

ПОСТРОЕНИЕ КЛАССИФИКАТОРОВ ДЛЯ РАСПОЗНАВАНИЯ ЛИЦ НА ОСНОВЕ ПОКАЗАТЕЛЕЙ СОПРЯЖЕННОСТИ

Н.Е. Козин, В.А. Фурсов

Институт систем обработки изображений РАН

Самарский государственный аэрокосмический университет имени академика С.П. Королева

Аннотация

Одной из наиболее широко используемых техник распознавания изображений лиц является метод главных компонент (Principal Component Analysis – PCA), также иногда называемый методом собственных лиц (Eigenfaces). Идея метода заключается в разложении векторов изображений по системе собственных векторов, соответствующих наибольшим собственным значениям. В работе рассматривается использование в качестве меры близости различных показателей сопряженности с подпространством, натянутым на обучающие векторы из распознаваемого класса. Показана эффективность использования данного критерия при наличии малого числа обучающих примеров. Приведены результаты экспериментов для стандартной ORL-базы данных лиц.

Введение

Задача распознавания лиц в последние годы стала одной из самых популярных [1]. Объясняется это, с одной стороны, возрастанием практического интереса к созданию распределенных систем видеоконтроля и видеонаблюдения. С другой стороны, интерес иницируется сложностью задачи, что делает ее прекрасным полигоном для проверки новых подходов и идей.

Наиболее широко для распознавания лиц в настоящее время применяются: метод главных компонент и метод сравнения образца и претендента [3], заключающийся в вычислении коэффициента корреляции. Обычно принятие решения осуществляется по одной реализации. В этом случае коэффициент корреляции имеет смысл сопряженности векторов, соответствующих сравниваемым изображениям.

В работах [4, 5] рассматривалось применение трех типов показателей сопряженности для разбиения обучающей выборки на классы. Там же приводился пример, иллюстрирующий эффективность подхода для решения задачи распознавания графических символов на изображениях.

В системах распознавания лиц размерности описывающих их векторов и соотношение между этой размерностью и числом обучающих образцов могут изменяться в широких пределах. При этом оказывается, что с вычислительной точки зрения в разных ситуациях целесообразно использовать разные показатели сопряженности.

В настоящей работе более подробно, чем прежде [4, 5], обсуждается связь между различными показателями сопряженности. Даются рекомендации по их применению при различных размерностях векторов признаков и обучающих выборок. Приводится пример обучения и классификации лиц из базы ORL.

1. Определение и связь показателей сопряженности

В работах [4, 5] рассматривались следующие три типа показателей сопряженности.

1. Показатель парной сопряженности:

$$r_{i,k}^2 = \frac{(\mathbf{x}_i^T \mathbf{x}_k)^2}{(\mathbf{x}_i^T \mathbf{x}_i)(\mathbf{x}_k^T \mathbf{x}_k)}, \quad (1)$$

где $\mathbf{x}_i, \mathbf{x}_k$ – $N \times 1$ -векторы, компонентами которых являются описания сравниваемых изображений в пространстве признаков.

2. Показатель сопряженности с подпространством, натянутым на векторы признаков объектов из заданного класса:

$$R_{i,k} = \frac{\mathbf{x}_i^T \mathbf{X}_k [\mathbf{X}_k^T \mathbf{X}_k]^{-1} \mathbf{X}_k^T \mathbf{x}_i}{\mathbf{x}_i^T \mathbf{x}_i}. \quad (2)$$

Здесь \mathbf{x}_j – $N \times 1$ -вектор, описывающий изображение, предъявленное для установления принадлежности к некоторому (в данном случае k -му) классу, а \mathbf{X}_k – $N \times M$ -матрица, составленная из векторов признаков объектов, принадлежащих k -му классу.

3. Показатель сопряженности с нуль-пространством того же пространства:

$$S_{i,k} = \frac{\mathbf{x}_i^T \mathbf{T}_{0,k} \mathbf{T}_{0,k}^T \mathbf{x}_i}{\mathbf{x}_i^T \mathbf{x}_i}, \quad (3)$$

где $\mathbf{T}_{0,k}$ – матрица, составленная из $N - M$ собственных векторов, соответствующих нулевым собственным значениям $N \times N$ -матрицы $\mathbf{X}_k \mathbf{X}_k^T$, а \mathbf{X}_k – $N \times M$ – матрица, составленная из векторов признаков объектов, принадлежащих k -му классу.

Нетрудно заметить, что показатель парной сопряженности (1) является частным случаем показателя сопряженности с подпространством векторов из k -го класса, если этот класс состоит из единственного вектора \mathbf{x}_k , т.е. матрица \mathbf{X}_k имеет размерность $N \times 1$.

Между показателями $R_{i,k}$, $S_{i,k}$ также существует связь:

$$R_{i,k} = 1 - S_{i,k}, \quad (S_{i,k} = 1 - R_{i,k}) \quad (4)$$

Покажем это. Для этого нам понадобится следующая лемма:

Лемма. Пусть \mathbf{X} – $N \times M$ -матрица: $N > M$ и $\text{Rank}(\mathbf{X}) = M$, а \mathbf{x} – $N \times 1$ -вектор, тогда

$$\mathbf{E}_N - \mathbf{X}_k [\mathbf{X}_k^T \mathbf{X}_k]^{-1} \mathbf{X}_k^T = \mathbf{T}_0 \mathbf{T}_0^T \quad (5)$$

где $\mathbf{T}_0 - N \times (N - M)$ – матрица, составленная из собственных векторов, соответствующих нулевым собственным значениям $N \times N$ – матрицы $\mathbf{X}\mathbf{X}^T$.

Равенство (5) доказывается с использованием свойств $\mathbf{T}_0^T \mathbf{X} = 0$ ($\mathbf{X}^T \mathbf{T}_0 = 0$), $\mathbf{T}_0^T \mathbf{T}_0 = \mathbf{E}_{N-M}$ путем умножения обеих частей слева на \mathbf{T}_0^T и справа на \mathbf{T}_0 .

Теперь, в соответствии с (2), запишем:

$$1 - R_{i,k} = \frac{\mathbf{x}_i^T \mathbf{x}_i}{\mathbf{x}_i^T \mathbf{x}_i} - \frac{\mathbf{x}_i^T \mathbf{X}_k [\mathbf{X}_k^T \mathbf{X}_k]^{-1} \mathbf{X}_k^T \mathbf{x}_i}{\mathbf{x}_i^T \mathbf{x}_i} = \frac{\mathbf{x}_i^T [\mathbf{E}_N - \mathbf{X}_k [\mathbf{X}_k^T \mathbf{X}_k]^{-1} \mathbf{X}_k^T] \mathbf{x}_i}{\mathbf{x}_i^T \mathbf{x}_i}$$

Откуда, с учетом (5), следует, что $1 - R_{i,k} = S_{i,k}$. Ясно, что $R_{i,k} = 1 - S_{i,k}$.

Таким образом, показатели (1)-(3), при заданных i и k , эквивалентны и выбор одного из них для построения классификатора должен определяться лишь удобством в организации вычислительного процесса и традиционным требованием уменьшения числа операций. Поэтому далее там, где особенности реализации не имеют значения, мы будем употреблять термин «показатель сопряженности», не указывая о каком из них, конкретно, идет речь.

2. Правила принятия решений

Для каждого класса объектов (в данном случае одного или нескольких изображений одного лица) формируются либо $N \times 1$ -вектор \mathbf{x}_k , либо $N \times M$ – матрица \mathbf{X}_k , соответственно. Каждый вектор-столбец в \mathbf{X}_k (либо вектор \mathbf{x}_k) соответствует одному предварительно обработанному изображению лица. Предварительная обработка изображения заключается в локализации (выделении) лица, нормализации гистограммы яркости и приведении к заданному масштабу (размерности векторов). Компонентами векторов могут быть не все, а лишь отобранные каким-либо образом наиболее информативные отсчеты изображения.

Каждый класс может быть разбит на подклассы так, как описано в работе [5]. При этом каждому классу будет соответствовать несколько векторов $\mathbf{x}_{k,j}$, и/или матриц $\mathbf{X}_{k,j}$, где j – номер подкласса.

Для принятия решения о принадлежности к k -му классу в данном случае осуществляется объединение результатов проверки принадлежности к каждому подклассу, например, так, как описано в работах [3, 4]. Поскольку проверка принадлежности к классу и подклассу в данном случае осуществляется по одинаковым правилам, далее для простоты мы будем говорить лишь об определении принадлежности к классу.

Предположим, что для каждого (k -го) класса в соответствии с соотношениями (2)-(3) сформирована одна из следующих $N \times N$ -матриц $\mathbf{Q}_{k,(\ast)}$:

$$\mathbf{Q}_{k,R} = \mathbf{X}_k [\mathbf{X}_k^T \mathbf{X}_k]^{-1} \mathbf{X}_k^T, \quad (6)$$

$$\mathbf{Q}_{k,S} = \mathbf{T}_{0,k} \mathbf{T}_{0,k}^T. \quad (7)$$

Тогда правило принятия решения с учетом связи (4) показателей сопряженности R и S можно сформулировать в следующем виде: вектор \mathbf{x}_i принадлежит m -му классу:

$$\text{если } R_m = \max_{\forall k} R_k, \text{ где } R_k = \frac{\mathbf{x}_i \mathbf{Q}_{k,R} \mathbf{x}_i}{(\mathbf{x}_i^T \mathbf{x}_i)}, \quad (8)$$

$$\text{либо } S_m = \min_{\forall m} S_k, \text{ где } S_k = \frac{\mathbf{x}_i \mathbf{Q}_{k,S} \mathbf{x}_i}{(\mathbf{x}_i^T \mathbf{x}_i)}. \quad (9)$$

В случае, когда используется показатель парной сопряженности $r_{i,k}^2$, также можно вычислить матрицу

$$\mathbf{Q}_{k,R} = (\mathbf{x}_k \mathbf{x}_k^T) (\mathbf{x}_k^T \mathbf{x}_k)^{-1}$$

и применить условие (8). Однако в данном случае, как с точки зрения экономии памяти, так и уменьшения вычислительных затрат, предпочтительнее воспользоваться непосредственно соотношением (1).

Для реализации описанных правил принятия решений для каждого класса с использованием обучающей выборки изображений лиц заранее должна быть рассчитана и храниться в памяти одна из $N \times N$ – матриц $\mathbf{Q}_{k,(\ast)}$. Из соотношений (6), (7) видно, что матрицу $\mathbf{Q}_{k,R}$ выгоднее использовать в случае, когда число обучающих векторов M сравнительно невелико. Если же наоборот M велико настолько, что размерность нуль-пространства $N - M$ оказывается сравнительно (с размерностью N) малой, целесообразно использовать матрицу $\mathbf{Q}_{k,S}$.

В случае использования показателя парной сопряженности (1) вероятность распознавания, как правило, выше, если он вычисляется с каждым элементом каждого класса (а возможно и подкласса). При этом решение о принадлежности вектора \mathbf{x}_i k -му классу принимается, если наиболее «близкий» в смысле показателя (1) вектор \mathbf{x}_i также из этого класса.

Применение показателей сопряженности не накладывает никаких ограничений на выбор системы признаков. Одним из возможных подходов к формированию \mathbf{x}_i может быть преобразование исходных векторов изображений с использованием ортонормированной матрицы, составленной из собственных векторов, соответствующих наибольшему собственным значениям ковариационной матрицы.

Этот подход известен, как метод главных компонент [5], в сочетании с решающими правилами, основанными на вычислении показателей сопряженности, позволяет снизить вычислительную сложность задачи без существенной потери полезной информации. Далее будет приведен пример реализации этого подхода.

3. Примеры и результаты

Экспериментальное исследование классификаторов, построенных на основе описанных выше показателей сопряженности, проводилось с использованием стандартной базы данных ORL. Данная база содержит изображения лиц сорока человек. Для каждого человека имеется 10 различных ракурсов с произвольной мимикой. Таким образом, база данных содержит 400 изображений. На рис. 1 приведены примеры изображений пяти лиц из базы данных ORL.

Размер изображений оригинальной базы данных ORL равен 112×96 . Однако в экспериментах, с целью снижения вычислительных затрат, использовались изображения более малого размера 80×70 . Таким образом, каждое изображение представлялось в виде вектора x , размерностью 5600×1 .

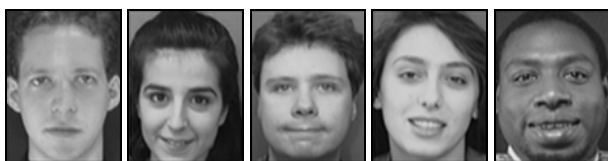


Рис. 1. Примеры изображений разных лиц (классов) из базы ORL

Из различных наборов этих векторов составлялись матрицы X_k , для каждого класса. С целью выявления зависимости вероятности распознавания от числа векторов в классе их число в матрице варьировалось от одного до пяти. На рис. 2 в качестве примера приведены пять изображений одного лица (класса), по которым была сформирована одна из матриц X_k размерности 5600×5 .



Рис. 2. Пример изображений, относящихся к одному классу

Для каждого класса с использованием матриц X_k по соотношениям (6), (7) формировались матрицы Q_{k*} для распознавания. Решение о принадлежности некоторого вектора из контрольной выборки m -му классу принималось по правилам (8), (9). В контрольную выборку были включены по пять изображений каждого лица, т.е. всего 200 лиц. Подчеркнем, что эти лица не использовались для формирования матриц Q_{k*} .

На рис. 3а и б представлены полученные в эксперименте зависимости вероятности распознавания от числа изображений, в классе (векторов в матрице X_k) для классификатора, построенного на критериях (1)-(3).

Векторы x_i в данном случае формировались путем строчной развертки изображений лиц. Как и следовало ожидать, при увеличении числа примеров, представляющих каждый класс, вероятность распознавания выше. Кроме того, из сравнения графиков видно, что при одинаковом числе образцов,

представляющих класс, качество распознавания также практически одинаково.

Следующий пример иллюстрирует подход, основанный на комбинации метода главных компонент и решающих правил, основанных на вычислении показателей сопряженности.

На рис. 4 приведен пример собственных векторов и соответствующих им собственных значений, вычисленных для указанного training set.

На рис. 5а, б приведен пример восстановления изображений лица по 40 главным компонентам, а на рис. 5в – собственное лицо, соответствующее собственному значению $\lambda = 0,228$. Приведенные примеры иллюстрируют, что переход к вектору главных компонент позволяет сохранить существенные черты лица. В то же время собственное лицо, соответствующее малому собственному значению ($\lambda = 0,228$), по существу является помехой.

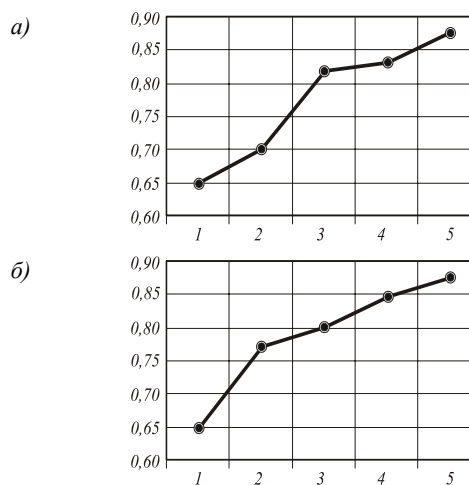


Рис. 3. Зависимость вероятности распознавания от объема обучающей выборки при использовании показателей а) парной сопряженности; б) сопряженности с пространством векторов-признаков из класса



$\lambda=8,745$ $\lambda=10,818$ $\lambda=16,859$ $\lambda=18,484$ $\lambda=23,283$

Рис. 4. Собственные лица и соответствующие им собственные значения

Для оценки качества восстановления изображений с использованием главных компонент проводился эксперимент, в котором их число варьировалось от 40 до 200 векторов. На рис. 6 показана зависимость средней ошибки восстановления от числа собственных векторов.

Эксперименты по оценке качества распознавания проводились для двух видов классификаторов.

1. Классификатор, использующий в качестве критерия близости векторов главных компонент евклидово расстояние. В этом случае вычислялись расстояния между предъявленным вектором и каждым вектором класса. Решение о принадлежности к классу принималось в случае, когда евклидово рас-

стояние до одного из векторов данного класса оказывалось минимально.

2. Классификатор, использующий в качестве критерия близости показатель сопряженности R (или S) по соотношениям (2), (3). Решение о принадлежности к классу принималось по правилу, описанному в пункте 3.

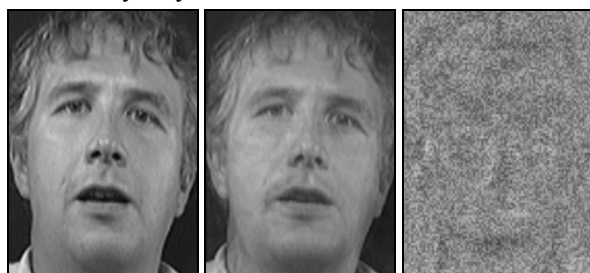


Рис. 5. Собственное лицо с малым собственным значением (а, б), (в) исходное изображение и его проекция на пространство из 40 собственных лиц

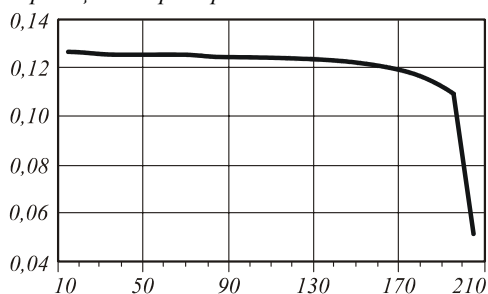


Рис. 6. Зависимости ошибки восстановления от числа собственных векторов

На рис. 7 представлен процент распознавания для критерия сопряженности (2) и критерия, основанного на евклидовом расстоянии в зависимости от числа обучающих примеров в классе. Как видно из рисунка, процент распознавания, полученный с использованием показателя сопряженности, выше в случае малого числа обучающих примеров в отдельном классе.

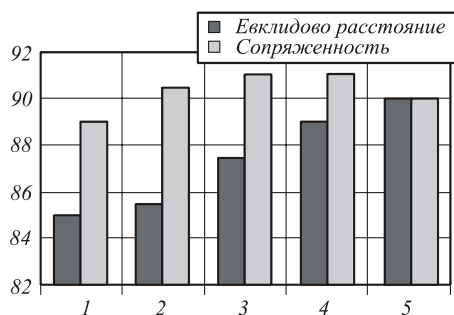


Рис. 7. Зависимость процента распознавания от числа обучающих векторов в классе

Заключение

Исследована возможность построения классификаторов для распознавания лиц, основанных на критериях сопряженности. Представлена связь показателей парной сопряженности, сопряженность с пространством столбцов матрицы признаков и нуль-пространством транспонированной матрицы признаков.

Построено унифицированное решающее правило принятия решений о принадлежности классу, учитывающее связь между показателями сопряженности.

Эксперименты подтвердили почти одинаковую эффективность всех показателей. Поэтому целесообразность применения конкретного показателя сопряженности в каждом случае определяется соотношением между размерностями N и M . В частности, если размерность N векторов признаков сравнительно невелика по сравнению с числом обучающих примеров M , т.е. размерность нуль-пространства меньше размерности пространства столбцов матрицы X_k выгоднее использовать матрицу $Q_{k,S}$ – (7).

Описанный метод распознавания может успешно применяться в распределенных системах видеоконтроля и видеонаблюдения. В данном случае может использоваться распределенная база данных лиц, которая пополняется без переобучения системы в целом. При реализации такой системы может использоваться идеология сервис-ориентированных систем, обсуждавшаяся в работе [6].

Благодарности

Работа выполнена при поддержке Министерства образования и науки РФ, правительства Самарской области и Американского фонда гражданских исследований и развития (CRDF Project SA-014-02) в рамках российско-американской программы «Фундаментальные исследования и высшее образование» (BRHE), а также при поддержке гранта Президента РФ № НШ-1007.2003.01 и грантов РФФИ № 05-01-08043 и № 04-07-90149.

Литература

1. Zhao W., Chellappa R., Rosenfeld A., Phillips P.J. Face Recognition // A Literature Survey, ACM Computing Surveys, 2003. P. 399-458.
2. Turk M.A., Pentland A.P. Face Recognition Using Eigenfaces, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Maui, Hawaii, USA, 3-6 June 1991. P. 586-591.
3. Belhumeur P.N., Hespanha J.P., Kriegman D.J. Eigenfaces vs. Fisherfaces: Recognition using Class Specific Linear Projection, Proc. of the 4th European Conference on Computer Vision, ECCV'96, 15-18 April 1996, Cambridge, UK, P. 45-58.
4. Fursov V.A., Kozin N.E. Stage-wise learning of radial neural networks // Proceedings of The 12th ISPE International Conference on Concurrent Engineering: Research and Applications, Focus Symposium Recursive Dynamics and Iterated Mappings in Service Modeling and Design, Ft. Worth/Dallas, USA, 25 – 29 July, pp. 391-396.
5. Fursov V.A., Kozin N.E. Algorithm for parallel learning of radial neural networks, Proceedings of The IASTED International Conference on Automation, Control and Applications (ACIT-ACA 2005), Novosibirsk, June 20-24, Russia. 2005. P. 481-485.
6. Turk A., Pentland A.P. Face Recognition Using Eigenfaces // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 3-6 June 1991, Maui, Hawaii, USA. P. 586-591.
7. Berger M., Sobolewski M. SILENUS // A Federated Service-oriented Approach to Distributed File Systems. P. 89-96.