

Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Факультет інформаційних радіотехнологій та технічного захисту інформації
(повна назва)

Кафедра медіаінженерії та інформаційних радіоелектронних систем
(повна назва)

КВАЛІФІКАЦІЙНА РОБОТА

Пояснювальна записка

рівень вищої освіти другий (магістерський)
(позначення документа)

Дослідження систем технічного зору в задачах розпізнавання образів

(тема)

Виконав:

студент 2 курсу, групи СТМм-21-1
Ярова Ю.Д.

(прізвище, ініціали)

Спеціальність 171 Електроніка

(код і повна назва спеціальності)

Тип програми освітньо-професійна
(освітньо-професійна або освітньо-наукова)

Освітня програма Системи, технології і комп'ютерні засоби мультимедіа

(повна назва освітньої програми)

Керівник проф. Колендовська М.М.

(посада, прізвище, ініціали)

Допускається до захисту

Зав. кафедри _____
(підпис)

Карташов В.М.
(прізвище, ініціали)

2021 р.

Харківський національний університет радіоелектроніки

Факультет Інформаційних радіотехнологій та технічного захисту інформації
Кафедра Медіаінженерії та інформаційних радіоелектронних систем

Рівень вищої освіти другий (магістерський)

Спеціальність 171 Електроніка
(код і повна назва)

Тип програми освітньо-професійна
(освітньо-професійна або освітньо-наукова)

Освітня програма "Системи, технології і комп'ютерні засоби мультимедіа"

ЗАТВЕРДЖУЮ:

Зав. кафедри _____
(підпис)

« ____ » _____ 20 ____ р.

ЗАВДАННЯ

НА КВАЛІФІКАЦІЙНУ РОБОТУ

Студентові Яровій Юлії Дмитрівні
(прізвище, ім'я, по батькові)

1. Тема роботи Дослідження систем технічного зору в задачах розпізнавання образів

затверджена наказом по університету від " 08 " 11 2021 р. № 1675

2. Термін подання студентом роботи 08.12.2021 р.

3. Вихідні дані до проекту (роботи) _____

1. Створити систему розпізнавання оптичного сигнал $A(x, y)$, який фокусується оптичною системою на матриці фотоелементів. Оптична система характеризується імпульсною характеристикою $H(x, y)$.

4. Перелік питань, що потрібно опрацювати в роботі

ВСТУП

1. Аналіз сучасних систем технічного зору

2. Розпізнавання образів за допомогою штучних нейронних мереж

3. Опис системи ідентифікації

4. Розробка системи розпізнавання

ВИСНОВКИ

ПЕРЕЛІК ПОСИЛАНЬ

ДОДАТКИ

5. Перелік графічного матеріалу із зазначенням обов'язкових креслеників, схем, плакатів, комп'ютерних ілюстрацій:

1. Визначення системи технічного зору

2. Постановка задачі

3. Принцип роботи систем технічного зору

4.	Класифікація систем розпізнавання
5.	Основні завдання при побудові системи розпізнавання
6.	Методи розпізнавання образів
7.	Статичні методи розпізнавання образів 1
8.	Синтаксичні методи розпізнавання образів
9.	Блок-схема системи синтаксичного розпізнавання образів
10.	Розпізнавання образів за допомогою штучних нейронних мереж
11.	Структура чотиришарової нейромережі
12.	Розробка системи розпізнавання
13.	Рішення завдання
14.	Колірна модель HSV
15.	Об'єкт та результат розпізнавання об'єкту за жовтим кольором
16.	Аналіз розробленого алгоритму
17.	Використання отриманої програми в промисловості
18.	Використання отриманої програми в сільському господарстві
19.	Висновки

КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів роботи	Термін виконання етапів роботи	Примітка
1.	Аналітичний огляд літературних джерел	08.11.21–15.11.20	
2.	Огляд сучасних систем технічного зору. Аналіз новітніх методів розпізнавання образів.	15.11.21–20.11.21	
3.	Огляд структурних методів розпізнавання образів. Дослідження розпізнавання образів за допомогою штучних нейронних мереж	21.11.21–27.11.21	
4.	Розробка системи розпізнавання. Реалізація алгоритму розпізнавання. Аналіз алгоритму.	28.11.21–02.12.21	
5.	Графічна частина роботи	25.11.21–02.12.21	
6.	Перевірка керівником	02.12.21–03.12.21	
7.	Перевірка на академічний плагіат	04.12.21	
8.	Перевірка завідувачем кафедри, рецензування	05.12.21–07.12.21	

Дата видачі завдання _____ 8.11.2021 р. _____

Студент _____
(підпис)

Керівник роботи _____
(підпис)

проф. Колендовська М.М.
(посада, прізвище, ініціали)

РЕФЕРАТ

Пояснювальна записка кваліфікаційної роботи має: 97 с., 26 рис., 4 табл., 2 додатка, 60 джерел.

СИСТЕМИ ТЕХНІЧНОГО ЗОРУ, СИСТЕМИ РОЗПІЗНАВАННЯ ОБРАЗІВ, АВТОНОМНЕ УПРАВЛІННЯ, НЕЙРОННІ МЕРЕЖІ, ШТУЧНИЙ ІНТЕЛЕКТ

Об'єкт дослідження – системи технічного зору.

Предмет дослідження – задачі розпізнавання образів мобільними роботами.

Мета кваліфікаційної роботи – розробити модель розпізнавання різноманітних об'єктів мобільними роботами.

Методи дослідження – теоретичний аналіз, числові розрахунки, математичне моделювання, статистична обробка даних, аналіз систем з використанням штучного інтелекту.

У даній роботі наведено класифікацію методів розпізнавання образів. Детальний аналіз систем технічного зору, перевірка статистичних гіпотез під час розпізнавання образів. Розглянуто розпізнавання образів за допомогою штучних нейронних мереж. Розроблено алгоритм і програмне забезпечення, в якому відповідно до певного кольору, камера знаходить відповідний об'єкт, розпізнає його і посилає сигнали в систему управління і переміщається навколо об'єкта; все йде в автономному режимі

ABSTRACT

The explanatory note of the qualification work has: 97 pages, 26 figures, 4 tables, 2 appendices, 60 sources.

TECHNICAL VISION SYSTEMS, IMAGE RECOGNITION SYSTEMS, AUTONOMOUS CONTROL, NEURAL NETWORKS, ARTIFICIAL INTELLIGENCE

The object of study - technical vision systems.

The subject of research - the tasks of pattern recognition by mobile robots.

The purpose of the qualification work is to develop a model of recognition of various objects by mobile robots.

Research methods - theoretical analysis, numerical calculations, mathematical modeling, statistical data processing, analysis of systems using artificial intelligence.

This paper presents a classification of image recognition methods. Detailed analysis of technical vision systems, testing of statistical hypotheses during pattern recognition. Image recognition using artificial neural networks is considered. An algorithm and software have been developed in which, according to a certain color, the camera finds the corresponding object, recognizes it and sends signals to the control system and moves around the object; everything goes offline

ПЕРЕЛІК УМОВНИХ СКОРОЧЕНЬ

CVV – Card Verification Value;
БД – база даних;
ОС – операційна система;
ПЗ – програмне забезпечення;
ПК – персональний комп'ютер
ПК – персональний комп'ютер;
РЕБ - радіоелектронна боротьба;
РЕЗ- радіоелектронні засоби;
РЕП - радіоелектронне подавлення;
С2С – Consumer-to-consumer;
СПО - спеціальне програмне забезпечення;
СТЗ – система технічного зору;
ТЗ - технічний засіб;
ЧТС - частотно тимчасовий сигнал;
ШІ – штучний інтелект;
ШНМ – штучна нейронна мережа;
ШПС - широкосмугові сигнали;
ШРД - широкосмуговий радіодоступ;
ШШС - шумоподібний широкосмуговий сигнал.

ЗМІСТ

Перелік умовних скорочень.....	6
Вступ.....	9
1 Аналіз існуючих сучасних систем технічного зору	11
1.1 Аналіз сучасних методів розпізнавання образів	11
1.2 Основні поняття.....	15
1.3 Ознаки образів	18
1.4 Аналіз методів розпізнавання образів	19
1.5 Статичні методи розпізнавання образів	21
1.6 Перевірка статистичних гіпотез під час розпізнавання образів. Ймовірність помилки під час перевірки гіпотез	23
1.7 Структурні методи у розпізнаванні образів	29
1.7.1 Лінійні класифікатори	31
1.8 Оцінювання ймовірності помилки	38
2 Аналіз методів розпізнавання образів за допомогою штучних нейронних мереж	41
2.1 Коннекціонізм.....	45
2.2 Універсальність навчаючих алгоритмів.....	47
2.3 Аналіз класифікацій базових нейроструктур. Типи навчання	48
2.3.1 Персептрони. Прототипи задач.....	49
2.3.2 Вибір функції активації	52
2.4 Багатошарові персептрони.....	55
2.4.1 Проблема функції заперечуючого АБО (XOR).....	56
2.5. Лінійна роздільність	59
2.6 Аналіз методів подолання обмеження лінійної роздільності	60

	8
3 Опис системи ідентифікації.....	65
3.1 Структурна схема системи розпізнавання відео	65
3.2 Опис оптичної системи	67
3.3 Опис матриці фотоелементів	70
3.4 Опис системи попередньої обробки	70
3.5 Опис комп'ютерної системи.....	75
4 Розробка системи розпізнавання	77
4.1 Реалізація алгоритму розпізнавання образів.....	78
4.2 Аналіз алгоритму розпізнавання образів	81
Висновки.....	85
Перелік посилань	88
Додатки.....	98
Додаток А	99
Додаток Б.....	109

ВСТУП

Розпізнавання образів (об'єктів, сигналів, ситуацій, явищ чи процесів) є чи не найпоширенішим завданням, яке доводиться вирішувати людині майже кожену секунду від першого до останнього дня свого існування.

Щоб вирішити цю проблему, люди використовують величезні ресурси свого мозку, включаючи близько 10 – 12 мільярдів нейронів одночасно. Це дозволяє миттєво впізнавати один одного, швидко читати друковані і рукописні тексти, безпомилково керувати автомобілем у складному потоці дорожнього руху, здійснювати відбраковування деталей, розгадувати коди.

У своєму повсякденному житті людина справляється з такими завданнями настільки легко, що це сприймається як належне. Тим часом, спроби імітувати ці високоінтелектуальні функції на комп'ютері стикалися з дуже серйозними труднощами.

В даний час найбільший прогрес був досягнутий в розпізнаванні друкованих символів. Широко використовуються відомі програми для розпізнавання текстової інформації – FineReader і CuneiForm.

Функції виявлення і розпізнавання військових об'єктів противника вже давно вбудовані в бортові комп'ютери літаків, ракет, кораблів і підводних човнів.

В даний час актуальною є задача створення технічного зору для інтелектуальних роботів, в ній відіграє чимале значення розпізнавання образів.

Мета кваліфікаційної роботи роботи – створити алгоритм і програмне забезпечення, в якому, відповідно до певного кольору, камера знаходить відповідний об'єкт, розпізнає його і посилає сигнали в систему управління і переміщається навколо об'єкта; все йде в автономному режимі

Складне завдання розпізнавання образів вимагає розробки гнучкої розпізнає системи, яка могла б класифікувати будь – який образ, наявний на зображенні. Цілями даної роботи є розробка такої системи з використанням

як давно відомих засобів, так і нових, нещодавно розроблених; втілити розроблену систему в програмному продукті. Результатом виконання роботи повинен стати прототип програмного засобу, що дозволяє проводити аналіз і тестування різних підходів до виконання поставленого завдання.

В цій роботі будуть розглянуті такі питання як огляд існуючих систем технічного зору, класифікація методів розпізнавання образів. Детальний аналіз систем технічного зору, перевірка статистичних гіпотез під час розпізнавання образів. Розглянуто розпізнавання образів за допомогою штучних нейронних мереж. Розроблено алгоритм і програмне забезпечення, в якому, відповідно до певного кольору, камера знаходить відповідний об'єкт, розпізнає його і посилає сигнали в систему управління і переміщається навколо об'єкта; все йде в автономному режимі.

Об'єкт дослідження – системи технічного зору.

Предмет дослідження – задачі розпізнавання образів мобільними роботами.

Мета кваліфікаційної роботи – розробити модель розпізнавання різноманітних об'єктів мобільними роботами.

Методи дослідження – теоретичний аналіз, числові розрахунки, математичне моделювання, статистична обробка даних, аналіз систем з використанням штучного інтелекту.

1 АНАЛІЗ ІСНУЮЧИХ СУЧАСНИХ СИСТЕМ ТЕХНІЧНОГО ЗОРУ

1.1 Аналіз сучасних методів розпізнавання образів

Система технічного зору СТЗ – це система яка забезпечує виявлення, автоматичний контроль та аналіз об'єктів по їх зображенню. СТЗ це сенсорні пристрої, які забезпечують зображення робочих сцен і об'єктів, їх перетворення, обробку та інтерпретацію за допомогою комп'ютера, а також передачу результатів на роботизований пристрій управління. Серед різних систем виявлення роботів СТЗ володіє найбільшою інформативністю, надаючи, за деякими оцінками, від 80% до 90% всієї інформації, необхідної для успішного функціонування роботів.

Принцип роботи СТЗ – СТЗ виявляє нерухомий або рухомий об'єкт, захоплює зображення цього об'єкта, перевіряє певні властивості цього зображення, а потім передає результати різних етапів контролю.

СТЗ в даний час є одним з головних засобів розвитку автоматичних систем управління рухом в умовах, коли обсяг апіорної інформації не достатній і для вирішення завдань керування необхідно аналізувати навколишнє середовище в режимі реального часу.

Наразі дуже стрімко йде розвиток систем автоматичного керування на базі СТЗ, оскільки цей напрямок є порівняно новим і має дуже великі можливості для подальшого використання.

Системи обробки зображень широко поширені сьогодні і все частіше використовуються в різних областях людської діяльності. Важливою особливістю сучасного підходу до створення систем обробки зображень є швидкий перехід від ідеї до математичного алгоритму, перевірка якого не вимагає значних витрат.

Основні завдання побудови систем розпізнавання.

Система розпізнавання – складна динамічна система, що складається у загальному випадку з колективу підготовлених фахівців та сукупності технічних засобів отримання та обробки інформації та призначена для вирішення завдань розпізнавання відповідних об'єктів та явищ на основі спеціально сконструйованих алгоритмів. Кожна система розпізнавання пристосована для розпізнавання лише цього виду об'єктів та явищ. Перелічимо основні завдання, що виникають при побудові та проектуванні системи розпізнавання [2].

Завдання 1. Завдання полягає у докладному та ретельному вивченні об'єктів, для розпізнавання яких призначена проектована система. Її мета – усвідомити особливості об'єктів, що вивчаються, і визначити, що ріднить і відрізняє їх один від одного.

Завдання 2. Це завдання полягає у проведенні класифікації об'єктів і явищ, що розпізнаються. Основне у цій задачі – вибір належного принципу класифікації. Цей принцип визначається вимогами до системи розпізнавання, які, у свою чергу, залежать від того, які рішення можуть прийматися за результатами розпізнавання системою невідомих об'єктів та явищ.

Завдання 3. Це завдання полягає у складанні словника ознак, використовуюваного як апріорного описи класів, так апостеріорного описи кожного невідомого об'єкта чи явища, що надходить на вхід системи і підлягає розпізнаванню.

При розробці словника ознак стикаються з низкою обмежень:

1) До словника можуть бути включені лише ті ознаки, щодо яких може бути отримана апріорна інформація, достатня для опису класів мовою цих ознак. Складений із цих ознак словник називатимемо апріорним. Інші ознаки, які або непотрібні, або малокорисні для поділу за класами, включати в словник недоцільно.

2) Наявність чи можливість створення технічних засобів спостережень, які забезпечують основі експериментів визначення попередньо відібраних ознак.

Ці обмеження часто перетворюють розробку словника ознак на складне завдання.

При проектуванні системи розпізнавання спочатку мовою ознак апріорного словника проводиться опис класів і після вибору алгоритмів розпізнавання оцінюється інформативність кожної ознаки. У результаті розгляду виключаються найменш корисні ознаки. Потім знову формується модель системи розпізнавання, і аналізуються якості частини ознак, що залишилися. Після цього, враховуючи обмеження, що накладаються створення технічних засобів отримання апостеріорної інформації, остаточно вирішується питання про створення робочого словника ознак системи розпізнавання.

Завдання 4. Завдання полягає в описі класів об'єктів мовою ознак. Вона не має однозначного рішення, і в залежності від обсягу апріорної інформації для її вирішення можуть бути використані методи безпосередньої обробки вихідних даних, навчання чи самонавчання. Розглянемо суть цього завдання.

Нехай у словнику міститься впорядкований набір параметрів об'єктів чи явищ – ознаки x_1, x_2, \dots, x_N . Величини x_1, x_2, \dots, x_N можна розглядати як складові вектора $x = \{x_1, x_2, \dots, x_N\}$, характеризує простір ознак.

Безліч векторів x утворює простір ознак розмірності N , а точки цього простору є об'єкти, що розпізнаються.

Нехай здійснено розбиття об'єктів на класи C_1, C_2, \dots, C_m . Потрібно виділити у просторі ознаки області $D_i, i=1, 2, \dots, m$, еквівалентні класам, тобто характеризуються наступною залежністю: якщо об'єкт характеризується набором ознак $x = \{x_1, x_2, \dots, x_N\}$ і належить до класу C_i , то точка, що представляє його в просторі ознак, належить області D_i .

Крім геометричної, існує і алгебраїчне трактування завдання, яке полягає в наступному. Потрібно побудувати розділяючі функції

$F_i(x_1, x_2, \dots, x_N), i=1, 2, \dots, m$, які мають наступну властивість: якщо об'єкт, що має ознаки $x^0 = \{x_1^0, x_2^0, \dots, x_N^0\}$, відноситься до класу C_i , то величина $F_i(x_1^0, x_2^0, \dots, x_N^0)$ має бути найбільшою. Вона має бути найбільшою й у всіх

інших значень ознак об'єктів, які стосуються класу C_i . Якщо через x_q позначити вектор ознак об'єкта, що відноситься до класу C_q , то для всіх значень вектора x_q

$$Fq(xq) > Fp(xp), q, p = 1, 2, \dots, m, q \neq p$$

Таким чином, у ознаковому просторі системи розпізнавання межа розбиття, яка називається вирішальною межею між областями D_i , виражається рівнянням

$$Fq(x) - Fp(x) = 0$$

Вироблення відомостей про об'єкти, що розпізнаються, і апіорний опис класів – дуже трудомістка частина у вирішенні класифікаційних завдань, що вимагає глибокого вивчення властивостей цих об'єктів.

Завдання 5. Завдання полягає в розробці алгоритмів розпізнавання, що забезпечують віднесення об'єктів, що розглядаються, до того чи іншого класу.

Алгоритми розпізнавання ґрунтуються на порівнянні тієї чи іншої міри близькості (подібності) об'єкта, що розпізнається, з кожним класом. При цьому, якщо обрана міра близькості L даного об'єкта ω з яким – небудь класом C_p , $p=1, 2, \dots, m$, перевищує міру його близькості з іншими класами, то приймається рішення про належність цього об'єкта до класу C_p , т.ч. $\omega \in C_p$, якщо

$$L(\omega, C_p) = \max \{L(\omega, C_i)\}, i = 1, 2, \dots, m$$

Завдання 6. Завдання полягає у розробці спеціальних алгоритмів управління роботою системи. Їхнє призначення полягає в тому, щоб процес функціонування системи розпізнавання був у певному сенсі оптимальним і

обраний критерій якості цього процесу досягав екстремального значення. Як подібний критерій може використовуватися, наприклад, вірогідність правильного розв'язання задачі розпізнавання, середній час її розв'язання, витрати, пов'язані з реалізацією процесу розпізнавання і т.д. Наприклад, мінімізація середнього часу вирішення завдання має здійснюватися за умови досягнення заданої ймовірності правильного розпізнавання.

Завдання 7. Завдання полягає у виборі показників ефективності системи розпізнавання та оцінці їх значень. Як показники ефективності можуть розглядатися ймовірність правильного розв'язання задачі розпізнавання, середній час, що витрачається на її вирішення, витрати, пов'язані з реалізацією процесу розпізнавання і т.д. Оцінка значень обраної сукупності показників ефективності, зазвичай, проводиться з урахуванням експериментальних досліджень чи реальної системи розпізнавання, чи з допомогою її фізичної чи математичної моделі.

1.2 Основні поняття

У розпізнаванні образів можна виділити 2 напрямки:

1) у вузькому сенсі – автоматичне розпізнавання таких образів, які можуть розпізнаватися за допомогою органів чуття. Цей напрямок будемо називати розпізнаванням на основі даних, що сприймаються органами чуттів (sense–data pattern recognition);

2) у широкому сенсі – автоматичне розпізнавання всіляких регулярностей, що зустрічаються, наприклад, в даних наукових досліджень; розпізнавання цих регулярностей досягається в результаті використання методів розпізнавання образів. При такій інтерпретації образу відповідають всі типи регулярностей (порядок, структура), що зустрічаються в складних даних. Цей напрямок розпізнавання будемо називати розпізнаванням на основі даних довільного характеру (general – data pattern recognition).

Під образом (об'єктом) в системі розпізнавання розуміється сукупність даних на вході системи. Дані можуть бути представлені різним чином: це може відеозображення, послідовність звуків, набір числових характеристик і т.д. кожен образ характеризується набором ознак – величин, на підставі яких система приймає рішення про приналежність об'єкта певному класу. Клас в даному випадку група образів, що володіють певною ознакою або значенням ознаки, який (яке) відрізняє дану групу образів від інших образів.

СТЗ реалізують складний процес виділення, розподілу, ідентифікації та перетворення відеоматеріалів та фотоматеріалів, що складається з шести основних етапів:

- 1) прийняття (сприйняття) інформації;
- 2) попередня обробка;
- 3) сегментація;
- 4) опис;
- 5) розпізнавання
- 6) інтерпретація.

Результатом визнання є класифікація конкретного об'єкту. Класифікація в такому випадку має на увазі призначення розглянутого об'єкта вірним і однозначним терміном, тобто в якості об'єкта порівняння одного з класів, відомих системі з цим об'єктом.

У процесі сортування людина не обов'язково повинна чітко визначати характеристики об'єкта, важливий тільки кінцевий результат процесу спостереження, сприйняття і розпізнавання. Автоматичні системи повинні виконувати ту ж класифікацію, що і людина, але повинні явно використовувати характеристики об'єкта. Класи, що виникають в результаті реалізації процесу розпізнавання, можуть бути дискретними (об'єкт або не є елементом класу), або розмитим (об'єкти присвоюються функціям приналежності до класу). У разі невизначених класів деякі об'єкти можуть бути охарактеризовані приналежністю до одного або декількох класів, значення яких може бути визначено в (0 та 1). Підходи, засновані на неясних

термінах, можуть бути корисними для розв'язання проблем розпізнавання, якщо класи незрозумілі. Вони також можуть відігравати роль посередника у розв'язанні проблем розпізнавання таких об'єктів, що характеризуються приналежністю до дискретних класів.

У більшості завдань розпізнавання класи є дискретними.

Тому бажано описати деякі особливості розпізнавання образів в порівнянні з таким випадком. Вхідні дані мають певний ступінь складності і генеруються з фактичного джерела. Вихід відносно простий – зводиться до визначення класу. Дуже важливо, отримання набору різних вхідних даних з різних джерел s_{i1}, s_{i2}, \dots , але з інформацією про одне і те ж зображення ω повинні бути відображені в один і той же клас C_i . Отже, розпізнавання образів являє собою однозначне відображення. Це означає, що всі вхідні дані, що надходять від s_{i1}, s_{i2}, \dots , і ті, що підлягають відображенню в один і той же клас, еквівалентні (хоча і різні) щодо відповідного образу. Мовою теорії множин це означає, що на множині $S_i = \{s_{i1}, s_{i2}, \dots\}$ має існувати деяке відношення еквівалентності « S_i ». Відношенням еквівалентності називається всяке відношення, що володіє властивостями рефлексивності, симетричності і транзитивності. Рефлексивність означає, що s_{ii} належить до того ж класу, що і s_{ii} . Симетричність вказує, що якщо s_{ii} належить до того ж класу, що і s_{ij} , то s_{ij} належить до того ж класу, що і s_{ii} . Транзитивність означає, що якщо s_{ii} належить до того ж класу, що і s_{ij} , а s_{ij} належить тому ж класу, що і s_{ik} , то і s_{ik} також належать до одного і того ж класу.

При розпізнаванні зображень ви можете ввести в систему механізм навчання, який зводиться до наступного. Розробник системи встановлює для кожного з образів i , які відрізняються один від одного задає приклади $s_{i1}, s_{i2}, \dots, s_{in}$, причому кожному образу ставиться у відповідність мітка, що вказує клас C_i , до якого він належить. Безліч таких прикладів називають навчальною множиною. У процесі навчання система має навчитися відображати вхідну інформацію у «правильні» класи. Навчальна множина

має бути репрезентативною щодо всіх можливих вхідних даних. Результат розпізнавання залежить також від розміру навчальної множини.

1.3 Ознаки образів

Двома суттєвими проблемами в розпізнаванні образів є наступні: які вхідні дані можна вважати доречними і яка попередня обробка вихідних даних (які зазвичай відрізняються надзвичайною надмірністю) призводить до отримання властивостей або ознак, що дійсно дозволяють проводити класифікацію. Так чи інакше, щодо ознак використовуються апріорні знання, інтуїція, метод спроб і помилок, досвід.

Залежно від специфіки завдання використовують безліч типів ознак. Деякі ознаки добре піддаються визначенню та легко інтерпретуються на об'єктах (наприклад, лінійні розміри). Більш складні ознаки засновані на формі, текстурі, статистичних зв'язках, розкладанні у ряд вихідних даних.

Коли ознаки тим чи іншим чином обрані, можна застосувати до них формальну схему, яка полягає у використанні наступних двох відображень :

1) $r_1: S_i \rightarrow F_i$, забезпечує отримання значень ознак $F_i = \{f_{i1}, f_{i2}, \dots\}$

2) $r_2: F_i \rightarrow C_i$ – процес класифікації, за допомогою якого вхідними даними ставляться у відповідність класи.

Основна проблема розпізнавання – мінливість образів. Вхідні дані, які повинні бути класифіковані як об'єкти одного класу, можуть відрізнитися досить сильно. Причин такої мінливості безліч.

1) мінливість, пов'язана з процесом вимірювання (шуми в датчиках);

2) мінливість, пов'язана з каналами зв'язку;

3) мінливість, властива власне образам (об'єкти одного класу можуть дуже сильно відрізнитися один від одного).

На даний час для автоматизації виробництва активно проводиться розробка СТЗ. Для підвищення ефективності функціонування СТЗ необхідно постійно поповнювати арсенал засобів попередньої обробки зображень та

побудова класифікаторів, які поєднують в собі задані показники швидкодії та перевіреної ідентифікації.

1.4 Аналіз методів розпізнавання образів

Існує цілий ряд методів прийняття рішень, придатних для віднесення об'єкта до одного або кількох класів за заданими ознаками та з урахуванням мінливості образів. Ці методи суттєво залежать від типів ознак та співвідношення власної та міжкласової мінливості. Можна виділити 2 великі сімейства способів прийняття рішення.

1) Методи статистичної теорії рішень

Ці методи застосовуються при розпізнаванні тих образів, ознаками яких є числові значення, причому ці значення у образів, що належать одному класу, різні. Апріорні відомості суттєві для вибору ознак; відомості про відхилення значень ознак об'єктів одного класу мають збиратися з використанням статистичного аналізу. Обидва ці різновиди відомостей можуть сприяти формуванню вирішальних функцій, що забезпечують класифікацію. Докладно цю групу методів розглянуто у роботі [3]

2) Структурні (лінгвістичні, синтаксичні) методи.

Ці методи використовуються при розпізнаванні тих образів, ознаками яких є непохідні елементи та їх відносини. Певна граматики визначає правила побудови образу з непохідних елементів. Виконання класифікації при заданих непохідних елементах та його відносинах вимагає синтаксичного аналізу [4].

Докладніше ці методи будуть розглянуті нижче.

Слід зазначити, що точний поділ статистичного та структурного підходів видається зручним лише з теоретичної точки зору. У практичних додатках їх відмінність менш велика. Завдання відбору та виділення ознак у статистичному підході та непохідних елементів у структурному за своєю природою схожі. Відмінність у тому, що непохідні елементи є підобрази, тоді

як ознаками може бути будь – яка безліч числових вимірів об'єкта. Тому ясно, що у практичних завданнях розпізнавання образів статистичний та структурний підходи часто доповнюють одне одного. Якщо явна структурна інформація про об'єкти не важлива, а завдання полягає скоріше тільки в класифікації, ніж в описі та класифікації, то немає потреби в структурних методах.

3) Метод базових точок.

Цей метод полягає у наступному. Кожен клас об'єктів характеризується набором точок, які певним чином розташовані у просторі. Рішення приймається на підставі аналізу співвідношень між координатами базових точок об'єкта, що розпізнається, і координатами базових точок еталонів.

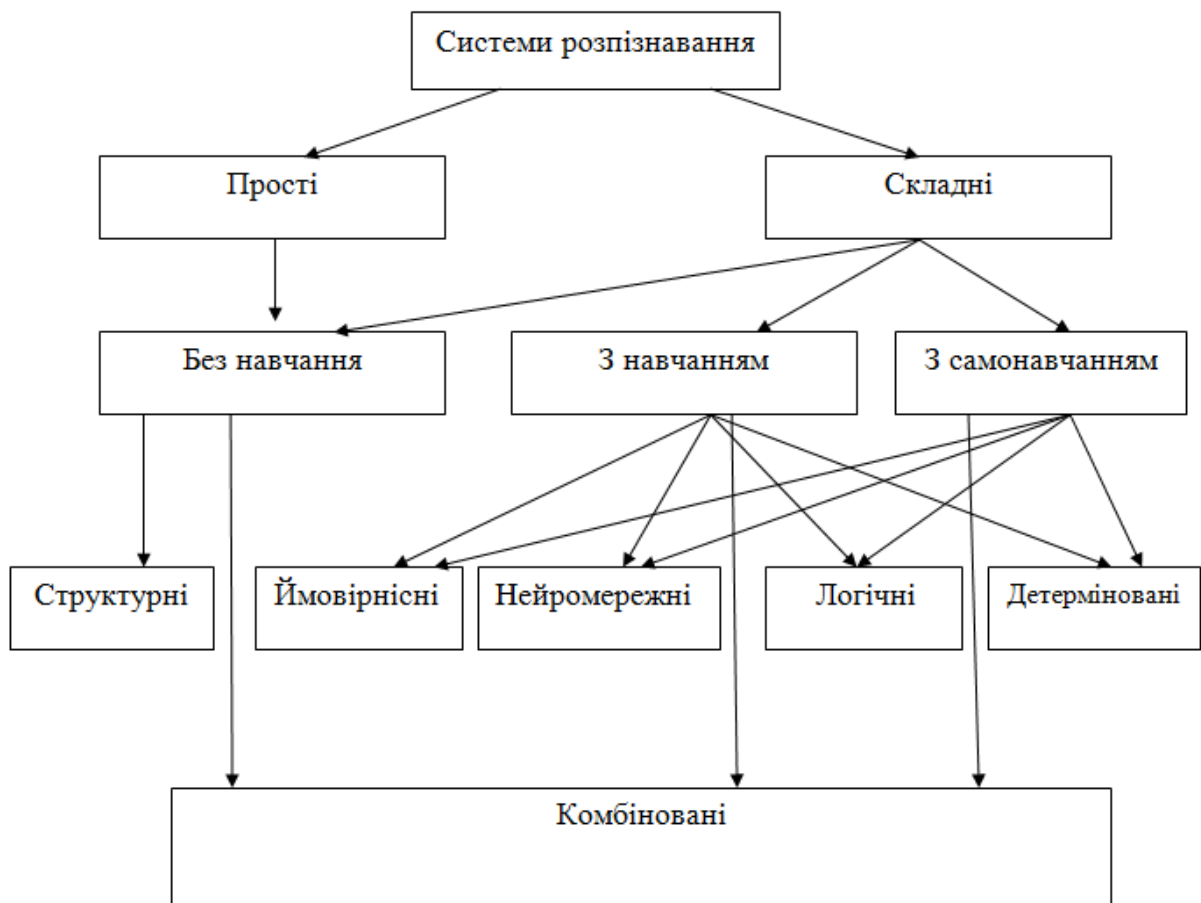


Рисунок 1.1– Класифікація систем розпізнавання

Системи розпізнавання можна розділити на прості і складні, в залежності від того, чи використовується фізично однорідна або фізично

різнорідна інформація для опису розпізнаваних об'єктів, чи мають ознаки, мовою яких описується алфавіт класів, єдину або різну фізичну природу.

Прості системи виявлення. До них відносяться, наприклад, автоматичні зчитувачі розпізнавання, де символи робочого словника представляють тільки певні лінійні розміри виявлених об'єктів; автомати для обміну монетами, де їх маса приймається за символ, який використовується при розпізнаванні монет.

Складні системи розпізнавання. До них відносяться, наприклад, системи медичної діагностики, в яких дані аналізу крові і кардіограми можуть використовуватися в якості ознак (симптомів); температура тіла і динаміка артеріального тиску і т. д.; системи виявлення зразків геологорозвідки, в яких різні фізичні і хімічні властивості використовуються в якості ознак.

1.5 Статичні методи розпізнавання образів

Вище вже було сказано, що статистичні методи розпізнавання образів застосовують у тих випадках, коли обрані ознаки характеризуються числовими значеннями, причому ці значення об'єктів одного класу можуть бути різними. Розглянемо тепер цю групу методів докладніше.

Позначимо через x вектор ознаки об'єкта. Кожну зі складових цього вектора вважатимемо значенням i ознаки, розподіленим відповідно до деякого закону. Отже вектор x буде випадковим вектором. Безліч таких векторів утворює n – вимірний простір образів, а кожна конкретна точка цього простору відповідає певному об'єкту. Класам образів при цьому відповідатимуть області цього простору .

Завдання розпізнавання образів зводиться, в такий спосіб, визначення меж областей, відповідних різним класам. На рисунку 1.2 наведено простий випадок двох розподілів. Якщо з минулого досвіду ці два розподіли вектора x відомі, можна встановити між ними кордон $g(x_1, x_2)$, яка поділяє двовимірний

простір на дві області. Таким чином, при розгляді нового образу в залежності від знака функції $g(x_1, x_2)$ можна вирішити, чи належить цей образ класу I чи класу II.

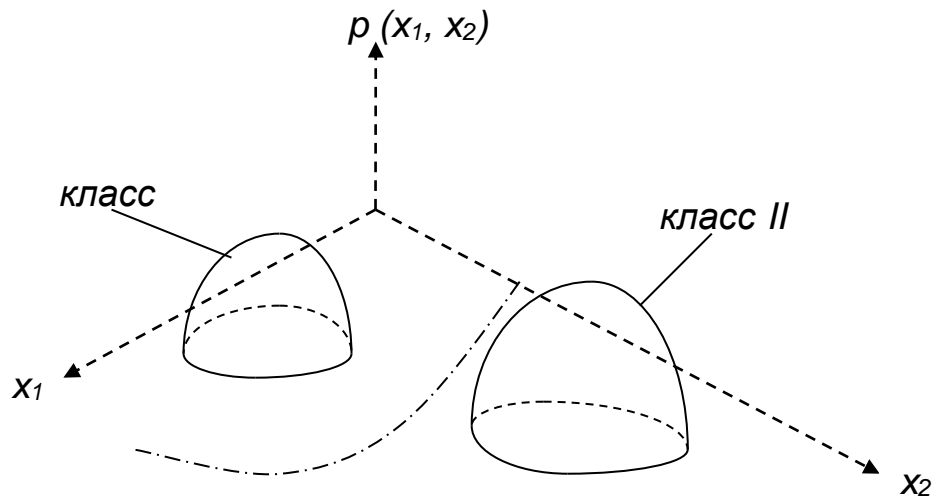


Рисунок 1.2 – Приклад простого випадку двох розподілів

Функцію $g(x_1, x_2)$ називають дискримінантною функцією, а технічний пристрій, що визначає знак $g(x_1, x_2)$ – блок розпізнавання образів або класифікатором. На рисунку 1.3 зображено блок - схему класифікатора в n – мірному просторі.

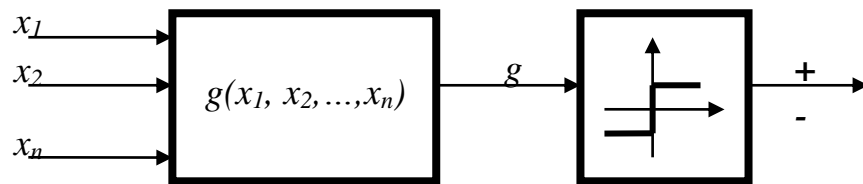


Рисунок 1.3 – Блок - схема класифікатора в n – мірному просторі

Щоб спроектувати класифікатор, потрібно вивчити характеристики розподілу вектора x кожного класу і визначити відповідну дискримінантну функцію. Вочевидь, що проектування класифікатора тим простіше, що менше ознак, використовуваних при розпізнаванні (розмірність простору ознак). Для того, щоб цього досягти, слід намітити деякі шляхи для вибору або вилучення істотних інформативних ознак із усієї сукупності

спостережуваних. Вибір ознак можна розглядати як відображення n -вимірного простору в простір меншої розмірності. У цьому необхідно зберегти властивість роздільності розподілів, відповідних різним класам.

Отже, розв'язання завдання розпізнавання образів методами статистичної теорії рішень і двох частин: вибору інформативних ознак і проектування класифікатора [3]. Насправді між цими частинами немає чіткої кордону. Дійсно, класифікатор можна представити як пристрій для вибору ознак, що відображає m ознак в одну (дискримінантна функція).

1.6 Перевірка статистичних гіпотез під час розпізнавання образів. Ймовірність помилки під час перевірки гіпотез

Припустимо, що вектор спостережень є випадковим вектором з умовною щільністю ймовірності, яка залежить від належності цього вектора певному класу. Якщо умовна щільність ймовірності відома кожному за класу, то завдання розпізнавання образів стає завданням перевірки статистичних гіпотез.

Розглянемо спочатку завдання перевірку гіпотез для двох альтернатив. Таке завдання виникає, якщо безліч класів, до яких може належати даний об'єкт, складається лише з двох класів C_1 та C_2 . Умовні щільності ймовірності та апіорні ймовірності вважатимемо відомими.

Практично всі критерії перевірки гіпотез засновані на порівнянні величини, яка називається ставленням правдоподібності, з пороговим значенням. Ставлення правдоподібності $l(x)$ визначається як відношення умовних густин ймовірності приналежності вектора ознак x класам C_1 та C_2 .

$$l(\vec{x}) = \frac{P(\vec{x}/C_1)}{P(\vec{x}/C_2)} \quad (1.1)$$

Порогове значення залежить від обраного вирішального правила. Розглянемо деякі найчастіше застосовувані правила ухвалення рішення.

1) Байєсовське правило прийняття рішення, що мінімізує функцію ризику.

Припустимо, що, приймаючи те чи інше рішення, ми зазнаємо певних втрат (штрафу). Величина цих втрат залежить від цього, якого класу належить об'єкт насправді. Можна ввести матрицю ризику $R = \|r_{ij}\|$, де r_{ij} – втрати у разі прийняття рішення на користь належності об'єкта класу j , тоді як насправді він належить класу i . Якщо записати вираз для середнього ризику і потім мінімізувати його, можна отримати, що значення порога дорівнює:

$$l(\vec{x}) = \frac{(r_{21} - r_{22})P(C_2)}{(r_{12} - r_{11})P(C_1)} \quad (1.2)$$

де r_{ij} – елементи матриці ризику. Тут $P(C_i)$ – апіорні можливості того, що об'єкт належить класу C_i .

2) Мінімаксне правило ухвалення рішення.

Байєсовський критерій, що мінімізує ризик, заснований на порівнянні відношення правдоподібності з пороговим значенням (1.2), яке є функцією апіорних ймовірностей $P(C_i)$, $i=1, 2$. Тому якщо апіорні ймовірності не змінюються, то байєсовське вирішальне правило завжди забезпечує мінімальний ризик. Однак якщо апіорні ймовірності змінюються або їхнє значення невідоме, то зафіксована величина порогу вже не забезпечує досяжний мінімум ризику. Мінімаксний критерій використовується для знаходження такої величини порога, при якій мінімізується максимум можливого ризику, навіть якщо апіорні ймовірності змінюються або невідомі.

Таким чином, при використанні мінімаксного критерію необхідно знайти значення $P^*(C_i)$, при яких досягається максимум середнього ризику, а потім підставити знайдені значення формулу (1.2).

3) Правило ухвалення рішення, засноване безпосередньо на ймовірностях.

Його можна записати так:

$$l(\vec{x}) \begin{cases} \geq \frac{P(C_2)}{P(C_1)} \\ < \frac{P(C_2)}{P(C_1)} \end{cases} \rightarrow \vec{x} \in \begin{cases} C_1 \\ C_2 \end{cases} \quad (1.3)$$

Рішення приймається на користь того класу, апіорна ймовірність приналежності якому більше. Вираз (1.3) називають байєсівським критерієм, що мінімізує помилку рішення.

Порівнюючи формули (1.2) і (1.3), видно, що це правило є окремим випадком байєсовського правила прийняття рішення, коли втрати пов'язані співвідношенням:

$$r_{12} - r_{22} = r_{21} - r_{11}.$$

Це так званий випадок симетричної функції штрафу, коли штрафом є можливість помилки, і критерій (1.3) її мінімізує.

Іноді замість відношення правдоподібності $l(x)$ зручно використовувати величину $\ln l(x)$. У цьому випадку вирішальне правило (1.3) набуде вигляду

$$-\ln p(\vec{x} / C_1) + \ln p(\vec{x} / C_2) \begin{cases} \geq \ln\{P(C_1)/P(C_2)\} \\ < \ln\{P(C_1)/P(C_2)\} \end{cases} \rightarrow \vec{x} \in \begin{cases} C_1 \\ C_2 \end{cases} \quad (1.4)$$

Ймовірність помилки під час перевірки гіпотез.

Зазвичай ніяке вирішальне правило не забезпечує безпомилкової класифікації. Тож оцінки якості правила прийняття рішення необхідно

обчислити ймовірність помилки, тобто. ймовірність того, що об'єкт помилково ставиться до цього класу.

Розрізняють помилки двох видів: помилки першого та другого роду. Помилка першого роду – помилка, коли ухвалено рішення про належність об'єкта класу C_1 , коли насправді він належить до класу C_2 . У свою чергу, помилка другого роду – помилка, коли приймається рішення на користь належності об'єкта класу C_2 , тоді як він належить до класу C_1 . Загальна ймовірність помилкового рішення визначається як виважена сума цих помилок:

$$\varepsilon = \varepsilon_1 P(C_1) + \varepsilon_2 P(C_2) \quad (1.5)$$

де ε_1 та ε_2 – ймовірності помилок першого та другого роду відповідно.

Знаходження ймовірності помилки зводиться сутнісно до обчислення n -мірного інтеграла то щільності ймовірності. Іноді зручніше інтегрувати щільність ймовірності відношення правдоподібності, яка є одновимірною. Інтеграл, які обчислюються в цьому випадку, мають вигляд

$$\varepsilon_1 = \int_0^{P(C)/p(C_1)} p\left(\frac{l}{C_1}\right) dl \quad (1.6)$$

$$\varepsilon_2 = \int_{P(C_2)/P(C_1)}^{+\infty} p\left(\frac{l}{C_2}\right) dl \quad (1.7)$$

Нижня межа інтегрування в (1.6) дорівнює нулю, оскільки ставлення правдоподібності завжди позитивне. На рис. 1.4 зображені щільності ймовірності вирішального правила $h(x)$.

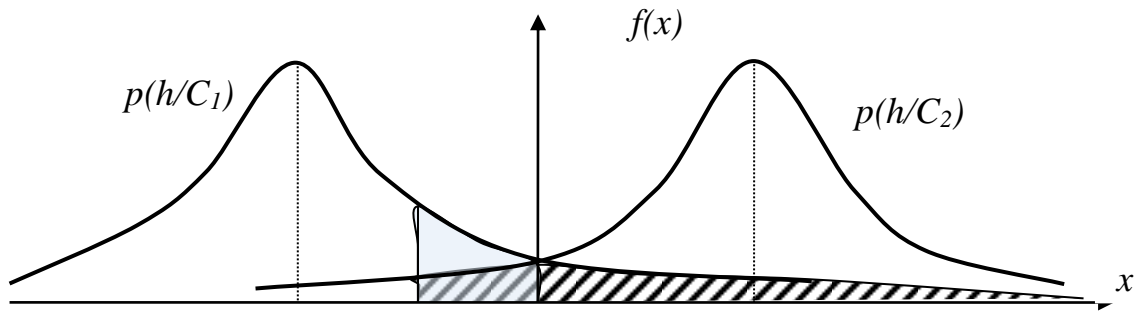


Рисунок 1.4 – Щільності ймовірності вирішального правила $h(x)$

Виділені площі відповідають ймовірностям помилки, зумовленим критерієм Байеса, який мінімізує помилку рішення.

Критерій Неймана – Пірсона оптимальності рішення.

Критерій Неймана – Пірсона використовується, коли невідома матриця втрат R та апіорні ймовірності $P(C_1)$ та $P(C_2)$. Вирішальне правило Неймана – Пірсона є вирішальним правилом, що мінімізує ймовірність помилки першого роду за умови, що ймовірність помилки другого роду є фіксована величина, наприклад ε_0 . Для визначення цього вирішального правила необхідно знайти мінімум виразу

$$r = \varepsilon_1 + \mu(\varepsilon_2 + \varepsilon_1) \quad (1.8)$$

де μ – множник Лагранжа, який є розв'язком рівняння

$$\varepsilon_2 = \int_{\Gamma_1} p\left(\frac{\vec{x}}{C_2}\right) d\vec{x} = \varepsilon_0 \quad (1.9)$$

Якщо використовувати густину ймовірності відношення правдоподібності, то рівняння для обчислення порогу μ має вигляд

$$\varepsilon_2 = \int_{\mu}^{+\infty} p\left(\frac{l}{C_2}\right) dl = \varepsilon_0 \quad (1.10)$$

Оскільки густина імовірності $p(l/C_2) \geq 0$, то ймовірність помилки ε_2 ,

визначається виразом (1.10), є монотонною функцією щодо μ . Інакше висловлюючись, коли поріг μ збільшується, ймовірність помилки ε_2 зменшується. Тому після обчислення значень ε_2 для кількох значень порога μ можна знайти таке μ , якому відповідає значення ε_2 , що дорівнює ε_0 . Однак одержати точне рішення рівняння (1.10) нелегко.

Можливість помилки для нормальних випадкових векторів.

Якщо розподіли є нормальними, то завжди можна визначити лінійне перетворення, що одночасно призводить до діагонального вигляду дві кореляційні матриці. Тому в перетвореній системі координат завжди буде виконано припущення незалежності випадкових векторів. З іншого боку, ймовірності помилки інваріантні щодо будь-якого перетворення, оскільки ставлення правдоподібності залежить від вибору системи координат.

Послідовний критерій Байєса.

Послідовний критерій можна також застосовувати для того, щоб скоротити кількість змінних, що спостерігаються, необхідних для прийняття рішення. Це доцільно у разі, коли кожен об'єкт окремі ознаки перевіряють послідовно. Однак на шляху застосування такого підходу трапляється багато труднощів [3].

1) Перемінні, що спостерігаються, взаємно корелювані і мають різні розподіли.

2) Значення порога має змінюватись при кожному спостереженні. На рис.5 зображено приклад класифікації двовимірного об'єкта на класи C_1 та C_2 . Якщо спостерігається випадковий вектор x_1 , то відношення правдоподібності можна встановити значення порогів рівними A_1 і B_1 . Однак якщо спостерігаються два випадкові вектори x_1 і x_2 , то A_2 має дорівнювати B_2 , так як не можна відкладати прийняття рішення до пред'явлення наступної змінної. Як правило, припускають, що в міру збільшення k значення порогів прагнуть одного й того самого значення.

3) Ймовірність помилки визначається умовними щільностями ймовірності $p(x/C_1)$ та $p(x/C_2)$ при спостереженні всього набору из n перемінних.

Послідовний критерій не зменшує ймовірність помилки, але скорочує кількість змінних, необхідних прийняття рішення. Тому така процедура може бути виправдана, якщо вартість спостереження за кожною змінною є суттєвим фактором.

1.7 Структурні методи у розпізнаванні образів

Структурний (синтаксичний) підхід було запропоновано у тому, щоб представляти ієрархічну інформацію, що у кожному образі, тобто. описувати образ за допомогою більш простих образів, а кожен підобраз знову описувати ще більш простими образами і т.д. [4]. Вочевидь, застосування цього підходу виправдано лише тому випадку, коли розпізнавати обрані найпростіші образи, звані непохідними елементами, легше, ніж самі образи. Мова, що забезпечує структурний опис образів у термінах непохідних елементів, іноді називають мовою опису образів. Правила композиції непохідних елементів зазвичай задають з допомогою так званої граматики мови описи образів. Процес розпізнавання здійснюється після ідентифікації непохідних елементів та складання опису об'єкта і полягає в граматичному розборі (синтаксичному аналізі) речення, що описує об'єкт.

Структурний підхід до розпізнавання образів дає можливість описувати велику кількість складних об'єктів шляхом використання небагатьох непохідних елементів та граматичних правил (правил підстановки). Граматичне правило може бути застосовано будь-яке число разів, тому виявляється можливим дуже компактно виразити основні структурні характеристики нескінченної множини речень. Різні відносини, визначені між подобрами, зазвичай можуть бути виражені логічними та (або) математичними операціями [4].

Система синтаксичного розпізнавання образів.

Система синтаксичного розпізнавання образів вважатиметься, що складається з трьох основних частин: блоку попередньої обробки, блоку

опису (уявлення) об'єкта та блоку синтаксичного аналізу. Зазвичай передбачається, що об'єкти на виході блоку передобробки представлені в "достатньо хорошій" якості. Потім кожен підданий обробці передобробці представляють у вигляді структури мовного типу. Цей процес представлення об'єкта складається, по-перше, із сегментації, і, по-друге, із виділення непохідних елементів (ознак). Іншими словами, кожен об'єкт після передобробки ділиться на частини та непохідні елементи на основі заздалегідь заданих синтаксичних операцій (або операцій композиції) і отримує своє уявлення через безліч непохідних елементів та певні синтаксичні операції. Рішення про те, чи є уявлення об'єкта синтаксично правильним (тобто належить він до класу об'єктів, що описуються цим синтаксисом або даною граматиною), приймається блоком синтаксичного аналізу. У процесі синтаксичного аналізу цей блок може давати повний синтаксичне опис об'єкта в термінах граматичних одиниць або дерева синтаксичного аналізу (якщо уявлення об'єкта синтаксично правильне). В іншому випадку об'єкт або виключають із розгляду, або аналізують на основі інших заданих граматики, які, можливо, описують інші можливі класи образів, що розглядаються [4].

Для того, щоб отримати граматику, що описує структурну інформацію класу образів, необхідно пристрій виведення граматики, що дозволяє відновити її по безлічі об'єктів у вигляді, подібному мовному. Функції цього пристрою аналогічні процесу навчання дискримінаційних системах розпізнавання образів. У результаті навчання на об'єктах, що розглядаються, формується структурний опис цього класу. Потім отриманий у вигляді граматики опис використовується для опису образів та синтаксичного аналізу.

Блок-схему системи синтаксичного розпізнавання образів зображено на рис.1.5.

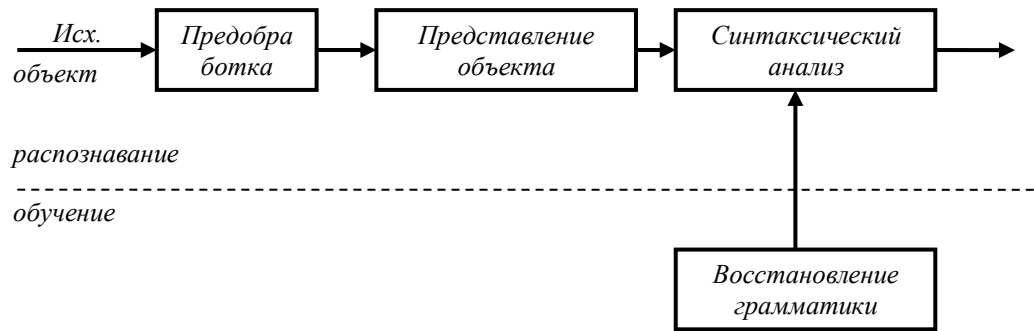


Рисунок 1.5 – Блок-схема системи синтаксичного розпізнавання образів

Хоча перевірка звичайного послідовного байєсівського критерію дуже складна, цей критерій використовують для послідовної перевірки ознак об'єкта. Байєсівський ризик можна представити наступним чином:

$$R = \sum_{i=1}^M p(C_i) \sum_{j=1}^n \int_{\Delta} \left[r_j + L(C_i, d_j(x_1, x_2, \dots, x_j)) \right] \times p\left(\frac{x_1, x_2, \dots, x_j}{C_c}\right) dx_1, dx_2, \dots, dx_j \quad (1.11)$$

де r_j – вартість спостережень x_1, x_2, \dots, x_j ; $L(C_i, d_j(x_1, x_2, \dots, x_j))$ – математичне очікування втрат для випадку, коли $(x_1, x_2, \dots, x_j) \in C_i$ використано вирішальне правило $d_j(x_1, x_2, \dots, x_j)$; Δ_i – область простору X , для якої послідовність випробувань закінчується при j -м спостереженні. У разі байєсівського критерію необхідно відшукати вирішальні правила $d_j(x_1, x_2, \dots, x_j)$, $j=1, 2, \dots, n$, мінімізуючі ризик R для заданої множини вартостей спостережень.

1.7.1 Лінійні класифікатори

Вище було сказано, що байєсівський класифікатор відношення правдоподібності оптимальний в тому сенсі, що він мінімізує ризик або ймовірність помилки. Однак для отримання відношення правдоподібності

необхідно розташовувати для кожного класу умовними щільностями ймовірності. У більшості додатків оцінка цих густин здійснюється за кінцевим числом вибірових векторів спостережень. Процедури оцінювання щільності ймовірності відомі, але вони є або дуже складними, або вимагають для отримання точних результатів великого числа векторів спостережень. У зв'язку з цим має сенс розглянути більш прості методи розробки класифікаторів. Проте, треба розуміти, що байєсівський Класифікатор у всіх випадках є найкращим, і ніякої лінійний Класифікатор не перевершує за якістю роботи Класифікатор, отриманий за критерієм відношення правдоподібності.

Найбільш простим і загальним видом є лінійний або кусочно–лінійний класифікатор.

Байєсівський лінійний класифікатор.

Для двох нормально розподілених випадкових величин байєсівське вирішальне правило можна представити у вигляді квадратичної функції щодо вектора спостережень \vec{x} наступним чином:

$$\frac{1}{2}(\vec{x} - M_1)^T \Sigma_1^{-1}(\vec{x} - M_1) - \frac{1}{2}(\vec{x} - M_2)^T \Sigma_2^{-1}(\vec{x} - M_2) + \frac{1}{2} \ln \frac{|\Sigma_1|}{|\Sigma_2|} \leq \ln \left\{ \frac{P(C_1)}{P(C_2)} \right\} \rightarrow \vec{x} \in \begin{cases} C_1 \\ C_2 \end{cases} \quad (1.12)$$

где Σ_i – кореляційні матриці випадкових величин, M_i – математичні очікування відповідно.

Якщо кореляційна матриця дорівнює одиничній, то можна вважати, що вектор \vec{x} являє собою спостереження, спотворене білим шумом. Компоненти вектора \vec{x} при цьому некорельовані і мають одиничну дисперсію, а байєсівське вирішальне правило набуває вигляду

$$(M_2 - M_1)^T \vec{x} + \frac{1}{2}(M_1^T M_1 - M_2^T M_2) \leq \ln \left\{ \frac{P(C_1)}{P(C_2)} \right\} \rightarrow \vec{x} \in \begin{cases} C_1 \\ C_2 \end{cases} \quad (1.13)$$

Добуток $M_i^T x$ являє собою коефіцієнт кореляції між векторами M_i і x . Легко бачити, що для прийняття рішення аналізований класифікатор порівнює різницю коефіцієнтів кореляції векторів x і M_1 і x і M_2 з вибраним порогом. Отже, його можна назвати кореляційним класифікатором [3].

Якщо помножити (1.13) на 2, а потім додати і відняти $x^T x$ з лівої частини, то можна отримати вирішальне правило

$$(\vec{x}^T \vec{x} - 2M_1^T \vec{x} + M_1^T M_1) - (\vec{x}^T \vec{x} - 2M_2^T \vec{x} + M_2^T M_2) \underset{>}{\overset{\leq}{>}} 2 \ln \left\{ \frac{P(C_1)}{P(C_2)} \right\} \rightarrow \vec{x} \in \begin{cases} C_1 \\ C_2 \end{cases} \quad (1.14)$$

Або

$$\|\vec{x} - M_1\|^2 - \|\vec{x} - M_2\|^2 \underset{>}{\overset{\leq}{>}} 2 \ln \left\{ \frac{P(C_1)}{P(C_2)} \right\} \rightarrow \vec{x} \in \begin{cases} C_1 \\ C_2 \end{cases} \quad (1.15)$$

Отриманому вирішальному правилу можна дати наступну інтерпретацію: порівнюються відстані між вектором x і векторами M_1 та M_2 з порогом.

У загальному випадку, коли кореляційні матриці не рівні одиничної, спостережуваний шум корельований, і його часто називають «забарвленим». У цьому випадку байєсівський Класифікатор так легко не інтерпретується. Однак все ще доцільно розглядати як вирішальне правило кореляційний Класифікатор або Класифікатор, заснований на обчисленні відстані. Для цього можна ввести декоррелююче перетворення $y = Ax$, яке переводить корельований («пофарбований») шум в білий:

$$A \Sigma A^T = I. \quad (1.16)$$

Зауважимо, що поки кореляційна матриця Σ є позитивно визначеною, матриця A існує і невироджена. Тому перетворення оборотне, і

спостереження вектора у можна класифікувати також ефективно, як і спостереження вектора x [3].

Лінійна розділяє функція, що мінімізує ймовірність помилки рішення.

Лінійні класифікатори являють собою найпростіші класифікатори, оскільки їх реалізація безпосередньо пов'язана з багатьма відомими методами класифікації, такими, як кореляційні методи або методи, засновані на обчисленні евклідових відстаней. Однак лінійні класифікатори оптимальні в байєсівському сенсі тільки для нормальних розподілів з рівними кореляційними матрицями. Для деяких додатків, таких, як виділення корисного сигналу в системах зв'язку, рівність кореляційних матриць є прийнятним припущенням, так як при зміні сигналу властивості шуму істотно не змінюються. Однак у багатьох інших додатках розпізнавання образів припущення про рівність коваріацій не виправдано.

Робилися різні спроби розробки лінійних класифікаторів для нормальних розподілів з нерівними кореляційними матрицями і для розподілів, відмінних від нормального. Зрозуміло, ці класифікатори не є оптимальними, однак у багатьох випадках їх простота служить достатньою компенсацією погіршення якості класифікації.

Розглянемо методи створення лінійних класифікаторів для цих більш складних випадків. Так як обумовлено, що незалежно то виду розподілів використовується лінійний Класифікатор, то вирішальне правило повинно мати вигляд

$$h(\vec{x}) = V^T \vec{x} + v_0 \begin{cases} \leq 0 \\ > 0 \end{cases} \rightarrow \vec{x} \in \begin{cases} C_1 \\ C_2 \end{cases} \quad (1.17)$$

Вираз $h(x)$ є лінійна розділяє функція щодо x . завдання синтезу класифікатора полягає в тому, щоб для заданих розподілів визначити коефіцієнти $V^T = [v_1, v_2, \dots, v_n]$ і значення порога v_0 , оптимальні за різними критеріями.

Якщо випадкова величина $h(x)$ розподілена по нормальному або

близькому до нього закону, то для обчислення ймовірності помилки можна використовувати її математичне очікування і дисперсію для класів C_1 та C_2 , а потім вибрати параметри V і v_0 так, щоб мінімізувати помилку рішення. Так як $h(x)$ є сумою, що складається з n доданків x_i , приходимо до висновку [3]:

1) якщо вектори x мають нормальний розподіл, то величина $h(x)$ також має нормальний розподіл;

2) навіть якщо вектори x розподілені не за нормальним законом, але при великому n виконані умови центральної граничної теореми, то розподіл величини $h(x)$ може бути близько до нормального.

Кусочно - лінійні розділяють функції.

Лінійні розділяють функції знаходять широке застосування в задачах розпізнавання образів для випадку двох класів, хоча при цьому неминуча деяка втрата якості розпізнавання. Однак при трьох або більше класах якість розпізнавання з використанням лінійної розділяє функції часто виявляється неприйнятним.

Якщо в цих випадках скористатися сукупністю лінійних розділяючих функцій, тобто скористатися кусочно-лінійною розділяючою функцією, то з'являються додаткові можливості поліпшення якості розпізнавання [3]. На рис. 1.6 зображено приклад завдання розпізнавання для випадку чотирьох класів.

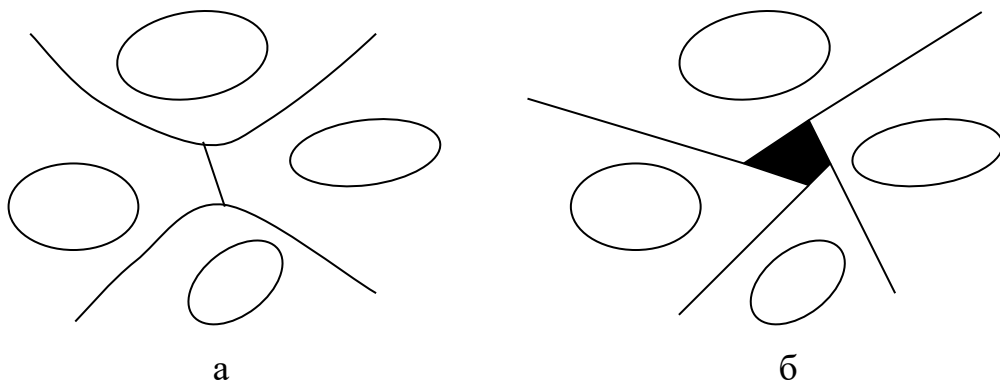


Рисунок 1.6 – Заміна байєсівських кордонів кусочно-лінійними кордонами

У завданнях з багатьма класами критерії перевірки багатьох гіпотез дають найкраще в байєсівському сенсі вирішальне правило, що забезпечує мінімум ризику або ймовірності помилки. Відповідно до критерію перевірки гіпотез щільність ймовірності або її логарифм слід порівнювати з густинами ймовірності інших класів, як випливає з (1.11). На рис. 1.6, а зображені отримані таким чином кордону. Якщо оцінювання щільності ймовірності є занадто складним завданням або межі, що визначаються відповідно до критерію перевірки гіпотез, занадто "химерні", то можна замінити ці складні межі безліччю простих лінійних кордонів. Така заміна, звичайно, призводить до деякого погіршення якості розпізнавання. Однак використання лінійних меж особливо ефективно для задач розпізнавання багатьох класів, тобто в тих випадках, коли складність меж швидко зростає зі збільшенням числа класів, і є бажаним деяке спрощення процедури синтезу класифікатора [3]. На рис. 1.6, б показана заміна байєсівських кордонів кусочно-лінійними кордонами.

Множина лінійних функцій, відповідна кусочно-лінійної розділяє функції, визначається наступним чином:

$$h_{ij}(x) = V_{ij}^T x + v_{i0}, i, j = 1, 2, \dots, M, i \neq j \quad (1.18)$$

де M – число класів. Знаки V_{ij} вибираються так, щоб розподіл класу i знаходилося в області позитивних значень лінійної функції $h_{ij}(x)$, а розподіл класу j – в області негативних значень. З цієї вимоги випливає, що

$$h_{ij}(x) = h_{ji}(x) \quad (1.19)$$

Припустимо, що для кожного класу відповідна область є опуклою, як зображено на рис. 1.6, б, тоді область класу i може бути просто визначена наступним чином:

$$h_{i1}(x) > 0, \dots, h_{iM}(x) > 0 \rightarrow x \in C_1 \quad (1.20)$$

Наявність темної області на рис. 1.6, б показує, що M областей, які визначаються умовами, не обов'язково покривають весь простір. Якщо об'єкт потрапляє в цю область, то кусочно-лінійний класифікатор не може прийняти рішення про його приналежність до певного класу; цю область називають областю відмов. Вирішальне правило складається з $M-1$ лінійних розділяючих функцій і логічного елемента AND з $M-1$ входами, які приймають значення $sign\{h_{ij}(x)\}$. Відповідна блок-схема зображена на рис. 1.7.

Якщо припущення про опуклість областей не виконується, то необхідно визначити взаємні перетину $M-1$ гіперплощин і будувати вирішальне правило відповідно до цих перетинами. Однак при цьому класифікатор стає досить складним для вирішення практичних завдань.

Імовірність помилки рішення для кожного класу ε_i можна виразити через $(M-1)$ -мірну функцію розподілу ймовірності:

$$\begin{aligned} \varepsilon_i &= 1 - Pr \left\{ h_{i1}(\vec{x}) > 0, \dots, h_{iM}(\vec{x}) > \frac{0}{\vec{x}} \in C_1 \right\} = \\ &= 1 - \int_0^\infty \dots \int_0^\infty p(h_{i1}, h_{i2}, \dots, h_{iM}/C_i) dh_{i1} dh_{i2} \dots dh_{iM} \quad (1.21) \end{aligned}$$

(функція h_{ii} виключена з розгляду).

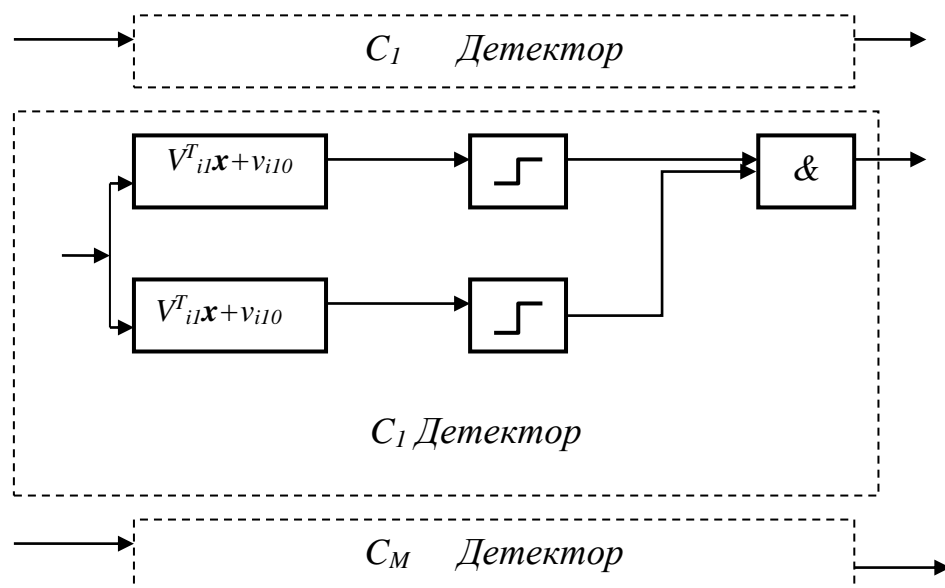


Рисунок 1.7– Двошарова машина

Оскільки кожна з паралельних ланцюгів складається з двох послідовно з'єднаних елементів, то такий кусочно-лінійний класифікатор іноді називають двошаровою машиною.

1.8 Оцінювання ймовірності помилки

Ймовірність помилки є основним показником якості розпізнавання образів, і тому її оцінювання є дуже важливим завданням. Ймовірність помилки є складна функція, що представляє собою n -кратний інтеграл від щільності ймовірності при наявності складної межі. Тому при її обчисленні часто доводиться звертатися до експериментальних методів.

При оцінці ймовірності помилки розглядають два завдання. Перша з них полягає в оцінюванні ймовірності помилки за наявною вибіркою в припущенні, що заданий класифікатор [3].

Друге завдання полягає в оцінці ймовірності помилки при заданих розподілах. Для цієї помилки ймовірність помилки залежить як від використовуваного класифікатора, так і від виду розподілу. Оскільки в розпорядженні є кінцеве число об'єктів, не можна побудувати оптимальний Класифікатор. Тому параметри такого класифікатора являють собою випадкові величини [3].

Оцінка ймовірності помилки для заданого класифікатора.

1) невідомі апріорні ймовірності – випадкова вибірка.

Припустимо, що задані розподілу обох класів і Класифікатор. Завдання полягає в оцінюванні ймовірності помилки за N об'єктами, отриманими відповідно до цих розподілів.

Коли невідомі апріорні ймовірності $P(C_i)$, $i=1, 2$, то можна випадково витягти N об'єктів і перевірити, чи дає даний Класифікатор правильні рішення для цих об'єктів. Такі об'єкти називають випадковою вибіркою.

Нехай τ – число об'єктів, неправильно класифікованих в результаті цього експерименту. Величина τ є дискретна випадкова величина. Позначимо

справжню ймовірність помилки через ε . Розподіл значень величини τ є біноміальним:

$$P(\tau) = C_N^\tau \varepsilon^\tau (1 - \varepsilon)^{N-\tau} \quad (1.22)$$

Оцінка максимальної правдоподібності $\hat{\varepsilon}$ дорівнює:

$$\hat{\varepsilon} = \tau/N \quad (1.23)$$

тобто оцінка максимальної правдоподібності дорівнює відношенню числа неправильно класифікованих об'єктів до загального числа об'єктів.

Математичне очікування і дисперсія біноміального розподілу відповідно рівні:

$$M(\hat{\varepsilon}) = \varepsilon \quad (1.24)$$

$$D(\hat{\varepsilon}) = \frac{\varepsilon(1-\varepsilon)}{N} \quad (1.25)$$

Таким чином, оцінка $\hat{\varepsilon}$ є незміщеною.

Оцінка ймовірності помилки, коли класифікатор заздалегідь не заданий.

Коли дані N об'єктів у разі відсутності класифікатора, то можна використовувати ці об'єкти як для проектування класифікатора, так і для перевірки його якості. Очевидно, оцінювана ймовірність помилки залежить від даних розподілів і використовуваного класифікатора.

Припустимо, що завжди використовується байєсівський класифікатор, що мінімізує ймовірність помилки. Тоді мінімальну ймовірність помилки байєсівського класифікатора, яку необхідно оцінити, можна розглядати як фіксований параметр при заданих розподілах. Крім того, ця ймовірність є мінімальною для даних розподілів.

Як правило, ймовірність помилки є функція двох аргументів:

$$\varepsilon(\theta_1, \theta_2) \quad (1.26)$$

де θ_1 – безліч параметрів розподілів, що використовуються для синтезу байєсівського класифікатора, а θ_2 , це – безліч параметрів розподілів, що використовуються для перевірки його якості.

Метод розбиття вибірки.

Для того, щоб розбити наявні об'єкти на навчальну та екзаменаційну вибірки, вивчимо, як це розбиття впливає на дисперсію оцінки ймовірності помилки.

Спочатку припустимо, що є нескінченне число об'єктів для синтезу класифікатора і N об'єктів для перевірки його якості. При нескінченному числі об'єктів синтезований класифікатор є класифікатором для істинних розподілів, і його внесок у дисперсію дорівнює нулю. Для фіксованого класифікатора організуємо селективну вибірку. У цьому випадку розподіл оцінки підпорядковується біноміальному закону з дисперсією:

$$D(\varepsilon) = \frac{1}{N} \sum_{i=1}^2 P(C_i) \varepsilon_i (1 - \varepsilon_i) \quad (1.27)$$

де ε_i – справжня ймовірність помилки для i -го класу

З іншого боку, якщо є N об'єктів для синтезу класифікатора і нескінченне число екзаменаційних об'єктів, то оцінка ймовірності помилки виражається наступним чином:

$$\hat{\varepsilon} = P(C_1) \int_{\Gamma_2} p(\vec{x}/C_1) d\vec{x} + P(C_2) \int_{\Gamma_1} p(\vec{x}/C_2) d\vec{x} \quad (1.28)$$

де Γ_i – область простору ознак, відповідна i -му класу.

У цьому випадку підінтегральні вирази постійні, але межа цих областей змінюється залежно від вибірки N об'єктів.

2 АНАЛІЗ МЕТОДІВ РОЗПІЗНАВАННЯ ОБРАЗІВ ЗА ДОПОМОГОЮ ШТУЧНИХ НЕЙРОННИХ МЕРЕЖ

Штучні нейронні мережі – це математична модель функціонування традиційних для живих організмів нейромереж, які представляють собою мережі нервових клітин. Як і в нервових клітинах, в штучних мережах основним елементом виступають нейрони, з'єднані між собою і утворюють шари, число яких може бути різним в залежності від складності нейромережі та її призначення (вирішуваних завдань).

Мабуть, найпопулярніша завдання нейромереж – розпізнавання візуальних образів. Сьогодні створюються мережі, в яких машини здатні успішно розпізнавати символи на папері і банківських картах, підписи на документах, детектувати об'єкти і т. д. Ці функції дозволяють значно полегшити роботу людини і підвищити надійність і точність різних робочих процесів, оскільки немає можливості здійснювати людські помилки.

Нейромережа – це математична модель у вигляді програмного і апаратного втілення, що будується на принципах функціонування біологічних нейромереж. Сьогодні такі мережі активно використовуються в практичних цілях, оскільки вони дозволяють не тільки розвиватися, а й навчатися. Вони використовуються для прогнозування, розпізнавання образів, машинного перекладу, розпізнавання звуку тощо. Перелічимо типові задачі для нейромереж:

- Апроксимація функцій за набором точок (регресія);
- Класифікація даних за заданим набором класів;
- Кластеризація даних з виявленням наперед невідомих класів-прототипів;
- Стиснення інформації;
- Відновлення (відтворення) втрачених даних;
- Асоціативна пам'ять;

– Оптимізація, оптимальне управління (керування).

Цей список можна продовжувати нескінченно. Зверніть увагу, що між цими відповідними рішеннями існує глибоке споріднене рішення. За ними стоїть єдиний прототип, який в якійсь мірі дає можливість звести ці завдання до одного.

Наприклад, візьмемо задачу апроксимації функції за відомим набором точок. Це типовий приклад неправильного завдання (завдання, яке не є єдиним рішенням). Щоб отримати єдине рішення, такі завдання необхідно регуляризувати – доповнити вимогою мінімізації деякого регуляризуючого функціоналу. Мінімізація такої функціональності є метою вивчення нейронної мережі. Завдання оптимізації також зводяться до мінімізації функцій цілі (цільових функцій) при заданому наборі обмежень. З другої сторони, класифікація – це не що інше, як апроксимація функції з дискретними значеннями. Завдання відновлення втрачених даних, в свою чергу, – це асоціативна пам'ять, що відтворює прообраз за його частиною. Такими зразками для наслідування в задачі кластеризації є центри кластерів. Адже якщо інформація може бути відтворена в якійсь її частині, то ми маємо ущільнення цієї інформації.

Представники різних наук займаються перерахованими вище завданнями і часто вважають, що нейромережі – це переклад відомих завдань на нову "мову", який нічого принципового не дає. Статистики стверджують, що нейромережі – це всього лише приватний спосіб обробки статистичних даних, а фахівці з оптимізації – методи навчання нейромереж давно відомі в їх області, теорія апроксимації функцій розглядає нейромережі поряд з іншими методами багатовимірної апроксимації.

Однак синтез різноманітних методів та ідей в єдиному нейромережевому підході є неоціненною перевагою нейросистем. Крім того, основною перевагою алгоритмів нейронної мережі для вирішення вищезазначених проблем є їх паралелізм.

У кожній галузі за детального розгляду можна знайти задачі для

нейромереж. Сфери застосування, де розв'язок таких задач має практичне значення, приведені нижче.

– Медицина: обробка медичних зображень, спостереження за пацієнтом. Діагностика, факторний аналіз ефективності лікування, очищення показань приладів від шуму. Наприклад, група “НейроКомп” з Красноярська розробила систему ранньої діагностики меланоми судинної оболонки ока. Цей вид раку становить 90% усіх внутрішньоочних пухлин ока і добре діагностується лише на пізніх стадіях хвороби. Метод заснований на непрямому вимірі вмісту меланіну в віях. Отримані дані спектрометрії і загальні характеристики піддослідного (стать, вік та ін.) подаються на вхідні синапси 43-нейронного класифікатора. Нейронна мережа вирішує, чи є у пацієнта пухлина, і якщо вона є, вона визначає ймовірність впевненості у відсотках.

– Економіка і бізнес: прогнозування ринків, оцінка ризиків дефолту за кредитами, прогнозування банкрутств, оцінка вартості нерухомості, оптимізація товарних і грошових потоків, автоматичне зчитування чеків і форм, безпека транзакцій на пластикових картах. Наприклад, програмне забезпечення компанії RETEK, дочірня фірма NHC Software, створила програмний продукт (січень 1998р.) Retek Predictive Enterprise Solution, це включає в себе нейронний блок для аналізу великих потоків даних, властивих великій різниці в торгівлі. Він також включає в себе блок прогнозів для оцінки конкретних рішень.

– Зв'язок: стиснення відео, швидке кодування і декодування, оптимізація мобільних мереж і схеми маршрутизації пакетів. Наприклад: нейромережі вже продемонстрували коефіцієнт стиснення 120:1 для чорно-білого відео. За допомогою спеціальної схеми колірного кодування ви можете отримати стиснення приблизно в два рази більше для кольорових відео.

– Інтернет: асоціативний пошук інформації, Електронні секретарі та агенти користувача в мережі, інформаційні фільтри, адресний маркетинг

для електронної комерції. Приклад: відповідно до прес-релізу, фірма Autonomy (фірма, котра відокремилась від фірми Nejdynamics) виробляє сімейство продуктів AGENTWARE, використання профілів інтересів користувачів у вигляді особистих автономних нейро-агентів.

– Авіоніка: самонавчальний автопілот, виявлення радіолокаційних сигналів, адаптивне пілотування серйозно пошкодженого літака. Наприклад: компанія McDonnell Douglas Electronic Systems розробляє автоматичний перемикач режиму польоту в режимі реального часу в залежності від типу пошкодження повітряного судна. Дані від 20-ти сенсорних датчиків і сигналів від пілота використовуються нейромережею для отримання понад 100 аеро-динамічних параметрів польоту. Сильна сторона мережі полягає в здатності адаптуватися до несподіваних 10 аеродинамічних режимів, таких як втрата частини крила і т.і.

– Автоматизація виробництва: оптимізація режимів виробництва, комплексна діагностика якості продукції (ультразвук, оптика, гамма-випромінювання), моніторинг та візуалізація багатовимірної інформації диспетчера, аварійне попередження, робототехніка. Приклад: Ford Motors Company запровадила в себе нейросистему для діагностики двигунів після невдалої спроби побудувати експертну систему, оскільки досвідчений механік може діагностувати помилки, але він не в змозі описати такий алгоритм виявлення. На вхід нейросистеми подають дані від 31-го сенсора. Нейромережею проаналізовано різноманітні види несправностей (868 прикладів). За опублікованими даними, після повного циклу навчання якість діагностики досягла рівня 11 кращих експертів, значно випереджаючи їх у швидкості.

– Політичні технології: аналіз та узагальнення соціологічних опитувань, прогнозування динаміки рейтингів, виявлення значущих факторів, об'єктивна кластеризація електорату.

– Безпека та охоронні системи: система ідентифікації, розпізнавання голосу та обличчя в натовпі, розпізнавання номерів,

аерокосмічний аналіз зображень, виявлення підробок. Введення і обробка інформації: обробка рукописних чеків, розпізнавання підписів, відбитків пальців і голосу. Введення у систему фінансових і податкових документів. Приклад: розроблені італійською фірмою RES Informatica нейромережеві пакети FlexRead використовуються для розпізнавання та автоматичного введення рукописних платіжних документів і податкових декларацій.

– Геологорозвідка: аналіз сейсмічних даних, асоціативні методи пошуку корисних копалин, оцінка ресурсів родовищ. Приклад: компанія "Амосо" використовує нейронні мережі для ізоляції характерних максимумів від сейсмічних показань. Надійність виявлення максимумів 95% по кожній сейсмолінії.

– Виявлення нових композиційних матеріалів і сплавів на основі попередніх оцінок зовнішнього збурення або спектральних властивостей з використанням асоціативної пам'яті, заснованої на нейромережі Гопфільда.

Приклади показують, що нейронні мережі є новим інструментом аналізу даних. Основними проблемами на шляху значного поширення нейронних технологій є нездатність широкого кола фахівців сформулювати свої проблеми в термінах, що пропонують просте нейронне рішення проблеми. Отже, спробуємо намітити основні особливості нейронної обробки інформації парадигми – нейрокомп'ютерів.

2.1 Коннекціонізм

Невід'ємною рисою нейронних мереж є глобальність зв'язків. Основними елементами штучних нейронних мереж є формальні нейрони, призначені для роботи з широкосмуговою інформацією. Кожен нейрон мережі зазвичай можна побачити з усіма нейронами попереднього шару (або взагалі з усіма нейронами), як показано на рисунку 2.1, де найбільш поширена архітектура в сучасних прикладних нейронних мережах багатозарового перцептрона проілюстрована. Це є відмінною рисою

формальних нейронів від базових елементів ЕОМ – логічних вентилів, котрі мають лише два входи та один вихід

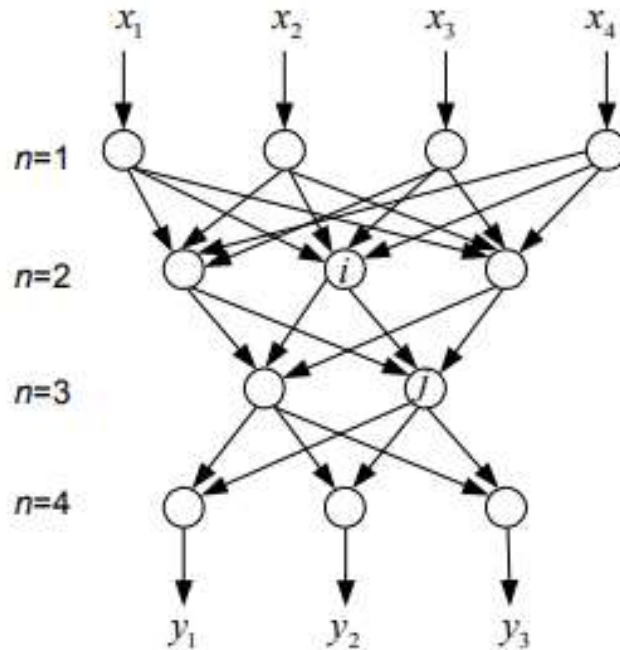


Рисунок 2.1 – Структура чотиришарової нейромережі

В результаті універсальні процесори мають складну архітектуру, засновану на ієрархії модулів, кожен з яких виконує зумовлену функцію. На відміну від них, Архітектура нейронних мереж проста і універсальна. Спеціалізація відносин виникає на етапі їх навчання під впливом конкретних даних.

Типовий формальний нейрон виконує просту операцію – зважає значення своїх входів зі своїми ж локальними вагами і здійснює щодо їх суми нелінійне перетворення:

$$y = f(u), u = w_0 + \sum w_i i$$

Нелінійність вихідної функції активації $f(u)$ принципова. Якби нейрони були лінійними елементами, довільна послідовність нейронів також виконувала б лінійне перетворення, щоб нейронна мережа в цілому відповідала одному нейрону (або шару нейронів, у випадку декількох

виходів). Нелінійність руйнує лінійну суперпозицію і призводить до того, що можливості нейронної мережі значно вище, ніж у окремих нейронів.

2.2 Універсальність навчаючих алгоритмів

Привабливою особливістю нейрокомп'ютерів є єдиний принцип навчання нейронних мереж – мінімізація емпіричних помилок. Функція помилки, яка оцінює цю конфігурацію мережі, встановлюється ззовні залежно від мети навчання. Потім мережа починає поступово змінювати свою конфігурацію-стан своїх синаптичних з'єднань, щоб звести до мінімуму цю помилку. Таким чином, в процесі навчання мережа буде все краще і краще справлятися з поставленим перед нею завданням. Образно цей процес можна представити як пошук мінімуму функції похибки $E(w)$, котра залежить від набору всіх синаптичних ваг мережі w (рис. 2.2).

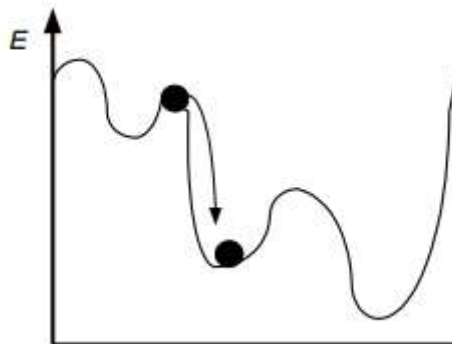


Рисунок 2.2 – Задача навчання нейромереж, як задача оптимізації

Базовою ідеєю усіх алгоритмів навчання є врахування локального градієнта в просторі конфігурацій для вибору траєкторії найшвидшого спуску по функції похибки. Функція похибки може мати множину локальних мінімумів, отож градієнтні методи, зазвичай, доповнюються елементами стохастичної оптимізації, щоб запобігти "застряганню" конфігурації мережі в таких локальних мінімумах. Ідеальний метод повинен знайти глобальний мінімум, котрий відповідає оптимальній конфігурації мережі.

2.3 Аналіз класифікацій базових нейроструктур. Типи навчання

Помилка мережі залежить від її конфігурації-сукупності всіх її синаптичних відносин. Однак цей зв'язок не є прямим, а опосередкованим, оскільки конкретні значення ваг приховані від зовнішнього спостерігача. Для нього мережа – це свого роду чорний ящик, і свою роботу він може оцінити тільки за своєю поведінкою – реакції на задані вхідні сигнали. Функція похибки набуває вигляду $E(w) = E\{x^a, y^a, y(x^a, w)\}$, де $\{x^a, y^a\}$ – набір прикладів (пар вхід-вихід), на яких навчається нейромережа, а $\{y(x^a, w)\}$ – реальні значення виходів нейромережі, що залежать від конкретних значень синаптичних ваг. Такий спосіб (тип) навчання, коли реальний вихід нейромережі порівнюють з еталонним, називають навчанням з учителем.

Іноді вихідна інформація відома не цілком. Наприклад, замість еталонних відгуків відомо лише краще чи гірше, така конфігурація мережі вирішує поставлену задачу (як у дитячій грі “тепло –холодно”). Цей тип навчання називають навчання з підтвердженням (reinforcement learning).

Можливий і такий режим навчання, коли бажані значення виходів загалом невідомі і мережа навчається лише на наборі вхідних даних $\{x^a\} : E(w) = E\{x^a, y(x^a, w)\}$. Такий режим навчання називають навчанням без вчителя. У цьому випадку самій мережі буде запропоновано знайти приховані шаблони в масиві даних. Як правило, надлишок даних передбачає стиснення інформації-мережу можна навчити знаходити найбільш компактне представлення даних, іншими словами-реалізувати оптимальне кодування такого типу вхідної інформації.

Великий вплив на те, як обробляється інформація, надає наявність або відсутність циклів зворотного зв'язку в мережі. Якщо між нейронами немає зворотного зв'язку (мережа має структуру послідовних шарів, де кожен нейрон отримує інформацію тільки з попередніх шарів), то обробка інформації в мережі є односпрямованою. Вхідний сигнал обробляється послідовністю шарів, і відповідь гарантовано буде отримана за кількістю

тактів, рівною кількості шарів.

Наявність зворотного зв'язку може зробити динаміку нейромереж (в цьому випадку її можна назвати рекурентною) непередбачуваною. У деяких випадках мережа може зупинитися і ніколи не видавати відповідь. І для будь-якої мережі немає алгоритму (по Тюрінгу), який дасть відповідь: чи досягли коли-небудь її елементи стану рівноваги (так звана проблема зупинки). У рекурентних мережах нейрони часто беруть участь в обробці інформації. Це дає змогу виконати глибоку обробку інформації. Зверніть увагу, що в цьому випадку слід застосовувати спеціальні заходи проти перевантаження мережі.

2.3.1 Перцептрони. Прототипи задач

Мережі такого типу сьогодні є найпоширенішими. Переважна кількість прикладних нейросистем передбачає використання багат шарових перцептронів (назва „перцептрон” походить з англійського *perceptron* – сприйняття, оскільки перші зразки таких структур призначались для моделювання зору). Популярність перцептронів зумовлене широким колом доступних для них задач. Загалом вони вирішують задачу апроксимації багатовимірних функцій, іншими словами – побудову багатовимірного відображення $F: x \Rightarrow y$, котре узагальнює заданий набір прикладів (еталонних пар даних) $\{x^a, y^a\}$.

Залежно від типу вихідних змінних (тип вхідних не має вирішального значення), апроксимація функції може набувати вигляду:

- Класифікації (дискретний набір вихідних значень);
- Регресії (неперервні вихідні дані).

Безліч практичних задач розпізнавання зображень, фільтрації шумів, передбачення часових рядів та інші зводяться до цих базових задач.

Розглянемо алгоритм навчання перцептрона на простій модельній задачі. Перцептрон навчають, подаючи сукупність (множину) зображень по одному на його вхід і змінюють ваги доти, доки для всіх зображень не буде

досягнуто необхідний вихід. Припустимо, що вхідні зображення нанесено на демонстраційні карти. Кожну карту розбито на квадрати і від кожного квадрата на перцептрон подається вхідний сигнал. Якщо в квадраті є лінія, то від неї подається одиниця, у протилежному випадку – нуль. Сукупність квадратів на карті задає сукупність нулів і одиниць, котрі подаються на входи перцептрона. Мета полягає в тому, щоб навчити перцептрон вмикати індикатор за умови подавання на нього сукупності входів, що задають непарне число, і не вмикати у випадку парного. На рисунку 2.3 показана така перцептронна конфігурація.

Для кінцевого числа кроків мережа навчиться ділити карти на парні і непарні, за умови, що сукупність чисел розділена лінійно. Таким чином, для всіх непарних карт вихід буде більше порога, а для всіх парних – нижче. Зверніть увагу, що це навчання є глобальним, тобто мережа навчається на всій можливій множині вхідних сигналів.

Причина популярності перцептронів в тому, що для свого кола задач вони є універсальними та ефективними з погляду обчислювальної складності пристроями.

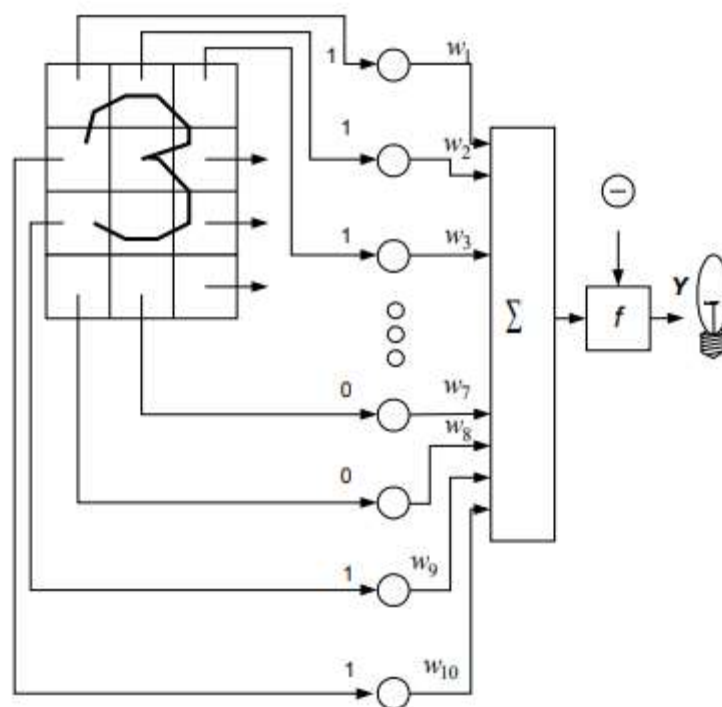


Рисунок 2.3 – Перцептронна система розпізнавання зображень

Можливості одношарових перцептронів

Вивчення можливостей багатшарових перцептронів зручно розпочати із властивостей головної компоненти і водночас найпростішого перцептрона – окремого нейрона.

Найпростішим пристроєм розпізнавання зображень, котрий належить до цього класу мереж, є одиночний нейрон, що перетворює вхідний вектор ознак у скалярну відповідь, яка залежить від лінійної комбінації вхідних змінних:

$$y = f\left(\sum_{j=1}^d w_j x_j + w_0\right) \equiv f\left(\sum_{j=0}^d w_j x_j\right)$$

важатимемо, що у кожного нейрона є додатковий одиничний вхід w_0 , значення якого $x_0 \equiv 1$. Це дає змогу спростити вирази, трактуючи усі синаптичні зв'язки w_j , у тім числі і поріг w_0 , однаково.

Скалярний вихід нейрона можна використовувати як дискримінантну функцію. Цим терміном у теорії розпізнавання зображень називають індикатор приналежності вхідного вектора до одного із заданих класів. Якщо вхідні вектори можуть належати одному з двох класів, нейрон здатний розрізняти тип входу.

Наприклад, якщо $f(x) \geq 0$, вхідний вектор належить першому класу, у всіх інших випадках – другому.

Оскільки дискримінантна функція залежить лише від лінійної комбінації входів, нейрон є лінійним дискримінатором. У деяких випадках лінійний дискримінатор є найкращим із можливих, зокрема, коли ймовірності приналежності вхідних векторів до класу C_i задаються гаусівськими розподілами. У цьому випадку границі, що розділяють області, де імовірність одного класу більша, ніж решта, є гіперплощинами. На рисунку 2.4 подано наочно випадок двох класів векторів.

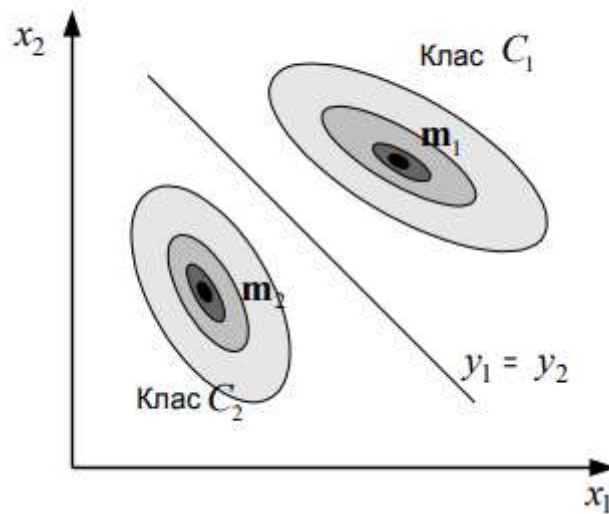


Рисунок 2.4 – Розділення гаусівських класів з математичними сподіваннями m_1 та m_2

Лінійний дискримінатор дає точне вирішення задачі, коли імовірності приналежності до різних класів – гаусівські з однаковою дисперсією і різними математичними сподіваннями m_i . В узагальненому випадку розділ множини вхідних векторів можна здійснити набором гіперплощин. Тоді для вирішення задачі необхідно декілька лінійних дискримінаторів – нейронів.

2.3.2 Вибір функції активації

Монотонні функції активації $f(s)$ не впливають на класифікацію. Однак їх можна вибрати таким чином, щоб вони могли інтерпретувати виходи нейронів як ймовірності приналежності до відповідного класу, що дає додаткову інформацію при класифікації. Можна проілюструвати у приведеному вище випадку гаусівського розподілу ймовірностей сигмоїдна функція активації нейрона $f(s) = \frac{1}{1 + \exp(-as)}$ дає імовірності приналежності до цього класу.

Розглянемо детальніше типи функції активації та їхні властивості. Вихід нейрона у визначається функцією активації $y = f(s)$, де s сума його n

входів $s = \sum_{j=1}^n x_j w_j$. Функції активації можуть набути різного вигляду. Однією з найпоширеніших функцій активації є логістична функція (сигмоїд):

$$f(s) = \frac{1}{1 + \exp(-as)}$$

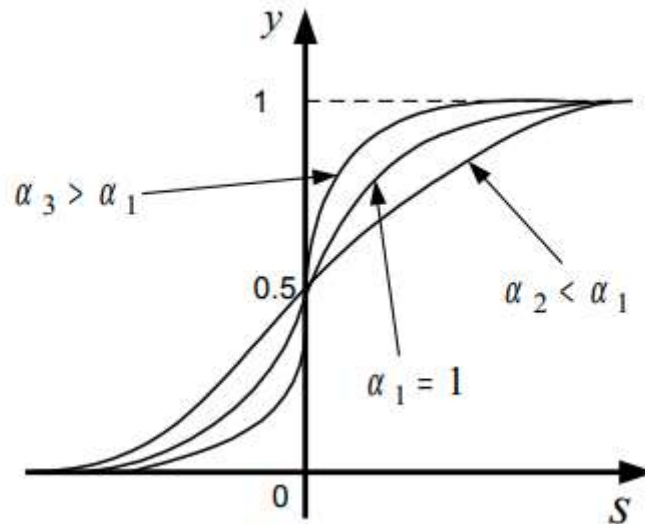


Рисунок 2.5 – Вигляд сигмоїдної функції за різних значень α

Зі зменшенням α сигмоїд стає дещо пологішим і в границі при $\alpha \rightarrow 0$ вироджується у горизонтальну лінію на рівні 0,5. Зі збільшенням α наближується за зовнішнім виглядом до функції одиничного стрибка (функції Хевісайда) з порогом у нульовій точці. Очевидно, що значення функції змінюється від нуля до одиниці при зміні аргументу від $-\infty$ до $+\infty$, отже кажуть, що вона має стискаючу властивість. Цінною властивістю функції є простий вираз для її похідної та її неперервність:

$$f'(s) = \alpha f(s)(1 - f(s))$$

У перших перцептронах використовували саме функцію Хевісайда, яка має розривну похідну, що суттєво гальмувало розвиток алгоритмів навчання нейромереж.

Часто для покращення збіжності алгоритму навчання або для врахування особливостей поставленою перед нейромережею задачі вводять поріг функції θ . Тоді функція задається залежністю $y = \frac{1}{1 + \exp(-a(s - \theta))}$. Вигляд функції при введенні порогу θ проілюстровано на рисунку 2. 6.

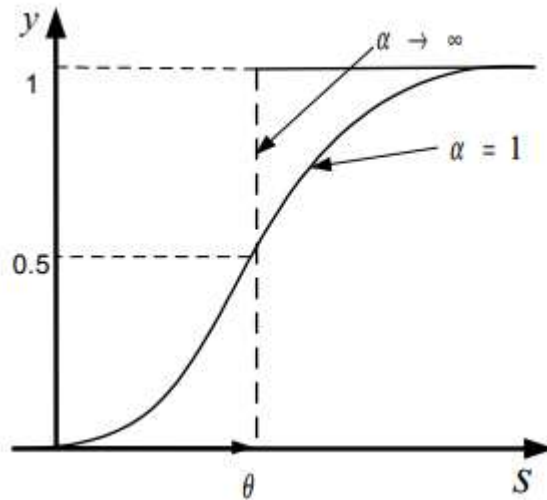


Рисунок 2.6 – Вигляд сигмоїдальної функції при введенні порогу θ

Ще однією функцією, котру використовують як функцію активації (вона теж володіє стискаючими властивостями) є тангенс гіперболічний: $y = \frac{\exp(as) - \exp(-as)}{\exp(as) + \exp(-as)}$. Вигляд функції проілюстровано на рисунку 2.7.

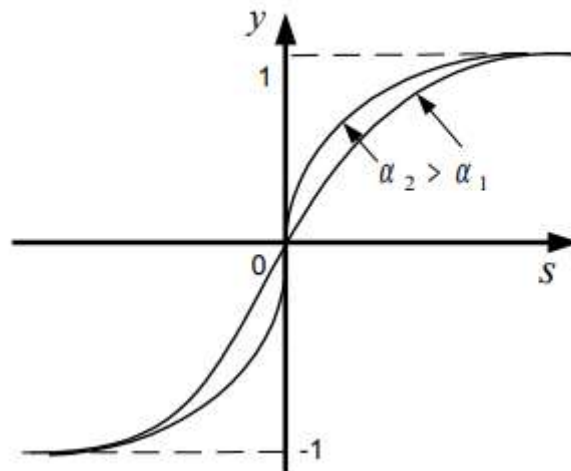


Рисунок 2.7 – Вигляд функції $\text{th}(s)$

Як і сигмоїдальна функція, тангенс гіперболічний має простий вираз для похідної: $tn(s)' = 1 - tn^2(s)$.

2.4 Багатошарові пепсетрони

Можливості лінійного дискримінатора обмежені. Він здатний класифікувати лише лінійно-роздільні множини векторів, які можна розділити гіперплощиною (рис. 2.8)

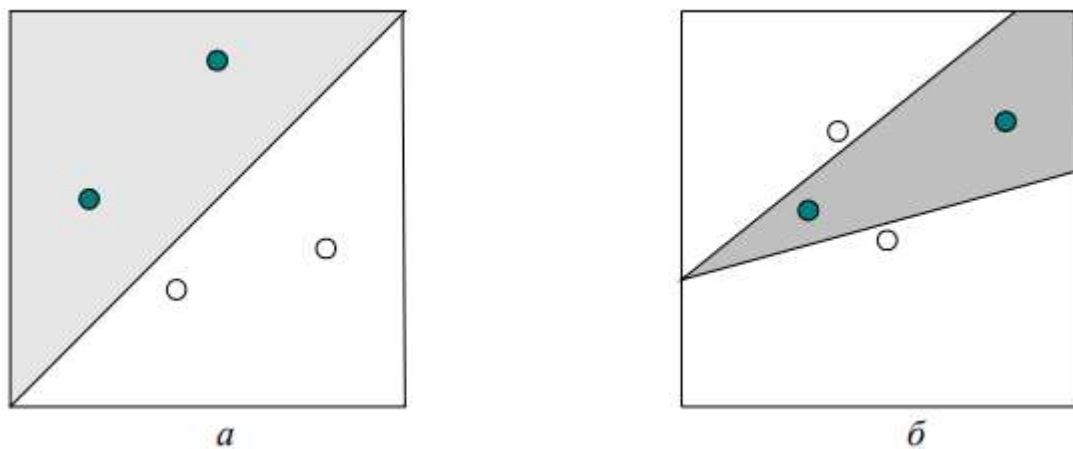


Рисунок 2.8 – Приклад лінійно роздільних (а) і лінійнонероздільних (б) множин

У d -мірному просторі гіперплощина може розділити довільним способом $d + 1$ точку. Наприклад, на площині можна довільно розділити на два класи три точки, проте розділити чотири точки в загальному випадку уже неможливо (див. рис. 2.8).

Якщо точок більше, ніж $d + 1$, завжди існують такі способи розбиття на два класи, котрі неможливо здійснити за допомогою однієї гіперплощини, проте це можна зробити за допомогою декількох гіперплощин.

Для вирішення складніших задач класифікації необхідно ускладнити мережу, вводячи додаткові (їх називають схованими) шари нейронів, котрі здійснюють проміжну обробку даних так, щоб вихідний нейрон – класифікатор отримував на свої входи вже лінійно-роздільні множини. Легко пересвідчитись, що завжди можна обійтись лише одним схованим шаром,

який містить достатньо велику кількість нейронів. Дійсно, збільшення схованого шару підвищує розмірність простору, в якому вихідний нейрон виконує дихотомію (розділення на дві частини). Отож завжди можна обійтись тришаровою нейромережею.

Унаслідок досліджень багатшарових персептронів зроблено такі висновки:

- Мережа з одним схованим шаром, який містить N нейронів зі сходинковою функцією активації, здатна здійснити довільну класифікацію Nd точок d вимірного простору (класифікувати Nd прикладів).

- Одного схованого шару нейронів з сигмоїдною функцією активації достатньо для апроксимації довільної границі між класами з будь-якою точністю.

Для задач апроксимації ці висновки можна сформулювати так:

- Одного схованого шару нейронів з сигмоїдною функцією активації достатньо для апроксимації довільної функції з як завгодно високою точністю. Така мережа може водночас апроксимувати і саму функцію, і її похідні.

Точність апроксимації зростає із зростанням кількості нейронів у схованому шарі. При N нейронах похибка оцінюється як $O\left(\frac{1}{N}\right)$

Розглянемо детальніше можливості багатшарового персептрона на прикладі вирішення проблеми функції заперечуючого АБО.

2.4.1 Проблема функції заперечуючого АБО (XOR)

Один з найпесимістичніших результатів дослідження одношарових персептронів, отриманих Мінським [6], засвідчує, що одношаровий персептрон не може відтворити таку просту функцію, як XOR .

Це функція від двох аргументів, кожен з яких може бути нулем чи одиницею. Вона набуває значення одиниці, коли один з аргументів дорівнює

одиниці. Проблему можна проілюструвати за допомогою одношарової однонейронної мережі з двома входами, поданої на рисунку 2.9. Позначимо один вхід через x_1 , а інший – через x_2 . Тоді всі їхні можливі комбінації входів налічуватимуть чотири точки на площині X_1X_2 , як проілюстровано на рисунку 2.3. Наприклад, точку $x_1 = 0$ і $x_2 = 0$ позначено на рисунку як точку A_0 . Таблиця 2.1 ілюструє необхідний зв'язок між входами і виходом, де вхідні комбінації, що повинні давати нульовий вихід, позначені A_0 і A_1 , одиничний вихід – B_0 і B_1 .

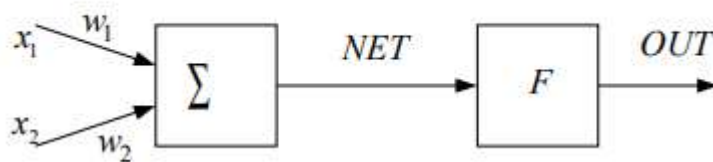


Рисунок 2.9 – Однонейронна система

У мережі на рисунку 2.9 функція F є функцією Хевісайда з порогом $\theta = 0,5$, отож $OUT = 0$, коли $NET < 0,5$, і одиниці у випадку, коли $NET \geq 0,5$.

Нейрон виконує наступне обчислення:

$$NET = x_1w_1 + x_2w_2 \quad (2.1)$$

Ми бачимо, що жодна комбінація значень двох ваг не може дати співвідношення між входом і виходом, задане таблицею 2.1. Щоб зрозуміти це обмеження, зафіксуємо NET на величині порогу 0,5. Мережа в цьому випадку описується рівнянням (2.2) Це рівняння лінійне щодо x_1 та x_2 . Отже, всі значення x_1 та x_2 , що задовольняють цьому рівнянню, будуть лежатимуть на деякій прямій у площині X_1X_2 .

$$x_1w_1 + x_2w_2 = 0,5 \quad (2.2)$$

Таблиця 2.1. Таблиця істинності для функції XOR

Точки	Значення x	Значення y	Необхідний вихід
A_0	0	0	0
B_0	1	0	1
A_1	0	1	1
B_1	1	1	0

Будь-які вхідні значення для x та y на цій лінії даватимуть граничне значення 0,5 для NET. Вхідні значення з одного боку прямої забезпечать значення NET більше порогу. Отже, $OUT = 1$. Вхідні значення з іншого боку прямої забезпечать значення NET менше граничного значення, роблячи OUT рівним нулю. Зміни значень w_1 , w_2 і порогу змінюватимуть нахил і положення прямої.

Щоб мережа реалізувала функцію XOR, задану таблицею 2.1, необхідно розташувати пряму так, щоб точки А були з одного боку прямої, а точки В – з іншого. Спробуємо намалювати таку пряму на рисунку 2.10. Переконаємося, що це неможливо. Отже, які б значення не приписувалися вагам і порогові, мережа не здатна відтворити співвідношення між входом і виходом, необхідне для представлення функції XOR.

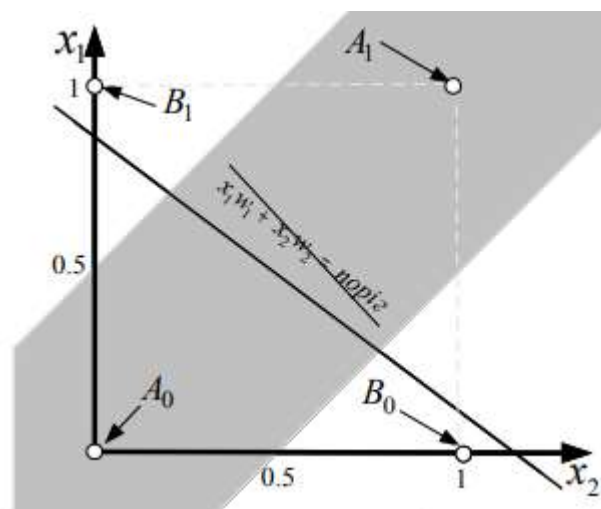


Рисунок 2.10 – Проблема XOR

2.5. Лінійна роздільність

Як бачимо, неможливо намалювати пряму лінію, що розділяє площину X_1X_2 так, щоб реалізовувалася функція XOR. На жаль, цей приклад не єдиний. Існує великий клас функцій, котрі неможливо розділити за допомогою одношарової мережі. Про ці функції говорять, що вони є лінійно нероздільними, і саме вони накладають вагомі обмеження на можливості одношарових мереж.

Лінійна роздільна здатність обмежує одношарові мережі завданнями класифікації, в яких кілька точок, відповідних вхідним значенням, можуть бути розділені гіперплощинами. У нашому випадку з двома входами розділова межа являє собою пряму лінію. Для трьох входів поділ відбувається площиною, яка розсікає тривимірний простір. Для чотирьох чи більше входів візуалізація неможлива і необхідно подумки представити n – вимірний простір, що розсікається «гіперплощиною» – геометричним об'єктом, що розсікає простір чотирьох чи більше вимірів.

Оскільки лінійне розділення обмежує можливості перцептронного представлення, важливо знати, чи дозволена зазначена функція лінійно. На жаль, немає простого способу визначити це у випадку великої кількості змінних.

У випадку нейронної мережі з n двійковими входами множина вхідних образів (векторів) є скінченою і має 2^n різних вхідних образів, що складаються з нулів та одиниць. Оскільки кожен вхідний образ може відповідати двом різним бінарним виходам (одиниця і нуль), то усього є 2^{2^n} функцій від n змінних.

Таблиця 2.2 – Кількість лінійно роздільних функцій у загальній кількості можливих двійкових функцій

№ з/п	Загальна кількість функцій 2^{2n}	Кількість лінійно роздільних функцій
1	4	4
2	16	14
3	256	104
4	65536	1882
5	$4,3 \times 10^9$	94572
6	$1,8 \times 10^{19}$	15028134

Як бачимо з таблиці 2.2, імовірність того, що випадково обрана функція виявиться лінійно роздільною, дуже мала навіть для помірної кількості змінних. Отож одношарові персептрони на практиці обмежені простими задачами.

2.6 Аналіз методів подолання обмеження лінійної роздільності

До кінця 60-х років проблема лінійної роздільності була добре зрозуміла. Вважали, що це серйозне обмеження одношарових мереж можна перебороти, додавши додаткові шари. Наприклад, ви можете отримати двошарові мережі, каскадуючи дві одношарові мережі. Вони здатні виконувати узагальнену класифікацію, відокремлюючи ті точки, що містяться в опуклих обмежених чи необмежених областях. Область називається опуклою, якщо лінія, що з'єднує її, повністю знаходиться в цій області для двох її точок. Область називається обмеженою, якщо її можна помістити в деяке коло. Необмежену область неможливо укласти всередину кола (наприклад, область між двома рівнобіжними лініями). Приклади опуклих обмежених і необмежених областей представлена на рисунку 2.11.

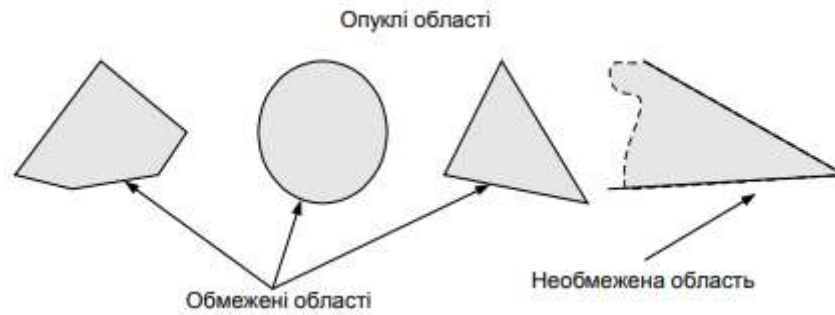


Рисунок 2.11 – Опуклі обмежені та необмежені області

Щоб уточнити вимогу опуклості, розглянемо просту двошарову мережу з двома входами, підведеними до двох нейронів першого шару і з'єднаними з єдиним нейроном у другому шарі (рис. 2.12). Нехай поріг вихідного нейрона дорівнює 0,75, а обидві його ваги рівні 0,5. У цьому випадку для того, щоб поріг був перевищений і на виході з'явилася одиниця, необхідно, щоб обидва нейрони першого рівня на виході мали одиницю. Отже, вихідний нейрон реалізує логічну функцію "І". На рисунку 2.13 кожен нейрон першого шару розбиває площину X_1X_2 на дві півплощини: один забезпечує одиничний вихід для входів нижче верхньої лінії, інший – для входів вище нижньої лінії. На рисунку 2.12 проілюстровано результат такого подвійного розбиття, де вихідний сигнал нейрона другого шару дорівнює одиниці тільки всередині V -подібної області. Аналогічним чином, в першому шарі можна використовувати три нейрони з подальшим поділом рівня і створенням трикутної області. Включення достатньої кількості нейронів у вхідний шар може викликати утворення опуклого багатокутника будь-якої бажаної форми. Точки, які не відносяться до опуклої області, не можуть бути відокремлені від інших точок на площині двошаровою мережею.

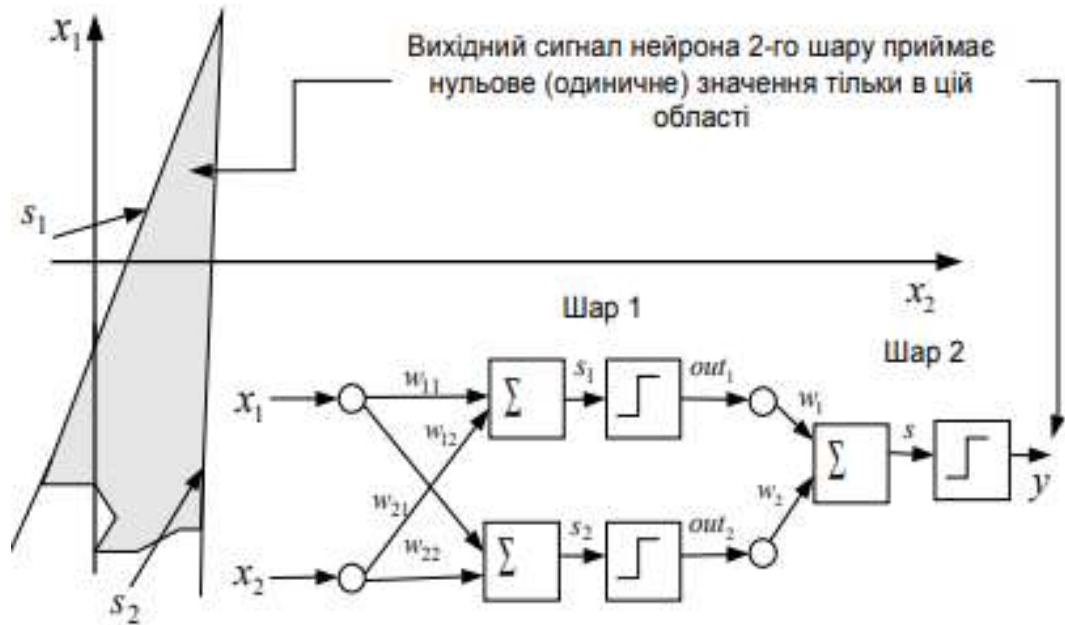


Рисунок 2.12 – Опукла необмежена область, виокремлена двошаровою мережею

Нейрон другого шару не обмежений функцією “Г”. Він може реалізовувати безліч інших функцій за певного вибору ваг і порогу. Наприклад, можна зробити так, щоб одиничний вихід одного з нейронів першого шару викликав одиницю на виході нейрона другого шару, реалізовуючи тим самим логічне “АБО”. Загалом існує 16 двійкових функцій від двох змінних. Якщо обирати певним способом ваги і поріг, то можна відтворити усі логічні функції за допомогою двошарової мережі.

Наведемо один з можливих варіантів двошарової мережі (див. рис. 2.12), що реалізує функцію *XOR* згідно з таблицею істинності 2.1.

Функцію *XOR* можна записати так:

$$y = (x_1 \text{ AND } \text{NOT } x_2) \text{ OR } (x_2 \text{ AND } \text{NOT } x_1).$$

Тоді функцію *NOT* можна реалізувати за допомогою від’ємного значення ваги, функцію *AND* за допомогою нейронів першого шару а функцію *OR* – другого шару. Обираючи за функцію активації функцію

Хевісайда (див. рис. 2.9 при $a \rightarrow \infty$) зі значенням порогу, що дорівнює 0,5, та значення ваг за таблицею 2.3, отримаємо мережу, що реалізує функцію XOR.

Таблиця 2.3 – Значення ваг нейромережі

Вага	w_{11}	w_{12}	w_{21}	w_{22}	w_1	w_1
Значення	1	-1	-1	1	1	1

Значення виходів нейронів першого та другого шарів і значення проміжних величин після операції сумування що здійснює кожний з нейронів подано у таблиці 2.4. Вони засвідчують виконання мережею логічної функції XOR, котра відповідає розбиттю площини вхідних векторів на дві лінійно нероздільні множини, як проілюстровано на рисунку 2.10 (сірим позначено область вхідних векторів, для котрих вихід мережі дорівнює нулю).

Таблиця 2.4 – Значення входів і виходів нейромережі

x_1	x_2	s_1	s_2	out_1	out_2	s	y
1	0	0.5	-1.5	1	0	0.5	1
0	1	-1.5	0.5	0	1	0.5	1
1	1	-0.5	-0.5	0	0	-0.5	0
0	0	-0.5	-0.5	0	0	-0.5	0

Входи не обов'язково є двійковими. Вихідний вектор може бути аналоговим і зображуватись довільною точкою на площині XY. У цьому випадку ми маємо справу зі здатністю мережі розбивати площину на неперервні області, а не з поділом дискретної множини точок.

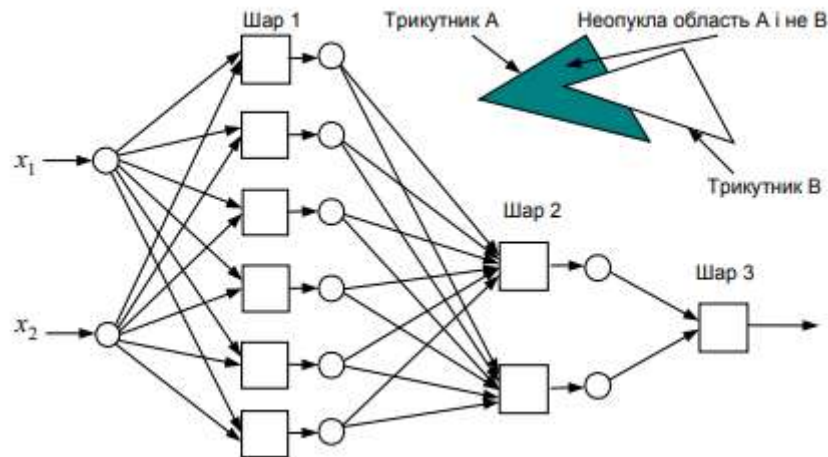


Рисунок 2.13 – Увігнута область, що задається тришаровою мережею

У тришаровій мережі можливості класифікації обмежені лише кількістю нейронів і ваг. Обмеження на опуклість області, що виокремлюється, відсутні. Нейрон третього шару вважає входом набір опуклих багатокутників, а їхня логічна комбінація може бути неопуклою. На рис. 2.13 проілюстровано випадок, коли дві трикутні області А і В, утворені за допомогою функції «А і не В», задають неопуклу область. За додавання нейронів і ваг кількість сторін багатокутників може необмежено зростати. Це дає змогу апроксимувати область будь-якої форми з будь-якою точністю. Зауважимо, що не всі вихідні області другого шару повинні перетинатися. Отже, можна поєднувати різні області, опуклі і неопуклі, видаючи на виході одиницю кожен раз, коли вхідний вектор належить до однієї з них.

Незважаючи на те, що можливості багатшарових мереж були відомі давно, протягом багатьох років не існувало теоретично обґрунтованого алгоритму для настроювання їхніх ваг.

3 ОПИС СИСТЕМИ ІДЕНТИФІКАЦІЇ

3.1 Структурна схема системи розпізнавання відео

Розглянемо блок-схему системи розпізнавання відео, яка показана на рисунку 3.1:

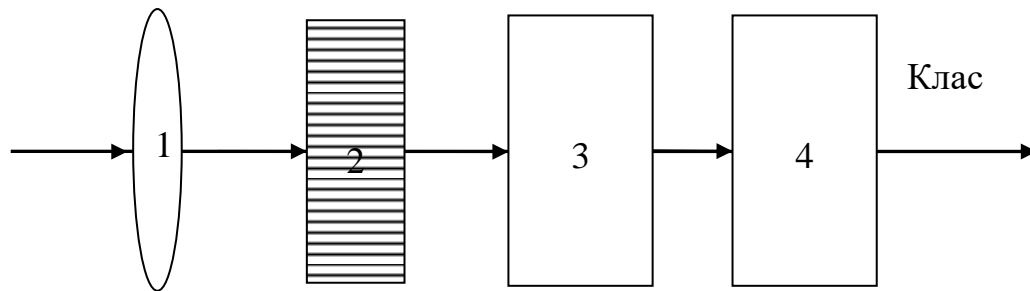


Рисунок 3.1 – Блок-схема системи розпізнавання

- 1 – оптична система;
- 2 – матриця фотоелементів;
- 3 – система попередньої обробки;
- 4 – комп'ютерна система (комп'ютер).

В систему розпізнавання надходить оптичний сигнал $A(x, y)$, який фокусується оптичною системою на матриці фотоелементів. Оптична система характеризується імпульсною характеристикою $H(x, y)$. Фізичний зміст імпульсної характеристики полягає в наступному: імпульсна характеристика — це відповідь системи на точкове джерело $\delta(x, y)$. Для системи розпізнавання функція $H(x, y)$ — одна з найважливіших характеристик, оскільки залежить від таких значень, як відстань, на якій система може з заданою надійністю розпізнавати об'єкти. Потім на виході оптичної системи вхідний сигнал перетворюється в сигнал такого типу:

$$B(x, y) = \iint_S A(\xi, \eta) H(x - \xi, y - \eta) d\xi d\eta \quad (3.1)$$

Матриця фотоелементів перетворює оптичний сигнал $B(x, y)$ в електричний. Слід враховувати, що фізично реалізовані лише матриці з кінцевою кількістю фотоелементів. Таким чином, на виході матриці фотоелементів буде багато електричних сигналів B_{ij} , $i=1, \dots, m$, $j=1, \dots, n$. Кількість елементів матриці визначає таку важливу характеристику системи, як роздільна здатність по горизонталі і вертикалі, а також якість зображення, що відображається на комп'ютерній системі.

Електричний сигнал від масиву фотоелементів передається в систему попередньої обробки зображення, призначену для покращення якості зображення, кодування та апроксимації.

Комп'ютерна система є основним блоком усієї системи розпізнавання, яка приймає рішення про належність об'єкта до певного класу. У нашому випадку буде лише два класи, а також клас «відмова від визнання», що означає, що прийняття рішення на користь того чи іншого класу з тих чи інших причин неможливе.

Режими роботи системи розпізнавання зображень:

1. фоновий режим.

У цьому режимі об'єкт не знаходиться в полі зору системи, режим «сплячий».

2. виділення об'єкта з фону (виявлення об'єкта).

Вибір об'єкта здійснюється на основі відмінностей між об'єктом і фоном (різний колір, яскравість, рух тощо).

3. розпізнавання об'єктів.

3.2 Опис оптичної системи

Як уже зазначалося, оптична система визначає багато важливих характеристик системи розпізнавання зображень, тому її слід враховувати при вивченні роботи системи розпізнавання зображень.

Характеристики оптичної системи визначаються її імпульсною характеристикою $H(x, y)$. Розглянемо, якими є імпульсні характеристики різних елементів оптичної системи.

Основними елементами оптичних систем є лінзи та керовані банери. Робота цих оптичних систем заснована на відомій властивості оптичної лінзи здійснювати просторове перетворення Фур'є від вхідного сигналу $A(x, y)$, що імпульсна характеристика лінзи має вигляд:

$$H_L(x, y) = \exp\{i(x + y)\} \quad (3.2)$$

При цьому перетворенні спектральні компоненти зображення (різні просторові частоти) поширюються під різними кутами до оптичної осі системи, а всі просторові спектральні компоненти зберігають свої амплітуди і фази. Друга лінза, вхідним сигналом якої є просторовий спектр сигналу $A(x, y)$, виконує перетворення Фур'є цього спектру. Таким чином, друга лінза відновлює вихідне зображення з тією лише різницею, що воно буде перевернуто.

У розглянутій оптичній системі можна виконати велику кількість різноманітних функціональних перетворень вхідних сигналів та зображень [5]. Для цього в оптичну систему вводяться два керованих банера. Один з них розміщується у вхідній площині системи, і в цьому випадку оброблене зображення буде добутком вхідного зображення на амплітудно-фазову функцію передачі керованого банера $T_1(x, y)$. Інший банер розміщується в спектральній площині, і тут також формується добуток вхідного спектру і передавальної функції банера $T_2(x, y)$.

Вибравши дві функції, реалізовані на обох банерах, можна виконати необхідне перетворення зображень. При цьому може бути реалізована велика кількість режимів роботи такої оптичної системи, при якій окремі банери можуть функціонувати як разом, так і окремо. Наприклад, якщо є лише другий банер, система може виконувати функції просторового фільтра, який усуває непотрібні просторові частоти або шум.

У загальному випадку оптичний пошук можна представити як процес отримання інформації про поле випромінювання, розподілене в просторі. Особливістю джерел такої інформації є їх двовимірність. Процеси отримання та обробки двовимірної інформації можуть виконуватися або шляхом паралельного перегляду всіх елементів поля без сканування (parallel viewing), або шляхом сканування простою точкою (послідовний перегляд), або шляхом сканування за допомогою складеної плями (послідовний-паралельний перегляд). Методи перегляду простору за допомогою простого скануючого плями представлені на рис 3.2 .

Букви вказують на такі методи огляду: а – паралельний; б – послідовний; в – послідовно-паралельний; г – паралельно-послідовний; е – послідовно-послідовний [5]

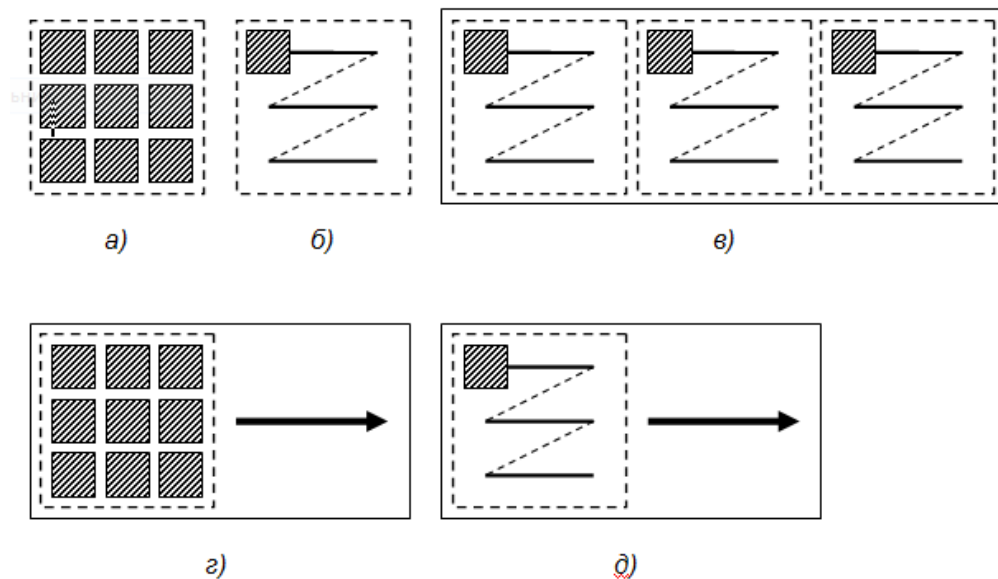


Рисунок 3.2 – Методи перегляду простору за допомогою простої скануючої плями

Системи паралельного сприйняття та аналізу візуальної інформації складаються з великої кількості елементарних паралельних каналів, що працюють одночасно. При цьому потік візуальної інформації може зберігати двовимірний характер на всіх етапах її обробки. Одночасний аналіз зображень дозволяє отримати вищу продуктивність порівняно з послідовним. Це пов'язано з тим, що багатоканальна система аналізу призначена для одночасного перетворення зображення в n одновимірних функцій часу за допомогою відповідного багатоканального перетворення світлового сигналу. У цьому випадку подальший аналіз прийнятих сигналів також може виконуватися одночасно на всіх каналах. Очевидно, що відсутність послідовного сканування зображення скорочує час його повного аналізу.

Однак зрозуміло, що наявність багатьох каналів призводить до значного ускладнення таких систем. Методи перегляду простору за допомогою композитної скануючої плями (комплексної апертури) показані на рис. 3.3. Букви позначають такі способи перегляду простору: а – послідовний; б – послідовно-паралельний; в – послідовно-послідовний [5].

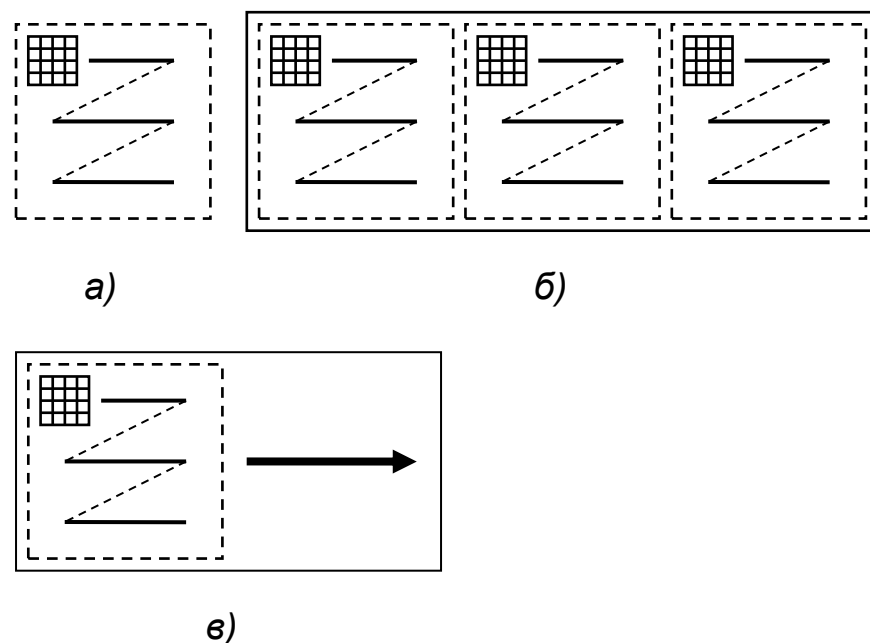


Рисунок 3.3 – Методи перегляду простору за допомогою композитної скануючої плями (комплексної апертури)

3.3 Опис матриці фотоелементів

Для одночасного перетворення зображення в набір сигналів використовуються різні багатоканальні мозаїчні фотоелектричні перетворювачі. Вихідний сигнал кожного елемента такої мозаїки визначається характеристиками падаючого випромінювання та параметрами елемента. Такі багатоканальні приймачі можуть працювати в режимі безперервного зчитування (без накопичення сигналу) і в режимі накопичення при імпульсному зчитуванні. Набір сигналів, що надходять від багатоканального фотоперетворювача, може оброблятися послідовно, паралельно і комплексно (паралельно-послідовно). Природно, що при використанні багатоканального датчика найбільш раціонально використовувати одночасний або паралельно-послідовний метод обробки сигналів, що надходять від окремих елементів.

3.4 Опис системи попередньої обробки

Кодування та апроксимація. Для обробки на комп'ютерах вхідні дані спочатку оцифровують, а потім представляють у більш компактному вигляді (стиск даних) за допомогою схем кодування або апроксимації [4].

Відомо, що якщо перетворення Фур'є функції (одновимірної чи двовимірної) відмінне від нуля в обмеженій області B частотного простору, то цю функцію можна точно відновити за її значеннями в кінцевому наборі опорних точок (теорема Котельникова).

На практиці перетворення Фур'є за межами області B лише близьке до нуля. Тому в результаті відновлення вибіркової функції буде отримано лише її наближення. В інших методах вибірки аналізована функція інтерполюється між опорними точками за допомогою простих типів функцій (наприклад, константних, лінійних або поліномів вищого порядку).

Інший метод апроксимації – квантування. У цьому випадку фактичне значення функції апроксимується за допомогою кінцевого набору значень, які називаються рівнями квантування. Найпростіший спосіб – використовувати рівні квантування з рівномірним розташуванням. Однак іноді краще використовувати рівні з нерівномірним розташуванням, наприклад, якщо рівні яскравості в певному діапазоні є більш поширеними на зображенні, ніж в інших діапазонах. Це збільшить середню щільність квантування без збільшення кількості використовуваних рівнів. Опорні точки або рівні квантування не потрібно встановлювати заздалегідь. Їх вибір може залежати від природи об'єктів, які підлягають апроксимації. У цьому випадку, звичайно, буде потрібно більш складна схема вибірки або квантування.

Фільтрація, відновлення та поліпшення. З метою фільтрації об'єкта з метою пошуку заданої конфігурації (порівняння зі стандартом) або відновлення вихідного об'єкта, який був пошкоджений в результаті апроксимації чи інших подібних операцій, а також для покращення якості зображення шляхом згладжування або, навпаки, для збільшення контрасту можна використовувати операції, які є інваріантними до часу і для зміни положення (тобто інваріантні до зсуву). Нехай F — набір функцій, кожна з яких визначає об'єкт, а T — операція перенесення (зсуву) у часі та просторі, визначена на F . Операція φ називається інваріантним щодо зміни моменту часу або положення в просторі, якщо він комутує з будь-якою операцією T в часі або просторі, тобто якщо

$$\varphi(T(f)) = T(\varphi(f)) \quad (3.3)$$

для всіх $f \in F$. Такі операції мають властивість, що їх вплив на значення функції в будь-який момент часу або в будь-яку дану точку простору не залежить від цього моменту і від положення цієї точки. До цієї категорії належать операції зсуву, точкові операції ($\varphi(f)$) в точці залежить

лише від значення f у цій точці) та локальних операцій ($\varphi(f)$ в точці залежить тільки від значень f в деякій околиці цієї точки). Типовим прикладом лінійної операції, інваріантною до зміни часу або положення, є перетворення Фур'є (одновимірне або двовимірне).

Часто необхідно визначити, наскільки два об'єкти узгоджуються один з одним (наскільки вони схожі) або знайти схожість між одним об'єктом і частиною іншого. Обчислення взаємної кореляції між двома функціями f і g є одним із найпростіших методів встановлення тотожності двох об'єктів аж до перенесення та множення на константу. Цей метод може бути дуже корисним для завдань розпізнавання зображень або їх непохідних елементів, для яких легко встановити прототипи або стандарти для цих зображень.

У цьому випадку завдання розпізнавання зводиться до розрахунку взаємної кореляції вихідного об'єкта з кожним із еталонів. Аналогічно задачу знаходження об'єкта f в об'єкті g можна вирішити, обчисливши відношення двох взаємних коефіцієнтів кореляції. Нехай об'єкт g містить шумний об'єкт f , і нам потрібно знайти цей об'єкт, обчисливши взаємну кореляцію стандарту f^* з об'єктом g . Ми можемо показати, що якщо шум є адитивним і не залежить від f , то найкращим стандартом f^* для виявлення f є сам об'єкт f . Це точно узгоджений фільтр.

Інший спосіб порівняння об'єкта з еталоном — обчислення не згортки, а добутку перетворень Фур'є об'єктів і подальшого обчислення оберненого перетворення Фур'є отриманого добутку. У цьому випадку процес порівняння визначається в частотному діапазоні або в просторових частотах. Одним з найбільш важливих застосувань такої фільтрації є завдання відновлення об'єкта. Припустимо, що існуючий об'єкт спотворений процесом міграції або наближення. Якщо спотворений об'єкт можна представити у вигляді згортки f^*g , де f — вихідний об'єкт, а g — деяка функція, то для відновлення необхідно обчислити обернене перетворення Фур'є функції $\frac{FG}{G}$ де FG — перетворення Фур'є згортки f^*g .

Процес згладжування часто використовується для придушення шуму в об'єкті. Згладжування ви можете зробити, просто встановивши значення функції в кожній точці на середнє значення навколо цієї точки. Розмір району може бути постійним або варіюватися від точки до точки. Ви також можете згладити, обчисливши безліч незалежних екземплярів об'єкта в середньому. Крім лінійних операцій, для усереднення можна використовувати нелінійні операції. Найпростішими операціями такого роду є поєднання усереднення з пороговою операцією. Інший спосіб згладити об'єкт — використовувати смугові фільтри.

Іноді може знадобитися не згладити об'єкт, а, навпаки, підкреслити його. Хоча згладжування вимагає інтеграції, природно застосовувати диференціацію в тій чи іншій формі, щоб підкреслити. При диференціації зображень використовується обчислення градієнта і лапласіана. Якщо ви хочете підкреслити зображення незалежно від напрямку, вам потрібно обчислити похідну в кожній точці в напрямку градієнта. Приблизно цю операцію можна реалізувати у вигляді згортки заданого зображення або шляхом застосування послідовного градієнтного фільтра.

Лапласіан — це корисна комбінація похідних, яка пропорційна різниці між рівнем яскравості в точці та середнім рівнем яскравості в кільці з центром у цій точці. Приблизне обчислення лапласіана зводиться до згортки функції зображення з опорною функцією, що має позитивний пік, оточений кільцем негативних значень. Ці додатні та від'ємні значення потрібно вибрати так, щоб інтеграл по області дорівнював нулю.

Сегментація зображення. Не існує загального методу сегментації зображень. Залежно від типу необхідного опису можуть бути корисні різні типи фрагментів зображень. Основний спосіб розмежування фрагмента-застосувати порогову операцію до цього зображення. Фрагмент може являти собою набір точок, яскравість яких більше або менше порогового значення або знаходиться між двома пороговими значеннями. Поріг операція може поєднуватися з іншими процедурами попередня обробка.

Значна складність, що виникає при використанні порогової операції, полягає у виборі значення (значень) порогового значення. Іноді ви можете знайти хороший поріг, вивчивши розподіл яскравості точок зображення. Наприклад, знаючи приблизну площу фрагмента, ви можете вибрати нижній (верхній) поріг таким чином, щоб область фрагмента, де яскравість перевищує поріг, була меншою (більшою), ніж частка площі фрагмента. вся картина. Цей метод називається методом процентиля.

Іншим підходом до вирішення проблеми виділення фрагментів є відстеження меж (контурів). Використання цього методу дозволяє не тільки виділити фрагменти, але й отримати інформацію про форму зображень. Простим методом відстеження меж у двоградацийному зображенні є систематичний перегляд його до тих пір, поки не зустрінеться пара точок або растрових осередків різної яскравості (наприклад, чорно-білі точки).

Цей метод можна узагальнити на випадок напівтонових зображень, якщо вказати критерій існування кордону. Таким критерієм може бути, наприклад, наявність сусідніх осередків, різниця яскравості яких перевищує певний поріг. Кордони можна відстежувати, навіть якщо є прогаліни. Для цього при виявленні кінця кордону слідкуйте за відстанню, що відповідає максимально допустимому зазору в попередньому напрямку.

Ви можете узагальнити цей метод на зображення на випадок напівтонових зображень, якщо вказати критерій існування кордону. Наприклад, таким критерієм може бути наявність сусідніх осередків, різниця в яскравості яких перевищує певний поріг. Межі можна відстежувати, навіть якщо є прогаліни. Для цього при виявленні кінця кордону зверніть увагу на відстань, рівну максимально допустимому відстані в попередньому напрямку.

3.5 Опис комп'ютерної системи

Комп'ютерна система отримує сигнал, який вже пройшов попередню обробку, включаючи вибірку зображення. Тому будемо вважати, що сигнал на вході обчислювальної системи представлений у вигляді матриці $\{C_{mn}\}$.

При описі обчислювальної системи виникає інтерес розглянути детальний розгляд режимів її роботи.

а) Фоновий режим.

Працюючи у цьому режимі система здійснює огляд простору з метою виявлення об'єкта. Розроблено різні методи оптичного огляду параметричних полів. Про способи огляду простору написано в розділі, присвяченому оптичній системі.

б) Режим виявлення об'єкта.

Виявлення об'єкта ґрунтується на визначенні відмінностей між об'єктом та фоном. Цими відмінностями можуть бути відмінності у фарбуванні, яскравості, рух об'єкта. Залежно від вибраного критерію, за яким об'єкт виділяється з фону, використовуються різні способи виявлення об'єкта.

1) Відмінності в яскравості фону і об'єкту;

За допомогою цього критерію розрізнення «об'єкта» можна вважати ділянку зображення, яскравість якої вище або нижче попередньо вибраного порогу яскравості. Крім того, можна вибрати об'єкт на основі перевищення заданого порогу за різницею в яскравості між сусідніми елементами зображення. Тоді має сенс знайти градієнт яскравості по всьому зображенню. Градієнт яскравості більш детально обговорюється в розділі про попередню обробку зображень.

2. Рух об'єкту;

Для виділення рухомого об'єкту зі статичного фону доводиться проводити порівняння двох кадрів, розділених деяким проміжком часу, і виявляти різницю між ними. У цьому випадку "об'єктом" вважається область

зображення, яка відрізняється у цих зображень.

3. Розпізнавання об'єкту;

Розпізнавання об'єкту – режим, заради якого, власне, будується вся система розпізнавання. Основне питання, яке необхідно розглянути в даному розділі – побудова алгоритму розпізнавання, за яким працюватиме система розпізнавання.

4 РОЗРОБКА СИСТЕМИ РОЗПІЗНАВАННЯ

Беручи до уваги існуючі алгоритми розпізнавання зображень, можна прийти до висновку про вибір методу розпізнавання в нашому випадку згідно поставлених задач.

Задачі класифікації:

1. Задача ідентифікації, яка полягає у тому, щоб ідентифікувати конкретний об'єкт серед йому подібних.

2. Присвоєння об'єкта певному класу. Це може бути, наприклад, завдання виявлення символів, або рішення про наявність дефекту в окремих технічних деталях.

3. Кластерний аналіз, який полягає в розділенні заданого набору об'єктів на класи – групи об'єктів, схожих один на одного за тим чи іншим критерієм. Цю задачу називають класифікацією без учителя, тому що, на відміну від завдання 2, класи не визначені апіорі.

Для дискримінантного розпізнавання, або для розпізнавання в просторі ознак об'єктів, характерним є те, що кожен об'єкт зображується окремою точкою в деякому просторі. Координатними осями цього простору є ознаки, за якими здійснюється розпізнавання. Отже, координати об'єктів – це значення відповідних характеристик. Далі виявлення здійснюється на основі аналізу близькості між об'єктами.

Існують різні типи знаків: дихотомічні (знак може бути присутнім або бути відсутнім; наприклад, є вуха або їх немає); Іменні (наприклад, колір: жовтий, білий, помаранчевий і т. д.); порядкові (наприклад, "великий" – "середній" – "маленький"); кількісні. Для кожного типу об'єктів ви можете ввести свої власні вимірювання відстані між об'єктами.

У нашому випадку ми використовуємо значення номінальних ознак-кольорів – в якості координат об'єкта.

Мета роботи – створити алгоритм і програмне забезпечення, в якому, відповідно до певного кольору, камера знаходить відповідний об'єкт,

розпізнає його і посилає сигнали в систему управління і переміщається навколо об'єкта; все йде в автономному режимі.

4.1 Реалізація алгоритму розпізнавання образів

У нашій роботі СТЗ використовує технологію ідентифікації символу / знаку / об'єкту за його кольором (за умови, що камера багатокольорова).

Для реалізації алгоритму було створено лабораторний стенд, до складу якого входить наступне обладнання:

- мікроконтролер Raspberry Pi 2B;
- мікроконтролер Arduino Mega;
- кроковий двигун з драйвером ULN2003;
- веб-камера Logitech;
- програмні середовища openCV, Python.

Підхід до вибору об'єктів на зображеннях з розподілу кольорової гами в основному досить простий, але його можна використовувати в ряді випадків.

Іноді цей підхід використовується в якості основного етапу обробки зображень для вирішення більш складних завдань, таких як розпізнавання осіб.

Такий підхід може застосовуватися і в системах трекінгу, якщо колір об'єкта відрізняється від фону.

Рішення завдання розбивається на кілька етапів:

- виділення пікселів, відповідних заданому об'єкту;
- виділення контурами знайдені об'єкти;
- знаходження контуру об'єкту;
- побудова прямокутника, до якого потрапляють усі точки контуру об'єкту.

За допомогою програмного середовища OpenCV та мови програмування Python ми створюємо відповідне програмне забезпечення.

OpenCV – (Open Computer Vision) – бібліотека комп'ютерного зору з відкритим вихідним кодом, який надає ряд типів даних і числових алгоритмів для обробки зображень з використанням алгоритмів комп'ютерного зору. Реалізована на C/C++.

Для виділення та пошуку об'єктів за кольором використовується колірна модель HSV, що описує простір кольору на основі трьох колірних характеристик: колірному тоні (Hue), насиченості (Saturation) і яскравості (Brightness, Value). Даний простір кольору є нелінійним. Простір HSV (HSI) можна вважати ідеальним інструментом для побудови алгоритмів обробки зображень, оскільки він заснований на природному і інтуїтивно зрозумілому описі кольору.

Отримати кольоровий тон для моделі HSV можна із простору RGB.

Лістинг коду програми для реалізації алгоритму розпізнавання об'єкту за кольором:

```
import cv2
import numpy as np
import os
#####
#####
def main():
    capWebcam = cv2.VideoCapture(0) #початок захвату відеозображення об'єкту та зв'язок з вебкамерою
    print «defaultresolution = « + str(capWebcam.get(cv2.CAP_PROP_FRAME_WIDTH)) + «x» + str(capWebcam.get(cv2.CAP_PROP_FRAME_HEIGHT)) #відображення оригінального розширення камери
    capWebcam.set(cv2.CAP_PROP_FRAME_WIDTH, 320.0) #трансформація розширення камери
    capWebcam.set(cv2.CAP_PROP_FRAME_HEIGHT, 240.0) #зменшення значення розширення камери для кращої швидкодії
    if capWebcam.isOpened() == False: #перевірка підключення камери
        print «error: cap Webcam not accessed successfully\n\n»
        os.system(«pause»)
        return
    # endif
    while cv2.waitKey(1) != 27 and capWebcam.isOpened(): #програма працює доти, доки не нажати кнопку ESC або не вимкнути камеру
        if not blnFrameReadSuccessfully or imgOriginal is None: #перевірка на наявність кадру з камери
            print «error: frame not read from webcam\n\n»
            os.system(«pause»)
            break
        # endif
#трансформація картинки з BGR в HSV простір
```

```

imgHSV = cv2.cvtColor(imgOriginal, cv2.COLOR_BGR2HSV)
#встановлення діапазону заданого кольору на зображенні
imgThreshLow = cv2.inRange(imgHSV, np.array([0, 170, 120]), np.array([20, 240, 255]))
imgThreshHigh = cv2.inRange(imgHSV, np.array([20, 70, 170]), np.array([40, 170, 255]))

imgThresh = cv2.add(imgThreshLow, imgThreshHigh)
#створення маски зображення заданого кольору
imgThresh = cv2.cvtColor(imgThresh, cv2.COLOR_GRAY2RGB)
#згладжування об'єкта методом Гаусса
imgThresh = cv2.GaussianBlur(imgThresh, (5, 5), 2)
#морфологічна операція по розширенню об'єкту для заповнення випадкового простору в об'єкті
imgThresh = cv2.dilate(imgThresh, np.ones((5,5),np.uint8))
#трансформація картинки в сірі відтінки
imgThresh = cv2.cvtColor(imgThresh, cv2.COLOR_BGR2GRAY)
#бінарізація зображення
_, imgThresh = cv2.threshold(imgThresh, 130, 255, cv2.THRESH_BINARY)
#морфологічні операції для покращення цільності об'єкта на зображенні
st1 = cv2.getStructuringElement(cv2.MORPH_RECT, (21, 21), (10, 10))
st2 = cv2.getStructuringElement(cv2.MORPH_RECT, (11, 11), (5, 5))
imgThresh = cv2.morphologyEx(imgThresh, cv2.MORPH_CLOSE, st1)
imgThresh = cv2.morphologyEx(imgThresh, cv2.MORPH_OPEN, st2)
#знаходження контурів об'єкта
img, contours, hierarchy = cv2.findContours(imgThresh, 2, 4)
#формування прямокутника (зеленого кольору) навколо об'єкта
cnt = contours[0]
a,b,w,h = cv2.boundingRect(cnt)
cv2.rectangle(imgOriginal,(a,b),(a+w,b+h),(0,255,0),2)
#знаходження центру об'єкта
x = (a+w)/2
y = (b+h)/2
#вивід картини на екран монітору
cv2.namedWindow(«imgOriginal», cv2.WINDOW_AUTOSIZE)
cv2.namedWindow(«imgThresh», cv2.WINDOW_AUTOSIZE)
cv2.imshow(«imgOriginal», imgOriginal)
cv2.imshow(«imgThresh», imgThresh)
# endwhile
cv2.destroyAllWindows() #очищення пам'яті від картини
return #знову на початок циклу
#####
#####
if __name__ == «__main__»:
main()

```

На рис. 4.1 та 4.2 зображено об'єкт та результат розпізнавання об'єкту за помаранчевим кольором.

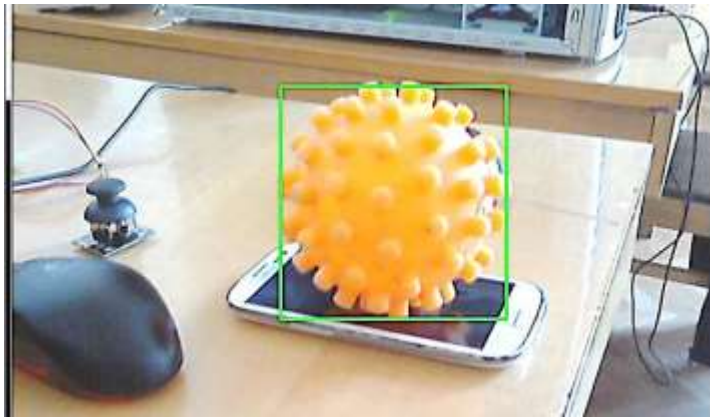


Рисунок 4.1 – Заданий об'єкт

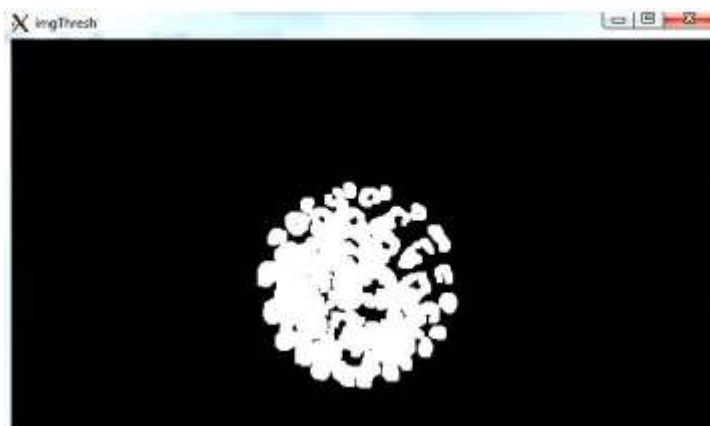


Рисунок 4.2 – Результат розпізнавання об'єкту

4.2 Аналіз алгоритму розпізнавання образів

Для представленої алгоритму використовувався метод просторової обробки зображень. Просторова область – це кількість пікселів, що складають зображення. Просторові методи – це процедури, які безпосередньо працюють зі значеннями пікселів зображення. Процеси просторової обробки описуються наступним рівнянням:

$$g(x,y) = T[f(x,y)]$$

де $f(x, y)$ – вхідне зображення, $g(x, y)$ – оброблене зображення, а T – оператор над f , визначений в деякому околі точки (x, y) .

Головний підхід у визначенні околу навколо точки (x, y) полягає у використанні квадратної або прямокутної області – чи є використання квадратної або прямокутної області – підмножина зображення центрується в одній точці (x, y) . Центр даної підмножини пересувається від пікселя до пікселя. Оператор T виконується в кожній точці (x, y) даючи в результаті вихідне значення g для даної точки. Процес використовує лише пікселі в області зображення, обмеженій зазначеною площею.

Збільшення околу дає методу набагато більшу гнучкість. Принцип полягає в тому, що для знаходження значення g в певній точці (x, y) , використовуються значення функції f усередині деякого околу заздалегідь заданої форми, що оточує точку (x, y) . Один з основних підходів у такій постановці задачі базується на використанні маски (фільтра) – невеликий двовимірний масив, значення елементів якого визначають суть процесу обробки, наприклад, різкість зображення. У графічному вигляді алгоритм буде виглядати наступним чином (рис. 3.6).

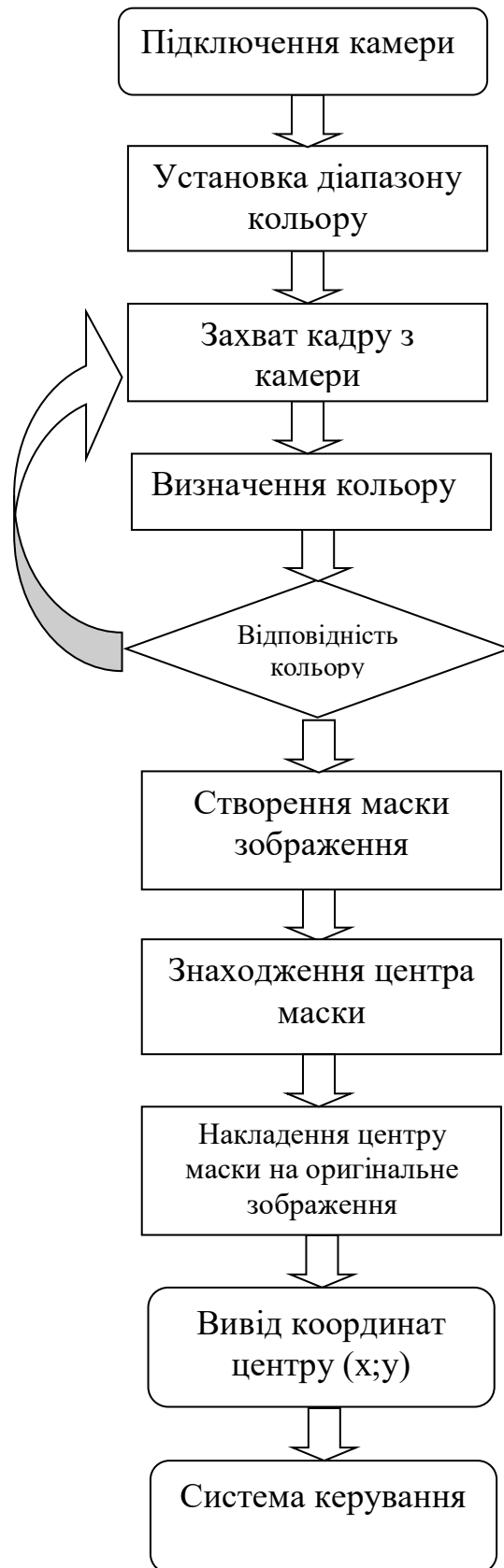


Рисунок 4.3 – Алгоритм програми

Для системи технічного зору з технологією розпізнавання за кольором були розроблені, реалізовані та протестовані алгоритм та програмне забезпечення.

Результати роботи можуть будуть використані на безпілотних летальних апаратах, типу квадрокоптер, з метою автономного керування пристроєм, на виробництві та при конструюванні роботів. Також даним алгоритмом можливо користуватися на виробництві для контролю якості продукції та сировини.

Знайдеться використання у сільському господарстві для контролю кількості поголів'я ВРХ, кіз, баранів, кролів, птиці коли поголів'я багато людині важко самотужки зробити підрахунки. Також така система може розпізнати, якщо окрім худоби на фермі є хижі тварини (лисиці, куниці, ласки) або шкідники (миші, пацюки), і допоможе вчасно зреагувати і врятувати фермерське господарство від значних втрат поголів'я та кормів. Відео сигнал може подаватися як з стаціонарних відеокамер так і з камер безпілотних літальних апаратів.

Контроль початку збору врожаю та врожайності теж стане легшим і простішим. Фермер зможе прогнозувати врожай, прорахувати скільки працівників завтра потрібно викликати на зміну і на які кількість врожаю можуть розраховувати в найближчий час оптові покупці.

Використання данного алгоритму та програмного забезпечення допоможе вийти сільському господарству на новій рівень.

ВИСНОВКИ

В роботі розглянуті системи технічного зору. СТЗ в даний час є одним з головних засобів розвитку автоматичних систем управління рухом в умовах, коли обсяг апіорної інформації не достатній і для вирішення завдань керування необхідно аналізувати навколишнє середовище в режимі реального часу. Також був зроблений ретельний аналіз сучасної ситуації в системах розпізнавання образів.

В першому розділі атестаційної роботи проведений огляд існуючих методів розпізнавання образів. Розглянуті основні завдання побудови систем розпізнавання. Візначили що двома суттєвими проблемами в розпізнаванні образів є наступні: які вхідні дані можна вважати доречними і яка попередня обробка вихідних даних (які зазвичай відрізняються надзвичайною надмірністю) призводить до отримання властивостей або ознак, що дійсно дозволяють проводити класифікацію. Так чи інакше, щодо ознак використовуються апіорні знання, інтуїція, метод спроб і помилок, досвід. Виділили 2 великі сімейства способів прийняття рішення: методи статистичної теорії рішень та структурні (лінгвістичні, синтаксичні) методи.

В другому розділі розглянули штучні нейронні мережі. Нейронні мережі – один з напрямків досліджень в області штучного інтелекту(ШІ), засноване на бажанні імітувати нервову систему людини. В тому числі її (нервової системи) здатність виправляти помилки і вчитися самостійно. Все це, хоча і дещо грубо, повинно дозволяти імітувати роботу людського мозку. Вони використовуються для прогнозування, розпізнавання образів, машинного перекладу, розпізнавання звуку тощо. проведений огляд існуючих методів розпізнавання образів.

В третьому розділі описано систему ідентифікації, наведено структурну схему системи розпізнавання. Комп'ютерна система є основним блоком усієї системи розпізнавання, яка приймає рішення про належність об'єкта до певного класу. У нашому випадку буде лише два класи, а також клас

«відмова від визнання», що означає, що прийняття рішення на користь того чи іншого класу з тих чи інших причин неможливе.

Оптична система визначає багато важливих характеристик системи розпізнавання зображень, тому її слід враховувати при вивченні роботи системи розпізнавання зображень. Описано матрицю фотоелементів та систему попередньої обробки зображень. Для обробки на комп'ютерах вхідні дані спочатку оцифровують, а потім представляють у більш компактному вигляді (стиск даних) за допомогою схем кодування або апроксимації. Якщо перетворення Фур'є функції (одновимірної чи двовимірної) відмінне від нуля в обмеженій області B частотного простору, то цю функцію можна точно відновити за її значеннями в кінцевому наборі опорних точок (теорема Котельникова).

На практиці перетворення Фур'є за межами області B лише близьке до нуля. Тому в результаті відновлення вибіркової функції буде отримано лише її наближення. В інших методах вибірки аналізована функція інтерполюється між опорними точками за допомогою простих типів функцій (наприклад, константних, лінійних або поліномів вищого порядку). Інший метод апроксимації – квантування. У цьому випадку фактичне значення функції апроксимується за допомогою кінцевого набору значень, які називаються рівнями квантування. Найпростіший спосіб – використовувати рівні квантування з рівномірним розташуванням.

В четвертому розділі для технічної системи зору з технологією розпізнавання кольору були розроблені, впроваджені і протестовані алгоритми і програмне забезпечення. Були розроблені та обґрунтовані базові принципи побудови даного типу програмних засобів.

Галузь технічного зору та застосування його до навігації є достатньо розвиненою та поступово втілюваною в технічні рішення. В ході роботи виявлено що процедура пошуку рішення задачі за допомогою мережі, що пройшла навчання, виявляється більш гнучкою, ніж процес програмування, оскільки штучні нейронні мережі ШНМ може підвищувати точність

результатів у міру накопичення нею досвіду і адаптуватися до змін, що відбуваються. Це пов'язано з тим, що сучасні ШНМ характеризуються такими властивостями, як: здатність до узагальнення і здатність до абстрагування. В результаті нейронні мережі перевершують послідовні машини у вирішенні тих же завдань, в яких людина перевершує обчислювальну техніку. Крім того, рішення завдання розпізнавання у великій мірі залежить від вибору вихідного опису об'єктів, тобто простору ознак. При вдалому виборі завдання розпізнавання може виявитися тривіальною і, навпаки, невдало обраний простір ознак здатне привести до дуже складної обробки інформації, або взагалі до відсутності рішення.

ПЕРЕЛІК ПОСИЛАНЬ

1. ДСТУ 3008–15. Документація. Звіти у сфері науки і техніки. Структура і правила оформлення [Текст]. – Введ. 2015-06-22. – К.: Держстандарт України, 2017. – 29 с.
2. Барский А. Б. Нейронные сети: распознавание, управление, принятие решений. – М.: Финансы и статистика, 2004. – 176 с.
3. Гридин В.Н., Титов В.С., Труфанов М.И. Адаптивные системы технического зрения, — СПб.: Наука, 2009. — 442 с.
4. Горбань А.Н., Дунин-Барковский В.Л., Кирдин А.Н. и др. Нейроинформатика. <http://oasis.peterlink.ru/~dap/neng/nlinks/book2/gl1.htm>
5. Городничий Дмитро Олегович Дослідження та розробка високопродуктивних повних нейромереж. 01.05.03 – Математичне і програмне забезпечення обчислювальних машин і систем. Автореферат дисертації на здобуття наукового ступеня кандидата фізико-математичних наук Київ -1997.
6. Данилов Е. П., Луцив В. Р.. Нейронные сети: современное состояние и перспективы // Оптико механическая промышленность, №4, 1991г., с20-32.
7. Заенцев И.В. Нейронные сети: основные модели. Учебное пособие к курсу “Нейронные сети” для студентов 5 курса магистратуры к. электроники физического ф-та Воронежского университета. (<http://neurnews.iu4.bmstu.ru/book/index.htm>)
8. Каллан Р. Основные концепции нейронных сетей / пер. с англ. А.Г. Сивака. – М.: Издательский дом «Вильямс», 2001. – 288 с.
9. Крилов В. С. Комп'ютерний зір: інноваційний віртуальний лабораторний практикум / В. С. Крилов – Науковий часопис НПУ імені М. П. Драгоманова. Серія 2 : Комп'ютерно-орієнтовані системи навчання. – 2012. – № 13. – С. 181 – 185.

10. Любунь З., Мандзій В.. Умови розпізнавання корельованих зображень у нейромережі Гопфілда. Збірник наукових праць. Теоретична електротехніка // Вип. 57, 2004
11. Милов В.Р. Гибридные алгоритмы обучения RBF-сетей // Нейроинформатика. – 2006. – Часть 2. – С. 195 – 202.
12. Миркес Е. М.. Учебное пособие по курсу нейроинформатика. Красноярск 2002. <http://www.softcraft.ru/neuro/index.shtml>.
13. Навігація мобільних наземних роботів у недетермінованих середовищах. Сергієнко О.Ю., Карташов В.М., Колендовська М.М. Харків: ХНУРЕ, 2020. 297 с.
14. Національна академія наук України Інститут проблем математичних машин і систем РІЗНИК Олександр Михайлович УДК 621.8:681.5 Неітеративні методи навчання нейромереж. 01.05.03- Математичне і програмне забезпечення обчислювальних машин і систем Автореферат дисертації на здобуття наукового ступеня доктора технічних наук Київ-1997
15. Пишем свою нейросеть: пошаговое руководство [Электронный ресурс] / Режим доступа : [www / URL: https://proglib.io/p/neural-nets-guide](http://www.proglib.io/p/neural-nets-guide) – Загл. с экрана.
16. Системы технического зрения (принципиальные основы, аппаратное и математическое обеспечение) / А.Н. Писаревский [и др.]; Под общ. ред. А.Н. Писаревского, А.Ф. Чернявского. Л.: Машиностроение. (Ленингр.отд-ние), 1988. 424 с.: ил.
17. Сомполинский Х. Статистическая механика нейронных сетей. Физика за рубежом- серия А(исследования).-М: Мир. 1991.
18. Сравнение библиотек компьютерного зрения для разработки систем охранного видеонаблюдения с возможностью видеоаналитики [Электронный ресурс] / Режим доступа : [www / URL: https://cyberleninka.ru/article/n/sravneniebibliotek-kompyuternogo-zreniya-](http://www.cyberleninka.ru/article/n/sravneniebibliotek-kompyuternogo-zreniya)

- dlya-razrabotki-sistem-ohrannogovideonablyudeniya-s-vozmozhnostyu-videoanalitiki/viewer – Загл. с экрана.
19. Радченко, Я.О. Связь нейронных сетей в автоматизированной системе управления [Текст] : тез. допв. науч.-практ. конф. (квіт. 2019) / конф.: Автоматизовані системи та комп'ютеризовані технології радіоелектронного приладобудовання / від. ред. Я. О. Радченко. – Х : 2019. – С. 23 – 24.
 20. Хайкин, Саймон. Нейронные сети: полный курс, 2-е издание. : Пер. с англ. – М.: Издательский дом «Вильямс», 2006. – 1104 с.
 21. Ярова Ю.Д Використання систем технічного зору в виробництві і системах розпізнавання образів [Текст] : тез. допв. науч. – практ. конф. (квіт. 2021) / конф.: Інформаційні радіо технології та технічний захист інформації / від. ред. Ю.Д Ярова – Х: 2021. – С. 77 – 78.
 22. Cohen M.A., Grossberg S.G. 1983. Absolute stability of global pattern formation and parallel memory storage by competitive neural networks. IEEE Transactions on System, Man and Cybernetics 13:815-26
 23. Loubyun Z. M., Mandzij V.M.. To the some properties of Hofield's neuronet. 26 Open scientific and technical conference of young scientist and specialists of the Karpenko Physico-Mechanical Institute of NAS of Ukraine. Lviv-2001.p.44.
 24. https://kipia.ru/upload/iblock/d78/Auton_sist_tech_zrn.pdf
 25. Watanabe H., Dettloff W., Yount E. A VLSI Fuzzy Logic Inference Engine for Real-time Process Control // IEEE Journal of Solid State Circuits, 1990. – V.25, N.2. P.376 – 382.
 26. A Comparative Example Between The Use Of Pca And Mds For Image Classification / Hernandez, W., Mendez, A., Flor-Unda, O., Camejo, I.M., Kolendovska, M. // IEEE International Symposium on Industrial Electronics, 29th IEEE International Symposium on Industrial Electronics, ISIE 2020; Delft; Netherlands; 17 June 2020 до 19 June 2020; Volume 2020-June, June 2020, № 9152565, Pages 1353-1358

27. Algorithm For Generating Refined Frequency Estimates In Atmospheric Radio Sounding Systems / Kartashov V., Hernandez W., Hernandez-Balbuena D., M. Kolendovska, Konovalenko O., Melnyk V. // IEEE International Symposium on Industrial Electronics, 29th IEEE International Symposium on Industrial Electronics, ISIE 2020; Delft; Netherlands; 17 June 2020 до 19 June 2020; Volume 2020-June, June 2020, № 9152562, Pages 79-82
28. Application of Fast Frequency Shift Measurement Method for INS in Navigation of Drones / D. Avalos-Gonzalez, D.H. Balbuena, V. Tyrsa, V.M. Kartashov, M. Kolendovska, S. Sheiko, O. Sergiyenko, V. Melnyk, F.N. Murrieta-Rico // IECON 2018 – 44th Annual Conference of the IEEE Industrial Electronics Society. – P. 3159–3164.
29. Avalos-Gonzalez, D., Sergiyenko, O., Hernandez-Balbuena, D., Tyrsa, V., Kartashov V.M., V., Rivas-Lopes, M., Murrieta-Rico, F.N. Constraints definition and application optimization based on geometric analysis of the frequency measurement method by pulse coincidence // Measurement: Journal of the International Measurement Confederation (USA). 2018, V.126. P. 184-193.
30. Book “Control and Signal Processing Applications for Mobile and Aerial Robotic Systems”, Hardback - Advances in Computational Intelligence and Robotics English. Edited by Oleg Sergiyenko, Moises Rivas-Lopez, Wendy Flores-Fuentes, Julio Cesar Rodríguez-Quñonez, Lars Lindner. Editorial IGI Global, Hershey, United States, January 2020, 340 páginas. ISBN10 152259924X, ISBN13 9781522599241
31. Cesar Sepulveda-Valdez ; Oleg Sergiyenko ; Vera Tyrsa ; Wendy Flores-Fuentes ; Julio César Rodríguez-Quñonez ; Fabian Natanael Murrieta-Rico ; Jesús Elías Miranda-Vega ; Paolo Mercorelli ; Marina Kolendovska. "Geometric analysis of a laser scanner functioning based on dynamic triangulation," 2020 IEEE 29th International Symposium on Industrial Electronics (ISIE), Delft, Netherlands, 17-19 of June 2020, pp. 1398-1403,

doi: 10.1109/ISIE45063.2020.9152268.

<https://ieeexplore.ieee.org/abstract/document/9152268>

32. Cuauhtémoc Mariscal-García; Wendy Flores-Fuentes; Daniel Hernández-Balbuena; Julio C. Rodríguez-Quiñonez ; Oleg Sergiyenko. "Classification of Vehicle Images through Deep Neural Networks for Camera View Position Selection," 2020 IEEE 29th International Symposium on Industrial Electronics (ISIE), Delft, Netherlands, 17-19 of June 2020, pp. 1376-1380, doi: 10.1109/ISIE45063.2020.9152440.

<https://ieeexplore.ieee.org/abstract/document/9152440>

33. Developing and Applying Optoelectronics in Machine Vision/ O. Sergiyenko, J.C. Rodriguez-Quiñonez, IGI Global, 2016; 341p.

34. Experimental estimation of direction finding to unmanned air vehicles algorithms efficiency by their acoustic emission, /Oleynikov, V., Zubkov, O., Kartashov, V., ...Sheiko, S., Babkin, S.//2019 IEEE International Scientific-Practical Conference: Problems of Infocommunications Science and Technology, PIC S and T 2019 - Proceedings, 2019, стр. 175-178, 9061337

35. Features of acoustic noise of small unmanned aerial vehicles / Semenets, V.V., Kartashov, V.M., Leonidov, V.I. //Telecommunications and Radio Engineering (English translation of *Elektrosvyaz and Radiotekhnika*), 2020, 79(11), стр. 985-995

36. Geometric Analysis Of A Laser Scanner Functioning Based On Dynamic Triangulation /Sepulveda-Valdez, C., Sergiyenko, O., Tyrsa, V, Mercorelli, P., Kolendovska, M.// IEEE International Symposium on Industrial Electronics, 29th IEEE International Symposium on Industrial Electronics, ISIE 2020; Delft; Netherlands; 17 June 2020 до 19 June 2020; Volume 2020-June, June 2020, № 9152268, Pages 1398-1403

<https://ieeexplore.ieee.org/abstract/document/9152255>

<https://ieeexplore.ieee.org/document/9161870>

- 37.I. Y. A. Corpus, L.Lindner, O.Sergiyenko. "Transimpedance Amplifier for Laser Scanning System Range Extension," 2020 IEEE 29th International Symposium on Industrial Electronics (ISIE), Delft, Netherlands, 17-19 of June 2020, pp. 1421-1426, doi: 10.1109/ISIE45063.2020.9152487.
<https://ieeexplore.ieee.org/abstract/document/9152487>
- 38.Ivanov, M., Sergiyenko, O., Mercorelli, P., Hernandez, W.c, Rodriguez Quinonez, J.C.d, Katashov V., Kolendovska, M., Iryna, T. Effective informational entropy reduction in multi-robot systems based on real-time TVS. IEEE International Symposium on Industrial Electronics, 2019-June,8781209, c. 1162-1167.
- 39.Jonathan J. Sanchez-Castro ; Julio C. Rodríguez-Quiñonez ; Luis R. Ramírez-Hernández ; Guillermo Galaviz ; Daniel Hernández-Balbuena ; Gabriel Trujillo-Hernández ; Wendy Flores-Fuentes ; Paolo Mercorelli ; Wilmar Hernández-Perdomo ; Oleg Sergiyenko ; Félix Fernando González-Navarro. "A Lean Convolutional Neural Network for Vehicle Classification," 2020 IEEE 29th International Symposium on Industrial Electronics (ISIE), Delft, Netherlands, 17-19 of June 2020, pp. 1365-1369, doi: 10.1109/ISIE45063.2020.9152274.
<https://ieeexplore.ieee.org/abstract/document/9152274>
- 40.Lindner, L., Sergiyenko, O., Rivas-López, M., (...), Gurko, A., Kartashov, V.M. Machine vision system for UAV navigation; IEEE, 2016 International Conference on Electrical Systems for Aircraft, Railway, Ship Propulsion and Road Vehicles and International Transportation Electrification Conference, ESARS-ITEC, 2016; pp.1–6. DOI: 10.1109/ESARS-ITEC.2016.7841356.
- 41.M. Ivanov, O. Sergiyenko, V. Tyrsa, P. Mercorelli, V. Kartashov, W. Hernandez, S. Sheiko, M. Kolendovska. Individual scans fusion in virtual knowledge base for navigation of mobile robotic group with 3D TVS // Proceedings of 44th Annual Conference of IEEE Industrial Electronics Society (IECON).. -2018. – Washington DC, USA. -S. 3187-3192. . ISBN 978-1-5090-6683-4/18/.

42. Murrieta-Rico, F.N., Petranovskii, V., Galvan, D.H., Sergiyenko, O., Yocupicio-Gaxiola, R.I., De Dios Sanchez-Lopez, J. Phase effect in frequency measurements of a quartz crystal using the pulse coincidence principle. 2020 IEEE 29th International Symposium on Industrial Electronics (ISIE), Delft, Netherlands, 17-19 of June 2020, pp. 185-190, 9152255, DOI: 10.1109/ISIE45063.2020.9152255
43. Oleksandr Sotnikov, Vladimir Kartashov, Oleksandr Tymochko, Oleg Sergiyenko, Vera Tyrza, Paolo Mercorelli, Wendy Flores-Fuentes. Methods for Ensuring the Accuracy of Radiometric and Optoelectronic Navigation Systems of Flying Robots in a Developed Infrastructure. Chapter 16// Machine Vision and Navigation; Springer, Cham. pp.537–578. Editors: Sergiyenko, Oleg, Flores-Fuentes, Wendy, Mercorelli, Paolo. DOI: 10.1007/978-3-030-22587-2_16.
44. Optical detection of unmanned air vehicles on a video stream in a real-time/Kartashov, V., Oleynikov, V., Zubkov, O., Sheiko, S.// 2019 International Conference on Information and Telecommunication Technologies and Radio Electronics, UkrMiCo 2019 - Proceedings, 2019, 9165362/
45. Principles Of Construction And Assessment Of Technical Characteristics Of Multi-Frequency Atmospheric Sodar In The Humidity Measurement Mode / Kartashov, V.M., Sidorov, G.I., Sheiko, S.A., Kolendovskaya, M.M., Sergienko, O.Yu. // Telecommunications And Radio Engineering (English Translation Of Elektrosvyaz And Radiotekhnika), 2020, ISSN Print: 0040-2508, ISSN Online: 1943-6009, DOI: 10.1615/TelecomRadEng.v79.i4.50, p. 323-333/
46. Research Of The Uncertainty Of Measurement Frequencies And Definitions Of The Frequency Signal In The Waveguide With Respect To Power / Semenets, V.Zakharov, I. Serhiienko, M., Kartashov, V.M., Kolendovska, M., Hernandez, W., Hipolito, J.I.N., Tyrza, V.// 45th Annual Conference of the IEEE Industrial Electronics Society, IECON 2019; Lisbon Congress

- CenterLisbon; Portugal; 14 October 2019 до 17 October 2019; CFP19IEC-ART; Код 155980, Volume 2019-October, October 2019, № 8927203, Pages 4674-4679
47. Spatial-Temporal Processing Of Acoustic Signals Of Unmanned Aerial Vehicles /Kartashov V.M., Oleinikov V.N., Zubkov O.V., Sheiko S.A., Kolendovska M.M.// Telecommunications And Radio Engineering (English Translation Of Elektrosvyaz And Radiotekhnika), 2020, ISSN Print: 0040-2508, ISSN Online: 1943-6009, DOI: 10.1615/Telecomradeng.v79.i9.40, p. 769-780
48. Stereoscopic Vision Systems In Machine Vision, Models, And Applications (Book Chapter)/ Ramírez-Hernández, L.R., Rodríguez-Quiñonez, J.C., Castro-Toscano, M.J., Kolendovska, M., Murrieta-Rico, F.N.// Machine Vision And Navigation, 2019 Machine Vision and Navigation 30 September 2019, Pages 241-265
49. Strelkova T., Kartashov V., Lytyuga A., Strelkov A. Theoretical Methods of Images Processing in Optoelectronic Systems. Chapter 16. // Biometrics: Concepts, Methodologies, Tools, and Applications; Oleg Sergiyenko and Julio C. Rodriguez-Quiñonez. (341p.), IGI Global, 2017; pp. 361-381. DOI: 10.4018/978-1-5225-0983-7.ch016.
50. Strelkova T., Kartashov V., Lytyuga A., Strelkov A. Theoretical Methods of Images Processing in Optoelectronic Systems. Chapter 6// Developing and Applying Optoelectronics in Machine Vision; Oleg Sergiyenko and Julio C. Rodriguez-Quiñonez. (341p.) – USA, Herhey, IGI Global, 2016; pp.180-205.
51. Sytnik O., Kartashov V. Methods and Algorithms for Technical Vision in Radar Introspection. Chapter 13// Optoelectronics in Machine Vision-Based Theories and Applications. IGI Global, 2019; pp. 373-391.
52. The Use of Factorization and Multimode Parametric Spectra in Estimating Frequency and Spectral Parameters of Signal/Semenets, V., Kartashov, V., Sergiyenko, O., ...Rodriguez-Quinonez, J.C., Flores-Fuentes, W.//IEEE

- International Symposium on Industrial Electronics, 2020, 2020-June, p. 215-219
53. Unda, O.F., Hernandez, W., Vargas, O., Mendez, A., Sergiyenko, O., Tyrsa, V. Construction of a robotic platform of differential type for first-year students of electronic engineering, 2020 International Symposium on Power Electronics, Electrical Drives, Automation and Motion, SPEEDAM 2020, 24-26 de junio de 2020, Sorrento, Italia, pp. 538-543, 9161870, DOI: 10.1109/SPEEDAM48782.2020.9161870
54. Use of Acoustic Signature for Detection, Recognition and Direction Finding of Small Unmanned Aerial Vehicles/Kartashov, V., Oleynikov, V., Koryttsev, I., ...Babkin, S., Selieznov, I.//Proceedings - 15th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering, TCSET 2020, 2020, p. 377-380/
55. V. Semenets; Vladimir Kartashov ; Oleg Sergiyenko; Vyacheslav Tikhonov ; Paolo Mercorelli ; Sergiy Sheiko ; Nataliya Chmelarova. "The Use of Factorization and Multimode Parametric Spectra in Estimating Frequency and Spectral Parameters of Signal," 2020 IEEE 29th International Symposium on Industrial Electronics (ISIE), Delft, Netherlands, 17-19 of June 2020, pp. 215-219, doi: 10.1109/ISIE45063.2020.9152238.
<https://ieeexplore.ieee.org/abstract/document/9152238>
56. Wilmar Hernandez ; Alfredo Mendez ; Omar Flor-Unda ; Vicente Gonzalez-Posada ; Jose Luis Jimenez ; Oleg Sergiyenko ; Julio C. Rodriguez-Quiñonez ; Mykhailo Ivanov ; Ivan Menes Camejo ; Marina Kolendovska. "A comparative example between the use of PCA and MDS for image classification," 2020 IEEE 29th International Symposium on Industrial Electronics (ISIE), Delft, Netherlands, 17-19 of June 2020, pp. 1353-1358, doi: 10.1109/ISIE45063.2020.9152565.
<https://ieeexplore.ieee.org/abstract/document/9152565>

- 57.Карташов В.М. и др. Обработка сигналов в радиоэлектронных системах дистанционного мониторинга атмосферы; Харьков: ХНУРЭ, 2014. 312 с.
- 58.Карташов В.М., Олейников В.Н., Колендовская М.М., Тимошенко Л.П., Капушта А.И., Рыбников Н.В. Комплексирование изображений при обнаружении беспилотных летательных аппаратов// Радиотехника. (Харьков). 2020. Вып. 201; С.120-129.
- 59.Карташов В.М., Посошенко В.А., Цехмистро Р.И., Тимошенко Л.П., Колендовская М.М. Методы ориентации, навигации и контроля мобильных робототехнических платформ// Радиотехника. (Харьков). 2019. Вып. 199. С. 38-44.
- 60.Ситнік О.В., Карташов В.М. Радіотехнічні системи. Навч. посібник. Х.: Сміт, 2009. 448 с.