

ОПТИМІЗАЦІЯ ПАРАМЕТРІВ НАВЧАННЯ ЗА ВИБОРКАМИ МАЛОГО ОБСЯГУ

А.С.Краснопоясовський, І.В.Шелехов

Сумський Державний Університет

Римського-Корсакова 2, м Суми Україна 40007

Тел. +38 0542 214084, факс +38 0542 334058

E-mail: kras@kpm.sumdu.edu.ua

Розглядається підхід до оптимізації параметрів навчання в рамках екстремально-інформаційного методу функціонально-статистичних випробувань (МФСВ), який ґрунтується на прямій оцінці інформаційної здатності системи розпізнавання.

ВСТУП

Основним недоліком відомих методів розпізнавання образів є відсутність у процесі навчання корекції вхідного математичного опису з метою побудови високоточного класифікатора [1]. Цей недолік усунено в МФСВ [2,3], в рамках якого розглянемо оптимізацію параметрів навчання за навчальними виборками малого обсягу, що має важливе практичне значення.

1. ПОСТАНОВЛЕННЯ ЗАДАЧІ

Нехай відомо алфавіт класів розпізнавання $\{X_m^o \mid m = \overline{1, M}\}$ і структурований вектор параметрів навчання $g = \langle g_1, \dots, g_\xi, \dots, g_\Xi \rangle$, які впливають на точнісні характеристики (ТХ) класифікатора. Такими параметрами є безпосередньо геометричні параметри контейнерів і параметри нормалізації образів. Під контейнером K_m^o класу X_m^o згідно роботи [4] тут розуміється його замкнена роздільна гіперповерхня, для якої деяким способом визначено геометричний центр. Для нечіткого розбиття $\tilde{\mathfrak{R}}^{|M|}$ бінарного простору ознак Ω_B , яке допускає перетиняємість класів розпізнавання, необхідно виконання таких умов:

- $(\forall X_m^o \in \tilde{\mathfrak{R}}^{|M|}) [X_m^o \neq \emptyset];$
- $(\exists X_k^o \in \tilde{\mathfrak{R}}^{|M|}) (\exists X_l^o \in \tilde{\mathfrak{R}}^{|M|}) [X_k^o \neq X_l^o \rightarrow X_k^o \cap X_l^o \neq \emptyset], k, l = \overline{1, M};$
- $(\forall X_k^o \in \tilde{\mathfrak{R}}^{|M|}) (\forall X_l^o \in \tilde{\mathfrak{R}}^{|M|}) [X_k^o \neq X_l^o \rightarrow$

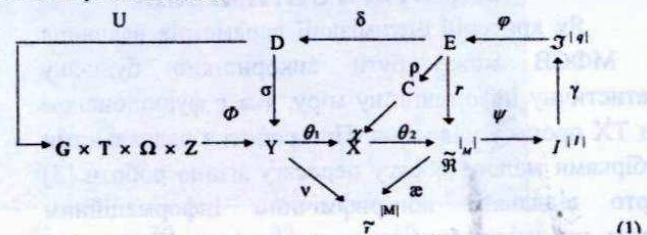
$\rightarrow \text{Ker} X_k^o \cap \text{Ker} X_l^o = \emptyset]$, тобто ядра класів не перетинаються;

- $\bigcup_{X_m^o \in \tilde{\mathfrak{R}}^{|M|}} X_m^o \subseteq \Omega_B \subset X$, де X – простір Хеммінга.

Так само відома дійсна або ціла навчальна матриця $\|y_{m,i}^{(j)}\|$. Треба шляхом нормалізації матриці $\|y_{m,i}^{(j)}\|$ для класу X_m^o знайти максимальне значення критерію функціональної ефективності (КФЕ) E_m^* , яке визначає оптимальні в інформаційному (точнісному) розумінні параметри навчання.

2. МАТЕМАТИЧНА МОДЕЛЬ ОПТИМІЗАЦІЇ ПАРАМЕТРІВ НАВЧАННЯ

Подано універсум випробувань у вигляді декартова добутка $W = G \times T \times \Omega \times Z$, де G – вхідна змінна; T – множина моментів зняття інформації; Ω – простір ознак; Z – множина станів. Введемо оператор $\Phi: W \rightarrow Y$ формування вхідного математичного опису Y класифікатора – навчальної матриці $\|y_{m,i}^{(j)}\|$ і оператор $\theta: Y \rightarrow \tilde{\mathfrak{R}}^{|M|}$ побудови розбиття $\tilde{\mathfrak{R}}^{|M|}$, який є композицією двох операторів: $\theta_1: Y \rightarrow X$ і $\theta_2: X \rightarrow \tilde{\mathfrak{R}}^{|M|}$, де X – бінарна вибіркова множина. Тоді діаграма відображень множин, які застосовуються в процесі навчання за МФСВ, має вигляд



Тут оператор ν апіорної побудови і оператор α апостеріорної побудови нечіткого покриття $\tilde{L}^{|M|}$,

яке задається алфавітом $\{X_m^o\}$, замикають комутативне кільце діаграми. Нехай оператор класифікації $\Psi: \tilde{\mathfrak{R}}^{l^{|M|}} \rightarrow I^{|l|}$ перевіряє основну статистичну гіпотезу про належність реалізацій класу X_m^o , де $I^{|l|}$ – множина виходів (гіпотез); l – кількість гіпотез. Для обчислення КФЕ навчання оператор $\gamma: I^{|l|} \rightarrow \mathfrak{Z}^{|q|}$ шляхом оцінки прийнятих гіпотез формує множину $\mathfrak{Z}^{|q|}$, де $q = l^2$ – кількість ТХ, а оператор $\varphi: \mathfrak{Z}^{|q|} \rightarrow E$ обчислює множину E значень КФЕ, який є функціоналом від ТХ. Ітераційний процес оптимізації параметрів розбиття реалізується оператором $r: E \rightarrow \tilde{\mathfrak{R}}^{l^{|M|}}$ шляхом пошуку максимуму критерію E_m^* . Завдяки принципу композиції кожний параметр навчання має відповідний контур оптимізації в діаграмі (1). При цьому обов'язковим є виконання умови повної композиції: множина E є загальною для всіх контурів оптимізації. Серед інших параметрів навчання, які впливають на достовірність класифікатора, розглянемо систему полів контрольних допусків $\{\delta_{K,i} | i = \overline{1, N}\}$ на значення ознак розпізнавання і рівні селекції $\{\rho_m\}$ координат еталонних векторів (ЕВ), які визначають центри контейнерів. Нехай ітераційний процес оптимізації системи контрольних допусків, яка утворює множину D , реалізується оператором $\delta: E \rightarrow D$ і оператором $\sigma: D \rightarrow Y$, який змінює значення реалізацій образу в процесі навчання. Нехай значення рівнів селекції утворюють множину C , а ітераційний процес їх оптимізації реалізується послідовно оператором $\rho: E \rightarrow C$ і оператором $\chi: C \rightarrow X$, який змінює бінарну навчальну матрицю. Оператор $U: E \rightarrow G \times T \times \Omega \times Z$ регламентує процес навчання.

3. КРИТЕРІЙ ОПТИМІЗАЦІЇ

Як критерій оптимізації параметрів навчання за МФСВ може бути використано будь-яку статистичну інформаційну міру, яка є функціоналом від ТХ процесу навчання. При роботі з навчальними вибірками малого обсягу перевагу згідно роботи [5] варто віддавати логарифмічним інформаційним мірам, що мають згладжувальний ефект. Наприклад, для двоальтернативної системи оцінок і рівномірних гіпотез згідно принципу неповноти

апріорної інформації Бернуллі-Лапласа можна застосовувати, наприклад, модифікацію нормованого критерію за Шенноном у вигляді [6]:

$$E = 1 + \frac{1}{2} \left(\frac{\alpha}{\alpha + D_2} \log_2 \frac{\alpha}{\alpha + D_2} + \frac{\beta}{D_1 + \beta} \log_2 \frac{\beta}{D_1 + \beta} + \frac{D_1}{D_1 + \beta} \log_2 \frac{D_1}{D_1 + \beta} + \frac{D_2}{\alpha + D_2} \log_2 \frac{D_2}{\alpha + D_2} \right), \quad (2)$$

де α , β , D_1 , D_2 – ТХ: помилки першого і другого роду, перша і друга достовірності відповідно. Критерій (2) є нелінійним і неоднозначним функціоналом ТХ, що вимагає оцінки в процесі навчання робочої області визначення його значень.

4. ОСОБЛИВОСТІ РЕАЛІЗАЦІЇ АЛГОРИТМУ ОПТИМІЗАЦІЇ ПАРАМЕТРІВ НАВЧАННЯ

Алгоритм навчання у рамках МФСВ полягає в послідовній реалізації ітераційних процедур оптимізації параметрів навчання шляхом пошуку максимуму КФЕ навчання для кожного контуру оптимізації в діаграмі (1). З метою спрощення класифікатора як наближення “точного” контейнера, наприклад, класу X_m^o може розглядатися гіперсфера, центром якої є вершина ЕВ $x_m \in X_m^o$, що відповідає прийнятній для задач контролю і управління гіпотезі компактності (як чіткої, так і нечікої) реалізацій образу. Тоді як параметри оптимізації за діаграмою (1) розглядаються радіуси $\{d_m\}$ контейнерів, контрольні допуски на ознаки розпізнавання і рівні селекції $\{\rho_m\}$ координат ЕВ.

5. ПРИКЛАД РЕАЛІЗАЦІЇ АЛГОРИТМУ НАВЧАННЯ ЗА МФСВ

Розглянемо задачу оптимізації геометричних параметрів контейнера у вигляді гіперсфери на прикладі автофокусування растрового електронного мікроскопа РЕМ-103 виробництва ВАТ SELMI (м.Суми, Україна) за зображенням зразка, що досліджується. Ідея автофокусування полягає в побудові на кожному кроці настроювання мікроскопа оптимального контейнера для початкового розфокусованого зображення (клас X_1^o) як за своїми реалізаціями, так і реалізаціями поточного зображення (клас X_S^o), отриманими на S -му кроці настроювання, шляхом пошуку максимуму інформаційної міри різноманітності цих класів. Як параметр настроювання розглядався струм управляючої обмотки об'єктивної лінзи. На рис.1а

наведено початкове розфокусоване зображення об'єкту «Гратка», а на рис.1б – зображення цього об'єкту, отриманого після першого кроку настроювання..

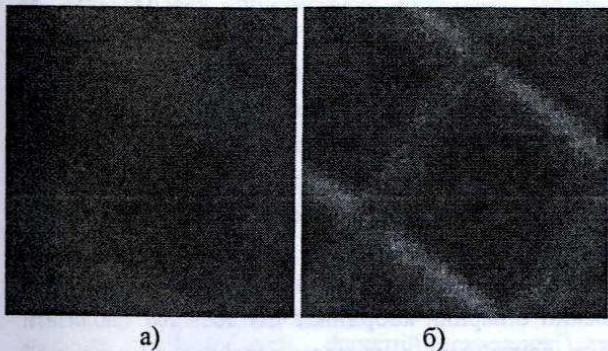


Рис.1. Зображення об'єкту «Гратка»: а)початкове розфокусоване зображення; б) поточне зображення

У процесі реалізації алгоритму навчання за діаграмою (1) було побудовано оптимальний контейнер $K_1^o \subset X_1^o$, який забезпечує максимальну асимптотичну достовірність розпізнавання в режимі екзамену реалізацій класу X_1^o . Оптимальний радіус d_1^* контейнера K_1^o , як видно з рис.2, дорівнює 35 кодовим одиницям відстані при значенні КФЕ $E_1^* = 0,809$. На рис.2 темна область графіка позначає робочу область його визначення, в якій $D_1, D_2 > 0,5$. В області I присутні тільки реалізації класу X_1^o , в області II присутні реалізації класів X_1^o і X_S^o , а в

області III – тільки реалізації класу X_S^o . В табл.1 наведено значення критерію E_1 , першої достовірності D_1 і помилки другого роду β , параметрів навчання ρ_1 , d_1 , δ і міжцентрова кодова відстань $d_c = d(x_1 \oplus x_S)$. Як видно з табл.1, оптимальне значення рівня селекції дорівнює $\rho_1^* = 0,54$, а оптимальне поле контрольних допусків для i -ої ознаки дорівнює $x_{1,i} \pm \delta^*$, де $\delta^* = 12$ градаціям яскравості.

Таблиця 1. Результати оптимізації контейнера K_1^o

ρ_1	E_1	D_1	β	d_1	δ	d_c
0,40	0,674989	0,80	0,00	39	12	55
0,42	0,710603	0,84	0,00	33	14	4
0,44	0,710603	0,84	0,00	33	14	38
0,46	0,710603	0,84	0,00	31	14	38
0,48	0,701025	0,83	0,00	30	14	43
0,50	0,754381	1,00	0,12	38	12	45
0,52	0,794417	1,00	0,09	37	12	48
0,54	0,809527	1,00	0,08	35	12	48
0,56	0,809527	1,00	0,08	35	12	51
0,58	0,765397	0,99	0,08	33	12	53
0,60	0,794417	1,00	0,09	33	12	51

Початкові координати ЕВ дорівнювали вибірковим середнім відповідних навчальних виборок. Оскільки навчальна матриця для кожного класу складалася тільки із 100 реалізацій, то центри розсіювання реалізацій не збігалися з вершинами початкових ЕВ.

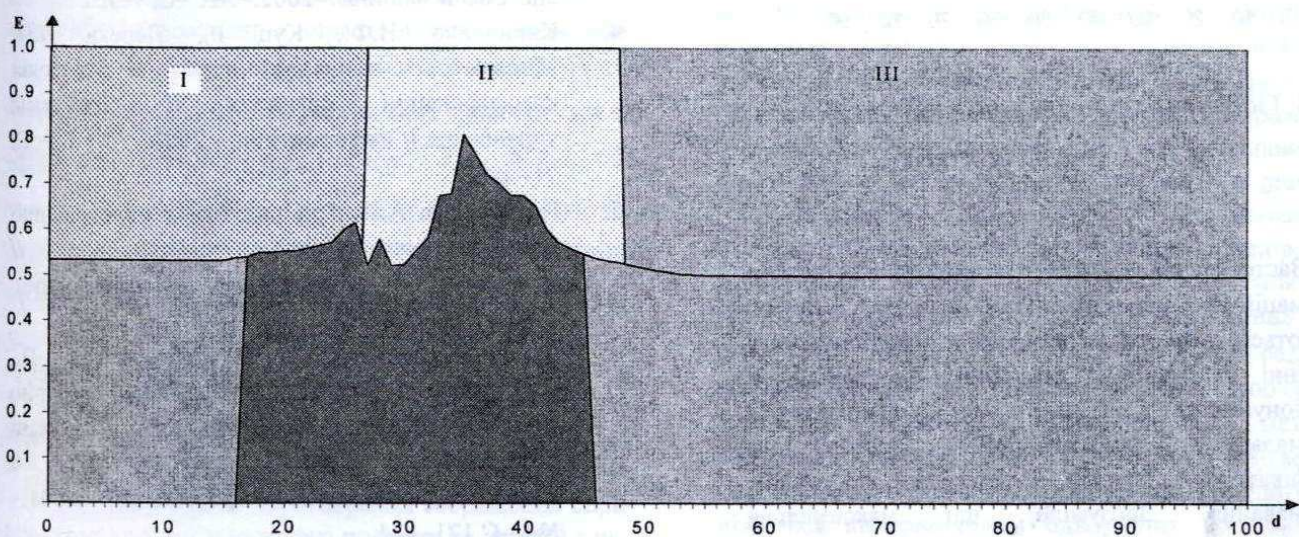
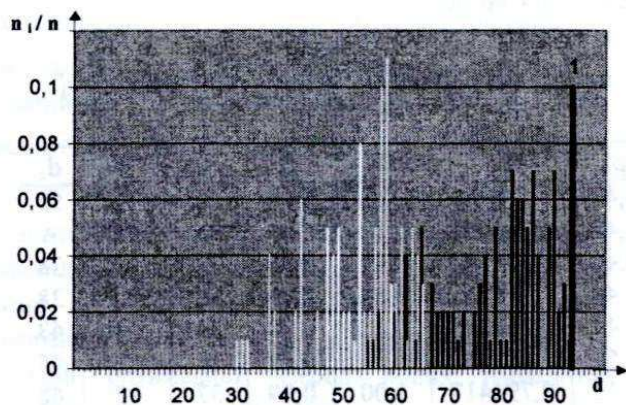
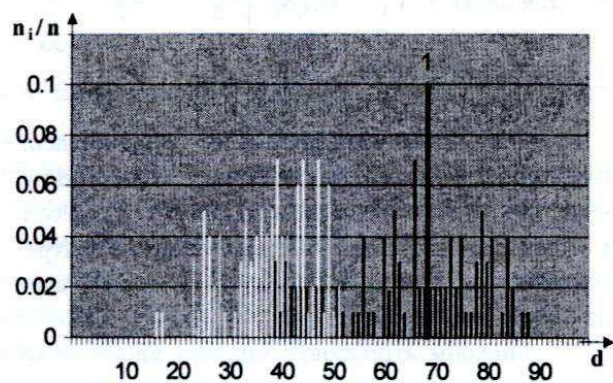


Рис. 2. Залежність КФЕ від радіуса контейнера K_1^o

На рис.3 показано гістограми розподілів реалізацій класів X_1^o (темні дискрети) і X_S^o (світлі дискрети). Тут мітка 1 позначає положення вершини початкового вектора x_1 при $\rho_1=0,50$ (рис.3а) і положення вершини цього вектора при оптимальному рівні селекції $\rho_1^* = 0,54$, яке, як видно з рис.3б, є більш наближеним до центру розсіювання реалізацій класу X_1^o .



а)



б)

Рис. 3, Гістограма розподілу реалізацій класів X_1^o і X_S^o : а) $\rho=0,50$; б) $\rho=0,54$

ПІДСУМОК

Застосування МФСВ для розв'язання задачі інформаційного синтезу систем розпізнавання, що навчаються, дозволяє оптимізувати в інформаційному розумінні просторово-часові параметри функціонування системи, які забезпечують максимальну асимптотичну достовірність класифікатора за умови перетиняємості класів розпізнавання. Зрозуміло, що максимальна асимптотична функціональна ефективність класифікатора збігається з фактичною тільки за умови однакової структурованості навчальної та

екзаменаційної матриць і однакових показників їх статистичної сталості та статистичної однородності. На відміну від відомих параметричних і непараметричних методів розпізнавання МФСВ дозволяє будувати ефективний класифікатор за репрезентативними в розумінні Паретто навчальними виборками малого обсягу, що важливо для розв'язання практичних задач контролю та управління. Основне обмеження на використання МФСВ полягає в наявності у контейнерів класів розпізнавання деяким способом визначеного геометричного центра. Показано, що вибір оптимального в інформаційному розумінні рівня селекції бінарних координат ЕВ дозволяє наблизити центр контейнера з центром розсіювання реалізацій відповідного класу, що призводить до підвищення достовірності класифікатора..

ЛІТЕРАТУРА

1. Васильев В. И. Распознающие системы. Справочник. 2е изд., перераб. и доп. – Киев: Наукова думка, 1983.– 422 с.
2. Красноясовський А. С., Заговора С. О. Розпізнавання електронограм за методом функціонально – статистичних випробувань // Оброблення сигналів і зображень, та розпізнавання образів: Праці п'ятої Всеукраїнської міжнародної конференції. – Київ, 2000. – С. 37 – 41.
3. Красноясовский А.С. Классификационная настройка сложной системы // Управляющие системы и машины.–2002.–№1.–С. 9–15.
4. Кириченко Н.Ф., Куц Р., Лепеха Н.П. Множества принадлежности в задачах классификации сигналов // Проблемы управления и информатики.– 2001. – № 5. – С. 71– 84.
5. Колмогоров А.Н. Три подхода к определению понятия «количество информации» // Проблемы передачи информации, 1961.–№3.–С.27–35.
6. Красноясовський А. С., Черниш А. В., Сластухевський О.Ю. Про вибір критерію функціональної ефективності системи розпізнавання, що навчається // Радиоэлектроника и информатика, 2001.– №4.–С.121–124.