Алгоритм поэкземплярного распознавания на основе метода опорных векторов

The single image-understanding algorithm based on the SVM method

Ключевые слова: распознавание – understanding; опорный вектор – support vector; гиперсфера – hypersphere.

В статье представлен алгоритм распознавания на основе метода опорных векторов с возможностью применения к задаче с количеством гипотез более двух, а также выделением новых объектов распознавания. Показана возможность использования ядра радиального типа. Предложена процедура настройки параметров классификаторов и ядра, позволяющая минимизировать вероятность ошибочных решений распознавания.

The article presents the understanding algorithm based on the SVM method with applying to the problem more than two hypotheses, as well as to discover new objects of recognition. The possibility of the radial type core is demonstrated. The procedure for setting parameters of classifiers and the core is suggested allowing to minimize the solution error probability.

СУББОТЕНКО / Александр Владимирович
(subbiki@yandex.ru)
Академия ФСО России,
научный сотрудник.
г. Орел

Основной целью систем распознавания является процесс отнесения исследуемого объекта к одному или нескольким из $k+1$ взаимоисключающих классов [1]. В настоящее время наиболее эффективным средством решения задач классификации объектов в данных условиях является одноклассовый метод опорных векторов (МОВ). Основным преимуществом одноклассового МОВ является возможность выявлять новые классы объектов [2].

Основная идея одноклассового МОВ состоит в том, чтобы получить границу сферической формы вокруг набора классифицируемых данных с минимальным радиусом и максимальным количеством объектов обучающей выборки внутри нее [2]. Соответствующая гиперсфера характеризуется двумя параметрами: центром $\bar{c}$ и радиусом R (рис. 1).

При этом целевая функция имеет вид

$$\min_{\bar{c},R,\xi_1,\dots,\xi_n} R^2 + \frac{1}{n\cdot\nu}\sum_{i=1}^{n}\xi_i \tag{1}$$

при условии:

$$\|\bar{x}_i - \bar{c}\|^2 \le R^2 + \xi_i,\ \xi_i \ge 0, i = 1,\dots,n. \tag{2}$$

где $\nu \in (0;1]$ – переменная, характеризующая долю объектов обучающей выборки, выходящих за границы гиперсферы;

ξ_i – объекты обучающей выборки, выходящие за границы гиперсферы.

Решение задачи (1) при условии (2) предполагает построение функции Лагранжа следующего вида [3]:

$$L(R,\bar{c},\xi,\lambda) = R^2 + \frac{1}{n\cdot\nu}\sum_{i}^{n}\xi_i - \\ -\sum_{i}^{n}\lambda_i\cdot(R^2+\xi_i-\|\bar{x}_i-\bar{c}\|^2), \tag{3}$$

где $\lambda_i \ge 0$ множители Лагранжа.

Из теоремы Куна-Таккера [3] следует, что оптимальное решение данной задачи равносильно поиску седловой точки лагранжиана. При этом из необходимых условий существования седловой точки следует:

$$\frac{\partial L}{\partial R} = 2R - 2R\sum_{i}^{n}\lambda_i = 0 \Rightarrow \sum_{i}^{n}\lambda_i = \\ = 1\ , \frac{\partial L}{\partial \bar{c}} = -2\sum_{i}^{n}\lambda_i\cdot x_i + 2\bar{c}\sum_{i}^{n}\lambda_i =$$

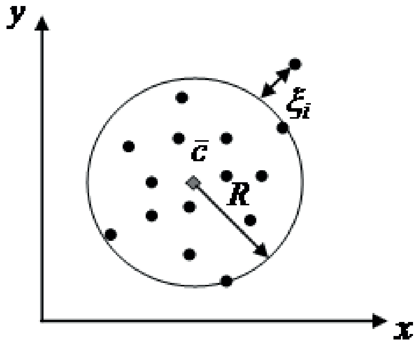

Рис. 1. Графическая интерпретация метода опорных векторов

$$= 0 \Rightarrow \bar{c} = \sum_{i}^{n} \lambda_i \cdot x_i \ ,$$

$$\frac{\partial L}{\partial \xi_i} = \frac{1}{n \cdot \nu} - \lambda_i = 0 \Rightarrow 0 \le \lambda_i \le \frac{1}{n \cdot \nu} \ ,$$

$$\text{т. к. } \lambda_i \ge 0 . \qquad (4)$$

Для того чтобы найти седловую точку лагранжиана необходимо минимизировать его относительно $(R; \bar{c}; \xi)$, а после этого максимизировать относительно множителей Лагранжа λ. Для этого необходимо подставить решения ограничений (4) в исходную функцию Лагранжа (3), получая при этом двойственную задачу максимизации функции следующего вида:

$$L = \max_{\lambda_1,\ldots,\lambda_n} \sum_{i=1}^{n} \lambda_i \cdot (\bar{x}_i^{\,T} \cdot \bar{x}_i) - \sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i \cdot \lambda_j \cdot (\bar{x}_i^{\,T} \cdot \bar{x}_j) \qquad (5)$$

при ограничениях

$$\sum_{i=1}^{n} \lambda_i = 1, \quad 0 \le \lambda_i \le \frac{1}{n \cdot \nu}, \quad i = 1,\ldots,n . \qquad (6)$$

Следовательно, для обучения классификатора, построенного по МОВ, можно решать как прямую задачу оптимизации, так и двойственную. При этом двойственная задача содержит меньше переменных, а граничные условия имеют простую форму. Таким образом, имеет место задача квадратичного программирования, решение которой в теории оптимизации может быть получено, например, градиентными методами или методом покоординатного спуска [3], и дает набор значений множителей Лагранжа λ, удовлетворяющих ограничению (6).

Центр гиперсферы определяется следующим выражением:

$$\bar{c} = \sum_{i=1}^{n} \lambda_i \cdot \bar{x}_i \ , \qquad (7)$$

причем для вычисления центра гиперсферы используются только те вектора из всего набора описывающих k-й класс, множитель Лагранжа при которых больше нуля $\lambda_i > 0$. Такие вектора называются опорными.

Для того чтобы вычислить радиус гиперсферы R, необходимо вычислить расстояние от центра гиперсферы до одного из опорных векторов согласно следующего выражения:

$$R^2 = \left\| \bar{x}_k - \bar{c} \right\|^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i \cdot \lambda_j \cdot (\bar{x}_i^{\,T} \cdot \bar{x}_j) -$$

$$- 2 \cdot \sum_{i=1}^{n} \lambda_i \cdot (\bar{x}_i^{\,T} \cdot \bar{x}_k) + (\bar{x}_k^{T} \cdot \bar{x}_k) \ , \qquad (8)$$

где $\bar{x}_k$ – любой из опорных векторов, множитель при котором удовлетворяет условию

$$0 < \lambda_i < \frac{1}{n \cdot \nu} .$$

Для определения принадлежности образа $\bar{x}$ объекта к классу с центром $\bar{c}$ используется следующее выражение (решающее правило):

$$f(\bar{x}) = R^2 - \left\| \bar{x} - \bar{c} \right\|^2 \ge 0 \ , \qquad (9)$$

согласно которому, если значение функции $f(\bar{x})$ больше или равно нулю, то образ $\bar{x}$ объекта расположен внутри гиперсферы и принадлежит данному классу, в противном случае – вне гиперсферы.

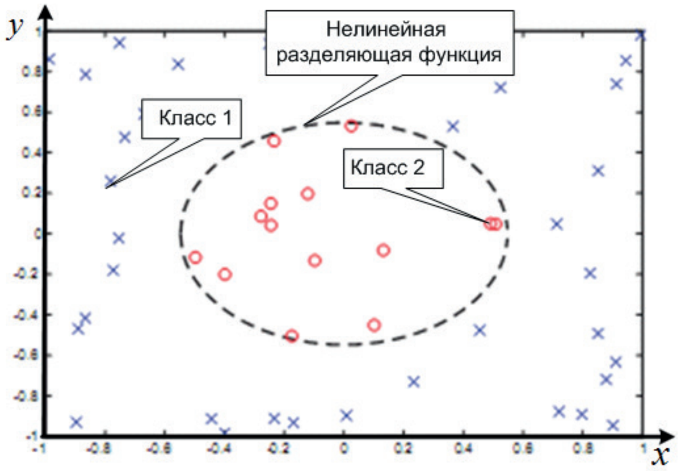

Рис. 2. Линейно-неразделимые классы

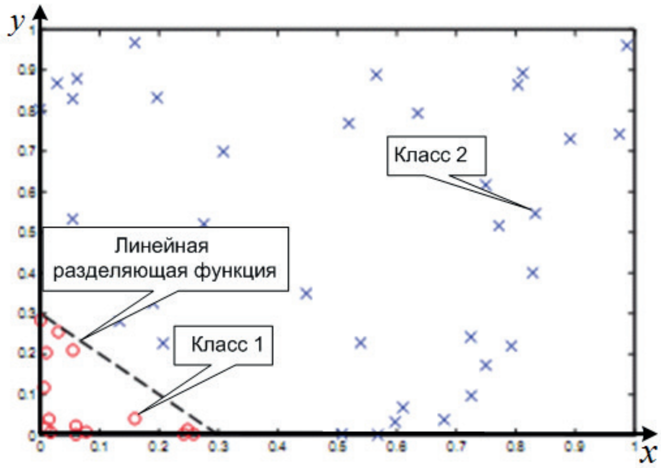

Рис. 3. Линейно-разделимые классы при переходе в пространство $(x_1^2 \cdot x_2^2)$

На практике, как правило, классы объектов могут быть линейно неразделимы, что приводит к ошибкам при распознавании (рис. 2). Линейный классификатор, построенный на основе МОВ, можно обобщить на случай построения нелинейных разделяющих поверхностей с помощью, так называемого ядрового перехода [2]. Векторы выборки $\{\bar{x}_1, \bar{x}_2, ..., \bar{x}_n\}$, принадлежащие пространству R^n, отображаются в пространство более высокой размерности с помощью некоторого нелинейного отображения, выбранного априори:

$$\{\bar{x}_1, \bar{x}_2, ..., \bar{x}_n\} \to \{\phi(\bar{x}_1), \phi(\bar{x}_2), ..., \phi(\bar{x}_n)\}. \quad (10)$$

Результатом отображения являются векторы $\{\phi(\bar{x}_1), \phi(\bar{x}_2), ..., \phi(\bar{x}_n)\}$ в пространстве признаков R^N. Исходное пространство R^n переходит при отображении в некоторое подмножество пространства признаков R^N, которое принято называть спрямляющим пространством.

Таким образом, получаемая гиперсфера в исходном пространстве R^n в общем случае имеет нелинейную поверхность и является прообразом линейной разделяющей гиперсферы, построенной в спрямляющем пространстве признаков R^N (рис. 3). В случае построения нелинейного классификатора в исходном пространстве скалярное произведение $(\bar{x}_i^T \cdot \bar{x}_j)$ представляется в следующем виде:

$$k(\bar{x}_i, \bar{x}_j) = \phi(\bar{x}_i)^T \cdot \phi(\bar{x}_j) \quad (11)$$

Функцию (11) будем в дальнейшем называть ядровой, а выражение (9) с учетом выражений (8) и (11) примет следующий вид:

$$f(\bar{x}) = \left(-2\sum_{i=1}^{n}\lambda_i \cdot k(\bar{x}_i, \bar{x}_k) + k(\bar{x}_k, \bar{x}_k)\right) - \\ -\left(-2\sum_{i=1}^{n}\lambda_i \cdot k(\bar{x}_i, \bar{x}) + k(\bar{x}, \bar{x})\right) \geq 0. \quad (12)$$

Анализ показал, что для построения нелинейного классификатора в исходном пространстве R^n с помощью линейного классификатора достаточно знать только функцию ядра. При построении классификаторов на основе одноклассового МОВ целесообразно использовать ядро вида:

$$k(\bar{x}_i, \bar{x}_j) = \mathrm{e}^{-\gamma \cdot (\|\bar{x}_i - \bar{x}_j\|)^2}, \quad (13)$$

где γ – параметр, характеризующий степень искривленности разделяющей гиперсферы. Данный тип ядра позволяет при корректном подборе параметров классификатора минимизировать ошибки классификации и строить замкнутые нелинейные разделяющие поверхности (рис. 4).

Качество сформированного правила принятия решения (12) полностью определяется вероятностями ошибочных решений, допускаемых при распознавании объектов. Ошибка распознавания возникает из-за пересечения классов в многомерном признаковом пространстве, когда объекты из обучающей выборки одного класса в результате проекции на гиперплоскость ошибочно относят к объектам другого класса. Снизить вероятность ошибочных решений распознавания возможно за счет настройки параметров классификатора в процессе обучения [4].

Применительно к задаче поэкземплярного распознавания передатчиков качество правила принятия решения будет определяться вероятностями ложной тревоги $P_{\text{л.т}}$ и пропуска цели $P_{\text{п.ц}}$, а «регулируемыми» параметрами, позволяющими осуществлять контроль за достоверностью распознавания по МОВ, будет выступать параметр – параметр классификатора, характеризующий долю объектов обучающей выборки, выходящих за границы гиперсферы; а в случае использования гауссова ядра – параметр γ, характеризующий степень искривленности поверхности разделяющей гиперсферы (рис. 5) [2]. Рациональный выбор последнего позволяет изменять количество опорных векторов, что в свою очередь позволяет регулировать значение вероятности правильного распознавания.

Исходя из этого, процесс определения параметров классификатора можно представить в виде алгоритма (рис. 6), осуществляющего подбор параметров ν и γ по критерию минимума ошибочных решений распознавания.

В результате экспериментальных исследований установлено, что $\nu \in (0;1]$ целесообразно изменять с шагом равным 0,01, параметр γ – в диапазоне значений $\gamma \in (1,20)$, а в качестве начального порогового значения вероятности ошибочного распознавания следует принять $P_{\text{ОР}}^{\text{пор}} = 0{,}5$. Кроме того, в качестве исходных данных задается количество классов *N*. Для каждого из используемых классификаторов в процессе обучения подсчитывается величина вероятности ошибочных решений $P_{\text{ОР}\,i}$. Вероятности ошибочных решений $P_{\text{ОР}\,i}$ для каждого из *i* используемых в процессе обучения классификаторов и значения параметров ν и γ рассчитывались методом статистического моделирования. По завершению обучения для классификаторов вычисляются значения их весовых коэффициентов значимости:

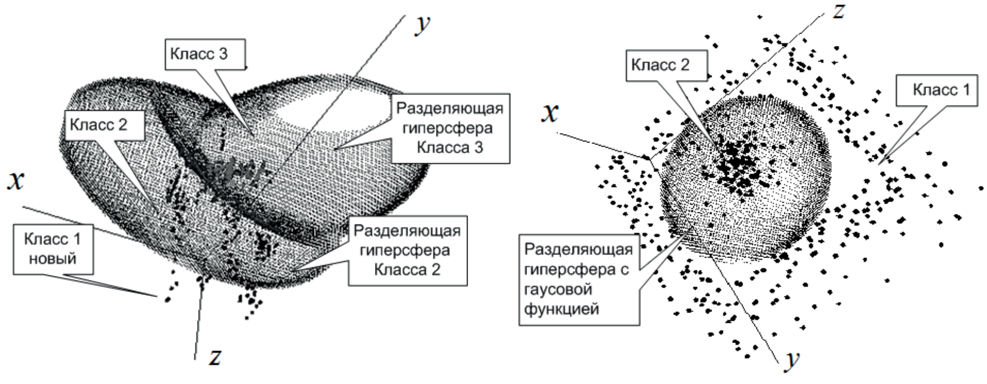

Рис. 4. Разделяющая гиперсфера, построенная при помощи радиальной функций ядра

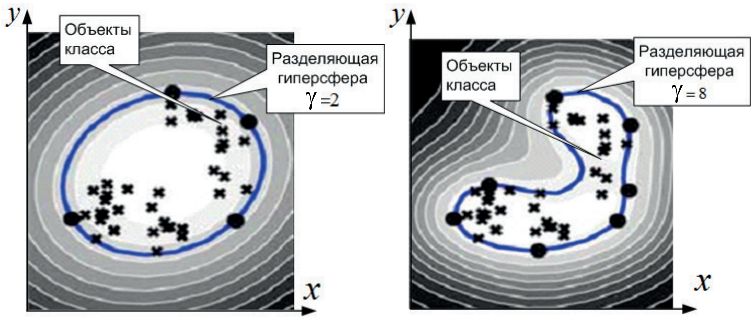

Рис. 5. Разделяющая гиперсфера, построенная при помощи радиальной функций ядра для $\gamma = 2$ (слева) и $\gamma = 8$ (справа)

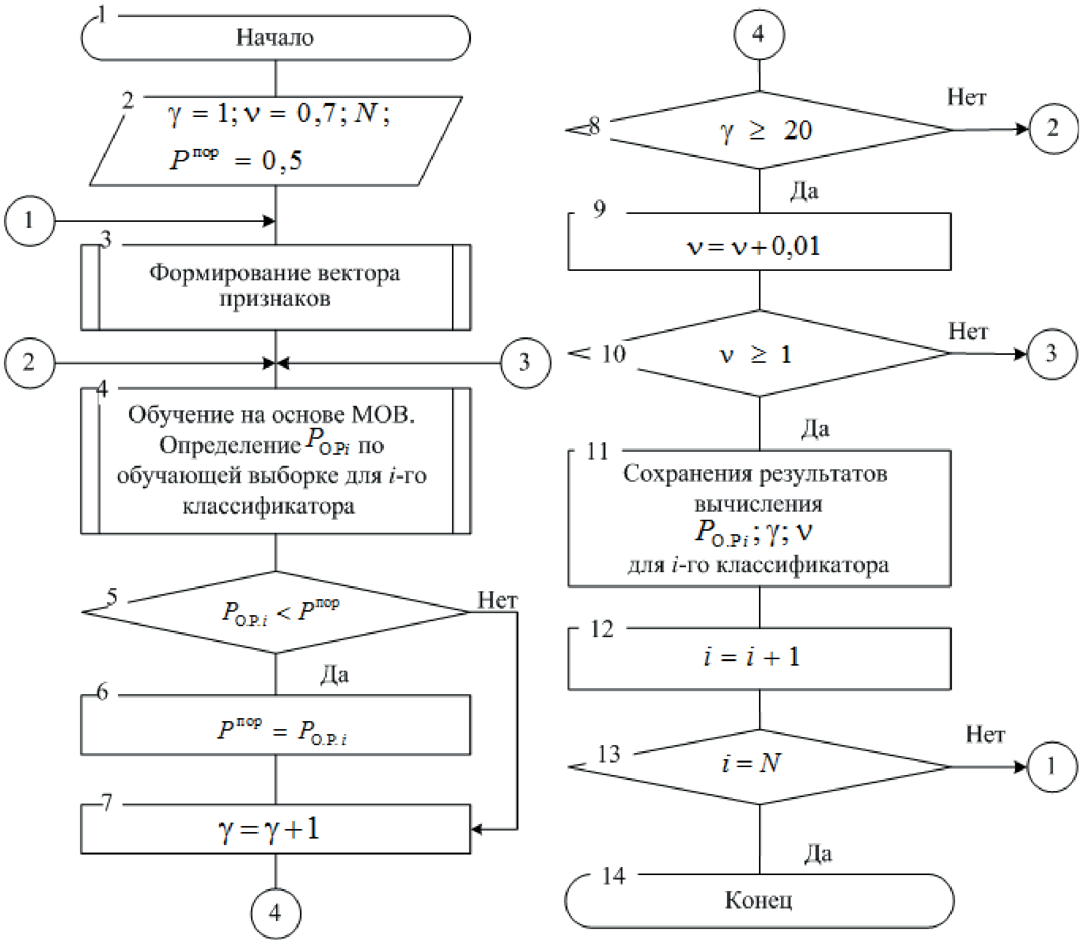

Рис. 6. Блок-схема алгоритма определения параметров классификаторов по критерию минимума ошибочных решений распознавания

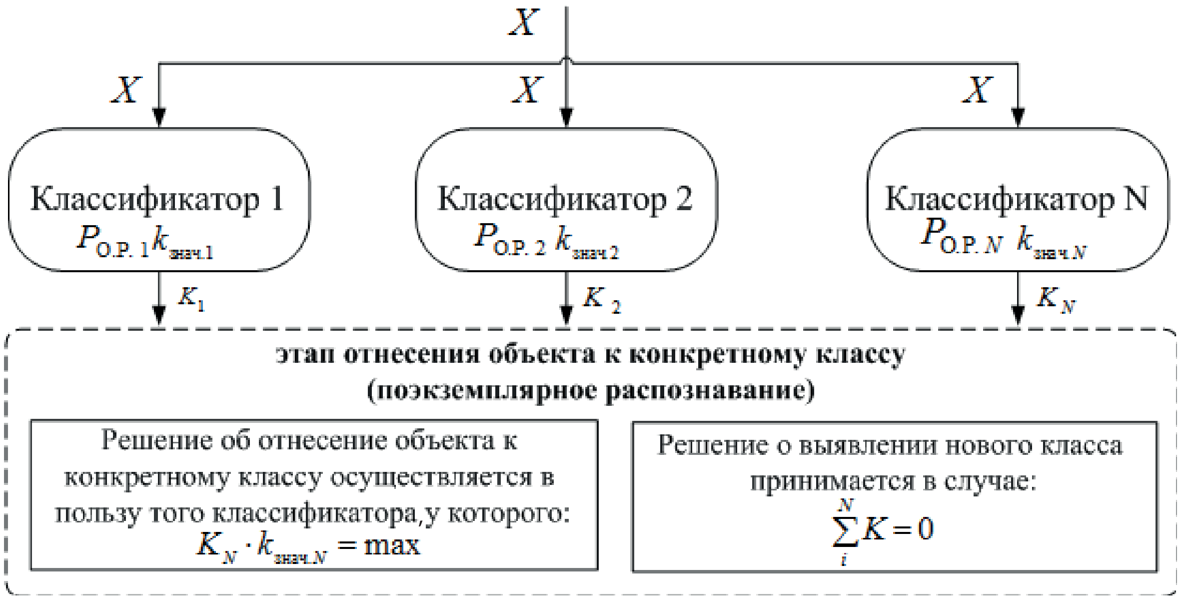

Рис. 7. Общая схема алгоритма поэкземплярного распознавания

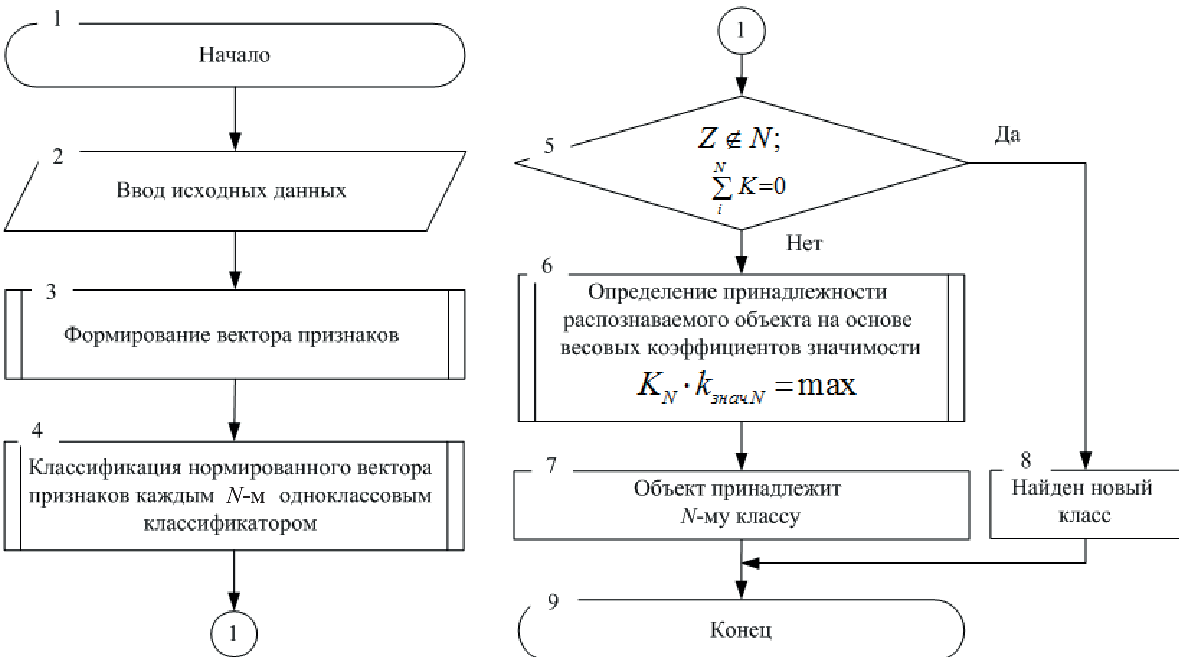

Рис. 8. Блок-схема алгоритма поэкземплярного распознавания

$$k_{\text{знач.}i} = 1 - \frac{P_{\text{О.Р.}i}}{\sum_{i=1}^{N} P_{\text{О.Р.}i}}, \quad (14)$$

где N – количество используемых классификаторов.

Весовые коэффициенты значимости являются вспомогательными коэффициентами, которые будут использованы на этапе поэкземплярного распознавания для разрешения противоречий в случае возникновения ситуаций срабатывания двух и более одноклассовых классификаторов (рис. 7).

Исходными данными для работы алгоритма поэкземплярного распознавания (рис. 8) являются данные, полученные в процессе обучения классификаторов. Сформированный вектор признаков классифицируется N обученными на предварительном этапе классификаторами одноклассового МОВ путем вычисления значения решающей функции (12). По результатам классификации каждый из классификаторов вычисляет значение коэффициента принадлежности вектора признаков к одному из N заданных классов $K_i = \{0,1\}$, при этом, когда классификатор делает вывод о том, что вектор принадлежит данному классу $f(\bar{x}) \geq 0$, он присваивает коэффициенту $K_i = 1$, в противном случае – $K_i = 0$.

Если значение всех

$$\sum_{i}^{N} K_i = 0,$$

то делается вывод о выявлении нового объекта (неизвестный класс). В случае отнесения объекта X к нескольким классам принятие решения о принадлежности вектора признаков к конкретному классу выполняется с использованием весовых коэффициентов значимости, полученных на этапе обучения классификаторов.

При этом вектор признаков относится в пользу того класса, произведение коэффициентов значимости и принадлежности у которого максимально:

$$K_i \cdot k_{\text{знач.}i} = \max. \quad (15)$$

Таким образом, обоснованно преимущество использования классификатора, построенного на основе метода опорных векторов, заключающееся в возможности выявления новых объектов, а также использования линейного классификатора для работы с линейно не разделимыми данными за счет радиальной функции ядра, которая позволяет строить замкнутые гиперсферы. Предложен алгоритм настройки параметров классификаторов, позволяющий по сформированной заранее обучающей выборке и заданном количестве классов распознавания получать значения параметров классификатора и используемого радиального ядра, соответствующие минимальному значению вероятности ошибочного распознавания.

Литература

1. *D. Tax and R. Duin.* Support vector data description. Machine Learning, 54:45–66, 2004.
2. *H. Xu and D.-S. Huang.* One class support vector machines for distinguishing photographs and graphics. In IEEE International Conference on Networking, Sensing and Control. – 2008. – P. 602–607.
3. *Кузнецов Ю.Н., Кузубов В.И., Волощенко А.Б.* Математическое программирование. – М.: «Высшая школа», 1980. – 300 с.
4. *Лепский А.Е., Броневич А.Г.* Математические методы распознавания образов : курс лекций. – Таганрог: Изд-во ТТИ ЮФУ, 2009. – 155 с.