

|  |           |
|--|-----------|
| <b>БИЛЕТ № 1</b>   | <b>4</b>  |
| 1. Частотная фильтрация.   | 4         |
| 2. Распознавание образов. Нейронные сети.  | 5         |
| <b>БИЛЕТ № 2</b>   | <b>7</b>  |
| 1. Видоизменение гистограммы. Эквализация.   | 7         |
| 2. Распознавание образов. Машины опорных векторов. (SVM, support vector machine).                | 9         |
| <b>БИЛЕТ № 3</b>   | <b>10</b> |
| 1. Первая и вторая производные в обработке изображений   | 10        |
| 2. Распознавание образов. Деревья решений.   | 11        |
| <b>БИЛЕТ № 4</b>   | <b>12</b> |
| 1. История цифровой обработки изображений.   | 12        |
| 2. Распознавание образов. Кластерный анализ  | 12        |
| <b>БИЛЕТ № 5</b>   | <b>14</b> |
| 1. Восстановление цифрового изображения. Подавление шумов с помощью частотной фильтрации.        | 14        |
| 2. Распознавание образов. Алгоритмы обучения с подкреплением.                                    | 16        |
| <b>БИЛЕТ № 6</b>   | <b>18</b> |
| 1. Области применения цифровой обработки изображений.  | 18        |
| 2. Распознавание образов. Сокращение размерности   | 18        |
| <b>БИЛЕТ № 7</b>   | <b>20</b> |
| 1. Восстановление цифрового изображения. Подавление шумов с помощью пространственной фильтрации. | 20        |
| 2. Распознавание образов. Показатели качества распознавания.                                     | 21        |
| <b>БИЛЕТ № 8</b>   | <b>23</b> |
| 1. Основные этапы цифровой обработки изображений.  | 23        |
| 2. Распознавание образов. Характерные признаки объектов.   | 24        |
| <b>БИЛЕТ № 9</b>   | <b>25</b> |
| 1. Восстановление цифрового изображения. Модели шума.  | 25        |
| 2. Распознавание образов. Скрытые Марковские модели.   | 27        |
| <b>БИЛЕТ № 10</b>  | <b>29</b> |
| 1. Связь зрительной системы человека с цифровой обработкой изображений.                          | 29        |
| 2. Распознавание образов. Глубокие нейронные сети.   | 31        |
| <b>БИЛЕТ № 11</b>  | <b>33</b> |
| 1. Описание цифрового изображения. Дескрипторы областей  | 33        |
| 2. Распознавание образов. Байесовы сети.   | 34        |

|  |           |
|--|-----------|
| <b>БИЛЕТ № 12</b>  | <b>35</b> |
| 1. Характеристики света в отношении формирования изображения.                          | 35        |
| 2. Распознавание образов. Нечеткие конечные автоматы.                                  | 35        |
| <b>БИЛЕТ № 13</b>  | <b>38</b> |
| 1. Описание цифрового изображения. Дескрипторы границ.                                 | 38        |
| 2. Распознавание жестов. Основные системы и математические модели распознавания.       | 39        |
| <b>БИЛЕТ № 14</b>  | <b>41</b> |
| 1. Способы захвата (регистрации) цифровых изображений.                                 | 41        |
| 2. Распознавание образов. Параллельный захват и отслеживание динамических жестов руки. | 41        |
| <b>БИЛЕТ № 15</b>  | <b>44</b> |
| 1. Описание цифрового изображения: коды, линии, границы, области.                      | 44        |
| 2. Распознавание образов. Нейронные сети.  | 46        |
| <b>БИЛЕТ № 16</b>  | <b>48</b> |
| 1. Равномерная пространственная дискретизация и квантование изображения.               | 48        |
| 2. Распознавание образов. Машины опорных векторов. (SVM, support vector machine).      | 49        |
| <b>БИЛЕТ № 17</b>  | <b>50</b> |
| 1. Сегментация. Обнаружение областей.  | 50        |
| 2. Основные взаимосвязи между пикселями.   | 52        |
| <b>БИЛЕТ № 18</b>  | <b>53</b> |
| 1. Сегментация. Обнаружение контуров.  | 53        |
| 2. Пространственная фильтрация.  | 54        |
| <b>БИЛЕТ № 19</b>  | <b>56</b> |
| 1. Цветовые модели.  | 56        |
| 2. Видоизменение гистограммы. Приведение (задание гистограммы).                        | 56        |
| <b>БИЛЕТ № 20</b>  | <b>60</b> |
| 1. Сегментация. Обнаружение моделей сегментов.   | 60        |
| 2. Преобразование Фурье и частотная область.   | 61        |
| <b>БИЛЕТ № 21</b>  | <b>62</b> |
| 1. Диаграмма цветностей.   | 62        |
| 2. Применение морфологических алгоритмов к полуточковым изображениям.                  | 63        |
| <b>БИЛЕТ № 22</b>  | <b>69</b> |
| 1. Обработка изображений в псевдоцветах.   | 69        |
| 2. Морфологические алгоритмы.  | 70        |
| <b>БИЛЕТ № 23</b>  | <b>72</b> |

|  |           |
|--|-----------|
| 1. Цветовые преобразования.                              | 72        |
| 2. Основы морфологической обработки                      | 77        |
| <b>БИЛЕТ № 24</b>  | <b>78</b> |
| 1. Представление знаний в системах компьютерного зрения. | 78        |
| 2. Сглаживание и повышение резкости цветных изображений. | 79        |
| <b>БИЛЕТ № 25</b>  | <b>82</b> |
| 1. Частотные фильтры повышения резкости.                 | 82        |
| 2. Цветовая сегментация.                                 | 82        |
| <b>БИЛЕТ № 26</b>  | <b>84</b> |
| 1. Сглаживающие частотные фильтры.                       | 84        |
| 2. Основные градационные преобразования                  | 85        |

## **БИЛЕТ № 1**

### **1. Частотная фильтрация.**

Методы обработки в частотной области основываются на модификации сигнала, формируемого путем применения к изображению преобразования Фурье.

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

Прямое дискретное Фурье преобразование функции  $f(x, y)$ , изображения размером  $M \times N$ .

#### **Фильтрация:**

- *пространственная* (относится к плоскости изображения как таковой, и данная категория объединяет подходы, основанные на прямом манипулировании пикселями изображения)
- *частотная* (основываются на модификации сигнала, формируемого путем применения к изображению преобразования Фурье)

#### **Виды фильтров:**

- *фильтр низких частот* (сглаживание изображения - плавное изменение яркости изображения)
- *фильтр высоких частот* (повышение резкости изображения - быстрое изменение яркости изображения (границы объектов))
- *полосовой фильтр*

#### **Структурная схема фильтрации:**

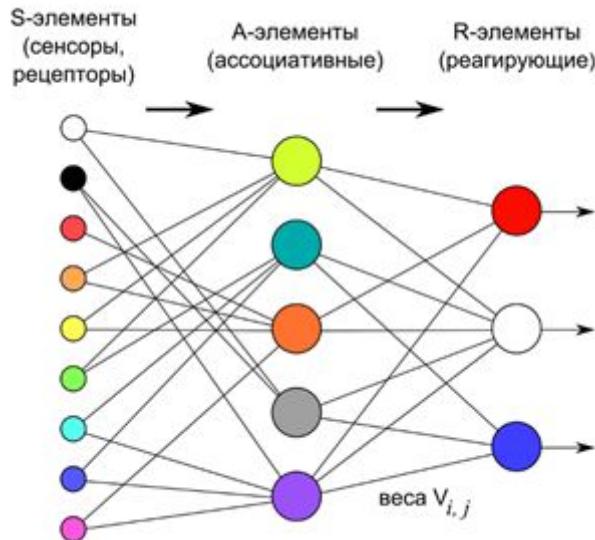
1. Исходное изображение умножается на  $(-1)^{x+y}$ , чтобы его Фурье преобразование оказалось центрированным.
2. Вычисляется прямое дискретное преобразование Фурье (ДПФ) исходного изображения, получаем  $F(u,v)$ .
3. Функция  $F(u,v)$  умножается на функцию фильтра  $H(u,v)$ .
4. Вычисляется обратное ДПФ от результата предыдущего шага.
5. Выделяется вещественная часть результата предыдущего шага.
6. Результат предыдущего шага умножается на  $(-1)^{x+y}$ .

## 2. Распознавание образов. Нейронные сети.

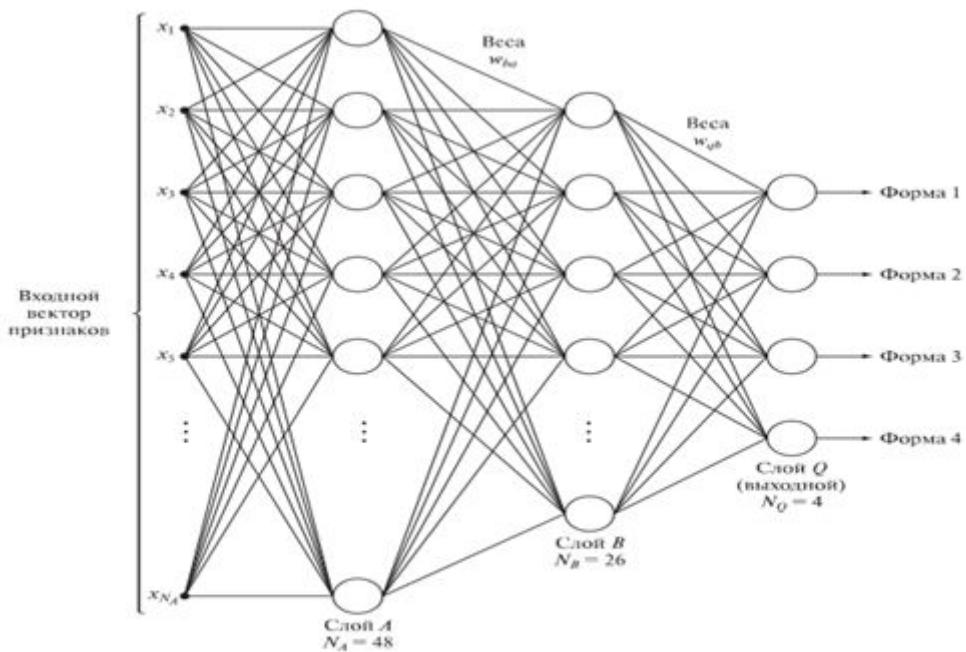
Под *образом* подразумевается некоторая упорядоченная совокупность дескрипторов. В литературе по распознаванию образов эти дескрипторы часто называют *признаками*. *Классом образов* называется совокупность образов, обладающих некоторыми общими свойствами. Будем обозначать классы символами  $\omega_1, \omega_2, \dots, \omega_W$ , где  $W$  — число классов. Под машинным *распознаванием образов* понимаются методы, позволяющие относить образы к тем или иным классам — автоматически или с минимальным вмешательством человека.

Нейронные сети позволяют решать задачу определения класса объекта. В процессе обучения сети ей предлагаются различные эталонные образцы с меткой того, к какому классу объекта их можно отнести. Эти образцы должны быть однозначны и непротиворечивы для того, чтобы нейронная сеть выдавала высокие вероятности принадлежности объекта к одному конкретному классу.

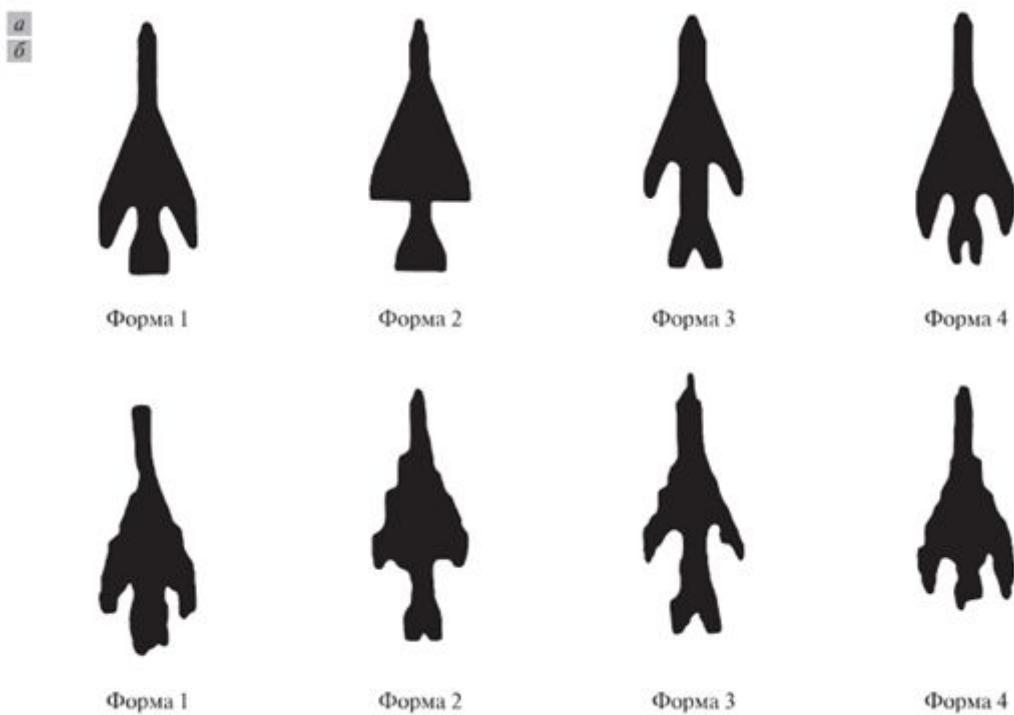
**Перцептрон** — математическая или компьютерная модель восприятия информации мозгом (кибернетическая модель мозга). Перцептрон стал одной из первых моделей нейросетей. Перцептрон состоит из трёх типов элементов, именно: поступающие от *сенсорных* элементов сигналы передаются *ассоциативным* элементам, а затем *реагирующими* элементам.



Пример применения нейронной сети



**Рис. 12.19.** Трехслойная нейронная сеть, использованная для распознавания фигур на рис. 12.18. (Изображение предоставлено д-ром Лалит Гупта, кафедра ЕСЕ, Южный университет шт. Иллинойс)

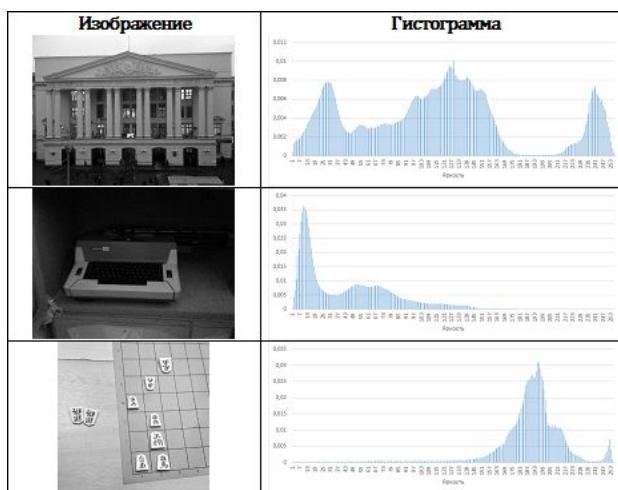


**Рис. 12.18.** (a) Эталонные формы и (б) типичные искаженные фигуры, использованные для обучения нейронной сети на рис. 12.19. (Изображения

## БИЛЕТ № 2

### 1. Видоизменение гистограммы. Эквализация.

Гистограммой цифрового изображения с уровнями яркости в диапазоне  $[0, L - 1]$  называется дискретная функция  $h(r_k) = n_k$ , где  $r_k$  есть  $k$ -й уровень яркости, а  $n_k$  — число пикселей на изображении, имеющих яркость  $r_k$ . Общей практикой является нормировка гистограммы путем деления каждого из ее значений на общее число пикселей в изображении, обозначаемое произведением  $MN$ , где, как обычно,  $M$  и  $N$  суть число строк и столбцов изображения, то есть его размеры. Тем самым значения нормированной гистограммы будут  $p(r_k) = n_k/MN$  для  $k = 0, 1, \dots, L - 1$ . Вообще говоря,  $p(r_k)$  есть оценка вероятности появления пикселя<sup>3</sup> со значением яркости  $r_k$ . Заметим, что сумма всех значений нормированной гистограммы равна единице.



По горизонтальной оси каждого графика отложены значения уровней яркости  $r_k$ . По вертикальной оси — значения гистограммы  $h(r_k) = n_k$  или  $p(r_k) = n_k/MN$ , если они нормированы. Тем самым гистограммы могут быть представлены просто как графики зависимостей  $h(r_k) = n_k$  от  $r_k$  или  $p(r_k) = n_k/MN$  от  $r_k$ .

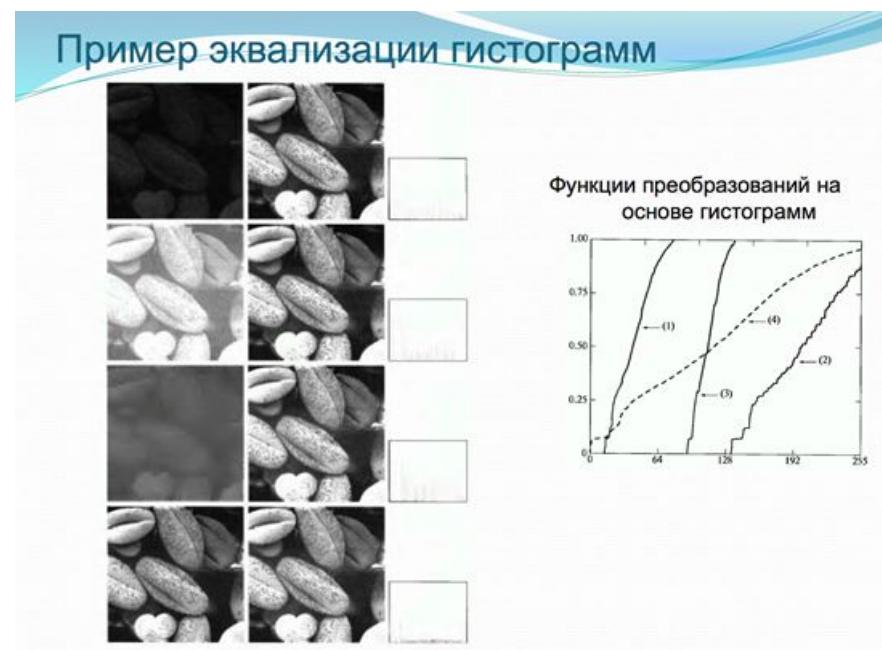
Эффективная методика повышения качества изображений - преобразование, при котором выходное изображение имеет равномерное распределение, называется эквализацией (выравниванием) гистограмм. Эквализация гистограммы автоматически находит функцию преобразования, которая стремится сформировать выходное изображение с равномерной гистограммой.

Легко видеть, что на гистограмме темного изображения ненулевые уровни сконцентрированы в области низких (темных) значений диапазона яркостей. Аналогично значимые уровни гистограммы светлого изображения смещены к верхней части диапазона.

Интуитивно можно сделать вывод, что наиболее удобным для восприятия человеком будет изображение, у которого гистограмма близка к равномерному распределению.

Т.е. для улучшения визуального качества к изображению надо применить такое преобразование, чтобы гистограмма результата содержала все возможные значения яркости и при этом в примерно одинаковом количестве. Такое преобразование называется эквализацией гистограммы. В результате эквализации гистограммы в большинстве случаев существенно расширяется динамический диапазон изображения, что позволяет отобразить ранее не замеченные детали.

**На второй пикче представьте, что гистограммы после эквализации как бы равномерно распределены по всем осям (их там не видно).**



## **2. Распознавание образов. Машины опорных векторов. (SVM, support vector machine).**

В исходном виде SVM представляет собой алгоритм, обучающийся ровно одной задаче: различению объектов двух классов. И этому SVM обучается очень быстро по сравнению, например, с нейронными сетями. Основная идея SVM такова. SVM учится только с учителем. Возьмем обучающее множество  $\mathbf{X} = \mathbf{X}_1, \dots, \mathbf{X}_N$  векторов признаков в евклидовом пространстве  $\mathbb{R}^d$ , разбитое на два класса, скажем, крестики и нолики. Попробуем разделить все пространство гиперплоскостью так, чтобы крестики попали в одно открытое полупространство, а нолики - в другое. Предположим, что такая гиперплоскость существует. Тогда она не единственна, так как множество разделяющих гиперплоскостей открыто. Попробуем найти в каком-то смысле самую лучшую гиперплоскость. И если мы ее нашли, то постановим все вектора признаков (уже не только из обучающего множества), попадающие в полупространство с крестиками, считать крестиками, а остальные - ноликами. А вероятность неправильной классификации оценивается как некоторая непрерывная убывающая функция от расстояния до разделяющей плоскости, равная 1/2 в нуле и стремящаяся к 0 на бесконечности.

## **БИЛЕТ № 3**

### **1. Первая и вторая производные в обработке изображений**

Производные дискретной функции определяются в терминах разностей. Эти разности можно задать различными способами, однако мы будем руководствоваться следующим.

Первая производная должна быть:

1. равной нулю на плоских участках (областиах с постоянным уровнем яркости);
2. ненулевой в начале и в конце ступеньки или склона яркости;
3. ненулевой на склонах яркости.

Аналогично, вторая производная должна быть:

1. равной нулю на плоских участках;
2. ненулевой в начале и в конце ступеньки или склона яркости;
3. равной нулю на склонах постоянной крутизны.

Первая производная одномерной функции  $f(x)$  определяется как разность значений соседних элементов:

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

Аналогично, вторая производная определяется как разность соседних значений первой производной:

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

**Первая производная используется в основном для выделения контуров, тем не менее она находит применение и в задачах улучшения.** (Первая производная – градиент изображения – производная от функции яркости изображения. Она используется, например, в операторе Собеля для нахождения границ изображения.)

**В большинстве приложений методов улучшения изображений (лапласиан, градиент) вторая производная оказывается более предпочтительной, чем первая, благодаря большему усилению мелких деталей.**

Сравнивая отклики первой и второй производных, можно отметить следующее:

1. Первая производная обычно дает в результате более толстые контуры.

2. Вторая производная дает больший по величине отклик на мелкие детали - как на отдельных точках, так и на тонких линиях.
3. Отклик на ступеньку у первой производной, как правило, выше, чем у второй.
4. На наклонных контурах вторая производная дает двойной отклик.

## **2. Распознавание образов. Деревья решений.**

**Дерево принятия решений** (также может называться деревом классификации или регрессионным деревом) — средство поддержки принятия решений, использующееся в статистике и анализе данных для прогнозных моделей. Структура дерева представляет собой «листья» и «ветки». На ребрах («ветках») дерева решения записаны атрибуты, от которых зависит целевая функция, в «листьях» записаны значения целевой функции, а в остальных узлах — атрибуты, по которым различаются случаи. Чтобы классифицировать новый случай, надо спуститься по дереву до листа и выдать соответствующее значение. Подобные деревья решений широко используются в интеллектуальном анализе данных. Цель состоит в том, чтобы создать модель, которая предсказывает значение целевой переменной на основе нескольких переменных на входе.

## **БИЛЕТ № 4**

### **1. История цифровой обработки изображений.**

В начале 1920-х годов была внедрена система «Бартлейн» для передачи изображений по кабелю, что позволило уменьшить время доставки иллюстраций через Атлантику с обычной недельной задержки до менее чем трех часов. Ранние системы «Бартлейн» были способны кодировать изображения с помощью пяти градаций яркости. В 1929 г. эти возможности были увеличены до 15 градаций.

Хотя в вышеприведенных примерах фигурируют цифровые изображения, их нельзя рассматривать как результат цифровой обработки изображений в контексте нашего определения, коль скоро при их получении не использовались компьютеры.

Первые компьютеры с мощностью, достаточной для выполнения осмысленных задач цифровой обработки изображений, появились в начале 1960-х гг. Рождение ЦОС прослеживается с момента возникновения таких машин и появления программ изучения космоса.

Работы по использованию вычислительной техники для улучшения визуального качества изображений, были развернуты в Лаборатории реактивного движения в Пасадене, шт. Калифорния, в 1964 г., когда переданные космическим аппаратом «Рейнджер-7» изображения лунной поверхности были подвергнуты компьютерной обработке для исправления различных искажений, обусловленных конструкцией бортовой телевизионной камеры.

Параллельно с космическими исследованиями в конце 1960-х — начале 1970-х гг. методы цифровой обработки изображений начали применяться в медицине, дистанционном исследовании земных ресурсов, астрономии. В начале 1970-х гг. была изобретена рентгеновская вычислительная томография, кратко называемая также компьютерной томографией (КТ).

С 1960-х гг. до настоящего времени область применения обработки изображений значительно расширилась. Помимо медицинских и космических приложений, методы цифровой обработки изображений сегодня используются в широком круге областей. В астрономии, биологии, медицинской радиологии, промышленности, в оборонной и правоохранительной сфере.

### **2. Распознавание образов. Кластерный анализ**

Кластерный анализ - математическая процедура многомерного анализа, позволяющая на основе множества показателей, характеризующих ряд состояний объектов (образов), сгруппировать их в классы (кластеры) таким образом, чтобы объекты, входящие в один класс (образ), были более однородными, сходными по сравнению с объектами, входящими в другие классы. На основе численно выраженных параметров объектов вычисляются расстояния между ними, которые могут выражаться в евклидовой метрике (наиболее употребимой), так и в других метриках.

k-средних — это итеративный метод, который используется, чтобы разделить изображение на  $K$  кластеров. Базовый алгоритм приведён ниже:

1. Выбрать  $K$  центров кластеров, случайно или на основании некоторой эвристики

2. Поместить каждый пиксель изображения в кластер, центр которого ближе всего к этому пикселью
3. Заново вычислить центры кластеров, усредняя все пиксели в кластере
4. Повторять шаги 2 и 3 до сходимости (например, когда пиксели будут оставаться в том же кластере)

Здесь в качестве расстояния обычно берётся сумма квадратов или абсолютных значений разностей между пикселям и центром кластера. Разность обычно основана на цвете, яркости, текстуре и местоположении пикселя, или на взвешенной сумме этих факторов.  $K$  может быть выбрано вручную, случайно или эвристически.

Этот алгоритм гарантированно сходится, но он может не привести к оптимальному решению. Качество решения зависит от начального множества кластеров и значения  $K$ .

## БИЛЕТ № 5

### 1. Восстановление цифрового изображения. Подавление шумов с помощью частотной фильтрации.

всегда основного средства улучшения изображений. В этом разделе мы обсудим более специализированные полосовые, режекторные, узкополосные и узкополосные режекторные фильтры как средства для уменьшения или устранения периодического шума.

Режекторные фильтры удаляют или ослабляют частоты в кольцевой полосе вокруг начала координат преобразования Фурье. Передаточная функция идеального режекторного фильтра задается выражением

$$H(u,v) = \begin{cases} 1 & \text{при } D(u,v) < D_0 - \frac{W}{2}; \\ 0 & \text{при } D_0 - \frac{W}{2} \leq D(u,v) \leq D_0 + \frac{W}{2}; \\ 1 & \text{при } D(u,v) > D_0 + \frac{W}{2}, \end{cases} \quad (5.4-1)$$

где  $D(u,v)$  — расстояние (4.3-3), измеряемое от центра частотного прямоугольника,  $W$  — ширина кольца, а  $D_0$  — радиус окружности, проходящей через его середину.

Аналогично, передаточная функция режекторного фильтра Баттерворта порядка  $n$  задается выражением

$$H(u,v) = \frac{1}{1 + \left[ \frac{D(u,v)W}{D^2(u,v) - D_0^2} \right]^{2n}}, \quad (5.4-2)$$

а передаточная функция режекторного гауссова фильтра задается выражением

$$H(u,v) = 1 - e^{-\frac{1}{2} \left[ \frac{D^2(u,v) - D_0^2}{D(u,v)W} \right]^2}. \quad (5.4-3)$$

На Рис. 5.15 представлены трехмерные перспективные изображения этих фильтров.

■ На Рис. 5.16 (а), который повторяет Рис. 5.5 (а), представлено изображение, сильно искаженное синусоидальным шумом различных частот. На Рис. 5.16 (б), где представлен фурье-спектр этого изображения, хорошо видны частотные компоненты шума в виде пар симметричных ярких точек. В рассматриваемом примере эти компоненты лежат приблизительно на окружности с центром в начале координат

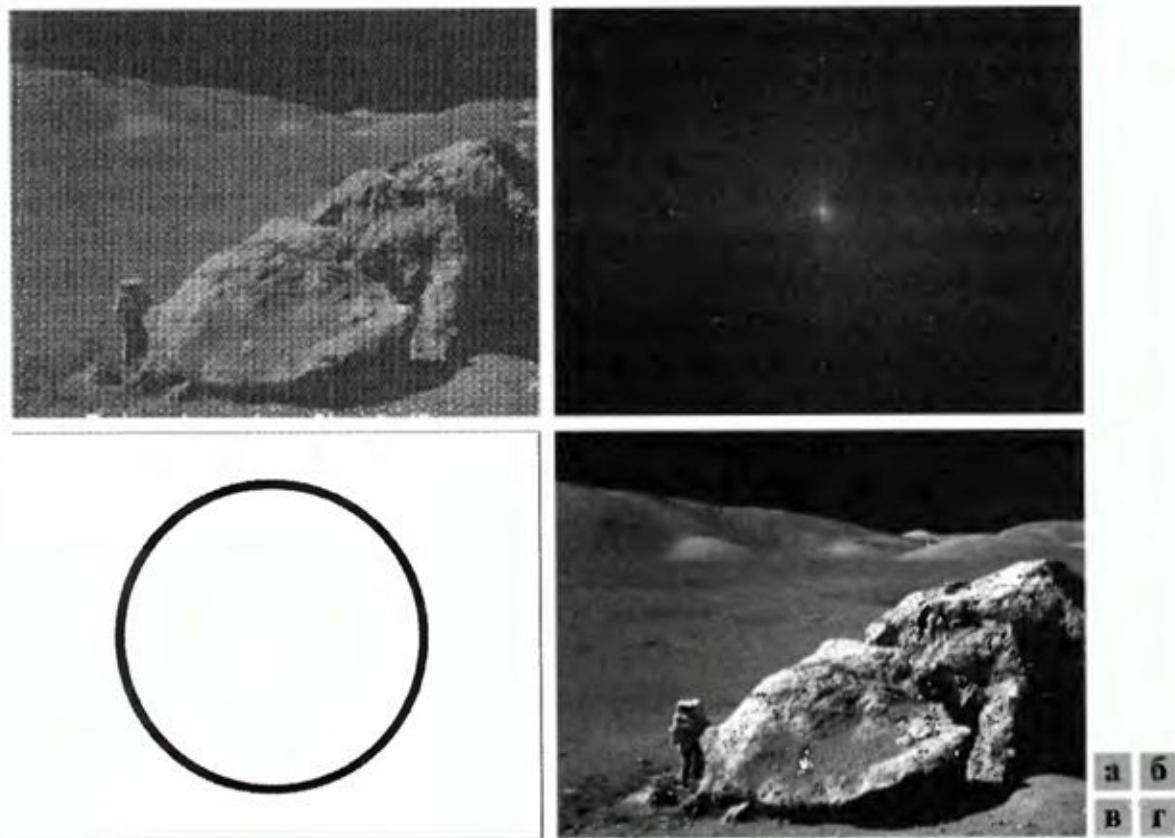


Рис. 5.16. (а) Изображение, искаженное синусоидальным шумом. (б) Спектр (а).  
(в) Режекторный фильтр Баттервортса (белым показаны точки со значением 1).  
(г) Результат фильтрации. (Изображение предоставлено Агентством NASA.)

частотного пространства, и, таким образом, использование центрально-симметричного режекторного фильтра представляется вполне оправданным. На Рис. 5.16 (в) представлен режекторный фильтр Баттерворта порядка 4, радиус и ширина которого выбраны так, чтобы шумовые импульсы полностью попадали в соответствующую область. Поскольку, вообще говоря, желательно удалять как можно меньшую часть фурье-преобразования, обычно используемые режекторные фильтры являются крутыми и узкими. Результат фильтрации изображения на Рис. 5.16 (а) с помощью выбранного фильтра представлен на Рис. 5.16 (г). Улучшение изображения вполне очевидно. Использованный простой метод фильтрации позволил эффективно восстановить даже мелкие детали и текстуры на изображении. Стоит отметить также, что прямой подход, основанный на фильтрации в пространственной области с использованием масок небольшого размера, не позволяет достичь подобного результата. ■

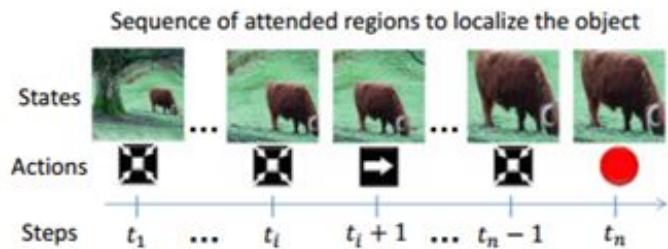
## **2. Распознавание образов. Алгоритмы обучения с подкреплением.**

Алгоритмы обучения с подкреплением => обучение нейронных сетей, машинное обучение.

- изучается, как агент должен действовать в окружении, чтобы максимизировать точность распознавания;
- ищется стратегия, приписывающая состояниям окружающей среды действия, которые должен предпринять агент в этих состояниях.

В ответ на каждую точку данных алгоритм выбирает действие. В ответ приходит сигнал, позволяющий алгоритму понять, насколько удачно было принято решение. На основе этого алгоритм изменяет свою стратегию для достижения лучшего результата. Накапливается база данных – база откликов среды, на их основе алгоритмы учатся и совершенствуются. Т.о. агент воздействует на среду, а среда воздействует на агента => система имеет обратную связь.

Пример в области распознавания образов: модель обнаружения объекта в сценах.



Агент фокусируется на предполагаемом объекте. Агент учится изменять ограничивающий прямоугольник с целью определения наиболее конкретного местоположения целевых объектов. Агент учится на своих ошибках и оптимизирует политику поиска объектов.

## **БИЛЕТ № 6**

### **1. Области применения цифровой обработки изображений.**

Классификация областей в соответствии с видами источников, формирующих соответствующие изображения:

**Изображения, полученные с помощью гамма-излучения**, используются главным образом в медицинской радиологии и астрономических наблюдениях. В медицинской радиологии применяется подход, при котором пациенту вводится радиоактивный изотоп, распад которого сопровождается гамма-излучением. Это излучение регистрируется детекторами гамма-излучения, сигналы которых используются для формирования изображения.

**Рентгеновские изображения** - Хорошо известно применение рентгеновских лучей для медицинской диагностики, однако они также широко используются в промышленности и других областях, в частности астрономии.

**Изображения в ультрафиолетовом диапазоне - Флуоресцентная микроскопия** — прекрасный метод исследования материалов, обладающих флуоресцирующими свойствами, либо в естественной форме, либо в результате обработки флуоресцирующими химикатами.

#### **Изображения в видимом и инфракрасном диапазонах**

|                       |           |   |
|-----------------------|-----------|---|
| Видимый синий цвет    | 0,45–0,52 | Максимальное проникание воды                    |
| Видимый зеленый цвет  | 0,52–0,60 | Измерение плотности растительного покрова       |
| Видимый красный цвет  | 0,63–0,69 | Различение формы растительности                 |
| Ближнее ИК-излучение  | 0,76–0,90 | Съемка береговой линии и распределения биомассы |
| Средний ИК-диапазон   | 1,55–1,75 | Содержание влаги в почве и растительности       |
| Тепловое ИК-излучение | 10,4–12,5 | Влажность почвы и температурная карта           |
| Средний ИК-диапазон   | 2,08–2,35 | Поиск полезных ископаемых                       |

**Изображения микроволнового диапазона** - применяются главным образом в радиолокации. Уникальным качеством радиолокации является возможность получения изображения любого района независимо от условий освещения и погоды.

**Изображения в диапазоне радиоволн** – основными областями применения таких изображений выступают медицина и астрономия. В медицине радиоволны используются для получения изображений методом ядерного магнитного резонанса (ЯМР). С помощью ЯМР можно получить срез вдоль любой плоскости.

### **2. Распознавание образов. Сокращение размерности**

Рассматривается возможность снижения размерности векторов признаков для задачи распознавания образов. Снижение размерности производится за счет отбрасывания

менее информативных признаков. В качестве такого критерия информативности предлагается показатель диагонально преобладания. Обсуждается эффективность данного показателя по сравнению с другими мерами мультиколлинеарности с вычислительной точки зрения. Предлагаются алгоритмы, реализующие предложенную идею при решении задачи распознавания изображений.

При построении классификаторов одним из важнейших этапов является выбор системы признаков. Удачный выбор признаков обеспечивает высокое качество распознавания. Обычно признаки стараются выбрать таким образом, чтобы высокое качество распознавания обеспечивалось при минимальной размерности признакового пространства. Это особенно актуально, если задача распознавания должна решаться с высокой оперативностью, возможно, в реальном времени. Для обеспечения этого требования в число признаков должны отбираться наиболее информативные. Известно два подхода к решению задачи отбора информативных признаков: наращивание числа признаков и редукция (сокращение) числа признаков. В последнем случае задается некоторая, заведомо избыточная система признаков, а затем осуществляется исключение из нее малоинформационных.

Идея алгоритма заключается в выделении наиболее информативных областей на изображениях путем вычисления простых достаточных оценок мультиколлинеарности векторов образов.

1-й шаг. Работа алгоритма начинается с присваивания переменным начальных значений размеров исходных изображений. Далее алгоритм реализуется в виде итерационной схемы, включающей следующую последовательность шагов.

2-й шаг. Значение  $k$  увеличивается на единицу. Исходные изображения обучающей выборки разбиваются на  $k^2$  одинаковых прямоугольных фрагментов путем деления каждой стороны на  $k$  частей, и вычисляется размерность соответствующего фрагменту изображения вектора.

3-й шаг. Осуществляется проверка условия  $K > M$ . Если да, то шаг 4, нет – шаг 6.

4-й шаг. Для каждого фрагмента изображения обучающей выборки, сформированного на шаге 2, составляется матрица с использованием которой формируется соответствующая информационная матрица.

5-й шаг. По информационной матрице вычисляется одна из мер мультиколлинеарности. Полученное значение присваивается всем отсчетам «своего» фрагмента на поле анализируемого изображения.

6-й шаг. Задается или определяется пороговое значение показателя (мультиколлинеарности). Все отсчеты (пиксели) изображения с равным ему или более высоким, чем заданный порог, значением включаются в число компонент всех векторов обучающих объектов.

## БИЛЕТ № 7

### 1. Восстановление цифрового изображения. Подавление шумов с помощью пространственной фильтрации.

Когда искажение изображения обусловлено исключительно наличием шума, равенства (5.1-1) и (5.1-2) приобретают вид

$$g(x, y) = f(x, y) + \eta(x, y) \quad (5.3-1)$$

и

$$G(u, v) = F(u, v) + N(u, v). \quad (5.3-2)$$

Слагаемое, описывающее шум, неизвестно, поэтому просто вычесть его из функции  $g(x, y)$  или  $G(u, v)$  невозможно. Обычно в случае периодического шума спектр  $G(u, v)$  дает возможность оценить величину  $N(u, v)$ , как было указано в Разделе 5.2.3. Тогда в целях построения приближения исходного изображения, величина  $N(u, v)$  может быть вычтена из функции  $G(u, v)$ . Однако этот случай является скорее исключением, чем правилом.

В тех ситуациях, когда на изображении присутствует только аддитивный шум, пространственная фильтрация является лучшим из возможных методов восстановления. Этот метод детально обсуждался в Разделах 3.5 и 3.6 как метод улучшения изображений. В подобном случае процедуры улучшения и восстановления становятся практически неразличимыми. За исключением вычислительной процедуры, характерной для использования некоторого особого фильтра, механизм применения всех нижеследующих фильтров в точности такой, как обсуждалось в Разделе 3.5.

### **Фильтр, основанный на вычислении среднего арифметического**

Такой фильтр, называемый *среднеарифметическим*, является простейшим среди усредняющих фильтров. Пусть  $S_{xy}$  обозначает прямоугольную окрестность (множество координат точек изображения) размерами  $m \times n$  с центром в точке  $(x,y)$ . Процедура фильтрации предполагает вычисление среднего арифметического значения искаженного изображения  $g(x,y)$  по окрестности  $S_{xy}$ . Значение восстановленного изображения  $\hat{f}$  в произвольной точке  $(x,y)$  представляет собой среднее арифметическое значений в точках, принадлежащих окрестности  $S_{xy}$ . Другими словами

$$\hat{f}(x,y) = \frac{1}{mn} \sum_{(s,t) \in S_{xy}} g(s,t). \quad (5.3-3)$$

Эта операция может быть реализована в виде свертки с маской, все коэффициенты которой равны  $1/mn$ . Как указывалось в Разделе 3.6.1, усредняющий фильтр просто сглаживает локальные вариации яркости на изображении. Уменьшение шума происходит в результате этого сглаживания.

### **Фильтр, основанный на вычислении среднего геометрического**

Изображение, восстановленное с использованием *среднегеометрического* фильтра, задается выражением

$$\hat{f}(x,y) = \left[ \prod_{(s,t) \in S_{xy}} g(s,t) \right]^{\frac{1}{mn}}. \quad (5.3-4)$$

Здесь значение восстановленного изображения в каждой точке  $(x,y)$  является корнем степени  $mn$  из произведения значений в точках окрестности  $S_{xy}$ . Как показывает Пример 5.2, применение среднегеометрического фильтра приводит к сглаживанию, сравнимому с тем, которое достигается при использовании среднеарифметического фильтра, но при этом теряется меньше деталей изображения.

## 2. Распознавание образов. Показатели качества распознавания.

Под *образом* подразумевается некоторая упорядоченная совокупность дескрипторов. В литературе по распознаванию образов эти дескрипторы часто называют *признаками*. *Классом образов* называется совокупность образов, обладающих некоторыми общими

свойствами. Будем обозначать классы символами  $\omega_1, \omega_2, \dots, \omega_W$ , где  $W$  — число классов. Под машинным *распознаванием образов* понимаются методы, позволяющие относить образы к тем или иным классам — автоматически или с минимальным вмешательством человека.

Основными показателями качества распознавания графических изображений являются:

- Вероятность правильного распознавания;
- Вероятность отказа от распознавания;
- Вероятность ложного распознавания.

Также выделяют условную вероятность ошибочной классификации, асимптотическую ожидаемую ошибку классификации, ожидаемую ошибку алгоритма классификации на выборке заданного объема.

## **БИЛЕТ № 8**

### **1. Основные этапы цифровой обработки изображений.**

1. Регистрация изображения
2. Улучшение изображения
3. Восстановление изображения
4. Обработка цветных изображений
5. Морфологическая обработка
6. Сегментация изображения
7. Описание изображения
8. Распознавание образов

В общем случае стадия регистрации изображения включает некоторую предобработку, например масштабирование.

Улучшение изображения — это процесс манипулирования изображением, в результате которого оно становится более подходящим для конкретного применения, чем оригинал.

Восстановление изображений — это область, также связанная с повышением визуального качества изображения. Методы восстановления изображений опираются на математические или вероятностные модели искажений изображения.

Обработка цветных изображений приобрела особую важность в связи со значительным расширением использования цветных изображений в Интернете. Существует ряд фундаментальных понятий, относящихся к цветовым моделям и основным видам цифровых преобразований цветов. Цвет также используется как основа для выделения из изображения некоторых интересующих признаков.

Морфологическая обработка связана с инструментами для извлечения таких компонентов изображения, которые могут быть полезны для представления и описания формы.

Сегментация разделяет изображение на составные части или объекты. В целом автоматическая сегментация принадлежит к числу самых трудных задач цифровой обработки изображений. Излишне подробная сегментация направляет процесс на длинный путь решения задачи, требуя идентификации объектов по отдельности. Недостаточно подробная или же неверная сегментация почти неизбежно приведет к возникновению ошибок на финальной стадии обработки. Чем точнее сегментация, тем больше шансов на успех при распознавании

Построение описания, иначе называемое выбором признаков, связано с выделением атрибутов, которые бы выражали интересующую информацию в количественном виде или бы могли служить основой для различения классов объектов.

Распознавание представляет собой процесс, который присваивает некоторому объекту идентификатор на основании его описателей. Сфера цифровой обработки изображений заканчивается разработкой методов распознавания отдельных объектов.

## 2. Распознавание образов. Характерные признаки объектов.

Под *образом* подразумевается некоторая упорядоченная совокупность дескрипторов. В литературе по распознаванию образов эти дескрипторы часто называют *признаками*. *Классом образов* называется совокупность образов, обладающих некоторыми общими свойствами. Будем обозначать классы символами  $\omega_1, \omega_2, \dots, \omega_W$ , где  $W$  — число классов. Под машинным *распознаванием образов* понимаются методы, позволяющие относить образы к тем или иным классам — автоматически или с минимальным вмешательством человека.

Важная задача распознавания образов связана с выделением характерных признаков или свойств из полученных исходных данных и снижением размерности векторов образов. Эту задачу часто определяют как задачу предварительной обработки и выбора признаков. Признаки класса образов представляют собой характерные свойства, общие для всех образов данного класса. Признаки, характеризующие различия между отдельными классами, можно интерпретировать как межклассовые признаки. Внутриклассные признаки, общие для всех рассматриваемых классов, не несут полезной информации с точки зрения распознавания и могут не приниматься во внимание. Выбор признаков считается одной из важных задач в теории распознавания образов и при построении распознающих систем. Проблема выбора признаков проявляется в том, что в конкретном приложении степень разделимости классов сильно зависит от выбора дескрипторов

Для большинства практических задач процедура выделения различающих признаков является чрезвычайно трудной, однако ее можно использовать для упрощения автоматического распознавания.

В практических задачах получили распространение три формы упорядоченного представления признаков: в виде *векторов признаков* (для количественных дескрипторов), в виде *символьных строк*, а также в виде *деревьев* (строки и деревья применяются для структурных описаний). Образы, представленные векторами признаков, обозначаются жирными строчными буквами, например  $x, y, z$ , и имеют форму

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad (12.1-1)$$

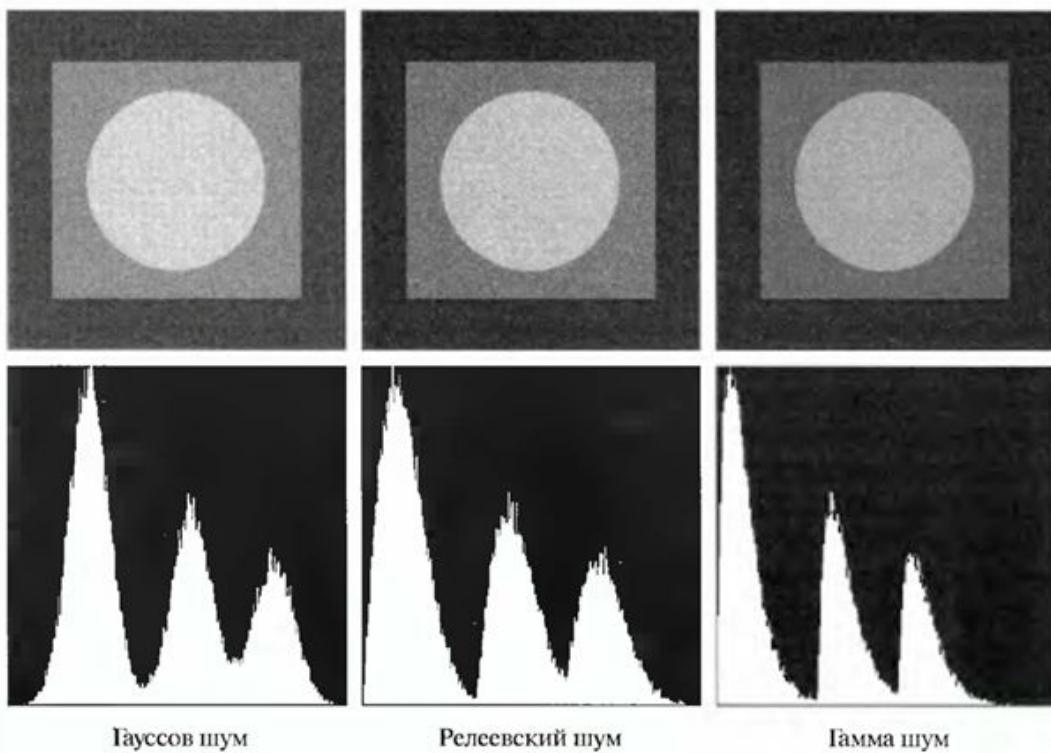
где каждая из компонент  $x_i$  представляет  $i$ -й дескриптор, а  $n$  — общее число дескрипторов, связанных с данным образом. Образы представляются вектор-столбцами (т. е. матрицами порядка  $n \times 1$ ) вида (12.1-1) или в эквивалентной форме  $x = (x_1, x_2, \dots, x_n)^T$ , где  $T$  — операция транспонирования. Мы уже использовали

## **БИЛЕТ № 9**

### ***1. Восстановление цифрового изображения. Модели шума.***

Основные источники шума на цифровом изображении — это сам процесс его получения (оцифровки), а также процесс передачи. Работа сенсоров зависит от различных факторов, таких как внешние условия в процессе видеосъемки и качество сенсоров. Например, в процессе получения изображения с помощью фотокамеры с ПЗС матрицей, основными факторами, влияющими на величину шума, являются уровень освещенности и температура сенсоров. В процессе передачи изображения могутискажаться помехами, возникающими в каналах связи.

Например, при передаче изображения с использованием беспроводной связи, оно может быть искажено в результате разряда молнии или других возмущений в атмосфере.



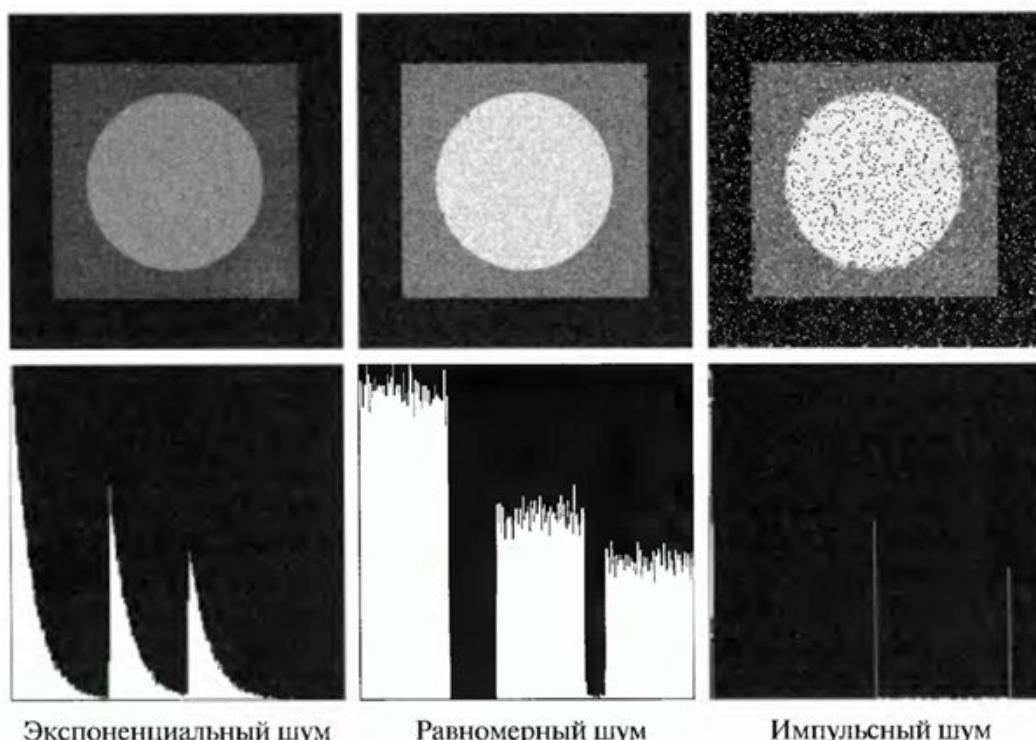
Гауссов шум

Релеевский шум

Гамма шум

**а б в**  
**г д е**

Рис. 5.4. Изображения и гистограммы, полученные в результате добавления гауссова, релеевского, гамма, экспоненциального, равномерного и импульсного шума к изображению на Рис. 5.3.



Экспоненциальный шум

Равномерный шум

Импульсный шум

**ж з и**  
**к л м**

Рис. 5.4 (продолжение). Изображения и гистограммы, полученные в результате добавления гауссова, релеевского, гамма, экспоненциального, равномерного и импульсного шума к изображению на Рис. 5.3.

## **2. Распознавание образов. Скрытые Марковские модели.**

Скрытая марковская модель (СММ) — статистическая модель, имитирующая работу процесса, похожего на марковский процесс с неизвестными параметрами, и задачей ставится разгадывание неизвестных параметров на основе наблюдаемых.

Пример, чтобы понять, как работает.

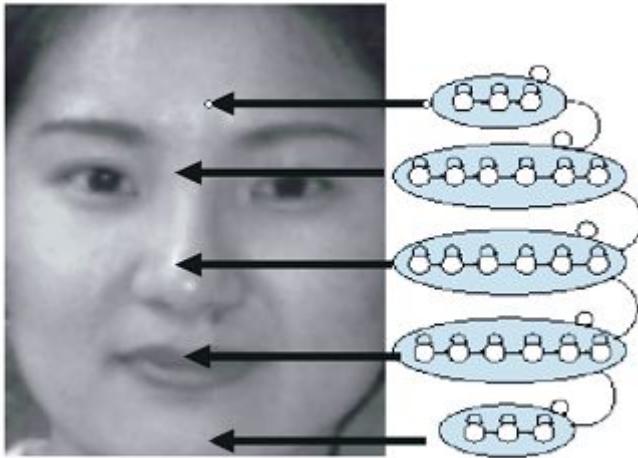
Представим двух друзей, обсуждающих каждый вечер по телефону, что они сегодня делали днём. Ваш друг может делать лишь три вещи: гулять в парке, ходить за покупками или убираться в комнате. Его выбор основывается лишь на погоде, которая была в момент принятия решения. Вы ничего не знаете о погоде в том регионе, где живёт ваш друг, но вы можете, основываясь на его решениях, попытаться угадать, какая погода была.

Погода представима в виде марковской цепи, она имеет два состояния: солнечно или дождливо, но вы не можете сами увидеть её, поэтому она скрыта от вас. Каждый день ваш друг принимает одно из трёх возможных решений: прогулка, покупки или уборка. Вы можете узнать о решении вашего друга, поэтому это наблюдаемое значение. В целом мы получаем СММ.

В обычной марковской модели состояние видимо наблюдателю, поэтому вероятности переходов — единственный параметр. В скрытой марковской модели мы можем следить лишь за переменными, на которые оказывает влияние данное состояние. Каждое состояние имеет вероятностное распределение среди всех возможных выходных значений. Поэтому последовательность символов, сгенерированная СММ, даёт информацию о последовательности состояний. Диаграмма, представленная ниже, показывает общую структуру СММ. Овалы представляют собой переменные со случайным значением. Случайная переменная  $x(t)$  представляет собой значение скрытой переменной в момент времени  $t$ . Случайная переменная  $y(t)$  — это значение наблюдаемой переменной в момент времени  $t$ . Стрелки на диаграмме символизируют условные зависимости.

Из диаграммы становится ясно, что значение скрытой переменной  $x(t)$  (в момент времени  $t$ ) зависит только от значения скрытой переменной  $x(t-1)$  (в момент  $t-1$ ). Это называется свойством Маркова. Хотя в то же время значение наблюдаемой переменной  $y(t)$  зависит только от значения скрытой переменной  $x(t)$  (обе в момент времени  $t$ ).

Каждая отдельная СММ представляет некоторый класс объектов. Тогда состояния модели определяют некоторые ключевые признаки класса. К примеру, СММ для поиска лиц состоит из 5 суперсостояний, соответствующих областям лица (лобовая часть, глаза, нос, рот, подбородок), каждое из которых делится на отдельные состояния.



Переход в следующее состояние возможно только после предыдущего, а переход в следующее суперсостояние возможно только после всех состояний текущего суперсостояния. Вероятность принадлежности некоторого объекта заданному классу, оценивается как вероятность генерации сигнала, соответствующего его вектору признаков.

Поиск изображений основан на поиске схожести в многомерном пространстве. При этом изображение определяется набором своих сигнатур. Мера схожести – это функция, которая вычисляет и возвращает значение, соответствующее схожести между двумя объектами согласно некоторым предопределенным критериям. Понятие меры сходства основывается на понятии метрики. Метрика – это функция расстояния  $d$ , определенная на метрическом множестве, для любых точек  $x, y, z$  которого выполняются условия:

1.  $d(x, y) = d(y, x)$ ,
2.  $d(x, y) = 0 \Leftrightarrow x = y$ ,
3.  $d(x, z) \leq d(x, y) + d(y, z)$ .

Индексирование и поиск изображений в коллекциях производится на основе треугольного дерева(triangle tree, trie дерево), также называемое Really Fixed Query Tree. Треугольное дерево – это структура данных, разработанная для поиска приблизительных совпадений в зависимости от усилий поиска. Оно ассоциируется с мерой расстояния, набором ключевых изображений и набором элементов базы данных. Форма trie – это форма дерева, в котором ребра следуют от корня до листьев, определяя индекс листа. Листья дерева содержат элементы базы. Каждое внутреннее ребро в дереве ассоциируется с неотрицательным числом. Каждый уровень дерева ассоциируется с одиночным ключом. Путь от корня до листьев представляет собой расстояние от элемента базы данных до каждого из ключей. Таким образом, сравнивая объекты из базы данных и объекты запроса с третьим ключевым объектом, может быть

получена нижняя граница дистанции между двумя объектами. На основе этой границы можно отсечь большую часть коллекции, как не удовлетворяющую запросу. Также необходимо отметить, что процесс поиска происходит в два этапа. На первом этапе определяется нижняя граница расстояния от изображения запроса до изображений коллекции. На этой стадии происходит отсечение части коллекции. На втором этапе происходит уточняющий поиск: определяются расстояния от изображения-запроса до каждого изображения, из найденных на первой стадии. Изображения, соответствующие максимальным расстояниям, составляют итоговый результат поиска.

## **БИЛЕТ № 10**

### **1. Связь зрительной системы человека с цифровой обработкой изображений.**

Большое количество информации представляется в виде статических или динамических изображений. Поскольку далее эти изображения рассматриваются и анализируются человеком, то важно знать механизмы зрительного восприятия. Эти знания являются мощным инструментом при разработке различных систем обработки изображений.

Отметим, что психофизические аспекты восприятия света изучены не полностью. Однако существует большое количество различных методов и подходов, объясняющих законы и свойства зрительного восприятия света человеком.

Глаз человека является уникальным механизмом, обеспечивающим адаптивную настройку в соответствии с внешними условиями.

Важной характеристикой зрительной системы является чувствительность, т.е. способность реагировать на внешние изменения. Чувствительность характеризуется верхним и нижним абсолютными порогами.

Существует несколько различных видов чувствительности. Световая чувствительность характеризует свойство глаза реагировать на максимально малый световой поток. Однако здесь следует отметить, что вероятность распознавания максимально малого светового потока зависит также и от других факторов, например угла зрения.

Зрительная система по-разному реагирует на излучения, которые равны по мощности, но излучаемые из различных диапазонов спектра. Такая чувствительность называется спектральной.

Способность глаза различать минимальные различия яркости смежных областей изображения характеризуется контрастной чувствительностью. Также зрительная система характеризуется различной чувствительностью к цветовому тону, т.е. к

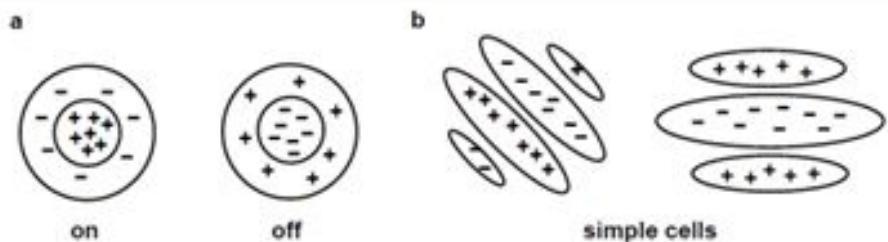
излучениям из различных участков спектра. Зрительная система характеризуется еще чувствительностью к насыщенности цвета.

Приведенные выше типы чувствительности зрительной системы не являются постоянными, а зависят от многих факторов, в частности, условий освещения. Например, при переходе из темной комнаты в светлую, нужно некоторое время для восстановления светочувствительности глаза. Этот процесс называется яркостной адаптацией глаза.

Цветоощущение характеризуется тремя основными характеристиками – светлота, цветовой тон и насыщенность. Для классификации цветов используются цветовые пространства.

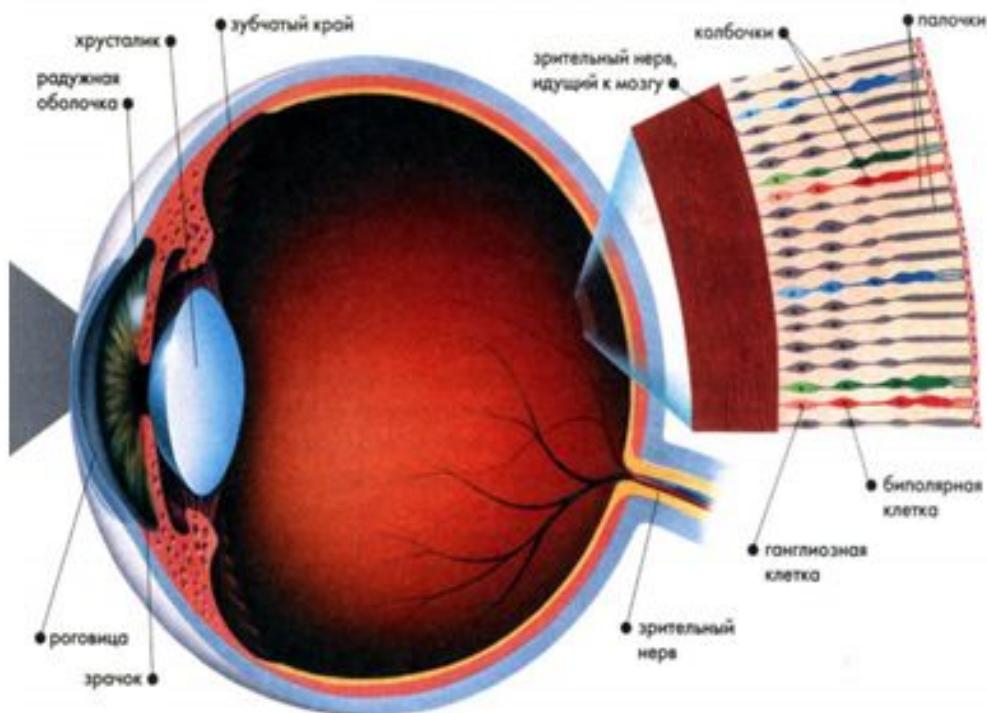
На основе свойств и характеристик зрительных систем создаются различные модели цветового зрения. Среди них следует выделить модель цветового зрения, предложенную Фреем. Особенностью этой модели является то, что зрительная система представлена тремя каналами, два из которых характеризуют цветность, а третий – яркость. Эта модель наиболее удачно согласуется со многими свойствами цветного зрения.

## Схема расположения чувствительных клеток



Typical classical receptive fields of neurons early in the visual pathway. Plus signs denote regions of the visual field where light causes excitation, minuses regions where light inhibits responses. **a)** Retinal ganglion and LGN neurons typically exhibit center-surround receptive field organization, in one of two arrangements. **b)** The majority of simple cells in V1, on the other hand, have oriented receptive fields.

### СТРОЕНИЕ ГЛАЗА



## 2. Распознавание образов. Глубокие нейронные сети.

Под *образом* подразумевается некоторая упорядоченная совокупность *дескрипторов*. В литературе по распознаванию образов эти дескрипторы часто называют *признаками*. *Классом образов* называется совокупность образов, обладающих некоторыми общими свойствами. Будем обозначать классы символами  $\omega_1, \omega_2, \dots, \omega_W$ , где  $W$  — число классов. Под *машинным распознаванием образов* понимаются методы, позволяющие относить образы к тем или иным классам — автоматически или с минимальным вмешательством человека.

**Глубинное обучение (deep learning)** – форма машинного обучения, которая предусматривает извлечение, или моделирование, признаков данных с использованием сложных многослойных фильтров. Поскольку глубинное обучение является весьма общим способом моделирования, оно способно решать сложные задачи, такие как компьютерное зрение и обработка естественного языка. Этот подход существенно отличен и от традиционного программирования, и от других методов машинного обучения.

Глубинное обучение не только может дать результат там, где другие методы не сработают, но и позволяет построить более точную модель или же сократить время на ее создание; расплачиваться же за это приходится еще большими вычислительными мощностями. Еще одним недостатком глубинного обучения является сложность интерпретации получаемых моделей.

Определяющая характеристика глубинного обучения — наличие более одного слоя между входом и выходом. Обычно, говоря о глубинном обучении, подразумевают использование глубоких нейронных сетей. Есть, однако, несколько алгоритмов, которые реализуют глубинное обучение, используя иные типы слоев.

Обучение в глубоких нейронных сетях происходит путем закрепления связи между двумя нейронами, когда оба они одновременно активны во время обучения. В современном ПО нейросетей это чаще всего реализуется путем увеличения значений веса связей между нейронами, используя правило, называемое обратным распространением ошибки.

### **Реальные глубокие нейронные сети**

Глубокая нейронная сеть для реальной задачи может иметь более десяти скрытых слоев. Ее топология может быть простой или довольно сложной.

Чем больше слоев в сети, тем больше характеристик она может распознать. Однако, чем больше слоев в сети, тем больше времени потребуется для расчета, и тем сложнее будет обучение.

### **Алгоритмы глубинного обучения**

Как говорилось выше, «наиболее глубинное» обучение осуществляется с помощью глубоких нейронных сетей. Сверточные нейронные сети (CNN) часто используются для компьютерного зрения. Рекуррентные нейронные сети (RNN) часто используются в задачах естественного языка и для обработки других последовательностей, как и сети с долгой краткосрочной памятью (Long short-term memory; LSTM) и нейронные сети с механизмом внимания. Случайные леса (они же — леса случайных решений), нейронными сетями не являющиеся, полезны для целого ряда задач классификации и регрессии.

## **БИЛЕТ № 11**

### **1. Описание цифрового изображения. Дескрипторы областей**

**Дескрипторы областей:**

A) простые дескрипторы

**Площадь** области определяется как число содержащихся в ней пикселей. **Периметр** области есть длина ее границы. Эти дескрипторы используются при вычислении **меры компактности** области, которая определяется как отношение квадрата периметра к площади. **Коэффициент округлости** определяется как отношение площади области к площади круга.

К числу других простых дескрипторов, применяемых для описания областей, относятся различные яркостные характеристики, такие как минимальное, максимальное и среднее значения, медиана яркостей элементов области, а также число пикселей со значениями яркости больше и меньше среднего значения.

B) топологические дескрипторы

В общих чертах **топология** изучает свойства фигур, на которые не влияют любые их непрерывные деформации, то есть такие, при которых не происходит разрывов и склеек (как будто плоскость изображения ведет себя аналогично листу резины). Если в качестве топологического дескриптора использовать число дырок внутри области, то это свойство, очевидно, будет инвариантным относительно растяжения или поворота. Однако, вообще говоря, число дырок будет меняться, если область складывается или разрывается.

Другое полезное для описания области топологическое свойство — это число ее связных компонент (**компонент связности**).

B) текстурные дескрипторы

Одним из важных подходов к описанию областей является количественное представление их **текстурных признаков**. Интуитивно ясно, что этот дескриптор является мерой таких свойств области, как гладкость, шероховатость и регулярность. В цифровой обработке изображений для описания текстуры области применяются три основных подхода: статистический, структурный и спектральный. **Статистические методы** позволяют охарактеризовать текстуру области как гладкую, грубую, зернистую и т. д. **Структурные методы** занимаются изучением взаимного положения простейших составляющих изображения, как например при описании текстуры из параллельных линий, проходящих с постоянным шагом. **Спектральные методы** основаны на свойствах Фурье-спектра и используются прежде всего для обнаружения глобальной периодичности в изображении по имеющим большую энергию узким выбросам на спектре.

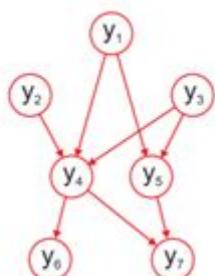
## 2. Распознавание образов. Байесовы сети.

Под *образом* подразумевается некоторая упорядоченная совокупность дескрипторов. В литературе по распознаванию образов эти дескрипторы часто называют *признаками*. *Классом образов* называется совокупность образов, обладающих некоторыми общими свойствами. Будем обозначать классы символами  $\omega_1, \omega_2, \dots, \omega_W$ , где  $W$  — число классов. Под машинным *распознаванием образов* понимаются методы, позволяющие относить образы к тем или иным классам — автоматически или с минимальным вмешательством человека.

**Байесовский классификатор** — классификатор, минимизирующий суммарную величину средних потерь на множестве всех решений. Байесовский классификатор относит объект  $x$  к классу  $C_i$ , если  $r_i(x) < r_j(x)$  ( $j \neq i$ ,  $r$  — условные средние потери). Принимается, что при правильном выборе класса величина потерь = 0, во всех остальных случаях = 1.

**Байесова сеть** — направленный, нециклический граф; модель, представляющая собой множество переменных и их вероятностных зависимостей по Байесу. В вершинах могут быть, например, гипотезы о принадлежности объектов к классам.

Пример:



Совместное распределение системы переменных задается выражением

$$p(Y) = p(y_1, y_2, y_3, y_4, y_5, y_6, y_7) = \\ p(y_1)p(y_2)p(y_3)p(y_4|y_1, y_2, y_3)p(y_5|y_1, y_3)p(y_6|y_4)p(y_7|y_4, y_5).$$

При использовании байесовского классификатора берут:

предполагаемые аналитические выражения общего вида для различных функций плотности распределения вероятностей.

Оценивают:

их параметры по выборке образов из каждого класса.

Байесовское решающее правило:  $x \in C_1$ , если  $\Lambda(x) > \eta$ .

$\Lambda(x) = p(x|H_1)/p(x|H_2)$  — отношение правдоподобия,  $\eta = \frac{P(H_2)(L_{21}-L_{22})}{P(H_1)(L_{12}-L_{11})}$ ,  $H$  — гипотезы,  $L$  — условные потери.

Пример использования: классификация данных дистанционного зондирования.

Сама формула Байеса (теорема гипотез):

$$p_i = p(H_i|x) = \frac{P(H_i)p(x|H_i)}{p(x)}$$

## **БИЛЕТ № 12**

### **1. Характеристики света в отношении формирования изображения.**

**Визуальное восприятие** — это результат взаимодействия визуального раздражителя с уже имеющимися в мозге наблюдателя знаний, целей и ожиданий. Нашему мозгу необходимо привести в соответствие то, что увидел глаз, с тем образцами, которые уже есть в нашей памяти.

Факторами, влияющими на визуальное восприятие человека, являются:

- *интенсивность сигнала* от наблюдаемого объекта (светлое и громкое воспринимается быстрее);
- *подвижность объекта* (движущиеся объекты воспринимаются в большей степени, чем неподвижные);
- *форма объекта* (необычные формы воспринимаются лучше);
- *размер объекта* (большие объекты воспринимаются проще, чем маленькие);
- *цвет объекта* (яркие цвета воспринимаются лучше, чем серые).

В первую очередь человеческий глаз замечает движение, затем форму, цвет, форму и контраст.

### **2. Распознавание образов. Нечеткие конечные автоматы.**

Под *образом* подразумевается некоторая упорядоченная совокупность дескрипторов. В литературе по распознаванию образов эти дескрипторы часто называют *признаками*. *Классом образов* называется совокупность образов, обладающих некоторыми общими свойствами. Будем обозначать классы символами  $\omega_1, \omega_2, \dots, \omega_W$ , где  $W$  — число классов. Под машинным *распознаванием образов* понимаются методы, позволяющие относить образы к тем или иным классам — автоматически или с минимальным вмешательством человека.

#### ***Автоматы как распознаватели строк символов***

До сих пор демонстрировалось, как формальные грамматики порождают образы. Далее мы рассмотрим задачу распознавания, которая состоит в выяснении, принадлежит ли некоторый образ (предложение) языку  $L(G)$ , порождаемому грамматикой  $G$ . Основные принципы, лежащие в основе синтаксических методов распознавания, можно проиллюстрировать на математических моделях вычислительных машин; такие модели принято называть *автоматами*. На основании поступившего на вход образа в виде строки символов автомат способен распознать, принадлежит ли этот образ тому языку, который соответствует данному автомату. Здесь мы будем рассматривать только *конечные автоматы*, которые являются распознавателями языков, порождаемых регулярными грамматиками.



По определению, *конечный автомат* — это пятерка

$$A_f = (Q, \Sigma, \delta, q_0, F), \quad (12.3-7)$$

где  $Q$  — конечное непустое множество *состояний автомата*,  $\Sigma$  — конечный входной алфавит,  $\delta$  — отображение множества  $Q \times \Sigma$  (множества упорядоченных пар, составленных из элементов множеств  $Q$  и  $\Sigma$ ) в множество всех подмножеств  $Q$ , тем самым  $\delta$  есть функция, определяющая (в зависимости от текущего символа входной строки) следующее состояние автомата,  $q_0$  — *начальное состояние*, а  $F$  (подмножество  $Q$ ) — *множество заключительных состояний*.

**Пример 12.10: Простой автомат.**

■ Рассмотрим автомат, определенный согласно (12.3-7), где  $Q = \{q_0, q_1, q_2\}$ ,  $\Sigma = \{a, b\}$ ,  $F = \{q_0\}$ , а отображение  $\delta$  действует по правилам  $\delta(q_0, a) = \{q_2\}$ ,  $\delta(q_0, b) = \{q_1\}$ ,  $\delta(q_1, a) = \{q_2\}$ ,  $\delta(q_1, b) = \{q_0\}$ ,  $\delta(q_2, a) = \{q_0\}$  и  $\delta(q_2, b) = \{q_1\}$ . Если, например, этот автомат находится в состоянии  $q_0$ , и на вход поступает символ  $a$ , то автомат переходит в состояние  $q_2$ . Аналогично, если затем поступает символ  $b$ , то автомат переходит в состояние  $q_1$ , и т.д. В данном случае начальное и заключительное состояния совпадают. ■

На Рис. 12.27 показана *диаграмма состояний* рассмотренного автомата. Она состоит из вершин, отвечающих каждому состоянию, и направленных дуг, показывающих возможные переходы между состояниями. Заключительное состояние показано двойным кругом, а каждая

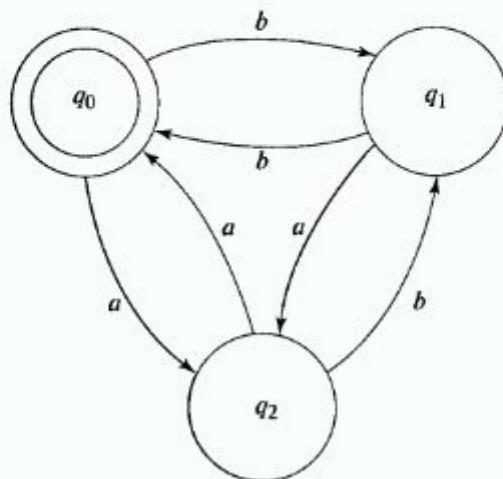


Рис. 12.27. Конечный автомат.

дуга помечена символом входного алфавита, при поступлении которого выполняется переход в состояние, куда направлена эта дуга. Как уже сказано, в данном случае начальное и заключительное состояния совпадают. Говорят, что автомат *допускает* или *распознает* строку терминальных символов  $w$ , если, начав с состояния  $q_0$  и по очереди подавая на вход символы  $w$  слева направо, после обработки последнего символа  $w$  автомат оказывается в заключительном состоянии. Например, показанный на Рис. 12.27 автомат распознает строку  $w = abbabb$ , но отвергает строку  $w = aabab$ .

Между регулярными грамматиками и конечными автоматами имеется взаимно-однозначное соответствие. Это значит, что язык распознается конечным автоматом в том и только в том случае, если он порожден регулярной грамматикой. Построение синтаксического распознавателя строк символов, работающего на основе изложенных принципов, состоит в формальной процедуре получения конечного автомата из данной регулярной грамматики. Пусть грамматика обозначена  $G = (N, \Sigma, P, X_0)$ , где  $X_0 \equiv S$ , и предположим, что множество  $N$  состоит из начального символа  $X_0$  и  $n$  добавочных нетерминальных символов  $X_1, X_2, \dots, X_n$ . Множество состояний  $Q$  конструируемого автомата строится из  $n+2$  состояний  $\{q_0, q_1, \dots, q_n, q_{n+1}\}$ , таких, что  $q_i$  соответствует  $X_i$  для  $0 \leq i \leq n$ , а  $q_{n+1}$  — заключительное состояние. Входной алфавит совпадает с множеством терминальных символов грамматики  $G$ . Отображение  $\delta$  строится из продукции  $G$  с помощью двух нижеследующих правил; а именно, для любых  $i$  и  $j$  таких, что  $0 \leq i \leq n$  и  $0 \leq j \leq n$ ,

1. Если в множестве  $P$  имеется правило  $X_i \rightarrow aX_j$ , то в множество  $\delta(q_i, a)$  включается  $q_j$ .
2. Если в множестве  $P$  имеется правило  $X_i \rightarrow a$ , то в множество  $\delta(q_i, a)$  включается  $q_{n+1}$ .

Наоборот, если задан конечный автомат  $A_f = (Q, \Sigma, \delta, q_0, F)$ , то соответствующая ему регулярная грамматика  $G = (N, \Sigma, P, X_0)$  строится следующим образом: множество  $N$  составляется из элементов  $Q$ , в качестве начального символа  $X_0$  принимается  $q_0$ , а продукции грамматики  $G$  получаются по правилам:

1. Если  $q_j \in \delta(q_i, a)$ , то в множество  $P$  включается правило  $X_i \rightarrow aX_j$ .
2. Если в  $\delta(q_i, a)$  входит состояние из  $F$ , то в множество  $P$  включается правило  $X_i \rightarrow a$ .

Множество терминальных символов  $\Sigma$  одно и то же в обоих случаях.

## БИЛЕТ № 13

### 1. Описание цифрового изображения. Дескрипторы границ.

Одним из простейших дескрипторов границы является ее *длина*. Общее число пикселей границы является грубым приближением ее длины. Для кривой, представленной цепным каком с единичными шагами дискретизации по обоим направлениям, сумма числа вертикальных, горизонтальных и умноженных на  $\sqrt{2}$  диагональных составляющих, дает точное значение длины границы.

*Диаметр* границы В определяется соотношением

$$\text{Diam}(B) = \max_{i,j} [D(p_i, p_j)],$$

где D — мера расстояния, а  $p_i$  и  $p_j$  точки границы.

Полезными дескрипторами границы являются значение ее диаметра и направление отрезка, соединяющего две экстремальные точки, которые определяют диаметр (этот отрезок называется *большой осью* границы). *Малая ось* границы определяется как отрезок, перпендикулярный большой оси и имеющий такую (минимальную) длину, что проведенный через концы обеих осей прямоугольник со сторонами, параллельными этим осям, полностью содержит в себе всю границу. Упомянутый прямоугольник называется *базовым прямоугольником*, а отношение длины большой оси к длине малой — *эксцентризитетом границы*, величина которого также является полезным дескриптором.

*Кривизна* определяется как скорость изменения угла наклона. В общем случае трудно надежно измерить кривизну в некоторой точке дискретной границы, потому что обычно на таких границах имеются локальные зазубрины. Тем не менее, часто оказывается полезным использовать разность углов наклона соседних сегментов границы (которые приближены отрезками ломаной) в качестве дескриптора кривизны границы в точке пересечения этих отрезков.

### Фурье-дескрипторы

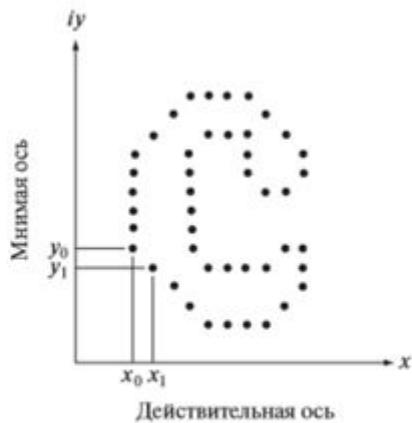
для  $k = 0, 1, 2, \dots, K-1$ . Таким образом,  $x$  и  $y$  рассматриваются как действительная и мнимая оси для последовательности комплексных чисел. Несмотря на изменившийся способ интерпретации этой последовательности, сущность границы осталась прежней. Конечно, такое представление имеет одно крупное преимущество: оно позволяет свести двумерную задачу к одномерной.

Согласно формуле (4.4-6), дискретное преобразование Фурье (ДПФ) конечной последовательности  $s(k)$  имеет вид

$$a(u) = \sum_{k=0}^{K-1} s(k) e^{-j2\pi uk/K} \quad (11.2-3)$$

для  $u = 0, 1, 2, \dots, K-1$ . Комплексные коэффициенты  $a(u)$  называются *Фурье-дескрипторами* границы. Обратное преобразование Фурье, примененное к этим коэффициентам, позволяет восстановить границу  $s(k)$ . То есть, согласно формуле (4.4-7),

$$s(k) = \frac{1}{K} \sum_{u=0}^{K-1} a(u) e^{j2\pi uk/K} \quad (11.2-4)$$



**Рис. 11.19.** Дискретная граница и ее представление в виде комплексной последовательности. Отмеченные точки  $(x_0, y_0)$  и  $(x_1, y_1)$  являются первыми двумя точками последовательности (начальная точка выбрана произвольно)

## 2. Распознавание жестов. Основные системы и математические модели распознавания.

Жест – это форма неверbalного общения, при которой некоторое действие или движение человеческого тела или его части передает информацию. Жест может содержать движение как одной, так и нескольких частей тела и нести эмоциональную нагрузку, что делает его очень информативным средством общения и взаимодействия.

## Система распознавания жестов.

Упрощенная структура системы распознавания жестов приведена на схеме 1. Она состоит из 3 блоков:

- Устройство ввода
- Блок распознавания
- Реагирующий блок



Схема 1. Структура системы распознавания жестов

Устройство ввода является точкой входа в систему. Оно преобразовывает жест в цифровую форму и передает его в блок распознавания, который сравнивает поступившую информацию с шаблонами, хранящимися в базе данных, и в качестве результата возвращает распознанный жест. В итоге реагирующий блок выполняет некоторые функции в соответствии с распознанными жестами.

### Популярные методы

По мере развития технологии распознавания жестов определились наиболее популярные методы:

- Скелетный метод
- 3D-распознавание
- 2D-распознавание

Скелетный метод использует математическую модель человеческой руки. Преимуществом этого метода является скорость работы, так как он использует высоко детализированную модель, которая описана с помощью длин сегментов и углов между ними. Примеры скелета человеческой ладони приведены на рисунках 1 и 2с.

Следующий метод использует 3D-модель для идентификации жеста (рис. 2а, 2б). Такие модели могут быть представлены в виде сложных трехмерных поверхностей и классифицироваться с помощью нейронных сетей. Недостатком этого метода является его ресурсосёмкость. Построение модели, обучение нейронной сети и ее использование могут потребовать значительных ресурсов.

Метод 2D-распознавания схож с предыдущим методом, но оперирует двумерными изображениями вместо объемных моделей (рис. 2д, 2е). Как следствие, снижается вычислительная сложность и отпадает необходимость в специальном оборудовании, так как для получения изображений может быть использована обычная веб-камера. Основным недостатком этого метода является низкая точность.

## **БИЛЕТ № 14**

### **1. Способы захвата (регистрации) цифровых изображений.**

В зависимости от способа их получения, принятой модели и структуры данных:

**Аналоговая форма** изображения предполагает его получение каким-либо образом на физическом носителе - на бумаге, фотобумаге, фотопленке и др. и используется с незапамятных времен. Во многих случаях такая форма является основной, особенно там, где важна юридическая значимость изображения либо если оно необходимо для использования в нестационарных условиях.

**Цифровая форма** изображения возникла в связи с потребностью его представления в памяти электронных вычислительных машин. В настоящее время она является одной из основных и связана не только с хранением, но и получением и обработкой изображения.

**Векторная форма** цифрового изображения находит применение в цифровой картографии; в фотограмметрии векторная форма цифрового изображения используется при создании цифрового оригинала. Элементы цифрового векторного изображения представляются в виде набора примитивов и их комбинаций - точек, векторов, граней, ребер и т. п., причем, положение точек задано в некоторой координатной системе, выбираемой пользователем в зависимости от характера решаемых задач. Внутренние связи элементов изображения соответствуют определенной структуре, выбор которой зависит от используемого принципа формирования и описания его элементов, способа доступа к ним, характера связи с окружающими и др.

Получение: рентгенография, узи, радиоизотопное излучение, Использование эффекта ядерного магнитного резонанса. При получении как цифровой, так и пленочной фотографии свет с фотографируемой сцены попадает на линзу и затем регистрируется либо на пленку, либо на сенсорную матрицу цифрового фотоаппарата

### **2. Распознавание образов. Параллельный захват и отслеживание динамических жестов руки.**

Под *образом* подразумевается некоторая упорядоченная совокупность *дескрипторов*. В литературе по распознаванию образов эти дескрипторы часто называют *признаками*. *Классом образов* называется совокупность образов, обладающих некоторыми общими свойствами. Будем обозначать классы символами  $\omega_1, \omega_2, \dots, \omega_W$ , где  $W$  — число классов. Под машинным *распознаванием образов* понимаются методы, позволяющие относить образы к тем или иным классам — автоматически или с минимальным вмешательством человека.

В зависимости от способа выполнения жесты были разделены на статические и динамические жесты. Статические жесты выполняются заданием определенного положения кисти и пальцев в пространстве, вне зависимости от времени.

Динамические жесты выполняются движением какой-либо части человеческого тела, чаще всего кистью, во времени и пространстве.

Описаны два алгоритма: алгоритм захвата и отслеживания областей интересов на сложном фоне, основанный на последовательном выделении объектов по перемещению, цвету и кластерам и алгоритм, основанный на параллельном каскадном детекторе, с использованием характерных признаков Хаара.

Основная идея первого алгоритма заключается в последовательной обработке и кластерном разбиении области интересов. Рассмотрим данный алгоритм по шагам. Кадр, получаемый видеокамерой в момент времени  $t$  и имеющий по горизонтали  $V$ , а по вертикали  $W$  пикселей, обозначим  $I_t(V, W)$ . Под областью интересов  $Ob_t(X, Y)$  понимается множество пикселей кадра  $I_t(V, W)$ , очерчивающих искомый объект. Захватом области интересов называется выделение ее в кадре в момент времени  $t$ . Отслеживанием области интересов называется процесс последовательного захвата в кадрах  $I_t(V, W), I_{t+1}(V, W), \dots, I_{t+k}(V, W)$  областей интереса  $Ob_t(X_t, Y_t), Ob_{t+1}(X_{t+1}, Y_{t+1}), \dots, Ob_{t+k}(X_{t+k}, Y_{t+k})$ .

*Шаг 1.* На первом шаге алгоритма кадр  $I_t(V, W)$  поступает от Web-камеры с разрешением  $640 \times 480 / 320 \times 240$ , 8 бит.

*Шаг 2.* Второй шаг алгоритма – это фильтрация входного кадра  $I_t(V, W)$ . Для выполнения фильтрации было отдано предпочтение медианному фильтру перед фильтром Гаусса. Так как, при применении фильтра Гаусса в кадре получалось некоторое “размытие” областей изображения, шум которых описывался распределением с нулевым математическим ожиданием.

*Шаг 3.* На третьем шаге алгоритма, используя кадры  $I_t(V, W)$  и  $I_{t+1}(V, W)$  и применяя алгоритм, основанный на вычитании соседних кадров, находится перемещающийся объект и осуществляется захват области интересов  $Ob_{t+1}(X_{t+1}, Y_{t+1})$ . Для выделения перемещающегося объекта используется фактор изменения яркости пикселей, относящихся к перемещающемуся объекту, в последовательности двух смежных  $I_t(V, W)$  и  $I_{t+1}(V, W)$  кадров. Если разность яркостей пикселей превышает заданный порог, то этот пиксель кадра  $I_{t+1}(V, W)$  считается принадлежащим перемещающемуся объекту  $Ob_{t+1}(X, Y)$ . В данном случае порог найден экспериментально и равен 20. Для получения явных значений яркости пикселя, каждый кадр переводился из цветового пространства  $RGB$  в полутоновое цветовое пространство.

*Шаг 4.* Чтобы выделить в полученной области интересов  $Ob_{t+1}(X, Y)$  только изображение кисти, на четвертом шаге ищутся пиксели, значение цвета которых совпадает со значением цвета кожи человека. В цветовом пространстве  $HSV$  (параметры  $H, S, V$  соответственно обозначают *Hue* (тон), *Saturation* (насыщенность), *Volume* (яркость)) для цвета кожи значения параметра  $H$  лежат в промежутке от 18 до 22, параметра  $S$  от 5 до 10.

*Шаг 5.* На пятом шаге алгоритма наложением четырехсвязной маски, отсекаются одиночные пиксели. Эти пиксели считаются шумом. Этих шагов достаточно, чтобы захватить и отследить только одну кисть человека

находящуюся в кадре. Но, чтобы найти правую, левую кисть и лицо пользователя управляющего системой издалека, необходимо применить алгоритм кластеризации *c*-средних.

*Шаг 6.* На шестом шаге, алгоритм кластеризации разбивает пиксели найденной области интересов  $Ob_{i+1}(X_{i+1}, Y_{i+1})$ , принадлежащие перемещающемуся объекту и распределению цвета кожи, на три кластера, соответственно: правая кисть, левая кисть, лицо (рис. 2).

*Шаг 7.* На заключительном, седьмом шаге алгоритма для каждого кластера вычисляется центр тяжести. Путем отслеживания этих центров, строятся траектории перемещения кистей, с помощью которых распознается выбранный алфавит жестов для каждой руки.



Рис. 2. (а) Входной кадр, (б) Найдены пиксели области интересов  $Ob_{i+1}(X_{i+1}, Y_{i+1})$ , принадлежащие перемещающемуся объекту и распределению цвета кожи человека,  
(в) Найдены и маркированы три кластера: правая кисть, левая кисть, лицо

В конце главы описан второй разработанный алгоритм захвата и отслеживания областей интересов, основанный на параллельном каскадном детекторе, с использованием характерных признаков Хаара, который способен, как и рассмотренный алгоритм, отслеживать кисти и лицо человека в видеопотоке.

Параллельный каскадный детектор состоит из трех параллельных каскадов. Каждый каскад это цепь классификаторов, основанных на характерных признаках Хаара. Каскады, состоящие из пятнадцати классификаторов, предназначены для захвата и отслеживания правой и левой кистей человека, каскад из двадцати восьми классификаторов предназначен для захвата и отслеживания лица человека. Кроме своей структуры, оригинальной особенностью параллельного каскадного детектора, является обучение его на специально сформированной обучающей выборке. Выборка состояла из изображений кисти, захваченных при разных условиях освещения, что позволило в дальнейшем повысить устойчивость работы детектора в неконтролируемых условиях.

Найденная верхняя оценка вычислительной сложности разработанных алгоритмов составила  $O(N)$ , где  $N$  – количество пикселей кадра  $I(V, W)$ .

## БИЛЕТ № 15

### 1. Описание цифрового изображения: коды, линии, границы, области.

#### A) прослеживание границы

Пусть задана двоичная область (или ее граница)  $R$ . Алгоритм прослеживания границы  $R$  состоит из следующих шагов.

1. В качестве начальной точки  $b_0$  выбирается *самая левая верхняя*<sup>1</sup> точка изображения, имеющая значение 1. Обозначим  $c_0$  левого соседа  $b_0$ , как показано на рис. 11.1(б); ясно, что  $c_0$  всегда является точкой фона. Рассмотрим восьмерку соседей  $b_0$ , начиная с  $c_0$  и двигаясь по часовой стрелке. Пусть  $b_1$  — *первая* встретившаяся точка со значением 1, а  $c_1$  — точка фона, непосредственно ей предшествующая в указанном порядке обхода. Запоминаем положения точек  $b_0$  и  $b_1$  для использования на шаге 5.
2. Пусть  $b = b_1$  и  $c = c_1$  (см. рис. 11.1(в)).
3. Начиная с точки  $c$  и двигаясь по часовой стрелке, обозначим восьмерку соседей точки  $b$  через  $n_1, n_2, \dots, n_8$ . Находим первую точку  $n_k$ , имеющую значение 1.
4. Кладем  $b = n_k$  и  $c = n_{k-1}$ .
5. Повторяем шаги 3 и 4 до тех пор, пока не получим, что  $b = b_0$  и следующая найденная точка границы —  $b_1$ . Последовательность найденных точек  $b$  в момент остановки алгоритма является множеством упорядоченных точек границы.

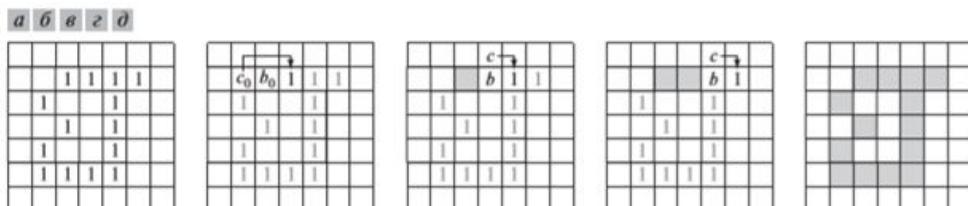


Рис. 11.1. Иллюстрация нескольких первых шагов алгоритма прослеживания границы. Точка, которая будет обрабатываться следующей, выделена жирным шрифтом, а уже найденные точки отмечены серыми квадратами

#### Б) цепные коды

С помощью цепных кодов граница представляется в виде последовательности соединенных отрезков, для которых указаны длина и направление. Как правило, такое представление основывается на отрезках с 4- или 8-связностью. Направление каждого отрезка кодируется числом в соответствии со схемой нумерации, изображенной на рис. 11.3. Границу, закодированную с помощью последовательности так занумерованных направлений, называют *цепным кодом Фримена*.

Как правило, регистрация и обработка цифровых изображений осуществляется на квадратной решетке, имеющей одинаковые шаги дискретизации в направлениях осей  $x$  и  $y$ , поэтому цепной код можно построить путем прослеживания границы, скажем, по часовой стрелке, присваивая отрезкам, соединяющим каждую пару пикселей, номера соответствующих направлений. В общем случае такой метод является неприемлемым по двум главным причинам: (1) получаемая цепочка кодов оказывается слишком длинной, и (2) любые малые возмущения вдоль границы области, вызванные наличием шума или несовершенством алгоритма сегментации, приводят к изменениям в кодовой последовательности, которые не связаны с общей формой границы.

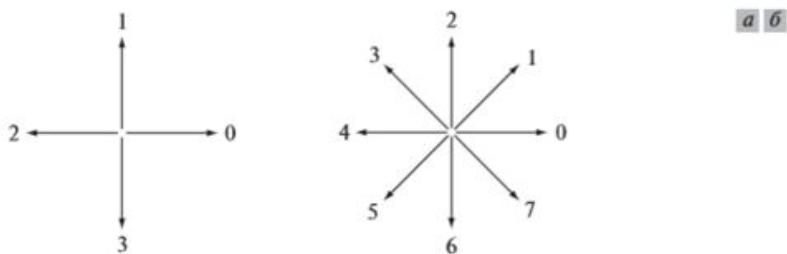


Рис. 11.3. Нумерация направлений для (a) 4-связного цепного кода и (б) 8-связного цепного кода

Чтобы обойти указанные проблемы, часто применяется подход, использующий вторичную дискретизацию границы по сетке с увеличенным шагом, как показано на рис. 11.4(a). После этого в процессе обхода границы строятся отрезки, соединяющие узлы укрупненной сетки, выбираемые по признаку их близости к первоначальной границе, что иллюстрируется на рис. 11.4(б). Полученная таким способом граница с пониженным разрешением может затем представляться с помощью 4- или 8-связного цепного кода. На рис. 11.4(в) показаны точки огрубленной границы с представлением в виде 8-связного цепного кода, который несложно преобразовать в 4-связный (задача 2.12). В качестве начальной точки на рис. 11.4(в) произвольно выбран левый верхний угол границы,

что приводит к цепному коду 0766...12. Как и следовало ожидать, точность представления границы полученным кодом определяется шагом дискретизации новой сетки.

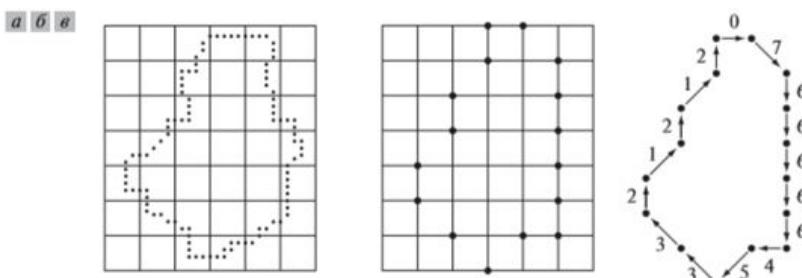


Рис. 11.4. (а) Дискретная граница с наложенной укрупненной сеткой дискретизации. (б) Результат новой дискретизации. (в) Представление границы 8-связным цепным кодом

## B) Аппроксимация ломаной линией минимальной длины

Сформированное множество элементов разбиения, содержащее дискретную границу (объекта), называется **клеточным комплексом**. В силу выбранного способа построения мы рассматриваем только границы без самопересечений и только связные клеточные комплексы. Исходя из такого предположения и присваивая выпуклым вершинам белый цвет (W), а зеркальным к вогнутым вершинам — черный (B), можно сформулировать следующие утверждения.

1. ЛМД, ограниченная связным клеточным комплексом, является не самопересекающейся.
2. Каждая выпуклая вершина ЛМД является белой, но не каждая белая вершина границы является вершиной ЛМД.
3. Каждая вогнутая вершина ЛМД является черной, но не каждая черная вершина границы является вершиной ЛМД.
4. Все черные вершины расположены на ЛМД или снаружи от нее, а все белые вершины расположены на ЛМД или внутри нее.
5. Левой верхней вершиной в последовательности вершин, заключенных внутри клеточного комплекса, всегда является белая вершина ЛМД.

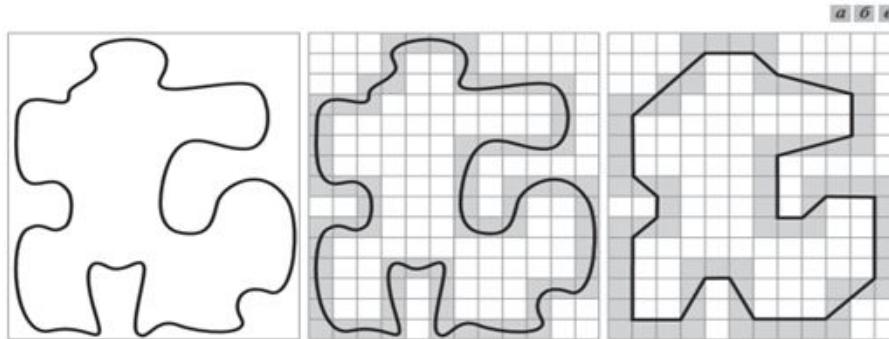


Рис. 11.6. (a) Граница объекта (черная кривая). (б) Граница объекта, заключенная внутри цепочки элементов (выделены темным цветом). (в) Ломаная минимальной длины, полученная в результате стягивания границы. Вершины ломаной совпадают с углами внутренней или внешней границы выделенной области

## 2. Распознавание образов. Нейронные сети.

Под *образом* подразумевается некоторая упорядоченная совокупность *дескрипторов*. В литературе по распознаванию образов эти дескрипторы часто называют *признаками*. *Классом образов* называется совокупность образов, обладающих некоторыми общими свойствами. Будем обозначать классы символами  $\omega_1, \omega_2, \dots, \omega_W$ , где  $W$  — число классов. Под машинным *распознаванием образов* понимаются методы, позволяющие относить образы к тем или иным классам — автоматически или с минимальным вмешательством человека.

Нейронные сети позволяют решать задачу определения класса объекта. В процессе обучения сети ей предлагаются различные эталонные образцы с меткой того, к какому классу объекта их можно отнести. Эти образцы должны быть однозначны и непротиворечивы для того, чтобы нейронная сеть выдавала высокие вероятности принадлежности объекта к одному конкретному классу.

**Перцептрон** — математическая или компьютерная модель восприятия информации мозгом (кибернетическая модель мозга). Перцептрон стал одной из первых моделей нейросетей. Перцептрон состоит из трёх типов элементов, именно: поступающие от *сенсорных* элементов сигналы передаются *ассоциативным* элементам, а затем *реагирующими* элементам.

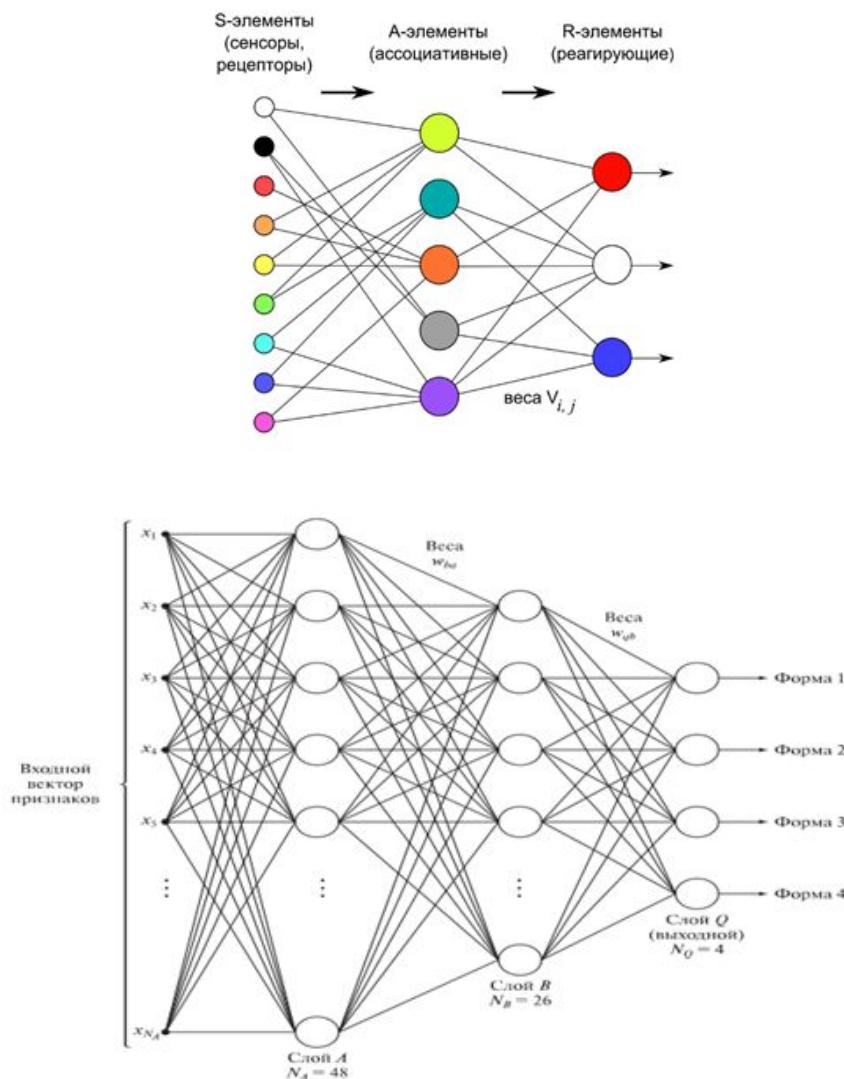


Рис. 12.19. Трехслойная нейронная сеть, использованная для распознавания фигур на рис. 12.18. (Изображение предоставлено д-ром Лалит Гупта, кафедра ECE, Южный университет шт. Иллинойс)

## БИЛЕТ № 16

### 1. Равномерная пространственная дискретизация и квантование изображения.

**ПРОСТРАНСТВЕННАЯ ДИСКРЕТИЗАЦИЯ** (< дискретный фр. discret< лат. discretus — отдельный, разделенный) — разделение *изображения* на участки, в пределах которых значения *тона* либо усредняются, либо представлены одним из значений участка; имеет место при *сканировании (растрировании)* изображения, модуляции и аналого-цифровом преобразовании видеосигнала.

При цифровой обработке изображений непрерывный динамический диапазон значений яркости делится на ряд дискретных уровней. Эта процедура называется **квантованием**. Квантователь преобразует непрерывную переменную  $x$  в дискретную переменную  $x_{\text{кв}}$ , принимающую конечное множество значений  $(r_1, \dots, r_L)$ . Эти значения называются уровнями **квантования**. В общем случае преобразование выражается **ступенчатой функцией** (рис. 1.5). Если яркость  $x$  отсчета изображения принадлежит интервалу  $(d_j, d_{j+1}]$  (т.е., когда  $d_j < x \leq d_{j+1}$ ), то исходный отсчет заменяется на уровень **квантования**  $r_j$ , где  $d_j, j = \overline{1, L+1}$  — пороги квантования. При этом полагается, что динамический диапазон значений яркости ограничен и равен  $[d_1, d_{L+1}]$ .

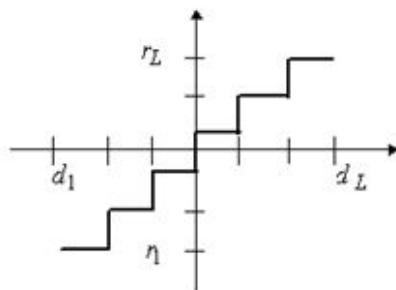


Рис. 1.5. Функция, описывающая **квантование**

Задача построения квантователя состоит в определении значений порогов  $d_j$  и уровней  $r_j$ . Простейший способ решения этой задачи состоит в разбиении динамического диапазона на одинаковые интервалы. Однако такое решение не является наилучшим. Если значения яркости большинства отсчетов изображения сгруппированы, например, в «темной» области и число уровней  $L$  ограничено, то целесообразно квантовать неравномерно. В «темной» области следует квантовать чаще, а в «светлой» реже. Это позволит уменьшить ошибку **квантования**  $\varepsilon = x - x_{\text{кв}}$ .

Таким образом, задачу построения квантователя можно сформулировать как задачу нахождения оптимальных значений  $d_j$  и  $r_j$ , удовлетворяющих некоторому критерию оптимизации. Обычно при фиксированном числе уровней квантователь оптимизируется по критерию минимальной среднеквадратической ошибки

$$E(\varepsilon^2) = E((x - x_{\text{кв}})^2) = \min_{\{d_j, r_j\}} \quad (1.12)$$

в предположении, что яркость  $x$  — случайная величина с известной плотностью вероятности  $w(x)$ .

## **2. Распознавание образов. Машины опорных векторов. (SVM, support vector machine).**

В исходном виде SVM представляет собой алгоритм, обучающийся ровно одной задаче: различению объектов двух классов. И этому SVM обучается очень быстро по сравнению, например, с нейронными сетями. Основная идея SVM такова. SVM учится только с учителем. Возьмем обучающее множество  $\mathbf{X} = \mathbf{X}_1, \dots, \mathbf{X}_N$  векторов признаков в евклидовом пространстве  $\mathbf{R}^d$ , разбитое на два класса, скажем, крестики и нолики. Попробуем разделить все пространство гиперплоскостью так, чтобы крестики попали в одно открытое полупространство, а нолики - в другое. Предположим, что такая гиперплоскость существует. Тогда она не единственна, так как множество разделяющих гиперплоскостей открыто. Попробуем найти в каком-то смысле самую лучшую гиперплоскость. И если мы ее нашли, то постановим все вектора признаков (уже не только из обучающего множества), попадающие в полупространство с крестиками, считать крестиками, а остальные - ноликами. А вероятность неправильной классификации оценивается как некоторая непрерывная убывающая функция от расстояния до разделяющей плоскости, равная 1/2 в нуле и стремящаяся к 0 на бесконечности.

## БИЛЕТ № 17

### 1. Сегментация. Обнаружение областей.

Сегментация изображения - это процесс разделения цифрового изображения на кратное число значимых областей или наборов пикселей относительно заданного критерия. Сегментация подразделяет изображение на составляющие его области или объекты.

Две основные цели сегментации:

1. Декомпозиция изображения на части, более удобные для дальнейшей обработки.
2. Изменение формы описания изображения.

Сегментация основывается на измерениях, проведенных над изображением, и может быть уровнем серого, цветом, текстурой, глубиной или движением. Результат сегментации изображения - ряд сегментов, покрывающие все изображение. Все пиксели в области подобны друг другу относительно некоторой характеристики или вычисленному значению, таких как цвет, интенсивность или текстура. Смежные области отличаются относительно тех же характеристик.

А) Выращивание областей представляет собой процедуру, которая группирует пиксели или подобласти в более крупные области по заранее заданным критериям укрупнения. Основной подход состоит в том, что вначале берется множество точек, играющих роль «центров кристаллизации», а затем на них наращиваются области путем присоединения к каждому центру тех пикселей из числа соседей, которые по своим заранее заданным свойствам близки к центру кристаллизации (например, имеют яркость или цвет в определенном диапазоне).

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

1. Найти все компоненты связности  $S(x, y)$  и применять 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  $f_Q$  по следующему правилу: если в точке  $(x, y)$  исходное изображение удовлетворяет заданному предикату  $Q$ , принимаем  $f_Q(x, y) = 1$ , иначе принимаем  $f_Q(x, y) = 0$ .
3. Сформировать изображение  $g$  путем присоединения к каждому центру кристаллизации в  $S$  всех единичных элементов  $f_Q$ , которые являются 8-связными с этим центром кристаллизации.
4. Пометить пиксели каждой связной компоненты  $S$  уникальной пометкой данной области (например числами 1, 2, 3...). Это и будет сегментированное изображение, полученное выращиванием областей.

А) Альтернативный подход состоит в том, чтобы провести первичное разбиение изображения на множество произвольных непересекающихся областей и в дальнейшем осуществлять слияние и/или разделение этих областей. Ниже излагаются основы метода разделения—слияния.

Пусть вся область изображения обозначена  $R$  и выбран предикат  $Q$ . Один из подходов к сегментации  $R$  состоит в том, чтобы последовательно разбивать ее на все более и более мелкие области-квадранты (т. е. четверти)  $R_p$ , пока для каждой из них не будет выполняться условие  $Q(R_p) = \text{TRUE}$  либо пока разбиение не будет остановлено в связи с достижением минимально допустимого размера области. Работа начинается со всей области изображения. Если  $Q(R) = \text{FALSE}$ , изображение делится на четверти вертикальной и горизонтальной пряммыми, проходящими через середину. Если для какой-то четверти предикат  $Q$  принимает значение  $\text{FALSE}$ , она аналогичным способом делится на более мелкие четверти и так далее. Такой метод разбиения удобно представлять в форме так называемого *квадродерева* (т. е. дерева, у которого вершины, не являющиеся листьями, имеют в точности четыре потомка), как показано на рис. 10.52 (изображения, соответствующие вершинам квадродерева, иногда называют *квадроизображениями* или *квадрообластями*). Отметим, что корень дерева соответствует целому изображению, а каждая другая вершина — какой-то из подобластей. В данном случае только область  $R_4$  подверглась дальнейшему разбиению.

Если использовать только операцию разделения, то в окончательном разбиении изображения обычно присутствуют смежные области с одинаковыми свойствами. Этот недостаток можно устранить, применив наряду с разделением также операцию слияния. Для соблюдения ограничений из раздела 10.1, касающихся сегментации, 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  $Q(R_j \cup R_k) = \text{TRUE}$ .

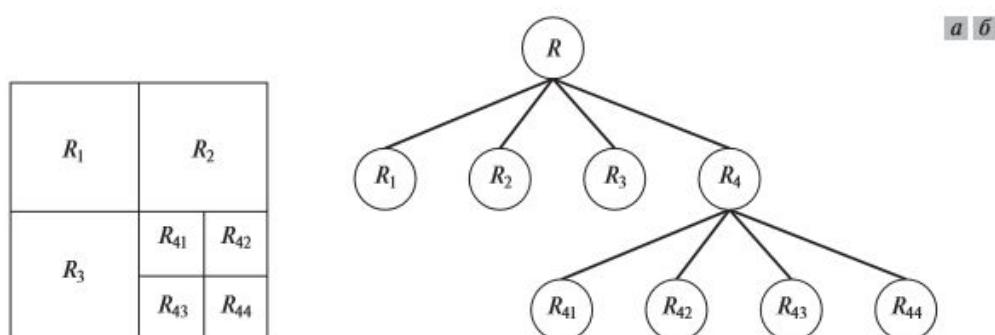


Рис. 10.52. (а) Изображение, разбитое на области. (б) Соответствующее квадродерево. Корень дерева  $R$  представляет всю область изображения

Проведенное обсуждение можно кратко суммировать в виде процедуры, на каждом шаге которой выполняются следующие действия:

1. Любая область  $R_p$ , для которой  $Q(R_p) = \text{FALSE}$ , разделяется на четыре непересекающиеся четверти.
2. Когда продолжать разбиение невозможно, любые две смежные области  $R_j$  и  $R_k$ , для которых  $Q(R_j \cup R_k) = \text{TRUE}$ , объединяются в одну.
3. Если невозможно выполнить ни одной операции слияния, то наступает окончание процедуры.

## **2. Основные взаимосвязи между пикселями.**

Пиксель может иметь несколько соседей. Обычно это 4 или 8, что обозначается как  $N_4(P)$  и  $N_8(P)$

Обычно выделяют следующие типы связи:

- 1) *четырехсвязный*. Два пикселя  $p$  и  $q$  с определенными значениями интенсивности являются четырехсвязными, если  $q$  относится к группе  $N_4(p)$ .
- 2) *восьмисвязный*. Два пикселя  $p$  и  $q$  со значениями интенсивности из  $Q$ , где  $Q$ -ряд значений интенсивности, являются восьмисвязными, если  $q$  относится к группе  $N_8(p)$ .
- 3) *m-связный (смешанная связь)*. Два пикселя  $p$  и  $q$  со значениями интенсивностей из  $Q$  являются *m*-связными, если:
  - а)  $q$  относится к группе  $N_4(p)$ ;
  - б)  $q$  относится к группе  $N_8(p)$  и множество  $N_4(p) \cap N_4(q)$  – пустое.

Иными словами, это множество пикселей, являющихся четырьмя соседними как по отношению к  $p$ , так и по отношению к  $q$  со значениями интенсивности из  $Q$ .

## БИЛЕТ № 18

### 1. Сегментация. Обнаружение контуров.

Сегментация изображения - это процесс разделения цифрового изображения на кратное число значимых областей или наборов пикселей относительно заданного критерия. Сегментация подразделяет изображение на составляющие его области или объекты.

Две основные цели сегментации:

1. Декомпозиция изображения на части, более удобные для дальнейшей обработки.
2. Изменение формы описания изображения.

Сегментация основывается на измерениях, проведенных над изображением, и может быть уровнем серого, цветом, текстурой, глубиной или движением. Результат сегментации изображения - ряд сегментов, покрывающие все изображение. Все пиксели в области подобны друг другу относительно некоторой характеристики или вычисленному значению, таких как цвет, интенсивность или текстура. Смежные области отличаются относительно тех же характеристик.

Выделение контуров - один из часто используемых методов в обработке цифровых изображений. Контур на изображении является значительным локальным изменением интенсивности изображения, и как правило, связаны с разрывом интенсивности изображения или первой производной интенсивности изображения.

Три шага в процессе выделение контуров - это а) фильтрация б) усиление и в) выделение.

Фильтрация применяется, если изображения были повреждены шумом, таких как соль и перец шума, импульсными помехами и Гауссовским шумом. Так же учитывается компромисс между прочностью края и снижением уровня шума при фильтрации.

Усиление подчеркивает пиксели там, где существует значительное изменение в локальной интенсивности и обычно выполняется путем вычисления градиента

Для нахождения в точке  $(x, y)$  изображения  $f$  одновременно величины перепада яркости и его направления обычно применяют градиент изображения  $\nabla f$ , определяемый как вектор

$$\nabla f \equiv \text{grad}(f) \equiv \begin{bmatrix} g_x \\ g_y \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix}. \quad (10.2-9)$$

Выделение. У многих точек на изображении есть ненулевое значение градиента, и не все из этих точек - края для заданного изображения. Пороговая обработка используется для обнаружения граничных точек.

Различные методы обнаружения края используются в операторах Собеля, Прюитта, Робертса, Кэнни, Лапласиана Гаусса, ЕМ алгоритме, OSTU и генетическом алгоритме.

Маски фильтров, применяемые для вычисления частных производных при нахождении градиента, часто называют *градиентными операторами, разностными операторами, операторами выделения контуров или детекторами контуров*.

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

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

Маски оператора Превитта

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

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

Маски оператора Собела

## 2. Проспранственная фильтрация.

Проспранственная фильтрация изображений — метод предварительной фильтрации изображения, при которой обработка происходит посредством применения некоторого оператора последовательно к каждой точке изображения.

Изображения, подлежащие распознаванию, могут быть зашумлены. Причиной тому могут быть искажения, вносимые в изображение предметами, активно отражающими свет (стекло, железо, водные поверхности), неравномерная прозрачность воздушного слоя, пыль, попавшая в объектив, качество используемой аппаратуры и др. Поэтому необходима дополнительная предварительная фильтрации изображения.

В общем виде модель искаженного изображения имеет вид:

- в пространственном представлении:  $g(x, y) = h(x, y) \cdot f(x, y) + \nu(x, y)$
- в частотном представлении:  $G(u, v) = H(u, v) \cdot F(u, v) + N(u, v)$

Если искажение представлено только шумом, то эти формулы принимают вид:

$$g(x, y) = f(x, y) + \nu(x, y)$$

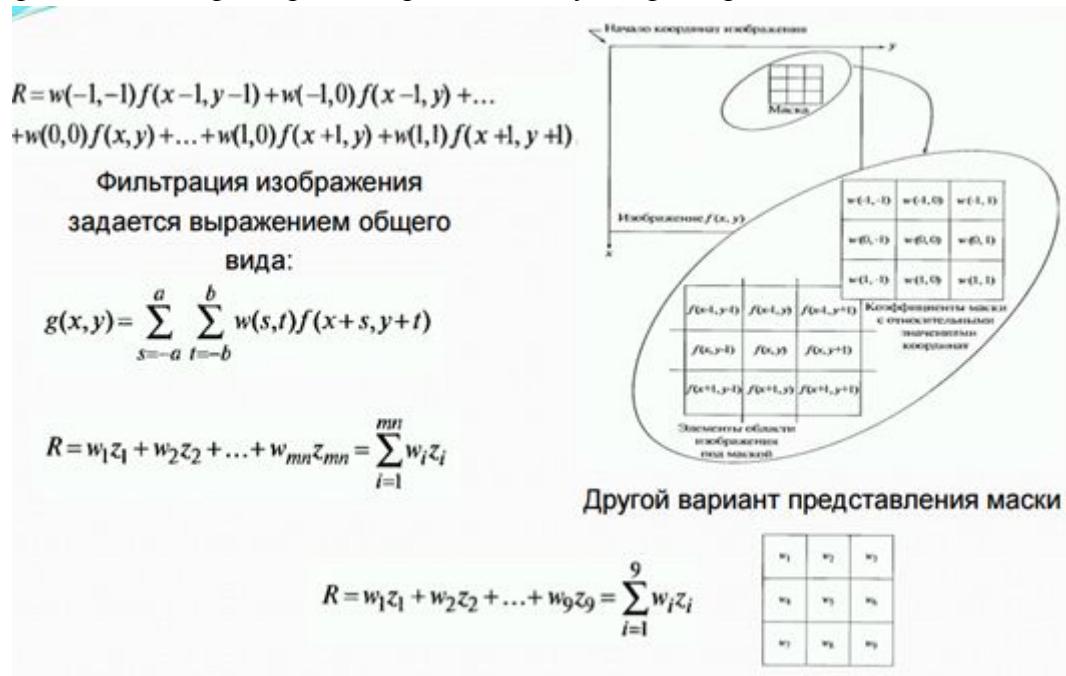
или

$$G(u, v) = F(u, v) + N(u, v)$$

Поскольку слагаемое, описывающее шум, неизвестно, то просто получить  $g(x, y)$  или  $G(u, v)$  невозможно. В таких случаях, когда присутствует только аддитивная составляющая шума, проспранственная фильтрация является лучшим решением.

## Механизмы пространственной фильтрации

Обработанное (фильтрованное) изображение возникает в процессе сканирования исходного изображения фильтром. Если оператор, выполняемый над пикселями исходного изображения, является линейным, то фильтр называют линейным пространственным фильтром. В противном случае фильтр является нелинейным.



w — коэффициенты фильтра размерами  $m \times n$ , а z — соответствующие значения яркостей пикселей изображения, покрываемых фильтром

### Сглаживающие пространственные фильтры

**Линейные сглаживающие фильтры**

$$R = \frac{1}{9} \sum_{i=1}^9 z_i$$

|   |   |   |
|---|---|---|
| 1 | 1 | 1 |
| 1 | 1 | 1 |
| 1 | 1 | 1 |

$\frac{1}{9} \times$

**Общая формула фильтрации изображения фильтром взвешенного среднего по окрестности**

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

|   |   |   |
|---|---|---|
| 1 | 2 | 1 |
| 2 | 4 | 2 |
| 1 | 2 | 1 |

$\frac{1}{16} \times$

## **БИЛЕТ № 19**

### **1. Цветовые модели.**

Цветовая модель — математическая модель описания представления цветов в виде кортежей чисел (обычно из трёх, реже — четырёх значений), называемых цветовыми компонентами или цветовыми координатами. Все возможные значения цветов, задаваемые моделью, определяют цветовое пространство.

Цветовая модель задаёт соответствие между воспринимаемыми человеком цветами, хранимыми в памяти, и цветами, формируемыми на устройствах вывода (возможно, при заданных условиях)

Цветовые модели:

**CMY(K)**( Cyan, Magenta, Yellow, Key color - субтрактивная (subtract, англ. - вычитать) схема формирования цвета, используемая в полиграфии для стандартной триадной печати. Обладает меньшим, в сравнении с RGB, цветовым охватом.);

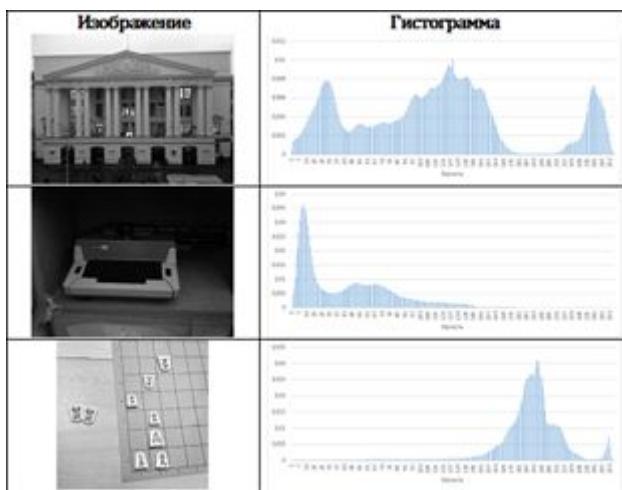
**RGB**(Red, Green, Blue);

**Lab** (В цветовом пространстве Lab значение светлоты отделено от значения хроматической составляющей цвета (тон, насыщенность). Светлота задана координатой L (изменяется от 0 до 100, то есть от самого темного до самого светлого), хроматическая составляющая — двумя декартовыми координатами a и b. Первая обозначает положение цвета в диапазоне от зеленого до пурпурного, вторая — от синего до желтого.);

**HSB** (HSB - модель, которая в принципе является аналогом RGB, она основана на её цветах, но отличается системой координат. Любой цвет в этой модели характеризуется тоном (Hue), насыщенностью (Saturation) и яркостью (Brightness). Тон - это собственно цвет. Насыщенность - процент добавленной к цвету белой краски. Яркость - процент добавленной чёрной краски.);

### **2. Видоизменение гистограммы. Приведение (задание гистограммы).**

*Гистограммой* цифрового изображения с уровнями яркости в диапазоне  $[0, L - 1]$  называется дискретная функция  $h(r_k) = n_k$ , где  $r_k$  есть  $k$ -й уровень яркости, а  $n_k$  — число пикселей на изображении, имеющих яркость  $r_k$ . Общей практикой является нормировка гистограммы путем деления каждого из ее значений на общее число пикселей в изображении, обозначаемое произведением  $MN$ , где, как обычно,  $M$  и  $N$  суть число строк и столбцов изображения, то есть его размеры. Тем самым значения нормированной гистограммы будут  $p(r_k) = n_k/MN$  для  $k = 0, 1, \dots, L - 1$ . Вообще говоря,  $p(r_k)$  есть оценка вероятности появления пикселя<sup>3</sup> со значением яркости  $r_k$ . Заметим, что сумма всех значений нормированной гистограммы равна единице.



По горизонтальной оси каждого графика отложены значения уровней яркости  $r_k$ . По вертикальной оси — значения гистограммы  $h(r_k) = n_k$  или  $p(r_k) = n_k/MN$ , если они нормированы. Тем самым гистограммы могут быть представлены просто как графики зависимостей  $h(r_k) = n_k$  от  $r_k$  или  $p(r_k) = n_k/MN$  от  $r_k$ .

В некоторых случаях улучшение, основанное на модели равномерной гистограммы (эквализация гистограммы стремится сформировать выходное изображение с равномерной гистограммой), не является наилучшим подходом. В частности, иногда полезно задать иную желаемую форму гистограммы для обрабатываемого изображения. Метод, позволяющий получить обработанное изображение с задаваемой формой гистограммы, называется методом приведения гистограммы или задания гистограммы.

Вернемся ненадолго к непрерывному представлению яркостей  $r$  и  $z$  (рассматриваемые непрерывные случайные переменные), которые обозначают уровни яркостей входного и выходного (обработанного) изображений, и пусть  $p_r(r)$  и  $p_z(z)$  соответственно означают их непрерывные плотности распределения вероятностей. Значения  $p_r(r)$  мы можем оценить по исходному изображению, в то время как  $p_z(z)$  является задаваемой плотностью распределения вероятностей, которую должно иметь выходное изображение.

(ПРИМЕР НА ФОТО НИЖЕ) На первый взгляд кажется, что эквализация гистограммы будет хорошим способом улучшения изображения, таким, что детали в темных областях станут более видимыми, но это не всегда так.

Предположим, что мы имеем изображение, на котором преобладают большие темные области, вследствие чего точки на изображении концентрируются вблизи темного края шкалы яркостей и на гистограмме появляется пик вблизи нуля. На первый взгляд кажется, что эквализация гистограммы будет хорошим способом улучшения изображения, таким, что детали в темных областях станут более видимыми, но это не так. Причина заключается в том, что большинство значений пикселей входного изображения сконцентрировано вблизи нулевых уровней гистограммы. Эффект применения данного преобразования к входному изображению с целью эквализации гистограммы сводится к тому, что очень узкий интервал темных пикселей

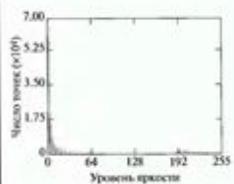
отображается в верхнюю часть яркостного диапазона выходного изображения. Поскольку большая часть пикселей входного изображения имеет значения как раз в этом интервале, следовало бы ожидать, что результатом будет светлое изображение, выглядящее как «вылинявшее». Поскольку проблемы с функцией преобразования возникли из-за высокой концентрации значений пикселей исходного изображения вблизи нуля, разумным подходом было бы модифицировать гистограмму таким образом, чтобы избежать этого обстоятельства. Например, поможет искусственно заданная функция, которая сохраняет основную форму исходной гистограммы, но имеет сглаженный переход уровней в темной области шкалы яркостей.

Если вкратце, то метод задания гистограммы в большинстве случаев является процессом проб и ошибок.

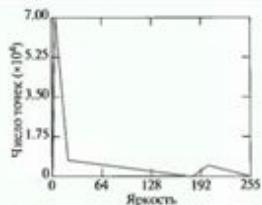


# Сравнение эквализации и приведения гистограмм

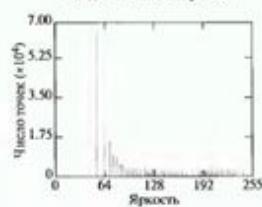
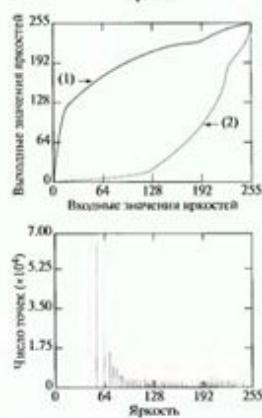
Исходное изображение



Приведение



Эквализация



## **БИЛЕТ № 20**

### **1. Сегментация. Обнаружение моделей сегментов.**

Сегментация изображения - это процесс разделения цифрового изображения на кратное число значимых областей или наборов пикселей относительно заданного критерия. Сегментация подразделяет изображение на составляющие его области или объекты.

Две основные цели сегментации:

1. Декомпозиция изображения на части, более удобные для дальнейшей обработки.
2. Изменение формы описания изображения.

Сегментация основывается на измерениях, проведенных над изображением, и может быть уровнем серого, цветом, текстурой, глубиной или движением. Результат сегментации изображения - ряд сегментов, покрывающие все изображение. Все пиксели в области подобны друг другу относительно некоторой характеристики или вычисленному значению, таких как цвет, интенсивность или текстура. Смежные области отличаются относительно тех же характеристик.

При обработке изображений и выделении контуров изображения может возникать задача замены выделенного контура на близкую ему математическую модель в виде прямой, дуги, окружности, эллипса и т.п. При таком подходе выполняется аппроксимация выделенного контура соответствующей **моделью**. Для подбора такой модели обычно используют метод наименьших квадратов. Например, при выявлении на изображении прямой или отрезка прямой с помощью преобразования Хафа, этот метод можно использовать для проверки пикселей, проголосовавших за данный элемент, на соответствие их уравнению прямой. При использовании метода наименьших квадратов должен быть известен метод, с помощью которого выполняют подбор соответствующей модели из большого числа возможных вариантов.

Кроме подбора модели прямых и отрезков прямых могут выполняться аналогичные приемы для кривых, плоскостей и иных сегментов. В простейших случаях предполагается, что известны вид модели и набор исходных точек. В более сложных случаях, когда контурные цепочки могут быть представлены сочленением набора различных форм, применяют сегментацию кривых с последующим поиском моделей. Реальные изображения могут содержать контурные сегменты в виде лент, углов и более сложные конфигурации.

## 2. Преобразование Фурье и частотная область.

Преобразование Фурье (символ  $\mathcal{F}$ ) — операция, сопоставляющая одной функции вещественной переменной другую функцию вещественной переменной. Эта новая функция описывает коэффициенты («амплитуды») при разложении исходной функции на элементарные составляющие — гармонические колебания с разными частотами (подобно тому, как музыкальный аккорд может быть выражен в виде амплитуд нот, которые его составляют).

Функция, заданная как рядом, так и преобразованием Фурье, может быть полностью, без потери информации, восстановлена (реконструирована) при помощи некоторой процедуры обращения. Это свойство является одним из наиболее важных свойств рассматриваемых представлений, поскольку оно позволяет работать в «Фурье-области», а затем вернуться в исходную область определения функции без потери какой-либо информации.

Преобразование Фурье функции вещественной переменной является интегральным и задаётся следующей формулой:

$$\hat{f}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) e^{-ix\omega} dx.$$

двумерному дискретному преобразованию Фурье (ДПФ):

$$F(u, v) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) e^{-i2\pi(ux/M+vy/N)}, \quad (4.5-15)$$

где  $f(x, y)$  — цифровое изображение размерами  $M \times N$ . Как и в одномерном случае, выражение (4.5-15) должно быть вычислено для всех значений дискретных переменных  $u$  и  $v$  в диапазонах  $u = 0, 1, 2, \dots, M - 1$  и  $v = 0, 1, 2, \dots, N - 1$ <sup>15</sup>.

Если имеется преобразование  $F(u, v)$ , можно получить  $f(x, y)$  при помощи обратного дискретного преобразования Фурье (обратного ДПФ)

$$f(x, y) = \frac{1}{MN} \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F(u, v) e^{i2\pi(ux/M+vy/N)} \quad (4.5-16)$$

для  $x = 0, 1, 2, \dots, M - 1$  и  $y = 0, 1, 2, \dots, N - 1$ . Выражения (4.5-15) и (4.5-16) составляют пару двумерных дискретных преобразований Фурье. Оставшаяся часть данной главы базируется на свойствах этих двух выражений и использовании их для фильтрации изображений в частотной области.

## БИЛЕТ № 21

### 1. Диаграмма цветностей.

Цветовая модель — математическая модель описания представления цветов в виде кортежей чисел (обычно из трёх, реже — четырёх значений), называемых *цветовыми компонентами* или *цветовыми координатами*. Все возможные значения цветов, задаваемые моделью, определяют цветовое пространство.

Человек является трихроматом — сетчатка глаза имеет три вида рецепторов (колбочек), ответственных за цветное зрение. Можно считать, что каждый вид колбочек даёт свой отклик на определённую длину волны видимого спектра. Важным свойством (для всех физически реализуемых цветов) является неотрицательность как функций отклика, так и результирующих цветовых координат для всех цветов. Системой, основанной на откликах колбочек человеческого глаза, является цветовая модель LMS. Исторически сложилось, что для измерения цвета используется другое цветовое пространство — XYZ. Это — эталонная цветовая модель, заданная в строгом математическом смысле организацией CIE (International Commission on Illumination — Международная комиссия по освещению) в 1931 году. Модель CIE XYZ является мастер моделью практически всех остальных цветовых моделей, используемых в технических областях.

$$X = \int_{380}^{780} I(\lambda) \bar{x}(\lambda) d\lambda$$

$$Y = \int_{380}^{780} I(\lambda) \bar{y}(\lambda) d\lambda$$

$$Z = \int_{380}^{780} I(\lambda) \bar{z}(\lambda) d\lambda$$

где  $I(\lambda)$  — спектральная плотность какой-либо энергетической фотометрической величины (например потока излучения, энергетической яркости и т. п., в абсолютном или относительном выражении).

Если формально построить сечение пространства XYZ плоскостью , то можно две

оставшиеся линейно независимыми координаты записать в виде:

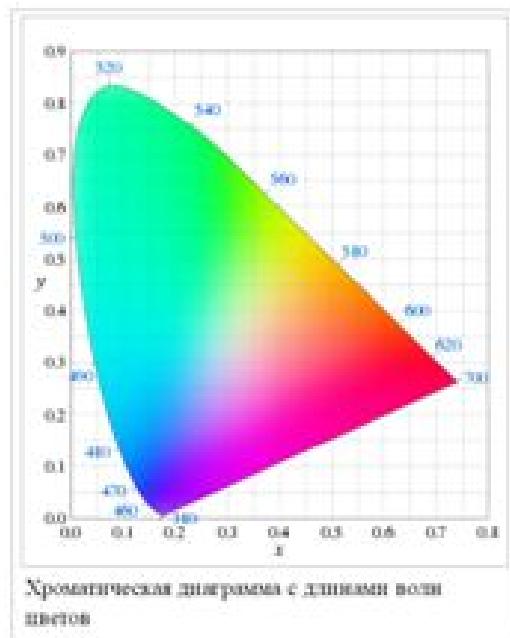
$$x = X/(X + Y + Z)$$

$$y = Y/(X + Y + Z).$$

аналогично, но необязательно:

$$z = Z/(X + Y + Z)$$

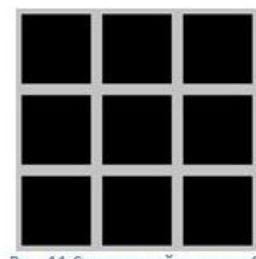
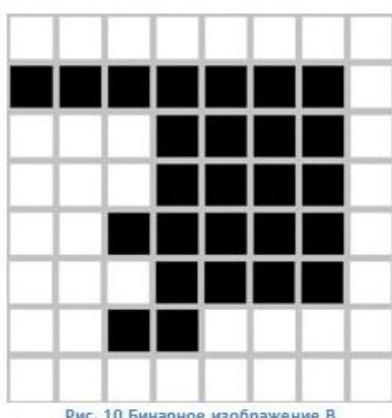
Такое сечение называется хроматической диаграммой (диаграммой цветности).



## 2. Применение морфологических алгоритмов к полутоноовым изображениям.

Наращивание.

Структурный элемент S применяется ко всем пикселам бинарного изображения. Каждый раз, когда начало координат структурного элемента совмещается с единичным бинарным пикселом, ко всему структурному элементу применяется перенос и последующее логическое сложение с соответствующими пикселями бинарного изображения. Результаты логического сложения записываются в выходное бинарное изображение, которое изначально инициализируется нулевыми значениями.



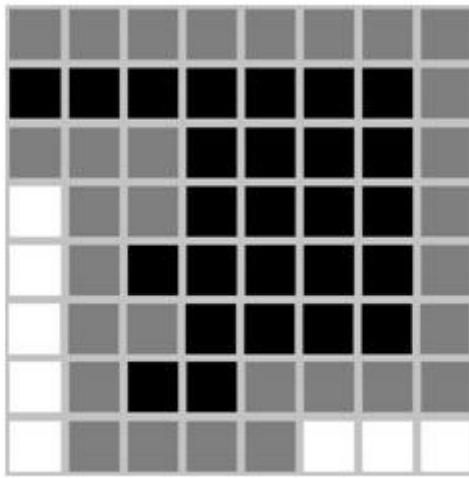


Рис. 12 Наращивание изображения В структурным элементом  $S^2$

### Эрозия

При выполнении операции эрозии структурный элемент тоже проходит по всем пикселям изображения. Если в некоторой позиции каждый единичный пиксель структурного элемента совпадет с единичным пикселом бинарного изображения, то выполняется логическое сложение центрального пикселя структурного элемента с соответствующим пикселом выходного изображения.

В результате применения операции эрозии все объекты, меньшие чем структурный элемент, стираются, объекты, соединённые тонкими линиями становятся разъединёнными и размеры всех объектов уменьшаются.

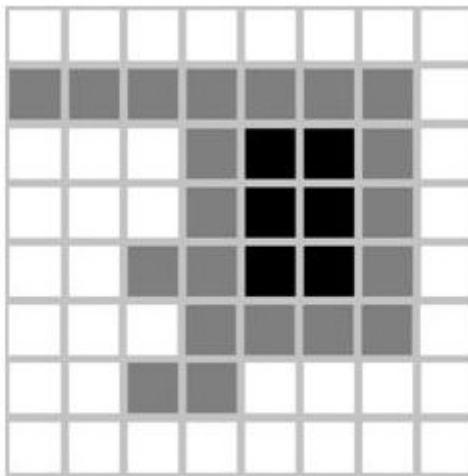


Рис. 13 Эрозия изображения В структурным элементом  $S^2$

### Размыкание

Операция эрозии полезна для удаления малых объектов и различных шумов, но у этой операции есть недостаток – все остающиеся объекты уменьшаются в размере.

Этого эффекта можно избежать, если после операции эрозии применить операцию наращивания с тем же структурным элементом.

Размыкание отсеивает все объекты, меньшие чем структурный элемент, но при этом помогает избежать сильноуменьшения размера объектов. Также размыкание идеально подходит для удаления линий, толщина которых меньше, чем диаметр структурного элемента. Также важно помнить, что после этой операции контуры объектов становятся более гладкими.

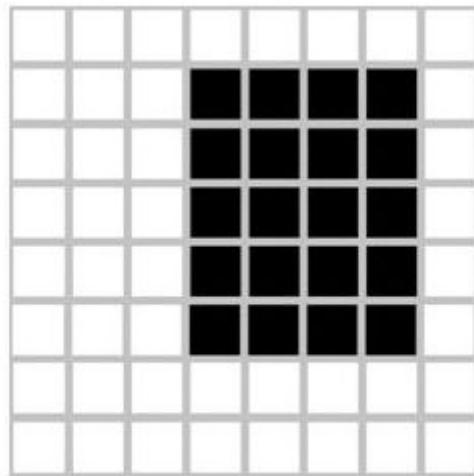


Рис. 14 Размыкание изображения В структурным элементом

### Замыкание

Если к изображению применить сначала операцию наращивания, то мы сможем избавиться от малых дыр и щелей, но при этом произойдёт увеличение контура объекта. Избежать этого увеличения позволяет операция эрозия, выполненная сразу после наращивания с тем же структурным элементом.

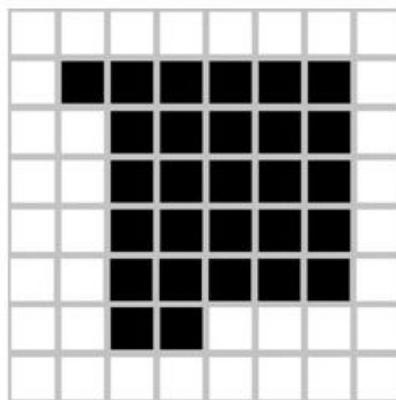


Рис. 15 Замыкание изображения В структурным элементом 5

## Условное наращивание

Одним из типичных применений бинарной морфологии является выделение на бинарном изображении компонент, у которых форма и размеры удовлетворяют заданным ограничениям. Во многих подобных задачах возможно построение структурного элемента, который после применения к бинарному изображению удаляет не удовлетворяющие ограничениям компоненты и оставляет несколько единичных пикселов, соответствующих удовлетворяющим ограничениям компонентам. Но для последующей обработки могут потребоваться компоненты целиком, а не только их фрагменты, оставшиеся после эрозии. Для решения этой проблемы была введена операция условного наращивания.

Множество полученное в результате эрозии циклически наращивается структурным элементом  $S$ , и на каждом шаге результат уменьшается до подмножества пикселов, которые имеют единичные значения на исходном изображении  $B$ . Операция условного наращивания пояснена на рисунке ниже. На этом рисунке бинарное изображение  $B$  было подвергнуто эрозии элементом  $V$  для выделения компонент, содержащих вертикальные фрагменты высотой 3 пикселя. На полученном изображении  $C$  есть две таких компоненты. Чтобы выделить эти компоненты целиком, изображение  $C$  условно наращивается элементом  $D$  относительно исходного изображения  $B$ .



Рис. 16 Структурный элемент V

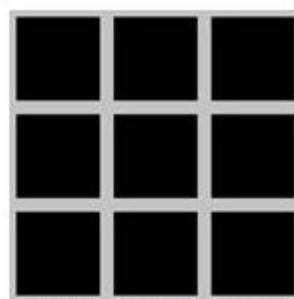


Рис. 17 Структурный элемент D

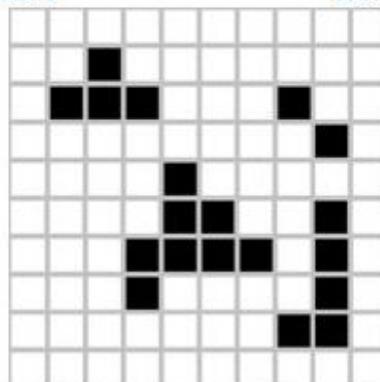


Рис. 18 Бинарное изображение В

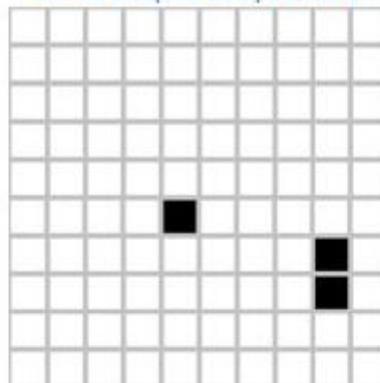


Рис. 19 Эрозия изображения В структурным элементом V

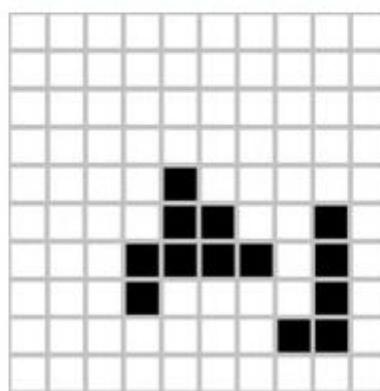


Рис. 20 Условное наращивание изображения структурным элементом D

## Выделение границ

Морфологические операции можно также использовать для выделения границ бинарного объекта. Это операция очень важна, потому что граница является полным, и в то же время весьма компактным описанием объекта.

Легко заметить, что граничные точки имеют как минимум один фоновый пиксель в своей окрестности. Таким образом, применив оператор эрозии с структурным элементом, содержащим все возможные соседние элементы, мы удалим все граничные точки... Тогда граница получится с помощью операции разности множеств между исходным изображением и изображением, полученным в результате эрозии.

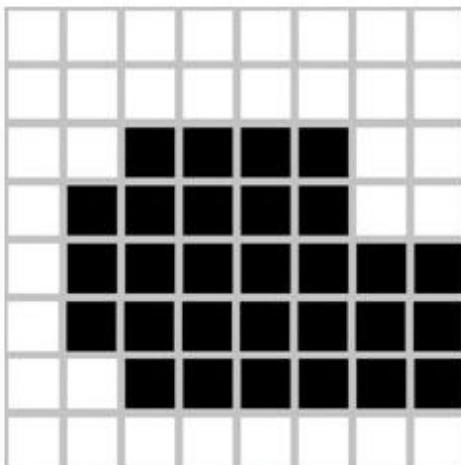


Рисунок 21 Исходное изображение

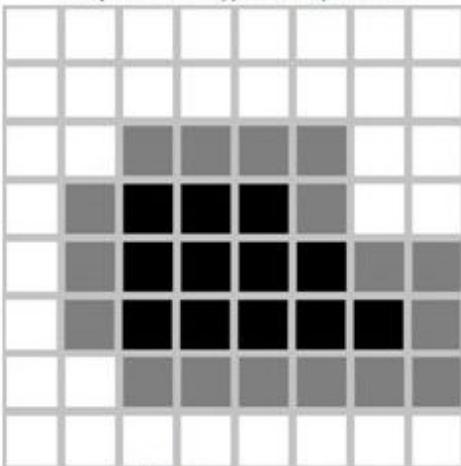


Рисунок 23 Изображение, после эрозии

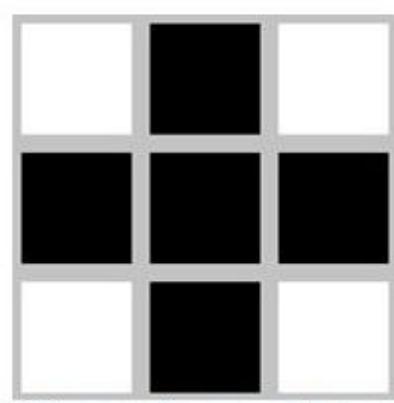


Рисунок 22 Структурный элемент для 4-компонентной окрестности

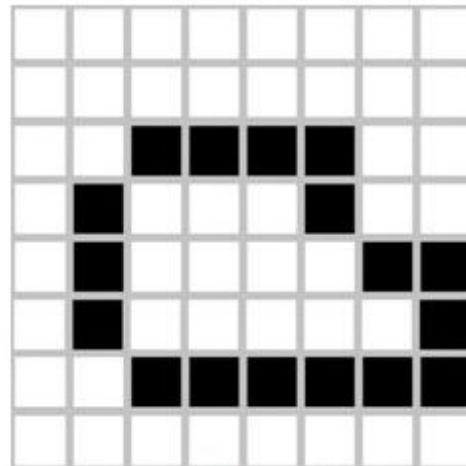


Рисунок 24 Границы объекта

## БИЛЕТ № 22

### 1. Обработка изображений в псевдоцветах.

Обработка изображения в псевдоцветах подразумевает присвоение цветов пикселям полутонового изображения на основе некоторого определенного правила. Термин псевдоцвета, или ложные цвета, используется для того, чтобы отличать цветные изображения, полученные в результате присвоения цветов точкам монохромного изображения, от изображений в натуральных цветах.

#### Квантование по яркости

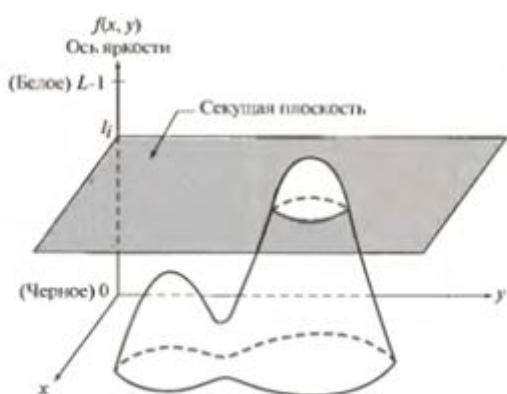


Рис. 6.18. Геометрическое объяснение метода квантования по яркости.

Если рассматривать изображение как поверхность в трехмерном пространстве. Обсуждаемый метод можно представить, как метод основанный на проведении плоскостей параллельный координатной плоскости изображения. Присвоим каждой из сторон показанной на рисунке плоскости свой цвет, и будем кодировать пиксели, которым отвечают точки поверхности, лежащие выше плоскости, одним цветом, а пиксели, которым отвечают точки поверхности, лежащие ниже плоскости, - другим. Точкам, которые соответствуют пересечению плоскости с поверхностью, можно присвоить любой из двух выбранных цветов. В результате мы получим двухцветное изображение, видом которого можно управлять, двигая секущую плоскость вверх и вниз вдоль оси яркости.

#### Преобразование яркости в цвет

Основная идея, лежащая в основе этого метода, состоит в том, чтобы осуществить три независимых преобразования значений яркости для каждого пикселя входного изображения. Затем три полученных изображения подаются в красный, зеленый и синий каналы цветного монитора. В результате формируется составное изображение, цветовое содержание которого определяется природой используемых функций преобразования. Отметим, что эти преобразования затрагивают лишь значения яркости изображения и не зависят от положения точки на изображении. Метод, может использовать гладкие и нелинейные функции, что, естественно, придает всей методике значительную гибкость.

## 2. Морфологические алгоритмы.

| Операция                          | Выражение   | Примечание<br>(римские цифры в скобках<br>означают номера примитивов<br>на рис. 9.33)  |
|-----------------------------------|---|--|
| Параллель-<br>ный перенос         | $(B)_z = \{w \mid w = b + z, b \in B\}$   | Сдвиг центра (начала координат)<br>множества $B$ в точку $z$   |
| Централь-<br>ное отраже-<br>ние   | $\hat{B} = \{w \mid w = -b, b \in B\}$  | Симметричное отражение всех<br>элементов $B$ относительно начала<br>координат  |
| Дополнение                        | $A^c = \{w \mid w \notin A\}$   | Множество точек, не входящих<br>в $A$  |
| Разность                          | $A \setminus B = \{w \mid w \in A, w \notin B\} = A \cap B^c$   | Множество точек, принадлежа-<br>щих $A$ , но не принадлежащих $B$  |
| Дилатация                         | $A \oplus B = \{z \mid (\hat{B})_z \cap A \neq \emptyset\}$   | «Расширение» границы множе-<br>ства $A$ (I)  |
| Эрозия                            | $A \ominus B = \{z \mid (B)_z \subseteq A\}$  | «Сужение» границы множества<br>$A$ (I)   |
| Размыкание                        | $A \circ B = (A \ominus B) \oplus B$  | Сглаживает контуры, разры-<br>вает узкие перешейки, убирает<br>небольшие островки и острые<br>выступы (I)  |
| Замыкание                         | $A \bullet B = (A \oplus B) \ominus B$  | Сглаживает контуры, заполняет<br>узкие разрывы, углубления и не-<br>большие дырки (I)  |
| Попадание/<br>пропуск             | $A \circledast B = (A \ominus B_1) \cap (A^c \ominus B_2) = \\ = (A \ominus B_1) \setminus (A \oplus \hat{B}_2)$  | Множество координат точек, в ко-<br>торых одновременно для $B_1$ есть<br>совпадение в $A$ , а для $B_2$ — в $A^c$                                  |
| Выделение<br>границы              | $\beta(A) = A \setminus (A \ominus B)$  | Множество граничных точек<br>множества $A$ (I)   |
| Заполнение<br>дырок               | $X_k = (X_{k-1} \oplus B) \cap A^c; \quad k = 1, 2, 3, \dots$   | Заполняет дырки в множестве<br>$A$ ; $X_0$ — массив нулей с единицей<br>в каждой дырке (II)  |
| Выделение<br>связных<br>компонент | $X_k = (X_{k-1} \oplus B) \cap A; \quad k = 1, 2, 3, \dots$   | Находит связные компоненты<br>в множестве $A$ ; $X_0$ — массив нулей<br>с единицей в каждой связной<br>компоненте (I)                              |
| Выпуклая<br>оболочка              | $X_k^i = (X_{k-1}^i \circledast B^i) \cup A; \quad i = 1, 2, 3, 4; \\ k = 1, 2, 3, \dots; \quad X_0^i = A \quad \text{и} \quad D^i = X_{\text{сход}}^i$ | Находит для множества $A$ вы-<br>пуклую оболочку $C(A)$ . Индекс<br>«сход» соответствует сходимости<br>в том смысле, что $X_k^i = X_{k-1}^i$ (III) |

**Эрозия** — приводит к сужению деталей на изображении. По существу, эрозию можно рассматривать как операцию морфологической фильтрации, которая отфильтровывает (удаляет) детали изображения с размерами меньше структурообразующего множества.

**Дилатация.** В отличие от эрозии, которая приводит к сужению или утончению, дилатация «расширяет» или «утолщает» объекты на двоичном изображении. Одно из простейших применений дилатации — устранение разрывов линий путем их перекрытия.

**Размыкание.** 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.

**Морфологическое преобразование «попадание/пропуск»** является основным инструментом для обнаружения объектов определенных размеров и формы.

## **БИЛЕТ № 23**

### **1. Цветовые преобразования.**

Предметом цветовых преобразований является обработка компонент цветного изображения в рамках одной отдельно взятой цветовой модели. Это отличает данные преобразования от преобразований координат цвета при переходе из одной цветовой модели в другую (таких как преобразования цветов из модели RGB в модель HSI и обратно).

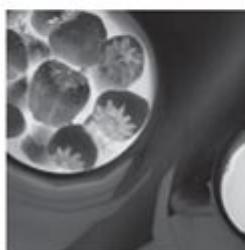
мы задаем преобразование цветных изображений следующим выражением:  $g(x,y) = T[f(x,y)],$

где  $f(x,y)$  — цветное изображение на входе,  $g(x,y)$  — преобразованное, или обработанное, цветное изображение на выходе и  $T$  — действующий на изображение  $f$  оператор обработки по пространственной окрестности точки  $(x,y)$ .

значение пикселя представляет собой трехмерный или многомерный вектор, т.е. набор из трех или более координат того цветового пространства, которое используется для представления изображения



Цветное изображение



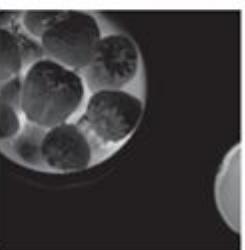
Голубой (C)



Пурпурный (M)



Желтый (Y)



Черный (K)



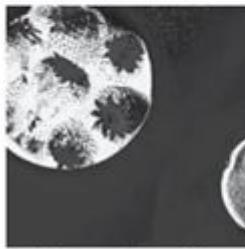
Красный (R)



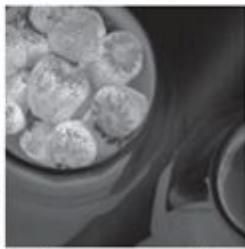
Зеленый (G)



Синий (B)



Цветовой тон (H)



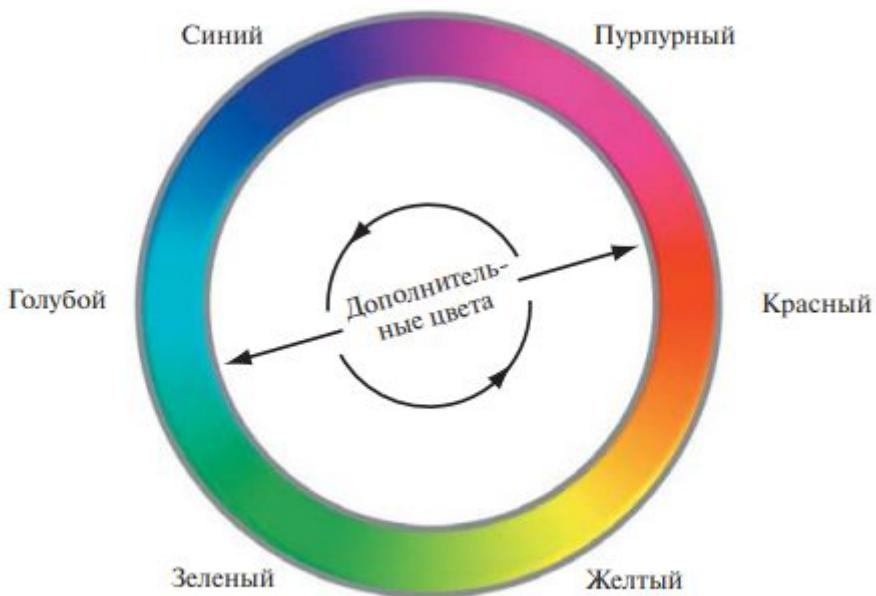
Насыщенность (S)



Интенсивность (I)

### 6.5.2. Цветовое дополнение

Цвета, расположенные друг напротив друга на изображенном на рис. 6.32 цветовом круге<sup>7</sup>, называются *дополнительными цветами*. Наш интерес к цветовому дополнению, т. е. к переходу от данных цветов к соответствующим дополнительным цветам, связан с тем обстоятельством, что эта операция аналогична преобразованию полутонового изображения в негатив (см. раздел 3.2.1). Как и в полутоновом случае, цветовое дополнение полезно для выявления деталей внутри темных областей цветного изображения, особенно когда размеры областей заметно превосходят размеры деталей.



**Рис. 6.32.** Дополнительные цвета на цветовом круге

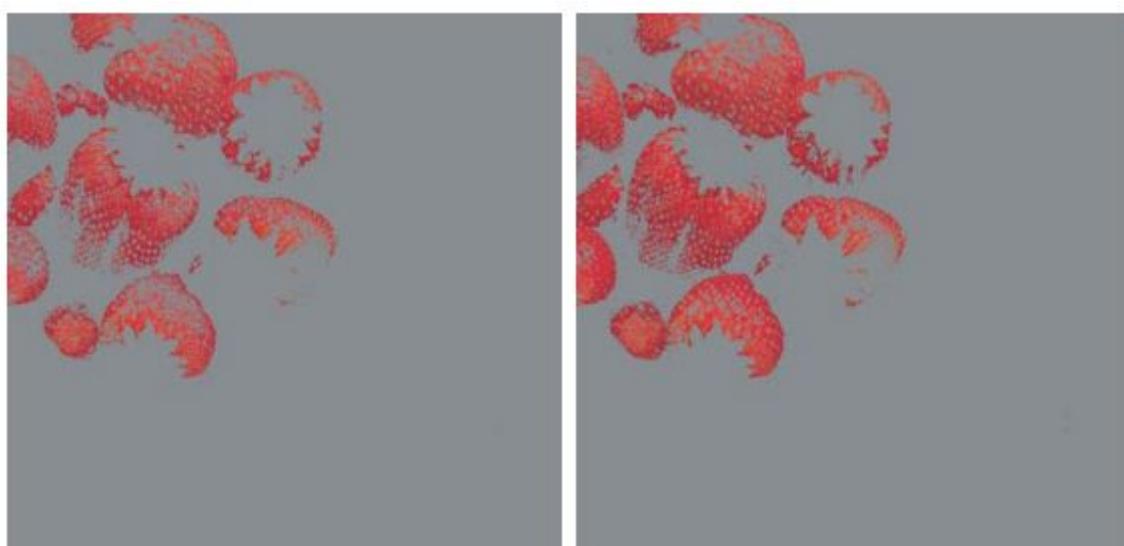


### 6.5.3. Вырезание цветового диапазона

Вырезание определенного диапазона цветов на изображении применяется для выделения некоторых объектов из их окружения. Основная идея заключается в том, чтобы либо (1) воспроизвести интересующие цвета так, чтобы они выступали на общем фоне, либо (2) использовать определяемые цветом области в качестве маски при дальнейшей обработке.

Однако, поскольку значение цветного пикселя представляет собой  $n$ -мерный вектор, функции преобразования для вырезания цветового диапазона являются более сложными, чем аналогичные функции для вырезания яркостного диапазона. В действительности интересующие нас преобразования превосходят по сложности все цветовые преобразования, рассмотренные до сих пор. Это связано с тем, что при любом применяемом на практике методе вырезания цветового диапазона каждая цветовая компонента преобразованного пикселя зависит от всех  $n$  цветовых компонент исходного пикселя.

Один из простейших способов «разделить» цветное изображение состоит в том, чтобы отобразить все цвета, лежащие вне области интереса, в некоторый нейтральный, не бросающийся в глаза цвет.



#### 6.5.4. Яркостная и цветовая коррекция

Вначале обычно производится коррекция яркостного диапазона изображения, а затем устраняется цветовой дисбаланс, такой как недостаточная или избыточная насыщенность цветов. Яркостной диапазон цветного изображения связан с общим распределением интенсивности его цветов. Как и в монохромном случае, часто желательно распределить значения интенсивности цветного изображения равномерно между наиболее светлыми и наиболее темными значениями. В приводимых ниже примерах рассматриваются различные цветовые преобразования, используемые для яркостной и цветовой коррекции.

Подобно системе HSI, система  $L^*a^*b^*$  превосходно разделяет интенсивность (которая представлена яркостью  $L^*$ ) и цветность (которая представлена двумя цветоразностями:  $a^*$  — красный минус зеленый и  $b^*$  — зеленый минус синий). Это свойство делает систему  $L^*a^*b^*$  весьма удобной как для улучшения изображений (тональной и цветовой коррекции), так и для их сжатия<sup>8</sup>.



- После того как яркостные свойства изображения правильно откорректированы, можно приступать к устранению цветового дисбаланса. Хотя объективно наличие цветового дисбаланса может быть установлено путем анализа на изображении какого-нибудь участка известного цвета с помощью колориметра, однако в том случае, когда на изображении имеются белые области, где значения RGB- или CMY(K)-компонент должны быть равны между собой, возможна и достаточно точная визуальная оценка. ]

Существуют различные способы коррекции, применяемые при наличии цветового дисбаланса. При изменении цветовых компонент изображения важно понимать, что каждое действие влияет на весь цветовой баланс на изображении. Другими словами, восприятие каждого цвета зависит от всех окружающих его цветов. Тем не менее, чтобы предсказать, как одна цветовая компонента влияет на остальные, можно использовать цветовой круг

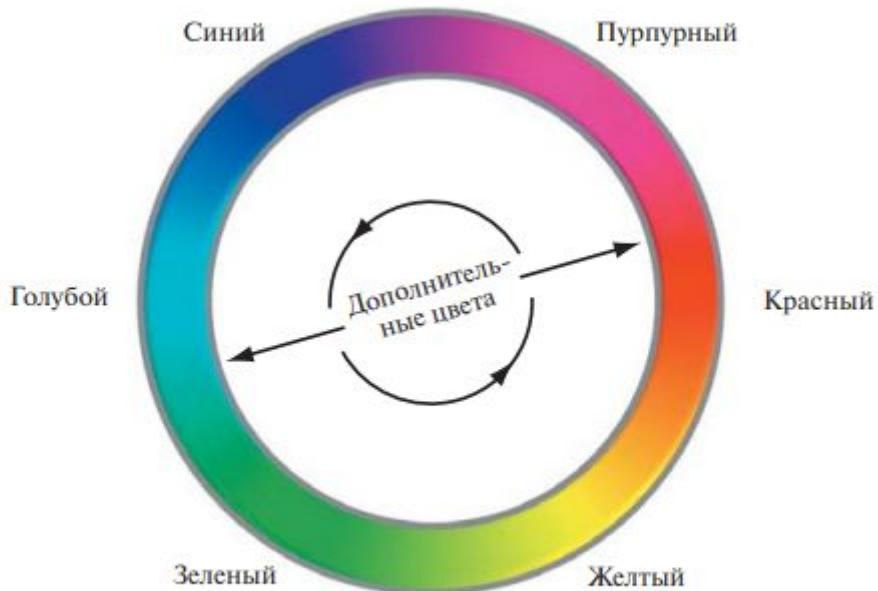


Рис. 6.32. Дополнительные цвета на цветовом круге

Основываясь на этом цветовом круге можно, например, прийти к заключению, что относительное количество некоторого цвета на изображении может быть увеличено, если уменьшить количество противоположного (дополнительного) цвета.



## **2. Основы морфологической обработки**

Морфологическая обработка связана с инструментами для извлечения таких компонентов изображения, которые могут быть полезны для представления и описания формы.

Морфологические операции можно использовать для построения фильтров, похожих по своему принципу на пространственные фильтры.

### ***Примеры морфологических операций:***

**Эрозия** — приводит к сужению деталей на изображении. По существу, эрозию можно рассматривать как операцию морфологической фильтрации, которая отфильтровывает (удаляет) детали изображения с размерами меньше структурообразующего множества.

**Дилатация.** В отличие от эрозии, которая приводит к сужению или утончению, дилатация «расширяет» или «утолщает» объекты на двоичном изображении. Одно из простейших применений дилатации — устранение разрывов линий путем их перекрытия.

**Размыкание.** 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.

**Морфологическое преобразование «попадание/пропуск»** является основным инструментом для обнаружения объектов определенных размеров и формы.

## **БИЛЕТ № 24**

### **1. Представление знаний в системах компьютерного зрения.**

Реализация систем компьютерного зрения сильно зависит от области их применения, аппаратной платформы и требований по производительности. Некоторые системы являются автономными и решают специфические проблемы детектирования и измерения, тогда как другие системы составляют подсистемы более крупных систем, которые уже могут содержать подсистемы контроля механических манипуляторов (роботы), информационные базы данных (поиск похожих изображений), интерфейсы человек-машина (компьютерные игры) и т. д. Однако, существуют функции, типичные для многих систем компьютерного зрения.

- **Получение изображений:** цифровые изображения получаются от одного или нескольких датчиков изображения, которые помимо различных типов светочувствительных камер включают датчики расстояния, радары, ультразвуковые камеры и т. д. В зависимости от типа датчика, получающиеся данные могут быть обычным 2D изображением, 3D изображением или последовательностью изображений. Значения пикселей обычно соответствуют интенсивности света в одной или нескольких спектральных полосах (цветные или изображения в оттенках серого), но могут быть связаны с различными физическими измерениями, такими как глубина, поглощение или отражение звуковых или электромагнитных волн, или ядерным магнитным резонансом.
- **Предварительная обработка:** перед тем, как методы компьютерного зрения могут быть применены к видеоданным с тем, чтобы извлечь определённую долю информации, необходимо обработать видеоданные, с тем чтобы они удовлетворяли некоторым условиям, в зависимости от используемого метода.
- **Выделение деталей:** детали изображения различного уровня сложности выделяются из видеоданных. Типичными примерами таких деталей являются:
  - Линии, границы и кромки
- **Детектирование/Сегментация:** на определённом этапе обработки принимается решение о том, какие точки или участки изображения являются важными для дальнейшей обработки. Примерами являются:
  - Выделение определённого набора интересующих точек
  - **Сегментация** одного или нескольких участков изображения, которые содержат характерный объект
- **Высокоуровневая обработка:** на этом шаге входные данные обычно представляют небольшой набор данных, например набор точек или участок изображения, в котором предположительно находится определённый объект. Примерами являются:

- Проверка того, что данные удовлетворяют условиям, зависящим от метода и применения
- Оценка характерных параметров, таких как положение или размер объекта
- Классификация обнаруженного объекта по различным категориям

## **2. Сглаживание и повышение резкости цветных изображений.**

### **6.6.1. Сглаживание цветных изображений**

процедура сглаживания черно-белого полутонового изображения может рассматриваться как операция пространственной фильтрации, при которой все коэффициенты фильтрующей маски имеют одинаковые значения. По мере того как маска перемещается по сглаживаемому изображению, значение каждого пикселя заменяется средним значением пикселей в окрестности, определяемой маской.

данная концепция легко распространяется на цветные изображения. Основная разница состоит в том, что вместо скалярных значений яркости мы должны оперировать компонентами векторов вида

$$\mathbf{c}(x,y) = \begin{bmatrix} c_R(x,y) \\ c_G(x,y) \\ c_B(x,y) \end{bmatrix} = \begin{bmatrix} R(x,y) \\ G(x,y) \\ B(x,y) \end{bmatrix}$$

Дальше будет бонусный пример про сглаживанию с помощью усреднения по окрестности, по сути он чисто для того, чтобы самим чуть лучше понять, вряд ли будет юзабельно на экзамене

Пусть  $S_{xy}$  обозначает совокупность координат, определяемых некоторой окрестностью с центром в точке  $(x,y)$  в плоскости цветного RGB-изображения. Среднее значение трехкомпонентного вектора RGB по этой окрестности равно

$$\bar{\mathbf{c}}(x,y) = \frac{1}{K} \sum_{(s,t) \in S_{xy}} \mathbf{c}(s,t).$$

$$\mathbf{c}(x,y) = \begin{bmatrix} c_R(x,y) \\ c_G(x,y) \\ c_B(x,y) \end{bmatrix} = \begin{bmatrix} R(x,y) \\ G(x,y) \\ B(x,y) \end{bmatrix}$$

Из выражения

и свойств векторного сложения следует, что

$$\bar{\mathbf{c}}(x,y) = \begin{bmatrix} \frac{1}{K} \sum_{(s,t) \in S_{xy}} R(s,t) \\ \frac{1}{K} \sum_{(s,t) \in S_{xy}} G(s,t) \\ \frac{1}{K} \sum_{(s,t) \in S_{xy}} B(s,t) \end{bmatrix}$$

Мы видим, что компоненты полученного среднего вектора представляют собой значения скалярных изображений, которые могут быть вычислены независимым сглаживанием каждой цветовой компоненты исходного RGB-изображения с использованием определенной «полутоновой» процедуры обработки

по окрестности. Таким образом, мы приходим к заключению, что сглаживание методом усреднения по окрестности может быть осуществлено отдельно в каждой цветовой плоскости (по каждой цветовой компоненте). При этом получаемый результат совпадает с результатом векторного усреднения, выполняемого в RGB-пространстве.

#### (4) Сглаживание и повышение резкости

Пространственные маски для полутоновых и цветных изображений

$$\bar{\mathbf{c}}(x,y) = \frac{1}{K} \sum_{(x,y) \in S_{xy}} \mathbf{c}(x,y) \quad (1)$$

$$\mathbf{c}(x,y) = \begin{bmatrix} c_R(x,y) \\ c_G(x,y) \\ c_B(x,y) \end{bmatrix} = \begin{bmatrix} R(x,y) \\ G(x,y) \\ B(x,y) \end{bmatrix} \quad \bar{\mathbf{c}}(x,y) = \begin{bmatrix} \frac{1}{K} \sum_{(x,y) \in S_{xy}} R(x,y) \\ \frac{1}{K} \sum_{(x,y) \in S_{xy}} G(x,y) \\ \frac{1}{K} \sum_{(x,y) \in S_{xy}} B(x,y) \end{bmatrix} \quad (2)$$



### 6.6.2. Повышение резкости цветных изображений

Повышение резкости изображения при помощи лапласиана.

Известно, что лапласиан вектора определяется как вектор, компоненты которого равны лапласиану от каждой отдельной скалярной компоненты входного вектора. В цветовом пространстве RGB лапласиан вектора с вида равен

$$\nabla^2 [\mathbf{c}(x, y)] = \begin{bmatrix} \nabla^2 R(x, y) \\ \nabla^2 G(x, y) \\ \nabla^2 B(x, y) \end{bmatrix}.$$

Это означает, что вычисление лапласиана цветного изображения может быть осуществлено путем вычисления лапласиана отдельно для каждой компоненты изображения.



## **БИЛЕТ № 25**

### **1. Частотные фильтры повышения резкости.**

Повышение резкости изображения может быть достигнуто в частотной области при помощи процедуры высокочастотной фильтрации, которая подавляет низкочастотные составляющие.

К фильтрам повышения резкости относятся:

идеальные высокочастотные фильтры, высокочастотные фильтры Баттервортса, и гауссовые высокочастотные фильтры.

Полученные в результате высокочастотной фильтрации изображения имеют одно общее свойство: среднее значение яркости фона близко к нулю. Это возникает вследствие того, что данные фильтры уничтожают нулевую компоненту их Фурье-преобразования.

Гомоморфные фильтры. Данный метод обработки изображений можно представить в виде следующей последовательности преобразований:

Обрабатываемое изображение раскладывается на составляющие, связанные с освещенностью и коэффициентом отражения, после чего гомоморфный фильтр действует на каждую из полученных составляющих по отдельности.

Использование гомоморфного фильтра дает возможность контролировать низкочастотную составляющую преобразования Фурье от логарифма изображения (освещенность) и высокочастотную (коэффициент отражения). Для этого требуется задать передаточную функцию так, чтобы фильтр по-разному воздействовал на низкочастотные и высокочастотные составляющие Фурье-преобразования. Результат работы данного фильтра заключается в одновременном сжатии динамического диапазона и увеличении контраста.

### **2. Цветовая сегментация.**

Сегментация цветного изображения это процесс выделения из пространства изображения одной или нескольких связных областей, удовлетворяющих критерию однородности, основанному на признаках, вычисляемых из значений нескольких цветовых компонентов.

Возможно, наиболее важный признак метода сегментации это способ определения области. Можно выделить 4 основных типа определений:

1. Область это связное подмножество множества пикселов, определенное порождающей функцией, заданной в пространстве цветов. Семейство алгоритмов, использующих такое определение области, называют пиксельными алгоритмами сегментации. Выделяется несколько групп пиксельных алгоритмов:

a). Пороговое отсечение. В данной группе алгоритмов строятся гистограммы цветовых признаков. Для автоматического поиска по гистограммам порогов, как правило, используются алгоритмы, применяемые при обработке полутоновых изображений, например, критерий Отсу.

b). Кластеризация в пространстве цветов. Значения пикселов собираются в группы (кластеры) с одним или более представителем, которые используются далее в процессе классификации.

c). Кластеризация в пространстве цветов с использованием аппарата нечеткой логики (fuzzy logic). Для всех пикселов вычисляются нечеткие функции принадлежности и определяются нечеткие кластеры. Четкие кластеры получаются в результате анализа степени принадлежности. Далее изображение разбивается на связные области.

2. Область это максимальное по площади связное множество пикселов, для которого выполняется свойство однородности. Алгоритмы, использующие такой подход, можно разделить на две группы:

a) Методы наращивания областей. Задается некоторое количество базовых однородных областей, после чего используются различные стратегии для объединения соседних областей.

b) Методы разбиения-слияния. Такие алгоритмы начинают работать с неоднородными областями, разбивая их до получения однородных областей, после чего применяют некоторые эвристические подходы для объединения с целью получить максимальные по площади области.

3. Область это связное множество пикселов, окруженное граничными пикселями, образующими цветовой контур. Области также однородны, т.к. они являются дополнением множества неоднородных областей, образованных граничными пикселями. Часто контур имеет разрывы и необходимо решать задачу замыкания контуров.

Локальные методы для определения граничных точек используют информацию из окрестности этой точки. Глобальные методы выполняют глобальную оптимизацию, и, таким образом, граничная точка может быть определена в результате многих шагов оптимизации, включающих изменения больших участков изображения.

4. Области соответствуют поверхностям или объектам из однородного материала, т.е. области представляют связный материал в сцене. Основное отличие данных методов от вышеперечисленных состоит в том, что учитываются отражающие свойства материалов сцены. Таким образом, применение данного класса алгоритмов ограничивается набором объектов, для которых хорошо известны и могут быть смоделированы их отражающие свойства.

## БИЛЕТ № 26

### 1. Сглаживающие частотные фильтры.

В частотной области сглаживание (размывание) достигается подавлением высоких частот, т. е. при помощи низкочастотной фильтрации.

Широко используются три вида низкочастотных фильтров: идеальный фильтр, фильтр Баттервортса и гауссов фильтр. Эти три фильтра покрывают диапазон от очень резких фильтров (идеальный) до очень гладких фильтров (гауссов). Фильтр Баттервортса характеризуется параметром, который называется порядком фильтра. При больших значениях этого параметра фильтр Баттервортса приближается по форме к идеальному фильтру. При малых значениях он имеет гладкую форму, похожую на форму гауссова фильтра. Таким образом, фильтр Баттервортса может рассматриваться как переходный между двумя «крайностями».

#### 4.8.1. Идеальные фильтры низких частот

Двумерный фильтр низких частот, который пропускает без затухания все частотные составляющие внутри круга радиусом  $D_0$  от начала координат и «обрезает» все частотные составляющие вне этого круга, называется *идеальным фильтром низких частот* (идеальный ФНЧ); он определяется следующей функцией:

$$H(u,v) = \begin{cases} 1 & \text{при } D(u,v) \leq D_0 \\ 0 & \text{при } D(u,v) > D_0, \end{cases} \quad (4.8-1)$$

где  $D_0$  — заданная положительная константа, а  $D(u,v)$  — расстояние от точки  $(u,v)$  частотной области до начала координат (центра частотного прямоугольника).

#### 4.8.2. Фильтры низких частот Баттервортса

Передаточная функция низкочастотного фильтра Баттервортса (ФНЧ Баттервортса) порядка  $n$  с частотой среза на расстоянии  $D_0$  от начала координат задается формулой

$$H(u,v) = \frac{1}{1 + [D(u,v)/D_0]^{2n}}, \quad (4.8-5)$$

#### 4.8.3. Гауссовые фильтры низких частот

Для одномерного случая гауссовые фильтры низких частот (ФНЧ Гаусса) были введены в разделе 4.7.4, где они использовались для того, чтобы установить некоторые важные взаимосвязи между пространственной и частотной областями. В двумерном случае эти фильтры задаются формулой

$$H(u,v) = e^{-D^2(u,v)/2\sigma^2}, \quad (4.8-6)$$

## 2. Основные градационные преобразования

Градационные преобразования изображений являются подгруппой группы пространственных методов улучшения изображений. Термин *пространственная область* относится к плоскости изображения как таковой, и, таким образом, все пространственные методы основаны на прямом манипулировании пикселями изображения.

Градационные преобразования – процедуры, оперирующие непосредственно значениями пикселей изображения. Процессы градационной обработки описываются уравнением:

$$s(x, y) = T[r(x, y)], \quad (1.1)$$

где  $s$  и  $r$  – переменные, обозначающие, соответственно, значения яркостей исходного и результирующего изображений в каждой точке  $(x, y)$ ;  $T$  – некий оператор преобразования над исходным изображением.

### 1. Преобразование изображения в негатив

Преобразование изображения в негатив с яркостями в диапазоне  $[0, L - 1]$  осуществляется с использованием негативного преобразования, определяемого выражением:

$$s = (L - 1) - r. \quad (1.2)$$

Подобный переворот уровней яркости изображения создает эквивалент фотографического негатива. Этот тип обработки особенно подходит для усиления белых или серых деталей на фоне темных областей изображения, особенно когда темные области имеют преобладающие размеры.

### 2. Логарифмическое преобразование

Общий вид логарифмического преобразования, выражается формулой:

$$s = c \lg(1 + r), \quad (1.3)$$

где  $c$  – константа и предполагается, что  $r > 0$ .

логарифмическая функция имеет важную особенность: позволяет сжимать динамический диапазон изображений, имеющих большие вариации в значениях пикселей.

### 3. Степенные преобразования

Степенные преобразования имеют вид:

$$s = cr^\gamma, \quad (1.4)$$

где  $c$  и  $\gamma$  являются положительными константами.

Амплитудная характеристика многих устройств, используемых для ввода, печати или визуализации изображений, соответствует степенному закону. Процедура, используемая для коррекции такой степенной характеристики, называется *гамма-коррекцией*.

Гамма-коррекция необходима, если требуется точное воспроизведение изображения на экране компьютера. Изображения, которые не откорректированы правильно, могут выглядеть или как выбеленные, или, что более вероятно, как слишком темные. Правильное воспроизведение цветов также требует некоторых знаний о гамма-коррекции, поскольку подобное преобразование меняет не только яркость, но также соотношения между красным, зеленым и синим цветами.

### 4. Кусочно-линейные функции преобразования

Подходом, дополняющим рассмотренные методы, является использование кусочно-линейных функций. Главное преимущество кусочнолинейных функций по сравнению с вышерассмотренными состоит в том, что их форма может быть сколь угодно сложной. На самом деле практическая реализация некоторых важных преобразований может быть осуществлена только с помощью кусочно-линейных функций. Основной недостаток кусочно-линейных функций заключается в том, что для их описания необходимо задавать значительно большее количество параметров. Усиление контраста. Одним из простейших случаев использования кусочно-линейных функций является преобразование, усиливающее контрасты. Усиление контраста достигается увеличением динамического диапазона яркостей на обрабатываемом изображении

### 5. Вырезание битовых плоскостей

## **Вырезание битовых плоскостей**

Вместо выделения диапазонов яркостей, может оказаться полезным выделение информации о вкладе тех или иных битов в общее изображение. Если каждый пиксель изображения кодирован 8 битами, то все изображение можно представить себе в виде 8-битовых плоскостей, ранжированных от плоскости 0 с наименее значащими битами до плоскости 7 с наиболее значащими. В терминах 8-битовых байтов плоскость 0 содержит все младшие биты, а плоскость 7 – все старшие биты из составляющих изображение байтов. Можно заметить, что старшие биты (главным образом первые четыре) содержат основную часть визуально значимых данных. Остальные битовые плоскости дают вклад в более тонкие детали изображения.

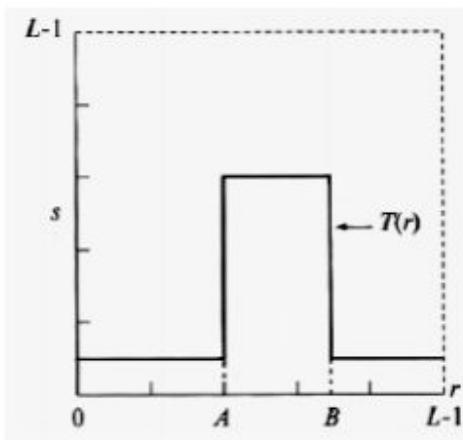


Рис. 1.4 (а). Вырезание диапазона яркостей  $[A, B]$  с приведением остальных значений к константе

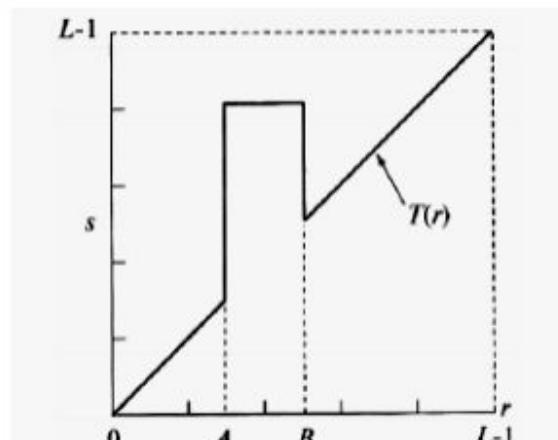


Рис. 1.4 (б). Вырезание диапазона яркостей  $[A, B]$  с сохранением остальных значений

Разделение цифрового изображения на битовые плоскости полезно для анализа относительной информативности, которую несет каждый бит изображения, что позволяет оценить необходимое число битов, требуемое для квантования каждого пикселя, например, при сжатии изображений. На основе извлечения битовых плоскостей из 8-битового изображения нетрудно показать, что (бинарное) изображение битовой плоскости 7 может быть получено обработкой исходного изображения пороговым градационным преобразованием, которое отображает все уровни изображения от 0 до 127 в некоторый один (например, 0), а все уровни от 128 до 255 – в другой (например, 255).

