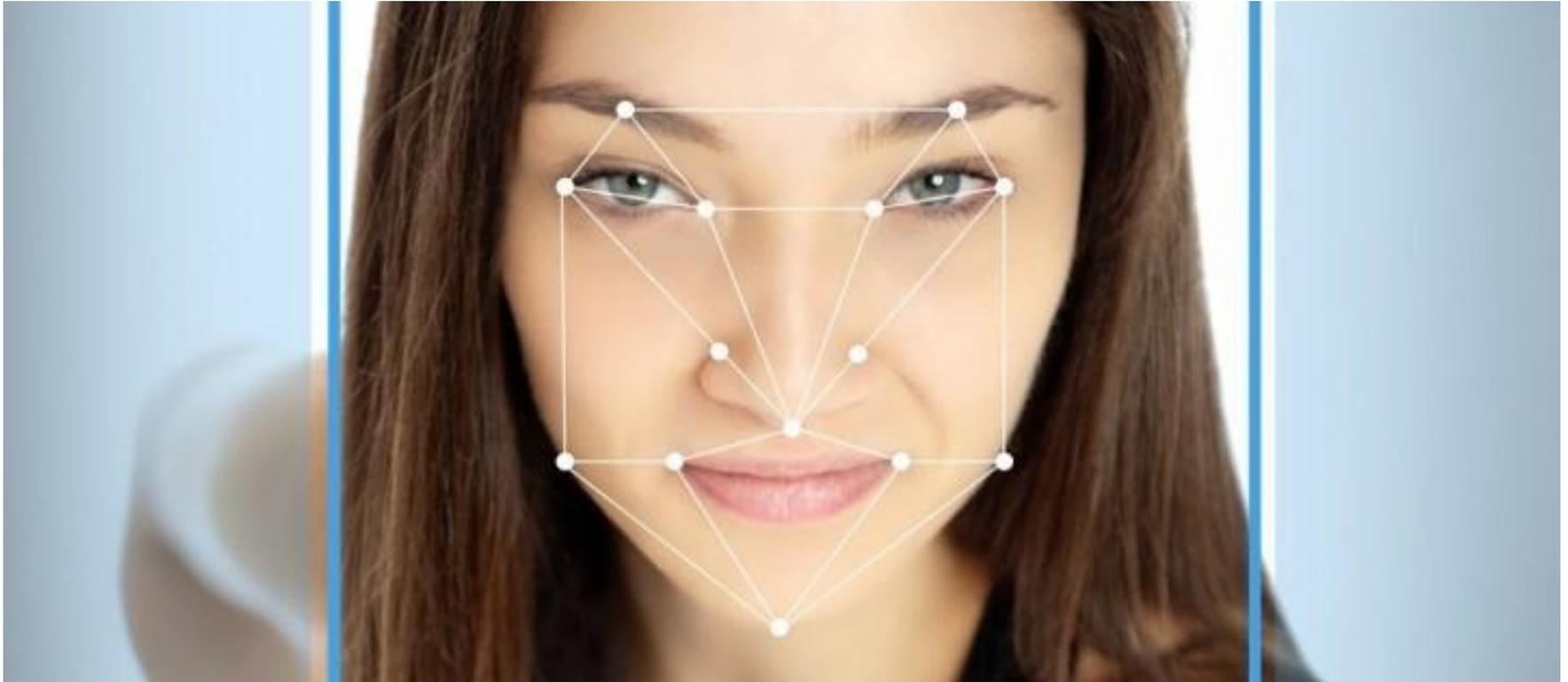


3. Распознавание лиц



3. Распознавание лиц

- Задача обнаружения (детектирования) лиц ставит перед собой цель локализации, определения и выделения лица на цифровом изображении или видео.
- Обнаружение лица является первым шагом в решении задач более высокого уровня: **распознавание лиц**, автоматический учёт числа посетителей в системах наблюдения и безопасности, автоматическая фокусировка на лице и стабилизация изображения в фототехнике.
-
- Практически все современные цифровые фотокамеры распознают при съёмке лица и наводят фокус на них.
- Механизм детектирования: для статичных изображений **метод определяет характерные черты** и рамкой изолирует лицо от прочих объектов на картинке; для видео - **отслеживает лицо в потоке видеок кадров**

3. Распознавание лиц

- Таким образом, при распознавании лиц ставится самое меньшее две разные задачи:
- 1) Задача обнаружения (выделения, детектирования) лица на цифровом изображении или видео.
- 2) Задача распознавания лица, в которую может входить определение параметров личности (пол, возраст, этническая принадлежность и др.) и идентификация личности по имеющимся данным (фотография, описание характерных признаков и др.)
-
- Это две разные задачи, которые решаются различными методами

3. Распознавание лиц

- Человеческий мозг справляется с обнаружением лиц быстро, затрачивая на задачу, как правило, не более одной секунды.
- Для компьютерного зрения процесс детектирования гораздо более сложный, так как лица человека представляют собой эластичные нефиксированные объекты с большим набором изменяющихся признаков: **размер, форма, цвет, текстура**. К тому же, задача усложняется тем, что локализации лица должна быть выполнена **независимо от масштаба, ориентации лица в плоскости изображения, условий освещения, положения и удалённости камеры**.
- Алгоритм детектирования выделяет лицо на фотоснимке или видеофайле. После детектирования обнаруженное лицо может быть распознано: устанавливается личность человека.

3. Распознавание лиц

- Задача выделения (детектирования) лица на цифровом изображении.
- Методы выделения и распознавания лиц всегда были интересны в связи с все практическими потребностями предприятий: системы охраны, верификация кредитных карточек, криминалистическая экспертиза и т.д.
- Распознавание лиц давно стало популярной областью исследований в компьютерном зрении и разработчики этих методов и алгоритмов всегда востребованы на ИТ предприятиях.
- Для решения задач выделения и распознавания лиц были предложены различные методики, среди которых подходы, основанные на нейронных сетях, на разложении Карунена-Поэва на алгебраических моментах, на линиях одиноко-

3. Распознавание лиц

- Метод Виолы-Джонса (2001г):
- Самый популярный метод для поиска области лица на изображении;
- Имеет высокую скорость и эффективность;
- Позволяет осуществлять поиск лица в режиме реального времени.
- Применяется как в RGB, так и полутоновым изображениям.
- Для RGB используются все 3 цветовых канала.
- В основе метода Виолы-Джонса по поиску лица лежат 3 идеи:
 - 1) интегральное представление изображения по признакам Хаара,
 - 2) метод построения классификатора на основе алгоритма адаптивного бустинга,
 - 3) метод комбинирования классификаторов в каскадную

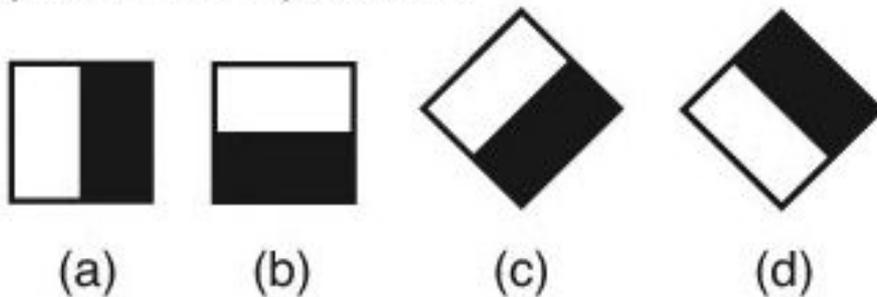
3. Распознавание лиц

- Интегральное представление и признаки Хаара.
- В этом представлении вместо исходного RGB изображения $I(i, j)$ размерности $m \times n$ (m строк и n столбцов) на протяжении всей процедуры используется матрица $L(i, j)$ той же размерности но с накопленными показателями яркости. Яркость пикселя $L(i, j)$ сумме яркостей всех пикселей прямоугольника с вершинами от $I(1, 1)$ до $I(i, j)$:
$$L(i, j) = \sum_{k=1}^i \sum_{l=1}^j I(k, l)$$
- Поскольку алгоритму приходится постоянно пересчитывать сумму интенсивностей в прямоугольных фрагментах исходного изображения, то накопленная интенсивность позволяет значительно снизить число операций сложения.

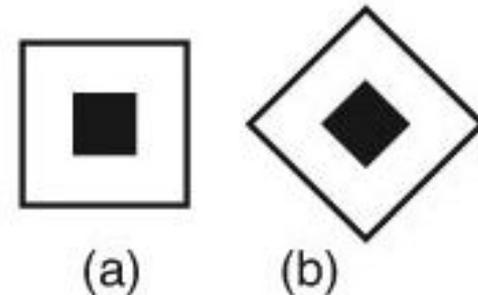
3. Распознавание лиц

• Признаки Хаара действуют на изображения как свертка, пробегая по нему. На каждом прямоугольнике Хаара вычисляет параметр, равный разности суммарных интенсивностей области рисунка, заметаемой светлой частью маски и

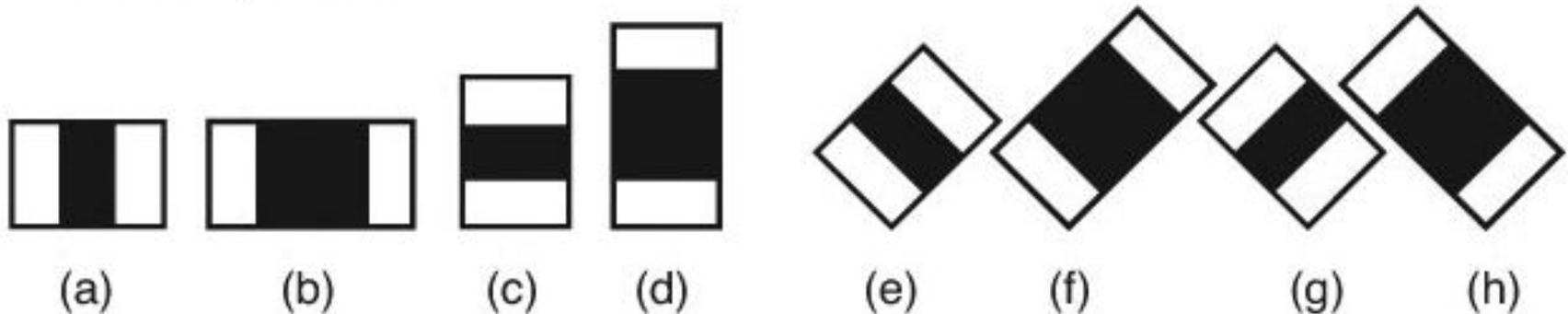
1. Граничные признаки



3. «Центральные» признаки



2. Линейные признаки



3. Распознавание лиц

- То есть, если это была бы обычная свертка, светлая часть маски заполнена числами 1, а темная числами -1.
- На самом деле, если стороны прямоугольника-маски горизонтальны и вертикальны (то есть ориентированы по рисунку), то использование накопленных интенсивностей $L(\cdot, \cdot)$ позволяет сократить число операций.
- Каждая маска пробегает по всему изображению, как обычно, с шагом 1 пиксель. Получаются массивы параметров, которые поступают на классификатор.

• **Алгоритм сканирования** окна с масками:

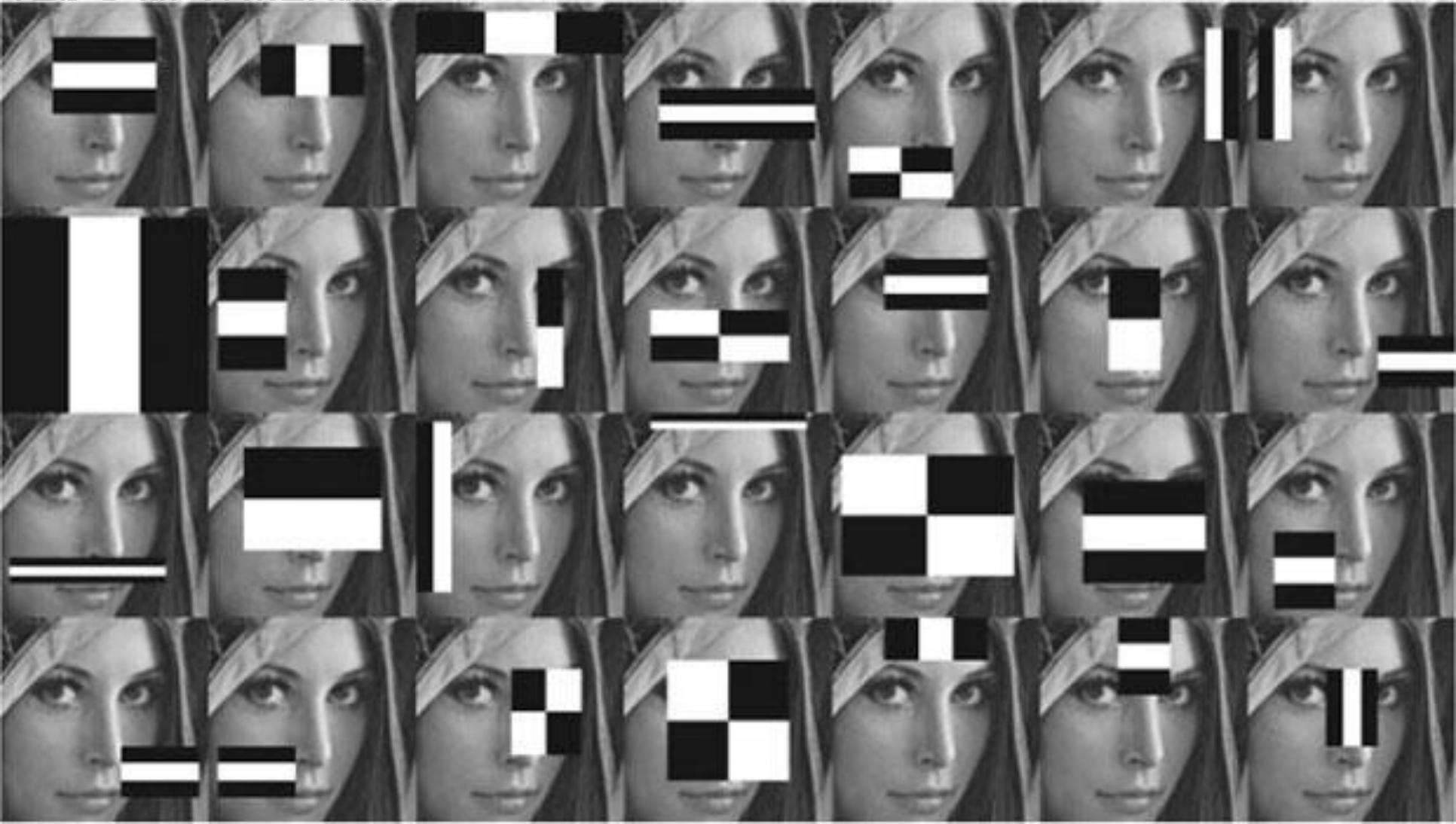
- 1) выбирается размер окна сканирования;
- 2) выбираются маски (возможно все);
- 3) далее окно сканирования начинает последовательно двигаться по изображению с шагом в 1 пиксель (пусть

3. Распознавание лиц

- 4) в каждом окне сканирования **вычисляется сотни тысяч** (напр 200 000) вариантов расположения масок (варьируются масштабы масок и их положения в окне сканирования, масштабируется не изображение, а сканирующее окно);
- все найденные признаки попадают к классификатору, который выносит решение.

3. Распознавание лиц

- Примеры применения признаков (фильтров) Хаара. Реально они применяются к интегрированному изображению (не к исходному)



3. Распознавание лиц

- Построение классификатора с использованием адаптивного бустинга.
- Частным случаем кластеризации является классификация, когда заранее известно число классов, на которые следует разделить рассматриваемое множество.
- При выделении лиц (носа, глаз и т.п.) число классов равно трем, они соответствуют решению модуля-классификатора:
 - 1) да - объект принадлежит классу (лицо, нос, глаз, ...);
 - 2) нет - объект не принадлежит классу;
 - 3) не определено (не получен вывод о принадлежности классу).
- Обучение классификатора выполняется на выборке из генеральной совокупности. Выборка должна быть репрезен¹²

3. Распознавание лиц

- Для обучающей выборки все объекты должны быть отнесены только к двум классам «Да» или «Нет».
- В классификаторе применяется технология **бустинга** (boosting - повышение, усиление, улучшение) - это комплекс методов для повышения точности методов.
- Бустинг строит последовательную процедуру композиции алгоритмов машинного обучения, когда каждый следующий алгоритм стремится компенсировать недостатки композиции всех предыдущих алгоритмов.
- Идея бустинга была предложена Р. Шапиро в 90-е гг.
- Он на основе плохих, незначительно отличающихся от случайных алгоритмов обучения создавал более сложный хороший. В основе идеи лежит построение **ансамбля** классификаторов, который он назвал **каскадом**, каждый следу¹³-

3. Распознавание лиц

- Один из первых алгоритмов Шапиро решал задачу машинного обучения. Он использовал **3 каскада**, **первый** обучался на всей обучающей выборке, **второй** - на выборке элементов, в половине из которых **первый каскад** дал правильные ответы, а **третий** каскад — на примерах, где ответы первых двух разошлись.
- Таким образом, имеет место **последовательная обработка элементов каскадом классификаторов**, при этом задача для каждого последующего становится труднее.
- Окончательный результат определяется путем простого голосования: пример относится к классу Да или Нет, который выдан большинством моделей каскада.
- Бустинг представляет собой **жадный алгоритм** построения композиции алгоритмов — это алгоритм, который на

3. Распознавание лиц

- В задачах кластеризации и классификации бустинг считается одним из **наиболее эффективных методов с точки зрения результата**.
- В экспериментах наблюдалось практически неограниченное уменьшение частоты ошибок на независимой тестовой выборке по мере наращивания композиции каскадов.
- Более того, качество на тестовой выборке часто продолжало улучшаться даже после достижения безошибочного распознавания всей обучающей выборки.
- (Лучшее - враг хорошего?)
- Это изменило существовавшее долгое время общее мнение о том, что для повышения обобщающей способности **необходимо ограничивать сложность алгоритмов**. То есть, на примере бустинга стало видно, что хорошим качеством

3. Распознавание лиц

- Алгоритмы классификации часто имеют такую структуру: сначала вычисляются оценки принадлежности объекта классам, затем решающее правило переводит эти оценки в номер класса.
- Обычно оценка - это вещественное число на отрезке $[0, 1]$
- Если на каскаде элемент получил оценку ≤ 0.5 , то его относят к классу «Нет», иначе к классу «Да».
- Решающие правила могут гораздо более сложными и иметь настраиваемые параметры.
- В алгоритме Виолы-Джонса вначале каждому элементу присваивается значение 0, которое повышается в процессе обработки.
- Алгоритм Виолы-Джонса после обучения проверяется на ¹⁶

3. Распознавание лиц

- Бустинг алгоритма Виолы-Джонса. Основа классификатора - признаки Хаара.
- Цель - выделить на каждом элементе выборки лица людей.
- 1-й каскад удаляет слабые прямоугольные (линейные) признаки классификации:
 - 1. Для каждого перемещения сканирующего окна вычисляется прямоугольные признаки.
 - 2. Выбирается наиболее подходящий порог для каждого признака.
 - 3. Отбираются лучшие признаки и лучший порог для каждого из них.
 - 4. Для каждого лица вычисляется его вес.

3. Распознавание лиц

- Далее к элементам применяются следующие каскады.
- 1. Для весов > 0.5 (?) применяется второй, более сложный каскад и т. д.
- 2. Отрицательный результат классификатора на любом этапе приводит к отбрасыванию признака, отбрасываются и окна тех размеров, которые дают слабую классификацию.
- 3. Классификаторы должны быть выбраны так, **чтобы ошибок становилось меньше.**
- Прямоугольные особенности не всегда достаточно хорошо могут описать характерные признаки.
- Можно строить их **линейные комбинации**, но это увеличивает время обработки.
- В статье Джонса 2004г в каскадах использовалось **4297**¹⁸

3. Распознавание лиц

- В задачах распознавания в зависимости от постановки задачи учитываются ошибки 1-го и 2-го родов.
- В реализациях алгоритма Виолы-Джонса применяются от 5 до 50 каскадов бустинга.
- Вместо каскадов признаков можно применять **дерево решений**. Это бинарное дерево, то есть из каждого узла исходит две дуги, одна с ответом «Нет», вторая с ответом «Да» - ответ касается признака



- Задача распознавания лиц обычно труднее, чем выделения лиц.

3. Распознавание лиц

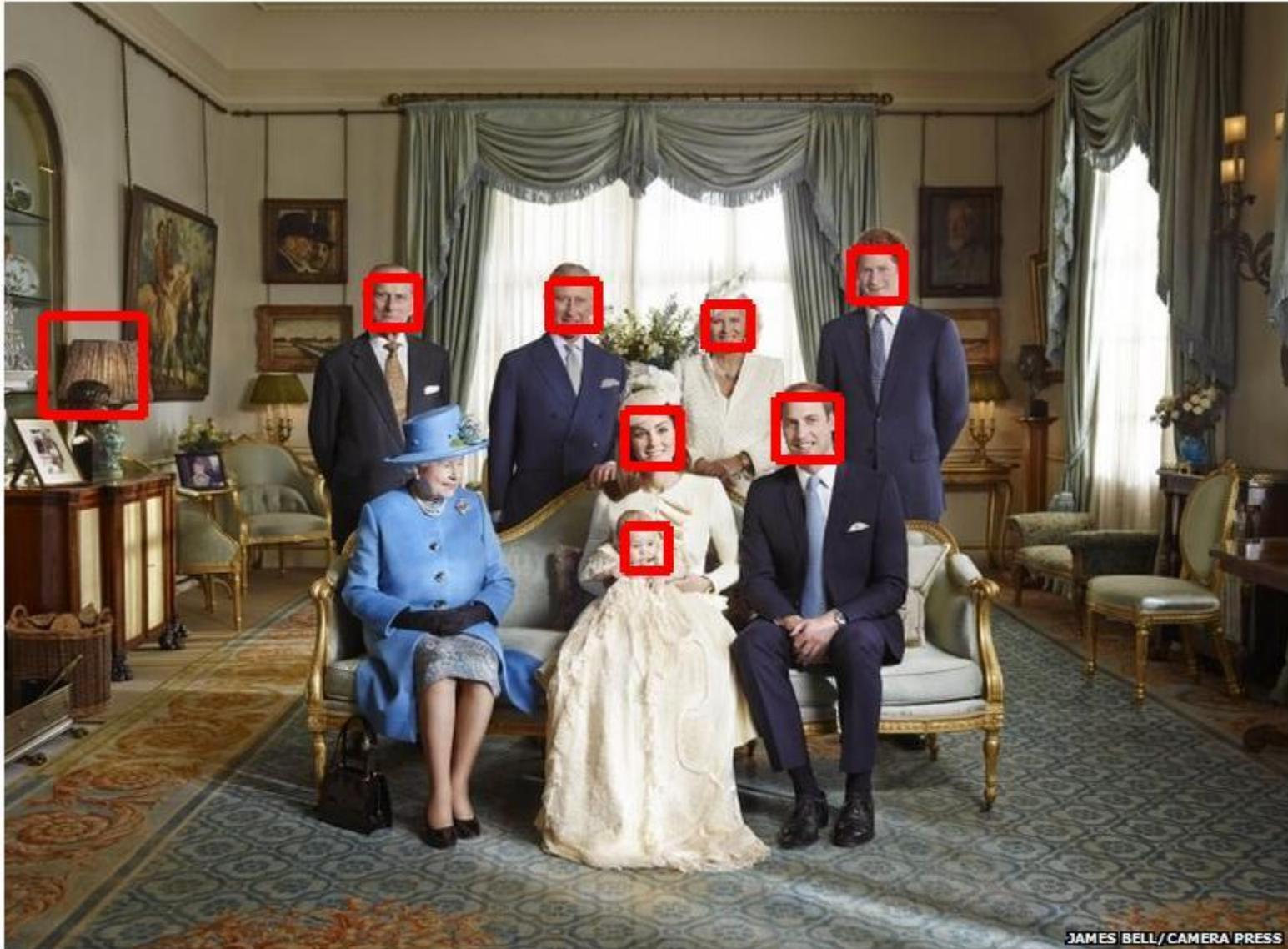
• Детекция лиц на фотографии. Исходное фото



3. Распознавание лиц

- Детекция лиц на фотографии. Есть ошибки.

Face Detection



3. Распознавание лиц

• Детекция носа. Много ошибок

Nose Detection



3. Распознавание лиц

• Детекция рта. Только ошибки

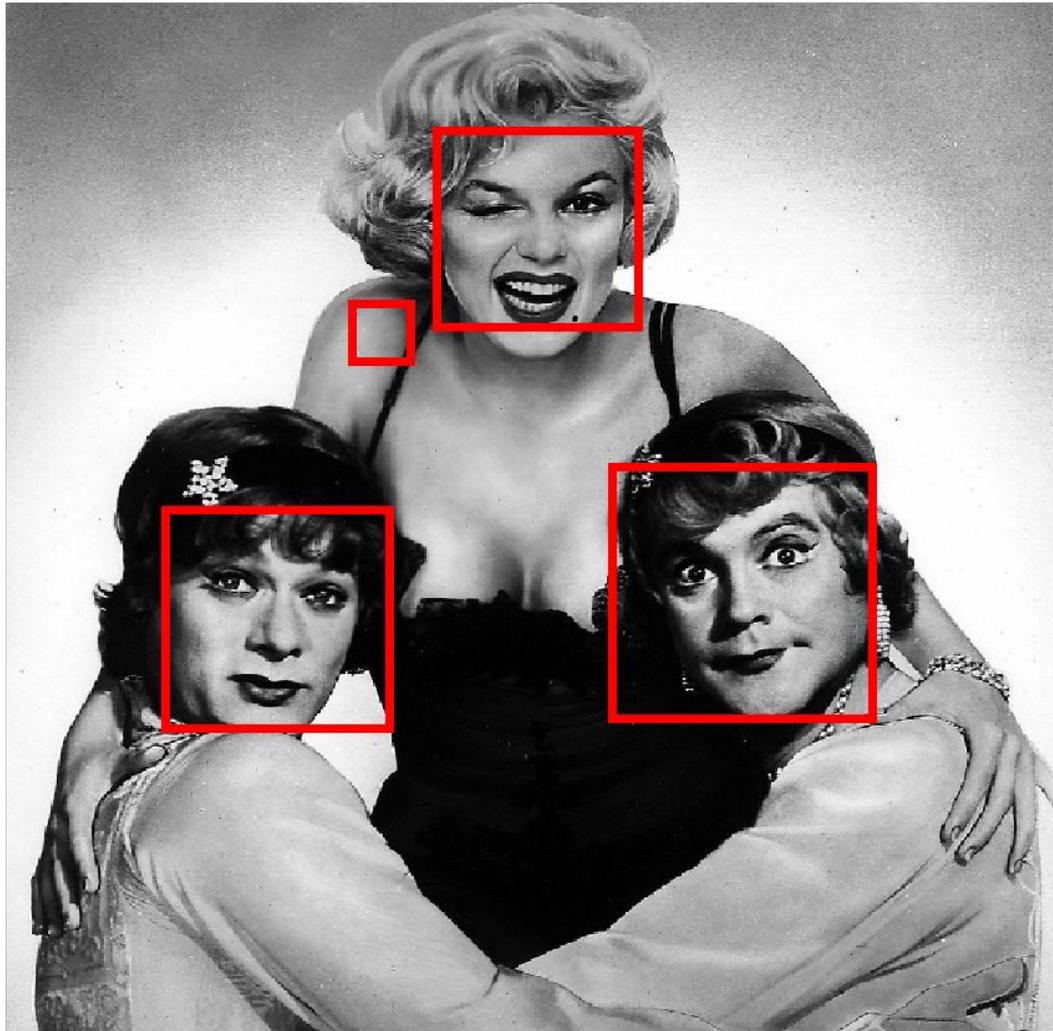
Mouth Detection



3. Распознавание лиц

- Детекция лиц. Ошибок нет.

Face Detection



3. Распознавание лиц

• Детекция носа. Ошибок нет.

Nose Detection



3. Распознавание лиц

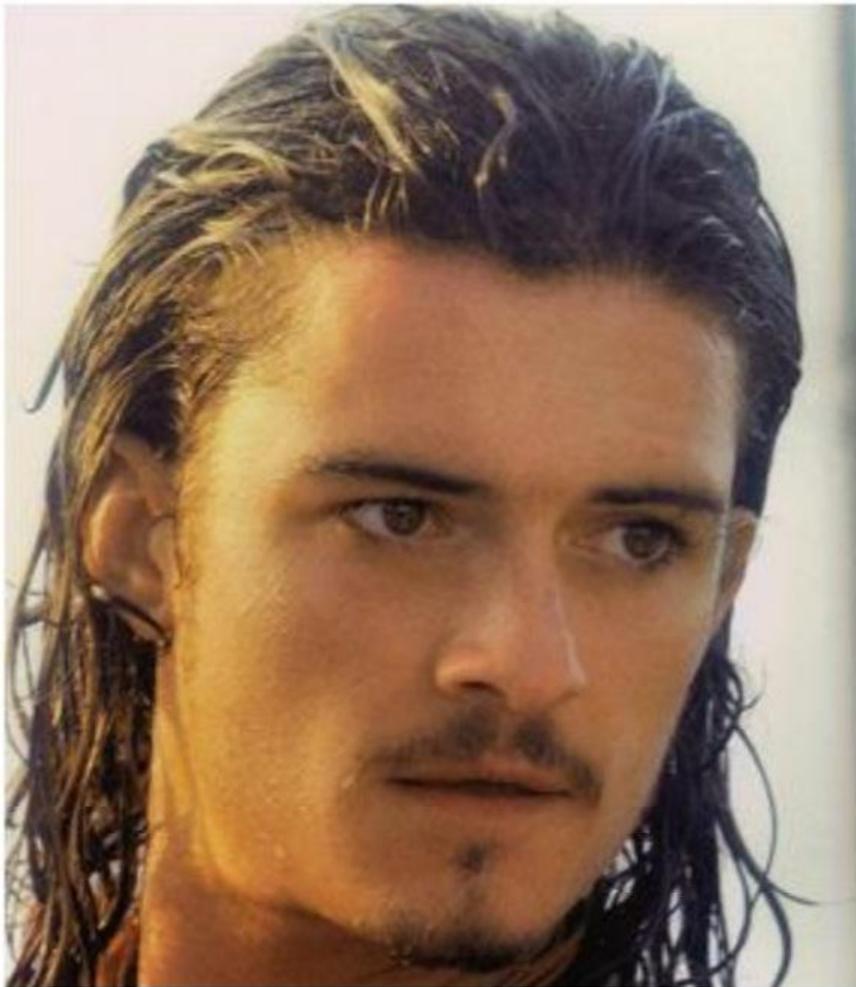
• Детекция глаз и рта. Ошибок нет.

Mouth Detection



3. Распознавание лиц

• Проблемы распознавания. Это один и тот же человек?



3. Распознавание лиц

- Задача распознавания личности, изображенной на фотографии по фотографиям из имеющейся базы данных – одна из возможных постановок задачи распознавания лица.
- На этапе предобработки производится выделение лица на изображении.
- На следующем этапе производится выравнивание изображения лица (геометрическое и яркостное),
- вычисление признаков
- и затем непосредственно распознавание:
 - – сравнение вычисленных признаков с заложенными в базу данных эталонами.
- Отличие всех методов такого рода состоит различии учитываемых признаков и различии способов их

3. Распознавание лиц

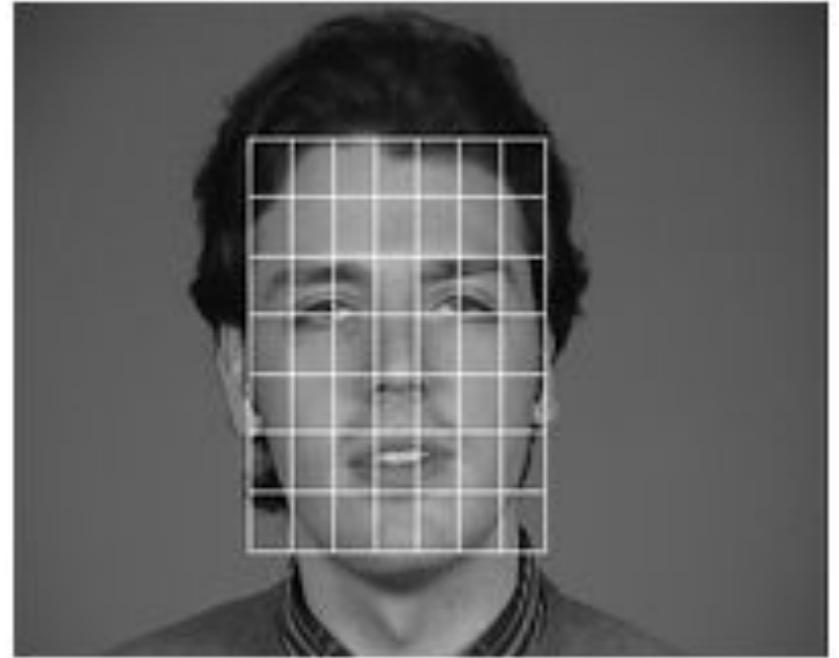
- Задача распознавания личности, изображенной на фотографии по фотографиям из имеющейся базы данных – одна из возможных постановок задачи распознавания лица.
- На этапе предобработки производится выделение лица на изображении.
- На следующем этапе производится выравнивание изображения лица (геометрическое и яркостное),
- вычисление признаков
- и затем непосредственно распознавание:
 - - сравнение вычисленных признаков с заложенными в базу данных эталонами.
- Отличие всех методов такого рода состоит различии учитываемых признаков и различии способов их

3. Распознавание лиц

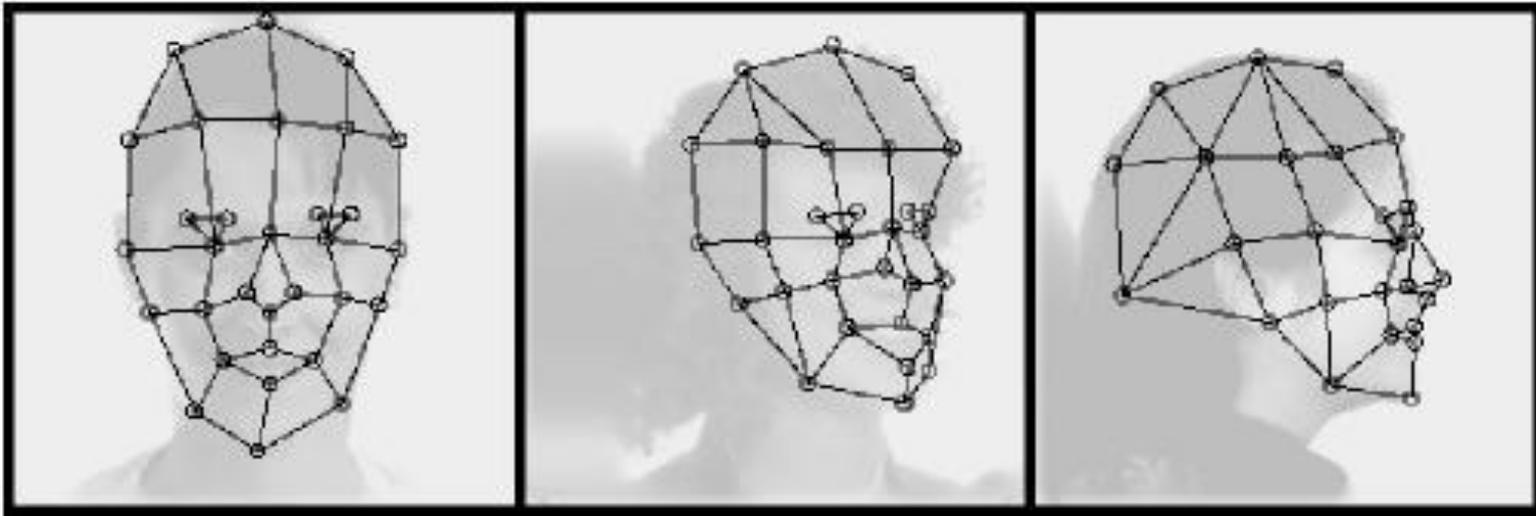
- Существует много методов распознавания лиц.
- **1. Метод гибкого сравнения на графах.**
- Метод эластично сопоставляет два графа, описывающих лицо на фото и лица из базы. Если графы «близки», то делается заключение об их идентичности.
- Черты лица представляются графом со взвешенными вершинами и взвешенными ребрами. Чем больше значимость маркера или расстояния между маркерами, тем больше их вес.
- В процессе работы алгоритма один из графов (эталонный) не изменяется, а другой деформируется для наилучшей подгонки к эталонному. В таких системах распознавания графы могут представлять собой как прямоугольную решетку, наложенную на лицо, так и структуру, образованную

3. Распознавание лиц

- Прямоугольная решетка



- Антропометрические маркеры лица (3D).

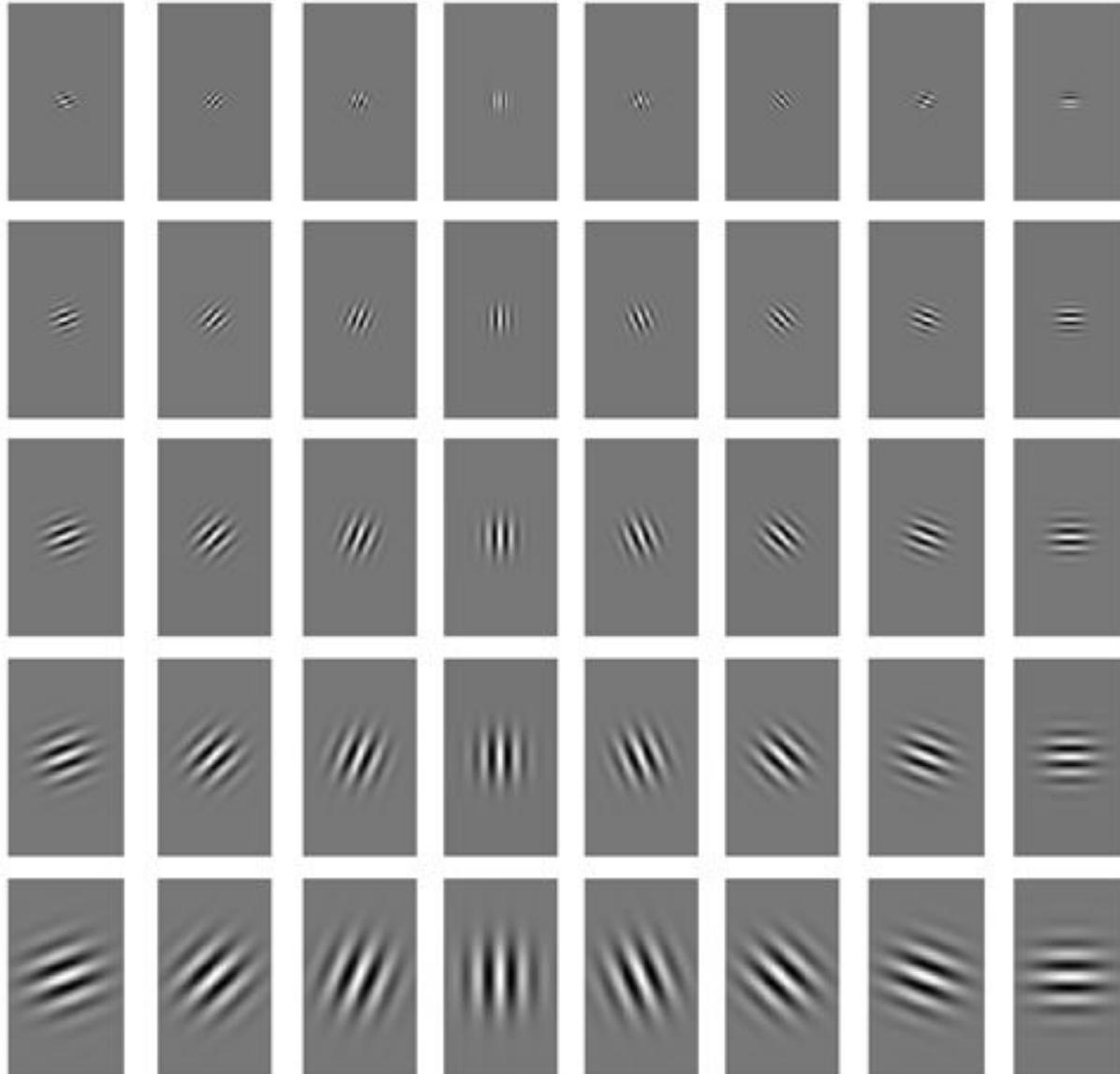


3. Распознавание лиц

- В узлах графа вычисляются значения признаков, это могут быть фильтры Габора однократно примененные к некоторой локальной области (такого же размера) узла графа. В результате получается упорядоченный набор комплексных изображений (применяется фильтр Габора, содержащий комплексные числа).
- Могут применяться вейвлеты Габора.

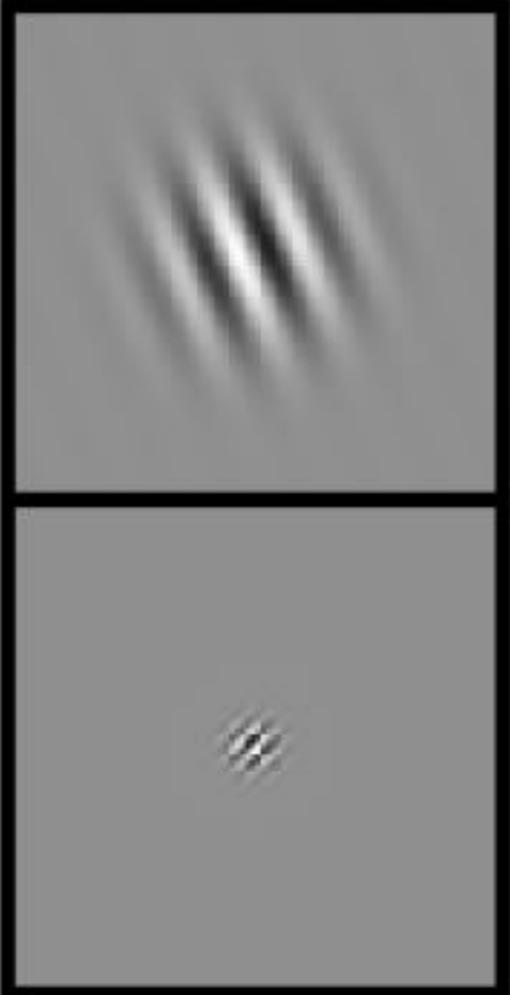
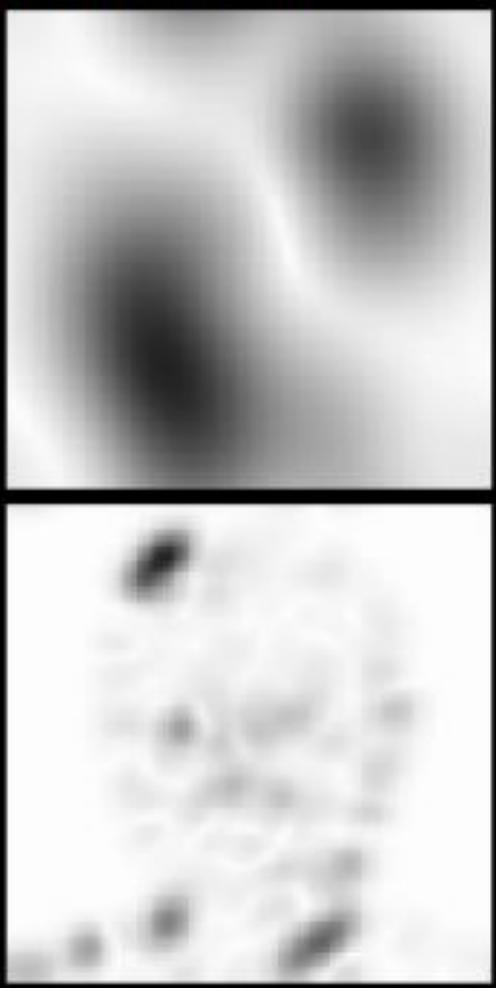
3. Распознавание лиц

• Набор фильтров Габора (маски Габора)



3. Распознавание лиц

• Пример применения двух фильтров Габора

Входное изображение	Пример двух фильтров Габора	Результат свертки входного изображения лица и фильтров Габора
		

3. Распознавание лиц

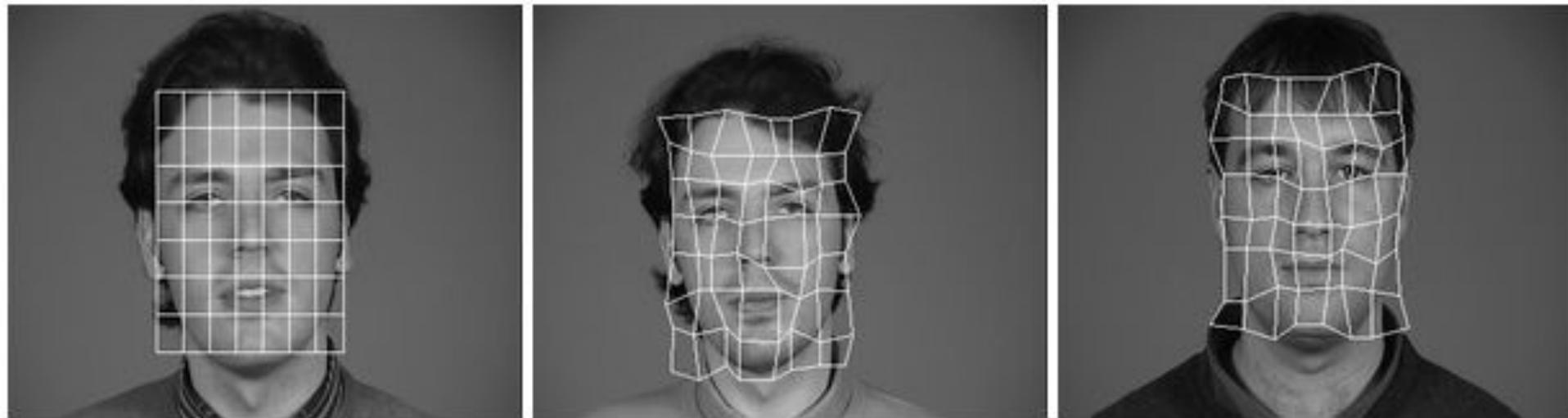
- При сравнении с эталоном решетка графа накладывается на эталон. В графе изменяются положения вершин и длины ребер. **Различие между двумя графами** (эталонным и на фото) вычисляется при помощи некоторой функции деформации, учитывающей результат, полученный сверткой с фильтром Габора различие между значениями признаков, в окрестности вершин и изменением длины ребер графа.
- При этом для уменьшения близости двух соответствующих вершин, вершина на фото может смещаться. Процедурой корректировки всех вершин достигается значение близости графов, **близкое к минимальному**.
- Процедура сравнения лица на фото с эталонами выполняется для **всех лиц из эталонного множества** и выбирается лицо с наименьшей мерой близости. Это и есть **решение**³⁵

3. Распознавание лиц

- При сравнении с эталоном решетка графа накладывается на эталон. В графе изменяются положения вершин и длины ребер. **Различие между двумя графами** (эталонным и на фото) вычисляется при помощи некоторой функции деформации, учитывающей результат, полученный сверткой с фильтром Габора различие между значениями признаков, в окрестности вершин и изменением длины ребер графа.
- При этом для уменьшения близости двух соответствующих вершин, вершина на фото может смещаться. Процедурой корректировки всех вершин достигается значение близости графов, **близкое к минимальному**.
- Процедура сравнения лица на фото с эталонами выполняется для **всех лиц из эталонного множества** и выбирается лицо с наименьшей мерой близости. Это и есть **решение**³⁶

3. Распознавание лиц

- Деформация графа при сравнении лиц



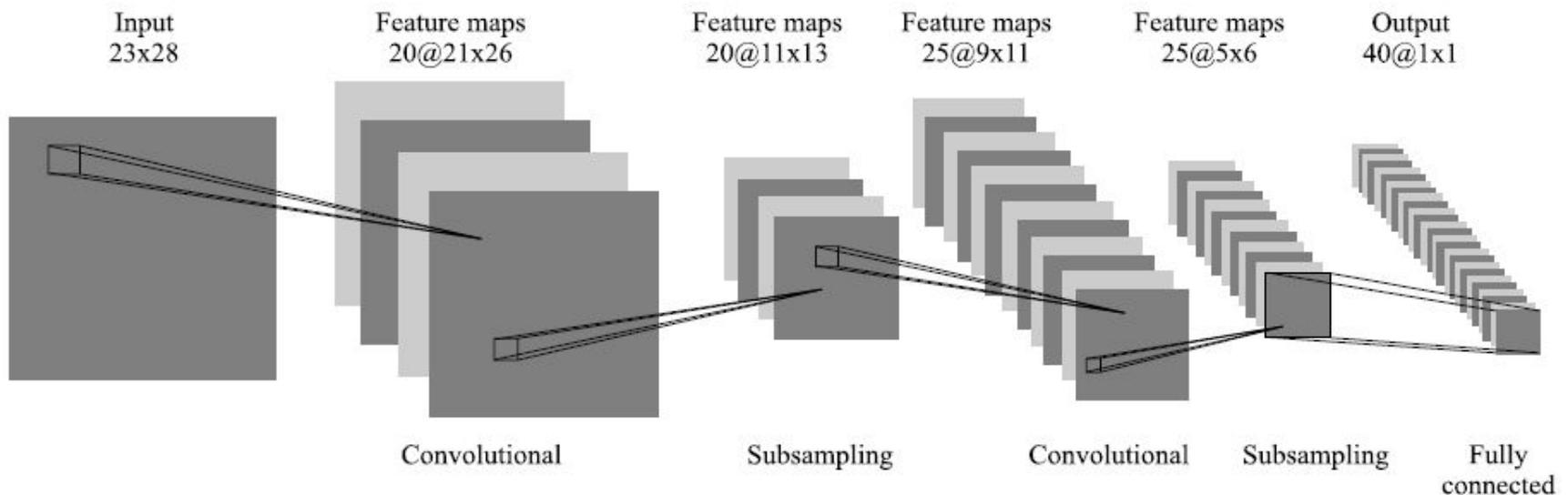
- Метод гибкого сравнения на графах требует значительных вычислительных ресурсов и время его работы линейно зависит от объема эталонного множества.

3. Распознавание лиц

- 2. Метод нейронных сетей в обработке изображений.
- При использовании нейронных сетей для машинного обучения первым шагом является обучение сети на обучающей выборке. Цель обучения сети состоит в поиске весов связей между узлами (нейронами) сети.
- Фактически в процессе обучения начально заданные веса корректируются методом градиентного спуска.
- В процессе обучения сети из обучающей выборки автоматически извлекаются ключевые признаки объектов, определяется их вес (важность) и устанавливается взаимосвязь между ними.
- При этом ожидается, что таким образом построенная

3. Распознавание лиц

- На практике оказалось, что наилучшие результаты по распознаванию лиц дает сверточная нейронная сеть.
- Такой результат объясняется учетом двумерной топологии лица на фотографии, то есть, принимается во внимание пиксель и соседние пиксели его двумерной окрестности.
- Кроме учета двумерной топологии, сверточная сеть использует иерархические разложения рисунка с **пространственным сжатием**

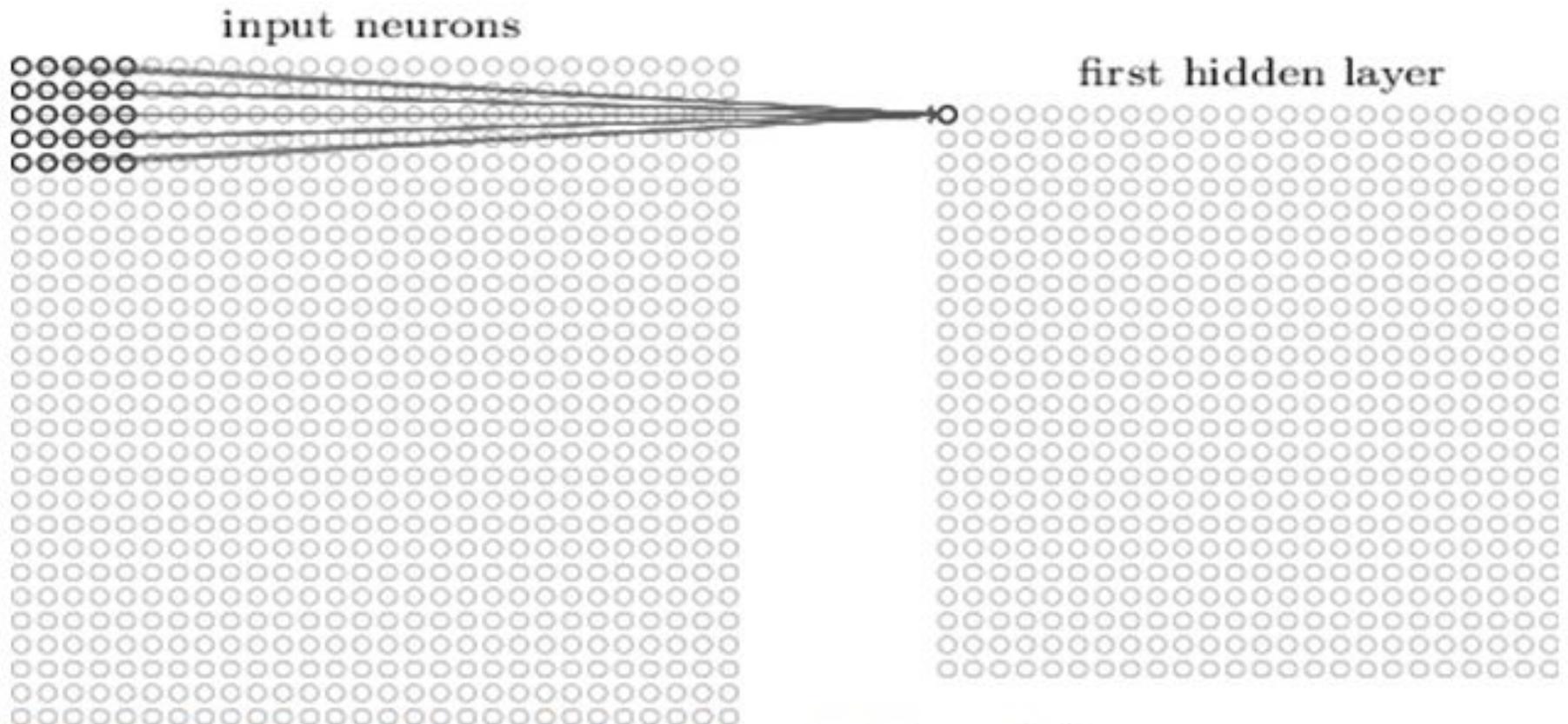


3. Распознавание лиц

- На практике оказалось, что наилучшие результаты по распознаванию лиц дает сверточная нейронная сеть.
- Такой результат объясняется учетом двумерной топологии лица на фотографии, то есть, принимается во внимание пиксель и соседние пиксели его двумерной окрестности.
- Кроме учета двумерной топологии сверточная сеть использует иерархические разложения рисунка с пространственными сэмплингом. Благодаря этому сверточная сеть обеспечивает частичную устойчивость к изменениям масштаба, смещениям, поворотам, смене ракурса и др. искажениям.

3. Распознавание лиц

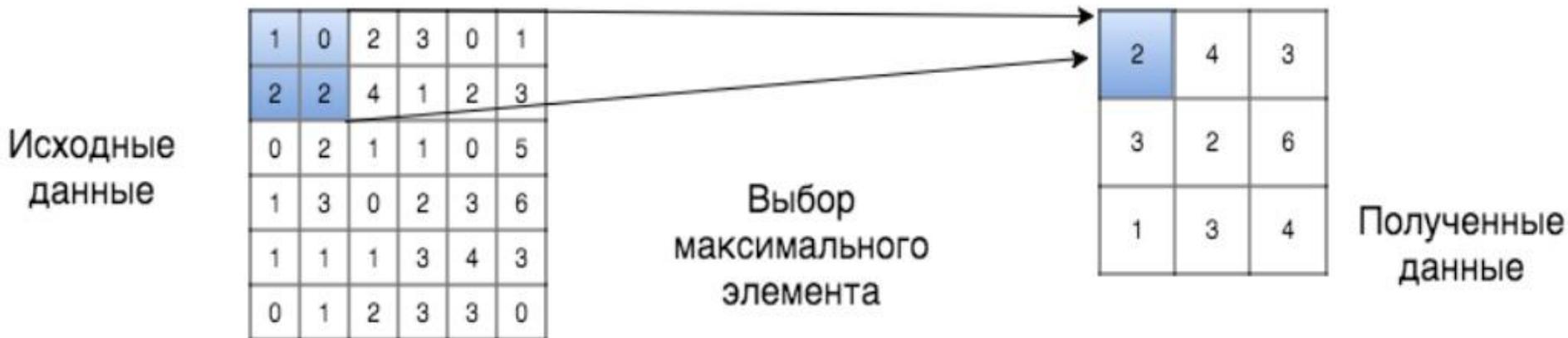
- Сверточная сеть операцией свертки (sampling - сэмплинг) уменьшает размер рисунка за счет граничных пикселей. Пиксели на границах, которые не меняют свою интенсивность при свертке, просто удаляются. Пример



Visualization of 5 x 5 filter convolving around an input volume and producing an activation map

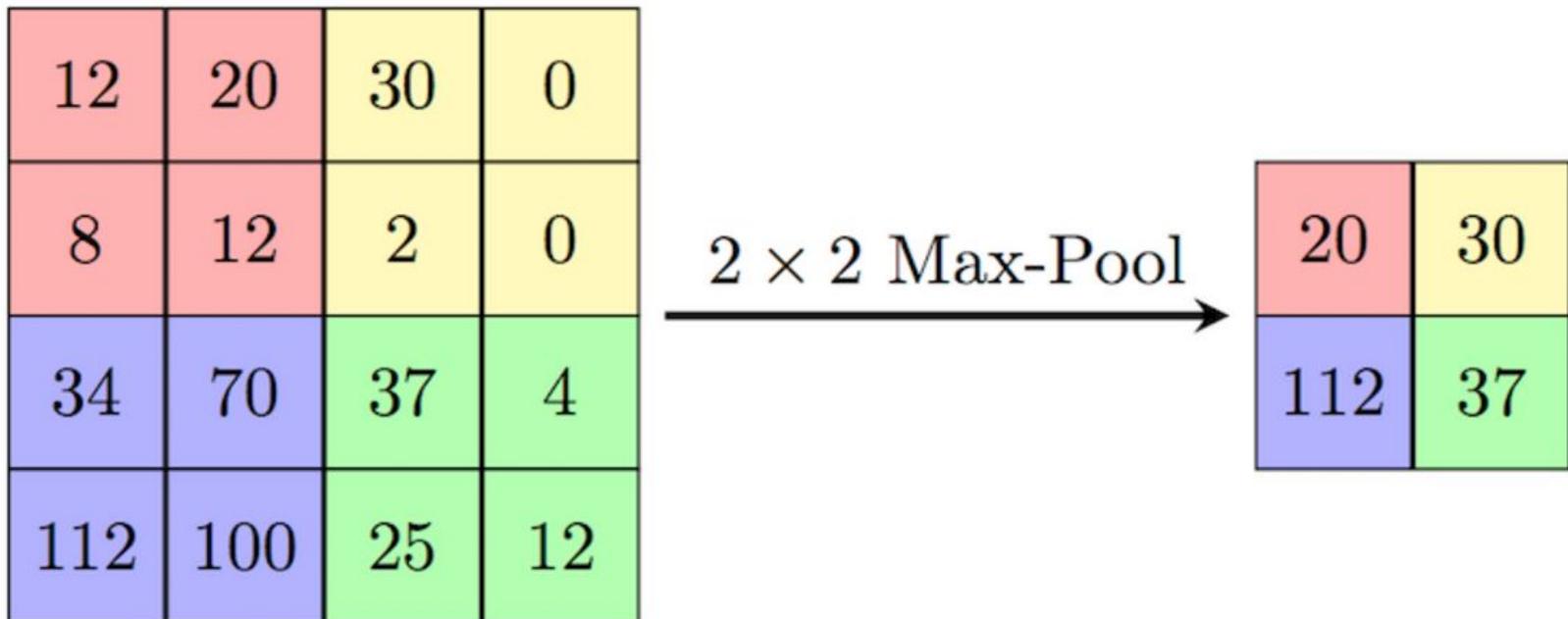
3. Распознавание лиц

- То есть, маска свертки проходит весь рисунок из верхнего левого угла в правый нижний угол с шагом в 1 пиксель. Маска идет по горизонтали, фильтруя заметаемые квадраты рисунка, до правого края. Затем она смещается на 1 пиксель вниз и снова проходит весь рисунок по горизонтали.
- В результате получаем измененный рисунок меньшего размера за счет отброшенных граничных пикселей, которые не изменили своей яркости.
- Следующий этап является операция субпикселизации



3. Распознавание лиц

- При субдискретизации (пулинге) маска размерности 2×2 пробегает по пиксели рисунка с шагом 2, создавая новый рисунок в 2 раза меньшей размерности, заменяя квадрат 2×2 его максимальным или средним значением его пикселей.
- Соседние пиксели рисунка отличаются незначительно, то есть такая операция изменяет рисунок и уменьшает его в 2 раза



3. Распознавание лиц

- Для сохранения всей информации с преобразованием сэмплинга на **выходе** пришлось бы строить **4 матрицы**: одну для левых верхних пикселей квадрата, вторую для правого верхнего пикселя квадрата, и т.д.
- Благодаря этому сверточная сеть обеспечивает частичную устойчивость к изменениям масштаба, смещениям, поворотам, смене ракурса и др. искажениям.

3. Распознавание лиц

- Ошибки в распознавании лиц по базе данных фотографий личностей.
- Компьютеры способны решать все более сложные задачи, но, тем не менее, они не являются универсальным средством.
- Человек часто не может узнать знакомого, изображенного на фотографии. Трудно ожидать от компьютера **надежной идентификации по фотографии**.
-
- Человеческий мозг приспособлен к распознаванию лиц, к примеру, маленькие дети запоминают лица гораздо лучше, чем любые другие формы.
- Помимо всего прочего, человеческий мозг обладает гораздо большей, чем компьютеры, способностью компенсации изменений освещенности и угла зрения. Лица имеют⁴⁵

3. Распознавание лиц

- Нередко и человеку, и компьютеру бывает **невозможно сопоставлять изображения** при наличии разницы в освещении, угла обзора камеры, не говоря уже изменения внешнего вида самого лица.
- Исследования программных средств распознавания по лицу установили **высокий % ложного распознавания** людей и идентификации их с фотографиями других лиц, находящихся в базе данных. Программные средства зачастую **неспособны распознать преступников по фото** в базе данных.
- Проблема заключается еще и в том, что в отличие от отпечатков пальцев или радужной оболочки, наши лица **меняются с течением времени**. Системы распознавания легко ошибаются из-за **изменения прически, растительности на лице или веса тела, из-за применения**

3. Распознавание лиц

- К примеру, исследование установило, что уровень ложной идентификации или нераспознавания субъектов, чьи фотографии были сделаны всего **18 месяцев назад, равен 43%**. При этом фотографии, использованные в исследовании, были отсняты в идеальных условиях, что весьма важно, так как программы распознавания по лицу очень плохо справляются с оценкой изменения освещенности или угла наклона камеры.
- Трудны для них и **фотографии с оживленным фоном**. Исследование пришло к заключению, что изменение угла **наклона камеры на 45 градусов** делает такие программы практически бесполезными.
-
- Наилучшим образом технология распознавания лица срабатывает в строго контролируемых условиях, когда

3. Распознавание лиц

- Другое исследование, проведенное Мин обороны США, обнаружило высокий уровень ложной идентификации даже при идеальных условиях. От снимков с зернистым изображением или старых фотографий, типа тех, которые хранятся в личных делах, очень мало пользы.
- Кроме этого, программы не всегда срабатывают для людей с небелым цветом кожи. Черты лица которых различаются программами, оптимизированы для съемки людей со светлой кожей.
- В распоряжении властей находится огромная база данных изображений лиц граждан - это фотографии паспортов и водительских удостоверений. Правительство имеет право использовать эти снимки для целей видеонаблюдения и идентификации.