

С. М. Аракелян, И. И. Зиновьев, М. М. Рожков

Владимирский государственный университет
им. Александра Григорьевича и Николая Григорьевича Столетовых
ул. Горького, 87, Владимир, 600000, Россия

E-mail: izinoviev@vlsu.ru

ПАКЕТНЫЙ СПОСОБ СРАВНЕНИЯ ИЗОБРАЖЕНИЙ ПО ИХ ФРАГМЕНТАМ *

Сравнение изображений некоторых объектов сильно осложняется их возможной изменчивостью. Одним из решений такой проблемы может стать выделение относительно неизменных участков изображений для их сравнения. В работе предложен проверенный на практике способ сравнения изображений путем комбинации их соседних участков. Полученные результаты свидетельствуют о его практической применимости, например, в процессе распознавания лиц людей.

Ключевые слова: сравнение изображений, пространства векторов атрибутов, соседние участки изображения.

Введение

Классификация объектов по их изображениям является интенсивно развивающимся направлением применения современных средств вычислительной техники. Особый интерес представляет задача распознавания людей по изображениям их лиц [1–3]. Однако существует ряд факторов, осложняющих решение этой задачи. Один из них заключается в изменчивости объекта распознавания. Это особенно актуально для изображений лиц людей: появление усов, бороды, очков, изменение прически, выражения лица, возможность частичного перекрытия другим объектом и т. п. Помехи такого рода ухудшают результаты распознавания [1]. Возможным решением проблемы может рассматриваться поиск относительно неизменных в таких условиях участков изображения лица и проведение сравнения на базе таких участков [4–6].

Одним из недостатков такого подхода является необходимость точной юстировки изображений, т. е. приведение их к стандартной системе координат, когда одинаковые участки на разных изображениях привязаны к одним и тем же координатам [4]. Традиционно такая предобработка осуществляется путем поиска на лице характерных точек: центры глаз, рта или носа, относительно которых производится приведение изображений к стандартному виду. Например, удаление поворотов объекта в плоскости изображения, удаление ненужной информации (фона).

Но задача поиска таких ключевых точек сама по себе нетривиальна. Даже проставление этих точек вручную не гарантирует их идеального соответствия друг другу на разных изо-

* Работа выполнена в рамках НИР по госзаданию «Наука» (№ 8.3303.2011 и 8.3534.2011 от 23.11.2011).

бражениях. Поэтому возможна неточность позиционирования изображения объекта после предобработки.

В работе предлагается способ сравнения изображений, который позволяет компенсировать возможные неточности юстировки путем применения пакетного сравнения [7] к участкам изображений.

Предлагаемый способ сравнения изображений

В работе [7] описан способ сравнения фотографий, в основе которого лежит построение групп размытых исходных изображений и их производных карт фильтрами Гаусса. Для обозначения этих групп было предложено понятие «пакеты трансформаций». Их сравнение ведется по векторам атрибутов, полученным с помощью инструментов сканирования-сжатия, основная задача которых заключается в корректном уменьшении пиксельной размерности изображений. Исходное изображение может быть подвергнуто некоторому преобразованию, результат которого в дальнейшем будем называть производной картой яркостей (или производной картой). Существенно то, что размерность полученной карты равна размерности исходного изображения. Комбинация такого преобразования и инструмента сжатия определяет процедуру «обработки изображения», в результате которой формируются подпространства векторов атрибутов для пакетов трансформаций. Расстояние между получаемыми подпространствами рассчитываются по метрике $d_{m,n}$ на базе определителя Грамма. Особенностью метрики является нахождение минимального расстояния между всеми возможными парами векторов атрибутов, образующих подпространства. В зависимости от количества сравниваемых векторов атрибутов возможны частные варианты метрики, которые мы будем рассматривать ниже: $d_{1,1}$ – для сравнения единичных векторов и $d_{3,3}$ – для сравнения подпространств, образованных тремя векторами. Применение соответствующего варианта метрики к полученным после процедуры «обработки изображения» подпространствам формирует «механизм сравнения».

Для сравнения изображений по их небольшим участкам воспользуемся описанным выше пакетным способом сравнения. Только для формирования пакетов сравнения будем использовать не размытия исходного изображения, а его небольшие соседние участки. Благодаря особенностям метрики, используемой при пакетном сравнении, перемещение сравниваемых участков в рамках пакета на разных изображениях не повлияет на результат сравнения, т. е. позволит компенсировать возможные неточности позиционирования объекта на изображении после юстировки.

Более формально преобразования, описанные выше, можно записать в виде

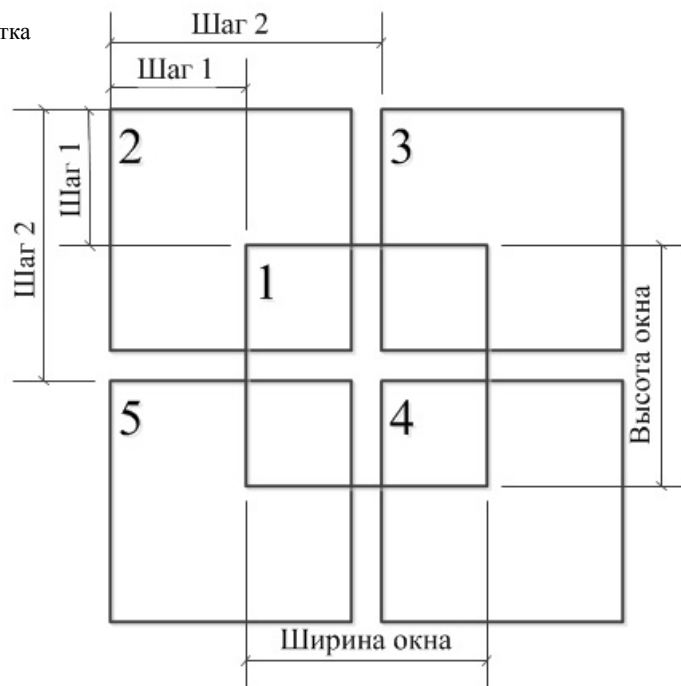
$$M \rightarrow t(M) \rightarrow P_i t(M) \rightarrow a_i \equiv a(P_i t(M)); i = 1, \dots, m,$$

где M – карта яркости изображения (исходное изображение, хранимое в памяти компьютера в виде матрицы значений интенсивности отдельных его пикселей); $t(M)$ – необязательное преобразование исходной карты, не изменяющее ее размерности, например, свертка фильтрами Собеля или преобразование с использованием текстурных карт Лавса. Задачей данного этапа является извлечение максимального объема информации об объекте, представленном на изображении, например о его текстуре или форме. Этап $P_i t(M)$ формирования пакетов для сравнения из участков изображения опишем подробнее.

Рассматриваемый участок изображения назовем окном интереса. Параметрами, определяющими форму окна, являются его высота и ширина. Предлагается вокруг исходного окна интереса (номер 1) взять еще четыре (рис. 1). Взаимное расположение выбранных окон определяется значением переменных «шаг 1» и «шаг 2». При таком сравнении участка изображения, представленного окном 1, будет компенсироваться возможное его смещение в пределах небольшой окрестности (окна 2–5). Полученный таким образом пакет будем называть «пакетом сравнения».

Расстояние между подпространствами векторов атрибутов a_i , полученных в результате применения инструментов сканирования-сжатия к сформированным пакетам $a(P_i t(M))$, определяется метрикой $d_{m,n}$.

Рис. 1. Общая схема анализируемого участка изображения



Проверим эффективность предложенного способа сравнения изображений в рамках задачи классификации фотографий людей с разными выражениями лиц.

Исследование предложенного способа сравнения изображений

Для эксперимента выбраны 865 пар изображений FA и FB из базы ColorFERET, для которых имеются координаты глаз и рта [8]. Данные наборы различаются выражением лиц людей на них. Проведена юстировка выбранных изображений (рис. 2). Из этих изображений были составлены 2 множества следующим образом: первое содержит пары изображений FA и FB одного класса (человека), а второе – фотографии разных классов (людей). Количество пар в обоих множествах одинаково и равно 865. Выбор комбинации изображений для пар второго множества при проведении эксперимента несущественен.

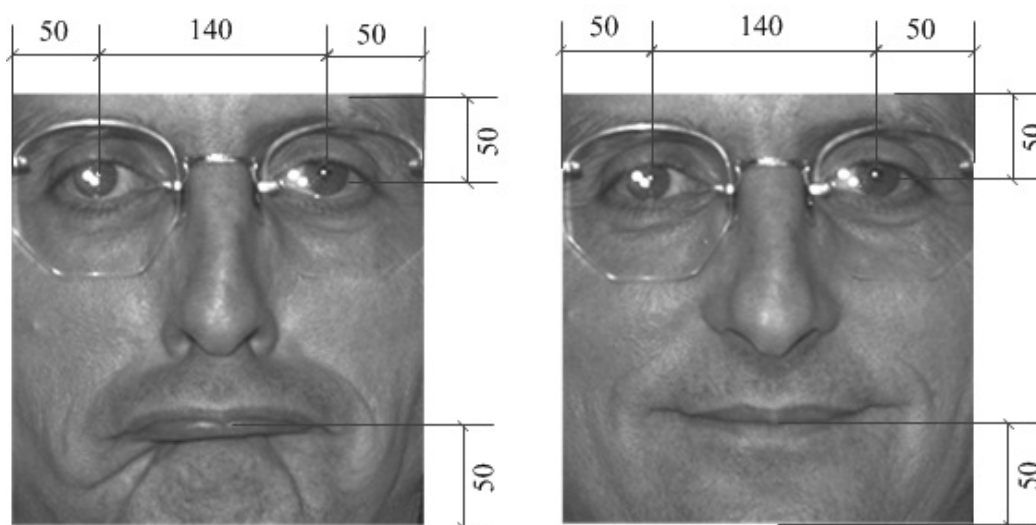


Рис. 2. Пример изображений после их юстировки

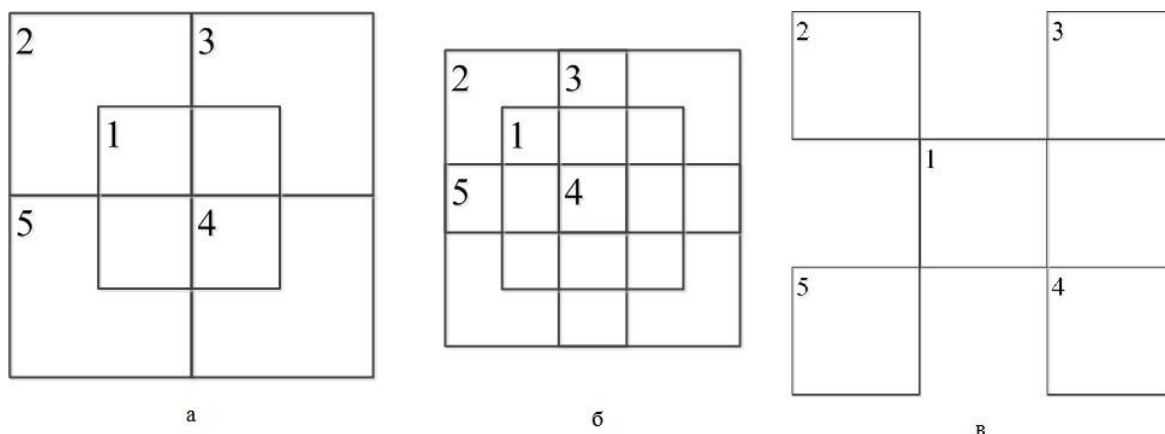


Рис. 3. Взаимное расположение окон в зависимости от выбранного шага:
a – шаг равен половине размера окна; *б* – шаг равен трети размера окна; *в* – шаг равен размеру окна

Таблица 1

Исследуемые комбинации параметров для построения «пакетов сравнений»

Идентификатор комбинации	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	<i>e</i>	<i>f</i>	<i>g</i>	<i>h</i>	<i>i</i>
Размер окна	24	24	24	36	36	36	48	48	48
Шаг	8	12	24	12	18	36	16	24	48

В качестве критерия оценки эффективности предлагаемого подхода возьмем отношение межклассового расстояния к внутриклассовому (K) [9]. Чем больше значение K , тем более эффективно возможно разделение на классы. В ходе эксперимента будем оценивать влияние таких параметров формирования «пакета сравнения», как размер окон и их взаимное расположение на значение коэффициента K для пакета.

Как описывалось выше, на процесс формирования пакетов из участков изображения оказывают влияние несколько параметров: ширина и высота окна, «шаг 1» и «шаг 2». Для эксперимента взято квадратное окно, т. е. ширина окна равна высоте окна. Обозначим это измерение как размер окна. Также для упрощения эксперимента был зафиксирован размер «шага 2» относительно «шага 1», а именно «шаг 2» = «шаг 1» * 2. Таким образом, для вариации процесса формирования пакета осталось два параметра: размер окна и шаг между окнами.

В ходе эксперимента использовались окна с размерами 24×24 , 36×36 и 48×48 пикселей, при этом шаг между ними был равен трети, половине или размеру окна (рис. 3). Все исследуемые комбинации этих параметров приведены в табл. 1.

Исследуемые в ходе эксперимента процедуры обработки изображения $t(M)$ приведены в табл. 2. При использовании инструментов сканирования бралось ограниченное число полученных в результате их работы коэффициентов. В связи с рассмотрением окон небольших размеров для данного эксперимента было выбрано значение в 50 коэффициентов.

Опишем подробнее последовательность эксперимента.

Для каждой процедуры обработки изображения из табл. 2 производились следующие действия.

1. Формировались пакеты сравнения, состоящие только из окна 1 (см. рис. 1). Выбор окна производился путем его перемещения по всему изображению с интервалом, равным половине размера окна. Проводились измерение расстояния для каждой пары изображений по метрике $d_{1,1}$ и расчет критерия оценки эффективности $K_{1,1}$. В результате для каждого из рассматриваемых размеров окон были отобраны по 10 участков изображения, на которых было достигнуто максимальное значение величины $K_{1,1}$.

Таблица 2

Исследуемые процедуры обработки изображения

Идентификатор	Название процедуры обработки изображения	Преобразование	Инструмент сканирования
1	Дискретное косинусное преобразование (ДКП)	Тождественно	Коэффициенты ДКП-преобразования
2	Полиномы Лежандра	Тождественно	Коэффициенты разложения по полиномам Лежандра
3	Карты Лавса	Преобразование с использованием текстурных карт Лавса	Коэффициенты ДКП
4	Маски Собеля	Свертка фильтрами Собеля	Коэффициенты ДКП
5	Карты Лавса + Полиномы Лежандра	Преобразование с использованием текстурных карт Лавса	Коэффициенты разложения по полиномам Лежандра
6	Маски Собеля + Полиномы Лежандра	Свертка фильтрами Собеля	Коэффициенты разложения по полиномам Лежандра

2. Для каждого из 10 выбранных участков были построены пакеты сравнения из трех соседних участков изображения следующим образом. Окно 1 считалось основным, соответствующим одному из 10 рассматриваемых участков. Для формирования пакета бралось еще два соседних к нему участка. Таким образом, возможны следующие комбинации соседних участков: 1 – 2 – 3, 1 – 2 – 4, 1 – 2 – 5, 1 – 3 – 4, 1 – 3 – 5 и 1 – 4 – 5. Для всех возможных комбинаций параметров формирования пакетов (см. табл. 1) проводились измерение расстояния для каждой пары изображений по метрике $d_{3,3}$ и расчет величины $K_{3,3}$.

3. Нас интересует отношение $K_m = K_{3 \times 3} / K_{1 \times 1}$ для каждого из анализируемых участков и сформированных на его основе пакетов. Этот критерий характеризует степень увеличения межклассового расстояния относительно внутриклассового при сравнении изображений предлагаемым нами способом сравнения по отношению к простому анализу независимых участков изображения. Его увеличение говорит о возможности более качественного различения двух сформированных нами тестовых множеств.

На рис. 4 приведены средние значения коэффициента K_m , полученные в ходе эксперимента для каждой процедуры обработки изображения. На рис. 5 – максимальные значения коэффициента K_m , которые были достигнуты в ходе эксперимента.

Эксперимент подтвердил возможность применения «пакетного сравнения» с формированием пакетов из пересекающихся участков карты изображения для сравнения изображений. Все комбинации карт изображений и инструментов сканирования показали ту или иную степень уменьшения внутриклассового расстояния относительно межклассового. Следует отметить, что наиболее эффективно пакетное сравнение реализуется при анализе изображения довольно крупным окном (порядка 48 пикселей) с достаточной степенью перекрытия соседних окон (шаг равен трети размера окна).

Полученные в ходе эксперимента результаты показали эффективность использования пакетов сравнения из соседних участков изображения по отношению к анализу отдельных его участков. Интерес представляет оценка эффективности описанного выше пакетного способа сравнения для извлечения признаков из изображения в процессе построения классификатора людей по изображениям их лиц. С этой целью был проведен следующий эксперимент.

Построение классификатора людей по изображениям их лиц

Для проведения эксперимента воспользуемся 865 парами изображений базы ColorFERET из описанного выше эксперимента.

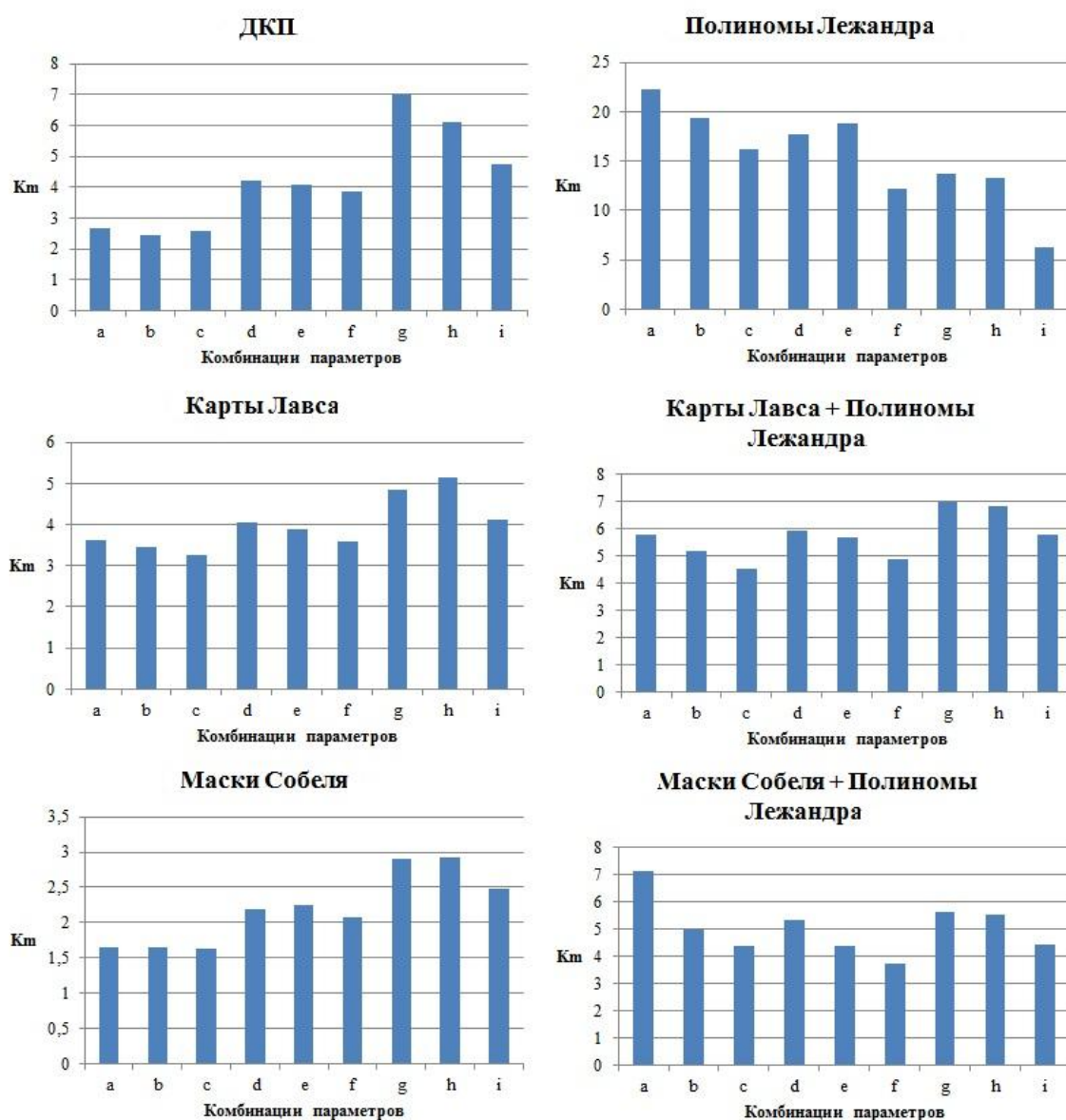


Рис. 4. Средние значения коэффициента K_m

Целью эксперимента является сравнение эффективности различных признаков для построения классификатора лиц людей.

Опишем исследуемые признаки.

1. «Глобальные признаки». Вектор атрибутов, полученный с использованием любой процедуры обработки изображений из табл. 2, примененной ко всему изображению в целом. Количество значащих коэффициентов инструментов сжатия равно 200. Вычисление расстояния между векторами атрибутов производится по метрике $d_{1,1}$.

2. «Локальные признаки 1×1 ». Вектор атрибутов, полученный с использованием процедуры обработки изображений из табл. 2, примененной к участку изображения. Перебор участков изображения осуществляется окном размером 48×48 пикселей, скользящим по всему изображению с шагом в 24 пикселя. Поскольку размер окна в 5 раз меньше размеров полного изображения, то и значащих коэффициентов инструментов сжатия можно использовать пропорционально меньше. Мы использовали 50 старших коэффициентов. Вычисление расстояния между векторами атрибутов производится по метрике $d_{1,1}$.

3. «Локальные признаки 3×3 ». Подпространство векторов атрибутов, полученных с использованием одной из процедур обработки изображений, описанной в табл. 2, примененной к участку изображения и его окрестности аналогично тому, как это описано в предыдущем эксперименте. Для выбора участка изображения используется окно размером 48×48 пикселей, скользящее по всему изображению с интервалом в 24 пикселя. Шаг, задающий расположение соседних окон, равен 16 (см. рис. 1). Количество значащих коэффициентов инструментов сжатия также равно 50. Вычисление расстояния между подпространствами векторов атрибутов производится по метрике $d_{3,3}$.

Введем понятие слабого классификатора $F(x)$, который состоит из метрики $M(x) \equiv d(\Phi, \Phi')$ и порога q , Φ – один из описанных выше признаков (глобальный, локальный 1×1 или локальный 3×3). Классификатор работает по принципу

$$F(x) = \begin{cases} +1, & M(x) \leq q \\ -1, & M(x) > q \end{cases}$$

Такие классификаторы могут использоваться для формирования сильного классификатора путем их взвешенного объединения, например, с помощью алгоритма усиления слабых классификаторов AdaBoost, предложенного Йоавом Фройндом (Yoav Freund) и Робертом Шапире (Robert Schapire) [10].

Особенностью классификатора, построенного по алгоритму AdaBoost, является его способность различать только два класса. Поэтому сведем задачу классификации лиц людей к задаче разделения на два класса. Для этого определим эти классы:

- 1) «свои» (+1) – класс, содержащий пары изображений одного человека;
- 2) «чужие» (-1) – класс, содержащий пары изображений разных людей.

В эксперименте используется реализация алгоритма из библиотеки OpenCV¹. Процесс обучения реализован в методе CvBoost::train(), процесс работы классификатора – в методе CvBoost::predict().

В ходе эксперимента производится обучение трех сильных классификаторов на базе одного из трех исследуемых признаков. Можно выделить следующие основные параметры, влияющие на процесс обучения:

- 1) вариант алгоритма AdaBoost (Discrete AdaBoost, Real AdaBoost, Gentle AdaBoost или LogitBoost) [11];
- 2) количество используемых слабых классификаторов;

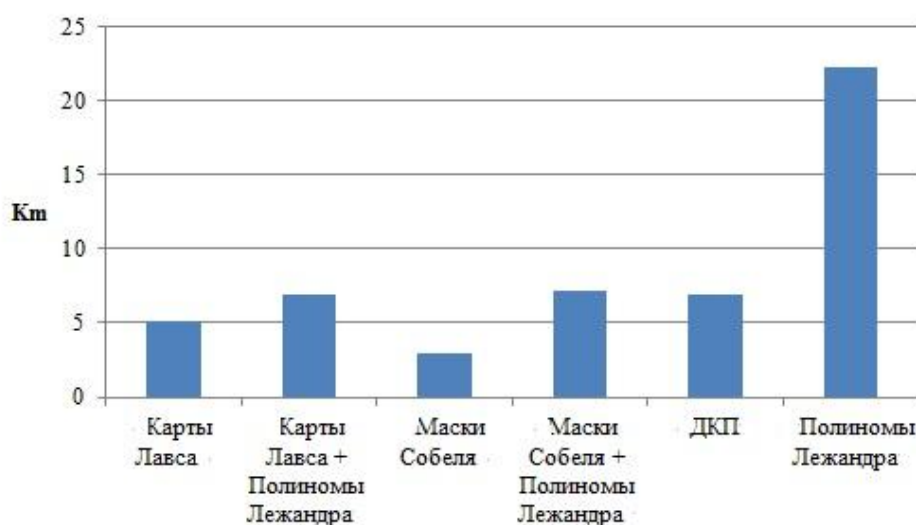


Рис. 5. Максимальные значения K_m , полученные в ходе эксперимента

¹ Документация к библиотеке OpenCV. URL: <http://opencv.willowgarage.com/documentation/cpp/boosting.html>

3) в алгоритме обучения, реализованном в библиотеке OpenCV, слабые классификаторы становятся узлами деревьев принятия решений, поэтому имеется возможность задать требуемую глубину этих деревьев;

4) для ускорения процесса обучения возможно исключение из рассмотрения «легких» обучающих примеров, т.е. тех, на которых обучаемый классификатор не ошибается. Для этого можно указать вес тренировочного экземпляра, ниже которого он не участвует в процессе обучения.

В нашем случае выбор этих параметров не принципиален. Мы их зафиксируем, чтобы обеспечить одинаковые условия сравнения. Для эксперимента будем использовать дискретный вариант алгоритма AdaBoost, число слабых классификаторов 20, глубину деревьев 5, тренировочный вес 0,95.

Обученные классификаторы при сравнении двух изображений способны определить, относятся ли они к одному классу или к разным. Для обучения классификатора использовались первые по порядку 432 пары изображений FA и FB одного человека и любые 432 пары изображений разных людей.

Множество, на котором проверялся предлагаемый подход, сформировано оставшимися 433 парами изображений. Изображения FA сформировали базу эталонов, изображения FB – базу распознаваемых изображений. Каждое изображение FB сравнивалось со всеми изображениями FA. В результате для каждого распознаваемого изображения формируется список эталонов, которые попали в класс «свои».

Для различения эталонов в этом списке введем понятие «степень соответствия» W_d :

$$W_d = \sum_{n:F_n(x)=d} \log \frac{1}{\beta_n},$$

где $\beta_n = \frac{\varepsilon_n}{1 - \varepsilon_n}$; ε_n – ошибка гипотезы F_n отнесения теста к тому или иному классу; $d = -1, +1$ – параметр отнесения к классу свои / чужие.

Тогда степень соответствия W эталона распознаваемому изображению можно задать следующей формулой:

$$W = \frac{W_+}{W_+ + W_-}.$$

Этот параметр принимает значение в отрезке $[0, 1]$ и может применяться для ранжирования полученного списка эталонов.

С целью оценки эффективности обученных классификаторов для каждого из них измерялись ошибки распознавания FAR и FRR при разных значениях «степени соответствия» для выбранных из базы эталонов. Под ошибкой распознавания FAR будем считать присутствие чужого эталона в результирующем списке. Под ошибкой распознавания FRR будем считать отсутствие нужного эталона в результирующем списке эталонов.

Результаты эксперимента представлены на рис. 6–8.

Сравнение суммарных ошибок классификации (рис. 9) полученных классификаторов показывает, что глобальные и локальные признаки 1×1 в ходе эксперимента показали примерно одинаковую минимальную ошибку классификации в 30 %. Локальные признаки 3×3 позволили снизить суммарную ошибку до 20 %. Что подтверждает эффективность их применения для извлечения признаков из изображения в процессе формирования классификатора.

Заключение

В работе предложен способ сравнения изображений по их небольшим участкам. Проведены эксперименты, которые подтвердили его применимость на практике. Предложенный подход позволяет снизить ошибки классификации фотографий людей с разными выражениями лиц. Достигается это благодаря объединению нескольких соседних участков изображения в пакет и сравнению изображений пакетным способом. В ходе проведенных экспериментов при сравнении изображений двух классов предложенным способом было зафиксировано увеличение межклассового расстояния относительно внутриклассового от 3 до 22 раз.

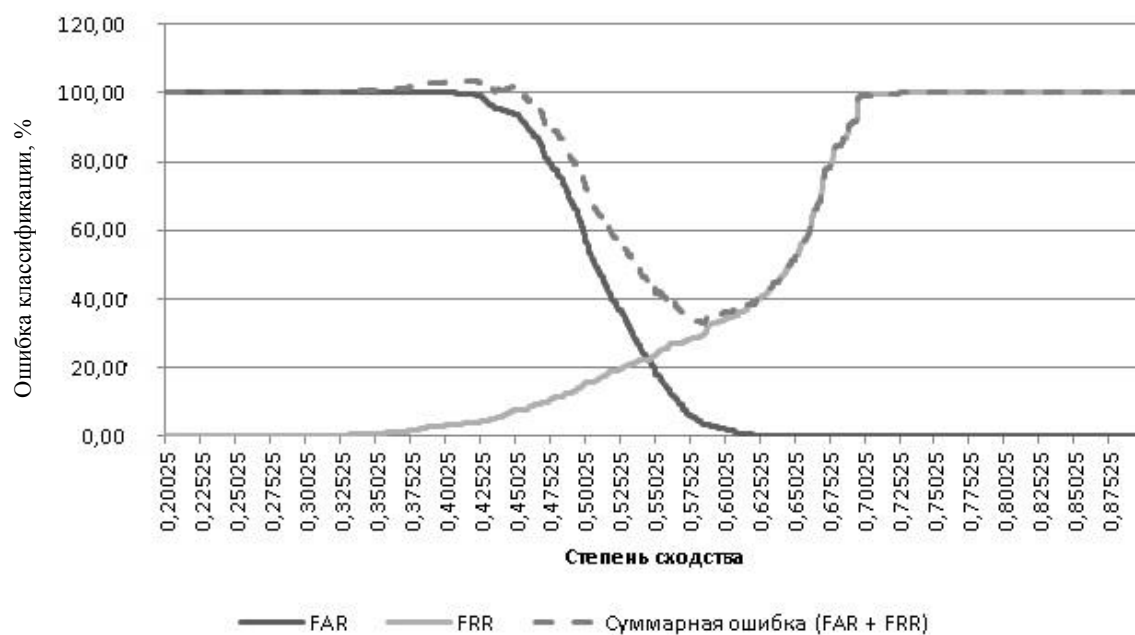


Рис. 6. Ошибка распознавания классификатором по «глобальным признакам»

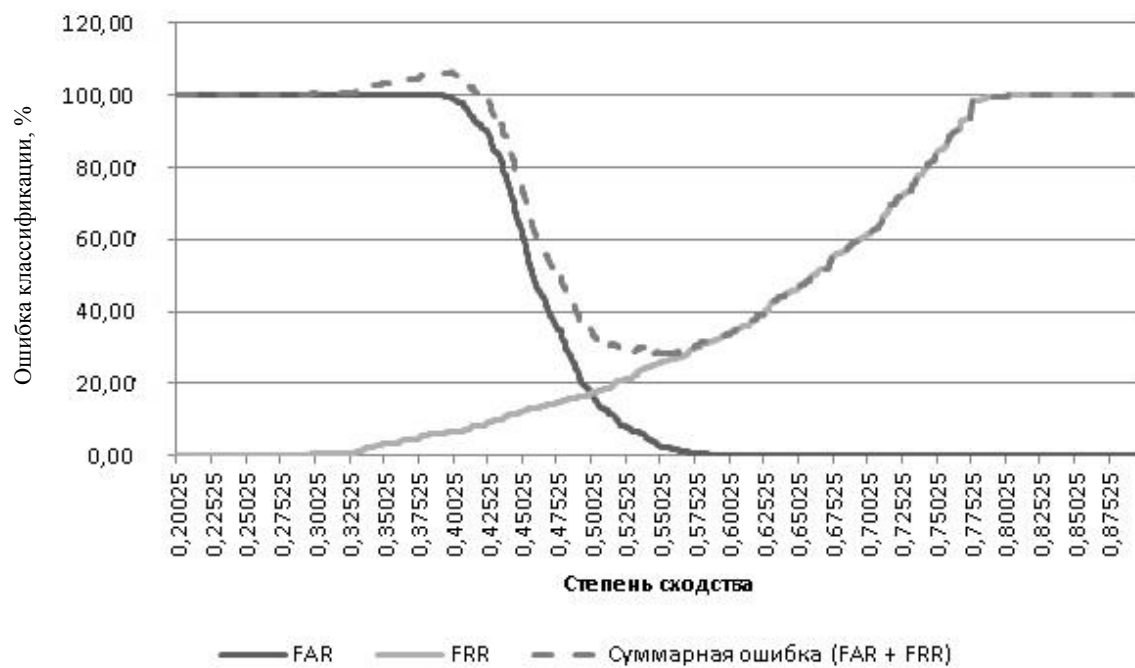


Рис. 7. Ошибка распознавания классификатором по «локальным признакам 1 × 1»

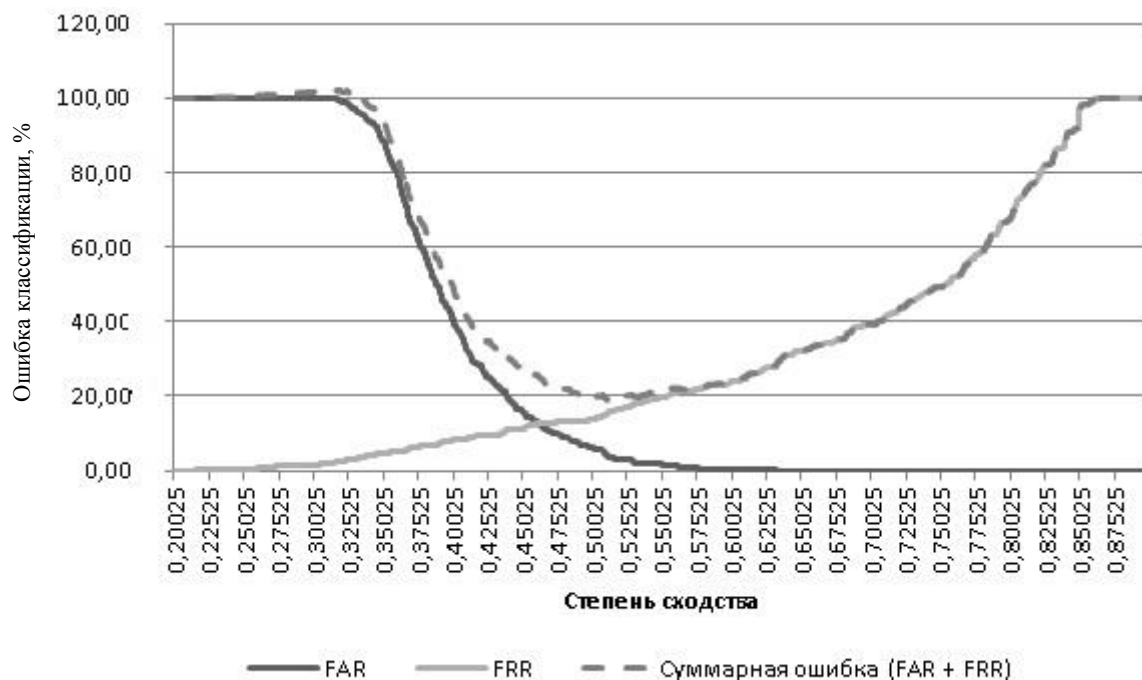


Рис. 8. Ошибка распознавания классификатором по «локальным признакам 3×3 »

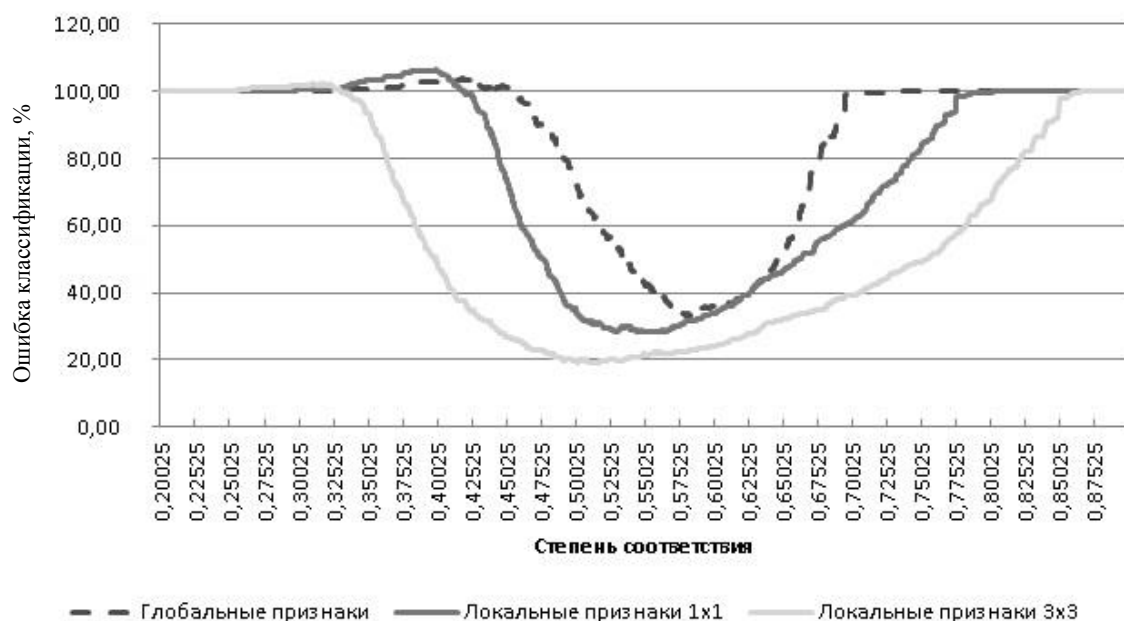


Рис. 9. Суммарные ошибки распознавания

Это свидетельствует о возможности более качественного разделения рассматриваемых классов и, следовательно, снижения ошибки классификации.

В качестве возможного применения предложенного способа сравнения приведена экспериментальная система классификации людей по изображениям их лиц. Было продемонстрировано преимущество предложенного подхода, как перед сравнением изображений целиком, так и по отдельным небольшим участкам.

Список литературы

1. *Tan X., Chen S. C., Zhou Z.-H., Zhang F.* Face Recognition from a Single Image per Person: A Survey // *Pattern Recognition*. 2006. Vol. 39. P. 1725–1745.
2. *Tolba A. S., El-Baz A. H., El-Harby A. A.* Face Recognition: A Literature Review // *International Journal of Signal Processing*. 2006. Vol. 2. P. 88–103.
3. *Zhao W., Chellappa R., Phillips P. J., Rosenfeld A.* Face Recognition: A Literature Survey // *ACM Computing Survey*. 2003. Vol. 35 (3). P. 399–458.
4. *Martinez A. M.* Recognizing Imprecisely Localized, Partially Occluded, and Expression Variant Faces from a Single Sample per Class // *IEEE Trans. Pattern Analysis and Machine Intelligence*. 2002. Vol. 25 (6). P. 748–763.
5. *Samaria F.* Face Segmentation for Identification Using Hidden Markov Models // *British Machine Vision Conference*. BMVA Press, 1993. P. 399–408.
6. *Tan X., Chen S. C., Zhou Z.-H., Zhang F.* Recognizing Partially Occluded, Expression Variant Faces from Single Training Image per Person with SOM and Soft kNN Ensemble // *IEEE Transactions on Neural Networks*. 2005. Vol. 16 (4). P. 875–886.
7. *Зиновьев И. И.* Пакетный способ сравнения изображений // *Прикладная информатика*. 2011. № 6 (36). С. 95–101.
8. *Phillips P. J., Wechsler H., Huang J., Rauss P. J.* The FERET Database and Evaluation Procedure for Face Recognition Algorithms // *Image and Vision Computing*. 1998. Vol. 16 (5). P. 295–306.
9. *Матвеев Ю. Н.* Основы теории систем и системного анализа: Учеб. пособие. Тверь, 2007. Ч. 1. С. 28–29.
10. *Freund Y., Schapire R.* A Decision-Theoretic Generalization of on-Line Learning and an Application to Boosting // *Journal of Computer and System Sciences*. 1997. Vol. 55 (1). P. 119–139.
11. *Friedman J. H., Hastie T., Tibshirani R.* Additive Logistic Regression: A Statistical View of Boosting // *Technical Report*. Dept. of Statistics. Stanford Univ., 1998.

Материал поступил в редколлегию 12.10.2012

S. M. Arakelyan, I. I. Zinoviev, M. M. Rozhkov

THE BATCH METHOD OF IMAGE COMPARISON BY THEIR AREAS

Comparison of images of some objects is complicated by their possible variability. One solution to this problem might be to find relatively constant parts of the images for their comparison. The authors propose a method for comparing images by combining their neighboring areas. Proposed in the algorithm has been approved in practice. The received results show its practical applicability, for example in the task of humans face recognition.

Keywords: image comparison, vector spaces of attributes, neighboring image areas.