https://orcid.org/0000-0001-9952-9438
https://orcid.org/0000-0001-6737-7134
https://orcid.org/0000-0002-1792-9401

УДК 004.93'1

Т.Б. МАРТИНЮК[*], Л.М. КУПЕРШТЕЙН[*], М.Д. КРЕНЦІН[*]

ОСОБЛИВОСТІ ПРОЦЕСУ КЛАСИФІКАЦІЇ ОБ'ЄКТІВ НА БАЗІ ДИСКРИМІНАНТНИХ ФУНКЦІЙ

[*]Вінницький національний технічний університет, м. Вінниця, Україна

Анотація. *У роботі запропоновано альтернативний підхід до обробки елементів лінійних дискримінантних функцій для класифікації об'єктів. В основі цього підходу використано метод обробки векторного масиву даних за різницевими зрізами, який розповсюджено на стовпці матричного представлення елементів дискримінантних функцій. Це, у свою чергу, свідчить про просторово-розподілений характер обробки елементів матриці розмірністю* $m \times n$*, де* m *– кількість класів,* n *– розмірність вектора вхідних даних. Лінійність представлення кожної дискримінантної функції у наборі дискримінантних функцій дозволяє не тільки подати їх доданки у вигляді елементів матриці, але й застосувати до їх обробки правила щодо таких властивостей, як асоціативність та комутативність складових (доданків) суми. В результаті замінюються такі складні процедури, як багатооперандне підсумовування при формуванні кожної дискримінантної функції з подальшим їх порівнянням між собою для виявлення серед них максимальної за значенням. У запропонованому методі у кожному циклі обробки базовою операцією є зменшення одночасно на величину подібності однойменних елементів у стовпцях матриці, що в результаті дозволяє визначити максимальну за значенням дискримінантну функцію. При цьому також задіяно операцію транспозиції (просування) до краю всіх поточних нульових елементів у кожному рядку матриці. Даний підхід застосовано до лінійних дискримінантних функцій як апарата статистичного методу класифікації об'єктів. У результаті за відомим вирішальним правилом максимальна за значенням дискримінантна функція визначає відповідний клас, до якого належить вхідний об'єкт, поданий у вигляді* n*-вимірного вектора ознак. Одним із варіантів апаратного застосування такого класифікатора є підсистема підтримки прийняття рішень при медичному діагностуванні.*

Ключові слова: *лінійна дискримінантна функція, класифікація об'єктів, різницевий зріз, просторово-розподілена обробка.*

Abstract. *An alternative approach to the processing of elements of linear discriminant functions for the classification of objects is proposed in the paper. This approach is based on the method of processing a vector array of data by differential slices which is distributed in the column of matrix representation of elements of discriminant functions. In its turn, it indicates the spatially distributed nature of the processing of matrix elements with dimension* $m \times n$*, where* m *is the number of classes,* n *is the dimension of the input data vector. The linearity of the representation of each discriminant function in a set of discriminant functions allows not only to represent their terms in the form of matrix elements, but also apply to their processing some rules concerning such properties as associativity and commutativity of components (terms) of the sum. As a result, such complex procedures as multi-operand summation in the formation of each discriminant function are replaced. Moreover, it is followed by their comparison with each other with the aim of identifying the one with the maximum value. In each processing cycle of the proposed method, the basic operation is a simultaneous reduction in the magnitude of the elements similarity of the same name in the columns of the matrix, which allows determining the maximum value of the discriminant function. The operation of transposition (advancement) to the edge of all current zero elements in each row of the matrix is also involved. This approach is applied to linear discriminant functions as an apparatus of the statistical method of object classification. Consequently, according to the known decision rule, the maximum discriminant function determines the corresponding class to which the input object, represented as an n-dimensional feature vector, belongs. One of the options for the hardware application of such classifier is the decision support subsystem for medical diagnosis.*

Keywords: *linear discriminant function, object classification, differential slice, spatially distributed processing.*

DOI: 10.34121/1028-9763-2021-3-81-87

1. Вступ

Серед базових процедур розпізнавання образів однією з найбільш затребуваних є класифікація об'єктів [1]. Це пов'язано з її широкою областю застосування. Найбільш показовими у цьому плані є ті області, де активно використовуються експертні технології прийняття рішень [2], оскільки класифікатори входять до складу підсистеми підтримки прийняття рішень, наприклад, при медичному діагностуванні [3, 4], а також у перспективній області ідентифікації користувачів [5].

Актуальність тематики. Особливий інтерес до створення та вдосконалення класифікаторів різного призначення пов'язаний з активним використанням нейротехнологій [6, 7]. Наприклад, нейромережевий підхід при медичному діагностуванні [8] дозволяє зробити процес формування рекомендацій підсистемою підтримки прийняття рішень більш гнучким, оскільки існує можливість навчання такої підсистеми через інтелектуальний аналіз даних, а також розширення переліку симптомів та їх уточнення [9–11]. При цьому добре зарекомендував себе при класифікації підхід із використанням дискримінантного аналізу, який базується на початкових даних з обмеженим статистичним описом, тобто без застосування ймовірнісних моделей, які потребують точних даних [12].

Метою статті є аналіз математичної моделі удосконаленого методу обробки елементів дискримінантних функцій у процесі класифікації об'єктів.

2. Перетворення матриці елементів дискримінантних функцій

Методи лінійного дискримінантного аналізу дозволяють вибрати проекцію простору об'єктів на просторі ознак таким чином, щоб мінімізувати внутрішньокласову та максимізувати міжкласову відстань у просторі ознак [10–12]. Отже, класифікатор, що використовує критерій мінімуму середньоквадратичної відстані між векторами образу X і середнім образом $\overline{X}_i$, який отримано за експертними даними, визначає i-й клас за максимумом величини i-ої дискримінантної функції (ДФ) $g_i(X)$ [10, 12].

Таким чином, у процесі класифікації використано n-вимірний вхідний образ X, m класів $C = \{C_1, ..., C_m\}$ та систему лінійних ДФ $g_i(X)$, $i = \overline{1,m}$ вигляду

$$
\begin{aligned}
g_i(X) &= w_{11}x_1 + ... + w_{1n}x_n - \theta_1, \\
&\dots \\
g_m(X) &= w_{m1}x_1 + \ldots + w_{mn}x_n - \theta_m,
\end{aligned}
\tag{1}
$$

де θ_i – вільний елемент i-ої ДФ, w_{ij} – i-та вага j-го елемента x_j вхідного образу X, $j = \overline{1,n}$. А результат класифікації формується за максимум i-ої лінійної ДФ, тобто

$$\max_i g_i(X) = g_l(X) \Rightarrow X \in C_l. \tag{2}$$

Формування системи лінійних ДФ (1) можна представити у вигляді вектор-стовпця $G(X)$ як результат множення матриці ваг $\boldsymbol{W}$ на вектор X, нехтуючи вільними елементами θ_i, $i = \overline{1,m}$, таким чином:

$$G(X)=\begin{bmatrix} g_1(X) \\ \ldots \\ g_m(X) \end{bmatrix} = \boldsymbol{W}\cdot X = \begin{bmatrix} w_{11}x_1+\ldots+w_{1j}x_j+\ldots+w_{1n}x_n \\ \ldots \\ w_{m1}x_1+\ldots+w_{mj}x_j+\ldots+w_{mn}x_n \end{bmatrix}. \quad (3)$$

У подальшому вибирається максимальний елемент $g_l(X)$ вектор-стовпця $G(X)$ за формулою (2) з визначенням класу C_l. Це приклад класичного варіанта класифікації за лінійними ДФ [10].

Разом з тим, використання методу обробки масиву чисел із формуванням різницевих зрізів (РЗ) дозволяє зробити цей процес класифікації двовимірним і просторово-розподіленим. Математичну модель методу різницево-зрізової (РЗ) обробки векторних масивів даних представлено в роботі [13], а ефективність такого підходу досліджено в роботі [14].

Для матричних масивів даних математичну модель обробки за РЗ для класифікації за лінійними ДФ можна описати і дослідити, якщо результат матрично-векторного множення (3) представити у вигляді початкової матриці $\boldsymbol{A}^0$ елементів ДФ:

$$\boldsymbol{A}^0 = \begin{bmatrix} w_{11}x_1\ldots w_{1j}x_j\ldots w_{1n}x_n \\ \ldots \\ w_{m1}x_1\ldots w_{mj}x_j\ldots w_{mn}x_n \end{bmatrix}. \quad (4)$$

У подальшому матрицю $\boldsymbol{A}^0$ (4) будемо розглядати як

$$\boldsymbol{A}^0 = \begin{bmatrix} a_{11}^0\ldots a_{1j}^0\ldots a_{1n}^0 \\ \ldots \\ a_{m1}^0\ldots a_{mj}^0\ldots a_{mn}^0 \end{bmatrix}, \quad (5)$$

де

$$a_{ij}^0 = w_{ij}x_j, i=\overline{1,m}, j=\overline{1,n}. \quad (6)$$

Крім того, для подальшої обробки елементів матриці $\boldsymbol{A}^0$ (5) доцільно виділяти її стовпці A_j^0, орієнтуючись на обробку за РЗ векторних даних [4, 13].

Отже, представлення елементів системи ДФ $g_1(X),..., g_m(X)$ у вигляді матриці $\boldsymbol{A}^0$ (5) з елементами a_{ij}^0 (6) дозволяє розглядати їх у трьох варіантах:

а) як доданки відповідних i-х ДФ $g_i(X)$ (1);

б) як елементи відповідних i-х рядків A_i^0, $i=\overline{1,m}$ матриці $\boldsymbol{A}^0$ (5);

в) як елементи однойменних j-х стовпців A_j^0, $j=\overline{1,n}$ матриці $\boldsymbol{A}^0$ (5).

Зважаючи на перший варіант, до всіх елементів a_{ij}^0 рядків A_i^0 матриці $\boldsymbol{A}^0$ (5) правомірно застосувати такі властивості суми доданків, як комутативність та асоціативність. У результаті всі елементи рядків A_i^0 можна реверсивно зсувати в межах кожного рядка, оскільки їх сума, а, отже, і величина відповідної i-ої ДФ не змінюється. Так реалізується властивість комутативності суми доданків.

Крім того, треба врахувати, що для визначення максимальної суми елементів кожного рядка A_i^0 як значення відповідної i-ої ДФ за виразом (2), не важлива величина кожної ДФ, а тільки співвідношення (>) між ними. Тому всі ДФ правомірно одночасно зменшувати на однакову величину, а саме зменшувати всі однойменні елементи кожного стовпця A_j^0 матриці $\boldsymbol{A^0}$ (5) на однакову величину, але різну у кожному стовпці. Так реалізується властивість асоціативності суми доданків у вигляді

$$x+(y \pm z)=(x+y) \pm z,$$

що не протирічить співвідношенню для величин: $A > B = A - C > B - C$.

Такі самі дії правомірно виконувати над елементами наступних сформованих матриць $\boldsymbol{A^t}$, $t=\overline{1,N}$, де N – кількість циклів обробки. Якщо поступово в кожному циклі зменшувати всі однойменні елементи в усіх стовпцях A_j^t поточних матриць $\boldsymbol{A^t}$, то першими обнуляться ті рядки A_i^t матриць $\boldsymbol{A^t}$, які відповідають ДФ із найменшою сумою елементів a_{ij}^0 початкової матриці $\boldsymbol{A^0}$ (4). Таким чином можна визначити максимальну за величиною ДФ $g_i(X)$, обнуляючи мінімальні ДФ.

Отже, запропонований підхід до обробки елементів ДФ у вигляді матриці, по-перше, представляє собою двовимірну просторово-розподілену обробку по стовпцях та рядках поточних матриць, а, по-друге, дозволяє в одному процесі обробки елементів матриці визначити максимальну ДФ. У класичному варіанті за виразами (1) і (2) такий результат можна отримати після двох послідовних процедур: спочатку сформувати всі ДФ, а потім у процесі їх порівняння між собою визначити максимальну ДФ.

Крім того, при будь-якому методі класифікації обов’язковим є вибір метрики, тобто міри подібності між багатовимірними об’єктами. У випадку використання методу обробки за РЗ такою мірою подібності можна прийняти мінімальний ненульовий елемент як загальну частину всіх елементів векторного масиву (РЗ) [13].

3. Процес просторово-розподіленої обробки матричного масиву

Відомий метод обробки за РЗ векторних масивів у даному випадку застосовується до стовпців A_j^{t-1} поточної матриці $\boldsymbol{A^{t-1}}$, $t=\overline{1,N}$ вигляду

$$A^{t-1}=[A_1^{t-1} \dots A_n^{t-1}]=\begin{bmatrix} a_{11}^{t-1} \dots a_{1j}^{t-1} \dots a_{1n}^{t-1} \\ \dots \\ a_{m1}^{t-1} \dots a_{mj}^{t-1} \dots a_{mn}^{t-1} \end{bmatrix}. \tag{7}$$

Ітераційний процес виконується одночасно над всіма n стовпцями A_j^{t-1}, які у даному випадку розглядаються як n РЗ і формують двовимірний РЗ. Починається процес обробки з початкової матриці $\boldsymbol{A^0}$ (5), причому в кожному t-му циклі виконуються три базові операції [4]:

1) визначення поточного порогу обробки (міри подібності) у кожному j-му стовпці A_j^{t-1} матриці $\boldsymbol{A}^{t-1}$ у вигляді мінімального значущого елемента:

$$q_j^{t-1}=\min a_{ij}^{t-1},\ t=\overline{1,N}; \tag{8}$$

2) формування поточного двовимірного РЗ у вигляді матриці $\overline{\boldsymbol{A}}^t$, в якій кожний елемент j-го стовпця $\overline{A}_j^t$ розраховується таким чином:

$$\overline{a}_{ij}^{t} = a_{ij}^{t-1} - q_{ij}^{t-1}; \tag{9}$$

3) формування впорядкованої матриці $\boldsymbol{A}^t$ в результаті транспозиції (просування) праворуч до краю всіх поточних нульових елементів у кожному i-му рядку $\overline{A}_i^t$ матриці $\overline{\boldsymbol{A}}^t$:

$$A_i^t = \overrightarrow{Tr}(\overline{A}_i^t). \tag{10}$$

Якщо перші дві базові операції реалізують властивість асоціативності, то третя базова операція, яка реалізує властивість комутативності, необхідна для усунення впливу нульових елементів, що з’являються у стовпцях поточних матриць $\boldsymbol{A}^{t-1}$, щодо визначення поточного порогу обробки q_j^{t-1} (8).

Отже, такі дії призводять до поступового зменшення елементів усіх рядків A_i^{t-1} поточних матриць $\boldsymbol{A}^{t-1}$, а отже, і сум доданків вигляду (3) на однакову (найменшу) величину однойменних елементів, і в результаті визначать найменшу поточну суму доданків, тобто найменшу відповідну ДФ $g_i(X)$ (1).

Для цього перед кожною транспозицією (10) необхідно виконати перевірку наявності хоча б одного нульового рядка $\overline{A}_i^t$:

$$\exists \overline{A}_i^t = 0, i = \overline{1,m}, \tag{11}$$

а також перевірку умови обнуління всіх рядків матриці $\overline{\boldsymbol{A}}^t$:

$$\forall \overline{A}_i^t = 0, t = \overline{1,N}. \tag{12}$$

Виконання умови (11) є ознакою мінімальної суми елементів i-го рядка A_i^0 початкової матриці $\boldsymbol{A}^0$ (4), який у подальшому із процесу обробки видаляється. При виконанні умови (12) для останнього нульового рядка $\overline{A}_l^N$ матриці $\overline{\boldsymbol{A}}^N$ можна сформувати одиничне значення відповідного елемента p_l вектора топологічних ознак P:

$$p_l = 1, l = \overline{1,m}. \tag{13}$$

Попередньо обнуленим рядкам $\overline{A}_i^t$ у цьому випадку відповідає нульове значення $p_l, i = \overline{1,m}, i \neq l.$

Отже, у запропонованому методі мірою подібності між елементами векторного масиву (РЗ) A_j^{t-1} є поточний поріг обробки q_j^{t-1} (8), а сам РЗ $\overline{A}_j^t$ з елементами вигляду (9) представляє собою вектор різниці між елементами цього РЗ в t-му циклі обробки. А це свідчить про використання векторної (лінійної в межах кожного стовпця $\overline{A}_j^t$ матриці $\boldsymbol{A}^t$) метрики подібності вигляду

$$Q_j = (q_j^0 \ldots q_j^{N-1}), \; j = \overline{1,n} \tag{14}$$

на відміну від скалярної лінійної метрики – Манхеттенської відстані.

Таким чином, запропонованим методом класифікації об'єктів реалізовано такий принцип класифікації за ДФ, що кожному рядку A_i^o початкової матриці $\boldsymbol{A}^0$ (4) ставиться у відповідність елемент p_i вектора топологічних ознак P:

$$A_i^0 \to p_i, i = \overline{1, m}, \tag{15}$$

тобто початковій матриці $\boldsymbol{A}^0$ розмірністю $m \times n$ елементів поставлено у відповідність вектор топологічних ознак P розмірністю m:

$$A^0 \to P. \tag{16}$$

Співвідношення (16) свідчить про стискання вхідного двовимірного простору даних $\boldsymbol{A}^0$ в одновимірний простір ознак P, що є характерною особливістю процесу розпізнавання образів [6, 7].

Отже, перехід від векторного представлення ДФ (3) до матричного представлення елементів ДФ (4) дає можливість організувати двовимірну просторово-розподілену обробку по стовпцях і рядках поточних матриць $\boldsymbol{A}^t$, $t = \overline{1, N}$, що, у свою чергу, дозволяє відмовитись від формування (накопичення у процесі багатооперандного підсумовування) елементів ДФ $g_i(X)$ вигляду (1). Цей процес замінюється зменшенням на величину подібності відповідних однойменних елементів у матриці $\boldsymbol{A}^t$ в кожному t-му циклі обробки. В результаті виключення найменших за величиною функцій $g_i(X)$ залишається остання функція $g_l(X)$, яка є максимальною за своєю величиною і визначає належність об'єкта X до l-го класу C_l за таким вирішальним правилом [4]:

$$(X \mid p_l = 1, l = \overline{1, m}) \Rightarrow C_l, \tag{17}$$

оскільки одиничне значення елемента p_l вектора топологічних ознак P відповідає максимальній ДФ $g_l(X)$ (2).

У роботі [15] наведено результати імітаційного моделювання процесу обробки лінійних ДФ за наведеним методом. Моделювання виконувалось на конкретних прикладах діагностування груп захворювань на апендицит [10]. Як початкові використані попередньо сформовані лінійні ДФ, тому матриця $\boldsymbol{A}^0$ має розмірність 4×7 [10]. У процесі імітаційного моделювання було також сформовано ранги чотирьох ДФ, що при медичному діагностуванні дозволить визначити не тільки найбільш вірогідний діагноз, тобто найбільшу за значенням ДФ (відповідно до найбільшого за рангом), але й найближчі до неї за рангами ДФ. Це, у свою чергу, у подальшому дозволить уточнювати діагноз. Результати імітаційного моделювання довели слушність та достовірність запропонованого підходу до класифікації об'єктів за ДФ.

4. Висновки

При методі обробки за різницевими зрізами у процесі класифікації об'єктів із формуванням дискримінантних функцій використовується векторна метрика подібності, що підкреслює двовимірний просторово-розподілений характер обробки за різницевими зрізами по всій матриці елементів дискримінантних функцій.

При цьому альтернативною процедурою накопиченню (багатооперандному підсумовуванню) елементів дискримінантних функцій є зменшення їх значень одночасно на величину подібності однойменних елементів у стовпцях матриці в кожному циклі обробки з остаточним визначенням максимальної за значенням дискримінантної функції, номер якої визначає клас, до якого належить вхідний об'єкт.

Як приклад області ефективного використання апаратної моделі класифікатора об'єктів із наведеним методом обробки елементів дискримінантних функцій можна визначити підсистему підтримки прийняття рішень при медичному діагностуванні.

СПИСОК ДЖЕРЕЛ

1. Шлезингер М., Главач В. Десять лекций по статистическому и структурному распознаванию. К.: Наукова думка, 2004. 236 с.
2. Гнатієнко Г.М., Снитюк В.Є. Експертні технології прийняття рішень: монографія. К.: ТОВ «Маклаут», 2008. 444 с.
3. Абакумов В.Г., Крылов В.Н., Антощук С.Г. Автоматизированное распознавание при обработке медицинских изображений. *Электроника и связь*. 2002. № 15. С. 124–127.
4. Мартынюк Т.Б., Буда А.Г., Хомюк В.В., Кожемяко А.В., Куперштейн Л.М. Классификатор биомедицинских сигналов. *Искусственный интеллект*. 2010. № 3. С. 88–95.
5. Олійник Г.Т., Степанушко І.В., Трегубенко І.Б. Побудова класифікаторів в задачах біометричної ідентифікації та аутентифікації користувачів. *Вісник Черкаського державного технологічного університету*. 2009. № 1. С. 37–40.
6. Куссуль Н.Н., Куссуль М.Э. Нейросетевая аппроксимация метода динамического программирования на основе классификаторов с преобразованием входного пространства. *Управляющие системы и машины*. 2001. № 1. С. 52–57.
7. Осовский С. Нейронные сети для обработки информации: пер. с польск. М.: Финансы и статистика, 2004. 344 с.
8. Мартинюк Т.Б., Запетрук Я.В. Нейромережевий підхід до медичної експрес-діагностики. *Вісник Вінницького політехнічного інституту*. 2019. № 6. С. 37–44.
9. Рангайян Р. М. Анализ биомедицинских данных. Практический подход / пер. с англ. М.: ФИЗМАТЛИТ, 2007. 440 с.
10. Юнкеров В.И., Григорьев С.Г. Математико-статистическая обработка данных медицинских исследований. СПб.: ВМедА, 2002. 266 с.
11. Новосёлова Н.А., Мастыкин А.С., Том И.Э. Эволюционный подход к выделению информативных признаков в задачах анализа медицинских данных. *Искусственный интеллект*. 2008. № 3. С. 105–112.
12. Бернюков А.К., Сушкова Л.Т. Распознавание биоэлектрических сигналов. *Зарубежная радиоэлектроника*. 1996. № 12. С. 47–51.
13. Мартынюк Т.Б., Хомюк В.В. Особенности математической модели дискретного SM-преобразования. *Математичні машини і системи*. 2010. № 4. С. 145–155.
14. Мартынюк Т.Б., Кожемяко А.В., Куперштейн Л.М. Эффективность посрезовой обработки векторных массивов данных. *Математичні машини і системи*. 2017. № 2. С. 60–67.
15. Мартинюк Т.Б., Медвідь А.В., Гуцол О.М. Моделювання процесу ранжування значень дискримінантних функцій. *Вісник Вінницького політехнічного інституту*. 2013. № 5. С. 74–80.

Стаття надійшла до редакції 02.07.2021