

КОНСОЛІДАЦІЯ ДАНИХ ТА ПРИЙНЯТТЯ РІШЕНЬ НА ОСНОВІ РАНЖУВАННЯ ГРУП ІДЕНТИЧНИХ КЛАСИФІКАЦІЙ В ГІБРИДНИХ СИСТЕМАХ РОЗПІЗНАВАННЯ

О. І. Захожай, В. О. Лифар, В. Г. Іванов

Східноукраїнський національний університет імені Володимира Даля

Анотація. Гібридні системи розпізнавання на основі використання сукупності образів одного об'єкту розпізнавання, що мають різну природу виникнення, дозволяють отримувати рішення в широкому діапазоні змін зовнішніх умов, рівня перешкод і викривлень вхідних даних. Однак, збільшення кількості інформаційних ознак, підвищує часову складність аналізу і прийняття рішення. Метою дослідження є розробка методу консолідації даних та прийняття рішення, який дозволяє усувати від аналізу менш інформативні дані, тим самим зменшуючи часову складність процесу обробки та прийняття рішення.

Ключові слова системи розпізнавання образів, методи прийняття рішень, консолідація даних, достовірність розпізнавання, часова складність алгоритмів розпізнавання образів, алгоритми обробки даних, програмна інженерія, інформаційні технології.

Вступ

Використання інтелектуальних алгоритмів обробки даних знаходить все більшого використання в сучасному світі, в різноманітних галузях людської діяльності. Це, насамперед, пов'язано з безсумнівними перевагами, пов'язаними з відсутністю необхідності повної формалізації задачі обробки, можливості прийняття рішень в умовах часткової невизначеності, використання різноманітних евристик. Це все значно спрощує процес прийняття рішення та дозволяє забезпечити результат з апіорно визначеним рівнем достовірності.

У випадку, коли об'єкт аналізу представляє собою складну систему, з певним рівнем стохастизму, методи і засоби інтелектуальної обробки даних стають майже єдиним інструментом отримання бажаного результату. В цьому аспекті, значну роль відіграє апарат розпізнавання образів, який останнім часом все частіше використовується для прийняття рішень в складних системах, у випадку неможливості створення математичних моделей, що описують об'єкти інформатизації [1-9].

Серед різновидів принципів і підходів до організації розпізнавання образів, відомі багатопараметричні комбіновані (гібридні) системи [8-15], які є подальшим розвитком концепції комбінованих систем [16-21], запропонованої Журавльовим Ю.І. На відміну від класичної концепції комбінованих систем, в гібридних системах об'єкт розпізнавання представляється сукупністю образів, які мають різну природу виникнення,

що дозволяє збільшити кількість ознак, на які одні й ті самі перешкоди та викривлення мають різний вплив. Це значно розширює можливість отримання достовірного рішення за умови широкої зміни умов отримання вхідних даних про об'єкт розпізнавання. Однак, важливо зазначити, що в цьому випадку збільшується кількість неоднорідних даних, що підлягають аналізу.

Розширення детальності опису об'єкта розпізнавання більш широким набором ознак дозволяє підвищити достовірність розпізнавання, що обґрунтовується загальноприйнятою глибокою статистичною теорією розпізнавання образів Вапніка-Червоненкіса [22].

Однак, збільшення кількості даних значно збільшує часову складність процесу обробки та зменшує оперативність прийняття рішень, що в багатьох прикладних застосуваннях носить критичний характер [1, 4, 5, 11-15]. В цьому аспекті, доцільним є здійснення консолідації даних з метою спрощення їхньої подальшої обробки та зменшення часової складності прийняття рішення щодо класифікації.

Відповідно до вищесказаного, можна зробити висновок, що подальше вдосконалення методів обробки даних і прийняття рішень щодо класифікації в системах розпізнавання є актуальним. Вимагає особливої уваги питання розробки ефективних методів консолідації неоднорідних даних в гібридних системах розпізнавання, використання яких дозволить підвищити оперативність обробки даних та прийняття рішень.

1. Аналіз питання та постановка завдання

В ряді робіт [23, 24] для консолідації і зменшення кількості даних, що підлягають обробці для гібридних систем розпізнавання образів про-

© Захожай О. І., Лифар В. О.,
Іванов В. Г., 2019

понується використання екстенціонально-інтенціонального методу обробки і прийняття рішення. Сутність цього методу полягає в тому, що прийняття рішення щодо класифікації може прийматися у два етапи: перший екстенціональний, коли рішення щодо класифікації приймається в результаті аналізу меншої кількості узагальнених характеристик об'єкта розпізнавання, і другий інтенціональний, який дозволяє за рахунок деталізованого аналізу ознак уточнити класифікаційне рішення, отримане на першому етапі. Мінімізація часової складності процесу прийняття рішення полягає в намаганні отримати рішення на етапі екстенціонального аналізу і не використовувати наступний – інтенціональний. Інтенціональний аналіз здійснюється тільки у випадку, коли на першому екстенціональному етапі не було отримано результату, що відповідає заданому рівню достовірності. Але, у випадку неможливості отримання достовірного результату на першому етапі, другий етап призводить до збільшення часу прийняття остаточного рішення щодо класифікації.

Аналіз цього метода вказує на те, що його ефективність, яка виражається у забезпеченні бажаного рівня достовірності при мінімізації часової складності, може бути забезпечена тільки у випадку здійснення консолідації даних шляхом усунення від аналізу найменш інформативних, за поточних умов спостереження за об'єктом розпізнавання. В цьому випадку, як на екстенціональному, так і інтенціональному етапах, прийняття рішення буде здійснюватися за меншою сукупністю, але більш інформативних даних. Це дозволить зменшити кількість даних що підлягають співставленню та мінімізувати час обробки та прийняття рішення щодо класифікації.

Також, в гібридних системах розпізнавання образів відомий метод селекції раціональної сукупності ознак [12, 14], згідно якого, в залежності від кількості даних за якими приймається рішення, може використовуватися строгий та нестрогий розподіл ознак за критерієм інформативності.

У випадку строгого розподілу, ознаки поділяються на дві групи: інформативні та неінформативні, відповідно до критерію

$$\forall P_n \in C_{\text{інф}} \Leftrightarrow D_n \leq D_R,$$

$$\forall P_n \in C_{\text{нінф}} \Leftrightarrow D_n > D_R,$$

де $C_{\text{інф}}$ і $C_{\text{нінф}}$ – відповідно, інформативний та неінформативний розподіли образів;

P_n – образ об'єкта розпізнавання, який відноситься до інформативного або неінформативного розподілу;

D_n і D_R – метрика образа P_n і репрезентативне значення метрики для розподілу, відповідно.

У випадку малої кількості даних за якими приймається рішення, в окремих випадках може бути доцільним використання нестроного розподілу, відповідно до критерію

$$\forall P_n \in C_{\text{інф}} \Leftrightarrow D_n < (D_R - \Delta D_R),$$

$$\forall P_n \in C_{\text{уінф}} \Leftrightarrow (D_R - \Delta D_R) \leq D_n \leq (D_R + \Delta D_R),$$

$$\forall P_n \in C_{\text{нінф}} \Leftrightarrow D_n > (D_R + \Delta D_R),$$

де ΔD_R – апріорно визначена норма допуску відстані для репрезентативних образів умовно інформативного розподілу $C_{\text{уінф}}$.

Такий метод дозволяє здійснювати розподіл образів за ознакою інформативності, однак ефективність його використання на пряму залежить від визначення репрезентативних мір відстані для кожного з розподілів $C_{\text{інф}}$, $C_{\text{уінф}}$, $C_{\text{нінф}}$. За відсутності формалізації цієї задачі, її успішне вирішення залежить виключно від експертної оцінки, яка повинна бути апріорно здійснена для кожної групи образів за ознакою інформативності. При цьому, формалізація задачі визначення норми відстані, для кожного з розподілів інформативності, дозволило би значно спростити використання цього методу.

Таким чином, на основі проведеного аналізу, була сформульована задача дослідження – розробка нового методу консолідації неоднорідних даних в гібридних системах розпізнавання, який дозволяє формалізувати процес визначення груп інформативних, умовно інформативних або неінформативних образів, а також здійснити селекцію ознак з метою зменшення кількості даних, що надаються для обробки і, тим самим, знизити часову складність процесу отримання рішення щодо класифікації.

2. Вирішення задачі

Виходячи з концепції гібридних систем розпізнавання, що об'єкт представляється сукупністю образів за кожним з яких він може бути віднесений до одного і того ж самого класу, пропонується для реалізації методу використовувати зважене голосування. В цьому випадку, результати класифікації за окремими образами будуть групуватися за принципом ідентичної класифікації і, в ідеальному випадку, коли відсутні прояви викривлень вхідних даних, що характеризують об'єкт, та за гарної розподіленості класів, усі образи будуть надавати ідентичну класифікацію до одного класу, до якого відноситься об'єкт. На практиці, за реальних умов, може спостерігатися неоднозначна класифікація, коли частина образів буде вказувати на інший варіант класифікації.

Таким чином, в процесі формування груп образів з ідентичною класифікацією буде визначатися матриця вигляду

$$N = \begin{bmatrix} N(C_1) \\ N(C_2) \\ \dots \\ N(C_m) \end{bmatrix}, \quad (1)$$

де $N(C_1) \dots N(C_m)$ – значення кількості класифікацій по кожному з m класів.

Тобто для кожного з m визначених класів визначається характеристика $N(C_m)$ яка відповідає кількості образів, що вказують на віднесення

об'єкту розпізнавання до цього класу. Остаточню об'єкт розпізнавання відноситься до того класу, для якого характеристика $N(C_m)$ має максимальне значення.

На рис. 1 проілюстрований метод консолідації даних на основі пошуку груп ідентичних класифікацій, в результаті якого отримується масив (1) результатів класифікації, а також здійснюється поєднання образів за ознакою ідентичності класифікації, тобто групуються образи, за якими об'єкт розпізнавання відноситься до одного й того самого класу з наявної сукупності $C_i, i \in [1, m]$.

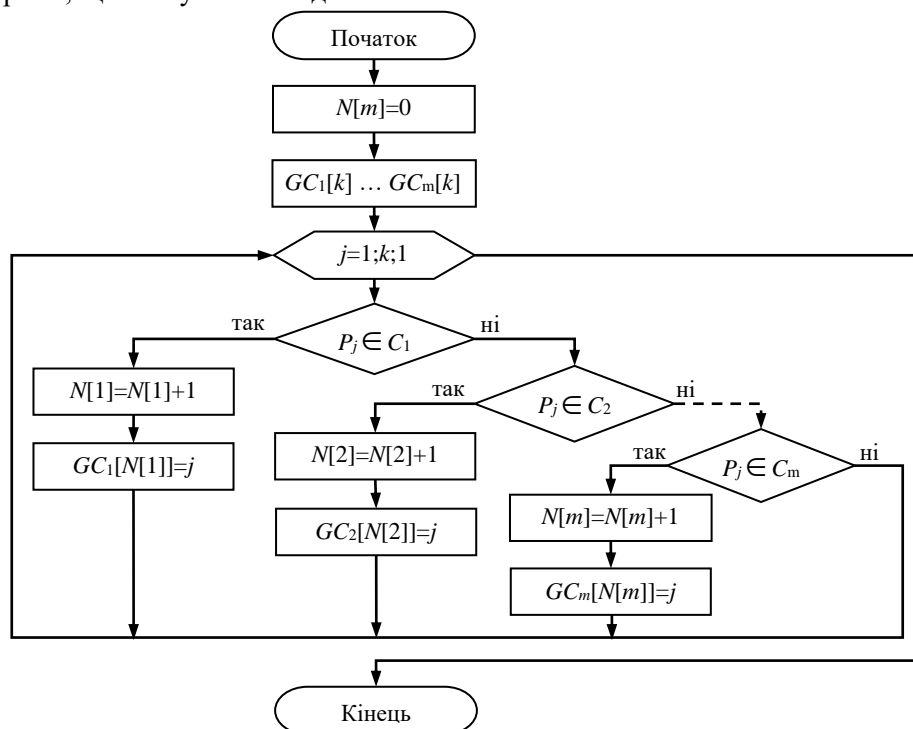


Рис. 2. Ілюстрація методу консолідації даних та прийняття рішення на основі пошуку груп ідентичних класифікацій

Згідно цього методу, спочатку здійснюється ініціалізація масиву $N[m]$, кожен елемент якого буде містити кількість образів, за якими здійснюється класифікація до кожного з m визначених класів. На початку виконання алгоритму, елементам цього масиву присвоюються нульові значення. Надалі, здійснюється ініціалізація m масивів $GC_1[k], GC_2[k], \dots, GC_m[k]$, в кожному з яких будуть міститися індекси образів, за якими об'єкт розпізнавання відноситься до одного з m передвизначених класів. Тобто ці масиви, в результаті реалізації методу, будуть містити групи образів, що відтворюють ідентичну класифікацію до кожного з сукупності m класів.

В процесі виконання алгоритму здійснюється циклічний перебір усієї сукупності k образів та у випадку класифікації до кожного з m класів

здійснюється інкрементування відповідного елемента масиву $N[m]$ та, відповідно до отриманого збільшеного значення, здійснюється розміщення відповідного індексу образу у відповідному масиві одного з m класів ($GC_1[k]-GC_m[k]$).

Максимальна кількість елементів в масивах $GC_1[k]-GC_m[k]$ дорівнює k . Тобто у випадку абсолютного розподілу класів, усі наявні класи дозволять віднести об'єкт розпізнавання до одного й того ж самого класу C_g , в цьому випадку буде досягнута максимальна достовірність розпізнавання. Критерій максимальної достовірності розпізнавання буде мати вигляд

$$D_{\max} \Leftrightarrow \begin{cases} \forall P_i \in \{P\}, P_i \in C_g, \\ N[g] = k, \\ \forall i \neq g, N[i] = 0. \end{cases} \quad (2)$$

Тобто, значення усіх елементів i масиву N , за винятком елементу g , дорівнюють 0, а значення елементу g дорівнює k . Таким чином, усі k образи вказують на класифікацію до одного класу C_g , і жоден з образів не вказує на інший варіант класифікації. В цьому випадку отримується однозначний результат розпізнавання. Однак, за реальних умов спостереження об'єкту інформаційного процесу, критерій (2) є асимптотичним і на практиці має місце недостатній опис об'єкта розпізнавання, викривлення відображення його характеристик у просторі образів, а також недосконалість систем реєстрації та вимірювання цих характеристик можуть призвести до неоднозначності класифікації. В цьому випадку, достовірність розпізнавання буде визначатися за умови

$$D \Leftrightarrow \begin{cases} \exists P_i \in \{P\}, P_i \in C_g, \\ N[g] < k, \\ \forall i \neq g, N[i] \neq 0. \end{cases}$$

В цьому випадку, найбільш достовірне рішення класифікації може бути знайдене за максимальним значенням характеристики $N(C_m)$ (у алгоритмі – значення відповідних елементів масиву $N[m]$).

Таким чином, в результаті реалізації представленої методи, окрім характеристик $N(C_m)$ визначаються групи GC_1 - GC_m образів, що дають ідентичну класифікацію за сукупністю класів C_1 - C_m . За максимумом характеристики $N(C_m)$ приймається остаточне рішення щодо розпізнавання об'єкта інформаційного процесу, а аналіз відповідної сукупності GC_m дає інформацію про сукупність найбільш інформативних образів, за яким здійснюється найбільш достовірна класифікація.

Особливий інтерес мають групи образів, що формують характеристики $N(C_m)$. З погляду на те, що кількість образів, за якими об'єкт розпізнавання відноситься до визначеного класу, на пряму впливає на достовірність, то очевидно, що групи цих образів за характеристикою $N(C_m)$ ранжуються за ознакою інформативності. Тобто, образи, що входять у найчисельнішу групу, за поточних умов спостереження об'єкта розпізнавання, мають найвищу інформативність. Таким чином, визначення характеристик $N(C_m)$ дозволяє вирішити задачу консолідації даних через селекцію образів за ознакою інформативності. У подальшому, образи, що утворюють групи з меншими характеристиками $N(C_m)$ та дають інші ре-

зультати класифікації, можуть бути усунені від розпізнавання.

Визначення груп найбільш інформативних образів (за отриманими характеристиками GC_1 - GC_m) дає можливість здійснити консолідацію даних, що надаються для аналізу, та, в подальшому, усунути від розпізнавання ті групи образів, що за поточних умов спостереження за об'єктом розпізнавання, не дозволяють отримати результат класифікації з апіорно заданим рівнем достовірності.

Якщо враховувати той факт, що кожній характеристиці $N(C_m)$ однозначно співставляється відповідна група образів з ідентичною класифікацією, то сортування $N[m]$ та отримання (1) призводить до упорядкування груп образів з ідентичною класифікацією у відповідності до рівня достовірності розпізнавання

$$N^S = \begin{bmatrix} \max N(C_m) \\ \dots \\ \min N(C_m) \end{bmatrix} \Leftrightarrow \begin{bmatrix} GC_m^I | \max N(C_m) \\ \dots \\ GC_m^m | \min N(C_m) \end{bmatrix} \Leftrightarrow \begin{bmatrix} \{P_i\}^I, i \in GC_m^I \\ \dots \\ \{P_j\}^m, j \in GC_m^m \end{bmatrix}, \quad (3)$$

де $GC_m^I | \max N(C_m)$ – група індексів образів, для якої виконується умова максимуму кількості елементів (група з найбільшою кількістю образів, що забезпечують ідентичну класифікацію);

$GC_m^m | \min N(C_m)$ – група індексів образів, для якої виконується умова мінімуму кількості елементів (група з найменшою кількістю образів, що забезпечують ідентичну класифікацію);

$\{P_i\}^I$ – сукупність образів, індекси i яких входять до групи GC_m^I (сукупність образів найчисельнішої групи, що забезпечують найбільшу достовірність розпізнавання);

$\{P_j\}^m$ – сукупність образів, індекси яких входять до групи GC_m^m (сукупність образів найменш чисельнішої групи, що забезпечують найменшу достовірність розпізнавання).

Таким чином, для гібридних систем розпізнавання образів запропонований новий метод консолідації даних та прийняття рішення щодо класифікації, який базується на пошуку та аналізі груп образів з ідентичною класифікацією.

Запропонований метод може також бути використаний для апріорного визначення груп образів об'єкту розпізнавання, що за поточних умов спостереження, дозволяють отримати найдостовірніший результат класифікації. Таким чином, консолідація даних здійснюється шляхом усунення менш інформативних образів, для яких отримані менші значення $N(C_m)$ (елементи масиву GC_m). З цією метою, під час синтезу системи обробки інформації і управління на основі гібридного розпізнавання, за різних умов навколишнього для об'єкта середовища пред'являють репрезентативні образи. В результаті, для кожних умов спостереження визначаються характеристики $N(C_m)$ і GC_m . Надалі, відповідно до (3) здійснюється ранжування груп образів за рівнем інформативності. В системі фіксується, які групи образів більш інформативні за поточних умов, або менш інформативні.

В подальшому, у життєвому циклі системи, та з урахуванням наявних умов спостереження, менш інформативні групи образів усуваються від розпізнавання.

Таким чином, у будь-який момент часу процес розпізнавання базується на аналізі меншої кількості образів найбільш інформативних груп, що забезпечує зменшення часової складності аналізу та прийняття управлінських рішень, при одночасному збереженні апріорно заданого рівня достовірності.

Висновки

Основні результати представлені в статті полягають в наступному.

1. Враховуючи неоднорідність даних при гібридному розпізнаванні образів є доцільним здійснення консолідації даних шляхом усунення від розпізнавання тих образів, які за поточних умов спостереження за об'єктом розпізнавання мають меншу інформативність.

2. Консолідація даних шляхом усунення від аналізу менш інформативних образів забезпечує зменшення кількості співставлень ознак при розпізнаванні та зменшує часову складність отримання результату класифікації.

3. Розроблений новий метод консолідації даних та прийняття рішень на основі пошуку груп ідентичних класифікацій для гібридних систем розпізнавання образів, який дозволяє здійснити консолідацію даних що надаються для аналізу через усунення менш інформативних образів, тим самим забезпечуючи прийняття достовірного рішення на основі меншої кількості, але найбільш інформативних даних.

4. Розглянутий метод може бути також використаний при апріорному аналізі для визна-

чення ступеня інформативності кожного з m образів при різних умовах отримання вхідних даних про об'єкт розпізнавання. Надалі, ці дані використовуються при апостеріорному аналізі та прийнятті рішення щодо класифікації.

5. Ефективність запропонованого методу була підтверджена експериментально для трьох різноманітних інформаційних систем: аналізу просторового розподілу температури коксового пирога [25], ультразвукового вимірювання лінійних відстаней [26] та перевірки текстових даних на унікальність [27].

Список використаної літератури

1. Симанков, В. С. Адаптивное управление сложными системами на основе теории распознавания образов [Текст] / В. С. Симанков, Е. В. Луценко. – Краснодар: Техн. ун-т Кубан. гос. технол. ун-та, 1999. – 318 с.

2. Leigh, W. Forecasting the NYSE composite index with technical analysis, pattern recognizer, neural network, and genetic algorithm: a case study in romantic decision support [Text] / W. Leigh, R. Purvis, J. M. Ragusa // Decision Support Systems. – 2002. – vol. 32. – pp. 361–377.

3. Wen, Q. Automatic stock decision support system based on box theory and SVM algorithm [Text] / Q. Wen, Z. Yang, Y. Song, P. Jia // Expert systems with Applications. – 2010. – vol. 37. – pp. 1015–1022.

4. Liu, Yang-Yu. Control principles of complex systems [Text] / Yang-Yu Liu, Albert-Laszlo Barabasi // Review of modern physics. – 2016. – vol. 88. – iss. 3. – 58 p.

DOI: 10.1103/RevModPhys.88.035006

5. Kyriakos, G. Vamvoudakis. Control of complex systems [Text] / G. V. Kyriakos, J. Saragapani. – Elsevier Inc., 2016. – 762 p.

DOI: 10.1016/C2015-0-02422-4

6. Fang, Le. Expectation propagation with stochastic kinetic model in complex interaction systems [Text] / Le Fang, Fan Yang, Wen Dong, Tong Guan, Chunming Qiao // 31st International Conference on Neural Information Processing Systems. – 2017. – pp. 2026–2036.

7. Yang, Fan. Optimal Control of Complex Systems through Variational Inference with a Discrete Event Decision Process [Text] / Fan Yang, Bo Liu, Wen Dong // 18th International Conference on Autonomous Agents and Multiagent Systems. – 2019. – pp. 296–304.

8. Меньяйленко, О. С. Основы синтеза классификаторов технических систем распознавания образов с использованием моделей эмоциональных процессов людини [Текст] / О. С. Меньяйленко, О. І. Захожай // Науково-технічний журнал «На-

уковий вісник Національного гірничого університету». – 2015. – № 1(145). – С. 120–126.

9. Загоруйко, Н. Г. Прикладные методы анализа данных и знаний [Текст] / Н. Г. Загоруйко. – Новосибирск: Изд-во Ин-та математики, 1999. – 270 с.

10. Рябенський, В. М. Комбіновані системи розпізнавання образів [Текст] / В. М. Рябенський, О. І. Захожай // Журнал «Проблеми інформаційних технологій». – 2011. – С. 156–160.

11. Захожай, О. І. Спільний аналіз інформаційних ознак у багатопараметричних комбінованих системах розпізнавання образів [Текст] / О. І. Захожай // Журнал «Електротехнічні та комп'ютерні системи». – 2019. – № 27(103). – С. 78–86.

12. Захожай, О. І. Селекція раціональної сукупності образів в комбінованих системах розпізнавання [Текст] / О. І. Захожай // Журнал «Електротехнічні та комп'ютерні системи». – 2013. – № 09(85). – С. 186–192.

13. Захожай, О. І. Критерії визначення інформативності та ранжування образів при прийнятті рішень в багатопараметричних комбінованих системах розпізнавання [Текст] / О. І. Захожай // Журнал «Електротехнічні та комп'ютерні системи». – 2018. – № 27(103). – С. 196–204.

14. Захожай, О. І. Концепція подальшого вдосконалення теоретико-методологічних основ синтезу інформаційних технологій автоматизованої обробки інформації і управління складними системами [Текст] / О. І. Захожай // Вісник Східноукраїнського національного університету імені Володимира Даля. – 2018. – № 6(246). – С. 49–55.

15. Захожай, О. І. Раздельный анализ информационных признаков в многопараметрических комбинированных системах распознавания образов [Текст] / О. І. Захожай, А. С. Меньяйленко, В. А. Лыфарь // Problemele Energeticii Regionale. – Chişinău, Republica Moldova. – 2019. – № 1-1(40). – С. 60–68. DOI: 10.5281/zenodo.3239140

16. Журавлев, Ю. И. Распознавание. Математические методы. Программная система. Практические применения. [Текст] / Ю. И. Журавлев, В. В. Рязанов, О. В. Сенько. – М.: Фазис, 2005. – 159 с.

17. Барабаш, Ю. Л. Коллективные статистические решения при распознавании [Текст] / Ю. Л. Барабаш. – М.: Радио и связь, 1983. – 224 с.

18. Горелик, А. Л. Методы распознавания. Изд. 2. [Текст] / А. Л. Горелик, В. А. Скрипкин. – М.: Высшая школа, 1984. – 219 с.

19. Garcia-Pedrajas, N. Improving multicast pattern recognition by the combination of two

strategies [Text] / N. Garcia-Pedrajas, D. Ortiz-Boyer // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2006. – vol. 28. – pp. 1001–1006.

20. Popovici, Eduard C. Combined use of pattern recognition algorithms for keystroke-based continues authentication system [Text] / Eduard C. Popovici, Liviu A. Stancu, Ovidiu G. Guta, Stefan C. Arseni, Octavian Fratu // 10th International Conference on Communications (COMM). – 2014. – pp. 91–109.

DOI: 10.1109/ICComm.2014.6866686

21. Цыпкин, Я. З. Основы информационной теории идентификации. [Текст] / Я. З. Цыпкин. – М.: Наука, 1984. – 520 с.

22. Cherkassky, V., Mulier F. Vapnik-Chervonenkis (VC) learning theory and its applications [Text] / V. Cherkassky, F. Mulier // IEEE Transactions on Neural Networks. – 1999. – vol. 10. – pp. 985–987. DOI: 10.1109/TNN.1999.788639

23. Захожай, О. І. Екстенціонально-інтенсіональний підхід до синтезу інформаційних технологій автоматизованої обробки інформації і управління на базі багатопараметричних комбінованих систем розпізнавання образів [Текст] / О. І. Захожай // Журнал «Проблеми інформаційних технологій». – 2015. – № 02(018). – С. 106–111.

24. Патент 100078 Україна, МПК (2015.01) G06K 9/00. Спосіб розпізнавання образів [Текст] / О. С. Меньяйленко, П. І. Бідюк, О. І. Захожай (Україна). – опубл. 10.07.2015, бюл. № 13.

25. Меньяйленко, О. С. Комбіновані системи розпізнавання образів при аналізі просторового розподілу температури коксового пирога [Текст] / О. С. Меньяйленко, О. І. Захожай // Журнал «Електротехнічні та комп'ютерні системи». – 2013. – № 12(88). – С. 147–154.

26. Меньяйленко, О. С. Інформаційна технологія автоматизованої обробки даних в багатоканальних системах ультразвукового вимірювання [Текст] / О. С. Меньяйленко, О. І. Захожай // Наукові вісті Національного технічного університету України «Київський політехнічний інститут». – 2014. – № 6 (98). – С. 62–67.

27. Меньяйленко, О. С. Підвищення достовірності перевірки унікальності текстів шляхом використання комбінованих систем розпізнавання образів [Текст] / О. С. Меньяйленко, О. І. Захожай // Журнал «Системні дослідження та інформаційні технології». – 2017. – № 4. – С. 29–37.

References

1. Simankov, V. S., Lutsenko, E. V., (1999), Adaptive management of complex systems based on the theory of pattern recognition [Adaptivnoe uprav-

lenie slozhnymi sistemami na osnove teorii raspoznavaniya obrazov], Techn. univ. Kuban state technol. univ., Krasnodar, 318 p.

2. Leigh, W., Purvis, R., Ragusa, J. M., (2002), Forecasting the NYSE composite index with technical analysis, pattern recognizer, neural network, and genetic algorithm: a case study in romantic decision support. *Decision Support Systems*, vol. 32, pp. 361–377.

3. Wen, Q., Yang, Z., Song, Y., Jia, P., (2010), Automatic stock decision support system based on box theory and SVM algorithm. *Expert systems with Applications*, vol. 37, pp. 1015–1022.

4. Yang-Yu, Liu, Albert-Laszlo, Barabasi, (2016), Control principles of complex systems. *Review of modern physics*, vol. 88, iss. 3., 58 p. DOI: 10.1103/RevModPhys.88.035006.

5. Kyriakos, G. Vamvoudakis, Sarangapani Jagannathan, (2016), Control of complex systems. Elsevier Inc., 762 p. doi: 10.1016/C2015-0-02422-4

6. Fang, Le, Yang, Fan, Dong, Wen, Guan, Tong, (2017), Expectation propagation with stochastic kinetic model in complex interaction systems. *31st International Conference on Neural Information Processing Systems*, pp. 2026–2036.

7. Yang, Fan, Liu, Bo, Dong, Wen, (2019), Optimal Control of Complex Systems through Variational Inference with a Discrete Event Decision Process. *18th International Conference on Autonomous Agents and Multiagent Systems*, pp. 296–304.

8. Menyaylenko, O. S., Zakhozhay, O. I., (2015), Synthesis fundamentals of classifiers for technical systems of patterns recognition with the use of human's models of emotional processes. *Scientific Bulletin of National Mining University [Osnovy syntezy klasyfikatoriv tehniknykh system rozpoznavannya obraziv z vykorystannyam modelej emotsiinykh protsesiv liudyny. Naukovii visnyk natsionalnogo girnychogo universytetu]*, vol. 1(145), pp. 120–126.

9. Zagoruyko, N. G., (1999), Applied methods of data and knowledge analysis [Prikladnye metody analiza dannykh i znaniy], Publ. Inst. of mathematics, Novosibirsk, 270 p.

10. Ryabenkyi, V. M., Zakhozhay, O. I., (2011), Combined systems of patterns recognition. *Problems of information technologies [Kombinovani systemy rozpoznavannya obraziv. Problemy informatsiinykh tekhnolohii]*, vol. 1(009), pp. 152–157.

11. Zakhozhay, O. I., (2019), Joint analysis of informational signs in multi-parametric combined patterns recognition systems. *Electrotechnical and computer systems [Spilnii analiz informatsiinykh oznak u bagatoparmetrychnykh kombinovanykh systemakh rozpoznavaniya obraziv. Elektrotekhnichni ta kompiuterni systemy]*, vol. 29(105), pp. 78–86.

12. Zakhozhay, O. I., (2013), The rational aggregate selection of informative patterns in the combined recognition systems. *Electrotechnical and computer systems [Seleksiia ratsionalnoi sukupnosti obraziv v kombinovanykh systemakh rozpoznavaniya. Elektrotekhnichni ta kompiuterni systemy]*, vol. 09(85), pp. 186–192.

13. Zakhozhay, O. I., (2015), Criteria for the determination of informativity and patterns ranking for making decisions in multi-parametric combined recognition systems. *Electrotechnical and computer systems [Kriterii vyznacheniya informatyvnosti ta ranzhuvaniya obraziv pry pryjniatti rishen v bagatoparmetrychnykh kombinovanykh systemakh rozpoznavaniya. Elektrotekhnichni ta kompiuterni systemy]*, vol. 27(103), pp. 196–204.

14. Zakhozhay, O. I., (2018), The concept of improving the theoretical and methodological foundations of synthesis information technologies for automated information processing and complex systems controlling. *Bulletin of Volodymyr Dahl East Ukrainian University [Kontsepsiia podalshogo vdoskonaleniya teoretyko-metodologichnykh osnov syntezy informatsiinykh tekhnologii avtomatyzovanoi obrobky informatsii i upravlinnia skladnymi systemamy. Visnyk Skhidnoukrainskogo natsionalnogo universytetu imeni Volodymyra Dalia]*, vol. 6(246), pp. 49–55.

15. Zakhozhay, O. I., Menyaylenko, A. S., Lyfar, V. A., (2019), Separate Analysis of Informational Signs in Multi-Parametric Combined Patterns Recognition Systems. *Problemele Energeticii Regionale [Razdelnii analiz informatsiinykh priznakov v mnogoparmetrycheskikh kombinovanykh sistemakh rozpoznavaniya obraziv. Problemy regionalnoj energetiki]*, vol. 1-1(40), pp. 60–68. DOI: 10.5281/zenodo.3239140

16. Zhuravlev, Yu. I., Ryazanov, V. V., Senko, O. V., (2005), Recognition. Mathematical methods. Programm System. Practical using [Metematicheskie metody. Programmnaya sistema. Practicheskie primeneniya], Phazis, Moscow, 159 p.

17. Barabash, Yu. L., (1983), Collective statistic solving in time recognition [Kollektivnye statisticheskie reshenia pri raspoznavanii], Radio and communication, Moscow, 224 p.

18. Gorelik, A. L., Skripkin, V. A., (1984), Recognition methods. Ed. 2. [Metody raspoznavaniya Izd. 2], High School, Moscow, 219 p.

19. Garcia-Pedrajas, N., Ortiz-Boyer, D., (2006), Improving multicast pattern recognition by the combination of two strategies. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, pp. 1001–1006.

20. Eduard, C. Popovici, Liviu A. Stancu, Ovidiu, G. Guta, Stefan, C. Arseni, Octavian, Fratu,

(2014), Combined use of pattern recognition algorithms for keystroke-based continues authentication system. 10th International Conference on Communications (COMM), pp. 91–109.

DOI: 10.1109/ICComm.2014.6866686

21. Tsypkin, Ya. Z., (1984), Base of information technology of identification [Osnovy informatsionoi tekhnologii identifikatsii], Science, Moscow, 520 p.

22. Devroye, L., Györfi, L., Lugosi, G., (1996), Vapnik-Chervonenkis Theory. A Probabilistic Theory of Pattern Recognition. Stochastic Modelling and Applied Probability, vol. 31, pp. 187–213.

DOI: 10.1007/978-1-4612-0711-5_12

23. Zakhzhay, O.I., (2015), The extensional-intensional approach to the synthesis of information technology of automated data processing and management based on multiparameter combined recognition systems. Journal “Problems of information technologies” [Ekstensionalno-intensionalnyi pidkhid do syntezy informatsiinykh tekhnologii avtomatyzovanoi obrobky informatsii i upravlinnia na bazi bagatoparametrynykh kombinovanykh system rozpoznavannia obraziv. Jurnal “Problemy informatsiinykh tekhnologii”], 02(018), pp. 106–111.

24. Menyaylenko, O. S., Zakhzhay, O. I., Bidiuk, P. I., Patent of Ukraine 100078 IPC (2015.01) G06 9/00 publ. 10.07.2015, bull. 13.

25. Menyaylenko, O. S., Zakhzhay, O. I., (2013), The Combined systems of patterns recognition in monitoring temperature's spatial distribution for coke pie. Electrotechnical and computer systems [Kombinovani systemy rozpoznavannia obraziv pry analizi prostorovoho rozpodilu temperatury koksovoho pyroha. Elektrotekhnichni ta kompiuterni systemy], vol. 12(88), pp. 147–154.

26. Menyaylenko, O. S., Zakhzhay, O. I., (2014), The information technology of automated data processing in the multi-channel ultrasonic measurement systems. Research bulletin of National technical university of Ukraine “Kyiv polytechnic institute” [Informatsiina tekhnolohiia avtomatyzovanoi obrobky danykh v bahatokanalnykh systemakh ultrazvukovoho vymiriuvannia. Naukovi visti Natsionalnoho tekhnichnoho universytetu Ukrainy “Kyivskiy politekhnichnyi instytut”], vol. 6(98), pp. 62–67.

27. Menyaylenko, O. S., Zakhzhay, O. I., (2017), Reliability increasing of texts unique verifying by using the combined patterns recognition systems. Journal “System researching and information technologies” [Pidvyschennia dosovirnosti perevirky unikalnosti tekstiv shlahkom vykorystannia kombinovanykh system rozpoznavannia obraziv. Jurnal “Systemni doslidzennia ta informatsiini tekhnologii”], vol. 4, pp. 29–37.

THE DATA CONSOLIDATION AND DECISION MAKING BY THE RANKING OF GROUPS WITH IDENTICAL CLASSIFICATION IN HYBRID RECOGNITION SYSTEMS

O. I. Zakhzhay, V. O. Lyfar, V. G. Ivanov

Volodymyr Dahl East Ukrainian National University

Abstract. Hybrid recognition systems based on the use of a set of patterns for single recognition object. Its patterns have a different origin and allow to obtain solutions in a wide range of changes in external conditions, noise level and distortions of the input data also their heterogeneity. High recognition result reliability is ensured by the presence of an extended set of signs by which a decision is made. However, in this case there is also a negative aspect - an increase the time complexity of analysis and decision making. Wherein, various signs may have different informational content for the formation of a reliable decision on the classification depending on the current conditions of observation of the object of recognition. Researches of the subject area showed that to ensure high reliability of the classification with minimal time spent, it is necessary to consolidate the data. The essence of consolidation, in this case, is to make a decision based on an analysis not of the entire set of data, but only of the set of the most informative ones. In this case, the information content of the same data may vary depending on the external conditions for obtaining data. Thus, the aim of the researching is developing a new method of heterogeneity data consolidation and decision making in hybrid combined recognition systems that allows to exclude less informative data from the analysis. It leads to reducing the time complexity of the decision-making process. The developed method is based on the formation and ranking of groups of patterns with identical classification, which allows to select the most informative of them for analysis and decision making. According to the developed method, classification is carried out according to different patterns of the recognition object and groups of patterns are formed that give a similar result. Obviously, in the ideal case, all images should give the same classification result, but in practice this is not always the case. Further, groups of images are ranked in quantitative terms and the most numerous group arrow to the most reliable result. Thus, the decision is made on the basis of the classification results provided, from the most reliable option (the most numerous group) to the less reliable.

Keywords *patterns recognition systems, decision making methods, data consolidation, recognition reliability, time complexity of patterns recognition algorithms, data processing algorithms, program engineering, information technologies.*

КОНСОЛИДАЦИЯ ДАННЫХ И ПРИНЯТИЕ РЕШЕНИЙ НА ОСНОВЕ РАНЖИРОВАНИЯ ГРУПП ИДЕНТИЧНЫХ КЛАССИФИКАЦИЙ В ГИБРИДНЫХ СИСТЕМАХ РАСПОЗНАВАНИЯ

О. И. Захожай, В. А. Лыфарь, В.Г. Иванов

Восточноевропейский национальный университет имени Владимира Даля

Аннотация. Гибридные системы распознавания на основе использования совокупности образов одного объекта распознавания, которые имеют различную природу возникновения, позволяют получать решения в широком диапазоне изменения внешних условий, уровня помех и искажений входных данных. Однако, увеличение количества информационных признаков, увеличивает временную сложность анализа и принятия решения. Целью исследования является разработка метода консолидации данных и принятия решения, который позволяет исключить из анализа менее информативные данные, тем самым снижая временную сложность процесса обработки и принятия решения.

Ключевые слова: системы распознавания образов, методы принятия решений, консолидация данных, достоверность распознавания, временная сложность алгоритмов распознавания образов, алгоритмы обработки данных, программная инженерия, информационные технологии.

Отримано 30.08.2019



Захожай Олег Ігорович, Східноукраїнський національний університет імені Володимира Даля, кандидат технічних наук, доцент, доцент кафедри програмування та математики. Просп. Центральний, 59-А, Сєвєродонецьк, Україна, E-mail: zakhozhay.oleg@gmail.com

Oleh Zakhozhay, Volodymyr Dahl East Ukrainian National University, PhD, Associate professor, associate professor of the Programming and Mathematics department, Central ave., 59-A, Severodonetsk, Ukraine, E-mail: zakhozhay.oleg@gmail.com

ORCID ID: 0000-0002-9078-3242



Лифар Володимир Олексійович, Східноукраїнський національний університет імені Володимира Даля, доктор технічних наук, доцент, завідувач кафедри програмування та математики. Просп. Центральний, 59-А, Сєвєродонецьк, Україна, E-mail: lyfarva61@ukr.net

Volodymyr Lyfar, Volodymyr Dahl East Ukrainian National University, ScD, associate professor, head of the Programming and Mathematics department, Central ave., 59-A, Severodonetsk, Ukraine, E-mail: lyfarva61@ukr.net

ORCID ID: 0000-0002-7860-9663



Іванов Віталій Геннадійович, Східноукраїнський національний університет імені Володимира Даля, кандидат технічних наук, доцент, доцент кафедри програмування та математики. Просп. Центральний, 59-А, Сєвєродонецьк, Україна, E-mail: vetgen75@gmail.com

Vitaliy Ivanov, Volodymyr Dahl East Ukrainian National University, PhD, Associate professor, Associate professor of the Programming and Mathematics Department, Central ave., 59-A, Severodonetsk, Ukraine, E-mail: vetgen75@gmail.com

ORCID ID: 0000-0002-8698-0972