

¹Д-р техн. наук, профессор, заведующий лабораторией Центра разработки программных продуктов и аппаратно-программных комплексов при Ташкентском Университете Информационных Технологий, Ташкент, Узбекистан

²Канд. техн. наук, старший научный сотрудник Центра разработки программных продуктов и аппаратно-программных комплексов при Ташкентском Университете Информационных Технологий, Ташкент, Узбекистан

³Младший научный сотрудник Центра разработки программных продуктов и аппаратно-программных комплексов при Ташкентском Университете Информационных Технологий, Ташкент, Узбекистан

ПОСТРОЕНИЕ РАСПОЗНАЮЩИХ ОПЕРАТОРОВ В УСЛОВИЯХ ВЗАИМОСВЯЗАННОСТИ ПРИЗНАКОВ

Рассмотрена задача распознавания образов, заданных в пространстве взаимосвязанных признаков. Предложен новый подход к построению модели распознающих операторов, учитывающий взаимосвязанность заданных признаков. При этом построение модели осуществлено для распознающих операторов типа потенциальных функций. Основная идея предлагаемого подхода заключается в формировании независимых подмножеств взаимосвязанных признаков и выделении предпочтительной модели зависимости для каждого подмножества сильносвязанных признаков. Анализ полученных результатов показывает, что рассмотренные распознающие операторы используются в тех случаях, когда между признаками объектов, принадлежащих к одному и тому же классу, существует некоторая зависимость. При слабом выражении этой зависимости используется классическая модель распознающих операторов. Основным преимуществом предложенных распознающих операторов является улучшение точности и значительное сокращение объема вычислительных операций при распознавании неизвестных объектов, что позволяет применить их при построении распознающих систем, работающих в режиме реального времени.

Ключевые слова: распознавание образов, модель распознающих операторов, потенциальная функция, зависимость признаков, подмножество сильносвязанных признаков, репрезентативный признак, предпочтительная модель зависимости.

НОМЕНКЛАТУРА

\mathfrak{S} – множество допустимых объектов;

\tilde{S}^m – множество выделенных допустимых объектов

$(\tilde{S}^m \subset \mathfrak{S}, \tilde{S}^m = \{S_1, \dots, S_i, \dots, S_m\})$;

\tilde{S}^q – множество контрольных объектов

$(\tilde{S}^q \subset \mathfrak{S}, \tilde{S}^q = \{S'_1, \dots, S'_q\})$;

\tilde{K}_j – множество эталонных объектов, принадлежа-

щих к классу K_j ($\tilde{K}_j \subset \tilde{S}^m$);

$|\tilde{K}_j|$ – мощность множества объектов \tilde{K}_j ;

V_o – обучающая выборка

$(V_o = \{< S_i; \chi(S_i) > : S_i \in \tilde{S}^m\})$;

V_k – контрольная выборка

$(V_k = \{< S_u; \chi(S_u) > : S_u \in \tilde{S}^q\})$;

$\chi(S_i)$ – индикатор объекта S_i

$(\chi(S_i) = -1, \text{ при } S_i \in K_1, \chi(S_i) = +1, \text{ при } S_i \in K_2)$;

X – пространства n -мерных признаков

$(X = (x_1, \dots, x_i, \dots, x_n))$;

A – произвольный алгоритм распознавания

$(A = B \circ C)$;

B – распознающий оператор,

$B(I_o, \tilde{S}^q) = (b_1, b_2, \dots, b_q)$;

C – решающее правило $C(b_i) = \lambda_i (\lambda_i \in \{-1, 0, +1\})$;

$\tilde{\pi}$ – вектор параметров распознающего оператора;

$\tilde{\pi}_0$ – вектор параметров распознающего оператора,

вычисленных по выборке V_o ;

$\varphi(\tilde{\pi})$ – функционал качества распознающего опера-

тора $B(\tilde{\pi})$;

n' – число подмножеств сильносвязанных признаков;

Ω_q – q -е подмножество сильносвязанных признаков;

Y – пространство n' -мерных репрезентативных признаков ($Y = (y_1, \dots, y_{n'})$).

ВВЕДЕНИЕ

В настоящее время все более широкий круг специалистов уделяет внимание проблеме распознавания образов, и число научных публикаций по данной тематике постоянно растет. Это связано с тем, что в последние годы распознавание образов находит все большее применение в науке, технике, производстве и повседневной жизни.

В проблеме распознавания образов центральное место занимает задача построения распознающих алгоритмов. Для рассмотрения разнообразных постановок этой задачи можно выделить три параметра, от способа задания которых во многом определяется метод решения. К ним относятся, во-первых, вид описания распознаваемых объектов (независимые и взаимосвязанные признаки), во-вторых, способ предъявления обучающей информации (единственная выборка и последовательность выборок) и, в-третьих, тип используемой модели, на основе которой формируется алгоритм распознавания (см. раздел 2).

Такой подход дает возможность выделить отдельные подклассы задач распознавания образов, описанных в пространстве взаимосвязанных признаков. Однако решение этих задач исследовано недостаточно (см. раздел 2).

Целью данной работы является разработка модифицированной модели распознающих операторов типа потенциальных функций [1–3], направленная на приспособление данной модели к решению задач распознавания образов, описанных в пространстве взаимосвязанных признаков.

Для достижения этой цели необходимо решить следующие задачи:

- проанализировать существующие модели распознающих операторов и определить круг решаемых задач;
- разработать модель алгоритмов распознавания, основанных на оценке взаимосвязанности признаков;
- провести экспериментальные исследования для оценки эффективности разработанных распознающих операторов.

Объектом исследования являются распознающие операторы типа потенциальных функций. Предмет исследования – модифицированные распознающие операторы, основанные на оценке взаимосвязанности признаков.

Основная идея предлагаемой модели распознающих операторов заключается в поиске некоторых зависимостей между признаками, характеризующими объекты, принадлежащие к одному и тому же классу.

В научном плане результаты данной работы представляют собой новое решение научной задачи, связанной с построением распознающих алгоритмов при условии большой размерности пространства признаков. Практическая значимость полученных результатов заключается в том, что разработанные алгоритмы и программы могут быть применены при решении прикладных задач в условиях большой размерности пространства признаков (например, при идентификации человека по фото-портрету).

1 ПОСТАНОВКА ЗАДАЧИ

Для простоты, но без ограничения общности, рассмотрим задачу распознавания образов в случае двух классов. Пусть набор объектов $\tilde{S}^q = \{S'_1, \dots, S'_q\}$ ($\tilde{S}^q \in \{S\}$) задан в пространстве признаков X большой размерности ($X = (x_1, \dots, x_i, \dots, x_n)$, $n > 200$).

Основная задача заключается в построении такого распознающего оператора B , который с применением решающего правила C (см. формула 2 в разделе 2) позволяет вычислить значения характеристической функции $\tilde{\lambda}_i = \chi(S'_i)$, $i = \overline{1, q}$ по выборке V_o :

$$B(\tilde{S}^q) = \|b_i\|_q, \quad \|b_i\|_q = (b_1, \dots, b_i, \dots, b_q)^T,$$

$$C(\|b_i\|_q) = (\tilde{\lambda}_1, \dots, \tilde{\lambda}_i, \dots, \tilde{\lambda}_q)^T, \quad \tilde{\lambda}_i \in \{-1, 0, 1\}.$$

Здесь $\tilde{\lambda}_i$ интерпретируется следующим образом: если $\tilde{\lambda}_i \in \{-1, +1\}$ ($\tilde{\lambda}_i = -1$ – объект S'_i входит в класс K_1 , $\tilde{\lambda}_i = +1$ – объект S'_i входит в класс K_2), то $\tilde{\lambda}_i$ есть значение характеристической функции $\lambda(S'_i)$ на допустимом объекте S'_i , вычисленное оператором B ; если $\tilde{\lambda}_i = 0$, то считается, что оператор B не смог определить значение характеристической функции $\lambda(S'_i)$.

2 ОБЗОР ЛИТЕРАТУРЫ

В течение длительного времени подавляющее большинство приложений теории распознавания образов

было связано с плохо формализованными областями – медициной, геологией, социологией, химией и т. д. Поэтому на первом этапе развития распознавания появилось множество алгоритмов, которые носили характер проектов различных технических устройств или алгоритмов для решения конкретных прикладных задач. Ценность разработанных алгоритмов определялась, прежде всего, достигнутыми экспериментальными результатами [2, 4].

Второй этап развития теории распознавания образов характеризуется переходом от отдельных алгоритмов к построению моделей – семейства алгоритмов для единого описания методов решения классификационных задач. Потребность в синтезе моделей алгоритмов распознавания образов определялась необходимостью фиксировать класс алгоритмов при выборе оптимальной или хотя бы приемлемой процедуры решения конкретной задачи.

На данном этапе развития Ю. И. Журавлев показал [2], что произвольный алгоритм распознавания можно представить как последовательное выполнение операторов B (распознающий оператор) и C (решающее правило):

$$A = B \circ C. \quad (1)$$

Из (1) следует, что каждый алгоритм A можно разделить на два последовательных этапа. На первом этапе распознающий оператор B осуществляет перевод допустимого объекта S_i в числовую оценку b_i : $B(S_i) = b_i$.

На втором этапе по числовой оценке b_i решающее правило C определяет принадлежность объекта S_i к классам K_1 и K_2 :

$$C(\mu(S_i)) = C(b_i), \quad C(b_i) = \tilde{\lambda}_i,$$

$$\tilde{\lambda}_i = \begin{cases} -1, & \text{если } \mu(S_i) < -c; \\ 0, & \text{если } c \leq \mu(S_i) \leq c; \\ +1, & \text{если } c < \mu(S_i), \end{cases} \quad (2)$$

где c – параметр решающего правила.

К настоящему времени построено и изучено несколько типов моделей. К их числу относятся [1–12]:

- модели, основанные на использовании разделяющих функций;
- модели, основанные на использовании математической статистики и теории вероятности;
- модели, основанные на использовании аппарата математической логики;
- модели, основанные на использовании принципа потенциалов;
- модели, основанные на использовании алгоритмов вычисления оценок.

Анализ литературных источников, в частности [1–12], показывает, что разработанные модели алгоритмов распознавания, в основном, ориентированы на решение задач, где объекты описаны в пространстве независимых признаков (или зависимость между признаками достаточно слабая).

На практике часто встречаются прикладные задачи распознавания образов, заданных в пространстве признаков большой размерности. При решении подобных

задач предположение о независимости признаков достаточно часто не выполняется [4, 13]. Следовательно, остается недостаточно решенным вопрос по созданию распознающих алгоритмов, которые могут быть применены для решения прикладных задач распознавания при больших размерностях признакового пространства и наличии взаимосвязанности признаков.

3 МАТЕРИАЛЫ И МЕТОДЫ

Для решения задачи, сформулированной в разделе 1, предлагается подход, который является логическим продолжением работ академика Ю. И. Журавлева и его учеников. На базе этого подхода разработана модель модифицированных распознающих операторов, основанных на выявлении независимых подмножеств взаимосвязанных признаков и выделении предпочтительной модели зависимости для каждого подмножества сильносвязанных признаков. В качестве исходной модели рассматриваются распознающие операторы типа потенциальных функций [1–3].

Предлагаемая модель распознающих операторов включает следующие основные этапы.

1. *Выделение подмножеств сильносвязанных признаков.* На этом этапе определяется система «независимых» подмножеств признаков, состав которой будет зависеть от параметра n' . В зависимости от способа задания меры близости между подмножествами сильносвязанных признаков (Ω_p и Ω_q) и функционала качества классификационного анализа можно получить разнообразные процедуры выделения независимых множеств сильносвязанных признаков. Процедура выделения подмножеств сильносвязанных признаков более подробно рассмотрена в [14].

2. *Формирование набора репрезентативных признаков.* В результате выполнения данного этапа получаем сокращенное пространство признаков, размерность которого намного меньше исходного ($n' < n$). Далее сформированное пространство признаков обозначим через Y ($Y = (y_1, \dots, y_{n'})$).

Процедура выделения набора репрезентативных признаков более подробно рассмотрена в [15].

3. *Определение моделей зависимости в каждом подмножестве признаков для класса K_j ($j = \overline{1, l}$).* Пусть x_i – произвольный признак, принадлежащий подмножеству Ω_q . Предполагается, что элементы Ω_q линейно упорядочены по индексу признаков (т. е. $x_i < x_j$, если $i < j$).

Далее, нулевым элементом (x_0) подмножества Ω_q считается y_q , остальные элементы обозначаются через x_i

$$(N_q = |\Omega_q|; i = 1, \dots, N_q - 1).$$

Пусть модель зависимости в Ω_q имеет вид

$$y_q = F(\bar{c}, x_i), \quad x_i \in \Omega_q \setminus y_q,$$

где \bar{c} – вектор неизвестных параметров, x_i – произвольный признак, принадлежащий подмножеству Ω_q , F – функция из некоторого заданного класса $\{F\}$.

Вычисленные значения вектора неизвестных параметров \bar{c} определяют модель зависимости в подмножестве признаков Ω_q для класса K_j ($j = \overline{1, l}$). В зависимости от задания параметрического вида $F(\bar{c}, x)$ и метода определения \bar{c} получаем разнообразные модели зависимости во множестве признаков Ω_q ($q = \overline{1, n'}$).

4. *Выделение предпочтительных моделей зависимости.* Поиск предпочтительной модели зависимости в Ω_q осуществляется на основе оценки доминированности рассматриваемой модели, которая разделяет объекты, принадлежащие множеству \tilde{S}^m , на два подмножества K_1 и K_2 [16]:

$$T_{qi} = \frac{L_1 \sum_{S \in \tilde{K}_2} |y_q - F_q(\bar{c}, x_i)|}{L_2 \sum_{S \in \tilde{K}_1} |y_q - F_q(\bar{c}, x_i)|}, \quad L_1 = |\tilde{K}_1|, \quad L_2 = |\tilde{K}_2|.$$

Чем больше величина T_{qi} , тем больше отдается предпочтение i -ой модели зависимости в подмножестве Ω_q . Если несколько моделей получают одинаковое предпочтение, то выбирается любая из них.

В результате выполнения данного этапа определяется предпочтительная модель зависимости для подмножества признаков Ω_q . Далее рассматриваются только эти модели зависимости.

5. *Определение функции различия между объектами S_u и S .* Пусть заданы два объекта S_u и S в пространстве $X = (x_1, \dots, x_i, \dots, x_n)$:

$$S_u = (a_{1u}, \dots, a_{nu}) \quad \text{и} \quad S = (b_1, \dots, b_n).$$

На данном этапе определяется расстояние между этими объектами по всем подмножествам Ω_q ($q = \overline{1, n'}$):

$$d(S_u, S) = \sum_{q=1}^{n'} \tau_q d_q(S_u, S),$$

$$d_q(S_u, S) = |\phi_q - \psi_q|, \quad \phi_q = (a_{ui} + b_i)/2,$$

$$\psi_q = (F_q(\bar{c}, a_{ui}) + F_q(\bar{c}, b_i))/2,$$

где τ_q – параметр распознающего оператора.

6. *Определение функции близости между объектами S_u и S .* На этом этапе функция близости между объектами S_u и S определяется как потенциальная функция второго типа, например [1, 3],

$$U(S_u, S) = \frac{1}{1 + \xi d(S_u, S)},$$

где ξ – параметр распознающего оператора.

7. *Вычисление оценки принадлежности к классу K_j .* Оценка принадлежности объекта S к классу K_j ($j = 1, 2$) определяется как функция оценки близости:

$$\mu(S) = \sum_{S_u \in \tilde{K}_2} \gamma_u U(S_u, S) - \sum_{S_u \in \tilde{K}_1} \gamma_u U(S_u, S),$$

где γ_u – параметр распознающего оператора.

Таким образом, определена модель распознающих операторов типа потенциальных функций, основанных на оценке взаимосвязанности признаков. Произвольный оператор B из этой модели полностью определяется заданием набора параметров $\tilde{\pi}$ [17]. Совокупность всех распознающих операторов из предлагаемой модели обозначаем через $B(\tilde{\pi}, S)$. Определение наилучшего распознающего оператора в рамках рассмотренной модели осуществляется в пространстве параметров $\tilde{\pi}$. Наилучший оператор $B(\tilde{\pi}_o, S)$ выбирается на основе поиска минимального значения функционала качества распознающего оператора:

$$\varphi(\tilde{\pi}) = \frac{1}{m} \sum_{S \in V_o} \theta \| \chi(S) - C(B(\tilde{\pi}, S)) \|, \quad \theta(x) = \begin{cases} 1, & \text{при } x = 0; \\ 0, & \text{при } x \neq 0, \end{cases} \quad (3)$$

где $\| \cdot \|$ – норма булевого вектора.

4 ЭКСПЕРИМЕНТЫ

Экспериментальное исследование работоспособности предложенной модели распознающих операторов осуществлено на примере решения модельной задачи.

Исходные данные распознаваемых объектов для модельной задачи сгенерированы в пространстве зависимых признаков. Количество классов в данном эксперименте равно двум. Объем исходной выборки – 1100 реализаций (по 550 реализаций для объектов каждого класса). Количество признаков в модельном примере равно 20. Число подмножеств сильносвязанных признаков – 5.

В качестве испытуемых моделей распознающих операторов были выбраны: классическая модель распознающих операторов типа потенциальных функций (B_1), модель распознающих операторов, основанных на вычислении оценок (B_2), и модель (B_3), предлагаемая в настоящей работе. Сравнительный анализ перечисленных моделей распознающих операторов при решении рассмотренной задачи проведен по следующим критериям:

- точность распознавания объектов контрольной выборки;
- время, затраченное на обучение;
- время, затраченное на распознавание объектов из контрольной выборки.

Для вычисления указанных критериев при решении рассматриваемой задачи, в целях исключения удачного (или неудачного) разбиения исходной выборки V на две равные части V_o и V_k ($V = V_o \cup V_k$, V_o – выборка для обучения, V_k – выборка для контроля), используется метод скользящего контроля [18]. В этом методе исходная выборка объектов случайным образом разбивается на 110 непересекающихся блоков, включающих по 10 объектов каждый. При этом требуется, чтобы во всех блоках сохранилась пропорция по количеству объектов, принадлежащих к разным классам. В результате получается, что в каждом блоке по 5 объектов каждого класса. Процесс скользящего контроля по этим блокам включает несколько шагов. На каждом шаге выбирают 10 из 110 блоков в качестве обучающей выборки, и на этой выборке обучаются распознающие операторы с заданными параметрами.

Обученный таким образом распознающий оператор проверяется на остальных 100 блоках (контрольной выборке). В результате каждой проверки определяется и фиксируется оценка качества распознающих операторов по указанным критериям. При выполнении каждого очередного шага для оценки качества распознающих операторов из контрольной и обучающей выборок выбирают по одному блоку и меняют их местами. В целях исключения повторного использования объектов обучающей выборки соответствующие блоки маркируются, и при выборе кандидатов для включения в обучающую выборку они не участвуют. После завершения процедуры скользящего экзамена точность распознавания и временные показатели определялись как средние. Эксперименты проводились на компьютере Pentium IV Dual Core 2,2 GHz с объемом оперативной памяти 1 Gb.

5 РЕЗУЛЬТАТЫ

Основным результатом данной статьи является модель распознающих операторов, основанных на оценке взаимосвязанности признаков (см. раздел 3). Построение данной модели выполнено в рамках модели алгоритмов типа потенциальных функций и заключалось в выявлении независимых подмножеств взаимосвязанных признаков с последующим выделением предпочтительной модели зависимости для каждого подмножества сильносвязанных признаков. Данная модель распознающих операторов имеет большое значение при решении задачи распознавания образов в условиях взаимосвязанности признаков и позволяет расширить область применения метода потенциальных функций.

В целях проверки работоспособности предложенной модели операторов проведено экспериментальное исследование при решении модельной задачи. Как уже отмечалось ранее, модельная задача решалась с применением классической модели распознающих операторов типа потенциальных функций (B_1), модели распознающих операторов, основанных на вычислениях оценок, (B_2) и предлагаемой модели (B_3). Точность распознавания в процессе обучения (см. формула 3 в разделе 3) для B_1 равна 94,8%, для B_2 – 97,6%, и, наконец, для B_3 – 99,1%. Результаты решения рассматриваемой задачи с применением B_1 , B_2 и B_3 в процессе контроля приведены в табл. 1.

Сравнение этих результатов показывает (см. табл. 1), что предложенная модель распознающих операторов $|B_3|$ позволила повысить точность распознавания объектов, описанных в пространстве взаимосвязанных признаков (более чем на 10% выше, чем по B_1 и B_2). Это объясняется тем, что в моделях B_1 и B_2 не учитывается взаимосвязанность признаков. Однако для модели B_3

Таблица 1 – Результаты решения задачи с помощью различных распознающих операторов

Распознающий оператор	Время (в сек.)		Точность распознавания (в %)
	обучения	распознавания	
B_1	4,141	0,014	80,5
B_2	6,127	0,012	82,7
B_3	9,041	0,002	93,4

имеет место некоторое увеличение времени обучения за счет реализации дополнительной процедуры формирования независимых подмножеств взаимосвязанных признаков.

6 ОБСУЖДЕНИЕ

Разработанные операторы отличаются от традиционных распознающих операторов типа потенциальных функций тем, что они основаны на оценке взаимосвязанности признаков. Поэтому эти алгоритмы целесообразно использовать в тех случаях, когда между признаками обнаруживается какая-нибудь зависимость. Несомненно, что эта зависимость должна отличаться для объектов каждого класса. Это позволяет описать объекты каждого класса индивидуальной моделью. Если зависимость между признаками слаба, то используется классическая модель распознающих операторов (например, модель, рассмотренная в [2, 3]). Следовательно, предложенная модель распознающих операторов не является альтернативной моделям распознающих операторов типа потенциальных функций, а только дополняет их.

В случае, когда между признаками всех рассматриваемых объектов обнаруживается достаточно сильная зависимость, то в процессе формирования набора репрезентативных признаков (описанных на первом и втором этапах) исключаются признаки, повторяющие одну и ту же информацию, что обеспечивает выбор признаков, достаточно хорошо представляющих все те признаки, которые не содержатся в данном наборе [4].

Результаты приведенного экспериментального исследования показывают, что предложенная модель распознающих операторов позволяет решать задачу распознавания образов более точно в условиях взаимосвязанности признаков. При этом показатель времени, затраченного на распознавание объектов, намного меньше, чем тот же показатель B_1 и B_2 (см. табл. 1). Это связано с тем, что в предложенных распознающих операторах для распознавания объекта используются только предпочтительные модели зависимости, что обусловило повышение скорости распознавания объектов. Поэтому эту модель можно использовать при разработке систем распознавания, функционирующих в режиме реального времени. Вместе с тем необходимо отметить тот факт, что время, израсходованное на обучение алгоритма, увеличилось, т. к. для построения оптимального распознающего оператора требуется оптимизировать большее число параметров, чем при использовании традиционной модели распознающих операторов, в частности B_1 и B_2 .

ВЫВОДЫ

1. Анализ существующих моделей распознающих операторов показал, что в основном они ориентированы на решение прикладных задач распознавания объектов с независимыми признаками. В большинстве прикладных задач распознавания, встречающихся в науке, технике и производстве, рассматриваемые образы характеризуются взаимосвязанными признаками. Хотя разработаны различные алгоритмы решения задач распознавания для таких образов, они оказались малоэффективными с точки зрения точности и сложности вычисления.

2. Предложен новый подход для построения модели распознающих операторов и на основе этого подхода построена модифицированная модель распознающих операторов типа потенциальных функций. Предложенная модель операторов опирается на оценку взаимосвязанности признаков. Она позволяет расширить модели распознающих операторов типа потенциальных функций и увеличивает область применения метода потенциальных функций. Данная модель ориентирована на решение задачи распознавания объектов, заданных в пространстве взаимосвязанных признаков.

3. Результаты решения модельной задачи показали, что предложенная модель распознающих операторов улучшает точность и значительно сокращает число вычислительных операций посредством распознавания неизвестного объекта, заданного в пространстве взаимосвязанных признаков. Она может быть использована при составлении различных программных комплексов, ориентированных на решение прикладных задач распознавания.

БЛАГОДАРНОСТИ

Работа выполнена в рамках гранта А5-ФА-Ф022 Государственного комитета по координации развития науки и технологий Республики Узбекистан.

СПИСОК ЛИТЕРАТУРЫ

1. Айзерман М. А. Метод потенциальных функций в теории обучения машин / М. А. Айзерман, Э. М. Браверманн, Л. И. Розоноэр. – М. : Наука, 1970. – 348 с.
2. Журавлев Ю. И. Избранные научные труды / Ю. И. Журавлев. – М. : Магистр, 1998. – 420 с.
3. Журавлев Ю. И. Распознавание. Математические методы. Программная система. Практические применения / Ю. И. Журавлев, В. В. Рязанов, О. В. Сенько. – М. : Фазис, 2006. – 159 с.
4. Камиллов М. М. Современное состояние вопросов построения моделей алгоритмов распознавания / М. М. Камиллов, Н. М. Мирзаев, С. С. Раджабов. // Научный журнал: Химическая технология. Контроль и управление. – 2009. – № 2. – С. 21–27.
5. Загоруйко Н. Г. Прикладные методы анализа данных и знаний / Н. Г. Загоруйко. – Новосибирск : ИМ СО РАН, 1999.
6. Лбов Г. С. Логические решающие функции и вопросы статистической устойчивости решений / Г. С. Лбов, Н. Г. Старцева. – Новосибирск : Изд-во ИМ СО РАН, 1999. – 211 с.
7. Шлезингер М. Десять лекций по статистическому и структурному распознаванию / М. Шлезингер, В. Главач. – К. : Наукова думка, 2004. – 545 с.
8. Потапов А. С. Распознавание образов и машинное восприятие / А. С. Потапов. – СПб. : Политехника, 2007. – 548 с.
9. Duda R. O. Pattern Classification, Second Edition / R. O. Duda, P. E. Hart, D. G. Stork. – New York : John Wiley, Inc., 2001. – 680 p.
10. Vapnik V. Statistical Learning Theory / V. Vapnik. – New York : John-Wiley Sons, Inc., 1998. – 732 p.
11. Theodoridis S. Pattern Recognition: Theory and Applications, 4th edition / S. Theodoridis, K. Koutroumbas. – New York : Academic Press, 2009. – 957 p.
12. Dougherty G. Pattern Recognition and Classification: An Introduction / G. Dougherty. – New York: Springer, 2013. – 196 p.
13. Фазылов Ш. Х. Построение модели алгоритмов вычисления оценок с учетом большой размерности признакового пространства / Ш. Х. Фазылов, Н. М. Мирзаев, С. С. Раджабов // Вестник СГТУ. – Саратов, 2012. – № 1 (64), Вып. 2. – С. 274–279.
14. Мирзаев Н. М. Алгоритмы выделения подмножеств сильно связанных признаков / Н. М. Мирзаев // Вопросы кибернетики : сб. науч. тр. – Ташкент : ИМИТ АН РУз, 2008. – Вып. 177. – С. 99–104.
15. Мирзаев О. Н. Выделение репрезентативных признаков при построении алгоритмов распознавания / О. Н. Мирзаев // Проблемы информатики и энергетики. – Ташкент, 2008. – № 6. – С. 23–27.

16. Модель распознающих операторов, основанных на принципе ближайшего соседа, в условиях взаимосвязанности признаков / Ш. Х. Фазылов, Н. М. Мирзаев, С. С. Раджабов, И. К. Каримов // Информатика и системы управления. – Благовещенск, 2012. – № 4. – С. 34–42.
17. Мирзаев Н. М. О параметризации моделей алгоритмов распознавания, основанных на оценке взаимосвязанности признаков / Фазылов Ш. Х.¹, Мирзаев Н. М.², Мирзаев О. Н.³
- Н. М. Мирзаев, С. С. Раджабов, Т. С. Жумаев // Проблемы информатики и энергетики. – Ташкент, 2008. – № 2–3. – С. 23–27.
18. Vorontsov K. V. Computable Combinatorial Overfitting Bounds / K. V. Vorontsov, A. I. Frey, E. A. Sokolov // Journal of Machine Learning and Data Analysis. – 2013. – Т. 1, № 6. – С. 734–743. Статья поступила в редакцию 09.11.2015. После доработки 02.12.2015.

¹Д-р техн. наук, профессор, завідувач лабораторії Центру розробки програмних продуктів і апаратно-програмних комплексів при Ташкентському Університеті Інформаційних Технологій, Ташкент, Узбекистан

²Канд. техн. наук, старший науковий співробітник Центру розробки програмних продуктів і апаратно-програмних комплексів при Ташкентському Університеті Інформаційних Технологій, Ташкент, Узбекистан

³Молодший науковий співробітник Центру розробки програмних продуктів і апаратно-програмних комплексів при Ташкентському Університеті Інформаційних Технологій, Ташкент, Узбекистан

ПОБУДОВА РОЗПІЗНАЮЧИХ ОПЕРАТОРІВ В УМОВАХ ВЗАЄМОП'ЯЗАНОСТІ ОЗНАК

Розглянуто задачу розпізнавання образів, заданих у просторі взаємозалежних ознак. Запропоновано новий підхід до побудови моделі розпізнавальних операторів, що враховує взаємозв'язок заданих ознак. При цьому побудова моделі здійснена для розпізнавальних операторів типу потенційних функцій. Основна ідея запропонованого підходу полягає у формуванні незалежних підмножин взаємозалежних ознак і виділенні кращої моделі залежності для кожної підмножини сильнозв'язаних ознак. Аналіз отриманих результатів показує, що розглянуті розпізнавальні оператори використовуються в тих випадках, коли між ознаками об'єктів, що належать до доного класу, існує деяка залежність. При слабкому вираженні цієї залежності використовується класична модель розпізнавальних операторів. Основною перевагою запропонованих розпізнавальних операторів є поліпшення точності і значне скорочення обсягу обчислювальних операцій при розпізнаванні невідомих об'єктів, що дозволяє застосувати їх при побудові розпізнавальних систем, які працюють у режимі реального часу.

Ключові слова: розпізнавання образів, модель розпізнавальних операторів, потенційна функція, залежність ознак, підмножина сильнозв'язаних ознак, репрезентативна ознака, краща модель залежності.

Fazilov Sh. Kh.¹, Mizraev N. M.², Mizraev O. N.³

¹D. Sc., Professor, Head of the Laboratory of the Centre for development of software and hardware-program complexes at Tashkent University of information technologies, Tashkent, Uzbekistan

²PhD, Senior Researcher at the Centre for Development of Software and Hardware-program Complexes at Tashkent University of information technologies, Tashkent, Uzbekistan

³Junior researcher at the Centre for Development of Software and Hardware-program Complexes at Tashkent University of information technologies, Tashkent, Uzbekistan

BUILDING OF RECOGNITION OPERATORS IN CONDITION OF FEATURES' CORRELATIONS

The problem of recognizing of patterns given in the space of correlated features is considered. The new approach to the building of model of recognition operators, which considers the correlation of given features, is proposed. The building of the model is carried out for potential function type recognition operators. The main idea of the proposed approach is formation of uncorrelated subsets of strongly correlated features and extracting preferred correlation model for each of subsets of strongly correlated features. Analysis of the results shows that the considered recognition operators are used in cases when there is a certain correlation between objects belonging to the same class. When the expression of this relationship is weak, classical model of recognition operators is used. The main advantage of the proposed recognition operators is to improve the accuracy and the significant reduction in the volume of computational operations in recognition of unknown objects, which allows them to use when building recognition systems working in real time.

Keywords: pattern recognition, model of recognition operators, potential function, features' correlations, subset of strongly correlated features, representative feature, preferred correlation model.

REFERENCES

1. Ayzeman M. A., Braverman E. M., Ayzeman M. A., Rozonoer L. I. Metod potentsialnyh funktsiy v teorii obucheniya mashin. Moscow, Nauka, 1970, 348 p.
2. Zhuravlev Yu. I. Izbrannyye nauchnyye trudy. Moscow, Magistr, 1998, 420 p.
3. Zhuravlev Yu. I., Ryazanov V. V., Senko O. V. Raspoznavaniya. Matematicheskiye metody. Programmnaya sistema. Prakticheskiye primeneniya. Moscow, Fazis, 2006, 159 p.
4. Kamilov M. M., Mirzayev N. M., Radjabov S. S. Sovremennoye sostoyaniye voprosov postroyeniya modeley algoritmov raspoznavaniya, *Nauchnyy zhurnal: Himicheskaya tehnologiya. Kontrol i upravleniye*, 2009, No. 2, pp. 21–27.
5. Zagoruyko N. G. Prikladnyye metody analiza dannyh i znaniy. Novosibirsk, IM SO RAN, 1999.
6. Lbov G. S., Startseva N. G. Logicheskiye reshayushiy funktsii i voprosy statisticheskoy ustoychivosti resheniy. Novosibirsk, Izd-vo IM SO RAN, 1999, 211 p.
7. Shlezinger M., Glavach V. Desyat lektsiy po statisticheskomu i strukturnomu raspoznavaniyu. Kiev, Naukova dumka, 2004, 545 p.
8. Potapov A. S. Raspoznavaniye obrazov i mashinnoye vospriyatiye. Sankt-Peterburg, Politehnika, 2007, 548 p.
9. Duda R. O. Hart P. E. Stork D. G. Pattern Classification, Second Edition. New York, John Wiley, Inc., 2001, 680 p.
10. Vapnik V. Statistical Learning Theory. New York, John-Wiley Sons, Inc., 1998, 732 p.
11. Theodoridis S., Koutroumbas K. Pattern Recognition: Theory and Applications, 4th edition. New York, Academic Press, 2009, 957 p.
12. Dougherty G. Pattern Recognition and Classification: An Introduction. New York, Springer, 2013, 196 p.
13. Fazylov Sh. H., Mirzayev N. M., Radjabov S. S. Postroyeniye modeli algoritmov vychisleniya otsenok s uchytom bolshoy razmernosti priznakovogo prostranstva, *Vestnik SGTU*, Saratov, 2012. – No. 1 (64), Vyp. 2, pp. 274–279.
14. Mirzaev N. M. Algoritmy vydeleniya podmnozhestv silnosvyazannyh priznakov, *Voprosy kibernetiki: Sb. nauch. tr.* Tashkent, IMIT AN RUz, 2008, Vyp. 177, pp. 99–104.
15. Mirzaev O. N. Vydeleniye perpezentativnyh priznakov pri postroyenii algoritmov raspoznavaniya, *Problemy informatiki i energetiki*. Tashkent, 2008, No. 6, pp. 23–27.
16. Fazylov Sh. H., Mirzayev N. M., Radjabov S. S., Karimov I. K. Model raspoznavayushih operatorov, osnovannyh na printsipe bliyayshego sosedya, v usloviyah vzaimosvyazannosti priznakov, *Informatika i sistemy upravleniya*. Blagoveshensk, 2012, No. 4, pp. 34–42.
17. Mirzayev N. M., Radjabov S. S., Zhumayev T. S. O parametrizatsii modeley algoritmov raspoznavaniya, osnovannyh na otsenke vzaimosvyazannosti priznakov, *Problemy informatiki i energetiki*, Tashkent, 2008, No. 2–3, pp. 23–27.
18. Vorontsov K. V., Frey A. I., Sokolov E. A. Computable Combinatorial Overfitting Bounds, *Journal of Machine Learning and Data Analysis*, 2013, Vol. 1, No. 6, pp. 734–743.