

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ  
ЛУГАНСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ  
ІМЕНІ ТАРАСА ШЕВЧЕНКА

Кваліфікаційна наукова  
праця на правах рукопису

ЗАХОЖАЙ ОЛЕГ ІГОРОВИЧ

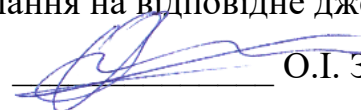
УДК 004.93'1:65.01:65.011.56

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

05.13.06 – інформаційні технології

Подається на здобуття наукового ступеня доктора технічних наук

Дисертація містить результати власних досліджень. Використання ідей, результатів і текстів інших авторів мають посилання на відповідне джерело

 О.І. Захожай

Науковий консультант:

Меняйленко Олександр Сергійович,  
доктор технічних наук, професор,  
заслужений діяч науки і техніки України

Старобільськ – 2020

## АНОТАЦІЯ

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

Дисертація на здобуття наукового ступеня доктора технічних наук за спеціальністю 05.13.06 «Інформаційні технології». – Луганський національний університет імені Тараса Шевченка Міністерства освіти і науки України, Старобільськ, 2020.

Дисертаційне дослідження спрямоване на вирішення науково-технічної проблеми, пов'язаної з неможливістю одночасного збільшення достовірності та оперативності прийняття рішень щодо класифікації в складних системах з неоднорідними даними та різними проявами перешкод та викривлень, так як ці показники є конкуруючими. З одного боку, збільшення кількості інформаційних ознак, використання більш складних алгоритмів, дозволяє отримати підвищення достовірності результату розпізнавання, але, з іншого боку, це призводить до збільшення часової складності процесу аналізу даних та прийняття рішення. 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 автоматики, аналіз і прогнозування газовиділення з підроблених пластів при розробці родовищ корисних копалин, тощо. Відповідно, основним аргументом щодо вибору напрямку досліджень був запит практики: створення точних і швидкодіючих систем обробки неоднорідних даних в складних системах за умови різнорідних перешкод і викривлень.

Основна мета дисертаційного дослідження полягає в забезпеченні 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.

Висунута гіпотеза наукового дослідження полягає в наступному. 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.

Основні результати наукового дослідження, що виносяться на захист полягають в подальшому розвитку теорії комбінованого розпізнавання в напрямку модифікації інформаційної моделі складної системи для обробки неоднорідних даних; обґрунтування принципу консолідації неоднорідних даних; на основі нової концепції гібридного розпізнавання, розробки моделей, методів та інформаційної технології обробки неоднорідних даних в складних системах з різноманітними проявами перешкод і викривлень.

Наукова новизна дисертаційного дослідження полягає в наступному.

*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 Аткінсоном-Шифріном, для складних систем з неоднорідними даними та різномірним 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 апіорно завданий рівень достовірності результату консолідованої обробки неоднорідних даних в складних системах з різномірним 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 апіорно 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.

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.

У вступі наводиться загальна характеристика роботи, 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 складних систем, в результаті якого визначається основна проблема, пов'язана з певним рівнем невизначеності та стохастичності впливу вхідних даних на стан об'єкта розпізнавання і, як наслідок, складність у виборі чи розробці математичної моделі, що забезпечує формалізований опис поведінки на основі апіорної

інформації. Таким чином, для обробки даних і прийняття рішень в складних системах доцільним є використання також апостеріорної інформації як про саму складну систему, так і про зовнішнє для неї середовище.

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 складних системах.

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

Четвертий розділ присвячений розробці інформаційної технології гібридного розпізнавання образів для консолідованої обробки даних в складних системах з різнорідними проявами перешкод і викривлень.

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 автоматки, а також в інформаційній системі визначення ступеня унікальності текстових даних.

У висновках наведено основні результати дисертаційного дослідження.

Додатки містять данні експериментальних досліджень прикладних інформаційних технологій, що були отримані підчас впровадження нових науково-технічних рішень, копії актів впровадження та перелік публікацій.

**Ключові слова:** комбіноване розпізнавання образів, гібридне розпізнавання образів, обробка даних в складних системах, часова складність процесів обробки даних, достовірність класифікації стану складних систем, селекція інформативних ознак, черга обробки даних.

## СПИСОК ОСНОВНИХ ПУБЛІКАЦІЙ ЗДОБУВАЧА

*Наукові праці, в яких опубліковані основні наукові результати дисертації*

1. Меньяйленко О. С., Захожай О. І. Основи синтезу класифікаторів технічних систем розпізнавання образів з використанням моделей емоційних процесів людини // Науково-технічний журнал «Науковий вісник Національного гірничого університету». – 2015. – № 1(145). – С. 120–126. *(Видання входить до міжнародних наукометричних баз: SciVerse Scopus, Ulrich's Periodicals Directory (ULRICHSWEB), ВИНИТИ, EBSCO, РИНЦ (eLIBRARY), Index Copernicus)*

2. Захожай О. І. Інформаційна технологія розпізнавання образів в задачах автоматизованої обробки інформації й управління складними системами // Журнал «Проблеми інформаційних технологій». – 2013. – № 01(013). – С. 61–68. *(Видання входить до міжнародних наукометричних баз: Open Academic Journals Index (OAJI), РИНЦ (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 // Журнал «Проблеми інформаційних технологій». – 2015. – № 02(018). – С. 106–111. *(Видання входить до міжнародних наукометричних баз: Open Academic Journals Index (OAJI), РИНЦ (eLIBRARY), Research Bible, AcademicKeys)*

4. Меньяйленко О. С. Захожай О. І. Комбіновані системи розпізнавання образів при аналізі просторового розподілу температури коксового пирога // Електротехнічні та комп'ютерні системи. – 2013. – № 12(88). – С. 147–154. *(Видання входить до міжнародних наукометричних баз: 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 // Електротехнічні та комп'ютерні системи. – 2018. – № 27(103). – С. 196–204. *(Видання входить до міжнародних наукометричних баз: Ulrich's Periodicals Directory (ULRICHSWEB), Electronic Journals Library, Google Scholar, ВИНИТИ)*

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

7. Захожай О. И., Меняйленко А. С., Лыфарь В. А. Раздельный анализ информационных признаков в многопараметрических комбинированных системах распознавания образов // Problemele Energeticii Regionale. – Chişinău, Republica Moldova, 2019. – № 1-1(40). – С. 60–68. *(Видання входить до міжнародних наукометричних баз: Emerging Sources Citation Index Web of Science of Thomson Reuters, Ulrich's Periodicals Directory (ULRICHSWEB), ProQuest LLC, Directory of Open Access Journal (DOAJ), РИИЦ (eLIBRARY), ВИНИТИ)*

8. Захожай О. И., Филимонцев В. В. Минимизация временной сложности в многокритериальных системах обработки информации // Журнал «Проблемы информационных технологий». – 2014. – № 01(015). – С. 161–166. *(Видання входить до міжнародних наукометричних баз: Open Academic Journals Index (OAJI), РИИЦ (eLIBRARY), Research Bible, AcademicKeys)*

9. Захожай О. І. Спільний аналіз інформаційних ознак у багатопараметричних комбінованих системах розпізнавання образів //

Електротехнічні та комп'ютерні системи. – 2018. – № 29(105) – С. 78–86. (Видання входить до міжнародних наукометричних баз: *Ulrich's Periodicals Directory (ULRICHSWEB), Electronic Journals Library, Google Scholar, ВИНІТИ*)

10. Рябенський В. М., Захожай О. І. Комбіновані системи розпізнавання образів // Журнал «Проблеми інформаційних технологій» – 2011. – № 01(009). – С. 156–160. (Видання входить до міжнародних наукометричних баз: *Open Academic Journals Index (OAJI), РИНЦ (eLIBRARY), Research Bible, AcademicKeys*)

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

12. Захожай О. І., Шевцова А. С. Інформаційна модель системи непрямого аналізу стану коксової печі засобами комп'ютерного зору // Збірник наукових праць ДонДТУ. – 2010. – № 30 – С. 184–188.

13. Меньяйленко О. С., Захожай О. І. Інформаційна технологія автоматизованої обробки даних в багатоканальних системах ультразвукового вимірювання // Наукові вісті Національного університету України «Київський політехнічний інститут». – 2014. – № 6(98). – С. 62–67. (Видання входить до міжнародних наукометричних баз: *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, IIJIF, InfoBase Index, Index Copernicus*)

14. Захожай О. І., Паеранд Ю. Е. Основні аспекти структурної організації комбінованих систем розпізнавання образів // Вісник ХНТУ. – 2012. – №1(44). – С.221–225.

15. Захожай О. І. Ефективні способи перетворення графічних зображень при їхньої сегментації та розпізнаванні засобами комп'ютерного зору // 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 // Сборник научных трудов ДонГТУ. – 2007 – № 25. – С. 243–249.

17. Захожай О. И., Солошенко А. С. Высокоскоростная обработка изображений с применением цифровых фильтров на базе FPGA // Сборник научных трудов ДонГТУ. – 2013. – №39. – С. 250–254.

18. Меньяйленко О. С., Захожай О. І., Бідюк П. І. Підвищення достовірності перевірки унікальності текстів шляхом використання комбінованих систем розпізнавання образів // Журнал «Системні дослідження та інформаційні технології» – 2017. – № 4. – С. 29–37. *(Видання входить до міжнародних наукометричних баз: Directory of Open Access Journal (DOAJ), EBSCO, Index Copernicus, РИИЦ (eLIBRARY), ВИНИТИ)*

19. Захожай О. І., Лифар В. О., Батурін О. І. Прийняття рішень на основі пошуку груп ідентичних класифікацій в багатопараметричних комбінованих системах розпізнавання образів // Вісник Східноукраїнського національного університету імені Володимира Даля. – 2019. – №2 (250). – С. 36–42. *(Видання входить до міжнародної наукометричної бази: Index Copernicus)*

20. Захожай О. І. Підвищення достовірності розпізнавання об'єктів за умови наявності викривлень їхнього відображення у просторі образів // 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 // Вісник Східноукраїнського національного університету імені Володимира Даля. – 2018. – № 6(247). –

С. 49–55. (Видання входить до міжнародної наукометричної бази: *Index Copernicus*)

22. Захожай О. И., Шевцова А. С. Использование модели температурного режима в автоматизированной системе управления коксовой печью // Сборник научных трудов ДонГТУ. – 2011. – № 33. – С. 365–370.

23. Захожай О. І., Лифар В. О., Іванов В. Г. Консолідація даних та прийняття рішень на основі ранжування груп ідентичних класифікацій в гібридних системах розпізнавання // Електротехнічні та комп'ютерні системи. – 2019. – № 31(107) – С. 95–103. (Видання входить до міжнародних наукометричних баз: *Ulrich's Periodicals Directory (ULRICHSWEB), Electronic Journals Library, Google Scholar, ВИНИТИ*)

24. Захожай О. І. Визначення черги обробки даних при гібридному розпізнаванні образів // Вісник Східноукраїнського національного університету імені Володимира Даля. – 2019. – № 7(255). – С. 111–116. (Видання входить до міжнародної наукометричної бази: *Index Copernicus*)

25. Захожай О. І. Інформаційна технологія гібридного розпізнавання образів для обробки неоднорідних даних в складних системах // Вісник Східноукраїнського національного університету імені Володимира Даля. – 2019. – № 8(256). – С. 141–147. (Видання входить до міжнародної наукометричної бази: *Index Copernicus*)

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

27. Меньяйленко О., Захожай О. Організація перевірки студентських і наукових робіт на плагіат // Науковий вісник Мелітопольського державного педагогічного університету. – 2017. – № 1(18). – С. 157–162.

*Наукові праці, які засвідчують апробацію матеріалів дисертації*

28. Захожай О. І. Використання системи технічного зору для управління гільйотинними ножицями металургійного виробництва // 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 // 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.

31. Захожай О. И., Шевцова А. С. Использование модели температурного режима коксования в автоматизированной системе управления коксовой печью // Труды 12-й международной научно-практической конференции «Современные информационные и электронные технологии – 2011». – Одесса: ОНПУ, 2011. – С. 127.

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

33. Zakhozhay O., Paerand Yu. Combined Systems of Patterns Recognition // Контроль і управління у складних системах (КУСС-2012). XI Міжнародна конференція. Тези доповідей. – Вінниця: ВНТУ, 2012. – С. 241.

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

35. Zakhochay O. Human emotion models in automated data processing and control of complex systems // 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 // IX Всеукраїнська науково-практична конференція «Електроніка та телекомунікації-2019» – Сєверодонецьк: СНУ, 2019. – С. 113–115.

*Наукові праці, які додатково відображають наукові результати дисертації*

37. Меньяйленко 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.

38. Бідюк П. І., Захожай О. І., Меньяйленко О. С. Спосіб розпізнавання образів. Патент на корисну модель № 100283 МПК (2015.01) G06К 9/00; опубл. 27.07.2015, бюл. №14.

39. Бідюк П. І., Захожай О. І., Меньяйленко О. С. Спосіб розпізнавання образів. Патент на корисну модель № 100078 МПК (2015.01) G06К 9/00; опубл. 10.07.2015, бюл. №13.

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

41. Меньяйленко О. С., Захожай О. І. Спосіб розпізнавання образів. Патент на корисну модель № 92493 МПК (2014.01) G06К 9/00; опубл. 26.08.2014, бюл. №16.

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.

## SUMMARY

***Zakhozhay O. I. Models, Methods and Information Technology of Hybrid Patterns Recognition for the Consolidate Data Processing in Complex Systems.*** – Qualifying scientific work as a manuscript.

Thesis for a Doctor of Technical Science degree in specialty 05.13.06 – Information technologies. – Luhansk Taras Shevchenko National University of Ministry of Education and Science of Ukraine, Starobilsk, 2020.

The dissertation research is directed to the solution of a scientific and technical problem, related with impossibility of simultaneous increase in reliability and efficiency of decision-making on classification in complex systems with non-uniform data and various manifestations of interferences and distortions as these indicators are competing. On the one hand, increase in amount of information signs, use of more difficult algorithms, allows to promote reliability of recognition's result, but, on the other hand, it leads to increase the time-complexity process of data analysis and decision making. The result obtaining and data analysis time increasing in many cases has critical character because the result, which was received with essential time delays, loses actuality for further using and becomes doubtful. Thus, attempts to maximize result reliability reduces its worth for further using. This situation has the most critical character for information processes where decision-making it has to be carried out in real time. From such processes it is possible to distinguish: analysis of temperature's spatial distribution for coke pie in coke-chemical industry, ultrasonic measurement of linear distance to objects for automation tools, analysis and prognosing gas emission from coal layers for mining industry, etc. Thus, the main argument for choice the direction of researches was practice inquiry: Creation of accurate and high-speed systems for heterogeneous data processing in complex systems with different interferences and distortions.

The main dissertation objective is in ensuring the set patterns recognition reliability level with simultaneous reduction of data`s quantity, that provided for comparison in complex systems with heterogeneous interferences and distortions.

The objective is achieved by solving such main tasks:

- analysis of the existing methods, models and information technologies of patterns recognition for data processing in complex systems;
- the consolidation non-uniform data principle development in complex systems with different interferences and distortions;
- substantiation of conception of hybrid patterns recognition for the consolidated data processing in complex systems with different interferences and distortions;
- the process formalization data presentation and developing of hybrid patterns recognition model;
- development of a model for determining the level of informativeness of signs in dynamics of their changing;
- the informative signs consolidation methods development and patterns recognition for complex systems with heterogeneous data, different interferences and distortions;
- developing of information technology of hybrid patterns recognition for the consolidate data processing in complex systems on the basis of the developed methods and models;
- implementation of the developed models, methods and information technology for various applied fields.

The hypothesis of scientific research is consisting in the following. If to present subject to recognition as set of patterns where signs have different nature of origin, to increase thereby amount of non-uniform signs on which interferences have a different influence, follow to carry out consolidation of data and select informative patterns and their signs, having defined those which in the current conditions have the highest informational content, then, after analysis, it is possible



to receive desirable reliability of result and minimize quantity of data, that provided for comparison. The selection of the most informative data will be provide obtaining reliable result in wide range of input data changing and different interferences and distortions.

The main results of scientific research, submitted for protection consist in further development of the theory of the combined recognition in the direction of complex system information model modification for non-uniform data processing; substantiation of the non-uniform data consolidation principle; according with the new conception of hybrid recognition, developing models, methods and information technology for non-uniform data processing in complex systems with heterogeneous interferences and distortions.

The scientific novelty of a dissertation research consists in the following.

*For the first time:*

- was constructed the situation-event model of hybrid patterns recognition, based on representation of characteristics of a complex system as the set of patterns, where signs have a different nature origin, plural of characteristics of external conditions – as manifestation of a situation, static component of a situation – in the form of plural of stationary informativity characteristics, dynamic component – in the form of plural of nonstationary informativity characteristics, and sets of classes, as result of performance the actions sequence for recognition;

- was developed the method of non-uniform data consolidation at hybrid recognition, which the bases on the theory of rough plurals and determination of stationary and non-stationary criteria of informational content, according to level of interferences and distortions, and allows to carry out selection of complex system patterns for decreasing the data comparisons quantity for decision-making;

- was developed a method and information model for determination data processing sequence, that based on generalization of human emotional processes model, that was presented by Atkins-Shifrin, for complex systems with heterogeneous data and different interferences and distortions, what allows to

reduce the data's quantity which are comparison subject to obtaining reliable result;

- was developed a method and information technology of hybrid patterns recognition, which are based on the separate analysis of informative signs and finding the groups of patterns with identical classification, what allows to provide a priori reliability level of heterogeneous data processing result in complex systems with different interferences and distortions;

- was developed the method of extensional-intensional signs analysis, which the based-on adoption of the fast decision on the generalized characteristics of a complex system and using, if there is a need, the detailed analysis of data for specification of result, what allows to reduce the number of comparisons of data when receiving a priori level of classification reliability;

- was developed information model of decision-making supporting, which in case of obtaining multi-alternative classification result, due to use of information about groups of patterns with identical classification, allows to present the received results of classification for the expert opinion (alternatives) are ranged as reduction of their reliability, also information on set of patterns which provided such classification.

The combined patterns recognition theory *was further developed*, where the recognition object description is carried out to signs of various nature of origin, but not various nature of information, what allows to carry out reliable classification on condition of wide changes of different interferences and distortions.

*It is improved* matrix information model of a complex system, where non-uniform data are packed not in one, but in different layers, what provides big flexibility in the choice of alternatives for data consolidation and methods of their processing.

The practical value of the received results consists in suitability for practical application of the developed approach, methods, information technology of data processing on the base of hybrid patterns recognition, what allows to make the

reliable and operative decision in a wide range of input data conditions changes and interferences and distortions level.

The Introduction contains general characteristic of work, what emphasizes and proves its relevance and chosen direction study degree, link with scientific programs and to subjects, the scientific problem is formulated, objective and tasks for researching, subject, object, hypothesis, methods of researching, scientific novelty and practical value of the received results.

The first chapter is devoted to the analysis of methods and means of data processing in complex systems. In chapter is carry out complex systems analysis, as result the main problem is defined. this problem is related to a certain level uncertainly and stochastically influence input data on recognition object state. As a result, there is a complexity in the choice or development of mathematical model that provides formalized the description of behavior on the basis of prior information. Thus, for data processing and decision-making in complex systems have a reason also a posteriori information about the complex system and external environmental condition.

The second chapter is devoted to development and substantiation of the new principle of non-uniform data consolidation, models and methods of signs informativity determination and hybrid patterns recognition for decision-making in complex systems.

In the third chapter, on the basis of the presented models, the developed methods of non-uniform data consolidation and hybrid patterns recognition for complex systems with diverse manifestations of interferences and distortions.

The fourth chapter is devoted to development of information technology of hybrid patterns recognition for the consolidated data processing in complex systems with diverse manifestations of interferences and distortions.

The fifth chapter is devoted to applied information technologies of hybrid patterns recognition. The models, methods and information technology were used at creation of three various information processing systems: in information system

of analysis of temperature's spatial distribution for coke pie, in information system of ultrasonic measurement of linear distance for automation tools, and also in information system of definitions of text data uniqueness level.

The conclusions describe the main results of the dissertation research.

The appendices contain the data of experimental researching of applied information technologies, what will be obtained intime new scientific-technical results implementation, and also the copies of implementation acts.

**Keywords:** combined patterns recognition, hybrid patterns recognition, data processing in complex systems, time complexity of data processing, classification reliability of complex systems state, informative signs selection, data processing queue.

## LIST OF PUBLICATIONS

*Publications where the main scientific results of the dissertation are published*

1. Menyaylenko O. S., Zakhzhay O. I. Osnovy syntezy klasyfikatoriv tehniknykh system rozpoznavannya obraziv z vykorystannyam modelej emotsiynykh protsessiv lyudyny // Scientific and technical journal “Scientific Bulletin of National Mining University”. – 2015. – № 1(145). – P. 120–126. *(The publication is included in the international scientometric bases: SciVerse Scopus, Ulrich's Periodicals Directory (ULRICHSWEB), EBSCO, eLIBRARY, Index Copernicus)*

2. Zakhzhay O. I. Informatsiyna tekhnologiya rozpoznavannya obraziv v zadachakh avtomatyzovanoi obrobky informatsii i upravlinnya skladnyymi systemami // Journal “Problems of information technologies”. – 2013. – № 01(013). – P. 61–68. *(The publication is included in the international scientometric*

bases: *Open Academic Journals Index (OAJI)*, *eLIBRARY*, *Research Bible*, *AcademicKeys*)

3. Zakhochay O. I. Ekstensionalno-intensionalnyj pidkhid do syntezy informatsiynykh tekhnologij avtomatyzovanoi obrobky informatsii i upravlinnya na bazi bagatoparmetrychnykh kombinovanykh system rozpiznavannya obraziv // Journal “Problems of information technologies”. – 2015. – № 02(018). – P. 106–111. *(The publication is included in the international scientometric bases: Open Academic Journals Index (OAJI), eLIBRARY, Research Bible, AcademicKeys)*

4. Menyaylenko O. S. Zakhochay O. I. Kombinovani systemy rozpiznavannya obraziv pry analizi prostorovogo rozpodilu temperatury koksovogo pyroga // *Electrotechnical and computer systems*. – 2013. – № 12(88). – P. 147–154. *(The publication is included in the international scientometric bases: Ulrich’s Periodicals Directory (ULRICHSWEB), Electronic Journals Library, Google Scholar)*

5. Zakhochay O. I. Kryterii vyznachennya informatyvnosti ta ranzhuvannya obraziv pru pryjnyatti rishen v bagatoparmetrychnykh kombinovanykh systemakh rozpiznavannya // *Electrotechnical and computer systems*. – 2018. – № 27(103). – P. 196–204. *(The publication is included in the international scientometric bases: Ulrich’s Periodicals Directory (ULRICHSWEB), Electronic Journals Library, Google Scholar)*

6. Zakhochay O. I. Seleksiya ratsionalnoi sukupnosti obraziv v kombinovanykh systemakh rozpiznavannya // *Electrotechnical and computer systems*. – 2013. – № 09(85). – P. 186–192. *(The publication is included in the international scientometric bases: Ulrich’s Periodicals Directory (ULRICHSWEB), Electronic Journals Library, Google Scholar)*

7. Zakhochay O. I., Menyaylenko A. S., Lyfar V. A. Razdelnyj analiz informatsionnykh priznakov v mnogoparmetrychnykh kombinovanykh sistemakh raspoznavaniya obrazov // *Problemele Energeticii Regionale*. – Chişinău, Republica Moldova, 2019. – № 1-1(40). – P. 60–68. *(The publication is*

*included in the international scientometric bases: Emerging Sources Citation Index Web of Science of Thomson Reuters, Ulrich's Periodicals Directory (ULRICHSWEB), ProQuest LLC, Directory of Open Access Journal (DOAJ), eLIBRARY)*

8. Zakhochay O. I., Filimontsev V. V. Minimizatsiya vremennoj slozhnosti v mnogokriterialnykh sistemakh obrabotki informatsii // Journal "Problems of information technologies". – 2014. – № 01(015). – P. 161–166. *(The publication is included in the international scientometric bases: Open Academic Journals Index (OAJI), PИИЦ (eLIBRARY), Research Bible, AcademicKeys)*

9. Zakhochay O. I. Spilnyj analiz informatsijnykh oznak u bagatoparmetrychnykh kombinovanykh sistemakh rozpiznavannya obraziv // Electrotechnical and computer systems. – 2018. – № 29(105) – P. 78–86. *(The publication is included in the international scientometric bases: Ulrich's Periodicals Directory (ULRICHSWEB), Electronic Journals Library, Google Scholar)*

10. Ryabenky V. M., Zakhochay O. I. Kombinovani systemy rozpiznavannya obraziv // Journal "Problems of information technologies". – 2011. – № 01(009). – P. 156–160. *(The publication is included in the international scientometric bases: Open Academic Journals Index (OAJI), eLIBRARY, Research Bible, AcademicKeys)*

11. Zakhochay O. I., Chibisova Yu. V. Uluchshenie izobrazhaniya daktiloskopicheskikh otpechatkov v sistemakh kontrolya dostupa // Scientific bulletin of DonSTU. – 2011. – № 33. – P. 353–358.

12. Zakhochay O. I., Shevtsova A. S. Informatsijna model systemy nepryamogo analiza stanu koksovoj pechi zasobamy kompyuternogo zoru // Scientific bulletin of DonSTU. – 2010. – № 30 – P. 184–188.

13. Menyaylenko O. S., Zakhochay O. I. Informatsijna tekhnologiya avtomatyzovanoj obrobky danykh v bagatokanalnykh sistemakh ultrazvukovogo vymiryuvannya // Research Bulletin of the National Technical University of

Ukraine "Kyiv Politechnic Institute". – 2014. – № 6(98). – 3. 62–67. (*The publication is included in the international scientometric bases: 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. Zakhohzhay O. I., Paerand Yu. E. Osnovni aspekty strukturnoj organizatsii kombinovanykh system rozpoznavannya obraziv // Bulletin of KhNTU. – 2012. – №1(44). – P.221–225.

15. Zakhohzhay O. I. Efektyvni sposoby peretvorennia grafichnykh zobrazen pry ikhnyoi segmentatsii ta rozpoznavanni zasobamy kompyuternogo zoru // Scientific bulletin of DonSTU. – 2008. – № 26. – P. 338–342.

16. Zakhohzhay O. I. Opredelenie proektivnoy struktury i dvizheniya po binolularnym sootvetstviyam v sistemakh tekhnicheskogo zreniya // Scientific bulletin of DonSTU. – 2007 – № 25. – P. 243–249.

17. Zakhohzhay O. I., Soloshenko A. S. Vysokoskorostnaya obrabotka izobrazhenij s primeneniem tsyfrovnykh filtrov na osnove FPGA // Scientific bulletin of DonSTU. – 2013. – №39. – P. 250–254.

18. Menyaylenko O. S., Zakhohzhay O. I., Bidyuk P. I. Pidvyschennya dostovirnosti perevirky unikalnosti tekstiv shlyakhom vykorystannya kombinovanykh system rozpoznavannya obraziv// Journal "System researches and information technologies" – 2017. – № 4. – P. 29–37. (*The publication is included in the international scientometric bases: Directory of Open Access Journal (DOAJ), EBSCO, Index Copernicus, eLIBRARY*)

19. Zakhohzhay O. I., Lyfar V. O., Baturin O. I. Pryjnyattya rishen na osnovi poshuku grup identychnykh klasyfikatsij v bagatoparmetrychnykh kombinovanykh sistemakh rozpoznavannya obraziv // Bulletin of Volodymyr Dahl

East Ukrainian National University. – 2019. – №2 (250). – P. 36–42. (*The publication is included in the international scientometric bases: Index Copernicus*)

20. Zakhochay O. I. Pidvyschennya dostovirnosti rozpoznavannya obektiv za umovy nayavnosti vykryvlen ikhnyogo vidobrazhennya u prostori obraziv // Bulletin KhNTU. – 2013. – №1 (46). – P.128-131.

21. Zakhochay O. I. Kontsepsiya vdoskonalennya teoretyko-metodologichnykh osnov syntezu informatsijnykh tekhnologij avtomatyzovanoi obrobky informatsii i upravlinnya skladnymy systemamy // Bulletin of Volodymyr Dahl East Ukrainian National University. – 2018. – № 6(247). – P. 49–55. (*The publication is included in the international scientometric bases: Index Copernicus*)

22. Zakhochay O. I., Shevtsova A. S. Ispolzovanie modeli temperaturnogo rezhima v avtomatizirovannoj susteme upravleniya koksovoj pechyu // Scientific bulletin of DonSTU. – 2011. – № 33. – P. 365–370.

23. Zakhochay O. I., Lyfar V. O., Ivanov V. G. Konsolidatsiya danykh ta pryjnyattya rishen na osnovi ranzhuvannya grup identychnykh klasyfikatsij v gibrydnykh systemakh rozpoznavannya // Electrotechnical and computer systems. – 2019. – № 31(107) – P. 95–103. (*The publication is included in the international scientometric bases: Ulrich's Periodicals Directory (ULRICHSWEB), Electronic Journals Library, Google Scholar*)

24. Zakhochay O. I. Vyznachennya chergy obrobky danykh pry gibrydnomu rozpoznavanni obraziv // Bulletin of Volodymyr Dahl East Ukrainian National University. – 2019. – № 7(255). – P. 111–116. (*The publication is included in the international scientometric bases: Index Copernicus*)

25. Zakhochay O. I. Informatsijna tekhnologiya gibrydnogo rozpoznavannya obraziv dlya obrobky neodnorodnykh danykh v skladnykh systemakh // Bulletin of Volodymyr Dahl East Ukrainian National University. – 2019. – № 8(256). – P. 141–147. (*The publication is included in the international scientometric bases: Index Copernicus*)



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

27. Menyaylenko O., Zakhohzhay O. Organizatsiya perevirky studentskykh I naukovykh robit na plagiat // Scientific bulletin of Melitopol State Pedagogical University. – 2017. – № 1(18). – P. 157–162.

*Publications that confirm the approbation of dissertation materials*

28. Zakhohzhay O. I. Vykorystannya systemy tekhnichnogo zoru dlya upravlinnya nozhytsyamy metalurgijnogo vyrobnytstva// Pratsi 8-i mizhnarodnoi naukovo-praktychnoi konferentsii “Suchasni informatsijni ta elektronni tekhnologii – 2007”. – Odessa: ONPU, 2007. – P. 81.

29. Zakhohzhay O. I. Informatsijna tekhnologiya rozpiznavannya obraziv pryanalizi temperaturnogo rezhymu koksovykh pechej // Pratsi 10-i mizhnarodnoi naukovo-praktychnoi konferentsii “Suchasni informatsijni ta elektronni tekhnologii – 2009”. – Odessa: ONPU, 2009. – P. 77.

30. Zakhohzhay O. I., Shevtsova A. S. Analiz stanu koksovoi pechi zasobamy kompyuternogo zoru // Pratsi 11-i mizhnarodnoi naukovo-praktychnoi konferentsii “Suchasni informatsijni ta elektronni tekhnologii – 2010” т.1. – Odessa: ONPU, 2010. – P. 34.

31. Zakhohzhay O. I., Shevtsova A. S. Ispolzovanie modeli temperaturnogo rezhyma koksovaniya v avtomatizirovannoj sisteme upravleniya koksovoj pechyu // Pratsi 12-i mizhnarodnoi naukovo-praktychnoi konferentsii “Suchasni informatsijni ta elektronni tekhnologii – 2011”. – Odessa: ONPU, 2011. – P. 127.

32. Zakhohzhay O. I. Kombinivani systemy rozpiznavannya obraziv // Pratsi 13-i mizhnarodnoi naukovo-praktychnoi konferentsii “Suchasni informatsijni ta elektronni tekhnologii – 2012”. – Odessa: ONPU, 2012. – P. 113.

33. Zakhohay O., Paerand Yu. Combined Systems of Patterns Recognition // Kontrol i upravlinnya v skladnykh systemakh (KUSS-2012). 11-a mizhnarodna konferentsiya. Thesis. – Vinnytsya: VNTU, 2012. – P. 241.

34. Zakhohay O. I. Kryterij selektsii ratsionalnoi sukupnosti obraziv v kombinovanykh systemakh rozpoznavannya // Systemnyj analiz. Informatyka. Upravlinnya. (SAIU-2013) Mizhnarodna naukovo-praktychna konferentsiya. – Zaporizhya: KPU, 2013. – P. 114–116.

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

36. Zakhohay O. I. Gibrydne rozpoznavannya obraziv dlya obrobky neodnorodnykh danykh v skladnykh systemakh // 9-a vseukrainska naukovo-praktychna konferentsiya “Elektronika ta telekomunikatsii-2019” – Severodonetsk: SNU, 2019. – P. 113–115.

*Publications that additionally reflect the scientific results of the dissertation*

37. Menyaylenko O. S., Zakhohay O. I. Sposib kontrolyu temperatury nagrivu koksovoi pechi. Patent of Ukraine № 91118 MPC (2014.01) C10B 17/00, C10B 23/00; publ. 25.06.2014, bull. №9.

38. Bidyuk P. I., Zakhohay O. I., Menyaylenko O. S. Sposib rozpoznavannya obraziv. Patent of Ukraine № 100283 MPC (2015.01) G06K 9/00; publ. 27.07.2015, bull. №14.

39. Bidyuk P. I., Zakhohay O. I., Menyaylenko O. S. Sposib rozpoznavannya obraziv. Patent of Ukraine № 100078 MPC (2015.01) G06K 9/00; publ. 10.07.2015, bull. №13.

40. Menyaylenko O. S., Zakhohay O. I. Prystrij rozpoznavannya obraziv. Patent of Ukraine № 90109 MPC (2014.01) G06/00; publ. 12.05.2014, bull. №9

41. Menyaylenko O. S., Zakhohay O. I. Sposib rozpiznavannya obraziv. Patent of Ukraine № 92493 MPC (2014.01) G06K 9/00; publ. 26.08.2014, bull. №16.

42. Zakhohay O. I., Menyaylenko O. S. Prystrij vymiryuvannya vidstani za dopomogoyu ultrazvuku. Patent of Ukraine № 93963 MPC (2006.01) G01S 7/52, G01S 15/14; publ. 27.10.2014, bull. №20.

## ЗМІСТ

ПЕРЕЛІК СКОРОЧЕНЬ.....	32
ВСТУП .....	33
РОЗДІЛ 1 АНАЛІЗ МЕТОДІВ І ЗАСОБІВ ОБРОБКИ ДАНИХ В СКЛАДНИХ СИСТЕМАХ .....	45
1.1 Підходи і проблеми обробки даних в складних системах .....	45
1.1.1 Загальна характеристика складних систем. ....	45
1.1.2 Методи обробки даних в складних системах.....	46
1.2 Розпізнавання образів в задачах обробки даних складних систем .....	52
1.2.1 Області використання методів та інформаційних технологій розпізнавання образів. ....	52
1.2.2 Методи та інформаційні технології розпізнавання образів.....	55
1.2.3 Методи розпізнавання образів для обробки неоднорідних даних	58
1.3 Інформаційна модель обробки даних в складних системах засобами розпізнавання образів.....	67
1.4 Методи аналізу інформативності неоднорідних даних.....	70
1.4.1 Метод визначення інформативності ознак на основі пошуку в навчальній виборці інформативної зони .....	72
1.4.2 Метод визначення інформативності ознак на основі кластеризації .....	74
1.4.3 Метод визначення інформативності ознак, заснований на статистичному кластер-алгоритмі.....	76
1.4.4 Методи визначення інформативності ознак на основі ентропії....	77
1.4.5 Метод визначення інформативності неоднорідних ознак на основі непараметричних оцінок щільності .....	80
1.4.6 Оцінка мінімальної кількості ознак, що забезпечують апіорно заданий рівень достовірності розпізнавання образів .....	81
1.5 Постановка завдань досліджень і розробок.....	84
Висновки до розділу 1.....	93

РОЗДІЛ 2 КОНСОЛІДАЦІЯ НЕОДНОРІДНИХ ДАНИХ ТА РОЗПІЗНАВАННЯ ОБРАЗІВ В СКЛАДНИХ СИСТЕМАХ З РІЗНОРІДНИМИ ПРОЯВАМИ ПЕРЕШКОД І ВИКРИВЛЕНЬ .....	96
2.1 Принцип консолідації неоднорідних даних в СС з різнорідними проявами перешкод і викривлень .....	96
2.2 Модифікована матрична інформаційна модель обробки неоднорідних даних в складних системах.....	99
2.3 Концепція гібридного розпізнавання образів для консолідованої обробки даних в СС з різнорідними проявами перешкод і викривлень..	104
2.4 Ситуаційно-подійна модель гібридного розпізнавання образів.....	107
2.5 Інформаційна модель і метод визначення послідовності обробки даних образів складної системи .....	117
Висновки до розділу 2.....	126
РОЗДІЛ 3 МЕТОДИ ГІБРИДНОГО РОЗПІЗНАВАННЯ ОБРАЗІВ ТА ІНФОРМАЦІЙНА МОДЕЛЬ ПІДТРИМКИ ПРИЙНЯТТЯ РІШЕННЯ ДЛЯ ВИЗНАЧЕННЯ СТАНУ СКЛАДНИХ СИСТЕМ З РІЗНОРІДНИМИ ПРОЯВАМИ ПЕРЕШКОД І ВИКРИВЛЕНЬ .....	129
3.1 Методи гібридного розпізнавання образів для СС з різнорідними проявами перешкод і викривлень .....	129
3.1.1 Спільний аналіз інформаційних ознак.....	129
3.1.2 Роздільний аналіз інформаційних ознак.....	140
3.1.3 Метод гібридного розпізнавання образів, оснований на роздільному аналізу ознак складної системи.....	142
3.1.4 Модифікований метод гібридного розпізнавання на основі пошуку груп образів з ідентичною класифікацією.....	149
3.1.5 Інформаційна модель підтримки прийняття рішення у випадку отримання багатоальтернативного результату .....	153
3.3 Метод консолідації образів СС та їх селекція на основі характеристик інформативності.....	155

3.4 Метод екстенціонально-інтенціонального аналізу ознак при гібридному розпізнаванні образів.....	164
Висновки до розділу 3.....	168
<b>РОЗДІЛ 4 ІНФОРМАЦІЙНА ТЕХНОЛОГІЯ ГІБРИДНОГО РОЗПІЗНАВАННЯ ОБРАЗІВ ДЛЯ КОНСОЛІДОВАНОЇ ОБРОБКИ ДАНИХ В СКЛАДНИХ СИСТЕМАХ З РІЗНОРІДНИМИ ПРОЯВАМИ ПЕРЕШКОД І ВИКРИВЛЕНЬ.....</b>	
4.1 Граничні умови використання інформаційної технології.....	171
4.1.1 Часові умови.....	171
4.1.2 Умови багатозадачності.....	173
4.2 Вимоги до апаратно-програмних засобів реалізації інформаційної технології.....	176
4.2.1 Загальні вимоги.....	176
4.2.2 Вимоги до апаратного забезпечення інформаційної технології.....	177
4.2.3 Вимоги до програмного забезпечення інформаційної технології.....	181
4.3 Реалізація інформаційної технології гібридного розпізнавання образів.....	182
4.3.1 Формування черги обробки даних при гібридному розпізнаванні.....	182
4.3.2 Етапи реалізації інформаційної технології гібридного розпізнавання образів.....	184
4.3.3 Апріорне визначення складових характеристик ситуаційно-подійної моделі гібридного розпізнавання.....	187
4.3.4 Структурно-логічна послідовність процесів обробки даних в інформаційній технології гібридного розпізнавання образів.....	190
4.4 Рекомендації щодо впровадження інформаційної технології гібридного розпізнавання образів.....	194
Висновки до розділу 4.....	195

РОЗДІЛ 5 ПРИКЛАДНІ ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ ГІБРИДНОГО РОЗПІЗНАВАННЯ ОБРАЗІВ .....	198
5.1 Інформаційна технологія аналізу просторового розподілу температури коксового пирога для коксохімічного виробництва .....	198
5.2 Інформаційна технологія ультразвукового вимірювання лінійних відстаней для систем автоматики .....	226
5.3 Інформаційна технологія визначення ступеня унікальності тестових даних .....	241
Висновки до розділу 5.....	254
ВИСНОВКИ.....	256
СПИСОК ЛІТЕРАТУРИ.....	260
ДОДАТОК А Таблиця лінійного розподілу температури коксових печей на одній горизонтальній осі .....	286
ДОДАТОК Б Графіки лінійних розподілів температури коксових печей на одній горизонтальній осі .....	289
ДОДАТОК В Таблиця розкиду характеристик між просторовим та еталонним лінійним розподілами температури коксових печей .....	293
ДОДАТОК Г Таблиця корегованого лінійного розподілу температури коксових печей на одній горизонтальній осі.....	295
ДОДАТОК Д Графіки корегованих лінійних розподілів температури коксових печей на одній горизонтальній осі.....	298
ДОДАТОК Е Результат співставлення отриманих температурних характеристик класу <i>C</i> .....	302
ДОДАТОК Ж Акти впровадження результатів дисертаційної роботи .....	305
ДОДАТОК И Список публікацій здобувача .....	310

**ПЕРЕЛІК СКОРОЧЕНЬ**

СС	– складна система
КСРО	– комбінована система розпізнавання образів
КВП	– колективи вирішальних правил
АОО	– алгоритми обчислення оцінок
ДНФ	– диз'юнктивна нормальна форма
ІП	– інформаційний потік
NSICh	– нестационарна характеристика інформативності
SICh	– стаціонарна характеристика інформативності
ICh	– характеристика інформативності
БРХ	– блок реєстрації характеристик
БФ	– блок формування
БК і ПР	– блок класифікації і прийняття рішення
СРЧ	– система реального часу
СЖРЧ	– система жорсткого реального часу
СМРЧ	– система м'якого реального часу
БК	– блок класифікації
ІТ	– інформаційна технологія



## ВСТУП

### **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, 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. Gonzalez, 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. Меньяйленко, 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 автоматки, 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 різнорідних 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 перешкод, 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 різноманітними 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 різноманітними проявами перешкод і викривлень;
- обґрунтування концепції гібридного розпізнавання образів для консолідованої обробки даних в складних системах з різноманітними проявами перешкод і викривлень;
- формалізація процесу представлення даних та розробка інформаційної 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 різноманітними 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 теорію 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 використовувались 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, представлені Аткінсоном-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 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 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 у 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 інформатики 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» (Сєверодонецьк, 2019).

### **Особистий внесок здобувача.**

Наукові положення, висновки, рекомендації та вимоги, викладені в дисертаційній роботі й пропоновані 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 рисунок, 7 таблиць, 8 додатків на 32 сторінках. Бібліографічний список містить 236 найменувань на 26 сторінках.

# РОЗДІЛ 1

## АНАЛІЗ МЕТОДІВ І ЗАСОБІВ ОБРОБКИ ДАНИХ В СКЛАДНИХ СИСТЕМАХ

### 1.1 Підходи і проблеми обробки даних в складних системах

#### *1.1.1 Загальна характеристика складних систем*

Складні системи (СС) мають структурну, обчислювальну, організаційну складність, а також характеризуються нелінійністю, елементами стохастичності, наявністю часткової інформаційної невизначеності [1-3]. Поведінку СС важко моделювати через залежності, відносини або взаємодію між їх частинами або між даною системою і її середовищем. Система визначається як складна [4, 5], якщо для побудови її адекватної моделі недостатньо апріорної інформації. Це означає, що поведінка складної системи істотним чином залежить від чинників, про вплив яких практично нічого невідомо. Тому у разі, коли об'єктом є СС, про її модель можна сказати лише те, що невідомі не лише коефіцієнти або параметри аналітичних виразів, що описують об'єкт, але й навіть сам вид цих виразів. Тому побудова змістовної аналітичної моделі СС дуже проблематична і, в цьому випадку, може йтися лише про побудову моделей інших класів.

Такі системи застосовуються в різних областях. Можна виділяти характерні типи СС [6, 7], загальні і приватні характеристики. СС включають цілий ряд спеціальних систем, найбільш поширеними серед яких є складні технологічні системи [8]. Прикладом технологічних систем є програмні і мультиагентні системи. Слід виділити складні технічні системи [9]. Велике значення в управлінні мають складні організаційно-технічні системи [10, 11]. До числа СС відносяться складні інформаційні системи [12, 13], СС

зберігання даних (сховища [14], багатовимірні бази даних [15], банки даних просторової інформації [16]). Найбільш складними слід вважати кіберфізичні системи [17].

Таке різноманіття СС викликає також і різноманіття методів і алгоритмів обробки інформації в цих системах. Однак особливістю є те, що обробка інформації не є головним завданням, а тільки інструментом досягнення головної мети – прийняття рішень різного характеру, в залежності від сфери застосування СС. Методи і алгоритми обробки даних в СС повинні, в першу чергу, відповідати вимогам аналізу і прийняття рішень і в другу чергу вирішувати завдання оптимальних обчислень.

### *1.1.2 Методи обробки даних в складних системах*

Методи та інформаційні технології обробки даних в СС спрямовані на вирішення таких основних завдань: зменшення інформаційної невизначеності, підготовка рішень, зниження ризику прийняття рішень [18].

Аналіз існуючих методів обробки даних в СС вказує на їхню розрізненість та направленість на вирішення вузькоспеціалізованої задачі. При цьому по теперішній час залишається актуальним, по можливості, об'єднання цих алгоритмів для поетапного вирішення загальної задачі, або створення нових комплексів методів, які можна використовувати в межах однієї інформаційної технології обробки даних та прийняття рішень в СС. Такі методи повинні мати властивості пов'язаності, цілісності, а, в окремих випадках, ємерджентності.

Реальна ситуація підтримки прийняття рішень [11] така, що в розпорядженні особи, що приймає рішення (ОПР) є лише частина інформації для прийняття рішень, а інша частина – невідома.

Найпростішим методом обробки даних є лінійна послідовність, яка має вигляд, наведений на рисунку 1.1. Довжина лінійної послідовності визначає час обчислень. Час послідовної обробки  $t_{носл}$  пропорційний числу етапів  $n$  обробки лінійної послідовності:  $t_{носл}=f(n)$ . В результаті переробки

здійснюється поетапне перетворення вхідного потоку даних  $\{D\}_{ex}$  у вихідний  $\{D\}_{вих}$ .

Схема на рисунку 1.1 називається також маршрутом обчислень [19]. Такий метод описує поетапне перетворення вхідного потоку даних  $\{D\}_{ex}$  в вихідний  $\{D\}_{вих}$  за допомогою функції (або послідовності функцій) перетворення  $\{D\}_{вих} = f(\{D\}_{ex})$ . Існує альтернативний опис методу у формі СС, що надається Месаровичем [20]:  $A_s : \{D\}_{ex} \rightarrow \{D\}_{вих}$ , де  $A_s = f_A(A_1, A_2, A_3, \dots, A_n)$  комплексний алгоритм для кожного з  $n$  етапів. Таким чином, комплексний алгоритм  $A_s$  здійснює перетворення вхідної множини  $\{D\}_{ex}$  в вихідну  $\{D\}_{вих}$ . Якщо функція  $f_A$  лінійна, то алгоритм  $A_s$  є лінійним. Однак, незалежно від виду функції  $f_A$ , вираз  $A_s$  описує прямий алгоритм.

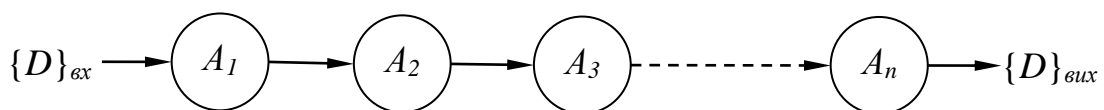


Рисунок 1.1 – Лінійна послідовність обробки даних

На практиці можливі ситуації, коли задається альтернатива вибору, що визначається змінною умовою. Залежно від змінної умови можуть формуватися різні альтернативи рішень. Ця інформаційна ситуація [21, 22] показана на рисунку 1.2.

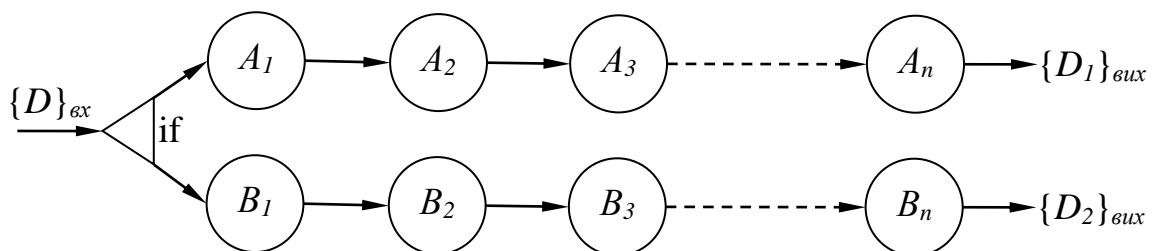


Рисунок 1.2 – Альтернативна обробка даних

Такий варіант містить два комплексних алгоритми  $A_S$  і  $B_S$ , поєднані логічною функцією «АБО»:  $A_S : \{D\}_{ex} \rightarrow \{D_1\}_{вих}$ ,  $A_S : \{D\}_{ex} \rightarrow \{D_2\}_{вих}$ .

Змінні умови і результат їх дії позначені перемикачем (if). Результат виконання методу є дві різних множини вихідних даних  $\{D_1\}_{вих}$  і  $\{D_2\}_{вих}$ . Очевидно, що при зростанні числа альтернатив, кількість ланцюжків буде зростати відповідно до кількості альтернатив. Однак, слід зазначити, що при цьому, час обчислення не буде значно збільшуватися, якщо довжини альтернативних ланцюжків обробки приблизно однакові.

При великому обсязі даних, що підлягають обробці, може бути використаний метод розпаралелювання [19]. Такий метод обробки інформаційної ситуації наведений на рисунку 1.3. На схемі (if1, if2, if3) – роздільники, (int) - інтегратор результатів обробки. Час паралельної обробки  $t_{обр}$  зменшиться в  $m$  разів у порівнянні з часом лінійної обробки:  $t_{нар} = t_{носл} / m$ .

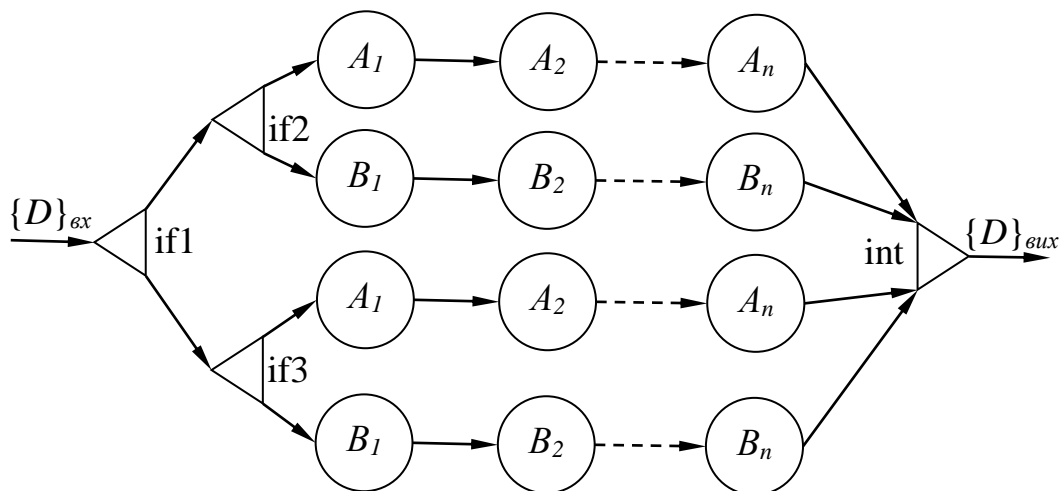


Рисунок 1.3 – Паралельна обробка даних

Схема на рисунку 1.3 представляє собою простий транспортний граф, в якому розгалуження здійснюється на початковому етапі, а в подальшому умови є стаціонарними. Слід ще раз відзначити що розглядаються схеми алгоритмів описують не тільки обчислювальний процес, але і процес



прийняття рішень. Для стаціонарних умов ланцюжки на рисунку 1.3 не перетинаються. Це означає, що маршрути обробки не змінюються.

У випадках, коли можлива зміна умов і наявні залежності між окремими етапами різних маршрутів, в алгоритм транспортного графа додаються перетини і цикли [18, 21]. Така інформаційна ситуація [21-22] показана на рисунку 1.4. Залежно від зміни зовнішніх умов, можливі переходи між різними маршрутами обчислень. Така ситуація вводить необхідність додаткової процедури оптимізації маршрутів для обробки даних та прийняття рішень.

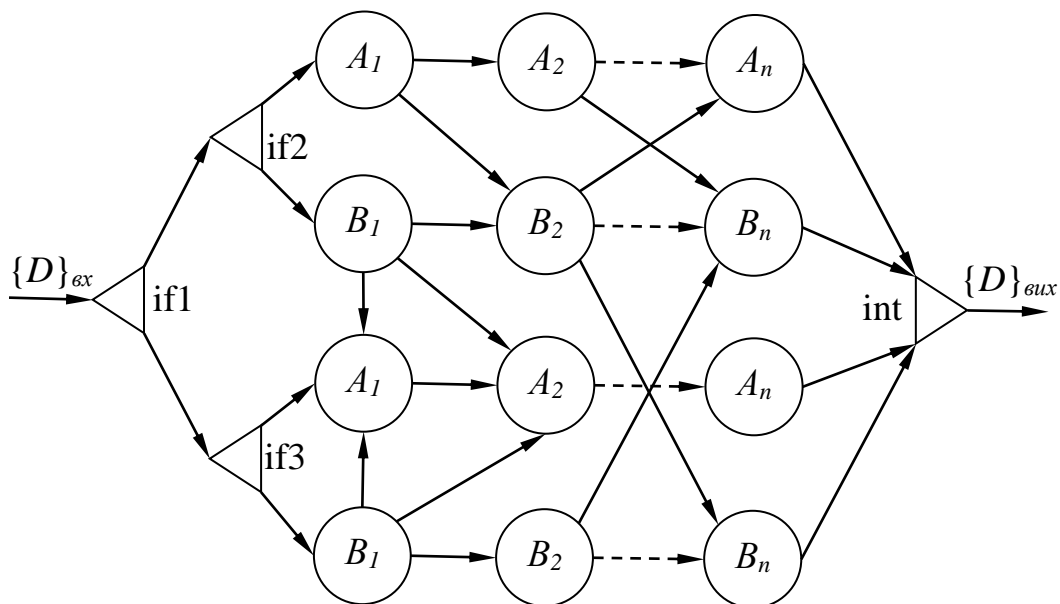


Рисунок 1.4 – Обробка даних при залежних основних маршрутах

Особливий характер процесу обробки даних в СС обумовлений не скільки збільшенням числа параметрів, скільки зростанням численних зв'язків між ними [21, 23]. Причому, моделі зв'язків, фактично, не можливо описати планарними графами. Параметри СС в нестаціонарній ситуації нерідко пов'язані між собою безліччю суперечливих вимог. Ця проблема стає значно вираженою у випадку, коли СС характеризується неоднорідними даними, для обробки яких достатньо складно (а частіше й взагалі неможливо)

використовувати уніфіковані методи і алгоритми. В такому випадку, використовуються алгоритми бустінгу [34-38], комбінаторика алгоритмів [39-42], однак все це значно ускладнює алгоритмічну базу процесу обробки та прийняття рішення, підвищує кількість співставлень даних при аналізі і, як наслідок, збільшує часову складність отримання результату.

Значна складність завдань обробки та прийняття рішень в СС призводить до того, що стає нормою, коли функцією мети є отримання не оптимального, а будь-якого рішення, яке відповідає заданим вимогам в припустимі строки [23]. Тобто процес оптимізації, який вимагає часу, не завжди дозволяє вкластися в задані терміни прийняття рішень, особливо, коли обробка даних повинна проводитися в режимі реального часу [24-26]. Це призводить до відмови від глобальної і переходу до локальної оптимізації та пошуку раціонального рішення, достовірність якого відповідає заданим вимогам. Одним з перспективних підходів до створення методів обробки даних в СС є застосування топологічних методів. В рамках даного підходу, описується завдання обробки в термінах предметної області та в подальшому здійснюється розробка методу рішення.

Вимога підвищення оперативності обробки і прийняття рішень дає підставу проводити алгоритмізацію СС двома етапами [27]. На першому етапі виконується структурне моделювання, виходячи з умов застосування системи і умов завдань, які повинні бути вирішені. На другому етапі проводять локальну оптимізацію простими методами, типу статистичних ігрових або методом Монте Карло. Даний підхід суттєво скорочує час прийняття рішень, оскільки глобальна оптимізація в складних системах може вимагати часових ресурсів, що на порядки перевищують час прийняття рішень або прийняття кількох альтернативних рішень. Крім цього, в окремих СС, процес пошуку оптимуму взагалі не можливий [5, 28, 29]. Тому одним з основних підходів в СС є побудова дозволених топологій в заданих зовнішніх умовах. В результаті, створюється можливість оперативного

пошуку наборів рішень з подальшим вибором серед них локально оптимального (або раціонального за поточних умов) [28-32].

Наведені вище факти тільки ускладнюють задачу створення гнучких, ефективних систем обробки даних в СС, в яких методи і алгоритми виробки управлінських рішень дозволяли б однаково враховувати змінні складові поведінки складного об'єкту.

З погляду на те, внутрішня організація та взаємний зв'язок між окремими параметрами СС як правило, невідомий, висновки щодо поведінки такої системи можна родити тільки на основі її реакції на вхідні данні, тобто спостерігаючи залежності між зміною вхідних параметрів та реакції на них вихідних параметрів [5, 6]. В цьому випадку СС представляється «чорним ящиком» [5], в якому відомий набір вхідних та вихідних параметрів, однак внутрішні процеси та однозначний зв'язок між параметрами невідомий. Такий підхід створює можливості щодо аналізу систем, внутрішня організація яких або взагалі невідома, або достатньо складна, для того, щоб за окремими властивостями її складових і окремим зв'язкам між ними можна було зробити висновки про поведінку системи в цілому. В цих умовах, достатньо широкої популярності для аналізу СС набуло використання елементів теорії ймовірності, коли на основі спостереження за складним об'єктом на протязі певного часу можна було визначити статистичні висновки щодо його поведінки. Однак такі методи тільки частково дозволяють вирішувати задачі системного аналізу СС, так як спостереження за СС протягом довгого часу не завжди дає можливість зафіксувати всі можливі варіанти її поведінки, особливо враховуючи достатньо високу ентропію її вихідних параметрів.

Враховуючи вказані складнощі, в ряді робіт [5, 6, 33] пропонується здійснювати аналіз СС не на основі окремих параметрів, а в результаті прояву їх стану. Визначення стану можливе без формування цілісного образу СС та зовнішнього для системи середовища і такий підхід є адекватним з

погляду на те, що усі параметри складного об'єкта однозначним, але не завжди відомим, образом пов'язані з відповідними станами.

З погляду на вищезгадане, можна зробити висновок, що аналіз СС не за параметрами, а за станом є найбільш раціональним, але вимагає додаткового дослідження існуючих методів і засобів обробки даних та прийняття рішення щодо визначення (класифікації) поточного стану СС для подальшого прийняття необхідних управлінських рішень.

З урахуванням зазначених вище особливостей СС, в ряді робіт [5, 12, 43-52] вказується на доцільність використання апарату розпізнавання образів для обробки даних в складних системах. При цьому, можливе вирішення цілої низки проблем, що мають місце в сучасних інформаційних системах зі складними об'єктами.

## **1.2 Розпізнавання образів в задачах обробки даних складних систем**

### *1.2.1 Области використання методів та інформаційних технологій розпізнавання образів*

Аналіз існуючих методів та інформаційних технологій розпізнавання образів вказує на існуючий потенціал по їхньому застосуванню для обробки даних та прийняття рішень в СС. Це, насамперед, пов'язано з тим, що інтелектуальні алгоритми розпізнавання образів не потребують створення вичерпних, переускладнених математичних моделей об'єкту розпізнавання, а рішення приймаються на основі окремих, найбільш характерних відомостей про нього. Цей підхід особливо актуальний для СС, де відсутня можливість повної структурованості поведінки, а також частково або повністю невідомий аналітичний зв'язок між вхідними та вихідними параметрами.

Одним з найбільш перспективних напрямків використання методів розпізнавання образів є створення адаптивних систем обробки даних в СС [5, 12, 53]. На практиці створення таких систем стримується недостатньою

розробленістю теорії, стандартних, перевірених і відпрацьованих моделей і методів обробки даних, а також недоліком відповідних інструментальних і програмних засобів реалізації відповідних інформаційних технологій [5, 45].

В інформаційних технологіях розпізнавання образів, для отримання адекватного результату класифікації немає необхідності врахування усього переліку наявних характеристик СС. Відмінною рисою є відмова від прагнення створити вичерпну математичну модель складного об'єкту, а замість цього задовольнитися приватним рішенням, на основі використання різноманітної сукупності евристик і апарату розпізнавання образів.

Перша вітчизняна робота в області розпізнавання образів була виконана одним з основоположників сучасної теорії інформації акад. А.А. Харкевичем [54, 55]. Значний внесок у розвиток теорії і практики розпізнавання внесли також всесвітньо відомі вчені-академіки В.М. Глушков, Ю.І. Журавльов, В.С. Михалевич, В.С. Пугачов, Я.З. Ципкін, В.Л. Матросов і К.В. Рудаков [56-62] та ін. Із зарубіжних вчених слід зазначити в першу чергу F. Rosenblatt [63], що запропонував у 1957 р. машину, що навчається розпізнаванню образів, названу їм як персептрон (від англ. to percept – сприймання), в якості простої моделі діяльності мозку, пов'язаної з розпізнаванням образів. Крім того, необхідно назвати видатних вчених R. Gonzalez, U. Grenander, R. Duda, G. Tu, K. Fu, P. Hart, основні роботи яких перекладені на багатьох мовах [64-70].

Розвиток обчислювальної техніки та інформаційних технологій забезпечив сприятливі умови для реалізації технічних систем розпізнавання, в яких закладені основні природні концепції, що мають прояв у біологічних системах.

Область застосування інформаційної технології розпізнавання образів досить широка. Фактично, будь які інформаційні процеси, в тому чи іншому прояві, використовують апарат розпізнавання. При цьому, процес обробки

даних та прийняття рішення не є винятком. Нижче наведений малий перелік існуючих застосувань розпізнавання образів.

- Розпізнавання зорових образів, і побудова систем технічного зору [47-49, 51, 70-71].

- Розпізнавання звукових образів [46, 72-74].

- Розпізнавання стану технічних об'єктів і системи технічної діагностики [5, 12, 45, 50, 52, 75, 76].

- Аналіз екологічної обстановки, прогнозування сейсмічної активності, рівня річок, динаміки та просторового розподілу забруднень тощо [33, 77-79].

- Прогнозування врожайності, системи підтримки прийняття рішень в задачах агропромислового комплексу [80, 81].

- Завдання розпізнавання нафтоносних і водоносних пластів за непрямыми геофізичними даними [82-83].

- Аналіз стану технічних систем і формування управлінських рішень в автоматизованих системах управління складними багатокomпонентними комплексами та системами [5, 12, 84-85].

- Аналіз економічних і геополітичних процесів [87-91].

- Системи соціологічних досліджень [92, 93].

Як правило, конкретний результат в цих завданнях може бути представлений в узагальненій формі, як приналежність об'єкта одному із заздалегідь фіксованих класів. Початкові дані цих завдань, як правило, завдаються у вигляді дискретних відомостей про об'єкти, що піддаються розпізнаванню, наприклад у вигляді набору об'єктів, що заздалегідь класифікуються. З математичної точки зору це означає, що розпізнавання образів є далекосяжним узагальненням ідеї екстраполяції функції [5, 12].

Слід зазначити, що в інформаційній технології розпізнавання образів, за умови значного ступеня невизначеності при прийнятті рішення щодо класифікації, можуть бути використані евристики, що притаманні людині, як об'єкту з раціонально-нераціональним мисленням [94, 95]. Базові

дослідження в цьому питанні проведені такими вченими, як Г. Райфа, М. Алле, А. Тверский, П. Словик, Б. Фишхоф, Д. Канеман, С. Лихтенштейн [96-99].

### *1.2.2 Методи та інформаційні технології розпізнавання образів*

За характером інформації про ознаки об'єктів розпізнавання, методи поділяються на детерміновані, імовірнісні, логічні і структурні [53, 65, 100-102].

Детермінованим методам розпізнавання образів приділяється належна увага у роботах таких вчених, як U. Grenander, J. Tu, R. Gonzalez, K. Fu, R. Duda, P. Hart, М.З. Згуровський, В.І. Васильєв, В.С. Симоненко, Е.В. Луценко, О.Л. Горелік, В.О. Скрипкін та ін. [5, 12, 64, 65, 100, 103-109]. В таких методах використовуються геометричні норми близькості, засновані на вимірі відстаней між об'єктом розпізнавання і еталонами класів. У загальному випадку застосування детермінованих методів розпізнавання передбачає наявність координат еталонів класів у просторі ознак або координат об'єктів, що належать відповідним класам.

Відносно імовірносних методів розпізнавання образів, які хронологічно були перші на шляху розвитку систем розпізнавання, існує достатньо розроблена науково-технічна база, включаючи глибоку статистичну теорію розпізнавання образів В.Н. Валніка – О.Я. Червоненкіса [109]. Крім цього, відомі роботи таких вчених: Ю.Л. Барабаш, В.А. Скрипкін, Р. Hart, Ф.Е. Темніков, J. Tu, R. Gonzalez, P. Winston, G.J. McLachlan, H. Niemann, V. Ripley та ін. [64, 65, 100, 110-115]. Імовірнісні методи засновані на теорії статистичних рішень. У загальному випадку застосування цих методів передбачає наявність імовірнісних залежностей між ознаками об'єктів розпізнавання і класами, до яких ці об'єкти відносяться.

Логічні методи розпізнавання знайшли відображення в роботах В.І. Левіна, Ю.С. Сагдулаєва, Л.А. Растригіна, Р.Х. Еренштейна, Л.Т. Кузіна, Ф.П. Тарасенка [101, 116-122]. Логічні методи розпізнавання засновані на

дискретному аналізі та елементах булевої алгебри і логічних висловлюваннях. У загальному випадку, застосування логічних методів розпізнавання передбачає наявність логічних зв'язків, виражених через систему булевих рівнянь, в якій змінні являють собою логічні ознаки об'єктів розпізнавання, а невідомі величини, це класи, до яких ці об'єкти відносяться. Форма представлення класів в логічних системах розпізнавання в загальному вигляді може бути представлена так:

$$\exists\{x_n \mid P\} \in C, B^n \rightarrow B, \quad (1.1)$$

де  $x_n$  – множина  $n$  ознак образу  $P$ ;

$C$  – клас, на відповідність якому перевіряється образ  $P$ ;

$B^n$  – деяка визначена булева функція  $n$  аргументів, що представляють ознаки  $x_n$  образу  $P$ ;

$B$  – булева множина  $\{0,1\}$ .

Відповідно до (1.1) за приналежність образу  $P$  до визначеного класу  $C$ , може свідчити одне із значень, булевої множини  $B$ . Очевидно, що інше значення буде свідчити, що  $P \notin C$ .

Структурні (лінгвістичні) методи розпізнавання відбиті в роботах таких вчених: К. Fu, David A. Forsyth, J. Ponce, Y. Park, В.І. Васильєв, В.М. Кляцкін, Е.В. Щепін, К.М. Зингерман, В.В. Лазарев [49, 66, 71, 103, 123-129]. В цих методах розпізнавання використовуються спеціальні граматики, які породжують мови, що складаються з речень. Кожне з таких речень описує об'єкти, що належать конкретному класу. Застосування структурних методів розпізнавання вимагає наявності сукупностей речень, що описують усю множину об'єктів. Ці об'єкти належать усім класам алфавіту системи розпізнавання. При цьому, множина речень має бути розділена на підмножини по числу класів системи. Елементами цих підмножин є речення,



що описують окремі об'єкти. Таким чином, апріорними описами класів є сукупності речень, кожне з яких відповідає конкретному об'єкту, що належить цьому класу. В структурному підході до розпізнавання широко використовується композиція образів у відповідні речення, до більш глобальних образів і речень, а також декомпозиція глобальних образів і речень на підобрази і відповідні їм речення [71, 123].

Різні автори (Ю.Л. Барабаш [139], В.І. Васильєв [103], А.Л. Горелик, В.А. Скрипкін [100], Р. Дуда, Р.Нат [64], Л.Т. Кузін [118], Ф.І. Перегудов, Ф.П. Тарасенко [4], Ф.Е. Темников [111], Дж. Ту, R.Gonsalez [65], Р.Winston [112], К. Фу [66], Я.З. Ципкін [60]) дають різну класифікацію методів розпізнавання образів. Одні автори розрізняють параметричні, непараметричні і евристичні методи, інші – виділяють групи методів, виходячи з історично сформованих шкіл і напрямів в даній області. Наприклад, в роботі [129], в якій дано достатньо повний огляд методів розпізнавання, використовується наступна типологія методів розпізнавання образів:

- методи, засновані на принципі поділу;
- статистичні методи;
- методи, побудовані на основі "потенційних функцій";
- методи, засновані на численні висловів, зокрема на апараті алгебри логіки;
- методи обчислення оцінок (голосування).

В основі цієї класифікації лежить відмінність у формальних методах розпізнавання образів і тому опущений розгляд евристичного підходу до розпізнавання, що отримав повний і адекватний розвиток в експертних системах. Евристичний підхід заснований на знаннях, що важко формалізуються, і інтуїції експерта. При цьому експерт сам визначає, яку інформацію і яким чином система повинна використовувати для досягнення необхідного ефекту розпізнавання.

### 1.2.3 Методи розпізнавання образів для обробки неоднорідних даних

Значна кількість існуючих СС характеризуються не тільки великою кількістю ознак, закономірності поведінки яких не піддається точному аналітичному опису, а й представляють собою неоднорідні (гетерогенні данні) [56, 130]. В останні роки при побудові систем розпізнавання використовується комбінація неоднорідних ознак різного характеру інформації [57, 130]. В цьому випадку, модель представлення СС образом  $P$ , з неоднорідними ознаками різного характеру інформації, представляється у вигляді:

$$P = \{X\}^d \cup \{X\}^p \cup \{X\}^l \cup \{X\}^s, \quad (1.2)$$

де  $\{X\}^d$  – множина детермінованих ознак образу  $P$ ;

$\{X\}^p$  – множина імовірнісних ознак образу  $P$ ;

$\{X\}^l$  – множина логічних ознак образу  $P$ ;

$\{X\}^s$  – множина структурних (лінгвістичних) ознак образу  $P$ .

Для здійснення класифікації образів, представлених моделлю (1.2), використовується комбіноване розпізнавання. Початок розвитку комбінованого розпізнавання був покладений Ю.І. Журавлевим [56, 57, 130-133] та продовжений в ряді робіт [134-140]. Прийняття рішення щодо класифікації при комбінованому розпізнаванні здійснюється колективами вирішальних правил (КВП), які будуються на алгоритмах обчислення оцінок (АОО) [56, 57]. Такий підхід орієнтований саме на обробку неоднорідних даних різного характеру інформації.

Основою КВП є використання комбінації окремих алгоритмів обробки для різнорідних ознак СС. Оскільки різні алгоритми розпізнавання проявляють себе по-різному на одній і тій самій вибірці об'єктів, то закономірно встає питання про синтетичне вирішальне правило, яке

адоптивно використовує сильні сторони цих алгоритмів. Такі алгоритми і отримали назву колективів вирішальних правил [110].

У синтетичному вирішальному правилі застосовується дворівнева схема розпізнавання. На першому рівні працюють приватні алгоритми розпізнавання, результати яких об'єднуються на другому рівні у блоці синтезу. Найбільш поширені способи такого об'єднання засновані на виділенні областей компетентності того або іншого приватного алгоритму. Простий спосіб знаходження областей компетентності полягає в апріорному розбитті простору ознак виходячи з професійних міркувань конкретної галузі застосування (наприклад розшарування вибірки за деякою ознакою). Тоді для кожної з виділених областей будується власний алгоритм розпізнавання. Інший спосіб базується на застосуванні формального аналізу для визначення локальних областей простору ознак як околиць об'єктів розпізнавання, для яких доведена успішність роботи якого-небудь приватного алгоритму розпізнавання.

Найзагальніший підхід до побудови блоку синтезу розглядає результуючі показники приватних алгоритмів як початкові ознаки для побудови нового узагальненого вирішального правила. Ефективними для вирішення завдання створення колективу вирішальних правил є алгоритми обчислення оцінок (АОО), покладені в основу так званого підходу алгебри, що забезпечує дослідження і конструктивний опис алгоритмів розпізнавання, в рамки якого укладаються усі існуючі типи алгоритмів.

Принцип реалізації АОО полягає у обчисленні пріоритетів (оцінок схожості), що характеризують ступінь близькості об'єкту розпізнавання до еталонних об'єктів за сукупностями ознак, що є системою підмножин заданої множини ознак.

Поєднання ознак (підпростори) прийнято називати опорними множинами або множинами часткових описів об'єктів. Вводиться поняття узагальненої близькості між розпізнаваним об'єктом і об'єктами навчальної

вибірки (з відомою класифікацією), які називають еталонними об'єктами. Ця близькість представляється комбінацією близькості об'єкту розпізнавання до еталонних об'єктів, що обчислюються на множинах часткових описів. Таким чином, АОО є розширенням класичного методу  $k$ -найближчих сусідів, в якому близькість об'єктів розглядається тільки в одному заданому просторі ознак.

Ще одним розширенням АОО є те, що в цих алгоритмах визначення схожості і відмінності об'єктів формулюється як параметричне завдання, а також виділяється етап налаштування АОО по навчальній вибірці, на якому підбираються оптимальні значення введених параметрів. Критерієм якості служить помилка розпізнавання, а параметризації підлягає буквально усе:

- правила обчислення близькості об'єктів за окремими ознаками;
- правила обчислення близькості об'єктів в підпросторах ознак;
- міра важливості того або іншого еталонного об'єкту як прецеденту;
- значущість вкладу кожної опорної множини ознак у підсумкову оцінку схожості об'єкту розпізнавання з яким-небудь діагностичним класом.

Параметри АОО задаються у вигляді значень порогів і (чи) як ваги вказаних складових.

В АОО множина об'єктів  $\{O\}$  розподіляється на класи  $C_i$ , де  $i=1, \dots, m$ , та для опису об'єктів використовуються ознаки  $x_j$ , де  $j=1, \dots, n$ . Тоді апріорна інформація представляється у вигляді таблиці, що містить описи на мові ознак  $\{x_1, \dots, x_n\}$  усіх об'єктів, що належать різним класам. Також наявний деякий об'єкт  $O_r$ , в якому сукупність ознак  $\{X\} = \{x_1, \dots, x_n\}$  приймає значення  $\{\beta_1, \dots, \beta_n\}$ . Таке співвідношення наглядно проілюстровано у таблиці 1.1.

Згідно алгоритму, порівнюється опис об'єкту розпізнавання з описами усіх об'єктів, що містяться в таблиці, і приймається рішення про те, до якого класу віднести об'єкт. Класифікація заснована на обчисленні міри схожості об'єкту розпізнавання, на об'єкти, приналежність яких до класів відома. Фактично, це розпізнавання з використанням прецедентних вибірок.

Таблиця 1.1 – Співвідношення ознак об’єктів розпізнавання

Класи	Об’єкти	Значення ознак			
		$x_1$	$x_2$	.....	$x_n$
$C_1$	$O_1$	$\alpha_1^1$	$\alpha_2^1$		$\alpha_n^1$
	$O_2$	$\alpha_1^2$	$\alpha_2^2$	...	$\alpha_n^2$
	...	...	...		...
	$O_{k_1}$	$\alpha_1^{k_1}$	$\alpha_2^{k_1}$		$\alpha_n^{k_1}$
$C_2$	$O_{k_1+1}$	$\alpha_1^{k_1+1}$	$\alpha_2^{k_1+1}$		$\alpha_n^{k_1+1}$
	$O_{k_1+2}$	$\alpha_1^{k_1+2}$	$\alpha_2^{k_1+2}$	...	$\alpha_n^{k_1+2}$
	...	...	...		...
	$O_{k_2}$	$\alpha_1^{k_2}$	$\alpha_2^{k_2}$		$\alpha_n^{k_2}$
	....	...	...	...	...
$C_m$	$O_{k_{m-1}+1}$	$\alpha_1^{k_{m-1}+1}$	$\alpha_2^{k_{m-1}+1}$		$\alpha_n^{k_{m-1}+1}$
	$O_{k_{m-1}+2}$	$\alpha_1^{k_{m-1}+2}$	$\alpha_2^{k_{m-1}+2}$	...	$\alpha_n^{k_{m-1}+2}$
	...	...	...		...
	$O_{k_m}$	$\alpha_1^{k_m}$	$\alpha_2^{k_m}$		$\alpha_n^{k_m}$
	$O_r$	$\beta_1^r$	$\beta_2^r$	...	$\beta_n^r$

Визначення схожості здійснюється у два етапи: спочатку підраховується оцінка для кожного об’єкту з таблиці, а потім отримані оцінки використовуються для отримання сумарних оцінок по кожному з класів  $C_i$ .

Досвід рішення завдань розпізнавання свідчить про те, що часто основна інформація знаходиться не в окремих ознаках, а в їх різних поєднаннях. Оскільки не завжди відомо, які саме поєднання інформативні, то в АОО міра схожості об’єктів обчислюється не послідовним зіставленням окремих ознак, а зіставленням усіх можливих поєднань ознак, що входять в опис об’єктів [100]. Цей аспект є суттєвим недоліком, що значно збільшує часову складність процесу розпізнавання та прийняття управлінських рішень.

По повній сукупності ознак  $\{X\}=\{x_1, \dots, x_n\}$  виділяється система підмножин множини ознак (система опорної множини алгоритму)  $S_1, \dots, S_q$ . В АОО, за наявності обмежень на систему опорних множин, зазвичай, розглядаються або усі підмножини множини ознак фіксованої довжини  $k$ ,  $k=2, \dots, n-1$ , або взагалі усі підмножини множини ознак.

Надалі, виділяється довільний піднабір ознак з рядків  $O_1, O_2, \dots, O_{k_1+1}, \dots, O_{k_{m-1}+1}, O_{k_m}, O_r$ . Отримані рядки можна позначити як  $\tilde{S}O_1, \tilde{S}O_2, \dots, \tilde{S}O_{k_1+1}, \dots, \tilde{S}O_{k_{m-1}+1}, \dots, \tilde{S}O_{k_m}, \tilde{S}O_r$ . Правило близькості, що дозволяє оцінити схожість рядка  $\tilde{S}O_r$ , який відповідає об'єкту розпізнавання  $O_r$  і рядка  $O_{k_{i-1}+v}$ , що відповідає довільному об'єкту початкової таблиці, полягає в наступному. Індекс  $v$ -го об'єкту класу  $C_i$ , є сумою порядкового номера останнього об'єкту попереднього класу  $k_i$  і порядкового номеру  $v$  даного об'єкту у поточному класі  $C_i$ . Природно, що  $1 \leq v \leq k_i - k_{i-1}$ . Припустимо, що «усічені» рядки містять  $\mu$  перших ознак, тобто

$$\tilde{S}O_{k_{i-1}+v} = (\alpha_1, \dots, \alpha_\mu), \quad (1.3)$$

$$\tilde{S}O_r = (\beta_1, \dots, \beta_\mu). \quad (1.4)$$

За умови визначення порогових характеристик  $\varepsilon_1, \dots, \varepsilon_\mu, \eta$ , строки  $\tilde{S}O_{k_{i-1}+v}$  і  $\tilde{S}O_r$  вважаються схожими, якщо виконується не менше ніж  $\eta$  нерівностей

$$|\alpha_j - \beta_j| \leq \varepsilon_j, \quad (1.5)$$

де  $j=1, \dots, \mu$ .



На підставі аналізу цих величин приймається рішення або про віднесення об'єкту  $O_r$  до одного з класів  $C_i$ ,  $i= 1, \dots, m$ , або про відмову від його розпізнавання. Вирішальне правило може приймати різні форми, зокрема розпізнаваний об'єкт може бути віднесений до класу, якому відповідає максимальна оцінка, або ця оцінка перевищуватиме оцінки усіх інших класів не менше ніж на певну порогову величину або значення відношення відповідної оцінки до суми оцінок для усіх інших класів буде не менше визначеного порогу.

Таким чином, інформаційна технологія комбінованого розпізнавання включає наступні етапи [100].

Етап 1. Виділяється система опорної множини алгоритму, по якій робиться аналіз розпізнаваних об'єктів.

Етап 2. Вводиться поняття близькості на множині частин описів об'єктів.

Етап 3. Задаються правила, за якими може прийматися рішення щодо класифікації.

Завдання правил здійснюється у наступній послідовності:

1) правила подібності еталонного і розпізнаваного об'єкту, що дозволяють за вчисленою оцінкою міри, вчислити величину, що називається оцінкою для пар об'єктів;

2) правила формування оцінок для кожного з класів по фіксованій опорній множині на основі оцінок для пар об'єктів;

3) правила формування сумарної оцінки для кожного з класів по усіх опорних підмножинах;

4) правила ухвалення рішення, яке на основі оцінок для класів забезпечує віднесення розпізнаваного об'єкту до одного з класів або відмовляє йому в класифікації.

Таким чином, спосіб вибору системи опорних множин, тип функції близькості, правила обчислення оцінок і вирішальні правила визначають



АОО, а завдання значень відповідних параметрів завдається конкретним алгоритмом, що використовується.

На практиці, як правило, варіюванням способом вибору опорних множин і параметрами, можна добитися визначення такого алгоритму обчислення оцінок, який забезпечує найкраще рішення задачі розпізнавання, наприклад, з точки зору мінімуму помилок і відмов від розпізнавання. Однак організація обчислювальної процедури, відповідно до приведеного алгоритму, при великій потужності системи опорних множин (як у випадку складних систем) вимагає значної кількості машинних операцій. У зв'язку з цим для обчислення оцінок, що визначають приналежність розпізнаваного об'єкту деякому класу, можна використати прості аналітичні формули, що замінюють складні перебірні процедури [141].

Практика розпізнавання показує, що в деяких випадках апріорі відомі піднабори ознак, які слід враховувати при зіставленні об'єкту розпізнавання з об'єктами навчальної таблиці (таблиця 1.1) і ці підмножини ознак не завжди співпадають з окремими, приватними, випадками; вони можуть мати різну довжину, виключати заборонені комбінації і т. п. Так наприклад, аналітичні формули можуть бути отримані для випадку довільних опорних великих кількостей [142].

У [142] показано, що складність формули обчислення оцінок при довільній  $S_A$  пропорційна складності ДНФ, що представляє характеристичну функцію системи опорної множини алгоритму. Це означає, що побудова простої формули для обчислення оцінок  $\Gamma_i(O_r)$  пов'язана із завданням мінімізації булевих функцій в класі ДНФ [143], а точніше – із завданням побудови найкоротшої ортогональної ДНФ або ДНФ, в якій кожен інтервал має невелику кількість перетинів з сусідніми. У загальному випадку завдання такого синтезу нерозв'язне і тому слід користуватися наближеними алгоритмами, що забезпечують отримання «досить простих» ортогональних ДНФ або ДНФ з невеликою кількістю взаємних перетинів інтервалів.

Зведення завдання побудови екстремальних АОО до відшукування екстремумів функції багатьох змінних було обґрунтоване Ю.І. Журавльовим [144]. Для проведення оптимізації можуть бути застосовані методи переборного типу, градієнтного типу або випадкового пошуку. Але ці методи раціонально використовувати у випадку малої кількості параметрів. У випадку же багатопараметричності та неоднородності даних, такі алгоритми або зовсім не дозволяють розв'язати задачу або потребують використання наближених алгоритмів. Крім цього, все це призводить до значного збільшення часової складності алгоритмів і ускладненню їхньої практичної реалізації, зокрема, у випадку використання комбінованих систем розпізнавання образів в СС, де швидкість рішення задачі класифікації напрямку впливає на актуальність практичного використання результату, особливо, коли використовуються системи реального часу.

Також, слід зазначити, що представлені методи достатньо вдало відбивають підходи щодо метрологічної бази обробки неоднорідних даних в СС на основі апарату розпізнавання образів. Однак, враховуючи той факт, що вирішення задач обробки інформації вимагає прийняття рішення та його представлення експерту в зрозумілій формі, така типологія не враховує одну дуже істотну характеристику, яка відбиває специфіку способу представлення знань про предметну область за допомогою якого-небудь формального алгоритму розпізнавання образів. В цьому випадку дуже важливим є спосіб представлення знань, що достатньо добре відбивається у концепції, яку запропонував Д.О. Поспелов [145-150], сутністю якої є поділ методів обробки даних за способом представлення знань при розпізнаванні. За таким принципом методи розпізнавання поділяються на інтенціональні та екстенціональні [5, 149, 150].

Інтенціональне представлення пов'язане з формуванням закономірності і зв'язків, якими пояснюється структура даних. Стосовно діагностичних завдань, така фіксація полягає у визначенні операцій над атрибутами

(ознаками) об'єктів, що призводять до необхідного діагностичного результату. Іntenсiональне представлення реалiзуються за допомогою операцiй над значеннями атрибутiв i не припускає операцiй над конкретними iнформацiйними фактами (об'єктами).

У свою чергу, екстенсiональне представлення знань пов'язане з описом i фiксацiєю конкретних об'єктiв з предметної області i реалiзуються в операцiях, елементами яких служать об'єкти як цiлiсні системи.

Іntenсiональний спосiб припускає представлення знань у виглядi схеми зв'язкiв мiж атрибутами (ознаками) об'єктiв розпiзнавання, а екстенсiональний – за допомогою конкретних фактiв (об'єктiв).

Необхiдно особливо пiдкреслити, що iснування саме цих двох (i тiльки двох) груп методiв розпiзнавання: що оперують з ознаками, i що оперують з об'єктами, глибоко закономірно. З цієї точки зору, жоден з цих пiдходiв, узятий окремо вiд iншого, не дозволяє сформуванню адекватного вiддзеркалення предметної області. Мiж цими методами iснує вiдношення додаткової в сенсi Н. Бора [151], тому перспективнi системи розпiзнавання повиннi забезпечувати реалiзацiю обох цих методiв, а не тiльки якого-небудь одного з них.

Таким чином, можна зробити висновок, що в основу концепцiї, запропонованої Д.О. Поспеловим, покладенi фундаментальнi закономірностi, що лежать в основi людського способу пiзнання взагалi, що природньо вказує на рацiональнiсть її використання не тiльки в бiологiчних, а й у технiчних системах розпiзнавання та прийняття управлiнських рiшень.

### **1.3 Інформаційна модель обробки даних в складних системах засобами розпізнавання образів**

З погляду на те, що побудувати змістовну аналітичну модель СС не представляється можливим через дефіцит апріорної інформації, залишається

варіант використання феноменологічної аналітичної моделі, наприклад регресійного типу, проте і для її побудови необхідна вихідна інформація, яка, як це впливає безпосередньо з визначення СС, може бути тільки апостеріорною. Загальноприйнятою стандартною формою подання вихідної інформації для аналізу є матрична інформаційна модель СС, яка забезпечує відображення взаємозв'язків між вхідними та вихідними параметрами СС і навколишнього середовища (факторами), з одного боку, і майбутніми станами СС – з іншого.

Відома матрична інформаційна модель СС [5], представлена в формі двовимірної матриці, стовпці якої відповідають можливим майбутнім, кінцевим станам СС (в тому числі - цільовим), а рядки – вхідним параметрам, тобто факторам. Така матрична інформаційна модель представлена в таблиці 1.2.

Таблиця 1.2 – Матрична інформаційна модель складної системи

Ознаки СС	Стани СС			Потужність ознаки
	$C_1$	...	$C_j$	
$x_1$	$I_1^1$		$I_1^j$	$\sigma_1$
...	...	...	...	...
$x_i$	$I_i^1$		$I_i^j$	$\sigma_i$

Елементами моделі є приватні критерії  $I_i^j$ , які відображають вплив  $i$ -го фактору на переведення СС в  $j$ -й стан.

Обрання конкретного виду приватних критеріїв  $I_i^j$  є одним з основних моментів у побудові інформаційної моделі СС. В якості  $I_i^j$  в [5] використовується семантична міра доцільності інформації, введена у 1960 році А.А. Харкевичем [54] на основі апостеріорного підходу. Сутність такого

підходу полягає в тому, що кількість інформації оцінюється за зміною ступеню доцільності поведінки системи в результаті сповіщення їй цієї інформації [145]. Фактично, такий процес і має місце у випадку аналізу СС за станом, а не за параметрами. У випадку здійснення керуванням станом СС (переведення СС з поточного стану до апріорно визначеного цільового), здійснюється керуючий вплив, який характеризується передаванням відповідної сукупності інформації на вхід СС. Керуючий вплив, в цьому випадку, є, насамперед, інформаційним впливом, тобто деяким повідомленням.

Поняття «доцільності поведінки» тісно пов'язано з поняттям «цільовий стан». Цільовим станом називається кінцевий стан СС, в який вона повинна перейти в результаті отримання управляючого впливу (повідомлення) [146].

Поведінка СС вважається «доцільною», якщо наближує її до цільового стану, і «недоцільною» – в іншому випадку. Таким чином, керуючий вплив може змінювати доцільність поведінки системи, якщо змінює ймовірність її переходу в цільовий стан. Множина цільових станів є підмножиною кінцевих станів СС. Для розглянутої інформаційної моделі, міра Харкевича має вигляд [5]:

$$I_i^j = \text{Log}_2(P_i^j / P^j). \quad (1.8)$$

Доцільність вибору саме цієї міри обумовлена самим смислом величин  $P_i^j$  і  $P^j$ , які позначають відповідно ймовірність переходу СС в  $j$ -й кінцевий стан під впливом  $i$ -го фактору та ймовірність випадкового (спонтанного, самовільного) переходу СС в той самий стан.

Така ж сама кількість інформації  $I_i^j$  може бути виражена в іншій еквівалентній формі:

$$I_i^j = \text{Log}_2(P_i^j / P_i), \quad (1.9)$$

де  $P_i^j$  і  $P_i$  позначають ймовірність класифікації стану СС до класу  $C_j$  в результаті співставлення  $i$ -ї ознаки і ймовірність такої ж класифікації на основі інших ознак, відмінних від  $i$ -ї.

Вирази (1.8) і (1.9) безпосередньо визначають, яка кількість інформації  $I_i^j$  отримується про стан СС, що настає: {СС перейде до  $j$ -го стану}, з повідомлення: {на СС впливає  $i$ -а ознака} [5]. Коли кількість інформації  $I_i^j > 0$  –  $i$ -а ознака сприяє переходу СС до  $j$ -го стану; коли  $I_i^j < 0$  – перешкоджає цьому переходу; коли  $I_i^j = 0$  – ніяк не впливає на стан.

#### 1.4 Методи аналізу інформативності неоднорідних даних

Отримання достовірного результату розпізнавання за мінімальних витрат часу можливе за рахунок проведення попереднього аналізу даних, що надаються для обробки інформації [133, 136]. Метою такого аналізу є оцінка основних інформативних характеристик (ознак) СС, що надаються для співставлення, та подальша селекція найбільш інформативних з них для обробки та прийняття рішення.

В прикладних застосуваннях, початкові описи об'єктів розпізнавання можуть містити усі доступні спостереженню (або вимірюванню чи обчисленню) характеристики або параметри, тому в описі об'єктів, як правило, беруть участь кілька десятків (сотень) величин. Однак велика кількість прецедентів часто вимагає дорогих і трудомістких підготовчих робіт по визначенню алфавіту класів, що значно ускладнює підготовку та впровадження таких інформаційних технологій обробки даних. У випадку неоднорідних даних, така задача стає ще актуальнішою, та вимагає

ретельнішого вибору сукупності найбільш інформативних прецедентів для отримання достовірного рішення щодо класифікації.

При прийнятті рішення про вибір класу, якому належить аналізований об'єкт, виникає проблема його оцінки за кількома ознаками і коректного обліку цих оцінок при їх узагальненні або спільному розгляді на етапі прийняття остаточного рішення [65, 133, 136]. Ця проблема, в свою чергу, поділяється на дві складові: встановлення виду вирішального правила (відповідно до наявного апарату КВП) та визначення вагових коефіцієнтів – коефіцієнтів приватних, характеристичних ознак [65, 133], що відображають властивості об'єктів. Аналіз робіт з цього напрямку вказує на актуальність дискретного підходу до оцінки інформативності різнорідних ознак і статистичних підходів, які представлені в [152-154]. Застосовність імовірнісних та статистичних підходів до оцінювання інформативності ознак при прийнятті рішень можливе в разі пропусків даних; великій розмірності завдання і ніяк не пов'язана з тим, якою шкалою вимірюються значення оцінюваних ознак.

Слід зазначити, що складність обчислювальних операцій також залежить від багатокomпонентності класифікатора. Так, якщо мається єдиний механізм класифікації, матриця розрізень, відповідно до представленої раніше матричної інформаційної моделі, вироджується до один стовбця, що відповідає традиційному поданню знань в задачах розпізнавання образів [133, 136, 152-154]

Вимоги до попередньої обробки даних для оцінки характеру розподілу в досліджуваній виборці також вказуються в ряді робіт [4, 5, 12, 23, 24, 133, 136, 152-154]. Основні операції попередньої обробки в таких інформаційних технологіях зводяться до наступного переліку:

– перевірка однорідності груп ознак, що спостерігаються, в тому числі контрольних, що може бути проведено або експертним шляхом, або

методами багатовимірної статистики (наприклад, за допомогою кластерного аналізу);

- нормалізація змінних, тобто усунення аномалій показників в матриці даних;
- зниження розмірності простору ознак (здійснюється формальними методами шляхом оцінки інформативності);
- стандартизований опис ознак;
- побудова класифікаційних шкал ознак, тобто процедура ідентифікації та встановлення меж розподілу даних та подання інформації в квантовій формі (кожному значенню ознаки відповідає певне кодове число).

З метою визначення підходів до математичної обробки та оцінювання інформативності неоднорідних даних, а також оцінювання імовірнісного розподілу мінімально необхідного числа інформативних ознак, що забезпечують апріорно заданий рівень достовірності розпізнавання (відповідно до основної мети дисертаційного дослідження), представляються п'ять підходів до відбору інформативних ознак: на основі дискретних методів пошуку в навчальній вибірці інформативної зони [152]; на основі методів кластеризації [65, 152, 154]; на основі припущення про нормальності розподілів об'єктів у кластерах [152-154]; на основі теоретико інформаційного поняття ентропії [12, 23, 60]; на основі непараметричних оцінок щільності [152, 154].

#### *1.4.1 Метод визначення інформативності ознак на основі пошуку в навчальній вибірці інформативної зони*

В [152, 154] описується методика попереднього аналізу навчальної інформації, заснована на перебуванні в навчальній вибірці інформативної зони (інформативними вважаються такі фрагменти описів, які дозволяють розрізнити об'єкти із різних класів або відрізнити даний об'єкт від усіх об'єктів, що не належать до того ж класу і типових для своїх класів об'єктів



(найбільш репрезентативних). Інформативна зона виділяється на основі оцінки типовості значень кожної ознаки.

Отримали найбільшу поширеність два способи виділення типових об'єктів [152-154]:

1) типовими вважаються об'єкти, описи яких складаються з типових значень ознак;

2) типовими є об'єкти, які правильно розпізнаються на ковзному контролі. Так, до нетипових (неінформативних) ознак відносяться ознаки, що приймають багато значень або значення яких рідко зустрічаються в усіх класах. Про такі ознаки не можна сказати, що вони мають високу інформативність. Кожен об'єкт навчальної вибірки, що лежить на границі між класами, також не є інформативним для свого класу, оскільки його опис схожий на опис об'єктів з інших класів.

Пошук інформативних фрагментів заснований на використанні апарату дискретної математики, зокрема булевої алгебри, теорії ДНФ, теорії покриттів булевих і цілочисельних матриць, що вже розглядалося раніше. Вирішенню цього завдання присвячені роботи [56, 57, 130, 133, 143, 144] Так, в розглянутих вище, АОО, знаходяться оцінки ансамблів ознак, які є узагальненнями коефіцієнтів інформативності, розглянутих у [153].

Для тестових алгоритмів розпізнавання образів в [153] вводиться показник інформативності ознаки – його інформаційна вага, як відношення числа входжень ознаки у всі остаточні класифікації до числа всіх остаточних класифікацій. Так в розглянутій матричній інформаційній моделі СС (таблиця 1.2) значення  $\sigma_i$  виражають числову оцінку інформативності ознаки (ваговий коефіцієнт ознаки).

Перелічені в [152, 153] методи оцінки інформативності ознак показують хороші результати в прикладних задачах. При цьому, під якістю розпізнавання розуміється якість алгоритму поза навчальної вибірки (здатність алгоритму до узагальнення або екстраполяції), яке оцінюється як

відсоток правильно розпізнаних об'єктів при проведенні процедури ковзного контролю [154].

При цьому слід зазначити, що застосування дискретного підходу виявляється у багатьох випадках складним у силу чисто обчислювальних труднощів перебірного характеру, що виникають на етапі пошуку інформативних фрагментів описів об'єктів. Експоненціальне зростання числа фрагментів при зростанні розмірності описів, рішення проблеми тільки за рахунок підвищення продуктивності обчислювальної техніки в цьому випадку не можна назвати раціональним, тому дискретні методи важко застосовні для попереднього аналізу великого обсягу неоднорідних даних.

#### *1.4.2 Метод визначення інформативності ознак на основі кластеризації*

Методи відбору ознак на основі кластеризації [65, 152, 154] дозволяють розбити вибірку ознак на кластери, що складаються зі схожих ознак, і виділити у кожній групі по одному найбільш типовому представнику (репрезентативна ознака). Для застосування методів кластеризації необхідно ввести метрику на множині ознак  $\{X\}$ . Розглянемо можливі метрики на ознаках.

Нехай  $x_1$  та  $x_2$  – дві ознаки з  $\{X\}$ . вибіркові вектори значень ознак для об'єктів вибірки  $O$  об'єму  $n$ . Позначимо, відповідно, через  $x_1=(x_1(O_1), x_1(O_2) \dots x_1(O_n))$  і  $x_2=(x_2(O_1), x_2(O_2) \dots x_2(O_n))$ . Розглянемо три варіанти визначення функції відстані  $d(x_1, x_2)$ .

$$d(x_1, x_2) = 1 - |r(x_1, x_2)|, \quad (1.10)$$

де

$$r(x_1, x_2) = \sum_{i=1}^n x_1'(O_i) \cdot x_2'(O_i), \quad (1.11)$$

де  $x_1'(O_i)$  і  $x_2'(O_i)$  – нормовані і центровані значення ознак  $x_1$  і  $x_2$ .

Відстань  $r(x_1, x_2) = 0$  тільки коли ознаки пов'язані лінійною залежністю.

Інший варіант – метрика Кендала-Кемени [154], застосована для порядкових ознак. Вона визначається як частка пар об'єктів  $O_i, O_j$  з різними порядковими відносинами між значеннями ознак  $x_1$  і  $x_2$ :

$$d(x_1, x_2) = \frac{1}{2C_n^2} \sum_{i=1}^{n-1} \sum_{j=i}^n \left| \text{sign}(x_1(O_i) - x_1(O_j)) - \text{sign}(x_2(O_i) - x_2(O_j)) \right|. \quad (1.12)$$

Ця відстань дорівнює нулю коли ознаки пов'язані монотонною залежністю, тобто існує монотонна функція  $f$ , така що  $x_1(O_j) = f(x_2(O_j))$  ( $j=1, \dots, n$ ).

Метрика Хеммінга, застосовна для номінальних ознак з однаковими множинами припустимих значень  $D_{x_1} = D_{x_2}$ , яка приймає нульове значення коли вектори  $x_1$  і  $x_2$  збігаються:

$$d(x_1, x_2) = \sum_{i=1}^{n-1} [x_1(O_i) \neq x_2(O_i)]. \quad (1.13)$$

Для номінальних ознак з різними множинами значень шукається відповідність  $\sigma: D_{x_1} \rightarrow D_{x_2}$ , при якій відстань Хеммінга мінімальна (без обмеження спільності передбачається, що  $|D_{x_1}| \geq |D_{x_2}|$ ).

Якщо необхідно знайти метрику між неоднорідними ознаками, обмірюваними в різних шкалах, то вони спочатку приводяться до однієї загальної шкали [152, 154].

Слід зазначити такі особливості застосування методів кластеризації. По-перше, якщо у вихідній множині ознак будуть присутні неінформативні ознаки, то можуть з'явитися кластери, що цілком складаються з неінформативних ознак. По-друге, вищезгадані метрики використовують інформацію про попарну подібність між ознаками, але не вирішують проблему мультилінійності ознак (набір попарно некорельованих ознак може виявитися лінійно залежним) [154]. По-третє, класифікація за принципом мінімальної відстані придатна, якщо кластери, відповідні різним класам, рознесені досить далеко один від одного (добра роз'єднувальність класів), тобто значення ознак, що описують об'єкти з різних образів, досить далекі одне від іншого.

#### *1.4.3 Метод визначення інформативності ознак, заснований на статистичному кластер-алгоритмі*

Алгоритм побудови кластерів в припущенні про нормальність розподілів об'єктів у кластерах (що є природним в багатьох прикладних застосуваннях) використовує кластер критерій, запропонований у [152], і властивість оцінок Мешалкіна [155] оцінювати параметри розподілу кластера, найбільшого за кількістю точок. Слід зазначити, що умови критерію К. Пірсона [156] перевірки гіпотези збігу центрів двох багатовимірних нормальних сукупностей з відомими коваріаціями, що різняться зсувом, не є придатними для кластер-аналізу у зв'язку з жорсткістю використовуваних припущень [154].

Ідея побудови кластер-алгоритму, що розбиває вибірку ознак на кластери, які складаються зі схожих ознак, полягає в наступному. Кластер-критерієм на рівні значущості  $\alpha$  перевіряється гіпотеза про приналежність вибірки одній функції нормального розподілу. Якщо гіпотеза приймається, то

вибірка з імовірністю  $1-\alpha$  містить один кластер. Якщо гіпотеза відкидається, то самий великий кластер вирізається еліпсоїдом на рівні значущості  $\alpha$  і до решти точок застосовується описана процедура в циклі. Процедура виділення кластерів закінчується, коли всі точки розподілені за кластерами або їх кількість невелика (тобто припускаються  $\alpha$  залишки від перевірки гіпотез).

На відміну від більшості алгоритмів кластерного аналізу даний статистичний кластер-алгоритм не вимагає апріорного визначення кількості класів, на яке треба поділити сукупність точок спостереження, і вирішує завдання зі складними накладеннями і перетинами кількох класів (випадок, коли ознаки можуть мати однакові значення для об'єктів з різних образів при поганій роз'єднувальності класів).

#### *1.4.4 Методи визначення інформативності ознак на основі ентропії*

*Метод на основі формалізму мультимножин і теоретико-інформаційному понятті ентропії.* Заснований на формалізмі мультимножин [154] метод визначення вагових коефіцієнтів ознак при прийнятті рішення в СС з використанням матричної інформаційної моделі враховує внесок ознак в розпізнавальну здатність тесту з урахуванням їх взаємозалежності і базується на представленні сукупністю всіх помітних пар об'єктів з різних образів для кожної ознаки  $x_i$ ,  $i = \overline{1, M}$ , у вигляді мультимножини [154].

Співставимо ознаці  $x_i$  сукупність  $i$ -ї ознаки пар об'єктів з різних образів. Тоді, можна стверджувати, що ознака  $x_i$  породжує мультимножину  $\{k_{P_i}(u) \square u \mid u \in U, k_{P_i}(u) \in X_+\}$ . Таким чином, представлені ознаки [154] є множинами з повторюваними елементами  $u \in U$  (мультимножинами), при цьому потужність мультимножини визначається як загальне число його елементів, де множина  $U$  – домен або універсальна множина [154], з якої отримуються можливі значення мультимножини, що породжені відповідними ознаками. Слід зауважити, що елементи домену  $u$  є

позначеннями пар об'єктів з різних образів, тому елементи  $(i-j)$  і  $(j-i)$  (пари номерів об'єктів) вважаються еквівалентними ( $i \neq j, i, j \in \{1, \dots, K\}$ ).

Позначимо через  $P(v/i)$  ймовірність прояву  $v$ -го елемента домену для  $i$ -го образу, тоді ймовірність  $v$ -го елемента домену для всіх  $K$  образів дорівнює [154]

$$P_v = \sum_{i=1}^K P(v/i). \quad (1.14)$$

З урахуванням умови нормування, частка  $i$ -го образу в цій сумі

$$r_i = P(v/i) / P(v) \quad (1.15)$$

і ентропія  $v$ -го елемента домену виражається наступним значенням [154]:

$$H_v = - \sum_{i=1}^K r_i \cdot \log_2 r_i. \quad (1.16)$$

З властивості адитивності ентропії випливає, що загальна невизначеність при розпізнаванні образів за ознакою  $x_i$  має вид [154]

$$H_{x_i} = - \sum_{v=1}^{N_U} H_v \cdot P_v. \quad (1.17)$$

Кількість інформації, одержуваної в результаті вимірювання ознаки  $x_i$ , дорівнює  $H_0 - H_{x_i}$ , де  $H_0$  – початкова невизначеність відносно образів  $H_0 = \log_2 K$ .

При якісному аналізі емпіричних даних, роль ймовірностей  $P(v/i)$  (ймовірність прояви  $v$ -го елементу домену для  $i$ -го образу) грають їх вибіркові оцінки (частоти, частки).

В реальних таблицях даних залежність між ознаками спостерігається дуже часто. А якщо ознаки залежні, то, при виборі найбільш інформативної підсистеми ознак, оцінками їх індивідуальної інформативності керуватися недоцільно.

*Метод аналізу взаємозв'язку ознак на основі ентропії.* Трактуювання статистичного зв'язку між ознаками  $x_i$  і  $x_j$  зводиться до оцінки кількості інформації  $I(x_i, x_j) = H(x_j) - H(x_j/x_i)$ , яке зменшує невизначеність того, яке значення прийме  $x_j$ , якщо відомо значення  $x_i$ . В якості оцінки міри спряженості приймається частка спряженого розмаїття величини  $x_j$ , включеної до системи  $(x_i, x_j)$ , порівняно з розмаїттям (ентропією)  $x_j$ , що розглядається окремо:  $I(x_i, x_j) / H(x_j)$ , де в явній формі величина  $I(x_i, x_j)$  визначається по формулі [154]

$$\begin{aligned} I(x_i, x_j) &= - \sum_{s=1}^l \sum_{k=1}^l p(x_s, y_k) \log_2 \frac{p(x_s | y_k)}{p(x_s)} = \\ &= - \sum_{s=1}^l \sum_{k=1}^l p(x_s, y_k) \log_2 \frac{p(x_s, y_k)}{p(x_s) \cdot p(y_k)}, \end{aligned} \quad (1.18)$$

де  $l$  – кількість градацій ознак.

Очевидно, що в силу симетричності вираження  $I(x_i, x_j)$  щодо  $x_i, x_j$ , значення кількості інформації, що міститься в ознаці  $x_i$  про ознаку  $x_j$  та заключеного в ознаці  $x_j$  про ознаку  $x_i$ , рівні, тобто середня кількість інформації є мірою відповідності двох ознак, характеристикою їхнього зв'язку, а не характеристикою однієї з ознак.

Взаємна інформація  $I(x_i, x_j)$  стає дорівнювати нулю лише тоді, коли ознаки  $x_i$  і  $x_j$  статистично незалежні. Максимальне значення взаємної

інформації, яка дорівнює  $H(x_i)$  або  $H(x_j)$ , відповідає функціональній залежності (повного зв'язку) ознак  $x_i$  і  $x_j$ , коли кожному значенню ознаки  $x_i$  відповідає єдине значення ознаки  $x_j$ . Слід зазначити, що використання коефіцієнта  $I(x_i, x_j) / H(x_j)$  для поділу ознак на "залежні" і "визначальні", взагалі, не має достатнього обґрунтування.

Побудова теоретико-інформаційної міри зв'язку може здійснюватися не тільки на основі ентропії. Важливим прикладом такої міри є коефіцієнт Валліса, що реалізує принцип «пропорційної предикації» [154], згідно якого, мірою зв'язку виступає відносне зменшення ймовірності помилки передбачення ознаки  $x_j$  при знанні ознаки  $x_i$  у порівнянні з ймовірністю помилки прогнозу  $x_j$  без знання  $x_i$ . Інтерпретація коефіцієнта Валліса вельми проста: якщо, наприклад, його значення дорівнює 0.5, то знання  $x_i$  зменшує число помилок прогнозування значення  $x_j$  у двічі. Однак, в ряді робіт [152-154, 156] вказується на доцільність використання не коефіцієнту Валліса, а інформаційних мір зв'язку, значимість яких може бути встановлена, наприклад, відповідно до критерія Пірсона [154, 156].

#### *1.4.5 Метод визначення інформативності неоднорідних ознак на основі непараметричних оцінок щільності*

Методи оцінювання щільності ймовірності у просторах загального вигляду запропоновані і спочатку вивчені в [157]. Зокрема, в задачах класифікації об'єктів з недетрмінованими ознаками, пропонується використовувати непараметричні ядерні оцінки щільності типу Парзена-Розенблата (цей вид оцінок і його назва вводиться в [157] по імені американських вчених Парзена і Розенблата, які раніше використовували подібні статистики у разі  $X = R^1$ ,  $\rho(x_i, x) = |x - ix|$ :

$$f_n(x) = \frac{1}{v_n(h_n, x)} \sum_{1 \leq i \leq n} K \frac{\rho(x_i, x)}{h_n}, \quad (1.19)$$



де  $K: R_+^1 \rightarrow R^1$  – ядерна функція;

$\{x_1, x_2, \dots, x_n\} \in X$  – вибірка, за якою оцінюється щільність;

$\rho(x_i, x)$  – відстань між елементом вибірки  $x_i$  та точкою  $x$ , в якій оцінюється щільність; послідовність показників розмитості  $\{h_n\}$  така, що при  $n \rightarrow \infty$   $h_n \rightarrow 0$  і  $nh_n \rightarrow \infty$ ;

$v_n(h_n, x)$  – нормуючий множник, що забезпечує виконання умови

$$\int_X f_n(x) d\mu = 1. \quad (1.20)$$

В [157] показано, що оцінка щільності типу Парзена-Розенבלата  $f_n(x)$  є самостійною  $f_n(x) \rightarrow f(x)$  за ймовірністю при  $n \rightarrow \infty$  та оцінена середньоквадратична швидкість збіжності ядерних оцінок  $\alpha_n = E(f_n(x) \rightarrow f(x))^2$ .

Оскільки простір різнотипних ознак – це декартовий добуток безперервних та дискретних просторів, то, у випадку фіксованої кількості градацій якісних ознак, непараметричну оцінку щільності можна звести до добутку частоти потрапляння в точку в просторі якісних ознак на класичну оцінку Парзена-Розенבלата в просторі кількісних ознак [154, 157]. Відстань  $\rho(x, y)$  можна розглядати як суму Евклідової відстані  $\rho_1$  між кількісними факторами, відстані  $\rho_2$  між номінальними ознаками ( $\rho_2(x, y) = 0$ , якщо  $x \neq y$  і  $\rho_2(x, y) = 1$ , якщо  $x = y$ ) і відстані  $\rho_3$  між порядковими змінними (якщо  $x$  і  $y$  – номери градацій, то  $\rho_3(x, y) = |x - y|$ ).

*1.4.6 Оцінка мінімальної кількості ознак, що забезпечують апріорно заданий рівень достовірності розпізнавання образів*

Очевидно, що інформативність ознак СС, що надаються для класифікації та прийняття рішення значною мірою впливає на якість здійснення самого інформаційного процесу обробки даних. Основу підходу

до оцінки якості процедури розпізнавання складає задана інформація про частоти проявлення ознак для кожного класу з наявного алфавіту.

Для отримання двосторонніх оцінок для ймовірності мінімального числа ознак, що забезпечують апіорно заданий рівень достовірності розпізнавання, доцільне використання ймовірнісної моделі перетину випадковим процесом певного рівня [154].

Якщо  $\xi_i = \xi_i(m)$  – випадкова величина, що означає кількість коректних класифікацій об'єктів на навчальній вибірці ознакою  $x_i$ ,  $i = \overline{1, n}$ , де  $n$  – кількість наявних ознак, а  $m$  – кількість класифікованих об'єктів). Випадкова величина

$$\Phi_n = \frac{\sum_{i=1}^n \xi_i}{n \cdot m} \quad (1.21)$$

представляє собою частку коректно розпізнаних об'єктів за  $n$  першими ознаками, ранжируваною будь-яким чином, наприклад, у порядку спадання вагових коефіцієнтів ознак, обчислених за навчальною вибіркою, і є числовою інтерпретацією міри достовірності розпізнавання (якості роботи алгоритму розпізнавання [154]).

Для моменту зупинки випадкового процесу  $\Phi_n$  вводиться випадкова величина [154]):

$$\tau = \inf\{n \geq 1 : \Phi_n > q_0\}, \quad (1.22)$$

де  $q_0$  – заданий рівень достовірності процедури розпізнавання образів.

Якщо відомий ймовірнісний розподіл ознак, то можна оцінити ймовірнісний розподіл  $R_n = P(\tau > n)$  величини  $\tau$ , або мінімально необхідного числа інформативних ознак, що забезпечує апіорно заданий рівень

достовірності розпізнавання. Визначений розподіл величини  $\tau$ , що є пороговим значенням якості процедури розпізнавання, дозволить:

- оцінити необхідну кількість ознак  $n$  для досягнення апріорно заданого рівня достовірності розпізнавання;
- оптимізувати процедуру розпізнавання за параметрами  $n$ ,  $m$ , а також за їхніми ваговими коефіцієнтами. Так, у багатьох прикладних застосуваннях актуальне вирішення питання про співвідношення довжини вибірки даних  $m$  і кількості інформаційних ознак  $n$ , які є конкуруючими.

Таким чином, вибір інформативних ознак є важливою операцією, оскільки для вирішення будь-якої класифікаційної задачі доцільне здійснення селекції тих ознак, які мають максимальну інформативність. Якщо мається кілька різних методів (алгоритмів) обробки даних, то раціональним є вибір того варіанту, який володіє найкращою узагальнюючою здатністю [154], тобто кращою результативністю на незалежній вибірці.

Слід зазначити, що при збільшенні числа використовуваних ознак середня помилка на навчальній вибірці, як правило, монотонно зменшується. При цьому, середня помилка на незалежних контрольних даних спочатку зменшується, потім проходить через точку мінімуму і далі тільки зростає. Таке явище є фактично, перенавчанням. Однак, отримання теоретичних оцінок ймовірності перенавчання (відхилення частот помилок алгоритму на контрольній та навчальній вибірках) поки залишається відкритою проблемою.

За оцінку інформативності ознаки в розпізнаванні образів прийнято вважати відношення якості розпізнавання контрольної вибірки у повному просторі ознак до якості розпізнавання, що проводиться без урахування оцінюваної ознаки [152-154]. Якщо відомий вид функції розподілу, інформативність ознак може оцінюватися через середньоквадратичні відхилення [154].

Крім цього, вибір методу аналізу доцільно здійснювати пропорційно ефективності відповідного методу на контрольній вибірці.

### 1.5 Постановка завдань досліджень і розробок

Найбільш значущим критерієм ефективності процесу розпізнавання образів є достовірність результатів, що отримуються, значення якої в ідеалі повинно прагнути 100%. Однак, на практиці, зазвичай, існують обмеження, які не дозволяють збільшувати обсяг даних, що обробляються, без збільшення часової складності процесу розпізнавання. Таким чином, є достатньо важливим знаходження оптимуму в процесі розпізнавання, який забезпечує бажаний (прийнятний для практичного застосування) рівень достовірності за мінімальних витрат часу (мінімальної часової складності процесу класифікації).

На достовірність результатів класифікації впливають репрезентативність даних (їхня інформативність), кількість ознак, що надаються для співставлення на отримання рішення; метод розпізнавання.

Існує статистична достовірність [100, 158], що розраховується за статистичними результатами серії класифікацій, і одинична достовірність, що розраховується для одного конкретного випадку.

В роботі [159] показано, що створення інформаційної технології розпізнавання (особливо в реальному часі), необхідна не тільки ідентифікація невідомих образів, але й оцінка достовірності та часової складності процесу обробки даних. Складність  $C$  визначається як ймовірність помилки розпізнавання образів  $\varepsilon$  [160], тобто  $C$  дорівнює  $\varepsilon$  (див. рисунок 1.5). На цьому рисунку області помилкового розпізнавання позначені як  $\varepsilon$ .

Складність розпізнавання визначається якістю ознак та їхньою кількістю, як це проілюстровано на рисунку 1.5. Якщо виходити з того, що

ознаки незалежні, то, при досить великій кількості циклів обробки, складність розпізнавання від декількох ознак буде визначатися виразом:

$$C(x_1, x_2, \dots, x_n) = \prod_{j=1}^n C(x_j), \quad (1.23)$$

де  $x_j$  – ознаки об'єкта розпізнавання;

$n$  – загальна кількість ознак об'єкта розпізнавання.

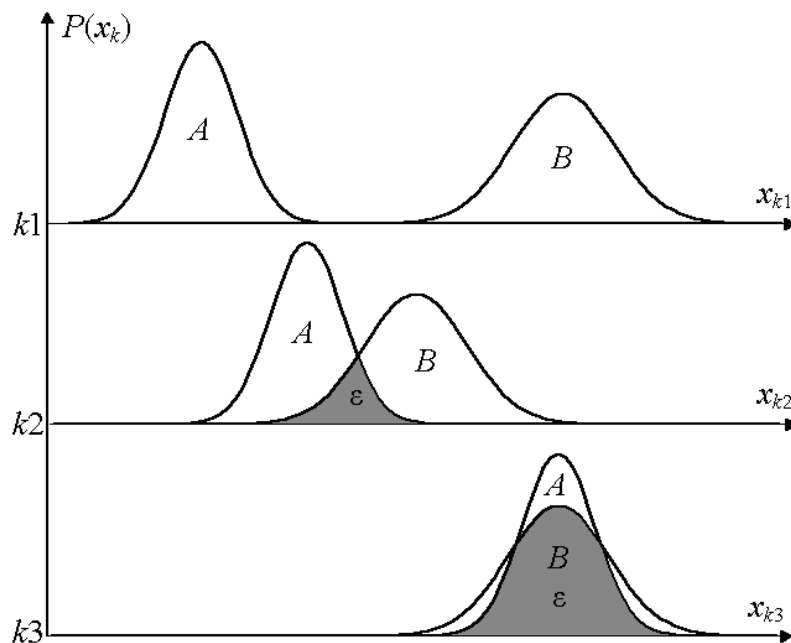


Рисунок 1.5 – Графіки розподілів щільності ймовірності значень ознак об'єкту розпізнавання

Для кожного методу розпізнавання, статистична достовірність визначається складністю розпізнавання та репрезентативністю даних:

$$D = F(C, N), \quad (1.24)$$

де  $N$  — обсяг вибірки, що характеризує репрезентативність.

На практиці, достатньо широко використовується визначення статистичної достовірності розпізнавання за виразом:

$$D = \frac{N_{cor}}{N_{com}}, \quad (1.25)$$

де  $N_{cor}$  – кількість правильних процесів розпізнавання;

$N_{com}$  – загальна кількість процесів розпізнавання.

Слід зазначити, що при великому значенні  $N$  одинична достовірність наближається до статистичної:

$$\lim_{N \rightarrow \infty} \{D_s\} = D, \quad (1.26)$$

де  $D_s$  – одинична достовірність результату розпізнавання образів.

Однак для малих значень  $N$  величина  $D_s$  може сильно відрізнятись від  $D$ . Таким чином, статистична достовірність точніше відбиває якість процесу обробки даних при більшій кількості процесів розпізнавання. В цьому аспекті є раціональним створення таких методів розпізнавання, які дозволяли би примати рішення на основі не одного, а декількох результатів розпізнавання. При чому, чим більша кількість таких результатів вказує на одну й ту саму класифікацію, тим вища статистична достовірність такого результату розпізнавання.

Проведений аналіз існуючих методів обробки даних на основі розпізнавання образів вказує на достатньо високу залежність часової складності цих методів від кількості даних на основі яких формується рішення щодо класифікації стану СС. У випадку неоднорідності даних, широкого поширення набули методи комбінованого розпізнавання, такі як АОО та КВП.

На відміну від інших методів, АОО принципово по-новому оперують описами об'єктів. Для таких алгоритмів, об'єкти існують одночасно в самих різних підпросторах простору ознак. Клас АОО доводить ідею використання ознак до логічного кінця: оскільки не завжди відомо, які поєднання ознак найбільш інформативні, то в АОО міра схожості об'єктів обчислюється при зіставленні усіх можливих або визначених поєднань ознак, що входять до опису об'єктів розпізнавання [124].

Теоретичні можливості АОО принаймні не нижче можливостей будь-якого іншого алгоритму розпізнавання образів, оскільки за допомогою АОО можуть бути реалізовані усі можливі операції з об'єктами розпізнавання.

Однак, розширення потенційних можливостей натрапляє на великі труднощі їхньої практичної реалізації, особливо на етапі побудови (налаштування) алгоритмів цього типу – для високорозмірних завдань виникають складні теоретичні питання і проблеми, пов'язані з організацією ефективного обчислювального процесу. Якщо спробувати використовувати можливості АОО в повному обсязі, вказані труднощі зростають багаторазово.

З іншого боку встановлено, що, відповідно до глибокої статистичної теорії Вапніка-Червоненкіска, збільшення кількості прецедентів призводить до збільшення імовірності отримання достовірної класифікації стану СС. Так, згідно асимптотичного обґрунтування методів розпізнавання, у випадку імовірності достовірної класифікації, що прагне одиниці, довірчий інтервал прагне нуля [109]. Тобто, збільшення кількості інформаційних каналів та прецедентів підвищує імовірність достовірного результату розпізнавання. Однак, в цьому випадку, збільшується часова складність аналізу великого масиву прецедентів, що негативно впливає на оперативність прийняття рішень щодо класифікації.

У випадку неоднорідності даних та динамічної зміни умов спостереження за характеристиками СС, отримання достовірного результату навіть за великою кількістю прецедентів, ускладнюється наявністю перешкод

і викривлень вхідного потоку даних. В цьому випадку, за поточних умов спостереження стану СС, частина інформаційних ознак може бути малоінформативною для прийняття достовірного рішення. Тоді наявність великої кількості даних, що підлягають співставленню не дозволяє однозначно підвищити достовірність, якщо інформативність цих даних сумнівна.

З точки зору теорії інформації, процес обробки даних та прийняття рішення щодо класифікації можна представити як канал передачі даних. Якщо вважати, що ідеальний канал дозволяє передати інформацію зі входу на вихід без втрат і викривлень, то трансформація вхідного потоку інформації  $I(t)$  у вихідний канал  $I_d(t)$  можна представити як внесення деякої перешкоди  $\xi$ . Тоді залежність між вхідним і вихідним інформаційними потоками може бути представлена у детермінованому вигляді як

$$I_d(t) = \xi \cdot I(t). \quad (1.27)$$

В залежності (1.27) характеристика перешкоди представлена скалярною величиною, однак таке припущення можливе тільки для нестохастичних систем. У випадку стохастичності, характеристика внесеної перешкоди буде мати імовірнісний характер і повинна бути представлена функціоналом. Зміна характеристики перешкоди у часі пов'язана з внутрішньою організацією СС, неструктурованості завдання на цільовий стан, а також відсутністю повного опису і врахування факторів, що визначають стан системи (система з елементами невизначеності). З урахуванням вищенаведеного, залежність (1.27) прийме вигляд:

$$I_d(t) = \xi(t) \cdot I(t). \quad (1.28)$$



З урахуванням того, що у більшості випадків у СС немає можливості аналітичного апріорного опису характеристики  $\zeta(t)$ , задача її ідентифікації може бути виконана на основі додання апостеріорної інформації і механізмів розпізнавання образів [5]. Цей факт додатково підтверджує, що у випадку СС доцільне використання концепції аналізу і класифікації стану системи, так як неповна формалізація задачі не дозволяє здійснювати аналіз за параметрами.

Залежність (1.28) ілюструє процес перетворення вхідного потоку даних, однак, в даному випадку, не враховується кінцевий час аналізу і класифікації, так як часова складність процесів обробки даних не дорівнює нуля. Кінцевий час протікання інформаційних процесів призводить до того, що вихідний результат, який базується на апостеріорній інформації  $I(t)$ , буде отриманий тільки у момент часу  $t+\Delta t$ , де  $\Delta t$  характеризує часову складність процесу обробки даних СС. Характеристика  $\Delta t$  є інтегральною та складається з часової складності кожного етапу співставлення ознак об'єкту розпізнавання:

$$\Delta t = \sum_{i=0}^{n-1} (t_{i+1} - t_i), \quad (1.29)$$

де  $n$  – кількість етапів порівняння ознак СС для отримання остаточного результату;

$t_0$  – відповідає моменту часу надходження вхідного потоку інформації.

Тоді, з урахуванням (1.29) залежність (1.28) прийме наступний вигляд:

$$I_d(t + \Delta t) = \xi(t) \cdot I(t). \quad (1.30)$$

З залежності (1.30) витікає, що з урахуванням часової складності процесу співставлення ознак СС буде спостерігатися невідповідність миттєвих характеристик вихідного і вхідного потоків. Тобто, у момент часу

$t+\Delta t$  вихідний потік фактично буде характеризувати перетворений стан вхідного потоку для моменту часу  $t$ . Таким чином, з урахуванням кінцевої часової складності процесу обробки даних та співставлення ознак СС, у будь-який момент часу вихідний потік буде не актуально характеризувати вхідні дані, що негативно відбивається оперативності прийняття рішення. Цей аспект особливо важливий для випадків, коли прийняття рішення повинно здійснюватися в режимі «м'якого» чи «жорсткого» реального часу.

В таких умовах стає актуальним не тільки створення достовірних методів і алгоритмів обробки неоднорідних даних в СС, але й мінімізація кількості співставлень ознак СС для мінімізації  $\Delta t$  та більш оперативного прийняття достовірного рішення.

В результаті проведеного аналізу обґрунтовується наявність суттєвої науково-технічної проблеми, яка виникає через наступне протиріччя. 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.

Рішення цієї проблеми може бути знайдене у двох основних напрямках:

- 1) збільшенні кількості інформаційних ознак, що характеризують стан СС при різноманітних умовах навколишнього середовища та впливу

різноманітних перешкод і викривлень вхідних даних;

2) за поточних зовнішніх для СС умов здійснення ситуативної селекції сукупності найбільш інформативних у поточний момент ознак СС, що надаються для співставлення та класифікації;

3) виділення та сумісна обробка ознак, що входять до логічних поєднань.

В цьому випадку, у будь-який момент часу, прийняття рішення буде здійснюватися на основі виділеної сукупності найбільш інформативних, за поточних зовнішніх умов, даних. Це дозволить отримувати достовірний результат за мінімальний проміжок часу. При чому, початкова сукупність характеристик СС повинна бути інваріантною до різноманітних проявів перешкод і викривлень. Тоді, за різних умов навколишнього для СС середовища, одні й ті самі перешкоди будуть мати різний вплив на різноманітні ознаки. Це забезпечить достатню кількість достовірних ознак для порівняння та формування результату за будь яких умов спостереження за об'єктом розпізнавання.

Також, важливо відзначити, що для переважної більшості практичних застосувань інформаційних технологій обробки даних в СС є достатнім отримання бажаного, апріорно завданого рівня достовірності результату, який задовольняє технічним вимогам.

Таким чином, обґрунтовується основна мета дослідження, яка полягає у забезпеченні заданого рівня достовірності розпізнавання образів, з одночасним зменшенням кількості даних, що надаються для співставлення в складних системах з різнорідними проявами перешкод і викривлень.

На основі проведеного аналізу вказаної вище проблеми була сформульована гіпотеза дослідження, яка полягає в наступному. 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.

Для перевірки гіпотези дослідження та досягнення вказаної мети роботи в розділі 1 було вирішене постановочне завдання аналізу існуючих методів, моделей та інформаційних технологій розпізнавання образів для обробки даних в СС. В подальших розділах розглядається наступний переліку завдань:

- 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 інформаційної моделі гібридного розпізнавання образів;
- розробка моделі визначення рівня інформативності ознак в динаміці їхньої зміни;
- розробка методів консолідації інформативних ознак та розпізнавання образів для складаних систем з неоднорідними даними, 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.

## **Висновки до розділу 1**

Перший розділ присвячений аналізу методів і засобів обробки даних в складних системах. Вирішене перше завдання дисертаційного дослідження, а саме: аналіз існуючих методів, моделей та інформаційних технологій розпізнавання образів для обробки даних в складних системах. Основні висновки полягають в наступному.

Для побудови алгоритмів обчислення оцінок можуть бути застосовані методи переборного типу, градієнтного типу або випадкового пошуку. Але ці методи раціонально використовувати у випадку малої кількості параметрів. У випадку же багатопараметричності та неоднородності даних, такі алгоритми або зовсім не дозволяють розв'язати задачу або потребують використання наближених алгоритмів. Крім цього, все це призводить до значного збільшення часової складності алгоритмів і ускладненню їхньої практичної реалізації, зокрема, у випадку використання комбінованих систем розпізнавання образів в СС, де швидкість рішення задачі класифікації на пряму впливає на актуальність практичного використання результату, особливо, коли використовуються системи реального часу.

Для обробки неоднорідних даних в СС даних використовується комбіноване розпізнавання та методи, засновані на АОО та КВП. Однак, ці методи раціонально використовувати у випадку малої кількості даних. У випадку багатопараметричності та неоднорідності даних, такі алгоритми або зовсім не дозволяють розв'язати задачу або потребують використання наближених алгоритмів. Крім цього, все це призводить до значного

збільшення часової складності алгоритмів і ускладненню їхньої практичної реалізації.

Аналіз існуючих підходів до комбінованого розпізнавання для обробки неоднорідних даних в СС показав, що швидкість рішення задачі класифікації напряму впливає на актуальність практичного використання результату, особливо, коли використовуються системи реального часу.

Обрання конкретного виду приватних критеріїв є одним з основних моментів у побудові інформаційної моделі СС. В відомій матричній інформаційній моделі СС в якості такого критерію використовується семантична міра доцільності інформації Харкевича А.А., яка заснована на апостеріорному підході. Сутність такого підходу полягає в тому, що кількість інформації оцінюється за зміною ступеню доцільності поведінки системи в результаті сповіщення їй цієї інформації. Фактично, такий процес і має місце у випадку аналізу СС за станом, а не за параметрами.

При збільшенні кількості прецедентів, за якими приймається рішення, достовірність результату може бути підвищена, але при цьому збільшується кількість співставлень даних і, як наслідок, підвищується часова складність інформаційного процесу класифікації. При цьому, наявність певного рівня перешкод і викривлень призводить до зменшення інформативності та доцільності використання певних ознак для класифікації. Тому, для забезпечення бажаного рівня достовірності та зменшення часової складності прийняття рішення щодо стану СС, інформаційна технологія обробки даних повинна забезпечувати у будь який момент часу прийняття рішення на основі виділеної кількості найбільш інформативних, за поточних зовнішніх умов, ознак. В цьому випадку достовірність результату буде забезпечуватися за мінімальною кількістю співставлень даних та у мінімальні проміжки часу.

Обґрунтована головна мета роботи, яка направлена на вирішення визначеної раніше науково-технічної проблеми. Мета полягає в забезпеченні заданого рівня достовірності розпізнавання образів з одночасним

зменшення кількості даних, що надаються для співставлення, в складних системах з різнорідними проявами перешкод і викривлень.

Обґрунтовані завдання дослідження, які підлягають вирішенню в подальшому. А саме: розробка принципу консолідації неоднорідних даних в СС з різнорідними проявами перешкод і викривлень; обґрунтування концепції гібридного розпізнавання образів для консолідованої обробки даних в СС з різнорідними проявами перешкод і викривлень; формалізація процесу представлення даних та розробка інформаційної моделі гібридного розпізнавання образів; розробка моделі визначення рівня інформативності ознак в динаміці їхньої зміни; розробка методів консолідації інформативних ознак та розпізнавання образів для СС з неоднорідними даними, різнорідними проявами перешкод і викривлень; на основі розроблених методів і моделей, створення інформаційної технології гібридного розпізнавання образів для консолідованої обробки даних в СС; практична реалізація розроблених моделей, методів та інформаційної технології для побудови систем обробки неоднорідних даних різноманітного спрямування.

## РОЗДІЛ 2

### КОНСОЛІДАЦІЯ НЕОДНОРІДНИХ ДАНИХ ТА РОЗПІЗНАВАННЯ ОБРАЗІВ В СКЛАДНИХ СИСТЕМАХ З РІЗНОРІДНИМИ ПРОЯВАМИ ПЕРЕШКОД І ВИКРИВЛЕНЬ

#### 2.1 Принцип консолідації неоднорідних даних в СС з різнорідними проявами перешкод і викривлень

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 рисунку 2.1 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 сприятиме 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 2.1 – Вплив різного рівня завад на достовірність прийняття рішення

При вирішенні задачі розпізнавання типу транспорту за візуальними характеристиками ефективність і точність розпізнавання знаходиться у значній залежності від умов спостереження за об'єктом. Так ясна погода і світлова частина доби створюють більш сприятливі умови для коректного розпізнавання, ніж наявність туману і темної частини доби. Або, для інших застосувань, наприклад, розпізнавання речових команд в системах автоматизованого управління простіше реалізується за умови відсутності високого рівня шуму. Відповідно до цього, прийняття класифікаційних рішень повинне здійснюватися з урахуванням імовірності появи додаткових перешкод під час отримання і обробки неоднорідних даних. Таким чином, для коректного вирішення задачі обробки даних СС в умовах наявності додаткових перешкод множина класифікуючих ознак  $\{X\}$  повинна містити ознаки, які інваріантні до можливих перешкод і викривлень.

Додатковим аспектом, який ускладнює ефективне рішення задач обробки даних в СС є використання класифікуючих ознак, що мають схожу природу і відповідно, існує значна імовірність, що поява помилки або



можуть бути звук двигуна, маса об'єкта тощо. В цьому випадку, серед сукупності отриманих образів об'єкту можна виділити ті, ознаки яких інваріантні до наявних перешкод при розпізнаванні. Так звук або маса автотранспорту не залежать від рівня туману або освітлення сцени спостереження. Таким чином, відповідно до нового принципу консолідації даних СС, завдяки введенню додаткових інформативних ознак, які знаходяться в малій залежності одна від іншої, дозволяє ефективно вирішувати задачі класифікації в умовах додаткових перешкод, або часткової невизначеності. Відповідно, якщо в поганих умовах спостереження візуальних ознак об'єктів автотранспорту досить складно співвіднести образ до одного з передвизначених класів (легковий автомобіль, вантажний автомобіль, велосипед, мотоцикл і т. д.), то використання запропонованого принципу спрощує таку задачу. Очевидно, що вказані класи об'єктів легко класифікуються, наприклад, за масою, і можуть бути розподілені у відповідності до ряду збільшення ознак ваги: велосипед → мотоцикл → легковий автомобіль → вантажний автомобіль.

## **2.2 Модифікована матрична інформаційна модель обробки неоднорідних даних в складних системах**

Розглянута раніше, матрична інформаційна модель СС не розрахована на обробку даних за окремими сукупностями образів, тому для її використання при побудові інформаційної технології обробки неоднорідних даних потрібна її модифікація.

В таблиці 2.1 представлена модифікована матрична інформаційна модель СС, в якій враховується запропонований принцип консолідації неоднорідних даних.

Критерій відповідності стану СС класу  $C_j$  в результаті співставлення ознаки  $x_l$  образа  $P_k$ , відповідно до міри Харкевича буде визначатися як

$$I_{kl}^j = \text{Log}_2(P_{kl}^j / P^j), \quad (2.2)$$

де  $P_{kl}^j$  – ймовірність переходу стану СС до стану  $C_j$  під впливом  $kl$ -го фактору;

$P^j$  – ймовірність спонтанного переходу СС до стану  $C_j$ .

Таблиця 2.1 – Модифікована матрична інформаційна модель складної системи

Ознаки СС		Стани СС			Потужність ознаки
		$C_1$	...	$C_j$	
$P_1$	$x_{1l}$	$I_{1l}^1$		$I_{1l}^j$	$\sigma_{1l}$
	...	...	...	...	...
	$x_{1n}$	$I_{1n}^1$		$I_{1n}^j$	$\sigma_{1n}$

---

Ознаки СС		Стани СС			Потужність ознаки
		$C_1$	...	$C_j$	
$P_k$	$x_{kl}$	$I_{kl}^1$		$I_{kl}^j$	$\sigma_{kl}$
	...	...	...	...	...
	$x_{kl}$	$I_{kl}^1$		$I_{kl}^j$	$\sigma_{kl}$

Слід зазначити, що критерій (2.2), за своєю математичною формою, нічим не відрізняється від залежності, що описує перевищення сигналу над завадою для інформаційного каналу [167]. З цього, на перший погляд, випадкового співпадіння можна зробити корисну інтерпретацію виразу еквівалентної форми міри Харкевича, де  $P_{kl}^j$  буде характеризувати ймовірність класифікації стану СС, відповідного класу  $C_j$  на основі

співставлення ознаки  $x_l$  образу  $P_k$ , а  $P^j$  – відповідатиме класифікації стану СС до класу  $C_j$  в результаті будь-яких інших співставлень з ознаками, відмінними від  $x_l$  образу  $P_k$ . Тоді, відповідно до теорії інформації, можна вважати, що виявлення деякої ознаки  $x_l$  образу  $P_k$  СС, відповідає отриманню сигналу, що містить деяку кількість інформації про те, що стан СС відповідає класу  $C_j$ :

$$D_{kl}^j = \text{Log}P_{kl}^j. \quad (2.3)$$

Для того, щоб оцінити достатність кількості цієї інформації для отримання достовірного результату класифікації, необхідно мати точку відліку чи базу для порівняння. В якості цієї бази використовується середня за всіма класами розпізнавання кількість інформації, що отримується у випадку виявлення цієї ознаки  $x_l$  образу  $P_k$ :

$$\bar{I}_{kl} = \text{Log}P_{kl}. \quad (2.4)$$

Тобто, якщо в процесі розпізнавання виявлена ознака  $x_l$  образу  $P_k$ , то для того, що сформулювати обґрунтоване рішення щодо класифікації стану СС до одного з передвизначених класів, необхідно мати інформацію про частоту виявлення класифікуючої ознаки відповідності стану СС відповідному класу. З урахуванням запропонованого принципу консолідації та моделі представлення СС (2.1) інформація про частоту виявлення класифікуючої ознаки може бути отримана на основі співставлення результатів класифікації стану СС за кожним окремим образом з сукупності  $P_1-P_k$ . Іншими словами, вирішальне правило щодо класифікації може бути побудоване наступним чином: стан СС відноситься до класу  $C_j$  з наявною передвизначеною сукупності  $\{C\}$ , якщо частота такої класифікації є переважною за більшістю

образів з множини  $\{P\}$ . Цей аспект дає обґрунтовану базу для розробки нового метода розпізнавання на основі обробки неоднорідних даних в СС, розробці якого приділена увага в наступному розділі 3.

Фактично, середню кількість інформації можна розглядати як деякий "інформаційний шум", який наявний у відповідній ознаці та не несе ніякої корисної інформації про приналежність стану СС до того чи іншого класу.

З урахуванням вище викладеного, критерію інформативності (2.2) можна надати більш класичний для теорії зв'язку вигляд:

$$I_{kl}^j = \text{Log}P_{kl}^j - \text{Log}P_{kl} = D_{kl}^j - \bar{I}_{kl}. \quad (2.5)$$

Ця форма інтерпретується як віднімання шуму з корисного сигналу. Ця операція є тривіальною в системах шумозаглушення. Якщо корисний сигнал вищий за рівень шуму, то його виявлення несе інформацію на користь належності стану СС відповідному класу. В протилежному випадку така класифікація вважається такою, що не відповідає дійсності.

Також слід відзначити, що сам А. А. Харкевич розглядав  $P_{kl}^j$  як ймовірність досягнення мети, за умови, що система отримала інформацію  $I_{kl}^j$ , а  $P^j$  – як ймовірність його досягнення за умови, що система цієї інформації не отримувала [5]. Необхідно також відмітити, що кожна ознака СС, як каналу зв'язку, може бути охарактеризована динамічним діапазоном, який дорівнює максимально можливому рівню сигналу в каналі і рівня перешкод в логарифмічному масштабі:

$$I_{kl \max}^j = \text{Log}P_{kl \max}^j - \text{Log}P_{kl} = \text{Log}(P_{kl \max}^j / P_{kl}). \quad (2.6)$$

Максимальна кількість інформації, яка може міститися в ознаці, поністю визначається кількістю класів розпізнавання  $N$  і дорівнює кількості інформації по Хартлі:  $I = \text{Log}_2 N$ . Динамічний діапазон ознаки є кількісною мірою її корисності (цінності для отримання достовірної класифікації стану СС) для розпізнавання. Для визначення ступеня інформативності (корисності) ознаки для прийняття достовірного рішення використовується середня кількість корисної для класифікації інформації в ознаці, тобто виправлене вибіркве середньоквадратичне відхилення інформативностей [5, 167], яка для представленої модифікованої інформаційної моделі приймає вигляд:

$$\sigma_{kl} = \sqrt{\frac{I}{N-1} \sum_{j=1}^N (I_{kl}^j - \bar{I}_{kl})^2} . \quad (2.7)$$

Виходячи з вище наведеного, ступінь інформативності ознак образів об'єкта розпізнавання напряду залежить від співвідношення значення цієї ознак та перешкод, які мають місце в інформаційних каналах отримання даних сукупності образів  $\{P\}$  різної природи виникнення. Це надає наукове обґрунтування тому, що для визначення критеріїв інформативності ознак СС для співставлення та тримання бажаної достовірності результату обробки даних, слід використовувати поточний рівень перешкод і викривлень, що мають місце в каналі отримання даних і, відповідно, є поточною характеристикою зовнішніх для СС умов. Цей аспект використовується в подальшому для розробці системи критеріїв та методу визначення інформативності образів СС.

### 2.3 Концепція гібридного розпізнавання образів для консолідованої обробки даних в СС з різномірними проявами перешкод і викривлень

На основі нового принципу консолідації даних та модифікованої матричної інформаційної моделі запропонована концепція гібридного розпізнавання образів [161-165], 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 різномірних перешкод і викривлень. Процес гібридного розпізнавання проілюстрований на рисунку 2.2.

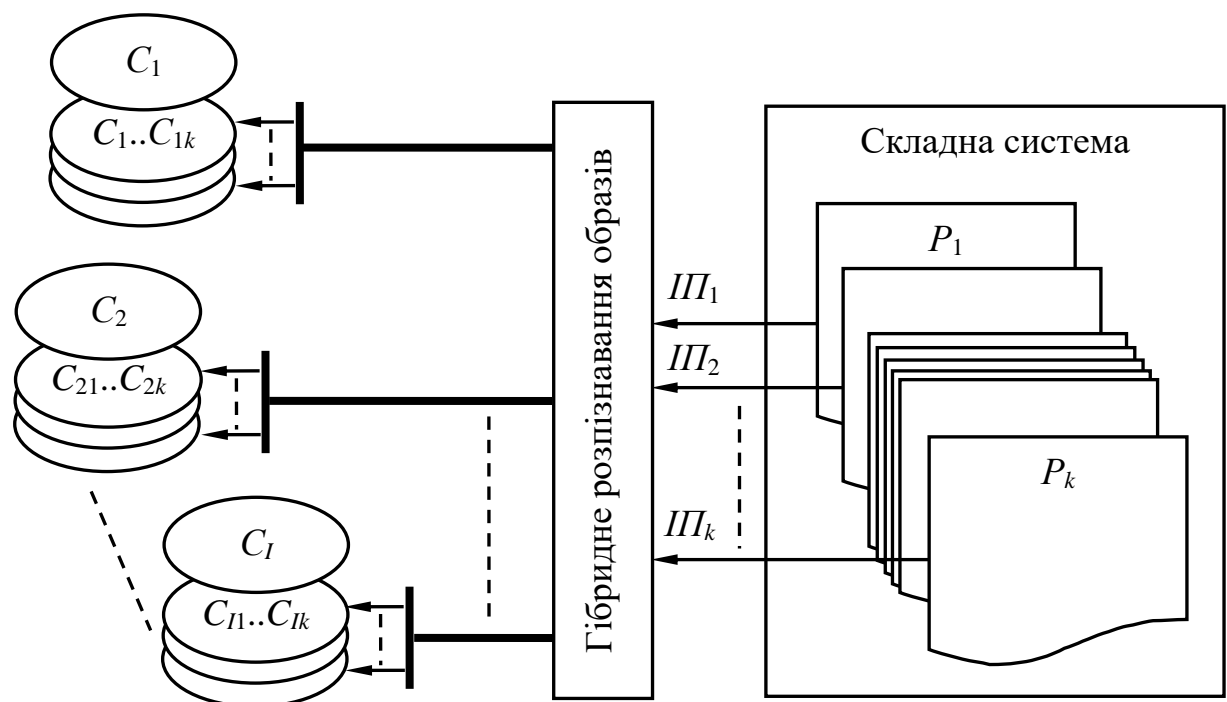


Рисунок 2.2 – Гібридне розпізнавання образів

При гібридному розпізнаванні, стан СС описується множиною характеристик різної природи виникнення. У відповідності до цього, формується  $k$  інформаційних потоків ( $III_1, III_2, \dots, III_k$ ), які використовуються для отримання даних СС та характеристик поточних зовнішніх умов, які



необхідні для визначення ступеня інформативності ознак об'єкта розпізнавання (згідно моделі, розглянутої у попередньому підрозділі). Аналіз цих інформаційних потоків дозволяє співвіднести об'єкт розпізнавання до одного з передвизначених класів ( $C_1, C_2, \dots, C_I$ ).

Особливістю опису передвизначених класів при гібридному розпізнаванні являється створення сукупності ознак, які дозволяють однозначно класифікувати стан  $CC$  до одного з класів  $C_1-C_I$  на підставі отримання інформації за кожним з  $k$  інформаційних потоків. У відповідності до того, що  $CC$  представляється сукупністю образів  $\{P\}$ , які відповідають кожному окремому інформаційному потоку, класи також повинні бути структуровані за класифікуючими ознаками, які відповідають кожному інформаційному потоку. В цьому випадку, кожен клас  $C_1-C_I$  повинен представляти собою сукупність підкласів  $C_{I_k}$ , за якими може здійснюватися класифікація. Таким чином, один з образів  $P_k$ , за сукупністю його класифікуючих ознак  $\{X_k\}_C$ , може бути співставлений з відповідним підкласом  $C_{I_k}$ , що дозволяє встановити відповідність стану  $CC$  одному з передвизначених класів  $C_I$ .

Представлена концепція гібридного розпізнавання дозволяє використати два окремих підходи до побудови методів співставлення даних та класифікації в  $CC$ : спільний або роздільний аналіз.

Спільний аналіз передбачає трансформацію моделі опису складної системи (2.1) до класичного вигляду (1.2), який передбачає обробку одного глобального образу. При цьому аналіз  $CC$ , представленої моделлю (2.1), повинен проводитися за умови

$$P_1 \cup P_2 \cup \dots \cup P_k. \quad (2.8)$$

Тоді систему (2.1) можна представити у вигляді

$$P = \{X_{11}, X_{12}, \dots, X_{1n}\} \cup \{X_{21}, X_{22}, \dots, X_{2m}\} \cup \dots \cup \{X_{k1}, X_{k2}, \dots, X_{kl}\}, \quad (2.9)$$

де  $P$  – глобальний образ, що включає усі ознаки відповідних образів ( $P_1$ - $P_k$ ).

З урахуванням об'єднаних множин, вираження (2.9) можна записати у вигляді

$$P = \{X\}, \quad (2.10)$$

що фактично підтверджує виродження моделі (2.1) до вигляду (1.2), тобто класичному представленню СС для комбінованого розпізнавання. Відповідно до цього, подальша селекція класифікуючих ознак СС, обробка даних і розпізнавання може проводитися з використанням традиційних підходів. Однак, з приводу того, що множина  $\{X\}$  містить ознаки, які сформовані використанням декількох інформаційних потоків, обумовлених різною природою виникнення, забезпечується більша кількість непов'язаних або слабо пов'язаних ознак, що дозволяє забезпечити ідентифікацію помилки або перешкоди та підвищити достовірність розпізнавання.

Таким чином, можна зробити висновок, що використання декількох інформаційних потоків для отримання ознак СС різної природи виникнення, дозволяє отримувати більш достовірний результат класифікації навіть у випадку використання класичних підходів комбінованого розпізнавання на основі АОО та КВП.

При роздільному аналізі, задача класифікації може вирішуватись шляхом  $k$  приватних рішень за кожним з визначених образів. В цьому випадку, здійснюється співвідношення кожного образу об'єкта до відповідного класу, як це проілюстровано на рисунку 2.2, а остаточне рішення приймається за урахуванням ймовірнісних характеристик. Тобто, остаточне рішення щодо класифікації отримується на підставі вагових коефіцієнтів кожного з інформаційних потоків, або через аналіз сукупності

образів і визначення їх класифікуючих ознак, що мають найменше розходження (малі значення метрики).

Одним з найпростіших алгоритмів прийняття рішень щодо класифікації при гібридному розпізнаванні є мінімізація метрики ознак за кожним інформаційним потоком (образом), або проста композиція результатів співвідношення образів до класів за кожним інформаційним потоком. У випадку раціонального визначення множини класифікуючих ознак  $\{X\}_C$  кластеризація буде здійснюватися з малим значенням метрики, що забезпечить добру збіжність результату розпізнавання.

Таким чином, використання саме роздільного аналізу при гібридному розпізнаванні дозволяє отримати найбільші переваги у напрямку забезпечення різноманіття стратегій обробки неоднорідних даних та прийняття рішень в СС. Враховуючи представлений принцип консолідації ознак за різною природою виникнення і створення сукупності образів  $P_1-P_k$ , класифікація за кожним з цих образів може бути використана як додатковий прецедент для підтвердження або спростовування отриманого результату класифікації. В цьому випадку, вирішальне правило може бути побудовано відповідно до наступної умови: стан СС визначається таким, що відповідає одному з передвизначених класів у випадку, коли така класифікація підтверджується за результатами співставлень більшості образів з сукупності  $P_1-P_k$ .

Методологічні аспекти реалізації сумісного то роздільного аналізу при гібридному розпізнаванні докладно розглянуті в розділі 3.

## **2.4 Ситуаційно-подійна модель гібридного розпізнавання образів**

Гібридне розпізнавання, забезпечуючи підвищення достовірності отримання результату обробки неоднорідних даних в СС з різнорідними проявами перешкод і викривлень, базується на збільшенні кількості

можливих ознак, що характеризують СС за різними зовнішніми умовами. При цьому, раніше вже вказувалося, що для отримання точного і оперативного результату аналізу даних СС, необхідно, у будь-який момент часу, здійснювати обробку обмеженої сукупності але найбільш інформативних даних. При цьому, зменшення кількості співставлень ознак для прийняття рішення забезпечить зниження часової складності процесу аналізу даних та класифікації стану СС.

Враховуючи той факт, що у будь який момент часу життєвого циклу СС набір інформативних ознак для співставлення і класифікації повинен змінюватися відповідно поточним зовнішнім умовам, процес обробки даних і гібридного розпізнавання буде мати ситуативний характер [149, 168-175].

Для забезпечення прийняття рішень на основі обмеженої кількості але найбільш інформативних ознак, було розроблено ситуаційно-подійну модель гібридного розпізнавання образів у вигляді кортежу [176]

$$SEMHPR = \langle P, EC, SICCh, NSICCh, C \rangle, \quad (2.11)$$

де  $P$  – множина образів складної системи, 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 складної системи;  $NSICCh$  – множина нестаціонарних характеристик інформативності образів складної системи;  $C$  – множина класів, що характеризують стан складної системи. 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. З цією метою на основі множини  $EC$

визначаються стаціонарні та нестаціонарні характеристики інформативності для кожного образу з  $\{P\}$ .

Селекція інформативних ознак СС, згідно моделі, здійснюється на підставі стаціонарних і нестаціонарних характеристик інформативності, які забезпечують кількісну оцінку ступеня викривлення ознак СС на основі поточного рівня і характеру перешкод у зовнішньому просторі.

Серед відомих підходів щодо виключення впливу викривлень відображення характеристик об'єкта розпізнавання у просторі образів, широкого розповсюдження набула концепція побудови класифікатора на основі архітектури неокогнітронну [177-179]. Стійкість неокогнітронного класифікатора до викривлень відображення об'єктів у просторі образів пов'язана з тим, що неокогнітрон представляє собою ієрархічну нейронну структуру, кожен шар якої складається з масиву площин. На більш високих шарах, вузли нейромережі реагують на більш складні і абстрактні образи, ніж на нижчих шарах. На ще більш високих рівнях ступень абстракції ще збільшується. У загальному випадку, вузли на більш високих рівнях отримують вхід від групи низькорівневих вузлів і, відповідно, реагують на більш широку область вхідного поля ознак. Таким чином, реакції вузлів більш високого рівня менш залежать і більш стійкі до локальних викривлень вхідного потоку інформаційних ознак, що надходять для розпізнавання.

Однак використання такої концепції пов'язане з низкою складнощів і недоліків, а саме:

- неокогнітронна структура надзвичайно складна, вимагає достатньо великого обсягу обчислень, що негативно впливає на часову складність процесу розпізнавання і ускладнює практичну реалізацію таких класифікаторів;

- значна часова складність процесу навчання;

- відсутність характеристики внесених викривлень і його впливу на отримання остаточного результату класифікації.

Для усунення вказаних складнощів, та розробки методу визначення інформативності образів СС (визначення характеристик  $SIC_h$  і  $NSIC_h$ ) використано показники точності системи відображення ознак і характеристики поточного рівня завад зовнішнього для СС середовища.

Метод полягає в наступному. Якщо представити, що образ СС  $P$  описується деякою функцією [180]

$$P = f(x, y), \quad (2.12)$$

де  $x, y$  – аргументи, що визначають характеристику СС, яка знаходиться у деякому просторі  $\Omega$ .

Тоді відображення цього об'єкта  $I$  визначається як

$$I = g(x', y'), \quad (2.13)$$

де  $x', y'$  – аргументи, що визначають характеристику образа, який знаходиться у деякому інформаційному просторі  $\Omega'$ .

У випадку ідеальної відображаючої системи, для будь-якої точки простору виконується умова

$$g = f. \quad (2.14)$$

В реальних умовах залежність (2.14) не може виконуватись повною мірою, за причини наявності викривлення відображення. Таким чином, для достовірного процесу розпізнавання завжди існує необхідність визначення функціонального зв'язку між  $g$  і  $f$ . Іншими словами, існує необхідність використання методів визначення закону внесення викривлень при відображенні характеристик СС в її образі (образах). Причому, ця

функціональна характеристика повинна мати просторову прив'язку. Це дозволить в подальшій обробці ознак образів здійснити компенсацію викривлень і підвищити достовірність класифікації.

Для спрощення представлення аналітичних залежностей представимо, що відображення образу здійснюється без масштабування. В цьому випадку, розподіли  $f$  і  $g$  мають однакову розмірність. Крім цього, не будемо враховувати взаємний вплив суміжних точок простору  $\Omega$ , що може мати місце в реальних умовах.

Розглянемо зв'язок між простором об'єкта і простором його образу [181]. В простір образу надходить інформація з простору об'єкту з урахуванням внесених викривлень системою відображення. Введемо функцію  $h(x', y', x, y)$ , яка описує просторовий зв'язок між об'єктом і його відображенням. Будемо враховувати, що перенесення інформації здійснюється від точених джерел. Тоді зображення для точки простору об'єкту з координатами  $(\alpha, \beta)$  буде визначатися як

$$g'(x', y') = h(x', y', \alpha, \beta, f'(\alpha, \beta)). \quad (2.15)$$

У вираженні (2.15) залежність розподілу відображення від рівня сигналу від точкового джерела враховується додатковим функціональним аргументом  $f'(\alpha, \beta)$ . Очевидно, що для іншого точкового джерела, який має таке ж саме розміщення

$$g''(x', y') = h(x', y', \alpha, \beta, f''(\alpha, \beta)). \quad (2.16)$$

У відповідності до принципу суперпозиції, отримані характеристики складаються.

$$\begin{aligned}
 g'(x', y') + g''(x', y') &= \\
 &= h(x', y', \alpha, \beta, f'(\alpha, \beta)) + h(x', y', \alpha, \beta, f''(\alpha, \beta)).
 \end{aligned}
 \tag{2.17}$$

Як витікає з (2.17), за рахунок нелінійної суперпозиції, сумі розподілів у просторі відображення не відповідає сума функцій у просторі об'єкта.

У випадку лінійності системи

$$g'(x', y') = h(x', y', \alpha, \beta) f(\alpha, \beta), \tag{2.18}$$

а суперпозиція буде мати наступний вигляд:

$$g'(x', y') + g''(x', y') = h(x', y', \alpha, \beta) [f'(\alpha, \beta) + f''(\alpha, \beta)]. \tag{2.19}$$

Таким чином, у випадку лінійності системи, складання функцій у просторі об'єкта приводить до складання розподілів у просторі відображення з точністю до однієї функції перетворення  $h$ .

В загальному випадку, для нелінійної системи, співвідношення зв'язку простору об'єкта і простору зображення буде мати наступний вигляд:

$$g(x', y') = \iint_{x \in X, y \in Y} h(x', y', f(x, y)) dx dy, \tag{2.20}$$

а для лінійної

$$g(x', y') = \iint_{x \in X, y \in Y} h(x', y') f(x, y) dx dy. \tag{2.21}$$

Якщо значення  $x$  і  $y$  не корелюють друг з другом, функція  $h$  буде мати наступний вигляд:



$$h(x', y', x, y) = h'(x', x)h(y', y). \quad (2.22)$$

Функція  $h$  враховує викривлення відображення об'єкта на просторі образу, а отримане відображення може бути представлено так:

$$g(x', y') = \iint_{x \in X, y \in Y} h(x', y', x, y) f(x, y) dx dy + \xi(x', y'), \quad (2.23)$$

де  $\xi(x', y')$  – характеристика розподілу викривлень відображення об'єкта у просторі образу.

Виходячи з вираження (2.23) можна зробити висновок, що, для підвищення точності відображення характеристик СС і достовірності подальшого розпізнавання, необхідне використання засобів мінімізації характеристики  $\xi(x', y')$  або її компенсації при обробці даних та прийнятті остаточного рішення.

У випадку нестационарності викривлень, коли їх прояв не постійний у часі, можливе використання серії послідовних замірів, після чого статистична обробка отриманих даних дозволяє виключити вплив завади. Такий підхід розглянутий у [182, 183] для аналізу зображення, що представляється послідовністю відеокadrів. Аналіз послідовності декількох кадрів зображення дозволяє виділити графічні елементи, що не проявляються на кожному з них, а відповідно, ідентифікуються в системі як завада.

У випадку стаціонарності викривлень задача істотно ускладнюється. При гібридному розпізнаванні ця задача значно спрощується, так як використання декількох інформаційних потоків, та формування сукупності образів, що мають різну природу виникнення, підвищує кількість незалежних ознак об'єктів, що не корелюють, або слабо корелюють між собою.



Характеристики перешкод  $\xi_1$ - $\xi_k$  змінюються впродовж життєвого циклу СС в залежності від змін зовнішніх умов. Ця сукупність є нестационарною та вимагає постійного контролю на кожному циклі обробки даних СС та здійснення класифікації. Тоді, за аналогією зі стаціонарною характеристикою інформативності, може бути визначена нестационарна характеристика інформативності кожного з  $k$  образів  $NSICH_k$  [180]

$$NSICH_k \sim \frac{1}{\xi_k(x', y')}. \quad (2.26)$$

Тоді, загальна характеристика інформативності може бути представлена у вигляді

$$ICH_k = SICH_k \cdot NSICH_k. \quad (2.27)$$

Для фактичного визначення чисельних значень стаціонарних та нестационарних характеристик інформативності образів для кожного конкретного застосування в системі необхідно провести нормування характеристик  $h$  і  $\xi$ .

На підставі (2.24) і враховуючи (2.25) і (2.26) можна зробити висновок, що, при гібридному розпізнаванні, для кожного образу, відповідно до запропонованої вище методики, може бути визначена функція  $h$  і, відповідно, окрема характеристика  $\xi$ .

Враховуючи різну природу образів СС, що отримуються різними відображаючими системами, на підставі статистичного аналізу з'являється можливість ідентифікації завади за окремими каналами.

Крім цього, аналіз значень характеристик  $\xi$  для кожного інформаційного каналу дозволяє визначити ті образи, що в поточних умовах

виміру не дозволяють адекватно відображати об'єкт розпізнавання. Для підвищення достовірності розпізнавання, такі образи мають відкидатися. Особливе значення в цьому випадку має використання таких інформаційних каналів отримання даних  $CC$ , для яких характеристики  $\xi$  мають малу коваріацію.

Представлені критерії інформативності використовуються для ранжування образів. Тобто, у будь який момент життєвого циклу  $CC$  найбільш інформативніший образ буде мати максимум характеристики інформативності. Тоді, найбільш достовірне рішення  $Rd$  отримується за результатом аналізу образу, що задовольняє критерію інформативності [180]

$$\exists P_k \in \{P\}, \max ICh_k \Rightarrow Rd. \quad (2.28)$$

Ранжування образів  $CC$  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 [180]

$$\forall P_k \in \{P\}' \mid \{P\}' \subset \{P\}, SICH_k \geq SICH^*. \quad (2.29)$$

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 апріорно 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.

Ці моменти докладно розглядаються у розділі 4 при розробці інформаційної технології гібридного розпізнавання образів для консолідованої обробки даних в СС.

## **2.5 Інформаційна модель і метод визначення послідовності обробки даних образів складної системи**

Отримання чітко-визначеної черги обробки даних вимагає селекції та ранжування не тільки множини образів  $\{P\}$  СС, про що згадувалося у попередньому підрозділі, а й ознак цих образів, які підлягають співставленню. З цією метою в роботі розроблено метод та інформаційну модель визначення послідовності даних, які є результатом узагальнення когнітивної моделі емоційних процесів людини Аткінсона-Шифріна для СС з неоднорідними даними та різнорідними проявами перешкод та викривлень

Як вже згадувалося раніше, системи розпізнавання образів позбавлені необхідності повного аналітичного представлення усіх закономірностей поведінки об'єктів для прийняття рішення. Крім цього, інтелектуальні

алгоритми розпізнавання, дозволяють використовувати різноманітні евристики, на кшталт людської поведінки, що також сприяє спрощенню інформаційних процесів обробки даних і прийняття рішень. Фактично, починаючи з робіт Уоррена Мак-Каллока і Уолтера Питтса по штучним нейронним мережам, що розвинуті у роботах Ф. Розенблатта по перцептроні, спостерігалася тенденція до використання моделей розумової діяльності людини. Значний внесок у цьому також був зроблений Р. Пенроузом, у своїй спірній теорії квантової свідомості. Таким чином, сучасні інтелектуальні методи і алгоритми розпізнавання та прийняття рішень використовують в тій чи іншій мірі моделі розумової активності людини [184, 185].

По мірі ускладнення задач, що вирішуються системами розпізнавання, також підвищуються вимоги щодо їхнього функціонування. Існує прагнення до створення універсальних систем розпізнавання, що на кшталт людини, могли б вирішувати широке коло задач з використанням відповідних спрощень, евристик. Для гібридного розпізнавання ці аспекти надважливі з погляду на те, що їх використання пов'язане зі значним збільшенням кількості неоднорідних даних, що надаються для співставлення та класифікації.

Відомо, що біологічним організмам притаманна можливість вирішення різноманітного кола задач розпізнавання у достатньо широкому спектрі змін умов спостереження за ознаками, а також у випадку їхньої неоднорідності.

Дослідження когнітивних здібностей біологічних організмів, зокрема людини, при достатньо високій універсальності та різноманітні вирішуваних задач розпізнавання, вказує на різний час прийняття рішень навіть за умови постійності вказаних вище факторів, чого не спостерігається у технічних системах. Так, наприклад, будь-який водій автотранспорту, за умови значної досвідченості, яка дозволяє судити про нього як про експерта, у ідентичних ситуаціях може приймати рішення різної достовірності. При цьому, не завжди відома система критеріїв за якими здійснювалося прийняття рішення.

Цей факт вказує на наявність додаткових умов здійснення інформаційного процесу аналізу ознак та класифікації образів, визначення яких можливе через аналіз моделі пам'яті біологічних організмів [185].

Аналіз існуючих на сьогоднішній день моделей пам'яті вказує на достатньо високу достовірність моделі Р. Аткинсона [186, 187], так як вона добре пояснює експериментальні результати по рішенню людиною завдань обробки інформації. Згідно цієї моделі, існують три види пам'яті: сенсорна, короткочасна і довготривала. Види пам'яті розрізняються часом утримання і об'ємом, а також способом кодування і рівнем організації інформації, що зберігається. Інформація з навколишнього середовища поступає в сенсорні реєстри, де зберігається біля третини секунди. Далі вона поступає в короткочасну пам'ять, де піддається кодуванню і може зберігатися до 30 секунд (а при повтореннях – істотно більше). Без повторень інформація або витісняється іншою інформацією, або згасає. Через короткочасну пам'ять інформація може надходити у довготривалу пам'ять. Останню можна уявити собі як невизначене за об'ємом сховище, в якому інформація може зберігатися скільки завгодно довго. Використання цієї моделі раціональне ще з приводу того, що виникла вона на основі, так званої, комп'ютерної метафори, що проводить паралель між процесами обробки інформації у комп'ютерних системах (введення інформації, обробка та зберігання у пам'яті, яка поділяється на оперативну та постійну) та людської системи обробки інформації.

В процесі прийняття рішень, насамперед, здійснюється обробка даних, що містяться у короткочасній пам'яті. Тобто, сам процес прийняття рішень відбувається у короткочасній пам'яті куди може надходити як апостеріорна вхідна інформація з сенсорної пам'яті, так і апріорна інформація, що міститься у довготривалій пам'яті.

Процес зберігання інформації пов'язаний з прагненням до глобалізації, тобто до укрупнення класів інформації, що зберігається. При цьому відомо,

що система пов'язаних між собою даних зберігається мозком ефективніше ніж окремі дані. Цей факт також пов'язаний з кінцевою пропускнуою здатністю мозку, яка була визначена ще Дж. Мілером [188]. Було встановлено, що людський мозок здатний запам'ятовувати комбінації даних у кількості  $7 \pm 2$ . Тобто людина не здатна запам'ятовувати довгі послідовності непов'язаних даних. Пов'язані дані групуються у одну смислову категорію, що прийнято називати чанком [189]. Тоді, максимальний обсяг інформації, яка може бути одночасно сприйнята, дорівнює  $7 \pm 2$  чанки.

Узагальнення когнітивного апарату людини та принципу ранжування даних за тривалістю зберігання, дозволило отримати інформаційну модель визначення послідовності обробки даних образів СС, як узагальнення когнітивного апарату людини.

На рисунку 2.3 представлена розроблена інформаційна модель визначення послідовності обробки даних, яка узагальнена для технічних інтелектуальних систем розпізнавання образів [185]. У відповідності до запропонованої 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. Для гібридних систем розпізнавання, аналогом чанків є образи об'єкта розпізнавання. Слід зазначити, що при гібридному розпізнаванні до одного чанку може бути віднесено кілька образів, які об'єднуються у групи ідентичної класифікації, так як апріорі відомо, що кожен образ характеризує один й той самий стан СС. В результаті комбінаторного аналізу формуються укрупнені смислові категорії (образи або 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 довгостроковій пам'яті на момент здійснення класифікації (ап'іорні відомості).

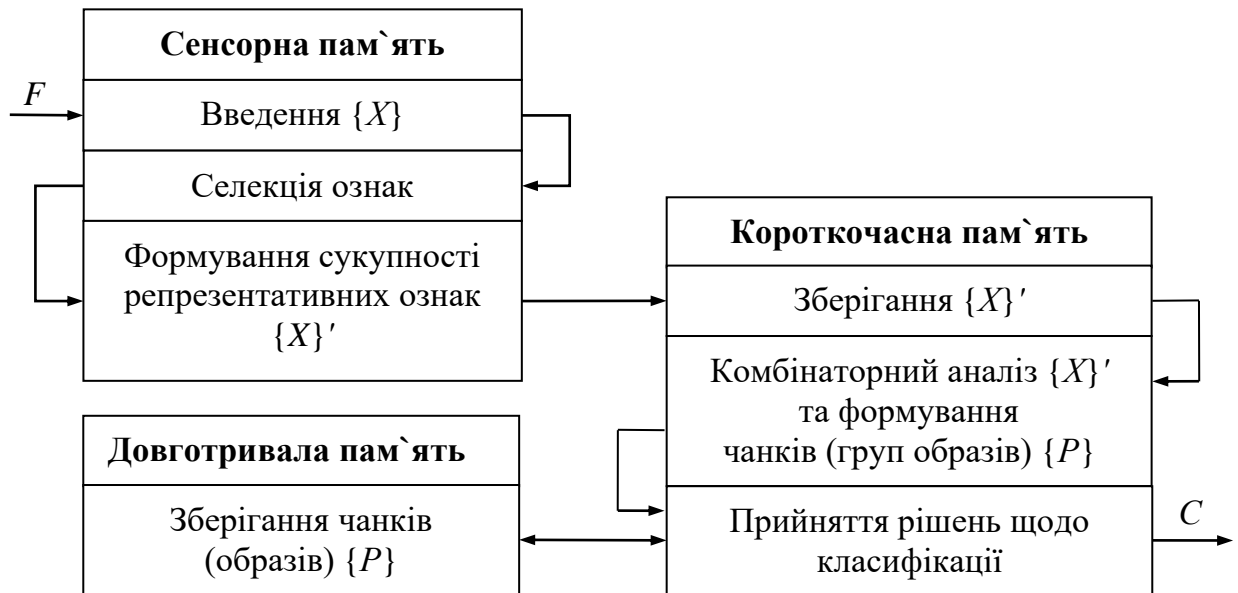


Рисунок 2.3 – Інформаційна модель визначення послідовності обробки даних

Представлена модель ілюструє процес обробки даних і прийняття рішень щодо класифікації, однак не пояснює вище розглянуту ситуацію різної ефективності прийняття рішень одним експертом. Пояснення цієї невідповідності можна знайти у тому, що технічні системи, фактично повністю відтворюючи когнітивні процеси біологічних організмів, не враховують емоційні процеси, такі як афект, емоція, почуття та настрої. При цьому відомо, що в біологічних організмах ці процеси істотно впливають на достовірність та швидкість прийняття рішень [190]. Так, під впливом емоційних процесів може як підвищуватися достовірність та швидкість розпізнавання так і зменшуватися. Відповідно до цього, можна зробити висновок, що відтворення відповідних емоційних процесів у технічних

системах розпізнавання може забезпечити збільшення достовірності класифікації та зниження часової складності аналізу даних.

Для зручності використання в технічних системах, емоційні процеси пропонується ранжувати за інтенсивністю таким чином:

1) афект – емоційний процес, що характеризується короткочасністю і високою інтенсивністю, та супроводжується різко вираженою зміною проявів станів системи;

2) емоція – емоційний процес середньої тривалості, що відбиває суб'єктивне оцінне відношення до існуючих або можливих ситуацій та станів системи;

3) настрій – досить тривалий емоційний процес невисокої інтенсивності, що утворює емоційний фон для процесів, які мають місце у СС та навколишньому середовищі.

Окрім емоційних процесів, що ранжуються за інтенсивністю, слід визначити почуття, як емоційний процес людини, що відбиває суб'єктивне оцінне відношення до реальних або абстрактних об'єктів.

Таким чином, ефективність сприйняття інформації, так як і ефективність її відтворення, багато в чому залежать від наявності та рівня емоційної складової.

Слід зазначити, що вказані вище емоційні процеси повинні мати різні механізми застосування у технічних системах розпізнавання образів. Так механізм афекту, як найбільш інтенсивний та короткочасний процес-реакція на входні параметри може бути застосований у сенсорній пам'яті. Однак, при цьому слід зазначити, що швидкоплинність може стати додатковим джерелом помилок. Тобто афект, як додатковий інструмент при розпізнаванні, потребує обов'язкової перевірки позитивності його впливу на достовірність класифікації через співставлення з іншими емоційними складовими. Наприклад, співставлення афекту з відповідною емоцією.

Емоція, як емоційний процес середньої тривалості є основною складовою для побудови «емоційних» класифікаторів [185], так як ці складові безпосередньо пов'язані з короткочасною пам'яттю, в якій, згідно з моделлю, представленою на рисунку 2.3, здійснюється формування результату класифікації.

Настрій, як тривалий емоційний процес, пов'язаний з довготривалою пам'яттю повинен відображати характерні стани СС, що характеризуються одним або групою класів. При цьому, для всіх класів складова настрою повинна бути однаковою.

Відчуття, як характеристика суб'єктивної оцінки відношення до реальних або абстрактних об'єктів (образів), як і афект, може бути джерелом додаткових помилок. Цей механізм може бути використаний разом з евристичними і вводиться у систему як додаткові вагові характеристики до відповідних ознак образів.

На основі розробленої інформаційної моделі, представленої на рисунку 2.3, розроблений метод визначення послідовності обробки даних. Цей метод базується на представленні емоційної складової у вигляді функції від вхідних факторів, що характеризують стан СС або навколишнього середовища та визначається складовою  $EC$  в ситуаційно-подійній моделі гібридного розпізнавання. Згідно методу, рівень «емоційності» визначається ступенем зміни цієї функції, тобто ступенем зміни вхідних даних про стан СС. Таким чином, 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.

Сутність метода полягає в наступному. Представимо процес гібридного розпізнавання образів (рисунок 2.4), що містить  $k$  вхідних інформаційних потоків  $F_1-F_k$  і в якому враховуються емоційні процеси. В кожній пам'яті можуть бути введені відповідні емоційні процеси (сенсорна пам'ять –





провести ранжирування і селекцію ознак образів СС, що дозволяє усунути визначити чергу співставлення даних для отримання класифікації.

Питання побудови черги обробки ознак СС докладно розглянуті у розділі 4 при розробці інформаційної технології гібридного розпізнавання образів для консолідованої обробки даних в СС.

## **Висновки до розділу 2**

В розділі 2 розглянуті і вирішені раніше поставлені завдання 2-5, а саме: розроблений принцип консолідації неоднорідних даних в СС з різнорідними проявами перешкод і викривлень; обґрунтована концепція гібридного розпізнавання образів для консолідованої обробки даних в СС з різнорідними проявами перешкод і викривлень; формалізовано процес представлення даних та розроблена ситуаційно-подійна модель гібридного розпізнавання образів; розроблена модель і метод визначення рівня інформативності ознак в динаміці їхньої зміни. Основні висновки полягають в наступному.

Сформульовано принцип консолідації неоднорідних даних в СС, 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.

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 різнорідних перешкод і викривлень.

Відповідно до теорії інформації та міри Харкевича, інформація про частоту виявлення класифікуючої ознаки може бути отримана на основі співставлення результатів класифікації стану СС за кожним окремим образом з сукупності  $P_1-P_k$ . Іншими словами, вирішальне правило щодо класифікації

може бути побудоване наступним чином: стан  $CC$  відноситься до класу  $C_j$  з наявною передвизначеною сукупністю  $\{C\}$ , якщо частота такої класифікації є переважною за більшістю образів з множини  $\{P\}$ . Цей аспект дає обґрунтовану базу для розробки нового метода розпізнавання на основі обробки неоднорідних даних в  $CC$ , розробці якого приділена увага в подальшому дослідженні.

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

Аналіз модифікованої матричної інформаційної моделі на основі теорії інформації з використанням міри Харкевича та теорії зв'язку дозволило надати наукове обґрунтування тому, що для розробки системи критеріїв та методу визначення інформативності образів  $CC$  доцільно використання співвідношення ознак та перешкод, які мають місце в інформаційних каналах отримання даних сукупності образів  $\{P\}$  різної природи виникнення.

Розроблена концепція гібридного розпізнавання образів, що базується на модифікованій матричній інформаційній моделі  $CC$  та новому принципу консолідації неоднорідних ознак в  $CC$ , яка дозволяє забезпечити різні варіанти стратегії обробки даних та прийняття рішень, що підвищує гнучкість адаптації інформаційної технології під окремі практичні застосування.

Для гібридного розпізнавання образів, методологія побудови алгоритмів співставлення ознак та класифікації базується на двох базових підходах: сумісному або роздільному аналізі.

При спільному аналізі, множина образів  $CC$  різної природи виникнення поєднується до однієї сукупності та обробляється традиційними для комбінованого розпізнавання методами, що базуються на використанні АОО та КВП.

При роздільному аналізі ознак СС обробка даних та класифікація здійснюється за кожним образом окремо, а результати класифікації можуть використовуватися для уточнення остаточного результату розпізнавання.

Для забезпечення прийняття рішень на основі обмеженої кількості але найбільш інформативних ознак, розроблено ситуаційно-подійну модель гібридного розпізнавання образів, засновану на поданні характеристик складної системи 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.

Розроблений метод визначення інформативності образів складної системи, згідно якого, стаціонарні характеристики інформативності  $SIC_h$  визначаються апріорно і є незмінними характеристиками точності відтворення ознак об'єкта розпізнавання за кожним образом з  $\{P\}$ . Нестационарні характеристики інформативності  $NSIC_h$  змінюються відповідно до зміни зовнішніх умов та поточного рівня різнорідних 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.

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.



### РОЗДІЛ 3

## МЕТОДИ ГІБРИДНОГО РОЗПІЗНАВАННЯ ОБРАЗІВ ТА ІНФОРМАЦІЙНА МОДЕЛЬ ПІДТРИМКИ ПРИЙНЯТТЯ РІШЕННЯ ДЛЯ ВИЗНАЧЕННЯ СТАНУ СКЛАДНИХ СИСТЕМ З РІЗНОРІДНИМИ ПРОЯВАМИ ПЕРЕШКОД І ВИКРИВЛЕНЬ

### 3.1 Методи гібридного розпізнавання образів для СС з різномірними проявами перешкод і викривлень

Як вказувалося у попередньому розділі, концепція гібридного розпізнавання образів створила основу для побудови методів обробки неоднорідних даних та прийняття рішення, яка базується на двох альтернативних підходах: сумісному або роздільному аналізі 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.

#### *3.1.1 Спільний аналіз інформаційних ознак*

При спільному аналізі інформативних ознак віднесення стану СС до одного з класів здійснюється за одним образом, що містить інформацію від різних інформаційних потоків (а, відповідно, і ознак різної природи виникнення) [193]. Спільний аналіз проілюстрований на рисунку 3.1.

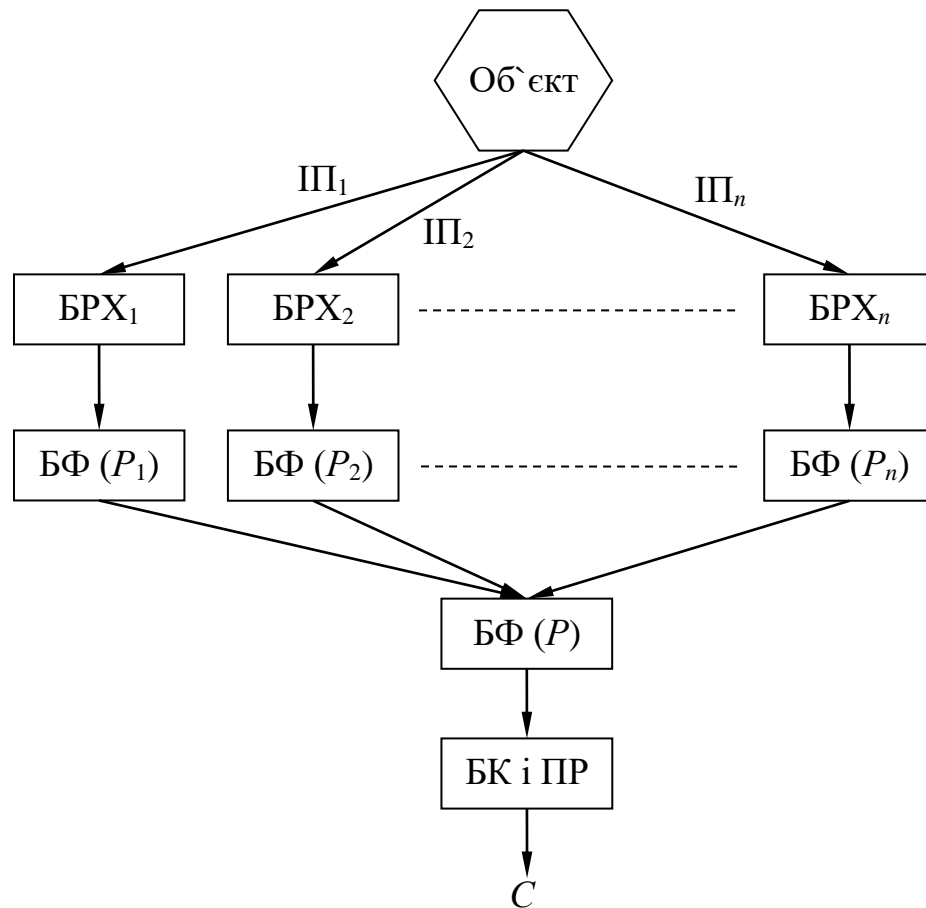


Рисунок 3.1 – Спільний аналіз ознак

Згідно рисунку, характеристики об'єкту розпізнавання за  $n$  інформаційними потоками  $\Pi_1$ - $\Pi_n$  реєструються сукупністю блоків реєстрації характеристик  $БРХ_1$ - $БРХ_n$ . Ці блоки представляють собою сукупність технічних засобів вимірювання детермінованих ознак  $СС$ , 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, що характеризують стан  $СС$ . Надалі, інформаційні ознаки цих образів поєднуються у спільний образ  $P$ , який надається до співставлення та формування рішення щодо класифікації. Результат класифікації представляється визначеним класом  $С$  до якого відноситься стан  $СС$ . Відповідно до (2.9) образ  $P$  буде містити повну інформацію про об'єкт

розпізнавання, що отримується за  $n$  інформаційними потоками (тобто містить ознаки, притаманні усієї сукупності образів  $P_1-P_n$ ). Таким чином, представлення об'єкту розпізнавання стає аналогічним традиційному (2.10).

Припустимо, що мається апріорно визначена сукупність класів стану  $CC C_1-C_m$ . Відповідно до репрезентативних характеристик цих класів, у просторі ознак необхідно виділити  $m$  областей, еквівалентних класам. В цьому випадку, якщо  $CC$ , 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 знайдено рішення щодо класифікації за критерієм [192]

$$P \in C_j, \max F_j(x_1, x_2, \dots, x_i). \quad (3.1)$$

Тобто стан  $CC$ , що характеризується спільним образом  $P$  може бути віднесений до класу  $C_j$  за умови максимального значення відповідної функції розподілу.

Однозначне класифікаційне рішення може бути прийнято, якщо для інших класів виконується умова [193]

$$\forall F_k(x_1, x_2, \dots, x_i), k \neq j \Rightarrow F_k(x_1, x_2, \dots, x_i) - F_j(x_1, x_2, \dots, x_i) < 0. \quad (3.2)$$

Відповідно, гранична умова для образів, що знаходяться на межі розподілу двох сусідніх класів буде мати наступний вигляд [193]:

$$F_k(x_1, x_2, \dots, x_i) - F_j(x_1, x_2, \dots, x_i) = 0. \quad (3.3)$$

В цьому випадку, при використанні спільного аналізу, алгоритми розпізнавання ґрунтуються на порівнянні тієї чи іншої міри близькості або міри схожості об'єкта розпізнавання з кожним класом. Тобто рішення задачі розпізнавання здійснюється класичним шляхом комбінованого розпізнавання [193]. При цьому, якщо міра близькості  $L$  даного 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 відповідного об'єкту до класу  $C_j$ , тобто виконується умова

$$P \in C_j, L(P, C_j) = \text{extr } L(P, C_i), \quad j = 1, \dots, m, \quad i \neq j. \quad (3.4)$$

Однак, з урахуванням того, що гібридне розпізнавання є подальшим розвитком комбінованого розпізнавання, де використовуються не тільки ознаки різної природи виникнення, але й різного характеру інформації, аналіз цих ознак з метою визначення ступенів подібності слід проводити різнопланово.

Для детермінованих ознак, в якості міри близькості доцільне використання середньоквадратичної відстані [193] між спільним образом  $P$  та репрезентативними образами  $P_{ji}$ , що характеризують відповідний клас  $C_j$ :

$$L(P, C_j) = \sqrt{k_j^{-1} \sum_{i=1}^{k_j} D^2(P, P_{ji})}. \quad (3.5)$$

Метод вимірювання відстані між образами  $D(P, P_{ji})$  обирається у кожному окремому випадку практичного застосування інформаційної технології гібридного розпізнавання образів.

Якщо ознаки образу мають різний вплив на прийняття остаточного рішення щодо класифікації, то залежність (3.5) може бути трансформована з

урахуванням вагових коефіцієнтів  $w_g$  ознак  $x_g$ ,  $g=1, \dots, N$ , спільного образу  $P$  та ознак  $x_{jig}$  репрезентативних образів  $P_{gi}$  класу  $C_j$ :

$$L(P, C_j) = \sqrt{k_j^{-1} \sum_{i=1}^{k_j} \sum_{g=1}^N w_g (x_g - x_{jig})}. \quad (3.6)$$

При використанні в алгоритмах розпізнавання ймовірнісних ознак, за міру близькості використовується характеристика ризику, пов'язаного з віднесенням стану  $CC$  до класу  $C_i$ ,  $i=1, \dots, m$ . Представимо, що завдані описи класів  $\{f_i(x), P(C_i)\}$ ,  $x=\{x_1 \dots, x_n\}$  і ризики коректних та помилкових класифікацій представляють елементи платіжної матриці виду

$$W = \begin{pmatrix} w_{11} & w_{12} & \dots & w_{1m} \\ w_{21} & w_{22} & \dots & w_{2m} \\ \dots & \dots & \dots & \dots \\ w_{m1} & w_{m2} & \dots & w_{mm} \end{pmatrix}. \quad (3.7)$$

Головна діагональ матриці  $W$  містить коефіцієнти, що характеризують втрати за умови коректної класифікації, а інші коефіцієнти характеризують втрати за умови помилкової класифікації. Якщо  $w_{ii} < 0$ ,  $i=1, \dots, m$ , то такі негативні втрати можна вважати як виграш при коректних рішеннях щодо класифікації.

Представимо, що в результаті вимірювання характеристик  $CC$  встановлено, що апостеріорні значення ознак образу  $P$  складають  $x_1 = x_1^{apost}$ ,  $x_2 = x_2^{apost}$ ,  $\dots$ ,  $x_n = x_n^{apost}$ . Позначимо цей етап як  $stage_n$ . Тоді значення ризику, пов'язаного з рішенням виду  $P \in C_j$  на етапі  $stage_n$ , буде визначатися як

$$W(P \in C_j | stage_n) = W(C_j / stage_n) = \sum_{i=1}^m w_{ij} p(C_i / stage_n), \quad (3.8)$$

де умовна апостеріорна імовірність того, що  $P \in C_i$ , у відповідності з теоремою гіпотез або формулою Байєса [194],

$$p(C_i / stage_n) = \frac{p(C_i) f_i(x_1^{apost}, x_2^{apost}, \dots, x_n^{apost})}{\sum_{i=1}^m p(C_i) f_i(x_1^{apost}, x_2^{apost}, \dots, x_n^{apost})}. \quad (3.9)$$

Очевидно, що у загальному випадку, рішення виду  $P \in C_j$  приймається, якщо

$$W(C_j / stage_n) = \min W(C_i / stage_n). \quad (3.10)$$

Для характеристик СС, які базуються на логічних ознаках не використовується міра близькості. Коли побудовано опис класів на мові логічних ознак у вигляді відповідних булевих співвідношень еквівалентності або імплікацій, при підстановці в ці співвідношення значень ознак, що характеризують стан СС, рішення щодо класифікації є результатом логічних операцій у створеному логічному висловлюванні. Використання такого різновиду ознак природньо легко піддається алгоритмізації та програмуванню.

В алгоритмах розпізнавання, що базуються на використанні структурних (лінгвістичних) ознак, поняття міри близькості також може не використовуватися. Коли побудовані мови, що описують класи у вигляді сукупностей висловлювань (речень), що характеризують структурні особливості об'єкта, що відносяться до кожного класу, то розпізнавання

невідомого об'єкта здійснюється ідентифікацією речення, що описує цей об'єкт, з одним з репрезентативних речень, що є елементами опису відповідного класу.

Об'єктивно, не враховуючи різного характеру інформації, що мають ознаки  $CC$ , їх поєднання при спільному аналізі, відповідно до залежності (2.9), призводить до збільшення розмірності простору ознак  $i$ , як наслідок, до збільшення достовірності розпізнавання. Цей факт підтверджується на основі статистичної теорії розпізнавання [109]. Статистичний підхід ґрунтується на ідеї, що початковий простір об'єктів представляє собою ймовірнісний простір, а ознаки об'єктів – це випадкові величини задані на ньому. На основі експертної оцінки слід висунути статистичну гіпотезу про розподіл ознак, а точніше про залежність класифікуючих ознак від інших. Статистична гіпотеза, як правило, може бути представлена у якості параметрично заданої множини функцій розподілу ознак. Тоді класифікаційне рішення може бути прийняте на основі критеріїв (3.1)-(3.4). Типовою і класичною статистичною гіпотезою є гіпотеза про нормальність цього розподілу [59, 109, 110, 140, 195]. На практиці, ця гіпотеза може бути перевірена на прецедентних даних. З цією метою, обирається деякий розподіл з апріорно заданої множини розподілів та робиться оцінка надійності такого вибору. Оцінку надійності отримання коректної класифікації, в цьому випадку, зручно здійснювати за значенням довірчого інтервалу. При цьому, відповідно до асимптотичного обґрунтування ймовірнісних методів [100, 130, 196], апріорно встановлюється функціонал якості вибору розподілу (довірчий інтервал) і показується, що при збільшенні числа прецедентів, рішення з імовірністю що прямує до 1 стає вірним у сенсі цього функціоналу. При цьому довірчий інтервал прагне 0. Такий підхід дозволяє, у випадку спільного аналізу з метою непрямої оцінки достовірності розпізнавання, використовувати значення довірчого інтервалу. Тоді, задача класифікації об'єкта

розпізнавання за визначеною сукупністю класів може бути здійснена з використанням підходу, запропонованого Р. Фішером [197].

Представимо, що у деякому  $n$ -вимірному просторі ознак  $\Omega$  мається деякий об'єкт, який характеризується образом  $P$  та може бути класифікований за одним з  $m$  класів. Об'єкти цих класів з'являються з невідомими імовірностями  $p_1-p_m$ . В процесі побудови інформаційної технології розпізнавання ці ймовірності можуть бути оцінені за навчальними вибірками. Надалі, в процесі навчання, за визначеними ймовірностями оцінюються відповідні математичні очікування  $\mu_m$  та матриці коваріації  $\Sigma_m$ . Якщо нормальні розподіли за  $m$  класами позначити як  $N(x, \mu_m, \Sigma_m)$ , то достовірний результат класифікації об'єкта розпізнавання за класом  $m$  може бути визначений як

$$\max p_m N(x, \mu_m, \Sigma_m) \Rightarrow P \in C_m. \quad (3.11)$$

Достовірність такого результату класифікації підтверджується тим, що за умови спільного аналізу ознак, імовірності  $p_m$  оцінюються з великою точністю за законом великих чисел, а характеристики  $\mu_m$  та  $\Sigma_m$ , відповідно, по частинам вибірок класів  $C_m$  за допомогою стандартних статистик, для яких, як вже було зазначено вище, довірчі інтервали прагнуть нуля при збільшенні прецедентів за якими здійснюється оцінка.

З іншого боку, використання надмірної кількості інформативних ознак створює негативні наслідки при розпізнаванні, що пов'язані з перенавчанням. Обґрунтування цього аспекту можна знайти в основі глибокої статистичної теорії розпізнавання образів Вапніка-Червоненкіса [109]. Згідно цієї теорії, у процесі навчання фіксується множина функцій

$$F \subset \{f : P \rightarrow C\}, \quad (3.12)$$



де  $P$  – множина образів об'єкта розпізнавання;

$C=\{0,1\}$  – множина класів.

Навчальна вибірка  $X^l$ , елементи якої послідовно та незалежно вибираються з множини  $P \times C$  згідно деякому невідомому розподілу  $\rho$ , позначається як

$$X^l = \{x_i, y_i\}_{i=1}^l. \quad (3.13)$$

В результаті використання навчальної вибірки (3.13) обирається деяка функція  $f \in F$ , яка і буде використовуватися в якості класифікатора.

Згідно теорії Вапніка-Червоненкіса [109], імовірність того, що помилка на контролі істотно перевищить помилку при навчанні, може бути оцінена, у випадку виконання умови

$$P(v(\mu(X^l), X^k) > v(\mu(X^l), X^l) + \varepsilon), \quad (3.14)$$

де  $X^k$  – контрольна вибірка.

Контрольна вибірка також незалежно обирається по  $\rho$  [109], при чому, ця оцінка отримується для будь-якого розподілу на  $P \times C$  наступним чином:

$$P(v(\mu(X^l), X^k) > v(\mu(X^l), X^l) + \varepsilon) \leq \Delta^F(2l) \cdot 1.5e^{-\varepsilon^2 l}, \quad (3.15)$$

Для простоти можна прийняти  $k=1$  (тобто довжина контролю дорівнює довжині навчання).

Фактично, функціонал (3.15) від (3.14) відрізняється двома множниками у правій частині. Множник  $1.5e^{-\varepsilon^2 l}$  експоненціально прагне нуля, якщо довжина вибірки прагне нескінченності. Інший множник  $\Delta^F(2l)$  у

теорії Вапніка-Червоненкіса називається складнішим. Для будь-якої підмножини  $M^{2l}$  об'єктів потужності  $2l$  вводиться  $S^F(M^{2l})$  для позначення множини різноманітних розбиттів  $M^{2l}$  на класи, що визначаються функціями  $F$ . Тоді

$$\Delta^F(2l) = \max S^F(M^{2l}). \quad (3.16)$$

Таким чином,  $\Delta^F(2l)$  характеризує різноманіття множини функцій  $F$ , обмежених на множині довжини  $2l$  та залежить тільки від множини функцій з яких визначається результат і не залежить від алгоритму його отримання.

Якщо припустити, що  $\Delta^F(2l)$  при  $2l \rightarrow \infty$  збільшується повільніше, ніж будь-яка експоненційна складова, то можна стверджувати, що

$$P(v(\mu(X^l), X^k) > v(\mu(X^l), X^l) + \varepsilon) \rightarrow 0, k = l - \infty. \quad (3.17)$$

Це забезпечує мале відхилення помилки на контролі від помилки на навчанні.

Таким чином, можна зробити висновок, що зменшення величини  $\Delta^F(2l)$  призводить до зменшення додаткової помилки, що додається до помилки при навчанні. З іншого боку, якщо складність  $F$  мала, то неможливо обрати  $f \in F$  з малою помилкою на навчальній вибірці. В цьому випадку складається ситуація, коли від збільшення складнісного множника  $\Delta^F(2l)$  збільшується достовірність прийняття рішення щодо класифікації, але надмірне його збільшення призводить до перенавчання.

Відповідно, вищенаведені докази показують, що для отримання достовірного результату розпізнавання, з одного боку слід збільшувати кількість даних, що надаються для співставлення та отримання рішення, з

іншого – надмірне його збільшення призводить до негативних наслідків перенавчання та збільшення часової складності аналізу даних та прийняття рішення.

Виходячи з вищезгаданого, для отримання достовірного результату класифікації за мінімальні часові втрати слід у будь-який момент часу здійснювати селекцію раціональної сукупності ознак, що забезпечить раціональну кількість функцій  $F$  за якими може бути отриманий результат.

Також, з метою зменшення часової складності аналізу даних та прийняття рішень, слід прагнути до ідентичності коваріаційних матриць наявної множини класів. В цьому випадку, може бути використана евристика стосовно лінеаризації вирішальних функцій [100]. Однак, для зменшення впливу можливих завад у випадку викривлення простору ознак об'єктів розпізнавання слід прагнути до малої коваріації між ними. В цьому випадку залежність ознак буде мінімальна, що призведе до мінімізації ймовірності однакового прояву завад на різних групах ознак (різних образах, що характеризують стан СС). Таким чином, додаткова помилка не буде набувати адитивного прояву від однієї завади на різних ознаках, що використовуються для класифікації.

Для здійснення селекції інформативних ознак в методах, що базуються на спільному аналізі, пропонується використання коваріаційної матриці. При цьому, перевагу в аналізі слід надавати її діагональним елементам, так як вони знаходяться у малій взаємозалежності. Зменшення кількості залежних ознак спрощує семантику прийняття рішень щодо класифікації та прийняття рішень. Крім цього, використання слабо пов'язаних між собою ознак дозволяє знизити вірогідність помилки при розпізнаванні, так як такі ознаки породжуються різною фізичною природою виникнення (відповідно, різними за принципами засобами реєстрації характеристик). Іншими словами, завада, яка впливає на значення однієї ознаки з малою долею ймовірності також буде відбиватися на значеннях слабо пов'язаних ознак.

Таким чином, у випадку використання гібридного розпізнавання образів зі спільним аналізом інформаційних ознак, представлення стану СС не тільки зводиться до класичного вигляду (вектору у просторі ознак у якого ознаки є координатами), а й значно розширюється кількість інформаційних ознак, особливо тих, що мають малий взаємний вплив. Все це сприяє забезпеченню високої достовірності розпізнавання.

### *3.1.2 Роздільний аналіз інформаційних ознак*

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, відповідно до (2.9) і (2.10). Тобто, прийняття рішення здійснюється в результаті окремої обробки консолідованих ознак кожного образу [198, 199]. Класифікаційне рішення може прийматися за одним або декількома образами з сукупності  $\{P\}$ .

При роздільному аналізі інформаційних ознак рішення щодо класифікації стану СС приймається за кожним з наявних образів. Ілюстрація принципу роздільного аналізу інформаційних ознак представлена на рисунку 3.2.

При роздільному аналізі, відомості про об'єкт розпізнавання за  $n$ -інформаційними каналами реєструються сукупністю блоків реєстрації характеристик  $БРХ_1$ - $БРХ_n$ . 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, що характеризують стан СС. Априорно відомо, що кожен з цих образів за різними сукупностями ознак характеризують один й той самий об'єкт розпізнавання. Відповідно, у будь-який момент часу, за кожним з цих образів можна здійснити класифікацію об'єкта розпізнавання та його віднесення до одного з передвизначених класів.

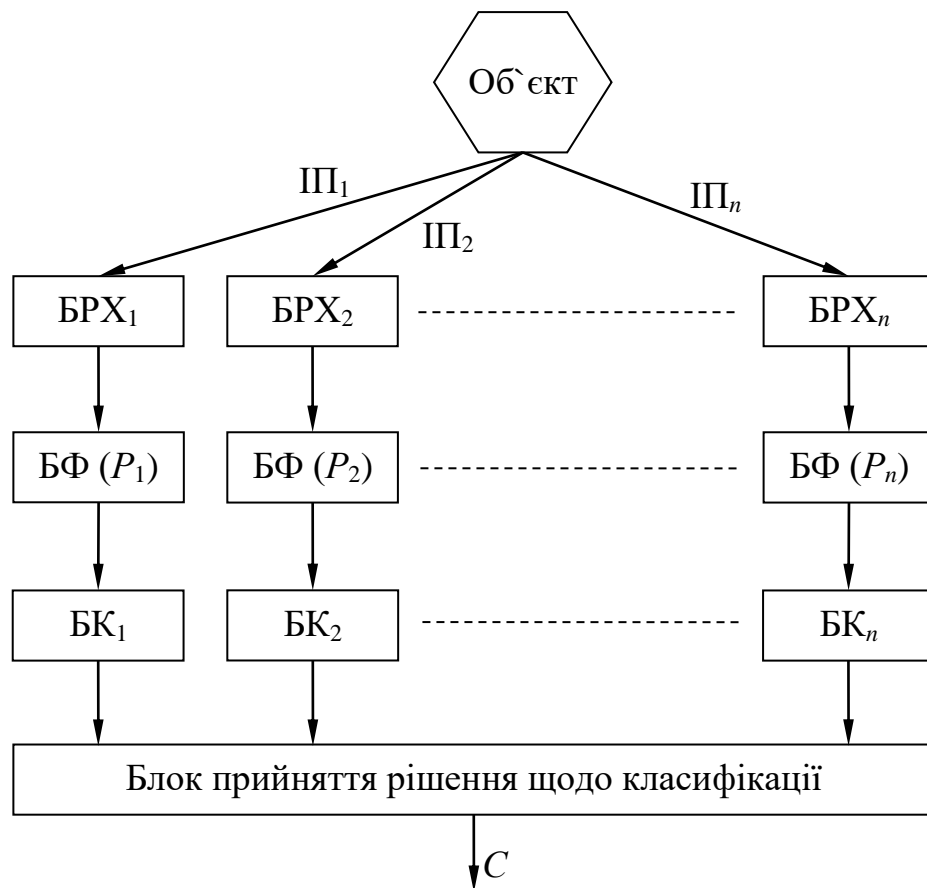


Рисунок 3.2 – Роздільний аналіз ознак

Використання роздільного аналізу дозволяє ефективно реалізувати задачі таксономії та здійснити розпізнавання об'єкту за принципом кластеризації. Також, саме роздільний аналіз дозволяє отримати максимальні переваги від гібридного розпізнавання, тому в подальшому дослідженні, саме на основі роздільного аналізу здійснюється розробка нових моделей і методів консолідованої обробки даних та прийняття рішення.

При цьому, роздільний аналіз має також істотний недолік, пов'язаний з паралельною обробкою та окремої класифікації за різними образами одного об'єкта розпізнавання. Звичайно, в ідеальному випадку, при гарній роздільності класів та відсутності перешкод і викривлень даних, що характеризують стан СС, за усіма образами повинне отримуватися однаковий результат класифікації до одного й того самого класу. Але, за реальних умов,

може виникати ситуація, коли за різними образами отримується різні результати класифікації. Надалі цей момент враховано при розробці методу гібридного розпізнавання та інформаційної моделі підтримки прийняття рішення у випадку багатоальтернативного результату класифікації.

### *3.1.3 Метод гібридного розпізнавання образів, оснований на роздільному аналізі ознак складної системи*

Роздільний аналіз покладено в основу розробки нового методу гібридного розпізнавання [198, 199], який проілюстрований на рисунку 3.3.

На етапі  $E_1$  здійснюється реєстрація характеристик СС з використанням різноманітних технічних засобів вимірювання. Згідно розробленої ситуаційно-подійної моделі гібридного розпізнавання (2.11), підлягають реєстрації не тільки ознаки СС, а й необхідні параметри зовнішнього середовища. На цьому етапі також здійснюється нормування отриманих характеристик з метою спрощення їхньої подальшої обробки.

На наступному етапі  $E_2$ , на основі визначених характеристик СС та навколишнього середовища, здійснюється формування сукупності  $k$  образів об'єкту розпізнавання (елементи  $P_1-P_k$ ), на основі інформації різної природи виникнення. Таким чином, інформація про стан СС представляється у традиційному, для гібридного розпізнавання, формі (2.1). Також, відповідно до моделі (2.11) здійснюється визначення їх характеристик інформативності  $SICH_k$  і  $NSICH_k$  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.

$C_1, C_2, C_3, \dots, C_m$  ілюструють області розподілу ознак еталонних образів попередньо визначених  $m$  класів, до яких може бути віднесений об'єкт розпізнавання. Для здійснення класифікації, ці класи повинні бути апріорно визначені.

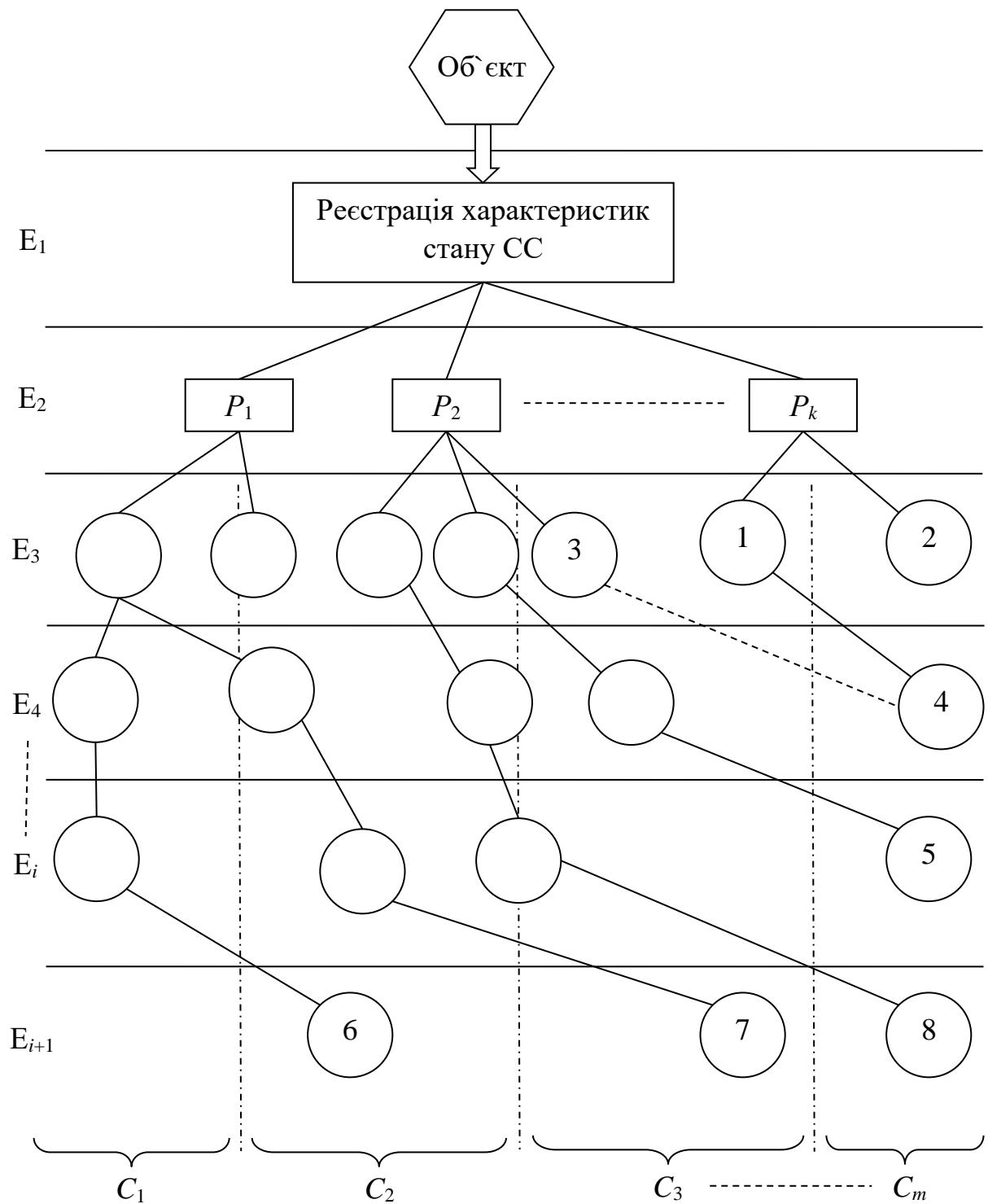


Рисунок 3.3 – Метод гібридного розпізнавання образів  
з роздільним аналізом ознак СС

Позначеннями  $E_2, E_3, E_4, \dots, E_i, E_{i+1}$  проілюстровані етапи порівняння інформативних ознак образів  $P_1-P_k$  з ознаками еталонних образів та

визначення характеристик відстані. На цих етапах проводять роздільний аналіз цих образів шляхом порівняння їхніх ознак з ознаками еталонних образів та визначають відстані цих образів від еталонних. Якщо, на поточному етапі, визначені відстані перевищуватиме наперед задане максимально припустиме значення, то здійснюється наступний етап аналізу з іншими ознаками наявної сукупності образів  $SS$ . На рисунку 3.3. представлений приклад розподілу образів  $SS$  за передвизначеними класами. Так, згідно цього прикладу, 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  $E_4$ , подальшу обробку інформативних ознак припиняють і достовірним результатом буде вважатися віднесення стану  $SS$  до класу  $C_m$ . Подальший аналіз ознак (на рисунку 3.3 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 часової складності аналізу ознак  $SS$ . В подальшому, аналіз інформативних ознак може тільки підтвердити зроблену класифікацію (наприклад визначення відстані 5, що також вказує на клас  $C_m$ ), або вказати на наявність викривлень інформаційного простору ознак та здійснення помилкового віднесення об'єкту розпізнавання до класів  $C_2$  і  $C_3$  за визначеними відстанями 6 і 7, відповідно.

Згідно представленого методу, на кожному етапі може бути прийняте остаточне рішення за наступними вирішальними правилами:



1) стан  $CC$  відноситься до того класу (з наявної сукупності передвизначених), 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

$$L = \begin{bmatrix} l(P_1, C_1) & l(P_1, C_2) & \dots & l(P_1, C_m) \\ l(P_2, C_1) & l(P_2, C_2) & \dots & l(P_2, C_m) \\ \dots & \dots & \dots & \dots \\ l(P_k, C_1) & l(P_k, C_2) & \dots & l(P_k, C_m) \end{bmatrix}, \quad (3.18)$$

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

Тоді, для першого вирішального правила, рішення щодо класифікації об'єкту розпізнавання здійснюється за умови виконання критерію

$$\exists P_k \in C_m, l(P_k, C_m) = \min L. \quad (3.19)$$

Тобто, рішення щодо класифікації отримується шляхом пошуку мінімального значення елемента матриці  $L$  який однозначно вказує на мінімальну відстань одного з образів до відповідного класу.

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

$$P_k \in \{P\}' \mid \{P\}' \subset \{P\}, l(P_k, C_m) \leq l^*, \quad (3.20)$$

де  $l^*$  – апріорно задане мінімальне значення відстані;

$l(P_k, C_m)$  – відстань образу  $P_k$  до класу  $C_m$ .

У випадку хорошої роз'єднувальності класів, якщо тільки один образ з множини  $\{P\}$  відповідає критерію (3.20), то цей приватний випадок аналогічний розглянутому вище. У випадку, коли декілька образів відповідають критерію (3.20), то в цьому випадку необхідне проведення додаткового аналізу отриманої множини  $\{P\}'$ . Тоді можливі дві ситуації:

– усі образи сукупності  $\{P\}'$  вказують на віднесення об'єкту розпізнавання до одного з передвизначених класів;

– образи сукупності  $\{P\}'$  вказують на віднесення об'єкту розпізнавання до різних класів.

Перша ситуація більш проста і дозволяє одразу сформулювати остаточне рішення щодо класифікації. Другий випадок, більш складний та вимагає прийняття класифікаційного рішення на основі додаткового аналізу інформативності отриманих класифікацій. Так, у відповідності до поточних умов спостереження за об'єктом розпізнавання для кожного образу з множини  $\{P\}'$  є визначена характеристика інформативності, за методом, запропонованим у підрозділі 2.4. З урахуванням вагових коефіцієнтів для кожного образу, що входить до репрезентативної сукупності, рішення щодо класифікації приймається за принципом максимуму характеристики

$$P_n \in C_g, \max \frac{Ich_n \cdot l(P_n, C_g)}{\sum_{i=1}^k \sum_{j=1}^m Ich_i \cdot l(P_i, C_j) - Ich_n \cdot l(P_n, C_g)}, \quad (3.21)$$

де  $Ich_n$  – характеристика інформативності для  $n$ -го класу  $P_n$  ( $n \in i=1, 2, \dots, k$ );

$Cg$  – 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 вагового коефіцієнту і відповідної відстані до визначеного класу на суму усіх інших добутків.

Для третього вирішального правила, пропонується використання зваженого голосування коли об'єкт розпізнавання відноситься до визначеного класу в тому випадку, коли така класифікація підтверджується за більшістю образів. Тобто для кожного з  $m$  визначених класів визначається характеристика  $N(C_m)$  яка відповідає кількості образів, що вказують на класифікацію стану  $CC$  до цього класу. Остаточно, об'єкт розпізнавання відноситься до того, класу для якого характеристика  $N_m$  має максимальне значення. На рисунку 3.4 представлений процес пошуку груп ідентичних класифікацій [200], в результаті якого отримується масив  $N[m]$  результатів класифікації у вигляді

$$N = \begin{bmatrix} N(C_1) \\ N(C_2) \\ \dots \\ N(C_m) \end{bmatrix}, \quad (3.22)$$

де  $N(C_1)$ - $N(C_m)$  – характеристики кількості ідентичних класифікацій по кожному з  $m$  класів.

Згідно рисунку 3.4, спочатку здійснюється ініціалізація масиву  $N$ , кожен елемент якого буде відповідати кількості образів, за якими здійснюється віднесення об'єкту розпізнавання до відповідного класу. Тобто, ці елементи характеризують  $N(C_1)$ - $N(C_m)$ . Очевидно, що розмірність масиву  $N$  дорівнює кількості класів  $m$ , за якими може бути здійснений розподіл. За

умовчанням, кожному елементу масиву привласнюється початкове нульове значення. В процесі виконання алгоритму, вони будуть збільшуватися до шуканих значень.

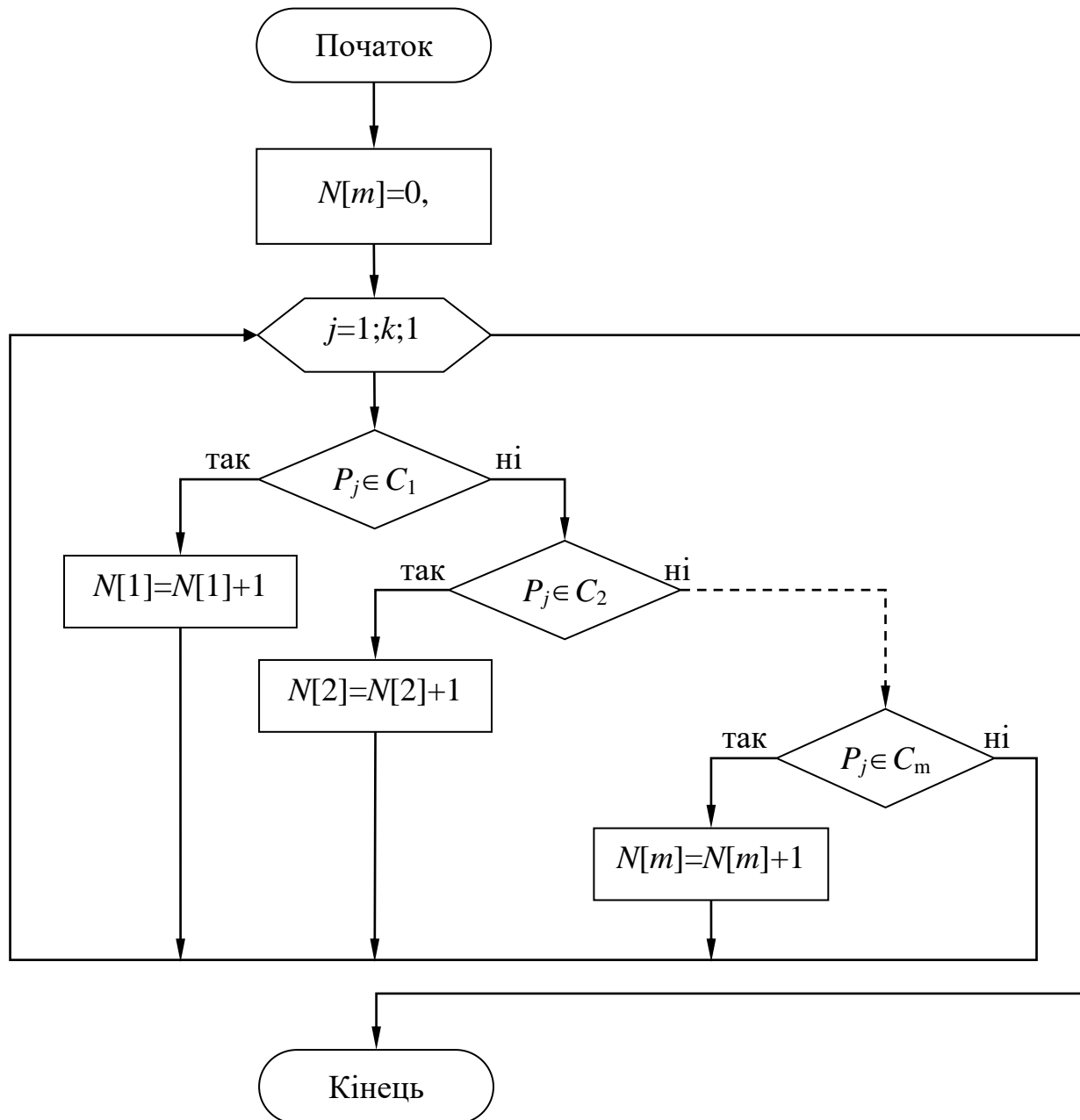


Рисунок 3.4 – Метод гібридного розпізнавання на основі пошуку груп образів з ідентичною класифікацією

Надалі, організовується цикл для перебирання усіх образів, за якими здійснюється класифікація. В цьому циклі, у відповідності до отриманої класифікації за поточним образом, збільшується один з елементів масиву  $N$ ,

що відповідає обраному класу. Таким чином, після завершення циклу кожен елемент масиву  $N$  буде містити кількість образів, що класифікують об'єкт розпізнавання до кожного класу з наявної сукупності. Діапазон значень, які можуть приймати елементи масиву

$$N[i] \in [0; k], i = 1, 2, \dots, m, \quad (3.23)$$

де  $k$  – кількість образів об'єкта розпізнавання;

$m$  – кількість передвизначених класів.

На основі отриманого масиву  $N$ , з використанням тривіального алгоритму пошуку максимального елемента, може бути знайдено остаточне рішення щодо класифікації, відповідно до критерія [200]

$$P_k \in C_m, \max N(C_m). \quad (3.24)$$

#### *3.1.4 Модифікований метод гібридного розпізнавання на основі пошуку груп образів з ідентичною класифікацією*

Фактично, у представленому вище методі гібридного розпізнавання, елементи масиву (3.22) характеризують ступень достовірності здійснення класифікації стану  $CC$ . Якщо в результаті використання алгоритму зворотного сортування матриці  $N$  отримується матриця  $N^S$  то її елементи будуть розміщатися у порядку зменшення достовірності відповідної класифікації

$$N^S = \begin{bmatrix} \max N(C_m) \\ \dots \\ \dots \\ \min N(C_m) \end{bmatrix}. \quad (3.25)$$

Тобто, найбільш достовірне рішення щодо класифікації буде відповідати елементу  $\max N(C_m)$ , а найменш достовірне – елементу  $\min N(C_m)$ .

Особливий інтерес мають групи образів, що формують характеристики  $N(C_m)$  [201]. З погляду на те, що кількість образів, за якою стан  $CC$  відноситься до визначеного класу, напряду характеризує достовірність отриманого результату, то очевидно, що групи цих образів за характеристикою  $N(C_m)$  ранжуються за ознакою інформативності. Тобто, образи, що входять у найчисельнішу групу, за поточних умов спостереження об'єкта розпізнавання мають найвищу інформативність. Таким чином, визначення характеристик  $N(C_m)$  дозволяє вирішити додаткову задачу ранжування образів за ознакою інформативності. У подальшому, образи, що утворюють групи з меншими характеристиками  $N(C_m)$  та дають інші результати класифікації, можуть бути усунені від розпізнавання. В цьому випадку, на основі методу, проілюстрованому на рисунку 3.4, пропонується модифікований метод пошуку груп образів з ідентичною класифікацією, представлений на рисунку 3.5 [201].

Згідно модифікованого методу, спочатку ініціалізується масив  $N$ , кожен елемент якого буде містити кількість образів, за якими здійснюється класифікація до визначеного класу. На початку виконання алгоритму елементам цього масиву присвоюються нульові значення.

Надалі, здійснюється ініціалізація  $m$  масивів  $GC_1[k]$ ,  $GC_2[k]$ , ...,  $GC_m[k]$  в кожному з яких будуть міститися індекси образів, за якими об'єкт розпізнавання відноситься до одного з  $m$  передвизначених класів. Тобто, ці масиви в результаті виконання алгоритму будуть містити групи образів, що відтворюють ідентичну класифікацію до кожного з  $m$  класів наявної сукупності  $\{C\}$ .

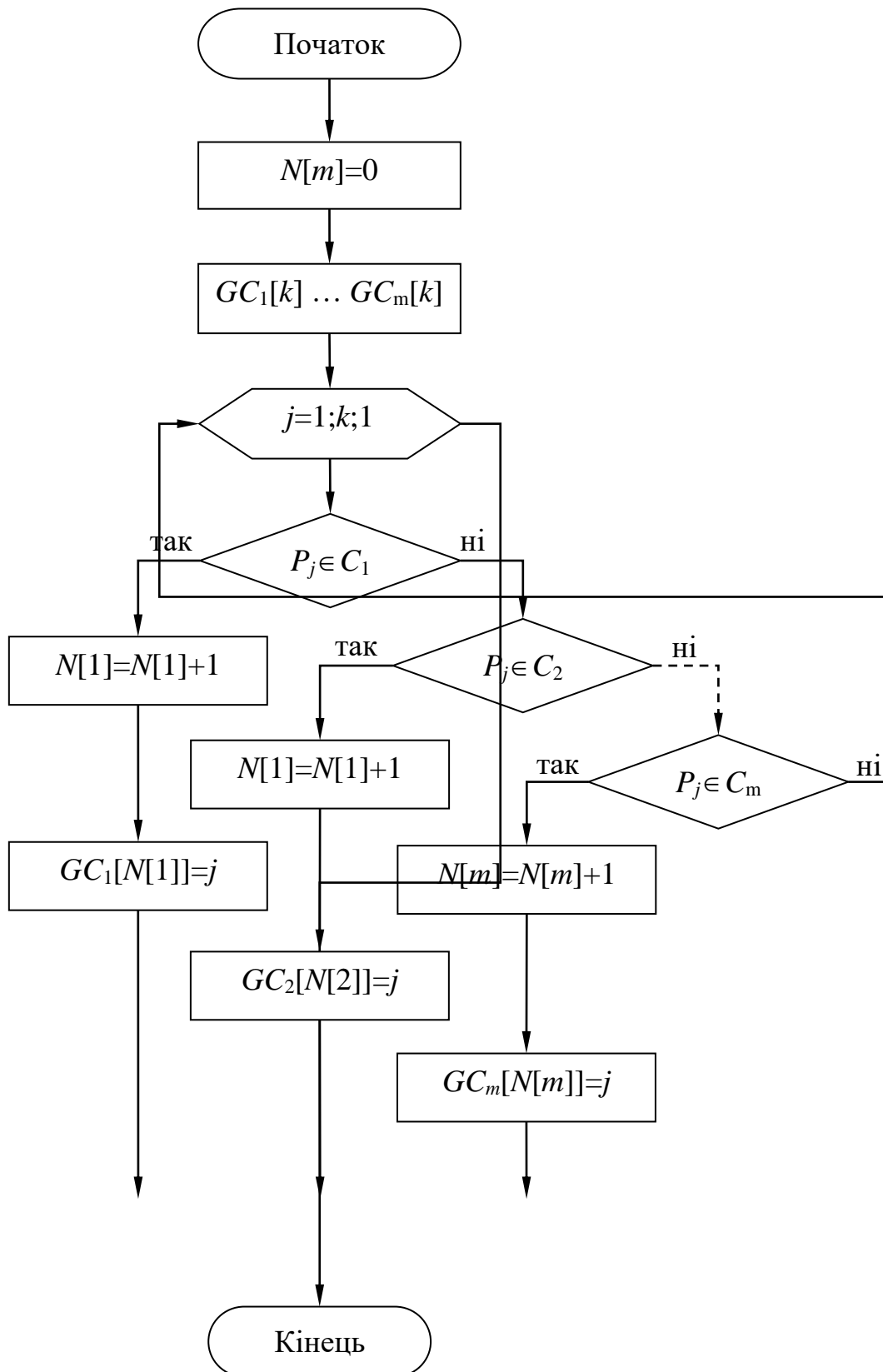


Рисунок 3.5 – Модифікований метод гібридного розпізнавання на основі пошуку груп образів з ідентичною класифікацією

В процесі подальшого циклічного перебору усієї сукупності  $k$  образів та, у випадку класифікації до класу  $C_g$ ,  $g \in \overline{1..m}$ , здійснюється інкрементування елементу масиву  $N[m]$  та, відповідно до отриманого збільшеного значення, здійснюється розміщення індексу образу у масиві  $GC_g[k]$ .

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

Критерій максимальної достовірності розпізнавання має вигляд [201]

$$D_{\max} \Leftrightarrow \begin{cases} \forall P_i \in \{P\}, P_i \in C_g, \\ N[g] = k, \\ \forall i \neq g, N[i] = 0. \end{cases} \quad (3.26)$$

Однак, за реальних умов наявності різноманітних перешкод і викривлень даних СС, критерій (3.26) є асимптотичним, так як недостатній опис об'єкта розпізнавання, наявні викривлення відображення його характеристик у просторі образів, а також недосконалість систем реєстрації та вимірювання цих характеристик, призводить до неоднозначності класифікації. В такому випадку, критерій достовірності розпізнавання визначатися за умови [201]

$$D \Leftrightarrow \begin{cases} \exists P_i \in \{P\}, P_i \in C_g, \\ N[g] < k, \\ \forall i \neq g, N[i] \neq 0. \end{cases} \quad (3.27)$$



Тобто, найбільш достовірне рішення щодо класифікації стану  $CC$  відповідає максимальному значенню характеристики  $N(C_m)$  (це максимальне значення серед елементів масиву  $N$ ).

Таким чином, після виконання модифікованого алгоритму пошуку груп ідентичних класифікацій, крім характеристик  $N(C_m)$  визначаються групи  $GC_1$ - $GC_m$  образів, що дають ідентичну класифікацію за сукупністю класів  $C_1$ - $C_m$ . За максимумом характеристики  $N(C_m)$  приймається остаточне рішення щодо розпізнавання об'єкта інформаційного процесу, а аналіз відповідної сукупності  $GC_m$  дає інформацію про сукупність найбільш інформативних образів, за якими здійснюється найбільш достовірна класифікація.

Визначення груп найбільш інформативних образів, за отриманими характеристиками  $GC_1$ - $GC_m$ , дає можливість здійснити селекцію цих образів та, в подальшому, усунути від розпізнавання ті групи образів, що за поточних зовнішніх умов та рівня перешкод і викривлень, не дозволяють отримати достовірного результату класифікації.

### *3.1.5 Інформаційна модель підтримки прийняття рішення у випадку отримання багатоальтернативного результату*

Аналіз запропонованого методу гібридного розпізнавання образів вказує на те, що за реальних умов здійснення класифікації може бути отримане багатоальтернативне рішення, коли декілька груп ідентичної класифікації будуть вказувати на різні альтернативи остаточного рішення. Для підтримки прийняття рішення в такому випадку була розроблена інформаційна модель, яка враховує той факт, що кожній характеристиці  $N(C_m)$  однозначно співставляється відповідна група образів з ідентичною класифікацією. Тоді, сортування  $N$  та отримання (3.25) призводить до упорядкування груп образів з ідентичною класифікацією у відповідності до рівня достовірності розпізнавання. З урахуванням цього, модель бути мати вигляд [201]:

$$N^S = \begin{bmatrix} \max N(C_m) \\ \dots \\ \dots \\ \min N(C_m) \end{bmatrix} \Leftrightarrow \begin{bmatrix} GC_m^1 | \max N(C_m) \\ \dots \\ \dots \\ GC_m^m | \min N(C_m) \end{bmatrix} \Leftrightarrow \begin{bmatrix} \{P_i\}^1, i \in GC_m^1 \\ \dots \\ \dots \\ \{P_j\}^m, j \in GC_m^m \end{bmatrix}, \quad (3.28)$$

де  $GC_m^1 | \max N(C_m)$  – 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);

$\{P_i\}^1$  – сукупність образів, індекси  $i$  яких входять до групи  $GC_m^1$  (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).

Слід зазначити, що метод гібридного розпізнавання на основі пошуку груп ідентичної класифікації та інформаційна модель підтримки прийняття рішення у випадку отримання багатоальтернативного рішення можуть також бути використані для апріорного визначення груп образів СС, що за поточних умов спостереження, дозволяють отримати найдостовірніший результат класифікації. Це дозволяє здійснити апріорну настройку інформаційної технології гібридного розпізнавання для обробки неоднорідних даних в СС. З цією метою, під час синтезу інформаційної технології, за різних зовнішніх для СС умов пред'являють репрезентативні образи. В результаті, для кожних умов визначаються характеристики  $N(C_m)$  і

$GC_m$ . Надалі, відповідно до (3.25), здійснюється ранжування груп образів за рівнем інформативності. В подальшому, в системі фіксується, які групи образів більш інформативні за поточних умов, або менш інформативні. Ступінь інформативності образів визначається в кількісному виразі чисельністю відповідної групи, до якої потрапив цей образ. В результаті, малоінформативні образи усуваються від розпізнавання. Таким чином, у будь-який момент часу, процес розпізнавання базується на аналізі меншої кількості образів найбільш інформативних груп, що забезпечує зменшення кількості співставлень ознак  $CC$  та, відповідно, зниженню часової складності процесу обробки даних та прийняття рішення, при одночасному збереженні високого рівня достовірності результату.

Аналіз принципу роздільного аналізу інформаційних ознак, та методу гібридного розпізнавання вказує на те, що край важливою є реалізація методів консолідації та селекції інформативних образів і їхніх ознак, з метою зменшення кількості співставлень даних при аналізі стану  $CC$  та мінімізації часової складності процесу класифікації. В подальшому це питання розглядається докладно.

### **3.3 Метод консолідації образів $CC$ та їх селекція на основі характеристик інформативності**

При гібридному розпізнаванні використовується велика кількість інформаційних образів (відповідно, і ознак) стану  $CC$  та навколишнього середовища (елементи  $P$  та  $EC$  раніше запропонованої ситуаційно-подійної моделі), які дозволяють отримати різну достовірність класифікації за різних умов спостереження за об'єктом розпізнавання. При цьому, один образ (або сукупність образів) дають кращий результат класифікації за одних умов, а інші – за других. Крім цього, додаткові образи можуть бути використані для остаточного уточнення результату класифікації. При цьому, для класифікації

може бути використана як уся сукупність образів, з обов'язковим визначенням їх вагових характеристик та впливу на остаточну класифікацію (принцип спільного аналізу інформаційних ознак) так і окремі найбільш інформативні, з урахуванням наявних умов спостереження за об'єктом розпізнавання, образи (принцип роздільного аналізу інформаційних ознак).

Однак, як при спільному, так і при роздільному аналізі залишається проблема надлишковості даних, що не тільки збільшує кількість порівнянь для класифікації та часову складність аналізу даних, але й збільшує ймовірність отримання помилкової класифікації. Все це дозволяє прийти до логічного висновку, що у багатопараметричних комбінованих системах розпізнавання образів рішення щодо класифікації повинні прийматися на меншій кількості але найбільш інформативних образів. Тільки в цьому випадку забезпечується висока достовірність та низька часова складність отримання результату розпізнавання. Таким чином, при гібридному розпізнаванні необхідне рішення задачі консолідації та селекції інформативних образів.

Відомі різноманітні підходи до аналізу інформативності образів. Серед них можна виділити п'ять основних груп:

- на основі дискретних методів пошуку у навчальній вибірці інформативної зони [191];
- на основі методів кластеризації [133];
- на основі припущення о нормальності розподілу об'єктів у кластерах [202];
- на основі теоретико-інформаційного поняття ентропії [196];
- на основі непараметричних оцінок щільностей [4].

Однак, наведені вище методи, направлені на визначення інформативності ознак одного образу але, при гібридному розпізнаванні, стає необхідним визначення критеріїв селекції окремих образів, а не їхніх ознак.

Такий підхід дозволяє виключати одразу групи найменш інформативних ознак, що відповідають усуненому від класифікації образу.

З метою зменшення кількості даних, що підлягають обробці, розроблений метод консолідації образів СС та їх селекції на основі характеристик інформативності [203-205]. 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, відповідно до ситуаційно-подійної моделі гібридного розпізнавання (2.11).

Припустимо, що у деякому просторі  $\Omega$  існує сукупність образів  $P_1-P_k$ , що описують стан СС та є категорією універсуму  $U$ :  $\{P\} \subseteq U$ ,  $U \neq \emptyset$ . В якості відношення еквівалентності пропонується використання відстані образів стану СС 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.

$$P^R \Leftrightarrow \max Ich. \quad (3.29)$$

Надалі здійснюється визначення узагальненої елементарної категорії  $U/P^R = \{\{G_{inf}\}, \{G_{ninf}\}\}$ , яка дозволяє здійснити строгу консолідацію образів СС за ознаками інформативності на дві групи:  $G_{inf}$  – інформативні образи та  $G_{ninf}$  – неінформативні [203, 205]. Консолідація образів, проілюстрована на рисунку 3.6, згідно якої елементарна категорія розділяє образи критеріальною характеристикою  $G=f(ICh)$  [204].

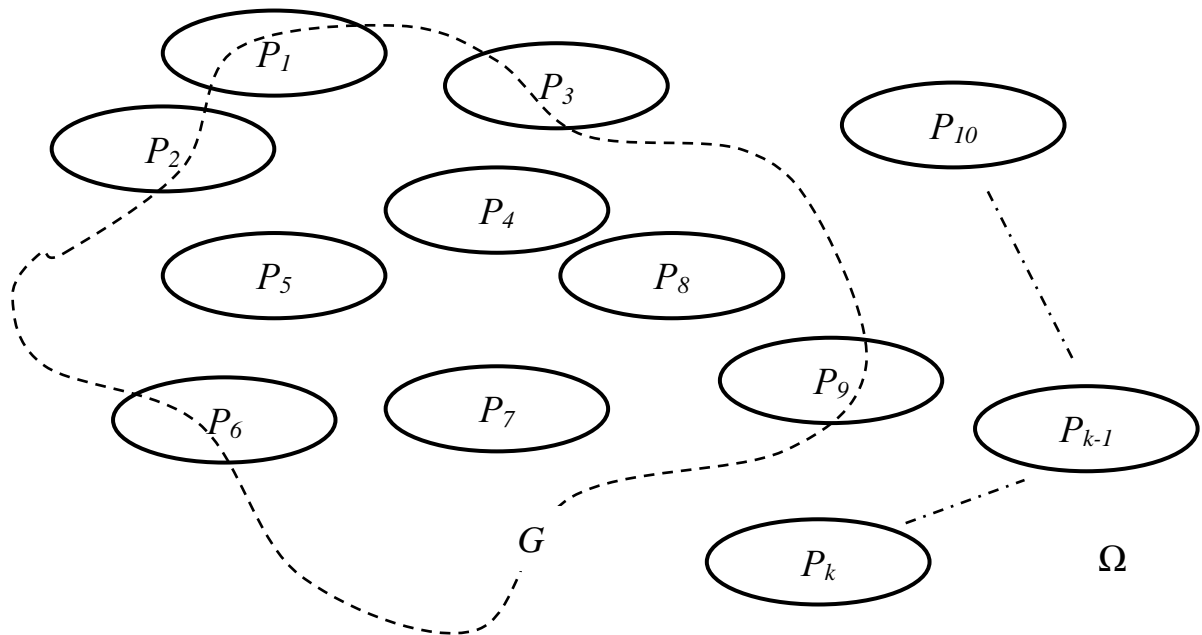


Рисунок 3.6 – Консолідація образів за ознаками інформативності

Елементи універсальної категорії, поділені характеристикою  $G$  за строгим розподілом:  $G_{інф} = \{P_5, P_4, P_7, P_8\}$  – образи що знаходяться у визначеній області  $G$  і відповідають визначеному критерію інформативності;  $G_{нінф} = \{P_1, P_2, P_3, P_6, P_9, P_{10-P_k}\}$  – образи, що не відповідають визначеному критерію  $G$ .

Для здійснення класифікації використовується перша сукупність, яка повністю відповідає критерію. Друга сукупність є неінформативною (або малоінформативною) і може бути усунена від аналізу.

У випадку обмеженого опису класів СС, може бути задіяна нестрога консолідація, згідно якої визначається елементарна категорія  $U/P^R = \{\{G_{інф}\}, \{G_{уінф}\}, \{G_{нінф}\}\}$ :  $G_{інф} = \{P_5, P_4, P_7, P_8\}$  – сукупність образів, що повністю відповідають обраному критерію  $G$ ;  $G_{уінф} = \{P_1, P_2, P_3, P_6, P_9\}$  – образи, що лежать на межі критеріальної характеристики  $G$ ;  $G_{нінф} = \{P_{10-P_k}\}$  – образи, що розташовані за межами обраної критеріальної характеристики  $G$ .

Серед визначених трьох груп, перша є найбільш інформативною і використовується для класифікації. Друга група, хоч і є менш інформативною, але може бути використана для додаткового уточнення

результатів класифікації. Це дозволяє ввести додаткову логічну семантику в процес розпізнавання. Третя група є найменш інформативною і усувається від розпізнавання.

Виходячи з того, що при гібридному розпізнаванні  $k$  образів характеризують стан СС та кожен з цих образів дозволяє здійснити класифікацію за одним і тим самим класом, ефективність використання того чи іншого образу для класифікації визначається через введення спеціальної метрики, яка є відображенням шуканого критерію консолідації. В якості такої метрики пропонується використання принципу мінімуму відстані. Тоді, визначення доцільності включення кожного образу до однієї з консолідованих груп буде визначатися за значенням метрики.

Припустимо, що мається сукупність  $k$  образів  $P_1-P_k$ , при чому  $\{P\} \in \Omega$  (де  $\Omega$  – деякий евклідовий простір). Крім цього, заданий репрезентативний образ  $P^R$ , згідно (3.29). Така ситуація проілюстрована на рисунку 3.7.

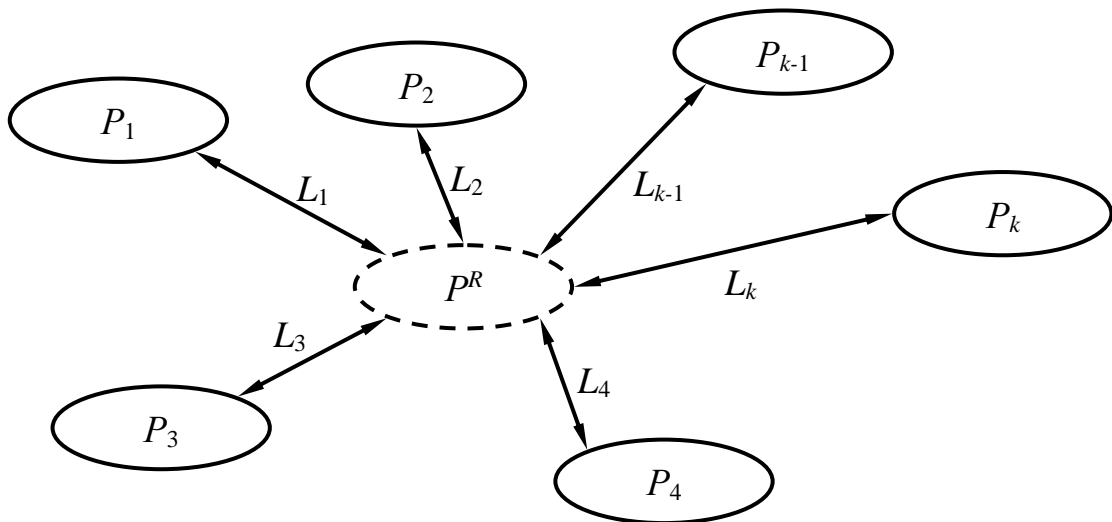


Рисунок 3.7 – Визначення відстані для  $k$  образів об'єкта розпізнавання

Для кожного з  $k$  образів визначається значення відстані  $L_k$ , відносно репрезентативного образу  $P^R$ , як проілюстровано на рисунку 3.7. Характеристика відстані для кожного образу визначається наступним чином:

$$L_k = \|P_k - P^R\| = \sqrt{(P_k - P^R)'(P_k - P^R)}, \quad (3.30)$$

де  $L_k$  – характеристика відстані для  $k$ -го образу об'єкту розпізнавання.

Залежність (3.30), для зручності, можна представити в наступному вигляді:

$$\begin{aligned} D_k^2 &= (P_k - P^R)'(P_k - P^R) = \\ &= P_k' P_k - 2P_k' P^R + P_k^R' P^R = P_k' P_k - 2\left(P_k' P^R - \frac{1}{2}P^R' P^R\right). \end{aligned} \quad (3.31)$$

Відповідно до отриманої залежності (3.31), селекція інформативної групи образів для класифікації здійснюється за мінімальним значенням відстані  $L_k$ . При чому, слід зазначити, що мінімальне значення метрики буде визначатися максимальним значенням від'ємника залежності (3.31):

$$D_k \sim P_k' P^R - \frac{1}{2}P^R' P^R. \quad (3.32)$$

Ті образи, що не відповідають заданій нормі відстані, відносяться до консолідованої групи  $G_{\text{інф}}$  та усуваються від подальшого співставлення і класифікації. Тоді, визначення критерію  $G$  зводиться до визначення  $L^*$  – максимально припустимої відстані образів від репрезентативного образу  $P^R$ . В цьому випадку, віднесення образу  $P_k$  до консолідованої групи інформативних образів  $G_{\text{інф}}$  буде здійснюватися за умови:

$$L_k \leq L^*. \quad (3.33)$$



На основі вищевикладеного, строга консолідація образів між двома групами  $G_{inf}$  і  $G_{ninf}$  буде визначатися наступним чином:

$$\forall P_k \in G_{inf} \Leftrightarrow L_k \leq L^*, \quad (3.34)$$

$$\forall P_k \in G_{ninf} \Leftrightarrow L_k > L^*. \quad (3.35)$$

У випадку використання нестрогої консолідації між групами  $G_{inf}$ ,  $G_{yinf}$ ,  $G_{ninf}$  необхідно додатково визначити норму допуску  $\Delta L^*$  для умовно інформативної групи образів  $G_{yinf}$ . Тоді критерій нестрогої консолідації образів буде мати вигляд:

$$\forall P_k \in G_{inf} \Leftrightarrow L_k < (L^* - \Delta L^*), \quad (3.36)$$

$$\forall P_k \in G_{yinf} \Leftrightarrow (L^* - \Delta L^*) \leq L_k \leq (L^* + \Delta L^*), \quad (3.37)$$

$$\forall P_k \in G_{ninf} \Leftrightarrow L_k > (L^* + \Delta L^*). \quad (3.38)$$

Таким чином, використання запропонованого методу консолідації образів та їх селекції дозволяє зменшити кількість даних, що надаються для співставлення під час класифікації. Таким чином, забезпечується у будь який момент часу та за умови поточного прояву перешкод і викривлень, використання для співставлення меншої сукупності але найбільш інформативних даних. Це підвищує часову складність процесу обробки даних. Використання цього методу дозволяє усунути менш значущі інформаційні ознаки, що негативно не впливає на точність класифікації.

Недоліком запропонованої методу є необхідність визначення репрезентативного образу, що на практиці часто пов'язано з істотними складнощами. Однак, цей недолік може бути виключений завдяки використанню запропонованого вище модифікованого методу пошуку груп ідентичних класифікацій. Визначення та ранжування груп образів об'єкту розпізнавання за ознаками інформативності може бути розглянуте з боку рішення задачі консолідації образів на основі кластеризації (фактично це здійснення таксономії). На основі модифікованого методу та у відповідності до інформаційної моделі (3.28), усі образи, що характеризують стан  $SS$ , розподіляються на групи ідентичних класифікацій. При цьому, значення характеристики  $N(C_m)$  визначає ступінь інформативності. Таким чином, строга і нестрога консолідація може здійснена для груп образів, які визначені на основі кластеризації модифікованим алгоритмом.

В цьому випадку, в якості міри відстані пропонується використання характеристики  $N(C_m)$ , тоді для строгої і нестрокої консолідації в якості граничних характеристик відстані повинні завдаватися граничні значення характеристик  $N^*(C_m)$  та допуску  $\Delta N^*(C_m)$ . Таким чином, використання критеріїв строгої і нестрокої консолідації можливе за умови

$$\begin{cases} L^* \sim N^*(C_m), \\ \Delta L^* \sim \Delta N^*(C_m), \end{cases} \quad (3.39)$$

де  $N^*(C_m)$  – апріорно задане граничне значення чисельності групи образів, що класифікують об'єкт розпізнавання до класу  $C_m$ ;

$\Delta N^*(C_m)$  – норма допуску заданого граничного значення чисельності групи образів, що відносять об'єкт розпізнавання до класу  $C_m$ .

З урахуванням (3.39) критерії строгої консолідації образів (3.34) і (3.35) приймуть вигляд

$$\forall GC_m \in G_{\text{інф}} \Leftrightarrow N(C_m) \leq N^*(C_m), \quad (3.40)$$

$$\forall GC_m \in G_{\text{нінф}} \Leftrightarrow N(C_m) > N^*(C_m). \quad (3.41)$$

Тоді критерії нестрокої консолідації образів (3.36)-(3.38) можна представити у вигляді

$$\forall GC_m \in G_{\text{інф}} \Leftrightarrow N(C_m) < (N^*(C_m) - \Delta N^*(C_m)), \quad (3.42)$$

$$\begin{aligned} & \forall GC_m \in G_{\text{уінф}} \Leftrightarrow \\ & \Leftrightarrow (N^*(C_m) - \Delta N^*(C_m)) \leq N(C_m) \leq (N^*(C_m) + \Delta N^*(C_m)), \end{aligned} \quad (3.43)$$

$$\forall GC_m \in G_{\text{нінф}} \Leftrightarrow N(C_m) > (N^*(C_m) + \Delta N^*(C_m)). \quad (3.44)$$

Таким чином, у випадку роздільного аналізу образів існує можливість аналізу даних і класифікації стану СС як на основі порівняння відстаней до репрезентативних образів, так і на основі кластеризації. При цьому, якщо у конкретному прикладному застосуванні можливе використання обидвох підходів, то у випадку великої кількості наявних образів СС доцільніше використання кластеризації, а за невеликої кількості – відстань до репрезентативних образів.

Здійснення консолідації образів та селекції за ознаками інформативності тільки частково вирішують задачу пріоритетної обробки найбільш значущих даних. Для побудови черги обробки даних слід також

визначитися з послідовністю обробки окремих ознак кожного образу, що потрапив до інформативної групи. Надалі представлений метод, що дозволяє вирішити таку задачу.

### **3.4 Метод екстенціонально-інтенціонального аналізу ознак при гібридному розпізнаванні образів**

У пункті 1.2.3 вказувалося на концепцію, запропоновану Д. О. Поспеловим Г.С., сутністю якої є розподіл методів за способом представлення знань при розпізнаванні. За таким принципом методи розпізнавання поділяються на інтенціональні (пов'язані з оперуванням детальними описами структури даних об'єктів інформаційного процесу) та екстенціональні (пов'язані з оперуванням загальними описами об'єктів інформаційного процесу) [145-150, 206, 207].

З точки зору гібридного розпізнавання, інтенціональний підхід припускає представлення знань у вигляді схеми зв'язків між ознаками об'єкту розпізнавання, а екстенціональний – представлення за допомогою опису окремих образів цього об'єкту.

Аналіз цих підходів показав [206, 208], що жоден з цих підходів, узятий окремо від іншого, не дозволяє сформувати адекватне віддзеркалення предметної області та отримати достовірний результат обробки даних в широкому діапазоні зміни характеристик навколишнього середовища. Аналіз обмежених категорій при екстенціональному підході, як правило, не дозволяє отримати бажану високу достовірність розпізнавання, а інтенціональний підхід передбачає більшу кількість співставлень даних та, відповідно, має більшу часову складність отримання результату класифікації.

З погляду на те, що екстенціональний та інтенціональний підходи знаходяться у відношенні доповнення один до одного, для гібридного розпізнавання образів пропонується використання екстенціонально-

інтенсіонального методу [206-208]. В цьому випадку, метод розпізнавання може повністю зводитися до екстенсіонального або інтенсіонального аналізу. Розроблений метод екстенсіонально-інтенсіонального аналізу ознак при гібридному розпізнаванні проілюстрований на рисунку 3.8.

Згідно рисунку 3.8 та у відповідності до (2.1), стан СС представляється 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  $\{C\}$ , що впливає з концепції гібридного розпізнавання образів.

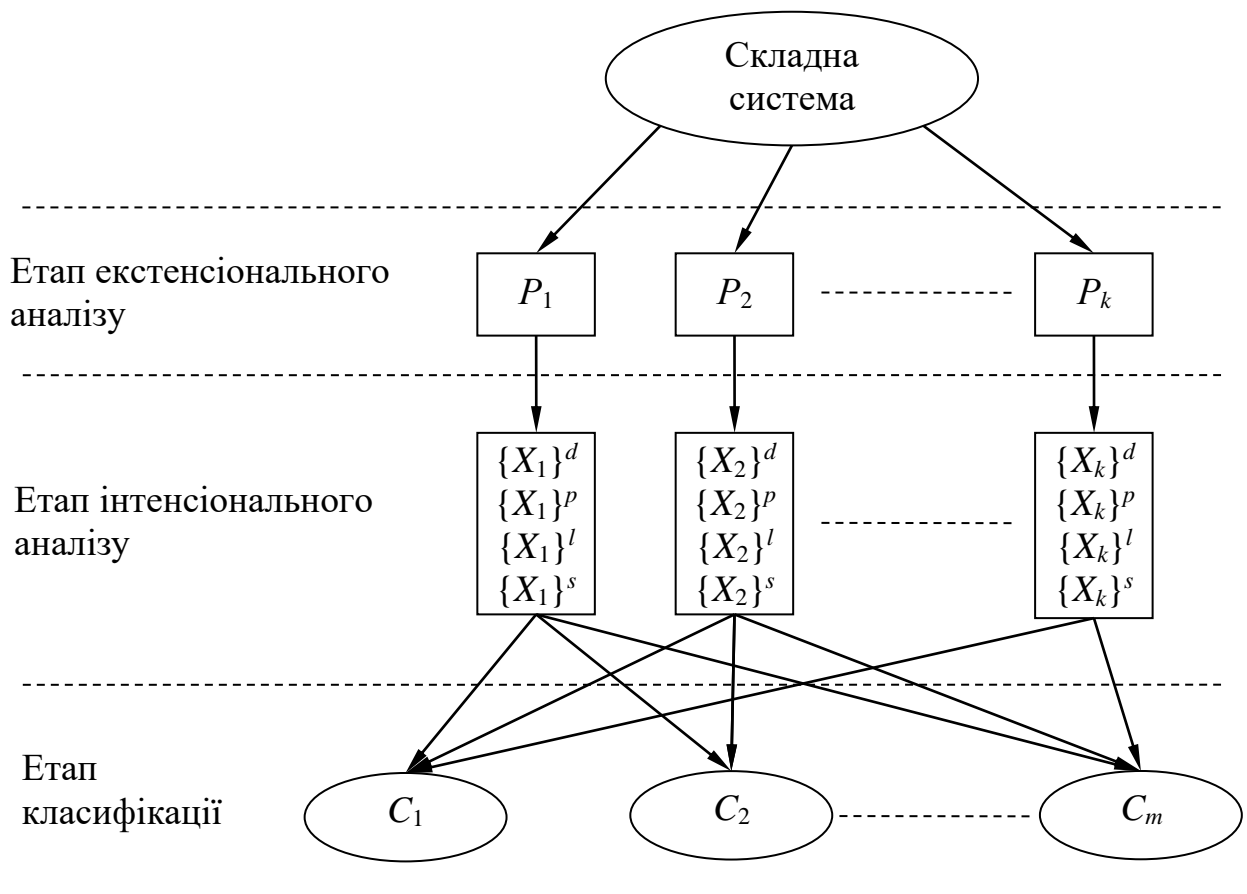


Рисунок 3.8 – Метод екстенсіонально-інтенсіонального аналізу ознак при гібридному розпізнаванні образів

Для забезпечення мінімізації часу класифікації, в екстенціонально-інтенціональному методі, спочатку, здійснюється етап екстенціонального аналізу, при якому виконується спроба класифікації об'єкта розпізнавання за узагальненими описами образів. У зв'язку з цим, особливо актуальним в системах розпізнавання стає завдання реалізації механізму узагальнення описів об'єктів, що відносяться до одного класу, тобто механізму формування компактних узагальнених образів. Ця задача вирішується через використання розробленого методу консолідації та селекції образів, в якому здійснюється пошук та аналіз груп образів з ідентичною класифікацією.

Оперування з узагальненими образами вимагає вирішення наступної послідовності операцій:

- визначення інформаційного вкладу ознак до інформаційного опису узагальненого образу;
- кластерно-конструктивний аналіз узагальнених образів;
- визначення семантичного навантаження ознак;
- семантичний кластерно-конструктивний аналіз ознак;
- змістовне порівняння узагальнених образів класів один з одним і ознак однієї з іншою, наприклад, за допомогою когнітивних діаграм, в тому числі діаграм Мерліна) [162, 203–205].

При цьому слід враховувати, що достовірність представлення стану  $CC$  за кожним з  $k$  образів буде змінюватися у випадку зміни зовнішніх умов. Таким чином, у кожному окремому випадку, інформативність образів з множини  $\{P\}$  буде різною. Для отримання достовірного результату розпізнавання за мінімальний час, необхідне здійснення селекції інформативних образів, за методикою, запропонованою вище, та опублікованою у [206, 207]. Надалі, класифікація може бути визначена на основі статистики, що отримується шляхом голосування: належність стану  $CC$  до класу  $C_i$  визнається достовірною, якщо ця класифікація підтверджується більшістю образів з множини  $\{P\}$ .

Якщо екстенсіональний аналіз не дозволяє отримати бажаний рівень достовірності класифікації, на наступному етапі здійснюється інтенсіональний аналіз. Крім цього, перехід від екстенсіонального знання до інтенсіонального доцільно здійснювати у випадку попередньої селекції та наявності обраної малої кількості інформативних ознак об'єкту розпізнавання.

Слід зазначити, що усунення від аналізу неінформативних обрізів СС фактично призводить до усунення відповідної сукупності інформаційних ознак. Тому, операція консолідації та селекції образів призводить до зменшення кількості співставлень даних як на екстенсіональному, так і на інтенсіональному етапах. При цьому, одразу усувається певна група менш інформативних ознак і перевірка інформативності окремих ознак в цих групах непотрібна.

Сама класифікація здійснюється методом гібридного розпізнавання на основі пошуку груп образів з ідентичною класифікацією. Для збільшення кількості інформативних показників за якими приймається рішення щодо класифікації, на етапі інтенсіонального аналізу використовується розподіл, який був отриманий на етапі екстенсіонального аналізу. В цьому випадку, інтенсіональний аналіз нібито «уточнює» результат класифікації за узагальненими описами об'єктів розпізнавання на екстенсіональному етапі.

Слід зазначити, що використання запропонованого методу не вимагає проведення повного аналізу інформативних ознак за кожним з  $k$  образів об'єкту розпізнавання. Такий аналіз може здійснюватися лише до моменту, коли за одним або декількома образами з множини  $\{P\}$  буде отримана класифікація, рівень достовірності якої досягає апріорно заданого значення. Також, у випадку використання СРЧ, співставлення може тривати до моменту вичерпання часу, відведеного на обробку та класифікацію.

Розроблений екстенціонально-інтенціональний метод не виключає можливості використання тільки екстенціонального етапу. Іntenціональний етап може бути пропущений, якщо на екстенціональному етапі отримана класифікація, що задовольняє апріорно визначеному рівню достовірності.

Використання екстенціонально-інтенціонального методу створює умови щодо повномірного використання при гібридному розпізнаванні різноманітних варіантів опису ознак СС. Це сприяє створенню гнучких, адаптивних інформаційних технологій обробки неоднорідних даних в СС на основі гібридного розпізнавання. Питання створення такої інформаційної технології розглядається в наступному розділі.

### **Висновки до розділу 3**

В розділі 3 вирішено шосте завдання дисертаційного дослідження, а саме: розробка методів консолідації інформативних ознак та розпізнавання образів для складаних систем з неоднорідними даними, різнорідними проявами перешкод і викривлень. Розробка методів базується на модифікованій матричній інформаційній моделі СС, нового принципу консолідації неоднорідних даних ситуаційно-подійної моделі гібридного розпізнавання. Основні висновки полягають в наступному.

При спільному аналізі інформаційних ознак здійснюється поєднання сукупності образів СС до одного глобального образу. В цьому випадку, задача розпізнавання вирішується на основі традиційних підходів і методів комбінованого розпізнавання. Однак, опис СС, в цьому випадку, отримує більше різноманіття ознак різної природи виникнення, на які неоднорідні перешкоди і викривлення мають різний вплив. Таким чином, забезпечується прийняття класифікаційного рішення на основі більшої кількості незалежних або слабо залежних ознак з малою коваріацією.



Роздільний аналіз інформаційних ознак дозволяє отримати усі переваги гібридного розпізнавання стосовно різноманіття побудови альтернати обробки неоднорідних даних на основі ситуаційного підходу в залежності від поточних зовнішніх умов. Однак, для зменшення кількості співставлень даних та зменшення часової складності процесу класифікації, необхідне використання методу консолідації та селекції образів за характеристикою їхньої інформативності при поточних зовнішніх умовах та рівня перешкод і викривлень.

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

Модифікація розробленого методу гібридного розпізнавання образів дозволяє отримувати здійснювати консолідацію образів з ідентичною класифікацією, що фактично є процесом кластеризації та може використовуватися під час синтезу інформаційних технологій обробки неоднорідних даних в СС для визначення характеристик інформативності образів за різних зовнішніх умов.

Модифікований метод гібридного розпізнавання образів дозволив створити інформаційну модель підтримки прийняття рішення у випадку багатоальтернативного результату класифікації. Розроблена інформаційна модель 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.

Розроблено метод екстенсiонально-інтенсiонального 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.

Розроблені методи гібридного розпізнавання образів, консолідації та селекції образів, а також екстенсiонально-інтенсiонального аналізу ознак адаптовані до використання у складі однієї інформаційної технології обробки неоднорідних даних в СС з різноманітними проявами перешкод і викривлень.

## РОЗДІЛ 4

### ІНФОРМАЦІЙНА ТЕХНОЛОГІЯ ГІБРИДНОГО РОЗПІЗНАВАННЯ ОБРАЗІВ ДЛЯ КОНСОЛІДОВАНОЇ ОБРОБКИ ДАНИХ В СКЛАДНИХ СИСТЕМАХ З РІЗНОРІДНИМИ ПРОЯВАМИ ПЕРЕШКОД І ВИКРИВЛЕНЬ

#### 4.1 Граничні умови використання інформаційної технології

##### 4.1.1 Часові умови

Виходячи з формулювання мети роботи, інформаційна технологія (ІТ) гібридного розпізнавання повинна забезпечувати заданий рівень достовірності класифікації, з одночасним зменшенням кількості даних, що надаються для співставлення в складних системах з різномірними проявами перешкод і викривлень. Тобто процес співставлення повинен здійснюватися не до моменту завершення повної обробки усієї сукупності даних, а до моменту отримання апріорно заданого рівня достовірності результату.

З іншого боку, на ІТ накладається додаткова умова часового обмеження процесу обробки та прийняття рішення. Наявність певних обмежень на час обробки даних та класифікації, вказує на доцільність реалізації інформаційної технології гібридного розпізнавання образів на основі системи реального часу [212-214].

В СРЧ маються жорсткі обмеження на часові (динамічні) характеристики виконання інформаційних процесів. Формально, це обмеження може бути записано у вигляді [214]:

$$t_{\min} < t < t_{\max}, \quad (4.1)$$

де  $t$  – деяка часова характеристика, що характеризує тривалість інформаційного процесу, тобто час співставлення ознак СС та отримання результату класифікації;

$t_{\min}$  і  $t_{\max}$  – відповідно, гранично припустимі значення тривалості інформаційного процесу.

В СРЧ вихід часу виконання інформаційного процесу за межі гранично припустимої тривалості вважається відмовою в роботі.

Існують системи «жорсткого» та «м'якого реального часу» [213, 214]. В системах «жорсткого» реального часу виконання умови (4.1) обов'язково, коли для систем «м'якого» реального часу, умова тривалості інформаційного процесу припускається у формі [214]

$$t_{\min} - \Delta t < t < t_{\max} + \Delta t, \quad (4.2)$$

де  $\Delta t$  – припустиме значення відхилення тривалості інформаційного процесу.

Тобто в системах «м'якого» реального часу припускається деяке перевищення граничних значень тривалості інформаційного процесу.

Таким чином, для забезпечення оперативної обробки даних, ІТ гібридного розпізнавання образів повинна бути адаптована для використання на базі як систем «жорсткого», так і «м'якого» реального часу.

Для реалізації ІТ в системі «жорсткого» реального часу, процес обробки не повинен перевищувати строго визначеного часу, тобто

$$t_{\text{обр}} \leq t_{\text{макс}}, \quad (4.3)$$

де  $t_{\text{макс}}$  – 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 цей рівень буде меншим за цільове значення.

В системах «м'якого» реального часу дозволяється затримка часу обробки на певне визначене значення, тобто

$$t_{обр} \leq t_{макс} + \Delta t, \quad (4.4)$$

де  $\Delta t$  - визначений проміжок часу, на який процес обробки може перевищувати заданий максимально припустимий час.

Тоді, в цьому випадку, процес обробки продовжується до моменту отримання заданого рівня достовірності, тобто  $t_{обр} < t_{макс}$ , або, якщо заданий 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.

#### *4.1.2 Умови багатозадачності*

Для задач критичного застосування, підвищення оперативності обробки даних може бути забезпечена використанням багатозадачності. При чому, як вже згадувалося раніше, роздільний аналіз і запропоновані моделі і методи гібридного розпізнавання, орієнтовані на реалізацію в режимі багатозадачності.

В СРЧ використовуються два основних режими багатозадачності [213, 214]:

- кооперативна багатозадачність;
- витісняюча багатозадачність.

Кооперативна багатозадачність – режим роботи операційної системи, при якому різні завдання самі здійснюють передачу управління одна одній в ті моменти часу, коли це необхідно [214]. Точніше, поточна задача віддає управління операційній системі, а вже та витягує іншу задачу з черги і віддає їй управління. Передача управління іншій задачі може також відбуватися з ініціативи користувача або зовнішнього пристрою, який порушив апаратне переривання. З одного боку, при правильному проектуванні програмного забезпечення з урахуванням всіх факторів, що впливають, такий підхід дозволяє досягати високої реактивності СРЧ на зовнішні події, передбачуваності її поведінки та ефективного використання ресурсів ЕОМ. З іншого боку, це ускладнює процес розробки, накладає підвищені вимоги на кваліфікацію розробника, служить потенційним джерелом помилок, зменшує стійкість системи. Зокрема, відмова обслуговування одного завдання призводить до неможливості передачі управління наступній задачі і, як наслідок, до відмови всієї системи.

Витісняюча багатозадачність – режим роботи, при якому перемикання задач здійснює сама операційна система. Кожній задачі виділяється короткий часовий інтервал (типове значення 0,02 секунд [214]), що називається квантом часу (timeslice). Після закінчення цього проміжку часу, завдання примусово призупиняється (витісняється) і управління передається завданню, що знаходиться далі в черзі готових задач. Після вичерпання черги, управління знову передається першому завданню в черзі і так далі. Такий режим роботи, в загальному випадку, не призводить до високої реактивності системи, але забезпечує гарантований відгук на зовнішні події в межах часової затримки, яку можна розрахувати заздалегідь. Програмне забезпечення, яке працює в умовах витісняючої багатозадачності, маю істотну стійкість до помилок, що пов'язані з відмовою обробки одного із

завдань. Така ситуація призведе лише до втрати квантів часу, виділених на виконання збійного завдання, але працездатність системи в цілому збережеться.

При гібридному розпізнаванні здійснюється поетапне співставлення інформаційних ознак СС і на кожному етапі здійснюється визначення отриманої достовірності. Перехід до наступного етапу порівняння здійснюється за умови, що не було отримано бажаного рівня достовірності, а час, визначений для класифікації ще не вичерпано. При цьому, формується ранжована черга обробки найбільш інформативних ознак СС, за якою здійснюється поетапне прийняття рішення. Тому, переваги від використання витісняючої багатозадачності повністю нівелюються тим, що до моменту завершення обробки попередньої ознаки і визначення отриманої достовірності результату класифікації відсутня необхідність переключення на обробку іншої ознаки. Визначення же при цьому найбільш раціональних інтервалів кванту часу, взагалі неможливий, так як час обробки неоднорідних ознак відрізняється. Визначення же кванту часу, на рівні найбільш тривалого процесу співставлення ознаки призведе до перевитрат часу на обробку інших ознак, для який обраний квант часу є надмірним. Тому, для організації ІТ гібридного розпізнавання образів доцільне використання корпоративної багатозадачності, коли процес обробки ознаки з черги і визначення поточного рівня достовірності класифікації буде передавати управління процесу обробки наступної ознаки з черги, якщо апріорно заданий рівень достовірності не досягнутий, а час, виділений на обробку та класифікацію за умовами «жорсткого» чи «м'якого» реального часу не вичерпаний. При цьому, корпоративна багатозадачність забезпечить передбачуваність процесу обробки в часовій області, та високу реактивність на зміну вхідних даних, що найбільш актуально для ситуаційного підходу до управління станом СС.

Таким чином, до інформаційної технології гібридного розпізнавання на основі СРЧ представляються наступні основні 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 даних 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, як вже вказувалося раніше, 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 в СРЧ кооперативну багатозадачність, яка дозволяє найбільш ефективно використовувати ресурси апаратно-програмного забезпечення.

## **4.2 Вимоги до апаратно-програмних засобів реалізації інформаційної технології**

### *4.2.1 Загальні вимоги*

Специфіка ІТ гібридного розпізнавання образів вимагає наступних основних вимог до апаратно-програмних засобів:

- надійність роботи пристроїв реєстрації ознак СС, стабільність їхніх характеристик точності, так як на їх основі визначаються стаціонарні характеристики інформативності  $ICh$ ;

- робота в широкому діапазоні зовнішніх умов, так як в ІТ використовується ситуаційний підхід, який забезпечує адаптивний процес



класифікації за умови широкого діапазону змін наявних перешкод і викривлень;

- наявність розвиненої системної служби часу, для забезпечення передбачуваності поведінки в часовій області та надійного контролю за тривалістю інформаційних процесів співставлення даних та класифікації;

- легка конфігурованості, для забезпечення можливості настройки ІТ під різноманітні (в тому числі такі, що змінюються з часом) умови навколишнього для СС середовища, а також можливості корекції моделі опису СС в просторі ознак у випадку змін завдань на класифікацію.

- простота системи команд і способів адресації, що забезпечує більшу швидкість виконання алгоритмів обробки, а також уніфікацію довжини команд та часу їх виконання, на кшталт архітектурі RISC [215].

- розвинена системи інтерфейсів та можливість підключення великої кількості зовнішніх пристроїв реєстрації характеристик, так як при гібридному розпізнаванні здійснюється оперування зі значною кількістю пристроїв реєстрації характеристик СС різної природи виникнення.

- наявність засобів попередньої обробки і нормування характеристик СС, що забезпечує приведення отриманих неоднорідних характеристик стану СС до форматів, що можуть оброблятися підчас гібридного розпізнавання, а також забезпечення нормування характеристик по значенню і рівню.

- розвинена системи переривань, що забезпечить максимальну реактивність ІТ на ситуацію зміни вхідних характеристик СС та зовнішнього середовища.

#### *4.2.2 Вимоги до апаратного забезпечення інформаційної технології*

З урахуванням вказаних вище вимог, була розроблена структурна організація системи обробки даних СС, яка наведена на рисунку 4.1. В цій структурі використовуються розроблені моделі та методи гібридного розпізнавання.

Підсистема виявлення і відображення характеристик СС є однією з основних в системі розпізнавання та представляє собою сукупність сенсорів, різної фізичної та інформаційної природи, які дозволяють отримати відображення стану СС у просторі образів. Стосовно фізичних явищ і параметрів, такі сенсори можна розділити на: сенсори візуальних ознак об'єктів; сенсори звукових ознак об'єктів; сенсори механічних ознак об'єктів; сенсори характеристичних ознак об'єктів; сенсори спеціальних ознак об'єктів.

Сенсори візуальних ознак дозволяють сформувати візуальне уявлення про об'єкт, його геометричну форму, розміри, колірні і текстурні характеристики, а також характеристики яскравості.

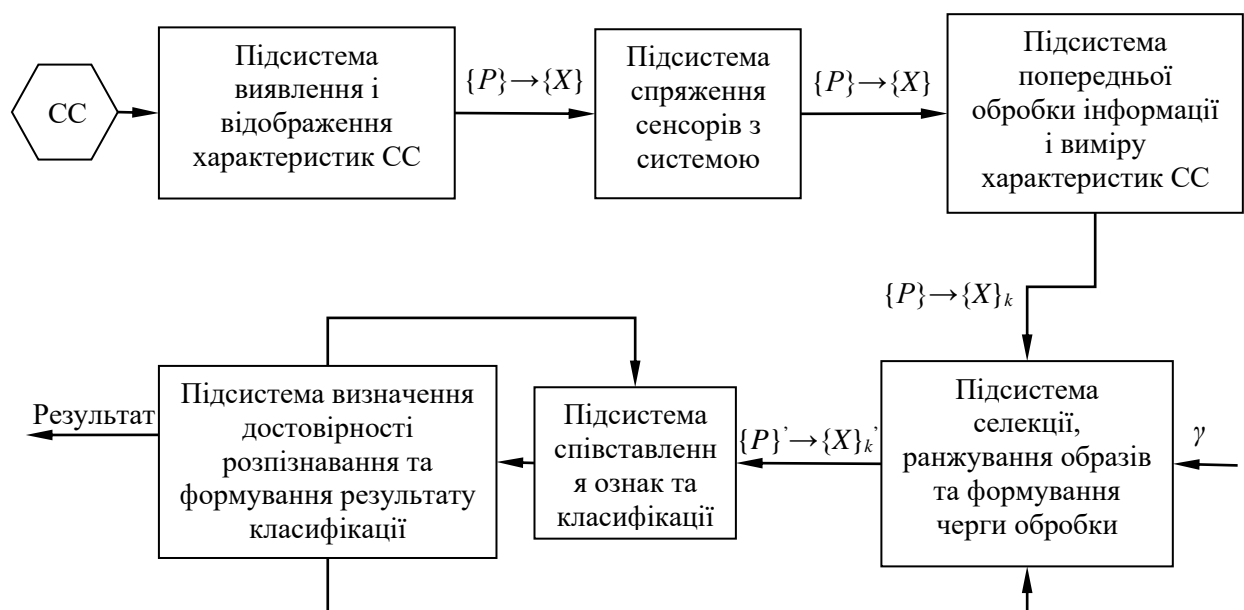


Рисунок 4.1 – Структурна організація системи обробки даних СС на основі гібридного розпізнавання образів

Сенсори звукових ознак дозволяють сформувати данні про інтенсивність звукового сигналу, його частотне наповнення, що в подальшому може бути використане для витягування додаткових

інформаційних ознак, наприклад, у випадку лінгвістичного аналізу і розпізнавання речі.

Сенсори механічних ознак дозволяють зафіксувати будь які механічні обурення, які відбиваються у просторі під впливом СС, або діють на СС (наприклад, сенсори вібрації, контактний і безконтактний контроль руйнації, дефектоскопи, і т. п.).

Сенсори характеристичних ознак дозволяють визначити додаткові фізичні параметри СС, такі наприклад, як температура, вологість, інтенсивність освітлення, електромагнітні характеристики та будь які інші рівні фізичних величин або просторові характеристики (координати положення у просторі, швидкість руху, прискорення і т. п.).

Сенсори спеціальних ознак представляють собою широкий спектр апаратних пристроїв і алгоритмів фіксації будь-яких додаткових фізичних, природних явищ, соціально-економічних характеристик і параметрів (для систем розпізнавання соціально-економічних і геополітичних факторів і явищ тощо).

Кінцевою задачею цієї підсистеми є формування сукупності образів різної природи виникнення, які характеризуються визначеною сукупністю ознак ( $\{P\} \rightarrow \{X\}$ ).

З урахуванням того, що для реалізації гібридного розпізнавання доцільне використання апаратно-програмного комплексу (що дозволяє гнучко оперувати різноманітними алгоритмами обробки даних, а також створювати ефективну композицію алгоритмів) задачею підсистеми спряження сенсорів з системою є реалізація і підтримка необхідних апаратних інтерфейсів і програмних протоколів пересилки даних. Такі апаратно-програмні засоби повинні чітко враховувати специфіку даних, що передаються, і мінімізувати внесені викривлення і перешкоди, з метою більш точного відображення об'єктів у просторі образів. Від ефективності реалізації апаратно-програмних інтерфейсів також істотно залежить

ефективність усієї системи розпізнавання образів. Підсистема спряження сенсорів з системою забезпечує передачу усієї сукупності інформаційних ознак відповідних образів  $\{P\} \rightarrow \{X\}$  до підсистеми попередньої обробки інформації і виміру характеристик СС.

Підсистема попередньої обробки інформації і виміру характеристик СС вирішує коло задач по корекції отриманої інформації  $\{P\} \rightarrow \{X\}$ , визначення викривлень і завад, а також приведення даних до необхідного формату представлення. В результаті, на виході підсистеми формується коректована множина інформаційних ознак образів СС  $\{P\} \rightarrow \{X\}_k$ . При гібридному розпізнаванні, протоколи попередньої обробки можуть базуватися на апаратно-програмному аналізі декількох інформаційних каналів з метою виявлення адекватності отриманої інформації про СС та зовнішні умови. Крім цього, попередня обробка полягає у консолідації даних для подальшого використання методів гібридного розпізнавання.

Підсистема селекції, ранжування та формування черги виконує задачі визначення сукупності даних, що за поточних зовнішніх умов  $\gamma$ , 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 методи гібридного 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.

#### *4.2.3 Вимоги до програмного забезпечення інформаційної технології*

Сучасне системне програмне забезпечення, зокрема СРЧ, підтримують два режими виконання програмного коду [214, 215]:

- привілейований режим, або режим ядра (kernel mode) чи режим супервізора;
- непривілейований режим, або режим додатків.

Основна відмінність між цими режимами полягає в тому, що програми, які працюють в режимі супервізора, мають безпосередній, необмежений доступ до усіх ресурсів комп'ютерної системи: до зовнішніх пристроїв, адресному простору, перериванням та ін. Програми непривілейованого режиму працюються у віртуальному середовищі, що формується програмним середовищем режиму супервізора. У зв'язку з цим, поділяють два класи архітектури СРЧ: монолітна та мікроядерна [214].

Всі компоненти монолітної СРЧ працюють в режимі супервізора в єдиному адресному просторі [214]. Головна перевага таких СРЧ – висока продуктивність, а головний недолік – відсутність можливості внесення будь-яких змін в структуру операційної системи в процесі її експлуатації, тобто погана масштабованість. Інший недолік – невисока реактивність системи, тому що якщо зовнішня подія, що вимагає негайної реакції, відбувається під час виконання завдань рівня ядра, то обробка цієї події затримується до повернення на рівень додатків.

Особливість мікроядерної СРЧ – наявність компактного і швидкодіючого ядра, що працює в режимі супервізора, а всі інші компоненти операційної системи, включаючи менеджери ресурсів, при цьому працюють в непривілейованому режимі [214]. Такий підхід забезпечує хорошу гнучкість і масштабованість операційної системи, малий час реакції на зовнішні події. З іншого боку, мікроядерні операційні системи відрізняються щодо невисокою

продуктивністю, тому що під час роботи відбуваються часті переключення з режиму в режим.

З урахуванням наведеного аналізу, для реалізації ІТ гібридного розпізнавання доцільне використання мікроядерної архітектури. В цьому випадку забезпечується малий час реакції та добра масштабованість при зміні конфігурації периферійних пристроїв реєстрації характеристик СС та зовнішніх умов. Недолік мікроядерної архітектури, пов'язаний з частими переключеннями режимів усувається використанням окремої, спеціалізованої комп'ютерної системи, на якій встановлюється СРЧ з мікроядерною архітектурою і яка здійснює виключно обробку сформованої черги ознак СС. Усі додаткові сервіси зберігання результатів, їхньої візуалізації та представлення експерту, генерації словника класів та ін. покладаються на іншу комп'ютерну систему. Таким чином, комп'ютерна система, що реалізовує гібридне розпізнавання є вбудованою комп'ютерною системою. В цьому випадку зменшується кількість переключень режимів СРЧ в процесі обробки даних, а також спрощується її програмне забезпечення та зменшується навантаження на багатозадачність за рахунок мінімізації спектра обчислювальних функцій. При цьому, також слід зазначити, що часові затримки в процесі переключення є чітко визначеними, так як для кожної СРЧ, зазвичай, вказуються числові значення граничного часу переключення з задачі на задачу та гранична затримка між виникненням переривання і початком його обробки.

### **4.3 Реалізація інформаційної технології гібридного розпізнавання образів**

#### *4.3.1 Формування черги обробки даних при гібридному розпізнаванні*

Для ефективною реалізації інформаційної технології гібридного розпізнавання та отримання достовірного рішення за мінімальних часових

витрат, у будь який момент часу повинна здійснюватися обробка найбільш інформативних даних. Слід зазначити, що розроблений метод роздільного аналізу ознак має переваги перед спільним завдяки спрямованості на використання багатозадачності і, насамперед, кооперативної. На рисунку 4.2 наведена схема паралельної обробки та формування черги при гібридному розпізнаванні образів [216].

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  $CC$ ).

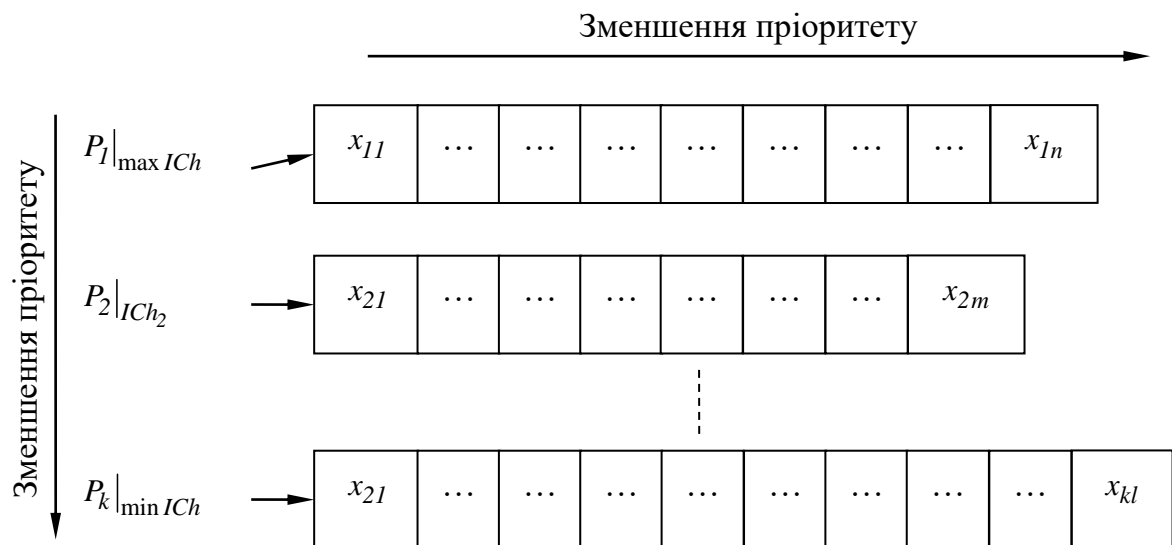


Рисунок 4.2 – Формування черги і послідовності обробки даних при гібридному розпізнаванні

Також, слід зазначити, що відповідно до розробленого екстенсіонально-інтенсіонального 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.

#### *4.3.2 Етапи реалізації інформаційної технології гібридного розпізнавання образів*

Серед етапів реалізації інформаційної технології гібридного розпізнавання образів для консолідованої обробки неоднорідних даних в СС можна відокремити апріорний етап (Етап 0) та апостеріорні етапи (Етап 1 – Етап 9) [217]. Апріорний етап виконується тільки під час створення інформаційної технології і направлений на визначення характеристик, що є незмінними в процесі реалізації інформаційної технології. Апостеріорні етапи здійснюються кожний раз, для кожного циклу обробки неоднорідних даних та гібридного розпізнавання образів. Однак, слід зазначити, що у залежності від поточної ситуації, рівня наявних перешкод і викривлень, деякі етапи можуть пропускатися, якщо бажаний рівень достовірності класифікації досягається на попередніх етапах. Надалі наведений опис кожного етапу розробленої інформаційної технології гібридного розпізнавання образів.

Етап 0. Апріорно визначаються стаціонарні характеристики інформативності  $SICH$  кожного інформаційного потоку в СС, 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 рисунку 4.2, позначена вертикальною стрілкою).

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 рисунку 4.2, позначена горизонтальною стрілкою).

Етап 7. Здійснюється екстенсіональний 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  $s$  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  $CC$  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 досягється бажане значення  $D(C)$ , то подальший аналіз інших ознак припиняється і результат класифікації стану  $CC$  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.

#### 4.3.3 Априорне визначення складових характеристик ситуаційно-подійної моделі гібридного розпізнавання

Інформаційні процеси, що мають місце при обробці неоднорідних даних в СС, є проявами ситуації, яка визначається поточним рівнем та характером наявних перешкод і викривлень характеристик об'єкта розпізнавання. Згідно розробленої ситуаційно-подійної моделі (2.11) прояв ситуації визначається сукупністю характеристик навколишнього середовища  $EC$ , що мають місце за поточних умов отримання даних для класифікації стану СС. Поточна ситуація ініціює виконання дій, які пов'язані з попередньою обробкою сукупності образів  $P$  СС через визначення їхніх нестационарних характеристик інформативності  $NSICH$  на основі поточного стану  $EC$ , а також отриманням априорно заданих стаціонарних характеристик інформативності  $SICH$  для кожного образу  $P$  та подальше визначення загальної характеристик інформативності образів, що відповідають поточній ситуації. На основі цих даних, здійснюється подальша селекція найбільш інформативних даних, формування з них черги та прийняття рішення щодо класифікації за одним з передвизначених класів з алфавіту  $C$ .

Функціонування такої ситуаційно-подійної моделі можливе за умови априорного визначення необхідних її складових:

- сукупності образів  $P$ , що характеризують стан СС;
- стаціонарних характеристик інформативності, що є проявом якості відображення даних СС та точності пристроїв реєстрації цих характеристик;
- алфавіту класів  $C$ , які характеризують усі можливі стани СС, які можуть бути визначені на етапі класифікації.

Логічна схема априорного визначення елементів ситуаційно-подійної моделі представлена на рисунку 4.3.

На основі експертного аналізу, визначаються джерела інформації, які найбільш повною мірою можуть відображати стан СС за будь яких можливих проявів перешкод і викривлень. Такий експертний аналіз повністю залежить

від специфіки СС, особливостей її функціонування, а також умов, в яких це функціонування реалізовується. З метою забезпечення максимальної кількості різнорідних ознак на які різні перешкоди і викривлення мають різний вплив, джерела інформації повинні бути різної природи виникнення. Тоді, ступінь незалежності ознак може бути забезпечений повною мірою. Слід зазначити, що від якості проведення такої експертної оцінки багато в чому залежить якість та ефективність подальшого використання інформаційної технології гібридного розпізнавання образів.



Рисунок 4.3 – Логічна схема формування ситуаційно-подійної моделі

Кінцевою метою експертного аналізу джерел неоднорідних даних є формування сукупності образів, різні комбінації яких можуть забезпечити прийняття достовірного результату класифікації стану СС за різних

зовнішніх умов. При чому, різним проявам умов відповідають різні поєднання образів, що надаються для співставлення та розпізнавання.

На основі характеристик засобів відображення інформації також визначаються стаціонарні характеристики інформативності  $SIC_h$ , які розміщуються в базі знань та не змінюються у впродовж усіх циклів гібридного розпізнавання. У випадку, якщо інформаційна технологія підлягає модернізації, заміні технічних засобів реєстрації ознак, то тільки в цьому випадку, на основі експертної оцінки, оновлені характеристики повинні бути завантажені до бази знань, замінивши характеристики тих засобів, що були виведені з експлуатації.

Також, відповідно до логічної схеми на рисунку 4.3, потребує апріорного визначення алфавіт класів, який описує можливі стани  $CC$ , які можуть бути ідентифіковані під час розпізнавання. Кожен клас повинен характеризуватися репрезентативними образами різної природи виникнення. Таким чином, кожний клас містить підкласи, відповідно до концепції гібридного розпізнавання образів, що представлена в підрозділі 2.3 та проілюстрована на рисунку 2.2. Якщо,  $C_{11} \dots C_{1k}$  підкласи класу  $C_1$ , які містять репрезентативні ознаки різної природи виникнення, то результат класифікації стану  $CC$  на основі співпадиння з одним чи декількома класами  $C_{11} \dots C_{1k}$  дозволяє з різною достовірністю стверджувати про відповідність стану складної системи класу  $C_1$ . Очевидно, що максимальне значення статистичної достовірності розпізнавання  $D=100\%$  буде у випадку, коли кожен підклас з сукупності  $C_{11} \dots C_{1k}$  буде підтверджувати ідентичну класифікацію (відповідно до метода гібридного розпізнавання на основі пошуку груп ідентичних класифікацій, що представлений у підрозділі 3.1). Кількість підкласів у кожному класі, відповідно до аналізу, проведеному у пункті 1.4.6, прямо пропорційна статистичній достовірності розпізнавання.

Для визначення сукупності найбільш інформативних образів, ситуативно-подійна модель вимагає також експертного аналізу можливих перешкод і викривлень. На основі цього аналізу та з урахуванням визначеної сукупності образів, що характеризують стан СС, формується сукупність показників зовнішнього середовища *ЕС*. На основі цих показників, модель дозволяє визначати інформативність кожного образу з сукупності *P* за поточних умов отримання їхніх ознак. Використовуючи поточну інформацію про стан навколишнього середовища та рівень перешкод і викривлень, модель дозволяє визначати раціональну сукупність неоднорідних даних для співставлення та отримання класифікаційного рішення. Результатом цієї операції є визначення сукупності нестационарних характеристик інформативності для кожного образу з сукупності *P*. Фактично, нестационарні характеристики інформативності пов'язують кожен образ СС з відповідним станом навколишнього середовища. Таким чином, ситуаційно-подійна модель дозволяє визначити для відповідного стану середовища та рівня перешкод і викривлень найбільш інформативні для аналізу образи.

Таким чином, ситуаційно-подійна модель орієнтована на отримання класифікаційного рішення на основі обробки виділеної мінімальної кількості образів СС, що мають максимальну інформативність за поточних умов спостереження.

#### *4.3.4 Структурно-логічна послідовність процесів обробки даних в інформаційній технології гібридного розпізнавання образів*

Процес гібридного розпізнавання образів для обробки неоднорідних даних в СС повною мірою реалізований в інформаційній технології, яка представлена в структурно-логічній послідовності процесів обробки даних, представленої на рисунку 4.4.



Рисунок 4.4 – Структурно-логічна послідовність процесів обробки даних в інформаційній технології гібридного розпізнавання образів

На основі вхідних даних про стан складної системи формується сукупність образів ознаки яких мають різну природу виникнення. Також, в момент отримання характеристик СС, додатково отримуються дані про

поточну ситуацію навколишнього середовища для подальшого визначення поточного ступеня інформативності кожного образу з сукупності  $P$ .

Отримана сукупність образів і характеристик навколишнього середовища, відповідно до ситуаційно-подійної моделі гібридного розпізнавання, використовується для визначення характеристик інформативності кожного образу за поточних умов спостереження за об'єктом розпізнавання. Характеристика інформативності складається з двох складових:  $SICH$  (стаціонарної) і  $NSICH$  (нестаціонарної). Стаціонарні характеристики інформативності апріорно визначені та витягуються з бази знань, а нестаціонарні характеристики визначаються відповідно до поточних зовнішніх умов  $EC$ .

Надалі, здійснюється селекція образів на основі їх поточних значень характеристик інформативності. Таким чином, від подальшого аналізу та прийняття рішення усуваються образи, які за поточних умов спостереження за об'єктом розпізнавання мають характеристики менші за попереднє завдане значення. Відповідно, зменшується потужність інформаційного поля подальшого аналізу, що дозволяє зменшити кількість співставлень ознак при прийнятті рішення щодо класифікації.

Виділені для аналізу найбільш інформативні образи підлягають подальшій селекції відповідно до значень їхніх характеристик інформативності. Таким чином, визначається, черга обробки образів  $CC$ , згідно етапу 5 інформаційної технології, опис якої представлений у пункті 4.3.2.

На наступному етапі здійснюється ранжування ознак образів  $CC$ , відповідно до запропонованих у підрозділі 2.5 інформаційної моделі та методу визначення послідовності обробки даних. З цією метою здійснюється визначення динаміки зміни ознак у порівнянні з попереднім етапом. Для цього з бази знань запитуються збережені на попередньому етапі дані. Так, ознаки, які з попереднього етапу класифікації набули більшої зміни



отримують пріоритет у обробці. Відповідно до цього, формується багатопоточна черга обробки даних, де кожен потік містить ознаки кожного з інформативних образів, а ознаки цих образів вибудовуються у відповідності до зменшення ступеня їхньої зміни, відповідно до попереднього етапу розпізнавання.

Сформована черга обробки неоднорідних даних надається для класифікації, в результаті якої визначається відповідність стану СС до передвизначених класів.

Слід зазначити, що відповідно до розробленого методу екстенсіонально-інтенсіонального аналізу, за кожним образом максимальний пріоритет в обробці мають узагальнені, екстенсіональні ознаки, що дозволяє здійснити спробу класифікації з завданням рівнем достовірності на основі мінімальної кількості співставлень. Якщо екстенсіональний аналіз не дає збігу класифікації із заданим рівнем достовірності, то надалі здійснюється інтенсіональний аналіз деталізованих ознак, в результаті чого уточнюється класифікація. В результаті класифікації за методом гібридного розпізнавання на основі пошуку груп ідентичних класифікацій, представленого у підрозділі 3.1, процес співставлення ознак здійснюється до моменту часу коли одна чи декілька груп ідентичних класифікацій не досягне апріорно заданої чисельності. Цей процес відповідає етапу 8 представленого у пункті 4.3.2 опису інформаційної технології гібридного розпізнавання образів.

Якщо в результаті класифікації за двома та більше класами отримується апріорно завданий рівень достовірності (чисельність груп ідентичних класифікацій) то такий багатоальтернативний результат, обробляється відповідно до розробленої інформаційної моделі підтримки прийняття рішення, яка представлена в пункті 3.1.5.

Таким чином, багатоальтернативний результат ранжується у порядку зменшення його достовірності та надається для подальшого експертного висновку та остаточного прийняття рішення. Також, до бази знань заносяться

відомості щодо поточного рівня ознак для того, щоб на наступному етапі ця інформація була запитана для ранжування ознак інформативних образів та корекції черги обробки даних.

#### **4.4 Рекомендації щодо впровадження інформаційної технології гібридного розпізнавання образів**

Розроблена ІТ гібридного розпізнавання образів для консолідованої обробки даних в СС може бути використана у будь-яких прикладних застосуваннях. Аналіз можливих застосувань гібридного розпізнавання образів показує достатньо широкий спектр вирішуваних задач в області технічних систем [47-49, 51, 70, 71, 46, 72-74, 5, 12, 45, 50, 52, 75, 76, 84, 85], екологічних систем [33, 77-79], агропромислового комплексу [80, 81], геології та геодезії [82-83], економічних та геополітичних систем [87-91], систем соціологічних досліджень [92, 93].

Основними вимогами до використання розробленої ІТ є наявність широкого спектру неоднорідних даних, що характеризують СС, а також, виділення тих характеристик, що мають різну природу виникнення. Як правило переважна більшість СС, відповідають цієї умові. Однак, у кожному окремому випадку, може ставати необхідність у додаванні характеристик, що реєструються для подальшого аналізу і класифікації. Це, відповідно, може вимагати розширення номенклатури використаних пристроїв реєстрації.

Однак, з іншого боку, перегляд наявного переліку технічних засобів реєстрації характеристик, може мати зворотний ефект: додавання до аналізу характеристик різної природи виникнення може створити умови до виключення частини характеристик однієї природи виникнення. У будь-якому випадку, для кожного практичного застосування ІТ підлягає експертній оцінці набір інформаційних характеристик СС з метою

визначення раціональної сукупності для побудови образів та класифікації засобами гібридного розпізнавання.

Для підтвердження ефективності розроблених технічних рішень, в розділі 5 здійснюється розробка прикладних інформаційних технологій гібридного розпізнавання різного застосування.

#### **Висновки до розділу 4**

В розділі 4 вирішено сьоме завдання дисертаційного дослідження, а саме: на основі розроблених методів і моделей, створення інформаційної технології гібридного розпізнавання образів для консолідованої обробки даних в складних системах. Розроблені умови використання розробленої ІТ, описані інформаційні процеси на кожному етапі ІТ, представлений процес формування черги обробки даних. Приділена увага рекомендаціям щодо впровадження розробленої ІТ. Основні висновки полягають в наступному.

Для варіанта використання системи «жорсткого» реального часу, обробка даних та класифікація здійснюється до моменту отримання бажаного рівня достовірності результату, або до завершення часу, виділеного на прийняття рішення.

Для варіанта використання системи «м'якого» реального часу, обробка даних здійснюється до моменту отримання бажаного достовірного рішення, або до моменту перевищення завданого максимального часу обробки на певну величину.

Для організації ІТ гібридного розпізнавання образів доцільне використання корпоративної багатозадачності, коли процес обробки ознаки з черги і визначення поточного рівня достовірності класифікації буде передавати управління процесу обробки наступної ознаки з черги, якщо апіорно заданий рівень достовірності не досягнутий, а час, виділений на

обробку та класифікацію за умовами «жорсткого» чи «м'якого» реального часу не вичерпаний.

Використання корпоративної багатозадачності забезпечує передбачуваність процесу обробки в часовій області, та високу реактивність на зміну вхідних даних, що найбільш актуально для ситуаційного підходу до управління станом СС.

Специфіка ІТ гібридного розпізнавання образів вимагає наступних основних вимог до апаратно-програмних засобів: надійність роботи пристроїв реєстрації ознак СС, робота в широкому діапазоні зовнішніх умов, наявності розвиненої системної служби часу, легкої конфігурованості, простоти системи команд і способів адресації, розвиненої системи переривань, широкої номенклатури інтерфейсів та можливості підключення великої кількості зовнішніх пристроїв реєстрації характеристик, наявності засобів попередньої обробки і нормування характеристик СС.

Розроблена структурна організація системи обробки даних, в якій реалізована інформаційна технологія гібридного розпізнавання образів для консолідованої обробки даних в СС.

Для реалізації ІТ гібридного розпізнавання доцільне використання мікроядерної архітектури. В цьому випадку забезпечується малий час реакції та добра масштабованість при зміні конфігурації периферійних пристроїв реєстрації характеристик СС та зовнішніх умов. Недолік мікроядерної архітектури, пов'язаний з частими переключеннями режимів усувається використанням окремої, спеціалізованої комп'ютерної системи, на якій встановлюється СРЧ з мікроядерною архітектурою і яка здійснює виключно обробку сформованої черги ознак СС.

На основі розроблених моделей і методів гібридного розпізнавання образів, представлений процес формування черги обробки неоднорідних даних, в якій образи і ознаки СС ранжуються у порядку зменшення їхньої інформативності. В цьому випадку, у будь який момент здійснення ІТ,

забезпечується обробка найбільш інформативних даних, що дозволяє швидше наближуватися до бажаного рівня достовірності класифікації.

Розроблено ІТ розпізнавання образів, що заснована на принципі гібридного розпізнавання та вирішує задачу консолідованої обробки неоднорідних даних в складних системах з різномірним проявом перешкод та викривлень. В розділі представлені етапи здійснення цієї ІТ.

Розроблена ІТ гібридного розпізнавання образів для консолідованої обробки даних в СС може бути використана у будь-яких прикладних застосуваннях.

Основними вимогами до використання розробленої ІТ є наявність широкого спектру неоднорідних даних, що характеризують СС, а також, виділення тих характеристик, що мають різну природу виникнення.

## РОЗДІЛ 5

### ПРИКЛАДНІ ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ ГІБРИДНОГО РОЗПІЗНАВАННЯ ОБРАЗІВ

#### 5.1 Інформаційна технологія аналізу просторового розподілу температури коксового пирога для коксохімічного виробництва

Коксові печі служать для випалювання коксу з кам'яного вугілля і отримання, як побічного продукту, коксового газу. Конструктивно, коксова піч представляє собою батарею з послідовним розміщенням камер спікання. На рисунку 5.1 наведений загальний вигляд коксової батареї.



Рисунок 5.1 – Загальний вигляд коксової батареї

Кожна камера спікання з обох боків оснащена дверима, через які здійснюється завантаження шихти та вивантаження готового коксу. Сторону завантаження шихти називають машинною стороною, а сторону вивантаження готового коксу – коксовою стороною. На рисунку 1.2 представлена коксова батарея з боку завантаження шихти (машинна сторона). На передньому плані відображається коксова машина, що переміщується уздовж коксової батареї і здійснює завантаження шихти до камер спікання. Вигляд камери спікання з відчиненою двір'ю наведено на рисунку 5.2. Фотографія зроблена з коксової сторони батареї. З цього боку здійснюється виштовхування готового коксового пирога до гасильного вагону.



Рисунок 5.2 –Вигляд камери спікання з відчиненою двір'ю

Режим роботи коксових печей циклічний: усі технологічні етапи (завантаження вугільної шихти, коксування і видача готового коксу) періодично повторюються.

Для нормального ходу процесу коксування потрібен певний температурний режим в огрівальних простінках і гідравлічний режим в камерах. Стабілізація гідравлічного режиму необхідна для підтримки постійних співвідношень тиску в камері коксування, простінках і довкіллі. Герметичність камери спікання запобігає горінню коксу в камері або витoku газу з камер.

Економічність використання палива на обігрів визначається оптимальним коефіцієнтом надлишку повітря для горіння опалювального газу.

Вихідним параметром, що характеризує гідравлічний режим газової мережі і роботи газодувки, може бути прийняте розрідження в мережі перед газодувкою [218].

У якості вихідних величин коксохімічного виробництва, для агрегатів уловлювання цінних хімічних продуктів з коксового газу, використовують визначені умови його протікання та температурні параметри процесу. Такий вибір вихідних величин пов'язаний з відсутністю можливості прямого контролю результатів процесу. У цьому спостерігається одна з основних труднощів організації ефективного технологічного процесу коксопікання: з одного боку температурний режим є основним і вирішальним для організації технологічного процесу, з іншого боку прямий контроль і фіксація температури неможливі. Крім цього, існує безліч факторів, що впливають на технологічні режими, при чому, зміна значної кількості цих факторів має стохастичний характер. У цьому випадку, будь які повні аналітичні залежності, що описують температурний режим, відсутні. Одним з можливих шляхів вирішення цієї проблеми може бути використання інформаційної системи аналізу і управління температурним режимом коксових печей, яка на



основі інтелектуальних алгоритмів дозволить визначити розподіл температури в процесі коксування і виробити управляючі впливи для наступного технологічного циклу.

При обігріві печей батареї коксовим і доменним газами за типовою схемою для стабілізації теплового режиму обігріву і температурного режиму печей регулюють співвідношення витрат опалювальних газів на усю батарею в цілому, а також їх втрати на машинну і коксову сторони кожної печі. Схема регулювання розходу коксового і доменного газів для обігріву коксових печей наведена на рисунку 1.4 [218].

Суміш хімічних продуктів (у вигляді парів і газів), що виділяються в процесі коксування, називають прямим коксовим газом [218]. З печей він поступає в газозбірник і далі за допомогою газодувки спрямовується на конденсацію і витягування цінних хімічних продуктів (аміаку, смоли, бензолу та ін.).

Розжарений кокс в гасильному вагоні охолоджують (гасять) під вежею гасіння водою або інертним газом («мокрим» або «сухим» способом)[218].

У процесі гасіння виділяється конденсат, який розділяють на аміачну воду і кам'яновугільну смолу. Охолоджений газ втягується потужними газодувками і нагнітається в систему уловлювання і очищення. Тут з газу уловлюють пари сірого бензолу, значну частину аміаку, що залишилася в газі, сірководень і ціаністи з'єднання. Сирій бензол, кам'яновугільну смолу і аміачну воду переробляють в товарні продукти [218].

Очищений коксовий газ, що називається зворотним [218], використовують як паливо, а також в якості сировини для синтезу аміаку.

Після завершення процесу гасіння кокс скидають на похилу рампу, де він висихає. Охолоджений і висохлий кокс по транспортеру направляють на сортування.

Після завантаження шихти до камери спікання здійснюється її багаточасовий (близько 13-18 годин) прогрів при температурі близько 1100-

1200 °С [218]. Прогрів здійснюється за допомогою пропускання коксового газу через міжстінний проміжок камер спікання.

Слід зазначити, що стабілізація температурного режиму коксової печі пов'язана з відсутністю можливості прямого контролю температури в камері спікання. В цьому спостерігається одна з основних проблем організації ефективного технологічного процесу кокоспікання: з одного боку температурний режим є основним і вирішальним для організації технологічного процесу, з іншого боку, прямий контроль і фіксація температури неможливі. Крім цього, існує безліч факторів, що впливають на технологічні режими, при чому, зміна значної кількості цих факторів має стохастичний характер. У цьому випадку, будь які повні аналітичні залежності, що описують температурний режим, відсутні.

Базова інформаційна технологія передбачає вимір температури коксового пирога на етапі його виштовхування з печі. Вимір здійснюється пірометром, що встановлюється на осі виштовхування. Цей процес проілюстрований на рисунку 5.3.

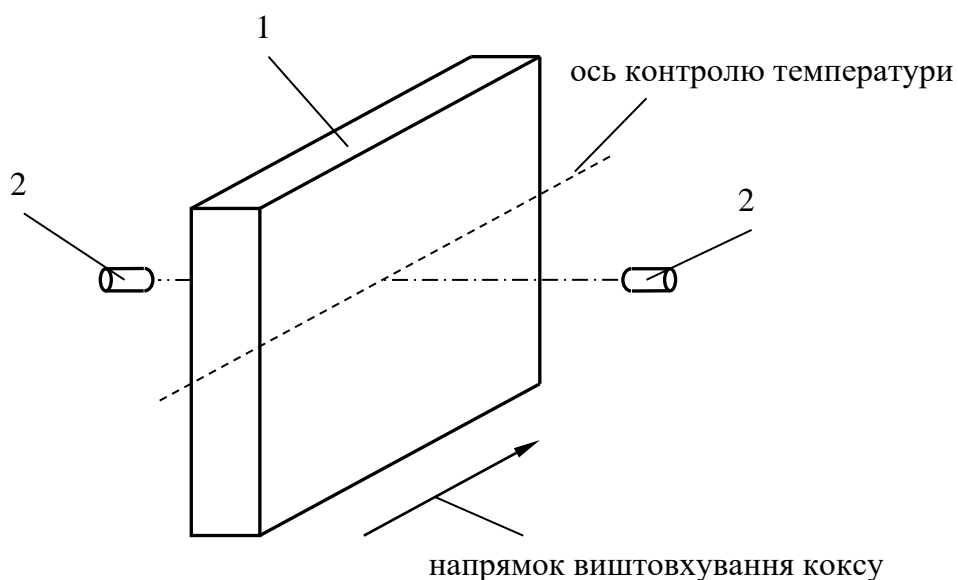


Рисунок 5.3 – Вимірювання температури коксового пирога

Зовнішній вигляд, встановленого пірометричного датчика на стороні виштовхування коксу з підключеними комунікаціями (канал охолодження та інтерфейсний кабель), представлений на рисунку 5.4.



Рисунок 5.4 – Зовнішній вигляд підключеного пірометричного датчика на стороні виштовхування коксу

Таким чином, кожен пірометр дозволяє сформувати сигнал струму

$$i = f(t), \quad (5.1)$$

де  $i$  – величина струму пірометра, що пропорційна температурі у кожній точці осі виміру;

$t$  – часова характеристика.

З іншого боку, враховуючи передавальну характеристику пірометра характеристика температури в кожній точці може бути визначена як

$$t = f(i), \quad (5.2)$$

Визначення координати точки з температурою  $t$  можливе з урахуванням швидкості виштовхування коксового пірога з камери спікання.

$$X_{x0} = \mathcal{G} \cdot (t - t_{t0}), \quad (5.3)$$

де  $X_{x0}$  – координата точки по осі виміру відносно початкової координати краю коксового пірога;

$\mathcal{G}$  – швидкість виштовхування коксового пірога з камери спікання;

$t$  – поточний час реєстрації точки, для якої визначається координата;

$t_{t0}$  – початковий час реєстрації температури уздовж осі виміру.

Фрагмент характеристики струму, зафіксованої стаціонарним самописцем, наведений на рисунку 5.5.

На цьому рисунку, цикли виміру температури уздовж осі, розділяються вертикальними спадами сигналу до нуля. Порівняння залежності сигналу пірометра на різних циклах вивантаження коксу вказує, що для однієї і тієї ж камери спікання, на різних технологічних циклах розподіл температури відрізняється. При чому, враховуючи значну кількість факторів, що впливають на температуру у кожній точці, залежність (5.1) фактично є функцією багатьох аргументів і аналітичне визначення її практично неможливе.

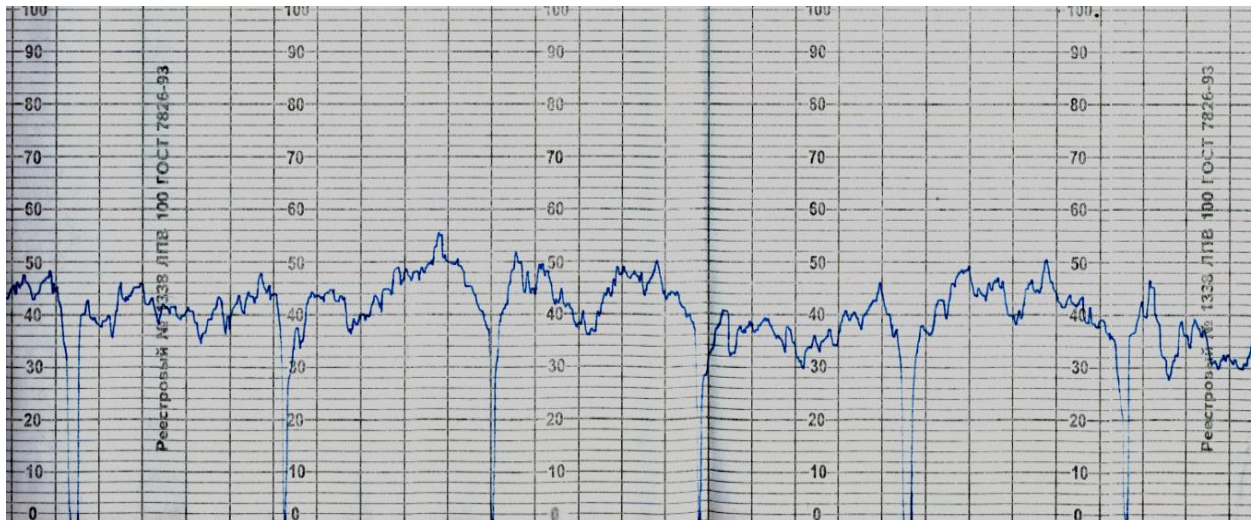


Рисунок 5.5 – Часова характеристика зміни сигналу струму пірометра

З урахуванням залежностей (5.1)-(5.3) проводиться попередня обробка даних і визначається часова прив'язка параметрів розподілу температури. Результати розрахунку одного циклу виміру, та прив'язка до реального часу, наведені на рисунку 5.6.

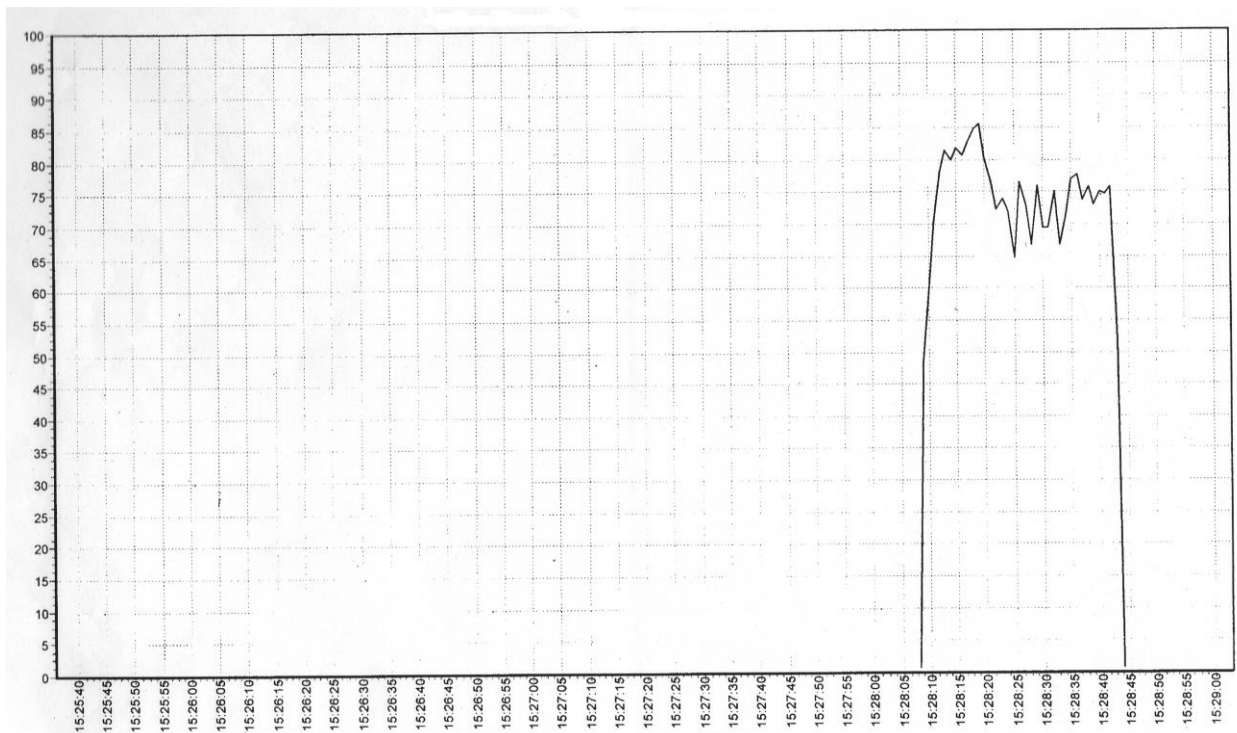


Рисунок 5.6 – Отримана характеристика розподілу температури

Для визначення кількісних і якісних показників режиму обігріву батареї проводяться регламентні заміри температур у вертикалі батареї. На підставі аналізу отриманих даних проводиться корекція режиму обігріву батареї. У таблиці середніх температур (рисунок 5.7) у всіх вертикалях групи простінків можна побачити заміри відповідних температур.

Средние температуры во всех вертикалях. Группа простенков 102-111

Изменить период \* Вид \* Вычисления \*

Отчет за период с : 17.02.2013  
по : 17.03.2013

Вычисления простенки в особом режиме выключены.

№ Простенка	Вертикали																												
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	КС/МС
102	1223	1251	1266	1259	1268	1264	1267	1263	1272	1268	1273	1274	1283	1277	1274	1278	1286	1293	1294	1312	1305	1317	1302	1289	1299	1317	1332	1317	54
103	1239	1261	1257	1251	1265	1257	1272	1263	1271	1250	1267	1237	1273	1265	1264	1264	1273	1270	1280	1280	1294	1291	1282	1272	1287	1289	1287	1281	28
104	1284	1302	1269	1266	1279	1279	1288	1285	1300	1294	1293	1298	1298	1296	1275	1299	1288	1302	1289	1309	1307	1309	1301	1292	1249	1297	1304	1307	24
105	1291	1300	1274	1231	1282	1282	1297	1293	1308	1302	1303	1309	1311	1307	1291	1307	1297	1301	1307	1319	1311	1307	1287	1250	1272	1279	1317	1289	14
106	1257	1276	1258	1252	1261	1257	1259	1267	1251	1257	1260	1258	1249	1256	1253	1267	1264	1263	1280	1292	1300	1298	1298	1284	1279	1297	1306	1302	31
107	1218	1240	1242	1234	1254	1252	1254	1259	1259	1264	1261	1263	1255	1263	1257	1266	1258	1267	1260	1268	1292	1291	1276	1266	1225	1269	1297	1277	32
108	1281	1299	1268	1254	1284	1213	1284	1291	1276	1289	1289	1289	1294	1305	1307	1314	1311	1318	1321	1332	1328	1330	1295	1299	1293	1299	1326	1297	39
109	1272	1281	1257	1254	1281	1280	1279	1287	1286	1290	1291	1286	1281	1283	1301	1302	1219	1308	1318	1319	1321	1313	1281	1269	1216	1288	1339	1298	26
110	1225	1269	1250	1258	1243	1269	1268	1273	1272	1279	1262	1279	1288	1289	1286	1291	1293	1302	1296	1311	1303	1306	1268	1262	1266	1282	1280	1296	33
111	1241	1263	1268	1263	1277	1278	1287	1284	1286	1276	1289	1288	1301	1292	1317	1311	1309	1312	1308	1313	1316	1288	1287	1283	1261	1295	1294	1261	4
T <sub>ср</sub>	1253	1274	1260	1252	1269	1263	1275	1276	1278	1276	1278	1278	1283	1283	1282	1289	1279	1293	1295	1305	1307	1305	1287	1276	1264	1291	1310	1292	28

Средняя температура по вертикалям: 1281°C

Пользователь: Оператор

АСУТП "Кокс" © 2005 ООО "Темпла МС"

Пользователь: Оператор

Смена №2 Бригада №2

17.03.2006 09:56:40

Рисунок 5.7 – Таблица отриманих середніх температур

Замір температур у вертикалі батареї здійснюється лише обслуговуючим персоналом батареї за допомогою оптичного пірометра. Після закінчення заміру дані передаються на операторську станцію. Програма моніторингу температурного режиму коксових печей, що працює на операторській станції, здійснює прийом даних заміру, їх обробку та

збереження у системній базі даних сервера. Приклад графіку середніх температур у всіх вертикалях групи простінків наведений на рисунку 5.8.

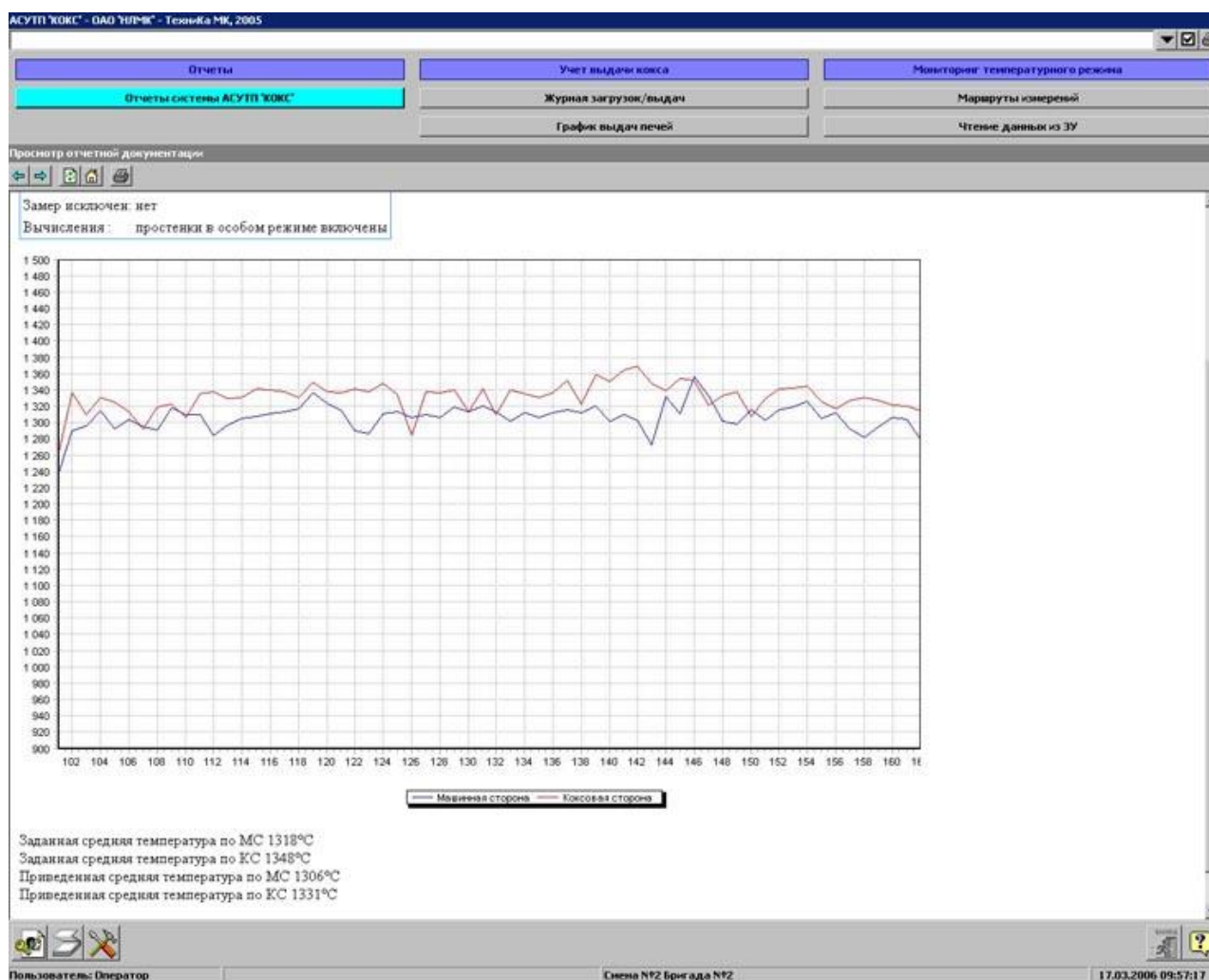


Рисунок 5.8 – Графік середніх температур у всіх вертикалях групи простінків

Проведений аналіз існуючої ІТ моніторингу теплового режиму коксових печей вказав на низку недоліків, основні з яких:

- контроль температури здійснюється тільки на декількох осях (горизонталях) коксового пирогу, що не дає можливості судити про просторовий розподіл температури по всій його площині;
- відсутність ефективних методів і засобів інформаційної підтримки операцій автоматизованого діагностування стану коксових печей та наявності в них дефектів, а також визначення динаміки їхнього розвитку.

У світовому досвіді, згідно наявних сучасних винаходів, існує прагнення до збільшення кількості горизонталей уздовж яких здійснюється контроль температури. Так, наприклад, китайські винахідники [219] пропонують збільшення осей контролю до 20-30 шт. При цьому збільшується інформативність процесу контролю температури так як звужуються неконтрольовані зони коксового пирога, що лежать між сусідніми контрольованими горизонталями. Однак такий підхід вимагає значного збільшення кількості пірометричних датчиків контролю температури, що значно підвищує витрати на впровадження такої системи, враховуючи значну вартість пірометрів для подібного застосування та необхідного додаткового обладнання. Вартість пірометрів такого класу лежить у діапазоні 45-60 тис. грн. за одиницю (з урахуванням вартості необхідного устаткування). Крім цього, наявність значної кількості пірометричних датчиків ускладнює конструкцію системи контролю температури, обмежує доступ персоналу до технологічного обладнання та ускладнює його обслуговування.

Для виключення вказаних вище недоліків, була розроблена прикладна ІТ гібридного розпізнавання образів, коли для фіксації просторового розподілу температури використовується система технічного зору [75, 76, 220]. В цьому випадку, в якості пристрою реєстрації використовується відеокамера, що встановлюються на дільниці видачі готового коксу. Це виключає необхідність встановлення значної кількості пірометрів для реєстрації температури у будь-якій точці коксового пирога.

Однак, оптична реєстрація поверхні коксового пирога здійснюється за умови наявності значного рівня перешкод, що пов'язані з задимленістю навколишнього середовища, атмосферними та погодними проявами тощо. Крім цього, точність реєстрації атрибутів зображення знаходяться у значній залежності від побічних джерел світлового випромінювання, а також структурованості освітлення. Все це значно впливає на отримання точних залежностей співставлення атрибутів зображення відповідним значенням



температури. Використання цього підходу можливе за умови використання гібридного розпізнавання, коли температурні характеристики коксового пирога будуть завдані двома образами різної природи виникнення. Ознаки першого образу отримуються за допомогою системи технічного зору, а другого – пірометричним датчиком. Перший образ надає інформацію про просторовий розподіл температури по поверхні коксового пирога, а другий надає температурні характеристики лінійного розподілу температури по осі виштовхування коксового пирога.

На відміну від базової ІТ контролю температури, розглянутого вище, інформація від пірометричного датчика використовується не як основний інформаційний потік, а як додатковий, який дозволяє засобами гібридного розпізнавання образів здійснити корекцію атрибутів зображення та уточнити дані щодо просторового розподілу температури коксового пирога. Основний інформаційний канал обумовлений системою технічного зору.

В цьому випадку, модель опису СС (2.1) описується системою двох образів  $P_1$  і  $P_2$  виду

$$\left\{ \begin{array}{l} P_1 = \begin{pmatrix} X_{1,1} & X_{1,2} & \dots & X_{1,n} \\ X_{2,1} & X_{2,2} & \dots & X_{2,n} \\ \dots & \dots & \dots & \dots \\ X_{m,1} & X_{m,2} & \dots & X_{m,n} \end{pmatrix}, \\ P_2 = (X(t)_1, X(t)_2, \dots, X(t)_k), \end{array} \right. \quad (5.4)$$

де  $\{X_{m,n}\}$  – множина інформаційних ознак образу  $P_1$ , що характеризують просторовий розподіл температури по поверхні коксового пирога;

$\{X(t)_k\}$  – множина інформаційних ознак образу  $P_2$ , що характеризують лінійний розподіл температури на одній з осей коксового пирога, який отримується за схемою, згідно рисунку 5.3.

Множина інформаційних ознак образу  $P_1$ , фактично, представляє собою матрицю растрового зображення, що отримується системою технічного зору. Значення  $n$  і  $m$  обумовлені роздільною здатністю матриці системи відеореєстрації. Для рішення задачі аналізу просторового розподілу температури, розмірності параметрів  $n$  і  $m$ , як правило, є надлишковими, що не сприятиме мінімізації часової складності процесу аналізу образу  $P_1$ . З метою зменшення розмірності інформаційного поля аналізу ознак образу  $P_1$  доцільне використання методу динамічної зміни деталізації на основі низькочастотної фільтрації, який запропонований у [24]. В результаті такого перетворення буде отриманий образ  $P_1'$ , в якому розмірності відповідають умовам

$$\begin{aligned} n' &\ll n, \\ m' &\ll m, \end{aligned} \quad (5.5)$$

де  $m'$  і  $n'$  – розмірності перетвореного образу  $P_1'$ .

В моделі (5.4), множина  $P_2$  представляє собою дискретизований у часі лінійний розподіл температури, що отримується за допомогою пірометричного датчика.

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.

Співставлення образів  $P_1'$  і  $P_2$  необхідно здійснювати шляхом часового суміщення ознак, що отримані у визначені проміжки часу в процесі виштовхування коксового пирога. Задача часового суміщення спрощується завдяки тому, що швидкість виштовхування є постійною і регламентується технологією коксування.

Співставленню з ознаками образу  $P_2$  піддаються ознаки образу  $P_1'$ , що характеризують розподіл температури на осі отримання лінійного розподілу. Цей процес проілюстрований на рисунку 5.9.

На рисунку 5.9 чорні сегменти ілюструють ознаки образів  $P_1'$  і  $P_2$ , що підлягають співставленню. Просторово вони характеризують зону коксового пирога, що лежить на осі отримання лінійного розподілу температури.

Множину ознак образу  $P_1'$ , що підлягають співставленню, позначимо як репрезентативну  $\{X_{m,n}^R\}$ , при чому

$$\{X_{m,n}^R\} \subset \{X_{m,n}\}. \quad (5.6)$$

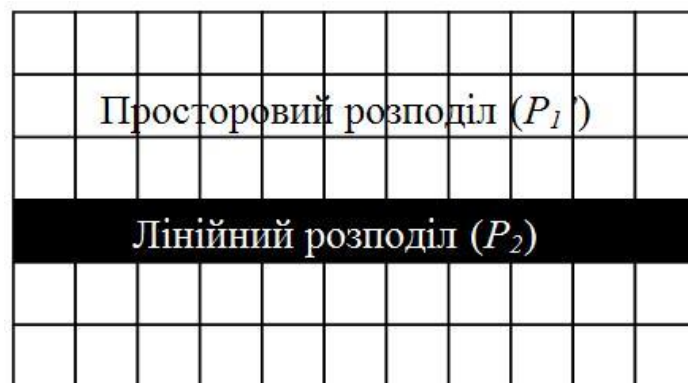


Рисунок 5.9 – Співставлення ознак образів  $P_1'$  і  $P_2$  на осі отримання лінійного розподілу температури

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

$$\forall X_{m,n}^R \in P_1', T(X_{m,n}^R) = T(X(t)_k). \quad (5.7)$$

Ця відповідність витікає з того, що співставленню підлягають ознаки двох образів, які характеризують одну область коксового пирога, а відповідно, обидва образи характеризують одну й ту ж саму температуру.

У випадку, якщо для окремих областей коксового пирога умова (5.7) не виконується, то температурна характеристика, що відповідає поточному атрибуту зображення, присвоюється рівною до температурної характеристики відповідної ознаки образу  $P_2$ . Так, наприклад, якщо для деякої області коксового пирога, що лежить на осі контролю лінійного розподілу температури  $T(X_{m,n}^R) = 1200$  °C, а відповідна характеристика образу  $P_2$ :  $T(X(t)_k) = 1150$  °C, то необхідно провести корекцію через присвоєння  $T(X_{m,n}^R) = T(X(t)_k) = 1150$  °C.

Після співставлення репрезентативних ознак образу  $P_1'$  з ознаками образу  $P_2$  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 [76].

На рисунку 5.10 представлена структурно-логічна схема послідовності операцій в прикладній інформаційній технології гібридного розпізнавання образів для аналізу просторового розподілу температури коксового пирога.

Згідно представленої структурно-логічної схеми, в інформаційній технології аналізу просторового розподілу температури коксового пирога, спочатку, здійснюється реєстрація даних про просторовий і лінійний розподіли температури та на основі їх формуються образи  $P_1$  і  $P_2$ , як це було представлено вище. Образ  $P_1$  містить ознаки, що характеризують просторовий розподіл температури, а  $P_2$  – лінійний розподіл температури.

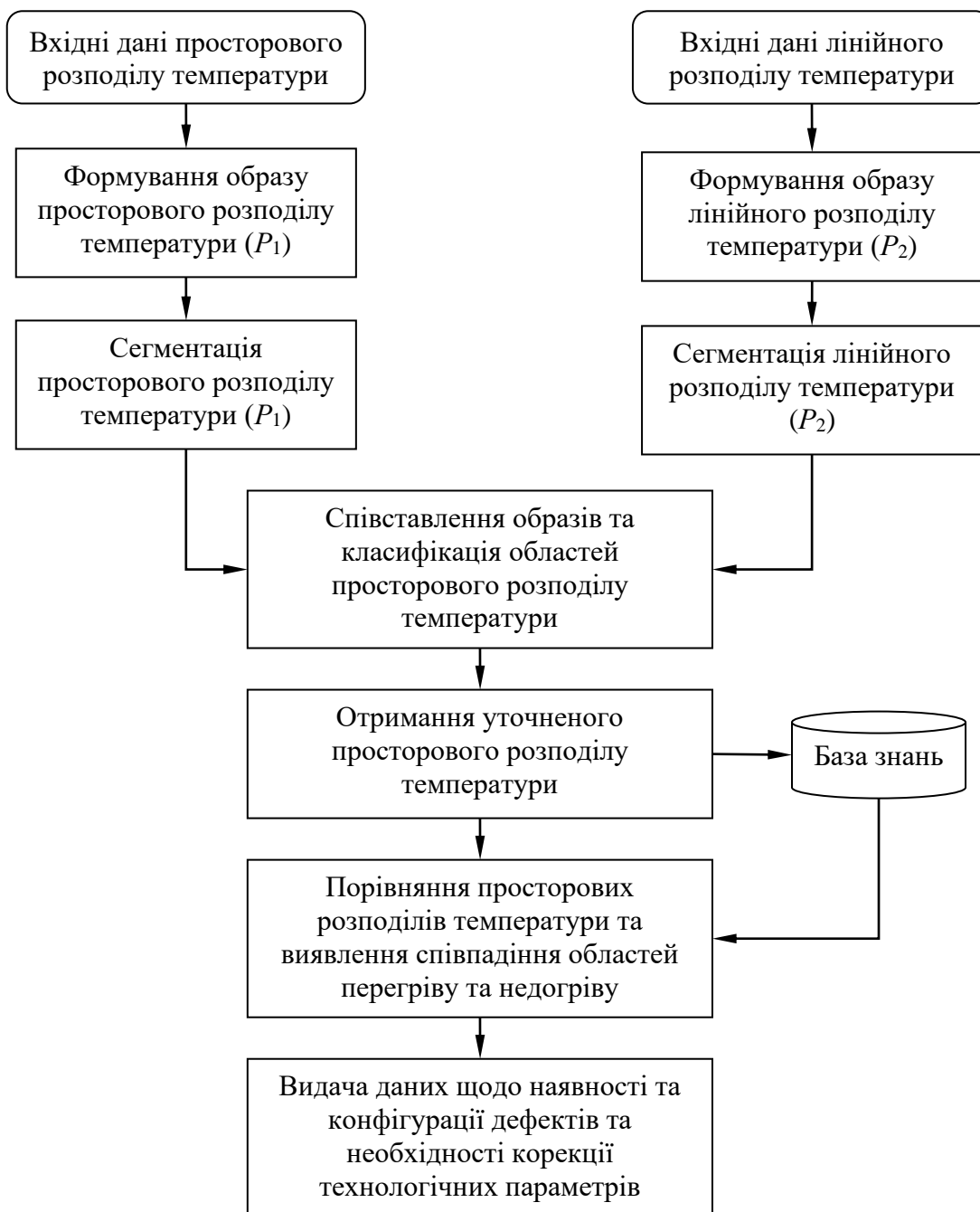


Рисунок 5.10 – Структурно-логічна схема послідовності операцій в прикладній інформаційній технології гібридного розпізнавання образів для аналізу просторового розподілу температури коксового пирога

Слід зазначити, що технічні засоби реєстрації температурних характеристик (система технічного зору та пірометрична система вимірювання температури) мають надлишкову роздільну здатність для такої практичної задачі. Це призводить до значного збільшення обчислювального

навантаження на апаратно-програмний комплекс та збільшує часову складність аналізу даних. Тому, в аспекті рішення цього питання, після формування образів просторового і лінійного розподілів здійснюється їхня сегментація для формування усереднених областей температурних характеристик. Алгоритм визначення температури в середині області є тривіальним визначенням середнього значення. Однак, основною вимогою до визначення областей є просторове співпадіння для просторового і лінійного розподілу, для того щоб реалізувати їхнє співставлення, відповідно до рисунку 5.9. Таким чином, для просторового і лінійного розподілу обирається однаковий шаг сегментації, як це показано на рисунку 5.11.

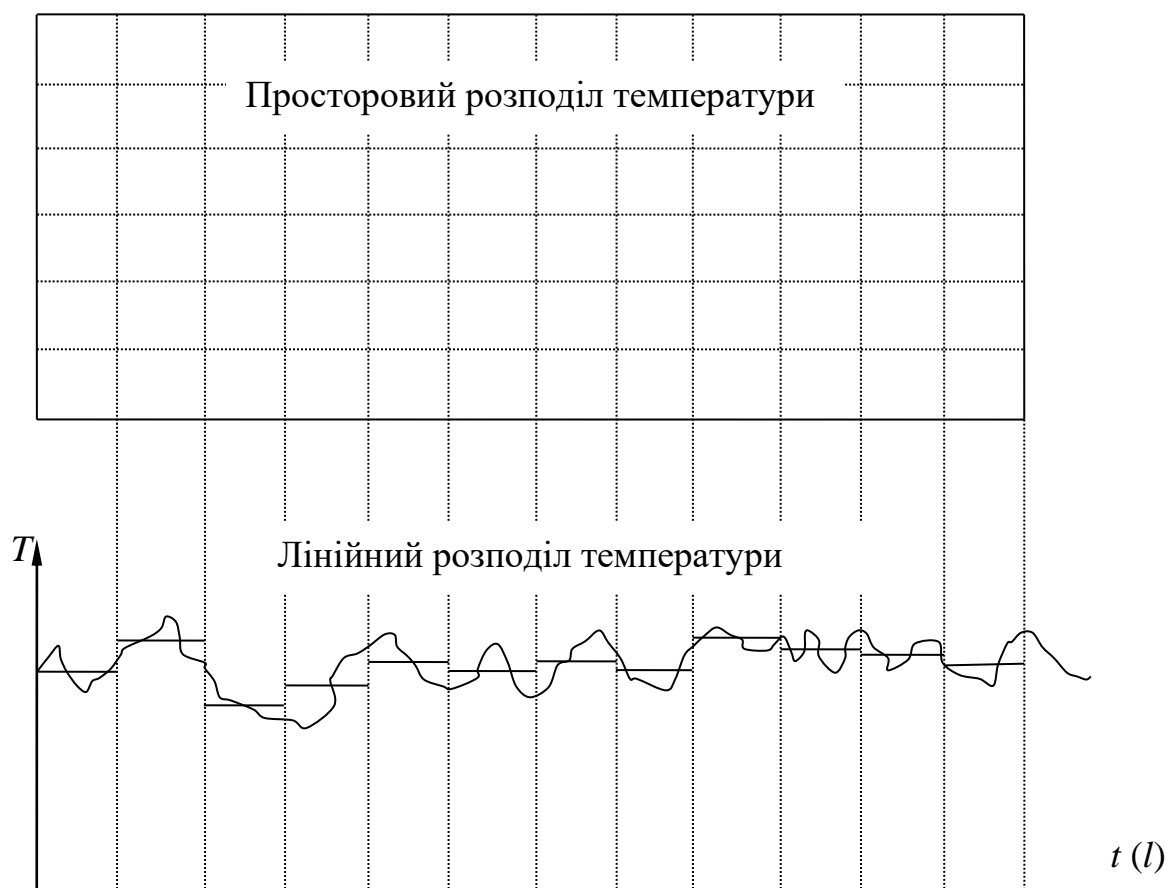


Рисунок 5.11 – Відповідність сегментації просторового та лінійного розподілу температури коксового пирога

Так на кожному шагу сегментації як лінійного так і просторового розподілів визначається середня температура. Надалі алгоритми обробляють не окремі точки розподілів, а їхні поєднання у вигляді областей, обмежених границями сегментації. Лінійний розподіл при цьому приймає кусочно-лінійчатий вигляд, де на кожному інтервалі сегментації визначається єдина середня температура, як це проілюстровано на рисунку 5.11.

Надалі, згідно структурно-логічній схемі, представленій на рисунку 5.10, попередньо оброблені образи  $P_1$  і  $P_2$  розподілу температур надаються для співставлення та класифікації просторового розподілу температури. Наглядно процес співставлення представлений на рисунку 5.9.

Після співставлення, відповідно до умови (5.7), здійснюється корекція просторового розподілу температури з урахуванням репрезентативних даних лінійного розподілу. Після цього, уточнена характеристика просторового розподілу температури поміщається до бази знань, для використання на наступних циклах виготовлення коксового пирога.

На наступному етапі інформаційної технології, відповідно до схеми на рисунку 5.10, здійснюється порівняння просторових розподілів температури та виявлення співпадіння областей перегріву та недогріву. З цією метою, з бази знань витягуються дані про просторовий розподіл температури на попередніх технологічних циклах та здійснюється їх співставлення. Якщо на температурних характеристиках різних технологічних циклів виявляються закономірні зони локального перегріву або недогріву і конфігурація цих зон має збіги, то можуть бути прийняті наступні висновки:

– стаціонарний характер зон локального недогріву вказує на недостатній розхід опалювального газу, що може бути виправлено корекцією технологічних параметрів на наступний технологічний цикл;

– наявність стаціонарних за конфігурацією зон локального перегріву вказує на дефекти в кладці коксової печі, що призводить до втрати

герметизації камери спікання. Цей аспект вказує на технічний дефект, який може бути виправлений тільки виконанням ремонтних робіт.

При цьому слід зазначити, що вказана інформаційна технологія спрощує аналіз і прогнозування розвитку дефектів коксових печей батареї, що дозволяє спростити розробку регламенту проведення сервісного обслуговування та ремонтних робіт. На сьогодні, така діагностика проводиться тільки підчас візуального експертного огляду кладки печі, що виконується із зупинкою технологічного процесу та призводить до фінансових втрат.

Перевірка точності отримання розподілу температур, відповідно до представленої ІТ, 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.

Експериментальна установка відрізняється наявністю другого пірометричного датчика, за яким додатково здійснюється перевірка точності корегування просторового розподілу, що характеризується образом  $P_2$ . Схема експериментальної установки наведена на рисунку 5.12. Цифрою 1 позначений коксовий піріг, за тепловим полем якого здійснюється моніторинг. Цифрою 2 позначена система технічного зору, яка здійснює реєстрацію ознак образу  $P_1$ . Пірометричний датчик 3 на горизонтальній осі 4 здійснює реєстрацію лінійного розподілу температури, який представляє образ  $P_1$ . Пірометричний датчик 5, що здійснює реєстрацію лінійного розподілу температури на горизонтальній осі 6, введений тільки для експериментального дослідження ефективності функціонування ІТ моніторингу стану коксової печі, і в робочому режимі він непотрібний. Лінійний розподіл температури, що отримується датчиком 5 використовується для порівняння результатів корекції просторового



розподілу. Точність корекції визначається співставленням лінійного розподілу 6 з відповідними ознаками корегованого образу  $P_1$ .

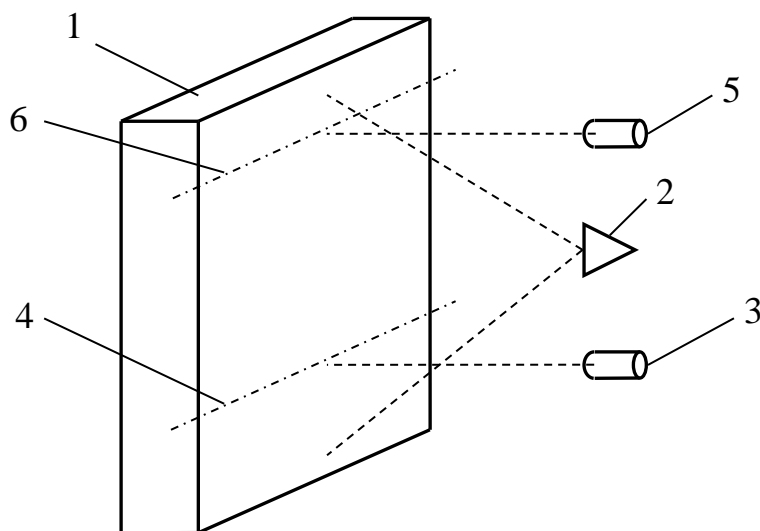


Рисунок 5.12 – Експериментальна схема моніторингу просторового розподілу температури коксового пирога

Експериментальне дослідження здійснювалося для п'яти коксових печей. У додатку А неведена таблиця відповідності температурних характеристик, що отримуються за двома незалежними інформаційними каналами, які формують образи  $P_1'$  і  $P_2$ . Для зменшення обсягу даних, в таблиці наведені ознаки образу  $P_2$  тільки ті, що відповідають лінійному розподілу, представленою образом  $P_1'$ . Також, для оцінки відтворюваності результатів, для кожної печі проводилося по 5 циклів вимірювання (у таблиці також наведений один варіант). Інші цикли експериментів показали добру відтворюваність результатів вимірювань.

Для наочності, в додатку Б, на рисунках Б.1-Б.5 наведено графічне представлення характеристик розподілу температури, що отримуються за двома незалежними каналами формування образів  $P_1'$  і  $P_2$ . Аналіз температурних характеристик, зображених на рисунках Б.1-Б.5, вказують значний розбіг відповідних значень температури, що отримуються за двома



$X(t)_k$  – репрезентативні ознаки еталонного образу  $P_2$ , що характеризує лінійний розподіл, який отримується з контрольного пірометричного датчика.

Для визначення ефекту використання гібридного розпізнавання при аналізі температурного режиму коксової печі, у додатку В наведений розбіг температурних параметрів між просторовим та еталонним лінійним розподілом температури за кожною з п'яти дослідних коксових печей.

Як витікає з додатку В, система технічного зору має тенденцію до заниження температурних характеристик, що пов'язано з додатковими перешкодами під час контролю, такими як дим, від продуктів горіння, а також неструктурованості та непостійності освітлення сцени спостереження. Для підвищення точності функціонування системи технічного зору може здійснюватися калібрування, однак, з погляду на нестаціонарність умов спостереження на кожному технологічному циклі, така задача стає край складною і не є раціональною в умовах промислового виробництва.

В результаті класифікації ознак образів  $P_1'$  і  $P_2$  була отримана корекція просторового розподілу температури коксової печі  $P_1''$ . У додатку Г наведена таблиця корегованих характеристик  $P_1''$  та відповідних характеристик репрезентативного образу  $P_2$ , а в додатку Д – графіки корегованих температурних розподілів. Порівняння результатів, наведених у додатках Б и Д вказує на те, що розбіг температур значно скоротився.

Враховуючи той факт, що відповідні характеристики еталонного і просторового розподілів температури на кожній з 28 вертикалей, в ідеальному випадку, повинні збігатися, то точність проведеної корекції може бути визначена через перевірку відповідності кожної характеристики апріорно визначеному класу  $C$  (5.8). При цьому, у випадку ідеального збігу, ознаки образу  $P_1''$ , що характеризує просторовий розподіл температури, будуть співпадати з ознаками репрезентативного образу  $P_2$ .

З погляду на те, що температурні характеристики представляють собою детерміновані ознаки, що мають однакову міру, для визначення ступеня подібності образу  $P_1''$  передвизначеному класу  $C$  використана метрика, на основі евклідової відстані. В цьому випадку, міра відстані образу  $P_1''$  може бути визначена наступним чином

$$d(P_1'', C) = \sum_{i=1}^k (T(X_{m,n}^R)_k - T(X(t)_k)) \quad (5.9)$$

де  $T(X_{m,n}^R)_k$  – репрезентативна характеристика температури коксового пирога, яка визначається за корегованим образом  $P_1''$ , та відповідає аналогічній характеристиці для  $k$ -ї вертикалі, яка визначається за еталонним образом  $P_2$ ;

$T(X(t)_k)$  – характеристика температури коксового пирога для  $k$ -ї вертикалі, яка визначається за еталонним образом  $P_2$ .

Для даних про температурні характеристики, які наведені у додатку Г, відповідно до (5.9), для кожної з п'яти експериментальних печей, міра відстані складатиме  $d_1(P_1'', C) = 80,5$ ;  $d_2(P_1'', C) = 60,2$ ;  $d_3(P_1'', C) = 76,5$ ;  $d_4(P_1'', C) = 62,3$ ;  $d_5(P_1'', C) = 52,6$ .

Середньоквадратичне відхилення міри відстані для кожної печі визначається наступним чином:

$$\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^n (d_i(P_1'', C) - \bar{d}(P_1'', C))^2} \quad (5.10)$$

де  $n$  – кількість коксових печей, за якими здійснюється експериментальне дослідження;

$d_i(P_1'', C)$  – значення евклідової метрики для  $i$ -ї коксової печі;

$\bar{d}(P_1'', C)$  – середнє значення евклідової метрики для усіх коксових печей,

за якими здійснюється експериментальне дослідження.

Для розглянутого випадку  $\sigma = 23,4$ , що знаходиться на рівні технологічних вимог, викладених у [218]. Однак, такий традиційний підхід до визначення достовірності розпізнавання повністю не відражає сутності проблематики поставленої задачі, згідно якої точність визначення розподілу температури повинна забезпечуватися для кожної окремо узятю точки коксового пирога. В цьому випадку, розглянута характеристика (5.10) дає тільки загальне уявлення, однак на вказує на точність отримання розподілу температури за окремими ділянками коксового пирога.

З метою визначення достовірності отримання просторового розподілу у кожній точці використана статистична достовірність розпізнавання, за умови розгляду результату на кожній з вимірювальних вертикалей коксового пирога.

Відомо [203], що статистична достовірність визначається за виразом:

$$D = \frac{N_{cor}}{N_{com}} 100\%, \quad (5.11)$$

де  $N_{cor}$  – кількість подій коректного розпізнавання образів;

$N_{com}$  – загальна кількість подій розпізнавання.

З урахуванням того, що збіг просторового розподілу з еталонним лінійним повинен спостерігатися для кожної вертикалі, то достовірність результату розпізнавання визначається для кожної групи еквівалентних характеристик. Таким чином, у випадку аналізу та співставлення температурних характеристик 28 вертикалей на визначеній горизонтальній

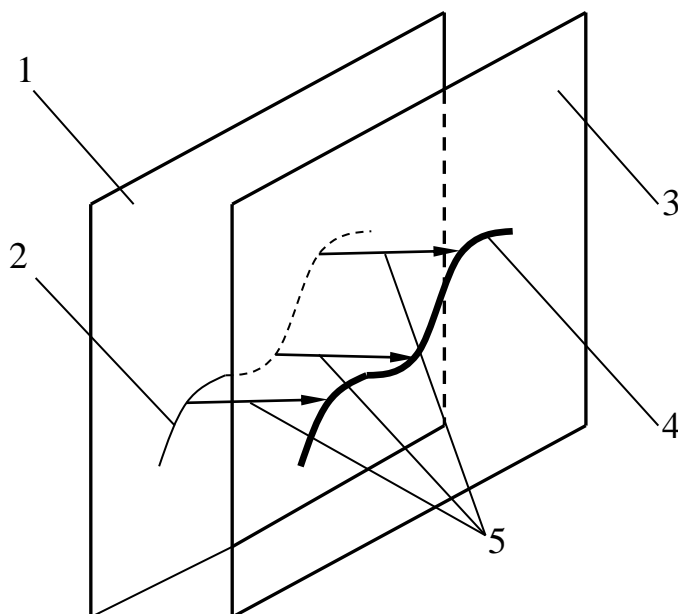
осі коксового пирога, загальна кількість подій розпізнавання буде дорівнювати 140 (по 28 для кожної з п'яти дослідних коксових печей).

З урахуванням передвизначеного класу  $C$  (5.8) для кожної точки вертикалі визначеної горизонтальної осі, здійснювалася перевірка відповідності класу  $C$ . Результати наведені у додатку Д. В таблиці додатку Д символами «1» позначені події відповідності образу класу  $C$ , а «0», відповідно, події невідповідності. Так, статистична достовірність розпізнавання образів для п'яти експериментальних печей отримані, відповідно:  $D_1= 82,1\%$ ;  $D_2= 96,4\%$ ;  $D_3= 85,7\%$ ;  $D_4= 92,9\%$ ;  $D_5= 100,0\%$ . Середня достовірність за результатами усіх експериментів (для усіх п'яти печей) складає 91.4%, що відповідно до вимог, викладених вище, задовольняє існуючим нормам [219].

Використання представленої ІТ аналізу просторового розподілу температури коксового пирога, окрім можливості більш точної стабілізації технологічних параметрів, дозволяє здійснювати непрямий аналіз стану коксової печі. Відповідний метод запропонований у [48], який полягає в тому, що у випадку виникнення дефектів у кладці коксової печі спостерігається додаткове потрапляння газу у робочу зону печі та виникнення локальної зони перегріву коксового пирога в області виникнення дефекту. При цьому, від масштабу дефекту залежить кількість газу, що додатково потрапляє до робочої зони та ступінь локального перегріву. Крім цього слід зазначити, що геометрична конфігурація зони локального перегріву коксового пирога буде відповідати геометричній комбінації дефекту. Цей процес проілюстрований на рисунку 5.13.

Згідно рисунку 5.13, потрапляння газу через порушення герметичності простінків коксової печі створює додатковий тепловий потік, що спричиняє додаткове прогрівання зон коксового пирога, що знаходяться у безпосередньої близькості від дефекту. Враховуючи малі зазори між простінками коксової печі та коксового пирога, зона локального перегріву

буде мати положення, форму та розміри наближені до конфігурації дефекту. Іншими словами, через утворення додаткового теплового потоку дефект нібито «відбивається» на поверхні коксового пирога.



- 1 – простінок коксової печі; 2 – дефект простінку коксової печі;  
 3 – поверхня коксового пирога; 4 – локальний перегрів поверхні;  
 5 – тепловий потік внаслідок виникнення дефекту печі;

Рисунок 5.13 – Процес формування зони локального перегріву коксового пирога внаслідок виникнення дефекту печі

Крім цього спостереження за конфігурацією та температурною характеристикою зон локального перегріву на різних технологічних циклах дозволяє здійснювати моніторинг динаміки розвитку дефекту [75, 220]. Так, якщо від одного технологічного циклу до іншого спостерігається збільшення температури та розмірів відповідних зон локального перегріву – це дозволяє стверджувати про розвиток дефекту. Швидкість змін характеристик зон локального перегріву буде відповідати швидкості розвитку дефекту.

Слід зазначити, що у [221] наголошується на тому, що перегрів окремих областей коксового пирога може бути наслідком неоднорідності

складу шихти та якості її трамбування у коксовому пирозі. Однак цей факт не впливає на ефективність запропонованого метода контролю стану коксової печі. Так, якщо конфігурація і положення зон локального перегріву коксового пирога на різних технологічних циклах носить стохастичний характер, то цей факт дозволяє стверджувати, що такі зони є наслідком неоднорідності шихти. Якщо ж на різних технологічних циклах спостерігається певна закономірність у положенні і конфігурації зон локального перегріву, то це однозначно вказує на наявність дефекту в коксовій печі.

Представлена IT аналізу просторового розподілу температури дозволяє наочно зафіксувати зони локального недогріву та перегріву коксового пирога, їхні розміри та конфігурацію. Побудова кольорової термограми здійснюється шляхом зазначення шкали відповідності значень температури кольоровим атрибутам. Формування такої шкали здійснюється за лінійним пропорційним законом. Приклад побудови термограми, проілюстрований на рисунку 5.14.

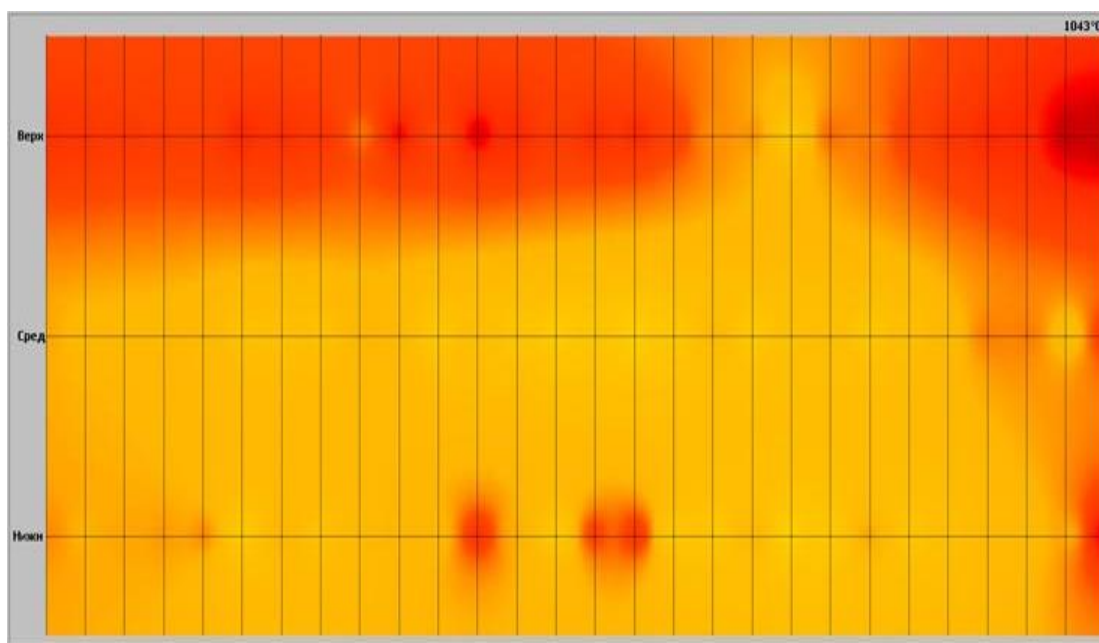


Рисунок 5.14 – Візуалізація отриманого просторового розподілу температури коксового пирога



Визначення статичних зон перегріву, що характеризують наявний дефект коксової печі може бути ідентифікований через порогову фільтрацію термограми, в результаті якої видаляються усі зони, температура в яких не перевищує заданого значення. Области просторового розподілу температури, які згідно рисунку 5.12 знаходяться безпосередньо напроти дефекту коксової печі, можуть бути виділені згідно умови

$$\begin{aligned} \exists T(X_{m,n}) \in P_1'', T(X_{m,n}) \geq T_G \rightarrow A(X_{m,n}) = \gamma', \\ \exists T(X_{m,n}) \in P_1'', T(X_{m,n}) < T_G \rightarrow A(X_{m,n}) = \gamma'', \end{aligned} \quad (5.12)$$

де  $T_G$  – порогове значення температурної характеристики, що вказує на дефект коксової печі;

$A(X_{m,n})$  – атрибут кольору точки термограми, що відповідає характеристиці просторового розподілу температури  $X_{m,n} \in P_1''$ ;

$\gamma'$  і  $\gamma''$  – значення атрибутів областей зображення, які символізують наявність та відсутність дефекту, відповідно.

Для здійснення порогової фільтрації зображення термограми, був розроблений додатковий програмний модуль, який інтегрується до складу інформаційної системи управління тепловим режимом коксовий печей. Приклад порогової фільтрації та виділення дефектної області наведений на рисунку 5.15.

На рисунку 5.15 зліва наведене зображення термограми, а справа – виділені області з підвищеною температурою.

У випадку, коли конфігурація виділених областей підвищеної температури зберігається від одного технологічного циклу до іншого, система визначає цю подію як наявність дефекту коксової печі [183].

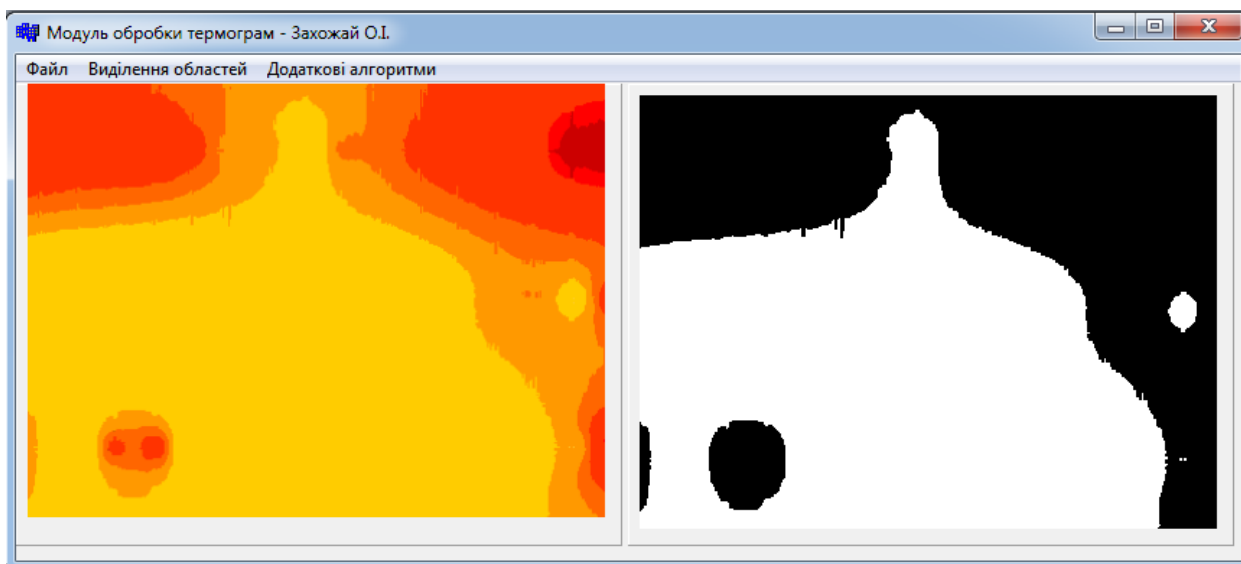


Рисунок 5.15 – Приклад роботи модулю обробки термограм

Представлена ІТ аналізу просторового розподілу температури коксового пирога була реалізована в якості попереднього модулю обробки даних, що надходять до АСУТП «КОКС» (розробник ТОВ «Техніка МК» [222]), яка є найбільш поширеною системою, що сертифікована для використання на коксохімічних виробництвах ряду держав, у тому числі України. Акт впровадження представлений у додатку Ж.

## **5.2 Інформаційна технологія ультразвукового вимірювання лінійних відстаней для систем автоматички**

Сучасні ультразвукові датчики, що використовуються у складі систем вимірювання, поділяються на два різновиди: з суміщеними і роздільними приймачем і передавачем [223, 224].

Приклад ультразвукового датчику з роздільними приймачем і передавачем, представлений на рисунку 5.16 (а), а зі суміщеними передавачем і приймачем – на рисунку 5.16 (б).



Рисунок 5.16 – Ультразвукові датчики з роздільними (а) та суміщеними (б) приймачем та передавачем

Вимірювальна схема з суміщеними приймачем і передавачем, більш компактна і зручна для експлуатації в системах автоматики, але має істотний недолік, який полягає в тому, що після випромінювання сигналу має пройти деякий час, перш ніж закіняться коливання мембрани випромінювача і стане можливим її робота на прийом зворотного сигналу. Цей інтервал часу називається «мертвим» часом датчика. Процес передачі та прийняття ультразвукового сигналу проілюстрований на рисунку 5.17.

Сигнал передачі можна умовно поділити на два етапи, тривалість яких позначена як  $T_{TR}$  – час формування сигналу передачі і  $T_D$  – «мертвий» час датчика. Інтервал часу між сигналом передачі та сигналом прийому, позначений як  $T_L$ , характеризує вимірювальну відстань до перепони, що викликала відбиття сигналу передачі. При цьому, відстань до об'єкта може бути визначена як

$$L = \frac{C \cdot T_L}{2}, \quad (5.13)$$

де  $C$  – швидкість поширення ультразвукової хвилі у просторі.

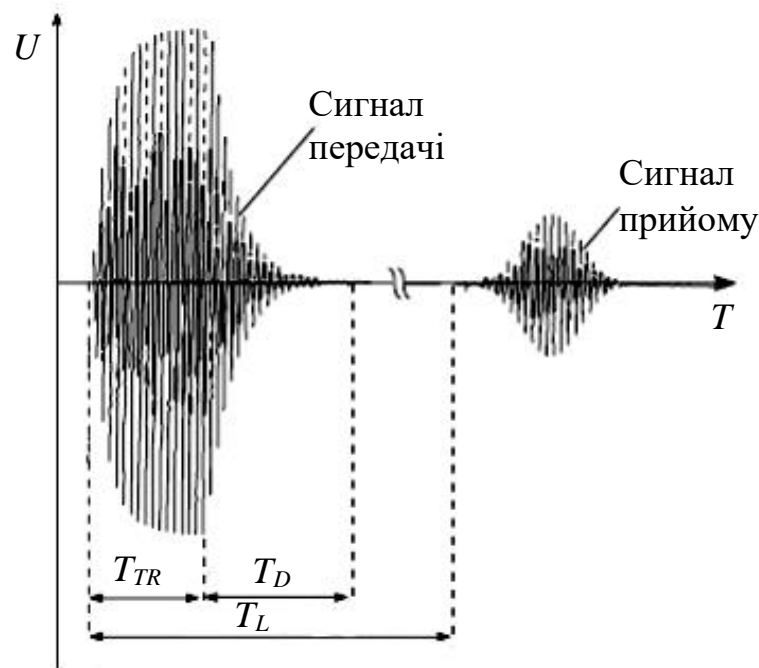


Рисунок 5.17 – Сигнали передачі та прийому ультразвукового датчика

При цьому, можливість вимірювання мінімальної відстані, обмежена необхідністю виконання умови

$$T_L > T_{TR} + T_D. \quad (5.14)$$

Виходячи з (5.13) і (5.14), мінімальна відстань, що може бути визначена датчиком з характеристикою  $T_D$  на основі управляючого сигналу довжиною  $T_{TR}$ , дорівнює

$$L_{\min} = \frac{C \cdot (T_{TR} + T_D)}{2}. \quad (5.15)$$

Виходячи з (5.15), основним недоліком систем з суміщеними приймачем і передавачем є наявність мінімальних значень відстані, що не можуть бути реєстровані датчиком. При чому, таке обмеження визначається конструктивними особливостями системи вимірювання та характеристиками конкретно обраного для цієї задачі ультразвукового датчика.

Основним недоліком систем вимірювання з роздільними приймачем і передавачем є складність реалізації, а також завищення результатів вимірювання на основі того, що для датчиків такої конструкції неможливе встановлення передавача та приймача на одній осі. Схема такої системи вимірювання проілюстрована на рисунку 5.18.

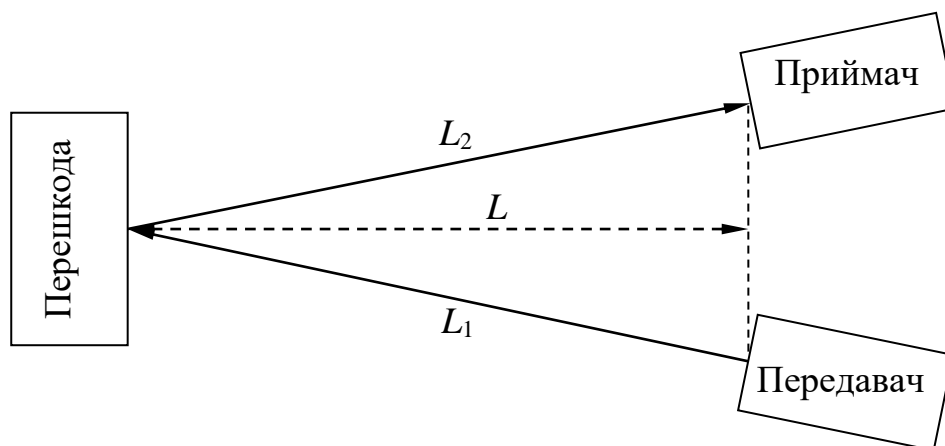


Рисунок 5.18 – Схема вимірювання з роздільними передавачем та приймачем

Виходячи з наведеної схеми, системою вимірювання буде отримане значення  $T_L$ , що відповідає часу проходження шляху  $L_1+L_2$ . Однак, дійсна дистанція до перешкоди представляє собою  $L$ . При чому, очевидно, що

$$2L \neq L_1 + L_2. \quad (5.16)$$

Відповідно до (5.16) створюється помилка визначення відстані, яка залежить від геометричних аспектів розташування приймача та передавача. Відповідно до схеми на рисунку 5.14, розбіжність у вимірюванні може бути легко розрахована і внесена у якості поправки до інформаційно-вимірювальної системи. Однак, таке рішення легко реалізується у випадку, коли взаємне розташування приймача та передавача залишається незмінним, а розташування перешкоди ортогональне до осі визначення дійсної відстані  $L$ . У випадку динамічної зміни взаємного положення елементів системи

вимірювання величина поправки не буде постійною, що значно ускладнює отримання точного результату вимірювання.

Максимально припустима відстань вимірювання обмежена кінцевою потужністю передавача, чутливістю приймача та характеристиками опору середовища прямому поширенню ультразвукової хвилі та відповідним коефіцієнтом загасання сигналу.

Таким чином, на основі вищезгаданого можна зробити висновок, що існуючі підходи до побудови схем вимірювання та інформаційні системи визначення достовірного значення відстані потребують вдосконалення у напрямку підвищення точності та розширення діапазону вимірювань.

З метою рішення цієї задачі пропонується використання ІТ ультразвукового вимірювання лінійних відстаней [46, 72] на основі гібридного розпізнавання образів. ІТ базується на використанні  $k$  ультразвукових датчиків з суміщеними передавачем та приймачем. В результаті чого, в системі формується  $k$  інформаційних каналів отримання результатів вимірювання відстані, які підлягають подальшій обробці з метою отримання у будь який момент часу найбільш достовірного результату вимірювання.

Виходячи з того, що характеристики ультразвукових датчиків фактично визначаються їхньою конструкцією, то для розширення діапазону вимірювань необхідне використання сукупності  $k$  датчиків, в яких діапазони вимірювань доповнюють один одного. Принцип суміщення діапазонів вимірювання датчиків проілюстрований на рисунку 5.19. Згідно рисунку, кожен датчик характеризується діапазоном відстані, в якому може бути отриманий достовірний результат вимірювання. Ці діапазони схематично представлені 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 спільної обробки результатів вимірювань, мінімальна відстань реєстрації буде співпадати з мінімальною

характеристикою «мертвої» зони одного з  $k$  датчиків (на рисунку 5.19 це  $\Delta_1$ ). При цьому, максимальне припустиме значення достовірного вимірювання відстані буде співпадати з максимальним значенням відстані для одного з датчиків (на рисунку 5.19 це  $\Delta L_k$ ). Таким чином, загальний діапазон достовірного вимірювання відстані системою ультразвукових датчиків буде визначатися як  $\Delta L$ , що значно ширший ніж будь який діапазон вимірювання кожного, окремо узятго датчика [46, 72].

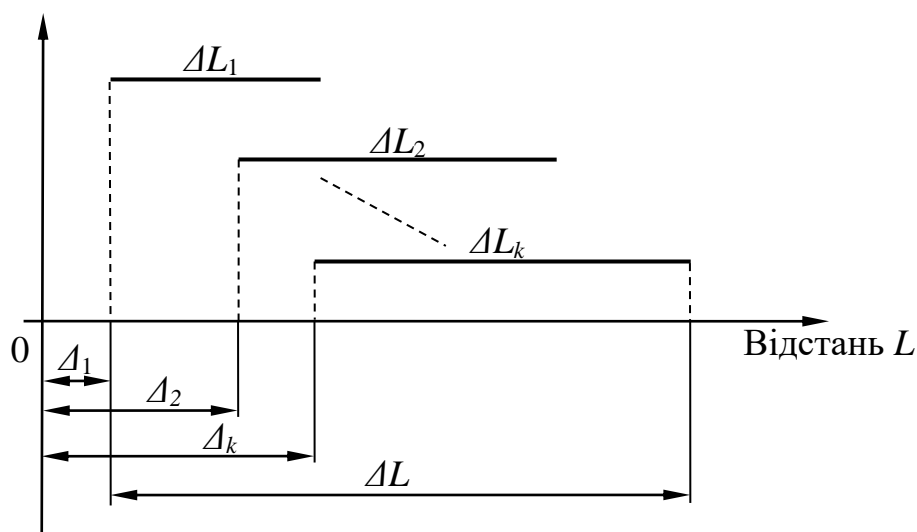


Рисунок 5.19 – Отримання загального діапазону вимірювання  $\Delta L$  з використанням системи ультразвукових датчиків

Однак, використання сукупності датчиків створює додаткові складнощі визначення достовірного результату вимірювання. Це пов'язано з двома факторами:

- нерівномірність характеристики точності вимірювання у всьому діапазоні для кожного датчику;
- наявність зон перекриття діапазонів вимірювання окремих датчиків, що створює неоднозначність (багатоальтернативність) у визначенні остаточного результату.

Використання гібридного розпізнавання образів дозволяє виключити неоднозначність результату вимірювання. На рисунку 5.20 показана структурна реалізація системи ультразвукового вимірювання лінійних відстаней на базі гібридного розпізнавання [46].

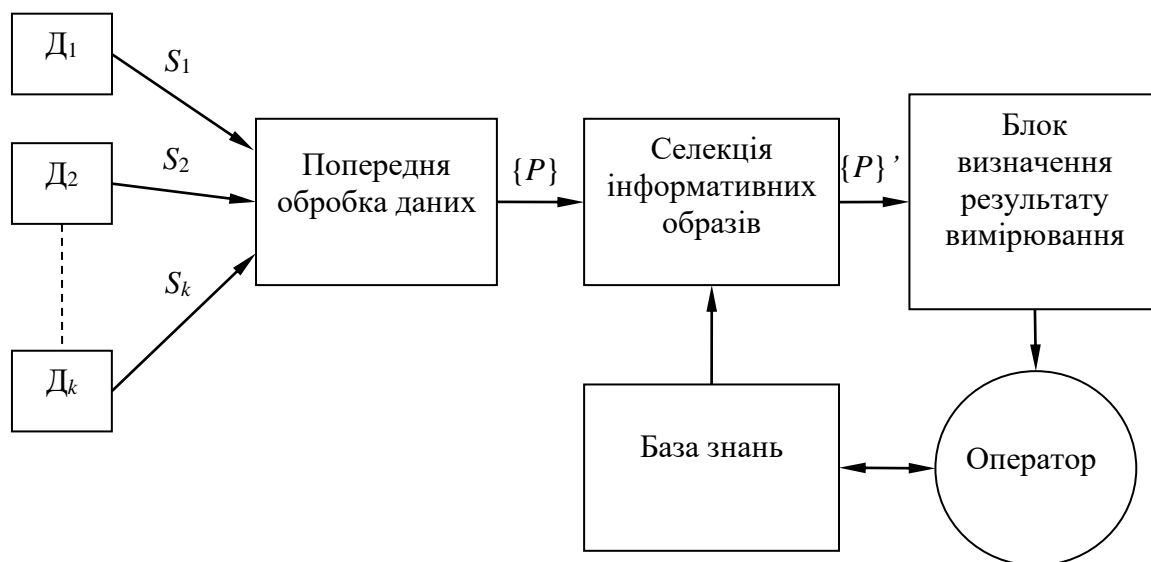


Рисунок 5.20 – Структурна реалізація системи ультразвукового вимірювання лінійних відстаней на базі гібридного розпізнавання

Відповідно до наведеної структурної реалізації данні виміру поточної відстані до об'єкта надходять по  $k$  інформаційним каналам ( $S_1-S_k$ ) від датчиків  $D_1-D_k$  та піддаються попередньої обробці. На цьому етапі здійснюється нормування даних та формування на їх основі  $k$  образів, що характеризують кожен канал вимірювання. Надалі, сформована множина образів  $\{P\}$  підлягає селекції та на її основі формується множина інформативних образів  $\{P\}'$ . Сукупність  $\{P\}'$  містить образи, що відповідають тим каналам виміру, що за наявних умов вимірювання забезпечують максимальну достовірність результату. При чому

$$\{P\}' \subset \{P\}. \quad (5.17)$$



Селекція множини  $\{P\}$  на дві групи – інформативні та неінформативні здійснюється за методикою, що викладена у розділі 3 та [203], на основі наступної апріорної інформації бази знань:

- дані про типи ультразвукових датчиків;
- розміри «сліпої» зони датчика  $\{L_k\}$ ;
- діапазони вимірювання датчиків  $\{\Delta L_k\}$ ;
- дані про розподіл точності вимірювання у всьому діапазоні датчика  $\{p_k(r)\}$ .

Можливість додання до бази знань характеристик датчиків, що використовуються у системі, надається оператору на основі його експертних знань та евристичних здібностей. В результаті чого, інформаційну модель системи розпізнавання представляється у вигляді

$$\begin{pmatrix} S_1 \\ S_2 \\ \dots \\ S_k \end{pmatrix} \Rightarrow \begin{pmatrix} P_1 \\ P_2 \\ \dots \\ P_k \end{pmatrix} = \begin{cases} R_1, \{Ch_1|G\}, \\ R_2, \{Ch_2|G\}, \\ \dots \\ R_k, \{Ch_k|G\}, \end{cases} \quad (5.18)$$

де  $R_1-R_k$  – сукупність апостеріорних даних про вимір відстані кожним з  $k$  датчиків;

$\{Ch_1|G\} - \{Ch_k|G\}$  – сукупність апріорних характеристик  $k$  датчиків за умови наявності критеріальної характеристики  $G$ , за якої може бути здійснена селекція достовірного (або достовірних) результатів вимірювання.

Відповідно до моделі (5.18) 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  $R_1 = R_2 = \dots = R_k$ . Однак, у випадку реальних умов, на

кожній відстані отримуються результати вимірювання з різною достовірністю, тому  $R_1 \approx R_2 \approx \dots \approx R_k$ . При чому ступінь наближення до достовірного результату для кожного датчика на різній відстані буде різним. Таким чином, для визначення достовірного результату в усьому загальному діапазоні системи вимірювання  $\Delta L$  використовуються характеристики розподілу точності вимірювання від відстані. Така характеристика визначається апріорно для кожного датчика, виходячи з його типу, та на основі його технічної документації. Проведені дослідження широкого спектру ультразвукових датчиків показав що розподіл точності вимірювання в усьому діапазоні, за незначним припущенням, можна представити як нормальний розподіл. Якщо ймовірнісні характеристики отримання достовірного результату для кожного датчика представити в одній системі, то характеристика розподілу ймовірності отримання достовірного результату вимірювання відстані для всієї системи з  $k$  ультразвукових датчиків буде представлена у вигляді, наведеному на рисунку 5.21.

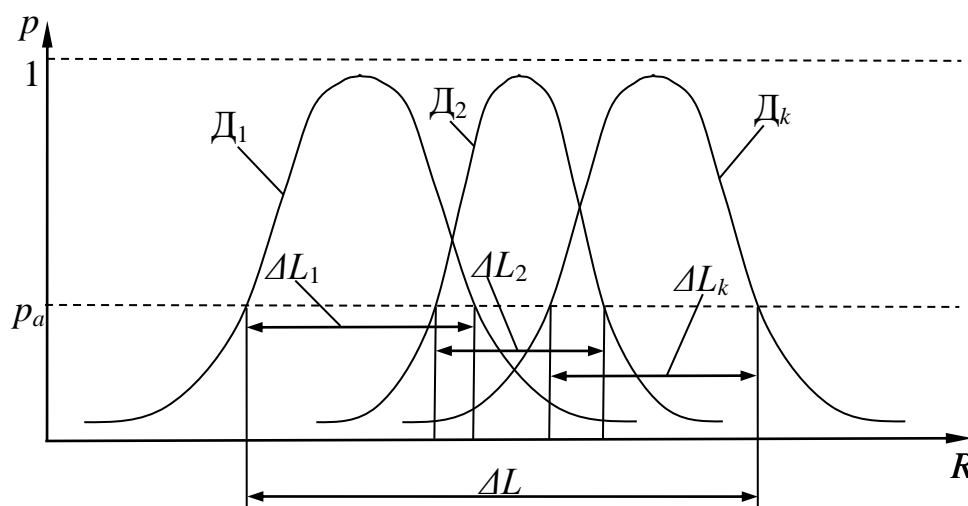


Рисунок 5.21 – Розподіл ймовірностей отримання достовірного результату вимірювань системою ультразвукових датчиків

За реальних умов реалізації ІТ, задачею експерта є визначення мінімально припустимого рівня достовірності результату вимірювання, що на рисунку 5.20 показаний як  $p_a$ . Відповідно до цього, стає можливим визначення припустимих діапазонів вимірювання для кожного датчика ( $\Delta L_1, \Delta L_2, \dots, \Delta L_k$ ). Максимальна точність для кожного датчика прагне одиниці. Таким чином, припустимі характеристики точності можуть бути отримані тільки у діапазоні  $[p_a, 1)$ .

В ІТ ультразвукового вимірювання, отримання найбільш достовірного результату (той що належить діапазону  $[p_a, 1)$ , та має найбільше можливе значення в цьому діапазоні) можливий завдяки апріорній побудові уздовж усього діапазону вимірювання лінії отримання достовірних результатів вимірювання. Ця лінія фактично визначає критеріальну характеристику  $G$  за якою здійснюється селекція достовірного значення вимірюної відстані. Визначення цієї лінії можливо через порівняння значень функцій розподілу  $p_k(R)$  для кожного датчика. Якщо достовірний результат вимірювань позначити як  $R$ , то критерій селекції цього значення з сукупності буде визначатися як

$$\forall R_i, R_i \in \Delta L \Leftrightarrow \exists R_i = R, p_i(R_i) = \max p_i(R). \quad (5.19)$$

Визначення лінії отримання достовірних результатів вимірювань проілюстровано на рисунку 5.22. Вона складається з відрізків  $\Delta L_1', \Delta L_2', \dots, \Delta L_k'$ , які є частинами визначених раніше діапазонів датчиків  $\Delta L_1, \Delta L_2, \dots, \Delta L_k$ . Відповідно до (5.19), серед отриманих вимірів  $R_1-R_k$  датчиків  $D_1-D_k$  найбільш достовірним значенням  $R$  буде вважатися вимір  $R_i$  датчика  $D_i$ , для якого забезпечується максимум функції розподілу.

Однак, в цьому випадку, побудова безперервної лінії отримання достовірних результатів вимірювань можлива тільки за умови перекриття локальними діапазонами вимірювань датчиків усього необхідного діапазону

$\Delta L$ . Так, згідно прикладу, представленому на рисунку 5.22, лінія отримання достовірних результатів буде проходити від точки  $A$  до точки  $B$ , далі від точки  $B$  до  $C$ , та від  $E$  до  $F$ . Таким чином, у діапазоні  $AB$  достовірний результат буде відповідати датчику  $D_1$ , у діапазоні  $BC$  – датчику  $D_2$ , а для діапазону  $EF$  – датчику  $D_k$ .

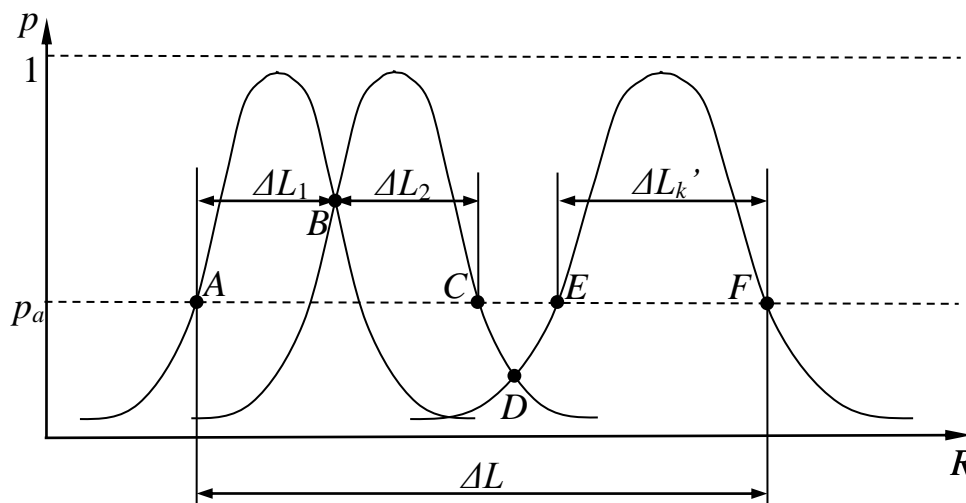


Рисунок 5.22 – Визначення лінії отримання достовірних результатів вимірювань у системі ультразвукових датчиків

Слід зазначити, що під час синтезу ІТ ультразвукового вимірювання може скластися ситуація, коли не буде безперервною лінія  $AF$  в усьому діапазоні  $\Delta L$ . Такий випадок також проілюстрований на рисунку 5.22. На рисунку видно, що на діапазоні  $CE$  неможливе отримання достовірного результату, так як область в околиці точці  $D$  лежить нижче мінімально припустимого рівня достовірності результату вимірювання. Цей факт дозволяє під час синтезу ІТ здійснити автоматичне визначення кількості необхідних датчиків з відповідним діапазоном.

У випадку коректного синтезу ІТ ультразвукового вимірювання необхідне обов'язкове виконання умови

$$p(A) \in [p_a, 1) \wedge p(B) \in [p_a, 1) \wedge p(C) \in [p_a, 1) \wedge \dots \wedge p(D) \in [p_a, 1) \wedge \dots \wedge p(E) \in [p_a, 1) \wedge p(F) \in [p_a, 1). \quad (5.20)$$

Якщо умова (5.20) не виконується, то обрана кількість датчиків не дозволяє з заданим рівнем  $p_a$  перекрити увесь діапазон  $\Delta L$ . Найбільш раціональним є введення до системи датчику, для якого максимально достовірний результат вимірювань співпадає з точкою мінімуму діапазону розриву лінії отримання достовірних результатів вимірювань  $R_{ecs}$ . Критерій визначення такої точки  $R_{ecs}$  буде мати вигляд

$$\forall R_i, p_i(R_i) \notin [p_a, 1) \Leftrightarrow \exists R_i = R_{ecs}, p_i(R_i) = \min p_i(R). \quad (5.21)$$

На рисунку 5.23 проілюстрований процес усунення розриву лінії отримання достовірних результатів вимірювань за рахунок введення  $k+1$ -го датчику з максимумом, що співпадає з точкою  $R_{ecs}$ .

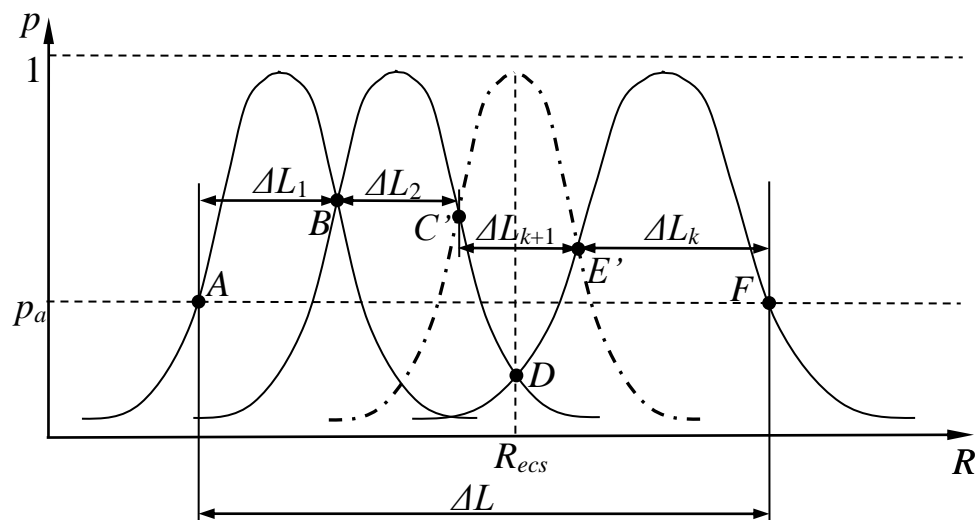


Рисунок 5.23 – Усунення розриву лінії отримання достовірних результатів вимірювань через введення додаткового датчику

Введення до системи  $k+1$ -го датчику призвело до зміни граничних точок  $C$  і  $E$  (згідно рисунку 5.23) на точки  $C'$  і  $E'$  відповідно. Таким чином, скорегована лінія отримання достовірних результатів вимірювань буде проходити через послідовність точок  $A, B, C', E', F$  та повністю відповідати умові безперервності у всьому діапазоні  $\Delta L$ , що, в свою чергу буде складатися з діапазонів вимірювань датчиків:  $\Delta L_1, \Delta L_2', \dots, \Delta L_{k+1}, \dots, \Delta L_k'$ .

На рисунку 5.24 представлена структурно-логічна схема послідовності виконання операцій в прикладній інформаційній технології гібридного розпізнавання образів для ультразвукового вимірювання лінійних відстаней.

Згідно цієї схеми, вхідні дані про лінійну відстань до об'єкту у просторі сприймаються системою ультразвукових датчиків, діапазоні вимірювань яких співвідносяться згідно рисунку 5.19. Ці дані реєструються  $k$  датчиками, що забезпечує формування сукупності з  $k$  образів, які містять ознаки вимірювальної відстані.

Так як динамічні характеристики датчиків відрізняються, це призводить до того, що, навіть якщо команда, що ініціює випромінювання ультразвукового сигналу надходить одночасно на усі датчики, сигнали будуть відіслані з різними часовими затримками. Приймання зворотного, відбитого сигналу також здійснюється з різними часовими затримками, так як накопичується похибка з моменту відправлення сигналу. Таким чином, для часової синхронізації процесів обробки даних, після формування образів, здійснюється етап прив'язки результатів вимірювання датчиків до певних часових відліків:

$$\forall l(t), t_{i-1} \leq t \leq t_i \rightarrow l(t) = l(t_i). \quad (5.22)$$

де  $l(t)$  – поточне значення, отримане в момент часу  $t$ , що підлягає часовій прив'язці;

$t_i$  і  $t_{i-1}$  – граничні значення часу для інтервалу прив'язки;

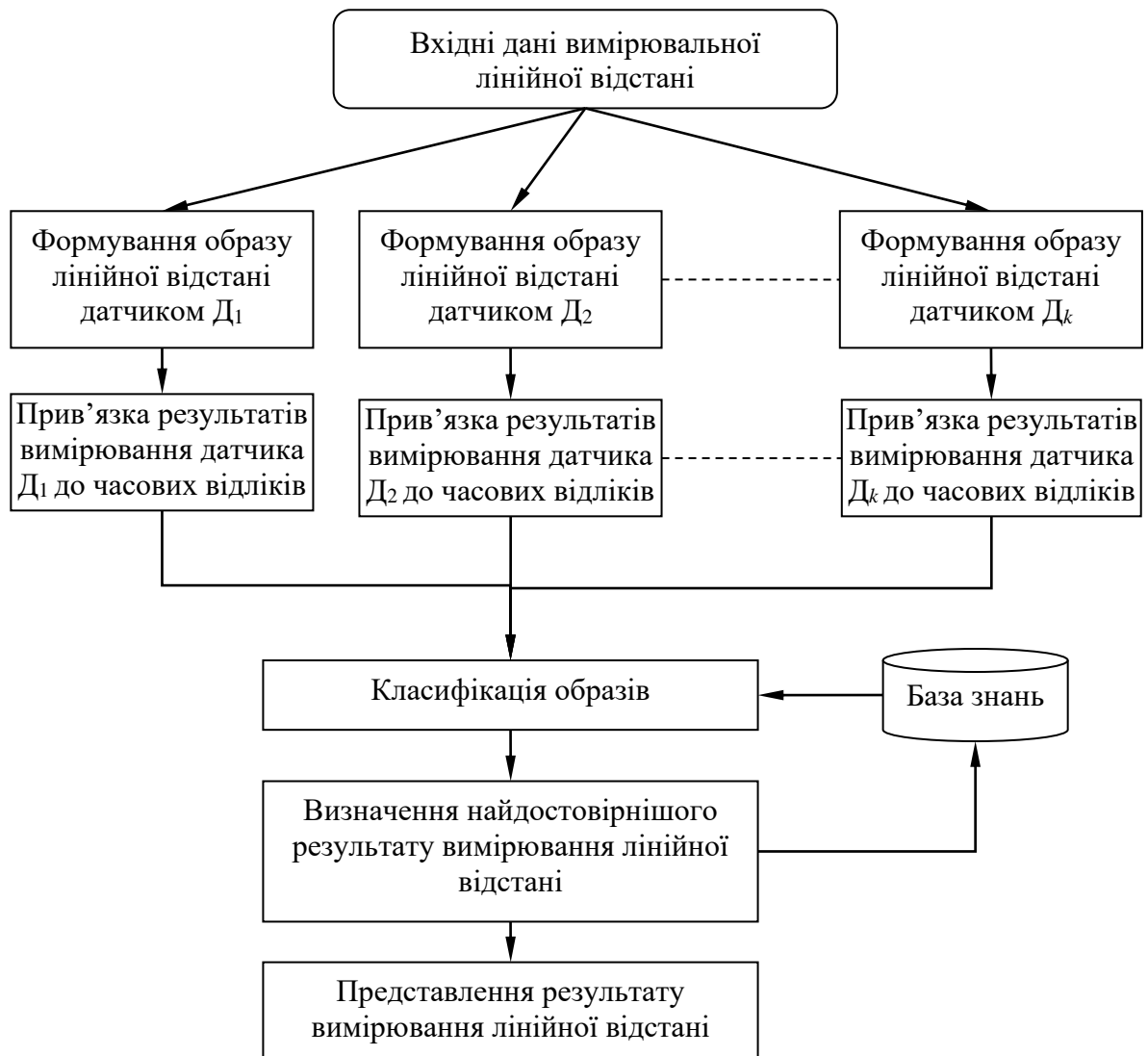


Рисунок 5.24 – Структурно-логічна схема послідовності операцій в прикладній інформаційній технології гібридного розпізнавання образів для ультразвукового вимірювання лінійних відстаней

Процес прив'язки проілюстрований на рисунку 5.25. Згідно рисунку, результати вимірювань  $l_1(t) \dots l_5(t)$  відповідно до (5.21) будуть прив'язані до наступних відліків:

- $l_1(t)$ ,  $l_2(t)$  і  $l_3(t)$  будуть вважатися результатами вимірювання  $i$ -го відліку  $(l_1(t_i), l_2(t_i), l_3(t_i))$ ;
- $l_4(t)$  і  $l_5(t)$  будуть вважатися результатами вимірювання  $i+1$ -го відліку  $(l_4(t_i), l_5(t_i))$ .

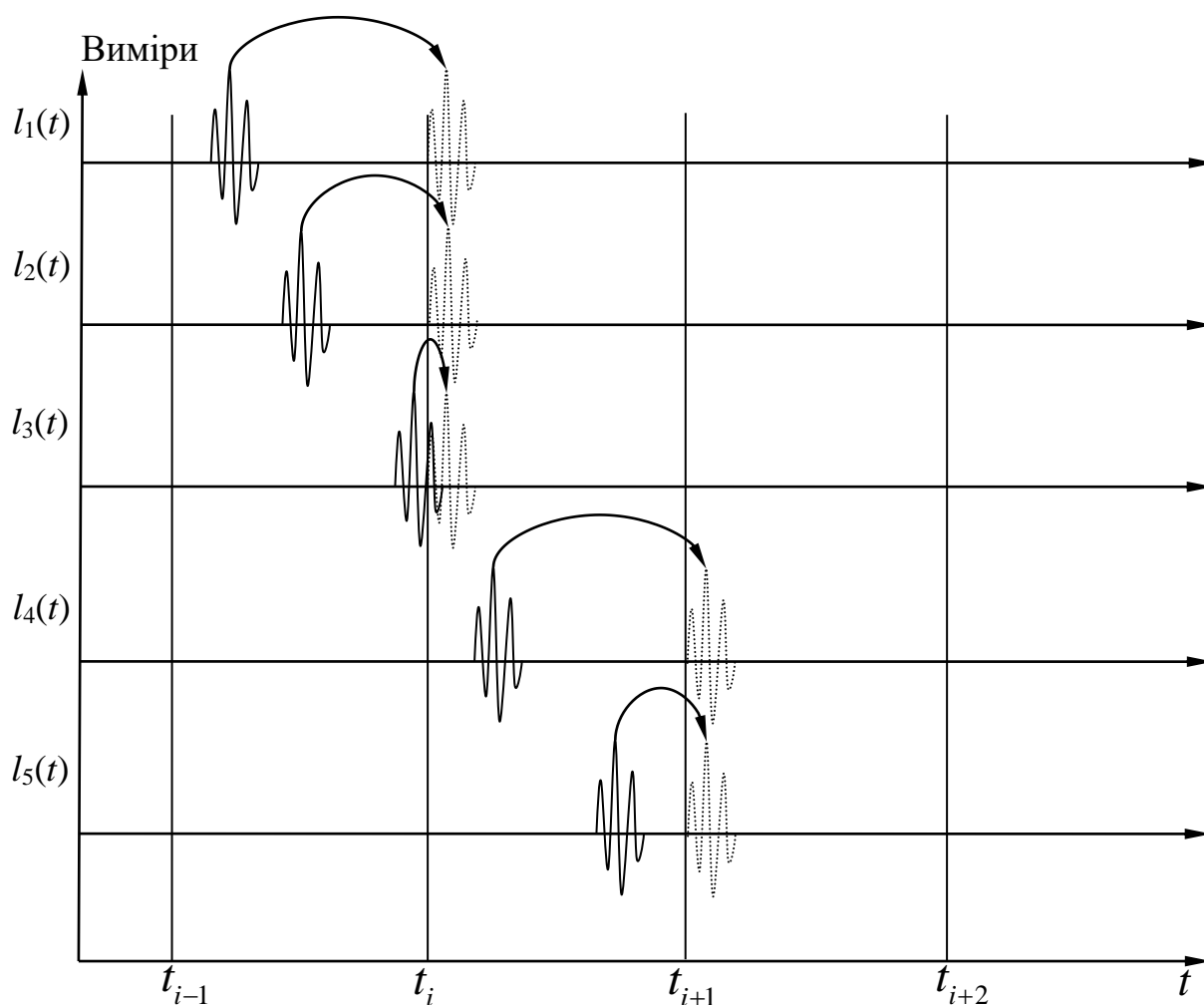


Рисунок 5.25 – Прив’язка результатів вимірювань сукупністю датчиків до послідовності часових відліків

Після прив’язки, здійснюється класифікація образів при якій співставляються результати одного відліку з характеристиками в базі даних. В результаті чого, визначається той результат, що був отриманий датчиком, який за поточних умов (тобто поточній відстані) має найвищу точність, відповідно до критерію (5.19). Крім цього, результати вимірювання датчиків порівнюються з результатами попередніх відліків на предмет значних розходжень. Якщо результат якогось датчика значно відрізняється від попереднього вимірювання або істотно різниться у порівнянні з результатами вимірювань іншими датчиками за тих самих умов і того ж відліку, то це



вказує на його некоректну роботу. При цьому, обслуговуючий персонал повинен приймати рішення щодо ремонту і заміни системи.

Надалі, отриманий результат вимірювання представляється для подальшого використання в системах контролю і управління.

Таким чином, використання ІТ гібридного розпізнавання образів для ультразвукового вимірювання лінійних відстаней дозволило забезпечити більш широкий діапазон вимірювання та визначення найдостовірнішого результату, автоматизоване визначення кількості необхідних датчиків в системі, у відповідності до їх характеристик, а також верифікацію адекватності отриманих результатів вимірювання за кожним датчиком та визначення моментів їх виходу з ладу.

ІТ ультразвукового вимірювання лінійних відстаней біла використана у складі засобів автоматизації, про що мається відповідний акт у додатку Ж.

### **5.3 Інформаційна технологія визначення ступеня унікальності тестових даних**

Нажаль, термін «плагіат» стає дедалі поширенішим у будь-яких сферах людської діяльності. Така ситуація спостерігається, зокрема, в області освіти, де під час підготовки різноманітних студентських робіт, доповідей, рефератів, курсових і дипломних робіт, а іноді й дисертацій, суб'єктами освітнього процесу робляться спроби запозичення результатів чужої інтелектуальної праці [225].

У зв'язку з цим, в сучасній науці й освіті стає дедалі поширенішим підхід до написання робіт, який полягає в звичайному копіюванні інформації з одного або декількох джерел, подальшої редакторської обробці та представлення отриманих матеріалів як результату особистої інтелектуальної праці [225, 226].

Одним зі шляхів протидії вказаним вище негативним аспектам є використання спеціальних засобів перевірки текстів на унікальність, які дозволяють за результатами аналізу представлених продуктів інтелектуальної праці, виявити запозичені матеріали, визначити їхній обсяг, а також ідентифікувати джерела запозичення [226-233]. Такі засоби отримали назву систем антиплагіату [226-228].

Принцип функціонування систем перевірки текстів на унікальність полягає у порівнянні представлених текстів з наявними у базі даних. В процесі роботи системи здійснюється попередній збір і фільтрація інформації про наявні продукти інтелектуальної праці [226, 228]. Після попередньої обробки, тексти індексуються і вносяться до бази даних. Надалі, ця база даних використовується для порівняння з представленими для аналізу текстами. Кожний документ, завантажений для перевірки, ставиться до черги для обробки. Після перевірки необхідного документу система формує звіт, в якому докладно представляються усі результати пошуку та рівень унікальності перевіреного матеріалу. Як правило, ті фрагменти, що не є унікальними, для наочності, виділяються в тексті [60].

Існує велике різноманіття програмних продуктів, які дозволяють перевірити текст на унікальність та ідентифікувати ознаки плагіату. Кожна з них може мати специфічні особливості, а також переваги та недоліки [226-233]. Аналіз таких систем дозволив також виявити основні функції, що притаманні більшості з них, а саме:

- визначення відсотку унікальності представленого тексту;
- пошук фрагментів тексту, що були запозичені, а також ідентифікація джерел інформації, що були використані в плагіаті;
- маркування запозичених фрагментів тексту кольором, для наочності, а також використання різного кольору для запозичених фрагментів з різних джерел;
- формування докладних звітів після перевірки тексту на унікальність.

Програмне забезпечення, як правило має можливість гнучкої настройки алгоритмів та інструментарію пошуку подібності (наприклад: визначення кількості слів у шинглі, кількість вибірок з тексту та ін.).

До додаткових функцій програмного забезпечення перевірки текстів на унікальність можна віднести [226]:

- перевірку контенту різноманітних web-ресурсів на унікальність з формуванням докладного звіту, що містить зафіксовані джерела плагіату;
- робота зі списком проксі;
- здійснення пакетної перевірки представлених текстових файлів;
- можливість повторної перевірки текстів після видалення неунікальних фрагментів;
- збереження історії перевірок та ведення логу подій під час функціонування алгоритму.

Але, незважаючи на існуючу достатньо велику різноманітність систем перевірки унікальності текстів, їх поєднує один негативний аспект – неоднозначність отриманих результатів аналізу, рівень достовірності якої знаходиться в істотній залежності від умов перевірки, а також, більшою мірою, від організації бази даних текстів за якими здійснюється перевірка. При цьому, можливість отримання достовірного результату аналізу тексту ускладнюються тим, що набули широкого поширення методи трансформації тексту, що дозволяють штучно підвищити унікальність тексту, що не дозволить системі представити адекватний результат перевірки [225, 232, 234].

З погляду на вищесказане, можна зробити висновок, що вдосконалення систем перевірки текстів на унікальність, особливо в частині логічної організації бази даних текстів та алгоритмів перевірки, є актуальною науково-технічною задачею, рішення якої дозволить ефективно протидіяти незаконному використанню запозичених продуктів інтелектуальної праці.

З метою визначення основних аспектів, що значною мірою впливають на достовірність результатів перевірки унікальності текстів, було здійснене планування та проведення експерименту, в якому різним програмам перевірки на плагіат представлялися текстові фрагменти, які повністю запозичені, тобто мають відсоток унікальності 0%.

В експерименті використовувалися чотири найбільш відомі та поширені системи перевірки на антиплагіат, назви яких, з метою недопущення реклами або антиреклами, не будуть розкриватися. Назвемо досліджені системи: «Система 1», «Система 2», «Система 3», «Система 4». Для перевірки достовірності експериментального дослідження, кожній з чотирьох систем представлялися для аналізу 20 фрагментів тексту, які є повністю запозиченими (при плануванні експерименту це 20 груп дослідів).

Визначення дисперсії середнього значення [235] розрахункового відсотку унікальності тексту здійснювалося як:

$$s_i^2 = \frac{\sum_{j=1}^m (y_{ji} - \bar{y}_i)^2}{m-1}, \quad (5.23)$$

де  $m$  – кількість рівнобіжних дослідів;

$y_{ji}$  – відгук  $j$ -го рівнобіжного дослідів;

$\bar{y}_i$  – середній відгук в даному досліді.

Оцінка відтворювальності середнього значення відсотку унікальності тексту проводилася за критерієм Кохрена [233]

$$G = \frac{s_{i \text{ макс}}^2}{\sum_{i=1}^n s_i^2}, \quad (5.24)$$

де  $n$  – кількість груп дослідів;

$s_i^2$  – дисперсія  $i$ -го дослідів, яка визначається по формулі (5.22);

$s_{i \text{ макс}}^2$  – максимальне значення дисперсії в групі дослідів.

Результати проведеного експерименту з аналізу текстових фрагментів, представлених на мові оригіналу зведено у таблицю 5.1.

Таблиця 5.1 – Результати аналізу текстових фрагментів, представлених на мові оригіналу та з очікуваним відсотком унікальності 0%

	Система 1	Система 2	Система 3	Система 4
Середнє значення відсотку унікальності тексту ( $\bar{y}_i$ )	4,65	5,60	10,35	8,65
Середня дисперсія значення відсотку унікальності тексту ( $\bar{s}_i^2$ )	2,33	4,36	2,72	1,74
Значення коефіцієнту Кохрена (при довірчому інтервалі 0,95)	0,37	0,39	0,32	0,31

Проаналізував результати дослідів, представлені в таблиці 1, можна зробити висновок, що усі системи практично однаково визначили представлений плагіат. Значення коефіцієнту Кохрена не перевищує табличного значення, відповідно до методики [235]. Такі результати можна назвати очікуваними, так як тексти бралися з загально доступних джерел і в базах даних досліджуваних систем однаковою мірою малися відомості про наявний плагіат.

Враховуючи той факт, що системи перевірки можуть аналізувати тексти за умови їхнього можливого перекладання на інші мови, наступним етапом досліджень було проведено автоматизоване перекладання текстів загальновідомими системами перекладу з української мови оригіналу на англійську та з української мови оригіналу на російську. Для збереження чистоти експерименту ручне редагування перекладених текстів не проводилася. Результати аналізу на унікальність текстів перекладених з української мови на англійську та з української мови на російську наведені в таблицях 5.2 і 5.3 відповідно.

Таблиця 5.2 – Результати аналізу текстових фрагментів, переведених з української мови на англійську та з очікуваним відсотком унікальності 0%

	Система 1	Система 2	Система 3	Система 4
Середнє значення відсотку унікальності тексту ( $\bar{y}_i$ )	65,95	62,85	56,18	67,34
Середня дисперсія значення відсотку унікальності тексту ( $\bar{s}_i^2$ )	6,38	8,03	7,64	5,12
Значення коефіцієнту Кохрена (при довірчому інтервалі 0,95)	0,35	0,36	0,34	0,30

Аналіз результатів, зведених в таблиці 5.2 і 5.3 показав, що, з достатньо високою відтворювальністю, усі чотири системи показали, що переклад тексту з відсотком плагіату 100% дозволяє значно підвищити розрахунковий відсоток унікальності, при чому, переклад на англійську мову дає більш високий результат унікальності. Отримані результати пояснюються тим, що усі системи перевірки текстів на унікальність, фактично, здійснюють пошук

послідовності символів, що співпадають з фрагментами, наявними в базі даних, на кшталт пошуковим системам.

Таблиця 5.3 – Результати аналізу текстових фрагментів, переведених з української мови на російську та з очікуваним відсотком унікальності 0%

	Система 1	Система 2	Система 3	Система 4
Середнє значення відсотку унікальності тексту ( $\bar{y}_i$ )	40,30	38,74	44,57	41,05
Середня дисперсія значення відсотку унікальності тексту ( $\bar{s}_i^2$ )	7,32	8,94	7,42	5,04
Значення коефіцієнту Кохрена (при довірчому інтервалі 0,95)	0,36	0,38	0,36	0,32

Таким чином, синтаксично, текст російською мовою буде мати більшу кількість збігів з текстом українською мовою завдяки більшій кількості однакових символів. Цей аспект може бути використаний плагіаторами для підвищення унікальності тексту, через автоматичну підміну на латинські символи, наприклад в українському тексті, букв, що мають однакове написання (тобто підміна «а», «Н», «К», «р» та інших символів кирилиці на відповідні символи латиниці). Візуально це буде не помітно, але при аналізі буде отриманий більш високий відсоток унікальності.

Очевидно, що найбільш ефективним і універсальним способом підвищення достовірності визначення унікальності тексту може бути повна відмова від синтаксичного аналізу текстів на користь логічного (сміслового) аналізу. Але, велике різноманіття текстів, їхньої тематики та направленості, а

також достатньо великий розбіг специфічних понять з різних областей знань не дозволить в сучасних умовах створити систему штучного інтелекту з такою розвинутою семантикою та забезпечити її ефективне функціонування.

Відповідно до вищенаведеного, на сучасному етапі розвитку ІТ має науково-практичну цінність вирішення задачі вдосконалення існуючих систем синтаксичного аналізу текстів на унікальність з метою підвищення достовірності результатів, що отримуються, особливо, за умови вірогідності представлення перекладеного або трансформованого тексту.

На основі проведеного аналізу було встановлено, що достовірність визначення відсотку унікальності текстових фрагментів знаходиться в істотній залежності від форми представлення матеріалу на аналіз, а також форми та вмісту бази даних текстових фрагментів з якими здійснюється порівняння [236]. При цьому, якщо для порівняння текстових фрагментів система повинна здійснювати його трансформацію (наприклад під час порівняння фрагментів на різних мовах) – достовірність результату значно зменшується. У відповідності до цього, підвищення достовірності аналізу можливе через усунення необхідності попередньої трансформації текстових фрагментів та забезпечення прямого порівняння наданих фрагментів з вмістом бази даних системи перевірки на плагіат. Такий варіант стає можливим у випадку спеціальної організації бази даних систем перевірки текстів на унікальність, що запропонований далі.

Для вирішення питання підвищення достовірності визначення відсотку унікальності текстів, пропонується використання гібридного розпізнавання образів [236]. Об'єктом розпізнавання є СС – текстовий фрагмент, що надається для аналізу, а фрагменти з якими здійснюється порівняння, є апріорно визначеними класами, ступінь подібності до яких буде визначати відсоток унікальності тексту, що перевіряється. Кожен з цих класів  $C_1-C_n$  характеризується відповідним образом  $P_1-P_n$ . Узагальнена схема порівняння



представленого тексту з наявними в базі даних фрагментами, з точки зору теорії розпізнавання образів, представлена на рисунку 5.26.

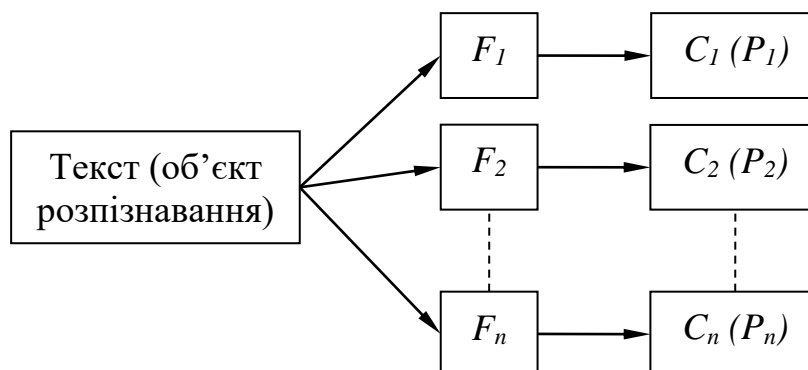


Рисунок 5.26 – Узагальнена схема порівняння тексту з наявними в базі даних фрагментами, з точки зору теорії розпізнавання образів

Представлена схема повністю відповідає традиційному підходу щодо побудови систем перевірки на унікальність текстової інформації. Відповідно до неї, для врахування можливих трансформацій тексту, що представляється для аналізу, вводяться функції перетворення  $F_1-F_n$ , що дозволяють співставити різноманітні форми представлення інформації з наявними у базі (з класами  $C_1-C_n$ , які характеризуються образами  $P_1-P_n$  відповідно) [236]. Наприклад, така трансформація має місце у випадку необхідності співставлення тексту з урахуванням можливих варіантів його перекладу іншими мовами.

Для виключення трансформації тексту, узагальнена схема на рисунку 5.26 перетворюється до варіанта використання гібридного розпізнавання образів та буде мати вигляд, проілюстрований на рисунку 5.27.

В ІТ визначення ступеня унікальності текстових даних  $C_1-C_n$  характеризуються множиною репрезентативних образів  $\{P^R\}$  причому, за кожним образом з цієї множини може бути здійснене віднесення наданого для аналізу тексту до відповідного передвизначеного класу.

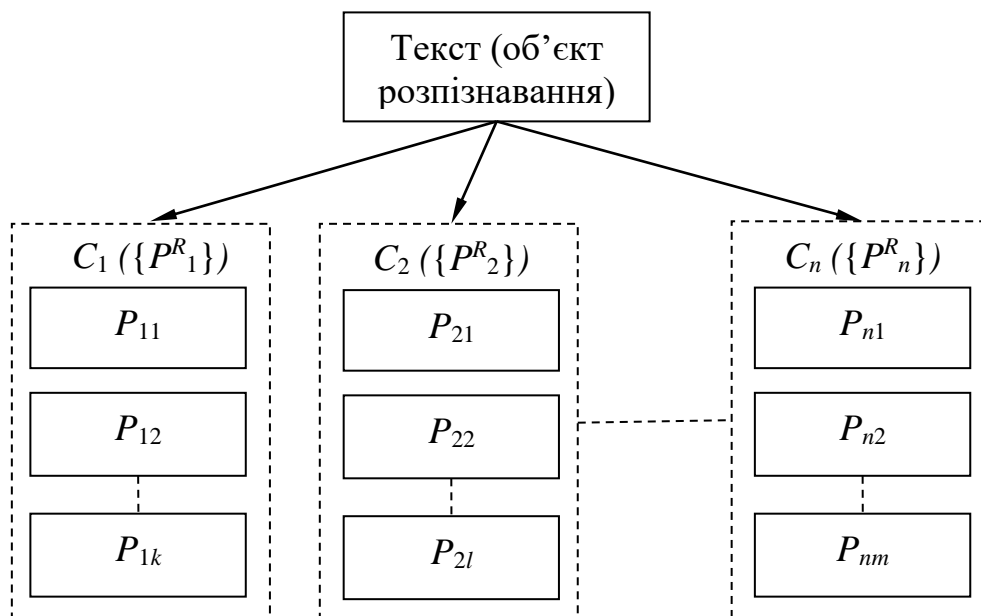


Рисунок 5.27 – Узагальнена схема порівняння тексту за наявними в базі даних фрагментами, з використанням гібридного розпізнавання образів

Відмінності ІТ визначення ступеня унікальності текстових з використанням гібридного розпізнавання образів полягають у тому, що у базі даних, кожний клас представляється сукупністю репрезентативних образів, які відповідають можливим, апріорно заданим формам представлення текстової інформації.

В процесі аналізу, фрагмент тексту, що аналізується, співставляється з кожним репрезентативним образом. Визначення відсотку унікальності тексту здійснюється відповідно до міри відстані, представленого для аналізу, образу від репрезентативних образів. Міра відстані для  $i$ -го образу відзначається як [236]

$$D_i = \|P_i - P^R\| = \sqrt{(P_i - P^R)'(P_i - P^R)}, \quad (5.25)$$

де  $D_i$  – характеристика відстані до  $i$ -го репрезентативного образу;

$P_i$  – образ, що надається для аналізу;

$P^R$  – репрезентативний образ.

Відсоток унікальності текстового фрагмента відзначається пропорційно значенню відстані, тобто

$$D_i \rightarrow 0 \Rightarrow y_i \rightarrow 0, \quad (5.26)$$

де  $y_i$  – відсоток унікальності текстового фрагменту.

На рисунку 5.28 представлена структурно-логічна схема послідовності виконання операцій в прикладній інформаційній технології гібридного розпізнавання образів для визначення ступеня унікальності текстових даних.

Реалізація інформаційної технології здійснюється через послідовне співставлення текстових даних з репрезентативними образами, що характеризують класи, згідно схеми на рисунку 5.27. На рисунку 5.28 представлені послідовні етапи співставлення, після яких здійснюється визначення відстані від класів. Якщо на етапі порівняння до жодного з класів не отримується апріорно заданої мінімальної відстані, то здійснюється наступний етап порівняння для уточнення результату. Таким чином, якщо в процесі співставлення з усіма репрезентативними образами, які представляють усі можливі варіанти представлення текстів з трансформаціями, не було досягнуто апріорно заданого мінімального значення метрики, то такий текст вважається унікальним та заноситься до бази знань для використання у наступних циклах аналізу текстових даних.

У випадку не унікальності текстових даних, апріорно задане мінімальне значення метрики може бути досягнуто на будь-якому етапі порівняння. Таким чином, аналіз текстових даних здійснюється до моменту доки не будуть виконані співставлення з усіма наявними репрезентативними образами, або доки за результатом поточного порівняння не буде отримане значення метрики, що не перевищує апріорно задане. В цьому випадку,

апріорно задане мінімальне значення метрики дозволяє встановити ступінь, який є пороговим для визначення унікальності тексту.

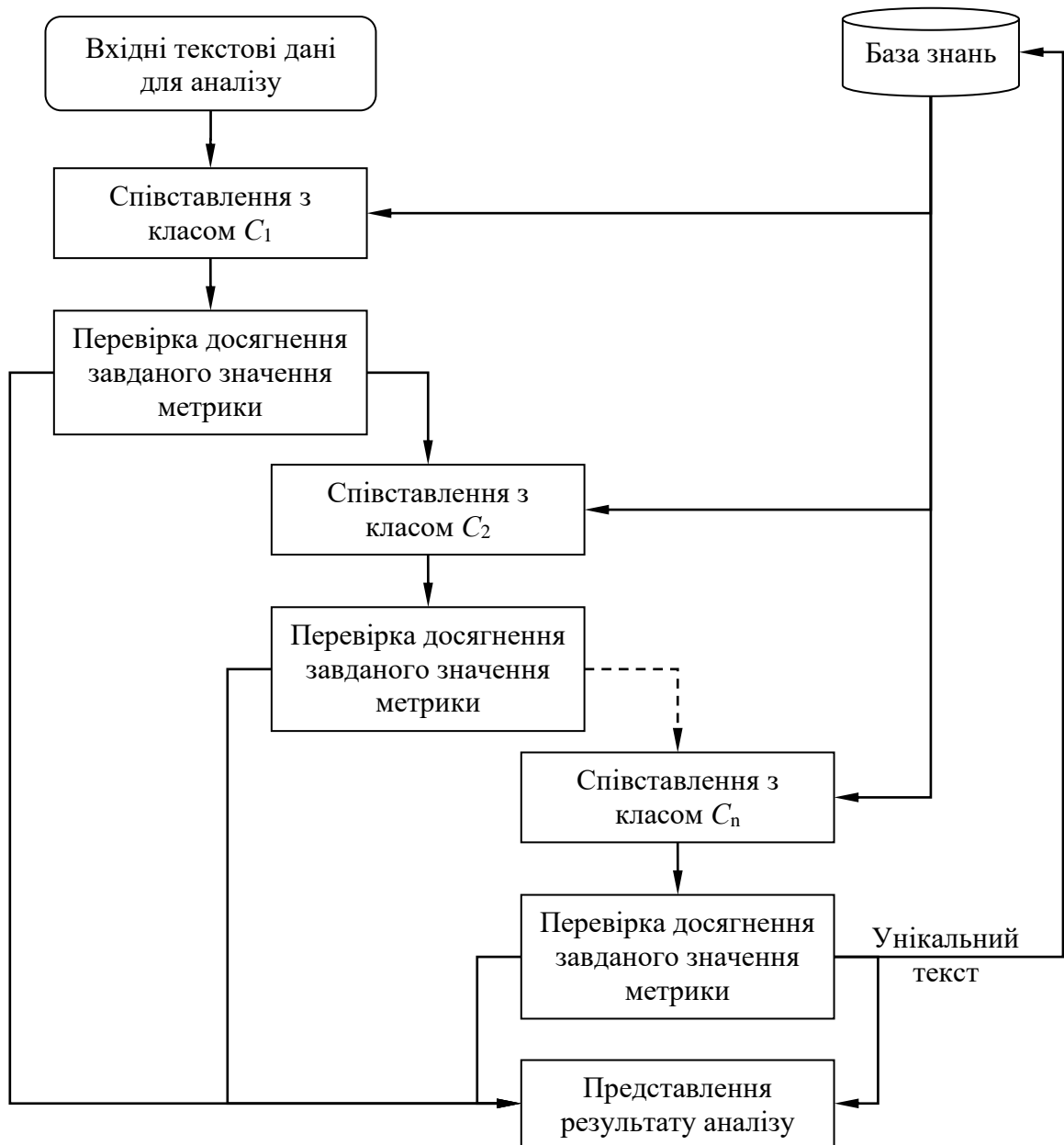


Рисунок 5.28 – Структурно-логічна схема послідовності виконання операцій в прикладній інформаційній технології гібридного розпізнавання образів для визначення ступеня унікальності текстових даних

Перевірка представленої ІТ щодо підвищення достовірності визначення відсотку унікальності текстової інформації проводилося на прикладі двох з

чотирьох раніше досліджених систем. На двох системах розробниками не передбачена можливість довільного формування бази даних за запропонованою схемою. Були повторно проведені дослід з аналізу тексту з очікуваним відсотком унікальності 0% за умови їхнього перекладу англійською і російською мовами. Результати дослідів зведені у таблицю 5.4.

Таблиця 5.4 – Результати аналізу текстових фрагментів, перекладених з української мови на англійську і російську та з очікуваним відсотком унікальності 0%

	Система 1	Система 2	Система 1	Система 2
	англійська мова		російська мова	
Середнє значення відсотку унікальності тексту ( $\bar{y}_i$ )	6,84	7,12	5,21	6,84
Середня дисперсія значення відсотку унікальності тексту ( $\bar{s}_i^2$ )	5,16	3,07	4,18	2,72
Значення коефіцієнту Кохрена (при довірчому інтервалі 0,95)	0,35	0,31	0,33	0,30

Аналіз даних таблиці 5.4 вказав на істотне збільшення достовірності визначення відсотку унікальності тексту. У порівнянні з результатами дослідів, представлених у таблицях 5.2 і 5.3 достовірність збільшилась у середньому на 54% [234, 236], що вказує на більш коректне порівняння та ефект від усунення попередньої трансформації текстових фрагментів та їхнього аналізу не на мові оригіналу.

Представлена ІТ визначення ступеня унікальності текстових даних була впроваджена в системі антиплагіату на базі вищого навчального закладу. Відповідний акт представлений у додатку Ж.

### **Висновки до розділу 5**

В розділі 5 вирішено восьме завдання дисертаційного дослідження, а саме: виконано практичну реалізацію розроблених моделей, методів та інформаційної технології для побудови систем обробки неоднорідних даних різноманітного спрямування. Розроблені нові науково-технічні рішення були використані в інформаційних технологіях: аналізу просторового розподілу температури коксового пирога для коксохімічного виробництва; ультразвукового вимірювання лінійних відстаней для систем автоматичної; визначення ступеня унікальності текстових даних. Основні висновки полягають в наступному.

Розроблена і впроваджена нова ІТ аналізу просторового розподілу температури коксового пирога з використанням гібридного розпізнавання образів, яка дозволяє здійснювати контроль без зупинки технологічного процесу, а також здійснювати аналіз наявності дефектів герметичності камери спікання. В результаті впровадження ІТ було отримана статистична характеристика достовірності результату на рівні 96,3%.

Розроблений новий метод аналізу просторового розподілу температури коксового пирога, який базується на використанні гібридного розпізнавання образів і не вимагає великої кількості пірометрів, а співставлення інформаційних ознак дозволяє здійснити корекцію температурної характеристики з метою зниження впливу наявних неоднорідних перешкод реєстрації.

Розроблена і впроваджена ІТ ультразвукового вимірювання лінійних відстаней, яка базується на використанні сукупності ультразвукових датчиків

з суміщеними передавачем та приймачем та гібридного розпізнавання образів, та дозволяє підвищити точність та розширити діапазон вимірювань.

Розроблена і впроваджена ІТ визначення ступеня унікальності текстових даних, яка завдяки використанню гібридного розпізнавання образів дозволяє підвищити достовірність отриманого результату за умови наявності неоднорідних перетворень текстової інформації.

## 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 компонується 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 Аткінсоном-Шифріном, для складних систем з неоднорідними даними та

різномірним 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 апіорно завданий рівень достовірності результату обробки неоднорідних даних.

10) З метою мінімізації кількості співставлень даних при аналізі та класифікації 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 при отриманні апіорно 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 різномірним 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.

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 ЛИТЕРАТУРЫ

1. Буравцев А. В. Серый управленческий анализ // Перспективы науки и образования. – 2017. – №5 (29). – С. 74–79.
2. Буравцев А. В. Алгоритмы обработки информации в сложных системах [Электронный ресурс] // Перспективы науки и образования. – 2018. – №1 (31). Режим доступа: <http://pnojournal.wordpress.com/archive18/18-01/>
3. Цветков В. Я. Информационная неопределенность и определенность в науках об информации // Информационные технологии. – 2015. – №1. – С. 3–7.
4. Перегудов Ф. И., Тарасенко Ф. П. Введение в системный анализ / Ф. И. Перегудов, Ф. П. Тарасенко. – М.: Высш. шк., 1989. – 367 с.
5. Симанков В. С. Адаптивное управление сложными системами на основе теории распознавания образов: Монография (научное издание) / В. С. Симанков, Е. В. Луценко. – Краснодар: Техн. ун-т Кубан. гос. технол. ун-та. – 1999. – 318 с.
6. Майоров А. А. Типизация сложных систем // Славянский форум. – 2014. – 1(5). – С. 131–137.
7. Цветков В. Я. Теория систем. Монография. М.: МАКС Пресс. – 2018. – 72 с.
8. Матчин В. Т. Обновление в сложной технологической системе // Славянский форум. – 2017. – № 3(17). – С. 62–68.
9. Цветков В. Я. Сложные технические системы // Образовательные ресурсы и технологии. – 2017. – 3 (20). – С. 86–92.
10. Корнаков А. Н. Модель сложной организационно-технической системы // Перспективы науки и образования. – 2015. – № 2. – С. 44–50.

11. Буравцев А. В. Функционирование сложной организационно-технической системы в транспортной сфере // Наука и технологии железных дорог. – 2017. – № 3(3). С. 48–58.
12. Симанков В. С., Луценко Е. В., Лаптев. В. Н. Системный анализ в адаптивном управлении: Монография (научное издание). / Под науч. ред. В. С. Симанкова. Ин-т совр. технол. и экон. – Краснодар. – 2001. – 258 с.
13. Монахов С. В. Методология анализа и проектирования сложных информационных систем / С. В. Монахов, В. П. Савиных, В. Я. Цветков. – М.: Просвещение. – 2005. – 264 с.
14. Спирли Э. Корпоративные хранилища данных. Планирование, разработка, реализация. Т. 1. Пер. с англ. – М.: Издательский дом «Вильямс». – 2001. – 400 с.
15. Педерсен Т. Технология многомерных баз данных / Т. Педерсен, К. Йенсен // Открытые системы. СУБД. – 2002. – № 1. – С. 45–50.
16. Дышленко С. Г. Специализированный банк геоданных / С. Г. Дышленко, В. Я. Цветков // Славянский форум. – 2016. – № 4(14). – С. 80–86.
17. Wolf W. Cyber-physical systems // Computer. – 2009. – vol. 42. – № 3. – p. 88–89.
18. Tsvetkov V. Ya. Information Asymmetry as a Risk Factor // European Researcher. Series A. – 2014. – vol. (86). – № 11-1. – pp. 1937–1943.
19. Моисеев Н. Алгоритмы развития. М.: Наука. – 1987. – 304 с.
20. Месарович М., Такахара Н. Общая теория систем: математические основы / М. Месарович, Н. Такахара. М.: Мир. – 1978. – 311 с.
21. Tsvetkov V. Ya. Information Situation and Information Position as a Management Tool // European researcher. Series A. – 2012, – vol. (36). – № 12-1. – p. 2166–2170.
22. Цветков В. Я. Систематика информационных ситуаций // Перспективы науки и образования. – 2017. – №5 (29). – С. 64–64.

23. Захожай О. И. Минимизация временной сложности в многокритериальных системах обработки информации / О. И. Захожай, В. В. Филимонцев // Журнал «Проблемы информационных технологий». – Херсон: ХНТУ. – 2014. – № 01 (015). – С. 161–166.

24. Захожай О. И. Высокоскоростная обработка изображений с применением цифровых фильтров на базе FPGA / О. И. Захожай, А. С. Солошенко // Сборник научных трудов ДонГТУ. – Алчевск: ДонГТУ. – 2013. – №39. – С. 250–254.

25. Клементьев К. Е. Системы реального времени. Самара: СГАУ. – 2008. – 52 с.

26. Олссон Г. Цифровые системы автоматизации и управления / Г. Олссон, Дж. Пиани. СПб: Невский диалект. – 2001. – 557 с.

27. Павлов А. А. Информационные технологии и алгоритмизация в управлении / А. А. Павлов, С. Ф. Теленик. – К.: Техніка. – 2002. – 344 с.

28. Трахтенгерц Э. А. Компьютерные системы поддержки принятия управленческих решений Текст. / Э. А. Трахтенгерц // Проблемы управления: ежемес. науч.-технический журн. / Ин-т проблем управления им. В. А. Трапезникова РАН. М.: «СенСиДат», ООО. – 2003. – №1. – С. 74–83.

29. Мачулин В. В. Эффективность систем обработки информации / В. В. Мачулин, А. П. Пятибратов. – М.: Изд-во «Советское радио». – 1972. – 280 с.

30. Finlay P. N. Introducing decision support systems. – Oxford, UK Cambridge, Mass., NCC Blackwell: Blackwell Publishers. – 1994.

31. Marakas G. M. Decision support systems in the twenty-first century. Upper Saddle River, N.J.: Prentice Hall. – 1999.

32. Power D. J. Web-based and model-driven decision support systems: concepts and issues. Americas Conference on Information Systems, Long Beach, California. – 2000.

33. Khokhlov V. N. Short-range forecast of atmospheric pollutants using non-linear prediction method / V. N. Khokhlov, A. V. Glushkov, N. S. Loboda, Yu. Ya. Bunyakova // Atmospheric Environment (Elsevier; The Netherlands). – 2008. – vol. 42. – P. 7284–7292.
34. Friedman J. Greedy Function Approximation: A Gradient Boosting Machine. — IMS 1999 Reitz Lecture.
35. Culp Mark. Ada: an R Package for Stochastic Boosting / Mark Culp, Kjell Johnson, George Michailidis // Journal of Statistical Software. – vol. VV. – Issue II. – 1997.
36. Freund Y., Shapire R. Experiments with a New Boosting Algorithm / Y. Freund, R. Shapire. – 1996.
37. Ratsch G. S. SVM and Boosting: One Class / G. Ratsch, B. Scholk, S. Mika, K.-R. Muller. Berlin: Springer. – 2000. – 243 p.
38. Ho Tin. The Random Subspace Method for Constructing Decision Forests. IEEE Transactions on Pattern Analysis and Machine Intelligence. – 1998.
39. Булдаков Н. С. Комбинированный алгоритм распознавания минных полей на аэрокосмических снимках // Математические методы в технике и технологиях: сб. трудов XVIII Междунар. науч. конф. Казань: Изд-во КГТУ. – 2005. – Т.6. – С. 60–63.
40. Круглов В. В. Интеллектуальные информационные системы: компьютерная поддержка систем нечеткой логики и нечеткого вывода / В. В. Круглов, М. И. Дли. – М.: Физматлит. – 2002. – 256 с.
41. Методы робастного, нейро-нечеткого и адаптивного управления / под ред. Н.В. Егунова. М.: Изд-во МГТУ им. Н.Э. Баумана. – 2002. – 744 с.
42. Большаков А. А. Синтез интеллектуальных систем дистанционного мониторинга и автоматизированной обработки снимков / А. А. Большаков, Н. С. Булдаков // Вестник Саратовского государственного технического университета. – 2012. – №1 (64). – Вып. 2. – С. 128–132.

43. Jain A. K. Statistical Pattern Recognition: A Review // IEEE Transaction on Pattern Analysis and Machine Intelligence / A. K. Jain, R. P. W. Duin. – 2000. – vol.22. – №1. – P.4–38.

44. Makridakis S. G., Wheelwright S. C. Forecasting methods for management. – New York: Wiley. –1989. – 164 p.

45. Захожай О. І. Інформаційна технологія розпізнавання образів в задачах автоматизованої обробки інформації й управління складними системами // Журнал «Проблеми інформаційних технологій» – Херсон: ХНТУ – 2013. – № 01 (013). – С. 61–68.

46. 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». – Київ: НТУУ «КПІ» – 2014. – № 6 (98) – С. 62–67.

47. Захожай О.И. Улучшение изображения дактилоскопических отпечатков в системах контроля доступа / О. И. Захожай, Ю. В. Чибисова // Сборник научных трудов ДонГТУ. – Алчевск: ДонГТУ. – 2011. – вып. 33 – С. 353–358.

48. 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. – Алчевськ: ДонДТУ. – 2010. – вип. 30 – С. 184–188.

49. Захожай О. І. Використання системи технічного зору для управління гільйотинними ножицями металургійного виробництва // Праці 8-ї міжнародної науково-практичної конференції “Сучасні інформаційні та електронні технології – 2007”. – Одеса. – 2007. – С. 81.

50. Захожай О. І. Інформаційна технологія розпізнавання образів при аналізі температурного режиму коксових печей // Праці 10-ї міжнародної



науково-практичної конференції «Сучасні інформаційні та електронні технології – 2009». – Одеса. – 2009. – С. 77.

51. 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. – 2010. – С. 34.

52. 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. – 2011. – С. 127.

53. Симанков В. С. Исследование эффективности управления обучением с применением адаптивной информационной модели / В. С. Симанков, Е. В. Луценко // Компьютерные технологии в науке и образовании XXI века: Тез. докл. Всероссийской конференции. – Ульяновск: УлГУ. – 1999. – 316 с.

54. Харкевич А. А. О ценности информации // Проблемы кибернетики: Сб. науч. трудов. Вып. 4. – М.: Физматгиз. – 1960. – С. 53–57.

55. Хартли Р. В., Передача информации. – В кн.: Теория информации и ее приложения / Пер. под ред. А. А. Харкевича. – М.: Физматгиз. – 1959. – 305 с.

56. Журавлев Ю. И. Алгоритмы распознавания, основанные на вычислении оценок / Ю. И. Журавлев, В. В. Никифоров // Кибернетика. – 1971. – №3. – С. 1–11.

57. Журавлев Ю. И., Избранные научные труды. – М.: Издательство Магистр. – 1998. – 420 с.

58. Глушков В. М. Кибернетика, вычислительная техника, информатика. Избранные труды. Том 3 из 3. Кибернетика и ее применение в

народном хозяйстве. Редкол.: Михалевич В. С. (отв. ред.) и др.; Ан усср. Ин-т кибернетики им. В. М. Глушкова. – Киев: Наукова думка. – 1990.

59. Пугачев В. С. Теория вероятностей и математическая статистика. – М.: Наука. – 1979. – 495 с.

60. Цыпкин Я. З. Основы информационной теории идентификации. – М.: Наука. – 1984. – 520 с.

61. Матросов В. Л. Синтез оптимальных алгоритмов в алгебраических замыканиях моделей алгоритмов распознавания // Распознавание, классификация, прогноз: Матем. методы и их применение. М.: Наука. – 1988. Вып.1. – С. 229–279.

62. Рудаков К. В. Об алгебраической теории универсальных и локальных ограничений для задач классификации // Распознавание, классификация, прогноз: Матем. методы и их применение. М.: Наука. – 1988. – Вып.1. – С.176–200.

63. Розенблатт Ф. Принципы нейродинамики (перцептрон и теория механизмов мозга). – М.: Мир. – 1965. – 480 с.

64. Дуда Р. Распознавание образов и анализ сцен: Пер. с англ / Р. Дуда, П. Харт. – М.: Мир. – 1978. – 510 с.

65. Ту Дж. Принципы распознавания образов: Пер. с англ / Дж. Ту, Р. Гонсалес. – М.: Мир. – 1978. – 410 с.

66. Фу К. Структурные методы в распознавании образов: Пер. с англ. – М.: Мир. – 1977. – 320 с.

67. Гренандер У. Лекции по теории образов. Том 1. Синтез образов. В 3-х томах. Пер. с англ. – М.: Мир. – 1979. – 383 с.

68. Гренандер У. Лекции по теории образов. Том 2. Анализ образов. В 3-х томах. Пер. с англ. – М.: Мир. – 1981. – 448 с.

69. Гренандер У. Лекции по теории образов. Том 3. Регулярные структуры. В 3-х томах. Пер. с англ. – М.: Мир. 1983. – 432 с.

70. Гонсалес Р., Вудс Р. Цифровая обработка изображений. 3-е издание, исправленное и дополненное. – М.: Техносфера. – 2012. – 1104 с.

71. Forsyth A. David, Ponce Jean. Computer vision. A modern approach. Second edition. London: Pearson. – 2012. – 761p.

72. Захожай О. І., Меняйленко 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.

73. Ли У. А. Методы автоматического распознавания речи: В 2-х книгах. Пер. с англ. / Под ред. У. Ли / У. А. Ли, Э. П. Нейбург, Т. Б. Мартин, Дж. Р. Уэлч, В. У. Зу, Р. М. Шварц, Дж. Е. Шуп, А. Р. Смит, М. Р. Самбур, Ф. Хейс-Роз, Г. Гудмэн, Р. Редди. – М.: Мир. – 1983. – Кн. 1. – 328 с.

74. Rabiner Lawrence. Fundamentals of speech recognition. Prentice-Hall, Inc. Upper Saddle River / Lawrence Rabiner, Biing-Hwang Juang. – NJ, USA. – 1993. – 507 p.

75. Меняйленко О. С., Захожай О. І. 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.

76. Меняйленко О. С., Захожай О. І. Спосіб контролю температури нагріву коксової печі. Патент на корисну модель № 91118 МПК (2014.01) C10B 17/00, C10B 23/00; опубл. 25.06.2014, бюл. №9.

77. Препелица Г. П. К вопросу о применении аппарата нейросетевого моделирования и распознавания образов в задачах прикладной экологии и гидрометеорологии / Г. П. Препелица, А. В. Лобода, О. Н. Грушевский, В. В. Буюджи // Вестник Одесского государственного экологического университета. – 2013. – Вып. 16. – С. 227–231.

78. Han Xiaozhuo. Pattern recognition and simulation in ecology Computational Ecology and Software / Xiaozhuo Han, Meng Gao, Cang Hui. –

Hong Kong: International Academy of Ecology and Environmental Sciences. – 2015 – 5(4). – pp. 271–275.

79. Maher Sean P. Pattern-recognition ecological niche models fit to presence-only and presence–absence data / Sean P. Maher, Christophe F. Randin, Antoine Guisan, John M. Drake // *Methods in ecology and evolution: British ecological society.* – 2014. URL: <https://doi.org/10.1111/2041-210X.12222>

80. Bauckhage Christian. Data Mining and Pattern Recognition in Agriculture / Christian Bauckhage, Kristian Kersting // *Künstliche Intelligenz.* – Berlin: Springer. – November 2013. – vol. 27. – Issue 4. – pp 313–324.

81. Raj M. P. Applications of Pattern Recognition Algorithms in Agriculture: A Review / M. P. Raj, P. R. Swaminarayan, J. R. Saini, D. K. Parmar. – *Int. J. Advanced Networking and Applications.* – 2015. – vol. 6. – pp. 2495–2502. URL: <http://www.ijana.in/papers/V6I5-13.pdf>

82. Rezania Mohammad. Application of evolutionary programming techniques in geotechnical engineering / Mohammad Rezania, A.A. Javadi // *Sixth European Conference on Numerical Methods in Geotechnical Engineering.* London: Taylor & Francis Group. – 2006. – pp. 677–682.

83. Ribeiro Antonio Júnior Alves. Proposition of a Geotechnical Mapping Based on Artificial Neural Networks for the Town of Caucaia, Ceará, Brazil for Paving Purposes / Antonio Júnior Alves Ribeiro, Carlos Augusto Uchôa da Silva, Suelly Helena de Araújo Barroso // *International Journal of Engineering & Technology IJET-IJENS.* – Al Mamourah Area (UAE): Science Publishing Corporation. – 2012. – vol. 12. – No. 05 – pp. 65–74.

84. Castiglia Massimina. Recognition of the Mechanical Properties for Soils in Complex Conditions: A Case Study / Massimina Castiglia, Filippo Santucci. – Campobasso (Italy): University of Molise. – 2016. – 283 p.

85. Петров Ю. А. Комплексная автоматизация управления предприятием. Информационные технологии теория и практика Текст.:

монография / Ю. А. Петров, Е. Л. Шлимович, Ю. В. Ирюпин. – М.: Финансы и статистика. – 2001. – 159 с.

86. Кукушкин А.А. Теоретические основы автоматизированного управления. Ч.2: Основы управления и построения автоматизированных информационных систем. – Орел: Изд-во ВИПС. – 1999. – 254 с.

87. Павлов А. А. Информационные технологии и алгоритмизация в управлении / А.А. Павлов, С.Ф. Теленик. – К.: Техніка. – 2002. – 344 с.

88. Информационные системы в экономике. Под ред. Г.А. Титоренко. – 2-е изд., перераб. и доп. – М.: ЮНИТИ-ДАНА. – 2008. – 463с.

89. Стешин А. И. Информационные системы в маркетинге. Учебное пособие. – СПб.: БГТУ «ВОЕНМЕХ». – 2009. – 107с.

90. Карминский А. М. Информационные системы в экономике: ч. 2. Практика использования / А. М. Карминский, Б. В. Черников. – М.: Финансы и статистика. – 2006. – 240с.

91. Окин А. А. Адаптация информационных систем // Экономика и Время. – 2004. – № 27 (314). – С. 46–59.

92. Zagoruiko N. G. Pattern Recognition Methods in Sociological Research / N. G. Zagoruiko, T. I. Zaslavskaja // In: Quantitative sociology. Blalock, H.M. at al. (Editors), Academic Press. – 1975. – pp. 429–440.

93. Adnan Amin. Advances in Pattern Recognition. / Amin Adnan, Dori Dov, Pudil Pavel, Freeman Herbert // Joint IAPR International Workshops SSPR'98 and SPR'098, Sydney, Australia. – Berlin; Heidelberg; New York; Barcelona; Budapest; Hong Kong; London; Milan; Paris; Singapore; Tokyo: Springer. – 1998. – 1052 p.

94. Фишхоф Б. Субъективная ожидаемая полезность: модель принятия решений / Б. Фишхоф, Б. Гойтейн, З. Шапиро // Процедуры оценивания многокритериальных объектов: Сб. тр. ВНИИСИ / Под. ред. О. И. Ларичева. – 1984 – № 9. – С. 121–127.

95. Kahneman D. Judgment under uncertainty: Heuristics and Biases / D. Kahneman, P. Slovic, A. Tversky. Cambridge: Cambridge University Press. – 1982. – 443 p.
96. Day R. H. Rational Choice and Economic Behavior // Theory and Decision. – 1997. – № 1. – pp. 437–449.
97. Garling T. Travel choice and the goal process utility distinction / T. Garling, K. Axhausen, M. Brydsten // Applied Cognitive Psychology. – 1996. – № 10.
98. Russell T. The Relevance of Quasi-Rationality in Competitive Markets. In: D. Bell, H. Raiffa, A. Tversky Decision Making: Descriptive, Normative and Prescriptive Interactions / T. Russell, R. Taylor. Cambridge: Cambridge University Press. – 1988.
99. Wu. G. Editing and Prospect Theory: Ordinal Independence and Outcome Cancellation // Working Paper of Harvard Business School. – 1993.
100. Горелик А. Л. Методы распознавания. Изд. 2 / А. Л. Горелик, В. А. Скрипкин. – М.: Высшая школа. – 1984. – 219 с.
101. Перегудов Ф. И. Введение в системный анализ / Ф. И. Перегудов, Ф. П. Тарасенко. – М.: Высш. шк. – 1989. – 367 с.
102. Журавлёв Ю. И. Об алгебраическом подходе к решению задач распознавания или классификации // Проблемы кибернетики. – 1978. – Вып. 33. – С. 5–68.
103. Васильев В. И. Распознающие системы: Справочник. – К.: Наукова думка. – 1983. – 230 с.
104. Fung L. W. An axiomatic approach to rational decision making in a fuzzy environment, in: Fuzzy Sets and their Applications to Cognitive and Decision Processes (Ed. by L. A. Zadeh, K. S. Fu, K. Tanaka, M. Shimura) / L. W. Fung, K. S. Fu. – New York: Academic Press. – 1975. – pp. 348–362.
105. Луценко Е. В. Автоматизированная система распознавания образов: математическая модель и опыт применения // В. И. Вернадский и

современность (к 130-летию со дня рождения): Материалы конференции. – Краснодар: Краснодарское книжное изд-во. – 1993. – С. 37–42.

106. Луценко Е. В. Разработка методологии синтеза адаптивных АСУ сложными объектами на основе применения моделей распознавания образов и принятия решений. Дисс. канд. техн. наук / КубГТУ. – Краснодар. – 1998. – 150 с.

107. Симанков В. С. Синтез адаптивных АСУ сложными системами с применением моделей распознавания образов / В. С. Симанков, Е. В. Луценко // Автоматизация и современные технологии. – 1999. – № 1. – С. 32–37.

108. Zgurovsky M. The Fundamentals of Computational Intelligence: System Approach / M. Zgurovsky, Yu. Zaychenko. – Berlin: Springer. – 2017. – 375 p.

109. Вапник В. Н. Теория распознавания образов / В. Н. Вапник, А. Я. Червонекис. – М.: Наука. – 1974. – 298 с.

110. Барабаш Ю. Л. Коллективные статистические решения при распознавании. – М.: Радио и связь. – 1983. – 224 с.

111. Темников Ф. Е. Теоретические основы информационной техники / Ф. Е. Темников, В. А. Афонин, В. И. Дмитриев. – М.: Энергия. – 1979. – 511 с.

112. Уинстон П. Искусственный интеллект: Пер. с англ. – М.: Мир. – 1980. – 520 с.

113. McLachlan J. Geoffrey. Discriminant analysis and statistical patterns recognition. A John Willey & Sons, Inc. – 2004. – 544p.

114. Niemann H. Pattern recognition and image understanding / H. Niemann, Yu. Zhuravlev, I. Gourevitch, I. Laptev. Amsterdam: IOS Press. – 2000. – 340 p.

115. Ripley B. D. Pattern recognition and Neural networks. Cambridge University Press. – 1996. – 416 p.

116. Левин Р. Практическое введение в технологию искусственного интеллекта и экспертных систем с иллюстрациями на Бейсике / Р. Левин, Д. Дранг, Б. Эделсон. – М.: Финансы и статистика. – 1990. – 239 с.

117. Растригин Л. А. Современные принципы управления сложными объектами. – М.: Советское радио. – 1980. – 120 с.

118. Кузин Л. Т. Основы кибернетики: Основы кибернетических моделей. Т. 2. – М.: Энергия. – 1979. – 584 с.

119. Сагдулаев Ю. С. Спектральная селекция и распознавание объектов / Ю. С. Сагдулаев, Т. Ю. Сагдулаев // Вопросы радиоэлектроники, серия Техника телевидения. – 2012. – вып. 2. – С. 97–106.

120. Сагдулаев Ю. С. Видеоинформационные технологии систем связи / Ю. С. Сагдулаев, Т. Ю. Сагдулаев, Ю. Б. Зубарев. – М.: Изд-во «Спутник+». – 2011. – 296 с.

121. Растригин Л. А. Адаптация сложных систем. – Рига: Зинатне. – 1981. – 375с.

122. Растригин Л. А. Метод коллективного распознавания / Л. А. Растригин, Р. Х. Эренштейн. – М.: Энергоатомиздат. – 1981. – 80 с.

123. Лазарев Ю. Н. Алгоритм решения многокритериальных задач управления Текст / Ю. Н. Лазарев, М. И. Гераськин // Известия Самарского научного центра Российской академии наук. – 2001. – Т.3. – № 1. – С. 80–85.

124. Forsyth R. Expert systems. Principles and case studies. – London: Chapman and Hall. – 1984. – 318 p.

125. Sugeno M. An approach to linguistic instruction-based learning / M. Sugeno, G. K. Park // Intern. J. of Uncertainty, Fuzziness and Knowledge-Based Systems. – 1993. – vol. 1. – No. 1. – pp. 19–56.

126. Shchepin E. V. Character recognition via critical points // Inter. J. of Imag. Systems and Technology. – 1991. – vol. 3. –pp. 141–157.



127. Kliatskine V. structural method for the recognition of complex historical tables / V. Kliatskine, G. Thorvaldsen, K. Zingerman, V. Lazarev, E. Shchepin // *History & Computing*. – 1997. – vol. 9. – No. 3. – pp. 58–77.

128. Кляцкин В. М. Структурные методы распознавания сложноорганизованных исторических табличных форм. Круг идей: модели и технологии исторической информатики / В. М. Кляцкин, К. М. Зингерман, В. В. Лазарев, Е. В. Щепин // *Труды III конференции Ассоциации «История и компьютер»*. – 1996. – С. 109–132.

129. Дюк В. А. Компьютерная психодиагностика. – СПб: Братство. – 1994. – 365 с.

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

131. Журавлёв Ю. И. Об алгебраическом подходе к решению задач распознавания или классификации. Проблемы кибернетики. – 1978. – Вып. 33. – С. 5–68.

132. Баскакова Л. В. Модель распознающих алгоритмов с представительными наборами и системами опорных множеств / Л. В. Баскакова, Ю. И. Журавлев // *Журн. вычисл. матем. и матем. физики*. – 1981. – Т. 21. – № 5. – С. 1264-1275.

133. Журавлев Ю. И. Распознавание образов и анализ изображений / Ю. И. Журавлев, И. Б. Гуревич // *Искусственный интеллект: Модели и методы*. – Москва: Радио и связь. – 1990. – 190 с.

134. Загоруйко Н. Г. Прикладные методы анализа данных и знаний. – Новосибирск: Изд-во Ин-та математики. – 1999. – 270 с.

135. Зувев Ю. А. Метод повышения надежности классификации при наличии нескольких классификаторов, основанный на принципе монотонности // *ЖВМиМФ*. – 1981. – Т.21. – № 1. – С.157–167.

136. Кочетков Д. В. Распознающие алгоритмы, инвариантные относительно преобразований пространства признаков // Распознавание, классификация, прогноз: Мат. методы и их применение. – М.: Наука. – 1989. – Вып. 11. – С. 178–206.

137. Кузнецов В. А. Распознавание нечетких систем по методу статистически взвешенных синдромов и его применение для иммуногематологической нормы и хронической патологии / В. А. Кузнецов, О. В. Сенько, А. В. Кузнецова // Химическая физика. – 1996. – т. 15. – №1. – С. 81–100.

138. Сенько О. В. Использование процедуры взвешенного голосования по системе базовых множеств в задачах прогнозирования // Вычисл. матем. и матем. физ. – 1995. – т. 35. – № 10. – С. 1552–1563.

139. Кузнецов Н. П. Опыт решения инженерно-технических задач на основе методов таксономии и распознавания. Применение прогрессивных технологий в добыче нефти на месторождениях Западной Сибири / Н. П. Кузнецов, А. Н. Ромасевич, В. В. Рязанов // Сб. научных трудов СибНИИНП. – 1988. – С. 62–66.

140. Барабаш Ю. Л. Коллективные статистические решения при распознавании. – М.: Радио и связь. – 1983. – 224 с.

141. Демьянов В. Ф. Введение в минимакс / В. Ф. Демьянов, В. Н. Малоземов. – М.: Наука. – 1972. – 311 с.

142. Горелик А. Л. Игровой подход к построению пространства признаков систем распознавания объектов и явлений // Кибернетика. – 1973 №5. – С. 124–138.

143. Журавлев Ю. И. Теоретико-множественные методы в алгебре логики. В кн.: Проблемы кибернетики, вып. 8. – М.: ГИФМЛ. – 1962. – 317 с.

144. Журавлев Ю. И. Экстремальные задачи, возникающие при обосновании эвристических процедур. – В кн.: Проблемы прикладной математики и механики. – М.: Наука. 1971. – 374 с.

145. Темников Ф. Е., Афонин В. А., Дмитриев В. И. Теоретические основы информационной техники / Ф. Е. Темников, В. А. Афонин, В. И. Дмитриев. – М.: Энергия. – 1979. – 511 с.
146. Методы и средства автоматизации психологических исследований / Под ред. Ю. М. Забродина. – М.: Наука. – 1982. – 300 с.
147. Экспертные системы: состояние и перспективы. Под ред. Д. А. Поспелова. – М.: Наука. – 1989. – 150 с.
148. Поспелов Д. А. Моделирование рассуждений. Опыт анализа мыслительных актов. – М.: Радио и связь. – 1989. – 184 с.
149. Поспелов Д. А. Ситуационное управление: теория и практика. – М.: Наука. – Гл. ред. физ.-мат. лит. – 1986. – 288с.
150. Наппельбаум Э. Л. Субъективное структурирование информации в задачах коллективного принятия решений / Э. Л. Наппельбаум, Д. А. Поспелов // Нормативные и дескриптивные модели принятия решений / Под ред. Б. Ф. Ломова, В. Ю. Крылова, Н. В. Крыловой. – М.: Наука. – 1981. – С.191–205.
151. Бор Н. Атомная физика и человеческое познание: Пер. с англ. – М.: Мир. – 1961. – 151 с.
152. Шурыгин А. М. Статистический кластер-алгоритм / А. М. Шурыгин // Математические методы распознавания образов: Сб. докл. 13-й Всерос. конф. Ленинградская обл., г. Зеленогорск, 30 сентября – 6 октября 2007 г. – М.: МАКС Пресс, 2007. – С. 241–242.
153. Колесникова С. И. Статистический подход к оцениванию зависимых признаков в интеллектуальных системах / С. И. Колесникова, А. Е. Янковская // Математические методы распознавания образов: Сб. докл. 13-й Всерос. конф. Ленинградская обл., г. Зеленогорск, 30 сентября 6 октября 2007 г. – М.: МАКС Пресс, 2007. – С. 143–146.

154. Колесникова С. И. Методы анализа информативности разнотипных признаков / С. И. Колесникова // Вестник томского государственного университета. – 2009. – № 1(6). – С. 69–80.

155. Meshalkin L. D. Some mathematical methods for the study of noncommunicable diseases / L. D. Meshalkin // Proc. 6-th Intern. Meeting of Uses of Epidemiol. in Planning Health Services. Yugoslavia, Primosten, 1971. V. 1. – P. 250–256.

156. Pearson K. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can reasonably be supposed to have arisen from random sampling / K. Pearson // Phil. Mag. 1900. – V. 50. – P. 157 – 175.

157. Орлов А.И. Анализ нечисловой информации в социологических исследованиях. М.: Наука, 1985. – С. 58–92.

158. Горелик А. Л. Методы распознавания / А. Л. Горелик, В. А. Скрипкин. – М.: Высш. шк, 2004. – 261 с.

159. Кручинин, А. Ю. Модель динамической оптимизации режимов мониторинга и диагностирования буровой скважины / А. Ю. Кручинин, Т. З. Аралбаев // Известия ОрелГТУ, серия «Информационные технологии». – 2006. – № 1(2). – С. 81–85.

160. Фукунага К. Введение в статистическую теорию распознавания образов / К. Фукунага. М.: Наука, 1979. – 368 с.

161. Рябенский В. М. Комбіновані системи розпізнавання образів / В. М. Рябенский, О. І. Захожай // Журнал «Проблеми інформаційних технологій» – Херсон: ХНТУ. – 2011. – № 01 (009). – С. 156–160.

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

163. 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.

164. 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.

165. Захожай О.І. Гібридне розпізнавання образів бля аналізу даних в складних системах / О. І. Захожай // Труды всеукраїнської конференції «Електроніка та телекомунікації-2019». – Сєверодонецьк: СНУ, 2019. – С. 113–114.

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

167. Яглом А. М. Вероятность и информация. Изд. 3-е, перераб. и доп. / А. М. Яглом, И. М. Яглом. – М.: Наука. – 1973. – 511 с.

168. Коваленко І. І. Моделі подання та виведення знань у системах ситуаційного управління / І. І. Коваленко, А. В. Швед, К. О. Антіпова. – Миколаїв: «Іліон». – 2018. – 91 с.

169. Витгих В. А. К определению понятия «ситуационное управление» / В. А. Витгих // Проблемы управления и моделирования в сложных системах: Труды XIV международной конф., 19-22 июня 2012 г., Самара. – Самара: РА СамНЦ РАН. – 2012. – С. 112–115.

170. Гюльмамедов Р. Г. Метод построения стратегии в системах ситуационного управления / Р. Г. Гюльмамедов // Информационно–управляющие системы. – 2011. – № 6. – С. 36–39.

171. Еременко Т. К. Использование СВР–подхода для баз знаний ситуационных центров / Т. К. Еременко, Ю. Г. Пилипенко // Системы підтримки прийняття рішень. – 2010. – Т. 1. – С. 151–153.

172. Скобелев П. О. Онтологии деятельности для ситуационного управления предприятиями в реальном времени / П. О. Скобелев // Онтология проектирования. – 2012. – №1(3). – С. 6–38.

173. Смирнов С. В. Онтологическое моделирование в ситуационном управлении / С. В. Смирнов // Онтология проектирования. – 2012. – № 2 (4). – С. 16–24.

174. Терехин Д. Э. Архитектура системы ситуационного управления / Д. Э. Терехин // Технологии Microsoft в теории и практике программирования: труды XIII Всероссийской научно-практической конф. студентов, аспирантов и молодых ученых, 22-23 марта 2016 г., Томск. – Томск: Изд-во ТПУ. – 2016. – С. 99–101.

175. Филиппович А. Ю. Ситуационные центры: определения, структура и классификация / А. Ю. Филиппович // PCWeek/RE. – 2003. – №26(392). – С. 21–22.

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

177. Satoh S. Recognition of Hand-written Patterns by Rotation-invariant Neocognitron / S. Satoh, J. Kuroiwa, H. Aso, S. Miyake // Proc. of ICONIP'98. – 1998 – vol. 1. – pp. 295–299.

178. Fausett L. Fundamentals of Neural Networks: Architectures, Algorithms and Applications. – Prentice Hall, New Jersey. – 2000.

179. Уоссермен Ф. Нейрокомпьютерная техника. Теория и практика. – М.: Мир. – 1992. – 260 с.

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

181. 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.

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

183. Захожай О.И. Определение проективной структуры и движения по бинокулярным соответствиям в системах технического зрения / О.И. Захожай // Сборник научных трудов Донбасского государственного технического университета. Вып. 25 – Алчевск, ДонГТУ, 2007 С. 243-249.

184. Zhongzhe Xiao. Classification of Emotional Speech Based on an Automatically Elaborated Hierarchical Classifier / Xiao Zhongzhe, Emmanuel Dellandrea, Weibei Dou, Liming Chen // International Scholarly Research Network ISRN Signal Processing. – 2011. – Article ID 753819. doi:10.5402/2011/753819.

185. 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». – Дніпропетровськ: НГУ. – 2015. – № 1 (145). – С. 120-126.

186. Аткинсон Р. Человеческая память и процесс обучения. – М.: Прогресс. – 1980. – 526 с.

187. Simon H. A. Information-processing models of cognition // J. Amer. Soc. Information Science. – Sept. 1981. – pp. 363–396.

188. Миллер Дж. А. Магическое число семь плюс или минус два. О некоторых пределах нашей способности перерабатывать информацию // Инженерная психология. – М.: Прогресс. – 1964. – 581 с.

189. Simon H. A. How big is a chunk // Science. – 1974. – №183. – pp. 482–488.

190. Эшби У. Росс Введение в кибернетику. Под ред. В.А. Успенского. – М.: Изд-во иностранной литературы. – 1959. – 429 с.

191. Захожай О. І. Основні аспекти структурної організації комбінованих систем розпізнавання образів / О. І. Захожай, Ю. Е. Паеранд // Вісник ХНТУ. – Херсон. – 2012. – №1 (44). – С. 221–225.

192. 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. – 195 p.

193. Захожай О. І. Спільний аналіз інформаційних ознак у багатопараметричних комбінованих системах розпізнавання образів / О.І. Захожай // Електротехнічні та комп'ютерні системи. – Одеса. – 2018. – № 29(105) – С. 78–86.

194. Моррис У. Наука об управлении. Байесовский подход. – М.: Мир. – 1971. – 282 с.

195. Гмурман В. Е. Теория вероятностей и математическая статистика. – М.: Высш. шк. – 2002. – 478 с.

196. Шурыгин А. М. Статистический кластер-алгоритм // Математические методы распознавания образов. – 2007. – С. 241–242.

197. Fisher R.A. The Use of Multiple Measurements in Taxonomic Problems // Annals of Eugenics. – 1936. – vol. 7. – pp. 179–188.

198. Захожай О. И. Раздельный анализ информационных признаков в многопараметрических комбинированных системах распознавания образов /



О. И. Захожай, А. С. Меньяйленко, В. А. Лыфарь // *Problemele Energeticii Regionale*. – Chişinău, Republica Moldova. – 2019. № 1-1(40).

199. 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.

200. 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*. – Сєверодонецьк: СНУ. – 2019. – № 2 (250). – С. 36-42.

201. 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*. – Одеса. – 2019. – № 31(107) – С. 95-103.

202. Воронцов К.В. Обзор современных исследований по проблеме качества обучения алгоритмов. // *Таврический вестник информатики и математики*. – 2004. – №1. – С. 5–24.

203. Захожай О.І. Селекція раціональної сукупності образів в комбінованих системах розпізнавання / О.І. Захожай // *Електротехнічні та комп'ютерні системи*. – Одеса. – 2013. – № 09(85). – С. 186–192.

204. 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

205. 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.

206. 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». – Херсон: ХНТУ. – 2015. – № 02 (018). – С. 106–111.

207. 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.

208. 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. – Сєверодонецьк: СХУ. – 2018. – № 6 (247). – С. 49–55.

209. Луценко Е. В. Когнитивная аналитическая система "ЭЙДОС-6.0", как инструмент исследования интегральной индивидуальности по Вольфу Мерлину // Актуальные проблемы социально-правовой подготовки специалистов и перспективы совершенствования системы комплектования ОВД: Сб. науч. тр. – Краснодар: КЮИ МВД РФ. – 1997. – С.136–141.

210. Мерлин В. С. Очерк интегрального исследования индивидуальности. – М. – 1986. – 187 с.

211. Третьяк В. Г. Учебная активность и индивидуальные особенности обучаемых юридического института МВД. – Краснодар: КЮИ МВД РФ. – 1996. – 110 с.

212. Олссон Г. Цифровые системы автоматизации и управления / Г. Олссон, Дж. Пиани. – СПб.: Невский Диалект. – 2001. – 557 с.

213. Зыль С. Н. QNX Momentics: основы применения / С. Н. Зыль. – СПб.: БХВ-Петербург. – 2005. – 256 с.

214. Климентьев К. Е. Системы реального времени. / К. Е. Климентьев. – Самара: СГАУ. – 2008. – 52 с.

215. Захожай О. І. Мікропроцесорна техніка / О. І. Захожай, В. Я. Жуйков, Ю. Е. Паеранд, Т. О. Терещенко. – Алчевськ: ВПЦ «Ладос». – 2012. – 499 с.

216. 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.

217. 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.

218. Макаров Г. Н. Химическая технология твердых горючих ископаемых / Г. Н. Макаров, Г. Д. Харлампович. М.: Химия. – 1986. – 496 с.

219. Patent CN 101070476B. Coke-oven heating-temperature intelligent control system / Liu Xiannan (刘显南), Wang Zhancheng (王占成), Zhao Guozhu (赵国柱); declarant Liu Xiannan (刘显南); declared 06.22.2007; published 05.19.2010.

220. Харлампович Г. Д. Технология коксохимического производства / Г. Д. Харлампович, О. А. Кауфман. – М: Нефть-газ. – 1995. – 384с.

221. Захожай О. И. Использование модели температурного режима в автоматизированной системе управления коксовой печью / О. И. Захожай, А. С. Шевцова // Сборник научных трудов ДонГТУ. – Алчевск: ДонГТУ. – 2011. – № 33. – С. 365-370.

222. «Мечел» – глобальная горнодобывающая и металлургическая компания. URL: <http://www.mechel.ru/>

223. Титов В. С. Автоматизированная система контроля на основе ультразвуковых датчиков / В. С. Титов, В. И. Иванов, М. В. Бобырь, Ю. Л. Бартенева // «Датчики и системы». – 2007, № 7. – С. 7–9.

224. Евтушенко Г. С. Ультразвуковой уровнемер для стационарных резервуаров нефтепродуктов / Г. С. Евтушенко, А. И. Солдатов, В. С. Макаров, П. В. Сорокин // «Датчики и системы». – 2010. – № 7. – С. 39–41.

225. Сичивиця О. Авторство і псевдоавторство в науці. Стаття друга. Плагіат і плагіатори / О. Сичивиця, В. К. Жданкин // Соціогуманітарні проблеми людини. – 2008. – № 3. – С. 39–47.

226. Ліннік І. Програмне забезпечення для виявлення плагіату: практичний аспект / І. Ліннік // Науковий блог НаУ «Острозька Академія». – 2013. – Режим доступу: <http://naub.oa.edu.ua/2013/prohramne-zabezpechennya-dlya-vuyavlennya-plahiatu-praktychnyj-aspekt>

227. Михайловський Ю. Б. Система Anti-Plagiarism як інструмент запобігання плагіату в навчальній та науковій діяльності / Ю. Б. Михайловський, Н. А. Длугунович // Вісник Хмельницького національного університету. Технічні науки. – 2013. – № 3. – С. 162–168.

228. Шарапова Е. В. Универсальная система проверки текстов на плагиат «Автор.net» / Е. В. Шарапова, Р. В. Шарапов // Информатика и её применения. – 2012. – № 3 (6). – С. 52–58.

229. Hariharan Sh. Automatic Plagiarism Detection Using Similarity Analysis / Sh. Hariharan // The International Arab Journal of Information Technology. – 2012. – № 4(9). – pp. 322–326 – Available from: <http://www.ccis2k.org/iajit/PDF/vol.9,no.4/2796-4.pdf>

230. Kharat R. Semantically Detecting Plagiarism for Research Papers / R. Kharat, P. M. Chavan, V. Jadhav, K. Rakibe // International Journal of Engineering Research and Applications (IJERA). – 2013. – № 3 (3). – pp. 77–80 – Available from: [http://www.ijera.com/papers/Vol3\\_issue3/P33077080.pdf](http://www.ijera.com/papers/Vol3_issue3/P33077080.pdf)

231. Shenoy M. Automatic Plagiarism Detection Using Similarity Analysis [online] / M. Shenoy, K. C. Shet, U. D. Acharya // Advanced Computing: An

International Journal (ACIJ). – 2012. – № 3 (3). – pp. 59-62 – Available from: <http://airccse.org/journal/acij/papers/0512acij06.pdf>

232. Singh R. Duplicity Detection System for Digital Documents / R. Singh, C. Dutta // International Journal of Soft Computing and Engineering (IJSCE). – 2012. – № 5 (2). – pp. 24-28 – Available from: [http://www.iaeng.org/publication/IMECS2011/IMECS2011\\_pp272-277.pdf](http://www.iaeng.org/publication/IMECS2011/IMECS2011_pp272-277.pdf)

233. Tschuggnall M. Detecting Plagiarism in Text Documents through Grammar-Analysis of Authors / M. Tschuggnall, G. Specht // 15th GI-Symposium Database Systems for Business, Technology and Web, 11th March-15th March, 2013. – 2013. – pp. 241-259.

234. 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.

235. Румшиский Л. З. Математическая обработка результатов эксперимента: Справочное руководство. М.: Изд-во «Наука». – 1971. – 192с.

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

## **ДОДАТОК А**

**Таблиця лінійного розподілу температури коксових печей  
на одній горизонтальній осі**

Таблиця А.1 – Значення ознак образів  $P_1'$  і  $P_2$  для кожної вертикалі п'яти експериментальних печей

Вертикаль	Піч 1		Піч 2		Піч 3		Піч 4		Піч 5	
	$P_1'$	$P_2$	$P_1'$	$P_2$	$P_1'$	$P_2$	$P_1'$	$P_2$	$P_1'$	$P_2$
1	1233	1138	1291	1194	1244	1105	1281	1202	1235	1233
2	1251	1184	1300	1202	1263	1150	1279	1211	1279	1251
3	1266	1167	1274	1179	1258	1152	1268	1187	1260	1266
4	1249	1161	1231	1176	1252	1144	1286	1184	1268	1249
5	1268	1175	1282	1181	1261	1164	1284	1211	1253	1268
6	1264	1197	1282	1189	1217	1162	1265	1210	1282	1264
7	1297	1205	1297	1198	1259	1144	1286	1187	1268	1297
8	1263	1173	1293	1195	1267	1169	1291	1217	1292	1263
9	1272	1201	1308	1188	1251	1169	1276	1206	1265	1272
10	1238	1160	1313	1204	1257	1174	1289	1220	1289	1238
11	1273	1177	1261	1155	1297	1191	1259	1197	1272	1273
12	1254	1147	1309	1208	1258	1173	1289	1216	1289	1254
13	1283	1183	1311	1209	1249	1165	1294	1211	1276	1283

Продовження таблиці А.1

Вертикаль	Піч 1		Піч 2		Піч 3		Піч 4		Піч 5	
	$P_1'$	$P_2$	$P_1'$	$P_2$	$P_1'$	$P_2$	$P_1'$	$P_2$	$P_1'$	$P_2$
14	1277	1175	1307	1206	1236	1173	1305	1213	1299	1277
15	1274	1174	1291	1185	1253	1137	1307	1231	1286	1274
16	1278	1164	1307	1175	1267	1176	1314	1232	1282	1278
17	1266	1183	1297	1198	1264	1168	1287	1198	1303	1266
18	1293	1153	1282	1212	1263	1177	1318	1238	1312	1293
19	1294	1190	1307	1199	1280	1170	1321	1248	1306	1294
20	1312	1190	1319	1219	1272	1178	1312	1249	1321	1312
21	1305	1197	1311	1217	1300	1202	1328	1251	1313	1305
22	1317	1201	1307	1219	1238	1142	1330	1243	1316	1317
23	1312	1192	1287	1211	1268	1186	1295	1211	1278	1312
24	1299	1183	1273	1202	1284	1176	1299	1199	1272	1299
25	1314	1197	1264	1159	1279	1135	1293	1185	1276	1314
26	1317	1199	1279	1207	1297	1179	1299	1218	1297	1317
27	1332	1197	1317	1215	1246	1132	1326	1219	1280	1332
28	1317	1191	1289	1217	1237	1117	1297	1204	1273	1317



## **ДОДАТОК Б**

**Графіки лінійних розподілів температури коксових печей  
на одній горизонтальній осі**

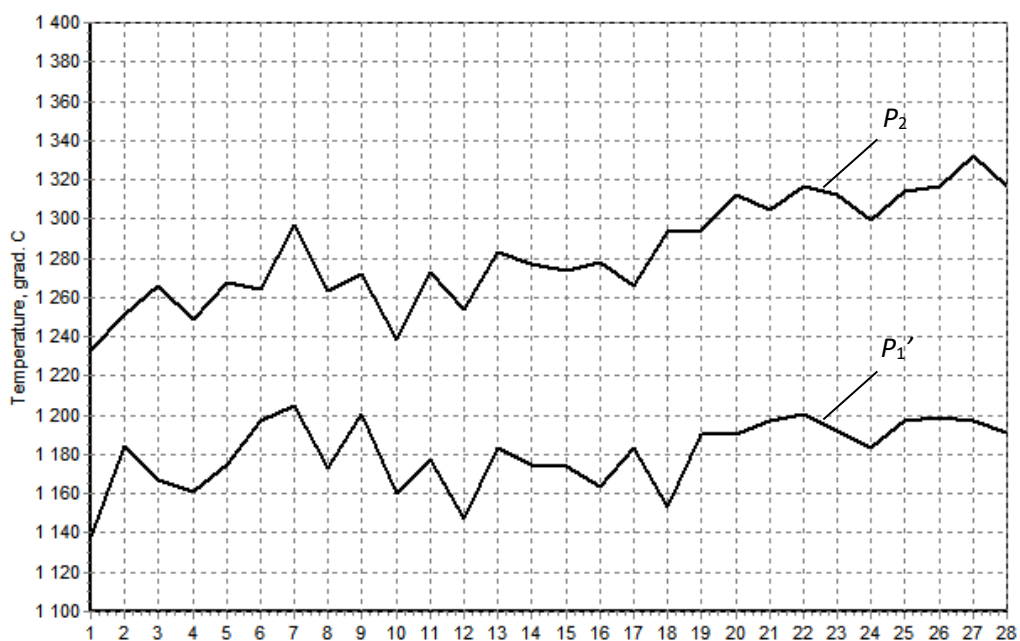


Рисунок Б.1 – Графік лінійних розподілів,  
що характеризуються образами  $P_1'$  і  $P_2$  для коксової печі 1

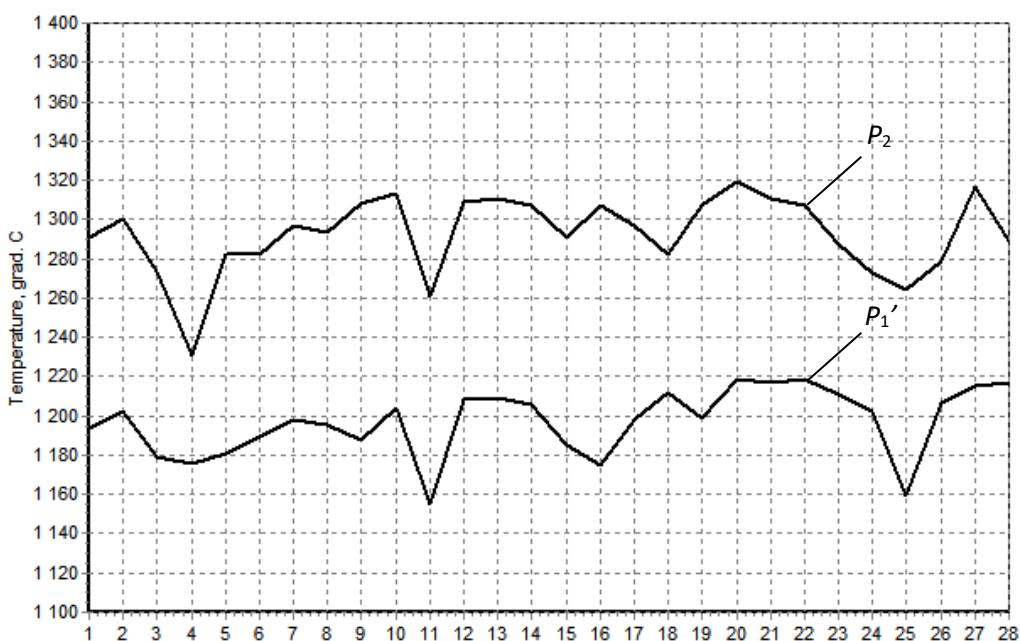


Рисунок Б.2 – Графік лінійних розподілів,  
що характеризуються образами  $P_1'$  і  $P_2$  для коксової печі 2

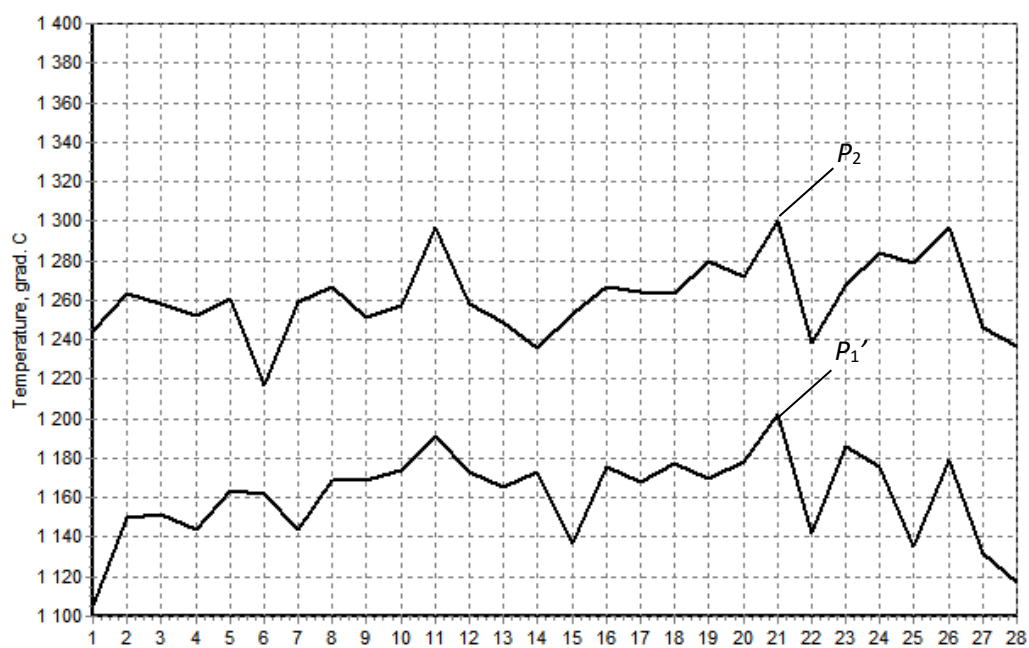


Рисунок Б.3 – Графік лінійних розподілів,  
що характеризуються образами  $P_1'$  і  $P_2$  для коксової печі 3

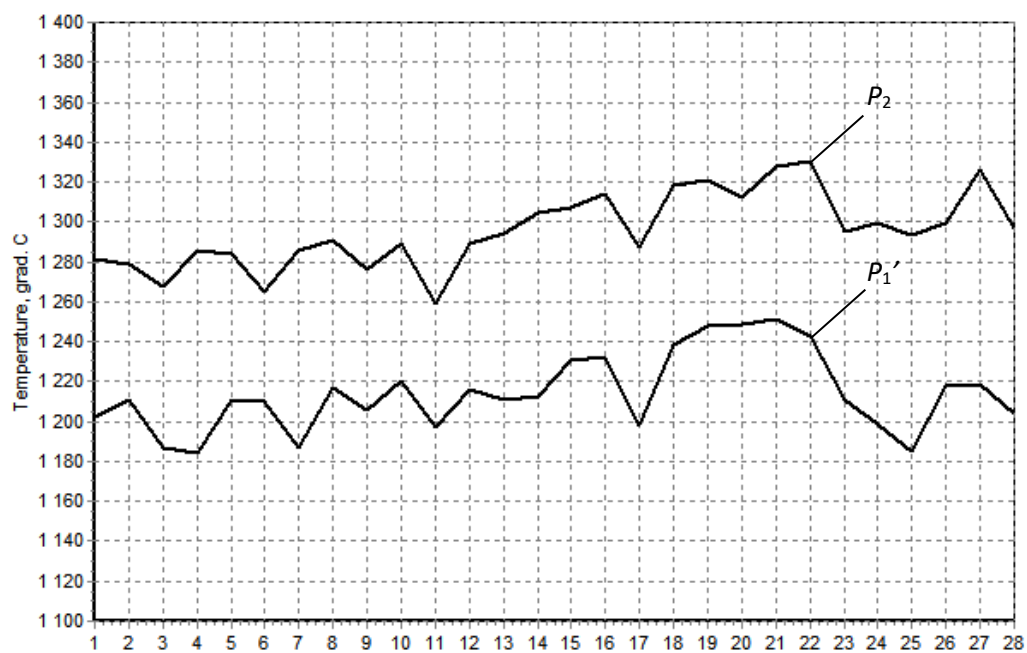


Рисунок Б.4 – Графік лінійних розподілів,  
що характеризуються образами  $P_1'$  і  $P_2$  для коксової печі 4

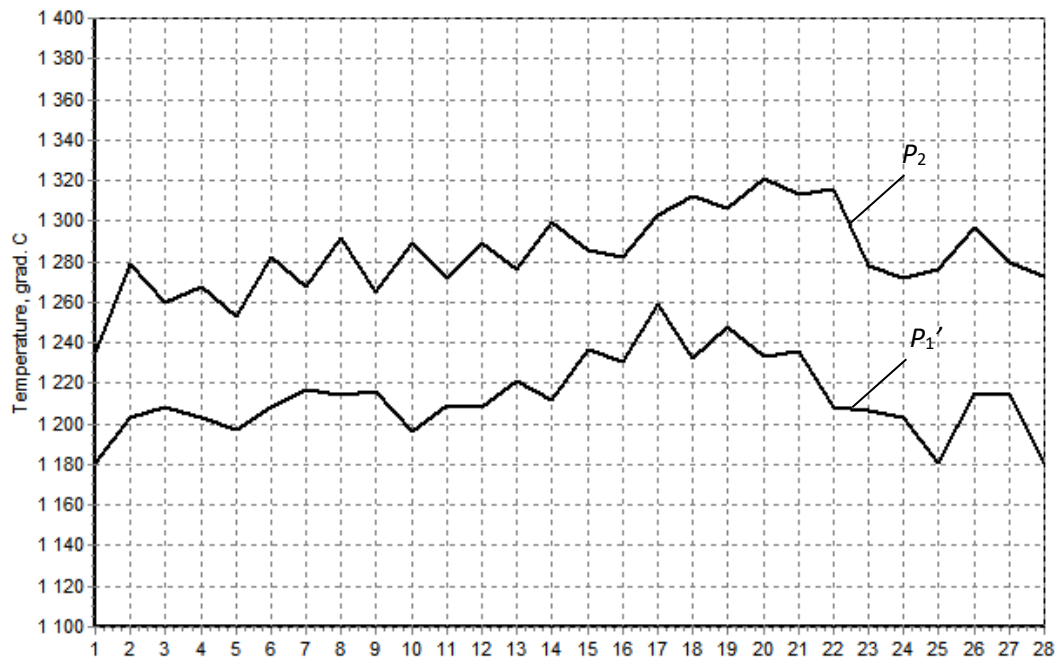


Рисунок Б.5 – Графік лінійних розподілів,  
що характеризуються образами  $P_1'$  і  $P_2$  для коксової печі 5

## **ДОДАТОК В**

**Таблиця розкиду характеристик між просторовим та еталонним  
лінійним розподілами температури коксових печей**

Таблиця В.1 – Розкид характеристик між просторовим ( $P_1$ ) та еталонним лінійним ( $P_2$ ) розподілами температури коксових печей

<b>Вертикаль</b>	<b>Піч 1</b>	<b>Піч 2</b>	<b>Піч 3</b>	<b>Піч 4</b>	<b>Піч 5</b>
1	-95	-97	-139	-79	-54
2	-67	-98	-113	-68	-76
3	-99	-95	-106	-81	-52
4	-88	-55	-108	-102	-65
5	-93	-101	-97	-73	-56
6	-67	-93	-55	-55	-74
7	-92	-99	-115	-99	-51
8	-90	-98	-98	-74	-78
9	-71	-120	-82	-70	-49
10	-78	-109	-83	-69	-93
11	-96	-106	-106	-62	-63
12	-107	-101	-85	-73	-81
13	-100	-102	-84	-83	-55
14	-102	-101	-63	-92	-87
15	-100	-106	-116	-76	-49
16	-114	-132	-91	-82	-51
17	-83	-99	-96	-89	-44
18	-140	-70	-86	-80	-80
19	-104	-108	-110	-73	-58
20	-122	-100	-94	-63	-88
21	-108	-94	-98	-77	-77
22	-116	-88	-96	-87	-108

**ДОДАТОК Г**

**Таблиця корегованого лінійного розподілу температури коксових печей  
на одній горизонтальній осі**

Таблиця Г.1 – Кореговані значення ознак образу  $P_1''$  і  $P_2$  для кожної вертикалі п'яти експериментальних печей

Вертикаль	Піч 1		Піч 2		Піч 3		Піч 4		Піч 5	
	$P_1''$	$P_2$	$P_1''$	$P_2$	$P_1''$	$P_2$	$P_1''$	$P_2$	$P_1''$	$P_2$
1	1219	1138	1279	1194	1217	1105	1269	1202	1247	1233
2	1258	1184	1286	1202	1245	1150	1272	1211	1287	1251
3	1250	1167	1264	1179	1246	1152	1259	1187	1271	1266
4	1228	1161	1249	1176	1244	1144	1267	1184	1272	1249
5	1259	1175	1264	1181	1275	1164	1280	1211	1259	1268
6	1273	1197	1274	1189	1248	1162	1271	1210	1274	1264
7	1283	1205	1301	1198	1253	1144	1282	1187	1272	1297
8	1275	1173	1286	1195	1272	1169	1283	1217	1281	1263
9	1255	1201	1287	1188	1261	1169	1264	1206	1283	1272
10	1248	1160	1298	1204	1263	1174	1283	1220	1277	1238
11	1265	1177	1273	1155	1285	1191	1265	1197	1265	1273
12	1231	1147	1298	1208	1246	1173	1278	1216	1280	1254
13	1269	1183	1306	1209	1233	1165	1276	1211	1280	1283
14	1265	1175	1289	1206	1228	1173	1291	1213	1294	1277
15	1257	1174	1279	1185	1230	1137	1311	1231	1301	1274



Продовження таблиці Г.1

Вертикаль	Піч 1		Піч 2		Піч 3		Піч 4		Піч 5	
	$P_1''$	$P_2$	$P_1''$	$P_2$	$P_1''$	$P_2$	$P_1''$	$P_2$	$P_1''$	$P_2$
16	1284	1164	1295	1175	1249	1176	1316	1232	1285	1278
17	1277	1183	1304	1198	1249	1168	1292	1198	1315	1266
18	1266	1153	1285	1212	1255	1177	1310	1238	1304	1293
19	1283	1190	1295	1199	1268	1170	1315	1248	1310	1294
20	1299	1190	1304	1219	1265	1178	1320	1249	1315	1312
21	1311	1197	1303	1217	1291	1202	1334	1251	1297	1305
22	1307	1201	1298	1219	1253	1142	1332	1243	1303	1317
23	1296	1192	1291	1211	1260	1186	1298	1211	1282	1312
24	1311	1183	1280	1202	1272	1176	1272	1199	1277	1299
25	1298	1197	1258	1159	1274	1135	1270	1185	1262	1314
26	1301	1199	1277	1207	1273	1179	1295	1218	1282	1317
27	1309	1197	1305	1215	1257	1132	1302	1219	1273	1332
28	1291	1191	1294	1217	1229	1117	1285	1204	1262	1317

## **ДОДАТОК Д**

**Графіки корегованих лінійних розподілів температури коксових печей  
на одній горизонтальній осі**

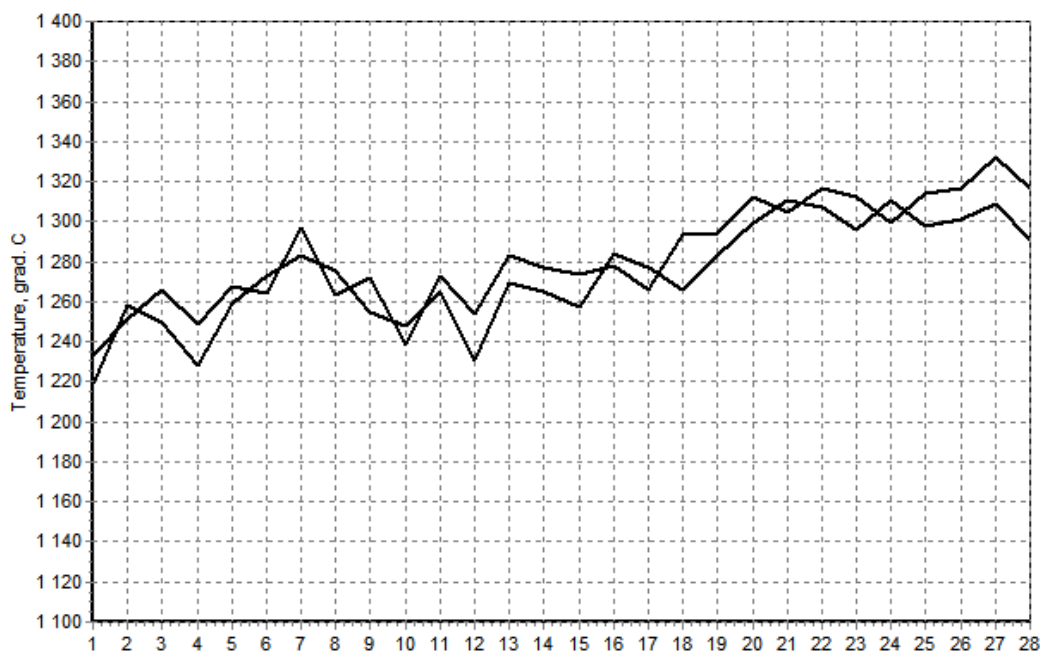


Рисунок Г.1 – Графік лінійних розподілів,  
що характеризуються образами  $P_1''$  і  $P_2$  для коксової печі 1

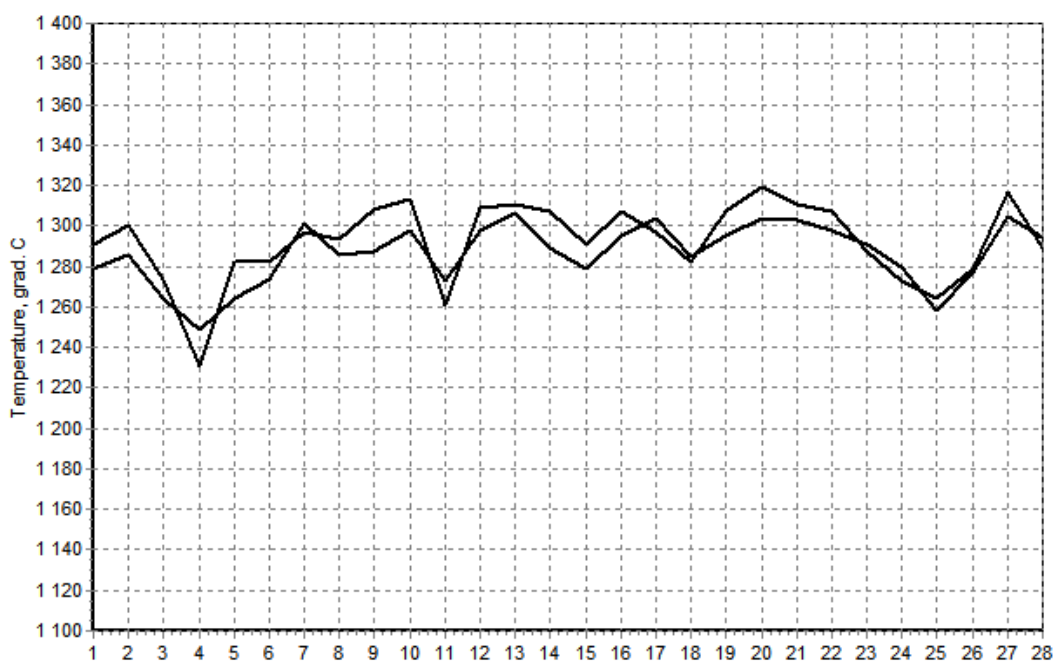


Рисунок Г.2 – Графік лінійних розподілів,  
що характеризуються образами  $P_1''$  і  $P_2$  для коксової печі 2

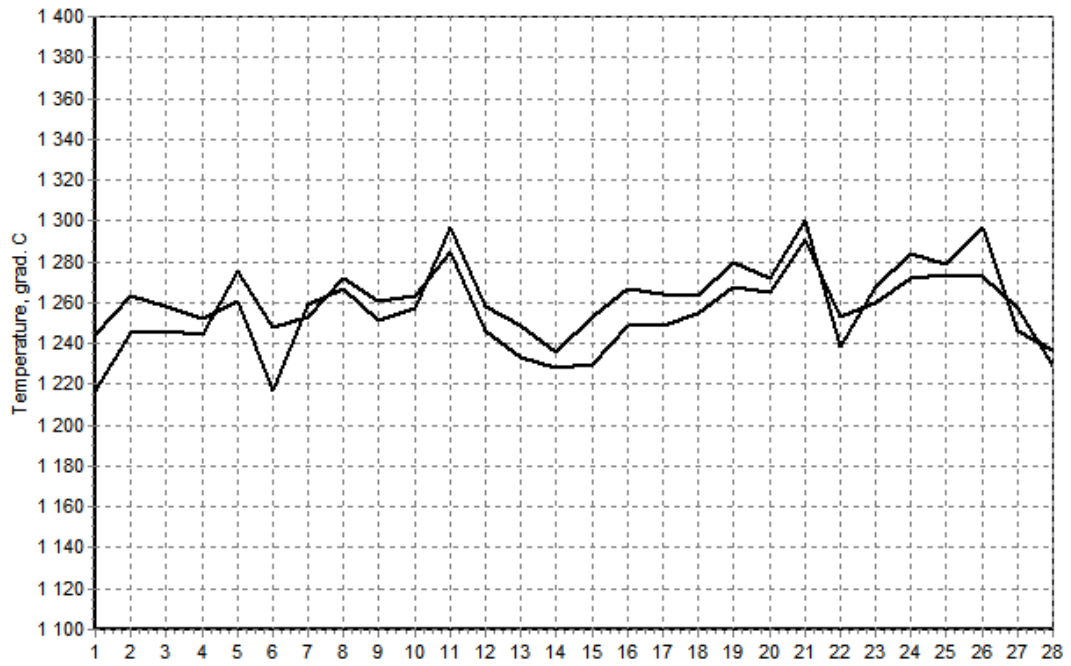


Рисунок Г.3 – Графік лінійних розподілів,  
що характеризуються образами  $P_1''$  і  $P_2$  для коксової печі 3

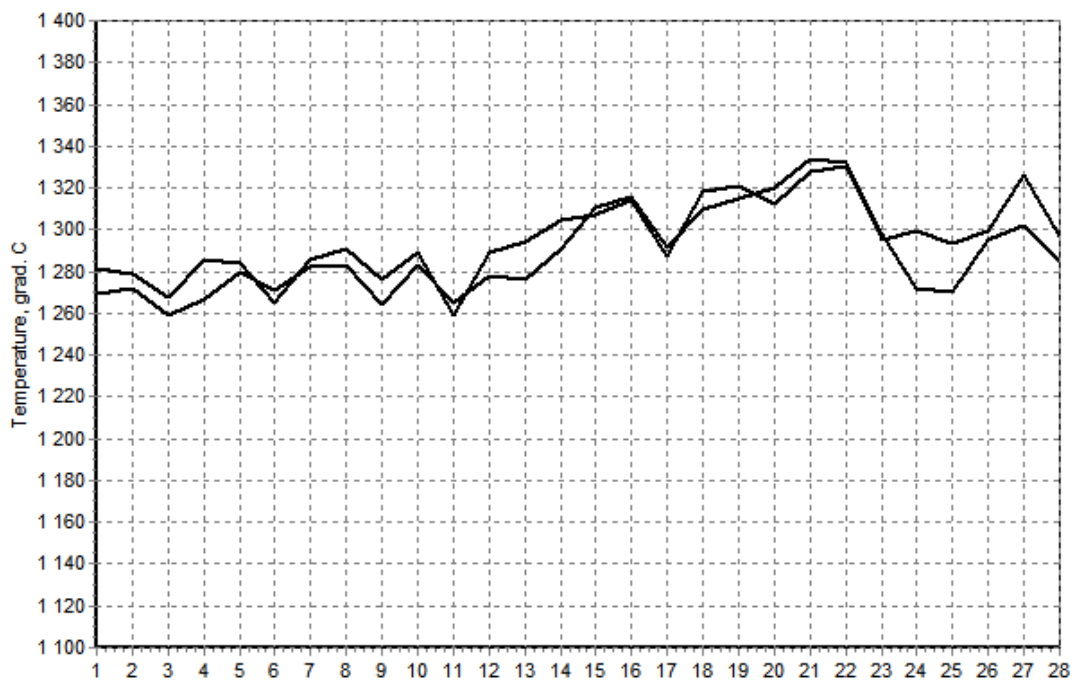


Рисунок Г.4 – Графік лінійних розподілів,  
що характеризуються образами  $P_1''$  і  $P_2$  для коксової печі 4

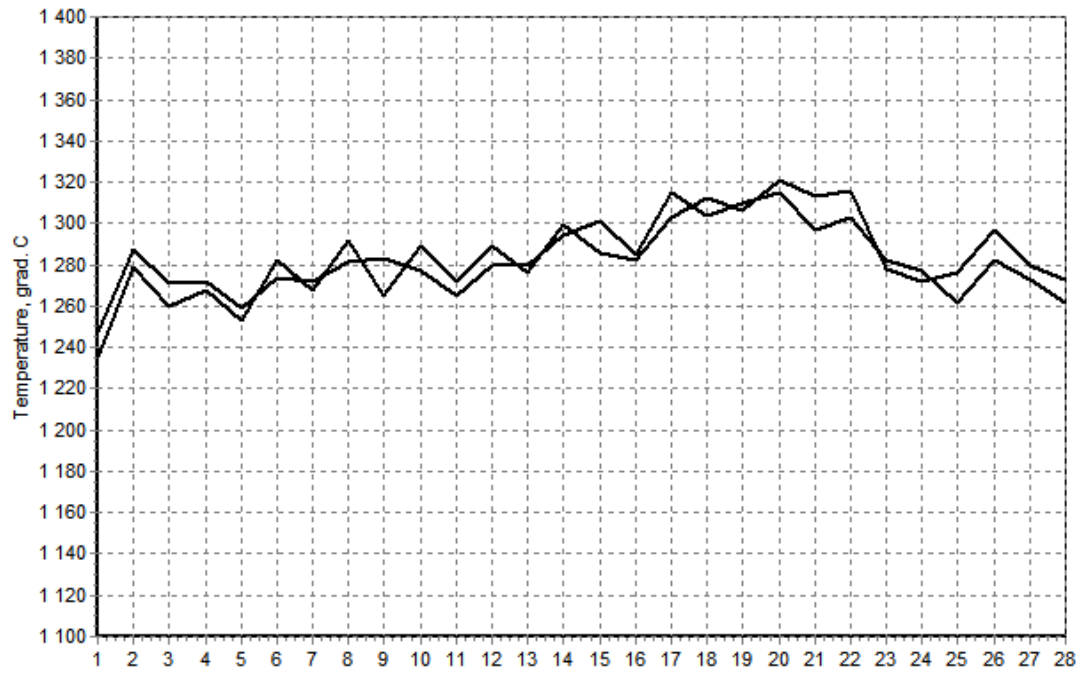


Рисунок Г.5 – Графік лінійних розподілів,  
що характеризуються образами  $P_1''$  і  $P_2$  для коксової печі 5

**ДОДАТОК Е**

**Результат співставлення отриманих температурних характеристик  
класу С**

Таблиця Е.1 – Відповідність температурних характеристик класу С та визначення статистичної достовірності розпізнавання

<b>Вертикаль</b>	<b>Піч 1</b>	<b>Піч 2</b>	<b>Піч 3</b>	<b>Піч 4</b>	<b>Піч 5</b>
1	1	1	0	1	1
2	1	1	1	1	1
3	1	1	1	1	1
4	0	1	1	1	1
5	1	1	1	1	1
6	1	1	0	1	1
7	1	1	1	1	1
8	1	1	1	1	1
9	1	0	1	1	1
10	1	1	1	1	1
11	1	1	1	1	1
12	0	1	1	1	1
13	1	1	1	1	1
14	1	1	1	1	1
15	1	1	0	1	1
16	1	1	1	1	1
17	1	1	1	1	1
18	0	1	1	1	1
19	1	1	1	1	1
20	1	1	1	1	1
21	1	1	1	1	1
22	1	1	1	1	1
23	1	1	1	1	1

Продовження таблиці Е.1

Вертикаль	Піч 1	Піч 2	Піч 3	Піч 4	Піч 5
24	1	1	1	0	1
25	1	1	1	0	1
26	1	1	0	1	1
27	0	1	1	0	1
28	0	1	1	1	1
$N_{cor}$	23	27	24	26	28
$N_{com}$	28	28	28	28	28
$D, \%$	82,1	96,4	85,7	92,9	100,0



## **ДОДАТОК Ж**

**Акти впровадження результатів дисертаційної роботи**



АКТ

про використання результатів наукової роботи

Захожая Олега Ігоровича

за напрямом «Інформаційні технології автоматизованої обробки інформації і управління на основі багатопараметричних комбінованих систем розпізнавання образів» в системі контролю температурного режиму коксової печі

Комісія у складі:

голови – начальника виробничо-технічного відділу, Августиновича О.А.

та членів:


начальника ділянки з ремонту електрообладнання цехів комбінату – Куліш С.І.;

інженера (проектно-організаційні роботи) 1 кат. ВТВ – Гапонова О.І.

у період з 14 по 18 жовтня 2013 р. розглянула впровадження у експериментальне дослідження результатів наукової роботи Захожая О.І. у напрямку “Інформаційні технології автоматизованої обробки інформації і управління на основі багатопараметричних комбінованих систем розпізнавання образів” на базі публічного акціонерного товариства «Алчевський металургійний комбінат».

Комісія встановила, що окремі положення наукової роботи використані при модернізації системи контролю температурного режиму нагрівальних печей ПАТ «АМК». Теоретичні проробки та експериментальні дослідження показали, що запропонований метод аналізу просторового розподілу температури на основі використання двох незалежних інформаційних каналів та подальшої обробки багатопараметричною системою розпізнавання образів дозволяє отримувати показники температури по всій нагрівальній зоні та реалізувати контроль наявності дефектів кладки печі та спостерігати за динамікою їхнього розвитку.

Перевірка точності отримання розподілу температур здійснювався з використанням еталонного пірометричного датчика, що сертифікований для використання у цьому технологічному процесі. В результаті, статистична характеристика подібності була отримана на рівні 96,3%, що задовольняє існуючим нормам.

Начальник виробничо-технічного відділу  О.А. АвгустиновичНачальник ділянки з ремонту ЕО цехів комбінату  С.І. КулішІнженер (проектно-організаційні роботи) 1 кат. ВТВ  О.І. Гапонов

«ЗАТВЕРДЖУЮ:»

Голова правління

АКВПТ «Укрчерметавтоматика»



Бриков В.П.

« 17 » 02 20 14 р.

**АКТ****впровадження результатів****дисертаційної роботи Захожая О. І.**

Комісія у складі: заступника директора з виробництва Тацохи Ю.І.; начальника ділянки з ремонту електрообладнання та автоматики Плугатаренка К.О. склала цей акт про те, що результати дисертаційної роботи Захожая О.І. були використані для вдосконалення систем ультразвукового контролю положення пересувного технологічного обладнання засобів автоматизації та безконтактного вимірювання відстані.

Використання запропонованого Захожаем О.І. методу вимірювання лінійної відстані на базі багатопараметричних комбінованих систем розпізнавання образів дозволив уніфікувати набір ультразвукових датчиків для широкого діапазону вимірювання та розширити цей діапазон. Ці аспекти дозволяють покращити експлуатаційні показники, спростити обслуговування вимірювального обладнання, а також забезпечити достатню точність в широкому діапазоні вимірювання.

Заступник директора з виробництва

Ю. Тацоха

Начальник ділянки з ремонту  
електрообладнання та автоматики

К. Плугатаренко

## «ЗАТВЕРДЖУЮ»



Ректор Луганського національного  
університету імені Тараса Шевченка

С.В. Савченко

09 20 19 р.

## АКТ

про впровадження результатів дисертаційної роботи

Захожая Олега Ігоровича

за напрямом: «Інформаційні технології автоматизованої обробки інформації і управління на основі багатопараметричних комбінованих систем розпізнавання образів», представленій на здобуття наукового ступеня доктора технічних наук

Цим актом підтверджується, що окремі результати наукової роботи Захожая Олега Ігоровича за напрямом: «Інформаційні технології автоматизованої обробки інформації і управління на основі багатопараметричних комбінованих систем розпізнавання образів» було використано в системі перевірки текстів на унікальність Луганського національного університету імені Тараса Шевченка. Використання технічних рішень, розроблених Захожаєм О.І., дозволило підвищити достовірність визначення унікальності текстів на 54 відсотки за рахунок більш повного і коректного представлення матеріалів для порівняння.

Проректор з науково-педагогічної роботи, д.пед.н., проф.

Л.Ц.Ваховський

Директор інституту психології і педагогіки, д.пед.н., проф.

О.Л. Караман

**«ЗАТВЕРДЖУЮ»**



Ректор Східноукраїнського національного  
університету імені Володимира Даля

О.В. Поркрян

« 17 » 06 20 19 р.

**АКТ**

**про впровадження в освітній процес результатів дисертаційної роботи**

Захожая Олега Ігоровича

за напрямом «Інформаційні технології автоматизованої обробки інформації і управління на основі багатопараметричних комбінованих систем розпізнавання образів», представленої на здобуття наукового ступеня доктора технічних наук

Цим актом підтверджується, що результати досліджень, які проводилися в межах наукового напрямку «Інформаційні технології автоматизованої обробки інформації і управління на основі багатопараметричних комбінованих систем розпізнавання образів», впроваджені на кафедрі програмування та математики Східноукраїнського національного університету імені Володимира Даля в процес підготовки фахівців зі спеціальностей 121 «Інженерія програмного забезпечення» та 123 «Комп'ютерна інженерія» для наступних дисциплін:

1) «Інженерія програмного забезпечення» – введений теоретичний матеріал за темою: «Базові технології, інструментальні та програмні засоби підтримки автоматизованої обробки інформації і управління складними системами»;

2) «Інтелектуальний аналіз даних» – доданий теоретичний матеріал стосовно методів і алгоритмів селекції інформаційних образів і ознак при інтелектуальній обробці даних та прийнятті рішень;

3) «Методи та засоби комп'ютерних інформаційних технологій» – доданий теоретичний матеріал за темою: «Комп'ютерні технології обробки даних на основі багатопараметричних комбінованих систем розпізнавання образів»;

4) «Інтелектуальні інформаційні системи» – введений теоретичний матеріал за темою: «Інформаційні системи автоматизованої обробки даних на основі апарату розпізнавання образів».

Узгоджено:

Завідувач кафедри програмування та математики  
Східноукраїнського національного університету  
імені Володимира Даля, д.т.н., доцент

В.О. Лифар

## **ДОДАТОК И**

### **Список публікацій здобувача**

*Наукові праці, в яких опубліковані основні наукові результати дисертації*

1. Меньяйленко О. С., Захожай О. І. Основи синтезу класифікаторів технічних систем розпізнавання образів з використанням моделей емоційних процесів людини // Науково-технічний журнал «Науковий вісник Національного гірничого університету». – 2015. – № 1(145). – С. 120–126. (Видання входить до міжнародних наукометричних баз: *SciVerse Scopus, Ulrich's Periodicals Directory (ULRICHSWEB), ВИНИТИ, EBSCO, РИНЦ (eLIBRARY), Index Copernicus*)

2. Захожай О. І. Інформаційна технологія розпізнавання образів в задачах автоматизованої обробки інформації й управління складними системами // Журнал «Проблеми інформаційних технологій». – 2013. – № 01(013). – С. 61–68. (Видання входить до міжнародних наукометричних баз: *Open Academic Journals Index (OAJI), РИНЦ (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 // Журнал «Проблеми інформаційних технологій». – 2015. – № 02(018). – С. 106–111. (Видання входить до міжнародних наукометричних баз: *Open Academic Journals Index (OAJI), РИНЦ (eLIBRARY), Research Bible, AcademicKeys*)

4. Меньяйленко О. С. Захожай О. І. Комбіновані системи розпізнавання образів при аналізі просторового розподілу температури коксового пирога // Електротехнічні та комп'ютерні системи. – 2013. – № 12(88). – С. 147–154. (Видання входить до міжнародних наукометричних баз: *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 // Електротехнічні та комп'ютерні системи. – 2018. – № 27(103). – С. 196–204. (Видання входить до міжнародних наукометричних баз: *Ulrich's Periodicals Directory (ULRICHSWEB)*, *Electronic Journals Library*, *Google Scholar*, *ВИНИТИ*)

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

7. Захожай О. И., Меняйленко А. С., Лыфарь В. А. Раздельный анализ информационных признаков в многопараметрических комбинированных системах распознавания образов // *Problemele Energeticii Regionale*. – Chişinău, Republica Moldova, 2019. – № 1-1(40). – С. 60–68. (Видання входить до міжнародних наукометричних баз: *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. И., Филимонцев В. В. Минимизация временной сложности в многокритериальных системах обработки информации // Журнал «Проблемы информационных технологий». – 2014. – № 01(015). – С. 161–166. (Видання входить до міжнародних наукометричних баз: *Open Academic Journals Index (OAJI)*, *РИНЦ (eLIBRARY)*, *Research Bible*, *AcademicKeys*)

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



10. Рябенський В. М., Захожай О. І. Комбіновані системи розпізнавання образів // Журнал «Проблеми інформаційних технологій» – 2011. – № 01(009). – С. 156–160. (Видання входить до міжнародних наукометричних баз: *Open Academic Journals Index (OAJI)*, *РИНЦ (eLIBRARY)*, *Research Bible*, *AcademicKeys*)

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

12. Захожай О. І., Шевцова А. С. Інформаційна модель системи непрямого аналізу стану коксової печі засобами комп'ютерного зору // Збірник наукових праць ДонДТУ. – 2010.– № 30 – С. 184–188.

13. Меняйленко О. С., Захожай О. І. Інформаційна технологія автоматизованої обробки даних в багатоканальних системах ультразвукового вимірювання // Наукові вісті Національного університету України «Київський політехнічний інститут». – 2014. – № 6(98). – С. 62–67. (Видання входить до міжнародних наукометричних баз: *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., Паєранд Ю. Е. Основні аспекти структурної організації комбінованих систем розпізнавання образів // Вісник ХНТУ. – 2012. – №1(44). – С.221–225.

15. Захожай О.І. Ефективні способи перетворення графічних зображень при їхньої сегментації та розпізнаванні засобами комп'ютерного зору // 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 // Сборник научных трудов ДонГТУ. – 2007 – № 25. – С. 243–249.

17. Захожай О. И., Солошенко А. С. Высокоскоростная обработка изображений с применением цифровых фильтров на базе FPGA // Сборник научных трудов ДонГТУ. – 2013. – №39. – С. 250–254.

18. Меньяйленко О. С., Захожай О. І., Бідюк П. І. Підвищення достовірності перевірки унікальності текстів шляхом використання комбінованих систем розпізнавання образів // Журнал «Системні дослідження та інформаційні технології» – 2017. – № 4. – С. 29–37. *(Видання входить до міжнародних наукометричних баз: Directory of Open Access Journal (DOAJ), EBSCO, Index Copernicus, РИИЦ (eLIBRARY), ВИНІТИ)*

19. Захожай О. І., Лифар В. О., Батурін О. І. Прийняття рішень на основі пошуку груп ідентичних класифікацій в багатопараметричних комбінованих системах розпізнавання образів // Вісник Східноукраїнського національного університету імені Володимира Даля. – 2019. – №2 (250). – С. 36–42. *(Видання входить до міжнародної наукометричної бази: Index Copernicus)*

20. Захожай О. І. Підвищення достовірності розпізнавання об'єктів за умови наявності викривлень їхнього відображення у просторі образів // 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 // Вісник Східноукраїнського національного університету імені Володимира Даля. – 2018. – № 6(247). – С. 49–55. *(Видання входить до міжнародної наукометричної бази: Index Copernicus)*

22. Захожай О. И., Шевцова А. С. Использование модели температурного режима в автоматизированной системе управления коксовой печью // Сборник научных трудов ДонГТУ. – 2011. – № 33. – С. 365–370.

23. Захожай О. І., Лифар В. О., Іванов В. Г. Консолідація даних та прийняття рішень на основі ранжування груп ідентичних класифікацій в гібридних системах розпізнавання // Електротехнічні та комп'ютерні

системи. – 2019. – № 31(107) – С. 95–103. (Видання входить до міжнародних наукометричних баз: *Ulrich's Periodicals Directory (ULRICHSWEB), Electronic Journals Library, Google Scholar, ВИНИТИ*)

24. Захожай О. І. Визначення черги обробки даних при гібридному розпізнаванні образів // Вісник Східноукраїнського національного університету імені Володимира Даля. – 2019. – № 7(255). – С. 111–116. (Видання входить до міжнародної наукометричної бази: *Index Copernicus*)

25. Захожай О.І. Інформаційна технологія гібридного розпізнавання образів для обробки неоднорідних даних в складних системах // Вісник Східноукраїнського національного університету імені Володимира Даля. – 2019. – № 8(256). – С. 141–147. (Видання входить до міжнародної наукометричної бази: *Index Copernicus*)

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

27. Меняйленко О., Захожай О. Організація перевірки студентських і наукових робіт на плагіат // Науковий вісник Мелітопольського державного педагогічного університету. – 2017. – № 1(18). – С. 157–162.

*Наукові праці, які засвідчують апробацію матеріалів дисертації*

28. Захожай О. І. Використання системи технічного зору для управління гільйотинними ножицями металургійного виробництва // 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 // 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.

31. Захожай О. И., Шевцова А. С. Использование модели температурного режима коксования в автоматизированной системе управления коксовой печью // Труды 12-й международной научно-практической конференции «Современные информационные и электронные технологии – 2011». – Одесса: ОНПУ, 2011. – С. 127.

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

33. Zakhzhay O., Paerand Yu. Combined Systems of Patterns Recognition // Контроль і управління у складних системах (КУСС-2012). XI Міжнародна конференція. Тези доповідей. – Вінниця: ВНТУ, 2012. – С. 241.

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

35. Zakhzhay O. Human emotion models in automated data processing and control of complex systems // 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 // IX Всеукраїнська науково-практична конференція «Електроніка та телекомунікації-2019» – Сєвєродонецьк: СНУ, 2019. – С. 113–115.

*Наукові праці, які додатково відображають наукові результати дисертації*

37. Меньяйленко О. С., Захожай О. І. Спосіб контролю температури нагріву коксової печі. Патент на корисну модель № 91118 МПК (2014.01) C10B 17/00, C10B 23/00; опубл. 25.06.2014, бюл. №9.

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.

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.

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

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.

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.