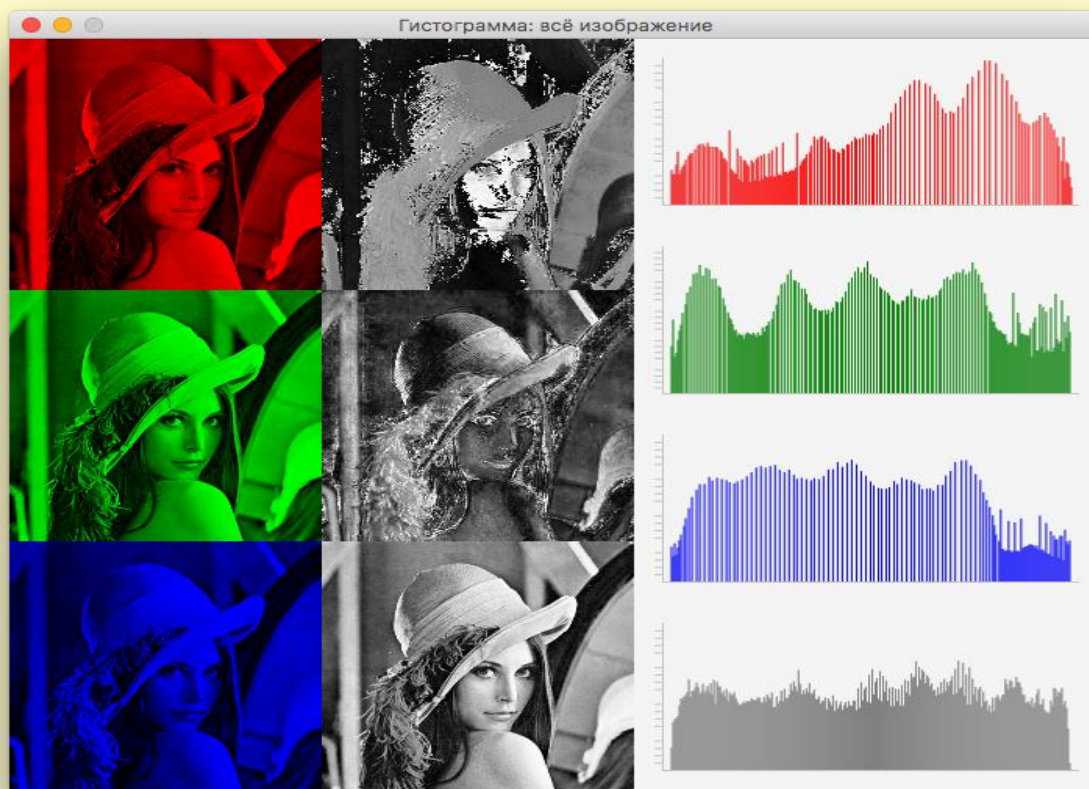


Основы построения систем обработки визуальной информации

- Лк- 20 ч
- Лаб – 28 ч
- Экзамен

ВВЕДЕНИЕ В ОБРАБОТКУ ИЗОБРАЖЕНИЙ



Основные свойства зрительной системы человека

Чувствительность - способность реагировать на внешние изменения.

- Чувствительность характеризуется верхним и нижним абсолютными порогами.
- Существует несколько различных видов чувствительности
 - Световая чувствительность
 - Спектральная чувствительность
- Способность глаза различать минимальные различия яркости смежных областей изображения характеризуется контрастной чувствительностью
- Чувствительность к насыщенности цвета.

Цветоощущение характеризуется тремя основными характеристиками

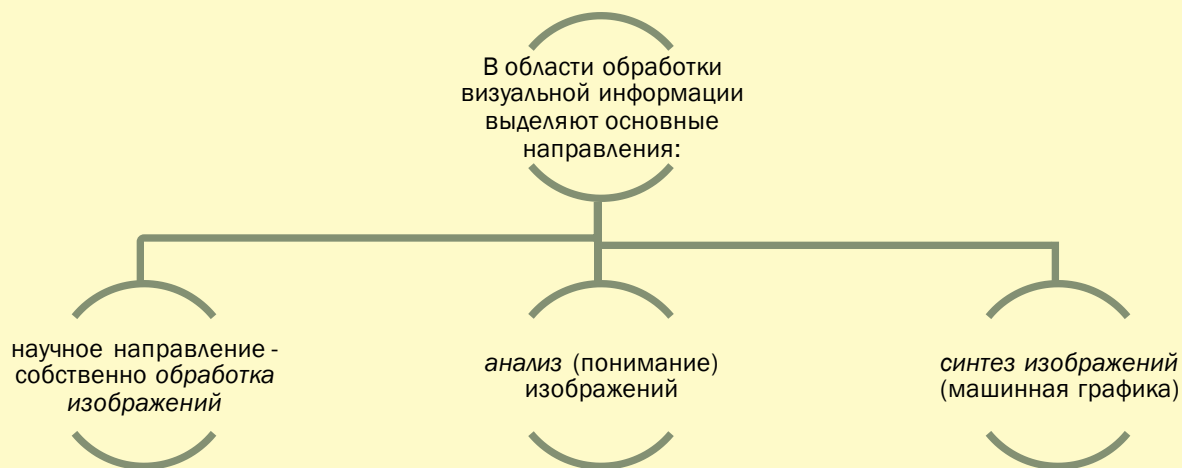
- Светлота
- Цветовой тон
- Насыщенность

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.

Современные задачи в области обработки изображений



Системы обработки изображений

- входные данные и результаты обработки представляются в изобразительной форме

Системы анализа (понимания) изображений

- входные данные представляются в изобразительной форме, а результат обработки – в неизобразительной форме, например в виде текстового описания сцены

Системы синтеза изображений (машинной графики).

- решаются задачи синтеза изображений в их изобразительной форме по некоторому их описанию или алгоритму построения

В качестве основных операций, реализуемых в системах обработки визуальной информации, можно выделить:

оцифровывание

кодирование и сжатие

декодирование и восстановление сжатых изображений

улучшение качества и реставрация изображений

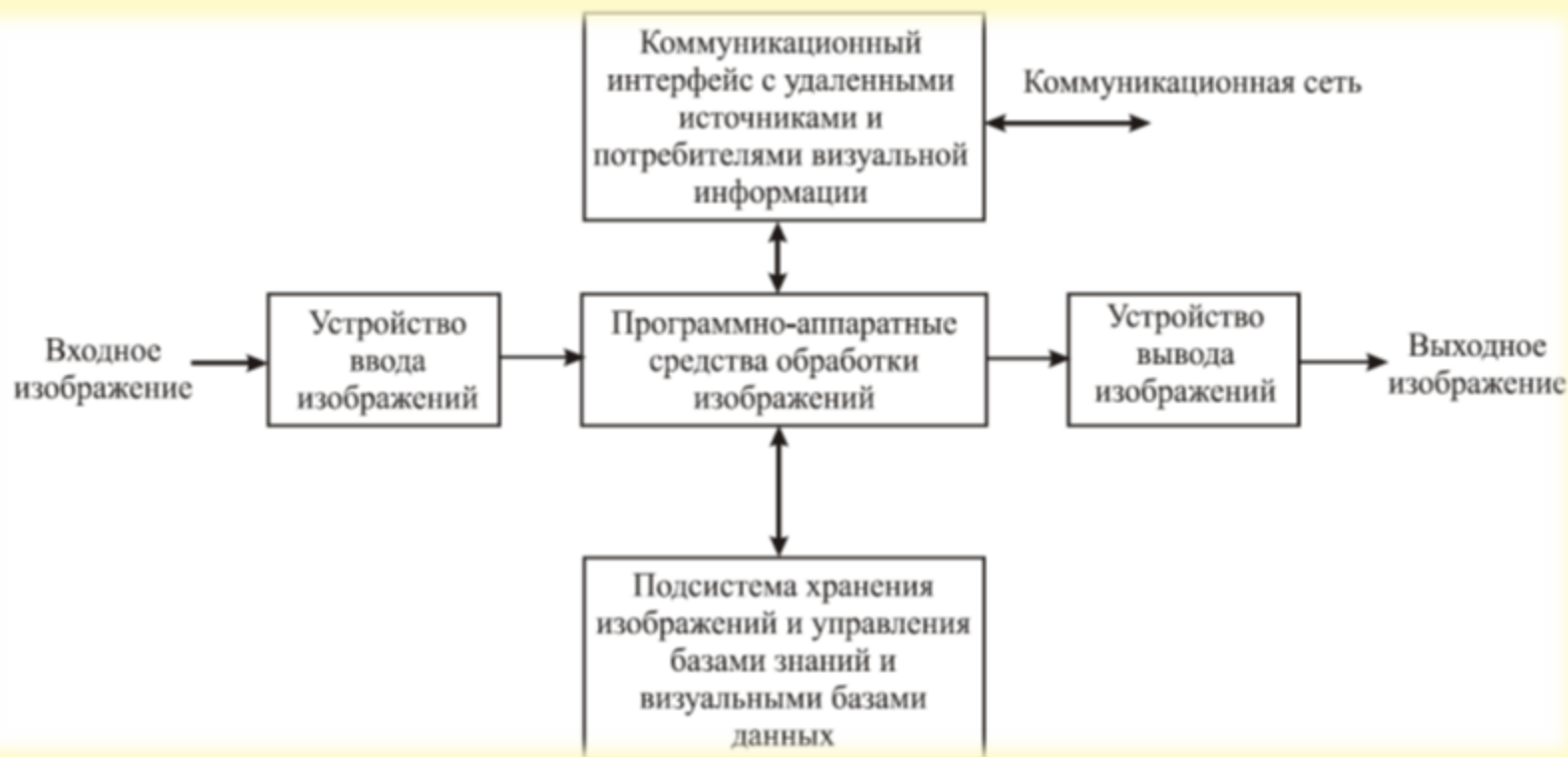
сегментация изображений

анализ изображений

составление описаний изображений

понимание изображений

Управление визуальной информацией



- На вход устройства ввода изображений поступает непрерывное изображение $f(x, y)$, задающее распределение яркости элементов некоторой сцены по пространственным координатам x, y . В устройстве ввода данное изображение затем подвергается дискретизации, т.е. формируется некоторое множество отсчетов функции $f(x, y)$. Далее полученные отсчеты подвергаются квантованию. Таким образом, цифровое изображение представляет собой двумерный массив квантованных отсчетов функции $f(x, y)$.
- Элемент такого массива – элемент цифрового изображения (пиксель)
- Цифровое изображение задается функцией $f(i, j)$ от дискретных аргументов i и j
- Результаты обработки исходных изображений могут представляться в различных формах, в частности списком сегментов изображения (для сегментированного изображения), геометрической моделью в виде совокупности геометрических фигур, текстовым описанием сцены и др.
- Наиболее сложными по своей реализации и выполняемым функциям являются системы понимания изображений.



Признаки исходного изображения

- Значение яркости отдельной точки – $f_{i,j}$. Данный признак является основным, и практически все остальные признаки определяются на его основе.
- Максимальное значение яркости изображения – f_{max} .
- Минимальное значение яркости изображения – f_{min} . Значения f_{max} и f_{min} позволяют определить диапазон изменения яркости изображения.
- Средняя яркость изображения определяется по следующей формуле:

$$f_{cp} = \frac{1}{MN} \sum_{i=1}^N \sum_{j=1}^M f_{ij}$$

- Количество точек определенного значения яркости K_f . Обычно составляется таблица

Значение яркости точек	x_0	x_1	x_2	x_z
Количество точек данной яркости	k_0	k_1	k_2	k_z

- Энергия изображения – E . Определяется следующим образом:

$$E = \sum_{i=1} p_i^2 .$$

- Математическое ожидание яркости изображения μ :

$$\mu = \sum_{i=0}^L p_i f_i ,$$

- Дисперсия значений яркости изображения - σ^2 :

$$\sigma^2 = \sum_{i=0}^L (f_i - \sum_{i=0}^L f_i p_i)^2 p_i ,$$

ТИПИЧНЫЕ ВИДЫ ГИСТОГРАММ

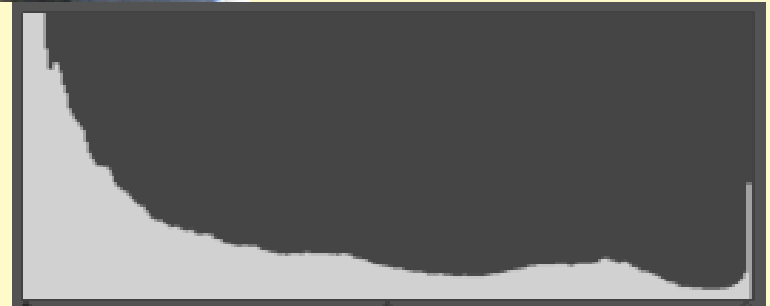
Гистограмма — это график статистического распределения элементов цифрового изображения с различной яркостью, в котором по горизонтальной оси представлена яркость, а по вертикали — относительное число пикселей с конкретным значением яркости.

Изучив гистограмму, можно получить общее представление о правильности экспозиции, контрасте и цветовом насыщении снимка, оценить требуемую коррекцию как при съёмке (изменение экспозиции, цветового баланса, освещения либо композиции снимка), так и при последующей обработке.

Гистограмма темного изображения



Имеет ярко выраженное смещение тонов в сторону теней. Ей характерен широкий и высокий пик в левой части графика. То есть изображение считается темным, если основная масса пикселей имеет низкие уровни яркости.



Гистограмма светлого изображения



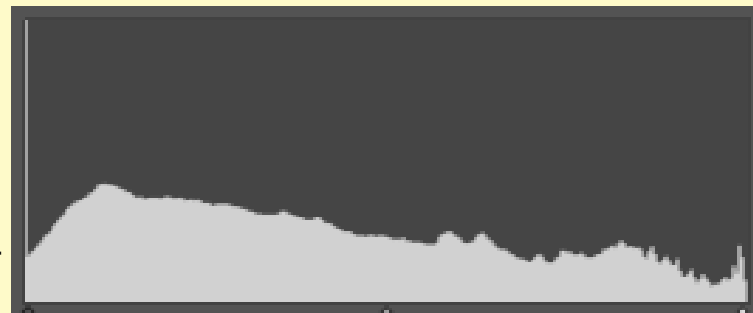
В некотором смысле светлая картинка — прямая противоположность темной. В ее гистограмме будет широкий и высокий «холм» в области светов, определяющий основную массу пикселей изображения, и экспоненциальный спад в тенях.



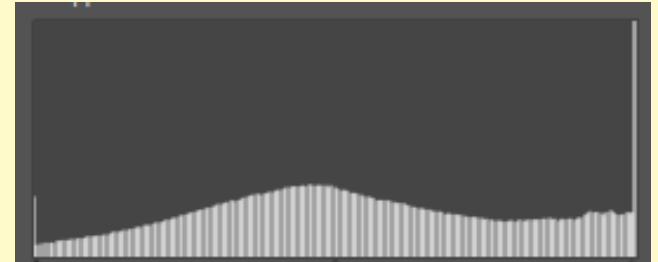
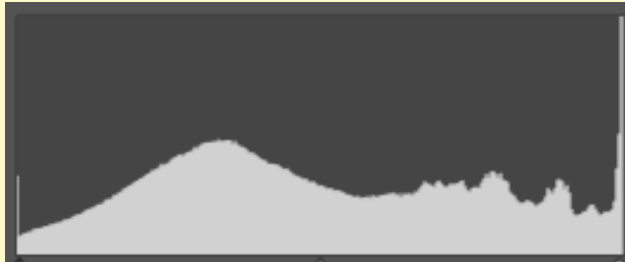
Гистограмма среднего по тону изображения



Изображение может быть средним по тону либо вследствие того, что большинство пикселей имеют среднюю яркость, либо из-за равного соотношения светлых и темных пикселей в целом



Гистограмма «вылета» в светах



При избыточной экспозиции может возникать гистограмма, показанная на рис. Срез справа в области высоких яркостей свидетельствует об утраченных деталях в светах. Яркая засветка привела к тому, что белые части стен превратились в плоское пятно на изображении. Утрата деталей — худшее, что может случиться с цифровым снимком.

Гистограмма изображения с зеркальными бликами

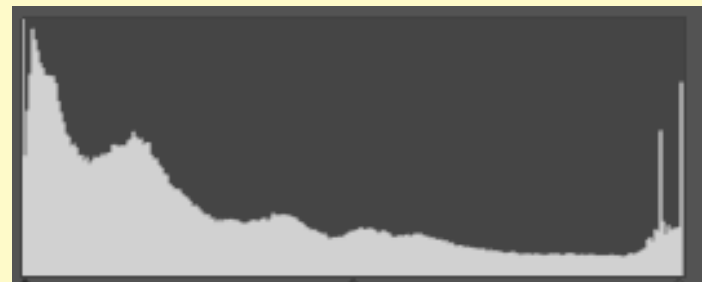
Во многих снимках есть легкий всплеск в области максимальных яркостей, повторяющий в миниатюре гистограмму со срезом в светах. Это не что иное, как следствие зеркальных бликов на поверхности запечатленных объектов. Зеркальный блик — область очень высокой яркости, возникающая в результате отражения света источника на блестящей поверхности. Зеркальные блики есть практически в любой сцене на стеклянных, металлических, гладких пластиковых объектах. Яркость их настолько велика, что никакие ухищрения не позволяют уместить их в доступный тоновый диапазон, и они естественным образом срезаются в области светов — возникает особый пик, который не следует учитывать при качественном анализе гистограмм.



Гистограмма изображения с источниками света в кадре



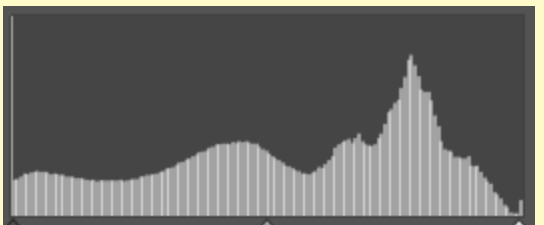
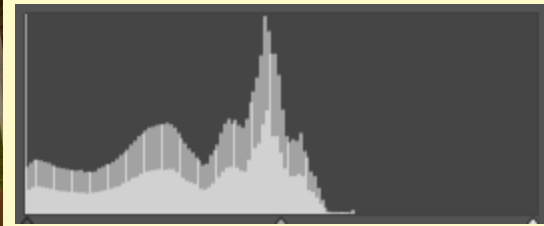
Этот вид во многом похож на гистограмму с зеркальными бликами, с той лишь разницей, что вместо бликов фигурируют сами источники. Как и в предыдущем случае, в области светов есть всплеск, только его размер больше, а сам он несколько шире.



Гистограмма изображения с узким тоновым диапазоном

При недостаточной экспозиции часто возникают изображения, где доступный тоновый диапазон не используется полностью (остаются значительные области в районе высоких яркостей). Из-за этого снимок выглядит темным, хотя следовало бы ожидать средней тональности.

Такая гистограмма свидетельствует о возможном недостаточном контрасте, но главное, что происходит чаще всего, такое изображение содержит всю необходимую информацию о деталях (в области теней наблюдается схождение к левому краю). Последующая тоновая коррекция значительно улучшает его вид



Гистограмма изображения, обработанного цифровыми методами

Такая гистограмма свидетельствует о проведенной коррекции фотографии цифровыми методами. При перераспределении уровней яркости некоторые значения оказываются практически неиспользуемыми, в результате получается «линейчатая» диаграмма.

Оптимальная гистограмма

Оптимальная гистограмма должна максимально эффективно использовать тоновый диапазон, плавно спадать до нуля к краям тонового диапазона (срезов ни в светах, ни в тенях быть не должно), быть относительно плавной, без разрывов. И главное, чтобы тоновый характер, определенный по гистограмме, отвечал ожидаемой тональности изображения.

УЛУЧШЕНИЕ ИЗОБРАЖЕНИЙ

Формирование цифрового изображения в памяти ЭВМ во многих случаях сопровождается его искажением за счет влияния различных фактов.

Наличие помех в исходном изображении существенно усложняет дальнейший его анализ. Поэтому возникают задачи путем **уменьшения**, а в лучшем случае путем **устранения** помех

Задачи улучшения изображения

```
graph TD; A[Задачи улучшения изображения] --> B[Регистрация (коррекция изображений - компенсация изображения, которую создает изображающая система таким образом, чтобы получить идеальное изображение (т. е. изображение при отсутствии этого изображения) или его некоторое (обычно максимальное) приближение.]; A --> C[Преобразование изображений для получения изображений в подходящей для пользователя форме (задачи препарирования изображений)];
```

Регистрация (коррекция изображений - компенсация изображения, которую создает изображающая система таким образом, чтобы получить идеальное изображение (т. е. изображение при отсутствии этого изображения) или его некоторое (обычно максимальное) приближение.

Преобразование изображений для получения изображений в подходящей для пользователя форме (задачи препарирования изображений)

Изменение контраста

Слабый контраст – наиболее распространенный дефект изображения. Обычно он вызван малым динамическим диапазоном изменения яркости либо нелинейностью характеристики передачи уровня яркости.

- При преобразовании градаций яркости выполняется фактически процесс переквантования яркости элемента исходного изображения в новую градацию яркости с целью повышения контраста изображения.

На многих изображениях уровни яркости объектов настолько близки к уровням фонов, что различить их очень трудно.



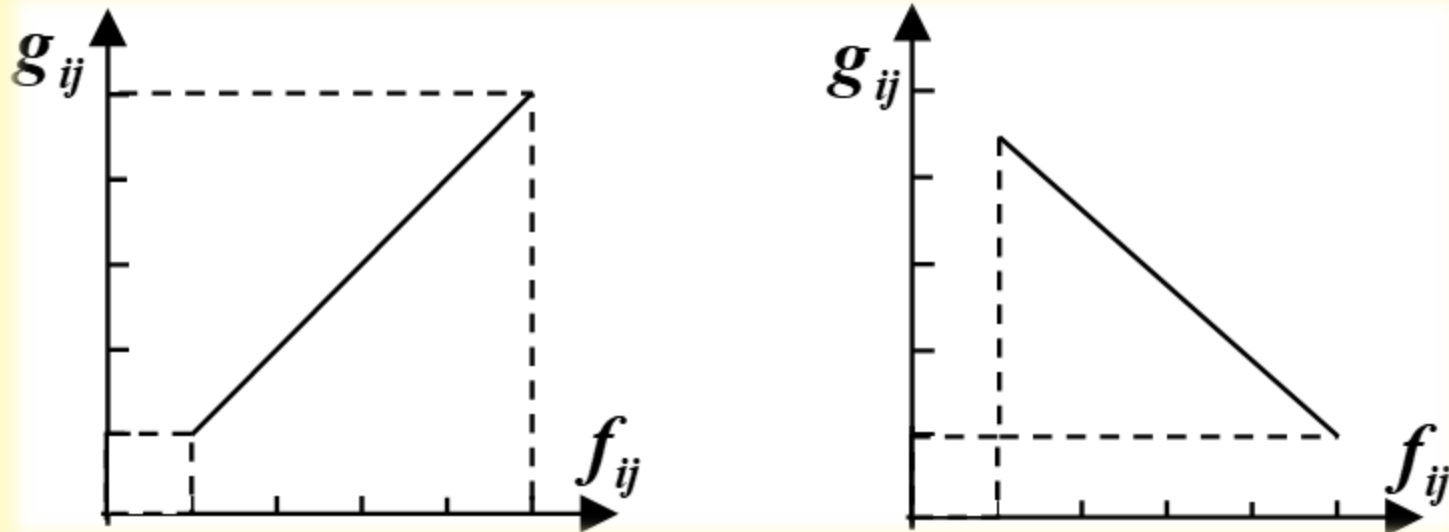
Следовательно, необходимо **улучшить контрастность**, чтобы **увеличить разницу** между градациями яркости объектов фона.

Функциональное отображение градации яркости

Простейшим методом контрастирования является функциональное отображение градации яркости

$$f_{ij} \rightarrow g_{ij}, \text{ т. е. } g_{ij} = R(f_{ij})$$

Для изменения контраста часто используют линейные функциональные отображения



Один из вариантов процедуры **линейного отображения** можно записать следующим образом.

Если z_{min} и z_{max} – соответственно минимальное и максимальное значения яркости изображения на исходном изображении, то преобразование **растягивания диапазона яркостей** $[z_{min}, z_{max}]$ на весь диапазон $[0, k]$ осуществляется следующим образом:

$$g_{ij} = R(f_{ij}) = \left[\frac{k}{z_{max} - z_{min}} \right] \cdot [f_{ij} - z_{min}]$$

$$g_{ij} = R(f_{ij}) = \left[\frac{g_{max} - g_{min}}{f_{max} - f_{min}} \right] \cdot [f_{ij} - f_{min}]$$

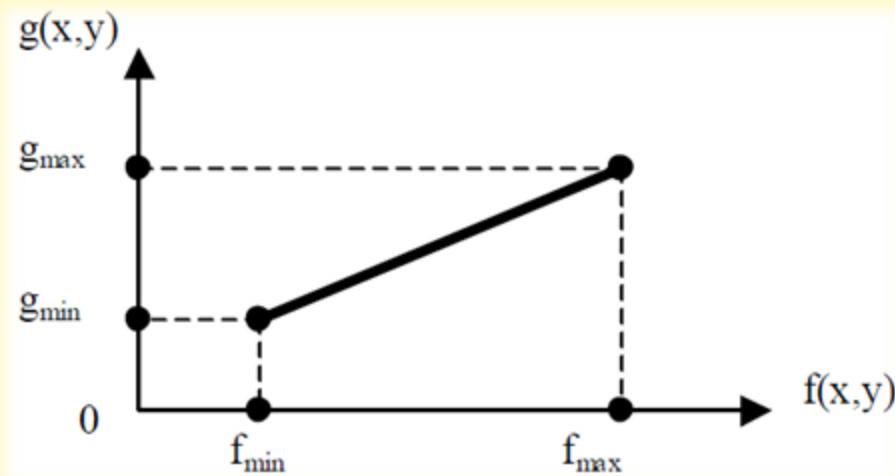
Графически это выражение представляется следующим образом.

→

Наклон линии определяется значением:

$$d = \frac{g_{max} - g_{min}}{f_{max} - f_{min}}$$

Изменение обеспечивает сжатие и растяжение диапазона яркостей. d





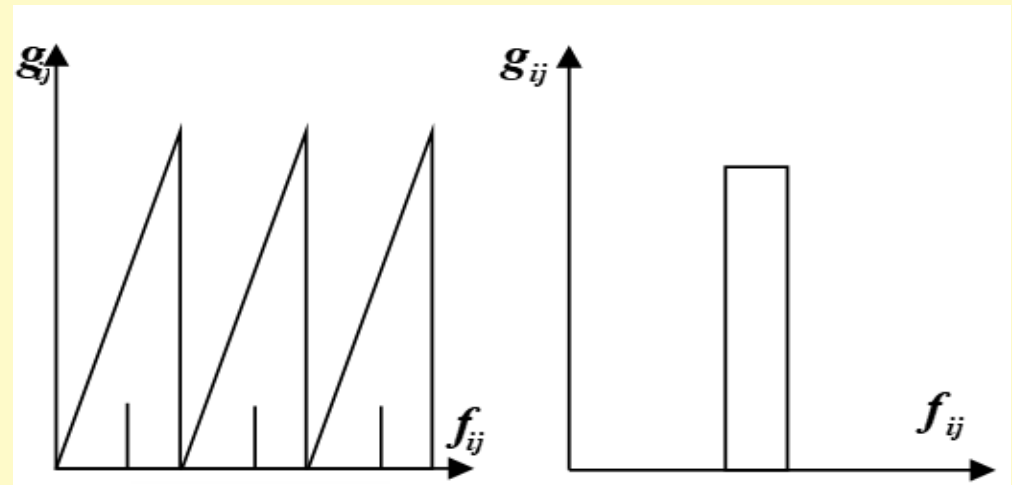
$$X_{min}=180 \quad X_{max}=240$$



$$Y_{min}=0 \quad Y_{max}=255$$

Оператор косочно-линейного контрастирования изображения

- График передаточной функции в этом случае имеет пилообразный вид →
- Характеристика передачи уровней для **яркостного среза** исходного изображения показана на втором рис



Это преобразование позволит выделить основной интервал диапазона яркостей входного изображения.

Широко используется метод яркостных срезов при визуальном анализе изображении в интерактивном режиме.

Соляризация

Суть **соляризация** заключается в перемножении яркостей, получающегося исходного изображения f_{ij} и его негатива $g_{ij}=k-f_{ij}$ с последующим растяжением диапазона яркостей получающегося изображения на весь диапазон от 0 до k .

Преобразование осуществляется по формуле:

$$g_{ij} = \frac{1}{k^2} f_{ij} [k - f_{ij}]$$

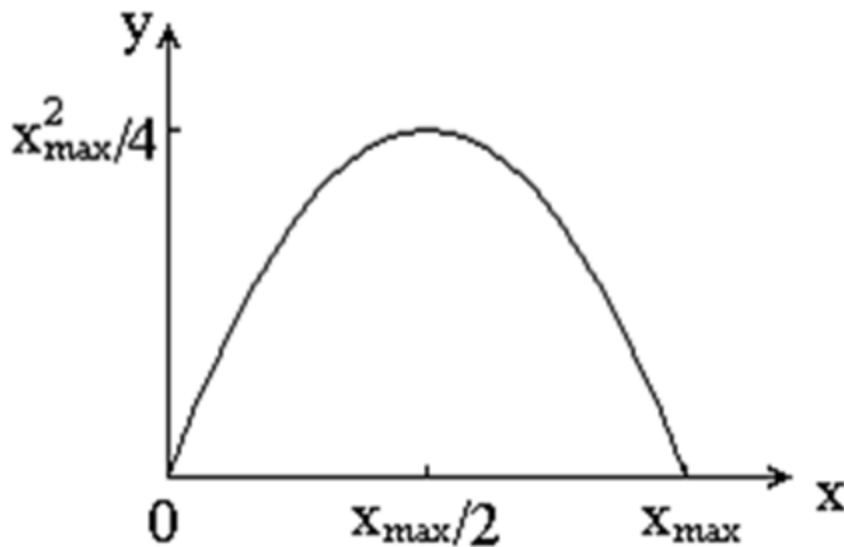
k - нормирующий множитель, приводящий диапазон яркостей исходного изображения и его негатива к отрезку $[0;1]$

Максимум, будет достигаться при $f_{ij}=k/2$. Это преобразование значительно увеличивает яркость точек исходного изображения, обладающих яркостью, близкой к $k/2$

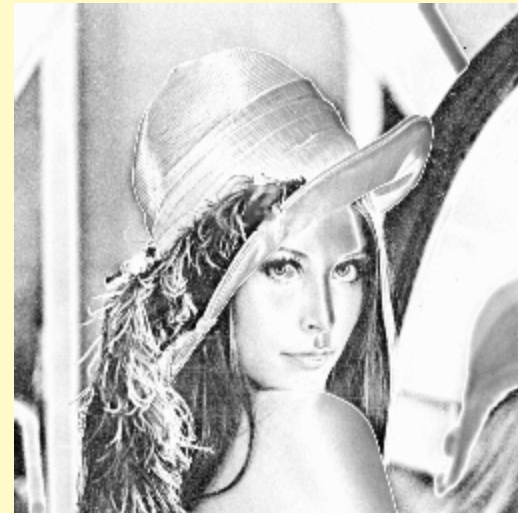
$$y = k \cdot x \cdot (x_{max} - x)$$

где x_{max} - максимальное значение исходного сигнала, a - константа, позволяющая управлять динамическим диапазоном преобразованного изображения.

Функция, описывающая данное преобразование, является квадратичной параболой, ее график при $k=1$ приведена ниже ↓. При $y_{max}=x_{max}$ динамические диапазоны изображений совпадают, что может быть достигнуто при $k=4/x_{max}$.



Смысл соляризации заключается в том, что участки исходного изображения, имеющие уровень белого или близкий к нему уровень яркости, после обработки имеют уровень черного. При этом сохраняют уровень черного и участки, имеющие его на исходном изображении. Уровень же белого на выходе приобретают участки, имеющие на входе средний уровень яркости (уровень серого)



Псевдостереоскопический эффект

- Это операция аналогична преобразованию **солиризации** с той лишь разницей, что негатив (позитив) сдвигается относительно позитива (негатива) перед наложением на несколько элементов диагонали, т. е. в результате этого преобразования получатся $g_{if} = R \left(\left(\frac{1}{k^2} \right) \cdot$

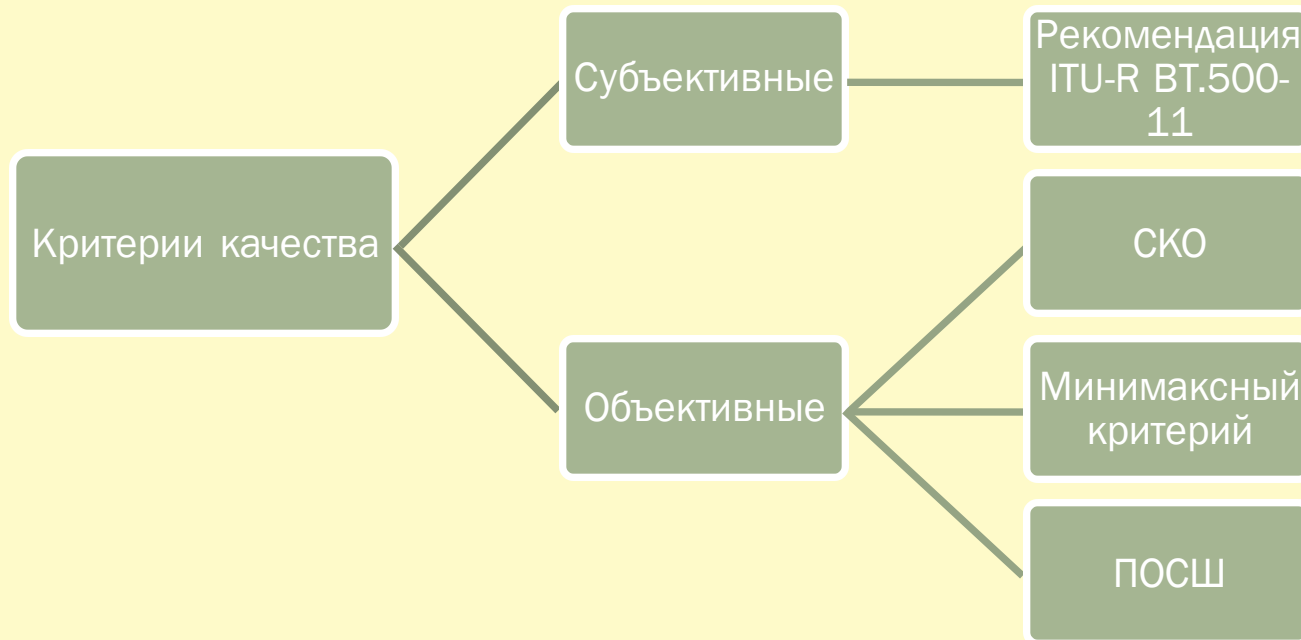
$$f_{i+\alpha, j+\alpha} \cdot [k - f_{ij}] \right)$$

- Обычно α не велико от одного до трех элементов

ОЦЕНКА КАЧЕСТВА ЦИФРОВЫХ ИЗОБРАЖЕНИЙ

На современном этапе развития цифровой обработки изображений существует обширное множество критериев оценки их качества

Классификация критериев оценки качества изображений



Качество

Характеристика самого изображения

Мера близости двух изображений



СУБЪЕКТИВНЫЕ КРИТЕРИИ КАЧЕСТВА

Два основных вида экспертных оценок

- абсолютные
- сравнительные

На результаты субъективных экспертных оценок влияют

- характер рассматриваемых изображений
- окружающая обстановка

Шкала погрешностей, составленная Мерцем, Фаулером и Кристофером (применяется, например, в методе DSIS (Double-stimulus impairment scale – метод двойного воздействия со шкалой искажений).

1. Незаметные
2. Еле заметные
3. Вполне заметные, но слабо ухудшающие изображение
4. Ухудшающие изображение, но допустимые
5. Несколько нежелательные
6. Определенно нежелательные
7. Крайне нежелательные

Результаты экспертных оценок обычно выражают с помощью среднего балла, определяемого как

$$\bar{C} = \frac{\sum_{k=1}^K n_k C_k}{\sum_{k=1}^K n_k},$$

где K – количество категорий, n_k – число изображений, отнесенных к k -й категории, а C_k – соответствующий ей балл.

главный недостаток субъективного критерия – отсутствие количественных оценок

ОБЪЕКТИВНЫЕ КРИТЕРИИ КАЧЕСТВА

Сложность и дороговизна субъективных экспертных оценок делают привлекательным автоматическое измерение качества с помощью определенных правил

- Объективные меры качества должны хорошо согласовываться с результатами субъективных оценок для широкого класса изображений, не подразумевая при этом проведение чересчур сложных вычислений.
- Желательно, чтобы эти метрики имели простую аналитическую форму, и их можно было бы применять в качестве критериев оптимальности при выборе параметров системы обработки изображений.

На данный момент наиболее популярной объективной мерой служит пиковое отношение сигнал/шум (ПОСШ)

$$ПОСШ = 20 \log_{10} \frac{255}{\sqrt{СКО}},$$
$$СКО = \frac{1}{N} \sum_{i=1}^N (x_i - y_i)^2,$$

где

- N – число пикселей в изображении,
- x_i и y_i – значения пикселей исходного и восстановленного изображений соответственно, а СКО – среднеквадратичная ошибка.

Минимаксный критерий, или критерий максимального правдоподобия (maximum likelihood)

$$ML = \max_{i=1, \dots, N} |x_i - y_i|.$$

Этот критерий используется в тех случаях, когда выдвигается требование высокой точности представления не изображения в целом, а каждой его точки (пикселя).

МОДЕЛИ ШУМОВ НА ИЗОБРАЖЕНИЯХ И ИХ ОСНОВНЫЕ ХАРАКТЕРИСТИКИ

Основные источники шума на цифровом изображении – это сам процесс его получения (оцифровки), а также процесс передачи

Важными являются параметры, определяющие пространственные характеристики шума, а также вопрос, коррелирует ли шум с изображением. Под частотными характеристиками понимают свойства спектра шума в смысле преобразования Фурье.

ТИПЫ ШУМОВ

Знания относительно системы получения изображений, визуальный анализ зашумленного изображения помогают в построении модели шума и оценке его статистических характеристик

Математически заданная модель шума помогает построить оптимальную для данного приложения систему восстановления изображений

Аддитивный шум описывается соотношением $g(x, y) = f(x, y) + \eta(x, y)$, где $f(x, y)$ – исходное изображение, $g(x, y)$ – зашумленное изображение, $\eta(x, y)$ – аддитивный и не зависящий от сигнала шум с гауссовым (или другим) распределением функции плотности вероятности.

Мультипликативный шум - описывается соотношением $g(x, y) = f(x, y) \cdot \eta(x, y)$. (зернистость фотопленки; изображения, полученные с помощью ультразвука, радиолокационные изображения)

Импульсный шум - описывается соотношением $g(x, y) = (1 - p) f(x, y) + p \cdot i(x, y)$, где $i(x, y)$ – модель импульсного шума, p – бинарный параметр, принимающий значения 0 и 1.

Шум квантования - является шумом, зависящим от сигнала, и характеризуется выбранным шагом квантования. Такой шум может производить нежелательные артефакты, например, ложные контуры вокруг объектов, или устранять низкоконтрастные детали на изображении

Спекл-шум - класс шумов, зависящих от самого сигнала изображения. Для объектов с низким разрешением обычно является **мультипликативным**. Одним из характерных примеров возникновения спекл-шума является получение цифровых изображений с помощью оптического сканера.

ПРИМЕРЫ ФУНКЦИЙ ПЛОТНОСТИ РАСПРЕДЕЛЕНИЯ ВЕРОЯТНОСТЕЙ ДЛЯ ОПИСАНИЯ МОДЕЛЕЙ ШУМА

Предполагается, что шум не зависит от пространственных координат и не коррелирует с самим изображением (т.е. между значениями элементов изображения и значениями шумовой составляющей нет корреляции)

В нашем случае имеем дело с описанием поведения шума в пространственной области, которое основано на статистических свойствах значений яркости компоненты шума. Эти значения яркости могут рассматриваться как случайные величины, характеризующиеся функцией распределения вероятностей

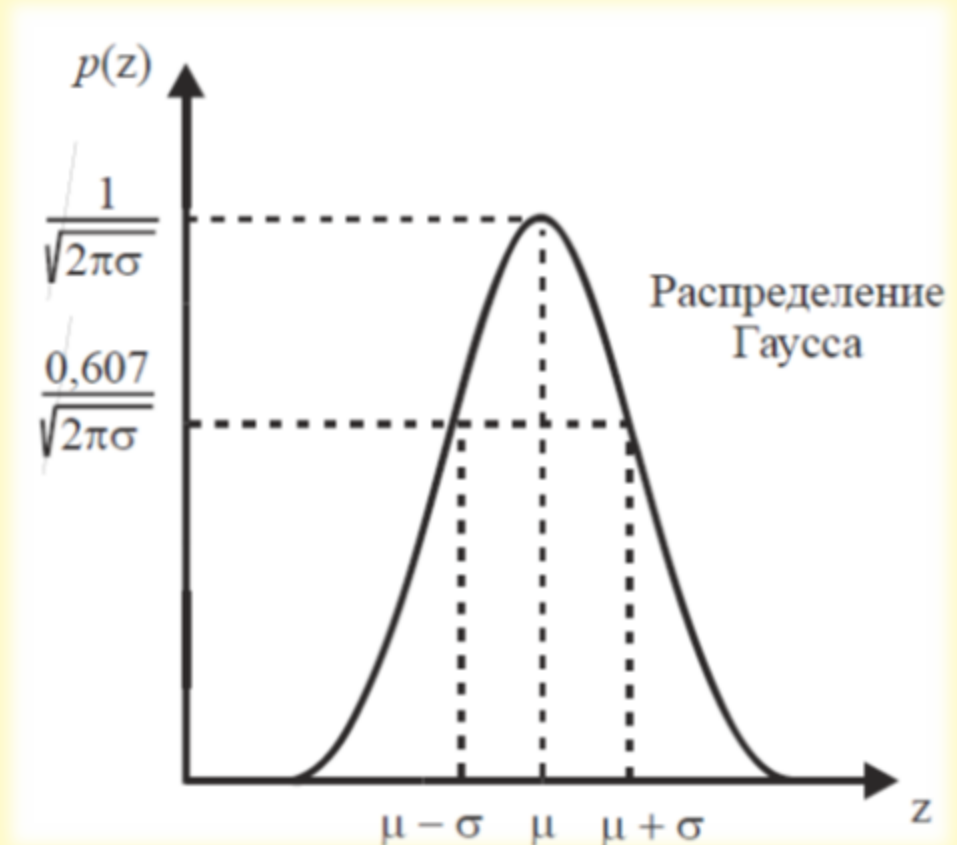
Гауссов шум

Гауссов шум возникает на изображении в результате возникновения таких факторов, как шум в электрических цепях, шум сенсоров – из-за недостатка освещения и/или высокой температуры.

Функция плотности распределения вероятностей гауссовой случайной величины z задается следующим выражением:

$$p(z) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(z-\mu)^2/2\sigma^2}$$

Если случайная величина z описывается этим выражением, то приблизительно 70% её значений сосредоточено в диапазоне $[(\mu - \sigma), (\mu + \sigma)]$ и около 95% в диапазоне $[(\mu - 2\sigma), (\mu + 2\sigma)]$.



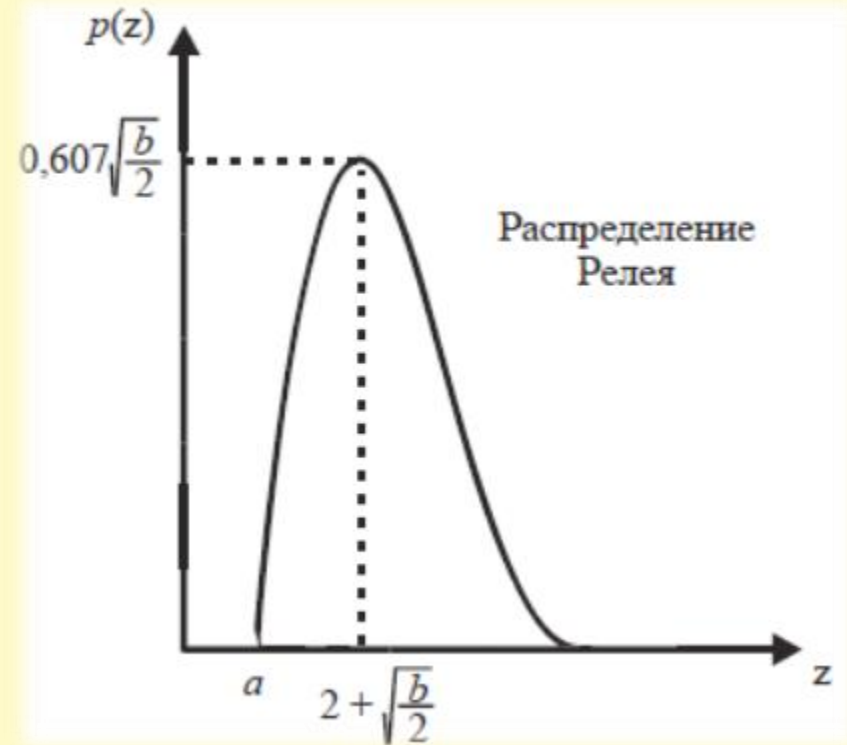
Шум Релея

- Распределение Релея полезно при моделировании шума, который возникает на снимках, снятых с большого расстояния.
- Функция плотности распределения вероятностей для шума Релея имеет вид

$$p(z) = \begin{cases} \frac{2}{b}(z-a) e^{-(z-a)^2/b}, & z \geq a \\ 0, & z < a \end{cases}$$

$$\mu = a + \sqrt{\pi b/4},$$

$$\sigma^2 = \frac{b(4-\pi)}{4}.$$

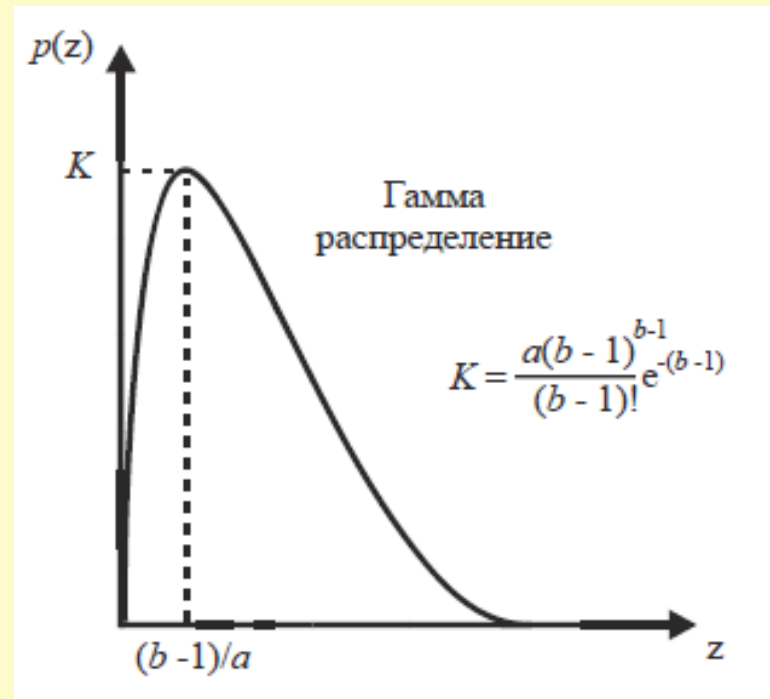


Шум Эрланга (гамма-шум)

$$p(z) = \begin{cases} \frac{a^b z^{b-1}}{(b-1)!} e^{-az}, & z \geq 0 \\ 0, & z < 0 \end{cases}$$

$$\mu = \frac{b}{a},$$
$$\sigma^2 = \frac{b}{a^2},$$

$$K = \frac{a(b-1)^{b-1}}{(b-1)!} e^{-(b-1)}$$



Экспоненциальный шум

$$p(z) = \begin{cases} ae^{-az}, & z \geq 0 \\ 0, & z < 0 \end{cases}$$

$$\mu = \frac{1}{a},$$

$$\sigma^2 = \frac{1}{a^2}.$$

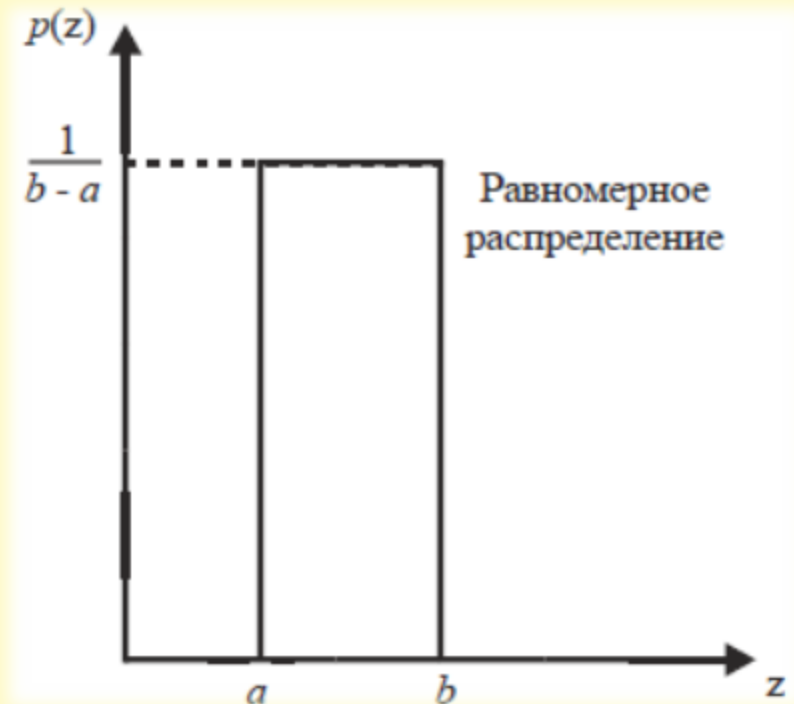


Белый шум

$$p(z) = \begin{cases} \frac{1}{b-a}, & z \in [a, b] \\ 0, & z \notin [a, b] \end{cases}$$

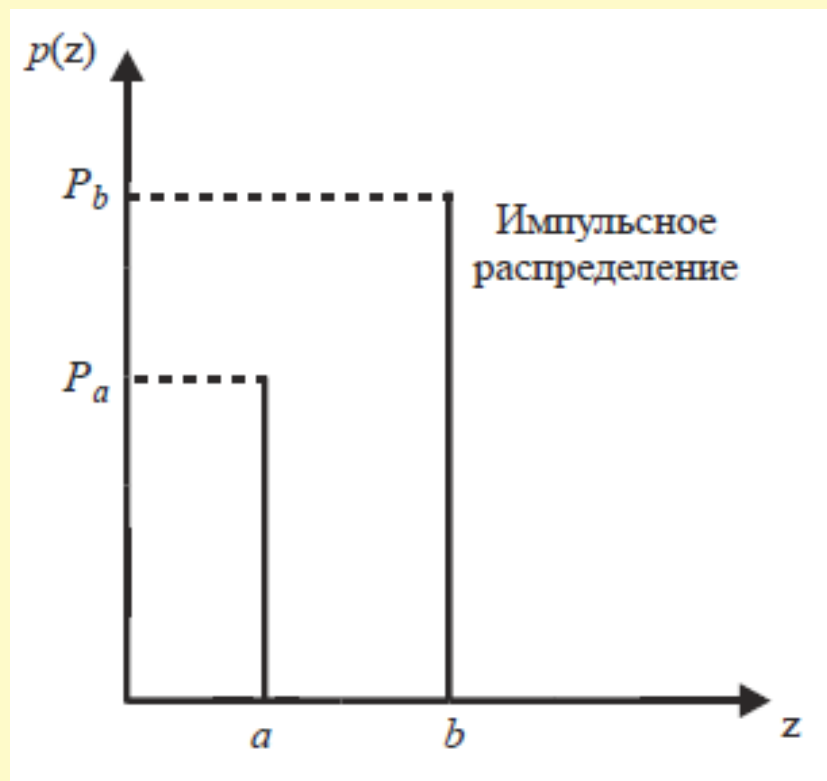
$$\mu = \frac{a+b}{2},$$

$$\sigma^2 = \frac{(b-a)^2}{12}.$$



Импульсный шум

$$p(z) = \begin{cases} P_a, & z = a \\ P_b, & z = b \\ 0, & z \neq a, z \neq b \end{cases}$$



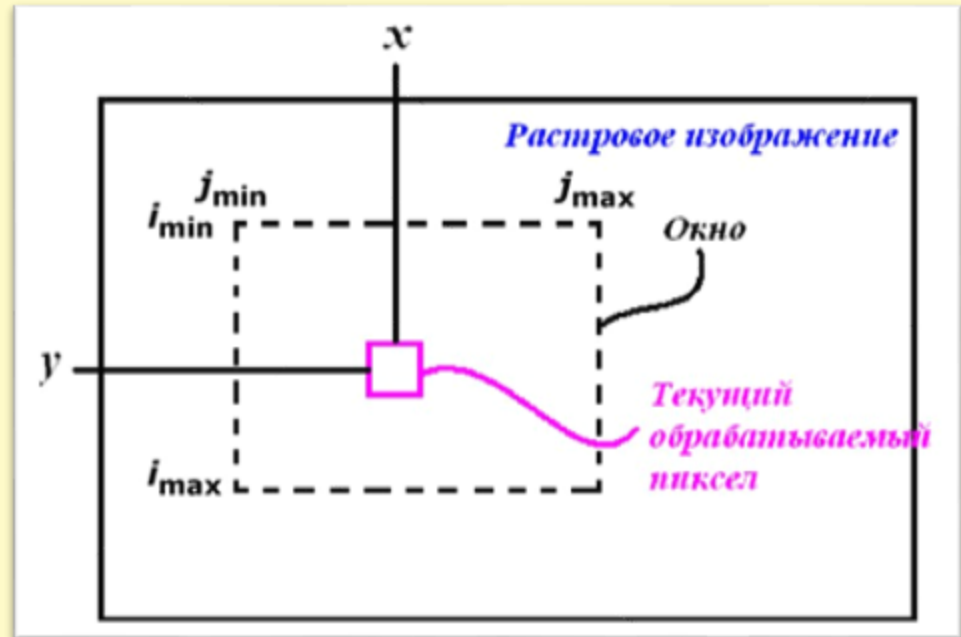
ЛОКАЛЬНАЯ ФИЛЬТРАЦИЯ ИЗОБРАЖЕНИЙ

Алгоритмы локальной фильтрации объединены общей идеей фильтрации скользящим окном, или апертурой.

Цели локальной фильтрации обычно состоят в улучшении качества изображения; чаще всего это устранение помех или повышение резкости, подчеркивание контуров.

- Берется небольшой, обычно прямоугольный, участок плоскости и на нем определяется некоторая функция. Упомянутый участок называется **апертурой**, или **окном**, а заданная на нем функция - **весовой функцией**, или **функцией окна**.
- Каждому элементу апертуры соответствует (присваивается пользователем) определенное число, называемое в дальнейшем **весовым множителем**. Совокупность всех весовых множителей и составляет **весовую функцию**. Апертура вместе с заданной на ней весовой функцией часто называется **маской**.

- Обычно используются апертуры небольшого размера 3×3 элемента; увеличение размеров апертуры существенно увеличивает объем вычислений, в то время как качество обработки улучшается незначительно.
- Линейные размеры апертуры обычно берутся нечетными, чтобы можно было однозначно указать ее центральный элемент.



Виды локальной фильтрации

В зависимости от того, куда (в какое поле) записывается отклик фильтра, различают **простые** и **рекурсивные** фильтры

- В **простых (нерекурсивных)** фильтрах отклик записывается в **выходное** изображение,
- В **рекурсивных** он записывается обратно в **исходное** изображение, изменяя значение пикселей непосредственно в процессе фильтрации. Иногда оказывается полезным многократное повторение процедуры фильтрации; в этих случаях говорят об итеративном применении фильтров (не путать с рекурсивными фильтрами!).

В локальной фильтрации используются **линейные** и **нелинейные** фильтры.

- В **линейных** фильтрах отклик является линейной функцией нескольких переменных, роль которых играют попавшие в окно пиксели. Весовые множители – это коэффициенты упомянутой линейной функции.
- Фильтры, в которых отклик не может быть выражен линейной функцией от значений элементов изображения, являются по определению **нелинейными**.

При линейной фильтрации с помощью **ОКОННЫХ** функций различают:

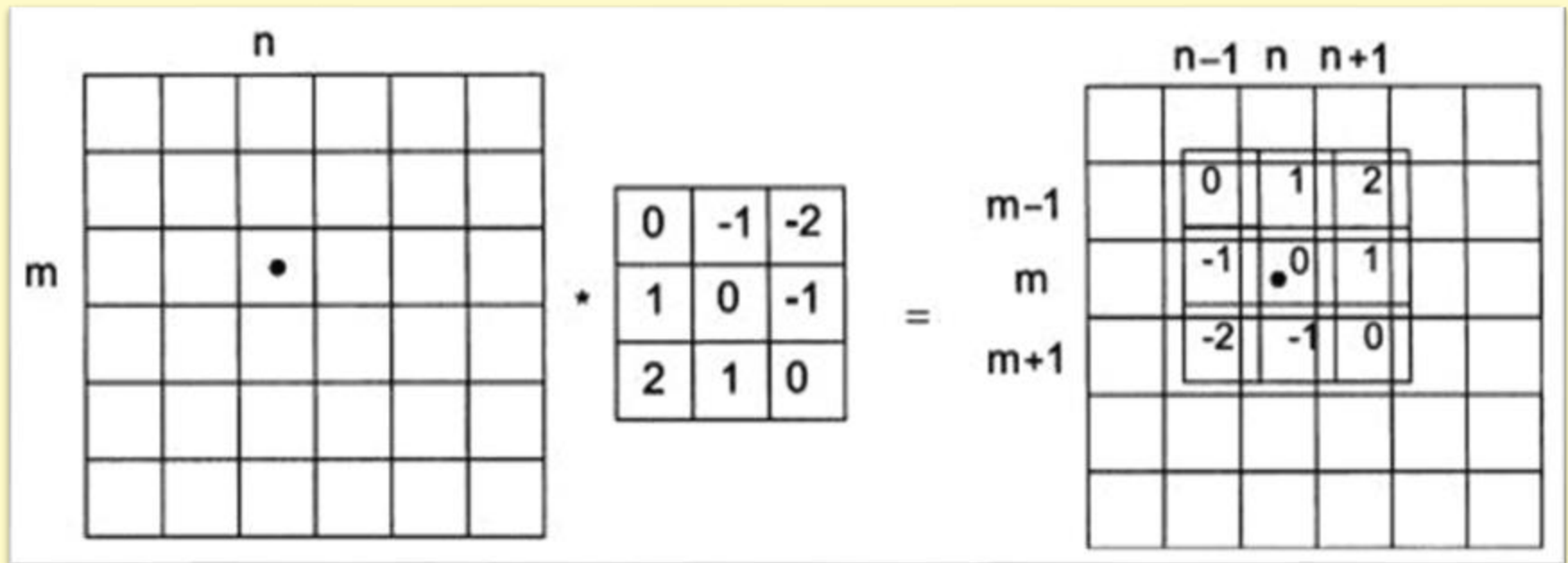
- 1. сумма коэффициентов SK оконной функции отлична от нуля. Импульсный отклик такого фильтра имеет **низкочастотный** характер и поэтому содержит постоянную составляющую, т.е. фоновую компоненту. Для получения этой компоненты значения коэффициентов следует **делить на величину SK** ;
- 2. сумма коэффициентов $SK = 0$. Фильтр имеет **высокочастотный** характер и служит для выделения верхних пространственных частот (повышения **резкости** изображений, **выделения контуров** и т.д.). Импульсный отклик содержит отрицательные составляющие, которые не могут быть представлены в результирующем изображении с положительными значениями полутонов. Поэтому при воспроизведении результирующего изображения необходимо сдвинуть уровень черного на величину $G/2$ и уменьшить динамический диапазон изменения яркости до $G/2$.

Существует так называемый **краевой эффект**, который редко упоминается: способ обработки пикселей, лежащих вблизи границ поля, когда маска фильтра простирается за пределы изображения

Особенно
ощутимо на
небольших
изображениях.

Если размер
изображения много
больше размера окна,
то доля площади
профильтрованного
поля, на которой
заметны краевые
эффекты, мала. В

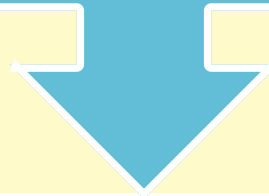
- Отклик линейного фильтра линейным образом **зависит от обрабатываемого изображения**.
- Алгоритм линейной фильтрации: последовательность значений выходного изображения Q формируется путем дискретной свертки входного поля f с функцией окна $h(i, j)$, для задания которой используются локальные операторы



Фильтрация производится путем перемещения слева направо (или сверху вниз) окна (апертуры) на один пиксель.



При каждом положении окна осуществляется перемножение весовых коэффициентов с соответствующими значениями исходного изображения, суммирование полученных произведений, нормирование результата суммирования (деления на нормирующий коэффициент) и присвоение центральному элементу окна полученного значения.



Центр здесь выбирается потому, что в ином случае (если, например, присваивать значение отклика пикселю, соответствующему левому верхнему элементу окна) фильтрация сопровождается **нежелательным сдвигом** профильтрованного изображения относительно исходного.

$$Q(P, Q) = \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} f(p - I + i, q - J + j)h(i, j)$$

Рекурсивные линейные фильтры

Фильтр, в котором отклик определяется только через входные значения, называется простым, или нерекурсивным. Именно такими были рассмотренные выше линейные фильтры.

Рекурсивным называется фильтр, в котором отклик определяется не только через входные значения, но и через выходные

В рекурсивных фильтрах могут использоваться те же весовые функции, что и в нерекурсивных.

- **Рекурсивным фильтром 1-го рода** называется такой фильтр, у которого отклик в каждом положении окна формируется так же, как у простого фильтра, но записывается обратно во входной массив, который теперь одновременно играет роль выходного

$$F(P, Q) = \sum_{i=1}^m \sum_{j=1}^n f(p - I + i, q - J + j)h(i, j)$$

- **Рекурсивным фильтром 2-го рода** называется комбинированный фильтр, у которого отклик формируется как взвешенная сумма откликов нерекурсивного фильтра и рекурсивного фильтра 1-го рода:

$$R(P, Q) = kQ(p, q) + (1 - k)F(p, q)$$

k - коэффициент, определяющий вклады каждого из двух упомянутых фильтров



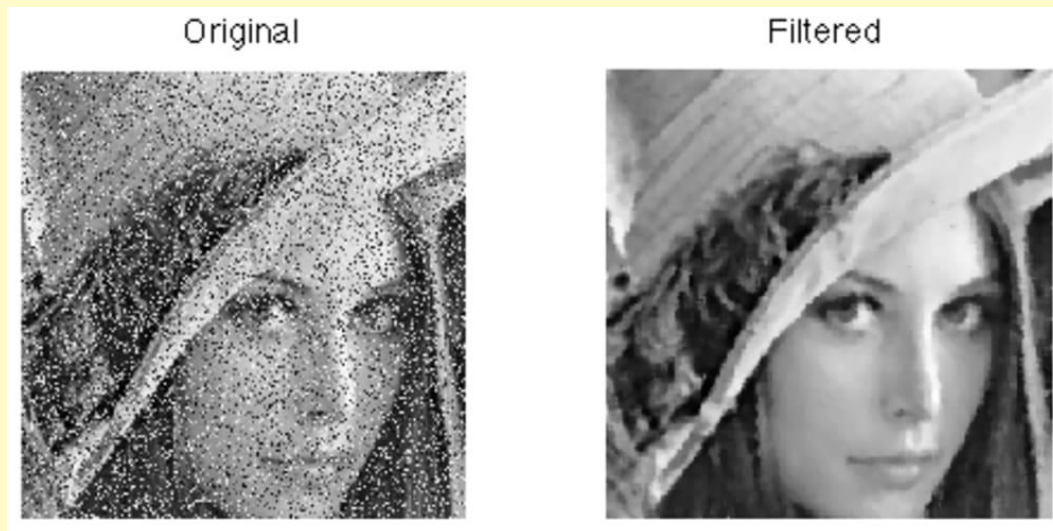
ЛИНЕЙНЫЕ ФИЛЬТРЫ



Подавление шумов и сглаживание

Равномерный фильтр-маска

- Поскольку шум пространственно декоррелирован, в его спектре, как правило, содержатся более **высокие пространственные частоты**, чем в спектре обычного изображения.
- Простая **низкочастотная** пространственная фильтрация может служить эффективным средством подавления шумов.



$$h_{s1} = \frac{1}{9} \cdot \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

Слаживающий оконный фильтр с положительными элементами

$$h_{s2} = \frac{1}{16} \cdot \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$$

$$\frac{1}{256} \cdot \begin{bmatrix} 1 & 4 & 6 & 4 & 1 \\ 4 & 16 & 24 & 16 & 4 \\ 6 & 24 & 36 & 24 & 6 \\ 4 & 16 & 24 & 16 & 4 \\ 1 & 4 & 6 & 4 & 1 \end{bmatrix}$$

Биномиальный фильтр (фильтр Гаусса)

Другие низкочастотные фильтры-маски

$$h_{S3} = \frac{1}{10} \cdot \begin{vmatrix} 1 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 1 \end{vmatrix}$$

$$y = 1/10 (x_{p-1,q-1} + x_{p-1,q} + x_{p-1,q+1} + x_{p,q-1} + 2x_{p,q} + x_{p,q+1} + x_{p+1,q-1} + x_{p+1,q} + x_{p+1,q+1})$$

где $x_{p,q}$ – элементы входного (исходного) изображения, попавшие в окно фильтра. Благодаря большему весовому коэффициенту центральный элемент изображения, попадающий в окно, выделяется сильнее.

$$h_{S4} = \frac{1}{16} \cdot \begin{vmatrix} 2 & 1 & 2 \\ 1 & 4 & 1 \\ 2 & 1 & 2 \end{vmatrix}$$

С помощью рассмотренной выше биномиальной маски h_{S2} , благодаря большим весам, подчеркиваются четырехсвязные элементы исходного изображения, т.е. горизонтальные и вертикальные линии.

Если необходимо подчеркнуть диагональные линии, то следует выделить восьмисвязные элементы исходного изображения, не являющиеся четырехсвязными. Для этого целесообразно применять оконную функцию h_{S4} .

МЕДИАННАЯ ФИЛЬРАЦИЯ

- Медианная фильтрация - нелинейный фильтр, представляющий собой технологию нелинейной обработки сигналов, которая может эффективно подавлять шумы на основе статистической теории сортировки.
- Основная идея – использование медианного значения в окрестности пикселя для замены значения пикселя.
- Метод может сохранить детали края изображения, устраняя импульсный шум..
- Медианный фильтр должен сначала сгенерировать шаблон фильтра, отсортировать значения пикселей в шаблоне и сгенерировать монотонно возрастающую или монотонно падающую двумерную последовательность данных.
- $g(x, y) = \text{med}\{f(x-k, y-l), (k, l \in w)\}$

где $f(x, y)$ и $g(x, y)$ - исходное изображение и обработанное изображение, w - входной двумерный шаблон

Принципиальная схема медианной фильтрации

- Сначала мы рассмотрим медиану.
- В серии чисел {1, 4, 6, 8, 9} число 6 является медианой ряда чисел. Из этого мы можем обратиться к обработке изображений.
- В матрице 3 * 3 в изображении, которая имеет 9 пикселей, мы сортируем 9 пикселей и, наконец, назначаем центр этой матрицы медиане этих девяти пикселей.

$$\text{mid} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

http://blog.csdn.net/weixin_37720172

$g = \text{median}[f(x-1,y-1) + f(x,y-1) + f(x+1,y-1) + f(x-1,y) + f(x,y) + f(x+1,y) + f(x-1,y+1) + f(x,y+1) + f(x+1,y+1)]$

10	9	4	7	2
4	31	12	23	30
5	43	1	32	8
54	7	3	43	4
23	76	56	45	9



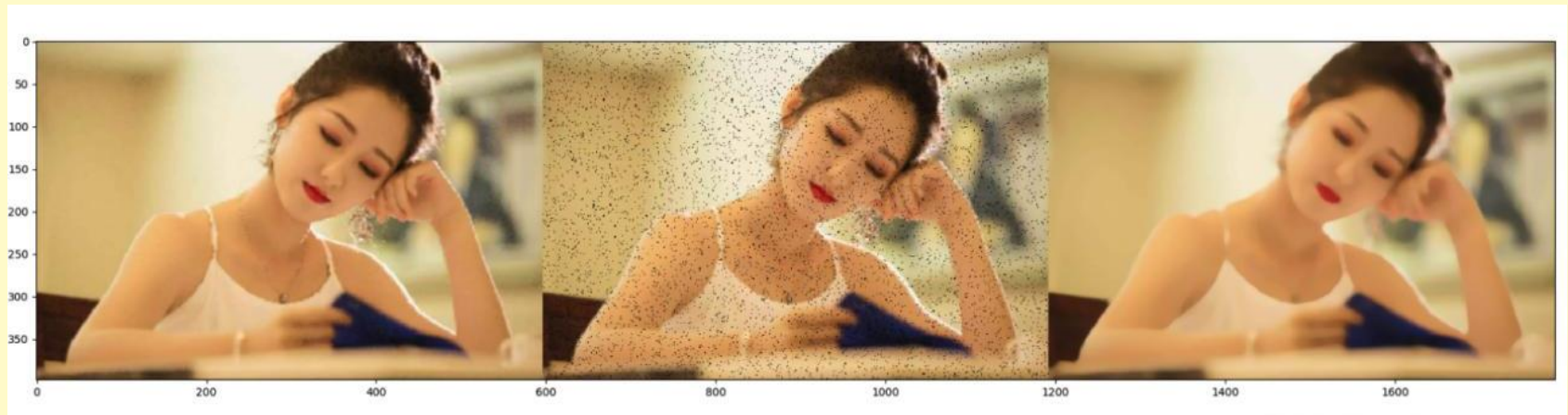
10	9	4	7	2
4	31	12	23	30
5	43	23	32	8
54	7	3	43	4
23	76	56	45	9

https://blog.csdn.net/Monster_H7

«Соль и перец»

- Соль и перец - это шум черно-белого света и темных пятен, создаваемый датчиком изображения, каналом передачи и обработкой декодирования.
- Наиболее часто используемым алгоритмом для устранения импульсных помех и шума соли и перца является медианная фильтрация.
- Большое количество экспериментальных исследований показало, что изображения, снятые камерой, более серьезно подвержены влиянию дискретных импульсов, солевого и перцового шума и гауссовского шума с нулевым средним.
- Существует два требования для обработки фильтра: одно состоит в том, что важная информация, такая как контур и края изображения, не может быть повреждена, а другое - в том, что изображение четкое и визуальный эффект хороший.

Отображение эффекта обработки соли и перца и медианной фильтрации



ЗАДАЧА РАСПОЗНАВАНИЯ ОБРАЗОВ

Предмет распознавания образов

Задача распознавания образов

Предмет распознавания образов

Примеры интеллектуальных компьютерных систем

- 1) Машинное зрение. Это системы, назначение которых состоит в получении изображения через камеру и составление его описания в символьном виде (какие объекты присутствуют, в каком взаимном отношении находятся и т.д.).
- 2) Символьное распознавание – это распознавание букв или цифр.
 - a. Optical Character Recognition (OCR);
 - b. Ввод и хранение документов;
 - c. Pen Computer;
 - d. Обработка чеков в банках;
 - e. Обработка почты.
- 3) Диагностика в медицине.
 - a. Маммография, рентгенография;
 - b. Постановка диагноза по истории болезни;
 - c. Электрокардиограмма.
- 4) Геология.
- 5) Распознавание речи.
- 6) Распознавание в дактилоскопии (отпечатки пальцев), распознавание лица, подписи, жестов.

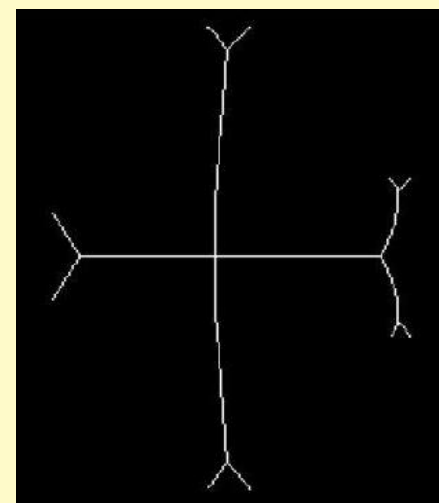
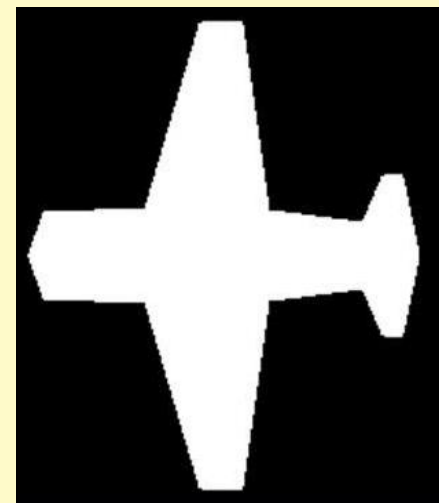
АЛГОРИТМЫ РАСПОЗНАВАНИЯ ОБРАЗОВ

1. Алгоритм скелетизации

Это метод распознавания одинарных бинарных образов, основанный на построение скелетов этих образов и выделения из скелетов ребер и узлов.

Далее по соотношению ребер, их числу и числу узлов строится таблица соответствия образам.

Так, например, скелетом круга будет один узел, скелетом буквы П - три ребра и два узла, причем ребра относятся как 2:2:1.



2. Нейросетевые структуры



ИНС (искусственные нейросети) – это математическая модель функционирования традиционных для живых организмов нейросетей, которые представляют собой сети нервных клеток.



Как и в биологическом аналоге, в искусственных сетях основным элементом выступают нейроны, соединенные между собой и образующие слои, число которых может быть разным в зависимости от сложности нейросети и ее назначения (решаемых задач).



Одна из самых популярных задач нейросетей – распознавание визуальных образов. Сегодня создаются сети, в которых машины способны успешно распознавать символы на бумаге и банковских картах, подписи на официальных документах, детектировать объекты и т.д.

Нейросеть – это математическая модель в виде программного и аппаратного воплощения, строящаяся на принципах функционирования биологических нейросетей. Сегодня такие сети активно используют в практических целях за счет возможности не только разработки, но и обучения. Их применяют для прогнозирования, распознавания образов, машинного перевода, распознавания аудио и т.д.

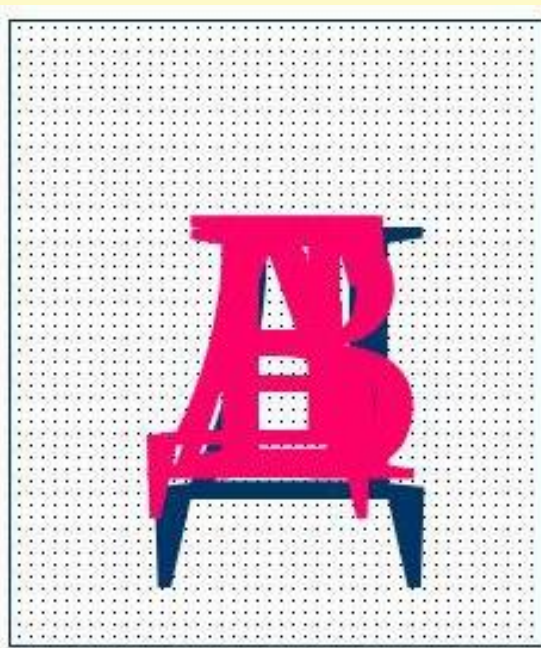
3. Инвариантные числа

Из геометрии образов можно выделить некоторые числа, **инвариантные** относительно размера и поворота образов, далее можно составить таблицу соответствия этих чисел конкретному образу(почти как в алгоритме скелетизации).

Примеры инвариантных чисел - число эллера, эксцентриситет, ориентация(в смысле расположения главной оси инерции относительно чего-нибудь тоже инвариантного)..

3. Поточечное процентное сравнение с эталоном

Некоторая предобработка, для получения инвариантности относительно размера и положения, затем осуществляется сравнение с заготовленной базой эталонов изображений - если совпадение больше чем какая-то отметка, то считаем образ распознанным.



А	а	А	а
Б	б	В	в
В	в	С	с
Г	г	Д	д
Д	д	Е	е
Е	е	Ё	ё
Ё	ё	Ж	ж
Ж	ж	З	з
З	з	И	и
...

Д – ошибка минимальная →
код 196 (Win-1251) =&11000100

Возможная последовательность шагов при распознавании образов



Бинаризация

- При работе с цветной камерой преобразование из цвета в ЧБ может идти по формуле:

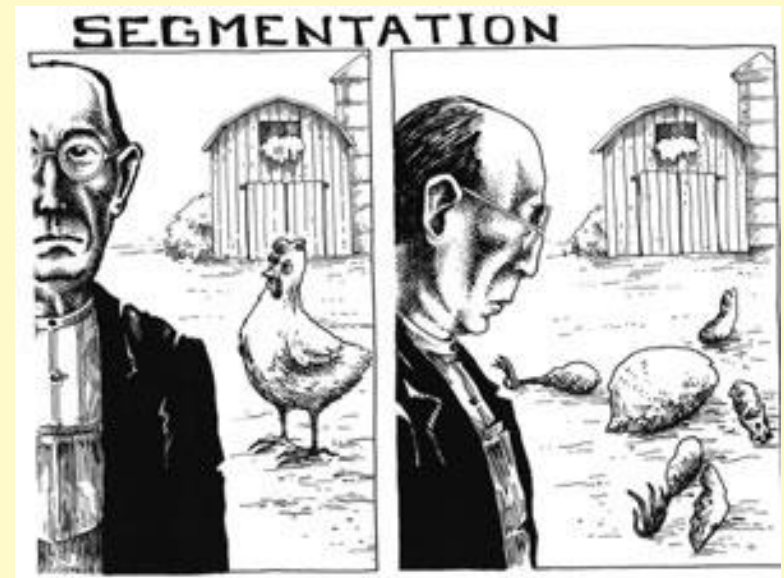
$$Y:=0.3*R+0.59*G+0.11*B$$

- Алгоритм: есть некоторая граница, если цвет оттенка серого выше - он считается белым, если ниже - считается черным.
- Для серьезного улучшения качества работы распознавания, и уменьшения времени работы последующих модулей, на этом месте лучше ввести некий фильтр для исключения последовательностей типа: 01010101.

Существует несколько алгоритмов бинаризации

Сегментация

- Все описанные выше алгоритмы распознавания образов работают с **единственным видимым образом**
- Если не предусматривать некоторое разбиение общего изображения на части, то ни один из описанных выше алгоритмов не сможет корректно работать.
- Существуют алгоритмы сегментации цветных и черно-белых изображений, так же в сегментации четко разделяются черно-белые изображения на **бинарные** и с **оттенками серого**.
- На стадии бинаризации должны преобразовать объект изображения в **бинарную матрицу данных**.



Сегментация осуществляется путем прохода по матрице изображения слева направо, сверху вниз. При проходе выполняются следующие правила:

$$\begin{array}{ccc} \begin{array}{cc} 0 & 0 \\ 0 & 1 \end{array} \rightarrow \begin{array}{cc} 0 & L \\ 0 & L \end{array} & \begin{array}{cc} 0 & 0 \\ L & 1 \end{array} \rightarrow \begin{array}{cc} L & L \\ M & 1 \end{array} & \begin{array}{cc} 0 & 0 \\ L & L \end{array} \rightarrow \begin{array}{cc} L & L \\ M & L \end{array} \quad (L = M) \end{array}$$

Далее, если уже меченные объекты сливаются, то должно происходить и слияние меток.

Для корректной работы алгоритма требуется переназначать метки во всей матрице изображения каждый раз, когда выполняется последнее правило.

Вариант 1

- ❑ Можно завести отдельно массив меток, в котором надо прописывать какая метка на какую ссылается, и все операции при проходе по матрице изображения проводить только с массивом ссылок меток. При этом в конце прохода должен иметься массив меток, ссылающихся друг на друга, и только несколько меток должны быть уникальными - которые собственно и **образуют объекты**.
- ❑ Что бы получить уникальные объекты, которые уже можно направить в модуль распознавания нужно нормализовать этот массив ссылок и переназначить все элементы матрицы на уникальные метки исходя из данных нормализованного массива ссылок.
- ❑ Но на сложных объектах может происходить разделение цельного объекта на части с потерей связи между частями объекта.

Вариант 2

- ❑ Рекуррентное нахождение уникальных предков каждой метки, и далее всю работу по переприсвоению производить уже с ними, таким образом связь в объекте не может быть потеряна, потому что рекурсия дает возможность выхватывать самые длинные «отростки» объекта.
- ❑ У объектов может быть слишком высокая детализация , в результате алгоритм работает не с идеальными картинками, а с довольно сложными, но это происходит только на стадии сегментации, затем эти лишние детали на больших объектах убираются

После выполнения алгоритма сегментации, исходная бинарная матрица элементов должна выглядеть примерно так:

[illegible]

В реальности элементы которые видит камера довольно сложны, и не идеальны, поэтому при проходе только лишь по началу первого элемента могут быть задействованы несколько номеров меток

Выделение образа, вычисление некоторых инвариантных чисел

После успешного завершения сегментации, каждый сегмент попадает в модуль распознавания

В модуле сегментации для каждого образа могут обозначаться его границы и вычисляться его финальная площадь.

- Границы нужны для ускорения работы модуля распознавания - т.к он ищет образ только в указанном месте.

Для того, что бы образы распознавались инвариантно относительно положения и поворота надо привязаться к их структуре(форме)

У каждого бинарного образа можно вычислить несколько признаков, не зависящих от его поворота или размера

Например, можно составить таблицу из нескольких инвариантных признаков и распознавать образы исходя только из этих данных.

- При вычислении ряда морфометрических признаков используются понятия механики твердого тела. В частности, это относится к длинам осей инерции объекта. Направления в теле, совпадающие с полуосями эллипсоида инерции, называют **главными осями инерции**.
- Пусть N - количество пикселей, относящихся к объекту.
- Все множество пикселей $p(x, y)$, относящихся к объекту, обозначим Q . Тогда координаты центра масс объекта вычисляются как

$$x_c = \frac{1}{N} \sum_{p(x,y) \in \Omega} x$$

$$y_c = \frac{1}{N} \sum_{p(x,y) \in \Omega} y$$

Вспомогательные величины:

$$U_x = \frac{1}{12} + \frac{1}{N} \sum_{p(x,y) \in \Omega} (x - x_c)^2$$

$$U_y = \frac{1}{12} + \frac{1}{N} \sum_{p(x,y) \in \Omega} (y - y_c)^2$$

$$C = \sqrt{(U_x - U_y)^2 + 4 \cdot U_x \cdot U_y}$$

- Тогда длины максимальной A_{\max} и минимальной A_{\min} осей инерции вычисляются как:

$$A_{\max} = 2\sqrt{2} \cdot \sqrt{U_x + U_y + C} \quad A_{\min} = 2\sqrt{2} \cdot \sqrt{U_x + U_y - C}$$

- Длины главных осей инерции используются для вычисления эксцентриситета и ориентации объекта.
- Эксцентриситет определяется с помощью соотношения

$$E = \frac{2 \cdot \sqrt{(0.5 \cdot A_{\max})^2 - (0.5 \cdot A_{\min})^2}}{A_{\max}}$$

- Ориентация определяется как угол в градусах между максимальной осью инерции и осью X. Если $U_y > U_x$, то ориентация O вычисляется с помощью формулы

$$O = \frac{180}{\pi} \cdot \arctg\left(\frac{U_y - U_x + C}{2 \cdot U_{xy}}\right)$$

в противном случае O вычисляется как

$$O = \frac{180}{\pi} \cdot \arctg\left(\frac{2 \cdot U_{xy}}{U_x - U_y + C}\right)$$

Поворот образа, повторное выделение образа

- Ориентация изображения уникальная для каждого образа.
- Теперь образ можно повернуть относительно центра масс, так что бы его ориентация была параллельна оси X.
- Это дает инвариантность относительно начального поворота.



Сжатие образа в матрицу заданного размера

- В бинарную матрицу записывается каждый отдельно сегментированный образ.
- Если нужна большая детализация, то лучше использовать матрицы большего размера
- Малое разрешение распознаваемой матрицы позволяет исправить некоторые возможные ошибки при повороте изображения (вычисления ориентации образа), т.к незначительное отклонение в 5 градусов, не будет заметно после сжатия образа.
 - *Когда какую-то часть образа закроет помеха или блик - центр масс смещается, ориентация, возможно, тоже - поворот будет не вполне корректен.*

ПРИЗНАКИ И КЛАССИФИКАТОРЫ

Измерения, используемые для классификации
образов, называются **признаками**

Признак - это некоторое количественное измерение объекта произвольной природы.

Совокупность признаков, относящихся к одному образу, называется **вектором признаков**.

Вектора признаков принимают значения в **пространстве признаков**.

Каждому образу ставится в соответствие единственное значение вектора признаков и наоборот: каждому значению вектора признаков соответствует единственный **образ**.

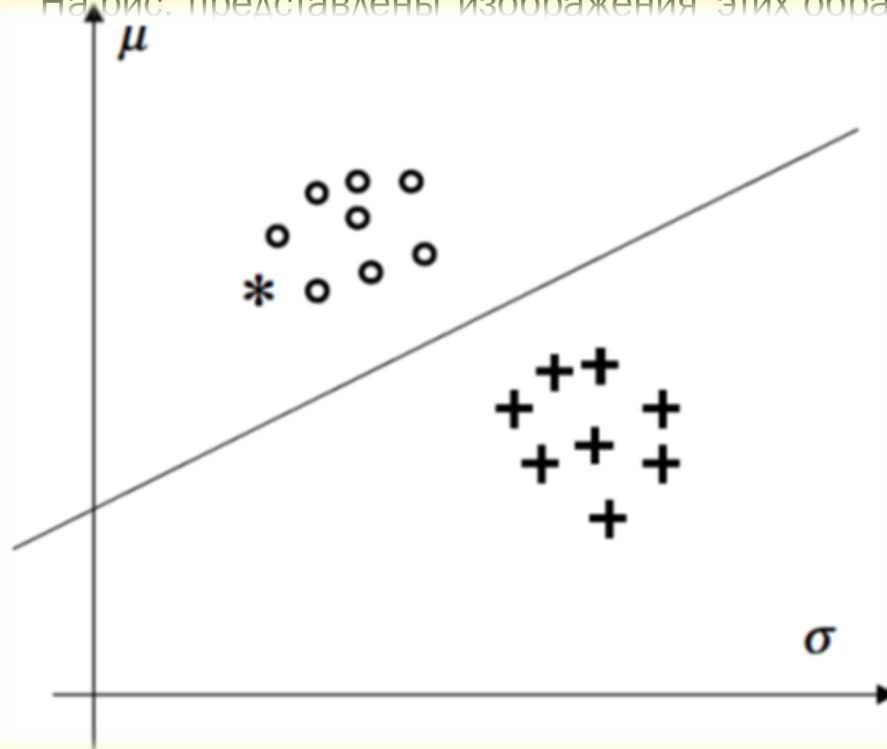
Классификатором или **решающим правилом** называется правило отнесения образа к одному из классов на основании его вектора признаков.

Пример иллюстрации понятий Признаков и Классификаторов

Задача диагностики печени по результатам инструментального исследования. Доброкачественные (левый рисунок – **класс А**) и злокачественные (правый рисунок – **класс В**) изменения дают разную картину.

Имеется несколько препаратов в базе данных, про которые известна их принадлежность к классам А и В (правильная классификация). Очевидно, что образцы отличаются интенсивностью точек изображения. В качестве **вектора признаков** выберем пару: среднее значение (m) и среднеквадратичное отклонение (s) интенсивности в изображении.

На рис. представлены изображения этих образов в пространстве признаков.



← Распределение векторов признаков прецедентах класса А (кружки) и класса В (крестики). Признаки - средние значения и средние отклонения яркости в образах. Прямая линия разделяет вектора из разных классов.

Практическая разработка системы классификации осуществляется по следующей схеме:



В процессе разработки необходимо решить следующие вопросы:

1. Как выбрать вектора признаков?

- **Задача генерации признаков** – это выбор тех признаков, которые с достаточной полнотой (в разумных пределах) описывают образ.

2. Какие признаки наиболее существенны для разделения объектов разных классов?


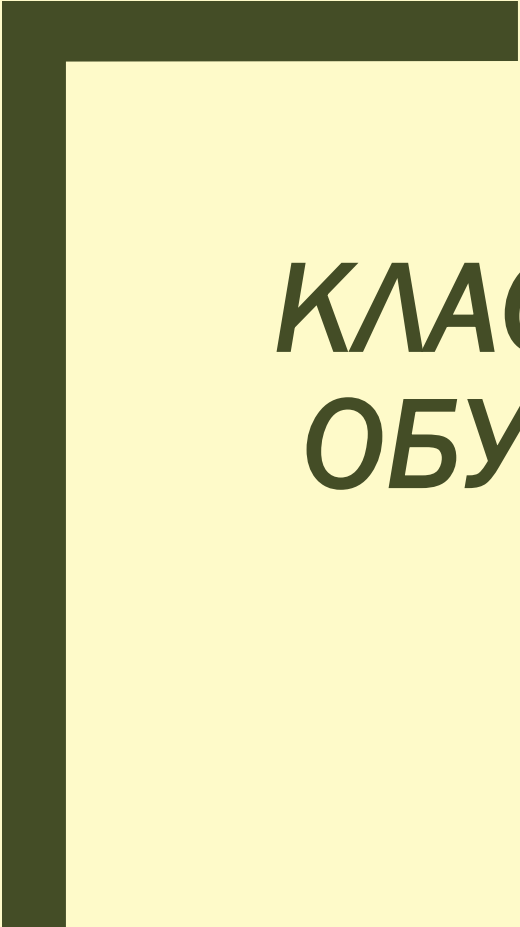
- **Задача селекции признаков** – отбор наиболее информативных признаков для классификации.

3. Как построить классификатор?

- **Задача построения классификатора** – выбор решающего правила, по которому на основании вектора признаков осуществляется отнесение объекта к тому или иному классу.

4. Как оценить качество построенной системы классификации?

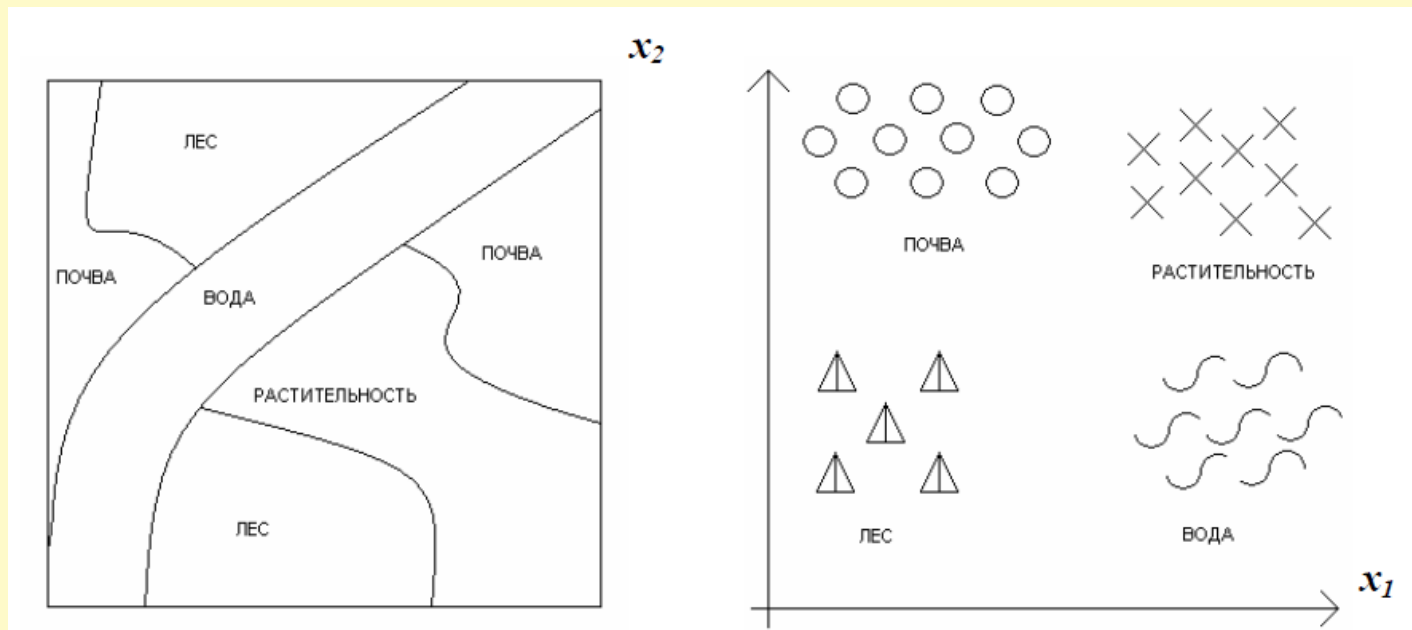
- **Задача количественной оценки системы** (выбранные признаки + классификатор) с точки зрения правильности или ошибочности классификации.



КЛАССИФИКАЦИЯ С ОБУЧЕНИЕМ И БЕЗ ОБУЧЕНИЯ

Пример

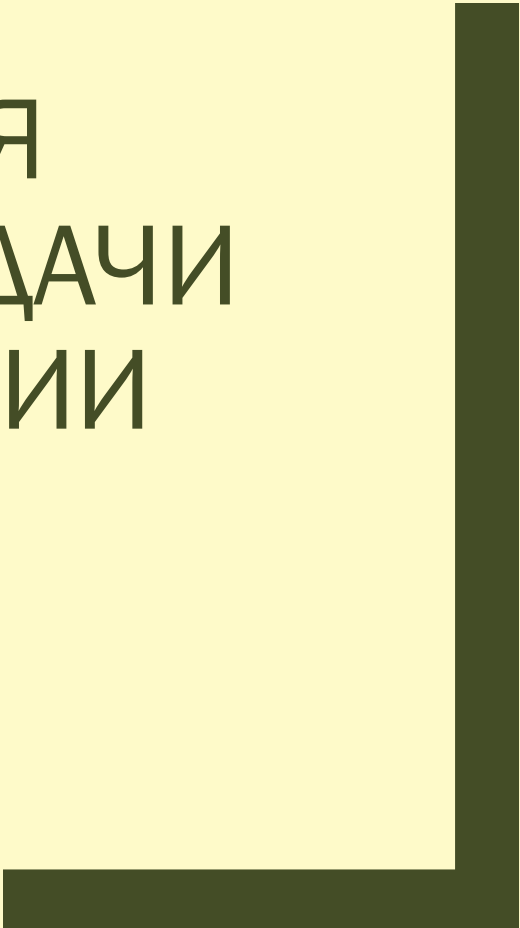

Съемка со спутника и классификацию поверхности по отраженной энергии (рис). На рисунке изображены снимок из космоса (слева) и результат кластеризации векторов признаков, рассчитанных для различных элементов изображения (справа). Распределение образов, изображенных точками (x_1, x_2) по классам осуществляется на основе анализа «скоплений» этих точек в пространстве признаков.



Изображение различных типов поверхности и кластеризация соответствующих векторов признаков

Пример

- Распознавание образов – в общественных (социальных) науках. **Целью** задачи является построение системы классификации государств для определения необходимости гуманитарной поддержки со стороны международных организаций.
- Необходимо выявить закономерности связей между различными, объективно измеряемыми параметрами, например, связь между **ВНП, уровнем грамотности и уровнем детской смертности**.
- Страны можно представить трехмерными векторами, а задача заключается в построении меры сходства этих векторов и дальнейшем построении схемы кластеризации (выбора групп) по этой мере.

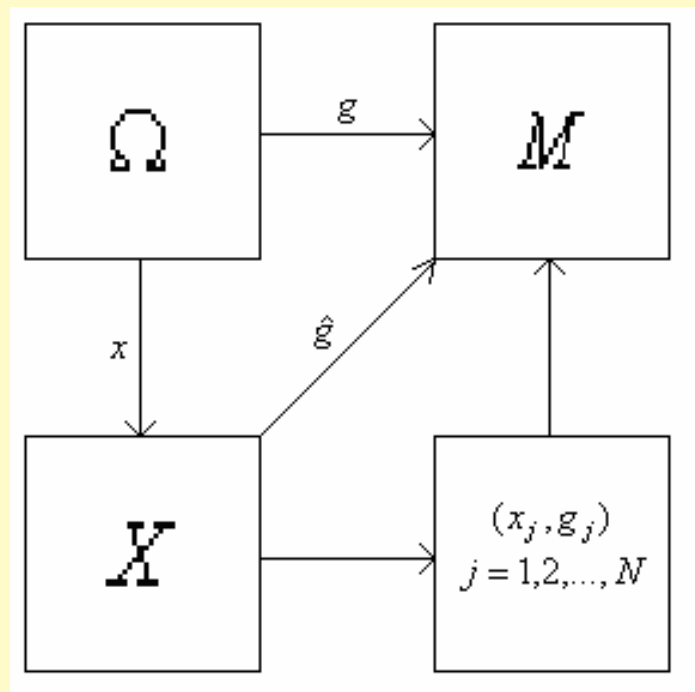


ФОРМАЛЬНАЯ ПОСТАНОВКА ЗАДАЧИ КЛАССИФИКАЦИИ

- Ω – множество объектов распознавания (пространство образов).
- $\omega : \varpi \in \Omega$ – объект распознавания (образ).
- $g(\omega) : \Omega \rightarrow M, M = \{1, 2, \dots, m\}$ – индикаторная функция, разбивающая пространство образов Ω на m непересекающихся классов $\Omega^1, \Omega^2, \dots, \Omega^m$. Индикаторная функция неизвестна наблюдателю.
- X – пространство наблюдений, воспринимаемых наблюдателем (пространство признаков).
- $x(\omega) : \Omega \rightarrow X$ – функция, ставящая в соответствие каждому объекту ω точку $x(\omega)$ в пространстве признаков. Вектор $x(\omega)$ – это образ объекта, воспринимаемый наблюдателем.
- В пространстве признаков определены непересекающиеся множества точек $K_i \subset X, i = 1, 2, \dots, m$, соответствующих образам одного класса.
- $\hat{g}(x) : X \rightarrow M$ – решающее правило – оценка для $g(\omega)$ на основании $x(\omega)$, т.е. $\hat{g}(x) = \hat{g}(x(\omega))$.

■ Пусть $x_j = x(\omega_j), j = 1, 2, \dots, N$ – доступная наблюдателю информация о функциях $g(\omega)$ и $x(\omega)$, но сами эти функции наблюдателю неизвестны. Тогда $(g_j, x_j), j = 1, 2, \dots, N$ – есть множество прецедентов.

■ Задача заключается в построении такого решающего правила $\hat{g}(x)$, чтобы распознавание проводилось с минимальным числом ошибок.



■ Обычный случай – считать пространство признаков евклидовым, т.е. $X = \mathbb{R}^l$.

■ Качество решающего правила измеряют частотой появления правильных решений. Обычно его оценивают, наделяя множество объектов Ω некоторой вероятностной мерой. Тогда задача записывается в виде

$$\min P\{ \hat{g}(x(\omega)) \neq g(\omega) \}$$

КЛАССИФИКАЦИЯ НА ОСНОВЕ БАЙЕСОВСКОЙ ТЕОРИИ РЕШЕНИЙ

Байесовский подход

- Задано M классов $\Omega_1, \Omega_2, \dots, \Omega_M$, а также $P(\Omega_i | x)$, $i = 1, 2, \dots, M$ – вероятность того, что неизвестный образ, представляемый вектором признаков x , принадлежит классу Ω_i .
- $P(\Omega_i | x)$ называется *апостериорной вероятностью*, поскольку задает распределение индекса класса после эксперимента (*a posteriori* – т.е. после того, как значение вектора признаков x было получено).
- Рассмотрим случай двух классов Ω_1 и Ω_2 . Естественным образом выбрать решающее правило таким образом: объект относим к тому классу, для которого апостериорная вероятность выше. Такое правило классификации по максимуму апостериорной вероятности называется **Байесовским**: если $P(\Omega_1 | x) > P(\Omega_2 | x)$, то x классифицируется в Ω_1 , иначе в Ω_2 . Таким образом, для Байесовского решающего правила необходимо получить апостериорные вероятности $P(\Omega_i | x)$, $i = 1, 2$. Это можно сделать с помощью формулы **Байеса**.

Пусть A_1, A_2, \dots, A_n – полная группа несовместных событий.

$$\bigcup_{i=1}^n A_i = \Omega. \quad A_i \cap A_j \neq \emptyset \text{ при } i \neq j.$$

Тогда апостериорная вероятность имеет вид:

$$P(A_i|B) = \frac{P(A_i)P(B|A_i)}{\sum_{i=1}^n P(A_i)P(B|A_i)}$$

где $P(A_i)$ – априорная вероятность события A_i , $P(B|A_i)$ – условная вероятность события B при условии, что произошло событие A_i .

Если $P(A)$ и $P(A|B)$ описываются плотностями $p(x)$ и $p(x|B)$, то

$$P(B|x) = \frac{p(x|B)P(B)}{P(x)} \Rightarrow P(\Omega_i|x) = \frac{p(x|\Omega_i)P(\Omega_i)}{P(x)}$$

При проверке классификации сравнение $P(\Omega_1|x)$ и $P(\Omega_2|x)$ эквивалентно сравнению $p(x|\Omega_1)P(\Omega_1)$ и $p(x|\Omega_2)P(\Omega_2)$