

О РЕШЕНИИ ПРОБЛЕМЫ ВЫБОРКИ МАЛОГО РАЗМЕРА ПРИ ИСПОЛЬЗОВАНИИ ЛИНЕЙНОГО ДИСКРИМИНАНТНОГО АНАЛИЗА В ЗАДАЧАХ РАСПОЗНАВАНИЯ ЛИЦ

В.В. Мокеев,

доктор технических наук, старший научный сотрудник,
заведующий кафедрой информационных систем Южно-Уральского
государственного университета (ЮУрГУ)

С.В. Томилов,

аспирант кафедры информационных систем
Южно-Уральского государственного университета (ЮУрГУ)

E-mail: mokeyev@mail.ru, tomilov_stas@mail.ru

Адрес: г. Челябинск, пр. Ленина, д. 76

Рассматриваются эффективность линейного дискриминантного анализа в задачах распознавания лиц для случая малого размера выборки. Предлагается для вычисления дискриминантных компонент использовать обобщенный метод Якоби, что позволяет не терять дискриминантную информацию. Эффективность предлагаемого подхода демонстрируется в ходе экспериментальных исследований на базе данных ORL.

Ключевые слова: распознавание лиц, выборка малого размера, линейный дискриминантный анализ, база ORL.

1. Введение

Большое внимание в настоящее время уделяется проблемам обработки изображений. Это обусловлено многообразием прикладных задач, в которых используется либо сами изображения, либо результат их обработки. Наиболее актуальной задачей в данной области является распознавание фотоизображений лиц с целью иден-

тификации личности человека. Одной из проблем с которой приходится сталкиваться, решая данную задачу, является отсутствие достаточного количества фотоизображений одного человека, которое бы хорошо описывало всю индивидуальную изменчивость данного субъекта.

Среди различных методов распознавания лиц широкое распространение сегодня нашли мето-

ды, базирующиеся на линейном дискриминантном анализе. Линейный дискриминантный анализ (LDA – Linear Discriminant Analysis) [1] заключается в выборе проекции пространства изображений на пространство признаков таким образом, чтобы было минимизировано внутриклассовое и максимизировано межклассовое расстояние в пространстве признаков.

Несмотря на то, что линейный дискриминантный анализ является полезным инструментом для классификации образов, алгоритмы на основе LDA, страдают от проблемы «малого размера выборки» (SSS – Small Sample Size) [2], которая существует в задачах, где изображения обладают высокой размерностью, в то время как количество изображений, описывающих одну персону, является небольшим. В случае малого размера выборки матрица внутриклассовых различий становится сингулярной. Это значит, что собственные значения матрицы становятся нулевыми. Во многих работах используется понятие нуль-пространства матрицы. Обычно под этим термином понимается пространство, образованное собственными векторами, которым соответствуют нулевые собственные значения. Матрица межклассовых различий также становится сингулярной, в случае если размерность изображений больше числа классов. Так как в традиционном LDA вычисление дискриминантных компонент сопровождается обращением либо матрицы внутриклассовых либо матрицы межклассовых различий, то сингулярность матриц ведет к противоречию, которое и называется проблемой SSS. Чтобы решить данную проблему были разработаны различные разновидности алгоритмов LDA.

Среди различных путей решения этой проблемы, наиболее успешными являются подходы, сочетающие линейный дискриминантный анализ с методом главных компонент [3,4]. Хотя эффективность такого подхода при распознавании лиц очевидна, а теоретические его основы были заложены в [5,6], уже в работе [4] мы находим, что применение PCA не может гарантировать успешное применение LDA, т.е. после преобразования ковариационная матрица внутриклассовых различий может все еще быть сингулярной.

В связи с этим появилось достаточно большое число работ, в которых исследователи ищут более эффективные дискриминантные подпространства [5-9]. В работах [7, 8] были разработаны методы регуляризации, которые устраняют

сингулярность матрицы путем добавления к ней скалярной матрицы. Однако такой подход дает ограниченный эффект и неэффективен в случае, когда вырождены сразу две матрицы. Поиск эффективных дискриминантных пространств показал, что существует ключевая дискриминантная информация в нуль-пространстве матрицы внутриклассовых различий. Такой тип дискриминантной информации называется нерегулярной дискриминантной информацией, в отличие от регулярной дискриминантной информации находящейся вне пределов нуль-пространства. К сожалению, многие из вышеперечисленных методов, используя подход «PCA+LDA» теряют дискриминантную информацию, которая содержится в нуль-пространстве матрицы внутриклассовых различий, при том, что эта дискриминантная информация является очень важной для решения проблемы SSS.

В работе [5] описывается прямой линейный дискриминантный анализ (DLDA – Direct Linear Discriminant Analysis), который удаляет нуль-пространство матрицы межклассовых различий, что позволяет устранить сингулярность этой матрицы. В работе [9] предлагается подход NLDA (null space LDA), который базируется на том, что нуль-пространство матрицы внутриклассовых различий содержит ценную дискриминантную информацию. Основная идея DLDA заключается в том, чтобы исключить нуль-пространство матрицы межклассовых различий, которое не содержит полезной информации и сохранить нуль-пространство матрицы внутриклассовых различий, которое содержит важную дискриминантную информацию. Это может быть достигнуто путем диагонализации сначала матрицы межклассовых различий, и только затем уже матрицы внутриклассовых различий. Процедура NLDA выполняет эти действия в обратном порядке. Общей чертой подходов NLDA и DLDA является то, что все они теряют некоторую дискриминантную информацию.

В статье предлагается алгоритм обобщенного линейного дискриминантного анализа, в котором вычисление дискриминантных компонент осуществляется с помощью обобщенного метода Якоби. Обобщенный метод Якоби использует вращения для одновременной диагонализации матрицы межклассовых и внутриклассовых различий, что позволяет не терять дискриминантную информацию.

2. Обобщенный линейный дискриминантный анализ

Пусть имеется набор изображений, каждое из которых описывается вектором x_i^k , где i – номер изображений ($i = 1, 2, 3, \dots, M_k$), k – номер класса ($k = 1, 2, \dots, K$). Размерность вектора x_i^k равняется числу пикселей образа (N). Таким образом, все изображения можно представить в виде матрицы X , столбцами которой являются векторы x_i . Размерность пространства признаков определяется произведением $N \times M$. Пространство признаков центрируется относительно среднелинейного вектора

$$m = \frac{1}{M} \sum_k \sum_i x_i^k.$$

$$\text{Здесь } M = \sum_k M_k -$$

суммарное число изображений во всех классах. При этом формируется матрица отцентрированных изображений X^0 , строками которого являются векторы

$$\bar{x}_i^k = x_i^k - m.$$

При распознавании изображений целью линейного дискриминантного анализа является поиск проекционной матрицы, которая максимизирует так называемый критерий Фишера. Прежде чем описать этот критерий, необходимо определить две ковариационные матрицы: межклассовых и внутриклассовых различий. Матрица межклассовых различий определяется по формуле:

$$A_b = \frac{1}{M} \sum_k M_k (m_k - m)(m_k - m)^T, \quad (1)$$

где m_k – среднелинейный вектор изображений k -го класса, определяется с помощью выражения

$$m_k = \frac{1}{M_k} \sum_{i=1}^{M_k} x_i^k.$$

Матрица внутриклассовых различий определяется выражением

$$A_\omega = \frac{1}{M} \sum_{k=1}^K \sum_{i=1}^{M_k} (x_i^k - m_k)(x_i^k - m_k)^T. \quad (2)$$

Проекционная матрица V_{lda} выбирается как матрица с ортонормальными столбцами, максимизирующими критерий Фишера

$$V_{lda} = \arg \max_{V \in R^{n \times r}} \frac{|V^T A_b V|}{|V^T A_\omega V|} \quad (3)$$

Матрица дискриминантных компонент V_{lda} определяется из решения обобщенной задачи собственных значений

$$(A_b - \lambda A_\omega) v_0^{lda} = 0. \quad (4)$$

Если существует обратная матрица к матрице A_ω , то решение уравнения (4) сводится к стандартной задаче собственных значений. Вектор новых факторов получается с помощью соотношения

$$y_i^k = V_{lda}^T x_i^k. \quad (5)$$

Однако, в задачах распознавания лиц, из-за большой размерности исходных данных, LDA страдает от двух типов трудностей: сингулярности и высокой размерности матриц внутриклассовых и межклассовых различий.

Метод главных компонент в рамках подхода «PCA + LDA» применяется для уменьшения размерности пространства признаков от N до r ($r \ll N$). В результате формируется матрица главных компонент V_{pca} , которая имеет размерность $N \times r$. В работах [10-12] описаны различные алгоритмы вычисления главных компонент. Далее следует применение линейного дискриминантного анализа для поиска представления свойств объектов в пространстве признаков меньшей размерности. В результате вычисляется матрица дискриминантных компонент V_{lda} . Таким образом, матрица преобразования W_{lda} в рамках подхода «PCA + LDA» определяется следующим образом

$$W_{lda} = \arg \max_w \frac{|W^T V_{pca}^T A_b V_{pca} W|}{|W^T V_{pca}^T A_\omega V_{pca} W|} = \arg \max_w \frac{|W^T A_b' W|}{|W^T A_\omega' W|}. \quad (6)$$

Столбцами матрицы W_{lda} являются собственные векторы w_{0i}^{lda} , которые получаются в результате решения уравнения

$$(A_b' - \lambda A_\omega') w_0^{lda} = 0. \quad (7)$$

Здесь $A_b' = V_{pca}^T A_b V_{pca}$, $A_\omega' = V_{pca}^T A_\omega V_{pca}$. Задача (7) является обобщенной задачей собственных значений. Традиционно уравнение (7) сводится к стандартной задаче собственных значений путем обращения одной из матриц, например, матрицы A_ω' . Естественно, это возможно только в случае если обращаемая матрица не вырождена. Уравнение (7) при этом преобразуется к виду

$$(A_\omega')^{-1} A_b' w_0^{lda} = \lambda I w_0^{lda}. \quad (8)$$

Решение (8) позволяет определить матрицу дискриминантных компонент, столбцами которой являются собственные векторы уравнения (8) с наибольшими собственными значениями. Количество дискриминантных компонент $p < r$.

Для распознавания изображений может использоваться либо классификатор по ближайшему центру класса, либо классификатор по ближайшему соседу [15].

3. Обобщенный метод Якоби

Для решения уравнения (7) предлагается использовать обобщенный метод Якоби, который не требует обращений матриц и не теряет дискриминантную информацию при вычислении дискриминантных компонент.

Обобщенный метод Якоби реализуется в виде многошагового процесса, на каждом шаге которого обнуляются внедиагональные элементы. Подробное описание алгоритма можно найти в литературе по проблеме собственных значений.

Применительно к задаче распознавания изображений в алгоритм обобщенного метода Якоби вычисления собственных значений внесены небольшие дополнения.

1. Если в процесс вращений диагональные коэффициенты матриц принимают значение близкие к машинному нулю, то величина коэффициентов устанавливается на несколько порядков выше машинного нуля. Например, для расчетов с двойной точностью эта величина находится в пределах $10^{-12} \dots 10^{-14}$.

2. Если в процессе диагонализации невозможно найти коэффициенты обнуляющее внедиагональные коэффициенты сразу двух матриц, процесс диагонализации не прерывается. При этом выполняется попытка обнулить внедиагональные коэффициенты только одной матрицы, Если и это не возможно, то осуществляется переход к следующему шагу.

4. Экспериментальные исследования

Экспериментальное исследование эффективности описанного выше обобщенного линейного дискриминантного анализа проводится с использованием базы изображений лиц «ORL» [16]. База ORL содержит 400 фронтальных изображений 40 человек, с различной мимикой, поворотами и на-

клонами головы. Все изображения в базе полутоновые, с 256 градациями яркости. Размер каждого изображения – 92 x 112 пикселей. *Рис. 1* показывает изображения нескольких субъектов базы данных ORL.

Исходный набор изображений делится на обучающую и тестовую выборки. Изображения обучающей выборки формируются из L изображений каждого класса. Все оставшиеся изображения составляют тестовую выборку.

В первой серии экспериментов сравниваются два классификатора: классификатор по ближайшему центру класса и классификатор по ближайшему соседу. Для исследования качества распознавания используется процедура кросс-валидации, усредняющая коэффициенты распознавания, полученных при различном делении набора изображений на учебную и тестовую выборки. Эксперименты проводятся для учебных выборок, содержащих 2, 3 и 4 изображений в каждом классе базы ORL, которые выбираются случайно. Все оставшиеся изображения составляют тестовую выборку. Таким образом, обеспечивается не пересечение учебной и тестовой выборок. Учебные выборки при этом содержат 80 (2x40), 120 (3x40) и 160 (4x40) изображений, а тестовые — 320, 280 и 140 изображений, соответственно.

В ходе эксперимента главные компоненты вычисляются по обучающей выборке и далее используются для сокращения размерности обучающих изображений. Редуцированные изображения используются для формирования межклассовых и внутриклассовых матриц и вычисления дискриминантных компонент. Количество дискриминантных компонент определяется как наименьшее число компонент, обеспечивающее максимальную точность распознавания. Варьируя число главных компонент, мы получаем различную точность распознавания. В ходе экспериментов для каждого числа изображений в классе ($L=2, 3, 4$, где L — число изображений в классе) было проведено 15 экспериментов, для каждого из которых учебная выборка формируется путем случайного выбора изображений из базы ORL. Оставшиеся изображения используются для формирования тестовой выборки. Результаты, полученные в ходе экспериментов, обрабатываются и представляются в виде усредненного коэффициента распознавания и среднеквадратического отклонения.

При проведении эксперимента дважды решается задача собственных значений. Первый раз при вычислении главных компонент решается стандартная задача собственных значений. Порядок матриц для различных обучающих выборок варьируется от 80 до 160. Так как порядок матриц небольшой и матрицы хорошо обусловлены, поэтому для вычисления главных компонент можно использоваться метод Хаусхолдера. Второй раз задача собственных значений возникает при вычислении дискриминантных компонент и представляет обобщенную задачу собственных значений. Порядок матриц равен числу главных компонент и не превышает 70. Когда число главных компонент меньше числа классов, матрицы внутриклассовых различий и межклассовых различий хорошо обусловлены. В этом случае все собственные значения уравнения (8) положительные. В случае, когда число главных компонент больше числа классов матрицы внутриклассовых и межклассовых различий становятся полуопределенными. В этом случае с решением уравнения (7) справляется обобщенный метод Якоби.

На рис. 2 — 4 представлены усредненные коэффициенты распознавания (k_{test}) тестовой выборки в зависимости от числа главных компонент (r) для трех значений L (2, 3, 4). Коэффициент распозна-

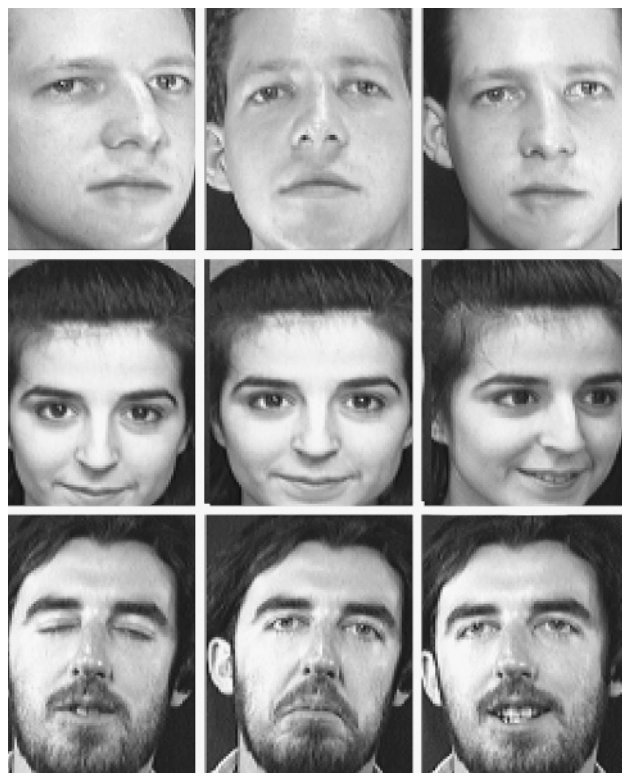


Рис. 1. Примеры лиц базы данных ORL

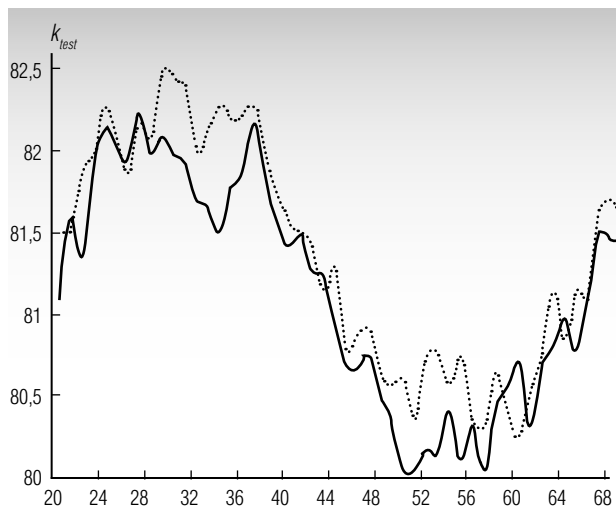


Рис. 2. Точность распознавания тестового набора ($L = 2$)

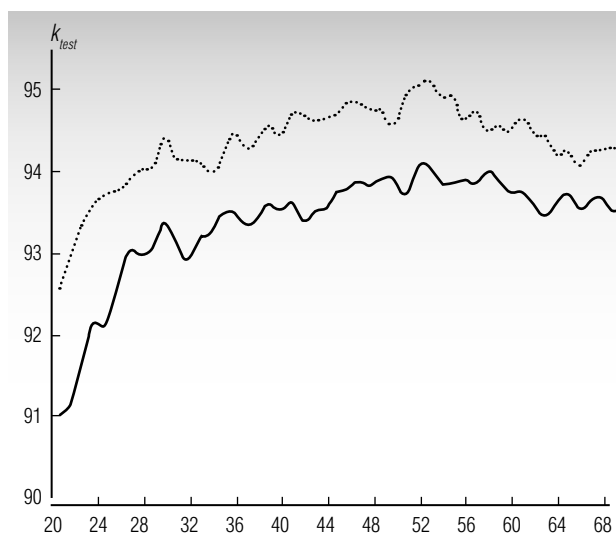


Рис. 3. Точность распознавания тестового набора ($L = 3$)

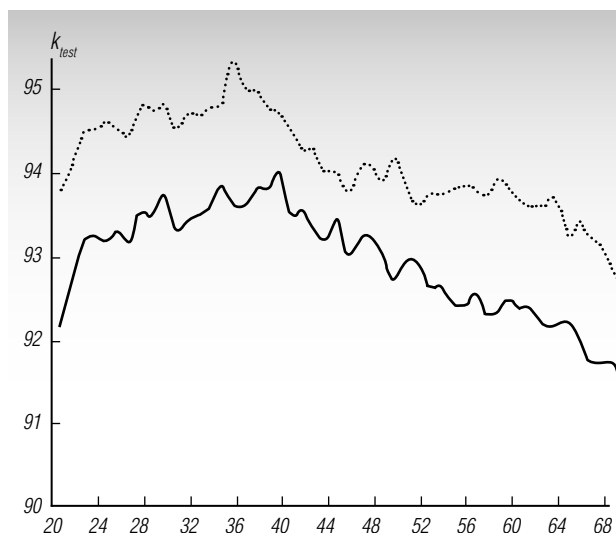


Рис. 4. Точность распознавания тестового набора ($L = 4$)

вания рассчитывается как отношение количества правильно распознанных тестовых изображений к общему числу тестовых изображений. Сплошной линией показаны результаты, полученные при использовании классификатора по ближайшему центру класса, а штриховой линией – результаты, полученные с помощью классификатора по ближайшему соседу.

Анализ полученных результатов показывает, что классификатор по ближайшему соседу дает более высокую точность распознавания.

Качество представленного алгоритма оценивается путем сравнения его с другими алгоритмами, такими как PCA, LDA, NLDA 2DPCA, RLPDA (Regularized Locality Preserving Discriminant Analysis), NDLPP (Null space Discriminant Locality Preserving Projections).

PCA, LDA, NLDA методы были описаны выше. 2DPCA представляет двумерный анализ главных компонент, который позволяет снизить трудоемкость обработки большого числа изображений, и показывает хорошие результаты для SSS проблемы. Описание его можно найти в работе [12]. Методы RLPDA, NDLPP описываются в работах [13, 14] и демонстрируют высокую эффективность при решении SSS проблемы.

Результаты распознавания изображений лиц из базы данных ORL с помощью алгоритма 2DPCA взяты из работы [12], а для алгоритмов PCA, LDA, RLPDA, NDLPP, NLDA – из работы [14]. В табл. 1 представлены результаты сравнения точности распознавания различных методов, где L – это количество образцов каждого класса в учебной выборке, N_{teach} – количество изображений лиц в учебной выборке и N_{test} – количество изображений лиц в тестовой выборке.

Таблица 1.

Результаты сравнения точности распознавания различных алгоритмов

| Алгоритм | Вариант ($L/N_{teach}/N_{test}$) | | |
|----------------|------------------------------------|------------------|------------------|
| | 2/80/320 | 3/120/280 | 4/160/240 |
| PCA [14] | 69,5±2,38 | 78,6±2,59 | 83,58±1,77 |
| LDA [14] | 80,1±2,44 | 87,9±2,23 | 91,5±1,85 |
| 2DPCA [12] | 82,9±3,42 | 90,2±1,32 | 92,2±1,53 |
| RLPDA [14] | 80,7±2,96 | 90,4±2,59 | 94,8±1,79 |
| NDLPP [14] | 83±2,31 | 91,3±1,57 | 94,7±1,64 |
| NLDA [14] | 81,8±2,61 | 91±2,11 | 94,4±1,21 |
| Обобщенный LDA | 82,5±1,55 | 91,4±2,59 | 95,2±1,06 |

Из табл. 1 видно, что с увеличением количества лиц в классе обучающего набора возрастает точность распознавания для всех алгоритмов. Обобщенный линейный дискриминантный анализ демонстрирует точность распознавания, которая в случае двух изображений в каждом классе обучающей выборки немного уступает алгоритмам 2DPCA и NDLPP, при этом в остальных случаях не уступает другим методам.

5. Заключение

В работе показано применение линейного дискриминантного анализа, позволяющему не терять дискриминантную информацию благодаря использованию обобщенного метода Якоби для вычисления дискриминантных компонент. Экспериментальные результаты на базе данных ORL показывают, что предлагаемый вариант LDA достигает более высокого качества распознавания, чем классический вариант LDA и не уступает другим методам распознавания, что говорит о его высокой эффективности. ■

Литература

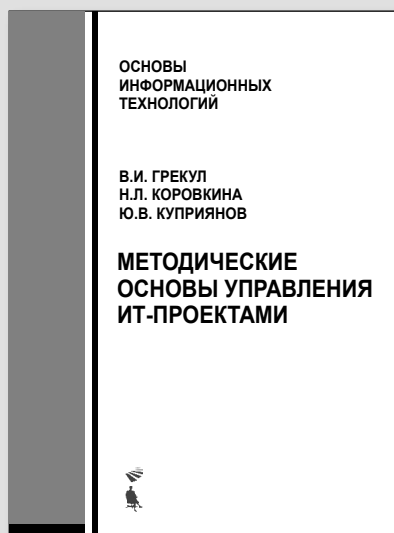
1. Etemad K. Chellappa R. Discriminant Analysis for Recognition of Human Face Images // Journal of the Optical Society of America. – 1997. – Vol. 14. – No. 8. – P. 1724–1733.
2. Raudys S.J., Jain A.K. Small sample size effects in statistical pattern recognition: recommendations for practitioners // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 1991. – 13. – P. 252–264.
3. Martinez A.M., Как А.С. PCA versus LDA // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2001. – 23 (2). – P. 228–233.
4. Belhumeur P.N., Hespanha J.P., Kriegman D.J. Eigenfaces vs Fisherfaces: recognition using class specific linear projection // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 1997. – 19. – P. 711–720.
5. Yu H., Yang J. A direct LDA algorithm for high-dimensional data with application to face recognition // Pattern Recognition. – 2001. – 34 (10). – P. 2067–2070.
6. Yang, J., Yang J.Y. Why can LDA be performed in PCA transformed space? // Pattern Recognition. – 2003. – 36 (2). – P. 563–566.

7. Dai D.Q., Yuen P.C. Regularized discriminant analysis and its applications to face recognition // Pattern Recognition. – 2003. – 36 (3). – P. 845–847.
8. Pima I., Aladjem M. Regularized discriminant analysis for face recognition // Pattern Recognition. – 2004. – 37 (9). – P. 1945–1948.
9. Chen L.F., HLiao.Y.M., Lin J.C., Kao M.D, Yu G.J. A new LDA based face recognition system which can solve the small sample size problem // Pattern Recognition. – 2000. – 33 (10). – P. 1713–1726.
10. Мокеев В.В. О повышении эффективности вычислений главных компонент в задачах анализа изображений // Цифровая обработка сигналов. – 2011. – №4. – С. 29-36.
11. Мокеев А.В. О точности и быстродействии метода синтеза главных компонент // Бизнес-информатика. – 2010. – № 3 (13). – С. 65-68.
12. Щеголева Н.Л., Кухарев Г.А. Применение алгоритмов двумерного анализа главных компонент для задач распознавания изображений лиц // Бизнес информатика. – 2011. – №4 (18). – С. 31-38
13. Yang L., Gong Wj., Gu X., et al. Null space discriminant locality preserving projections for face recognition // Neurocomputing. – 2008. – 71 (16). – P. 3644-3649.
14. Gu X., Gong W., Yang L. Regularized locality preserving discriminant analysis for face recognition // Neurocomputing. – 2011. – 74 (17). – P. 3036-3042.
15. Форсайт Дж., Понс Ж. Компьютерное зрение современный подход. – М.: Изд. дом «Вильямс», 2004.
16. The ORL Database of Faces – база изображений лиц, подготовленная в научно-исследовательской лаборатории компании Olivetti (Olivetti Research Laboratory), содержит 400 фронтальных изображений лиц 40 человек.

**МЕТОДИЧЕСКИЕ ОСНОВЫ УПРАВЛЕНИЯ
ИТ-ПРОЕКТАМИ**
Учебник

В.И. Грекул, Н.Л. Коровкина, Ю.В. Куприянов

*Интернет-университет информационных технологий: Бином.
Лаборатория знаний, 2011.*



При создании ИТ-решений перед всеми сторонами, вовлеченными в жизненный цикл проекта, возникает целый ряд вопросов, связанных с определением и детальным структурированием необходимых работ, с распределением прав и обязанностей, с управлением и контролем за исполняемыми работами. Одним из действенных инструментов для решения данных вопросов является использование унифицированных подходов, закрепленных в современных международных и российских стандартах и методологиях управления проектами. Представленный учебник содержит детальное описание процедур управления проектами внедрения информационных технологий. Отличительной особенностью данной книги является изложение материала с привязкой к этапам жизненного цикла создаваемого продукта, а не к фазам некоторого абстрактного проекта. Это позволяет читателю сформировать целостное представление о необходимых в ИТ-проекте управленческих процедурах, а также использовать материал последовательно во времени, по мере перехода от одного этапа технологического цикла создания продукта к другому.