

<https://doi.org/10.26160/2474-5901-2021-25-51-55>

АЛГОРИТМЫ РАСПОЗНАВАНИЯ ЛИЦ

Половинченко М.И., Елисеев В.С.

Ключевые слова: лица, распознавание, алгоритм, компьютерное зрение.

Аннотация. Системы распознавания лиц могут работать в основном в двух режимах: верификация или аутентификация изображения лица: он в основном сравнивает входное изображение лица с изображением лица, связанным с пользователем, который требует аутентификации. Это в основном сравнение один к одному. Идентификация или распознавание лиц в основном происходит так: сравнивается входное изображение лица со всеми изображениями лица из набора данных с целью найти пользователя, который соответствует этому лицу. Это в основном сравнение один ко многим.

FACE RECOGNITION ALGORITHMS

Polovinchenko M.I., Eliseev V.S.

Keywords: faces, recognition, algorithm, computer vision.

Abstract. Face recognition systems can operate mainly in two modes: face image verification or authentication: it basically compares the input face image with the face image associated with the user that requires authentication. It's basically a one-to-one comparison. Face identification or face recognition basically works like this: the input face image is compared with all the face images from the dataset in order to find a user who matches that face. It's basically a one-to-many comparison.

В информатике распознавание лиц – это в основном задача распознавания человека на основе его изображения лица. Этот алгоритм распознавания стал очень популярным в последние два десятилетия, главным образом из-за новых разработанных методов и высокого качества современных видеокамер.

Стоит обратить внимание, что распознавание лиц сильно отличается от других распознаваний, так как оно имеет цель найти лица (местоположение и размер) на изображении и, вероятно, извлечь их для использования алгоритмом распознавания лиц. А также когда изображения лиц уже извлечены, обрезаны, изменены и обычно преобразованы в оттенки серого, алгоритм распознавания лиц отвечает за поиск характеристик, которые лучше всего описывают изображение.

Системы распознавания лиц могут работать в основном в двух режимах: верификация или аутентификация изображения лица: он в основном сравнивает входное изображение лица с изображением лица, связанным с пользователем, который требует аутентификации. Это в основном сравнение один к одному [1].

Идентификация или распознавание лиц в основном происходит так: сравнивается входное изображение лица со всеми изображениями лица из набора данных с целью найти пользователя, который соответствует этому лицу. Это в основном сравнение один ко многим.

Существуют различные типы алгоритмов распознавания лиц, например:
– собственные поверхности (1991);

- локальные бинарные паттерны гистограммы (LBPH) (1996);
- рыбацкие лица (1997);
- масштабно-инвариантное преобразование признаков (SIFT) (1999);
- ускоренные надежные функции (SURF) (2006).

Каждый метод имеет свой подход к извлечению информации об изображении и выполнению сопоставления с входным изображением. Однако методы Eigenfaces и Fisherfaces имеют сходный подход, а также методы SIFT и SURF.

Обратимся к одному из старейших и более популярных алгоритмах распознавания лиц: локальных бинарных паттернах гистограмм (LBPH).

Цель исследования в данной статье состоит в том, чтобы объяснить LBPH как можно проще, показывая этот метод шаг за шагом.

Local Binary Pattern (LBP) – это простой, но очень эффективный текстурный оператор, который помечает пиксели изображения путем порогового значения окрестности каждого пикселя и рассматривает результат как двоичное число.

Впервые он был описан в 1994 году (LBP) и с тех пор был признан мощным средством классификации текстур. Кроме того, было установлено, что при объединении LBP с гистограммами дескриптора ориентированных градиентов (HOG) значительно улучшается производительность обнаружения на некоторых наборах данных. Используя LBP в сочетании с гистограммами, имеется возможность представить изображения лиц простым вектором данных. Поскольку LBP является визуальным дескриптором, он также может быть использован для задач распознавания лиц.

Теперь, рассмотрим и изучим на шаги алгоритма.

1. Параметры: LBPH использует 4 параметра.

1) Радиус. Радиус используется для построения кругового локального двоичного шаблона и представляет собой радиус вокруг центрального пикселя. Обычно он равен 1.

2) Соседи. Количество выборочных точек для построения кругового локального бинарного паттерна. Имейте в виду: чем больше точек выборки вы включаете, тем выше вычислительные затраты. Обычно он установлен на 8.

3) Сетка X: количество ячеек в горизонтальном направлении. Чем больше ячеек, тем тоньше сетка, тем выше размерность результирующего вектора признаков. Обычно он установлен на 8.

4) Сетка Y: количество ячеек в вертикальном направлении. Чем больше ячеек, тем тоньше сетка, тем выше размерность результирующего вектора признаков. Обычно он установлен на 8.

2. Обучение алгоритму: Во-первых, нам нужно обучить алгоритм. Для этого нам нужно использовать набор данных с изображениями лиц людей, которых мы хотим узнать. Нам также нужно установить идентификатор (это может быть номер или имя человека) для каждого изображения, поэтому алгоритм будет использовать эту информацию, чтобы распознать входное

изображение и выдать вам выход. Изображения одного и того же человека должны иметь один и тот же идентификатор. С уже построенным обучающим набором давайте рассмотрим вычислительные шаги LBPН.

3. Применение операции LBP: Первым вычислительным шагом LBPН является создание промежуточного изображения, которое лучше описывает исходное изображение, выделяя характеристики лица. Для этого алгоритм использует концепцию скользящего окна, основанную на параметрах радиуса и соседей. На рисунке 1 показана эта процедура.

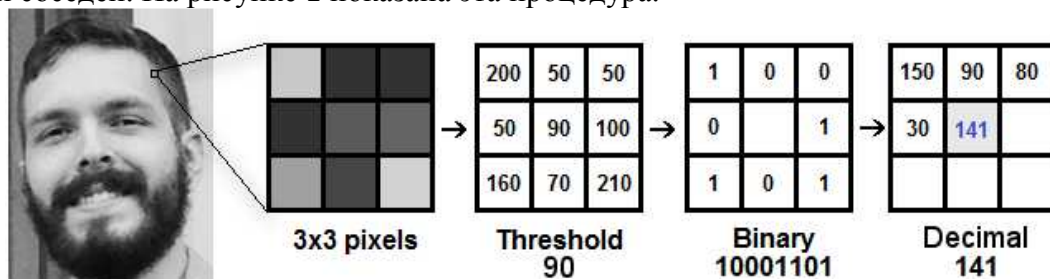


Рис. 1. Применение операции LBP

Основываясь на рисунке 1, разделим его на несколько небольших шагов, чтобы можно было наиболее легко понять его. Предположим, у нас есть изображение лица в оттенках серого. Мы можем получить часть этого изображения в виде окна размером 3x3 пикселя. Оно также может быть представлено в виде матрицы 3x3, содержащей интенсивность каждого пикселя (0~255). Затем необходимо взять центральное значение матрицы, которое будет использоваться в качестве порога [2]. Это значение будет использоваться для определения новых значений из 8 соседей. Для каждого соседа центрального значения (порога) мы устанавливаем новое двоичное значение. Мы устанавливаем 1 для значений, равных или превышающих порог, и 0 для значений, меньших порога. Теперь матрица будет содержать только двоичные значения (игнорируя центральное значение). Нам нужно объединить каждое двоичное значение из каждой позиции матрицы строка за строкой в новое двоичное значение.

Затем мы преобразуем это двоичное значение в десятичное и устанавливаем его в центральное значение матрицы, которое на самом деле является пикселем исходного изображения. В конце этой процедуры (процедуры LBP) мы получаем новое изображение, которое лучше отражает характеристики исходного изображения.

Процедура LBP была расширена, чтобы использовать другое количество радиусов и соседей, она называется круговой LBP, что представлено на рисунке 2.

Это можно сделать с помощью билинейной интерполяции. Если какая-нибудь точка данных находится между пикселями, она использует значения из 4 ближайших пикселей (2x2) для оценки значения новой точки данных.

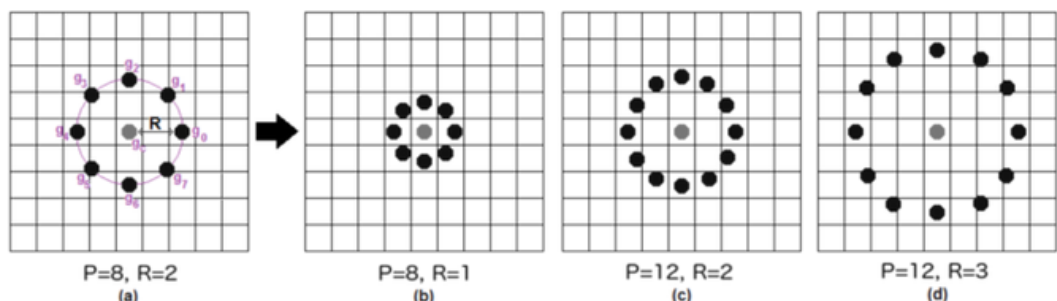


Рис. 2. Круговой LBP

4. Извлечение гистограмм: Теперь, используя изображение, сгенерированное на последнем шаге, мы можем использовать параметры сетки x и сетки y для деления изображения на несколько сеток, как это видно на рисунке 3.

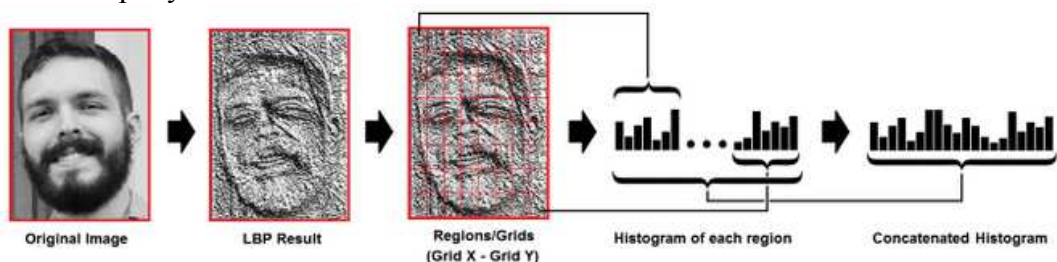


Рис. 3. Разбиение изображения по сетке

Основываясь на приведенном выше изображении, мы можем извлечь гистограмму каждой области следующим образом.

Поскольку у нас есть изображение в оттенках серого, каждая гистограмма (из каждой сетки) будет содержать только 256 позиций (0~255), представляющих вхождения интенсивности каждого пикселя.

Затем нам нужно объединить каждую гистограмму, чтобы создать новую и большую гистограмму. Предположим, что у нас есть сетки 8×8 , мы будем иметь $8 \times 8 \times 256 = 16384$ позиции в конечной гистограмме. Итоговая гистограмма представляет характеристики изображения исходного изображения.

5. Выполнение распознавания лиц: На этом этапе алгоритм уже обучен. Каждая созданная гистограмма используется для представления каждого изображения из обучающего набора данных. Итак, учитывая входное изображение, мы снова выполняем шаги для этого нового изображения и создаем гистограмму, что проиллюстрировано на рисунке 3.

Таким образом, чтобы найти изображение, соответствующее входному изображению, нам просто нужно сравнить две гистограммы и вернуть изображение с ближайшей гистограммой. Мы можем использовать различные подходы для сравнения гистограмм (вычислить расстояние между двумя гистограммами), например: евклидово расстояние [3], хи-квадрат, абсолютное

значение и т.д. В этом примере мы можем использовать евклидово расстояние (которое достаточно известно), основанное на следующей формуле

$$D = \sqrt{\sum_{i=1}^n (hist1_i - hist2_i)^2}.$$

Таким образом, на выходе алгоритма получаем идентификатор изображения с ближайшей гистограммой. Алгоритм также должен возвращать вычисленное расстояние, которое может быть использовано в качестве «доверительного» измерения. Затем мы можем использовать порог, чтобы автоматически оценить, правильно ли алгоритм распознал изображение. Можно предположить, что алгоритм успешно распознал, если доверительная вероятность ниже определенного порога.

В заключение можно сделать вывод о том, что LBRH-один из самых простых алгоритмов распознавания лиц. Он может представлять локальные объекты на изображениях. С помощью него можно получить отличные результаты, (в основном в контролируемой среде). Он устойчив к монотонным преобразованиям серой шкалы. Это предусмотрено благодаря библиотеке OpenCV (библиотека с открытым исходным кодом по компьютерному зрению).

Список литературы / References

1. Ahonen, Timo, Abdenour Hadid, and Matti Pietikainen. Face description with local binary patterns: Application to face recognition // IEEE transactions on pattern analysis and machine intelligence 28.12 (2006): 2037-2041.
2. Ojala, Timo, Matti Pietikainen, and Topi Maenpaa. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns // IEEE Transactions on pattern analysis and machine intelligence 24.7 (2002): 971-987.
3. Ahonen, Timo, Abdenour Hadid, and Matti Pietikainen. Face recognition with local binary patterns // Computer vision-eccv 2004 (2004): 469-481.

Половинченко Марина Игоревна – магистрант, m.polovin4enko@yandex.ru	Polovinchenko Marina Igorevna – master's student, m.polovin4enko@yandex.ru
Елисеев Владислав Сергеевич – магистрант	Eliseev Vladislav Sergeevich – master's student
Донской государственный технический университет, Ростов-на-Дону, Россия	Don state technical University, Rostov-on-don, Russia

Received 08.05.2021