

# ПОСТРОЕНИЕ ОПОРНЫХ ПОДПРОСТРАНСТВ В ЗАДАЧАХ РАСПОЗНАВАНИЯ ФРАКТАЛЬНЫХ ИЗОБРАЖЕНИЙ

В.А. Фурсов, Е.Ю. Минаев

Самарский государственный аэрокосмический университет имени академика С.П. Королёва (национальный исследовательский университет) (СГАУ), Самара, Россия

В статье решается задача распознавания фрактальных изображений. Рассматривается подход, основанный на использовании опорных подпространств, образованных подмножеством векторов распознаваемых классов. Предлагается новый, основанный на ортогональном разложении, алгоритм вычисления показателя сопряженности, с использованием которого формируются опорные подпространства, обеспечивающие повышение качества распознавания. Приводятся результаты экспериментов, показывающие, что качество распознавания растет даже при заметном уменьшении числа векторов в опорном подпространстве.

**Ключевые слова:** цифровая обработка изображений, распознавание фрактальных изображений, показатель сопряженности, бинарная и множественная классификация

## Введение

При анализе и распознавании изображений часто их представляют в виде вектора, компонентами которого являются значения яркости пикселей. Этот подход широко применяется в компьютерном зрении, в частности, при распознавании фрактальных изображений [1]. При переходе от обычных изображений к их фрактальным представлениям обычно удается существенно снизить требования к объему памяти для хранения исходных данных, при сохранении качества распознавания [2]. Например, при формировании фрактального образа размером  $16 \times 16$  из исходного изображения размером  $128 \times 128$  размерность вектора признаков уменьшается в 64 раза. Однако размерность вектора, представляющего фрактальное изображение, все-таки, остается достаточно высокой:  $256 \times 1$ .

Тем не менее, в интересах обеспечения высокого качества распознавания обычно применяют методы, в которых некоторая, достаточно простая, процедура распознавания строится на исходном или даже расширенном признаковом пространстве, например, метод опорных векторов (SVM) [3]. Метод SVM в настоящее время признается большинством исследователей наиболее эффективным в случае линейной разделимости классов. С использованием, так называемых, ядерных функций метод может использоваться также для классификации в случае отсутствия свойства линейной разделимости. Однако пока отсутствуют регулярные методы выбора наиболее подходящих ядерных функций. Другой проблемой метода является то, что опорные векторы определяются на этапе его настройки в результате решения оптимизационной задачи, что часто требует большого числа итераций и значительных вычислительных ресурсов.

В работе [4], в качестве альтернативы методу SVM, предложен метод опорных подпространств и приведены результаты сравнительных исследований этого метода с методом опорных векторов. Показано, что этот метод обеспечивает более высокое качество распознавания. Вместе с тем, оказалось, что качество распознавания существенным образом зависит как от способа формирования опорного подпространства, так и от количества включаемых в него векторов.

При распознавании фрактальных изображений, задача осложняется тем, что фрактальные изображения имеют вариации, в зависимости от числа итераций, на которых они получены. Поэтому возникает проблема формирования опорных подпространств с учетом фактора вариативности обучающих векторов.

### Постановка задачи

Каждое фрактальное изображение представляется  $N \times 1$ -вектором

$$\mathbf{x} = [x_1, x_2, \dots, x_N]^T, \quad (1)$$

компонентами которого являются числовые значения яркости, в  $N = W \times H$  точках, где  $W$ ,  $H$  – размеры изображений. Предполагается, что имеется  $M$  различных фрактальных вариантов изображений для каждого из  $K$  объектов. Векторы, соответствующие фрактальным изображениям одного объекта, составляют класс. Совокупность векторов, представляющих фрактальные изображения известных классов, образует обучающую выборку.

Для построения классификатора будем использовать подход, описанный в работе [4]. Для каждого ( $k$ -го) класса из принадлежащих ему  $M$  обучающих векторов  $\mathbf{x}_j(k)$ ,  $j = \overline{1, M}$ ,  $k = \overline{1, K}$  составляется –  $N \times M$  -матрица:

$$\mathbf{X}_k = [\mathbf{x}_1(k), \mathbf{x}_2(k), \dots, \mathbf{x}_j(k), \dots, \mathbf{x}_M(k)], \quad k = \overline{1, K} \quad (2)$$

и вычисляется  $N \times N$  -матрица  $k$ -го класса:

$$\mathbf{Q}_k = \mathbf{X}_k [\mathbf{X}_k^T \mathbf{X}_k]^{-1} \mathbf{X}_k^T, \quad k = \overline{1, K}, \quad (3)$$

которую далее мы будем называть решающей.

На этапе распознавания решение о принадлежности вектора  $\mathbf{x}_j$  к  $m$ -му классу принимается, если

$$R_m(\mathbf{x}_j) = \max_{k=1, \overline{K}} R_k(\mathbf{x}_j), \quad (4)$$

$$\text{где } R_k(\mathbf{x}_j) = \mathbf{x}_j^T \mathbf{Q}_k \mathbf{x}_j (\mathbf{x}_j^T \mathbf{x}_j)^{-1}, \quad k = \overline{1, K}. \quad (5)$$

– показатели сопряженности текущего вектора  $\mathbf{x}_j$  с каждым из распознаваемых классов.

Нетрудно заметить, что в данном методе информация о классах содержится в матрицах  $\mathbf{Q}_k$   $k = \overline{1, K}$ , вычисленных по матрицам  $\mathbf{X}_k(M)$ . В работе [4] для формирования этих матриц предложено использовать небольшое число обучающих векторов, образующих так называемые опорные подпространства классов. В указанной работе эти векторы выбирались из обучающего множества путем перебора всех возможных вариантов по критерию качества распознавания.

В настоящей работе исследуется схема формирования опорных подпространств, учитывающая специфику задачи распознавания фрактальных изображений. В частности, опорные подпространства каждого класса формируются на множестве различных фрактальных изображений одного объекта, полученных при различном числе итераций. При этом одновременно определяется число итераций формирования фрактального изображения, при котором обеспечивается заданное качество распознавания на обучающей выборке.

В работе решаются задачи бинарной и множественной классификации. Приводятся сравнительные результаты экспериментов на учебных наборах данных, полученных моделированием. Рассматриваются также вопросы снижения вычислительной сложности алгоритма на этапе распознавания.

### Построение опорных подпространств

Вначале рассмотрим простой случай различения двух классов фрактальных изображений. Предположим, что для исходных изображений двух объектов сформированы множества их фрактальных образов, имеющих некоторые различия, вследствие разного числа итераций и влияния случайных шумов. Задача состоит в том, чтобы из этих множеств сформировать опорные подпространства для классов  $k$  и  $l$ , представляемых  $N \times M$ -матрицами  $\mathbf{X}_k$  и  $\mathbf{X}_l$  с использованием которых далее решается задача бинарной классификации (4), (5), т.е. определения принадлежности вектора  $\mathbf{x}$  одному из этих двух классов –  $k$  или  $l$ .

При построении опорных подпространств мы будем использовать понятие разделимых по показателю сопряженности классов. Классы  $k$  и  $l$  назовем разделимыми по показателю сопряженности, если для любых векторов  $\mathbf{x}_{k,i} \in \mathbf{X}_k, i = \overline{1, M}$   $\mathbf{x}_{l,i} \in \mathbf{X}_l, i = \overline{1, M}$  из этих классов выполняются условия:

$$R(\mathbf{x}_{k,i}, \mathbf{Q}_k) > R(\mathbf{x}_{k,i}, \mathbf{Q}_l), \quad (6)$$

$$R(\mathbf{x}_{l,i}, \mathbf{Q}_l) > R(\mathbf{x}_{l,i}, \mathbf{Q}_k), \quad (7)$$

где  $R(\mathbf{x}_k, \mathbf{Q}_k)$ ,  $R(\mathbf{x}_k, \mathbf{Q}_l)$  – показатели сопряженности вектора  $\mathbf{x}_k \in \mathbf{X}_k$  со «своим» ( $k$ -м) и «чужим» – ( $l$ -м) классами соответственно,  $R(\mathbf{x}_l, \mathbf{Q}_k)$ ,  $R(\mathbf{x}_l, \mathbf{Q}_l)$  – показатели сопряженности вектора  $\mathbf{x}_l \in \mathbf{X}_l$  из  $l$ -го класса с теми же классами, а  $\mathbf{Q}_k$ ,  $\mathbf{Q}_l$  –  $N \times N$ -матрицы (3) для  $k$ -го и  $l$ -го классов соответственно.

Сопоставляя условия (6), (7) с решающим правилом (5), нетрудно заметить, что в случае их выполнения все векторы из обоих классов будут правильно распознаны. К сожалению, в действительности эти условия, как правило, не выполняются. Однако есть основания предположить, что качество распознавания будет выше, если условия (6), (7) будут выполняться, по крайней мере, для построенных в ходе обучения опорных подпространств.

Зададимся вопросом – всегда ли возможно построение разделимых по показателю сопряженности опорных подпространств. Пусть два класса заданы своими  $N \times M$ -

матрицами  $\mathbf{X}_k^{(M)}$  и  $\mathbf{X}_l^{(M)}$  так, что среди входящих в них векторов нет совпадающих. Выберем из этих матриц пару векторов  $\mathbf{x}_{k,M} \in \mathbf{X}_k^{(M)}$ ,  $\mathbf{x}_{l,M} \in \mathbf{X}_l^{(M)}$  таких, что

$$R(\mathbf{x}_{k,M}, \mathbf{Q}_k(M)) = \max_{i=1,M} R(\mathbf{x}_{k,i}, \mathbf{Q}_k(M)), \quad (8)$$

$$R(\mathbf{x}_{l,M}, \mathbf{Q}_l(M)) = \max_{i=1,M} R(\mathbf{x}_{l,i}, \mathbf{Q}_l(M)). \quad (9)$$

Если для этих векторов условия (6), (7) выполняются, то ясно, что они будут выполняться для всех остальных векторов. Следовательно, опорные пространства образованные множеством обучающих векторов, из которых составлены матрицы  $\mathbf{X}_k^{(M)}$  и  $\mathbf{X}_l^{(M)}$  разделимы по показателю сопряженности и эти матрицы могут непосредственно использоваться для формирования решающих матриц а  $\mathbf{Q}_k(M)$  и  $\mathbf{Q}_l(M)$  классов.

Если условия (6) и (7), или хотя бы одно из них, не выполняются, тогда исключим эти векторы из матриц  $\mathbf{X}_k^{(M)}$  и  $\mathbf{X}_l^{(M)}$ . Далее для векторов из (8), (9) проверим выполнение неравенств (6), (7) для вновь образованных матриц. Если условия выполнены, строим решающие матрицы  $\mathbf{Q}_k(M-1)$  и  $\mathbf{Q}_l(M-1)$ , если нет – продолжаем описанный процесс исключения векторов до тех пор, пока условия (6), (7) выполняются.

В самом неблагоприятном случае в результате процедуры исключения в каждом классе может остаться по одному вектору. Поскольку по предположению в исходных классах совпадающие векторы отсутствуют, эти два вектора и образуют искомые опорные подпространства и решающие матрицы  $\mathbf{Q}_k(1)$  и  $\mathbf{Q}_l(1)$ . Заметим, что в этом, последнем, случае каждая решающая  $N \times N$ -матрица имеет только одно, отличное от нуля, собственное значение. Это, конечно, сопровождается потерей информации о классе и может привести к потере качества распознавания.

Построенные описанным способом опорные подпространства допускают простую процедуру множественной классификации. В данном случае для отнесения распознаваемого вектора к одному из классов достаточно определить класс, с которым показатель сопряженности максимален. При этом на этапе построения опорных подпространств должна быть обеспечена парная разделимость по показателю сопряженности всех классов между собой. Заметим, что число обучающих векторов в опорных подпространствах разных классов может быть различным, поскольку число исключаемых векторов определяется классами, которые имели наибольшее «пересечение».

### Вычислительные аспекты

Как следует из (5) на этапе распознавания необходимо распознаваемый вектор умножать на  $N \times N$  –матрицу (3)  $\mathbf{Q} = \mathbf{X}[\mathbf{X}^T \mathbf{X}]^{-1} \mathbf{X}^T$ . Обычно  $N$  велико, поэтому требуются как большие вычислительные ресурсы, так и память. Построим вычислительную схему, позволяющую снизить эти требования.

С учетом (3) представим матрицу  $\mathbf{Q}$  в виде

$$\mathbf{Q} = \mathbf{T}_\lambda \mathbf{T}_\lambda^T, \quad (10)$$

где

$$\mathbf{T}_\lambda = \mathbf{X} [\mathbf{X}^T \mathbf{X}]^{-1/2}. \quad (11)$$

С учетом (10) вычислительную схему определения показателя сопряженности нормированного вектора  $\bar{\mathbf{x}}_j = \mathbf{x}_j (\mathbf{x}_j^T \mathbf{x}_j)^{-1/2}$  ( $\|\bar{\mathbf{x}}_j\| = 1$ ), в предположении, что  $N \times M$ -матрица  $\mathbf{T}_\lambda$  определена на этапе обучения, можно представить в виде следующих двух шагов:

1) вычисление  $M \times 1$ -вектора:

$$\mathbf{t}_\lambda = \mathbf{T}_\lambda^T \bar{\mathbf{x}}_j; \quad (12)$$

2) вычисление показателя сопряженности:

$$R(\bar{\mathbf{x}}_j) = \mathbf{t}_\lambda^T \mathbf{t}_\lambda. \quad (13)$$

Нетрудно заметить, что в данном случае потребуется  $(N+1)M$  умножений ( $N \times M$  умножений в (12) и  $M$  умножений в (13)). В то же время для вычисления  $R(\bar{\mathbf{x}}_j)$  по соотношению (4) потребовалось бы  $(N+1)N$  умножений.

Можно показать, что для вычисления матрицы  $\mathbf{T}_\lambda = \mathbf{X} [\mathbf{X}^T \mathbf{X}]^{-1/2}$  справедливо соотношение:

$$\mathbf{T}_\lambda = \mathbf{X} \mathbf{F} \mathbf{\Lambda}^{-1/2}, \quad (14)$$

где  $M \times M$ -матрица  $\mathbf{F}$  такая, что

$$\mathbf{X}^T \mathbf{X} = \mathbf{F} \mathbf{S}^T \mathbf{S} \mathbf{F}^T = \mathbf{F} \mathbf{\Lambda} \mathbf{F}^T,$$

а  $\mathbf{\Lambda} = \text{diag}(\lambda_1, \lambda_1, \dots, \lambda_M)$  – диагональная матрица, составленная из  $M$  собственных значений  $\lambda_i \geq 0$ ,  $i = \overline{1, M}$ , а векторы-столбцы матрицы  $\mathbf{F}$  являются соответствующими им собственными векторами.

Вычислительный выигрыш достигается за счет того, что решающее правило «спроектировано» в столбцовое пространство матрицы  $\mathbf{X}$ . Для реализации этой схемы требуется увеличение вычислительных ресурсов на этапе обучения. Однако это не является серьезным недостатком, т.к. обучение осуществляется один раз, кроме того во многих случаях, например, в мобильных системах технического зрения, критичным является быстродействие именно на этапе распознавания.

### Результаты экспериментов

Для проведения экспериментов в качестве исходных данных была выбрана общеизвестная база радиолокационных изображений MSTAR (moving and stationary target acquisition

and recognition). Использовались объекты BMP2, BTR70, T72, ZIL131, ZSU234, по каждому объекту 100 изображений использовались для обучения, и 100 изображений в качестве контрольных для распознавания. Для каждого исходного изображения был получен фрактальный образ, методом описанным в [1]. Примеры исходных изображений (128×128) из базы MSTAR и соответствующих им фрактальных изображений (16×16) приведены на рис. 1.

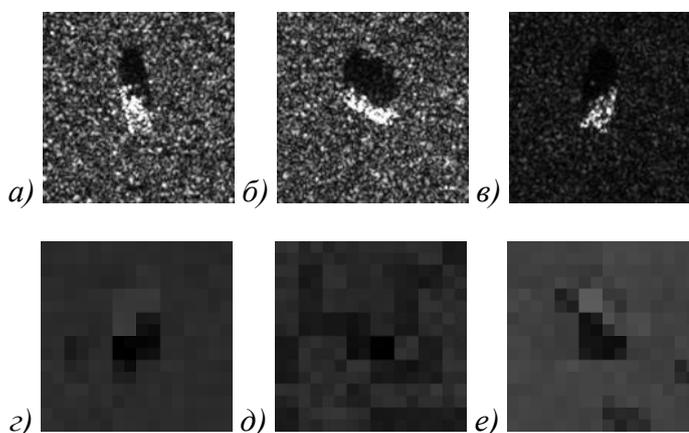


Рис. 1. Примеры изображений: исходных из базы MSTAR (а,б,в); соответствующих им фракталов (г,д,е)

В первом эксперименте рассматривалось распознавание из двух классов объектов (BMP2, T72). Для каждого класса было построено опорное подпространство, из которого последовательно исключались вектора по условиям (10), (11). Цель эксперимента – определить зависимость качества распознавания от количества исключаемых векторов. Данный эксперимент был повторен для случаев, когда исходные фрактальные образы зашумлены случайным шумом с разной амплитудой для ухудшения разделимости по показателю сопряженности. Результаты эксперимента приведены на рис. 2. По оси X отмечено количество векторов в опорном пространстве (из 100 исходных), по оси Y вероятность распознавания.

Получено, что при уменьшении опорного подпространства в некотором диапазоне качество распознавания увеличивается. Для определения достаточного количества векторов в опорном пространстве с использованием условий (6), (7) было определено максимальное количество векторов, при котором оба условия выполняются одновременно.

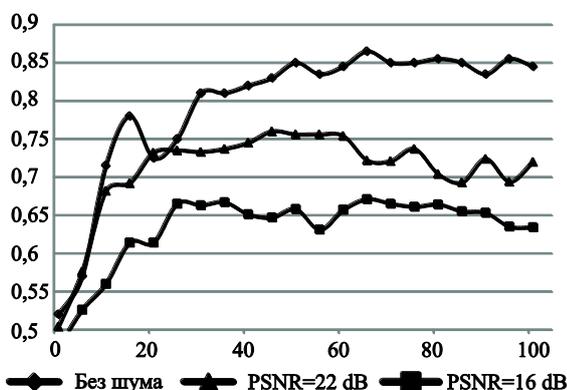


Рис. 2. Зависимости точности распознавания от количества векторов в опорном подпространстве

Построены опорные подпространства с учетом найденного количества векторов и результаты точности распознавания сравнены с результатами распознавания по исходному полному опорному подпространству.

Результаты распознавания представлены в табл. 1. Здесь PSNR - пиковое отношение сигнала к шуму в децибелах,  $P_{SVM}$ ,  $P_0$ ,  $P_{on}$  - вероятности распознавания методом SVM, с использованием исходного пространства и опорного подпространства соответственно,  $n_{on}$  - число векторов, оставшихся в опорном подпространстве. Видно, что при существенном уменьшении количества векторов в опорном подпространстве (на 21- 48% в зависимости от зашумленности данных), во всех случаях вероятность распознавания выросла на 1-3%. По сравнению с методом SVM, метод опорных подпространств обеспечивает существенно более высокое качество распознавания для неискаженных изображений. Для искаженных изображений метод опорных подпространств обеспечивает вероятность распознавания на 1-3% лучше SVM.

Табл. 1. Результаты двухклассовой классификации

PSNR	$P_{SVM}$	$P_0$	$P_{on}$	$n_{on}$
<b>Без шума</b>	0,755	0,84	0,855	79
<b>28 dB</b>	0,751	0,77	0,781	74
<b>22 dB</b>	0,714	0,72	0,754	60
<b>18 dB</b>	0,673	0,66	0,703	57
<b>16 dB</b>	0,651	0,63	0,658	51

Во втором эксперименте была рассмотрена процедура множественного распознавания на примере 5 классов объектов (BMP2, BTR70, T72, ZIL131, ZSU234). Исходные фрактальные образы были случайно искажены с PSNR=28 dB. Предварительно было найдено число векторов обеспечивающих попарную разделимость для каждой пары классов, и наименьший результат был выбран для количества векторов в опорном подпространстве для каждого класса.

В табл. 2 приведены результаты распознавания без исключения векторов.

Табл. 2. Результаты распознавания с помощью показателя сопряженности

	<b>BMP2</b>	<b>BTR70</b>	<b>T72</b>	<b>ZIL131</b>	<b>ZSU234</b>
<b>BMP2</b>	<b>0,741</b>	0,152	0,091	0	0,016
<b>BTR70</b>	0,162	<b>0,813</b>	0,025	0	0
<b>T72</b>	0,097	0,035	<b>0,757</b>	0,057	0,054
<b>ZIL131</b>	0	0	0,034	<b>0,829</b>	0,137
<b>ZSU234</b>	0	0	0,093	0,114	<b>0,793</b>

В табл. 3 приведены результаты распознавания, полученные для количества векторов в опорном подпространстве, при котором обеспечивается попарная разделимость по показателю сопряженности.

Табл. 3. Результаты распознавания при обеспечении делимости по показателю сопряженности

	<b>BMP2</b>	<b>BTR70</b>	<b>T72</b>	<b>ZIL131</b>	<b>ZSU234</b>
<b>BMP2</b>	<b>0,754</b>	0,149	0,097	0	0
<b>BTR70</b>	0,15	<b>0,825</b>	0,025	0	0
<b>T72</b>	0,096	0,026	<b>0,767</b>	0,051	0,06
<b>ZIL131</b>	0	0	0,03	<b>0,83</b>	0,14
<b>ZSU234</b>	0	0	0,081	0,119	<b>0,8</b>

Таким образом, показано, что построение опорных подпространств предложенным методом позволяет повысить точность распознавания при существенном снижении размерности исходных данных.

## Заключение

Существенное снижение размерности опорных подпространств не приводит к потере информативности выборки, более того качество распознавания даже возрастает. Предложенная схема вычисления показателя сопряженности позволяет также снизить вычислительные затраты. Если условия формирования изображений на этапе распознавания сохраняются, метод позволяет также прогнозировать число итераций для формирования фрактального изображения, обеспечивающего высокое качество распознавания. Подчеркнем, что в отличие от метода SVM упрощается процедура множественной классификации.

Работа выполнена при поддержке Министерства образования и науки РФ и гранта РФФИ №16-00729.

## Литература

1. Минаев, Е.Ю. Детектирование и распознавание объектов в системах автомобильной безопасности на основе фрактального анализа./ Е. Ю. Минаев, А. В. Никоноров // Компьютерная оптика, том 36, № 1, 2012, с. 124-130.
2. Минаев, Е.Ю. Высокоточная реконструкция пространственного положения пассивного цветового маркера по видеопоследовательности / Е. Ю. Минаев, А. В. Никоноров // Компьютерная оптика. – 2012. – Т. 36. – №. 4.
3. Лисицын, С.О. Распознавание дорожных знаков с помощью метода опорных векторов и гистограмм ориентированных градиентов./ С.О. Лисицын, О.А Байда. // Компьютерная оптика, том 36, № 2, 2012, с. 289-295.
4. Жердев, Д.А. Распознавание объектов по диаграммам рассеяния электромагнитного излучения на основе метода опорных подпространств / Д.А. Жердев, Н.Л. Казанский, В.А. Фурсов // Компьютерная оптика. – 2014. – № 38. – № 3 – С. 503-510.