

Поиск портретных изображений по содержанию

А. Гончаров
ag.tsure@gmail.com

А. Горбань
alex.gorban@gmail.com

А. Каркищенко
karkish@tsure.ru

А. Лепский
lepский@mail.ru

Таганрогский технологический институт Южного Федерального Университета
Лаборатория Математических проблем искусственного интеллекта

Аннотация

Проблема поиска портретных изображений по содержанию состоит из двух основных задач: задачи обнаружения лица на изображении (детекция лиц) и задачи сравнения лица-запроса с лицевыми изображениями из базы графических файлов (идентификация лиц). В данной работе рассматриваются методы детекции и идентификации лиц, основанные на переходе от исходного изображения к специальному представлению — так называемой матрице изменения яркостей, элементами которой являются знаки частных производных от функции яркости изображения. Показана устойчивость метода к изменениям яркостной составляющей изображения. Приведены результаты тестирования предложенных алгоритмов детекции и идентификации лиц на распространенных базах изображений лиц, позволяющие сделать вывод о перспективности предложенных методов.

1. Введение

На сегодняшний день проблема идентификации человека по изображению лица привлекает внимание большого количества исследовательских групп в области обработки изображений и распознавания образов. Эта задача получила бурное развитие за последние десять лет, о чем свидетельствуют опубликованные 27 марта 2007 г. результаты широкомасштабных исследований современных алгоритмов биометрической идентификации человека [22], проведенных Национальным институтом стандартов и технологий США (National Institute of Standards and Technology). В рамках исследования рассматривались три подхода биометрической идентификации, а именно идентификация по изображению радужной оболочки глаза, по трехмерной модели головы человека и по обычному изображению лица.

Задача биометрической идентификации человека по изображению лица состоит из двух основных задач: задачи обнаружения лица на изображении и задачи распознавания лица, т.е. вычисления меры близости между различными изображениями лиц. При этом возникают ошибки двойного характера. Ошибка первого рода (FAR — False Accept Rate) соответствует вероятности того, что алгоритм распознавания ошибочно отнесет неизвестное лицо к одному из известных лиц. Ошибка второго рода (FRR — False Reject Rate) характеризует вероятность того, что алгоритм распознавания известное лицо ошибочно посчитает неиз-

вестным. В практических приложениях обычно фиксируют допустимый уровень ошибки первого рода и стараются минимизировать ошибку второго рода.

Согласно приведенным результатам современные системы распознавания лиц способны идентифицировать лица с высокой надежностью, при этом надежность распознавания существенно зависит от размеров лица, его расположения и условий освещенности. Под размером лица понимается количество пикселей между центрами зрачков. Так, при работе лучших алгоритмов в условиях контролируемой освещенности для изображений со сверхвысоким разрешением [22] (размер лица 400 пикселей и более) при ошибке первого рода 0,1 % ошибка второго рода составляет всего 1 %. В условиях неконтролируемой освещенности на изображениях сверхвысокого разрешения при ошибке первого рода 0,1 % ошибка второго рода достигает 10–15 %.

Отметим, что приведенные результаты достигаются на фронтальных лицах из нормализованных баз изображений лучшими на сегодняшний день алгоритмами распознавания лиц [22].

В настоящее время для решения задачи распознавания лиц применяются алгоритмы, основанные на статистическом анализе изображений, Марковских цепях, эластичных графах, вейвлет-анализе, анализе характерных точек и алгоритмы, основанные на нейронных сетях. Дадим краткую характеристику наиболее популярным методам.

Метод главных компонент (Principal Component Analysis, PCA) [21, 19, 18] является, по сути, статистическим методом и применяется для сжатия информации без существенных потерь. Основная идея метода заключается в линейном ортогональном преобразовании входного вектора большой размерности в выходной вектор малой размерности с некоррелированными координатами. В терминах задачи распознавания лиц входные векторы представляют собой центрированные и приведенные к единому масштабу изображения лиц. Для каждого изображения лица вычисляются его главные компоненты, при этом процесс распознавания заключается в сравнении главных компонент известного изображения с компонентами всех остальных изображений. Основным недостатком этой группы методов заключается в высоких требованиях к условиям съемки и необходимости качественной предварительной обработки изображений.

Линейный дискриминантный анализ (линейный дискриминант Фишера, Linear Discriminant Analysis, LDA) [18, 11, 17] заключается в выборе проекции пространства изображений на пространство признаков таким образом, чтобы минимизировать внутриклассовое и максимизировать межклассовое расстояние в пространстве признаков. Для данной группы методов свойственна высокая точность распознавания.

Методы на основе скрытых Марковских моделей [20, 15] являются одним из наиболее мощных инструментов для решения задачи распознавания лиц. Впервые идея применения скрытой одномерной Марковской модели была предложена Samaria [24]. Каждая модель представляет собой набор состояний, между которыми возможны переходы. В каждый момент времени система находится в строго определенном состоянии. В наиболее распространенных Марковских моделях первого порядка полагается, что каждое последующее состояние зависит только от текущего состояния. В распознавании образов скрытые Марковские модели применяются следующим образом. Каждому классу изображений соответствует своя модель. Распознаваемый образ представляется в виде последовательности наблюдений, и для каждой модели вычисляется вероятность того, что эта последовательность могла быть сгенерирована именно этой моделью. Недостатком этого подхода является то, что скрытые Марковские модели обладают плохой различающей способностью, т. е. алгоритм обучения максимизирует отклик на свои классы, но не минимизирует отклик на другие классы.

Итак, мы бегло описали лишь основные подходы к решению задачи распознавания лиц. На самом деле, существует большое число оригинальных подходов и методов, заслуживающих отдельного внимания. Подробные обзоры современных методов можно найти в следующих работах [15, 5, 28].

Большинство работ по распознаванию лиц направлены на решение задачи биометрической идентификации человека в охранных системах и системах ограничения доступа. Специфика этих задач заключается в том, что алгоритму распознавания предъявляется изображение, полученное при определенных и заранее известных условиях съемки, поскольку не составляет труда оборудовать контрольно-пропускной пункт необходимым количеством источников освещения и видеокамер, обеспечивающих требуемое качество изображения. Такая свобода построения конфигураций позволяет решать задачу восстановления трехмерного изображения лица, а также осуществлять более точную идентификацию не по одному снимку, а по видеоряду.

Рассмотрим задачу распознавания лиц с точки зрения поиска портретных изображений в Интернете. Под портретными изображениями будем понимать изображения (как цветные, так и полутонные), содержащие произвольное количество фронтальных лиц, не скрытых за посторонними предметами и охватывающие основные лицевые признаки, такие как брови, глаза, нос и рот. При

обработке изображений из Интернета, как правило, мы имеем дело с изображениями низкого разрешения (размер лица 75–110 пикселей) и сверхнизкого разрешения (размер лица менее 75 пикселей), полученными при неконтролируемых условиях освещения. Кроме того, относительно редко встречаются строго фронтальные лица и изображения с нейтральным выражением лица. Наличие эмоций и отклонение от фронтального вида существенно усложняют задачу распознавания лица. Таким образом, ожидать столь высокой надежности от алгоритмов распознавания лиц в контексте Интернет поиска не приходится. В данной постановке целесообразно говорить о поиске нескольких лиц, наиболее похожих на лицо-запрос. Примером реализации подобной концепции поиска портретных изображений может служить популярный проект myheritage.com, позволяющий загрузить изображение лица и поискать наиболее похожие на него фотографии знаменитостей.

Основная проблема при работе с изображениями низкого и сверхнизкого разрешения связана с невозможностью аккуратной локализации основных лицевых признаков (таких как глаза, брови, нос и рот) из-за погрешностей дискретизации растрового изображения, наличия бликов, теней или помех, что, в свою очередь, не позволяет применять достаточно точные алгоритмы, оперирующие с этими признаками. В этом случае наиболее распространенной и робастной техникой является переход от исходного изображения к его специальному представлению [15]. Примерами такого перехода могут служить применение преобразования Фурье [25] или вейвлет-преобразования [13].

В данной работе рассматривается новый подход, предполагающий переход от исходного изображения к так называемой *матрице изменения яркостей*, что обеспечивает устойчивость алгоритмов детекции и идентификации лиц к вариации условий освещенности и не требует этапов предварительной обработки изображения, таких, как выравнивание гистограммы и нормализация яркости.

2. Мера близости на изображениях

В задачах детекции и распознавания лиц ключевым моментом является введение меры близости на изображениях. При этом мера близости по возможности должна удовлетворять аксиомам метрики и, кроме того, соответствовать особенностям восприятия зрительной системой человека. При анализе изображения лица, человек обращает слабое внимание на искажение яркостной составляющей изображения, на наличие теней и бликов, а анализирует характер изменения яркости, придавая большее значение качественным различиям, а не количественным. Классические метрики, такие как L_1 и L_2 , плохо подходят для сравнения изображений лиц, даже в том случае, когда осуществлен переход от исходного изображения к какому-либо специальному представлению.

Примером подходящей меры близости для сравнения изображений лиц может служить семейство

псевдометрик, определяемых следующим выражением:

$$\rho_w(I, J) = \sum_{x_1=1}^n \sum_{x_2=1}^m w_{x_1, x_2} \prod_{i=1}^2 \left| \operatorname{sgn} \frac{\partial I}{\partial x_i} - \operatorname{sgn} \frac{\partial J}{\partial x_i} \right|, \quad (1)$$

где $I = I(x_1, x_2)$, $J = J(x_1, x_2)$ – полутоновые изображения размером $n \times m$, θ – функция Хэвисайфа, w_{x_1, x_2} – весовые коэффициенты, $w_{x_1, x_2} \geq 0$.

Можно показать, что введенная мера близости удовлетворяет следующему набору аксиом:

1. $\rho_w(I, J) \geq 0$
2. $I = J \Rightarrow \rho_w(I, J) = 0$;
3. $\rho_w(I, J) = \rho_w(J, I)$;
4. $\rho_w(I, J) \leq \rho_w(I, X) + \rho_w(X, J)$.

Меру близости, удовлетворяющую указанному набору аксиом в классической литературе [8] принято называть *псевдометрикой*, в силу того, что вторая аксиома выполняется в одну сторону, т. е. из равенства меры близости нулю, вообще говоря, не следует равенство изображений. В то же время, в некоторых современных работах [9] меры близости, удовлетворяющие аксиомам 1–4 называют *метрикой*. В рамках данной работы будем придерживаться классической трактовки понятия метрики, и в дальнейшем называть меру близости (1) псевдометрикой.

Как следует из выражения (1), введенная псевдометрика оперирует со знаками частных производных, что придает устойчивость мере близости относительно вариации яркостной составляющей изображения.

Рассмотрим важное свойство введенной псевдометрики. Пусть I, J полутоновые изображения размера $n \times m$, т. е. $I, J \in \mathbb{N}^{n, m}$. Введем отношение эквивалентности \sim между изображениями. Будем говорить, что изображение I эквивалентно J , если $\rho_w(I, J) = 0$. Определим множество σ_I , состоящее из всех эквивалентных I изображений $\sigma_I = \{J \mid J \sim I, J \in \mathbb{N}^{n, m}\}$. Множество σ_I будем называть *классом эквивалентности* по отношению \sim .

Тогда имеет место следующее **утверждение**:

Изображение $J \in \sigma_I$ тогда и только тогда, когда $J(x_1, x_2) = f(I(x_1, x_2))$, где $f(t)$ – строго монотонно

возрастающая функция на \mathbb{N} , т. е. для любых $a, b \in \mathbb{N}$, $a < b$ выполняется неравенство $f(a) < f(b)$.

Рассмотрим некоторые примеры использования данного свойства. Если у исходного изображения изменить яркость и/или контрастность, то с точки зрения введенной псевдометрики исходное и искаженное изображения будут не различимыми. Кроме того, ограничениям, накладываемым на функцию f , удовлетворяют монотонные искажения яркостной характеристики изображения, которые имеют место при формировании растровых изображений цифровыми фотосенсорами [6] и соответствуют особенностям восприятия зрительной системой человека [7]. В большинстве случаев эти искажения носят логарифмический характер.

3. Представление изображения в виде матрицы изменения яркостей

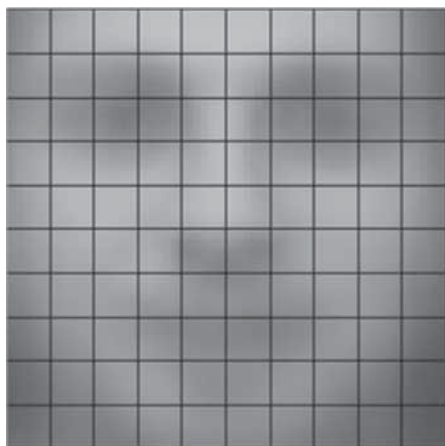
Для решения задач детекции и идентификации лиц с помощью псевдометрики (1) удобно осуществить переход от исходного изображения к *матрице изменения яркостей*, элементами которой являются пары чисел, соответствующие знакам частных производных от яркости исходного изображения в каждой точке:

$$M_{i,j} = \left[\operatorname{sgn} \frac{\partial I}{\partial x_1} \Big|_{(x_i, x_j)}, \operatorname{sgn} \frac{\partial I}{\partial x_2} \Big|_{(x_i, x_j)} \right],$$

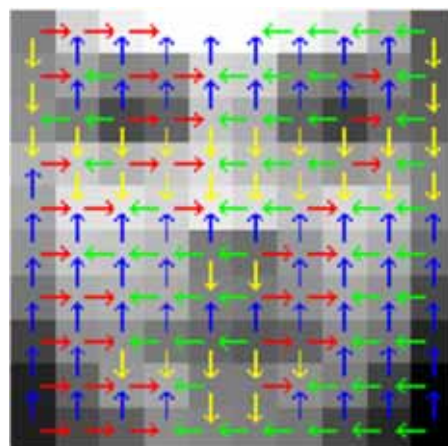
$$\frac{\partial I}{\partial x_1} \Big|_{(x_i, x_j)} = \frac{I(x_{i+1}, x_j) - I(x_i, x_j)}{h_1}.$$

На рис. 1 изображены шаблон лица и соответствующая ему матрица изменения яркостей.

Отметим, что с вычислительной точки зрения построение матрицы изменения яркостей сводится к выполнению простейшей операции поэлементного сравнения двух массивов, а вычисление псевдометрики (1) фактически соответствует вычислению метрики Хэмминга для матриц изменения яркостей, т. е. подсчету количества совпадающих элементов.



а)



б)

Рис. 1. Шаблон лица (а) и визуальное представление соответствующей ему матрицы изменения яркостей (б). Стрелками показаны направления изменения яркостей в каждой точке

4. Детекция лиц на изображениях

Прежде чем решать задачу распознавания лиц, необходимо детектировать лицо на изображении, т. е. найти область изображения, содержащую основные лицевые признаки, такие как глаза, брови, рот и нос, и не содержащую элементов фона.

В рамках данной работы предполагается осуществлять процедуру детекции лиц путем сравнения участков входного изображения с шаблоном лица на основе введенной псевдометрики (1). В качестве шаблона лица выступает «усредненное» лицо (см. рис. 1а), построенное по обучающей выборке следующим образом:

$$T = \frac{1}{N} \sum_k F_k,$$

где F_k — изображения лиц, а N — количество лиц в обучающей выборке.

Поскольку на изображениях могут находиться лица разного размера, то для осуществления детекции необходимо просмотреть входное изображение на нескольких масштабах.

Как и в задаче распознавания лиц, в задаче детекции также встречаются ошибки двоякого характера, при этом ошибке первого рода соответствует вероятность не найти лицо (FRR), а ошибке второго рода — вероятность ошибочной классификации изображения как лицевого (FAR).

Как следует из выражения (1), одним из параметров при вычислении меры близости является размер сравниваемых изображений. Очевидно, что, варьируя размер изображений посредством масштабирования, мы, вообще говоря, будем получать различный результат сравнения. Кроме того, от размера сравниваемых изображений зависит и объем вычислительных затрат, необходимых для вычисления псевдометрики (1). Таким образом, возникает задача выбора оптимального размера шаблона лица с точки зрения вычислительных затрат при ограничении на ошибки детекции первого и второго рода.

Для оценки зависимости ошибок FRR и FAR от размера классификатора рассмотрим две выборки, составленные из изображений лиц Ω_F и изображений, не являющихся лицами Ω_N . Ввиду того, что элементы выборок случайны, значения меры близости между шаблоном лица T и изображениями из выборок можно считать случайными величинами:

$$\xi = \rho_w(T, I), I \in \Omega_F, \eta = \rho_w(T, I), I \in \Omega_N.$$

Поскольку на изображения влияют множество независимых случайных факторов, и мера близости вычисляется как сумма случайных событий, будем считать, что соответствующие случайные величины ξ, η распределены по нормальному закону:

$$f_F(x) = \frac{1}{\sqrt{2\pi}\sigma_F} \exp\left(-\frac{(x-m_F)^2}{2\sigma_F^2}\right),$$

$$f_N(x) = \frac{1}{\sqrt{2\pi}\sigma_N} \exp\left(-\frac{(x-m_N)^2}{2\sigma_N^2}\right),$$

где в качестве параметров распределения используются статистические оценки:

$$m = \frac{1}{N} \sum_{i=1}^N \xi_i, \sigma^2 = \frac{1}{N-1} \sum_{i=1}^N (\xi_i - m)^2.$$

В качестве оценок для ошибок детекции первого и второго рода возьмем следующие величины:

$$\alpha = \int_{T_F}^1 f_F(x) dx, \beta = \int_0^{T_F} f_N(x) dx, 0 < T_F < 1,$$

где T_F — пороговая величина принятия лица, выбираемая из условия равенства ошибки первого рода заданной величине.

На рис. 2 представлена зависимость ошибки второго рода от размера классификатора при заданном уровне ошибки первого рода.

Оценка зависимости ошибок детекции лиц от размера классификатора проводилась на изображениях базы BioID [16, 1], краткое описание которой дано в пункте 5.

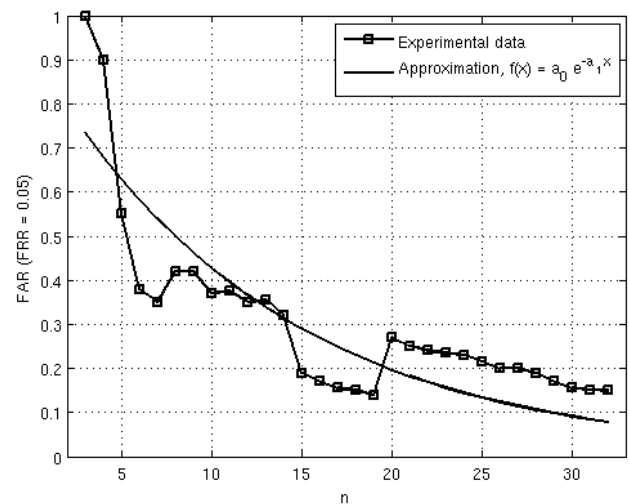


Рис. 2. Зависимость ошибки второго рода FAR от размера классификатора n при ошибке первого рода FAR = 0.05, и аппроксимация экспериментальных данных экспоненциальной функцией

Как видно из рис. 2, с увеличением размера классификатора ошибка детекции снижается. Немонотонное поведение графика ошибки объясняется влиянием существенной неопределенности входных данных, например, случайным характером дискретизации, обусловленным положением лица относительно растровой сетки при масштабировании изображений и др.

Необходимо отметить, что количество арифметических операций, требуемых для вычисления меры близости между одним участком изображения и шаблоном лица, соответствует $O(n^2)$, при этом, для поиска лица на изображении размером $w \times h$ потребуется $O(n^2hw)$ операций. Считая размер изобра-

жения постоянным, и выражая его через размер шаблона, получим, что сложность алгоритма детекции равна $O(n^4)$.

Одним из способов сокращения вычислительных затрат является применение каскадной классификации [27, 12]. Суть этого подхода заключается в последовательном применении классификаторов к изображению (рис. 3). При этом каждый последующий классификатор вычисляется только для той области изображения, в которой отклик предыдущего классификатора превосходит заданное пороговое значение.

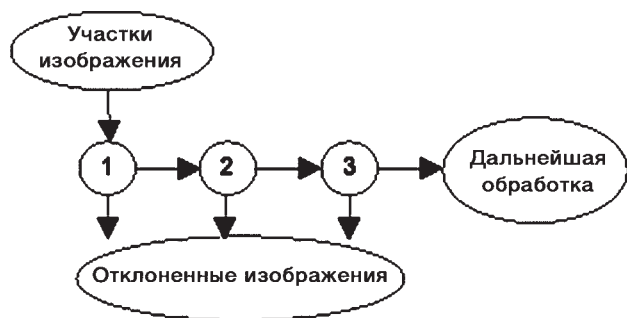


Рис. 3. Схематическое описание каскадной классификации. Первый классификатор отсеивает большое количество участков изображения, не являющихся лицами. Последующие классификаторы осуществляют более тонкую проверку для оставшихся участков изображения. Участки изображения, преодолевшие все стадии классификации, считаются лицами

Для устранения путаницы, необходимо обратить внимание на то, что по тем или иным причинам задачу распознавания образов иногда описывают в терминах мер близости, а иногда в терминах фильтров и классификаторов. В первом случае, близким образом соответствует маленькая мера близости, в то время как в терминологии классификаторов искомому образу соответствует высокое значение отклика классификатора. Поскольку в данном случае речь идет о каскадной классификации, здесь используется терминология классификаторов.

Ввиду того, что выбором весовых коэффициентов можно осуществить нормировку псевдометрики (1), соответствующий ей классификатор можно записать следующим образом:

$$H(I) = 1 - \rho_w(I, T).$$

На основе полученной зависимости ошибок детекции от размера классификатора рассмотрим задачу минимизации вычислительных затрат при ограничениях на ошибку второго рода при фиксированной ошибке первого рода.

Можно показать, что количество арифметических операций, необходимых для каскадной детекции, выражается следующим образом:

$$Q(s) = c \sum_{j=1}^p s_j^4 \prod_{k=1}^{j-1} \beta(s_k),$$

где $s = (s_1, s_2, \dots, s_p)$ — вектор размеров классификаторов, $s_j \geq 2$, $\beta(s)$ — зависимость ошибки второго рода

от размера классификатора при фиксированной ошибке первого рода, c — постоянная величина.

Поскольку при каскадной детекции лиц каждый последующий классификатор применяется лишь к той области, для которой отклик предыдущего классификатора превосходит заданное пороговое значение, ограничение на ошибку детекции накладывается только на последний классификатор в каскаде:

$$\beta(s_p) \leq \beta_0.$$

Таким образом, необходимо решить следующую оптимизационную задачу:

$$\begin{cases} \sum_{j=1}^p s_j^4 \prod_{k=1}^{j-1} \beta(s_k) \rightarrow \min \\ \beta(s_p) \leq \beta_0, \\ s_j \in \mathbb{N}, \\ s_j \geq 2. \end{cases} \quad (2)$$

Как следует из (2), целевая функция является нелинейной, поэтому ограничение на дискретность решения существенным образом усложняет задачу. Поэтому, найдем непрерывное решение поставленной задачи и в окрестности точки минимума выберем такое дискретное решение, которому соответствует минимальное значение целевой функции. Кроме того, экспериментально было показано, что зависимость ошибки второго рода от размера классификатора $\beta(s)$ экспоненциально убывает с ростом s , поэтому для решения задачи (2) аппроксимируем зависимость функцией

вида $f(x) = a_0 e^{-a_1 x}$ методом наименьших квадратов по экспериментальным данным (см. рис. 2). Тогда оптимизационная задача примет вид:

$$\begin{cases} \sum_{j=1}^p s_j^4 a_0 \exp\left(-a_1 \sum_{k=1}^{j-1} s_k\right) \rightarrow \min \\ a_0 e^{-a_1 s_p} \leq \beta_0, \\ s_j \geq 2. \end{cases} \quad (3)$$

Решение непрерывной задачи (3) было получено с помощью пакета Optimization Toolbox системы MATLAB. На рис. 4 представлены графики зависимости ускорения вычислений процедуры детекции лиц от количества классификаторов в каскаде.

5. Результаты тестирования алгоритма детекции лиц

Предложенный алгоритм детекции лиц на основе псевдометрики был протестирован на таких базах изображений лиц, как ORL [2, 23], BioID [16, 1], Yale [4], Yale B [14, 3], Libor Spacek [26], а также на базе изображений из Интернета, предоставленной компанией Яндекс.

Дадим краткое описание используемых баз изображений.

База ORL состоит из изображений 40 персон, по 10 изображений для каждой персоны. На каждом

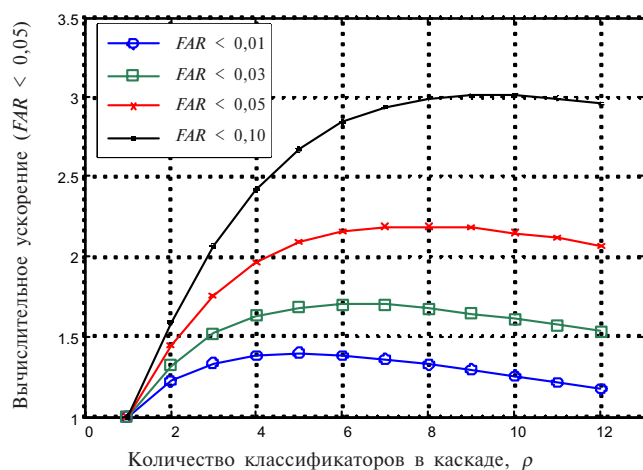


Рис. 4. Зависимость ускорения вычислений от количества классификаторов при каскадной детекции (при ограничении на ошибку второго рода $FAR \leq 0.05$)

изображении присутствует одно лицо на однородном фоне, все изображения сделаны в одинаковых условиях освещения при различных положениях головы. Некоторые лица существенно отклонены от фронтального вида и содержат ярко выраженные эмоции.

Изображения базы BioID содержат лица различного масштаба, оснащенные дополнительной информацией о координатах центров зрачков на изображении, что позволяет упростить оценку надежности алгоритмов детекции. Основным недостатком этой базы является слабая репрезентативность, поскольку она содержит изображения лишь 23 персон, кроме того, количество изображений для каждой персоны может существенно изменяться.

База изображений Yale содержит по 11 изображений 15 человек при различных условиях освещения и предназначена для тестирования алгоритмов распознавания лиц. Каждый человек сфотографирован в очках и без очков, с различными эмоциями на лице.

База изображений Yale В предназначена для тестирования алгоритмов детекции лиц при различных условиях освещенности. Изображения лиц 10 персон сделаны в студийных условиях для 65 различных направлений освещения и 9 положений головы, итого 5850 изображений. При тестировании алгоритма детекции из базы было выбрано 650 фронтальных изображений лиц, оснащенных дополнительной информацией о координатах центров зрачков для автоматической проверки качества детекции.

Изображения из базы, предоставленной Либором Спасеком (Libor Spacek), на наш взгляд, являются

наиболее репрезентативными. База содержит по 20 изображений лиц 375 персон, как на однородном, так и на неоднородном фоне при различных условиях освещения. При этом имеют место изображения лиц со значительными отклонениями от фронтального вида и наличием эмоций.

База изображений, предоставленная компанией Яндекс, представляет собой набор уменьшенных копий изображений, встречающихся в Интернете. При этом все изображения отмасштабированы таким образом, чтобы их размер не превышал 120×120 пикселей. Количество изображений, содержащих фронтальные и близкие к фронтальным лица, составляет примерно 3–5 % от общего числа изображений базы. Для тестирования алгоритма детекции лиц использовалась выборка, содержащая 20000 изображений, на которых вручную помечены координаты центров зрачков.

На сегодняшний день нет общепринятой методики оценки ошибок детекции лиц. В некоторых работах для выражения ошибки второго рода (FAR) в процентах, количество неправильно классифицированных изображений делится на количество изображений в базе. В других работах при расчете того же показателя в знаменателе стоит не количество изображений, а количество сравнений шаблона лица с участками анализируемого изображения. Например, при анализе изображения размером 640×480 шаблоном лица размером 20×20 для поиска лиц произвольного размера необходимо просмотреть 12–15 масштабов, что в итоге сводится примерно к 100 тыс. сравнений шаблона лица с участками изображения.

В табл. 1 представлены результаты тестирования алгоритма детекции лиц в виде абсолютных значений ошибок, т.е. количество участков изображений, ошибочно принятых за лица для ошибки второго рода (FAR), и количество не найденных лиц для ошибки первого рода (FRR).

В таблице используются следующие обозначения:

N_{img} — количество изображений в базе;

N_F — количество лиц на изображениях;

F_{sz} — средний размер лица (количество пикселей между центрами зрачков);

FAR — количество участков изображений, ошибочно распознанных как лица;

FRR — количество не найденных лиц.

6. Распознавание лиц на основе псевдометрики

Задача распознавания лиц заключается в вычислении меры близости между лицом-запросом и всеми ли-

Таблица 1

Результаты тестирования алгоритма детекции лиц

База	Параметры изображений	N_{img}	N_F	F_{sz}	FAR	FRR
ORL	92×112 , 8Gray	400	400	32	3	3
BioID	384×286 , 8Gray	1521	1521	55	168	78
Yale B	640×480 , 8Gray	650	650	92	151	138
Libor Spacek	200×200 , 24RGB	7900	7900	38	27	45
Яндекс	120×120 , 24RGB	20000	553	13	823	27

цами из базы изображений и поиске наиболее близких лиц, либо таких лиц, для которых мера близости не превышает заданного порогового значения. В качестве меры близости, как и при решении задачи детекции, использовалась введенная псевдометрика (1).

Основные отличия в применении псевдометрики в задачах детекции и идентификации лиц связаны с изменением размеров сравниваемых образов и используемых весовых коэффициентов. Очевидно, что для процедуры детекции лиц достаточно маленького размера изображений (размер лица 20–30 пикселей), поскольку требуется уловить только основные особенности лица. Что касается процедуры идентификации лиц, то здесь нам необходимо отличать одно лицо от другого, следовательно, требуется более подробное изображение (размер лица более 30 пикселей).

В рамках данной работы будем ориентироваться на задачу распознавания лиц в контексте поиска наиболее похожих изображений из базы лиц на изображение-запрос. Тогда работа алгоритма поиска портретных изображений состоит из следующих основных этапов:

Детекция лиц на изображении-запросе.

Вычисление матрицы изменения яркостей подходящего размера для всех найденных лиц.

Сравнение матриц изменения яркостей изображения-запроса со всеми матрицами изменения яркостей из базы изображений лиц.

Предоставление пользователю первых N наиболее релевантных изображений для каждого найденного лица на изображении-запросе.

Необходимо отметить, что одним из недостатков методов распознавания на основе низкоуровневых признаков (в данном случае речь идет о знаках частных производных от функции яркости изображения) является отсутствие аффинной инвариантности, т.е. сдвиг одного изображения относительно другого приводит к существенному искажению результата сравнения. Такое смещение изображений главным образом связано со случайным характером дискретизации, обусловленным расположением лица относительно растровой сетки при масштабировании изображения. Кроме того, изменение ракурса съемки и наклон головы также вносят свой вклад в неточность положения лица внутри детектируемой области. На

рис. 5 представлены примеры неточностей детекции, которые в результате могут оказывать значительное влияние на качество распознавания.

Наличие неточностей положения лица на изображении приводит к тому, что алгоритм сравнения изображений реагирует не столько на изменение во внешности, сколько на положение лица в кадре, что в свою очередь является дополнительным источником ошибок распознавания.

Для устранения этого негативного фактора можно предложить два подхода. Первый заключается в дополнительной локализации лица в кадре, например посредством поиска координат центров зрачков и последующей корректировкой изображения. Второй способ состоит в сравнении двух изображений для различных относительных смещений и углов наклона и выборе минимального расстояния (наилучшего совпадения) между сравниваемыми образами.

Очевидным недостатком первого подхода является невозможность точного определения координат центров зрачков на изображениях низкого разрешения (см. рис. 5).

Второй подход влечет за собой дополнительную вычислительную нагрузку. Так для сравнения лиц для трех различных углов наклона и относительных смещений в три пикселя по каждой координате потребуется осуществить примерно полтысячи обычных сравнений, а если сюда добавить еще и небольшую вариацию в масштабе, то количество сравнений может возрасти до нескольких тысяч.

Естественным шагом в данном случае является уход от полного перебора к методам направленного поиска. На рис. 6 представлена зависимость меры близости (1) от относительного смещения изображений.

Идея метода направленного поиска глобального минимума заключается в вычислении меры близости для смещений в окрестности нулевого отклонения и дальнейшее движение в направлении максимального убывания меры близости (в соответствии с методом наискорейшего спуска). Как известно, градиентные методы не всегда сходятся к глобальному экстремуму, в силу чего происходит некоторое снижение надежности распознавания (рис. 7).

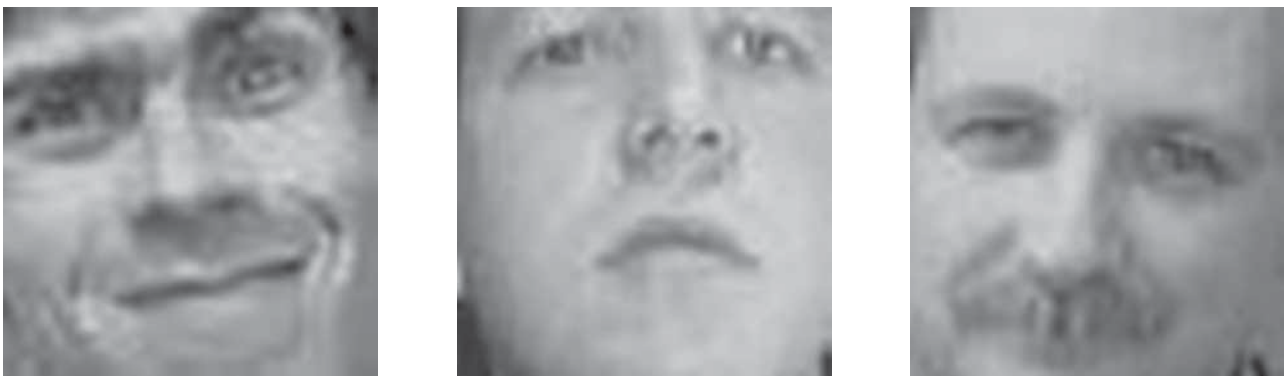


Рис. 5. Пример неточностей детекции лиц

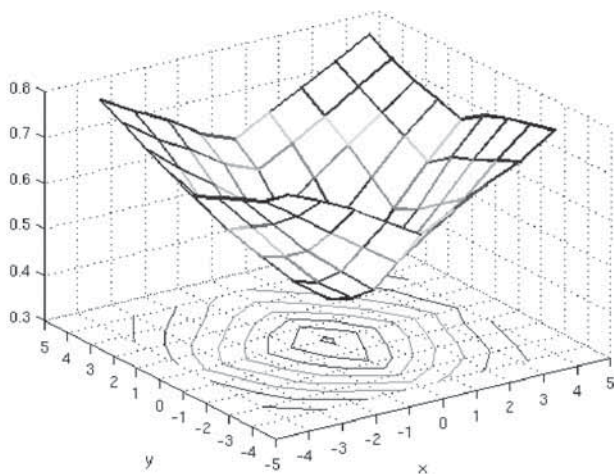


Рис. 6. Зависимость меры близости от относительного смещения изображений

Результаты вычислительных экспериментов, проведенных на изображениях из базы ORL, свидетельствуют о 150-кратном ускорении алгоритма распознавания лиц с применением метода наискорейшего спуска для поиска оптимального совпадения сравниваемых лиц. При этом сохраняются удовлетворительные показатели надежности.

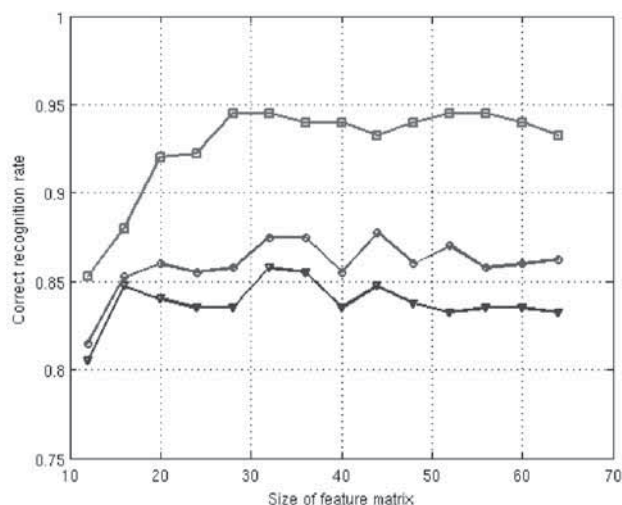


Рис. 7. Зависимость надежности распознавания лиц от размера классификатора для рассмотренных алгоритмов идентификации: полный перебор по параметрам сдвига, наклона и масштаба — квадратный маркер; сравнение по параметрам наклона, сдвига и масштаба с применением направленного поиска — круглый маркер; обычное сравнение на основе псевдометрики — треугольный маркер

7. Результаты тестирования алгоритма распознавания лиц

Предложенные алгоритмы распознавания лиц были протестированы на базе изображений ORL и на трех выборках из базы Либора Спасека.

Для оценки надежности алгоритма распознавания

лиц изображения из тестовых баз были распределены по классам, состоящим из изображений лиц одного человека. Далее, для всей базы была вычислена матрица близости, которая строится следующим образом:

$$D_{i,j} = \rho(F_i, F_j),$$

где F_i, F_j — изображения лиц. Таким образом, каждая i -я строка матрицы близости содержит расстояния от лица F_i до элементов базы $F_j, j = 1, 2, \dots, N$, где N — количество элементов в тестовой выборке. По матрице близости для каждого лица F_i находится ближайший элемент F_k , такой, что $D_{i,k} = \min_{j, j \neq i} D_{i,j}$. Если изображения лиц F_i и F_j

принадлежат одному классу, то считается, что алгоритм распознавания лиц отработал успешно.

Рассчитанная подобным образом ошибка распознавания лиц называется *равной нормой ошибки* (ERR — *Equal Error Rate*) и соответствует ситуации, когда ошибка первого рода (FAR) равна ошибке второго рода (FRR). В табл. 2 приведены результаты тестирования алгоритма распознавания лиц на основе псевдометрики (1) на изображениях из базы ORL и базы Либора Спасека.

Таблица 2
Результаты тестирования алгоритма распознавания лиц

База лиц	N_{img}	N_p	N_0	N_{sz}	ERR, %
ORL	400	40	10	32	5,5
Yale	165	15	11	48	4,0
Libor Spacek	1000	200	5	38	5,0
Libor Spacek	1125	375	3	38	5,3
Libor Spacek	7190	375	20	38	4,0

В таблице используются следующие обозначения:

N_{img} — количество изображений в базе;

N_p — количество персон;

N_0 — количество изображений для каждой персоны;

N_{sz} — средний размер лица;

ERR — уровень ошибки распознавания.

Сравнение ошибки распознавания предложенного метода с известными результатами на базе изображений ORL [5] представлены в табл. 3.

Таблица 3
Надежность распознавания лиц современных методов

Алгоритм	Надежность, %
Метод главных компонент	80
Линейный дискриминант Фишера	91
Одномерная Марковская модель	84
Двумерная Марковская модель	99,5
Вейвлеты Габора	95,5
Матрица изменения яркостей	94,5

Как следует из табл. 3, предложенный алгоритм распознавания лиц, основанный на переходе от исходного изображения к матрице изменения яркостей с мерой близости (1), демонстрирует удовлетворительную надежность распознавания. Основным отличием предложенного метода является то, что он не требует нормализации базы лиц и предварительной обработки изображения, по вычислительным затратам сравнимой с самой процедурой распознавания.

8. Выводы

Как следует из результатов проведенных исследований, предложенные алгоритмы детекции и идентификации лиц демонстрируют вполне приемлемое качество работы на изображениях сверхнизкого разрешения (когда размер лица меньше 75 пикселей). При этом на изображениях из баз Либора Спасека и ORL получены достаточно хорошие оценки надежности с учетом специфики этих изображений. Кроме того, удовлетворительные результаты получены для изображений, предоставленных Яндекс, что позволяет использовать предложенные алгоритмы в качестве основы для поиска портретных изображений по содержанию в Интернете.

9. Благодарности

Авторы выражают благодарность компании Яндекс за поддержку проведенных исследований, а также Либору Спасеку, лаборатории Olivetty Research Laboratory, Yale University и www.bioid.com за предоставленные базы изображений лиц.

10. Литература

- [1] База изображений лиц BioID. Режим доступа: <http://www.bioid.com/downloads/facedb/>
- [2] База изображений лиц ORL. Режим доступа: <http://www.cam-orl.co.uk/facedatabase.html>
- [3] База изображений лиц Yale B. Режим доступа: <http://cvc.yale.edu/projects/yalefacesB/yalefacesB.htm>
- [4] База изображений лиц Yale. Режим доступа: <http://cvc.yale.edu/projects/yalefaces/yalefaces.html>
- [5] Волочков М. П. Об автоматическом распознавании лиц / М. П. Волочков, И. Ю. Самоненко // Интеллектуальные системы. 2005. Т. 9. Вып. 1–4. — С. 135–156.
- [6] Гоголь А. Телевидение: учебник для вузов / Гоголь А., Джакобия В. 2 изд. — М.: Радио и связь. 2004. — 616 с.
- [7] Гонсалес Р. Цифровая обработка изображений / Р. Гонсалес, Р. Вудс; Пер. с англ. — М.: Техносфера, 2006. — 1072 с.
- [8] Математическая энциклопедия. Т. 3. — М.: Сов. Энциклопедия. 1982.
- [9] Рыжов А. П. Элементы теории нечетких множеств и измерения нечеткости / А. П. Рыжов. — М.: Диалог-МГУ. 1998.
- [10] Eickeler S. Comparison of Confidence Measures for Face Recognition / Eickeler S., Jabs M., Rigoll G. // Gerhard-Mercator-University Duisburg, Germany. 2000. 6 p.
- [11] Etemad K. Discriminant Analysis for Recognition of Human Face Images / K. Etemad, R. Chellappa // Journal of the Optical Society of America A. 1997. Vol. 14. No. 8. P. 1724–1733.
- [12] Froba B. Real-time face detection using edge-orientation matching / B. Froba, C. Kulbeck // Proc. AVBPA 2001. 2001. P. 78–83.
- [13] Garcia C. Wavelet packet analysis for face recognition / C. Garcia, G. Zikos, G. Tziritas // Image and Vision Computing. 2000. 18. P. 289–297.
- [14] Georghiades A. S. From Few to Many: Illumination Cone Models for Face Recognition under Variable Lighting and Pose / A. S. Georghiades, P.N. Belhumeur, D. J. Kriegman // IEEE Trans. Pattern Anal. Mach. Intelligence. 2001. vol. 23. No. 6. P. 643–660.
- [15] Hjelm E. Face detection: A survey / E. Hjelm and B. K. Low // Computer Vision and Image Understanding. 2001. 83(3). P. 236–274.
- [16] Jesorsky O. Robust Face Detection Using the Hausdorff Distance / O. Jesorsky, K. Kirchberg, R. Frischholz // J. Bigun and F. Smeraldi, editors // Audio and Video based Person Authentication — AVBPA. Springer. 2001. P. 90–95.
- [17] Lu J. Face Recognition Using LDA-Based Algorithms / J. Lu, K. N. Plataniotis, A. N. Venetsanopoulos // IEEE Trans. on Neural Networks. 2003. Vol. 14. No. 1. P. 195–200.
- [18] Martinez A. M. PCA versus LDA / A. M. Martinez, A. C. Kak // IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 23. No. 2. 2001. P. 228–233.
- [19] Moon H. Computational and Performance aspects of PCA-based Face Recognition Algorithms / H. Moon, P. J. Phillips // Perception, Vol. 30. 2001. P. 303–321.
- [20] Nefian A. V. Embedded Bayesian networks for face recognition / A. V. Nefian // Proc. of the IEEE International Conference on Multimedia and Expo. 2002. Vol. 2. Lusanne, Switzerland. P. 133–136.
- [21] Pentland A. View-Based and Modular Eigenspaces for Face Recognition / A. Pentland, B. Moghaddam, T. Starner // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 21–23 June 1994, Seattle, Washington, USA. P. 84–91.
- [22] Phillips P. J. FRVT 2006 and ICE 2006 Large-Scale Results. / P. J. Phillips, W. T. Scruggs, A. J. O'Toole, P. J. Flynn, K. W. Bowyer, C. L. Schott, M. Sharpe: National Institute of Standards and Technology, Gaithersburg, 2007. Режим доступа: <http://www.frv.org/FRVT2006/>
- [23] Samaria F. Parameterisation of a stochastic model for human face identification / F. Samaria, A. Harter // 2nd IEEE Workshop on Applications of Computer Vision. Sarasota (Florida). 1994.
- [24] Samaria F. HMM-Based Architecture for Face Identification / F. Samaria, S. Young // Image and Vision Computing. 1994. vol. 12. no. 8. P. 537–543.
- [25] Saradha A. A Hybrid Feature Extraction Approach for Face Recognition Systems / A. Saradha, S. Annaduri. Режим доступа: <http://www.icgst.com/gvip/v5/P1150515002.pdf>.
- [26] Spacek L. База изображений лиц / Spacek L. Режим доступа: <http://cswww.essex.ac.uk/mv/allfaces/index.html>
- [27] Viola P. Robust real-time object detection / P. Viola, M. Jones // 2nd International Workshop on Statistical and Computational Theories of Vision-Modelling, Learning, Computing, and Sampling. Vancouver. 2001.
- [28] Zhao W. Face Recognition: A Literature Survey / W. Zhao, R. Chellappa, A. Rosenfeld, P. J. Phillips // ACM Computing Surveys. 2003. P. 399–458.

Content Based Facial Image Search

A. Goncharov, A. Gorban, A. Karkishchenko,
A. Lepskiy

The task of searching facial images consists of two problems: face detection and face recognition. This paper describes face detection and recognition methods based on special image representation — matrix of brightness variation formed from derivative signs of image intensity function. The partial invariance of algorithms to illumination conditions have shown. All proposed algorithms were tested on the proper facial image bases. Experimental results confirm efficiency of proposed methods.