УДК 681.513

РАЗРАБОТКА И ИССЛЕДОВАНИЕ ДВУХУРОВНЕВЫХ НЕПАРАМЕТРИЧЕСКИХ СИСТЕМ КЛАССИФИКАЦИИ*

А. В. Лапко, В. А. Лапко

Учреждение Российской академии наук
Институт вычислительного моделирования Сибирского отделения РАН,
630036, г. Красноярск, Академгородок, 50, стр. 44
E-mail: lapko@icm.krasn.ru

Предлагается методика синтеза двухуровневых непараметрических систем распознавания образов, основанных на принципах декомпозиции обучающих выборок по их объему и использовании технологии параллельных вычислений. Проводится анализ их свойств по результатам вычислительных экспериментов.

Ключевые слова: непараметрическая статистика, системы распознавания образов, декомпозиция обучающей выборки, асимптотические свойства.

Введение. Использование непараметрических классификаторов, основанных на оценках плотности вероятности типа Розенблатта — Парзена [1, 2], является одним из активно развивающихся направлений теории распознавания образов. Однако при усложнении условий классификации появляются методические и вычислительные трудности применения традиционных непараметрических алгоритмов распознавания образов, что наблюдается особенно при обработке статистических данных большого объема. Подобные условия часто встречаются при анализе аэрокосмической информации, исследовании медико-биологических, экологических и экономических процессов.

Перспективное направление «обхода» возникающих проблем состоит в применении принципов декомпозиции обучающих выборок по их объему и последовательных процедур формирования решений.

Цель данной работы — на основе анализа асимптотических свойств непараметрической оценки смеси плотностей вероятности обосновать эффективность использования принципов декомпозиции при решении задач распознавания образов в условиях больших выборок и разработать методику синтеза многоуровневых непараметрических систем классификации, обеспечивающих применение параллельных вычислительных технологий.

Непараметрическая оценка смеси плотностей вероятности и ее свойства. Пусть $V = (x^i,\ i = \overline{1,\ n})$ — выборка из n независимых наблюдений одномерной случайной величины x с плотностью вероятности $p(x)$, вид которой априори неизвестен.

Разобьем выборку V на T групп наблюдений: $V_j = (x^i, i \in I_j)$, $j = \overline{1,T}$. Здесь I_j — множество номеров наблюдений x, составляющих j-ю группу, причем $\bigcup_{j=1}^{T} I_j = I = (i = \overline{1,n})$ и $V_j \cap V_\lambda = \emptyset,\ j, \lambda = \overline{1,T},\ j \neq \lambda$.

На основе каждой выборки V_j построим непараметрические оценки плотности вероятности [1]

$$\bar{p}_j(x) = \frac{1}{n_j c_j} \sum_{i\,\in\, I_j} \Phi\Big(\frac{x - x^i}{c_j}\Big), \quad j = \overline{1,T}, \tag{1}$$

*Работа выполнена при поддержке Российского фонда фундаментальных исследований (грант № 07-01-00006).

где $\Phi(\cdot)$ — ядерные функции, удовлетворяющие условиям нормированности, положительности и симметричности; $n_j = |I_j|$ — количество элементов множества I_j; $c_j = c(n_j)$ — коэффициенты размытости ядерных функций, значения которых убывают с ростом n_j.

В качестве приближения $p(x)$ по статистической выборке V примем смесь непараметрических оценок плотностей вероятности типа (1)

$$\bar{p}(x) = \frac{1}{T}\sum_{j=1}^{T}\bar{p}_j(x). \tag{2}$$

Статистика (2) допускает использование технологии параллельных вычислений при оценивании плотности вероятности в условиях больших выборок.

Асимптотические свойства $\bar{p}(x)$ определяются следующим утверждением.

Теорема. Пусть $p(x)$ и первые две ее производные ограничены и непрерывны; ядерные функции $\Phi\,(u)$ удовлетворяют условиям

$$\Phi(u) = \Phi(-u), \qquad 0 \le \Phi(u) < \infty,$$
$$\int \Phi(u)du = 1, \qquad \int u^2\Phi(u)du = 1,$$
$$\int u^m\Phi(u)du < \infty, \quad 0 \le m < \infty;$$

последовательности $c_j = c_j(n_j)$ коэффициентов размытости ядерных функций таковы, что при $n_j \to \infty$ значения $c_j \to 0$, а $n_j c_j \to \infty$, $j = \overline{1,T}$. Здесь и далее бесконечные пределы интегрирования опускаются.

Тогда при конечных значениях T непараметрическая оценка (2) плотности вероятности $p(x)$ обладает свойством асимптотической несмещенности и состоятельности.

Доказательство.

1. По определению

$$\mathrm{M}(\bar{p}(x)) = \frac{1}{T}\sum_{j=1}^{T}\mathrm{M}(\bar{p}_j(x)) =$$
$$= \frac{1}{T}\sum_{j=1}^{T}\frac{1}{c_j}\int \Phi\Big(\frac{x-t}{c_j}\Big)p(t)dt = \frac{1}{T}\sum_{j=1}^{T}\int \Phi(u)p(x-c_ju)du,$$

где М — знак математического ожидания.

Разложив $p(x-c_ju)$ в ряд Тейлора и ограничиваясь первыми двумя членами ряда, имеем

$$W_1 = \mathrm{M}(\bar{p}(x) - p(x)) \sim \frac{p^{(2)}(x)}{2}\,\frac{1}{T}\sum_{j=1}^{T}c_j, \tag{3}$$

где $p^{(2)}(x)$ — вторая производная плотности вероятности $p(x)$ по x.

Отсюда при выполнении условий $c_j \to 0$, если $n_j \to \infty$, следует свойство асимптотической несмещенности статистики (2).

2. Рассмотрим выражение

$$\mathrm{M}\int(\bar{p}(x)-p(x))^2dx =$$

$$= \frac{1}{T^2}\,\mathrm{M}\left(\sum_{j=1}^{T}\int(p(x)-\bar{p}_j(x))^2dx + \sum_{j=1}^{T}\ \sum_{\substack{t=1\\ t\neq j}}^{T}\ \int(p(x)-\bar{p}_j(x))(p(x)-\bar{p}_t(x))dx\right). \qquad (4)$$

С учетом результатов исследования асимптотических свойств традиционной непараметрической оценки Розенблатта — Парзена [2] выражение (4) при достаточно больших значениях n_j, $j=\overline{1,\,T}$, принимает вид

$$\mathrm{M}\int(\bar{p}(x)-p(x))^2dx \sim$$

$$\sim \frac{1}{T^2}\left(\sum_{j=1}^{T}\Big(\frac{\|\Phi(u)\|^2}{n_jc_j}+\frac{c_j^4\|p^{(2)}(x)\|^2}{4}\Big)+\frac{\|p^{(2)}(x)\|^2}{4}\sum_{j=1}^{T}\ \sum_{\substack{t=1\\ t\neq j}}^{T}\ c_j^2c_t^2\right), \qquad (5)$$

где $\|\Phi(u)\|^2=\displaystyle\int\Phi^2(u)du$.

Нетрудно заметить, что при выполнении условий $c_j\to 0$ и $n_jc_j\to\infty$ при $n_j\to\infty$, $j=\overline{1,T}$, оценка плотности вероятности (2) сходится к $p(x)$ в среднеквадратическом, а с учетом свойств асимптотической несмещенности является состоятельной.

Сравнение аппроксимационных свойств статистики $\bar{p}(x)$ и оценки плотности вероятности Розенблатта — Парзена $\tilde{p}(x)$. Для получения аналитически значимых результатов примем $n_j=n/T$, $j=\overline{1,T}$. В этом случае можно полагать, что $\bar{c}=\bar{c}(n_j)=\bar{c}(n/T)$, $j=\overline{1,T}$.

На этой основе сравним отношения асимптотических смещений, среднеквадратических отклонений и дисперсий непараметрических оценок $\bar{p}(x)$ и $\tilde{p}(x)$ при оптимальных значениях коэффициентов размытости ядерных функций.

Асимптотическое выражение смещения для непараметрической оценки Розенблатта — Парзена $\tilde{p}(x)$ имеет вид [2]

$$W_2=\mathrm{M}(\tilde{p}(x)-p(x))\sim\frac{p^{(2)}(x)}{2}\,c^2(n),$$

а для оценки $\bar{p}(x)$ определяется выражением (3).

Вычислим отношение W_2/W_1 при оптимальных значениях коэффициентов размытости

$$c^*=\Big[\frac{\|\Phi(u)\|^2}{n\|p^{(2)}(x)\|^2}\Big]^{1/5},\quad \bar{c}^*=\Big[\frac{T\|\Phi(u)\|^2}{n\|p^{(2)}(x)\|^2}\Big]^{1/5}$$

для оценок $\tilde{p}(x)$ и $\bar{p}(x)$ соответственно.

После несложных преобразований получим

$$\frac{W_2}{W_1} = \frac{1}{T^{2/5}} < 1,$$

т. е. смещение статистики $\bar{p}(x)$ больше смещения $\tilde{p}(x)$.

При разбиении исходной выборки на T групп точек одинакового объема n/T выражение (5) при оптимальных коэффициентах $\bar{c}^*$ преобразуется к виду

$$W_3 \sim \frac{\|\Phi(u)\|^2}{n\bar{c}} + \frac{\bar{c}^4\|p^{(2)}(x)\|^2}{4} = \Big(\Big(\frac{\|\Phi(u)\|^2}{n}\Big)^4 \|p^{(2)}(x)\|^2\Big)^{1/5}\Big(\frac{4+T}{4T^{1/5}}\Big). \qquad (6)$$

Заметим, что при T=1 выражение (6) совпадает со среднеквадратическим отклонением W_4 для традиционной оценки Розенблатта — Парзена:

$$W_4 \sim \frac{5}{4}\Big(\Big(\frac{\|\Phi(u)\|^2}{n}\Big)^4 \|p^{(2)}(x)\|^2\Big)^{1/5}.$$

Рассмотрим отношение

$$\frac{W_4}{W_3} = \frac{5T^{1/5}}{4+T} \le 1,$$

которое подтверждает несколько большую эффективность в среднеквадратическом непараметрической оценки $\tilde{p}(x)$ по сравнению с предлагаемой $\bar{p}(x)$.

Однако дисперсия оценки $\bar{p}(x)$ меньше, чем для традиционной статистики $\tilde{p}(x)$, в чем нетрудно убедиться, если сравнить их дисперсионные составляющие

$$W_5 = \frac{\|\Phi(u)\|^2}{n\bar{c}^*}, \quad W_6 = \frac{\|\Phi(u)\|^2}{nc^*}$$

в среднеквадратических отклонениях $\bar{p}(x)$ и $\tilde{p}(x)$.

Их отношение

$$\frac{W_6}{W_5} = \frac{\bar{c}^*}{c^*} = T^{1/5} \ge 1.$$

Синтез непараметрической системы классификации для решения двухальтернативной задачи распознавания образов. Пусть $V = (x^i, \sigma(x^i), i = \overline{1,n})$ — обучающая выборка объема n, составленная из признаков $x^i = (x^i_j, j = \overline{1,k})$ классифицируемых объектов и соответствующих «указаний учителя» об их принадлежности к одному из двух классов:

$$\sigma(x^i) = \begin{cases} -1, & \text{если } x^i \in \Omega_1, \\ 1, & \text{если } x^i \in \Omega_2. \end{cases}$$

Условные плотности вероятностей распределения признаков x в области определения классов неизвестны. Причем объем выборки достаточно большой, что снижает вычислительную эффективность традиционных непараметрических классификаторов.

При решении задачи распознавания образов в данных условиях предлагаемый подход состоит из следующих действий:

1. Осуществление декомпозиции исходной выборки V на однородные части

$$V_j(1) = (x^i,\, \sigma(x^i),\, i \in I_j), \quad j = \overline{1,T}.$$

Основанием для выполнения данной операции является формирование выборок $V_j(1)$ в соответствии с их принадлежностью различным временным интервалам наблюдения за признаками классифицируемых объектов, разбиение выборки на группы по уровню заполненных пропусков данных либо использование случайной стратегии при декомпозиции выборки.

2. Построение решающих правил по полученным данным:

$$\bar{m}_j(x) : \begin{cases} x \in \Omega_1, & \text{если } \bar{f}_{12}^j(x) \le 0, \\ x \in \Omega_2, & \text{если } \bar{f}_{12}^j(x) > 0,\, j = \overline{1,T}, \end{cases} \tag{7}$$

где непараметрические оценки уравнений разделяющих поверхностей определяются выражением

$$\bar{f}_{12}^j(x) = \frac{1}{n_j \prod\limits_{v=1}^{k} c_v} \sum_{i \in I_j} \sigma(x^i) \prod_{v=1}^{k} \Phi\Big(\frac{x_v - x_v^i}{c_v}\Big), \quad j = \overline{1,T}. \tag{8}$$

Оптимизация частных решающих правил (7) по коэффициентам размытости $c_v, v = \overline{1,k}$, осуществляется в режиме «скользящего экзамена» из условия минимума статистической оценки вероятности ошибки распознавания образов:

$$\bar{\rho}_j = \frac{1}{n_j} \sum_{t \in I_j} 1(\sigma(x^t),\, \bar{\sigma}(x^t)),\ j = \overline{1,T}, \quad 1(\sigma(x^t),\, \bar{\sigma}(x^t)) = \begin{cases} 0, & \text{если } \sigma(x^t) = \bar{\sigma}(x^t), \\ 1, & \text{если } \sigma(x^t) \ne \bar{\sigma}(x^t), \end{cases}$$

где $\bar{\sigma}(x^t)$ — «решение» алгоритма (7) о принадлежности ситуации x^t к одному из двух классов. При формировании решения $\bar{\sigma}(x^t)$ ситуация x^t исключается из процесса обучения в непараметрической статистике (8).

3. Формирование обучающей выборки второго уровня распознающей системы на основе непараметрической оценки решающих функций (8):

$$V(2) = (\bar{f}_{12}^j(x^i),\ j = \overline{1,T},\ \sigma(x^i),\ i \in I(2)),$$

где $I(2)$ — множество номеров ситуаций из исходных данных V, составляющих выборку $V(2)$. В качестве подобной контрольной выборки может быть принята одна из выборок V_j, $j = \overline{1,T}$, пользующаяся наибольшим доверием исследователя.

4. Построение решающего правила в пространстве значений $\bar{f}_{12}(x) = (\bar{f}_{12}^j(x),\, j = \overline{1,T})$ по полученной выборке $V(2)$:

$$\bar{m}(\bar{f}_{12}(x)) : \begin{cases} x \in \Omega_1, & \text{если } \bar{F}_{12}(\bar{f}_{12}(x)) \le 0, \\ x \in \Omega_2, & \text{если } \bar{F}_{12}(\bar{f}_{12}(x)) > 0, \end{cases} \tag{9}$$

где непараметрическая оценка обобщенной решающей функции имеет вид

$$\bar{F}_{12}(\bar{f}_{12}(x)) = \frac{1}{n(2) \prod\limits_{j=1}^{T} c_j} \sum_{i \in I(2)} \sigma(x^i) \prod_{j=1}^{T} \Phi\Big(\frac{\bar{f}_{12}^j(x) - \bar{f}_{12}^j(x^i)}{c_j}\Big).$$

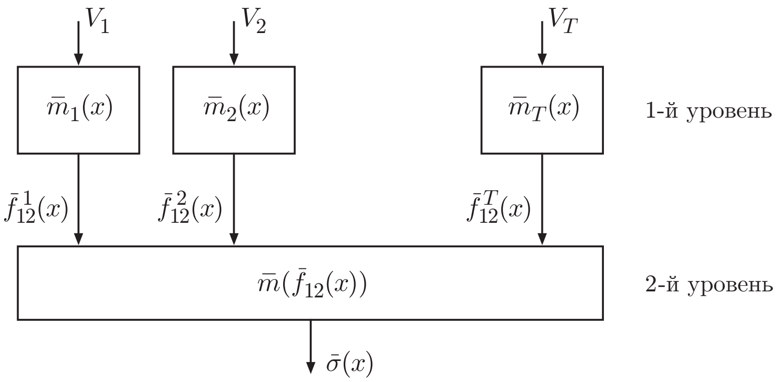

Рис. 1

Структура предлагаемой непараметрической распознающей системы, основанной на декомпозиции обучающей выборки по ее объему, представлена на рис. 1.

На первом уровне структуры системы классифицируемая ситуация x преобразуется в значения непараметрических оценок $\bar{f}^j_{12}(x)$, $j = \overline{1, T}$, в пространстве которых правилом (9) принимается решение $\bar{\sigma}(x)$ о принадлежности ситуации x к тому или иному классу.

Методика синтеза непараметрической системы для решения многоальтернативной задачи распознавания образов. Пусть $V = (x^i,\, \sigma(x^i),\ i = \overline{1, n})$ — обучающая выборка объема n, составленная из признаков $x^i = (x^i_j,\, j = \overline{1, k})$ классифицируемых объектов и соответствующих указаний учителя $\sigma(x^i)$ об их принадлежности к одному из M классов.

Предлагаемый подход при решении задачи распознавания образов в данных условиях выполняет следующие действия:

1. Осуществление декомпозиции исходной выборки V на однородные части

$$V_j = (x^i,\, \sigma(x^i),\, i \in I_j),\, j = \overline{1, T},$$

где I_j — множество номеров ситуаций их V, составляющих j-ю группу.

2. Построение решающих правил по полученным данным:

$$\bar{m}_j(x):\ x \in \Omega_t, \quad \text{если } \bar{p}^j_t(x) = \max_{\gamma=\overline{1, M}} \bar{p}^j_\gamma(x), \quad j = \overline{1,\, T}, \tag{10}$$

где непараметрические оценки плотности вероятности $\bar{p}^j_\gamma(x)$ распределения значений признаков x в классах Ω_γ, $\gamma = \overline{1, M}$, определяются выражением

$$\bar{p}^j_\gamma(x) = \frac{1}{|I_j(\gamma)| \prod\limits_{v=1}^{k} c_v(j)} \sum_{i \in I_j(\gamma)} \prod_{v=1}^{k} \Phi\Big(\frac{x_v - x^i_v}{c_v(j)}\Big), \quad \gamma = \overline{1, M},\ j = \overline{1, T}. \tag{11}$$

Здесь $I_j(\gamma)$ — множество номеров точек класса Ω_γ в выборке V_j, а $|I_j(\gamma)|$ — их количество.

Оптимизация частных решающих правил (10) по коэффициентам размытости $c_v(j)$, $v = \overline{1, k}$, $j = \overline{1, T}$, ядерных функций осуществляется в режиме скользящего экзамена из условия минимума статистической оценки вероятности ошибки распознавания образов [3].

3. Формирование решающего правила второго уровня структуры системы классификации:

$$\bar{m}(x): \ x \in \Omega_t, \qquad \text{если } \bar{\bar{p}}_t(x) = \max_{\gamma=\overline{1,M}} \bar{\bar{p}}_\gamma(x),$$

где $\bar{\bar{p}}_\gamma(x)$ — оценка смеси плотностей вероятности:

$$\bar{\bar{p}}_\gamma(x) = \frac{1}{T}\sum_{j=1}^{T} \bar{p}_\gamma^j(x), \quad \gamma = \overline{1,M}.$$

Сравним вычислительную эффективность предложенной двухуровневой системы классификации и традиционного непараметрического алгоритма. Будем полагать, что объем частных выборок V_j, $j = \overline{1,T}$, одинаков и равен n/T, время расчета одной ядерной функции составляет τ.

Максимальное время, необходимое для принятия решения традиционным непараметрическим алгоритмом и предлагаемой системой при использовании технологии параллельных вычислений, составляет $t_{\text{тр}} \approx nk\tau$ и $t_{\text{п}} \approx n\tau k/T$. Тогда их отношение $t_{\text{тр}}/t_{\text{п}} = T$, т. е. вычислительная эффективность предлагаемой двухуровневой системы классификации возрастает по мере роста T.

Исследование свойств двухуровневой непараметрической системы распознавания образов. На основании данных вычислительного эксперимента сравнивались эффективности предлагаемой системы классификации и хорошо зарекомендовавшего себя на практике традиционного непараметрического алгоритма распознавания образов [3].

Традиционный непараметрический алгоритм основывался на решающей функции типа (10) и строился по полной обучающей выборке $(x_v^i,\ v = \overline{1,k},\ \sigma(i),\ i = \overline{1,n})$.

Исследования осуществлялись при решении многоальтернативной задачи распознавания образов Ω_j, $j = \overline{1,\,5}$, в k-мерном пространстве признаков ($k = \overline{2,\,10}$). Законы распределения признаков в области первого класса формировались в соответствии с датчиком случайных чисел

$$x_v = m + \sigma_1\Big(\sum_{i=1}^{p} \varepsilon^i - 0{,}5p\Big)\frac{6}{\sqrt{3p}}, \quad v = \overline{1,k},$$

при $p = 5$; $m = 3$; $\varepsilon \in [0,\ 1]$ — случайная величина с равномерным законом распределения; $\sigma_1 = 0{,}6$ или $\sigma_1 = 0{,}9$ соответственно для нечетных и четных значений $v = \overline{1,k}$.

Значения признаков второго и третьего классов генерировались с использованием датчиков случайных чисел:

$$\Omega_2: \ x_{v+1} = (x_v)^2 - 6x_v + 10 + \sigma_2\Big(\sum_{i=1}^{p} \varepsilon^i - 0{,}5p\Big)\frac{6}{\sqrt{3p}},$$

$$\Omega_3: x_{v+1} = -(x_v)^2 + 6x_v - 3{,}5 + \sigma_2\Big(\sum_{i=1}^{p_1} \varepsilon^i - 0{,}5p_1\Big)\frac{6}{\sqrt{3p_1}},$$

где $x_v = a + \varepsilon(b - a)$ ($v \in I_{\text{н}}$ — множество нечетных чисел, меньших k, $a = 1{,}5$, $b = 4{,}5$); $\sigma_2 = 0{,}7$; $p = 5$; $p_1 = 12$.

Значения признаков четвертого и пятого классов формировались на основе датчиков случайных чисел:

$$\Omega_4 : x_{v+1} = -(x_v)^2 + 6x_v - 1{,}8 + \sigma_2\Big(\sum_{i=1}^{p_1} \varepsilon^i - 0{,}5p_1\Big)\frac{6}{\sqrt{3p_1}},$$

$$\Omega_5 : x_{v+1} = (x_v)^2 - 6x_v + 8{,}2 + \sigma_2\Big(\sum_{i=1}^{p_1} \varepsilon^i - 0{,}5p_1\Big)\frac{6}{\sqrt{3p_1}},$$

где $x_v = a + \varepsilon(b - a)$ ($v \in I_{\text{н}}$, $a = 1{,}2$, $b = 4{,}8$).

Априорные вероятности классов Ω_j, $j = \overline{1, 5}$, считались одинаковыми.

Вычислительный эксперимент при фиксированных условиях исследований осуществлялся 10 раз, полученные результаты расчетов усреднялись. Для оценивания вероятностей ошибок распознавания образов использовалась контрольная выборка объемом $N = 5000M$ (здесь M — количество классов), что обеспечивало представительность получаемых результатов.

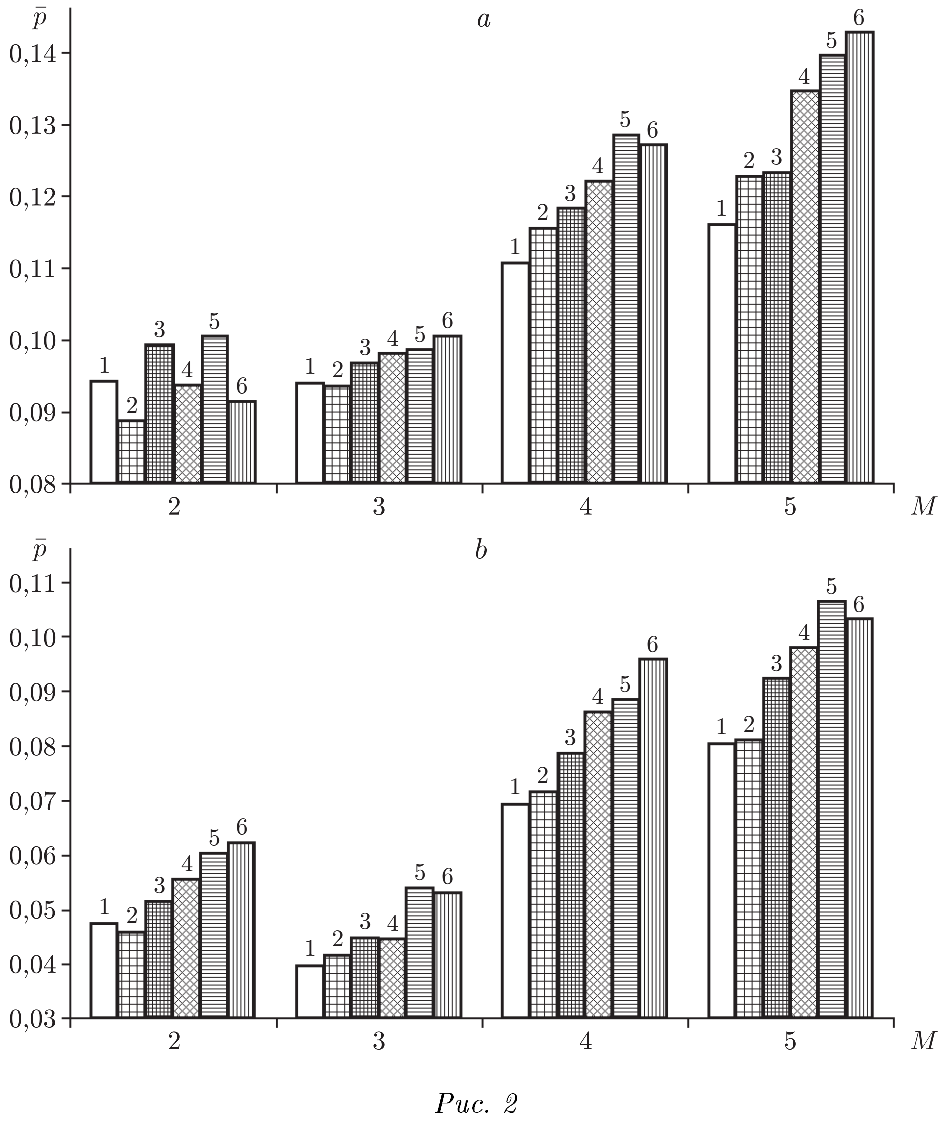

Рис. 2

Анализ результатов вычислительного эксперимента показал, что законы распределения оценок вероятностей ошибок классификации сравниваемых алгоритмов распознавания образов близки к симметричным. Поэтому средние значения оценок исследуемых показателей эффективности соответствуют их наиболее вероятным значениям.

На рис. 2 приведены зависимости средних значений оценок вероятности ошибки распознавания образов $\bar{p}$ традиционного (столбцы 1) и двухуровневых непараметрических классификаторов (столбцы 2–6) от количества классов M решаемой задачи распознавания образов в пространстве признаков размерности $k = 4$ (рис. 2, *a*), $k = 10$ (рис. 2, *b*).

Наблюдается повышение средних значений оценок вероятностей ошибок распознавания образов сравниваемых классификаторов при увеличении количества классов M, что объясняется естественным ростом сложности границ между ними (см. рис. 2, *a*). В данных условиях при фиксированном объеме обучающей выборки ($n = 400M$) появляются трудности в построении непараметрических решающих функций.

Использование дополнительных признаков сохраняет отмеченную тенденцию при относительно меньших значениях средних оценок вероятностей ошибки распознавания образов (см. рис. 2, *b*).

Существуют условия классификации, когда эффективность сравниваемых непараметрических алгоритмов сопоставима. Это особенно характерно для двух- и трехальтернативных задач распознавания образов: на протяжении диапазонов изменения $k \in [4; 10]$ и $T \in [2; 10]$, при которых расхождение между оценками вероятностей ошибок распознавания образов не превышает значений 0,010–0,015. В многоальтернативном случае при $M > 3$ подобные расхождения соответствуют значениям $T < 8$.

Заключение. На основе анализа асимптотических свойств непараметрической оценки смеси плотностей вероятности в предлагаемой работе обоснована возможность декомпозиции исходных статистических данных при синтезе непараметрических статистик в условиях больших выборок. С этих позиций предложена методика построения двухуровневых непараметрических систем распознавания образов, характеризующихся высокой вычислительной эффективностью.

Преимущество данной системы перед традиционным непараметрическим алгоритмом распознавания образов парзеновского типа заключается в возможности использования технологий параллельных вычислений, что позволяет в T раз сократить время формирования решений при соизмеримых ошибках классификации. Показано существование достаточно широких условий классификации, при которых оценки вероятностей ошибки распознавания образов сравниваемых алгоритмов практически не отличаются.

СПИСОК ЛИТЕРАТУРЫ

1. **Parzen E.** On estimation of a probability density function and mode // Ann. Math. Statist. 1962. **33**, N 3. P. 1065–1076.

2. **Епанечников В. А.** Непараметрическая оценка многомерной плотности вероятности // Теория вероятностей и ее применения. 1969. **14**, Вып. 1. С. 156–161.

3. **Лапко А. В., Лапко В. А., Соколов М. И., Ченцов С. В.** Непараметрические системы классификации. Новосибирск: Наука, 2000. 240 с.

Поступила в редакцию 15 апреля 2008 г.