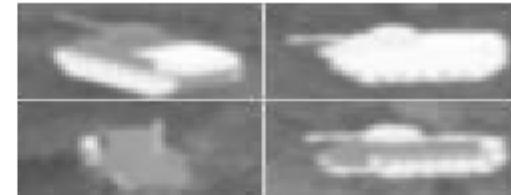


# Распознавание образов

# Примеры задач

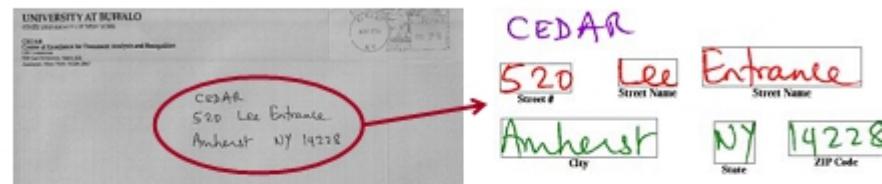
## ■ Машинное зрение

- Визуальная проверка
- Опознавание целей



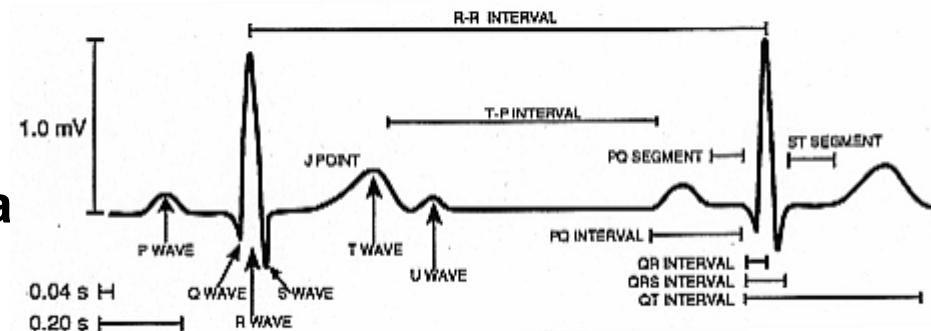
## ■ Распознавание символов

- Автоматическая сортировка почты
- Обработка налоговых деклараций



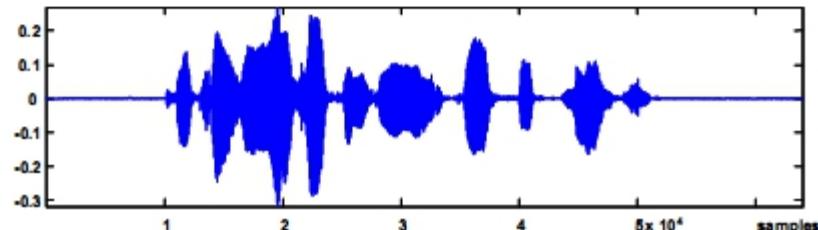
## ■ Компьютерная диагностика

- Медицинские изображения
- ЭКГ, ЭЭГ и т.п.



## ■ Распознавание речи

- Человеко-машинные интерфейсы
- Универсальный доступ

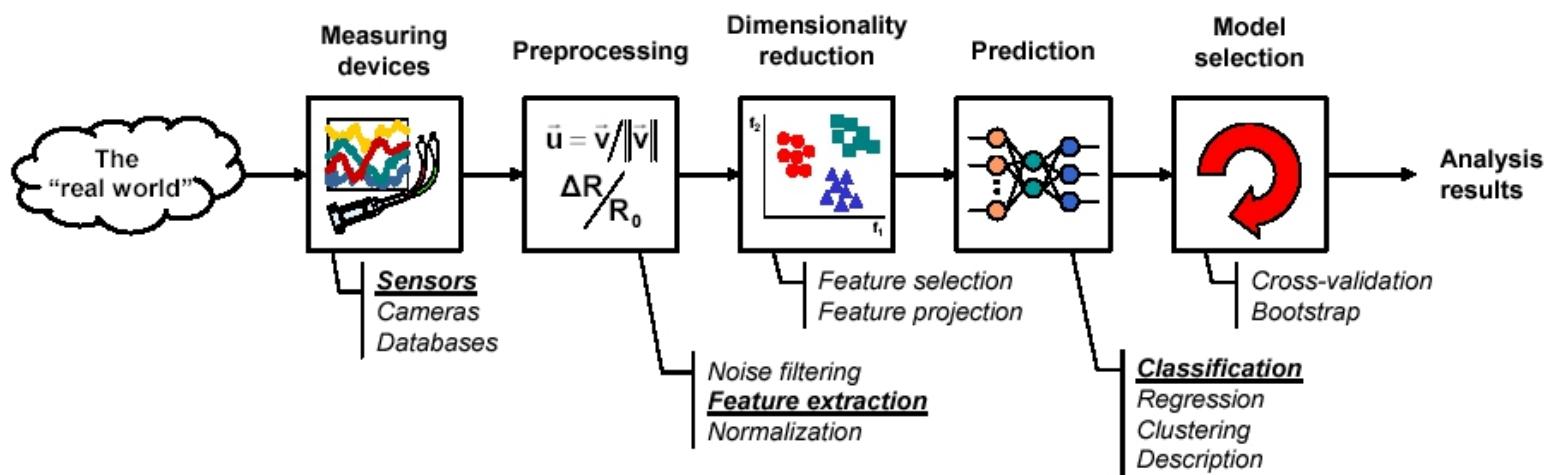


# Области и приложения, связанные с РО

- **Близкие области**
    - Адаптивная обработка сигнала
    - Машинное обучение
    - Нейронные сети
    - Роботостроение и машинное зрение
    - Когнитивные науки
    - Мат. статистика
    - Нелинейная оптимизация
    - Разведочный анализ данных
    - Нечеткие и генетические системы
    - Теория оценивания
    - Формальные языки
    - Структурное моделирование
    - Биологическая кибернетика
    - Нейро-вычисления
  - **Приложения**
    - Обработка изображений
    - Машинное зрение
    - Распознавание речи
    - Поиск целей
    - Распознавание текстов
    - Сейсмический анализ
    - Диагностика человека и механизмов
    - Дактилоскопия
    - Промышленный контроль
    - Финансовый прогноз
    - Медицинский диагноз
    - Обработка сигналов
- 3

# Структура типичной системы РО

- Базовая система РО содержит
  - Датчик(и)
  - Механизм предобработки
  - Механизм извлечения признаков
  - Алгоритм классификации
  - Набор описанных (классифицированных) образцов (обучающую выборку)



# Типы задач предсказания

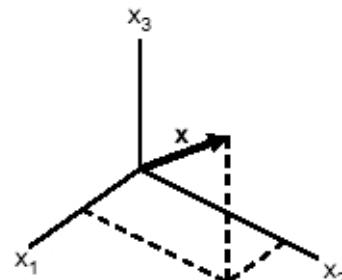
- Классификация
    - Задача РО – отнесение образца к некоторому классу
    - На выходе из системы РО – целочисленная метка
  - Регрессия
    - Обобщение задачи классификации
    - На выходе из системы РО – действительное число
  - Кластеризация
    - Задача объединения объектов в осмысленные группы
    - Система выдает (иногда иерархическую) группировку объектов
  - Описание
    - Задача представления объекта в терминах набора примитивов
    - Система выдает структурное синтаксическое описание объекта
- 5

# Признаки и образы

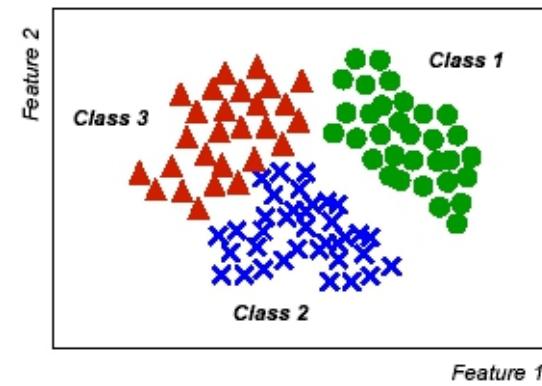
- **Признак**

- Признак – любая отличительная характеристика, качественная или количественная
- Определения
  - Комбинация  $d$  характеристик, представленная в виде  $d$ -вектора называется **вектором характеристик**
  - $d$ -мерное пространство, определяемое векторами характеристик, называется **пространством признаков**
  - Объекты представляются точками пространства признаков.

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_d \end{bmatrix}$$



Feature vector



Feature space (3D)

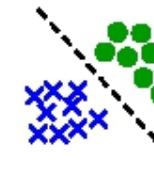
Scatter plot (2D)

# Признаки и образы

- **Образ**
  - Образ есть совокупность признаков, характеризующих объект
  - В задачах классификации образ есть пара переменных  $\{x, \omega\}$ , где
    - $x$  есть набор наблюдений (вектор признаков)
    - $\omega$  есть понятие за пределами наблюдения (метка наблюдения)

# Признаки и образы

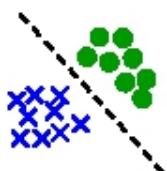
- Что такое «хороший» вектор признаков?
  - Качество вектора признаков связано с его разделяющей способностью для образов из разных классов
    - Образцы из одного класса должны иметь подобные вектора признаков
    - Образцы из разных классов должны иметь различные вектора признаков
- Свойства, зависящие от векторов признаков



*“Good” features*



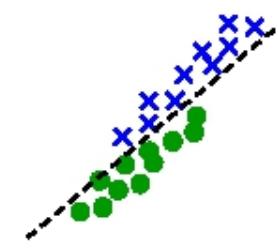
*“Bad” features*



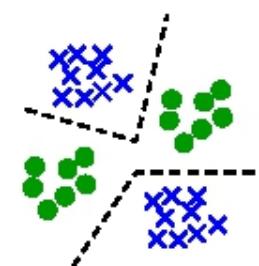
*Linear separability*



*Non-linear separability*



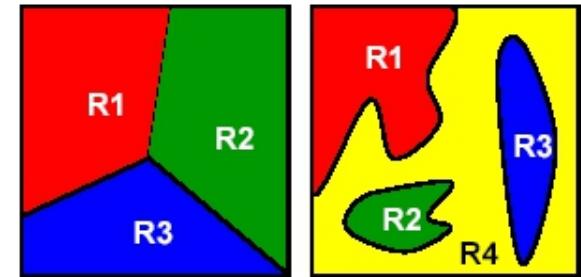
*Highly correlated features*



*Multi-modal*

# Классификаторы

- Назначение классификатора – разбиение пространства признаков на области классов
  - Границы между классами именуются решающими правилами
  - Классификация образа  $x$  состоит в определении класса, которому принадлежит  $x$  и отнесении его к этому классу

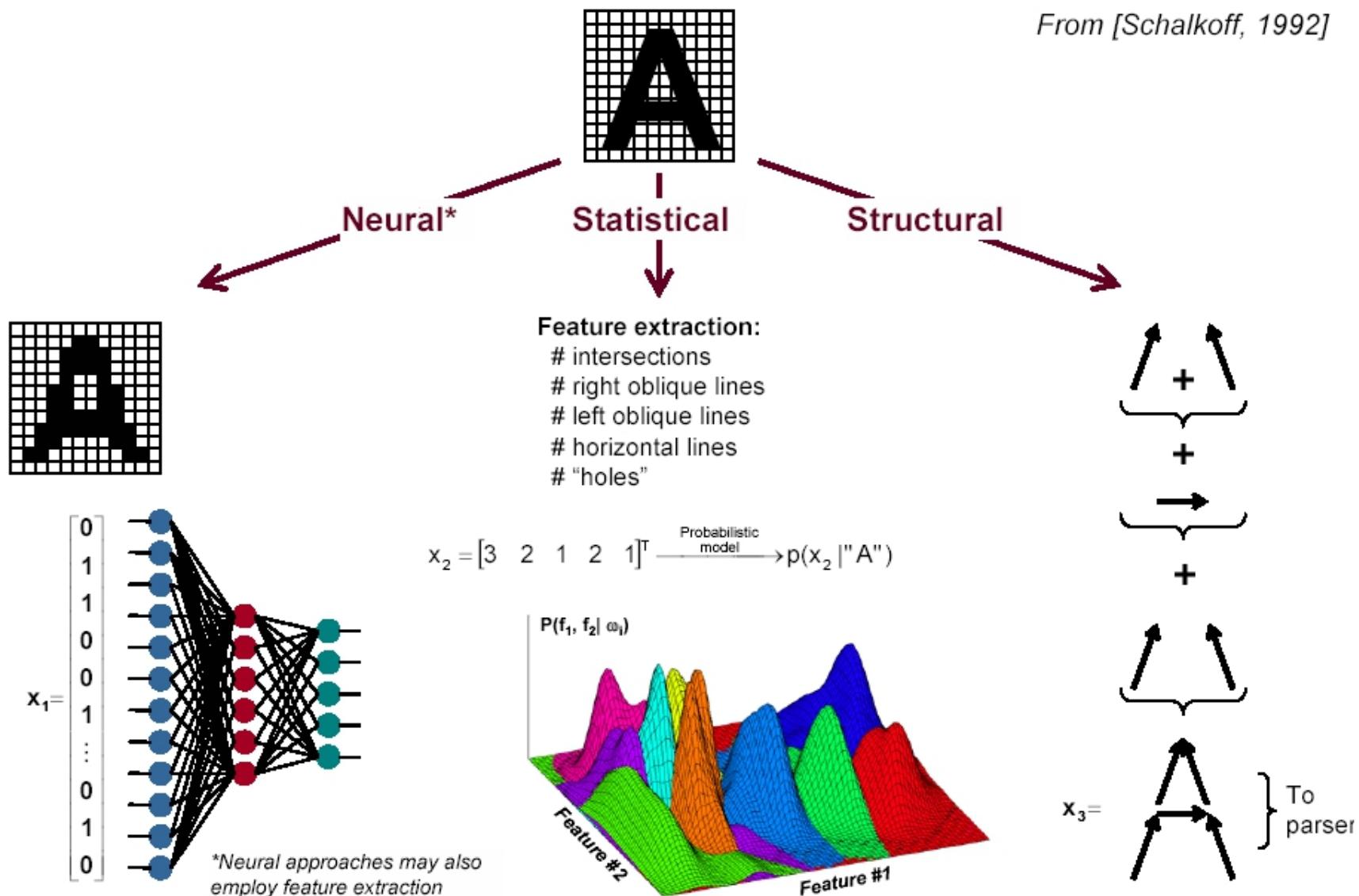


# Подходы к РО

- Статистический
  - Классификация образов основана на статистической модели признаков
- Нейросетевой
  - Классификация основана на отклике сети нейронов для заданного стимула
- Синтаксический
  - Основа классификации – структурное подобие образов
- Прочие
  - Нестатистические методы анализа данных
  - Нечеткие (fuzzy) методы
  - Деревья решений
  - ...

# Подходы к РО. Пример

From [Schalkoff, 1992]



# Цикл построения системы РО

- Сбор данных
  - Самый затратный по времени этап
  - Как много образцов нужно собрать?
- Выбор признаков
  - Критично для успешности проекта в целом
  - Требует априорной информации
- Выбор модели
  - Статистическая, нейронная, структурная, другая
  - Выбор параметров
- Обучение
  - Адаптация модели для описания данных по обучающей выборке
  - С учителем, без учителя
- Оценка результатов
  - Насколько хорошо данные описываются моделью?
  - Переобучение vs. Обобщение

# Цикл построения системы РО

- Ситуация
  - Автоматизация процесса сортировки рыбы (окунь/лосось)
    - Предприятию, занимающемуся упаковкой морепродуктов, требуется автоматизировать процесс сортировки по видам рыбы, идущей по конвейеру

**Морской окунь**

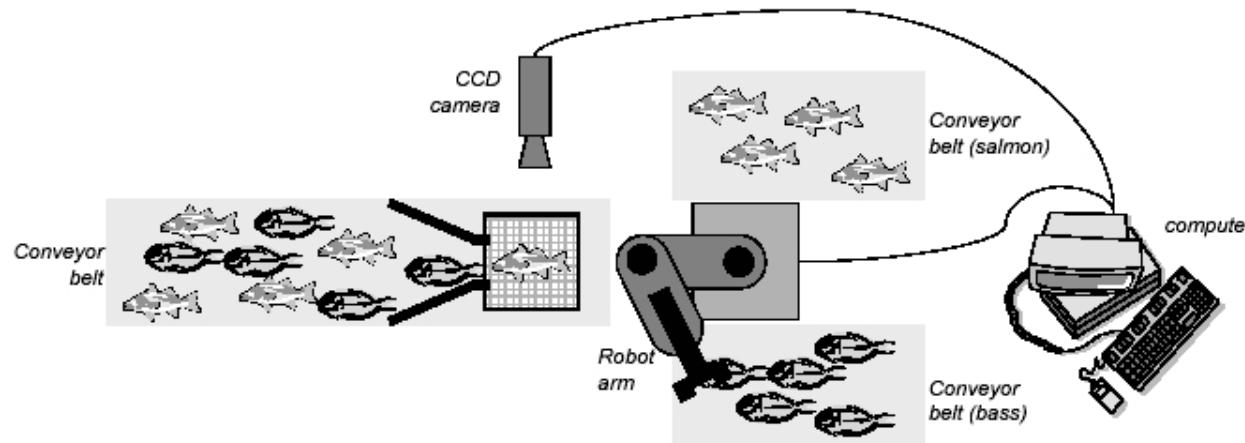


**Лосось**



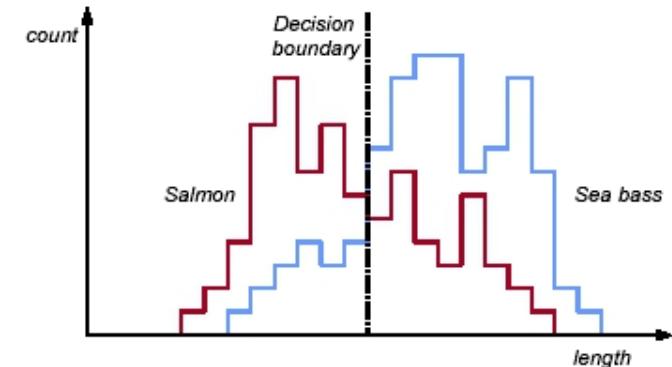
# Цикл построения системы РО

- Ситуация
  - Автоматизация процесса сортировки рыбы (окунь/лосось)
  - Система содержит следующие элементы
    - конвейерная лента с входящими продуктами
    - две конвейерные ленты с разделенными продуктами
    - роботизированная рука (взять/положить)
    - видео-система
    - компьютер для анализа данных и управления роботом



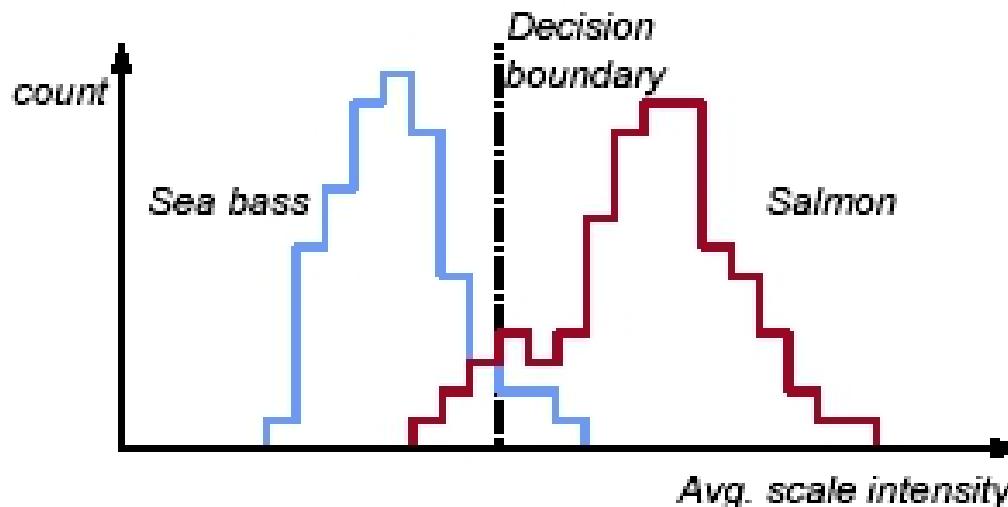
# Цикл построения системы РО

- Датчик
  - Камера захватывает изображение рыбы
- Предобработка
  - Обработка изображения (контрастирование, сегментация и т.п.)
- Выделение признаков
  - Например, известно, что в среднем окунь больше лосося
  - По изображению оценивается длина рыбы
- Классификация
  - Собрать выборку для двух видов
  - Вычислить распределение длин для обоих классов
  - Определить разделяющую границу (порог), минимизирующий ошибку классификации
  - После оценки вероятности ошибки получим никудышный результат – 40%  
15
  - Что дальше?



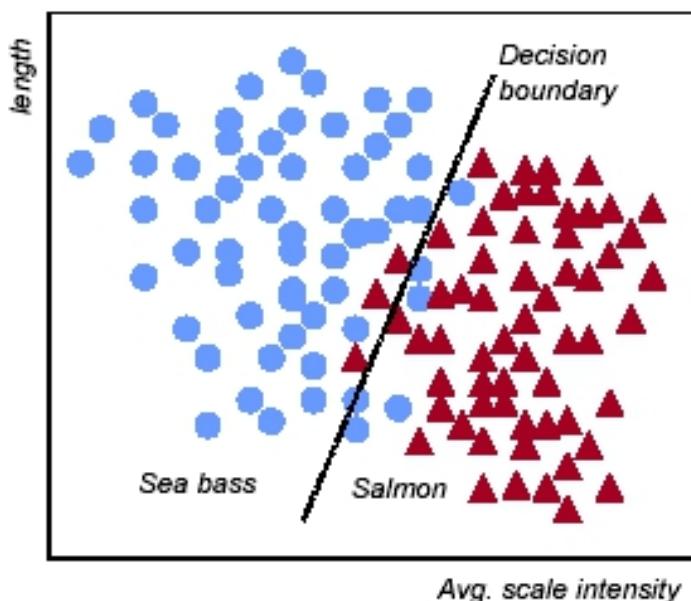
# Цикл построения системы РО

- Повышение эффективности системы РО
  - Для достижения правильной классификации в 95% случаев, перебираем признаки
    - Ширина, Площадь, Положение глаз по отношению ко рту и т.п...
    - ...только для того, чтобы выяснить, что ни один из них не обеспечивает разделения
  - В конце концов найден «хороший» признак – средний размер чешуйки



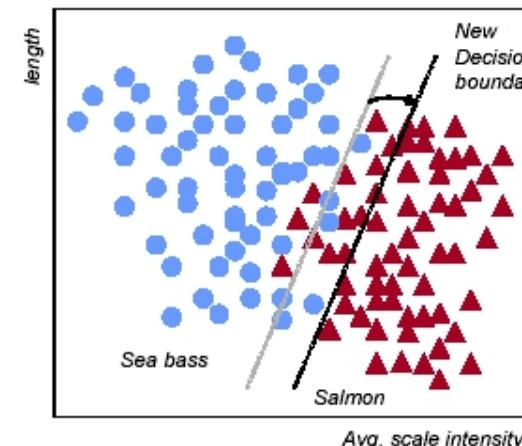
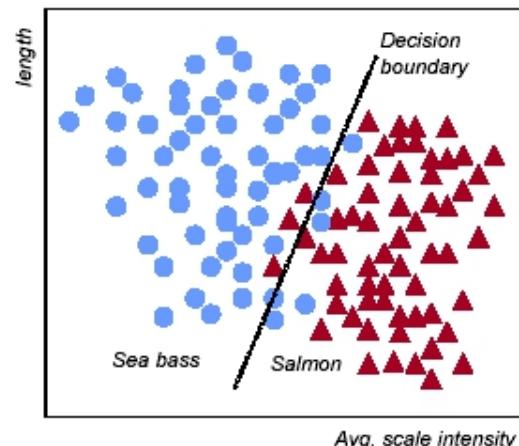
# Цикл построения системы РО

- Повышение эффективности системы РО
  - Комбинируя *длину* рыбы со *средним размером чешуеки*, улучшаем разделяющую способность классификатора
  - Построив линейную разделяющую функцию, получим точность классификации 95.7%



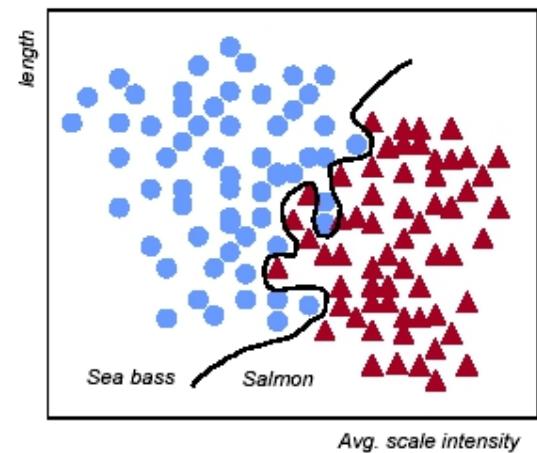
# Цикл построения системы РО

- Цена ошибки и точность классификации
  - Линейный классификатор построен так, чтобы минимизировать общую ошибку классификации
  - Является ли этот выбор функции наилучшим для автоматизируемого предприятия?
    - Плата за неправильную классификацию - иногда покупатель будет обнаруживать, что куплен лосось по цене дешевого окуня, и наоборот
  - Интуитивно, мы должны минимизировать именно цену ошибки выбирая разделяющую функцию



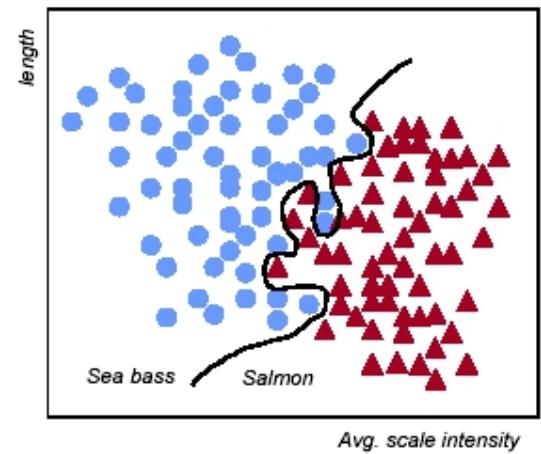
# Цикл построения системы РО

- Обобщающая способность
  - Точность построенного линейного классификатора – 95.7% отвечает поставленным требованиям, но мы решили улучшить этот результат
    - Мы строим нейросеть с 5 скрытыми слоями, комбинацией логистической функции и гиперболического тангенса, обучаем ее методом Левенберга-Марквардта и получаем впечатляющую точность классификации – 99.9975% и следующую разделяющую функцию
  - Удовлетворенные таким классификатором мы внедряем его в систему на предприятии



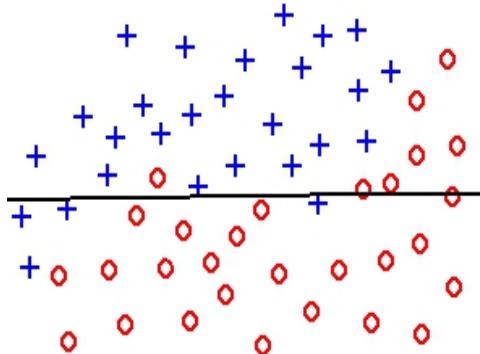
# Цикл построения системы РО

- Обобщающая способность
  - Точность построенного линейного классификатора – 95.7% отвечает поставленным требованиям, но мы решили улучшить этот результат
    - Мы строим нейросеть с 5 скрытыми слоями, комбинацией логистической функции и гиперболического тангенса, обучаем ее методом Левенберга-Марквардта и получаем впечатляющую точность классификации – 99.9975% и следующую разделяющую функцию
  - Удовлетворенные таким классификатором мы внедряем его в систему на предприятии
    - Через несколько дней менеджер предприятия жалуется, что система ошибочно классифицирует в среднем 25% рыбы.
    - Что случилось?

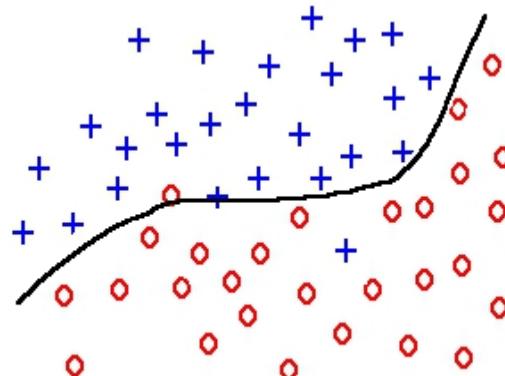


# Недо- и переобучение

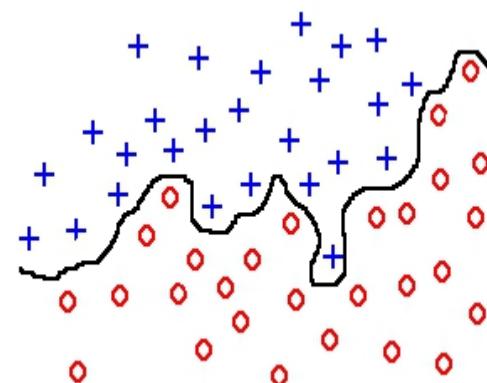
- **Задача:** насколько гибок должен быть классификатор?



излишнее обобщение,  
недообучение,  
underfitting



хорошая модель  
good fit



излишняя  
избирательность,  
переобучение,  
overfitting

- **Проблема обобщения:** малый эмпирический риск  $R_{\text{emp}}$  не означает, что истинный ожидаемый риск  $R$  будет мал



# Статистический анализ изображений

# Статистический анализ изображений

- Изображение – реализация случайной величины
- Простые статистические характеристики изображений
- Использование статистических характеристик при улучшении изображений
- Контрастирование
- Эквализация гистограмм
- Попиксельные преобразования изображений

# Изображение – реализация случайной величины

- Изображения могут рассматриваться как реализации случайной величины
- Цель статистического анализа изображений -- оценивание по имеющейся реализации свойств случайной величины (закона распределения).

# Случайная величина

- *Случайной величиной* (с.в.) называется числовая функция от случайных событий.
- Примеры
  - Число очков, выпавших при бросании игральной кости  
(дискретная с.в.: принимает значения из дискретного множества  $M = \{1, 2, 3, 4, 5, 6\}$  )
  - Рост случайно выбранного из учебной группы студента  
(непрерывная с.в.: принимает значения из непрерывного числового множества – из промежутка числовой прямой  $I = [100, 230]$ )

# Функция распределения

- Если  $\xi$  – случайная величина, то функция

$$F(x) = F_\xi(x) = P(\xi \leq x)$$

называется *функцией распределения* (ф.р.) случайной величины  $\xi$ .

- Приращения функции распределения имеют простой смысл:

$$F_\xi(b) - F_\xi(a) = P(\xi \leq b) - P(\xi \leq a) = P(a < \xi \leq b)$$

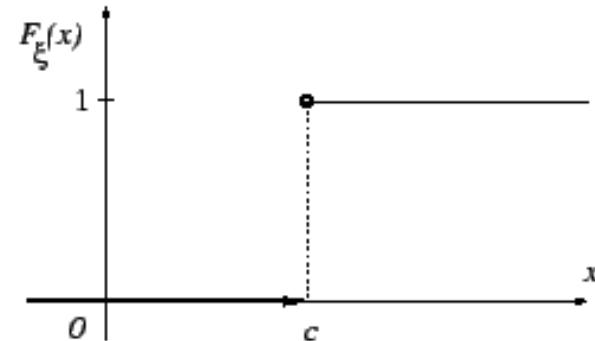
# Функция распределения

- Функция распределения случайной величины обладает следующими свойствами:
  1.  $F(x)$  определена на всей числовой прямой  $R$ ;
  2.  $F(x)$  не убывает;
  3.  $F(-\infty) = 0, F(+\infty) = 1$ ;
  4.  $F(x)$  непрерывна справа.

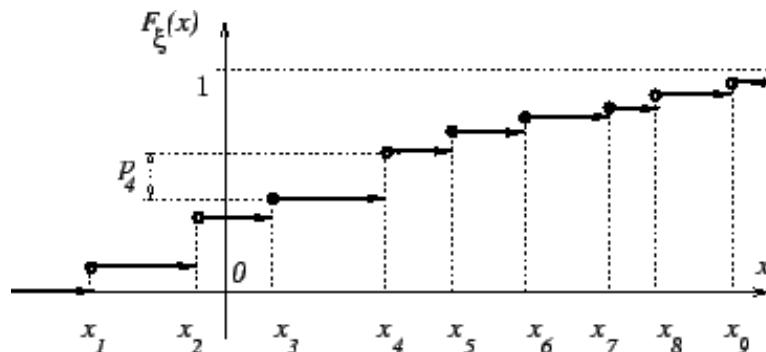
# Функция распределения

- Примеры

- $$F_{\xi}(x) = \begin{cases} 0, & x < c, \\ 1, & x \geq c. \end{cases}$$



- Дискретная с.в. со значениями  $x_1 < \dots < x_k < \dots$ , принимаемыми с вероятностями  $p_1, \dots, p_k, \dots$  соответственно:



# Плотность распределения

- Если ф.р.  $F(x)$  дифференцируема, то *плотность распределения*  $p(x)$  связана с ней следующим образом:

$$F_{\xi}(x) = \int_{-\infty}^x p_{\xi}(t)dt, \quad p_{\xi}(x) = \frac{dF_{\xi}(x)}{dx}$$

- Очевидно, что

$$\int_{-\infty}^{+\infty} p_{\xi}(x)dx = 1$$

и

$$P(a < x < b) = F_{\xi}(b) - F_{\xi}(a) = \int_a^b p_{\xi}(x)dx$$



# Характеристики случайной величины

- Вариация
- Математическое ожидание
- Стандартное отклонение
- Коэффициент асимметрии
- Коэффициент эксцесса
- Энтропия



Функция  
распределения

# Моменты

- Начальный момент  $m_k$   $k$ -го порядка

$$m_k = \sum_i x_i^k p_i$$

# Дисперсия и отклонение

- Дисперсия

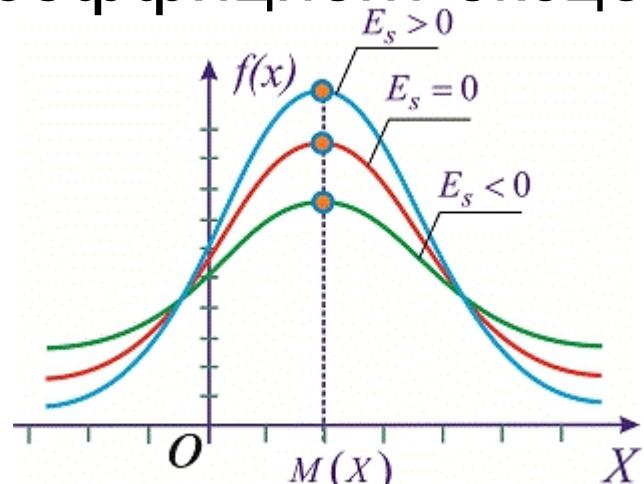
$$D\xi = u_2 = m_2 - m_1^2$$

- Среднеквадратичное отклонение

$$\sigma = \sqrt{u_2}$$

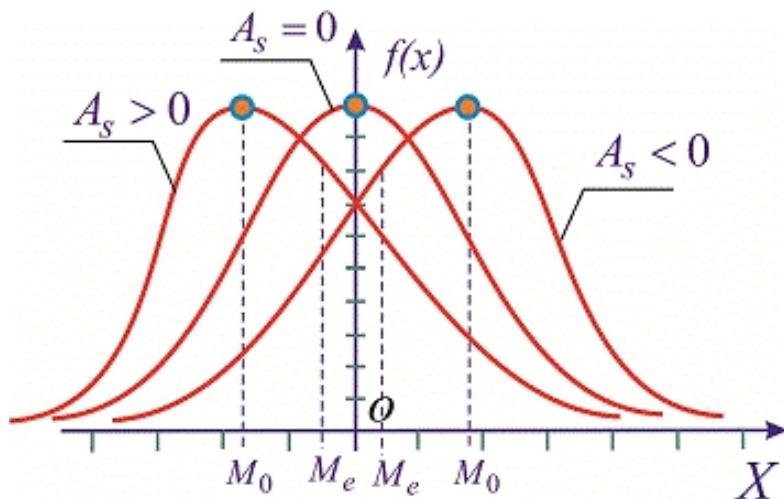
# Эксцесс и асимметрия

- Коэффициент эксцесса (крутисти)



$$E_s = g_1 = \frac{u_3}{\sigma^3}$$

- Коэффициент асимметрии (скошенности)



$$A_s = g_2 = \frac{u_4}{\sigma^4} - 3$$

# Энтропия

- Энтропия (анти-информация Шеннона)

$$E = - \sum_i p_i \log_2 p_i$$

$$E_{\max} = \log_2 (h_{\max} - h_{\min} + 1)$$

- Степень близости распределения к равномерному (относительная энтропия)

$$D = 1 - E/E_{\max}$$



Улучшение изображений  
с помощью  
попиксельных преобразований

# Улучшение изображений

- **Цель:** преобразовать изображение таким образом, чтобы его было проще использовать при решении определенных задач.
- Общей теории улучшения изображений не существует.
- Набор специализированных средств, применяемых на усмотрение пользователя.
- Субъективность восприятия.
- **Методы:**
  - В пространственной области
    - Попиксельные операции
    - Преобразование интенсивностей пикселов: не зависит от данных
    - Преобразование гистограмм: зависит от данных
    - Арифметические операции
  - Пространственная фильтрация
  - Фильтрация в частотной области

# Преобразование интенсивностей пикселов

- Осуществляется попиксельно:

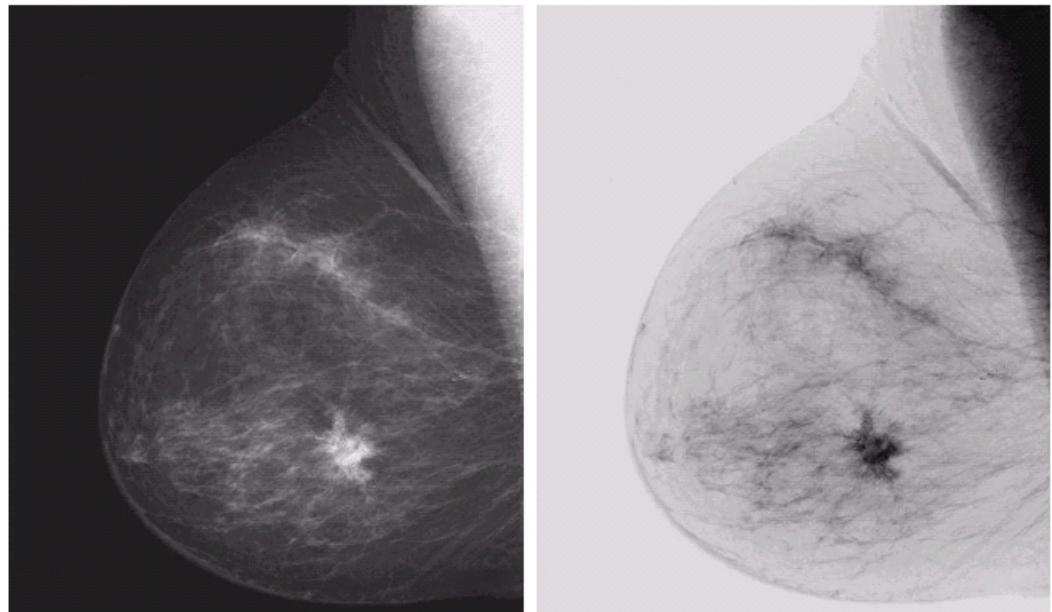
$$s = T(r),$$

где  $r$  – исходная интенсивность пикселя,  
 $s$  – результирующая интенсивность

- Универсальный метод: не учитывает специфику обрабатываемых данных.
- Классы преобразований
  - Обращение (негативизация)
  - Логарифмические
  - Степенные (полиномиальные)
  - Кусочно-линейные

# Негатив изображения

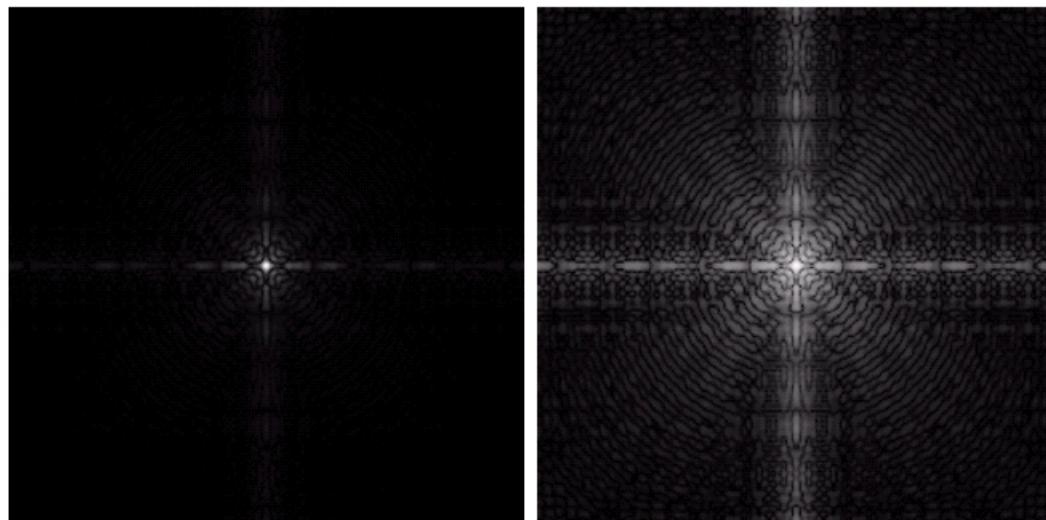
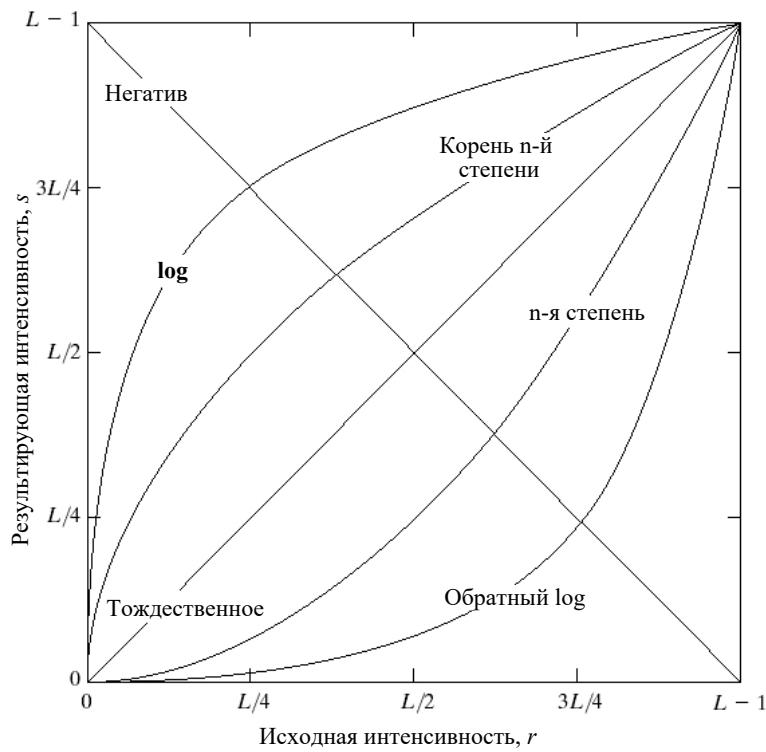
- $s = T(r) = L - 1 - r$
- Аналогично фотонегативу.
- Применимо для улучшения белых или серых деталей на темном фоне.



# Логарифмические преобразования

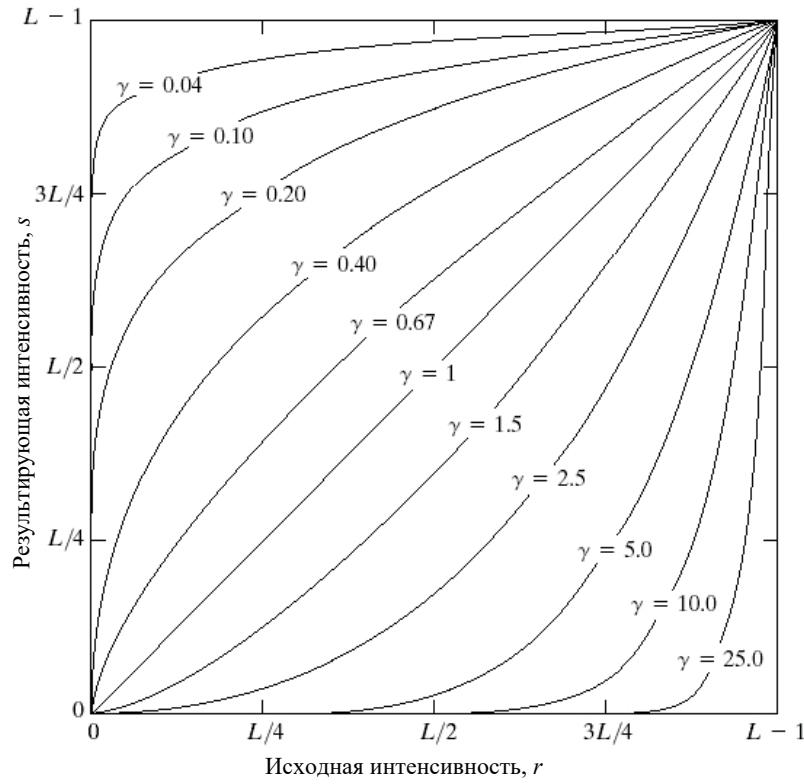
- $s = T(r) = c \log(1+r)$

- Контрастирование «темных» значений интенсивностей улучшает восприятие деталей

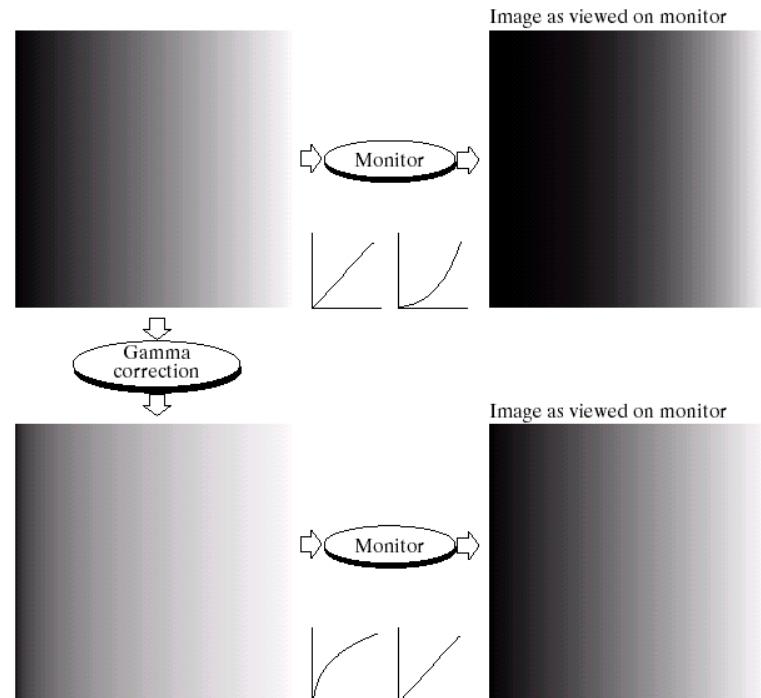


# Степенное преобразование

$$\bullet \quad s = T(r) = c s^\gamma$$



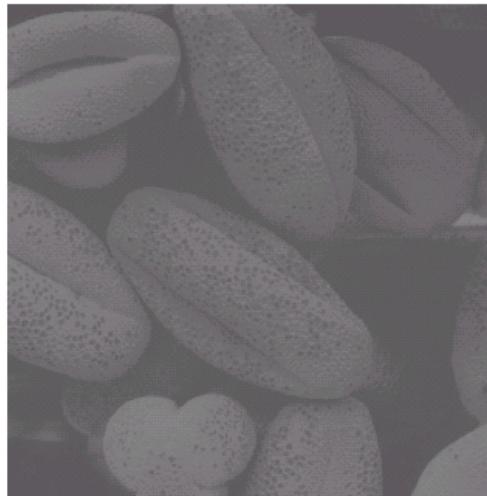
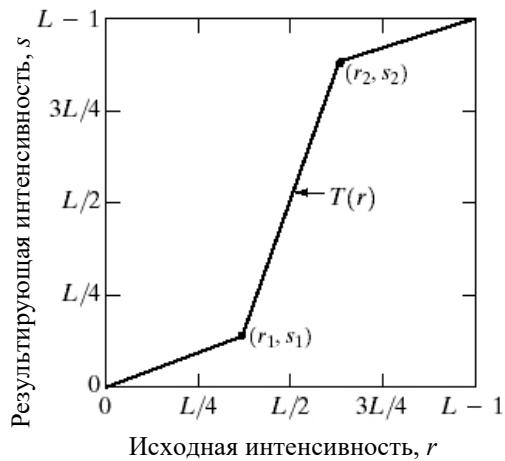
- Гамма-коррекция используется для адаптации изображений к особенностям дисплеев



# Кусочно-линейные преобразования

- Позволяют конструировать сложные функции преобразования  $T(r)$ .
  - Контрастирование
  - Выделение поддиапазона серого
  - Выделение битовых плоскостей
  - Бинаризация
  - и т.д.

# Контрастирование

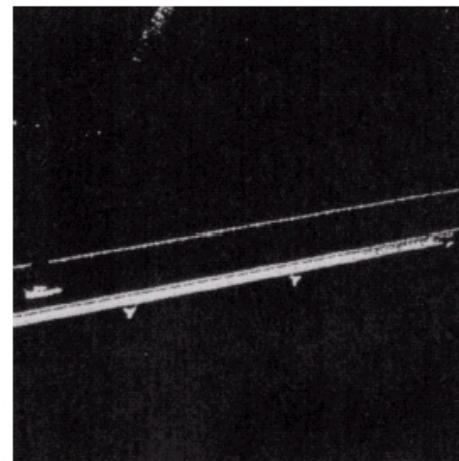
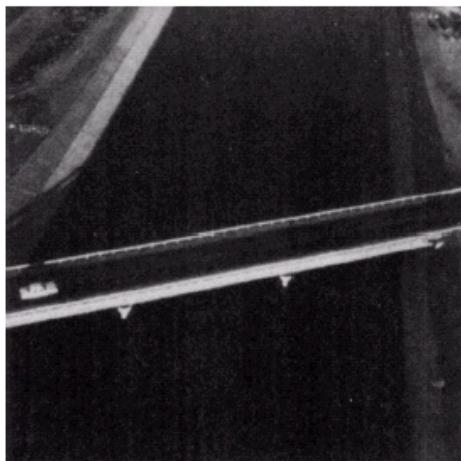
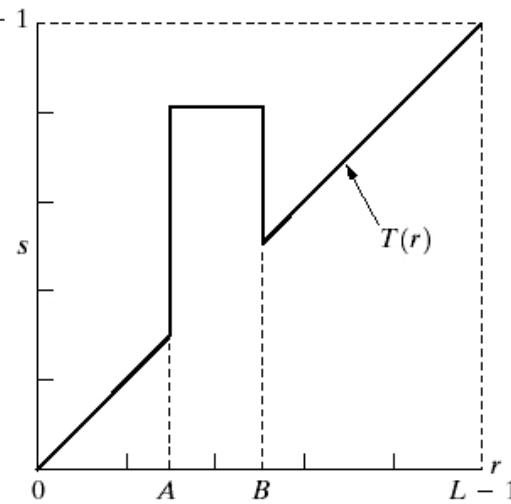
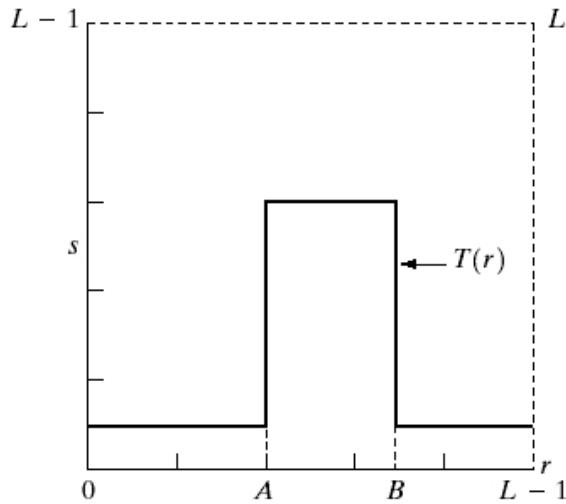


a	b
c	d

## Контрастирование

- (a) Вид преобразования
- (b) Неконтрастное изображение
- (c) Результат контрастирования
- (d) Результат пороговой бинаризации

# Выделение поддиапазонов серого



a	b
c	d

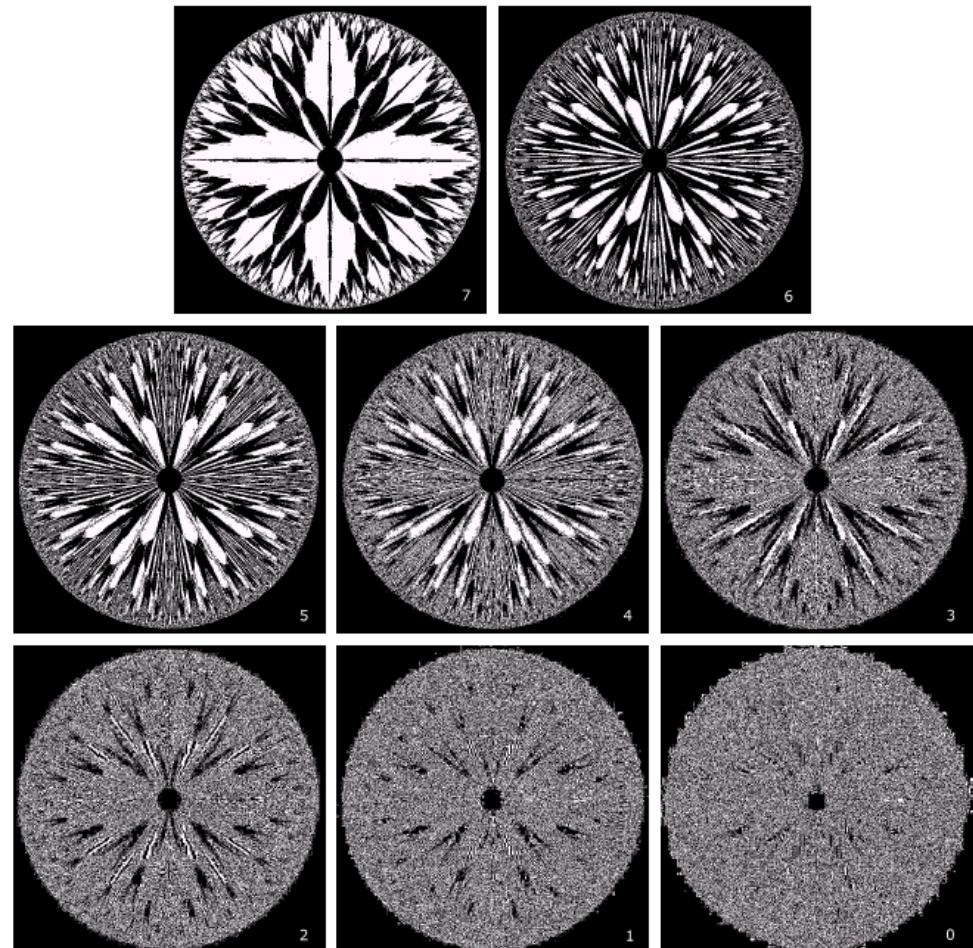
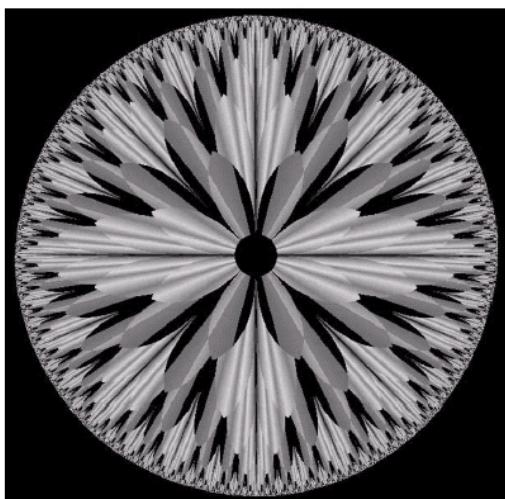
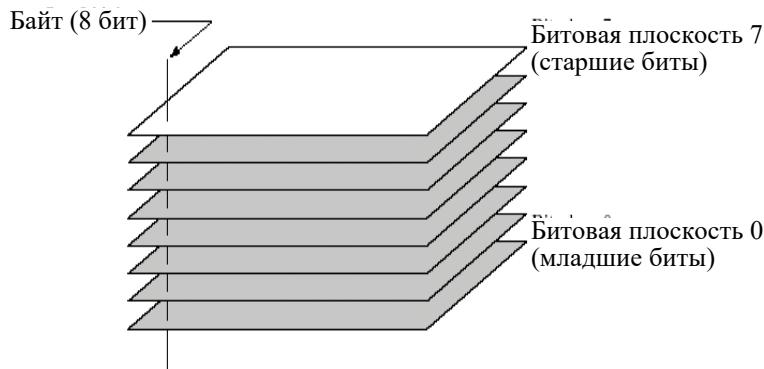
(a) Преобразование, осветляющее пиксели с интенсивностью из диапазона  $[A,B]$  и затемняющее все прочие пиксели до константного уровня

(b) Преобразование, осветляющее пиксели с интенсивностью из диапазона  $[A,B]$  и сохраняющее значения прочих пикселов

(c) Исходное изображение

(d) Результат преобразования вида (a)

# Выделение битовых плоскостей



# Выделение битовых плоскостей



Пошаговое удаление битовых плоскостей (оригинал)

# Выделение битовых плоскостей



Удалены 4 битовых плоскости

# Выделение битовых плоскостей



Удалены 6 битовых плоскостей

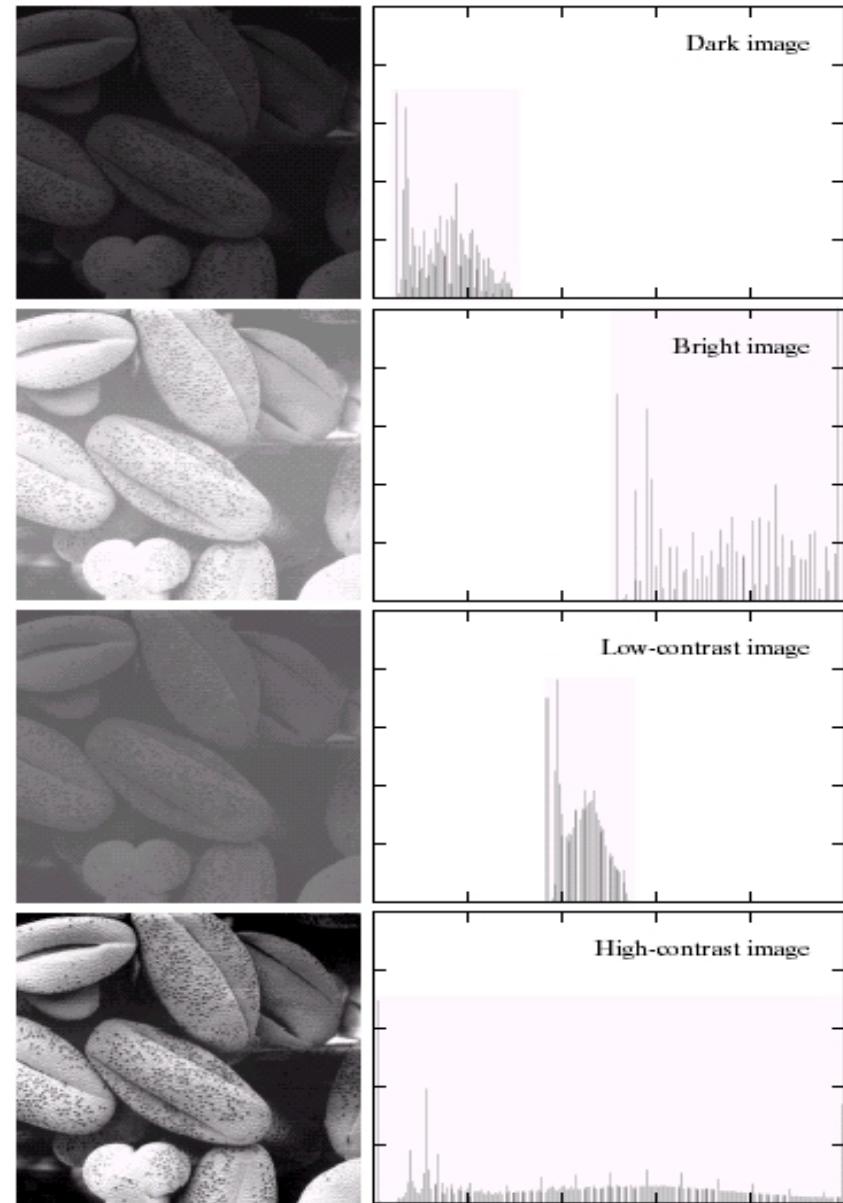
# Выделение битовых плоскостей



Удалены 7 битовых плоскостей

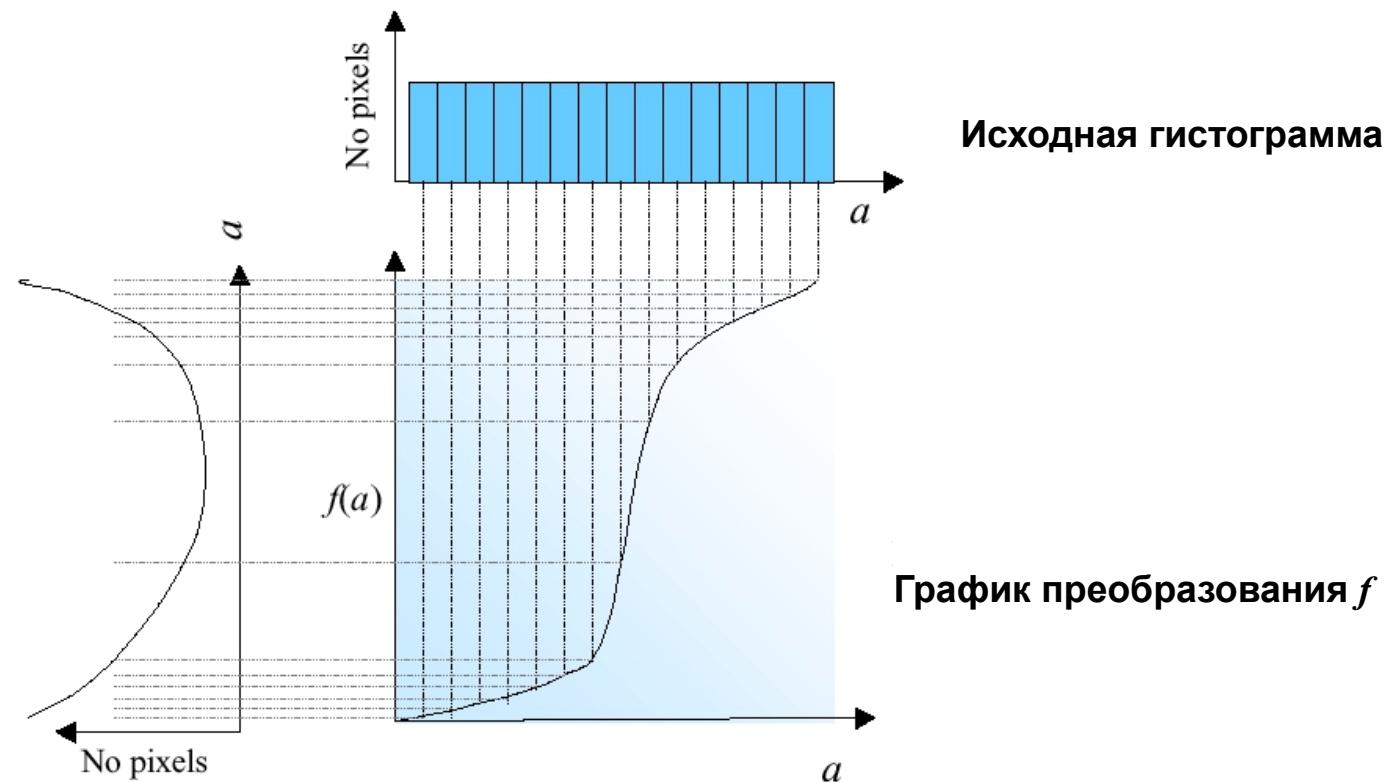
# Преобразование гистограмм

- Попиксельный метод улучшения изображения, зависящий от данных.
- Гистограмма  $\approx$  функция плотности распределения значений пикселов.
  - Предположение: значения пикселов изображения – одинаково распределенные случайные величины
  - На рисунках приведены примеры изображений и соответствующие им гистограммы



# Преобразование гистограмм

Под влиянием попиксельного преобразования изменяется гистограмма изображения



Преобразованная гистограмма (лежит на боку)

# Прямая и обратная задачи статистического анализа

## ПРЯМАЯ:

Определение вероятностных  
характеристик (плотности  
вероятности) изображений,  
прошедших попиксельную  
обработку

## ОБРАТНАЯ:

Построение преобразования,  
переводящего изображения с  
известной плотностью  
вероятности  $p_x(x)$  в  
изображение с наперед  
заданной плотностью  
вероятности  $p_y(y)$

# Прямая и обратная задачи статистического анализа

$y = T(x)$  – попиксельное преобразование

$x = \tau(y)$ ,  $\tau = T^{-1}$  – обратная функция

$p_x(x)$ ,  $p_y(y)$  – плотности вероятности исходного и преобразованного изображений

$p_x(x)dx = p_y(y)dy$  – события  $x$   $dx$  и  $y$   $dy$  равновероятны

$p_y(y) = p_x(x) \frac{dx}{dy} = p_x(\tau(y)) \frac{d\tau(y)}{dy}$  – прямая задача

$$dy = \frac{p_x(x)dx}{p_y(y)}$$

$y = \frac{1}{p_y(y)} \int p_x(x)dx = \frac{F(x)}{p_y(y)}$  – обратная задача

# Эквализация гистограмм

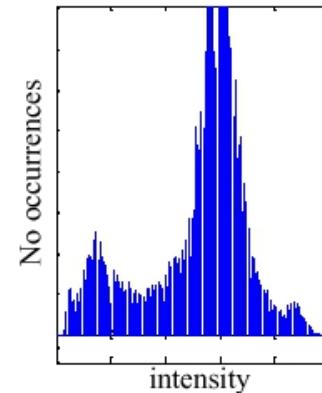
- Попиксельное преобразование, равномерно перераспределяющее уровни серого на весь динамический диапазон
  - Полностью автоматическое
  - Неоднородное (зависит от данных)
  - Контрастирующее (как правило)
- Как правило, не обеспечивает *точного* равномерного распределения уровней серого на результирующем изображении (из-за дискретности)

# Эквализация гистограмм

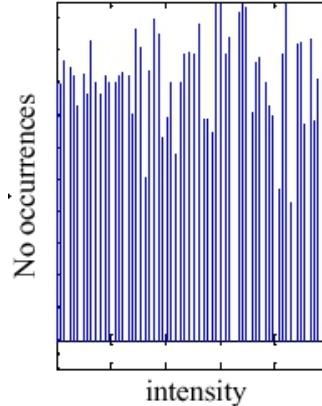
$$p_y(y) = \begin{cases} 1/(y_{\max} - y_{\min}), & y \in [y_{\min}, y_{\max}] \\ 0, & \text{иначе} \end{cases}$$

$$y = (y_{\max} - y_{\min}) * F(x) + y_{\min}$$

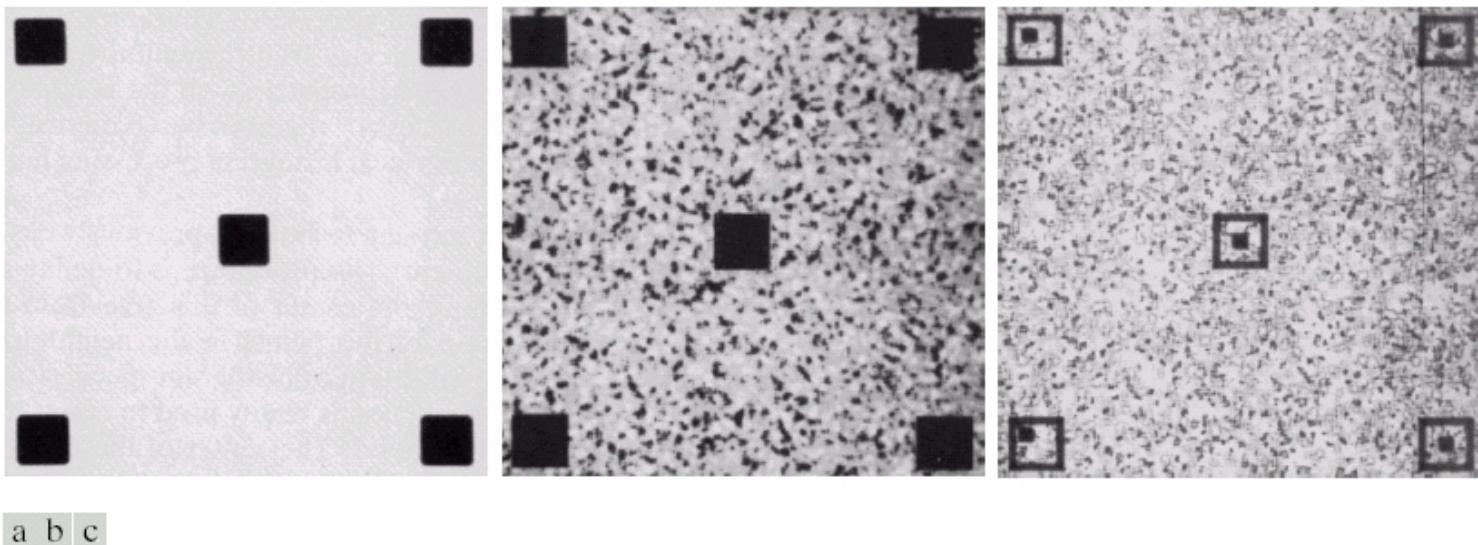
Исходные гистограмма и изображение



После эквализации



# Глобальная и локальная эквализация

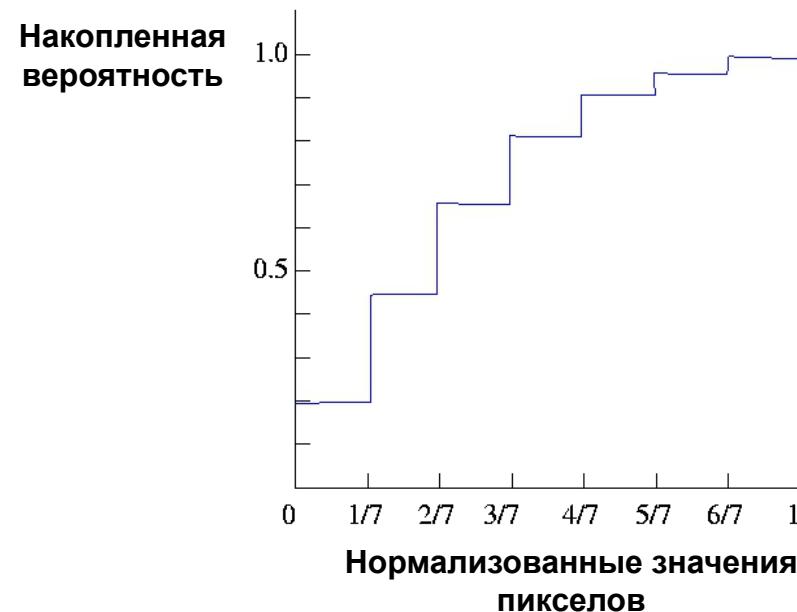
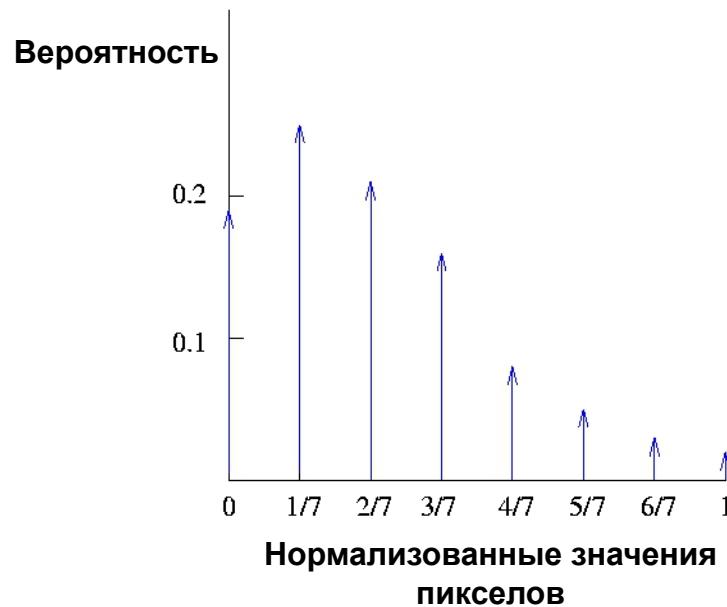


(а) Исходное изображение, (б) Результат глобальной эквализации, (с) Результат локальной эквализации в окрестности размера  $7 \times 7$  для каждого пикселя

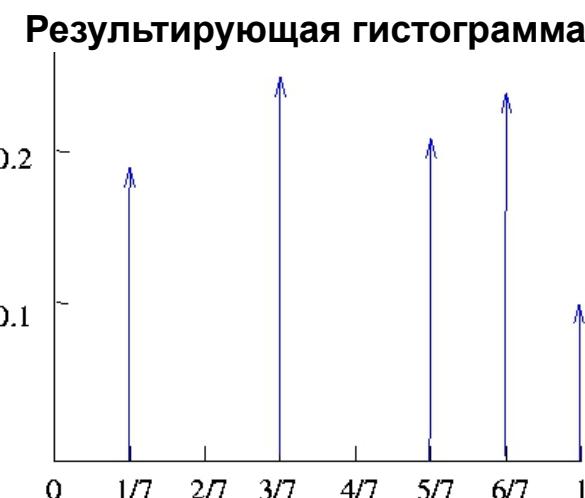
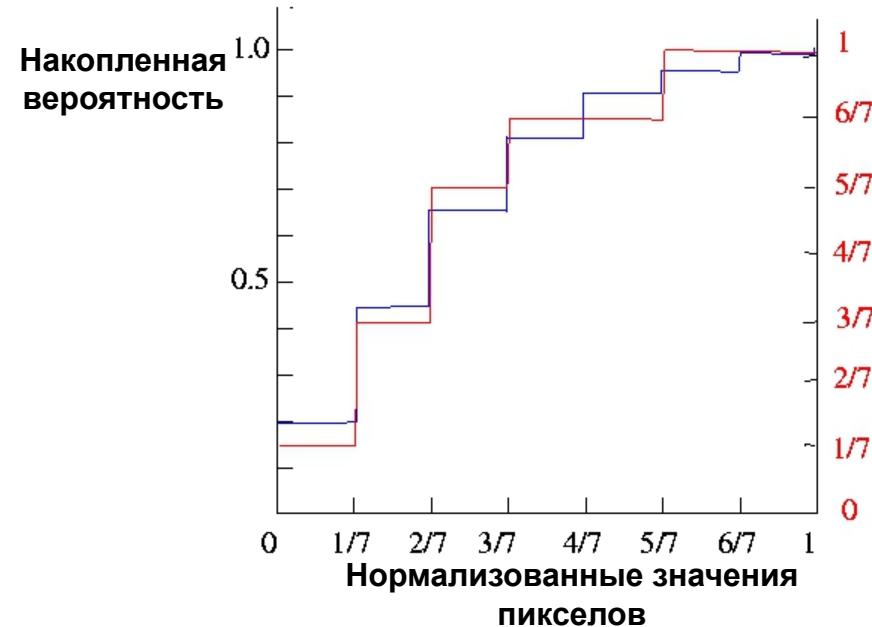
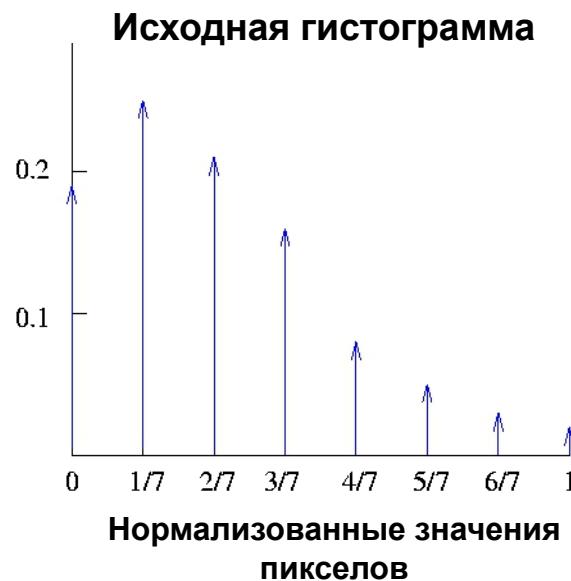
# Дискретная эквализация. Пример

Рассмотрим изображение с размерами  $64 \times 64$ , и с 8-ю градациями серого от 0 до 7 (3 бита). Общее количество пикселов  $N = 64 \times 64 = 4096$ .

Нормализован- ные значения пикселов	Количество пикселов	Вероятность	Накопленная вероятность
0	790	0.19	0.19
1/7	1023	0.25	0.44
2/7	850	0.21	0.65
3/7	636	0.16	0.81
4/7	329	0.08	0.89
5/7	245	0.06	0.95
6/7	122	0.03	0.98
1	81	0.02	1.00



# Дискретная эквализация. Пример



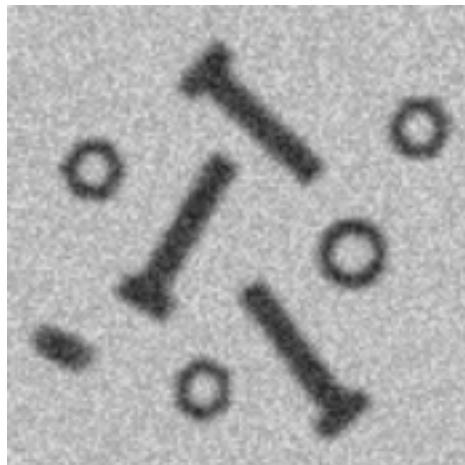
Соответствие уровней

0/7	-	1/7
1/7	-	3/7
2/7	-	5/7
3/7	-	6/7
4/7	-	6/7
5/7	-	7/7
6/7	-	7/7
7/7	-	7/7

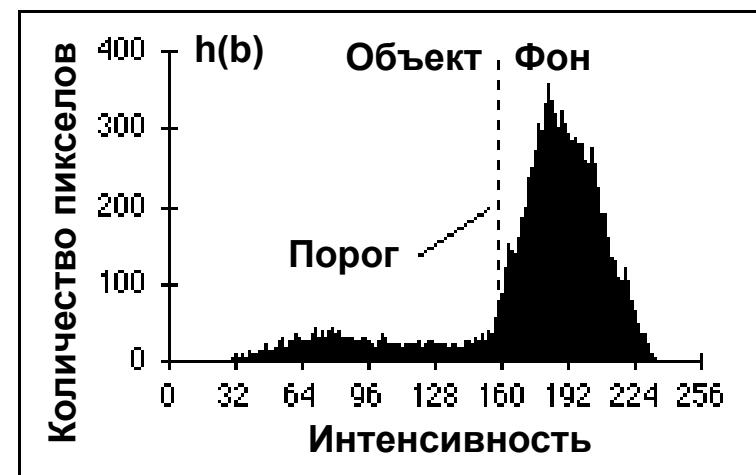
# Пороговая бинаризация

- Пиксели с интенсивностью ниже порогового значения помечаются, как принадлежащие объекту, остальные пиксели – как принадлежащие фону.
- Здесь и далее рассматриваются изображения, содержащие темные объекты на светлом фоне.

Изображение

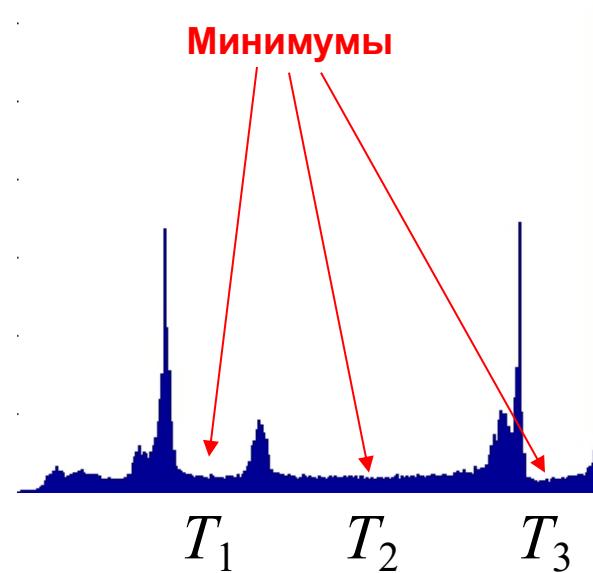
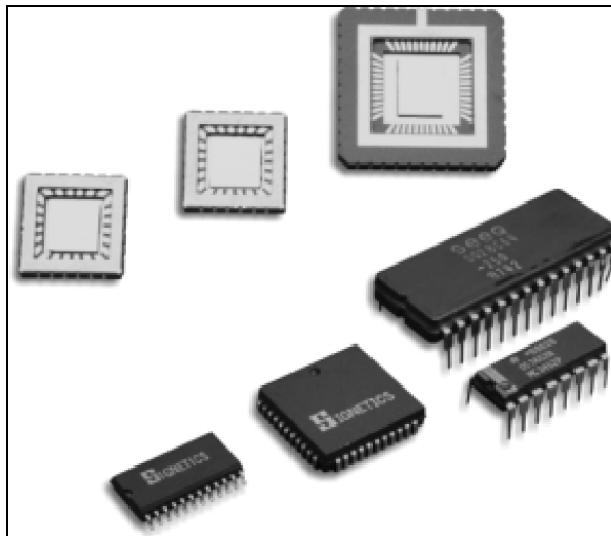


Гистограмма



# Пороговая бинаризация

Возможно обобщение на случай многопороговой сегментации

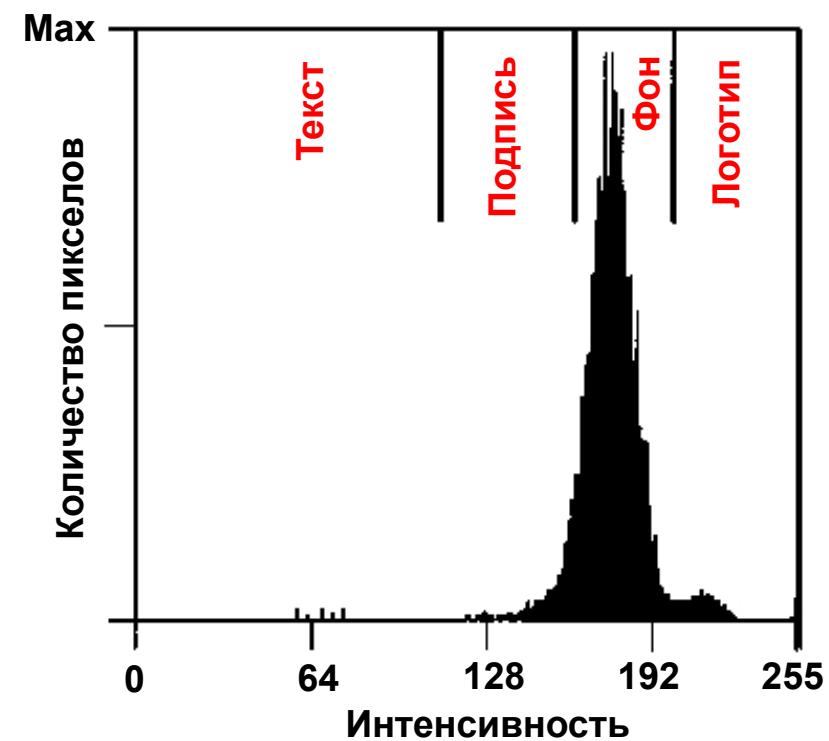
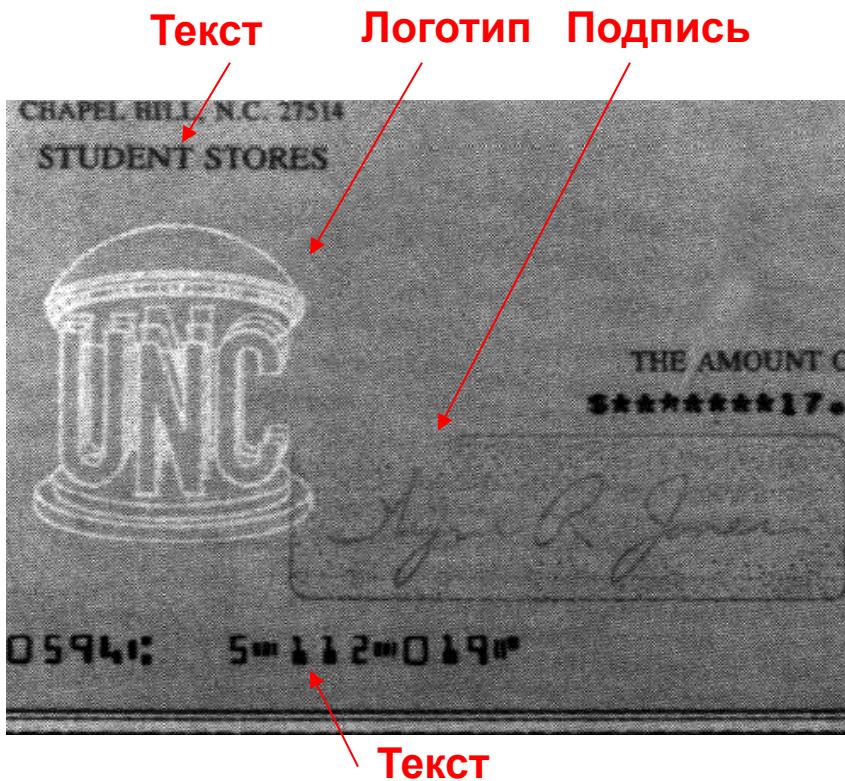


$$\begin{array}{ll} I(x,y) < T_1, & I(x,y) = 1 \text{ (объект 1)} \\ T_1 < I(x,y) < T_2, & I(x,y) = 2 \text{ (объект 2)} \\ T_2 < I(x,y), & I(x,y) = 3 \text{ (фон)} \end{array}$$

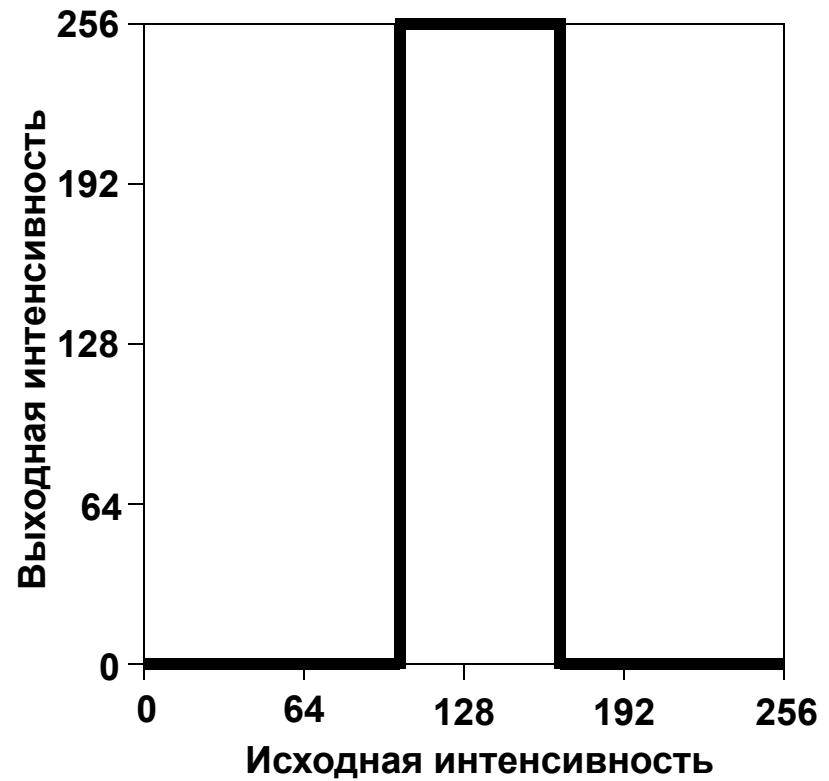
# Пороговая бинаризация

- Основная задача, решаемая при пороговой бинаризации – автоматический выбор значения порога.
- Известно несколько алгоритмов выбора значения порога по гистограмме
  - Алгоритм Изодата (Isodata)
  - Алгоритм симметрии фона
  - Алгоритм треугольника

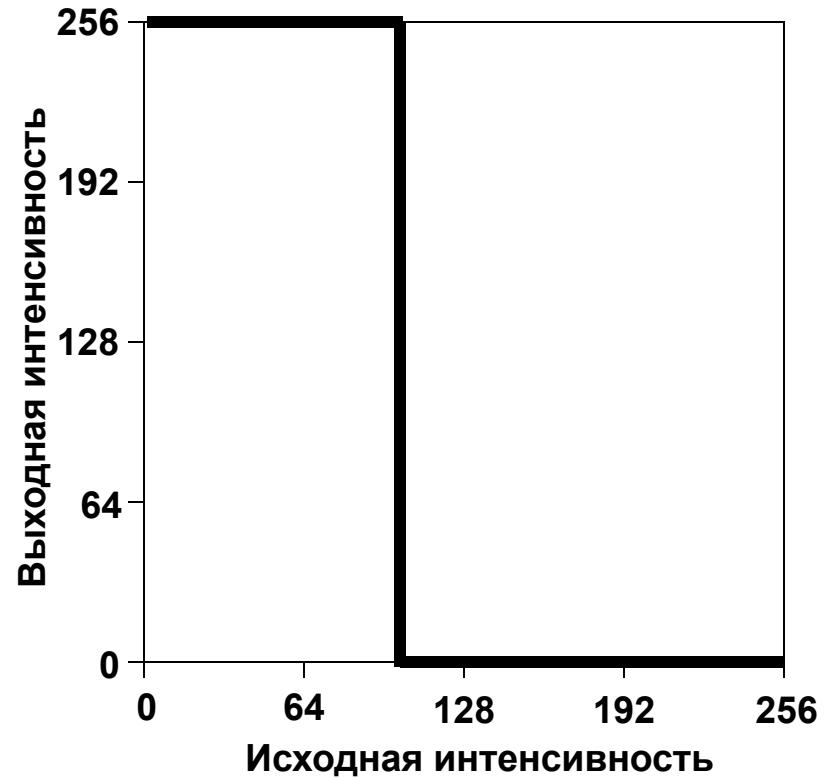
# Пороговая сегментация для многопиковых гистограмм



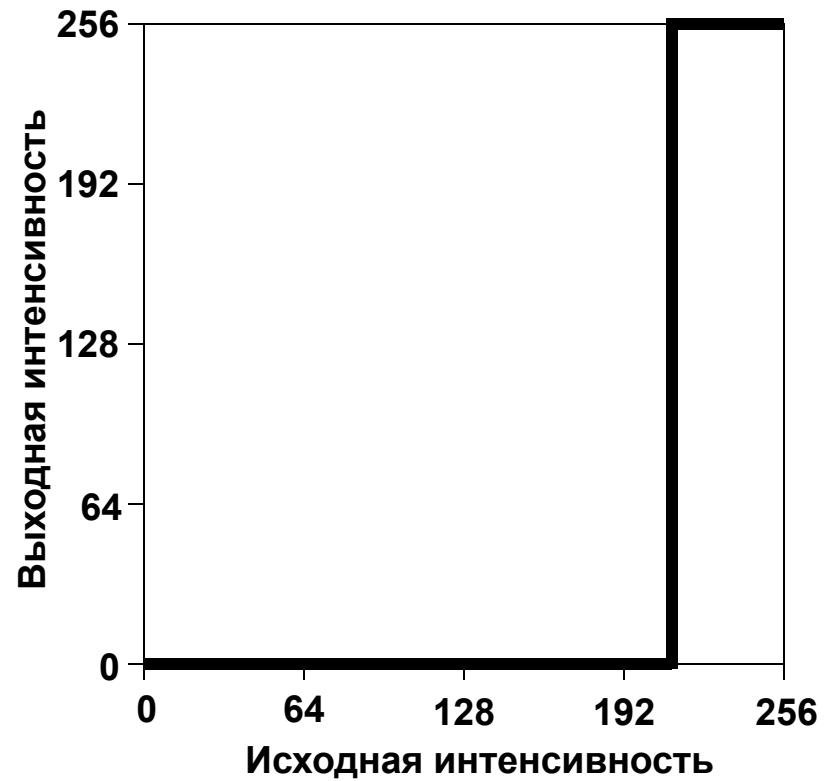
# Препарирование изображения и пороговая сегментация



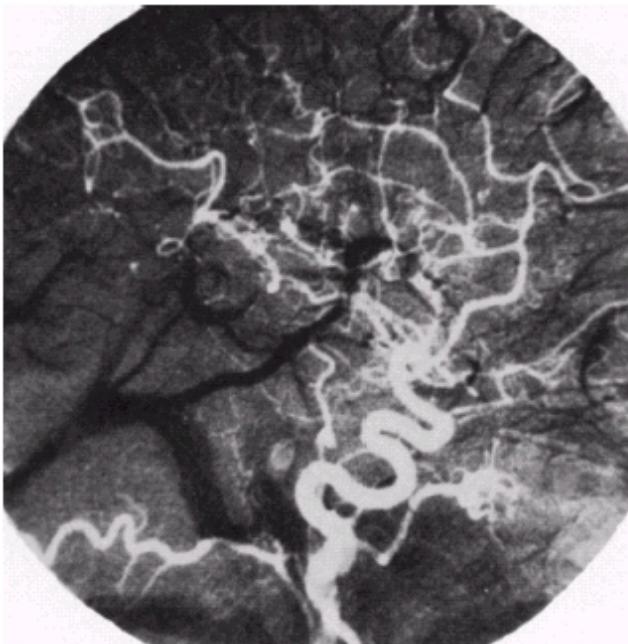
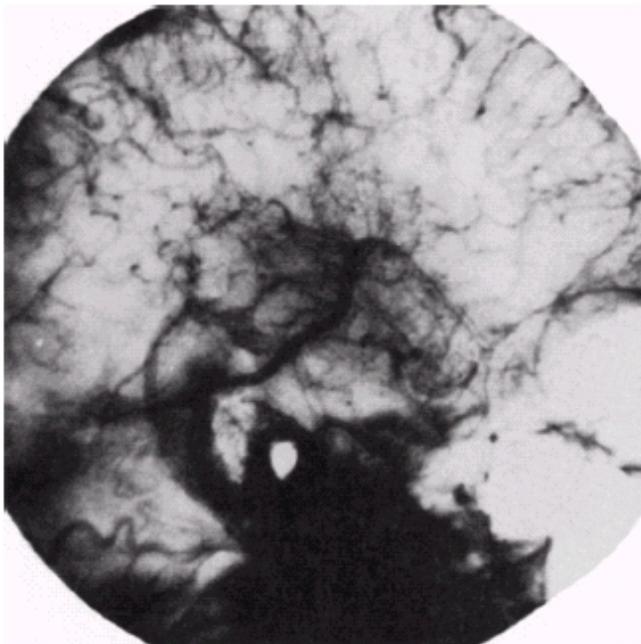
# Препарирование изображения и пороговая сегментация



# Препарирование изображения и пороговая сегментация



# Вычитание изображений

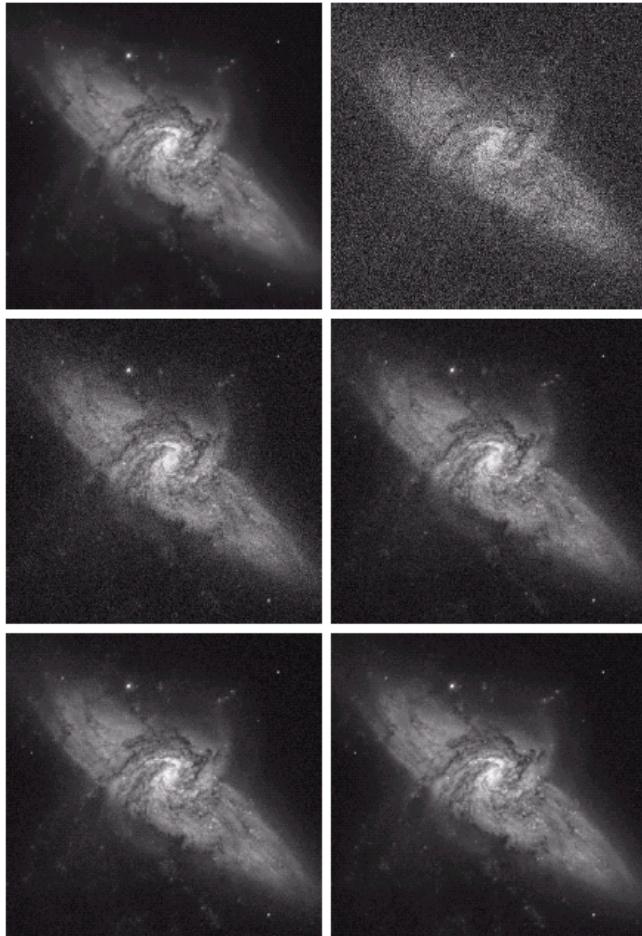


a b

**Улучшение изображений с помощью вычитания**  
(а) Изображение-маска,  
(б) Разность изображения, полученного после введения в сосуды контрастного вещества, и изображения-маски

Контрастная радиография

# Усреднение изображений



- Тот же сигнал, но разные реализации шума.
- Усреднение нескольких таких изображений улучшает отношение *сигнал/шум*.

# Табличный метод попиксельных преобразований

- Прямое выполнение попиксельных преобразований для больших изображений требует многократного вычисления функции  $s=T(r)$ , что достаточно трудоемко.
- Табличный метод состоит в предварительном построении таблицы значений  $s=T(r)$  для всех уровней  $r_k$  и ее использовании для осуществления преобразования

$r$	$r_1$	$r_2$	...	...	...	...	...	$r_{n-1}$	$r_n$
$s = T(r)$	$s_1$	$s_2$	...	...	...	...	...	$s_{n-1}$	$s_n$