

Математические модели обработки сигналов

# Тема 12: Обработка изображений

Лектор: Кривошеин А.В.

## Обработка изображений

Под обработкой изображений можно понимать любое отображение из множества матриц  $\mathbb{R}^{N \times M}$  в то или иное множество. Можно провести некую классификацию методов обработки: обработка низкого, среднего и высокого уровня.

**Методы низкого уровня** — это, как правило, некие операторы в пространстве изображений, то есть результатом метода будет тоже изображение, изменённое некоторым образом. Например, целью обработки является уменьшение шума, изменение контраста, улучшение резкости и прочие методы улучшения зрительного восприятия изображений человеком.

**Методы среднего уровня** — это, как правило, выделение признаков из изображения. Эти методы включают сегментацию изображения, выделение объектов в изображении, а также классификация (или распознавание) отдельных объектов, кроме того извлечение особенностей изображения (границ, контуров, особых точек).

**Методы высокого уровня** включают в себя, например, “осмысление” распознанных объектов.

Обсудим ряд низкоуровневых методов обработки изображений.

## Обработка в пространственной области

Методы обработки изображений в пространственной области осуществляют манипуляции с пикселями изображений. Можно также выделить обработку изображений в частотной области, где происходит манипуляция с результатами преобразования Фурье.

Общий вид методов пространственной обработки изображений можно задать как оператор

$$J = T(I),$$

где  $I$  — это входное изображение,  $J$  — это обработанное изображение, а  $T$  — это оператор в пространстве изображений. Вообще говоря, размеры входного и выходного изображений могут отличаться. Оператор может обрабатывать и набор изображений:

$$J = T(I_1, \dots, I_K).$$

Поскольку цифровое изображение является матрицей (или совокупностью матриц в случае цветных изображений), то базовые операции с матрицами (сложение/вычитание, умножение, поэлементное умножение/деление, поэлементное возведение в степень) переносятся на изображения.

## Базовые операции над пикселями

По умолчанию будем работать с изображениями в серых тонах. Для цветных изображений аналогичные операции определяются для каждого цветового канала отдельно.

Пусть  $V$  — это некое множество уровней яркости,  $V^{M \times N}$  — это пространство изображений размера  $M \times N$ .

Пусть  $I_A, I_B$  некоторые изображения размера  $M \times N$ . Тогда сложение\вычитание двух изображений определяется равенством

$$I_{\text{out}}(i, j) := I_A(i, j) \pm I_B(i, j), \quad i = 1, \dots, M, \quad j = 1, \dots, N,$$

$I_{\text{out}}$  — это матрица на выходе. Аналогично можно определить поэлементное умножение.

Также можно определить прибавление константы к изображению (что значит увеличение\уменьшение яркости):

$$I_{\text{out}}(i, j) := I_A(i, j) \pm C, \quad C \in \mathbb{R}.$$

Важно понимать, что изображение — это матрица с определённым набором значений пикселей. Ясно, что элементы матрицы  $I_{\text{out}}$  могут принимать значения, выходящие за рамки диапазона допустимых значений  $V$ . Поэтому после указанных выше операций, требуется как-то **обработать возможное переполнение диапазона значений**. Есть два базовых варианта:

1. присвоить пикселям, имеющим значения выходящие сверху или снизу за рамки диапазона, максимальное или минимальное значение диапазона соответственно;
2. масштабировать значения в нужный диапазон.

## Базовые операции над пикселями

Операция НЕ (инверсия). Для бинарных изображений белый пиксель становятся чёрным и наоборот. Для изображений в серых тонах (8 бит на пиксель) или цветных (8 бит на один цвет), формула операции имеет вид

$$I_{\text{out}}(m, n) = \text{MAX} - I(m, n), \text{ где } \text{MAX} = 255.$$

Для изображений в иных форматах, MAX означает максимально возможное значение пикселя для этого формата. Эта операция используется, например, для создания негатива изображения.

Операция OR/XOR (ИЛИ/исключающее ИЛИ). Применяется для бинарных изображений. Используется, например, для детектирования движущихся объектов.

$$I_{\text{out}}(m, n) = I_A(m, n) \text{ OR } I_B(m, n)$$

$$I_{\text{out}}(m, n) = I_A(m, n) \text{ XOR } I_B(m, n)$$

Операция AND (И). Применяется для бинарных изображений. Используется для детектирование различий бинарных изображений, подсвечивания целевых областей с помощью бинарной маски.

$$I_{\text{out}}(m, n) = I_A(m, n) \text{ AND } I_B(m, n)$$

## Обработка в пространственной области

Рассмотрим общий оператор обработки изображения в пространственной области  $J = T(I)$ . Многие операторы действуют локально, то есть пиксель с координатами  $(m, n)$  выходного изображения зависит от значений пикселей в некоторой окрестности пикселя  $(m, n)$  входного изображения.

Рассмотрим частный случай таких локальных операторов, когда для формирования значения каждого пикселя обработанного изображения  $J$  используется только один пиксель входного изображения  $I$  (или входных изображений), то есть значение  $J(m, n)$  зависит только от  $I(m, n)$ .

Иными словами, яркость пикселя  $r = J(m, n)$  зависит только от яркости пикселя  $s = I(m, n)$ , то есть оператор пространственной обработки по сути имеет вид

$$r = T(s).$$

Применение оператора  $T$  ещё называют **функцией преобразования яркости**.

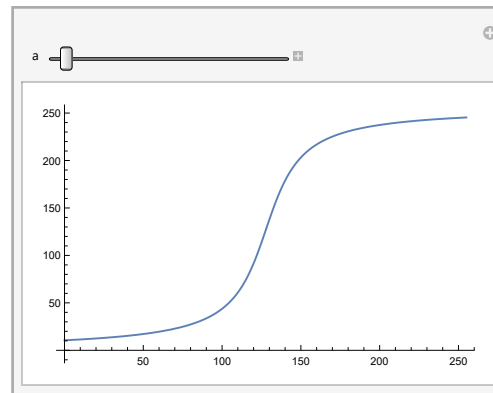
## Функция преобразования яркости

**Функция преобразования яркости** отображает диапазон допустимых значений яркости изображения  $V$  в себя  $T: V \rightarrow V$ .

Эта функция полезна для изменения динамического диапазона изображения.

При узком динамическом диапазоне изображение будет иметь низкий контраст. Для улучшения контраста надо расширить динамический диапазон.

Типичный график функции  $T$  для расширения динамического диапазона имеет вид:



Рассмотрим подробнее основные функции преобразования яркости. Среди них можно выделить три типа:

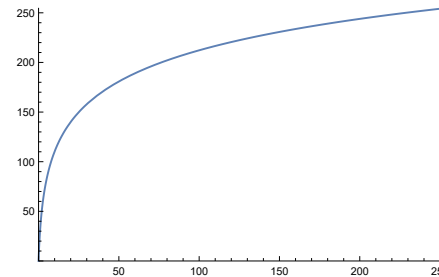
1. линейное;
2. логарифмическое;
3. степенное.

## Функция преобразования яркости

Логарифмическая функция преобразования яркости действует по правилу:

$$I_{\text{out}}(i, j) = C \ln(1 + I_{\text{in}}(i, j)),$$

где  $C$  некоторая константа, которая используется для нормировки, чтобы попасть в тот или иной диапазон. График соответствующей функции  $T$  имеет вид:



Для темных пикселей выделяется больший диапазон значений, а для светлых — меньший. То есть темное изображение осветляется. Типичный случай для исправления динамического диапазона — это фотография, сделанная перед источником света. Логарифмическое преобразование также полезно для отображения амплитуда спектра ДПФ, которые может иметь очень широкий диапазон.

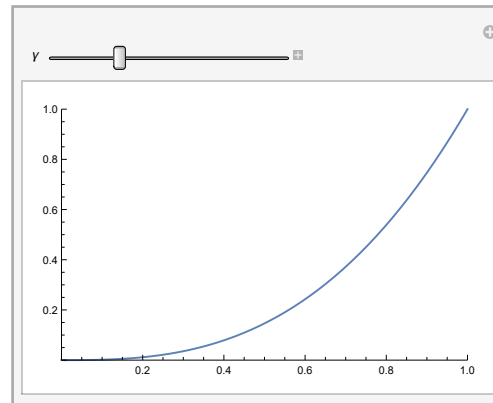


## Функция преобразования яркости

Степенная функция преобразования яркости действует по правилу:

$$I_{\text{out}}(i, j) = C (I_{\text{in}}(i, j))^{\gamma}$$

где  $C$  некоторая константа для нормировки, а параметр  $\gamma$  влияет вид графика.



Параметр  $\gamma > 1$  улучшает контраст в областях изображения с высокими значениями за счёт ухудшения его для регионов с пикселями с низкими значениями. При  $\gamma < 1$  наоборот.

Такое преобразование известно как гамма-коррекция.

## Гистограмма

**Гистограмма изображения** — это дискретная функция  $h$  сопоставляющая значению яркости число пикселей, имеющих эту яркость.

Например, для изображения  $I \in V^{M \times N}$ , функция  $h_I : V \rightarrow \mathbb{N} \cup \{0\}$  действует по правилу: для уровня яркости  $r \in V$

$$h_I(r) = \# \{(m, n) : I(m, n) = r\}.$$

Пример изображения и его гистограммы:

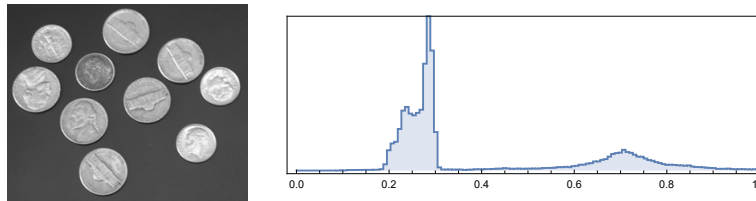


График гистограммы позволяет в ряде случаев сразу обнаружить разницу в распределении уровней яркости для фона изображения и переднего плана.

Для светлого изображения — график гистограммы смещён в правую часть, для тёмного — в левую. Низкоконтрастное изображение имеет узкую гистограмму, а у высококонтрастного гистограмма близка к равномерному распределению. Одна из распространённых операций — это выравнивание гистограммы.

## Выравнивание гистограммы

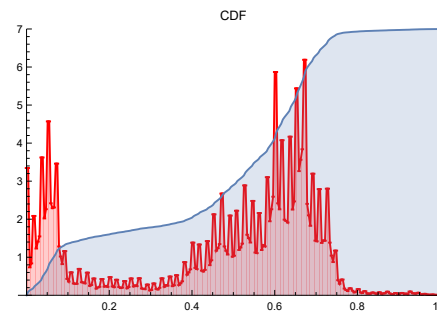
Рассмотрим серое изображение  $I$  размера  $N \times M$ . Пусть  $h_I$  является гистограммой этого изображения. Вычислим так называемую функцию распределения яркости (англ. cumulative distribution function)

$$\text{cdf}_I(k) = \sum_{i=0}^k h_I(i), \quad k \in V$$

То есть  $\text{cdf}_I(k)$  — это количество пикселей в изображении, имеющих значения яркости от 0 до  $k$ .

$$\text{cdf}_I(255) = NM.$$

Пример графика этой функции приведён ниже.

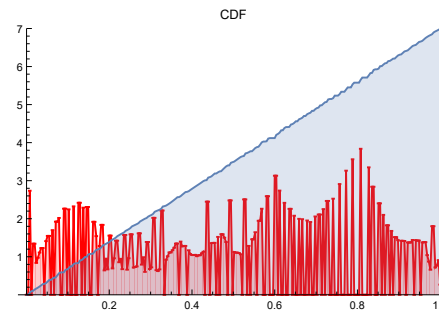


## Выравнивание гистограммы

Суть выравнивания гистограммы в том, чтобы функция распределения  $\text{cdf}_I$  была как можно более похожей на линейную функцию. Новые значения яркости по старым можно получить из следующей формулы:

$$k_{\text{new}} = T(k) = \text{round}\left(255 \frac{\text{cdf}_I(k) - 1}{NM - 1}\right)$$

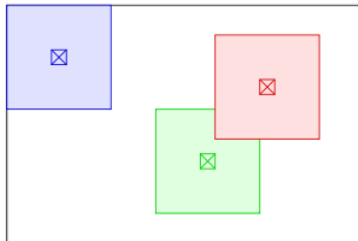
Новая гистограмма и функция распределения после выравнивания будут иметь вид



## Адаптивное выравнивание гистограммы

Описанное выше выравнивание гистограммы использует для всех пикселей единое преобразование, которое получено из гистограммы всего изображения. Этот метод хорошо работает, когда распределение значений пикселей одинаково по всему изображению. Однако, если изображение содержит области, которые значительно светлее или темнее большей части изображения, контраст в этих областях будет недостаточно усилен.

**Адаптивное выравнивание гистограммы** (англ. AHE, adaptive histogram equalization) преобразует каждый пиксель с помощью функции преобразования, которая формируется по некоторой небольшой области около текущего пикселя. В простейшем случае такой областью является квадратная окрестность текущего пикселя. Вывод функции преобразования из гистограммы точно такой же, как и для обычного выравнивания гистограммы: функция преобразования пропорциональна кумулятивной функции распределения (CDF) значений пикселей в окрестности.

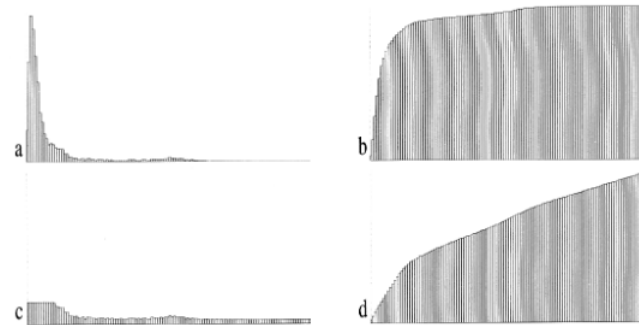


Для обработки пикселей вблизи границ можно считать, что изображение зеркально продолжено за границы изображения.

## Адаптивное выравнивание гистограммы

Проблема такого подхода, что в областях с почти постоянным значением яркости гистограмма будет иметь вид высокого пика и после выравнивания в таких областях усилится шум (пиксели со значением чуть ниже среднего будут принимать сильно меньшее значение, пиксели со значением чуть больше среднего будут принимать сильно большее значение).

Контрастно ограниченное адаптивное выравнивание гистограммы (англ. Contrast Limited ANE) решает эту проблему. Суть в том, что после построения гистограммы проверяется, превосходят ли значения гистограммы некоторый порог. Если да, то эти значения перераспределяются по всей гистограмме.



Это делает функцию распределения более пологой.

## Бинаризация

Бинаризация изображения по сути является предельным усилением контраста, до тех пор пока не останется два уровня — белые и чёрные пиксели.

Простейший метод бинаризации изображения заключается в следующем: выбирается (глобальный) порог  $T$  применяется формула:

$$I_B(i, j) = \begin{cases} 1, & I(i, j) \geq T \\ 0, & I(i, j) < T \end{cases}.$$

Глобальный порог можно выбирать из различных соображений. Один из способов связан с **методом Оцу**.

Пусть для изображения  $I$  вычислена гистограмма,  $p_I(k)$  это оценка вероятности появления уровня яркости  $k$ , то есть  $p_I(k) = \frac{h_I(k)}{NM}$ .

Пусть  $s$  — это порог, пиксели с меньшим уровнем яркости отнесём к классу  $C_1$  или к фону, а с большим к классу  $C_2$  или к объекту.

Порог  $s$  равен такому числу, что **минимальна внутриклассовая дисперсия** множеств  $C_1$  и  $C_2$ , то есть минимальна величина:

$$\sigma_{\text{in}}^2(s) = P_1(s) \sigma_1^2(s) + P_2(s) \sigma_2^2(s),$$

где  $P_1(s), P_2(s)$  суммарные вероятности яркостей классов, разделённых порогом  $s$ .

## Бинаризация

Метод глобального порога плохо работает, если в гистограмме нет двух выделенных пиков или если объект на переднем плане очень мал по сравнению с фоном.

Более успешным будет **поиск адаптивного порога** на основе ближайших соседних пикселей. Метод обосновывается тем, что хотя освещённость может изменяться в разных точках изображения, но она будет приблизительно равномерной в достаточно малой окрестности.

Около каждого пикселя рассмотрим окрестность некоторого размера, подсчитаем среднее. Далее составляется правило вида:

$$I_B(i, j) = \begin{cases} 1 & I(i, j) > \text{mean}(i, j) + C \\ 0 & \text{else} \end{cases},$$

где  $\text{mean}(i, j)$  это среднее в окрестности пикселя  $(i, j)$ ,  $C$  константа.

Под окрестностью обычно понимается квадратная окрестность размера  $N \times N$ .

Выбор величины  $N$  важен, и  $N$  должно быть достаточно большим, чтобы охватить достаточно много пикселей фона и переднего плана около каждой точки, но не слишком большим, чтобы глобальные изменения освещённости не влияли на порог.

Также, когда в окрестность попало недостаточно много пикселей переднего плана, то порог будет плох (среднее, медиана или разность максимума и минимума будут сильно зависеть от шума в изображении). Для этого требуется добавить отступ в виде константы  $C$ , что позволит установить порог выше, чем вариации шума в выбранной окрестности.