

DOI 10.23859/1994-0637-2019-4-91-2
УДК004.932.2

© Трофименко Я. М., Виноградова Л. Н., Ершов Е. В., 2019

Трофименко Ярослав Максимович

Аспирант,
Череповецкий государственный университет
(Череповец, Россия)
E-mail: trofimenko.y.m@yandex.com

Trofimenko Yaroslav Maksimovich

Post-graduate student,
Cherepovets State University
(Cherepovets, Russia)
E-mail: trofimenko.y.m@yandex.com

Виноградова Людмила Николаевна

Кандидат технических наук, доцент,
Череповецкий государственный университет
(Череповец, Россия)
E-mail: lnvinogradova@bk.ru

Vinogradova Lyudmila Nikolayevna

PhD in Technical Sciences,
Associate Professor,
Cherepovets State University
(Cherepovets, Russia)
E-mail: lnvinogradova@bk.ru

Ершов Евгений Валентинович

Доктор технических наук, профессор,
Череповецкий государственный университет
(Череповец, Россия)
E-mail: ershov_ev@mail.ru

Ershov Evgeniy Valentinovich

Doctor of Technical Sciences, Professor,
Cherepovets State University
(Cherepovets, Russia)
E-mail: ershov_ev@mail.ru

**АЛГОРИТМЫ ПРЕДОБРАБОТКИ
ИЗОБРАЖЕНИЙ В СИСТЕМЕ
ИДЕНТИФИКАЦИИ ЛИЦ
В ВИДЕОПОТОКЕ**

**ALGORITHMS FOR IMAGE
PRE-PROCESSING IN THE FACE
IDENTIFICATION SYSTEM
IN THE VIDEO STREAM**

Аннотация. Рассмотрены методы обработки изображения для системы идентификации лиц в видеопотоке. Реализован алгоритм предобработки изображений для минимизации влияния изменений освещения и положения лица в кадре на точность идентификации. Данный алгоритм адаптирован для обработки кадров из видеопотока, поступающего с веб-камеры в реальном времени. Приведены результаты экспериментальной проверки работы алгоритма.

Abstract. Image processing techniques for the system of face identification in the video stream are reviewed. An image pre-processing algorithm is implemented to minimize the effect of changes in the illumination and position of a face in a frame on the accuracy of identification. This algorithm is adapted for processing frames from a video stream coming from a webcam in real time. The results of experimental verification of the algorithm are given.

Ключевые слова: обработка изображений, идентификация лиц, детектирование лиц, минимизация влияния освещения, выравнивание идентифицируемого объекта, видеопоток

Keywords: image processing, face identification, face detection, minimizing the effects of illumination, alignment of the identified object, video stream

Введение

Идентификация лиц является сложным многоуровневым процессом, для реализации которого требуются значительные вычислительные ресурсы. В настоящее время

они становятся все более доступными благодаря развитию компьютеров и устройств ввода-вывода, таких как камеры. Идентификация лиц на изображениях является актуальной темой разработок приложений, внедряемых в повседневную жизнь (начиная от подтверждения платежей в банкоматах и заканчивая медицинскими приложениями, позволяющими определить некоторые генетические заболевания по чертам лица) [5].

Несмотря на возрастающую популярность и расширяющуюся область применения идентификации лиц на изображениях и видео, данные технологии и алгоритмы еще имеют ряд проблем и недостатков, которые только предстоит решить разработчикам приложений и информационных систем данного направления.

Основная часть

Общими проблемами, связанными со спецификой обработки изображений, являются следующие [3]:

А. Проблема освещенности, заключающаяся в том, что результаты идентификации сильно зависят от положения источника света относительно лица и общей освещенности в кадре. Так, при недостаточном освещении почти невозможно выделить лицо на изображении, а при интенсивном могут теряться некоторые черты лица.

В. Проблема цветопередачи изображения, заключающаяся в том, что для точной цветопередачи необходимо отрегулировать баланс белого камеры в соответствии с цветовой температурой используемого источника света. В большей степени это касается камер, установленных вне помещений, поскольку при наружном видеонаблюдении цветовая температура меняется в течение дня.

С. Проблема разрешения изображения, которая может быть представлена в виде вопроса: Насколько высокое разрешение необходимо для наиболее точной идентификации без чрезмерного увеличения размеров изображения?

Д. Проблема расположения камеры. В области обзора не должно находиться объектов, попадающих между камерой и лицом, иначе возникнут помехи, и идентификация не выполнится или выполнится с ошибкой. Кроме того, видеозаписывающее устройство следует установить примерно на одном уровне с лицом, чтобы объект был расположен перпендикулярно ему.

Е. Проблема положения лица относительно камеры. То, под каким углом к камере повернута голова, может серьезно влиять на идентификацию, особенно в том случае, когда база данных обладает ограниченным набором изображений для конкретного лица.

Процесс идентификации лица заключается в нахождении его на изображении и сопоставлении с эталонами, имеющимися в базе данных. Сложность детектирования и идентификации лица заключается в том, что оно является слабоконтрастным объектом. Лица имеют чрезвычайно сложный рисунок и часто отличаются друг от друга только неуловимыми деталями, поэтому нередко как человеку, так и компьютеру бывает невозможно сопоставить изображения при наличии разницы в освещении, угле обзора камеры, не говоря уже об изменениях внешнего вида, и это при том, что человеческий мозг обладает гораздо большей, чем компьютеры, способностью компенсации изменений освещенности и угла зрения.

Автоматическое детектирование лица, предшествующее идентификации, строится за счет обнаружения на его изображении областей с перепадами яркости, характерными для глаз, губ, носа, носогубной складки. При этом на конечном этапе детектирования учитывается положение областей относительно друг друга, поскольку человеческое лицо должно иметь определенные пропорции.

Благодаря тому, что известен набор основных частей лица, а также их положение и пропорции, в алгоритмы детектирования добавляют проверки на соответствие результата этим величинам, что позволяет не включать те результаты, в которых расположение и пропорции частей лица не соответствуют характерным для человека.

Математическая модель процесса идентификации основывается на определении набора признаков и их извлечении из изображения лица с целью сопоставления данных значений с эталонными.

Имеется множество лиц I , которое используется в процессе идентификации. Множество I состоит из отдельных непересекающихся подмножеств L , представляющих собой наборы лиц F для конкретных объектов идентификации. Каждое такое подмножество L является отдельным классом, которое представляет конкретное лицо. Все изображения лиц G , еще не прошедшие идентификацию, принадлежат подмножеству N , а N в свою очередь – множеству I . При этом N не пересекается ни с одним из подмножеств L .

Для каждого лица G из подмножества N и лица F из подмножества L имеется набор признаков $P(F) = \{p_1, p_2, p_3 \dots p_n\}$, которые являются элементами некоторого конечного множества признаков $A = \{a_1, a_2, a_3 \dots a_n\}$, используемых при идентификации лиц. Каждый признак представляет собой пару параметров: название, позволяющее выделить данный признак из числа других, и его значение / величина. Значение признака должно лежать в допустимых для него пределах, при этом для каждого пределы значений могут быть свои.

В процессе идентификации выполняется проверка попадания значений признаков $P(G)$ в пределы, полученные из множества значений признаков $\{P(F), P(F_1) \dots P(F_K)\}$ для каждого конкретного L . В случае, если значения признаков $P(G)$ попадают в эти пределы, можно утверждать, что изображение лица G относится к одному из классов L множества I .

Использование описанной модели идентификации должно осуществляться с учетом применения искусственной нейронной сети (ИНС). Особенностью данного подхода является то, что для обучения ИНС необходимо составить обучающую выборку, в которой будут присутствовать наборы лиц F для каждого из подмножеств L . Принадлежность каждого лица F к конкретному классу L известна «учителю». При обучении на вход в ИНС подается вектор значений, характеризующих изображение и ожидаемое значение выхода – принадлежность образца конкретному классу L . В процессе обучения нейронная сеть извлекает из входного вектора признаки и умножает их на установленные случайным образом веса связей между нейронами разных слоев [1].

Признаки $P(F)$ извлекаются из входного вектора в процессе обучения. В результате первый скрытый слой выделяет низкоуровневое пространство признаков входных данных, второй слой детектирует пространство признаков более высокого уровня абстракции и т. д. [4]. После этого происходит сопоставление полученного ре-

зультата с ожидаемым, и в случае несовпадения производится корректировка весов. Для проверки точности обучения требуется сформировать тестовую выборку изображений лиц G . Каждое из них, согласно оценке «учителя», должно принадлежать конкретному классу L . При этом в тестовой выборке не может быть изображений лиц, которые были в обучающей выборке. В результате успешного обучения на вход ИНС можно подавать образцы, чья принадлежность к одному из классов заранее неизвестна.

Обработка изображения лица осуществляется при помощи следующих действий:

- 1) масштабирование изображения – метод ближайшего соседа;
- 2) обесцвечивание изображения;
- 3) выравнивание гистограммы.

Масштабирование осуществляется за счет переноса значений пикселей исходного изображения на пустое нужного размера с объединением соседних схожих значений в том случае, если исходное изображение было больше необходимого размера. При интерполяции методом ближайшего соседа в качестве промежуточного значения выбирается ближайшее известное значение функции, данный метод является самым простым методом интерполяции, не требующим больших ресурсов.

Обесцвечивание изображения осуществляется при помощи умножения каждой цветовой составляющей RGB на соответствующий коэффициент по следующей формуле (1):

$$P_{color} = R \cdot 0,30 + G \cdot 0,59 + B \cdot 0,11, \quad (1)$$

где P_{color} – цвет обесцвеченного пикселя, R – красная составляющая пикселя, G – зеленая составляющая пикселя, B – синяя составляющая пикселя.

Выравнивание гистограммы выполняется для того, чтобы все уровни яркости имели одинаковую частоту, а гистограмма соответствовала равномерному закону распределения. Так, для изображения с разрешением $N \times M$ пикселей количество уровней квантования яркости пикселей (число бинов) составляет J . Тогда среднее количество пикселей на каждый уровень яркости должно рассчитываться по формуле (2):

$$n_{avg} = \frac{M \times N}{J}. \quad (2)$$

Базовая математика лежит в сопоставлении двух распределений. Пусть x, y – случайные величины, описывающие изменение интенсивности пикселей на изображениях, $w_x(x)$ – плотность распределения интенсивности на исходном изображении, $w_y(y)$ – желаемая плотность распределения. Необходимо найти преобразование плотностей распределения $y=f(x)$, которое позволило бы получить желаемую плотность (3):

$$w_y(y) = \begin{cases} \frac{1}{y_{max}-y_{min}}, & y_{min} \leq y \leq y_{max} \\ 0, & \text{в противном случае} \end{cases}. \quad (3)$$

Обозначим через $F(x)$ и $F(y)$ интегральные законы распределения случайных величин x и y . Из условия вероятностной эквивалентности следует, что $F_x(x) = F_y(y)$.

Распишем интегральный закон распределения по определению и получим формулу (4).

$$F_x(x) = F_y(y) = \int_{y_{min}}^y w_y(y) dy = \frac{y - y_{min}}{y_{max} - y_{min}}. \quad (4)$$

Отсюда получаем следующую формулу (5):

$$y = (y_{max} - y_{min})F_x(x) + y_{min}. \quad (5)$$

Далее требуется определить, как оценить интегральный закон распределения $F_x(x)$. Для этого необходимо сначала построить гистограмму исходного изображения, затем нормализовать полученную гистограмму, разделив величину каждого бина на общее количество пикселей $N \times M$. Значения бинов можно рассматривать как приближенное значение функции плотности распределения $w_x^*(x)$, $0 \leq x \leq 255$. Таким образом, значение интегральной функции распределения можно представить как сумму следующего вида (6):

$$F_x^*(x) = \sum_{j=0}^x w_x^*(j). \quad (6)$$

Построенную оценку можно использовать для вычисления новых значений интенсивности. Перечисленные преобразования гистограмм можно применять не только ко всему изображению, но и к отдельным его частям.

При применении выравнивания гистограммы последовательно к отдельным частям изображения можно говорить о методе адаптивного выравнивания. Существует разновидность данного метода, называемая контрастно-ограниченным адаптивным выравниванием гистограммы (Contrast Limited Adaptive Histogram Equalization CLANE) [2]. Помимо применения выравнивания гистограммы к отдельным частям, данный метод ограничивает усиление контраста, что помогает избавиться от проблемы усиления шума в определенных границах контрастности.

Алгоритм адаптивного выравнивания гистограммы работает в пять этапов:

1. Выделение области сканирующего окна на исходном изображении (8-битном одноканальном). Преобразование изображения из трехканального RGB в одноканальное по формуле 1.

2. Вычисление гистограммы H в сканирующем окне.

3. Нормализация гистограммы обеспечивает растяжку не всего диапазона изменения интенсивностей, а только его наиболее информативной части. Под информативной частью понимается набор пиков гистограммы (интенсивности), которые чаще остальных встречаются в области. Бины, соответствующие редко встречающимся интенсивностям, в процессе нормализации отбрасываются. Нормализация осуществляется посредством деления величины каждого бина гистограммы на общее количество пикселей.

4. Построение интегральной гистограммы: $H'_i = \sum 0 \leq j \leq H_j$.

5. Определение нового значения интенсивности пикселя (7):

$$dst(x, y) = H'_i(src(x, y)). \quad (7)$$

Данный алгоритм повторяется для всех пикселей изображения, что позволяет осветлить слишком темные участки и затемнить слишком светлые в пределах скани-

рующего окна, тем самым выровнять яркость даже в случае значительных перепадов на разных участках.

Для уменьшения влияния наклона лица относительно вертикали в кадре имеет смысл применить алгоритм поворота изображения на определенный вычисляемый угол. Для этого нужно:

- 1) выполнить детектирование глаз на исходном изображении, проверив, что обнаружено не менее двух глаз;
- 2) вычислить координаты центров для каждого глаза;
- 3) вычислить разницу между соответствующими координатами (8, 9):

$$dX = C_1(x) - C_2(x), \quad (8)$$

$$dY = C_1(y) - C_2(y), \quad (9)$$

где $C_1(x)$ – координата центра левого глаза по X и $C_2(x)$ – координата центра правого глаза по X , $C_1(y)$ и $C_2(y)$ – соответствующие координаты по оси Y .

- 4) с учетом полученных данных вычислить необходимый угол для поворота изображения (10):

$$\theta = \tan(dY, dX). \quad (10)$$

- 5) значение угла нужно перевести в градусы и выполнить аффинное преобразование, результатом которого будет изображение лица, выровненное относительно вертикали.

Блок-схема алгоритма поворота изображения представлена на рис. 1.

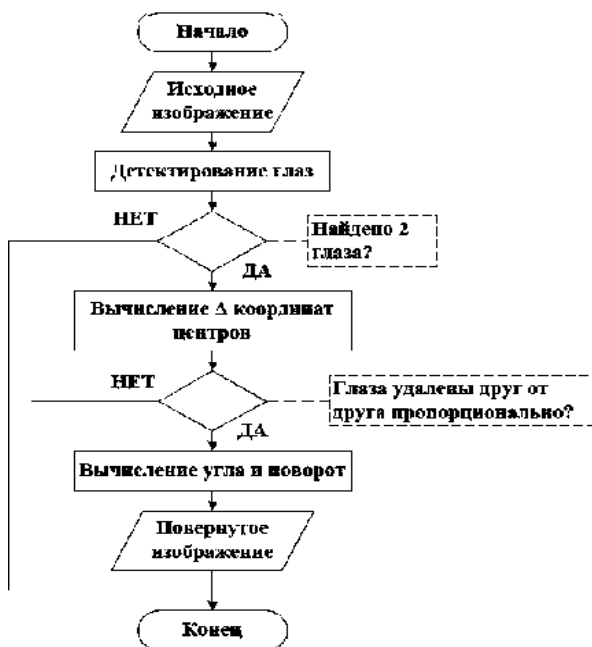


Рис. 1. Алгоритм поворота изображения

В случаях, когда на изображении найден только 1 глаз или же расстояние между центрами найденных глаз является непропорциональным, т. е. не попадающим в установленные пределы, алгоритм выравнивания прекратит свою работу в связи с невозможностью дальнейших вычислений и преобразований.

Идентификация в системе может осуществляться как автоматически с заданным интервалом, пока есть видеопоток с камеры, так и однократно по команде пользователя. В первом случае общий алгоритм работы системы идентификации лиц будет представлять собой цикл, во втором – единичный проход по шагам алгоритма.

Блок-схема обобщенного алгоритма идентификации представлена на рис. 2.

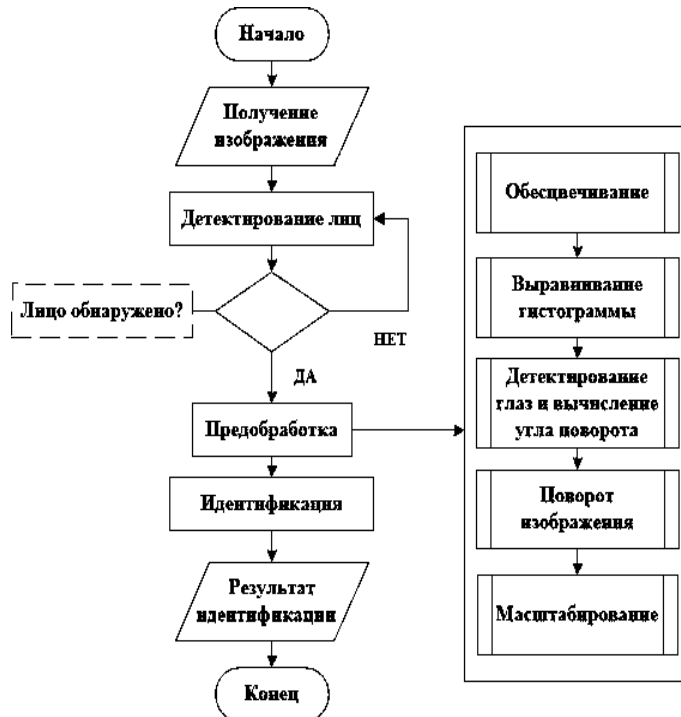


Рис. 2. Обобщенный алгоритм предобработки и идентификации

Обобщенный алгоритм идентификации состоит из следующих этапов:

- 1) получение видеопотока с камеры;
- 2) получение отдельного кадра из видеопотока;
- 3) применение алгоритма Виолы-Джонса для детектирования лиц в кадре, полученном из видеопотока;
- 4) в случае обнаружения лица выполнить предобработку, включающую в себя:
 - а) обесцвечивание изображения;
 - б) выравнивание гистограммы;
 - в) использование метода Виолы-Джонса для детектирования глаз, вычисление угла поворота изображения для выравнивания глаз на одном уровне;
 - г) поворот изображения;
 - д) масштабирование изображения к размеру 90 на 100 пикселей;

- 5) преобразование обработанного изображения лица в вектор;
- 6) подача вектора на вход обученной искусственной нейронной сети, которая классифицирует лицо;
- 7) вывод результата идентификации.

Для экспериментальной проверки была составлена выборка из 200 лиц, для каждого из них имелось по 5 различных изображений с разным освещением и положением головы, общая выборка состояла из 1000 изображений. Для обучающей выборки было выбрано 600 изображений (по 3 для каждого лица), а оставшиеся 400 (или по 2 изображения для каждого лица) вошли в тестовую выборку. Все изображения лиц были взяты из открытых баз от Dr. L. Spacek Description of the Collection of Facial Images. В результате экспериментальной проверки установлено, что 94 % лиц из тестовой выборки были безошибочно идентифицированы, несмотря на изменения в освещении и положении лица, 5 % были идентифицированы неверно, и 1 % лиц не был идентифицирован. При этом алгоритм идентификации без предобработки распознавал только 80 % лиц из выборки верно, 15 % – неверно и 5 % лиц не были идентифицированы. Экспериментальная проверка выполнялась с использованием веб-камеры с разрешением 640x480 пикселей.

Выводы

Одна из областей применения созданного алгоритма идентификации лиц в видеопотоке – разработка программного обеспечения для социальных роботов и роботов-компаньонов. В данном случае распознавание лиц может быть усовершенствовано идентификацией по голосу, во внешних условиях, которые могут изменяться в определенных пределах. Также есть возможность применять данный алгоритм в системе подсчета посетителей для какого-либо объекта, такого как аудитория или вход в офис.

При использовании больших выборок, когда одному человеку соответствует набор фотографий лиц (при этом изображения регулярно обновляются и соответствуют по внешним условиям местам, где производится идентификация), возможно повышение точности распознавания. При отсутствии актуальных фотографий или маленьких выборках точность идентификации со временем будет снижаться из-за изменчивости человеческого лица.

Литература

1. Брилюк Д. В., Старовойтов В. В. Распознавание человека по изображению лица нейросетевыми методами. – Минск: Институт технической кибернетики Национальной академии наук Беларуси, 2002. – 54 с.
2. Гонсалес Р., Вудс Р. Цифровая обработка изображений. – Москва: Техносфера, 2012. – 1104 с.
3. Идентификация и распознавание. – URL: <https://www.axis.com/ru/ru/learning/web-articles/identification-and-recognition> (дата обращения: 10.04.2019).
4. Hinton G., Salakhutdinov R. Reducing the dimensionality of data with Neural Networks // Science. – 2006. – № 313. – Vol. 5786. – P. 504–507.

5. How facial recognition technology is creeping into your life // The Week. – URL: <http://theweek.com/articles/737750/how-facial-recognition-technology-creeping-into-life> (дата обращения: 10.04.2019).

References

1. Briliuk D. V., Starovoitov V. V. *Raspoznavanie cheloveka po izobrazheniiu litsa neirosetevymi metodami* [Person identification by face images using neural network methods]. Minsk: Institut tekhnicheskoi kibernetiki Natsional'noi akademii nauk Belarusi, 2002. 54 p.
2. Gonsales P., Vuds R. *Tsifrovaia obrabotka izobrazhenii* [Digital image processing]. Moscow: Tekhnosfera, 2012. 1104 p.
3. *Identifikatsiia i raspoznavanie* [Identification and recognition]. Available at: <https://www.axis.com/ru/ru/learning/web-articles/identification-and-recognition> (accessed: 10.04.2019).
4. Hinton G., Salakhutdinov R. Reducing the dimensionality of data with Neural Networks. *Science*, 2006, no. 313, vol. 5786, pp. 504–507.
5. How facial recognition technology is creeping into your life. *The Week*. Available at: <http://theweek.com/articles/737750/how-facial-recognition-technology-creeping-into-life> (accessed: 10.04.2019).

Для цитирования: Трофименко Я. М., Виноградова Л. Н., Ершов Е. В. Алгоритмы предобработки изображений в системе идентификации лиц в видеопотоке // Вестник Череповецкого государственного университета. – 2019. – № 4 (91). – С. 21–29. DOI: 10.23859/1994-0637-2019-4-91-2

For citation: Trofimenko Ya. M., Vinogradova L. N., Ershov E. V. Algorithms for image pre-processing in the face identification system in the video stream. *Bulletin of the Cherepovets State University*, 2019, no. 4 (91), pp. 21–29. DOI: 10.23859/1994-0637-2019-4-91-2