

# НЕЙРОІНФОРМАТИКА ТА ІНТЕЛЕКТУАЛЬНІ СИСТЕМИ

## НЕЙРОІНФОРМАТИКА И ИНТЕЛЛЕКТУАЛЬНЫЕ СИСТЕМЫ

### NEUROINFORMATICS AND INTELLIGENT SYSTEMS

УДК 681.518:004.93.1'

Москаленко В. В.<sup>1</sup>, Коробов А. Г.<sup>2</sup>

<sup>1</sup>Канд. техн. наук, старший викладач кафедри комп'ютерних наук Сумського державного університету, Суми, Україна  
<sup>2</sup>Аспірант кафедри комп'ютерних наук Сумського державного університету, Суми, Україна

#### ІНФОРМАЦІЙНО-ЕКСТРЕМАЛЬНИЙ АЛГОРИТМ ФУНКЦІОНУВАННЯ СИСТЕМИ РОЗПІЗНАВАННЯ ОБ'ЄКТІВ НА МІСЦЕВОСТІ З ОПТИМІЗАЦІЄЮ ПАРАМЕТРІВ ЕКСТРАКТОРА ОЗНАК

**Актуальність** теми статті полягає в тому, що питання вибору параметрів екстрактора ознак та алгоритму класифікаційного аналізу за умов апріорної невизначеності, ресурсних та інформаційних обмежень є недостатньо дослідженим і в повній мірі досі не вирішене.

**Мета статті** – підвищити в інформаційному та вартісному сенсі ефективність функціонування автономної системи розпізнавання об'єктів на місцевості, що функціонує в режимах навчання та екзамєну за умов обмежених обсягів навчальної вибірки шляхом оптимізації параметрів екстрактора ознак та класифікатора спостережень.

**Методи дослідження** базуються в алгоритмах цифрової обробки та аналізу зображень для формування дескрипторів об'єктів інтересу, на принципах математичної статистики та теорії інформації для оцінки функціональної ефективності вирішальних правил, положеннях теорії популяційних алгоритмів для оптимізації параметрів системи та сканування зображень.

**Результати** – розроблений алгоритм машинного навчання з грубим двійковим кодуванням спостережень та модифікація ройового алгоритму оптимізації параметрів функціонування системи розпізнавання дозволяє отримати за вибірками малого обсягу вирішальні правила з достовірністю, що наближається до граничного максимального значення. При цьому експериментально показано перевагу використання ройового алгоритму сканування зображень, яка полягає у трикратному підвищенні швидкодії порівняно з відомими алгоритмами RASW та ESS.

**Висновки.** Запропоновано метод синтезу інформаційно-екстремального класифікатора зображень з грубим двійковим кодуванням розріджених гістограм частоти появи візуальних слів, що дозволяє одержати обчислювально ефективні вирішальні правила. Удосконалено метод популяційного пошуку для настроювання параметрів екстрактора ознак, що дозволяє отримати оптимальні в інформаційному та вартісному сенсі параметри функціонування системи розпізнавання за декілька ітерацій роботи алгоритму. Практична цінність результатів полягає в отриманні робочих алгоритмів проектування здатної навчатися системи розпізнавання зображень, що функціонує за умов ресурсних та інформаційних обмежень.

**Ключові слова:** словник ознак розпізнавання, алфавіт класів розпізнавання, інформаційний критерій, оптимізація, машинне навчання, рій частинок.

#### НОМЕНКЛАТУРА

$C$  – трудомісткість вирішальних правил;

$C_{\min}$  – стартова трудомісткість вирішальних правил;

$c_1$  – перша константа прискорення частинки ройового пошуку оптимуму;

$c_2$  – друга константа прискорення частинки ройового пошуку оптимуму;

$d_m$  – кодовий радіус гіперсферичного контейнера класу  $X_m^o$ ;

$d(x_m^* \oplus x)$  – кодова відстань від центру контейнера класу  $X_m^o$  до вектора  $x$ ;

$E_{\max}$  – граничне значення інформаційного машинного навчання;

$\bar{E}$  – усереднений за алфавітом класів КФЕ машинного навчання;

$\langle e_1, \dots, e_{\xi_1}, \dots, e_{\Xi_1} \rangle$  – генотипні параметри функціонування, які впливають на параметри алгоритмів екстракції ознак розпізнавання;

$\langle f_1, \dots, f_{\xi_3}, \dots, f_{\Xi_3} \rangle$  – фенотипні параметри функціонування, які впливають на геометрію контейнерів класів розпізнавання;

$\langle g_1, \dots, g_{\xi_2}, \dots, g_{\Xi_2} \rangle$  – генотипні параметри функціонування, які впливають на параметри розподілу векторів-реалізацій образу в бінарному просторі ознак;

$G$  – область допустимих значень параметрів функціонування системи;

$G_{x,y}$  – область допустимих значень координат скануючого вікна;

$G_{\mu}$  – область допустимих значень функції належності об'єкта інтересу;

$J_m$  – критерій ефективності машинного навчання;

$K$  – коефіцієнт крутизни (за замовчуванням  $K=4$ );

$m$  – номер поточного класу розпізнавання;

$M$  – потужність алфавіту (кількість класів) системи розпізнавання об'єктів;

$N$  – кількість ознак у векторах-реалізаціях навчальної вибірки;

$n$  – кількість векторів – реалізацій у навчальній вибірці;

$N_W^{\min}, N_W^{\max}$  – мінімальна та максимальна потужність алфавіту візуальних слів відповідно;

$P_j, Pbest_j$  – поточна та найкраща персональна позиція  $j$ -го агента відповідно;

$\Delta_H(P_j), \Delta_H(Pbest_j)$  – параметр порогу відкидання локальних ознак поточної та найкращої персональної позицій  $j$ -го агента відповідно;

$\Delta_W(P_j), \Delta_W(Pbest_j)$  – параметр потужності словника візуальних слів поточної та найкращої персональної позицій  $j$ -го агента відповідно;

$T_H^{\min}, T_H^{\max}$  – мінімальний та максимальний поріг відкидання локальних ознак відповідно.

$w$  – константа інерції частинки ройового алгоритму оптимізації;

$x_m$  – геометричний центр його гіперсферичного контейнера класу  $X_m^o$ ;

$U_{i,\max}$  – максимальне значення  $i$ -ї ознаки в навчальній вибірці;

$\alpha_m$  – помилка першого роду;

$\beta_m$  – помилка другого роду;

$\delta_{\max}$  – граничне значення ширини поля контрольних допусків;

$\delta_i$  – ширина поля контрольних допусків для  $i$ -ї ознаки розпізнавання;

$\varepsilon$  – будь-яке мале, наближене до нуля, додатне число.

$\mu_m(x,y)$  – значення функції належності вектора-реалізації, обчислена у вікні з координатами  $(x,y)$  до  $m$ -го класу із алфавіту  $\{X_m^o\}$ ;

ІЕІ-технологія – інформаційно-екстремальна інтелектуальна технологія;

КФЕ – критерій функціональної ефективності;

СКД – система контрольних допусків.

### ВСТУП

Відомі алгоритми глибинного машинного навчання для ідентифікації об'єктів інтересу у відеопотоці даних є мало придатними для використання в автономних системах аеро моніторингу території, охоронних системах відеоспостереження та роботехнічних системах різного при-

значення, оскільки потребують використання значного обсягу обчислювальних ресурсів та навчальних даних. У задачах співставлення орієнтирів, прив'язки зображень до ортофотопланів, детектування та супроводження об'єктів набуло поширення використання локальних ознак і дійсночислових (з комою) або бінарних дескрипторів для їх опису, які характеризуються стійкістю до геометричних та фотометричних перетворень [1, 2]. При цьому сформовані на їх основі глобальні дескриптори характеризуються відносно невисокою обчислювальною складністю та достатньою узагальнюючою здатністю для забезпечення репрезентативності навчальних вибірок малого обсягу [2, 3]. Проте питання визначення оптимальних параметрів локальних та глобальних дескрипторів досі є малодослідженим і на практиці його вирішення зводиться до ручного налаштування.

Одним із шляхів одночасного забезпечення максимальних достовірності розпізнавання та оперативності функціонування системи в режимах навчання та пошуку об'єктів інтересу у відеопотоці даних є розробка інтегральних критеріїв ефективності системи та застосування алгоритмів пошукової оптимізації. При цьому найбільший інтерес представляють ройові алгоритми, які характеризуються високою оперативністю і низькою ймовірністю застрягання в локальних екстремумах фітнес-функції. Проте загальна оперативність функціонування залежить як від структури та параметрів екстрактора ознак так і від обраного алгоритму класифікаційного аналізу.

У статті розглядається задача інформаційного синтезу системи розпізнавання об'єктів на місцевості з оптимізацією параметрів екстрактора ознак, оснований на локальних та глобальних дескрипторах зображень. При цьому для оптимізації в інформаційному та вартісному розумінні параметрів вирішальних правил системи розпізнавання та для знаходження координат об'єкта інтересу на кадрі зображення запропоновано використовувати алгоритм рою частинок.

### 1 ПОСТАНОВКА ЗАДАЧІ

Нехай сформовано алфавіт класів розпізнавання  $\{X_m^o \mid m = \overline{1, M}\}$ , що характеризують об'єкти інтересу та фонову місцевість, і навчальна матриця у вигляді двовимірного масиву кольорових зображень  $\{y_m^j \mid j = \overline{1, n_m}, m = \overline{1, M}\}$ , де  $n_m$  – кількість зображень класу  $X_m^o$ .

Дано структурований вектор просторово-часових параметрів функціонування системи ідентифікації об'єктів інтересу на місцевості, який у загальному випадку має структуру

$$g = \langle e_1, \dots, e_{\xi_1}, \dots, e_{\Xi_1}, g_1, \dots, g_{\xi_2}, \dots, g_{\Xi_2}, f_1, \dots, f_{\xi_3}, \dots, f_{\Xi_3} \rangle,$$

$$\Xi_1 + \Xi_2 + \Xi_3 = \Xi. \tag{1}$$

При цьому відомі обмеження на відповідні параметри функціонування:

$$R_{\xi_1}(e_1, \dots, e_{\xi_1}, \dots, e_{\Xi_1}) \leq 0; \quad R_{\xi_2}(g_1, \dots, g_{\xi_2}, \dots, g_{\Xi_2}) \leq 0;$$

$$R_{\xi_3}(f_1, \dots, f_{\xi_3}, \dots, f_{\Xi_3}) \leq 0.$$

Необхідно в процесі машинного навчання визначити оптимальні значення координат вектора (1), які забезпечують максимум усередненого за алфавітом класів розпізнавання інформаційного критерію функціональної ефективності (КФЕ) навчання системи та мінімальну обчислювальну трудомісткість вирішальних правил

$$g^* = \arg \max_G \left\{ \frac{\bar{E}}{E_{\max}} \cdot \frac{C_{\min}}{C} \right\}.$$

При функціонуванні системи розпізнавання безпосередньо в робочому режимі необхідно за максимумом функції належності вектора-реалізації, обчисленим у вікні пошуку, до одного з класів із заданого алфавіту  $\{X_m^o\}$  визначити координати об'єкта інтересу в кадрі відеопотоку

$$\langle x^*, y^*, m^* \rangle = \arg \max_{G_{xy} \cap G_\mu} \{ \mu_m(x, y) \}, m = \overline{1, M}.$$

## 2 ОГЛЯД ЛІТЕРАТУРИ

При розробці здатних до навчання автономних мобільних систем розпізнавання візуальних образів високої розмірності набуває важливого значення розробка невибагливих до обчислювальних ресурсів алгоритмів аналізу даних з адаптацією до умов спостереження. У працях [1–3] було запропоновано підходи, основані на застосуванні принципів лінгвістики до аналізу зображення за його найбільш помітними фрагментами, які асоціюють з візуальними словами. При цьому визначення координат таких фрагментів відбувається за результатами роботи детекторів ключових точок, таких як Harris-Laplace, Hessian-Laplace, Fast Hessian, SUSAN та DoG, а їх векторний опис здійснюється за допомогою дійсничислових чи бінарних дескрипторів, найбільш популярними серед яких є SIFT, SURF, MSER, BRIEF та FREAK [4, 5]. Проте досі не існує формалізованих правил щодо вибору значень порогових параметрів, що відповідають за врахування чи відкидання локальних ознак, та параметрів, що впливають на розмірність кожного рівня подання ознак в глобальних дескрипторах.

Формування глобальних дескрипторів зображення у вікні пошуку згідно принципів лінгвістики полягає у використанні мішка візуальних слів (Bag of Visual Words, BoW) та фраз (Bag of Visual Phrase, BoP), де словник візуальних слів формується за результатами кластер-аналізу дескрипторів знайдених ключових точок [6, 7]. При цьому багаторівневе представлення ознакового опису в екстраторі, основаному на BoP, може бути побудоване за принципами  $n$ -грамм, де фрази формуються шляхом об'єднання  $k$ -найближчих ключових точок, або на основі локальних графів, які формуються за допомогою триангуляції Delaunay [7–9].

Ефективність застосування екстратора ознак визначається обраним алгоритмом класифікаційного аналізу. У працях [8, 10] досліджується використання методу класифікації результатів екстракції ознак на основі опорних векторів, однак даний метод потребує ручного налаштування параметрів алгоритму, є чутливим до шуму і йому характерне уповільнення навчання та зниження точності вирішальних правил у випадку перетину класів

в просторі ознак. У працях [11, 12] показано, що перспективним напрямком синтезу здатних навчатися класифікаторів є застосування грубого двійкового кодування ознак на основі системи контрольних допусків (СКД), та логарифмічних інформаційних мір ефективності вирішальних правил. Це пов'язано з можливістю побудови в рамках геометричного підходу простих вирішальних правил в бінарному просторі ознак, перехід до якого відбувається за допомогою простих операцій порівняння. При цьому згладжуючий ефект логарифмічної інформаційної міри підвищує ймовірність знаходження глобально-го максимуму її функції.

Для визначення координат об'єкта інтересу на кадрі відеопотоку можуть бути використані різноманітні алгоритми сканування, серед яких найбільш простим у реалізації є алгоритм ковзного вікна, що полягає у послідовному скануванні кадру з заданим кроком по координатах  $x$  та  $y$  [13]. При цьому за умов апріорної невизначеності щодо масштабу зображення високої розмірності алгоритм ковзного вікна є трудомістким і не придатним для практичного використання у системах реального часу. У праці [14] було запропоновано алгоритм адаптивного ковзного вікна (RASW), де крок переміщення скануючого вікна пошуку змінюється залежно від значення функції належності класифікатора, що дозволяє дещо підвищити оперативність алгоритму за рахунок збільшення кроку сканування на фонових ділянках зображення. У праці [15] для локалізації об'єкта інтересу на кадрі було запропоновано алгоритм ефективного пошуку підвікна (ESS), оснований на алгоритмі гілок та границь, що є придатним для використання у багатьох практичних застосуваннях реального часу. Проте ускладнення екстрактора ознак обумовлює актуальність дослідження нових підходів до підвищення оперативності сканування зображення. Одним з перспективних підходів до сканування зображень є популяційні алгоритми пошуку, однак їх реалізація потребує розробки фітнес-функції на основі функції належності використаного класифікатора та визначення налаштувань пошукового алгоритму [16].

Таким чином, для підвищення оперативності та точності алгоритмів розпізнавання об'єктів на місцевості актуальними є розробка і дослідження популяційних алгоритмів оптимізації параметрів функціонування екстрактора ознак, класифікатора та сканера зображень як в інформаційному так і вартісному сенсах. При цьому перспективним напрямком синтезу системи класифікаційного аналізу зображень є застосування двійкового кодування ознак на основі контрольних допусків в процесі максимізації інформаційної міри ефективності вирішальних правил.

## 3 МАТЕРІАЛИ ТА МЕТОДИ

Для формування вхідного математичного опису системи розпізнавання, основаного на пошуці візуальних слів та фраз, пропонується використовувати SURF алгоритм екстракції локальних ознак [4, 5]. В цьому методі ключові точки шукаються за допомогою матриці частинних похідних – матриці Гессе. Детермінант матриці Гессе досягає екстремуму в точках максимальної зміни градієнту яскравості, що дозволяє успішно детектувати

плями, кути і краї ліній. При цьому інваріантність до масштабу досягається шляхом використання різномасштабних фільтрів. Після кластер-аналізу дескрипторів ключових точок, що описують поведінку градієнтів навколо точок, для кожного зображення може бути сформована гістограма частот як візуальних слів так і фраз. Візуальні фрази довжиною  $k+1$  пропонується формувати шляхом пошуку для кожної точки найближчих в системі координат  $(x, y)$   $k$ -візуальних слів. Для скорочення розмірності алфавіту візуальних фраз у роботі пропонується виконати видалення фраз, що з'являються більше ніж в одному класі розпізнавання.

Алгоритм грубого двійкового кодування векторів спостережень полягає у порівнянні значення  $i$ -ї ознаки з відповідним нижнім  $A_{L,i}$  та верхнім  $A_{U,i}$  межами несиметричного поля контрольних допусків класу розпізнавання, які розраховуються за формулами

$$A_{L,i} = y_{i,\max} \left[ 1 - \frac{\delta_i}{\delta_{\max}} \right],$$

$$A_{U,i} = y_{i,\max}.$$

Формування бінарної навчальної матриці  $\{x_{m,i}^{(j)} \mid i = \overline{1, N}; j = \overline{1, n_m}; m = \overline{1, M}\}$  здійснюється за правилом

$$x_{m,i}^{(j)} = \begin{cases} 1, & \text{if } A_{L,i} \leq y_{m,i}^{(j)} \leq A_{U,i}; \\ 0, & \text{else.} \end{cases}$$

Обчислення значень координат двійкового еталонного вектора  $x_m$ , відносно якого відбувається побудова в радіальному базисі контейнерів класів, здійснюється за правилом

$$x_{m,i} = \begin{cases} 1, & \text{if } \frac{1}{n_m} \sum_{j=1}^{n_m} x_{m,i}^{(j)} > \frac{1}{n} \sum_{m=1}^M \sum_{j=1}^{n_m} x_{m,i}^{(j)}; \\ 0, & \text{if else.} \end{cases}, \quad i = \overline{1, N \cdot M}$$

Як критерій ефективності машинного навчання класифікатора розглядається модифікація інформаційної міри Кульбака [11, 12]:

$$J_m = \frac{1 - (\alpha_m^{(k)} + \beta_m^{(k)})}{\log_2(2 + \varepsilon) + r \log_2 10} * \log_2 \left[ \frac{2 - (\alpha_m^{(k)} + \beta_m^{(k)}) + \varepsilon}{(\alpha_m^{(k)} + \beta_m^{(k)}) + \varepsilon} \right]. \quad (2)$$

Оптимізація параметрів поля контрольних допусків  $\{\delta_{m,i}\}$  та інших генотипних параметрів полягає в пошуку екстремуму функції КФЕ (2) в гіперпросторі рішень. При цьому як пошуковий алгоритм в даній роботі пропонується використати рій частинок (Particle Swarm Optimization, PSO), який характеризується простотою реалізації та інтерпретабельністю [16]. Оптимізація словника ключових фрагментів та фенотипних параметрів вирішальних правил (радіуси контейнерів класів) може здійснюватися методом прямого перебору з заданим кроком, оскільки кількість кроків такого пошуку є відносно малою.

У режимі екзамену рішення про належність вектора реалізації  $x$  одному з класів алфавіту  $\{X_m^o\}$  приймається шляхом обчислення геометричної функції належності:

$$\mu_m^*(x) = \max_{\{m\}} \{\mu_m(x)\},$$

в якій  $\mu_m(x)$  являє собою функцію належності вектора  $x$  до контейнера класу  $X_m^o$ , яка обчислюється за правилом:

$$\mu_m(x) = 1 - \frac{d(x_m^* \oplus x)}{d_m^*}.$$

Трудомісткість інформаційно-екстремального машинного навчання збільшується швидше квадрату кількості навчальних векторів, тому для підвищення оперативності навчання використовується зведення багатокласової класифікації до серії двохкласових за принципом «кожен проти кожного», де будується  $M^*(M-1)/2$  двохкласових класифікаторів [10]. Результат класифікації класифікатором, що розрізняє  $m$ -й та  $k$ -й класи подається у вигляді значення дискримінантної функції  $\mu_{mk}$ , використовуючи рівність  $\mu_{mk} = -\mu_{km}$  для скорочення обчислень. При цьому дискримінантна функція двохкласового інформаційно-екстремального класифікатора має вигляд:

$$\mu_{mk}(x) := \begin{cases} \mu'_m(x) & \text{якщо } \mu'_m(x) > \mu'_k(x), \\ -\mu'_k(x) & \text{якщо інакше.} \end{cases}$$

в якій  $\mu'_m(x)$  є нормалізованим значення функції належності та обчислене за формулою:

$$\mu'_m(x) = \frac{\exp(\mu_m(x) - \mu_{\max}(x))}{\sum_{k=1}^M \exp(\mu_k(x) - \mu_{\max}(x))},$$

при цьому  $\mu_{\max}$  визначає максимальне значення функції належності та розраховується за правилом

$$\mu_{\max}(x) = \max_m \{\mu_m(x)\}.$$

Віднесення екзаменаційного вектора ознак  $x$  до одного з класів алфавіту  $\{X_m^o\}$  відбувається шляхом головування відповідних двохкласових класифікаторів. Результуючу функцію належності до  $m$ -го класу можна подати у вигляді

$$\mu_m^n(x) = \sum_{k \neq m} \frac{1}{1 + \exp(-K * \mu_{mk}(x))}.$$

Таким чином, алгоритм навчання системи ідентифікації об'єктів на місцевості полягає в оптимізації параметрів формування вхідного математичного опису та параметрів вирішальних правил шляхом пошуку глобального максимуму інформаційного КФЕ (2). При цьому для підвищення оперативності навчання використовується зведення багатокласової класифікації до серії двохкласових.

#### 4 ЕКСПЕРИМЕНТИ

Розглянемо процедуру настроювання потужності алфавіту візуальних слів  $N_W$  та порогу відкидання ло-

кальних ознак  $T_H$  зображення як процедуру популяційного пошуку максимуму фітнес-функції в двовимірному просторі рішень. При цьому агент популяції містить параметри  $\Delta_W$  та  $\Delta_H$  для розрахунку параметрів алгоритму екстракції ознак за правилами

$$N_W = N_W^{\min} + \frac{\Delta_W}{\Delta_{MAX}} (N_W^{\max} - N_W^{\min});$$

$$T_H = T_H^{\max} - \frac{\Delta_H}{\Delta_{MAX}} (T_H^{\max} - T_H^{\min}).$$

Ефективність кожної частинки популяційного алгоритму, тобто її близькість до глобального оптимуму, вимірюється за допомогою наперед визначеної фітнес-функції, роль якої в даному випадку виконує функція КФЕ. Кожна  $j$ -та частинка крім її позиції  $P_j$  зберігає наступну інформацію:  $V_j$  – поточна швидкість частинки,  $Pbest_j$  – краща персональна позиція частинки. Краща персональна позиція  $j$ -ї частинки – це позиція  $j$ -ї частинки, в якій значення фітнес функції для частинки було максимальним на поточний момент часу. Крім цього, з метою пошуку глобального екстремуму фітнес-функції найкраща частинка шукається в усьому рої, а її позиція позначається як  $Gbest$ .

Проте розглянутий вище ройовий алгоритм пошуку спрямований переважно на видалення заважаючих ознак розпізнавання і відповідно на підвищення усередненого за алфавітом класів КФЕ навчання. З метою додаткового скорочення потужності словника ознак шляхом видалення решти неінформативних ознак потрібно виконати деяку модифікацію ройового алгоритму пошуку. Для цього слід модифікувати процедуру оновлення значень найкращої персональної  $Pbest_j$  позиції агентів пошуку за правилом (3), в якому цільова функція  $E(\dots)$  є усередненим значенням функції КФЕ:

$$if \left| E(P_j) - E(Pbest_j) \right| < \varepsilon,$$

$$if \sqrt{\Delta_H^2(P_j) + \Delta_W^2(P_j)} < \sqrt{\Delta_H^2(Pbest_j) + \Delta_W^2(Pbest_j)},$$

$$Pbest_j := P_j. \quad (3)$$

Аналогічно потрібно модифікувати процедуру оновлення значень найкращої глобальної  $Gbest_j$  позиції

агентів пошуку

$$if \left| E(Pbest_j) - E(Gbest) \right| < \varepsilon,$$

$$if \sqrt{\Delta_H^2(Pbest_j) + \Delta_W^2(Pbest_j)} < \sqrt{\Delta_H^2(Gbest) + \Delta_W^2(Gbest)},$$

$$Gbest := Pbest_j. \quad (4)$$

Для оптимізації параметрів екстрактора ознак, СКД та пошукового сканування зображення необхідно задати параметри настройки популяційного алгоритму. Запропоновані настройки алгоритму рою частинок приведені у табл. 1.

Пошук максимуму інформаційного критерію функціональної ефективності машинного навчання повинен здійснюватися в робочій області визначення його функції, обмежена нерівностями  $\alpha_m < 0,5$ ,  $\beta_m < 0,5$  та  $d_m < d(x_m \oplus x_c)$ , де  $d_m$  – радіус гіперсферичного контейнера класу  $X_m^o$ ,  $d(x_m \oplus x_c)$  – кодова відстань між усередненим вектором класу  $X_m^o$  та найближчого до нього усередненим вектором класу  $X_c^o$ . На графіках оптимізації радіусу контейнерів класів розпізнавання робоча область заштриховується.

Оцінку ефективності алгоритму сканування зображення пропонується здійснювати шляхом вимірювання середнього часу пошуку об'єкта інтересу та підрахунку відсотку правильних спрацювань. Порівняння розробленого алгоритму сканування з відомими алгоритмами RAWs та ESS буде здійснюватися за цими оцінками.

Як вхідні дані для алгоритму навчання розглядаються фотографії розміром 100x100 пікселів для 3-х об'єктів інтересу та фоновій місцевості без об'єктів інтересу. Для алгоритму сканування використовуються фотографії місцевості формату Full HD (1920 x 1080 пікселів) із зображенням об'єктів інтересу, отримані на тій же висоті, на якій робилися навчальні зображення. Опис локальних особливостей зображень пропонується здійснювати дескрипторами SURF [4].

## 5 РЕЗУЛЬТАТИ

Відповідно до параметрів пошукового алгоритму, заданих у табл. 1, було виконано настроювання екстрактора ознак за вибірковими даними 4-х класів розпізнавання з обсягом вибірки  $n_m = 100$ . На рис. 1 показано кращі рішення, отримані за ройовим алгоритмом з урахуванням модифікацій (3) та (4) протягом 5-ти ітерацій настроювання екстрактора.

Таблиця 1 – Параметри ройового алгоритму оптимізації для різних контурів системи

Контур	Максимальна швидкість частинок, $V_{\max,i}$	Перша константа прискорення, $c_1$	Друга константа прискорення, $c_2$	Кількість агентів, $n_a$	Вага інерції, $w$	Кількість ітерацій, $K_{ITER}$
Настроювання екстрактора	2	1	1	5	0,95	5
Оптимізація СКД	2	1	1	100	0,95	50
Сканування зображення	2	1	1	10	0,95	5

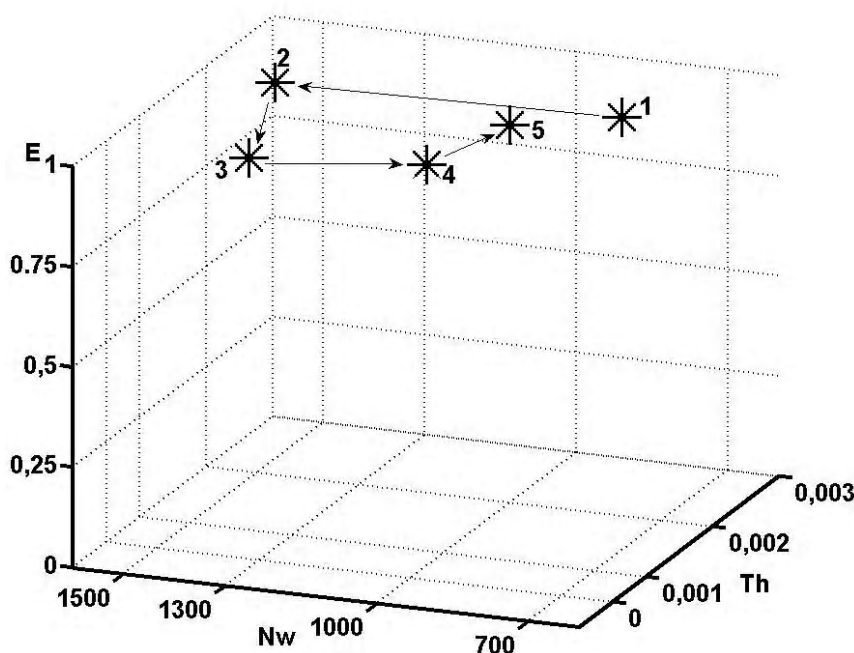


Рисунок 1 – Тривимірний графік залежності інформаційного критерію (8) від кращих позицій агентів рою протягом 5-ти ітерацій роювого алгоритму настроювання параметрів екстрактора ознак

Аналіз рис. 1 показує, що кращі позиції частинок протягом 5-ти ітерацій спочатку розміщувалися в області з більшою кількістю кластерів і високим порогом відкидання локальних ознак, але остаточно алгоритм збігся до більш компромісного варіанту, який забезпечує граничне значення інформаційного критерію і характеризується середніми значеннями параметрів настройки екстрактора. При цьому оптимальна кількість кластерів становить  $N_w^* = 1000$ , а оптимальний поріг відкидання локальних ознак рівний  $T_H^* = 0,002$ .

В табл. 2 показано динаміку зміни інформаційного критерію та точнісних характеристик вирішальних правил при оптимальних настройках екстрактора ознак в процесі роювої оптимізації полів контрольних допусків на значення ознак.

Аналіз табл. 2 показує, що практично з першої ітерації роювого пошуку було отримано високодостовірні вирішальні правила, однак граничне значення інформаційного критерію було отримано на третій ітерації.

Результати сканування 100 тестових зображень за відомими та запропонованими алгоритмами показано в табл. 3. При цьому роювий та RAWs алгоритми використовують незмінне значення розміру скануючого вікна (100x100 пікселів), розраховане на основі матриці внутрішніх калібровочних параметрів камери та відомої висоти спостереження [12].

Таблиця 2 – Результати популяційного алгоритму оптимізації полів контрольних допусків

Номер ітерації	Значення критерію (8)	Помилка першого роду	Помилка другого роду
1	0,750	0	0,009
2	0,898	0	0,002
3	1,000	0,000	0,000
4	1,000	0,000	0,000
5	1,000	0,000	0,000

Таблиця 3 – Результати сканування тестових зображень

Алгоритм сканування	Середній час обробки одного зображення, с	Відсоток хибних спрацювань, %	Відсоток пропусків, %
RAWs	0,687	0	5
ESS	0,811	0	0
Ройовий алгоритм	0,208	0	1

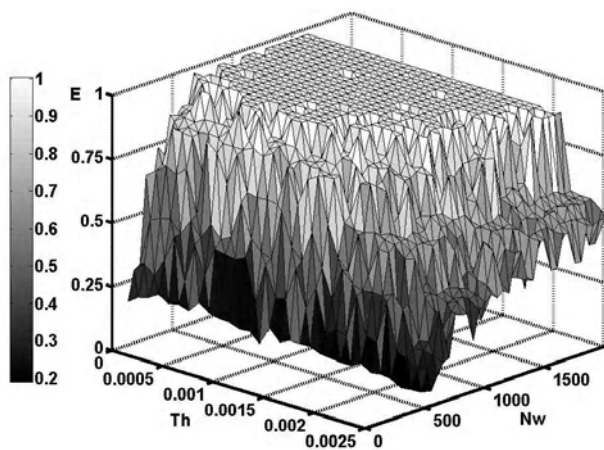
Аналіз табл. 3 показує, що використання апріорної інформації про умови спостереження за об'єктом інтересу та роювого алгоритму сканування зображення дозволяє з прийнятною для практичного використання точністю детектувати об'єкти інтересу. При цьому роювий алгоритм більше ніж втричі швидше за інші алгоритми здійснює пошук, забезпечуючи швидкість сканування Full HD відеопотоку до 5 кадрів/с.

Таким чином, запропонована модифікація алгоритму пошуку роєм частинок дозволяє за невелику кількість кроків знайти оптимальні у вартісному та інформаційному сенсах вирішальні правила. При цьому застосування роювої оптимізації контрольних допусків на значення ознак та роювого сканування зображення дозволяє здійснювати обробку відопотоку у форматі Full HD з прийнятною для практичного використання точністю та швидкістю.

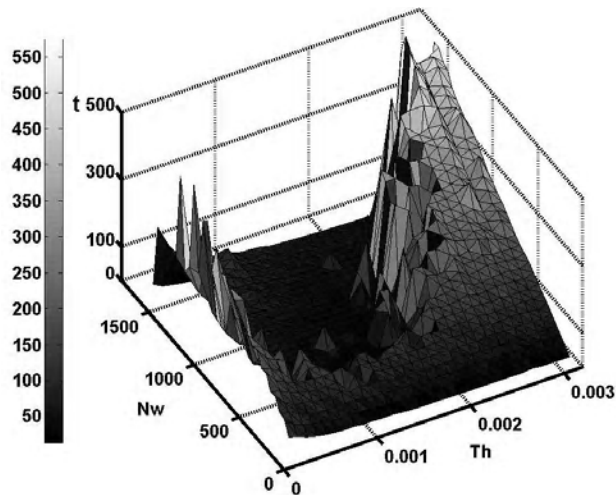
## 6 ОБГОВОРЕННЯ

Оцінка ефективності процедури настроювання в процесі машинного навчання потребує більш детального аналізу складових цієї ефективності. Для цього розглянемо область значень параметрів настройки екстрактора ознак та відповідні оцінки функціональної ефективності і часу навчання системи розпізнавання (рис. 2).

Аналіз рис. 1 та рис. 2 показує, що оптимальні значення параметрів екстрактора ознак, що визначаються на кожному кроці настроювання, знаходяться біля межі, вихід за яку супроводжується зниженням оперативності



а



б

Рисунок 2 – Тривимірні графіки залежності мір продуктивності навчання від параметрів настройки екстрактора ознак: а – інформаційний критерій функціональної ефективності вирішальних правил (8); б – час машинного навчання в секундах

та функціональної ефективності машинного навчання системи розпізнавання об'єктів на місцевості. При цьому вирішальні правила мають геометричний зміст і відновлюються у бінарному просторі Хеммінга в процесі максимізації інформаційного критерію (2).

Таким чином, запропоновані алгоритми оптимізації параметрів функціонування системи ідентифікації об'єктів на місцевості дозволяють підвищити ефективність системи як в інформаційному так і вартісному сенсах, що важливо за умов обмежених обчислювальних ресурсів та навчальних даних.

#### ВИСНОВКИ

Наукова новизна одержаних результатів:

– вперше запропоновано метод синтезу інформаційно-екстремального класифікатора з грубим двійковим кодуванням розріджених гістограм частоти появи візуальних слів на вхідному зображенні, що дозволяє одержати обчислювально ефективні вирішальні правила;

– удосконалено метод популяційного пошуку методом рою частинок для настроювання параметрів екстрактора ознак шляхом модифікації процедур оновлення персональної на глобальній кращих позицій агентів рою, що дозволяє отримати оптимальні в інформаційному та вартісному сенсі параметри функціонування системи розпізнавання за декілька ітерацій роботи алгоритму.

Практична цінність отриманих результатів для безпilotної авіації полягає у формуванні сучасної науково-методологічної основи проектування здатних навчатися автономних систем розпізнавання образів, що функціонують за умов ресурсних та інформаційних обмежень. При цьому результати імітаційного моделювання підтверджують високу ефективність отриманих вирішальних правил при розпізнаванні об'єктів інтересу на кадрах відеопотоку.

#### ПОДЯКИ

Робота виконана на базі лабораторії інтелектуальних систем кафедри комп'ютерних наук Сумського державного університету при частковій фінансовій підтримці держбюджетних науково-дослідних робіт МОН України ДР № 0109U006555 та ДР № 52.17.02-01.15/16 ЗП.

#### СПИСОК ЛІТЕРАТУРИ

1. 3D Object Recognition Based on Image Features / [K. Alhamzi, A. Survey, M. Elmogy, S. Barakat] // International Journal of Computer and Information Technology. Faizabad, India: Research and Publication Unit. – 2014. – Vol. 3, Issue 3. – P. 651–660.
2. Хуршудов А. А. Обнаружение локальных пространственных структур для распознавания изображений / А. А. Хуршудов // Научно-технические ведомости СПбГУ. Информатика. Телекоммуникации. Управление. – 2014. – № 5(205). – С. 72–82.
3. Siddhartha C. Learning Hierarchical Bag of Words using NaiveBayes Clustering / C. Siddhartha, K. Shailesh, C. V. Jawahar // 11th Asian Conference on Computer Vision, Daejeon, Korea, November 5–9, 2012: proceedings. – Springer Berlin Heidelberg, 2013. – P. 382–395. DOI: 10.1007/978-3-642-37331-2\_29.
4. SURF: Speeded Up Robust Features / [H. Bay, A. Ess, T. Tuytelaars, L. V. Gool] // Computer Vision and Image Understanding (CVIU). – 2008. – Vol. 110, No. 3. – P. 346–359.
5. Kachikian S. A Review of detector descriptors' on Object Tracking / S. Kachikian, M. Emadi // International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering. – 2016. – Vol. 5, Issue 7. DOI: 10.15662/IJAREEIE.2016.0507002.
6. Shiliang Z. Descriptive Visual Words and Visual Phrases for Image Applications / Z. Shiliang, Q. Tian, G. Hua, Q. Huang, S. Li // 17th ACM international conference on Multimedia, Beijing, China October 19–24, 2009: proceedings. – ACM New York, NY, USA, 2009. – P. 75–84. DOI: 10.1145/1631272.1631285.
7. Yongtao Y. Bag-of-visual-phrases and hierarchical deep models for traffic sign detection and recognition in mobile laser scanning data / [Y. Yongtao, L. Jonathan, W. Chenglu and other] // ISPRS Journal of Photogrammetry and Remote Sensing. – 2016. – Vol. 113. – P. 106–123. DOI: 10.1016/j.isprsjprs.2016.01.005.
8. Glauco V. P. From Bag-of-Visual-Words to Bag-of-Visual-Phrases using *n*-Grams / V. P. Glauco, J. Agma, M. Traina // SIBGRAPI '13 Proceedings of the 2013 XXVI Conference on Graphics, Patterns and Images, Arequipa, Peru, August 05–08, 2013 : proceedings. – IEEE Washington, DC, USA, 2013. – P. 304–311. DOI: 10.1109/SIBGRAPI.2013.49.
9. Multi-Layer Local Graph Words for Object Recognition / [S. Karaman, J. Benois-Pineau, R. Mé gret, A. Bugeau] // Advances in Multimedia Modeling: 18th International Conference on MultiMedia Modeling, Klagenfurt, Austria. January 2012: proceedings. – Springer Berlin Heidelberg, 2012. – P. 29–39. DOI: 10.1007/978-3-642-27355-1\_6.

10. Nadhir B. H. Bag of Words Based Surveillance System Using Support Vector Machines / B. H. Nadhir, H. Osama // International Journal of Security and Its Applications – 2016. – Vol. 10, № 4. – P. 331–346. DOI: 10.14257/ijisa.2016.10.4.30.
11. Москаленко В. В. Інформаційно-екстремальне машинне навчання системи ідентифікації об'єктів на місцевості / В. В. Москаленко, А. Г. Коробов // Журнал інженерних наук. – 2016. – Т. 3, № 1. – С. Н1–Н7.
12. Москаленко В. В. Оптимізація параметрів функціонування інтелектуальної системи ідентифікації об'єктів на місцевості / В. В. Москаленко, А. Г. Коробов // Радіоелектронні і комп'ютерні системи. – 2016. – № 2. – С. 32–39.
13. Zoran S. S. Sliding window object detection without spatial clustering of raw detection responses / S. S. Zoran, K. I. Kovač ek // 10th IFAC Symposium on Robot Control International Federation of Automatic Control Dubrovnik, Croatia September 5–7: proceedings. – IFAC. Elsevier Ltd, 2012 – Vol. 45, Issue 22. – P. 114–119. DOI: 10.3182/20120905-3-HR-2030.00192.
14. RASW: A run-time adaptive sliding window to improve Viola-Jones object detection / [F. Comaschi, S. Stuijk, T. Basten, H. Corporaal] // Seventh International Conference on Distributed Smart Cameras, Palm Springs, California, USA, October 29. – November 1, 2013: proceedings. – IEEE Washington DC, USA, 2013. DOI: 10.1109/ICDSC.2013.6778224
15. Efficient algorithms for subwindow search in object detection and localization / [A. Senjian, P. Peursum, L. Wanquan, S. Venkatesh] // IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Miami, Florida, USA, June 20–25, 2009 : proceeding, IEEE, Washington, D. C., USA, P. 264–271. DOI: 10.1109/CVPR.2009.5206822.
16. Москаленко В. В. Інтелектуальна система підтримки прийняття рішень для функціонального діагностування на гамма-камері / В. В. Москаленко, А. С. Рижова, А. С. Довбиш // Радіоелектроніка, інформатика, управління. – 2015. – № 4. – С. 52–58. DOI: 10.15588/1607-3274-2015-4-8

Стаття надійшла на редакції 02.01.2017.

Після доробки 17.01.2017.

Москаленко В. В.<sup>1</sup>, Коробов А. Г.<sup>2</sup>

<sup>1</sup>Канд. техн. наук, старший преподаватель кафедры компьютерных наук Сумского государственного университета, Сумы, Украина

<sup>2</sup>Аспирант кафедры компьютерных наук Сумского государственного университета, Сумы, Украина

### ІНФОРМАЦІОННО-ЕКСТРЕМАЛЬНИЙ АЛГОРИТМ ФУНКЦІОНУВАННЯ СИСТЕМИ РОЗПОЗНАВАННЯ ОБ'ЄКТІВ НА МЕСТНОСТІ С ОПТИМІЗАЦІЄЮ ПАРАМЕТРІВ ЕКСТРАКЦІЇ ПРИЗНАКІВ

**Актуальність** теми статті состоит в тому, що питання вибору параметрів екстрактора ознак і алгоритма класифікаційного аналізу в умовах априорної неопределенності, ресурсних і інформаційних обмежень недостатньо досліджено і в повному обсязі не вирішено до цих пор.

**Ціль** статті – підвищити в інформаційному і стоимісному сенсі ефективність функціонування автономної системи розпізнавання об'єктів на місцевості, яка функціонує в режимах навчання і екзамінації в умовах обмежених обсягів навчальної вибірки шляхом оптимізації параметрів екстрактора ознак і класифікатора спостережень.

**Методи дослідження** базуються на алгоритмах цифрової обробки і аналізу зображень для формування дескрипторів об'єктів інтересу, на принципах математическої статистики, теорії інформації для оцінки функціональної ефективності рішення правил, положеннях теорії популяційних алгоритмів для оптимізації параметрів системи сканування зображень.

**Результати** – розроблений алгоритм машинного навчання з грубим двоичним кодуванням спостережень і модифікацією роєвого алгоритму оптимізації параметрів функціонування системи розпізнавання зображень дозволяють отримувати по виборкам малого обсягу рішення правила з достовірністю, яка наближається до граничного максимального значення. При цьому експериментально показано перевагу використання роєвого алгоритму сканування зображень, яке состоит в троекратному підвищенні швидкості в порівнянні з відомими алгоритмами RASW і ESS.

**Висновки.** Предложено метод синтеза информационно-экстремального классификатора изображений с грубым двоичным кодированием разреженных гистограм частоты появления визуальных слов, что позволяет получить вычислительно эффективные решающие правила. Улучшено метод популяционного поиска для настройки параметров экстрактора признаков, что позволяет получить оптимальные в информационном и стоимісному смысле параметры функционирования системы распознавания изображений в несколько итераций работы алгоритма. Практическая ценность результатов состоит в получении рабочих алгоритмов проектирования обучающейся системы распознавания изображений, которая функционирует в условиях ресурсных и информационных ограничений.

**Ключевые слова:** словарь признаков распознавания, алфавит классов распознавания, информационный критерий, оптимізація, машинное обучение, рої частин.

Moskalenko V. V.<sup>1</sup>, Korobov A. G.<sup>2</sup>

<sup>1</sup>Ph.D., Senior Lecturer of Computer Science Department of Sumy State University, Sumy, Ukraine

<sup>2</sup>Post-graduate Student of Computer Science Department of Sumy State University, Sumy, Ukraine

### ІНФОРМАЦІОННО-ЕКСТРЕМАЛЬНИЙ АЛГОРИТМ СИСТЕМИ РОЗПОЗНАВАННЯ ОБ'ЄКТІВ НА МІСЦЕВОСТІ З ОПТИМІЗАЦІЄЮ ПАРАМЕТРІВ ЕКСТРАКЦІЇ ПРИЗНАКІВ

**Context** of this article topics is that the issue of the choice of parameters of the extractor features and classification analysis algorithm in conditions of a priori uncertainty, resource constraints, and information is not enough studied and in full has not been decided so far.

**Objective:** to increase efficiency of functioning autonomous system of recognition in information and cost sense which functions in the modes of training and examination in the conditions of limited volumes of the training dataset by optimization of parameters feature extractor and classifiers.

**Methods of a research** are based on algorithms of digital processing and analysis of images for formation descriptors of interest objects, on the principles of mathematical statistics, the theory of information for assessment of functional efficiency of decision rules, provisions on the theory based on population search algorithms for optimization parameters of scanning images system.

**Results:** the developed machine learning algorithm with rough observations binary coding and modification swarm optimization algorithm recognition system operating parameters allow to obtain for small volume dataset decision rules with reliability which approaches boundary maximum value. This experiment shows the advantage of the use swarm algorithm for scanning images, which is three-fold increase in performance compared to known algorithms RASW and ESS.

**Conclusions.** Proposed method for the synthesis of information extreme classifier of images with rough binary encoding of sparse histogram of frequency of occurrence of visual words, to provide a computationally efficient decision rules. Improved method based on population search to adjust parameters of the extractor features that allows you to get the best value in the information and cost meaning of



the parameters functioning system recognition of images in a few iterations of the algorithm work. The practical value of the results is to obtain well-functioning designing algorithms capable of learning image recognition, which operates under conditions of resource limitations and information.

**Keywords:** Dictionary features recognition, alphabet recognition class, information criterion, optimization, machine learning, particle swarm.

## REFERENCES

1. Alhamzi K. Survey A., Elmogy M., Barakat S. 3D Object Recognition Based on Image Features, *International Journal of Computer and Information Technology. Faizabad, India: Research and Publication Unit*, 2014, Vol. 3, Issue 3, pp. 651–660.
2. Hurshudov A. A. Obnaruzhenie lokal'nyh prostranstvennyh struktur dlja raspoznavanija izobrazhenij, *Nauchno-tehnicheskie vedomosti SpbGU. Informatika. Telekommunikacii. Upravlenie*, 2014, № 5(205), pp. 72–82.
3. Siddhartha C., Shailesh K., Jawahar C. V. Learning Hierarchical Bag of Words using NaiveBayes Clustering, *11th Asian Conference on Computer Vision, Daejeon, Korea, November 5–9, 2012: proceedings*. Springer Berlin Heidelberg, 2013, pp. 382–395. DOI: 10.1007/978-3-642-37331-2\_29.
4. Bay H., Ess A., Tuytelaars T., Gool L. V. SURF: Speeded Up Robust Features, *Computer Vision and Image Understanding (CVIU)*, 2008, Vol. 110, No. 3, pp. 346–359.
5. Kachikian S. Emadi M. A Review of detector descriptors' on Object Tracking, *International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering*, 2016, Vol. 5, Issue 7. DOI: 10.15662/IJAREEIE.2016.0507002.
6. Shiliang Z., Tian Q., Hua G., Huang Q., Li S. Descriptive Visual Words and Visual Phrases for Image Applications, 17th ACM international conference on Multimedia, Beijing, China October 19–24, 2009: proceedings. ACM New York, NY, USA, 2009, pp. 75–84. DOI: 10.1145/1631272.1631285.
7. Yongtao Y., Jonathan L., Chenglu W., Haiyan G., Huan L., Cheng W. Bag-of-visual-phrases and hierarchical deep models for traffic sign detection and recognition in mobile laser scanning data, *ISPRS Journal of Photogrammetry and Remote Sensing*, 2016, Vol. 113, pp. 106–123. DOI: 10.1016/j.isprsjprs.2016.01.005.
8. Glauco V. P., Agma J., Traina M. From Bag-of-Visual-Words to Bag-of-Visual-Phrases using n-Grams, *SIBGRAPI '13 Proceedings of the 2013 XXVI Conference on Graphics, Patterns and Images, Arequipa, Peru, August 05–08, 2013: proceedings*. IEEE Washington, DC, USA, 2013, – pp. 304–311. DOI: 10.1109/SIBGRAPI.2013.49.
9. Karaman S., Benois-Pineau J., Mégret R., Bugeau A. Multi-Layer Local Graph Words for Object Recognition, *Advances in Multimedia Modeling: 18th International Conference on MultiMedia Modeling*. Klagenfurt, Austria, January 2012: proceedings. Springer Berlin Heidelberg, 2012, pp. 29–39. DOI: 10.1007/978-3-642-27355-1\_6.
10. Nadhir B. H., Osama H. Bag of Words Based Surveillance System Using Support Vector Machines, *International Journal of Security and Its Applications*, 2016, Vol. 10, No. 4, pp. 331–346. DOI: 10.14257/ijisia.2016.10.4.30.
11. Moskalenko V. V., Korobov A. G. Informacijno-ekstremal'ne mashynne navchannja systemy identyfikacii' ob'jektiv na misceivosti, *Zhurnal inzhenernyh nauk*, 2016, Vol. 3, No. 1, pp. H1–H7.
12. Moskalenko V. V., Korobov A. G. Optyimizacija parametriv funkcionuvannja intelektual'noi' systemy identyfikacii' ob'jektiv na misceivosti, *Radioelektronni i komp'juterni systemy*, 2016, No. 2, pp. 32–39.
13. Zoran S. S., Kovaček K. I. Sliding window object detection without spatial clustering of raw detection responses, 10th IFAC Symposium on Robot Control International Federation of Automatic Control Dubrovnik, Croatia September 5–7: proceedings. IFAC. Elsevier Ltd, 2012, Vol. 45, Issue 22. pp. 114–119. DOI: 10.3182/20120905-3-HR-2030.00192.
14. Comaschi F., Stuijk S., Basten T., Corporaal H. RASW: A run-time adaptive sliding window to improve Viola-Jones object detection, *Seventh International Conference on Distributed Smart Cameras, Palm Springs, California, USA, October 29–November 1, 2013: proceedings*. IEEE Washington DC, USA, 2013. DOI: 10.1109/ICDSC.2013.6778224.
15. Senjian A., Peursum P., Wanquan L., Venkatesh S. Efficient algorithms for subwindow search in object detection and localization, *IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Miami, Florida, USA, June 20–25, 2009: proceeding*, IEEE, Washington, D. C, USA, pp. 264–271. DOI: 10.1109/CVPR.2009.5206822.
16. Moskalenko V. V., Ryzhova A. S., Dovbysh A. S. Intelektual'na systema pidtrymky pryjnjattja rishen' dlja funkcional'nogo diagnostuvannja na gamma-kameri, *Radio Electronics, Computer Science, Control*, 2015, No. 4, pp. 52–58. DOI: 10.15588/1607-3274-2015-4-8.