

# ОБРАБОТКА И РАСПОЗНАВАНИЕ ИЗОБРАЖЕНИЙ

Леонид Моисеевич Местецкий  
профессор

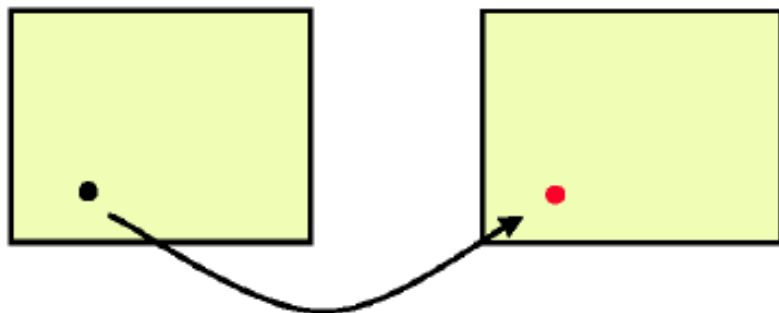
кафедра математических методов  
прогнозирования ВМК МГУ

кафедра интеллектуальных систем МФТИ

# ОПЕРАЦИИ НАД ИЗОБРАЖЕНИЯМИ

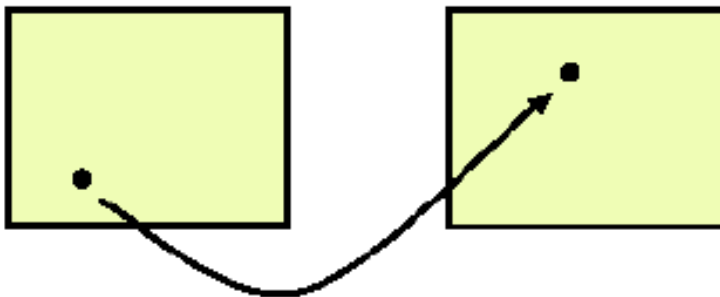
- Точечные
- Пространственные
- Геометрические
- Алгебраические
- Межкадровые

# Точечные операции



1. Результат зависит только от яркости пикселя и не зависит от его положения
2. Результат не зависит от окружающих пикселей
3. Минимальный расход памяти
4. Пример:  $I(x,y) = a I(x,y) + b$

# Геометрические операции



1. Результат зависит только от координат пикселя
2. Результат не зависит от окружающих пикселей
3. Пример:  $I(x,y) = I(x+a, y+b)$

# Алгебраические операции

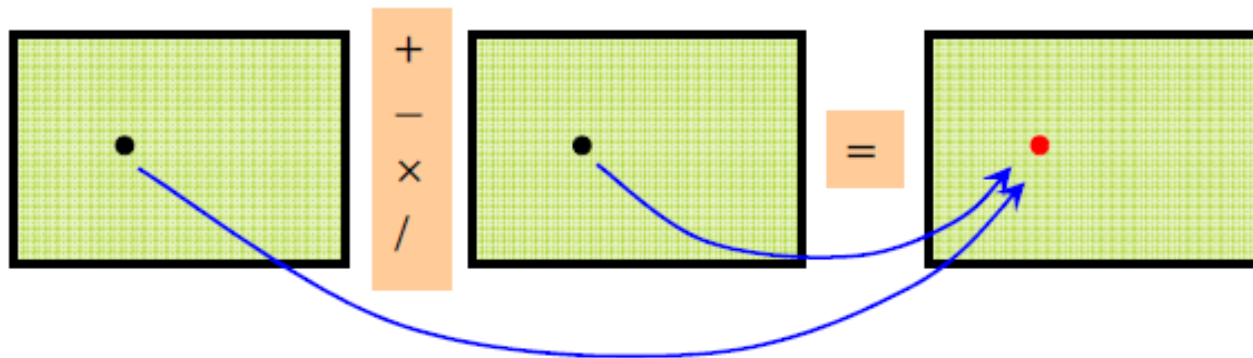
Составляют новое изображение из поточечных сумм, разностей, произведений и частных двух исходных изображений.

Сумма:  $C(x, y) = A(x, y) + B(x, y)$

Разность:  $C(x, y) = A(x, y) - B(x, y)$

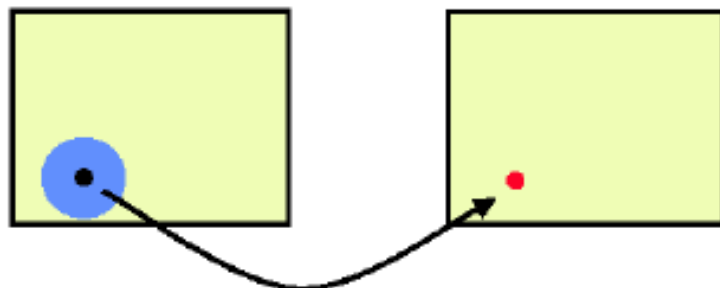
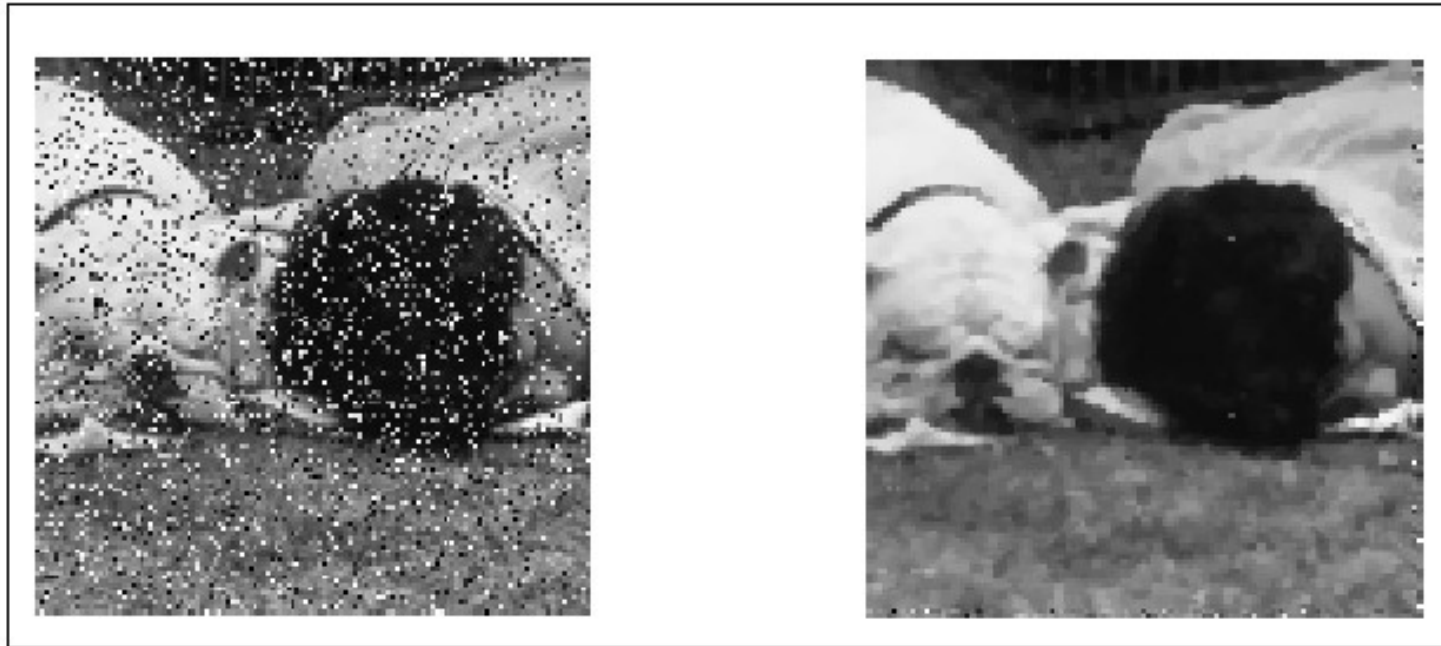
Произведение:  $C(x, y) = A(x, y) \times B(x, y)$

Частное:  $C(x, y) = A(x, y) \div B(x, y)$



# Пространственные

операции



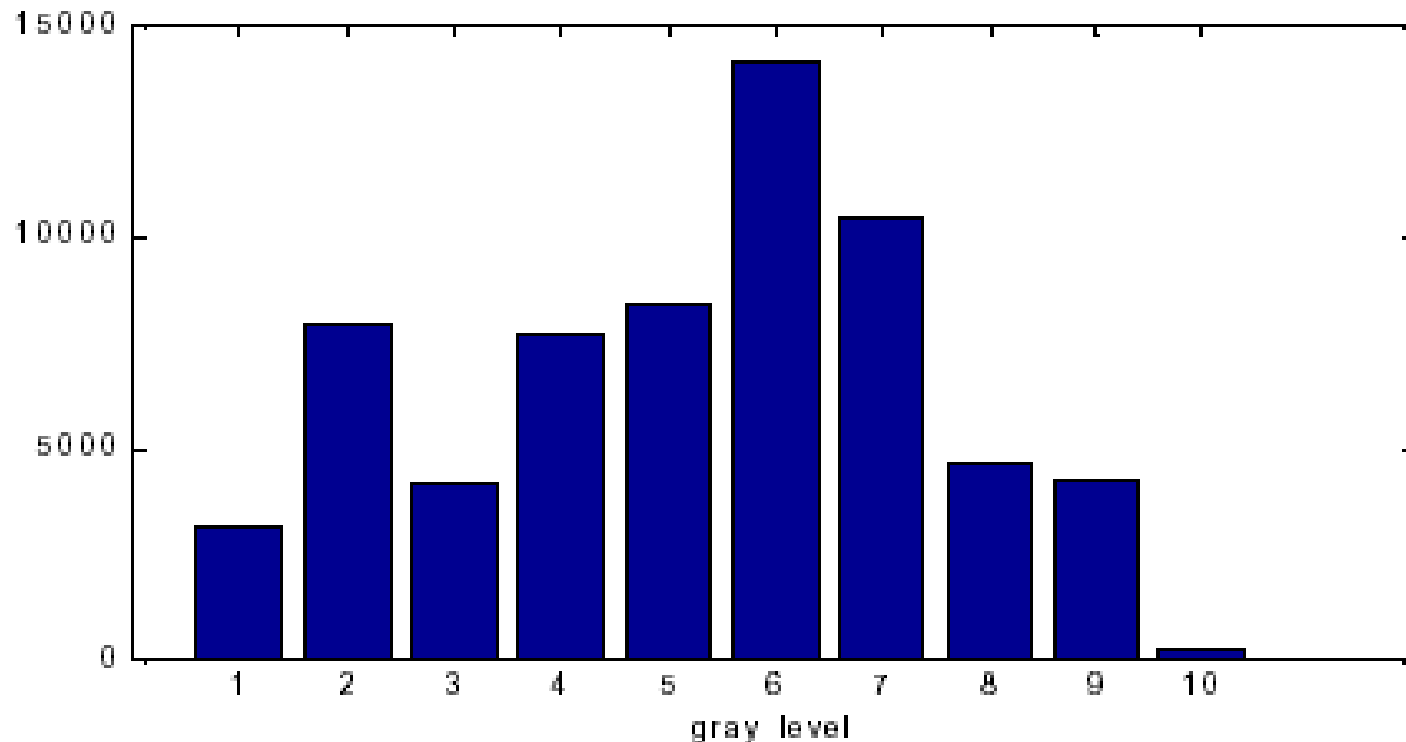
$$I'(x,y) = \sum_{(u,v) \in N(x,y)} I(u,v) / |N(x,y)|$$

$N(x,y)$  – окрестность  $(x,y)$

- Результат зависит от координат пикселя
- Результат зависит от окружающих пикселей

# Гистограмма изображения

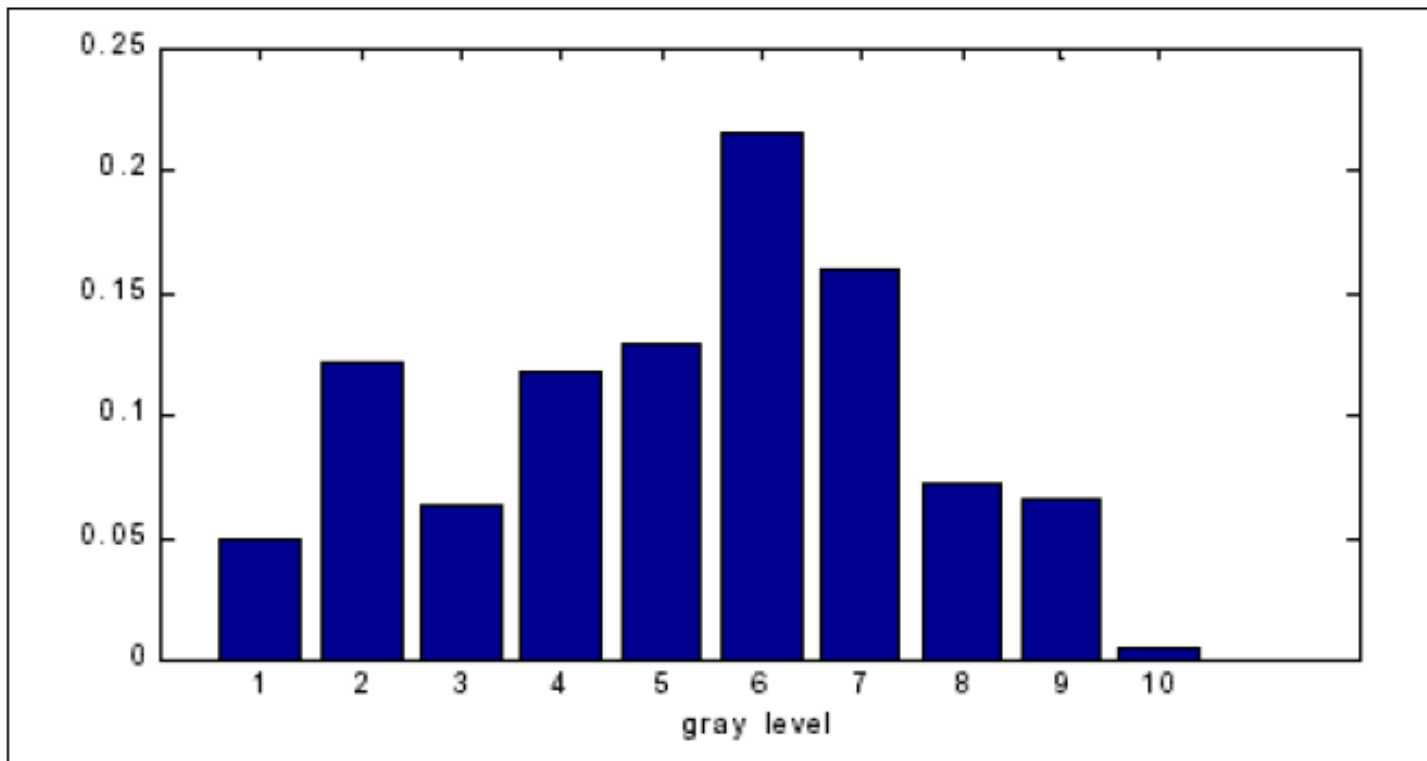
– количество пикселей с яркостью  $k$



# Нормализованная гистограмма

=

- количество пикселей в изображении
- вероятность получения яркости  $k$  при случайном выборе пикселя

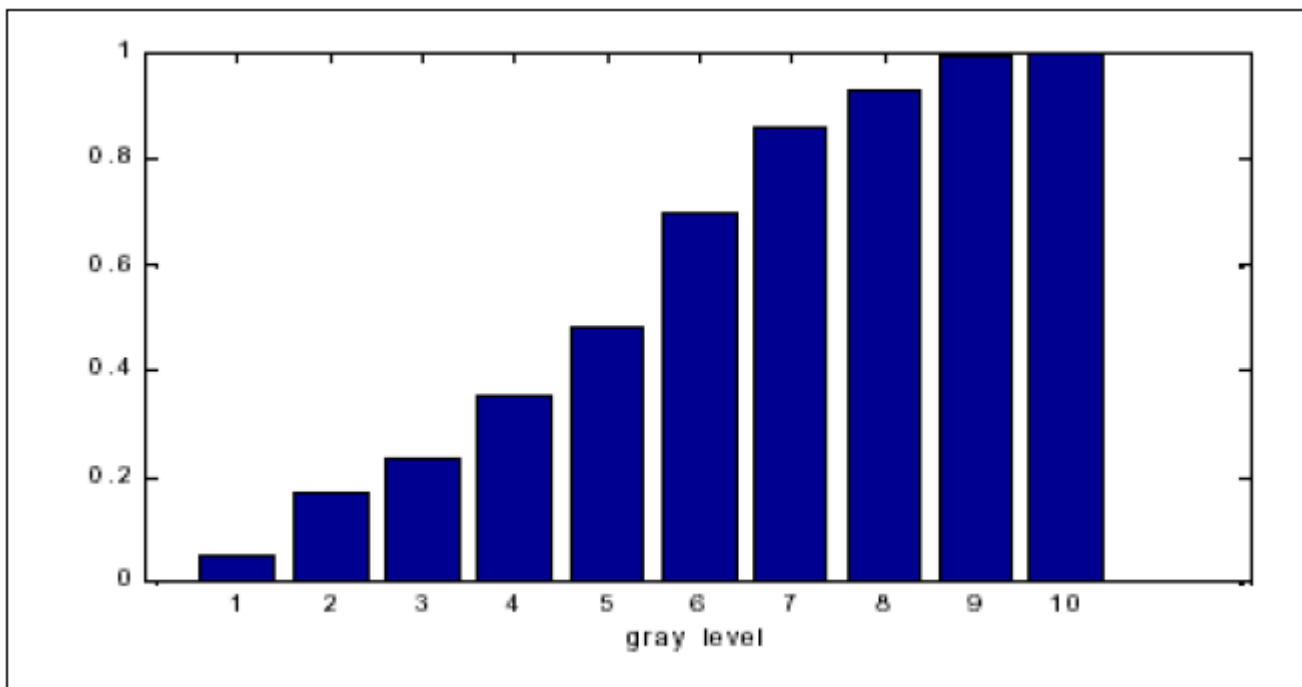




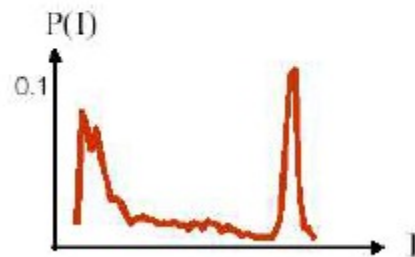
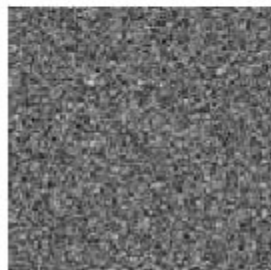
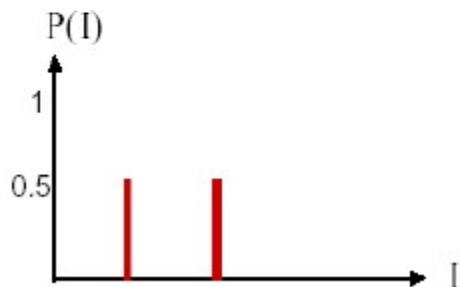
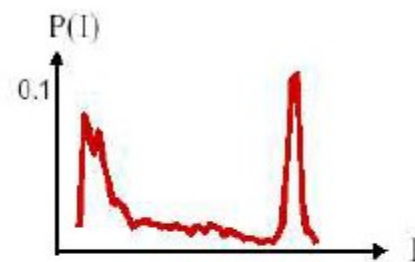
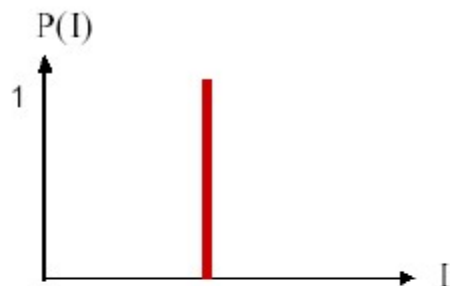
# Накопительная гистограмма

=

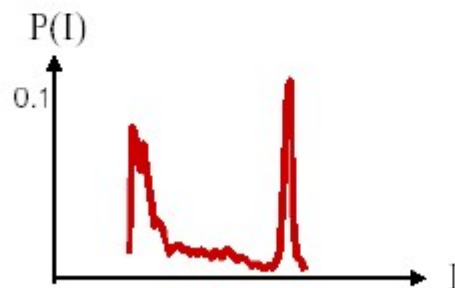
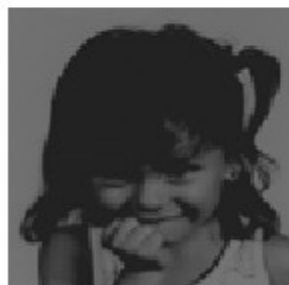
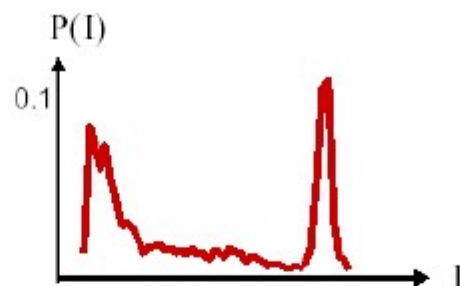
– вероятность получения яркости  
не больше при случайном выборе пикселя  
Заметим, что =



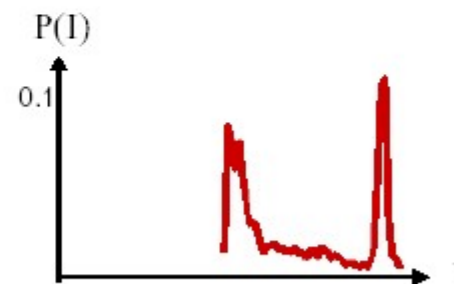
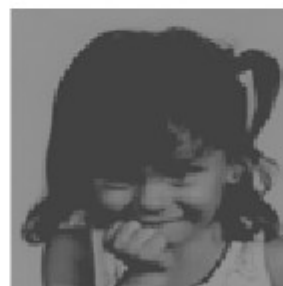
# Примеры гистограмм



# Примеры гистограмм



Уменьшение контраста



Просветление изображения

# Свойства гистограммы

## Средняя яркость

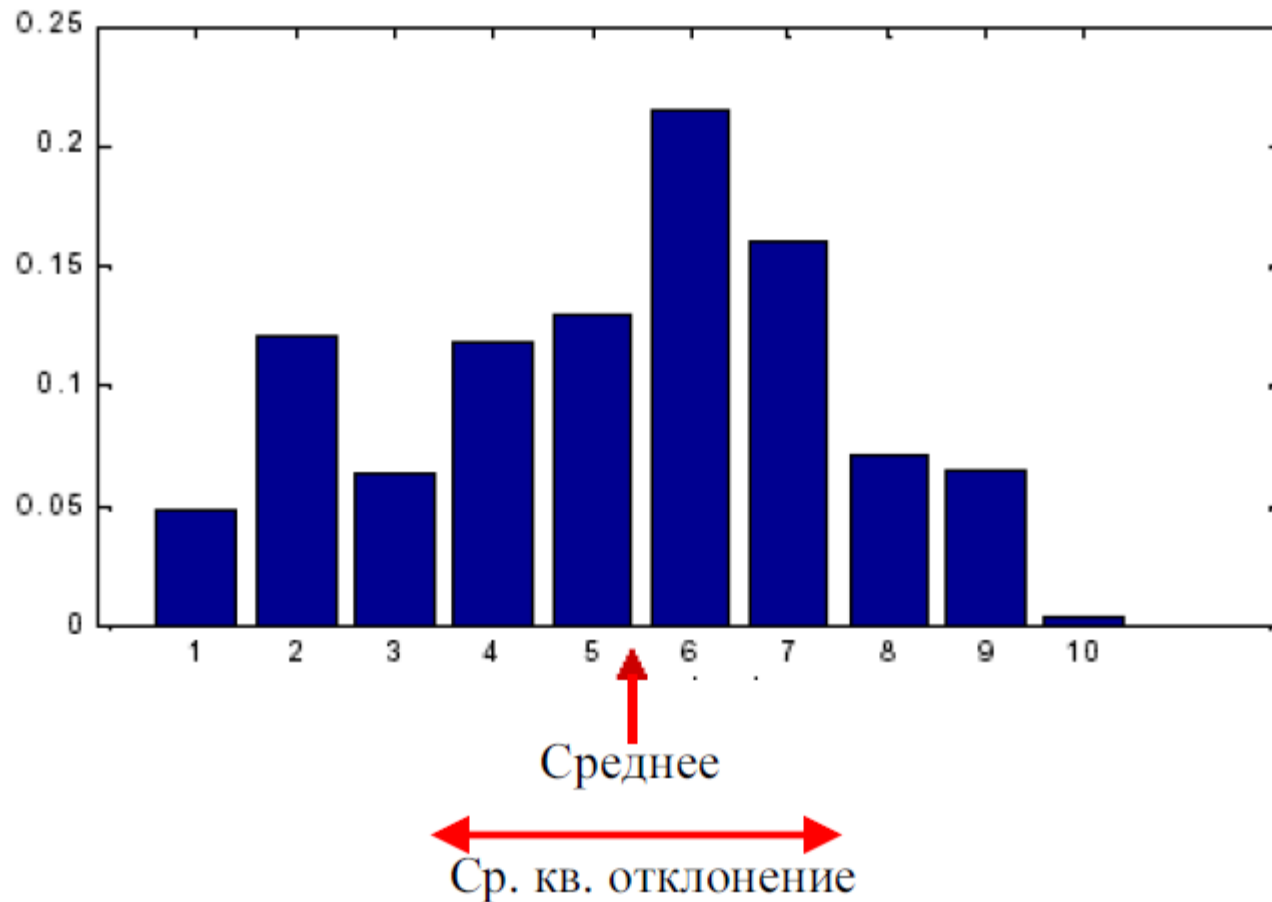
$$E(I) = \frac{\sum_{i,j} I(i,j)}{N} = \frac{\sum_k k H_I(k)}{N} = \sum_k k P_I(k)$$

$$E(I^2) = \frac{\sum_{i,j} I^2(i,j)}{N} = \sum_k k^2 P_I(k)$$

## Среднеквадратичное отклонение

$$S(I) = \frac{\sqrt{\sum_{i,j} (I(i,j) - E(I))^2}}{N} = \sqrt{E(I^2) - E^2(I)}$$

# Свойства гистограммы



# Применение гистограмм

- Оценка параметров изображения:  
среднее, вариация, энтропия,  
контрастность,  
площадь (для заданного уровня  
яркости)
- Выбор порога бинаризации
- Мера различия изображений
- Улучшение изображений:
  - Эквиализация гистограмм
  - Гистограммное растягивание
  - Гистограммное выравнивание

# Точечные операции

Каждый пиксель выходного изображения зависит от одного соответствующего пикселя входного изображения.

Точечная операция – это преобразование яркости.

$A(x,y)$  – исходное изображение

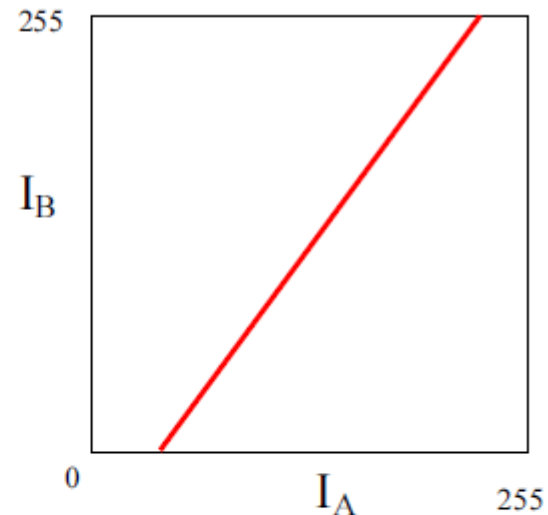
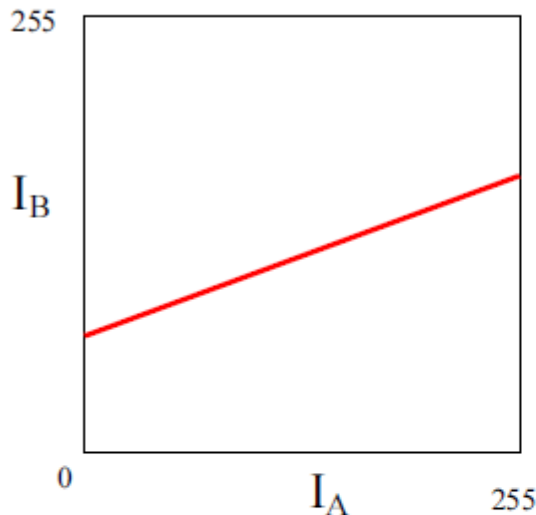
$B(x,y)$  – преобразованное изображение

$$B(x,y) = f[A(x,y)]$$

Точечная операция полностью задаётся функцией преобразования яркости  $f[k]$

# Линейные точечные

$$I_B = f(I_A) = \alpha \cdot I_A + \beta$$

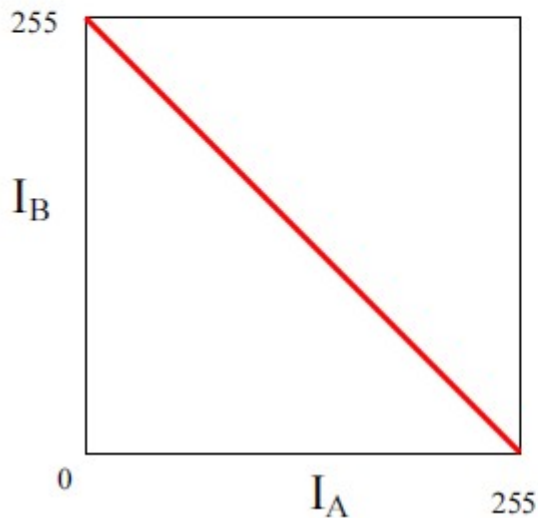


- при  $\alpha > 1$  контрастность усиливается,
- при  $\alpha < 1$  контрастность уменьшается,
- при  $\alpha = 1$  и  $\beta > 0$  изображение светлее,
- при  $\alpha = 1$  и  $\beta < 0$  изображение темнее,
- при  $\alpha < 0$  получаем дополнение (негатив в частности)



# Негативное изображение

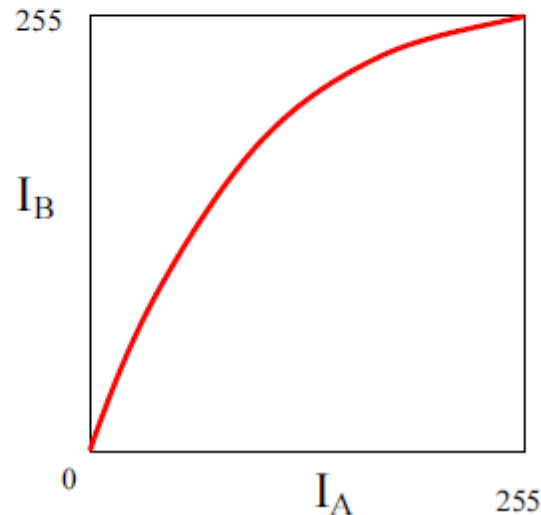
$$I_B = f(I_A) = I_{\max} - I_A$$



# Нелинейные точечные операции

$$I_B = f(I_A) = I_A + \alpha \cdot I_A \cdot (I_{\max} - I_A)$$

Оставляет тёмные и светлые точки почти без изменения, а средние увеличивает

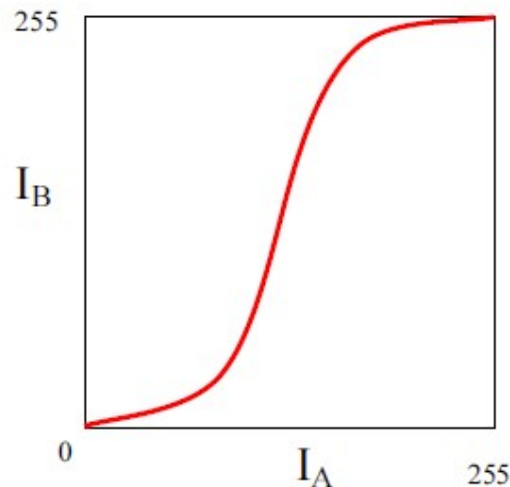


# Нелинейные точечные операции

$$I_B = f(I_A) = \frac{I_{\max}}{2} \cdot \left[ 1 + \frac{1}{\sin\left(\alpha \cdot \frac{\pi}{2}\right)} \cdot \sin\left(\alpha \cdot \pi \left(\frac{I_A}{I_{\max}} - \frac{1}{2}\right)\right) \right]$$

$$0 < \alpha < 1$$

Увеличивает контраст средних значений

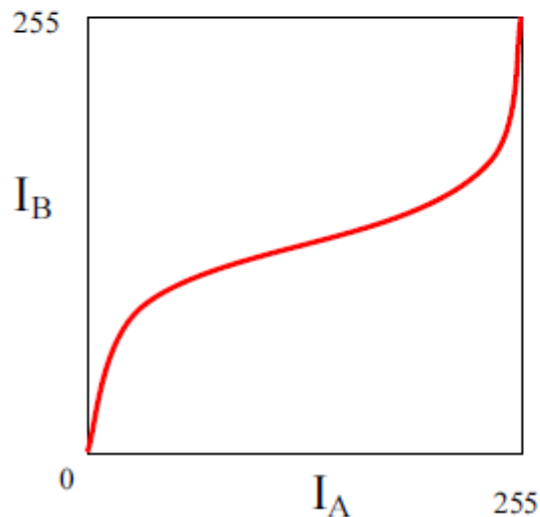


# Нелинейные точечные операции

$$I_B = f(I_A) = \frac{I_{\max}}{2} \cdot \left[ 1 + \frac{1}{\operatorname{tg}\left(\alpha \cdot \frac{\pi}{2}\right)} \cdot \operatorname{tg}\left(\alpha \cdot \pi \left(\frac{I_A}{I_{\max}} - \frac{1}{2}\right)\right) \right]$$

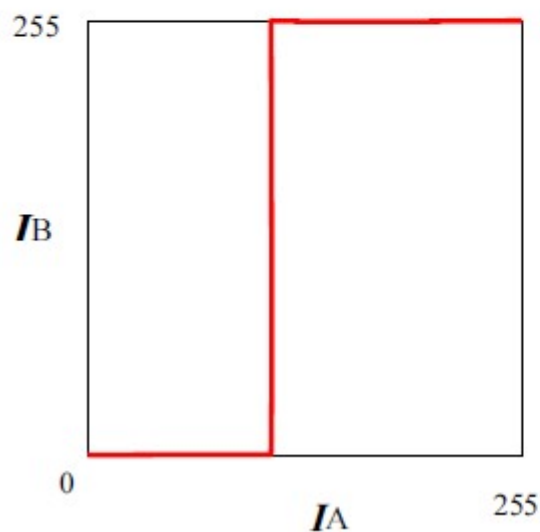
$$0 < \alpha < 1$$

Уменьшает контраст средних значений и увеличивает для малых и больших яркостей



# Бинаризация

Преобразование серого изображения в бинарное (двухцветное)



# Проблема выбора порога бинаризации



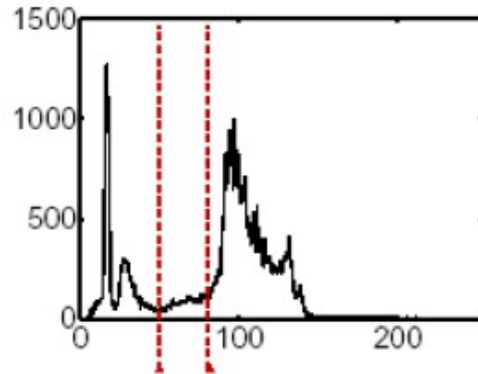
Исходное и бинарное  
изображения



Завышение и занижение  
порога бинаризации

# Выбор порога по ГИСТОГРАММЕ

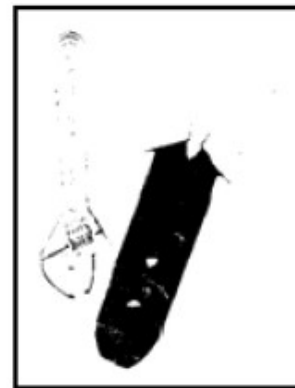
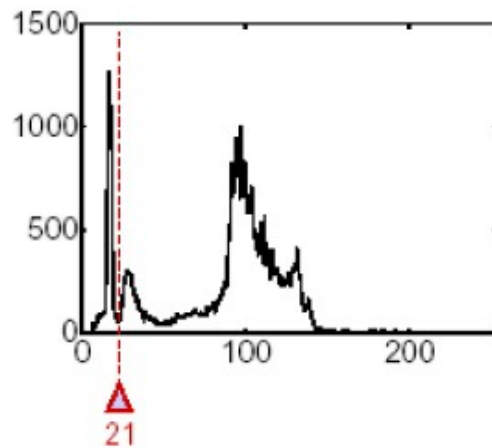
Изображение и его гистограмма



Порог 50



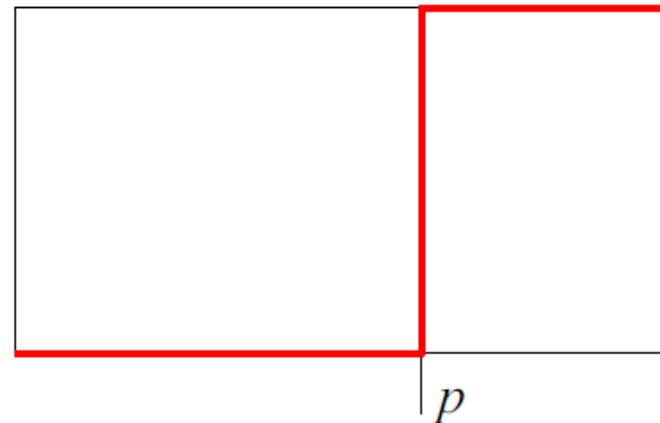
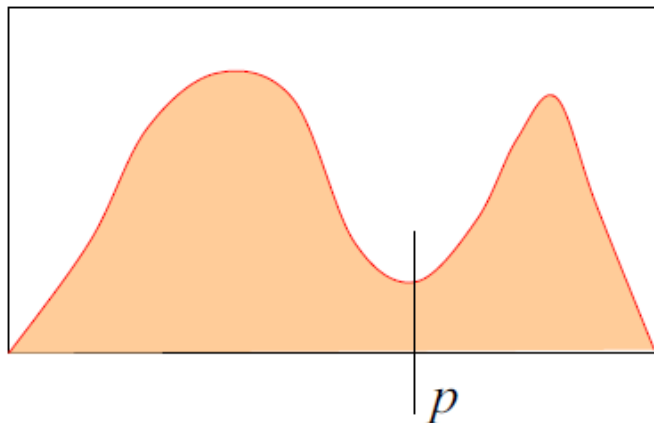
Порог 75



# Выбор порога бинаризации

Пусть  $H(D)$ ,  $D = 0, 1, \dots, 255$  - гистограмма яркости, имеющая выраженную двухмодальную структуру (двугорбый верблюд).

Выбираем порог бинаризации  $p$ . Вся картинка (все пиксели) разбивается на два подмножества: в первом яркость всех пикселей  $\leq p$ , а во втором -  $> p$ .





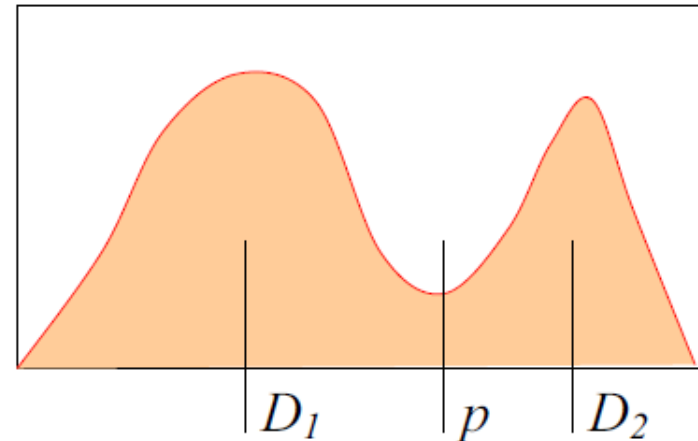
# Выбор порога бинаризации

Вычислим среднюю яркость для первого подмножества:

$$D_1 = \frac{\sum_{k=0}^p k \cdot H(k)}{\sum_{k=0}^p H(k)} \quad (1)$$

и для второго подмножества:

$$D_2 = \frac{\sum_{k=p+1}^{255} k \cdot H(k)}{\sum_{k=p+1}^{255} H(k)}. \quad (2)$$



# Модель двух классов

Теперь рассмотрим величины  $D_1$  и  $D_2$  в качестве центров двух кластеров. Смысл в этом такой. Задачу бинаризации мы рассмотрим, как задачу классификации. Нам нужно для каждого пиксела с яркостью  $D$  принять решение, к какому классу его отнести: к 0 или к 255. Используем алгоритм 1NN – одного ближайшего соседа. Если  $D$  ближе к  $D_1$ , то полагаем, что это класс 0, а если ближе к  $D_2$ , то полагаем класс 255.

Очевидно, что решающее правило имеет вид

$$\text{Класс}(D) = \begin{cases} 0 & \text{при } D \leq \frac{D_1 + D_2}{2} \\ 255 & \text{при } D > \frac{D_1 + D_2}{2} \end{cases}$$

Получаем, что новый порог есть не  $p$ , а  $\frac{D_1 + D_2}{2}$ .

# Уравнение для выбора порога

Но поскольку  $D_1 = D_1(p)$  и  $D_2 = D_2(p)$ , возникает естественное желание найти такое  $p$ , для которого эти пороги совпадают, т.е.

$$\frac{D_1(p) + D_2(p)}{2} = p.$$

Решаем уравнение

$$F(p) = D_1(p) + D_2(p) - 2p = 0.$$

Решение всегда есть, так как

$$F(0) = D_1(0) + D_2(0) - 2 \cdot 0 > 0 \text{ и}$$

$$F(255) = D_1(255) + D_2(255) - 2 \cdot 255 < 0.$$

# Алгоритм вычисления порога

Алгоритм поиска порогового значения  $p$  состоит в переборе значений  $F(p)$  по  $p = 0, 1, \dots$  до момента смены знака  $F(p)$  с плюса на минус.

Для того, чтобы не вычислять для каждого  $p$  суммы в выражениях (1) и (2), строим инкрементный алгоритм, исходя из следующих соотношений:

$$D_1(p) = \frac{S(p)}{R(p)}, \quad S(p) = \sum_{k=0}^p k \cdot H(k),$$

$$R(p) = \sum_{k=0}^p H(k),$$

$$D_2(p) = \frac{S(255) - S(p)}{R(255) - R(p)},$$

$$S(p+1) = S(p) + (p+1) \cdot H(p+1),$$

$$R(p+1) = R(p) + H(p+1).$$

# Метод Оцу

Метод Оцу ищет порог, уменьшающий дисперсию внутри класса, которая определяется как взвешенная сумма дисперсий двух классов:

$$\sigma_w^2(t) = \omega_1(t)\sigma_1^2(t) + \omega_2(t)\sigma_2^2(t) ,$$

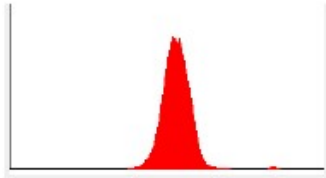
где веса  $w_i$  — это вероятности двух классов, разделенных порогом  $t$ ,  $\sigma_i$  — дисперсия этих классов.

Минимизация дисперсии *внутри* класса равносильна максимизации дисперсии *между* классами

$$\sigma_b^2(t) = \sigma^2 - \sigma_w^2(t) = \omega_1(t)\omega_2(t)[\mu_1(t) - \mu_2(t)]^2$$

где  $\mu_i$  — средние значения в классах.

# Пример бинаризации Оцу



Полутоновое изображение



Растяжение яркости



Бинаризация Оцу

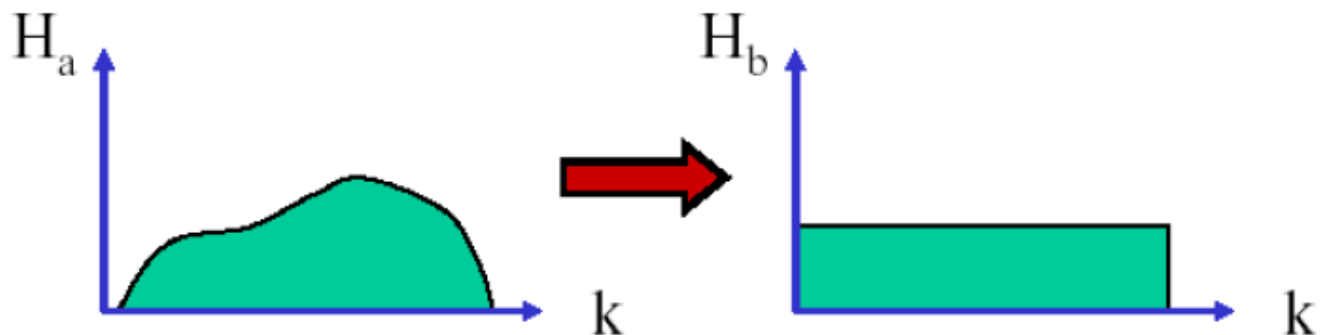
# Эквализация гистограмм

Для наилучшего визуального разрешения можно переопределить распределение яркости в изображении более равномерно.

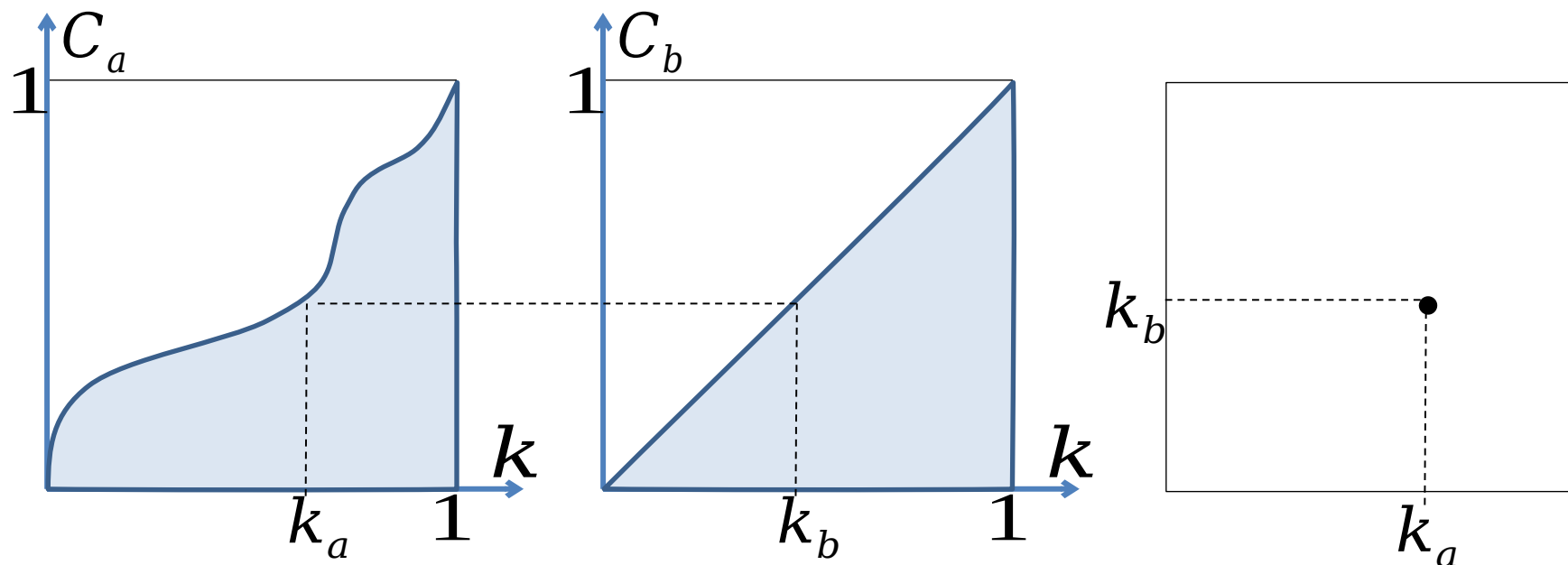
Зададим преобразование яркости  $I_B = f(I_A)$

таким образом, чтобы:

- гистограмма  $H_B$  была как можно более близка к равномерной плотности;
- функция  $f$  - монотонно возрастает.



# Эквализация гистограмм

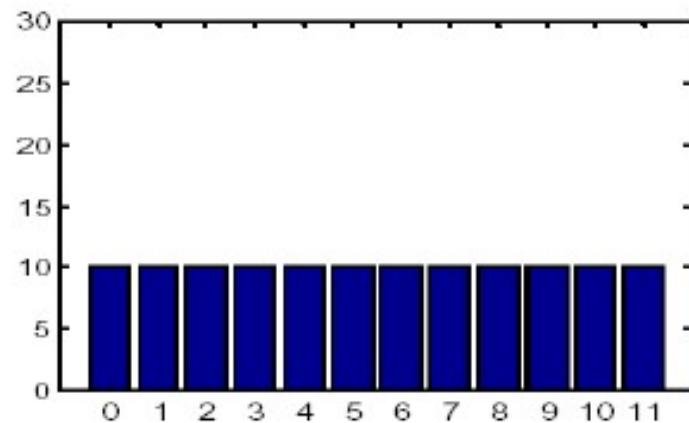
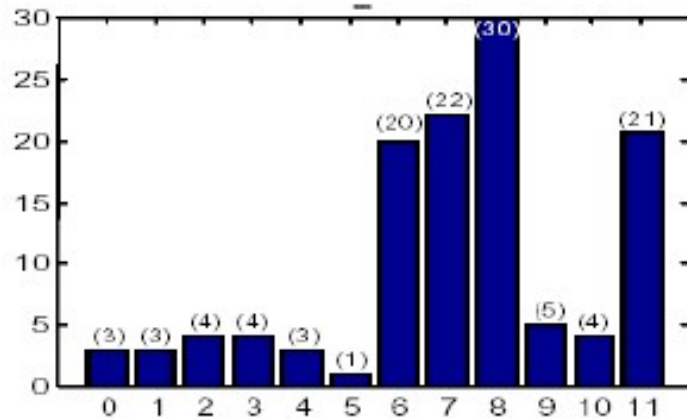


Нужно привести накопительную гистограмму к линейной функции

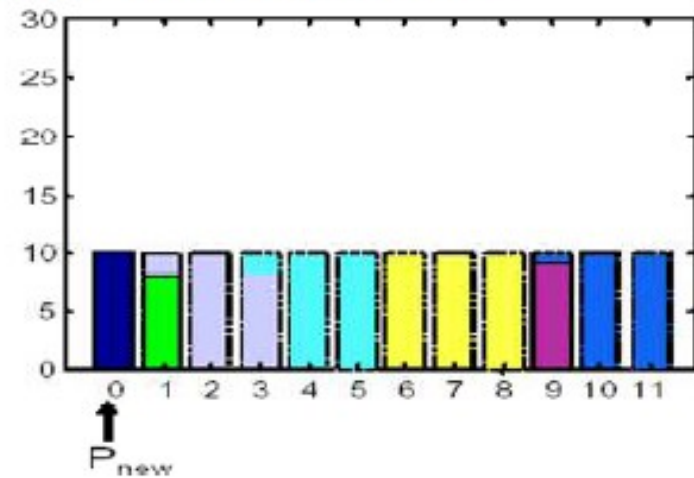
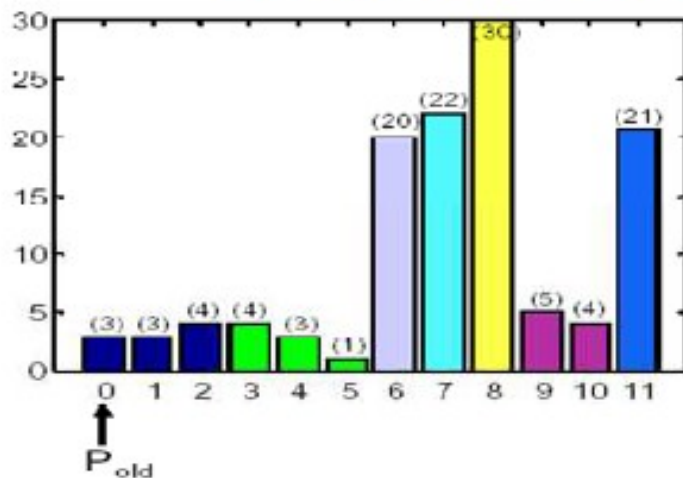
-



# Алгоритм двух указателей



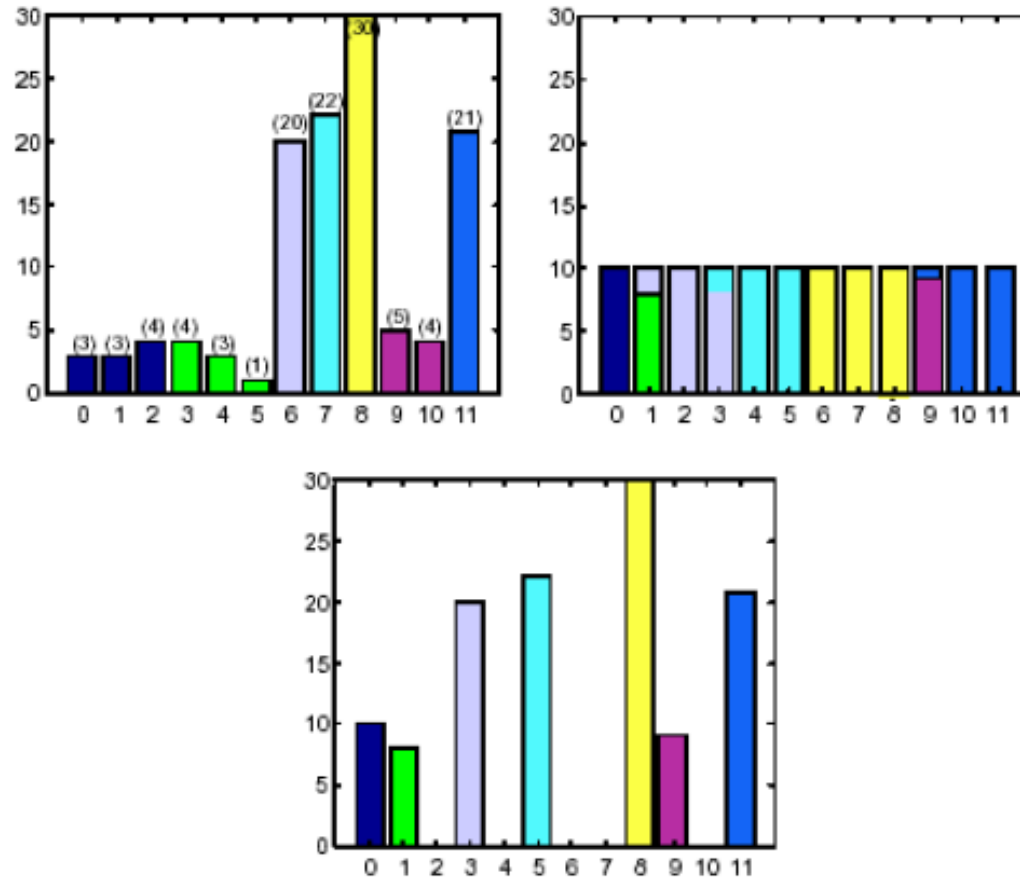
Исходная и выходная гистограммы



↓

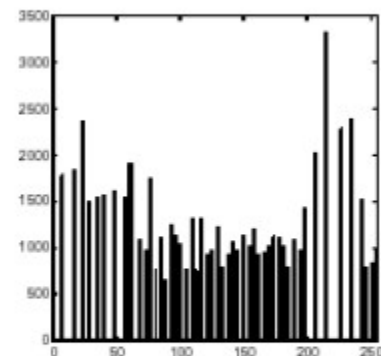
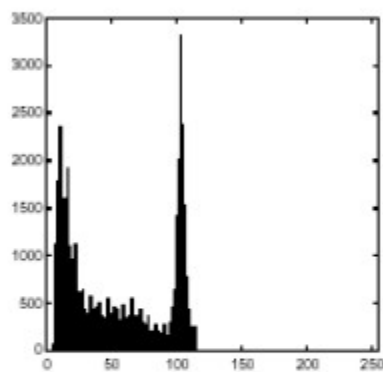
old	new
0	0
1	0
2	0
3	1
4	1
5	1
6	3
7	5
8	8
9	9
10	9
11	11

# Алгоритм двух указателей

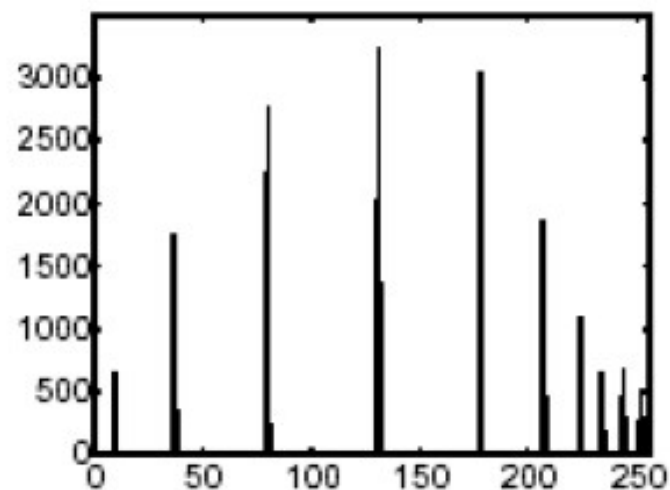
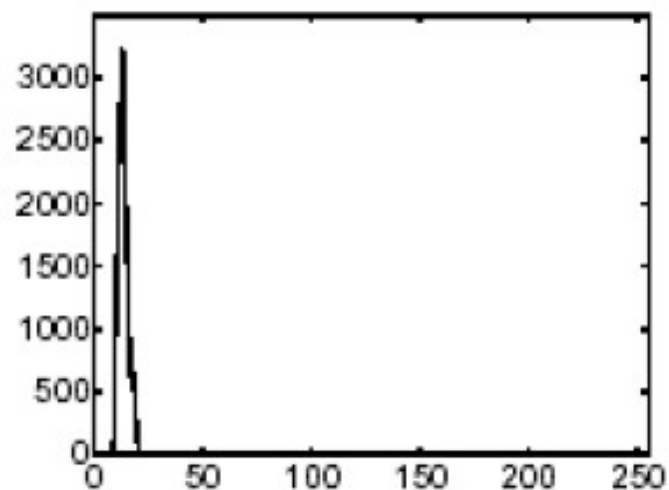


↓  
 0 1 2 3 4 5 6 7 8 9 10 11 old  
 0 0 1 1 1 3 5 8 9 9 11 new

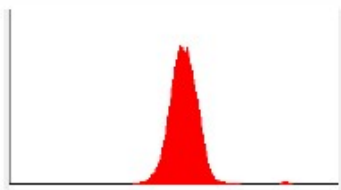
# Пример эквализации



# Пример эквализации



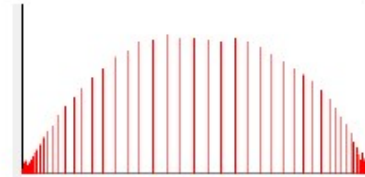
# Сравнение растяжения и эквализации



Полутоновое изображение



Растяжение яркости

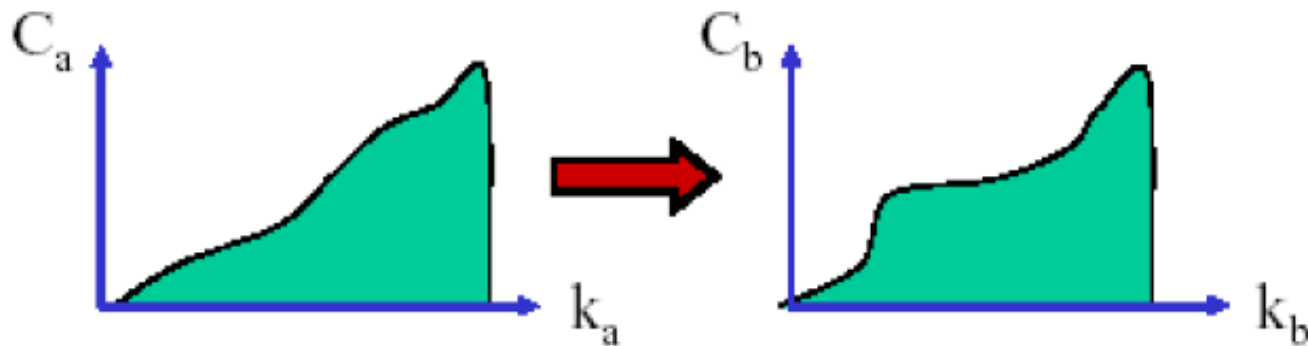


Эквализация

# Выравнивание гистограмм

Преобразование изображения  $A$  так, чтобы его гистограмма совпала с гистограммой  $B$

Используется при сравнении изображений одной и той же сцены, полученных при разном освещении



$$k_B = f(k_A) = C_B^{-1}(C_A(k_A))$$