

УДК 537.8.029.6;621.35.029.6; 537.86.029.65/.79

КП

№ держреєстрації 0112U000208

Інв.

Міністерство освіти і науки України
Харківський національний університет радіоелектроніки
(ХНУРЕ)

61166, м. Харків, пр. Леніна, 14;

тел. (057) 702 13 97

ЗАТВЕРДЖУЮ

Проректор з наукової роботи

д.ф-м. н., професор

_____ М.І. Сліпченко

2014 р.

ЗВІТ

ПРО НАУКОВО-ДОСЛІДНУ РОБОТУ

**«Фізичні моделі та інструментарій для 3-D візуалізації взаємодії
низькоінтенсивного електромагнітного поля з мікро- та наноб'єктами різної
фізичної природи та біосередовищами (кров)»**

(заключний)

Керівник НДР
д.ф-м.н, професор

Г.І. Чурюмов

(підпис)

2014

Рукопис закінчено 16.11.2014 р.

Результати цієї роботи розглянуто Науково-методичною радою ХНУРЕ,
протокол № 44 від « 21 » листопада 2014 р.

СПИСОК ВИКОНАВЦІВ

Керівник НДР докт. фіз.-мат.наук, професор	Г.І. Чурюмов (вступ, розділ 1, заключення)
Відповідальний виконавець НДР, молодший науковий співробітник	А.І. Екезлі (розділ 1)
Керівник розділу 1 докт. фіз.-мат.наук, професор	Г.І. Чурюмов (розділ 1)
Керівник розділу 2 докт. техн.наук, професор	Ю.П. Мачехін (розділ 2, висновки)
Керівник розділу 3 докт. техн.наук, професор	О.Г. Руденко (розділ 3, висновки)
Керівник розділу 4 докт. техн.наук, професор	Є.В. Бодянский (розділ 4, висновки)
Керівник розділу 5 канд. техн.наук, професор	О.І. Дохов (розділ 5, висновки)
Керівник розділу 6 докт. техн.наук, професор	В.О. Стороженко (розділ 6, висновки)
Керівник розділу 7 докт. техн.наук, професор	І.П. Захаров (розділ 7, висновки)
Виконавець розділу 1 докт. фіз.-мат.наук, професор	Є.М. Одаренко (розділ 1)
Виконавець розділу 1, канд. техн.наук, с.н.с.	І.Г. Перова (розділ 1)
Виконавець розділу 1, Аспірант	С.Г. Галстян (розділ 1)
Виконавець розділів 1 та 2, м.н.с.	О.Б. Ісаєва (розділ 1,2)

Виконавець розділу 1 канд. фіз.-мат.наук, доцент	Т.І. Фролова (розділ 1,2)
Виконавець розділу 2, аспірант	Л.С. Хоролец (розділ 2)
Виконавець розділу 2, аспірант	А.В. Гнатенко (розділ 2)
Виконавець розділу 3, канд. техн.наук, п.н.с.	О.М. Сотніков (розділ 3)
Виконавець розділу 3, с.н.с.	С.Г. Єлаков (розділ 3)
Виконавець розділу 3, с.н.с.	А.О. Елісеев (розділ 3)
Виконавець розділу 4, канд. техн.наук, п.н.с.	І.П. Плісс (розділ 4)
Виконавець розділу 4, канд. техн.наук, с.н.с.	А.І. Долото (розділ 4)
Виконавець розділу 4, канд. техн.наук, с.н.с.	Н.О. Тесленко (розділ 4)
Виконавець розділу 5, наук.співр.	В.Й.Стрільченко (розділ 5)
Виконавець розділу 5, провідний науковий співробітник, канд. техн. наук	Г.М. Валєвахін (розділ 5)
Виконавець розділу 5, п.н.с., канд. техн. наук	О.Я. Контарь (розділ 5)
Виконавець розділу 5, г.н.с., докт. техн. наук	І.І. Зима (розділ 5)
Виконавець розділу 6, канд. техн. наук, доцент	С. М. Мешков (розділ 6)

Виконавець розділу 6, канд. техн. наук, доцент	Р. Орел (розділ 6)
Виконавець розділу 6, н.с.	О. Мягкий (розділ 6)
Виконавець розділу 6, м.н.с.	О. Бандурян (розділ 6)
Виконавець розділу 7, канд. техн.наук, доцент	М.П. Сергієнко (розділ 7)
Виконавець розділу 7, аспірант	О.І. Бочарова (розділ 7)
Виконавець розділу 7, аспірант	О.А. Боцюра (розділ 7)
Виконавець розділу 7, аспірант	К.А. Клімова (розділ 7)
Виконавець розділу 7, аспірант	Ю.Г. Жарко (розділ 7)
Виконавець розділу 7, аспірант	Альравашдех Бакер (розділ 7)
Нормоконтролер	Т.Г. Сілантьєва

В виконанні НДР прийняли участь студенти студенти Подюк А.В. та Григоров Н.С. (гр. ЛОЕТ 11-1) та магістр Усенко А.А. (гр. ЛОЕТ 11-1).

РЕФЕРАТ

Звіт о НДР: 336 стор., 97 рис., 10 табл., 116 джерел.

Об'єкт дослідження – процес взаємодії електромагнітного випромінювання з мікро- і нанооб'єктами та середовищами різної фізичної природи.

Предмет дослідження – фізичні (феноменологічні) та математичні моделі фізичних процесів взаємодії електромагнітного поля з середовищем (кров), в тому числі і на мікрорівні, а також технічні засоби для експериментального дослідження особливостей впливу поля на властивості та параметри середовищ.

Метою проекту є подальший розвиток теорії та практики взаємодії електромагнітного поля з мікро- та нанооб'єктами, в тому числі біологічними (мікроби, бактерії, віруси), розробка технічних засобів, інформаційного та метрологічного забезпечення для автоматизації процесу виявлення особливостей впливу електромагнітної енергії на біологічне середовище (кров) та змін його властивості. Для цього в рамках загальної проблеми взаємодії електромагнітного поля з активними та пасивними середовищами вирішувалася конкретна фундаментальна задача взаємодії низькоінтенсивного електромагнітного поля широкого діапазону частот (від міліметрового до оптичного) на мікро- та нанооб'єкти, а також середовища (в тому числі біологічні), виявлення фізичних умов та особливостей впливу поля на властивості об'єктів та середовищ в цілому, проведено аналіз експериментальних і теоретичних досліджень впливу електромагнітного випромінювання на біологічні системи і об'єкти живої природи. Розглядаються особливості та специфіка теплового (потужність випромінювання більше 10 Вт) і нетеплового (інформаційного) (потужність випромінювання менше 1 Вт) застосування електромагнітної енергії в лікувальних цілях. Наводяться варіанти побудови математичних моделей електромагнітного поля і біологічних систем

на основі самоузгоджених систем диференціальних рівнянь в приватних похідних для електромагнітного поля (рівняння Максвелла, хвильове рівняння) і рівняння теплопровідності. Отримала подальший розвиток математична модель протибактериального імунної відповіді з урахуванням дії зовнішнього електромагнітного поля, в тому числі моделювання дії енергії поля на мікробну, вірусну та грибкову середу в крові.

Для аналізу дії електромагнітного поля на біологічні системи (кров) проведено вибір структурної схеми та розроблення макету цифрового 3-D голографічного мікроскопу; розроблення та дослідження методів обробки зображень під час введення та під час класифікації об'єктів, що досліджуються в реальному часі; дослідження просторових та частотних характеристик оптичного випромінювання, яке забезпечує ефективний вплив на наноб'єкти; дослідження умов формування просторових характеристик оптичного випромінювання необхідних для маніпулювання наноб'єктами та контролю за переміщенням у просторі; розроблення статистичних правил прийняття рішень ідентифікації виявлених неоднорідностей в середовищі; розробка та вдосконалення методів, засобів та процедур метрологічного забезпечення нанометрології; розробка рекомендацій для практичного застосування отриманих результатів.

Особливістю данної роботи є розроблення нейромережових методів відновлення вихідних даних, що враховують нелінійність та нестационарність об'єктів; розроблення нейро-фаззи-моделей та методів їх навчання, орієнтованих на малу навчальну вибірку; дослідження змін властивостей об'єктів в результаті комплексного впливу різних фізичних полів;

ЕЛЕКТРОМАГНІТНЕ ПОЛЕ, БІОЛОГІЧНА СИСТЕМА,
МАТЕМАТИЧНА МОДЕЛЬ, МІКРОХВИЛЬОВА ДІАГНОСТИКА,
ОБЧИСЛЮВАЛЬНИЙ ІНТЕЛЕКТ, ІНТЕРФЕРЕНЦІЙНИЙ МІКРОСКОП,
ОБРОБКА ІНФОРМАЦІЇ

ЗМІСТ

ВСТУП	12
1 МЕДИЧНІ АСПЕКТИ ВПЛИВУ ЕЛЕКТРОМАГНІТНОГО ВИПРОМІНЮВАННЯ НА БІОЛОГІЧНІ СИСТЕМИ І ОБ'ЄКТИ.....	15
1.1 Аналіз сучасного стану проблеми.....	15
1.2 Особливості застосування математичних моделей	22
1.3 Постановка завдання впливу ЕМВ.....	26
1.4 Математичні моделі електромагнітного поля.....	28
1.5 Математичні моделі біологічних систем.....	30
1.5.1 Випадок теплового впливу ЕМП	31
1.5.2 Випадок інформаційного (нетеплового або специфічного) впливу ЕМП	33
1.6 Моделювання впливу ЕМВ на біологічні об'єкти.....	35
1.7 Цифровий голографічний інтерференційний мікроскоп	39
1.8 Висновки к розділу 1	42
2 ВЗАЄМОДІЯ ЛАЗЕРНОГО ВИПРОМІНЮВАННЯ 3 БІОЛОГІЧНИМИ МІКРО ТА НАНО ОБ'ЄКТАМИ	43
2.1 Дослідження фізичних умов механічного впливу лазерного випромінювання на біологічні об'єкти	43
2.1.1 Основні принципи роботи лазерного пінцету (стислий історичний екскурс)	45
2.1.2 Фізичні умови і вихідні теоретичні умови.....	48
2.2 Основні принципи оптичного захоплення	58
2.2.1 Сили оптичного захоплення.....	58
2.3 Основні типи оптичних пасток	65
2.3.1 Однопучкова 2Д-пастка.....	65
2.3.2 Двопучкова 3Д-пастка	67
2.3.3 Однопучкова 3Д-пастка.....	68

2.3.4	Інтерференційна пастка	69
2.4	Основна схема оптичного пінцета	70
2.5	Області застосування оптичного захоплення мікрооб'єктів	72
2.5.1	Маніпулювання біоб'єктами	72
2.5.2	Фізика колоїдів.....	74
2.5.3	Мікромеханіка.....	74
2.6	Динамічне управління мікрооб'єктом.....	76
2.6.1	Методи управління об'єктом	76
2.6.2	Метод управління за допомогою пересування оптики.....	80
2.6.3	Метод управління за допомогою предметного столу.....	82
2.6.4	Вибір конструкції координатного столика	83
2.6.5	Аналіз профілю інтенсивності лазера	87
2.7	Розробка основних вузлів оптичного пінцета	96
2.7.1	Система візуалізації мікрооб'єктів	96
2.7.2	Система захоплення і утримання мікрооб'єктів	107
2.7.3	Система переміщення мікрооб'єктів	112
2.7.4	Охолоджені атоми в фотонних кристалах.....	122
2.7.5	Оптичний лазерний маніпулятор з фемтосекундними імпульсами.....	132
2.7.6	Рух частки в потенційній ямі під дією світла	134
2.7.7	Ефект руху потенційної ями	135
2.8	Виводи до розділу 2	139
3	ЗАСОБИ ОБРОБКИ ВІДЕОІНФОРМАЦІЇ В РЕАЛЬНОМУ ЧАСІ НА БАЗІ НЕЙРОННОЇ МЕРЕЖІ СМАС	140
3.1	Загальні відомості	140
3.2	Етапи обробки інформації	143
3.2.1	Кодування інформації	143
3.2.2	Вибір базисних функцій.....	146
3.2.3	Хешування інформації	150

3.2.4 Хеш-колізії	152
3.3 Методи навчання.....	154
3.3.1 Алгоритм навчання Альбуса	155
3.3.2 Методи скорочення процесу навчання мережі.....	165
3.3.3 Навчання за допомогою ковзаючого середнього (AV).....	167
3.3.4 Навчання в точці з максимальною помилкою (ME)	168
3.3.5 Часткова оптимізація довжини кроку в останньому перпендикулярному напрямі (OS)	170
3.3.6 Алгоритм навчання з самоналагоджувальними коефіцієнтами навчання.....	171
3.4 Застосування нейронної мережі СМАС для обробки відеоінформації	173
3.4.1 Застосування мережі СМАС при мікшуванні зображень	174
3.4.2 Виконання колірної корекції зображень з використанням мережі СМАС	177
3.4.3 Застосування мережі СМАС для стиснення і фільтрації зображень	179
3.5 Модифіковані архітектури ШНМ СМАС	183
3.6 Адаптивна дискретизація вхідних сигналів.....	186
3.7 Мережа Low Dimensional Based СМАС	192
3.8 Ієрархічна мережа НСМАС.....	193
3.9 Мінімізація структури мережі НСМАС	200
3.9.1 Архітектура МНСМАС.....	200
3.9.2 Порівняння структур НСМАС і МНСМАС.....	204
3.9.3 Самоорганізація вхідного простору МНСМАС	206
3.9.4 Правило навчання ШНМ МНСМАС.....	208
3.10 Апаратна реалізація нейронної мережі СМАС: загальний підхід.....	215
3.10.1 Кодування змінних	215
3.10.2 Обчислення значення функції	221

3.11 Апаратна реалізація мережі СМАС у разі послідовного перебору значень змінних	223
3.12 Стиснення і фільтрація зображень.....	332
4 НЕЙРО-ФАЗЗИ СИСТЕМИ ТА АЛГОРИТМИ НАВЧАННЯ ДЛЯ ОБРОБКИ НЕОДНОРІДНИХ ПОЛІВ СПОСТЕРЕЖЕНЬ ЗА УМОВ МАЛОЇ ВИБІРКИ ДАНИХ	238
5 ДОСЛІДЖЕННЯ ЗМІН ВЛАСТИВОСТЕЙ ОБ'ЄКТІВ ПІД ВПЛИВОМ КОМБІНОВАНИХ ПОЛІВ РІЗНОЇ ФІЗИЧНОЇ ПРИРОДИ ТА ЇХ ПРАКТИЧНЕ ЗАСТОСУВАННЯ	252
5.1 Вплив електромагнітного випромінювання на біологічні об'єкти	252
5.2 Вплив резонансної магнітоакустичної обробки на фізичні властивості води.....	253
5.2.1 Вплив окремих видів резонансної магнітоакустичної обробки на властивості води	255
5.3 Висновки до розділу 5	264
6 ПОБУДОВА ТЕПЛОФІЗИЧНИХ МОДЕЛЕЙ ПРОЦЕСУ АКТИВНОГО ТЕПЛОВОГО КОНТРОЛЮ.....	265
6.1 Теплофізична модель композитних стільникових конструкцій	265
6.2 Програмна реалізація методу	271
6.3 Результати розрахунків та їх аналіз.....	273
6.4 Теплофізична модель судів крові.....	275
6.5 Висновки до розділу 6	277
7 ФОРМУВАННЯ ЗАСАД МЕТРОЛОГІЧНОГО ЗАБЕЗПЕЧЕННЯ ЦИФРОВОЇ ГОЛОГРАФІЧНОЇ МІКРОСКОПІЇ	278
7.1 Дослідження існуючих методів, технічних засобів та їх метрологічного забезпечення для виконання вимірювань у нанодіапазоні.....	278

7.1.1 Програми та стратегії в області нанометрології	278
7.1.2 Стандарти в області нанометрології.....	284
7.2 Вдосконалення методів та засобів нановимірювань та їх метрологічного забезпечення	286
7.2.1 Аналіз основних особливостей нановимірювань та оцінювання їх невизначеності	286
7.2.2 Особливості оцінювання невизначеності вимірювань в стандарті ISO 14253-2	290
7.3 Підвищення достовірності статистичних оцінок невизначеності вимірювань	293
7.4 Оцінювання розширеної невизначеності вимірювань	300
7.5 Практичні питання щодо оцінювання невизначеності нановимірювань	307
7.5.1 Вибір мінімальної кількості результатів спостережень при складанні методики виконання вимірювань	307
7.5.2 Оцінювання невизначеності вимірювань диференційної функції	309
7.6 Висновки до розділу 7	312
ЗАКЛЮЧЕННЯ.....	315
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ.....	318
Додаток А1.....	329
Додаток А2.....	330
Додаток А3.....	331
Додаток А4.....	332
Додаток А5.....	333
Додаток А6.....	334
Додаток А7.....	335
Додаток А8.....	336

ВСТУП

Розвиток нанотехнологій вимагає розробки нових підходів і методів спостереження явищ і процесів, що протікають в різних середовищах і об'єктах, а також вивчення відбуваються в них структурних змін в результаті впливу зовнішніх факторів. Особлива роль у дослідженні мікро- і наномасштабних явищ і процесів відводиться методам класичної оптичної мікроскопії [1]. Однак до істотних недоліків даних методів належать обмеження, пов'язані з граничним збільшенням просторової роздільної здатності в силу явища дифракції хвиль (формула Аббе), а також неможливістю (або технічною складністю) реалізації тривимірної (3-D) візуалізації фазових мікрооб'єктів. Більш перспективним напрямком досліджень і вивчення поведінки фазових мікрооб'єктів є голографічна мікроскопія [2-4], можливості якої в поєднанні з методами комп'ютерної обробки оптичної інформації (цифрова голографічна мікроскопія) дозволяють здійснити динамічну 3-D візуалізацію фазових мікрооб'єктів. В даний час цифрова голографічна мікроскопія забезпечує формування зображень з осьової роздільною здатністю ~ 5 нм, дозволяє досліджувати прозорі середовища та об'єкти, формувати 3-D зображення, проводити моніторинг біологічних об'єктів і мікроорганізмів без їх руйнування або пошкодження (неінвазивний метод) [5]. У той же час розроблена 3-D технологія побудови зображення, в основному, являє собою алгоритм для візуалізації оптичної інформації і не передбачає автоматизовану її обробку (кластеризацію) для автоматичного виявлення можливих зміни і що з'явилися відмінностей фазових мікрооб'єктів та їх структури в результаті впливу навколишнього середовища або впливу зовнішніх факторів (електромагнітного поля, різних лікарських препаратів і т.п.).

Метою даної роботи є подальший розвиток і вдосконалення методу 3-D цифрової голографії, пов'язане з автоматизацією процесу обробки 3-D зображень за допомогою методів обчислювального інтелекту.

Необхідно зазначити, що ідея голографічного методу запису і відтворення інформації для подальшого її візуального аналізу не є новою і відома давно [6]. Використовуючи даний метод вдалося реалізувати різні схеми запису і відтворення голограм, які лягли в основу створення голографічних мікроскопів (безлінзовий і комбінованих). Дані мікроскопи на відміну від класичних оптичних мікроскопів після запису інформації про динамічні фазових мікрооб'єктах на фотопластинку (запис голограм) дозволяють апостеріорно переглядати і аналізувати їх збільшене відновлене об'ємне 3-D зображення. З іншого боку, в силу здатності голограм зберігати запис хвильового фронту та відновлювати його в довільний момент часу, голографічний мікроскоп можна розглядати одночасно і як інтерференційний його аналог. Їх спільною особливістю є можливість дослідження фазових мікрооб'єктів, тобто мікрооб'єктів, які слабо впливають на інтенсивність проходить (або відображеної) електромагнітної хвилі, але здатні робити істотний вплив на її фазу. У підсумку отримана інтерференційна картина (інтерферограмма), яка являє собою результат складання опорної та сигнальної хвиль, містить в собі повну інформації, включаючи зміни амплітуди і фази сигнальної хвилі. Цього виявляється достатньо, щоб при відновленні голограми візуалізувати найдрібніші деталі досліджуваних фазових мікрооб'єктів. Необхідно відзначити той факт, що вибір схеми реалізації голографічного мікроскопа, наприклад, схеми на відбиття або на просвіт залежить від особливостей досліджуваних середовищ і мікрооб'єктів. Для вивчення властивостей прозорих середовищ і мікрооб'єктів (наприклад, дослідження живих клітин, бактерій і т.п.) оптимальним для застосування є схема на просвіт. У той же час для дослідження непрозорих мікрооб'єктів і середовищ доцільно застосовувати схему на відображення.

Можливості голографічних мікроскопів, пов'язані з підвищенням інформативності та достовірності інтерпретації експериментальних даних, можна вдосконалювати і покращувати за допомогою цифрових методів

обробки оптичної інформації. Для реєстрації (запису) цифровий голограми, що містить інформацію про мікрооб'єктах, використовується цифрова камера, з якої через USB-порт цифровий сигнал передається в комп'ютер для 3-D візуалізації (відновлення) і автоматизованої обробки зображень за допомогою розробленого спеціалізованого програмного забезпечення (ВПЗ) в режимі реального часу. Особливістю застосовуваного СПО є той факт, що розроблено воно на основі методів обчислювального інтелекту і дозволяє проводити порівняння різних властивостей фазових мікрооб'єктів (наприклад, форми, розмірів і т.п.) до і після зовнішнього впливу. За своєю суттю "порівнянні" піддаються цифрові інтерферограми фазових мікрооб'єктів, що дозволяє фіксувати динаміку їх змін в відмінності.

Таким чином, запропонований підхід розширює можливості цифрової 3-D голографічного мікроскопії і дозволяє автоматизувати процес обробки інтерферограми фазових мікрооб'єктів в режимі реального часу. Для досягнення цього використовується математична обробка зображень за допомогою методів обчислювального інтелекту. На основі запропонованих алгоритмів розроблено СПО для автоматизації процесу обробки та аналізу фазових мікрооб'єктів до і після зовнішнього впливу.

Метою проекту є подальший розвиток теорії та практики взаємодії електромагнітного поля з мікро- та нанооб'єктами, розробка технічних засобів, інформаційного та метрологічного забезпечення для автоматизації процесу виявлення особливостей впливу електромагнітної енергії на біосередовище та його властивості.

1 МЕДИЧНІ АСПЕКТИ ВПЛИВУ ЕЛЕКТРОМАГНІТНОГО ВИПРОМІНЮВАННЯ НА БІОЛОГІЧНІ СИСТЕМИ І ОБ'ЄКТИ

1.1 Аналіз сучасного стану проблеми

Результати впливу електромагнітного випромінювання (ЕМВ) на об'єкти і середовища біологічної природи і наслідки такого впливу мають широке практичне застосування [див., Напр., 7-9]. Актуальність даних питань не викликає сумнівів, так як затребувана всім прогресом розвитку нових методів діагностики та лікування різних захворювань [10].

В основі впливу ЕМВ на біологічні системи лежить ефект перетворення енергії електромагнітного поля (ЕМП) в інші види енергій, зокрема, в теплову (тепловий вплив) [7, 10], а також стимулювання життєдіяльності організмів зовнішнім ЕМІ та встановлення закономірностей електромагнітного зв'язку між організмами (нетепловий або інформаційний вплив) [7, 11, 12]. Це відкриває нові можливості в біології [13] (квантова біологія [14]), медицині [15] (квантова медицина [16]), біофізики, сільському господарстві та інших галузях науки.

Розглядаючи вплив ЕМВ на біологічні системи, зокрема, на організм людини, слід зазначити, що умови розповсюдження і характер взаємодії ЕМП з біологічними об'єктами істотно змінюються залежно від частоти та інтенсивності випромінювання [17].

На рис. 1.1 схематично представлені задіяні частоти ЕМВ для різних терапевтичних застосувань з повного спектру електромагнітних коливань. Слід зазначити, що в обраних для застосування діапазонах частот задіяні лише окремі частоти, застосування яких узгоджено з Міжнародної комісії з радіочастот для промислового, побутового та медичного застосувань.

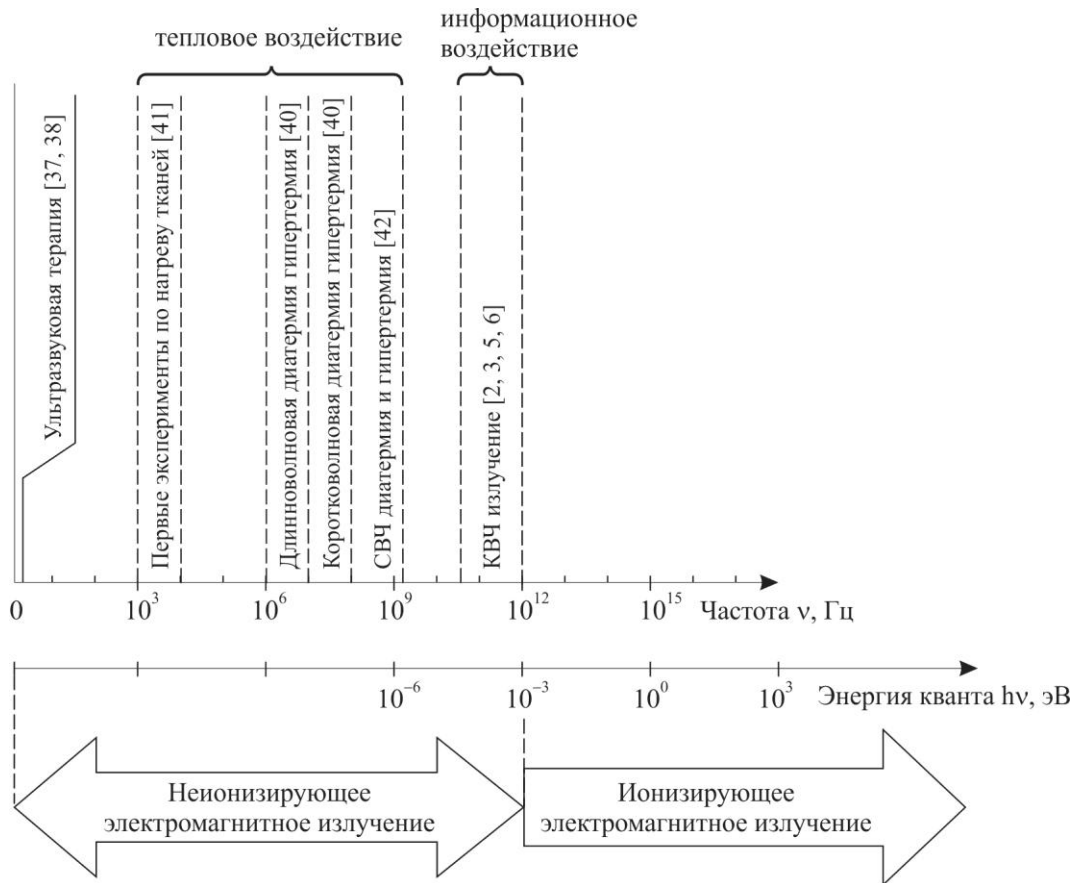


Рисунок 1.1 – Схема використання спектру електромагнітних коливань в медичних цілях

Аналіз впливу ЕМВ на біологічні системи можна охарактеризувати як теплове (або високоенергетичне дія) та інформаційне (низькоенергетичний дія). У разі високоенергетичного дії ЕМВ (з інтенсивністю поля більше 100 В / м або потужності випромінювання більше 10 Вт) починають помітно проявлятися теплові ефекти, завдяки процесу перетворення електромагнітної енергії у внутрішню енергію об'єктів різної біологічної структури, що викликає підвищення їх температури. Це дозволяє використовувати електромагнітну енергію в терапевтичних цілях для нагрівання уражених тканин (довгохвильова, короткохвильова і СВЧ діатермія або мікрохвильова терапія [10]) і лікування онкологічних захворювань (гіпертермія [18]). У цьому випадку мова йде про застосування властивості поглинання електромагнітної енергії в

живих тканинах для контрольованого і тимчасового підвищення температури тіла, окремого органу або частини органу, ураженого патологічним процесом, понад 39°C до 42°C .

У роботах [10, 18] здійснено огляд застосування існуючих методів термічної терапії в медицині (діатермія і гіпертермія), можливих біологічних ефектів від теплового впливу ЕМП різної довжини хвилі, схем і конструкцій апаратури для практичної реалізації такого впливу.

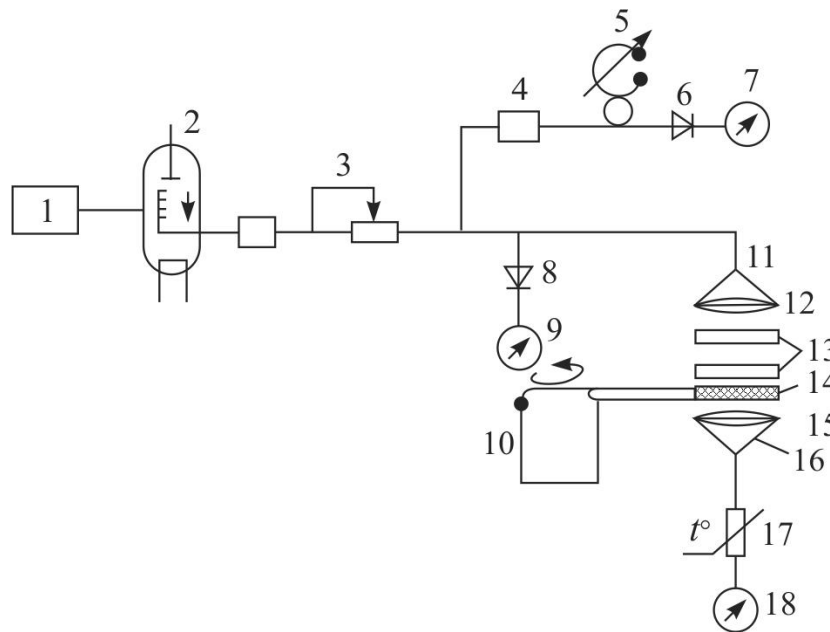
Останнім часом методи гіпертермії ефективно і широко застосовуються в онкології. Зокрема, в [19] розглядається сучасний стан проблеми застосування гіпертермії в клінічній онкології. Відзначаються позитивні результати застосування гіпертермії на підставі експериментальних робіт та досліджень європейських вчених. У роботі розглядаються позитивні аспекти застосування гіпертермії. При температурі $39-41^{\circ}\text{C}$ в тканинах організму активується мікроциркуляція, посилюються харчування і метаболізм клітин і тканин, прискорюється провідність в периферичних нервах, знижується тонус м'язів. Також в [19] відзначена реактивация хронічного запалення, яке стає гострим, що ініціює процеси самовідновлення в організмі, зокрема, відновлення нормальних імунних реакцій.

У разі використання ЕМВ низької інтенсивності (з інтенсивністю поля менше 10 В/м і потужністю менше 1 Вт) має місце особливий процес взаємодії ЕМП з біологічними середовищами та об'єктами, який протікає без помітного збільшення середньої кінетичної енергії складових їх молекул (об'єктів). Такий вплив ЕМП на біологічні середовища та об'єкти (наприклад, окремі клітини) в літературі визначається як специфічне чи інформаційне дія, яка має явно виражену біологічну ефективність і викликає зміна функціонального стану фізіологічної системи на клітинному і молекулярному рівнях [7, 9, 11-16]. У його основі лежить гіпотеза про резонансне взаємодії ЕМВ НВЧ діапазону з біологічними системами, які супроводжують механізм обміну енергією між ЕМП і біологічними системами з різною молекулярною організацією

внутрішньої структури. З іншого боку, не менш важливим моментом інформаційного впливу на біологічні об'єкти є облік модуляційно-часових параметрів впливає ЕМП, що дозволяє розглядати такий процес як передачу інформації між ЕМП і біологічною системою (середина її дискретні об'єкти) або всередині самої системи (між окремими об'єктами системи) [7, 9]. У цьому процесі для підвищення стимулюючого фактора від впливу ЕМП слід визначити залежність біологічної ефективності випромінювання КВЧ діапазону від його інтенсивності і частоти, підбираючи експериментально для кожного конкретного випадку «амплітудні» і «частотні» вікна, вид і форму модулюючого сигналу для управління процесами, що протікають в біологічній системі [17, 20].

Досить докладно питання інформаційного впливу ЕМВ КВЧ діапазону на живі організми досліджувалися в роботах, виконаних у НДІ «Істок» (м. Фрязіно) під керівництвом акад. Девяткова М.Д. Дані результати є найбільш фундаментальними і складають основу сучасних розробок в цьому напрямку. Так, в роботі [9] зазначається, що енергія квантів КВЧ діапазону, з одного боку, менше кінетичної енергії теплового руху атомів і молекул, що припадає на одну ступінь свободи, $E_T = kT$ де $k = 1,38 \cdot 10^{-23} \text{ Дж} \cdot \text{K}^{-1}$ – постійна Больцмана, $T = 293,15 \text{ K}$ – кімнатна температура за шкалою Кельвіна, а з іншого - багато менше енергії навіть слабких водневих зв'язків в живих організмах. Відсутність шкідливої дії на організм в сукупності з великою інформативною ємністю може сприяти використанню розглянутих діапазонів хвиль живими організмами для зв'язку і управління при невисокому рівні потужності керуючих сигналів.

На рис. 1.2 представлена структурна схема однієї з типових лабораторних установок для впливу ЕМП на живі організми [9].



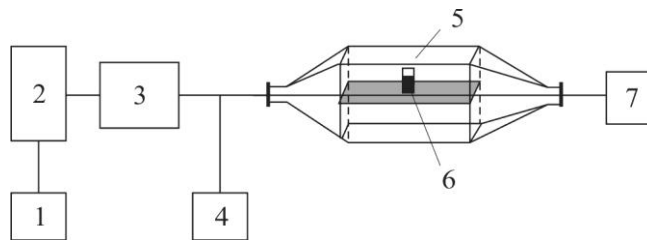
1 - блок живлення; 2 - лампа зворотної хвилі; 3- ослабителі;
 4 - вимірювальна лінія; 5 - хвилемір; 6, 8 - детекторні головки;
 7 - стрілочний прилад; 9 - вимірювач потужності падаючого випромінювання; 10 - електромотор, що обертає і перемішуючий середу (в деяких експериментах з мікроорганізмами); 11, 16 - рупори; 12, 15 - коригувальні лінзи; 13 - трансформатор; 14 - об'єкт; 17 - термісторний головка; 18 - вимірювач потужності проходить випромінювання.

Рисунок 1.2 - Схематичне зображення установки для впливу ЕМВ на живі організми [9]

Особливо слід відзначити дослідження, спрямовані на вивчення впливу фізичних полів і, зокрема, ЕМП на кров та її формені освіти (еритроцити, лейкоцити, тромбоцити). Так, в роботі [21] розглядається роль КВЧ випромінювання в підвищенні біологічної активності крові, що призводить до поліпшення неврологічного стану хворих, зменшенню агрегативної тромбоцитів і активізації антиоксидантної системи при опроміненні крові на довжині хвилі 7,1 мм.

В експериментальній роботі [22] описуються електричні і магнітні властивості формених елементів крові одного виду - еритроцитів. На основі кулонівської взаємодії представлені наближені моделі електростатичного взаємодії еритроцитів у кровотоці, яке виникає між двома зарядженими частинками, а також описується наближена модель магнітної взаємодії еритроцитів у крові та гігроскопічні ефекти, що виникають при русі еритроцитів.

Вивчення впливу ЕМП і парамагнітних частинок на еритроцити проводилися в роботі [23], в якій представлена експериментальна установка для опромінення еритроцитів ЕМП з щільністю потужності $0,6 \text{ мВт} / \text{см}^2$ і частотою 418 МГц (рис. 1.3).



1 - джерело живлення; 2 - генератор НВЧ; 3 - пристрій узгодження;
4 - вимірювач НВЧ потужності; 5 - СВЧ резонатор; 6 - дослідницький зразок;
7 - вимірювач прохідної НВЧ потужності

Рисунок 1.3 - Система для електромагнітної експозиції

Питання впливу низькоінтенсивного ЕМВ НВЧ діапазону на процеси мікроциркуляції крові у людини і тварин (щурів), а також можливі механізми цього впливу детально розглянуті в [31, 34]. Практичний інтерес до отриманих результатів важливий і полягає в тому, що кров об'єднує і пов'язує між собою багато органів, що забезпечують життєдіяльність організму людини. Дослідження системи крові, яка включає безпосередньо кров, органи

кровотворення і кроверазрушення пов'язана з успіхами, які отримані на молекулярному рівні, в галузі фундаментальної імунології, гематології, біології і біофізики. Облік впливу низькоінтенсивного ЕМВ НВЧ діапазону на систему крові й розуміння механізму такої дії на формені її елементи дозволяє по новому підійти до терапії ряду захворювань системи крові, а також багатьох інших захворювань.

Необхідно відзначити, що існуюча практика дослідження впливу ЕМВ КВЧ діапазону на різні біологічні системи й об'єкти живої природи застосовує, в основному, емпіричний метод дослідження. Враховуючи, складність і неоднозначність досліджуваних явищ і процесів, що протікають при взаємодії ЕМВ з біологічними системами, виникає необхідність у більш поглиблених теоретичних дослідженнях і, зокрема, застосування математичних моделей і моделювання для опису і вивчення біофізичних процесів, що відбуваються на різних рівнях організації живого організму. Об'єднання теорії та експериментальних досліджень дозволяє отримувати об'єктивні і достовірні результати, з яких шляхом узагальнення можна встановлювати загальні закономірності взаємодії ЕМП з біологічними системами і об'єктами, підтверджувати або спростовувати основні положення висунутих раніше гіпотез і концепцій. Такий підхід також служить основою для постановки і прискорення нових клінічних досліджень, розробки та впровадження в медичну практику серії нових приладів і апаратів для діагностики і терапії різних захворювань. На цьому шляху існує ряд обмежень, пов'язаних з розвитком і застосуванням експериментальних методів дослідження. Зокрема, однією з таких експериментальних проблем є існуючий обмежений доступ до використання частотного діапазону, а також досить невеликий вибір вакуумних і твердотільних приладів НВЧ, які забезпечили б необхідний рівень вихідної потужності в зазначених областях частотного спектру і необхідну стабільність частоти генерації. Необхідно відзначити важливість параметра стабільності частоти, так як від його величини залежить рівень неконтрольованої модуляції

несучої частоти і можливі закономірності її зміни. Все це посилює позиції і роль методу математичного моделювання як потужного інструменту теоретичного дослідження біофізичних процесів, що відбуваються на різних рівнях організації біологічних систем при їх взаємодії з ЕМВ з прогнозованими законами зміни амплітуди і частоти.

1.2 Особливості застосування математичних моделей

У роботах [34, 35] відзначаються важливість застосування математичних методів і моделей, використання обчислювального експерименту як інструменту для вивчення біології клітини, автоволнових процесів, які порушуються в сильно нелінійних біологічних середовищах та математичного моделювання просторової організації макромолекул, макромолекулярних комплексів і внутрішньоклітинних структур, а також різних біохімічних реакцій, що відбуваються в організмі людини.

При побудові математичних моделей впливу ЕМВ на біологічні системи, як показано в роботах [26-29], слід виходити з особливостей вибору і використання математичного апарату для опису біологічних процесів і реакцій. Для цього необхідно розуміння й оцінка біологічної проблеми, адекватного математичного подання та формалізації принципово важливих біологічних явищ, а також біологічної інтерпретації отриманих медико-біологічних результатів моделювання. Одним з вдалих таких прикладів є підхід до розробки математичних моделей в імунології, запропонований акад. Марчуком Г.І. і детально представлений в монографії [29]. При цьому імунна система людини розглядається як своєрідний бар'єр, що перешкоджає розмноженню і поразці організму людини різними патогенними мікроорганізму або антигенами (вірусами, бактеріями, грибками) і таким чином захищає організм від руйнівного впливу зовнішніх факторів. В якості основного різновиду імунітету розглядається гуморальний імунітет (захисні функції виконують молекули, що

знаходяться в плазмі крові), пов'язаний з боротьби організму з патогенними антигенами шляхом вироблення антитіл, специфічних до захворювання. У цьому випадку, як показано в [29], процес утворення антитіл пов'язаний з роботою трьох типів клітин: макрофаги, В-лімфоцита і Т-лімфоцита. Головною кліткою, що подає антиген лімфоцитам, є макрофаг. Антигени V, зустрічаючись з макрофагами M, поглинаються ними, переробляються, після чого макрофаги формують набір з антигенних детермінант на своїй поверхні. Схематичне зображення даного процесу представлено на рис. 1.4.

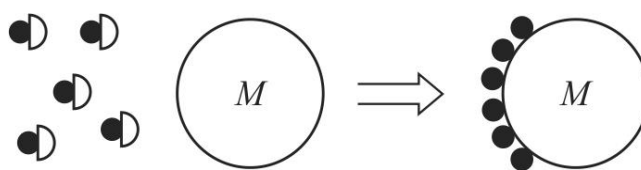


Рисунок 1.4 - Формування набору антигенних детермінант на поверхні макрофага

Макрофаги M подають цей набір B – лімфоцитів, внаслідок чого відбувається зв'язування антигенних детермінант з поверхневими рецепторами B-лімфоцитів іммуноглобулінової природи. Даний механізм представлений на рис. 1.5.

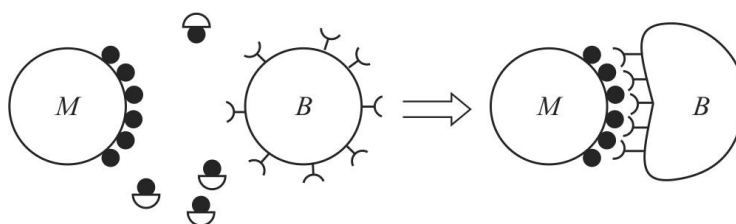


Рисунок 1.5 – Презентування макрофагами детермінант B- лімфоцити

У присутності Тн - помічника, активованого антигеном відбувається стимуляція B-лімфоцита, який починає ділитися і диференціюватися в сторону плазматичних клітин, як показано на рис. 1.6.

Такий каскадний процес утворення клону плазматичних клітин триває від декількох годин до декількох днів.

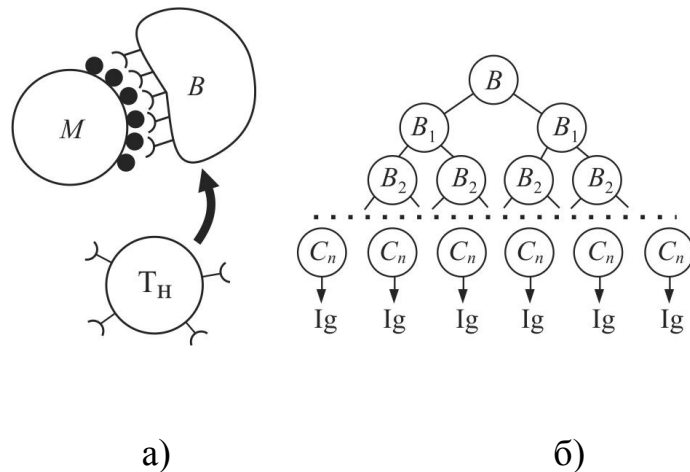


Рисунок 1.6 - Схема процесу стимуляції В-лімфоцита (а) і схема каскадного процесу поділів і освіти клону плазматичних клітин (б)

Розглянута фізіологічна модель механізму імунної реакції організму і вироблення антитіл, специфічних до захворювання, дозволяє представити модель протибактеріального імунної відповіді у вигляді системи зв'язаних диференціальних рівнянь першого порядку з запізнілим аргументом [29]:

1. Рівняння балансу для числа бактерій

$$\frac{dV}{dt} = (\beta - \gamma F)V ; \quad (1.1)$$

2. Рівняння балансу для плазматичних клітин

$$\frac{dC}{dt} = \xi(m)\alpha V(t - \tau)F(t - \tau) - \mu_c(C - C^*) ; \quad (1.2)$$

3. Рівняння для антитіл

$$\frac{dF}{dt} = \rho_F C - (\mu_f + \eta \mathcal{N})F ; \quad (1.3)$$

4. Рівняння для характеристики стану ураженого органу

$$\frac{dm}{dt} = \sigma V - \mu_m m, \quad (1.4)$$

де $V(t)$ – концентрація патогенних розмножуються антигенів; $C(t)$ – концентрація плазматичних клітин; $F(t)$ – концентрація антитіл; $m(t)$ – відносна характеристика ураженого органу; β – коефіцієнт розмноження антигенів; γ – коефіцієнт, що визначає ймовірність нейтралізації антигену антитілами при зустрічі з ним; $m = 1 - M'/M$; M – характеристика здорового органу (маса або площа); M' – відповідна характеристика здорової частини ураженого органу; $\xi(m)$ – функція, що характеризує працездатність імунологічних органів залежно від тяжкості хвороби; α – коефіцієнт, що враховує ймовірність зустрічі антиген - антитіло, збудження каскадної реакції і число утворюваних нових клітин; μ_c – коефіцієнт, що дорівнює зворотній величині їх часу життя; ρ_f – швидкість виробництва антитіл однієї плазматичної кліткою; μ_f – коефіцієнт, обернено пропорційний часу розпаду антитіл; η – кількість антитіл, яке потрібно для нейтралізації одного антигену; σ – константа, яка характеризує захворювання; μ_m – коефіцієнт, що характеризує зворотну величину періоду відновлення органу в е раз.

В якості початкових умов для розв'язання вихідної системи рівнянь (1.1) - (1.4) при $t = t_0$ виходимо з припущення, що $V(t) = 0$ при $t < t_0$, тобто до моменту зараження при $t = t_0$ віруси в організмі відсутні. Тоді, остаточно, початкові умови мають вигляд

$$V(t_0) = V_0; \quad C(t_0) = C_0; \quad F(t_0) = F_0; \quad m(t_0) = m_0. \quad (1.5)$$

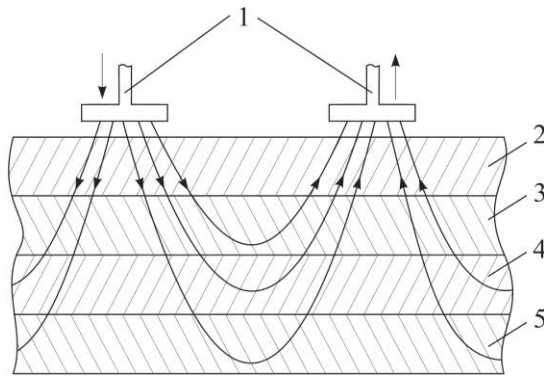
Представлена система нелінійних диференціальних рівнянь (1.1) - (1.4) з початковими умовами (1.5) представляє найпростішу математичну модель захворювання. Використання створеної математичної моделі дає можливість зрозуміти особливості механізму поразки і побудувати оптимальний алгоритм лікування різних інфекційних захворювань (наприклад, вірусного гепатиту С, деструктивної пневмонії та грипозної інфекції [29]).

Таким чином, на основі узагальнення результатів побудови фізичних і фізіологічних моделей різних захворювань організму людини розвивається підхід до побудови математичних моделей, що описують еволюційний цикл змін, що відбуваються в організмі людини при впливі різноманітних зовнішніх факторів на макро і мікро рівнях будови матерії, включаючи вплив патогенних мікроорганізмів і ЕМВ різної інтенсивності і частоти (НВЧ і КВЧ діапазони).

1.3 Постановка завдання впливу ЕМВ

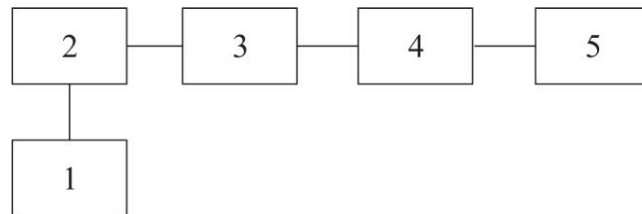
Розглянемо загальну постановку задачі впливу ЕМВ на біологічні системи і об'єкти. Узагальнення практичного досвіду багаторічних емпіричних досліджень дозволило визначити основні схеми, що реалізують механізм електромагнітного впливу (рис. 1.7). Залежно від частоти та інтенсивності ЕМП і особливостей біологічного об'єкта (розміру, форми і стану) застосовуються різні схеми збудження ЕМП для отримання оптимального його розподілу. Це можуть бути випромінювачі (антени) у вигляді аплікаторів, які накладаються на тіло людини або розташовуються поблизу нього (рис. 1.7, а). Місця розташування таких випромінювачів і їх кількість розраховуються таким чином, щоб створювати максимум інтенсивності ЕМП в місці розташування вогнища захворювання всередині або на поверхні тіла людини. Такі схеми застосовуються при термічному варіанті для терапевтичного застосування (різні види діатермії [10]) і лікування деяких видів онкологічних захворювань (гіпертермія [18]). У цьому випадку частоти ЕМВ відповідають десяткам і

сотня МГц, а їх вибір залежить від глибини розташування вогнища захворювання. З ростом частоти ЕМП (до 10 ГГц) і його інтенсивності (потужність понад 10 Вт) доцільно використовувати для концентрації ЕМП резонаторні системи, в яких формується поле з заданим розподілом (рис. 1.7, б).



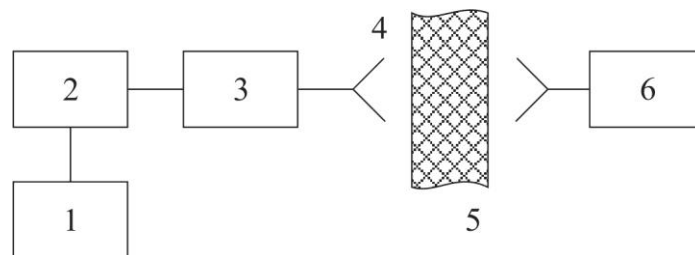
1 - випромінювачі ЕМП; 2 - шкірний покрив; 3 - жировий шар;
4 - м'язовий шар; 5 - кісткові утворення

а)



1 - джерело живлення; 2 - генератор НВЧ; 3 - пристрій узгодження;
4 - СВЧ резонатор; 5 - вимірювач пройшла НВЧ потужності

б)



1 - джерело живлення; 2 - генератор НВЧ; 3 - пристрій узгодження;
4 - СВЧ випромінювач; 5 - досліджуваний зразок; 6 - вимірювач пройшла НВЧ
потужності

в)

Рисунок 1.7 - Схеми впливу ЕМВ на біологічні системи

У КВЧ і терагерцевому діапазонах через малість геометричні розміри волноведущих структур застосовується відкритий спосіб опромінення біологічних систем за допомогою спеціальних типів випромінювачів (рупорні антени) (рис. 1.7, в). В цьому випадку в якості досліджуваних зразків можуть розглядатися або окремі частини і органи тіла людини або окремі біологічні об'єкти (клітини живої матерії або різні патогенні мікроорганізми).

Аналіз представлених схем показує, що для адекватного розуміння біологічної проблеми та математичного описи природи біологічного впливу ЕМІ слід вирішити два основні завдання. З одного боку, розглядається електродинамічна задача збудження ЕМП з вибирається необхідний розподіл поля [32], а з іншого - розглядається безпосередньо біологічне середовище і її уявлення на макро- або мікро рівнях з визначенням відповідних електрофізичних та біофізичних параметрів [33].

1.4 Математичні моделі електромагнітного поля

Для впливу ЕМВ на біологічні системи і об'єкти використовується широкий діапазон довжин хвиль: від акустичних до хвиль радіо і СВЧ діапазонів і далі до КВЧ і терагерцевого діапазонів (рис. 1.1). Основними параметрами ЕМВ, яка поширюється у вакуумі, є довжина хвилі і її частота. При поширенні ЕМВ у довільній середовищі дані параметри пов'язані між собою співвідношенням

$$\lambda = \frac{1}{f \cdot \sqrt{\varepsilon_0 \mu_0 \varepsilon(\vec{r}, t) \mu(\vec{r}, t)}} = \frac{c}{f \cdot \sqrt{\varepsilon(\vec{r}, t) \mu(\vec{r}, t)}}, \quad (1.6)$$

де $c = \frac{1}{\sqrt{\varepsilon_0 \mu_0}}$ – швидкість світла; $\varepsilon_0 = \frac{10^{-9}}{36\pi}$ – діелектрична постійна вакууму (Ф/м);

$\mu_0 = 4\pi \cdot 10^{-7}$ – магнітна постійна вакууму (Гн/м); $\varepsilon(\vec{r}, t)$ и $\mu(\vec{r}, t)$ – відносні діелектрична і магнітна проникності.

Для визначення розподілу напруженостей ЕМП в біологічному середовищі зі складною макроструктурою слід скористатися повною системою рівнянь Максвелла, яка може бути представлена у вигляді рівнянь Максвелла для вихрового ЕМП:

$$\nabla \times \vec{H}(\vec{r}, t) - \frac{\partial \vec{D}(\vec{r}, t)}{\partial t} = \vec{J}_{\text{стор}}, \quad (1.7)$$

$$\nabla \times \vec{E}(\vec{r}, t) = -\frac{\partial \vec{B}(\vec{r}, t)}{\partial t} \quad (1.8)$$

і матеріальних рівнянь для середовища:

$$\vec{D}(\vec{r}, t) = \varepsilon_0 \cdot \varepsilon(\vec{r}, t) \cdot \vec{E}(\vec{r}, t), \quad (1.9)$$

$$\vec{B}(\vec{r}, t) = \mu_0 \cdot \mu(\vec{r}, t) \cdot \vec{H}(\vec{r}, t), \quad (1.10)$$

де $\vec{E}(\vec{r}, t)$ – напруженість електричного поля (В / м); $\vec{H}(\vec{r}, t)$ – напруженість магнітного поля (А / м); $\vec{D}(\vec{r}, t)$ – індукція електричного поля (К / м²); $\vec{B}(\vec{r}, t)$ – індукція магнітного поля (В / м²); $\vec{J}_{\text{стор}}(\vec{r}, t)$ – щільність стороннього струму (А / м²).

Рішення системи рівнянь (1.7) - (1.10) слід проводити при відповідних граничних і початкових умовах. У повній постановці дана задача не має аналітичного рішення і вимагає застосування сіткового звичайно-різницевого методу (наприклад, методу кінцевих різниць [36,37]). Результатом такого рішення є визначення розподілу інтенсивності сумарного ЕМП (без поділу на можливі види коливання) у вигляді кількісного значення напруженостей і.

Серед більш простих підходів до моделювання ЕМП у вільному просторі, довільної середовищі або в електродинамічній структурі (резонаторі) є рішення хвильових рівнянь для електричної та магнітної компонент поля

$$\Delta \vec{E} + k_0^2 \vec{E} = 0, \quad (1.11)$$

$$\Delta \vec{H} + k_0^2 \vec{H} = 0, \quad (1.12)$$

де $k_0 = \frac{\omega}{c} \cdot \sqrt{\varepsilon(\vec{r}, t) \cdot \mu(\vec{r}, t)}$ – постійна поширення в біологічному середовищі;
 $\omega = 2\pi f$ – кругова частота.

Характер рішення рівнянь (1.11) і (1.12) залежить від параметрів електродинамічної системи і складності її граничних умов, а також електрофізичних параметрів самого середовища розповсюдження (біологічної системи).

1.5 Математичні моделі біологічних систем

Розробка математичних моделей біологічних систем являє собою складну біофізичну задачу, для вирішення якої слід узагальнити багатий практичний досвід, нові наукові ідеї і знання в галузі інформатики, біології, біофізики та медицини, а потім представити їх у вигляді зручному для математичного опису.

Для вибору параметрів, що характеризують електричні властивості тканин живих організмів (шкірний покрив, жирова тканина, м'язи, кров, печінку, кістки тощо), скористаємося результатами досліджень поведінки відносної діелектричної проникності і питомої електропровідності для різних частот [38-40].

Кількісні значення даних параметрів для різних частот ЕМП представлені на Таблиці.

Розглянемо особливості побудови математичних моделей біологічних систем для різних випадків впливу ЕМВ.

1.5.1 Випадок теплового впливу ЕМП

Розгляд теплового дії ЕМП проводиться на макро рівні. Ефективність перетворення електромагнітної енергії в теплову енергію і нагрівання біологічних тканин залежить від електрофізичних і фізіологічних параметрів тканини, а також від внутрішнього особливостей її будови з урахуванням патологічних змін, викликаних захворюванням. Для визначення внутрішньої структури тканини з вогнищем захворювання слід скористатися можливостями томографічного дослідження для побудови 2-D та / або 3-D зображень області її поразки. В результаті визначаються межі патологічного ураження, що дозволяє коректно задати граничні умови для вирішення електродинамічної задачі знаходження розподілу напруженостей $\vec{E}(\vec{r},t)$ і $\vec{H}(\vec{r},t)$ ЕМП і рішення рівняння теплопровідності для визначення температури нагрівання тканини.

Рішення електродинамічної задачі залежить від складності граничних умов досліджуваної області тканини і проводиться з використанням однієї з математичних моделей ЕМП, тобто чисельне або аналітичне рішення рівнянь (1.6) - (1.9) або (1.10) і (1.11) дозволяє отримати розподіл електричного поля $\vec{E}(\vec{r})$ в досліджуваній області для різних моментів часу.

Для отримання рівняння теплопровідності скористаємося законом збереження кількості тепла, що виділяється в досліджуваній області поразки. Наявність зображення тканини дає можливість встановити взаємозв'язок між кількістю поглиненої електромагнітної енергії та механізмами охолодження тканини за рахунок теплопровідності різних її областей, кровообігу і т.п. Для визначення залежності між енергією і швидкістю зміни в часі температури на одиницю об'єму опромінюваної тканини розглянемо вираз

$$\frac{d(\Delta T)}{dt} = \frac{1}{c} \cdot (P_{mo} + P_{ногл} - P_m - P_k), \quad (1.13)$$

де P_{mo} – потужність метаболічного теплоутворення (Вт / кг); $P_{ногл}$ – потужність, поглинена в тканинах (Вт / кг); P_m – потужність, що розсіюється за рахунок теплопровідності (Вт / кг); P_k – потужність, що відводиться системою кровообігу (Вт / кг); c_0 – питома теплоємність тканини (Дж / кг град)); $\Delta T = (T - T_0)$ – різниця між поточною температурою тканини і вихідної її температурою до початку нагрівання T_0 (кімнатна температура).

Для визначення потужності метаболічного теплоутворення скористаємося виразом [10]

$$P_{mo} = P_0 \cdot (1,1)^{\Delta T}, \quad (1.14)$$

де P_0 – початкова потужність метаболічного теплоутворення (Вт / кг).

Потужність, що поглинається тканиною при впливі зовнішнього ЕМП, дорівнює

$$P_{ногл} = \frac{\sigma}{\rho} \cdot |E(\vec{r})|^2, \quad (1.15)$$

де σ – питома провідність тканини в заданому обсязі (См / м); ρ – щільність тканини (кг / м³); $E(\vec{r})$ – середньоквадратичне значення напруженості електричного поля в об'ємі тканини (В / м).

Відводиться тепло від тканин за рахунок явища теплопровідності і наявності кровообігу в досліджуваній області. Потужність, відведена за рахунок теплопровідності тканини дорівнює

$$P_m = \frac{\lambda_T}{\rho} \cdot \Delta T, \quad (1.16)$$

де λ_T – питома теплопровідність тканини (Вт / (м град)).

Вважаючи, що кров надходить у досліджувану область з температурою $T_{к0}$, а покидає область з температурою $T_к$, то для потужності, що відводиться за рахунок кровообігу, можна скористатися виразом, отриманим в [10]:

$$P_к = \frac{k \cdot v_к \cdot c_к}{\rho_к} \cdot \Delta T_к, \quad (1.17)$$

де $\Delta T_к = T_к - T_{к0}$; $c_к$ – питома теплоємність крові; $\rho_к$ – об'ємна щільність крові, кг / м³; $v_к$ – швидкість течії крові; $k = 0,698$ – постійна величина.

Таким чином, для відомих форми і тканинного складу досліджуваної області організму з урахуванням фізичних і теплових характеристик тканин проводиться самоузгоджене рішення прямої електродинамічної задачі і задачі теплопровідності. Результатом рішення даної задачі є визначення температурного поля $T(\vec{r})$ в тканинах досліджуваної області.

1.5.2 Випадок інформаційного (нетеплового або специфічного) впливу ЕМП

Зниження інтенсивності ЕМВ до рівня, коли щільність потужності не перевищує 10 мВт / см² викликає зменшення теплового навантаження на біологічні тканини. В результаті температура в опромінюваних тканинах практично не змінюється (нагрівання опромінюваних об'єктів відповідає часткам градуса) [7]. З іншого боку, при впливі ЕМП такої інтенсивності КВЧ діапазону в організмі спостерігаються прояви «біологічних ефектів», яке розглядається як специфічний вплив (за визначенням акад. Девяткова Н.Д. [12])

з явно вираженим пороговим ефектом при щільності потужності $P_{пор} = 100$ мкВт / см². Ідея застосування ЕМВ НВЧ діапазону в своїй основі зводиться до його використання з метою передачі інформації (або керуючих сигналів) між клітинами всередині живих організмів [7,9,11-13,20]. У цьому випадку клітини живого організму розглядаються не тільки як об'єкти впливу ЕМП, але і як активні елементи організму, що формують відгук організму людини (теплове випромінювання) на дію зовнішнього ЕМП КВЧ діапазону. Вважаючи, що щільність потужності випромінювання живого організму в КВЧ діапазоні дорівнює $P_m = 10^{-18} - 10^{-19}$ Вт / см², і той факт, що $P_{пор} \gg P_m$ підтверджує ефективність впливу зовнішнього ЕМВ на біологічні об'єкти [41]. Зазначений факт доповнюється яскраво вираженою частотною залежністю біологічного відгуку і проявом резонансного ефекту. За сукупністю експериментально виявлених закономірностей взаємодії низькоінтенсивного ЕМВ НВЧ діапазону з біологічною системою можна розглядати як нетеплове вплив. У цьому випадку основну увагу слід приділяти модуляційно-часових параметрів і формі сигналів, що впливають на біологічні об'єкти живого організму, тобто вважаючи, що впливає ЕМП представляється у вигляді узагальненої функції

$$U(t) = \bar{U} \cdot [1 + \alpha(t)] \cdot \cos(\bar{\omega}t + \varphi(t)) , \quad (1.18)$$

де \bar{U} і $\bar{\omega}$ – середні постійні значення амплітуди і частоти ЕМП; $\alpha(t)$ – функція, що визначає закон зміни частоти і фази поля і є повільно змінюючою порівняно з $\cos(\bar{\omega}t)$ [42]. Як видно, найбільш перспективними сигналами для застосування в якості впливає неіонізуючого ЕМВ низької інтенсивності є частотно-модульовані коливання з програмованим законом модуляції. Застосування таких сигналів вимагає використання високостабільних по частоті електромагнітних коливань і застосування відповідних джерел таких коливань. Істотним недоліком сучасних експериментальних методів дослідження впливу ЕМВ НВЧ діапазону є не облік розгляду та впливу нестабільності частоти

випромінювання генераторів СВЧ на біофізичні процеси і по суті зводиться до застосування сигналів з випадковим поведінкою частоти (фази). Незважаючи на накопичений колосальний експериментальний досвід по впливу низькоінтенсивного ЕМВ різних частот і амплітуд, зараз неможливо дати однозначну відповідь про закони модуляції основного коливання з точки зору отримання прогнозованого поведінки біологічного об'єкта. У цьому плані роль математичних методів і моделювання для створення загальної теорії прогнозування ефектів впливу і відгуку біологічного об'єкта (клітини) стає визначальною. Розглядаючи клітину як самостійний біологічний об'єкт, як показано в [41], слід досліджувати амплітудно-частотні характеристики ЕМП, що генерується здоровими клітинами організму. Це дозволить порівняти ЕМП різних клітин (як здорових, так і хворих) і при необхідності проводити корекцію ЕМІ хворих клітин за допомогою зовнішнього КВЧ випромінювання, підбираючи потрібний закон модуляції його амплітуди і частоти (фази). У такому випадку проводиться синхронізація роботи складної коливальної системи, в якості якої розглядається живий організм як сукупність величезного числа клітин. Для математичного подання таких складних систем доцільно використовувати систему пов'язаних звичайних диференціальних рівнянь, в якій формалізований і визначено характер зв'язків між окремими складовими їх біологічними об'єктами (клітинами).

1.6 Моделювання впливу ЕМВ на біологічні об'єкти

Існують різні підходи до математичного моделювання впливу ЕМІ на біологічні системи і об'єкти [25-30]. Одним з таких підходів може служити стимуляція імунної системи організму зовнішнім низькоінтенсивних ЕМВ НВЧ діапазону. В якості математичної моделі такого процесу можна розглядати подальший розвиток математичної моделі противобактеріального імунної відповіді у вигляді системи звичайних диференціальних рівнянь (1.1) - (1.4) і

початкових умов (1.5) з урахуванням впливу додаткових факторів у вигляді ЕМВ [29].

На рис. 1.8 показано схематичне зображення імунної системи людини з урахуванням дії зовнішнього чинника у вигляді ЕМВ з фіксованою амплітудою і частотою.

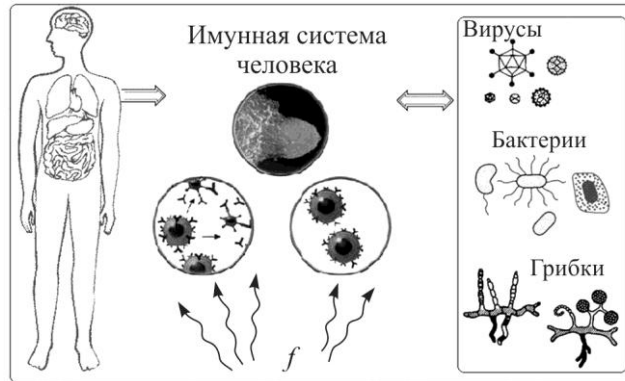


Рисунок 1.8 - Схематичне зображення імунної системи людини

Просту математичну модель інфекційного захворювання, що описує зміни в організмі людини з урахуванням дії різноманітних зовнішніх факторів, включаючи вплив патогенних мікроорганізмів і низкоінтенсивного ЕМВ НВЧ діапазону з заданими значеннями амплітуди E_m і частоти ω можна представити у вигляді:

$$\frac{dV}{dt} = (\beta - \gamma F)V + \alpha_v(E_m, \omega) \cdot V; \quad (1.19)$$

$$\frac{dC}{dt} = \xi(m)\alpha V(t - \tau)F(t - \tau) - \mu_c(C - C^*) + \alpha_c(E_m, \omega) \cdot C; \quad (1.20)$$

$$\frac{dF}{dt} = \rho_F C - (\mu_f + \eta\gamma V)F + \alpha_F(E_m, \omega)F; \quad (1.21)$$

$$\frac{dm}{dt} = \sigma V - \mu_m m + \alpha_m(E_m, \omega) \cdot m. \quad (1.22)$$

В якості початкових умов вирішення даної системи звичайних диференціальних рівнянь використовуємо висловлювання (1.5). Коефіцієнти α_V , α_C , α_F , і α_m в правій частині системи (1.19) - (1.22) є функціями амплітуди E_m і частоти ω ЕМП. Фізичний зміст даних коефіцієнтів і аналітичне подання їх функціональних залежностей для кількісного визначення значень вимагають окремого, більш докладного і глибокого вивчення. Система рівнянь формується окремо для кожного захворювання з урахуванням характеру проникнення антигенів в організм людини і дії ЕМВ.

Розглянута математична модель може бути використана як для моделювання та пошуку шляхів руйнування специфічних антигенів, так і для стимуляції імунної системи людини.

Для концептуального вивчення впливу ЕМВ НВЧ і КВЧ діапазонів різної інтенсивності на кров і її формені об'єкти (еритроцити, лейкоцити, тромбоцити) розглянемо загальний алгоритм побудови детермінованої математичної моделі такого процесу на мікро- і макро- рівнях. Розглянемо феноменологічну модель системи кровообігу людини, яка схематично представлена на рис. 1.9.

Як видно, кров у процесі циркуляції в тілі людини взаємодіє з іншими органами (взаємодія на макрорівні, рис. 1.9, а). З іншого боку, об'єкти крові мають електричним зарядом, реалізуючи електромагнітний механізм взаємодії (взаємодія на мікро- рівні, рис. 1.9, б). Залежно від поставленої мети така взаємодія може бути визначальним, а може бути другорядним. Ступінь зв'язку визначається коефіцієнтами зв'язку на макро D_{ij} і мікро C_{ij} рівнях, де $i = 1, 2, 3, \dots, M_{\text{об'єкт}}$ поточний індекс, що характеризує номер об'єкта крові як форменого, так і патогенних мікроорганізмів (віруси, бактерії, грибки); $j = 1, 2, 3, \dots, M_{\text{орган}}$ - Поточний індекс, що характеризує номер органу, з яким розглядається взаємодія.

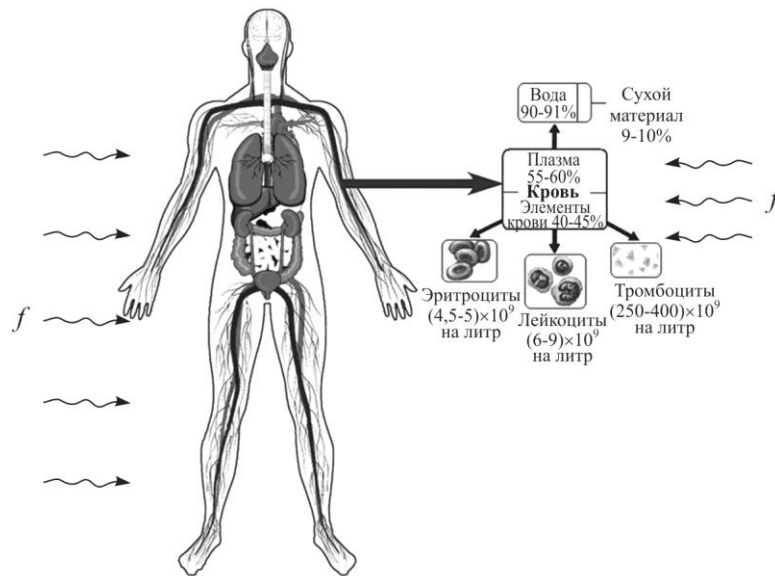


Рисунок 1.9 - Схематичне зображення кровеносної системи людини на макро (а) і мікро (б) рівнях

У результаті впливу ЕМВ на кров та її об'єкти, а також подальша взаємодія об'єктів крові з іншими органами призводить до зміни як загальної кількості таких об'єктів, так і їх розмірів і форми. У першому випадку для визначення динаміки зміни числа об'єктів крові розглядається система пов'язаних диференціальних рівнянь для кожного об'єкта. У цьому випадку вирішується задача Коші для системи пов'язаних звичайних диференціальних рівнянь. У початковий момент часу при $t = t_0$ задається початкове (початкове) кількість об'єктів крові і розглядається їх еволюція з урахуванням фактора зовнішнього впливу. Остаточно, система пов'язаних звичайних диференціальних рівнянь для формених елементів і можливих патогенних мікроорганізмів крові має вигляд:

$$\frac{dN_1}{dt} = (C_{11} + D_{11} + D_{12} + D_{13} + D_{14} + D_{15} + D_{16}) \cdot N_1 + C_{12} \cdot N_2 + C_{13} \cdot N_3 + C_{1x} \cdot N_x ; \quad (1.23)$$

$$\frac{dN_2}{dt} = C_{21} \cdot N_1 + (C_{22} + D_{21} + D_{22} + D_{23} + D_{24} + D_{25} + D_{26}) \cdot N_2 + C_{23} \cdot N_3 + C_{2x} \cdot N_x ; \quad (1.24)$$

$$\frac{dN_3}{dt} = C_{31} \cdot N_1 + C_{32} \cdot N_2 + (C_{33} + D_{31} + D_{32} + D_{33} + D_{34} + D_{35} + D_{36}) \cdot N_3 + C_{3x} \cdot N_x; \quad (1.25)$$

$$\frac{dN_x}{dt} = C_{x1} \cdot N_1 + C_{x2} \cdot N_2 + C_{x3} \cdot N_3 + (C_{xx} + D_{x1} + D_{x2} + D_{x3} + D_{x4} + D_{x5} + D_{x6}) \cdot N_x, \quad (1.26)$$

В системі рівнянь (1.23) - (1.26) вводиться додаткове рівняння для x - об'єкта крові. Під даним об'єктом розуміється будь-який зовнішній об'єкт крові (вірус, мікроб тощо), наявність якого вимагає проведення терапевтичного дій.

Як коефіцієнтів зв'язку D_{ij} і C_{ij} розглядаються складні функції багатьох змінних, включаючи і параметри ЕМП (амплітуду і частоту). Визначення коефіцієнтів зв'язку у вигляді аналітичної залежності являє собою складну задачу, що вимагає окремого розгляду.

Рішенням системи рівнянь (1.23) - (1.26) є тимчасові залежності числа об'єктів крові. Динаміка їх зміни дозволяє прогнозувати досягнення позитивного терапевтичного ефекту.

Для дослідження змін розмірів і форми об'єктів крові як наслідок зовнішнього електромагнітного впливу слід використовувати додатково математичні моделі самих об'єктів крові (наприклад, модель еритроцита [33] і т.д.).

1.7 Цифровий голографічний інтерференційний мікроскоп

Як згадувалося в роботі [4], одна з переваг голографічних методів вивчення фазових мікрооб'єктів полягає в можливості реалізації всіх методів інтерферометрії в одному пристрої. Модифікована оптична схема цифрового голографічного інтерференційного мікроскопа наведена на рис. 1.10.

По аналогії з [4], в якості когерентного джерела випромінювання в мікроскопі використовується HeNe лазер з довжиною хвилі випромінювання 0,63 мікромметра. Випромінювання лазера 1 напівпрозорим дзеркалом 5 ділиться на два пучка: об'єктний і опорний. Об'єктний пучок проходить через об'єктив мікроскопа 11 і направляється на фотопластинку (голограму) 12. Опорний пучок проходить через коліматор 9 і також прямує на голограму 12. Для повороту променів в схему введені дзеркала 3 і 4. Опорний пучок грає у схемі допоміжну роль. Він необхідний для запису і відновлення по голограмі об'єктної хвилі. На фотопластинці 12 реєструється голограма об'єктного пучка в відсутність досліджуваного препарату перед об'єктивом мікроскопа. Проявлена і відфіксована голограма повертається у вихідне положення. Спеціальний тримач дозволяє переміщати голограму щодо її первісного стану. Голограма 12 розташована між об'єктивом 11 і окуляром 13. Так як на голограмі реєструється неспокійованого зображення, то таке розташування голограми дозволяє збільшити поле зору мікроскопа за рахунок можливості переміщення окуляра в площині, ортогональної напрямку спостереження і проводити додаткову фокусування по глибині спостережуваної сцени. Досліджуваний препарат поміщається перед об'єктивом мікроскопа. При одночасному поширенні відновленої по голограмі за допомогою опорного пучка не обуреної об'єктної хвилі і обуреної об'єктної хвилі, що пройшла через досліджуваний препарат, за допомогою окуляра 13 можна спостерігати картину їх інтерференції, тобто інтерферограма. Необхідний період фонових інтерференційних смуг для отримання інтерферограми в нескінченно широкій смузі або в смугах кінцевої ширини формують поперечним зсувом голограми з її вихідного положення за допомогою мікрометричних гвинтів держателя.

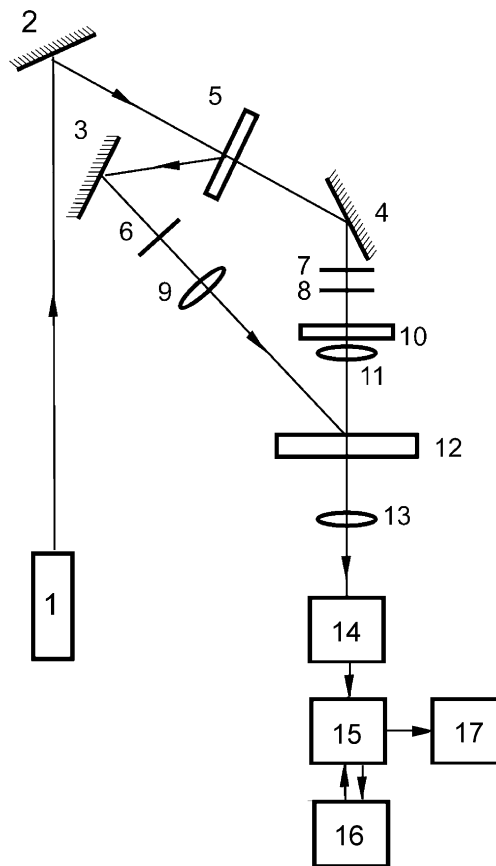


Рисунок 1.10 – Оптична схема цифрового голографічного інтерференційного мікроскопу: 1 - лазер; 2, 3, 4 - дзеркала; 5 - напівпрозоре дзеркало; 6, 7, 8 - поляризатори; 9 - коліматор; 10 - досліджуваний препарат; 11 - об'єктив; 12 - голограма; 13 - окуляр; 14 - блок реєстрації зображень (цифрова відео камера); 15 – блок порівняння; 16 – блок; 17 – блок

Зареєстрована для даного об'єктиву голограма може використовуватися постійно при роботі з даним об'єктивом. Як відомо, якість і контраст голограми і інтерферограми визначаються такою величиною, як глибина модуляції інтерференційної картини $\mu = \frac{2 \cdot \sqrt{b} \cdot \cos \phi}{1 + b}$, де b – це відношення інтенсивностей інтерферируючих хвиль, а ϕ – кут між напрямками лінійної поляризації цих хвиль. Як випливає з цього виразу, максимальне значення $\mu = 1$ отримуємо в тому випадку, коли інтерферируючі хвилі мають однакові інтенсивності і поляризації. Для вирівнювання інтенсивностей і поляризацій взаємодіючих хвиль в схему введені поляризатори 6, 7, 8. Поляризатори 6 і 8 задають однакові

напрямки площині лінійної поляризації взаємодіючих хвиль, а обертанням площини поляризації поляризатора γ вирівнюються інтенсивності взаємодіючих хвиль.

Інтерферограмми досліджуваних мікрооб'єктів реєструються за допомогою блоку реєстрації зображень 14, що містить цифрову камеру. Оцифровані інтерферограмми передаються в комп'ютер для подальшої обробки.

1.8 Висновки к розділу 1

Таким чином, представлений докладний аналіз впливу ЕМІ широкого діапазону частот і інтенсивностей на біологічні системи і об'єкти. Проведений аналіз фундаментальних літературних джерел показав проблеми, які супроводжують експериментальним методам дослідження впливу ЕМІ на біологічні об'єкти. Показана і обґрунтована необхідність в математичному моделюванні процесів впливу ЕМІ різної інтенсивності на біологічні системи та об'єкти з урахуванням їх подання на макро і мікрорівнях. Показані алгоритми побудови математичних моделей різних захворювань з урахуванням зовнішнього електромагнітного впливу. Отримала подальший розвиток математична модель інфекційного захворювання, що враховує вплив на організм людини різних зовнішніх факторів і особливості реакції імунної системи.

Розглянуто сучасні математичні моделі крові. На основі пророблених досліджень була запропонована концепція побудови математичної моделі крові, що відбиває еволюційний цикл змін, що відбувається в організмі людини при впливі зовнішніх факторів різної природи. Запропоновано базова феноменологическая модель кровоносної системи і отримано математичну її модель, що описує поведінку крові як активної системи на мікро- і макро рівнях за умови зовнішнього електромагнітного впливу. Застосування даної математичної моделі дозволить кількісно визначити еволюцію об'єктів крові в процесі впливу ЕМВ і спрогнозувати наслідки такого впливу.

2 ВЗАЄМОДІЯ ЛАЗЕРНОГО ВИПРОМІНЮВАННЯ З БІОЛОГІЧНИМИ МІКРО ТА НАНО ОБ'ЄКТАМИ

2.1 Дослідження фізичних умов механічного впливу лазерного випромінювання на біологічні об'єкти

Для реалізації дослідження фізичних явищ, процесів та технічних умов, які забезпечують механічну взаємодію лазерного випромінювання з діелектричними нанооб'єктами необхідно провести:

- Дослідження фізичних умов формування просторових характеристик оптичного випромінювання необхідних для маніпулювання нанооб'єктами та контролю їх переміщенням у просторі;
- Дослідження фізичних умов ефективного механічного впливу лазерного випромінювання на нанооб'єкти.
- Розробка експериментального макету лазерного «пінцету» для досліджень умов механічного пересування мікро- та нанооб'єктів в різноманітних середовищах.

Протягом останніх років були ретельно вивчені теоретично фізичні процеси механічної взаємодії лазерного випромінювання з об'єктами біологічного походження. Подальший логічний розвиток цієї тематики полягає в звертанні уваги на фізичні процеси, що дозволяють реалізувати взаємодію лазерного випромінювання з діелектричними об'єктами. Для проведення таких досліджень необхідні джерела лазерного випромінювання. Для використання необхідно провести дослідження умов та технічні реалізації формування оптичного пучка.

Таким чином, суттєвою сучасною проблемою при будівництві лазерного «пінцету» є використання оптичних приладів з наступними параметрами:

- Потужність лазерного випромінювання, не менш ніж – 100 мВт;

- Довжина хвилі лазерного випромінювання – 0,44 мкм; 0,53 мкм; 0,64 мкм;

Робота по проекту була виконана у 3 етапи:

- на першому етапі було виконано розвиток моделі взаємодії біологічного середовища з лазерним випромінюванням, умови взаємодії лазерного випромінювання з нанооб'єктами;

- другий етап роботи включав дослідження фізичних умов реалізації захвату нанооб'єктів лазерним випромінюванням. Були проведені дослідження методів формування лазерного випромінювання, які забезпечують механічну дію на діелектричні нанооб'єкти;

- на фінальному етапі була проведена розробка експериментального зразка лазерного «пінцету» для переміщення нанооб'єктів у просторі та методика його застосування.

За результатами виконання робіт, в яких приймали участь студенти, за темою НДР, були захищені 3 бакалаврські та 4 магістерські роботи.

В процесі виконання НДР планується отримати нові наукові знання у галузі взаємодії імпульсного та неперервного лазерного випромінювання з нанооб'єктами.

По-перше, будуть отримані результати по формуванню градієнтної сили для лазерного випромінювання.

По-друге, будуть розраховані умови управління у просторі діелектричними об'єктами за допомогою лазерного випромінювання.

По-третє, буде отримано зв'язок стабільності лазерного проміння з умовами захвату нанооб'єкта лазерним пучком. Для цього буде розроблена модель мікростабільності осі лазерного пучка в оптичній системі.

У четвертих, стабільність градієнтної сили залежить від стабільності потужності випромінювання. При дослідженнях будуть знайдені нові умови стабілізації потужності випромінювання, встановлені також умови впливу нестабільності частоти (довжини хвилі) випромінювання за короткий інтервал

часу на нестабільність градієнтної сили. Буде встановлена залежність між поляризацією та градієнтною силою, та буде встановлено зв'язок між тривалістю імпульсів і ефективною градієнтною силою. Таким чином, у результаті виконання проекту буде встановлене зв'язок між параметрами імпульсного фемтосекундного випромінювання та умовами захвату нанооб'єктів.

2.2.1 Основні принципи роботи лазерного пінцету (стислий історичний екскурс)

У загальному випадку при взаємодії світла з речовиною (віддзеркалення, заломлення або поглинання світла), світло передає частину свого імпульсу матеріального тіла, що еквівалентно тиску світла на цей матеріальний об'єкт. У 1899 році, за допомогою крутильних терезів, які помістили у вакуумну камеру, російському фізику П. Н. Лебедєву вдалося на фоні конвекційних потоків залишкового газу виділити і виміряти величину тиску світла на механічний пристрій. Багато десятиліть експерименти Лебедєва залишалися єдиним підтвердженням реальності умов, за яких оптичне випромінювання впливає на матеріальні об'єкти. При цьому використовувалось природне світло з малим рівнем щільності потужності. В експериментах П.Н. Лебедєва було зафіксовано тиск світлового поля, багато десятиліть ці результати розглядалися тільки як підтвердження фундаментальним фізичним законам. Практичного застосування ці результати не знайшли, що було вочевидь в силу малості реєстрованих величин.

Майже через сімдесят років, в 1968 році, В.С. Летохов [49] теоретично показав, за яких умов і за якими параметрами лазерного випромінювання, можливо переміщати атоми в просторі. Спочатку, Летохов аналізував ситуацію, яким чином за допомогою лазерного випромінювання можливо було б охолоджувати температуру атомів або іонів. Це була фундаментальна задача,

вирішення якої дозволило б створювати природні частотні репери, властивості яких не залежать від теплових швидкостей атомів (іонів).

У 1987 році А. Ешкін [50] теоретично і експериментально продемонстрував можливість утримання діелектричної частинки в заданій області простору за допомогою сфокусованого пучка лазерного випромінювання. Свої досліди він проводив з маленькими пластмасовими частками (кульками з латексу), що знаходились у воді.

Дослідження явища взаємодії лазерного випромінювання з мікро- і нанооб'єктами пов'язане з необхідністю враховувати співвідношення між довжиною хвилі застосовуваного випромінювання і характерним розміром об'єкта. З електродинаміки відомо, що при вивченні процесів розсіювання електромагнітних хвиль на локальних матеріальних об'єктах, слід виділяти два граничних випадки. Перший, режим Мі-розсіювання, коли довжина хвилі випромінювання багато менше розмірів частинки і можна використовувати наближення геометричної оптики. Другий, відповідає умові, коли довжина хвилі багато більше розмірів об'єкта. Якщо розмір частки менше довжини хвилі електромагнітного випромінювання (Релеєвське розсіювання), то її можна розглядати як точковий диполь в електромагнітному полі.

Умови захоплення оптичним випромінюванням мікрочастинок, розміри яких можуть бути кілька десятків мікрон, описуються режимом Мі-розсіювання. В основі механічного впливу оптичного випромінювання на матеріальні об'єкти лежить закон збереження імпульсу. На частку, яка знаходиться в пучку випромінювання лазера, діють радіальні сили світлового тиску. Оскільки показник заломлення частки звичайно вище, ніж у навколишнього середовища, тому заломлені промені відхиляються ближче до її осі, змінюючи напрямок свого поширення. За законом збереження імпульсу повинна виникнути сила, що компенсує це відхилення променя (воно еквівалентно передачі імпульсу частці або тиску світла). Якщо показник заломлення частинки буде менше

показника заломлення середовища, градієнтна сила буде направлена від осі пучка, і частка виштовхуватиметься з нього.

Сила, з якою світло діє на навколишні об'єкти, невелика, але її виявляється достатньо, щоб утримати наночастинку у фокусі лазерного променя. Як тільки частинка опинилася у фокусі, її можна переміщати за допомогою зміщення лазерного променя. З іншого боку, можливість захоплення частки і її утримання в заданому місці тривимірного простору дозволяє припустити, що можна побудувати просторовий розподіл наночастинок, які можуть застосовуватися в якості елементів оптичної пам'яті. Тут слід зазначити, що застосовувати можна як діелектричні частинки, так і металеві. Вочевидь, що для діелектричних частинок, через малість величини діелектричної проникності, сила впливу менше, ніж для металевих частинок. Що стосується розмірів частинки, то чим менше частка, тим менше атомів вносять свій внесок у індукований дипольний момент. Тим менше її поляризованість, а значить, і менше сила, яка діє на частку. Тому, потужність лазера, необхідна для полонення діелектричних частинок, зростає зі зменшенням їх розміру [49].

При взаємодії електромагнітного випромінювання з металевими наночастинками, електрони провідності узгоджено зміщуються відносно позитивно заряджених іонів решітки. Флуктуації щільності заряду створюють електричне поле, що прагне відновити електронейтральність. Величина повертаючої сили пропорційна величині зсуву, так само як і для типового осцилятора, тому можна казати про наявність власної частоти колективних коливань електронів у частці. Якщо частота коливань падаючого світла збігається з власною частотою коливань вільних електронів поблизу поверхні металевої частки, спостерігається різке збільшення амплітуди коливання «електронної плазми», квантовим аналогом якої є квант.

Спочатку вважалося, що управляти металевими наночастинками за допомогою лазерного пінцета неможливо через високий коефіцієнт відбиття. У

1992 році К. Свобода і С. Блок провели експеримент [51], який показав, що 40 нм золоті наночастинки можливо захопити оптичним пінцетом з довжиною хвилі випромінювання, яка не співпадає з довжиною хвилі поверхневого плазмонного резонансу. У порівнянні з діелектричними частками такого ж розміру, золоті поляризуються сильніше (зумовлено високою величиною діелектричної проникності), а значить, для їх утримання оптичним пінцетом потрібні більш сильні поля. Для того щоб знизити потужність лазерних джерел для утримання металевих наночастинок, можливо спробувати змінити іммерсійну рідину, в якій знаходяться частинки. Зі збільшенням показника заломлення іммерсійної рідини жорсткість пастки зростає, це означає, що можна буде використовувати менш потужні лазери, але при цьому утримувати частку пінцетом з тією ж силою. Слід зазначити, що знаходження в полі лазера безперервної дії використання сильно поглинаючих металевих частинок призводить до значних теплових ефектів.

Грунтуючись на отриманих до теперішнього часу знаннях про механічні взаємодії лазерного випромінювання з мікрооб'єктами, у цій статті обговорюються теоретичні результати створення двовимірної оптичної сітки для фіксації положення мікрочастинок. Подібна сітка із захопленими мікрочастками може стати основою для системи запам'ятовування і зберігання інформації в оптичному комп'ютері.

2.1.2 Фізичні умови і вихідні теоретичні умови

Створення оптичного комп'ютера до теперішнього часу не знайшло свого вирішення. Це пов'язано з проблемою формування оптичної системи запам'ятовування. Геометрично правильне розміщення мікрооб'єктів. Основні ідеї формування. Оскільки захоплення і управління становищем мікрооб'єктами здійснюється за допомогою сфокусованого лазерного пучка, то виникає припущення, що можливе формування просторової оптичної сітки, яка буде

стаціонарно утримувати мікрочастинки. Можливості застосування захоплених мікрочасток можуть бути різними, в тому числі і для формування пам'яті оптичного комп'ютера.

Частка повинна знаходитися в рідині, щоб броунівська сила взаємодії з рідиною доповнювала градиентну силу, сформовану в пучку.

Умовою виникнення плазмонного резонансу є звернення в нуль дійсної частини знаменника поляризуємості металевої наночастки. Якщо частота лазера точно налаштована на частоту ППР, відбудеться різке збільшення частоти поверхневої моди частинки, таким чином резонансно збільшуючи силу впливу зовнішнього поля на частку. На частоті плазмонного резонансу металева частка виштовхується з області фокуса. Це пояснюється перевагою однієї з двох сил – від'ємної градиентної сили або сили розсіювання. Мода може відчувати зміну фази коливань на 180° , таким чином миттєво змінюючи напрямок дії сили. Зменшуючи частоту лазера, можна домогтися затягування металевої наночастки в область фокуса або її виштовхування.

Перш ніж пояснити прагнення частинок до фокусу, згадаємо, що промінь світла - це електромагнітна хвиля, і чим більше інтенсивність світла, тим більше напруженість електричного поля в поперечному перерізі променя. Тому у фокусі середньоквадратична величина напруженості електричного поля може збільшуватися у багато разів. Таким чином, електричне поле фокусованого світлового променя стає неоднорідним, збільшуючись по інтенсивності за мірою наближення до фокусу.

Нехай частинка, яку ми бажаємо утримати за допомогою оптичного пінцета, зроблена з діелектрика. Відомо, що зовнішнє електричне поле діє на молекулу діелектрика, переміщуючись всередині неї різнойменні заряди в різні боки, в результаті чого ця молекула стає диполем, який орієнтується уздовж силових ліній поля. Це явище називають *поляризацією* діелектрика. При поляризації діелектрика на його протилежних по відношенню до зовнішнього

поля поверхнях з'являються різнойменні і рівні за величиною електричні заряди, звані *пов'язаними*.

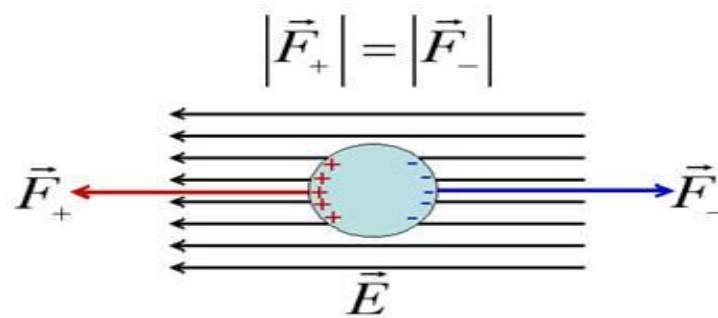


Рисунок 2.1 - Схематичне зображення кулястої частки, що знаходиться в однорідному електричному полі напруженістю E

Знаками "+" і "-" показані пов'язані заряди, що виникли на поверхні частинки при її поляризації. Електричні сили, що діють на позитивні ($F+$) і негативні ($F-$) пов'язані заряди, однакові.

Нехай наша частка з діелектрика знаходиться в світловому промені далеко від фокусу. Тоді можна вважати, що вона знаходиться в однорідному електричному полі (див. рис. 2.1). Так як напруженість електричного поля ліворуч і праворуч від частки одна і та ж, то й електричні сили, що діють на позитивні ($F+$) і негативні ($F-$) пов'язані заряди, теж однакові. В результаті, частка, що знаходиться в однорідному електричному полі залишається нерухомою.

Нехай тепер наша частка знаходиться поруч з областю фокусу, де напруженість електричного поля (густина силових ліній) поступово зростає (крайня ліва частка на рис. 2.2) при русі зліва направо. У цьому місці частинка теж буде поляризована, але електричні сили, що діють на позитивні ($F+$) і негативні ($F-$) пов'язані заряди, будуть різні, тому напруженість поля зліва від частки менше, ніж справа. Тому на частку буде діяти результуюча сила, спрямована вправо, до області фокусу.

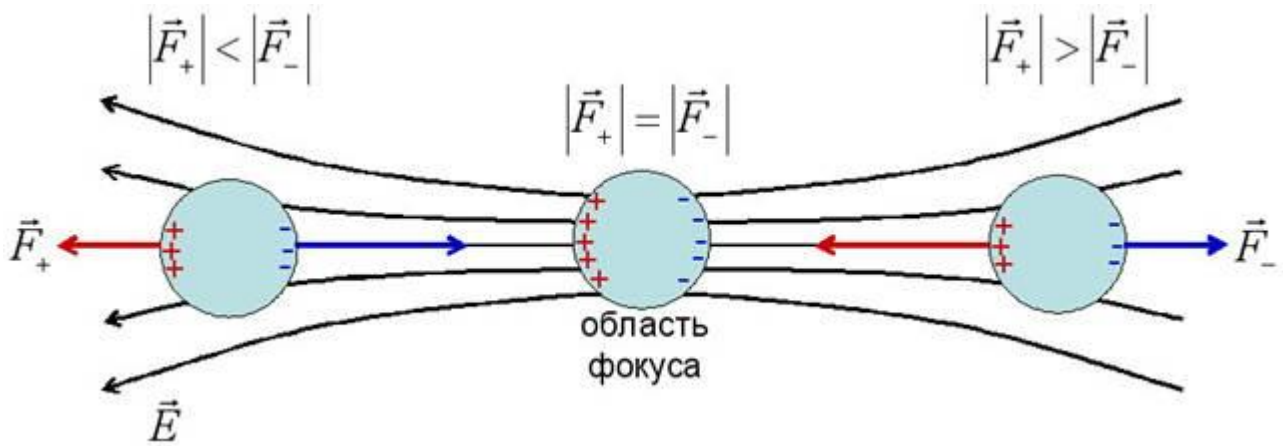


Рисунок 2.2 - Схематичне зображення трьох кулястих часток, що знаходяться в неоднорідному електричному полі сфокусованого світлового променя поблизу області фокуса

Знаками "+" і "-" показані пов'язані заряди, що виникли на поверхні частинок при їх поляризації. Електричні сили, що діють на позитивні (F_+) і негативні (F_-) пов'язані заряди, викликають рух часток у напрямку до області фокуса.

До краю правої частки (див. рис. 2.2), що знаходиться з іншого боку фокуса, діятиме результуюча, спрямована вліво, до області фокуса. Таким чином, всі частки, що опинилися в сфокусованому промені світла, будуть прагнути до його фокусу, (як маятник прагне до положення рівноваги). Якщо діаметр частки набагато більше довжини хвилі світла, то для такої частки стають справедливими закони геометричної оптики, а саме, частка може заломлювати світло, тобто змінювати його напрямок. У той же час, відповідно до закону збереження імпульсу сума імпульсів світла (фотонів) і частки повинна залишатися постійною. Іншими словами, якщо частка заломлює світло, наприклад, праворуч, то сама вона повинна рухатися ліворуч.

Слід зазначити, що інтенсивність світла в лазерному промені максимальна уздовж його осі і поступово падає при віддаленні від неї. Тому, якщо частка знаходиться на осі світлового пучка, то число фотонів, відхилених

нею ліворуч і праворуч, однаково. В результаті, частка залишається на осі (див. Рис. 2.3, b).

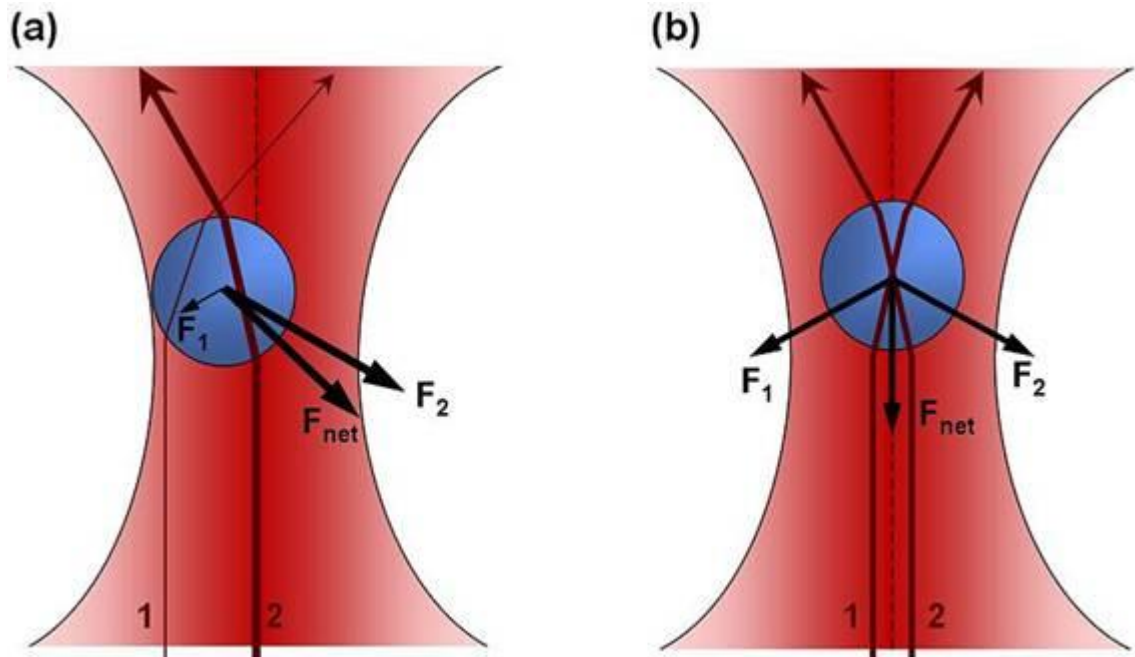


Рисунок 2.3 - Схематичне зображення кулястої частки, що знаходиться в зфокусованому пучку світла ліворуч від його осі (a) і на його осі (b).

Інтенсивність червоного забарвлення відповідає інтенсивності світла в даній області променя. 1 і 2 - промені, заломлення яких відображено на малюнку, а товщина відповідає їх інтенсивності. F_1 і F_2 - сили, що діють на частку згідно закону збереження імпульсу, при зламленості променів 1 і 2, відповідно. F_{net} - результуюча F_1 і F_2 .

У випадках, коли частка зміщена ліворуч відносно осі світлового променя (див. Рис. 2.3, a), число фотонів, відхиляючихся ліворуч (див. промінь 2 на рис. 2.3, a), перевищує їх число, відхиляємих праворуч (див. промінь 1 на рис. 3a). Тому виникає складова сили F_{net} , спрямована до осі променя, праворуч.

Вочевидь, що на частку, зміщену праворуч від осі променя, діятиме результуюча, спрямована ліворуч, і знову до осі даного променя. Таким чином, всі частки, які опинилися не на осі променя, будуть прагнути до його осі.

Одним з актуальних напрямків сучасної оптики та лазерної фізики є оптична маніпуляція. Цей напрямок засновано на силах тиску лазерного

випромінювання, і, в першу чергу, пов'язано з можливістю гострого фокусування (фокусування в пляму, розміри якого порівнювані з розмірами довжини світлової хвилі) лазерного пучка. Виконавши прості оцінюючі розрахунки, в 1970 Артур Ешкін показав, що сила світлового тиску, вироблена лазерним пучком, досить велика для захоплення, утримання і переміщення в різних середовищах нейтральних часток мікронних розмірів [49]. Сьогодні ці розрахунки наводяться в підручниках оптики [50]: аргонний лазер ($\lambda = 0.5$ мкм) безперервної дії потужністю $P = 1$ Вт можна сфокусувати на майданчик діаметром від λ до $10 \cdot \lambda$; якщо помістити у фокусі дзеркально відображуючу кульку радіусом $r = \lambda$, то сила світлового тиску, що діє на неї дорівнює

$$F = 2P/c = 2 \cdot 1 \text{ [Н} \cdot \text{м/с]} / 3 \cdot 10^8 \text{ [м/с]} \sim 10^{-8} \text{ [Н]}.$$

Нехай щільність матеріалу ρ , з якого зроблена кулька, дорівнює 1 г / см^3 , тоді її маса буде дорівнюється

$$m = 4/3 \pi \lambda^3 \rho \sim 4 \cdot (0.5 \cdot 10^{-6} \text{ м})^3 \cdot 10^3 \text{ кг/м}^3 \sim 10^{-15} \text{ кг}.$$

Отже, згідно 2-му закону Ньютона під впливом сили світлового тиску кулька придбає величезне прискорення $a = F / m \sim 10^7 \text{ м/с}^2 = 10^6 \cdot g$. Таким чином, за допомогою світлового тиску лазерного пучка дійсно можна управляти рухом мікроскопічних часток. Ешкін здійснив і експериментальну перевірку своїх розрахунків - він спостерігав за рухом прозорих латексних часток, зважених у воді, викликаним тиском лазерного випромінювання потужністю 10 мВт.

Частки рухалися в напрямку світлового пучка зі швидкостями, які добре узгоджуються з проведеними чисельними оцінками сил світлового тиску. У цих же експериментах було виявлено та наявність якоїсь поперечної сили, яка притягувала частинку до центру пучка - в область найбільшої інтенсивності (коли частинка, рухаючись у напрямку світлового пучка, досягала межі комірки, вона залишалася в центрі світлового пучка; якщо світловий пучок перекривався, частинки починала йти з центру; потім світло вмикалось знову, і,

якщо частка опинялась в смузї світлового пучка, вона тут же притягувала до його центру). Пояснити принцип виникнення поперечної сили захоплення можливо, в рамках геометричної оптики, виходячи із закону збереження імпульсу [див. рис. 2.4]. Світловий пучок являє собою потік квантових часток - фотонів, кожен з яких несе імпульс $p = h / \lambda$, де h - постійна Планка, λ - довжина хвилі випромінювання. Коли світловий пучок падає на частку, показник заломлення якої більше показника заломлення навколишнього середовища, світлові промені заломлюються (частки діють подібно сферичній лінзі), тобто напрямок світлових променів змінюється, а значить, змінюється і напрямок імпульсу. Зміна напрямку імпульсу світлового потоку за законом збереження імпульсу (для системи "світло-частка") і згідно з другим законом Ньютона призводить до виникнення сил, що діють на частку. Ці сили спрямовані уздовж зміни напрямку імпульсу. Так як інтенсивність світла на осі пучка вище, ніж на периферії, то $F_a > F_b$, і, отже, рівнодіюча сила спрямована до осі пучка.



Рисунок 2.4 - Схема виникнення радіальних сил захоплення [49]

Таким чином, за допомогою двох пересічних пучків можна реалізувати так звану "світлову пастку". Такі світлові пастки першим реалізував А.Ешкін. Артур Ешкін виконав і перші експерименти з оптичною левітації, в яких сила світлового тиску підіймає і утримує частки мікронних розмірів у повітрі проти сили тяжіння, а також розробив ряд інших методів захоплення і утримання лазерними пучками часток мікронних розмірів [49]. Однак справжній прорив

був здійснений в 1986 році, коли А. Ешкін, С. Чу і їхні колеги з Bell Telephone Laboratories в США створили тривимірну оптичну пастку, використовуючи один (єдиний) точно сфокусований лазерний пучок [51,52]. Сьогодні така однопучкова пастка широко відома як оптичний або лазерний пінцет.

Принцип дії тривимірної оптичної пастки заснований на комбінації поперечної і поздовжньої сил захоплення. Причому поздовжня сила спрямована до фокусу пучка. Дійсно, тепер, коли на частку спрямований остросфокусований пучок світла, прозора частка частково колімірує випромінювання що проходить. Зміна напрямку променів, що поширюються не вздовж осі пучка, призводить до того, що їх імпульс також змінюється. В результаті виникають сили, що діють на сферу у напрямку до фокусу (див. Рис. 2.4).

Однак атом - все-таки куля вельми своєрідна. Тісно притулившись один до одного в центрі протони і нейтрони немов би броньовані, оточені багаточисловою електронною оболонкою. І передати імпульс руху атому лазерний пучок, відповідно до законів квантової фізики, можливо лише в тому випадку, коли той поглине фотон. Частка світла як би проб'є броню.

Електрон, що поглинув фотон, перескакує на іншу, вищу, орбіту, тримається там певний час, а потім повертається назад, віддаючи отриману енергію знову-таки у вигляді кванта. Стало бути, атом повинен зупинитися, подібно до того як зупиняється куля на сукні більярдного столу через тертя або натрапивши на іншу кулю. Але оскільки і поглинання і випускання фотона - процеси дуже швидкоплинні і різноспрямовані, ми можемо й не помітити, зрушився чи стіл з місця. Скільки б не атакували його квантами світла, він би компенсував отриманий імпульс викидом такої ж кількості енергії в протилежний бік.

На наше щастя, на практиці справа йде не настільки безнадійно. Атакувати атом-кулю ми можемо, обстрілюючи його з лазерної "гармати" з певної позиції. Випускання ж фотона відбувається у випадковому напрямку,

оскільки електрони безладно обертаються по своїх орбітах. І безліч фотонів, обсіпаючи, немов з кулемета, що рухається, припустимо, їм назустріч атомна куля, здатні пригальмувати його рух.

Отже, завдання начебто вирішено. Залишилося відповісти на питання, яким чином вибрати з безлічі атомів той єдиний, який рухається строго назустріч променю, зосередити фотонний потік саме на ньому? Реально адже жоден снайпер не зміг би зловити атом на приціл.

Однак ще Летохов і його колеги з'ясували, що "стріляти" по атомній кулі має сенс лише світловим випромінюванням строго певної довжини хвилі або частоти, що потрапляють в резонанс із власними коливаннями електронів на своїх орбітах. Інакше фотони будуть поглинатися електронами, а тому, атом і не отримає імпульсів гальмування. Причому спектральний резонанс заодно здатний забезпечити і надійне прицілювання по тих атомах, які рухаються тільки назустріч лазерному пучку. Допоможе тут і ефект Доплера.

Останній, як ви пам'ятаєте, в простому випадку полягає в наступному. Гудок локомотиву, який не рухається, що рухається нам назустріч або віддаляється від нас, буде здаватися нам різної висоти, або частоти. А коли так, то і частоту лазерного пучка можна підібрати з таким розрахунком, що фотони, які випускаються, будуть сприйматися, поглинатися лише атомами, що рухаються тільки назустріч. І ніякими іншими! На цій дотепній здогадці в 1975 році американськими фізиками і була побудована лазерна пастка для нейтральних атомів. Вони поставили попарно шість лазерів, розташували їх по трьох взаємно перпендикулярних осях, і пригальмували таким чином відразу безліч атомів, що рухаються в різні боки.

А ще 10 років потому, удосконаливши подібну систему, А.Ешкін, Л.Холлберг, Дж.Бьорхольм, А.Кейбл і С.Чу з'ясували, що атоми натрію в експерименті охолоджуються до 0,00024К! Причому в пастці утворювалася середина, що працювала як своєрідний клей - атоми загрузли в ній, немов мухи в цукровому сиропі, або, як назвали вчені, - в "оптичній патоці". Однак така

"патока", як незабаром з'ясувалося, виявилася не дуже гарним "клеєм" - атоми-кулі затримувалися в пастці-лузі не довше, ніж на 0,5 с. Довелося продовжити роботу. За справу взялися вчені різних країн. Незабаром, наприклад, У.Філіпс і його колеги виявили, що при певних умовах "оптичну патоку" можна використовувати для охолодження атомів до температур майже цілковитого нуля.

Таке відкриття, в свою чергу, спонукало групу К.Коен-Таннуджи з французького колежа і групу С.Чу з Станфорда створити нову теорію лазерного охолодження, засновану на ще більш складному взаємозв'язку між атомами і їх взаємодією зі світловими полями. А це знову-таки дозволило удосконалити пастку, знизивши до 1995 року температуру в ній до 2 - 3 мікрокельвінів.

Зупинені в своєму русі атоми стали настільки чутливі до світлового тиску, що тепер лазерними пучками можна діяти немов щипцями, просуваючи мікрооб'єкти, що називається, поштучно. На малюнку, наприклад, ви можете побачити схему "оптичного пінцета" - установки, що дозволяє переміщати мікроби, навіть окремі частини у середині живої клітини, не руйнуючи її! Крім того, нові методи виявилися вельми перспективні для створення унікальних атомних інтерферометрів, особливо точних атомних годин, гіроскопічних датчиків нового покоління для космічної навігації.

За все це Нобелівський комітет і присудив премію Стівену Чу з Станфордського університету, Вільяму Філіпсу з Мерілендського національного інституту стандартів і технологій, а також Клоду Коен-Таннуджи з Коллежу де Франс.

2.2 Основні принципи оптичного захоплення

2.2.1 Сили оптичного захоплення

2.2.1.1 Імпульс світла. Спосіб оптичного захоплення і маніпулювання малими нейтральними частинками за допомогою лазерів заснований на силах тиску випромінювання. Світлове поле представляється у вигляді квантів, званих фотонами. Кожен фотон несе в собі імпульс, рівний $p = \hbar k$, де k - хвильовий вектор $|k| = k = 2\pi / \lambda$; \hbar - постійна Планка. Взаємодія між випромінюванням і речовиною може призводити до обміну (зміни) імпульсу. Основний принцип дії пасток може бути пояснений за допомогою другого і третього законів Ньютона. Так як світло має імпульс, то зміна напрямку світла означає, що повинна існувати сила, пов'язана з цією зміною. При проходженні крізь прозору частинку світло відхиляється. Сила, що виникає в результаті зміни напрямку руху світла, діє на частинку так, що вона рухається в область найбільшої інтенсивності світла.

З точки зору передачі імпульсу можна уявити процес розсіювання як поглинання фотонів, що рухаються в напрямку падаючого світла, і випромінювання нових фотонів в напрямку розсіяного світла. Результуюча передача імпульсу описується вектором різниці імпульсів, відповідних поглинанню та випромінюванню фотонів.

2.2.1.2 Типи частинок. З теоретичної точки зору захоплювані об'єкти можна розділити на дві групи: а) радіус об'єктів більше $\lambda / 20$ [49], коли можна застосувати узагальнену теорію розсіювання Лоренца-Мі для опису сил взаємодії частинки і лазерного пучка; б) радіус об'єктів менше $\lambda / 20$ (релеєвської частки), вони можуть розглядатися як електричні диполі і до них може бути застосована релеєвська теорія розсіювання.

Під «наночастинками» розуміються частинки, радіус яких менше $\lambda / 20$ (релеєвської апроксимація). У роботах [50, 51] продемонстровано успішне

захоплення таких частинок за допомогою однопучкової пастки. У зв'язку з броунівським рухом кінетична енергія таких малих об'єктів збільшується зі зменшенням їх розміру. Тому для стабільного захоплення релєєвських частинок необхідно використовувати більш потужні пастки для створення більш глибокої потенційної ями. Чим більше потужність, тим меншого розміру релєєвські частинки можна захопити [50].

При розгляді як мікронних частинок ($d \leq \lambda$), так і субмікронних, включаючи і атоми ($d \ll \lambda$), часто поділяють повну силу на дві компоненти:

1. Одна називається силою розсіяння (кеплерівська сила тиску світла). Вона пропорційна розсіючій поверхні частинки і інтенсивності світла, для параксіальних пучків спрямована уздовж вектора Пойнтінга, тобто в напрямку падіння світла. Ця сила виникає через зміни імпульсу електромагнітної хвилі при розсіюванні на диполі;

2. Інша називається градієнтною силою. Ця сила виникає в присутності просторового градієнта енергії світла [52]. Її часто розглядають як електричну силу на оптично індукційованому диполі частинки в градієнті інтенсивності поля. Для релєєвського (субмікронних) часток градієнтна сила зазвичай називається дипольною силою. Градієнтна сила дорівнює силі Лоренца, що діє на створений (індукційований) диполь [50, 51], і збігається з напрямком градієнта інтенсивності.

Обидві силові компоненти досліджувалися в параксіальному наближенні [53]; було показано, що сила, яка діє в напрямку, протилежному поширенню світлового пучка (тобто осьова градієнтна сила, спрямована у фокус пучка), більше сили розсіювання тільки для дуже малих часток.

Об'єкти з показником заломлення вищим, ніж навколишнє середовище, притягуються до областей високої інтенсивності світла, а менш щільні відштовхуються від цих областей. Так, бульбашки повітря (створені збовтуванням) і непрозорі частки будуть виштовхуватися з фокальної області.

Металеві об'єкти відбивають світло і зазвичай не розглядаються як кандидати на захоплення в класичній гауссовській пастці. Однак в [51] показано, що для малих релєєвських ($\lambda \gg a$, a - радіус частинки) металевих часток таке захоплення можливе. У масштабі таких розмірів картина розсіяння для металів і діелектриків однакова, а велика поляризованість металів визначає велику силу захоплення.

Більшість біологічних об'єктів мають показник заломлення вище, ніж у води, що дозволяє їх стабільно захоплювати в сильно сфокусованому пучку.

2.2.1.3 Променева теорія оптичного захоплення. Розглянемо діелектричний сферичний об'єкт в сфокусованому світловому пучку. Якщо малі непрозорі об'єкти при віддзеркаленні світла від їх поверхні відчують поштовх, то при заломленні світла прозорими об'єктами може відбуватися їх притягання в область високої інтенсивності.

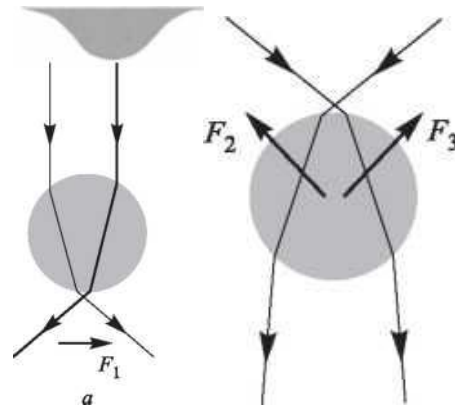
Заломлення або відхилення променів на діелектричній поверхні змінює напрямок імпульсу світла (сфера діє як мікролінза).

І як наслідок з третього закону Ньютона, сфера відчуває «рівне і спрямоване у зворотному напрямку» дію, пов'язану зі зміною імпульсу.

На рис. 2.5, а розглянуто поперечний рух [54] і показано, що сфера притягається в область максимальної інтенсивності пучка. Для лазерного пучка з гауссовским профілем найбільша інтенсивність на осі пучка, і частинка утримується в поперечному центрі пучка.

На рис. 2.5, бачимо, як зміна імпульсу призводить до появи сили, що штовхає сферу назад в фокальну площину пучка. Так як фокус пучка є максимумом інтенсивності в поздовжньому розподілі пучка, то частка притягається в фокус, тобто за відсутності інших сил об'єкт буде рухатися в бік фокуса (поздовжній захват). Щоб зробити цю силу достатньою для подолання різних сил, виштовхуючих об'єкт з фокусу (тиск світла, сила тяжіння), світло повинне бути сильно сфокусованим, що досягається при використанні об'єктивів з великою числовою апертурою NA. Максимальне значення NA

дорівнює показнику заломлення середовища n , що оточує лінзу. Для повітря $NA_{max} = 1$, для іммерсійних об'єктивів NA_{max} досягає 1,4. Збільшення потужності лазерів також збільшує захоплюючу силу.



а) в поперечному напрямку;

б) в поздовжньому напрямку.

Рисунок 2.5 - Ілюстрація до променевої теорії оптичного захоплення частинок

Комбінація поперечного та поздовжнього захоплення дозволяє створити 3Д-пастку.

2.2.1.4 Релєєвська теорія оптичного захоплення. В [53] були отримані аналітичні вирази в наближенні Релєя для сил тиску випромінювання, що діють на діелектричну частку з радіусом $a < \lambda / 20$ при розсіюванні на ній гауссовського пучка (фундаментальної моди лазерного випромінювання) з радіусом $\omega_0 > 2\lambda$.

Загальна сила розглядалася у вигляді суми двох складових:

$$F = F_{\nabla} + F_S, \quad (2.1)$$

де F_{∇} - сила, що виникає при наявності градієнта інтенсивності світлового поля; F_S - сила, що виникає при розсіянні світла на частці.

Кожна з компонентів була виражена через вектор Пойнтінга:

$$F_{\nabla} = 2\pi a^3 \frac{\sqrt{\varepsilon_0}}{c} \left(\frac{\varepsilon - \varepsilon_0}{\varepsilon + 2\varepsilon_0} \right) \nabla |S|, \quad (2.2)$$

$$F_s = \frac{8}{3} \pi \left(\frac{2\pi a}{\lambda} \right)^4 a^2 \frac{\sqrt{\varepsilon_0}}{c} \left(\frac{\varepsilon - \varepsilon_0}{\varepsilon + 2\varepsilon_0} \right)^2 S, \quad (2.3)$$

де ε - діелектрична постійна частки; ε_0 - діелектрична постійна навколишнього середовища.

Сила F_{∇} тягне частку в область максимальної інтенсивності світла, а сила F_s штовхає її в напрямку поширення світла. Для 3Д-захоплення частки градієнтна сила повинна бути досить великою, щоб подолати силу розсіювання.

Дані вирази можуть бути застосовані до металевих або поглинаючих часток, якщо врахувати комплексний характер показника заломлення.

2.2.1.5 Узагальнена теорія Лоренца-Мі оптичного захоплення. Точний розрахунок сил, що діють на частинку в оптичній пастці, досить складний. Як правило, використовуються наближення, коли $a \gg \lambda$ або $a \ll \lambda$, а також приймається, що частка однорідна і має сферичну форму. Однак особливий, практичний інтерес представляють частки розміром $a \sim \lambda$ і несферичної форми, так як саме вони використовуються в безлічі досліджень із застосуванням оптичних пасток.

В [55] теоретично і експериментально досліджуються радіальні сили в оптичній пастці, що діють на частинки розміром від 0,5 до 7,5 мкм. Це відповідає проміжному діапазону між релєєвського ($a < \lambda/20$) і геометричним ($a \ll \lambda$) наближеннями. Оцінка діючих сил була отримана на основі електромагнітного підходу за допомогою узагальненої теорії Лоренца-Мі [56-58], застосовної для будь-яких розмірів часток. Коректність отриманих виразів в [53] також перевірялася на основі узагальненої теорії Лоренца-Мі.

В [59] розглядаються різні методи, в тому числі і електромагнітний підхід, для розрахунку сил, що діють на частки довільної форми з розмірами порядку мкм.

2.2.1.6 Взаємодія світла з частками. Дипольний момент визначається як сума множень зарядів системи на радіуси-вектори цих зарядів. Для одних молекул такий дипольний момент існує і при відсутності будь-якого зовнішнього поля. Для інших він дорівнює нулю і утворюється тільки в присутності зовнішнього електричного поля через перерозподіл заряду.

Взаємодія світла з зарядженими частками. Взаємодія диполя з потужним полем лазерного випромінювання призводить до обертання або коливання молекулярної осі. При постійному дипольному моменті і постійному електричному полі молекули орієнтуються вздовж напрямку поля. Електричне поле електромагнітної хвилі змушує заряджені частки осцилювати. Тоді лоренцевська взаємодія з магнітним полем хвилі штовхає заряд у напрямку поширення хвилі. Так як хвиля дає імпульс заряду, то це визначає, що вона сама має імпульс.

Тенденція молекул вибудовуватися в присутності змінного лазерного поля добре відома в нелінійній оптиці [60]. Вибудовування полярних молекул відповідає за орієнтаційний ефект Керра: оптично-ізотропна середа під дією постійного електричного поля стає анізотропною і двоякопроменезаломлюючою.

Взаємодія світла з нейтральними частинками. Для неполярних молекул в світловому полі має місце оптичний ефект Керра: атоми і молекули, початково не мають постійного дипольного моменту, набувають його під дією поля, тобто - поляризуються. Причина поляризації полягає в тенденції електронів зрушити в напрямку, протилежному напрямку напруженості електричного поля, в той час як важкі позитивно заряджені ядра практично не зміщуються. Негативно заряджені електрони прагнуть зайняти область, протилежну напрямку вектора напруженості зовнішнього електричного поля. При цьому дипольний момент

пропорційний напруженості поля, а коефіцієнт пропорційності називають поляризуємостю молекули α .

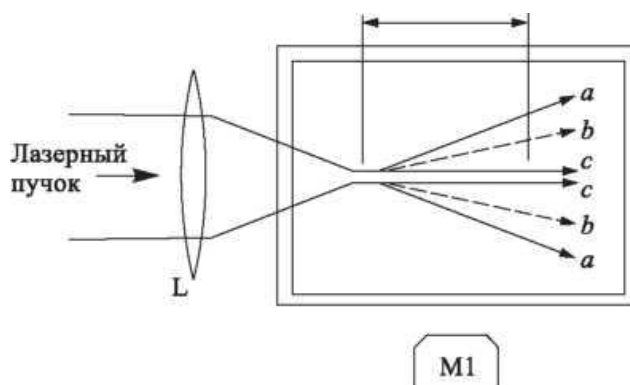
Практично в будь-яких молекулярних середовищах виникає оптичний ефект Керра, що призводить до залежності показника заломлення від напруженості поля лазерного випромінювання та анізотропії цього показника заломлення щодо направлення поляризації випромінювання [61].

Самофокусування. В [62] були описані результати дії сил тиску випромінювання від сфокусованого імпульсного лазерного пучка на вільну поверхню прозорої рідини діелектричного середовища. Виявлено, що при вході або виході світла з рідини на поверхні виникає сила, спрямована назовні, яка викликає сильний лінзовий ефект і розсіювання на поверхні. Це явище пов'язане з імпульсом світла в діелектричному середовищі, пондеромоторною силою електромагнітної хвилі в діелектричному середовищі і самофокусуванням лазерного світла в рідинах.

Пізніше в [63] були описані експерименти з самофокусування і самозахоплення лазерних пучків в штучному середовищі Керра, що складається з розчину субмікронних діелектричних часток (рис. 2.6) [63]. Вперше спостерігалося «схлопування» (самозахоплення) безперервного лазерного пучка в нитку малого (до 2 мкм) діаметру, а також самозахоплення гауссовських мод високих порядків.

Поняття самофокусування використовується також при описі поширення світла у хвилеводах з градієнтним профілем показника заломлення. Наявність поляризуємості у молекул призводить, крім ефекту Керра, ще до одного важливого ефекту - можливості фокусувати пучки молекул в полі неоднорідності (градієнта) розподілу інтенсивності лазерного випромінювання, яка, наприклад, є у фокусі на осі лазерного пучка. Сила, що діє на молекулу в такому потенціалі, спрямована до осі лазерного пучка. Сучасні лазери легко можуть прискорити пучки нейтральних молекул водню до кінетичної енергії в

4 eV, що значно перевищує теплову енергію молекули ($\sim 0,025$ eV при кімнатній температурі) [61].



a - початковий гауссовський пучок; b - самофокусування при підвищенні потужності випромінювання; c - стискання при досягненні критичної потужності і вище.

Рисунок 2.6 - Хід променів

Таким чином, нейтральні молекули, завдяки поляризованості, можуть ефективно взаємодіяти з інтенсивним лазерним випромінюванням. Вибудовування молекул в полі лазерного випромінювання відбуватиметься краще для більш високих інтенсивностей лазерного випромінювання і більш легких молекул.

2.3 Основні типи оптичних пасток

2.3.1 Однопучкова 2Д-пастка

Переміщення і прискорення частки. Теоретичні розрахунки показують [64], що сфокусований пучок потужністю 1 Вт діє на частку радіуса λ , за умови, що вона ідеально відображає весь імпульс падаючого на неї світла, з силою, рівною 10^{-3} дин. В абсолютному значенні це мала величина, але якщо

врахувати малу масу частки, вона дає прискорення $F/m \sim 10^5 g$, де g - прискорення сили тяжіння.

Перші експерименти зі спостереження прискорення мікронних часток, зважених в рідині і газі, описані в [64] (рис. 2.7). В експерименті використовувалися латексні ($n_p = 1,58$) сферичні частки 0,59, 1,31 і 2,68 мкм в діаметрі, розчинені у воді ($n_s = 1,33$). TEM_{00} - мода аргонного лазера з радіусом перетяжки $\omega_0 = 6,2$ мкм горизонтально фокусувалася на окрему частку крізь скляну кювету товщиною 120 мкм.

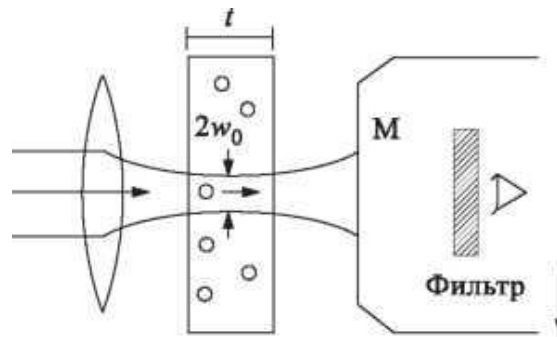


Рисунок 2.7 - Схема експерименту щодо прискорення руху частки

Спостереження велося за допомогою мікроскопа. Частинка розміром 2,68 мкм потрапляла в пучок потужністю 19 мВ не по центру, притягалася до осі пучка і прискорювалася в напрямку руху світла (швидкість ~ 26 мкм / с) поки не впиралася в стіну кювети, залишаючись захопленою в пучку. Після відключення лазера частка починала броунівський рух.

Експериментально показано, що частки з більш високим, ніж навколишнє середовище, показником заломлення притягуються в області високої інтенсивності. Менш щільні, ніж навколишнє середовище, частки (наприклад, бульбашки повітря розміром 8 мкм, створені збовтуванням) виштовхувалися з пучка.

Левітація. Перша пастка була створена в AT & T Bell Labs в 1970 р [64]. У розробленій тоді «левітаційної» пастці використовувався тиск випромінювання,

що є джерелом потоку фотонів, спрямованого вгору і чинного проти сили гравітаційного тяжіння (рис. 2.8).

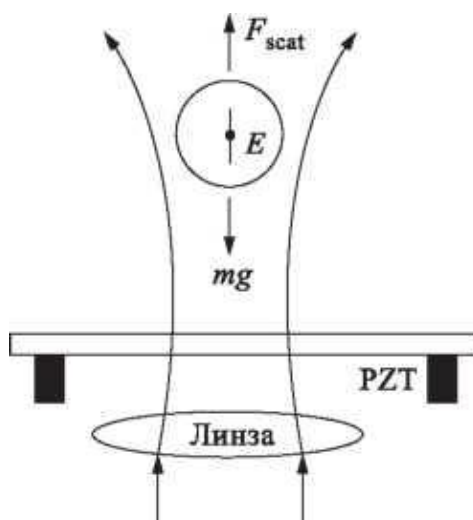


Рисунок 2.8 - Схема експерименту з левітації

Сила тиску світла мала в порівнянні з іншими силами і, як правило, повинна діяти проти сили тяжіння. Величина сили, створеної лазером потужністю в кілька мілліватт, зазвичай знаходиться в діапазоні декількох піконьютонів. Однак частки з розмірами порядку мкм можуть бути підняті і утримані проти сили тяжіння.

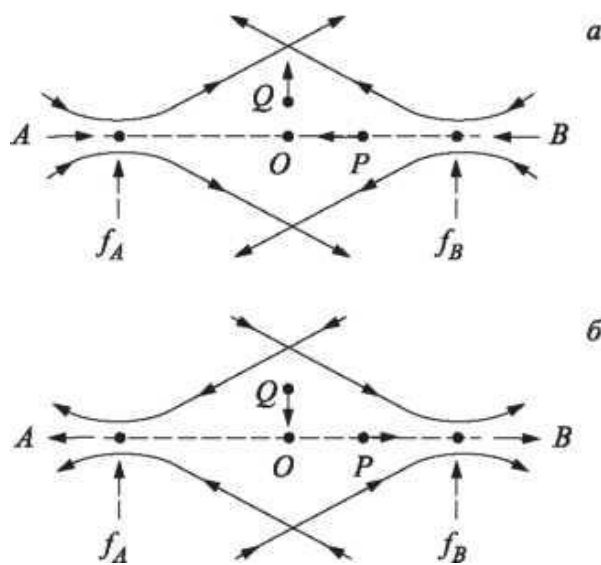
З використанням оптичної левітації спостерігалася зміна сили тиску випромінювання на прозорі діелектричні сферичні об'єкти залежно від довжини хвилі і розміру [65]. При цьому точність вимірювання розмірів і показника заломлення частинок збільшена на 2-3 порядки в порівнянні з методом розсіювання в дальній зоні.

2.3.2 Двопучкова 3Д-пастка

Для стабільного захоплення вздовж оптичної осі була запропонована [64] двопучкова пастка, яка використовує два зустрічних гауссових пучка. Експериментально спостерігалася, як частка, що знаходиться поблизу одного з

пучків, притягалася в точку рівноваги. При відключенні одного з лазерів частка починала швидко прискорюватись в напрямку діючого пучка і повільно поверталася в точку рівноваги, коли лазер знову включався. У зв'язку з такими пастками була обґрунтована оптична теорема Ерншоу [66].

Схема такої пастки складається з двох гауссовських пучків (TEM_{00} -мода), спрямованих назустріч один одному. Точка рівноваги розташовується між їх фокусами в дальній зоні. Зміна напрямку проводиться за допомогою електрооптичного поляризаційного модулятора [67] (рис. 2.9).



а) пучки спрямовані один до одного, тоді аксіальне положення (P) стабільно, а поперечне (Q) - нестабільно; б) (через півциклу) напрямок пучків переключасться на зворотне (в схемі це ортогональна поляризація) і стабілізується поперечне положення, а аксіальне - перестає бути стабільним.

Рисунок 2.9 - Схема двухпучкової пастки

2.3.3 Однопучкова 3Д-пастка

У попередніх «левітаційних» пастках [68] стабільність осьового захоплення залежала від балансу розсіючої сили і сили гравітації, так як осьова градієнтна сила мала. В [50] описана однопучкова пастка, в якій створено

поздовжній градієнт інтенсивності за рахунок сильного фокусування пучка. За допомогою цієї пастки стабільно захоплені діелектричні частки розміром від 0,025 до 10 мкм у водному розчині. Експериментально продемонстровано існування негативного тиску випромінювання, або силовий компонент (який пов'язаний з осьовим градієнтом інтенсивності), що притягає частку назад до фокусу.

Фізична природа зворотної градиентної сили особливо наочна при розгляді часток, діаметр яких істотно перевищує довжину хвилі ($\geq 10\lambda$). Тоді для опису можна використовувати променеву оптику. Для стабільного захоплення релеєвських часток (діаметр яких значно менше λ) необхідно, щоб градієнтна сила, що притягає частку назад у фокус пучка, перевищувала силу розсіювання, що штовхає її вперед.

Альтернативою однопучкової пастки з сильним фокусуванням є двопучкова пастка, описана раніше і здатна виробляти більш потужну захоплюючу силу при меншій числовій апертурі МЛ. Основний недолік такого пристрою - технічна складність.

2.3.4 Інтерференційна пастка

Незважаючи на великий успіх, однопучкові пастки мають певні недоліки, якщо використовуються в схемі класичного мікроскопа (тобто світло спрямоване зверху вниз). Тоді для отримання досить великої сили, що діє проти напрямку поширення світла, необхідно сильно фокусувати пучок. При цьому осьова сила завжди слабкіше поперечної в три і більше разів. Таким чином, захоплена частинка легко звільняється з пастки поблизу дзеркала, так як в цьому випадку поверхневі сили долають оптичні [69].

Щоб подолати ці недоліки, в [70] пропонується модифікувати пастку - дно кювети покрити сильно відображуючим шаром. Тоді в результаті інтерференції падаючої і відбитої хвиль утворюється стояча гауссовська хвиля,

і нанооб'єкти будуть локалізовані в пучностях (розділених $\lambda/2$ вздовж оптичної осі) такої пастки.

Теоретичні розрахунки показали, що навіть «дуже мале» дзеркальне відображення може виробляти градиентну силу того ж порядку, що і сила розсіювання. Навіть без спеціального, сильно відбиваючого покриття дна кювети, градиентна сила дає превалюючий внесок у загальну силу випромінювання, і можливе захоплення релеєвського частинок більшого розміру, ніж за допомогою однопучкової пастки.

Інший тип інтерференційних пасток було розглянуто в [71], де досліджуються сили, що виникають при взаємодії розповсюджених назустріч один одному коротких π -імпульсов лазерного випромінювання, які можуть направляти атоми в невеликі області перекриття імпульсів. Така пастка може бути скомбінована з лазерними пучками, які здійснюють доплеровске охолодження.

2.4 Основна схема оптичного пінцета

У своєму класичному уявленні схема оптичного пінцета може бути представлена як на рис. 2.10.

Конструктивно лазерний маніпулятор можна розділити на чотири окремі, але пов'язані між собою, функціональні частини:

- лазер з оптикою, формуючий його випромінювання в перетяжку з необхідними характеристиками;
- схема візуалізації процесу маніпуляції;
- схема, що забезпечує можливість переміщення досліджуваного мікрооб'єкту;
- керуюча електроніка.

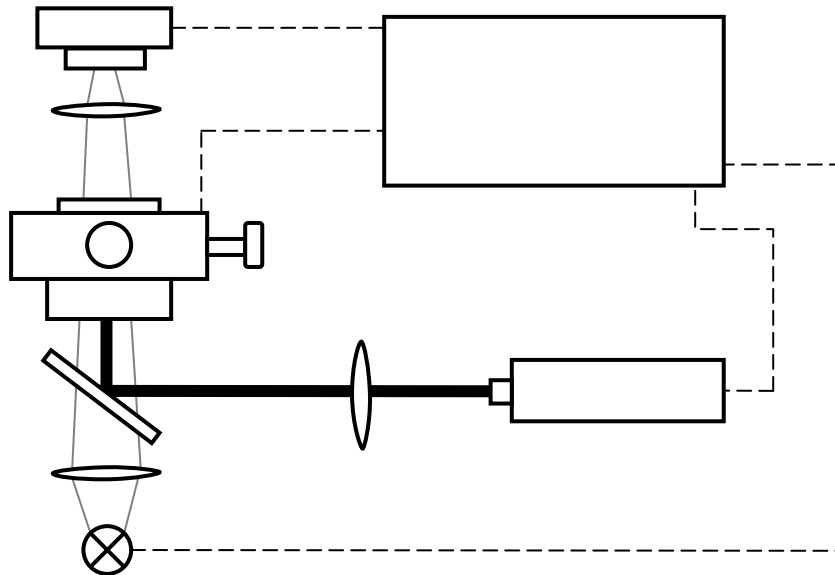


Рисунок 2.10 - Схема оптичного пінцета

Для лазерних пінцетів застосовуються різні лазери. В основному своє застосування тут знаходять напівпровідникові, фемтосекундні (волоконні) і гелій-неонові лазери. В якості формуючої оптики використовуються різні об'єктиви.

Схема візуалізації являє собою мікроскоп, що складається з об'єктива і окуляра, що забезпечують необхідне збільшення, веб-камеру та програмне забезпечення на комп'ютері. Основне завдання схеми візуалізації - отримання якісного зображення на дисплеї ПК, яке дозволить розрізнити досліджувані мікрооб'єкти.

Схема, що забезпечує можливість переміщення досліджуваного мікрооб'єкту, конструктивно може бути представлена двома різними варіантами. Перший варіант реалізації: за допомогою переміщення досліджуваного середовища щодо лазерного променя. Другий варіант: відхилення променя відносно нерухомої досліджуваної середовища з мікрооб'єктами.

2.5 Области застосування оптичного захоплення мікрооб'єктів

Оптичні пастки - безконтактний інструмент, який використовує один або декілька лазерних пучків для генерації сил порядку пН, досить потужних для маніпулювання об'єктами мезоскопічного масштабу (лазерне випромінювання потужністю 1 мВт достатньо для підйому і транспортування 10-мкм діелектричної частки). Ця можливість представляє особливий інтерес в ряді областей: атомна фізика, біологія, мікрохірургія, нанотехнологія і мікромеханіки, а також при вивченні властивостей колоїдів, мікропотоків і світлових пучків.

Оптичне маніпулювання за допомогою лазерів застосовно до об'єктів від атомів до великих молекул і невеликих діелектричних частинок, що варіюються в розмірі від десятків нанометрів до десятків і сотень мікрометрів, а також до таких біологічних об'єктів, як віруси, окремі живі клітини і внутрішньоклітинні включення.

2.5.1 Маніпулювання біооб'єктами

У біології оптичні пастки відкривають можливості безконтактного впливу (без оптичного руйнування) на живі клітини. Оптичний «пінцет» дозволяє керувати окремими молекулами з точністю до нанометрів і вимірювати сили, що діють на ці молекули, з точністю 1-100 пН. Поява оптичних пасток відкрило кілька важливих областей дослідження в біофізиці: маніпуляції з живими вірусами і бактеріями [72], окремими молекулами ДНК [73] і людськими гаметами [74]; клітинна мікрохірургія [75], дослідження рухових протеїнів [76], митотичних хромосом [77] та ін.

Серед унікальних додатків - вивчення рухових молекул і механоензимів (mechanoenzymes). Завдяки оптичній пастці було детально відстежено їх пошговий рух з точністю ~ 10 нм/крок, коли вони рухаються уздовж субмікронної трубки.

В останніх, що вражають увагу, експериментах - прямі вимірювання сил, створюваних РНК-полімеразой при русі вздовж молекули ДНК. Нові експерименти по маніпулюванню окремими молекулами дозволили перевірити фізичні моделі, що описують властивості ДНК, способом, який був раніше неможливий. З'явилася можливість спостерігати активність окремого ензиму, який діє на молекулу ДНК, і простежити сили, якими молекули рухомих тканин (органів) впливають на біологічні молекули. Стало можливим напряму контролювати вплив зв'язування белків ДНК на свойства ДНК і використовувати ці виміри, щоб визначити енергію взаємодії [78]. Детальні знання індивідуальних взаємодій між молекулами важливі для розуміння складних механізмів, що діють в реальних біологічних процесах.

Інша велика область - вимірювання механічних (пружних) властивостей частей клітин цитоплазми, жгутіков бактерій, світлочутливих оболонки клітин еритроцитів, окремих мікротрубок ендоплазматичної мережі, окремих світлочутливих волокон, мембран нервових клітин, довгих спіралей окремих молекул ДНК.

Можливість розділення живих клітин була застосована для пошуку і клонування нових високотемпературних анаеробних бактерій «archaea», що має важливе значення в відкритті нових високотемпературних ензимів. Ведуться дослідження з розпізнавання клітин, зварюванню клітин, вивченню руху хромосом під час поділу клітин.

Комбінкування оптичних пасток з іншими лазерними пучками дозволяє використовувати їх в мікрохірургії. Наприклад, захоплюється хромосома інфрачервоним (1064 нм) лазером-пасткою і розрізається на невеликі частки зеленим (532 нм) лазером. Такий устрій носить назву оптичний «скальпель». Це можливо, оскільки більшість біологічних структур практично не поглинають в інфрачервоній області, але значно поглинають світло в зеленій частині оптичного діапазона довжин хвиль.

2.5.2 Фізика колоїдів

Оптичні пастки дозволяють досліджувати сили, які виникають при взаємодії коллоїдів [79], полімерів [80], а також вибудовувати кілька коллоїдних частинок в мезоскопічні структури [81].

В [79] використання оптичної пастки дозволило провести точне вимірювання з високим дозволом динаміки введення коллоїдних часток у розчині за дуже короткий час.

Ентропійні сили взаємодії між двома великими (розміром порядку мкм) сферичними частками, поміщеними в розчин колоїдних часток набагато меншого розміру, були виміряні за допомогою лінійно-скануючої оптичної пастки [50]. При цьому частки «нанизані» на продовжений (лінійний) фокус: вільно переміщуються в одному напрямку – вздовж фокальної лінії і обмежені в будь-якому перпендикулярному напрямку. Така методика дозволила провести вимірювання функціональної форми потенційної ями з високою роздільною здатністю (15 нм) [82]. Вимірювання показали, що при низькій концентрації колоїдного розчину сила тяжіння великих частинок є монотонною, а при високій концентрації великі частинки дуже рідко притягувалися один до одного, що можна розглядати як новий спосіб стабілізації суспензії тільки на основі ентропії.

В [82] представлено пристрій, за допомогою якого, на основі механізму оптичної пастки, окремі мікрочастинки кристалів можуть бути відібрані і переміщені в розчин для вирощування.

2.5.3 Мікромеханіка

Останнім часом тенденція до мікромініатюризації в дослідженнях і розробках робить актуальним виробництво мікромашин і молекулярних двигунів, розробку мікромеханізмів і мікрокомпонент [82].

Лазерні пастки дають можливість безконтактного маніпулювання крихкими об'єктами, зменшуючи ризик їх розлому, забруднення і псування. Нещодавно була показана можливість захоплення, маніпулювання, орієнтування та обертання несферичних часток. Мікромеханізми можуть бути різної форми, залежно від завдання, яке вони виконують. Складність механізму визначається кількістю деталей і етапів зборки.

В [82] розглядався оптичне захоплення в області перетяжки сильно сфокусованого гауссовського пучка часток, що мають різні форми: сферичну, циліндричну, кільцеву, зіркоподібну, кубічну та ін., розміром 1-10 мкм. Безконтактне маніпулювання і орієнтування таких об'єктів забезпечує можливість спрощення збірки (монтажу), активації і, можливо, ремонту мікромеханізмів.

В [82] використовується оптичний обертальний момент (порядку $10^{-15} \text{ Н} \cdot \text{м}$) для управління обертанням мікромашиного елемента, захопленого в оптичній пастці. В експерименті мікромашиний двопротинепреломлюючий елемент (квадратна пластинка) з вуглекислого кальцію (CaCO_3) розміром 1 мкм захоплено і обертається (з частотою до 350 Гц) сфокусованим лазерним пучком (1064 нм, потужність 300 мВт) з круговою поляризацією. При цьому оптичний обертальний момент передається другому мікромашиному елементу (шестерня з 6 зубцями) діаметром 10 мкм (зубці розміром близько 1 мкм) з аморфного діоксид кремнію (SiO_2 , $n = 1,46$) через рух навколишньої рідини (вода). Другий елемент був виготовлений методом фотолітографії. Матеріал був обраний так, щоб елемент був досить прозорим, щоб не піддаватися оптичному руйнуванню, і показник заломлення був вище, ніж у навколишнього середовища.

Максимальний обертальний момент, який може бути отриманий при падінні світла а з круговою поляризацією на дволучепреломляючого кристал, досягається при товщині $\lambda / 2$. Наприклад, частки CaCO_3 товщиною 1,4 мкм будуть $\lambda / 2$ -пластина для світла з довжиною хвилі. $\lambda = 488 \text{ нм}$. При цьому

оптичний обертальний момент не дуже великий (для лазера з $\lambda = 488 \text{ нм}$ і потужністю 1 Вт обертальний момент дорівнює $6,1 \cdot 10^{-16} \text{ Н} \cdot \text{м}$), проте дозволяє обертати частинки розміром порядку мкм з дуже високою швидкістю. При обертанні ісландського шпату з частотою 100-200 Гц кремнієва шестерня обертається з частотою $\sim 0,2 \text{ Гц}$ у зворотний бік [82] (рис. 2.11). Ефективність передачі імпульсу між мікроелементами $\sim 50\%$.

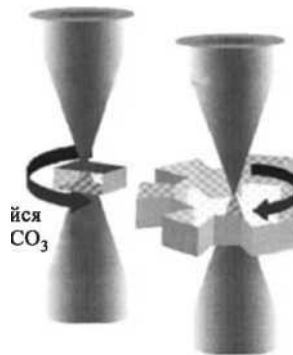


Рисунок 2.11 - Передача обертального моменту від двопротинепреломляючого кристала кремнієвої шестеринці

Однак простіше й ефективніше обертати безпосередньо діелектричні елементи мікромеханізмів, наприклад, за рахунок передачі їм орбітального кутового моменту вихрового лазерного пучка. Формування таких пучків можливо за допомогою дифракційних оптичних.

2.6 Динамічне управління мікрооб'єктами

2.6.1 Методи управління об'єктом

Предметний столик та інші юстувальні вузли оптичного пінцета складають основу його механічної частини. Предметний столик, як один з основних вузлів лазерного пінцета, призначений для виконання таких функцій

як забезпечення кріплення препарату і легкий доступ до нього спостерігача, а також точне переміщення препарату по двох координатах. В даний час застосовується автоматичне переміщення кареток столика по двох координатах, кероване від ПК.

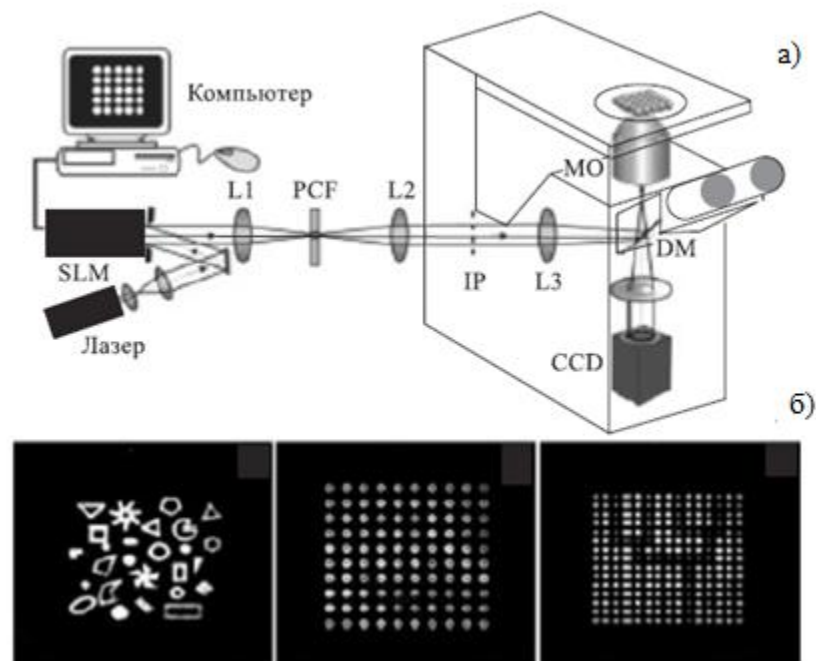
Зазвичай захоплену частинку рухають механічним переміщенням пучка або за допомогою переміщення платформи мікроскопа. Існує також оптичний спосіб для направлення захоплюючого пучка з використанням адаптивних дзеркал або акустооптичних модуляторів. Однак такі системи стають досить складними, якщо необхідно реалізувати 3D-захоплення або кілька пасток [57].

За допомогою синтезованих на комп'ютері голограм, записаних на рідкокристалічній дисплей (LCD), можна досить просто динамічно управляти кількістю, позиціями і формою оптичних пасток як в 2D-, так і в 3D-просторі [49]. Тепер НОТ доповнюються програмованими просторовими модуляторами світла на рідких кристалах (SLM) і набувають динамічні властивості за рахунок проектування послідовності голограм, покроково перебудовують картину.

Використання безлічі динамічних пучків є перспективним засобом для конструювання та удосконалення мікромашин, мікровіхревих механізмів, насосів, клапанів, вирощування колоїдних кристалів, формування полімерних фотонних проводів.

Інтерактивне оптичне маніпулювання також може бути інтегроване в мікросистеми, що використовують мікропотоки і «lab-on-a-chip» - технології, забезпечуючи точне сортування і аналіз певних часток або біологічних клітин.

Одним з варіантів є також оптичне захоплення і переміщення колоїдних часток в інтерактивному режимі реального часу з використанням програмованих модуляторів світла. Колімірований пучок з лазерного діода потужністю 200 мВт (робоча довжина хвилі $\lambda = 830$ нм) відбивається від фазового модулятора (Hamamatsu Photonics), розподіл пропускання на якому кодується комп'ютером. На рис. 2.12 демонструються різні картини, які можна формувати таким способом.



а) оптична схема динамічного управління оптичними пастками;

б) картини, які можна формувати у такий спосіб.

Рисунок 2.12 - Оптичний захоплювач з використанням програмованих модуляторів світла

На рисунку 2.12 абрєвіатура SLM - це просторовий модулятор світла; PCF - фазовий контрастуючий фільтр; IP - площина зображення; MO - мікрооб'єкт; DM - діхромне дзеркало; CCD - ПЗС-камера; L1-L3 - лінзи;

У літературі розглядаються методи по формуванню великих масивів оптичних пасток з довільною 3Д-конфігурацією, і їх динамічному перестроювання, під управлінням комп'ютера [58]. При цьому пастки можуть бути різних типів - гаусові, вихреві, бesselевські та ін.

Ще одним джерелом пропонується динамічний дифракційний елемент, що представляє собою матрицю $N \times N$ програмованих фазових решіток, доповнити матрицею $N \times N$ мікролінз. Фазовий модулятор світла на рідких

кристалах (480 x 480 осередків) у відбивючому режимі забезпечує для світла з довжиною хвилі $\lambda = 830$ нм динамічний діапазон фазової затримки $[0, 3\pi]$, що дозволяє реалізувати матрицю пасток 2 x 2 (відстань між пастками $d = 76$ мкм), в яку захоплюються частки розміром 5 мкм в діаметрі. При цьому 8-рівнева решітка забезпечує максимально можливе зміщення пучка пастки в площині захоплення частки до 2,53 мкм (в експерименті було до 1,5 мкм). При збільшенні числа пасток ($N \geq 3$) зменшується відстань між пастками (наприклад, для $N = 3$ $d = 38$ мкм) і максимально можливе зміщення пучка пастки. Також необхідно міняти масив мікролінз [49].

Основними недоліками просторових модуляторів світла на рідких кристалах поки залишаються низька дифракційна ефективність (сильний дифракційний шум через велику дискретність модуляторів) та недостатнє для реалізації складних фазових розподілів дозвіл матриці пікселів. Кінцевий розмір пікселів також обмежує максимальний розкид дифракційних порядків (при високих несучих просторових частотах відбувається бінаризація фазового профілю і дифракційна ефективність зменшується).

Вимірювання показали, що після рідкокристалічного модулятора залишається 15% енергії падаючого пучка [59]. Втрати енергії пов'язані з наступним: 1) непрозора частина панелі (основні втрати, до 65%), 2) структура схожа на решітку, яка виробляє високі порядки (втрати 54%), 3) неможливість сконцентрувати всю енергію в корисному порядку через те, що модулятор має максимальний фазовий зсув менше 2π (максимально досягнуте співвідношення між першим і нульовим порядком 2: 1), 4) невідповідність квадратної апертури панелі круглому профілю падаючого пучка (8%).

Таким чином, при використанні просторових модуляторів світла на рідких кристалах виникає необхідність вирішувати дві основні проблеми: розробка швидких алгоритмів розрахунку фазових функцій по малій кількості пікселів, а також підвищення дифракційної ефективності.

Розрахунок фазового розподілу, що формує задану конфігурацію різного типу пасток, не є простим процесом. Наприклад, розглядаючи ітеративний розрахунок фази ДОЕ, формуючий масив пасток, локалізованих в 3Д-обсязі, через узагальнене перетворення Фур'є. Для прискорення розрахунків при використанні динамічних голограм використовується метод прямої відповідності фази на модуляторі й інтенсивності в Фур'є-площині, аналогічний методу фазового контрастування. Для підвищення ефективності пропонується доповнити електрично керований модулятор фільтром низьких частот і оптично записувати інформацію, яку несе нульовий порядок на оптично адресований просторовий модулятор світла (OALCD), який буде працювати в відбиваючому режимі вже з іншим джерелом світла, забезпечуючи 85% дифракційної ефективності (в експерименті 53%).

Також є пропозиції використовувати оптично адресуємий (ні піксельний) фазовий модулятор, який управляється рідкокристалічним проєкційним елементом (640 x 480 осередків) через відеовихід комп'ютера [49].

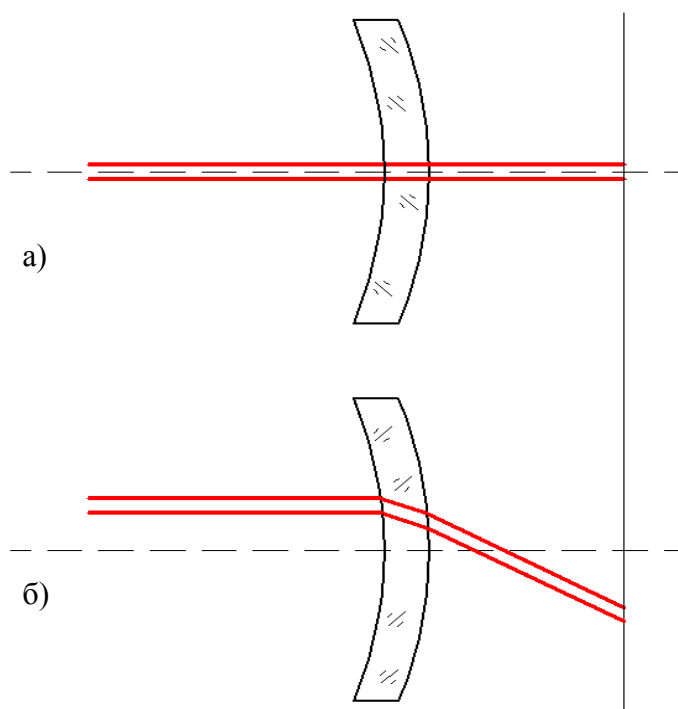
Ми ж розглянемо два методи більш доступних для управління об'єктом через їх простоти конструкції і можливості їх здійснення.

2.6.2 Метод управління за допомогою пересування оптики

Завдяки використанню даного методу управління променем лазера можна здійснювати за допомогою пересування вихідної лінзи або меніска (рис. 2.13).

На рис. 2.13, а видно, що при проходженні променя через оптичну вісь меніска промінь не відхиляється від осі. У випадку, коли меніск пересунули (рис. 2.13, б і оптична ось променя не збігається з осю меніска, промінь відхиляється, ґрунтуючись на цьому ефекті можна переміщати промінь в межах поля зору мікроскопа з дуже високою точністю.

На основі цього можна побудувати схему управління променем лазера (рис. 2.14) для нашого оптичного пінцета



- а) промінь проходить не через оптичну ось меніска
 б) промінь проходить через оптичну ось меніска

Рисунок 2.13 - Схема відхилення променя за допомогою меніска

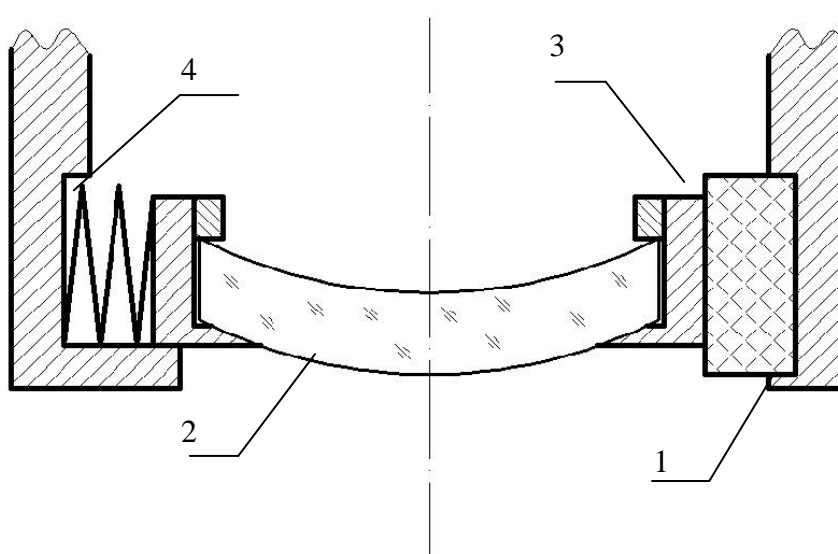


Рисунок 2.14 - Схема керування променем лазера

На даній схемі наведено приклад управління пересуванням меніска 2 щодо оптичної осі об'єктиву за допомогою п'єзоелемента 3, вся конструкція монтується в корпус об'єктива лазера 1, пружина 4 призначена для повернення меніска у початкове положення. П'єзоелементи встановлюються так щоб була можливість переміщати меніск в двох площинах. Великою відхилення меніска від оптичної осі можна управляти за допомогою подачі напруги на п'єзоелемент, тим самим можна забезпечити точне управління нашої відхиляючої системи за допомогою ПК, регулюючи подачу напруги мікроконтролером. Проте внесення ще одного оптичного елемента ускладнює фокусування лазера і вносить свої втрати і оптичні дефекти, які не припустимі у зв'язку з необхідністю отримання гауссового розподілу в площині фокуса лазерного пучка [60].

2.6.3 Метод управління за допомогою предметного столу

Даний метод полягає у переміщенні об'єкта щодо променя за допомогою пересування предметного столу із зразком у двох площинах, при цьому промінь залишається нерухомим. Схематично конструкція даного столу представлена на рис. 2.15.

Мікрометричні гвинти 1 виконують пересування предметного скла із зразком по двох координатах, пересуваючи платформи 3, пружини 2 служать для повернення у вихідне положення платформ 3, отвір 4 призначено для установки предметного скла, на яке розміщується зразок.

Пересування за допомогою мікрометричних гвинтів здійснюється вручну з точністю до мікрона, для збільшення точності можна поставити п'єзоелементи між мікрометрами і платформою або безпосередньо на місце установки предметного скла, тим самим підвищивши точність пересування предметного скла до двох нанометрів, та отримавши можливість керувати точною настройкою переміщення з ПК.

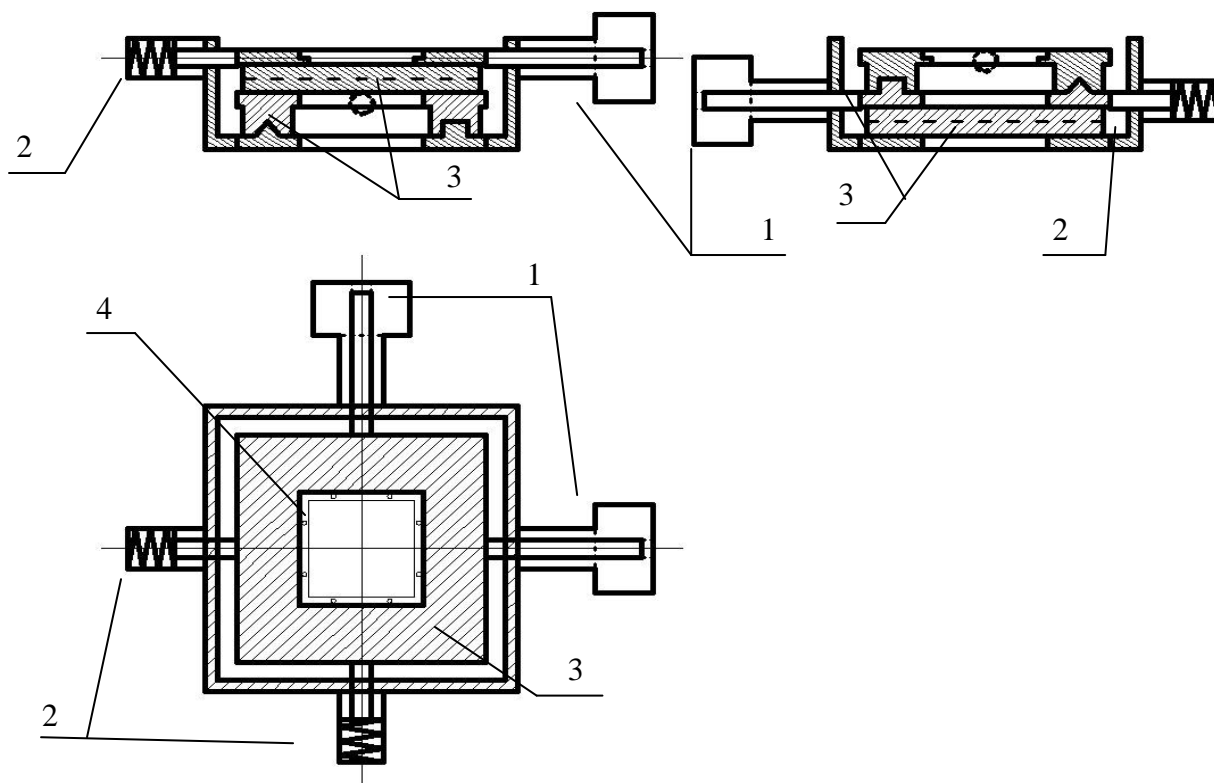


Рисунок 2.15 - Схематичне креслення пересуваючого предметного столика в трьох проекціях

У зв'язку меншою похибкою при переміщенні, а також у зв'язку з простотою і зручністю конструкції, як пристрій переміщення об'єктів був обраний предметний столик з можливістю переміщення в двох площинах.

2.6.4 Вибір конструкції координатного столика

Предметний столик представляє пристрій, що складається з направляючих для точного переміщення. Направляючими називають пристрої для переміщення встановлених на них елементів по заданій траєкторії. По виду руху напрямні поділяють на дві групи:

- напрямні для поступального руху,

- напрямні для обертального руху,
- напрямні для обох видів руху.

В залежності від роду тертя розрізняють напрямні:

- з тертям ковзання;
- з тертям кочення;
- з молекулярним (пружним) тертям;
- з повітряним або рідинним тертям.

Для характеристики направляючих використовують й інші ознаки:

- наявність або відсутність силового або кінематичного замикання;
- форму тіл кочення або переміщених деталей;
- використовувані матеріали і т.д.

У загальному обсязі приладу напрямні складають значну частину, іноді до 25 - 30% деталей, а в окремих приладах визначають точність їх функціонування.

У нашому пристрої використовуються напрямні для поступального руху з тертям ковзання.

На тертя ковзання впливають багато фізико-механічних, конструктивних та технологічних чинників, детермінованого і випадкового характеру. Це пов'язано не тільки з особливостями самого тертя, але і з особливостями роботи механізмів (перехід від спокою до руху, випадковий характер зміни геометрії і шорсткості деталей при переміщенні тощо). Для зменшення тертя використовують мастила різного типу і консистенції.

Деталі напрямних виготовлені з латуні. Швидкості переміщення напрямних знаходяться в межах від 0,0002 до 15 мм / сек [61].

Направляючі ковзання мають достатньо плавний рух і витримують великі навантаження. До числа їх недоліків відносяться:

- велике тертя і великі зусилля рушання;
- збільшені осьові розміри;
- чутливість до змін температури.

Розглянемо напрямні ковзання типу "ластівчин хвіст", конструкція яких визначається технологією їх виготовлення. Дана конструкція застосовується для переміщення верхньої платформи предметного столика.

На малюнку 2.16 представлені два типи напрямних ковзання «ластівчин хвіст» [61].

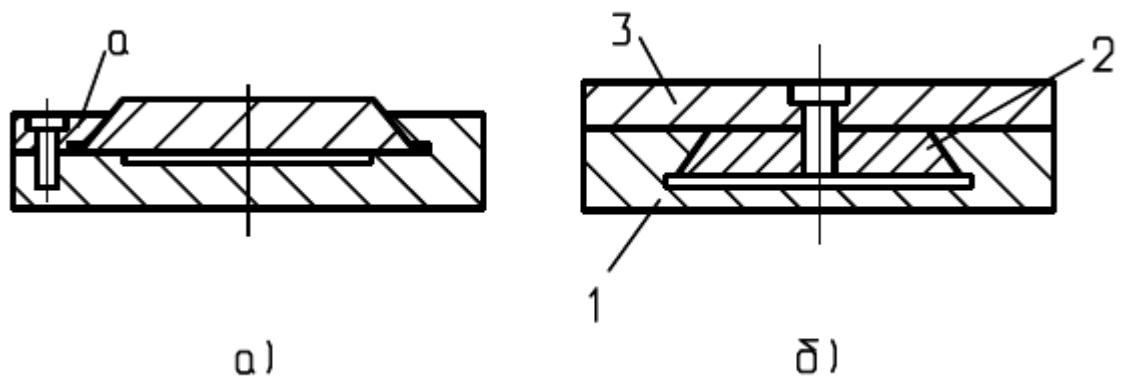


Рисунок 2.16 - Направні ковзання

У звичайних напрямних типу «ластівчин хвіст» (рис. 2.16, а) зйомна бічна планка «а» використовується для регулювання зазорів між деталями при складанні. Таке регулювання зазорів проводиться вручну.

Технологія і конструкція направляючих (рис. 2.16, б), усуває ручні операції і підвищує точність переміщення. У корпусі 1 рухається клин 2, скріплений з планкою 3. З креслення видно, що ця конструкція не має деталей для пригону. Працездатність направляючої забезпечується технологією виготовлення.

Суттєвою причиною, що впливає на точність наведення, є стрибкоподібний характер руху, викликаний багатьма причинами, частина з яких має випадковий характер. Цей процес аналогічний поняттю «шум» в

радіоелектроніці. Тому перспективним є застосування, що дозволяє зменшити величину стрибків (знизити «шум» руху напрямних) і демпфируючих точно установний рух. Зважаючи багатофакторности процесів, що відбуваються при терті.

Для вибору мастила були взяті результати досліджень, які були виконані на шести предметних столиках з використанням мастил ф. Nye США, які рекомендуються для використання в пристроях, де потрібне повільне контрольоване переміщення елементів. В експерименті використовувалися мастила малої, середньої і високої в'язкості.

Отримані дані показали, що точність наведення збільшується при зростанні в'язкості. При випробуваннях з'ясувалося, що мастила з високою в'язкістю значно підвищують зусилля рушання напрямних і моменти на приводах. Повторні випробування столиків після року зберігання показали, що точність наведення збереглася в межах помилки вимірювання. Через два і три роки істотних змін також не відбулося.

Виходячи з отриманих результатів як мастило був обраний Солідол С ОКП02 54110102 з коефіцієнтом в'язкості $190 \text{ Па} \cdot \text{с}$, при такій в'язкості точність переміщення становить менше 1 мкм .

Для переміщення нижньої платформи предметного столика застосовуються напрямні кочення.

Направляючі кочення мають значно менше тертя і зусилля рушання. Вони нечутливі до коливань температури і невимогливі до вибору мастила. Але вони часто мають великий габарит і є більш складними у виготовленні на відміну від напрямних з тертям ковзання.

У напрямних з тертям кочення використовують кулькові, роликові або голчасті тіла кочення. Мертвий хід в них не перевищує 2 мкм .

Різні конструкції, які використовуються у вузлах з тертям кочення, показані на рис. 2.17.

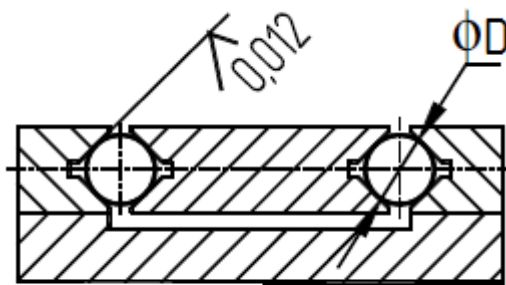


Рисунок 2.17 - Напрямні кочення

Кулькові призматичні напрямні, виготовляються досить ретельно. Їх бігові доріжки мають шорсткість $Ra = 0,012$ мкм, кульки упорядковано по овальності, огранності і відхиленню від номінального значення діаметра до 0,5 мкм.

Для таких кулькових направляючих середнє значення «шуму» становить 0,2 мкм на інтервалі в 500 мкм, що є достатньою точністю переміщення, необхідною для виконання поставлених завдань.

2.6.5 Аналіз профілю інтенсивності лазера

Для управління мікрооб'єктами необхідно отримати градієнт електричного поля і тому належить вибрати лазер з розподілом інтенсивності пучка близьким до гауссова, для цього необхідно вивчити профіль інтенсивності променя лазера і проаналізувати отримані результати на відсутність дефектів (інтерференції, додаткові моди).

Для експерименту було вибрано два твердотільних лазера, випромінюючих у синій і зеленій області спектра. Для проведення експерименту був побудований лабораторний стенд рис. 2.18.

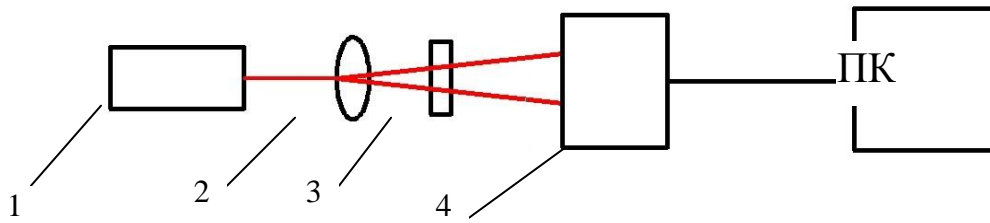
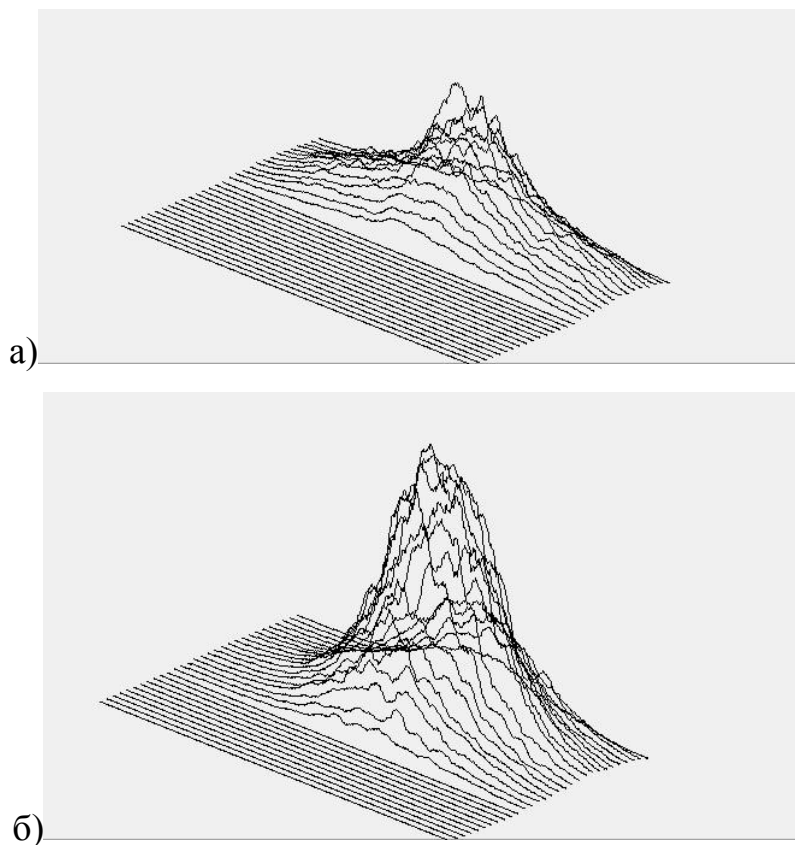


Рисунок 2.18 - Схема лабораторного стенду для перевірки профілю інтенсивності лазерного випромінювання

Луч з лазера 1 проходить через розсіючу лінзу 2, потрапляє на світлофільтр 3 призначений для зменшення потужності випромінювання з метою уникнення блюмінга (розтікання надлишкового заряду в сусідні осередки), через високу потужність лазера. Далі розсіяне і отфільтроване світло потрапляє на ПЗС лінійку 4, де фіксується одним з 3000 елементів в горизонтальній площині, дані передаються на комп'ютер у вигляді 2Д графіка, після чого робиться переміщення лінійки на фіксований крок у вертикальній площині і знову знімаються дані інтенсивності в кожній точці ПЗС лінійки. Таким чином проводиться сканування по всій площині розсіяного пучка лазера.

На рис. 2.19 представлений 3х мірний графік побудований для синього лазера.

З метою більш достовірної картини профілю інтенсивності сканування здійснювалося за двома координатами, коли лазер встановлений прямо (рис. 2.19, а), і лазер встановлений на бік (рис. 2.19, б), внаслідок експерименту видно, що розподіл близько до гауссова, але на рис. 2.19, а видно що лазерний пучок більш розтягнутий в одній з площин, і на краях спостерігаються деякі піки.



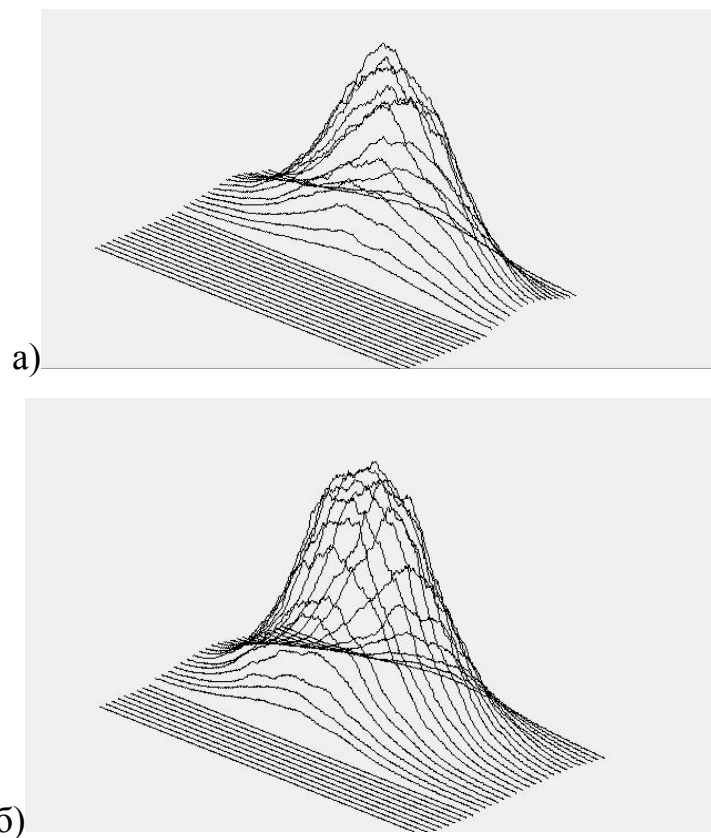
а) лазер встановлений прямо

б) лазер встановлений боком

Рисунок 2.19 - Об'ємна модель профілю інтенсивності синього лазера складається з двовимірних графіків

Також була виміряна потужність пучка лазера до і після світлофільтру потужність лазера складала 18 мВт, а потужність після світлофільтру становила 9 мкВт

Аналогічний експеримент був проведений і для зеленого лазера (рис. 2.20).



а) лазер встановлений прямо

б) лазер встановлений боком

Рисунок 2.20 - Об'ємна модель профілю інтенсивності зеленого лазера,
яка складається з двовимірних графіків

Міжмолекулярні сили визначають не тільки фізичні властивості твердих тіл, рідин, газів, але також відіграють важливу роль у самоорганізації біологічних об'єктів [53]. Природу біомолекулярних взаємодій, що беруть участь в організації живих організмів, визначає набір слабких, нековалентних міжмолекулярних сил. Так, динаміку цитоскелета клітини, а отже, і зміни форми, орієнтації і напрямку руху клітини можна описати, якщо відомі всі міжмолекулярні сили. Знання сил, що виникають між молекулами і біомолекулярними утвореннями в клітці, дозволяє розкрити природу таких процесів, як поділ, зміна форми і орієнтації, зростання, пересування,

самоорганізація клітини, процесів транскрипції і трансляції ДНК, а також процесів, пов'язаних з перенесенням органел і молекул усередині клітини. В даний час активно розробляються динамічні методи для визначення міжмолекулярних взаємодій. Так, за допомогою атомносилової спектроскопії, оптичного пінцета, мікроголок за останні десятиліття проведено ряд унікальних робіт з визначення міжмолекулярних сил [60].

Оптичний пінцет дозволяє управляти цими процесами. Оптичне маніпулювання мікронними і субмікронними частками і навіть окремими атомами стало можливим з 1970 р, коли вперше був проведений такого роду експеримент. Але застосування оптичних методів для дослідження біологічних об'єктів почалося лише в 1987р. Оптичні пінцети використовувалися в експериментах з бактеріями, вірусами, клітинами, органеллами і ДНК. За допомогою оптичного пінцета можна вивчати міжклітинні взаємодії, взаємодії типу клітина-поверхня, антиген-антитіло, рецептор-ліганд.

У комбінації з оптичним скальпелем пінцет також використовується для вивчення злиття клітин [53]. Оптичним скальпелем видаляються частини зовнішніх стінок двох клітин, і за допомогою пінцета здійснюється їх контакт. Використання пінцета в комбінації з оптичним скальпелем для маніпулювання хромосомами під час поділу клітини дозволяє отримати додаткову інформацію про процеси мітозу.

Одним з важливих переваг оптичних пінцетів є те, що вони дозволяють вивчати біомеханічну природу окремих молекулярних моторів. Так, в еукаріотічеській клітці рух і перенесення молекул здійснюються не тільки за допомогою дифузійних механізмів, але і за допомогою спеціальних молекул. Спеціальні білки рухаються по нитках і трубочках, які складають цитоскелет клітини, доставляючи прикріплені до них молекули і транспортні пухирці до різних частин клітини. Ці спеціальні білки називаються "білки-переносники". Мембрана практично непроникна для полярних речовин і великих молекул без білків-переносників, тому їх роль у процесах обміну речовин між клітинами і

органеллами всередині клітини важко переоцінити. Вивчення таких білків-переносників також можливо завдяки оптичним пінцетам. Один з перших експериментів був направлений на вивчення обертальних моторів, що беруть участь в механічних рухах бактерій. Механічна взаємодія молекулярних моторів з клітинними мікротрубочками і філаментами генерує сили, відповідальні за рух клітини, скорочення м'язової тканини, пересування органел всередині клітини.

У роботах вивчається взаємодія молекулярного мотора - молекули кінезин - з мікротрубочкам при наявності АТФ. У цих роботах кінезин, закріплений на сферичній частинці, яка може бути захоплена в оптичну пастку, за допомогою пінцета поміщається на мікротрубочки, де він активується АТФ. Схематично експеримент зображений на рис. 2.21. Було показано, що молекула кінезин створює сили в межах 5-7 пН і пересувається з кроком 8 нм.

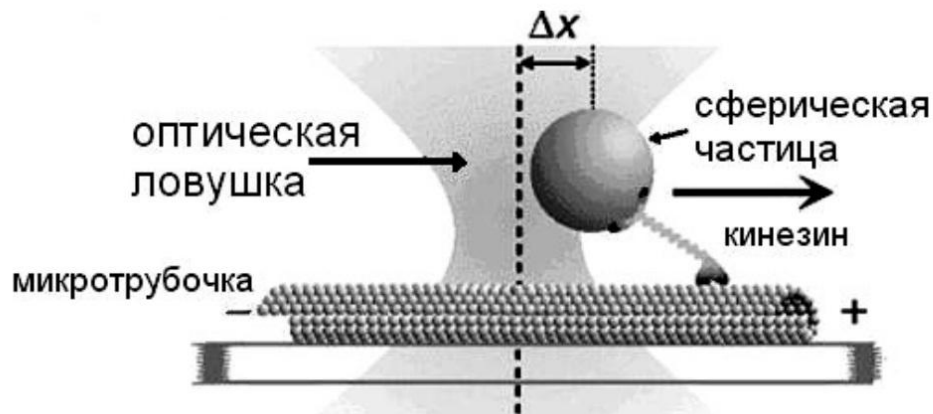


Рисунок 2.21 - Схема експерименту з вивчення взаємодії молекулярного мотора - молекули кінезин з мікротрубочками

На рисунку 2.21 - величиною Δx позначення зсув сферичної частинки відносно центру оптичної пастки.

Використовуюючи подібні методи, Ашкін в роботі виміряв силу, створювану декількома молекулярними моторами - молекулами дінеїна, закріпленими на мітохондріях і рухаючимися уздовж мікротрубочки. Основні зусилля були і залишаються спрямованими на вивчення лінійних рухових білків, які рухаються вздовж полімеру. Наприклад, РНК-полімераза - молекулярний мотор, який бере участь в транскрипції ДНК. У роботі автори показали, що РНК-полімераза при транскрипції впливає на ДНК з силою 21-27 пН (див. Рис. 2.22).

Оптичні пінцети використовуються не тільки для дослідження міжмолекулярних сил, але також для вивчення еластичних властивостей окремих полімерів і білків, згортання-розгортання білка, в'язких властивостей мембран.

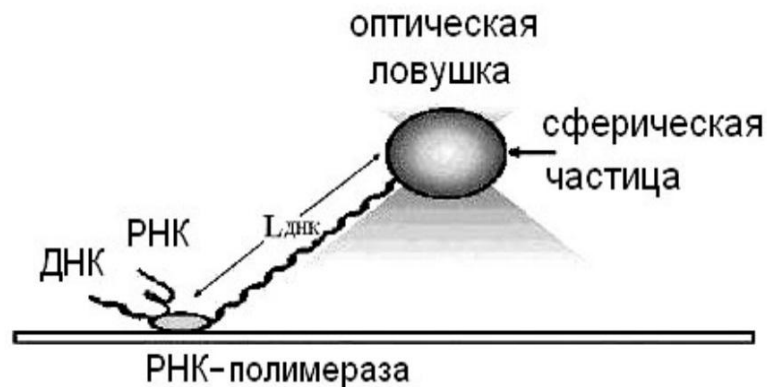


Рисунок 2.22 - Закріплена рецептор-лигандним зв'язком з поверхнею РНК-полімераза транскрибує ДНК

При вивченні процесів згортання-розгортання білка тітіна, в цьому експерименті були знайдені значення сил, при яких білок зберігає згорнуту форму або розгортається. Схема цього експерименту показана на рис. 2.23.

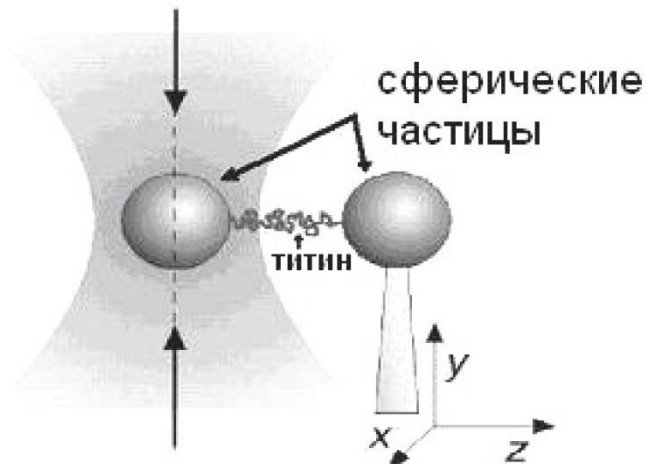


Рисунок 2.23 - Схема експеримента определения максимальной и минимальной сил разгортания и згортания белка титина

В эксперименте, показаному на рис. 2.24, визначається коефіцієнт в'язкості штучної мембрани, що складається з двох ліпідних шарів.



Рисунок 2.24 - Схематическое изображение эксперимента с визначення коефіцієнта в'язкості штучної мембрани

Відстань між сферичною часткою і оптичною пасткою скорочується за експоненціальним законом від часу, в аргумент експоненти входить коефіцієнт в'язкості мембрани

Оптичний пінцет вже зараз використовують для пересадки генів у клітини, а також при штучному заплідненні в пробірці.

Вельми цікаві експерименти виконані в угорському Біологічному дослідному центрі [61]. Там розроблена методика отримання мікроскопічних об'єктів довільної форми в результаті полімеризації клейкої маси під дією світла. Оптичний пінцет на основі інфрачервоного ($\lambda = 0,994$ мкм) напівпровідникового лазера захоплював і утримував у фокусі мікрочастинки. Далі використовувалася так звана двухфотонная методика: клей висвітлювали ультрафіолетовим лазером, генеруючим дві злегка розрізнячі довжини хвилі поблизу $0,340$ мкм, а необхідна для полімеризації інтенсивність досягалася фокусуванням в потрібній точці випромінювання аргонного лазера ($\lambda = 0,514$ мкм). У результаті впливу світла утворювався твердий полімер. Високоточний трьохкоординатний п'єзоелектричний маніпулятор, керований комп'ютером, переміщував матеріал щодо фокусу, створюючи мікроскопічні деталі - ротори, шестерінки, пропелери. А далі починається найцікавіше.

Було виявлено, що при зсуві точки фокусу світло, відхиляючись від частки, приводить її в обертання. Величина і напрям моменту обертання залежать від орієнтації ротора або шестерінки у фокусі. Якщо ротор оснащувався центральною осю, стійкість його захоплення в пінцеті підвищувалася, а при збільшенні числа зубців шестірні обертання ставало більш рівномірним. При потужності випромінювання 20 мВт конструкція рівномірно оберталася з частотою до декількох обертів в секунду. Звідси - один крок до створення діючих мікромашин, керованих світлом. Автори сконструювали два зчеплені шестерінки, що сидять на фіксованих осях, і вільно плаваючий ротор (рис 2.25). Ротор захоплювали лазерним пінцетом, приводили в обертання і потім підводили до пари шестерень, змушуючи їх крутитися.

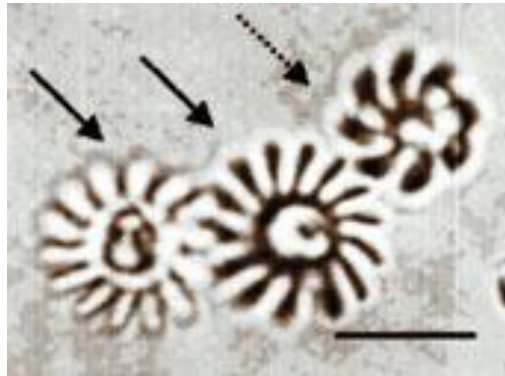


Рисунок 2.25 - Машина, що працює на енергії світла

Дві шестерінки, наведені в зачеплення, сидять на фіксованих осях (показані суцільними стрілками), а вільний обертовий ротор (пунктирна стрілка) підводиться до них лазерним пінцетом і приводить їх до руху.

Винахід оптичного пінцета вчинила справжню революцію в МІКРОТЕХНІКИ. Зараз в безлічі лабораторій ведеться відпрацювання методів його використання в різних областях.

2.7 Розробка основних вузлів оптичного пінцета

2.7.1 Система візуалізації мікрооб'єктів

Однією з основних функцій лазерного пінцета є візуалізація мікрочастинок, що переміщуються. Для досягнення мети створення цілісного пристрою маніпуляції мікрооб'єктів, необхідно розробити і створити систему, яка дозволить детектувати і розрізняти мікрооб'єкти розмірами від 1 мкм і більше. Сучасна мікроскопія пропонує ряд рішень даної проблеми. Існують такі види мікроскопії:

- оптична мікроскопія;
- рентгенівська мікроскопія;
- електронна мікроскопія;
- скануюча зондова мікроскопія.

Оптична мікроскопія дозволяє спостерігати мікрооб'єкти при загальному збільшенні до 2000 - 3000 разів [83], що дає можливість виявити елементи мікроструктури розміром 0,2 мкм і більше. Оптичні мікроскопи мають дозволяючу здатність до 150 нм.

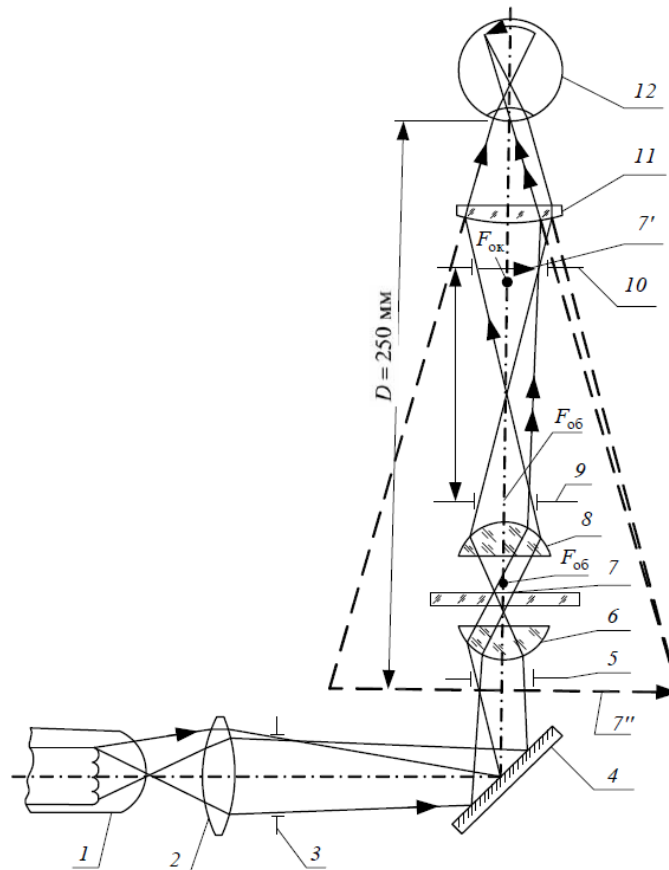
Рентгенівська мікроскопія призначена для дослідження дуже малих об'єктів, розміри яких порівнянні з довжиною рентгенівської хвилі. Вона заснована на використанні електромагнітного випромінювання з довжиною хвилі від 0,01 нм до 1 нм. Рентгенівські мікроскопи по спроможності знаходяться між електронними та оптичними мікроскопами. Теоретично дозволяюча здатність рентгенівського мікроскопа досягає 2-20 нм, що на порядок більше роздільної здатності оптичного мікроскопа. В даний час існують рентгенівські мікроскопи з роздільною здатністю близько 5 нм.

В електронній мікроскопії для побудови зображення замість світлових променів використовується пучок електронів. Це дозволяє збільшити роздільну здатність електронного мікроскопа порівняно зі світловим в сотні разів. Електронні мікроскопи дозволяють отримувати зображення об'єктів з максимальним збільшенням до 10^6 раз, завдяки використанню пучка електронів з енергіями 200 кеВ ÷ 400 кеВ і більше. Роздільна здатність електронного мікроскопа в 1000 ÷ 10000 разів перевершує дозвіл світлового мікроскопа і для кращих сучасних приладів може бути менше одного ангстрема. Для отримання зображення в електронному мікроскопі використовуються спеціальні магнітні лінзи, що керують рухом електронів в колоні приладу за допомогою магнітного поля.

Скануючий зондовий мікроскоп - мікроскоп для отримання зображення поверхні і її локальних характеристик. Процес побудови зображення заснований на скануванні поверхні зондом. У загальному випадку дозволяє отримати тривимірне зображення поверхні (топографію) з високою роздільною здатністю.

Виходячи із завдань лазерного маніпулятора мікрооб'єктами, найбільш придатною є оптична мікроскопія, яка за співвідношенням складності виконання конструкції до необхідних результуючим параметрам, не тільки не поступається, а й перевершує інші рішення.

Розглянемо схему найбільш типового оптичного мікроскопа, представлену на рис. 2.26.



- 1 - лампа освітлення; 2 - колімуюча лінза; 3 - польова освітлювальна діафрагма; 4 - дзеркало; 5 - ірисова діафрагма; 6 - мікрооб'єктив;
7 - прозорий предметний столик; 8 - мікрооб'єктив; 9 - апертурна діафрагма;
10 - діафрагма окуляра; 11 - лінза окуляра; 12 - окуляр.

Рисунок 2.26 - Принципова оптична схема мікроскопа

Умовно мікроскоп можна розбити на чотири вузли: освітлювальна система - позиції 1-6 рисунка, препарат 7, з площиною препарату пов'язана

польова діафрагма окуляра 10 і польова освітлювальна діафрагма 3, апертурна діафрагма 9, з якою пов'язана ірисова діафрагма 5, окуляр 12. Розглянемо хід променів у мікроскопі. Об'єкт 7, позначений стрілочкою, розташований на предметному столику 7 перед мікрооб'єктивом 8, на відстані трохи більшем його фокусної відстані $F_{об}$. Він буде дійсне, збільшене і перевернуте зображення 7' в площині діафрагми 10. Дане проміжне зображення розташоване за переднім фокусом окуляра 12.

Тому окуляр буде уявне збільшене зображення 7'' на відстані $D = 250$ мм від ока спостерігача. Така відстань вважається найкращею з погляду фізіологічних властивостей ока. Якщо ми хочемо отримати дійсне зображення об'єкта на телевізійному приймачі або фотоплівці, то достатньо окуляр зрушити так, щоб зображення 7' виявилось перед фокусом окуляра. Основні характеристики мікроскопа: видиме збільшення - Γ , лінійне поле зору - 2α , числова апертура - NA.

Числова апертура мікроскопа визначає його основні характеристики: світосилу і роздільну здатність. Числовою апертурою називають множення показника заломлення середовища, в якому знаходиться об'єкт, на синус апертурного кута.

Основними вузлами мікроскопа є освітлювальна система, мікрооб'єктив, предметний столик і окуляр.

Освітлювальна система мікроскопа складається з джерела оптичного випромінювання та оптичної системи, що забезпечує рівномірне освітлення об'єкту. Складнощі, що виникають при створенні освітлювальних систем, викликані тим, що в мікроскопах, особливо з великим збільшенням, необхідно застосовувати яскраві джерела світла. Тіло накалу таких джерел, наприклад нитку лампи, має нерівномірну яскравість, тому воно проектується оптичною системою, яка називається конденсором, що не точно в площину предмета, а ближче або далі. При такому освітленні утворюється багато розсіяного світла, який погіршує умови спостереження. Іншою проблемою є те, що числова

апертура конденсора повинна бути більше або дорівнює числовій апертурі об'єктива. В іншому випадку знижується роздільна здатність мікроскопа. Вирішення цих завдань призвело до створення різних типів освітлювальних систем мікроскопа. Найбільшого поширення набула схема освітлення по Келлеру, яка представлена на рис. 2.26 позиціями 1-6. Вона застосовується у всіх вітчизняних мікроскопах. Не вдаючись у подробиці принципу її роботи, можна сказати, що дана схема освітлення дозволяє отримати рівномірне освітлення поля мікроскопа від джерела з нерівномірною яскравістю тіла, яке світиться без погіршення якості зображення. У сучасних мікроскопах, наприклад конфокального, а також при використанні лазерних джерел випромінювання застосовуються і інші схеми освітлення.

Мікрооб'єктив є найважливішою частиною мікроскопа, що визначає якість зображення і світлосилу. Він являє собою багатолінзову систему, в якій виправлені аберації, які спотворюють зображення. Залежно від виду виправлених аберацій об'єктиви ділять на такі основні види: ахромати, апохромати, планахромати, планапохромати. Основними характеристиками об'єктива мікроскопа є збільшення і числова апертура. Остання визначає роздільну здатність мікроскопа - чим більше числова апертура, тим краще роздільна здатність. Залежно від збільшення і числової апертури об'єктиви мікроскопів можна розбити на три групи, представлені в табл. 2.1, яка визначає якість зображення і світлосилу.

Таблиця 2.1 – Групи об'єктивів мікроскопів

Групи об'єктивів	Збільшення	Апертура
Малих збільшень і апертур	$\Gamma \leq 10x$	$NA \leq 0,2$
Середніх збільшень і апертур	$\Gamma \leq 40x$	$NA \leq 0,65$
Великих збільшень і апертур	$\Gamma > 40x$	$NA > 0,65$

Числова апертура об'єктива залежить від показника заломлення середовища, в якому знаходиться препарат. Тому для збільшення NA використовують іммерсію: воду або масло, показник заломлення яких більше одиниці. Мікрооб'єктиви, які можуть працювати в таких умовах, називають імерсійними. Сучасні мікрооб'єктиви мають збільшення до $100\times$ і числову апертуру до 1,6.

Окуляр являє собою більш просту оптичну систему, ніж об'єктив. Його роль у формуванні зображення мікрооб'єкту набагато менш важлива і тому вимоги до нього нижчі. Основною характеристикою окуляра є збільшення, яке змінюється залежно від типу від 4 до 15 крат.

Важливим механічним вузлом мікроскопа є пристрій, що забезпечує його фокусування. Як правило, фокусування мікроскопа виробляється пересуванням тубоутримувача, в якому кріпиться об'єктив і окуляр, за допомогою грубого і мікротрендного механізмів. За допомогою них тубус, як єдине ціле, наближається або віддаляється від об'єкта до тих пір, поки оператор не приймає рішення про найкраще фокусування на препарат.

Однією з основних і найбільш відповідальних складових частин мікроскопа є об'єктив, тому від його числової апертури і корекції аберацій залежать роздільна здатність і якість зображення в цілому.

Об'єктиви мікроскопа класифікуються за спектральною областю, для якої вони розраховані і застосовуються; способу освітлення спостережуваного об'єкта; довжині тубуса, на яку вони розраховані; наявності покривного скла і іммерсионної рідини; конструктивним особливостям (лінзові, зеркальнолінзові і дзеркальні).

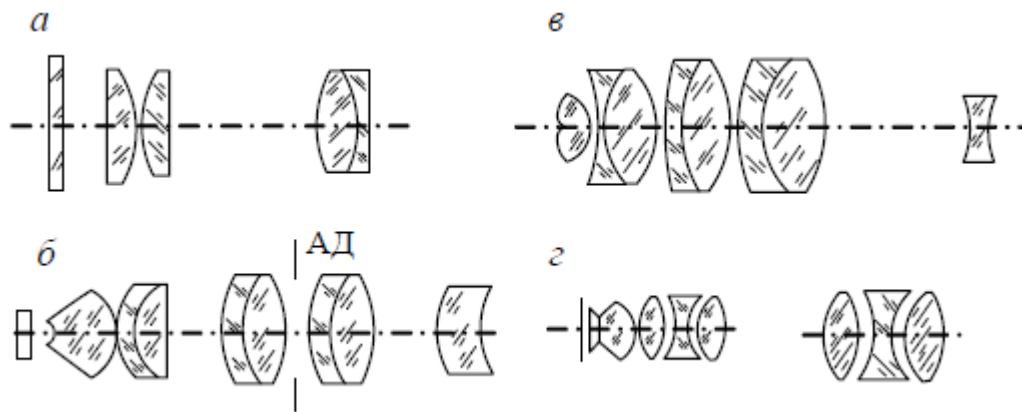
Великого поширення набули лінзові об'єктиви. Це викликано насамперед великими технологічними можливостями їх виготовлення в порівнянні з дзеркально-лінзовими. Для лінзових об'єктивів допуски задаються значно ширше, ніж для дзеркально-лінзових. Крім того, у лінзових об'єктивів відсутнє центральне екранування, що знижує контраст у зображенні.

За ступенем корекції розрізняють такі типи об'єктивів: монохромати, ахромати, апохромати, суперапохромати, планоб'єктиви. Монохромати - це об'єктиви, у яких аберації виправлені для вузької спектральної області. Об'єктиви, у яких ахроматизація виконана для двох довжин хвиль, називаються ахроматами. У апохроматов спектральна область розширена і ахроматизація виконується для трьох довжин хвиль. У суперапохроматов ахроматизація виконана для видимій і ультрафіолетовій області спектру. У об'єктивів з плоскою поверхнею зображення істотно збільшено поле.

У монохроматів в першу чергу виправляються сферична аберация, кома і астигматизм. У ахроматів виправленню підлягають: сферична аберация, кома і астигматизм, хроматизм положення і частково сферохроматичеськая аберация. У апохроматов і суперапохроматов крім абераций, що виправляються в ахроматов, значно краще коригуються вторинний спектр і сферохроматична аберация.

Планоб'єктиви за ступенем корекції діляться на планмонохромати, планахромати і планапохромати. Вимоги до корекції абераций для точки на осі планоб'єктивів такі ж, як і для відповідних монохроматов, ахроматов і апохроматов. Але на відміну від останніх у планоб'єктивів істотно краще виправлені кривизна поверхні і астигматизм. На рис. 2.27 наведені оптичні схеми деяких об'єктивів.

Оцінка якості зображення об'єктивів мікроскопа виробляється у хвильовій міри. Для основної довжини хвилі для точки на осі у всіх об'єктивів виконується критерій Релея, тобто хвильова аберация менш $0,25\lambda$. У ахроматов для всієї спектральної області, на яку вони розраховані, аберация не більше $0,5\lambda$. Для лінії G' хвильова аберация порядку $(1-2)\lambda$. У планоб'єктивів хвильові аберация для внеосевій точок предмета в межах всього поля не перевищують $0,5\lambda$. Остаточний хроматизм збільшення у сильних об'єктивів досягає 1,5-2,0% і компенсується застосуванням спеціально розрахованих компенсаційних окулярів.



а - ахромат (ОХ-22); б - планахромати (ОПХ-25);

в - планахромати (ОПА-60); г – суперахромат

Рисунок 2.27 - Принципові оптичні схеми лінзових об'єктивів різних типів

Наведена класифікація та кількісні критерії оцінки корекції аберацій є до певної міри умовними і наближеними. Не можна провести різкої межі, наприклад, між ахроматом і апохроматом, ахроматом і планахроматом і т. п.

До основних оптичних і механічних характеристик об'єктивів відносяться наступні: збільшення або фокусна відстань, числова апертура, довжина тубуса (механічна), висота об'єктива, робоча відстань, лінійне поле (зазвичай у просторі зображення), наявність покривного скла і імерсії.

Також важливою задачею є вибір методу світлової мікроскопії. Основними методами є:

- метод світлого поля в прохідному світлі;
- метод темного поля в прохідному світлі;
- метод ультрамікроскопії;
- поляризаційна мікроскопія;
- метод фазового контрасту;
- метод інтерференційного контрасту;
- метод дослідження в світлі люмінесценції;
- метод спостереження в УФ променях.

У розробленій системі візуалізації застосовується такий метод світлової мікроскопії, як метод світлого поля в прохідному світлі.

Цей метод застосовується при вивченні прозорих препаратів з включеними в них абсорбуючими (поглинаючими світло) частками і деталями. Це можуть бути, наприклад, тонкі забарвлені зрізи тваринних і рослинних тканин, тонкі шліфи мінералів і т. п. У відсутність препарату пучок світла з конденсора, проходячи через об'єктив, дає поблизу фокальної площини окуляра рівномірно освітлене поле. При наявності в препараті абсорбуючого елемента відбувається часткове поглинання і часткове розсіювання падаючого на нього світла, що й обумовлює появу зображення. Можливе застосування методу і при спостереженні неабсорбуючих об'єктів, але лише в тому випадку, якщо вони розсіюють висвітлюючий пучок настільки сильно, що значна частина його не потрапляє в об'єктив.

На підставі тих даних, які пропонує нам сучасна мікроскопія, для вирішення завдання візуалізації мікрооб'єктів з метою їх подальшої маніпуляції була розроблена система, схема якої показана на рис. 2.28.

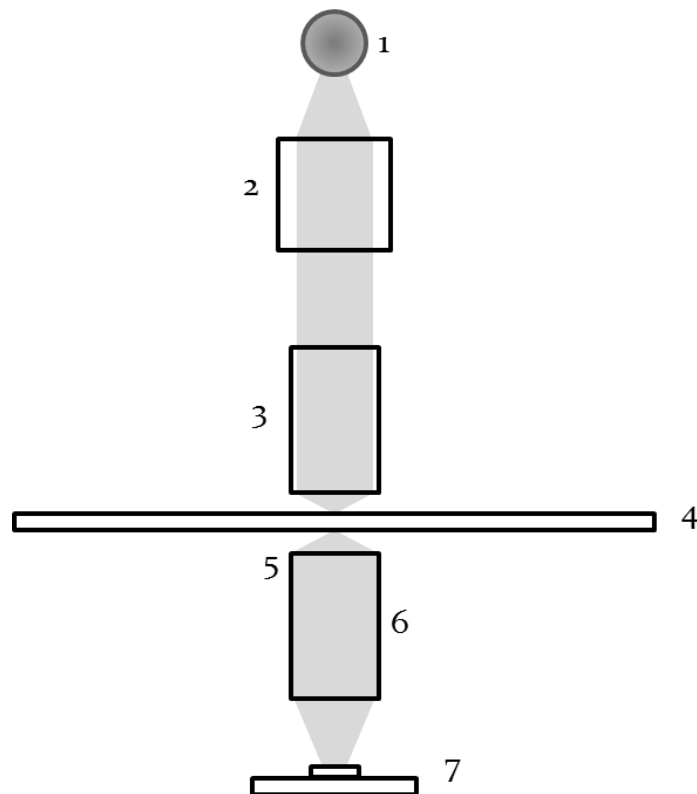


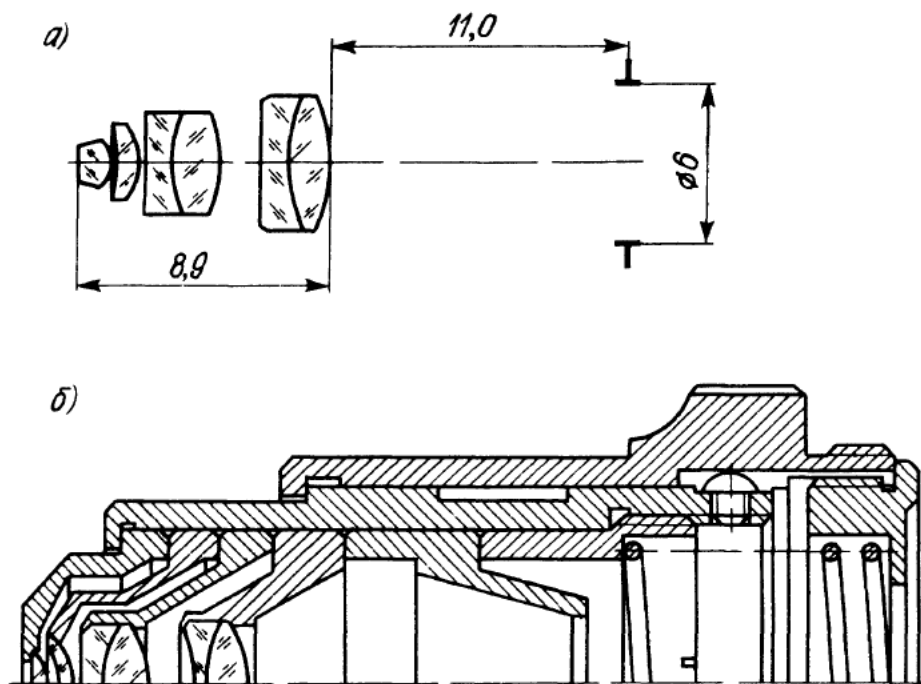
Рисунок 2.28 - Система візуалізації оптичного пінцета

У розробленій системі візуалізації оптичного пінцета світло, від джерела освітлення 1 потрапляє на коліматор 2, перетворюється в паралельний пучок і, проходячи через об'єktiv фокусування лазерного випромінювання 3, потрапляє на осередок, який досліджується. Проходячи через осередок, який досліджується, світло потрапляє на оптику КМОП - матриці 6, яка формує зображення на КМОП - матриці 7. Також перед оптикою КМОП - матриці встановлений світлофільтр 5, який запобігає засвіченню матриці лазерним освітленням, яке проходить.

Як джерело освітлення було вибрано світлодіодне освітлення широкого спектра з оптикою, розробленою за схемою освітлення Келлера.

Світлофільтр підібраний виходячи з того, що він найбільшою мірою повинен відсікати ту довжину хвилі, на якій випромінює лазерне джерело.

Оптична схема і конструкція підбраного об'єктива показані на рис. 2.29.



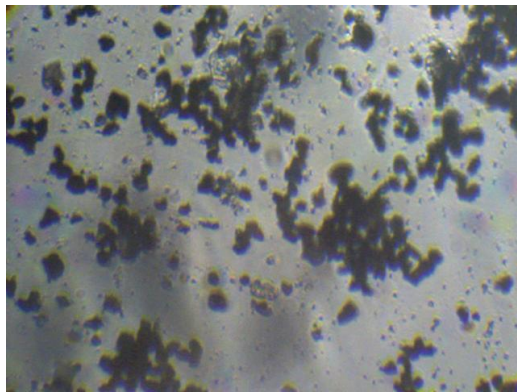
а) оптична схема; б) конструкція.

Рисунок 2.29 - Об'єktiv ахромат (90 x 1,25) масляної іммерсії

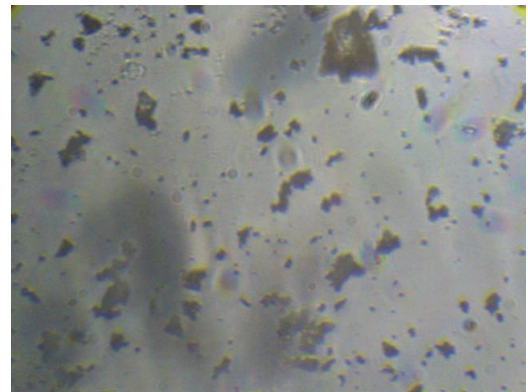
Об'єктив в розробленій системі представляє собою ахромат з числовою апертурою 1,25, збільшенням 90 крат, фокусною відстанню 1,96 мм і роздільною здатністю 0,22 мкм. Конструкція об'єктива, складається з чотирьох компонентів: двох позитивних лінз і двох склеювань з позитивних і негативних лінз.

В якості реєстратора зображення досліджуваного середовища була обрана КМОП - матриця з роздільною здатністю 0,3 Мп.

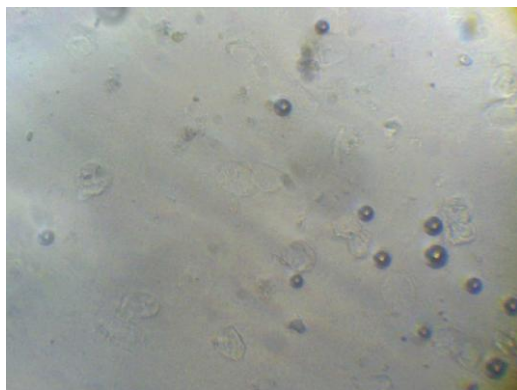
На розробленій системі були проведені експерименти з вивчення мікрооб'єктів. Результати проведених експериментів, у вигляді зображень представлені на рис. 2.30.



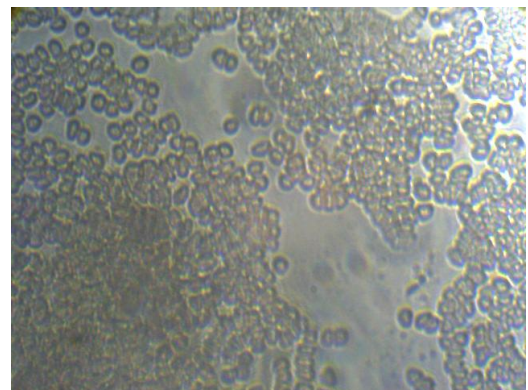
а)



б)



в)



г)

а) тонер; б) цемент; в) порошок; г) еритроцити.

Рисунок 2.30 - Досліджувані мікрочастинки

2.7.2 Система захоплення і утримання мікроб'єктів

З метою розробки системи захоплення і утримання мікроб'єктів лазерного пінцета було проведено ряд експериментів, спрямованих на вивчення властивостей сфокусованого випромінювання лазерних джерел.

Виходячи з теорії взаємодії високо сфокусованого лазерного випромінювання з мікро- і нанооб'єктами, сила впливу цього випромінювання на частку прямопропорційно залежить від величини градієнта потужності випромінювання в поперечному перетині лазерного пучка, як показано у формулі

$$F = \frac{1}{2} \alpha \cdot \nabla E^2. \quad (2.4)$$

Грунтуючись на цих даних можна зробити висновок, що для виконання поставленого завдання по захопленню й утриманню мікроб'єкта в оптичній пастці в першу чергу потрібно якісно вивчити розподіл потужності в поперечному перетині пучка.

Для виконання цього завдання в рамках дипломного проекту була розроблена і створена експериментальна установка, схема якої показана на рис. 2.31.

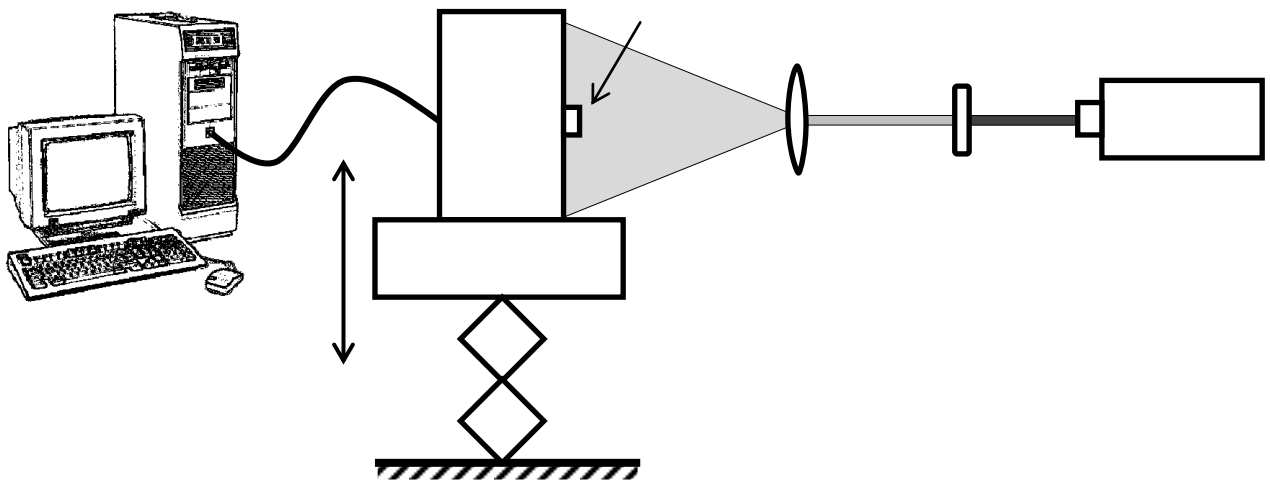


Рисунок 2.31 - Схема експериментальної установки

Експериментальна установка являє собою конструкцію, в якій випромінювання лазера (1), проходить через світлофільтр (2) і розсіюється на лінзі (3); розсіяне випромінювання частково потрапляє на ПЗС-лінійку, як показано на рис. 2.32. Пристрій ПЗС-лінійки 4 закріплено на координатному столику (6), який дозволяє переміщати ПЗС-лінійку у трьох напрямках. Також пристрій з ПЗС-лінійкою підключено до ПК, на якому в режимі реального часу відображається графік інтенсивності випромінювання, що потрапив на кожен елемент ПЗС-лінійки.

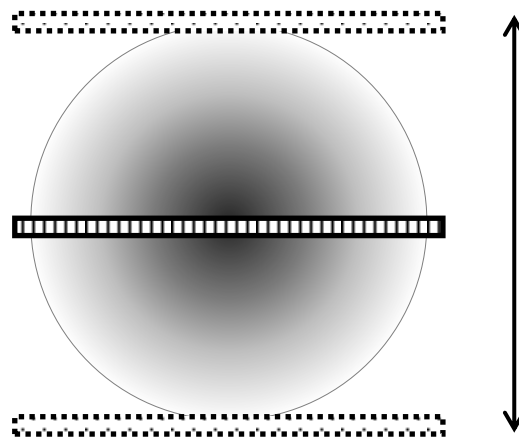


Рисунок 2.32 - Випромінювання лазера, що потрапляє на ПЗС - лінійку

В якості лазерів були випробувані два твердотільних лазера з діодною накачкою з довжинами хвиль 440 нм і 532 нм.

Вимірюваним приладом служить ПЗС-лінійка LPC1-USB фірми Thorlabs. Вона працює в діапазоні довжин хвиль від 350 до 1100 нм і складається з 3000 пікселів, розташованих в ряд.

Вид вікна програмного забезпечення для роботи з ПЗС-лінійкою показаний на рис. 2.33.

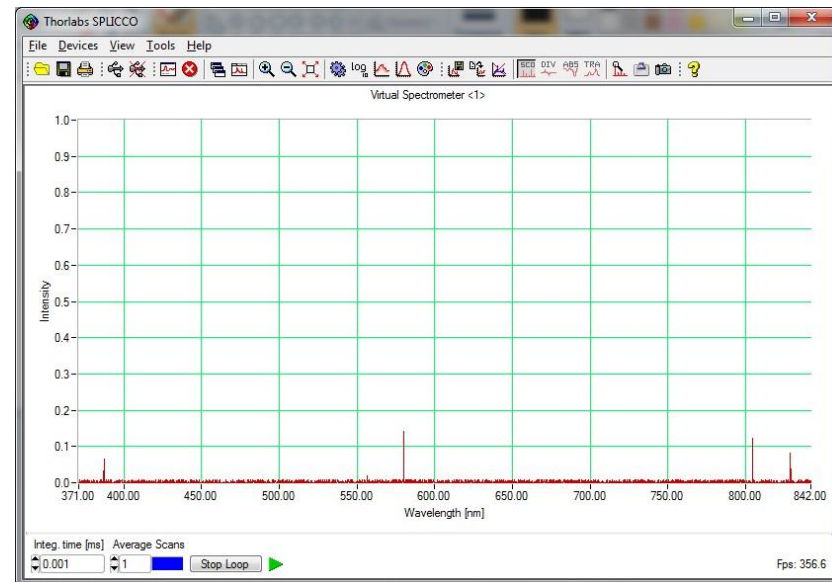


Рисунок 2.33 - Програмне забезпечення до ПЗС-лінійки

Розробку експериментальної установки, а також проведення експериментів, умовно можна розділити на кілька етапів.

На першому етапі була зібрана експериментальна установка за схемою, показаною на рис. 2.31.

Була забезпечена можливість переміщення ПЗС-лінійки, шляхом розміщення її на координатному столику (рис. 2.34).

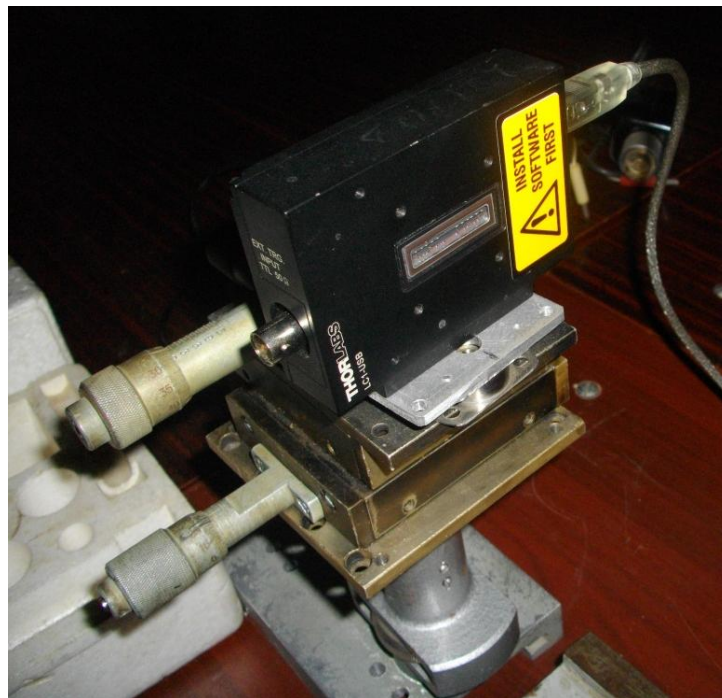
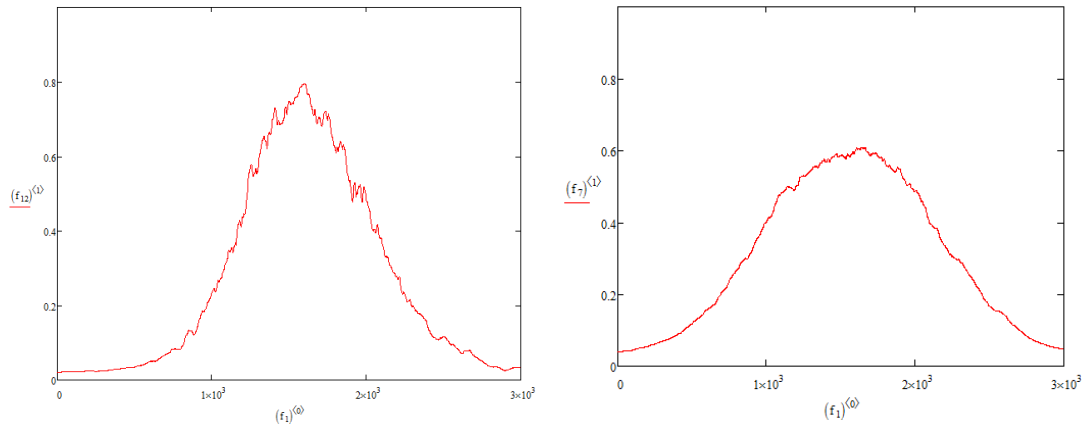


Рисунок 2.34 - ПЗС-лінійка, закріплена на координатному столику

На другому етапі шляхом сканування ПЗС-лінійки по перетину лазерного променя був знятий ряд графіків, які представляють собою залежність інтенсивності від номера пікселя в лінійці. Приклад таких графіків представлений на рис. 2.35.



а) для лазера з довжиною хвилі $\lambda = 473\text{нм}$;

б) для лазера з довжиною хвилі $\lambda = 532\text{нм}$.

Рисунок 2.35 - Графіки залежності потужності лазерного випромінювання від номера пікселя ПЗС-лінійки

У ході виконання експерименту було отримано велику кількість графіків (рис. 2.35) для експериментів з різними параметрами. Для проведення аналізу отриманих результатів на третьому етапі експерименту було розроблено програмне забезпечення, яке дозволяє відображати розподіл потужності лазерного променя в поперечному перетині у вигляді тривимірного графіка. Результати таких побудов представлений на рис. 2.36.

За підсумками проведених експериментів, можна зробити висновок, що отримані результати свідчать про те, що вивчені лазери з довжинами хвиль 473 нм і 532 нм володіють розподілом профілю інтенсивності близькому до розподілу Гаусса. Також, виходячи з отриманих тривимірних графіків розподілу потужності лазерного випромінювання в поперечному перетині променя, можна зробити висновок, що лазерне джерело з довжиною хвилі 473

нм володіє великим градієнтом потужності в поперечному перетині, ніж джерело з довжиною хвилі 532 нм.

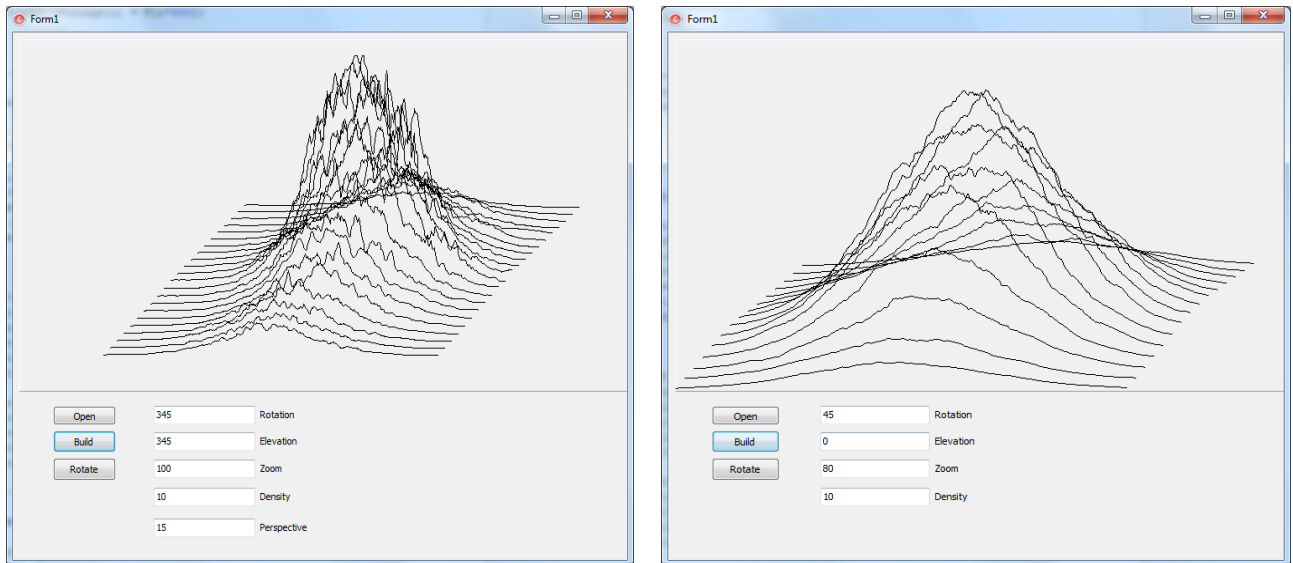


Рисунок 2.36 - Тривимірні графіки розподілу потужності лазерного випромінювання в поперечному перерізі променя

Подальший експеримент був спрямований на вивчення фізичних розмірів формованої перетяжки, що також важливо. Актуальність проведення даного експерименту обумовлена тим, що фізичні розміри перетяжки визначають той мінімальний розмір частки, при якому вона може бути захоплена.

Для проведення експерименту була розроблена і створена експериментальна установка, схема якої показана на рис. 2.37.

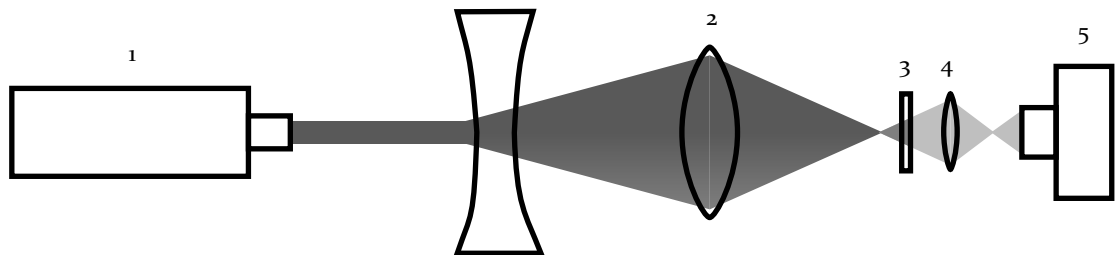


Рисунок 2.37 - Експериментальна установка

Згідно зі схемою на рис. 2.37 випромінювання лазера (1) фокусувалася за допомогою об'єктива (2) і проходячи через світлофільтр (3) і мікрооб'єктив (4) потрапляло на ПЗС-матрицю (5).

Суть експерименту полягає в підстановці в експериментальну установку різних об'єктів та реєстрації розподілу потужності лазерного випромінювання в поперечному перетині пучка за допомогою ПЗС-матриці на ПК.

Таким чином, в ході проведення експерименту були вивчені властивості перетяжок, формованих різними об'єктивними.

За підсумками проведеного експерименту була обрана схема об'єктиву, при використанні якої виходить найменша перетяжка, а також встановлена межа мінімальної частки, яка може бути захоплена за допомогою даного об'єктиву. Мінімальний розмір частки виявився рівним 1 мкм.

2.7.3 Система переміщення мікрооб'єктів

Переміщення часток - одне з основних завдань оптичного пінцета. Існує кілька основних способів переміщення частинки відносно предметного столика:

- переміщення частки;
- переміщення предметного столика.

Перший спосіб полягає в захопленні частки і переміщенні її шляхом сканування променя відносно нерухомого предметного столика. Другий спосіб полягає в захопленні і утриманні частки, в той час як здійснюється переміщення предметного столика.

Реалізація першого методу полягає в розробці такої оптичної схеми, яка дозволить нам здійснювати переміщення вихідної лінзи або меніска в площині перпендикулярній оптичній осі системи фокусування лазерного випромінювання. Схема реалізації показана на рисунку 2.2.

Реалізація другого методу полягає в розробці механічної системи, яка дозволить переміщати предметний столик в площині, перпендикулярній оптичній осі системи фокусування лазерного випромінювання з необхідною точністю.

Одним з найбільш відповідальних вузлів такого механізму є направляючі для точних переміщень.

Направляючими називають пристрої для переміщення встановлених на них елементів по заданій траєкторії.

По виду руху направляючі поділяють на три групи:

- направляючі для поступального руху;
- направляючі для обертального руху;
- направляючі для обох видів руху.

В залежності від роду тертя розрізняють направляючі:

- з тертям ковзання;
- з тертям кочення;
- з молекулярним (пружним) тертям;
- з повітряним або рідинним тертям.

Для характеристики направляючих використовують й інші ознаки:

- наявність або відсутність силового або кінематичного замикання;
- форму тіл кочення або деталей, які переміщуються;
- використовувані матеріали і т.д.

У загальному обсязі приладу направляючі складають значну частину - до 25 - 30% деталей і також визначають точність їх функціонування.

Розглянемо направляючі з тертям ковзання. На тертя ковзання впливають багато фізико-механічні, конструктивні та технологічні чинники, детермінованого і випадкового характеру. Це пов'язано не тільки з особливостями самого тертя, але і з особливостями роботи механізмів (перехід від спокою до руху, випадковий характер зміни геометрії і шорсткості деталей

при переміщенні тощо). Для зменшення тертя використовують мастила різного типу і консистенції.

Тертя ковзання використовується в направляючих столиків, лінзових систем, об'єктивів та інших механізмів. Деталі направляючих виготовляються зазвичай з латуні або поєднання латуні з алюмінієвим сплавом. Швидкості переміщення направляючих знаходяться в межах від 0,0002 до 15 мм / сек.

Направляючі ковзання мають достатньо плавний рух і витримують великі навантаження. До числа їх недоліків відносяться:

- велике тертя і великі зусилля рушання;
- збільшені осьові розміри;
- чутливість до змін температури.

Розглянемо направляючі ковзання типу "ластівчин хвіст", конструкція яких визначається технологією їх виготовлення. Вони застосовуються для переміщення призм, дзеркал, предметних столиків та інших вузлів і деталей приладів.

У звичайних направляючих типу "ластівчин хвіст" (рис. 2.38, а) знімна бічна планка (а) використовується для регулювання зазорів між деталями при складанні. Таке регулювання зазорів проводиться вручну і досить трудомістке, тому що пов'язане з ручною шабровкой і притиранням деталей.

Тому розроблена технологія і конструкція направляючих (рис. 2.38, б), які усувають ручні операції і підвищують точність переміщення. У корпусі (1) рухається клин (2), скріплений з планкою (3). З кресленника видно, що ця конструкція не має деталей для пригону. Працездатність направляючої забезпечується технологією виготовлення.

Обробка деталей проводиться на спеціальному трьохшпіндельном верстаті. Спочатку обробляють різцями (4) (рис. 2.38, в) і (4') похилі площини в корпусі (1). Різці встановлені в шпіндельних голівках (5) і (6), які при установці ширини паза "а" переміщуються в напрямку стрілок А. Потім обробляють клин (2) (рис. 2.38, г) тими ж різцями, попередньо розсунувши

шпинделі (5) і (6) в напрямку Б. При цьому ширину клина за розміром "а" встановлюють на 0,2 ... 0,4 мм менше, ніж ширина паза в деталі (1). Нарешті, деталь (2) вставляють у паз деталі (1) і притискають її знизу, в напрямку стрілки В (рис. 2.38, д). Оскільки розмір "а" клина (1) менше ширини паза в корпусі (2), то клин виступає над верхньою поверхнею корпусу. Різцями (7), встановленими у вертикальній шпиндельній голівці (8), знімають верхній шар товщиною 0,1 - 0,2 мм у обох деталях. В результаті утворюється загальна площина деталей (1) і (2), на яку встановлюється планка (3) (рис. 2.38, б). Зазори в таких направляючих не перевищують 3,5 мкм, а самі направляючі практично не вимагають пригону.

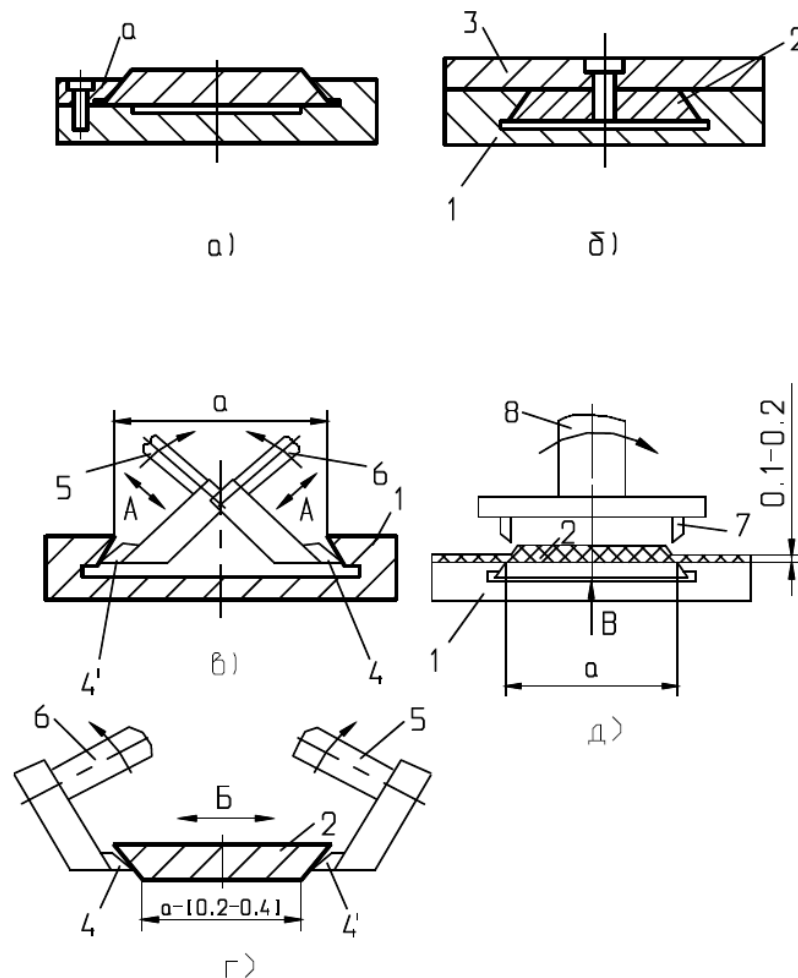


Рисунок 2.38 – Направляючі ковзання

На рис. 2.39 показана конструкція з тертям ковзання, використовувана в столику, що має обертання навколо осі Z . Конструкція має загальну єдину поверхню тертя (А) для поворотної частини і підстави. Таке рішення дозволяє зменшити розфокусування столика по осі Z при обертанні поворотної частини і зменшує габарити столика по висоті.

Найбільш відповідальні механізми використовуються в режимі позиціонування (наводки, суміщення). Позиціонування вздовж оптичної осі Z або в площині, перпендикулярній до неї, за координатами X і Y з точністю до 0,0001 - 0,0002 мм необхідно при мікроспектрофотометрії клітин, виготовленні та контролі мікросхем, лінійних і кутових вимірах, суміщенні міток в полі зору і в інших випадках.

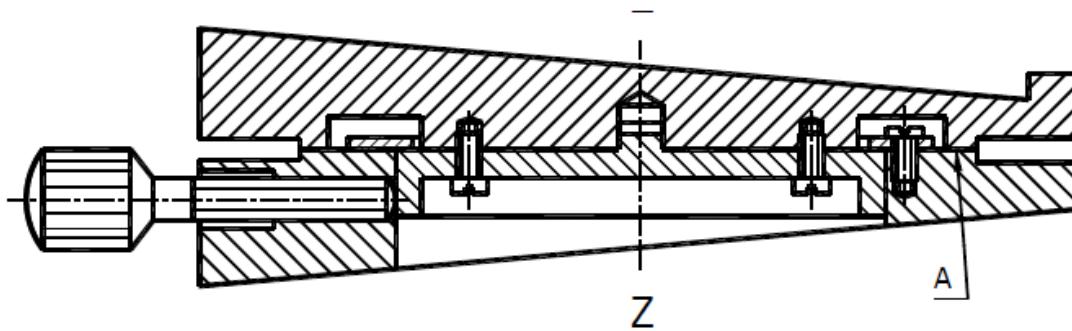


Рисунок 2.39 - Направляючі ковзання для обертювих деталей

Для оцінки точності позиціонування по осях X і Y використовується методика, заснована на вимірюванні зони розкиду. При підведенні еталонного об'єкта до штриху окулярного мікрометра з одного боку спостерігач припиняє переміщення об'єкта до штриху в момент їх найкращого суміщення. Потім він вимірює величину недовода або переходу за штрих, визначаючи тим самим зону розкиду.

Суттєвою причиною, що впливає на точність наведення, є стрибкоподібний характер руху, викликаний багатьма причинами, частина з яких має випадковий характер. Цей процес аналогічний поняттю «шум» в радіоелектроніці. Тому перспективним є застосування гелієвих мастил, що

дозволяють зменшити величину стрибків (знизити «шум» руху направляючих) і демпфуючих точний настановний рух.

Розглянемо також направляючі кочення. Направляючі кочення мають значно менше тертя і зусилля рушання. Вони нечутливі до коливань температури і невимогливі до вибору мастила. Але вони часто мають великий габарит і більш високу вартість виготовлення на відміну від направляючих з тертям ковзання.

У направляючих з тертям кочення використовують кулькові, роликові або голчасті тіла кочення. Мертвий хід в них не перевищує 2.4 мкм.

Різні конструкції, використовувані у вузлах з тертям кочення, показані на рис. 3.40.

На рис. 2.40, а зображена типова конструкція направляючих, у яких і рухома каретка (1) і нерухомі бічні напрямні (2) виготовлені у вигляді монолітних деталей зі сталі 20 або ШХ - 15 з подальшою цементацією і гартуванням до твердості НЯС 55.60. Така конструкція характеризується тривалим циклом виготовлення деталей та їх високою трудомісткістю.

На рис. 2.40, б наведені збірні направляючі, в яких використовуються різні вкладиші, виконані у вигляді кутників (1), дротиків(2), дротиків зі скосами (3) і плоских прокладок (5). У таких конструкціях направляючі (4) і каретки (6) можуть бути виготовлені з алюмінієвих сплавів або сталі без термообробки, що істотно спрощує і здешевлює виготовлення. Відхилення від прямолінійності і неплоскості не повинні перевищувати 0,005 - 0,01 мм. Для цього, щоб не допустити великого розкиду розміру "а", обидві направляючі часто шліфують разом, попарно, розташовуючи їх на верстаті послідовно притиснутими до загальної плоскості. Відстань L між кріпильними деталями на бічних направляючих (4) можна визначити за формулою:

$$L = 5,5 \cdot \sqrt[4]{h}, \quad (2.5)$$

де L - висота направляючих, мм.

Направляючі повинні збиратися з натягом в 2 ... 3 мкм, який створюється за допомогою гвинтів (7) (див. рис. 2.40, а). Використовується і спеціальна технологія збірки, яка показана на рис. 2.40, в. Основа направляючих (1) разом з бічними планками деформується за допомогою скоби (2) і болта (3).

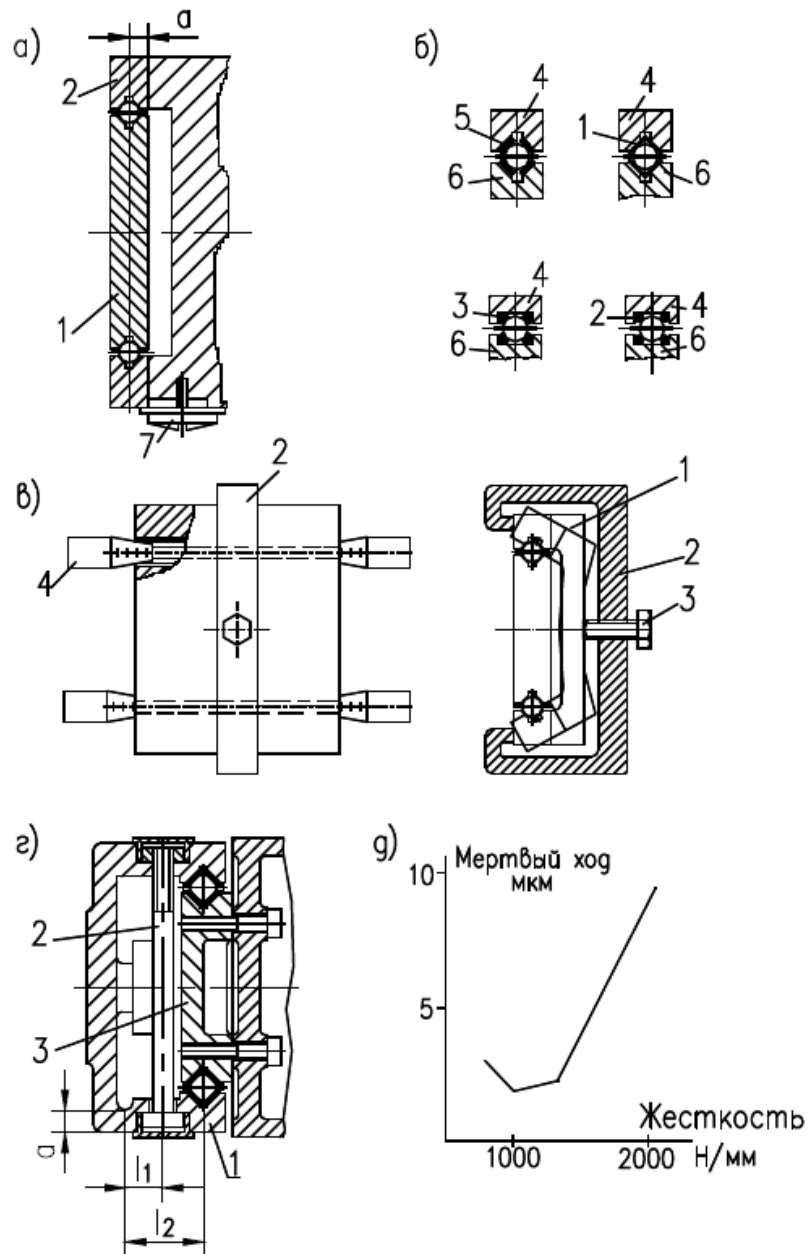


Рисунок 2.40 - Направляючі кочення для поступального руху

Величина деформації контролюється за допомогою конічних пробок (4), на які нанесені поділкі, відповідні зміні діаметрів конусів через 2...5 мкм. Після установки направляючих, в канавки насипають кульки, заздалегідь розсортовані по діаметру. Величина деформації підбирається такою, щоб після

видалення скоби (2) направляючі були зібрані з необхідним натягом кульок, який попередньо визначається експериментально.

На рис. 2.40, г наведена конструкція з пружною бічною направляючою (1). Стиснення цієї направляючої і регулювання натягу проводиться за допомоги гвинтів (2). Вигин відбувається завдяки деформації тонкої перемички товщиною "а", що дорівнює 3 ... 5 мм. Корпус з направляючими (1) виготовляється з алюмінієвих сплавів, а каретка (3) - зі сталі, латуні або алюмінієвих сплавів. Конструкція прокладок під кульки виконана по одному з варіантів, показаному на рис. 2.40, б. Дослідження показали, що за допомоги гвинтів (2) (див. рис. 2.40, в) можна оптимізувати співвідношення між мертвим ходом і жорсткістю конструкції направляючих (рис. 2.40, д).

Виконаємо розрахунок такої конструкції. На рис. 2.40, г позначено:

l_1 - відстань від перемички до гвинта, 30 мм;

l_2 - відстань від перемички до кульки, 40 мм;

а - товщина перемички, 4 мм.

Можна прийняти, що деформація пружної направляючої (5) знаходиться в межах $\delta = \pm 0,2$ мм, оскільки допуск на відстань між кульками становить у кожній деталі $\pm 0,1$ мм.

Переміщення гвинта δ_B для отримання такої деформації дорівнює $\delta_B = \delta \cdot l_1 / l_2 = 0,2 \cdot 30 / 40 = 0,15$ мм.

Зусилля Р для вигину направляючої можна розрахувати за формулою:

$$P = \frac{\delta_B \cdot l \cdot a^3 \cdot E}{4 \cdot l_1^3}, \quad (2.6)$$

де l - довжина направляючої, 80 мм; E - модуль пружності алюмінієвих сплавів, рівний 70000 Н / мм^2 .

$$P = \frac{0,15 \cdot 80 \cdot 4^3 \cdot 70000}{4 \cdot 30^3} = 498 \text{ Н.}$$

Перевіримо напругу вигину в перемичці і розрахуємо його за формулою:

$$\sigma_{II} = \frac{6P \cdot l_1}{l \cdot a^2}. \quad (2.7)$$

$$\sigma_{II} = \frac{6 \cdot 498 \cdot 30}{80 \cdot 4^2} = 70 \text{ (Н / мм)},$$

що допустимо.

Перевірка корпусу на зріз різьби і стискування показали, що напруги не перевищують допустимих.

Розглянемо направляючі з молекулярним тертям та інші типи направляючих.

У випадках, коли потрібен мінімальний мертвий хід, висока відтворюваність траєкторії руху, підвищена зносостійкість і відсутність необхідності мастила, використовують направляючі з тертям пружності - пружинні направляючі (рис. 2.41). Зазвичай використовують направляючі з плоскими пружинами у вигляді простого паралелограма (рис. 2.41,а) або подвійного паралелограма (рис. 3.41,б), запропонованого Клеєм в 1888 р для вимірювання вібрацій. При невеликих кутах вигину траєкторія, що описується кінцем пружини, близька до дуги кола з радіусом $R = 3/41$. При вигині пружин у паралелограмі відбувається зсув основи / на величину Δl . При застосуванні подвійного паралелограма зсув Δl , що виникає при згині пружини (2) в одному напрямку, компенсується таким же зміщенням, але спрямованим в іншу сторону іншого паралелограма, утвореного пружинами (3).

Для підвищення точності переміщення використовують примусове переміщення другого паралелограма в два рази повільніше за допомогу важільного чи іншого механізму. У механізмі точної подачі (рис. 2.41,в), що забезпечує переміщення з "шумами" не більше 0,01 мкм на довжині 100 мкм, примусовий рух на паралелограми у відношенні 1: 2 передається стержнем. При обертанні гвинта (4) через важіль (5) пружини (7) створюють силове замикання механізму. На точність механізму великий вплив робить розкид

довжини пружин l і розмірів d , їх товщини, ширини, прогину, напрямку прокату та інших властивостей пружин. Для виконання вимог по точності переміщення застосовують спільну обробку деталей за розмірами d , виготовляють пружини з одного рулону стрічки і використовують інші технологічні прийоми. Пружини виготовляються зі сталі 70С2ХА товщиною 0,2 ... 0,5 мм і шириною 20 ... 30 мм.

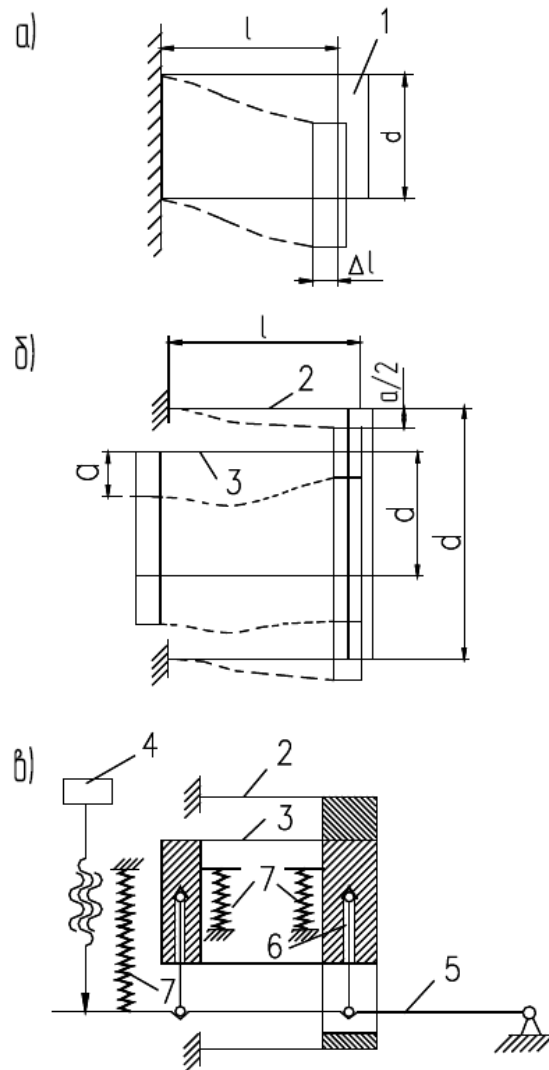


Рисунок 2.41 - Пружинні направляючі

Виходячи з розглянутих вище методів реалізації направляючих точного переміщення, був розроблений предметний столик, з двома ступенями свободи.

Його креслення показано на рисунку 2.4. Розроблений пристрій дозволяє здійснювати високоточне ручне переміщення захопленої мікрочастини.

2.7.4 Охолоджені атоми в фотонних кристалах

Оптичний частотний репер являє собою середовище, атоми або молекули якого мають лінію поглинання, частота переходу якої збігається з частотою випромінювання лазера. Активна стабілізація частоти здійснюється по відношенню до частоти репера. Перші стандарти частоти використовували частотні реperi на основі газоподібних середовищ, поміщених в оптичні осередки. З розширенням номенклатури лазерів, частота яких повинна бути стабілізована, розширювався список оптичних реперів. У цей список стали включатися охолоджені атоми і іони. Їх утримання здійснюється у вакуумних камерах з використанням охолоджувальних лазерних променів. Щоб сформувати частотні реperi, які представляють собою компактні пристрої необхідно використовувати нові технічні рішення на основі нових фізичних явищ. У даній роботі розглядаються умови використання фотонних кристалів в якості оптичного середовища, в якому розміщуються і утримуються поглинаючі атоми, які в свою чергу повинні взаємодіяти з частотою, яка стабілізується. Обґрунтування використання фотонних кристалів в роботі було вивчено починаючи з умов роботи лазерного пінцета.

У 1987 році А. Ешкін [49] продемонстрував можливість утримання частки в заданій області простору за допомогою сфокусованого пучка лазерного випромінювання. Завдяки великому поперечному градієнту інтенсивності в сфокусованих лазерних пучках, частка, яка перебуває на осі цього пучка, залишається тривалий час і при переміщенні лазерного пучка в просторі частка переміщується разом з останнім. Грунтуючись на цих результатах, були розроблені конструкції лазерних пінцетів. Для біомедичних експериментів утримання часток за допомогою сфокусованого лазерного пучка являє собою

дуже ефективний оптичний спосіб без механічного травмування, вибору і перенесення в просторі біологічних клітин. Цей спосіб вже реалізований в лазерних пінцетах, що серійно випускаються і які використовуються в біомедичних експериментах. Слідуючи вивченим умовам захвата і утримання наночастинок, можна припустити, що за допомогою будь-якого іншого поля, спеціальним чином сформованого з неоднорідною інтенсивністю в поперечному напрямку, також можна реалізувати умови захвата і утримання наночастинок.

Теорія фотонних кристалів дозволяє вирішувати завдання з управління поширенням оптичного випромінювання на відстанях порядку довжини хвилі випромінювання. Зараз це самостійний розділ оптики, який ідентифікується як нанофотоніка і забезпечує подальший розвиток оптичного зв'язку.

Фотонні кристали залежно від їх внутрішньої структури можна розглядати як провідники, ізолятори, напівпровідники або надпровідники. Головна особливість такого поділу полягає в існуванні дозволених і заборонених зон. Фотонні «провідники» мають широкі дозволені зони. Це прозорі тіла, в яких світло поширюється на великі відстані, практично не поглинаючись. Інший клас фотонних кристалів - фотонні ізолятори - володіють широкими забороненими зонами (наприклад, широкодіапазонні багат шарові діелектричні дзеркала), в яких світло не згасає, перетворюючись на тепло, а створюються умови для непропускання світла. Фотонні напівпровідники мають більш вузькі в порівнянні з ізоляторами заборонені зони.

Інша можливість, яка реалізована в фотонних кристалах, [50] полягає в тому, що при наявності дефектів у структурі фотонного кристала, в них формуються поля з дуже великими поперечними градієнтами, оскільки за межі дефекту поле не поширюється. Такий розподіл поля є необхідною умовою для захвата і утримання мікро і нано часток. Із захваченими наночастинами, молекулами і атомами в дефектах фотонних кристалів може взаємодіяти інше інформаційне електромагнітне поле. В таких умовах, коли наночастки

знаходяться в захваченому стані, з ними можна працювати в умовах оптичного комп'ютера, оптичного стандарту частоти або будь-якого іншого засобу вимірювань, в основі якого використовується спектр поглинання поодиноких атомів або молекул.

У загальному випадку, для наведених практичних завдань, необхідно забезпечити утримання наночасток за допомогою електромагнітного поля не обурюючи, при цьому, їх спектр поглинання. Взаємодія іншого електромагнітного поля буде носити інформаційний характер, який полягає в тому, що взаємодія поля з часткою буде забезпечувати передачу інформації від одного об'єкта іншому. У цьому звіті розглядається можливість використання захвачених наночасток для створення частотних реперів оптичних стандартів частоти. У цьому зв'язку, метою цієї роботи була постановка задачі, яка пов'язана із застосуванням фотонних кристалів для утримання наночасток.

В роботі [51] була досліджена можливість використання фотонних кристалів з дефектом для лазерного управління охолодженими атомами. У роботі обговорювалося використання пучка атомів цезію, які охолоджені в поперечному напрямку. Розглядалася можливість локалізації атомів в дефекті фотонного кристала, було розраховано поле в дефекті в умовах проникнення лазерного випромінювання в останній. Проте в роботі задача була орієнтована на виконання квантових обчислень.

В даний час розвиток оптичних стандартів частоти обмежене тільки технічною можливістю реалізації частотних реперів. Однією з найбільш перспективних технологій формування частотного репера є лазерне охолодження атомів і іонів. Здійснити таку технологію можна тільки в умовах вакууму і тільки при наявності не менше шести лазерних променів, частота випромінювання яких автономно управляється залежно від температури сформованої оптичної патоки [52]. Надалі, після охолодження патоки до 10^{-6} К, її можна використовувати в якості частотного репера, у якого доплеровське зрушення частоти буде мінімальним. Головна проблема застосування в такому

виконанні частотного репера полягає в тому, що патока повинна постійно перебувати у вакуумі та під дією потужного лазерного випромінювання. Таким чином, технічний частотний репер являє собою складний, габаритний пристрій.

Справжня робота спрямована на пошук умов, за яких можна здійснити захват охолоджених атомів в фотонному кристалі. Основна умова полягає в тому, що використовується два поля. Одне поле, частота якого забезпечує локалізацію цього поля в дефекті і друге поле, частота якого збігається з частотою поглинаючого переходу і взаємодіє з охолодженими атомами.

Для розрахунку поля в дефекті в роботі розглядається можливість застосування чисельної процедури, заснованої на прямому інтегруванні рівнянь Максвелла методом кінцевих різниць (техніка FDTD). Цей метод дозволяє досліджувати розподіл поля в двовимірних фотонних кристалах. Розрахунок поля при наявності дефекту в кристалі був проведений в [53]. Там же наведена схема (рис. 2.41) очікуваного експерименту, який реалізує взаємодію поля з охолодженими атомами. Теоретичне дослідження умов поширення електромагнітних хвиль у фотонному кристалі передбачає облік граничних і початкових умов.

У роботі обговорювалося питання використання кремнієвих технологій для виготовлення двовимірних і тривимірних фотонних структур з дефектами. Інтенсивність світла в центрі дефекту більш ніж на два порядки перевищує інтенсивність на кордоні каналу, що забезпечує дуже високий градієнт поля в поперечному напрямку, який впливає на утримання атомів в центрі дефекту.

Якщо в фотонному кристалі сформований лінійний дефект, то пучок атомів орієнтується уздовж дефекту, і буде поширюватися вздовж нього. Оскільки інтенсивність світла в центрі дефекту більш ніж на два порядки перевищує периферійну інтенсивність, то градієнтна сила буде утримувати атоми в центрі дефекту.

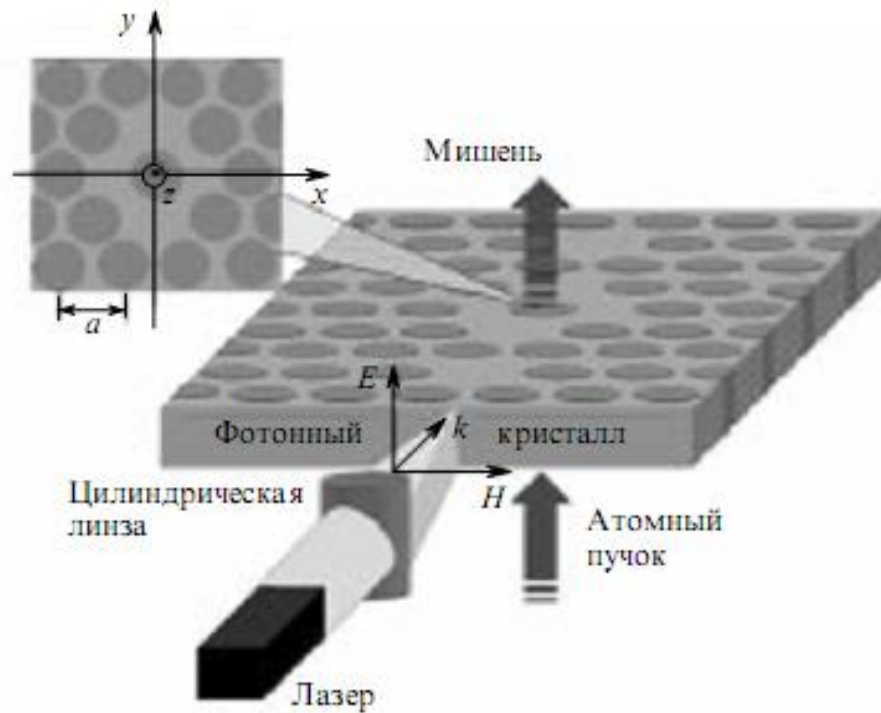


Рисунок 2.41 - Схема взаємодії лазерного випромінювання сконцентрованого в дефекті фотонного кристала з охолодженими атомами

Таким чином, в даній роботі пропонується схема реалізації частотного репера на основі атома, який захвачений в дефекті фотонного кристала одним полем і взаємодіє з другим полем (рис. 2.42). Перший лазер, як було показано вище, створює в дефекті фотонного кристала такий розподіл поля, який, завдяки високому градієнту інтенсивності, дозволяє утримати атом або частку. Цей атом або частка використовується безпосередньо для стабілізації іншого лазера. Суть конструкції полягає в тому, що параметри фотонного кристала підібрані так, що він є абсолютно прозорим для другого поля. Таким чином, взаємодія поля лазера, частота якого стабілізується, обмежується лише взаємодією з захваченим атомом.

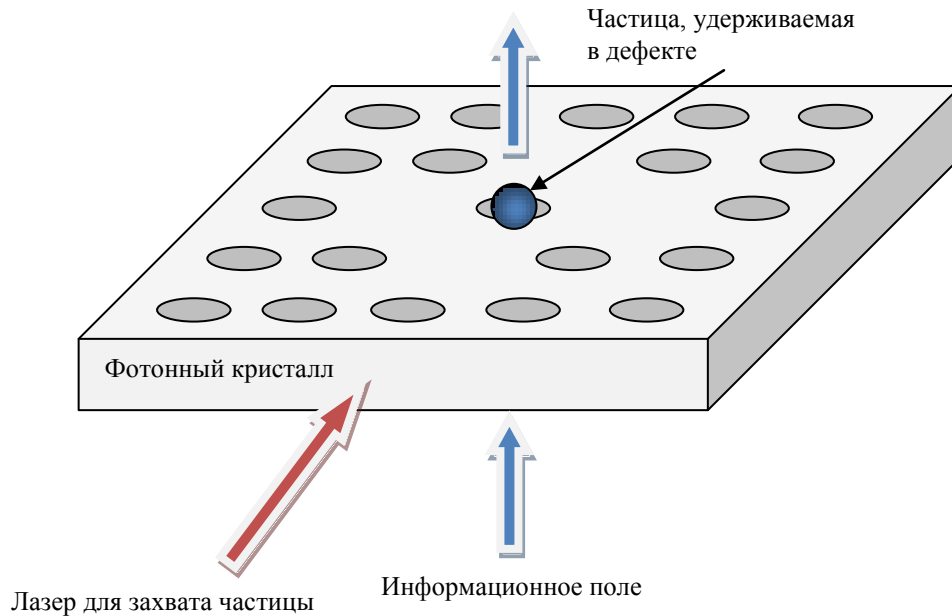


Рисунок 2.42 - Схема взаємодії частки в дефекті фотонного кристала з двома полями

Така конструкція реалізації частотного репера має значні переваги по відношенню до вже відомих і використовуваним конструкціям. По-перше, для стабілізації використовуються охолоджені атоми, що означає мінімальне доплеровське зрушення частоти поглинання у поперечному напрямку. По-друге по відношенню до існуючих конструкцій, які використовують охолоджені атоми, конструкція на фотонних кристалах має явну перевагу в габаритних параметрах, а також у відсутності необхідності використовувати велике число лазерів, що значно спрощує і конструкцію, і умови її експлуатації.

Оптичний (або лазерний) пінцет (рис. 2.43) - це пристрій, що використовує сфокусований промінь лазера для пересування мікроскопічних об'єктів або для утримання їх у певному місці. Поблизу точки фокусування лазерного променя світло тягне до фокусу все, що знаходиться навколо.

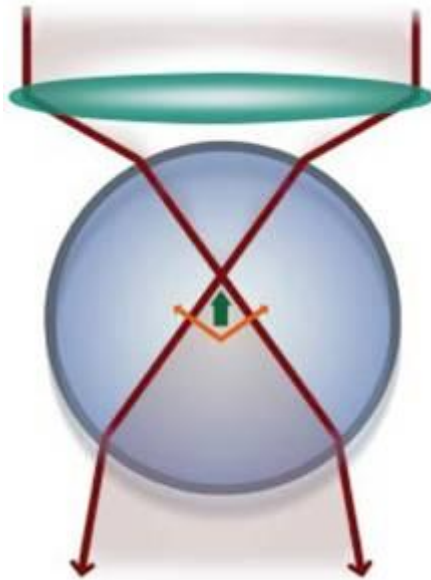


Рис.2.43 - Схематичне зображення оптичного пінцета.

Промінь лазера, що падає згори на лінзу, фокусується всередині краплі. При цьому на кожну частку, що знаходиться у воді, діють сили (помаранчеві стрілки), результуюча яких (зелена стрілка) завжди спрямована до фокусу.

У випадках, коли діаметр спійманої в пастку частки значно менше, ніж довжина хвилі світла, виконується умова розсіювання Релея, і частку можна розглянути як точковий диполь в неоднорідному електромагнітному полі. Сила, що діє на заряджену частку в електромагнітній області, відома як сила Лоренца:

$$\mathbf{F}_1 = q \left(\mathbf{E}_1 + \frac{d\mathbf{x}_1}{dt} \times \mathbf{B} \right).$$

Сила, що діє на диполь, обчислюється як сума сил, діючих на окремі заряди:

$$\mathbf{F} = \mathbf{F}_1 + \mathbf{F}_2,$$

$$\mathbf{F} = q \left(\mathbf{E}_1(x, y, z) - \mathbf{E}_2(x, y, z) + \frac{d(\mathbf{x}_1 - \mathbf{x}_2)}{dt} \times \mathbf{B} \right).$$

Через малість відстані між $\mathbf{d} = \mathbf{x}_1 - \mathbf{x}_2$ зарядами в диполі можна розкласти електричне поле поблизу першого заряду:

$$\mathbf{F} = q \left(\mathbf{E}_1(x, y, z) + (\mathbf{d} \cdot \nabla) \mathbf{E} - \mathbf{E}_1(x, y, z) + \frac{d\mathbf{d}}{dt} \times \mathbf{B} \right).$$

Відмітимо, що \mathbf{E}_1 скорочується. Розкриваємо дужки і замінюємо множення заряду q на відстань \mathbf{d} поляризацією диполя $\mathbf{P} = q\mathbf{d}$ і отримуємо:

$$\mathbf{F} = (\mathbf{p} \cdot \nabla) \mathbf{E} + \frac{d\mathbf{p}}{dt} \times \mathbf{B} = \alpha \left[(\mathbf{E} \cdot \nabla) \mathbf{E} + \frac{d\mathbf{E}}{dt} \times \mathbf{B} \right],$$

де у другому рівненстві припущено, що поляризація частки є лінійною функцією електричного поля, тобто

$$\mathbf{p} = \alpha \mathbf{E}.$$

Тепер, якщо використовувати два додаткових рівняння: рівняння з векторного аналізу і одне з рівнянь Максвелла

$$\begin{aligned} (\mathbf{E} \cdot \nabla) \mathbf{E} &= \nabla \left(\frac{1}{2} E^2 \right) - \mathbf{E} \times (\nabla \times \mathbf{E}) \\ \nabla \times \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t} \end{aligned}$$

отримаємо:

$$\mathbf{F} = \alpha \left[\frac{1}{2} \nabla E^2 + \frac{d}{dt} (\mathbf{E} \times \mathbf{B}) \right].$$

Другий додаток в останньому рівненстві - похідна за часом величини, яка пов'язана через постійний множник з вектором Пойнтінга, який описує потужність випромінювання, яке проходить через одиничну площадку.

Припускаючи, що потужність лазера не залежить від часу, похідна цього додатка - нуль, і сила запишеться у вигляді:

$$\mathbf{F} = \frac{1}{2} \alpha \nabla E^2.$$

Сила, з якою світло діє на навколишні об'єкти, невелика, але її виявляється достатньо, щоб ловити наночастки у фокус лазерного променя. Як тільки частка опинилася у фокусі, її можна рухати разом з лазерним променем. За допомогою оптичного пінцета можна пересувати частки розміром від 10 нм до 10 мкм і збирати з них різні структури (див. рис. 2.44). Є всі підстави вважати, що надалі лазерний пінцет стане одним з потужних інструментів нанотехнологій.

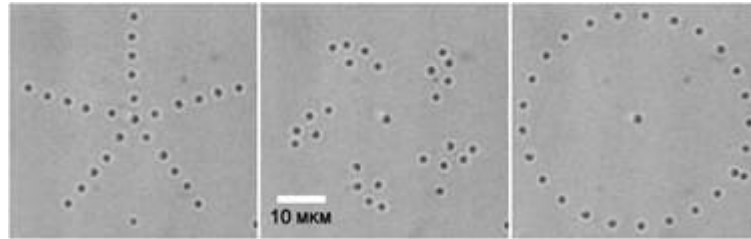


Рисунок 2.44 - Складені за допомогою лазерного пінцета різні візерунки з гелевих наночасток

Чому деякі частки, опинившись в лазерному промені, прагнуть в ту область, де інтенсивність світла максимальна, тобто у фокус (рис. 2.45) Для цього існують, принаймні, дві причини, наведені у вступі.

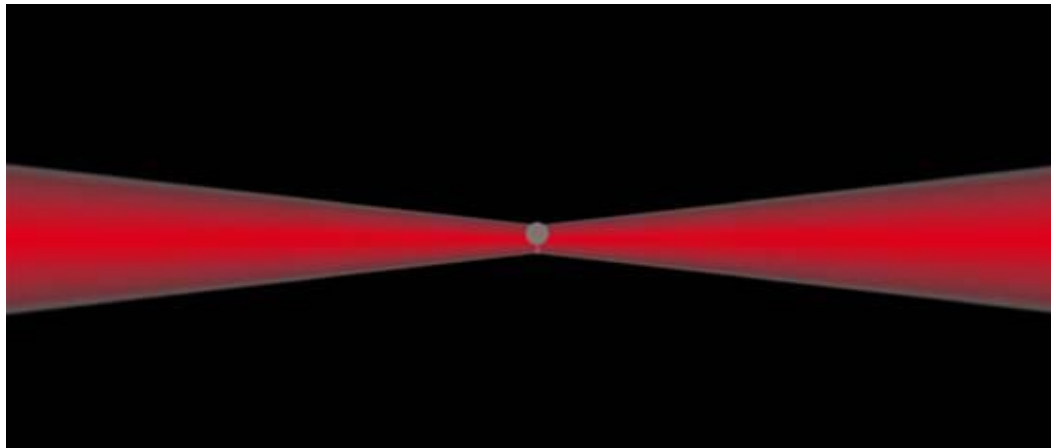
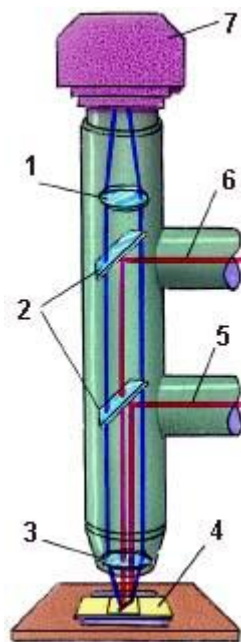


Рисунок 2.45 - Схематичне зображення червоного променя, що сходиться до фокусу і розходиться після нього

У місці фокусування променя видно сіру кулясту частку. Щоб оптичний пінцет використовував сили, описані вище в "причини I", необхідно, щоб

частка поляризувалася в зовнішньому електричному полі, і на її поверхні з'являлися пов'язані заряди. При цьому пов'язані заряди повинні створювати поле, спрямоване в протилежну сторону. Тільки в цьому випадку частки спрямуються до області фокуса. Якщо ж діелектрична постійна середовища, в якій плаває частка, більше діелектричної постійної речовини частки, то поляризація частки буде зворотною, і частка буде прагнути втекти з області фокуса. Так, наприклад, поводять себе повітряні бульбашки, які плавають в гліцерині.

Такі ж обмеження стосуються й "причини II". Якщо абсолютний показник заломлення матеріалу частки буде менше, ніж у середовища, в якому вона знаходиться, то частка буде відхиляти світло в інший бік, а значить, прагнути відійти подалі від осі променя. Прикладом можуть бути ті ж повітряні бульбашки в гліцерині. Тому оптичний пінцет працює краще, якщо відносний показник заломлення матеріалу частки більше.



1 і 3 - лінзи; 2 - дзеркала; 4 - досліджуваний зразок;
5 - освітлюючий пучок; 6 - захоплюючий пучок; 7 - відеокамера.

Рисунок 2.46 - Конструкція оптичного пінцета

2.7.5 Оптичний лазерний маніпулятор з фемтосекундними імпульсами

Фізика захвату залежить від відношення розміру частки і довжини хвилі випромінювання, за допомогою якого проводиться захват. Існує два граничних випадки:

- 1) частка багато менше довжини хвилі;
- 2) частка багато більше довжини хвилі.

У першому випадку вважають, що частка у фокусі об'єктива під дією лазерного випромінювання стає однорідно поляризованою і на наступних етапах розгляду представляється як точковий диполь. На таку частку діють сили обумовлені:

- розсіюванням ($F_{\text{расс}}$);
- поглинанням ($F_{\text{погл}}$);
- градієнтом інтенсивності випромінювання ($F_{\text{град}}$).

Сила $F_{\text{град}}$ пропорційна поляризованості і завжди спрямована уздовж градієнта інтенсивності до обмеженої дифракцією фокальної плями:

$$F_{\text{град}} = \frac{|\alpha|}{2} \nabla \langle E^2 \rangle .$$

Ця сила може перевершувати інші згадані сили, що призводить до ефекту «пастки». Частка потрапляє в потенційну яму, яку можна описати гармонійним потенціалом:

$$U = \frac{1}{2} k \langle (x - x_0)^2 \rangle$$

де k – характеристична константа захвата;

x_0 - центр «пастки».

Поведінка прозорих часток, розміри яких багато більше довжини хвилі падаючого випромінювання, можуть бути описані класичними теоріями заломлення і віддзеркалення.

Прикладом може стати розгляд прозорої кульки, геометричний центр якої зміщений відносно положення фокуса об'єктива (рис. 2.47). На рисунку пунктир - кінцеве положення кульки, суцільна лінія - вихідне зміщене положення. Два променя (a) і (b) відчують переломлення в кульці і відхиляються від початкового напрямку поширення. Таким чином, кулька змінює імпульс фотонів, асоційованих з променями (a) і (b).

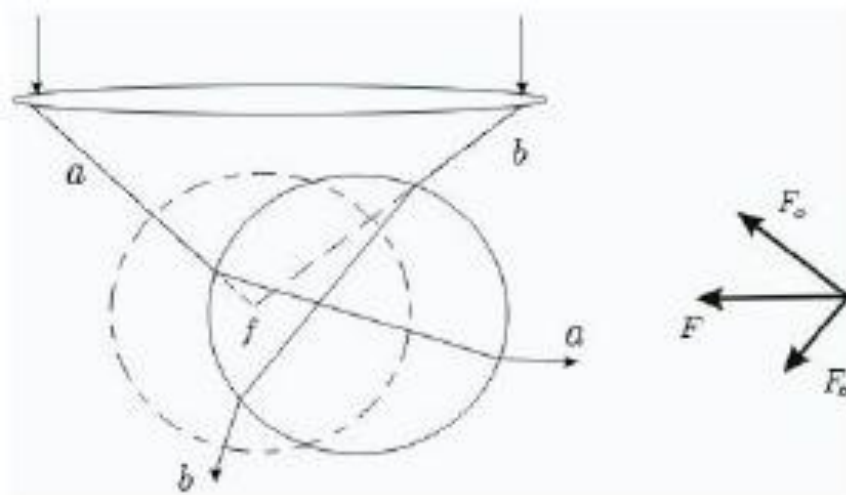


Рисунок 2.47 – Прозора кулька зі зміщеним геометричним центром відносно положення фокуса об'єктива

З рисунка виходить, що кулька «штовхає» промінь (a) праворуч і нагору. Отже, відхилені фотони надають імпульс кульці, спрямований у протилежний бік, тобто ліворуч і вниз. Аналогічно можна прийти до висновку, куди штовхає кулька промінь (b). Таким чином, результуюча сила, яка створена полем направляє кульку у бік фокуса. При досить сильному полі випромінювання частці важко змінити положення поблизу фокусу, і вона опиняється в «пастці».

2.7.6 Рух частки в потенційній ямі під дією світла

Завдання, яке поставлене в цьому розділі, відноситься до використання фемтосекундного лазера для оптичного захвата в якості джерела імпульсного випромінювання.

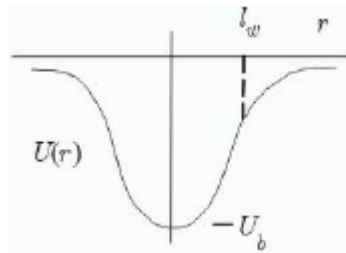


Рисунок 2.48 – Потенційна яма

У разі безперервного випромінювання поведінка захваченої частки в потенційній ямі $U(r)$ глибини U_b і ширини l_w . Потенціал обумовлено взаємодією частки з електричним полем, яке пропорційно квадрату електричного поля лазерного випромінювання та поляризуємості частки α

$$U(r) \approx -\alpha E^2(r) \approx -\alpha W(r).$$

Ширина і глибина ями визначаються довжиною хвилі випромінювання лазера λ , потужністю випромінювання W і параметрами об'єктива. У тому випадку, коли існує модуляція напруженості електричного поля випромінювання лазера, то розглянутий вище потенціал має тимчасову залежність U тому випадку, коли маємо справу з імпульсами випромінювання, тобто використовуємо наступні параметри: E_p - амплітуда імпульсів, τ_p - тривалість імпульсів, τ_d - час між імпульсами, при цьому $\tau_d \gg \tau_p$. У межі швидкої модуляції $E(r, t)$, коли за час «модуляції» частка практично не зміщується усередині ями $D\tau \ll l_w^2$, де $\tau = \tau_p + \tau_d$, завдання істотно спрощується.

У цій межі кінетика частки може бути описана за допомогою функції розподілу, усередненої по деякому періоду T_e , який задовольняє нерівностям $\tau \ll T_e \ll l_w^2 / D$:

$$\bar{\rho}(r, t) = T_e^{-1} \int_{t-T_e/2}^{t+T_e/2} dt_1 \rho(r, t_1).$$

Функція розподілу підпорядковується усередненому рівнянню Смолуховського. Диференціальне рівняння, що описує еволюцію розподілу ймовірностей для просторового положення броунівської частки має вигляд

$$\dot{\bar{\rho}} = D \nabla (\nabla \bar{\rho} + \bar{\rho} \nabla \bar{u}),$$

Опускаючи всі деталі математичного обґрунтування уявлення поведінки захваченої частки, як броунівської частки, що знаходиться в потенційній ямі, можна сформулювати. Для оптичного захвату броунівської частки необхідно використовувати імпульсний лазер з високою частотою повторення імпульсів, щоб за час між імпульсами частка не встигала значно зміститися з точки захвату. Захват фемтосекундними імпульсами можливий для широкого діапазону розмірів об'єктів.

2.7.7 Ефект руху потенційної ями

Кінетика руху броунівської частки в потенційній ямі, яка переміщається в просторі зі швидкістю $v_p = \dot{r}_p$ (тут $r_p(t)$ - положення точки мінімуму вихідної потенційної ями до моменту часу t) описується за допомогою рівняння Смолуховського, але із заміною потенціалу $\bar{U}(r)$ на $\bar{U}(r - r_p(t))$.

Тимчасова залежність потенційної ями моделює тимчасову нестійкість лазерного пучка, який захоплює частинку.

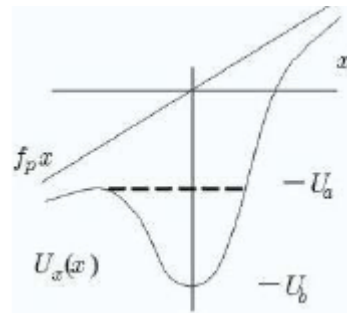


Рисунок 2.49 – Приклад здійснення оптичного захоплення

В одній з робіт виконаних співробітниками Московського фізико-технічного інституту спільно з фахівцями онкологами і клініцистами (Праці МФТІ, 2009, Т.1, №1, С. 53-58), представлені результати експериментального дослідження установки для оптичного захвату. Установка являє собою оптико-механічний електронний пристрій, який дозволяє здійснювати переміщення лейкоцита в полі лазерного випромінювання. Для забезпечення роботи установки використовуються серійні прилади та елементи:

- Мікроскоп Olympus IX71,
- Об'єктиви ЛОМО М-ФЛЮАР, Olympus LCАch N40x,
- МАІ ТАІ, 690-100 нм, 3 Вт,
- Sony ExwaveHAD камера.

Експерименти по маніпуляції окремими і клітинними агломератами. Захвату підлягали лейкоцити і ракові клітини. Для переміщення одиначної клітини розміром від 5 до 40 мкм використовувалася потужність 25 мВт.

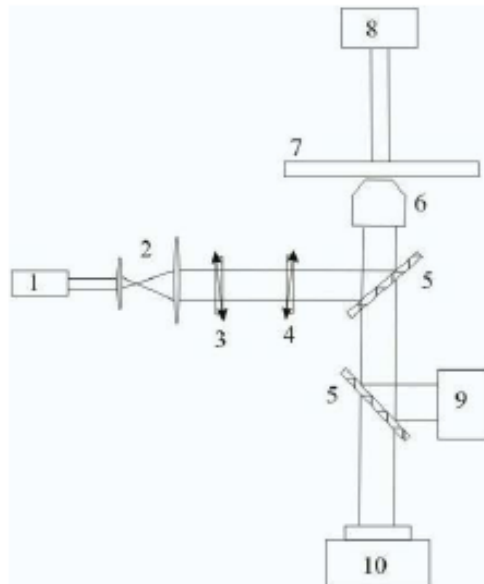


Рис. 4. Схема установки: 1 – лазер; 2 – линзы, расширяющие пучок; 3, 4 – поляризатор, анализатор; 5 – дихроическое зеркало; 6 – объектив; 7 – предметный столик; 8 – освещение; 9 – спектрометр/счётчик фотонов; 10 – камера

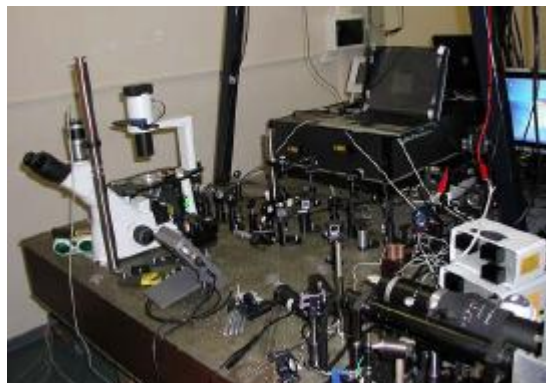


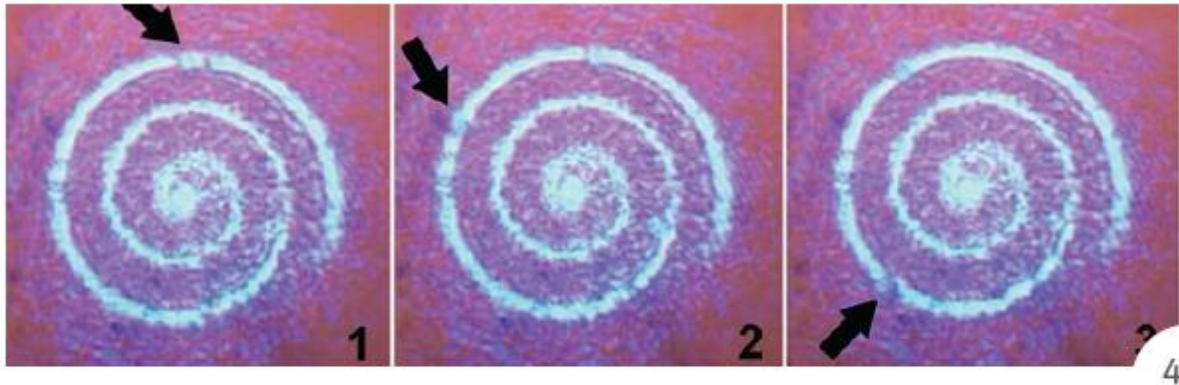
Рисунок 2.50 - Фемтосекундный лазер «пинцет-скальпель»

З часу винаходу першого лазерного пінцета, заснованого на одному гаусовому пучку (фундаментальна лазерна мода TEM_{00}) А. Ашкином в 1986 [58], концепція одномодових лазерних пучків розвинулася за рахунок використання лазерних мод високого порядку, тобто пучків ерміт-гаусових (TEM_{nm}) лагер-гаусових пучків (LG, TEM_{pl}) і беселевих пучків (J_n).

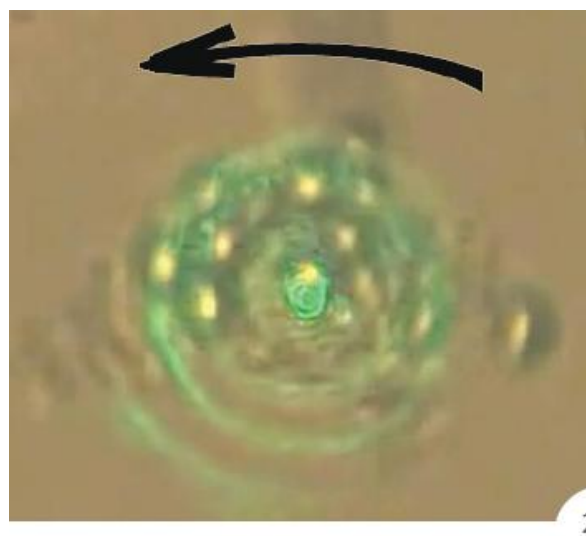
Лагер-гаусові пучки мають унікальну властивість втягувати в оптичну пастку оптично відображають і поглинають частки. Пучки з круговою поляризацією імеютспіновий орбітальний момент і можуть обертати частинки.

У Лагера-гауссових пучків також є власний кутовий момент, який може обертати частки навколо центру пучка. Цей ефект спостерігається без зовнішньої механічної або електричної регуляції променя. Крім Лагера-гауссових пучків, бesselевих пучки як нульового, так і вищих порядків мають орбітальний момент, а також унікальна властивість одночасно утримувати багато частинок на деякій відстані.

У цих роботах використовується інший спосіб застосування лазерного випромінювання пов'язаний з використанням складних в поперечному перерізі мод. Вони забезпечують ненульовий кутовий момент (спіральні пучків; світлових полів, отриманих при використанні тільки фазової частини від спірального пучка).



Перемещение частицы латекса размером 3,2 мкм вдоль спирали Архимеда, сформированной дифракционным элементом



Микромиксер на основе спирали Архимеда

Рисунок 2.51 – Реалізація результатів роботи

2.8 Виводи до розділу 2

Наукові результати, отримані в результаті виконання НДР, планується використати при підготовці монографій та навчальних посібників (з грифом Міністерства), докторських та кандидатських дисертацій, публікації у закордонних та фахових вітчизняних виданнях (за переліком ВАК України), а також статей у журналах, що входять до наукометричних баз даних, або мають імпакт-фактор, та отримання патентів України.

Таким чином, використання наукових результатів планується в подальшій науковій роботі, в навчальному процесі та в науково-дослідній роботі студентів. В цьому плані щорічно планується захист бакалаврських робіт, дипломних роботах спеціалістів та магістрів. Також отримані результати знайдуть застосування при підготовці нових та модернізації існуючих лабораторних робіт, а також будуть впроваджені в нових та існуючих навчальних курсах (таких, наприклад, як «Лазери в мікротехнологіях», «Взаємодія електромагнітного поля з середовищами» та «Використання лазерного випромінювання у оптоелектронній техніці та приладобудуванні») для їх удосконалення та приведення до рівня, який відповідає сучасним світовим вимогам.

3 ЗАСОБИ ОБРОБКИ ВІДЕОІНФОРМАЦІЇ В РЕАЛЬНОМУ ЧАСІ НА БАЗІ НЕЙРОННОЇ МЕРЕЖІ СМАС

3.1 Загальні відомості

Нейронні мережі використовуються для моделювання і/або контролю технологічного процесу з метою уникнення математичного моделювання, що пов'язане з такими його особливостями:

1. Математично орієнтоване моделювання вимагає великих витрат інженерної думки для створення опису процесу.

2. Існують процеси, де традиційні способи моделювання не дозволяють отримати результат, тоді як управління людиною таким процесом протікає нормально.

3. Математична модель процесу може бути настільки комплексною, або сам процес настільки швидким, що час обчислень стає критичним.

4. Математична модель прототипу може стати непридатною внаслідок старіння або зміни об'єкта.

Незважаючи на те, що можна використовувати найбільш відомі і популярні нейромережні рішення, такі як мережі з прямим зв'язком, що навчаються за алгоритмом зворотного розповсюдження помилки або його модифікаціями, дослідження показали, що швидкість обчислень і навчання таких мереж дуже мала для застосування в управлінні. Водночас багато дослідників намагалися застосувати спеціалізовані нейронні мережі для вирішення таких завдань.

Нейронна мережа СМАС була запропонована Дж. Альбусом (J. Albus). Абревіатура СМАС означає "Cerebellar Model Articulation Controller" (церебральна модель контролера артикуляції). Мережа так названа, оскільки була моделлю мозочка і використовувалася для управління рукою робота. Теоретичні переваги бачилися у високій збіжності і можливостях модифікації та удосконалення. Стверджувалося, що мережа СМАС забезпечує високу збіжність до будь-якої комбінації вхід-вихід, що є важливим у завданнях

управління і швидкого синтезу функцій. Крім того, прогнозувалася можливість апаратної реалізації мережі СМАС, і, відповідно, отримання високої швидкості навчання і роботи.

Структура простої класичної мережі СМАС наведена на рис. 3.1. Вона може розглядатися як асоціативна пам'ять, яка здійснює два послідовні відображення.

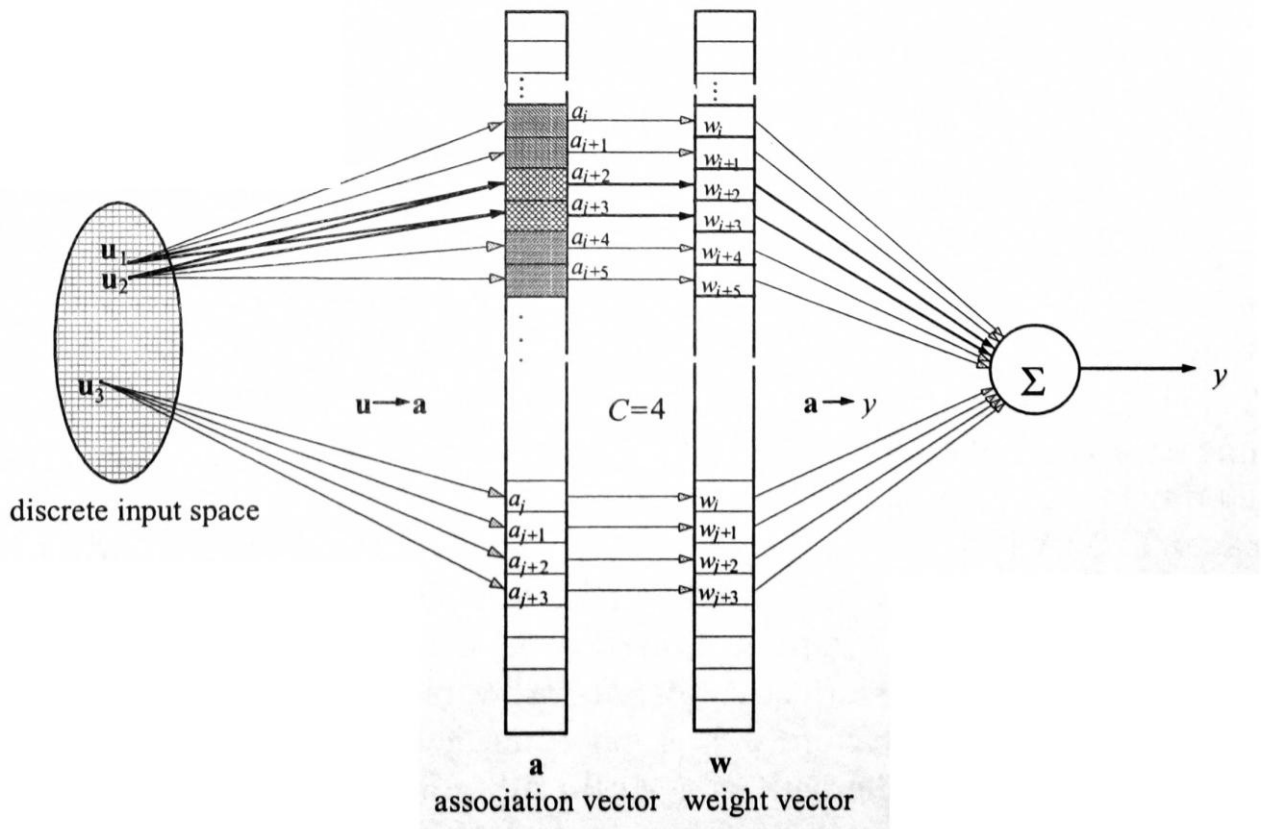


Рисунок 3.1 – Узагальнена структура нейронної мережі СМАС

Хай ϵ дискретний простір u вхідних сигналів u_j . Такий простір може утворюватися змінними при синтезі функцій або даними, що отримуються після аналого-цифрового перетворення. Основною особливістю цього простору є кінцевий набір значень. Перше відображення є нелінійним і полягає в тому, що кожному значенню u_j з простору u ставиться у відповідність унікальний двійковий вектор асоціацій a , що містить ρ одиничних біт. Решта біт вектора асоціацій, відповідного даному вхідному значенню, рівні нулю. Слід зазначити,

що значення ρ значно менше довжини вектора асоціацій. Цей параметр визначає узагальнювальні властивості мережі і часто називається узагальнювальним параметром. Кожному біту вектора асоціацій ставиться у відповідність двійкова базисна функція, яка рівна одиниці для ρ вхідних сигналів. З точки зору математики вона може задаватися так:

$$f_n(u) = \begin{cases} 1, & u \in [u_k, u_{k+1}, u_{k+2}, \dots, u_{k+\rho-1}] \\ 0, & u \notin [u_k, u_{k+1}, u_{k+2}, \dots, u_{k+\rho-1}] \end{cases} \quad (3.1)$$

У наведеній залежності u_k і u_{k+1} відрізняються на величину кроку квантування вхідних сигналів, який при двійковому поданні сигналів рівний одиниці молодшого розряду. Набір значень вхідних сигналів, на яких базисна функція рівна 1, називається її рецептивним полем.

Відображення вхідного простору у вектор асоціацій має такі властивості:

1. Два сусідні вхідні значення відображуються у вектори асоціацій, що відрізняються лише декількома елементами.

2. Із зростанням кодової відстані між вхідними значеннями кількість загальних активних бітів у векторі асоціацій зменшується. Якщо це відстань більша ρ , вектори асоціацій не повинні мати загальних одиничних бітів.

Друге відображення обчислює вихід нейронної мережі як скалярний твір вектора асоціацій a на вектор ваг w . З урахуванням того, що вектор асоціацій є двійковим,

$$y(u) = \sum_{i:a_i(u)=1} w_i \quad (3.2)$$

Так, проста нейронна мережа СМАС для своєї реалізації не вимагає виконання ресурсомістких і таких, що вимагають часових витрат, операцій добутку, що робить привабливою її апаратну реалізацію.

Далі детальніше розглянемо наведену структуру, покажемо можливість вибору значень базисних функцій, відмінних від одиничних, вкажемо шляхи скорочення довжини вектора вагових коефіцієнтів, розглянемо навчання мережі. У завершенні розділу будуть розглянуті можливості застосування нейронної мережі СМАС для обробки зображень в реальному часі.

3.2 Етапи обробки інформації

3.2.1 Кодування інформації

Різде зменшення об'єму пам'яті, необхідної для зберігання оброблюваної інформації, є особливістю мережі СМАС. Ця особливість реалізується за рахунок застосування в мережі спеціального кодування.

Кодування інформації здійснюється в шарах L1 і L2 і полягає в тому, що кожному N-вимірному вхідному вектору $x(i)$ ставиться у відповідність n-вимірний вектор асоціацій $a(i)$. При цьому лише $\rho \ll n$ елементів даного вектора мають відмінні від нуля значення. Іншими словами, для обчислення функції, відповідної вхідному вектору $x(i)$, необхідно прочитати вміст ρ елементів пам'яті, що зберігає значення ваг, що настроюються. Таку ж кількість ваг необхідно обчислити і записати в пам'ять при навчанні мережі на вхідному векторі $x(i)$.

Кодування здійснюється в два етапи. На першому етапі відбувається дискретизація безперервної безлічі вхідних сигналів (квантування за рівнем), після чого отримана дискретна множина аналізується нейронами вхідного шару на предмет входження в область визначення їх функцій активації. При цьому на кожному ступені квантування активізується один нейрон для кожного елемента вектора $x(i)$. Якщо вхідні сигнали дискретні за своєю природою (наприклад, координати точок, яскравість яких обчислюється мережею), операція дискретизації втрачає необхідність. На другому етапі кодування асоціативні нейрони другого шару аналізують, які нейрони першого шару активні і

формують вектор асоціацій. Схема кодування інформації однозначно визначається розподілом нейронів вхідного шару за ступенями квантування.

Розглянемо причину скорочення об'єму інформації. За відсутності кодування на вхідному шарі кожному рівню квантування однієї з вхідних змінних необхідно зіставити всі можливі значення інших змінних. Для кожного такого зіставлення необхідно зберігати значення функції. Кількість значень рівна добутку кількостей рівнів квантування вхідних змінних. Під час використання кодування зіставляються не рівні квантування, а області. При цьому на кожній області на j -му ступені кодування однієї змінної зіставляються всі області квантування інших змінних, що належать лише j -му ступеню. Кількість областей квантування на ступені визначається відношенням кількості рівнів квантування R змінної до рівнів квантування, що об'єднуються областю. Ступінь може містити до двох неповних областей на краях, поява яких пов'язана із зміщенням розподілу областей на сусідніх ступенях. Ці міркування пояснюють суть залежності, що визначає максимальну кількість зіставлень при даних параметрах вхідних змінних і самої мережі.

Отже, значення ваги, що зберігається в даному елементі асоціативної пам'яті мережі СМАС, використовується при різних комбінаціях вхідних сигналів (близьких в плані відстані Хеммінга), тоді як за відсутності кодування один елемент пам'яті відповідає єдиній комбінації вхідних змінних. Таке кодування припускає, що близькі (у плані відстані Хеммінга) вхідні змінні мають корельовані відповідні значення вихідних функцій. Це твердження справедливе для яскравостей і кольорів сусідніх точок на повнокольорових зображеннях, для функцій присутності, використовуваних при мікшуванні зображень, для функцій колірної корекції. В результаті цього мережу СМАС можна запропонувати для стиснення зображень, для синтезу функцій присутності і колірної корекції.

Для оцінки скорочення об'єму необхідної пам'яті для мережі СМАС у порівнянні з табличним методом розглянемо приклад на рис. 3.2. Для табличного способу кодування необхідне $R^2 = 19^2 = 361$ елементів пам'яті. Водночас для мережі СМАС

– максимум 106 комірок, тобто більш ніж в три рази менше. Для випадку 3-х вхідних змінних, що мають розрядність 8 біт ($R = 256$), при $\rho = 8$ економія складе:

$$\left\lceil \frac{256^3}{8 \cdot \left(\frac{256-1}{8} + 1 \right)^3} \right\rceil = 59 \text{ раз.}$$

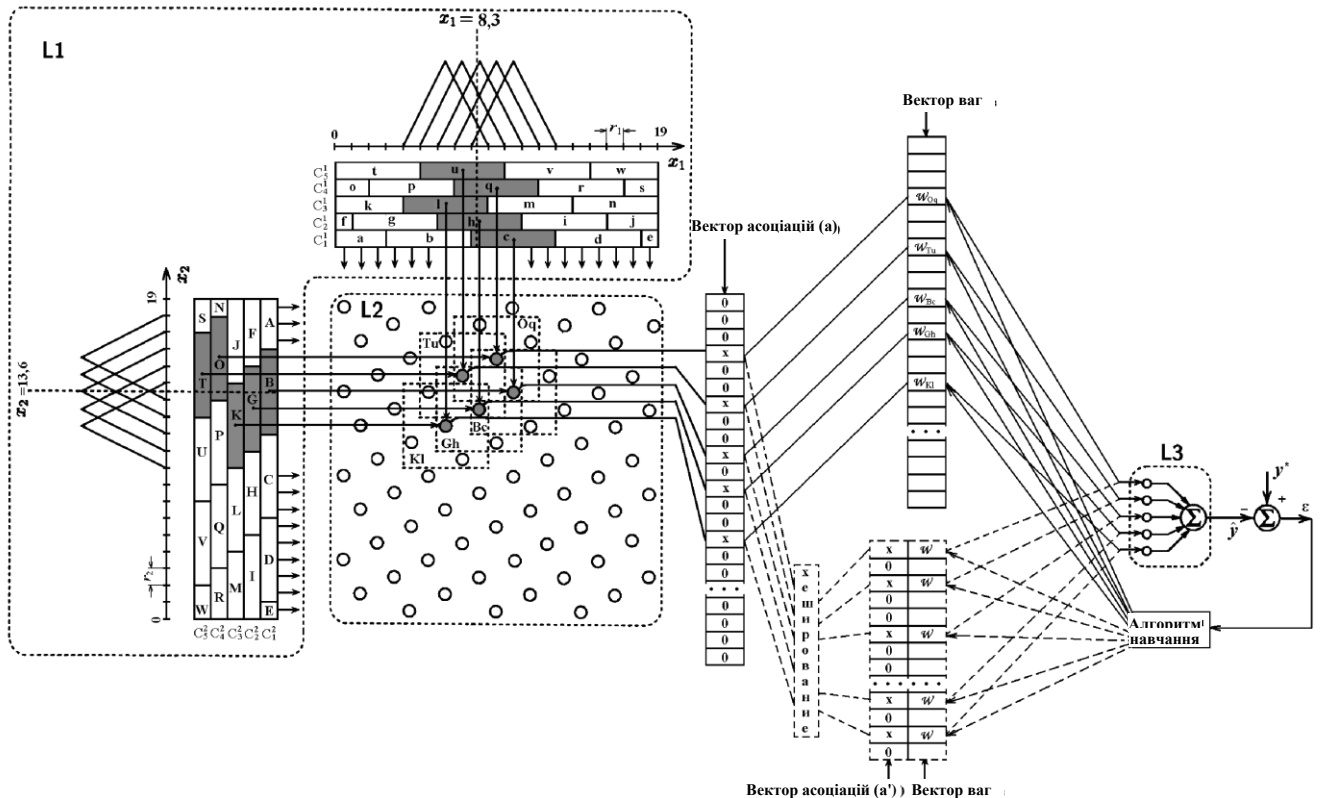


Рисунок 3.2 – Структура нейронної мережі СМАС

Останній приклад показує економію пам'яті під час використання для колірної корекції зображення, наведеного в колориметричній системі RGB, мережі СМАС у порівнянні з табличним методом.

Подальше скорочення об'єму пам'яті можливе під час використання хешування вектора асоціацій.

Зрушення областей на ступенях квантування обумовлюється тим, що лише $(\rho - 1)$ комірок вектора асоціацій можуть співпадати при кодуванні різних

вхідних сигналів. На сусідніх ступенях області квантування можуть зміщуватися як на один рівень, так і на декілька. Величини зрушень для кожного ступеня і змінної визначають схему кодування у вхідному шарі.

3.2.2 Вибір базисних функцій

Як уже зазначалося, класична мережа СМАС використовує нейрони вхідного шару з прямокутною функцією активації і, відповідно, виконує кусково-лінійну апроксимацію. Проте, як показують дослідження, на апроксимуючі властивості мережі істотно впливає вибір цих функцій. На рис. 3.3 наведено різні види базисних функцій нейронів першого шару.

При виборі базисних функцій прямокутної форми (рис. 3.3,а) обчислювальні витрати будуть мінімальними. Це обумовлено тим, що у вхідному шарі замість обчислень значень базисних функцій для кожного вхідного сигналу достатньо визначити, які нейрони збуджені, і подати на їхні виходи сигнал, що дорівнює 1. У шарі асоціативних нейронів також не відбувається обчислень, а визначаються збуджені нейрони і з їх виходів знімається сигнал “1”. Внаслідок цього значно скорочується час реакції мережі на вхідний сигнал, що поступив. Компоненти вектора асоціацій в цьому випадку можуть приймати значення “0” або “1”. Швидкість навчання мережі при виборі базисних функцій прямокутної форми буде максимальною. Істотним також є той факт, що в мережі СМАС з базисними функціями прямокутної форми схема кодування інформації не впливає на точність апроксимації, тобто будь-яка схема кодування є прийнятною.

На рис. 3.3 показано активаційні функції збуджених нейронів s , h , l , q , u (див. приклад, наведений на рис. 3.2), що беруть участь у кодуванні координати $x = (0,83 \ 1,36)^T$, подані у вигляді В-сплайнів першого (рис. 3.3,а), другого (рис. 3.3,б) і четвертого (рис. 3.3,в) порядків. Безперечною перевагою В-

сплайнів є можливість рекурентного обчислення як самих В-сплайнів, так і їх похідних δ -го порядку.

На рис. 3.3,г показано активаційні функції Гауса цих же збуджених нейронів, що мають вигляд:

$$\Phi_i(x_j) = \exp\left\{-\frac{(x_j - \mu_i)^2}{\sigma^2}\right\}. \quad (3.3)$$

Легко побачити, що вони також допускають достатньо просте обчислення похідних $\frac{\partial \Phi_i}{\partial x_j}; \frac{\partial^2 \Phi_i}{\partial x_j^2} \dots$.

Проте, хоча функція Гауса і має властивість локального збудження, достатньо складно чітко виділити її межі збудження, що є важливим при реалізації кодування інформації в СМАС. Для усунення цього недоліку можна використати модифіковану функцію Гауса, яка має вигляд:

$$\Phi_i(x) = \begin{cases} \exp\left\{-\frac{(\lambda_2 - \lambda_1)^2 / 4}{(x - \lambda_1)(\lambda_2 - x)}\right\} & \text{при } x \in (\lambda_1, \lambda_2); \\ 0 & \text{інакше.} \end{cases} \quad (3.4)$$

Як видно з (3.4), дана функція точно визначена в інтервалі (λ_1, λ_2) , що спрощує процедуру масштабування базисних функцій при зміні таких параметрів мережі, як R і ρ . На рис. 3.3, д показано модифіковані активаційні функції Гауса збуджених нейронів с, h, l, q, u.

Аналогічний вигляд матимуть і базисні функції збуджених нейронів Т, О, К, G, В, що беруть участь в кодуванні другої координати x_2 (у нашому прикладі $x_2 = 1,36$). Оскільки значення компонент вектора асоціацій, що з'являються на виході L2, отримується перемножуванням відповідних значень базисних функцій збуджених нейронів.

Як видно з наведених рисунків, дані базисні функції мають поблизу межі області квантування (зони чутливості нейрона вхідного шару) порівняно

невеликі значення, що негативно позначається на результатах апроксимації. Для уникнення вказаного недоліку як базисні можливе використання тригонометричних функцій, наприклад, косинусоїдальної, що має вигляд:

$$\phi_i(x_j) = \begin{cases} \cos\left(\frac{\pi}{\rho r_j}(x_j - \lambda_i)\right) & \text{якщо } x_j \in (\lambda_i - \frac{\rho r_j}{2}, \lambda_i + \frac{\rho r_j}{2}); \\ 0 & \text{інакше;} \end{cases} \quad (3.5)$$

де λ_i – i -й центр області квантування; r_j – крок квантування після j -ї компоненти вхідного сигналу.

Вид даної функції наведено на рис. 3.3, е. Тригонометричні функції також дозволяють легко обчислювати їх похідні. З рисунка видно, що мінімальні значення базисних функцій збільшилися.

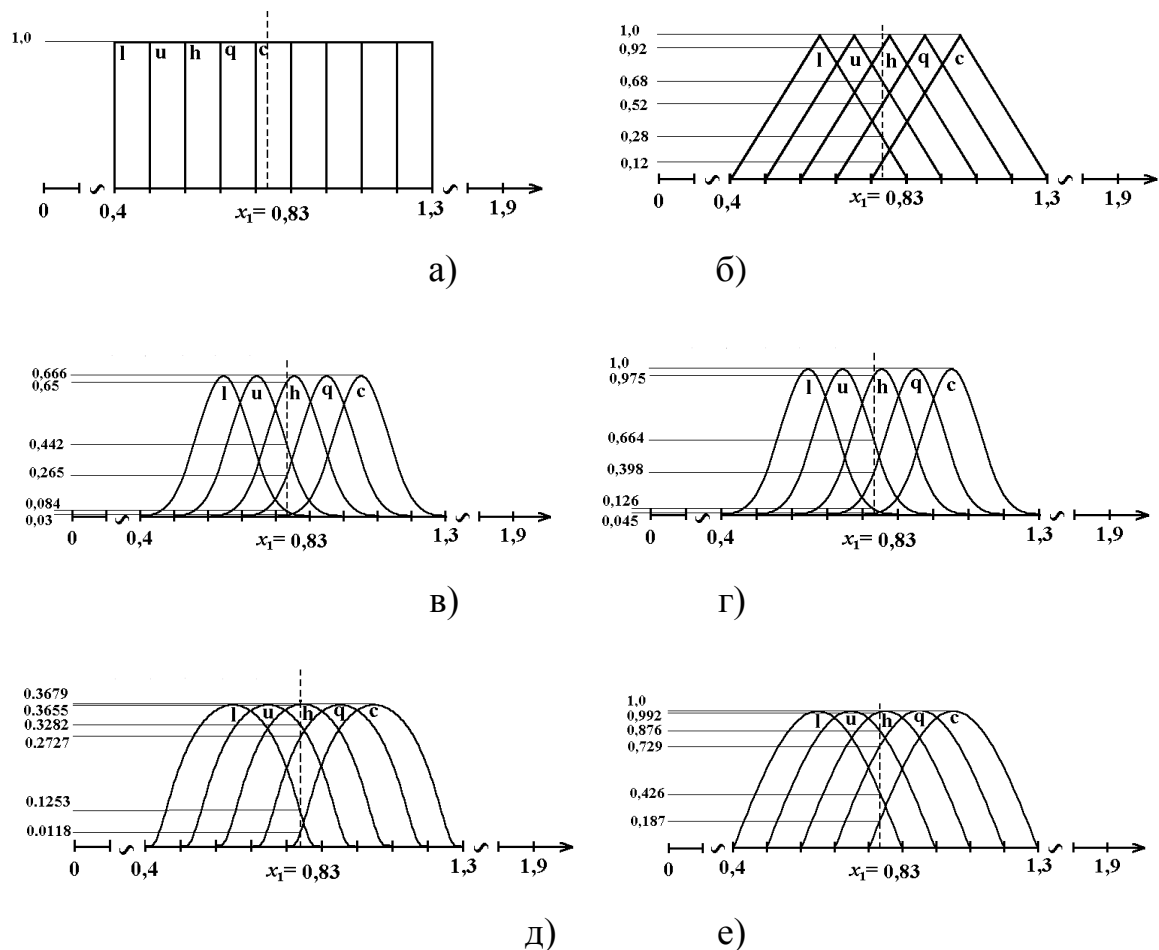


Рисунок 3.3 – Різні види базисних функцій

Як було зазначено вище, ненульові компоненти вектора асоціації утворюються шляхом перемножування відповідних значень базисних функцій збуджених нейронів шару L1. На рис. 3.4 показано, як здійснюється добуток відповідних базисних функцій, і який вигляд має функція, що є результатом такого добутку.

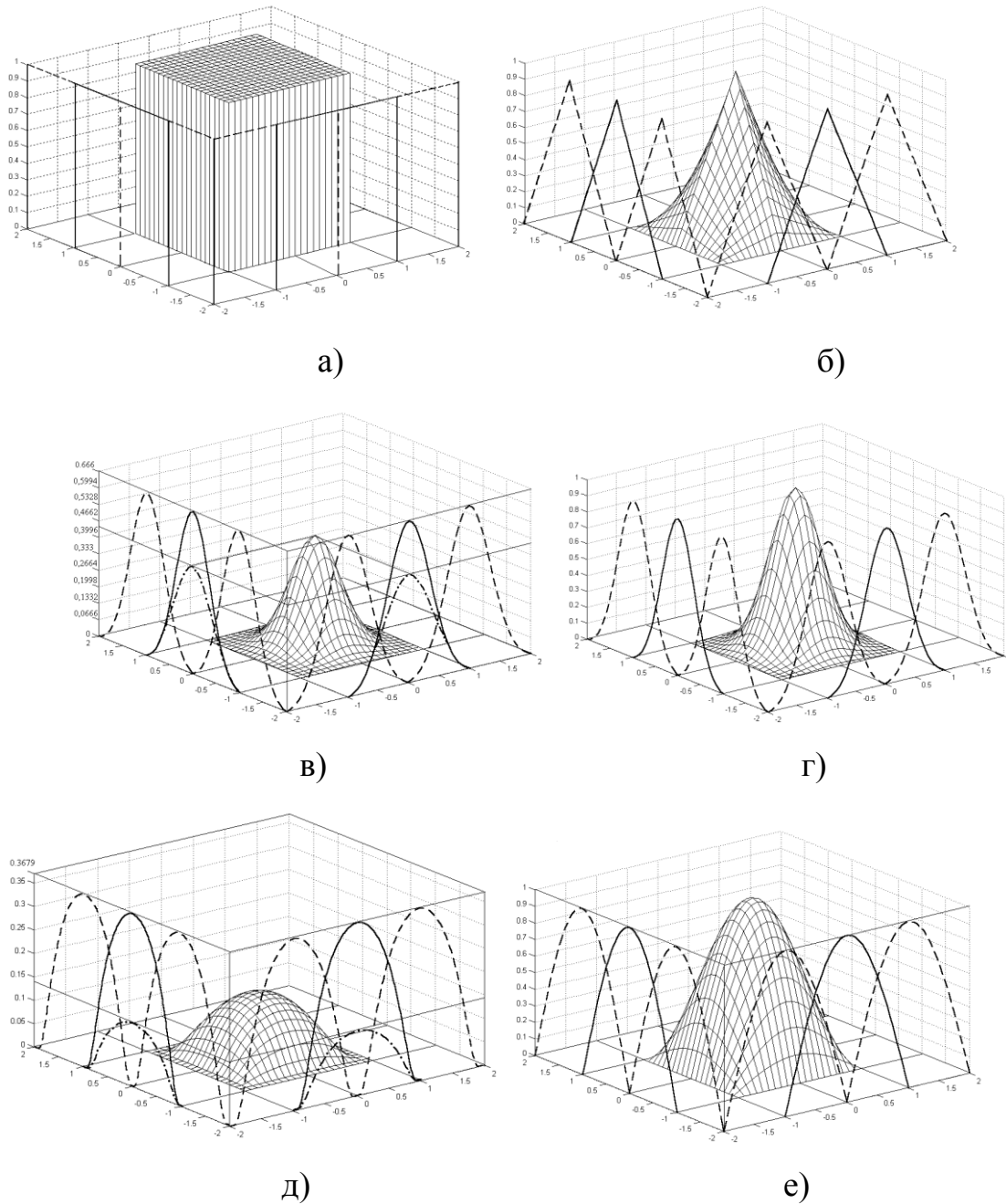


Рисунок 3.4 – Результати перемножування базисних функцій різного вигляду

З рис. 3.4 видно, що при виборі базисних функцій вигляду, відмінного від прямокутного, ці функції визначають, в яку зону області, утвореної базисними функціями нейронів першого шару, виходи яких поступають на один і той самий асоціативний нейрон другого шару, входить даний вхідний вектор. Чим ближче до краю області даний вхідний сигнал, тим менше значення має добуток базисних функцій, і тим менший вплив надаватиме вага, відповідно даному асоціативному нейрону на вихідний сигнал мережі. Під час навчання мережі ця вага буде найменш змінена.

3.2.3 Хешування інформації

Для скорочення об'єму пам'яті, необхідного для зберігання ваг, можливе застосування хешування інформації, яке полягає у відображенні більшої, необхіднішої пам'яті в меншу, яка простіше реалізується фізично і швидше ініціалізується. Найпростішим прикладом хешування інформації є обчислення залишку від ділення великого числа на деяке відоме, відносно коротке число. В цьому випадку, яким би не було перше число, результат його хешування буде не більш, ніж модуль (число, на яке проводиться ділення). Очевидно, що в результаті хешування може виникнути ситуація, коли різні початкові дані призводять до одного результату – виникає так звана хеш-колізія, що носить імовірнісний характер. Вірогідність хеш-колізії залежить від співвідношення довжин інформаційного поля і модуля.

У мережі СМАС початковою для операції хешування є адреса ваги в пам'яті, що отримується з вектора асоціацій і що лежить в інтервалі $[0, M-1]$. Результат – адреса ваги у фізичній пам'яті, що лежить в інтервалі $[0, m-1]$, де $m < M$. Функція перетворення, або хеш-функція, має відповідати вимозі

$$0 \leq H(k) \leq m - 1, \forall k : 0 \leq k \leq M - 1.$$

Вибір хеш-функції $H(k)$ є дуже трудомістким завданням, проте численні дослідження показали, що хороші результати забезпечує застосування двох основних типів хеш-функцій, один з яких заснований на діленні, а інший – на множенні.

Перший випадок відповідає розглянутому прикладу хешування і аналітично може бути записаний у вигляді

$$H(k) = k \bmod m, \quad (3.6)$$

де функція \bmod означає ділення за модулем, відповідне обчисленню залишку від цілочисельного ділення k на m .

Така функція рівномірно розподіляє велику пам'ять в меншу. У другому випадку хеш-функція обчислюється за формулою

$$H(k) = m \left[\text{frac} \left(\frac{F}{w} k \right) \right], \quad (3.7)$$

де функція frac означає узяття дробової частини виразу, F – деяка ціла константа, взаємно проста з розміром машинного слова w , $\lceil \cdot \rceil$ означає округлення у бік найближчого більшого цілого числа.

Як F можна використовувати числа Фібоначчі. Хеш-функція, заснована на множенні, «тасує» елементи пам'яті.

Розглянемо вплив застосування алгоритмів хешування на результат апроксимації на прикладі функції «Виключне АБО». На рис. 3.5, а наведений результат реалізації даної функції за допомогою класичної мережі СМАС із такими параметрами: $R = 50$; $\rho = 20$; $n = 236$. На рис. 3.5, б–д наведено реалізацію функції «Виключне АБО» із застосуванням різних алгоритмів хешування.

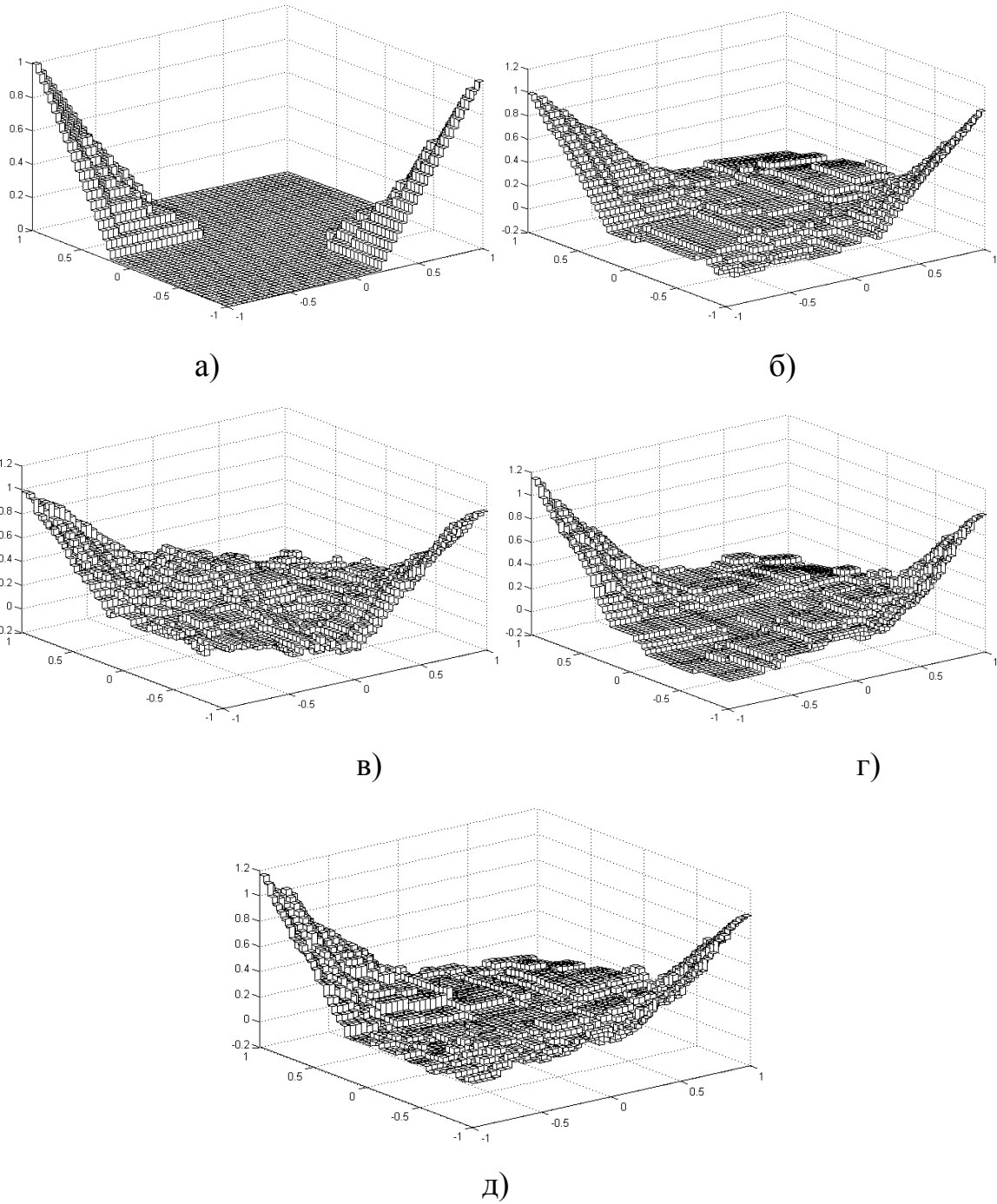


Рисунок 3.5 – Реалізація функції «Виключне АБО» з використанням різних алгоритмів хешування

3.2.4 Хеш-колізії

Який би метод хешування не застосовувався, неможливо повністю усунути проблему хеш-колізій. При хешуванні інформації в мережі СМАС можливе виникнення двох типів колізій. Колізії першого типу виникають при

активації одним вхідним сигналом однієї комірки пам'яті більше одного разу. Імовірність того, що один вхідний сигнал звернеться два й більше рази до однієї комірки фізичної пам'яті, можна визначити за формулою

$$P = \sum_{i=2}^{\rho-1} \frac{i}{m}, \quad (3.8)$$

де m – об'єм меншої пам'яті.

Як уже зазначалося, незалежно від використовуваного методу хешування, повністю усунути хеш-колізії неможливо. Колізії першого типу виникають при активації одним вхідним сигналом одного елемента пам'яті більше одного разу. Вірогідність того, що один вхідний сигнал звернеться два і більше разів до одного елемента фізичної пам'яті, можна визначити за формулою

$$P = \sum_{i=2}^{\rho-1} \frac{i}{m}. \quad (3.9)$$

Причиною виникнення колізій другого типу є активація одного й того ж елемента пам'яті різними далеко розташованими один від одного (у значенні відстані Хеммінга) вхідними сигналами. Ці колізії призводять до серйознішої проблеми – появи небажаної кореляції між двома далеко розташованими (у значенні відстані Хеммінга) вхідними сигналами. Ці колізії не можуть істотно вплинути на роботу мережі, поки виконується умова

$$a'_1 \wedge a'_2 \ll \rho, \quad (3.10)$$

де a'_1 і a'_2 – перетворені хеш-функцією вектори асоціацій двох різних вхідних сигналів.

Виникнення хеш-колізій призводить до того, що на одну і ту ж вагу впливають далекі в значенні відстані Хеммінга сигнали. Для таких сигналів значення обчислюваної функції слабко корельовані або не корельовані взагалі. Проте хеш-колізія внесе кореляцію і додасть шум до вихідного сигналу мережі.

У плані апаратної реалізації найбільш очевидний і простий алгоритм хешування (3.6). Оскільки використовувана пам'ять має об'єм рівний, як правило, степені числа 2, операція ділення за модулем зводиться до взяття як результату обчислення g – молодші розряди початкової адреси ваги ($m = 2^g$).

3.3 Методи навчання

Навчання мережі СМАС, як практично і всіх інших ШНС, полягає в настроюванні вектора її вагових параметрів w розмірності $n \times 1$, тобто в даній мережі настроюються лише ваги зв'язків між асоціативними нейронами і нейронами вихідного шару. Використання спеціального кодування дозволяє зберігати в ρ елементах пам'яті ваги мережі. Якби різні вхідні вектори при кодуванні не мали загальних комірок, то ваги мережі могли визначатися і без її навчання, аналітично, так:

$$w_{ik} = \frac{y(k)}{\rho}. \quad (1.10)$$

Наявність загальних областей пам'яті для зберігання різних векторів не дозволяє обмежитися таким простим обчисленням ваг і обумовлює необхідність навчання мережі. Компоненти вектора ваг мають фізичні адреси і вибираються так, щоб середнє арифметичне значення активних ваг, тобто ваг, відповідних одиничним компонентам вектора асоціацій, було рівне значенню функції, що запам'ятовувалася, $y(i)$.

Наявність загальних елементів пам'яті для різних векторів призводить до колізій, тобто явища, при якому мережа обчислює для двох різних вхідних векторів однакові адреси.

3.3.1 Алгоритм навчання Альбуса

Під час навчання мережі подають навчальні $(x(i), y^*(i))$, $i = \overline{1, M}$, де $y^*(i)$ – необхідне значення функції, і коректуються значення вагових параметрів так, щоб із зростанням кількості подавань ця послідовність ваг, що настроюються, прагнула до стаціонарного значення w^* .

Вихідний сигнал мережі на k -му такті навчання можна записати так:

$$y(k) = wT(k)a(k), \quad (3.11)$$

де $w(k) = (w_1(k), w_2(k), \dots, w_n(k))^T$ – вектор оцінок вагових параметрів.

Як вже наголошувалося, ваги $w(k)$ вибираються так, щоб їх середнє арифметичне було рівне значенню функції, що запам'ятовувалася. А оскільки на кожному такті навчання (при поданні вхідного вектора) тільки ρ компонент вектора асоціацій a відмінні від нуля (активними виявляються ρ елементів пам'яті $\rho \ll n$), то вираз (3.11) може подаватися так:

$$y(k) = \frac{1}{\rho} \sum_{i=1}^n w_i(k), \quad (3.12)$$

де підсумовування відбувається по всіх i , для яких $a_i(k) = 1$.

З урахуванням цього запропоновано рекурентний алгоритм навчання СМАС, згідно з яким при поданні чергового вхідного сигналу коректуються лише ті ρ ваг мережі, які відповідають одиничним компонентам вектора асоціацій для даного сигналу. Аналітично алгоритм навчання Альбуса можна записати у такій формі:

$$w_i(k) = \begin{cases} w_i(k-1) + \left(y(k) - \sum_i w_i(k-1) \rho^{-1} \right) \\ \forall i, a_i(k) = 1, a_i(k-1) = 1; \forall j, a_j(k) = 1, \\ w_i(k) = w_i(k-1) \quad \forall i, a_i(k) = 1. \end{cases} \quad (3.13)$$

Отже, корекція вагового вектора $w(k)$, обчисленого за k вимірювань, при набутті $k+1$ -го значення вхідного вектора $x(k+1)$ і відповідного значення функції $y^*(k+1)$ відбувається за таким правилом: за вхідним вектором $x(k)$ шляхом кодування визначається ρ -вимірний вектор асоціацій $a(k)$, визначаються відповідні активні асоціативні елементи вектора ваг $w(k)$; із значення функції $y^*(k)$, що запам'ятовується, віднімається середнє арифметичне значення активних ваг вектора $a(k-1)$ (оскільки всі ваги при виборі прямокутних активаційних функцій є рівноправними), а отримана різниця додається до значень активних ваг; неактивні ж значення вагового вектора не коректуються.

Метою алгоритму Альбуса є визначення вектора w , який прагне до постійного вектора w^* , такого, як

$$\hat{a}_i w^* = \sqrt{\rho} y_i^* \quad (3.14)$$

для кожного i ($i=1, 2, \dots, N$). Оскільки вектор w^* містить ρ елементів, (3.14) подає N рівнянь з ρ невідомими й існування розв'язання (або розв'язань) цієї системи визначається відомими результатами з лінійної алгебри. Запишемо (3.14) в матричній формі

$$Aw^* = h, \quad (3.15)$$

де i -й рядок матриці A розміру $N \times \rho$ складається з елементів ρ -вимірного вектора $\hat{a}_{(i)}^T$ (єдинична нормаль до гіперплощини, вибрана на i -му кроці), а $(h-N)$ -вимірний вектор-стовпець, i -й елемент якого рівний $\sqrt{\rho} y^*(i)$. Щоб система рівнянь (3.14) була сумісна, матриця A і розширена матриця $[A, h]$ повинні мати однаковий ранг, рівний, наприклад R . Для існування єдиної точки x^* потрібно, щоб $N \geq \rho$ і, більш того, ранг R має досягати свого найбільшого значення, тобто $R = \rho$. Якщо $N > \rho$, то $N - \rho$ надмірних рівнянь можуть виключатися з системи, то ρ рівнянь, що залишилися, записуються у

вигляді $A^1 w^* = h^1$, де A^1 — матриця розміру $p \times p$, ранг якої рівний p , а h^1 — p -вимірний вектор-стовпець, який дозволяє при відомій зворотній матриці A^1 однозначно визначити вектор $w^* = (A^1)^{-1} h^1$.

Існують й інші можливості під час розв'язання системи рівнянь, наприклад у тому випадку, коли система рівнянь (3.15) несумісна. Ці можливі ситуації наведені в табл. 3.1, де вказаний їх вплив на поведінку алгоритму Альбуса (3.13). Наведені в табл. 3.1 випадки 3, 5 і 7, коли система рівнянь несумісна, можуть зустрічатися на практиці.

Зазначимо, що розв'язання системи рівнянь (3.15) має вигляд

$$w^* = A^+ h^1, \quad (3.16)$$

де A^+ — псевдо звернена матриця за Муром-Пенроузом, визначена виразами

$$A^+ = \begin{cases} (A^T A)^{-1} A^T & \text{при } N \geq \rho; \\ A^T (A A^T)^{-1} & \text{при } N < \rho. \end{cases} \quad (5.17)$$

$$(5.18)$$

У тому випадку, коли система рівнянь несумісна (випадки 5 і 7), це розв'язання, крім того, має таку бажану властивість, як розв'язання з мінімальною нормою. Основною проблемою, що виникає під час реалізації даного розв'язання, є істотна обчислювальна складність, обумовлена великою розмірністю матриць A $N \times p$, які на практиці складаються з тисяч рядків і стовпців, так що обчислення відповідно до формул (3.17), (3.18) виявляються фактично неможливими. Тому в обчислювальному відношенні алгоритм Альбуса, що використовує тільки обчислення скалярного твору два p -вимірних векторів, причому один з цих векторів \hat{a}_i^T містить всього ρ ($\rho \ll p$) відмінних від нуля елементів, має величезну перевагу.

Таблиця 3.1 – Різні варіанти розв'язання рівняння $Aw^* = h$ (3.15) і їх вплив на алгоритм Альбуса

Випадок	Співвідношення між розмірами p и N	Ранг матриці	Ранг розширеної матриці $[A, h]$	Перетин гіперплощин, визначених (1.14)	Властивості збіжності алгоритму	
					циклічне навчання	випадкове навчання
1	$p \geq N$	N (максимально-можливий)	N	Перетином є гіперплощина розмірності $p - N$	Збіжність до єдиної точки, якщо початкова точка x_0 лінійно залежить від \hat{a}_i , інакше – до точки, яка залежить від x_0	
2	$p \geq N$	$R < N$	$R < N$	Перетином є гіперплощина розмірності $p - R$	Як і у випадку 1	
3	$p \geq N$	$R < N$	$R + 1 < N$	Рівняння (1.14) несумісні. Не існує єдиної гіперплощини перетину	Граничний цикл існує, його форма залежить від порядку циклічного навчання, але не залежить від x_0 . Розташування не залежить від x_0 , якщо x_0 лінійно залежить від \hat{a}_i	Точки, що породжуються алгоритмом (1.15), лежать в мінімальній зоні захоплення з вірогідністю 1

4	$N \geq p$	p (максимально -можливий)	p	Перетином ϵ гіперплощина розмірності 0, тобто єдина точка	Збіжність в цю єдину точку з будь-яких початкових умов x_0	
5	$N \geq p$	p (максимально -можливий)	$p + 1$	Рівняння (1.14) несумісні. Не існує єдиної гіперплощини перетину	Граничний цикл існує, залежить від циклічного порядку навчання, але не залежить від $x_0 x_0$	Точки, що породжуються алгоритмом (1.15), лежать в мінімальній зоні захоплення з вірогідністю 1
6	$N \geq p$	$R < p$	$R < p$	Перетином ϵ гіперплощина розмірності $p - R$	Збіжність до єдиної точки, якщо початкова точка x_0 лінійно залежить від \hat{a}_i , інакше – до точки, яка залежить від x_0	
7	$N \geq p$	$R < p$	$R + 1 \leq p$	Рівняння (1.14) несумісні. Не існує єдиної гіперплощини перетину	Граничний цикл існує, як і у випадку 3	Як у випадку 5

Оскільки на кожному такті навчання $\|a(k)\|^2 = \sum_{i=1}^n a_i^2(k) = \rho$, алгоритм (3.13) можна записати в еквівалентній векторній формі

$$w(k+1) = w(k) + \frac{\rho y^*(k) - w^T(k)a(k)}{\|a(k)\|^2} a(k), \quad (3.19)$$

що є алгоритмом Качмажа, і відомий в теорії ШНС як алгоритм Уїдроу і Хоффа. Під час використання алгоритму (9) на кожному такті навчання всі ρ коректованих ваг змінюються на однакову величину.

Слід зазначити, що алгоритм Качмажа, у свою чергу, є окремим випадком (3.18), оскільки в ньому використовується операція псевдо звернення вектора $a(k)$.

На відміну від алгоритму Альбуса, в алгоритмі Качмажа елементи вектора можуть бути довільними дійсними числами, тоді як елементи вектора $a(k)$ в алгоритмі Альбуса – бінарні числа. Отже, алгоритм навчання Альбуса – окремий випадок алгоритму Качмажа, і, отже, до алгоритму Альбуса застосовні всі відомі до теперішнього часу висновки щодо алгоритму Качмажа.

Властивість збіжності послідовності $w(k)$ досліджені в роботах Мілітцера і Паркса. У цих роботах було вивчено два способи навчання: циклічне навчання і випадкове.

Циклічне навчання – це така форма навчання мережі, при якій процес обходу N точок навчання повторюється в заданому порядку. При невеликих значеннях N можна отримати матричне співвідношення, яке пов'язує вагові вектори на початку і кінці циклу, та досліджувати збіжність алгоритму за поведінкою власних значень відповідної квадратної матриці.

Під час дослідження питань збіжності алгоритму (3.13) використовувалася функція Ляпунова

$$V(k) = y^T(k-1)L(I_N + L + L^T)^{-1}L^T y(k-1), \quad (3.20)$$

$$\text{де } L = \begin{bmatrix} 0 & 0 & \dots & 0 \\ \hat{a}_2^T \hat{a}_1 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & 0 \\ \hat{a}_N^T \hat{a}_1 & \hat{a}_N^T \hat{a}_{N-1} & \dots & 0 \end{bmatrix},$$

I_N – одинична матриця розмірності $N \times N$.

Як випливає з результатів цих робіт, необхідною і достатньою умовою збіжності алгоритму (3.13) є позитивна визначеність матриці $(I_N + L^T)(I_N + L + L^T)(I_N + L)$ або, що еквівалентно, матриці $(I_N + L + L^T)$.

Оскільки

$$(I_N + L + L^T) = \begin{bmatrix} 1 & \hat{a}_1^T \hat{a}_2 & \dots & \hat{a}_1^T \hat{a}_N \\ \hat{a}_2^T \hat{a}_1 & 1 & \dots & \hat{a}_2^T \hat{a}_N \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \hat{a}_{N-1}^T \hat{a}_n \\ \hat{a}_N^T \hat{a}_1 & \hat{a}_N^T \hat{a}_2 & \dots & 1 \end{bmatrix}$$

є матрицею Грама, відповідно властивості цієї матриці необхідна і достатня умова збіжності, яка полягає в тому, щоб одиничні нормальні вектори \hat{a}_i ($i=1, 2, \dots, N$) були лінійно незалежні. Отже, необхідно, щоб $p \gg N$ і щоб ранг матриці A дорівнював N . Відповідає випадку 1 з табл. 3.1.

При випадковому навчанні на відміну від заданого циклічного порядку, розглянутого вище, кроки навчання здійснюються випадково з множини цілих чисел $1, 2, \dots, N^1$.

Вибираючи функцію Ляпунова вигляду

$$V(k+1) = (w(k+1) - w^*)^T (w(k+1) - w^*), \quad (3.21)$$

можна показати, що якщо одиничні вектори $\hat{a}_1, \hat{a}_2, \dots, \hat{a}_N$ лінійно незалежні,

$$\Delta V(k+1) = V(k+1) - V(k) = (w(k+1) - w^*)^T (w(k+1) - w^*) - (w(k) - w^*)^T (w(k) - w^*) \leq 0.$$

Тут $\mu(k+1) = \sqrt{\rho} y^*(k+1) - a^T(k+1)w(k)$.

Тобто алгоритм (3.13) збігається до фіксованої точки

$$w^* = A^T (I_N + L + L^T)^{-1} Y(k), \quad (3.22)$$

що знаходиться на перетині всіх N^1 гіперплощин.

Це підтверджує той факт, що коли на $k+1$ -му кроці опускається псевдоперпендикуляр з точки $w(k)$ в точку $w(k+1)$, лежачи на гіперплощині $a^T(k)w = \sqrt{\rho} y^*(k)$, то в прямокутному трикутнику ABC з верхівками $A = w^*, B = w(k), C = w(k+1)$, гіпотенуза AB довша за катет AC , прилеглий до прямого кута ACB (див. рис. 3.6, а). Отже, алгоритм породжуватиме послідовність точок $w(k)$, які прагнутимуть до w^* при $k \rightarrow \infty$ (цей висновок є правильним і для циклічного навчання).

Фіксована точка w^* існує і у разі, коли одиничні нормальні вектори не є лінійно незалежними, а система рівнянь (3.15) сумісна. Для несумісної системи рівнянь (3.15) не існує такої точки w^* , яка була отримана вище, і при випадковому навчанні не можна навести числові приклади, які можна було зробити у разі циклічного навчання. Тому були висловлені тільки деякі припущення, які були підтверджені для невеликих значень N^1 і які, мабуть, справедливі і в більш загальному випадку.

У даній роботі використовується поняття зон захоплення N^1 наборів точок, кожна з яких лежить в одній з N^1 гіперплощин, визначених рівнянням $Ax=h$ так, що кожен набір точок ортогонально проектується на інші N^1-1 набори точок. Мінімальна зона захоплення – така зона захоплення, яка не може

бути далі зменшена в розмірах без втрати властивостей ортогональності зони захоплення.

Тепер можна припустити, що для заданого набору N^1 несумісних рівнянь $Ax=h$ існує мінімальна зона захоплення Z , що має таку властивість, що будь-яка випадкова початкова послідовність $(x_0, x_1, \dots, x_k, \dots)$ майже напевно потрапляє в зону Z при $k \rightarrow \infty$. Звичайно, як тільки якийсь вектор x потрапить у зону Z , всі подальші вектори опиняться в цій зоні. Виразимо це припущення в строгій імовірнісній формі

$$p\{x_k \in Z\} = 1 - \varepsilon(k),$$

де $\varepsilon(k) \rightarrow 0$ при $k \rightarrow \infty$.

У загальному випадку вірогідність існування послідовностей, які не потрапляють в зону захоплення при випадковому навчанні, дорівнює $\varepsilon(k) \leq Ac^k$, де A і $c < 1$ – константи, залежні від початкової точки x , і геометрії мінімальної зони захоплення.

Розмір і форма мінімальної зони захоплення виявляються дуже великими і складними навіть у простих випадках при невеликих значеннях p і N^1 , і тому може виявитися важливішим при теоретичних дослідженнях обмежитися гіперсферою, центр якої $x^* = A^+h$ визначається псевдозверненням за Муром – Пенроузом. У цьому випадку гіперсфера S мінімального радіусу R описує зону Z , і тоді можна сказати, що $p\{x_k \in Z\} = 1 - \varepsilon(k)$, де $\varepsilon(k) \rightarrow 0$ при $k \rightarrow \infty$. Мінімальна помилка, тобто найбільша довжина перпендикуляра x_k до N^1 гіперплощин, буде тоді менше $2R_0$.

На рис. 3.7 показано прості зони захоплення і відповідні мінімальні зони захоплення.

З погляду класичного алгоритму Качмажа зони захоплення відповідають областям збіжності за наявності перешкод вимірювань. Так, за наявності перешкод вимірювань $\xi(k) \sim N(0, \sigma_\xi^2)$ у вихідному сигналі $y(k+1)$, тобто

$$\tilde{y}(k+1) = y(k+1) + \xi(k+1), \quad (3.23)$$

алгоритм Качмажа збігається в область, визначену формулою

$$\lim_{k \rightarrow \infty} \mathbb{M} \left\| w(k+1) - w^* \right\|^2 = \frac{N}{N-2} \frac{\sigma_\xi^2}{\sigma_x^2}, \quad (3.24)$$

де $k \rightarrow \infty$, $\mathbb{M} \bullet$ – символ математичного очікування;

N – розмірність вектора вхідного сигналу $x \sim N(0, \sigma_x^2)$.

При цьому

$$\lim_{k \rightarrow \infty} \mathbb{M} \left\| w(k+1) - w^* \right\|^2 = 0,$$

тобто оцінка вектора ваг є незміщеною.

Зменшити розмір області можна, модифікуючи алгоритм Качмажа так:

$$w(k+1) = w(k) + \gamma \frac{y(k+1) - w^T(k)x(k+1)}{\|x(k+1)\|^2} x(k+1), \quad (3.25)$$

де $\gamma \in 0, 2$ – параметр, що впливає на швидкість навчання.

Введення параметра γ у алгоритм змінює траєкторію оцінки. Відповідні траєкторії за відсутності перешкод вимірювань показано на рис. 3.1. Тут рис. 3.1, б відображує випадок $0 < \gamma < 1$, а рис. 1.1, в – випадок $1 < \gamma < 2$.

За наявності ж перешкод вимірювань така модифікація дозволяє забезпечити область збіжності, розмір якої визначається виразом

$$\lim_{k \rightarrow \infty} \mathbb{M} \left\| w(k+1) - w^* \right\|^2 = \frac{\gamma}{(2-\gamma)} \frac{N}{N-2} \frac{\sigma_\xi^2}{\sigma_x^2}. \quad (3.26)$$

Звідси видно, що приймаючи $\gamma \ll 1$, можна зменшити розмір області збіжності, а вибираючи параметр γ змінним, тобто задовольняючим умови Дворецького

$$\lim_{k \rightarrow \infty} \gamma(k) = 0; \sum_{k=1}^{\infty} \gamma(k) = \infty; \sum_{k=1}^{\infty} \gamma^2(k) < \infty, \quad (3.27)$$

можна отримати збіжність в точку, тобто

$$\lim_{k \rightarrow \infty} \mathbb{M} \left\| w(k+1) - w^* \right\|^2 = 0.$$

Модифікований алгоритм Альбуса, відповідний (3.25), матиме вигляд

$$w(k+1) = \begin{cases} w(k) + \gamma \frac{y(k+1) - w^T(k)x(k+1)}{\|x(k+1)\|^2} x(k+1); \\ w(k), \forall i, a_i(k) = 0. \end{cases} \quad (3.28)$$

3.3.2 Методи скорочення процесу навчання мережі

Однією з найважливіших характеристик алгоритму Качмажа є його швидкість збіжності, що визначає кількість кроків навчання (кількість ітерацій), необхідних для знаходження розв'язань системи рівнянь (3.15) із заданою точністю. Збіжність алгоритму істотно залежить від властивостей вхідної послідовності $a(k)$. Так, зокрема, якщо послідовність $a(k)$ є послідовністю взаємно ортогональних векторів, то розв'язання w^* можна знайти за кінцеве p число ітерацій, де p – розмірність вектора w . В інших випадках швидкість збіжності алгоритму Качмажа може виявитися недостатньою для знаходження і відстеження закону зміни розв'язання системи рівнянь (3.15) $w^*(k)$ у нестационарному випадку. У зв'язку з цим виникає необхідність в підвищенні швидкості збіжності алгоритму.

Швидкість збіжності алгоритму Альбуса визначається найбільшим за модулем власним значенням матриці $B = -(I_N + L)^{-1} L^T$ порядку N . При великих N ці власні значення обчислити важко. Функція Ляпунова дозволяє знайти оцінку швидкості збіжності у вигляді $1 - 1/\lambda_{\max}(P)^{1/2}$, де $\lambda_{\max}(P)$

позначає найбільше (дійсне, позитивне) власне значення матриць $P = (I_N + L^T)(I_N + L + L^T)^{-1}(I_N + L)$, яке також важко обчислити при великих N . Єдиною перевагою тут є те, що матриця P симетрична, тоді як матриця B такою не є. Результати чисельного моделювання і знаходження власних значень при невеликих N дозволяють припустити, що у багатьох випадках власні значення матриці B близькі $k + 1$, залишаючись усередині одиничного кола, і тому швидкість збіжності цього алгоритму виявляється невеликою.

У зв'язку з цим виникає завдання прискорення процесу навчання мережі. При розгляді цього питання необхідно відзначити таке.

У загальному випадку вхідним сигналом для СМАС є вектор $x \in R^n$, що кодується в нормалізований вектор асоціацій \mathbf{a} . Проте для подальшого вивчення істотним є визначення вхідного сигналу як скаляр x , тобто $x \in R^1$. Тоді код x буде записаний так

$$\hat{\mathbf{a}} = \hat{\mathbf{a}}(w) = (\hat{a}_1, \hat{a}_2, \dots, \hat{a}_p)^T,$$

де елементи $\hat{\mathbf{a}}(w)$ отримані з

$$\hat{a}_i = \begin{cases} \frac{1}{\sqrt{p}}, & \text{при } w+1 \leq i \leq w+p, \\ 0 & \end{cases}$$

Порівняльний аналіз різних алгоритмів навчання мережі СМАС шляхом тестування був проведений в роботі. У прийнятому в даній роботі підході, тестування кожного алгоритму навчання починається з того, що мережу навчають відновлювати функцію одного змінного, тобто $n = 1$, незважаючи на те, що даний тест є достатньо простим в порівнянні з реальними системами з великими значеннями n , N і p . Такі тести проводилися й іншими авторами. Ті алгоритми, що пройшли цей тест, потім тестуються на n -вимірних вхідних сигналах x з $n = 2, 3, 4$.

Вихідний сигнал мережі на вхідний сигнал x , що поступив, після k кроків навчання обчислюється як середнє значення активних ваг, які визначаються ненульовими елементами вектора $\hat{a}(s)$, за допомогою такого виразу

$$y(x, k) = \frac{1}{\sqrt{\rho}} \hat{a}_i^T(s) * w_k.$$

У роботі зроблений порівняльний аналіз наступних п'яти різних алгоритмів, базовим серед яких є алгоритм навчання Альбуса (AL).

3.3.3 Навчання за допомогою ковзаючого середнього (AV)

У даному алгоритмі значення вектора ваг $w(k)$ обчислюється як середнє значення корекцій, які отримано в базовому алгоритмі AL.

Для обчислення середнього використовується вектор-лічильник $c = (c_1, c_2, \dots, c_p)^T$ такої ж розмірності, як і вектор ваг $w(k)$. Початкове значення $c_0 = 0$.

Значення лічильника оновлюється в кінці k -го кроку навчання для того, щоб показати, як часто кожна вага була модифікована, так

$$c(k+1) = c(k) + \sqrt{\rho} \hat{a}(k+1),$$

де $\sqrt{\rho} \hat{a}(k+1)$ містить одиницю у позиції активної ваги і нуль – інакше.

Настройка ваг здійснюється так

$$w(k) = w(k-1) + \mu b(k), \quad (3.29)$$

де $\mu(k) = \sqrt{\rho} r(k) - \hat{a}^T(k) w(k-1)$,

$$b(k) = (b_1(k), \dots, b_p(k))^T$$

$$b_i(k) = \frac{\hat{a}_i(k)}{c_i(k)},$$

де $\hat{a}^T(k)$ – скорочений запис $\hat{a}^T(x(k))$, а $\hat{a}_i^T(k)$ – i -й елемент даного вектора.

Відзначимо, що: 1) випадковий шум з нульовим математичним очікуванням, який часто зашумляє навчальні дані y_j і присутній в реальних системах, – відфільтрований; 2) якщо система рівнянь (3.15) несумісна, алгоритм збігається в точку, а не до обмеженого циклу, як в алгоритмі AL. Проте формального доказу збіжності даного алгоритму на сьогодні немає.

Алгоритм AV (3.25) збігається швидше, ніж AL при циклічному навчанні, але надалі збіжність стає дуже повільною для незашумлених тестів. У разі зашумленого навчального сигналу результати, отримані за допомогою алгоритму AV, найкращі. Обчислювальні витрати невеликі і практично такі ж, як для алгоритму AL.

3.3.4 Навчання в точці з максимальною помилкою (ME)

Цей метод застосовується разом з нормальною процедурою навчання в тій точці вхідного простору $x(j)$, для якої бажаний вихідний сигнал $y^*(j)$, отриманий з найбільшою помилкою серед заданої кількості l , останніх навчальних пар. Шукана довжина l вибирається в інтервалі $1 \leq l \leq k$. Минулі і справжні навчальні пари розміщуються в тимчасовій пам'яті. Відповідна помилка в точці $x(j)$ обчислюється для $j = \overline{k-l+1, k}$ так

$$\begin{aligned} e_j(k+1) &= f(x_{(j)}(k)) - y_{(j)}^* = \\ &= \frac{1}{\sqrt{\rho}} \hat{a}_i^T(s_{(j)})x(k) - y_{(j)}^* \end{aligned}$$

або, використовуючи скорочений запис,

$$y_j(k) = \frac{1}{\sqrt{\rho}} \hat{a}_{(j)}^T x_{k-1} - y_{(j)}^*.$$

Індекс $(j) = m$ найбільшої помилки визначається так, що

$$|e_{mk}| \geq |e_{(j)k}| \text{ для } j = k - l + 1, \dots, k.$$

Якщо існує декілька найбільших помилок, тобто

$$|e_{m_1 k}| = |e_{m_2 k}| = \dots = |e_{m_\mu k}| \geq |e_{jk}|,$$

то вибирається найближча, тобто помилка m_j , відповідна найбільшому значенню j .

Корекція ваг для точки x_m відбувається так

$$w(k+1) = w(k) + \mu(k) \hat{a}_m, \quad (3.30)$$

$$\mu_k = \sqrt{\rho} r_m - \hat{a}_m^T w_{k-1}.$$

Відзначимо, що при $l = 1$ цей алгоритм також ідентичний AL. Подальше вивчення швидкості збіжності даного алгоритму порівняно з AL, циклічним і випадковим навчанням, було проведено Еллісоном, який також рекомендує даний ME алгоритм.

Алгоритм ME (3.30) показує гарні властивості збіжності практично для всіх тестів. Початкове значення коефіцієнта навчання достатньо велике, іноді навіть більш, ніж для алгоритму GS (див. нижче). Для випадку відсутності перешкод, кращі властивості збіжності виходять при більшому значенні l . Проте велике значення l не завжди підходить для зашумлених даних. Обчислювальні витрати помірні і зростають лінійно з l .

3.3.5 Часткова оптимізація довжини кроку в останньому перпендикулярному напрямі (OS)

У даному алгоритмі, як і в алгоритмі AL, використовується корекція вектора ваг у напрямі перпендикуляра до $\hat{a}(k)$, проте довжина кроку $\mu(k)$ вибирається не такою, що відразу забезпечує досягнення k -ї гіперплощини $\hat{a}^T(k)x = \sqrt{\rho}r_{(k)}$, а такою, що мінімізує суму квадратів перпендикулярів до гіперплощини, визначеної останніми l частинами навчальних даних, тобто критерій, що мінімізує

$$I_t = \sum_{i=k-j+1}^k (\hat{a}_{(j)}^T x_k - \sqrt{\rho}r_{(j)})^2 \rightarrow \min,$$

де l – ціле, що лежить в інтервалі $1 \leq l \leq k$.

Перпендикулярна відстань пропорційна відповідній помилці СМАС. Останні l частин навчальних даних $(x(j), y^*(j)), j = \overline{k-l+1, k}$ (або в еквівалентному записі $(w(j), y^*(j))$) записуються в тимчасову пам'ять.

Ваги настроюються так

$$w(k+1) = w(k) + \mu(k+1)u(k+1),$$

$$\mu(k+1) = \frac{\sum_{i=k-j+1}^k (\sqrt{\rho}r(k+1) - \hat{a}_{(j)}^T x(k)) \hat{a}^T(j) \hat{a}(k)}{\sum_{i=k-j+1}^k (\hat{a}^T(j) \hat{a}(k))^2}. \quad (3.31)$$

Відзначимо, що базовий алгоритм AL – окремий випадок даної процедури з $l = 1$. Для цього алгоритму також немає формального доказу його збіжності.

Алгоритм OS (3.31) добре підходить для зашумлених гладких тестових функцій. За його допомогою можна отримати в цьому випадку високу точність

при великому значенні l . Для незашумленого випадку швидкість збіжності відносно повільна і зменшуватиметься зі збільшенням l . Обчислювальні витрати більші, ніж для алгоритму ME (3.30).

Цікаві результати можна отримати при комбінації різних алгоритмів, наприклад, ME+AV. Загальну збіжність можна прискорити обмеженням значення лічильника, наприклад, $c_i \leq c_{\max}$. Ця ідея досліджуватиметься в подальших роботах.

Сьогодні рекомендується використовувати алгоритм ME (3.27), оскільки він досягає хорошої збіжності для широкого спектру тестів з помірними обчислювальними витратами.

3.3.6 Алгоритм навчання з самоналагоджувальними коефіцієнтами навчання

У роботі запропоновано покращений алгоритм з самоналагоджувальним коефіцієнтом навчання, який реалізується так:

а) знайти межі збіжності і отримати коефіцієнт навчання, що призводить до максимальної швидкості збіжності β_{\max} :

1) встановити максимальну кількість епох пошуку max_epoch1 і початкове значення коефіцієнта навчання β_0 , потім $\beta \leftarrow \beta_0$ і обчислити початкову вихідну помилку pre_Mse1 . Встановити $epoch = 1$ і $flag = 0$;

2) оцінити поточну вихідну помилку СМАС $Mse1$, потім хай $\beta_0 \leftarrow \beta$;

3) якщо $Mse1 \geq pre_Mse1$, то продовжити, інакше перехід до 5);

4) якщо $flag = 0$, хай $\beta \leftarrow \frac{\beta + \beta_0}{2}$; якщо $flag = 1$, хай $flag \leftarrow 1$. Перехід

до 6);

5) якщо $flag = 0$, хай $\beta \leftarrow \frac{\beta + \beta_0}{2}$, $flag \leftarrow 1$, і перехід до б); якщо $flag = 1$, хай $\beta_{\max} \leftarrow \beta$, вихід - β_{\max} , завершити крок а) і перехід на крок б);

б) $epoch \leftarrow epoch + 1$. Якщо $epoch \leq \max_epoch1$, перехід на 2); інакше закінчити програму, видати повідомлення про помилку, збільшити кількість епох пошуку;

б) пошук оптимального значення коефіцієнта навчання β_{best} в інтервалі $0, \beta_{\max}$.

Послідовність x_i , $i = 0, 1, 2, \dots$, що згенерована за допомогою Logistic mapping, має вигляд

$$x_{n+1} = 4x_n(1 - x_n), n = 0, 1, 2, \dots \quad (3.32)$$

та буде випадково розподілена в інтервалі $[0, 1]$ при $0 \leq x_0 \leq 1$ [84]. Встановити h , і використовуючи вираз (3.20), згенерувати послідовність x_i , $i = 0, 1, 2, \dots, h-1$ з початковим значенням $x_0 = 0.1$, і виконати перетворення $x_i = \beta_{\max} x_i$, потім, послідовність після перетворення буде випадково розподілена в інтервалі збіжності $0, \beta_{\max}$. Встановити параметр \max_epoch2 , відповідно використовуючи x_i , $i = 0, 1, 2, \dots, h-1$ як коефіцієнт навчання для пошуку \max_epoch2 епох. Порівнюючи з вихідними помилками відповідними x_i , вибрати x_i , яке робить помилку мінімальною і виконати $\beta_{best} \leftarrow x_i$, і потім перейти до кроку в);

в) використовуючи β_{best} як коефіцієнт навчання, продовжити традиційне навчання СМАС впродовж \max_epoch2 епох. Закінчити програму, досягши бажаного рівня помилки або після закінчення максимальної кількості епох.

Рекомендовані значення параметрів для запропонованого алгоритму такі: $\beta_0 = 2$; $\max_epoch1 = 100$; $h = 20$; и $\max_epoch2 = 100$.

3.4 Застосування нейронної мережі СМАС для обробки відеоінформації

Будь-яке плоске зображення за своєю суттю є деякою функцією колориметричної характеристики точки від її координат. Під колориметричною характеристикою можна розуміти яскравість і кольоровість точки в системі YUV, яскравість колірних складових у системі RGB, яскравість для зображень у відтінках сірого кольору. Математично задати будь-яке зображення реальних об'єктів практично неможливо внаслідок його складності. Класичні методи апроксимації статичними поліномами не дадуть бажаної якості, крім того, обчислення коефіцієнтів цих поліномів – у край складне і ресурсомістке завдання. Табличне подання зображень забезпечить максимальна якість, але водночас зажадає великих витрат пам'яті і пропускну здатності каналів зв'язку для зберігання і передачі. Вихід був знайдений в стисненні зображень методами архівації без втрат і з втратами, розглянутими в розділі 1. Проте такі методи або добре працюють тільки на вузьких класах зображень, або вимагають великих обчислювальних витрат для компресії і декомпресії. Ці методи засновані на особливості зображень, пов'язаній із залежністю сусідніх елементів. Подібна властивість безлічі вхідних сигналів використовується під час обчислення функцій розглянутою нейронною мережею СМАС, що дозволяє запропонувати її як засіб для стиснення і фільтрації зображень. Крім того, особливості структури мережі СМАС, такі як вибірка обмеженої кількості ваг з пам'яті на даному кроці, істотно зниження об'єму пам'яті для зберігання параметрів, роблять привабливим використання мережі СМАС при апаратній реалізації обробки зображень.

Далі розглянемо застосування мережі СМАС для реалізації обробки зображень. При цьому необхідно врахувати, що стандартне телевізійне зображення має розміри 720x576 точок, тобто кількість рівнів квантування для вхідних змінних, відповідних координатам точки (x,y) в зображенні, значення інтенсивності, в якій обчислює мережу, неоднакове. В цьому випадку

залежність (3.3) для розрахунку об'єму пам'яті, необхідного для кодування зображення мережею СМАС, набуде вигляду:

$$n_{\max} = \left\lceil \rho \left(\frac{R_x - 1}{\rho} + 1 \right) \left(\frac{R_y - 1}{\rho} + 1 \right) \right\rceil, \quad (3.33)$$

де R_x і R_y – кількість рівнів квантування для змінних R_x і R_y відповідно. Кількість ступенів квантування вважатимемо однаковою для обох змінних.

3.4.1 Застосування мережі СМАС при мікшуванні зображень

Операцію мікшування зображень в колориметричній системі RGB можна описати виразом

$$\begin{cases} R_O(x, y) = K_{R1}(x, y)R_{I1}(x, y) + K_{R2}(x, y)R_{I2}(x, y) \\ G_O(x, y) = K_{G1}(x, y)G_{I1}(x, y) + K_{G2}(x, y)G_{I2}(x, y), \\ B_O(x, y) = K_{B1}(x, y)B_{I1}(x, y) + K_{B2}(x, y)B_{I2}(x, y) \end{cases} \quad (3.34)$$

де буквений індекс визначає приналежність до колірної шини, а цифровий індекс – до шини даних.

Для системи YUV замість інтенсивностей колірних складових необхідно підставити значення яскравості і кольороорізаційних сигналів. Ця операція лежить в основі переважної кількості переходів між сюжетами у відеопотоці. Форма переходу, фігура або шторка, задається функцією присутності $K(x, y)$ і від можливостей системи обробки за реалізацією цієї функції залежить різноманітність можливих переходів. Табличне задання цієї функції вимагає витрат часу для зміни шторки або фігури, апаратне формування функцій присутності обмежує коло можливостей відеопроектора. Одним з можливих шляхів скорочення об'єму пам'яті, потрібної для задання функції присутності, є застосування для її апроксимації нейронної мережі СМАС.

За своєю суттю функція присутності – це деяке зображення, яке назвемо «фігура мікшування», за розмірами таке ж, як мікшоване зображення,

яскравість точок якого визначає ступінь присутності початкових зображень в результуючому. Так, чим яскравіше точка фігури, тим більше в результуючому зображенні в цій же точці присутнє перше зображення і відсутнє друге. На рис. 3.6 наведено мікшоване зображення а) і б), фігура мікшування в), задана у відтинках сірого, і результат мікшування г) початкових зображень з використанням фігури. Геометрично функція присутності для задання зображеної на рис. 3.6 фігури є конічною поверхнею, яку аналітично можна подати у вигляді

$$\begin{cases} \frac{K(x, y) - K_0}{c^2} = \frac{(x - x_0)^2}{a^2} + \frac{(y - y_0)^2}{b^2}, \\ K(x, y) = 0, K(x, y) = 255 \end{cases}, \quad (3.35)$$

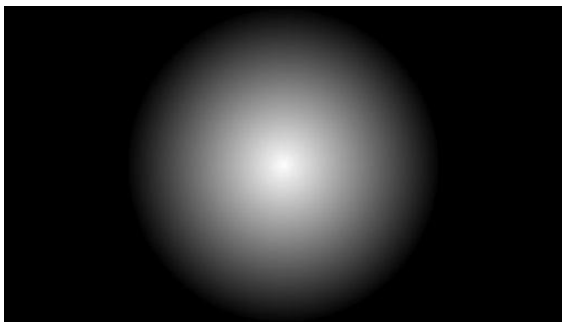
де параметри a , b , c визначають еліптичність і гостроту конічної поверхні, x_0, y_0 , K_0 – положення осі і вершини конічної поверхні.



а)



б)



в)



г)

Рисунок 3.6 – Приклад мікшування зображень

Площини $K(x,y) = 0$ та $K(x,y) = 255$ обмежують область припустимих значень функції присутності. Тут і далі вважатимемо, що повній присутності зображення 1 і відсутності зображення 2 відповідає значення функції присутності 255, а повній відсутності зображення 1 і присутності зображення 2 – значення 0.

Перевагою використання мережі СМАС для апроксимації функцій присутності є гнучкість, властива табличному способу задання функції і при цьому істотно менший необхідний об'єм пам'яті. Якщо зображення фігури має розмір 720×576 точок, об'єм пам'яті, необхідний для табличного її задання складе $720 \cdot 576 = 414720$ байт. Під час використання мережі СМАС з кількістю ступенів

$\rho = 4$ цей об'єм складе 104655 байт, а при $\rho = 8$ – 52981 байт, що, відповідно, в 3,96 і в 7,83 разів менше, порівняно з табличним способом. Наведені значення можуть бути зменшені під час використання хешування в мережі СМАС.

Під час використання мережі СМАС для синтезу функцій присутності в реальному для відеопотоку часі не потрібне навчання мережі безпосередньо в процесі використання. Його можна провести на математичній моделі мережі, реалізованої у вигляді програми на управляючій ЕОМ. Під час навчання як бажані значення функції присутності мережі можуть подаватися як обчислені за математичними залежностями значення функції присутності, так і отримані шляхом цифрування елементи розроблених дизайнером фігур. Отримані в результаті навчання вагові коефіцієнти передаються інтерфейсом з ЕОМ у відеопроцесор, де зберігаються в пам'яті, адреси читання якої формуються при апаратному кодуванні координат.

3.4.2 Виконання колірної корекції зображень з використанням мережі СМАС

У розділі розглянуто спосіб колірної корекції зображень для випадку залежності передавальних функцій коректора тільки від інтенсивностей колірної складової в даному каналі. Головний недолік такого способу корекції в тому, що він не дозволяє усувати такі спотворення, як появи небажаних відтінків на ділянках зображень певного кольору. Для такої корекції необхідно мати передавальні функції, залежні від всіх колірних складових

$$\begin{aligned} R_{out} &= f_R(R_{in}, G_{in}, B_{in}) \\ G_{out} &= f_G(R_{in}, G_{in}, B_{in}), \\ B_{out} &= f_B(R_{in}, G_{in}, B_{in}) \end{aligned} \quad (3.36)$$

де R_{in} , G_{in} , B_{in} – вхідні, а R_{out} , G_{out} , B_{out} – вихідні значення інтенсивностей кольорів коректора.

При цьому табличний метод вимагає $2^{3 \cdot 8} = 16777216$ байт пам'яті на один колірний канал при поданні інтенсивностей кольорів 8-бітовими кодами. При синтезі передавальних функцій мережею СМАС необхідно 1085874 і 284242 байти пам'яті при $\rho = 4$ і $\rho = 8$. Економія складе 15,45 і 59 разів відповідно.

Запропоновано використовувати нейронну мережу СМАС для колірної корекції зображень, що поступають з сканера на принтер. Розглянуто блокові структури для такої корекції. Проте детальну роботу блоків задання функцій корекції не розкрито. Крім того, реальний час для розглянутого завдання складає декілька десятків секунд на кадр (час проходження сканера). Внаслідок цього, модель колірної корекції зображень, що використовує нейронну мережу СМАС для апроксимації передавальних характеристик, потребує подальшого розвитку, пов'язаного з її адаптацією для роботи в реальному для відеопотоку часі.

Безпосередньо у колірному коректорі необхідно реалізувати мережу СМАС, що має три вхідні змінні і пам'ять ваг, що заповнюється значеннями, обчисленими при навчанні моделі мережі на управляючій ЕОМ. Задання навчальних дій у формі, зрозумілій для користувача відеопроцесора, є складним завданням, оскільки візуально необхідно подати функцію від трьох змінних. Можна запропонувати такий спосіб, який розглянемо на прикладі червоного колірного каналу:

1. Будується передавальна характеристика $f_R(R_{in}, 0, 0)$ шляхом графічного задання додаткових точок на графіці в системі координат (R_{in}, f_R) так, як наведено на рис. 3.7 (на рис. 3.7 осі позначено R_i і R_o відповідно);

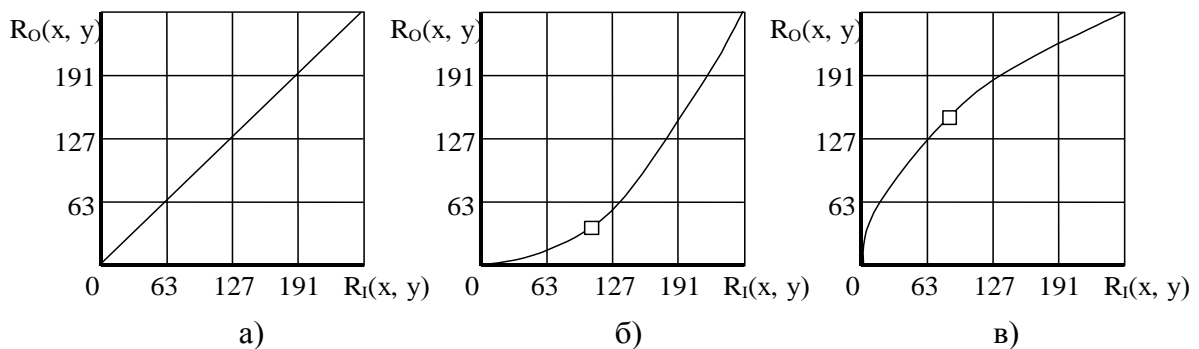


Рисунок 3.7 – Задання передавальних характеристик колірного коректора

2. У системі координат (G_{in}, B_{in}) задаємо положення точок, в яких інтенсивність червоного кольору має залежати від значень інтенсивності зеленого і синього кольорів.

3. У системах координат (G_{in}, dR) (B_{in}, dR) задаються величини і знак залежності червоного від зеленого і синього кольорів відповідно.

4. Отримуємо декілька точок передавальної характеристики для каналу червоного кольору:

$$R_{out} = f_R(R_{in}, 0, 0) + dR(G_{in}) + dR(B_{in}). \quad (3.37)$$

5. Подаємо отримані точки як навчальні програмної моделі мережі СМАС.

Зважаючи на згладжуючі властивості мережі СМАС, даний метод задання передавальних характеристик колірною коректора має забезпечити плавні передавальні характеристики, що не мають розривів і різких стрибків.

3.4.3 Застосування мережі СМАС для стиснення і фільтрації зображень

Подібно до того, як функцію присутності $K(x,y)$ можна подати у вигляді зображення, що має в точці (x,y) значення яскравості, рівне значенню функції присутності, будь-яке зображення можна подати у вигляді функцій, що визначають значення його колірних компонентів в кожній точці

$$\begin{cases} R = f_R(x, y) \\ G = f_G(x, y) \\ B = f_B(x, y) \end{cases} \quad \text{або} \quad \begin{cases} Y = f_Y(x, y) \\ U = f_U(x, y) \\ V = f_V(x, y) \end{cases}. \quad (3.38)$$

Вид функцій $f(x,y)$ визначається змістом зображення і математично, в переважній більшості випадків, не подається. Табличне подання зображень вимагає великих об'ємів пам'яті і швидкостей вводу-виводу та передачі для зберігання і передачі відеоінформації, хоч і забезпечує максимальну її якість. Всі існуючі алгоритми стиснення припускають, що елементи зображень залежні між собою, а сусідні кадри у відеопотоці відрізняються положенням деяких об'єктів. Багато хто з них видаляє, вважаючи надмірними, високі частоти, присутні в зображенні, хоча ці складові впливають на різкість і чіткість меж об'єктів. Алгоритми, що мають високий ступінь стиснення, або істотно погіршують якість зображень, або вимагають великих обчислювальних витрат.

Якщо розглядати зображення як деяку функцію від координат, задачу стиснення можна розв'язати шляхом апроксимації цієї функції, способом, що вимагає менше пам'яті для зберігання коефіцієнтів. Таку властивість має дана нейронна мережа СМАС.

Метод стиснення зображення мережею СМАС полягає в такому. Вхідними сигналами мережі є координати (x,y) точок зображення, що стискається, на екрані. Ці значення кодується вхідним шаром мережі, формуючим вектор асоціацій. Відповідно до цього вектора, з пам'яті зчитуються ρ ваг, які з урахуванням значень базисних функцій першого шару підсумовуються в третьому шарі. Набуте значення функції порівнюється із значенням інтенсивності в даному колірному каналі зображення. Отримана помилка використовується як критерій підстроювання ваг алгоритмом навчання, якому необхідно модифікувати ті ваги, які були використані під час обчислення мережею значення функції.

Для опису процесу стиснення зображення позначимо через $I_{in}(x,y)$ функцію, що виражає залежність від координат інтенсивності однієї з складових початкового зображення. В результаті апроксимації отримуємо зображення $I_1(x,y)$, пов'язане з початковим співвідношенням

$$I_{in}(x, y) = I_1(x, y) + e_1(x, y), \quad (3.39)$$

де $e_1(x,y)$ – помилка, що отримується при апроксимації. Цю помилку можна подати як деяке різницеве зображення, що отримується шляхом віднімання початкового і апроксимованого зображень. Це зображення, подібно до початкового, може апроксимуватися мережею СМАС, внаслідок чого отримуємо зображення $I_2(x,y)$ і помилку $e_2(x,y)$ у вигляді

$$e_1(x, y) = I_2(x, y) + e_2(x, y). \quad (3.40)$$

Під час використання n мереж СМАС, і, відповідно, n кроків апроксимації, маємо співвідношення, що отримується з підстановки (3.40) в (3.39) і узагальнення на n кроків

$$I_{in}(x, y) = \sum_{j=0}^n I_j(x, y) + e_n(x, y). \quad (3.41)$$

Особливістю (3.41) є те, що воно визначає перетворення початкового зображення без втрат. Якщо на n -му кроці апроксимації помилкою нехтувати, отримаємо співвідношення

$$I_{in}(x, y) = \sum_{j=0}^n I_j(x, y). \quad (3.42)$$

Кількість кроків n і кількість ступенів квантування на кожному кроці апроксимації визначають якість і ступінь стиснення зображення.

Особливістю пропонованого методу стиснення зображень є те, що якість апроксимації залежатиме від кількості поданих під час навчання мережі точок зображення. Ці точки можуть вибиратися в початковому зображенні як за будь-яким критерієм (межею, областю, об'єктом), так і випадково за допомогою будь-якого з алгоритмів формування псевдовипадкової або випадкової послідовності. Максимальна якість буде досягнута після подання мережі всіх точок зображення.

Слід відзначити, що під час обробки зображення розгортається в часі рядками і стовпчиками. Проте таке систематизоване подання навчальних точок призводить до появи направлених «змазувань» у відновленому зображенні. Для усунення цього недоліку необхідно мати проміжний буфер, який записується послідовно, а читається випадково.

Структуру кодера-декодера зображень, заснованого на застосуванні апроксимації функцій мережею СМАС, наведено на рис. 3.8.

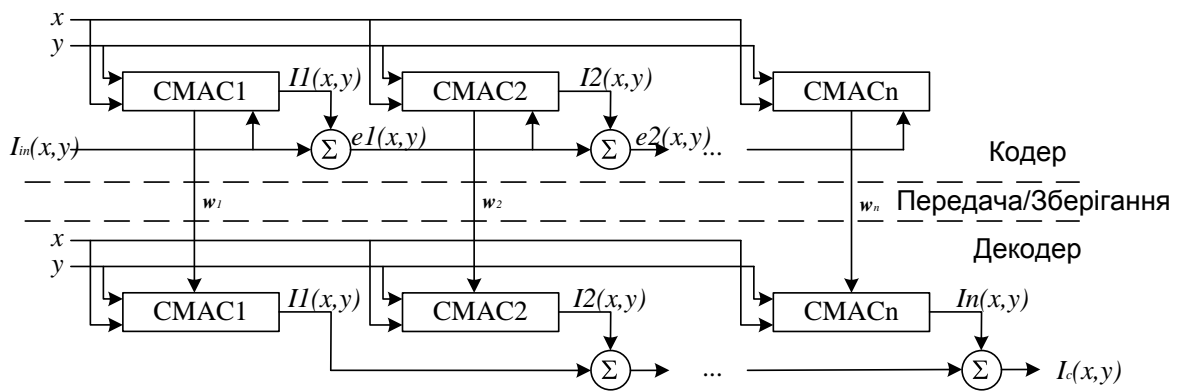


Рисунок 3.8 – Ієрархічна структура СМАС для кодування-декодування зображень

На наведеному рисунку мережі СМАС є вищезрозглянуті структури, що складаються з вхідного кодуючого шару, асоціативного шару і вихідного обчислюючого шару. Відмінність мереж СМАС кодера полягає в тому, що вони мають засоби для настройки вагових коефіцієнтів. Ці коефіцієнти і є закодоване (стисле) зображення, яке можна зберігати, або передавати лінією зв'язку. У декодері ці вагові коефіцієнти заносяться в асоціативну пам'ять відповідних мереж і беруть участь в обчисленні значень функцій. Отже, декодуюча сторона простіша і дешевша в реалізації, що є одним з важливих атрибутів алгоритму стиснення у зв'язку з тим, що декодування виконується в апаратурі користувачів відеоінформації, яких, як правило, в декілька тисяч разів більше, ніж об'єктів, що формують відеосигнал.

Іншою важливою особливістю використання мережі СМАС для апроксимації зображень є можливість фільтрації зображень. Розглянемо залежність (3.39). Якщо припустити, що в цій залежності зображення $I_1(x,y)$ містить частина частот початкового зображення, то в цьому випадку помилка $e_1(x,y)$ міститиме ті частоти, які присутні в початковому зображенні і відсутні в обчисленому зображенні $I_1(x,y)$. В цьому випадку мережа СМАС буде фільтром тих або інших частот, а отже, може використовуватися для згладжування і виділення контурів, очищення від шуму, виділення деталей, що з'явилися.

Частоти, що виділяються мережею, визначаються параметром ρ і точками початкового зображення, поданими мережі для навчання. Чим вище ρ – тим більше низькочастотні складові виділить мережа, при цьому для навчання мережі необхідно подати в ρ^2 менше точок, ніж в зображенні. Для фільтрації може застосовуватися структура, наведена на рис. 3.8. Ця структура дозволить виділяти ті або інші групи частот (кожна мережа відповідає за певну групу частот, відповідно до якої вибираються її параметри і проводиться навчання).

3.5 Модифіковані архітектури ШНМ СМАС

ШНМ СМАС була запропонована Альбусом ще в 1975 р. У цей час мережі СМАС успішно застосовуються в багатьох сферах життєдіяльності суспільства, таких як роботизоване керування, обробка сигналів, збір даних у глобальній комп'ютерній мережі, а також розпізнавання образів. Поширенню нейронних мереж СМАС чимало сприяла здатність до швидкого навчання, гарна здатність до узагальнення й простота апаратної реалізації.

Попередній досвід вивчення ШНМ СМАС полягав у спробах створення розширених алгоритмів навчання СМАС, оптимізації структури мережі СМАС, виборі параметрів навчання для підвищення швидкості збіжності, апаратної реалізації моделі СМАС, а також застосуванні нейронних мереж СМАС у різних сферах. Властивості збіжності СМАС доведені математично.

Для навчання ШНМ СМАС застосовується локальний навчальний підхід, а саме: кожна точка навчальних даних асоціюється з кінцевим набором ваг, спільно використовуваних сусідніми точками. Отже, між вагами, настроєними по одній точці навчальних даних, і сусідніми точками можуть виникати колізії.

В СМАС Альбуса використовуються константні (одиничні) БФ для виконання локально-локально-зваженої апроксимації сигналів. У результаті застосування константних БФ настроювання вагових коефіцієнтів мережі здійснюється за допомогою рівномірного розподілу похибки на асоційованому

наборі ваг. У межах кожного дискретного стану виходом мережі СМАС є константа, у той час як інформація про похідні не зберігається.

Запропонованій Альбусом моделі СМАС властиві два основних недоліки: по-перше, для моделювання функцій багатьох змінних або при класифікації багатомірних образів потрібно мати надмірну кількість пам'яті; по-друге, існує проблема рівномірності дискретизації вхідного простору навчальних наборів.

Як показано вище, при моделюванні багатомірних функцій з ростом розмірності простору вхідних змінних обсяг пам'яті, необхідний для зберігання вагових коефіцієнтів класичної СМАС, зростає експоненційно. Крім того, збільшується кількість необхідних точок навчальних даних і зростає складність кодування інформації. Обмежені ж об'єми й швидкодія пам'яті в прикладних системах значною мірою звужують галузь практичного застосування СМАС у реальних застосуваннях.

У свою чергу, при реалізації рівномірного квантування вхідного простору навчальних сигналів доводиться ігнорувати проблему нерівномірності розподілу навчальних образів, у результаті чого виділяються невикористовувані згодом області пам'яті й знижується продуктивність навчання як наслідок неоптимального використання пам'яті для зберігання отриманих у процесі навчання даних.

Серед відомих способів вирішення описаної проблеми найбільш відомим є гешування, що полягає у відображенні більшої необхідної пам'яті в меншу, фізично реалізовану. Застосування геш-кодування істотно знижує вимоги ШНМ СМАС до пам'яті при рішенні деяких задач. Але при цьому неминуче виникає проблема геш-колізій, що полягає в появі небажаної кореляції між вхідними сигналами, що досить негативно позначається на апроксимуючих властивостях мережі, що може призвести до зниження швидкості збіжності й навіть розбіжності алгоритму навчання.

Для зниження вимог до пам'яті при моделюванні багатомірних задач була запропонована модульна СМАС, для якої з метою оптимізації обчислень

Альбусом застосований метод тимчасової інверсії. В СМАС з тимчасовою інверсією для вирішення багатомірних задач використовуються підключені послідовно СМАС меншої розмірності. При цьому модулі СМАС вихідного шару породжують тимчасові точки навчальних даних для настроювання параметрів СМАС малої розмірності, розташованих у нижніх шарах.

Так, для рішення багатомірних завдань Лином і Чи була запропонована неймережева структура, що складається з малих двовимірних мереж СМАС. На практиці така структура показала недієздатність у випадку неправильного вибору архітектури модулів мережі. Фактично, метод Ліна й Чи схильний до проблеми перестановки вхідних змінних при вирішенні деяких задач класифікації образів з високою деталізацією вимірів (таких, як набір даних Лунга з 56-ю вхідними ознаками).

Для моделювання багатомірних гладких функцій Хангом і Яном використана нейронна мережа MS_СМАС. Мережа такого типу поєднує множину одномірних ШНМ СМАС у деревовидну структуру, що навчається за допомогою методу тимчасової інверсії. Крім того, Хангом і Яном розроблено схему квадратичного сплайна для згладжування значень виходів одномірної мережі СМАС за допомогою трансформації ступінчастих ваг у безперервно-гладкі ваги. Однак як недолік нейронної мережі MS_СМАС необхідно вказати нездатність вирішувати задачі класифікації багатомірних образів, оскільки її деревовидна структура швидко росте зі збільшенням розмірності вхідних векторів.

Спроби вирішення задачі пошуку оптимальних схем дискретизації вхідного простору сигналів у ШНМ СМАС здійснюються протягом досить тривалого часу, але дотепер її вирішення не втрачає актуальності. Так, з метою адаптивного розподілу пам'яті пропонується застосування в мережах СМАС декількох шарів дискретизації, що мають різну роздільну здібність.

Також використовується підхід з ітеративним зворотним зв'язком на виході СМАС для настроювання вхідної функції дискретизації й, як наслідок,

рівномірного розподілу цільових вихідних сигналів системи. Не слід забувати, що для використання ітеративного зворотного зв'язку необхідна похідна цільової функції за часом для корекції функції дискретизації, а це вкрай ускладнює процес навчання. Крім зазначеного недоліку, розглянутий підхід також не може бути застосовано для вирішення задач класифікації образів. З метою усунення описаного недоліку Бергером був запропонований адаптивний бінарний метод розподілу вхідного простору на прийнятні діапазони дискретизації вхідного простору при мінімізації помилки навчання.

У свою чергу, досліджене застосування кластеризації для одержання адаптивного вирішення дискретизації вхідного простору. На закінчення можна відзначити, що зазначені підходи завжди передбачають компроміс між економією сховища даних і обчислювальною складністю розглянутої моделі.

3.6 Адаптивна дискретизація вхідних сигналів

Адаптивне настроювання схеми дискретизації вхідних сигналів у нейронній мережі СМАС можуть здійснюватися на основі обчислення й аналізу коефіцієнта подібності Грея, що використовується для визначення міри подібності послідовностей.

Якщо задано еталонну послідовність $x_0 = \{x_0^1, x_0^2, \dots, x_0^m\}$, з якою порівнюють m послідовностей виду $x_i = \{x_i^1, x_i^2, \dots, x_i^m\}$, $i = 1, 2, \dots, m$, то коефіцієнт подібності Грея між x_0 і x_i на k -ому кроці визначається як

$$r(x_0, x_i) = \frac{\Delta_{\min} + \zeta \cdot \Delta_{\max}}{\Delta_{0i} + \zeta \cdot \Delta_{\min}}, \quad (3.43)$$

де $\Delta_{0i} = |x_0 - x_i|$, $\zeta \in [0, 1]$ – коефіцієнт відмінності Δ_{\max} й Δ_{\min} , $\Delta_{\max} = \max_i \max_j \Delta_{0i}$, $\Delta_{\min} = \min_i \min_j \Delta_{0i}$. Як витікає з (3.43),

$$0 < r(x_0, x_i) \leq 1.$$

Після визначення всіх коефіцієнтів подібності Грея можна обчислити їх зважене середнє, що називають ступенем подібності Грея

$$r_{(0), x_i} = \sum_{k=1}^n w_k \cdot r_{(0), x_i}^{(k)} \quad (3.44)$$

де w_k – ваговий фактор, $\sum_{k=1}^n w_k = 1$ (звичайно вибирають $w_k = 1/n$ для всіх k).

При використанні адаптивної дискретизації навчання СМАС починається із заданою рівномірною схемою дискретизації. Нехай вхідний простір на початку розділено на n частин і бажана функція \hat{y} відома. Позначимо реакцію мережі (вихідний сигнал) на вхідний сигнал x_k ($k = 1, 2, \dots, n$), що надійшов, як $y^{(k)}$. Об'єднання n бажаних виходів дозволяє сформувати еталонну послідовність $\hat{Y} = \{\hat{y}^{(1)}, \hat{y}^{(2)}, \dots, \hat{y}^{(n)}\}$.

Для аналізу подібності Грея між бажаними й реальними вихідними сигналами мережі, порівнювана послідовність $Y = \{y^{(1)}, y^{(2)}, \dots, y^{(n)}\}$ генерується реальними виходами СМАС на кожному кроці навчання. Відповідно до (1.43), коефіцієнт подібності Грея для сигналу x_k для однієї порівнюваної послідовності може бути визначений як

$$r_{(0), \hat{y}^{(k)}} = \frac{\Delta_{\min} + \zeta \cdot \Delta_{\max}}{\Delta_{y^{(k)}} + \zeta \cdot \Delta_{\min}}, \quad (3.45)$$

де $\Delta_{y^{(k)}} = |y^{(k)} - \hat{y}^{(k)}|$, $\Delta_{\max} = \max_k \Delta_{y^{(k)}}$, $\Delta_{\min} = \min_k \Delta_{y^{(k)}}$ і $\zeta \in [0, 1]$ – коефіцієнт відмінності.

Внаслідок того, що при рівномірній дискретизації не враховується швидкість зміни бажаної функції \hat{y} , результати навчання можуть бути далекими від оптимальних. Пропонований метод адаптивної дискретизації

базується на мінімізації вихідної помилки СМАС, що виникла в результаті дискретизації вхідних сигналів. Як ступінь зміни бажаної функції можливе використання, наприклад, її похідної, що можна отримати заздалегідь. Однак, недоліком такого підходу є те, що коли бажана функція змінюється, то значення похідної також змінюється. З іншого боку, при розподілі вхідного простору на основі аналізу вихідної помилки виникає проблема визначення адекватної стратегії дискретизації.

Метод адаптивної дискретизації, що використовує коефіцієнт подібності Грея, не тільки відображає співвідношення між еталонною й порівнюваною послідовністю в стані x_k , але й характеризує вихідну помилку – чим більше коефіцієнт Грея, тим менше значення вихідної помилки.

Основною перевагою використання коефіцієнта подібності Грея є те, що він легко обчислюється наприкінці кожної ітерації навчання мережі. Якщо результат незадовільний, то відбувається корекція кроку дискретизації з використанням коефіцієнта подібності Грея у відповідності зі значенням функції ухвалення рішення про дискретизацію $Q(\epsilon)$ як показано на рис. 3.9.

Якщо в деякій області коефіцієнти подібності Грея менше деякого фіксованого значення r_c (межа збільшення розділення), то крок дискретизації в цій області зменшується, якщо ж вони більше, ніж r_d (межа зменшення розділення) – то збільшується. Якщо коефіцієнти подібності Грея знаходяться усередині інтервалу $[r_c, r_d]$, відповідні інтервали дискретизації не змінюються.

Нехай $r(\epsilon)$ – коефіцієнт подібності Грея в стані x_k . Якщо існує кілька послідовних станів x_i, x_{i+1}, \dots, x_j , що задовольняють умові, що всі їхні коефіцієнти подібності менше r_c , тобто $0 < r(\epsilon) < r_c$ для $k = i, i+1, \dots, j$, то число дискретизації в області, яка поєднує ці стани, позначене як $R_{ij} = [x_i^L, x_j^R]$, повинно збільшуватися для усунення помилки навчання.

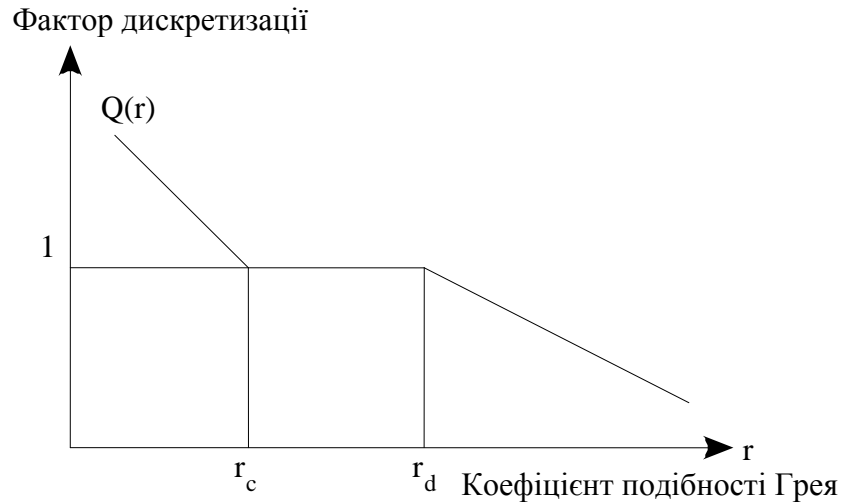


Рисунок 3.9 – Функція ухвалення рішення про дискретизацію

Нехай r_{ij} – середнє r для $k = i, i + 1, \dots, j$, тоді число дискретизації в об'єднаній області R обчислюється за допомогою функції ухвалення рішення про дискретизацію як

$$\Omega_{ij} = \lceil Q(r_{ij}) N_x \rceil, \quad (3.46)$$

де $\lceil \cdot \rceil$ позначає округлення до найближчого цілого; $Q(r_{ij}) \geq 1$ – значення функції ухвалення рішення про дискретизацію; $N_x = j - i + 1$ – представляє число послідовних станів, що задовольняють наведеним вище вимогам.

З виразу (3.46) видно, що $\Omega_{ij} > N_x$, тобто розділення в області R_{ij} збільшується. Також, крок дискретизації в області R_{ij} після адаптивного настроювання стає рівним

$$d_{ij} = \frac{x_j^R - x_i^L}{\Omega_{ij}}. \quad (1.47)$$

Внаслідок того, що дискретизаційне число в області R_{ij} змінюється з N_x на Ω_{ij} , число станів у вхідному просторі X стає рівним $n - N_x + \Omega_{ij}$, на відміну від вихідного n .

Інакше, якщо $r \ll r_d$ для $k = i, i+1, \dots, j$, то вихідні сигнали мережі в цих станах близькі або дуже близькі до їх відповідних бажаних значень. У цьому випадку крок дискретизації в об'єднаній області R_{ij} повинен бути збільшений для зменшення розділення в цій області. Відповідно до функції ухвалення рішення про дискретизацію на рис. 3.9, можна помітити, що в розглянутому випадку $0 < Q_{ij} \lesssim 1$, а тому дискретизаційне число Ω_{ij} , визначене в (3.46), повинно задовольняти $\Omega_{ij} < N_x$, що відповідає зменшенню розділення в області R_{ij} . У цьому випадку кількість станів у вхідному просторі X після адаптивного настроювання дискретизації стане рівним $n - N_x + \Omega_{ij}$, що менше вихідного n . Крок дискретизації обчислюється також за допомогою (3.47).

Слід відзначити, що наведені вище ствердження справедливі тільки до послідовностей станів, а не до одиничних станів. При розгляді одиничного стану можна припустити, що $i = j$. Якщо коефіцієнт подібності Грея в цьому стані менше r_c , то метод адаптивної дискретизації може бути застосований для збільшення розділення дискретизації. Але збільшити інтервал дискретизації для зменшення розділення в цьому випадку неможливо, тому що для одного стану дискретизаційне число вже мінімально й дорівнює одиниці. Крім того, неможливо зменшити розділення області, якщо коефіцієнт Грея більше r_d . Таким чином, використовуючи вирази (3.46) і (3.47), можна визначити дискретизаційне число Ω_{ij} в розглянутій області R_{ij} й крок дискретизації в тій же області.

На рис. 3.10 показана схема адаптивної дискретизації змінної x_2 , тобто тут $r_1 = 1$, r_2 – змінна.

У випадку використання адаптивної дискретизації, алгоритм навчання мережі СМАС матиме такий вигляд:

- 1) Вибрати адекватне число ступенів дискретизації n .
- 2) Зробити настроювання вагових коефіцієнтів за допомогою будь-якого алгоритму навчання мережі СМАС.
- 3) Обчислити вихідні значення мережі СМАС для всіх наборів вхідних навчальних даних.
- 4) Обчислити коефіцієнти подібності Грея між реальними й бажаними вихідними сигналами за допомогою (3.45).

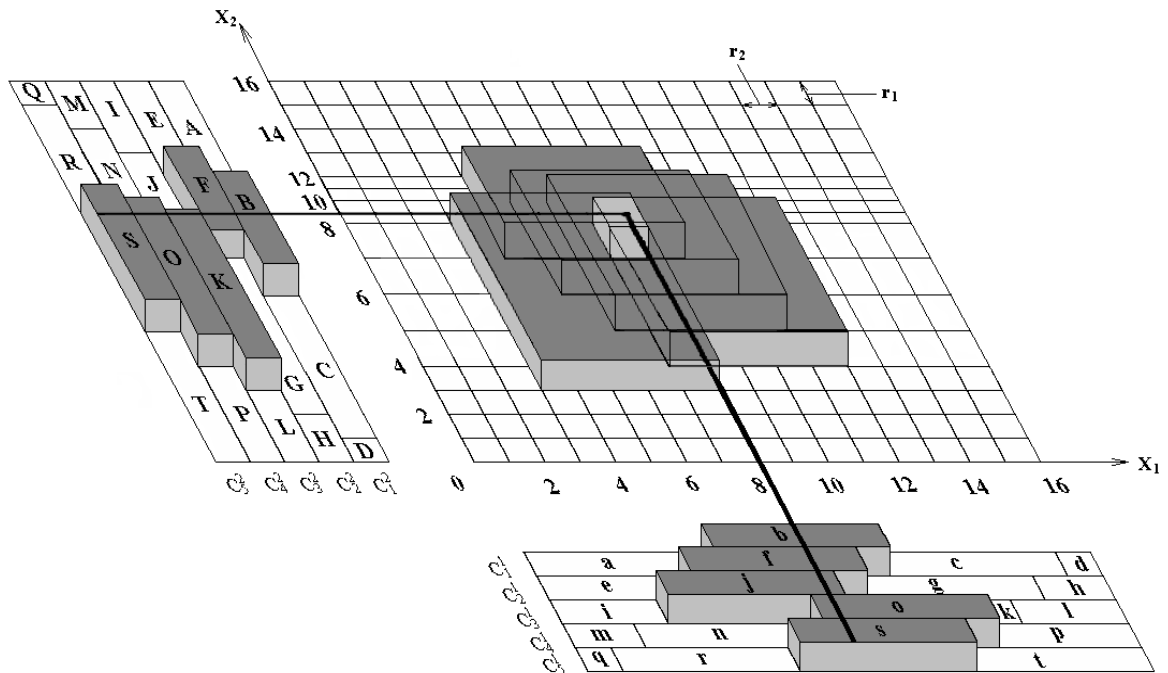


Рисунок 3.10 – Схема дискретизації СМАС після корекції

5) Скорегувати схему дискретизації вхідного простору відповідно до функції ухвалення рішення про дискретизацію (3.46).

6) Перерозподілити вміст пам'яті у відповідності зміною схеми квантування.

7) Якщо загальна помилка зростає, і швидкість навчання стала менше або пройшли всі кроки навчання, необхідно тимчасово призупинити навчання; інакше повторювати із кроку 2 поки результати не будуть задовільними.

3.7 Мережа Low Dimensional Based CMAC

Одним з ефективних способів мінімізації об'єму необхідної пам'яті мережі CMAC при роботі з багатомірними об'єктами є застосування запропонованих мереж Low Dimensional Based CMAC (LDB CMAC). Використання подібного підходу звільняє розроблювача від деяких проблем, що супроводжують реалізацію алгоритму зворотного поширення помилки при дослідженні багатомірних функцій.

Приклад структури мережі LDB CMAC показаний на рис. 3.11, де n – кількість базових нейронних мереж CMAC, а N – число вхідних змінних. В основі такої структури лежить безліч малорозмірних (базових) нейронних мереж CMAC (наприклад, двовимірних), на які подаються всі можливі парні комбінації вхідних сигналів.

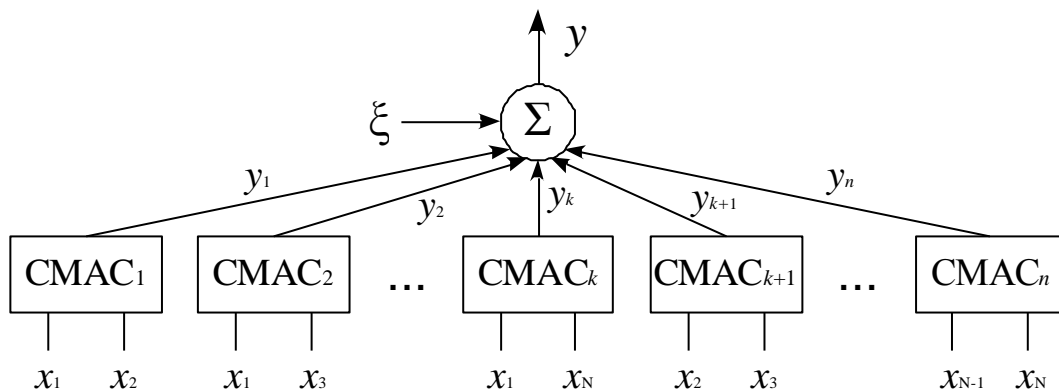


Рисунок 3.11 – Структура мережі LDB CMAC

Зважені вагами z_i ($i = 1, 2, \dots, n$) виходи цих CMAC формують загальний вихід мережі

$$y = \sum_{i=1}^n y_i z_i, \quad (3.48)$$

де $y_i^{(k)}$ – вихідне значення базової мережі $СМАС_i$; z_i – вага відповідного виходу $y_i^{(k)}$, $i = 1, 2, \dots, n$.

Крім цього, розглянута архітектура дає можливість створення неповних структур, у яких беруть участь не всі можливі комбінації пар вхідних змінних, а лише частина з них.

Процедура навчання нейронної мережі LDB $СМАС$ полягає в наступному:

- 1) Настроювання ваг z_i , наприклад, відповідно алгоритму
- 2)

$$z_i^{(k+1)} = z_i^{(k)} + \frac{y_i^{(k)} - \hat{y}^{(k)}}{\|Y^{(k)}\|} Y_i^{(k)}, \quad (3.49)$$

де $Y^{(k)} = \{y_1^{(k)}, y_2^{(k)}, \dots, y_n^{(k)}\}$ – вектор виходів базових мереж $СМАС$.

2) Настроювання параметрів базових мереж $СМАС_i$, $i = 1, 2, \dots, n$ за допомогою будь-якого алгоритму навчання мережі $СМАС$.

Слід зазначити, що на відміну від ШНМ $НСМАС$, розглянутої далі, LDB $СМАС$ не вимагає використання диференційовних БФ і дозволяє застосовувати градієнтний метод навчання при виборі БФ будь-якої форми, включаючи найпростішу прямокутну.

3.8 Ієрархічна мережа $НСМАС$

Як більш ефективний спосіб рішення описаних вище проблем пропонується застосування ієрархічної структури $СМАС$ (Hierarchical $СМАС$ - $НСМАС$), що складається з декількох простіших модулів, наприклад, двовимірних $СМАС$ з неконстантними диференційовуваними гавсівськими БФ ($ГСМАС$).

Запропонована $НСМАС$ вирішує проблему підвищеної вимоги до пам'яті вихідної моделі $СМАС$ за допомогою декомпозиції багатомірної задачі на

декілька двовимірних підзадач. Можливе застосування структури, що самоорганізується, вхідного простору, заснованої на мірі ентропії Шеннона і методі пошуку золотим перерізом для визначення оптимальної схеми дискретизації вхідного простору, основою якої є розподіл вхідних навчальних даних.

У порівнянні з іншими підходами до дискретизації вхідних просторів структурі, що самоорганізується, не потрібна інформація про похідну. Цей спосіб більше підходить для розв'язання задач класифікації образів, але, у свою чергу, не може бути застосован для подання дискретизації вхідного простору в випадках, що не можуть бути категоризовані, наприклад при розв'язанні задач апроксимації функцій.

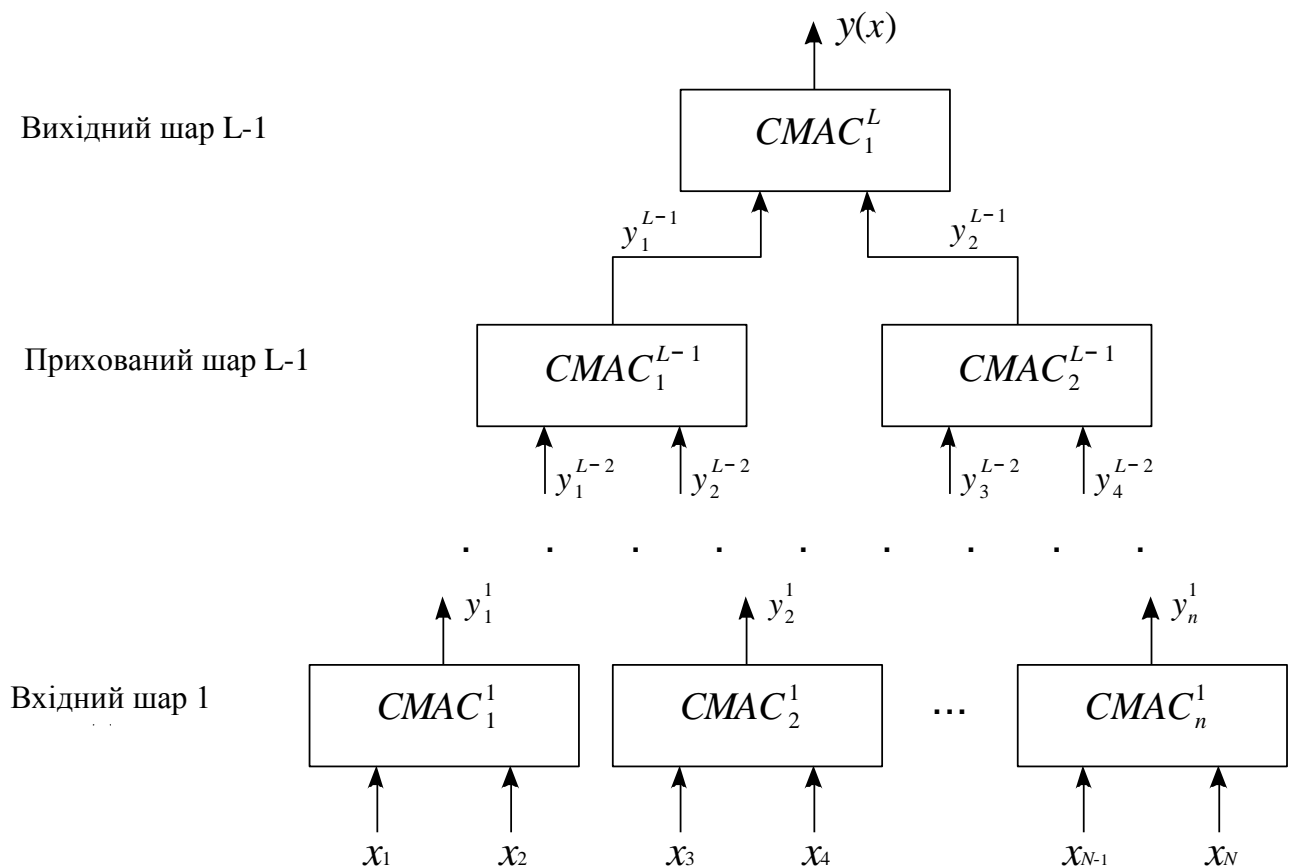
Після того, як вхідний простір розбивається на дискретні області, виділення пам'яті в нейронній мережі НСМАС відбувається автоматично. Для настроювання параметрів мережі НСМАС, що самоорганізується, пропонується градієнтно-спадне навчальне правило. Там же показано, що застосування класифікатора, що самоорганізується, на основі НСМАС дозволяє досягти гарних результатів у класифікації багатомірних образів і дозволяє домогтися значної економії пам'яті в порівнянні з багатомірною класичною мережею СМАС.

Структура нейронної мережі НСМАС, що самоорганізується, ієрархічно розкривається відповідно до топології повного бінарного дерева відповідно до розмірності вхідних векторів. При цьому можливо створення надлишкових вузлів ГСМАС у випадку, коли розмірності вхідних векторів у розв'язуваній задачі не дорівнюють ступеням двійки й, відповідно, не збігаються із числом входів НСМАС. Для того щоб обминути обмеження, що накладає на розмірності вхідних векторів розв'язуваних задач в роботі запропоновано надлишкові вхідні змінні мережі НСМАС дорівнювати нулю. Але при цьому, за рахунок наявності надлишкових вузлів ГСМАС, значно збільшується об'єм

необхідної пам'яті, що у свою чергу негативно впливає на швидкість і збіжність процесу навчання НСМАС.

У зв'язку із цим актуальним є розв'язання задачі пошуку мінімальної структури нейронної мережі НСМАС, що самоорганізується, яка має таку ж розмірність вхідного простору, як і розв'язувана нею задача. Для одержання мінімальної структури мережі СМАС, що самоорганізується, насамперед коротко проаналізуємо архітектуру мережі НСМАС і підхід до самоорганізації вхідного простору.

На рис. 3.12 наведено найменшу топологію нейронної мережі НСМАС, у якій кожна мережа ГСМАС містить у собі два вхідних параметри, а значення на виході ГСМАС другого рівня служать як вхідні значення для ГСМАС першого рівня.



Тут використовуються такі позначення: x_i ($i=1, 2, \dots, N$) – i -й вхід нейронної мережі НСМАС; y_j^l ($j=1, 2, \dots, n; l=1, 2, \dots, L$) – вихід j -ї ГСМАС шару l ; n – кількість мереж ГСМАС у шарі l ; y – вихідний сигнал НСМАС для вхідного сигналу x .

На рис. 3.13 наведено топологію НСМАС, яка враховує, що кожна мережа ГСМАС містить два входи, а вихідний сигнал ГСМАС першого шару є вхідним сигналом для ГСМАС другого шару й т.п. Завдяки тому, що мережа НСМАС використовує топологію повного бінарного дерева, то, наприклад, задачі класифікації образів із трьома й чотирма вхідними ознаками будуть вирішуватися за допомогою НСМАС однакових архітектур. У цьому випадку, при розв'язанні задачі із трьома вхідними вимірами буде створений один надлишковий вузол ГСМАС ($ГСМАС_3$), тим самим займаючи більший обсяг пам'яті.

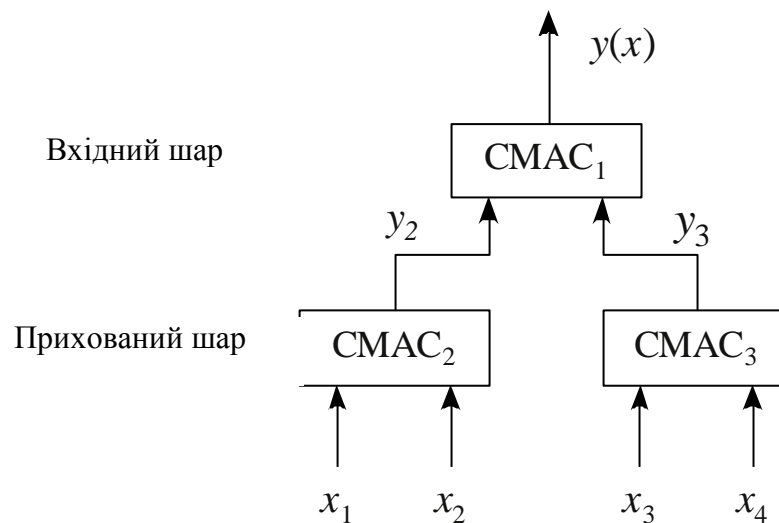


Рисунок 3.13 – Топологічна структура ШНМ НСМАС

Очевидно, що витрати пам'яті значно зростуть при розв'язанні задач із більшими розмірностями вхідного простору, відмінними від ступеня двійки.

Зокрема, розв'язання задачі класифікації образів з 513 вхідними ознаками призведе до створення структури з 1024 вхідними змінними, тим самим генеруючи 511 надлишкових вузлів GCMAC.

Для виводу правила навчання мережі HCMAC розглянемо представлену на рис. 3.14 детальну архітектуру мережі HGCMAC (Hierarchical Gaussian CMAC). Тут як БФ застосовуються диференційовувані гавсівські функції. Відповідно до рис. 3.14, у процесі навчання HGCMAC мають настроюватися три види параметрів для кожної вхідної до неї двовимірної GCMAC: ваги w , радіуси σ й центри m .

При виборі диференційовуваних БФ для навчання розглянутої мережі (настроювання її параметрів) застосуємо градієнтний алгоритм зворотного розповсюдження помилки. У цьому випадку спочатку провадиться настроювання параметрів GCMAC вихідного шару L , після чого помилка розповсюджується у зворотному напрямку від L -го шару до $(L-1)$ -го прихованого шару для настроювання параметрів GCMAC $(L-1)$ -го шару й так далі.

Якщо як функціонал, що мінімізують, обрано квадратичний функціонал помилки

$$E = \frac{1}{2} \sum_{i=1}^N (\hat{y}_i - y_i)^2, \quad (3.50)$$

де \hat{y}_i й y_i – необхідний й реальний вихідні сигнали мережі HCMAC для вхідного сигналу x відповідно, то узагальнена процедура навчання нейронної мережі можна подати так:

1) Настроювання GCMAC вихідного шару. Настроювання параметрів мережі GCMAC вихідного шару практично нічим не відрізняється від настроювання звичайної двовходової мережі CMAC, за винятком того, що вхідними сигналами для GCMAC вихідного шару є виходи мереж GCMAC

попереднього шару. Настойка здійснюється за допомогою будь-якого алгоритму навчання мережі СМАС.

2) Настроювання мереж GСМАС прихованих шарів. Для виводу алгоритма навчання мереж GСМАС у прихованому шарі з використанням алгоритма зворотного розповсюдження помилки в першу чергу необхідно обчислити

частинні похідні $\frac{\partial y_{h/2}^{l+1}}{\partial y_h^l}$ за допомогою виразу

$$\frac{\partial y_{h/2}^{l+1}}{\partial y_h^l} = \sum_{i=1}^p a_i w_{hi}^l \cdot \left[\prod_{j=2}^{h/2} \Phi_j^l \right] \cdot \frac{\partial \Phi_i^l}{\partial y_h^l}, \quad (3.51)$$

де $h=1, 2, \dots, n$ – номер мережі СМАС у поточному шарі; а $\lceil \cdot \rceil$ – позначає округлення до найближчого більшого цілого числа.

Далі запишемо алгоритм навчання мереж GСМАС у прихованому шарі

$$w_{hi}^l \leftarrow w_{hi}^l + \gamma \cdot \left[y_{h/2}^{l+1} - y_{h/2}^l \right] \cdot \frac{\partial y_{h/2}^{l+1}}{\partial y_h^l} \cdot \Phi_i^l \cdot a_i, \quad (3.52)$$

де w_{hi}^l – значення ваги в i -ій комірці пам'яті h -ої мережі СМАС шару l ; γ – коефіцієнт навчання.

Приведена вище процедура повторюється доти, поки не буде досягнуто задане значення критерію помилки (3.50), або задане максимальне число ітерацій навчання.

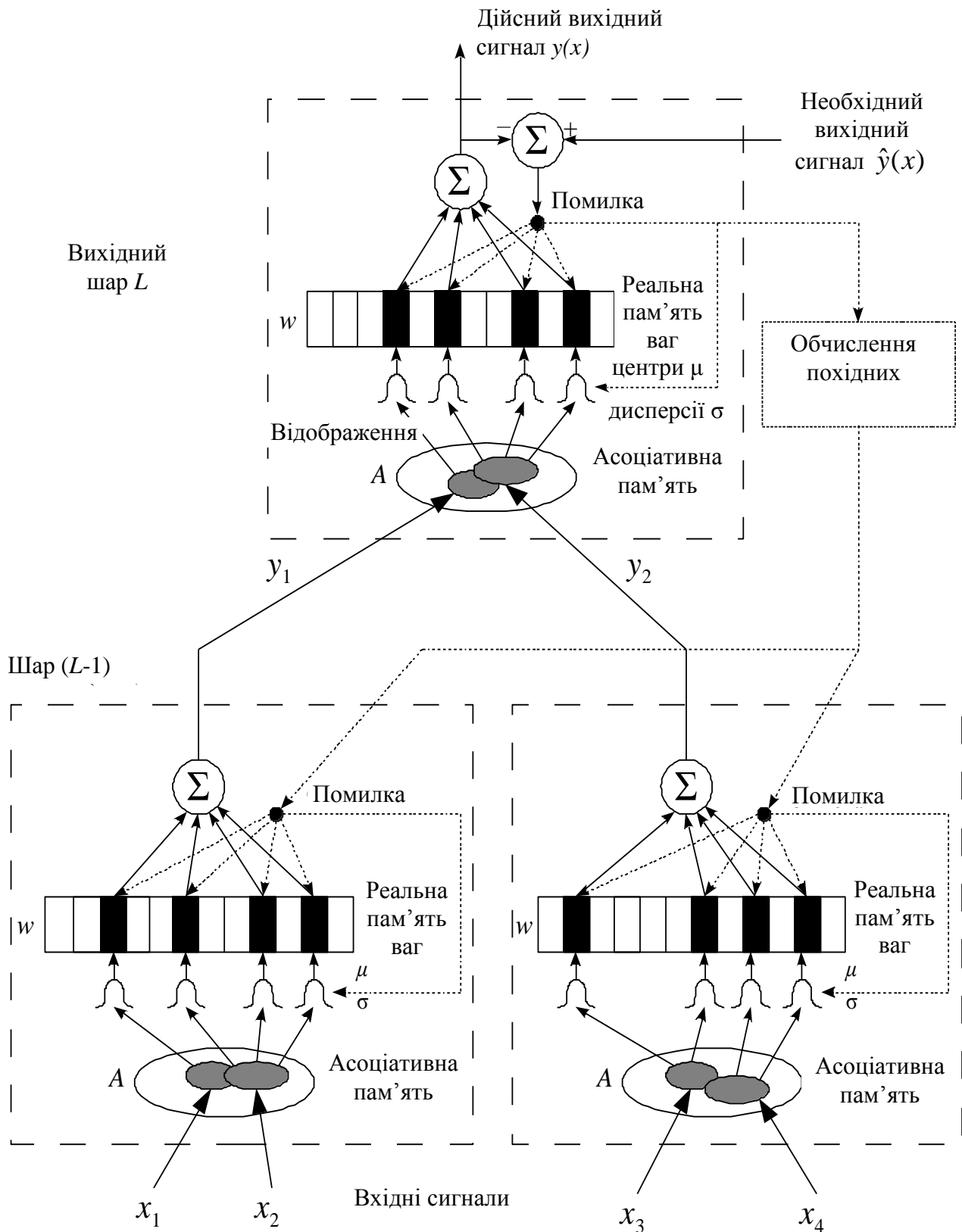


Рисунок 3.14 – Докладна архітектура ШНМ НСМАС

3.9 Мінімізація структури мережі НСМАС

3.9.1 Архітектура МНСМАС

З метою усунення перевитрати пам'яті й поліпшення характеристик навчання НСМАС, пропонується мережа МНСМАС, що є НСМАС мінімальної конфігурації з розмірністю вхідного вектора, рівною числу вхідних змінних розв'язуваної прикладної задачі. Приводиться опис структури нейронної мережі МНСМАС, механізму самоорганізації вхідного простору, порівняння МНСМАС зі звичайною мережею, що самоорганізується, НСМАС і докладний опис алгоритму навчання нейронної мережі МНСМАС.

Нейронна мережа, що самоорганізується, МНСМАС, як і НСМАС, є ієрархічною структурою, що складається із двовимірних диференційовуваних мереж ГСМАС, але може розширюватися довільним образом, використовуючи для масштабування топологію точного бінарного дерева. Таким чином, МНСМАС має ту ж розмірність вхідного простору, що й розв'язувана прикладна задача, тобто не містить надлишкових входів і вузлів.

На рис. 3.15 наведено кілька нейронних мереж МНСМАС з різними значеннями розмірностей вхідних векторів.

Тут x_i позначає задану вхідну послідовність ознак. Докладний опис алгоритму автоматичного призначення вхідних ознак входам нейронної мережі МНСМАС наведено далі.

Очевидно, що структура нейронної мережі МНСМАС буде збігатися зі структурою НСМАС у випадку, коли вхідні вектори розв'язуваної задачі мають у точності 2^n вимірів, де $n \in Z$. Навпроти, у випадку, коли вхідні вектори мають розмірність відмінну від 2^n , структура мережі МНСМАС буде відрізнятися від структури НСМАС.

Топології нейронних мереж МНСМАС, наведені на рис. 3.15, дозволяють вивести загальне правило автоматичного визначення структури МНСМАС. При

побудові структури МНСМАС у першому шарі створюється кореневий вузол GСМАС, позначуваний як $GСМАС_1$. Потім, у наступному шарі створюються другий і третій вузли GСМАС, позначувані відповідно як $GСМАС_2$ і $GСМАС_3$, і так далі. У структурі МНСМАС кожному вузлу GСМАС присвоюються послідовні номери, які можуть використатися для визначення зв'язків між батьківськими й дочірніми вузлами GСМАС у структурі МНСМАС. Кореневий вузол GСМАС ($GСМАС_1$) перебуває у вихідному шарі L , тоді як інші вузли розташовані в прихованих шарах $\llcorner - 1 \lrcorner$, $\llcorner - 2 \lrcorner$ і т.д.

Кількість вузлів GСМАС, необхідна для побудови структури одиночної МНСМАС, обчислюється на основі числа вимірів вхідного вектора відповідно до виразу

$$N_{GСМАСs} = N - 1, \quad (3.53)$$

де $N_{GСМАСs}$ – кількість вузлів GСМАС, необхідна для побудови повної структури нейронної мережі МНСМАС із числом входів N .

Для розв'язання задач класифікації з множиною вихідних категорій сумарна кількість вузлів GСМАС, необхідна для побудови повної структури МНСМАС, обчислюється як

$$N_{GСМАСs} = N_y \cdot \llcorner - 1 \lrcorner, \quad (3.54)$$

де N_y – кількість вихідних категорій у розв'язуваній задачі.

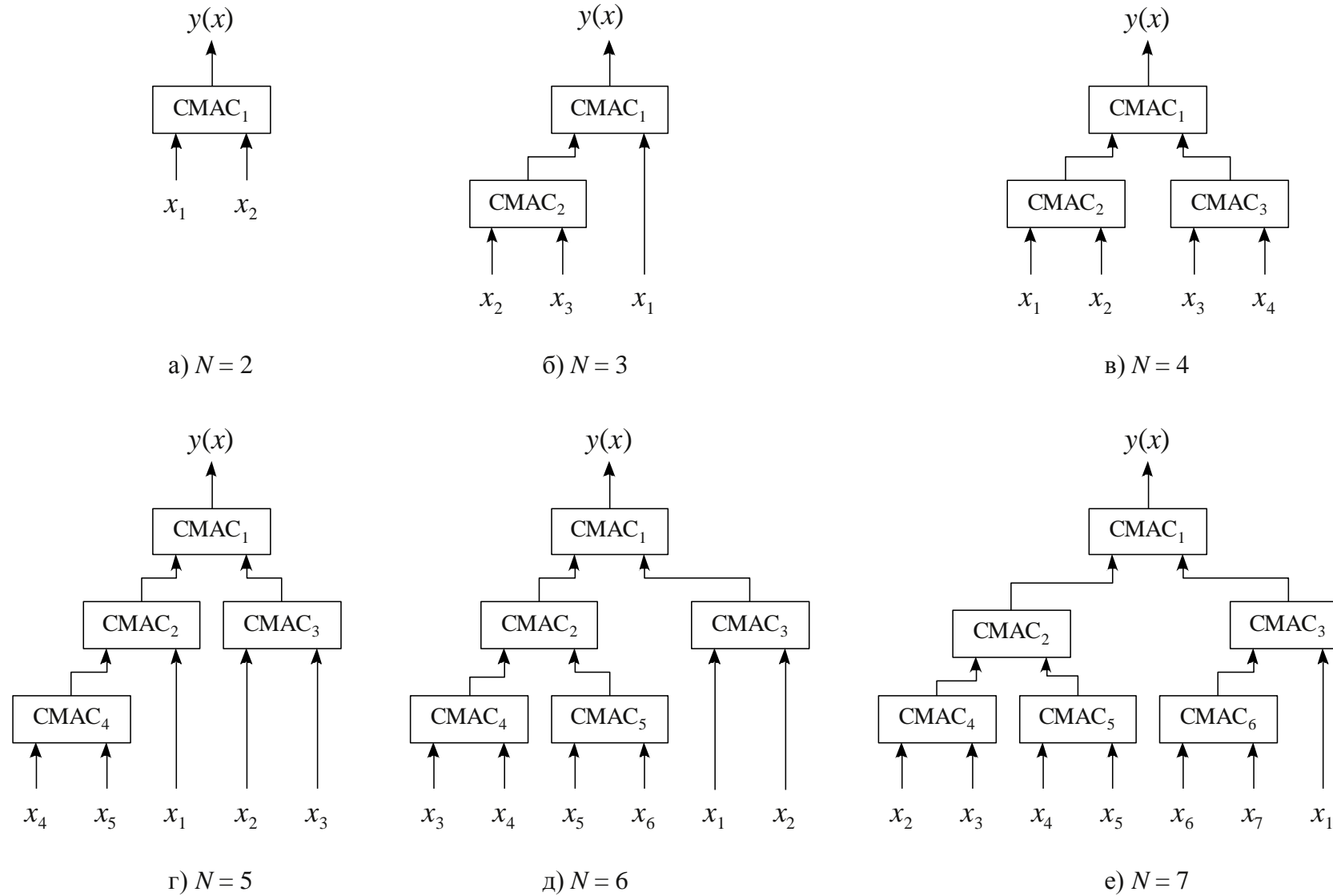


Рисунок 3.15 – Мінімальна структура ШНМ НСМАС з різною кількістю вхідних змінних

Батьківський вузол i -ої GCMAC у структурі нейронної мережі МНСМАС можна визначити, використовуючи вираз

$$P_{GCMAC_i} = \lfloor i/2 \rfloor, \quad (3.55)$$

де P_{GCMAC_i} – номер вузла, що є батьківським для вузла $GCMAC_i$, $i = 2, 3, 4, \dots, N_{GCMACs}$, $\lfloor \bullet \rfloor$ – округлення до найменшого найближчого цілого.

Дочірній вузол i -ої GCMAC обчислюється як

$$C_{GCMAC_i} = i \bmod 2, \quad (3.56)$$

де $C_{GCMAC_i} = 0$ позначає, що вузол $GCMAC_i$ є лівим дочірнім вузлом для свого батьківського вузла; а $C_{GCMAC_i} = 1$ виходить, що вузол $GCMAC_i$ є правим дочірнім вузлом для свого батьківського вузла.

Номер прихованого шару структури МНСМАС, у якому знаходиться вузол $GCMAC_i$, можна обчислити за формулою

$$L_{GCMAC_i} = L - \lfloor \log_2(i) \rfloor, \quad (3.57)$$

де L_{GCMAC_i} – номер шару, що містить вузол $GCMAC_i$, $i = 2, 3, 4, \dots, N_{GCMACs}$.

У запропонованій структурі МНСМАС у випадку, коли число вхідних змінних у розв'язуваній задачі є непарним числом, структура мережі МНСМАС буде мати один вузол GCMAC, на один із входів якого подається вихід його дочірнього вузла GCMAC, а на іншій - безпосередньо одна із вхідних змінних. Визначити номер цього вузла в структурі МНСМАС можна, використовуючи вираз

$$N_U = \frac{N_{GCMACs}}{2}, \quad (3.58)$$

де N_U – номер вузла GCMAC, на входи якого буде подаватися вихід його дочірнього вузла GCMAC і вхідна змінна.

Відмітимо, що відповідно до формули (3.53), у задачах з непарним числом вхідних змінних N_{GCMACs} завжди буде числом парним.

Якщо ж розв'язувана задача має у своєму розпорядженні парне число вимірів, тоді лівий і правий входи кожного задіяного вузла GCMAC на нижньому шарі відповідають вхідним змінним розв'язуваної задачі.

3.9.2 Порівняння структур НСМАС і МНСМАС

Таблиця 3.2 дозволяє порівняти запропоновану ШНМ НСМАС, з розглянутою ШНМ МНСМАС, що само організується, з погляду витрат пам'яті, топологічної структури й підходу до присвоєння вхідних змінних розв'язуваної прикладної задачі входам нейронної мережі.

У таблиці 3.2 прийняті такі позначення: N – кількість вхідних ознак, N_y – число вихідних категорій розв'язуваної задачі, а $\lceil \bullet \rceil$ позначає округлення убік найближчого більшого цілого.

Так як структура мережі НСМАС, що самоорганізується, розкривається на основі топології повного бінарного дерева незалежно від числа вхідних змінних, в НСМАС при розмірності вхідних векторів, відмінної від ступеня двійки, неминуче виникають надлишкові листові вузли GCMAC і надлишкові входи, на які подаються константні нульові значення. У свою чергу, МНСМАС позбавлена зазначеного недоліку. Завдяки використанню топології точного бінарного дерева витрати пам'яті для мережі МНСМАС зростають лінійно з

ростом кількості вхідних змінних, у той час як витрати пам'яті НСМАС ростуть зі ступенем двійки граничного логарифма кількості вхідних вимірів.

Таблиця 3.2 – Порівняння ШНМ НСМАС і МНСМАС

Моделі Елементи	НСМАС	МНСМАС
Загальна кількість вузлів ГСМАС, необхідна для побудови топологічної структури мережі	$N_y \times \lceil \log_2 N - 1 \rceil$	$N_y \times \lceil V - 1 \rceil$
Топологічна структура	Структура повного бінарного дерева	Структура точного бінарного дерева
Алгоритм призначення вхідних змінних розв'язуваної задачі входам ШНМ	Довільний порядок призначення ознак входам листових вузлів ГСМАС	Довільний порядок призначення, або пропонується метод автоматичного присвоєння вхідних ознак
Алгоритм призначення входів надлишкових листових вузлів ГСМАС	Присвоєння нульових значень входам надлишкових вузлів ГСМАС	Відсутність надлишкових листових вузлів ГСМАС

Вхідні змінні розв'язуваної прикладної задачі можуть присвоюватися довільним входам листових вузлів ГСМАС як у що самоорганізується НСМАС, так і в МНСМАС. Проте, з метою спрощення апаратної реалізації МНСМАС у цьому розділі пропонується метод автоматичного присвоєння вхідних змінних відповідним їм входам і вузлам ГСМАС у топологічній структурі МНСМАС.

3.9.3 Самоорганізація вхідного простору МНСМАС

Формалізація порядку призначення вхідних змінних розв'язуваної задачі входам мережі МНСМАС має вирішальне значення при її практичній реалізації, зокрема, для розробки комп'ютерної програми.

Загальне правило автоматичного присвоєння вхідних ознак вхідним змінним і створення структури МНСМАС виводиться на основі аналізу властивостей набору вхідних змінних. Для виводу загального правила самоорганізації вхідного простору мережі МНСМАС скористаємося прикладами структур, наведеними на рис. 3.15.

Введемо параметри K_j й M_j , які визначають номер вузла ГСМАС і вхід, якому буде привласнена j -а вхідна змінна x_j . Значення K_j й M_j визначаються відповідно за допомогою наступних виразів

$$K_j = \left\lceil \frac{j + N_{ГСМАС}}{2} \right\rceil, \quad (3.59)$$

$$M_j = \left\lfloor \frac{j + N_{ГСМАС}}{2} \right\rfloor \bmod 2, \quad (3.60)$$

де K_j вказує на те, що j -а вхідна змінна x_j присвоюється лівому входу K_j -го вузла $ГСМАС_{K_j}$ в структурі МНСМАС у випадку, якщо $M_j = 0$, і, відповідно, правому входу вузла $ГСМАС_{K_j}$, якщо $M_j = 1$.

Наведене правило є універсальним для побудови нейронних мереж МНСМАС, тому що з його допомогою можна автоматично призначати вхідні змінні розв'язуваної задачі вхідним сигналам МНСМАС як для парних розмірностей вхідних векторів, так і для непарних.

Тому що класичні ШНМ СМАС використовують рівномірну схему дискретизації вхідного простору сигналів, вони не здатні точно відобразити

фактичний розподіл навчальних даних, призводячи тим самим до додаткових витрат пам'яті й зниженню продуктивності навчання нейронної мережі.

У МНСМАС для усунення зазначеного недоліку шляхом адаптивного визначення оптимальної схеми дискретизації для кожного виміру вхідного простору запропонований метод вхідних просторів, що самоорганізуються, заснований на аналізі міри ентропії Шеннона набору навчальних даних і пошуку золотим перерізом. Оскільки вхідний простір кожної двовимірної мережі GCMAC дискретизується в кінцевий набір інтервалів, то обчислення міри ентропії Шеннона для кожного інтервалу дозволяє нам одержати інформацію про розподіл вхідних образів. Тому що точки адаптивного поділу невідомі, те необхідна ефективна стратегія пошуку оптимального розподілу простору вхідних сигналів. Для розв'язання цієї задачі пропонується застосувати метод пошуку золотим перерізом для одержання точок адаптивного поділу вхідного простору.

Як альтернатива також може бути використаний запропонований вище метод адаптивного настроювання схеми дискретизації, заснований на аналізі подібності Грея. На підставі інформації про схему дискретизації простору вхідних сигналів автоматично будуються вектори асоціативної адресації й фактична структура пам'яті МНСМАС.

У загальному вигляді алгоритм автоматичної побудови структури нейронної мережі, що самоорганізується, МНСМАС можна представити так:

1) визначення схеми дискретизації вхідного простору сигналів за допомогою запропонованого методу вхідних просторів, що самоорганізуються, заснованого на аналізі міри ентропії Шеннона набору навчальних даних, або метод, заснований на аналізі подібності Грея;

2) автоматична побудова структури МНСМАС із використанням ознак K_j і M_j , що обчислюють за формулами (3.59) і (3.60).

Слід зазначити, що методи самоорганізації вхідного простору застосовуються тільки для листових вузлів GCMAC, що знаходяться у вхідному

шарі, тому що в цьому випадку вони ґрунтуються на апіорно відомій інформації про розподіл вхідних навчальних даних. Для інших вузлів GСМАС, розташованих у прихованих і вихідних шарах структури МНСМАС, використовується схема рівномірної дискретизації вхідного простору, тому що розподіл вхідних сигналів вузлів GСМАС у прихованих шарах постійно змінюється в процесі навчання, і інформація про кінцевий розподіл може бути отримана лише по закінченню процесу навчання нейронної мережі.

3.9.4 Правило навчання ШНМ МНСМАС

Для спрощення виводу правила навчання мережі МНСМАС розглянемо приклад навчання МНСМАС із п'ятьма вхідними змінними, топологічна структура якої наведена на рис. 3.16.

Дана мережа має п'ять вхідних сигнали, позначені як x_i ($i = 1, 2, \dots, 5$), y_j ($j = 1, 2, 3, 4$) є виходом j -го вузла GСМАС $_j$, при цьому y_2 , y_3 і y_4 – виходи вузлів GСМАС прихованого шару, а y_1 – вихід МНСМАС. Крім того, для узагальнення й спрощення процедури виводу навчальних правил, вхідні ознаки x_i ($i = 1, 2, \dots, 5$) також позначимо як y_j , де $j = 5, 6, \dots, 9$. Виведемо навчальне правило на підставі методу градієнтного спуску й властивості диференційовності вузлових GСМАС.

Тому що GСМАС $_1$ L -го шару є вихідним вузлом, і його вихідне значення є виходом мережі МНСМАС, те для настроювання вагових коефіцієнтів GСМАС $_1$ застосуємо правило навчання звичайної нейронної мережі GСМАС. У якості функції вартості, що мінімізують, виберемо квадратичний функціонал помилки (3.50).

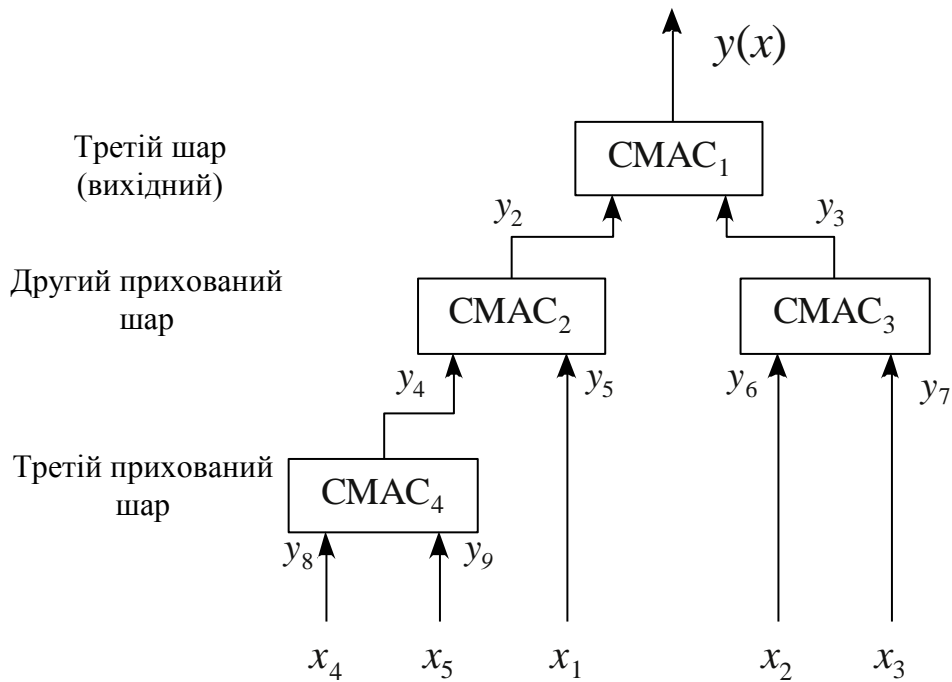


Рисунок 3.16 – Топологічна структура МНСМАС з п'ятьма вхідними змінними

У кожному вузлі $GCMAC$ налаштуванню підлягають три параметри, а саме: ваги w , радіуси σ й центри БФ t . Якщо на якій-небудь ітерації процесу навчання фактичні й шукане вихідні значення мережі МНСМАС різні, то в першу чергу налаштовуються параметри $GCMAC_1$ у вихідному шарі. Далі на підставі інформації про похідні, помилка розповсюджується зворотно по дереву МНСМАС від першого шару до другого, третього й наступного шарів для корекції параметрів $GCMAC_2$, $GCMAC_3$ і $GCMAC_4$.

Алгоритм навчання для наведеної на рис. 3.16 п'ятивходової нейронної мережі МНСМАС у загальному вигляді може бути представлений у такий спосіб.

Крок 1 – Корекція $GCMAC_1$ вихідного шару

Вагові коефіцієнти w_j в пам'яті $GCMAC_1$ налаштовуються по формулі

$$w_j \leftarrow w_j \leftarrow \frac{\alpha}{Ne} \frac{\partial E}{\partial w_j} = \frac{\alpha}{Ne} \cdot \left[\prod_{i=2}^3 e^{-\left(\frac{y_i - m_{jk}}{\sigma_{jk}}\right)^2} \right], \quad (3.61)$$

де Δw_j – це значення виправлення до значення j -го вагового елемента w_j в пам'яті $GCMAC_1$; m_{jk} – центр k -го виміру гавсівської БФ в j -му відображеному гіперкубі, k визначається за допомогою i : якщо $i=2$, тоді $k=1$, якщо $i=3$, то $k=2$; σ_{jk} – це радіус k -го виміру гавсівської БФ в j -му відображеному гіперкубі; α – коефіцієнт швидкості навчання, $a_j \leftarrow$ є j -м елементом вектора асоціативного вибору пам'яті для окремого вхідного стану x ; а Ne – це номер відображеного гіперкуба для вхідного стану x .

Настроювання радіусів БФ здійснюються в такий спосіб

$$\begin{aligned} \Delta \sigma_{ji} &= -\frac{\alpha}{Ne} \cdot \frac{\partial E}{\partial \sigma_{ji}} = \\ &= \frac{\alpha}{Ne} \cdot \left[\prod_{i=2}^3 e^{-\left(\frac{y_i - m_{jk}}{\sigma_{jk}}\right)^2} \right] \cdot \frac{2(y_i - m_{jk})^2}{\sigma_{jk}^3}, \end{aligned} \quad (3.62)$$

де $\Delta \sigma_{ji}$ – коригувальне виправлення до i -го виміру радіуса σ_{ji} БФ для j -го відображеного гіперкуба вхідного стану x в $GCMAC_1$

У свою чергу, центри m_{jk} БФ $GCMAC_1$ настроюються відповідно до правила

$$\begin{aligned} \Delta m_{ji} &= -\frac{\alpha}{Ne} \cdot \frac{\partial E}{\partial m_{ji}} = \\ &= \frac{\alpha}{Ne} \cdot \left[\prod_{i=2}^3 e^{-\left(\frac{y_i - m_{jk}}{\sigma_{jk}}\right)^2} \right] \cdot \frac{2(y_i - m_{jk})}{\sigma_{jk}^2}, \end{aligned} \quad (3.63)$$

де Δm_{ji} – значення виправлення i -го виміру центра m_{ji} БФ $GCMAC_1$ для j -го відображеного гіперкуба для вхідного стану x .

Крок 2 – Корекція $GCMAC_2$ й $GCMAC_3$ перший прихований шар

Тому що інформація про оптимальні значення \hat{y}_2 й \hat{y}_3 не доступна, то для виводу навчальних правил настроювання параметрів $GCMAC_2$ і $GCMAC_3$ в першому прихованому шарі спочатку необхідно обчислити

похідні $\frac{\partial y_1}{\partial y_2}$ і $\frac{\partial y_1}{\partial y_3}$ як

$$\frac{\partial y_1}{\partial y_2} = \sum_{j=1}^{N_h} a_j \cdot w_j \cdot \left[\prod_{i=2}^3 e^{-\left(\frac{y_i - m_{jk}}{\sigma_{jk}}\right)^2} \right] \cdot \frac{-2(y_2 - m_{j1})}{\sigma_{j1}^2} \quad (3.64)$$

та

$$\frac{\partial y_1}{\partial y_3} = \sum_{j=1}^{N_h} a_j \cdot w_j \cdot \left[\prod_{i=2}^3 e^{-\left(\frac{y_i - m_{jk}}{\sigma_{jk}}\right)^2} \right] \cdot \frac{-2(y_3 - m_{j2})}{\sigma_{j2}^2} \quad (3.65)$$

відповідно.

У наведених вище рівняннях w_j позначає j -й ваговий коефіцієнт у пам'яті $GCMAC_1$; m_{jk} є центром k -го виміру БФ для j -го відображеного гіперкуба вхідного стану x в $GCMAC_1$; σ_{jk} – радіус k -го виміру БФ для j -го відображеного гіперкуба вхідного стану x в $GCMAC_1$; N_h – це спільний фактичний розмір пам'яті $GCMAC_1$; а k визначається за допомогою i : якщо $i=2$, то $k=1$, якщо $i=3$, то $k=2$.

Далі, правила настроювання ваг w_j , радіусів σ_{ji} і центрів m_{ji} БФ $GCMAC_2$ у першому прихованому шарі МНСМАС виводяться в такий спосіб

$$\begin{aligned} \Delta w_j &= -\frac{\alpha}{Ne} \cdot \frac{\partial E}{\partial w_j} = \\ &= \frac{\alpha}{Ne} \cdot \left[\prod_{i=1}^n y_i \right] \left(\frac{\partial y_1}{\partial y_2} \right) \cdot a_j \cdot \left[\prod_{i=4}^5 e^{-\left(\frac{y_i - m_{jk}}{\sigma_{jk}} \right)^2} \right] \end{aligned} \quad (3.66)$$

$$\begin{aligned} \Delta \sigma_{ji} &= -\frac{\alpha}{Ne} \cdot \frac{\partial E}{\partial \sigma_{ji}} = \\ &= \frac{\alpha}{Ne} \cdot \left[\prod_{i=1}^n y_i \right] \left(\frac{\partial y_1}{\partial y_2} \right) \cdot a_j \cdot w_j \cdot \left[\prod_{i=4}^5 e^{-\left(\frac{y_i - m_{jk}}{\sigma_{jk}} \right)^2} \right] \cdot \frac{2(y_i - m_{jk})}{\sigma_{jk}^3}, \end{aligned} \quad (3.67)$$

$$\begin{aligned} \Delta m_{ji} &= -\frac{\alpha}{Ne} \cdot \frac{\partial E}{\partial m_{ji}} = \\ &= \frac{\alpha}{Ne} \cdot \left[\prod_{i=1}^n y_i \right] \left(\frac{\partial y_1}{\partial y_2} \right) \cdot a_j \cdot w_j \cdot \left[\prod_{i=4}^5 e^{-\left(\frac{y_i - m_{jk}}{\sigma_{jk}} \right)^2} \right] \cdot \frac{2(y_i - m_{jk})}{\sigma_{jk}^2}. \end{aligned} \quad (3.68)$$

У рівностях (3.66), (3.67) і (3.68) використані символічні позначення аналогічні прийнятим у виразах (3.61), (3.62) і (3.63), ваги w_j , радіуси σ_{jk} й радіуси σ_{ji} належать вузлу $GCMAC_2$, а k може бути визначене за допомогою i так: якщо $i = 4$, то $k = 1$, якщо $i = 5$, то $k = 2$.

За аналогією на підставі обчислених похідних $\frac{\partial y_1}{\partial y_3}$ виводиться навчальне правило для настроювання $GCMAC_3$ в першому прихованому шарі $\mathcal{C} - 1$ МНСМАС.

Крок 3 – Корекція $GCMAC_4$ другого прихованого шару

Правило настроювання параметрів вузла $GCMAC_4$ другого прихованого шару $\mathcal{C} - 2$ аналогічно алгоритму настроювання вузлів $GCMAC_3$ і $GCMAC_3$

шаруючи $\mathcal{L} - 1$. Але тепер необхідна інформація про похідну $\frac{\partial y_1}{\partial y_4}$, що може бути знайдена за допомогою виразу

$$\frac{\partial y_1}{\partial y_4} = \frac{\partial y_1}{\partial y_2} \cdot \frac{\partial y_2}{\partial y_4}, \quad (3.69)$$

де

$$\frac{\partial y_2}{\partial y_4} = \sum_{j=1}^{N_h} a_j \cdot w_j \cdot \left[\prod_{i=4}^5 e^{-\left(\frac{y_i - m_{jk}}{\sigma_{jk}}\right)^2} \right] \cdot \frac{-2(y_4 - m_{j1})}{\sigma_{j1}^2}, \quad (3.70)$$

де, у свою чергу, w_j позначає j -й ваговий коефіцієнт у пам'яті $GCMAC_2$; m_{jk} – центр k -го виміру БФ для j -го відображеного гіперкуба вхідного стану x в $GCMAC_2$; σ_{jk} – радіус k -го виміру БФ для j -го відображеного гіперкуба вхідного стану x в $GCMAC_2$; N_h – спільний фактичний розмір пам'яті $GCMAC_2$; а k – індекс, що визначається за допомогою i : якщо $i = 4$, то $k = 1$, якщо $i = 5$, то $k = 2$.

Отримавши похідні $\frac{\partial y_1}{\partial y_2}$ й $\frac{\partial y_2}{\partial y_4}$, нескладно вивести формули для настроювання ваг w_j , радіусів σ_{ji} і центрів m_{ji} БФ $GCMAC_4$ у другому прихованому шарі $\mathcal{L} - 2$ структури МНСМАС. У цьому випадку градієнт помилки між шуканим і фактичним виходами розповсюджується у зворотному напрямку від виходу мережі МНСМАС y_1 до першого прихованого шару $\mathcal{L} - 1$ згідно виразу (3.64), після чого аналогічним образом помилка розповсюджується до другого прихованого шару $\mathcal{L} - 2$ відповідно до виразів (3.69) і (3.70). У результаті всіх підстановок виразу для настроювання параметрів вузла $GCMAC_4$ приймають вигляд

$$\begin{aligned} \Delta w_j &= -\frac{\alpha}{Ne} \cdot \frac{\partial E}{\partial w_j} = \\ &= \frac{\alpha}{Ne} \cdot \left[\sum_{i=1}^5 y_i \left(\frac{\partial y_1}{\partial y_2} \right) \cdot \left(\frac{\partial y_2}{\partial y_4} \right) \cdot a_j \cdot \left[\prod_{i=4}^5 e^{-\left(\frac{y_i - m_{jk}}{\sigma_{jk}} \right)^2} \right] \right], \end{aligned} \quad (3.71)$$

$$\begin{aligned} \Delta \sigma_{ji} &= -\frac{\alpha}{Ne} \cdot \frac{\partial E}{\partial \sigma_{ji}} = \frac{\alpha}{Ne} \cdot \left[\sum_{i=1}^5 y_i \left(\frac{\partial y_1}{\partial y_2} \right) \cdot \left(\frac{\partial y_2}{\partial y_4} \right) \times \right. \\ &\quad \left. \times a_j \cdot w_j \cdot \left[\prod_{i=4}^5 e^{-\left(\frac{y_i - m_{jk}}{\sigma_{jk}} \right)^2} \right] \cdot \frac{2(y_i - m_{jk})}{\sigma_{jk}^3} \right], \end{aligned} \quad (3.72)$$

$$\begin{aligned} \Delta m_{ji} &= -\frac{\alpha}{Ne} \cdot \frac{\partial E}{\partial m_{ji}} = \frac{\alpha}{Ne} \cdot \left[\sum_{i=1}^5 y_i \left(\frac{\partial y_1}{\partial y_2} \right) \cdot \left(\frac{\partial y_2}{\partial y_4} \right) \times \right. \\ &\quad \left. \times a_j \cdot w_j \cdot \left[\prod_{i=4}^5 e^{-\left(\frac{y_i - m_{jk}}{\sigma_{jk}} \right)^2} \right] \cdot \frac{2(y_i - m_{jk})}{\sigma_{jk}^2} \right]. \end{aligned} \quad (3.73)$$

У рівняннях (3.71), (3.72) і (3.73) використані символічні позначення, подібні прийнятим у виразах (3.61), (3.62), (3.63), з тією лише відмінністю, що ваги w_j , радіуси σ_{jk} й радіуси σ_{ji} належать вузлу $GCMAC_4$, а k визначається за допомогою i : якщо $i = 8$, то $k = 1$, якщо $i = 9$, то $k = 2$.

Отже, вирази (3.66), (3.67) і (3.68) показують, що помилка передається зворотно з $GCMAC_1$ у вихідному шарі до вузлів $GCMAC$ першого прихованого шару відповідно до інформації про похідні $\frac{\partial y_1}{\partial y_2}$ і $\frac{\partial y_1}{\partial y_3}$. І далі, рівняння (3.71), (3.72) і (3.73) показують, що помилка також розповсюджується назад з $GCMAC_1$ у вихідному шарі до вузлів $GCMAC$ другого прихованого шару відповідно до обчислених похідними $\frac{\partial y_1}{\partial y_2}$ і $\frac{\partial y_1}{\partial y_3}$. Отже, представляється можливим поширити це правило навчання на довільну мережу МНСМАС, що має будь-яку кількість вузлів і вхідних змінних.

3.10 Апаратна реалізація нейронної мережі СМАС: загальний підхід

Апроксимація функцій будь-якого вигляду нейронною мережею СМАС зводиться до виконання таких операцій:

- кодування змінних, в результаті якого формуються вектор асоціацій, що є адресою читання ваг з оперативної пам'яті, і вектор коефіцієнтів, – значень активаційних функцій вхідного шару нейронів, що беруть участь у формуванні чергового значення функції;
- хешування вектора асоціацій, яке служить для скорочення об'єму пам'яті, необхідного для зберігання ваг;
- вибірка значень ваг і обчислення значення функції;
- порівняння набутого значення функції з потрібним і коректування ваг за одним з алгоритмів навчання.

3.10.1 Кодування змінних

У цифровій системі обробки інформації змінні подаються у формі двійкових слів. Вважатимемо, що змінні x_i , від яких залежить функція y , мають однакову розрядність, рівну n . Вхідний шар мережі містить ρ ступенів квантування, таких, що для будь-якого ступеня існує інший ступінь, на якому області квантування зміщені щодо даного ступеня на один рівень квантування вхідної змінної x_i . Оскільки ця змінна є двійковим числом, таке зміщення відповідає одиниці молодшого розряду вхідної змінної. Привласнимо ступені квантування змінної x_i індекс $j_i = \overline{0, \rho - 1}$.

Областю квантування є нейрон, що має унікальний для даного ступеня номер $NNS_i^{j_i}$ (Neuron Number on Stair j_i , variable i) і функцію активації $\Phi(u_i^{j_i})$. Область визначення цієї функції $u_i^{j_i} \in \overline{0, \rho - 1}$, оскільки область квантування містить ρ рівнів квантування вхідної змінної. Приймемо, що всі нейрони вхідного шару мають функції активації одного вигляду. Аргументом цієї функції буде значення

$$u_i^{j_i} = (x_i + j_i) \bmod \rho, \quad (3.74)$$

де \bmod – функція обчислення залишку від цілочислового ділення.

Номер активного нейрона на ступені j_i можна визначити як

$$NNS_i^{j_i} = (x_i + j_i) \operatorname{div} \rho, \quad (3.75)$$

де div – функція цілочислового ділення.

Апаратна реалізація операції ділення на довільне число є або ресурсомісткою, або повільною. При виборі значень ρ , рівних 2^m , операцію ділення можна замінити операцією зсуву на m розрядів управо. Залишком від ділення, в цьому випадку, є m молодших розрядів вхідної змінної, які можуть використовуватися як аргумент активаційної функції.

У класичному варіанті мережі СМАС активаційні функції нейронів областей квантування мають прямокутну форму. Проте, як показують дослідження, кращий результат досягається під час використання функцій вищих порядків (лінійних, квадратичних, кубічних), а також тригонометричних функцій і функцій степені. Апаратна реалізація широкого спектра активаційних функцій (яка необхідна для структури, що моделює мережу СМАС) складна і ресурсоемна. Тому пропонується таблично задавати функцію активації в ОЗП, що містить 2^m комірок. Розрядність комірок ОЗП h визначається областю допустимих значень функції активації, а також необхідною точністю задання значень функції активації. Наприклад, якщо $m = 8$, то необхідне ОЗП місткістю 256 комірок. Блоки ОЗП такого об'єму є в більшості сучасних ПЛІС, наприклад, сімейства ACEX1K фірми ALTERA. Така реалізація функцій активації є простою і гнучкою, оскільки дозволяє задати функцію будь-якого вигляду.

Структуру блока кодування змінної x_i наведено на рис. 3.17. На вхід цього блока значення змінної поступають в паралельному коді з частотою f . Для кожного значення, що поступило, необхідно отримати ρ значень NNS і функцій активації. Тобто кодування має здійснюватися з частотою $f\rho$. Таку частоту можна отримати за допомогою схеми фазового автоматичного підстроювання частоти (ФАПЧ), що складається з генератора, керованого напругою (ГКН), фазового детектора (ФД), фільтра низьких частот (ФНЧ). Напруга, ГКН, що управляє, залежить від того, скільки періодів формованої частоти потрапило в один період опорної. Воно формується ФД, який порівнює фази опорної частоти f і сигналу переповнювання лічильника з модулем ρ , що підраховує імпульси частоти, яка синтезується. Напруга управління ГКН фільтрується за допомогою ФНЧ.

Код змінної x_i захищується в регістрі RGX, звідки поступає на суматор $\Sigma 1$. На другий вхід цього суматора поступає m -розрядне значення номера ступеня j , на якому в даний момент відбувається кодування. Це значення формується з частотою $f\rho$ пристроєм вибірки ступеня, який служить для усунення діагональної структури асоціативного поля, і, відповідно, поліпшення апроксимуючих властивостей мережі. З виходу суматора n -розрядний код поступає на логічний зрушувач (LSh – Logical Shifter), що дозволяє змінювати значення m . При фіксованому значенні m логічний зрушувач як такий зникає. Значення $u_i^{j_i}$ визначають m молодших розрядів вихідного коду LSh, а значення $NNS_i^{j_i}$ – відповідно старші $n - m$ розрядів. Значення $u_i^{j_i}$ поступає на адресні входи ОЗП зберігання активаційної функції. З виходу ОЗП h -розрядне значення функції активації поступає на регістр RgVFS, а з регістра – на вихід VFS (Value of Function on Stair) блока кодування. Значення $NNS_i^{j_i}$ поступає на регістр RgNNS, а звідти на вихід NNS (Number of Neuron on Stair) блока кодування. Тактові імпульси на регістри, а також сигнали управління ОЗП формуються з частотою $f\rho$ пристроєм управління, загальним для всієї структури. Мультиплексор MX_1 і буфер B_1 необхідні для ініціалізації ОЗП значеннями

активаційної функції. У разі фіксованого виду цієї функції необхідність в цих блоках відпадає.

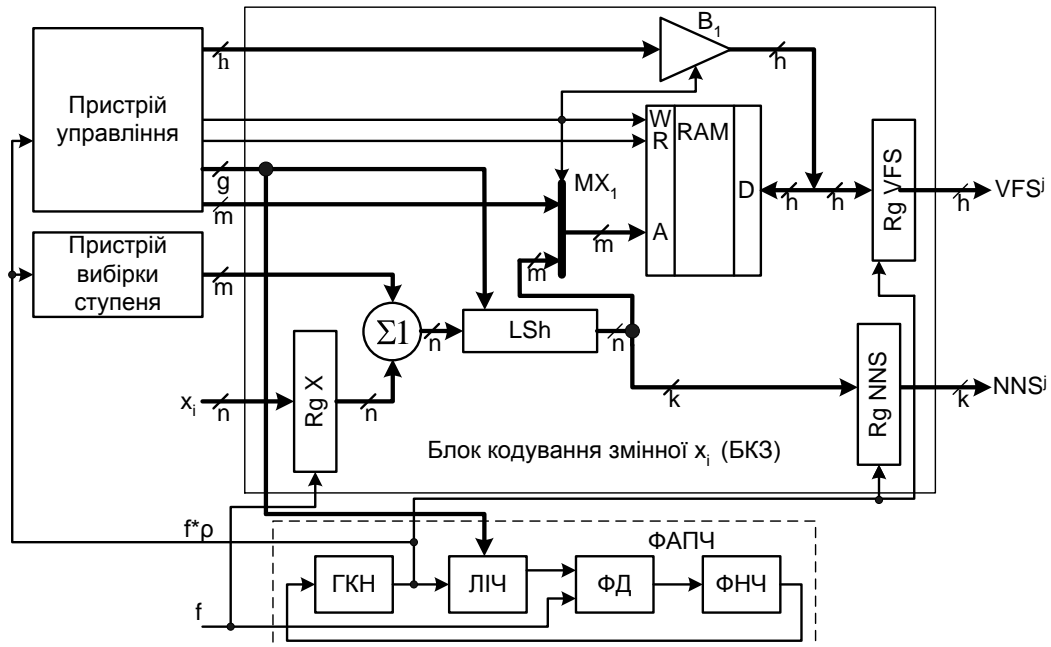


Рисунок 3.17 – Блок кодування змінної x_i

При малих значеннях m (4–6) можлива реалізація активаційної функції безпосередньо на логічних елементах ПЛІС, що є так званими пошуковими таблицями LUT (Look-up Table). Зокрема, для вже відміченого сімейства ПЛІС ACEX1K LUT дозволяє реалізувати будь-яку двійкову функцію від 4 двійкових змінних. Так, при $m=4$ і 8-розрядному поданні активаційної функції буде потрібно всього 8 комірок ПЛІС. Функція синтезується на етапі проектування спецпроцесора і її не можна змінити під час його роботи.

Блоки кодування змінних працюють паралельно і формують вектор асоціацій і вектор коефіцієнтів, компоненти яких послідовно розподілені в часі. Вектор асоціацій можна подати у вигляді

$$a_k = \begin{pmatrix} 0 : NNS_r^0 : \dots : NNS_2^0 : NNS_1^0, \\ 1 : NNS_r^1 : \dots : NNS_2^1 : NNS_1^1, \\ \dots, \\ (\rho - 1) : NNS_r^{\rho-1} : \dots : NNS_2^{\rho-1} : NNS_1^{\rho-1} \end{pmatrix}, \quad (3.76)$$

де k – номер вхідної комбінації змінних; r – кількість вхідних змінних. Залежність (3.76) відображує тільки ненульові компоненти вектора асоціацій і визначає, до яких елементів пам'яті ваг буде проведено звернення під час обчислення функції.

Елемент вектора асоціацій є комбінацією номера кроку кодування і номерів нейронів на ступенях, що беруть участь в кодуванні на даному кроці. Кількість кроків кодування рівна кількості ступенів ρ . Проте номер кроку не обов'язково відповідає номеру ступеня для i -ї змінної. Як вже було зазначено, для усунення діагональної структури асоціативного поля ступеня із зрушеними на один рівень квантування областями можуть знаходитися на послідовних кроках кодування тільки у однієї із змінних. Для цієї змінної номер кроку кодування рівний номеру ступеня j . Для решти змінних номери ступенів, які беруть участь в кодуванні на даному кроці, формує пристрій вибірки ступеня.

Розрядність q_{av} елемента вектора асоціацій рівна

$$q_{av} = m + r(n - m), \quad (3.77)$$

де m , n і r визначено вище.

Вважатимемо, що змінні подаються однаковою кількістю розрядів.

Значення q_{av} визначить об'єм пам'яті V_{wc} , необхідний для зберігання вагових коефіцієнтів мережі

$$V_{wc} = 2^{q_{av}}. \quad (3.78)$$

Елементи вектора коефіцієнтів формуються як добуток значень активаційних функцій для всіх змінних на даному кроці кодування. Весь вектор коефіцієнтів (Ψ_k) має вигляд

$$\Psi_k = \begin{pmatrix} \Phi(u_r^0) * \dots * \Phi(u_2^0) * \Phi(u_1^0), \\ \Phi(u_r^1) * \dots * \Phi(u_2^1) * \Phi(u_1^1), \\ \dots, \\ \Phi(u_r^{\rho-1}) * \dots * \Phi(u_2^{\rho-1}) * \Phi(u_1^{\rho-1}) \end{pmatrix}. \quad (3.79)$$

Розрядність елементів вектора коефіцієнтів на відміну від вектора асоціацій визначається не тільки розрядністю значень $\Phi(u_i^{j_i})$, яка рівна h . Це відбувається унаслідок того, що область допустимих значень активаційної функції нормується на інтервалі $[0..1)$. Тобто є дробовим числом. При цьому недоцільно використовувати розряд для цілої частини, і всі h розрядів відводяться під дробову частину значення активаційної функції. Добуток значень, що лежать в інтервалі $[0..1)$, також лежить в інтервалі $[0..1)$. Кількість розрядів q_{cv} , необхідне для подання цього здобутку, визначається бажаною точністю обчислень. Максимальна точність буде при $q_{cv} = r \cdot h$.

Під час використання мережі СМАС для обробки зображень, кількість вхідних змінних визначається або кількістю координат (дві) точки, або кількістю кольорних складових (три) точки в зображенні. Крім того, необхідність виконання ρ кроків при кодуванні для кожної точки збільшує необхідну для роботи тактову частоту. Так, при $\rho=32$ і частоті проходження точок у розгортці зображення $f=13,5$ МГц, значення робочої тактової частоти має складати $f \cdot \rho = 13,5 \cdot 32 = 432$ МГц. Ця величина є граничною для доступних на сьогоднішній

день ПЛІС, а, отже, розгляд значень ρ , більш ніж 32, позбавлено прикладного сенсу.

Для формування значень вектора коефіцієнтів необхідна реалізація операції множення, витрати на яку великі і пропорційні кількості множників. Водночас множники є вибраними з ОЗП значеннями базисної функції. Розрядність адрес для звернення до цього ОЗП рівна m . При $\rho = 32$, $m = 5$. Тобто ОЗП містить всього 32 комірки. Пропонується замінити обчислення добутку базисних функцій вибіркою значень цього добутку із загального для всіх змінних ОЗП, яке міститиме ρ^N комірок. При кількості змінних $N = 2$ і $\rho = 32$ об'єм ОЗП складе 1024 комірки, при $N = 3$ – 32768 комірок. ОЗП такого об'єму є в структурах приступних ПЛІС. Перевагами такої модифікації є:

1. Зниження апаратних витрат, пов'язане з усуненням помножувача і структур для ініціалізації ОЗП базисних функцій для кожної змінної.
2. Усунення похибок, пов'язаних з обмеженою розрядністю окремих ОЗП і помножувача.
3. Скорочення об'єму обчислень при кодуванні змінних.

3.10.2 Обчислення значення функції

Сформований в результаті кодування вектор асоціацій (3.8) за необхідності обробляється алгоритмом хешування, в результаті роботи якого формується вектор адрес a'_k вибірки ваг з пам'яті. В результаті вибірки з пам'яті отримаємо вектор ваг $w_k(a'_k)$, який спільно з вектором коефіцієнтів Ψ_k визначить значення функції для даної комбінації вхідних

$$y_k = w_k^T(a'_k)\Psi_k. \quad (3.80)$$

На рис. 3.18 наведено загальну структуру апаратної реалізації мережі СМАС, що містить дві вхідні змінні.

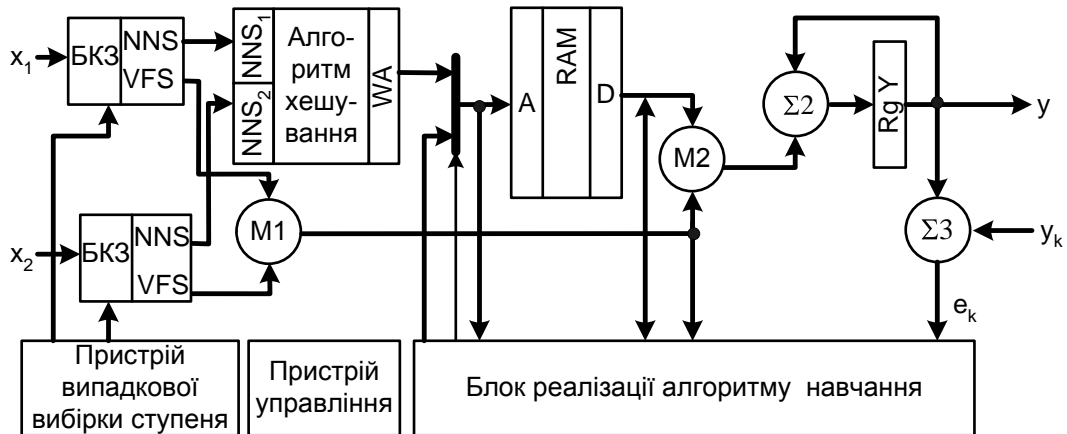


Рисунок 3.18 – Загальна структура апаратної реалізації мережі СМАС

Мережа обробляє дві вхідні змінні, кожна з яких кодується блоками БКП, що містять рівну кількість ступенів квантування. Функціонування такої структури полягає в такому. На j -му такті блоки БКЗ формують значення номерів нейронів у поточних ступенях (NNS_1 і NNS_2) і значення їх функцій активації (VFS_1 і VFS_2). Значення VFS поступають на помножувач $M1$, що формує елементи вектора коефіцієнтів. Значення цих елементів поступають на помножувач $M2$, що обчислює добуток ваги і відповідного йому коефіцієнта. Адреса ваги в ОЗП ваг визначається алгоритмом хешування, вхідним параметром якого є вектор асоціацій. Накопичуючий суматор $\Sigma 2$ спільно з помножувачем $M2$ обчислює (3.80) і після виконання ρ тактів у регістрі RgY формує значення функції, відповідне вхідним змінним. За допомогою суматора $\Sigma 3$ формується код помилки e_k , який визначається як різниця набутого значення функції y і потрібного y^* значень функції. Значення помилки поступає на блок реалізації алгоритму навчання, який під час ρ тактів обчислення накопичував адреси вибірки ваг, значення ваг і добутків функцій активації. Набувши значення помилки, алгоритм навчання модифікує ті, що брали участь в

обчисленні ваги і записує їх в ОЗП ваг за відповідними адресами. Це відбувається під час обчислення значення функції для наступної комбінації вхідних змінних. Пам'ять має підтримувати одночасне читання і запис за різними адресами. Крім того, необхідно враховувати ситуацію, коли для поточної комбінації вхідних змінних необхідна вага, яка була модифікована під час навчання на попередній комбінації. Координацію роботи вузлів мережі здійснює пристрій управління.

3.11 Апаратна реалізація мережі СМАС у разі послідовного перебору значень змінних

Головним недоліком розглянутого способу апаратної реалізації нейронної мережі СМАС є його застосовність лише для значень ρ , рівних степеням числа 2. Причина в необхідності виконання ресурсомісткої операції ділення. Проте для оптимальної апроксимації функцій складного вигляду, якими є функції присутності, а також самі зображення при їх стисненні, необхідно мати можливість задавати значення ρ , відмінні від степенів числа 2.

Розглянемо реалізацію мережі СМАС для завдань синтезу функцій присутності і стиснення зображень. Відмітною особливістю таких завдань є той факт, що значення змінних, відповідних координатам точок зображення на екрані, перебираються послідовно. Це відбувається унаслідок відрядкового розгортання зображення на екрані. Колір зображення може задаватися в різних колориметричних системах і визначається сукупністю складових (червоною, зеленою і синьою в системі RGB, яскравістю і кольорорізницевами складовими в системі YUV). Ці складові, у свою чергу, в загальному вигляді можуть подаватися у вигляді функцій $I(x,y)$ від координат точки на екрані x і y . При цьому координата x задається у вигляді тактової частоти f виведення точок в рядку, а координата y – у вигляді рядкових синхронізуючих імпульсів H-Sync. Функція присутності має такий же вигляд і розгортається аналогічно зображенню.

Розглянемо процес кодування змінних при їх послідовному переборі. У основі міркувань лежить той факт, що

$$\begin{aligned}x_{k,p} &= k, \\ y_{k,p} &= p.\end{aligned}\tag{3.81}$$

У зв'язку з тим, що дана функція залежить від двох змінних, приймемо двоіндексне позначення комбінацій цих змінних. При цьому індекс k відповідає за зміну змінних уздовж рядка з номером p . Метою аналізу є виявлення закономірностей при кодуванні змінних і визначенні змінної, для якої вибір номера ступеня, рівного номера кроку кодування, дасть кращий результат з огляду мінімуму апаратних витрат.

Обчислимо значення $u_x^{j_x}$, $NNS_x^{j_x}$ і $u_y^{j_y}$, $NNS_y^{j_y}$ відповідно на різних кроках кодування для різних значень k і p . Результати зведемо в табл. 3.3 для випадків рівності номера кроку кодування і номера ступеня кодування для змінної x і y . Приймемо $\rho = 5$.

Таблиця 3.3 – Розрахунок значень $u_x^{j_x}$, $NNS_x^{j_x}$ і $u_y^{j_y}$, $NNS_y^{j_y}$

(k,p)	j	$X(k,p)$			$Y(k,p)$			$X(k,p)$			$Y(k,p)$		
		j_x	$u_x^{j_x}$	$NNS_x^{j_x}$	j_y	$u_y^{j_y}$	$NNS_y^{j_y}$	j_x	$u_x^{j_x}$	$NNS_x^{j_x}$	j_y	$u_y^{j_y}$	$NNS_y^{j_y}$
(0,0)	0	0	0	0	1	1	0	1	1	0	0	0	0
	1	1	1	0	4	4	0	4	4	0	1	1	0
	2	2	2	0	2	2	0	2	2	0	2	2	0
	3	3	3	0	0	0	0	0	0	0	3	3	0
	4	4	4	0	3	3	0	3	3	0	4	4	0
(1,0)	0	0	1	0	1	1	0	1	2	0	0	0	0
	1	1	2	0	4	4	0	4	0	1	1	1	0
	2	2	3	0	2	2	0	2	3	0	2	2	0
	3	3	4	0	0	0	0	0	1	0	3	3	0

(k,p)	j	$X(k,p)$			$Y(k,p)$			$X(k,p)$			$Y(k,p)$		
		j_x	$u_x^{j_x}$	$NNS_x^{j_x}$	j_y	$u_y^{j_y}$	$NNS_y^{j_y}$	j_x	$u_x^{j_x}$	$NNS_x^{j_x}$	j_y	$u_y^{j_y}$	$NNS_y^{j_y}$
	4	4	0	1	3	3	0	3	4	0	4	4	0
(2,0)	0	0	2	0	1	1	0	1	3	0	0	0	0
	1	1	3	0	4	4	0	4	1	1	1	1	0
	2	2	4	0	2	2	0	2	4	0	2	2	0
	3	3	0	1	0	0	0	0	2	0	3	3	0
	4	4	1	1	3	3	0	3	0	1	4	4	0
(3,0)	0	0	3	0	1	1	0	1	4	0	0	0	0
	1	1	4	0	4	4	0	4	2	1	1	1	0
	2	2	0	1	2	2	0	2	0	1	2	2	0
	3	3	1	1	0	0	0	0	3	0	3	3	0
	4	4	2	1	3	3	0	3	1	1	4	4	0
(4,0)	0	0	4	0	1	1	0	1	0	1	0	0	0
	1	1	0	1	4	4	0	4	3	1	1	1	0
	2	2	1	1	2	2	0	2	1	1	2	2	0
	3	3	2	1	0	0	0	0	4	0	3	3	0
	4	4	3	1	3	3	0	3	2	1	4	4	0
(k = p)	0	0	0	1	1	1	0	1	1	1	0	0	0
	1	1	1	1	4	4	0	4	4	1	1	1	0
	2	2	2	1	2	2	0	2	2	1	2	2	0
	3	3	3	1	0	0	0	0	0	1	3	3	0
	4	4	4	1	3	3	0	3	3	1	4	4	0
(6,0)	0	0	1	1	1	1	0	1	2	1	0	0	0
	1	1	2	1	4	4	0	4	0	2	1	1	0
	2	2	3	1	2	2	0	2	3	1	2	2	0
	3	3	4	1	0	0	0	0	1	1	3	3	0
	4	4	0	2	3	3	0	3	4	1	4	4	0
...													
(9,0)	0	0	4	1	1	1	0	1	1	2	0	0	0
	1	1	0	2	4	4	0	4	4	1	1	1	0
	2	2	1	2	2	2	0	2	2	2	2	2	0
	3	3	2	2	0	0	0	0	0	2	3	3	0

(k,p)	j	$X(k,p)$			$Y(k,p)$			$X(k,p)$			$Y(k,p)$		
		j_x	$u_x^{j_x}$	$NNS_x^{j_x}$	j_y	$u_y^{j_y}$	$NNS_y^{j_y}$	j_x	$u_x^{j_x}$	$NNS_x^{j_x}$	j_y	$u_y^{j_y}$	$NNS_y^{j_y}$
	4	4	3	2	3	3	0	3	3	2	4	4	0
(10,0)	0	0	0	2	1	1	0	1	1	2	0	0	0
(k = 2p)	1	1	1	2	4	4	0	4	4	2	1	1	0
	2	2	2	2	2	2	0	2	2	2	2	2	0
	3	3	3	2	0	0	0	0	0	2	3	3	0
	4	4	4	2	3	3	0	3	3	2	4	4	0
	0	0	1	2	1	1	0	1	2	2	0	0	0
(11,0)	1	1	2	2	4	4	0	4	0	3	1	1	0
	2	2	3	2	2	2	0	2	3	2	2	2	0
	3	3	4	2	0	0	0	0	1	2	3	3	0
	4	4	0	3	3	3	0	3	4	2	4	4	0
...													
	0	0	0	0	1	2	0	1	1	0	0	1	0
(0,1)	1	1	1	0	4	0	1	4	4	0	1	2	0
	2	2	2	0	2	3	0	2	2	0	2	3	0
	3	3	3	0	0	1	0	0	0	0	3	4	0
	4	4	4	0	3	4	0	3	3	0	4	0	1
	0	0	1	0	1	2	0	1	2	0	0	1	0
(1,1)	1	1	2	0	4	0	1	4	0	1	1	2	0
	2	2	3	0	2	3	0	2	3	0	2	3	0
	3	3	4	0	0	1	0	0	1	0	3	4	0
	4	4	0	1	3	4	0	3	4	0	4	0	1
...													
	0	0	0	0	1	3	0	1	1	0	0	2	0
(0,2)	1	1	1	0	4	1	1	4	4	0	1	3	0
	2	2	2	0	2	4	0	2	2	0	2	4	0
	3	3	3	0	0	2	0	0	0	0	3	0	1
	4	4	4	0	3	0	1	3	3	0	4	1	1
(1,2)	0	0	1	0	1	3	0	1	2	0	0	2	0
	1	1	2	0	4	1	1	4	0	1	1	3	0

(k,p)	j	$X(k,p)$			$Y(k,p)$			$X(k,p)$			$Y(k,p)$		
		j_x	$u_x^{j_x}$	$NNS_x^{j_x}$	j_y	$u_y^{j_y}$	$NNS_y^{j_y}$	j_x	$u_x^{j_x}$	$NNS_x^{j_x}$	j_y	$u_y^{j_y}$	$NNS_y^{j_y}$
	2	2	3	0	2	4	0	2	3	0	2	4	0
	3	3	4	0	0	2	0	0	1	0	3	0	1
	4	4	0	1	3	0	1	3	4	0	4	1	1
...													
(k = ρ, p = ρ)	0	0	0	1	1	1	1	1	1	1	0	0	1
	1	1	1	1	4	4	1	4	4	1	1	1	1
	2	2	2	1	2	2	1	2	2	1	2	2	1
	3	3	3	1	0	0	1	0	0	1	3	3	1
	4	4	4	1	3	3	1	3	3	1	4	4	1
...													
(k = ρ, p = 2ρ)	0	0	0	1	1	1	2	1	1	1	0	0	2
	1	1	1	1	4	4	2	4	4	1	1	1	2
	2	2	2	1	2	2	2	2	2	1	2	2	2
	3	3	3	1	0	0	2	0	0	1	3	3	2
	4	4	4	1	3	3	2	3	3	1	4	4	2
...													
(k = 2ρ, p = 2ρ)	0	0	0	2	1	1	2	1	1	2	0	0	2
	1	1	1	2	4	4	2	4	4	2	1	1	2
	2	2	2	2	2	2	2	2	2	2	2	2	2
	3	3	3	2	0	0	2	0	0	2	3	3	2
	4	4	4	2	3	3	2	3	3	2	4	4	2
...													

З табл. 3.3 для випадку рівності номерів кроку квантування і ступеня для змінної x маємо такі закономірності:

- $u_x^{j_x+1}(k, p) = (u_x^{j_x}(k, p) + 1) \bmod \rho$, що з апаратної точки зору відповідає лічильнику CT_x^1 з модулем рахунку, рівним ρ , інкремент якого відбувається за кожним імпульсом тактової частоти кодування f_k ;

- $u_x^0(k+1, p) = (u_x^{\rho-1}(k, p) + 2) \bmod \rho$, що може реалізуватися шляхом уведення додаткового такту рахунку для лічильника CT_x^1 для кожного значення k , що у свою чергу, викличе необхідність синтезу частоти кодування $(\rho+1) \cdot f$ за допомогою системи ФАПЧ;

- $NNS_x^{j_x+1}(k, p) = NNS_x^{j_x}(k, p) + (u_x^{j_x}(k, p) + 1) \operatorname{div} \rho$, що відповідає лічильнику CT_x^2 , інкремент якого відбувається за сигналом переповнювання лічильника CT_x^1 ;

$$- NNS_x^0(k+1, p) = \begin{cases} NNS_x^{\rho-1}(k, p) - 1, & k+1 \neq s \cdot \rho, s \in \text{int} \\ NNS_x^{\rho-1}(k, p), & k+1 = s \cdot \rho, s \in \text{int} \end{cases},$$

що може реалізуватися шляхом декремента лічильника CT_x^2 у додатковому такті рахунку, якщо в цьому такті відсутнє переповнювання лічильника CT_x^1 ;

- $u_x^{j_x}(k, p) = u_x^{j_x}(k, p+1)$ і $NNS_x^{j_x}(k, p) = NNS_x^{j_x}(k, p+1)$ – тобто немає залежності від номера рядка;

- $u_x^0(0, p) = 0$, $NNS_x^0(0, p) = 0$, отже, лічильники CT_x^1 і CT_x^2 повинні мати вхід скидання по рядковому синхронізуючому імпульсу H-Sync;

- $u_y^{j_y}(k, p) = (u_y^0(k, p) + j_y) \bmod \rho$ – відповідає двійковому суматору SM_y^1 з корекцією результату за модулем ρ . Цей блок може працювати відповідно до

$$\text{правила } u_y^{j_y}(k, p) = \begin{cases} (u_y^0(k, p) + j_y) - \rho, & (u_y^0(k, p) + j_y) - \rho < 0 \\ (u_y^0(k, p) + j_y) & (u_y^0(k, p) + j_y) - \rho \geq 0 \end{cases};$$

- $u_y^0(k, p) = u_y^0(k+1, p) = u_y^0(0, p)$ та $NNS_y^0(k, p) = NNS_y^0(k+1, p) = NNS_y^0(0, p)$, що означає, що значення $u_y^0(0, p)$ і $NNS_y^0(0, p)$ можуть розраховуватися один раз для всього рядка;

- $u_y^0(k, p+1) = (u_y^0(k, p) + 1) \bmod \rho$, що може реалізовуватися на лічильнику CT_y^1 з модулем рахунку, рівним ρ , інкремент якого відбувається за кожним рядковим синхронізуючим імпульсом H-Sync;

- $NNS_y^{j_y}(k, p) = NNS_y^0(k, p) + (u_y^0(k, p) + j_y) \operatorname{div} \rho$, що реалізується шляхом додавання 1 до значення $NNS_y^0(k, p)$ на даному кроці кодування, якщо на ньому є перенесення в суматорі SM_y^1 ;

- $NNS_y^0(0, p+1) = NNS_y^0(0, p) + (u_y^0(0, p) + 1) \operatorname{div} \rho$, що реалізується на лічильнику CT_y^2 шляхом його інкремента по рядковому синхронізуючому імпульсу H-Sync у разі наявності перенесення в лічильнику CT_y^1 ;

Отже, апаратні витрати на реалізацію блока кодування обох змінних складуться:

- лічильник CT_x^1 з кількістю розрядів, входами $m_x^1 = \lceil \log_2(\rho + 1) \rceil$ для інкремента і скидання та виходу переповнювання (для формування сигналу переповнювання необхідно $m_x^1/4$ комірок ПЛІС для випадку фіксованого ρ і $m_x^1/2$ – для випадку того, що змінюється ρ). Тут і далі вираз $\lceil * \rceil$ означає округлення до найближчого більшого цілого значення;

- лічильник CT_x^2 з кількістю розрядів $m_x^2 = \lceil \log_2(x_{\max} \operatorname{div} \rho) \rceil$ і входами для інкремента, декремента і скидання;

- суматор SM_y^1 з кількістю комірок LUT $3 \cdot m_y^1 = \lceil \log_2(\rho) \rceil$ (m_y^1 – для реалізації $(u_y^0(k, p) + j_y)$, m_y^1 – для реалізації, $(u_y^0(k, p) + j_y) - \rho$, m_y^1 – для вибору необхідного значення);

- лічильник CT_y^1 з кількістю розрядів $m_y^1 = \lceil \log_2(\rho) \rceil$, входами для інкремента і скидання та виходом переповнювання;

- лічильник CT_y^2 з кількістю розрядів $m_y^2 = \lceil \log_2(y_{\max} \operatorname{div} \rho) \rceil$ і входами для інкремента, декремента і скидання;

- система ФАПЧ, що синтезує частоту $f_k = (\rho+1):f$.

Аналогічний аналіз проведемо для випадку рівності номера кроку кодування номера ступеня кодування для змінної u :

- $u_x^{j_x}(k, p) = (u_x^0(k, p) + j_x) \bmod \rho$ – що відповідає двійковому суматору SM_x^1 з корекцією результату за модулем ρ . Цей блок може працювати

відповідно до правила $u_x^{j_x}(k, p) = \begin{cases} (u_x^0(k, p) + j_x) - \rho, & (u_x^0(k, p) + j_x) - \rho \geq 0 \\ (u_x^0(k, p) + j_x) & (u_x^0(k, p) + j_x) - \rho < 0 \end{cases}$;

- $u_x^0(k+1, p) = (u_x^0(k, p) + 1) \bmod \rho$ – інкремент з частотою f лічильника CT_x^1 , що зберігає значення $u_x^0(k, p)$;

- $NNS_x^{j_x}(k, p) = NNS_x^0(k, p) + (u_x^0(k, p) + j_x) \text{div } \rho$, що реалізується шляхом додавання 1 до значення $NNS_x^0(k, p)$ на даному кроці кодування, якщо на ньому є перенесення в суматорі SM_x^1 ;

- $NNS_x^0(k+1, p) = NNS_x^0(k, p) + (u_x^0(k, p) + 1) \text{div } \rho$, що реалізується на лічильнику CT_x^2 шляхом його інкремента з частотою f у разі наявності перенесення в лічильнику CT_x^1 ;

- $u_x^{j_x}(k, p) = u_x^{j_x}(k, p+1)$ и $NNS_x^{j_x}(k, p) = NNS_x^{j_x}(k, p+1)$ – тобто немає залежності від номера рядка;

- $u_x^0(0, p) = 0$, $NNS_x^0(0, p) = 0$, отже, лічильники CT_x^1 і CT_x^2 повинні мати вхід скидання по рядковому синхронізуючому імпульсу H-Sync;

- $u_y^{j_y+1}(k, p) = (u_y^{j_y}(k, p) + 1) \bmod \rho$ – лічильник CT_y^1 з модулем рахунку, рівним ρ , інкремент якого відбувається за кожним імпульсом тактової частоти кодування f_k ;

- $u_y^0(k+1, p) = (u_y^{\rho-1}(k, p) + 1) \bmod \rho$, що автоматично реалізується лічильником CT_y^1 ;

- $u_y^0(0, p+1) = (u_y^{\rho-1}(x_{\max} - 1, p) + 1) \bmod \rho$, що реалізується додатковим інкрементом лічильника CT_y^1 на початку нового рядка;
- $NNS_y^{j_y+1}(k, p) = NNS_y^{j_y}(k, p) + (u_y^{j_y}(k, p) + 1) \operatorname{div} \rho$, що відповідає інкременту лічильника CT_y^2 , який відбувається за сигналом переповнювання лічильника CT_y^1 ;
- $NNS_y^0(k, p) = NNS_y^0(k+1, p) = NNS_y^0(0, p)$, що може бути отримане шляхом декремента лічильника CT_y^2 після останнього ступеня кодування, тобто за сигналом переповнювання лічильника кроків кодування. При цьому, якщо виникає переповнювання лічильника CT_y^1 , ні інкремент, ні декремент лічильника CT_y^2 не проводиться.

Як видно з наведених міркувань, апаратні витрати за виключенням декількох вентилів в обох розглянутих випадках однакові. Однак другий випадок не вимагає додаткового такту кодування, і, відповідно, дозволяє знизити необхідну для кодування частоту. У першому випадку також можна обійтися без введення додаткового такту кодування. Це потребує додаткового лічильника для зберігання значення u_x^0 , інкремент якого має відбуватися за кожним імпульсом частоти f . Крім того, значення, що зберігається цим лічильником, має з частотою f записуватися в лічильник CT_y^1 . Тобто збільшаться апаратні витрати. Оскільки в цілому немає принципової різниці, для якої змінної вибрати рівність кроку кодування і номера ступеня, синтез структури блока кодування змінних проводимо для другого випадку.

Структура блока кодування змінних для випадку послідовного перебору значень змінних і рівності номера кроку кодування номеру ступеня кодування для змінної u наведена на рис. 3.19.

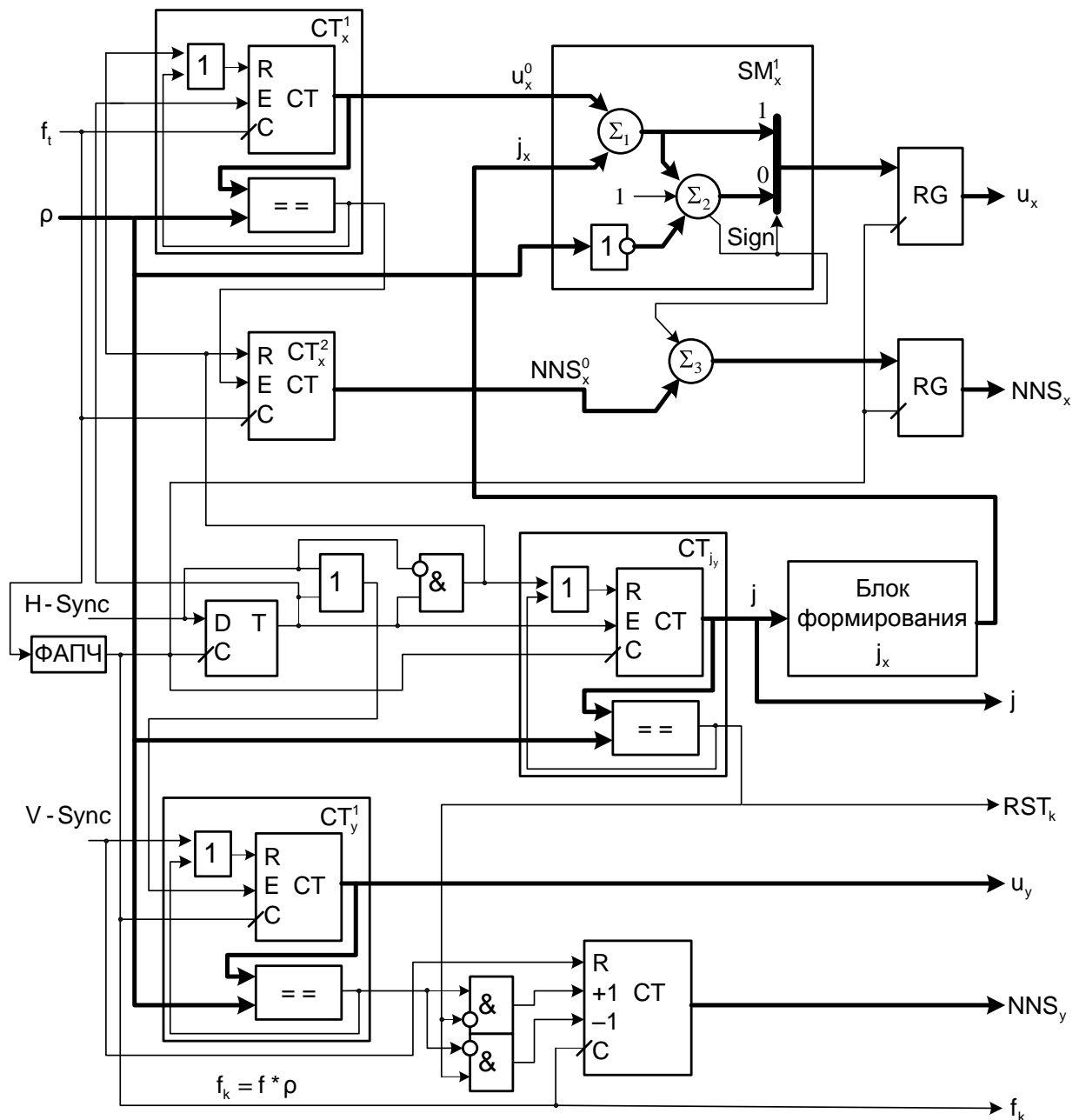


Рисунок 3.19 – Структура блока кодування змінних з послідовно перебираними значеннями

3.12 Стиснення і фільтрація зображень

Основною відмінністю застосування мережі СМАС для стиснення і фільтрації зображень є необхідність навчання в реальному масштабі часу. Це вимагає наявності засобів для проведення навчання, а також обумовлює необхідність заміни послідовного перебору змінних довільним, що, у свою чергу, викликає необхідність переробки кодуєчих блоків. Структура мережі

СМАС, яка може використовуватися як самостійно, так і в ієрархічній структурі, наведено на рис. 3.20.

У пропонованій структурі для збереження можливості вибору довільних значень ρ у заданому інтервалі пропонується як кодуючі блоки використовувати ОЗП відповідного розміру, що містять таблиці, за допомогою яких виконується перекодування змінних x і y в комбінації $u_x:NNS_x$, і $u_y:NNS_y$. Для стандартного телевізійного кадру розмірами 720 на 576 точок вхідні змінні подаються 10-розрядним двійковим кодом, а, отже, ОЗП для кодування мають містити по 1024 комірки, що мають розрядність 14 бітів для можливості задання ρ у діапазоні від 2-х до 32-х. Такі ОЗП можуть реалізовуватися в сучасних ПЛІС. Мультиплексори $MX1x$ і $MX1y$ і буферні елементи $B1x$ і $B1y$ необхідні для ініціалізації кодуючих ОЗП.

Обчислення добутку базисних функцій здійснюється аналогічно розглянутим вище структурам. Ваги, що настроюються, зберігаються в ОЗП ваг, яке, на відміну від структур, що не вимагають навчання, має реалізовувати одночасне читання і запис за різними адресами.

Дана структура припускає навчання згідно з залежністю, яку запишемо нижче у вигляді, зручному для апаратної реалізації

$$w_j(k+1) = w_j(k) + \gamma \left(\frac{I_{in}(x(k), y(k)) - \sum_{i=0}^{\rho-1} \Psi_i(x(k))w_i(k)}{\sum_{n=0}^{\rho-1} (\Psi_n(x(k)))^2} \Psi_j(x(k)) \right). \quad (3.82)$$

Сума, що стоїть в чисельнику (3.82), є по суті значенням функції, що обчислюється мережею до настройки ваг. Вона обчислюється помножувачем $M1$ і накопичуючим суматором $\Sigma 1$. Після проходження ρ ступенів кодування в регістрі $RG1$ формується значення функції, обчислене мережею до навчання,

яке за допомогою суматора $\Sigma 2$ віднімається з потрібного при даній вхідній комбінації змінних значення функції $I_{in}(x(k), y(k))$.

Значення помилки зберігається в регістрі RG2, з якого поступає на помножувач M2, що обчислює добуток помилки на коефіцієнт γ , який визначає швидкість навчання мережі. Значення норми вектора коефіцієнтів, що стоїть в знаменнику Ψ обчислюється помножувачем M3 і накопичуючим суматором $\Sigma 3$ -RG3 паралельно із значенням вихідної функції до навчання. Обчислене значення норми після завершення ρ ступенів кодування зберігається в регістрі RG4. Вектор коефіцієнтів, використовуваний для розрахунку значення функції до навчання, зберігається для подальшого використання в процесі навчання в ланцюжку зсувних регістрів. Аналогічно зберігаються адреси і значення ваг. Ці ваги будуть модифіковані в процесі навчання і записані в ОЗП за відповідними адресами. Довжина ланцюжків зсувних регістрів рівна ρ . Ці ланцюжки можуть реалізовуватися на вбудованих в ПЛІС блоках пам'яті, що включаються в двопортовий режим. Зміщення між адресами запису і читання має бути рівне ρ . Після затримки на ρ тактів, елементи вектора коефіцієнтів нормуються шляхом ділення на значення норми вектора коефіцієнтів в дільнику D1. Нормовані значення елементів вектора коефіцієнтів поступають на помножувач M4 спільно із значенням добутку помилки на коефіцієнт γ , який обчислює значення, що ко-ректус, для ваги, яка модифікується на даному кроці навчання. Це значення, поступаючи на суматор $\Sigma 4$, додається до початкового значення ваги. Модифіковане значення ваги зберігається в регістрі RG5. З регістра RG5 спільно з власною адресою це значення поступає на лінії запису ОЗП ваг. Крім того, це значення поступає на помножувач M5 і накопичуючий суматор $\Sigma 5$ -RG6, функції, що обчислюють значення, після навчання мережі. Це значення зберігається в регістрі RG7 і надалі може використовуватися як результат фільтрації початкового зображення, або для навчання наступної в ієрархічному ланцюжку мережі.

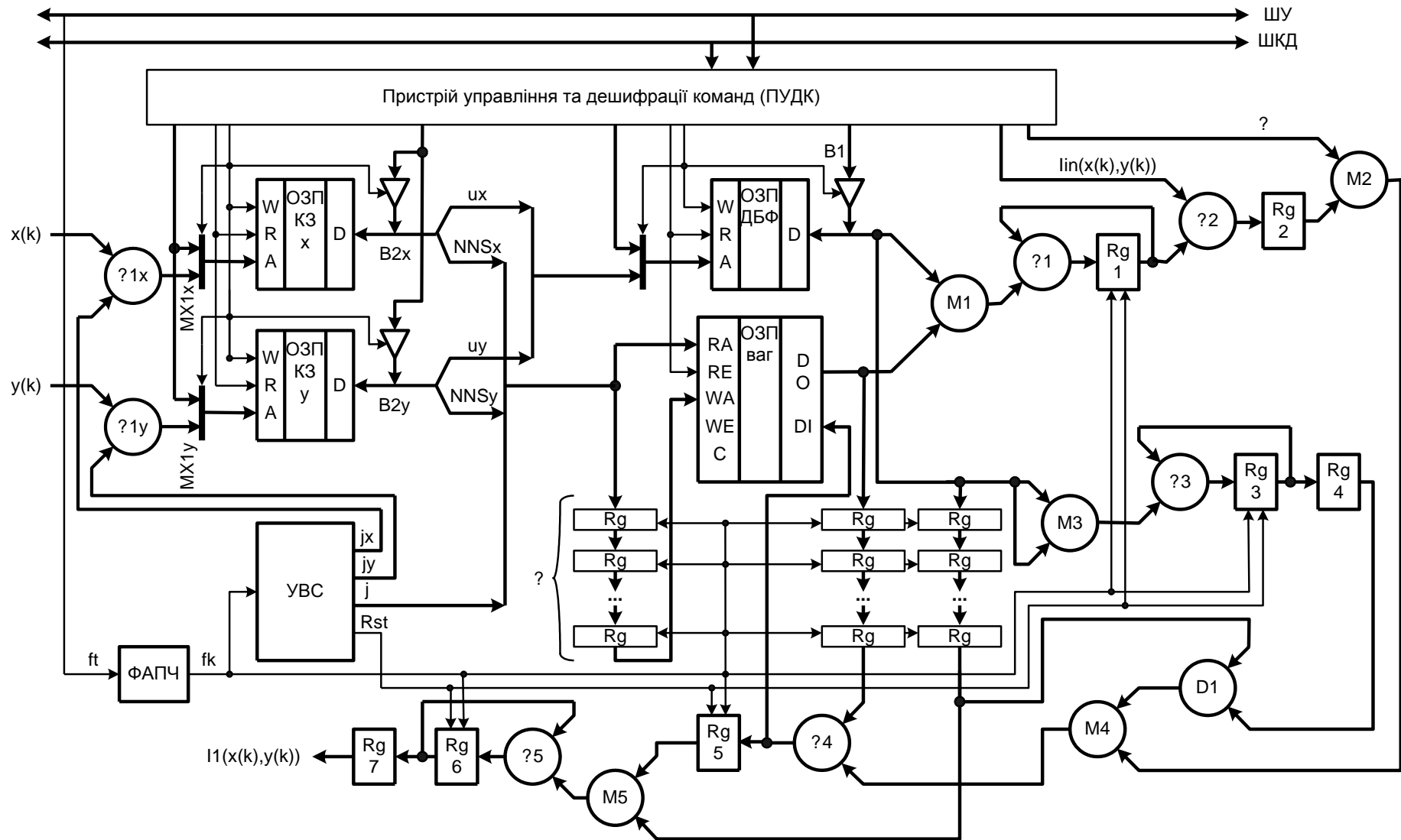


Рисунок 3.20 – Структура апаратної реалізації мережі СМАС, що вимагає навчання в реальному часі

Необхідно зазначити збільшення складності апаратної реалізації мережі СМАС за необхідності її навчання в реальному масштабі часу. Це пов'язано із збільшеною кількістю операцій множення і появою операції ділення при розрахунку величин корекції ваг. Існують алгоритми спільного виконання операцій множення і ділення, проте у зв'язку з їх ітераційним характером вони в даній структурі непридатні. Потрібна їх адаптація до необхідної швидкості обчислень. Ці питання виходять за рамки даної дисертаційної роботи. Безпосереднє виконання операцій множення і ділення можливо на ПЛІС, що мають в своєму складі засоби цифрової обробки сигналів. Такі ПЛІС є на сучасному ринку, хоча все ще мають високу ціну. Внаслідок того, що кодуючої апаратури потрібно значно менше, ніж декодуючої, це не є серйозним недоліком.

При фільтрації зображень можливе навчання мережі на частини точок початкового зображення, що дозволяє знизити швидкість виконання обчислень поправок ваг і використовувати вищезазначені ітераційні алгоритми.

Мережі СМАС у декодуючій частині розглянутої ієрархічної структури не вимагають навчання і мають структуру, подібну наведеній на рис. 3.20. У цих мережах вхідні змінні перебираються послідовно, оскільки розгортка зображення відбувається в правильному порядку. Ці мережі ініціалізувалися вагами, що поступили каналом зв'язку і базисними функціями того ж вигляду, що використовуються на кодуючій стороні.

Наведемо порівняльну таблицю об'ємів пам'яті, необхідних для зберігання початкового зображення і стислого з використанням однієї мережі СМАС (табл. 3.4).

З таблиці видно, що скорочення об'єму інформації залежно від параметра мережі ρ може складати від 3,9 до 29,1 разів. За необхідності підвищення якості стислого зображення необхідно застосовувати ієрархічну структуру. В цьому випадку об'єм зображення дорівнюватиме сумі об'ємів пам'яті ваг, використовуваних у мережах ієрархічної структури.

Таблиця 3.4 – Об'єми зображень після стиснення

Нестисле зображення розмірами 720·576 точок	Стисле з використанням мережі СМАС зображення, комірок			
	$\rho = 4$	$\rho = 8$	$\rho = 16$	$\rho = 32$
414720	104655	52981	27150	14246

При розгляді реалізації колірної корекції, синтезу функцій присутності, стиснення і фільтрації зображень не проводиться аналіз отримуваної похибки обчислень, яка залежить як від властивостей самої мережі СМАС і вибору її параметрів, так і від кількості розрядів при обчисленнях, вироблюваних у розглянутих структурах.

У зв'язку з тим, що пропонувані структури мають можливості вибору оптимального співвідношення (об'єм інформації)/(якість) і використовуються в області обробки відеоінформації, де критерієм якості є суб'єктивна оцінка, пропонується залишити за користувачем настройку параметрів структур через деякий інтерфейс і, відповідно, візуальну оцінку якості обробки.

4 НЕЙРО-ФАЗЗИ СИСТЕМИ ТА АЛГОРИТМИ НАВЧАННЯ ДЛЯ ОБРОБКИ НЕОДНОРІДНИХ ПОЛІВ СПОСТЕРЕЖЕНЬ ЗА УМОВ МАЛОЇ ВИБІРКИ ДАНИХ

4.1 Загальна інформація

Проблема оброблення багатовимірних спостережень досить часто зустрічається в багатьох задачах, пов'язаних з інтелектуальним аналізом даних. Традиційний підхід до вирішення цих завдань допускає, що кожне спостереження може відноситися тільки до одного кластеру, хоча більш природною є ситуація, коли вектор ознак, який надходить на обробку, з різними рівнями належності або ймовірності, або можливості може належати одночасно декільком класам. Така ситуація є предметом розгляду нечіткого (фаззи) кластерного аналізу, який в нинішній час досить інтенсивно розвивається.

Традиційно вихідною інформацією для завдання кластеризації є вибірка спостережень, сформована з N n -вимірних векторів ознак $X = x(1), x(2), \dots, x(k), \dots, x(N)$, $x(k) = x_1(k), \dots, x_n(k)^T \in R^n$, $k = 1, 2, \dots, N$, а результатом роботи алгоритму є поділ вихідного масиву даних на m класів з деяким рівнем $u_j(k)$ належності k -ого вектора ознак j -ому кластеру.

У той же час існує широкий клас задач, коли вихідна інформація надходить не в векторній, а в матричній формі, тобто $x(k) = \{x_{i_1 i_2}(k)\}$; $i_1 = 1, 2, \dots, n_1$; $i_2 = 1, 2, \dots, n_2$; $k = 1, 2, \dots, N$. Така ситуація є характерною, наприклад, при обробці зображень, коли вихідна $(N_1 \times N_2)$ -матриця розбивається на $N = N_1 \cdot N_2 \cdot (n_1 \cdot n_2)^{-1}$ $(n_1 \times n_2)$ -матриць-фрагментів, які підлягають кластеризації, в результаті якої формуються однорідні в деякому розумінні сегменти цього зображення. Традиційно це завдання вирішується шляхом попередньої векторизації фрагментів і використання вже відомих

процедур, найбільш популярною з яких є метод кластеризації нечітких с–середніх.

Для обробки матричних даних необхідно ввести матричні методи кластеризації. Вводиться в розгляд матричний метод нечітких с–середніх, який є узагальненням FCM, дозволяє уникнути зайвих операцій векторизації-девекторизації при обробці даних, які задані у формі двовимірних масивів, і забезпечує обробку інформації в online режимі.

Матрична модифікація методу кластеризації нечітких с–середніх

Нехай задана вибірка спостережень $x(k) = \{x_{i_1 i_2}(k)\} \in R^{n_1 \times n_2}$, $k = 1, 2, \dots, N$, при цьому для зручності подальшої обробки ці дані попередньо відцентровані щодо середнього:

$$\bar{x} = \frac{1}{N} \sum_{k=1}^N x(k) \quad (4.1)$$

і нормовані на свою сферичну норму (Frobeniusnorm):

$$\|x(k)\| = \sqrt{\text{Tr } x(k)x^T(k)}. \quad (4.2)$$

В якості цільової функції кластеризації використовується імовірнісний критерій

$$\begin{aligned} E(u_j(k), c_j) &= \sum_{k=1}^N \sum_{j=1}^m u_j^\beta(k) D^2(x(k), c_j) = \\ &= \sum_{k=1}^N \sum_{j=1}^m u_j^\beta(k) \text{Tr}(x(k) - c_j)(x(k) - c_j)^T \end{aligned} \quad (4.3)$$

при наявності обмежень

$$\sum_{j=1}^m u_j(k) = 1, \text{ або } \sum_{j=1}^m u_j(k) - 1 = 0, k = 1, 2, \dots, N, 0 < \sum_{j=1}^m u_j(k) < N, j = 1, 2, \dots, m,$$

де $u_j(k) \in [0, 1]$ – рівень належності спостереження $x(k)$ j -му кластеру, c_j – прототип (центроїд) j -го кластера, β – невід’ємний параметр, званий «фаззіфікатором», $D^2(x(k), c_j)$ – міра відстані (квадрат норми) між $x(k)$ та c_j . Результатом кластеризації є $(N \times m)$ -матриця $U = \{u_j(k)\}$, яка називається матрицею нечіткого розбиття.

Вводячи функцію Лагранжа

$$\begin{aligned} L(u_j(k), c_j, \lambda(k)) &= \sum_{k=1}^N \sum_{j=1}^m u_j^\beta(k) D^2(x(k), c_j) + \sum_{k=1}^N \lambda(k) \left(\sum_{j=1}^m u_j(k) - 1 \right) = \\ &= \sum_{k=1}^N \left(\sum_{j=1}^m u_j^\beta(k) D^2(x(k), c_j) + \lambda(k) \left(\sum_{j=1}^m u_j(k) - 1 \right) \right), \end{aligned} \quad (4.4)$$

де $\lambda(k)$ – невизначений множник Лагранжа, і вирішуючи систему рівнянь Каруша-Куна-Таккера

$$\begin{cases} \frac{\partial L(u_j(k), c_j, \lambda(k))}{\partial u_j(k)} = \beta u_j^{\beta-1}(k) D^2(x(k), c_j) + \lambda(k) = 0, \\ \frac{\partial L(u_j(k), c_j, \lambda(k))}{\partial \lambda_j(k)} = \sum_{j=1}^m u_j(k) - 1 = 0, \\ \left\{ \frac{\partial L(u_j(k), c_j, \lambda(k))}{\partial c_j(k)} \right\} = -2 \sum_{k=1}^N u_j^\beta(k) (x(k) - c_j) = 0, \end{cases}$$

де $\left\{ \frac{\partial L(u_j(k), c_j, \lambda(k))}{\partial c_j(k)} \right\}$ $-(n_1 \times n_2)$ -матриця, утворена частковими похідними $\frac{\partial L(u_j(k), c_j, \lambda(k))}{\partial c_{j_1 j_2}}$, O – матриця тієї ж розмірності, утворена нулями, приходимо до кінцевого результату:

$$\left\{ \begin{array}{l} u_j(k) = \frac{(D^2(x(k), c_j))^{\frac{1}{1-\beta}}}{\sum_{l=1}^m (D^2(x(k), c_l))^{\frac{1}{1-\beta}}}, \\ \lambda(k) = -\left(\sum_{l=1}^m (\beta D^2(x(k), c_l))^{\frac{1}{1-\beta}}\right)^{1-\beta}, \\ c_j = \frac{\sum_{k=1}^N u_j^\beta(k) x(k)}{\sum_{k=1}^N u_j^\beta(k)}. \end{array} \right. \quad (4.5)$$

Отримана система породжує широкий клас процедур кластеризації. Так, вибираючи $\beta = 2$, отримуємо простий і ефективний алгоритм матричної кластеризації, який є узагальненням популярної процедури Дж. Бездека [85]:

$$\left\{ \begin{array}{l} u_j(k) = \frac{(Tr(x(k) - c_j)(x(k) - c_j)^T)^{-1}}{\sum_{l=1}^m (Tr(x(k) - c_l)(x(k) - c_l)^T)^{-1}}, \\ c_j = \frac{\sum_{k=1}^N u_j^2(k) x(k)}{\sum_{k=1}^N u_j^2(k)}, \end{array} \right. \quad (4.6)$$

де Tr – символ сліду матриці.

Функціонування алгоритму кластеризації починається з завдання початкової (зазвичай випадкової) матриці нечіткого розбиття U^0 . На її основі

розраховується початковий набір прототипів c_j^0 , які далі використовуються для обчислення нової матриці U^1 . Дані в пакетному режимі перераховуються $c_j^1, U^2, \dots, c_j^t, U^t$, поки різниця $\|U^t - U^{t-1}\|$ чи не стане менше деякого апріорі заданого порогу ε . Таким чином весь масив спостережень обробляється кілька разів.

4.2 Адаптивний матричний алгоритм кластеризації нечітких s -середніх

Алгоритм (4.6) може бути розширений на випадок коли дані на обробку надходять послідовно в online режимі. Для цього, застосовуючи до лагранжіана (4.4) процедуру пошуку сідлової точки Ерроу-Гурвіца-Удзави (Arrow-Hurwitz-Uzawa), при надходженні $(k+1)$ -ого спостереження оцінки рівнів належності і центроїд можуть бути уточнені за допомогою рекурентних співвідношень

$$\left\{ \begin{array}{l} u_j(k+1) = \frac{(D^2(x(k+1), c_j(k))^{\frac{1}{1-\beta}})}{\sum_{l=1}^m (D^2(x(k+1), c_l(k))^{\frac{1}{1-\beta}})}, \\ c_j(k+1) = c_j(k) - \eta(k) \left\{ \frac{\partial L(u_j(k+1), c_j, \lambda(k+1))}{\partial c_j} \right\} = \\ = c_j(k) + \eta(k) u_j^\beta(k+1) (x(k+1) - c_j(k)) \end{array} \right. \quad (4.7)$$

для довільного значення фаззифікатора β (тут $\eta(k)$ – параметр кроку пошуку (learningrateparameter)) та

$$\begin{cases} u_j(k+1) = \frac{(\text{Tr}(x(k+1) - c_j(k))(x(k+1) - c_j(k))^T)^{-1}}{\sum_{l=1}^m (\text{Tr}(x(k+1) - c_l(k))(x(k+1) - c_l(k))^T)^{-1}}, \\ c_j(k+1) = c_j(k) + \eta(k)u_j^2(k+1)(x(k+1) - c_j(k)) \end{cases} \quad (4.8)$$

для $\beta = 2$.

Нескладно побачити, що співвідношення (4.7) є адаптивною версією процедури (4.5), а (4.8) - відповідає (4.6). Відзначимо також, що процедури (4.7), (4.8) є матричним узагальненням рекурентних алгоритмів нечіткої кластеризації і відрізняються простотою чисельної реалізації.

В якості програмного засобу для реалізації модифікацій алгоритму нечітких с-середніх була обрана об'єктно-орієнтована мова програмування Java і інтегрована у середовище розробки IntelliJ IDEA 11.

Існує велика кількість різноманітних об'єктно-орієнтованих мов програмування, а також математичних пакетів, що знаходять широке застосування в сучасних дослідженнях, зокрема в інтелектуальному аналізі даних. В першу чергу це пакети Matlab, Statistica, Matcad, Scilab, Octave і мови C # .NET, C ++, Python. Вони володіють багатими можливостями візуалізації, роботи з векторними і матричними об'єктами і деякі з них могли б використовуватися для вирішення поставленого завдання.

Вибір мови Java обумовлений декількома причинами. По-перше, Java може здійснювати обмін даними з іншими середовищами і мовами, в тому числі з Matlab і Scilab. По-друге, Java є кросплатформною мовою, а також є безкоштовним, як і середовище розробки IntelliJ IDEA. По-третє, мова має безліч вбудованих бібліотек, що забезпечують просту і ефективну роботу з графічними об'єктами, має широку базу математичних функцій. Також в мові Java високий рівень абстракції, що дозволяє використовувати класи і їхні методи, не знаючи деталей їх реалізації.

Для тестування матричних модифікацій алгоритму кластеризації нечітких c -середніх були використані вибірки UCI-репозиторію Iris (ірис), GlassIdentification (ідентифікація типів скла), Pen-BasedRecognitionofHandwrittenDigits (розпізнавання рукописних цифр), BreastCancerWisconsinOriginal (рак молочної залози) [86], а також цифрові зображення, в тому числі сателітні знімки міста Харкова. Вибірки не містять пропущених атрибутів і є повністю числовими. Властивості вибірок вказані в табл. 4.1.

Таблиця 4.1 – Характеристики вибірок даних

Назва вибірки	Кількість спостережень	Кількість атрибутів	Кількість класів
Iris	150	4	3
GlassIdentification	214	9	6
Pen-Based Recognition of Handwritten Digits	10992	16	10
BreastCancerWisconsinOriginal	699	9	2

Після запуску програми в оперативну пам'ять середовища завантажується файл даних, з якого формується вибірка векторних або матричних спостережень. Потім задаються параметри алгоритму: кількість кластерів m , поріг зупину алгоритму і параметр, що відповідає за відсоток спостережень вибірки, на якому відбуватиметься навчання алгоритму. Кожному спостереженню присвоюється прапор, який вказує, чи буде алгоритм навчатися на ньому. Потім всі об'єкти центруються відносно середнього і нормуються на свою сферичну норму.

Після того, як вибірка спостережень завантажена і попередньо оброблена, задається випадкова матриця нечіткого розбиття U^0 . Потім алгоритм циклічно в пакетному режимі обчислює нові значення прототипів c_j , де j – номер класу, і матриці нечіткого розбиття U , поки різниця $\|U^t - U^{t-1}\|$ не стає менше порога ε .

Адаптивний алгоритм працює за схожим принципом, з тією лише різницею, що кожен об'єкт обробляється один раз, і відбувається це не в пакетному режимі, а в режимі online, коли об'єкти на обробку надходять по одному.

Після отримання кінцевого набору прототипів, для спостережень, які брали участі в процесі навчання, обчислюються векториналежностей за формулою (4.6) для пакетного алгоритму кластеризації і за формулою (4.8) для адаптивного алгоритму.

Результатом роботи алгоритму є кінцева матриця нечіткого розбиття для всіх об'єктів вибірки і прототипи класів.

При обробці цифрових зображень, об'єкти (матриці або вектори однакової розмірності) створюються з фрагментів цього зображення, а кожен піксель з колірної моделі RGB (Red-Green-Blue) переводиться в модель Grayscale, де яскравість пікселя виражається скалярним значенням з інтервалу $[0, 1]$. Переклад з моделі RGB в модель Grayscale здійснюється за формулою $Y = (0.299R + 0.587G + 0.114B) / 255$, де Y – яскравість світіння пікселя, R , G , B – яскравості світіння червоного, зеленого і синього тонів відповідно, значення яких лежать на інтервалі $[0, 255]$.

Набори спостережень, сформовані з цифрових зображень, обробляються за тим же принципом, що і стандартні кількісні вибірки.

Після обробки зображень, кожному кластеру присвоюється колір моделі Grayscale, а кожен об'єкт забарвлюється в колір найближчого кластеру.

Приклад цифрового зображення, яке надходить на обробку, вказаний на рис. 4.1, а приклад обробленого зображення - на рис. 4.2.

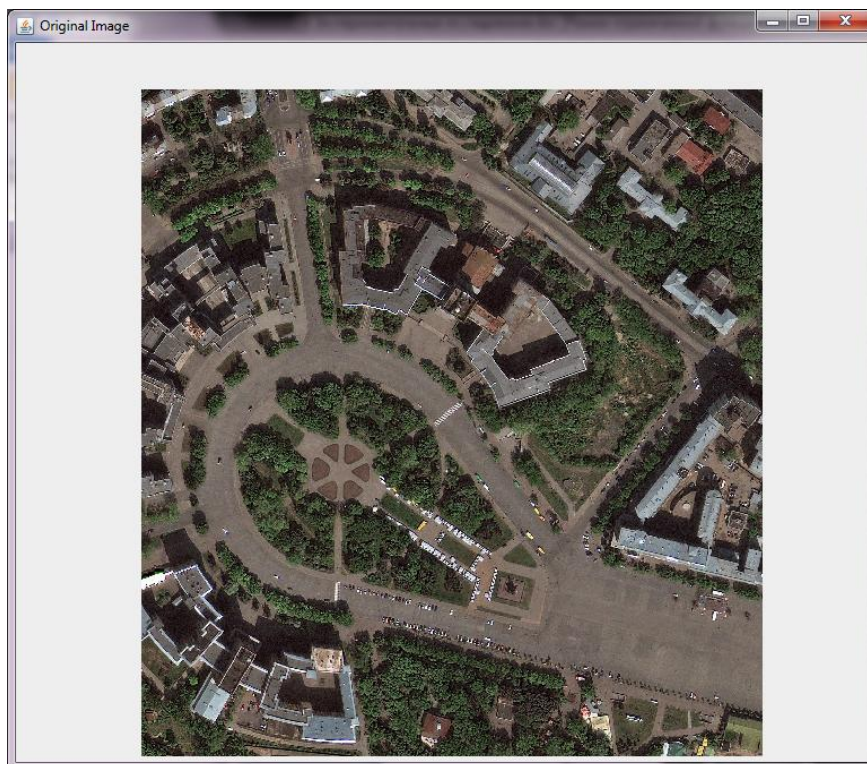


Рисунок 4.1 – Цифрове зображення, яке надходить на обробку

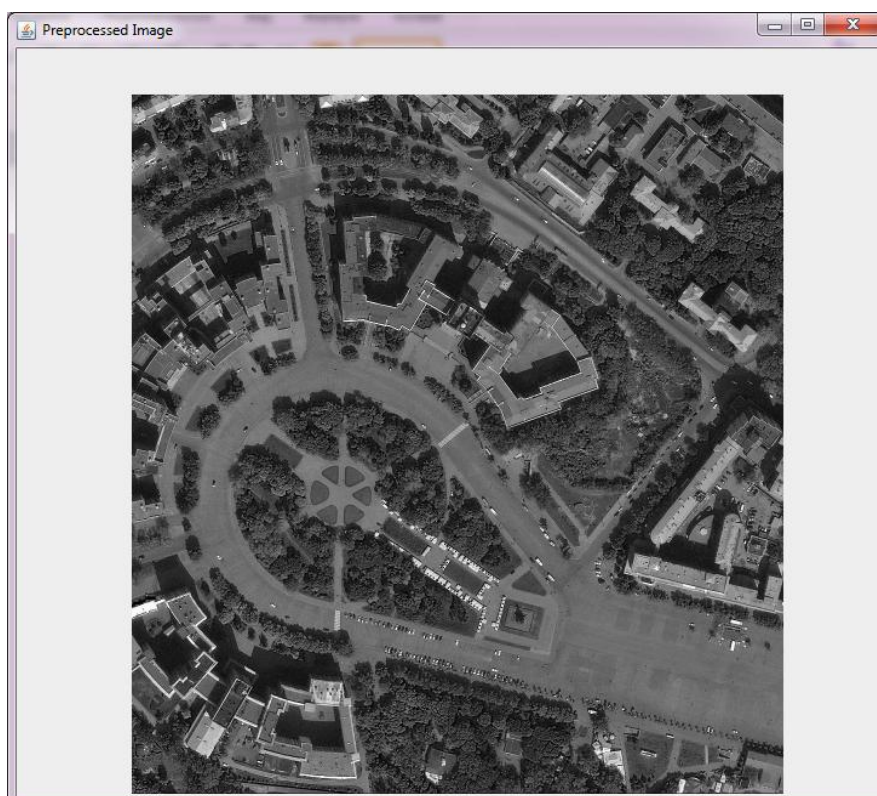


Рисунок 4.2 – Передоброблене зображення

Опис програмної реалізації

Модифікації алгоритму нечітких с середніх реалізовані на мові Java. При цьому розроблені наступні класи:

- `DataObject.java` – являє собою матрицю розміром $n \times m$, а також може представляти скалярну або векторну величину. Містить методи для роботи з матрицями;

- `DataSet.java` – набір спостережень (об'єктів класу `DataObject.java`) однакової розмірності. Містить методи обробки всієї вибірки;

- `DataLoader.java` – клас-завантажувач даних для аналізу;

- `CentroidsSet.java` – являє собою набір прототипів класів. Містить методи для роботи з центроїдами кластерів;

- `BelongVector.java` – представляє вектор чисел з плаваючою точкою для вказівки приналежності спостереження класу;

- `BelongingMatrix.java` – включає в себе набір об'єктів класу `BelongVector.java` для представлення матриці нечіткого розбиття;

- `ClusteringAlgorithm.java` – абстрактний клас алгоритму кластеризації. Містить методи для кластерного аналізу в пакетному режимі;

- `MatrixFCMAAlgorithm.java` – реалізація матричного алгоритму нечітких с середніх;

- `AdaptiveClusteringAlgorithm.java` – абстрактний клас алгоритму кластеризації. Містить методи для кластерного аналізу в режимі online;

- `AdaptiveMatrixFCMAAlgorithm.java` – реалізація адаптивного матричного алгоритму нечітких с середніх;

- `ExcelWriter.java` – клас, що містить статичні методи запису результатів кластеризації у вигляді електронної таблиці у файл формату `*.xls`;

- `ImagePanel.java` – клас, який розширює `JFrame` (віконна форма) і відображає цифрові зображення.

- Листинги вищевказаних класів містяться в додатку А.

Для оцінки якості роботи алгоритму використовувалися такі критерії: PartitionCoefficient (PC), ClassificationEntropy (CE), PartitionIndex (PI).

У табл. 4.2 представлені результати точності і часу роботи алгоритмів кластеризації на вибірці Iris, а в табл. 4.3 - на сателітному цифровому зображенні міста Харкова. Час вказано в середньому для однієї ітерації з урахуванням операції векторизації-девекторизації.

Таблиця 4.2 - Результати кластерного аналізу на вибірці Iris

Алгоритм кластеризації	PC	CE	PI	Час, с
Нечіткихс-середніх	0.531	0.811	12.19	0.003
Матричний алгоритм нечітких с-середніх	0.531	0.811	12.19	0.0025

Таблиця 4.3 – Результати кластерного аналізу на цифровому зображенні

Алгоритм кластеризації	PC	CE	PI	Час, с
Нечіткихс-середніх	0.697	0.419	8.23	1.9
Матричний алгоритм нечітких с-середніх	0.697	0.419	8.23	1.8

На рис. 4.3 – 4.6 відповідно представлено вихідне зображення, оброблена вибірка (20% об'єктів), результат кластерного аналізу і процес роботи алгоритму.

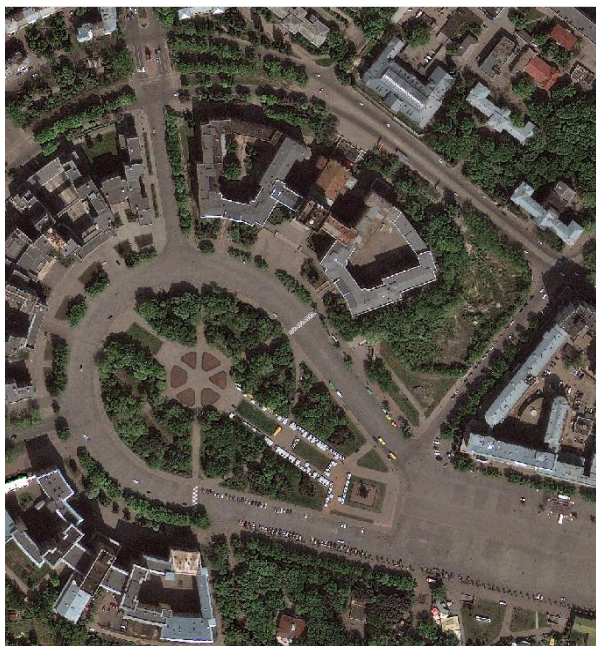


Рисунок 4.3 – Початкове цифрове зображення для кластеризації

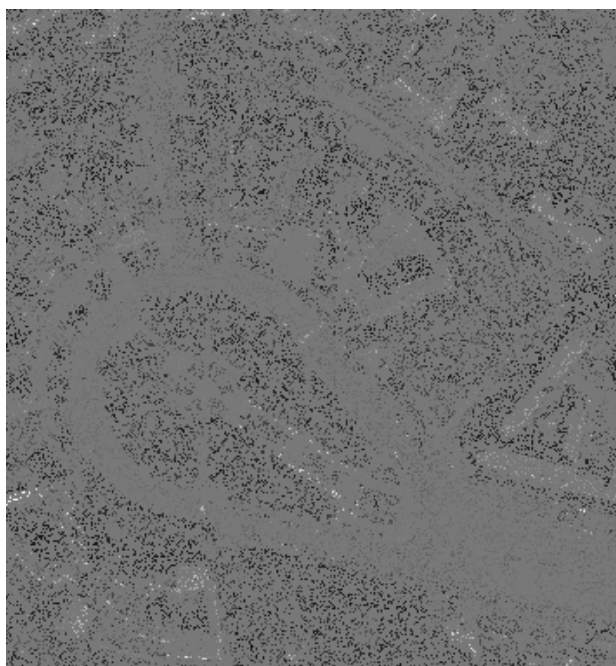


Рисунок 4.4 – Передоброблена вибірка (20% об'єктів) для кластеризації



Рисунок 4.5 – Вихідне зображення кластерного аналізу

```

-----
Iteration: 1   ε: 24,154805   Time: 1,86880 s
Iteration: 2   ε: 1,835973   Time: 1,90209 s
Iteration: 3   ε: 0,628801   Time: 2,60592 s
Iteration: 4   ε: 0,353193   Time: 2,27173 s
Iteration: 5   ε: 0,215733   Time: 2,23635 s
Iteration: 6   ε: 0,150384   Time: 2,26352 s
Iteration: 7   ε: 0,109181   Time: 2,35191 s
Iteration: 8   ε: 0,080619   Time: 2,26822 s
Iteration: 9   ε: 0,060460   Time: 2,62709 s
Iteration: 10  ε: 0,046043   Time: 2,48717 s
Iteration: 11  ε: 0,035485   Time: 2,28367 s
Iteration: 12  ε: 0,027542   Time: 2,19080 s
Iteration: 13  ε: 0,021493   Time: 2,24285 s
Iteration: 14  ε: 0,016868   Time: 2,26129 s
Iteration: 15  ε: 0,013294   Time: 2,39989 s
Iteration: 16  ε: 0,010520   Time: 2,19403 s
Iteration: 17  ε: 0,008349   Time: 2,18732 s

```

Рисунок 4.6 – Процес роботи алгоритму кластерного аналізу

На рис. 4.7 наведений результат кластеризації цифрового зображення адаптивним матричним алгоритмом нечітких-середніх.

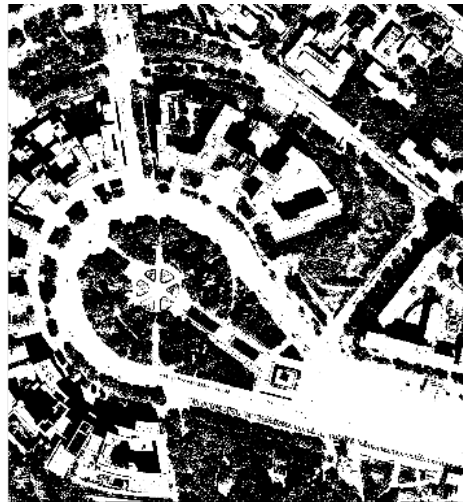


Рисунок 4.7 – Вихідне зображення адаптивного кластерного аналізу

На рис. 4.8 представлена частина кінцевої матриці нечіткого розбиття, а на рис. 4.9 –кінцеві значення прототипів класів, отримані в результаті кластеризації вибірки Iris.

	A	B	C	D
1		Cluster # 0	Cluster # 1	Cluster # 2
2	Object # 0	0.000496325	0.00067804	0.99882563
3	Object # 1	0.006529864	0.01019459	0.98327555
4	Object # 2	0.003900382	0.0059351	0.99016452
5	Object # 3	0.009136773	0.01450453	0.97635869
6	Object # 4	0.000934492	0.00126564	0.99779986
7	Object # 5	0.015379749	0.01832698	0.96629327
8	Object # 6	0.004997068	0.00735954	0.98764339
9	Object # 7	0.000170108	0.00024099	0.99958889
10	Object # 8	0.01783695	0.03004381	0.95211924
11	Object # 9	0.004897069	0.00754933	0.9875536
12	Object # 10	0.008003043	0.01008681	0.98191015
13	Object # 11	0.00259657	0.00378084	0.99362259
14	Object # 12	0.007549432	0.01196828	0.98048229
15	Object # 13	0.013171975	0.02146973	0.96535829
16	Object # 14	0.029986294	0.03415834	0.93585537
17	Object # 15	0.04761057	0.0500123	0.90237713
18	Object # 16	0.013894244	0.01671872	0.96938704
19	Object # 17	0.000628996	0.00085571	0.99851529
20	Object # 18	0.023032024	0.02694006	0.95002792

Рисунок 4.8 – Матриця нечіткого розбиття для вибірки Iris

```
Centroid # 0
[0,3631    -0,0708
 0,8561     0,3609    ]
Centroid # 1
[-0,2986   -0,8378
 0,4570     0,0063    ]
Centroid # 2
[-0,3095    0,1376
 -0,8692   -0,3602    ]
```

Рисунок 4.9 – Значення прототипів класів для вибірки Iris

5 ДОСЛІДЖЕННЯ ЗМІН ВЛАСТИВОСТЕЙ ОБ'ЄКТІВ ПІД ВПЛИВОМ КОМБІНОВАНИХ ПОЛІВ РІЗНОЇ ФІЗИЧНОЇ ПРИРОДИ ТА ЇХ ПРАКТИЧНЕ ЗАСТОСУВАННЯ

5.1 Вплив електромагнітного випромінювання на біологічні об'єкти

Технологічне використання електромагнітних полів НВЧ діапазону досить широке, наприклад, нагрівання, сушіння, активізація хімічних процесів та інше, і давно перейшло з стадії досліджень в промислову, що і висвітлено в різних публікаціях [87–96].

Останнім часом значно підвищився інтерес до вивчення впливу електромагнітних полів на біологічні об'єкти з метою виявлення як факторів корисних для практичного використання в технологіях, так і негативних впливів.

Дослідженням впливу електромагнітних полів на біологічні об'єкти присвячена численна бібліографія, проте до сих пір не має теоретичного обґрунтування, які давали б цілісну картину біохімічних змін під дією полів. Встановлено, що вплив електромагнітного поля на біологічні об'єкти має комплексний характер та впливає на біологічні об'єкти навіть при дуже низькій інтенсивності, і діапазон ефективних частот дуже широкий. Багато авторів, наприклад [97, 98], розглядає зміни структуризації води під впливом енергії електромагнітних коливань, як один з найважливіших факторів механізму впливу електромагнітних хвиль на біологічні об'єкти. Інші в своїх роботах [100 – 102] відмічаючи широкий діапазон ефективних частот електромагнітних коливань пояснюють це явище їх резонансним впливом на молекули білка, які виступають в ролі рецепторів. Відмічається, що взаємодія з електромагнітним полем приводить до збудження коливань складних просторових молекул білку з численними боковими відгалуженнями і тому діапазон її власних частот широкий. Дисипативний зв'язок між осциляторами забезпечує синхронізацію цих коливань і за рахунок цього зростає амплітуда цих коливань навіть у випадку низької інтенсивності електромагнітного випромінювання. В результаті

накопичення енергії виникає конформаційний перехід біомолекули. Функціональна активність молекул залежить від того, в яких із конформаційних станів вони перебувають.

В ряді робіт пояснюється велику ширину ефективних частот дисипативним та стохастичним резонансом [103, 104].

В роботі [105] наведені результати експериментальних досліджень впливу електромагнітного випромінювання на активність амілази в проростках зерна.

Найбільш важливі з погляду біології й медицини ефекти електромагнітного випромінювання на молекулярному й клітинному рівнях наведено в таблиці 5.1 [106].

5.2 Вплив резонансної магнітоакустичної обробки на фізичні властивості води

В розділі розглянутий механізм впливу на воду і водне середовище резонансної магнітоакустичної обробки (РМАО), яка може включати впливи акустичних, магнітних, електричних коливань та вплив кавітації, що приводить до зміни її фізичних властивостей. При цьому припускається, що дія РМАО на воду в першу чергу пов'язана з впливом на молекулярні асоціати молекул води, зв'язаних слабкими дипольними зв'язками. При впливі змінюється середнє число водневих зв'язків на одну молекулу та число молекул в іонних асоціатах, порушуються зовнішні іонні та ковалентні зв'язки. Утворення й розрив ковалентних та іонних зв'язків супроводжується перебудовою електронних оболонок взаємодіючих атомів. При замиканні й розриві дипольних і водневих зв'язків, відбувається не перебудова електронних оболонок, а головним чином їхня деформація.

Можна припустити, що РМАО впливає на поліморфні модифікації води й на їхні поліморфні перетворення. При впливі РМАО у воді інтенсифікуються фазові переходи другого роду, при яких об'єм і густина речовини не змінюються, а стрибкоподібно змінюються теплоємність, в'язкість, теплопровідність,

коефіцієнт термічного розширення в результаті чого у воді відбуваються зміни процесів теплообміну. Резонансна магнітоакустична обробка приводить до зсуву температурних переходів води та зміні її діелектричних і механічних властивостей.

Таблиця 5.1

Об'єкт впливу	Діапазон частот	Наслідку впливу електромагнітного випромінювання
Біомакромолекули	1 – 150 ГГц	<ol style="list-style-type: none"> 1. Конфірмаційні переходи в білкових молекулах. 2. Зміна функціональної активності: <ol style="list-style-type: none"> а) ферментів; б) транспортних білків; в) іонних каналів. 3. Розрив ланцюгів ДНК. 4. Зміна третинної структури хромосом. 5. Модифікація траскрибірування генетичної інформації
Мембрани	1 – 150 ГГц	<ol style="list-style-type: none"> 1. Зміна розподілу зарядів на поверхні мембран. 2. Модуляція стійкості до зовнішніх факторів. 3. Зміна спорідненості до біомолекул. 4. Зміна проникності для іонів і біомолекул.
Клітини	0,1 – $2 \cdot 10^{11}$ Гц	<ol style="list-style-type: none"> 1. Зміна метаболізму. 2. Модуляція росту й розподілу. 3. Модифікація функціональної активності. 4. Ініціація апоптоза.
Амілаза	1 – 150 ГГц	Зміна активності в процесі пророщування насіння.

Проведені експерименти підтвердили зміни діелектричної проникності, електричної провідності, водневого показника, коефіцієнту теплопровідності та порогу кавітації.

При різних видах РМАО одночасно з реструктуризацією й предисоціацією рідини відбувається порушення її термодинамічної рівноваги й зміна больцманівського розподілу населеності магнітних енергетичних рівнів. Змінюється співвідношення концентрацій право- і лівообертальних магнітних моментів і результуюча магнітна поляризація рідини. Ефект обертання площини магнітної поляризації може бути використаний для оперативного контролю й оцінки ефективності РМАО.

Дослідження впливу РМАО на воду та водні середовища розглянуті в роботі автора [23] та практичне використання такої обробки в різних галузях, наприклад, в сільському господарстві при передпосівній обробці насіння злакових культур, зрошуванні сільсько-господарчих культур та інше.

5.2.1 Вплив окремих видів резонансної магнітоакустичної обробки на властивості води

5.2.1.1 Резонансна обробка акустичними хвилями звукових частот. При обробці акустичними хвилями звукових частот відбуваються механічні та спін-фононні впливи на рідину.

Механічний вплив спричиняє коливання зі звуковою частотою всіх часток, молекул і асоціатів рідини в доповненні до їх теплових коливань. Це приводить до температурних змін в'язкості, щільності та інших параметрів рідини. Механічний вплив змінює структуру води та процеси дисоціації. Відбувається впорядкування структури рідини, зменшення розмірів іонних асоціатів, утвореної водневими зв'язками. Очевидно, що механічний вплив ефективний при малих вихідних температурах рідини й маскується при її нагріванні.

Спін-фононний вплив спричиняє коливання спінів протонів спінами звукових фононів. При цьому коливання спінів протонів збуджує прецесію їхніх

магнітних моментів. Максимальна ефективність такого впливу має місце при збігу частот збудження із частотою протонного магнітного резонансу. При цьому магнітні моменти протонів мають максимальний кут прецесії, а міцність їхніх магнітних зв'язків з іншими частками - мінімальна. Предисоціація сприяє відриву протонів від молекул, посиленню процесів дисоціації й змінам молекулярної сітки, утвореної водневими зв'язками. При цьому за рахунок розриву слабких і утворення нових осередків відбувається перебудова сітки. Крім того в рідині утворюються кислі й лужні іони.

5.2.1.2 Резонансна обробка акустичними хвилями ультразвукових частот.

При обробці акустичними хвилями ультразвукових частот також відбуваються механічні та спін-фононні впливи на рідину. Ультразвуковий механічний вплив еквівалентно додатковому нагріванню рідини й впливає на всі зв'язки її молекул, змінюючи й структуру, і процеси дисоціації. Однак при цьому енергія ультразвукових фононів приблизно на два порядки більше, що обумовлює їх більшу теплову добавку й ефективність. При спільному використанні ультразвукового й звукового впливів їхні теплові добавки мають однаковий знак і аддитивно складаються.

Ультразвуковий спін-фононний вплив збуджує коливання магнітних моментів електронів ультразвуковими фонами. Максимальна ефективність такого впливу має місце при збігу частот збуджуючих коливань із електронного магнітного резонансу. При цьому магнітні моменти електронів мають максимальний кут прецесії, а міцність їхніх магнітних зв'язків з іншими частками – мінімальна. Така предисоціація сприяє ослабленню міцності ковалентних зв'язків ОН і посиленню процесів дисоціації по гетеролітичному способі з утворенням іонів H^+ і OH^- . Вона також сприяє відриву електронів від іонів OH^- і перетворенню їх у радикали ОН.

5.2.1.3 Резонансна акустична кавітаційна обробка.

Суть кавітації полягає в тім, що під дією досить інтенсивної акустичної хвилі рідина розривається, утворюючи дрібні бульбашки, заповнені газом і паром.

Утворенню кавітаційних бульбашок сприяють зародки – тверді включення, пухирці газу та інше, які завжди присутні в рідині. Час життя кавітаційних бульбашок невеликий: за фазою розрідження рідини слідує фаза стискування, що приводить до колапсу кавітаційних бульбашок. При цьому в середині бульбашки виникає величезний тиск, що породжує в рідині ударну хвилю. Колапс окремої бульбашки не дає відчутного ефекту, але кавітаційних бульбашок багато й всі вони в сукупності здатні впливати на воду й водне середовище та виявляти руйнуючий вплив на дисперсні тверді включення.

При кавітації в рідині виникають сильні гідродинамічні збудження й інтенсивне шумове випромінювання акустичних хвиль та виділяється теплова енергія. Тому кавітаційна обробка рідини є різновидом акустичної ультразвукової обробки, і спричиняє механічні та спін-фононне впливи. Так як і при ультразвуковій обробці, вплив кавітації еквівалентний додатковому нагріванню рідини й впливає на всі зв'язки її молекул, змінюючи й структуру, і процеси дисоціації.

Додатково, за рахунок колапсу бульбашок дія кавітації локально супроводжується розігрівом рідини, деструкцією кластерів і молекул, диспергуванням твердих часток та ерозією їхньої поверхні, виділенням газу, що містить атомарну та іонізованих компонент та інше. Змінюючи умови протікання кавітації, можна підсилювати або послабляти різні кавітаційні ефекти.

Кавітаційна обробка також сприяє протіканню й прискоренню хімічних реакцій у рідині. Наприклад, у роботі [25] наведені результати досліджень кавітуючої води методом спінових уловлювачів, які безпосередньо довели наявність у продуктах її розщеплення радикалів Н та ОН, а також продуктів їхньої рекомбінації у вигляді перекису водню H_2O_2 і кислот: азотистої HNO_2 і азотної HNO_3 .

5.2.1.4 Резонансна обробка змінними магнітними полями. При обробці змінними магнітними полями відбуваються магнітні впливи на магнітні моменти парамагнітних часток рідини, що викликають їхню нутацію й прецесію.

Максимальна ефективність такого впливу має місце при збігу частоти порушення із частотою протонного або (і) електронного магнітного резонансу в діапазоні звукових або (і) ультразвукових частот. При цьому, як і у випадку спин-фононного впливу, магнітні моменти протонів і електронів одержують максимальний кут прецесії, а міцність їхніх магнітних зв'язків з іншими частками стає мінімальною. Відбувається предисоціація рідини, що при незмінних температурі й тиску, сприяє ослабленню міцності водневих і ковалентних зв'язків і посиленню процесів дисоціації. Крім того, нутація магнітних моментів часток викликає їхню магнітну поляризацію й структурну перебудову водневих зв'язків рідини. У речовині з водневими зв'язками магнітний вплив може викликати, наприклад, перескок протонів з одного положення в інше. У вуглеводнях та ін., може мати місце поворот (перекидання) не цілих молекул, а окремих їхніх частин.

Слід зазначити, що магнітний вплив на магнітні моменти є прямим і, отже, більше ефективним, у частині предисоціації, чим непряме спин-фононне. Крім того, при магнітній обробці одночасно з відзначеними змінами рідини, відбувається порушення її термодинамічної рівноваги й зміна больцмановського розподілу населеності магнітних енергетичних рівнів. Однак при обробці змінними магнітними полями відсутній механічний фактор. Тому акустичну та магнітну обробки доцільно вважати такими, що взаємно доповнюють.

5.2.1.5 Резонансна обробка змінними електричними полями. При резонансній обробці змінними електричними полями передбачається електромеханічний вплив на заряджені парамагнітні частки. При збігу частот впливу й магнітного резонансу часток обробка змінним електричним полем стає резонансною (спін-кулонним). Як було показано вище, умова протонного резонансу виконується в діапазоні звукових частот, а електронного - у діапазоні ультразвукових частот.

Електромеханічний вплив обумовлений дипольною та іонною поляризацією рідини в електричному полі.

Дипольна поляризація характерна для полярних рідин. Сутність цього виду поляризації полягає в повороті (орієнтації) у напрямку електричного поля молекул і агрегатів, що мають постійний електричний момент. При резонансній обробці змінними електричними полями подібно механічному акустичному впливу дипольна поляризація еквівалентна додатковому нагріванню рідини. Одночасно вона вносить деяку впорядкованість у положення полярних молекул і структуру дипольних асоціатів рідини.

Іонна поляризація полягає в зсуві друг щодо друга різнойменно заряджених іонів рідини під дією змінного електричного поля резонансної частоти. Коливання іонів приводять до додаткового нагрівання рідини подібно коливанням дипольних молекул. Цей тепловий вплив впливає на всі зв'язки рідини, змінюючи, і її структуру й процеси дисоціації. Крім того іонна поляризація приведе до міграції та видаленню іонів з області взаємодії, істотно змінюючи хімічний склад і властивості обробленої рідини.

Одночасно змінне електричне поле впливає на полярні молекули й іоні-радикали рідини, що мають некомпенсовані магнітні моменти. До таких об'єктів відносяться протони, електрони, гідроксильні групи, мономолекули води та ін. Переміщення полярних молекул і іонів-радикалів у геомагнітному полі викликає коливання їхніх магнітних моментів. Оскільки ці коливання відбуваються із частотою прецесії, те подібно спин-фононному впливу вони супроводжуються предисоціацією рідини, що при незмінних температурі й тиску, сприяє ослабленню міцності водневих і ковалентних зв'язків і посиленню процесів дисоціації й структурної перебудови.

Слід зазначити, що резонансна магнітоакустична обробка води змінними електричними полями має технологічні особливості, обумовлені тим, що нержавіюча сталь в апаратах з електродами не стійка до анодного розчинення. При подачі електричного струму, анод розчиняється й іони нікелю, хрому, ванадію, молібдену, заліза та ін., переходячи у воду, отруюють її. Анод повинен

бути виконаний або з матеріалів, що не руйнуються, типу платини, або інших екологічно чистих матеріалів.

5.2.1.6 Зміни параметрів води при резонансній магнітоакустичній обробці електромагнітними хвилями. Резонансна обробка електромагнітними хвилями викликає зміни у воді, обумовлені сумарним впливом змінних магнітних і електричних полів. Оскільки ці види обробки впливають на магнітні й дипольні моменти часток, той їхній сумарний вплив є найбільш ефективним для предисоціації та структурної перебудови рідини.

Як було показано вище, при резонансній магнітоакустичній обробці води змінюються: діелектрична проникність, електрична провідність, водневий показник, коефіцієнт теплопровідності та поріг кавітації.

5.2.1.7 Зміни діелектричної проникності. Молекула води несиметрична й має постійний електричний момент $P_{\text{п}} = 6.1 \cdot 10^{-30}$ Кл·м, здатний до переорієнтації в електричному полі. Якщо в одиниці об'єму рідини містяться N незв'язаних молекул з поляризованістю α кожна, то її діелектрична проникність рівна [110]

$$\varepsilon = 1 + N\alpha/\varepsilon_0, \quad (5.1)$$

де $\varepsilon_0 = 8,854 \cdot 10^{-12}$ Ф/м — електрична постійна.

Відомо, що ε у воді може змінюватися в межах 1,78...81.

Зміна ε при РМАО свідчить або про зміну поляризованості α молекул та інших часток, або про зміну їхньої концентрації в одиниці об'єму рідини N . Зростання ε також може свідчити про збільшення щільності рідини при РМАО.

При зміні ε також змінюються такі параметри речовини:

1. Електрична сприйнятливість

$$\chi = \varepsilon - 1 = N\alpha/\varepsilon_0, \quad (5.2)$$

Для води $\varepsilon = 80...0$.

2. Поляризованість

$$P = NaE, \quad (5.3)$$

де E – напруженість електричного поля.

3. Ємність ділянки ізоляції

$$C = \varepsilon_0 \varepsilon A [\Phi], \quad (5.3)$$

де A - наведена довжина ділянки ізоляції [м].

По зміні ємності конденсатора, заповненого рідиною, звичайно визначають значення ε .

4. Електромагнітна енергія, що накопичується ділянкою ізоляції

$$W = \varepsilon_0 \varepsilon A U^2 / 2 [\text{Дж}], \quad (5.4)$$

де U – напруга між електродами [В].

4. Швидкість електромагнітної хвилі в рідині

$$V = \frac{c}{\sqrt{\varepsilon\mu}}, \text{ при } \mu = 1 \quad \varepsilon \quad v = \frac{c}{\sqrt{\varepsilon}}, \quad (5.5)$$

де c - швидкість світла.

5. Показник переломлення рідини

$$n^2 = \varepsilon\mu, \text{ при } \mu = 1 \quad n^2 = \varepsilon. \quad (5.6)$$

Для води $n^2 = 1,78...81$.

7. Хвильовий опір (імпеданс)

$$z_o = \frac{E}{H} = 377 \sqrt{\frac{\mu}{\epsilon}}, \text{ при } \mu = 1 \quad Z_0 = 377 \sqrt{1/\epsilon} \quad z_o = 377 \sqrt{\frac{1}{\epsilon}}. \quad (5.7)$$

Для води $Z_0 = 202,75 \dots 41,5$ Ом.

5.2.1.8 Зміни електричної провідності. Електропровідність рідин обумовлена іонами, що утворюються при дисоціації молекул самої рідини або наявністю в ній домішок.

Температурні залежності концентрацій іонів n і питомої провідності γ рідин мають однаковий вигляд

$$n = n_0 \exp(-W/k) = 10^{-p}, \quad (5.8)$$

$$\gamma = Nq(\mu_+ + \mu_-) = A \exp(-W/k), \quad (5.9)$$

де W - енергія дисоціації; q - заряд іона; T - температура, К; k - постійна Больцмана; μ_+ і μ_- — рухливості позитивних і негативних іонів; n_0 і A — константи.

Питомий опір рідин $\rho = 1/\gamma$ з ростом температури зменшується

$$\rho = B \exp(W/k), \quad (5.10)$$

де B - константа.

За аналогічним законом змінюється й в'язкість рідини. Тому залежність $\rho(T)$ рідин пояснюється як дисоціацією молекул, так і зміною в'язкості.

Таким чином, зміни питомої провідності рідини при РМАО свідчить про зміни в'язкості й концентрації іонів дисоціації n .

При зміні n також змінюється водневий показник рідини рН – величини, що характеризує концентрацію водневих іонів H^+ і чисельно рівна десятковому логарифму цієї концентрації зі зворотним знаком

$$pH = - \lg[H^+] \text{ [моль/л]}. \quad (5.11)$$

5.2.1.9 Зміни коефіцієнту теплопровідності. Експерименти показують, що при РМАО змінюється коефіцієнт теплопровідності води λ , що визначається як кількість тепла, що пройшло через одиницю площі за одиницю часу при градієнті температури один ДО/м.

Можна припустити, що це обумовлено взаємозалежними змінами в'язкості й теплопровідності рідини, що відбуваються в результаті впорядкованості орієнтації молекул при РМАО. Тому що мікроскопічні механізми теплопровідності й в'язкості дуже схожі, то при великій в'язкості завжди виникає більша теплопровідність. Це обумовлено тим, що теплопровідність визначається швидкістю передачі теплової енергії, а в'язкість – швидкістю передачі кількості руху.

В'язкість рідини звичайно зменшується з підвищенням температури

$$\eta = A(T) \exp(W/k), \quad (5.12)$$

де $A(T)$ – слабо залежний від температури множник; W – енергія активації в'язкої течії.

За аналогічним законом зменшується й теплопровідність рідини. Отже, якщо при експериментах швидкість передачі теплової енергії після РМАО росте, то її вплив еквівалентно охолодженню рідини й упорядкуванню її структури. Експериментально це припущення підтверджується зниженням ефективності РМАО при підвищенні температури рідини вище 35.

Крім того, як показано в [26], існує залежність між коефіцієнтом теплопровідності рідини λ , щільністю d , питомою теплоємністю c , швидкістю звуку V і середньою відстанню між молекулами L за допомогою вираження

$$\lambda = c d V L. \quad (5.13)$$

Ця формула відображає залежність λ від агрегатного стану рідини, її атомно-молекулярної будови, складу та ін.

Таким чином, крім в'язкості в рідини в результаті РМАО міняється щільність d , коефіцієнт теплопровідності λ , питома теплоємність c , швидкість звуку V і середня відстань між молекулами L .

5.2.1.10 Зміна порогу кавітації. Поріг кавітації це границя між двома режимами гідродинамічних процесів у рідині – безкавітаційним і кавітаційним. Його зміни характеризують зміну структури рідини.

Поріг акустичної кавітації характеризують мінімальним значенням амплітуди звукового тиску P_k , при якому виникає кавітаційний процес. Його величина залежить від температури, тиску, вмісту включень в рідині, турбулентності та інших параметрів.

Експерименти показують, що при РМАО водопровідної води зі збільшенням часу експозиції поріг кавітації зростає. Можна припустити, що це пов'язане зі зменшенням числа зародків кавітації й зміною структури води, викликаних вмістом включень в її складі.

5.3. Виводи до розділу 5

1) Розроблений електродний нагрівач зі стабільними характеристиками. Новизна й актуальність підтверджені патентом України.

2) В результаті досліджень електромагнітного впливу на біологічні середовища, що містять водяні включення розроблено конструкцію складного випромінювача у вигляді вібраторної антени для формування електромагнітного поля.

3) Розроблено, обґрунтовано і виготовлений мікрохвильовий калориметричний перетворювач, який здатний працювати на підвищених рівнях потужності з підвищеною швидкодією. Конструкція запатентована.

6 ПОБУДОВА ТЕПЛОФІЗИЧНИХ МОДЕЛЕЙ ПРОЦЕСУ АКТИВНОГО ТЕПЛОВОГО КОНТРОЛЮ

6.1 Теплофізична модель біологічних середовищ

Метою теплової контролю (ТК) біологічних об'єктів (БО) є виявлення дефектів їх внутрішньої структури. Критерієм визначення дефекту є поява на поверхні БО локальної теплової аномалії ΔT , надлишкового температурного поля, яке є наслідком неоднорідності структури БО.

Існуючі теплофізичні моделі містять ряд спрощувань, які знижують їх адекватність (мають обмеження по температурі нагрівання T_{\max} , не враховується теплопровідність дефекту, або не у повній мірі враховуються особливості внутрішньої структури біологічних середовищ та ін.). При цьому критерієм оптимізації є максимум корисного сигналу. Дослідження показують, що при наявності шумів типу флуктуацій випромінювальної спроможності $\Delta \epsilon$ об'єкта контролю та нерівномірності нагрівання Δq критерієм оптимізації повинно бути відношення сигнал/шум. Тому потрібне удосконалення теплофізичної моделі та оптимізація ТК по критерію відношення сигнал/шум на основі цієї моделі.

Була розроблена геометрична модель БО у вигляді трьохшарової пластини, яка складається з обкладенок, між котрих знаходиться біологічний матеріал із неоднорідністю (рис. 6.1). Дефект типу «непроклій» між обкладенкою та сотопластом моделюється повітряним шаром. При описанні процесу ТК була використана циліндрична система координат (r, φ, z) , де r - радіальна координата, z - вертикальна координата. Углова координата φ - виключена, тому що представлена модель ОК має центральну симетрію.

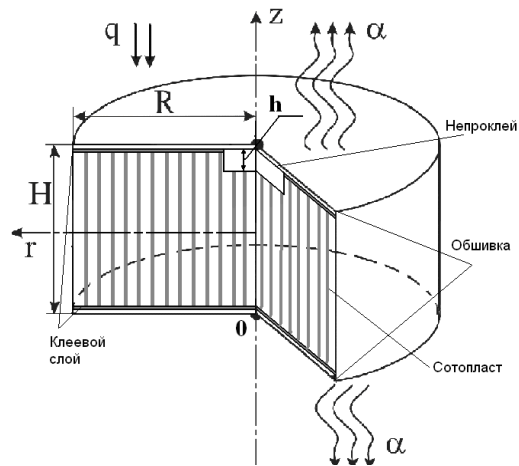


Рисунок 6.1. Объект контролю: R – радіус БО; H – товщина ОК; q – потужність нагріву; α – коефіцієнт тепловіддачі поверхні, h – товщина дефекту.

У запропонованій теплофізичній моделі неоднорідний по структурі шар сотопласту (рис. 6.2) був змінений однорідним еквівалентним шаром з усередненими теплофізичними характеристиками (ТФТ): питомою теплоємністю C' , густиною ρ та коефіцієнтом теплопровідності λ' . Для визначення усереднених значень були використані наступні співвідношення:

$$c' = \langle c \rangle = \frac{V_1 C_1 + V_2 C_2}{V_1 + V_2},$$

$$\rho' = \langle \rho \rangle = \frac{V_1 \rho_1 + V_2 \rho_2}{V_1 + V_2},$$

$$\lambda' = \langle \lambda \rangle = \frac{V_1 \lambda_1 + V_2 \lambda_2}{V_1 + V_2},$$

де V_1 - об'єм полімера; V_2 - об'єм повітря; c_1, λ_1, ρ_1 - ТФТ полімера; c_2, λ_2, ρ_2 - ТФТ повітря.

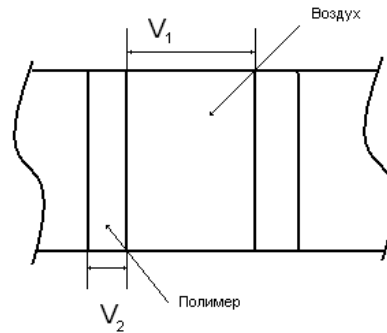
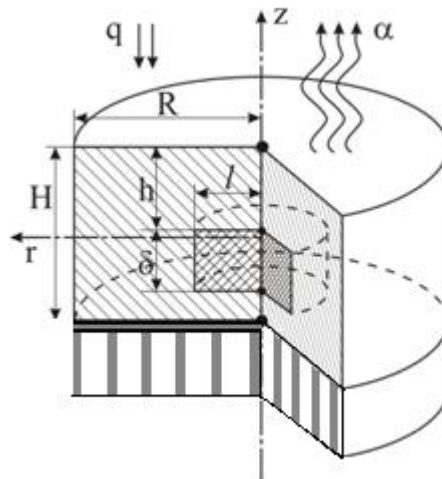


Рисунок 6.2. Схема шару сотопласта, який складається з осередків з полімерними стінками, які заповнені повітрям

Характерною особливістю даної моделі є те, що ТФХ залежать від просторових аргументів не тільки скачкоподібно при перетинанні скрізь границю шарів, але і вони різняться у залежності від напрямку (нормальний, радіальний напрямок).

У випадку розглядання дефекту типу «розшарування» була розроблена наступна теплофізична модель (рис. 6.3).



H - товщина пластини, м; R - радіус пластини, м; l - радіус дефекту, м; густина потоку тепла, $\text{Вт}/\text{м}^2$; коефіцієнт тепловіддачі, $\text{Вт}/\text{м}^2$; q - густина речовини, $\text{кг}/\text{м}^3$;

δ - розкриття дефекту, м; h - глибина залягання дефекту, м.

Рисунок 6.3. Теплофізична модель моделі БО з дефектом типу «розшарування»

Об'єкт контролю представляє собою необмежену пластину з неоднорідністю (дефектом), представленою у циліндричній системі координат, при цьому продольний радіус пластини значно більше її товщини. Дефект являє собою область циліндричної форми, яка відрізняється від основного матеріалу по ТФТ. На зовнішніх границях ОК діють граничні умови 2-го та 3-го роду, які відображають реальні умови контролю: нагрівання ОК зовнішнім тепловим потоком q та його охолодження за рахунок теплообміну з навколишнім середовищем (α - коефіцієнт тепловіддачі).

Математичний опис моделі базується на рішенні диференціального рівняння нестационарної теплопровідності, яке в загальному випадку має вигляд

$$\operatorname{div}(\lambda(\vec{r}, T)\nabla T(\vec{r}, t)) + q(\vec{r}, t) = c\rho \frac{\partial T(\vec{r}, t)}{\partial t}, \quad (6.1)$$

де $T(\vec{r}, t)$ – температура об'єкту контролю, яка залежить від координат точки M і часу t ; $\lambda(\vec{r}, T)$ – коефіцієнт теплопровідності (у загальному випадку він може залежити від температури T); $q(\vec{r}, t)$ – функція внутрішніх джерел тепла [$\text{Вт}/\text{м}^2$]; c – питома теплоємність [$\text{Дж}/\text{кг}\cdot\text{К}$]; ρ – густина речовини [$\text{кг}/\text{м}^3$].

Для реального процесу теплового контролю (тепловий дефектоскопії) рівняння (6.1) можна спростити з урахуванням таких факторів: внутрішні джерела відсутні, а коефіцієнт теплопровідності не залежить від температури, тому прогріву ОК не перевищує 100°C . З урахуванням цього отримуємо

$$\left(\lambda(\vec{r}) \frac{\partial^2 T}{\partial x^2} + \frac{\partial \lambda(\vec{r})}{\partial x} \frac{\partial T}{\partial x} \right) + \left(\lambda(\vec{r}) \frac{\partial^2 T}{\partial y^2} + \frac{\partial \lambda(\vec{r})}{\partial y} \frac{\partial T}{\partial y} \right) + \left(\lambda(\vec{r}) \frac{\partial^2 T}{\partial z^2} + \frac{\partial \lambda(\vec{r})}{\partial z} \frac{\partial T}{\partial z} \right) = c\rho \frac{\partial T(\vec{r}, t)}{\partial t} \quad (6.2)$$

Рівняння (6.2) є однорідним лінійним диференціальним рівнянням другого порядку параболічного типу (так як $\lambda(r) \geq 0$). Для перетворення диференціального рівняння нестационарної теплопровідності (6.2) у циліндричні координати використовуємо:

пряме перетворення

$$\begin{cases} r = \sqrt{x^2 + y^2} \\ \varphi = \operatorname{arctg}\left(\frac{y}{x}\right) \\ z = z \end{cases}$$

зворотне перетворення

$$\begin{cases} x = r \cos \varphi \\ y = r \sin \varphi \\ z = z \end{cases}$$

Після проведення перетворень рівняння (6.2) приймає вигляд

$$\left(\lambda \frac{\partial^2 T}{\partial r^2} + \frac{\partial \lambda}{\partial r} \frac{\partial T}{\partial r} + \frac{\lambda}{r} \frac{\partial T}{\partial r} \right) + \left(\frac{\lambda}{r^2} \frac{\partial^2 T}{\partial \varphi^2} + \frac{1}{r^2} \frac{\partial \lambda}{\partial \varphi} \frac{\partial T}{\partial \varphi} \right) + \left(\lambda \frac{\partial^2 T}{\partial z^2} + \frac{\partial \lambda}{\partial z} \frac{\partial T}{\partial z} \right) = c\rho \frac{\partial T(\vec{r}, t)}{\partial t} \quad (6.3)$$

Таке рівняння адекватним чином описує обрану модель (рис 6.3), за умови його вирішення при вірно обраних граничних умовах, тобто відповідних реальної процедури теплового контролю.

Вище вже зазначалося, що такими є граничні умови 2-го і 3-го роду на зовнішніх поверхнях БО

$$z = H \quad \lambda \frac{\partial T(\vec{r}, t)}{\partial r} = \alpha (\vec{r}, t) T_{\text{середы}} - q(\vec{r}, t) \quad (6.4)$$

$$z = 0 \quad -\lambda \frac{\partial T(\vec{r}, t)}{\partial r} = -\alpha (\vec{r}, t) T_{\text{середы}} \quad (6.5)$$

де $T(r, t)$ - температура об'єкта контролю; λ - коефіцієнт теплопровідності (у загальному випадку може залежати від температури), Вт / м • К; $q(r, t)$ - щільність потоку тепла, Вт / м²; α - коефіцієнт тепловіддачі, Вт / м² • К; ρ - щільність речовини, кг / м³.

Забезпечення урахування теплопровідності через дефект досягається використанням умови неперервності на кордонах дефектної області

$$-\lambda_1(\vec{r}, T, t) \left(\frac{\partial T_1(\vec{r}, t)}{\partial n} \right) = -\lambda_2(\vec{r}, T, t) \left(\frac{\partial T_2(\vec{r}, t)}{\partial n} \right). \quad (6.6)$$

У зв'язку з тим, що рішення диференціального рівняння нестационарної теплопровідності (6.6) в аналітичному виді неможливо (без звуження

спільності), особливо з урахуванням особливостей створеної моделі, був використаний метод кінцевих різниць.

При побудові кінцево-різницевої схеми був обраний неявний варіант, за рахунок його більш високої точності.

Для цього будується просторово-часова сітка з кроком за часом і координатами відповідно $h1, h2$ (рис. 6.4).

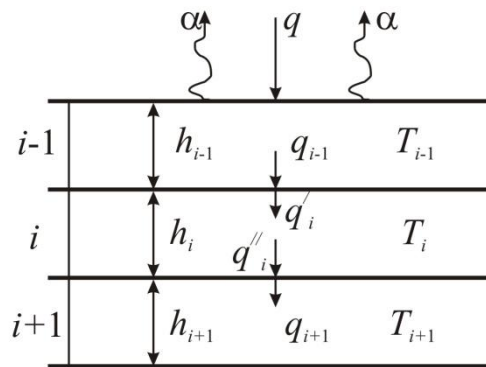


Рисунок 6.4 Схема вузлової сітки h - координата по полярній осі; $T_{i+1}^{t+\Delta t}$ - температура у вузлі $i+1$ в момент часу $t+\Delta t$; T_i^t - температура у вузлі $i+1$ в момент часу t

Відповідно обраному методу рівняння (6.6) апроксимується в кінцево-різницеву схему

$$\begin{aligned} & \frac{\Delta t}{c_i \rho_i} \frac{2\lambda_i}{h_i} \frac{\lambda_{i-1}}{\lambda_i h_{i-1} + \lambda_{i-1} h_i} T_{i-1}^{t+\Delta t} - \left(1 + \frac{\Delta t}{c_i \rho_i} \frac{2\lambda_i}{h_i} \left(\frac{\lambda_{i+1}}{\lambda_{i+1} h_i + \lambda_i h_{i+1}} + \frac{\lambda_{i-1}}{\lambda_i h_{i-1} + \lambda_{i-1} h_i} \right) \right) T_i^{t+\Delta t} + \\ & + \frac{\Delta t}{c_i \rho_i} \frac{2\lambda_i}{h_i} \frac{\lambda_{i+1}}{\lambda_{i+1} h_i + \lambda_i h_{i+1}} T_{i+1}^{t+\Delta t} = -T_i^t - \frac{\Delta t}{c_i \rho_i} q_i, \end{aligned} \quad (6.7)$$

де λ_i - коефіцієнт теплопроводності у вузлі сітки i (λ_{i+1} у вузлі $i+1$, λ_{i-1} у вузлі $i-1$), c_i - теплоємність ОК у вузлі i (c_{i+1} у вузлі $i+1$, c_{i-1} у вузлі $i-1$); ρ_i - щільність ОК у вузлі i (ρ_{i+1} у вузлі $i+1$, ρ_{i-1} у вузлі $i-1$); h_i - величина шага по координаті h у вузлі i (h_{i+1} у вузлі $i+1$, h_{i-1} у вузлі $i-1$), де крок може змінюватися як лінійно, так і логарифмічно; $T_{i+1}^{t+\Delta t}$ - температура у вузлі $i+1$ у момент часу $t+\Delta t$; T_i^t - температура у вузлі $i+1$ у момент часу t .

Таким чином, задача зводиться до чисельного вирішення диференціального рівняння в циліндричеській системі координат з непостійними коефіцієнтами по радіусу.

6.2. Програмна реалізація методу

Для проведення розрахунків за наведеною в підрозділі 1 кінечно-різницевої схемою (6.1) - (6.7) була розроблена оригінальна програма «ТермоPro_3.0_dif». Особливістю даного програмного продукту є можливість моделювання полів градієнтів. Необхідність створення нової спеціалізованої програми для моделювання процедури АТК обумовлена відсутністю в універсальних програмних пакетах для розрахунку теплових полів можливості автоматичного варіювання параметрів моделі. Для реалізації розроблених алгоритмів в якості алгоритмічної мови програмування було обрано середу MicrosoftVisualStudio.NET. Це середовище поєднує в собі можливості, гнучкість і зручності розробки таких мов, як C #, Delphi та ін. Найбільш добре підходить для програмування в операційній системі Microsoft Windows XP. Крім того, дана середа найбільш повно поєднує в собі основні принципи об'єктно-орієнтованого підходу, володіє потужною візуальною системою розробки графічних інтерфейсів користувача за стандартом АКС і системою створення та підтримки контекстно-незалежної допомоги. Середа MicrosoftVisualStudio.NET дозволяє створювати додатки, що функціонують в багатозадачних і мережевих операційних системах, і містить широкий набір математичних операцій і додаткових модулів з чисельними методами. В основу програми покладений алгоритм, що відображає реальну процедуру активного теплового контролю, який полягає в наступному.

Користувач задає параметри об'єкта (a_0, λ_0, H, R) , параметри дефекту, що задається $(a_d, \lambda_d, \delta, l, h)$ і коефіцієнт теплообміну α . Крім того, в число вихідних даних входить гранично допустима для матеріалу заданого об'єкта температура нагріву, що вкрай важливо для запобігання його руйнування.

Метою розрахунку є визначення величини температурного перепаду на поверхні ОК, викликаного наявністю дефекту (рис. 6.1), а також тимчасові

параметри режиму контролю, що впливають на величину. У їх числі: час нагрівання об'єкта τ_n до T_{\max} і час формування (час затримки) τ_s максимального температурного перепаду ΔT .

Таким чином, програма дозволяє розрахувати не тільки корисний сигнал від дефекту, тобто ΔT , але й визначити оптимальний режим проведення контролю за критерієм максимізації параметра ΔT .

Крім того, варіюючи значення параметрів дефекту ($a_o, \lambda_o, \delta, l, h$), можна проаналізувати вплив кожного з них на величину корисного сигналу, а, в кінцевому рахунку, - оцінити чутливість теплового методу в цілому.

Алгоритм програми являє собою подієво залежну систему, в якій кожна дія, яка виконується визначається поточним запитом користувача. Оператор, використовуючи систему ієрархічних меню, може повністю управляти обчислювальним процесом.

Структурна схема розробленої програми моделювання процесу активного теплового контролю «ТермоPro_3.0_dif» наведена на (рис. 6.5).

У програмі передбачена можливість моделювання процесів теплового контролю для різних об'єктів, при цьому використана можливість застосування стандартних програмних засобів операційної системи Windows XP, Windows 7, що дещо спрощує структуру програми і дозволяє полегшити процес створення моделі. Ця функція реалізована в підсистемі моделювання.

Реалізація алгоритмів АТК виконана в підсистемі обробки даних. На цьому етапі дані про температурний поле на поверхні контрольованого об'єкта проходять стадію обробки (прорахунок температурного поля для наступного моменту часу), результати якої передаються в підсистему виводу.

При цьому передбачено вибір способу відображення результатів контролю.

Підсистема виводу (відображення) призначена для реалізації можливості виведення інформації на екран у вигляді графіків і чисельних значень шуканих параметрів. Результати роботи програми можуть бути також збережені у файл у вигляді числових значень і таблиць.

6.3 Результати розрахунків та їх аналіз

Для оцінки працездатності моделі був обраний конкретний біологічний матеріал – кров с характерними дефектами судів.

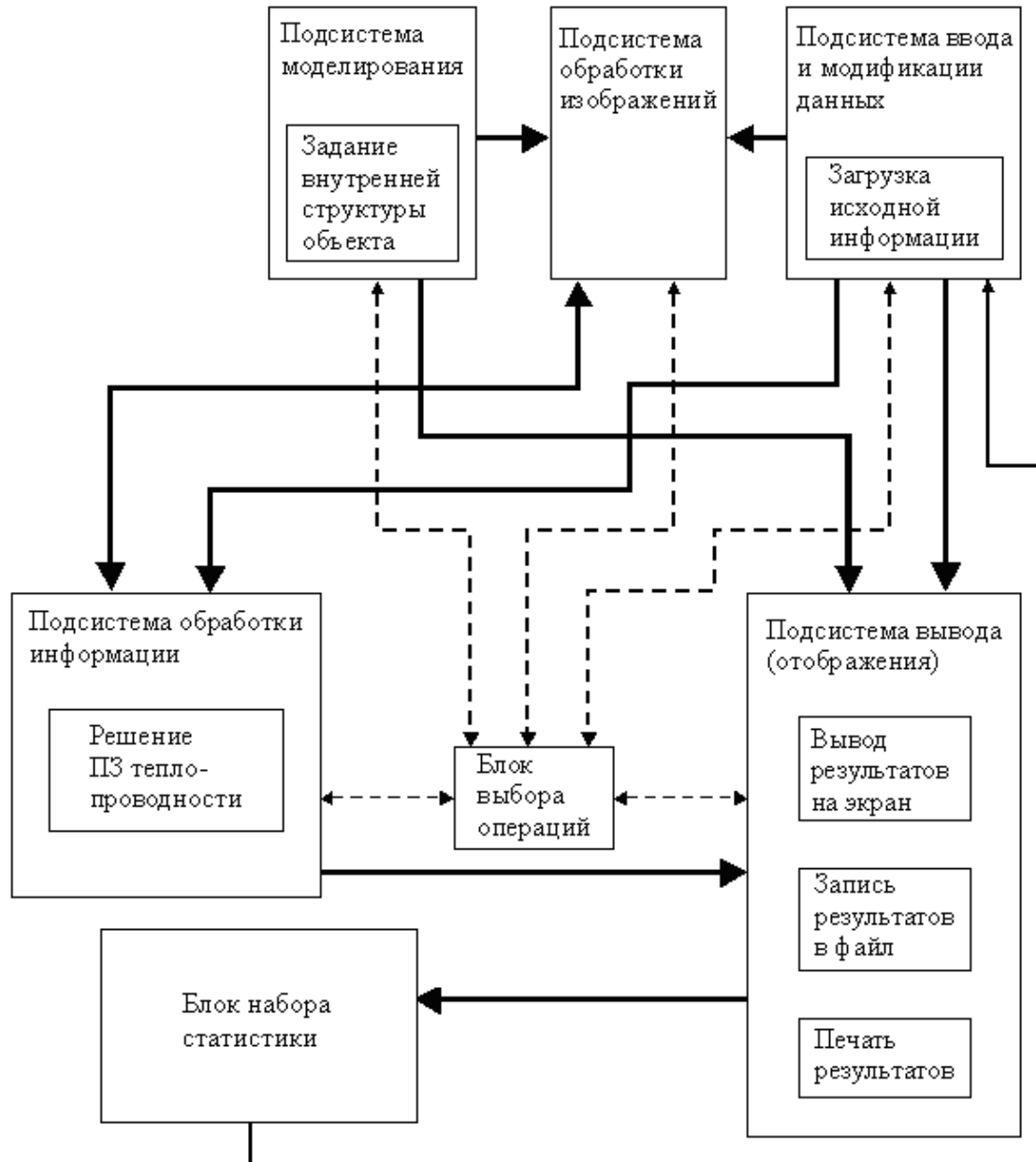
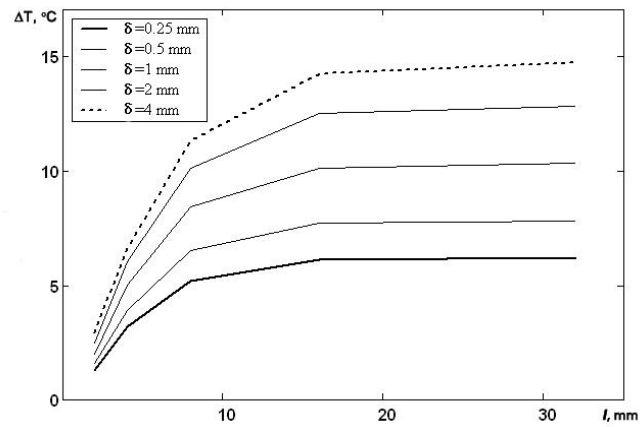
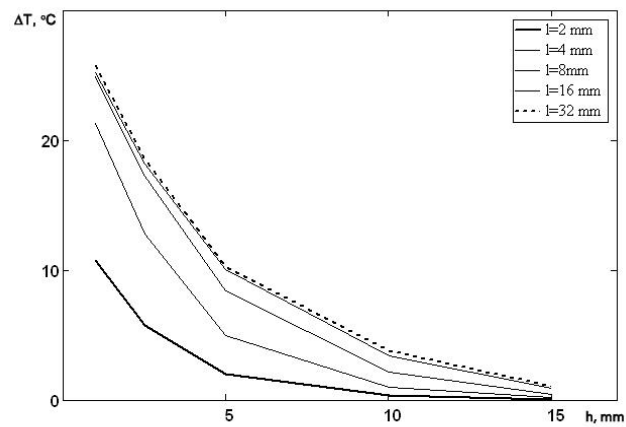


Рисунок 6.5 Структурна схема розробленої програми «TermoPro_3.0_dif». (пунктирною лінією позначені додаткові опції)

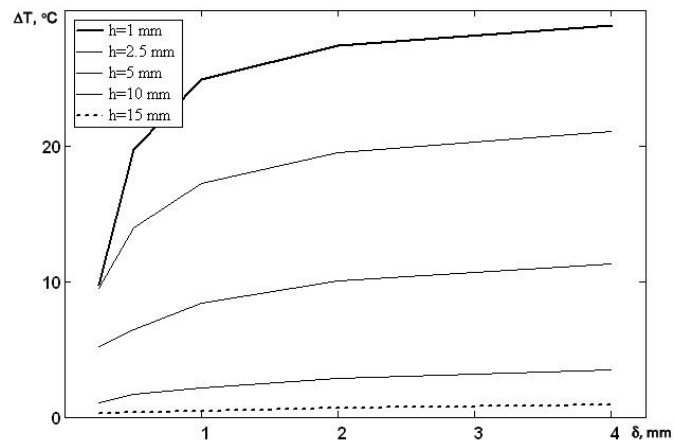
При моделюванні процедури АТК були досліджені залежності величини корисного сигналу (температурного перепаду ΔT , викликаного дефектом) від параметрів дефекту h , δ та l . Результати розрахунків наведено на рис. 6.6 (а-в).



а)



б)



в)

а) від поперечного розміру дефекту при глибині залягання $h = 5$ мм;

б) від глибини залягання дефекту при його розкритті $\delta = 1$ мм;

в) величини розкриття дефекту при радіусі дефекту $l = 8$ мм

Рисунок 6.6 Залежності величини температурного перепаду над дефектом

Отримані залежності збігаються з відомими закономірностями АТК (рис. 6.5), що підтверджує працездатність моделі. Отримані дані свідчать про те, що найбільший вплив на величину температурного перепаду надає глибина залягання дефекту, а поперечний розмір дефекту і його розкриття значно впливають на ΔT тільки при їх малих значеннях.

6.4 Теплофізична модель судів крові

Розроблений і реалізований алгоритм має досить високою гнучкістю при реалізації розрахунків по ТФ моделям, що дозволяє модифікувати процес під велику кількість об'єктів таких як потік крові в судинах.

Для перевірки гнучкості алгоритму для модифікацій була розроблені і випробувані теплофізичні моделі потоку крові в судинах, з локальним дефектом типу утоньшення стінки, і модель з таким поширеним відхиленням як наявність рідкої фази у крові, що значно знижує ефективне розтин судин.

Теплофізична модель потоку крові в судинах наведена на рис. 6.7.

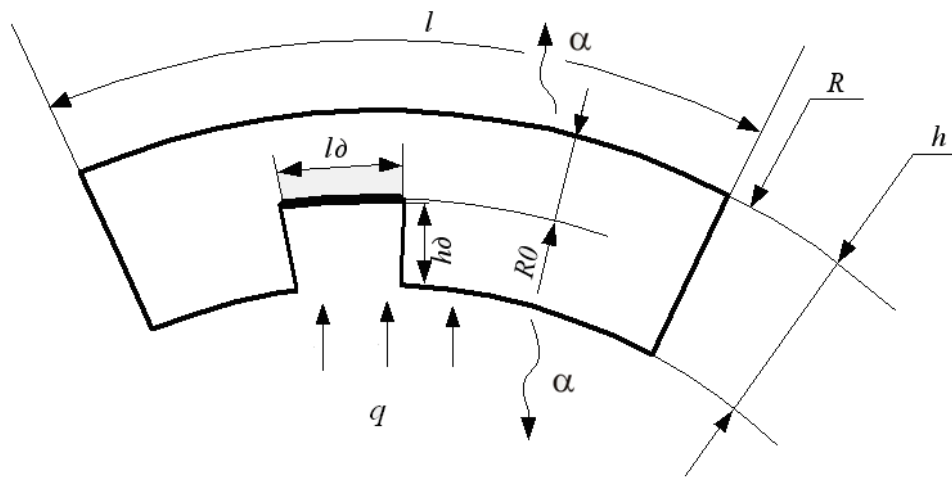


Рисунок 6.7 Об'єкт контролю

Згідно обраної моделі рішення здійснюється при граничних умовах 2-го і 3-го роду на зовнішніх поверхнях БО

$$h = H \quad \left(\lambda(\vec{r}, t) \frac{\partial T(\vec{r}, t)}{\partial n} \right) \Big|_S = \alpha \left(T(\vec{r}, t) \Big|_S - T_{\text{среды}} \right) - q(\vec{r}, t) \quad (6.8)$$

$$h = 0 \quad - \lambda(\vec{r}, t) \frac{\partial T(\vec{r}, t)}{\partial n} \Big|_S = -\alpha \left(T(\vec{r}, t) \Big|_S - T_{\text{среды}} \right) \quad (6.9)$$

де $T(\vec{r}, t)$ - температура об'єкта контролю; $\lambda(\vec{r}, t)$ - коефіцієнт теплопровідності (у загальному випадку може залежати від температури), Вт / м • К; $q(\vec{r}, t)$ - щільність потоку тепла, Вт / м²; α - коефіцієнт тепловіддачі, Вт / м² • К; ρ - щільність речовини, кг / м³; R_0 - Радіус на якому знаходиться дефект, м; R - радіус на якому знаходиться, м; h - товщина ОК, м.

Співвідношення (6,8) – (6,9) відображають реальні умови проведення ТДС, тобто нагрів ОК зовнішнім джерелом тепла і наявність теплообміну з навколишнім середовищем.

Математична модель процесу ТДС ґрунтується на рішенні диференціального рівняння нестационарної теплопровідності, записаного для обраної циліндричної системи координат:

$$\operatorname{div}(\lambda(\vec{r}, t) \nabla T(\vec{r}, t)) + q(\vec{r}, t) = c\rho \frac{\partial T(\vec{r}, t)}{\partial t}, \quad (6.10)$$

де c - питома теплоємність, Дж / кг.

Товщина труби була обрана 3 мм і 5 мм. Температура носія становила 36 0С, 39 0С, 42 0С, що відповідає реальним вимогам.

Результати комп'ютерного моделювання показані у табл.6.1.

Таблиця 6.1. Результати оцінки чутливості теплового контролю

$\Delta h, \text{мм}$	3				5					
$T_H, \text{°C}$	42				36	38	42			
$\Delta h, \text{мм}$	0,2	0,4	0,6	0,8	0,8	0,8	0,2	0,4	0,6	0,8
$\Delta T_c, \text{°C}$	2,3	3,28	4,11	6,92	0,56	1,12	0,34	0,61	1,22	2,83

Чутливість сучасних тепловізорів становить 0,1 - 0,2 °С. Достовірне визначення оператором зон з аномальною температурою, які відповідають дефекту, можливо при 3-5 кратному перевищенні сигналу над шумом.

6.5 Виводи к розділу 6

В результаті проведених досліджень можна зробити такі висновки.

1. Побудовані моделі активного теплового неруйнівного контролю для
 - біологічних матеріалів, які становлять основу живого організму;
 - судин живого організму, по яким розповсюджується кров.

Розроблені моделі враховують всі особливості названих БО і дозволяють знаходити оптимальні режими процедури активного теплового контролю. Оптимальний режим дозволяє отримати якісні зображення температурного поля (термограми), які несуть інформацію о схованих неоднорідностях БО.

2. Розроблена програма «TermoPro_3.0_dif», яка дозволяє моделювати поля температурних градієнтів для визначення величини температурного перепаду на поверхні БО, викликаного наявністю дефекту, тобто розраховує оптимальні температурні режими. Програма дозволяє варіювати теплофізичні характеристики об'єктів контролю у широких діапазонах.

3. Проведені розрахунки оптимальних режимів ТК для БО, включаючи судини.

В цілому тепловий метод завдяки своїм великим перевагам (безконтактність, велика швидкодія, безпечність для оператора, інформативність) швидко розвивається і має великі перспективи у багатьох галузях промисловості і науки.

7 ФОРМУВАННЯ ЗАСАД МЕТРОЛОГІЧНОГО ЗАБЕСПЕЧЕННЯ ЦИФРОВОЇ ГОЛОГРАФІЧНОЇ МІКРОСКОПІЇ

7.1 Дослідження існуючих методів, технічних засобів та їх метрологічного забезпечення для виконання вимірювань у нанодіапазоні

7.1.1 Програми та стратегії в області нанометрології

Міжнародною організацією зі стандартизації (International organization for standardization - ISO) створений Технічний комітет ISO / TC 229 «Нанотехнології». Своє перше засідання Комітет провів 9 листопада 2005 в Лондоні. Організатор засідання – Британська організація по стандартизації. Одним з найважливіших завдань ISO / TC 229 буде створення контрольованого словника для нанотехнологій, з набором однакових визначень та номенклатури, для полегшення комунікації, а також розміщення правових і комерційних операцій, стандартів і правил на міцну основу.

Конкретні завдання в першу чергу включають розробку трьох категорій стандартів:

- термінологія і номенклатура – стандарти забезпечують спільну мову для наукових, технічних, комерційних і регуляторних процесів;
- вимірювання і визначення характеристик – стандарти забезпечують міжнародно визнану основу для кількісної наукової, комерційної та нормативної діяльності;
- здоров'я, безпека та вплив на навколишнє середовище – стандарти безпеки праці, споживачів та охорони навколишнього середовища, стимулювання належної практики в галузі виробництва, використання та утилізації наноматеріалів, нанотехнологій і продукції нанотехнологій з підтримкою систем і продуктів.

Стандартизація в галузі нанотехнологій включає в себе один або обидва з наступних дій:

1. Розуміння і контролювання матеріалу і процесів на нанорівні (як правило, менше 100 нм) в одному або більше вимірюваннях, де виникнення розмірно-залежного явища зазвичай запускає в роботу нове застосування.

2. Використання властивостей нанорозмірних матеріалів, які відрізняються від властивостей окремих атомів, молекул і речовини в обсязі, для створення більш досконалих матеріалів, приладів і систем, які використовують ці нові властивості.

ISO / TC 229 встановив три Робочих Групи (WG – Working Groups) для розвитку певних областей стандартизації в галузі нанометрології:

WG 1: Терміни та визначення.

WG 2: Вимірювання та визначення характеристик.

WG 3: Здоров'я, безпека, вплив на навколишнє середовище нанотехнологій.

Вирішення цих завдань, на думку фахівців, дасть потужний імпульс розвитку нанотехнологій і їх практичного застосування в різних галузях економіки. В рамках Технічного комітету ISO / TC 229, секретаріат якого веде Британський інститут стандартів, діяльність підкомітету з метрології, методам вимірювань та випробувань координує Японія, підкомітету з термінам і визначенням - Канада, підкомітету по здоров'ю, безпеці та навколишньому середовищу - Сполучені Штати Америки.

В рамках фінансованого ЄС проекту, Co-nanomet, в січні 2011 року був підготовлений документ, який розглядає сучасні вимоги нанометрології і майбутні стратегії з розвитку європейської нанотехнології, сформульованих поруч Європейських Нанометрологічних Груп Дій. Були розглянуті питання єдності вимірювання, стандартизації, національних і міжнародних програм, регулювання та навички розвитку.

В області координатної метрології розглянуті ключові вимірювальні методи (рис. 7.1).

Одними з основних видів інструментів, що використовуються в нанометрології є скануючий зондовий мікроскоп (СЗМ), і, наприклад, атомний силовий мікроскоп (АСМ).

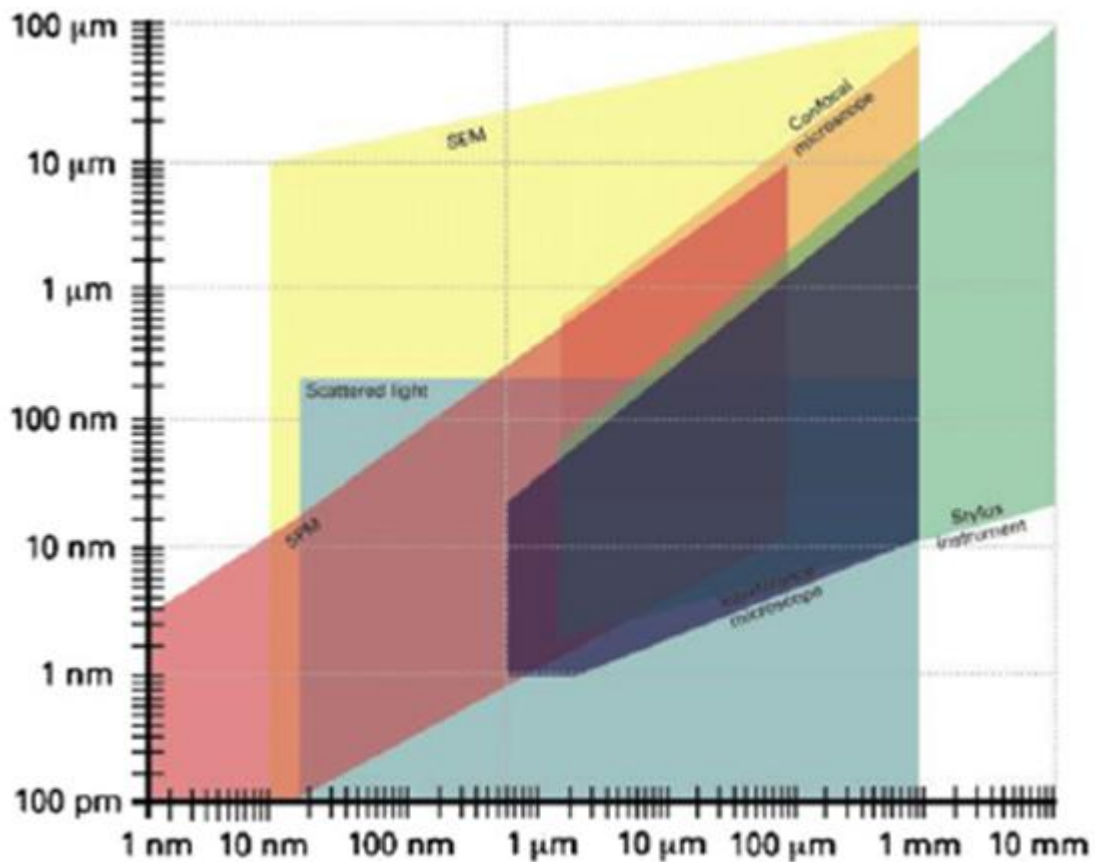


Рисунок 7.1 – Вимірювальні технології та їх типові динамічні діапазони у горизонтальному (довжина хвилі) та вертикальному (амплітуда) напрямках

Є багато комерційно доступних СЗМ і більшість НМІ використовують їх, або розробляють метрологічні СЗМ. Co-nanomet виділив ряд областей СЗМ досліджень, які необхідно вирішити.

- Горизонтальний діапазон сканування з СЗМ повинен бути збільшений.

В даний час діапазон сканування СЗМ по горизонталі в кращому випадку складає сотні мікрометрів, але повинен бути збільшений до декількох десятків міліметрів, щоб дозволити проводити вимірювання в більшій площі структурованих поверхонь, кремнієвих пластин і оптики.

Це вимагатиме збільшення швидкості сканування і використання інтелектуального зондування та системи управління, вибірки стратегій, які можуть включати в себе поєднання СЗМ з іншими приладами, такими, як оптична мікроскопія для огляду і «грубі» сканування. Інтелектуальні мультисенсорні концепції з поліпшеними алгоритмами злиття даних і автоматизованим визначенням можливих цікавлять областей допоможе заповнити великий динамічний діапазон і підвищити ефективність.

- Роздільна здатність СЗМ, як метрологічне і не метрологічне, потребує поліпшення, щоб виміряти ще більш дрібні структури з більш високою точністю. Це вимагає, серед іншого, кращого теоретичного уявлення про поверхню взаємодії зонда і підвищення якості характеристик зонда.

- По суті СЗМ є видом виміру, тобто вимірюються малі відхилення від площини, але все більше і більше потрібні справжні тривимірні (3D) вимірювання. Особливо це потрібно в напівпровідникової промисловості, де необхідно виміряти високе співвідношення сторін структур і критичних розмірів, для розвитку MEMS (microelectromechanical systems - мікроелектромеханічні системи) і NEMS (nanoelectromechanical systems - наноелектромеханічні системи). Повинні бути розроблені нові стратегії зондування з можливістю проведення 3D зондування та сканування (за аналогією з методами контрольно-вимірювальних машин (КІМ)), яке зможе досягти бічних стінок, канавок, отворів і т.д. Взаємодія між зондом СЗМ і поверхнею має важливе значення для метрології та має бути усвідомлене краще, тому можуть бути створені відповідні поправки або допомоги для справжніх квантового і координатного вимірювань.

- Всі вище досліджені області СЗМ потребують підтримки відповідної інфраструктури єдності вимірювання. Це включає в себе подальший розвиток метрологічних СЗМ, передачу еталонів, оптичні інтерферометри, щоб забезпечити маршрут єдності вимірювання.

Існує більш широке використання впливу структурованих поверхонь на

функціональні частини. Це вимагає швидкої і точної галузі вимірювань і характеристики методів. СЗМ зазвичай використовується для вимірювання структури з горизонтальними масштабами менше ніж 500 нм, а для більш великих масштабів, стилус і оптичні прилади використовуються з суб-нанометровою вертикальною роздільною здатністю. Перші технічні стандарти у цій галузі були опубліковані і в даний час розробляється єдність вимірювання інфраструктури.

Рефлектометрія (або дифрактометрія) є методом, який широко використовується в напівпровідниковій промисловості для вимірювання ціни поділки і структурованих поверхонь. Питання, які досі необхідно розглянути, включають: ефект неповноти, як виміряти все менші області і час обчислення. Спектроскопічні еліпсометрія (SE) є окремим випадком рефлектометрії, що має важливе значення для вимірювання нанометрових тонких плівок, в тому числі, що складаються з декількох шарів. Необхідні подальші дослідження для стандартизації методів рефлектометрії і порівняння їх з іншими оптичними та контактними методами.

Скануюча електронна мікроскопія (SEM) є дуже різнобічним і усталеним інструментом, який має великий динамічний діапазон і можливість великого збільшення. Останнім часом нові 3D еталони передачі, застосовні до стереоскопічного SEM (а також СЗМ і багатьом оптичним методам), були розроблені з відповідним аналітичним програмним забезпеченням.

Трансмісійний електронний мікроскоп (ПЕМ) також добре усталений метод, але дослідження і раніше вимагають збільшення своєї універсальності. Хоча методи електронної мікроскопії високорозвинені і внесли великий вклад у різних областях досліджень матеріалів, взаємодія міжнародно визнаних лабораторій з додатковими изученнями експертів обмежений. Інші методи пучка частинок, такі як сфокусований пучок іонів (FIB) і гелію іонної мікроскопії починають широко використовуватися, але маршрут єдності вимірювання знаходиться тільки в їх ранній стадії розвитку і такі методи, як правило,

використовуються тільки в режимі візуалізації. Поєднання нанорозмірною обробки та метрології на своєму місці (часто з використанням вимірювального зонда), ймовірно, буде розвинутою областю досліджень.

Метод оптичної інтерферометрії часто використовується для отримання прямої калібрування зміщення в інших інструментах, наприклад метрологічних СЗМ. Фінансований EMRP проект NANOTRACE включає в себе кілька НМІ і спрямований на розробку низки інтерферометров суб-нанометрового дозволу, які мають нелінійності, визначені за допомогою рентгенівського інтерферометричному пристрою. NANOTRACE буде брати участь у виконанні вимог до наступного покоління оптичних інтерферометрів, але завдання полягатиме у поширенні цих вимог для користувачів за межами НМІ і в створенні зміщення вимірювань з пікометричним дозволом, а також підвищенні точності.

У Росії функції державної метрологічної служби покладено на Федеральне агентство з технічного регулювання і метрології – Ростехрегулювання і підвідомчі йому організації. Для вирішення цих завдань у галузі високих технологій, включаючи нанотехнологію, в Ростехрегулюванні створений Технічний комітет з стандартизації ТК 441 "Наукоємні технології", функції організаторської діяльності секретаріату якого покладено на Науково-дослідний центр з вивчення властивостей поверхні і вакууму Ростехрегулювання (НІЦПВ).

У ТК 441 входять підкомітети:

- ПК 1 «Нанотехнології»;
- ПК 2 «Квантоворозмірні ефекти в наукоємних технологіях»;
- ПК 3 «Терміни та визначення»;
- ПК 4 «Методи і засоби забезпечення єдності вимірювань в нанотехнологіях»;
- ПК 5 «Нанотехнології в мікроелектроніці»;
- ПК 6 «Матеріали, структури та об'єкти нанотехнологій»;
- ПК 7 «Нанотехнології та наноіндустрія».

В рамках Міжнародної організації СОOMET з європейсько-азіатському співпраці був створений проект «Метрологічне забезпечення нанотехнологій» в галузі метрології, присвячений вирішенню фундаментальних проблем метрології в нанотехнологіях. Країни-учасники: Росія, Білорусь, Україна, Словаччина, Німеччина. На даний момент розроблені проект концепції метрологічного забезпечення нанотехнологій і технологія передачі розміру одиниць фізичних величин в нанодіапазоні.

7.1.2 Стандарти в області нанометрології

З нанометрологією пов'язані теоретичні та практичні аспекти забезпечення єдності вимірювань в нанодіапазоні.

В першу чергу – це еталони фізичних величин і еталонні установки, стандартні зразки складу, структури і властивостей для забезпечення передачі розміру в нанодіапазоні.

По-друге – атестовані або стандартизовані методи вимірювань фізико-хімічних параметрів і властивостей об'єктів нанотехнології, методи калібрування (повірки) застосовуваних засобів вимірювань.

По-третє – метрологічне супроводження процесів виробництва продукції нанотехнології.

Стандарти, які опубліковані ISO:

ISO / TR 27628: 2007 - Workplace atmospheres - Ultrafine, nanoparticle and nano-structured aerosols - Inhalation exposure characterization and assessment.

ISO / TR 27628: 2007 - Workplace atmospheres - Ultrafine, nanoparticle and nano-structured aerosols - Inhalation exposure characterization and assessment

ISO / TS 27687: 2008 - Nanotechnologies - Terminology and definitions for nano-objects - Nanoparticle, nanofibre and nanoplate

ISO / TS 80004-1: 2010 - Nanotechnologies - Vocabulary - Part 1: Core terms

ISO / TS 80004-3: 2010 - Nanotechnologies - Vocabulary - Part 3: Carbon nano-objects

ISO / TR 11360: 2010 - Nanotechnologies - Methodology for the classification and categorization of nanomaterials.

Ще понад 20 стандартів знаходяться в розробці.

У Росії для забезпечення нормативної бази нанометрології, розроблені і введені в дію сім стандартів:

ГОСТ Р 8.628-2007. Заходи рельєфні нанометрового діапазону. Вимоги до геометричних форм, лінійним розмірам і вибору матеріалу для виготовлення.

ГОСТ Р 8.629-2007. Заходи рельєфні нанометрового діапазону з трапецеїдальним профілем елементів. Методика повірки.

ГОСТ Р 8.630-2007. Мікроскопи скануючі зондові атомно-силові вимірювальні. Методика повірки.

ГОСТ Р 8.631-2007. Мікроскопи електронні растрові вимірювальні. Методика повірки.

ГОСТ Р 8.635-2007. Мікроскопи скануючі зондові атомно-силові. Методика калібрування.

ГОСТ Р 8.636-2007. Мікроскопи електронні растрові. Методика калібрування. ГОСТ Р 8.644-2008. Заходи рельєфні нанометрового діапазону з трапецеїдальним профілем елементів. Методика калібрування.

Європейська політика сусідства відкриває для України нові можливості для партнерства, економічної інтеграції й співробітництва, а в перспективі дозволить Україні одержати право повноправного членства в ЄС.

19 травня 2010 р. Кабінетом міністрів України була прийнята постанова № 1070-р «Про затвердження плану першочергових заходів щодо реформування системи технічного регулювання».

Дане розпорядження передбачало затвердити план першочергових заходів щодо реформування системи технічного регулювання, зокрема підготувати проекти законів України «Про державний ринковий нагляд», «Про загальну

безпеку продукції», «Про відповідальність постачальника за дефектну продукцію», «Про внесення змін у деякі закони України» і ін., а також забезпечити розробку не менш чим 1500 стандартів, необхідних для впровадження технічних регламентів.

Необхідно зазначити відсутність вітчизняної нормативної документації, що розглядає питання оцінювання невизначеності нановимірювань. Прийняття таких стандартів заплановане Держспоживстандартом, однак, такі стандарти планують приймати методом «обкладинки», що може викликати труднощі з інтерпретацією й ряд питань у користувачів такими стандартами.

7.2 Вдосконалення метрологічного забезпечення методів та засобів нановимірювань

7.2.1 Аналіз основних особливостей оцінювання невизначеності нановимірювань

Вирішення проблеми забезпечення єдності вимірювань потребує відомостей щодо характеристик похибок або невизначеності вимірювань із заданою ймовірністю. Підписання Україною Угоди про взаємне визнання національних еталонів одиниць і свідоцтв калібрувань і вимірювань, які видаються національними метрологічними інститутами (MRA), введення в дію стандартів ДСТУ ISO/IEC 17025:2006 та ДСТУ РМГ 43:2006 поклало початок законодавчому використанню концепції невизначеності (КН) в Україні. При втіленні КН, як здобутку процесу міжнародної стандартизації оцінювання точності вимірювань, необхідно забезпечити, щоб оцінки невизначеності, що одержують в різних лабораторіях були отримані не тільки за однаковою процедурою, але й були б достовірними. Але, незважаючи на існування великої кількості підходів до оцінювання невизначеності вимірювань, оцінки останніх, що отримують при використанні стандартизованих методик, в деяких випадках

не є достатньо достовірними. Це обумовлено спрощеннями при знаходженні розширеної невизначеності вимірювань, які нехтують законами розподілу вхідних величин модельного рівняння (рівняння вимірювань), наявністю (чи відсутністю) кореляції між результатами багатократних спостережень, нелінійністю модельної функції тощо. Такий стан не може вважатися задовільним, особливо коли отримані оцінки застосовуються під час калібрування засобів вимірювальної техніки (ЗВТ) для забезпечення простежуваності результатів вимірювань до еталонів SI.

Невідповідність вимог реальних практичних завдань метрологічного забезпечення сучасному стану розвитку КН робить нагальним вдосконалення теоретичних і прикладних основ оцінювання характеристик невизначеності вимірювань. В такій постановці тема дисертаційного дослідження є актуальною і відповідає потребам України, особливо в світі прийняття нової редакції Закону про метрологію та метрологічну діяльність та тривання процесу акредитації за ДСТУ ISO/IEC 17025 низки калібрувальних лабораторій. Актуальність цього завдання підтверджується здійсненням ВІРМ ревізії базового документа КН – Настанови з подання невизначеності вимірювань (GUM) з метою забезпечення правильного балансу між точністю та простотою використання.

Питання оцінювання невизначеності вимірювань розглядаються у книгах [111-114].

Залежно від цілей оцінювання невизначеності повинні прийматися в увагу різні категорії її джерел. Основні категорії джерел невизначеності, що мають істотне значення, зведені в табл. 7.1.

У процесі оцінювання невизначеності повинен бути зібраний повний список відповідних джерел невизначеності. З метою виявлення джерел невизначеності й величин, що впливають, варто розглянути повну специфікацію й кожне положення стандарту як можливе джерело невизначеності або величини, що впливає. Крім того, кожний крок у процедурі виміру представляє, у принципі, можливе джерело невизначеності. Для того щоб спростити завдання

виявлення всіх можливих джерел невизначеності й уникнути подвійного обліку джерел пропонується використовувати причинно-наслідкову діаграму (діаграма «риб'ячий скелет») при складанні списку джерел невизначеності, із вказівкою їхнього взаємозв'язку й впливу на невизначеність результату вимірювань.

Таблиця 7.1. Основні категорії джерел невизначеності

Обумовлені випробувальною лабораторією	Обумовлені стандартом випробування	Обумовлені виробництвом і відбором випробуваних зразків
<ul style="list-style-type: none"> • навички оператора; • аналіз і розрахунки; • протоколювання; • ступінь впровадження стандарту процедури вимірювання програмне забезпечення; • система якості. 	<ul style="list-style-type: none"> • установлені технічні вимоги вимірювань; • характеристики вимірювального обладнання, включаючи калібрування й верифікацію; • опис процедури вимірювань; • умови навколишнього середовища; • розміщення випробуваного обладнання; • функціонування випробуваного обладнання; • тип випробуваного обладнання. 	<ul style="list-style-type: none"> • виробничий допуск; • відбір зразків; • нерепрезентативна вибірка.

Важливим етапом також є перетворення кожного джерела невизначеності в одну або більше величин, що впливають. Тут можна виділити деяку особливість: величини, що впливають, які важко визначити і які не можуть бути визначені зовсім (невстановлені величини, що впливають) повинні бути також включені в бюджет невизначеності, незважаючи на ці труднощі. Це можна зробити, якщо

прийняти діапазон значень для розглянутої величини, що впливає, або розглянути діапазон імовірностей для джерела невизначеності. На рис. 7.2 показаний зв'язок між джерелами невизначеності, відповідними величинами, що впливають, і результуючими невизначеностями.

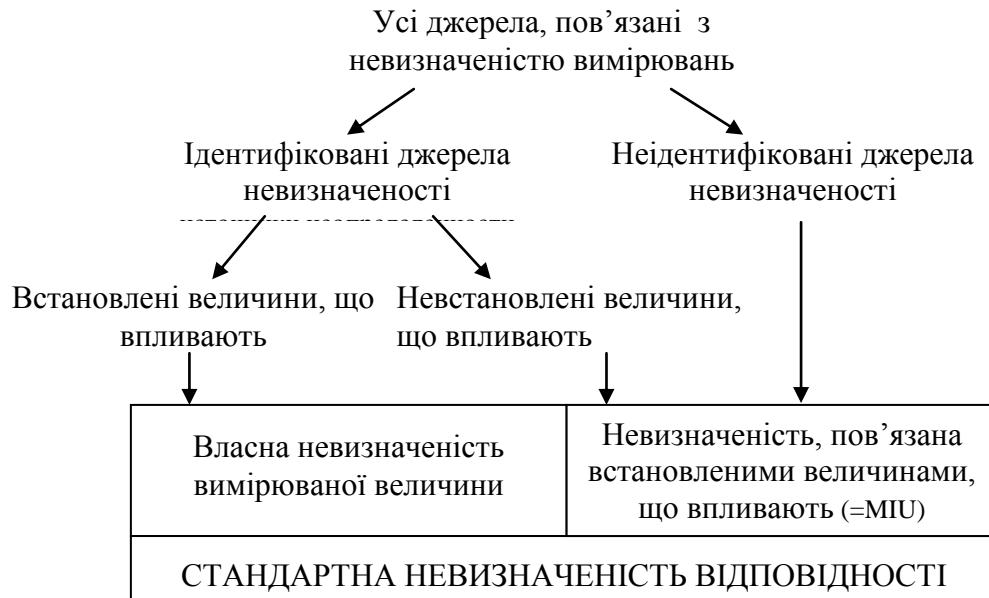


Рисунок 7.2 – Зв'язок між джерелами невизначеності, величинами, що впливають, й категоріями невизначеності

Після ідентифікації встановлених і невстановлених величин, що впливають, і пов'язаних з ними допусків, повинна бути визначена невизначеність результату вимірювань. Це може бути зроблено шляхом моделювання стандартизованого методу вимірювань або за допомогою проведення експерименту. Слід зазначити, в якості оцінки загальної невизначеності крім підсумовування дисперсій допускається приводити сумарну невизначеність, розраховану у вигляді повного диференціала.

При оцінюванні невизначеності необхідна модель, що описує зв'язок між вимірюваною величиною й всіма встановленими й невстановленими величинами, що впливають. Така модель може бути аналітичною або чисельною.

7.2.2 Особливості оцінювання невизначеності вимірювань в стандарті ISO 14253-2 [115]

Після видання Настанови з подання невизначеності вимірювань (GUM) з'явилася необхідність у створенні аналогічних нормативних документів для різних видів вимірювань. Так, для калібрувальних лабораторій було створено EA 4/02, в області аналітичних вимірювань – QUAM-P1:2000 і т.д. Для GPS (geometrical product specification) вимірювань була розроблена технічна специфікація ISO/TS 14253-2: 1999 «Геометричні характеристики виробів - Контроль при вимірюванні оброблюваних деталей і засоби вимірювання. Частина 2: Настанова з оцінки невизначеності в GPS вимірюваннях, при калібруванні засобів вимірювання та верифікації продукції» (далі - стандарт).

Слід зазначити, що цей стандарт має ряд особливостей в оцінюванні невизначеності вимірювань в порівнянні з GUM, включаючи процедуру менеджменту невизначеності (PUMA), які можна ефективно застосовувати не тільки для геометричних, але й для будь-яких інших видів вимірювань. На жаль, стандарт досі не перекладений українською мовою, незважаючи на те, що ДСТУ ISO / TS 14253-2: 2006 є чинним в Україні з 01.10.2007.

Алгоритм оцінювання невизначеності в стандарті включає наступні операції.

1. Складання модельного рівняння

У стандарті розглядаються два крайніх виду модельного рівняння:

- прозорий ящик

$$Y = f(X_1, X_2, \dots, X_{p+r}), \quad (7.1)$$

де Y – вимірювана величина; f – модельна функція; X_i – i -а вхідна величина, $i = 1, 2, \dots, p + r$ (ця модель відповідає модельному рівнянню GUM);

- чорний ящик:

$$Y = X + \sum_{i=1}^{p+r} C_i, \quad (7.2)$$

де X – показання вимірювального приладу; C_i – i -а адитивна поправка на відому (невиключену) систематичну похибку (похибка калібрування, температурна похибка, похибка через деформацію і т.п.).

У виразах (7.1), (7.2) символи i позначають відповідно кількість некоррельованих і корельованих джерел невизначеності.

2. Оцінювання невизначеностей вхідних величин

Здійснюється в стандарті двома способами.

1. За типом А, на основі відхилень результатів багаторазових спостережень від їх середнього арифметичного

$$u_A = u(\bar{x}) = h \sqrt{\frac{1}{n(n-1)} \sum_{i=1}^n (x_i - \bar{x})^2}, \quad (7.3)$$

де h – коефіцієнт надійності, що вводиться при малому числі спостережень (n) (в табл. 7.2).

Таблиця 7.2. Значення коефіцієнтів надійності

n	2	3	4	5	6	7	8	9	10
h	7,0	2,3	1,7	1,4	1,3	1,3	1,2	1,2	1,0

2. За типом В: визначається через відомі межі a невиключеної систематичної похибки (НСП) і коефіцієнт розподілу b , що характеризує закон розподілу похибки всередині цих меж

$$u_B = a \cdot b. \quad (7.4)$$

Цей коефіцієнт наведений у стандарті для трьох найпоширеніших видів законів розподілу:

- розподіл Гаусса, $b = 0,5$;
- рівномірний розподіл, $b = 0,6$;
- U-розподіл (арксінусний), $b = 0,7$.

Якщо закон розподілу НВВ невідомий, то слід вибрати U – розподіл (з найбільшим b).

3. Облік кореляції

У стандарті розглядається два випадки:

- складові некорреліровані, коефіцієнт кореляції $\rho = 0$;
- складові суворо корреліровані, тоді приймається $\rho = 1, -1$.

4. Обчислення сумарної стандартної невизначеності

Підсумовування некорреліованих ($\rho = 0$) складових невизначеності здійснюється геометрично (корінь квадратний із суми квадратів), а сильно корельованих складових – арифметично:

- для моделі у вигляді прозорого ящика (7.1)

$$u_c = \sqrt{\sum_{i=1}^p \left(\frac{\partial Y}{\partial X_i} u(X_i) \right)^2 + \left(\sum_{i=1}^r \frac{\partial Y}{\partial X_i} u(X_i) \right)^2}, \quad (7.5)$$

де $\frac{\partial Y}{\partial X_i}$ – коефіцієнт чутливості; $u(X_i)$ – стандартна невизначеність i -ї

вхідної величини;

- для моделі у вигляді чорного ящика (7.2)

$$u_c = \sqrt{\sum_{i=1}^p u(C_i)^2 + \left(\sum_{i=1}^r u(C_i) \right)^2}. \quad (7.6)$$

У моделі чорного ящика передбачається, що складові невизначеності адитивні, тому, коефіцієнти чутливості індивідуальних складових невизначеності дорівнює 1.

5. Складання бюджету невизначеності

У стандарті (на відміну від GUM) передбачено складання бюджету невизначеності, що включає в себе інформацію для кожної вхідної величини про тип оцінки, тип розподілу, кількість багаторазових спостережень, межі розсіювання НСП, коефіцієнт кореляції, коефіцієнт розподілу, вклад в сумарну невизначеність, а також інформацію про сумарну стандартну та розширену невизначеності.

6. Обчислення розширеної невизначеності

Здійснюється в стандарті для ймовірності 0,95 за формулою

$$U = k \cdot u_c, \quad k=2. \quad (7.7)$$

7.3 Підвищення достовірності статистичних оцінок невизначеності вимірювань

Реалізаціями модельного підходу є закон розповсюдження невизначеності (базовий алгоритм GUM) та закон розповсюдження розподілів (Додаток 1 до GUM). У відповідності до базового алгоритму оцінювання невизначеності вимірювань, розширена невизначеність результату вимірювань визначається як добуток сумарної стандартної невизначеності на коефіцієнт покриття. В загальному випадку цей коефіцієнт розраховується як коефіцієнт Стюдента для ймовірності 0,95 та ефективного числа степенів свободи ν_{eff} , яке визначається за формулою Велча-Саттерсвейта та представляє собою дробове число. GUM рекомендує «зменшувати отримане значення ν_{eff} до найближчого цілого числа або інтерполювати». У випадку, коли отримане значення ν_{eff} мале, похибка визначення коефіцієнта Стюдента при його округленні може бути значною. Тому треба розрахувати значення коефіцієнта Стюдента для дробового числа степенів свободи та отримати інтерполюючу формулу для подальшого практичного використання.

Були отримані значення коефіцієнта Стьюдента $t_p(v)$ для дробових чисел степенів свободи v (табл. 7.3) та розроблено інтерполюючу формулу для вирішення аналітичних задач оцінювання невизначеності вимірювань [116], а також для автоматизації оцінювання невизначеності вимірювань при вирішенні практичних задач у випробувальних та калібрувальних лабораторіях з прийнятною точністю. В основу розрахунку покладено метод послідовного наближення при виконанні операції чисельного інтегрування. Оскільки для малих значень v та Δt , час обчислення $t_p(v)$ стає неприпустимо великим, було застосовано ітераційний алгоритм.

Таблиця 7.3 – Значення коефіцієнта Стьюдента

v	p					
	0,6827	0,95	0,9545	0,98	0,99	0,9973
1,0	1,837409	12,70620	13,96781	31,82052	63,65674	235,8698
1,1	1,724314	10,27690	11,20320	23,70698	44,54138	146,518200
1,2	1,634356	8,648912	9,363620	18,64018	32,241698	99,034480
1,3	1,568548	7,500548	8,073710	15,26986	26,059392	71,402950
1,4	1,512838	6,657068	7,130930	12,91299	21,22623	54,141040
1,5	1,466860	6,016736	6,418240	11,19748	17,821458	42,728140
1,6	1,428298	5,517236	5,864485	9,907410	15,331599	34,827400
1,7	1,395508	5,118824	5,424160	8,910206	13,453983	29,146060
1,8	1,367308	4,794932	5,067150	8,121790	12,001167	24,930010
1,9	1,342804	4,527092	4,772745	7,486028	10,851552	21,715930
2,0	1,321315	4,302653	4,526551	6,964557	9,924843	19,209532
2,1	1,302332	4,112182	4,318069	6,530974	9,166100	17,215377
2,2	1,285435	3,948744	4,139501	6,165466	8,535110	15,601632
2,3	1,270307	3,807179	3,985100	5,854012	8,004078	14,276083
2,4	1,256682	3,683531	3,850407	5,585964	7,552258	13,172894
2,5	1,244352	3,574672	3,732029	5,353204	7,163988	12,243560
2,6	1,233140	3,478210	3,627236	5,149506	6,827460	11,452765
2,7	1,222901	3,392167	3,533896	4,969920	6,533367	10,773477
2,8	1,213216	3,315002	3,450280	4,810596	6,274575	10,184986
2,9	1,204883	3,245450	3,374984	4,668388	6,045344	9,671378
3,0	1,196913	3,182446	3,306830	4,540703	5,840909	9,219850
3,1	1,189538	3,125168	3,244940	4,425720	5,658148	8,820430
3,2	1,182691	3,072848	3,188450	4,321490	5,493436	8,465110
3,3	1,176317	3,024896	3,136710	4,226680	5,344480	8,147275

v	p					
	0,6827	0,95	0,9545	0,98	0,99	0,9973
3,4	1,170370	2,980796	3,089165	4,140110	5,209192	7,861605
3,5	1,164808	2,940104	3,045335	4,060770	5,085856	7,603690
3,6	1,159596	2,902460	3,004805	3,987830	4,973008	7,369855
3,7	1,154701	2,867540	2,967225	3,920570	4,869400	7,157055
3,8	1,150096	2,835056	2,932290	3,858350	4,773976	6,962665
3,9	1,145755	2,804756	2,899740	3,800650	4,685848	6,784515
4,0	1,141655	2,776445	2,869315	3,746947	4,604095	6,620720
4,1	1,137782	2,749960	2,840879	3,697005	4,528432	6,469696
4,2	1,134113	2,725088	2,814188	3,650310	4,457890	6,330064
4,3	1,130632	2,701712	2,789111	3,60660	4,392082	6,200640
4,4	1,127327	2,679696	2,765504	3,565610	4,330564	6,080352
4,5	1,124185	2,658928	2,743247	3,527100	4,272934	5,968336
4,6	1,121193	2,639304	2,722223	3,490850	4,218850	5,863792
4,7	1,118341	2,620736	2,702339	3,456675	4,168000	5,766000
4,8	1,11562	2,603144	2,683505	3,424400	4,120102	5,674368
4,9	1,113021	2,586456	2,665640	3,393880	4,074928	5,588336
5,0	1,110533	2,570582	2,648654	3,364930	4,032143	5,507445
5,1	1,108156	2,555510	2,632535	3,337558	3,991855	5,431245
5,2	1,105878	2,541140	2,617171	3,311530	3,953590	5,359350
5,3	1,103692	2,527442	2,602527	3,286780	3,917287	5,291415
5,4	1,101594	2,514362	2,588554	3,263218	3,882802	5,227140
5,5	1,09958	2,501870	2,575206	3,240766	3,850003	5,166240
5,6	1,097644	2,489924	2,562444	3,219346	3,818770	5,108460
5,7	1,095782	2,478482	2,550229	3,198892	3,789001	5,053590
5,8	1,093988	2,467526	2,538528	3,179338	3,760591	5,001390
5,9	1,092262	2,457014	2,527310	3,160624	3,733456	4,951710
6,0	1,090595	2,446912	2,516528	3,142668	3,707428	4,904355
6,1	1,08899	2,437235	2,506205	3,125528	3,682680	4,859110
6,2	1,08744	2,427915	2,496268	3,109049	3,658896	4,816000
6,3	1,085942	2,418955	2,486710	3,093227	3,636093	4,774690
6,4	1,084496	2,410325	2,477511	3,078023	3,614213	4,735270
6,5	1,083096	2,402015	2,468649	3,063401	3,593201	4,697470
6,6	1,081742	2,394000	2,460108	3,049331	3,573011	4,661200
6,7	1,080432	2,386270	2,451870	3,035780	3,553589	4,626460
6,8	1,079162	2,378805	2,443920	3,022724	3,534903	4,593070
6,9	1,077932	2,371600	2,436242	3,010133	3,516901	4,561030
7,0	1,076739	2,364624	2,428809	2,997952	3,499483	4,530160

По розрахованим значенням коефіцієнту Стьюдента для ймовірності 0,95 було отримано аналітичний вираз, який апроксимує значення коефіцієнта Стьюдента в діапазоні від 2,37 до 12,7

$$t_{0,95}(v) = \frac{2,348551}{1 - 1,45153 \cdot \exp(-0,576295 \cdot v)}. \quad (7.8)$$

Відносна похибка апроксимації отриманих значень $t_p(v)$ залежністю (7.8) не перевищує $\pm 1\%$ в діапазоні v від 1 до 6. Слід сказати, що відомі апроксимуючі та інтерполюючі формули використовуються в діапазоні обмеженому значеннями $v \geq 2$ та мають відносну похибку до 5%.

Було проведено аналіз методів оцінювання сумарної стандартної невизначеності за типом A непрямих корельованих та некорельованих вимірювань. Загальним недоліком закону розповсюдження невизначеності є зміщення результату вимірювань, обумовлене нехтуванням нелінійності модельного рівняння. Цей недолік для корельованих вимірювань можливо усунути реалізацією методу приведення (редукції). Для оцінювання сумарної стандартної невизначеності за типом A некорельованих вхідних величин запропоновано застосування методу транспозицій даних, який також дозволяє усунути зазначений недолік. Нажаль, відома реалізація цього методу працює тільки при однаковій кількості спостережень вхідних величин. Суть методу транспозицій даних полягає в отриманні всіх можливих значень вимірювальної величини y_i шляхом перебору всіх значень вхідних величин, які підставляються до модельного рівняння

$$y_i = f(x_{1i}, x_{2i}, \dots, x_{mi}). \quad (7.9)$$

Очевидно, що при m вхідних величинах процес отримання масиву можливих значень вимірюваної величини y_i буде виглядати як m -мірна таблиця з загальною кількістю комірок

$$N = \prod_{i=1}^m n_i. \quad (7.10)$$

За одержаним масивом вимірюваної величини можна отримати незміщений результат вимірювання

$$\bar{y} = \frac{1}{N} \sum_{i=1}^N y_i \quad (7.11)$$

та дисперсію y_i

$$s^2(y_i) = \frac{1}{(N-1)} \sum_{i=1}^N (y_i - \bar{y})^2. \quad (7.12)$$

Для переходу від дисперсії y_i до стандартної невизначеності за типом A , необхідно визначити еквівалентне число спостережень n_{eq} , яке характеризує невизначеність вимірюваної величини. Ця задача не була вирішена раніш, а використання для цієї цілі загальної кількості варіантів перебору N не є правомочним, оскільки воно не характеризує число реально проведених результатів вимірювань, та його використання призведе до заниження оцінки невизначеності. В результаті вирішення цієї задачі була отримана формула

$$n_{\text{eq}} = \frac{\sum_{j=1}^m n_j c_j^2 u_A^2(x_j)}{\sum_{j=1}^m c_j^2 u_A^2(x_j)}, \quad (7.13)$$

яка для двох вхідних величин буде мати вигляд

$$n_{\text{eq}} = \frac{n_1 c_1^2 u_A^2(x_1) + n_2 c_2^2 u_A^2(x_2)}{c_1^2 u_A^2(x_1) + c_2^2 u_A^2(x_2)} = \frac{n_1 w_{1,2}^2 + n_2}{w_{1,2}^2 + 1}, \quad (7.14)$$

де $w_{1,2} = c_1 u_A(x_1) / c_2 u_A(x_2)$ – відношення внесків невизначеності вимірювань першої та другої вхідних величин.

Залежність n_{eq} від $w_{1,2}$, n_1 та n_2 наведена на рис. 7.2, з якого видно, що при будь-яких $w_{1,2}$, значення n_{eq} знаходиться між n_1 та n_2 , та стає рівним $n_{eq} = n$ при $n_1 = n_2 = n$.

Таким чином, сумарна стандартна невизначеність за типом A дорівнює

$$u_A(y) = s(y_i) / \sqrt{n_{eq}}, \quad (7.15)$$

Оскільки формула (7.6) дає в загальному випадку дробові числа степенів свободи, для пошуку коефіцієнту покриття застосовувався вираз (7.15).

Для рівного числа спостережень всіх вхідних величин вираз (7.15) переходить в відомий вираз

$$u_A(y) = s(y_i) / \sqrt{n}. \quad (7.16)$$

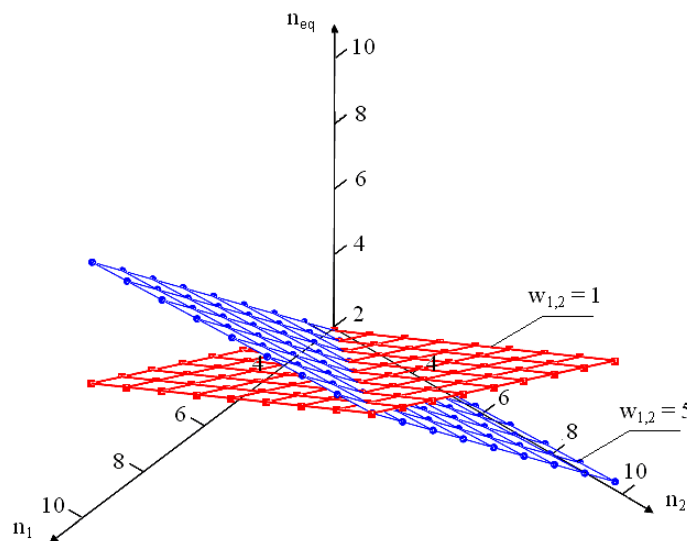


Рисунок 7.3 – Залежність еквівалентного числа вимірювань від відношення внесків невизначеності вимірювань вхідних величин та числа їх спостережень

При опрацюванні даних методом транспозицій виникає питання, чи не буде впливати на оцінку невизначеності вимірювань хибна (уявна) кореляція,

що має місце при обмеженому числі спостережень між двома наборами даних.

З нерівності Стьюдента, що визначає істотність значення коефіцієнту кореляції, була отримана залежність максимального значення коефіцієнта уявної кореляції (рис. 7.4), яка може спостерігатися між двома наборами некорельованих даних з рівним числом спостережень n . З цієї залежності видно, що для $p=0,95$ навіть при $n=15$, абсолютне значення r_{\max} може сягати 0,5.

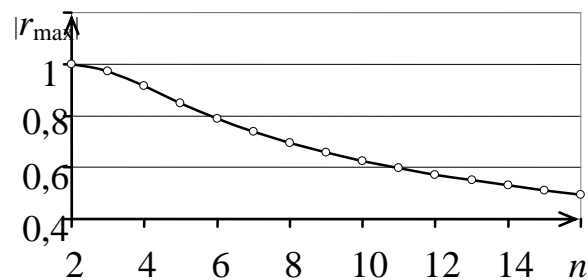


Рисунок 7.4 – Залежність $|r_{\max}|$ від числа спостережень n

Оскільки при перестановці даних в будь-якому з наборів уявна кореляція буде залишатися (буде змінюватися лише коефіцієнт кореляції), це може суттєво впливати на оцінку невизначеності вимірювань при реалізації методу транспозицій.

Для доведення, що при використанні методу транспозицій уявна кореляція вихідних даних не впливатиме на обробку результатів непрямих некорельованих вимірювань, було взято дві групи незалежно отриманих результатів вимірювання двох вхідних величин: $x_{11}, x_{12}, \dots, x_{1n}$; $x_{21}, x_{22}, \dots, x_{2n}$. Показано, що кореляційний момент для всіх даних, опрацьованих методом транспозицій, буде остаточно визначатися за формулою

$$K = \frac{1}{n^2 - 1} [(x_{21} - \bar{x}_2) + (x_{22} - \bar{x}_2) + \dots + (x_{2n} - \bar{x}_2)] [(x_{11} - \bar{x}_1) + (x_{12} - \bar{x}_1) + \dots + (x_{1n} - \bar{x}_1)], \quad (7.17)$$

з якої видно, що перший та другий множники його правої частини є сумою відхилень результатів спостережень кожної вхідної величини від їх середніх значень, і кожний з них дорівнює нулю, таким чином кореляційний момент (і коефіцієнт кореляції) для даних, опрацьованих методом транспозицій, буде також дорівнювати нулю.

Була також досліджена оцінка достовірності сумарної стандартної невизначеності типу A , яка отримана методом транспозицій. Показано, що дисперсія квадрата оцінки стандартної невизначеності типу A , що отримана традиційним методом буде в декілька разів більше дисперсії, яка отримана методом транспозицій. З ростом числа спостережень та кількості числа вхідних величин це відношення різко зростає.

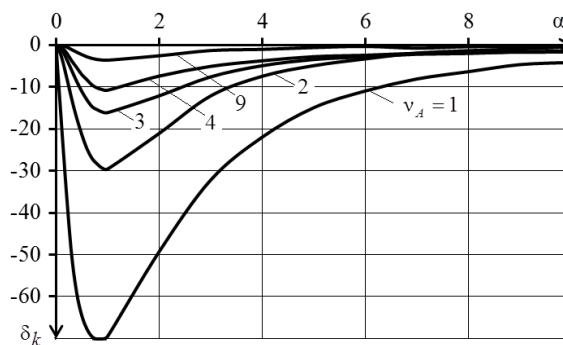
7.4 Оцінювання розширеної невизначеності вимірювань

Було досліджено достовірність оцінки розширеної невизначеності на основі ефективного числа степенів свободи, загальною моделлю розрахунку якого було наступне представлення формули Велча-Саттерсвейта

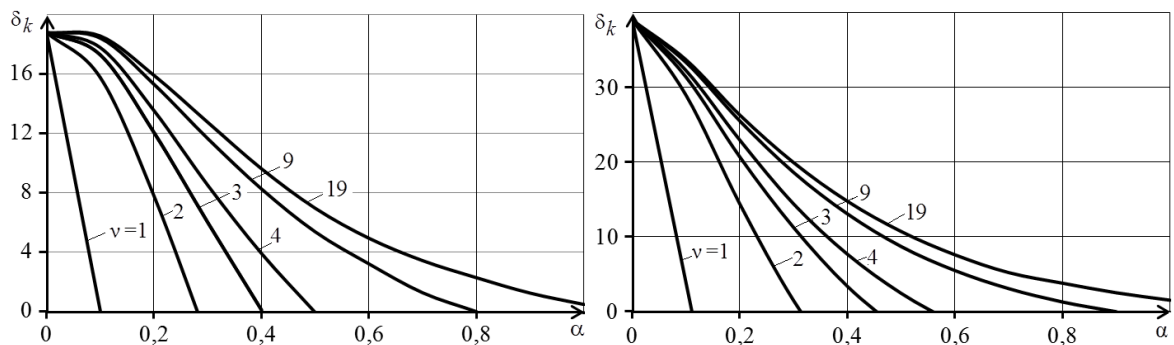
$$v_{\text{eff}} = v_A \left(1 + \frac{1}{\alpha^2} \right)^{-2}, \quad (7.18)$$

де $\alpha = u_A(y)/u_B(y)$; $u_A(y)$ та $u_B(y)$ – відповідно сумарні стандартні невизначеності типу A та B ; v_A – ефективне число степенів свободи, що відповідає сумарній стандартній невизначеності типу A . Оскільки формула (7.18) дає в загальному випадку дробові числа степенів свободи, для пошуку коефіцієнту покриття застосовувався вираз (2.8). Для дослідження похибки оцінювання коефіцієнта покриття за методикою GUM було проведено чисельний експеримент на основі ММК. В ньому досліджувалася композиція закону розподілу Стюдента з заданим числом степенів свободи v_A та

нормального, рівномірного, арксинусного законів розподілу, що відповідають внескам невизначеності типу B . Коефіцієнт покриття за методикою GUM знаходився як коефіцієнт Стюдента від ефективного числа степенів свободи, розрахованого з використанням виразу (7.8). На рис. 7.5 показано залежності відносної похибки δ_k оцінювання коефіцієнту покриття за методикою GUM від параметру α , що змінюється в діапазоні $0 \dots 10$ для чисел степенів свободи $\nu_A = 1 \dots 19$.



а – для нормального закону;



б – для рівномірного закону;

в – для закону арксинуса.

Рисунок 7.5 – Залежності відносної похибки δ_k , % оцінювання коефіцієнту покриття за методикою GUM для різних законів розподілу внесків типу B

З рисунків видно, що максимальна негативна похибка δ_k для нормального закону розподілу (рис. 3,а) сягає -70% для $\nu_A = 1$; -30% для $\nu_A = 2$; -16% для $\nu_A = 3$; -11% для $\nu_A = 4$. Дослідження показали, що й для інших законів розподілу в цій області максимальна негативна похибка має практично ті ж самі значення. При малих значеннях α для законів розподілу,

відмінних від нормального, буде мати місто значна позитивна похибка: до 19 % в випадку рівномірного закону розподілу (рис. 7.5,б) та до 39 % в випадку розподілу за законом арксинуса (рис. 7.5,в).

Такі суттєві значення похибок оцінювання коефіцієнту покриття та розширеної невизначеності ставлять задачу отримання достовірних оцінок коефіцієнта покриття.

В основу вирішення цієї задачі було положено метод ексцесів, розроблений в книзі П. Тойберта, в якій методом функціонального аналізу доведено, що інтеграл згортки, який використовується для отримання закону розподілу суми декількох складових похибки вимірювань можна трансформувати в правило підсумовування ексцесів розподілів. В термінах концепції невизначеності це правило має наступний вигляд

$$\eta(y) = \sum_{j=1}^m \eta(x_j) u_j^4(y) / u_c^4(y), \quad (7.19)$$

де $\eta(y)$, $\eta(x_j)$ – ексцеси вимірюваної y та j -ї вхідної x_j величин, відповідно; $u_j(y)$ – внесок невизначеності j -ї вхідної величини в невизначеність вимірюваної величини; $u_c(y)$ – сумарна стандартна невизначеність, розрахована за формулою

$$u_c(y) = \sqrt{\sum_{j=1}^m u_j^2(y)}. \quad (7.20)$$

Залежність між коефіцієнтами покриття та ексцесами симетричних розподілів наведена у табл. 7.4 та на рис. 7.5, а.

Тому, знаючи ексцеси розподілу вхідних величин, можна знайти ексцес розподілу результату вимірювання за формулою (7.12), а знаючи загальну залежність $k = \varphi(\eta)$, можна обчислити розширену невизначеність як

$$U = k u_c(y). \quad (7.21)$$

Таблиця 7.4 – Значення ексцесів та коефіцієнтів покриття для законів розподілу

Закон розподілу	Ексцес η	Коефіцієнт покриття k для ймовірності p	Значення k для $p=0,95$
Арсинусний	-1,5	$\sqrt{2} \sin(p\pi / 2)$	1,40985
Рівномірний	-1,2	$p\sqrt{3}$	1,64545
Трикутний	-0,6	$1 - \sqrt{1-p} \sqrt{6}$	1,90177
Нормальний	0	$t_p(\infty)$	1,95996

Однак при дослідженні композицій законів розподілу, зазначених у табл. 2, стрункість цієї теорії порушується навіть при обмеженні оцінювання невизначеності складовими типу B (див. рис. 7.4). Тому в роботі було отримано аналітичні вирази та апроксимуючі формули для композицій законів розподілу внесків типу B , які найчастіше застосовуються на практиці (табл. 1). Результати проведених досліджень зведені до табл. 7.5, в якій α позначає співвідношення між внесками невизначеності першого та другого закону розподілу.

При наявності двох внесків невизначеності, один з яких розподілений нормально, а інший – рівномірно, композиція їх законів розподілу має вигляд:

$$f(x) = \frac{1}{2\alpha u_n^2 \sqrt{6\pi}} \int_{-\alpha u_n \sqrt{3}}^{\alpha u_n \sqrt{3}} e^{-\frac{x-y}{2}} dy. \quad (2.22)$$

Значення інтегральної функції розподілу в точці розширеної невизначеності U_{np} для ймовірності 0,95 дорівнює:

$$F_{np}(U_{np}) = \frac{1}{2} + \frac{1}{2\alpha u_n^2 \sqrt{6\pi}} \int_0^{U_{np}} \int_{-\alpha u_n \sqrt{3}}^{\alpha u_n \sqrt{3}} e^{-\frac{x-y}{2}} dy dx = 0,975. \quad (2.23)$$

З цього виразу методом послідовного наближення було обчислено значення $U_{нр}$, а потім і $k=U_{нр}/u_{нр}$.

Таблиця 7.5 – Ексцеси та коефіцієнти покриття для композицій законів розподілу

Закони розподілу	Формула для ексцесу η	Значення ексцесу	Формула для коефіцієнта покриття k для ймовірності p	Значення k для $p=0,95$
2 рівномірних	$\eta = -1,2 \frac{1 + \alpha^4}{1 + \alpha^2}$	-1,2...-0,6	$\sqrt{3} \frac{1 + \alpha - \sqrt{4(1-p)\alpha}}{\sqrt{1 + \alpha^2}};$ $-0,3604\alpha^2 + 0,6754\alpha + 1,5831$	1,65...1,9
3 рівномірних	–	-0,4	$3 - 2\sqrt{3(1-p)}$	1,93734
4 рівномірних	–	-0,3	$\sqrt{3} 2 - \sqrt[4]{12(1-p)}$	1,93970
5 рівномірних	–	-0,24	–	1,94247*
6 рівномірних	–	-0,2	–	1,9453*
7 рівномірних	–	-0,1714	–	1,9476*
Рівномірний та нормальний	$\eta = -1,2 \frac{\alpha^4}{1 + \alpha^2}$	-1,2...0	$-0,197\alpha^2 + 0,5258\alpha + 1,5947$	1,65...1,96
2 арксинусних	$\eta = -1,5 \frac{1 + \alpha^4}{1 + \alpha^2}$	-1,5...-0,75	–	1,41...1,84
Арксинусний та нормальний	$\eta = \frac{-1,5}{1 + \alpha^2}$	-1,5...0	–	1,41...1,96
Арксинусний та рівномірний	$\eta = \frac{-1,5 - 1,2\alpha^4}{1 + \alpha^2}$	-1,5...-0,67	–	1,41...1,88

* розрахунки значень k виконувались методом Монте-Карло

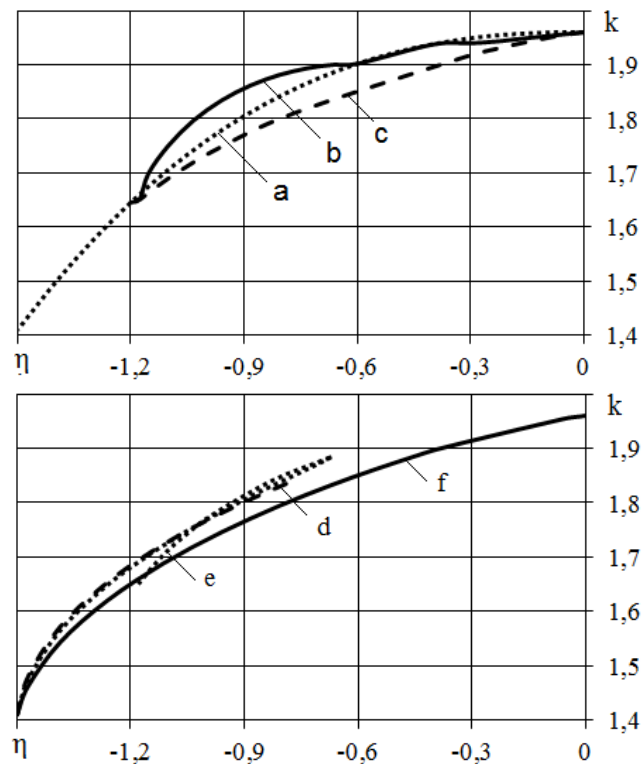
Обчислення параметрів внесків невизначеності типу A виконувалося з розрахунку, що закон розподілу випадкової похибки – нормальний. Тоді ексцеси для розподілу внесків типу A будуть дорівнювати нулю. Крім того,

стандартні невизначеності цих внесків обчислювалися з поправкою на коефіцієнт надійності, який при малій кількості вимірювань n дорівнює

$$h = t_p(\nu) / t_p(\infty), \quad (7.24)$$

де $t_p(\nu)$ – коефіцієнт Стюдента для числа степенів свободи $\nu = n - 1$ та рівня довіри p .

Були проведені розрахунки залежності коефіцієнтів покриття від ексцесів розподілу (рис. 7.6).



a – побудована за даними табл. 2; b – композиція рівномірних законів; c – композиція рівномірного та нормального законів; d – композиція арксинусних законів; e – композиція арксинусного та рівномірного законів; f – композиція арксинусного та нормального законів.

Рисунок 7.6 – Залежності коефіцієнту покриття k від ексцесу η

Оскільки, як видно з рис. 7.5, немає однозначної залежності між коефіцієнтом покриття та ексцесом розподілу, була отримана апроксимуюча

залежність $k=\varphi(\eta)$, яка забезпечує мінімальну систематичну похибку оцінювання невизначеності вимірювань для всіх можливих практичних випадків

$$k=0,1085\eta^3,+0,1\eta+1,96, \quad (7.25)$$

Вираз (7.25) забезпечує апроксимацію наведених вище залежностей з відносною похибкою не більше $\pm 2,5 \%$, а без урахування арксинусного закону – не більше $\pm 1,3 \%$ для всіх чисел степенів свободи.

Було розроблено бюджет невизначеності вимірювань (табл. 7.5), який враховує закони розподілу внесків невизначеності типу *B* і легко автоматизується в програмі Excel.

Порівняно з бюджетом невизначеності, що використовується при оцінюванні невизначеності у відповідності з методикою GUM, він має додатковий стовпець з ексцесами внесків розподілів, які задаються виходячи з табл. 7.6. Сумарна стандартна невизначеність обчислюється з урахуванням коефіцієнтів надійності

$$u_c(y) = \sqrt{\sum_{j=1}^m c_j^2 h_j^2 u^2(x_j)}. \quad (7.26)$$

Коефіцієнт покриття обчислюється за формулою (7.25), а розширена невизначеність – за формулою (7.14).

Таблиця 7.6 – Бюджет невизначеності з ексцесами розподілів

Вхідні величини	Значення вхідних величин	Стандартні невизначеності і вхідних величин	Ексцеси розподілів вхідних величин	Число степенів свободи	Коефіцієнт и чутливості	Внески невизначеності
X_1	x_1	$u(x_1)$	η_1	ν_1	c_1	$h_1 c_1 u(x_1)$
X_2	x_2	$u(x_2)$	η_2	ν_2	c_2	$h_2 c_2 u(x_2)$
...
X_m	x_m	$u(x_m)$	η_m	ν_m	c_m	$h_m c_m u(x_m)$

Вимірювана величина	Результат вимірювання	Сумарна стандартна невизначеність	Експрес розподілу вимірюваної величини	Рівень довіри	Коефіцієнт покриття	Розширена невизначеність
Y	y	$u_c(y)$, (19)	η_c , (12), (13)	$p=0,95$	k , (18)	U , (14)

7.5 Практичні питання щодо оцінювання невизначеності нановимірювань

7.5.1 Вибір мінімальної кількості результатів спостережень при складанні методики виконання вимірювань

При калібруванні ЗВТ часто зустрічаються задачі опрацювання результатів багатократних вимірювань. При розробці методики виконання вимірювань одним з основних питань є питання необхідного числа n проведених вимірювань. При цьому збільшення кількості вимірювань призводить до зменшення розширеної невизначеності, проте збільшує трудомісткість робіт. Це ставить задачу визначення мінімального числа спостережень, достатнього для забезпечення заданої розширеної невизначеності $U(y)$.

В розділі показано, що коли внесками невизначеностей типу B $u_B(y)$ ігнорувати не можна, будемо мати залежність (рис. 7.7)

$$\gamma = U(y)/u_B(y) = t_{0,95} (n-1) \left[1 + n/\alpha^2 \right]^2 \cdot \sqrt{1 + \alpha^2/n}, \quad (7.27)$$

де $\alpha = s/u_B(y)$, а s – СКВ окремих спостережень.

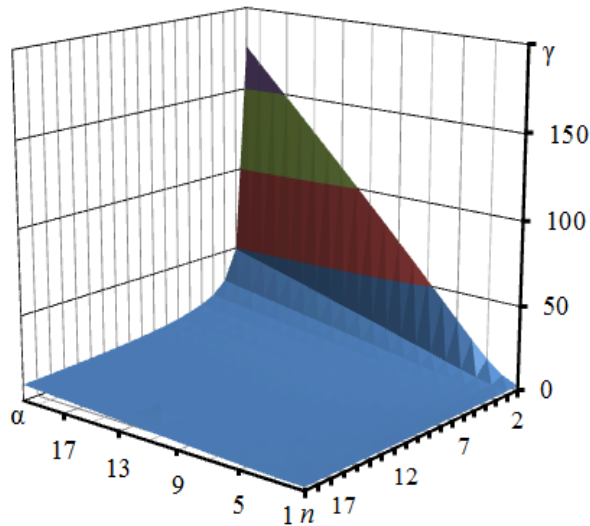


Рисунок 7.7 – Залежність $\gamma = U(y)/u_B(y)$ від α та n

При визначенні цієї залежності отримуємо дробові числа степенів свободи, тому для обчислення значень коефіцієнта Стюдента для ймовірності 0,95 застосовувався вираз (7.8). Використовуючи залежність $\gamma = \varphi(n, \alpha)$, можна для заданого значення $U(y)$ і відомих значень $u_B(y)$ та s , визначити необхідну кількість спостережень n .

У випадку, коли внески невизначеностей типу B можна ігнорувати, розширена невизначеність $U(y)$ буде повністю визначатися їх стандартною невизначеністю типу A $u_A(y)$. З урахуванням $u_A(y) = s/\sqrt{n}$, при відомому s , що визначається попередньо, та заданому $U(y)$ була розрахована залежність відношення $\beta = U(y)/s$ від n та побудована зворотна залежність $n = f(\beta)$. Для практичного використання була отримана апроксимація цієї залежності $n = f(\beta)$ на інтервалі $n = 2, 10$ з абсолютною похибкою не більше $\pm 0,046$:

$$n = 3,1/\beta^2 + 1,5/\beta + 1,8. \quad (7.28)$$

Розраховане за формулою (7.28) або отримане з рис. 7.6 дробове значення n округляється у більшу сторону до найближчого цілого числа.

7.5.2 Оцінювання невизначеності вимірювань диференційної функції

Широко розповсюдженою метрологічною задачею є обробка результатів вимірювань з використанням диференційної функції. Вона застосовується, наприклад, при знаходженні ступенів еквівалентності національних еталонів під час ключових звірень; при калібруванні двох вимірювальних приладів за допомогою засобу порівняння; під час реалізації диференційного методу вимірювань; під час проведення міжлабораторних звірень для виявлення наявності систематичного зсуву в результатах вимірювання двох лабораторій, вимірювальних приладів або операторів; при перевірці відповідності результатів випробувань вимогам нормативних документів; при валідації методик калібрування ЗВТ тощо. Зазначені задачі зводяться до опрацювання результатів двох груп вимірювань. Їх особливістю є те, що результат вимірювань порівняний, а інколи й суттєво менший за його невизначеність, при цьому ситуація може посилюватися наявністю кореляції між результатами вимірювання вхідних величин.

Розглянута диференційна функція, яка є поширеним модельним рівнянням в метрології. Її особливістю є те, що результат вимірювання може дорівнювати, а іноді бути істотно меншим його невизначеності, при цьому ситуація може посилюватися наявністю кореляції між результатами вимірювання вхідних величин X_1, X_2 . Результати дослідження невизначеності диференційної функції наведено в табл. 7.7.

Найбільші труднощі викликає випадок, коли значення величин X_1, X_2 є результатами багатократних корельованих вимірювань оскільки, при наявності кореляції формула Велча-Саттерсвейта для розрахунку ефективного числа степенів свободи дає недостовірні результати. Дійсно, визначаючи розширену невизначеність за методикою GUM, будемо мати ефективне число степенів свободи

$$v_{eff}^* = (n-1) \left[u^4(y) / u_A^4(x_1) + u_A^4(x_2) \right]. \quad (7.29)$$

Таблиця 7.7 – Невизначеність для диференційної модельної функції

Параметр	Величина X_2 не має невизначеності	X_1, X_2 є результатами багатократних некорельованих вимірювань	X_1, X_2 є результатами багатократних корельованих вимірювань
$u_A(y)$	$u_A(x_1)$	$\sqrt{u_A^2(x_1) + u_A^2(x_2)}$	$\sqrt{u_A^2(x_1) - 2ru_A(x_1)u_A(x_2) + u_A^2(x_2)}$
$u_B(y)$	$u_B(x_1)$	$\sqrt{u_B^2(x_1) + u_B^2(x_2)}$	
$u^2(y)$	$\sqrt{u_A^2(y) + u_B^2(y)}$		
$U(y)$	$ku(x_1)$	$ku(y)$	
k	$t_{0,95}(v_{eff})$		
v_{eff}	$(n_1 - 1) \left[\frac{u(x_1)}{u_A(x_1)} \right]^4$	$\frac{u^4(y)}{\frac{u_A^4(x_1)}{(n_1 - 1)} + \frac{u_A^4(x_2)}{(n_2 - 1)}}$	$v_{eff} = (n - 1) \left[\frac{u(y)}{u_A(y)} \right]^4$

Відомо, що надійну оцінку розширеної невизначеності при наявності кореляції між вхідними величинами можна отримати, застосувавши метод приведення (редукції). У цьому випадку сумарна стандартна невизначеність типу A результату вимірювання $u_A(y)$ має число степенів свободи $\nu = n - 1$. З урахуванням цієї обставини, формула Велча-Саттерсвейта для достовірного оцінювання числа степенів свободи при наявності корельованих вхідних величин буде мати наступний вигляд

$$v_{eff} = (n - 1) \frac{u(y)}{u_A(y)}^4. \quad (7.30)$$

В розділі визначена відносна похибка застосування виразу (7.22) для оцінювання розширеної невизначеності для двох корельованих вхідних величин:

$$\delta = \left[t_{0,95}(v_{eff}^*) / t_{0,95}(v_{eff}) - 1 \right] \cdot 100 \% . \quad (7.31)$$

Розрахунок δ проводився для діапазонів вимірювань $n \in 3; 20$, $r \in -1; 1$, а також співвідношень $u_A(x_2)/u_A(x_1) \in 0; 1$ та $u_B(Y)/u_A(X_1) \in 0; 10$. При розрахунках отримували дробові значення ефективних чисел степенів свободи, тому використовувалася апроксимація коефіцієнта Стьюдента формулою (7.8).

Результати розрахунку екстремальних значень похибки наведені на рис. 7.8, з якого видно, що для негативної кореляції похибка використання формули Велча-Саттерсвейта при наявності корельованих вхідних величин змінюється від 0 до -51% і досягає абсолютного максимуму при кількості повторних спостережень $n=3$ та коефіцієнті кореляції $r = -1$, при відсутності внесків типу B та рівних внесках стандартної невизначеності типу A $u_A(x_1) = u_A(x_2)$.

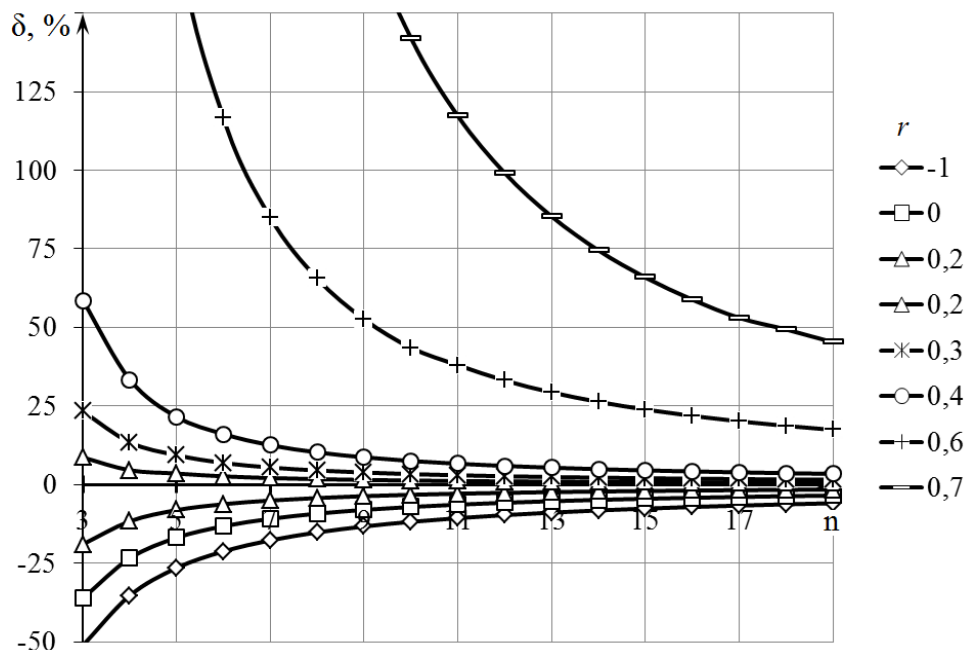


Рисунок 7.8 – Залежність екстремальних значень похибки δ від значення коефіцієнта кореляції r і числа вимірювань n

В останньому випадку

$$u(y) = |u_A(x_1) + u_A(x_2)|, \quad (7.32)$$

тому $v_{eff}^* = 8(n-1)$ та $v_{eff} = (n-1)$.

При наявності внесків типу B та зменшенні співвідношення $\alpha = u_A(X_2)/u_A(X_1)$ від 1 до 0, значення похибки δ зменшується.

При наявності позитивної кореляції ефективне число степенів свободи може приймати значення менше одиниці (до нуля) і похибка використання формули Велча-Саттерсвейта прямує до нескінченності (рис. 6)

У разі повної позитивної кореляції ($r=1$):

$$u_A(y) = |u_A(x_1) - u_A(x_2)| \quad (7.33)$$

і при $u_A(x_1) = u_A(x_2)$ невизначеність типу A вимірюваної величини $u_A(y) = 0$. Тоді вираз (7.23) обертається в нескінченність і закон розподілу вимірюваної величини буде композицією законів розподілу невизначеностей типу B вхідних величин.

7.6 Висновки до розділу 7

1. Показано, що міжнародна стандартизація буде відігравати важливу роль у забезпеченні того, щоб потенціал нанотехнологій здійснився в повній мірі і нанотехнології були надійно інтегровані в суспільство. Стандарти допоможуть створити плавний перехід від лабораторії до ринку, посуваючи прогрес уздовж ланцюжка нанотехнологій – від нанорозмірних матеріалів, які формують будівельні блоки для компонентів і пристроїв до інтеграції цих пристроїв у функціональні системи – для полегшення світової торгівлі.

2. На основі проведеного аналізу нормативної документації та науково-технічної літератури виявлено недоліки існуючих методів обробки результатів та оцінювання невизначеності вимірювань, сформульовано шляхи їх подолання в напрямку підвищення достовірності отримуваних оцінок невизначеності вимірювань за рахунок обліку законів розподілу вхідних величин, кореляції між результатами їх спостережень та нелінійності модельного рівняння.

3. На основі отриманих значень коефіцієнта Стюдента для дробових чисел степенів свободи розроблена точна інтерполуюча формула, яка дозволила підвищити достовірність оцінювання невизначеності вимірювань для малих значень степенів свободи та удосконалити розглянуті в дисертації методи. Удосконалено метод транспозицій даних, який дозволяє отримати незміщену оцінку результату вимірювання та достовірну оцінку його невизначеності. Отримана формула для розрахунку еквівалентного числа вимірювань при різних кількостях спостережень вхідних величин. Доведено, що при використанні методу транспозицій уявна кореляція вхідних даних не впливатиме на обробку результатів багатократних вимірювань. Показано, що дисперсія квадрата оцінки стандартної невизначеності типу A , що отримана традиційним методом буде в декілька разів більше дисперсії, яка отримана методом транспозицій.

4. Розроблено загальний метод оцінювання результату та стандартної невизначеності вимірювань за типом A для корельованих та некорельованих оцінок вхідних величин на основі композиції методів транспозицій та редукції, який дозволяє отримувати незміщені оцінки чисельних значень результату та достовірні оцінки невизначеності багатократних вимірювань для нелінійних модельних рівнянь.

5. Отримано вирази для оцінювання сумарної стандартної невизначеності вимірювань за типом B , які дозволяють враховувати закони розподілів внесків невизначеностей цього типу та їх співвідношення.

6. Удосконалено метод ексцесів для оцінювання невизначеності

вимірювань з урахуванням законів розподілу внесків невизначеності типу *B* і коефіцієнтів надійності для внесків невизначеності типу *A*. Показано, що максимальна відносна похибка використання удосконаленого методу не перевищує $\pm 2,5\%$, для всіх законів розподілу, чисел степенів свободи і співвідношень між внесками невизначеності. Запропоновано бюджет невизначеності, що дозволяє враховувати закони розподілу внесків типу *B* та числа степенів свободи внесків типу *A* для оцінювання розширеної невизначеності.

7. Отримано залежності необхідного числа вимірювань від заданої розширеної невизначеності для випадків відсутності та наявності вкладів невизначеності типу *B*, які дозволяють мінімізувати час виконання випробувань та калібрувань.

8. Отримано вирази для стандартних і розширених невизначеностей вимірювань різниці двох вхідних величин з урахуванням їх невизначеностей та кореляції між результатами їх спостережень. Показано, що застосування формули, рекомендованої GUM для оцінки числа степенів свободи при наявності кореляції, може, за певних умов, привести до нескінченної похибки оцінювання розширеної невизначеності вимірювань. Запропонована формула, яка усуває позначений недолік.

ЗАКЛЮЧЕННЯ

В результаті проведених досліджень отримані найважливіші наукові і науково-технічні результати:

1. Створені математичні моделі біологічних систем з урахуванням впливу на них ЕМП різних довжин хвиль та інтенсивності. В якості біологічних систем використовувалися імунна система людини та кров, як цілісна система формених елементів (еритроцитів, лейкоцитів, тромбоцитів).

2. Для концептуального вивчення впливу ЕМП мм діапазону та різної інтенсивності на кров і її формені об'єкти (еритроцити, лейкоцити, тромбоцити) був вибраний загальний алгоритм побудови детермінованої математичної моделі такого процесу на мікро- і макро- рівнях. Так, була створена математична модель інфекційного захворювання, що описує зміни в організмі людини з урахуванням дії різноманітних зовнішніх факторів, включаючи вплив патогенних мікроорганізмів і низкоінтенсивного ЕМП НВЧ діапазону.

3. Створена математична модель крові, заснованої на принципі декомпозиції і побудові системи диференціальних рівнянь, що враховують особливості формених її елементів. При цьому важливим питанням при дослідженні впливу ЕМП НВЧ є здійснення візуального контролю за змінами, які виникатимуть в середовищі (крові). Використання для дослідження субмікронних біологічних об'єктів, зокрема крові, методів класичної оптичної мікроскопії має свої недоліки та обмеження, пов'язані з граничним збільшенням просторової роздільної здатності в силу наявності явища дифракції хвиль (формула Аббе), складністю забезпечення високого контрасту в порівнянні з фоном при вивченні прозорих біологічних об'єктів (фазових об'єктів), а також неможливість (або технічну складність) реалізації 3-D візуалізації фазових мікроб'єктів. Більш перспективним напрямком у сфері дослідження субмікронних біологічних фазових об'єктів є 3-D цифрова голографічна мікроскопія (*Inkyu M., Mehdi D., Arun A., Bahram J. Cell*

Identification with Computational 3-D Holographic Microscopy. Optics & Photonics News, V. 22, № 6, June 2011, p.p. 18-23).

4. Для отримання об'ємних зображень фазових мікрооб'єктів крові використовується 3-D цифровий голографічний мікроскоп. В якості робочої схеми мікроскопа використовується схема на просвіт. Когерентне лазерне випромінювання з довжиною хвилі 534 нм розділяється на опорну і об'єктну хвилі. Опорна хвиля накладається на об'єктну хвилю, що пройшла через досліджуваний мікрооб'єкт. В результаті виникає інтерференційна картина, яка містить повну інформацію про оптичні властивості біологічного середовища. Сформовані голограми реєструються і записуються за допомогою цифрової камери і вводяться в комп'ютер для подальшої автоматизованої обробки.

5. Пропонований підхід розширює можливості цифрової 3-D голографічного мікроскопії і дозволяє автоматизувати процес обробки інтерферограмма фазових мікрооб'єктів в режимі реального часу. Для досягнення цього використовується математична обробка зображень за допомогою запропонованих методів нечіткої кластеризації.

6. На основі запропонованих алгоритмів пропонується програмне забезпечення для автоматизації процесу обробки та аналізу біологічних об'єктів крові до і після впливу зовнішнього електромагнітного поля.

7. Розроблено нові нейро-фаззі системи спеціального типу, в яких паралельно реалізуються як процеси навчання, так і самонавчання в залежності від рівня невизначеності щодо вхідної інформації.

8. Розроблено нейро-фаззі систему та метод її навчання, що призначені для обробки неоднорідних полів спостережень, що описують явища різної природи, включаючи зображення, задані у цифровій формі. Введено матричну форму метода нечітких с-середніх, що призначено для вирішення задач сегментації та кластеризації за умов класів, що перетинаються. Запропоновано метод послідовного оброблення даних, що надходять у систему у вигляді «ковзного» вікна. Показані переваги цього підходу перед

традиційними, коли поля спостережень перед обробленням перетворюються у векторну форму, а після розрахунків девекторизується.

8. Розроблено метод гібридного навчання багатошарових нейро-фаззі систем побудови математичних моделей нелінійних нестационарних об'єктів. Особливістю введеного методу є те, що він призначений для навчання не лише на основі традиційного квадратичного критерія, але і на основі мінімізації емпіричного ризику. При цьому в систему введено додатковий (вихідний) шар, що обраховує рівні належності отриманих результатів до деякого гіпотетичного оптимального вихідного сигналу, а також передбачена можливість настроювання не лише синаптичних ваг, але й центрів і параметрів ширини функції активації-належності. При цьому у випадку нестационарних об'єктів в контурі навчання передбачено «придушення» застарілої інформації.

9. В ситуації, коли кількість настроюваних параметрів є співрозмірною з обсягом навчальної вибірки було введено багатошарову класифікуючу нейро-фаззі систему навчання якої відбувається за принципом «нейрони в точках даних». При цьому центри функцій належності встановлюються у відповідних координатах вхідних багатовимірних сигналів, а замість синаптичних ваг використовуються відповідні відліки навчального сигналу.

10. Проведені дослідження є новими, аналоги таких досліджень не відомі.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Борн М., Вольф Э. Основы оптики. – М.: Наука, 1973. – 719 с.
2. Сорока Л.М. Голография и интерференционная обработка информации. УФН, т. 90, вып. 1, 1966, с. 3 – 45.
3. Соколов М.Э. Современные методы визуализации фазовых объектов. Голографические методы и аппаратура, применяемые в физических исследованиях. – М.: Наука, 1987. 259 с.
4. Тишко Т.В., Титарь В.П., Тишко Д.Н. Голографическая микроскопия. Трехмерная визуализация фазовых микрообъектов. // Прикладная радиоэлектроника. 2009. – Т. 8. – № 1. – С. 40-45.
5. Inkyu M., Mehdi D., Arun A., Bahram J. Cell Identification with Computational 3-D Holographic Microscopy. Optics & Photonics News, V. 22, № 6, June 2011, p.p. 18-23.
6. Gabor D. A New Microscopic Principle. Nature. – 1948.- V. 161, p.p. 777 – 778.
7. Пресман А.С. Электромагнитные поля и живая природа. – М.: Наука, 1968. – 227 с.
8. Шван Х.П. Воздействие высокочастотных полей на биологические системы: Электрические свойства и биофизические механизмы // ТИИЭР, т. 68, № 1, 1980. – С. 121 – 132.
9. Девятков Н.Д., Голант М.Б., Бецкий О.В. Миллиметровые волны и их роль в процессах жизнедеятельности. – М.: Изд-во «Радио и связь», 1988. – 63 с.
10. Гай, Леманн, Стоунбридж. Применение электромагнитной энергии в терапии // ТИИЭР, т. 62, № 1, 1974. – С. 66 – 93.
11. Макри Д.И. Исследование нетепловых резонансных эффектов мм-излучения как начало новой биофизике // ТИИЭР, т. 68, № 1, 1980. – С. 40 – 48.

12. Нетепловые эффекты миллиметрового излучения / Под ред. Н.Д. Девяткова. – М.: ИРЭ АН СССР, 1981. – 338 с.
13. Бецкий О.В., Голант М.Б., Девятков Н.Д. Миллиметровые волны в биологии. – М.: Изд-во «Знание», 1988. – 63 с.
14. Кватново-биологическая теория. / Под. ред. В.В. Бойка и М.О. Красноголовца. – Х.: «Факт», 2003. – 968 с.
15. Радиофизические основы и опыт применения СВЧ электромагнитных полей в хирургии. / Под. ред. проф. В.В. Бойка. – Х.: Новое слово, 2009. – 166 с.
16. Москаленко В.Ф., Сітько С.П., Горбань Є.М., Грубник Б.П., Яненко О.П. Квантова медицина: від фундаментальних основ до практичного використання. Український медичний часопис. - № 2 (28), 2002, 106 -109 с.
17. Эйди У.Р. Частотные и энергетические окна при воздействии слабых электромагнитных полей на живую ткань // ТИИЭР, т. 68, № 1, 1980. – С. 128 – 147.
18. Riadh W.Y. Nabash, Rajeev Bansal, Daniel Krewski, Hafid T. Alhafid Thermal Therapy, Part 1: An introduction to Thermal Therapy, Critical Reviews in Biomedical Engeneering, 34 (6). – 2006, p-p 459-489
19. С.П. Осинский Гипертермия в клинической онкологии: современное состояние проблемы (по итогам 20-й ежегодной конференции Европейского общества гипертермической онкологии (ESHО)), Онкология, т.4 № 4, 2002, 288 – 292.
20. Бецкий О.В., Кислов В.В., Лебедев Н.Н. Миллиметровые волны и живые системы. – М.: «Сайнс-пресс», 2004. – 107 с.
21. О.И. Белоус, Б.Г. Емец, В.А. Малахов и др. Роль КВЧ излучения в повышении биологической активности крови, Радиофизика и электроника, т.13, №4, 2008, с. 556-561
22. А.Л. Чижевский Электрические и магнитные свойства эритроцитов, Киев «Наукова думка». – 1973, - 94 с.

23. C. NĂDEJDE, D.E. CREANGĂ, C. GOICEANU Radiofrequency electromagnetic wave and paramagnetic particle effects on the red blood cells Rom. Journ. Phys. vol. 54, nos. 1-2, P. 105-114, Bucharest, 2009
24. Г.Р. Иваницкий Математическая биофизика клетки/В.И. Кринский, Е.Е. Сельков, М.: Наука. - 1978.- 310 с.
25. Ризниченко Г.Ю. Математические модели в биофизике и экологии. – Москва-Ижевск: Институт компьютерных исследований.- 2003.-184 с.
26. J.D. Murray, Mathematical Biology I. An Introduction. -3rd ed. Springer – 2001.- 551 p.
27. Погожев И.Б. Применение математических моделей заболеваний в клинической практике. 1988
28. Зуев С.М. Математические модели и методы анализа медико-биологических данных. 1990.
29. Марчук Г.И. Математические модели в иммунологии. Вычислительные методы и эксперименты. – 3-е изд., перераб. и доп. – М.: Наука. Гл.ред. физ.-мат. лит., 1991. – 304.
30. Galstyan S.G., Perova I.G., Churyumov G.I. Mathematical model of the influence of electromagnetic radiation on biological nano-objects / VI Международная научная конференция «Функциональная база нанoeлектроники». Сборник научных трудов. – Харьков: ХНУРЭ. 2013. с. 121-124
31. Чуян Е.Н., Трибрат Н.С. Влияние низкоинтенсивного электромагнитного излучения крайне высокой частоты на процессы микроциркуляции. Ученые записки Таврического национального университета им. В.И. Вернадского. Серия «Биология, химия». Том 21 (60). 2008. № 1. С. 156 – 166.
32. Jackson J.D. Classical Electrodynamics. Third Edition. John Wiley & Sons, Inc. 2001. – 807 с.
33. Березовский В.А., Колотилов Н.Н. Биофизические характеристики тканей человека. – Киев: Наукова думка. 1990. – 224 с.

34. Tzirtzilakis E.E. A mathematical model for blood flow in magnetical field. *Physics of fluids* 17, 2005. 077103-1-15.
35. Махонина М.М., Чуян Е.Н., Бержанский В.Н., Попов В.В. Изменение содержания серотонина в лейкоцитах крови крыс при действии низкоинтенсивного электромагнитного излучения крайне высокой частоты. *Ученые записки Таврического национального университета им. В.И. Вернадского. Серия «Биология, химия»*. Том 19 (58). 2006. № 4. С. 151 – 160.
36. Taflove A., Hagness S.C. *Computational Electrodynamics. The Finite Difference Time Domain Method*. – 3rd Edition. – 2005. – 997с.
37. Kunz K.S., Luebbers R.J. *The Finite Difference Time Domain Method for Electromagnetics*. CRC Press. 1993. – 446 с.
38. Schwan H.F. *Electrical Properties of Tissues // Adv. Biol. And Med. Phys.* – 1957. - # 5. – p. 147.
39. Schwan H.F., Li K. *Capacity and Conductivity of Body Tissues of UHF // Proc. IRE.* – 1953. – 41. - # 12. – p. 1735 - 1740.
40. Шван Н. *Электрические свойства тканей*. – В кн. СВЧ-энергетика. – М.: Мир. – 1971. – 227 с.
41. Бецкий О.В., Котровская Т.И., Лебедева Н.Н. *Миллиметровые волны в биологии и медицине. III Всероссийская конференция «Радиолокация и радиосвязь» – ИРЭ РАН. 2009.*
42. Малахов А.Н. *Флуктуации в автоколебательных системах*. –М.: Наука. 1968. – 660 с.
43. Lehmann J. F. *Ultrasound Therapy // in Therapeutic Heat and Cold*, S. Licht, Ed. New Haven, Conn.: Licht, 1965, sec. 13, p.p. 321 – 386.
44. // <http://www.health.org>. История развития ультразвуковой дигностикию
45. Буц В.А., Скибенко К.Н. Один из механизмов биологического действия ультразвука. *Зарубежная радиоэлектроника*, 1996, № 12, с.с. 52 – 56.
46. Licht S., *History of Therapeutic heat / in Therapeutic Heat and Cold*, S. Licht, Ed. New Haven, Conn.: Licht, 1965, sec. 6, p.p. 196 – 231.

47. Krusen F.H. and other. Microwave therapy: Preliminary report of experimental studies of the heating effect of microwave (radar) in living tissues. *Proc. Staff Meet. Maya Clin.*, v. 22, 1947, pp. 209 – 224.
48. Schwan H.P., Piersal G.M. The Absorption of Electromagnetic Energy in Body Tissues. Pt. 1. *Amer. J. Phys. Med.*, vol. 2, 1954, pp. 371 – 404.
49. Kerker, M. The Scattering of Light and Other Electromagnetic Radiation [Text] / M. Kerker. - N. Y.: Academic Press, 1969. – 666 p.
50. Ashkin, A. Observation of a single-beam gradient force optical trap for dielectric particles [Text] / A. Ashkin // *Opt. Lett.* - 1986. - V. 11, No. 5. - P. 288-290.
51. Svoboda, K. Optical trapping of metallic Rayleigh particles [Text] / K. Svoboda, S. M. Block // *Opt. Lett.* - 1994. - V. 19, No. 13. - P. 930-932.
52. Allen, L. Optical Resonance and Two-Level Atoms [Text] / L. Allen, J. H. Eberly. - N.Y.: Dover, 1987. – 256 p.
53. Moothoo, D. N. Beth's experiment using optical tweezers [Text] / D. N. Moothoo // *Am. J. Phys.* - 2001. - V. 69, No. 3. - P. 271-276.
54. Malagnino, N. Measurements of trapping efficiency and stiffness in optical tweezers [Text] / N. Malagnino // *Opt. Commun.* - 2002. - V. 214. - P. 15-24.
55. Maheu, B. A concise presentation of the generalized Lorenz-Mie theory for arbitrary location of the scatterer in an arbitrary incident profile [Text] / B. Maheu, G. Gouesbet, G. Grehan // *J. Opt. (Paris)*. - 1988. - V. 19, No. 2. - P. 59-67.
56. Ren, K. F. Prediction of the reverse radiation pressure by generalized Lorenz-Mie theory [Text] / K. F. Ren, G. Grehan, G. Gouesbet. // *Appl. Opt.* - 1996. - V. 35. - P. 2702-2710.
57. Wohland, T. Theoretical determination of the influence of the polarization on forces exerted by optical tweezers [Text] / T. Wohland, A. Rosin, E. H. K. Stelzer // *Optik*. - 1996. - V. 102, No. 4. - P. 181-190.
58. Nieminen, T. A. Calculation and optical measurement of laser trapping forces on non-spherical particles [Text] / T. A. Nieminen, H. Rubinsztein-Dunlop, N.

- R. Heckenberg // *J. Quantitative Spectroscopy & Radiative Transfer*. - 2001. - V.70. - P. 627-637.
59. Делоне, Н. Б. Основы нелинейной оптики атомарных газов [Текст] / Н. Б. Делоне, В. П. Крайное. - М.: Наука, 1986. – 184 с.
60. Крайное, В. П. Ориентация и фокусировка молекул полем лазерного излучения [Текст] / В. П. Крайное // *Соросовский образовательный журн.* - 2000. - Т. 6, №4. - С. 90-95.
61. Ashkin, A. Radiation pressure on free liquid surface [Text] / A. Ashkin, J. M. Dziedzic // *Phys. Rev. Lett.* - 1973. - V. 30, No. 4. - P. 139-142.
62. Ashkin, A. Continuous-wave self-focusing and self-trapping of light in artificial Kerr media [Text] / A. Ashkin, J. M. Dziedzic, P. W. Smith // *Opt. Lett.* - 1982. - V. 7, No. 6. - P. 276-278.
63. Ashkin, A. Acceleration and trapping of particles by radiation pressure [Text] / A. Ashkin // *Phys. Rev. Lett.* - 1970. - V. 24, No. 4. - P. 156-159.
64. Ashkin, A. Observation of resonances in the radiation pressure on dielectric spheres [Text] / A. Ashkin, J. M. Dziedzic // *Phys. Rev. Lett.* - 1977. - V. 38, No. 23. - P. 1351-1354.
65. Ashkin, A. Stability of radiation-pressure particle traps: an optical Earnshaw theorem [Text] / A. Ashkin, J. P. Gordon // *Opt. Lett.* - 1983. - V. 8, No. 10. - P. 511-513.
66. Ashkin, A. Observation of radiation-pressure trapping of particles by alternating light beams [Text] / A. Ashkin, J. M. Dziedzic // *Phys. Rev. Lett.* - 1985. - V. 54, No. 12. - P. 1245-1248.
67. Ashkin, A. Optical levitation by radiation pressure [Text] / A. Ashkin, J. M. Dziedzic // *Appl. Phys. Lett.* - 1971. - V. 19. - P. 283-285.
68. Ghislain, L. P. Scanning-force microscope based on an optical trap [Text] / L. P. Ghislain, W. W. Webb // *Opt. Lett.* 1993. - V. 18. - P. 1678-1680.
69. Zemanek, P. Optical trapping of Rayleigh particles using a Gaussian standing wave [Text] / P. Zemanek // *Opt. Commun.* - 1998. - V. 151. - P. 273-285.

70. Freearde, T. Confinement and manipulation of atoms using short laser pulses [Text] / T. Freearde, J. Walz, T. W. Hansch // *Opt. Commun.* - 1995. - V. 117. - P. 262-267.
71. Ashkin, A. Optical trapping and manipulation of single cells using infrared laser beams [Text] / A. Ashkin, J. M. Dziedzic, T. Yamane // *Nature.* - 1987. - V. 330. - P. 769-771.
72. Chu, S. Laser manipulation of atoms and particles [Text] / S. Chu // *Science.* - 1991. - V. 253. - P. 861-866.
73. Tadir, Y. Micromanipulation of gametes using laser microbeams [Text] / Y. Tadir // *Human Reprod.* - 1991. - V. 6. - P. 1011-1016.
74. Steubing, R. W. Laser induced cell fusion in combination with optical tweezers: the laser cell fusion trap [Text] / R. W. Steubing // *Cytometry.* - 1991. - V. 12. - P.505-510.
75. Ashkin, A. Force generation of organelle transport measured in vivo by an infrared laser trap [Text] / A. Ashkin // *Nature.* - 1990. - V. 348, No. 6299. - P. 346-348.
76. Liang, H. Directed movement of chromosome arms and fragments in mitotic newt lung cells using optical scissors and optical tweezers [Text] / H. Liang // *Exp. Cell Res.* - 1994. - V. 213. - P. 308-312.
77. Crocker, J. C. Methods of digital video microscopy for colloidal studies [Text] / J. C. Crocker, D. G. Grier // *J. Colloid Interf. Sci.* - 1996. - V. 179. - P. 298-310.
78. Smith, S. B. Direct mechanical measurements of the elasticity of single DNA molecules using magnetic beads [Text] / S. B. Smith, L. Finzi, C. Bustamante // *Science.* - 1992. - V. 258. - P. 1122-1126.
79. Misawa, H. Multibeam laser manipulation and fixation of microparticles [Text] / H. Misawa // *Appl. Phys. Lett.* - 1992. - V. 60. - P. 310-312.
80. Crocker, J. C. Entropic attraction and repulsion in binary colloids probed with a line optical tweezer [Text] / J. C. Crocker // *Phys. Rev. Lett.* - 1996. - V. 82, No. 21. - P. 352-355.

81. Bancel, P. A. Manipulating crystals with light [Text] / P. A. Bancel, V. B. Cajipe, F. Rodier // *J. Crystal Growth*. - 1999. - V. 196. - P. 685-690.
82. Friese, M. E. J. Optically driven micromachine elements [Text] / M. E. J. Friese // *Appl. Phys. Lett.* - 2001. - V. 78, No. 4. - P. 547-549.
83. Власов, А. И. Оптическая микроскопия [Текст] : учеб. пособие / А. И. Власов. – М. : Изд-во МГТУ им. Н. Э. Баумана, 2011. – 184с.
84. Яковлев Ю.С. Применение многоагентной технологии для разработки компьютерных систем / Ю.С. Яковлев // Засоби комп'ютерної техніки з віртуальними функціями і нові інформаційні технології: зб. наук. праць. - Київ: ІК НАНУ. - 2002. - С. 65-71
85. Кнут Д.Э. Искусство программирования, том 3. Сортировка и поиск, 2-е изд. / Кнут Д.Э. – М.: Вильямс, 2000. – С. 549-597.
86. Luo Zhong. A study of the convergence of the CMAC learning process / Luo Zhong, Xie Yong-Bin, Zhu Chong-Guang // *Acta Automatica Sinica*. – 1997. – №23(4). – P. 455 – 461.
87. НВЧ-энергетика . Под редакцией Э. Окресса. Т. 2. Применение энергии сверхвысоких частот в промышленности. Т.3. Применение энергии сверхвысоких частот в медицине, науке и технике. – М.: Издательство «Мир», 1971.
88. Архангельский Ю.С., Девяткин И.И. Сверхвысокочастотные нагревательные установки для интенсификации технологических процессов. – Саратов: Издат. Саратов. ун-та, 1983.
89. Рогов И.И. Электрофизические методы обработки пищевых продуктов. М.: Агропромиздат, 1988.
90. Бородин И.Ф., Шарков.Г.А., Гарин А.Д. Применение НВЧ-энергии в сельском хозяйстве. – М.: ВНИИТЭИагропром, 1987.
91. Нетепловые эффекты миллиметрового излучения. – М.: РотапринтИРЭ, 1981.
92. Некрутман С.В. Аппараты НВЧ в общественном питании. - М.: Экономика, 1973. Явчуновский В.Я. Микроволновая и комбинированная

- сушка. Физические основы технологии и оборудования. – Саратов: Изд-во Саратовского университета. 1992.
93. Пюшнер Г. Нагрев энергией сверхвысоких частот. – М.: Энергия. 1968.
94. Диденко А.Н., Зверев Б.В. НВЧ-энергетика. – М.: Наука. 2000.
95. Морозов Г.А. Микроволновые технологии в промышленности и сельском хозяйстве: современные подходы// Антенны. Выпуск 1(40), 1988, С. 88-97.
96. Jangu's Thuery. Microwave Industrial, Scientific and Medical Applications Edited by Edward M. Grant. Larstin Arteda House/ Boston/ London. 1992.
97. Мосин О.В. Вплив електромагнітних хвиль низької інтенсивності на воду й біологічні об'єкти. http://mmbio.ru/pdf/st_4.pdf.
98. Коваленко О. И., Кивва Ф. В., Калиниченко С. В., Коворотный А. Л. Особенности прямого и опосредованного воздействия электромагнитных полей низкой интенсивности на семена растений и микроорганизмы // Радиофизика и электроника, том 12, №1, 2007, С. 273-282
99. Горбунов В.В., Диденко Н.П. и др. Влияние электромагнитного излучения миллиметрового диапазона на быструю динамику молекулы гемоглобина. Тезисы 7-го Всесоюзного семинара “Использование низкоинтенсивного электромагнитного излучения в биологии и медицине”, Москва, 1989 , с. 74.
100. Чернавский Д.С. и др. О кооперативных (когерентных) явлениях в биологических явлениях (концепция «когерентного возбуждения» и «белок-машина»). Препринт ФИАН, 1986, № 185.
101. Бетский О.В., Яременко Ю.Г. Кожа и электромагнитные волны. Миллиметровые волны в биологии и медицине.1998, №1(11), сс. 3-15.
102. Диденко Н.П. и др. Исследование многорезонансного взаимодействия электромагнитных колебаний с молекулой гемоглобина с помощью мессбауэровской спектроскопии, Труды НИИ ядерной физики, Томск. 1983, №10, сс. 77 -81

103. Бецкий О.В., Лебедева Н.Н., Котровская Т.И. Стохастический резонанс в медицине и биологии // Биомедицинские технологии и радиоэлектроника, 2003. № 1.С. 3–9.
104. Карнаухов А.В. Диссипативный резонанс и его роль в механизмах действия электромагнитного излучения на биологические и физико-химические системы // Биофизика, 1997, Т. 42, вып 4. С. 971 – 979.
105. Мазец Ж.Э., Кайзинович К.Я., Пушкина Н.В., Родионова В.Н., Спиридович Е.В. Влияние низкоинтенсивного электромагнитного излучения на активность амилазы в проростках *lupinus angustifolius* L // Труды БГУ, том 8, часть 2. 2013. С.95-101.
106. Ча В.А. Глава 2. Влияние электромагнитного излучения на биологические объекты различного уровня организации // <http://portal.tpu.ru:7777/SHARED/v/VCHA/Tab/glava-2.doc>.
107. В.А. Киреев. Курс физической химии . Изд.»Химия» М. 1975 . 775 с.
108. Н.А. Глинка. Общая химия. Л.: Изд.»Химия»,1980, 719 с.
109. Зима И.И. Роторный геомагнетизм. Немного фантазии. – Харьков: Знаки, 2014. 228 с.
110. Маргулис М.А. Основы звукохимии.- М.: Высшая школа, 1984.
111. Захаров І.П., Павленко Ю.Ф., Гусельніков В.К., Кондрашев С.І., Забезпечення єдності електрорадіовимірювань // Під редакцією д.т.н., проф. Ю.Ф. Павленка Харків: Підручник НТУ «ХПІ» 2014, 236 с.
112. Захаров И.П., Сергієнко М.П. Метрологическая идентификация динамических характеристик средств измерительной техники. Харьков: Смит, 2012, 232 с.
113. Захаров И.П. Неопределенность измерений для чайников и ... начальников. Харьков: Смит, 2013, 36 с.
114. Захаров И.П. Неопределенность измерений для чайников и ... начальников. Издание 2-е перераб. и доп. С.-Петербург: Политехника-Сервис, 2014, 52 с.

115. Бочарова А.И., Водотыка С.В., Захаров И.П. Особенности оценивания неопределенности измерений в стандарте ISO 14253-2 // Системи обробки інформації, 2010, вип. 4 (85), с. 2 – 6.

116. Захаров И.П., Водотыка С.В., Климова Е.А., Шевченко Н.С. Некоторые случаи оценивания неопределенности измерений // Измерительная техника, 2013, №6, с. 12-15.

