

# КОМБИНИРОВАННЫЙ АЛГОРИТМ ПОВЫШЕНИЯ ВЕРОЯТНОСТИ РАСПОЗНАВАНИЯ ДВУМЕРНЫХ ИЗОБРАЖЕНИЙ

В.С. Ситников, Д.М. Шатохин

Одесский национальный политехнический университет,  
просп. Шевченко, 1, Одесса, 65044, Украина; e-mail: sitnvs@mail.ru

Предложен комбинированный алгоритм распознавания, объединяющий линейные и нелинейные признаки. Для получения признаков и первоначального понижения размерности пространства используется линейный и нелинейный метод главных компонент. Признаки объединяются в комплексное число, и далее производится распознавание по алгоритму Fisherfaces. Рассмотрено влияние коэффициентов нелинейного метода, а также метрик для определения принадлежности к определенному классу.

**Ключевые слова:** статистический метод, метод главных компонент, линейный дискриминантный анализ, комбинированный метод, «ядерные» методы

## Введение

При создании современного автомобиля на одно из первых мест выходят вопросы безопасности. Существующие электронные системы позволяют оптимизировать работу агрегатов и облегчить водителю вождение автомобилем, однако, при разработке новых систем, большое внимание уделяют системам, которые бы, анализируя ситуацию в автомобиле и вокруг него, выступали в роли советчика при принятии решения водителем [1].

При движении в сложных погодных условиях такие системы могли бы предупреждать водителя о состоянии дороги, о движении впереди идущей машины, о дорожных знаках, установленных на дороге и т.п. Повышение производительности вычислительной техники позволяет активно использовать системы технического зрения для распознавания двумерных образов. Например, системы распознавания дорожных знаков (СРДЗ) могут предупредить водителя во время движения о действующих ограничениях или иных ситуациях за счет считывания дорожных знаков [2].

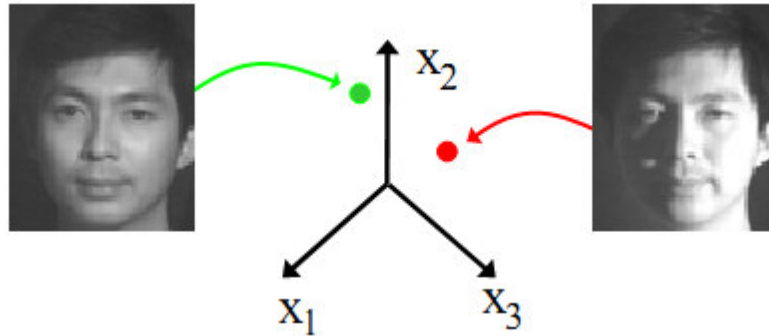
В области обработки изображений сложился общепринятый подход к обработке двумерных изображений в виде последовательности алгоритмов [3]:

- алгоритм локализации знака на видеоизображении;
- алгоритм определения формы и цвета, позволяющий произвести первичную классификацию изображения;
- алгоритм распознавания изображений;
- алгоритм формирования решений.

Одним из наиболее сложных алгоритмов является алгоритм распознавания изображения. Широко используемыми методами для решения подобных задач (в частности распознавание лиц) являются статистические методы. Семейство статистических методов предполагает наличие обучающей выборки, по которой строятся все возможные варианты и предположения о распознаваемых объектах. Для удобства описания обработки вводится понятие многомерного пространства за счет разворачивания изо-

браження в вектор  $\vec{T} = n \times m$ , где  $n$  и  $m$  ширина и высота изображения. Тогда, каждое изображение  $x$  есть точка в многомерном пространстве  $R$  размерностью  $N$ ,  $x \in R^N$  (рис. 1).

В каждом методе имеются свои правила выбора базиса для проектирования, но общим является то, что они позволяют уменьшить размерность базиса, оставив только наиболее значимые компоненты. Поэтому для каждого изображения выбирается вектор характерных признаков в выбранном базисе.



**Рис.1.** Понятие многомерного пространства

При распознавании широкое распространение получили линейные и нелинейные методы. Используя их или их сочетание, можно повысить степень распознавания изображений. К линейным методам следует отнести алгоритмы *Eigenfaces* (на основе метода главных компонент – PCA) и *Fisherfaces* (на основе линейного дискриминантного анализа – LDA) [4]. Применяя к этим алгоритмам нелинейные ядра, получим ядерный метод главных компонент (Kernel PCA или KPCA) и ядерный дискриминант Фишера (KFD или KLDA). На их основе созданы алгоритмы *Kernel Eigenfaces* и *Kernel Fisherfaces* [5, 6].

Рассмотрим задачу повышения распознавания изображений на основе линейных и нелинейных статистических методов.

**Метод главных компонент**

Пусть  $M$  набор образцов  $x_1, x_2, \dots, x_M$  в  $R^n$ , тогда общая ковариационная матрица набора будет определяться как

$$C = \frac{1}{M} \sum_{j=1}^M (x_j - \bar{x})(x_j - \bar{x})^T, \tag{1}$$

где  $\bar{x}$  – вектор средних значений по всем образцам.

Ортонормированный вектор собственных значений  $w_1, w_2, \dots, w_m$  матрицы  $C$ , соответствующий  $m$  наибольшим собственным значениям, позволяет выбрать их как оси проекции. В общем случае он может быть вычислен непосредственно. Однако, при распознавании образов размерность образца (вектора изображения) всегда велика (например, для изображения 50x50 это 2500 пикселей) и его вычисление занимает много времени. Для сокращения времени вычислений собственных векторов воспользуемся сингулярным разложением.

Пусть  $Q = [x_1 - \bar{x}, \dots, x_M - \bar{x}]$ , тогда ковариационная матрица представится в виде

$$C = \frac{1}{M} Q Q^T,$$

а матрица  $R = Q^T Q$  будет размером  $M \times M$ . Так как число образцов много меньше числа пикселей в изображении образца, то размерность матрицы  $R$  много меньше размерности матрицы  $C$ , поэтому вычислить ее собственный вектор будет проще. В соответствии с сингулярным разложением можно записать [7]

$$w_j = \frac{1}{\sqrt{\lambda_j}} Q u_j, j = 1, \dots, m, \quad (2)$$

где  $u_1, u_2, \dots, u_m$  собственные вектора в порядке соответствия  $m$  наибольшим собственным значениям  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m$  матрицы  $R$ .

Тогда для проекции образца на вычисленный нами базис собственных векторов  $w_j$  получим

$$y_j = w_j^T x = \frac{1}{\sqrt{\lambda_j}} Q^T u_j^T x, j = 1, \dots, m. \quad (3)$$

Вычисленные признаки каждого изображения  $y_1, y_2, \dots, y_m$  формируют вектор-столбец линейных (РСА) признаков  $Y = (y_1, y_2, \dots, y_m)^T$  для образца  $x$ .

### Ядерный метод главных компонент

Если нелинейное преобразование  $\Phi$  (ядро) проецирует пространство входных данных  $R^n$  в пространство признаков  $\mathfrak{Z}$

$$\begin{aligned} \Phi : R^n &\rightarrow \mathfrak{Z} \\ x &\rightarrow \Phi(x) \end{aligned} \quad (4)$$

то идея ядерного преобразования (КРСА) позволяет выполнить алгоритм РСА в пространстве признаков  $\mathfrak{Z}$ , в которое данные проецируются по нелинейному закону  $\Phi$ .

Таким образом, в пространстве  $\mathfrak{Z}$  матрица ковариации будет выглядеть как

$$C = \frac{1}{M} \sum_{j=1}^M (\Phi(x_j) - \bar{\Phi})(\Phi(x_j) - \bar{\Phi})^T, \quad (5)$$

где  $\bar{\Phi} = \frac{1}{M} \sum_{j=1}^M \Phi(x_j)$ .

Однако, центрировать данные в пространстве признаков  $\mathfrak{Z}$  из-за большого объема вычислений сложно. Поэтому рассмотрим не центрированную матрицу ковариации

$$\tilde{C} = \frac{1}{M} \sum_{j=1}^M \Phi(x_j) \Phi(x_j)^T, \quad (6)$$

В алгоритме КРСА вводится понятие ядра, что позволяет избежать сложных вычислений скалярного произведения [3]

$$k(x, y) = (\Phi(x_i) \cdot \Phi(x_j)). \quad (7)$$

Тогда, если  $Q = [\Phi(x_1), \dots, \Phi(x_M)]$ , то выражение (6) можно представить в виде

$$\tilde{C} = \frac{1}{M} Q Q^T.$$

Сформируем матрицу  $\tilde{R} = Q^T Q$ , используя выражение (7)

$$\tilde{R} = \Phi(x_j)^T \cdot \Phi(x_j) = (\Phi(x_i) \cdot \Phi(x_j)) = k(x_i, y_j). \quad (8)$$

Вычисляя собственные значения  $u_1, u_2, \dots, u_m$  собственных векторов в порядке соответствия  $m$  наибольшим собственным значениям  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m$  матрицы  $\tilde{R}$  можно показать, что собственные вектора матрицы  $\tilde{C}$  на основе сингулярного разложения будут иметь вид

$$w_j = \frac{1}{\sqrt{\lambda_j}} Q u_j, j = 1, \dots, m, \quad (9)$$

где  $w_1, w_2, \dots, w_m$  – собственные вектора  $\tilde{C}$ .

Для вычисления собственных значений и вектора матрицы ковариации  $C$  (5), на основе матрицы  $\tilde{R}$ , необходимо ее центрировать и произвести расчет в соответствии с (9). Для центрирования воспользуемся выражением

$$R = \tilde{R} - 1_M \tilde{R} - \tilde{R} 1_M + 1_M \tilde{R} 1_M, \quad (10)$$

где  $1_M$  – матрица, состоящая из единиц.

Тогда для проекции образца на вычисленный базис собственных векторов  $w_j$  запишем

$$y_j = w_j^T \Phi(x) = \frac{1}{\sqrt{\lambda_j}} Q^T u_j^T \Phi(x) = \frac{1}{\sqrt{\lambda_j}} u_j^T [k(x_1, x), k(x_2, x), \dots, k(x_M, x)], \quad (11)$$

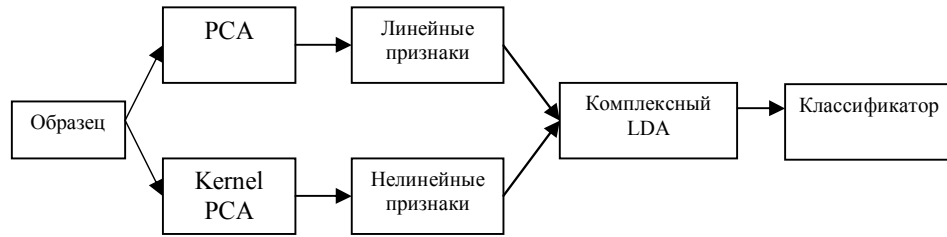
$$j = 1, \dots, m$$

Вычисленные признаки каждого изображения  $y_1, y_2, \dots, y_m$  формируют вектор-столбец нелинейных признаков  $Y = (y_1, y_2, \dots, y_m)^T$  для образца  $x$ .

### Комплексный LDA

В результате проведенных преобразований получено два пространства признаков – линейное (РСА) и нелинейное (КРСА), обозначим их соответственно как  $A$  и  $B$ . Пусть наши образцы принадлежат пространству  $\Omega$ . Получив их разложения на линейном наборе признаков  $\alpha \in A$  и нелинейном наборе признаков  $\beta \in B$ , объединим их в

некоторый комплексный вектор  $\gamma = \alpha + i\beta \in C$ . Вектора  $\alpha$  и  $\beta$  должны быть одинаковой размерности, либо должны быть приведены к этому, рисунок 2.



**Рис.2.** Комбинированный алгоритм

Таким образом, в новом пространстве между классовой, внутри классовой и общей матрицы разброса будут иметь следующий вид

$$S_b = \sum_{i=1}^c N_i (m_i - m_0)(m_i - m_0)^H, \quad (12)$$

$$S_w = \sum_{i=1}^c \{ \sum (X - m_0)(X - m_0)^H |w_i\}, \quad (13)$$

$$S_t = S_b + S_w = \sum (X - m_0)(X - m_0)^H, \quad (14)$$

где  $N_i$  – число образцов класса  $i$ ,  $m_i$  – среднее значение по классу  $i$ ,  $m_0$  – среднее значение по всей выборке.

Из выражений (12)-(14) следует, что все три матрицы  $S_b, S_w$  и  $S_t$  являются полуположительно определёнными эрмитовыми матрицами. Кроме того, матрицы  $S_w$  и  $S_t$  являются положительно определёнными в случае если  $S_w$  несингулярна. Пусть матрица  $S_w$  несингулярна.

Функцию дискриминантного критерия Фишера объявим как:

$$J(\varphi) = \frac{\varphi^H S_b \varphi}{\varphi^H S_w \varphi}, \quad (15)$$

где  $\varphi$  –  $n$ -размерный ненулевой вектор.

Для положительно определённой матрицы  $S_w$  и полуположительно определённой матрицы  $S_b$ , для любого  $\varphi$ , имеем  $\varphi^H S_b \varphi \geq 0$  и  $\varphi^H S_w \varphi > 0$ . Это означает что значения дискриминантного критерия  $J(\varphi)$  являются неотрицательными действительными числами. В физическом понимании это означает, что критерий Фишера в комплексном пространстве аналогичен критерию в евклидовом пространстве.

Вектор  $\varphi^*$ , соответствующий максимальному значению критерия  $J(\varphi)$ , называют оптимальным направлением проекции. При проекции образцов в  $\varphi^*$  отношение между классового и внутри классового рассеивания будет максимальным. Вектор  $\varphi^*$  можно выбрать как обобщенный вектор собственных значений выражения  $S_b X = \lambda S_w X$ , соответствующий максимальным обобщенным собственным значениям.

Как показано в [4] оптимальное число осей  $\mathbf{P} = (\varphi_1, \varphi_2, \dots, \varphi_d)$  выбирается из соотношения  $d \leq c - 1$ , где  $c$  – число классов. Далее базис  $\mathbf{P}$  используется для проецирования в комплексное пространство признаков. Для вычисления признаков по комплексному алгоритму LDA используется формула

$$Z = P^H Y, \quad (16)$$

где  $Y$  – проекция исходного образца в комплексное пространство,  $Z$  – вектор признаков по методу LDA.

### Постановка эксперимента и анализ результатов

Для проверки работоспособности комбинированного алгоритма распознавания проведено тестирование, для которой создана база изображений знаков ограничения скорости от 5, 10, 20 и далее до 120 км/ч, всего 13 классов знаков. Обучающую выборку для каждого класса составляют 10 изображений, которые сделаны фронтально, под углом слева и справа, сверху и снизу и в их комбинациях. Кроме того одно изображение с ухудшенным освещением, рисунок 3. Каждое изображение обрезано по размеру знака и приведено к разрешению  $50 \times 50$  пикселей. Над каждым изображением проведено автоматическое контрастирование. Аналогичные преобразования выполнены для тестовых изображений, которые были сняты с произвольных позиций.

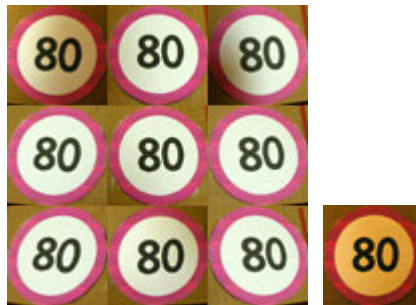


Рис. 3. Обучающая выборка для одного класса

Каждое изображение сохранено в формате JPEG и ему присвоено название в соответствии с маской <класс>-<номер>.jpg (например, 80-2541.jpg), что в дальнейшем позволило провести проверку правильности, путем сравнения классов.

При тестировании комбинированный метод сравнивался с методом *Eigenfaces*, *Fisherfaces*, *Kernel Eigenfaces*. На первом этапе тестирования использовалось 100 главных компонент (для *Eigenfaces* и *Kernel Eigenfaces*), в комбинированном методе и методе *Fisherfaces* количество компонент ограничено числом классов и равно  $d = c - 1 = 12$ . Для метода *Kernel Eigenfaces* использовалось полиномиальное ядро вида  $k(x, y) = (x \cdot y + 1)^q$ , которое давало наилучшие результаты [5].

Оценим оптимальную степень полиномиального ядра  $q$ . Воспользуемся для этого метрика Евклидоваго расстояния и косинуса угла между двумя векторами

$$g_i(Z) = \sqrt{(Z - \mu_i)^H (Z - \mu_i)}, \quad (17)$$

$$d_l(Z) = -\frac{Z \cdot Y}{\|Z\| \|Y\|}, \quad (18)$$

где  $Z$  – тестовый образец,  $\mu_l$  – вектор среднего по классу  $l$ ,  $Y$  – образец в обучающей выборке.

При этом, если для образца  $Z$

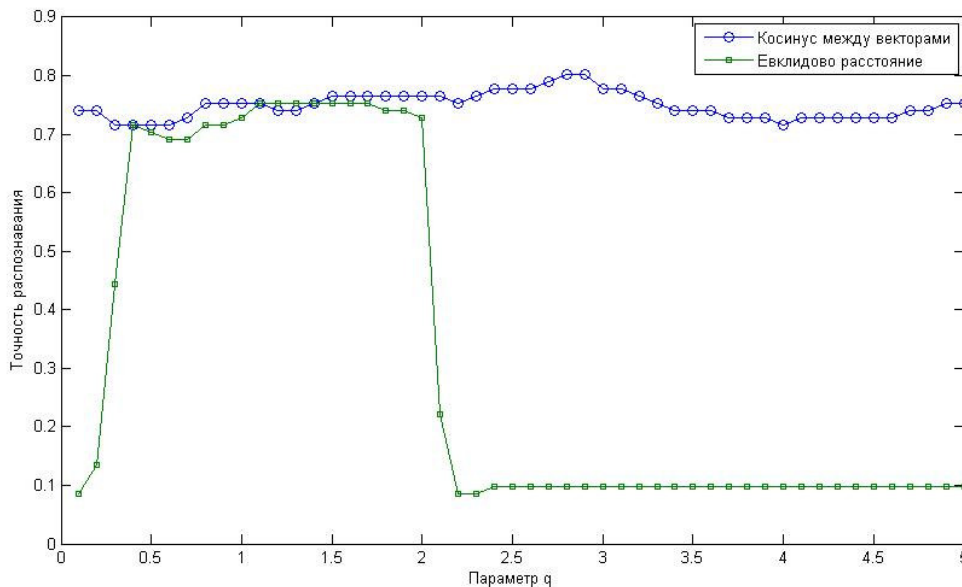
$$g_k(Z) = \min_l (g_l(Z)) \text{ или } d_k(Z) = \min_l (d_l(Z)),$$

то образец  $Z$  принадлежит классу  $k$ .

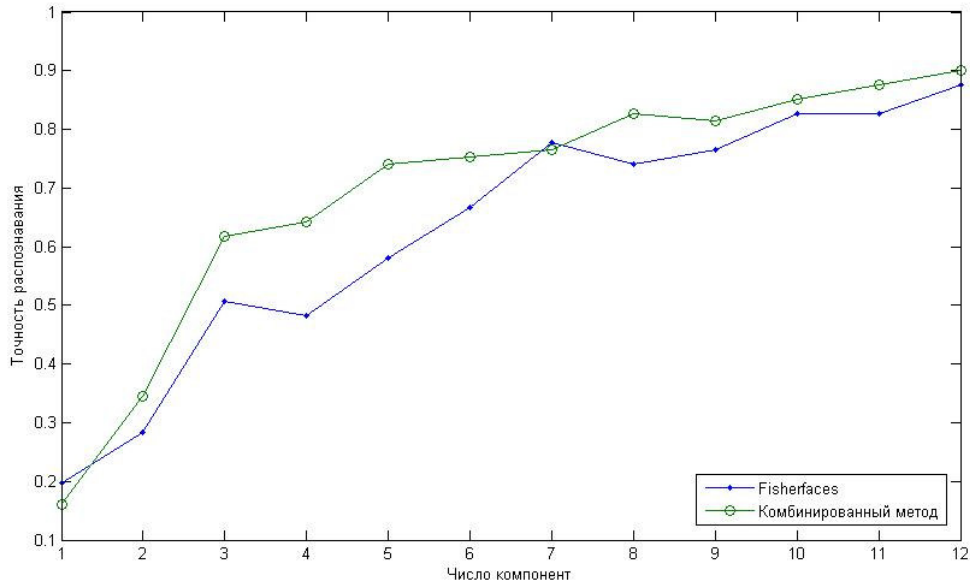
Результаты оценивания степени полиномиального ядра  $q$ , приведены на рисунке 4. Лучшие результаты получены по формуле (18) при параметре  $q = 2.8$ . Для евклидова расстояния оптимальное значение  $q = 1.8$ . В дальнейшем во всех исследованиях использовалась метрика по косинусу угла.

Анализ вероятности распознавания изображений показал, что чем больше число используемых компонент, тем выше точность распознавания. В общем зависимость точности распознавания от числа компонент близка к экспоненциальной и при большом числе компонент она стабилизируется, что нельзя сказать о алгоритме *Fisherfaces*, рисунок 5.

Результаты тестирования вероятности распознавания изображений приведены в таблице. Следует отметить, что комбинированный алгоритм дает прирост вероятности распознаваний изображений на 2.5%.



**Рис 4.** Сравнение метрик для метода *Kernel Eigenfaces* в зависимости от параметра  $q$  полиномиального ядра



**Рис 5.** Сравнение предложенного метода и метода *Fisherfaces* в зависимости от числа компонент

**Таблица 1.**  
Результаты тестирования вероятности распознавания изображений (%)

<i>Eigenfaces</i>	<i>Kernel Eigenfaces</i>	<i>Fisherfaces</i>	Комбинированный алгоритм
62.96	72.84	87.65	90.12

Предложенный комбинированный алгоритм распознавания изображений можно рекомендовать для использования в системе технического зрения распознавания дорожных знаков.

### Список литературы

1. Fang C.Y. et al. An automatic road sign recognition system based on a computational model of human recognition processing // *Computer Vision and Image Understanding*. – 2004. – Vol.96, Iss.2. – PP. 237-268.
2. Система распознавания дорожных знаков [Электронный ресурс] / Системы современного автомобиля. – Режим доступа: [http://systemsauto.ru/active/traffic\\_sign\\_recognition.html](http://systemsauto.ru/active/traffic_sign_recognition.html)
3. Escalera S. et al. Traffic-Sign Recognition System // *SpringerBriefs in Computer Science 2011*. – Springer, 2011. – 96 p., 34 illus.
4. Belhumeur P.N., Hespanha J.P., Kriegman D.J. Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection // *Proc. of the 4<sup>th</sup> European Conference on Computer Vision*. – Cambridge, UK. – 1996. – PP. 45-58.
5. Yang M.H. Kernel Eigenfaces vs. Kernel Fisherfaces: Face Recognition Using Kernel Methods // *Proc. of the 5<sup>th</sup> IEEE International Conference on Automatic Face and Gesture Recognition*. – Washington D.C., USA. – 2002. – PP. 215-220.
6. Schölkopf B., Smola A., Müller K.-R. Nonlinear component analysis as a kernel eigenvalue problem // *Neural Computation*. – 1998. – Vol.10, No.5. – PP. 1299-1319.
7. Голуб Дж. Матричные вычисления / Голуб Дж., Ван Лоун Ч. – М.: Мир, 1999. – 548 с.



В.С. Ситников, Д.М. Шатохин

## КОМБІНОВАНИЙ АЛГОРИТМ ПІДВИЩЕННЯ ВІРОГІДНОСТІ РОЗПІЗНАВАННЯ ДВОВИМІРНИХ ЗОБРАЖЕНЬ

Запропоновано комбінований метод розпізнавання двовимірних образів на основі лінійних і нелінійних статистичних методів. Для отримання ознак і початкового зниження розмірності простору використовується лінійний та нелінійний метод головних компонент. Ознаки об'єднуються в комплексне число, і далі проводиться розпізнавання за алгоритмом Fisherfaces. Розглянуто вплив коефіцієнтів нелінійного методу, а також метрик для визначення належності до певного класу.

**Ключові слова:** статистичний метод, метод головних компонент, лінійний дискримінантний аналіз, комбінований метод, «ядерні» методи

V. Sitnikov, D. Shatokhin

## COMBINED ALGORITHM FOR INCREASING THE RECOGNITION RATE OF RECOGNITION OF TWO-DIMENSIONAL IMAGES

Proposed combined recognition algorithm combining linear and nonlinear features. To obtain the original features and dimension reduction used linear and nonlinear principal component analysis. The features are combined into a complex number, and then made recognition algorithm Fisherfaces. Considered influence of coefficients of nonlinear method, and metrics for metrics for determining the belonging to the particular class.

**Keywords:** statistical techniques, principal component analysis, linear discriminant analysis, combined method, kernel methods