не случайно. Большинство средств визуализации по техническим причинам имеют искажения, которые корректируются с помощью степенного преобразования. Поэтому выбор такого обозначения сложился исторически. Процедуру такой коррекции принято называть *гамма-коррекцией*. Степенные преобразования оказываются очень эффективными для повышения контрастности изображений.

**Кусочно-линейное преобразование**

Кусочно-линейное преобразование задается в общем случае следующим образом. Полный диапазон яркостей $[0, L-1]$ разбивается определенным образом $k$ точками на $k+1$ поддиапазонов, в каждом из которых задается линейный закон изменения яркостей. Таким образом, преобразование имеет вид:

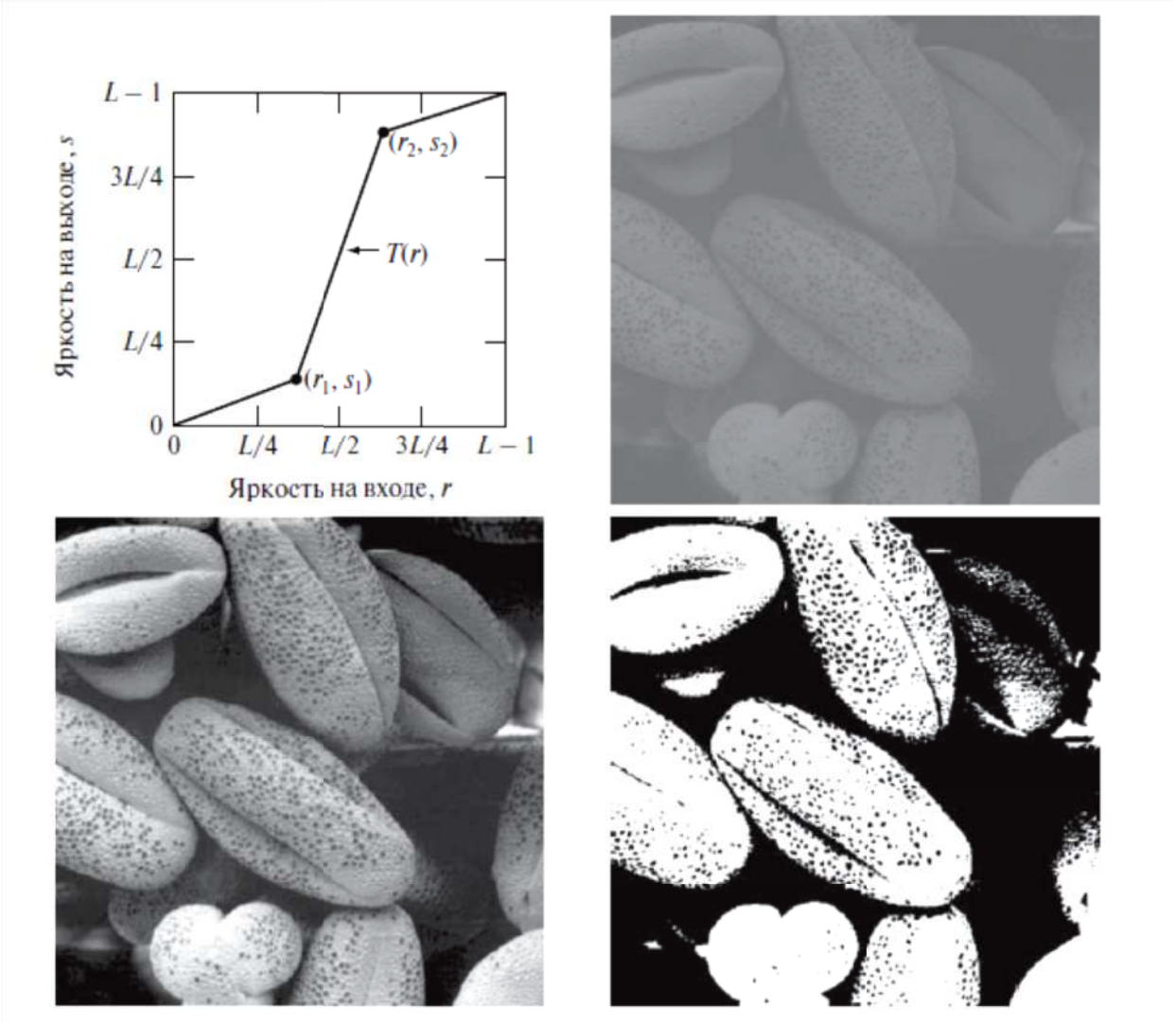
$$s = \begin{cases} a_0 r + b_0, & 0 \le r < r_1, \\ a_1 r + b_1, & r_1 \le r < r_2, \\ \dots & \dots \\ a_k r + b_k, & r_k \le r \le L-1. \end{cases}$$

Здесь $a_i, b_i$ и $r_i$, $i = 0,1,2,...,k$, – параметры кусочно-линейного преобразования.

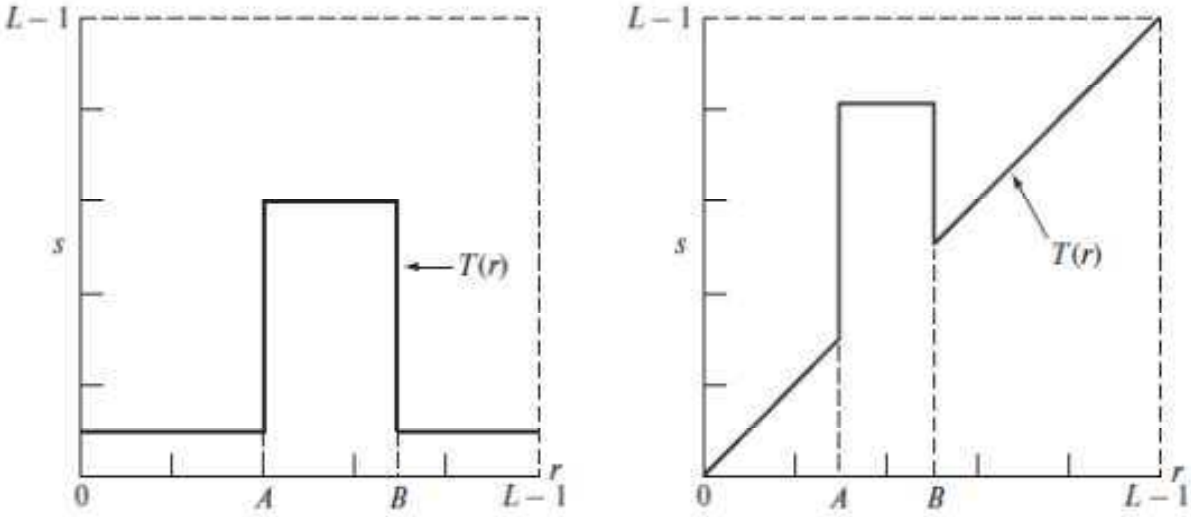
С помощью такого преобразования можно задать сколь угодно сложный закон изменения яркостей. Однако для задания такого преобразования требуется значительно большее число параметров. Типичные примеры применения такого преобразования – следующие.

**Усиление контраста.** Как правило, низкий контраст – это следствие недостаточно широкого диапазона яркостей. Соответственно, усиление контраста это расширение диапазона яркостей. Следующий пример иллюстрирует применение кусочно-линейного преобразования для коррекции контраста.

**Вырезание диапазона яркостей.** В этом случае можно либо бинаризовать изображение, как показано на следующем рисунке слева, либо увеличить (или уменьшить) яркость пикселей из некоторого диапазона, оставив неизменными яркости остальных пикселей, как на рисунке справа.




*Усиление контраста*




*Вырезание диапазона яркостей*

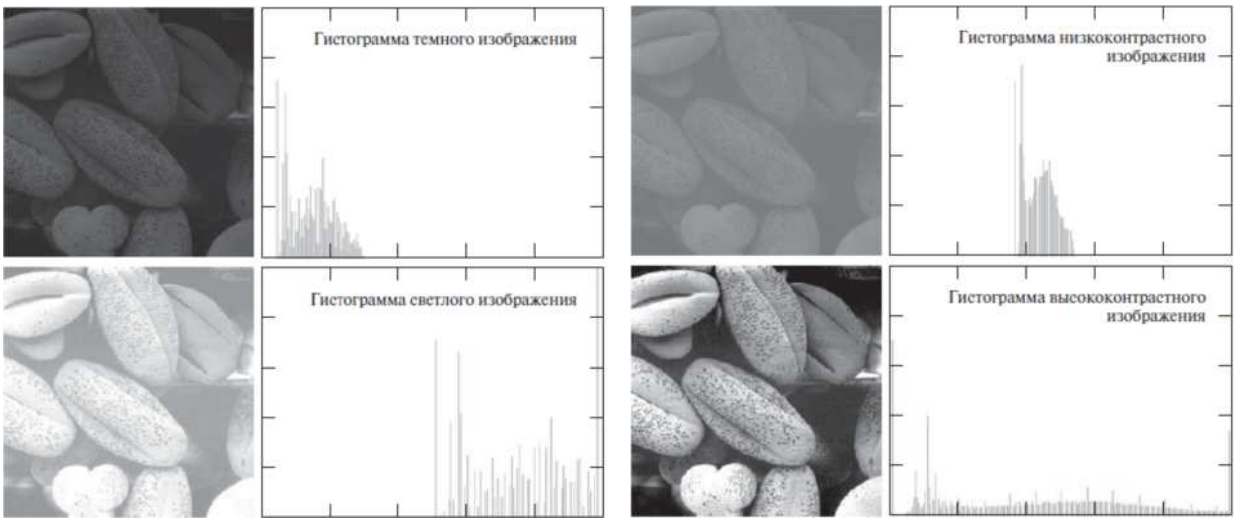
# Глава 4. Гистограммные преобразования

## 4.1. Основные понятия и определения

| | |
|---|---|
| **Определение** | *Гистограммой* изображения называется дискретная функция $h(r_k) = n_k$, $k = 0,1,2,....L-1$, где $r_k$ – $k$-й уровень яркости, $n_k$ – число пикселей изображения, имеющих яркость $r_k$. |

| | |
|---|---|
| **Определение** | *Нормированной гистограммой* изображения называется дискретная функция $p(r_k) = \frac{n_k}{MN}$, $k = 1,2,....L-1$, где $r_k$ – $k$-й уровень яркости, $n_k$ – число пикселей изображения, имеющих яркость $r_k$, $M$ и $N$ – число строк и столбцов изображения соответственно. |

Значение $p(r_k) = \frac{n_k}{MN}$ можно интерпретировать как оценку вероятности появления пикселя с яркостью $r_k$. Очевидно, $\sum_{k=0}^{L-1} \frac{n_k}{MN} = 1$. Оба вида гистограмм являются стандартными методами пространственной обработки изображений для улучшения восприятия изображений. Следующий пример показывает, как на гистограммах отражаются особенности градационных преобразований.



## 4.2. Эквализация гистограмм

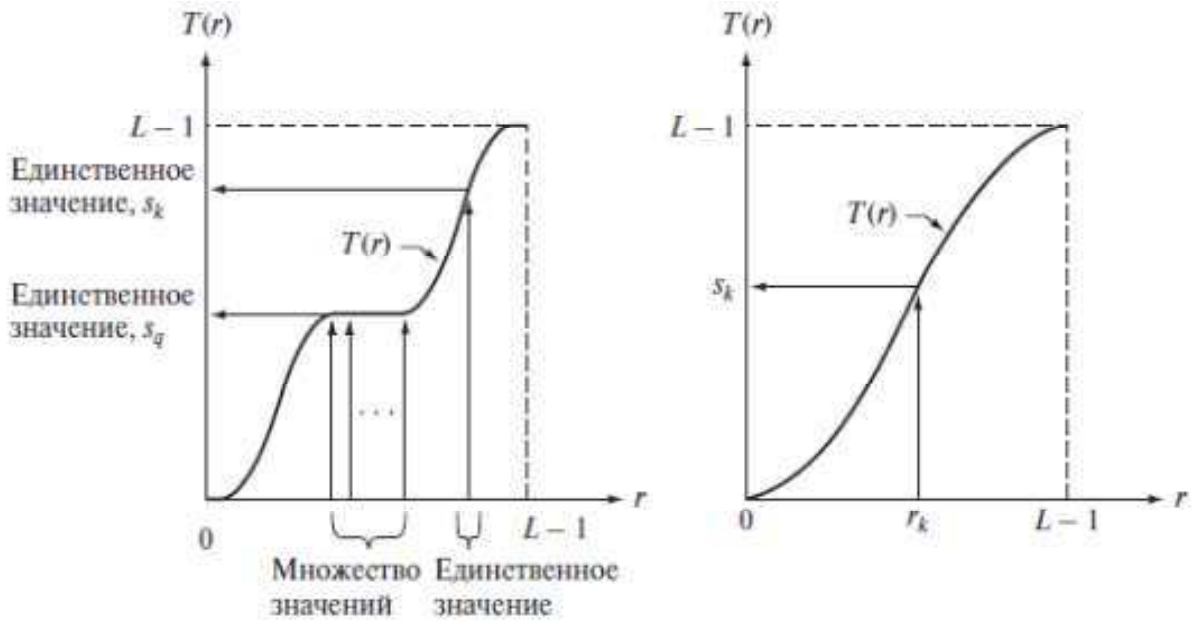
Рассмотрим важную в гистограммных методах процедуру эквализации гистограмм. Предположим, что яркости пикселей принимают непрерывные значения на отрезке $[0, L-1]$ и рассмотрим на этом отрезке *монотонно неубывающее* преобразование

$$s = T(r),$$

такое, что для всех $0 \le r \le L-1$ выполняются неравенства $0 \le T(r) \le L-1$. Если данное преобразование является *строго монотонно возрастающим*, то существует обратное преобразование

$$r = T^{-1}(s).$$

На следующем рисунке показаны примеры монотонно неубывающего и строго монотонно возрастающего преобразований.




В случае не непрерывных, а дискретных значений яркости, добиться строгой монотонности преобразования, очевидно, практически невозможно. Поэтому рассматриваются округления функции яркости до ближайших целых значений.

Заметим теперь, что яркости пикселей можно рассматривать как значения, принимаемые случайной величиной на отрезке $[0, L-1]$. Тогда нормированную гистограмму можно рассматривать как оценку *функции плотности вероятности* этой случайной величины. Пусть есть две случайные величины

$r$ и $s$ с плотностями вероятностей $p_r(r)$ и $p_s(s)$ соответственно, связанные соотношением $s = T(r)$. Как известно, если функция $T(r)$ является непрерывной и дифференцируемой на $[0, L-1]$, то функции плотностей вероятностей связаны выражением:

$$p_s(s) = p_r(r)\left|\frac{dr}{ds}\right|.$$

Рассмотрим теперь преобразование вида

$$s = T(r) = (L-1)\int_0^r p_r(w)dw.$$

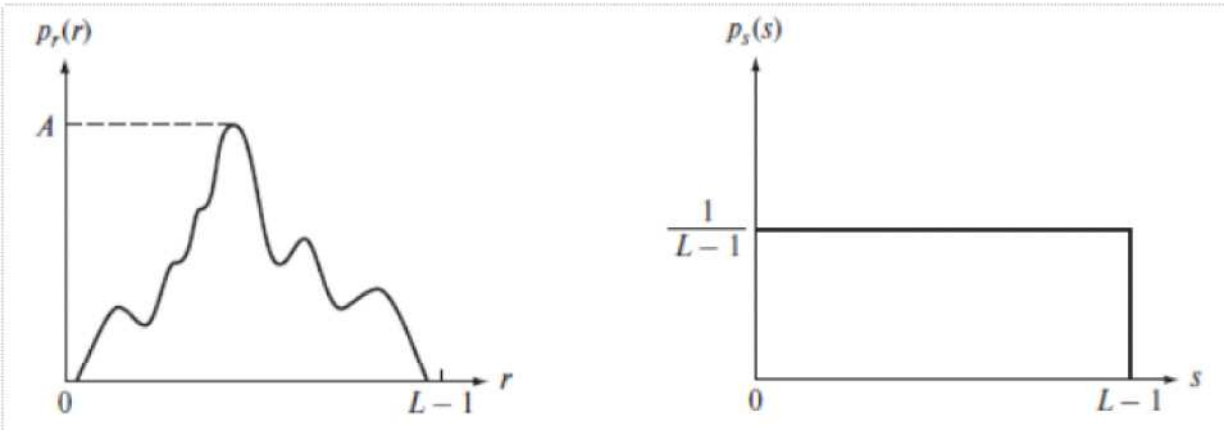
Нетрудно видеть, что справа стоит кумулятивная функция распределения вероятностей случайной величины $r$. Нетрудно убедиться, что данное преобразование удовлетворяет всем требованиям, предъявленным выше к преобразованию $T(r)$. Найдем $\frac{dr}{ds}$. Поскольку из равенства функций следует равенство их дифференциалов, т.е. $ds = d\left(T(r)\right) = T'(r)dr$, то получаем, что

$$\frac{ds}{dr} = T'(r) = \left((L-1)\int_0^r p_r(w)dw\right)' = (L-1)p_r(r).$$

Следовательно,

$$p_s(s) = p_r(r)\left|\frac{dr}{ds}\right| = p_r(r)\left|(L-1)p_r(r)\right|^{-1} = \frac{1}{L-1}.$$

На следующем рисунке показан пример такого преобразования функций плотности вероятности.



Важно отметить, что полученная плотность вероятности случайной величины $s$ соответствует равномерному распределению вероятностей. Особо надо подчеркнуть, что плотность равномерного распределения получается при указанном задании преобразования $T(r)$ и при *любой* плотности распределения $p_r(r)$.

Предыдущие рассуждения относились к случаю непрерывных яркостей. В реальности они принимают конечное множество значений $0,1,2,...,L-1$. Поэтому приходится иметь дело не с непрерывными функциями плотности вероятности, а с дискретными, а также не непрерывными случайными величинами, а с дискретными. Поэтому плотность вероятности задается лишь оценкой

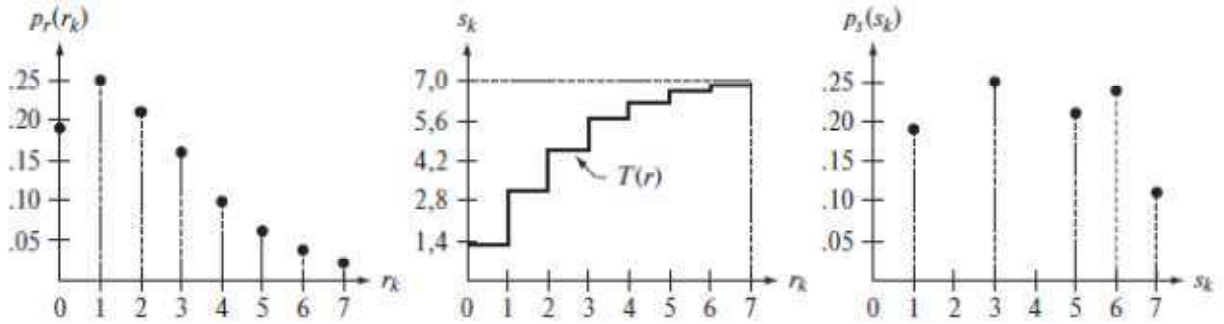
$$p(r_k) = \frac{n_k}{MN},\ k = 0,1,2,....L-1,$$

ее значений в точках, т.е. нормированной гистограммой. При этом выражение для функции задается своим дискретным аналогом:

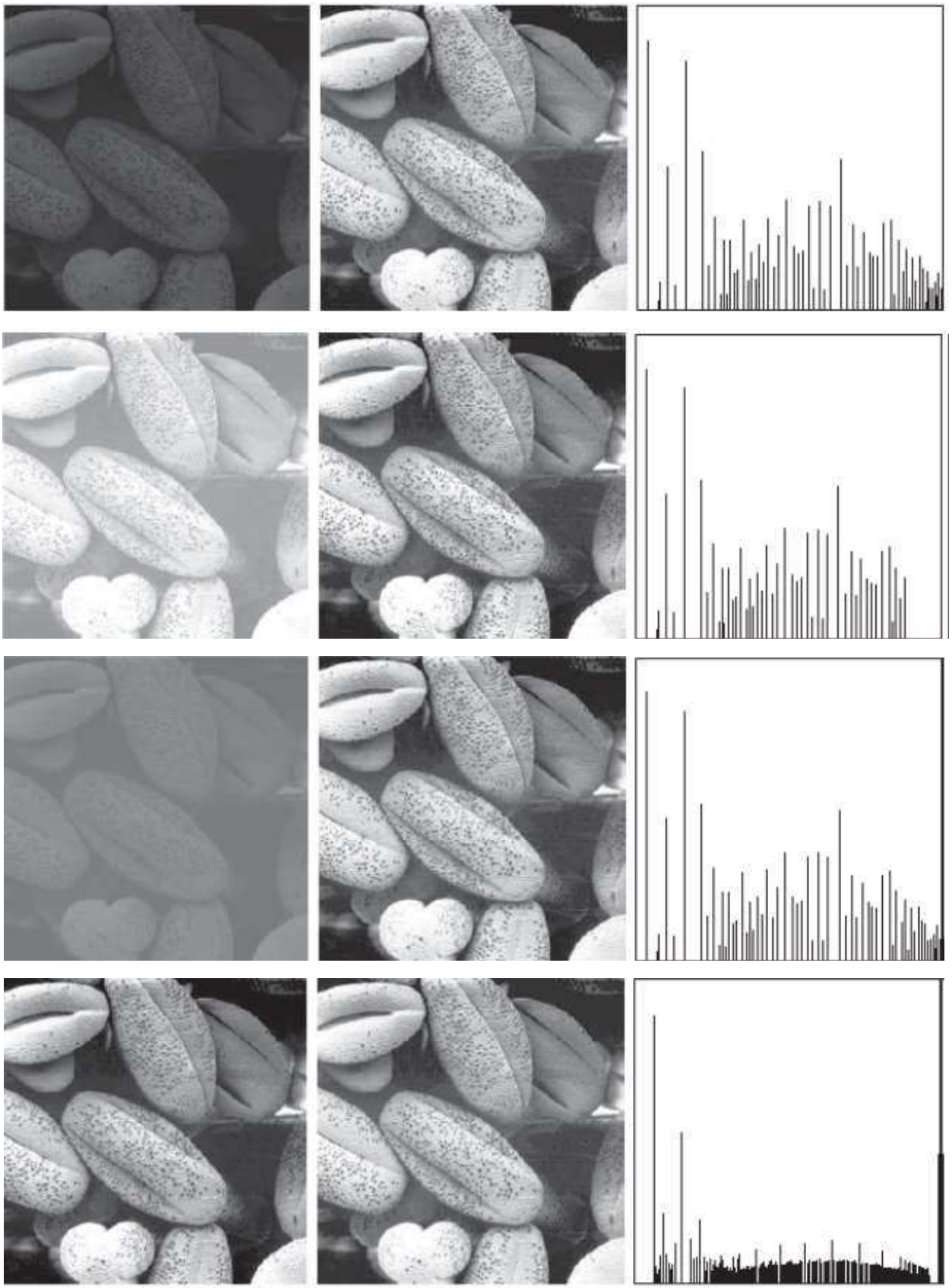
$$s_k = T(r_k) = (L-1)\sum_{j=0}^{k} p(r_j) = \frac{L-1}{MN}\sum_{j=0}^{k} n_j, \quad k = 0,1,2,....L-1.$$

Это выражение называется *эквализацией* или *линеаризацией гистограммы*.

Ввиду дискретности изображений идеально ровной гистограммы, соответствующей равномерному распределению, не получается. Однако распределение яркостей получается достаточно близким к нему. На следующем рисунке показан пример эквализации в случае квантованного изображения.




На следующих рисунках показаны реальные изображения, результат их эквализации и соответствующие им гистограммы.



## 4.3. Моделирование гистограммного распределения

Рассмотренная выше эквализация осуществляется с помощью преобразования, которое дает равномерное (или близкое к нему в случае дискретных значений) распределение яркостей пикселей. Однако в некоторых случаях необходимо сформировать не равномерное распределение, а наперед заданное. Покажем, как это можно сделать.

Рассмотрим снова непрерывный случай. Пусть заданы две непрерывные случайные величины $r$ и $z$, имеющие, как и раньше, смысл яркостей пикселей изображения. Случайная величина $r$ имеет плотность вероятности $p_r(r)$ и характеризует яркости исходного изображения. А $z$ имеет плотность вероятности $p_z(z)$, которая задана и по условию должна характеризовать распределение яркостей нового изображения.

Применим к $r$ преобразование эквализации. Получим новую случайную величину $s$ с равномерным распределением:

$$s = T(r) = (L-1)\int_0^r p_r(w)dw .$$

Если теперь применить такое же преобразование к заданному распределению $p_z(z)$, то мы тоже получим равномерно распределенную случайную величину, которая будет совпадать с $s$:

$$s = G(z) = (L-1)\int_0^z p_z(t)dt .$$

Отсюда следует, что $G(z) = T(r)$, поэтому $z = G^{-1}(s) = G^{-1}\left[T(r)\right]$. Заметим, что $T(r)$ получается по исходному изображению, а $G(z)$ – по последней формуле, поскольку предполагается, что плотность $p_z(z)$ задана аналитически (или еще каким либо образом).

**Пример (для непрерывных распределений)**. Пусть распределение исходного изображения имеет вид:

$$p_r(r) = \frac{2r}{(L-1)^2},\ 0 \le r \le L-1.$$

Требуется найти преобразование, которое даст случайную величину $z$, распределенную по закону

$$p_z(z) = \frac{3z^2}{(L-1)^3},\ 0 \le r \le L-1.$$

**Решение.** Выполним эквализацию изображения:

$$s = T(r) = (L-1)\int_0^r p_r(w)dw = (L-1)\int_0^r \frac{2w}{(L-1)^2}dw = \frac{r^2}{L-1}.$$

Теперь проведем эквализацию с заданной плотностью $p_z(z)$:

$$G(z) = (L-1)\int_0^z p_z(t)dt = (L-1)\int_0^z \frac{3t^2}{(L-1)^3}dt = \frac{z^3}{(L-1)^2}.$$

Таким образом, $\frac{z^3}{(L-1)^2} = \frac{r^2}{L-1}$, откуда находим:

$$z = \sqrt[3]{(L-1)r^2}.$$

Изображение, полученное с помощью последнего градационного преобразования, будет иметь заданное распределение. ■

В дискретном случае, так же как при эквализации, будет достигаться лишь некоторое приближение к заданному распределению, но оно будет вполне достаточным для получения желаемого результата.

В дискретном случае эквализация исходного изображения осуществляется по формуле:

$$s_k = T(r_k) = (L-1)\sum_{j=0}^{k} p_r(r_j) = \frac{L-1}{MN}\sum_{j=0}^{k} n_j, \quad k = 0,1,2,....L-1.$$

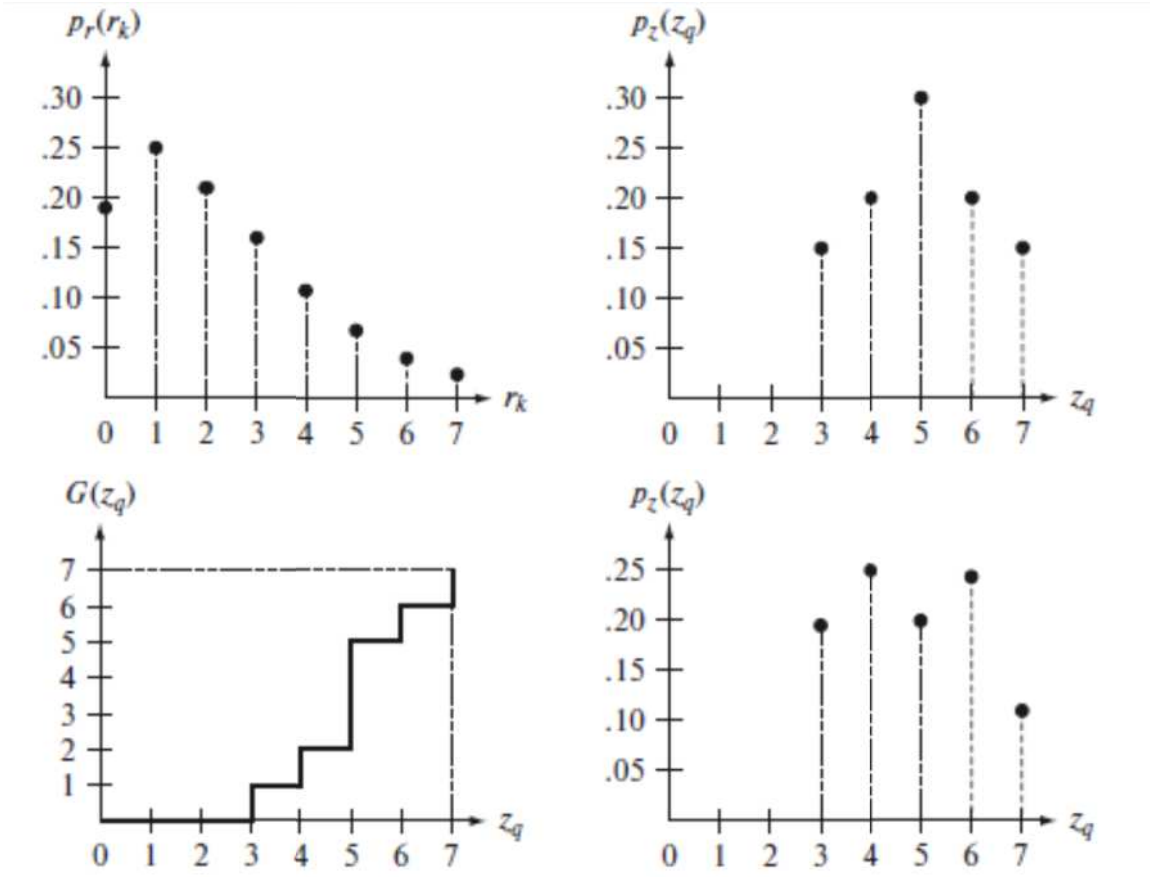
Аналогично

$$s_q = G(z_q) = (L-1)\sum_{i=0}^{q} p_z(z_i), \quad q = 0,1,2,....L-1.$$

В дискретном варианте преобразование $G(z_q)$ можно представить в виде таблицы соответствия значений $s_q$ и $z_q$. Тогда вычисление обратного преобразования сводится фактически к «переворачиванию» этой таблицы и применению ее к значениям $s_k$, т.е. $z_q = G^{-1}(s_k)$.

На следующем рисунке показаны графики преобразования распределения яркостей. Слева вверху – гистограмма исходного изображения (трехбитовое изображение с восемью градациями яркости). Справа вверху – гистограмма заданного распределения. Слева внизу – преобразование. Наконец,

справа внизу – результат применения преобразования к исходному изображению, т.е. гистограмма полученного изображения.




Данный пример демонстрирует приемлемое совпадение результата с заданным распределением.

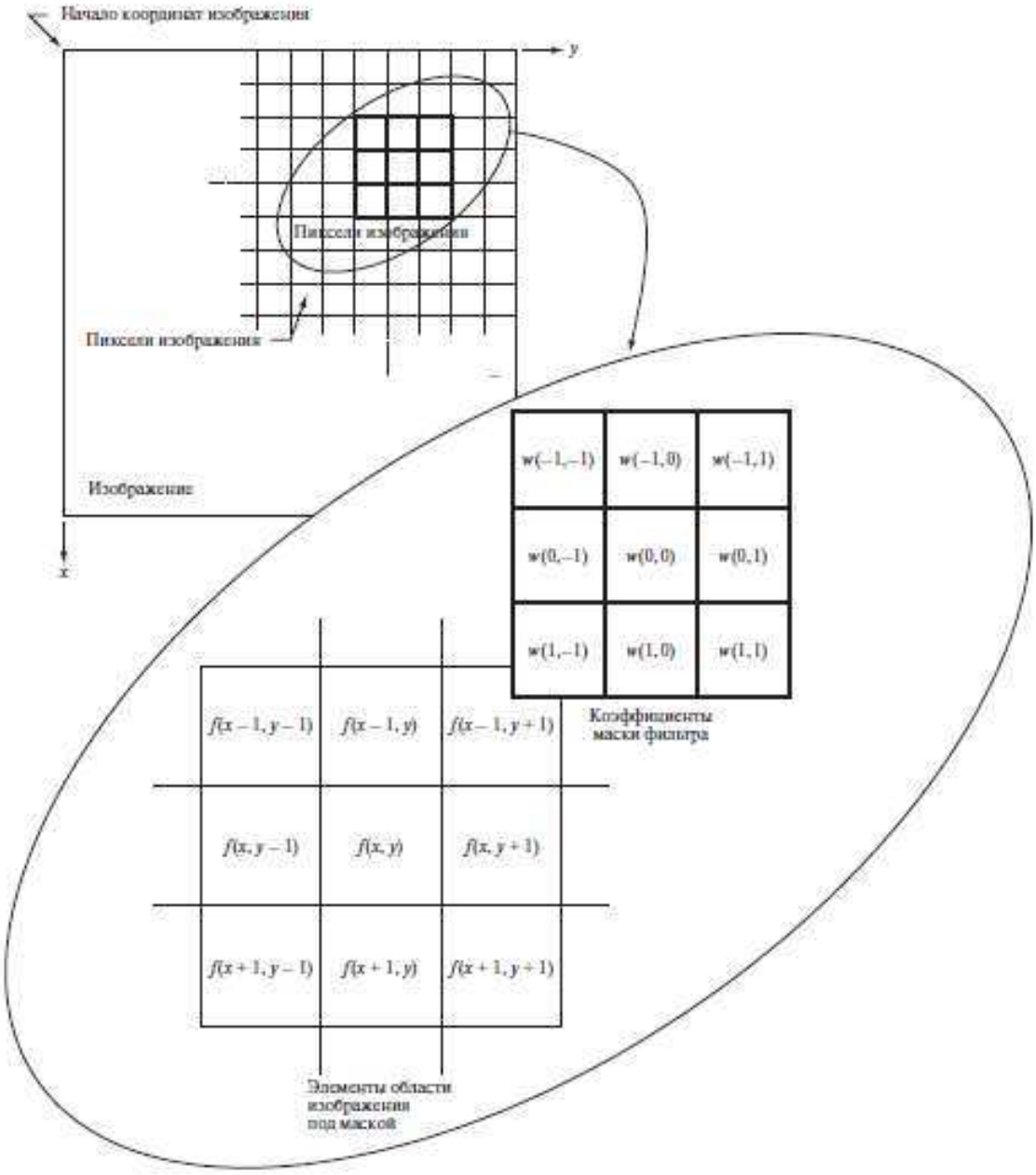
# Глава 5. Пространственная фильтрация изображений

Пространственная фильтрация – важный этап обработки изображений, который осуществляется с помощью так называемых пространственных фильтров. Ввиду не устоявшейся терминологии синонимичными терминами к слову *пространственный фильтр* являются *маска*, *ядро*, *шаблон*, *окно*.

## 5.1. Принципы пространственной фильтрации

| | |
|---|---|
| **Определение** | *Пространственным фильтром* называется совокупность *окрестности* (как правило, квадратной или прямоугольной формы) и *операции*, которая выполняется над пикселями из этой окрестности. |

Результатом фильтрации является новая яркость пикселя, который находится в центре окрестности. Процесс пространственной фильтрации состоит в последовательном движении окрестности по всему изображении и вычислении новых яркостей пикселей. Если операция, которая выполняется над пикселями из окрестности, является линейной, то и фильтр называется *линейным*. В противном случае фильтр называется *нелинейным*. На рисунке ниже представлен общий принцип линейной пространственной фильтрации.




Новое значение яркости пикселя в случае маски размером $3\times3$ вычисляется по формуле:

$$g(x,y) = w(-1,-1)f(x-1,y-1) + w(-1,0)f(x-1,y) + w(-1,1)f(x-1,y+1) + \ldots$$

$$... + w(0,0)f(x,y) + ... + w(1,0)f(x+1,y) + w(1,1)f(x+1,y+1).$$

Если же маска является прямоугольной и имеет размеры $m \times n$, где $m$ и $n$ – нечетные числа, удобно записывать их в виде $m = 2a+1$ и $n = 2b+1$. Тогда формула линейной фильтрации может быть записана в общем виде следующим образом:

$$g(x,y) = \sum_{s=-a}^{a} \sum_{t=-b}^{b} w(s,t) f(x+s, y+t).$$

## 5.2. Корреляция и свертка

Формулы линейной пространственной фильтрации распадаются на два основных класса – корреляция и свертка.

| | |
|---|---|
| **Определение** | *Корреляцией* фильтра $w(x,y)$ с изображением $f(x,y)$ называется новое изображение, получаемое по формуле $$g(x,y) = w \odot f = \sum_{s=-a}^{a} \sum_{t=-b}^{b} w(s,t) f(x+s, y+t).$$ |

Таким образом, рассмотренная в предыдущем разделе пространственная фильтрация является корреляцией.

| | |
|---|---|
| **Определение** | *Сверткой* фильтра $w(x,y)$ с изображением $f(x,y)$ называется новое изображение, получаемое по формуле $$g(x,y) = w * f = \sum_{s=-a}^{a} \sum_{t=-b}^{b} w(s,t) f(x-s, y-t).$$ |

Нетрудно видеть, что свертка и корреляция могут быть сведены друг к другу путем последовательного «переворота» фильтра симметрично относительно горизонтальной оси и симметрично относительно вертикальной оси. Поэтому выбор корреляции или свертки как способа реализации пространственной фильтрации – дело вкуса. Однако многие задачи естественным образом приводят к формулировке линейной фильтрации как корреляции. А в задачах частотной фильтрации, основанной на преобразовании Фурье, – к свертке. По этой причине ввиду неустоявшейся терминологии термин «свертка» не обязательно используется именно для свертки; под ним может

подразумеваться и скользящая корреляция, и, наоборот, под корреляцией может пониматься вычисление по формуле свертки.

В ряде случаев более удобной оказывается одноиндексная форма записи корреляции и свертки. В этом случае коэффициенты фильтра и соответствующие значения функции яркости нумеруются в некотором фиксированном порядке одним индексом. В результате формулы для корреляции и свертки принимают вид:

$$g(x,y) = w_1 z_1 + w_2 z_2 + \ldots + w_{mn} z_{mn} = \sum_{k=1}^{mn} w_k z_k = \mathbf{w}^T \mathbf{z},$$

где $m \times n$ – размеры фильтра.

## 5.3. Формирование фильтров

Для формирования пространственного фильтра размера $m \times n$ достаточно задать маску фильтра, т.е. $mn$ определяющих его коэффициентов. Например, для фильтра размера $3 \times 3$:

| | | |
|---|---|---|
| $w_1$ | $w_2$ | $w_3$ |
| $w_4$ | $w_5$ | $w_6$ |
| $w_7$ | $w_8$ | $w_9$ |

Коэффициенты данного линейного фильтра будут полностью определять те изменения, которые будут производиться над изображением. Например, для фильтра размером $3 \times 3$, если $w_1 = w_2 = \ldots = w_9 = \frac{1}{9}$, то формула пространственной фильтрации приобретает вид $\sum_{i=1}^{9} w_i z_i = \frac{1}{9} \sum_{i=1}^{9} z_i$, и сводится к усреднению по маске, т.е. значение яркости пикселя заменяется средней яркостью по окрестности этого пикселя размера $3 \times 3$.

Более сложные примеры фильтров предполагают, что коэффициенты фильтра различные. Это можно интерпретировать так, что яркости пикселей из рассматриваемой окрестности по-разному учитываются при формировании новой яркости. В частности, более удаленные пиксели могут учитываться с меньшим весом, чем более близкие.

Примером фильтра, который *не является линейным*, может служить фильтр, который присваивает пикселю максимальное (или минимальное) значение среди яркостей всех пикселей из окрестности.

# Глава 6. Сглаживающие пространственные фильтры

## 6.1. Линейные сглаживающие фильтры

Сглаживающие фильтры применяются, как правило, для

– удаления мелких деталей;

– уменьшения шума;

–расфокусировки изображения.

Такие фильтры называют также *усредняющими*, *сглаживающими*, *низкочастотными* или *интегральными* фильтрами.

Простейший усредняющий фильтр был приведен в качестве примера в предыдущем разделе. В общем случае для окрестности размера $m \times n$ он может быть записан в виде

$$z = \frac{1}{mn} \sum_{i=1}^{mn} z_i \, .$$

Маска, соответствующая такому фильтру показана на следующем рисунке слева. Более сложный фильтр может иметь вид, показанный на рисунке справа.

| | | |
|---|---|---|
| $\frac{1}{9}$ | $\frac{1}{9}$ | $\frac{1}{9}$ |
| $\frac{1}{9}$ | $\frac{1}{9}$ | $\frac{1}{9}$ |
| $\frac{1}{9}$ | $\frac{1}{9}$ | $\frac{1}{9}$ |

| | | |
|---|---|---|
| $\frac{1}{16}$ | $\frac{2}{16}$ | $\frac{1}{16}$ |
| $\frac{2}{16}$ | $\frac{4}{16}$ | $\frac{2}{16}$ |
| $\frac{1}{16}$ | $\frac{2}{16}$ | $\frac{1}{16}$ |

Второй фильтр учитывает яркости пикселей, покрываемых маской непропорционально: наибольшее значение на результирующую яркость пикселя оказывает яркость этого же пикселя на исходном изображении. В то же

время пиксели из окрестности влияют на формирование яркости обратно пропорционально их удаленности от модифицируемого пикселя.

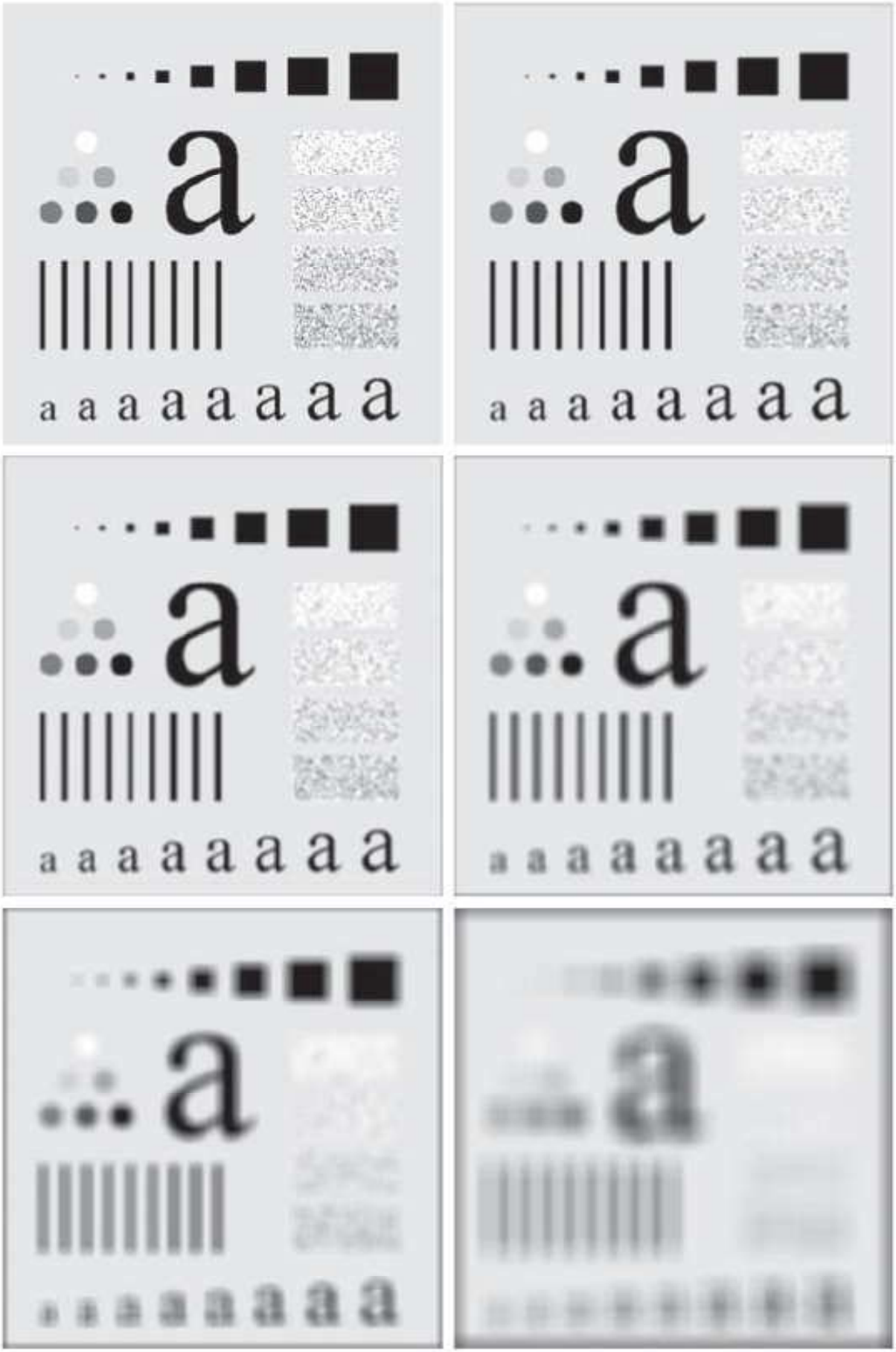
При этом важно отметить, что, как правило, требуется, чтобы коэффициенты фильтра были положительны, и их сумма была равна единице. Это гарантирует, что результирующая яркость будет заключена между минимальной и максимальной яркостями пикселей из окрестности, покрываемой фильтром. Последнее требование можно, очевидно снять, если записать формулу фильтрации в наиболее общем виде следующим образом:

$$g(x,y)=\frac{\sum_{s=-a}^{a}\sum_{t=-b}^{b} w(s,t)f(x+s,y+t)}{\sum_{s=-a}^{a}\sum_{t=-b}^{b} w(s,t)}.$$

Фильтры второго типа называют также взвешенными усредняющими фильтрами.

Важно подчеркнуть, что степень сглаживания изображения существенно зависит от размеров фильтра. На следующем рисунке показаны примеры применения сглаживающего фильтра к одному и тому же изображению, но при разных размерах фильтра. Исходное изображение приведено слева вверху. Далее последовательно слева направо и сверху вниз показаны изображения, полученные применением к исходному изображению фильтров с квадратной окрестностью размеров $m=3,5,9,15,35$ элементов изображения. По многочисленным деталям изображения можно увидеть, что с ростом размера фильтра степень расфокусировки изображения быстро растет.

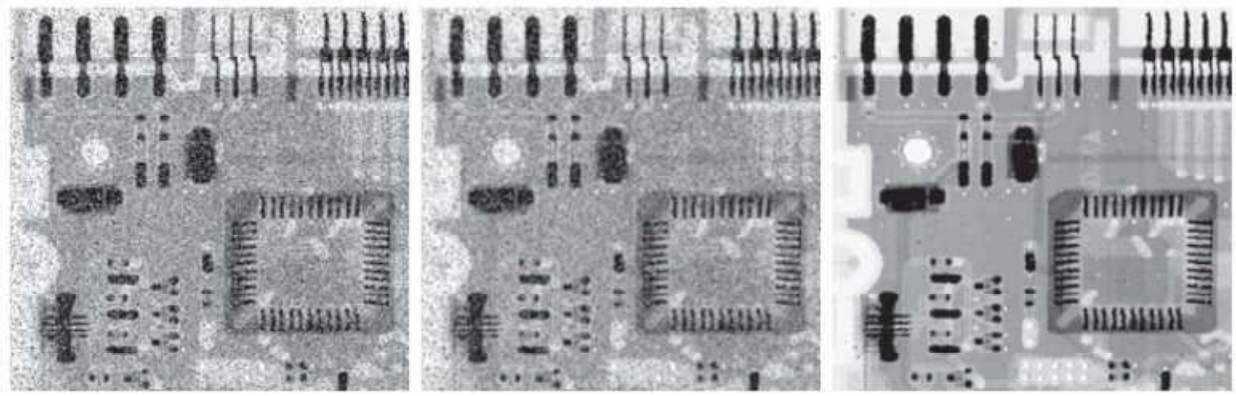
Следует отметить, что с ростом расфокусировки происходит размывание не только шума, но и мелких деталей изображения. Поэтому эту процедуру в совокупности с последующим пороговым градационным преобразованием часто применяют для выделения наиболее существенных в смысле яркости частей изображения.



## 6.2. Пример нелинейного фильтра. Медианный фильтр

Так называемый медианный фильтр основан на следующем алгоритме фильтрации. Все яркости пикселей из окрестности упорядочиваются по возрастанию. Затем в этом наборе определяется такая яркость $\alpha$ , что ровно половина пикселей имеет яркость, которая меньше, либо равна $\alpha$ , а половина – больше, либо равна $\alpha$ . После этого яркость $\alpha$ присваивается обрабатываемому пикселю. Число $\alpha$ называется *медианой* окрестности.

Очевидно, что данный фильтр является нелинейным. Несмотря на свою простоту, он демонстрирует очень хорошие сглаживающие свойства. На следующем снимке можно сравнить результат обработки изображения, показанного слева, сначала усредняющим фильтром размера $3\times3$ (в центре) и медианным фильтром размера $3\times3$ (справа).



Медианные фильтры оказываются очень эффективными при удалении помех, представляющих собой резко выделяющиеся по яркости пиксели среди пикселей окрестности. Это так называемый шум «соль и перец». При этом, что особенно характеризует медианную фильтрацию, не наблюдается эффекта сильной расфокусировки изображения.

# Глава 7. Дифференциальные фильтры

## 7.1. Основные принципы построения дифференциальных фильтров

Дифференциальные фильтры в отличие от усредняющих фильтров предназначены для повышения резкости изображения. Усредняющие фильтры были основаны на суммировании и усреднении яркостей пикселей, входящих в окрестность обрабатываемого пикселя и приводят к расфокусировке

изображения. Операция суммирования является дискретным аналогом интегрирования функции. Поэтому можно предполагать, что задача повышения резкости должна решаться с помощью фильтров, основанных на дискретном аналоге дифференцирования функций. По этой причине такие фильтры называются *дифференциальными*.

Конечно-разностная аппроксимация первой производной имеет вид:

$$\frac{df}{dx} = \frac{f(x+\Delta x) - f(x)}{\Delta x}.$$

Учитывая, что расстояние между соседними пикселями по горизонтали и по вертикали принимается равным единице, то $\Delta x = 1$, поэтому

$$\frac{df}{dx} = f(x+1) - f(x).$$

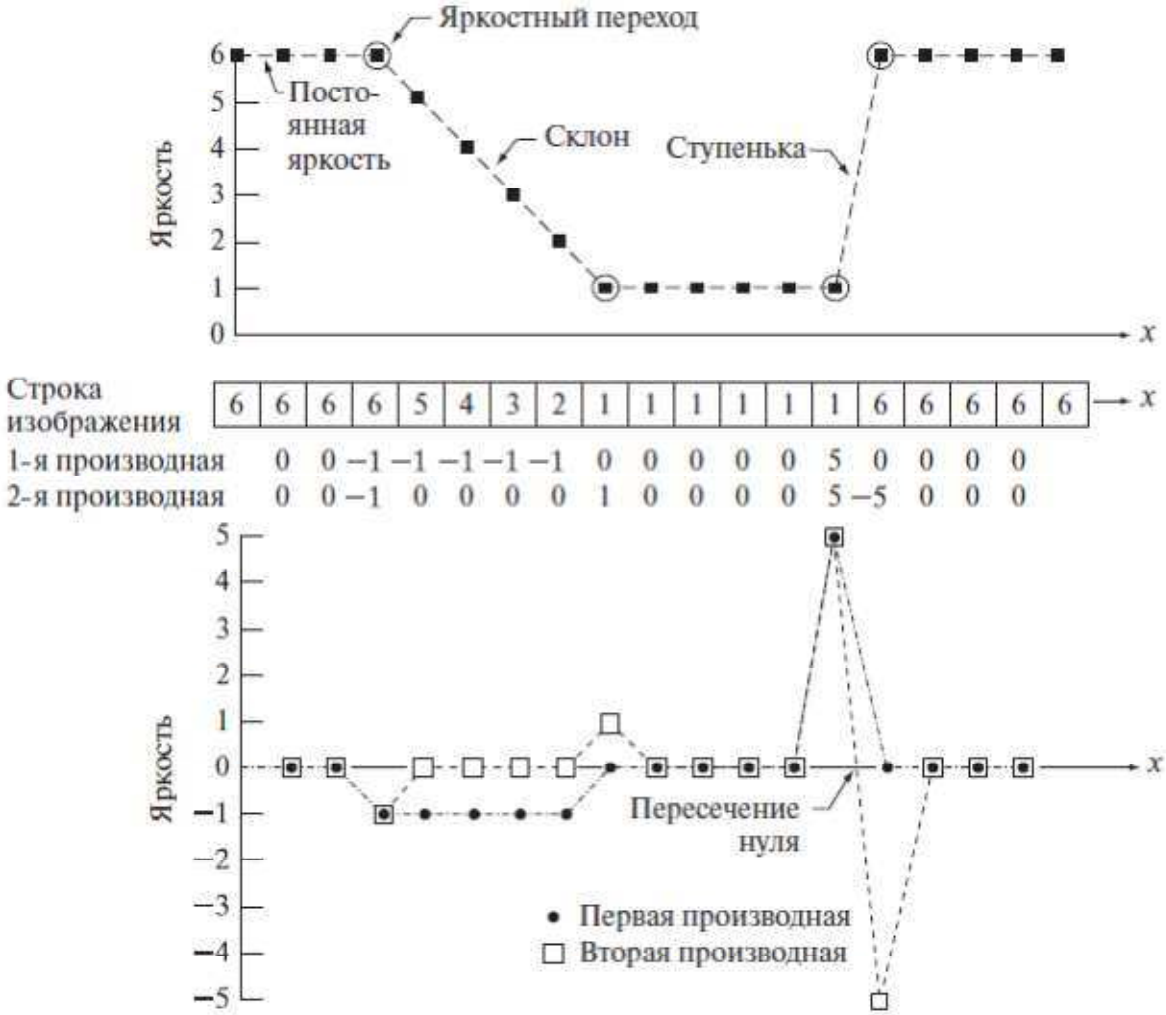
Аналогично, аппроксимация второй производной равна

$$\frac{d^2 f}{dx^2} = \big(f(x+1) - f(x)\big) - \big(f(x) - f(x-1)\big) = f(x+1) + f(x-1) - 2f(x).$$

Исходя из геометрического смысла первой производной, можно утверждать, что на участках относительно незначительного изменения яркостей первая производная будет принимать значения близкие к нулю. В то же время в точках, где происходит существенное изменение яркости, она будет принимать большие значения, пропорциональные величине этого изменения. При этом знак производной будет зависеть от того, происходит ли увеличение или уменьшение значения яркости. На участках, где наблюдается более или менее монотонное изменение яркости, первая производная будет примерно постоянна, и ее значение будет определяться тангенсом угла наклона функции яркости в соответствующем направлении (по горизонтали или по вертикали, если вычисляется производная по $y$).

Аналогично, вторая производная будет примерно равна нулю на участках примерно равномерной яркости и участках с примерно линейным изменением яркости. Но она будет отлична от нуля в точках, в которых происходит резкое изменение яркости изображения.

На следующем рисунке приведена иллюстрация, показывающая поведение первой и второй производных на модельном примере одномерного изображения.




## 7.2. Повышение резкости изображений на основе лапласиана

Важную роль при повышении резкости изображений играет так называемый *оператор Лапласа*, который называют просто *лапласианом*. Он определяется для непрерывной дважды дифференцируемой функции следующим образом:

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}.$$

Данный оператор, очевидно, является линейным. Учитывая приведенные выше аппроксимации вторых производных, лапласиан может быть представлен в виде конечно-разностной аппроксимации следующим образом:

$$\nabla^2 f(x,y) = f(x+1,y) + f(x-1,y) + f(x,y+1) + f(x,y-1) - 4f(x,y).$$

В задачах фильтрации изображения к лапласиану формально добавляются также вторые производные по направлениям, определяемым векторами $(1,1)$ и $(-1,1)$. В результате аппроксимация лапласиана «расширяется» и принимает вид:

$$\nabla^2 f(x,y) = f(x+1,y) + f(x-1,y) + f(x,y+1) + f(x,y-1) +$$

$$+ f(x+1,y+1) + f(x-1,y-1) + f(x+1,y-1) + f(x-1,y+1) - 8f(x,y).$$

На следующем рисунке показаны маски фильтров, соответствующих исходной и расширенной аппроксимациям оператора Лапласа.

| | | |
|---|---|---|
| 0 | 1 | 0 |
| 1 | –4 | 1 |
| 0 | 1 | 0 |

| | | |
|---|---|---|
| 1 | 1 | 1 |
| 1 | −8 | 1 |
| 1 | 1 | 1 |

Следует обратить внимание на то, что оператор Лапласа обладает свойством *изотропности*, т.е. он *инвариантен* к поворотам на углы, кратные $\frac{\pi}{4}$. В частности, это означает, что корреляция и свертка с оператором Лапласа будут совпадать.

Результат применения оператора Лапласа проявляется в том, что на участках изображения с мало изменяющейся яркостью или примерно линейно меняющейся яркостью новые пиксели будут иметь яркость близкую к нулю, т.е. изображение будет в таких местах почти черным. Напротив, в местах, где яркость быстро изменяется, на изображении лапласиана будут светлые полосы и контуры. Типичная схема применения оператора Лапласа для увеличения резкости изображения состоит в сложении исходного изображения и изображения, полученного после обработки лапласианом. Это можно выразить следующей формулой:

$$g(x,y) = f(x,y) + \nabla^2 f(x,y).$$

При этом, как правило, после преобразования изображения оператором Лапласа результат этой обработки подвергают градационному преобразованию,

так, чтобы растянуть диапазон яркостей на всю шкалу $[0, L-1]$. Данное преобразование было описано раньше.

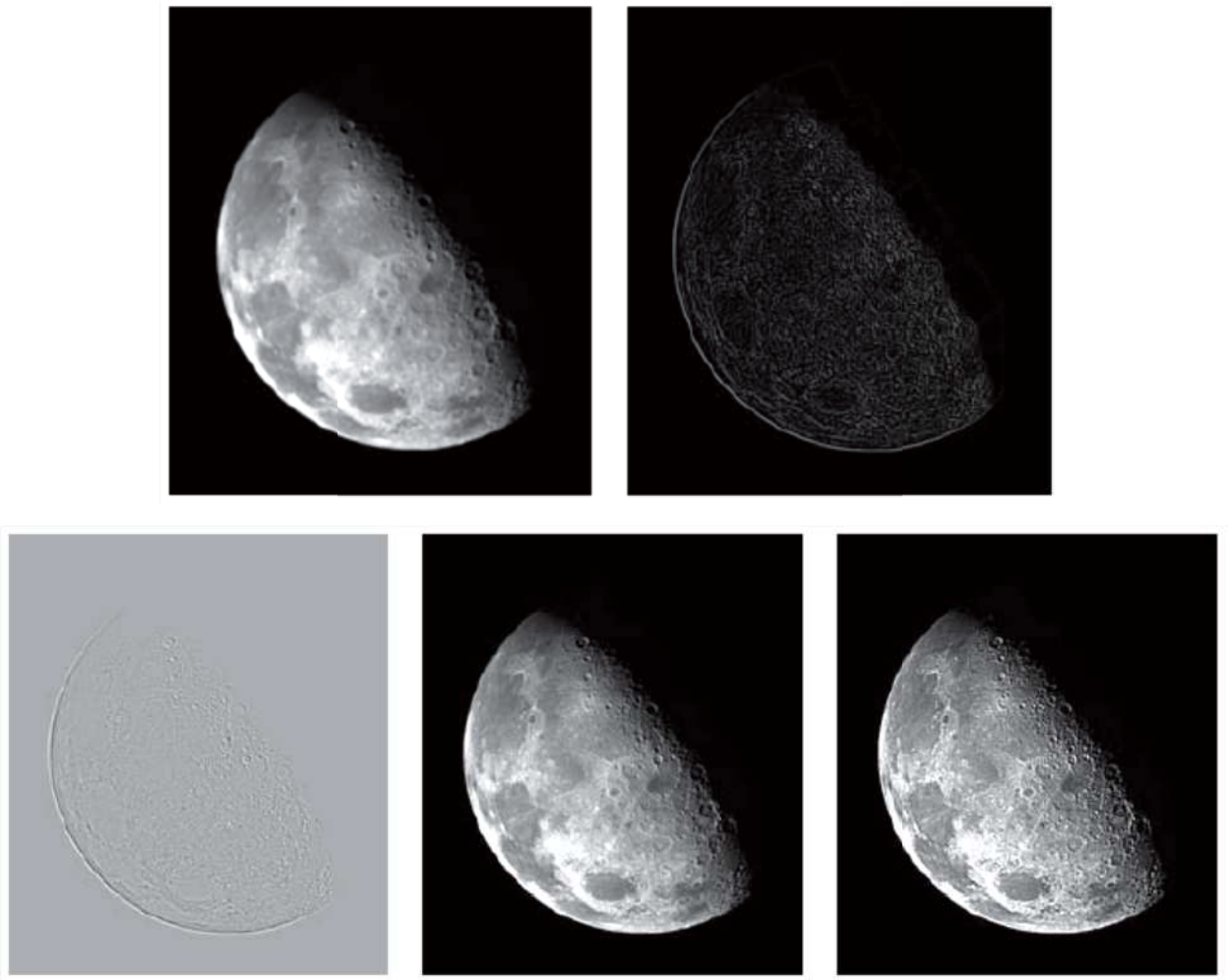
На следующем рисунке показан пример применения лапласиана для повышения резкости изображения. На первом рисунке дано исходное изображение части поверхности Луны. Далее последовательно показаны:

– результат его обработки лапласианом;

– градационное преобразование лапласиана;

– изображение с увеличенной резкостью при применении первого лапласианов;

– изображение с повышенной резкостью при применении второго лапласиана.



## 7.3. Повышение резкости: операторы Робертса и Собела

Популярным подходом для повышения резкости изображения является использование дифференциальных фильтров на основе первых производных функции двух переменных. Как и в предыдущем случае на цифровых изо-

бражениях можно говорить лишь об аппроксимациях этих производных. В основе такого подхода лежат не сами производные, а основанное на них понятие градиента функции двух переменных:

$$\nabla f(x,y) = \operatorname{grad} f(x,y) = \begin{pmatrix} \dfrac{\partial f}{\partial x} & \dfrac{\partial f}{\partial y} \end{pmatrix}^T .$$

Напомним, что градиент, как вектор, показывает направление наибольшего изменения значений функции в точке, в которой он вычисляется. Соответственно величина скорости изменения функции в данной точке измеряется нормой (длиной) этого вектора:

$$M(x,y) = \left\| \nabla f(x,y) \right\|_2 = \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2} .$$

где $\|*\|_2$ обозначает эвклидову норму. В некоторых случаях, когда требуется снижать вычислительные затраты, применяют более простую в вычислительном отношении норму (метрика $L_1$):

$$M(x,y) = \left\| \nabla f(x,y) \right\|_1 = \left|\frac{\partial f}{\partial x}\right| + \left|\frac{\partial f}{\partial y}\right| .$$

Если построить изображение, в котором каждому пикселю присвоить значение яркости равное $M(x,y)$, то такое изображение называется *градиентным*. Таким образом, задача сводится к аппроксимации производных и подстановке их в выражение для нормы.

Простейшая аппроксимация производных, вычисляемых вдоль диагональных направлений, выглядит так:

$$\frac{\partial f}{\partial x} = f(x+1, y+1) - f(x,y), \qquad \frac{\partial f}{\partial y} = f(x+1, y) - f(x, y+1) .$$

Соответствующие маски выглядят так, как показано на рисунке ниже.

| | | |
|---|---|---|
| 0 | 0 | 0 |
| 0 | –1 | 0 |
| 0 | 0 | 1 |

| | | |
|---|---|---|
| 0 | 0 | 0 |
| 0 | 0 | -1 |
| 0 | 1 | 0 |

Данный фильтр называется *перекрестным оператором Робертса* или просто *оператором Робертса*. Использование этого оператора несколько неудобно, поскольку он несимметричный.

Поэтому чаще используется так называемый *оператор Собела*:

$$\frac{\partial f}{\partial x} = \left[ f(x+1, y-1) + 2f(x+1, y) + f(x+1, y+1) \right] -$$

$$-\left[ f(x-1, y-1) + 2f(x-1, y) + f(x-1, y+1) \right],$$

$$\frac{\partial f}{\partial y} = \left[ f(x-1, y+1) + 2f(x, y+1) + f(x+1, y+1) \right] -$$

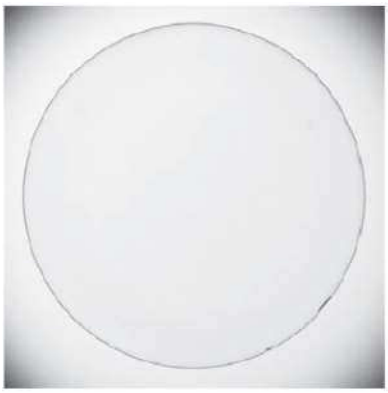
$$-\left[ f(x-1, y-1) + 2f(x, y-1) + f(x+1, y-1) \right],$$

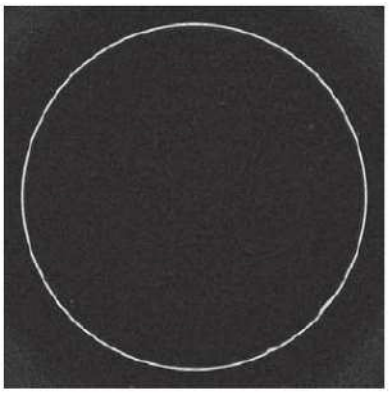
маски которого представлены ниже:

| | | |
|---|---|---|
| -1 | -2 | -1 |
| 0 | 0 | 0 |
| 1 | 2 | 1 |

| | | |
|---|---|---|
| -1 | 0 | 1 |
| -2 | 0 | 2 |
| 1 | 0 | 1 |

Нетрудно понять, что оператор Собела представляет собой некоторое усреднение аппроксимаций первых частных производных по окрестности $3 \times 3$. Кроме того, по маскам видно, что оператор Собела симметричный. Следует также отметить, что этот оператор не является линейным из-за вычисления нормы градиента. Ниже показан пример исходного и градиентного изображения контактной линзы.





# Раздел 3.

# ЧАСТОТНЫЕ МЕТОДЫ УЛУЧШЕНИЯ ИЗОБРАЖЕНИЙ

## Глава 8. Введение в Фурье-анализ функций одной переменной

### 8.1. Предварительные сведения

Напомним вначале некоторые определения из Фурье-анализа для функции одной переменной. Пусть $f(t)$ – периодическая с периодом $T$ функция. Тогда она, как известно, может быть представлена в виде *ряда Фурье* по синусам и косинусам, либо в *экспоненциальной форме* по ортогональному базису из экспонент:

$$f(t) = \sum_{n=-\infty}^{\infty} c_n e^{i\frac{2\pi n}{T}t},$$

где коэффициенты $c_n$ определяются выражениями

$$c_n = \frac{1}{T}\int_{-T/2}^{T/2} f(t)\,\mathrm{e}^{-i\frac{2\pi n}{T}t}\,dt, \qquad n = 0, \pm 1, \pm 2, \ldots$$

*Преобразованием Фурье* непрерывной функции $f(t)$ называется выражение

$$F(\mu) = \mathfrak{F}\left[f(t)\right] = \int_{-\infty}^{\infty} f(t)\,\mathrm{e}^{-i2\pi\mu t}\,dt.$$

Поскольку интегрирование осуществляется по переменной $t$, то преобразование Фурье зависит только от $\mu$. Достаточным условием существования преобразования Фурье является условие

$$\int_{-\infty}^{\infty} \left|f(t)\right| dt < \infty.$$

Особым свойством, которое делает преобразование Фурье чрезвычайно важным инструментов, в том числе и для обработки изображений, является

его обратимость. А именно, зная $F(\mu)$, можно восстановить исходную функцию $f(t)$ по формуле

$$f(t) = \mathfrak{F}^{-1}\left[F(\mu)\right] = \int_{-\infty}^{\infty} F(\mu)\mathrm{e}^{i2\pi\mu t}\, d\mu .$$

Выражения для $f(t)$ и $F(\mu)$ образуют так называемую *пару преобразований Фурье*, а сами функции $f(t)$ и $F(\mu)$ называются *Фурье-парой*.

Если функция $f(t)$ является действительной, то Фурье-образ $F(\mu)$ этой функции является, вообще говоря, комплексной функцией. Заметим также, что $F(\mu)$ зависит только от $\mu$. Эта переменная имеет смысл *частоты колебаний* (достаточно выразить экспоненту через синусы и косинусы по формуле Эйлера). Поэтому область определения Фурье-образа $F(\mu)$ называют *частотной областью* в отличие от области определения $f(t)$, которую называют *пространственной областью*.

Рассмотрим важный пример.

**Пример.** Построим Фурье-образ прямоугольной функции:

$$f(t) = \begin{cases} A, & -\frac{W}{2} \le t \le \frac{W}{2}, \\ 0, & \text{в противном случае}. \end{cases}$$

$$F(\mu) = \int_{-\infty}^{\infty} f(t)\mathrm{e}^{-i2\pi\mu t}\, dt = \int_{-\frac{W}{2}}^{\frac{W}{2}} A\mathrm{e}^{-i2\pi\mu t}\, dt = -\frac{A}{i2\pi\mu}\left[e^{-i2\pi\mu t}\right]_{-\frac{W}{2}}^{\frac{W}{2}} =$$

$$= -\frac{A}{i2\pi\mu}\left[e^{-i\pi\mu W} - e^{i\pi\mu W}\right] = \frac{A}{i2\pi\mu}\left[e^{i\pi\mu W} - e^{-i\pi\mu W}\right] = AW\frac{\sin \pi\mu W}{\pi\mu W} .$$

Заметим, что Фурье-образ прямоугольной функции оказался действительнозначной функцией. Дробь в последнем выражении известна как *кардинальный синус*:

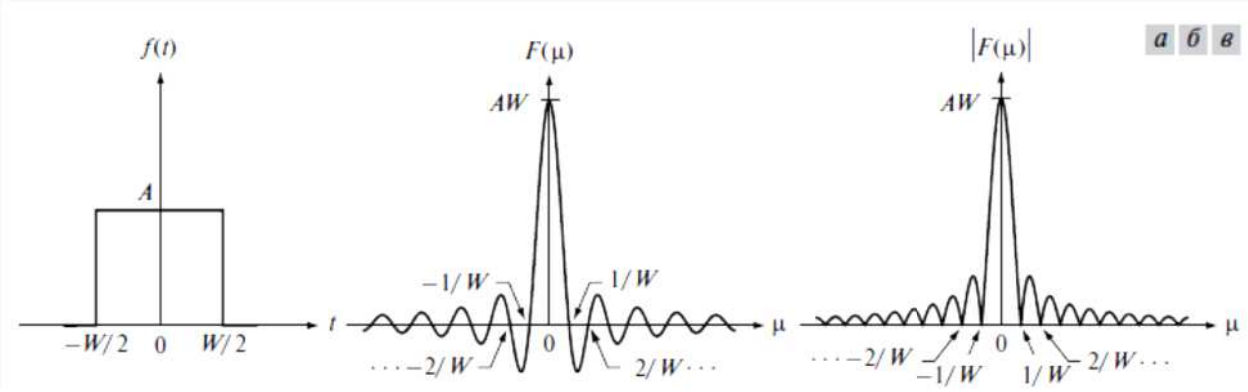
$$\mathrm{sinc}(m) = \frac{\sin \pi m}{\pi m} .$$

При этом $\mathrm{sinc}(0) = 1$ и $\mathrm{sinc}(m) = 0$ для всех целых не равных нулю $m$ .

Как уже отмечалось, образ является, вообще говоря, комплексной функцией. Для визуализации обычно используют амплитуду преобразования $F(\mu)$, которая является действительной величиной:

$$\left|F(\mu)\right| = AW\left|\frac{\sin \pi\mu W}{\pi\mu W}\right| = AW\left|\mathrm{sinc}(\mu W)\right|.$$

Амплитуду называют также *Фурье-спектром* или *частотным спектром*. На рисунке ниже показаны график функции $f(t)$, график Фурье-образа и частотного спектра.




Следует обратить внимание на несколько характерных свойств преобразования Фурье:

– расстояние между нулями функции $F(\mu)$ обратно пропорционально ширине $W$ прямоугольной функции;

– высота лепестков уменьшается по мере удаления от начала координат;

– сама функция $F(\mu)$ продолжается в бесконечность для положительных и отрицательных значений $\mu$.

## 8.2. Свертка

| | |
|---|---|
| **Определение** | *Сверткой* двух функций $f(t)$ и $h(t)$ называется выражение $$f(t)*h(t) = \int_{-\infty}^{\infty} f(\tau)h(t-\tau)d\tau\,.$$ |

Вычислим Фурье-преобразование от свертки двух функций:

$$\mathfrak{F}\left[f(t)*h(t)\right] = \int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} f(\tau)h(t-\tau)d\tau\right]e^{-i2\pi\mu t}dt = \int_{-\infty}^{\infty} f(\tau)\left[\int_{-\infty}^{\infty} h(t-\tau)e^{-i2\pi\mu t}dt\right]d\tau\,.$$

Рассмотрим квадратную скобку:

$$\int_{-\infty}^{\infty} h(t-\tau)e^{-i2\pi\mu t}dt = \int_{-\infty}^{\infty} h(s)e^{-i2\pi\mu(s+\tau)}ds = e^{-i2\pi\mu\tau}\int_{-\infty}^{\infty} h(s)e^{-i2\pi\mu s}ds = H(\mu)e^{-i2\pi\mu\tau},$$

где $H(\mu)$ – Фурье-образ от $h(t)$. Но тогда

$$\mathfrak{F}\left[f(t)*h(t)\right] = \int_{-\infty}^{\infty} f(\tau)\left[H(\mu)e^{-i2\pi\mu\tau}\right]d\tau = H(\mu)\int_{-\infty}^{\infty} f(\tau)e^{-i2\pi\mu\tau}d\tau = H(\mu)F(\mu).$$

Таким образом, Фурье-образ свертки двух функций в пространственной области равен произведению Фурье-образов этих функций в частотной области. Наоборот, обратное преобразование Фурье произведения двух Фурье-образов в частотной области может быть представлено в виде свертки двух функций в пространственной области. Иными словами, $f(t)*h(t)$ и $H(\mu)F(\mu)$ образуют Фурье-пару. Это соответствие принято записывать в виде:

$$f(t)*h(t) \Leftrightarrow H(\mu)F(\mu).$$

Аналогично можно получить и вторую зависимость:

$$f(t)h(t) \Leftrightarrow H(\mu)*F(\mu),$$

т.е. произведению функций в пространственной области соответствует свертка их Фурье-образов в частотной области.

## 8.3. Преобразование Фурье дискретных функций

При обработке на компьютере приходится иметь дело с дискретными и квантованными функциями. Поэтому необходимо понимать, как осуществляется преобразование Фурье дискретных функций.

Пусть имеется непрерывная функция $f(t)$, которая преобразуется в дискретную функцию с шагом $\Delta T$. Рассмотрим функцию

$$\delta(t) = \begin{cases} \infty, & t = 0, \\ 0, & \textit{в противном случае.} \end{cases}$$

Это так называемая *импульсная функция* или *дельта-функция Дирака*, которая, как известно, обладает свойством:

$$\int_{-\infty}^{\infty} \delta(t)dt = 1.$$

В частности, для любого $n$ функция $\delta(t-n\Delta T)=\infty$ при $t=n\Delta T$ и равна 0 во всех остальных точках. Можно показать также, что $\int\limits_{-\infty}^{\infty} f(t)\delta(t)dt = f(0)$ и, соответственно, $\int\limits_{-\infty}^{\infty} f(t)\delta(t-n\Delta T)dt = f(n\Delta T)$.

В дискретном случае единичная импульсная функция определяется как

$$\delta(t)=\begin{cases}1, & t=0,\\ 0, & \textit{в противном случае.}\end{cases}$$

Соответственно, имеют место дискретные аналоги приведенных выше выражений. Например, для переменной $x$, принимающей дискретные значения, получается: $\sum\limits_{x=-\infty}^{\infty} f(x)\delta(x-n\Delta T) = f(n\Delta T)$.

Тогда функция вида

$$\tilde{f}(t)=\sum_{n=-\infty}^{\infty} f(t)\delta(t-n\Delta T)$$

уже является *дискретной функцией*, принимающей ненулевые значения в точках, кратных $\Delta T$, и равные $f(n\Delta T)$. Посмотрим, как получается преобразование Фурье такой дискретной функции.

**Утверждение.** $\mathfrak{F}\left[\sum\limits_{n=-\infty}^{\infty} f(t)\delta(t-n\Delta T)\right] = F(\mu)*S(\mu)$, *где* $F(\mu)$ – *Фурье-образ функции* $f(t)$, *а* $S(\mu)=\dfrac{1}{\Delta T}\sum\limits_{n=-\infty}^{\infty}\delta\left(\mu-\dfrac{n}{\Delta T}\right)$.

*Доказательство*. Обозначим через $F(\mu)$ Фурье-образ непрерывной функции $f(t)$. Заметим теперь, что функцию $\tilde{f}(t)$ можно представить в виде произведения двух функций:

$$\tilde{f}(t)=f(t)\cdot\sum_{n=-\infty}^{\infty}\delta(t-n\Delta T).$$

В силу теоремы о свертке получаем:

$$\tilde{F}(\mu)=\mathfrak{F}\left[\tilde{f}(t)\right]=\mathfrak{F}\left[f(t)\cdot\sum_{n=-\infty}^{\infty}\delta(t-n\Delta T)\right]=F(\mu)*S(\mu),$$

где $S(\mu)$ Фурье-образ функции $\sum_{n=-\infty}^{\infty}\delta(t-n\Delta T)$. Вычислим $S(\mu)$. Для этого заметим, что последовательность импульсов в выражении $\sum_{n=-\infty}^{\infty}\delta(t-n\Delta T)$ можно рассматривать как периодическую функцию с периодом $\Delta T$. Значит, ее можно представить рядом Фурье

$$\sum_{n=-\infty}^{\infty}\delta(t-n\Delta T)=\sum_{n=-\infty}^{\infty}c_n e^{i\frac{2\pi n}{\Delta T}t},$$

где коэффициенты $c_n$ определяются выражением:

$$c_n=\frac{1}{\Delta T}\int_{-\frac{\Delta T}{2}}^{\frac{\Delta T}{2}}\sum_{n=-\infty}^{\infty}\delta(t-n\Delta T)e^{i\frac{2\pi n}{\Delta T}t}dt=\frac{1}{\Delta T}\int_{-\frac{\Delta T}{2}}^{\frac{\Delta T}{2}}\delta(t)e^{i\frac{2\pi n}{\Delta T}t}dt=\frac{1}{\Delta T}e^{0}=\frac{1}{\Delta T}.$$

Таким образом, $\sum_{n=-\infty}^{\infty}\delta(t-n\Delta T)=\frac{1}{\Delta T}\sum_{n=-\infty}^{\infty}e^{i\frac{2\pi n}{\Delta T}t}$. Следовательно, принимая во внимание линейность преобразования Фурье, получаем

$$S(\mu)=\mathfrak{F}\left(\sum_{n=-\infty}^{\infty}\delta(t-n\Delta T)\right)=\frac{1}{\Delta T}\sum_{n=-\infty}^{\infty}\mathfrak{F}\left[e^{i\frac{2\pi n}{\Delta T}t}\right].$$

Заметим, если $F(\mu)$ – Фурье-образ функции $f(t)$, то Фурье-образом функции $F(t)$ будет функция $f(-\mu)$. Это вытекает из того, что прямое и обратное преобразования Фурье отличаются лишь знаком у показателя экспоненты. Далее легко проверяется, что следующие функции образуют Фурье-пару:

$$\delta\left(t-\frac{n}{\Delta T}\right)\Leftrightarrow F(\mu)=\mathrm{e}^{-i\frac{2\pi n}{\Delta T}\mu}.$$

Следовательно, в силу сделанного выше замечания получаем:

$$\mathfrak{F}\left[\mathrm{e}^{i\frac{2\pi n}{\Delta T}t}\right]=\delta\left(\mu-\frac{n}{\Delta T}\right).$$

С учетом этого, получаем окончательно:

$$S(\mu)=\mathfrak{F}\left(\sum_{n=-\infty}^{\infty}\delta(t-n\Delta T)\right)=\frac{1}{\Delta T}\sum_{n=-\infty}^{\infty}\delta\left(\mu-\frac{n}{\Delta T}\right),$$

что завершает доказательство утверждения. ■

Доказанное утверждение показывает, что Фурье-преобразование последовательности импульсов с периодом $\Delta T$ также является последовательностью импульсов, но уже с периодом $\frac{1}{\Delta T}$. Эта обратная пропорциональность является важной особенностью Фурье преобразования в частотную область.

Вычислим теперь свертку Фурье-образов $F(\mu)$ и $S(\mu)$ и получим явное выражение для преобразования дискретной функции:

$$\tilde{F}(\mu) = F(\mu) * S(\mu) = \int_{-\infty}^{\infty} F(\tau) S(\mu - \tau) d\tau = \frac{1}{\Delta T} \int_{-\infty}^{\infty} F(\tau) \sum_{n=-\infty}^{\infty} \delta\left(\mu - \tau - \frac{n}{\Delta T}\right) d\tau =$$

$$= \frac{1}{\Delta T} \sum_{n=-\infty}^{\infty} \int_{-\infty}^{\infty} F(\tau) \delta\left(\mu - \tau - \frac{n}{\Delta T}\right) d\tau = \frac{1}{\Delta T} \sum_{n=-\infty}^{\infty} F\left(\mu - \frac{n}{\Delta T}\right).$$

Из последней формулы видно, что Фурье-образ дискретной функции $\tilde{f}(t)$ представляет собой бесконечную сумму копий функции $F(\mu)$, которая является Фурье образом исходной непрерывной функции $f(t)$. Причем эти функции последовательно сдвинуты на величину $\frac{1}{\Delta T}$.

*Фурье-образ дискретной функции является непрерывной периодической функцией!*

## 8.4. Дискретное преобразование Фурье функций одной переменной

Как выше было отмечено, преобразование Фурье дискретной функции, определенной от $-\infty$ до $\infty$, является непрерывной функцией, также заданной на всей оси от $-\infty$ до $\infty$. В реальности функции всегда заданы конечным числом своих значений, поэтому необходимо использовать конечномерные аналоги преобразования Фурье для отображения в частотную область функции, заданной в пространственной области. Это делается с помощью так называемого *дискретного преобразования Фурье*.

Заметим, что полученное выше преобразование Фурье $\tilde{F}(\mu)$ дискретной функции $\tilde{f}(t)$ выражается через Фурье-образ $F(\mu)$ соответствующей непрерывной функции $f(t)$. Получим выражение для $\tilde{F}(\mu)$ непосредственно

через значения дискретной функции, пользуясь определением преобразования Фурье:

$$\tilde{F}(\mu)=\int_{-\infty}^{\infty}\tilde{f}(t)\,\mathrm{e}^{-i2\pi\mu t}\,dt=\int_{-\infty}^{\infty}\sum_{n=-\infty}^{\infty}f(t)\delta(t-n\Delta T)\,\mathrm{e}^{-i2\pi\mu t}\,dt=$$

$$=\sum_{n=-\infty}^{\infty}\int_{-\infty}^{\infty}f(t)\delta(t-n\Delta T)\,\mathrm{e}^{-i2\pi\mu t}\,dt=\sum_{n=-\infty}^{\infty}f_n\,\mathrm{e}^{-i2\pi\mu n\Delta T}\ .$$

Здесь $f_n$ – это значения функции $\tilde{f}(t)$ в точках $n\Delta T$, $n=0,\pm1,\pm2,\ldots$. Таким образом, преобразование Фурье дискретной функции является непрерывной функцией. Причем эта функция периодическая с периодом $\dfrac{1}{\Delta T}$. Это значит, что для описания всей функции достаточно рассмотреть лишь один период.

Пусть на отрезке $\left[0,\dfrac{1}{\Delta T}\right]$ равном одному периоду преобразования Фурье требуется получить в точности $M$ значений, вычисленных в точках $\mu=\dfrac{m}{M\Delta T}$, $m=0,1,\ldots,M-1$. Подставим эти значения в последнее выражение, в результате получим совокупность из $M$ значений Фурье-образа:

$$F_m=\sum_{n=0}^{M-1}f_n\,\mathrm{e}^{-\frac{i2\pi mn}{M}},\qquad m=0,1,\ldots,M-1.$$

Полученное выражение и есть *дискретное преобразование Фурье*.

Аналогично может быть получено и *обратное дискретное преобразование Фурье*, которое позволяет восстановить значение исходной функции по набору значений Фурье-образа:

$$f_n=\frac{1}{M}\sum_{m=0}^{M-1}F_m\,\mathrm{e}^{i\frac{2\pi mn}{M}},\qquad n=0,1,\ldots,M-1.$$

Два последних выражения составляют так называемую *пару дискретных преобразований Фурье* или *ДПФ-пару*. Следует отметить, что полученные выражения применимы к *любому* конечному набору значений.

Обычно при записи ДПФ-пары принято использовать несколько другие обозначения. А именно,

$$F(u)=\sum_{x=0}^{M-1}f(x)\,\mathrm{e}^{-\frac{i2\pi ux}{M}},\qquad u=0,1,\ldots,M-1.$$

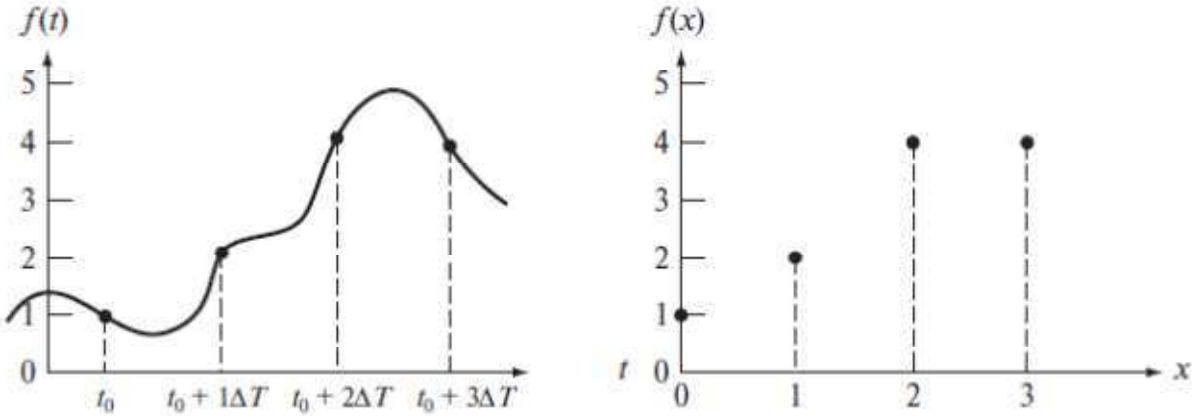
$$f(x)=\frac{1}{M}\sum_{u=0}^{M-1}F(u)\mathrm{e}^{i\frac{2\pi ux}{M}}, \quad x=0,1,...,M-1.$$

При этом понимается, что $F(u)=F_m$ и $f(x)=f_n$. Нетрудно проверить, что и прямое, и обратное дискретное преобразования Фурье являются периодическими функциями с периодом $M$, т.е. для любого целого $k$

$$F(u)=F(u+kM),$$

$$f(x)=f(x+kM).$$

**Пример.** Пусть имеется четыре равноотстоящих значения функции $f(t)$, взятых с интервалом $\Delta T$, как показано на рисунке. Справа эти отсчеты представлены в дискретном пространстве.




Тогда

$$F(0)=\sum_{x=0}^{3}f(x)=f(0)+f(1)+f(2)+f(3)=1+2+4+4=11,$$

$$F(1)=\sum_{x=0}^{3}f(x)\mathrm{e}^{-i\frac{2\pi\cdot 1\cdot x}{4}}=1\cdot\mathrm{e}^{-i\frac{2\pi\cdot 1\cdot 0}{4}}+2\cdot\mathrm{e}^{-i\frac{2\pi\cdot 1\cdot 1}{4}}+4\cdot\mathrm{e}^{-i\frac{2\pi\cdot 1\cdot 2}{4}}+4\cdot\mathrm{e}^{-i\frac{2\pi\cdot 1\cdot 3}{4}}=-3+2i,$$

$$F(2)=-1,$$

$$F(3)=-3-2i.$$

Нетрудно проверить аналогично, что значения исходной функции восстанавливаются по найденным значениям дискретного Фурье-образа.

# Глава 9. Введение в Фурье-анализ функций двух переменных

## 9.1. Двумерные непрерывные преобразования Фурье

Аналогом функции Дирака в двумерном случае является функция

$$\delta(t,z)=\begin{cases}\infty, & t=z=0,\\ 0, & \textit{в противном случае.}\end{cases}$$

Данная функция обладает свойством

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\delta(t,z)dtdz=1,$$

которое аналогично соответствующему свойству одномерной функции Дирака. В двумерном случае также выполняются свойства:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}f(t,z)\delta(t,z)dtdz=f(0,0),\qquad \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}f(t,z)\delta(t-t_0,z-z_0)dtdz=f(t_0,z_0).$$

Аналогичным образом определяется функция Дирака для дискретных переменных $x$ и $y$:

$$\delta(x,y)=\begin{cases}1, & x=y=0,\\ 0, & \textit{в противном случае.}\end{cases}$$

В этом случае также справедливы дискретные аналоги предыдущих соотношений:

$$\sum_{x=-\infty}^{\infty}\sum_{y=-\infty}^{\infty}f(x,y)\delta(x,y)=f(0,0),\qquad \sum_{x=-\infty}^{\infty}\sum_{y=-\infty}^{\infty}f(x,y)\delta(x-x_0,y-y_0)=f(x_0,y_0).$$

Если $f(t,z)$ – непрерывная функция двух переменных, то пара взаимно обратных двумерных непрерывных преобразований Фурье определяется следующим образом:

$$F(\mu,\nu)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}f(t,z)\mathrm{e}^{-i2\pi(\mu t+\nu z)}dtdz,$$

$$f(t,z)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}F(\mu,\nu)\mathrm{e}^{i2\pi(\mu t+\nu z)}d\mu d\nu.$$

Как и в одномерном случае, переменные $\mu$ и $\nu$ относятся к *частотной области* преобразования Фурье, а переменные $t$ и $z$ – к *пространственной*.

**Пример.** В качестве примера рассмотрим двумерное преобразование Фурье прямоугольной функции, представленной на рисунке ниже:

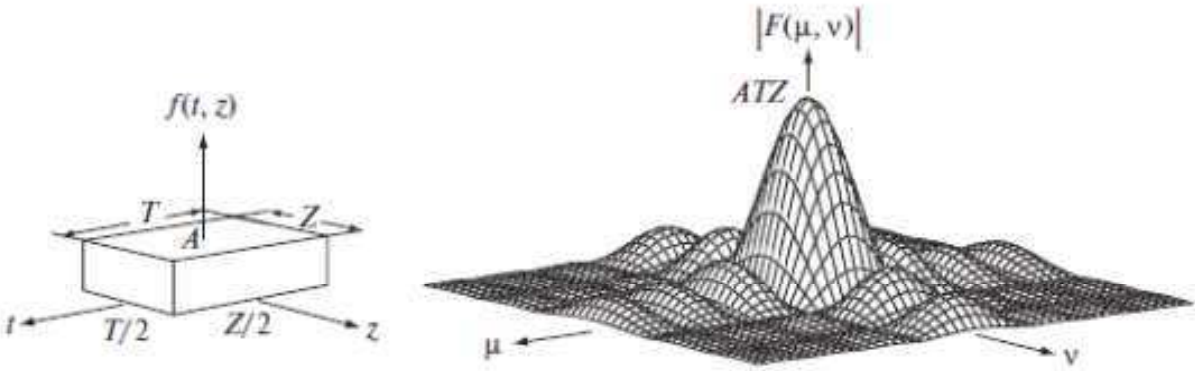
$$F(\mu,\nu)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(t,z)\,\mathrm{e}^{-i2\pi(\mu t+\nu z)}\,dtdz=\int_{-\frac{T}{2}}^{\frac{T}{2}}\int_{-\frac{Z}{2}}^{\frac{Z}{2}} A\,\mathrm{e}^{-i2\pi(\mu t+\nu z)}\,dtdz=$$

$$=ATZ\left(\frac{\sin\pi\mu T}{\pi\mu T}\right)\left(\frac{\sin\pi\nu Z}{\pi\nu Z}\right),$$

где в скобках записаны выражения кардинальных синусов. Соответственно *спектр* (амплитуда) определяется выражением

$$\left|F(\mu,\nu)\right|=ATZ\left|\frac{\sin\pi\mu T}{\pi\mu T}\right|\cdot\left|\frac{\sin\pi\nu Z}{\pi\nu Z}\right|.$$

На рисунке справа показан график спектра этого преобразования Фурье.

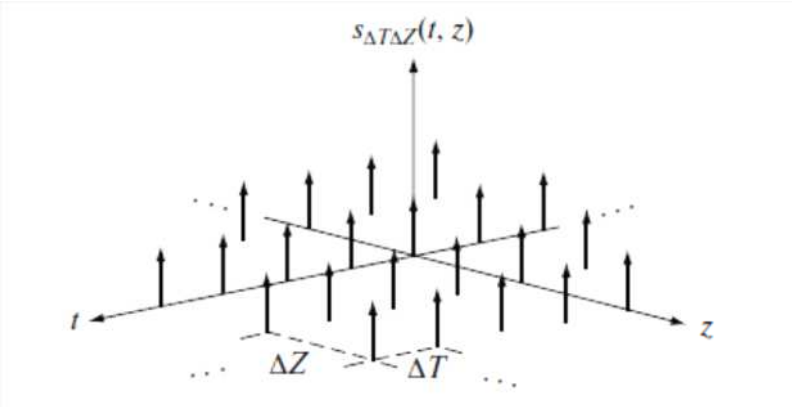



## 9.2. Двумерная дискретизация и двумерное дискретное преобразование Фурье

Как и в одномерном случае для случая двух переменных рассматривается двумерная последовательность импульсов, которая может быть представлена следующим образом:

$$\sum_{m=-\infty}^{\infty}\sum_{n=-\infty}^{\infty}\delta(t-m\Delta T,z-n\Delta Z)\,.$$

Здесь, как раньше, $\Delta T$ и $\Delta Z$ – шаги между отсчетами функции соответственно по осям $t$ и $z$. Графически такая система импульсов выглядит так, как показано на следующем рисунке.




| | |
|---|---|
| **Определение** | Функция $f(t,z)$ называется *функцией с ограниченной полосой частот*, если ее Фурье-образ равен нулю за пределами некоторого прямоугольника, заданного отрезками $[\mu_{\min}, \mu_{\max}]$ и $[\nu_{\min}, \nu_{\max}]$. |

**Теорема Котельникова.** *Непрерывная двумерная функция $f(t,z)$ с ограниченной полосой частот может быть точно восстановлена из набора своих отсчетов, если шаги дискретизации по каждой переменной удовлетворяют неравенствам:*

$$\Delta T < \frac{1}{2\mu_{\max}}, \qquad \Delta Z < \frac{1}{2\nu_{\max}}.$$

Данная теорема может быть интерпретирована следующим образом. Если функция с ограниченной полосой частот задается набором своих отсчетов, расстояние между которыми удовлетворяет приведенным выше неравенствам, то потери информации при дискретизации не происходит.

Двумерное дискретное преобразование Фурье может быть получено по аналогии с одномерным дискретным преобразованием Фурье. Оно имеет вид:

$$F(u,v) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x,y) \mathrm{e}^{-i2\pi\left(\frac{ux}{M} + \frac{vy}{N}\right)}.$$

Здесь $f(x,y)$ – цифровое изображение с размерами пиксельного поля $M \times N$. Последнее выражение должно быть вычислено для всех дискретных значений переменных $u$ и $v$, а именно, $u = 0,1,...,M-1$, $v = 0,1,...,N-1$.

Обратное преобразование Фурье получается аналогичным образом и имеет вид:

$$f(x,y) = \frac{1}{MN} \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F(u,v) \mathrm{e}^{i2\pi\left(\frac{ux}{M}+\frac{vy}{N}\right)},$$

для всех $x = 0,1,...,M-1$, $y = 0,1,...,N-1$. Данные выражения, как и в одномерном случае, представляют *пару двумерных дискретных преобразований Фурье*.

Отметим некоторые свойства двумерного дискретного преобразования Фурье.

**Свойство 1.** *Связь пространственных и частотных интервалов.*

Интервалы $\Delta T$ и $\Delta Z$ между отсчетами в пространственной области и соответствующие интервалы $\Delta u$ и $\Delta v$ в частотной области связаны соотношениями:

$$\Delta u = \frac{1}{M \Delta T}, \quad \Delta v = \frac{1}{N \Delta Z}.$$

Таким образом, интервалы между отсчетами в частотной области обратно пропорциональны величине интервалов в пространственной области и количеству отсчетов.

**Свойство 2.** *Сдвиг и поворот.*

$$f(x,y) \mathrm{e}^{i2\pi\left(\frac{u_0 x}{M}+\frac{v_0 y}{N}\right)} \Leftrightarrow F(u-u_0, v-v_0),$$

$$f(x-x_0, y-y_0) \Leftrightarrow F(u,v) \mathrm{e}^{-i2\pi\left(\frac{x_0 u}{M}+\frac{y_0 v}{N}\right)}.$$

**Свойство 3.** Периодичность.

$$F(u,v) = F(u+k_1 M, v) = F(u, v+k_2 N) = F(u+k_1 M, v+k_2 N),$$

$$f(x,y) = f(x+k_1 M, y) = f(x, y+k_2 N) = f(x+k_1 M, y+k_2 N).$$

**Свойство 4.** *Симметричность.*

Если $f(x,y)$ – *действительная* функция, то

$$F^*(u,v) = F(-u,-v),$$

т.е. Фурье-образ действительной функции является симметрично сопряженным.

*Доказательство.* $F^*(u,v) = \left[\sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x,y)\mathrm{e}^{-i2\pi\left(\frac{ux}{M}+\frac{vy}{N}\right)}\right]^* =$

$$= \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f^*(x,y)\mathrm{e}^{i2\pi\left(\frac{ux}{M}+\frac{vy}{N}\right)} = \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x,y)\mathrm{e}^{-i2\pi\left(\frac{(-u)x}{M}+\frac{(-v)y}{N}\right)} = F(-u,-v).$$

## 9.3. Фурье-спектр и Фурье-фаза

Напомним, что двумерное дискретное преобразование Фурье является в общем случае величиной комплексной, т.е.

$$F(u,v) = R(u,v) + iI(u,v),$$

поэтому его всегда можно представить в экспоненциальном виде:

$$F(u,v) = \left|F(u,v)\right| e^{i\varphi(u,v)},$$

где $\left|F(u,v)\right| = \sqrt{R^2(u,v) + I^2(u,v)}$ – амплитуда или *Фурье-спектр*, называемый также *частотным спектром*, а $\varphi(u,v) = \mathrm{arctg}\dfrac{I(u,v)}{R(u,v)}$ – *Фурье-фаза*.

Из свойства симметрии следует, в частности, что Фурье-спектр имеет центральную симметрию:

$$\left|F(u,v)\right| = F(-u,-v),$$

а Фурье-фаза имеет центральную антисимметрию:

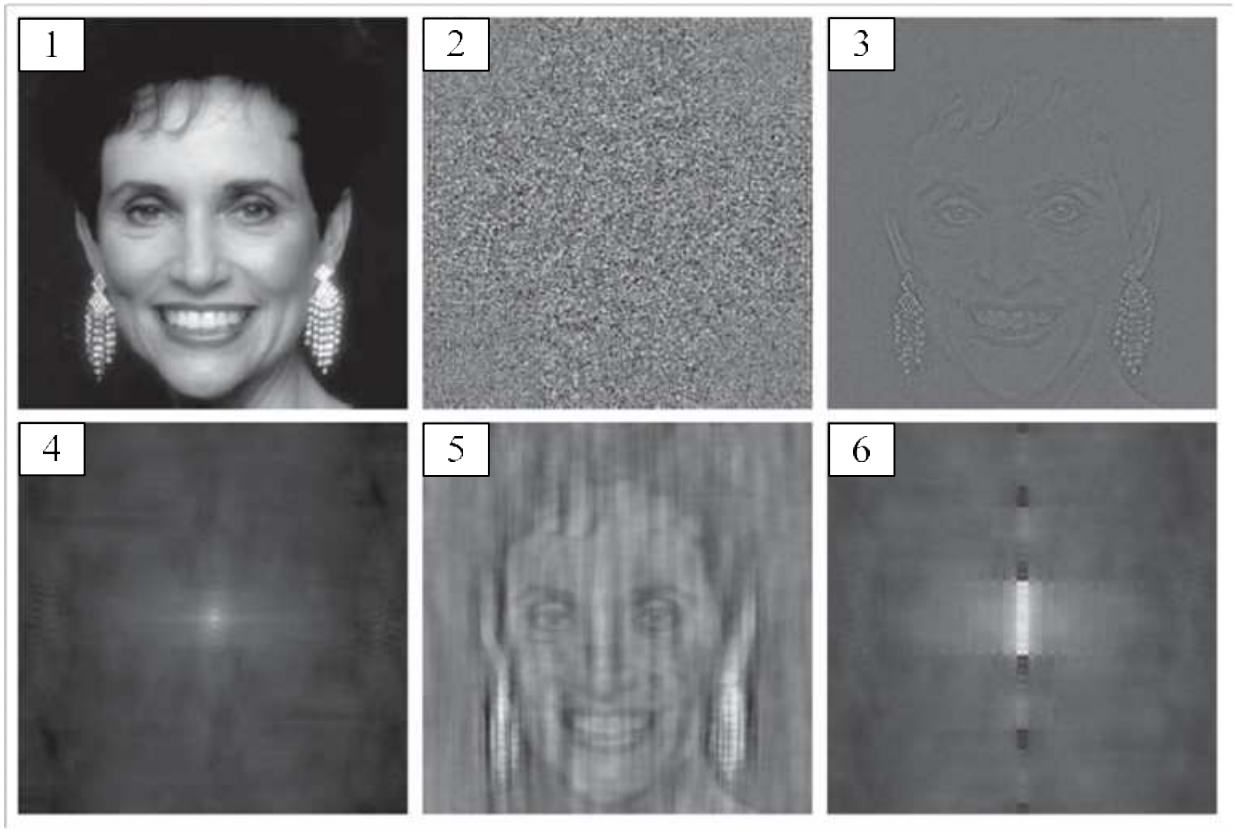
$$\varphi(u,v) = -\varphi(-u,-v).$$

Легко видеть также, что $F(0,0) = \sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x,y) = MN \cdot \overline{f}(x,y)$. Иными словами, член, который соответствует нулевой частоте, пропорционален среднему значению $\overline{f}(x,y)$ функции $f(x,y)$. Следовательно,

$$\left|F(0,0)\right| = MN\left|\overline{f}(x,y)\right|.$$

Как правило, за счет множителя $MN$ он достаточно велик, он является самой большой компонентой в спектре. Его называют *постоянной составляющей* преобразования Фурье.

Фурье-спектр и Фурье-фаза дискретного преобразования Фурье интерпретируются следующим образом. Компоненты спектра дают амплитуды синусоид, с помощью которых формируется изображение. Чем больше соответствующая амплитуда, тем больше вклад в формирование изображения синусоиды соответствующей частоты. В то же время фаза характеризует смещение каждой синусоиды относительно начала координат. Обычно для анализа изображений (и не только!) используется спектр дискретного преобразования Фурье. Фаза используется намного реже. Тем не менее, фазовая характеристика несет большой объем информации, очень важный для восстановления и понимания изображения. На следующем рисунке показана роль, которую играют эти две характеристики в восстановлении исходного изображения.




1 – исходное изображение;

2 – визуализированная Фурье-фаза;

3 – изображение, восстановленное только по фазовой характеристике (Фурье-спектр полагался тождественно равным единице);

4 – изображение, восстановленное только по спектру (Фурье-фаза полагалась тождественно равной нулю);

5 – изображение, восстановленное по фазе *исходного* изображения и спектру *другого* изображения (прямоугольника);

6 – изображение, восстановленное по спектру исходного изображения и фазе другого изображения (прямоугольника).

Данный пример показывает, насколько важную роль играет фаза в восстановлении исходного изображения.

## 9.4. Основы частотной фильтрации

По аналогии с одномерным случаем можно записать уравнение двумерной круговой свертки:

$$f(x,y)*h(x,y)=\sum_{m=0}^{M-1}\sum_{n=0}^{N-1}f(m,n)h(\mathrm{x}-m,y-n),$$

где $x=0,1,...,M-1$, $y=0,1,...,N-1$. Аналогичная теорема о свертке в двумерном случае задается соотношениями:

$$f(x,y)*h(x,y)=F(u,v)H(u,v),$$

$$f(x,y)h(x,y)=F(u,v)*H(u,v).$$

Эти соотношения являются основными в теории линейной фильтрации.

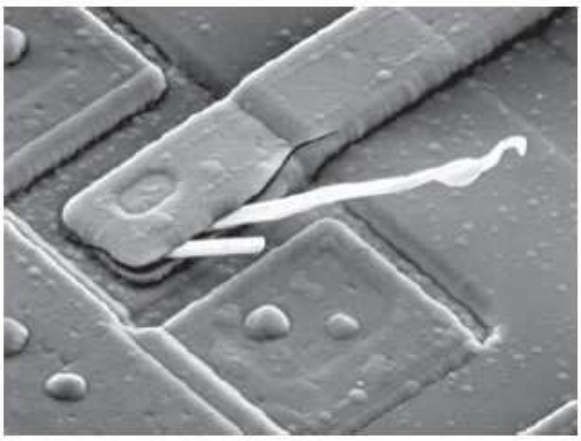
Методы фильтрации в частотной области основаны на преобразовании Фурье исходного изображения, затем изменении Фурье-образа для достижения конкретной цели модификации изображения, и последующем обратном дискретном преобразовании Фурье для получения в пространственной области уже модифицированного изображения. Эту схему обработки можно выразить в виде следующего так называемого уравнения частотной фильтрации:

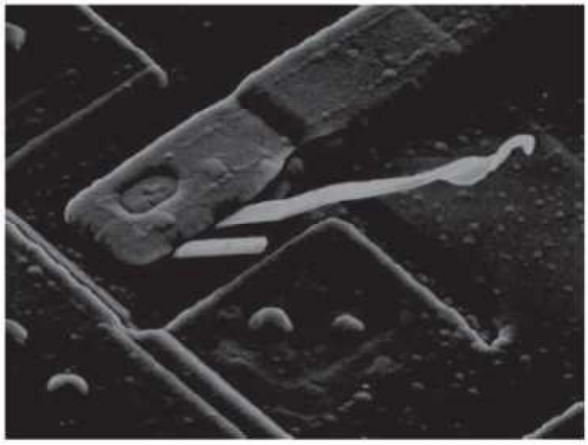
$$g(x,y)=\mathfrak{F}^{-1}\left[H(u,v)F(u,v)\right].$$

В этом уравнении $F(u,v)$ – дискретное преобразование Фурье исходного изображения $f(x,y)$, $H(u,v)$ – фильтр-функция (просто фильтр или передаточная функция фильтра), $\mathfrak{F}^{-1}$ – обратное преобразование Фурье, $g(x,y)$ – изображение, полученное в результате фильтрации.

Рассмотрим процесс частотной фильтрации более подробно.

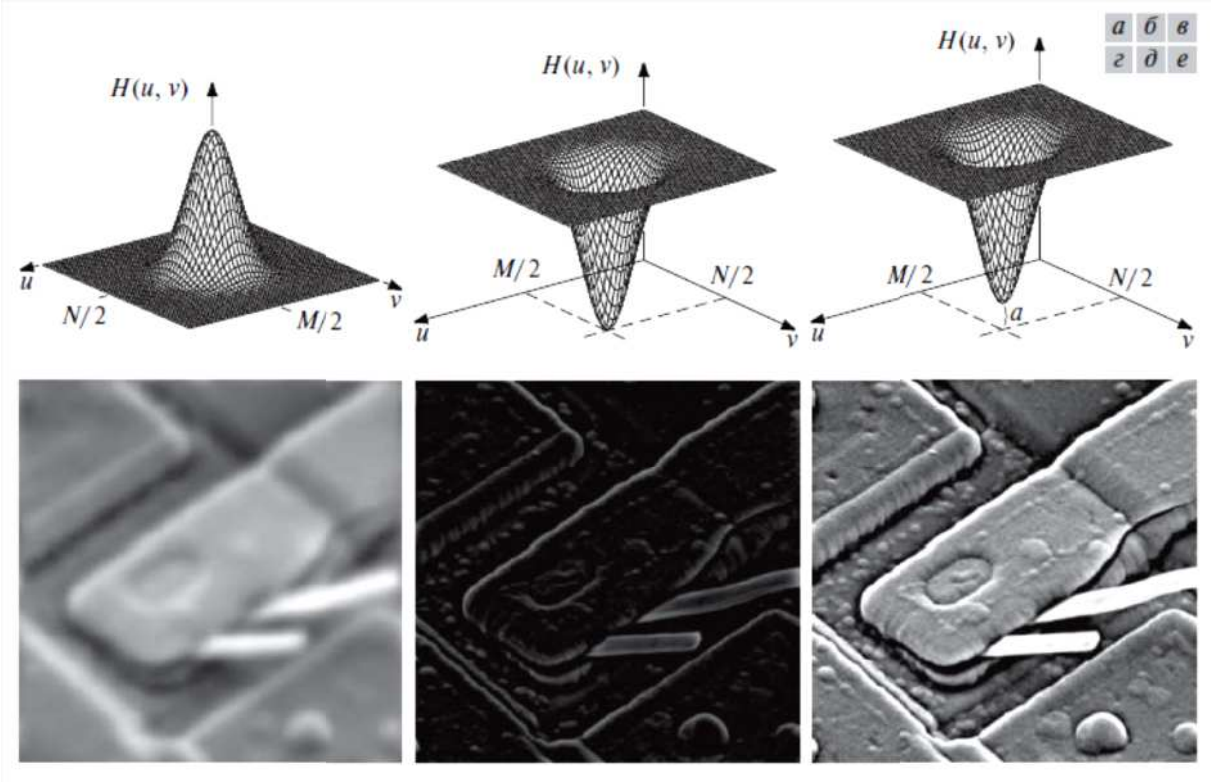
Простейшим фильтром является фильтр $H(u,v)$, который равен 0 в точке $u=v=0$ и 1 во всех остальных точках. При умножении $F(u,v)H(u,v)$ такой фильтр будет аннулировать член, который несет информацию о средней яркости, т.е. постоянную составляющую, и оставлять неизменными все остальные частоты. В этом случае средняя яркость изображения будет равна нулю. Разумеется, в этом случае будут получаться отрицательные значения яркостей. Однако при отображении такие значения будет заменяться нулевыми, поэтому отображаемая картинка будет лишь условно соответствовать результату преобразования. На рисунке ниже показан пример такой фильтрации.





Низкие частоты в преобразовании Фурье связаны с областями медленно меняющейся яркости. Это могут быть, например, дома на улице, стены комнаты, небо при хорошей погоде. Высокие частоты, наоборот, порождаются относительно резкими изменениями яркости, такими как контуры и шум. По этой причине можно ожидать, что фильтр $H(u,v)$, подавляющий высокие частоты и пропускающий низкие, будет размывать изображение. Такой фильтр называется *фильтром низких частот*. Фильтр с противоположными характеристиками, который называется *фильтром высоких частот*, будет повышать контраст резких деталей. Типы фильтров и получаемые эффекты показаны на следующем рисунке.

Следует обратить внимание, что на последнем рисунке постоянная составляющая подавляется не полностью, а лишь частично (определяется параметром $a$), поэтому изображение выглядит более светлым, чем на предыдущем рисунке.



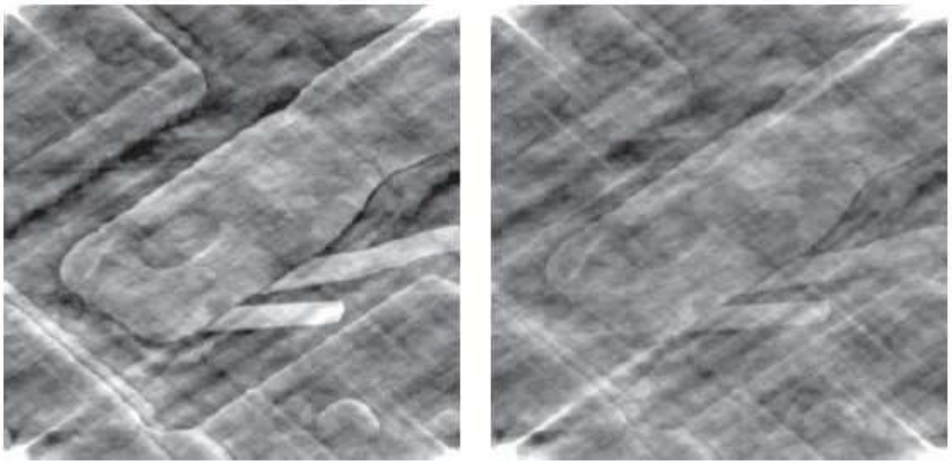

В заключение рассмотрим, что происходит с Фурье-фазой при фильтрации в частотной области. Поскольку преобразование Фурье является в общем случае комплексным, его можно представить в виде:

$$F(u,v) = R(u,v) + iI(u,v),$$

тогда

$$g(x,y) = \mathfrak{F}^{-1}\left[H(u,v)R(u,v) + iH(u,v)I(u,v)\right].$$

Значит, выражение для фазы, очевидно, не изменится, поскольку множитель $H(u,v)$ сократится. Фильтры, которые не меняют фазу, называются фильтрами *с нулевым сдвигом фазы*. Изображение, в котором была скалярным образом изменена фаза, показано на нижеследующем рисунке.



В заключение отметим, что в частотной области, как правило, проще изучать соотношение между изображением и его частотными свойствами. Более того, некоторые задачи трудно формулируются и решаются в пространственной области, в то время как в частной области они оказываются достаточно простыми. Решив задачу в частотной области, можно иногда достаточно просто подобрать соответствующий фильтр в пространственной области, что позволяет решить соответствующую задачу обработки с меньшими вычислительными затратами. При этом соответствующие пространственные фильтры представляют собой небольшие маски, реализующие преобразование изображения, аналогичное преобразованию в частотной области.

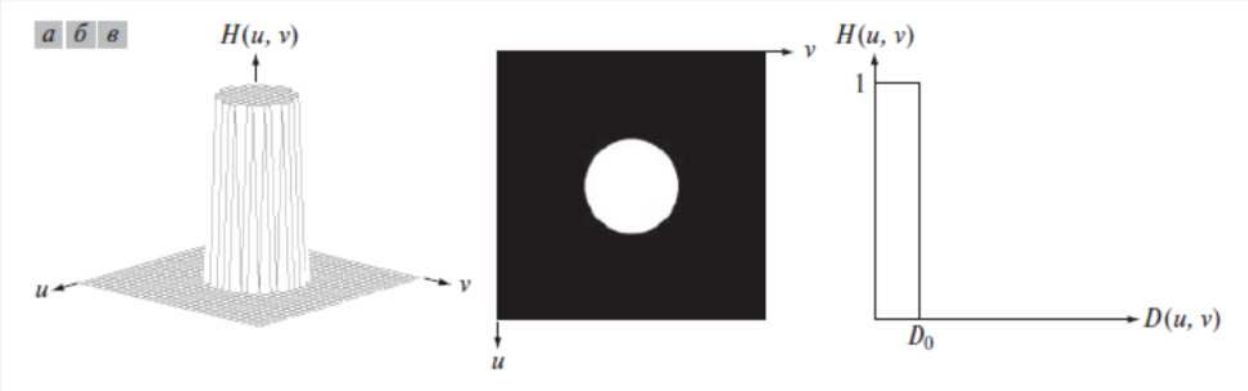
# Глава 10. Низкочастотная фильтрация изображений

Основной вклад в высокие частоты преобразования Фурье рассматриваемого изображения дают контуры на изображении, резкие изменения яркости, а также шумы. Поэтому для размывания изображения в частотной области используются *фильтры низких частот*, а сам процесс называется *низкочастотной фильтрацией*. Рассмотрим некоторые из таких фильтров.

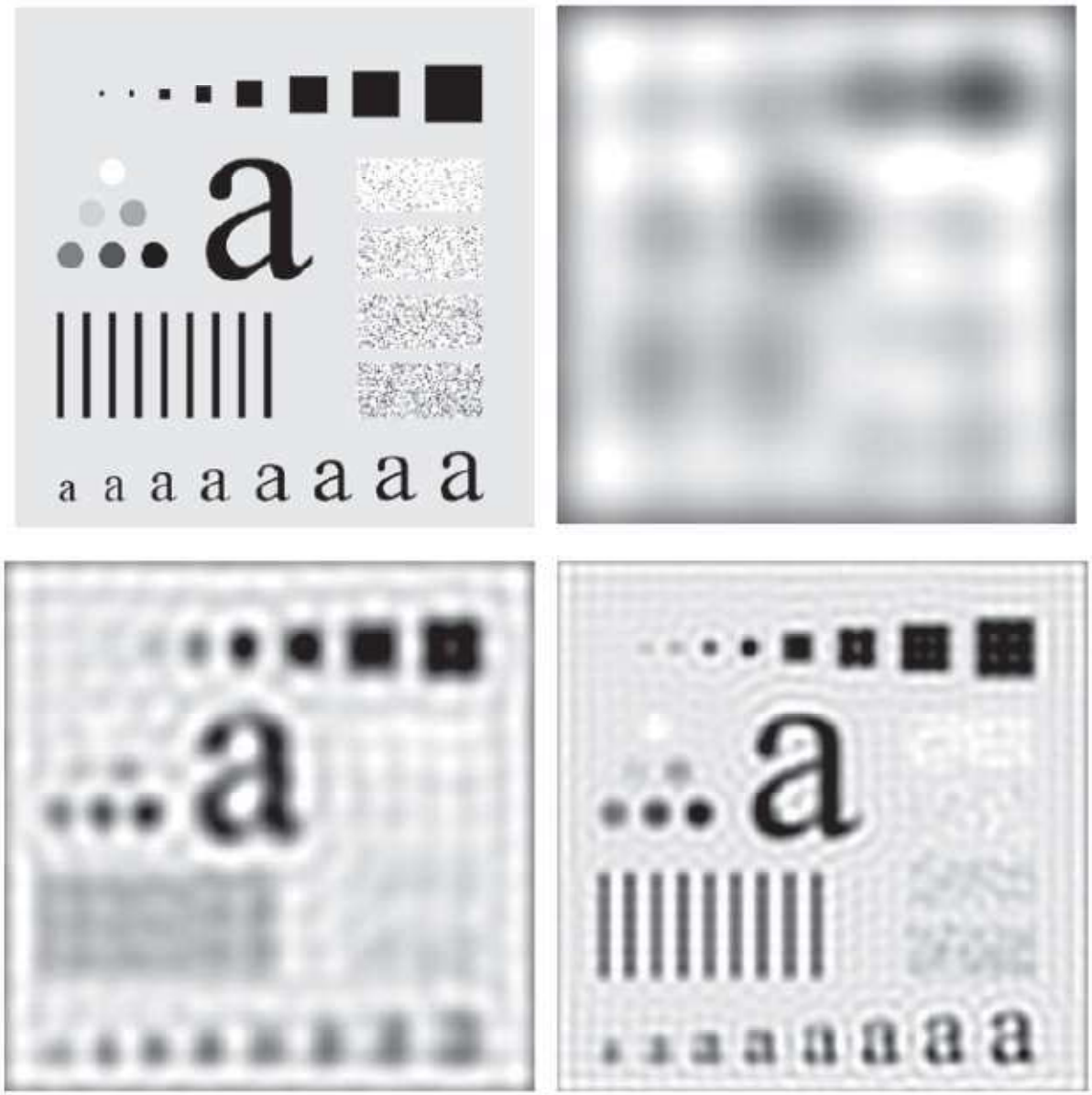
## 10.1. Идеальный фильтр низких частот

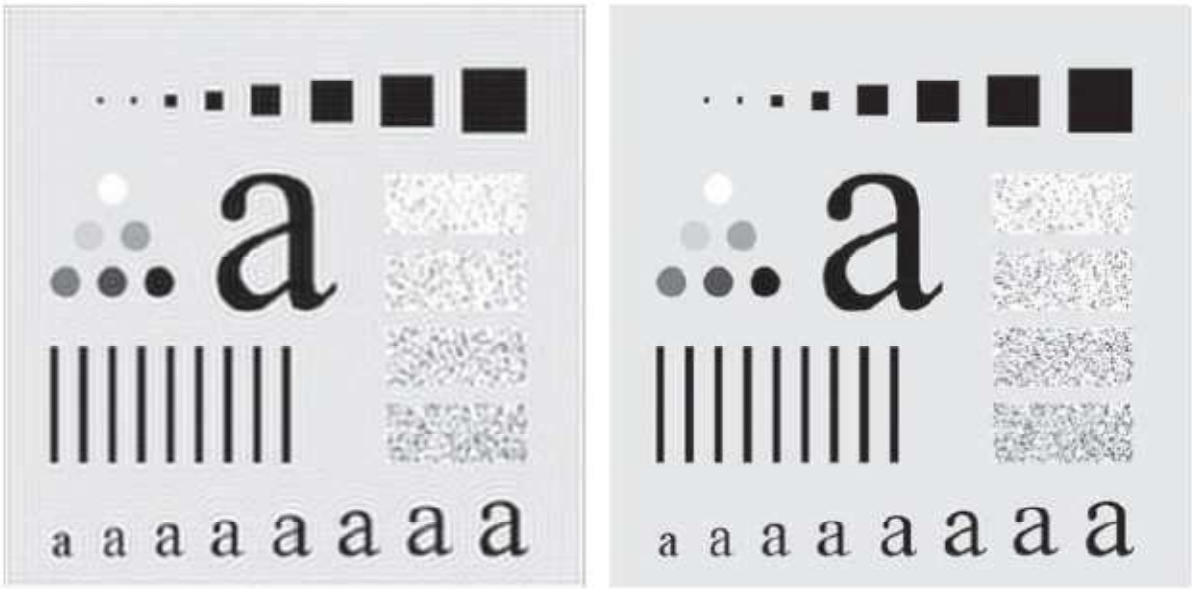
| | |
|---|---|
| **Определение** | *Идеальным фильтром низких частот* называется двумерный фильтр, который пропускает без затухания все частоты, находящиеся внутри круга радиуса $D_0$ с центром в начале координат, и удаляет все частоты вне этого круга, т.е. $$H(u,v)=\begin{cases}1, & при\ D(u,v)\le D_0,\\ 0, & при\ D(u,v)>D_0,\end{cases}$$ где $D_0$ – положительная константа, $D(u,v)$ – расстояние от точки $(u,v)$ в частотной области до начала координат. |

Все частоты внутри круга в идеальном фильтре проходят без изменения, а частоты за его пределами полностью подавляются. Точка $D_0$ называется *частотой среза*. Реализовать идеальный фильтр физически в электронных устройствах невозможно по техническим причинам, однако такой фильтр вполне может быть реализован программно в процессе компьютерных вычислений. Действие идеального фильтра поясняется графически на следующем рисунке.




Чем меньше частота среза $D_0$, тем большему огрублению подвергается изображение. Наоборот, если частота $D_0$ велика, то практически весь частотный спектр оказывается сохраненным идеальным фильтром. Следующий рисунок демонстрирует размытие изображения в зависимости от частоты среза.





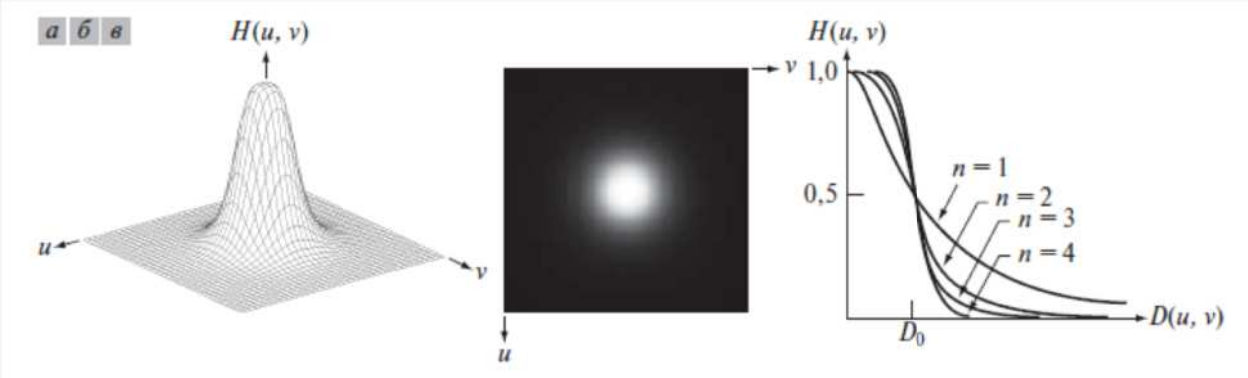
В левом верхнем углу приведено исходное изображение. Слева направо и сверху вниз показаны результаты действия идеального фильтра с частотами среза соответственно 10, 30, 60, 160, 460.

## 10.2. Фильтры низких частот Баттерворта

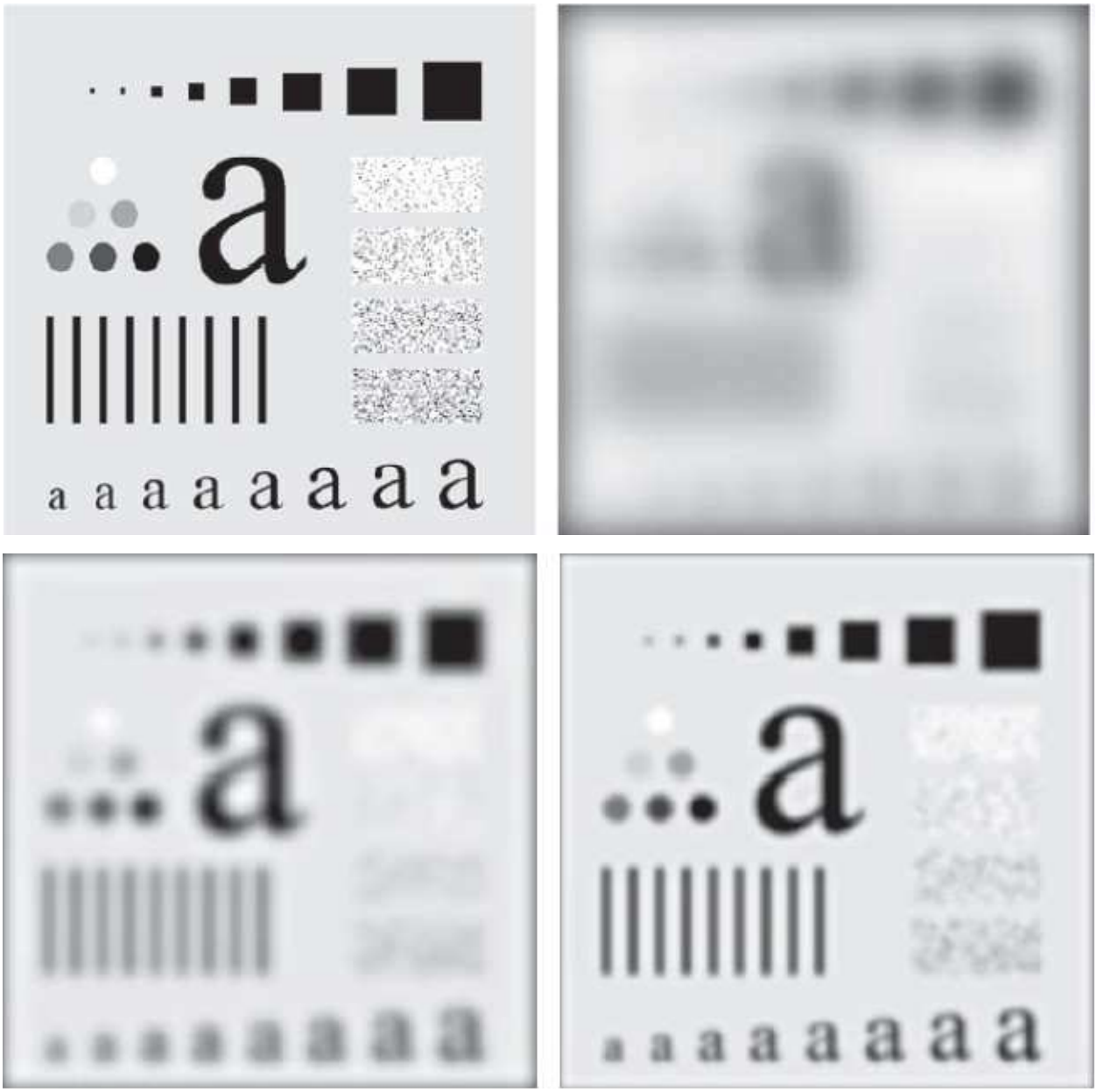
| | |
|---|---|
| **Определение** | *Низкочастотным фильтром Баттерворта порядка* $n$ с частотой среза $D_0$ от начала координат называется фильтр, задаваемый формулой $$H(u,v)=\frac{1}{1+\left(\frac{D(u,v)}{D_0}\right)^{2n}}.$$ |

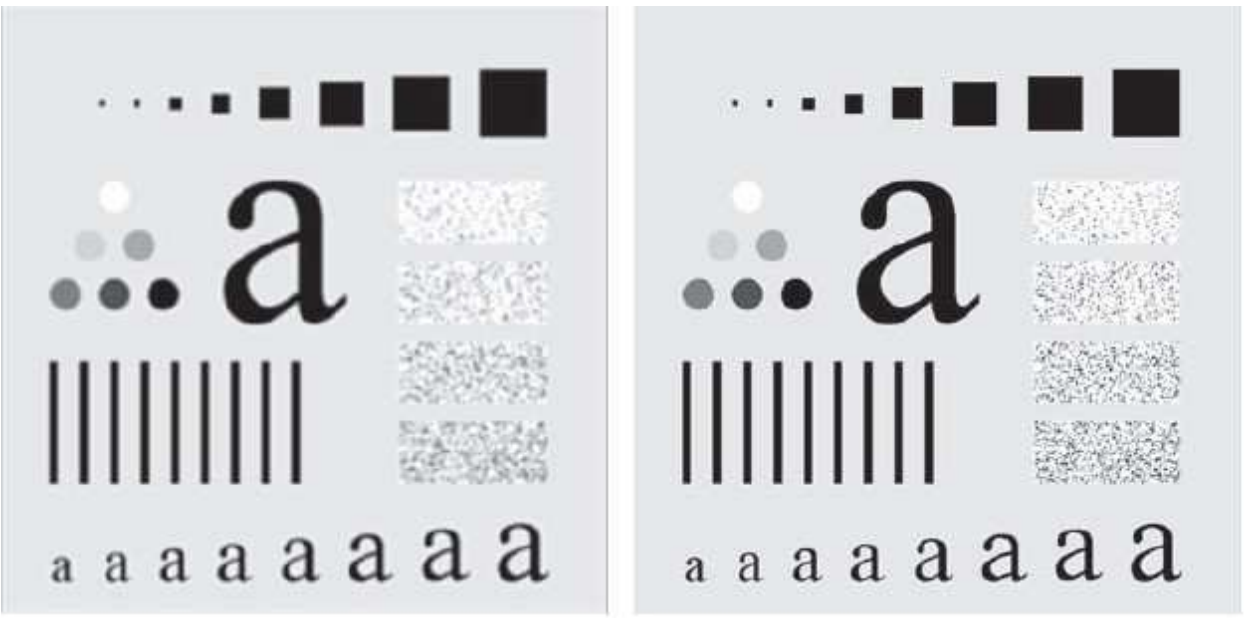
На следующем рисунке показан трехмерный график, его полутоновое изображение, а также радиальные профили функции, определяемой фильтром низких частот Баттерворта.

Заметим, что фильтр Баттерворта в отличие от идеального фильтра является гладкой функцией, т.е. не имеет точек разрыва. Возникает вопрос, что понимать под частотой среза такого фильтра. Обычно за частоту среза принимают точку $D_0$, в которой функция $H(u,v)$ становится меньше заранее определенной доли ее максимального значения. В приведенном выше определении, когда $D(u,v)$ становится равным $D_0$, функция принимает значение, равное половине ее максимального значения, которое, очевидно равно 1.




На следующем рисунке показаны примеры применения фильтра Баттерворта с частотами среза, такими же, как и для идеального фильтра в предыдущем примере.

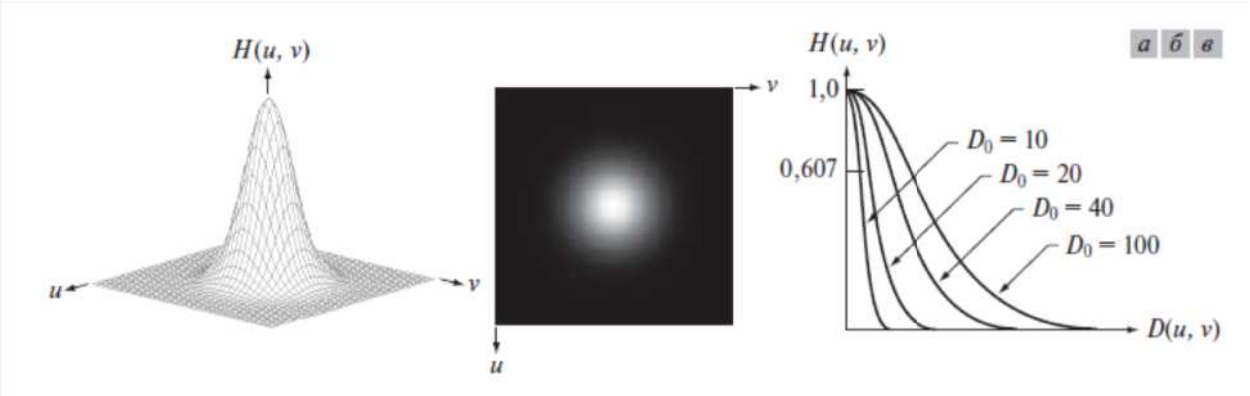




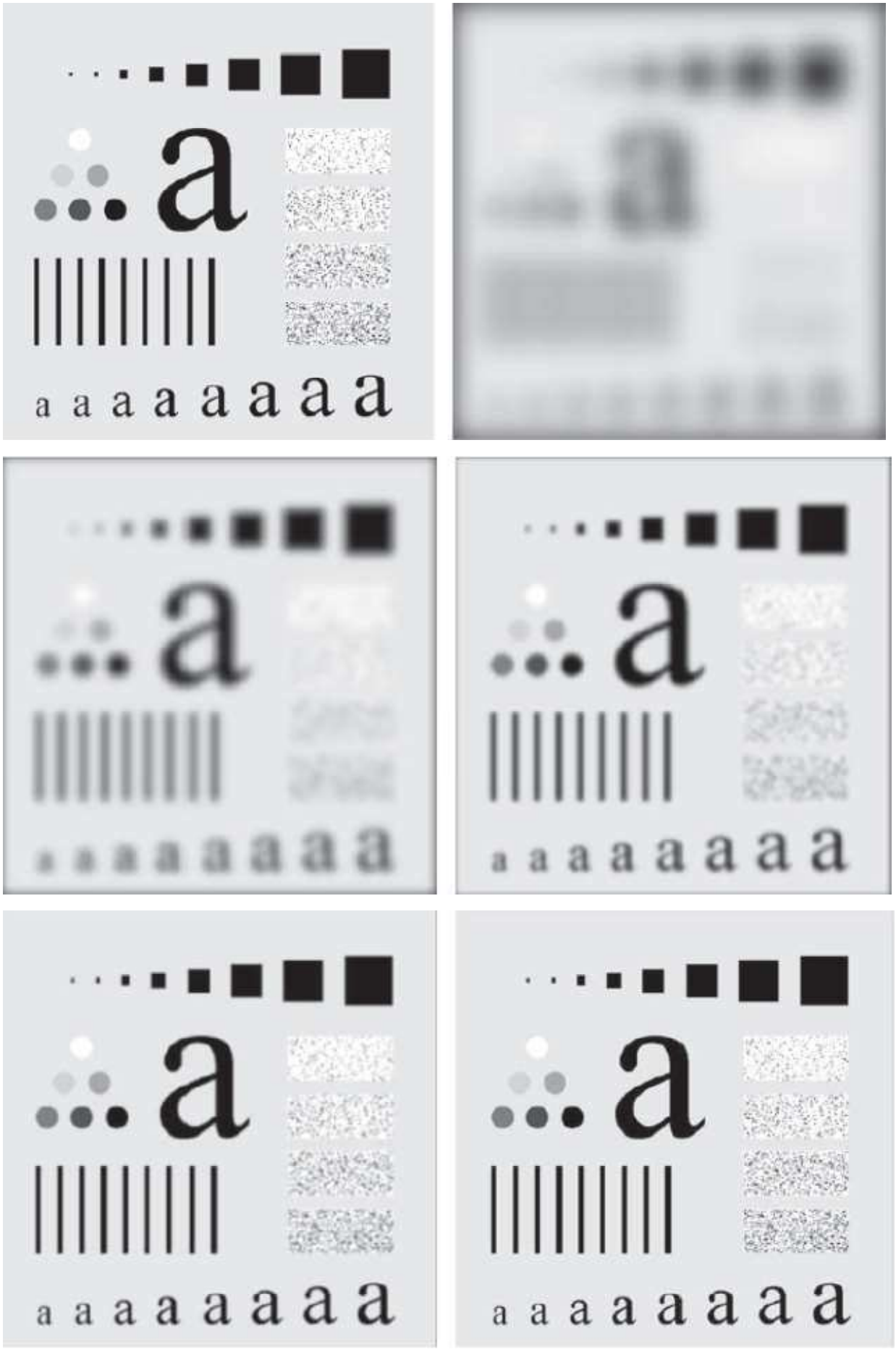
## 10.3. Гауссовы фильтры низких частот

| | |
|---|---|
| **Определение** | *Гауссовым низкочастотным фильтром* называется фильтр, задаваемый формулой $$H(u,v)=e^{-\frac{D^2(u,v)}{2D_0^2}},$$ где $D^2(u,v)$ – расстояние от центра прямоугольника. |

На следующем рисунке показан трехмерный график, его полутоновое изображение, а также радиальные профили функции, определяемой фильтром низких частот Гаусса.




Параметр $D_0$ задает размах функции вокруг центра. Этот параметр задает частоту среза. При $D(u,v)=D_0$ значение принимает значение равное 0.607 от своего максимального значения.



Результаты применения низкочастотного фильтра Гаусса к тому же изображению при тех же самых частотах среза показаны на предыдущем рисунке. Можно видеть, что результаты сопоставимы с аналогичными результатами применения фильтра Баттерворта 2-го порядка. Однако фильтр Гаусса обладает рядом преимуществ, которые связаны с тем, что он не дает некоторых специфических искажений, присущих другим фильтрам.

# Глава 11. Высокочастотная фильтрация изображений

В отличие от низкочастотной фильтрации при высокочастотной фильтрации подавляются низкочастотные составляющие изображения, не изменяя или почти не изменяя область высоких частот. Практически во всех случаях передаточная функция $H_H(u,v)$ высокочастотного фильтра может быть получена из передаточной функции $H_L(u,v)$ любого низкочастотного фильтра с помощью выражения

$$H_H(u,v) = 1 - H_L(u,v).$$

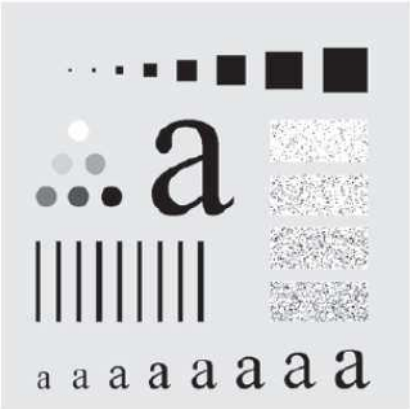
Из этого выражения видно, что содержательно действие высокочастотного фильтра инвертируется по отношению к низкочастотному фильтру.

## 11.1. Идеальный фильтр высоких частот

| | |
|---|---|
| **Определение** | *Идеальным фильтром высоких частот* называется двумерный фильтр, который удаляет все частоты, находящиеся внутри круга радиуса $D_0$ с центром в начале координат, и пропускает без затухания все частоты вне этого круга, т.е. $$H(u,v) = \begin{cases} 0, & при\ D(u,v) \le D_0, \\ 1, & при\ D(u,v) > D_0, \end{cases}$$ где $D_0$ – положительная константа, $D(u,v)$ – расстояние от точки $(u,v)$ в частотной области до начала координат. |

Здесь, как и раньше $D_0$ – частота среза. Действие высокочастотного фильтра прямо противоположно действию низкочастотного фильтра. Действие высокочастотное фильтра иллюстрируется на следующем рисунке. Пер-

вый рисунок – это исходное изображение. Следующие три рисунка – это результат фильтрации идеальным фильтром с частотой среза $D_0 = 30, 60, 160$.

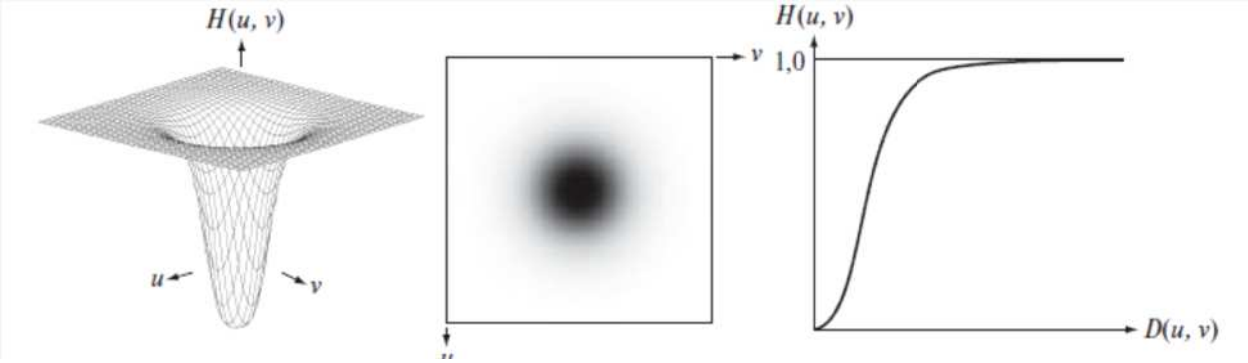


## 11.2. Фильтр высоких частот Баттерворта

| | |
|---|---|
| **Определение** | *Низкочастотным фильтром Баттерворта порядка* $n$ с частотой среза $D_0$ от начала координат называется фильтр, задаваемый формулой $$H(u,v) = \frac{1}{1 + \left(\frac{D_0}{D(u,v)}\right)^{2n}}.$$ |

Изображение и профиль передаточной функции Баттерворта представлены на рисунке ниже.

На следующем рисунке показаны результаты фильтрации с помощью высокочастотного фильтра Баттерворта с теми же частотами среза $D_0 = 30, 60, 160$.



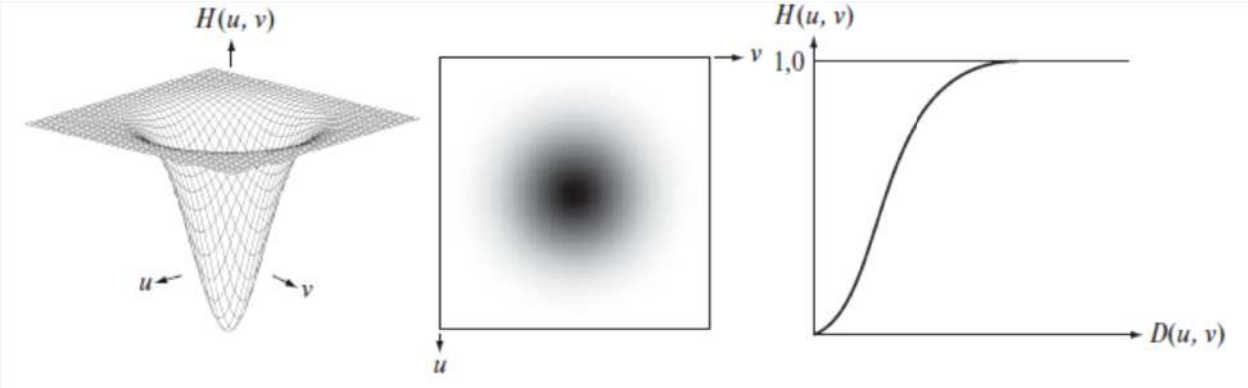

Фильтр Баттерворта демонстрирует более гладкое поведение по сравнению с идеальным фильтром.

## 11.3. Гауссовы фильтры высоких частот

| | |
|---|---|
| **Определение** | *Гауссовым низкочастотным фильтром* называется фильтр, задаваемый формулой $$H(u,v)=1-e^{-\frac{D^2(u,v)}{2D_0^2}},$$ где $D^2(u,v)$ – расстояние от центра прямоугольника. |

Изображение и профиль Гауссовой передаточной функции представлены на рисунке ниже.

На следующем рисунке показаны также результаты фильтрации с помощью высокочастотного Гауссова фильтра с аналогичными частотами среза $D_0 = 30, 60, 160$.



Гауссовы фильтры дают очень хорошее качество фильтрации для небольших объектов и тонких полос. Например, они являются очень удобными для обработки изображений папиллярных узоров при анализе дактилоскопических изображений.

# РАЗДЕЛ 4.

# МОРФОЛОГИЧЕСКАЯ ОБРАБОТКА ИЗОБРАЖЕНИЙ

## ГЛАВА 12. БАЗОВЫЕ ПОНЯТИЯ И ОПЕРАЦИИ МОРФОЛОГИЧЕСКОЙ ОБРАБОТКИ ИЗОБРАЖЕНИЙ

### 12.1. Основные определения

Математическая морфология – это совокупность методов извлечения некоторых компонент изображения, которые полезны для его представления и описания. Такими компонентами могут быть границы объектов на изображении, выпуклые оболочки, характерные точки, топологические характеристики, скелеты объектов. К морфологии изображений относят также некоторые методы обработки, такие как утончение, усечение и др.

В отличие от предыдущих методов, в которых на вход поступало изображение, а на выходе также было изображение, но измененное некоторым образом, в математической морфологии на вход также поступает изображение, а на выходе получаются некоторые его атрибуты, которые извлекаются из этого изображения. В некотором смысле морфология служит для получения таких характеристик, которые могут помочь в интерпретации изображения, понимании его смысла.

Основным языком для описания морфологических понятий и методов является язык алгебры множеств. При этом исторически математическая морфология развивалась применительно к двоичным изображениям. Мы в дальнейшем также будем подразумевать только двоичные изображения, если не будет оговариваться иное.

Основные морфологические операции строятся на двух простых базовых преобразованиях.

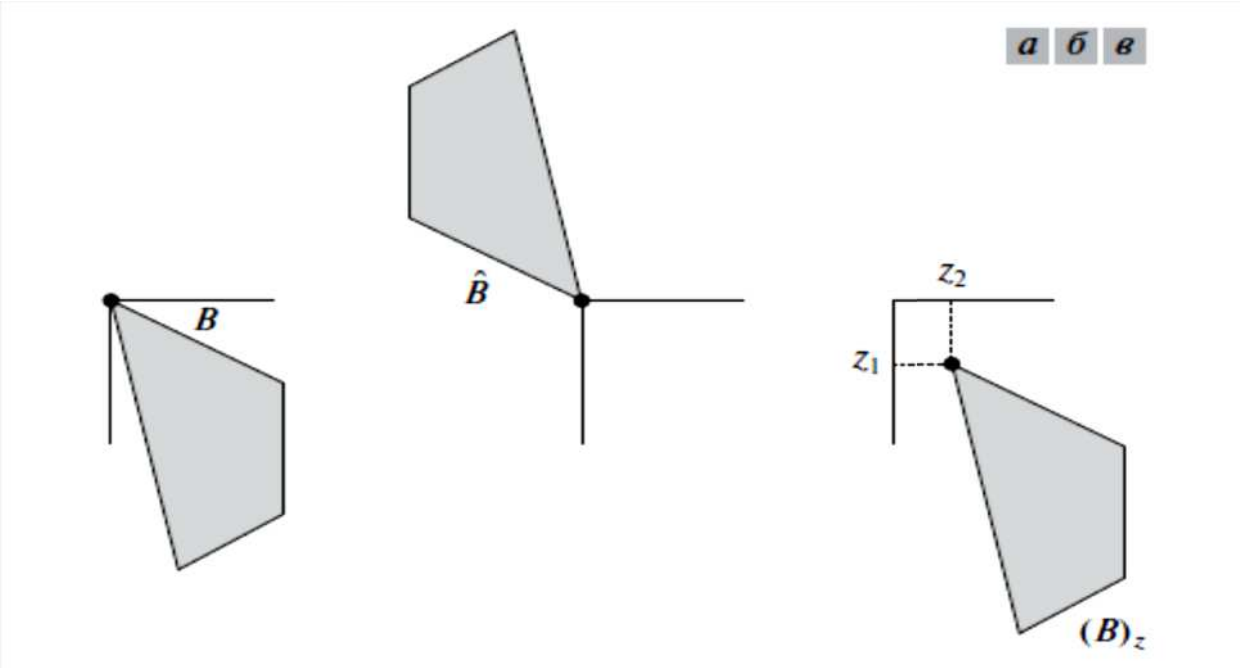
| | |
|---|---|
| **Определение** | *Центрально-симметричным отражением* множества $B = \{(x, y)\}$ называется множество $$\hat{B} = \{(-x, -y) \mid (x, y) \in B\}.$$ |

Множество $\hat{B}$ получается, таким образом, отражением каждой точки множества $B$ относительно начала координат.

| | |
|---|---|
| **Определение** | *Параллельным переносом* или *сдвигом* множества $B=\{(x,y)\}$ в точку $z=(z_1,z_2)$ называется множество $$B_z=\{(x+z_1,y+z_2)\,\vert\,(x,y)\in B\}.$$ |

Множество $B_z$ получается из $B$ путем смещения каждой точки на вектор $(z_1\ z_2)$.

Данные преобразования иллюстрируются на следующем рисунке.



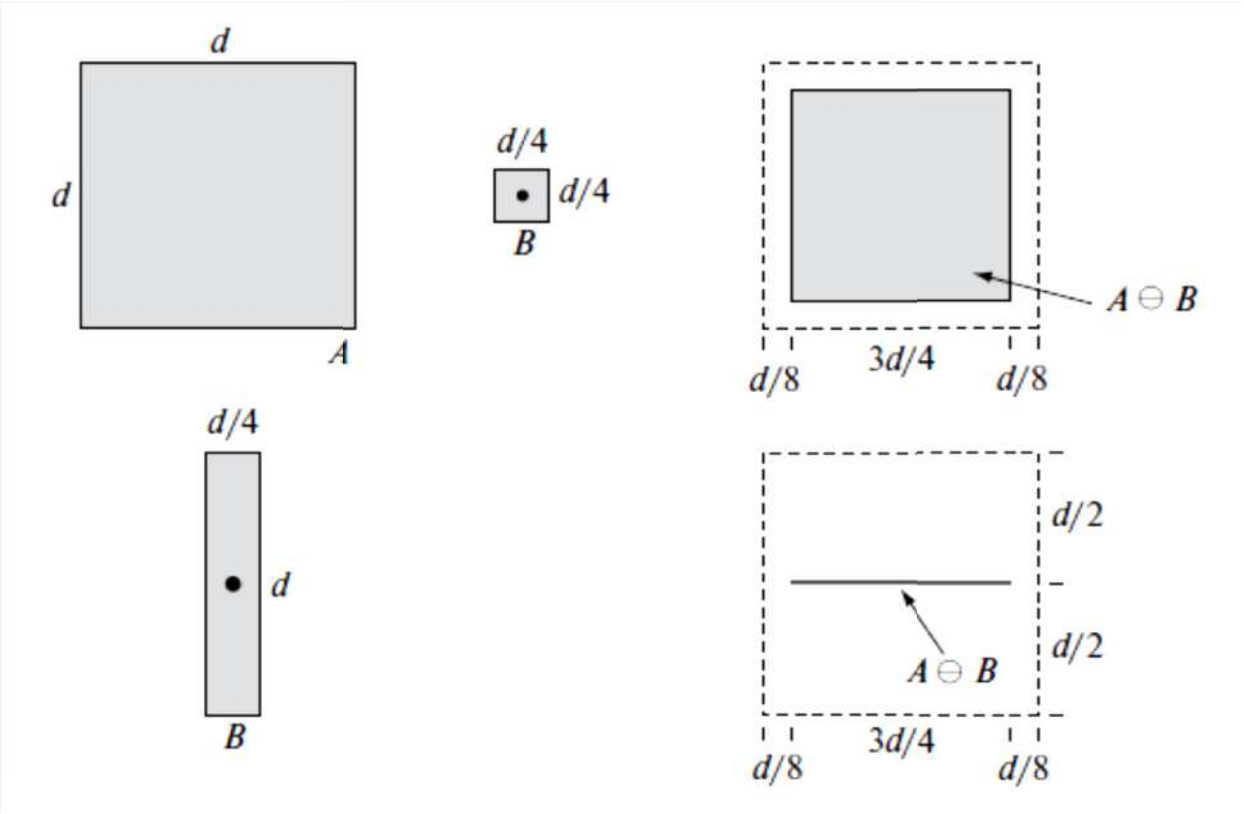

Рассмотрим основные морфологические операции.

## 12.2. Эрозия и дилатация

Эрозия и дилатация – две основные базовые морфологические операции, на которых строятся многие алгоритмы морфологической обработки изображений.

| | |
|---|---|
| **Определение** | *Эрозией* множества $A$ по множеству $B$ называется множество, обозначаемое $A\ominus B$ и определяемое следующим образом: $$A\ominus B=\{z=(x,y)\,\vert\,B_z\subseteq A\}.$$ |

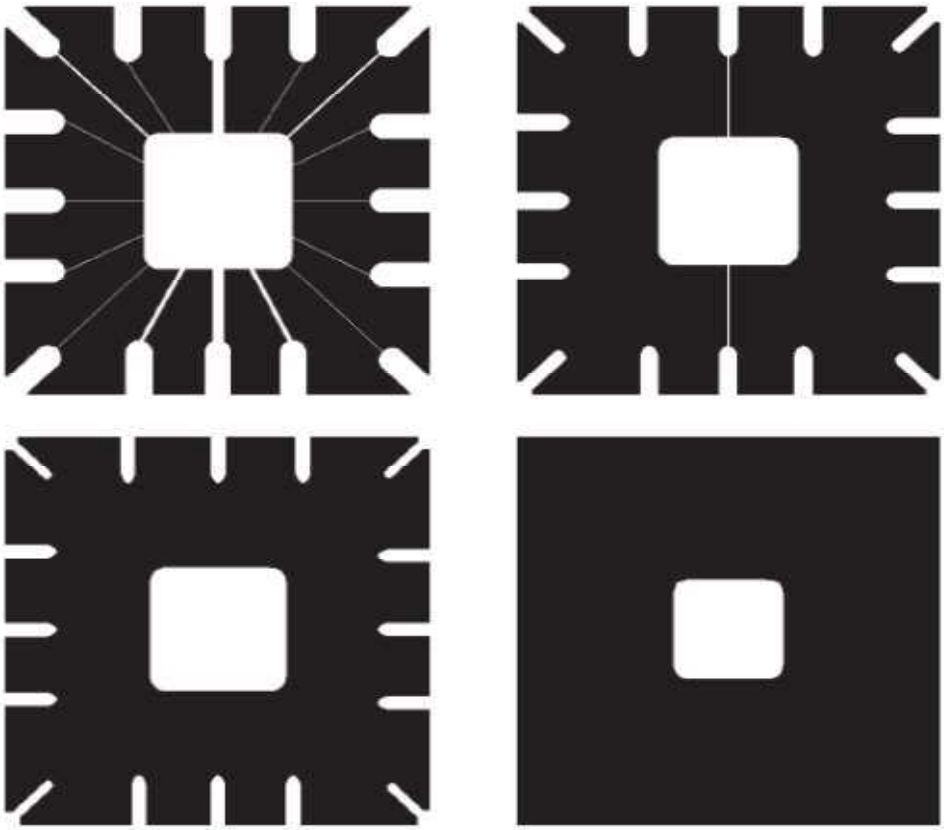
Иными словами, эрозия множества $A$ по множеству $B$ – это множество всех таких точек $z=(x,y)$, что если множество $B$ сдвигается в это точку, то оно полностью содержится во множестве $A$. Следующий рисунок демонстрирует два примера применения операции эрозии.



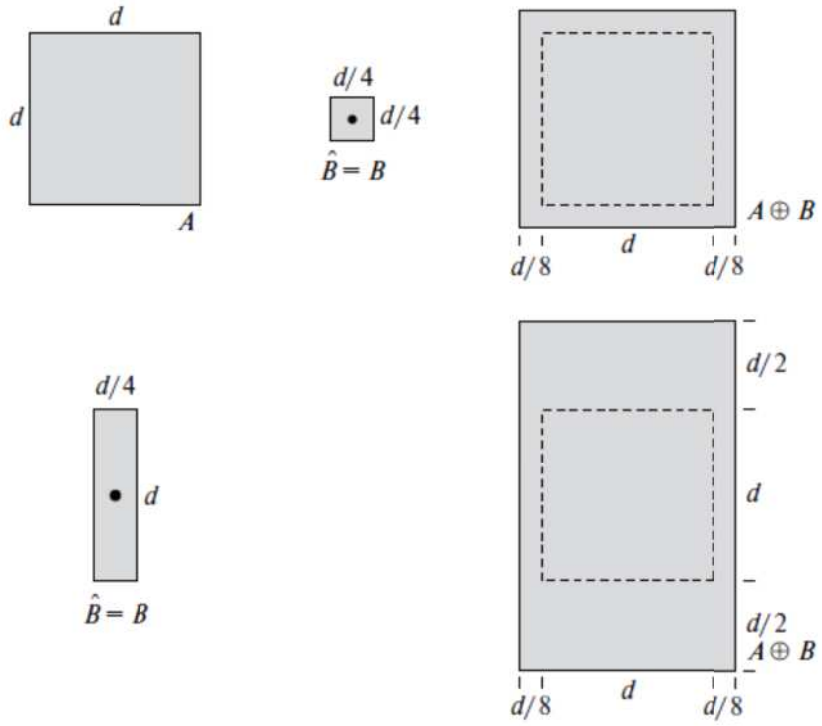

На следующем рисунке показаны примеры применения операции эрозии для удаления мелких и тонких деталей. Вначале показан исходный рисунок микросхемы. Следующие три рисунка – результаты эрозии с квадратным примитивом соответственно размеров 11x11, 15x15, 45x45.

Как показывает этот пример, эрозия приводит к тому, что объекты на двоичном изображении становятся меньше и тоньше. По существу, эрозию можно рассматривать как операцию *морфологической фильтрации*, которая отфильтровывает (удаляет) детали изображения с размерами меньше структурообразующего множества.

| | |
|---|---|
| **Определение** | *Дилатацией* множества $A$ по множеству $B$ называется множество, обозначаемое $A \oplus B$ и определяемое следующим образом: $$A \oplus B = \left\{ z=(x,y) \mid \hat{B}_z \cap A \neq \varnothing \right\}.$$ |



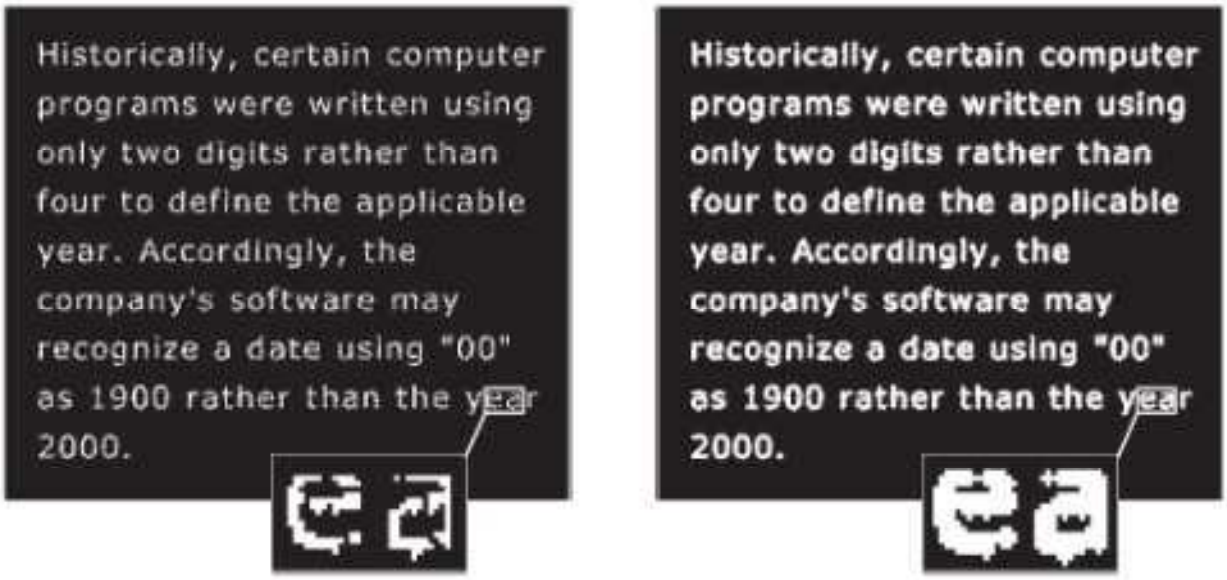
Таким образом, результат дилатации – это множество точек таких, что множество $\hat{B}$, центрально-симметричное к множеству $B$, помещенное в эту точку имеет непустое пересечение с множеством $A$. Следующий рисунок демонстрирует два примера применения операции дилатации.



В отличие от эрозии, которая приводит к сужению или утончению, дилатация «расширяет» или «утолщает» объекты на двоичном изображении. Конкретный вид и степень такого утолщения определяются формой используемого примитива. На предыдущем рисунке приведено то же множество, что и в примере для случая эрозии. Изображен примитив, который симметричен относительно центра, поэтому его центрально-симметричное отражение совпадает с ним самим, т.е. $\hat{B} = B$.

Пунктирной линией на рисунке для сравнения показаны границы исходного множества, а сплошная линия соответствует тем предельным значениям $z$, при выходе центра множества $\hat{B}$ за которые пересечение этого множества с множеством $A$ оказывается пустым. Следовательно, множество всех точек, находящихся внутри обозначенной сплошной линией границы и на ней, образует дилатацию множества $A$ по множеству $B$. В нижней строке на этом рисунке изображен примитив, построенный с таким расчетом, чтобы добиться большей дилатации в вертикальном направлении, чем в горизонтальном. Внизу справа показан результат дилатации по такому примитиву.

На следующем рисунке приведен пример практической дилатации для устранения разрывов линий текста путем их перекрытия с помощью показанного ниже примитива.




| | | |
|---|---|---|
| 0 | 1 | 0 |
| 1 | 1 | 1 |
| 0 | 1 | 0 |

Отметим, что введенные операции эрозии и дилатации являются, очевидно, примерами нелинейных операций.

Отметим в заключение, что операции эрозии и дилатации, как показывают предыдущие примеры, являются в некотором смысле *двойственными* операциями над изображениями. Эта связь выражается следующими соотношениями:

$$\overline{A \ominus B} = \overline{A} \oplus \hat{B},$$

$$\overline{A \oplus B} = \overline{A} \ominus \hat{B}.$$

Если примитив $B$ симметричен относительно своей центральной точки, то последние соотношения оказываются еще проще:

$$\overline{A \ominus B} = \overline{A} \oplus B,$$

$$\overline{A \oplus B} = \overline{A} \ominus B.$$

## 12.3. Размыкание и замыкание

Рассмотрим две важные морфологические операции, основанные на операциях эрозии и дилатации. Это операции размыкания и замыкания.

| | |
|---|---|
| **Определение** | *Размыканием* множества $A$ по множеству (примитиву) $B$ называется множество, обозначаемое $A \circ B$ и определяемое следующим образом: $$A \circ B = \left(A \ominus B\right) \oplus B.$$ |

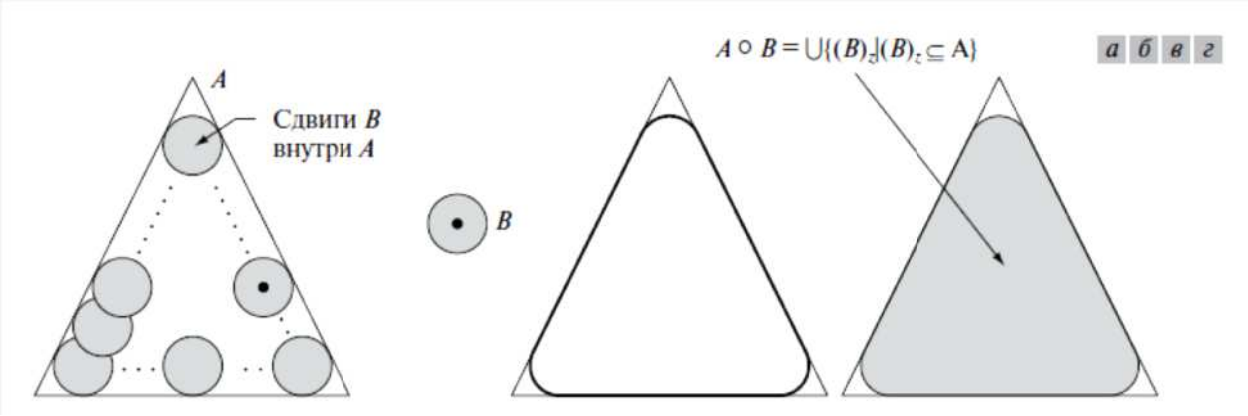
Иными словами, размыкание множества $A$ по примитиву $B$ – это эрозия $A$ по $B$ с последующей дилатацией результата по $B$.

| | |
|---|---|
| **Определение** | *Замыканием* множества $A$ по множеству (примитиву) $B$ называется множество, обозначаемое $A \bullet B$ и определяемое следующим образом: $$A \bullet B = \left(A \oplus B\right) \ominus B.$$ |

Таким образом, замыкание – это дилатация $A$ по $B$ с последующей эрозией результата по $B$.

В общем случае *размыкание* сглаживает контуры объекта, обрывает узкие перешейки и ликвидирует выступы небольшой ширины. *Замыкание* также приводит к сглаживанию участков контуров объектов, но, в отличие от размыкания, в общем случае «заливает» узкие промежутки и длинные углубления малой ширины, а также ликвидирует небольшие «дырки» и заполняет разрывы контура.

Операция размыкания допускает простую геометрическую интерпретацию, показанную на следующем рисунке. Для простоты примитив $B$ имеет форму круга. Тогда границу множества $A \circ B$ образуют максимально близкие к границе $A$ точки множества $B$ при «обкатывании» примитивом $B$ этой границы изнутри. На третьем и четвертом рисунках показаны соответственно граница и все множество, являющееся результатом построения размыкания.




Аналогичным образом можно построить геометрическую интерпретацию операции замыкания. В этом случае граница нового множества получается при обкатывании круглым примитивом границы множества снаружи.

Отметим без доказательства некоторые свойства введенных операций.

Свойства операция размыкания:

1. $A \circ B \subseteq A$.
2. Из $C \subseteq D$ следует, что $C \circ B \subseteq D \circ B$.
3. $(A \circ B) \circ B = A \circ B$.

Свойства операция замыкания:

1. $A \subseteq A \bullet B$.
2. Из $C \subseteq D$ следует, что $C \bullet B \subseteq D \bullet B$.
3. $(A \bullet B) \bullet B = A \bullet B$.

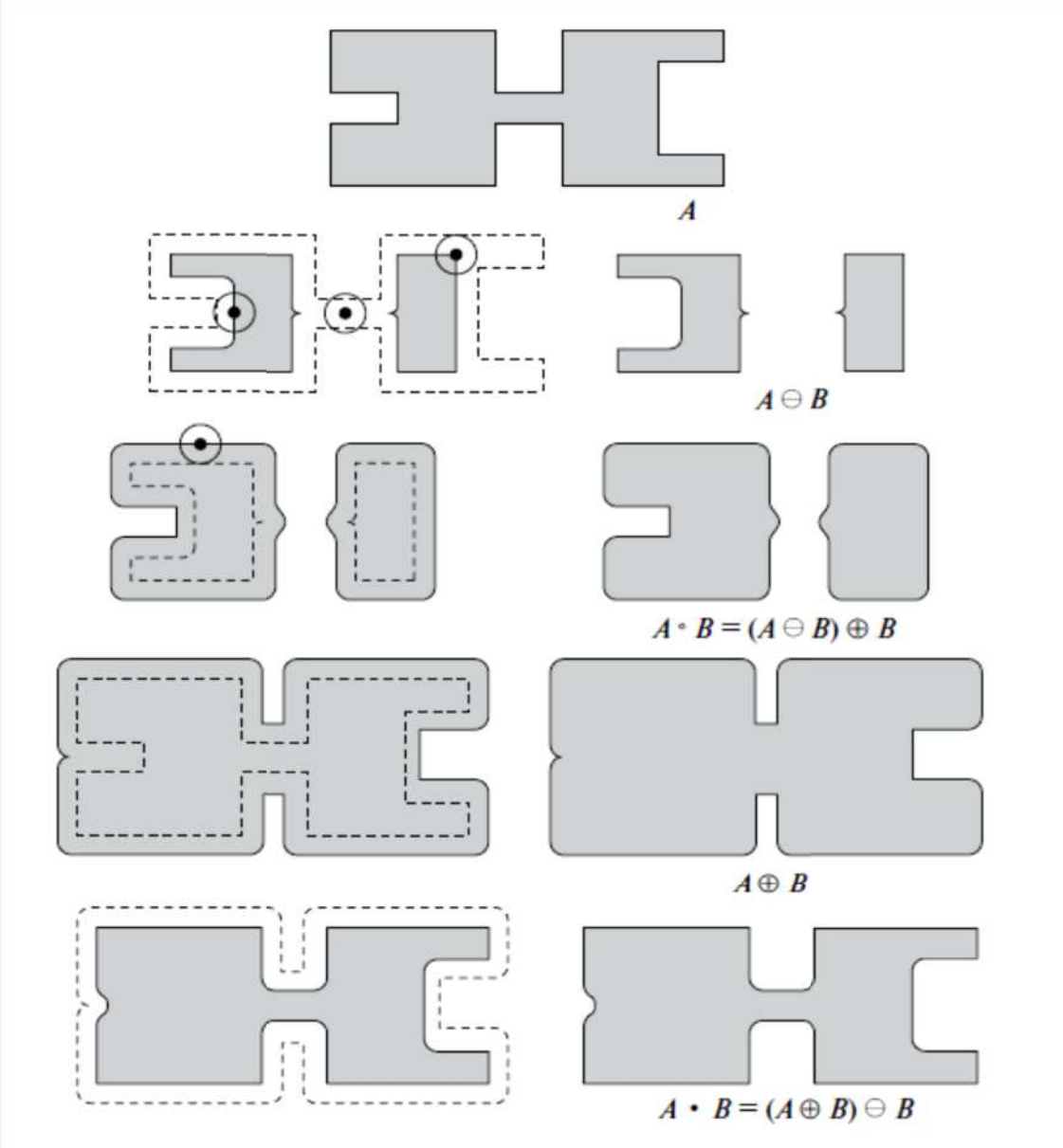
Введенные операции размыкания и замыкания также являются *двойственными*, что проявляется в справедливости следующих свойств:

$$\overline{A \bullet B} = \overline{A} \circ \hat{B},$$
$$\overline{A \circ B} = \overline{A} \bullet \hat{B}.$$

Соответственно для случая симметричного примитива эти соотношения принимают более простой вид:

$$\overline{A \bullet B} = \overline{A} \circ B,$$
$$\overline{A \circ B} = \overline{A} \bullet B.$$

На следующем рисунке приведен пример применения операций размыкания и замыкания к одному и тому же объекту (множеству) $A$ с помощью кругового примитива $B$.



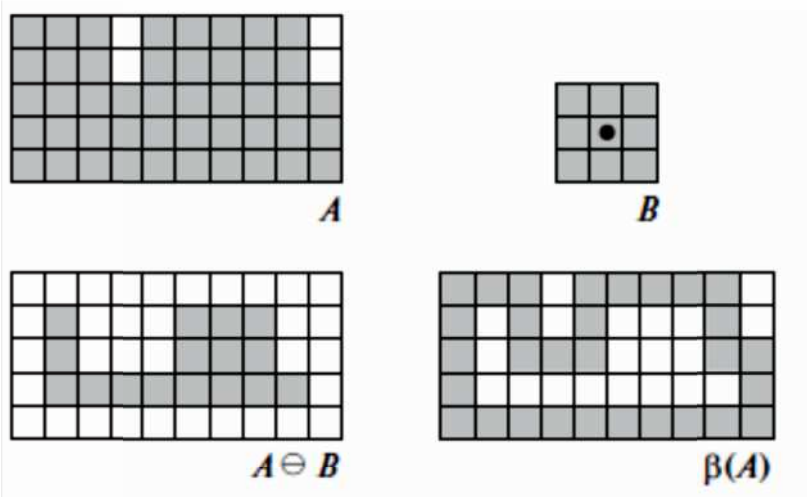
# Глава 13. Основные морфологические алгоритмы

## 13.1. Выделение границ

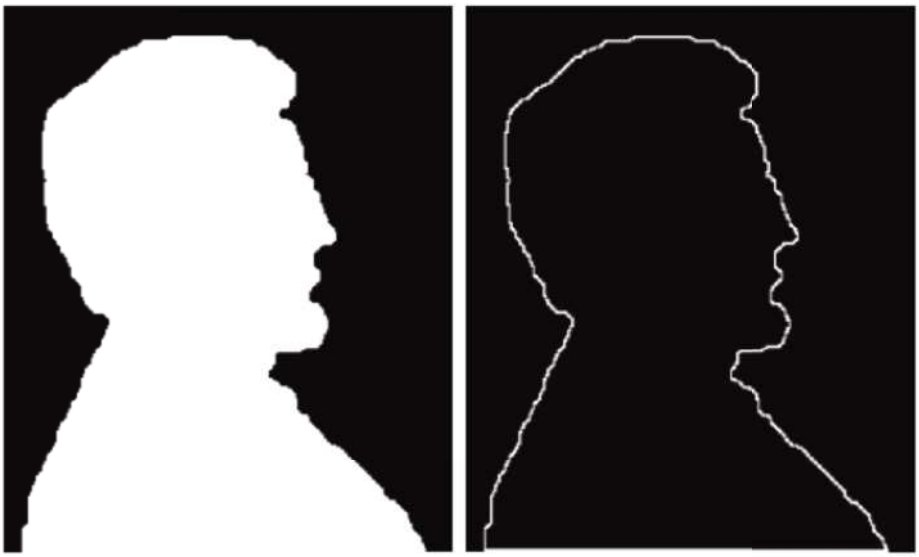
Граница $\gamma(A)$ множества $A$ может быть получена в результате следующей операции:

$$\gamma(A) = A \setminus \left(A \ominus B\right),$$

т.е. вначале к множеству $A$ применяется эрозия с некоторым примитивом $B$, а затем полученный результат теоретико-множественно вычитается из $A$, что иллюстрируется на следующем рисунке.




Следует заметить, что варьируя размер примитива, можно строить границы разной толщины. Например, если примитив имеет размеры 5x5, то граница будет иметь ширину 2. На следующем рисунке показан пример применения операции выделения границ.

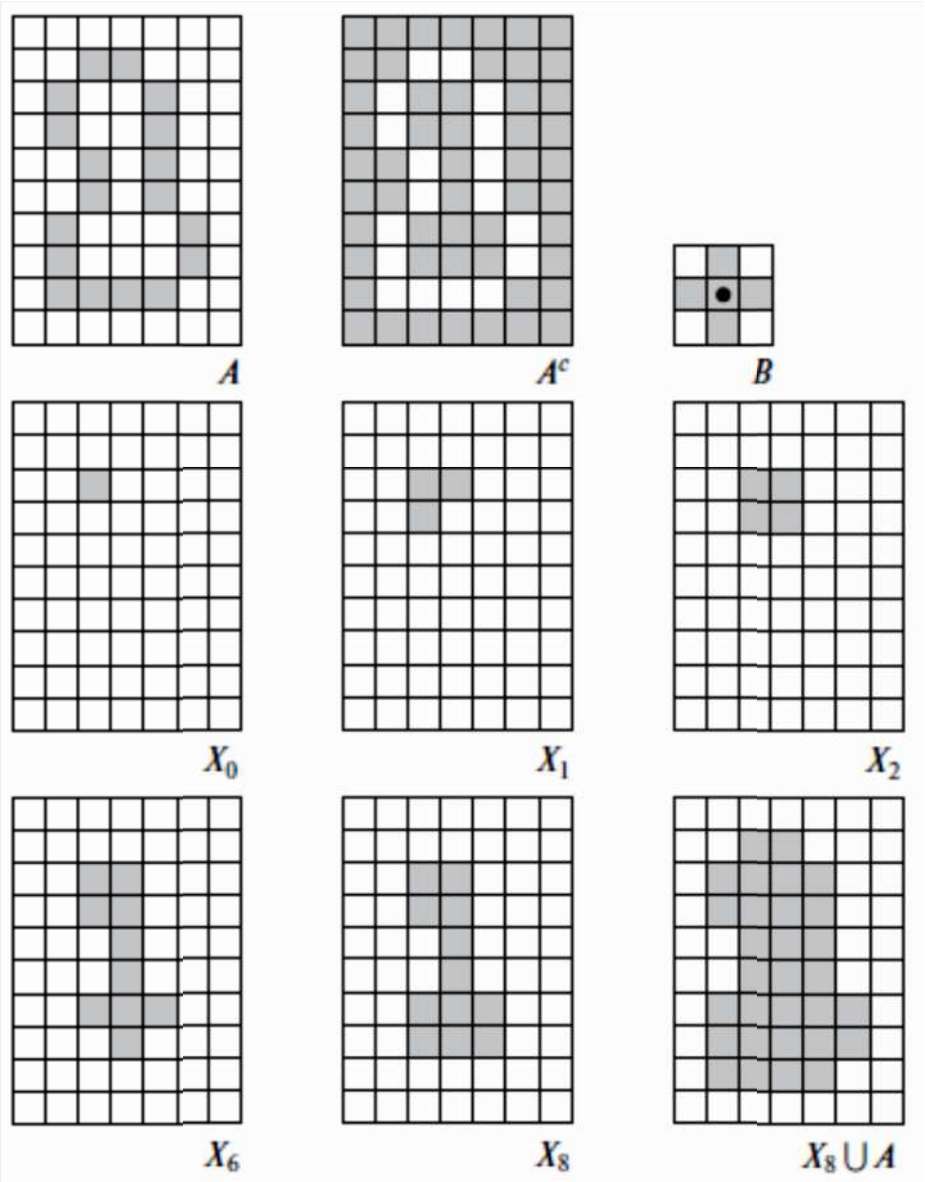


## 13.2. Заполнение дырок

Объекты на изображении могут иметь «дырки», т.е. области на объекте, которые имеют яркость фона. Иногда требуется устранить эти дырки. Такая операция может быть осуществлена также с помощью морфологических операций. Пусть $A$ – объект, внутри которого имеется одна или несколько дырок. Построим вначале массив $X_0$, который имеет такие же размеры, что $A$, но заполнен нулями, кроме одного пикселя внутри каждой дырки, который равен 1. Тогда заполнение дырок реализуется с помощью рекуррентной процедуры:

$$X_k = \left(X_{k-1} \oplus B\right) \cap \overline{A}, \; k = 1, 2, \ldots.$$

где $B$ – примитив, с помощью которого реализуется эта процедура. Этот процесс иллюстрируется следующим рисунком.



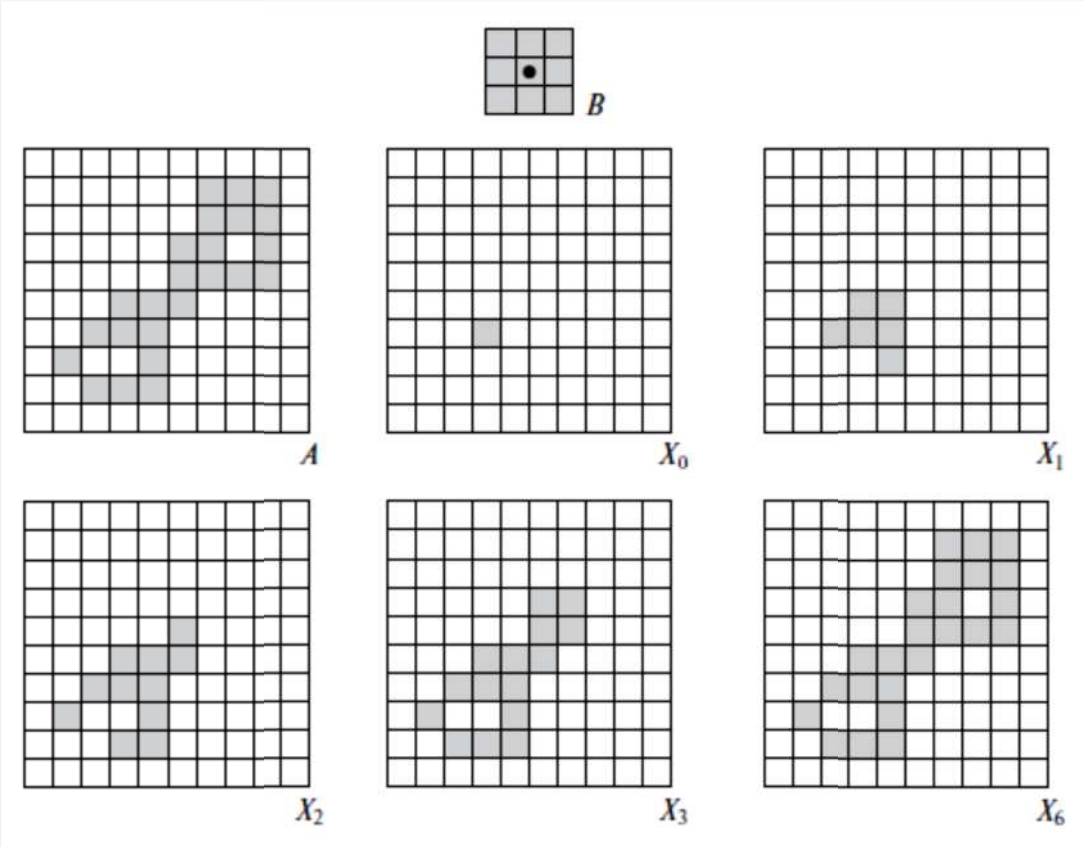
Работа алгоритма останавливается, когда $X_k = X_{k-1}$. Нетрудно понять, что пересечение в формуле дилатации с дополнением множества $A$ ограничивает распространение заполнения дырки на все изображение. Очевидно, что этот процесс может быть одновременно реализован для многих дырок на изображении.

## 13.3. Выделение связных компонент

Для выделения на двоичном изображении связных компонент применяется процедура во многом похожая на процедуру заполнения дырок. Предположим, что $A$ – множество, состоящее из одной или более связных компонент. Построим массив $X_0$ таких же размеров, что и $A$, но со всеми элементами равными 0 за исключением нескольких точек внутри связных компонент, которые равны 1. Результирующий массив связных компонент строится последовательно, начиная с $X_0$, в виде рекуррентной процедуры:

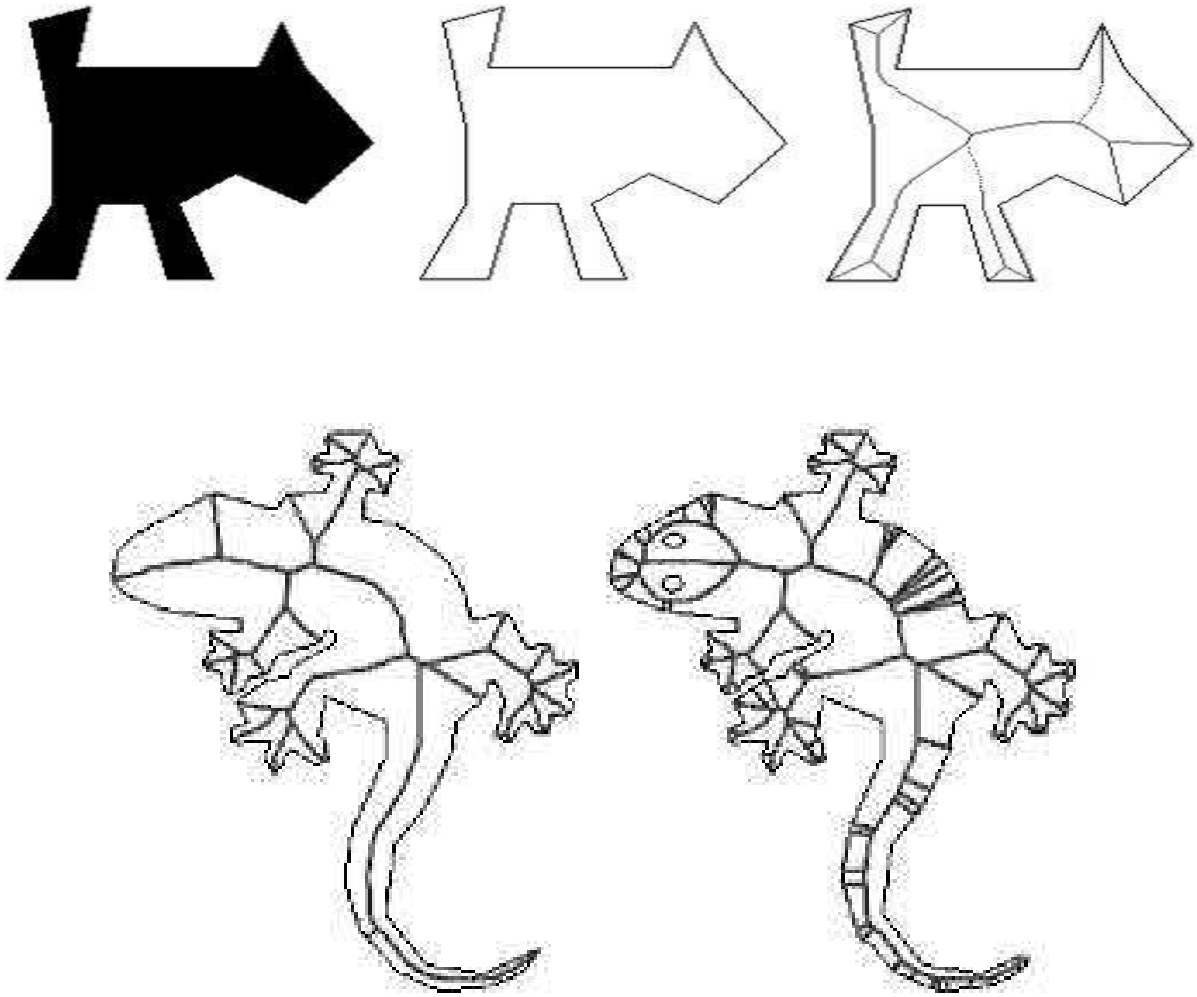
$$X_k = \left(X_{k-1} \oplus B\right) \cap A,\ k = 1, 2, \ldots.$$

где, как и раньше, $B$ – примитив, с помощью которого реализуется эта процедура. Работа алгоритма останавливается, когда выполняется условие $X_k = X_{k-1}$. Работа алгоритма показана на рисунке ниже.



## 13.4. Построение скелета

Для описания фигур на изображении часто используется плоский граф, который называется *скелетом* фигуры. Интуитивное понимание скелета фигуры дают следующие иллюстрации. На третьем рисунке в первом ряду, а также на остальных рисунках, внутренние линии в контуре фигуры и являются ее скелетом.

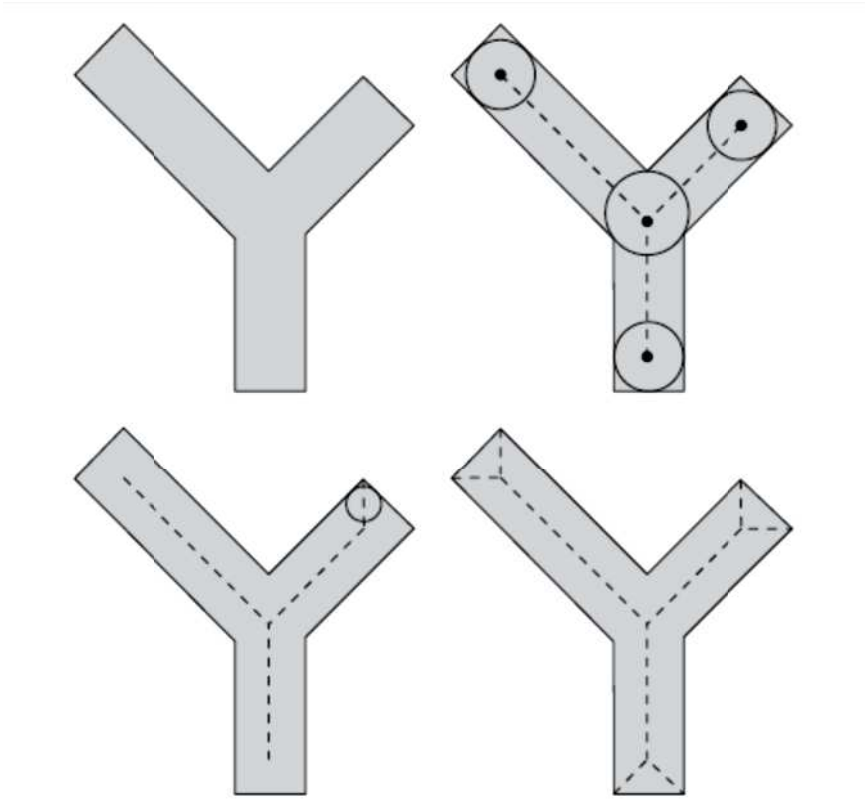


Пусть имеется двоичное изображение, на котором имеется фигура, заданная множеством пикселей $A$, яркости которых равны 1.

| | |
|---|---|
| **Определение** | *Скелетом* (или *остовом*) $S(A)$ множества $A$ называется геометрическое место точек $z$, удовлетворяющих следующим условиям:<br>1) если $D_z$ – наибольший круг с центром в точке $z$, целиком лежащий внутри множества $A$, то не существует круга с |

| | |
|---|---|
| | большим диаметром, также лежащего в $A$ и содержащего внутри себя $D_z$. Такой круг называется *максимальным кругом*;<br>2) круг $D_z$ касается границы множества $A$ в двух или более различных точках. |

Иными словами, скелет фигуры образуется множеством центров всех максимальных кругов, лежащих внутри этой фигуры. Следующий рисунок иллюстрирует данное определение.



Оказывается, множество точек, образующих скелет фигуры, может быть выражено с помощью введенных ранее морфологических операций. Пусть $B$ – примитив (структурный элемент). Обозначим через $A \ominus kB$ результат $k$ последовательных применений к множеству $A$ операции эрозии с примитивом $B$. Например, $A \ominus 3B = \left(\left(\left(A \ominus B\right) \ominus B\right) \ominus B\right)$. Определим теперь множество

$$S_k(A) = \left(A \ominus kB\right) \setminus \left(\left(A \ominus kB\right) \circ B\right),$$

Нетрудно понять, что в результате дилатации $A \ominus kB$ при некотором $k$ результирующее множество окажется пустым. Обозначим через $K$ наибольшее значение $k$, при котором множество $A \ominus kB$ еще не является пустым, т.е.

$$K = \max\{k \mid A \ominus kB \neq \varnothing\}.$$

Тогда скелет фигуры может быть выражен следующим образом:

$$S(A) = \bigcup_{k=1}^{K} S_k(A).$$

Более того, множество $A$ может быть восстановлено по множествам $S_k(A)$, если известен примитив $B$

$$A = \bigcup_{k=1}^{K} \left( S_k(A) \oplus kB \right).$$

В последнем выражении $S_k(A) \oplus kB$ обозначает последовательное применение дилатации к множеству $S_k(A)$ в точности $k$ раз.

Данное построение скелета фигуры не является удобным, поскольку результирующий скелет не всегда получается тонким, и не всегда оказывается связным, как на приведенных выше рисунках. Между тем, существуют достаточно эффективные в вычислительном смысле алгоритмы построения скелета фигуры, которые широко используются в практических задачах.

Заметим в заключение, что рассмотренные выше базовые морфологические операции дилатации, эрозии, размыкания и замыкания обобщаются на полутоновые изображения и используются для построения основных алгоритмов обработки полутоновых изображений.

# РАЗДЕЛ 5.
# СЕГМЕНТАЦИЯ ИЗОБРАЖЕНИЙ

## ГЛАВА 14. СЕГМЕНТАЦИЯ: ОСНОВНЫЕ ПОНЯТИЯ И БАЗОВЫЕ ОПЕРАЦИИ

### 14.1. Основные понятия и определения

Сегментация подразделяет изображение на составляющие его области или объекты. Детализации, до которой доводится такое разделение, зависит от решаемой задачи. Это следует понимать так, что сегментацию следует прекратить, когда интересующие объекты или области обнаружены.

Сегментация представляет собой одну из наиболее важных и в то же время сложных задач обработки изображений. Результат автоматической сегментации изображений во многом определяется точностью сегментации.

Большинство алгоритмов сегментации изображений основываются на одном из двух базовых свойств сигнала яркости: *разрывности* и *однородности*. В первом случае подход состоит в разбиении изображения на основании резких изменений сигнала, таких как перепады яркости на изображении. Вторая категория методов использует разбиение изображения на области, однородные в смысле заранее выбранных критериев. Примерами таких методов могут служить пороговая обработка, выращивание областей, слияние и разбиение областей.

Рассмотрим и проиллюстрируем некоторые из таких подходов и покажем, что улучшения качества сегментации можно достичь комбинацией методов из разных категорий.

| | |
|---|---|
| **Определение** | Под *сегментацией* понимается процесс разбиения всей области изображения на *связные* подобласти $R_i$, $i = 1, 2, ..., n$: $$\bigcup_{i=1}^{n} R_i, \quad R_i \cap R_j = \varnothing \text{ для любых } i \neq j,$$ причем все пиксели из одной и той же области удовлетворяют некоторому условию, а пиксели из разных областей этому условию не удовлетворяют. |

Алгоритмы сегментации монохромных изображений обычно относятся к одной из двух основных категорий, исходя из свойств значений яркости – *наличия разрывов* и *близости значений*. В первой категории предполагается, что края областей достаточно сильно отличаются как от фона изображения, так и друг от друга, что позволяет обнаруживать границу на основании локальных разрывов яркости. Основной подход в первой категории – *сегментация на основе контуров*. Ко второй категории относятся методы *сегментации на основе областей*, которые осуществляют разбиение изображения на области, обладающие внутренним сходством в соответствии с набором наперед заданных критериев.

## 14.2. Обнаружение точек, линий и перепадов яркостей

Большая группа методов сегментации основана на детекции участков с резким изменением яркости изображения. При этом, как правило, выделяют три группы типичных для этих методов объектов – изолированные точки, линии и так называемые перепады.

Локальные изменения яркости легко обнаруживаются с помощью первой и второй производных, о чем говорилось выше. При этом на изображениях можно говорить только об аппроксимациях производных. Аппроксимацию производных можно производить многими способами, но от любой аппроксимации требуется выполнение следующих условий:

аппроксимация первой производной должна быть:

1) равна нулю в областях с постоянной яркостью;

2) ненулевой в начале или конце ступенчатого или линейного изменения яркости;

3) ненулевой на участке линейного изменения яркости;

аппроксимация второй производной должна быть:

1) равна нулю в областях с постоянной яркостью;

2) ненулевой в начале или конце ступенчатого или линейного изменения яркости;

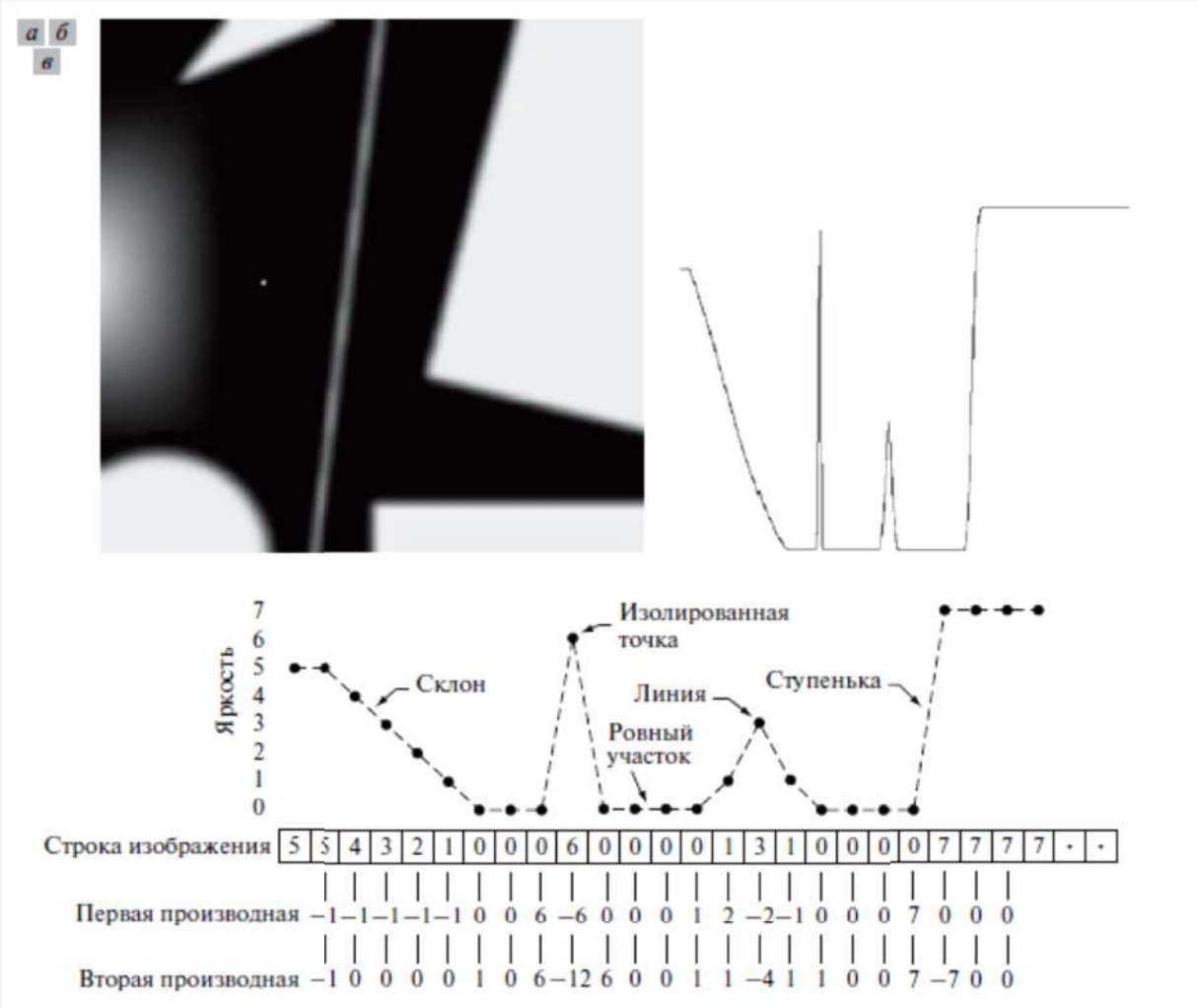
3) равна нулю на участке линейного изменения яркости.

Указанным условиям, разумеется, удовлетворяют рассмотренные выше аппроксимации

$$\frac{df}{dx} \approx f(x+1) - f(x),$$

$$\frac{d^2 f}{d^2 x} \approx f(x+1) - 2f(x) + f(x-1).$$

На следующем рисунке показано изображение и его горизонтальный профиль, проходящий через изолированную яркую точку. Ниже показаны отдельные характерные участки этого профиля и значения первой и второй производных в каждой точке.



Следует особо отметить, что вторая производная значительно «активнее» реагирует на мелкие детали (точки, тонкие линии), чем первая производная. Как следствие, вторые производные более чувствительны и к шумовым составляющим.

С алгоритмической точки зрения вычисление аппроксимаций первой и второй производных сводится к использованию пространственных фильтров вида

| | | |
|---|---|---|
| $w_1$ | $w_2$ | $w_3$ |
| $w_4$ | $w_5$ | $w_6$ |
| $w_7$ | $w_8$ | $w_9$ |

где коэффициенты определяются используемой аппроксимацией, т.е.

$$w_1 z_1 + w_2 z_2 + ... + w_9 z_9 = \sum_{k=1}^{9} w_k z_k ,$$

где $z_k$ – значение яркости пикселя, соответствующего коэффициенту $w_k$.

**Изолированные точки.** Изолированные точки наилучшим образом находятся с помощью лапласиана

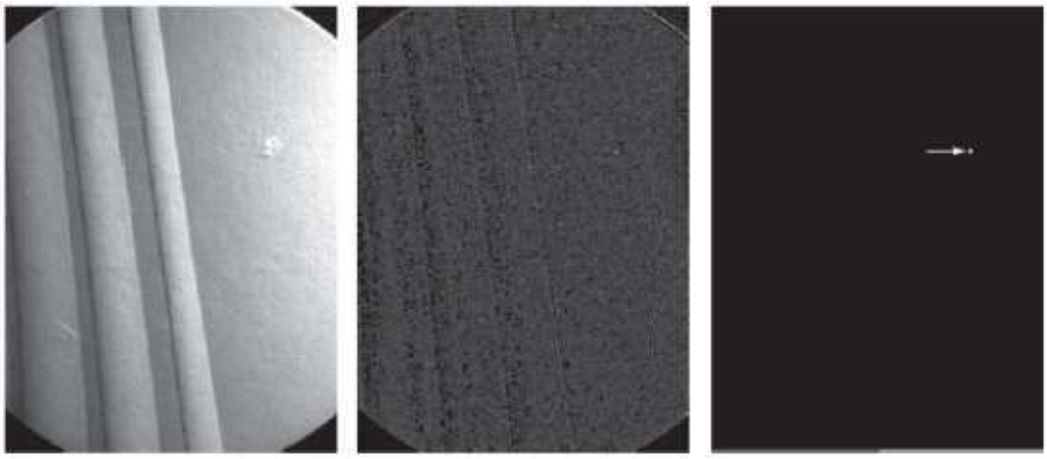
$$\nabla^2 f(x,y) = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} \approx f(x+1,y) + f(x-1,y) + f(x,y+1) + f(x,y-1) - 4f(x,y).$$

Указанная детекция реализуется с помощью маски

| | | |
|---|---|---|
| 0 | 1 | 0 |
| 1 | -4 | 1 |
| 0 | 1 | 0 |

При этом изолированная точка обнаруживается, если абсолютная величина отклика превышает некоторое пороговое значение. Если таким точкам присвоить значение 1, а всем остальным – 0, то получится двоичное изображение, на котором будут видны только изолированные точки. Следует также отметить, что сумма коэффициентов маски равна нулю, это гарантирует то,

что на областях постоянной яркости отклик фильтра будет равен нулю. На следующем рисунке приведен пример обнаружения изолированной черной точки.



**Линии.** Для обнаружения линий можно применять также маску лапласиана. Однако лапласиан может принимать отрицательные значения, поэтому требуется градационное преобразование, например, сдвиг нулевого значения яркости, так что он будет соответствовать серому значению яркости. В этом случае происходит эффект удвоения линии.

Следует также отметить, что используемый лапласиан обладает свойством изотропности, т.е. он инвариантен относительно направления. В некоторых случаях может требоваться обнаружение линий определенного направления. В таком случае применяются маски, изображенные на следующем рисунке:

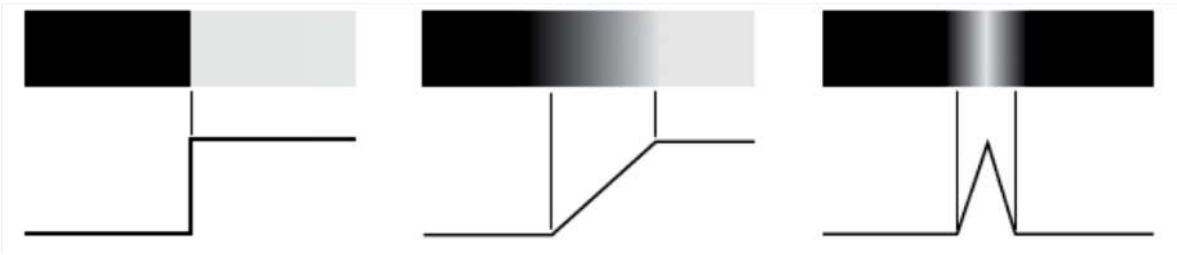
| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| −1 | −1 | −1 | 2 | −1 | −1 | −1 | 2 | −1 | −1 | −1 | 2 |
| 2 | 2 | 2 | −1 | 2 | −1 | −1 | 2 | −1 | −1 | 2 | −1 |
| −1 | −1 | −1 | −1 | −1 | 2 | −1 | 2 | −1 | 2 | −1 | −1 |
| Горизонтальная | | | +45° | | | Вертикальная | | | −45° | | |

Анизотропность приведенных масок позволяет выделять горизонтальные и вертикальные линии, а также линии, расположенные под углами $+45^\circ$ и $-45^\circ$ .

**Перепады яркостей.** Обнаружение перепадов яркости является наиболее часто встречающейся задачей при сегментации изображений. Форма области соединяющей разные яркости на изображении может различной, но чаще всего такие перепады представляют некоторой идеализированной моделью одного из трех типов.

Первый тип – это *ступенчатый перепад*. Это идеальное представление перепада, которое редко встречается в реальности. Однако в случае искусственно синтезированных изображений, а также изображений очень высокого качества эта модель оказывается приемлемой и удобной.

Второй тип – *наклонный перепад*. Он отражает присутствующее на изображениях размытие, вызванное оптикой, а также электронными шумами. Крутизна наклонного участка обратно пропорциональна степени размытия границы разных яркостей. В связи с этой моделью возникает необходимость определиться, что считать истинной точкой перепада яркостей. Обычно эта задача решается с помощью введения некоторого критерия качества границы, разделяющей яркости областей в целом.



Третий тип – *треугольный импульс*. Такая модель соответствует некоторой линии, разделяющей области с одинаковой яркостью. Ширина основания треугольного импульса определяется параметрами разделяющей линии. Если ширина равна одному пикселю, то это просто тонкая линия, проходящая через область с одинаковой яркостью. Примером таких линий с формой треугольного импульса могут быть дороги, реки на аэрофотоснимках, провода, далеко находящиеся столбы и пр.

Для детекции на изображении указанных перепадов яркости также могут применяться маски, задающие аппроксимации первых и вторых производных. Поведение производных на границе перепада яркости иллюстрируется на нижеследующем рисунке.

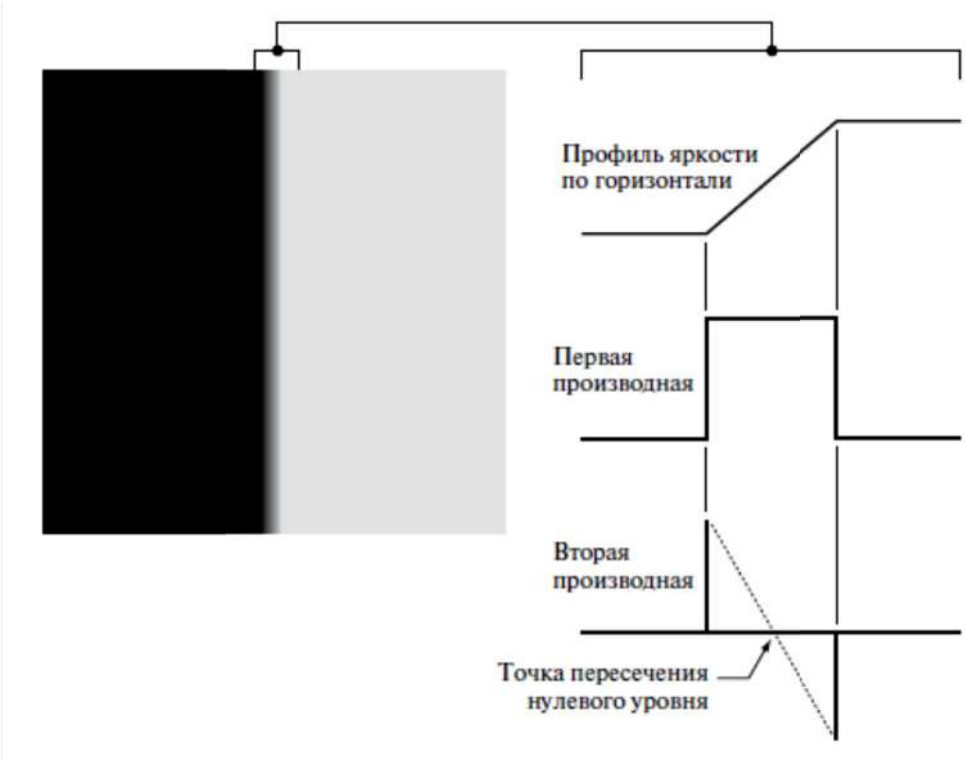
Важно помнить, что использование масок на основе дифференциальных фильтров оказывается весьма чувствительным к шумам на изображении.

По этой причине, а также по причинам указанным выше, при определении перепадов яркости, как правило, выполняются следующие три шага.

1) *Сглаживание изображения с целью уменьшения шума*. Необходимость вызвана отмеченной чувствительностью дифференциальных фильтров к шумам.

2) *Обнаружение точек перепада*. Это операция, которая выделяет на изображении все точки, которые могут оказаться точками реального перепада яркостей.

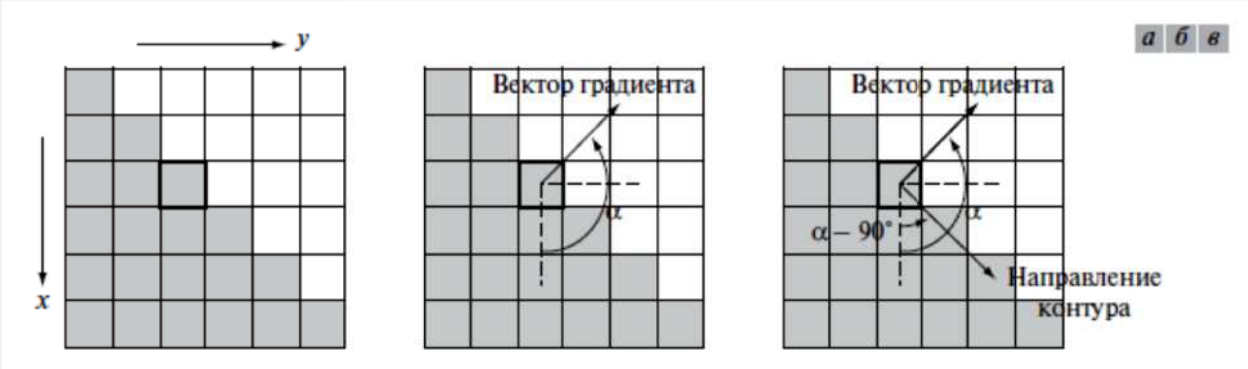
3) *Локализация перепада*. В результате этого шага среди выделенных точек, полученных на предыдущей шаге, определить те точки, которые действительно входят в множество точек, вместе составляющих контурный перепад.




## 14.3. Простые методы обнаружения контурных перепадов

С помощью градиента изображения $\nabla f = \left( \frac{\partial f}{\partial x} \ \frac{\partial f}{\partial y} \right)$ можно в каждой точке изображения найти одновременно величину перепада яркости и его направление. Рассмотрим пример.

Каждый квадрат представляет собой один пиксель. Темные пиксели имеют значение 0, а светлые – 1. Вычисление градиента будем осуществлять с помощью маски 3x3, центр которой находится в пикселе, в котором мы хотим найти градиент.




Из суммы значений пикселей в нижней строчке окрестности 3x3 вычтем сумму значений пикселей в верхней строчке, в результате получим оценку значения производной по $x$. Аналогично находим приближенное значение производной по $y$. Для приведенного примера получаем конкретно, что $\frac{\partial f}{\partial x} = -2$, а $\frac{\partial f}{\partial y} = 2$. Значит, модуль градиента равен $|\nabla f| = 2\sqrt{2}$. При этом направление вектора градиента равно $\alpha = \mathrm{arctg}\left(\frac{\partial f}{\partial y} \Big/ \frac{\partial f}{\partial x}\right) = \frac{3\pi}{4}$ относительно положительного направления оси $X$. Поскольку, как уже отмечалось, направление перепада яркости перпендикулярно направлению градиента, то направление линии перепада в этом примере равно $\frac{\pi}{4}$. По этой причине градиент называют также нормалью к контурному перепаду яркостей.

Вычисления градиента производится с помощью очевидных простых масок. Для вычисления перепадов, идущих в диагональных направлениях, часто применяю *перекрестный градиентный оператор Робертса*:

$$\frac{\partial f}{\partial x} = z_9 - z_5, \qquad \frac{\partial f}{\partial y} = z_8 - z_6,$$

где значения пикселей обозначены так, как это делалось и раньше (следующий рисунок).

Однако чаще применяются маски, симметричные относительно центрального элемента:

$$\frac{\partial f}{\partial x} = (z_7 + z_8 + z_9) - (z_1 + z_2 + z_3), \qquad \frac{\partial f}{\partial y} = (z_3 + z_6 + z_9) - (z_1 + z_4 + z_7).$$

Оператор, соответствующий данным маскам, показанным на следующем рисунке, называется *оператором Превитта*.

| | | |
|---|---|---|
| $z_1$ | $z_2$ | $z_3$ |
| $z_4$ | $z_5$ | $z_6$ |
| $z_7$ | $z_8$ | $z_9$ |

| | | | | | |
|---|---|---|---|---|---|
| −1 | 0 | | 0 | −1 | |
| 0 | 1 | | 1 | 0 | |

Маски оператора Робертса

| | | | | | |
|---|---|---|---|---|---|
| −1 | −1 | −1 | −1 | 0 | 1 |
| 0 | 0 | 0 | −1 | 0 | 1 |
| 1 | 1 | 1 | −1 | 0 | 1 |

Маски оператора Превитта

Модификацией оператора Превитта является *оператор Собела*, в котором больший вес придается средним точкам. Это несколько уменьшает эффект сглаживания, присущий оператору Превитта.

| | | | | | |
|---|---|---|---|---|---|
| −1 | −2 | −1 | −1 | 0 | 1 |
| 0 | 0 | 0 | −2 | 0 | 2 |
| 1 | 2 | 1 | −1 | 0 | 1 |

Маски оператора Собела

Следует подчеркнуть, что во всех масках сумма коэффициентов равна нулю. Это значит, что на областях равной яркости они будут давать нулевое значение.

Для обнаружения диагональных участков контуров перепада яркости оказываются более удобными модификации операторов Превитта и Собела, которые показаны на рисунке ниже.

| | | |
|---|---|---|
| 0 | 1 | 1 |
| −1 | 0 | 1 |
| −1 | −1 | 0 |

| | | |
|---|---|---|
| −1 | −1 | 0 |
| −1 | 0 | 1 |
| 0 | 1 | 1 |

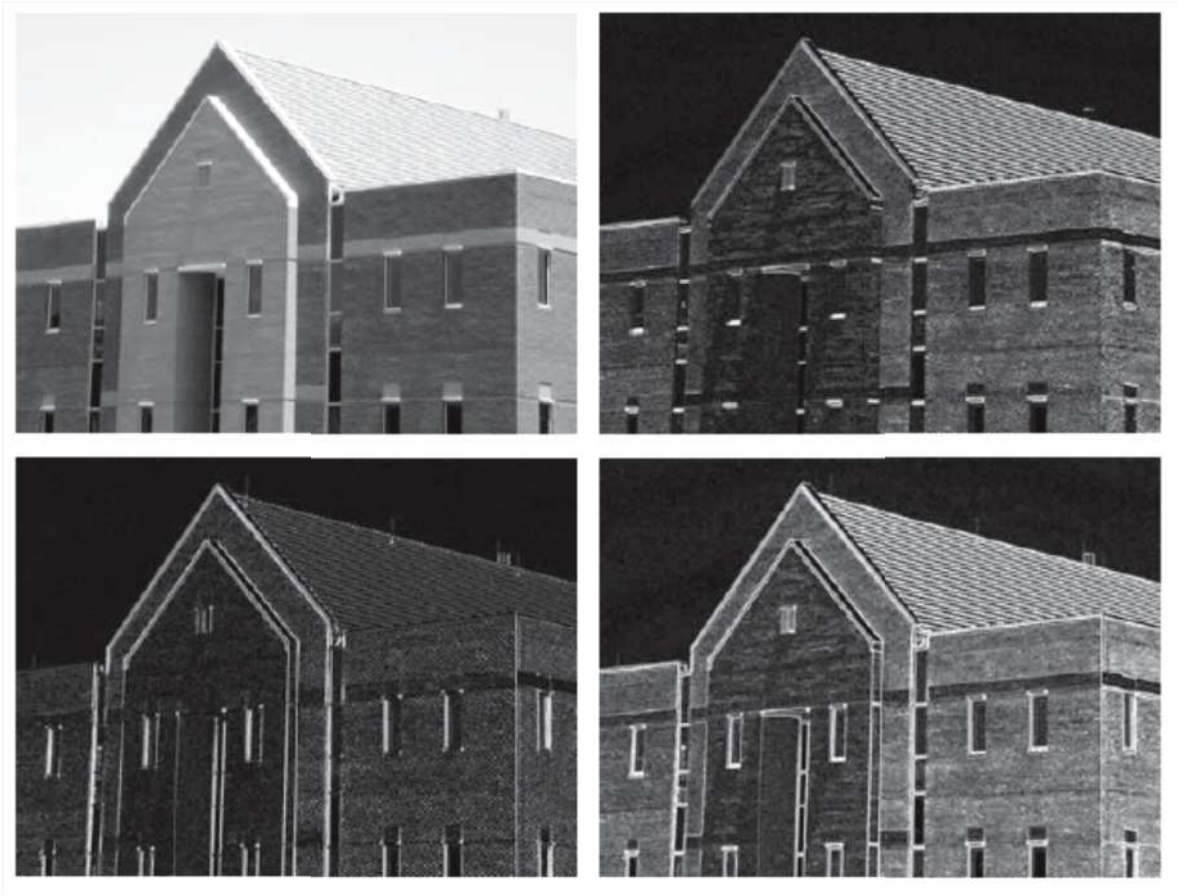
Маски оператора Превитта

| | | |
|---|---|---|
| 0 | 1 | 2 |
| −1 | 0 | 1 |
| −2 | −1 | 0 |

| | | |
|---|---|---|
| −2 | −1 | 0 |
| −1 | 0 | 1 |
| 0 | 1 | 2 |

Маски оператора Собела

На следующем рисунке приведен пример выделения перепадов яркости с помощью оператора Собела. Левый верхний рисунок является исходным. Вверху справа показан рисунок производной вдоль оси $X$, а внизу слева – производной вдоль оси $Y$. Внизу справа показан рисунок, на котором задается значение $\left|\frac{\partial f}{\partial y}\right| + \left|\frac{\partial f}{\partial x}\right|$, которое является достаточно грубым приближенным значением модуля градиента.



На приведенных рисунках нетрудно заметить особенности откликов каждой из составляющих градиента.

# Глава 15. Обнаружение линий и контуров

## 15.1. Детектор контуров Кэнни

Идея алгоритма Кэнни состоит в следующем. Изображение вначале сглаживается с помощью круговой двумерной гауссовой функции. Затем вычисляется градиент, его модуль и направление градиента. С их помощью определяется величина и направление перепада в каждой точке изображения.

Пусть $f(x,y)$ – исходное изображение, а $G(x,y)$ – гауссова функция

$$G(x,y)=e^{-\frac{x^2+y^2}{2\sigma^2}}.$$

Тогда сглаженное изображение $f_s(x,y)$ можно получить с помощью свертки $f(x,y)$ с $G(x,y)$:

$$f_s(x,y)=G(x,y)*f(x,y).$$

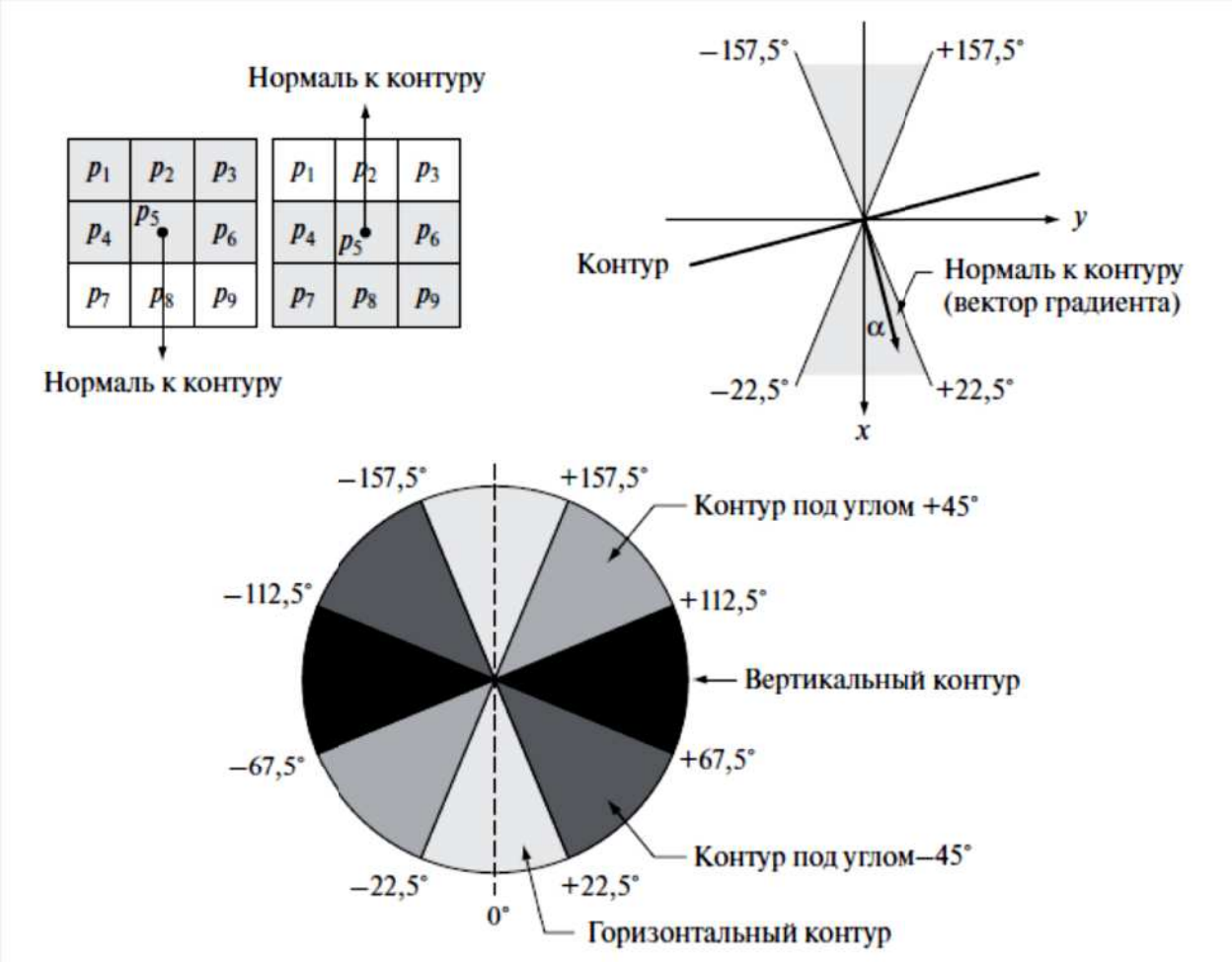
Соответственно, находятся модуль $M(x,y)=\sqrt{\left(\frac{\partial f_s}{\partial x}\right)^2+\left(\frac{\partial f_s}{\partial y}\right)^2}$ и направление градиента $\alpha(x,y)=\mathrm{arctg}\left(\frac{\partial f_s}{\partial y}\Big/\frac{\partial f_s}{\partial x}\right)$. Производные вычисляются с помощью любой из масок, рассмотренных ранее. $M(x,y)$ и $\alpha(x,y)$ являются массивами такого же размера, что и $f(x,y)$.

Свертка $f_s(x,y)=G(x,y)*f(x,y)$ вычисляется с помощью маски размера $n\times n$.

Размеры гауссова фильтра выбираются из следующих соображений. В теории вероятностей хорошо известно «правило трех сигм», в соответствии с которым приблизительно 99,7 % объема под двумерной гауссовой поверхностью лежит в окрестности $\pm 3\sigma$ с центром в точке математического ожидания. Поэтому при выборе размеров $n\times n$ дискретного гауссова фильтра можно взять в качестве $n$ наименьшее нечетное целое число, равное или большее $6\sigma$. Использование меньшего $n$ приведет к «усечению» функции, причем степень такого усечения обратно пропорциональна размеру маски. Использование маски с большими размерами не оказывает заметного влияния на результат.

Поскольку изображение $M(x,y)$ формируется как модуль градиента, оно обычно содержит широкие гребни вдоль линии локальных максимумов.

Следующий шаг состоит в утончении таких гребней. Один из применяемых методов использует *подавление немаксимальных точек*, которое может выполняться несколькими способами, но сущность метода в том, что выбирается несколько дискретных направлений нормали к контуру (т.е. направлений вектора градиента). Например, в области размерами $3 \times 3$ пикселя для контура, проходящего через центральный элемент области, можно определить четыре направления: горизонтальное, вертикальное, +45° и –45°.




На предыдущем рисунке слева вверху показаны две возможные ориентации горизонтального контура. Поскольку требуется привести все возможные направления контура к четырем фиксированным значениям, необходимо указать диапазон для каждого из выбранных направлений, например в каком из случаев контур будет считаться горизонтальным. Направление контура в точке определяется по направлению нормали к нему, которое определяется непосредственно из изображения, полученного с помощью соответствующей формулы. Как видно на рисунке справа, если угол нормали к контуру находится в интервале от –22,5° до +22,5° или от –157,5° до +157,5°, такой контур

считается горизонтальным. На нижнем рисунке показаны диапазоны углов, соответствующих всем четырем рассматриваемым направлениям контуров.

Обозначим через $d_1$, $d_2$, $d_3$ и $d_4$ четыре только что определенных базовых направления контура в окрестности $3\times3$: соответственно горизонтальное, –45°, вертикальное и +45°. Можно сформулировать следующую схему подавления немаксимальных точек для окрестности $3\times3$ с центром в *каждой* точке $(x,y)$ изображения $\alpha(x,y)$ и формирования изображения $g_N(x,y)$ с утонченными контурами:

1. Находим направление $d_k$, ближайшее к $\alpha(x,y)$.
2. Если значение $M(x,y)$ меньше, чем хотя бы у одного из соседей в направлении $d_k$, принимаем $g_N(x,y)=0$ (подавление); в противном случае $g_N(x,y)=M(x,y)$.

Изображение $g_N(x,y)$ содержит только утонченные контуры; оно отличается от $M(x,y)$ тем, что все немаксимальные точки контуров подавлены.

Заключительная операция состоит в пороговой обработке изображения $g_N(x,y)$ с целью уменьшения числа ложных контурных точек. Обычно используется пороговое преобразование с одним порогом, при котором устанавливаются в 0 значения во всех точках, где они меньше этого порога. Если установлен слишком низкий порог, то останутся некоторые ложные контуры. Если установить слишком высокий порог, то исчезнут истинные контурные точки. В алгоритме Кэнни делается попытка улучшить такое положение с помощью порогового преобразования с *гистерезисом*, в котором используются два порога – нижний $T_L$ и верхний $T_H$.

Результаты порогового преобразования можно визуализировать, сформировав еще два изображения

$$g_{NH}(x,y)=g_N(x,y)\geq T_H,$$

$$g_{NL}(x,y)=g_N(x,y)\geq T_L,$$

где первоначально все элементы $g_{NH}(x,y)$ и $g_{NL}(x,y)$ имеют нулевые значения. После такого порогового преобразования в изображении $g_{NH}(x,y)$ будет в общем случае меньше ненулевых пикселей, чем в $g_{NL}(x,y)$, но все ненулевые пиксели $g_{NH}(x,y)$ повторятся и в $g_{NL}(x,y)$, так как последнее изображе-

ние сформировано с меньшим значением порога. Удалим на изображении $g_{NL}(x,y)$ все пиксели с ненулевыми значениями в $g_{NH}(x,y)$, полагая

$$g_{NL}(x,y) := g_{NL}(x,y) - g_{NH}(x,y).$$

Теперь ненулевые пиксели изображений $g_{NH}(x,y)$ и $g_{NL}(x,y)$ можно рассматривать как соответственно «сильные» и «слабые» контурные точки.

После описанных пороговых преобразований все сильные пиксели в $g_{NH}(x,y)$ можно считать точками истинных контуров и сразу их пометить соответствующим образом. В зависимости от величины порога $T_H$ контуры на изображении $g_{NH}(x,y)$ обычно будут прерывистыми. Более длинные контуры формируются по следующему алгоритму:

**(а)** Найти очередной необработанный контурный пиксель $p$ на изображении $g_{NH}(x, y)$.

**(б)** Пометить как истинные контурные пиксели все слабые пиксели в изображении $g_{NH}(x,y)$, которые являются связными с $p$, скажем, в смысле 8-связности.

**(в)** Если обработаны все ненулевые пиксели $g_{NH}(x,y)$, перейти к шагу (г), в противном случае перейти к шагу (а).

**(г)** Установить в 0 значения всех пикселей $g_{NL}(x,y)$, которые не были помечены как истинные.

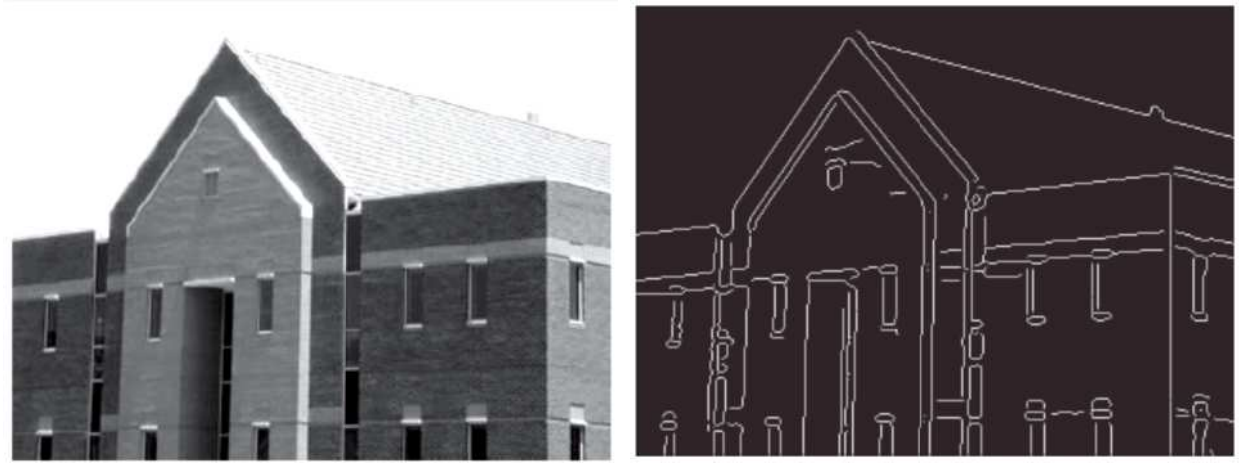
По окончании изложенной процедуры окончательным результатом алгоритма Кэнни будет изображение, получаемое добавлением к $g_{NH}(x,y)$ всех оставшихся ненулевых пикселей $g_{NL}(x,y)$.

Таким образом, алгоритм выделения контуров Кэнни состоит из следующих основных шагов:

1. Сгладить исходное изображение гауссовым фильтром.
2. Сформировать изображения модуля и направления градиента.
3. Применить подавление немаксимальных точек к изображению модуля градиента.
4. Выполнить преобразование с двойным порогом и анализ связности для обнаружения и связывания контуров.

На следующем рисунке приведен пример сегментации, выполненной с помощью алгоритма Кэнни. Как показал опыт практического использования,

алгоритм Кэнни обеспечивает наилучшие показатели сегментации по толщине линий, непрерывности и прямолинейности. По этой причине этот алгоритм стал наиболее часто применяемым методом выделения контуров.



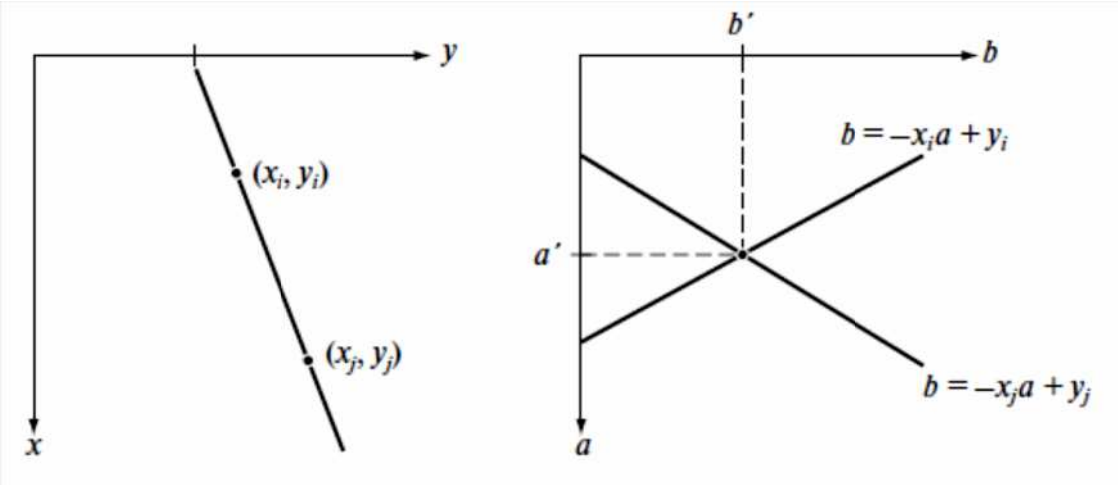
## 15.2. Преобразование Хафа

Суть алгоритма Хафа состоит в выяснении, лежит ли множество пикселей на некоторой кривой, форма которой заранее задана. Подобная задача очень часто возникает в задачах анализа и распознавания объектов на изображении. Прямолинейное решение оказывается весьма трудоемким, особенно при большом количестве анализируемых точек. Хаф (Hough) предложил более простое решение, которое очень широко используется в практических задачах и носит название *преобразование Хафа*.

Пусть на плоскости изображения имеется $n$ точек. Выберем одну из этих точек, например, точку с координатами $(x_i, y_i)$. Рассмотрим уравнение прямой на плоскости $XOY$ в форме с угловым коэффициентом $y = ax + b$. Очевидно, что через точку $(x_i, y_i)$ проходит бесконечно много прямых, удовлетворяющих уравнению $y_i = ax_i + b$ при различных значениях $a$ и $b$.

Перепишем это уравнение в виде $b = -x_i a + y_i$ и рассмотрим плоскость $AB$, которую будем называть *пространством параметров*. Тогда для заданной пары $(x_i, y_i)$ получаем уравнение *единственной* прямой. Другой точке $(x_j, y_j)$ также соответствует своя прямая в пространстве параметров, и если они не параллельны друг другу, эти две прямые пересекаются в некоторой точке $(a', b')$, такой, что $a'$ есть угловой коэффициент, а $b'$ – точка пересече-

ния с осью $Y$ прямой, проходящей через точки $(x_i, y_i)$ и $(x_j, y_j)$ в плоскости $XOY$. На самом деле каждой точке прямой, проходящей через точки $(x_i, y_i)$ и $(x_j, y_j)$, в пространстве параметров соответствует своя прямая линия, причем все они пересекаются в точке $(a', b')$, что показано на следующем рисунке.



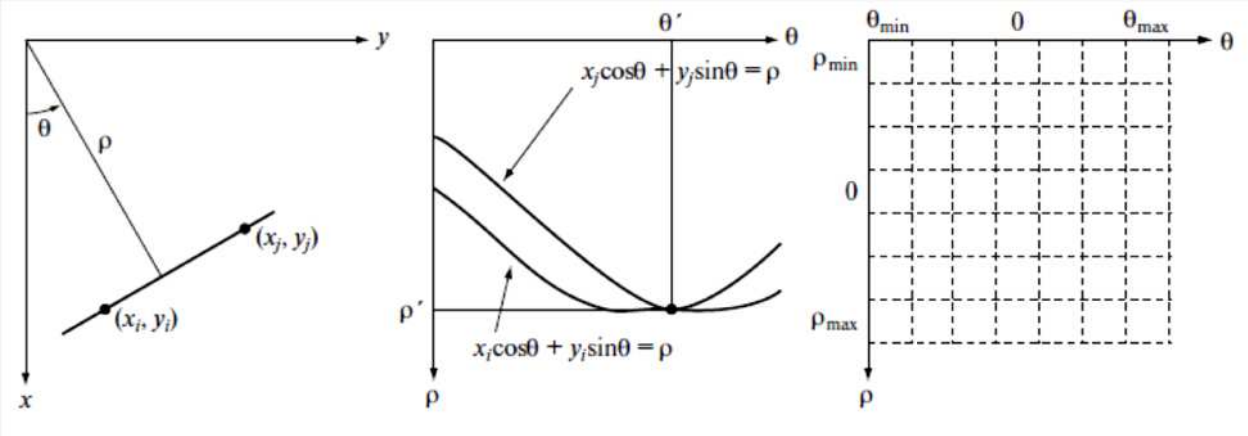

Можно изобразить в пространстве параметров прямые, соответствующие всем точкам $(x_k, y_k)$ заданного множества, и основные линии на исходной плоскости можно найти по тем точкам в пространстве параметров, где пересекается большое число линий. Однако с использованием на практике этого подхода возникают затруднения, поскольку если прямая близка к вертикали, то ее угловой коэффициент $a$ стремится к бесконечности. Обойти эту трудность можно с помощью иного способа задания прямой линии, а именно, с помощью *нормального* уравнения прямой:

$$x\cos\theta + y\sin\theta = \rho .$$

Геометрическая интерпретация параметров ρ и θ показана на следующем рисунке слева. Каждая синусоида на центральном рисунке представляет семейство прямых, проходящих через конкретную точку $(x_k, y_k)$ на плоскости $XOY$. Точка пересечения $(\rho', \theta')$ на соответствует прямой, проходящей через точки $(x_i, y_i)$ и $(x_j, y_j)$ на левом рисунке.

Алгоритмическая часть преобразования Хафа с точки зрения вычислений основана на так называемой идее *голосования*. Суть ее состоит в следующем. Вся область пространства параметров $\rho\theta$ разбивается на так называемые *ячейки накопления*, как показано на рисунке справа, где $(\rho_{\min}, \rho_{max})$ и

$(\theta_{\min}, \theta_{max})$ – задаваемые диапазоны возможных значений параметров: $-90^\circ \le \theta \le 90^\circ$ и $-D \le \rho \le D$, где $D$ – расстояние между углами изображения по диагонали. В ячейке с координатами $(p,q)$ накапливается значение $A(p,q)$ для прямоугольника в пространстве параметров, соответствующего точке $(\rho_p, \theta_q)$.




Первоначально значения во всех ячейках накопления равны нулю. Затем для каждой точки $(x_k, y_k)$ из заданного множества точек в плоскости $XOY$ полагаем параметр $\theta$ равным поочередно каждому разрешенному дискретному значению на оси $\theta$ и, подставляя это значение в уравнение $\rho = x_k \cos\theta + y_k \sin\theta$, находим соответствующее ему значение $\rho$. После этого найденное значение $\rho$ округляется до ближайшего разрешенного дискретного значения на оси $\rho$. Если выбор $\theta_q$ приводит к значению $\rho_p$, увеличиваем накопленное значение в соответствующей ячейке: $A(p,q) := A(p,q) + 1$. После выполнения описанной процедуры для всех анализируемых точек $(x_i, y_i)$ записанное в ячейке $(p,q)$ значение $A(p,q) = P$ означает, что в плоскости $XOY$ имеется $P$ точек, лежащих на прямой $\rho_p = x\cos\theta_q + y\sin\theta_q$. Точность попадания точек на эту прямую определяется размерами ячеек накопления на плоскости $\rho\theta$.

Можно показать, что вычислительная сложность изложенного метода линейна по $n$, т.е. по числу точек в заданном множестве на плоскости $XOY$.

На следующем рисунке приведен простой пример, иллюстрирующий основную идею преобразования Хафа. На этом рисунке сверху показано изображение размерами 101x101 пиксель, на котором отмечены пять точек – одна в центре, и четыре на углах квадратного изображения. Для каждой из них

построен образ (отрезок синусоиды) в плоскости $\rho\theta$ с дискретными целочисленными отсчетами по обеим осям, как показано на нижнем рисунке. Диапазон значений θ составляет [–90°, +90°], а диапазон значений $\rho$ равен $[-2D, 2D]$, где $D$ – расстояние между углами изображения по диагонали. Как видно, каждая из синусоид имеет свою амплитуду и фазу. Горизонтальная линия, полученная при отображении точки 1, представляет собой синусоиду с нулевой амплитудой.

В точке *A* пересекаются кривые, которые соответствуют точкам 1, 3 и 5 в плоскости *xy* исходного изображения. Координаты точки $A$ показывают, что упомянутые три точки лежат на прямой линии, проходящей через начало координат (т.е. $\rho = 0$) под углом 45°. Аналогично пересечение кривых в точке $B$ указывает, что точки 2, 3 и 4 лежат на прямой, проходящей под углом –45° на расстоянии от начала координат $\rho = 71$ (округленная до ближайшего целого значения половина длины диагонали от начала координат к правому нижнему углу изображения).

До этих пор рассматривались прямые линии на изображении. Однако преобразование Хафа можно применить также к любой функции вида $g(\mathbf{v}, \mathbf{c}) = 0$, где $\mathbf{v}$ – вектор координат, а $\mathbf{c}$ – вектор коэффициентов. Например, точки, лежащие на окружности

$$(x - c_1)^2 + (\text{y} - c_2)^2 = c_3^2,$$

также могут обнаруживаться описанным выше методом. Основное отличие состоит в увеличении числа параметров до трех ($c_1$, $c_2$ и $c_3$), что приводит к трехмерному пространству параметров с кубическими ячейками, накапливаемые значения в которых имеют вид $A(p, q, r)$. Процедура состоит в последовательном присваивании параметрам $c_1$ и $c_2$ всех сочетаний допустимых дискретных значений, нахождением для каждой пары значения $c_3$, которое бы удовлетворяло предыдущему уравнению, и увеличением накопленного значения в ячейке, отвечающей тройке параметров ($c_1$, $c_2$, $c_3$).

Сложность преобразования Хафа пропорциональна числу координат и коэффициентов в данном функциональном представлении.

Известны дальнейшие обобщения преобразования Хафа на случаи обнаружения кривых, не имеющих простого аналитического представления в случае применения этого преобразования к полутоновым изображениям.

# Глава 16. Структурные методы сегментации

## 16.1. Выращивание областей

Рассмотрим методы, основанные на непосредственном поиске областей, порождающих сегментацию изображения.

*Выращивание областей* – это метод группировки пикселей или подобластей в более крупные области по некоторым заранее определенным критериям. Вначале берется множество точек, которые играют роль «центров кристаллизации», а затем на них наращиваются области путем присоединения к ним тех пикселей из числа соседних, которые по определенным свойствам близки к центру кристаллизации (например, имеют близкую яркость или цвет).

Выбор первоначального множества, как правило, основан на существе решаемой задачи. Для каждого пикселя вычисляются одни и те же свойства, на основании которых он будет отнесен к той или иной области в процессе выращивания.

Выбор критериев сходства пикселей зависит не только от конкретной рассматриваемой задачи, но и от типа самого изображения и представленных в нем данных. Например, задача анализа данных спутниковой съемки земной поверхности существенно основана на использовании цвета. Если изображения монохромные, анализ областей проводится с использованием дескрипторов, основанных на значениях яркости и пространственных характеристиках, например, связанных с окрестностью этого пикселя.

Важно отметить, что использование при выращивании областей одних лишь характеристик пикселей может привести к ошибочным результатам, если это делается без учета информации о связности областей, так как если строить области, группируя вместе пиксели с одинаковой яркостью и не обращать внимания на их связность, то это приведет к сегментации, лишенной смысла.

Принципиальным условием при выращивании областей является формулировка правила остановки этого процесса. Выращивание некоторой области следует прекратить, когда в изображении больше нет пикселей, удовлетворяющих критериям присоединения к данной области. Такими критериями могут быть, например, размеры и форма выращиваемой области, а также сходство между пикселем-кандидатом и пикселями, объединенными к данному моменту (скажем, путем сравнения значений яркости нового канди-

дата и средней яркости уже выращенной области). Использование дескрипторов такого типа основано на предположении, что имеется хотя бы грубая модель ожидаемых результатов.

Пусть $f(x, y)$ — исходное изображение. Обозначим $S(x, y)$ массив, в котором единицами отмечено положение центров кристаллизации, а остальные элементы нулевые. Обозначим $Q$ предикат (условие, критерий), который будет применяться в каждой точке $(x, y)$. Предполагается, что массивы $f$ и $S$ имеют одинаковые размеры. Следующий простой алгоритм является примером метода выращивания областей на основе 8-связности.

1. Найти все компоненты связности $S(x, y)$ и применить к каждой из них операцию эрозии до тех пор, пока не останется один пиксель. Пометить все такие пиксели значением 1; остальные пиксели $S$ помечаются нулевым значением.

2. Сформировать изображение $f_0$ по следующему правилу: если в точке $(x, y)$ исходное изображение удовлетворяет заданному предикату $Q$, принимаем $f_0(x, y) = 1$, иначе принимаем $f_0(x, y) = 0$.

3. Сформировать изображение $g$ путем присоединения к каждому центру кристаллизации в $S$ всех единичных элементов $f_0$, которые являются 8-связными с этим центром кристаллизации.

4. Пометить пиксели каждой связной компоненты $g$ уникальной пометкой данной области (например, числами 1, 2, 3...). Это и будет сегментированное изображение, полученное выращиванием областей.

Проиллюстрируем действие этого алгоритма на примере. Здесь сверху вниз и слева направо показаны:

1. Рентгеновское изображение сварного шва с дефектами;

2. Гистограмма исходного изображения;

3. Изображение начальных областей кристаллизации;

4. Найденные центры кристаллизации;

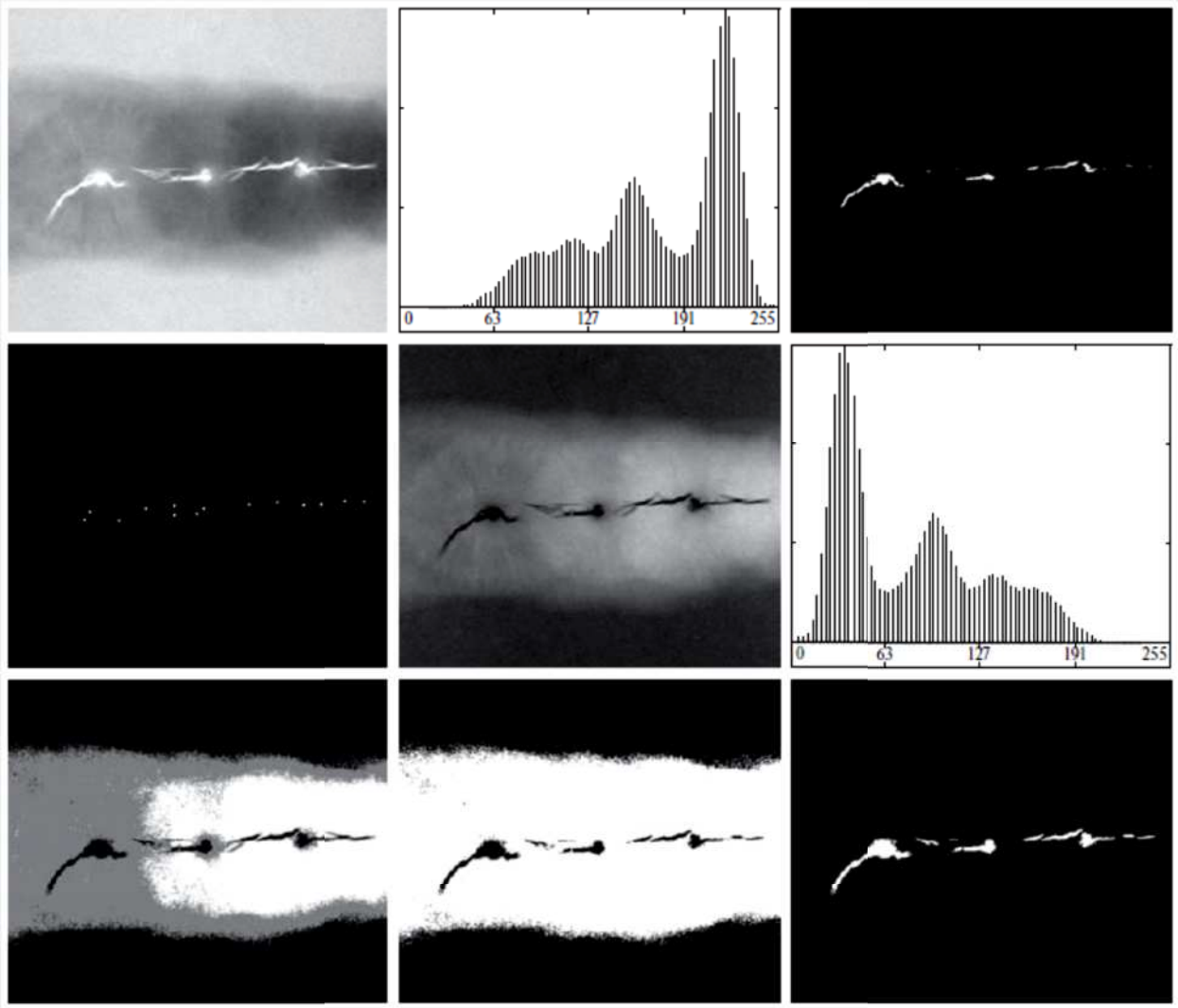
5. Абсолютная величина разности константы 255 (значение в центре кристаллизации) и исходного изображения;

6. Гистограмма последнего изображения;

7. Результат применения преобразования с двойным порогом к разностному изображению;

8. Результат порогового разделения разностного изображения по меньшему из двух порогов;

9. Результат сегментации, полученный методом выращивания областей.

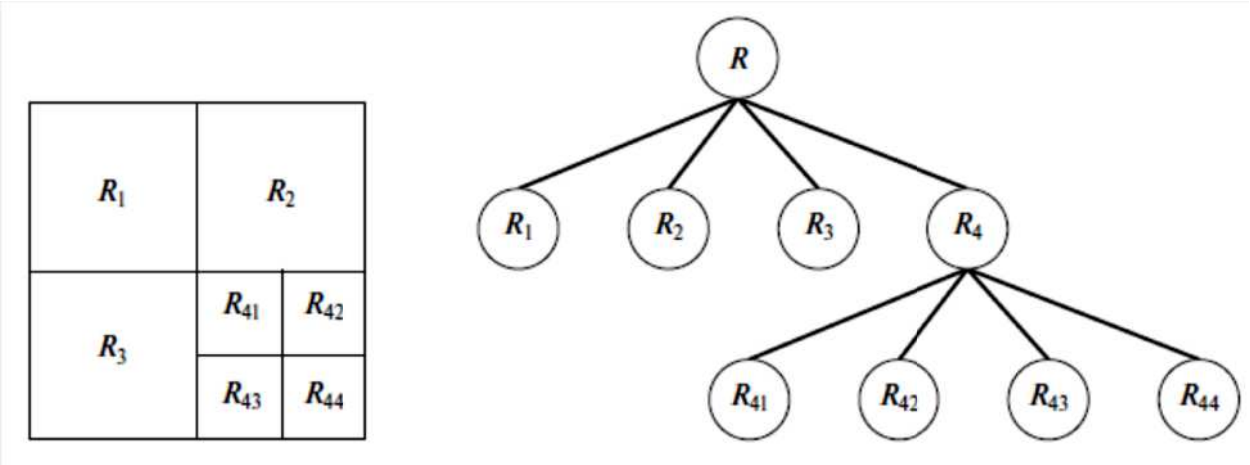



## 16.2. Разделение и слияние областей

Обратим внимание еще раз, что рассмотренная выше процедура выращивания областей была основана на увеличении областей, начиная с одного или небольшого множества пикселей, которые были названы центрами кристаллизации. Метод разделения и слияния областей представляет собой в некотором смысле противоположную процедуру. Вначале все изображение разбивается на непересекающееся множество областей, а затем в результате анализа каждая область либо разделяется на две или более подобластей, либо две или более соседних областей, напротив, сливаются в одну.

Предположим, что вся область изображения обозначена $R$ и, исходя из конкретной задачи, выбран предикат $Q$. Один из наиболее естественных подходов к сегментации $R$ состоит в том, чтобы последовательно разбивать эту область на все более и более мелкие области-квадранты (т.е. четверти) $R_i$, пока для каждой из них не будет выполняться условие $Q(R_i) = TRUE$, либо пока разбиение не будет остановлено в связи с достижением минимально допустимого размера области. Работа алгоритма начинается со всей области изображения. Если $Q(R_i) = FALSE$, изображение делится на четверти вертикальной и горизонтальной прямыми линиями, проходящими через середину. Если для какой-то четверти предикат $Q$ принимает значение $FALSE$, она аналогичным способом делится на более мелкие четверти и так далее. Такое разбиение удобно представлять в форме так называемого *квадродерева* (т.е. дерева, у которого вершины, не являющиеся листьями, имеют в точности четыре потомка), как показано на следующем рисунке. Отметим, что корень дерева соответствует целому изображению, а каждая другая вершина – какой-то из подобластей. В приведенном примере дальнейшему разбиению подверглась только область $R_4$.

А теперь следует обратить внимание на то, что если использовать только операцию разделения, то в окончательном разбиении изображения могут присутствовать смежные области с одинаковыми свойствами. Этот недостаток можно устранить, применяя наряду с разделением также операцию слияния. При этом требуется, чтобы слиянию подвергались только такие смежные области, пиксели которых в совокупности удовлетворяют предикату $Q$. Иначе говоря, две смежные области $R_j$ и $R_k$ сливаются только в том случае, если $Q(R_j \cup \mathrm{R}_k) = TRUE$.



Описанный метод можно описать в виде следующей формальной процедуры, на каждом шаге которой выполняются следующие действия:

1. Любая область $R_i$, для которой $Q(R_i) = FALSE$, разделяется на четыре непересекающиеся четверти.

2. Когда продолжать разбиение невозможно, любые две смежные области $R_j$ и $R_k$, для которых $Q(R_j \cup \mathrm{R}_k) = TRUE$, объединяются в одну.

3. Если невозможно выполнить ни одной операции слияния, то наступает окончание процедуры.

При этом обычно устанавливаются минимальные размеры квадрообласти, по достижении которых дальнейшее разделение уже не производится.

Следует отметить, что существуют различные варианты изложенной основной схемы. Например, можно получить существенное упрощение, если на шаге 2 допускать слияние любых двух смежных областей $R_j$ и $R_k$ в случае, когда каждая из них по отдельности удовлетворяет предикату $Q$. В результате такого допущения алгоритм становится значительно проще и быстрее, поскольку проверка предиката выполняется только для каждой квадрообласти в отдельности.

Рассмотрим **пример** сегментации методом разделения и слияния областей, в котором как раз используется такое упрощение.

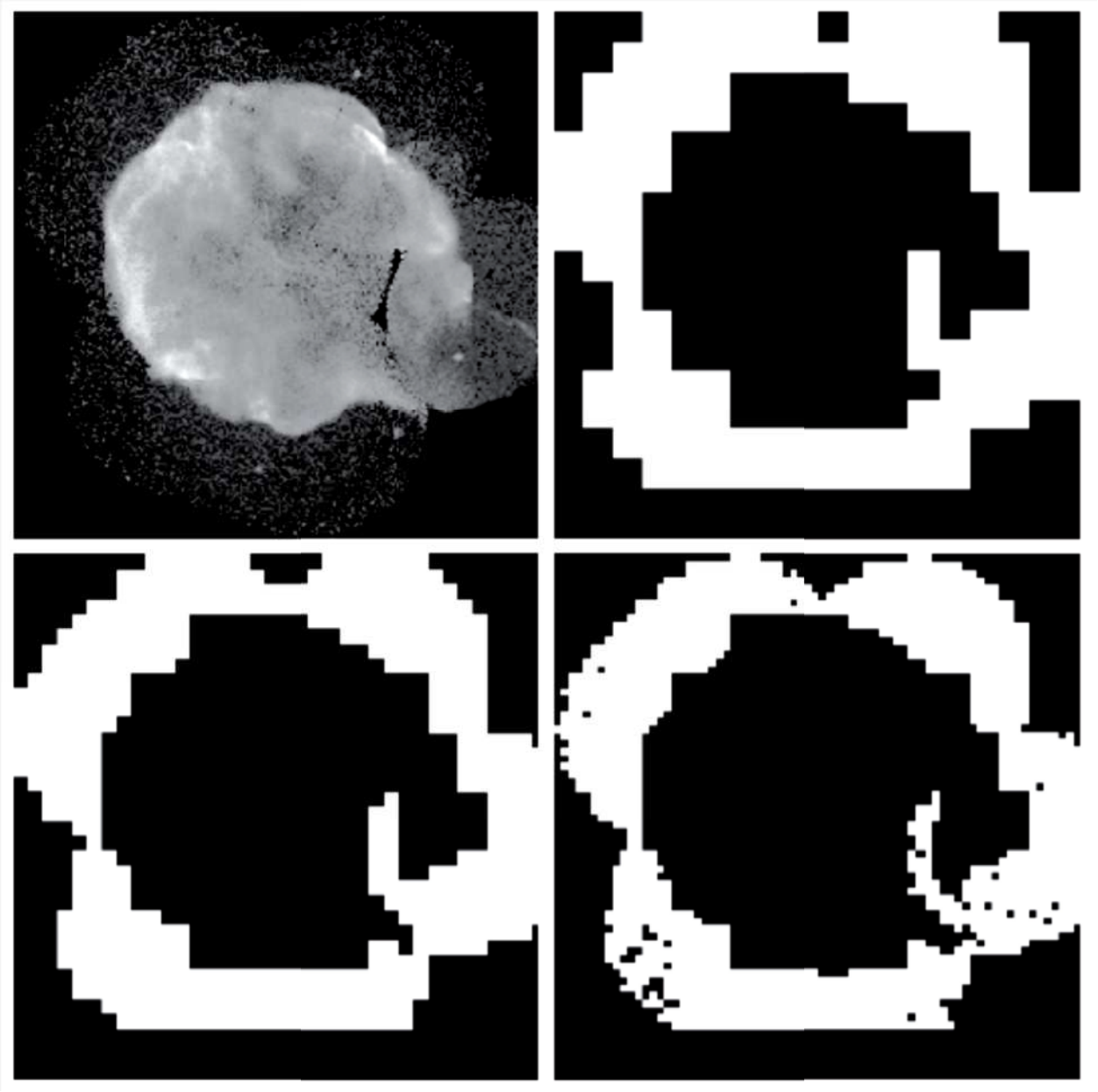
На следующем рисунке приведено изображение результата взрыва сверхновой звезды в созвездии Лебедя (Петля Лебедя), полученное в рентгеновском диапазоне. Требуется выделить «кольцо» разреженной материи, окружающее более плотную центральную область.

Данная область обладает рядом свойств, которые могут помочь ее сегментации. Прежде всего, как нетрудно заметить, яркости в этой области носят случайный характер; следовательно, стандартное отклонение по области должно быть выше, чем для области фона, значение которого близко к 0, или для центральной области, являющейся относительно гладкой (не совсем случайной). Среднюю яркость интересующей области также легко охарактеризовать – она должна быть больше, чем яркость области фона, и меньше, чем яркость центральной области. Таким образом, можно ожидать, что сегментировать данную область можно с помощью следующего предиката:

$$Q = \begin{cases} TRUE, & \text{если } \sigma > a \;\; И \;\; 0 < m < b, \\ FALSE, & \text{в противном случае,} \end{cases}$$

где $m$ и $\sigma$ – соответственно среднее значение и стандартное отклонение для пикселей в квадрообласти, а $a$ и $b$ – константы.

Анализ нескольких участков интересующей области показал, что средняя яркость на них не превышает 125, а стандартное отклонение всегда больше 10. На остальных трех рисунках представлены результаты, полученные с такими значениями $a$ и $b$ при минимально допустимом размере квадрообласти от 32 до 8. Пиксели квадрообластей, удовлетворяющих этому предикату, помечены белым цветом, а прочие пиксели – черным. Наилучшее соответствие форме интересующей области достигается при минимальных размерах квадрообластей 16×16. Мелкие черные квадраты на последнем рисунке – это квадрообласти размерами 8×8, для которых предикат не выполняется.



Использование квадрообластей меньшего размера приведет к еще большему числу таких черных квадратиков. Если использовать более крупные области, чем показано на рисунке, результаты сегментации будут иметь вид более грубых блоков.

Следует отметить, что во всех случаях сегментированные области (белого цвета) полностью разделяют внутреннюю сглаженную область и фон изображения. Таким образом, в результате сегментации исходное изображение фактически разделено на три области в соответствии с их основными характеристиками: фон, плотная и разреженная области.

Использованные в предыдущем примере свойства, основанные на среднем значении и стандартном отклонении яркости пикселей внутри некоторой области, представляют собой попытку количественно охарактеризовать *текстуру* области. Идея *сегментации по текстуре* основана на использовании количественных текстурных признаков для построения предикатов $Q(R_i)$. Иными словами, сегментация по текстуре может осуществляться с помощью любого из рассмотренных в этом разделе методов, если предикаты основаны на характеристиках текстуры.

# БИБЛИОГРАФИЧЕСКИЙ СПИСОК


1. Гонсалес Р., Вудс Р. Цифровая обработка изображений. – М.: Техносфера, 2005.
2. Гонсалес Р., Вудс Р., Эддинс С. Цифровая обработка изображений в среде Matlab. – М.: Техносфера, 2006.
3. Шапиро Л., Стокман Дж. Компьютерное зрение. – М.: Бином. Лаборатория знаний, 2006.
4. Визильтер Ю.В., Желтов С.Ю., Бондаренко А.В., Ососков М.В., Моржин А.В. Обработка и анализ изображений в задачах машинного зрения. Курс лекций и практических занятий. – М.: Физматкнига, 2010.
5. Яне Б. Цифровая обработка изображений. – М.: Техносфера, 2007.
6. Фисенко В.Т., Фисенко Т.Ю. Компьютерная обработка и распознавание изображений: учебное пособие. – СПб: СПбГУ ИТМО, 2008. –192 с.
7. Визильтер Ю. В. и др. Обработка и анализ цифровых изображений с примерами на LabVIEW IMAQ Vision. – М.: ДМК Пресс, 2007.
8. Крашенинников В. Р. Основы теории обработки изображений: учебное пособие. – Ульяновск: УлГТУ, 2003.
9. Методы компьютерной обработки изображений. Под ред. В. А. Сойфера. – М.: Физматлит, 2001. 784 с.
10. Цифровая обработка изображений. Под ред. Спектора А.А. – Новосибирск: НГТУ, 2002.
11. Косых В.П. Цифровая обработка изображений: учебное пособие. – Новосибирск: НГУ, 2006.