

УДК 004.932

РАСПОЗНАВАНИЕ ЛИЦ НА ФОТОГРАФИИ ПУТЕМ АНАЛИЗА ХАРАКТЕРНЫХ ОБЛАСТЕЙ

Е.Л. Столов, А.В. Шлянников

Аннотация

В работе предлагается подход к задаче распознавания лиц на фотографии, основанный на гипотезе, согласно которой в основе распознавания лица человеком лежит анализ взаимного расположения наиболее ярких областей на изображении. При сравнении двух фотографий одного лица, сделанных в разных ракурсах, производится искусственное совмещение ракурсов с помощью метода, предложенного авторами ранее. К совмещенным изображениям применяется специальное вейвлет-преобразование, выделяющее наиболее яркие области. После этого преобразованное изображение заменяется своим сжатым изображением, выполненным на основе преобразования Фурье, и обычным образом сравниваются коэффициенты преобразования.

Введение

В настоящее время задача идентификации лица по его фотографии привлекает большое внимание. Она лежит на стыке двух дисциплин: цифровой обработки изображений и искусственного интеллекта. Актуальность данной проблемы не вызывает сомнения, достаточно отметить работы по созданию систем для контроля доступа. Ни один из алгоритмов, предложенных к настоящему времени, не дает полного решения задачи [1]. Первая проблема, с которой начинается исследование, – обнаружение лица на фотографии. Это отдельная задача, которой посвящен обзор [1]. В данной статье предполагается, что уже выделен регион, внутри которого находится изображение только одного лица, поэтому данная задача здесь не рассматривается. На следующем этапе происходит идентификация фотографии. В общей форме задача выглядит следующим образом. Имеется база данных, содержащая некоторое количество фотографий, и конкретный образец. Требуется выделить в базе ту из фотографий (или целый класс фотографий), обладающую наибольшей схожестью с представленным образцом. Сложность задачи обусловлена тем обстоятельством, что иногда в базе данных содержится лишь одна фотография, сделанная в определенном ракурсе, а предъявленная фотография сделана в другом ракурсе. Кроме того, меняются условия освещения, играющие определяющую роль в процессе обработки.

Все используемые методы используют один и тот же прием: тем или иным способом понижают размерность задачи, после чего решение принимается на основе значений некоторой дискриминантной функции. Наибольшей популярностью пользуется метод, известный в статистике как метод главных компонентов (см., например, [2]). Типичные примеры применения этого метода содержатся в работах [3, 4]. В качестве исходного множества векторов берется множество матриц, представляющих цифровую фотографию, «вытянутых» в одну строку, либо векторы, полученные умножением строк этих матриц на некоторую фиксированную матрицу. Перед построением указанных матриц проводится предварительная обработка, состоящая в низкочастотной фильтрации изображения и удалении отдельных пятен. Недостатком указанного метода является то, что сравниваемые

векторы должны иметь одинаковую длину. С этой целью перед началом работы приходится искусственно выравнивать длины векторов, меняя разрешение изображения. Вместо специального базиса, используемого в методе главных компонентов, можно взять любой другой базис и сравнивать коэффициенты разложения. При этом используются не все коэффициенты, а лишь имеющие наибольшие по модулю значения. Разновидность указанного метода применена в данной работе. С математической точки зрения вместо базиса выбирается некоторое множество векторов, как правило, линейно зависимых, а разложение строится на основе проекций исходного вектора (изображения) на вектора этого множества.

В данной работе затрагивается лишь один аспект проблемы: сравнение двух фотографий, сделанных в разных ракурсах. В работе [5] приведены результаты исследования механизма распознавания изображений человеком. В частности, утверждается, что существенное значение имеет взаимное расположение определенных областей (лоб, глаза, брови, нос, рот) и овал лица. Последний параметр легко изменяется одеждой, и поэтому основное внимание уделяется определению указанных областей. В работе [5] установлено, что система восприятия человека легко компенсирует искажения типа пропорционального горизонтального или вертикального растяжения. Компьютерная система распознавания должна компенсировать подобные искажения, поэтому перед сравнением образца и предъявленного объекта следует определить параметры искажений.

В этой связи содержание работы разбито на две части: определение параметров искажения и собственно сравнение образца с изображением, полученным после компенсации искажения. Для решения обеих задач используется метод неортогонального вейвлет-преобразования, предложенный в [6]. В статье [7] показано, что на основе вейвлет-преобразования и последующего преобразования Фурье в принципе можно подсчитать коэффициенты поворота и растяжения, необходимые для последующего сравнения. Там же предлагалось использовать в качестве дискриминантной функции расстояние Хаусдорфа между двумя сжатыми изображениями. В данной работе предлагается параллельно использовать два вейвлет-преобразования: одно – для определения параметров искажения и второе – для сравнения образов после вейвлет-преобразования. С этой целью вводится новая дискриминантная функция, на основе которой принимается решение об идентичности двух фотографий, и находятся ошибки первого и второго рода, возникающие в процессе распознавания.

1. Набор шаблонов характерных областей

Будем рассматриваться только тоновые изображения, поэтому каждое такое изображение определяется вещественной матрицей. В большинстве методов, используемых для сжатия изображения, применяется разложение соответствующей матрицы по некоторому ортогональному базису с последующим исключением членов, несущих мало информации. С другой стороны, в задачах распознавания осуществляется поиск особых областей в изображении, своих в каждом случае, и на основе результатов этого поиска принимается решение. В силу этого обстоятельства набор найденных областей представляет сжатую форму исходного изображения. Этот подход использован, например, в [8], где предлагается создать набор параметрических шаблонов, после чего искать на изображении области, близкие к одному из шаблонов с указанием соответствующих параметров. Примерами таких шаблонов являются окружности, области, получающиеся в результате пересечения двух кругов, области, ограниченные кривыми Безье, и другие. Фактически, речь идет об использовании вейвлет произвольной природы. Недостаток такого подхода

заключается в том, что в отличие от ортогонального случая, поиск шаблонов с последующим подбором параметров требует значительных ресурсов.

В работах [6, 7] использован компромиссный подход к решению аналогичной задачи. В качестве шаблона применяются матрицы ранга 1, построенные по следующему правилу. Строятся векторы X длины N с компонентами $x_k = \sin(\pi k/N)$, $k = 0, \dots, N - 1$, и Y длины M с компонентами $y_k = \sin(\pi k/M)$, $k = 0, \dots, M - 1$. После этого подсчитывается матрица $Mat = X^T \cdot Y$. По построению эта матрица имеет ранг 1. Этой матрице отвечает поверхность, представленная на рис. 1. На основе построенной матрицы строится шаблон $Template(N, M, a, b)$, где N, M определены выше, а a, b задают центр шаблона. Каждый такой шаблон строится следующим образом. Берется нулевая матрица, размеры которой совпадают с размерами исходного изображения. В эту матрицу помещаются элементы матрицы Mat так, чтобы вершина поверхности Mat имела координаты a, b в этой матрице. Варьируя значения параметров N, M , получаем шаблоны, вытянутые вдоль одной или другой оси координат, а изменяя координаты a, b , перемещаем поверхность по изображению. В результате получаем целое семейство шаблонов, с помощью набора которых аппроксимируется исходное изображение. Следует отметить, что рассмотренные шаблоны образуют линейно зависимую систему в линейном пространстве всех матриц заданного размера. Теперь поиск характерных областей на изображении осуществляется следующим образом. Для каждого из наборов чисел N, M из определенного заранее множества строится шаблоны, которые перемещаются по рисунку. Ищется тот из шаблонов, для которого имеет место наилучшее

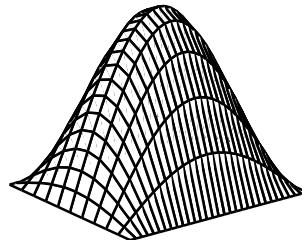


Рис. 1. Вид поверхности, определенной шаблоном

соответствие между шаблоном и накрытой им частью рисунка. Как обычно, в качестве меры совпадения берется корреляция между значениями в точках шаблона и в соответствующих точках изображения. В результате получаем набор значений вида $S = (N_s, M_s, a_s, b_s, c_s)$, $s = 1, \dots, R$, где c_s – коэффициент корреляции, а R – общее число использованных шаблонов. Каждому набору $s \in S$ ставится в соответствие поверхность P_s . Для этого матрица $Template(N_s, M_s, a_s, b_s)$ умножается на коэффициент c_s , после чего все построенные матрицы складываются. Другими словами, исходное изображение заменяется своим приближением P , вычисленным согласно:

$$P = \sum_{s=1}^R P_s, \quad (1)$$

а сам набор шаблонов представляет собой множество наиболее ярких областей изображения. Хотя шаблоны не образуют базу в пространстве матриц, они порождают все множество матриц. Для получения однозначной версии разложения (1) предложено две версии создания аппроксимации: в первом случае на каждом шаге строится новая поверхность, получаемая из предыдущей вычитанием матрицы P_s ,



Рис. 2. Исходное изображение

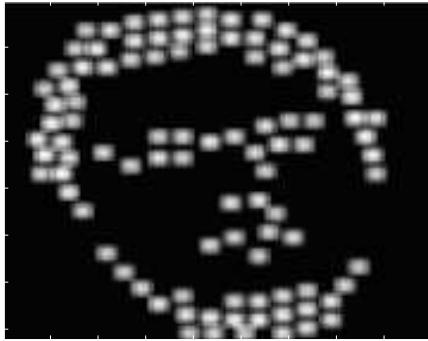


Рис. 3. Результат преобразования набором шаблонов малого размера

найденной на предыдущем шаге; во втором случае вычитание не производится, а найденные поверхности ранжируются по значениям c_s .

Поверхности, выбранные в качестве шаблонов, являются гладкими во всех точках, кроме четырех точек, лежащих в основании. Это обстоятельство гарантирует гладкость поверхности, полученной в результате аппроксимации. Хотя отыскание шаблонов требует значительных затрат ресурсов, в [6] показано, что выбор шаблонов указанного вида позволяет значительно сократить время, необходимое для проведения вычислений. Более того, там же предложен вариант параллельного алгоритма для решения задачи.

2. Отыскание коэффициентов искажения

Если фотография одного лица сделана в разных ракурсах, то для сравнения этих фотографий нужно предварительно произвести преобразования, чтобы совместить их максимально возможным образом. Предлагаемые шаблоны могут быть использованы для этой цели. Соответствующий алгоритм приведен в [7]. Кратко изложим идею этого алгоритма. Выравнивание изображений осуществляется с помощью переносов, поворотов и сжатий. С этой целью для каждого изображения строится представление в виде (1), но при этом используются шаблоны малого размера с $M = N = 3$. В результате изображение заменится набором точек (см. рис. 2 и 3).

Для отыскания параметров преобразования выравнивания используются изображения типа, представленного на рис. 3. Каждое из этих изображений заменяется набором точек – координатами центров шаблонов и в дальнейших преобразованиях рассматривается как бинарное множество. После этого второе изображение совмещается наилучшим способом с первым изображением параллельным переносом, поворотом и шкалированием. В качестве критерия качества совмещения выбрано расстояние Хаусдорфа между этими множествами. Параллельный перенос осуществляется с помощью совмещения центров медиан каждого из рисунков. Для определения угла поворота, необходимого для совмещения изображений, подсчитывается преобразование Фурье от матрицы P в (1) в полярных координатах:

$$F(r, u) = \frac{1}{R} \sum P(x_i, y_j) \exp(-2\pi r(x_i \cos(u) + y_j \sin(u))). \quad (2)$$



Рис. 4. Вариант исходного изображения

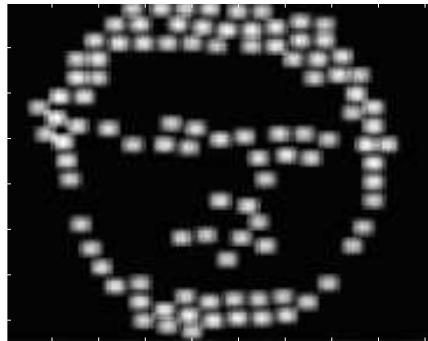


Рис. 5. Результат преобразования набором шаблонов малого размера для варианта фотографии

Заметим, что матрица P в (2) не является бинарной. Пусть $F_1(r, u)$, $F_2(r, u)$ – два преобразования Фурье вида (2), построенные по двум изображениям. В идеальном случае, если одно изображение получено из другого поворотом на угол t , то имеет место соотношение $F_1(r, u) = F_2(r, u - t)$. Эта формула положена в основу определения угла поворота для совмещения. Подсчитывается корреляция между коэффициентами преобразований Фурье для первого и второго изображения, и значение t , для которого эта корреляция достигает максимального значения, выбирается в качестве искомого угла поворота. Коэффициент шкалирования подсчитывается на последнем этапе. Он изменяется с малым шагом до тех пор, пока расстояние Хаусдорфа между множествами продолжает уменьшаться. На самом деле применяется итеративная процедура, в которой совмещение центров медиан и подсчет коэффициента шкалирования чередуются друг с другом. Процедура останавливается, как только перестает происходить дальнейшее уменьшение расстояния Хаусдорфа между множествами. Пример варианта фотографии одной персоны, использованной в экспериментах, и результат ее преобразования представлены на рис. 4 и 5

Для обнаружения таких областей используется то же самое неортогональное вейвлет-преобразование, но теперь шаблон имеет вытянутую форму в горизонтальном или вертикальном направлении. Согласно высказанной гипотезе это соответствует реальной ситуации. Исходным материалом в этом случае является базовая фотография, хранящаяся в базе, и изображение, полученное из предъявленной фотографии после применения преобразований поворота и шкалирования, построенных на предыдущем шаге. Непосредственное сравнение двух фотографий не имеет смысла. Даже малейшее смещение одной фотографии относительно другой приводит к большой величине невязки. В силу этого следует выбирать критерий, инвариантный относительно операции смещения. Модули коэффициентов преобразования Фурье обладают нужным свойством, поэтому они были использованы в процедуре сравнения. Преобразование Фурье применяется не к оригинальным фотографиям а к матрице P в (1), полученной в результате преобразования с «вытянутыми» шаблонами. Примеры результатов преобразований с шаблонами, вытянутыми в горизонтальном направлении, приведены на рис. 6 и 7. Специальный вид шаблона, принятый в работе, позволяет найти преобразование Фурье в явном виде. Действительно, любое слагаемое в (1) имеет вид $P_s = \text{Mat} = X^T \cdot Y$, поэтому $P_s(i, j) = X(i) \cdot Y(j)$. В силу этого двумерное преобразование Фурье от P_s сводится к одномерному преобразованию Фурье. В свою очередь, если исход-

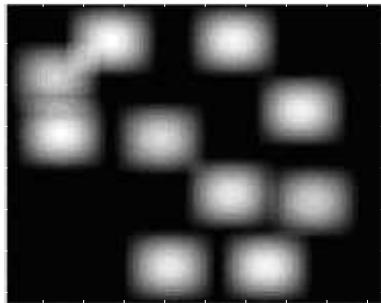


Рис. 6. Результат вейвлет-преобразования фотографии из базы

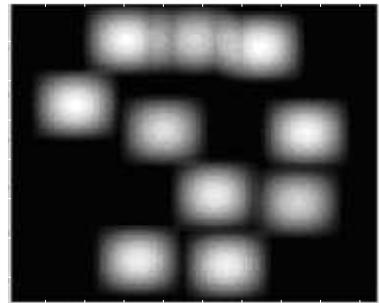


Рис. 7. Результат вейвлет-преобразования варианта фотографии

ное изображение имело размеры $U \times V$, то для нахождения всех сумм достаточно вычислить выражения вида

$$c(m) = \sum_{k=0}^M \exp(-2i\pi km/U) \sin(m\pi/M), \quad d(n) = \sum_{k=0}^N \exp(-2i\pi kn/V) \sin(n\pi/N).$$

Коэффициенты Фурье для остальных слагаемых получаются из $c(m)$ и $d(n)$ умножением на комплексные числа с единичным модулем.

В качестве критерия принадлежности двух фотографий одной персоне выбиралось евклидово расстояние между векторами, состоящими из модулей коэффициентов Фурье, подсчитанных по указанной выше методике. Если это расстояние не превышало некоторый порог T , то принималось решение об идентичности лиц, представленных на двух фотографиях.

Табл. 1

Результаты эксперимента										
ОПР	30	50	50	0	35	40	10	65	55	45
ОВР	10	2	2	1	10	6	9	4	6	4

Для тестирования была выбрана база, содержащая 1012 фотографий. Структура базы сформирована таким образом, что для каждого лица в ней содержится по 20 разных фотографий. Эти фотографии отличаются ракурсом, с которого была выполнена съемка. Для определения порогового значения метрики выбирался тестовый набор фотографий одного и того же лица, но с некоторыми искажениями (поворот лица, различная освещенность, ручное искажение фотографии – кадрирование, масштабирование, поворот в плоскости изображения), содержащий 15 фотографий. Затем выбиралась еще одна фотография того же лица, не присутствующая в этом наборе, и фиксировалась как базовое изображение. Далее это исходное изображение попарно сравнивалось с каждым изображением из тестового набора, и определялись значения метрики, полученные для каждой пары изображений при сравнении. В качестве порога T выбиралось максимальное из полученных значений метрики. Таким образом, при выбранном пороге и данном наборе фотографий ошибки первого рода отсутствуют.

Результатом каждого сравнения являлось значение метрики. Итоги нескольких экспериментов приведены в табл. 1. При вычислении ошибки первого рода (ОПР)

использовались все 20 фотографий одной персоны, при вычислении ошибки второго рода (ОВР) принимались во внимание все оставшиеся 992 фотографии. Ошибки первого и второго рода приведены в процентах.

Из приведенных результатов следует, что предложенная техника дает неплохие результаты для ошибок второго рода. Полученные оценки ошибок первого рода оказались значительными. С точки зрения статистики это связано прежде всего с малым количеством фотографий одной персоны по сравнению с объемом всей базы. Такая ситуация типична при практическом применении автоматического распознавания. Увеличив значение порога T , можем уменьшить ОВР, но при этом возрастет значение ошибок второго рода. Поскольку окончательное заключение об идентификации делает человек, такое увеличение не имеет практического смысла. Анализ фотографий, для которых получены неверные результаты, показывает, что предложенная методика плохо компенсирует повороты головы вокруг вертикальной оси на значительный угол. Эта проблема является предметом дальнейших исследований.

Summary

E.L. Stolov, A.V. Shlynikov. Recognition face in picture by means of analysis of special regions.

A new approach to face recognition problem is suggested. The approach is based on the hypothesis that human beings recognize person in picture by means of an analysis of brightest spots in picture. Comparing two pictures of the same person one has to level sizes of the pictures. To this end a special procedure that was developed by the authors earlier is implemented. After the leveling is completed the brightest regions in the picture are detected by the technique of non-orthogonal wavelet transform. Each transformed picture is replaced by its Fourier transform. The final decision based on a compare of the found Fourier coefficients.

Литература

1. *Ming-Hsuan Y., Kriegman D.J., Ahuja N.* Detecting faces in images: a survey // IEEE Trans. Pattern Analysis and Machine Intelligence. – 2002.– V. 24. – P. 34–57.
2. *Айвазян С.А. и др.* Классификация и снижение размерности. – М.: Финансы и статистика, 1989. – 607 с.
3. *Song H., Chung S.J., Park Y.H.* An online face recognition system using multiple compressed images over the internet // Web Information Systems Engineering – WISE 2005. Lecture Notes in Computer Science. V. 3806. – Berlin-Heidelberg: Springer, 2005. – P. 569–576.
4. *Yang J. et al.* Two-dimensional PCA: a new approach to appearance-based face representation and recognition // IEEE Trans. Pattern Analysis and Machine Intelligence. – 2004. – V. 26 – P. 131–137.
5. *Sinha P., Balas B., Ostrovsky Y., Russell R.* Face Recognition by Humans: Nineteen Results All Computer Vision Researchers Should Know About // Proc. IEEE. – 2006. – V. 94. – P. 1948–1961.
6. *Столов Е.Л.* Параллельный алгоритм для отыскания водяных знаков в изображении // Вестн. Казан. гос. техн. ун-та им. А.Н. Туполева. – 2006. – № 3. – С. 37–40.
7. *Столов Е.Л., Шляников А.В.* Алгоритм распознавания лиц на фотографии с помощью вейвлет преобразования специального вида // Моделирование процессов. Тр. Казан. науч. семинара «Методы моделирования». – 2007. – № 3. – С. 87–95.

-
8. *Candes E.J., Donoho D.L.* New tight frames of curvelets and optimal representations of objects with piecewise C^2 singularities // Comm. Pure Appl. Math. – 2004. – V. 57. – P. 219–266.

Поступила в редакцию
29.08.07

Столов Евгений Львович – доктор технических наук, профессор кафедры системного анализа и информационных технологий Казанского государственного университета.
E-mail: *Yevgeni.Stolov@ksu.ru*

Шлянников Александр Вадимович – аспирант кафедры системного анализа и информационных технологий Казанского государственного университета.