

Козин Никита Евгеньевич

**ПОКАЗАТЕЛИ СОПРЯЖЕННОСТИ
И МУЛЬТИКОЛЛИНЕАРНОСТИ
В ЗАДАЧАХ АНАЛИЗА
И РАСПОЗНАВАНИЯ ИЗОБРАЖЕНИЙ**

Специальность:

05.13.17 – Теоретические основы информатики

Автореферат
диссертации на соискание ученой степени
кандидата технических наук

Самара - 2008

Работа выполнена в Самарском государственном аэрокосмическом университете имени академика С.П.Королева и Институте систем обработки изображений РАН

Научный руководитель: доктор технических наук, профессор
Фурсов Владимир Алексеевич

Официальные оппоненты: доктор технических наук, профессор
Ташлинский Александр Григорьевич

доктор физико-математических наук, доцент
Мясников Владислав Валерьевич

Ведущая организация: Уфимский государственный авиационный
технический университет

Защита диссертации состоится «13» февраля 2009 г. в 10 ч. на заседании диссертационного совета Д 212.215.07 в Самарском государственном аэрокосмическом университете имени академика С.П.Королева, по адресу: 443086, г. Самара, Московское шоссе, д. 34.

С диссертацией можно ознакомиться в библиотеке Самарского государственного аэрокосмического университета имени академика С.П.Королева.

Автореферат разослан «31» декабря 2008 г.

Ученый секретарь
диссертационного совета,
д.т.н., профессор

И.В. Белоконов

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность.

Системы распознавания образов в настоящее время получили широкое распространение. Трудно назвать сферу деятельности, где такие системы не используются. Особенно широкое распространение получили системы распознавания и понимания изображений. Связано это с тем, что информация о многих объектах и явлениях в настоящее время регистрируется и хранится в виде цифровых изображений.

Распознавание образов как научное направление возникло и развивается с конца 50-х годов прошлого века. Большой вклад в развитие теории распознавания образов внесли отечественные ученые С.А. Айвазян, М.А. Айзерман, М.М. Бонгард, Э.М. Браверманн, В.Н. Вапник, К.В. Воронцов, В.М. Глушков, А.Л. Горелик, Ю.И. Журавлев, Н.Г. Загоруйко, А.Г. Ивахненко, В.А. Ковалевский, Г.С. Лбов, Л.И. Розоноэр, К.В. Рудаков, В.А. Скрипкин, А.А. Харкевич, Я.З. Цыпкин, А.Я. Червоненкис, М.И. Шлезингер, и др. За рубежом основоположником работ в области распознавания образов является Ф. Розенблатт, предложивший модель деятельности мозга – перцептрон. Большой вклад в развитие теории распознавания внесли также зарубежные ученые: Ф. Гонсалес, Р. Дуда, Дж. Ту, К. Фукунага, К. Фу, П. Харт и др.

Системы распознавания образов предназначены для классификации входных изображений на некоторые группы. Первые системы разрабатывались для читающих автоматов, в которых решалась задача распознавания знаков, изображающих букву или цифру. В последние годы повышенное внимание со стороны исследователей получило распознавание лиц. Связано это, с одной стороны, с тем, что распознавание лиц, является одним из наиболее сложных приложений анализа и понимания изображений, с другой стороны, с бурным ростом спроса на автоматические системы видеоконтроля и видеонаблюдения.

Несмотря на широкую коммерциализацию рынка программных продуктов распознавания и доступность ряда работающих технологий, интенсивность исследований в области распознавания не снижается, т.к. требуемый уровень надежности таких систем пока еще недостаточен. Актуальность проблемы подтверждается продолжающимся ростом числа конференций по распознаванию, таких как ICAFG (International Conference on Automatic Face and Gesture Recognition) или AVBPA (Audio- and Video-based Biometric Person Authentication), созданием систематических эмпирических тестов для проверки качества методов распознавания, например, FERET (Face Recognition Technology) или FRVT (Face Recognition Vendor Test) и др.

Способ классификации изображений, основанный на вычислении мер близости между ними, является одним из самых первых подходов к решению задачи распознавания образов. Экспериментальные исследования различных методов распознавания, использующих эту идею, подтверждают ее эффективность. Часто такие эксперименты осуществляются в пространстве признаков, где в качестве значений признаков используются значения яркостей отсчетов цифрового изображения.

В рамках этого направления наиболее широко используются следующие меры близости: евклидово расстояние, манхэттенская метрика, расстояние Махаланобиса. Сравнительные исследования показывают, что качество распознавания при применении различных мер близости может существенно различаться. Более того, эти различия существенным образом зависят также от конкретных особенностей задачи (характера искажений, взаимного расположения векторов образов в классе и др.). В связи с этим возникает вопрос о возможности применении других мер близости, которые в определенных условиях могут дать лучший результат, по сравнению с широко используемыми.

Известно, что взаимная ориентация (степень сопряженности) векторов образов в классе оказывает существенное влияние на качество распознавания. В задачах линейной регрессии в качестве меры почти линейной зависимости векторов независимых переменных широко используются так называемые меры мультиколлинеарности: определитель, минимальное собственное число, показатели парной и максимальной сопряженности. В задачах поиска и распознавания изображений эти меры пока не нашли заметного применения. В работе *Training in Pattern Recognition from a Small Number (ISPR – 2000, Barcelona, Spain. Vol. 2, p. 119-121)* для анализа свойств признакового пространства впервые предложено использовать так называемые показатели максимальной сопряженности и сопряженности с нуль-пространством.

Важной отличительной чертой указанных показателей сопряженности является то, что они характеризуют близость не с отдельным вектором, являющимся представителем класса (например, с вектором, являющимся средним значением векторов класса), а с пространством, образованным всеми векторами анализируемого класса. Представляется, что это должно приводить к более полному учету всей имеющейся информации о классе в каждой конкретной ситуации. Указанное обстоятельство послужило мотивом для постановки задачи исследования эффективности мер сопряженности в задачах распознавания. При этом в качестве одной из основных ставилась задача выявить условия, при которых показатели сопряженности «работают» лучше.

Другой важной проблемой распознавания и анализа изображений является формирование признакового пространства. В задачах распознавания изображений в качестве признакового пространства часто используются непосредственно сами отсчеты значений яркости. При этом высокое разрешение изображений приводит к большим размерностям пространства признаков и значительным вычислительным затратам. Известным способом преодоления этой трудности является отбор наиболее информативных признаков, например, путем использования матрицы весов или перехода к системе признаков меньшей размерности, например, с помощью разложения Карунена-Лозва (Principal Component Analysis – PCA), что также требует значительных вычислительных затрат. Поэтому актуальна задача построения и исследования эффективных и простых в вычислительном отношении алгоритмов отбора информативных признаков. В настоящей работе исследуется возможность использования для этой цели мер мультиколлинеарности.

Следует подчеркнуть, что более чем за 40-летнюю историю развития теории распознавания образов разработан огромный арсенал эффективных методов и алгоритмов, реализуемых на различных этапах распознавания и обеспечивающих повышение качества классификации. Поэтому, предпринимая исследование, связанное с изучением эффективности некоторой меры близости в задачах распознавания, следует отдавать отчет в том, что такое исследование должно быть проведено в сочетании с наиболее общепризнанными процедурами и технологиями, получившими распространение в области распознавания изображений.

Такими «типовыми» процедурами, направленными на существенное повышение качества распознавания являются разбиение классов на подклассы (кластеризация образов), а также применение в качестве признаков инвариантов. Поэтому наряду с исследованиями эффективности обычных схем принятия решений, представляет интерес исследование мер мультиколлинеарности и сопряженности в сочетании с указанными известными алгоритмами. В частности, актуально проведение исследований алгоритма кластеризации образов, построенного на основе показателей сопряженности, а также исследование эффективности показателей сопряженности в пространстве инвариантов. Более того, эти исследования целесообразно провести на популярных в последнее время 3-D моделях лиц.

Таким образом, актуальной является задача построения решающих правил и алгоритмов распознавания изображений, в которых в качестве мер близости при принятии решений о при-

надлежности классу и отборе информативных признаков используются показатели сопряженности и мультиколлинеарности, а также исследование эффективности этих алгоритмов в сочетании с известными процедурами кластеризации и в пространстве инвариантов.

Исследования по теме диссертации выполнялись при поддержке российско-американской программы «Фундаментальные исследования и высшее образование», а также грантов РФФИ (гранты №01-01-00097, №03-01-00109, №05-01-08043-офи_а, № 06-08-01024).

Цель и задачи исследований.

Целью работы является достижение более высоких показателей качества в задачах распознавания и анализа изображений за счет применения для отбора информативных данных и принятия решений о принадлежности классу показателей сопряженности и мультиколлинеарности, и выявление условий, при которых они более эффективны. В соответствии с поставленной целью в рамках диссертационной работы решаются следующие задачи.

1. Построение решающих правил, основанных на использовании в качестве мер близости показателей сопряженности, и установление диапазона значений показателя мультиколлинеарности векторов образов, при которых достигается повышение качества распознавания изображений.

2. Сравнительное исследование вычислительной сложности показателей сопряженности и разработка методики их выбора с учетом числа обучающих объектов и размерности пространства признаков.

3. Исследование мер мультиколлинеарности в качестве критериев отбора информативных данных на изображениях и построение алгоритмов формирования векторов признаков по этим критериям.

4. Исследование возможности повышения качества распознавания за счет применения показателей сопряженности в алгоритмах кластеризации обучающих объектов.

5. Исследование возможности повышения качества распознавания с применением показателей сопряженности в пространстве суммирующих инвариантов, в т.ч. трехмерных изображений лиц.

Методы исследований.

В диссертационной работе используются методы теории распознавания образов, цифровой обработки изображений, а также математического анализа, линейной алгебры и теории групп.

Научная новизна работы.

В диссертации получены следующие новые научные результаты.

1. Разработаны новые решающие правила принятия решений в задачах распознавания на основе показателей сопряженности с пространством, натянутым на векторы анализируемого класса, и/или нуль-пространством соответствующей транспонированной матрицы, обеспечивающие повышение качества распознавания в значительном диапазоне значений показателя мультиколлинеарности векторов образов.

2. Предложена и обоснована методика выбора одного из показателей (сопряженности с пространством и/или с нуль-пространством) в зависимости от размерности пространства признаков и числа обучающих объектов.

3. Разработан новый алгоритм формирования признакового пространства, для случая использования в качестве признаков значений отсчетов яркости изображений, основанный на отборе информативных областей изображений по показателям мультиколлинеарности.

4. Показана возможность повышения качества распознавания за счет применения показателей сопряженности в алгоритмах кластеризации обучающих объектов.

5. Показана возможность повышения качества распознавания, в т.ч. трехмерных изображений лиц, при использовании решающих правил на основе показателей сопряженности в пространстве суммирующих инвариантов.

Апробация работы.

Основные результаты работы докладывались на следующих конференциях:

- Международной конференции «The 12th ISPE International Conference on Concurrent Engineering: Research and Applications», Даллас, США, 25-29 июля 2005;
- Международной конференции «The IASTED International Conference on Automation, Control, And Applications», Новосибирск, Россия, 20-24 июня, 2005;
- Всероссийской научной конференции «Математическое моделирование и краевые задачи», Самара, Россия, 29-31 мая, 2006;
- Международной конференции «The 3th International Conference on Pattern Analysis (ICPA 2006)», Будапешт, Венгрия, 26-28 мая, 2006;
- Международной конференции «The International Conference on Machine Learning and Data Mining MLDM 2007», Лейпциг, Германия, 18-20 июля, 2007;
- Международной конференции «The IEEE International Conference on Advanced Video and Signal based Surveillance», Лондон, Великобритания, 5-7 Сентября 2007, а также представлялись на следующих выставках:
- Третья окружная выставка «Российским инновациям – российский капитал», первый приз в категории «информационные технологии», Самара, 14-15 апреля 2005;
- Пятая межрегиональная выставка «Промышленный салон – 2006», Самара, 10-13 октября, 2006;
- Российская национальная выставка в Китае, Пекин, Китай, 17-22 ноября, 2006.

Основные положения диссертации, выносимые на защиту:

1. Решающие правила принятия решений в задачах распознавания, основанные на использовании показателей сопряженности с пространством, натянутым на векторы анализируемого класса, и/или нуль-пространством соответствующей транспонированной матрицы, обеспечивающие повышение качества распознавания в значительном диапазоне значений показателя мультиколлинеарности векторов образов.

2. Методика выбора одного из показателей (сопряженности с пространством и/или с нуль-пространством) в зависимости от размерности пространства признаков и числа обучающих объектов.

3. Алгоритм формирования признакового пространства, для случая использования в качестве признаков значений отсчетов яркости изображений, основанный на отборе информативных областей изображений по показателям мультиколлинеарности.

4. Результаты экспериментов, показывающие возможность повышения качества распознавания в значительном диапазоне значений показателя мультиколлинеарности векторов образов, за счет применения показателей сопряженности, в т.ч. в задаче кластеризации, в пространстве суммирующих инвариантов и трехмерных изображений лиц.

Публикации.

По теме диссертации опубликовано шестнадцать работ, в том числе два изобретения и четыре статьи в рецензируемых научных журналах и изданиях, определенных высшей аттестационной комиссией.

Структура и объем работы.

Диссертация состоит из введения, четырех глав, заключения, списка литературы и приложений. Она изложена на 127 страницах машинописного текста, содержит 51 рисунок, список использованных источников из 96 наименований.

КРАТКОЕ СОДЕРЖАНИЕ ДИССЕРТАЦИИ

Во введении обоснована актуальность работы, сформулированы ее цель и задачи. Проведен анализ текущего состояния проблемы распознавания лиц и подходов к ее решению. Приводятся перечень научных результатов и положения, выносимые на защиту.

В первой главе приведена формулировка задачи распознавания изображений. Предполагается, что имеется M изображений каждого из K объектов. Каждое изображение представляется вектором $\mathbf{x} = [x_1, x_2, \dots, x_N]^T$ размерности N , где x_1, x_2, \dots, x_N – признаки. Векторы, соответствующие изображениям одного объекта, составляют *класс*. Совокупность векторов признаков всех классов образует *обучающую выборку*. Решение задачи распознавания состоит в конструировании *решающей функции* $f: R^N \mapsto \{0, 1, 2, \dots, K\}$, которая каждому вектору \mathbf{x} ставит в соответствие некоторый класс. Для уменьшения числа неправильных классификаций вводится также класс с номером 0, соответствующий *отказу* в распознавании.

Из множества $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_M\}$ векторов каждого класса составляется $N \times M$ -матрица

$$\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_M]. \quad (1)$$

Этой матрице ставятся в соответствие так называемая информационная $M \times M$ -матрица:

$$\mathbf{A} = \mathbf{X}^T \mathbf{X} \quad (2)$$

и ковариационная $N \times N$ -матрица

$$\mathbf{B} = \mathbf{X} \mathbf{X}^T. \quad (3)$$

Предполагается, что $\text{rank} \mathbf{A} = M$. Известно также, что собственные значения $\lambda_i(\mathbf{A})$, $i = \overline{1, M}$ матрицы \mathbf{A} совпадают с ненулевыми собственными значениями матрицы \mathbf{B} , а собственные векторы матрицы \mathbf{B} , соответствующие ненулевым собственным значениям, образуют ортогональный базис (разложение Карунена-Лоэва).

Одним из наиболее важных факторов, который оказывает существенное влияние на качество распознавания, является взаимная ориентация (сопряженность) векторов образов и векторов признаков в соответствующих пространствах. Известно, что для характеристики почти линейной зависимости векторов принято использовать так называемые меры мультиколлинеарности: определитель информационной матрицы; спектральное число обусловленности; минимальное собственное значение и показатель максимальной сопряженности (в регрессионном анализе). В задачах идентификации предложено использовать также показатель диагонального преобладания:

$$\phi = \frac{\text{tr}^2 \mathbf{A}}{\text{tr} \mathbf{A}^2} = \left(\sum_{i=1}^M a_{ii} \right)^2 / \sum_{i,j=1}^M a_{ij}^2 = \left(\sum_{i=1}^M \lambda_i \right)^2 / \sum_{i=1}^M \lambda_i^2, \quad 1 < \phi \leq M. \quad (4)$$

Меры, в которых фигурируют собственные числа информационной матрицы, являются наиболее полными характеристиками мультиколлинеарности. Однако их вычисление представляет трудности, известные как проблема собственных значений. С учетом этого в работе сформулирована задача построения алгоритмов отбора информативных признаков (информативных областей на изображениях) на основе наиболее простой в вычислительном отношении меры мультиколлинеарности (4).

Показатель максимальной сопряженности, в том виде как он используется в регрессионном анализе, к сожалению, не может быть применен в задачах распознавания. Поэтому в работе, опираясь на идею анализа сопряженности векторов, сформулированы задачи построения соответствующих показателей и решающих правил для классификации изображений. При этом одной из основных задач работы, является установлению условий, при которых решающие правила, основанные на вычислении показателей сопряженности, более эффективны по сравнению с известными.

Во второй главе строятся решающие правила и исследуются алгоритмы распознавания, основанные на использовании в качестве мер близости показателей сопряженности. В рассмотрение вводится так называемый показатель сопряженности с подпространством, натянутым на векторы признаков образов объектов из заданного класса:

$$R_k = \frac{\mathbf{x}^T \mathbf{X}_k [\mathbf{X}_k^T \mathbf{X}_k]^{-1} \mathbf{X}_k^T \mathbf{x}}{\mathbf{x}^T \mathbf{x}}. \quad (5)$$

Здесь \mathbf{x} – вектор признаков неизвестного образа, предъявленный для установления близости к k -му классу, а \mathbf{X}_k – $N \times M$ -матрица, составленная из векторов образов, принадлежащих k -му классу.

Наряду с указанным, в работе рассматривается также показатель сопряженности с нуль-пространством транспонированной матрицы \mathbf{X}_k , который вычисляется как

$$S_k = \frac{\mathbf{x}^T \mathbf{T}_k \mathbf{T}_k^T \mathbf{x}}{\mathbf{x}^T \mathbf{x}}. \quad (6)$$

Здесь \mathbf{T}_k – матрица, составленная из собственных векторов, соответствующих нулевым собственным значениям матрицы $\mathbf{B}_k = \mathbf{X}_k \mathbf{X}_k^T$, а \mathbf{X}_k – $N \times M$ -матрица, та же, что и в (5). Между показателями R_i и S_i существует связь:

$$R_i = 1 - S_i. \quad (7)$$

С использованием указанных показателей сопряженности предложены новые решающие правила. В предположении, что для каждого (k -го) класса сформирована одна из следующих $N \times N$ -матриц $\mathbf{Q}_{(s)}$:

$$\mathbf{Q}_{k,R} = \mathbf{X}_k [\mathbf{X}_k^T \mathbf{X}_k]^{-1} \mathbf{X}_k^T \quad (8)$$

$$\text{или } \mathbf{Q}_{k,S} = \mathbf{T}_k \mathbf{T}_k^T, \quad (9)$$

соответствующая решающая функция $f(\mathbf{x})$ строится следующим образом. Вектор \mathbf{x} принадлежит m -му классу, то есть $f(\mathbf{x}) = m$, $m = 1, 2, \dots, K$,

$$\text{если } R_m = \max_k R_k, \text{ где } R_k = \frac{\mathbf{x}^T \mathbf{Q}_{k,R} \mathbf{x}}{(\mathbf{x}^T \mathbf{x})}, \quad (10)$$

$$\text{либо } S_m = \min_k S_k, \text{ где } S_k = \frac{\mathbf{x}^T \mathbf{Q}_{k,S} \mathbf{x}}{(\mathbf{x}^T \mathbf{x})}. \quad (11)$$

При использовании порогового значения T_0 , решающая функция дополняется правилом

$$f(\mathbf{x})=0, \text{ если } R_m \leq 1-T_0 \text{ или } S_m \geq T_0. \quad (12)$$

Показатели (5), (6) связаны соотношением (7), поэтому применение одного из них в каждой конкретной ситуации может определяться только объемом вычислительных затрат. Поскольку, вдобавок, решающие правила (10) и (11) отличаются только матрицами $\mathbf{Q}_{k,R}$ и $\mathbf{Q}_{k,S}$, достаточно сравнивать вычислительную сложность формирования только этих матриц.

В разделе 2.3 приведены полученные в работе аналитические оценки числа необходимых арифметических операций n^* для вычисления матриц $\mathbf{Q}_{k,R}$ и $\mathbf{Q}_{k,S}$:

$$n(\mathbf{Q}_{k,R}) = 2NM^2 + \frac{M(M-1)^2}{2} + 2M^3 + NM(2M-1) + N^2(2M-1), \quad (13)$$

$$n(\mathbf{Q}_{k,S}) = \sum_{i=M}^{N-1} \left[\frac{i(i^2+3i-1)}{3} + M(N-(i+1)) \right] + N^2(2N-2M-1). \quad (14)$$

На рис. 1, слева в качестве примера приведены построенные по полученным соотношениям графики зависимости числа арифметических операций для показателей R_k (возрастающий) и S_k (убывающий) в зависимости от M при $N=500$. На рис. 1, справа для сравнения приведены графики фактических временных затрат, полученные при тех же исходных данных в эксперименте.

На рис. 2 в координатах M, N приведена построенная по совокупности точек, соответствующих точкам пересечения кривых на графиках (рис. 1), для которых имеет равенство правых частей (13), (14), кривая равных вычислительных затрат. В области ниже кривой более выгодно использовать решающее правило (15), построенное с использованием показателя сопряженности с нуль-пространством.

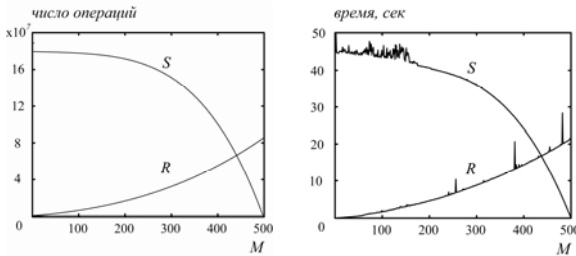


Рис. 1. Зависимость объема вычислений от M при $N = 200$: аналитическая оценка (слева); численный эксперимент (справа)

построенных решающих правил начинает проявляться лишь при числе векторов в классе более трех. Этого следовало ожидать. Ясно, что в случае или двух векторов в классе результат не может существенно отличаться от случая одного эталона.

Известно, что решение о принадлежности образа может оказаться ошибочным в случае, если векторы класса обучающей выборки сильно отличаются друг друга. Известный путь преодоления этой проблемы в разбиении обучающих классов на подклассы –

В разделе 2.2 диссертации приведен сравнительный пример распознавания лиц, взятых из известной базы данных ORL, с использованием решающих правил (10)-(12). Сравнение проводилось с методами распознавания, основанными на следующих мерах близости: евклидово расстояние; Манхэттенская метрика; косинус угла между векторами (частный случай меры (5) для класса, состоящего из одного вектора) и расстояние Махаланобиса. Результаты представлены на рис. 3. Заметим, что преимущество

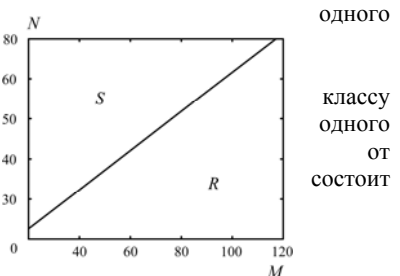


Рис. 2. Кривая равных объемов вычислений

кластеризации. В диссертации проведено исследование эффективности применения в качестве меры близости показателей сопряженности в составе одного из наиболее распространенных алгоритмов формирования кластеров по принципу максимального расстояния.

Для класса представленного в обучающей выборке множеством $\{x_1, x_2, \dots, x_M\}$ векторов образ алгоритм состоит в следующей последовательности шагов: выбор двух наиболее удаленных (по косинусу угла между векторами) образа, например, x_1 и x_2 (инициализация матриц $X_1 = [x_1]$ и $X_2 = [x_2]$) (1); произвольный выбор вектора x_i из числа оставшихся и вычисление показателя сопряженности со столбцовыми пространствами матриц X_1 и X_2 (2); добавление этого вектора x_i в качестве нового столбца к матрице, соответствующей ближайшему классу (3). Работа алгоритма прекращается, если все векторы класса распределены по кластерам (при этом сумма числа столбцов матриц X_1 и

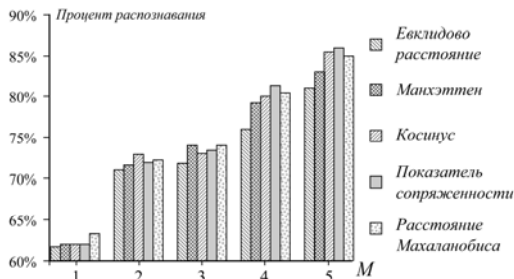


Рис. 3. Вероятность распознавания для различных мер в зависимости от числа векторов в обучающей выборке каждого класса

M), иначе повторяются последние два алгоритма. Алгоритм может применяться к му из полученных кластеров для дальнейшего разбиения на подклассы.

На рис. 4 приведен, полученный с использованием той же базы данных лиц сравнительный пример качества распознавания при использовании в описанном алгоритме кластеризации в качестве меры близости евклидова расстояния и показателей сопряженности. Процент распознавания в случае применения показателей сопряженности при возрастании числа векторов в классе растет быстрее.

Третья глава посвящена задаче отбора наиболее информативных областей на изображениях при формировании векторов признаков из значений яркости отсчетов изображений. Предполагается, что все изображения обучающей выборки имеют одинаковые размеры $N_1 \times N_2$, так что каждое изображение представляется $N_1 N_2 \times 1$ -вектором, компонентами которого являются значения яркости пикселей изображения «развернутые» по строкам или столбцам. Предположим также, что число изображений в каждом классе обучающей выборки равно M , а из векторов обучающей выборки строится $N \times M$ -матрица X . При этом задача отбора информативных признаков состоит в построении из X матрицы меньшей размерности $n \times M$ ($n < N$) путем исключения из нее строк, являющихся источником сильной мультиколлинеарности. В работе предложен следующий итерационный алгоритм отбора информативных отсчетов изображений.

При $k=1$ переменным n_k и m_k присваиваются начальные значения: $n_k = n_1 = N_1$, $m_k = m_1 = N_2$ $K = n_k \times m_k$. Далее алгоритм реализуется в виде следующей итерационной схемы.

X_2

равна шага каждого последующего

ORL,

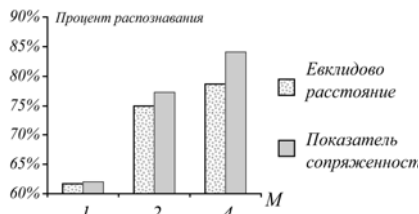


Рис. 4. Качество распознавания в зависимости от числа кластеров в классе

Шаг 1. Проверяется условие $K > M$. Если условие выполняется, значение k увеличивается на единицу и осуществляется переход к следующему шагу, если нет – к шагу 5.

Шаг 2. Изображения разбиваются на k^2 прямоугольных фрагментов, путем деления каждой стороны на k частей, и вычисляется размерность соответствующего изображению вектора: $K = n_k \times m_k$ где $n_k = n_{k-1}/k$, $m_k = m_{k-1}/k$. Результат округляется до ближайшего целого.

Шаг 3. Для каждого фрагмента, например, q -го изображения обучающей выборки, сформированных на шаге 2, составляется $K \times M$ -матрица $\mathbf{X}_{q,k}$, с использованием которой формируется соответствующая информационная $n \times n$ -матрица $\mathbf{A}_{k,q} = \mathbf{X}_{k,q}^T \mathbf{X}_{k,q}$.

Шаг 4. По информационной матрице вычисляется принятая мера мультиколлинеарности. Полученное значение присваивается всем отсчетам своего фрагмента на поле анализируемого изображения. Переход к шагу 1.

Шаг 5. Задается или определяется пороговое значение показателя мультиколлинеарности. Все пиксели изображения с равным ему или более высоким, чем заданный порог, значением яркости, включаются в число компонентов всех векторов \mathbf{x}_i обучающих объектов.

Схема формирования векторов образов из оставшихся отсчетов изображений (порядок обхода) не имеет значения. Важно, чтобы эта схема была одинаковой для всех M векторов. В результате реализации описанного алгоритма в зависимости от выбранного порога может быть достигнуто существенное снижение размерности векторов образов. Важнейшим с точки зрения вычислительной сложности и эффективности отбора данных в описанном алгоритме является выбор меры мультиколлинеарности. В работе проведены сравнительные исследования указанных в первой главе пяти мер мультиколлинеарности. Для сравнимости результатов в работе предложена модификация этих мер, заключающаяся в приведении их к одинаковому диапазону: $[0,1]$, путем следующих преобразований:

1. Нормировка информационной матрицы $\mathbf{A} = \{a_{ij}\}$ такая, что $\text{tr} \mathbf{A} = M$ (для этого матрица \mathbf{A} слева и справа умножается на диагональную матрицу $\mathbf{D} = \{d_{ii}\}$, где $d_{ii} = 1/\sqrt{a_{ii}}$).

2. Вместо спектрального числа обусловленности вводится в рассмотрение обратное спектральное число обусловленности: $K^{-1}(\mathbf{A}) = \lambda_{\min}/\lambda_{\max}$. Близость этого числа к 0 свидетельствует о плохой обусловленности матрицы \mathbf{A} .

3. Вместо показателя диагонального преобладания (4) рассматривается приведенный (к интервалу $[0,1]$) показатель диагонального преобладания $\bar{\phi} = (\phi - 1)/(M - 1)$.

Заметим, что близость всех мер к единице соответствует уменьшению степени мультиколлинеарности векторов образов.

На рис. 5 приведены графики изменения модифицированных показателей мультиколлинеарности в зависимости от размерности векторов образов (пороговых значений для отбора

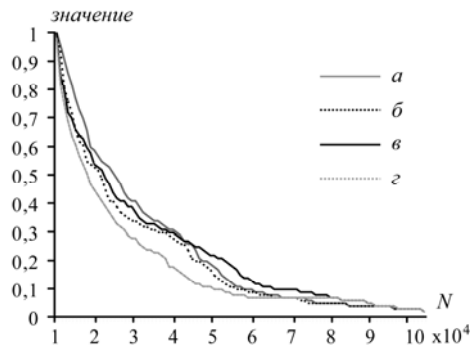


Рис. 5. Зависимость показателей $\det(\mathbf{A})$ (а), $K^{-1}(\mathbf{A})$ (б), $\lambda_{\min}(\mathbf{A})$ (в), $\bar{\phi}$ (г) от размерности векторов \mathbf{x}

информативных данных), сформированных с помощью описанного выше алгоритма. Нетрудно заметить, что характер изменения всех указанных показателей мультиколлинеарности при уменьшении размерности векторов образов примерно одинаков.

На рис. 6 для сравнения приведены также бинарные изображения полей информативности, полученные с использованием указанных 4-х показателей при достижении одинаковой размерности матрицы X . Областям, соответствующим наибольшим суммарным значениям показателей мультиколлинеарности, были присвоены соответственно большие значения яркости, поэтому они выглядят как более светлые.

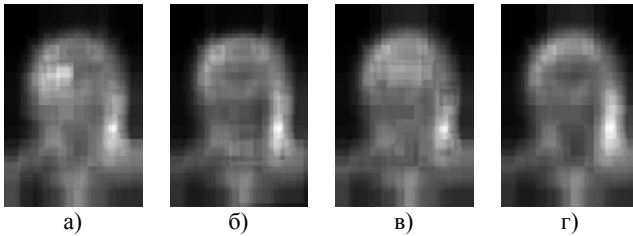


Рис. 6. Поля информативности, полученные с использованием $\det(A)$ (а), $K^{-1}(A)$ (б), $\lambda_{\min}(A)$ (в), $\bar{\phi}(z)$ (г)

Как и следовало ожидать, области фона, оказались наиболее темными, что соответствует представлениям об их малой информативности. Поскольку вид полей информативности при использовании различных показателей мультиколлинеарности отличается незначительно, целесообразно использовать показатель, обладающий наименьшей вычислительной

сложностью, в данном случае это показатель диагонального преобладания. В работе получены оценки для достижимого значения показателя диагонального преобладания на k шагах работы алгоритма: $\phi_{\max}(A) < \sqrt{(N_1 N_2) / M}$.

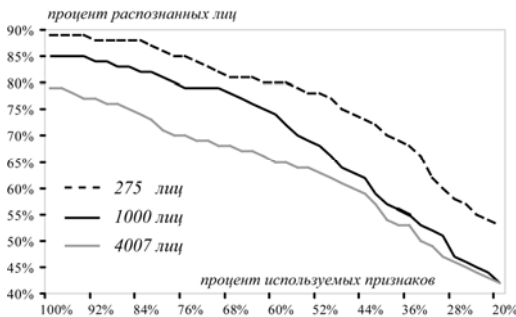


Рис. 7. Зависимость процента распознавания от доли используемых признаков

На рис. 7 представлены графики изменения процента распознавания при снижении размерности пространства признаков с использованием в качестве критерия информативности показателя диагонального преобладания. Результаты получены при различных объемах обучающей выборки. Как видно из графиков, при малом объеме обучающих данных, размерность можно снизить почти вдвое без существенной потери качества классификации.

Четвертая глава посвящена экспериментальным исследованиям эффективности решающих правил и алгоритмов анализа и распознавания изображений,

основанных на показателях сопряженности и мультиколлинеарности.

В разделе 4.1 диссертации приведены результаты по распознаванию цифр. В эксперименте использовались десять (0,...,9) выполненных пятью различными шрифтами цифр 20x16 пикселей. Для каждой цифры формировалось множество случайных реализаций путем сдвига (до трех пикселей), поворота (до 25 градусов), размытия и наложения гауссова шума.

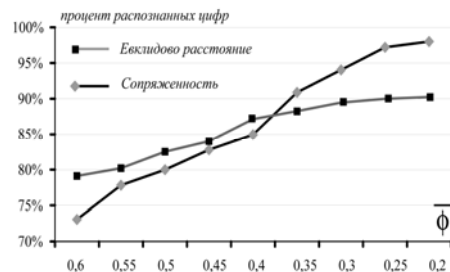


Рис. 8. Зависимость качества распознавания цифр от диагонального преобладания

Тестовые (распознаваемые) векторы (по 1000 для каждой цифры, т.е. всего 10000) формировались с использованием той же модели искажений.

На рис. 8 приведена одна из полученных характерных зависимостей числа распознаваемых объектов тестовой выборки от величины диагонального преобладания расширенной (построенной с использованием матрицы X и распознаваемого вектора) матрицы $\bar{A} = \bar{X}^T \bar{X}$. Приведенная зависимость, кроме прочего, позволяет обоснованно назначить диапазон значений показателя диагонального преобладания матрицы \bar{A} , при котором применение показателей сопряженности предпочтительнее.

В разделе 4.2. описана запатентованная информационная технология распознавания разрывов струи полимера в лабораторном анализе. На этапе предварительной обработки кадров изображений в технологии для выделения струи используется разработанный в 3 диссертации алгоритм определения информативных областей на изображении. На рис. 9 показан фрагмент сформированной области.

В разделе 4.3 приведены результаты экспериментов, показывающие эффективность применения показателей сопряженности в пространстве суммирующих инвариантов. Эксперимент проводился на базе трехмерных изображений лиц Face Recognition Grand Challenge (FRGC), включающей 275 человек с общим числом 4007 изображений, в т.ч. 943 обучающих.

Для каждой точки изображения инварианты вычислялись в размере 21×17 с использованием всех точек для трехмерного инварианта и для «горизонтального» и «вертикального» сечения внутри окна в случае двумерного инварианта. По вычисленным инвариантам для каждого изображения составлялся вектор признаков x . Далее классификация осуществлялась по той же схеме, что и в предшествующих экспериментах.

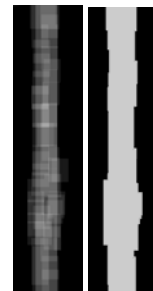


Рис. 9. Выделенный фрагмент изображения струи

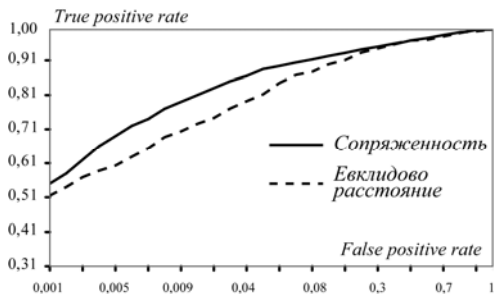


Рис. 10. Сравнение качества распознавания для одного инварианта

ния при использовании показателей сопряженности в качестве мер близости в пространстве суммирующих инвариантов.

На втором этапе наилучший результат, полученный для евклидова расстояния, сравнивался с результатами распознавания, полученными для того же инварианта с использованием показателя сопряженности.

Эксперименты проводились для 2-D евклидовых преобразований в биометрической среде распознавания ВЕЕ v.2.0 и оформлены в виде ROC-диаграмм (receiver operating characteristic), показанных на рис. 10. Сравнение полученных ROC-кривых показывает возможность достижения более высокого качества распознавания

ЗАКЛЮЧЕНИЕ

В диссертационной работе проведены исследования, связанные с применением показателей сопряженности и мультиколлинеарности в задачах анализа и распознавания изображений, в ходе которых получены следующие результаты.

1. Построены новые решающие правила, основанные на использовании в качестве мер близости показателей сопряженности с подпространством натянутым на векторы образов анализируемого класса и/или с нуль-пространством соответствующей транспонированной матрицы.
2. Построены аналитические оценки вычислительной сложности и на их основе предложена методика выбора конкретного показателя сопряженности в зависимости от размерностей пространства наблюдений и пространства признаков.
3. Исследована эффективность применения показателей сопряженности в известных алгоритмах кластеризации образов, показана возможность существенного повышения качества распознавания.
4. Предложен алгоритм, реализующий отбор информативных признаков по показателям мультиколлинеарности в случае, когда признаками являются значения яркостей пикселей изображения, проведены эксперименты, подтвердившие возможность выделения наиболее информативных участков изображений.
5. Обоснована возможность использования в предложенном алгоритме в качестве меры мультиколлинеарности показателя диагонального преобладания информационной матрицы. Получены оценки сверху для его достижимого значения на k шагах.
6. Экспериментально показана возможность повышения качества распознавания изображений с использованием показателей сопряженности в широком диапазоне значений показателя мультиколлинеарности векторов образов.
7. Показана возможность практического применения алгоритма анализа изображений с целью отбора наиболее информативных областей в информационной технологии распознавания разрывов струи полимера.
8. Экспериментально показана возможность повышения качества распознавания лиц с применением решающих правил на основе показателей сопряженности в пространстве суммирующих инвариантов.

Основные результаты опубликованы в следующих работах:

в ведущих рецензируемых научных изданиях определенных ВАК:

1. Козин, Н.Е. Поэтапное обучение радиальных нейронных сетей [Текст] / Н.Е. Козин, В.А. Фурсов // Компьютерная оптика. – 2004. – № 26. – С. 136-139.
2. Козин, Н.Е. Построение классификаторов для распознавания лиц на основе показателей сопряженности [Текст] / Н.Е. Козин, В.А. Фурсов // Компьютерная оптика. – 2006. – № 28. – С. 160-163.
3. Kozin, N. Building of classifiers based on conjugation indices [Text] / Kozin N., Kulagina I., Fursov V. // Optical Memory & Neural Networks. – 2007. – Vol. 16, No. 3. – pp. 136-141.
4. Козин, Н.Е. Снижение размерности векторов признаков по критериям мультиколлинеарности [Текст] / Н.Е. Козин, В.А. Фурсов // Компьютерная оптика. – 2008. – Том 3, № 32. – С. 307-311.
5. Козин, Н.Е. Распознавание лиц по показателям сопряженности в пространстве суммирующих инвариантов [Текст] / Н.Е. Козин, В.А. Фурсов // Компьютерная оптика. – 2008. – Том 4, № 32.

в других изданиях:

6. Козин, Н.Е. Автоматизированный анализ параллелизма программ [Текст] / Козин Н.Е., Фурсов В.А. // Сборник трудов 4-го Международного научно-практического семинара и Всероссийской молодежной школы «Высокопроизводительные параллельные вычисления на кластерных системах». – Самара, 2004.

7. Kozin, N. Stage-wise learning of radial neural networks [Text] / Kozin N., Fursov V. // The 12th ISPE International Conference on Concurrent Engineering: Research and Applications. Conference Proceeding. – Ft. Worth/Dallas, USA, 2005. – pp. 391-396.

8. Kozin, N. Algorithm for parallel learning of radial neural networks [Text] / Kozin N., Fursov V. // The IASTED International Conference on Automation, Control, And Applications (ACIT-ACA 2005). Conference Proceedings. Novosibirsk, 2005. – pp. 481-485.

9. Устройство для распознавания печатных и рукопечатных символов [Текст]: пат. 2285952 Рос. Федерация / Фурсов В.А., Казанский Н.Л., Козин Н.Е., Соيفер В.А.; заявитель и патентообладатель Институт систем обработки изображений Российской академии наук; опубл. 20.10.2006, Бюл. № 29.

10. Способ распознавания разрывов струи раствора на изображении [Текст]: пат. 2336563 Рос. Федерация / Фурсов В.А., Казанский Н.Л., Козин Н.Е., Попов С.Б.; заявитель и патентообладатель Институт систем обработки изображений Российской академии наук; опубл. 20.10.2008, Бюл. № 29.

11. Kozin, N. Constructing of Classifier for Face Recognition on the Basis of the Conjugation Indexes [Text] / Kozin N., Fursov V. // The Third International Conference on Pattern Analysis (ICPA 2006). Conference Proceedings. Budapest, Hungary, 2006.

12. Козин, Н.Е. Анализ вычислительной сложности алгоритмов по показателям сопряженности [Текст] / Козин Н.Е. // Сборник трудов Третьей Всероссийской научной конференции «Математическое моделирование и краевые задачи». – Самара, 2006.

13. Козин, Н.Е. Исследование алгоритмов распознавания на основе метода собственных лиц [Текст] // Козин Н.Е. // Сборник трудов Научно-технической конференции с международным участием «Перспективные информационные технологии в научных исследованиях, проектировании и обучении» (ПИТ-2006). – Самара, 2006.

14. Программа автоматического распознавания индексов «ПАРИ» [Текст]: свид. об официальной регистрации программы для ЭВМ № 2005611237 Рос. Федерация / Казанский Н.Л., Козин Н.Е., Попов С.Б., Фурсов В.А.

15. Kozin, N. Building of Classifiers Based on Conjugation Indices [Text] / Kozin N., Fursov V. // Proceedings of International Conference on Machine Learning and Data Mining (MLDM 2007). Conference proceedings. Leipzig, Germany, 2007.

16. Kozin, N. Recognition Through Constructing the Eigenface Classifiers using Conjugation Indices [Text] / Kozin N., Fursov V. // IEEE International Conference on Advanced Video and Signal based Surveillance. Conference proceedings. London, United Kingdom, 2007.