

# Обзор алгоритмов распознавания лиц

М. А. Сметанин

Владимирский государственный университет, г. Владимир, Россия

**Аннотация:** актуальность данной темы обусловлена тем, что в современном мире технология распознавания лиц используется во многих приложениях, которые вошли в обиход человека: проверка личности, интернет общение и компьютерные развлечения. Целью данной статьи ставится сравнение существующих на данный момент подходов к распознаванию лиц. Задача данной статьи состоит в обзоре разных подходов с указанием их достоинств и недостатков.

**Ключевые слова:** распознавание лиц, классификация, эластичные графы, обработка изображения лица.

Распознавание лиц - проблема визуального распознавания образов. Лицо, как трёхмерный объект, на который влияет освещение, поза и выражение лица должно быть идентифицировано на основе его двухмерного изображения. Общая процесс распознавания лиц включает в себя четыре модуля, как это показано на рисунке 1: обнаружение лица, его выравнивание, выделение признаков и сопоставление, где локализация и нормализация представляют собой этапы обработки перед распознавание лица. [1]

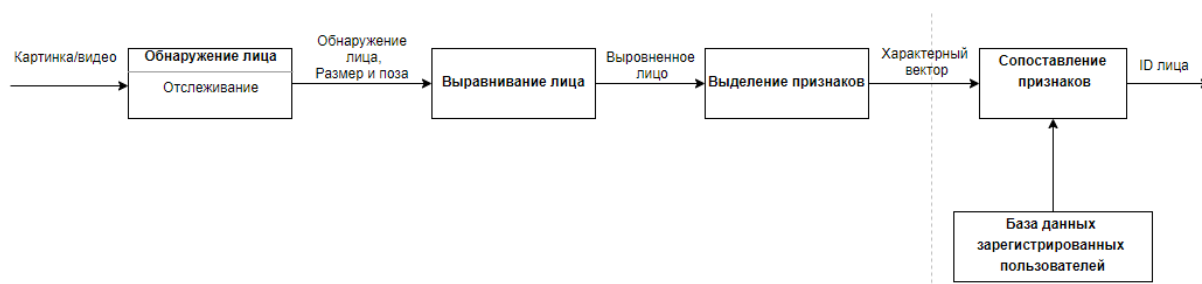


Рисунок 1. Общий процесс обработки изображения лица при распознавании.

На первом этапе происходит отделение лица от фона. Второй этап заключается в том, что происходит более детальная локализация и нормализация лица, так как первый этап обеспечивает лишь грубые оценки положения и масштаба лица. После того, как на предыдущем этапе грани лица были нормализованы геометрически и фотометрически, выполняется выделение признаков, чтобы дать информацию, которая нужна для различения лиц разных людей и стабильна относительно геометрических и фотометрических изменений.

### **Метод гибкого сравнения на графах (Elastic Bunch Graph Matching) [2]**

Метод впервые упоминается в статье “IEEE transactions on computers, vol 1993”. Эту статью можно считать “рождением” данного способа. В дальнейшем метод только расширялся и модернизировался, как и информационные решения, которые в нём использовались. В статье раскрывался способ создания программы распознавания лиц, которая основывается на представлении гибких графов: вершины отмечаются крупномасштабными по отношению к местному спектру мощности, а рёбра - как геометрические векторы расстояния. Для того чтобы вычислить значения признаков в вершинах графов используют комплексные значения фильтров Габора или их упорядоченных наборов, которые вычисляются в локальной области вершины графа путём свертки значений яркости пикселей с фильтрами Габора. Волны Габора, в свою очередь, дают значение для каждой волны изображения. И выходит, что со стандартными параметрами дискретных изображений способ даёт 80 значений в любой позиции пикселя. Набор значений для одного пикселя называется струйным. Рёбра графа взвешиваются расстояниями между смежными вершинами и различие между ними вычисляется при помощи ценовой функции деформации. Она учитывает различие между значениями признаков,

вычисленными в вершинах и степень деформации рёбер. Деформация графа - смещение каждой вершины относительно её первоначального положения и выбора такой позиции, в которой разница между значениями признаков деформируемого графа и соответствующей ей вершине эталонного графа - минимально. Операция будет выполняться для всех вершин до тех пор, пока не будет получено наименьшее различие признаков эталонного и деформируемого графа. Процедура должна выполняться для всех эталонных лиц, заложенных в базу данных системы.

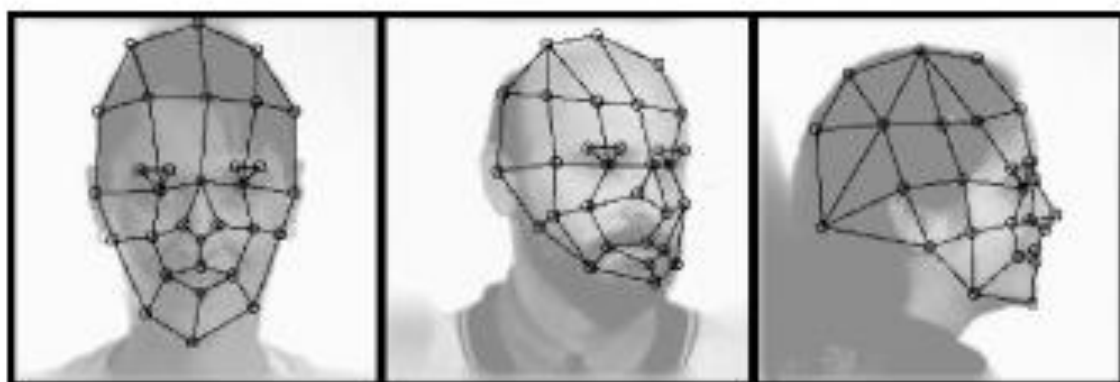


Рисунок 2. Объектно-адаптированные сетки для разных поз.

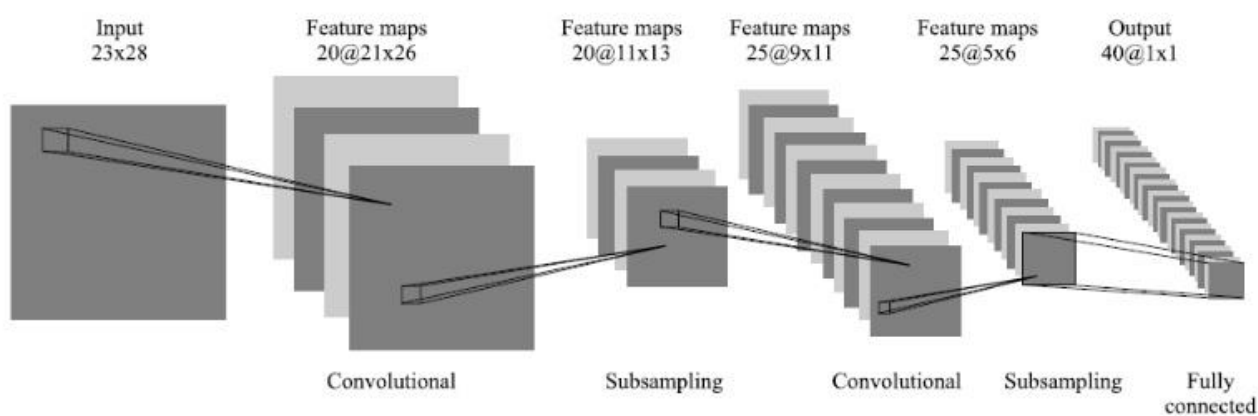
**Недостатки:** высокая вычислительная сложность процедуры распознавания, низкая технологичность при запоминании новых эталонов, линейная зависимость времени работы от размера базы данных.

**Достоинства:** быстроедействие (на современных мобильных устройствах метод обрабатывается примерно за 2 секунды), низкая нагрузка приложения на платформу и процессор устройства.

## Нейронные сети

Один из самых используемых вариантов - сеть, построенная на многослойном перцептроне. Эта сеть позволяет классифицировать изображение в соответствии с предварительным обучением сети. Сети обучаются на наборе общих примеров. Обучение сводится к настройке весов межнейронных связей,

которые настраиваются в процессе решения оптимизационной задачи методом градиентного спуска. Во время обучения нейронная сеть автоматически извлекает ключевые признаки, определяет их важность и строит взаимосвязь между ними. Предполагается, что нейронная сеть сможет применить полученные во время обучения навыки к неизвестным образам. Наилучший результат в этой области показала **сверточная нейронная сеть** (Convolutional Neural Network). Данная сеть – логическое развитие такого архетипа нейронных сетей, как когнитрона и неокогнитрона. Благодаря тому, что сеть может учитывать двумерную топологию изображения, она стала успешной в данной области. Благодаря своим отличительным особенностям: общим весам, иерархичной организации с пространственной подвыборкой (spatial



subsampling) и локальными рецепторными полями, сеть обеспечивает частичную устойчивость к изменениям масштаба, смене ракурса и иным искажениям.

Рисунок 3. Схематичное изображение архитектуры.

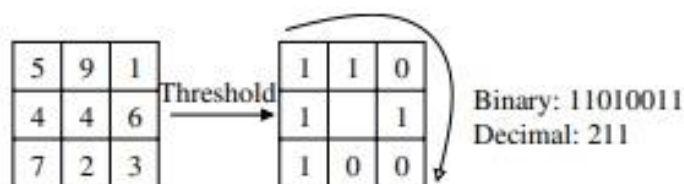
Данная сеть получила развитие в разработке DeepFace, технологии, которую использует Facebook для распознавания лиц пользователей своих социальных сетей. Особенности архитектуры носят закрытый характер.

**Недостатки:** добавление нового эталонного лица в базу данных требует полного переобучения сети, обширный набор математических проблем,

связанные с обучением и трудно формализуемый этап выбора архитектуры сети.

### Local Binary Pattern(LBP) [4]

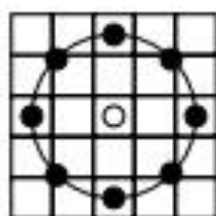
Оригинальный оператор LBP, представленный Ojala [5], является мощным средством описания текстуры. Оператор помечает пиксели изображения, устанавливая порог  $3 \times 3$  – окрестности каждого пикселя центральным



значением и рассматривает результат как двоичное число. В таком случае гистограмма меток может использоваться как дескриптор текстуры.

Рисунок 4. Основной оператор LBP.

Использование круговых окрестностей и билинейная интерполяция значений



пикселей допускает и количество пикселей в окрестности.

Рисунок 5. Круговая окрестность. Значение пикселей билинейно интерполируется всякий раз, когда точка выборки не в центре.

Локальный двоичный шаблон называется равномерным, если он содержит не более двух побитовых переходов от 0 до 1 или наоборот, когда двоичная строка считается круговой. Например, 00000000, 00011110 и 10000011 являются унифицированными шаблонами.

### Концепция:

- 1) Разделить исследуемое окно на ячейки
- 2) Для каждого пикселя в ячейке сравнивать пиксель с каждым из его соседей
- 3) Если значение центрального пикселя больше значения соседнего элемента, напишите «0». В противном случае «1». Это даст 8-значное двоичное число
- 4) Вычислить гистограмму по ячейке частоты каждого встречающегося числа.
- 5) Можно нормализовать гистограмму (необязательно).
- 6) Нормализовать гистограммы всех клеток. Это даст вектор признаков для всего окна.

Вектор функция теперь может быть обработана с использованием какого-либо алгоритма машинного обучения для классификации изображения. Такие классификации могут использоваться для распознавания лиц или анализа текстур.