УДК 004.93

ФОРМИРОВАНИЕ СЛОЖНЫХ ПРИЗНАКОВЫХ ПРОСТРАНСТВ R-ГО РАНГА, ОБЕСПЕЧИВАЮЩИХ КАЧЕСТВО И НАДЕЖНОСТЬ РАСПОЗНАВАНИЯ

[1]***Бекмуратов К.А.***, [2]***Ахатов А.Р.***, [1]***Бекмуратов Д.К.***

bekmurodov1958@mail.ru; akmalar@rambler.ru; bekmurodov_d@mail.ru

[1]Самаркандский филиал Ташкентского университета информационных технологий, Самарканд, Ибн Сино 2; [2]Самаркандский государственный университет, Самарканд, Университетский бульвар 15

Рассматривается решение задачи определения простых признаков, присущих конкретному классу из заданного свойства объектов эталонной выборки, и формирования сложных признаковых пространств из выделенных простых признаков. Приводится методика выбора простых признаков из исходных свойств и нахождения предельно-допустимой размерности пространства сложных признаков. Указаны основные параметры, которые играют существенную роль в процессе нахождения предельно-допустимой размерности пространства сложных признаков: вероятность ошибочной классификации объектов; вероятность ошибки, достигаемой с заданной надежностью; количество объектов эталонной выборки и количество признаков, выделенных для конкретного класса. Определены минимальные и реальные разделяющие силы сложных признаков на основе найденной предельно-допустимой размерности пространства. Предложены принципы формирования системы сложных признаков первого и второго типов для каждого класса из простых признаков второго и третьего типов с учетом минимальной и реальной разделяющей силой. На основе предложенной методики разработан алгоритм и программное обеспечение. Проведены вычислительные эксперименты, результаты которых приведены в виде решающих правил, используемых для распознавания объектов. Также приведены выводы по проведенному исследованию в целом.

Ключевые слова: объект, признак, класс, эталонная выборка, контрольная выборка, объем эталонной выборки, системы сложных признаков, качество и надежность распознавания, вероятность ошибки, предельная размерность пространства, сумма голосов, решающее правило

Цитирование: *Бекмуратов К.А., Ахатов А.Р., Бекмуратов Д.К.* Формирование сложных признаковых пространств r-го ранга, обеспечивающих качество и надежность распознавания // Проблемы вычислительной и прикладной математики. — 2019. — № 1(19). — С. 24–38.

1 Введение

В работах [1, 2], связанных с проблемой распознавания, появилась гипотеза компактности, которая гласит: классам соответствуют компактные множества точек в пространстве свойств. Однако эту гипотезу не всегда удавалось подтвердить, но те задачи, в рамках которых она выполнялась, все без исключения, находили простое решение. В некоторых случаях, когда гипотеза не подтверждалась, тогда задача либо совсем не решалась, либо решалась с большим трудом.

Поэтому в [3] введена новая формулировка гипотезы компактности: если некоторое множество объектов представляет собой классы, то обязательно существует такое пространство, в котором объектам каждого класса соответствуют компактные множества точек. А это значит, что центр тяжести проблемы распознавания должен быть перенесен на синтез таких пространств, в которых классы компактны. Отсюда следует, что если заданы классы некоторым набором объектов с известной принадлежностью, т.е. (X_i, V_j), то возникает задача, состоящая в выявлении признаков, определяющих сходство между объектами одного и различие между объектами разных классов, т.е. задача синтеза такого пространства, в котором классы компактны. В нахождении таких признаков и состоит задача распознавания в излагаемой далее постановке.

Для обеспечения требуемого качества распознавания необходимо из эталонной выборки выбирать такие признаки, с которыми предстоит работать распознающей системе, а эталонную выборку (ЭВ) следует выбирать так, чтобы признаки в полной мере отразились на ЭВ.

Следовательно, распознающая система должна обладать не только требуемым качеством, но и надежностью его достижения. Задача построения распознающей системы сводится к тому, чтобы в процессе обучения по ЭВ было достигнуто требуемое качество с надежностью, не ниже требуемой при выделении конкретного класса от всех остальных классов.

С целью достижения требуемого качества распознающей системы производится предварительное сокращение объема ЭВ за счет количества признаков и объектов. Из первоначального алфавита признаков выбираются лишь несколько наиболее существенных простых признаков или их сочетания для конкретного класса. С выделением существенных простых признаков или их сочетаний одновременно сокращается количество объектов рассматриваемого класса за счет идентичных объектов. Сокращение признаков и объектов ЭВ считается полезным в двух аспектах: во-первых, снижается объем вычислений, а во-вторых, с удалением из ЭВ несущественных признаков и идентичных объектов повышается качество и надежность распознавания.

В [3–7] получены теоретические результаты, основанные на соотношениях, связывающих такие параметры, как количество объектов и признаков ЭВ, а также качество и надежность распознавания. В этих работах показано, что чем ниже размерность признакового пространства, тем меньше вероятность ошибочных ответов при распознавании.

Эти выводы можно рассматривать как обоснование основной цели настоящей статьи.

2 Постановка задачи

Существуют различные методы определения признаков $x_i \in V$ $(i = \overline{1, n})$ [8–14]. Однозначно отдать предпочтение одному из них невозможно. Методы, эффективно работающие для одной задачи, неэффективны для другой, вместе с тем, возможность их реализации зависит от имеющихся технических средств. Исходя из таких соображений, выбор определенного метода зависит от реальной задачи и конкретных практических возможностей. Более того, учитывается инвариантность многих неизвестных методов определения информативных наборов x_i $(i = overline 1, n)$ относительно существующих критериев информативности.

В связи с изложенным обстоятельством выбор критериев информативности для сокращения признакового пространства является весьма актуальным.

В [4] получен теоретический результат, смысл которого состоит в том, что если из N решающих правил выбирается одно, которое безошибочно разделяет ЭВ длины m, то с вероятностью $(1-\eta)$ можно утверждать, что вероятность ошибочной классификации с помощью этого правила составит величину, меньшую ε, где

$$\varepsilon = \frac{\ln N - \ln \eta}{m} \tag{1}$$

В работах [5–7, 15–18] до обучения определяются предельные значения размерности пространства n_0 ($n_0 < n$) с учетом ε, η и объема ЭВ (m – количество объектов, n – количество исходных свойств), а само пространство в процессе обучения формируется из простых признаков (ПП) двух типов относительно каждого V_q так, чтобы заранее выбранное решающее правило (РП) безошибочно разделило ЭВ.

В данной статье, в отличие от работ [5–7, 15–18], при определении ПП первого, второго и третьего типа относительно V_q вводится новое требование, т.е. в качестве ПП первого, второго и третьего типа относительно V_q учитывается не только присутствие, а одновременно, и отсутствие данного признака $x_i \in V$ $(i = \overline{1,n})$ на некоторых или на всех объектах. Далее из отобранных x_i $(i = \overline{1,n})$ второго и третьего типа соответственно формируются сложные признаки первого типа (СППТ) и сложные признаки второго типа (СПВТ) r-го ранга.

Пусть задана ЭВ $V = V_1, \ldots, V_r$ $(V_i \cap V_j = \varphi, i \neq j)$, где каждый объект $X_j^p \in V_p$ $(p = \overline{1,l};\ j = \overline{1,m_l})$ является n-мерным вектором числовых признаков, т.е. $X_j^p = \left(x_{j1}^p, x_{j2}^p, \ldots, x_{jn}^p\right)$ $(p = \overline{1,m};\ j = \overline{1,m_l};\ i = \overline{1,n})$. Обозначим через V_q любого класса $V_j \in V$, т.е.$V_q = \forall V_j$, а через V_p все остальные классы, кроме V_q, т.е. $V_p = V \setminus V_q$.

Требуется, используя V, определить ПП первого, второго и третьего типа относительно каждого класса $V_q \in V$, сформировать из определенных ПП системы пространств сложных признаков (СП) в пределах n_0 и построить в этом пространстве РП $R^q(X)$ относительно каждого $V_q \in V$, а также указать, используя РП, принадлежность новых объектов к $V_q \in V$ с вероятностью ошибки, не превышающей ε и достигаемой надежностью $(1-\eta)$.

3 Метод решения

Допустим, что на $X = \left(\widetilde{X}_l, \ldots, \widetilde{X}_m\right) \in V$ задано некоторое свойство x_i. Значения x_i могут быть бинарные и числовые. В качестве примера для бинарного x_i можно привести:

$$x_i = \begin{cases} 1, & \text{если объект } X \text{ обладает свойством } x_i; \\ 0, & \text{в противном случае,} \end{cases} \tag{2}$$

Если x_i числовое, то его можно кодировать

$$x_i = \begin{cases} 1, & \text{если } x_i \leqslant \mu_i, j = \overline{1,l}; \\ 0, & \text{в противном случае,} \end{cases} \tag{3}$$

или

$$x_i = \begin{cases} 1, & \text{если } \delta_{min}^i \leqslant x_i \leqslant \delta_{max}^i; \\ 0, & \text{иначе,} \end{cases}$$

где μ_i заданные или определяемые алгоритмом пороги; δ_{min}^i, δ_{max}^i определяются

$$\delta_{max}^i = \max(x_i),$$

$$\delta^i_{min} = \min(x_i).$$

Введем следующее обозначение

$$x_i = \begin{cases} x_i, & \text{если } x_i = 1; \\ \overline{x_i}, & \text{если } x_i = 0. \end{cases}$$

Если для $x_i \in V$ выполняется

$$\begin{aligned} \big[\big(\forall X \in V_q : x_i = 1\big) \wedge \big(\forall X \in V_p : x_i = 0\big)\big] \vee \\ \big[\big(\forall X \in V_q : x_i = 0\big) \wedge \big(\forall X \in V_p : x_i = 1\big)\big] = 1, \end{aligned} \tag{4}$$

то x_i считается ПП первого типа (ПППТ) относительно V_q. Обозначим через x^1_{qi} ПППТ.

Из (2) видно, что каждый отобранный x^1_{qi} относительно V_q самостоятельно участвует при распознавании объектов, так как в этом случае $\forall X \in V_q$ безошибочно разделяются от $\forall X \in V_p$. Поэтому x^1_{qi} не участвует в качестве ПП при формировании СП.

Если для $x_i \in V$ выполняется

$$\begin{aligned} \big[\big(\forall X \in V_q : x_i = 1\big) \wedge \big(\exists X \in V_p : x_i = 0\big)\big] \vee \\ \big[\big(\forall X \in V_q : x_i = 0\big) \wedge \big(\exists X \in V_p : x_i = 1\big)\big] = 1, \end{aligned} \tag{5}$$

то x_i считается ПП второго типа (ППВТ) относительно V_q. Обозначим через x^2_{qi} ППВТ.

Из (5) видно, что каждый отобранный x^1_{qi} не может безошибочно разделить $\forall X \in V_q$ от $\forall X \in V_p$, так как они правильно распознают $\forall X \in V_q$ и ошибаются только на $\forall X \in V_p$. Поэтому x^2_{qi} при распознавании объектов самостоятельно не используется, а в качестве ПП участвует при формировании СППТ.

Если для $x_i \in V$ выполняется

$$\begin{aligned} \big[\big(\exists X \in V_q : x_i = 1\big) \wedge \big(\forall X \in V_p : x_i = 0\big)\big] \vee \\ \big[\big(\exists X \in V_q : x_i = 0\big) \wedge \big(\forall X \in V_p : x_i = 1\big)\big] = 1, \end{aligned} \tag{6}$$

то x_i считается ПП третьего типа (ППТТ) относительно V_q. Обозначим через x^3_{qi} ППТТ.

Из (6) видно, что отобранный каждый x^3_{qi} не может безошибочно разделить $\forall X \in V_q$ от $\forall X \in V_p$, так как они правильно распознают $\forall X \in V_p$ и ошибаются только на $\forall X \in V_q$. Поэтому x^3_{qi} так же, как и x^2_{qi}, при распознавании объектов самостоятельно не используется, а в качестве ПП участвует при формировании СПВТ.

Пусть путем последовательного просмотра из $x_i \in V$ $(i = \overline{1,n})$ отобраны ПППТ $V^1_q = x^1_{q1}, x^1_{q2}, \ldots, x^1_{qn_1}$ $(n_1 \leqslant n)$ согласно (4), ППВТ $V^2_q = x^2_{q1}, x^2_{q2}, \ldots, x^2_{qn_2}$ $(n_2 \leqslant n)$ согласно (5) и ППТТ $V^3_q = x^3_{q1}, x^3_{q2}, \ldots, x^3_{qn_3}$ $(n_3 \leqslant n)$ согласно (6) относительно V_q. Из (4) видно, что каждый отобранный ПППТ $x^1_{qi} \in V^1_q$ безошибочно разделяет $\forall X \in V_q$ от $\forall X \in V_p$ и самостоятельно участвует при распознавании объектов. Поэтому СП формируется только из ПП типа $x^2_{qi} \in V^2_q$ и $x^3_{qi} \in V^3_q$, так как каждый отобранный $x^2_{qi} \in V^2_q$ и $x^3_{qi} \in V^3_q$ самостоятельно не обеспечивает безошибочного разделения $\forall X \in V_q$ от $\forall X \in V_p$.

Обозначим через $x_{qr}^{c(1)}$ СППТ и $x_{qr}^{c(2)}$ СПВТ r-го ранга относительно V_q, формируемые из $x_{qi}^2 \in V_q^2$ и $x_{qi}^3 \in V_q^3$ соответственно.

Теперь рассмотрим процедуры последовательного выбора $x_{qr}^{c(1)}$ и $x_{qr}^{c(2)}$ из $x_{qi}^2 \in V_q^2$ и $x_{qi}^3 \in V_q^3$. Согласно (5) и (6) каждый отобранный $x_{qi}^2 \in V_q^2$ $(i = \overline{1, n_2})$ и $x_{qi}^3 \in V_q^3$ $(i = \overline{1, n_3})$ в отдельности не разделяет безошибочно $\forall X \in V_q$ от $\forall X \in V_p$. Для того чтобы достичь безошибочного разделения $\forall X \in V_q$ от $\forall X \in V_p$ на V, из V_q^2 и V_q^3 формируются $x_{qr}^{c(1)}$ и $x_{qr}^{c(2)}$. При формировании $x_{qr}^{c(1)}$ и $x_{qr}^{c(2)}$ необходимо учитывать количество $x_{qi}^2 \in V_q^2$ $(i = \overline{1, n_2})$ и $x_{qi}^3 \in V_q^3$ $(i = \overline{1, n_3})$ входящих в состав $x_{qr}^{c(1)}$ и $x_{qr}^{c(2)}$. Для того чтобы установить ограничение на количества $x_{qi}^2 \in V_q^2$ $(i = \overline{1, n_2})$ и $x_{qi}^3 \in V_q^3$ $(i = \overline{1, n_3})$, отбираемых из V_q^2 и V_q^3, которые участвуют в составе $x_{qr}^{c(1)}$ и $x_{qr}^{c(2)}$, необходимо заранее определить предельно-допустимое значение размерности пространства n_0^1 и n_0^2 для набора СП $V_q^{c(2)} = x_{qr_1}^{c(1)}, x_{qr_2}^{c(1)}, \ldots, x_{qr_v}^{c(1)}$ $(v = n_0^1)$ и $V_q^c(2) = = x_{q1}^{c(2)}, x_{q2}^{c(2)}, \ldots, x_{qn_1}^{c(2)}$ $(v = n_0^2)$, а затем в пределах n_0^1 и n_0^2 наложить ограничение на ранг r каждого выбираемого $x_{qr}^{c(1)}$ и $x_{qr}^{c(2)}$, который формируется из V_q^2 и V_q^3. Если найденные $V_q^{c(1)}$ и $V_q^{c(2)}$ обеспечивают безошибочное разделение $\forall X \in V_q$ от $\forall X \in V_p$ в пространствах n_0^1 и n_0^2, то при распознавании объектов гарантируется требуемая вероятность ошибки ε и надежность $(1 - \eta)$ при заданных m, η, а также выбранных n_2, n_3. Для этого необходимо определить значения n_0^1 и n_0^2.

Допустим, что на V определены $x_{qi}^2 \in V_q^2$ $(i = \overline{1, n_2})$ относительно V_q, для которых выполняются (5). Предположим, что до процесса обучения заранее заданы ε и η. Тогда из (1) можно получить функциональную зависимость

$$\ln N = \varepsilon m + \ln \eta. \tag{7}$$

Согласно (5) $\forall X \in V_q$ считается одинаковыми и в качестве представительного выбирается любой объект. Тогда в (7) m заменяется на $m^* = (m - m_q) + 1$ (где m_q – количество объектов класса V_q).

Учитывая $m^* = (m - m_q) + 1$ из (7), получим

$$\ln N = \varepsilon((m - m_q) + 1) + \ln \eta. \tag{8}$$

Если выбраны $x_{qi}^2 \in V_q^2$ $(i = \overline{1, n_2})$, то число всевозможных $x_{qr}^{c(1)}$, получаемых из $x_{qi}^2 \in V_q^2$ $(i = \overline{1, n_2})$, составит величину

$$N = 2^{n_0^1} C_{n_2}^{n_0^1}, \tag{9}$$

где N - число всевозможных $x_{qi}^{c(1)}$, получаемых из $x_{qi}^2 \in V_q^2$ $(i = \overline{1, n_2})$.

Для того чтобы найти n_0^1 логарифмируем (9)

$$\ln N = \ln 2^{n_0^1} + \ln C_{n_2}^{n_0^1}.$$

Учитывая $C_m^n \leqslant \frac{m^n}{2^n}$, из вышеприведенной формулы получим

$$\ln N = \ln 2^{n_0^1} + \ln C_{n_2}^{n_0^1} = n_0^1 \ln 2 + \ln \frac{n_2^{n_0^1}}{2^{n_0^1}} = n_0^1 \ln n_2. \tag{10}$$

Учитывая (1), (8) и (10), получим конкретную функциональную зависимость для n_0^1

$$n_0^1 = \frac{\varepsilon((m - m_q) + 1) + \ln \eta}{\ln n_2}. \tag{11}$$

Смысл (11) состоит в том, что если из N выбирается одно $x_{qr}^{c(1)}$ ($r \leqslant n_0^1$) или несколько $V_q^{c(1)} = x_{qr_1}^{c(1)}, x_{qr_2}^{c(1)}, \ldots, x_{qr_v}^{c(1)}$ ($r_i < (v = n_0^1)$) и они безошибочно разделяют $\forall X \in V_q$ от $\forall X \in V_p$ на V, то гарантируются заданные ε, η при распознавания объектов.

Так же, как и в n_0^1, при вычислении n_0^2 используются соотношения (7–11). В отличие от n_0^1, при вычислении n_0^2, в соотношении (7) m заменяется на $m^{**} = (m - m_p) + 1$ (где m_p – количество объектов класса V_p), так как $\forall X \in V_p$ считаются одинаковыми, и в качестве представительного выбирается любой объект.

Тогда

$$n_0^2 = \frac{\varepsilon((m - m_p) + 1) + \ln \eta}{\ln n_3}. \tag{12}$$

Далее, для того, чтобы в найденном пространстве n_0^1 и n_0^2 РП $R^q(X)$, построенном на основе найденных $V_q^{c(1)}$ и $V_q^{c(2)}$, безошибочно разделило бы $\forall X \in V_q$ от $\forall X \in V_p$ на V, и при этом $R^q(X)$ гарантировало бы требуемое ε, η при распознавании новых объектов, необходимо при вводе $x_{qr}^{c(1)}$ и $x_{qr}^{c(2)}$ в пространстве n_0^1 и n_0^2 наложить дополнительные требования для каждого выбираемого $x_{qr}^{c(1)}$ и $x_{qr}^{c(2)}$ из V_q^2 и V_q^3.

Введем следующие обозначения: $x_{qr}^{c(\alpha)}$ ($\alpha = \overline{1,2}$)- СП α-го типа (если $\alpha = 1$, то первого, и если $\alpha = 2$, то второго типа) относительно V_q; W_{qi}^{α} ($\alpha = \overline{1,2}$)-подмножество правильно классифицированных объектов с помощью $x_{qr}^{c(\alpha)}$; $|w_i^{\alpha}|$ -мощность подмножества W_{qi}^{α}.

Тогда РС для $x_{qr}^{c(\alpha)}$ определяется

$$f\left(x_{qr}^{c(\alpha)}\right) = \frac{|w_i^{\alpha}|}{m}. \tag{13}$$

Из этого соотношения видно, что $0 \leqslant f\left(x_{qr}^{c(\alpha)}\right) \leqslant 1$.

Зная предельную размерность пространства n_0^{α} ($\alpha = \overline{1,2}$) ((11) и (12)), можно оценить минимально-допустимую РС каждого СП $x_{qr}^{c(\alpha)}$, используемого при построении пространства n_0^{α} для выделения V_q

$$f_{min}\left(x_{qr}^{c(\alpha)}\right) \geqslant \frac{r}{n_0^{\alpha}}. \tag{14}$$

Для того, чтобы гарантировать безошибочное разделение $\forall X \in V_q$ от $\forall X \in V_p$ в найденном пространстве n_0^{α}, необходимо вычислить реальную РС каждого выбираемого $x_{qr}^{c(\alpha)}$.

Предположим, что в процессе обучения выбран $x_{qr}^{c(\alpha)}$, для которого выполняется (14). Тогда для выбора следующего $x_{qr_{(i+1)}}^{c(\alpha)}$ должно выполняться

$$f_{min}\left(x_{qr_{(i+1)}}^{c(\alpha)}\right) = \left|f_{min}\left(x_{qr_{(i+1)}}^{c(\alpha)}\right) - f_{min}\left(x_{qr_i}^{c(\alpha)}\right)\right| \geqslant \frac{r}{n_0^{\alpha}}. \tag{15}$$

Приведенные выше соотношения (13–15) позволяют в пространстве n_0^{α} формировать СП $V_q^{c(\alpha)}$, каждый из которых обладает требуемой реальной РС (15). При этом в формированным пространстве n_0^{α} с помощью $V_q^{c(\alpha)}$ безошибочно разделяются

$\forall X \in V_q$ от $\forall X \in V_p$ и гарантируется ε с надежностью $(1-\eta)$ с учетом заданных η, m и выбранных n_2 и n_3.

4 Алгоритм решения задачи

Рассмотрим алгоритм формирования пространства $V_q^{c(\alpha)} = x_{qr_1}^{c(\alpha)}, x_{qr_2}^{c(\alpha))}, \ldots, x_{qr_v}^{c(\alpha))}$ $(v = n_0^\alpha))$, обеспечивающий разделение $\forall X \in V_q$ от $\forall X \in V_p$ в пространстве n_0^α и гарантирующий ε с $(1-\eta)$ при распознавании новых объектов.

В алгоритме до обучения для $V_q^{c(\alpha)}$ отдельно находятся предельные значения размерности пространства n_0^α. Следовательно, каждое пространство в процессе обучения отдельно формируется из СП α-го типа относительно V_q. При этом в процессе последовательного выбора $V_q^{c(\alpha)}$ учитываются их реальные РС в найденном пространстве n_0^α.

Алгоритм включает в себя следующие основные этапы.

1. В оперативную память заносятся объекты $X_1, X_2, \ldots, X_m$, их признаки $x_1, x_2, \ldots, x_n$ и классы $V_1, V_2, \ldots, V_l$ в виде V.

2. В оперативную память заносятся $X_1, X_2, \ldots, X_b$ $b < m$ и значения их $x_1, x_2, \ldots, x_n$ в виде контрольной выборки V^*.

3. $q = 1$. В качестве V_q выбирается любой $V_j \in V$, т.е. $V_q = \forall V_j$ и в качестве V_p $(q \neq p)$ все остальные, кроме V_q, т.е. $V_p = V \setminus V_q$.

4. $i = 1$. Из оперативной памяти выбирается x_i, рассматриваемое как претендент на ППВТ.

5. Проверяется соотношение (2) или (3) для x_i. Если x_i логическое, то оно преобразуется с помощью (2), если же x_i непрерывное, то – с помощью (3).

6. Проверяется соотношение (4) для x_i. Если для x_i выполняется (4), то x_i является ПППТ x_{qi}^2 относительно V_q и x_{qi}^1 вводится в массив V_q^1, как ПППТ x_{qi}^1 и алгоритм переходит к шагу 9, в противном случае x_i не вводится в массив V_q^1 и алгоритм переходит к шагу 7.

7. Проверяется соотношение (5) для x_i. Если для x_i выполняется (5), то x_i является ППВТ x_{qi}^2 относительно V_q и x_{qi}^2 вводится в массив V_q^2, как ПППТ x_{qi}^2 и алгоритм переходит к шагу 9, в противном случае x_i не вводится в массив V_q^2 и алгоритм переходит к шагу 8.

8. Проверяется соотношение (6) для x_i. Если для x_i выполняется (6), то x_i как x_{qi}^3 вводится в массив V_q^3, и алгоритм переходит к шагу 9, в противном случае свойство x_i исключается из дальнейшего рассмотрения и алгоритм переходит к шагу 9.

9. $i = i + 1$. Если $i \leqslant n$, то алгоритм переходит к шагу 4, в противном случае – к шагу 10.

10. Формируются массивы $V_q^1 = x_{q1}^1, \ldots, x_{qn_1}^1$, $V_q^2 = x_{q1}^2, \ldots, x_{qn_2}^2$ и $V_q^3 = x_{q1}^3, \ldots, x_{qn_3}^3$ соответственно для ПППТ, ППВТ и ППТТ. Согласно (4) каждый отобранный ПППТ $x_{qi}^1 \in V_q^1$ безошибочно разделяет $\forall X \in V_q$ от $\forall X \in V_p$ и самостоятельно участвует при распознавании новых объектов. Поэтому ПППТ не используется при формировании СП.

11. $\alpha = 1$. Вычисляются предельные значения размерности пространства n_0^α для СП α-го типа. Если $\alpha = 1$, то n_0^α вычисляется соотношением (11) и алгоритм переходит к шагу 12, в противном случае – к шагу 35.

12. Выбирается в качестве представителя $X_z = \forall X \in V_q$, так как согласно (5) $\forall X \in V_q$ идентичные (одинаковые).

13. $t = 1$. Устанавливается t-й порядок расположения признаков в массиве $V_q^2 = x_{q1}^2, \ldots, x_{qn_2}^2$:

$$V_q^2 = x_{q1}^{2(t)}, x_{q2}^{2(t)}, \ldots, x_{qn_2}^{2(t)} \tag{16}$$

14. Фиксируется n_2-мерный булевый вектор $b_t^0 = \underbrace{111\ldots111}_{n_2}$ над V_q^2.

15. $k = 1$. Если $k = 1$, то осуществляется переход к шагу 16, если же $k > 1$, то – к шагу 18.

16. *1-й шаг поиска.* Присваивается $b_t^0 = \underbrace{111\ldots111}_{n_2}$ следующее значение $b_t^1 = \underbrace{011\ldots111}_{n_2}$. Если в b_t^1 векторе $b_k = 0$, то $X_z \in V_q$ не сравнивается $\forall X \in V_p$ по $x_{qk}^{2(t)}$ соответствующего разряда b_k. Если в b_t^1 векторе $b_k = 1$, то $X_z \in V_q$ сравнивается $\forall X \in V_p$ по $x_{qk}^{2(t)}$ соответствующего разряда b_k. Для сравнения используется следующее правило

$$d(X_z, X_j) = \begin{cases} 1, & \text{если } x_{zk} = x_{jk},\ k = \overline{1, n}; \\ 0, & \text{в противном случае.} \end{cases} \tag{17}$$

17. Если для $X_z \in V_q$ и $\forall X \in V_p$ выполняется $x_{q2}^{2(t)} \neq x_{p2}^{2(t)}$, $x_{q3}^{2(t)} \neq x_{p3}^{2(t)}$, ..., $x_{qn_2}^{2(t)} \neq x_{pn_2}^{2(t)}$ при $b_t^1 = \underbrace{011\ldots111}_{n_2}$, то $b_t^1 = \underbrace{011\ldots111}_{n_2}$ заменяется с $b_t^2 = \underbrace{001\ldots111}_{n_2}$ и осуществляется переход к шагу 18. Если же для $X_z \in V_q$ и $\forall X \in V_p$ выполняется $x_{q2}^{2(t)} = x_{p2}^{2(t)}$, $x_{q3}^{2(t)} = x_{p3}^{2(t)}$, ..., $x_{qn_2}^{2(t)} = x_{pn_2}^{2(t)}$ при $b_t^1 = \underbrace{011\ldots111}_{n_2}$, то $b_t^1 = \underbrace{011\ldots111}_{n_2}$ не заменяется с $b_t^2 = \underbrace{001\ldots111}_{n_2}$ и осуществляется переход к шагу 18.

18. *k-й шаг поиска.* Из (16) выбирается следующее k-й $x_{qk}^{2(t)}$. Если для $X_z \in V_q$ и $\forall X \in V_p$ выполняется $x_{q2}^{2(t)} \neq x_{p2}^{2(t)}$, $x_{q3}^{2(t)} \neq x_{p3}^{2(t)}$, ..., $x_{qn_2}^{2(t)} \neq x_{pn_2}^{2(t)}$ при $b_t^k = \underbrace{011\ldots0\ldots111}_{n_2}$, то $b_t^k = \underbrace{011\ldots0\ldots111}_{n_2}$ заменяется с $b_t^{k+1} = \underbrace{011\ldots00\ldots111}_{n_2}$ и осуществляется переход к шагу 19. Если же для $X_z \in V_q$ и $\forall X \in V_p$ выполняется $x_{q2}^{2(t)} = x_{p2}^{2(t)}$, $x_{q3}^{2(t)} = x_{p3}^{2(t)}$, ..., $x_{qn_2}^{2(t)} = x_{pn_2}^{2(t)}$ при $b_t^k = \underbrace{011\ldots0\ldots111}_{n_2}$, то $b_t^k = \underbrace{011\ldots0\ldots111}_{n_2}$ не заменяется с $b_t^{k+1} = \underbrace{011\ldots10\ldots111}_{n_2}$ и осуществляется переход к шагу 19.

19. $k = k + 1$. Если $k \leqslant n_2$, то осуществляется переход к шагу 15, противном случае – к шагу 20.

20. Определяется СППТ

$$x_{qr_t}^{c(1)} = x_{q1}^2, \overline{x}_{q2}^2, x_{q3}^2, \ldots, x_{q(k-1)}^2, \overline{x}_{qk}^2, x_{q(k+1)}^2, \ldots, x_{qr}^2,$$

соответствующее вектору $b_t^k = \underbrace{111\ldots111\ldots111}_{n_2}$.

21. Вычисляются $f\left(x_{qr_t}^{c(1)}\right)$ и $f_{min}\left(x_{qr_t}^{c(1)}\right)$ СП $x_{qr_t}^{c(1)}$ в виде (13) и (14). Если $f\left(x_{qr_t}^{c(1)}\right) \geqslant f_{min}\left(x_{qr_t}^{c(1)}\right)$, то $x_{qr_t}^{c(1)}$ оставляется, в противном случае $x_{qr_t}^{c(1)}$ исключается.

22. $t = t + 1$. Если $t \leqslant N$ осуществляется переход к шагу 13, в противном случае – к шагу 23.

23. Формируется СППТ соответственно для каждого t-го ($t = \overline{1, N}$) порядка расположения ППВТ $V_q^2 = x_{q1}^{2(t)}, x_{q2}^{2(t)}, \ldots, x_{qn_2}^{2(t)}$ в виде:

$$x_{qt}^{c(1)} = x_{q1}^2, \overline{x}_{q2}^2, x_{q3}^2, \ldots, x_{q(k-1)}^2, \overline{x}_{qk}^2, x_{q(k+1)}^2, \ldots, x_{qr_t}^2, (t = \overline{1, N}; r_t \leqslant n_0^1).$$

24. Фиксируется вектор соответственно для каждого полученного $x_{qt}^{c(1)}$ ($t = \overline{1, N}$) в виде $b_{qt}^1 = \underbrace{011\ldots0\ldots111}_{n_2}$ ($t = \overline{1, N}$).

25. Упорядочиваются $x_{qt}^{c(1)}$ ($t = \overline{1, N}$) по убыванию $x_{qt}^{c(1)} \geqslant \ldots \geqslant x_{qN}^{c(1)}$ относительно $f\left(x_{qr_t}^{c(1)}\right) \geqslant \ldots \geqslant f\left(x_{qr_t}^{c(1)}\right)$.

26. $t = 1$. Из $f\left(x_{qr_t}^{c(1)}\right)$ выбирается $x_{qr_{\tilde{\alpha}}}^{c(1)}$, обладающее максимальной РС согласно (13).

27. $t = 2$. Вычисляется $f\left(x_{qr_t}^{c(1)}\right)$ СП $x_{qr_t}^{c(1)}$ в виде (15). Если $f\left(x_{qr_t}^{c(1)}\right)$ СП $x_{qr_t}^{c(1)}$ удовлетворяет (15), то $x_{qr_t}^{c(1)}$ включается в $V_q^{c(1)} = x_{qr_1}^{c(1)}$ и осуществляется переход к шагу 28, в противном случае $x_{qr_t}^{c(1)}$ не включается в $V_q^{c(1)}$ и осуществляется переход к шагу 28.

28. $t = t + 1$. Если $t \geqslant N$, то алгоритм переходит к шагу 27, в противном случае – к шагу 29.

29. Формируется массив СППТ $V_q^{c(1)} = x_{qr_1}^{c(1)}, x_{qr_1}^{c(1)}, \ldots, x_{qr_v}^{c(1)}$ ($r_r \leqslant (v = n_0^2)$).

30. Для V_q обобщаются полученные векторы и СППТ в виде

$$V_q = \bigvee_{t=1}^{N} b_{qt}^1; \quad V_q = \bigvee_{\substack{t=1 \\ \left(b_{qt}^1 = 1\right)}}^{N} V_{qt}^{c(1)}.$$

31. $q = q + 1$. Если $q \geqslant l$, то осуществляется переход к шагу 3, противном случае – к шагу 32.

32. В результате для $\left(V_1, V_2, \ldots, V_l\right) \in V$ соответственно получим системы векторов и системы СППТ в виде:

$$V_j = \bigvee_{t=1}^{N} b_{1t}^1, (j = \overline{1, l}); \quad V_j = \bigvee_{\substack{t=1 \\ \left(b_{1t}^1 = 1\right)}}^{N} V_{1t}^{c(1)}; \tag{18}$$

33. На основе системы векторов (18) каждый новый $X \in V^*$ сравнивается с $X \in V_q$ ($q = \overline{1, l}$) с помощью правило (17) по найденной системе СППТ (18) и вычисляется сумма голосов схожести распознаваемого X к V_j ($j = \overline{1, l}$):

$$\Gamma_\Sigma\left(X \in V_j\right) = \sum_{\substack{t=1 \\ \left(b_{1t}^1 = 1\right)}}^{N} V_{1t}^{c(1)}, \quad (j = \overline{1, l}).$$

34. Если $\alpha = 1$, то алгоритм переходит к шагу 53, в противном случае – к шагу 35.

35. $\alpha = \alpha + 1$. Если $\alpha \leqslant 2$, то алгоритм переходит к шагу 11, в противном случае – к шагу 36.

36. Вычисляются предельные значения размерности пространства n_0^α для СП α-го типа вычисляется с соотношением (12).

37. Выбирается в качестве представителя $X_z = \forall X \in V_p$, так как согласно (6) $\forall X \in V_p$ идентичные.

38. $j = 1$. Выбирается $X_j \in V_q$.

39. Фиксируется n_3-мерный булевый вектор $b^0_{ji} = \underbrace{111\ldots111}_{n_3}$ над V^3_q.

40. Сравнивается $X_j \in V_p$ с $X_z \in V_q$ в виде (16) и вычисляется новое значение вектора

$$b^1_{ji} = \begin{cases} 1, & \text{если } x_{ji} \neq x_{zi},\ i = \overline{1, n_3}; \\ 0, & \text{если } x_{ji} = x_{zi},\ i = \overline{1, n_3}. \end{cases} \tag{19}$$

41. Определяется $b^1_{ji} = \underbrace{01011\ldots101}_{n_3}$. Если в b^1_{ji} векторе $b_i = 0$, то $X_j \in V_q$ схожи с $X_z \in V_p$ по x^3_{ji}, если же $b_i = 1$, то $X_j \in V_q$ несхожи с $X_z \in V_p$ по x^3_{ji}.

42. Определяется СПВТ

$$x^{c(2)}_{qj} = x^3_{q2}, \overline{x}^3_{q3}, x^3_{q5}, \ldots, x^3_{q(k-1)}, \overline{x}^3_{qk}, x^3_{q(k+1)}, \ldots, x^3_{qr}, \quad (r \leqslant n_3),$$

соответствующее вектору $b^1_{ji} = \underbrace{01011\ldots111\ldots1}_{n_3}$.

43. $j = j + 1$. Если $j \leqslant m_q$, то алгоритм переходит к шагу 38, в противном случае – к шагу 44.

44. Формируется СПВТ соответственно для каждого $X_j \in V_q$ $(j = \overline{1, m_q})$ r-го ранга в виде:

$$x^{c(2)}_{qj} = x^3_{q1}, \overline{x}^3_{q2}, x^3_{q3}, \ldots, x^3_{q(k-1)}, \overline{x}^3_{qk}, x^3_{q(k+1)}, \ldots, x^3_{qr_j}, \quad (r_j \leqslant n_3; j = \overline{1, m_q}).$$

45. Фиксируется вектор соответственно для каждого полученного СПВТ в виде $b^1_{ji} = \underbrace{011\ldots0\ldots111}_{n_3}$ $(j = \overline{1, m_q})$.

46. Объединяются полученные СПВТ $x^{c(2)}_{q1}, x^{c(2)}_{q2}, \ldots, x^{c(2)}_{qm_q}$ и векторы $b^1_{1i}, b^1_{2i}, \ldots, b^1_{m_q i}$ для V_q в виде:

$$V_q = x^{c(2)}_{q1} \vee x^{c(2)}_{q2} \vee \cdots \vee x^{c(2)}_{qm_q}, \quad V_q = b^1_{1i} \vee b^1_{2i} \vee \cdots \vee b^1_{m_q i}. \tag{20}$$

(20) является дизъюнктивная нормальная форма (ДНФ).

47. Используя законы $K_j = K_{j+1}$ и $K_j \vee K_{j+1} = K_j$ математической логики, из (20) удаляются лишние СПВТ ($K_j = x^{c(2)}_{qj}$ конъюнкция) и в результате для V_q получится сокращенная ДНФ:

$$V_q = x^{c(2)}_{q1} \vee x^{c(2)}_{q2} \vee \cdots \vee x^{c(2)}_{qt} \quad (t \leqslant m_q), \tag{21}$$

и соответственно

$$V_q = b^1_{1i} \vee b^1_{2i} \vee \cdots \vee b^1_{ti}. \tag{22}$$

48. Вычисляются $f\big(x^{c(2)}_{qj_t}\big)$, $f_{min}\big(x^{c(2)}_{qj}\big)$ и реальная РС СП $x^{\tilde{n}(2)}_{q2}$ соответственно в виде (13), (14) и (15). Если для $x^{\tilde{n}(2)}_{q2}$ выполняется (15), то $x^{\tilde{n}(2)}_{q2}$ оставляется, в противном случае исключается из (21).

49. В результате получим системы СП типа $x^{\tilde{n}(2)}_{q2}$, каждый из которых состоит не более n^2_0 ППТТ:

$$V_q = x^{c(2)}_{q1} \vee x^{c(2)}_{q2} \vee \cdots \vee x^{c(2)}_{qs} \quad (s \leqslant t),$$

где

$$x_{qi}^{c(2)} = \underbrace{x_{q1}^{3}, \overline{x}_{q2}^{3}, x_{q3}^{3}, \ldots, x_{q(k-1)}^{3}, \overline{x}_{qk}^{3}, x_{q(k+1)}^{3}, \ldots, x_{qr_1}^{3}}_{n_0^2}, \quad (r_i \leqslant n_0^2, i = \overline{1, s}).$$

Соответственно фиксируется

$$V_q = b_{1i}^1 \vee b_{2i}^1 \vee \cdots \vee b_{si}^1.$$

50. $q = q + 1$. Если $q \leqslant l$, то переход к шагу 3, противном случае к шагу 51.

51. В результате для $V_1, V_2, \ldots, V_l \in V$ получим системы СПВТ:

$$V_j = x_{j1}^{c(2)} \vee x_{j2}^{c(2)} \vee \cdots \vee x_{js_j}^{c(2)} \quad (s_j \leqslant t, j = \overline{1, l}). \tag{23}$$

и соответственно булевые векторы:

$$V_j = b_{j1}^1 \vee b_{j2}^1 \vee \cdots \vee b_{js_j}^1 \quad (s_j \leqslant t, j = \overline{1, l}). \tag{24}$$

52. На основе системы векторов (24) каждый новый $X \in V^*$ сравниваются с $X \in V_q$ $(q = \overline{1, l})$ с помощью (16) по найденной системе СПВТ (23) и вычисляется сумма голосов схожести распознаваемого X к V_q $(q = \overline{1, l})$:

$$\Gamma_{\Sigma}\big(X \in V_j\big) = \sum_{\substack{t=1 \\ \left(b_{1t}^1 = 1\right)}}^{m_1} x_{1j}^{c(2)} \quad (j = \overline{1, l}).$$

53. Для распознавания $X \in V^*$ используется РП

$$F(X) = \begin{cases} X \in V_j, & \text{если } \max\limits_{1 \leqslant j \leqslant l} \Big\{\Gamma_{\Sigma}\big(X \in V_1\big), \Gamma_{\Sigma}\big(X \in V_2\big), \ldots, \Gamma_{\Sigma}\big(X \in V_l\big)\Big\}; \\ X \notin V_1, \ldots, X \notin V_l & \text{если } \Gamma_{\Sigma}\big(X \in V_1\big) = \Gamma_{\Sigma}\big(X \in V_2\big) = \ldots, = \Gamma_{\Sigma}\big(X \in V_l\big). \end{cases}$$

54. Вывод результатов, относящих $X \in V^*$ по сумме голосований в один из $V_1, V_2, \ldots, V_l \in V$, или указывающее для $X \in V^*$ отказ от распознавания.

5 Программный комплекс

Создан комплекс программ на основе разработанного алгоритма. Общий вид интерфейсного окна программы приведен на Рис. 1.

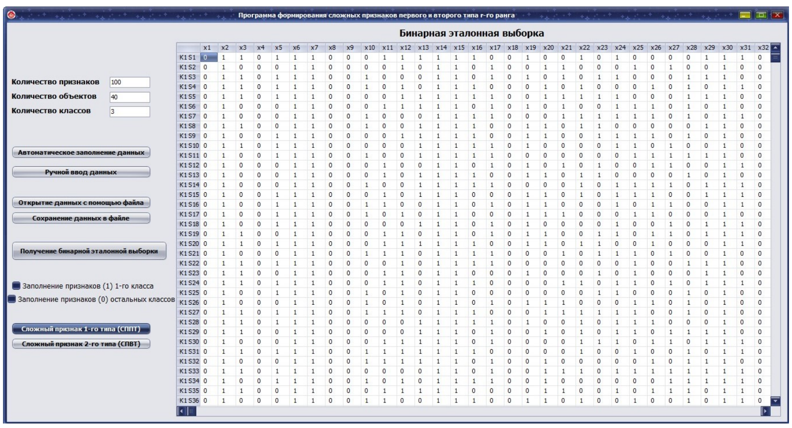

Рис. 1 Общий вид интерфейсного окна

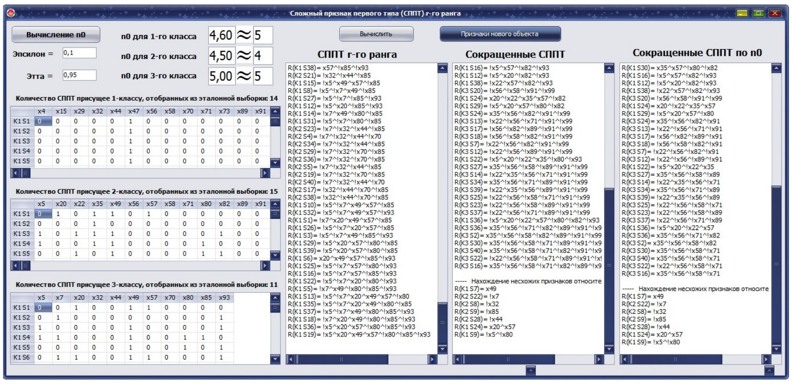

Рис. 2 Отобранные и сокращенные СППТ

Модуль вычисления n_0^1, отбора и сокращения СППТ представлен на рис. 2.

Модуль распознавания объектов с помощью сокращенных СППТ представлен на рис. 3.

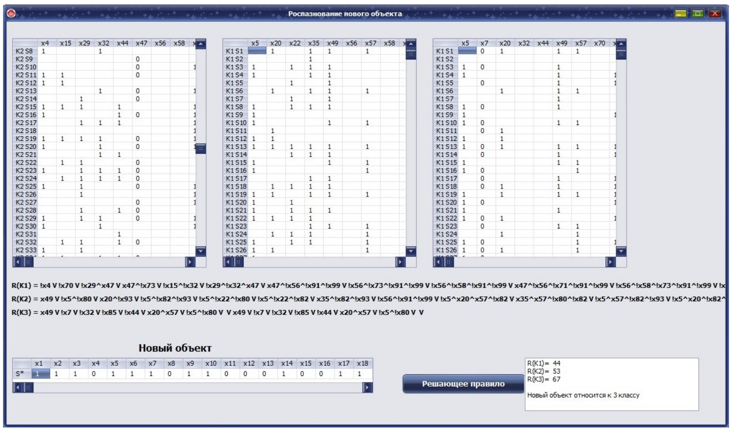

Рис. 3 Результаты распознавания с помощью сокращенных СППТ

Модуль вычисления n_0^2, отбора и сокращения СПВТ представлен на рис. 4.

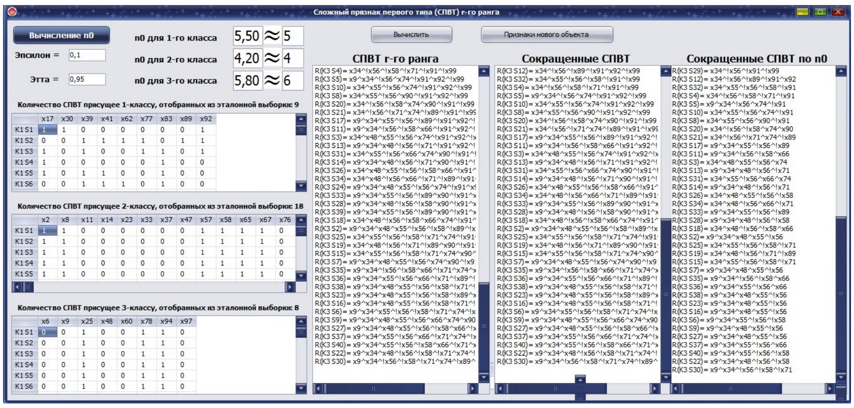

Рис. 4 Отобранные и сокращенные СПВТ

Модуль распознавания объектов с помощью сокращенных СПВТ представлен на рис. 5.

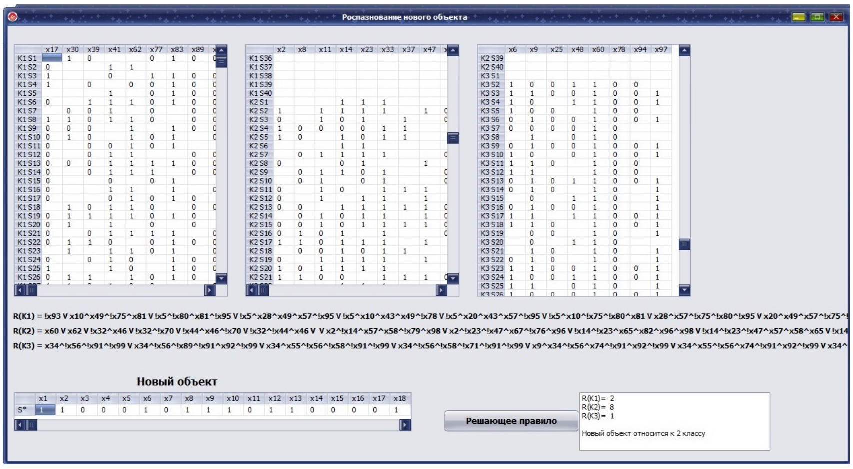

Рис. 5 Результаты распознавания с помощью сокращенные СПВТ

Проведено испытание по оценке работоспособности и эффективности предложенного алгоритма и программного комплекса применительно к задачам распознавания. Полученные результаты подтверждают то, что разработанный алгоритм и программный комплекс можно применять при решении практических задач распознавания объектов, касающихся медицинской, технической, археологической, гидрогеологической, геологической, сейсмологической, биологической и других сфер.

6 Заключение и выводы

В отличие от алгоритмов, приведенных в [6–8, 16–18], в данном алгоритме:

- при определении ПППТ, ППВТ и ППТТ вводится новое требование, т.е. при определении ПП учитывается не только присутствие, а одновременно, и отсутствие $x_i \in V$ $i = \overline{1, n}$ на $X_j \in V_q$ $j = \overline{1, m_q}$;
- до начала процесса формирования пространства СППТ и СПВТ r-го ранга вычисляются предельные значения размерности $n_0^1 < n$ и $n_0^2 < n$ соответственно с учетом выбранных n_2 и n_3, а также при заданных ε, η, m;
- при формировании n_0^1 и n_0^2 учитывается реальная разделяющая сила каждого выбираемого СППТ и СПВТ. Это приводит к резкому сокращению множества системы СППТ и СПВТ, а также позволяет при распознавании новых объектов использовать только те СП, которые входят в пространства n_0^1 и n_0^2;
- вычисляется сумма голосов схожести $\Gamma_\Sigma(X \in V_j)$ в пространствах n_0^1 и n_0^2, а также распознаются $X \in V*$ по сумме результатов голосования, одному из заданных $V_q \in V$;
- резко снижается объем вычислений на компьютере, так как:
  - на 6-8 этапах алгоритма удаляется из дальнейшего рассмотрения лишние $x_i \in V$, не являющиеся ПП согласно (4)-(6);
  - на 12 и 37 этапах алгоритма при формировании СППТ и СПВТ участвует не все объекты, а только один представительный $X_z \in V_q$ и $X_z \in V_p$ согласно (5) и (6);
  - на 27 и 48 этапах алгоритма удаляется из дальнейшего рассмотрения лишние СППТ и СПВТ, не удовлетворяющие требованиям реальной РС согласно (15);
  - при распознавании новых используются не все исходные $X \in V^*$, а только те, для которых выполняются (5), (6) и (15).

Необходимо отметить, что в алгоритме сильно сокращается количество исходных $x_1, x_2, \ldots, x_n$. Сокращение количества исходных $x_1, x_2, \ldots, x_n$ представляется полезным в двух аспектах: во-первых, снижается объем вычислений, а во-вторых, с удалением лишних из состава СП повышается надежность распознавания. Одновременно, за счет сокращения исходных $x_1, x_2, \ldots, x_n$ и СП уменьшается класс РП, что приводит зачастую к снижению надежности распознавания в целом. Поэтому, в этом алгоритме, рассматривается их функциональная зависимость (11) и (12), чтобы качество и надежность распознавания, а также объем (количество СП и объектов) ЭВ находился в требуемом разумном интервале.

Литература

[1] *Айзерман М. А., Браверман З. М., Розоноэр Л. И.* Метод потенциальных функций в теории обучения машин. –М.: Наука, 1970. 384 с.

[2] *Браверман Э. М.* Методы экстремальной группировки параметров и задачи выделения существенных факторов // Автоматика и телемеханика, 1970. – Т. 11, № 1. – С. 123–132.

[3] *Васильев В. И.* Проблема обучения распознаванию классов. Принципы, алгоритмы, реализация. –Киев: Выща школа, 1989. 64 с.

[4] *Вапник В. Н., Червоненкис А. Я.* Теория распознавания классов (статистические проблемы обучения). –М.: Наука, 1974. 464 с.

[5] *Васильев В. И.* О простоте решающих функций в проблеме обучения распознаванию образов // Автоматика, 1984. – № 2. – С. 14–23.

[6] *Васильев В. И., Бекмуратов К. А., Овсянникова Ф. Н.* Использование полипризнаков в задачах обучения распознаванию образов методом предельных упрощений // Автоматика, 1989. – № 2. – С. 13–16.

[7] *Васильев В. И., Бекмуратов К. А. и др.* Выбор бинарных псевдопризнаков в процессе обучения распознаванию образов // Автоматика, 1989. – № 3. – С. 23–28.

[8] *Айвазян С. А. и др.* Прикладная статистика: Классификация и снижение размерности. –М.: Финансы и статистика, 1989. 607 с.

[9] *Журавлев Ю.И. и др.* Пакет прикладных программ для решения задач распознавания и классификации (ПАРК). –М.: ВЦ АН СССР, 1981. 24 с.

[10] *Мазуров В. Д., Хачай М. Ю.* Комитеты систем линейных неравенств // Автоматика и телемеханика, 2004. Вып. 3. – С. 43–54.

[11] *Обухов А. С., Рязанов В. В.* Применение релаксационных алгоритмов при оптимизации линейных решающих правил // Доклады 10-й Всероссийской конференции "Математические методы распознавания образов "(ММРО-10)". –М.: Издательство, 2001. – С. 102–104.

[12] *Абдукаримов Р. Т.* Алгоритмы распознавания, основанные на поиске признаков классов // Вопросы кибернетики, 1976. Вып. 38.

[13] *Васильев В. И.* Распознающие системы. –Киев: Наукова думка, 1983. 421 с.

[14] *Журавлев Ю.И., Камилов М.М, Тулягaнов Ш.Е.* Алгоритмы вычисления оценок и их применение. –Ташкент: ФАН, 1974. 119 с.

[15] *Бекмуратов К. А., Васильев В. И., Бекмуратов Д. К.* Нахождение предельно-допустимых значений размерности признаковых пространств из обучающей выборки // Доклады международной научно-практической конференции "Современное состояние и перспективы развития информационных технологий". –Ташкент: Издательство ИМИТ АН РУз, 2011. – С. 309–312.

[16] *Бекмуратов К. А., Бекмуратов Д. К.* Последовательный выбор признаков, обладающих требуемой разделяющей силой // Доклады XI Международной научно-практической конференции "Научные перспективы XXI века. Достижения и пер-

спективы нового столетия". –Новосибирск: Международный научный институт "EDUCATIO 2015. – С. 9–13.

[17] *Bekmuratov D. Q., Umarov F.Sh.* Procedures for the formation of the k-th type feature space // 2nd International Scientific and Practical Conference "Topical researches of the World Science"Proceedings. – Dubai, UAE, 2016.

[18] *Бекмуратов Д. К.* Разработка алгоритма формирования системы опорных множеств признаков, обеспечивающих качество и надежность распознавания // Проблемы вычислительной и прикладной математики, 2017. – № 5. – С. 74–79.

Поступила в редакцию 01.12.2018

UDC 004.93

FORMATION OF COMPLEX FEATURE SPACES OF R RANKS, WHICH PROVIDE QUALITY AND RELIABILITY OF RECOGNITION

[1]***Bekmuratov Q.A.***, [2]***Akhatov A.R.***, [1]***Bekmuratov D.Q.***

bekmurodov1958@mail.ru; akmalar@rambler.ru; bekmurodov_d@mail.ru

[1]Samarkand branch of Tashkent university of information technologies, Samarkand, Ibn Sino 2;

[2]Samarkand state university, Samarkand, University boulevard 15

We consider the solution of the problem of determining simple features inherent in a particular class from a given property of objects of the reference sample and the formation of complex attribute spaces from the selected simple features. A technique for selecting simple features from the original properties, finding the maximum allowable dimension of the space of complex features, is given. The main parameters that play a significant role in the process of finding the maximum allowable dimension of the space of complex features are indicated: the probability of an erroneous classification of objects; probability of error achieved with a given reliability; the number of objects of the reference sample and the number of features allocated to a particular class. The minimal and real separating forces of complex features are determined on the basis of the found maximum permissible dimension of space. The principles of the formation of a system of complex signs of the first and second types are proposed for each class of simple signs of the second and third types, taking into account the minimal and real separating force. Based on the proposed methodology, an algorithm and software have been developed. Computational experiments were carried out on a computer, the results of which are given in the form of decision rules that are used to recognize objects. The findings of the study as a whole are also given.

Keywords: object, attribute, class, reference sample, control sample, volume of reference sample, systems of complex features, recognition quality and reliability, error probability, marginal dimension of space, amount of votes, decision rule

Citation: Bekmuratov Q.A., Akhatov A.R., Bekmuratov D.Q. 2019. Formation of complex feature spaces of r ranks, which provide quality and reliability of recognition. *Problems of Computational and Applied Mathematics.* 1(19): 24–38.