

Девіз «ГІБРИДИЗАЦІЯ»

Науковий напрям
«Інформатика, обчислювальна техніка та автоматизація»

**гібридна інтелектуальна система керування
вирощуванням сцинтиляційних монокристалів**

ЗМІСТ

ВСТУП.....	3
1 АНАЛІЗ ПРОБЛЕМИ ТА ПОСТАНОВКА ЗАДАЧІ	4
1.1 Аналітичний огляд методів автоматизації вирощування монокристалів. ..	4
1.2 Аналітичний огляд методів селекції ознак розпізнавання.	8
1.3 Постановка задачі аналізу та синтезу гібридної інтелектуальної.....	9
системи автоматизації вирощування сцинтиляційних монокристалів.	
2 ОПИС МЕТОДУ ДОСЛІДЖЕННЯ	11
2.1 Математична модель навчання інтелектуальної гібридної системи	11
автоматизації вирощування сцинтиляційних монокристалів	
2.2 Базовий алгоритм навчання системи розпізнавання	13
2.3 Алгоритм функціонування СППР у режимі екзамену	15
3 ІНФОРМАЦІЙНЕ ТА ПРОГРАМНЕ ЗАБЕЗПЕЧЕННЯ ГІБРИДНОЇ ІНТЕЛЕКТУАЛЬНОЇ СИСТЕМИ КЕРУВАННЯ ТЕХНОЛОГІЧНИМ ПРОЦЕСОМ ВИРОЩУВАННЯ СЦИНТИЛЯЦІЙНИХ МОНОКРИСТАЛІВ	17
3.1 Вхідний математичний опис.....	17
3.2 Гібридний алгоритм селекції словника ознак.	18
3.3 Короткий опис програмного забезпечення	23
3.4 Результати фізичного моделювання.....	25
ВИСНОВКИ	29
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ ІНФОРМАЦІЇ	30

ВСТУП

Для вирішення сучасних завдань приладобудування, екологічного захисту навколишнього середовища від різного роду випромінювань, впровадження нових ідей і напрямів у галузі фізики високих енергій, медицини, астрофізики, космонавтики характерно широке використання сцинтиляційних монокристалів (МК). Перспективи отримання високоякісних великогабаритних СМК, що користуються підвищеним попитом на світовому ринку, пов'язані з керованим вирощуванням монокристалів за методом Чохральського. Вирощування бездефектних МК високої оптичної та сцинтиляційної якості забезпечується не тільки рішенням широкого спектра задач фізико-хімічних досліджень та вдосконалення на цій основі технологічної оснастки ростових процесів, а й розвитком принципів їх автоматизації. Тому тема конкурсної наукової роботи є актуальною, оскільки вона присвячена підвищенню функціональної ефективності системи керування вирощуванням монокристалів.

Метою наукової роботи є розроблення інформаційного та програмного забезпечення інтелектуальної гібридної системи класифікаційного керування технологічним процесом вирощування крупно габаритних лужно-галоїдних сцинтиляційних монокристалів.

У роботі в рамках прогресивної інформаційно-екстремальної інтелектуальної технології аналізу і синтезу інтелектуальних систем на основі машинного навчання та розпізнавання образів розроблено інформаційно-екстремальний метод навчання гібридної системи керування технологічним процесом вирощування монокристалів за методом Чохральського, який дозволяє здійснювати оптимізацію словника ознак розпізнавання з використанням методу групового врахування аргументна (МГВА).

1 АНАЛІЗ ПРОБЛЕМИ ТА ПОСТАНОВКА ЗАДАЧІ

1.1 Аналітичний огляд методів автоматизації вирощування монокристалів.

Проведені протягом кількох десятиліть в НТК «Інститут монокристалів» (м. Харків) дослідження сцинтиляторів підтверджують, що досягнуті значення їх показників якості не є оптимальними і можуть бути істотно поліпшені, в тому числі, і шляхом впровадження нових автоматизованих технологій [1-5]. Широкий огляд методів та засобів контролю, регулювання та моделювання процесів вирощування МК в період до початку дев'яностих років минулого століття представлений в працях [6,7]. Автори оглядової інформації [6], прийшли до висновку, що автоматизація процесів вирощування МК у цей період здійснювалася в двох напрямках. Перший напрямок стосується підготовчих операцій (підготовки і завантаження сировини в ростові апарати тощо) та операцій знімання і механічної обробки МК. Другий напрямок - автоматизація власне процесу вирощування МК.

Роботи з удосконалення технологічних процесів (ТП) вирощування МК, їх фізико-математичного моделювання, формалізації задач керування цими складними динамічними системами здійснені під керівництвом Л.Г. Ейдельмана, Б.В. Гриньова, В.І. Горілецького та інших дослідників. Перспективи розвитку систем керування для методу Чохральського автори праці [7] пов'язують із забезпеченням функцій контролю та регулювання таких параметрів, які досить повно характеризують якість кристалу, що вирощується в режимі реального часу [8]. Історично можна виділити чотири періоди автоматизації ростових технологічних систем у сфері вирощування МК. На початковому етапі (1950-1970 р.) використовувалися аналогові засоби регулювання [6,7], які не відповідали вимогам мобільності, гнучкості, резервування, перебудови алгоритмів управління. При цьому основою управління було програмне регулювання температури.

Для другого етапу (70-80-ті роки минулого століття) характерно поступове впровадження мікропроцесорів і вітчизняних мікро-ЕОМ (ПЕОМ "ДЗ-28", "ПК-01", "Партнер", "Искра") як найпростіших керуючих пристроїв і блоків знімання інформації для обґрунтованого прийняття рішень. Завдання моделювання вирішувалися на великих ЕОМ, які перебували поза виробничих приміщень. Теоретичні розробки стосувалися досить складних алгоритмів керування з урахуванням все більшої кількості факторів, що впливають на ТП вирощування МК з розплаву [8]. Застосування стандартних модулів типу КТЗ ЛУС, КАМАК, керуючих ЕОМ серії М-6000 дозволяло досить гнучко перебудовувати алгоритми управління. Поступове оснащення на третьому етапі (80-90-і роки) лабораторних, а потім і промислових установок (наприклад, типу "КРИСТАЛ") персональними комп'ютерами ІВМ РС істотно змінило технічне та програмне забезпечення систем автоматизації, В Інституті монокристалів (м. Харків) у цей час був впроваджений перший варіант установки "РОСТ-5", оснащеної керуючою мікро-ЕОМ "Електроніка-60" (Україна, м. Луганськ). Четвертий сучасний етап відрізняється пошуком шляхів побудови математичного забезпечення СК, виходячи з найбільш об'єктивних уявлень про реальні фізико-технічних обмеження для моделей об'єкта керування, що є основою побудови адаптивних, робастних, ненадлишкових, "енергетично" ефективних алгоритмів керування [8]. Для технічного забезпечення СК характерно широке використання мікроконтролерних засобів та мережевого обладнання [4]. Паралельно процесу впровадження нових технічних засобів автоматизації ТП вирощування МК уточнюються алгоритми керування, виходячи з найбільш загальних припущень про об'єкт керування щодо розподіленості параметрів, нестационарності, стохастичності, динамічності, дискретності, нелінійності, взаємозв'язку характеристик). Нові алгоритми і методи синтезу мікроконтролерних пристроїв управління стали основою розробки комплексу апаратно-програмних засобів для СК вирощування великогабаритних СМК.

На рис. 1.1 приведено структурну схему керування процесом вирощування лужно-галоїдних крупно габаритних монокристалів за методом Чохральського, що впроваджена на установці промислового вирощування монокристалів типу «РОСТ» (НТК «Інститут сцинтиляційних матеріалів» НАН України, м. Харків).

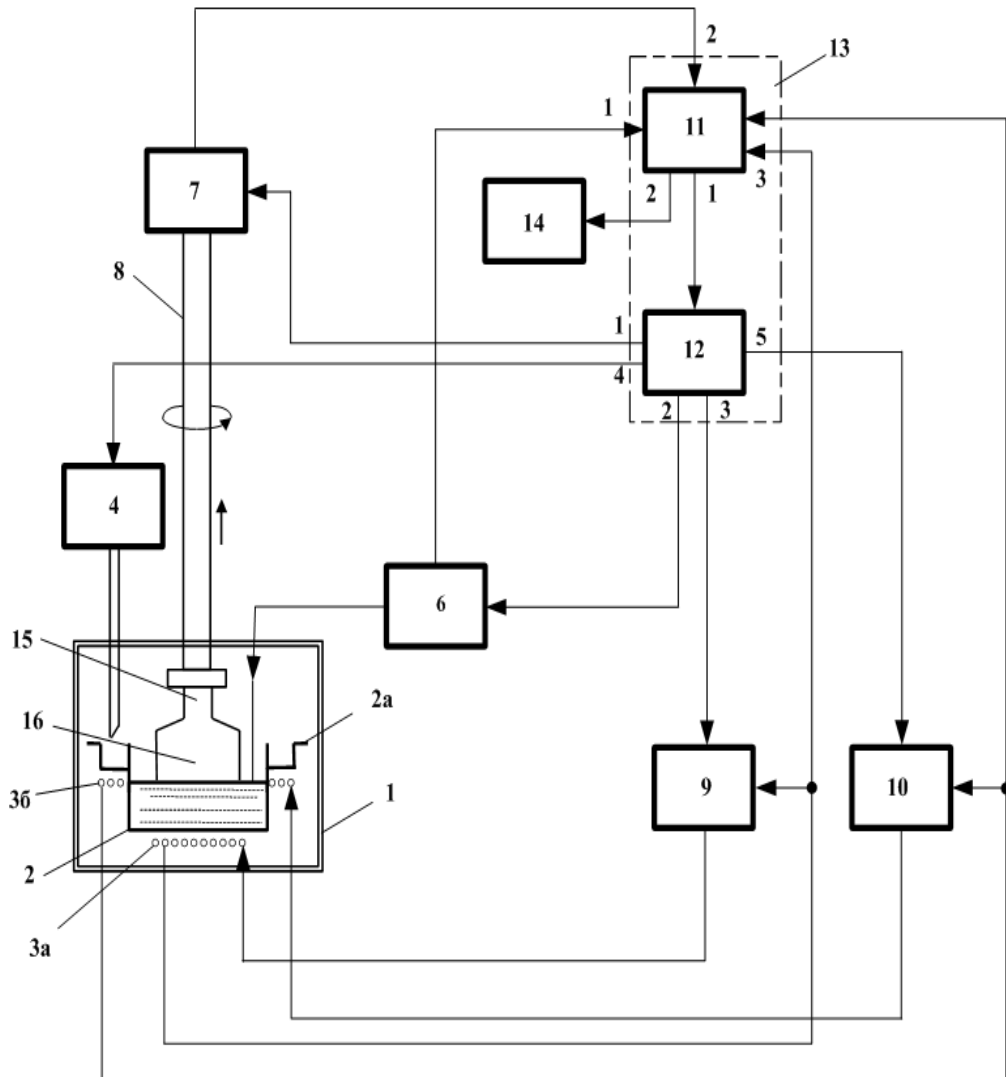


Рисунок 1.2 – Структурна схема СУ вирощування МК: 1 - ростова піч, 2 - тигель з кільцевою порожниною 2а для плавлення сировини, 3а - донний і 3б - бічний нагрівачі, 4 - підживлювач з транспортною трубкою 5, 6 - датчик рівня розплаву, 7 - блок управління переміщенням 8 кристалотримача, 9, 10 - блоки корекції температури донного і бічного нагрівачів відповідно, 11 - блок розрахунку параметрів і 12 - блок управління регулятора 13 діаметра МК, 14 - пристрій відображення даних, 15 - затравка вирощуваного МК 16

Одним із перспективних напрямків підвищення функціональної ефективності є застосування ідей і методів інформаційно-екстремальної інтелектуальної технології аналізу і синтезу систем керування на основі машинного навчання і розпізнавання образів [9,10].

На рис. 1.2 показано узагальнену функціональну схему здатної навчатися інтелектуальної системи керування (ІСК), Тут функціональні блоки, що знаходяться зверху пунктирної лінії відносяться до системи підтримки прийняття рішень (СППР), яка конструктивно являє собою спеціалізований управляючий електронно-обчислювальний комплекс.

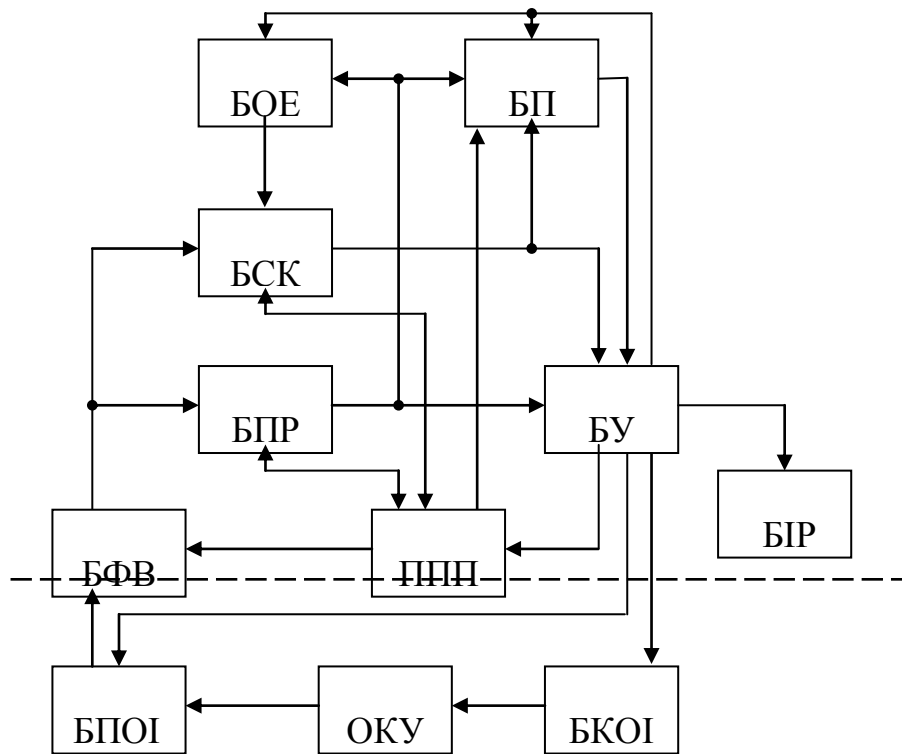


Рисунок 1.2 – Узагальнена функціональна схема ІСК, що навчається

Наведена структура дозволяє функціонувати системі як у режимі навчання, так і в режимі екзамену. Робота системи в режимі навчання регламентується програмно блоком управління (БУ), який при кожному випробуванні видає команду на блок кінцевої обробки інформації (БКОІ) для вироблення ініціюючої дії на виконавчі органи об'єкту контролю та управління (ОКУ). Блок первинної обробки інформації (БПОІ) перетворює інформацію про фун-

кціональний стан ОКУ до вигляду, зручного для подальшого оброблення ЕОМ, і передає її у блок формування векторів розпізнавання (БФВР).

Із пристрою постійної пам'яті (ППП) у БФВР за командами БУ надходять значення контрольних допусків на ознаки розпізнавання, які порівнюються із вимірними значеннями відповідних ознак. У результаті кожного випробування на виході БФВР формується структурований випадковий двійковий вектор-реалізація образу, координати якого є предикатами, що приймають значення "1", якщо значення відповідної ознаки знаходиться в полі контрольних допусків, і "0", якщо – не знаходиться. Вектор розпізнавання, який є реалізацією образу, надходить у блок прийняття рішень (БПР), де визначається його належність деякому класу розпізнавання, і в блок статистичної корекції (БСК), який визначає оптимальні значення параметрів функціонування ІСУ шляхом пошуку максимуму КФЕ навчання системи розпізнавати реалізації образу. КФЕ обчислюється блоком оцінки ефективності (БОЕ).

Режим екзамену виконується безпосередньо в робочому режимі функціонування системи. При цьому БСК і БОЕ не функціонують, але може підключатися блок прогнозування (БП), призначений для прогнозування зміни функціональної ефективності та надійності ІСУ. З цією метою в БП із ППП за командою БУ надходять значення одномірних статистик навчальних матриць, які порівнюються з поточними статистиками екзаменаційної матриці. Для спостереження і реєстрації інформації служить блок індикації і реєстрації (БІР), який складається із стандартних пристроїв виводу та відображення інформації. Таким чином, розглянута структура ІСУ характеризується широкими функціональними можливостями і дозволяє розв'язувати задачі навчання і екзамену, включаючи задачу прогностичної класифікації.

1.2 Аналітичний огляд методів селекції ознак розпізнавання.

Задача селекції ознак розпізнавання (ОР) (feature selection) полягає в доборі підмножини ознак (feature subset) такої, яка дозволяє максимально розділити класи розпізнавання, що характеризують функціональні стани ТП.

Така процедура дає змогу не тільки зменшити обчислювальні витрати необхідні для розпізнавання, скорочуючи кількість ОР, реалізації яких необхідно встановити, але забезпечує підвищену точність класифікації завдяки обмеженому обсягу вибірки. Більш загальні методи створення нових ОР за допомогою зміни або комбінування ОР вихідного словника мають назву виділення ОР (feature extraction).

Актуальність розвитку методів оптимізації словника ОР пов'язана з використанням значної кількості ОР у наступних видах задач:

- системи, що використовують комбіновані дані, отримані з різномірних сенсорів.;
- інтегрування різномірних моделей, коли параметри декількох математичних моделей об'єднуються з метою класифікації;
- системи інформаційної проходки, мета яких встановити приховані взаємозв'язки серед значної кількості ОР.

1.3 Постановка задачі аналізу та синтезу гібридної інтелектуальної системи автоматизації вирощування сцинтиляційних монокристалів.

Нехай відомі функціональні стани СПР, що задають алфавіт класів розпізнавання $\{X_m^0\}$, $m = \overline{1, M}$. Дано навчальну матрицю типу об'єкт-властивість

$$\|y_{m,i}^{(j)}\|, \quad i = \overline{1, N}, \quad j = \overline{1, n}, \quad (1.3.1)$$

де N, n - кількість ознак розпізнавання і реалізацій образу одного класу.

Визначено структурований вектор параметрів функціонування СПР $g = \langle g_1, \dots, g_\xi, \dots, g_\Xi \rangle$ з відповідними обмеженнями на ці параметри $R_\xi(g_1, \dots, g_z) \leq 0$.

Необхідно в процесі навчання побудувати оптимальне в інформаційно-му розумінні розбиття простору ознак розпізнавання на класи за умови, що інформаційний критерій функціональної ефективності (КФЕ) навчання системи набуває максимального значення в робочій області визначення його функції

$$E^* = \max_{\{m\}} E_m,$$

де E_m – інформаційний критерій функціональної ефективності навчання системи розпізнавати реалізації класу X_m^0 .

На етапі екзамену необхідно визначити з наближеною до асимптотичної повною достовірністю належність реалізації образу, що розпізнається, до одного із класів із сформованого на етапі навчання алфавіту.

Об'єктом дослідження є слабо формалізований процес вирощування сцинтиляційних монокристалів

Предметом дослідження є гібридні алгоритми та математичні моделі селекції словника ознак інтелектуальної системи автоматизації вирощування монокристалів у рамках інформаційно-екстремальної інтелектуальної технології (ІЕІТ).

У роботі необхідно розробити інформаційно-екстремальний метод синтезу гібридної системи автоматизації вирощування монокристалів.

Для досягнення мети необхідно виконати такі завдання:

- 1) сформулювати математичну модель оптимізації словника ознак на основі базової моделі навчання СППР за ІЕІТ;
- 2) розробити та програмно реалізувати алгоритм класифікаційного керування вирощуванням монокристалів;
- 3) перевірити функціональну ефективність гібридної інтелектуальної системи на реальних даних, отриманих в результаті моніторингу технологічного процесу вирощування сцинтиляційних монокристалів.

2 АНАЛІЗ І СИНТЕЗ ІНТЕЛЕКТУАЛЬНОЇ СИСТЕМИ КЕРУВАННЯ

2.1 Математична модель навчання інтелектуальної системи керування, що навчається

Основним функціональним блоком інтелектуальної системи керування технологічним процесом вирощування монокристалів є СППР, що навчається.

Відповідно до предметної галузі та постановки задачі, допускається перетин між класами розпізнавань, що утворюють алфавіт класів розпізнавання. Обмежимося в рамках даної роботи побудовою однорівневих вирішальних правил. Введемо узагальнений інформаційний критерій функціональної ефективності (КФЕ) \mathring{A}_δ , який є функціоналом від точнісних характеристик процесу навчання СК. Його значення оцінимо середнім арифметичним значень КФЕ для всіх класів розпізнавання. Також введемо параметр оптимальності навчальної матриці при додаванні нової ознаки в словник у вигляді

$$\overline{E}_i = \frac{1}{m} \sum_{j=1}^m E_{j\text{Prov}} * \frac{1}{m} \sum_{j=1}^m E_{j\text{Test}}, \quad (2.1.1)$$

де $\mathring{A}_{j\text{Prov}}$ -значення усередненого КФЕ, отриманого при синтезі вирішальних правил за перевіркою матрицею, $\mathring{A}_{j\text{Test}}$ -значення усередненого КФЕ, отриманого при синтезі вирішальних правил за тестовою матрицею.

Категорійну модель навчання СППР у вигляді діаграми відображень множин в процесі навчання СК показано на рис. 2.1.

Кожен параметр навчання в діаграмі відображення множин (рис. 2.1) має свій контур оптимізації при виконанні умови повної композиції: термножина E , елементи якої складають значення інформаційного КФЕ, є загальною для всіх контурів оптимізації.

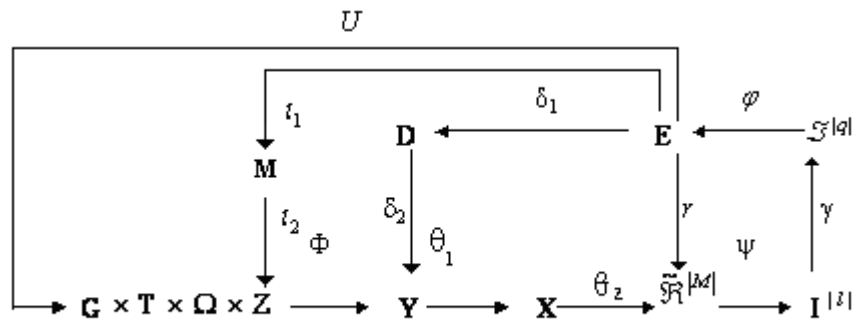


Рисунок 2.1 – Категорійна модель навчання СК

Серед параметрів навчання, які суттєво впливають на достовірність класифікатора, розглянемо поля контрольних допусків $\{\delta_{K,i} \mid i = \overline{1, N}\}$ на значення ознак розпізнавання. Нехай контур оптимізації СКД, яка утворює множину D , замикається послідовно оператором $\delta_1: E \rightarrow D$ і оператором $\delta_2: D \rightarrow Y$, який змінює реалізації образу в процесі максимізації критерію $E_m \in E$. Оператор ι_1 реалізує синтез нових ознак на основі опорних функцій, множина M містить значення мультикритерію функціональної ефективності, а оператор ι_2 змінює вихідний словник ОР згенерованими ознаками та поповнює словник на одну оптимальну ознаку за алгоритмом селективної оптимізації словника ознак.

У діаграмі (рис. 2.1) контур, показаний на рис. 2.2, оптимізує словник ознак розпізнавання.

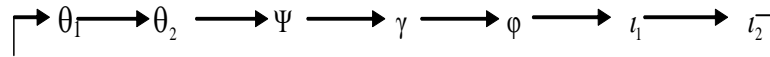


Рисунок 2.2 – Контур оптимізації словника ознак розпізнавання

Контур, показаний на рис. 2.3, оптимізує систему контрольних допусків (СКД) на ознаки розпізнавання.

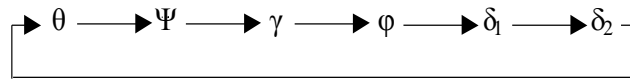


Рисунок 2.3 – Контур оптимізації СКД на ознаки розпізнавання

Згідно з принципом відкладених рішень [15] категорійна модель може мати додаткові контури оптимізації, які дозволять при необхідності побудувати на етапі навчання безпомилкові за навчальною матрицею вирішальні правила.

2.2 Базовий алгоритм навчання СППР

Призначенням базового алгоритму навчання [13] є оптимізація геометричних параметрів контейнерів класів розпізнавання, які відновлюються на кожному кроці навчання в радіальному базисі простору ознак розпізнавання. Вхідною інформацією для навчання за базовим алгоритмом у загальному випадку є дійсний масив реалізацій образу $\{y_m^{(j)} \mid m = \overline{1, M}; j = \overline{1, n}\}$; система полів контрольних допусків $\{\delta_{K,i}\}$ і рівні селекції $\{\rho_m\}$ координат двійкових еталонних векторів-реалізацій образу, які за замовчанням дорівнюють 0,5 для всіх класів розпізнавання.

Розглянемо основні етапи реалізації алгоритму:

1) формування бінарної навчальної матриці $\|x_{m,i}^{(j)}\|$, елементи якої дорівнюють

$$x_{m,i}^{(j)} = \begin{cases} 1, & \text{if } y_{m,i}^{(j)} \in \delta_{K,i}, \\ 0, & \text{if } y_{m,i}^{(j)} \notin \delta_{K,i}. \end{cases} \quad (2.2.1)$$

2) формування масиву еталонних двійкових векторів $\{x_{m,i} \mid m = \overline{1, M}, i = \overline{1, N}\}$, елементи якого визначаються за правилом

$$x_{m,i} = \begin{cases} 1, & \text{if } \frac{1}{n} \sum_{j=1}^n x_{m,i}^{(j)} > \rho_m, \\ 0, & \text{if } \text{else}, \end{cases} \quad (2.2.2)$$

де ρ_m – рівень селекції координат вектора $x_m \in X_m^o$.

3) розбиття множини еталонних векторів на пари найближчих "сусідів": $\mathfrak{R}_m^{[2]} = \langle x_m, x_l \rangle$, де x_l – еталонний вектор сусіднього класу X_l^o , здійснюється за таким алгоритмом:

а) структурується множина еталонних векторів, починаючи з вектора x_1 базового класу X_1^o , який характеризує найбільшу функціональну ефективність СПР;

б) будується матриця кодових відстаней між еталонними векторами розмірності $M \times M$;

в) для кожного рядка матриці кодових відстаней знаходиться мінімальний елемент, який належить стовпчику вектора, найближчого до вектора, що визначає строку. За наявності декількох однакових мінімальних елементів вибирається з них будь-який, оскільки вони є рівноправними;

г) формується структурована множина елементів попарного розбиття $\{\mathfrak{R}_m^{[2]} \mid m = \overline{1, M}\}$, яка задає план навчання;

4) здійснюється пошук максимуму КФЕ в робочій області його визначення: $E_m^* = \max_{\{d\}} E_m$, де $\{d\} = \{0, 1, \dots, d < d(x_m \oplus x_l)\}$ – множина радіусів

концентрованих гіперсфер, центр яких визначається вершиною $x_m \in X_m^o$.

5) визначається оптимальна кодова відстань – радіус контейнера, що відновлюється в радіальному базисі бінарного простору ознак розпізнавання за процедурою

$$d_m^* = \arg \max_{\{d\}} E_m^*.$$

При цьому $E_m(0) = 0$.

Таким чином, алгоритм навчання є ітераційною процедурою пошуку глобального максимуму інформаційного КФЕ в робочій області визначення його функції.

2.3 Алгоритм функціонування СППР у режимі екзамену

Параметри навчання СППР є вхідними даними алгоритму екзамену, тобто безпосереднього розпізнавання в робочому режимі СППР.

На рис. 2.4 показано структурну схему алгоритму екзамену.

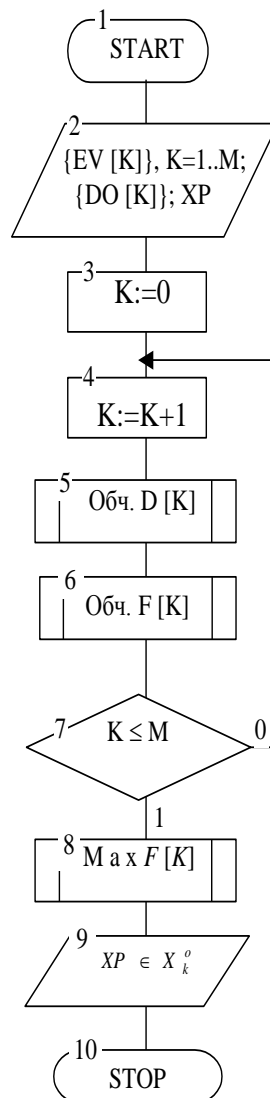


Рисунок 2.4– Структурна схема алгоритму екзамену

Алгоритм має такі вхідні дані: $\{EV[K]\}$ – масив еталонних двійкових векторів: $K = \overline{1, M}$ – змінна числа класів розпізнавання; $\{DOPT[K]\}$ – цілий масив оптимальних радіусів контейнерів класів розпізнавання у кодовій відстані Хеммінга; XP – двійкова реалізація образу, що розпізнається. Виходом алгоритму є повідомлення про належність реалізації, що розпізнається, деякому класу X_k^o із сформованого на етапі навчання алфавіту класів $\{X_m^o\}, k, m = \overline{1, M}$.

На рис.2.4 блок 5 обчислює, починаючи з базового класу, кодову відстань $d\{EV[K] \oplus XP\}$ між поточним еталонним вектором і реалізацією XP . Блок 6 для кожного класу обчислює значення функції належності $F[K]$, яка для гіперсферичного класифікатора має вигляд

$$F[K] = 1 - \frac{D[K]}{DOPT[K]}. \quad (2.3.1)$$

Після виходу із циклу блок 8 визначає клас, до якого належить реалізація XP за максимальним значенням функції належності (2.21).

3 ІНФОРМАЦІЙНЕ ТА ПРОГРАМНЕ ЗАБЕЗПЕЧЕННЯ ГІБРИДНОЇ ІНТЕЛЕКТУАЛЬНОЇ СИСТЕМИ КЕРУВАННЯ ТЕХНОЛОГІЧНИМ ПРОЦЕСОМ ВИРОЩУВАННЯ СЦИНТИЛЯЦІЙНИХ МОНОКРИСТАЛІВ

3.1 Вхідний математичний опис

Оскільки функціональним призначенням СППР є рекомендаційний режим, та на підставі аналізу збоїв обладнання та даних функціональної діагностики, то важливою для оператора є інформація щодо тенденції зміни діаметру монокристалу в процесі його вирощування, яка визначає вихідні властивості продукту. Тому то за мету функціонування СППР взято класифікаційне визначення напрямку зміни росту діаметра монокристалу на рівні мениску розплаву, що дозволить контролювати параметри регуляторів підсистем стабілізації росту та не допустити пере регулювання таких критичних параметрів, як температура бокового нагрівача тощо.

Алфавіт класів розпізнавання складався з двох класів, що характеризували функціональні стани технологічного процесу вирощування монокристалу: клас X_1^o – відхилення діаметра монокристалу в межах технологічного режиму і клас X_2^o – відхилення діаметра монокристалу перевищують допустимі.

Враховуючи нестационарність та часову параметричність тепломасообмінних та кристалізаційних процесів, квазістационарним періодом обрано проміжок часу довжиною в 1 день, статистику по котрому було відсортовано по значенню діаметру монокристалу та сформовано вектори-реалізації двох класів розпізнавання.

Навчальну матрицю побудовано за реальними даними, які було отримано в результаті моніторингу процесу вирощування крупногабаритного лужно-галоїдного монокристалу Cz I (цезій йод) на установці типу «РОСТ» технологічного парку НТК «Інститут монокристалів» (м. Харків). Навчальна матриця для кожного класу складалася із 40 векторів-реалізацій

Структурована реалізація функціонального стану технологічного процесу складалася із значень 25 ознак розпізнавання, що фіксують технологічні характеристики процесу вирощування, відповідно до праці [12].

3.2 Гібридний алгоритм селекції словника ознак.

Процедура ітераційної оптимізації параметрів навчання у рамках ІЕІ-технології спрямована на побудову в радіальному базисі оптимальних контейнерів класів розпізнавання, які забезпечують максимальну асимптотичну достовірність класифікатора в режимі екзамену.

В практичних задачах керування мають місце такі умови,

- 1) прийняття гіпотези нечіткої компактності, тобто апріорно допускається перетин множин ознак для різних функціональних станів;
- 2) обмеженість словника ознак розпізнавання, зумовлена особливостями технологічного процесу, технічними складнощами, тощо;
- 3) взаємна кореляція між групами ознак, що впливає на інформативність окремих ознак, інтерпретованість та в загальному випадку достовірність функціонування інтелектуальної системи керування, що працює в класифікаційному режимі, котра базується на теорії розпізнавання образів та машинному навчанні.

Таким чином, мають місце не лише взаємкореляції в рамках одного класу розпізнавання, а й кореляції окремих ознак з сусідніх функціональних станах, наслідком котрої і є нечітка компактність та відносно низька інформативність вхідного математичного опису у вигляді «об'єкт-властивість».

Задача селекції, або ж оптимізації словника ознак, є важливим питанням, здатним не лише введенням додаткових ознак, суперпозицій на основі існуючих, вирішити так звану проблему багато вимірності, але і інтерпретувати ступінь інформаційної складової у взаємодії факторів, що описують ознаки.

У рамках роботи розглянуто алгоритм введення додаткових інформативних ознак до існуючого словника ознак розпізнавання.

Гібридний алгоритм поєднує в собі принципи метода групового врахування аргументів та інформаційно-екстремальної інтелектуальної технології. Інструментом генерації додаткових ознак виступає МГВА, в той час як в рамках ІЕІ-технології відбувається оцінка інформативності нових ознак та синтез вирішального правила для функціонування інтелектуальної системи автоматизації вирощування сцинтиляційних монокристалів .

Розглянемо основні принципи методу:

1. Принцип неповноти навчальної вибірки, на основі теореми Геделя, фактично, стверджує про обмеженість навчальної вибірки по відношенню до генеральної сукупності випробувань.

За принципами МГУА генерація додаткових ознак має селекційний (еволюційний) характер, при цьому нові ознаки генеруються як результат застосування функцій певного виду , аргументами від котрих є попередні результати селекції або елементи вихідного словника ознак. Оскільки навчальна матриця є обмеженою, то використання функцій певного виду може гарантувати результат (інформативність) лише для обмеженої навчальною вибіркою кількості реалізацій, проте мати можливі неадекватні результати по відношенню до інших реалізацій з генеральної сукупності. Такими чином, введення перевіркової та тестової навчальних матриць та відповідного мультикритерію функціональної ефективності дозволить знайти найоптимальніші ознаки, здатні якнайкраще компенсувати проблему неповноти та гарантувати більш високу достовірність правильного розпізнавання на етапі екзамену (функціонування СППР).

2. Принцип зворотного зв'язку при побудові оптимальної моделі.

Відповідно до МГВА на кожному кроці будується певна кількість нових моделей на основі заданих аргументів невідомої моделі, причому за певним критерієм оптимальності відбираються ті з них, що найкраще описують шукану модель. Принцип зворотнього зв'язку дозволяє на наступному рівні

будувати нові моделі на основі оптимальних, отриманих на попередньому кроці, а також додатково будувати комбінації моделей на основі суперпозиції оптимальних та решти моделей, отриманих на попередньому кроці. Такий підхід дозволяє виявити складні кореляції не лише між окремими ознаками, але й групами ознак, тобто будується ієрархічна багат шарова структура кореляційних залежностей, що у випадку оцінки інформаційним критерієм інтерпретуються як можливість знайти оптимальні інформативні ознаки, що описують сутність функціональних станів.

3. Вибір опорних функцій.

Відповідно до [3] теплофізичні моделі ростових процесів та тепломасообмінних процесів мають параметричний в часі періодичний характер. Тому в якості опорної функції для генерації нових інформативних ознак обрано тригонометричні функції $\sin(x)$ $\cos(x)$, що є опорними для ряду Фур'є, котрий широко використовується для опису періодичних моделей.

4. Мультикритерій функціональної ефективності.

На основі принципу 1 будуються тестова та перевірна навчальні матриці, причому тестова є частиною перевіркової. Згенеровані ознаки підставляються в тестову матрицю, після чого обирається та, що максимально підвищує різноманітність відносно тестової матриці класа-сусіда. далі оптимальна ознака підставляється в перевірку матрицю і після навчання розраховується глобальний максимум усередненого КФЕ. Мультикритерій у вигляді добутків оптимальних усереднених КФЕ по та перевіркової навчальних матрицях є інтегральним критерієм надійності та інформативності отриманої ознаки, поєднуючи фактично апріорну (генерація по тестовій матриці) та апостеріорну (підстановка в навчальну матрицю, по котрій не відбувався відбір ознаки) оцінки достовірності СППР.

Змінений алгоритм навчання має ітеративно-рекурсивний характер та складається з трьох контурів. Внутрішній контур передбачає оптимізацію кроку параметра поля системи контрольних допусків (СКД) на ОР з метою пошуку оптимального розбиття простору ознак на класи еквівалентності.

Проміжний контур передбачає генерацію моделей на кожному рівні селекції (додавання нової ознаки до навчальної матриці) оцінку інформативності та добір найоптимальнішої ознаки.

Зовнішній контур передбачає розрахунок мультикритерію функціональної та перевірку умови переходу на наступний рівень побудови шару моделей.

Розглянемо паралельний алгоритм навчання, котрий оптимізує параметри контейнерів класів розпізнавання за умови ітераційної процедури визначення для базового класу X_1^o оптимальних контрольних допусків на всі ознаки одночасно. Вхідні дані такі самі як і для базового алгоритму, але за область визначення параметра δ приймається інтервал $[1; \delta_H / 2]$, де δ_H – ширина нормованого поля допусків. Розглянемо кроки реалізації цього алгоритму:

- 1) обнулюється лічильник кроків зміни параметра δ : $l:=0$;
- 2) запускається лічильник: $l:=l+1$ і обчислюються нижні та верхні контрольні допуски для всіх ознак: $\{A_{HK,i}[l] := y_{m,i} - \delta[l]\}$ і $\{A_{BK,i}[l] := y_{m,i} + \delta[l]\}$, $i = \overline{1, N}$, відповідно;
- 3) реалізується базовий алгоритм навчання;
- 4) якщо $E_1^*[l] \geq E_1^*[l-1]$, то виконується пункт 5, інакше пункт 6.
- 5) якщо $\delta \leq \delta_H / 2$, то виконується пункт 2, інакше пункт 6.
- 6) $\{A_{HK,i}^* := A_{HK,i}[l-1]\}; \{A_{BK,i}^* := A_{BK,i}[l-1]\}, i = \overline{1, N}; E_1^* := E_1^*[l-1]$ і “ЗУПИН”.

Вхідними даними для гібридного алгоритму селекції словника ознак є початковий словник ознак розпізнавання $\{Z\}, z = \overline{1, h}$, де h -потужність алфавіту класів розпізнавання, опорні функції, початкове значення мультикритерію функціональної ефективності \bar{A}_0 , розраховане по вихідній

навчальній матриці без додавання нових ознак. Розглянемо основні етапи гібридного алгоритму навчання з оптимізацією словника ознак розпізнавання:

- 1) лічильник кроків оптимізації словника ознак встановлюється в нуль: $i := 0$;
- 2) генерується вектор $\sin \cos$, що містить значення тригонометричної функції, яку треба використати як опорну при генерації нової ознаки;
- 3) на основі вектора будуються $2z-3$ нові ознаки. З них $z-2$ – нові ознаки у вигляді суперпозицій на основі сусідніх існуючих ознак до останньої включно $(1-2, 2-3, 3-4, \dots, (n-2)-(n-1))$ та $z-1$ комбінації останньої ознаки з попередніми $(1-(n), 2-(n), \dots, (n-1)-(n))$;
- 4) нові ознаки по черзі підставляють ся в тестову навчальну матрицю, при цьому відбувається побудова вирішального правила, шукаються оптимальні параметри функціонування СППР та обчислюється усереднене по всім класам розпізнавання значення КФЕ;
- 5) із множини згенерованих ознак обирається така, при котрій усереднене значення КФЕ прийняло максимальне значення в робочій області визначення його функції;
- 6) оптимальна ознака в процесі навчання СППР підставляється в перевірочну матрицю і розраховується усереднене по класам значення КФЕ;
- 7) обчислені по тестовій та перевірочній матрицям усереднені значення КФЕ підставляються в формулу (2.1.1) для обчислення узагальненого КФЕ;
- 8) сума лічильника збільшується на одиницю: $i := i + 1$;
- 9) якщо $\bar{E}_i > \bar{E}_{i-1}$, тобто вдалося збільшити достовірність розпізнавання шляхом введення нової ознаки, то потужність словника ОР збільшується на 1, тобто до навчальної матриці додається нова згенерована ознака, то виконується пункт 10, інакше – пункт 8;

10) потужність вектора опорних функцій збільшується на 1 і здійснюється перехід на пункт 2 до тих пір, поки не буде побудовано безпомилкові за тестовою та перевіркою матрицями вирішальні правила.

3.3 Короткий опис програмної реалізації

У рамках роботи створено програмну реалізацію алгоритму, що складається з вікна з декількома вкладками, що розділюють функціональні елементи реалізації методу та відображають етапи навчання та оптимізації ознак розпізнавання.

На рис 3.1 зображено інтерфейс загрузочного вікна.

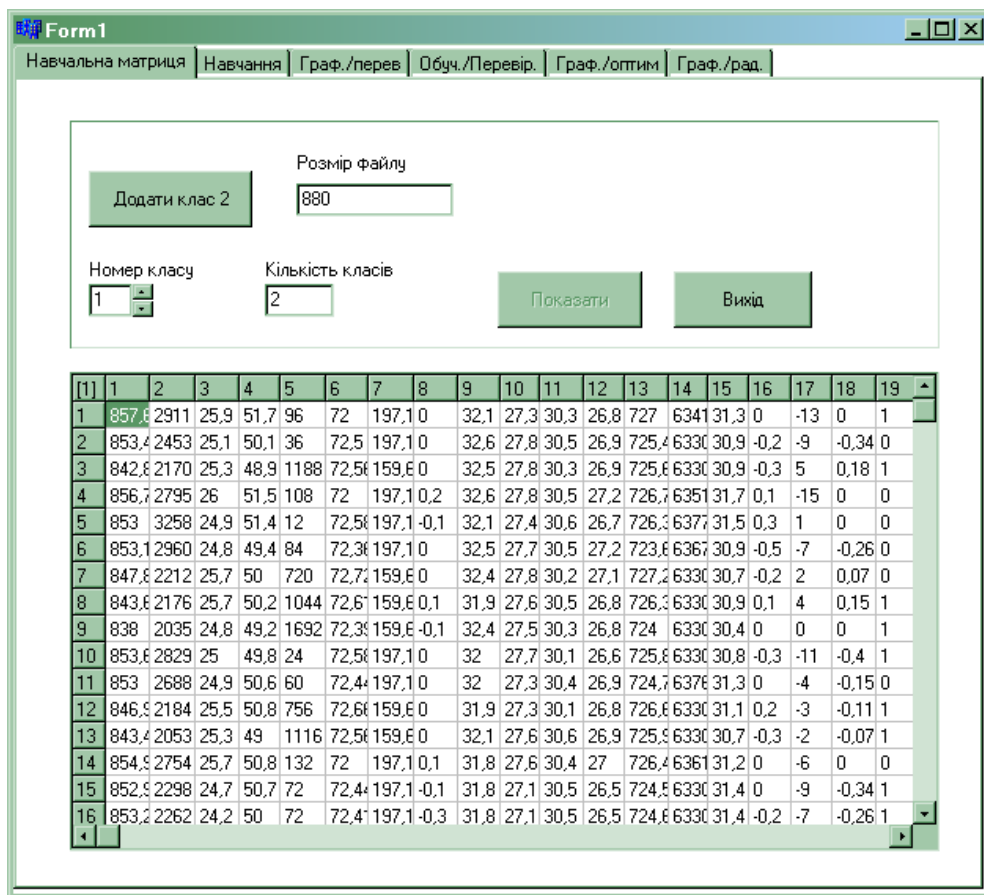


Рисунок 3.1 – Загрузочне вікно програмної реалізації.

Інтерфейс реалізовано у вигляді вкладок, кожна з яких має своє функціональне призначення. Перша вкладка реалізує ініціалізацію вхідного мате-

матичного опису, тобто додавання матриць для процесів навчання та оптимізації. Вона дозволяє зчитати з файлу необхідну інформацію, визначити розмір файлу, та кількість доданих матриць до програми. Наступна вкладка «Навчання» (рис.3.2) реалізує поділ матриць на тестові та навчальні для кожного класу розпізнавання та навчання матриць і визначення оптимального критерію. Вкладка «Навчання» дозволяє задати значення кроку квантування рівнів селекції ρ та оглянути оптимальне значення мультикритерію функціональної ефективності.

The screenshot shows a software window titled 'Form1' with several tabs: 'Навчальна матриця', 'Навчання', 'Граф./перев.', 'Обуч./Перевір.', 'Граф./оптим.', and 'Граф./рад.'. The 'Навчання' tab is active. On the left, there is a control panel with a 'Навчання' button and three input fields: 'Рівень селекції' (0.5), 'E'' (0.107131282115), and 'd' (69). The main area contains two tables: 'Test matrix' and 'Verification matrix'. Both tables have columns for 'delta', '0', '1', and 'E'' and rows numbered 0 to 13. The 'Test matrix' table shows zeros in all cells. The 'Verification matrix' table shows numerical values in the 'E'' column for rows 2 through 13.

delta	0	1	E'
0	0	0	0
1	0	0	0
2	0	0	0
3	0	0	0
4	0	0	0
5	0	0	0
6	0	0	0
7	0	0	0
8	0	0	0
9	0	0	0
10	0	0	0
11	0	0	0
12	0	0	0
13	0	0	0

delta	0	1	E'
0	0	0	0
1	0	0	0
2	0	0.01675	0.00839
3	0	0.01133	0.00566
4	0	0	0
5	0	0	0
6	0	0	0
7	0	0	0
8	0	0.02225	0.01112
9	0	0.00411	0.00206
10	0	0.07476	0.03738
11	0	0.07476	0.03738
12	0	0.02258	0.01129
13	0	0.00722	0.00361

Рисунок 3.2 – Вікно навчання початкових тестових та навчальних матриць

Дві наступні вкладки дозволяють оглянути графіки, які отримані в результаті навчання та оптимізації і остання вкладка «Обуч./Перевір.», показана на рис. 3.3, дозволяє провести генерацію ознак, додавання їх до матриць навчальної та перевіркової, навчання та виведення оптимального мультикритерію.

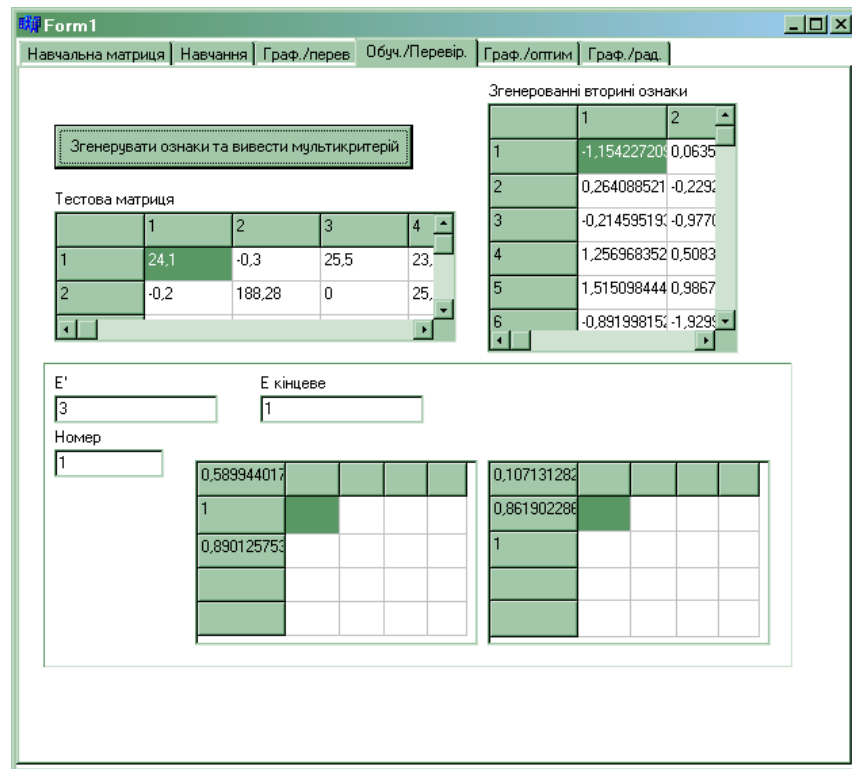


Рисунок 3.3 – Вікно генерації ознак

Програмну реалізацію виконано в середовищі розробки C++Builder6. При моделюванні алгоритму було використано принципи об'єктно-орієнтованого програмування, як поля використано параметри, які необхідні для функціонування алгоритму, а як методи – етапи навчання, оптимізації, генерації та інші допоміжні методи.

3.4 Результати фізичного моделювання

Результатом навчання СППР є однорівневі (неієрархічні) вирішальні правила.

Як критерій оптимізації параметрів функціонування СППР розглянуто модифіковану ентропійну міру Шеннона, при цьому в контурі оптимізації параметра дельта системи контрольних допусків на ознаки розпізнавання було використано узагальнений КФЕ

$$E_{\delta} = \frac{1}{m} \sum_{i=1}^m E_i, \quad (3.4.1)$$

де m - потужність алфавіту класів розпізнавання, $E_i, i = \overline{1, m}$ - значення максимального в робочій області значення інформаційного критерію для i - того класу.

В контурі оптимізації словника ознак як критерій оптимальності розглянуто мультикритерій (2.1.1).

Початковим кроком гібридного алгоритму оптимізації словника ознак є визначення мультикритерію (3.6) по навчальній матриці без додавання нових ознак. З метою підвищення достовірності функціонування в контур оптимізації словника ознак введено додатковий контур оптимізації параметра дельта за паралельним алгоритмом оптимізації СКД на ОР. Залежність параметра дельта від усередненого ентропійного КФЕ при навчанні за перевіркою матрицею показано на рис. 3.4

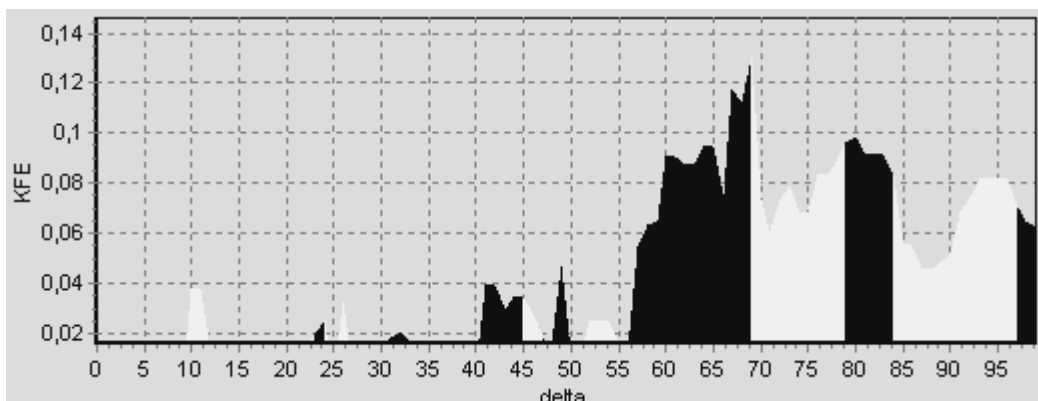


Рисунок 3.4 – Графік залежності усередненого КФЕ від параметра $delta$

На рис. 3.4 білим кольором позначено робочу область визначення КФЕ, а чорним кольором показано область, де значення першої та другої достовірності менше за 0,5. Аналіз рис. 3.4 показує, що при про максимальному значенні усередненого КФЕ ($\bar{E} = 0,13$) оптимальний параметр поля контрольних допусків дорівнює $delta^* = 69$ (у відносних одиницях поля контрольних допусків). Відповідно до визначеного алгоритму селекції словника ознак розпізнавання залежність значення мультикритерію (2.1.1) від кроку роботи алгоритму показано на рис. 3.6

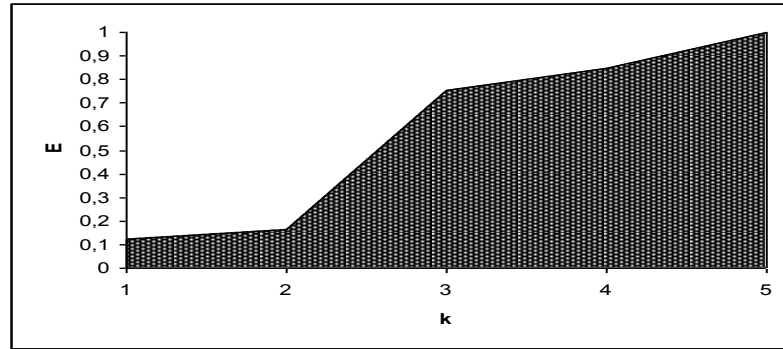


Рисунок 3.6 – Залежність узагальненого КФЕ від кроків оптимізації словника ознак

Аналізуючи рис. 3.6 можна зробити висновок, що після 4-х ітерацій роботи алгоритму (додавання 4-х інформативних ознак), було отримано безпомилковий за навчальною матрицею класифікатор і для тестової, і для перевіркової навчальних матриць, оскільки нормований узагальнений КФЕ (2.1.1) досягає свого максимального граничного значення.

На рисунку 3.7 зображено графік залежності параметра *delta* для перевіркової матриці за оптимальним словником ознак.

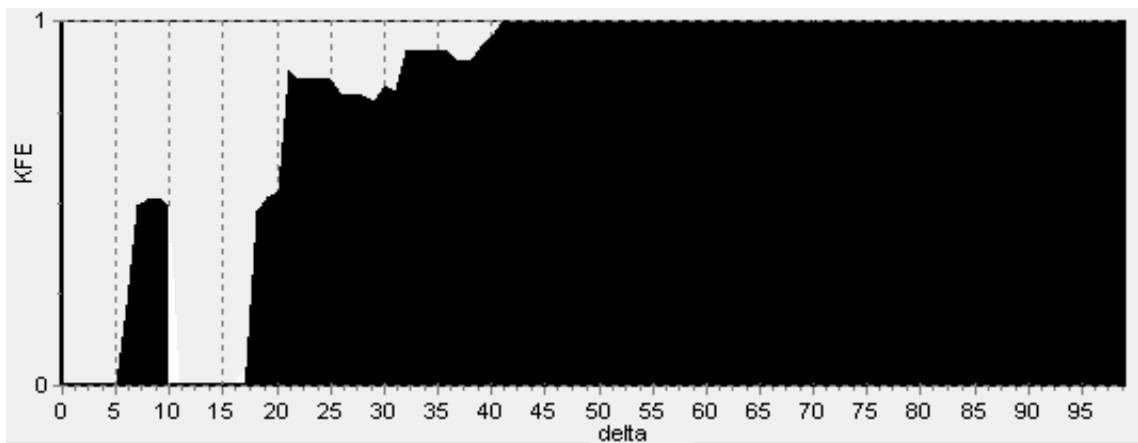
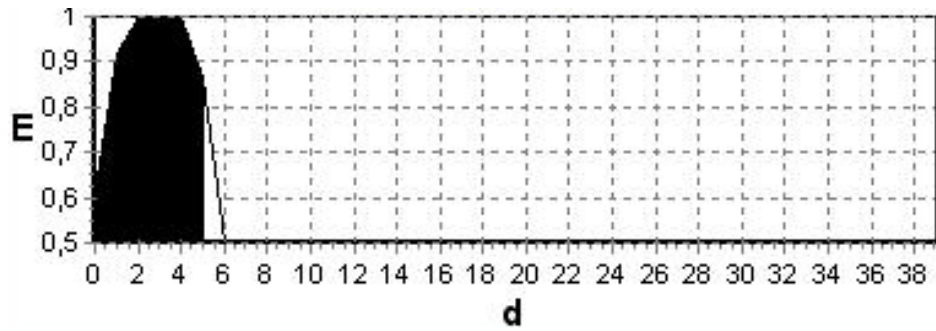


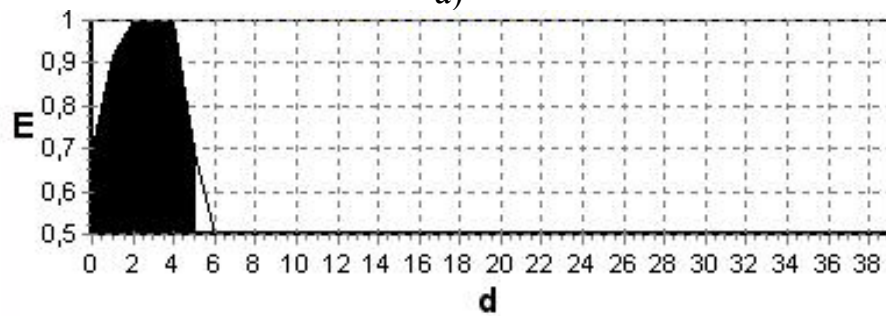
Рисунок 3.7 – Графік залежність усередненого КФЕ від параметра *delta*

Аналіз рисунку 3.7 свідчить про досягнення асимптотично максимального значення при $delta^* = 44$ та широку робочу область типу плато, де одночасно вдалося побудувати безпомилкові вирішальні правила для всіх класів розпізнавання.

Оскільки вирішальні правила задаються в просторі ознак розпізнавання геометричними параметрами контейнерів класів розпізнавання, які відновлюються в процесі навчання СППР, то на рис. 3.8 показано графіки оптимізації радіусів контейнерів класів розпізнавання при оптимальному за узагальненим КФЕ значенням параметра δ .



а)



б)

Рисунок 3.8 – Оптимізація радіусів контейнерів класів розпізнавання:
а) клас X_1^0 ; б) клас X_2^0 ;

Аналіз рис. 3.8 свідчить про досягнення асимптотичного максимуму функції ентропійного КФЕ в робочій області його визначення.

Таким чином, в результаті навчання СППР за гібридним алгоритмом побудовано безпомилкові вирішальні правила

Реалізований в рамках роботи алгоритм екзамену (функціонування) гібридної інтелектуальної системи вирощування сцинтиляційних монокристалів підтвердив надійність розробленого інформаційного і програмного забезпечення.

ВИСНОВКИ

1. Розроблено інформаційне та програмне забезпечення СППР для класифікаційного керування вирощуванням монокристалів.

2. Реалізовано гібридні алгоритм та математичну модель селекції словника ознак на основі методу групового врахування аргументів та інформаційно-екстремальної інтелектуальної технології.

3. Розроблено алгоритм класифікаційного керування в рекомендаційному режимі на основі алгоритму екзамену СППР.

4. В рамках розробленого алгоритму побудовано безпомилкове за навчальною вибіркою вирішальне правило, що дозволяє з високою достовірністю приймати рішення про належність поточної вектор-реалізації до того чи іншого функціонального стану системи керування вирощуванням монокристалу.

5. Одержані результати підтверджено фізичним моделюванням функціонування та навчання гібридної інтелектуальної системи на основі реальних даних, отриманих в процесі моніторингу технологічного процесу вирощування лужно-галоїдних неорганічних сцинтиляційних матеріалів на основі сполуки CzI , активованої Ta (талій) на базі промислової установки «РОСТ»Ю яке підтверджено.

6. Результати роботи впроваджено в НТК «Інститут монокристалів» (м. Харків).

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ ІНФОРМАЦІЇ

1. Гринев Б.В., Дубовик М.Ф., Толмачев А.В. Оптические монокристаллы сложных оксидных соединений. - Харьков: Институт МК, 2002. - 251 с.
2. Атрощенко Л.В., Бурачас С.Ф., Гальчинецкий Л.П., Гринев Б.В. и др. Кристаллы сцинтилляторов и детекторы ионизирующих излучений на их основе. – Киев: Наукова думка, 1998. - 310 с.
3. Гектин А.В., Гринев Б.В. Галоидные сцинтилляторы. Научные и инженерные разработки / Функциональные материалы для науки и техники. – Харьков, 2001. – С. 28-47.
4. Blumberg H., Wilke K. Die Automatisierung von Czochralski-Anlagen / Kristal und Technik, 1974. - Vol. 9. - №5. - S. 447-455.
5. Горилецкий В.И., Радкевич А.В., Эйдельман Л.Г. Принципы управляемого вытягивания МК из расплава / Обз.инф. - М.: НИИТЭХИМ. - 1977. - 37 с.
6. Методы и средства контроля, регулирования и управления процессов выращивания МК / Аннот. библиограф. указ. – Харьков: ВНИИ МК, 1975 - 156 с.
7. Методы и средства автоматизации и моделирования процессов выращивания МК / Аннот. библиограф. указ. – Харьков: ВНИИ МК, 1983 - 157 с.
8. Бурачас С.Ф., Стадник П.Е., Тиман Б.Л. Автоматизация процессов роста кристаллов/ Обз. инф. НИИТЭИ // Хим. пром-сть. Сер.: Монокристаллы и особо чистые вещества - М., 1984. - 69 с.
9. Краснополюсовський А.С. Інформаційний синтез інтелектуальних систем керування: Підхід, що ґрунтується на методі функціонально-статистичних випробувань. – Суми: Видавництво СумДУ, 2004. – 261с.
10. Довбиш А.С. Основи проектування інтелектуальних систем: Навчальний посібник.– Суми: Видавництво Сум ДУ, 2009.– 171 с.