Лекция 10

«Общая характеристика задач распознавания образов и их типы»

Автор: д.т.н Джомартова Ш.А.

Направления искусственного интеллекта



Область применения распознавания образов



Что такое распознавание образов (РО)?

Определения из литературы

- Отнесение физического объекта или события к одной из предопределенных категорий – *Дуда и Харт (Duda and Hart)*
- Задача оценивания функций плотности распределения в многомерном пространстве и разбиение пространства на области категорий и классов – *Фукунага (Fukunaga)*
- Автоматическое принятие решений для потока будущих образцов по некоторым примерам сложных сигналов и правильных решений для них – *Рипли (Ripley)*
- Наука, изучающая описание или классификацию (распознавание) измерений – *Шалькофф (Shalkoff)*
- Процесс присвоения имен w наблюдениям *x* – *Шурман (Shurmann)*
- Распознавание образов отвечает на вопрос «**Что это?**» -- *Морзе (Morse)*

Определение

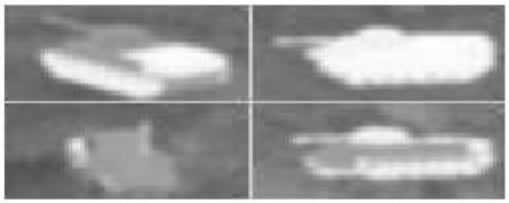
- **Распознавание образов** – это научная дисциплина, целью которой является разделение объектов по нескольким категориям или классам.

Примеры задач

Машинное зрение

- Визуальная проверка
- Опознавание целей

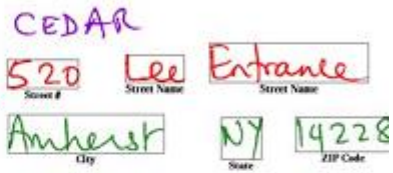




Распознавание символов

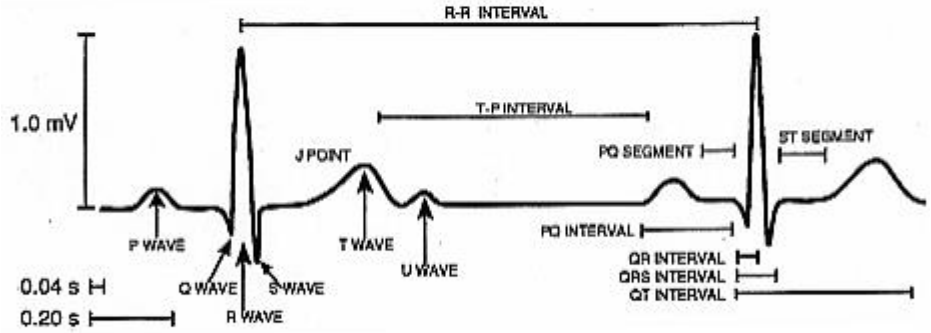
- Автоматическая сортировка почты
- Обработка налоговых деклараций





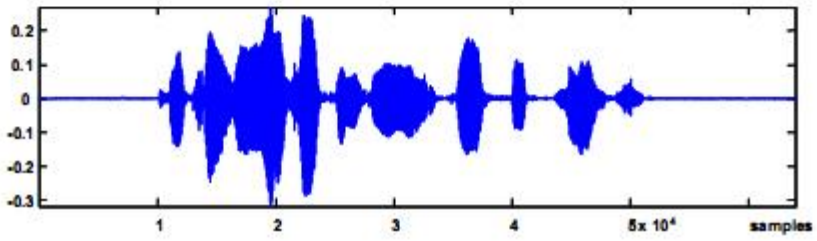
Компьютерная диагностика

- 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

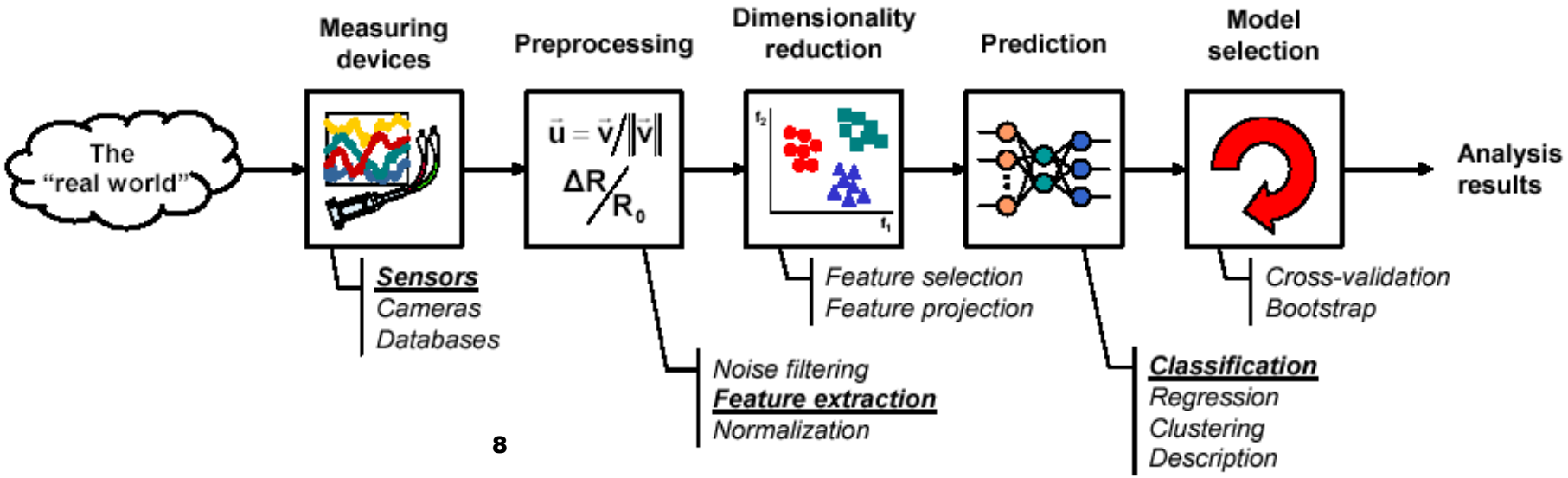
Приложения

- 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

Структура типичной системы РО

Базовая система РО содержит

- Датчик(и)
- Механизм предобработки
- Механизм извлечения признаков
- Алгоритм классификации
- Набор описанных (классифцированных) образцов (обучающую выборку)



Системы распознавания образов





Приложение AISBERG LLC





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

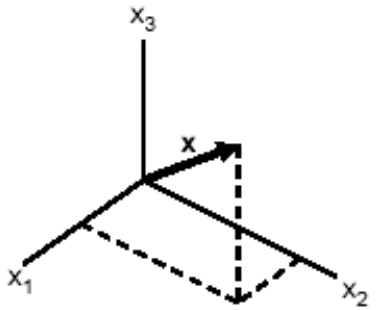
Признаки и образы

Признак

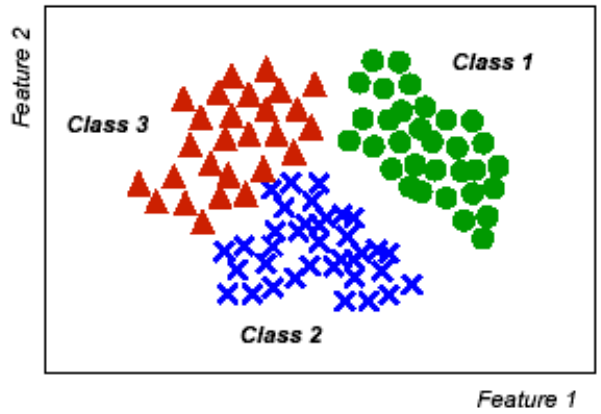
- Признак – любая отличительная характеристика, качественная или количественная
- Определения
  - Комбинация *d* характеристик, представленная в виде *d*-вектора называется **вектором характеристик**
  - *d*-мерное пространство, определяемое векторами характеристик, называется **пространством признаков**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

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

Feature vector



Feature space (3D)

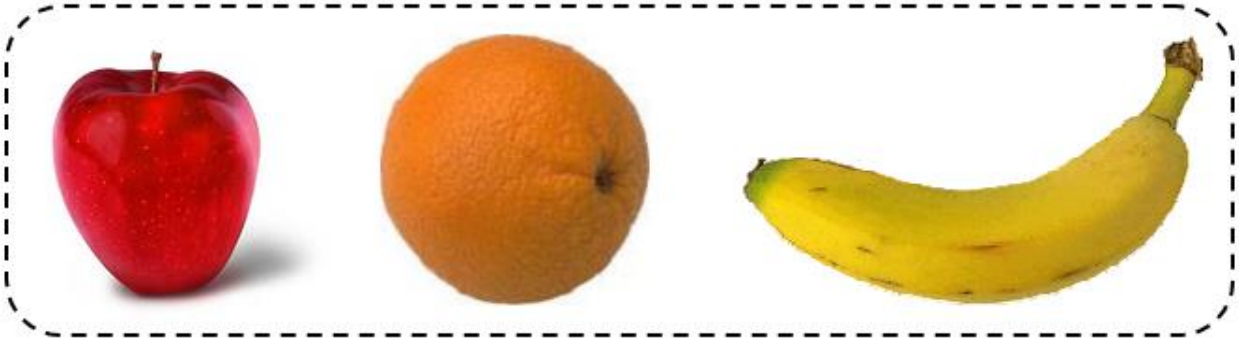


Scatter plot (2D)

Признаки и образы

Образ

- Образ есть совокупность признаков, характеризующих объект
- В задачах классификации образ есть пара переменных {*x*,w}, где
  - ***x*** есть набор наблюдений (вектор признаков)
  - w 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 признак



"Good" features



"Bad" features

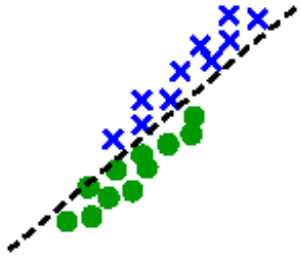
Свойства, зависящие от векторов признаков



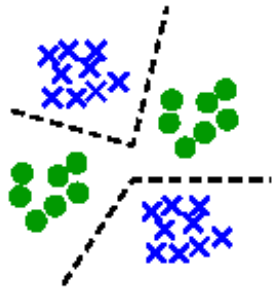
Linear separability



Non-linear separability



Highly correlated features



Multi-modal

Подходы к РО

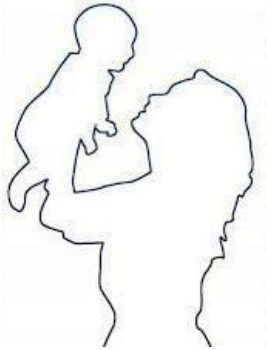
Статистический
- Классификация образов основана на статистической модели признаков

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
- …

Типы признаков объектов



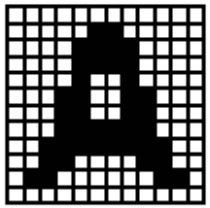
- Признаки объектов
  - Дескрипторы границ
    - Фурье дескрипторы
    - Кривизна
    - Сигнатура
    - Цепной код
    - Стат. характеристики
  - Дескрипторы областей
    - Текстура
      - Статические
        - гистограмма интенсивности
      - Структурные
        - 
      - Спектральные
        - Фурье дескрипторы
    - Цвет
      - гистограмма интенсивности
      - стат. характеристики
    - Геометрические характеристики
      - -площадь
      - -компактность
      - -число эйлера
      - -скелет

Подходы к РО. Пример

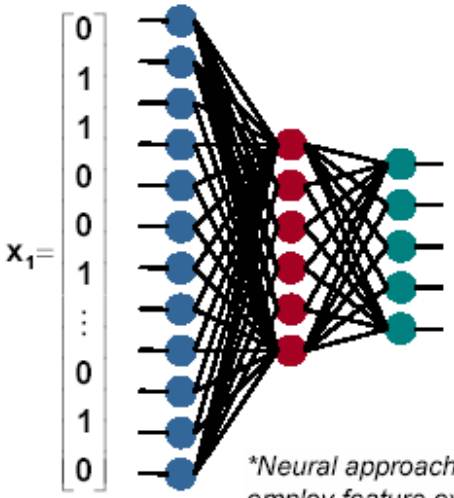


From [Schalkoff, 1992]

Neural* **Statistical** **Structural**

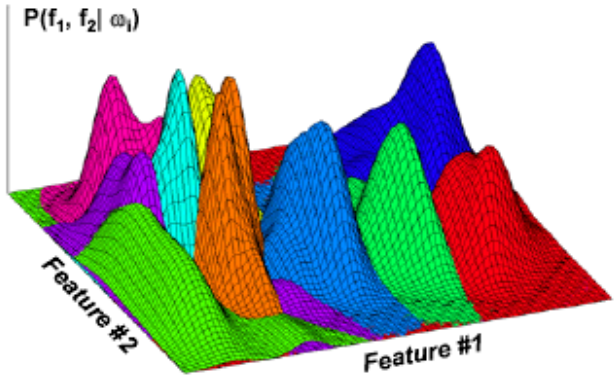


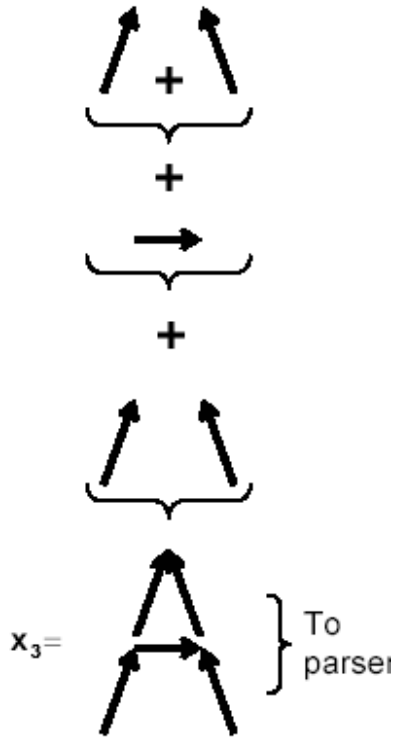
Feature extraction:
- # intersections
- # right oblique lines
- # left oblique lines
- # horizontal lines
- # "holes"



**Neural approaches may also employ feature extraction*

$$x_2 = [3 \quad 2 \quad 1 \quad 2 \quad 1]^T \xrightarrow{\text{Probabilistic model}} p(x_2 \mid \text{"A"})$$





Методы классификации

Обучение по подобию - методы, использующие приближение и расстояние

-метод ближайшего соседа

-метод опорных векторов

Статистические методы – основаны на вычислении вероятности принадлежности классу

-Байесовский классификатор

Нейронные сети

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

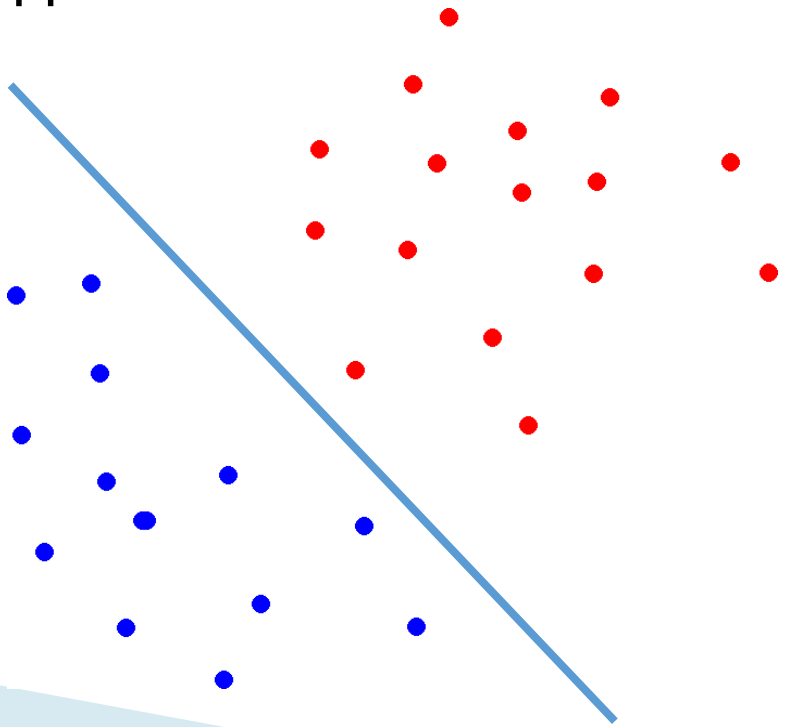
Обучение
- Адаптация модели для описания данных по обучающей выборке
- С учителем, без учителя

Оценка результатов
- Насколько хорошо данные описываются моделью? Переобучение vs. Обобщение

Обучение по подобию - ты то, на что похож

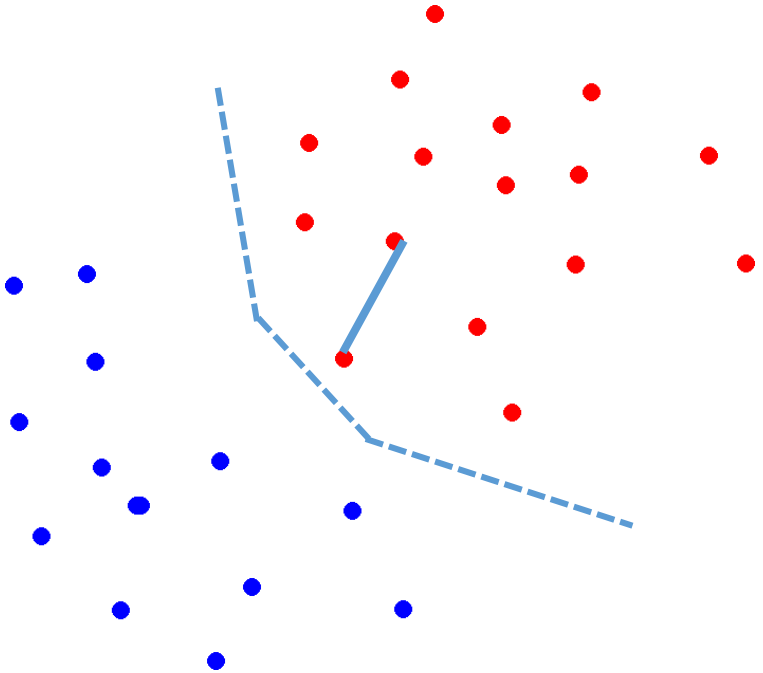
Объект = вектор признаков (точка в пространстве признаков)

Задача построить решающую функцию, которая по заданному вектору давала ответ кому принадлежит объект

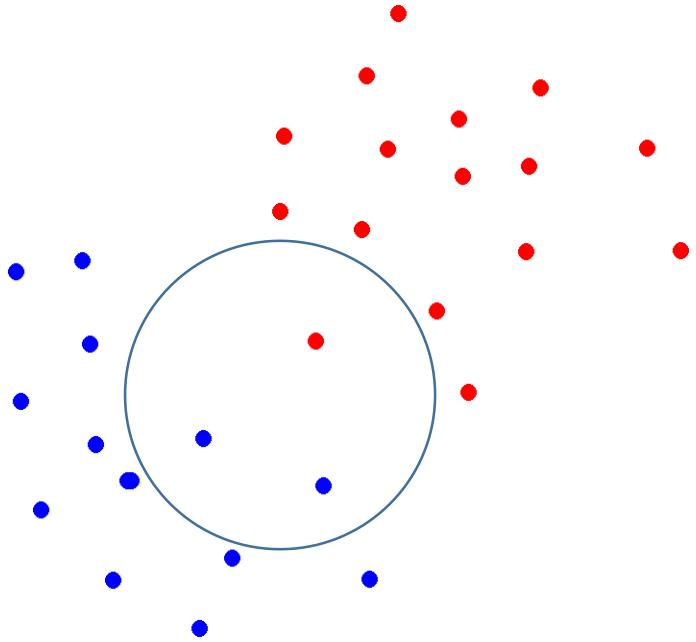


Обучение по подобию: методы ближайших соседей

Метод ближайших соседей — объект относится к тому классу, которому принадлежат ближайшие к нему объекты обучающей выборки.

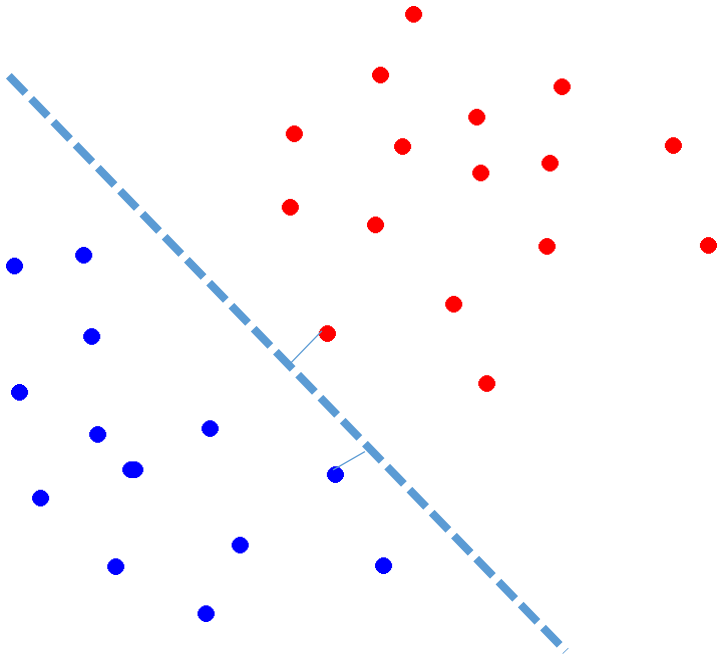


Метод k-ближайших соседей - классифицируется путем нахождения k- ближайших соседей, которые после этого голосуют.



Обучение по подобию: метод опорных векторов

Метод опорных векторов – выбирается плоскость, которая максимизирует зазор м/ду наборами данных.



Распознавание лиц нейронными сетями

Labeled face in the wild
– 13 000 фотографий



Точность распознавания



99.2%
Обрезанные изображения- 97.53%

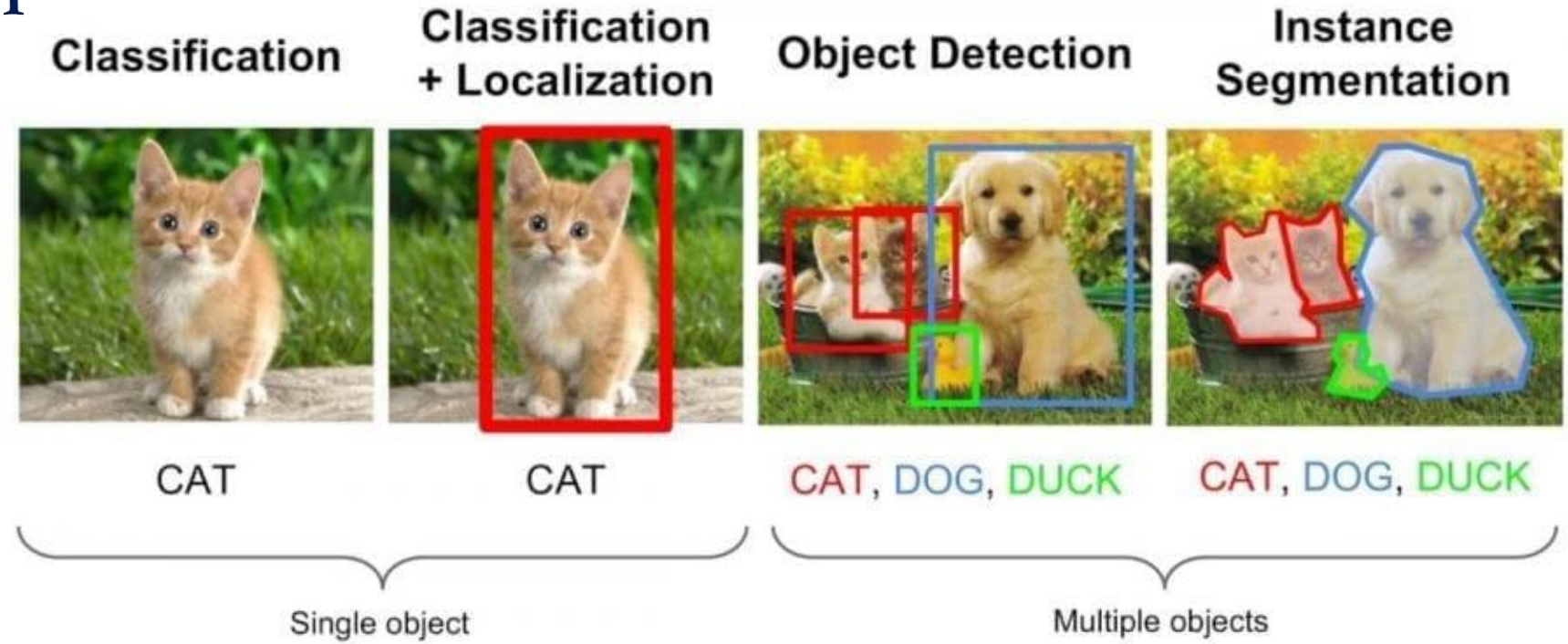


Deepface - 97.35 (Taigman et al., 2014)



99.15 (Sun Y. et al., 2014)

Пример обучения с учителем — классификация (слева), и дальнейшее ее использование для сегментации и распознавания объектов

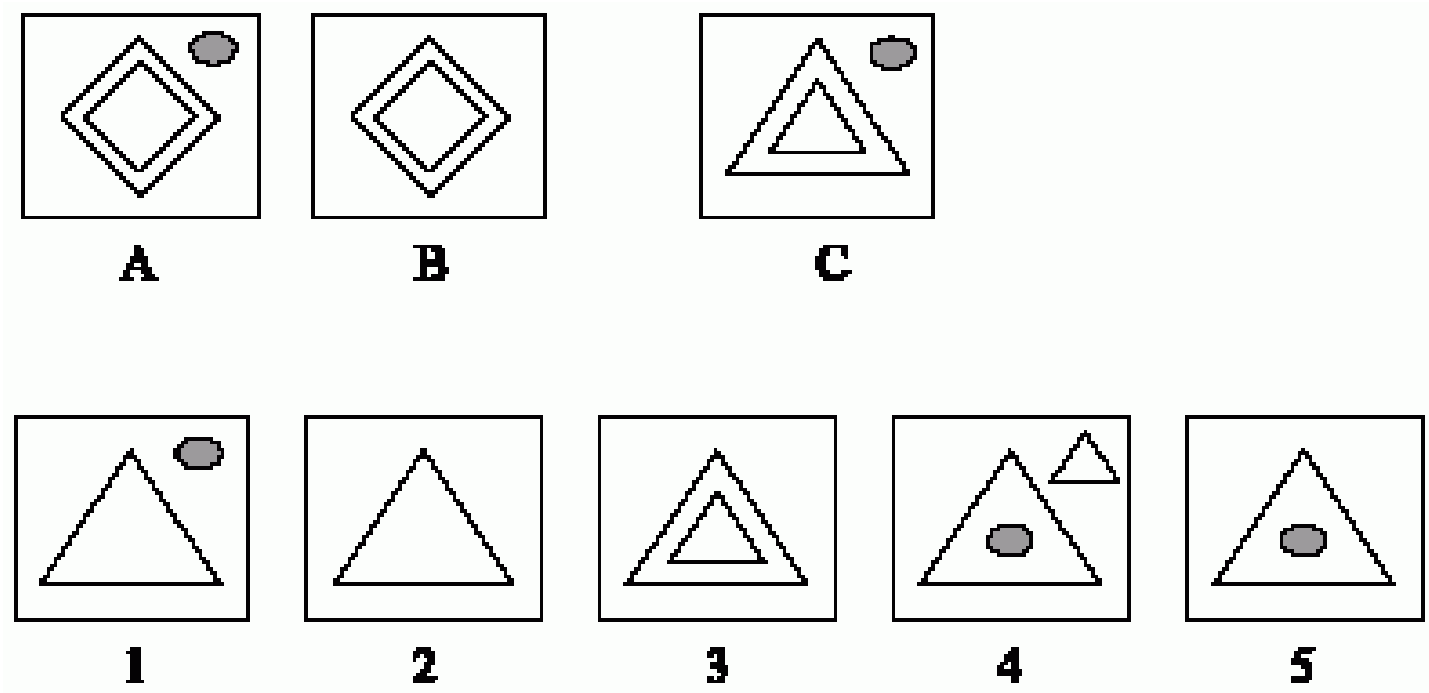


Обучение без учителя

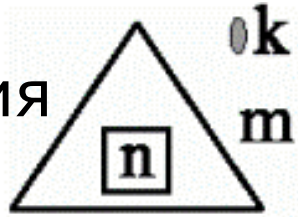
Кластеризация данных на основе общих признаков



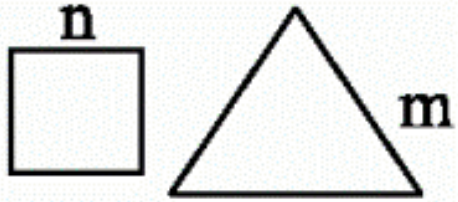
Распознавание по методу аналогий



Правила преобразования



Исходное изображение



Результирующее изображение

Актуальные задачи распознавания

- Оптическое распознавание символов
- Распознавание штрих-кодов
- Распознавание автомобильных номеров
- Распознавание изображений
- Распознавание локальных участков земной коры, в которых находятся месторождения полезных ископаемых
- Классификация документов

Особенности распознавания объектов человеком

∞

Большое количество примеров одного класса (богатый опыт распознавания)



Самолет? Нет, вид ч/з кружку

Контекст

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



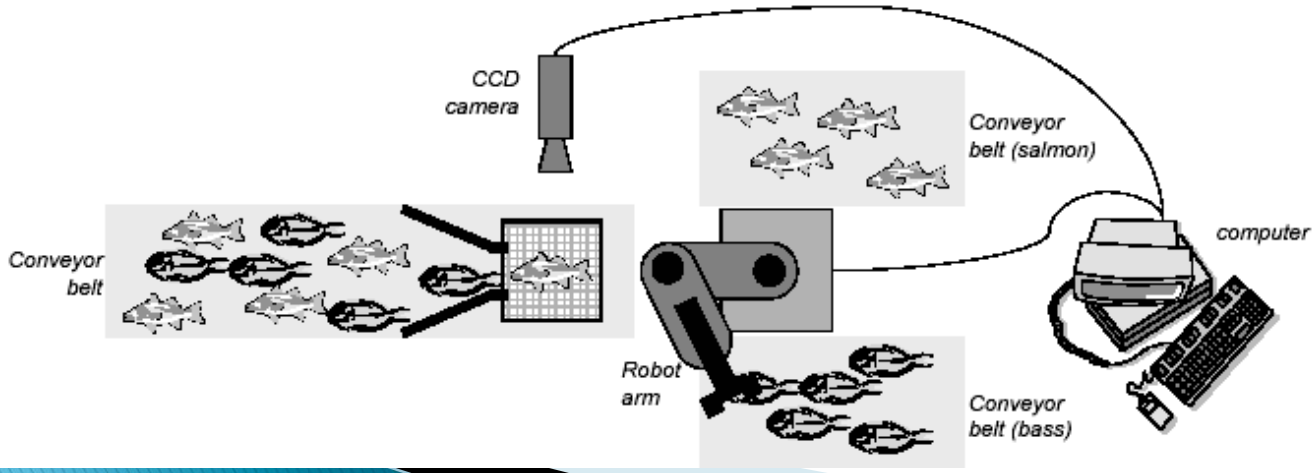
Лосось



Цикл построения системы РО

Ситуация

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



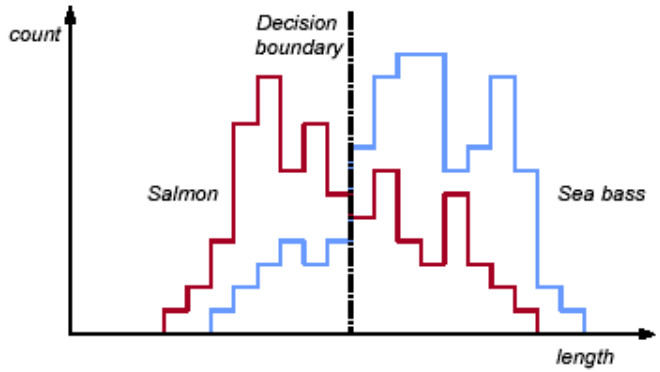
Цикл построения системы РО

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

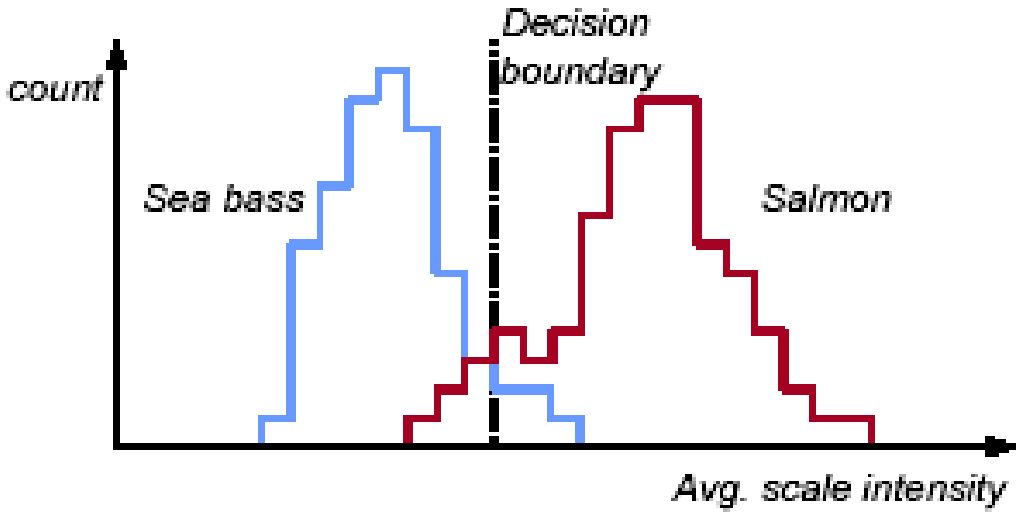
Классификация
- Собрать выборку для двух видов
- Вычислить распределение длин для обоих классов
- Определить разделяющую границу (порог), минимизирующий ошибку классификации
- После оценки вероятности ошибки получим никудышный результат – 4(
- Что дальше?



Цикл построения системы РО

Повышение эффективности системы РО

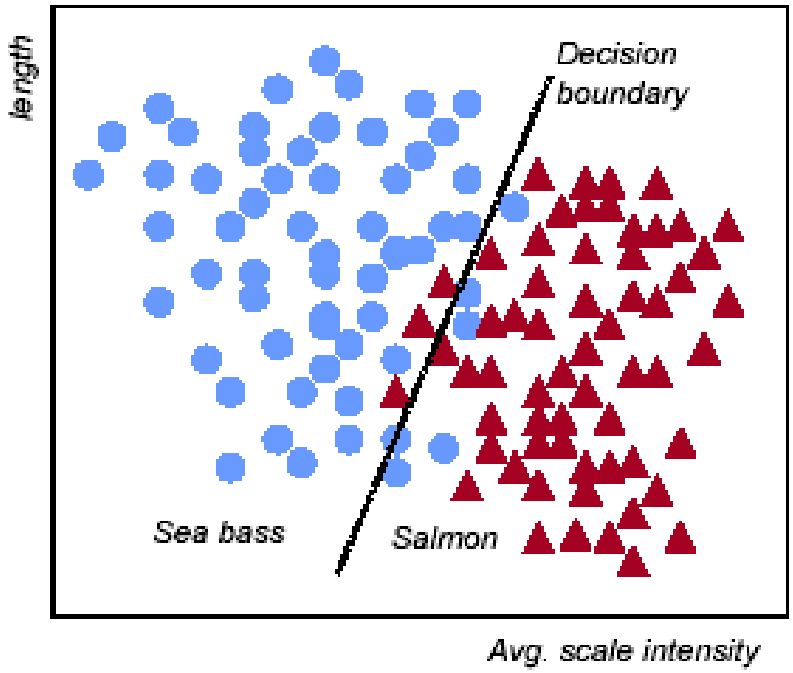
- Для достижения правильной классификации в 95% случаев, перебираем признаки
  - Ширина, Площадь, Положение глаз по отношению ко рту и т.п…
  - …только для того, чтобы выяснить, что ни один из них не обеспечивает разделения
- В конце концов найден «хороший» признак – средний размер чешуйки



Цикл построения системы РО

Повышение эффективности системы РО

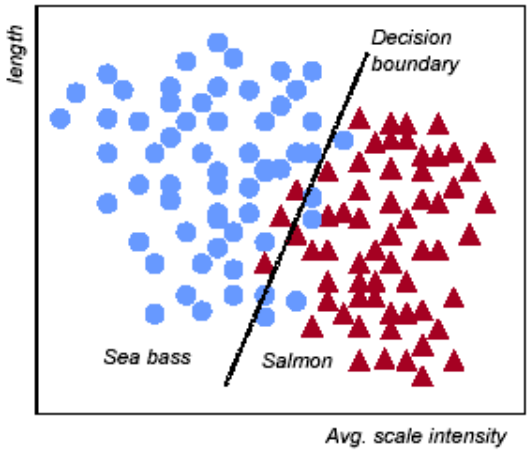
- Комбинируя *длину* рыбы со *средним размером чешуйки*, улучшаем разделяющую способность классификатора
- Построив линейную разделяющую функцию, получим точность классификации 95.7%

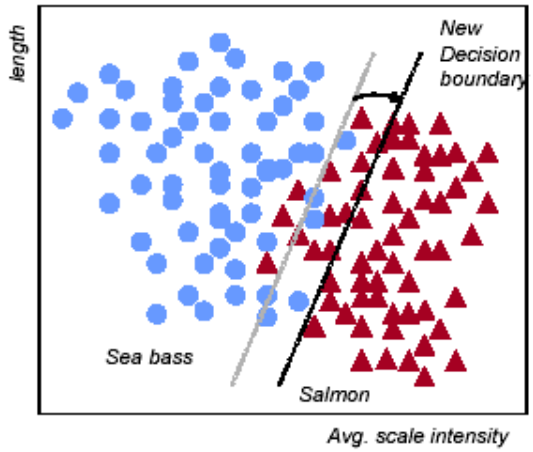


Цикл построения системы РО

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 лосося в качестве окуня -- иногда конечный покупатель будет обнаруживать, что куплен вкусный лосось по цене дешевого окуня
  - Плата за ошибочное отнесения окуня к классу лосося – покупатель будет очень огорчен, когда поймет что за цену лосося купил окуня
- Интуитивно, мы должны минимизировать именно цену ошибки выбирая разделяющую функцию

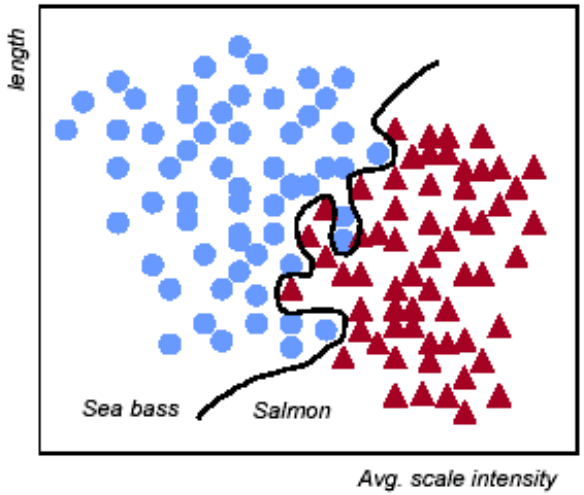




Цикл построения системы РО

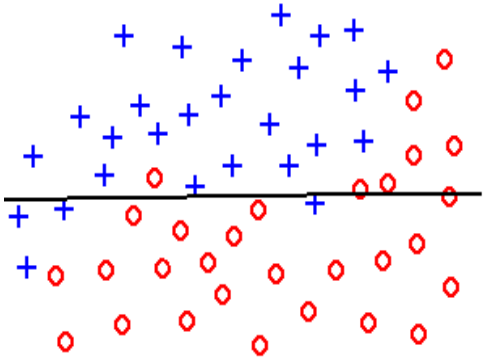
Обобщающая способность

- 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
  - Через несколько дней менеджер предприятия жалуется. что система ошибочно классифицирует в среднем 25% рыбы.
  - Что случилось?

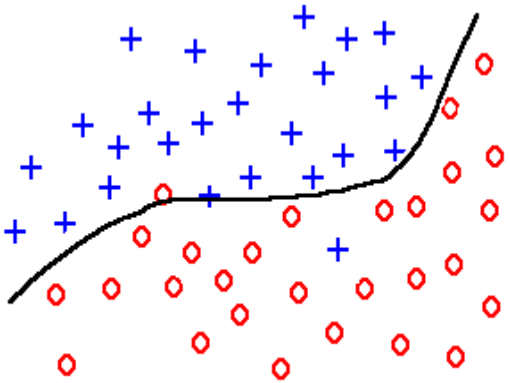


Недо- и переобучение

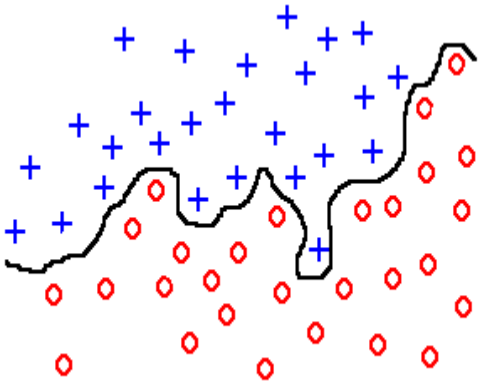
- **Задача:** насколько гибок должен быть классификатор?



излишнее обобщение, недообучение, underfitting



хорошая модель good fit



излишняя избирательность, переобучение, overfitting

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

Обработка и анализ изображений

Зачем обрабатывать изображения?

Для эффективного хранения и передачи
- Компактное представление изображения в камере
- Передача изображения с Марса на Землю

Для подготовки к показу или печати
- Приведение изображения к нужному размеру
- Полутоновое представление

Для улучшения и восстановления изображений
- Удаление царапин на кадрах старых фильмов
- Повышение видимости новообразований на медицинских снимках

Для получения информации с изображений
- Прочитать почтовый индекс на конверте
- Измерить уровень влажности поверхности по аэроснимку

Обработка изображений

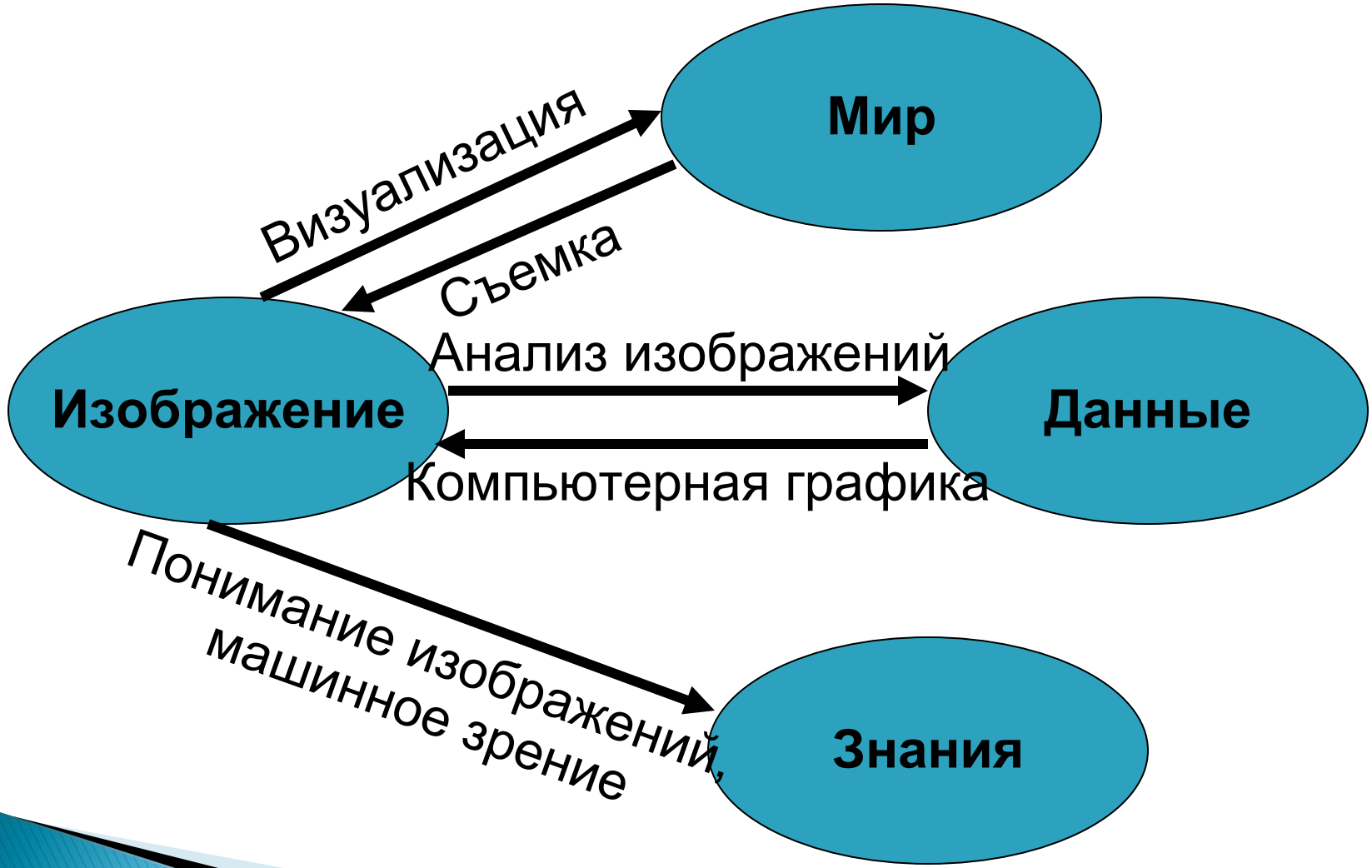
Улучшение изображений (image enhancement)
- Улучшение визуального качества изображений для восприятия человеком (сглаживание, контрастирование, эквализация гистограмм и т.п.)

Восстановление изображений
(image restoration/reconstruction)
- Устранение смазов, размытия и т.п.
- Воссоздание изображения – результата косвенных измерений (интерферометрия, радары, томография)

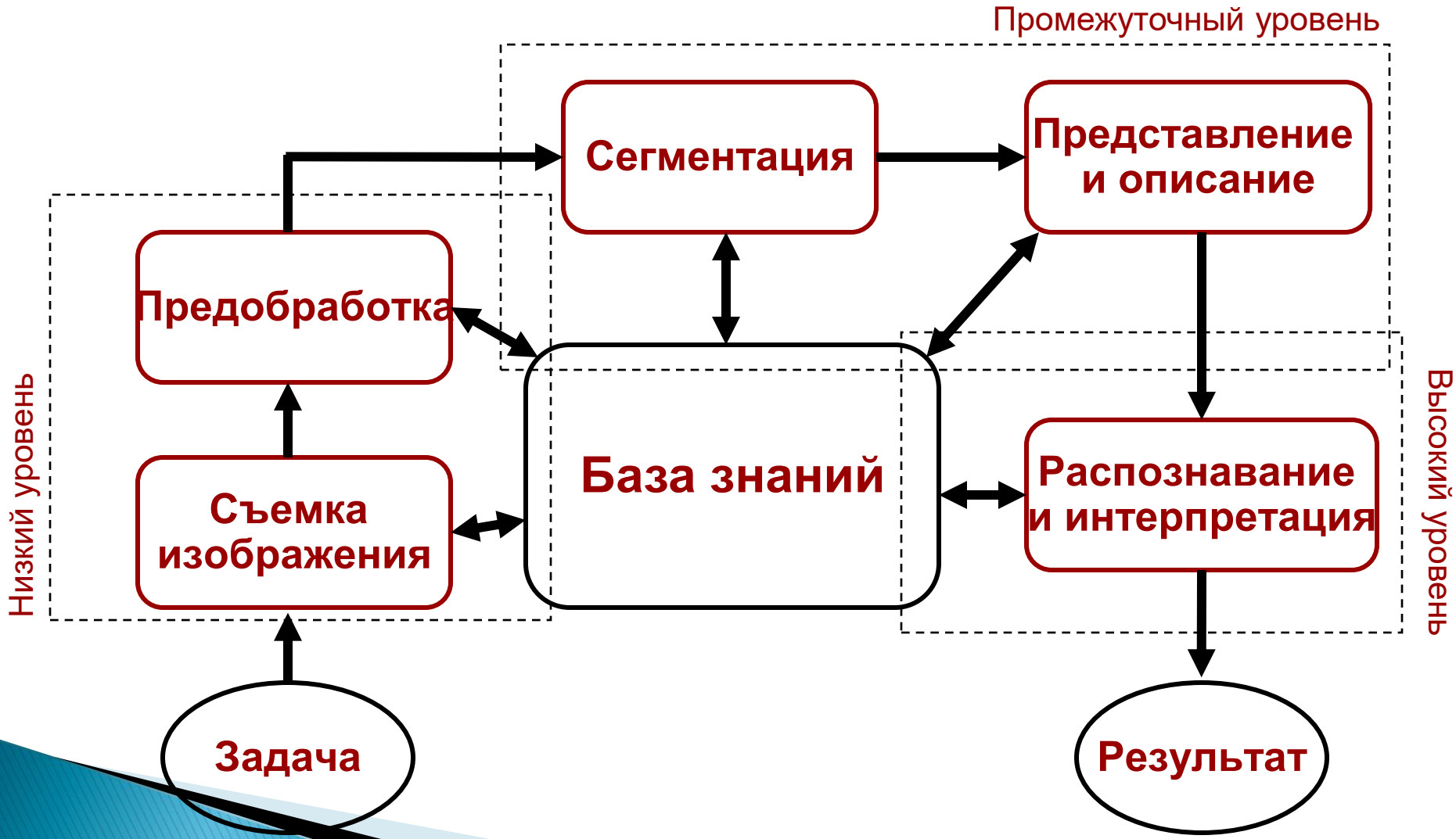
Анализ изображений (image analysis)
- Автоматическое извлечение информации из изображений (машинное зрение, медицинские приложения)

Сжатие изображений (image compression)
- Эффективное хранение и передача изображений с потерями или без в зависимости от приложения

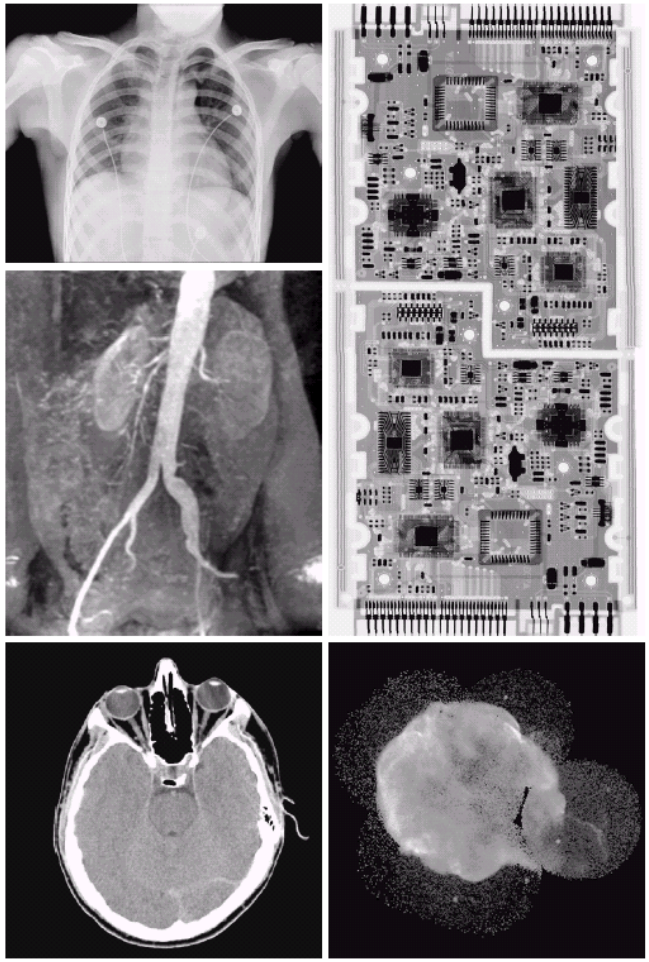
Анализ изображений



Анализ изображений

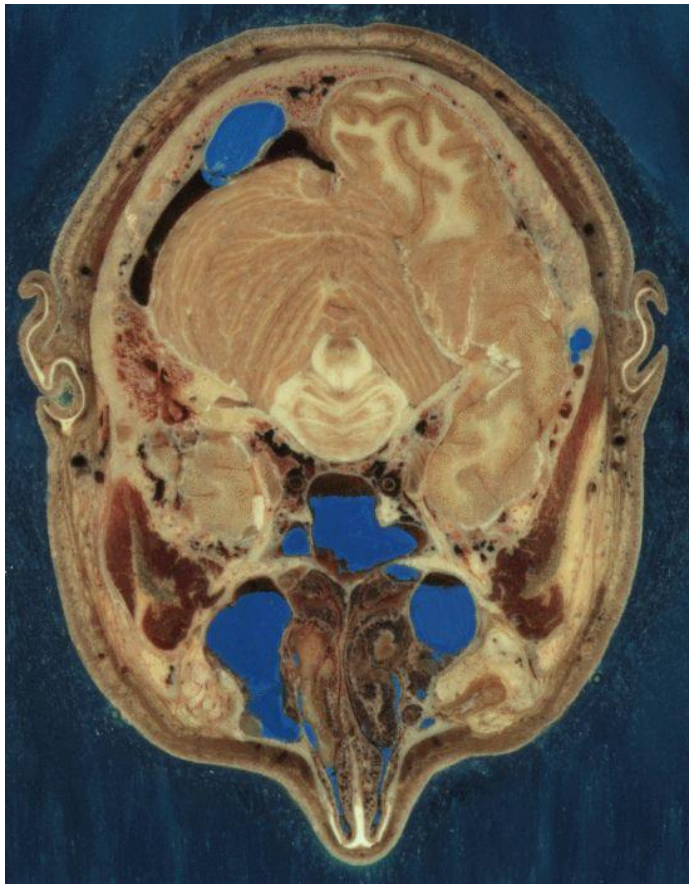


Рентгеновские изображения



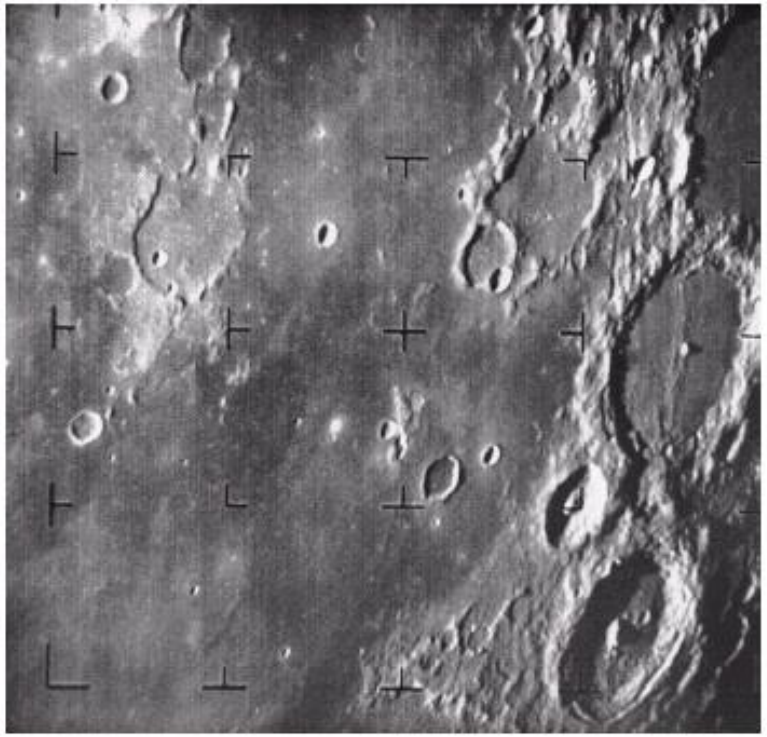
- Снимок грудной клетки человека
- Ангиограмма аорт
- Томограмма головы
- Снимок печатных плат

Компьютерная томография



- На изображении: сечение головы человека
- Изображение из проекта “Visible Human Project”

Астрономические изображения



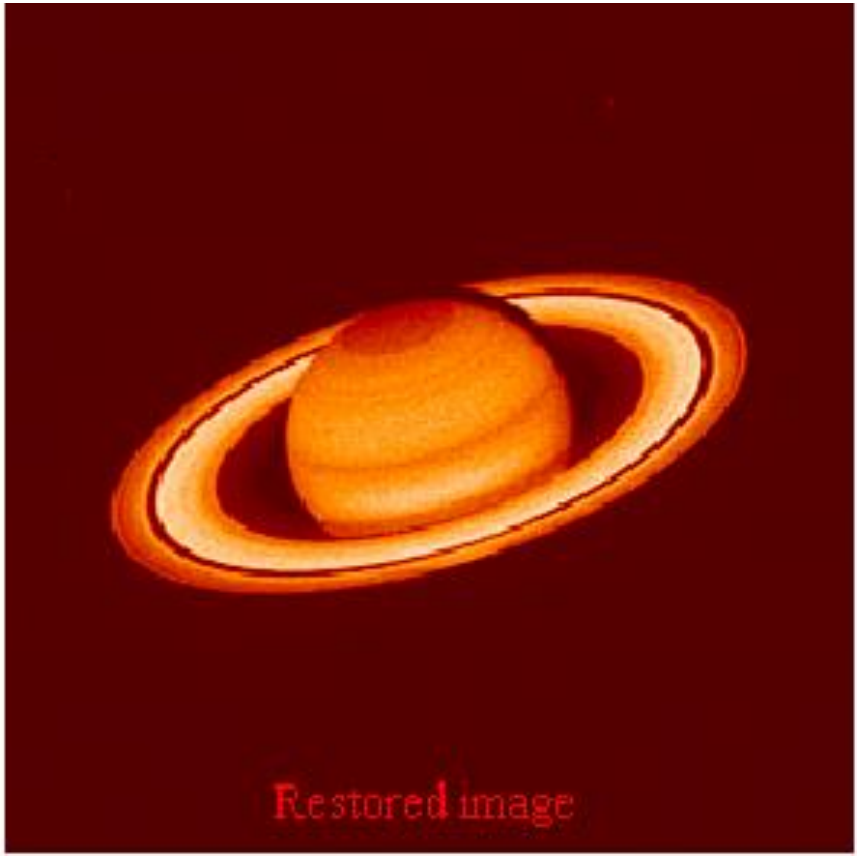
- Поверхность Луны

- Изображение, сделанное американским космическим аппаратом *Ranger 7* за 17 минут до прилунения, 31.07.1964

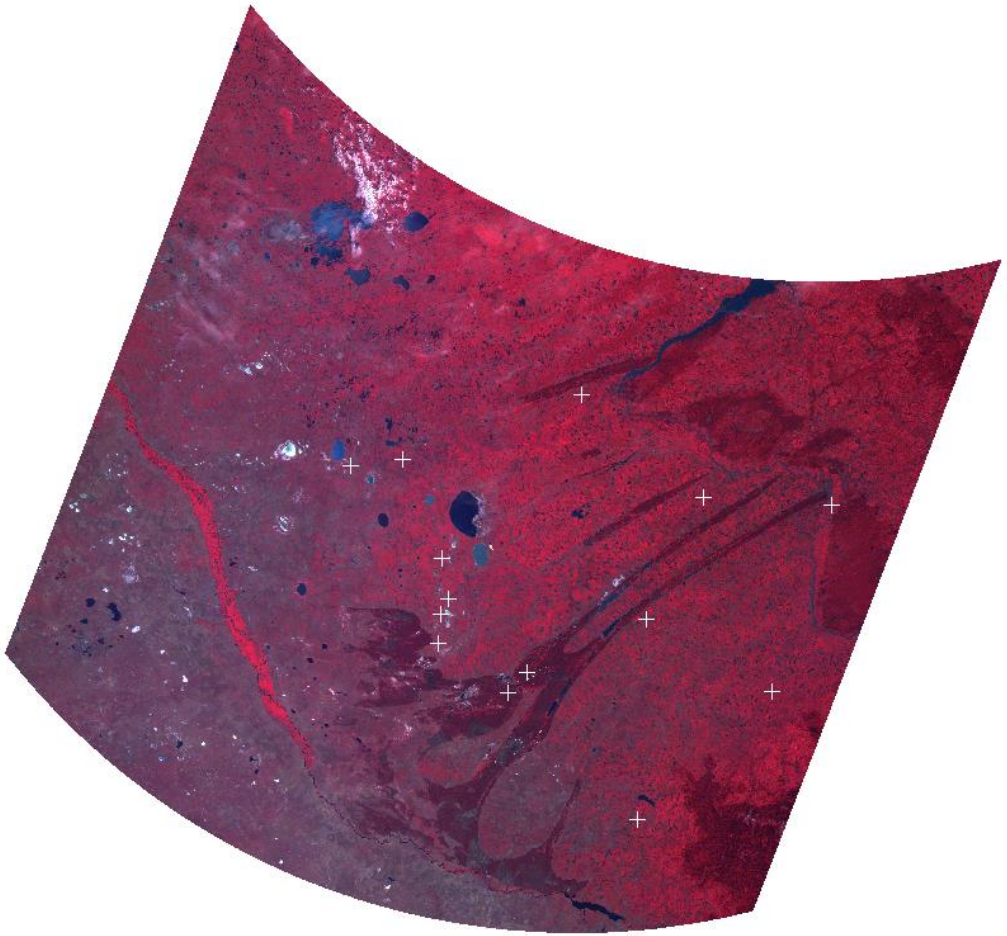
Астрономические изображения

Восстановление изображения, полученного орбитальным телескопом Hubble





Анализ данных дистанционного зондирования



- Изображение территории Алтайского края, сделанное советским орбитальным сканером МСУ-СК, 04.07.1999

- Многоканальное изображение в условных цветах

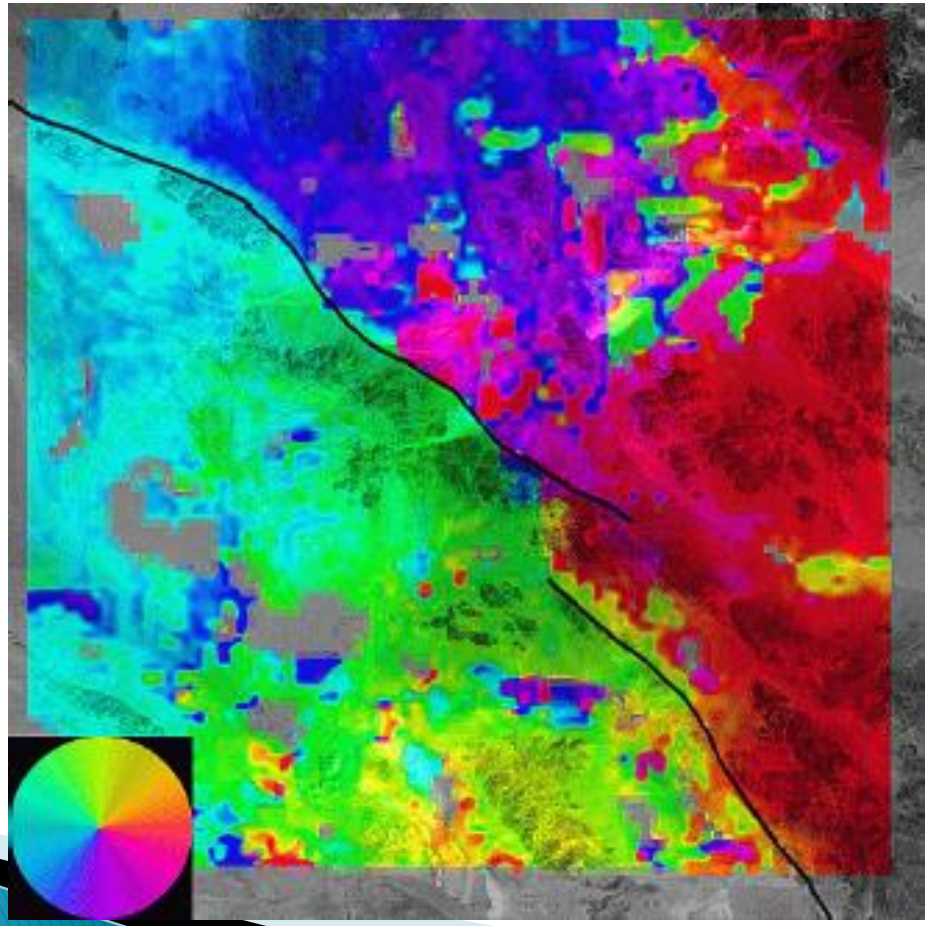
Анализ данных дистанционного зондирования

- Выделение территории поселения на аэроснимке



Анализ данных дистанционного зондирования

Анализ последствий землетрясения по спутниковым изображениям



Реставрация изображений

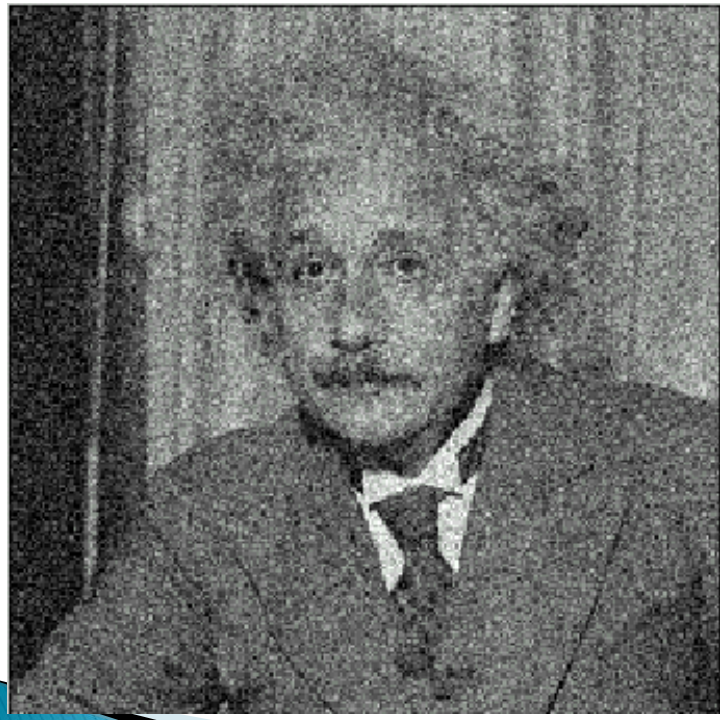
Реставрация фотографий

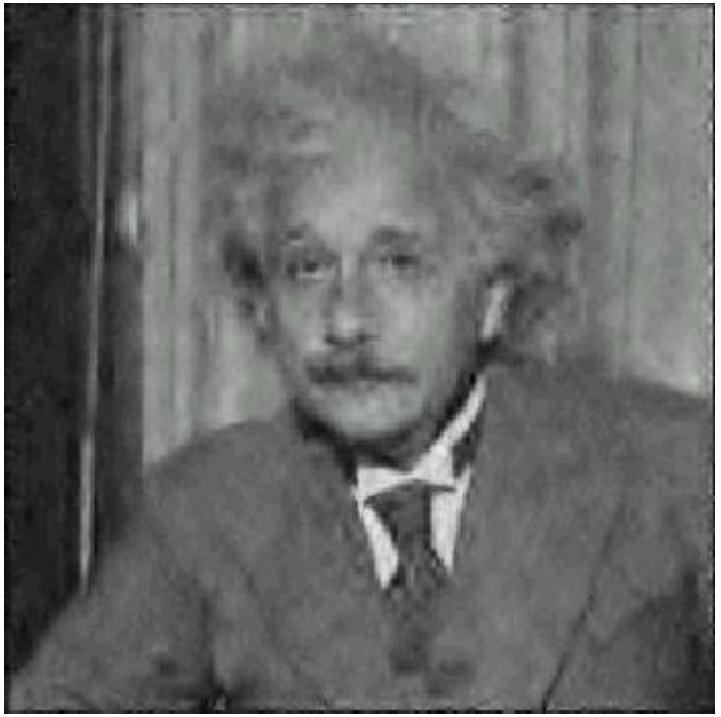




Реставрация изображений

Подавление шумов





Выделение лиц на изображении

Выделение лиц на портретных фото





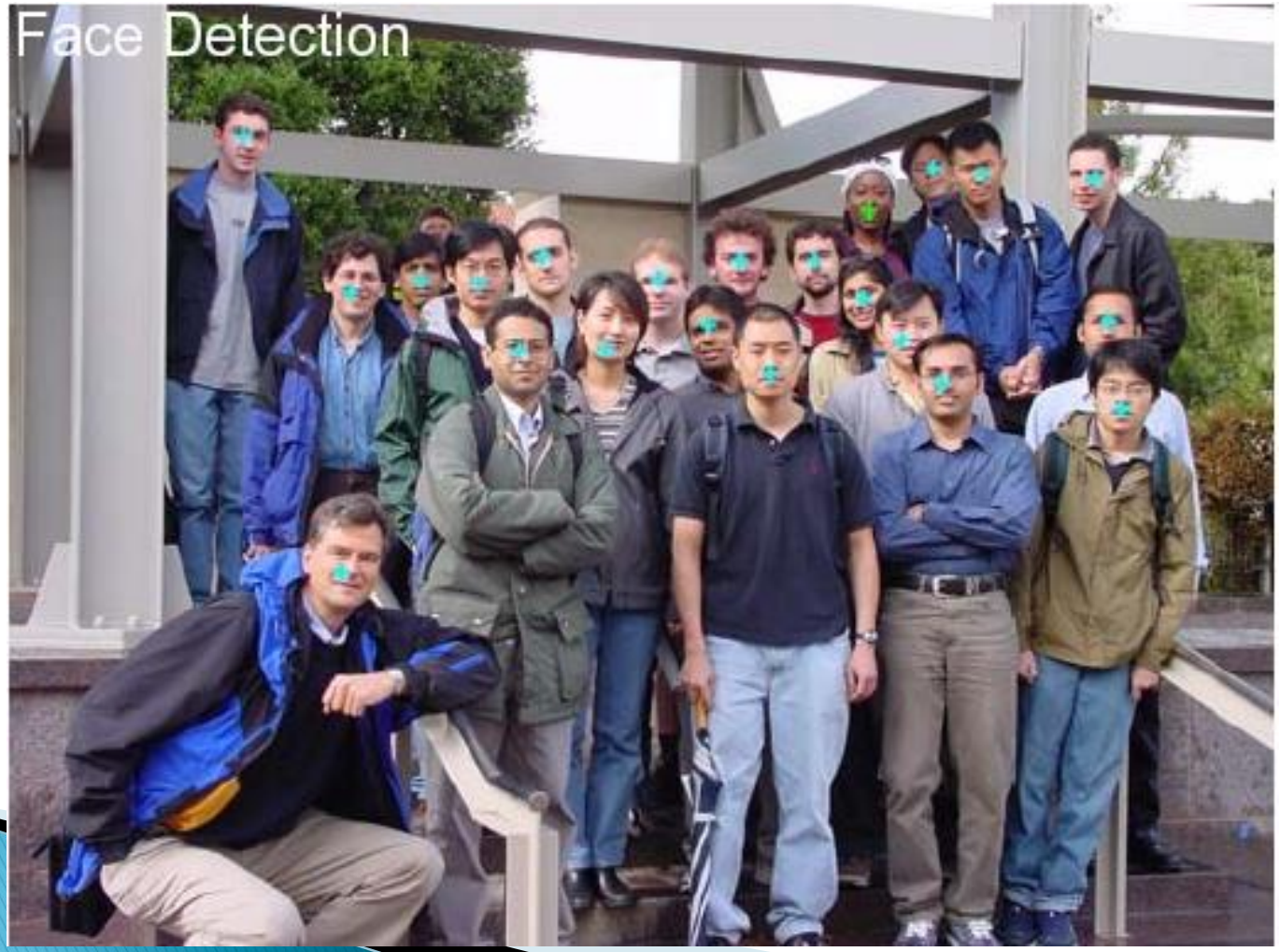






Выделение лиц на изображении

Выделение лиц на групповых фото



Выделение лиц на изображении

Отслеживание лиц



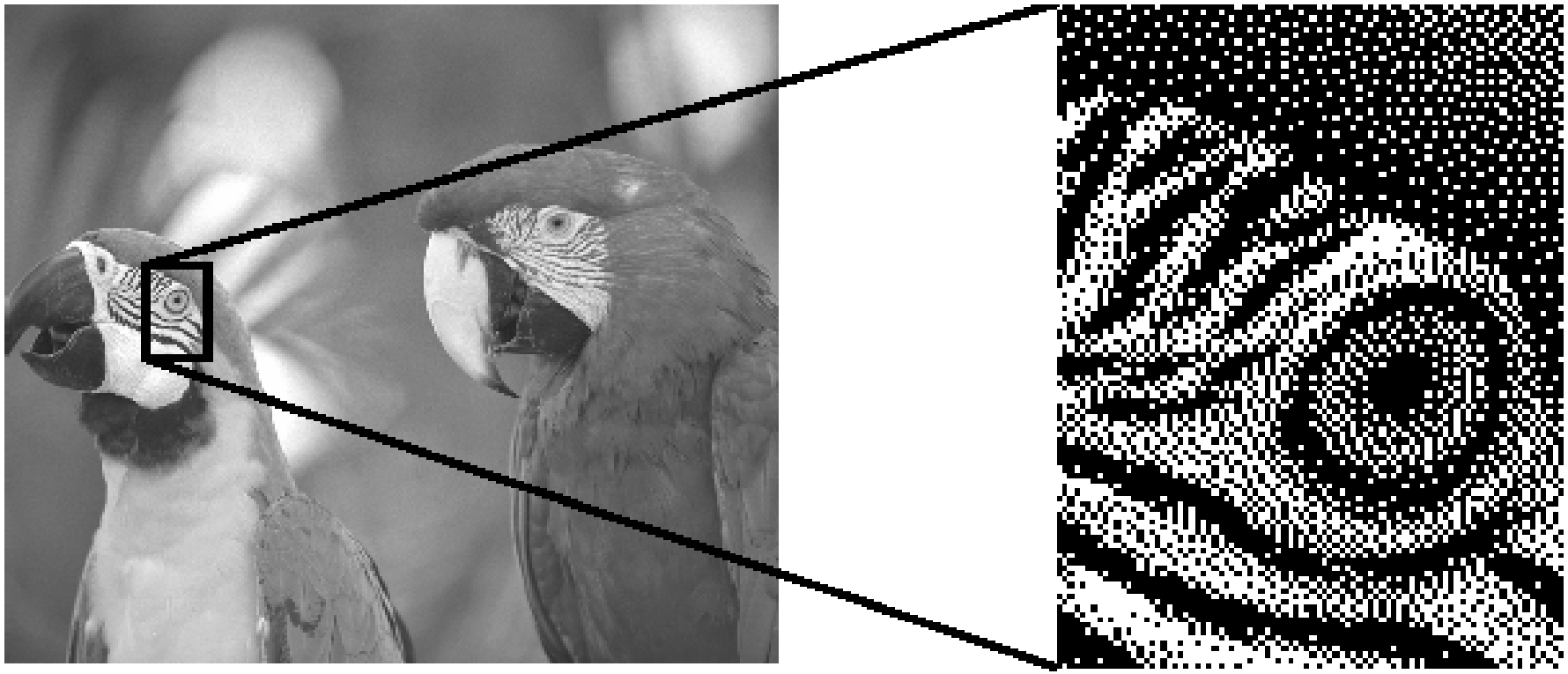
Выделение лиц на изображении

Поиск лиц в БД



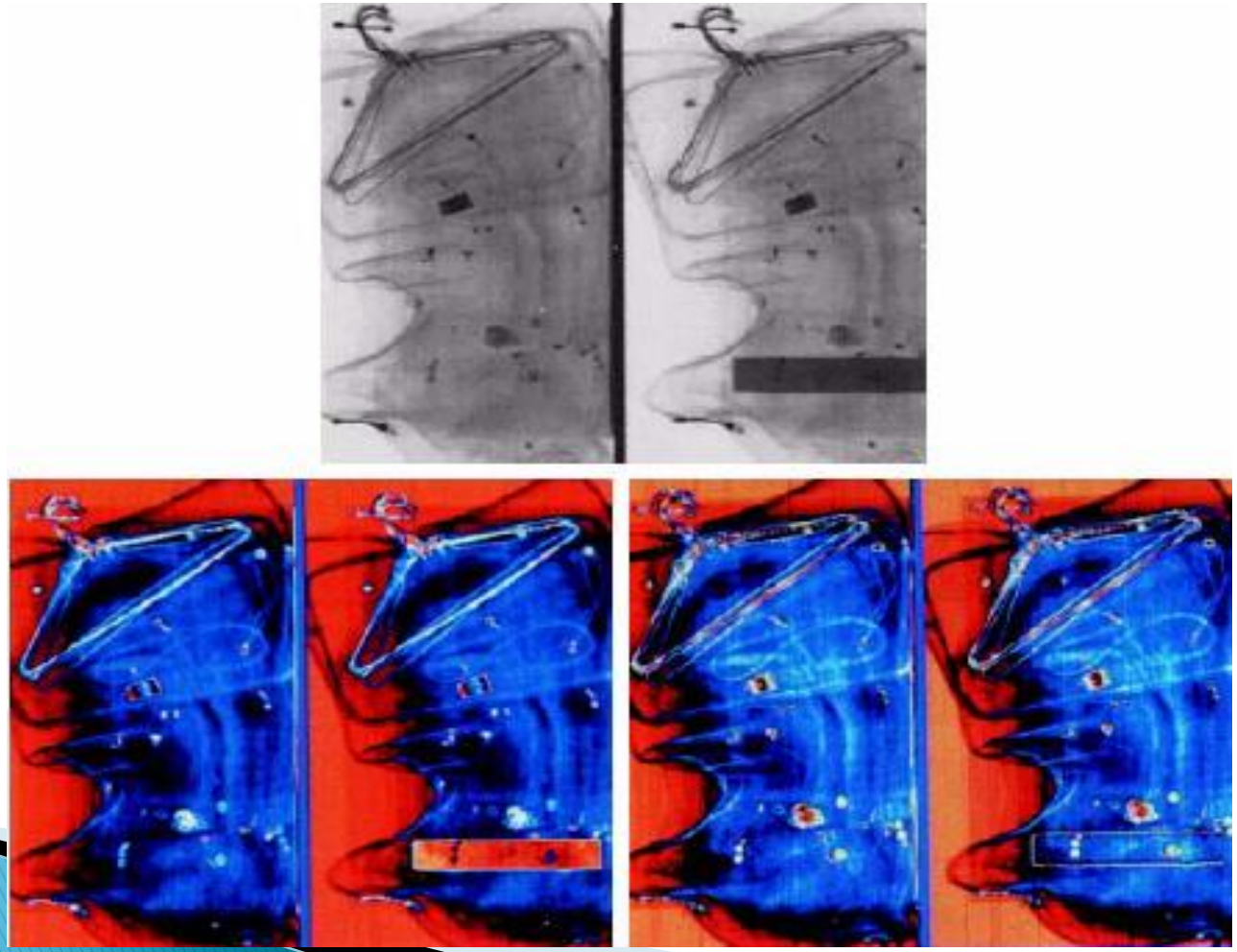
Подготовка к показу или печати

Полутоновое представление



Подготовка к показу или печати

Представление в псевдоцветах в системах безопасности



Подготовка к показу или печати

Спецэффекты



Фото



Имитация карандашного рисунка



Имитация рисунка маслом

Подготовка к показу или печати

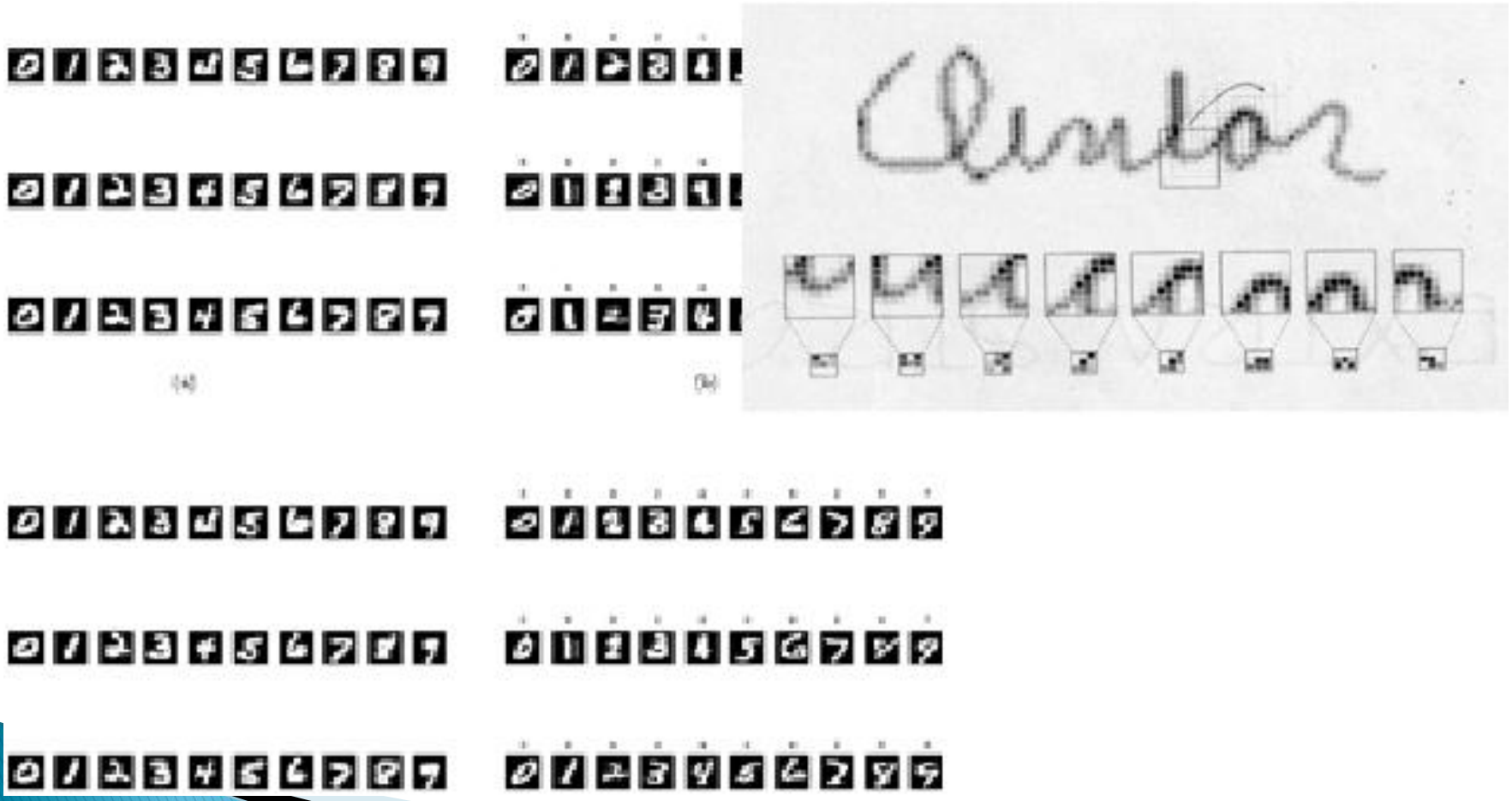
Морфинг





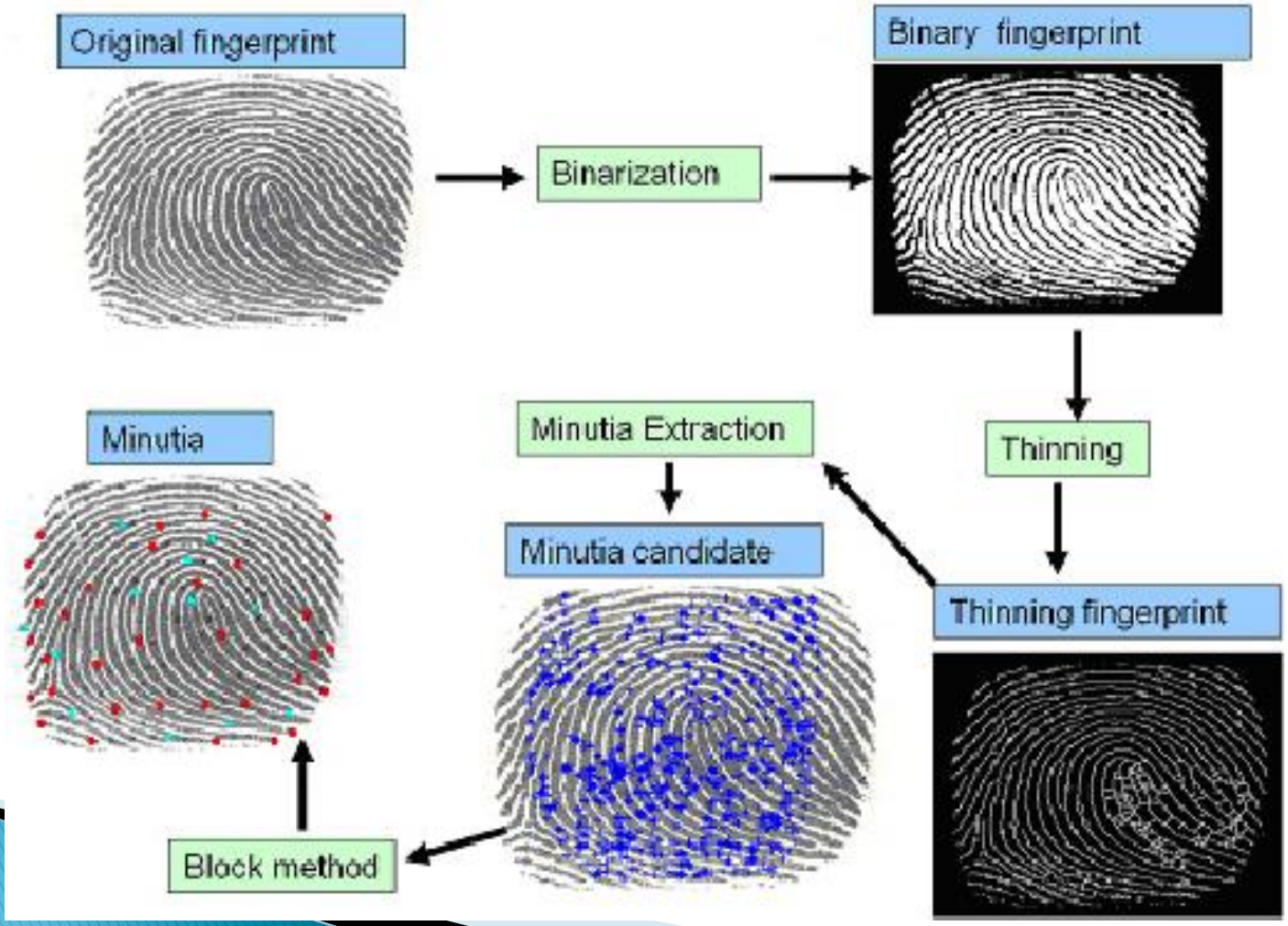
Распознавание изображений

Распознавание рукописного текста



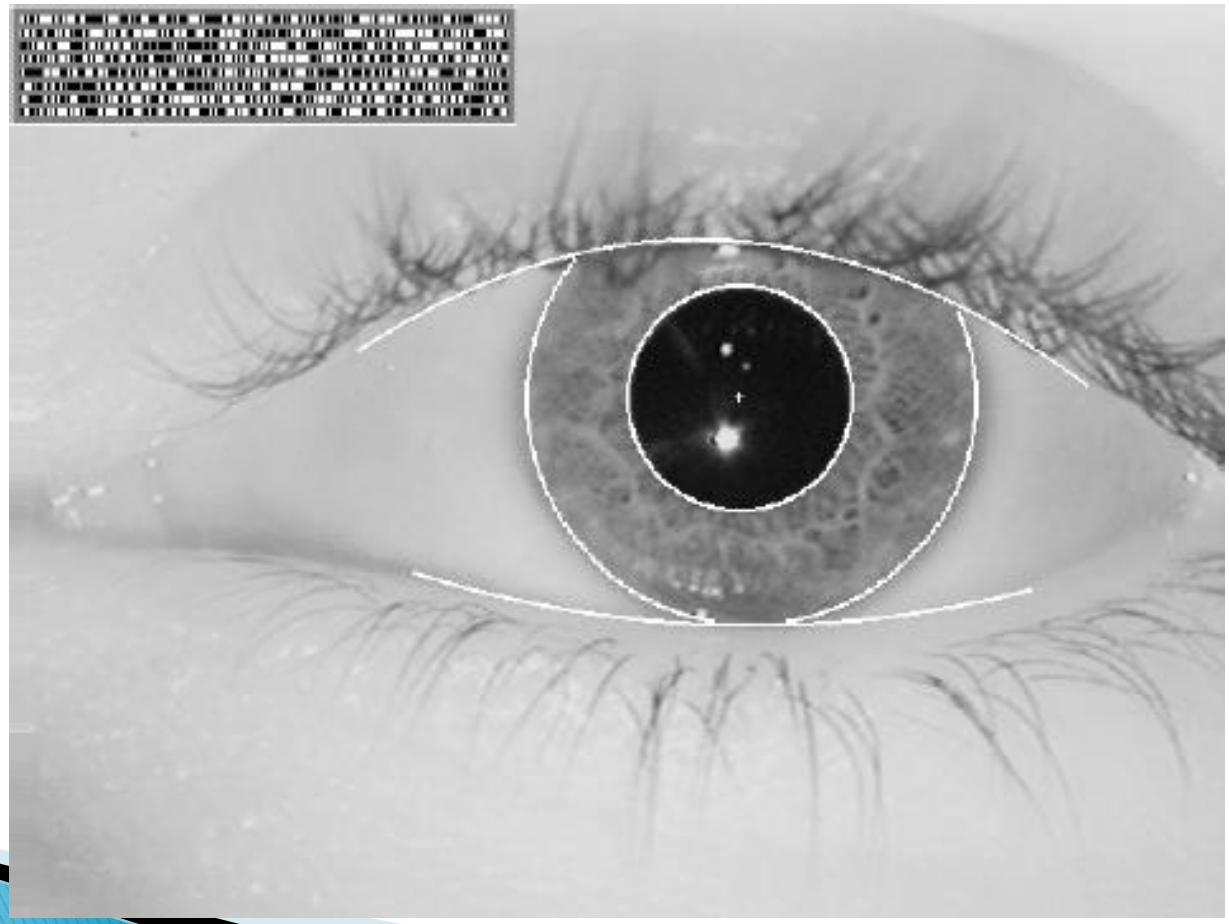
Распознавание изображений

Биометрия: распознавание отпечатков пальцев

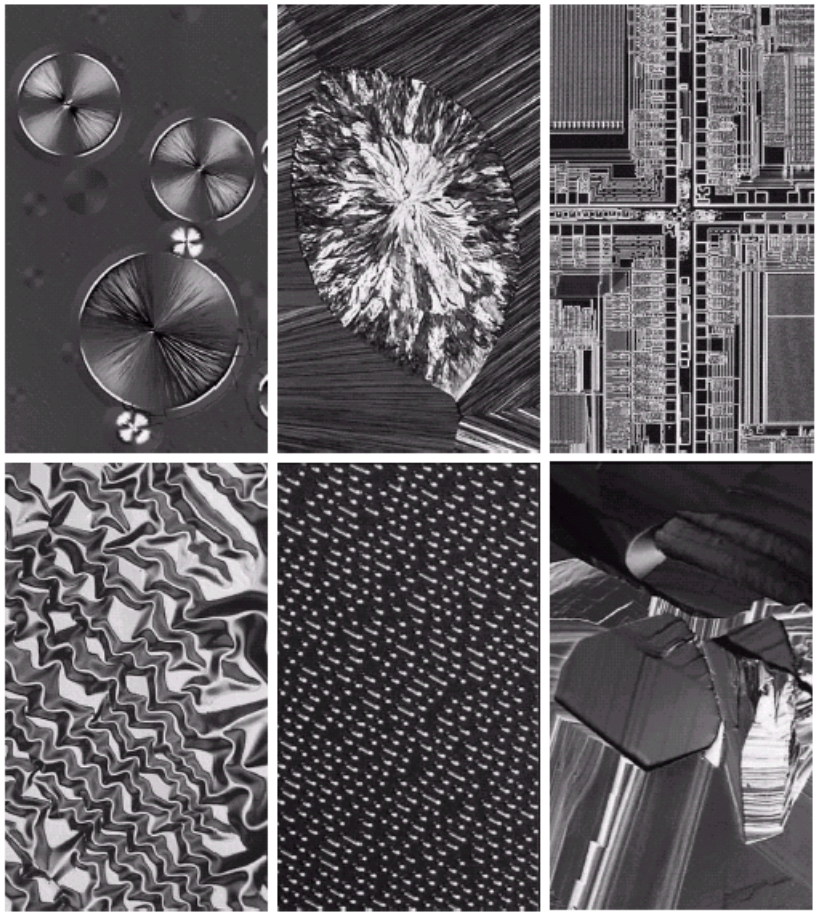


Распознавание изображений

Биометрия: распознавание радужки



Микроскопные изображения



- Таксол (антираковый реагент) увеличение 250×
- Холестерин увеличение 40×
- Микропроцессор увеличение 60×
- Пленка из оксида никеля увеличение 600×
- Поверхность аудио CD увеличение 1750×
- Органический сверхпроводник увеличение 450×

Что мы видим ?









Что мы видим ?









Что мы видим ?







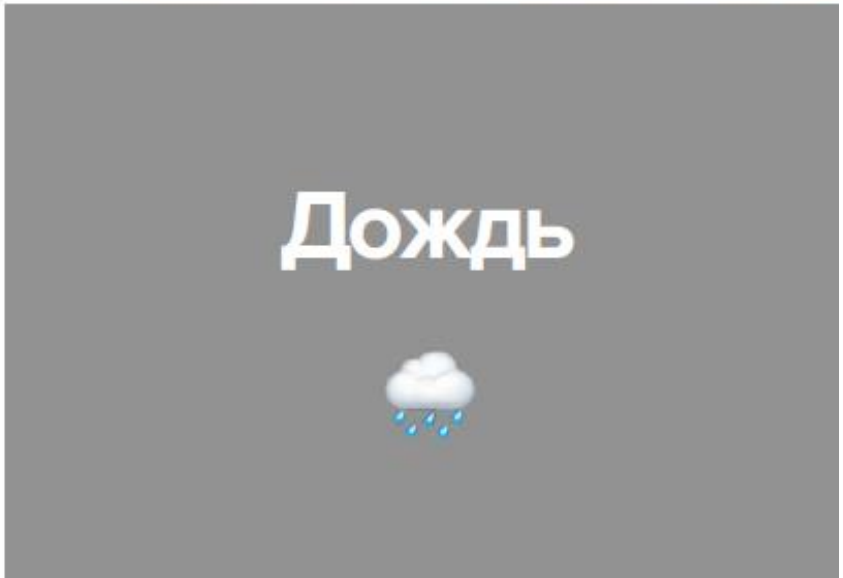


Что мы видим ?









Что мы видим ?







Праздник

Рисовать картины и обрабатывать графику

Какая из картин никогда не принадлежала кисти Рембранта ?



Рисовать картины и обрабатывать графику

Проект Гентского университета и Microsoft

- Полотно напечатано на 3D принтере
- Состоит более чем 148M пикселей
- Основано более чем на 168 000 фрагментов оригинальных картин Рембранта



Рисовать картины и обрабатывать графику

Какая из картин является автопортретом художника Винсента Ван Гога?







Рисовать картины и обрабатывать графику









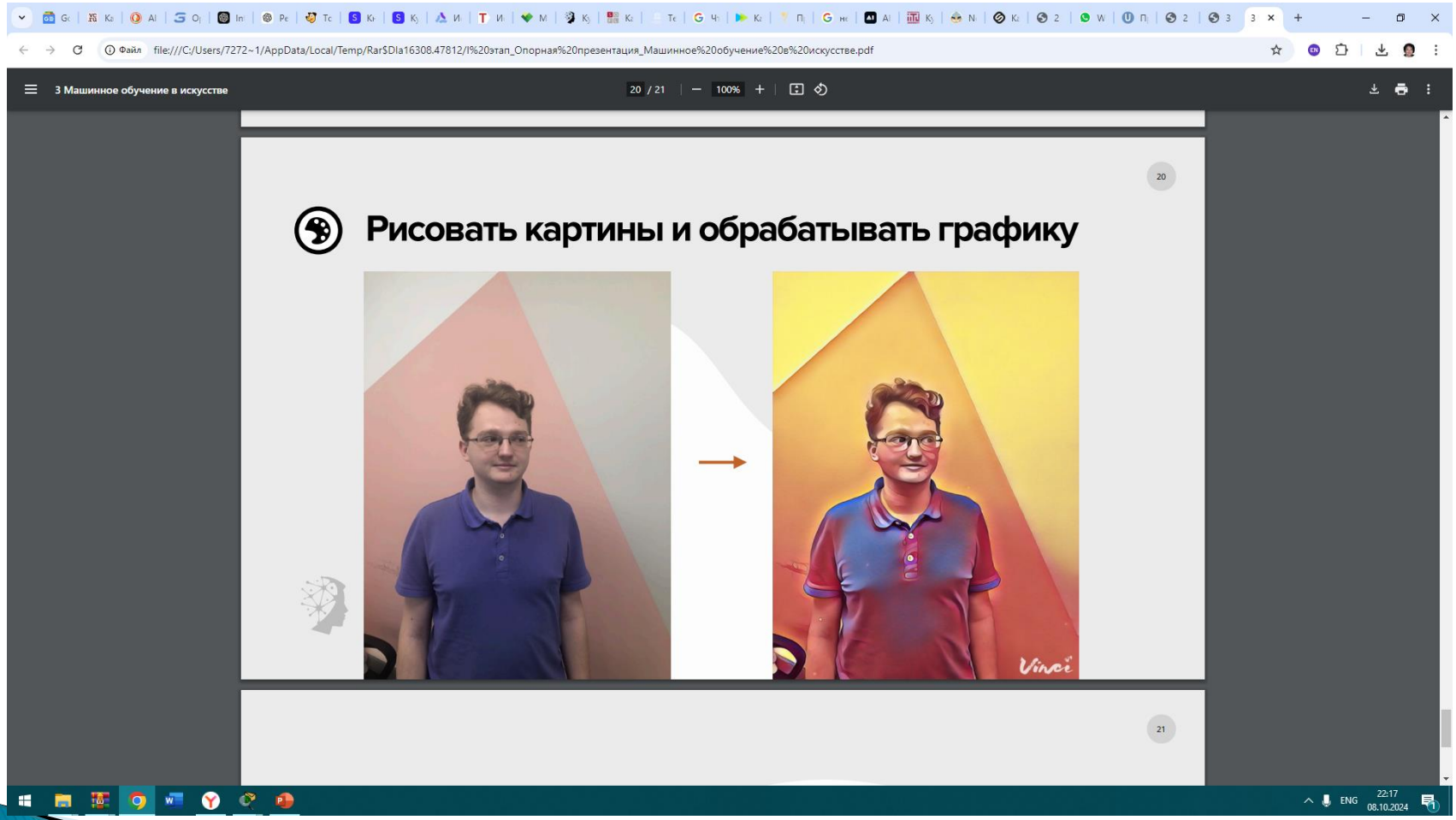
Рисовать картины и обрабатывать графику

Какая картина принадлежит алгоритму?



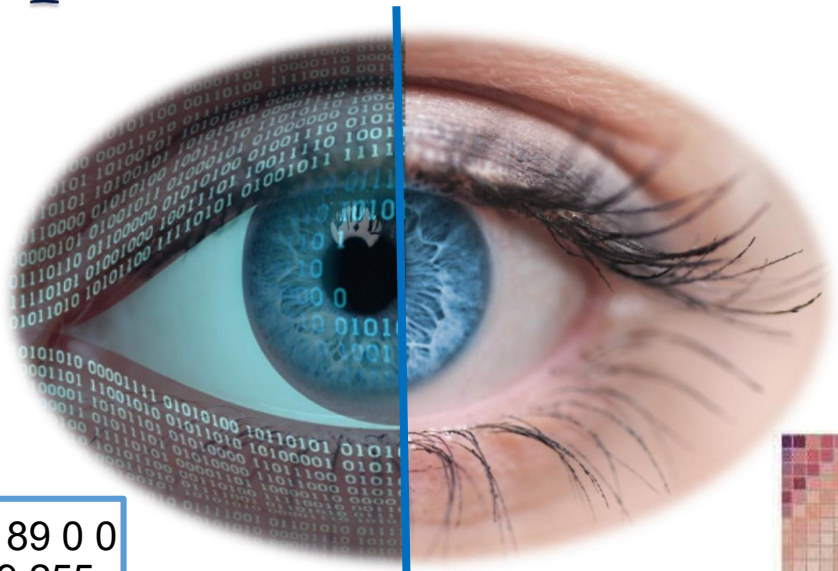


Рисовать картины и обрабатывать графику



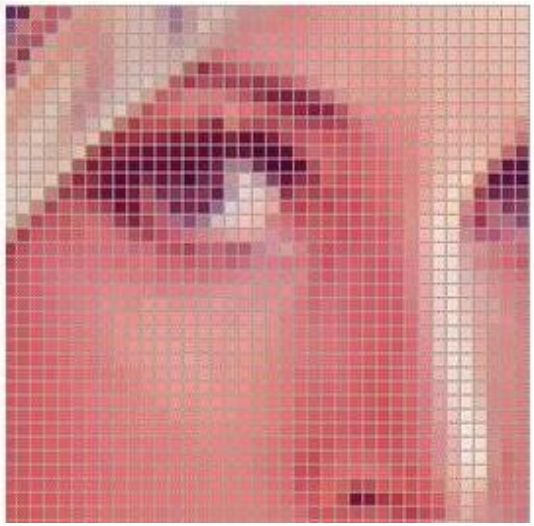
Трудности анализа изображений

Компьютер



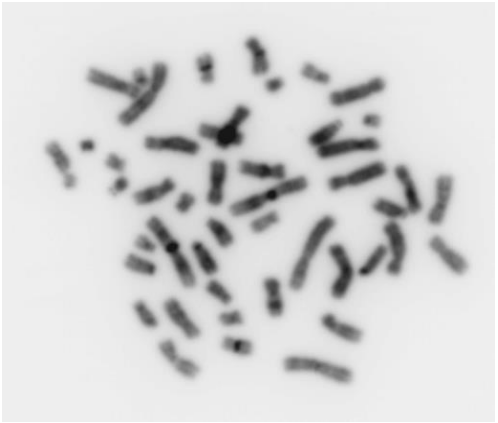
Человек

0 0 95 128 255 96 98 68 89 0 0
78 86 0 255 89 80 89 89 255
255 255 89 80 89 89 255 253
255 89 80 89 89 252 250 0 0
0 89 80 89 89 255 253 3 5 98
255 249 248 89 80 89 89 252

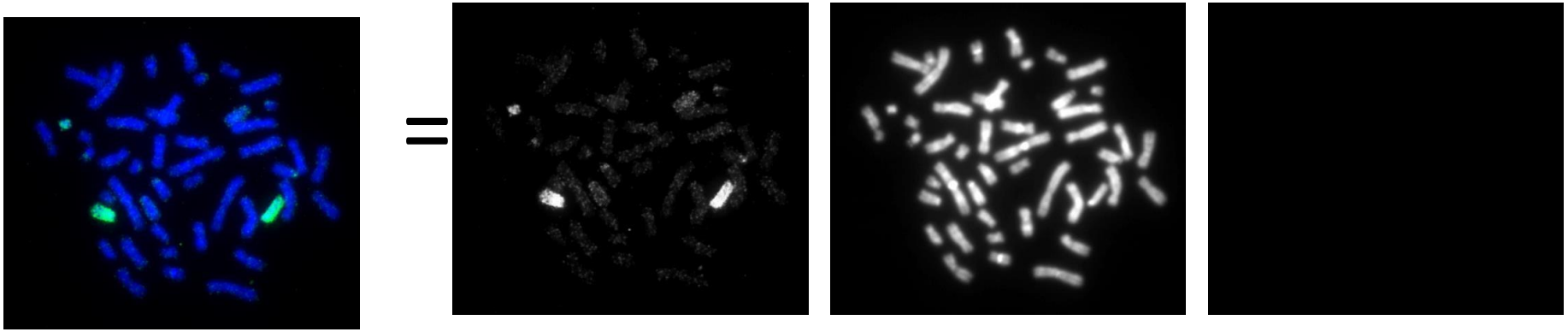


Цветные изображения

Монохромные изображения – двумерный массив чисел



Цветное изображение - серия монохромных изображений (серия массивов)



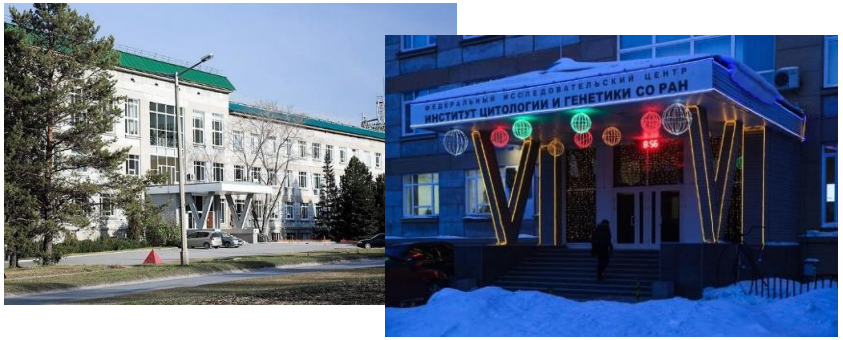
Зеленый Синий Красный

Трудности анализа изображений

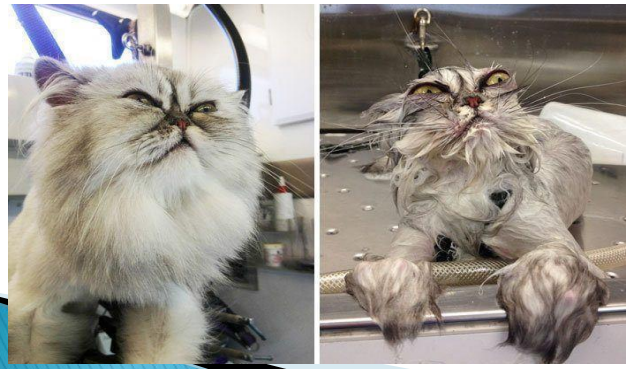
Изображение одних и тех же объектов могут сильно различаться при разных условиях.



Разные ракурсы (точки наблюдения)



Освещение



Форма объектов



Движение

Трудности анализа изображений

Изображение одних и тех же объектов могут сильно различаться при разных условиях.



Фон



Перекрытие



Внутриклассовая изменчивость

Генерируем идеи



Компьютерное зрение в нашем доме...

Что хотим усовершенствовать?	**Распознавание текста** (Необходимо найти изображение текста на картинке и представить его в виде текстовых данных, с которыми можно будет работать, например, в текстовом редакторе)	**Биометрия** (Для распознавания людей может использоваться изображение лица, радужная оболочка глаза, отпечатки пальцев. Однако, в основном, компьютерное зрение занимается распознаванием лиц)	**Видеоаналитика** (В мире устанавливается все больше камер: на дорогах для регистрации движения автомобилей или в общественных местах для отслеживания потоков людей и детектирования аномалий (например, оставленные вещи, нелегальные действия)
Холодильник 	Считываются • штрих-коды товаров, анализируется срок годности и выдается информация о просроченных товарах; • цифровые коды пищевых добавок, на основании чего выводится информация рекомендуемых пределах и максимальных суточных количествах продуктов, содержащих эту добавку и т.п.	Анализирует, кто открывает холодильник и, в соответствии с этим, дает рекомендации. • “Вы можете съесть яблоко. Оно лежит в правом ящике в нижней части” • “Сейчас уже поздно. Поэтому рекомендуется кефир. Он на полке в двери холодильника” • “Помните, что Вы на диете…” • “Мама не разрешает Вам…”	Анализируется наличие продуктов, даются рекомендации по заказу, например, в Интернет-магазине...

Что хотим усовершенствовать?	**Распознавание текста** (Необходимо найти изображение текста на картинке и представить его в виде текстовых данных, с которыми можно будет работать, например, в текстовом редакторе)	**Биометрия** (Для распознавания людей может использоваться изображение лица, радужная оболочка глаза, отпечатки пальцев. Однако, в основном, компьютерное зрение занимается распознаванием лиц)	**Видеоаналитика** (В мире устанавливается все больше камер: на дорогах для регистрации движения автомобилей или в общественных местах для отслеживания потоков людей и детектирования аномалий (например, оставленные вещи, нелегальные действия)
Класс 			
Коридор 			

Что хотим усовершенствовать?	**Распознавание текста** (Необходимо найти изображение текста на картинке и представить его в виде текстовых данных, с которыми можно будет работать, например, в текстовом редакторе)	**Биометрия** (Для распознавания людей может использоваться изображение лица, радужная оболочка глаза, отпечатки пальцев. Однако, в основном, компьютерное зрение занимается распознаванием лиц)	**Видеоаналитика** (В мире устанавливается все больше камер: на дорогах для регистрации движения автомобилей или в общественных местах для отслеживания потоков людей и детектирования аномалий (например, оставленные вещи, нелегальные действия)
Торговый зал 			
Склад магазина 			

Спасибо за внимание!