О РЕШЕНИИ ЗАДАЧИ РАСПОЗНАВАНИЯ ИЗОБРАЖЕНИЙ МЕТОДОМ ГЛАВНЫХ КОМПОНЕНТ И ЛИНЕЙНЫМ ДИСКРИМИНАНТНЫМ АНАЛИЗОМ

Мокеев В.В., Томилов С.В.
Южно-Уральский государственный университет
(национальный исследовательский университет)

Аннотация

В работе рассматриваются некоторые аспекты применения метода главных компонент (МГК) и линейного дискриминантного анализа (ЛДА) для решения задачи распознавания изображений. Основная идея такого подхода заключается в том, что сначала изображения лиц проецируются из исходного пространства признаков в редуцированное подпространство главных компонент, а затем для разделения классов изображений используется линейный дискриминантный анализ. В статье исследуется эффективность применения МГК и ЛДА к задаче распознавания изображений лиц без их предварительной нормализации. Если число изображений в классе невелико, предлагается дополнять учебную выборку изображениями, полученными путём поворота, масштабирования и зеркалирования. На изображениях баз данных ORL и Feret изучается влияние расширения учебной выборки на качество распознавания ненормализованных изображений лиц. Также рассматривается задача повышения эффективности расчёта главных компонент для больших наборов изображений. Метод линейной конденсации представляет новую технологию расчёта главных компонент больших матриц. Для повышения эффективности метода линейной конденсации предлагается использовать алгоритм блочно-ортогональной конденсации. Оценивается точность и быстродействие разработанного алгоритма.

<u>*Ключевые слова:*</u> распознавание лиц, метод главных компонент, линейный дискриминантный анализ, метод линейной конденсации.

Введение

Большинство систем распознавания изображений основывается на проекционных методах построения пространства признаков меньшей размерности. Одним из широко распространенных методов сокращения размерности изображений является метод главных компонент (МГК) [1]. В настоящее время для решения задачи поиска и распознавания лиц предлагается множество алгоритмов, использующих МГК.

Технология распознавания изображений с использованием проекционных методов включает два этапа. На первом этапе выполняется построение классификатора с использованием учебного набора изображений. На втором этапе осуществляется распознавание неизвестных изображений с помощью построенного классификатора.

Для построения классификатора используются различные методы, из которых можно отметить линейный дискриминантный анализ (ЛДА) [2, 3]. ЛДА позволяет преобразовать исходное пространство изображений в низкоразмерное пространство признаков, в котором изображения классов группируются вокруг их центров, а центры классов удаляются друг то друга настолько, насколько это возможно.

В работах [4, 5] предлагается подход для решения задач распознавания изображений, который основан на совместном использовании МГК и ЛДА. Для вычисления главных компонент формируется учебная выборка, состоящая из изображений, сгруппированных в классы. Изображения одного класса описывают лицо одного человека. Один класс может содержать десятки и даже сотни изображений одного лица. Сначала для сокращения размерности изображений используется метод главных компонент, а затем применяется ЛДА для разделения классов изображений.

При распознавании изображений обычно выполняется предварительная обработка изображений, приводящая их к стандартной форме (масштаб, центрирование, отсечение фона, выравнивание яркости). Нежелательно также наличие очков, бороды, мимики и т.п., так как алгоритм при распознавании начинает реагировать больше, например, на наличие очков, чем на межклассовые отличия. Нормализация изображений требует дополнительных вычислений, что на этапе построения классификатора не является критичным, но на втором этапе не всегда приемлемо в системах, работающих в режиме реального времени. Кроме того, нормализация изображений может приводить в некоторых случаях к потере информативности изображений, что ограничивает рост качества распознавания. Поэтому одной из задач, представляющих интерес, является разработка систем распознавания лиц без предварительной обработки изображений (нормализации). В работе рассматривается технология распознавания ненормализованных изображений лиц с помощью подхода МГК+ЛДА. Одним из факторов, повышающих качество распознавания лиц, является увеличение числа изображений в классе. Если число изображений в классе небольшое, то обучающую выборку можно расширить изображениями, полученными путём зеркального отражения, поворота и масштабирования исходных изображений лиц. Такой приём слабо влияет на качество распознавания в случае нормализации изображений лиц, так как процедура нормализации направлена на уменьшение подобных различий между изображениями. Однако в случае ненормализованных изображений такой приём может оказаться полезным, когда в учебной выборке содержится небольшое число изображений. Предлагаемая технология распознавания ненормализованных изображений обеспечивает хорошее качество распознавания только при большом числе изображений в классе, что увеличивает временные затраты на построение классификатора, но уменьшает затраты при распознавании неизвестных изображений, так как не требует процедуры нормализации.

В связи с тем, что на этапе построения классификатора учебная выборка может достигать больших размеров, трудоёмкость вычисления главных компонент может существенно расти. Поэтому в работе рассматривается задача повышения эффективности расчёта главных компонент для большого количества изображений. В работах [6–8] описываются подходы, снижающие трудоёмкость вычисления главных компонент больших наборов изображений. Алгоритмы двумерного анализа главных компонент исследуются в работе [6]. Главные компоненты двумерного МГК отражают различия между строками и столбцами изображений и не учитывают различия между отдельными точками изображения, которые могут быть полезны при распознавании изображений. Основное преимущество двумерного МГК заключается в том, что он снижает размерность матриц при вычислении главных компонент. Метод синтеза главных компонент базируется на делении набора изображений на части, получении частных решений и синтезе главных компонент из частных решений [7]. Метод линейной конденсации [8] использует понижение порядка матрицы при вычислении главных компонент. Идеи понижения порядка матриц при вычислении собственных значений показали высокую эффективность для решения обобщённой задачи собственных значений методами частотно-динамической [9] и частотной конденсации [10]. Метод линейной конденсации адаптирует эти идеи для стандартной задачи собственных значений. В данной работе для вычисления главных компонент больших наборов высокоразмерных изображений предлагается использовать алгоритм блочно-ортогональной конденсации, который является развитием алгоритма многоуровневой линейной конденсации [8].

1. Основные соотношения подхода МГК и ЛДА

Метод, основанный на МГК и ЛДА, состоит из двух шагов: сначала мы проецируем изображение лица из исходного пространства признаков в подпространство собственных лиц с помощью метода главных компонент, затем мы используем ЛДА, чтобы получить линейный классификатор. Допустим, существует набор изображений, каждое из которых описывается вектором x_i (i=1,2,3 ,..., m), где m – число различных изображений в обучающем наборе. Размерность n вектора x_i равна числу пикселей изображения. Таким образом, все изображения могут быть представлены в виде матрицы, строками которой являются векторы x_i. Средний вектор обучающих изображений $\nu = \dfrac{1}{m}\sum_i^m x_i$ вычитается из каждого изображения в обучающем наборе. Таким образом, получается новое пространство $\mathbf{X}^0$ размерностью $m\times n$, строками которого являются векторы $x_i^0 = x_i - \nu$.

Подход МГК+ЛДА можно рассматривать как поэтапное линейное преобразование пространства исходных изображений в проекцию пространства меньшей размерности. На первом этапе выполняется преобразование, позволяющее уменьшить размерность каждого изображения с n до p ($p \ll n$). Метод главных компонент является техникой снижения размерности, основанной на извлечении желаемого числа главных компонент из многомерных данных. Первая главная компонента представляет линейную комбинация исходных признаков (пикселей), которая имеет максимальную дисперсию, а i-я главная компонента является линейной комбинацией исходных признаков (пикселей) с самой высокой дисперсией среди m–i+1 главных компонент и ортогональной i–1 первым главным компонентам.

Известно, что матрица $\mathbf{X}^0$ может быть представлена в виде сингулярного разложения.

$$\mathbf{X}^0 = \mathbf{U}_{pca}\,\mathbf{\Lambda}\,\mathbf{V}_{pca}^T, \quad (1)$$

где $\mathbf{U}_{pca}(m\times p)$ и $\mathbf{V}_{pca}$ $(n\times p)$ – матрицы левых и правых собственных векторов $\mathbf{X}^0$, $\mathbf{\Lambda}$ – диагональная матрица $(p\times p)$, диагональные элементы которой $\eta_1, \eta_2, ..., \eta_p$ являются положительными собственными значениями матрицы $\mathbf{X}^0$. Здесь p – число собственных векторов, обычно p не превышает ранга матрицы $\mathbf{X}^0$.

Ключевым моментом МГК является вычисление матрицы главных компонент $\mathbf{V}_{pca}$. Матрица главных компонент $\mathbf{V}_{pca}$ формируется из правых собственных векторов v_{0i}^{pca}, которым соответствуют наибольшие собственные значения.

Матрица собственных векторов $\mathbf{V}_{pca}$ может определяться как

$$\mathbf{V}_{pca}^T = \mathbf{\Lambda}^{-1}\mathbf{U}_{pca}^T\mathbf{X}^0. \quad (2)$$

Матрица левых собственных векторов $\mathbf{U}_{pca}$ образуется из собственных векторов уравнения

$$(\mathbf{A} - \lambda\,\mathbf{I})\mathrm{u}_0^{pca} = 0, \quad (3)$$

где $\mathbf{I}$ – единичная матрица с порядком m, u_0^{pca} – собственный вектор, а λ – собственное значение, $\mathbf{A} = \frac{1}{m}(\mathbf{X}^0)^T\mathbf{X}^0$ – ковариационная матрица.

Главные компоненты изображений вычисляются по формуле

$$\mathbf{Z} = \mathbf{X}^0\,\mathbf{V}_{pca},$$

где $\mathbf{Z}$ – матрица $(m\times p)$ главных компонент изображений, i-я строка которой представляет вектор главных компонент i-го изображения. Будем обозначать далее вектор главных компонент изображения как z_i^k, где k – номер класса, i – номер изображения.

Обозначим средний вектор главных компонент изображений, принадлежащих классу k, как ν_k, а среднее значение главных компонент всех изображений как ν:

$$\nu_k = \frac{1}{m_k}\sum_{i=1}^{m_k}\mathrm{z}_i^k,\ \nu = \frac{1}{K}\sum_k^K\frac{1}{m_k}\sum_{i=1}^{m_k}\mathrm{z}_i^k. \quad (4)$$

Здесь K – число классов, а m_k – число изображений лиц в классе k.

Матрица межклассовых различий $\mathbf{A}_b$ может быть вычислена как

$$\mathbf{A}_\omega = \frac{1}{m}\sum_{k=1}^{K}\sum_{i=1}^{m_k}(z_i^k - \nu_k)(z_i^k - \nu_k)^T. \quad (5)$$

Матрица внутриклассовых различий определяется как

$$\mathbf{A}_b = \frac{1}{m}\sum_{k}^{K} m_k(\nu_k - \nu)(\nu_k - \nu)^T. \quad (6)$$

ЛДА пытается найти такое преобразование, которое максимизирует отношение межклассовых различий к внутриклассовым различиям, как показано ниже:

$$\mathbf{V}_{lda} = \arg\max_{\mathbf{V}\in R^{n\times r}} \frac{\left|\mathbf{V}^T\mathbf{A}_b\mathbf{V}\right|}{\left|\mathbf{V}^T\mathbf{A}_\omega\mathbf{V}\right|}. \quad (7)$$

Чтобы определить $\mathbf{V}_{lda}$, решается задача собственных значений

$$(\mathbf{A}_b - \lambda\,\mathbf{A}_\omega)\,\mathrm{v}_0^{lda} = 0. \quad (8)$$

Решение уравнения (8) представляет обобщённую задачу собственных значений. В работе [11] показано, что для решения уравнения (8) эффективно использовать обобщённый метод Якоби.

Комбинируя МГК и ЛДА, мы получаем матрицу линейного преобразования, которая проецирует изображение: сначала в подпространство главных компонент $\mathbf{Z}$, а потом в пространство классификации:

$$\mathbf{W} = \mathbf{V}_{lda}\mathbf{V}_{pca}, \quad (9)$$

где $\mathbf{V}_{lda}$ – линейное дискриминантное преобразование в пространстве главных компонент. После такого линейного преобразования распознавание изображения выполняется в пространстве дискриминантных компонент с использованием различных метрик, например, Евклидова расстояния.

2. Исследование качества распознавания ненормализованных лиц подходом МГК+ЛДА

База данных ORL содержит изображения 40 человек, каждый из которых описывается 10 различными изображениями лиц, имеющих различную мимику, ракурс, прорисовку деталей. Все изображения в базе полутоновые, с 256 градациями яркости. Размер каждого изображения – 92×112 пикселей. На рис. 1 приведены примеры изображений восьми лиц из базы данных ORL.

Размер изображений лиц базы данных FERET равен 512×768 пикселей [12]. Однако в экспериментах с целью снижения вычислительных затрат используются изображения размером 128×128 пикселей. Таким образом, каждое изображение представлялось в виде вектора размерностью 16384. Изображения лиц выделяются из изображений базы FERET и описывают 121 человека. Изображения лиц выделяются с помощью метода Виолы–Джонса, но не проходят процедуру геометрической нормализации. Для каждого человека имеется 10 изображений различного размера, полученных при различном освещении в разных ракурсах с произвольной мимикой.

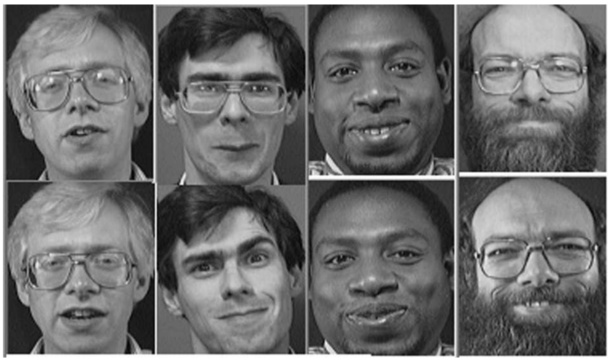


Рис. 1. Примеры лиц, выделенных из изображений базы данных ORL

Общее число изображений в учебной выборке составляет: 1210. На рис. 2 показаны изображения восьми лиц из базы данных FERET.

Рис. 2. Примеры лиц, выделенных из изображений базы данных FERET

Подход МГК+ЛДА используется для распознавания изображений и включает два этапа.

На первом этапе вычисляются главные компоненты учебной выборки, что позволяет сократить размерность каждого изображения до p главных компонент. Для формирования главных компонент используются собственные векторы уравнения (1) с наибольшими собственными значениями. Далее строятся матрицы внутриклассовых и межклассовых различий ($p\times p$) и вычисляются $r \le p$ дискриминантных компонент. При распознавании используется классификатор по ближайшему центру классов. В качестве тестовой выборки используются изображения, не попавшие в учебную выборку.

Качество распознавания людей по изображениям лиц описывается коэффициентом распознавания, который численно равен проценту правильно распознанных изображений из общего числа предъявленных. Коэффициенты распознавания K_{test} определяются для тестовой выборки.

Первая часть экспериментальных исследований выполняется на изображениях базы данных ORL. Эксперименты проводятся для исследования влияния количества изображений в классе учебной выборки на качество распознавания. Первые три серии экспериментов проводятся для учебных выборок, содержащих 2, 3 и 4 изображения в каждом классе, которые отбираются случайно. При использовании метода МГК+ЛДА для распознавания применяется классификатор по ближайшему соседу [11].

При проведении исследований используется процедура кросс-валидации, усредняющая коэффициенты распознавания, полученные при различном делении набора изображений на учебную и тестовую выборки. Изображения учебной выборки выбираются из базы ORL случайным образом. Все оставшиеся изображения составляют тестовую выборку. В каждой серии проводится 10 экспериментов.

В работе [13] экспериментальные исследования с нормализованными изображениями базы данных ORL выполняются различными методами: МГК, ЛДА, DLA (Discriminative Locality Alignment [14]), NDLPP (Null space Discriminant Locality Preserving Projections[15]), RLPDA (Regularized Locality Preserving Discriminant Analysis [13]).

В табл.1 представлены коэффициенты распознавания, полученные с помощью перечисленных выше методов [13]. В этой же таблице представлены коэффициенты распознавания, полученные с помощью классификатора, построенного подходом МГК+ЛДА. При построении классификатора использовались ненормализованные изображения лиц. Полученные результаты хорошо согласуются с данными экспериментальных исследований других авторов.

Таблица 1. Сравнение коэффициентов распознавания изображений базы ORL

Метод	Число изображений в классе		
	2	3	4
МГК [13]	69,6	78,6	83,6
ЛДА [13]	80,1	88,0	91,5
DLA	73,3	87,1	92,6
NDLPP	83,0	91,3	94,6
RLPDA	80,7	90, 4	94,8
МГК+ЛДА	82,2	90	94,2

Как видно из таблицы, коэффициенты распознавания, полученные подходом МГК+ЛДА без нормализации изображений, имеют достаточно высокие значения, хотя и уступают таким методам, как NDLPP и RLPDA.

Следующие серии экспериментов проводятся на учебных выборках первых трех серий испытаний, в которые добавляются изображения, полученные путём поворота, масштабирования и зеркалирования исходных изображений учебной выборки. Полученные учебные выборки используются для построения классификаторов подходом МГК+ЛДА.

В табл. 2 представлены коэффициенты распознавания, полученные помощью построенных классификаторов. В скобках указано число изображений в классе, полученное в результате расширения выборки. Первая строка содержит результаты, полученные с помощью учебной выборки, дополненной зеркалированными изображениями. Число изображений в классе увеличивается в два раза, также немного растут и коэффициенты распознавания.

Вторая строка содержит результаты, полученные при расширении учебной выборки изображениями, уменьшенными и увеличенными на 5 %. Число изображений в классе увеличивается в три раза, качество распознавания лиц заметно улучшается.

Таблица 2. Сравнение коэффициентов распознавания при расширении учебной выборки базы ORL

Способ расширения учебной выборки	Число исходных изображений в классе		
	2	3	4
Зеркалирование	84,3 (4)	91,1 (6)	94,5 (8)
Масштабирование	85,6 (6)	96,4 (9)	95,8 (12)
Поворот	85,4 (6)	91,8 (9)	94,6 (12)
Зеркалирование и поворот	85,7 (12)	92,2 (18)	95,1(24)
Масштабирование, поворот	85,6(18)	96,4 (27)	96, 5 (36)
Зеркалирование, поворот, масштабирование	85,6 (36)	96, 3 (54)	94,1 (72)

Третья строка содержит коэффициенты распознавания, полученные при расширении учебной выборки изображениями, повёрнутыми по часовой и против часовой стрелки на 4˚. Число изображений в классе увеличивается в 3 раза.

Четвёртая строка таблицы представляет результаты, полученные при расширении учебной выборки изображениями, которые сначала зеркалируются, а потом поворачиваются на 4˚ по часовой и против часовой стрелки. Число изображений в классе в этом случае увеличивается в 6 раз, коэффициенты распознавания немного выше результатов, полученных при расширении учебной выборки только зеркалированными или только повёрнутыми изображениями.

Пятая строка содержит результаты, полученные при следующем расширении учебной выборки. Сначала добавляются изображения, полученные из исходных изображений путём их уменьшения и увеличения на 5 %. Затем в расширенную выборку добавляются изображения, полученные путём поворота изображений по часовой и против часовой стрелке на 4˚. В результате число изображений в классе увеличивается в 9 раз, а коэффициенты распознавания самые большие.

Комбинация зеркалирования изображений, их масштабирования и поворота приводит к максимальному расширению учебной выборки. Однако классификатор, построенный на основе этой выборки, показывает не самое высокое качество распознавания лиц.

Вторая часть экспериментальных исследований выполняется на выборке изображений из базы данных FERET. Выборка содержит 1210 изображений лиц 121 человека. Для каждого человека имеется 10 изображений лиц различного размера и освещения, полученные при различных ракурсах с произвольной мимикой.

Проводятся эксперименты, в которых варьируется число изображений лиц в классе. Первые три серии экспериментов проводятся для учебных выборок, содержащих 3, 4 и 5 изображений в каждом классе базы FERET. При проведении исследований также используется процедура кросс-валидации. Заданное число изображений учебной выборки выбирается из базы FERET случайным образом. Все оставшиеся изображения составляют тестовую выборку. В каждой серии проводится 10 экспериментов.

В работе [16] представлены результаты исследования точности распознавания лиц, выполненные на нескольких базах данных, в том числе и на базе данных FERET. Изображения лиц проходят процедуру нормализации, включая центрирование, вращения, масштабирование, выравнивание по яркости и т.п. Для распознавания используются различные методы: МГК, ЛДА, NPE (Neighborhood Preserving Embedding [17]), MFA (Marginal Fisher Analysis [18]), LSDA (Locality Sensitive Discriminant Analysis [19]), RLPDE (Regularized Locality Preserving Discriminant Embedding [16]).

В табл. 3 представлены коэффициенты распознавания, вычисленные методами МГК, ЛДА, NPE, MFA, LSDA, RLPDE и МГК+ЛДА. Коэффициенты распознавания, полученные данными методами, сравниваются с результатами подхода МГК+ЛДА без нормализации изображений.

Таблица 3. Сравнение коэффициентов распознавания изображений базы FERET

Метод	Число изображений в классе		
	3	4	5
МГК	59,7	60	60
ЛДА	64	63,7	64
NPE	61,3	60	61,3
MFA	66,4	61,3	63,3
LSDA	64	63,6	65,6
RLPDE	75,4	75,4	75,4
МГК+ЛДА	52,6	59,3	65,6

Как видно из таблицы, увеличение числа изображений в классе не приводит к существенному увеличению качества распознавания лиц, если используются методы МГК, ЛДА, NPE, MFA, LSDA, RLPDE. Это объясняется тем, что процедура нормализация не только устраняет различия в лицах, но и может вносить искажения в изображения лиц. Для классификаторов, построенных с помощью ненормализованных изображений подходом МГК+ЛДА, качество распознавания лиц заметно повышается с ростом числа изображений в классе.

Последующие три серии экспериментов проводятся для учебных выборок ненормализованных изображений, которые увеличиваются путём добавления изображений, полученных поворотом, масштабированием и зеркалированием исходных изображений. В каждой серии проводятся 10 экспериментов, результаты которых усредняются. В табл. 4 представлены коэффициенты распознавания, полученные помощью классификаторов ЛДА+МГК, построенных на расширенных учебных выборках. В скобках указано число изображений в классе, полученное в результате расширения выборки.

В таблице представлены результаты, полученные, когда учебная выборка расширяется зеркалированными изображениями (первая строка), увеличенными и уменьшенными на 5 % изображениями (вторая строка), изображениями, повёрнутыми по часовой и против часовой стрелки на 4˚ (третья строка).

Таблица 4. Сравнение коэффициентов распознавания при расширении учебной выборки базы FERET

Способ расширения учебной выборки	Число изображений в классе		
	3	4	5
Зеркалирование	58,2 (6)	64,2 (8)	70 (10)
Масштабирование	64,4 (9)	70,5 (12)	75,3 (15)
Поворот	64 (9)	70 (12)	72 (15)
Масштабирование, поворот	65,7 (27)	73 (36)	76 (45)
Зеркалирование и поворот	65,5 (18)	69,9(24)	74,5 (30)
Масштабирование, поворот, зеркалирование	67,7 (54)	72,8 (72)	75,6 (90)

Четвёртая строка содержит результаты, полученные при расширении учебной выборки изображениями, которые сначала увеличиваются и уменьшаются на 5 %, а затем все изображения расширенной учебной выборки поворачиваются на 4˚ по часовой и против часовой стрелки. Как видно из таблицы, классификатор, построенный на такой учебной выборке, даёт наилучшие результаты распознавания.

Комбинация поворота изображений на 4˚ по часовой и против часовой стрелки и зеркалирования всех полученных изображений позволяет получить результаты, представленные в пятой строке таблицы.

Комбинация, сочетающая все три способа расширения учебной выборки, приводит к максимальному увеличению числа изображений в классе, однако классификатор, построенный по такой выборке, демонстрирует не самые лучшие результаты.

Как видно из таблицы, расширение учебной выборки производными изображениями, полученными путём поворота и масштабирования исходных изображений, показывает стабильный рост коэффициента распознавания. Зеркалирование изображений в сочетании с масштабированием и поворотом не даёт повышения качества распознавания. Это может быть связано с тем, что такое расширение учебной выборки приводит росту похожих изображений, что снижает информативность учебной выборки и качество построенных на её основе классификаторов.

Следует отметить, что расширение учебной выборки изображениями, полученными в результате их поворота, масштабирования и зеркалирования, целесообразно только в случае отсутствия большого числа оригинальных изображений.

В целом, подводя итог, можно сказать, что увеличение числа изображений в классе позволяет построить достаточно качественные классификаторы на основе методов МГК+ЛДА без применения процедуры нормализации изображений.

Рост числа изображений в классах приводит к увеличению вычислительных затрат при построении классификатора, что связано с тем, что трудоёмкость вычисления главных компонент существенно растёт с увеличением числа изображений в учебной выборке.

3. Вычисление главных компонент больших наборов изображений

Основная трудоёмкость расчёта главных компонент связана с вычислением собственных векторов уравнения (3). В случае большого числа изображений в учебной выборке порядок матриц уравнения (3) становится большим и для вычисления главных компонент требуются значительные вычислительные ресурсы.

Среди методов решения проблемы собственных значений можно выделить следующие группы методов: итерационные и преобразования подобия [20]. К группе итерационных методов относятся степенной метод, метод Ланцоша и т.д. Главный недостаток итерационных методов заключается в том, что на скорость сходимости решений влияет отношение величины искомого собственного значения к ближайшему собственному значению.

Методы преобразований подобия применяются с целью получить из исходной матрицы новую матрицу с теми же собственными значениями, но более простого вида. Наиболее известными являются методы Якоби, Гивенса и Хаусхолдера. Метод Хаусхолдера позволяет получить требуемый результат быстрее, чем метод Гивенса, так как связан с выполнением меньшего числа, хотя и более сложных преобразований. Методы понижения порядка матриц (методы конденсации) относятся к группе методов преобразования подобия, так как их целью является получение матрицы меньшего порядка, которая была бы подобна исходной матрице в том смысле, что собственные значения этих матриц в заданном диапазоне совпадали бы с заданной точностью.

В связи с тем, что собственные значения находятся достаточно близко друг к другу, особенно в верхней части спектра, итерационные методы являются малоэффективными при вычислении собственных векторов больших матриц. В работе [8] показывается, что метод Хаусхолдера демонстрирует более высокую скорость, чем степенной метод.

Эффективным способом решения неполной задачи собственных значений для больших матриц является понижение порядка матриц. Так как число требуемых собственных векторов обычно намного меньше порядка матриц, такой подход является особенно эффективным в случае матриц больших размеров. В работе [9] предлагается метод частотно-динамической конденсации, который понижает порядок матриц при решении обобщённой задачи собственных значений. В дальнейшем метод широко использовался для решения различных задач инженерного анализа [10]. Идея метода частотно-динамической конденсации послужила основой метода линейной конденсации, который при решении стандартной задачи собственных значений понижает порядок матриц, сохраняя собственные значения в заданном интервале [8].

Представим уравнение (3) в виде

$$(\mathrm{I}-\mu A)\,\mathrm{u}_0^{pca}=0\,, \tag{10}$$

где $\mu=1/\lambda$. Тогда задача формулируется следующим образом: определить наименьшие собственные значения и соответствующие им собственные векторы уравнения (10).

В работе [8] описывается алгоритм многоуровневой линейной конденсации для вычисления собственных значений в интервале от 0 до μ_2 и соответствующих им собственных векторов. Алгоритм включает пять шагов. Первый шаг представляет собой процедуру многоуровневого понижения порядка матриц, которая начинается с того, что все элементы вектора u_0^{pca}, которые в дальнейшем будем называть признаками, сортируются в порядке убывания диагональных коэффициентов ковариационной матрицы **A**. На каждом уровне процедуры понижения порядка матриц выбирается группа признаков с минимальными диагональными коэффициентами. Решение об исключении выбранных признаков принимается при выполнении условия

$$\mu_{\min} > k_c\mu_2\,. \tag{11}$$

Здесь $\mu_{\min}=1/\lambda_{\max}$ – это наименьшее собственное значение блока удаляемых переменных $\mathbf{A}_{ss}^k$, k_c называется параметром отсечения. Если условие (11) выполняется, то осуществляется понижение порядка матрицы. Условие (11) может быть переписано в виде $\lambda_{\max}<\lambda_2/\kappa_c$.

Как известно, сумма диагональных коэффициентов матрицы равна сумме собственных значений λ. Поэтому набор удаляемых признаков формируется из признаков, которым соответствуют наименьшие диагональные коэффициенты матрицы **A**. По сути, делается предположение о том, что чем меньше сумма собственных значений λ, тем меньше максимальное собственное значение анализируемой матрицы. В процессе понижения матрицы мы будем последовательно получать матрицы $\mathbf{A}_1$, $\mathbf{A}_2,\ldots,$ $\mathbf{A}_k=\begin{bmatrix}\mathbf{A}_{bb}^k & \mathbf{A}_{bs}\\ \mathbf{A}_{sb}^k & \mathbf{A}_{ss}^k\end{bmatrix}$, где k – уровень понижения порядка матрицы **A**.

Эффективность алгоритма многоуровневой линейной конденсации зависит от того, насколько сильно понижается порядок матриц, однако степень понижения не всегда оказывается достаточно высокой. В работе [8] демонстрируется быстродействие алгоритма многоуровневой линейной конденсации на примере вычисления 67 главных компонент наборов различной размерности (от 500 до 5000). Однако дальнейшие исследования показали, что при расчёте большого числа главных компонент (несколько сотен) не удаётся существенно снизить порядок матриц. Это обусловлено тем, что собственные значения в районе верхней границы μ_2 лежат очень плотно и не удаётся выбрать исключаемые признаки, не нарушая условия (11).

Таким образом, алгоритм многоуровневой линейной конденсации перестаёт быть эффективным. Для повышения быстродействия расчёта собственных векторов предлагается использовать алгоритм блочно-ортогональной конденсации.

Алгоритм блочно-ортогональной конденсации базируется на процедуре многоуровневого понижения порядка матрицы. Однако на каждом уровне блок признаков (кандидатов на удаление) приводится к диагональной форме с помощью ортогонального преобразования.

Пусть уравнение (10) на k-м уровне понижения порядка матриц имеет вид

$$(\mathbf{I}_k - \mu \mathbf{A}_k) u_{k-1}^{pca} = 0\,. \tag{12}$$

Представим матрицу $\mathbf{A}_k$ и вектор u_{k-1}^{pca} в форме

$$\mathbf{A}_k = \begin{bmatrix} \mathbf{A}_{bb}^k & \mathbf{A}_{bs} \\ \mathbf{A}_{sb}^k & \mathbf{A}_{ss}^k \end{bmatrix},\ u_{k-1}^{pca} = \begin{Bmatrix} u_b^{k-1} \\ u_s^{k-1} \end{Bmatrix}.$$

Здесь индекс b относится к удерживаемым признакам, и индекс s – к удаляемым признакам.

Диагонализация матрицы удаляемых признаков осуществляется с помощью ортогонального преобразования

$$\mathbf{P}_s^T \mathbf{A}_{ss}^k \mathbf{P}_s = \mathbf{\Sigma}_{ss}\,. \tag{13}$$

Матрица $\mathbf{P}s$ является ортогональной матрицей и состоит из собственных векторов матрицы $\mathbf{A}_{ss}^k$, полученных при решении уравнения

$$(\mathbf{I}_s - \mu \mathbf{A}_{ss}^k) p_s = 0\,. \tag{14}$$

Таким образом,

$$u_k^{pca} = \begin{bmatrix} \mathbf{I}_b & 0 \\ 0 & \mathbf{P}_s \end{bmatrix} \begin{Bmatrix} u_b^{k-1} \\ u_s^{k-1} \end{Bmatrix} = \mathbf{P} u_{k-1}^{pca}\,. \tag{15}$$

Отметим, что $\mathbf{P}s$ является ортогональной матрицей, поэтому справедливо следующее соотношение

$$\mathbf{P}_s^T \mathbf{I}_s \mathbf{P}_s = \mathbf{P}_s^T \mathbf{P}_s = \mathbf{I}_s\,. \tag{16}$$

Подставляя (15) в уравнение (12) и умножая справа на матрицу $\mathbf{P}^T$ с учётом (16), получим

$$(\mathbf{I}_k - \mu \mathbf{A}_k^*) u_k^{pca} = 0\,, \tag{17}$$

где $\mathbf{A}_k^* = \mathbf{P}^T \mathbf{A}_k \mathbf{P}$.

С учётом блочного деления и соотношения (13) матрица $\mathbf{A}^k$ может быть представлена в виде

$$\begin{aligned} \mathbf{A}_k^* &= \begin{bmatrix} \mathbf{I}_b & 0 \\ 0 & \mathbf{P}_s \end{bmatrix}^T \begin{bmatrix} \mathbf{A}_{bb}^k & \mathbf{A}_{bs} \\ \mathbf{A}_{sb}^k & \mathbf{A}_{ss}^k \end{bmatrix} \begin{bmatrix} \mathbf{I}_r & 0 \\ 0 & \mathbf{P}_s \end{bmatrix} = \\ &= \begin{bmatrix} \mathbf{A}_{bb}^k & \mathbf{A}_{bs}^k \mathbf{P}_s \\ \mathbf{P}_s^T \mathbf{A}_{sb}^k & \mathbf{\Sigma}_{ss} \end{bmatrix}. \end{aligned} \tag{18}$$

Если диагональные коэффициенты матрицы $\sum_{ss}$ упорядочены по убыванию, то обратная величина первого диагонального коэффициента матрицы $\sum_{ss}$ будет равна наименьшему собственному значению блока удаляемых признаков ($\mu_{\min}$). Проверка условия (11) позволяет определить, можно ли удалять признаки, входящие в блок удаляемых признаков. Если удалять признаки нельзя, условие (11) проверяется для второго диагонального коэффициента. По сути это означает, что мы уменьшаем блок удаляемых признаков на единицу и принимаем решение об удалении уменьшенного блока признаков. Если условие (11) не выполняется, то мы снова уменьшаем блок удаляемых признаков на единицу, т.е. условие (11) проверяется для третьего диагонального коэффициента матрицы $\sum_{ss}$ и т.д. Если условие (11) выполняется, то это означает, что мы определили блок признаков, которые можно удалять. Если условие (11) не выполняется, то это означает, что процесс понижения порядка закончен.

Остальные шаги соответствуют шагам алгоритма многоуровневой линейной конденсации [8]. Единственное отличие, возникающее на шаге восстановления собственных векторов, связано с учётом ортогональных преобразований (15) при вычислении значений удаленных признаков.

Алгоритм блочно-ортогональной конденсации, так же как и алгоритм многоуровневой линейной конденсации, даёт приближённое решение задачи собственных значений. Дело в том, что собственные векторы уравнения (4) образуют ортогональный базис $m \times m$, в то время как собственные векторы, вычисленные алгоритмом блочно-ортогональной конденсации, составляют ортогональный базис размерностью $m \times p$. Таким образом, погрешность собственных векторов, вычисленных алгоритмом блочно-ортогональной конденсации, связана тем, что они не являются ортогональными собственным векторам, начиная с p+1 и выше. Погрешность получаемых решений можно уменьшить с помощью параметра отсечения. Увеличивая параметр отсечения, мы уменьшаем погрешность собственных векторов. Однако при этом уменьшается степень сжатия матрицы и увеличивается время решения задачи.

Погрешность собственных векторов, вычисленных алгоритмом блочно-ортогональной конденсации, предлагается оценивать по формуле

$$e_i = \frac{\sum\limits_{k=1}^{N} (\mathrm{v}_{0ik} - \mathrm{v}_{0ik}^*)^2}{\sum\limits_{k=1}^{N} \mathrm{v}_{0ik}^{\ 2}}\,,$$

где v_{0i}^* – i-й собственный вектор, вычисленный алгоритмом блочно-ортогональной конденсации, v_{0i}– i-й собственный вектор, полученный методом Хаусхолдера.

Исследование точности вычисления собственных векторов выполняется на базе данных ORL. Собственные векторы вычисляются с использованием алгоритма блочно-ортогональной конденсации и метода Хаусхолдера. Точность вычисления собственных векторов алгоритмом блочно-ортогональной конденсации, как было показано выше, зависит от параметра отсечения. В табл. 5 показано, как меняется погрешность вычисления собственных векторов в зависимости от параметра отсечения.

Таблица 5. Погрешность вычисления собственных векторов в зависимости от параметра отсечения

Параметр отсечения	Погрешность собственных векторов, %	
	80 ГК	200 ГК
1,1	75,06	82,5
1,5	20,98	26,7
2	3,12	2,56
2,5	1,02	1,32
3	0,29	0,46

Как видно из таблицы, погрешность вычисления собственных векторов тем больше, чем меньше параметр отсечения. Однако для любого значения параметра отсечения собственные векторы, вычисленные алгоритмом блочно-ортогональной конденсации, представляют набор ортогональных векторов, поэтому могут использоваться в качестве главных компонент.

Исследование эффективности использования главных компонент, вычисленных алгоритмом блочно-ортогональной конденсации, для решения задачи распознавания лиц проводится на базе данных ORL. Изображения, хранящиеся в базе данных, делятся на обучающую и тестовую выборки. Для исследования точности распознавания лиц используется процедура кросс-валидации, усредняющая коэффициенты распознавания, полученные по различным учебным выборкам. Обучающая выборка формируется из 8 изображений каждого класса базы данных ORL, которые выбираются случайно. Все оставшиеся изображения составляют тестовую выборку. Для распознавания лиц используется классификатор по ближайшему соседу, который даёт более высокие значения коэффициентов распознавания, чем классификатор по ближайшему центру классов [11].

Выполняется 10 экспериментов с главными компонентами, полученными методом Хаусхолдера, и 10 экспериментов с главными компонентами, вычисленными алгоритмом блочно-ортогональной конденсации при различных значениях параметра отсечения. Качество распознавания в каждом испытании оценивается по коэффициенту распознавания тестовой выборки. В результате обработки результатов экспериментов получаются средние значения $K_{ср}$ и среднеквадратические отклонения $K_{скв}$ коэффициентов распознавания тестовой выборки.

В табл. 6 представлены результаты, полученные с использованием главных компонент, вычисленных алгоритмом блочно-ортогональной конденсации (параметр отсечения равен 1.1, 2 и 3). Результаты, полученные методом Хаусхолдера, полностью совпадают с коэффициентами распознавания, вычисленными методом блочно-ортогональной конденсации с параметром отсечения, равным 3.

Как видно из табл. 6, несмотря на то, что погрешность вычисления собственных векторов при параметре отсечения 1,1 наибольшая, их использование в качестве главных компонент позволяет получать даже более высокую точность распознавания.

В ряде экспериментов главные компоненты, вычисленные алгоритмом блочно-ортогональной конденсации с параметром отсечения 1,1, демонстрируют более высокое качество распознавания, что говорит о перспективности такого подхода вычисления главных компонент.

Исследование быстродействия алгоритма блочно-ортогональной конденсации осуществляется путём сравнения времени вычисления главных компонент наборов изображений, хранящихся в базе данных FERET. Под временем вычисления понимается время расчёта главных компонент u_0^{pca} путём решения уравнения (3) на персональном компьютере.

Таблица 6. Средние значения и среднеквадратические отклонения коэффициентов распознавания тестовой выборки

Число главных компонент	Параметр отсечения					
	1,1		2		3	
	$K_{ср}$	$K_{скв}$	$K_{ср}$	$K_{скв}$	$K_{ср}$	$K_{скв}$
24	95,5	0,87	95,2	0,98	95,2	0,98
26	96,1	0,92	96,1	0,92	96,1	0,92
28	97,2	1,15	97,1	1,02	97,1	1,02
30	97,2	0,79	96,8	0,65	96,7	0,87
32	97,3	0,40	97,2	0,98	97,1	0,84
34	97,2	0,79	97,2	0,79	97,2	0,79
36	97,6	0,92	97,6	0,92	97,6	0,92
38	97,6	0,71	97,5	0,83	97,5	0,83
40	97,6	0,92	97,6	0,92	97,6	0,92
42	98,1	0,66	98,0	0,87	98,0	0,87
44	98,0	0,65	97,8	0,84	98,0	0,64
46	97,6	0,71	97,3	0,92	97,3	0,92
48	97,7	0,79	97,38	0,922	97,3	0,92
50	97,6	0,71	97,75	0,791	97,6	0,92

Для демонстрации быстродействия алгоритма блочно-ортогональной конденсации вычисляются 290 главных компонент матриц различных размеров, а полученные результаты сравниваются с результатами, полученными методом Хаусхолдера.

В табл. 7 показывается, как меняется время вычисления 290 главных компонент матриц различного размера. Для сравнения представлены результаты, полученные методом Хаусхолдера, алгоритмами многоуровневой линейной конденсации блочно-ортогональной конденсации.

Таблица 7. Относительное время вычисления главных компонент в зависимости от размера матриц

Порядок матрицы	Метод Хаусхолдера	Многоуровневая линейная конденсация	Блочно-ортогональная конденсация
1210	61	12	12
4080	4742	776	477
4832	9000	1283	879
5793	19103	2298	1122
6899	35411	3467	1320

Как видно из таблицы, алгоритм блочно-ортогональной конденсации превосходит по быстродействию не только метод Хаусхолдера, но и алгоритм многоуровневой линейной конденсации.

Заключение

Рассмотрены некоторые аспекты применения метода главных компонент и линейного дискриминантного анализа для решения задачи распознавания людей. Предложена технология построения классификаторов с использованием учебной выборки ненормализованных изображений, что уменьшает затраты на этапе распознавания неизвестных изображений за счёт исключения процедуры нормализации.

Проведено исследование качества распознавания ненормализованных изображений лиц в зависимости от числа образцов в классе на базах данных ORL и FERET.

Проведённые исследования показывают, что использование ненормализованных изображений лиц при построении классификаторов подходом МГК+ЛДА является перспективным направлением при построении систем распознавания лиц, так как позволяет снизить затраты распознавания новых изображений за счёт частичного или полного отказа от процедуры нормализации лиц. Хорошее качество распознавания лиц по ненормализованным изображениям может быть достигнуто при увеличении числа изображений в классе. Если число изображений в классе невелико, то учебная выборка может быть дополнена изображениями, полученными путём их поворота, масштабирования и зеркалирования.

Также рассмотрена задача повышения эффективности вычисления главных компонент больших наборов изображений. В этом случае для вычисления главных компонент предлагается использовать алгоритм блочно-ортогональной конденсации. Показано, что данный алгоритм позволяет существенно уменьшить трудоёмкость вычисления главных компонент.

Литература

1. **Kirby, M.** Application of the KL procedure for the characterization of human faces / M. Kirby, L. Sirovich // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 1990. – Vol. 12(1). – P. 103-108.
2. **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(8). – P. 1724-1733.
3. **Lu, J.** Face Recognition Using LDA-Based Algorithms / J. Lu, K.N. Plataniotis, A.N. Venelsanopoulos // IEEE Trans, on Neural Networks. – 2003. – Vol. 14(1). – P. 195-200.
4. **Belhumeur, P.N.** Eigenfaces vs Fisherfaces: recognition using class specific linear projection / P.N. Belhumeur, J.P. Hespanha, D.J. Kriegman // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 1997. – Vol. 19. – P. 711-720.
5. **Martinez, A.M.** PCA versus LDA / A.M. Marline, A.C. Kak // IEEE Trans, on Pattern Analysis and Machine Intelligence. – 2001. – Vol. 23(2). – P. 228-233.
6. **Кухарев, Г.А.** Алгоритмы двумерного анализа главных компонент для задач распознавания изображений лиц / Г.А. Кухарев, И.Л. Щёголева // Компьютерная оптика. – 2010. – Т. 34, № 4. – С. 545-551.
7. **Мокеев, А.В.** О точности и быстродействии метода синтеза главных компонент / А.В. Мокеев // Бизнес-информатика. – 2010. – № 3(13). – С. 65-68.
8. **Мокеев, В.В.** О повышение эффективности вычислений главных компонент в задачах анализа изображений / В.В. Мокеев // Цифровая обработка сигналов. – 2011. – № 4. – С. 29-36.
9. **Гриненко, Н.И.** О задачах исследований колебаний конструкций методом конечных элементов / Н.И. Гриненко, В.В. Мокеев // Прикладная механика. – 1985. – Т. 21, № 3. – С. 25-30.
10. **Мокеев, В.В.** О задаче нахождения собственных значений и векторов больших матричных систем / В.В. Мокеев // Журнал вычислительной математики и математической физики. – 1992. – № 32(10). – С. 1652-1657.
11. **Мокеев, В.В.** О решении проблемы выборки малого размера при использовании линейного дискриминантного анализа в задачах распознавания лиц / В.В. Мокеев, С.В. Томилов // Бизнес информатика. – 2013. – № 1. – С. 37-43.
12. **Phillips, P.J.** The FERET evaluation methodology for face recognition algorithms / P.J. Phillips, H. Moon, P.J. Rauss, S. Rizvi // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2000. – Vol. 22(10). – P. 1090-1104.
13. **Gu, X.** Regularized locality preserving discriminant analysis for face recognition / X. Gu, W. Gong, L. Yang // Neurocomputing. – 2011. – Vol. 74. – P. 3036-3042
14. **Zhang, T.** Patch alignment for dimensionality reduction /T. Zhang, D. Tao, X. Li [et al.] // IEEE Transaction on Knowledge and Data Engineering. – 2009. – Vol. 21(9). – P. 1299-1313.
15. **Yang, L.** Null space discriminative locality preserving projections for space recognition / L. Yang, T. Zhang, G.X. Gu [et al.] // Neurocomputing. – 2008. – Vol. 71. – P. 2045-2054.
16. **Pang, Y.H.** Regularized locality preserving discriminant embedding for face recognition / Y.H. Pang, A.B.J. Teoh, F.S. Abas // Neurocomputing. – 2012. – Vol. 77. – P. 156-166
17. **Li, S.Z.** Non linear mapping from multi-view face patterns to a gaussian distribution in a low dimensional space / S.Z. Li, X. Rong, L.Z. Yu, Zhang HongJiang // in: Proceedings of the IEEE ICCV Workshop on Recognition, Analysis, and Tracking of Face sand Gestures in Real-Time Systems. – 2001. – P. 47–54.
18. **Yan, S.C.** Graph embedding and extensions: a general framework for dimensionality reduction / S.C. Yan, Dong Xu, Benyu Zhang, Hong-Jiang Zhang, Qiang Yang, S. Lin // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2007. – Vol. 29(1). – P. 40-51.
19. **Cai, D.** Locality sensitive discriminant analysis / D. Cai, X. He, K. Zhou, J. Han, H. Bao // in: Proceedings of the International Joint Conference Artificial Intelligence, 2007. – P. 708-713.
20. **Parlett, B.N.** The symmetric eigenvalue problem / B.N. Parlett. – New Jersey, Englewood Cliffs: Prentice-Hall, 1980.

References

1. **Kirby, M.** Application of the KL procedure for the characterization of human faces / M. Kirby, L. Sirovich // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 1990. – Vol. 12(1). – P. 103-108.
2. **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(8). – P. 1724-1733.
3. **Lu, J.** Face Recognition Using LDA-Based Algorithms / J. Lu, K.N. Plataniotis, A.N. Venelsanopoulos // IEEE Trans, on Neural Networks. 2003. – Vol. 14(1). – P. 195-200.
4. **Belhumeur, P.N.** Eigenfaces vs Fisherfaces: recognition using class specific linear projection / P.N. Belhumeur, J.P. Hespanha, D.J. Kriegman // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 1997. – Vol. 19. – P. 711-720.
5. **Martinez, A.M.** PCA versus LDA / A.M. Marline, A.C. Kak // IEEE Trans, on Pattern Analysis and Machine Intelligence. – 2001. – Vol. 23(2). – P. 228-233.
6. **Kukharev, G.A.** Algorithms of two-dimensional principal component analysis for face recognition / G.A. Kukharev, N.L. Shchegoleva // Computer Optics. – 2010. – Vol. 34(4). – P. 545-551. – (In Russian).
7. **Mokeev, A.V.** On accuracy and performance principal component synthesis method / A.V. Mokeev // Business Informatics. – 2010. – Vol. 3(18). – P. 64-68. – (In Russian).
8. **Mokeyev, V.V.** On effectiveness increase of principal components computation in image analysis problem / V.V. Mokeyev // Digital Signal Processing. – 2011. – Vol. 4. – P. 29-36. – (In Russian).
9. **Grinenko, N.I.** Problems of studying vibrations of structures by the finite-element method / N.I. Grinenko, V.V. Mokeev // International Applied Mechanics. – 1985. – Vol. 21(3). – P. 25-30. – (In Russian).
10. **Mokeev, V.V.** On the problem of finding the eigenvalues and eigenvectors of large matrix system, arising in use a finite element method / V.V. Mokeev // Computational Mathematics and Mathematical Physics. – 1992. – Vol. 32(10). – P. 1652-1657. – (In Russian).
11. **Mokeyev, V.V.** On solution of small sample size problem with linear discriminant analysis in face recognition / V.V. Mokeyev, S.V. Tomilov // Business Informatics. – 2013. – Vol. 1(18). – P. 37-43. – (In Russian).

12. **Phillips, P.J.** The FERET evaluation methodology for face recognition algorithms / P.J. Phillips, H. Moon, P.J. Rauss, S. Rizvi // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2000. – Vol. 22(10). – P. 1090-1104.
13. **Gu, X.** Regularized locality preserving discriminant analysis for face recognition / X. Gu, W. Gong, L. Yang // Neurocomputing. – 2011. – Vol. 74. – P. 3036-3042
14. **Zhang, T.** Patch alignment for dimensionality reduction / T. Zhang, D. Tao, X. Li [et al.] // IEEE Transaction on Knowledge and Data Engineering. – 2009. – Vol. 21(9). – P. 1299-1313.
15. **Yang, L.** Null space discriminative locality preserving projections for space recognition / T. Zhang, G.X. Gu [et al.] // Neurocomputing. – 2008. – Vol. 71. – P. 2045-2054.
16. **Pang, Y.H.** Regularized locality preserving discriminant embedding for face recognition / Y.H. Pang, A.B.J. Teoh, F.S. Abas // Neurocomputing. – 2012. – Vol. 77. – P. 156-166.
17. **Li, S.Z.** Non linear mapping from multi-view face patterns to a gaussian distribution in a low dimensional space / S.Z. Li, X. Rong, L.Z. Yu, Zhang HongJiang // in: Proceedings of the IEEE ICCV Workshop on Recognition, Analysis, and Tracking of Face sand Gestures in Real-Time Systems, 2001. – P. 47-54.
18. **Yan, S.C.** Graph embedding and extensions: a general framework for dimensionality reduction / S.C. Yan, Dong Xu, Benyu Zhang, Hong-Jiang Zhang, Qiang Yang, S. Lin // IEEE Transactions on Pattern Analysis and Machine Intelligence – 2007. – Vol. 29(1). – P. 40-51.
19. **Cai, D.** Locality sensitive discriminant analysis / D. Cai, X. He, K. Zhou, J. Han, H. Bao // in: Proceedings of the International Joint Conference Artificial Intelligence, 2007. – P. 708-713.
20. **Parlett, B.N.** The symmetric eigenvalue problem / B.N. Parlett. – New Jersey, Englewood Cliffs: Prentice-Hall, 1980.

ON THE SOLUTION OF THE IMAGE RECOGNITION PROBLEM BY A PRINCIPAL COMPONENT METHOD AND LINEAR DISCRIMINANT ANALYSIS

V.V. Mokeyev , S.V. Tomilov
South Ural State University (National Research University)

Abstract

In the paper, some aspects of image analysis based on PCA (Principal Component Analysis) and LDA (Linear Discriminant Analysis) are considered. The main idea of this approach is that, firstly, we project the face image from the original vector space to a face subspace via PCA, secondly, we use LDA to obtain a linear classifier. In the paper, the efficiency of application of PCA and LDA to a problem of recognition of face images without their preliminary normalization is investigated. When the number of images in a class is not large, it is proposed that the training set is supplemented by images obtained by rotating, scaling and mirroring. In the images from the ORL and Feret databases, the influence of the training set expansion on the quality of recognition of unnormalized face images is studied. Also, a problem of increasing the efficiency of principal component calculation for large image samples is addressed. A linear condensation method is used as a new technique to calculate the principal components of a large matrix. The accuracy and performance of the developed algorithm are evaluated.

Key words: face recognition, principal component analysis, linear discriminant analysis, linear condensation method.

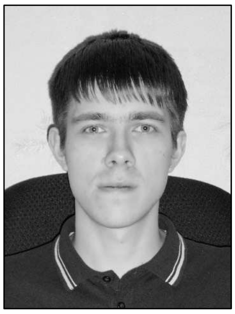
Сведения об авторах

Мокеев Владимир Викторович 1952 года рождения. В 1988 году защитил кандидатскую диссертацию в Челябинском политехническом институте, докторскую – в 1999 году в Южно-Уральском государственном университете.

В настоящее время является заведующим кафедрой информационных систем Южно-Уральского государственного университета (национальный исследовательский университет). Является автором более 100 научных работ по разработке и применению численных методов. Последние исследования связаны с анализом и распознаванием изображений.

E-mail: *mokeyev@mail.ru* .

Vladimir Victorovich Mokeyev (born in 1952) defended PhD thesis in Chelyabinsk Polytechnic Institute in 1988, Doctoral thesis –In South Ural State University in 1999. At present he is a head of Information Systems department at South Ural State University (National Research University). He is an author and co-author of more than 100 scientific papers on development and application of numerical method in different areas of knowledge. Recent researches are related to areas of computer vision, including face detection and face recognition.



Томилов Станислав Владимирович 1988 года рождения. В 2010 году окончил Южно-Уральский институт управления и экономики. Аспирант кафедры информационных систем Южно-Уральского государственного университета (национальный исследовательский университет). Область научных интересов: цифровая обработка и распознавание изображений.

E-mail: *tomilov_stas@mail.ru* .

Stanislav Vladimirovich Tomilov (born in 1988) graduated from South Ural Institute of Management and Economics in 2010. He is a postgraduate student of Information Systems department of South Ural State University (National Research University). His research interests are digital image processing and image recognition.

Поступила в редакцию 13 апреля 2013г.