

Компьютерное зрение и обработка изображений

Лекция 14

Распознавание лиц

Я.М.Демяненко
dem@math.sfedu.ru

Южный федеральный университет
Институт математики, механики и компьютерных наук

2019

Содержание

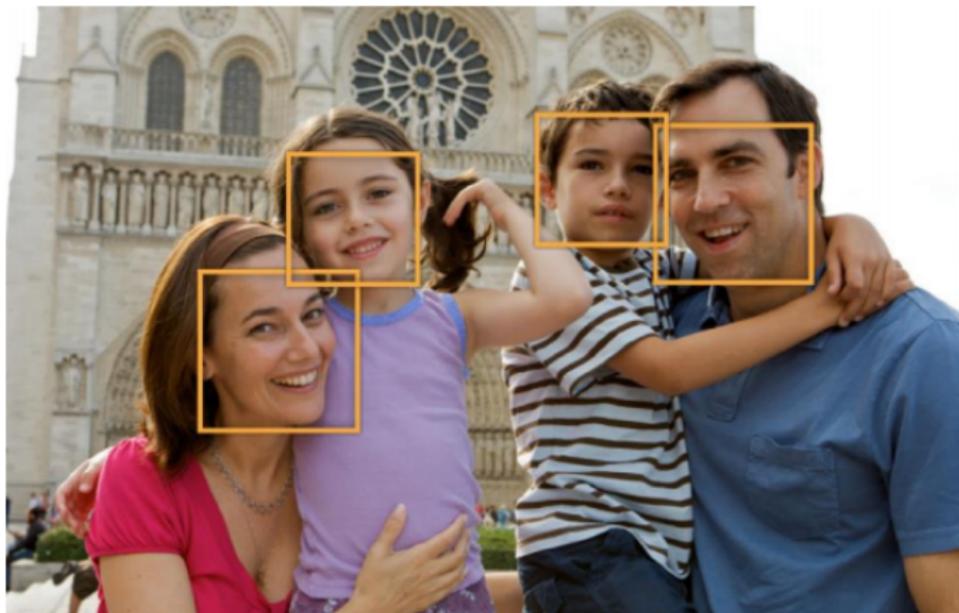
- 1 Обнаружение и локализация лица
 - Некоторые результаты исследований, проведенных MIT
 - Эмпирическое распознавание
 - Как частный случай распознавания
- 2 Метод Виолы-Джонса (Viola-Jones)
 - Принципы метода Виолы-Джонса
 - Признаки Хаара
 - Сканирующее окно
 - Классификаторы в методе Виолы-Джонса
- 3 Различные модификации Виолы-Джонса
- 4 И другие подходы

Содержание

- 1 Обнаружение и локализация лица
 - Некоторые результаты исследований, проведенных MIT
 - Эмпирическое распознавание
 - Как частный случай распознавания
- 2 Метод Виолы-Джонса (Viola-Jones)
 - Принципы метода Виолы-Джонса
 - Признаки Хаара
 - Сканирующее окно
 - Классификаторы в методе Виолы-Джонса
- 3 Различные модификации Виолы-Джонса
- 4 И другие подходы

Обнаружение лица

- На входе: изображение.
- На выходе: координаты найденных прямоугольников лица в системе координат изображения.



Задача обнаружения (face detection)

- является как предобработкой в задачах узнавания лица, распознавания выражения лица и т.д.,
- так и самостоятельной задачей

Модификации задачи обнаружения лица

- face detection
- face localization
- face tracking

Задача	Исходные данные	Результат
Обнаружение лица	Изображение	Вынесение решения о наличии (и, возможно, количестве) лиц на изображении, определение их положения
Локализация лица	Изображение (или его фрагмент), содержащее ровно одно лицо	Положение лица на изображении
Отслеживание лица	Текущий кадр видеопотока, положение лица на предыдущих кадрах	Позиция лица в текущем кадре видео

Проблемы при автоматическом обнаружении

- Сильно варьирующийся внешний вид у разных людей
- Ориентация лица относительно камеры
- Присутствие индивидуальных особенностей (усы, борода, очки, чёлка, макияж, маска)
- Выражение лица
- Лицо может быть частично перекрыто на изображении
- Условия съёмки (освещение, искажения, привносимые оптикой)

Две категории алгоритмов обнаружения

- Методы, опирающиеся на опыт человека.
- Методы, являющиеся частными случаями общей задачи распознавания

Попробуйте узнать



Точность узнавания при сглаживании



Здесь изображены: 1 — Майкл Джордан, 2 — Вуди Аллен, 3 — Голди Хоун, 4 — Билл Клинтон, 5 — Том Хэнкс, 6 — Саддам Хуссейн, 7 — Элвис Пресли, 8 — Джей Лено, 9 — Дастин Хофман, 10 — Принц Чарльз, 11 — Шер, 12 — Ричард Никсон.

Контур



Обработка: фрагментарно или целостно. Кто это?



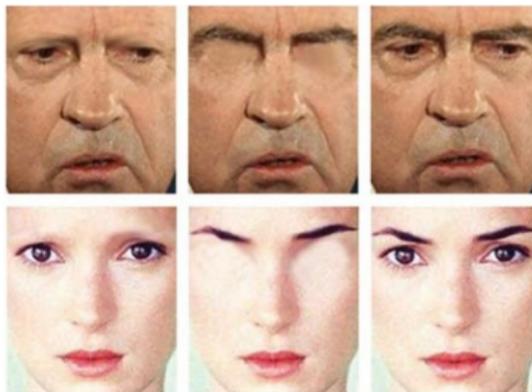
Черты лица обрабатываются как единое целое



Верхняя часть лица принадлежит Вуди Аллену, а нижняя — Опре Уинфри.

Самые важные черты лица для узнавания?

- Чаще всего, результаты экспериментов показывают, что: в порядке убывания, глаза, рот и нос.
- Однако, недавние эксперименты с цифровым стиранием бровей показали, что брови явно недооценены. В частности, процент узнаваемых лиц со стертymi бровями оказался значительно ниже, чем процент узнавания исходных портретов.

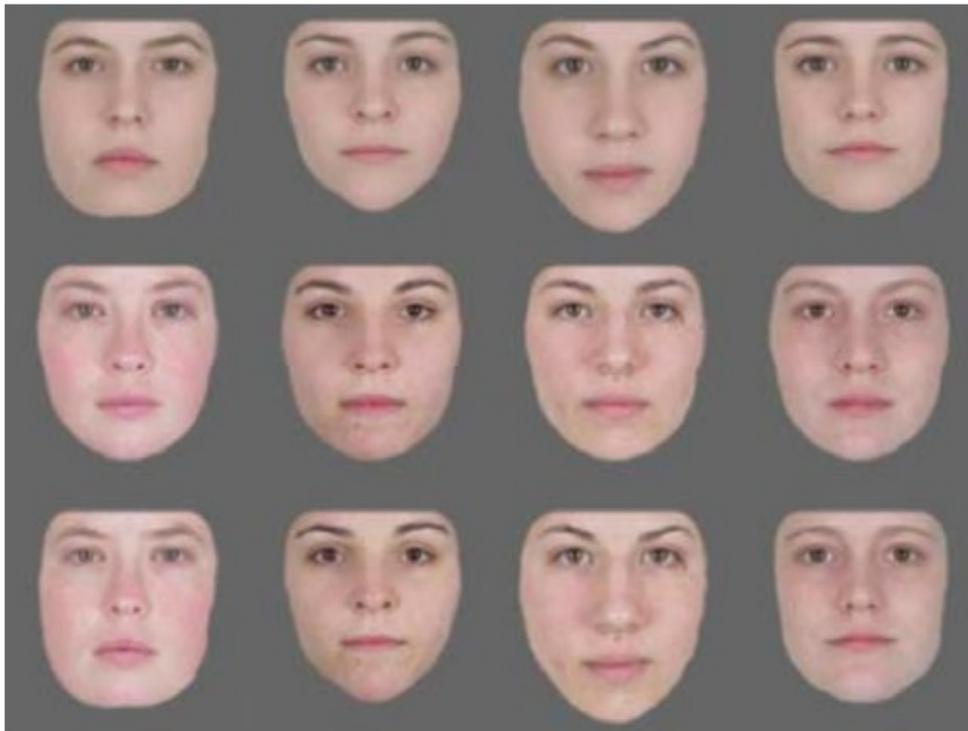


Сильные искажения по ширине



Даже сильные искажения по ширине (25 процентов от исходной) не мешают узнаванию лиц знаменитостей.

Форма или пигментация



Форма или пигментация



Лица в нижнем ряду — лазерные сканы лиц, различающиеся как формой, так и пигментацией. Лица в среднем ряду различаются только пигментацией, но не формой. Лица в верхнем ряду различаются формой, но не пигментацией.

Негатив



Освещение



Одно и то же лицо, освещаемое слева и справа.

Края

- Лицо обычно представляет эллипс
- Резкие перепады яркости часто соответствуют чертам лица

Яркость

- Черты лица зачастую темнее, чем кожа
- Локальные минимумы яркости как потенциальные черты лица

Цвет

- Цвет кожи различных людей занимает достаточно небольшую ограниченную подобласть цветового пространства
- Причём основное отличие в яркости

Характерная форма черт

- Симметричность
- Близость

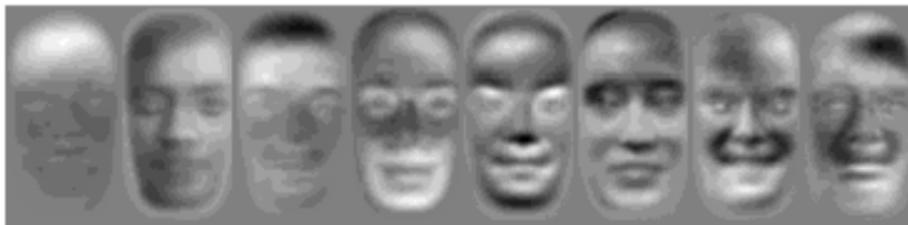
Типичный метод распознавания

- При всём многообразии различных алгоритмов и методов распознавания изображений, типичный метод распознавания состоит из трёх компонент
 - преобразование исходного изображения в начальное представление (может включать в себя как предобработку, так и математические преобразования, например вычисление главных компонент);
 - выделение ключевых характеристик (например берётся первые n главных компонент или коэффициентов дискретного косинусного преобразования);
 - механизм классификации (моделирования): кластерная модель, метрика, нейронная сеть и т.п.
- Построение метода распознавания опирается на априорную информацию о предметной области (характеристики лица человека), и корректируется экспериментальной информацией, появляющейся по ходу разработки метода

Некоторые распространённые подходы

- Метод главных компонент (Principal Component Analysis, PCA)
- Линейный дискриминантный анализ (линейный дискриминант Фишера [Fisher, Face23], Linear Discriminant Analysis, LDA)
- Гибкие контурные модели лица (Flexible Models)
- Сравнение эластичных графов (Elastic Bunch Graph Matchin)
- Методы, основанные на геометрических характеристиках лица

Метод главных компонент (Principal Component Analysis, PCA)



Пример изображений собственных векторов (собственные лица)

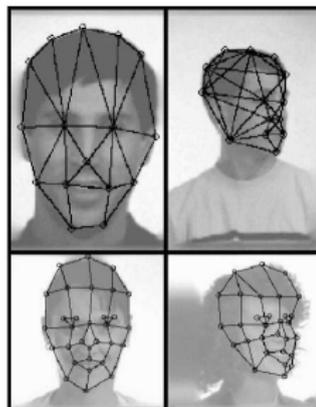
Недостатки PCA

- Метод собственных лиц требует для своего применения идеализированных условий, таких как единые параметры освещённости, нейтральное выражение лица, отсутствие помех вроде очков и бород.
- При несоблюдении этих условий главные компоненты не будут отражать межклассовые вариации

Линейный дискриминантный анализ (Linear Discriminant Analysis, LDA)

- Линейный дискриминантный анализ (линейный дискриминант Фишера [Fisher, Face23], Linear Discriminant Analysis, LDA), выбирает проекцию пространства изображений на пространство признаков таким образом, чтобы минимизировать внутриклассовое и максимизировать межклассовое расстояние в пространстве признаков
- Предполагается что классы линейно разделимы

Эластичный граф, покрывающий изображение лица (Elastic Bunch Graph Matching) [EIGraph]

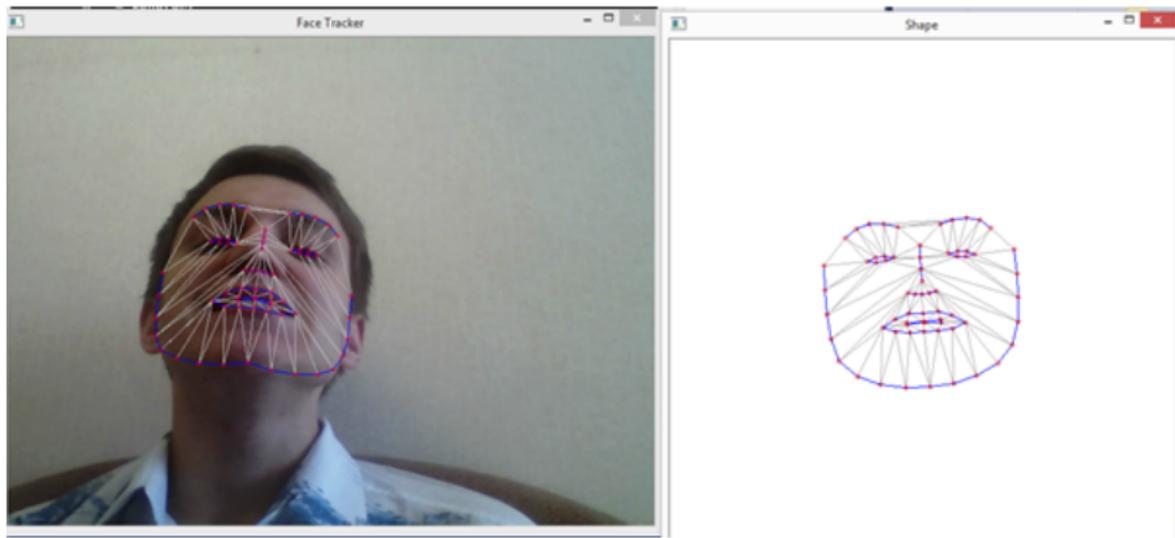


Лицо представляется в виде графа, вершины которого расположены на ключевых точках лица, таких как контуры головы, губ, носы и их крайних точках. Каждая грань помечена расстояниями между её вершинами.

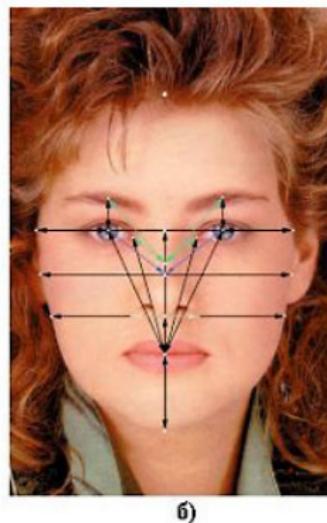
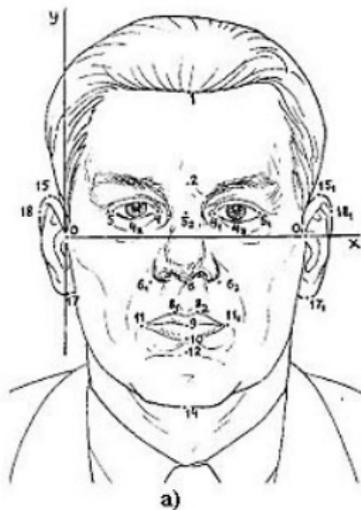
Джеты эластичного графа

- В каждой такой точке вычисляются коэффициенты Габоровых функций.
- Набор таких коэффициентов называется джетом (jet).
- Джеты характеризуют локальные области изображений и служат для двух целей.
- Во-первых, для нахождения точек соответствия в заданной области на двух различных изображениях.
- Во вторых – для сравнения двух соответствующих областей различных изображений.

Распознавание эмоций



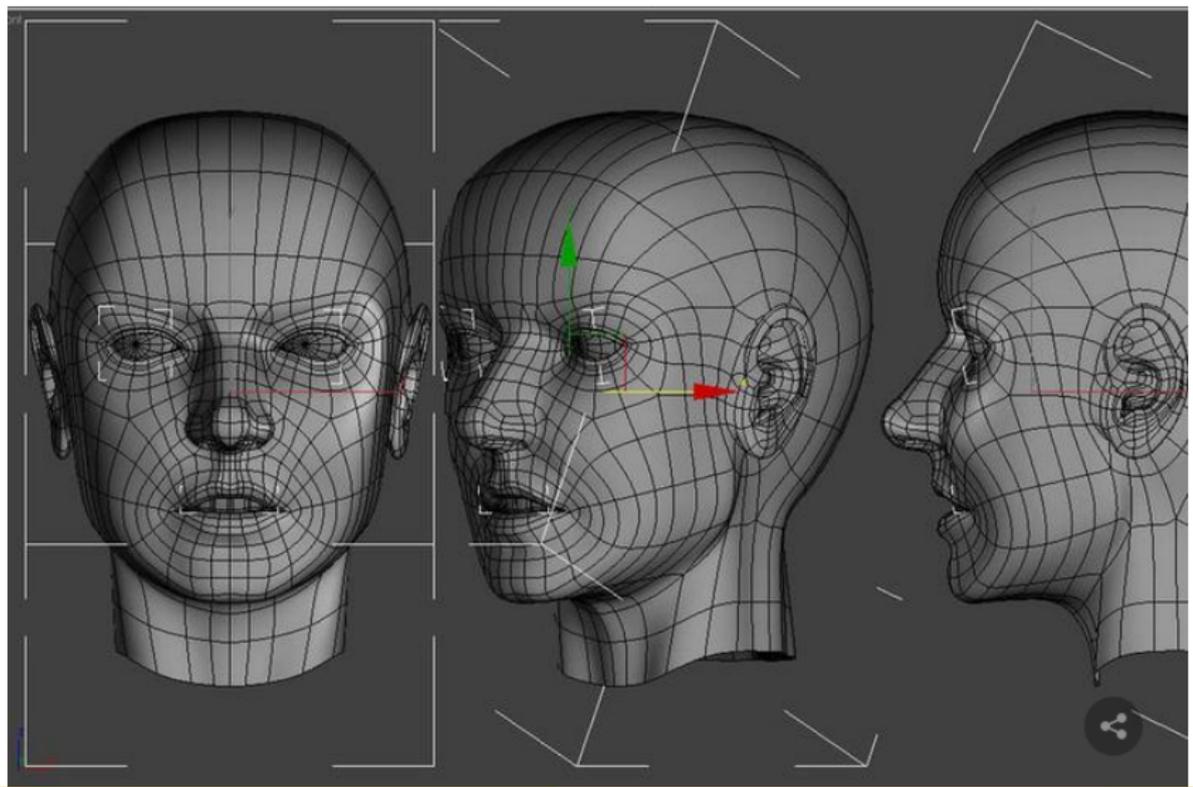
Методы, основанные на геометрических характеристиках лица



Множество отличительных признаков

- К примеру, программное обеспечение Facelt определяет эти признаки как узловые точки.
- Каждое лицо содержит примерно 80 узловых точек: расстояние между глазами, ширина носа, глубина глазных впадин, форма подбородка, длина челюсти.

3D-модель



3D-модель: построение

- Обнаружение: получение снимка при помощи цифрового сканирования существующей фотографии (2D) или видео для получения живой картинке субъекта (3D).
- Центровка: определив лицо, система отмечает положение головы, размер и позу.
- Измерение: система измеряет кривые на лице с точностью до миллиметра и создает шаблон.
- Репрезентация: система переводит шаблон в уникальный код. Этот код задает каждому шаблону набор чисел, представляющих особенности и черты лица.

3D-модель: распознавание

- Сопоставление: если снимок в 3D и база данных содержит трехмерные изображения, сопоставление пройдет без изменений снимка. Но если же база данных состоит из двумерных снимков, трехмерное изображение раскладывается на разные составляющие (словно сделанные под разными углами двумерные снимки одних и тех же черт лица), и они конвертируются в 2D-изображения. И затем находится соответствие в базе данных.
- Верификация или идентификация: в процессе верификации снимок сравнивается только с одним снимком в базе данных (1:1). Если целью же стоит идентификация, снимок сравнивается со всеми снимками в базе данных, что приводит к ряду возможных совпадений (1:N). Применяется тот или иной другой метод по необходимости.

Тенденции

- Использование Convolutional Neural Networks (CNN) где только можно.
- Попытки понять, на какие изображения / области изображений реагирует нейронная сеть.
- Вместо перебора различных классификаторов и подбора их параметров => перебор различных готовых архитектур CNN (AlexNet, ResNet и VGGNet) с подбором их параметров:
 - количество тех или иных слоёв;
 - размерности слоёв;
 - функции активации.

Содержание

- 1 Обнаружение и локализация лица
 - Некоторые результаты исследований, проведенных MIT
 - Эмпирическое распознавание
 - Как частный случай распознавания
- 2 Метод Виолы-Джонса (Viola-Jones)
 - Принципы метода Виолы-Джонса
 - Признаки Хаара
 - Сканирующее окно
 - Классификаторы в методе Виолы-Джонса
- 3 Различные модификации Виолы-Джонса
- 4 И другие подходы

Метод Виолы-Джонса (Viola-Jones)

- Метод был разработан и представлен в 2001 году Полом Виолой и Майклом Джонсом,
- он является основополагающим для поиска объектов на изображении в реальном времени

Основные принципы метода Виолы-Джонса

- используются изображения в интегральном представлении;
- используются признаки Хаара, с помощью которых происходит поиск нужного объекта (в данном контексте, лица и его черт);
- используется бустинг (от англ. boost – улучшение, усиление) для выбора наиболее подходящих признаков для искомого объекта;
- все признаки поступают на вход классификатору, который даёт результат «верно» либо «ложь»;
- используются каскады признаков для быстрого отбрасывания окон, где не найдено лицо.

Характеристики метода

- Обучение классификаторов идет очень медленно, но результаты поиска лица очень быстры.
- Виола-Джонс является одним из лучших по соотношению показателей эффективность распознавания/скорость работы.
- Обладает крайне низкой вероятностью ложного обнаружения лица.
- Алгоритм хорошо работает и распознает черты лица под небольшим углом, примерно до 30 градусов. При угле наклона больше 30 градусов процент обнаружений резко падает.

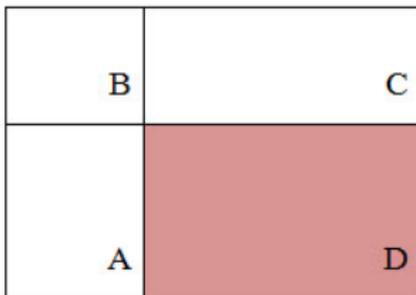
Принцип сканирующего окна

- $rectangle = \{x, y, w, h, a\}$,
- где x, y – координаты центра прямоугольника, w – ширина, h – высота, a – угол наклона прямоугольника к вертикальной оси изображения.
- применяется классификатор к каждому положению окна

Интегральное представление изображений

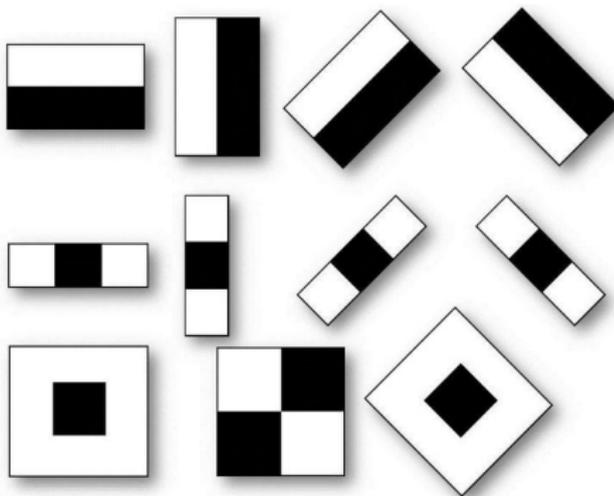
- позволяет быстро рассчитывать суммарную яркость произвольного прямоугольника на данном изображении

$$L(x, y) = \sum_{i=0, j=0}^{i \leq x, j \leq y} I(i, j)$$



Примитивы Хаара

- Признаковые описания допустимо отождествлять с самими объектами.
- В стандартном методе Виолы – Джонса используются прямоугольные признаки — примитивы Хаара:



Вычисляемым значением такого признака будет

- $F = X - Y$,
- где X – сумма значений яркостей точек закрываемых светлой частью признака, а Y – сумма значений яркостей точек закрываемых темной частью признака.
- Для их вычисления используется понятие интегрального изображения.

Примитивы Хаара

- В расширенном методе Виолы – Джонса, используемом в библиотеке OpenCV используются дополнительные признаки:

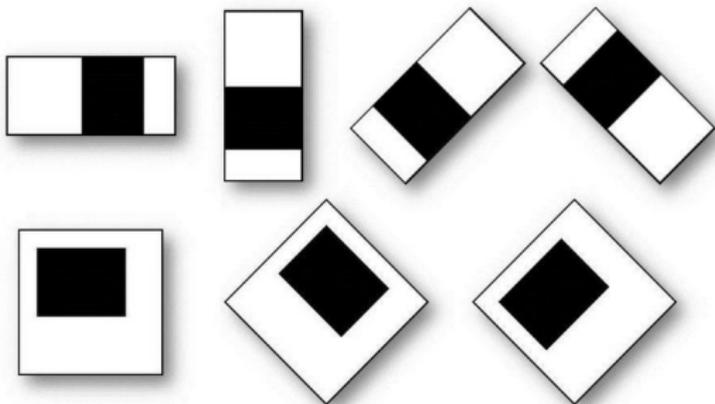
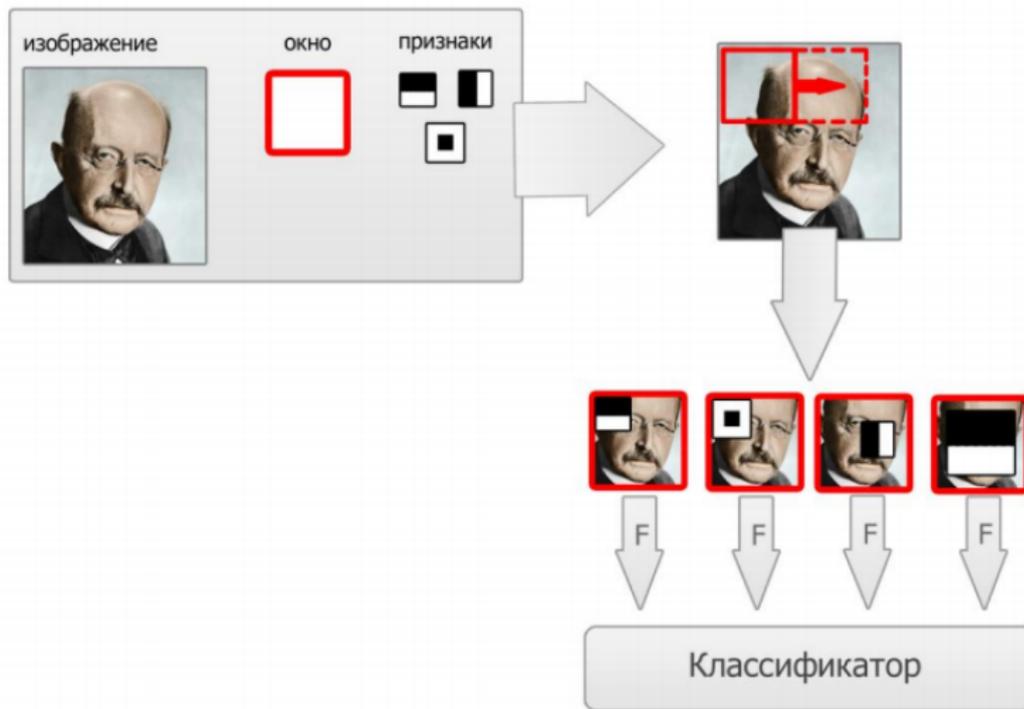
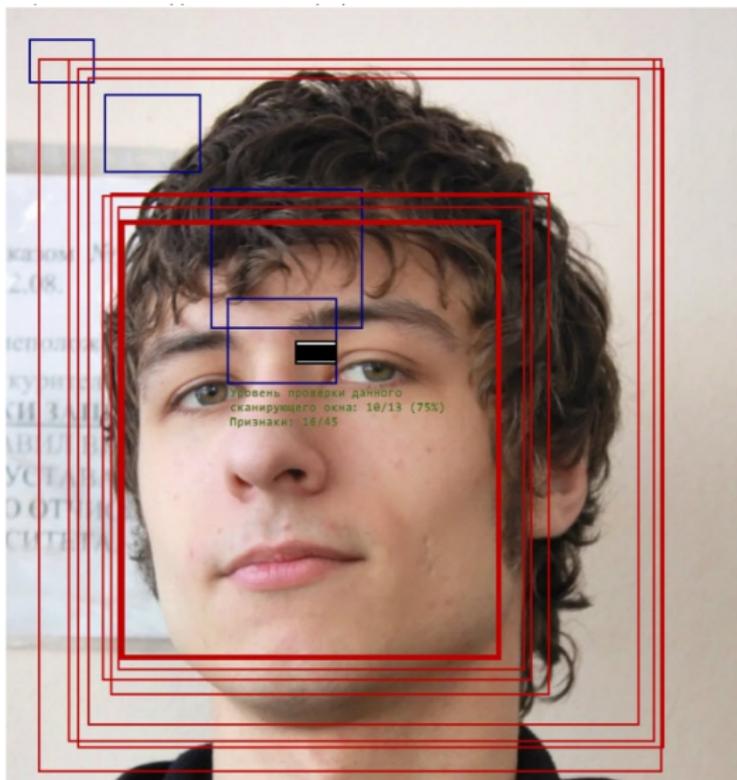


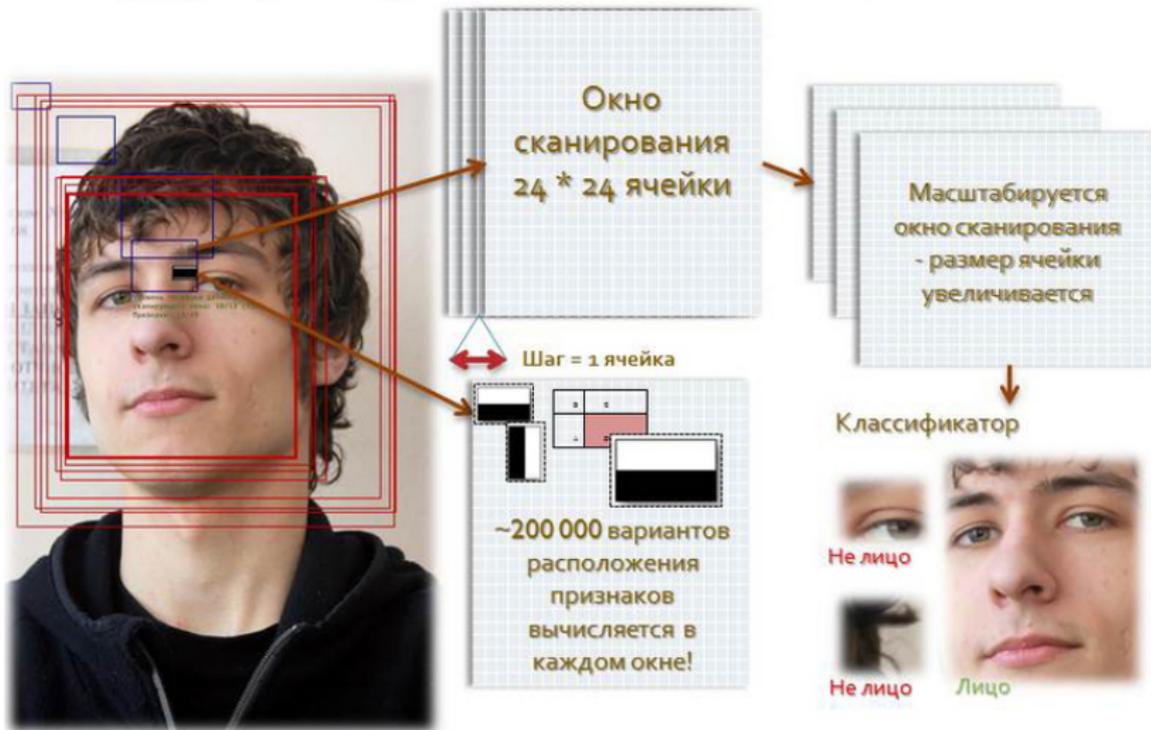
Схема работы сканирующего окна



Сканирующие окна



Алгоритм сканирования окна



Вычисление признаков

- В процессе поиска вычислять все признаки на ПК мало реально.
- Классификатор должен реагировать только на определенное, нужное подмножество всех признаков.
- Нужно обучить классификатор нахождению лиц по данному определенному подмножеству.
- В случае алгоритма Виолы-Джонса для идентификации и распознавания лица классификация является двухклассовой.

Постановка задачи классификации

- Есть X – множество, в котором хранится описание объектов,
- Y – конечное множество номеров, принадлежащих классам.
- Между ними есть зависимость – отображение $Y^*: X \Rightarrow Y$.
- Обучающая выборка представлена $X_m = \{(x_1, y_1), \dots, (x_m, y_m)\}$.
- Конструируется функция f от вектора признаков X , которая выдает ответ для любого возможного наблюдения X и способна классифицировать объект $x \in X$.
- Данное простое правило должно хорошо работать и на новых данных.

Бустинг

- Бустинг (от англ. boosting – повышение, усиление, улучшение) означает дословно «усиление» «слабых» моделей.
- Бустинг – это процедура последовательного построения композиции алгоритмов машинного обучения, когда каждый следующий алгоритм стремится компенсировать недостатки композиции всех предыдущих алгоритмов.

Идея бустинга

- предложена Робертом Шапиром (Schapire) в конце 90-х: имея множество плохих (незначительно отличающихся от случайных) алгоритмов обучения, получить один хороший.
- Построение цепочки (ансамбля) классификаторов, который называется каскадом, каждый из которых (кроме первого) обучается на ошибках предыдущего.
- Например, один из первых алгоритмов бустинга Boost1 использовал каскад из 3-х моделей, первая из которых обучалась на всем наборе данных, вторая – на выборке примеров, в половине из которых первая дала правильные ответы, а третья — на примерах, где «ответы» первых двух разошлись.
- Результат определяется путем простого голосования: пример относится к тому классу, который выдан большинством моделей каскада.

Бустинг

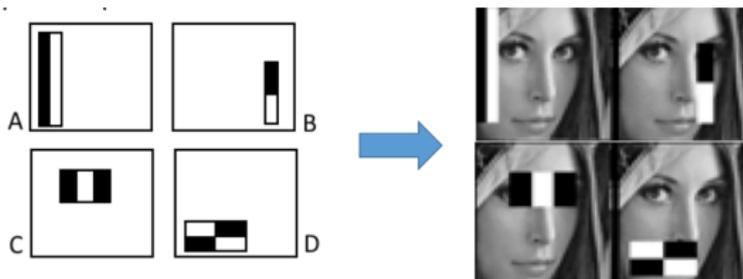
- Бустинг представляет собой жадный алгоритм построения композиции алгоритмов (greedy algorithm)
- Бустинг над решающими деревьями считается одним из наиболее эффективных методов с точки зрения качества классификации.
- Практически неограниченное уменьшение частоты ошибок на независимой тестовой выборке по мере наращивания композиции
- Хорошим качеством могут обладать сколь угодно сложные композиции, если их правильно настраивать
- Всё таки иногда бывает переобучение при наличии значительного уровня шума в данных

AdaBoost

- Развитием данного подхода явилась разработка более совершенного семейства алгоритмов бустинга (AdaBoost: adaptive boosting – адаптированное улучшение), предложенная Йоавом Фройндом (Freund) и Робертом Шапиром (Schapire) в 1999 году
- AdaBoost может использовать произвольное число классификаторов и производить обучение на одном наборе примеров, поочередно применяя их на различных шагах.

Слабые классификаторы

- Слабый классификатор содержит признак Хаара (прямоугольную область, заданную в координатах относительно скользящего окна)



- Ответ слабого классификатора:
 - 1, если разность интенсивностей в чёрных и белых областях больше порога ;
 - 0, иначе.
- Порог определяется на этапе обучения.

Сильные классификаторы

- Сильный классификатор – функция от слабых классификаторов. В зависимости от ответов слабых классификаторов, возвращает ответ вида:
 - точно не лицо;
 - возможно, лицо.
- Сильные классификаторы применяются последовательно в порядке:
 - уменьшения «подозрительности»;
 - увеличения количества используемых слабых классификаторов.

Пример каскадной модели сильных классификаторов



При обучении можно варьировать

- Для каждого слабого классификатора:
 - Вид признака Хаара
 - Координаты признака Хаара
 - Масштаб признака Хаара
 - Весовой коэффициент
- Количество слабых классификаторов для каждого сильного классификатора
- Количество сильных классификаторов

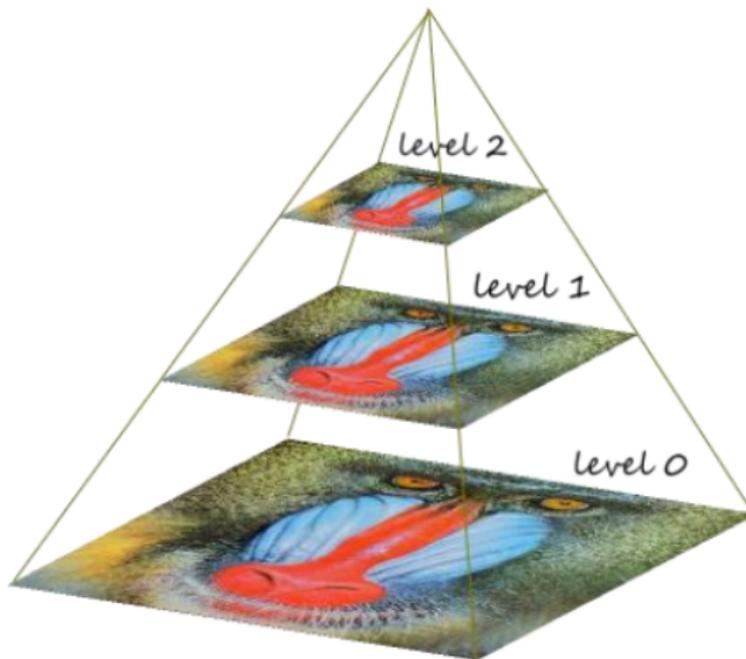
Алгоритм

На входе: изображение I (оттенки серого).

1. Построение интегрального изображения I_{integr} .
2. Обработка изображения I_{integr} скользящим окном размера $N \times N$ ($N=24$):
 - 2.1. Координаты верхнего левого угла скользящего окна: $(x,y) := (0,0)$.
 - 2.2. Последовательное применение сильных классификаторов к скользящему окну. Для каждого сильного классификатора:
 - 2.2.1. Применение слабых классификаторов.
 - 2.2.2. Если общий результат для данного классификатора отрицательный, перейти к Шагу 2.4.
 - 2.3. Запомнить координаты текущего окна, как *координаты прямоугольника лица*.
 - 2.4. Сдвиг скользящего окна:
 - 2.4.1. $x:=x+1$. Если окно **не вышло** за границы изображения, перейти к Шагу 2.2.
 - 2.4.2. $x:=0$. $y:=y+1$. Если окно **вышло** за границы изображения, перейти к Шагу 3.
3. Уменьшение изображения $I := Reduce(I)$. Если ширина или высота уменьшенного изображения больше фиксированного порога δ ($\delta \approx 100$), возврат на Шаг 1.
4. Определение наложенных друг на друга найденных *прямоугольников лица*. Для каждой области наложения:
 - 4.1. Вычисление количества прямоугольников лица, принадлежащих области.
 - 4.2. Если количество прямоугольников больше заданного порогового значения ν , *результатирующий прямоугольник лица* вычисляется путём усреднения координат соответствующих точек прямоугольников лица, принадлежащих данной области наложения.

На выходе: координаты результирующих прямоугольников лица.

Масштабирование



Адаптивное скользящее окно в реальном времени (RASW)

- 1 Подбор шага скользящего окна в зависимости от номера отклонённого сильного классификатора



- 2 F. Comaschi et al., 2013

Ссылки

- 1 Оригинальная работа по алгоритму Виолы-Джонса: P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in Proc. CVPR (Kauai, HI, 2001), pp. 511–518.
- 2 Метод RASW (адаптивный шаг скользящего окна): F. Comaschi, S. Stuijk, T. Basten and H. Corporaal "Rasw: A runtime adaptive sliding window to improve viola-jones object detection Distributed Smart Cameras (ICDSC), 2013 Seventh International Conference.

Содержание

- 1 Обнаружение и локализация лица
 - Некоторые результаты исследований, проведенных MIT
 - Эмпирическое распознавание
 - Как частный случай распознавания
- 2 Метод Виолы-Джонса (Viola-Jones)
 - Принципы метода Виолы-Джонса
 - Признаки Хаара
 - Сканирующее окно
 - Классификаторы в методе Виолы-Джонса
- 3 Различные модификации Виолы-Джонса
- 4 И другие подходы

Модификация 1

- 1 Поиск областей, содержащих изображение кожи человека.
- 2 Уточнение найденных областей: морфологическая обработка изображения эрозия и дилатация + цепного кода Фримена для контуров.
- 3 Применение алгоритма Виолы – Джонса.



Модификация 2



Трекинг лица

- 1 Предлагаемый подход основан на комбинировании метода Виолы-Джонса и модификации обобщенного метода Хафа для поиска произвольных контуров.
- 2 Поиск происходит итеративно по данным о контуре лица из предыдущего кадра. Незначительное изменение контура между кадрами позволяет применять метод поиска с шаблоном для нахождения нового положения лица. При этом происходит корректировка контура.

Результаты экспериментов

Количество кадров видео	5240	3120	4850	3650	6240
Среднее отклонение позиций	2,55%	3,1%	2,8%	3,05%	3%
Процент кадров потерянных методом Виолы-Джонса	5%	7%	6%	15%	12%
Процент кадров потерянных В.-Дж. и найденных комбинированным методом	60%	56%	61%	50%	48%

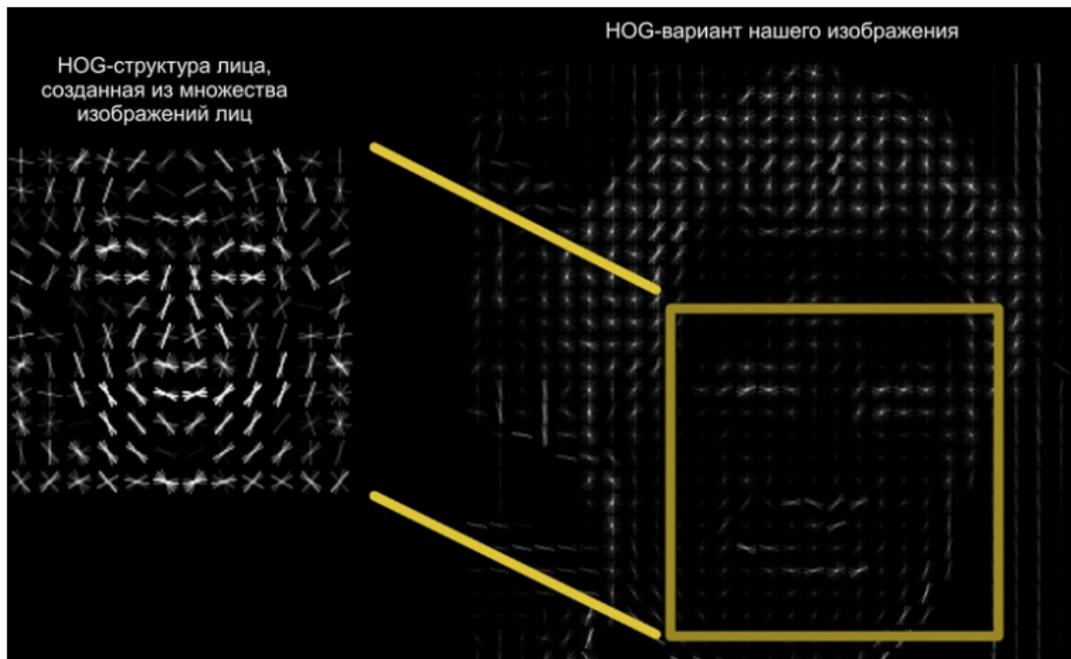
Каскадный алгоритм для трекинга лица

- 1 в виде последовательности алгоритмов сопровождения, которые сменяют друг друга в случае ухудшения качества отслеживания.
- 2 В качестве базовых алгоритмов используется кросскорреляция шаблонов и алгоритм на основе сопоставления областей по цвету с использованием фильтра частиц.
- 3 Предлагается дополнить каскад новым алгоритмом для отслеживания лица с использованием маски переднего плана.

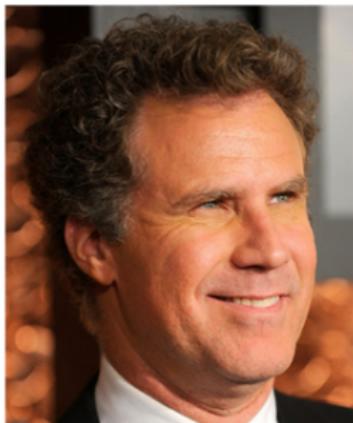
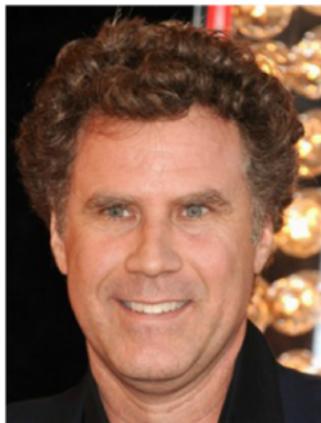
Содержание

- 1 Обнаружение и локализация лица
 - Некоторые результаты исследований, проведенных MIT
 - Эмпирическое распознавание
 - Как частный случай распознавания
- 2 Метод Виолы-Джонса (Viola-Jones)
 - Принципы метода Виолы-Джонса
 - Признаки Хаара
 - Сканирующее окно
 - Классификаторы в методе Виолы-Джонса
- 3 Различные модификации Виолы-Джонса
- 4 И другие подходы

HOG-дескриптор



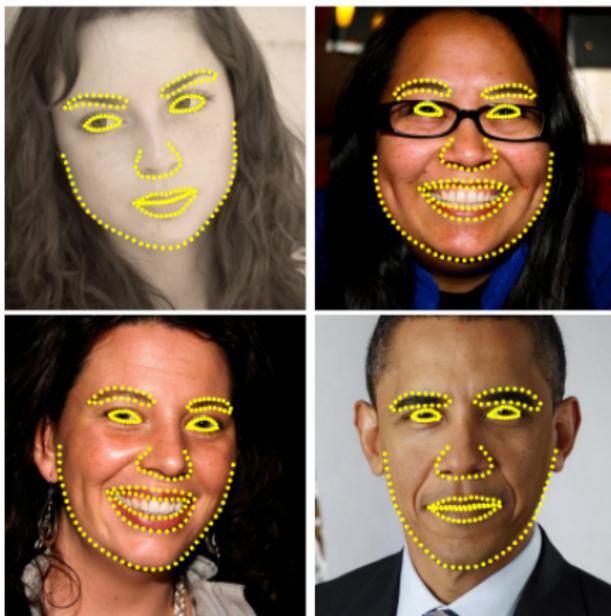
Разные ракурсы



Люди могут легко увидеть, что оба изображения относятся к актёру Уиллу Ферреллу, но компьютеры будут рассматривать их как лица двух разных людей.

Оценка антропометрических точек

- подход, предложенный в 2014 году Вахидом Кэземи и Джозефином Салливаном.



Оценка антропометрических точек



- Основная идея в том, что выделяется 68 специфических точек (меток), имеющихя на каждом лице, — выступающая часть подбородка, внешний край каждого глаза, внутренний край каждой брови и т.п. Затем происходит настройка алгоритма обучения машины на поиск этих 68 специфических точек на каждом лице

Аффинные преобразования, но не 3D-деформации

- Теперь, когда мы знаем, где находятся глаза и рот, мы будем просто вращать, масштабировать и сдвигать изображение так, чтобы глаза и рот оказались отцентрованы как можно лучше.

