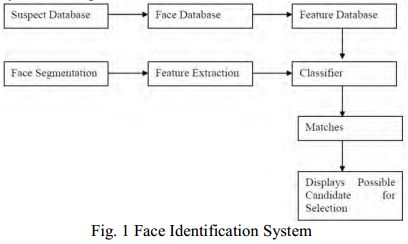
АНАЛЬТЕРНАТИВНЫЕ МЕТОДЫ РЕШЕНИЯ

Метод главных компонент

Метод главных компонент (МГК) включает математический метод, который преобразует ряд вероятностно коррелированных переменных в меньшее количество некоррелированных переменных, называемых основными компонентами. МГК является очень известным подходом, который используется для расчета набора функций для распознавания лица. Это может относиться к любому лицу:

1. Лицо может быть представлено минимальным набором характеристик.

2. Может быть построено с использование небольшой коллекции собственных фотографий. Изображение лица проецируется на некоторое количество шаблонов лиц, называемое собвственными лицами, затем вычисляется разница между этими лицами, которая будет считаться отличительной чертой для изображения. Когда набор собственных лиц вычислен, изображения лица может быть реконструировано взвешенной комбинацией собственных лиц. Веса образуют вектор признаков для представления лица и распознавания. Когда подается новое тестовое изображение, веса высчитаны благодяря проекции изображения на вектор собственных лиц. Классификация между изображениями выходит из сравнения расстояний между векторами весов тестового изображения и обучающего множества (входных данных). Как уже отмечено, можно реконструировать исходное изображение из собственных лиц так, что входное изображение должно точно соответствовать оригинальному изображению, используя все собственные лица, извлеченные из оригинальных изображений. МГК является математическим методом, который основан на преобразовании переменных, с помощью которого можно преобразовать набор коррелированных переменных в набор некоррелированных переменных. Переменные, которые некоррелируют друг с другом – основные компоненты. Количество основных компонентов означает, что некоррелированных должно быть меньше или равно количеству исходных переменных. В такого рода трансформации первые главные компоненты будут иметь самый высокий приоритет, который покажет максимальное отклонение. Это поможет вычислить точность данных. Если набор данных распределен нормально, только тогда главные компоненты могут быть независимыми от других переменных. Снижение размеров может привести к потере информации, но какое-то количество информации так или иначе останется.



Предобработка: (МГК) – на этом этапе сокращается размерность данных. Этап классификации - уменьшенные выходные векторы МГК идут на вход нейросети с обратным распространением ошибки для обучения данных, которые буду использованы для распознавания изображения. Распознавание лица – биометрическая технология используемая, к примеру, с целью наблюдения, поиска разыскиваемых преступников, пропавших без вести детей. Распознавание лица имеет три стадии:

1. обнаружение лица;
2. извлечение характеристик;
3. непосредственно классификация.

АЛГОРИТМЫ И ПРЕДЛОЖЕННЫЙ МЕТОД

Предобработка: (МГК) – на этом этапе сокращается размерность данных. Этап классификации - уменьшенные выходные векторы МГК идут на вход нейросети с обратным распространением ошибки для обучения данных, которые буду использованы для распознавания изображения. Распознавание лица – биометрическая технология используемая, к примеру, с целью наблюдения, поиска разыскиваемых преступников, пропавших без вести детей. Распознавание лица имеет три стадии:

1. обнаружение лица;
2. извлечение характеристик;
3. непосредственно классификация.

МГК-АЛГОРИТМ

Пусть имеется такой набор лиц для обучения:

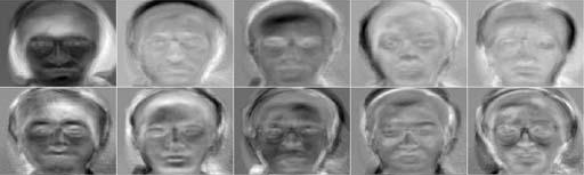
Среднее лицо по набору может быть определено:

Каждое лицо отличается от среднего вектором признаков:

Матрица ковариаций строится благодаря вектору:

Где матрица Множество больших векторов затем подвергают МГК . Получаем вектор весов изображения лица. Лицо проецируется на пространство лиц:

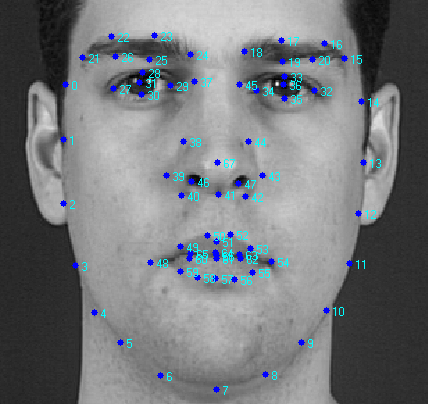
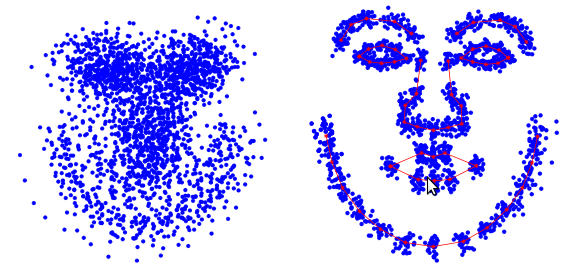
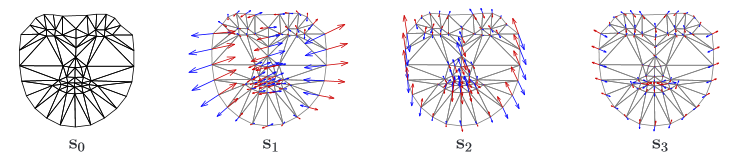
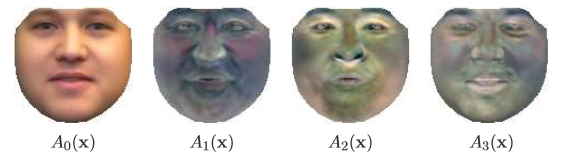
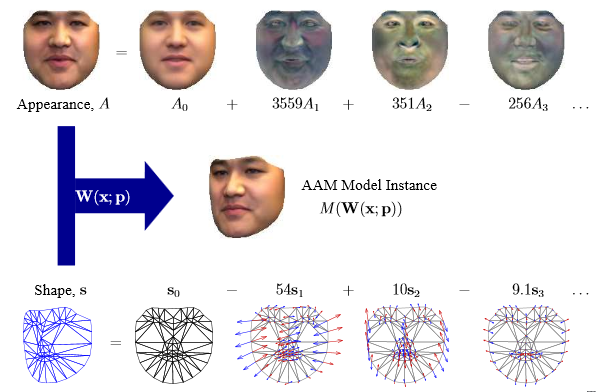
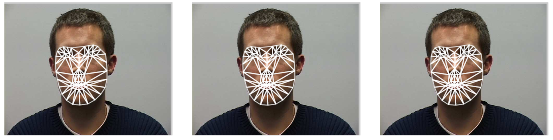
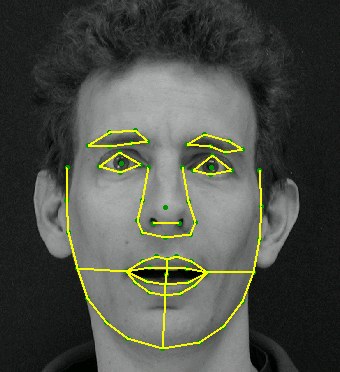
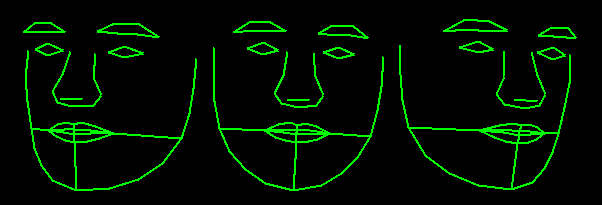
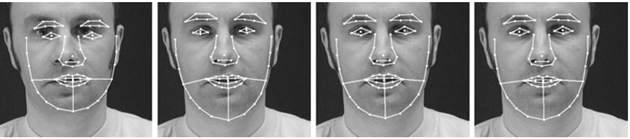
Для k = 1, M', где M'<= M это номер собственного лица, используемого для распознавания. Веса формируют вектор .

  
*Пример первых десяти собственных векторов (собственных лиц), полученных на обучаемом наборе лиц*

##### **Метод гибкого сравнения на графах (Elastic graph matching)** Суть метода сводится к эластичному сопоставлению графов, описывающих изображения лиц. Лица представлены в виде графов со взвешенными вершинами и ребрами. На этапе распознавания один из графов – эталонный – остается неизменным, в то время как другой деформируется с целью наилучшей подгонки к первому. В подобных системах распознавания графы могут представлять собой как прямоугольную решетку, так и структуру, образованную характерными (антропометрическими) точками лица. а)https://habrastorage.org/files/afe/bd2/804/afebd280435d4f7ba308d3b37dbb2623.png б)https://habrastorage.org/files/bc4/c72/440/bc4c724406d34056994b26d78926e1eb.jpg *Пример структуры графа для распознавания лиц: а) регулярная решетка б) граф на основе антропометрических точек лица.* В вершинах графа вычисляются значения признаков, чаще всего используют комплексные значения фильтров Габора или их упорядоченных наборов – Габоровских вейвлет (строи Габора), которые вычисляются в некоторой локальной области вершины графа локально путем свертки значений яркости пикселей с фильтрами Габора. https://habrastorage.org/files/4fd/5d1/826/4fd5d18265de4e2298c3847ef24724e8.png  *Набор (банк, jet) фильтров Габора* https://habrastorage.org/files/557/0ac/b02/5570acb0233d4310a44ce3e8c6263040.png*Пример свертки изображения лица с двумя фильтрами Габора* Ребра графа взвешиваются расстояниями между смежными вершинами. Различие (расстояние, дискриминационная характеристика) между двумя графами вычисляется при помощи некоторой ценовой функции деформации, учитывающей как различие между значениями признаков, вычисленными в вершинах, так и степень деформации ребер графа. Деформация графа происходит путем смещения каждой из его вершин на некоторое расстояние в определённых направлениях относительно ее исходного местоположения и выбора такой ее позиции, при которой разница между значениями признаков (откликов фильтров Габора) в вершине деформируемого графа и соответствующей ей вершине эталонного графа будет минимальной. Данная операция выполняется поочередно для всех вершин графа до тех пор, пока не будет достигнуто наименьшее суммарное различие между признаками деформируемого и эталонного графов. Значение ценовой функции деформации при таком положении деформируемого графа и будет являться мерой различия между входным изображением лица и эталонным графом. Данная «релаксационная» процедура деформации должна выполняться для всех эталонных лиц, заложенных в базу данных системы. Результат распознавания системы – эталон с наилучшим значением ценовой функции деформации.  https://habrastorage.org/files/306/381/7e7/3063817e70834273add4ee5c99d93d8b.png  *Пример деформации графа в виде регулярной решетки* В отдельных публикациях указывается 95-97%-ая эффективность распознавания даже при наличии различных эмоциональных выражениях и изменении ракурса лица до 15 градусов. Однако разработчики систем эластичного сравнения на графах ссылаются на высокую вычислительную стоимость данного подхода. Например, для сравнения входного изображения лица с 87 эталонными тратилось приблизительно 25 секунд при работе на параллельной ЭВМ с 23 транспьютерами [15] (Примечание: публикация датирована 1993 годом). В других публикациях по данной тематике время либо не указывается, либо говорится, что оно велико.  **Недостатки:** высокая вычислительная сложность процедуры распознавания. Низкая технологичность при запоминании новых эталонов. Линейная зависимость времени работы от размера базы данных лиц.

##### **3. Скрытые Марковские модели (СММ, HMM)** Одним из статистических методов распознавания лиц являются скрытые Марковские модели (СММ) с дискретным временем [32-34]. СММ используют статистические свойства сигналов и учитывают непосредственно их пространственные характеристики. Элементами модели являются: множество скрытых состояний, множество наблюдаемых состояний, матрица переходных вероятностей, начальная вероятность состояний. Каждому соответствует своя Марковская модель. При распознавании объекта проверяются сгенерированные для заданной базы объектов Марковские модели и ищется максимальная из наблюдаемых вероятность того, что последовательность наблюдений для данного объекта сгенерирована соответствующей моделью. На сегодняшний день не удалось найти примера коммерческого применения СММ для распознавания лиц.  **Недостатки:**  — необходимо подбирать параметры модели для каждой базы данных; — СММ не обладает различающей способностью, то есть алгоритм обучения только максимизирует отклик каждого изображения на свою модель, но не минимизирует отклик на другие модели.

##### **5. Active Appearance Models (AAM) и Active Shape Models (ASM) (Хабраисточник)**

**Active Appearance Models (AAM)**  
Активные модели внешнего вида (Active Appearance Models, AAM) — это статистические модели изображений, которые путем разного рода деформаций могут быть подогнаны под реальное изображение. Данный тип моделей в двумерном варианте был предложен Тимом Кутсом и Крисом Тейлором в 1998 году [17,18]. Первоначально активные модели внешнего вида применялись для оценки параметров изображений лиц.  
Активная модель внешнего вида содержит два типа параметров: параметры, связанные с формой (параметры формы), и параметры, связанные со статистической моделью пикселей изображения или текстурой (параметры внешнего вида). Перед использованием модель должна быть обучена на множестве заранее размеченных изображений. Разметка изображений производится вручную. Каждая метка имеет свой номер и определяет характерную точку, которую должна будет находить модель во время адаптации к новому изображению.  
  
   
*Пример разметки изображения лица из 68 точек, образующих форму AAM.*  
  
Процедура обучения AAM начинается с нормализации форм на размеченных изображениях с целью компенсации различий в масштабе, наклоне и смещении. Для этого используется так называемый обобщенный Прокрустов анализ.   
  
   
*Координаты точек формы лица до и после нормализации*  
  
Из всего множества нормированных точек затем выделяются главные компоненты с использованием метода PCA.   
  
   
*Модель формы AAM состоит из триангуляционной решетки s0 и линейной комбинации смещений si относительно s0*  
  
Далее из пикселей внутри треугольников, образуемых точками формы, формируется матрица, такая что, каждый ее столбец содержит значения пикселей соответствующей текстуры. Стоит отметить, что используемые для обучения текстуры могут быть как одноканальными (градации серого), так и многоканальными (например, пространство цветов RGB или другое). В случае многоканальных текстур векторы пикселов формируются отдельно по каждому из каналов, а потом выполняется их конкатенация. После нахождения главных компонент матрицы текстур модель AAM считается обученной.   
  
  
  
Модель внешнего вида AAM состоит из базового вида A0, определенного пикселями внутри базовой решетки s0 и линейной комбинации смещений Ai относительно A0  
  
  
  
Пример конкретизации AAM. Вектор параметров формы  
p=(p\_1,p\_2,〖…,p〗\_m )^T=(-54,10,-9.1,…)^T используется для синтеза модели формы s, а вектор параметров λ=(λ\_1,λ\_2,〖…,λ〗\_m )^T=(3559,351,-256,…)^Tдля синтеза внешнего вида модели. Итоговая модель лица 〖M(W(x;p))〗^ получается как комбинация двух моделей – формы и внешнего вида.  
  
Подгонка модели под конкретное изображение лица выполняется в процессе решения оптимизационной задачи, суть которой сводится к минимизации функционала  
  
https://habrastorage.org/files/ac4/33b/41e/ac433b41e1bd4ad3af3700fd967e662f.png  
  
методом градиентного спуска. Найденные при этом параметры модели и будут отражать положение модели на конкретном изображении.   
  
  
   
*Пример подгонки модели на конкретное изображение за 20 итераций процедуры градиентного спуска.*  
  
С помощью AAM можно моделировать изображения объектов, подверженных как жесткой, так и нежесткой деформации. ААМ состоит из набора параметров, часть которых представляют форму лица, остальные задают его текстуру. Под деформации обычно понимают геометрическое преобразование в виде композиции переноса, поворота и масштабирования. При решении задачи локализации лица на изображении выполняется поиск параметров (расположение, форма, текстура) ААМ, которые представляют синтезируемое изображение, наиболее близкое к наблюдаемому. По степени близости AAM подгоняемому изображению принимается решение – есть лицо или нет.   
  
***Active Shape Models (ASM)***  
  
Суть метода ASM [16,19,20] заключается в учете статистических связей между расположением антропометрических точек. На имеющейся выборке изображений лиц, снятых в анфас. На изображении эксперт размечает расположение антропометрических точек. На каждом изображении точки пронумерованы в одинаковом порядке.   
  
  
  
*Пример представления формы лица с использованием 68 точек*  
  
Для того чтобы привести координаты на всех изображениях к единой системе обычно выполняется т.н. обобщенный прокрустов анализ, в результате которого все точки приводятся к одному масштабу и центрируются. Далее для всего набора образов вычисляется средняя форма и матрица ковариации. На основе матрицы ковариации вычисляются собственные вектора, которые затем сортируются в порядке убывания соответствующих им собственных значений. Модель ASM определяется матрицей Φ и вектором средней формы s ̅.   
Тогда любая форма может быть описана с помощью модели и параметров:  
  
https://habrastorage.org/files/8a1/b92/6a1/8a1b926a18a842378dbd46ef8fd3d04c.png  
  
Локализации ASM модели на новом, не входящем в обучающую выборку изображении осуществляется в процессе решения оптимизационной задачи.   
  
   
а) б) в) г)  
*Иллюстрация процесса локализации модели ASM на конкретном изображении: а) начальное положение б) после 5 итераций в) после 10 итераций г) модель сошлась*  
  
Однако все же главной целью AAM и ASM является не распознавание лиц, а точная локализация лица и антропометрических точек на изображении для дальнейшей обработки.   
  
Практически во всех алгоритмах обязательным этапом, предваряющим классификацию, является выравнивание, под которым понимается выравнивание изображения лица во фронтальное положение относительно камеры или приведение совокупности лиц (например, в обучающей выборке для обучения классификатора) к единой системе координат. Для реализации этого этапа необходима локализация на изображении характерных для всех лиц антропометрических точек – чаще всего это центры зрачков или уголки глаз. Разные исследователи выделяют разные группы таких точек. В целях сокращения вычислительных затрат для систем реального времени разработчики выделяют не более 10 таких точек [1].  
  
Модели AAM и ASM как раз и предназначены для того чтобы точно локализовать эти антропометрические точки на изображении лица.