

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ЧОРНОМОРСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
ІМЕНІ ПЕТРА МОГИЛИ



ЗАХОЖАЙ ОЛЕГ ІГОРОВИЧ

УДК 004.93'1:65.01:65.011.56

**МОДЕЛІ, МЕТОДИ ТА ІНФОРМАЦІЙНА ТЕХНОЛОГІЯ ГІБРИДНОГО
РОЗПІЗНАВАННЯ ОБРАЗІВ ДЛЯ КОНСОЛІДОВАНОЇ ОБРОБКИ
НЕОДНОРІДНИХ ДАНИХ В СКЛАДНИХ СИСТЕМАХ**

05.13.06 – інформаційні технології

АВТОРЕФЕРАТ
дисертації на здобуття наукового ступеня
доктора технічних наук

Миколаїв – 2020

Дисертацією є рукопис.

Робота виконана в Луганському національному університеті імені Тараса Шевченка Міністерства освіти і науки України.

Науковий консультант: заслужений діяч науки і техніки України, доктор технічних наук, професор,
МЕНЯЙЛЕНКО Олександр Сергійович,
Луганський національний університет імені Тараса Шевченка, проректор з науково-педагогічної роботи.

Офіційні опоненти: доктор технічних наук, професор,
КОВАЛЕНКО Ігор Іванович,
Чорноморський національний університет імені Петра Могили, м. Миколаїв, професор кафедри інженерії програмного забезпечення;

доктор технічних наук, професор,
ДАНИЛОВ Валерій Якович,
НТУУ «КПІ імені Ігоря Сікорського», професор кафедри математичних методів системного аналізу

доктор технічних наук, професор,
ШОСТАК Ігор Володимирович,
Національний аерокосмічний університет імені М.Є. Жуковського «ХАІ», професор кафедри інженерії програмного забезпечення

Захист дисертації відбудеться 29 травня 2020 року об 11⁰⁰ годині на засіданні спеціалізованої вченої ради Д 38.053.05 у Чорноморському національному університеті імені Петра Могили за адресою: вул. 68 Десантників, 10, м. Миколаїв, 54003.

З дисертацією можна ознайомитися у бібліотеці Чорноморського національного університету імені Петра Могили за адресою: вул. 68 Десантників, 10, Миколаїв, 54003, та на сайті <https://chmnu.edu.ua/disertatsiyi/>.

Автореферат розісланий 28 квітня 2020 року.

Вчений секретар
спеціалізованої вченої ради Д 38.053.05



Євген ДАВИДЕНКО

ЗАГАЛЬНА ХАРАКТЕРИСТИКА РОБОТИ

Актуальність теми. Вдосконалення інформаційних технологій обробки даних було і залишається одним з пріоритетних напрямів науки і техніки. Це пов'язано, насамперед, з тим, що методи і засоби обробки інформації знаходять своє використання у будь-яких сферах людської діяльності та від якості організації цього процесу багато в чому залежить кінцевий результат виконання суспільно важливих завдань в різних сферах господарювання.

Існуючі складні системи (СС) функціонують в умовах наявності цілого ряду факторів, які характеризуються різною природою, неоднорідністю та невизначеністю, що значно ускладнює інформаційні процеси, що протікають в них. Традиційно прийнято, що система визначається як складна, якщо для побудови її адекватної моделі недостатньо апріорної інформації. Це означає, що поведінка СС істотним образом залежить від факторів, які носять стохастичний характер. За цієї причини, побудова змістовної аналітичної моделі СС або взагалі неможлива, або її складність обмежує можливості її використання для задач обробки даних.

В цьому випадку, для рішення задач обробки даних широкого використання набули методи і засоби розпізнавання образів, які дозволяють формувати результат за неповним описом складного об'єкта інформатизації. Для значної кількості СС може бути ефективно використана як пряма, так і зворотна задача розпізнавання, коли через обробку вхідних даних може бути класифікований стан СС, або за характеристиками того чи іншого стану можуть бути визначені вхідні параметри, що дозволять перевести СС до бажаного стану.

Однак, окрім стохастичної поведінки та неможливості повного аналітичного опису процесів, СС характеризуються, як правило, достатньо великою кількістю неоднорідних характеристик, а також динамічністю змін умов їхнього спостереження та реєстрації. Це значною мірою впливає на час обробки даних та прийняття рішень в СС. При цьому, слід зазначити, що достовірність процесу обробки даних залежить не тільки від використаних методів, а й від ступеня інформативності даних, що обробляються, їхнього ступеня достатності чи надлишковості, а також поточного рівня завад чи перешкод.

Очевидно, що від достовірності результату розпізнавання напряду залежить точність і ефективність процесу обробки інформації. З метою підвищення достовірності розпізнавання постійно вдосконалюється методологічна і алгоритмічна база розпізнавання, використовуються алгоритми бустингу, збільшується кількість даних, за якими отримується рішення.

Питаннями теорії і методології обробки даних в СС, в тому числі на основі апарату розпізнавання образів, у різні часи займалися вчені: А. А. Харкевич, Ф. Розенблатт, В. М. Глушков, Ю. І. Журавльов, Д. О. Поспелов, В. С. Михалевич, Л. А. Растрингін, Я. З. Ципкін, К. В. Рудаков, В. Н. Вапнік, О. Я. Червоненкіс, Ю. Л. Барабаш, В. О. Скрипкін, М. З. Згуровський, І. І. Коваленко, І. В. Шостак, Р. Hart, Ф. Е. Темніков, J. Tu,

R. Gonsalez, P. Winston, G. J. McLachlan, H. Niemann, B. Ripley, U. Grenander, K. Fu, R. Duda, В. І. Васильєв, П. І. Бідюк, В. Я. Данилов, О. С. Меняйленко, В. С. Симанко, Е. В. Луценко, О. Л. Горелік, Ф. П. Тарасенко, D. A. Forsyth, J. Ponce, Е. В. Щепін, К. М. Зингерман, В. В. Лазарев та ін.

Аналіз предметної галузі показав, що вдосконалення інформаційних технологій розпізнавання образів для обробки даних в СС ускладнюється наявністю **суттєвої проблеми**, яка викликана наступним протиріччям. З одного боку, збільшення кількості інформаційних ознак, використання більш складних алгоритмів, дозволяє отримати підвищення достовірності результату аналізу, але, з іншого боку, це призводить до збільшення часової складності процесу аналізу даних та прийняття рішення. При цьому слід зазначити, що збільшення часу отримання результату аналізу даних, в багатьох випадках, носить критичний характер, так як, отриманий з істотними часовими затримками, результат втрачає свою актуальність для подальшого використання і, фактично, стає недостовірним. Таким чином намагання максимізувати достовірність результату призводить до зворотного ефекту та зменшує його цінність для подальшого використання. Ця ситуація носить найбільш критичний характер для інформаційних процесів, де прийняття рішення повинно здійснюватися в режимі реального часу. Серед таких процесів можна виділити аналіз просторового розподілу температури коксового пирога для коксохімічного виробництва, ультразвукове вимірювання лінійної відстані до рухомих об'єктів для засобів автоматизації, аналіз і прогнозування газовиділення з підроблених пластів при розробці родовищ корисних копалин, тощо.

Одним з підходів подолання вказаної проблеми є використання більш потужних обчислювальних комплексів, що дозволяють зменшити час обробки даних, не зменшуючи часову складність алгоритмів. Однак, такий підхід не дає універсального й ефективного рішення тому, що, по-перше, це пов'язано з істотним збільшенням витрат на впровадження таких систем, а по-друге – значно ускладнює створення компактних, портативних і автономних систем. Ці аспекти особливо важливі враховуючи сучасні тенденції до мініатюризації обчислювальних систем і комплексів, що значно поширює можливості їхнього повсюдного використання.

Іншим підходом забезпечення високої достовірності результату аналізу за умови мінімальних часових витрат є використання вузькоспеціалізованих систем обробки даних, які забезпечують бажані якісні показники за чітко визначених умов. В цьому випадку, спеціалізація вирішуваної задачі дозволяє зменшити часову складність алгоритму аналізу за рахунок необхідності обробки меншої кількості альтернатив. Але це значно обмежує використання таких систем за умови динамічної зміни характеристик об'єкта інформаційного процесу, умов його спостереження, а також наявності великої кількості різнорідних перешкод і викривлень.

Існуючі підходи не дозволяють повною мірою отримати рішення вказаної проблеми та створювати системи обробки даних, які забезпечують високий рівень достовірності результату за мінімальних витрат часу на процес обробки. Крім цього, особливо важливим є забезпечення достовірного результату за умови динамічної зміни умов отримання вхідних даних, а також наявності

різномірних завад і викривлень інформаційного простору в СС. В цьому аспекті **актуальним** є створення нових та вдосконалення існуючих інформаційних технологій обробки даних для комплексного вирішення висвітленої проблеми та забезпечення якісного результату за умови широкої зміни завад і викривлень, а також неоднорідності даних.

Таким чином, вдосконалення інформаційних технологій обробки даних в СС доцільно здійснювати у двох напрямках: 1) вдосконалення методологічної та алгоритмічної бази; 2) консолідації даних, яка дозволить у будь який момент часу надавати для обробки і прийняття рішень тільки найбільш інформативні, за поточних умов отримання, дані.

Зв'язок роботи з науковими програмами, планами, темами. Робота виконувалась відповідно до пріоритетних напрямів науково-дослідних робіт Луганського національного університету імені Тараса Шевченка, в рамках науково-дослідних робіт: «Розробка інформаційної технології моделювання та оцінювання фінансово-економічних ризиків із врахуванням невизначеностей різної природи» (№ ДР 0113U000650); «Розробка методології системного аналізу моделювання та оцінювання фінансових ризиків» (№ ДР 0115U000356); «GreenCo – Green Computing and Communication» (530270-TEMPUS-1-2012-1-UK-TEMPUS-JPCR); «Розробка і дослідження методів синтезу індивідуальних стратегій управління в адаптивних навчальних системах на основі World wide web-технологій» (№ ДР 0109U007908); «Розробка та дослідження методів синтезу адаптивних комп'ютерних мультимедіа систем (№ ДР 0198002639).

Мета і задачі дослідження. Метою дисертаційної роботи є забезпечення заданого рівня достовірності розпізнавання образів, з одночасним зменшенням кількості даних, що надаються для співставлення в складних системах з різнорідними проявами перешкод і викривлень.

Досягнення поставленої мети здійснюється розв'язанням наступних **основних задач:**

- аналіз існуючих методів, моделей та інформаційних технологій розпізнавання образів для обробки даних в складних системах;
- розробка принципу консолідації неоднорідних даних в складних системах з різнорідними проявами перешкод і викривлень;
- обґрунтування концепції гібридного розпізнавання образів для консолідованої обробки даних в складних системах з різнорідними проявами перешкод і викривлень;
- формалізація процесу представлення даних та розробка моделі гібридного розпізнавання образів;
- розробка моделі визначення рівня інформативності ознак в динаміці їхньої зміни;
- розробка методів консолідації інформативних ознак і розпізнавання образів для складаних систем з неоднорідними даними та різнорідними проявами перешкод і викривлень;
- на основі розроблених методів і моделей, створення інформаційної технології гібридного розпізнавання образів для консолідованої обробки даних в складних системах;

– практична реалізація розроблених моделей, методів та інформаційної технології для побудови систем обробки неоднорідних даних різноманітного спрямування.

Об’єктом дослідження є процеси консолідованої обробки неоднорідних даних в складних системах.

Предметом дослідження є моделі, методи, інформаційні технології консолідованої обробки неоднорідних даних в складних системах на основі гібридного розпізнавання образів.

Гіпотеза дослідження. Якщо об’єкт розпізнавання представити як сукупність образів, ознаки яких мають різну природу виникнення, збільшити тим самим кількість неоднорідних ознак, на які завади мають різний вплив, надалі здійснити консолідацію даних та виділити інформативні образи та їх ознаки, визначивши ті, які за поточних умов отримання даних мають найбільшу інформативність, то, в результаті аналізу, можна отримати бажану достовірність результату та мінімізувати кількість даних, що надаються для співставлення, при чому, селекція найбільш інформативних даних буде забезпечувати отримання достовірного результату в широких межах змін вхідних даних та проявів перешкод і викривлень.

Методи дослідження. В основу методології роботи покладено системний підхід, що охоплює процес підтримки прийняття рішень з використанням інформаційних технологій та його складові частини (елементи). При цьому використовувалися аналітичні методи теорії ймовірностей, розпізнавання образів, методи теорії інформації, системного аналізу та загальної теорії систем, теорії множин, зокрема теорії грубих множин, експериментальні методи, в тому числі планування експерименту, чисельні дослідження, методи експертних оцінок. Рішення оптимізаційних задач здійснювалося у просторі Парето. Для програмної реалізації запропонованих технічних рішень використовувалися реляційні бази даних, низькорівневе та об’єктно-орієнтоване програмування.

Наукова новизна отриманих результатів.

Вперше:

– побудовано ситуаційно-подійну модель гібридного розпізнавання образів, засновану на поданні характеристик складної системи у вигляді сукупності образів, ознаки яких мають різну природу виникнення, множини характеристик зовнішніх умов – як проява ситуації, статичної складової ситуації – у вигляді множини стаціонарних характеристик інформативності, динамічної складової – у вигляді нестаціонарних характеристик інформативності, та сукупності класів, як результату виконання послідовності дій по розпізнаванню;

– розроблено метод консолідації неоднорідних даних при гібридному розпізнаванні, який оснований на теорії грубих множин та визначенні стаціонарних і нестаціонарних критеріїв інформативності, в залежності від наявних перешкод і викривлень, та дозволяє здійснити селекцію образів складної системи з метою зменшення кількості співставлень даних для отримання рішення;

– розроблено метод та інформаційну модель визначення послідовності обробки даних, які базуються на узагальненні моделі емоційних процесів

людини, представленої Аткинсоном-Шифріном, для складних систем з неоднорідними даними та різнорідним проявом перешкод і викривлень, що дозволяє в динаміці системи зменшити кількість даних, які підлягають співставленню та отриманню достовірного результату;

– розроблено метод та інформаційну технологію гібридного розпізнавання образів, які основані на роздільному аналізі інформативних ознак та пошуку груп образів з ідентичною класифікацією, що дозволяє забезпечити апріорно завданий рівень достовірності результату консолідованої обробки неоднорідних даних в складних системах з різнорідним проявом перешкод і викривлень;

– розроблено метод екстенціонально-інтенціонального аналізу ознак, який оснований на прийнятті швидкого рішення по узагальненим характеристикам складної системи та використання, в разі необхідності, деталізованого аналізу даних для уточнення результату, що дозволяє зменшити кількість співставлень даних при отриманні апріорно визначеного рівня достовірності класифікації;

– розроблена інформаційна модель підтримки прийняття рішення, яка у випадку отримання багатоальтернативного результату класифікації стану складної системи, за рахунок використання інформації про групи образів з ідентичною класифікацією, дозволяє представити для експертного висновку отримані результати класифікації (альтернативи), ранжовані у порядку зменшення їхньої достовірності, а також інформацію про сукупності образів, на основі яких ці класифікації були отримані.

Отримала подальший розвиток теорія комбінованих систем розпізнавання образів, в якій опис об'єкту розпізнавання здійснюється ознаками різної природи виникнення, а не різного характеру інформації, що дозволяє здійснювати достовірну класифікацію за умови широкої зміни різнорідних перешкод і викривлень.

Удосконалено матричну інформаційну модель складної системи, в якій неоднорідні дані компонуються не в одному, а в різних шарах, що забезпечує більшу гнучкість у виборі альтернатив консолідації даних та методів їхньої обробки.

Достовірність наукових результатів підтверджується коректним використанням математичного апарату, експериментальними чисельними дослідженнями, експериментальними дослідженнями в умовах лабораторії, теоретичними й експериментальними результатами, отриманими в галузі автоматизованих систем переробки даних на основі гібридного розпізнавання образів, перевіркою запропонованих технічних рішень в трьох реальних системах автоматизованої обробки інформації різної структурної організації та призначення.

Практичне значення отриманих результатів полягає в придатності для практичного застосування розробленого підходу, методів, інформаційної технології переробки даних на базі гібридного розпізнавання образів, що дозволяє приймати достовірне і оперативне рішення в широких межах змін умов отримання вхідних даних та рівня перешкод і викривлень.

Результати досліджень були використані у коксохімічному виробництві для вдосконалення автоматизованої системи контролю просторового розподілу температури коксової печі на базі Алчевського коксохімічного заводу, для

систем ультразвукового вимірювання лінійної відстані для засобів автоматизації на базі ДП «Укрчорметавтоматика», в системі перевірки унікальності тестових даних на базі ДЗ «Луганський національний університет імені Тараса Шевченка».

Результати наукових досліджень також впроваджені в освітній процес на базі Східноукраїнського національного університету при викладанні навчальних курсів: «Технології проектування комп'ютерних систем», «Людино-машинна взаємодія», «Комп'ютерне моделювання процесів і систем», «Інженерія програмного забезпечення», «Комп'ютерні технології в науці та виробництві», «Основи системного аналізу об'єктів і процесів комп'ютеризації», «Інтелектуальний аналіз даних», «Методи та засоби комп'ютерних інформаційних технологій», а також використовуються при виконанні студентами дипломних і магістерських кваліфікаційних робіт.

Основні результати роботи можна використовувати для побудови різноманітних систем автоматизованої переробки інформації і управління, де об'єкт інформатизації представляє собою складну, багатопараметричну систему з неоднорідними даними, для якої неможлива побудова повної математичної моделі поведінки, або така модель значно переускладнена та не може використовуватися для прийняття достовірного та оперативного рішення.

Апробація результатів дисертації. Основні положення дисертаційної роботи доповідалися й обговорювалися на наукових семінарах та засіданнях кафедри інформаційних технологій та систем Луганського національного університету, наукових семінарах Донбаського державного технічного університету; VIII міжнародній науково-технічній конференції «Сучасні інформаційні та електронні технології-2007» (Одеса, 2007); X міжнародній науково-технічній конференції «Сучасні інформаційні та електронні технології-2009» (Одеса, 2009); XI міжнародній науково-технічній конференції «Сучасні інформаційні та електронні технології-2010» (Одеса, 2010); XII міжнародній науково-технічній конференції «Сучасні інформаційні та електронні технології-2011» (Одеса, 2011); XIII міжнародній науково-технічній конференції «Сучасні інформаційні та електронні технології-2012» (Одеса, 2012); міжнародній науково-технічній конференції «Інформаційні технології в освіті і управлінні-2012» (Нова Каховка, 2012); міжнародній науково-технічній конференції «Контроль і управління у складних системах-2012» (Вінниця, 2012); міжнародній науково-технічній конференції «Системний аналіз. Інформатика і управління-2013» (Запоріжжя, 2013); міжнародній науково-технічній конференції «Інформаційні технології в освіті і управлінні-2013» (Нова Каховка, 2013); науково-практичній міжнародній конференції «Інформаційні технології в освіті і управлінні-2014» (Нова Каховка, 2014), міжнародній конференції «Контроль і управління в складних системах-2014» (Вінниця, 2014); науково-практичній міжнародній конференції «Інформаційні технології в освіті і управлінні-2015» (Нова Каховка, 2015), міжнародній науковій конференції «Електротехнічні та комп'ютерні системи-2018» (Одеса, 2018), Electrotechnical and Computer Systems ELTECS-2019 (Одеса, 2019), всеукраїнській науково-практичній конференції «Електроніка та телекомунікації-2019» (Севєродонецьк, 2019).

Особистий внесок здобувача. Наукові положення, висновки, рекомендації та вимоги, викладені в дисертаційній роботі й запропоновані до захисту, розроблені особисто здобувачем. У роботах, виконаних у співавторстві, особистий внесок становить: [8, 11, 17, 32] – формуються наукові ідеї щодо вирішення проблеми, обираються та обґрунтовуються варіанти вирішення задач, аналізуються концептуальні основи використання систем розпізнавання образів для консолідованої обробки неоднорідних даних в СС; [10, 14, 40, 41] – розроблений новий підхід до обробки неоднорідних даних на базі гібридного розпізнавання образів; [1, 7, 19, 22, 23, 38, 39] – розроблені критерії, моделі, методи та інформаційна технологія консолідації і обробки даних в СС на базі гібридного розпізнавання; [4, 12, 13, 18, 29, 30, 36, 37, 42] – розроблені прикладні інформаційні технології переробки даних на базі гібридного розпізнавання для аналізу температурного режиму коксових печей, ультразвукового вимірювання лінійних відстаней для засобів автоматизації, перевірки унікальності текстових даних.

Публікації. За тематикою дисертаційного дослідження опубліковано 42 наукові праці, в тому числі: 10 праць у фахових виданнях, рекомендованих МОН України для опублікування результатів наукових робіт; 16 статей в фахових виданнях, рекомендованих МОН України та таких, що входять до наукометричних баз даних Scopus та ін.; 1 стаття в іноземному виданні, що входить до наукометричних баз даних Web of Science та ін.; 9 тез доповідей на міжнародних та всеукраїнських конференціях; отримано 6 патентів України.

Структура та обсяг роботи. Дисертаційна робота складається зі вступу, п'яти розділів, висновків, списку використаної літератури та додатків. Робота викладена на 317 сторінках машинописного тексту, 259 сторінок основного тексту, містить 49 рисунків, 7 таблиць, 8 додатків на 32 сторінках. Бібліографічний список містить 236 найменувань на 26 сторінках.

ОСНОВНИЙ ЗМІСТ ДИСЕРТАЦІЙНОЇ РОБОТИ

У вступі наводиться загальна характеристика роботи, що підкреслює та обґрунтовує її актуальність та ступінь дослідженості обраного напрямку, зв'язок з науковими програмами і темами, формулюється наукова проблема, мета і завдання досліджень, предмет і об'єкт досліджень, гіпотеза дослідження, методи дослідження, наукова новизна й практичне значення отриманих результатів.

Перший розділ присвячений аналізу методів і засобів обробки даних в складних системах.

В розділі проводиться аналіз СС, в результаті якого визначається основна проблема, пов'язана з певним рівнем невизначеності та стохастичності впливу вхідних даних на стан СС і, як наслідок, складність у виборі чи розробці математичної моделі, що забезпечує формалізований опис поведінки на основі апріорної інформації. Таким чином, для обробки даних і прийняття рішень в СС доцільним є використання також апостеріорної інформації як про саму СС, так і про зовнішнє для системи середовище.

З урахуванням того, що в СС фактично відсутній формалізований опис закономірностей впливу окремих параметрів на поведінку системи, обґрунтована доцільність підходу, коли для прийняття рішень, використовується ідентифікація не окремих параметрів, а стану. В цьому випадку для рішення задач обробки даних та прийняття рішень ефективним є використання методів і засобів розпізнавання образів, коли алфавіт класів, за якими здійснюється розподіл об'єкта розпізнавання, описує сукупність можливих станів СС. При чому, методи і алгоритми розпізнавання реалізують інтелектуальний аналіз даних, при якому не має необхідності у побудові повної математичної моделі опису СС, що вирішує раніше зазначену проблему. Таким чином, для СС найефективнішим підходом є аналіз не параметрів, а стану, на основі якого приймаються рішення.

Для обробки даних на основі розпізнавання відома матрична інформаційна модель СС, яка представлена в таблиці 1. В цій моделі елементами матриці є приватні критерії I_i^j , що відбивають вплив ознаки x_i на перехід СС до j -го стану, який характеризується класом C_j . В якості I_i^j використовується семантична міра доцільності інформації, введена А. А. Харкевичем на основі апостеріорного підходу: $I_i^j = \text{Log}_2(P_i^j / P^j)$, де P_i^j і P^j позначають, відповідно, ймовірність переходу СС до j -го стану під впливом i -ї ознаки та ймовірність випадкового (спонтанного) переходу в той самий стан.

Таблиця 1 – Матрична інформаційна модель складної системи

Ознаки СС	Стани СС			Потужність ознаки
	C_1	...	C_j	
x_1	I_1^1		I_1^j	σ_1
...
x_i	I_i^1		I_i^j	σ_i

Складні системи описуються значною кількістю неоднорідних (гетерогенних) ознак x_i . Відповідно до концепції, запропонованої Ю. І. Журавлевым, модель опису складної системи може бути представлена сукупностями ознак різного характеру інформації: $CS = \{X\}^d \cup \{X\}^p \cup \{X\}^l \cup \{X\}^s$, де $\{X\}^d$, $\{X\}^p$, $\{X\}^l$ та $\{X\}^s$ - ознаки, відповідно, детермінованого, ймовірнісного, логічного і структурного характеру інформації. В цьому випадку, для обробки даних та прийняття рішення використовується комбіноване розпізнавання, яке базується на використанні алгоритмів обчислення оцінок чи колективами вирішальних правил.

Показано, що якість інформаційного процесу обробки даних в складних системах залежить від двох основних характеристик: достовірності розпізнавання та часової складності аналізу даних. При цьому неоднорідність даних в СС створює додаткові ускладнення алгоритмічної бази прийняття рішень щодо класифікації станів СС.

Для формулювання проблеми обробки даних в СС використано елементи теорії інформації та представлення процесу розпізнавання як каналу передачі

даних. В цьому випадку, модель обробки даних має вигляд $I_d(t) = \xi(t) \cdot I(t)$, де процес обробки представляється характеристикою перешкоди $\xi(t)$, а результат обробки $I_d(t)$ є наслідком викривлення вхідного потоку даних $I(t)$. З урахуванням скінченності часу кожного етапу обробки даних, результат обробки отримується із затримкою $\Delta t = \sum_{i=0}^{n-1} (t_{i+1} - t_i)$, де n – кількість етапів

обробки даних. Тоді, за реальних умов, модель обробки даних прийме вигляд: $I_d(t + \Delta t) = \xi(t) \cdot I(t)$, тобто результат обробки даних, які характеризують СС у момент часу t може бути отриманий тільки із запізненням, у момент часу $t + \Delta t$.

Таким чином, проблема обробки даних в СС викликана протиріччям, коли ускладнення алгоритмів обробки даних, а також збільшення кількості ознак СС, за якими здійснюється обробка, дозволяє збільшити достовірність отриманого рішення, але при цьому збільшується час обробки і отримане достовірне рішення стає неактуальним, а значить і недостовірним для виконання управлінських рішень. Крім цього зазначається, що на часову складність прийняття рішення впливає також кількість даних, що підлягають обробці.

Також в цьому розділі висувається гіпотеза, стосовно якої поставлена мета може бути досягнута завдяки використанню нового принципу консолідації неоднорідних даних, а також нової інформаційної технології гібридного розпізнавання образів.

Другий розділ присвячений розробці та обґрунтуванню нового принципу консолідації неоднорідних даних, моделям і методам визначення інформативності ознак та гібридного розпізнавання образів для прийняття рішень в СС.

В розділі обґрунтовано, що в СС з різноманітними проявами перешкод і викривлень складність прийняття достовірного рішення пов'язана зі зміною рівня інформативності кожної ознаки в залежності від наявних зовнішніх впливів. Так одна й та сама ознака може бути інформативною для прийняття рішення за одних умов, але бути взагалі не інформативною за інших умов. Складність прийняття рішення, в цьому випадку, буде пов'язана з необхідністю обробки великої кількості даних, інформативність яких змінюється під впливом зовнішніх факторів. В процесі такої обробки частина ознак буде сприяти отриманню достовірного результату, а частина – навпаки, заважатиме цьому, за причини впливу на них зовнішніх факторів, перешкод та викривлень. На рис. 1 в якості прикладу, представлена ситуація з прийняття рішення стосовно визначення типу транспортного засобу за умови наявності різного рівня зовнішніх перешкод. Модель представлення об'єкту розпізнавання буде включати образ P з n ознаками візуального характеру: $P = (x_1, x_2, \dots, x_n)$.

Очевидно, що навіть у випадку обмеженого словника класів (2 класи: легковий або вантажний), не забезпечується достовірний результат в випадку різного прояву перешкод і викривлень. При чому, збільшення розмірності вектору ознак $\{x\}$, їх консолідація для визначення та підкреслення контурів або підвищення деталізації опису об'єкту розпізнавання, в цьому випадку, тільки збільшує кількість порівнянь, необхідних для отримання класифікаційного рішення, але не сприяє підвищенню достовірності, так як інформативність

візуальних ознак зменшується зі збільшенням рівня відповідних перешкод. Подібна проблема має місце у будь-яких СС з різномірними проявами перешкод і викривлень, незважаючи на природу виникнення цих ознак (візуальні, чи будь-які інші).



Рисунок 1 – Вплив різного рівня завад на достовірність прийняття рішення

Сутність нового принципу консолідації неоднорідних даних в СС полягає у формуванні не одного, а сукупності образів об'єкту розпізнавання, ознаки яких мають різну природу виникнення. В цьому випадку, розглянута вище матрична інформаційна модель підлягає трансформації, відповідно до таблиці 2. Тоді, на відміну від класичної теорії комбінованих систем розпізнавання, модель опису СС представляється сукупністю ознак не різного характеру інформації, а різної природи виникнення, на які різномірні перешкоди і завади будуть мати різний вплив:

$$CS = \begin{cases} P_1(X_{11}, X_{12}, \dots, X_{1n}), \\ P_2(X_{21}, X_{22}, \dots, X_{2m}), \\ \dots \\ P_k(X_{k1}, X_{k2}, \dots, X_{kl}), \end{cases} \quad (1)$$

де X – ознаки образів $\{P\}$, що характеризують об'єкт розпізнавання.

Таблиця 2 – Модифікована матрична інформаційна модель складної системи

Ознаки СС		Стани СС			Потужність ознаки
		C_1	...	C_j	
P_1	x_{11}	I_{11}^1		I_{11}^j	σ_{11}

	x_{1n}	I_{1n}^1		I_{1n}^j	σ_{1n}

Ознаки СС		Стани СС			Потужність ознаки
		C_1	...	C_j	
P_k	x_{k1}	I_{k1}^1		I_{k1}^j	σ_{k1}

	x_{kl}	I_{kl}^1		I_{kl}^j	σ_{kl}

Тоді, в розглянутому раніше прикладі розпізнавання типу транспортного засобу, необхідно використати k інформаційних потоків, кожен з яких дозволить отримувати ознаки різної природи виникнення (наприклад: візуальні ознаки, характеристики маси, звукові характеристики двигуна та ін.), що

дозволить сформулювати опис об'єкта, відповідно до (1) та забезпечити прийняття рішення на основі ознак, на які різні прояви перешкод та викривлень мають різний вплив. Так, у випадку наявності значних викривлень візуальних ознак, достовірне прийняття рішення можливо на основі тих ознак, на які візуальні мають малий вплив, наприклад за масою транспортного засобу чи звуком двигуна.

На основі нового принципу консолідації даних та модифікованої матричної інформаційної моделі запропонована концепція гібридного розпізнавання образів, яка є подальшим розвитком теорії комбінованих систем розпізнавання образів і, за рахунок опису об'єкту ознаками різної природи виникнення, а не різного характеру інформації, дозволяє здійснювати достовірну класифікацію за умови широкої зміни різнорідних перешкод і викривлень. Процес гібридного розпізнавання проілюстрований на рис. 2.

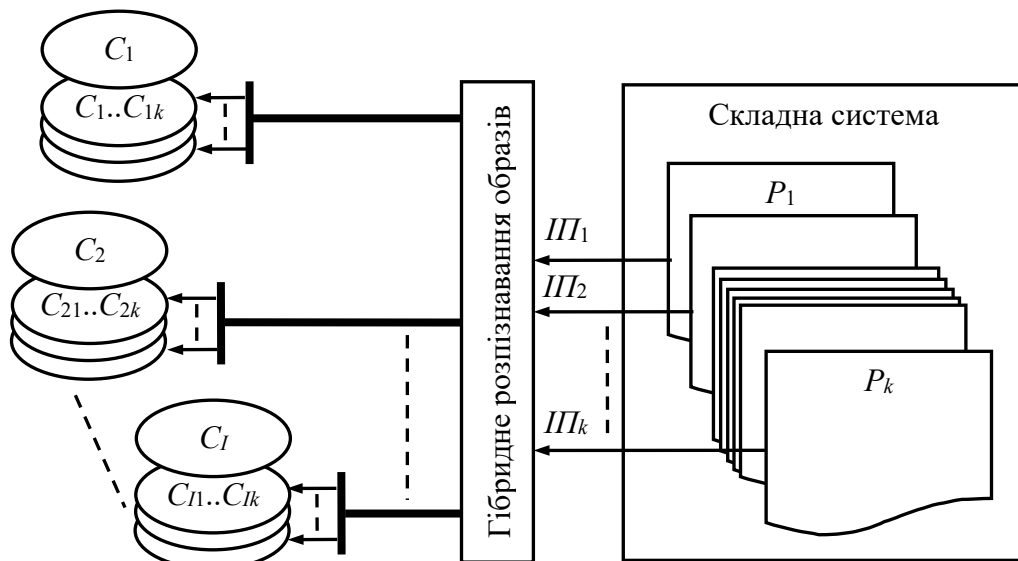


Рисунок 2 – Гібридне розпізнавання образів

При гібридному розпізнаванні об'єкт представляється сукупністю образів P_1-P_k , ознаки яких мають різну природу виникнення. Очевидно, що у випадку ідеальної, абсолютно достовірної класифікації, за кожним з k образів об'єкт розпізнавання може бути віднесений до одного з апріорно визначених класів C_1-C_l .

Особливістю опису алфавіту класів при гібридному розпізнаванні є формування сукупності репрезентативних ознак, які дозволяють однозначно співвіднести об'єкт до одного з класів C_1-C_l на підставі отримання інформації за кожним з k інформаційних потоків III_1-III_k (див. рис. 2). Відповідно до того, що CC представляється сукупністю образів $\{P\}$, які відповідають кожному окремому інформаційному потоку, класи також структуруються за класифікуючими ознаками, які відповідають кожному інформаційному потоку. В цьому випадку, кожен клас C_1-C_l повинен представляти собою сукупність підкласів C_{lk} , за якими може здійснюватися співвідношення. Таким чином, один з образів P_k за сукупністю його класифікуючих ознак $\{X_k\}_C$ може бути співставлений з відповідним підкласом C_{lk} , що дозволяє встановити відповідність об'єкта розпізнавання одному з передвизначених класів C_l .

З точки зору підвищення достовірності прийняття рішень, згідно статистичної теорії розпізнавання образів доведено, що збільшення кількості прецедентів призводить до збільшення імовірності отримання достовірної класифікації об'єкта розпізнавання. Так, відповідно до асимптотичного обґрунтування методів розпізнавання, якщо імовірність достовірної класифікації прагне одиниці, довірчий інтервал прагне нуля. Тобто, збільшення кількості інформаційних каналів (ознак об'єкту розпізнавання) та прецедентів підвищує імовірність достовірного результату розпізнавання, однак, збільшує часову складність прийняття рішень щодо класифікації. Таким чином, забезпечення високої достовірності розпізнавання та обмеження часу прийняття рішень можливо за умови збільшення кількості інформаційних ознак, які повною мірою описують СС, та подальшої їх селекції з метою визначення найбільш інформативної сукупності для визначення остаточного результату розпізнавання.

Для забезпечення прийняття рішень на основі обмеженої кількості але найбільш інформативних ознак, було розроблено ситуаційно-подійну модель гібридного розпізнавання образів вигляду

$$SEMHPR = \langle P, EC, SICCh, NSICCh, C \rangle, \quad (2)$$

де P – множина образів СС, які формуються на основі сукупності інформаційних джерел різної природи виникнення; EC – множина характеристик зовнішніх умов, що характеризують поточну ситуацію; $SICCh$ – множина стаціонарних характеристик інформативності образів СС; $NSICCh$ – множина нестаціонарних характеристик інформативності образів СС; C – множина класів, що характеризують стан СС. Згідно моделі, поточний стан СС визначається на підставі отримання сукупності її образів, ознаки яких мають різну природу виникнення. Крім цього, отримується множина характеристик зовнішніх умов, що надалі дозволяє визначити ступінь інформативності кожного образу задля прийняття достовірного рішення. Для цього, на основі множини EC визначаються стаціонарні та нестаціонарні характеристики інформативності для кожного образу з $\{P\}$.

Розроблений метод визначення інформативності образів складної системи, згідно якого, стаціонарні характеристики інформативності $SICCh$ визначаються апріорно і є незмінними характеристиками точності відтворення ознак об'єкта розпізнавання за кожним образом з $\{P\}$. Нестационарні характеристики інформативності $NSICCh$ змінюються відповідно до зміни зовнішніх умов та поточного рівня різномірних перешкод і викривлень. Таким чином, за поточних умов отримання характеристик СС інформативність кожного образу змінюється, що дозволяє усунути від подальшого аналізу образи з характеристиками інформативності меншими за апріорно задані значення. Якщо представити, що образ об'єкта розпізнавання P описується деякою функцією $P = f(x, y)$, де x, y – аргументи, що визначають характеристику об'єкта, який знаходиться у деякому просторі Ω . Тоді відображення цього об'єкта I визначається як $I = g(x', y')$, де x', y' – аргументи, що визначають характеристику образу, який знаходиться у деякому інформаційному просторі Ω' . У випадку ідеальної відображаючої системи, для будь-якої точки простору

$$NSICH_k = \frac{1}{\xi_k(x', y')}. \quad (6)$$

Загальна характеристика інформативності може бути представлена у вигляді $ICH_k = SICH_k \cdot NSICH_k$.

Для фактичного визначення чисельних значень стаціонарних та нестаціонарних характеристик інформативності образів для кожного конкретного застосування в системі необхідно провести нормування характеристик h і ξ . Враховуючи різну природу образів об'єкту, що отримуються різними відображаючими системами, на підставі статистичного аналізу з'являється можливість ідентифікації завади за окремими каналами. Крім цього, аналіз значень характеристик ξ для кожного каналу дозволяє визначити ті образи, що в поточних умовах виміру не дозволяють адекватно відображати об'єкт розпізнавання. Для підвищення достовірності розпізнавання такі образи виключаються з обробки.

Представлені критерії інформативності використовуються для ранжування образів. Тобто, у будь який момент життєвого циклу СС найбільш інформативніший образ буде мати максимум характеристики інформативності. Тоді, найбільш достовірне рішення Rd отримується за результатом аналізу образу, що задовольняє критерію інформативності $\exists P_k \in \{P\}, \max ICH_k \Rightarrow Rd$.

Ранжування образів СС здійснюється у два етапи: апіорний і апостеріорний. При апіорному ранжуванні, на основі експертної оцінки, визначається мінімально припустимий рівень стаціонарної характеристики інформативності $SICH^*$. За розглянутою вище методикою, визначаються стаціонарні характеристики інформативності для кожного з k образів ($SICH_1$ - $SICH_k$). Надалі, з множини $\{P\}$ виділяють підмножину $\{P\}'$ інформативних образів, за критерієм

$$\forall P_k \in \{P\}' \left| \{P\}' \subset \{P\}, SICH_k \geq SICH^*. \quad (7)$$

В результаті такого апіорного ранжування, визначається сукупність найбільш інформативних образів, за якими здійснюється подальша обробка та прийняття рішень. Така селекція даних має особливе значення для систем, в яких час отримання рішення має критичний характер, зокрема, коли обробка даних здійснюється в режимі реального часу. Той факт, що при гібридному розпізнаванні ставиться мета отримання апіорно завданого рівня достовірності, вказує на можливість реалізації підтримки прийняття рішень на основі систем як жорсткого так і м'якого реального часу. При цьому, для гібридного розпізнавання найбільш раціональним є реалізація концепції кооперативної багатозадачності, що дозволяє досягти високої реактивності системи реального часу (СРЧ) на зміну зовнішніх умов та стану (СС), а також передбачуваності її поведінки та ефективного використання обчислювальних ресурсів. Таким чином, результатом процесу консолідації даних для прийняття рішень є отримання чітко-визначеної черги, ранжованих у порядку зменшення рівня інформативності, даних. Обробка цієї черги дозволяє отримати бажану достовірність прийняття рішення при мінімальній кількості співставлень даних, а саме, вирішити головну мету дослідження.

Отримання чітко-визначеної черги обробки вимагає селекції та ранжування не тільки множини образів $\{P\}$ СС, а й ознак цих образів, які підлягають співставленню. З цією метою в роботі розроблено метод та інформаційну модель визначення послідовності обробки даних, які є результатом узагальнення когнітивної моделі емоційних процесів людини Аткинсона-Шифріна для СС з неоднорідними даними та різнорідними проявами перешкод та викривлень (див. рис. 3). Згідно цієї моделі, вхідний інформаційний потік F сприймається сенсорною пам'яттю, в результаті чого формується сукупність інформаційних ознак $\{X\}$. Надалі, ці ознаки піддаються селекції, де визначається сукупність більш значущих, репрезентативних ознак $\{X\}'$. Ця сукупність передається до короткострокової пам'яті, де для підвищення інформативності та ефективності запам'ятовування піддається комбінаторному аналізу та визначенню чанків (відповідно до концепції, запропонованої Дж. Мілером, чанк – це данні, пов'язані в одну смислову категорію). Для гібридного розпізнавання, аналогом чанків є образи $\{P\}$ СС. В результаті комбінаторного аналізу формуються укрупнені смислові категорії (групи образів що дають ідентичне рішення щодо класифікації), які піддаються класифікації та зберігаються у довготривалій пам'яті. Слід зазначити, що представлена модель пам'яті пояснює процес самонавчання, коли класифікація здійснюється на основі апостеріорних відомостей, що надходять від сенсорної пам'яті, а також репрезентативних образів, що вже мають у довгострокової пам'яті на момент здійснення класифікації (ап'юрні відомості).

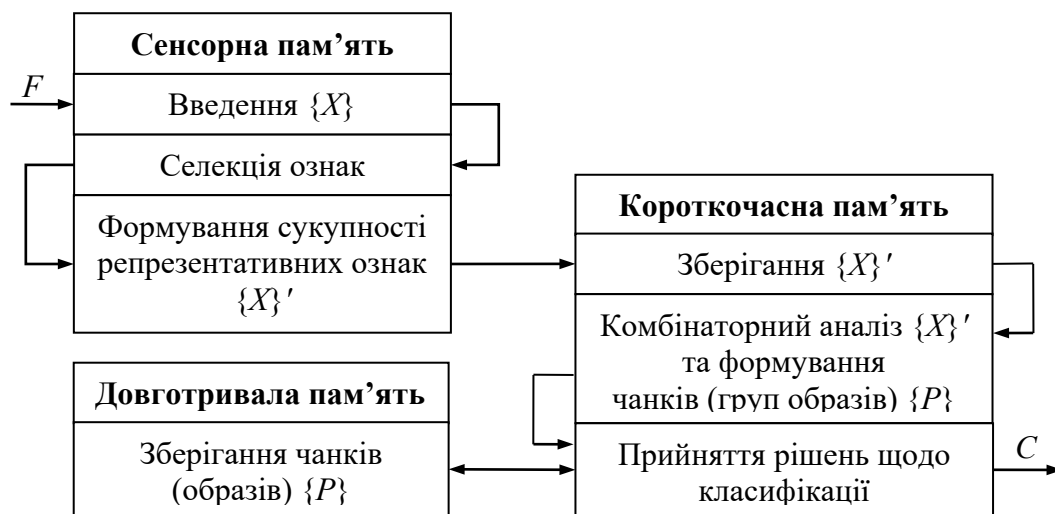


Рисунок 3 – Інформаційна модель визначення послідовності обробки даних

Емоція, як емоційний процес середньої тривалості є основною складовою для побудови «емоційних» класифікаторів, так як ці складові безпосередньо пов'язані з короткочасною пам'яттю, в якій, згідно з моделлю, представленою на рис. 3, безпосередньо здійснюється прийняття рішення. В технічних системах, представлення емоційної складової можна здійснити у вигляді функції від вхідних факторів, що характеризують стан СС та зовнішніх умов. Тоді про рівень «емоційності» можна судити виходячи зі ступеня зміни цієї функції. Тобто, для кількісної оцінки емоційної поведінки пропонується використання диференціальних характеристик. При чому, чим більше факторів визначають ступінь емоційних процесів, тим достовірніший вплив емоції на

формування остаточного результату класифікації. Для гібридного розпізнавання, що здійснюється на основі k вхідних інформаційних потоків F_1 - F_k , визначена сукупність ознак кожного з k потоків створює систему функціональних залежностей, які визначають відповідність стану СС одному з передвизначених класів:

$$P \sim \begin{cases} F_1(x_{11}, x_{12}, \dots, x_{1l}), \\ F_2(x_{21}, x_{22}, \dots, x_{2m}), \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots \\ F_k(x_{k1}, x_{k2}, \dots, x_{kn}). \end{cases} \quad (8)$$

Тоді емоційні характеристики, як ступень зміни стану, будуть визначатися у диференціальному вигляді як:

$$\begin{aligned} E_1 &= F_1'(x_{11}, x_{12}, \dots, x_{1l}), \\ E_2 &= F_2'(x_{21}, x_{22}, \dots, x_{2m}), \\ &\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ E_k &= F_k'(x_{k1}, x_{k2}, \dots, x_{kn}). \end{aligned} \quad (9)$$

Із залежності (9) витікає, що більша ступінь «емоційності» буде відповідати такому процесу, в якому спостерігається більший ступінь зміни цільової функції.

Таким чином, множина характеристик $\{E\}$ використовується для визначення черги обробки даних, при чому, по аналогії з біологічними системами, пріоритет в аналізі ознак і пошуку рішення щодо класифікації буде у характеристик, що мають найвищу характеристику «емоційності». Відповідно до цього, пріоритет розміщення даних у короткочасній пам'яті (тобто в черзі обробки) також повинен бути у характеристик з більшою «емоційною реакцією». Ця концепція повністю відповідає законам поведінки когнітивного апарату людини, коли в короткостроковій пам'яті розташовуються факти, що мають найбільший емоційний прояв.

В третьому розділі, на основі представлених моделей, розроблені методи консолідації неоднорідних даних та гібридного розпізнавання образів для СС з різнорідними проявами перешкод і викривлень.

В роботі запропоновано в якості основи для побудови методів консолідації неоднорідних даних та подальшого гібридного розпізнавання використовувати сумісний або роздільний аналіз інформаційних ознак стану СС. При спільному аналізі, здійснюється поєднання ознак СС до одного, спільного образу. В цьому випадку, задача обробки даних та прийняття рішення вирішується згідно традиційних підходів комбінованого розпізнавання, коли стан СС характеризується одним образом, однак більшою кількістю ознак, що отримуються від джерел різної природи виникнення. При роздільному аналізі, усі образи множини $\{P\}$ обробляються окремо, що забезпечує широкі можливості щодо реалізації різноманітних стратегій консолідації та обробки даних, а також використання в СРЧ з кооперативною багатозадачністю.

При спільному аналізі (рис. 4) характеристики стану СС за n інформаційними потоками Π_1 - Π_n реєструються сукупністю блоків реєстрації характеристик БРХ₁-БРХ _{n} . Ці блоки представляють собою сукупність технічних засобів вимірювання детермінованих ознак об'єктів розпізнавання, або реєстрації ознак імовірнісного, логічного або структурного характеру. За результатами реєстрації ознак блоками формування БФ(P_1)-БФ(P_n) формується сукупність образів P_1 - P_n різної природи виникнення,

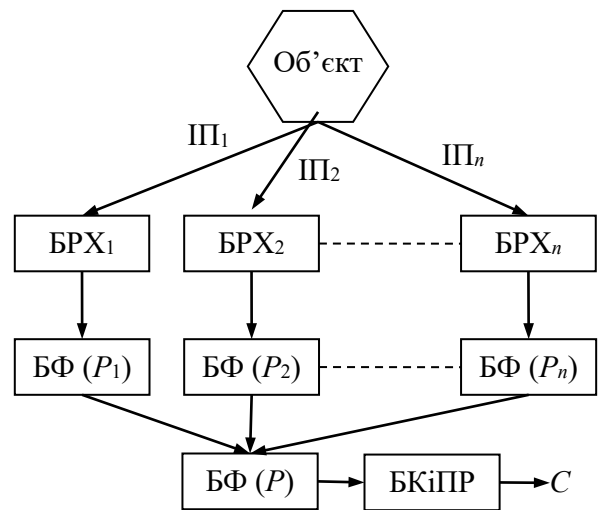


Рисунок 4 – Спільний аналіз ознак

що характеризують об'єкт розпізнавання. Надалі, ці образи поєднуються у блоці формування спільного образу БФ(P), після чого до блоку класифікації і прийняття рішення БКіПР надається спільний образ P . Результат класифікації представляється визначенням класу C , який характеризує стан СС. В цьому випадку, якщо стан СС, характеризується спільним образом $P=(x_1, x_2, \dots, x_i)$, що відноситься до класу C_m , то у просторі ознак відповідний вектор належить області, відповідної поділеної області C_m . Поділ простору ознак зводиться до побудови функцій розподілу $F_j(x_1, x_2, \dots, x_i)$, $j=1, \dots, m$, за якими може бути знайдене рішення щодо класифікації за критерієм $P \in C_j, \max F_j(x_1, x_2, \dots, x_i)$. Класифікаційне рішення може бути прийнято, якщо для інших класів виконується умова $\forall F_k(x_1, x_2, \dots, x_i), k \neq j \Rightarrow F_k(x_1, x_2, \dots, x_i) - F_j(x_1, x_2, \dots, x_i) < \theta$.

В цьому випадку, алгоритми розпізнавання ґрунтуються на порівнянні тієї чи іншої міри близькості або міри схожості об'єкта з кожним класом. Тобто рішення задачі розпізнавання здійснюється класичним шляхом. Якщо міра близькості L спільного образу P з яким-небудь класом $C_j, j=1, \dots, m$, перевищує міру його близькості з іншими класами, то приймається рішення про приналежність до класу C_j , тобто виконується умова $P \in C_j, L(P, C_j) = \text{extr } L(P, C_i), m = 1, \dots, m, i \neq j$. Міра близькості для ознак різного характеру інформації визначається наступним чином. Для

детермінованих ознак $L(P, C_j) = \sqrt{k_j^{-1} \sum_{i=1}^{k_j} \sum_{g=1}^N w_g (x_g - x_{jig})^2}$, де w_g – вагові

коефіцієнти ознак $x_g, g=1, \dots, N$, спільного образу P та ознак x_{jig} репрезентативних образів P_{gi} класу C_j . Для імовірнісних ознак, апостеріорна вірогідність того, що $P \in C_i$ визначатиметься відповідно до вираження

$$p(C_i / \text{stage}_n) = \frac{p(C_i) f_i(x_1^{\text{apost}}, x_2^{\text{apost}}, \dots, x_n^{\text{apost}})}{\sum_{i=1}^m p(C_i) f_i(x_1^{\text{apost}}, x_2^{\text{apost}}, \dots, x_n^{\text{apost}})}, \text{ де } \{x^{\text{apost}}\} - \text{ апостеріорні}$$

значення ознак образу P . Для характеристик об'єкта інформаційного процесу,

які базуються на логічних ознаках міра близькості також не використовується. В цьому випадку здійснюється ідентифікація речення, що описує об'єкт, з одним з репрезентативних речень, які є елементами опису відповідного класу. Оцінка достовірності класифікації, здійснюється за значенням довірчого інтервалу, так як при збільшенні числа прецедентів, рішення з імовірністю що прямує до 1 стає вірним, при цьому довірчий інтервал прагне 0. Тоді, задача класифікації об'єкта розпізнавання за визначеною сукупністю класів може бути здійснена з використанням підходу, запропонованого Р. Фішером. Якщо представити, що стан СС характеризується образом P , та може бути класифікований за одним з m класів при невідомих ймовірностях появи ознак цих класів p_1-p_m , то під час рішення задачі розпізнавання ці ймовірності можуть бути оцінені за навчальними вибірками. Надалі, в процесі навчання, за визначеними ймовірностями оцінюються відповідні математичні очікування μ_m та матриці коваріації Σ_m . Якщо нормальні розподіли за m класами позначити як $N(x, \mu_m, \Sigma_m)$, то достовірний результат класифікації об'єкта розпізнавання за класом C_m визначається як $\max p_m N(x, \mu_m, \Sigma_m) \Rightarrow P \in C_m$.

В основі роздільного аналізу лежить парадигма відповідності усіх образів одному об'єкту розпізнавання. При цьому, на відміну від спільного аналізу, не здійснюється перетворення простору ознак (див. рис. 5). Тобто прийняття рішення здійснюється в результаті окремої обробки консолідованих ознак кожного образу. Класифікаційне рішення може прийматися за одним або декількома образами з сукупності $\{P\}$.

Роздільний аналіз покладено в основу розробки нового методу гібридного розпізнавання, який проілюстрований на рис. 6. Згідно цього методу, на етапі E_1 здійснюється реєстрація та нормування характеристик стану СС. На наступному етапі E_2 , здійснюється формування сукупності k образів об'єкту розпізнавання (елементи P_1-P_k) на основі інформації різної природи виникнення. Також, відповідно до моделі (2) здійснюється визначення їх характеристик інформативності SIC_h і $NSIC_h$ з подальшим усуненням від обробки тих образів, що не відповідають апріорно заданому рівню інформативності. Послідовність обробки даних визначається ранжуванням образів, які відповідають заданому рівню інформативності. C_1, C_2, C_3-C_m ілюструють області розподілу ознак еталонних образів попередньо визначених m класів до яких може бути здійснена класифікація. Позначеннями E_2, E_3, E_4-E_i ,

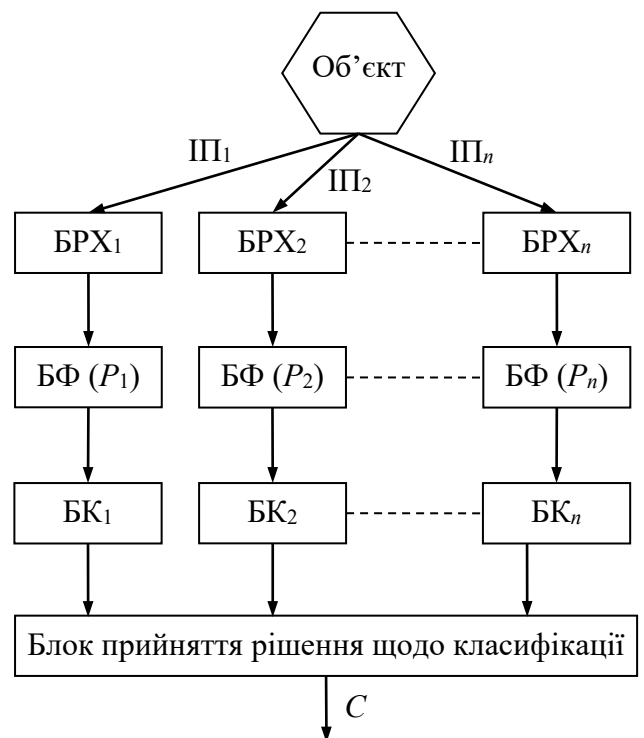


Рисунок 5 – Роздільний аналіз ознак

E_{i+1} проілюстровані етапи порівняння ознак образів P_1-P_k з еталонними ознаками класів та визначення характеристик відстані.

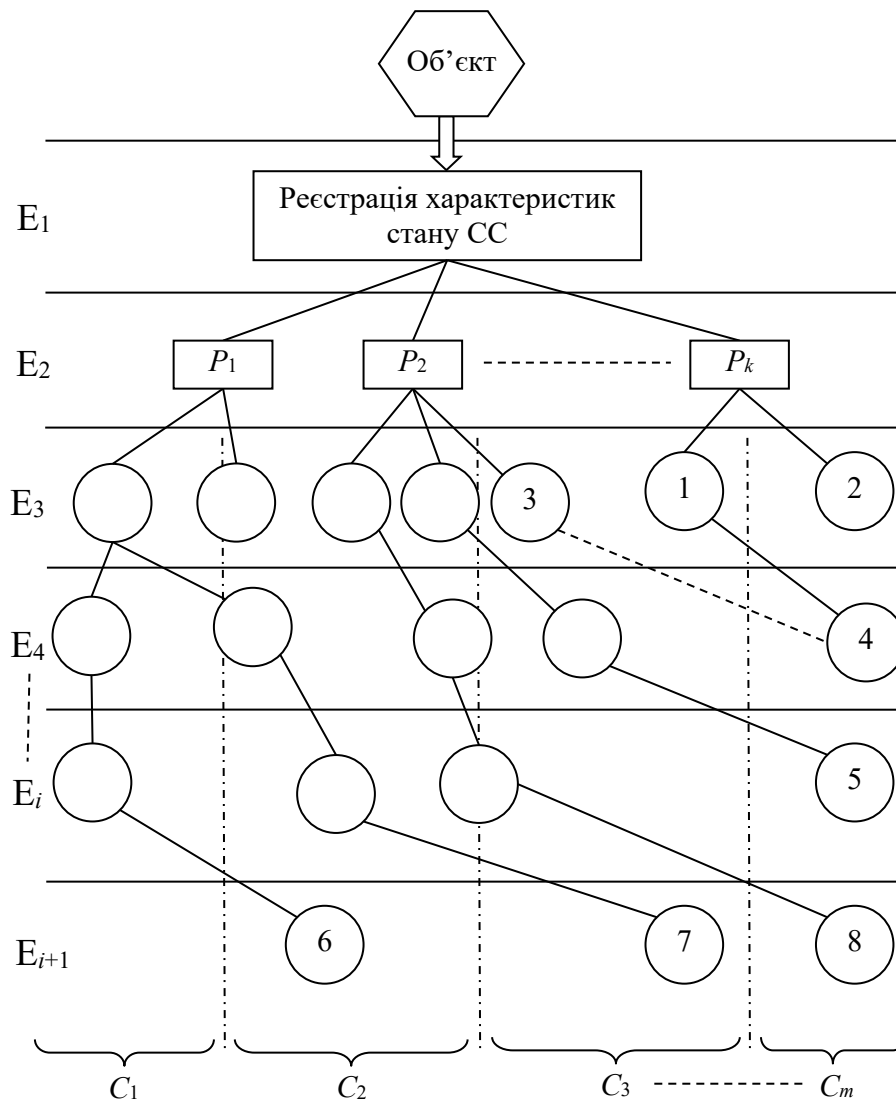


Рисунок 6 – Метод гібридного розпізнавання з роздільним аналізом ознак СС

Якщо на поточному етапі визначені відстані перевищуватиме наперед задане значення, то здійснюється наступний етап аналізу з іншими ознаками наявної сукупності образів стану СС, відповідно до черги, визначеної ранжуванням. Так, згідно рис. 6, на етапі E_3 аналіз ознак за всіма k образами не дає необхідного збігу класифікації, так як отримані відстані, включаючи відмічені цифрами 1-3, не дають однозначного збігу класифікації. Тому, надалі здійснюється етап 4, на якому подальша обробка ознак образу P_k призводить до визначення відстані 4, яка однозначно відповідає класу C_m . Крім цього, результат порівняння ознак образу P_2 дозволяє визначити відстань 4, яка також вказує на віднесення об'єкту розпізнавання до класу C_m . Таким чином, на етапі E_4 , припиняється подальша обробка і достовірним результатом вважається відповідність стану СС класу C_m . Подальший аналіз ознак (на рис. 6 етапи E_i-E_{i+1}) та отримання відстаней, таких як 5, 6, 7 або 8, не є раціональним і пов'язаний тільки з додатковими витратами машинного часу і збільшенням часу прийняття рішення.

Прийняття рішення може здійснюватися за наступними вирішальними правилами: 1) стан CC відноситься до того класу, до якого за одним з k образів отримується мінімальна відстань; 2) стан CC відноситься до визначеного класу, якщо за одним або декількома образами з сукупності k буде досягнуте значення відстані, яке не перевищує наперед заданого значення; 3) стан CC відноситься до того класу з сукупності передвизначених, до якого вказує класифікація за більшістю образів з сукупності k . В результаті порівняння ознак образів CC з репрезентативними ознаками кожного з класів отримується матриця відстані L , вигляду

$$L = \begin{bmatrix} l(P_1, C_1) & l(P_1, C_2) & \dots & l(P_1, C_m) \\ l(P_2, C_1) & l(P_2, C_2) & \dots & l(P_2, C_m) \\ \dots & \dots & \dots & \dots \\ l(P_k, C_1) & l(P_k, C_2) & \dots & l(P_k, C_m) \end{bmatrix}, \quad (10)$$

де $l(P_k, C_m)$ – відстань k -го образу об'єкту розпізнавання до m -го класу.

Для першого вирішального правила, рішення приймається за умови виконання критерію

$$\exists P_k \in C_m, l(P_k, C_m) = \min L. \quad (11)$$

Для другого вирішального правила, коли рішення приймається на основі досягнення за одним або декількома образами значення відстані, яке не перевищує наперед заданого значення, спочатку необхідне виділення з множини $\{P\}$ підмножини образів $\{P\}'$, що відповідають критерію

$$P_k \in \{P\}' \mid \{P\}' \subset \{P\}, l(P_k, C_m) \leq l^*, \quad (12)$$

де l^* – апріорно задане мінімальне значення відстані; $l(P_k, C_m)$ – відстань образу P_k до класу C_m . Надійшлиший результат визначається з урахуванням характеристик інформативності кожного образу, відповідно до критеріїв, розглянутих вище. З урахуванням вагових коефіцієнтів для кожного образу, що входить до репрезентативної сукупності, рішення щодо класифікації приймається за принципом максимуму характеристики

$$P_n \in C_g, \max \frac{Ich_n \cdot l(P_n, C_g)}{\sum_{i=1}^k \sum_{j=1}^m Ich_i \cdot l(P_i, C_j) - Ich_n \cdot l(P_n, C_g)}, \quad (13)$$

де Ich_n – характеристика інформативності для n -го класу P_n ($n \in i=1, 2, \dots, k$); C_g – клас, для якого визначається відносна оцінка ($g \in j=1, 2, \dots, m$). Тобто, об'єкт розпізнавання відноситься до того класу, для якого отримується максимальне значення відношення добутку інформаційної характеристики та відповідної відстані до визначеного класу на суму усіх інших добутків.

Для третього вирішального правила, використовується зважене голосування, коли стан СС відноситься до визначеного класу в тому випадку, коли така класифікація підтверджується за більшістю образів. Тобто для кожного з m визначених класів визначається характеристика N_m яка відповідає кількості образів, що вказують на відношення об'єкту до цього класу. Остаточню стан СС відноситься до того класу для якого характеристика N_m має максимальне значення. На рис. 7 представлений процес пошуку груп ідентичних класифікацій, в результаті якого отримується масив $N[m]$ результатів класифікації. На основі отриманої матриці N , з використанням тривіального алгоритму пошуку максимального елемента масиву, може бути знайдене остаточне рішення щодо класифікації, відповідно до критерія $P_k \in C_m, \max N(C_m)$.

З погляду на те, що кількість образів, за якими стан СС відноситься до визначеного класу, напряму вказує на достовірність, то очевидно, що групи цих образів за характеристикою $N(C_m)$ ранжуються за ознакою інформативності. Тобто, образи, що входять у найчисельнішу групу, вказують на достовірніший результат класифікації. Таким чином, визначення характеристик $N(C_m)$ дозволяє вирішити додаткову задачу ранжування результатів класифікації в порядку зменшення їхньої достовірності.

Також, для отримання додаткової інформації по сукупності класів за якими був отриманий результат класифікації в методі розраховуються m масивів $GC_1[k], GC_2[k], \dots, GC_m[k]$ в кожному з яких містяться індекси образів, за якими об'єкт розпізнавання відноситься до одного з m класів. В процесі виконання методу, здійснюється циклічний перебір сукупності k образів та у випадку класифікації до кожного з m класів здійснюється інкрементування відповідного елемента масиву N та, згідно отриманого збільшеного значення, здійснюється розміщення індексу образу у відповідному масиві, що відповідає одному з m класів ($CG_1[k]-CG_m[k]$).

В розробленому методі, критерій достовірності розпізнавання визначається за умови

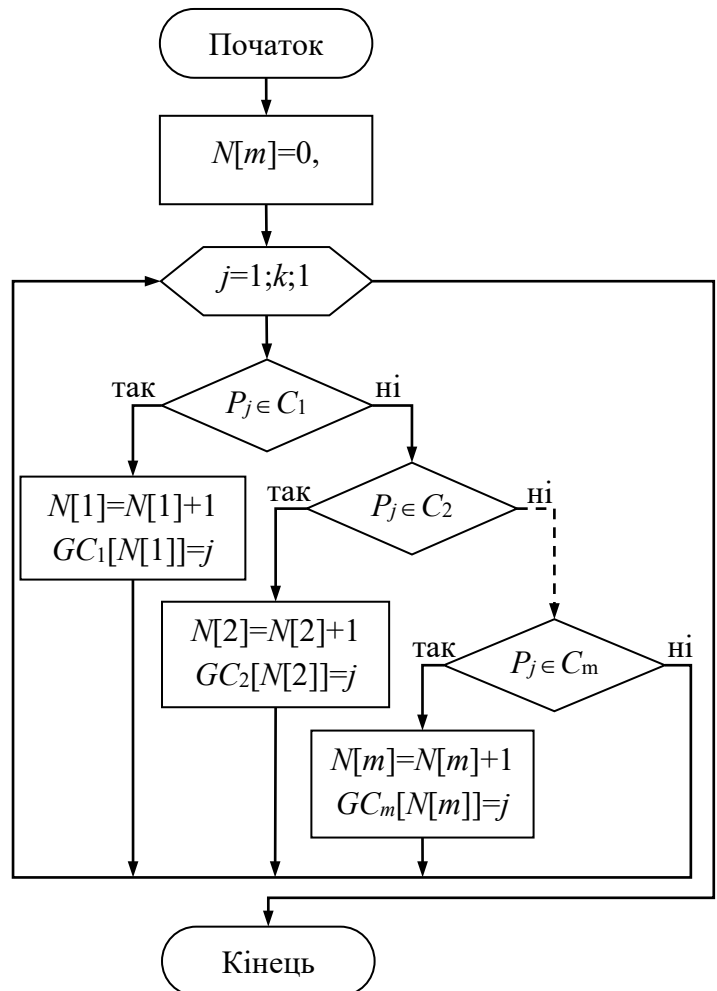


Рисунок 7 – Пошук груп ідентичних класифікацій

$$D \Leftrightarrow \begin{cases} \exists P_i \in \{P\}, P_i \in C_g, \\ N[g] < k, \\ \forall i \neq g, N[i] \neq 0. \end{cases} \quad (14)$$

Відповідно, найбільш достовірне рішення може бути знайдене за максимальним значенням характеристики $N(C_m)$ (в методі – значення відповідних елементів масиву N). Якщо враховувати той факт, що кожній характеристиці $N(C_m)$ однозначно співставляється відповідна група образів з ідентичною класифікацією, то сортування матриці N призводить до упорядкування груп образів з ідентичною класифікацією у відповідності до рівня достовірності розпізнавання. Таким чином отримується інформаційна модель прийняття рішення

$$N^S = \begin{bmatrix} \max N(C_m) \\ \dots \\ \dots \\ \min N(C_m) \end{bmatrix} \Leftrightarrow \begin{bmatrix} GC_m^l | \max N(C_m) \\ \dots \\ \dots \\ GC_m^m | \min N(C_m) \end{bmatrix} \Leftrightarrow \begin{bmatrix} \{P_i\}^l, i \in GC_m^l \\ \dots \\ \dots \\ \{P_j\}^m, j \in GC_m^m \end{bmatrix}, \quad (15)$$

де $GC_m^l | \max N(C_m)$ – 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); $\{P_i\}^l$ – сукупність образів, індекси i яких входять до групи GC_m^l (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).

Розроблена інформаційна модель дозволяє забезпечити підтримку прийняття остаточного рішення щодо класифікації стану СС, зокрема у випадку отримання багатоальтернативного результату класифікації. Згідно моделі, для остаточного експертного висновку надаються отримані результати класифікації (альтернативи), ранжовані у порядку зменшення їхньої достовірності, а також інформація про сукупності образів, на основі яких ці класифікації були отримані.

З метою зменшення кількості даних, що підлягають обробці, розроблений метод консолідації образів СС та їх селекції на основі характеристик інформативності. Використання елементів теорії грубих множин для розробки цього метода дозволило уникнути неоднозначності розподілу образів за консолідованими групами. Метод дозволяє здійснити однозначний розподіл образів СС на групи за показниками стаціонарної та нестаціонарної інформативності. Припустимо, що у деякому просторі Ω існує сукупність

образів P_1-P_k , що описують стан CC та є категорією універсуму $U: \{P\} \subseteq U$, $U \neq \emptyset$. В якості відношення еквівалентності пропонується використання відстані образів стану CC від репрезентативного образу P^R . Для визначення репрезентативного образу в методі здійснюється пошук максимального значення характеристики інформативності. Репрезентативним вважається той образ, характеристика інформативності якого максимальна за поточних умов зовнішнього середовища. Тобто $P^R \Leftrightarrow \max ICh$. Надалі здійснюється визначення узагальненої елементарної категорії $U/P^R = \{\{G_{інф}\}, \{G_{уінф}\}, \{G_{нінф}\}\}$, яка дозволяє здійснити консолідацію образів CC за ознаками інформативності на три групи: $G_{інф}$ – інформативні образи, $G_{уінф}$ – умовно інформативні та $G_{нінф}$ – неінформативні. Відносно репрезентативного образу P^R , для кожного з k образів визначається відстань $L_k = \|P_k - P^R\| = \sqrt{(P_k - P^R)' (P_k - P^R)}$. Тоді умови входження до кожної групи визначаються наступним чином:

$$\begin{aligned} \forall P_k \in G_{інф} &\Leftrightarrow L_k < (L^* - \Delta L^*), \\ \forall P_k \in G_{уінф} &\Leftrightarrow (L^* - \Delta L^*) \leq L_k \leq (L^* + \Delta L^*), \\ \forall P_k \in G'_{нінф} &\Leftrightarrow L_k > (L^* + \Delta L^*). \end{aligned} \quad (16)$$

де L^* і ΔL^* – апріорно задані максимально припустима відстань образів стану CC від репрезентативного образу P^R та припустимий допуск відстані.

В результаті такої консолідації, образи, які увійшли до групи $G_{нінф}$ усуваються від аналізу, як неінформативні. Це дозволяє зменшити кількість співставлень при обробці даних, що дозволяє вирішити поставлену мету роботи в аспекті зменшення часової складності аналізу і класифікації стану CC . образи групи $G_{уінф}$ знаходяться в зоні граничної інформативності, тому можуть використовуватися для підтвердження класифікації, яка здійснюється за образами групи $G_{інф}$.

Слід зазначити, що замість визначення L^* може бути використаний алгоритм пошуку груп ідентичних класифікацій. Тоді, значення характеристики $N(C_m)$ визначатимуть рівень інформативності. В цьому випадку, (16) приймає вигляд:

$$\begin{aligned} \forall GC_m \in G_{інф} &\Leftrightarrow N(C_m) < (N^*(C_m) - \Delta N^*(C_m)), \\ \forall GC_m \in G_{уінф} &\Leftrightarrow (N^*(C_m) - \Delta N^*(C_m)) \leq N(C_m) \leq (N^*(C_m) + \Delta N^*(C_m)), \\ \forall GC_m \in G'_{нінф} &\Leftrightarrow N(C_m) > (N^*(C_m) + \Delta N^*(C_m)). \end{aligned} \quad (17)$$

Окрім селекції образів за ознакою інформативності, для зменшення кількості співставлень даних підчас класифікації, розроблений метод екстенсіонально-інтенсіонального аналізу ознак, який включає етап аналізу і класифікації стану CC по узагальненим описам і етап деталізованого аналізу, який, за необхідності, може використовуватися для уточнення отриманої класифікації. Метод проілюстрований на рис. 8. Згідно цього рисунку, CC представлена сукупністю k образів, кожен з яких дозволяє здійснити розподіл до одного з наперед визначених класів C_1-C_m . При цьому, у випадку достовірної

класифікації, кожен з k образів повинен забезпечувати віднесення об'єкту розпізнавання до одного й того ж самого класу з множини $\{C\}$.

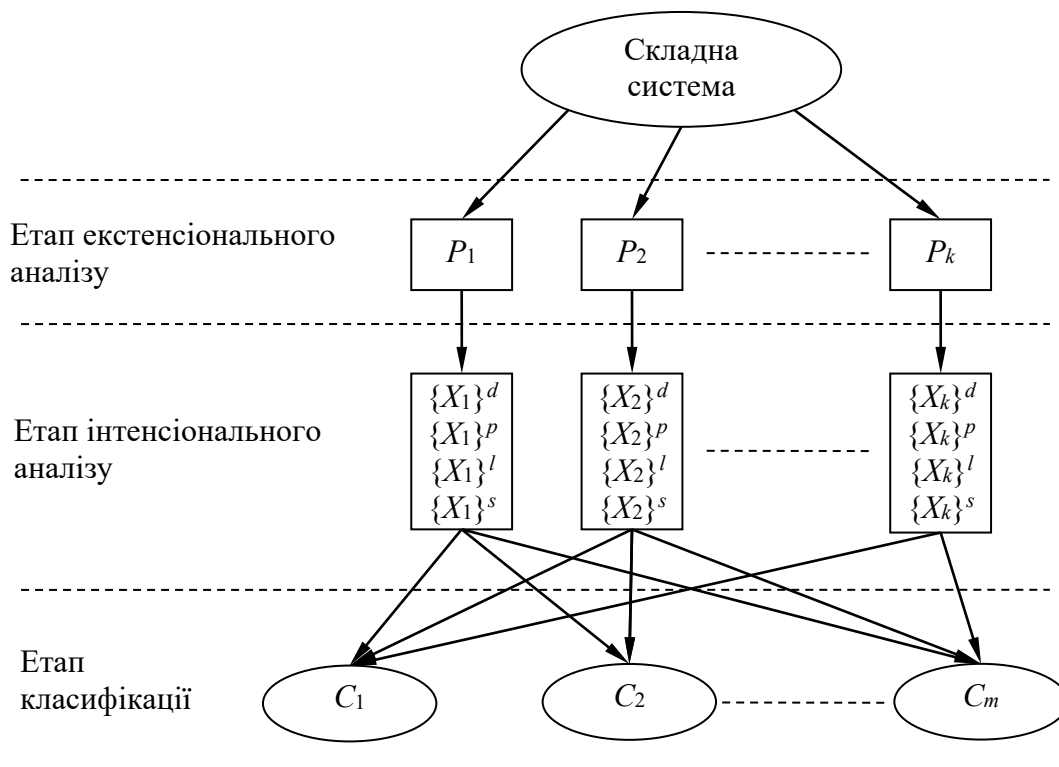


Рисунок 8 – Метод екстенціонально-інтенціонального аналізу ознак при гібридному розпізнаванні образів

Для зменшення кількості співставлень даних, в екстенціонально-інтенціональному методі аналізу ознак, спочатку, здійснюється етап екстенціонального аналізу, при якому виконується спроба класифікації стану СС за узагальненими описами (найбільш характерні ознаки стану СС). У зв'язку з цим, в інформаційній технології гібридного розпізнавання повинен бути реалізований механізм узагальнення описів об'єктів, що відносяться до одного класу, тобто механізму формування компактних узагальнених образів. Надалі, класифікація визначається на основі статистики, що отримується шляхом голосування: належність стану СС до класу C_i вважається достовірною, якщо ця класифікація підтверджується більшістю образів з множини k . Якщо екстенціональний аналіз не дозволяє отримати достовірну класифікацію, на наступному етапі здійснюється інтенціональний аналіз. При цьому слід зазначити, що усунення неінформативних за поточних умов образів зменшує кількість даних, наданих для аналізу, як на екстенціональному так й інтенціональному етапі. Це обумовлено тим, що виключення з аналізу найменш інформативних образів, призводить до автоматичного виключення і відповідних ознак.

Слід зазначити, що використання запропонованого екстенціонально-інтенціонального методу не вимагає проведення повного аналізу інформативних ознак за кожним з k образів стану СС. Такий аналіз здійснюється лише до моменту, коли за одним або декількома образами з множини $\{P\}$ буде отримана класифікація, рівень достовірності якої досягає апріорно заданого значення.

Четвертий розділ присвячений розробці інформаційної технології (ІТ) гібридного розпізнавання образів для консолідованої обробки даних в СС з різнорідними проявами перешкод і викривлень.

В розділі визначені граничні умови та вимоги до використання ІТ гібридного розпізнавання, які пов'язані з критичним застосуванням обробки даних в умовах реального часу. Вказано, що для оперативної обробки даних, ІТ гібридного розпізнавання образів адаптована для використання на базі як систем «жорсткого», так і «м'якого» реального часу.

В системах «жорсткого» реального часу, процес обробки не повинен перевищувати строго визначеного часу, тобто $t_{обр} \leq t_{макс}$, де $t_{макс}$ - максимальне порогове значення часу, що відводиться на обробку даних. З урахуванням цієї граничної умови, процес класифікації здійснюється або до моменту отримання заданого рівня достовірності ($t_{обр} < t_{макс}$), або до закінчення часу, відведеного на аналіз. При закінченні часу, відведеного для аналізу ($t_{обр} = t_{макс}$), результат розпізнавання представляється з досягнутим на цей час рівнем достовірності (очевидно, що меншим за цільове значення).

В системах «м'якого» реального часу дозволяється затримка часу обробки на певне визначене значення, тобто $t_{обр} \leq t_{макс} + \Delta t$, де Δt - визначений проміжок часу, на який процес обробки може перевищувати заданий максимально припустимий час. Тоді, в цьому випадку, процес обробки продовжується до моменту отримання завданого рівня достовірності, тобто $t_{обр} < t_{макс}$, або, якщо завданий рівень достовірності не досягнутий, то час обробки може бути продовжений до моменту $t_{макс} + \Delta t$. Якщо по досягненню граничного часу обробки цільове значення достовірності не було досягнуто, то результат розпізнавання надається з отриманим показником достовірності, який менше за попередньо заданий.

Таким чином, до ІТ гібридного розпізнавання в реальному часі представляються наступні вимоги: 1) передбачуваність поведінки в часовій області; 2) масштабованість (можливість отримувати надкомпактні, варіанти обробки мінімальної кількості даних для отримання бажаної достовірності результату); 3) реалізація алгоритмів обробки в режимі корпоративної багатозадачності. У зв'язку з цим, розроблені моделі і методи гібридного розпізнавання спрямовані на виділення мінімальної кількості ознак, які за поточних для СС зовнішніх умов мають максимальну інформативність та дозволяють отримати завданий рівень достовірності. Окрім цього, консолідація ознак при гібридному розпізнаванні спрямована на отримання черги даних для обробки, яка ранжується у порядку зменшення їхньої інформативності. Тобто в умовах обмеженого часу, пріоритет в співставленні мають найінформативніші ознаки образів $\{P\}$ стану СС. Наявність такої ранжованої черги дозволяє використати кооперативну багатозадачність, яка дозволяє найбільш ефективно займати системні ресурси.

Крім цього, слід зазначити, що розроблений метод роздільного аналізу ознак має переваги перед спільним завдяки спрямованості на використання багатозадачності і, насамперед, кооперативної. На рис. 9 наведена схема

паралельної обробки та формування черги при гібридному розпізнаванні образів.

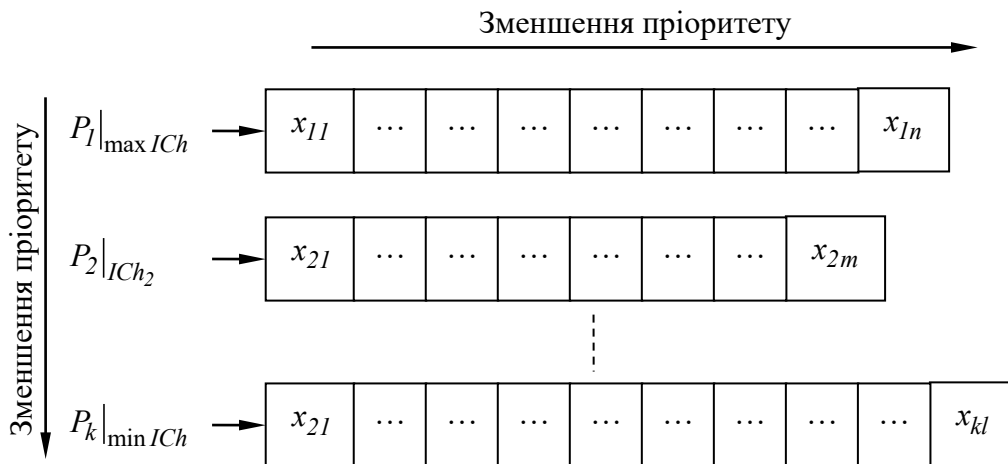


Рисунок 9 – Формування черги і послідовності обробки даних при гібридному розпізнаванні

Кожен образ відповідає окремому потоку обробки даних. Послідовність обробки образів відповідає зменшенню їх інформативних характеристик Ich . Послідовність співставлення ознак $\{x\}$ кожного образу здійснюється відповідно до розроблених метода та інформаційної моделі визначення послідовності даних на основі динаміки їх зміни (тобто ступеня їхньої зміни у порівнянні з попереднім циклом класифікації стану CC). Також, слід зазначити, що відповідно до розробленого екстенціонально-інтенціонального методу, на початку черги обробки за кожним образом розміщуються узагальнені характеристики (фрейми), а надалі – деталізовані ознаки (слоти). Згідно представленої черги, на кожному етапі порівняння здійснюється розрахунок достовірності класифікації до кожного класу. Процес обробки може бути завершений на будь якому етапі співставлення даних за умови отримання бажаного рівня достовірності класифікації або спливання максимального часу, виділеного на обробку.

Також, в розділі сформульовані вимоги до апаратно-програмних засобів реалізації інформаційної технології гібридного розпізнавання образів, які полягають у надійності та роботі в широкому діапазоні зовнішніх умов, наявності розвиненої системної служби часу, легкої конфігурованості, простоти системи команд і способів адресації, розвиненої системи інтерфейсів та можливості підключення великої кількості зовнішніх пристроїв реєстрації характеристик, наявності засобів попередньої обробки і нормування характеристик CC та розвиненої системи переривань.

Етапи реалізації інформаційної технології гібридного розпізнавання образів для консолідованої обробки неоднорідних даних в складних системах представлені далі.

Етап 0. Априорно визначаються стаціонарні характеристики інформативності $SIch$ кожного інформаційного потоку в CC , за якими отримуються образи різної природи виникнення. Враховується точність технічних засобів реєстрації характеристик. $\{SICh\}$ залишається постійною протягом усіх циклів гібридного розпізнавання. Якщо змінюється структура

системи, або замінюються пристрої реєстрації – $\{SICH\}$ підлягає корекції. Також на цьому етапі визначається алфавіт класів, що характеризують можливі стани СС.

Етап 1. Реєстрація характеристик стану СС та поточних зовнішніх умов.

Етап 2. Нормування отриманих характеристик та формування сукупності образів стану СС та зовнішніх умов. (Відповідно до *SEMHPR* визначення P і EC).

Етап 3. На основі методу визначення інформативності розраховуються нестационарні характеристики інформативності $NSICH$ для кожного образу P , а також, на основі результатів етапу 0, визначаються загальні характеристики ICh .

Етап 4. Згідно методу консолідації неоднорідних даних здійснюється селекція інформативних образів СС та усуваються від обробки ті образи, що за поточних зовнішніх умов EC не задовольняють заданим значенням ICh .

Етап 5. Відібрані для аналізу образи ранжуються відповідно до їх значень інформативності. Таким чином, формується багатопоточна черга образів для подальшої обробки (на рис. 9 – вертикальна стрілка).

Етап 6. Використовується метод визначення послідовності обробки даних на основі динаміки змін значень ознак СС. В результаті, пріоритет в обробці для кожного образу надається ознакам, які мають максимальну динаміку зміни, порівняно з попереднім етапом гібридного аналізу і класифікації. Таким чином, формуються ранжована черга обробки ознак кожного образу (на рис. 9 – горизонтальна стрілка).

Етап 7. Здійснюється екстенціональний аналіз, в якому методом гібридного розпізнавання, оснований на роздільного аналізу та пошуку груп ідентичних класифікацій, здійснюється співставлення ознак СС. За кожним образом, обчислюється ступінь подібності до кожного класу C множини S

$S(P,C) = 1 - \frac{L}{L_{max}}$, де L – відстань ознаки від репрезентативного значення

відповідного класу; L_{max} – сума відстаней, отриманих по кожному класу на поточному етапі порівняння ознак. Достовірність класифікації визначається після кожного циклу порівнянь та обраховується з використанням міри подібності Хеммінга:

$D(C) = \frac{N(C)}{N}$, де $N(C)$ – чисельність групи образів з

ідентичною класифікацією до відповідного класу C . На цьому етапі, якщо за одним або групою образів отримується наперед заданий рівень достовірності, то етап 8 пропускається і надається отриманий результат класифікації – тобто відповідність поточного стану СС одному з наявних класів. Якщо час, виділений для аналізу даних спливає (відповідно до умов систем «жорсткого» і «м'якого» реального часу, описаних вище) або для кожного класу не отримане задовільне значення $D(C)$ – здійснюється перехід на етап 8.

Етап 8. Здійснюється інтенціональний аналіз, в якому уточнюється результат класифікації, отриманої на попередньому етапі. Якщо в результаті такого уточнення після співставлення досягається бажане значення $D(C)$, то подальший аналіз інших ознак припиняється і результат класифікації стану СС надається для використання. Якщо за результатами аналізу усіх наявних ознак

не отримується заданий рівень достовірності $D(C)$ або в результаті класифікації для декількох класів отримані значення $D(C)$ що відповідають завданому значенню або час, визначений для аналізу даних та класифікації спливає – здійснюється перехід до етапу 9.

Етап 9. Відповідно до розробленої інформаційної моделі підтримки прийняття рішення формується інформація про групи ідентичних класифікацій з розрахованими останніми значеннями отриманої достовірності. Ця інформація надається в ранжованому вигляді, що дозволяє на основі експертної оцінки визначити остаточний результат класифікації, який, з великою долею ймовірності, буде визначатися першими значеннями виведеної ранжованої послідовності.

Також, в розділі розроблена структурна організація системи обробки даних СС (рис. 10), в якій використовуються розроблені моделі, методи та інформаційна технологія гібридного розпізнавання.

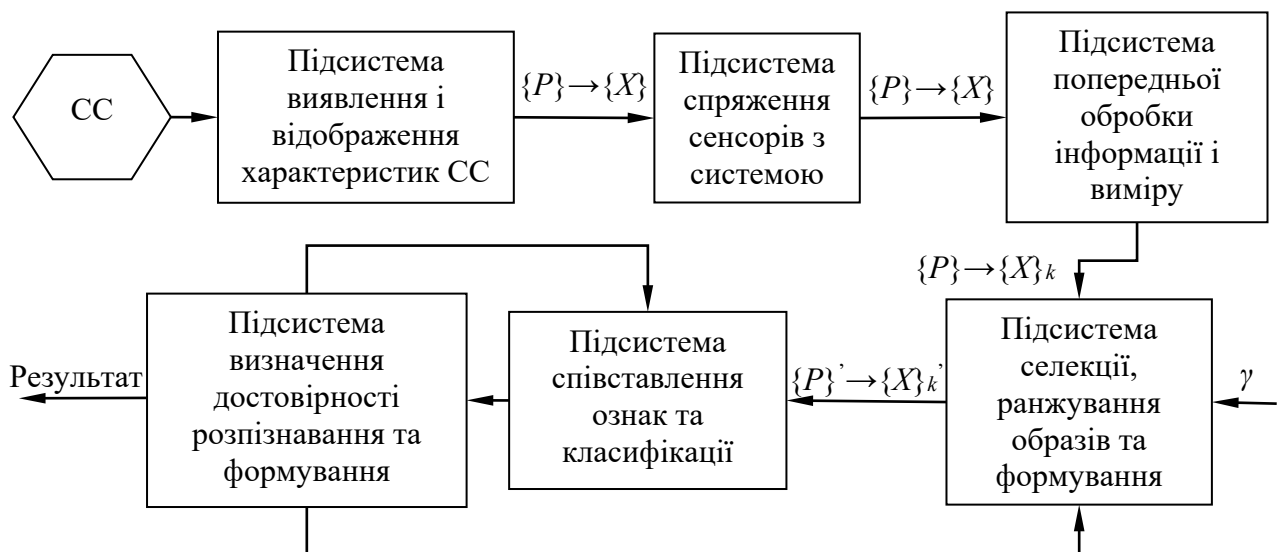


Рисунок 10 – Структурна організація системи обробки даних СС на основі гібридного розпізнавання образів

Підсистема виявлення і відображення характеристик СС представляє собою сукупність сенсорів, різної фізичної та інформаційної природи, які дозволяють отримати відображення об'єкту розпізнавання у просторі образів. Кінцевою задачею цієї підсистеми є формування сукупності образів різної природи виникнення, які характеризуються визначеною сукупністю ознак ($\{P\} \rightarrow \{X\}$). Підсистема спряження сенсорів з системою представляє собою сукупність інтерфейсів передачі даних, що дозволяють забезпечити взаємодію сенсорів з системою. Підсистема попередньої обробки інформації і виміру характеристик СС здійснює корекцію отриманої інформації $\{P\} \rightarrow \{X\}$, визначення викривлень і завад, нормалізацію даних. В результаті, на виході підсистеми формується коректована множина інформаційних ознак відповідних образів $\{P\} \rightarrow \{X\}_k$. Підсистема селекції, ранжування та формування черги виконує задачі визначення сукупності даних, що за поточних зовнішніх умов γ , можуть забезпечити більш ефективну і точну класифікацію. Надалі, ці дані ранжуються в порядку зменшення інформативності і формують чергу обробки. Результатом роботи цієї підсистеми є ранжована підвибірка $\{P\}' \rightarrow \{X\}'_k$ з

сукупності $\{P\} \rightarrow \{X\}_k$. Підсистема співставлення ознак і класифікації є основною підсистемою, в якій, безпосередньо, реалізуються алгоритми розпізнавання. Підсистема визначення достовірності розпізнавання та формування результату класифікації призначена для поточного обрахування достовірності класифікації та аналізу отриманих класифікацій відповідно до існуючих умов і критеріїв. У випадку, коли задані критерії не досягнуті, а граничні умови не порушені – здійснюється наступний цикл співставлення. В іншому випадку забезпечується виведення отриманих результатів для експертної оцінки та визначення остаточного результату класифікації.

П'ятий розділ присвячений прикладним інформаційним технологіям гібридного розпізнавання образів.

Розроблені моделі, методи та інформаційна технологія були використані при побудові трьох різноманітних систем переробки інформації: в інформаційній системі контролю просторового розподілу температури коксової печі, в інформаційній системі ультразвукового вимірювання лінійних відстаней для засобів автоматизації, а також в інформаційній системі визначення ступеня унікальності текстових даних.

Для коксохімічного виробництва, розроблена інформаційна технологія дозволила створити систему просторового розподілу температури коксових печей, яка забезпечила контроль температури по всій поверхні коксового пирога без використання великої кількості пірометрів. В такому прикладному застосуванні, образами, що надаються до розпізнавання, є температурні характеристики коксового пирога, що отримуються на етапі його виштовхування з коксової печі. Така система гібридного розпізнавання поєднує два образи різної природи виникнення: образ зображення поверхні коксового пирога за допомогою системи технічного зору та образ лінійного розподілу температури, що отримується з пірометра, який встановлюється уздовж горизонтальної осі коксового пирога. Схема співставлення образів P_1' і P_2 представлена на рис. 11.

Множина інформаційних ознак образа P_1 представляє собою матрицю растрового зображення, що отримується системою технічного зору, однак для зменшення деталізації, здійснюється сегментація зображення, що знижує його роздільну здатність. Таким чином, на основі образа P_1 формується менш деталізований образ P_1' . P_2 представляє собою дискретизований у часі лінійний

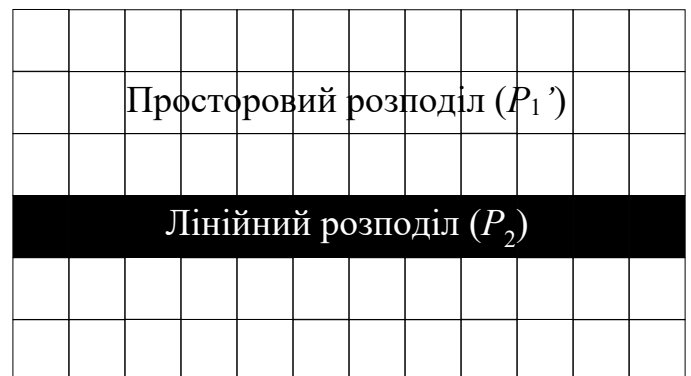


Рисунок 11 – Співставлення образів P_1' і P_2

розподіл температури, що отримується за допомогою пірометричного датчика в процесі виштовхування готового коксового пирога. При цьому, система технічного зору надає інформацію про просторовий розподіл температури по всій поверхні коксового пирога, а пірометр – більш точну інформацію про температуру. З метою корекції даних про просторовий розподіл температури

коксового пирога та підвищення його точності, підчас роздільного аналізу ознак необхідно провести співставлення образів P_1 і P_2 , так як вони обидва, в тій чи іншій мірі, характеризують один об'єкт розпізнавання – просторовий розподіл температури коксового пирога.

В цьому випадку, модель СС буде описуватися системою двох образів P_1 і P_2 вигляду

$$\left\{ \begin{array}{l} P_1 = \begin{pmatrix} X_{1,1} & X_{1,2} & \dots & X_{1,n} \\ X_{2,1} & X_{2,2} & \dots & X_{2,n} \\ \dots & \dots & \dots & \dots \\ X_{m,1} & X_{m,2} & \dots & X_{m,n} \end{pmatrix} \\ P_2 = (X(t)_1, X(t)_2, \dots, X(t)_k), \end{array} \right. \quad (18)$$

де $\{X_{m,n}\}$ – ознаки образу P_1 , що характеризують просторовий розподіл температури по поверхні коксового пирога; $\{X(t)_k\}$ – ознаки образу P_2 , що характеризують лінійний розподіл температури на одній з осей коксового пирога. Множина ознак образу P_2 , що підлягають співставленню, є репрезентативною $\{X_{m,n}^R\}$ (так як точність визначення температури пірометром більша). Якщо температурну характеристику визначеної області коксового пирога, яка характеризується ознакою X , позначити як $T(X)$, то умова корекції атрибутів зображення, що представлені образом P_1 , буде мати вигляд $\forall X_{m,n}^R \in P_1, T(X_{m,n}^R) = T(X(t)_k)$. У випадку, якщо для окремих областей коксового пирога умова не виконується, то температурна характеристика, що відповідає поточному атрибуту зображення, присвоюється рівною до температурної характеристики відповідної ознаки образу P_2 . Після співставлення репрезентативних ознак образу P_2 з ознаками образу P_1 і корекції температурних характеристик, необхідно провести корекцію інших характеристик просторового розподілу температури, що характеризуються ознаками образу P_1 для інших областей зображення коксового пирога. Такий підхід фактично виключає необхідність попереднього калібрування системи технічного зору з метою отримання достовірних значень розподілу температури.

Перевірка ефективності такого рішення і достовірності отримання розподілу температур здійснювалася з використанням еталонного пірометричного датчика, що сертифікований для використання у цьому технологічному процесі. Аналіз розбіжності температурних характеристик здійснювався на лінії встановленого еталонного датчика та відповідної осі просторового розподілу температури коксової печі. Виміри здійснювалися для п'яти коксових печей. Середня достовірність за результатами усіх вимірів склала 91.4%, що задовольняє існуючим технологічним нормам.

Також, використання гібридного розпізнавання дозволила не тільки визначати просторовий розподіл температури, а й наявність та динаміку розвитку дефектів коксової печі. Для візуалізації цієї інформації був створений програмний модуль аналізу термограм, який за наявності характерних зон

постійного перегріву коксового пирога на різних технологічних циклах надає можливість визначати наявні дефекти, їх конфігурацію і розміри. Вікно модулю, з бінаризованою обробкою фрагменту термограми, наведене на рис. 12. Розроблена система пройшла апробацію в умовах діючого коксохімічного підприємства, про що мається відповідний акт.

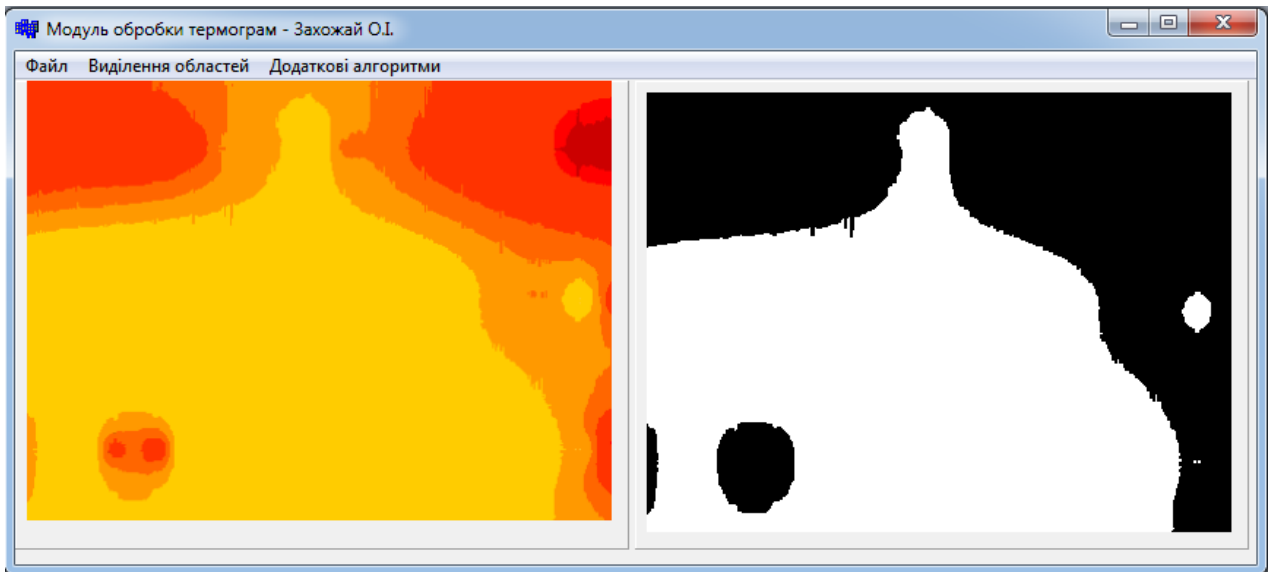


Рисунок 12 – Приклад роботи модулю обробки термограм

В інформаційній системі ультразвукового вимірювання лінійних відстаней гібридне розпізнавання образів використано для підвищення точності і розширення діапазону вимірювань, так як будь-який ультразвуковий датчик має конструктивні обмеження щодо достовірного виміру в широкому діапазоні відстані. Для вирішення цієї проблеми необхідне використання сукупності k датчиків, в яких діапазони вимірювань доповнюють один одного. Принцип суміщення діапазонів вимірювання датчиків проілюстрований на рис. 13.

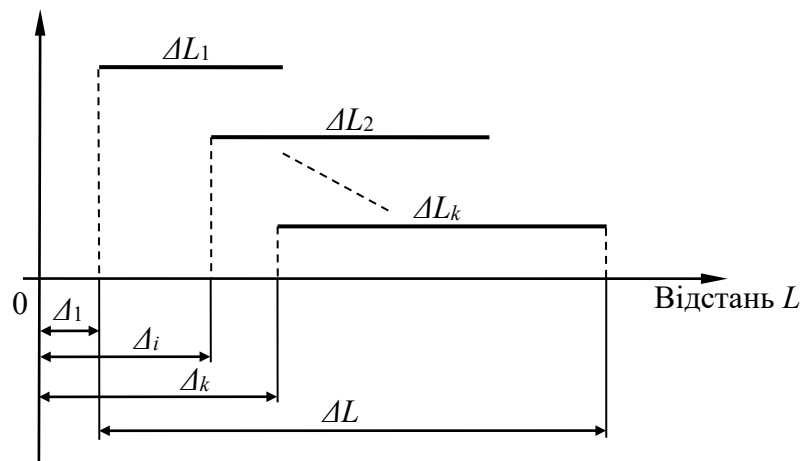


Рисунок 13 – Суміщення діапазонів вимірювання датчиків

Згідно рис. 13, кожен датчик характеризується діапазоном відстані, в якому може бути отриманий достовірний результат вимірювання. Ці діапазони схематично представлені на рисунку лініями з позначеннями $\Delta L_1, \Delta L_2, \dots, \Delta L_k$. При цьому, кожному датчику характерна «мертва» зона з характеристиками $\Delta_1, \Delta_2, \dots, \Delta_k$ відповідно. Очевидно, що у випадку сумісної обробки результатів вимірювань, мінімальна відстань реєстрації буде співпадати з мінімальною

характеристикою «мертвої» зони одного з k датчиків (на рис. 13 це Δ_1). Максимальне припустиме значення достовірного вимірювання відстані буде співпадати з максимальним значенням відстані для одного з датчиків (на рис. 13 це ΔL_k). Таким чином, загальний діапазон достовірного вимірювання відстані системою ультразвукових датчиків буде визначатися як ΔL , що значно ширший ніж будь який діапазон вимірювання кожного, окремо узятото датчика.

Однак, використання багатоканальної системи вимірювання створює додаткові складнощі визначення достовірного результату вимірювання. Це пов'язано з двома факторами: нерівномірністю характеристики точності вимірювання у всьому діапазоні для кожного датчику та наявності зон перекриття діапазонів вимірювання окремих датчиків, що створює неоднозначність у визначенні результату. Наявність цих двох аспектів вимагає в процесі обробки даних вирішення додаткової задачі виключення надлишковості даних та визначення найдостовірнішого результату вимірювань. Згідно принципу гібридного розпізнавання формується множина k образів $\{P\}$, яка підлягає селекції та на її основі формується множина інформативних образів $\{P'\}$. Сукупність $\{P'\}$ містить образи, що відповідають тим каналам виміру, що за поточних зовнішніх умов забезпечують максимальну достовірність результату. При чому $\{P'\} \subset \{P\}$. Визначення характеристик інформативності здійснюється на основі експертних знань, які визначаються апіорно: дані про типи ультразвукових датчиків, розміри «сліпої» зони датчика $\{\Delta_k\}$, діапазони вимірювання датчиків $\{\Delta L_k\}$, дані про розподіл точності вимірювання у всьому діапазоні датчика $\{p_k(r)\}$. Визначення критеріальної характеристики G розподілу множини $\{P\}$ на два класи – інформативні та неінформативні здійснюється за методом, розглянутим вище. В результаті чого, інформаційну модель системи гібридного розпізнавання для багатоканального ультразвукового вимірювання відстані сукупністю датчиків $\{S\}$ можна представити у вигляді

$$\begin{pmatrix} S_1 \\ S_2 \\ \dots \\ S_k \end{pmatrix} \Rightarrow \begin{pmatrix} P_1 \\ P_2 \\ \dots \\ P_k \end{pmatrix} = \begin{cases} R_1, \{Ich_1|G\}, \\ R_2, \{Ich_2|G\}, \\ \dots \\ R_k, \{Ich_k|G\}, \end{cases} \quad (19)$$

де R_1-R_k – сукупність апостеріорних даних про вимір відстані кожним з k датчиків; $\{Ich_k|G\}-\{Ich_k|G\}$ – представляє собою сукупність апіорних характеристик k датчиків за умови наявності критеріальної характеристики G , за якої може бути здійснена селекція достовірного (або достовірних) результатів вимірювання. Згідно моделі (19) кожна ознака R_1-R_k характеризує одне й теж саме значення виміру відстані. Якщо представити, що система складається з k ідеальних ультразвукових датчиків, кожен з яких, на всьому діапазоні вимірювання, дозволяє отримати абсолютно достовірний результат, то очевидно, що $R_1 = R_2 = \dots = R_k$. У випадку реальних датчиків, на кожній відстані отримуються результати вимірювання з різною достовірністю,

тому $R_1 \approx R_2 \approx \dots \approx R_k$. Якщо достовірний результат вимірювань позначити як R , а функцію розподілу точності вимірювання для i -го датчика як $p_i(R)$, то критерій селекції цього значення з наявної сукупності буде визначатися як $\forall R_i, R_i \in \Delta L \Leftrightarrow \exists R_i = R, p_i(R_i) = \max p_i(R)$.

Розроблена система ультразвукового вимірювання пройшла апробацію в умовах реального промислового підприємства про що мається відповідний акт.

При розробці системи визначення унікальності текстових даних, брався до уваги той факт, що достовірність отриманого результату багато в чому залежить від форми представлення матеріалу на аналіз, а також форми та вмісту бази даних текстових фрагментів з якими здійснюється порівняння.

Відповідно до теорії розпізнавання образів, об'єктом розпізнавання є текстовий фрагмент, що надається для аналізу, а фрагменти з якими здійснюється порівняння, є апріорно визначеними класами C_1-C_n , ступінь подібності до яких буде визначати відсоток унікальності тексту, що перевіряється. Для врахування можливих трансформацій тексту, що представляється для аналізу, в системі вводяться функції перетворення F_1-F_n , що дозволяють співставити різноманітні форми представлення інформації з наявними у базі класами C_1-C_n . Наприклад, така трансформація має місце у випадку необхідності співставлення тексту з урахуванням можливих варіантів його перекладу іншими мовами.

На рис. 14 наведена схема порівняння текстових даних з наявними в базі фрагментами. Класи C_1-C_n характеризуються репрезентативними образами $\{P^R\}$ причому, за кожним з них може бути здійснене віднесення наданого для аналізу тексту до відповідного класу. Відмінності схеми організації порівняння тексту полягають у тому, що у базі даних, кожний клас

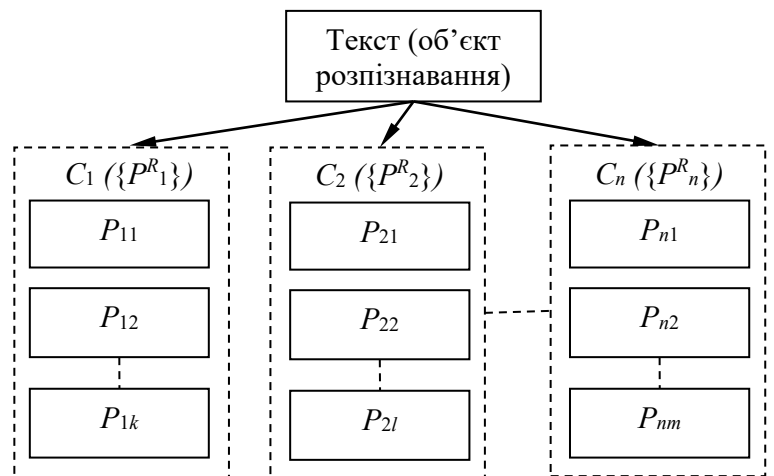


Рисунок 14 – Схема порівняння тексту за фрагментами в базі даних

представляється сукупністю репрезентативних образів, які відповідають можливим, апріорно заданим формам представлення текстових даних. В процесі аналізу, текст, що аналізується, співставляється з кожним репрезентативним образом. Визначення відсотку унікальності тексту здійснюється відповідно до міри відстані, яка визначається як

$D_i = \|P_i - P^R\| = \sqrt{(P_i - P^R)'(P_i - P^R)}$, де D_i – відстань до i -го репрезентативного образу; P_i – образ, що надається для аналізу; P^R – репрезентативний образ. Відсоток унікальності текстового фрагмента відзначається пропорційним значенню відстані, тобто $D_i \rightarrow 0 \Rightarrow y_i \rightarrow 0$, де y_i – відсоток унікальності текстового фрагменту.

Перевірка запропонованого технічного рішення щодо підвищення достовірності визначення відсотку унікальності текстової інформації проводилося на прикладі двох систем. В результаті достовірність перевірки збільшилась у середньому на 54%, що вказує на більш коректне порівняння та ефект від усунення попередньої трансформації текстових фрагментів та їхнього аналізу не на мові оригіналу. Запропоноване рішення було використане для побудови системи перевірки на плагіат вищого навчального закладу, про що мається відповідний акт.

ВИСНОВКИ

В дисертаційній роботі вирішена важлива науково-технічна проблема забезпечення завданого рівня достовірності обробки неоднорідних даних в складних системах при одночасному зменшенні кількості співставлень даних, що надаються для обробки.

Отримані в дисертаційній роботі результати повністю підтвердили гіпотезу досліджень, яка була зазначена в загальній характеристиці роботи.

В результаті дисертаційного дослідження були отримані наступні основні результати та висновки.

1) Аналіз існуючих моделей і методів обробки неоднорідних даних в СС з різномірним проявом перешкод і викривлень показав направленість на обробку усього набору характеристик та намагання прийняти рішення за усією сукупністю ознак, при цьому взагалі не враховувалося, що інформативність кожної ознаки СС може значно змінюватися у випадку змін умов навколишнього середовища. При цьому наявність великої кількості даних, що підлягають обробці, призводило до значних втрат часу, що в багатьох застосуваннях носить критичний характер. Для вирішення вказаних недоліків в роботі запропонована нова концепція обробки неоднорідних даних в СС через гібридне розпізнавання образів.

2) В процесі вирішення вказаної основної проблеми, отримала подальший розвиток теорія комбінованих систем розпізнавання образів, яка, за рахунок опису об'єкту розпізнавання ознаками різної природи виникнення, дозволяє здійснювати достовірну класифікацію за умови широкої зміни різномірних перешкод і викривлень.

3) Аналіз висунутої гіпотези дослідження дозволив вдосконалити матричну інформаційну модель СС, в якій неоднорідні дані компонуються в різних шарах, що забезпечує більшу гнучкість у виборі альтернатив консолідації неоднорідних даних та методів їх обробки.

4) На основі вдосконаленої інформаційної моделі складної системи сформульовано принцип консолідації неоднорідних даних, який базується на формуванні сукупності образів об'єкту розпізнавання, ознаки яких мають різну природу виникнення, та дозволяє отримувати достовірний результат аналізу за умови різного прояву перешкод і викривлень.

5) Новий принцип консолідації неоднорідних даних дозволив розробити ситуаційно-подійну модель гібридного розпізнавання образів, яка заснована на поданні характеристик СС у вигляді сукупності образів, множини

характеристик зовнішніх умов – як прояв ситуації, статичної складової ситуації – у вигляді множини стаціонарних характеристик інформативності, динамічної складової – у вигляді нестаціонарних характеристик інформативності, та сукупності класів, як результату виконання послідовності дій по розпізнаванню.

6) Для визначення параметрів ситуаційно-подійної моделі розроблено метод визначення інформативності образів СС, який базується на використанні стаціонарної і нестаціонарної складових характеристики інформативності та дозволяє, за поточних умов реєстрації ознак складної системи, визначати найбільш достовірні дані для обробки.

7) На основі стаціонарних та нестаціонарних характеристик інформативності та теорії грубих множин розроблено метод консолідації неоднорідних даних при гібридному розпізнаванні, який дозволяє здійснити селекцію образів СС з метою зменшення кількості співставлень даних для отримання рішення.

8) З метою визначення черги обробки ознак СС розроблено метод та інформаційну модель визначення послідовності обробки даних, які базуються на узагальненні моделі емоційних процесів людини, представленої Аткинсоном-Шифріном, для складних систем з неоднорідними даними та різнорідним проявом перешкод і викривлень, що дозволяє в динаміці системи зменшити кількість даних, що підлягають співставленню та отриманню достовірного результату.

9) Розроблено метод гібридного розпізнавання образів, який оснований на роздільному аналізі інформативних ознак та пошуку груп образів з ідентичною класифікацією, що дозволяє забезпечити апріорно завданий рівень достовірності результату обробки неоднорідних даних.

10) З метою мінімізації кількості співставлень даних при аналізі та класифікації розроблено метод екстенсіонально-інтенсіонального аналізу ознак, який оснований на прийнятті швидкого рішення по узагальненим характеристикам складної системи та використання, в разі необхідності, деталізованого аналізу даних для уточнення результату, що дозволяє зменшити кількість співставлень при отриманні апріорно визначеного рівня достовірності класифікації.

11) З метою зручного представлення результату класифікації для прийняття експертного рішення розроблена інформаційна модель підтримки прийняття рішення, яка у випадку отримання багатоальтернативного результату класифікації стану СС, за рахунок використання інформації про групи образів з ідентичною класифікацією, дозволяє представити для експертного висновку отримані результати класифікації (альтернативи), ранжовані у порядку зменшення їхньої достовірності, а також інформацію про сукупності образів, на основі яких ці класифікації були отримані.

12) На основі розроблених моделей і методів переробки даних розроблено інформаційну технологію розпізнавання образів, що заснована на принципі гібридного розпізнавання та вирішує задачу консолідованої обробки неоднорідних даних в складних системах з різнорідним проявом перешкод та викривлень.

13) Розроблено прикладну інформаційну технологію аналізу просторового розподілу температури коксового пирога для коксохімічного виробництва, яка базується на концепції гібридного розпізнаванні образів. Використання запропонованих технічних рішень дозволило створити систему, яка з заданою точністю визначає просторовий розподіл температури, а також забезпечую контроль наявності та розвитку дефектів простіноків коксової батареї.

14) Розроблено прикладну інформаційну технологію багатоканального ультразвукового вимірювання, яка дозволяє за рахунок гібридного розпізнавання образів отримувати достовірний результат виміру лінійної відстані в широкому діапазоні.

15) Розроблено прикладну інформаційну технологію перевірки текстових даних на унікальність, яка за рахунок використання гібридного розпізнавання дозволяє підвищити достовірність результату аналізу.

16) Розроблені моделі, методи та інформаційна технологія можуть бути використані для організації обробки неоднорідних даних в СС будь-якого прикладного застосування, у випадку можливості отримання ознак різної природи виникнення. Використання запропонованих технічних рішень дозволяє забезпечити отримання завданого рівня достовірності при одночасному зменшенні кількості співставлень ознак, що значною мірою зменшує часову складність аналізу і класифікації.

17) Отримані в рамках дисертаційного дослідження результати повністю підтверджують висунуту гіпотезу стосовно можливості обробки неоднорідних даних на основі ситуаційного управління через представлення груп даних різної природи виникнення, подальшої їхньої селекції та багатопоточної обробки.

СПИСОК ОПУБЛІКОВАНИХ ПРАЦЬ ЗА ТЕМОЮ ДИСЕРТАЦІЇ

Наукові праці, в яких опубліковані основні наукові результати дисертації

1. Меньяйленко О. С. Основи синтезу класифікаторів технічних систем розпізнавання образів з використанням моделей емоційних процесів людини / О. С. Меньяйленко, О. І. Захожай // Науково-технічний журнал «Науковий вісник Національного гірничого університету». – 2015. – № 1(145). – С. 120-126.; **внесок автора:** запропоноване узагальнення моделі емоційних процесів людини для технічних систем; **база(и):** *SciVerse Scopus, Ulrich's Periodicals Directory (ULRICHSWEB), ВИНИТИ, EBSCO, ПИНЦ (eLIBRARY), Index Copernicus.*

2. Захожай О. І. Інформаційна технологія розпізнавання образів в задачах автоматизованої обробки інформації й управління складними системами / О. І. Захожай // Журнал «Проблеми інформаційних технологій». – 2013. – № 01(013). – С. 61-68.; **база(и):** *Open Academic Journals Index (OAJI), ПИНЦ (eLIBRARY), Research Bible, AcademicKeys.*

3. Захожай О. І. Екстенціонально-інтенціональний підхід до синтезу інформаційних технологій автоматизованої обробки інформації і управління на базі багатопараметричних комбінованих систем розпізнавання образів / О. І. Захожай // Журнал «Проблеми інформаційних технологій». – 2015. –

№ 02(018). – С. 106-111.; **база(u):** *Open Academic Journals Index (OAJI), РИНЦ (eLIBRARY), Research Bible, AcademicKeys.*

4. Меняйленко О. С. Комбіновані системи розпізнавання образів при аналізі просторового розподілу температури коксового пирога / О. С. Меняйленко, О. І. Захожай // Електротехнічні та комп'ютерні системи. – 2013. – № 12(88). – С. 147-154.; **внесок автора:** розроблена прикладна інформаційна технологія для аналізу просторового розподілу температури коксового пирога; **база(u):** *Ulrich's Periodicals Directory (ULRICHSWEB), Electronic Journals Library, Google Scholar, ВИНИТИ.*

5. Захожай О. І. Критерії визначення інформативності та ранжування образів при прийнятті рішень в багатопараметричних комбінованих системах розпізнавання / О. І. Захожай // Електротехнічні та комп'ютерні системи. – 2018. – № 27(103). – С. 196-204.; **база(u):** *Ulrich's Periodicals Directory (ULRICHSWEB), Electronic Journals Library, Google Scholar, ВИНИТИ.*

6. Захожай О. І. Селекція раціональної сукупності образів в комбінованих системах розпізнавання / О. І. Захожай // Електротехнічні та комп'ютерні системи. – 2013. – № 09(85). – С. 186-192.; **база(u):** *Ulrich's Periodicals Directory (ULRICHSWEB), Electronic Journals Library, Google Scholar, ВИНИТИ.*

7. Захожай О. І. Раздельный анализ информационных признаков в многопараметрических комбинированных системах распознавания образов / О. І. Захожай, А. С. Меняйленко, В. А. Лыфарь // Problemele Energeticii Regionale. – Chişinău, Republica Moldova. – 2019. – № 1-1(40). – С. 60-68.; **внесок автора:** запропонований метод роздільного аналізу ознак; **база(u):** *Emerging Sources Citation Index Web of Science of Thomson Reuters, Ulrich's Periodicals Directory (ULRICHSWEB), ProQuest LLC, Directory of Open Access Journal (DOAJ), РИНЦ (eLIBRARY), ВИНИТИ.*

8. Захожай О. І. Минимизация временной сложности в многокритериальных системах обработки информации / О. І. Захожай, В. В. Филимонцев // Журнал «Проблемы информационных технологий». – 2014. – № 01(015). – С. 161-166.; **внесок автора:** зроблений аналіз методів мінімізації часової складності обробки даних в багатокритеріальних складних системах; **база(u):** *Open Academic Journals Index (OAJI), РИНЦ (eLIBRARY), Research Bible, AcademicKeys.*

9. Захожай О. І. Спільний аналіз інформаційних ознак у багатопараметричних комбінованих системах розпізнавання образів / О. І. Захожай // Електротехнічні та комп'ютерні системи. – 2018. – № 29(105) – С. 78-86.; **база(u):** *Ulrich's Periodicals Directory (ULRICHSWEB), Electronic Journals Library, Google Scholar, ВИНИТИ.*

10. Рябенкий В. М. Комбіновані системи розпізнавання образів / В. М. Рябенкий, О. І. Захожай // Журнал «Проблеми інформаційних технологій» – 2011. – № 01(009). – С. 156-160.; **внесок автора:** запропонована концепція консолідованої обробки неоднорідних даних; **база(u):** *Open Academic Journals Index (OAJI), РИНЦ (eLIBRARY), Research Bible, AcademicKeys.*

11. Захожай О. І. Улучшение изображения дактилоскопических отпечатков в системах контроля доступа / О. І. Захожай, Ю. В. Чибисова // Сборник научных трудов ДонГТУ. – 2011. – № 33. – С. 353-358.; **внесок**

автора: проведений аналіз методів виділення перешкод і викривлень образів та шляхів підвищення достовірності розпізнавання.

12. Захожай О. І. Інформаційна модель системи непрямого аналізу стану коксової печі засобами комп'ютерного зору / О. І. Захожай, А. С. Шевцова // Збірник наукових праць ДонДТУ. – 2010. – № 30 – С. 184-188.; **внесок автора:** запропонована інформаційна модель аналізу просторового розподілу температури коксового пирога.

13. Меняйленко О. С. Інформаційна технологія автоматизованої обробки даних в багатоканальних системах ультразвукового вимірювання / О. С. Меняйленко, О. І. Захожай // Наукові вісті Національного університету України «Київський політехнічний інститут». – 2014. – № 6(98). – С. 62-67.; **внесок автора:** розроблена прикладна інформаційна технологія гібридного розпізнавання для ультразвукового вимірювання лінійної відстані; **база(и):** *Ulrich's Periodicals Directory, Directory of Open Access Journal (DOAJ), EBSCO, WorldCat, J-Gate, Google Scholar, Open Academic Journal Index (OAJI), AcademicKeys, ResearchBib, Turkish Education Index, Eurasian Scientific Journal Index, Cosmos Impact Factor, Miar, WCOSJ, I2OR, Scholarsteer, SIS, IJIF, InfoBase Index, Index Copernicus.*

14. Захожай О. І. Основні аспекти структурної організації комбінованих систем розпізнавання образів / О. І. Захожай, Ю. Е. Паеранд // Вісник ХНТУ. – 2012. – №1(44). – С.221-225.; **внесок автора:** представлена структурна організація системи обробки даних на основі гібридного розпізнавання.

15. Захожай О. І. Ефективні способи перетворення графічних зображень при їхньої сегментації та розпізнаванні засобами комп'ютерного зору / О.І. Захожай // Збірник наукових праць ДонДТУ. – 2008. – № 26. – С. 338-342.

16. Захожай О. И. Определение проективной структуры и движения по бинокулярным соответствиям в системах технического зрения / О. И. Захожай // Сборник научных трудов ДонГТУ. – 2007 – № 25. – С. 243-249.

17. Захожай О. И. Высокоскоростная обработка изображений с применением цифровых фильтров на базе FPGA / О. И. Захожай, А. С. Солошенко // Сборник научных трудов ДонГТУ. – 2013. – №39. – С. 250-254.; **внесок автора:** представлен анализ путей снижения часовой сложности обработки данных та представленный метод багатопоточної обробки даних.

18. Меняйленко О. С. Підвищення достовірності перевірки унікальності текстів шляхом використання комбінованих систем розпізнавання образів / О. С. Меняйленко, О. І. Захожай, П. І. Бідюк // Журнал «Системні дослідження та інформаційні технології» – 2017. – № 4. – С. 29-37. **внесок автора:** розроблена прикладна інформаційна технологія аналізу текстових даних на унікальність; **база(и):** *Directory of Open Access Journal (DOAJ), EBSCO, Index Copernicus, РИНЦ (eLIBRARY), ВИНІТИ.*

19. Захожай О. І. Прийняття рішень на основі пошуку груп ідентичних класифікацій в багатопараметричних комбінованих системах розпізнавання образів / О. І. Захожай, В. О. Лифар, О. І. Батурін // Вісник Східноукраїнського національного університету імені Володимира Даля. – 2019. – №2 (250). – С. 36-42.; **внесок автора:** розроблений метод та інформаційна модель прийняття

рішення при розпізнаванні у випадку отримання багатоальтернативного результату; **база(u):** *Index Copernicus*.

20. Захожай О. І. Підвищення достовірності розпізнавання об'єктів за умови наявності викривлень їхнього відображення у просторі образів / О. І. Захожай // Вісник ХНТУ. – 2013. – №1 (46). – С.128-131.

21. Захожай О. І. Концепція вдосконалення теоретико-методологічних основ синтезу інформаційних технологій автоматизованої обробки інформації і управління складними системами / О. І. Захожай // Вісник Східноукраїнського національного університету імені Володимира Даля. – 2018. – № 6(247). – С. 49-55.; **база(u):** *Index Copernicus*.

22. Захожай О. И. Использование модели температурного режима в автоматизированной системе управления коксовой печью / О. И. Захожай, А. С. Шевцова // Сборник научных трудов ДонГТУ. – 2011. – № 33. – С. 365-370.; **внесок автора:** розроблений метод аналізу просторового розподілу температури коксового пирога на основі гібридного розпізнавання образів.

23. Захожай О. І. Консолідація даних та прийняття рішень на основі ранжування груп ідентичних класифікацій в гібридних системах розпізнавання / О. І. Захожай, В. О. Лифар, В. Г. Іванов // Електротехнічні та комп'ютерні системи. – 2019. – № 31(107) – С. 95–103.; **внесок автора:** розроблений метод консолідації даних при гібридному розпізнаванні, оснований на ранжуванні груп образів з ідентичною класифікацією; **база(u):** *Ulrich's Periodicals Directory (ULRICHSWEB), Electronic Journals Library, Google Scholar, ВИНТИ*.

24. Захожай О. І. Визначення черги обробки даних при гібридному розпізнаванні образів / О. І. Захожай // Вісник Східноукраїнського національного університету імені Володимира Даля. – 2019. – № 7(255). – С. 111-116.; **база(u):** *Index Copernicus*.

25. Захожай О. І. Інформаційна технологія гібридного розпізнавання образів для обробки неоднорідних даних в складних системах / О. І. Захожай // Вісник Східноукраїнського національного університету імені Володимира Даля. – 2019. – № 8(256). – С. 141-147.; **база(u):** *Index Copernicus*.

26. Zakhzhay O. I. Situational-event model of the hybrid patterns recognition for heterogeneous data processing in complex systems / O. I. Zakhzhay // Mathematical modeling in economy. – 2019. – № 4. – P. 16-25.

27. Меньяйленко О. Організація перевірки студентських і наукових робіт на плагіат / О. Меньяйленко, О. Захожай // Науковий вісник Мелітопольського державного педагогічного університету. – 2017. – № 1(18). – С. 157-162. **внесок автора:** розроблена прикладна інформаційна технологія гібридного розпізнавання образів для аналізу ступеня унікальності текстових даних.

Наукові праці, які засвідчують апробацію матеріалів дисертації

28. Захожай О. І. Використання системи технічного зору для управління гільйотинними ножицями металургійного виробництва / О. І. Захожай // Праці 8-ї міжнародної науково-практичної конференції «Сучасні інформаційні та електронні технології – 2007». – Одеса: ОНПУ, 2007. – С. 81.

29. Захожай О. І. Інформаційна технологія розпізнавання образів при аналізі температурного режиму коксових печей / О. І. Захожай // Праці 10-ї міжнародної науково-практичної конференції “Сучасні інформаційні та електронні технології – 2009». – Одеса: ОНПУ, 2009. – С. 77.

30. Захожай О. І. Аналіз стану коксової печі засобами комп'ютерного зору / О. І. Захожай, А. С. Шевцова // Праці 11-ї міжнародної науково-практичної конференції “Сучасні інформаційні та електронні технології – 2010» т.1. – Одеса: ОНПУ, 2010. – С. 34.; **внесок автора:** запропонований підхід до аналізу просторового розподілу температури коксового пирога на основі апарату розпізнавання образів.

31. Захожай О. И. Использование модели температурного режима коксования в автоматизированной системе управления коксовой печью / О. И. Захожай, А. С. Шевцова // Труды 12-й международной научно-практической конференции «Современные информационные и электронные технологии – 2011». – Одесса: ОНПУ, 2011. – С. 127.; **внесок автора:** представлена прикладна інформаційна технологія аналізу просторового розподілу температури коксового пирога.

32. Захожай О. І. Комбіновані системи розпізнавання образів / О. І. Захожай // Праці 13-ї міжнародної науково-практичної конференції “Сучасні інформаційні та електронні технології – 2012». – Одеса: ОНПУ, 2012. – С. 113.

33. Zakhozhay O. Combined Systems of Patterns Recognition / O. Zakhozhay, Yu. Paerand // Контроль і управління у складних системах (КУСС-2012). XI Міжнародна конференція. Тези доповідей. – Вінниця: ВНТУ, 2012. – С. 241.; **внесок автора:** запропонований підхід до обробки даних в складних системах на основі сукупності образів різної природи виникнення.

34. Захожай О. І. Критерій селекції раціональної сукупності образів в комбінованих системах розпізнавання / О. І. Захожай // Системний аналіз. Інформатика. Управління. (САГУ-2013) Міжнародна науково-практична конференція. – Запоріжжя: КПУ, 2013. – С. 114-116.

35. Zakhozhay O. Human emotion models in automated data processing and control of complex systems / O. Zakhozhay // XII International Conference “Measurement and control in complex systems” (MCCS - 2014) – Vinnytsia: VSTU, 2014. – P. 195.

36. Захожай О. І. Гібридне розпізнавання образів для обробки неоднорідних даних в складних системах / О. І. Захожай // IX Всеукраїнська науково-практична конференція «Електроніка та телекомунікації-2019» – Сєвєродонецьк: СНУ, 2019. – С. 113-115.

Наукові праці, які додатково відображають наукові результати дисертації

37. Меняйленко О. С., Захожай О. І. Спосіб контролю температури нагріву коксової печі. Патент на корисну модель № 91118 МПК (2014.01) С10В 17/00, С10В 23/00; опубл. 25.06.2014, бюл. №9.; **внесок автора:** розроблений спосіб контролю температури коксової печі на основі розпізнавання образів.

38. Бідюк П. І., Захожай О. І., Меньяйленко О. С. Спосіб розпізнавання образів. Патент на корисну модель № 100283 МПК (2015.01) G06K 9/00; опубл. 27.07.2015, бюл. №14.; *внесок автора*: запропонований спосіб гібридного розпізнавання образів на основі роздільного аналізу ознак.

39. Бідюк П. І., Захожай О. І., Меньяйленко О. С. Спосіб розпізнавання образів. Патент на корисну модель № 100078 МПК (2015.01) G06K 9/00; опубл. 10.07.2015, бюл. №13.; *внесок автора*: запропонований спосіб екстенсіонально-інтенсіонального аналізу ознак при гібридному розпізнаванні.

40. Меньяйленко О. С., Захожай О. І. Пристрій розпізнавання образів. Патент на корисну модель № 90109 МПК (2014.01) G06/00; опубл. 12.05.2014, бюл. №9.; *внесок автора*: запропонована структурна організація гібридного розпізнавання образів.

41. Меньяйленко О. С., Захожай О. І. Спосіб розпізнавання образів. Патент на корисну модель № 92493 МПК (2014.01) G06K 9/00; опубл. 26.08.2014, бюл. №16; *внесок автора*: запропонований спосіб гібридного розпізнавання образів на основі аналізу ознак різної природи виникнення.

42. Захожай О. І., Меньяйленко О. С. Пристрій вимірювання відстані за допомогою ультразвуку. Патент на корисну модель № 93963 МПК (2006.01) G01S 7/52, G01S 15/14; опубл. 27.10.2014, бюл. №20. *внесок автора*: запропонована структурна організація ультразвукового вимірювання відстані на основі гібридного розпізнавання образів.

АНОТАЦІЯ

Захожай О. І. Моделі, методи та інформаційна технологія гібридного розпізнавання образів для консолідованої обробки неоднорідних даних в складних системах. – На правах рукопису.

Дисертація на здобуття наукового ступеня доктора технічних наук за спеціальністю 05.13.06 – інформаційні технології. – Луганський національний університет імені Тараса Шевченка, Старобільськ, 2020.

Дисертаційне дослідження присвячене вирішенню важливої науково-технічної проблеми забезпечення завданого рівня достовірності обробки неоднорідних даних в складних системах при одночасному зменшенні кількості співставлень даних, що надаються для обробки і, як наслідок, зменшення часової складності процесу аналізу. Вказана проблема була викликана наступним протиріччям. З одного боку, збільшення кількості інформаційних ознак та використання більш складних алгоритмів дозволяє отримати підвищення достовірності результату аналізу, але, з іншого боку, це призводить до збільшення часової складності процесу аналізу даних та прийняття рішення.

Аналіз і дослідження в предметній області дозволив вдосконалити матричну інформаційну модель складної системи, а також розробити принцип консолідації неоднорідних даних для забезпечення швидкої та достовірної переробки даних в складних системах з різнорідними проявами завад та викривлень.

Розроблені методи визначення інформативності образів, а також їхньої селекції та ранжування для виключення з аналізу тих даних, що за поточних

зовнішніх умов є неінформативними, а також побудови черги даних для багатопоточної обробки в режимі реального часу.

Розроблено методи, модель та інформаційну технологію гібридного розпізнавання образів, які дозволяють забезпечити завданий рівень достовірності класифікації стану складної системи при одночасному зменшенні кількості співставлень даних.

Розроблено інформаційну модель підтримки прийняття рішень для визначення остаточної класифікації у випадку багатоальтернативного результату аналізу.

Розроблені та пройшли апробацію прикладні інформаційні технології аналізу просторового розподілу температури коксового пирога для коксохімічного виробництва, ультразвукового вимірювання відстані для засобів автоматизації та визначення ступеня унікальності текстових даних, в яких використання запропонованих моделей і методів дозволило підвищити достовірність отриманого результату аналізу.

Розроблені моделі, методи та інформаційна технологія можуть бути використані для організації обробки неоднорідних даних в складних системах будь-якого прикладного застосування, у випадку можливості отримання ознак різної природи виникнення.

SUMMARY

Zakhozhay O. I. Models, Methods and Information Technology of Hybrid Patterns Recognition for the Consolidate Data Processing in Complex Systems. – Manuscript.

Thesis for a Doctor of Technical Science degree in specialty 05.13.06 – Information Technologies. – Luhansk Taras Shevchenko National University, Starobilsk, 2020.

The dissertation researching is dedicated to solving an important scientific and technical problem of ensuring a priority heterogeneous data processing reliability level in complex systems. Wherein the number of processed data comparisons are reducing. It allows reducing the time complexity of the analysis process. This problem is related to the following contradiction. On the one hand, an increasing the number of information signs allows to increase the reliability of the analysis result, but on the other hand, it requires algorithms complication and increasing the time complexity of data analysis and the decision-making process.

Analysis and research in the subject area create possibility for complex systems improve the matrix information model and develop the principle of heterogeneous data consolidation for ensure fast and reliable data processing in complex systems with heterogeneous interference and distortion.

Methods have been developed for determining the patterns informativity and their selection and ranking for exclude from analysis the data that is uninformative under current external conditions. Also, it allows the constructing a data queue for multithreaded processing in real time.

Methods, models, and information technology for hybrid pattern recognition have been developed for make it possible to provide a priory level of complex system state classifying reliability and reducing the amount of data being compared.

An information model of decision support was developed to determine the final classification in the case of a multi-alternative analysis result.

Applied information technologies were developed and tested for analyze the spatial distribution of the coke cake temperature in coke-chemical production, ultrasonic distance measurement for automation equipment, and to determine the uniqueness of text data. In this case was use of the proposed models and methods allowed to increase the analysis result reliability.

The developed models, methods and information technology can be used to organize the of heterogeneous data processing in the complex systems for different applications, if the signs of different nature can be obtained.

Підписано до друку 23.04.2020 р.
Гарнітура «Таймс». Друк цифровий.
Умов. друк. арк. 1,9
Замовлення № 3714
Тираж 120 прим.

Надруковано з готових оригінал-макетів
ФОП Свірщевська Маргарита Іванівна
Код ЄДРПОУ 2186109863
63408, м. Северодонецьк, пр-т Хіміків, буд. 22, кв. 8