

ПОДГОТОВКА ИЗОБРАЖЕНИЙ ЛИЦ В ВИДЕОПОТОКЕ К РАСПОЗНАВАНИЮ И ФИЛЬТРАЦИЯ НЕИНФОРМАТИВНЫХ ИЗОБРАЖЕНИЙ

С. Г. Небаба, С. Ю. Андреев, М. А. Макаров

Томский политехнический университет¹,
634050, Томск, Россия

УДК 303.732.4

Выделены основные задачи подготовки изображения к процессу распознавания лиц и критерии оценки качества изображения. Предложен метод трекинга лица в видеопотоке, сформулирован критерий поиска похожих изображений, проведен анализ существующих алгоритмов оценки контрастности и резкости изображения. Рассмотрены метод компенсации освещения, а также метод контроля ракурса на основе алгоритма POSIT.

Ключевые слова: компьютерное зрение, трекинг, распознавание лиц, контроль ракурса, контрастность, резкость.

There have been specified the principal tasks of the image preparation for the face recognition and criteria of the image quality evaluation. There has been proposed a method of the face tracking in a video stream, a search criterion of the similar images formulated and the current estimators of the contrasting effect and image sharpness analyzed. The method of lighting compensation and the angle control method on the basis of POSIT algorithm have been studied.

Key words: computer vision, tracking, facial recognition, face position control, contrast, sharpness.

Введение. Работы по решению задачи идентификации личности на основе анализа изображений лица ведутся с самых ранних этапов развития компьютерного зрения [1, 2, 3]. В последнее время в различных сферах деятельности возрастает потребность быстрой и правильной идентификации личности человека в видеопотоке.

Задача поиска лиц на изображении на сегодняшний день успешно решена и применяется во многих технических устройствах. Например, в фототехнике поиск лиц используется для автоматического выбора точки фокусировки.

Задача распознавания (идентификации) лиц является более сложной, и на данный момент не существует алгоритма, который бы приблизился к точности распознавания лиц человеком.

Эффективность идентификации может быть оценена на основе двух вероятностных характеристик:

- 1) вероятность ложноположительного срабатывания (ложное опознание человека — обознались);
- 2) вероятность ложноотрицательного срабатывания (пропуск искомой личности — не узнали).

¹Работа выполнена за счет средств субсидии в рамках реализации Программы повышения конкурентоспособности ТПУ.



Рис. 1. Пример трека лица в видеопотоке

В системах идентификации эти вероятности являются взаимосвязанными величинами, находящимися в обратно пропорциональной зависимости. Помимо этого, все современные системы идентификации личности имеют четкую зависимость обоих видов ошибок от следующих факторов: 1) ракурс съемки идентифицируемого лица; 2) качество изображения лица, находящегося как в базе данных, так и регистрируемого камерами видеонаблюдения; 3) условия освещения.

Жесткие требования к ракурсу съемки опознаваемого лица обусловлены используемыми алгоритмами, самые лучшие из которых не позволяют получать удовлетворительные результаты распознавания даже при небольших углах поворота фронтальной плоскости лица по отношению к оптической оси объектива регистрирующей камеры. Как правило, существуют рекомендации, что угол поворота не должен превышать 15 градусов. При этом надо учитывать, что наилучшие результаты получаются при угле 0 градусов, а при угле, превышающем 15 градусов, результаты могут расцениваться как неудовлетворительные.

Изображения анализируемых лиц должны обладать минимально необходимым качеством, чаще всего требования к изображению включают такие пункты, как: 1) отсутствие эффекта смазанного изображения, вызванного движением идентифицируемого объекта; 2) изображение анализируемых лиц не должно быть меньше минимально необходимого размера, расстояние между зрачками не должно быть меньше 80 точек растра.

Условия освещения должны позволять регистрировать изображения лиц, на которых должны быть различимы основные черты лица (достаточный контраст). Для всех современных систем идентификации важно наличие прямого или близкого к нему угла падения лучей света на фронтальную плоскость лица, а также наличие рассеянного источника света (условия, приближенные к условиям получения фотографий для официальных документов). Крайне негативное влияние на качество идентификации оказывает наличие боковой засветки лица.

Учитывая вышеперечисленные свойства методов идентификации личности по растровому изображению лица, можно сделать вывод о необходимости предварительной фильтрации поступающих изображений с целью отбраковки непригодных для анализа.

Трекинг. Как правило, одно и то же лицо присутствует в поле зрения камеры видеонаблюдения в течение некоторого временного интервала, поэтому его изображение может быть найдено в последовательности кадров. Это позволяет осуществить трекинг анализируемого объекта.

Трекинг (tracking) — это определение местоположения движущегося объекта в видеопотоке. Алгоритм анализирует кадры видеопотока и выдает положение движущихся целевых объектов относительно кадра, позволяя составить последовательность изображений конкретного объекта, которая называется треком (track). Пример трека лица в видеопотоке представлен на рис. 1.

Видеопоток можно представить в виде последовательности кадров (изображений)

$$V_s = [f_1, f_2 \dots f_N],$$

где N — общее число кадров, f_i — текущий кадр. Любое лицо, найденное в кадре, описывается как набор

$$Fc_{ij} = \{If_{ij}, Rf_{ij}\},$$

где If_{ij} — изображение лица, Rf_{ij} — его положение в кадре, i — номер кадра, j — номер лица в кадре. При этом Rf можно представить как набор из двух точек

$$Rf = \{p_1 = (x_1; y_1), p_2 = (x_2; y_2)\},$$

где p_1 — верхняя левая точка прямоугольной области, p_2 — нижняя правая точка, x и y — их координаты.

Тогда трек лица k в видеопотоке — это последовательность лиц Fc_{ik} , соответствующая одному человеку:

$$Tr_k = [Fc_{1k}, Fc_{2k} \dots Fc_{Nk}].$$

Основная задача в трекинге состоит в сопоставлении положений целевого объекта на последовательности кадров. Таким образом, системы трекинга обычно используют модель движения, которая описывает, как может изменяться изображение целевого объекта при всевозможных различных его движениях [4]. В случае слежения за положением лица предполагается, что объекты в следующем кадре меняют свою позицию незначительно. В связи с этим предложен следующий метод трекинга: найденное изображение лица принадлежит треку Tr_k , если точка-центр прямоугольника Rf , $p_c = (\frac{x_1 + x_2}{2}; \frac{y_1 + y_2}{2})$, лица Fc_{ij} принадлежит области прямоугольника Rf лица $Fc_{i-1,j}$, найденного на предыдущем кадре и принадлежащего треку Tr_k .

Предложенный метод не может гарантировать абсолютную точность распознавания трека, но является очень простым с вычислительной точки зрения и достаточно эффективным для решаемой задачи.

Информативность изображения. В большей части систем видеонаблюдения используется стандартная частота поступления кадров, равная 25 кадрам в секунду. Очевидно, что изображения лица одного человека на двух соседних кадрах практически не должны отличаться друг от друга. Биометрический анализ одинаковых изображений неоправдан с точки зрения загрузки вычислительных ресурсов. В связи с этим целесообразно проводить отбор кадров, на которых изображения одного лица существенно отличаются друг от друга.

Корреляция изображений. Оценка схожести двух изображений может производиться путем вычисления коэффициента корреляции и исключения из трека Tr_k изображений с высоким уровнем корреляции

$$R(x, y) = \frac{\sum_{x', y'} (If_{ik}(x', y') \cdot If_{i+1, k}(x + x', y + y'))}{\sqrt{\sum_{x', y'} If_{ik}(x', y')^2 \cdot \sum_{x', y'} If_{i+1, k}(x + x', y + y')^2}}, \quad (1)$$

где If_{ik} — исходное изображение, $If_{i+1, k}$ — изображение для сравнения. Оценка, получаемая по формуле (1), лежит в диапазоне от 0 до 1, где 0 соответствует максимально разным изображениям, а 1 — максимально похожим.

Этот метод достаточно корректно отображает степень похожести двух изображений, но плохо применим для оценки кадра целиком, так как требователен к вычислительным ресурсам. Однако его можно применять для поиска серии лиц на изображении, принадлежащих одному человеку, и выбору из этой серии только отличающихся друг от друга изображений.

Оценка контрастности. Аналогично условием информативности можно указать контрастность изображения, так как этот показатель учитывает освещенность и различимость объектов на изображении. Этот параметр недостаточен, чтобы можно было признать изображение качественным, так как боковая засветка объектов дает высокую оценку контраста, но затрудняет выделение деталей изображения.

Тем не менее, в случае, если изображение лица удастся нормализовать, контрастность становится более полезной оценкой. Так как при анализе лиц используются черно-белые изображения (в градациях серого), в качестве оценки контрастности допустимо взять яркостную контрастность, которая вычисляется как дисперсия яркости пикселей изображения [5]:

$$\sigma^2 = \frac{1}{N} \sum_{p=1}^N (Y_p - Y)^2, \quad (2)$$

где Y — среднее значение яркостной контрастности всего изображения, Y_p — значение яркостной контрастности в точке p , N — общее число точек изображения.

Оценка контрастности (2) может быть нормирована путем вычисления отношения среднеквадратического отклонения к максимально возможному значению яркости:

$$C = \frac{2\sigma}{Y_{max}}. \quad (3)$$

C изменяется в диапазоне $[0; 1]$. Значение 0 соответствует однотонному изображению, значение 1 — максимально контрастному. Оптимальное значение контрастности (3) зависит от типа объекта, представленного на изображении. На рис. 2 приведены примеры изображений и их оценка контрастности.

Как можно видеть на рис. 2, контрастность изображения не может являться точным показателем качества изображения, например, изображения а) и г) размытые и менее информативные, но при этом имеют более высокую оценку контрастности, чем изображение б), из-за большого количества точек максимальной и минимальной яркости.

Оценка резкости. Еще одним критерием информативности изображения выступает резкость, которая условно может быть определена наличием контуров контрастных переходов. Однозначного метода для определения резкости не существует, но есть несколько алгоритмов, которые позволяют вычислить оценку резкости из различных характеристик изображения.



Рис. 2. Примеры оценки контрастности изображений
а) $C = 0,42$; б) $C = 0,28$; в) $C = 0,57$; г) $C = 0,52$

Главное различие всех предлагаемых методов состоит в формулах, которые определяют наличие и количественную оценку контрастности изображения. Первая группа методов для этого вычисляет разностную характеристику напрямую из исходного изображения, например с помощью оператора Лапласа, вторая группа вычисляет частотные характеристики изображения (преобразование Фурье, либо Wavelet-преобразование) и на основании их ищет разностную характеристику изображения.

Идея метода, использующего оператор Лапласа, заключается в вычислении среднего значения яркости от преобразованного оператором изображения. Такой метод чувствителен к качеству изображения, а также к характеру изображенных объектов, то есть он может дать высокую оценку резкости для плохого изображения со сложным объектом и меньшую — для хорошего изображения с малым числом контрастных переходов.

В работе [6] рассмотрен метод оценки резкости на основе преобразования Фурье, суть которого заключается в вычислении эксцесса изображения в частотной области:

$$\gamma_2 = \frac{\mu_4}{\mu_2^2} = \frac{\mu_4}{\sigma^4}, \quad (4)$$

где μ_2 — второй центральный момент изображения, μ_4 — четвертый центральный момент, σ — стандартное отклонение.

В работе [7] также используется преобразование изображения в частотную область, но в качестве критерия оценки резкости предлагается отношение

$$FM = \frac{T_H}{M * N}, \quad (5)$$

где T_H — число значимых точек амплитудного спектра, $M * N$ — размер изображения амплитудного спектра. Значимые точки определяются как точки, значение в которых превышает некоторую долю максимума амплитуды спектра.

Исследование методов показало, что оценка резкости с использованием оператора Лапласа дает заметный разброс результатов при схожих условиях освещения и ракурса, что связано с отсутствием нормирования результата и способом выделения области лица. Оценка (4) показывает более стабильные результаты, но практически не реагирует на фактическое снижение резкости, полученное размытием по Гауссу. Оценка (5) наиболее точно соответствует изначальной задаче выделения поиска резких изображений, она корректно реагирует на изменения резкости при движении и отсутствие фокуса, и помимо этого имеет строгую нормировку, позволяющую по заданному уровню отсеивать изображения неудовлетворительного качества.

На рис. 3 приведены примеры изображений, их амплитудных спектров и полученных оценок по формулам (4) и (5).

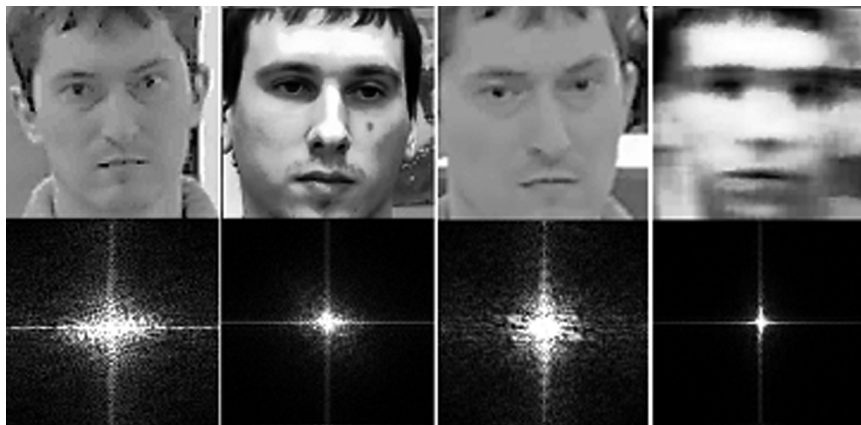


Рис. 3. Примеры оценки резкости изображений:
а) $\gamma_2 = 0,0035$, FM = 0,1301; б) $\gamma_2 = 0,0025$, FM = 0,0388;
в) $\gamma_2 = 0,004$, FM = 0,1203; г) $\gamma_2 = 0,0062$, FM = 0,0111

Несмотря на достоинства оценки (5), она отчасти подвержена тем же проблемам, что и оценка с использованием Лапласиана. Основная причина этого в том, что резкость изображения оценивается количественно и потому связана с общим числом контрастных переходов, а это число зависит от объекта на изображении. Исходя из этого, можно говорить только об относительной резкости изображения, которая оценивается для одного объекта в треке (серии изображений). При этом можно предположить, что метод даст более точную оценку, если при оценке используются изображения тех участков лица, которые в подавляющем большинстве случаев имеют сходные показатели резкости для всех людей, например глаза, нос, губы.

Компенсация освещения. В случае задачи подготовки изображения к распознаванию лиц проблема разных условий съемки может быть решена при помощи объединения трекинга лица в видеопотоке и вычисления оценок для серии изображений одного и того же лица, с последующим выделением кадра с наилучшими значениями. Повысить применимость методов (1)–(5) также может компенсация изображения лица по условиям освещения и ракурса.

Для нормирования исходного изображения по условиям освещения существует метод, описанный в [8]. Он включает в себя гамма-коррекцию, нормирование и DoG-фильтрацию и позволяет максимально выровнять изображение при дисбалансе яркости по освещению, например в случае, когда присутствует боковая засветка лица.

Контроль ракурса. Контроль положения лица человека относительно оптической оси камеры является важной проблемой в связи с чувствительностью алгоритмов распознавания к ракурсу. Наиболее известным методом оценки ракурса является POSIT (Pose from Orthography and Scaling with Iterations), он отличается высокой производительностью и быстрой сходимостью.

Алгоритм POSIT вычисляет матрицу поворота и вектор параллельного переноса так, чтобы ошибка между проекцией характерных точек трехмерной модели на плоскость изображения и соответствующими им размеченными точками на изображении была минимальна. Метод POSIT комбинирует два алгоритма. Первый, POS (Pose from Orthography and Scaling), аппроксимирует перспективную проекцию с помощью масштабированной ортографической проекции и вычисляет матрицу поворота и вектор параллельного переноса

объекта посредством решения СЛАУ. Второй, POSIT (POS with Iterations), в цикле использует результаты работы POS для нахождения более точной масштабированной ортографической проекции характерных точек, и затем применяет POS к полученным точкам, а не к исходным. В итоге алгоритм POSIT сходится к точному решению за несколько итераций [9].

Проблема использования алгоритма на практике заключается в том, что для его работы необходимо изображение, на котором должны быть отмечены характерные точки лица, что в общем случае не является автоматизированной процедурой. Существующие методы автоматической расстановки характерных точек допускают серьезные погрешности и на малых углах поворота не позволяют оценивать ракурс лица с требуемой точностью [10].

При верной расстановке характерных точек лица алгоритм оценки ракурса лица позволяет получить коэффициенты для задачи коррекции ракурса изображений лиц и вычислить нормированное изображение.

Заключение. В результате проведенной работы решены основные задачи подготовки изображения к процессу распознавания лиц и найдены критерии оценки качества изображения.

Предложенный метод трекинга лица в видеопотоке позволяет проводить группировку изображений лиц для последующего сравнительного анализа.

Метод исключения похожих изображений, основанный на оценке уровня корреляции, позволяет исключить из трека изображения лица, не несущие новой информации. Как показала практика, этот метод сокращает число анализируемых изображений приблизительно в 3 раза, сокращая соответственно и ресурсы, требуемые для функционирования системы распознавания лиц.

Проведен анализ существующего алгоритма оценки контраста изображения. Этот алгоритм подходит для проверки условий освещения, позволяя предотвратить использование системы распознавания лиц для малоконтрастных изображений и, как следствие, снизить вероятность ложноположительных срабатываний.

Проведены обзор и сравнение алгоритмов оценки резкости изображения. Оценка резкости позволяет ранжировать изображения в треке по количеству контрастных переходов, в общем случае давая возможность выбрать из всего количества изображений в треке более качественные.

Рассмотрены метод компенсации условий освещения для изображений с неравномерной яркостью, а также метод контроля ракурса на основе алгоритма POSIT.

Применение всех рассмотренных методов и алгоритмов снижает ресурсоемкость систем распознавания личности по изображению лица в несколько раз, а также устраняет влияние неблагоприятных для процесса распознавания факторов.

Список литературы

1. ADINI Y. Face recognition: The problem of compensating for changes in illumination direction // IEEE Transactions on Pattern Analysis and Machine Intelligence. 1997. N 19. P. 721–732.
2. BRONSTEIN A. 3D face recognition using geometric invariants // Proceedings of International Conference on Audio- and Video-Based Person Authentication, 2003.
3. BUI T. T. T. Face Recognition Based on Combination of Wavelet Transforms and Principal Component Analysis // Proceedings of International Forum on Strategic Technology, 2014.

4. ГАГАНОВ В. Сегментация движущихся объектов в видеопотоке // Компьютерная графика и мультимедиа. 2004. № 2(3). [Режим доступа]: <http://cgm.computergraphics.ru/content/view/67>.
5. ФИСЕНКО В. Т. Компьютерная обработка и распознавание изображений: учебное пособие. СПб: СПбГУ ИТМО, 2008.
6. МОНИЧ Ю. И. Оценки качества для анализа цифровых изображений. Минск: Государственное научное учреждение ОИПИ НАН Беларуси, 2008.
7. KANJAR D. Image Sharpness Measure for Blurred Images in Frequency Domain // International Conference on Design and Manufacturing. Procedia Engineering, 2013. P. 149–158.
8. TAN X. Enhanced Local Texture Feature Sets for Face Recognition Under Difficult Lighting Conditions // IEEE Transactions on image processing. 2010. V. 19. N 6. P. 1635–1650.
9. DEMENTHON D. F. Model-based object pose in 25 lines of code // International Journal of Computer Vision. 1995. V. 15. I. 1–2. P. 123–141.
10. MILBORROW S. Multiview Active Shape Models with SIFT Descriptors for the 300-W Face Landmark Challenge // The IEEE International Conference on Computer Vision (ICCV) Workshops. 2013. P. 378–385.

*Небаба Степан Геннадьевич —
аспирант Института кибернетики Томского
политехнического университета,
e-mail: stepan-lfx@mail.ru;*

*Андреев Сергей Юрьевич —
зав. научно-учебной лаборатории
3D моделирования Института кибернетики
Томского политехнического университета;*

*Макаров Максим Александрович —
аспирант Института кибернетики
Томского политехнического университета.*

Дата поступления — 10.11.2014