

Міністерство освіти і науки України  
Харківський національний університет радіоелектроніки

Факультет Комп'ютерних наук  
(повна назва)

Кафедра Інформаційних управляючих систем  
(повна назва)

## КВАЛІФІКАЦІЙНА РОБОТА

### Пояснювальна записка

рівень вищої освіти другий (магістерський)  
(рівень вищої освіти)

Дослідження методів та технологій реалізації  
систем доповненої реальності  
(тема)

Виконала: студент 2 курсу, групи ІУСТм-20-1

Акулова А.О.  
(прізвище, ініціали)

Спеціальність 122 Комп'ютерні науки  
(код і повна назва спеціальності)

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

Освітня програма Інформаційні управляючі системи та технології  
(повна назва освітньої програми)

Керівник доц. Кудрявцева М.С.  
(посада, прізвище, ініціали)

Допускається до захисту  
зав. кафедри

Петров К.Е.  
(прізвище, ініціали)

2021 р.

Харківський національний університет радіоелектроніки

Факультет Комп'ютерних наук  
(повна назва)

Кафедра Інформаційних управляючих систем  
(повна назва)

Рівень вищої освіти другий (магістерський)

Спеціальність 122 Комп'ютерні науки  
(код і повна назва)

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

Освітня програма Інформаційні управляючі системи та технології  
(повна назва освітньої програми)

ЗАТВЕРДЖУЮ:

Зав. кафедри \_\_\_\_\_  
(підпис)

« \_\_\_\_\_ » \_\_\_\_\_ 2021 р.

## ЗАВДАННЯ

### НА КВАЛІФІКАЦІЙНУ РОБОТУ

студентові Акуловій Анні Олегівні  
(прізвище, ім'я, по батькові)

1. Тема роботи Дослідження методів та технологій реалізації систем доповненої реальності

затверджена наказом по університету від "05" листопада 2021 р. № 1645 Ст

2. Термін здачі студентом закінченої роботи 6 грудня 2020 р.

3. Вихідні дані до роботи Науково-технічні публікації, дані статей, експериментальних досліджень по технологіям Augmented Reality та Virtual Reality

4. Перелік питань, що потрібно опрацювати в роботі

1. Огляд і аналіз існуючих технологій, методів і систем віртуальної реальності та доповненої реальності

2. Дослідження технологій і методів реалізації систем доповненої реальності

3. Створення нейронної мережі для реалізації системи доповненої реальності

4. Практична реалізація

5. Перелік графічного матеріалу із зазначенням обов'язкових рисунків

Рисунок 1 – Схема архітектури згорткової нейронної мережі

Рисунок 2 – Схема архітектури рекурентної нейронної мережі

Рисунок 3 – Структурна схема нейронної мережі RNN-CNN для реалізації системи доповненої реальності

Рисунок 4 – Структура застосування доповненої реальності

Рисунок 5 – Приклад класифікації об'єктів за допомогою запропонованих нейронних мереж

Рисунок 6–12 Приклади екранних форм

## КАЛЕНДАРНИЙ ПЛАН

Номер	Назва етапів роботи	Термін виконання етапів роботи	Примітка
1	Отримання завдання на кваліфікаційну роботу	05.11.2021	
2	Аналіз предметної галузі та постановка задачі	06.11.2021 – 14.11.2021	
3	Огляд і аналіз існуючих технологій, методів і систем доповненої реальності та віртуальної реальності	15.11.2021 – 17.11.2021	
4	Дослідження технологій і методів реалізації систем доповненої реальності	18.11.2021 – 20.11.2021	
5	Створення нейронної мережі для реалізації системи доповненої реальності	21.11.2021 – 22.11.2021	
6	Практична реалізація	23.11.2021 – 24.11.2021	
7	Підготовка пояснювальної записки та графічного матеріалу	25.11.2021 – 30.11.2021	
8	Надання пояснювальної записки на перевірку керівнику	01.12.2021 – 05.12.2021	
9	Захист перед ЕК	09.12.2021	

Дата видачі завдання 05 листопада 2021 року

Студент \_\_\_\_\_  
(підпис)

Керівник роботи \_\_\_\_\_  
(підпис)

доцент кафедри ІУС Кудрявцева М.С.  
(посада, прізвище, ім'я, по батькові)

## РЕФЕРАТ

Пояснювальна записка містить 103 сторінки, 28 рисунків, 40 джерел.

АЛГОРИТМ, ВІРТУАЛЬНА РЕАЛЬНІСТЬ, ДОПОВНЕНА РЕАЛЬНІСТЬ, ЗМІШАНА РЕАЛЬНІСТЬ, МАРКЕР, СИСТЕМА ДОПОВНЕНОЇ РЕАЛЬНОСТІ, ШТУЧНА НЕЙРОННА МЕРЕЖА

Об'єктом дослідження є система доповненої реальності, як різновид систем віртуальної реальності.

Предметом дослідження є методи машинного навчання та нейронні мережі.

Метою даної роботи є дослідження методів та технології реалізації систем доповненої реальності, дослідження методів класифікації об'єктів у режимі реального часу та методів побудови маркерної системи доповненої реальності для швидкого, мобільного і бюджетного розміщення AR-контенту, динамічної візуалізації об'єктів, занурення користувача у фізичний світ з реальними і доповненими об'єктами.

Методи дослідження – методи системного та об'єктно-орієнтованого аналізу, методи машинного навчання.

В роботі представлена модель штучної нейронної мережі, побудована за допомогою комбінації згорткової та рекурентної архітектури. Використання нейронної мережі дозволяє побудувати систему доповненої реальності, що у режимі реального часу дозволяє створити контент, необхідний користувачу – фото або відео з ефектами доповненої реальності, який поєднує реальні фото або відеозйомку з об'єктами доповненої реальності.

Виконане моделювання розробленої штучної нейронної мережі за допомогою бібліотеки комп'ютерного зору OpenCV (Open Source Computer Vision Library) та створена програмна реалізація системи доповненої реальності

## ABSTRACT

The thesis note contains 103 pages, 28 figures, 40 references.

ALGORITHM, ARTIFICIAL NEURAL NETWORK, AUGMENTED REALITY, AUGMENTED REALITY SYSTEM, MARKER, MIXED REALITY, VIRTUAL REALITY

The object of study is the augmented reality system as a kind of virtual reality systems.

The subject of research is machine learning methods and neural networks.

The purpose of this work is to study the methods and technologies of augmented reality systems, research methods of classification of objects in real time and methods of building an augmented reality marker system for fast, mobile and budget AR content, dynamic visualization of objects, user immersion in physical world with real and augmented objects.

Research methods – methods of systematic and object-oriented analysis, methods of machine learning.

The paper presents a model of an artificial neural network, built using a combination of convolutional and recurrent architecture. Using a neural network allows you to build an augmented reality system that allows you to create real-time content needed by the user - photos or videos with augmented reality effects, which combines real photos or video with augmented reality objects.

Modeling of the developed artificial neural network with the help of the OpenCV computer vision library (Open Source Computer Vision Library) was performed and a software implementation of the augmented reality system was created.

## ЗМІСТ

Вступ.....	7
1 Огляд і аналіз існуючих технологій, методів і систем віртуальної реальності та доповненої реальності.....	11
1.1 Аналіз сфер використання технологій доповненої реальності та актуальність роботи.....	11
1.2 Аналіз технологій віртуальної, доповненої та змішаної реальності.....	15
1.3 Класифікація систем доповненої реальності.....	21
1.4 Аналіз типових проблем систем доповненої реальності.....	25
1.5 Аналіз технічних засобів доповненої реальності.....	28
1.6 Аналіз мобільних систем доповненої реальності.....	33
1.7 Постановка задачі.....	34
2. Дослідження технологій і методів реалізації систем доповненої реальності.....	35
2.1 Дослідження напрямів використання технології доповненої реальності та обґрунтування використання маркерів в AR.....	35
2.2 Дослідження методів побудови систем доповненої реальності.....	38
2.3 Вимоги до реалізації системи доповненої реальності.....	48
3. Створення нейронної мережі для реалізації системи доповненої реальності.....	51
3.1 Обґрунтування вибору архітектури нейронної мережі для реалізації системи доповненої реальності.....	51
3.2 Архітектура комбінованої нейронної мережі для реалізації системи доповненої реальності.....	57
3.3 Методи навчання нейронних мереж RNN-CNN.....	59
3.4 Модель маніпуляції AR об'єктами.....	63
4 Практична реалізація.....	66
4.1 Алгоритм роботи системи доповненої реальності.....	66
4.2 Структура застосування доповненої реальності.....	69
4.3 Розробка системи доповненої реальності.....	70
Висновки.....	79
Перелік джерел посилань.....	81
Додаток А «Графічні матеріали».....	85

## ВСТУП

У високотехнологічному світі вже нікого не здивуєш стільниковими телефонами, ноутбуками та іншими електронними пристроями, без яких вже неможливо уявити повсякденне життя людини ХХІ століття. Завдяки інформаційним технологіям (ІТ), всілякі девайси будуть головними помічниками людей і надалі, а ці технології будуть невід'ємною частиною розвитку нашого суспільства.

Швидкий стрибок розвитку науки та інформаційних технологій призвів до того, що в сучасній світовій економіці інформація стала основоположним поняттям, а інформаційний сектор економіки – найбільш прибутковим і динамічно розвиваємим. Інформаційні технології перетворилися на незамінний спосіб взаємодії всіх суб'єктів ринку і інструмент ведення бізнесу, без яких неможливо здійснити більшість бізнес-процесів компаній [1].

Одним з перспективних напрямків ІТ-сфери є технології віртуальної реальності (Virtual Reality, VR) і її різновиду – доповненої реальності (Augmented Reality, AR). Ця актуальна технологія представляє новий спосіб отримання даних. Ефект появи повноцінних систем доповненої реальності здатний зрівнятися з ефектом появи Інтернету. І якщо «глобальна павутина» робить інформацію залежною від контексту, тобто визначається оточуючими умовами, то в досліджуваній галузі створюються нові проблеми. Стає важливою не просто наявність інформації, а добуття з інформації необхідних користувачеві даних, пов'язаних з об'єктом в режимі реального часу.

Першим, хто вжив термін «доповнена реальність» став Томас Престон Коделл. Дослідник лабораторії Боїнга в 1992 році застосував принципи технології в системі, яка була створена для допомоги робочим в монтажі електричних кабелів в літаках.

Доповнена реальність – це технологія накладення інформації у візуальній формі (текст, графіка, аудіо, відео та ін.) на реальні об'єкти в режимі реального часу. Саме взаємодія комп'ютеризованих пристроїв з картинкою реального світу відрізняє доповнену реальність від віртуальної. Їх головна відмінність в тому, що доповнена реальність вносить окремі штучні елементи в сприйняття реального

спостережуваного світу, а віртуальна реальність конструює новий (абсолютно уявний) штучний світ.

Якщо віртуальні об'єкти мають безпосередній зв'язок з реальним світом, більший, ніж просто глобальне положення, наприклад, віртуальна будівля, побудована на реальному пустирі, то для такої доповненої реальності необхідна додаткова інформація, така як кордони пустиря і його розміри. Отримання цієї додаткової інформації зазвичай досягається за допомогою спеціальних маркерів або за допомогою спеціальних функцій розпізнавання.

У мобільній доповненої реальності користувачі дивляться на пряме зображення, отримане з відеокамери на їх мобільному пристрої. Сцени, які вони бачать (тобто реальний світ), доповнюються інтегрованими тривимірними віртуальними об'єктами (тобто доповненої реальності). Ця технологія має величезний потенціал в таких галузях, як реклама, розваги, культурно-виставкова сфера і т.ін.

Якщо віртуальний об'єкт просто накладається на реальне зображення, а не інтегрується в нього, то для створення середовища доповненої реальності можуть бути використані додаткові сенсори, присутні в сучасних мобільних пристроях, такі як акселерометр, компас, GPS. Використовуючи інформацію про місцезнаходження, користувач може переміщатися по світу доповненої реальності.

Сучасний розвиток цих технологій вже дозволяє пристроям демонструвати абсолютно нові властивості об'єктів і отримувати нові відчуття від звичних реальних речей. Тому кількість областей, де можна застосовувати ці технології, дійсно безмежна.

Технології доповненої реальності в наш час вже не є картинкою з фантастичного фільму, вони реалізуються навіть в мобільних додатках для смартфонів і планшетів. Проте, ці технології зараз використовуються досить обмежено. Поява і широке поширення мобільних пристроїв з великою обчислювальною потужністю та камерами високого дозволу дозволило перейти до технологій обробки відеопотоку з одночасним додаванням віртуальних графічних або текстових об'єктів. Це ще більше розширює сферу застосування методів штучного інтелекту, технологій 3D-моделювання в напрямках створення нових принципів і

інтерфейсів створення людино-машинної взаємодії на основі розширеної реальності.

В сучасних бойових літаках і вертольотах часто використовується нашоломна система індикації. Вона дозволяє пілоту отримувати найбільш важливу інформацію про об'єкти прямо на тлі спостережуваної їм обстановки, не відволікаючись на приладову панель. Це дозволяє заощадити дорогоцінні секунди, наприклад, під час маневреного повітряного бою. В автомобільній промисловості досить давно практикують проектування показань приладів на лобове скло, так, щоб вони гармонійно накладалися на візуально сприйнятту водієм обстановку [2]. В кінотеатрах відтворення об'ємної звукової картини спільно з відеорядом також є елементом доповненої реальності в певному сенсі.

Однак, найбільші надії у користувачів пов'язані зі створенням пристроїв «hands-free». Така технологія буде затребувана в багатьох сферах від міліцейського патрулювання та орієнтування на місцевості до домогосподарок і мам, у яких руки постійно зайняті дітьми. Вважається, що рівень розвитку сучасної електроніки поки не досяг необхідних висот, і створення серійних «hands-free» пристроїв доповненої реальності поки розглядається лише в майбутньому часі.

Представлений прототип Google Glass окулярів з доповненою реальністю так і не отримав впровадження в масове виробництво, конструкція виявилася занадто складною і дорогою, а апаратна частина прихована від розробників. Аналітиками передбачається, що окуляри сильно визначили свій час і, в той же час, не запропонували будь-яких унікальних функцій для споживачів. Тим не менш, деякі прототипи використовуються в реальному житті, поліцейськими в ОАЕ і американськими хірургами в лікарнях.

Ідея розширення видимої реальності за допомогою підручних засобів, таких як мобільні телефони, здається вже не настільки далекою і вимагає технічного прогресу [3]. У сучасному світі інформація може представлятися в найцікавіших формах. Причини для цього різні і не завжди мають стеганографічний спосіб передачі або зберігання інформації (сокриття факту передачі інформації).

Завдяки технологіям доповненої реальності, що постійно розвиваються, те що раніше було фантастикою, сьогодні стає реальністю. Для створення проектів

доповненої реальності виділяються величезні гроші, і втілюють ці проекти компанії-фахівці. На даний момент технологію доповненої реальності можна реалізувати в альтернативному варіанті, набагато більш економічному, і не вимагаючому величезних знань програмування. Цим варіантом є QR-код.

Більшість сучасного суспільства сьогодні користується мобільними телефонами. За допомогою мобільного телефону можна не тільки дзвонити, передавати різні повідомлення, виходити в інтернет. У мобільного телефону є ще одне призначення: за допомогою фотокамери і спеціального встановленого додатку можна розшифрувати чорно-біле піксельне зображення QR-коду.

# 1 ОГЛЯД І АНАЛІЗ ІСНУЮЧИХ ТЕХНОЛОГІЙ, МЕТОДІВ І СИСТЕМ ВІРТУАЛЬНОЇ РЕАЛЬНОСТІ ТА ДОПОВНЕНОЇ РЕАЛЬНОСТІ

## 1.1 Аналіз сфер використання технологій доповненої реальності та актуальність роботи

Мобільні пристрої безперервно удосконалюються, щоб надавати додаткову і необхідну в конкретний момент часу інформацію: від інформації про пам'ятку архітектури до розкладу найближчих авіарейсів.

Прогрес у цій галузі в значній мірі визначається розвитком відповідних технологій. Одним з головних прикладів цього можна назвати концепцію «доповненої реальності», яка є дуже перспективною не тільки в найближчому майбутньому, але і в сьогоденні.

Доповнена реальність – термін, що відноситься до всіх проектів, спрямованих на додавання реальності будь-якими віртуальними елементами. Будь-який бажаючий зможе дізнатися, що за будівля стоїть перед ним, просто наводячи на цю будівлю фотокамеру свого смартфона. Після цього на екрані смартфона відразу з'явиться текстове вікно з докладною інформацією про цю споруду [4].

Телероботи, телеоператори і дистанційно керовані апарати відносяться до класу машин, використовуваних для досягнення завдань віддалено, без необхідності присутності людини на місці. Вони, як правило, використовуються в ситуації, які є дуже небезпечними для здоров'я людини або виживання, наприклад, в глибокій воді, космічному просторі, отруйному навколишньому середовищу. Зростаюче число таких пристроїв використовується для таких ситуацій, де було б занадто дорого або необхідно дуже багато часу, щоб відправити туди людей, наприклад, в галузі телемедицини або управління, де потрібний висококваліфікований персонал, що володіє спеціальними навичками [5].

Доповнена реальність є технологією, яка дозволяє користувачеві відображати віртуальні об'єкти в реальному середовищі. Це не обмежується тільки телеробототехнікою.

Наприклад, виробники літаків створюють модель складальної станції (Sims, 1994), де в дисплей доповненої реальності дивиться технік і створює систему проводки. Проблема полягала в складності прокладки тисяч кабелів. Це завдання вимагало бездоганною точності і безпомилковості. Робочим доводилося постійно звірятися з кресленнями, що істотно уповільнювало терміни виконання робіт. Співробітників компанії оснастили спеціальними пристроями, які накладали на видиму очима частину реального простору віртуальні об'єкти, написи і схеми, що надають необхідну інформацію в реальному часі. Таким чином, робочі могли бачити, який кабель в якому місці треба прокласти, які деталі і як з'єднати. Економити час, який раніше використовувалося для звернення і постійної звірки з паперовим схемами [6].

Дане завдання успішно вирішено шляхом створення переносних дисплеїв, які виводять потрібну інформацію. Проблема залишилася в ергономіці такого апаратного рішення – шоломи були дуже важкими, а їх тривале використання стомлювало очі. Після кількох перетворень подібних технічних комплексів стало ясно, що для масового застосування така реалізація доповненої реальності не годиться.

Але, це не означало, що доповнена реальність перестала бути цікавою технологією для різних галузей промисловості, медицини та дослідних підприємств.

Використання технологій доповненої реальності з кожним днем стає все більш і більш актуальним, має великий потенціал і активно розвивається, все частіше використовується в різних галузях. У суспільстві існує думка про віртуальну реальність, як реальність тільки для розваг, але це не так.

Маркетинг розглядається як основна галузь адаптації і каталізатор розвитку технологій доповненої реальності. Рекламна сфера стає найбільш швидкозростаючою галуззю застосування доповненої реальності в короткостроковій перспективі. Однак пізніше основною галуззю її застосування, на думку вчених, стане промислова і військова сфери [6]. Такі галузі, як роздрібна торгівля, місцеві органи управління і освіти, а також галузі важкої промисловості (включаючи видобуток природних ресурсів, металургію, конструювання, машинобудування, хімічну промисловість і т.ін.) також отримають великі вигоди від застосування технології, так як вони в

більшій мірі залежать як від фізичних активів, так і від інформаційних даних.

Системи розширеної реальності поступово проникають в багато сфер. Майбутнє доповненої реальності чекає в медицині, де лікар зможе накласти дані рентгена на конкретного пацієнта, або відображення місця розташування пухлини, щоб було легше проводити біопсію.

Доповнена реальність також використовується в автомобільній промисловості, коли водію на лобовому склі надаються дані про швидкість, вказівки по напрямку руху, а також у різних конструкторських роботах, де технології доповненої реальності допомагають створювати тривимірні моделі об'єктів. Доповнена реальність може накладатися на обчислювальні і виробничі дані і процеси і міняти уявлення реального світу, що дає кінцевому користувачу абсолютно новий досвід.

Але доповнена реальність має і проблеми, такі, що властиві новим технологіям, – це відсутність визнаних стандартів і істотної конкуренції. На даний момент існують різні оцінки обсягу ринку доповненої реальності, але всі одностайні в тому, що ринок буде рости величезними темпами. На думку аналітиків, до 2022 року доповнена реальність уже ввійде на пік свого розвитку і наблизиться до рівня проникнення в 50%, а до 2025 року подвоїть своє досягнення і досягне зони з майже 100% проникненням.

Найпривабливішим з усіх сегментів є сегмент планшетних комп'ютерів. На цьому ринку вже активно працюють такі стартапи, як Total Immersion (Франція), Metaio (Німеччина), Wikitude (Австрія), Zugara (США) і Layar (Нідерланди). Вважається, що до 2025 року більше половини всіх сервісних компаній (найбільш наближених до цифрового середовища) будуть застосовувати в тому чи іншому вигляді технологію доповненої реальності [7].

Перспективи розвитку технологій доповненої реальності позначені на рисунку 1.1.

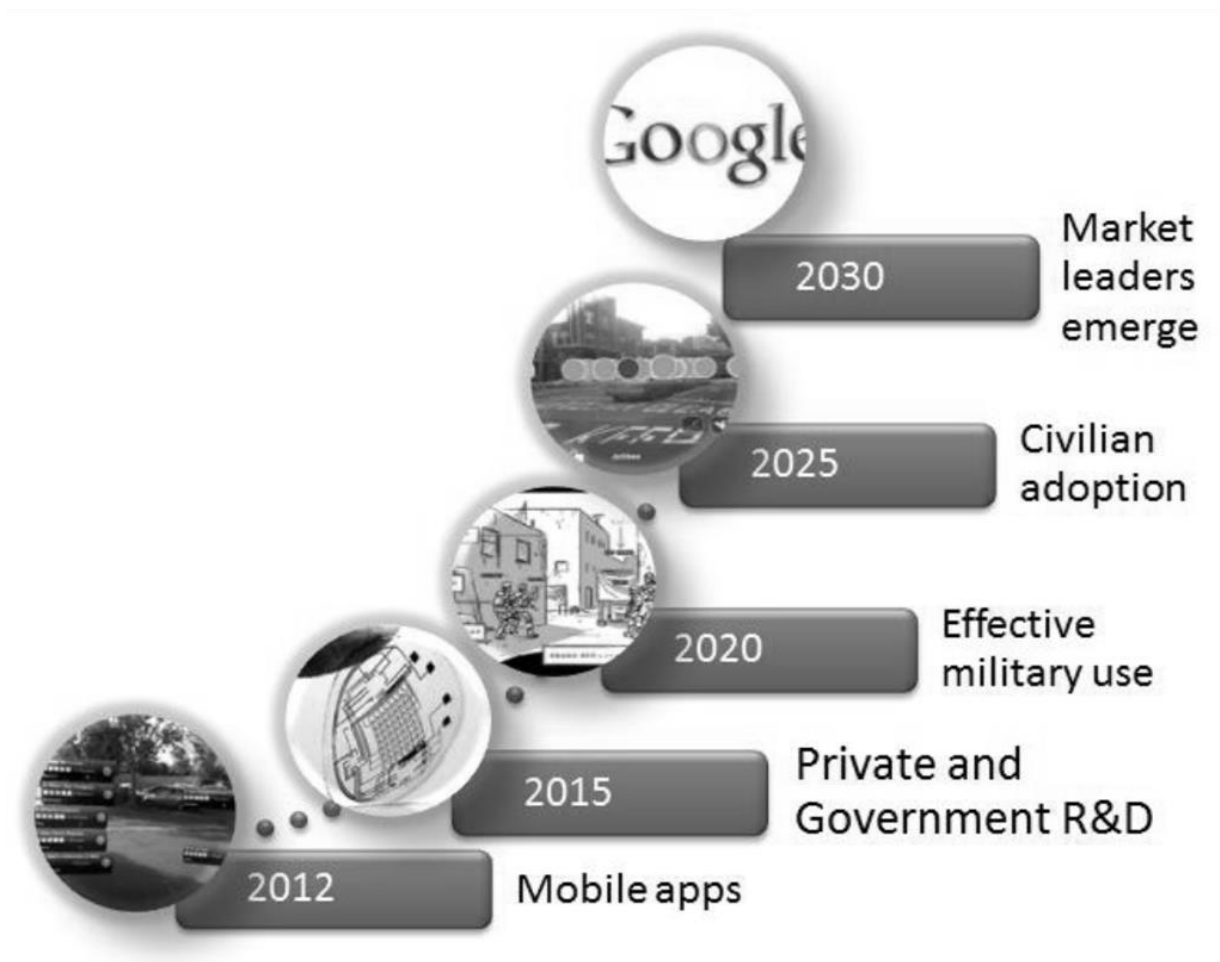


Рисунок 1.1 – Перспективи розвитку технологій доповненої реальності

Доповнена реальність почала отримувати поширення з підвищенням характеристик дисплеїв, камер, і обчислювальних потужностей мобільних пристроїв. Сучасні мобільні пристрої надають можливості по застосуванню доповненої реальності в різних галузях.

Великим потенціалом також володіють навчальні програми, інструкції, які в онлайн режимі показують необхідні дії користувача в конкретному контексті, наприклад, ілюстрація правильного підключення роз'ємів до пристрою.

У найближчі роки буде швидко рости кількість мобільних пристроїв, готових для використання технології доповненої реальності. Завдяки жорсткій конкуренції виробники смартфонів і планшетів будуть активніше впроваджувати такі компоненти, як гіроскопи і потужні процесори, які в підсумку стануть стандартним набором для мобільних пристроїв. В результаті, до 2022 року у користувачів буде не

менше 2 млрд пристроїв, готових до використання технології доповненої реальності.

Існує принаймні, шість можливих областей застосувань доповненої реальності, які вже були вивчені: медична сфера, візуалізація, технічне обслуговування та ремонт, анотації, планування шляху роботів, розваги, військова навігація літаків і орієнтування. Проте, сфера досліджень досить нова, тому навіть вчені плутаються в термінах і не охоплюють всі потенційні галузі застосування цієї технології, а деякі потенційні сфери застосування та розвитку цих засобів не вивчені досі.

Серед окремих компонентів технології доповненої реальності можна виділити наступні перспективні напрямки для розробки: засоби відображення інформації та інтуїтивного управління, тематичні бази даних, здатні працювати з різними операційними системами пристроїв, а також створення краудсорсингової платформи для створення контенту доповненої реальності самими користувачами.

Зазначені обставини визначають актуальність і практичну значимість даної роботи.

## 1.2 Аналіз технологій віртуальної, доповненої та змішаної реальності

Терміном доповненої або розширеної реальності західні аналітики називають симбіоз реального світу і віртуальної (комп'ютерної) реальності. Таким чином, доповнена реальність – це візуальне доповнення реального світу, шляхом проектування і введення будь-яких віртуальних, уявних об'єктів на даний простір (на екрані комп'ютера, телефону і подібних пристроїв) [8].

AR, як будь-яка система, визначається трьома характеристиками:

- поєднує в собі реальні та віртуальні об'єкти;
- інтерактивна в реальному часі;
- зареєстрована в трьох вимірах.

Але існують інші технології, які зберігають в собі основні компоненти

доповненої реальності. Наприклад, в таких фільмах, як «Парк Юрського періоду» особливістю є те, що фотореалістичні віртуальні об'єкти легко змішуються з реальною навколишнього середовища в 3D, але вони не є інтерактивним медіа. 2D віртуальні накладки на живому відео можливо здійснити в інтерактивному режимі, але накладки не в поєднанні з реальним світом в 3D.

Доповнена реальність є різновидом віртуальної реальності. Розглянемо поняття віртуальної і змішаної реальності.

Віртуальна реальність – це створений технічними засобами мир об'єктів і суб'єктів, який передається людині через його відчуття: зір, слух, нюх, дотик і інші. Віртуальна реальність імітує як вплив, так і реакції на вплив. Для створення переконливого комплексу відчуттів реальності комп'ютерний синтез властивостей і реакцій віртуальної реальності проводиться в реальному часі.

Об'єкти віртуальної реальності зазвичай ведуть себе близько до поведінки аналогічних об'єктів матеріальної реальності. Користувач може впливати на ці об'єкти в злагоді з реальними законами фізики (гравітація, властивості води, зіткнення з предметами, відображення і т. ін.). Однак часто в розважальних цілях користувачам віртуальних світів дозволяється більше, ніж можливо в реальному житті.

Відчуття занурення може бути досягнуто кількома способами, у яких є одна загальна особливість: все, що бачить користувач – комп'ютерна симуляція. Коли людина дивиться на монітор, то вона бачить безліч сторонніх речей, так як поряд з «нормальним» зором у людини є ще і так зване «бічний» зір – можливість бачити боком, «кутом ока».

Компанія Facebook зробила велику ставку на віртуальну реальність зі своїм недавнім придбанням Oculus VR, розробником потужної гарнітури віртуальної реальності Rift. Прилад Rift безпосередньо стимулює частини кори головного мозку, що відповідають за зір, занурюючи користувача в гіперреальність. Для геймерів це, звичайно, цікавий пристрій, але кінцева мета Facebook – створити на основі гарнітури нову грандіозну платформу після смартфонів і планшетів, яку можна буде поширити по школам, лікарням і місцях для розваг [9, 10]. В такому сценарії єдина реальність,

яка буде відчуватися, –це штучно змодельований світ всередині шоломів і очок.

Електронні віртуальні реальності також класифікуються як умовні, прожективні і прикордонні. До умовного типу віртуальних реальностей можна віднести систему, розроблену М. Крюгером, в якій зображення людини комбінується з комп'ютерної картинкою середовища. Такі віртуальні реальності моделюють певні ситуації або дії. До прожективного класу віртуальних реальностей відносяться всі реальності, спроектовані на основі деяких ідей. Наприклад, до класу прожективних віртуальних реальностей відносяться реальності, створені на основі наукових теорій. Як приклад можна привести роботу фахівців компанії «Digital Equipment Corporation», які допомагають хімікам моделювати сили молекулярного притягання і відштовхування. Своєю метою вчені поставили розробити за два роки таку систему, яка дасть можливість хімікам, в прямому сенсі слова, руками відчуті ці сили, будуючи об'ємні моделі молекул в віртуальному просторі. Прикордонні віртуальні реальності являють собою поєднання звичайної реальності і віртуальної. Їх створення дозволяє «розширювати свідомість» фахівця, озброюючи його «баченням» і знаннями, якими він не може володіти.

Традиційно вважається, що оточення віртуальної реальності таке, що учасник-спостерігач повністю занурений і взаємодіє з повністю штучним світом. Такий світ може копіювати властивості деяких реальних оточень, існуючих або вигаданих; він так само може вийти за межі фізичної реальності, створюючи світи, де фізичні закони, що регулюють простір, час, механіку, матеріальні властивості і т.ін. не підтримуються [11].

Аналітики сфокусувалися на приватному підкласі технологій віртуальної реальності, який включає об'єднання реальних і віртуальних світів, який визначають, як змішану реальність Mixed Reality (MR).

На рисунку 1.2 наведено діапазон технології змішаної реальності від реального світу до віртуального середовища.

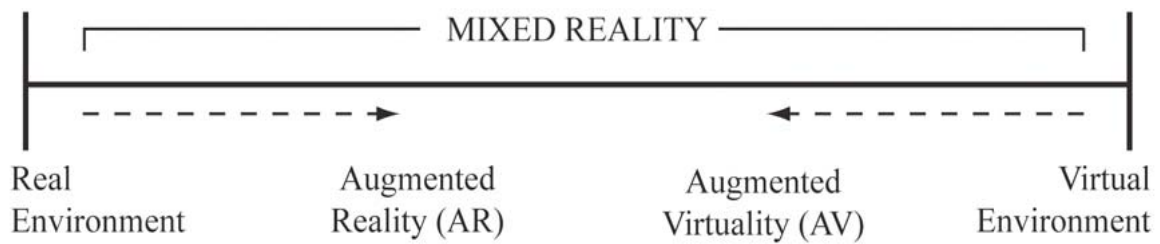


Рисунок 1.2 – Діапазон технологій змішаної реальності

Змішана (гібридна) реальність – є об'єднанням реального і віртуальних світів для створення нових оточень і візуалізацій, де фізичний і цифровий об'єкти співіснують і взаємодіють в реальному часі [12]. Об'єкти існують не тільки у реальному або віртуальному вигляді, але як суміш реальної і віртуальної реальності, охоплюючи доповнену реальність і доповнену віртуальність.

Інтегральна реальність, вбудована за допомогою непомітних цифрових компонентів в об'єктах, не відокремлює від реального світу, а гарантує емоційно привабливий досвід на його основі. Частково це вже реалізовано з інтелектуальною побутовою технікою, а також носимими (wearable) технологіями, які спостерігають за здоров'ям або фізичними навантаженнями. Споживачі вперше стикаються з ненав'язливими датчиками, які збирають дані і роблять життя більш комфортним, зручним, безпечним.

Змішана, віртуальна та доповнена реальності стали доступними завдяки певному набору розвинутих технологій. Датчики положення, системи позиціонування пристроїв і людини, енергоефективні обчислювальні платформи з дуже невеликими розмірами зробили можливими створення платформ, таких як Google Glass або Oculus Rift.

Головна подібність віртуальної і доповненої реальності в тому, що вони обидві збагачують життя віртуальними даними, хоча віртуальна максимально відокремлює людину від фізичного оточення, а доповнена навпаки змушує з ним взаємодіяти.

В цьому і головна відмінність технологій. Абстрагуючись від фізичного світу, віртуальна реальність замикається в собі, як фізичний світ без комп'ютерів, і ігнорує потреби людини як біологічної особи, фізіологічні потреби якої поза галуззю

віртуального світу, поки свідомість не відокремлена від тіла.

Істотною відмінністю доповненої реальності від віртуальної є збереження фізичного світу як контексту, в якому представлені віртуальні об'єкти і з яким вони взаємодіють. Віртуальна реальність повністю абстрагується від фізичного світу, щоб помістити користувача повністю в віртуальний світ.

Таким чином, різниця між технологіями в тому, що об'єкт і способи доповнення різні. Віртуальна реальність – це цифровий світ, створений технічно, і сприйняття його середовища передається людині через почуття, тобто фізичні здібності реальної людини доповнюють світ віртуальної реальності. Доповнена реальність – навпаки, додає до користувачів з реального світу відчуттів уявні, допоміжньо-інформативні об'єкти.

Віртуальна реальність використовує спеціальні позиційні трекери з дисплеями (окуляри віртуальної реальності), які динамічно оновлюють видимий користувачем простір у віртуальному середовищі.

Після того, як людина одягає на себе такі окуляри – все, що вона бачить, – це віртуальний світ. Це головна відмінність окулярів віртуальної реальності від окулярів доповненої реальності. Окуляри віртуальної реальності – це все одно, що навушники, але тільки для очей.

Важливо розуміти, що доповнена реальність повністю змінює цю парадигму, і в підсумку, віртуальні об'єкти розміщуються у реальному оточенні користувача [12].

Об'єкти і суб'єкти віртуального світу, включаючи людину, як введений в нього персонаж, реагують один на одного в онлайн режимі, і різниця між доповненою і віртуальною реальністю виходить розмитою: вона полягає лише в тому, що доповнена реальність – це поєднання реального і віртуального світу з відносинами в реальному часі, а віртуальна – середовище для поміщення в неї людської свідомості окремо від решти фізичного світу.

Функціонал доповненої реальності дозволяє дивитися на навколишнє середовище через камеру мобільного девайса і отримувати різноманітну корисну інформацію. Неможливо переоцінити значення цієї технології: це міст між середовищами онлайн і оффлайн, новий рівень візуалізації простору.

Доповнена реальність має потенціал зробити відносини людей з інформацією більш ергономічними. Дані будуть автоматично доставлятися користувачам у необхідному контексті для різних ситуацій в повсякденному житті, таким чином, технологія підніме взаємодію людини з інформацією на принципово інший рівень.

Таким чином, доповнена реальність – це технологія, що дозволяє доповнювати зображення реальних об'єктів різними об'єктами комп'ютерної графіки, а також поєднувати зображення, отримані від різних джерел комп'ютерного середовища: відеокамер, акселерометрів, компасів і т.ін. Схема середовища доповненої реальності представлена на рисунку 1.3.

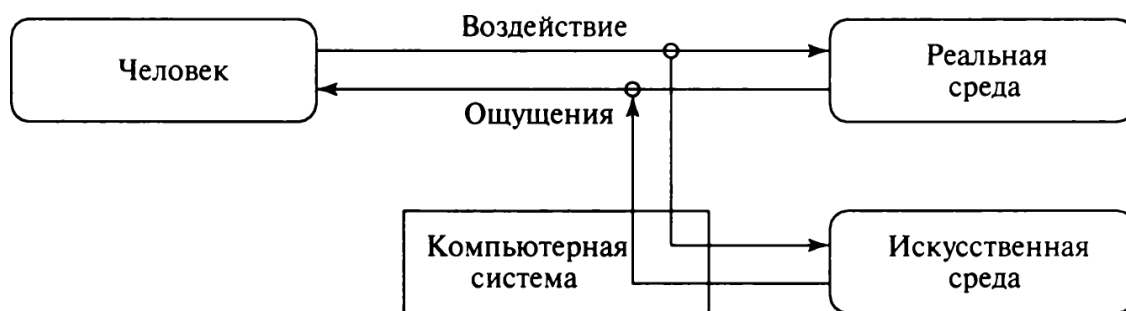


Рисунок 1.3 – Схема середовища доповненої реальності

Хоча AR системи можуть також використовувати інші людські почуття, як слуховий апарат або тактильні відчуття, в переважній більшості існуючих систем реалізований тільки візуальний канал.

Доповнена реальність не просто використовує фізичний світ, а й збагачує його і можливості людини щодо взаємодії з іншими об'єктами і суб'єктами.

Апаратно-програмні системи візуального пошуку, комп'ютерного зору, зайві в віртуальній реальності, в доповненій реальності працюють не тільки в окулярах або інших електронних пристроях, а й у робототехніці, рекламі, системах безпеки, будівництві, картографії та багатьох інших галузях. А алгоритми побудови доповненої реальності, пошуку, візуалізації, реалізації стають актуальними в обох реальностях і інших техносферах.

Доповнена реальність допомагає поєднати віртуальний і реальний світи за допомогою накладання додаткової інформації на існуючий досвід. Гнучкий пристрій Google Glass є пристроєм доповненої реальності, проте дані окуляри не перебувають постійно на вершині нашого досвіду і доповненої реальності. Вони просто надають інформацію і функціонал, який можна отримати за допомогою смартфонів. Всі дані просто виводяться на екран, що знаходиться на пристрої окулярів. Компанія Google зняла з виробництва свої окуляри і зайнялася доопрацюванням пристрою. Більш амбітні пристрої, такі як гарнітура HoloLens від компанії Microsoft, роблять віртуальні об'єкти і наведену інформацію частиною нашого сьогодення.

Сьогодні технології доповненої реальності застосовуються для їх впровадження в спеціальні окуляри і контактні лінзи. З огляду на це можна припустити, що очікується збільшення використання технологій доповненої і віртуальної реальності і в інших напрямках і галузях. Нові технології розвиваються для того, щоб вносити істотний внесок в інші сегменти і галузі, такі як мобільні обчислення, ігри, збройні сили, охорону здоров'я, маркетинг і т.ін.

На основі проведеного аналізу, можна прийти до висновку, що саме термін «доповнена реальність» дуже точний і відображає головну властивість середовища – доповненість. AR робить світ більш інформативним і багатшим, VR здатна відкрити такі горизонти у створенні цілих штучних екосистем, моделюванні, розвагах, відпочинку та навчанні, які зараз важко собі уявити, але на сьогоднішній день людині цінно саме його фізичне оточення, і люди не готові піти з нього повністю. Саме тому дослідження технологій AR є найбільш перспективним.

### 1.3 Класифікація систем доповненої реальності

Людина отримує уявлення про навколишній простір за допомогою великого набору органів почуттів. Система доповненої реальності, будучи посередником між людиною і реальністю, повинна створювати сигнал для одного з таких органів. Таким

чином, за типом подання інформації системи доповненої реальності класифікують на [13]:

- візуальні – в їх основі лежить зорове сприйняття людини. Завдання таких систем – створити зображення, яке буде використано людиною. Оскільки зображення для людини є більш інформативним і зрозумілим, такий вид систем є більш поширеним;

- аудіо – системи, орієнтовані на слухове сприйняття, найчастіше такі системи використовуються в навігації. Наприклад, вони видають спеціальні сигнали, коли людина досягає певного місця. Можливе використання стереоскопічного ефекту, що дозволяє людині йти в потрібному напрямку, орієнтуючись на джерело звуку. Прикладом такої системи є Hear & There;

- аудіовізуальні – це комбінація двох попередніх типів, однак, аудіоінформація в них має лише допоміжний характер.

Системи доповненої реальності завжди мають потребу в інформації, одержуваної з навколишнього середовища. Саме на основі цих даних будуються віртуальні об'єкти. Кожна з таких систем має певний набір сенсорів – пристроїв, що дозволяють збирати інформацію з навколишнього середовища: звукові і електромагнітні коливання, прискорення і т.ін.

В даний момент існує кілька типів пристроїв введення для роботи з доповненою реальністю. У разі смартфонів, телефон сам по собі може бути використаний в якості вказівного пристрою [13]. Сучасні пристрої оснащені цифровими камерами, GPS, акселерометром, магнітометрами, гіроскопами, які і є пристроями відстеження. Від даних, одержуваних з цих пристроїв, залежить результат роботи програми AR.

Система доповненої реальності повинна володіти потужним процесором і мати достатній обсяг оперативної і відео пам'яті для обробки зображень з камери. В наш час питання потужності не настільки актуальне, так як сучасні пристрої є досить потужними, щоб задовольнити більшу частину призначених для користувача запитів.

Для класифікації має сенс розділяти сенсори не по типам реєстрованих фізичних величин, а за їх призначенням, оскільки подібні за своєю природою сигнали можуть нести різну інформацію. За типом сенсорів можна виділити наступні системи:

– геопозиційні – орієнтуються, перш за все, на сигнали систем позиціонування GPS. На Застосування до приймачів таких сигналів геопозиційні системи можуть використовувати компас і акселерометр для визначення кута повороту щодо вертикалі і азимута;

Приклад такої системи – проект компанії Microsoft Nearby. Застосування наочно демонструє практичність доповненої реальності в смартфонах. Сервіс покликаний серйозно полегшити жителям великих міст і туристам. Використовуючи вбудовану відеокамеру, GPS-приймач і гіроскопічні датчики, програма відстежує, куди дивиться об'єктив камери і накладає на відображається картинка мітки з дистанцією до об'єкта, його описом і посиланням на докладнішу інформацію.

– оптичні – системи, що обробляють зображення, отримане з камери, які можуть переміщатися разом з системою або незалежно від неї.

Системи доповненої реальності можна розрізняти за ступенем взаємодії з користувачем. У деяких системах користувач грає пасивну роль, він лише спостерігає за реакцією системи на зміни в навколишньому середовищі. Інші ж системи вимагають активного втручання користувача – він може управляти як роботою самої системи для досягнення результатів, так і змінювати віртуальні об'єкти. За цією ознакою системи діляться на:

– автономні – не вимагають втручання користувача. Завдання таких систем зводиться до надання інформації про об'єкти. Наприклад, подібні системи можуть аналізувати об'єкти, що знаходяться в полі зору людини і видавати довідкову інформацію про них;

Також системи такого типу використовуються в медицині. Наприклад, система Gait Aid для людей з порушеннями опорно-рухового апарату. Вона шляхом використання віртуальних об'єктів надає мозку додаткову інформацію, яка допомагає координувати рухи.

– інтерактивні системи засновані на взаємодії з користувачем. На різні дії користувач отримує різні відповіді. У подібних системах необхідний пристрій введення інформації, в якості пристрою може застосовуватися сенсорний екран мобільного телефону, планшет або спеціальний маніпулятор.

Вибір пристроїв введення залежить від специфіки системи. У разі простих дій з віртуальним об'єктом досить простого вказування. Якщо ж необхідна імітація будь-яких реальних процесів і виконання складних маніпуляцій з об'єктами використовуються спеціальні маніпулятори, які мають різну кількість ступенів свободи. Прикладом можуть служити пристрої PHANTOM.

Інтерактивність виражається в різному ступені. Є системи, що дозволяють користувачеві активно змінювати віртуальне середовище. Зазвичай це системи-симулятори будь-яких реальних дій. Вони використовуються в разі, коли використання реальних об'єктів неможливо, наприклад, спеціалізовані медичні тренажери, які дозволяють початківцям лікарям відпрацьовувати необхідні навички.

Існують інші системи, де користувачеві не потрібно змінювати віртуальне середовище. Замість цього користувач вибирає, які віртуальні об'єкти він хоче побачити. Користувач також має можливість маніпулювати віртуальними об'єктами, але не на рівні структури, а на рівні відображення, тобто застосовувати, наприклад афінні перетворення типу повороту, переміщення і т.ін. До цієї групи можна віднести різні архітектурні системи, що дозволяють побачити, як впишеться в реально існуючу обстановку нова споруда або його частина, а також навігаційні і геоінформаційні системи. Подібні системи можуть показувати частини об'єктів, приховані іншими будівлями, додаткову інформацію по обраних об'єктах і т.ін.

За ступенем мобільності системи доповненої реальності можна класифікувати як [14]:

- стаціонарні системи – призначені для роботи в фіксованому місці, переміщення таких систем означає часткове або повне припинення їх працездатності;
- мобільні системи – можуть вільно переміщатися; часто таке переміщення і лежить в основі виконуваної ними функції.

Належність до того чи іншого типу визначається функціями системи. Так, симулятор хірургічного столу не повинен бути мобільним, оскільки його завдання - відтворити для людини спеціальні умови, максимально наближені до реальних. У той же час навігаційна система повинна бути якомога більш мобільною, щоб вона могла

переміщатися разом з транспортним засобом або людиною, не створюючи додаткових витрат на її переміщення.

#### 1.4 Аналіз типових проблем систем доповненої реальності

Для кожного типу систем доповненої реальності характерний свій набір проблем. Для аудіосистем основною проблемою є трекінг голови людини, оскільки саме від того як повернута або нахилена голова буде залежати відтворений звук.

Існує кілька підходів до визначення положення голови: визначення положення щодо фіксованої платформи; визначення положення щодо Землі, що обертається; визначення положення щодо довільної рухомої платформи.

Для кожного з перерахованих вище випадків вводяться системи координат і описуються вираження для визначення цих координат. Однак, яка б координатна система не використовувалась, на голову людини необхідно закріпити датчики, на основі показань яких координати і будуть обчислюватися.

Для візуальних систем основною проблемою є розпізнавання образів на зображенні. Це необхідно для ідентифікації навколишнього середовища, в залежності від якого будуються віртуальні об'єкти. Людина мислить категоріями – якимись колективними образами, з якими вона зіставляє видимі об'єкти. Те ж саме необхідно зробити і для систем доповненої реальності.

На теперішній час досі немає систем, які здатні в повній мірі та безпомилково розпізнавати об'єкти реального світу. Кожна конкретна система здатна ідентифікувати тільки певну групу об'єктів, яка визначається призначенням системи.

Процес ідентифікації зводиться до наступних етапів:

- побудова контуру (форми) об'єкта;
- побудова одновимірної функції з двовимірної форми об'єкта;
- порівняння отриманої функції з еталоном для її ідентифікації.

Однак тут виникає і ще одна проблема – класифікація об'єктів для завдання

еталонів. Наприклад, візьмемо такий об'єкт, як будинок. До цієї категорії можна віднести і невеликий одноповерховий склад, і хмарочос. При цьому зразки таких об'єктів будуть значно відрізнятися.

При розробці мобільних систем також виникають специфічні проблеми. Тут є дві тенденції: прагнення зробити пристрій якомога компактніше і забезпечити обчислювальну потужність, необхідну для забезпечення роботи системи в реальному часі.

Причому ці тенденції є протидіючими. Для отримання високопродуктивної системи доведеться змиритися зі збільшенням її розмірів і навпаки. Тому для мобільних систем використовуються менш ресурсомісткі алгоритми, навіть якщо це веде до зниження точності.

Також є ряд проблем, характерних для всіх систем доповненої реальності. Це проблеми, пов'язані з побудовою віртуальних об'єктів:

- організація сховища віртуальних об'єктів і засобів доступу до нього;
- забезпечення необхідного ступеня реалістичності віртуальних об'єктів;
- узгодження віртуальних об'єктів зі сценою.

Система доповненої реальності може оперувати цілим набором віртуальних об'єктів, які відтворюються в залежності від конкретної ситуації. Тому необхідно організувати сховище об'єктів таким чином, щоб система могла отримати до них швидкий доступ. Об'єкт може зберігатися в будь-якій формі, придатній для подальшої інтерпретації: функція або опис методу побудови, список полігонів для графічних об'єктів, зображення, звуковий файл.

Спосіб представлення об'єктів багато в чому залежить від призначення системи і її обмежень. Вибір або розробка способу представлення також є однією з проблем систем доповненої реальності. Системи різного призначення накладають свої обмеження на необхідний ступінь реалістичності об'єктів. Зразкова ситуація – коли віртуальні об'єкти не відрізняються від реальних, але на сьогодні рівень розвитку технології поки що не дозволяє цього зробити в повній мірі. Тому в даний час системи доповненої реальності оперують об'єктами в тій чи іншій мірі наближеними до реальних. У якій би мірі реалістичним ні віртуальний об'єкт, він повинен вписуватися

в сцену.

Для графічних об'єктів це означає добірку правильної перспективи, масштабу, яскравості; для звукових об'єктів – добірку гучності, синхронність з іншими звуками. Крім цього, необхідно накласти об'єкт на реальне зображення або звук таким чином, щоб він не випадав із загальної картини. До того ж сцена не є статичною, вона змінюється в залежності від дій користувача. Тому система доповненої реальності повинна забезпечувати відстеження положення реальних об'єктів на сцені і слідом за ними переміщати пов'язані віртуальні.

В результаті виникає ще одна проблема – трекінг елементів зображення. При цьому може виникати необхідність відслідковувати не один, а відразу кілька об'єктів, які можуть переміщатися не тільки щодо користувача, але і відносно один одного.

Для вирішення багатьох з цих проблем можна використовувати системи високої продуктивності. Однак вони не мобільні і часто доступні не кожному, тобто безпосередньо їх використовувати не можна. Але на базі високопродуктивних систем можна організувати обчислювальний сервіс, який вирішував би такі завдання:

- ідентифікація об'єктів за фрагментами зображення;
- пошук необхідних віртуальних об'єктів в сховище і підбір їх параметрів;
- злиття віртуальних об'єктів з фрагментами сцени;
- зміна уявлення об'єкта для забезпечення меншою ресурсоемності при його відтворенні на конкретному кінцевому пристрої.

При такому використанні систем високої продуктивності в завдання мобільних пристроїв будуть входити фіксація навколишнього середовища і відтворення доповненої реальності для користувача. Робота системи в цілому буде складатися з наступних етапів:

- визначення контурів об'єктів в цікавій для користувача галузі зображення на мобільному пристрої. Кожному з таких контурів буде присвоюватися мітка, положення якої пристрій буде відслідковувати;
- відсилання контурів і їх вмісту на сервер для обробки;
- обробка сервером отриманих даних, ідентифікація об'єктів і побудова віртуальних об'єктів. Кожен з таких об'єктів є якоюсь послідовністю байт в пам'яті

пристрою, на основі якої можна побудувати зображення об'єкта. Це може бути функція для процедурно описуваних об'єктів, вершини і текстури для полігональних об'єктів і т.ін.

- відсилання готових об'єктів, супроводжуваних ідентифікатором мітки на оригінальному зображенні;
- відтворення віртуальних об'єктів.

З проведеного дослідження випливає, що для систем доповненої реальності вузьким місцем є обчислювальна потужність пристроїв, так як вирішувані завдання вимагають реалізації складних алгоритмів, а вимоги до мобільності пристроїв не дозволяють забезпечити їх високу продуктивність на сучасному етапі розвитку техніки.

Можливим рішенням у такій ситуації є використання обчислювальних сервісів, що дозволить використовувати мобільні пристрої тільки для отримання і відтворення даних.

## 1.5 Аналіз технічних засобів доповненої реальності

Основними пристроями, що використовуються в системах доповненої реальності є дисплеї, пристрої введення, пристрої відстеження і комп'ютер.

Окуляри доповненої реальності являють собою електронний пристрій, який, крім оптичних здібностей, дозволяє також з метою подальшого доповнення за допомогою спеціального видошукача вводити в поле сприйняття різних сенсорних даних різну інформацію про навколишнє середовище, геолокацію, медичні показники і іншу інформацію.

Окуляри і шоломи віртуальної реальності мають інший принцип роботи: у пристрої встановлено відеоекрани, які створюють окреме зображення для кожного ока – подібна технологія і створює відчутний ефект тривимірного віртуального простору. Окуляри і шоломи віртуальної реальності можуть бути оснащені

датчиками, що відстежують поворот голови, гіроскопом, а також системами об'ємного звуку для створення ще більшої реалістичності.

Перші прототипи очок були розроблені ще в 1960-тих і використовувалися військовими льотчиками. За допомогою додаткового дисплея штурман міг бачити різну додаткову інформацію: кількість палива в баку і курс руху літака.

Технологічний стрибок на початку 2010-х дозволив безлічі компаній почати розробку власних рішень, що підтримують технологію доповненої реальності. На сьогоднішній день з'являється безліч технологічних стартапів, які запускають краудфандінгові кампанії на Kickstarter, IndieGoGo, Boomstarter та інших подібних інтернет-майданчиках.

Існує три основних типи дисплеїв, використовуваних в доповненої реальності: head mounted displays (HMD), ручні дисплеї і просторові [14].

HMD є пристроєм, який закріплюється на голові користувача, або на спеціальному шоломі і поміщає зображення реальної і віртуальної середовища перед очима користувача. HMD може бути або відео-прозорим або оптико-прозорим

Відео-прозорі системи більш вимогливі, ніж оптико-прозорі так, як вони вимагають, щоб користувач мав дві камери на голові, а отже обробки даних з обох камер, щоб відобразити як «реальну частину» доповненої сцени, так і віртуальні об'єкти. В цей час оптико-прозорі системи використовують технологію половини срібного дзеркала, що дозволяє дивитися на реальний фізичний світ через спеціальну «лінзу», яка накладає додаткову графічну інформацію.

Сцена, а також реальний світ в такій системі сприймаються більш природно. З іншого боку, в відео-прозорих системах доповнена реальність вже спочатку поєднана з реальною, що дає набагато більше контролю над результатом. Таким чином, контроль над часом кінцевої сцени може бути досягнутий шляхом синхронізації віртуального зображення до його відображення. В оптично-прозорому варіанті відображення реального світу не може бути загальмовано, однак це призводить до затримки при введенні в систему, графіки та обробки зображень.

Під ручними дисплеями розуміються невеликі обчислювальні пристрої з дисплеєм, які користувач може тримати в руках. Вони використовують відео-прозорі

методи для накладення графіки на реальне середовище, а так само використовують додаткові датчики, такі як компас, GPS і акселерометр.

Для реалізації систем доповненої реальності часто використовуються системи розпізнавання маркерів, такі як ARToolKit, або методи комп'ютерного зору, такі як SLAM. Такими пристроями можуть бути: смартфони, КПК і планшетні ПК. Смартфони портативні і широко поширені пристрої, що поєднують потужний процесор, камеру, набір необхідних сенсорів – це робить їх дуже перспективною платформою для AR систем.

Планшетні ПК набагато потужніші, ніж смартфони, але вони значно дорожче і занадто важкі для тривалого використання. Однак з розвитком технологій планшети стають все більш перспективною платформою для роботи з доповненою реальністю.

Просторова доповнена реальність (SAR) з використанням відеопроєкторів, оптичних елементів, голограм, і інших технологій надає графічну інформацію безпосередньо на фізичні об'єкти, не вимагаючи від користувача носити або тримати дисплей.

Просторові дисплеї відокремлюють більшість технологій від користувача і інтегрують їх у навколишнє середовище. Це дозволяє SAR природно масштабуватися до груп користувачів, що дозволяє здійснювати їх спільну роботу, підвищуючи інтерес до таких систем доповненої реальності в університетах, лабораторіях, музеях. Існують три різних підходи до SAR, які в основному відрізняються за способом доповнення навколишнього середовища: відео-прозорий, оптико-прозорий і прямого доповнення. Відео-прозорі SAR будуються на основі екрану, загального для використання. Оптично-прозорі просторові дисплеї генерують зображення, які поєднані в межах фізичного середовища.

Просторові оптичні технології, такі як плоскі або вигнуті дзеркала, прозорі екрани або оптичні голограми є важливими компонентами таких систем. Як і відео-прозорі дисплеї, побудовані на базі оптико-прозорих технологій, системи не підтримують мобільних додатків за рахунок просторово-вирівняної оптики і технології дисплеїв.

Також варто проаналізувати пристрої введення для AR систем. Деякі системи

використовують рукавички. Інші, такі як ReachMedia використовують бездротові браслети. У разі смартфонів, телефон сам по собі може бути використаний в якості вказівного пристрою, наприклад, в додатку Google Sky Map, Android телефон вимагає від користувача направити камеру в напрямку зірки або планети, про яку він хоче дізнатися. Вибір пристроїв введення в значній мірі залежить від типу системи, що розробляється і типу обраного дисплея. Наприклад, якщо Застосування вимагає, щоб руки користувача були вільними, пристрої введення будуть обрані відповідно. Аналогічним чином, якщо система використовує портативний дисплей, розробники можуть використовувати в якості пристрою введення сенсорний екран.

Пристроями стеження є цифрові камери або інші оптичні датчики, GPS, акселерометри, компаси, бездротові датчики і т.ін. Кожна з цих технологій має різну ступінь точності і багато в чому залежить від типу системи, що розробляється. Основними технологіями відстеження доповненої реальності на даний момент є: механічна, магнітна, GPS, ультразвукова, інерціальна і оптична.

Система доповненої реальності повинна володіти потужним процесором і мати достатній обсяг оперативної і відео пам'яті для обробки зображень з камери. З розвитком технологій і появою нових портативних і одночасно потужних пристроїв, таких як смартфони і планшетні ПК, проблема нестачі потужності відійшла на другий план.

Один з найбільш важливих аспектів при створенні систем доповненої реальності – це створити відповідний інтуїтивно зрозумілий інтерфейс між користувачем і віртуальними об'єктами в системі. Існують чотири основні способи взаємодії в додатках AR: матеріальний AR інтерфейс, спільний AR інтерфейс, гібридний AR інтерфейс і новий мультимодальних інтерфейс.

Матеріальні інтерфейси підтримують пряму взаємодію з реальним світом, з використанням реальних фізичних об'єктів і інструментів. Класичним прикладом матеріальних інтерфейсів, призначених для користувача є VOMAR Застосування, розроблений компанією Kato, який дозволяє людині вибрати і переставити меблі в кімнаті в доповненої реальності за допомогою реального, фізичного, інтуїтивно зрозумілого жесту. Жести на основі команд, таких як «зачерпнути» об'єкт, щоб

вибрати його для руху або «удару» по ньому, щоб скасувати операцію.

Ще один приклад такого інтерфейсу AR є TaPuMa. TaPuMa є настільним інтерфейсом, що використовує фізичні об'єкти для взаємодії з цифровим світом. Використовуються реальні об'єкти, які користувач носить з собою, в якості запитів для пошуку місць або інформації на карті. Перевага такого додатка є використання об'єктів в якості ключових слів, що усуває мовний бар'єр звичайного графічного інтерфейсу (хоча більшість з них мають мультимовну версію, вони часто бувають неправильно переведені). З іншого боку, використання об'єктів як ключових слів, може бути неоднозначним, через велику кількість можливих трактувань.

Спільні AR інтерфейси включають в себе використання декількох дисплеїв для підтримки віддаленої спільної діяльності. Для створення спільної робочої галузі використовуються 3D інтерфейс. У віддаленому обміні, AR може легко інтегруватися з кількох пристроїв з декількох місць для проведення телеконференцій.

Приклад такого інтерфейсу – Studierstube. При першому поданні Studierstube розробники представили інтерфейс, який використовує спільну доповнену реальність для об'єднання декількох користувальницький інтерфейсів: багатокористувальницького, контекстного та регіонального, а також додатків, 3D хостів, дисплейних платформ і операційних систем.

Віддалений обмін може бути використаний для поліпшення телеконференцій. Такі інтерфейси можуть бути інтегровані з медичними додатками для виконання діагностики, операцій.

Гібридні інтерфейси поєднують особливості різних, але взаємодоповнюючих інтерфейсів, а також можливість взаємодіяти за допомогою широкого спектру пристроїв. Вони забезпечують гнучку платформу для незапланованої, щоденної взаємодії, коли невідомо заздалегідь, який тип дисплею або пристрою буде використовуватися.

## 1.6 Аналіз мобільних систем доповненої реальності

Мобільні системи доповненої реальності включають в себе мобільні додатки для телефонів. Мобільні AR використовують різні мобільні інтерфейси для взаємодії користувача з віртуальними даними, які доповнюють реальний світ. Використання мобільних телефонів для доповненої реальності має як переваги так і недоліки. Більшість мобільних пристроїв в даний час обладнано камерами, що робить мобільний телефон однією з найбільш зручних платформ для реалізації систем доповненої реальності [15].

Крім того, більшість стільникових телефонів мають додаткові вбудовані датчики такі як: акселерометри, магнітометри і GPS-приймачі, які можуть поліпшити роботу AR додатку. Обчислювальна потужність телефонів для великих програм на сьогоднішній день є також достатньою. Але в деяких випадках для підвищення ефективності в додатках використовується клієнт-серверна архітектура, при якій дані передаються на віддалений комп'ютер, який виконує обчислення і відправляє результат назад на мобільний пристрій. Але при такому підході може виникнути проблема обмеженої пропускну здатності, і це може бути критичним для складних AR систем. Проте, з урахуванням швидкого розвитку мобільних технологій, ця проблема з успіхом вирішується використанням потужних телефонів, що обробляють дані для AR локально в реальному часі.

Успішною мобільною AR системою, як додатку, є система, яка дозволяє користувачеві зосередитися на самому функціоналі системи, що реалізує взаємодію з пристроєм в натуральному і соціально прийнятному вигляді, а також надає користувачеві додаткову корисну інформацію [16].

Це вказує на необхідність розробки в легких, портативних, мобільних пристроях, які володіють достатньою потужністю для складних обчислень і високими характеристиками датчиків для надійного стеження і розпізнавання.

## 1.7 Постановка задачі

Об'єктом дослідження є системи доповненої реальності, як різновид систем віртуальної реальності.

Предметом дослідження є методи машинного навчання та нейронні мережі.

Метою даної роботи є дослідження методів та технології реалізації систем доповненої реальності, дослідження методів класифікації об'єктів у режимі реального часу та методів побудови маркерної системи доповненої реальності для швидкого, мобільного і бюджетного розміщення AR-контенту, динамічної візуалізації об'єктів, занурення користувача у фізичний світ з реальними і доповненими об'єктами.

Для реалізації цієї мети в роботі необхідно:

- дослідити технологію віртуальної реальності та її різновиди (технологію доповненої реальності, змішаної реальності);
- дослідити напрями використання технології доповненої реальності та обґрунтувати використання маркерів в AR;
- дослідити методи побудови систем доповненої реальності, обґрунтувати використання нейронних мереж для реалізації системи доповненої реальності;
- сформулювати вимоги до реалізації системи доповненої реальності;
- обрати тип і архітектуру нейронних мереж, обґрунтувати вибір алгоритмів навчання;
- побудувати штучну нейронну мережу (у вигляді комбінації двох мереж) для реалізації системи доповненої реальності та виконати її навчання;
- представити структурну схему комбінації нейронних мереж для реалізації системи доповненої реальності та візуалізації об'єктів;
- представити модель маніпуляції AR об'єктами;
- виконати моделювання розробленої штучної нейронної мережі та представити програмну реалізацію системи.

## 2. ДОСЛІДЖЕННЯ ТЕХНОЛОГІЙ І МЕТОДІВ РЕАЛІЗАЦІЇ СИСТЕМ ДОПОВНЕНОЇ РЕАЛЬНОСТІ

### 2.1 Дослідження напрямів використання технології доповненої реальності та обґрунтування використання маркерів в AR

Доповнена реальність – одна з багатьох технологій взаємодії людини і комп'ютера, специфіка якої полягає в тому, що програмним чином візуально поєднуються два спочатку незалежних простори: світ реальних об'єктів і віртуальний світ, відтворений на комп'ютері.

Нове віртуальне середовище утворюється шляхом накладення запрограмованих віртуальних об'єктів поверх відеосигналу з камери, і стає інтерактивним шляхом використання спеціальних маркерів.

Основа технології доповненої реальності – це система оптичного (камера) трекінгу (маркери). Камера розпізнає маркери в реальному світі, «переносить» їх у віртуальне середовище, накладає один шар реальності на інший і таким чином створює світ доповненої реальності.

Технологія доповненої реальності це, в основі своїй, програмне забезпечення. Тобто це спеціальні математичні алгоритми, які пов'язують камеру, маркери (мітки) і комп'ютер в єдину інтерактивну систему.

Існують три основні напрямки використання цієї технології [17]:

- безмаркерна технологія AR;
- AR технологія на базі маркерів;
- просторова технологія.

Безмаркерна технологія працює за особливими алгоритмам розпізнавання, де на навколишній ландшафт, знятий камерою, накладається віртуальна сітка. На цій сітці програмні алгоритми знаходять якісь опорні точки, за якими визначають точне місце, до якого буде «прив'язана» віртуальна модель. Перевага такої технології в тому, що об'єкти реального світу служать маркерами самі по собі і для них не потрібно створювати спеціальних візуальних ідентифікаторів.

Технологія на базі спеціальних маркерів, або міток, зручна тим, що вони простіше розпізнаються камерою і дають їй більш жорстку прив'язку до місця для віртуальної моделі.

Крім маркерної та безмаркерної, існує технологія доповненої реальності, заснована на просторовому розташуванні об'єкта. У ній використовуються дані спеціальних датчиків: GPS, акселерометра, гіроскопа, компаса, вбудованих в мобільний телефон. Місце віртуального об'єкта визначається координатами в просторі. Активація програми доповненої реальності відбувається при збігу координати, закладеної в програмі, з координатами користувача.

Під маркером розуміється об'єкт, розташований в навколишньому просторі, який перебуває і аналізується спеціальним програмним забезпеченням для подальшого відтворення віртуальних об'єктів. На основі інформації про стан маркера в просторі, програма може досить точно спроектувати на нього віртуальний об'єкт, від чого буде досягнуто ефект його фізичної присутності в навколишньому просторі. Використовуючи додаткові графічні фільтри і високоякісні моделі, віртуальний об'єкт може стати практично реальним і важко відмінним від інших елементів інтер'єру або екстер'єру.

Найчастіше в ролі маркера виступає аркуш паперу з деяким спеціальним зображенням. Тип рисунку може варіюватися досить сильно і залежить від алгоритмів розпізнання зображень. У вигляді маркерів може бути будь що: геометричні фігури простої форми (наприклад, коло, квадрат), об'єкти у формі прямокутного паралелепіпеда, і навіть очі та обличчя людей.

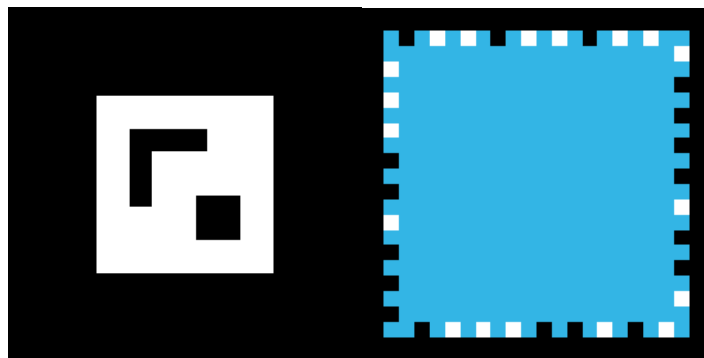


Рисунок 2.1 – Приклади маркерів

Для виключення технологічних ризиків при розробці прототипу програмного комплексу доцільно використовувати маркерну технологію доповненої реальності – надійну, перевірену, що працює практично без збоїв.

Однак, маркерна технологія також має ряд можливих проблем в роботі з мітками: при русі мітки об'єкт може «зіскочити» з неї або зовсім зникнути з екрану. Це означає, що камера просто перестала «бачити» мітку.

Перша проблема пов'язана з освітленням. Затемнена зона, занадто яскраве спрямоване освітлення, лампа денного світла, світлочутливість камери – всі ці параметри безпосередньо впливають на рівень розпізнавання мітки.

Друга проблема – це розташування реальної мітки в просторі по відношенню до камери. Оскільки камера повинна чітко і цілком бачити рамку мітки, вона не зможе розпізнати її, якщо мітка буде під нахилом або якщо область рамки буде закрита, наприклад, рукою. Ще одна причина – занадто швидке переміщення мітки з боку в бік. Більшість аматорських камер просто не встигає відстежити її переміщення по частоті кадрів в секунду і «втрачає» мітку разом з моделлю.

Ці дві причини легко усунути, дотримуючись інструкції по застосуванню.

Третя, більш серйозна проблема, пов'язана з калібруванням камери. Калібрування потрібне, щоб побудувати модель реальної камери в комп'ютерному просторі. Для того, щоб додати перспективу і глибину в 2D картинку, яка відображається з камери на екран, потрібно визначити параметри перспективної проекції для камери. Це можна зробити, використовуючи «шахівницю» і спеціальне програмне забезпечення.

Ще одна проблема, яка часто відноситься до web-камер, – це низька здатність камери. Любительська оптика, тим більше вбудовані камери на ноутбуках, як правило, не володіють хорошими об'єктивами з високою роздільною здатністю. Тому вони дають більше нелінійних спотворень і проблем в роботі з мітками доповненої реальності. Наприклад, якщо мітка буде перебувати занадто далеко від камери або на кордоні її видимості, то камера її просто «не побачить». Це питання вирішується використанням камери з більш високою роздільною здатністю і її наступним калібруванням.

І остання проблема – це програмне забезпечення. Деякі алгоритми розпізнавання можуть мати помилки і давати похибки під час розпізнавання рамки і «читання» картинки мітки. В цьому випадку моделі можуть відобразитися некоректно або зовсім зникати з екрану.

## 2.2 Дослідження методів побудови систем доповненої реальності

Математичні алгоритми технології доповненої реальності дозволяють камері побачити і розпізнати мітку (маркер) в навколишньому просторі, а потім визначити, яка саме модель програмно «прив'язана» до мітки, а потім спроектувати цю модель на мітку таким чином, щоб віртуальний 2D (фото) або 3D (відео) об'єкт повторював будь-який рух реальної мітки.

Основне завдання системи – визначити тривимірне положення реальної мітки по її знімку, отриманому за допомогою камери. Процес розпізнавання образу відбувається поетапно. Спочатку знімається зображення з камери. Потім програма розпізнає плями на кожному кадрі відео в пошуках заданого шаблону – рамки мітки. Оскільки відео передається в форматі 2D, то і знайдена на кадрі рамка мітки визначається як 2D контур. Як тільки камера «знаходить» в навколишньому просторі рамку, її наступне завдання – визначити, що саме зображено всередині рамки. Як тільки зроблений останній крок, завдання системи – побудувати віртуальну 3D модель в двомірній системі координат зображення камери і прив'язати її до мітки.

Після цього, як би не пересувалася мітка в реальному просторі, віртуальна 2D або 3D модель на ній буде точно слідувати за рухом мітки. Таким чином, основними функціями системи доповненої реальності є:

- розпізнавання зображення;
- по 2D моделі, отриманої з камери, визначити маркер, побудова 2D або 3D моделі об'єкта і прив'язування його до мітки.

Розглянемо методи вирішення цих функцій задачі побудови системи доповненої реальності.

Спочатку галузь, пов'язана з розпізнавання образів, описувалася алгоритмами, призначеними для вирішення задачі класифікації (наприклад, для визначення присутності цільового об'єкта на зображенні), що не повертає будь-яких даних про точне місцезнаходження. При цьому локалізація об'єкта (яка може бути визначена за допомогою обрамляемого прямокутника, шляхом вказівки центру і характерного розміру, завдання контуру і т.ін.) є важливим завданням в галузі автоматичної обробки цифрових зображень і знаходить своє застосування в таких галузях, як сегментація зображень виду об'єкт / фон, аналіз просторових взаємин об'єктів на зображенні, визначення та відновлення траєкторій руху і т.ін. Для випадку, коли локалізація об'єктів визначається за допомогою обрамляемого прямокутника, використовують метод скануючого (ковзного) вікна (sliding window, scan window) [18 – 19]. Даний метод полягає в послідовному аналізі різноманітних прямокутних подвікон зображення, взятих з різним зрушенням і різним масштабом, і наступним об'єднанням результатів класифікації для визначення локалізації об'єктів. Вузьким місцем для даного методу є жорстка вимога до обчислювальної складності використовуваного детектора, яке перешкоджає практичному застосуванню алгоритмів пошуку об'єктів в системах розпізнавання зображень, де найчастіше потрібне швидке (за частки секунди) прийняття рішення.

Для усунення недоліків методу ковзного вікна для детектування об'єктів американськими вченими Полом Віолою і Майклом Джонсом запропонований алгоритм пошуку особи на зображеннях, що забезпечує обробку до 15 кадрів VGA в секунду [20– 22]. Суть запропонованого алгоритму полягає в об'єднанні в єдину конструкцію чотирьох підходів:

- використання обчислювально легковажних і володіючих узагальнюючою здатністю ознак Хаара (прості прямокутні комбінації, звані функціями Хаара, які використовується для візуального виявлення об'єкта, рисунок 2.2);
- навчання класифікаторів за допомогою бустінга (техніка послідовного навчання наступної моделі на помилках попередньої);

- забезпечення високої продуктивності за рахунок використання каскаду класифікаторів;
- застосування методу ковзного вікна для визначення локалізації об'єктів.

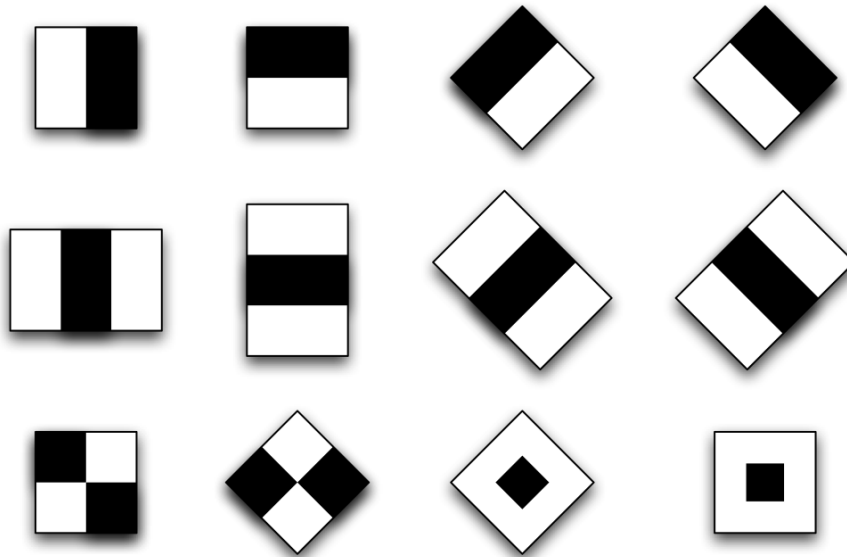


Рисунок 2.2 – Примітиви Хаара

Незважаючи на те, що даний метод спочатку розроблявся для пошуку осіб, він виявився застосовним для детекції широкого класу об'єктів, що володіють жорсткою геометрією. В результаті алгоритм Віоли і Джонса став класичним підходом для вирішення задачі локалізації об'єктів на зображенні [23].

На сьогоднішній день ступінь розробленості методу Віоли і Джонса досить висока, дослідниками по всьому світу запропоновано безліч модифікацій оригінального алгоритму і окремих його частин [23].

Незважаючи на наявність численних модифікацій алгоритму Віоли і Джонса, застосування його для вирішення великого діапазону прикладних задач в системах розпізнавання часто виявляється скрутним.

По-перше, значення класичних ознак Хаара виявляються неінваріантними до зміни освітленості, а модифікації простору ознак, які оперують з граничними точками, виявляються або обчислювально трудомісткими, або чутливими до шумових викидів, або непристосованими до масштабування.

По-друге, алгоритм Віоли і Джонса, як і представлені модифікації, вирішує завдання пошуку об'єктів в «лабораторній» постановці: набори прецедентів відомі, зафіксовані і не вимагають додаткової кластеризації на окремі підтипи, детекція об'єктів виконується на окремих стаціонарних зображеннях. Однак сучасні системи розпізнавання пред'являють до алгоритмів пошуку об'єктів додаткові вимоги. Так, наприклад, в якості джерела даних часто виступає відеокамера, яка забезпечує замість окремих стаціонарних зображень корельовану послідовність кадрів, які можуть бути використані для підвищення продуктивності алгоритму. В якості навчальних прикладів часто виступають дані, що вимагають попереднього поділу на підтипи. У відповідності з оригінальним методом навчання детектора проводиться в «пакетному режимі» (навчальна вибірка відома і зафіксована), на практиці регулярно виникає задача «донавчання» детектора у зв'язку з виникненням нових даних.

Виявлення об'єктів з використанням каскадних класифікаторів на основі ознак Хаара – це ефективний метод виявлення об'єктів.

Спочатку алгоритм вимагає досить багато позитивних зображень (зображень осіб) і негативних зображень (зображень без осіб) для навчання класифікатора [24]. Далі необхідно витягти з нього особливості. Для цього використовуються ознаки Хаара, наведені на рисунку 2.2. Кожен об'єкт являє собою одне значення, отримане шляхом віднімання суми пікселів під білим прямокутником з суми пікселів під чорним прямокутником [24, 25].

Ідея полягає в тому, що каскад Хаара витягує об'єкти з зображень за допомогою якогось фільтра. Ці фільтри називаються функціями Хаара і на рисунку 2.3 можна побачити теоретичну модель обличчя.

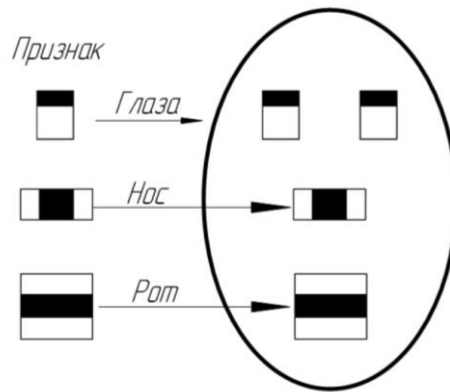


Рисунок 2.3 – Теоретична модель обличчя

Ознаки Хаара складаються з суміжних прямокутних областей. Вони позиціонуються на зображенні, далі сумуються інтенсивності пікселів в областях, після чого обчислюється різниця між сумами. Ця різниця і буде значенням певної ознаки визначеного розміру певним чином поміщеної на зображення [26].

Технологія комп'ютерного зору (Computer vision) є основоположною для розвитку технологій доповненої реальності, і перш за все в галузі використання маркерів. Основний напрямок даної дисципліни – це аналіз і обробка зображень (в тому числі і відеопотоку). Алгоритми комп'ютерного зору дозволяють виділяти ключові особливості на зображенні (кути, межі області), проводити пошук фігур і об'єктів в реальному часі, виконувати 3D реконструкцію з кількох фотографій і багато іншого.

В області доповненої реальності алгоритми комп'ютерного зору використовуються для пошуку в відеопотоці спеціальних маркерів. Залежно від завдання, в якості маркера можуть виступати як спеціально сформовані зображення, так і особи людей. Після знаходження маркера в відеопотоці і обчисленні його місця розташування, з'являється можливість побудови матриці проєкції і позиціонування віртуальних моделей. За допомогою них можна накласти віртуальний об'єкт на відеопотік таким чином, що буде досягнутий ефект присутності. Основна складність якраз і полягає в тому, щоб знайти маркер, визначити його місце розташування в кадрі і спроектувати відповідним чином віртуальну модель.

За останнє десятиліття створена велика теоретична база в сфері обробки зображень та пошуку на ньому різних об'єктів. Перш за все, це стосується методів контурного аналізу, метод сопоставлення шаблонів (template matching), метод виявлення ознак або особливостей (feature detection) і генетичних алгоритмів. Для побудови доповненої реальності найчастіше використовуються feature detection і генетичні алгоритми.

Генетичні алгоритми – це евристичні алгоритми пошуку, які використовуються для вирішення задач оптимізації та моделювання шляхом випадкового підбору, комбінування і варіації параметрів з використанням механізмів, що нагадують біологічну еволюцію.

У комп'ютерному зорі вони використовуються для пошуку об'єкта деякого заданого класу на статичному зображенні або відео потоці. Спочатку необхідно провести навчання алгоритму за допомогою двох різних наборів зображень:

- «добрі» – містять потрібний об'єкт (наприклад, зображення з обличчям);
- «погані» – помилкові зображення без об'єкта (наприклад, зображення без особи).

При цьому для навчання використовується велика кількість зображень, і чим їх більше, тим краще буде працювати сам алгоритм. Для кожної картинки проводиться виділення різних ключових особливостей: кордону, лінії, центральних елементів.

За ним проводиться побудова статистичної моделі, яка потім і використовується для пошуку об'єкта на зображенні.

Прикладом використання даного підходу може служити алгоритм розпізнавання облич і очей на відеопотоці. Поступово навчаючи алгоритм, можна домогтися високих результатів знаходження заданого класу об'єктів. Однак необхідність навчання якраз і робить використання генетичних алгоритмів досить проблематичним. Для їх хорошої роботи потрібно значне число різних зображень (як «добрих», так і «поганих»), і час побудови класифікатора для кожного об'єкта може займати тривалий час.

Метод виявлення ознак (feature detection) в комп'ютерному зорі відноситься до методів, які націлені на обчислення абстракцій зображення і виділення на ньому

ключових особливостей. Дані особливості можуть бути як у вигляді ізольованих точок, так і кривих або пов'язаних областей. Не існує суворого визначення того, що таке ключова особливість зображення. Кожен алгоритм розуміє під цим свої ознаки (кути, межі, області тощо).

Найчастіше для пошуку маркерів використовуються алгоритми, які виконують пошук і порівняння зображень по ключових точках. Ключова точка – це певна ділянка картинки, яка є характерною для заданого зображення (рисунок 2.4). Те, що саме приймається за дану точку, безпосередньо залежить від використовуваного алгоритму.

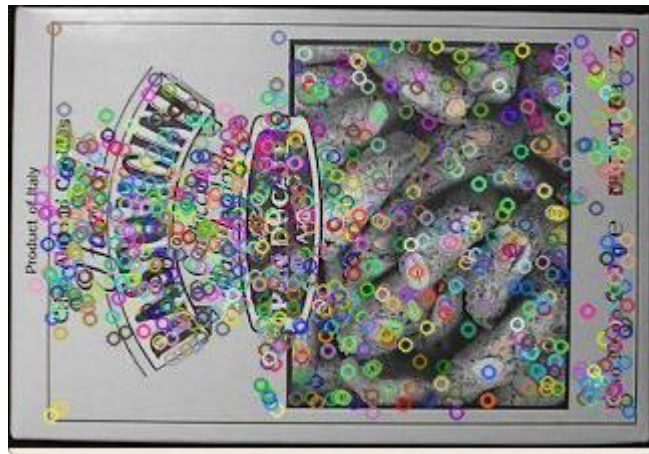


Рисунок 2.4 – Приклад ключових точок на зображенні.

Для їх знаходження і подальшого порівняння використовуються три складові:

- детектор (feature detector) – здійснює пошук ключових точок на зображенні;
- дескриптор (descriptor extractor) – виробляє опис знайдених ключових точок, оцінюючи їх позиції через опис навколишніх областей;
- матчер (matcher) – здійснює побудову відповідностей між двома наборами точок.

Спочатку за допомогою детектора проводиться пошук ключових точок шаблонного зображення. Отримані точки потім описуються за допомогою дескриптора. Дана інформація зберігається в окремий файл (або базу даних), щоб не виконувати цей процес повторно. При обробці відеопотоку з метою пошуку заданого

шаблону описаний процес виконується для кожного кадру (за винятком збереження даних). Для встановлення відповідності між ключовими точками і дескрипторами застосовується матчер (рисунок 2.5).



Рисунок 2.5 – Відповідності між точками шаблону та тестової зображення.

Різні алгоритми працюють з різною швидкістю та ефективністю. В умовах застосування їх для побудови доповненої реальності необхідно використовувати тільки ті, які показують високу швидкість роботи при досить хорошій якості відстеження позицій ключових точок, інакше можна отримати помітні відставання у знімаємих відеоданих.

Для підвищення швидкості роботи алгоритмів feature points detection застосовуються різні способи фільтрації точок, щоб мінімізувати їх число і відсіяти зовсім погані поєднання. Таким чином, можна домогтися не тільки підвищення швидкості роботи алгоритмів, але і якості трекінгу маркерів.

Розвиток робототехніки і технологій комп'ютерного зору, віртуальної і доповненої реальності, медичного і промислового сканування призвело до необхідності вирішення завдань, пов'язаних з роботою з 3D-даними. Для вирішення складних завдань, в яких необхідно не тільки працювати зі звичними зображеннями і файлами, але і класифікувати тривимірні об'єкти, відновлювати відсутню інформацію про них, перетворювати і створювати нові об'єкти використовуються алгоритми

машинного навчання (machine learning). Різноманітність завдань, пов'язаних з роботою з 3D-даними, призводить до посиленої інтеграції машинного навчання в робочі процеси.

Доповнена реальність є однією з ключових сфер застосування технологій нейронних мереж в області розпізнавання і класифікації зображень і роботи з 2D / 3D моделями. Це обумовлено тим, що нейронні мережі та використовувані інструменти спрямовані на вирішення певних завдань, які можна умовно розділити на роботу з зображеннями і відео, їх перетворенням 3D / 2D / 3D і роботу з 3D-файлами [27].

Нейронні мережі при роботі з 3D даними мають безліч напрямків застосування, представлених на рисунку 2.6.

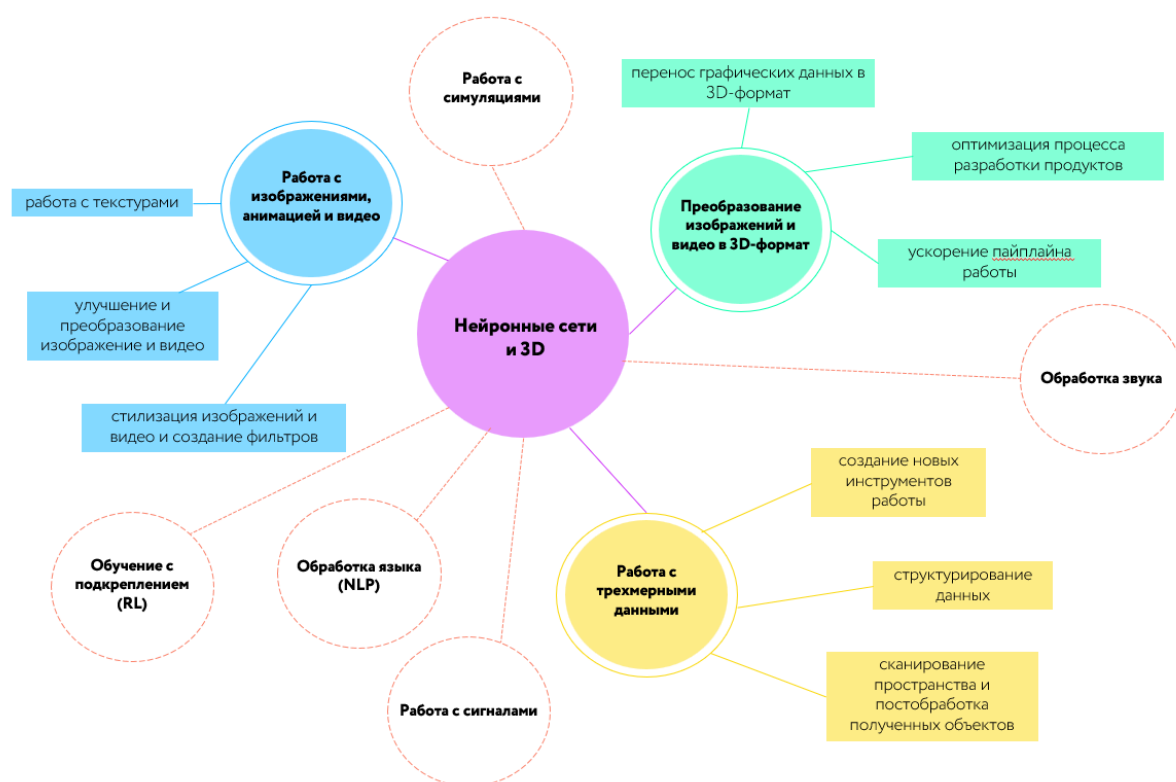


Рисунок 2.6 – Напрями застосування нейронних мереж при роботі з 3D даними

Нейронні мережі при роботі з зображеннями, анімацією та відео дозволяють [27]:

- працювати з текстурами і змінювати їх (сегментація об'єктів). При розпізнаванні об'єктів нейронна мережа може віднести об'єкт до певної групи і накласти на об'єкт обмежувальну рамку, на яку поміщається назва об'єкта;
- поліпшувати анімацію завдяки добудові кадрів;
- генерувати кадри і навіть нові зображення (за допомогою нейронних мереж можна створювати повноцінні особи і генерувати їх руху);
- створювати фільтри для зображень і відео (спеціальні інструменти дозволяють додавати різні ефекти зображень і домагатися потрібної картинки або відео, підлаштовуючись під будь-які завдання і бажання. Найбільше застосування такі нейронні мережі знаходять в мистецтві, проте так само мають великий потенціал використання в різних сферах, при інтеграції з іншими технологіями, наприклад, AR / VR);
- аналізувати інформацію на фотографіях і відео.

Сфера перетворення з 3D / 2D / 3D має великий потенціал. За допомогою нейронних мереж роботу з такими даними можна полегшити, оскільки вони дозволяють оптимізувати ресурси, що витрачаються на проекти і обробку даних.

Нейронні мережі, перетворюючи зображення, анімації та відео в 3D-об'єкти, допомагають:

- спростити процес перенесення будь-яких графічних даних в 3D-моделі (наприклад, для відновлення маски особи і моделі людини використовуються інструменти, що дозволяють розпізнати 3D-структуру особи і 3D-модель людини по зображенню або відео);
- оптимізувати процеси створення нових продуктів при програмуванні та розробці;
- раціоналізувати витрату фінансових і людських ресурсів;
- прискорювати пайплайн роботи (ланцюжок процесів перетворення сценарію (вихідних даних) в 2D (двомірну) або 3D (стереоскопічну) картинку або в послідовність картинок (відео)).

Нейронні мережі та інструменти дозволяють працювати із зображеннями, відео і 3D-об'єктами, класифікувати їх, відновлювати відсутню інформацію, добудовувати

їх, перетворювати стилі і текстури, а також породжувати нові об'єкти на основі існуючих даних. Завдяки цьому можна оптимізувати фінансові, тимчасові і людські ресурси, прискорювати процес роботи і структурувати дані.

Багато проектів вимагають роботи з 3D-даними, які представлені не структуровано, в зв'язку з чим зростає необхідність в інструментах і технологіях, які допомогли б спростити процес роботи з ними, і цю задачу також з успіхом вирішують різні нейронні мережі.

Виходячи з проведеного дослідження, доцільно використовувати нейронні мережі для побудови доповненої реальності.

### 2.3 Вимоги до реалізації системи доповненої реальності

На підставі визначених недоліків маркерних систем доповненої реальності наведених в пункті 2.1, визначимо основні вимоги до розроблюваної системи:

1) помітність: AR-маркери повинні легко і швидко знаходитися у відео-поточці. Дана вимога дозволяє врахувати асинхронну багатозадачність – властивість операційної системи пристрою, на якому інстальовано Застосування доповненої реальності, забезпечувати можливість паралельної обробки декількох процесів. При цьому процес виявлення маркерів виконується постійно, отже він повинен використовувати мінімальну кількість обчислювальних ресурсів і виконуватися з частотою, що співмірна частоті отримання кадрів відео-поточку (а в ідеалі дорівнювати їй). Вимога помітності також потребує мінімізації кількості помилок – у випадку пропуску кадра, на якому присутні маркери;

2) простота: AR-маркери повинні бути побудовані з використанням мінімальної кількості кольорів (або градацій сірого). Дана вимога тісно пов'язана з вимогами помітності та універсальності, оскільки мінімальна кількість кольорів дозволяє спростити і уніфікувати алгоритми виявлення. При виконанні вимоги простоти маркер буде суттєво відрізнятися від фону;

3) інформативність: AR-маркери повинні містити додаткову інформацію (повідомлення). Наявність додаткової інформації дозволяє суттєвим чином розширити сферу застосування маркерів, в першу чергу завдяки розширенню інтерактивності і універсальності застосувань. В такому випадку маркери можуть застосовуватися, як в системах доповненої реальності, так і в системах, де традиційно використовуються матричні коди (двовимірні штрих-коди або QR-коди), що призводить до повної інтеграції цих технологій;

4) орієнтованість: AR-маркери повинні дозволяти визначати положення камери в момент отримання кожного кадру, де виявлено маркер. Ця вимога є основою для тривимірного представлення об'єктів в процесі комбінування реального та віртуального світу, що є ключовими ознаками доповненої реальності. З урахуванням вимоги універсальності маркери повинні дозволяти проводити внутрішнє калібрування камери по самому зображенню маркера; таким чином, орієнтованість означає можливість визначення, як зовнішніх, так і внутрішніх параметрів калібрування. Дана вимога дозволяє розширити можливу галузь застосування маркерів для їх використання як калібраційних. Традиційні візуальні маркери дозволяють визначати тільки матрицю зовнішнього калібрування (розташування камери у просторі), що призводить до використання спрощених алгоритмів тривимірного представлення об'єктів і це представлення не є фізично точним;

5) робастність: AR-маркери повинні задовольняти вимоги інформативності та робастності при втраті до 50% інформації про маркер (наприклад, при закритті зображення маркеру завадою чи фізичній відсутності будь-якої частини маркеру). Вимога дозволяє масово використовувати інформаційні технології на основі маркерів на практично всіх мобільних сучасних пристроях. Вимога дозволяє з меншими витратами задовільнити вимогу помітності і проводити обробку маркерів, оскільки дозволяє спростити алгоритми обробки за рахунок втрати частини інформації;

6) криптостійкість: повідомлення, що закодоване у AR-маркері повинно бути стійким до криптоатак. Вимога дозволяє ширше використовувати маркери у військовій, медичній та інших галузях, де необхідно забезпечити конфіденційність даних. Для звичайних цивільних галузей дана вимога не висувається;

7) універсальність: технології обробки AR-маркерів повинні використовувати мінімум апріорної інформації, це незалежність алгоритмів від даних – при переході до кожної конкретної реалізації маркерів, реалізація основних методів виявлення маркерів, декодування повідомлення, тривимірного представлення об'єктів та інших повинна залишатися без змін.

Жоден стандартний тип AR-маркера не відповідає всім вимогам в повній мірі. Найкращим варіантом є 2D штрих-кодові маркери (barcode markers), але вони не задовольняють вимогам робастності і тільки частково вимогам орієнтованості та універсальності – для фізично коректного тривимірного представлення об'єктів необхідно провести попереднє калібрування камери для кожного пристрою, на якому стоїть Застосування, що використовує дані технології доповненої реальності.

### 3. СТВОРЕННЯ НЕЙРОННОЇ МЕРЕЖІ ДЛЯ РЕАЛІЗАЦІЇ СИСТЕМИ ДОПОВНЕНОЇ РЕАЛЬНОСТІ

#### 3.1 Обґрунтування вибору архітектури нейронної мережі для реалізації системи доповненої реальності

В даний час завдання аналізу відео та зображень, побудови на основі отриманих даних системи доповненої реальності є однією з найскладніших та найзатребуваніших. Це завдання вирішується різними методами машинного навчання з допомогою нейронних мереж різних архітектур. Нейронні мережі навчаються на тренувальному наборі даних, а потім відбувається оцінка точності моделі на тестовій вибірці.

Розпізнавання зображення відноситься до завдання введення зображення в нейронну мережу та присвоєння будь-якої мітки (регіону) для цього зображення. Мітка, яку виводить мережа, буде відповідати заздалегідь визначеному класу. Може бути присвоєний лише один клас (завдання розпізнавання) або кілька класів, тоді таке завдання розпізнавання кількох класів називається класифікацією і у даній роботі.

Класифікація зображень – це процес визначення об'єктів, при якому певні екземпляри об'єктів ідентифікуються як такі, що належать до певного класу. Прикладом такої класифікації є рішення найпоширенішої капчі — ReCaptcha v2 від Google, в якій із набору картинок необхідно вибрати лише ті, що належать до вказаного в описі класу.

Класично це завдання може бути вирішено за допомогою згорткових нейронних мереж (Convolutional Neural Network) та рекурентних нейронних мереж (Recurrent Neural Network).

Розглянемо архітектуру згорткової нейронної мережі [28– 31].

Згортка (convolution). Перший шар згорткової нейронної мережі приймає всі пікселі у зображенні. Після того, як всі дані введені в мережу, до зображення

застосовують різні фільтри, які формують розуміння різних частин зображення. Це витяг ознак, що створюють «карти ознак».

Цей процес отримання ознак із зображення виконується за допомогою «згорткового шару», і згортка просто формує уявлення частини зображення. Саме з цієї концепції згортки слід термін згорткова нейронна мережа - тип нейронної мережі, що найчастіше використовується в класифікації та розпізнаванні зображень.

Вилучення ознак за допомогою фільтрів. Цифрові зображення відображаються у вигляді висоти, ширини і деякого значення RGB, яке визначає колір пікселя, тому «глибина», що відстежується, – це кількість колірних каналів, які має зображення. Зображення в градаціях сірого (не кольорове) мають лише 1 кольоровий канал, тоді як кольорові зображення мають глибину 3 канали (RGB).

Таким чином, для фільтра розміром 3, застосованого до повнокольорового зображення, підсумкові розміри цього фільтра будуть  $3*3*3$ . Для кожного пікселя, охопленого цим фільтром, мережа помножує значення фільтра на значення самих пікселів, щоб одержати числове уявлення цього пікселя. Потім цей процес виконується для всього зображення, щоб отримати повне уявлення.

Функції активації. Після того, як карта ознак зображення створена, значення, що представляють зображення, передаються через активацію або шар активації. Функція активації набуває цих значень, які завдяки згортковому шару перебувають у лінійній формі (список чисел) і збільшує їх нелінійність, т.к. самі зображення є нелінійними. Типовою функцією активації, що використовується для досягнення цієї мети є випрямлена лінійна одиниця.

Поєднання шарів (pooling). Після активації дані відправляються через шар, що об'єднує. Об'єднання «спрощує» зображення: бере інформацію, яка представляє зображення, і стискає її. Процес об'єднання в пул робить мережу гнучкішою і здатною краще розпізнавати об'єкти та зображення на основі відповідних функцій.

Об'єднуючий шар CNN дозволяє позбутися непотрібних частин зображення, залишивши тільки ті частини, які є релевантними в залежності від заданого розміру об'єднуючого шару.

Стиснення (filter concatenate). Останні щільно пов'язані шари вимагають, щоб дані були представлені у формі вектора для подальшої обробки. Для цього значення стискаються у довгий вектор або стовпець послідовно впорядкованих чисел.

Повнозв'язаний шар (fully bonded output layer). Кінцевий шар CNN представляє щільно пов'язаний шар. Помилка або різниця між розрахованими значеннями та очікуваним значенням у навчальному наборі розраховується за допомогою даного шару. Кінцевий повністю пов'язаний шар представляє собою вектор значень. Отримавши вихідні дані попереднього шару, він надає вирогідність кожному з класів у межах одиниці (у сукупності).

Архітектура згорткової нейронної мережі представлена рисунку 3.1. та у таблиці 3.1

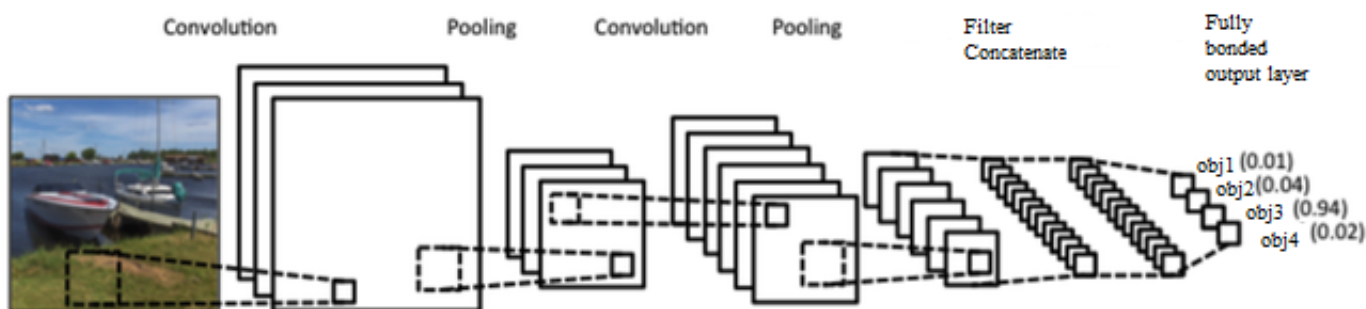


Рисунок 3.1 – Архітектура згорткової нейронної мережі

Таблиця 3.1. – Архітектура згорткової нейронної мережі

Номер шару	Тип шару	Розмір шару (входу)	Розмір ядра згортки
1	Згортковий (convolution)	100*100	5*5
2	Поєднуючий (pooling)	25*25	5*5
3	Згортковий (convolution)	16*16	3*3
4	Поєднуючий (pooling)	8*8	3*3
5	Повнозв'язаний (filter concatenate)	16	-
6	Повнозв'язаний (filter concatenate)	30	-
7	Кінцевий повністю пов'язаний шар (fully bonded output layer)	10	-

Таким чином, згорткова мережа представляє чергуванням шарів:

- згортки (convolution) – отримання карти ознак, для різних типів об'єктів передбачаються різні карти ознак;
- підвиборки зменшення розмірності (pooling, поєднання в пул).

Далі виконується об'єднання шарів за допомогою об'єднуючого фільтра (filter concatenate) і об'єднання у кінцевий повністю пов'язаний шар (fully bonded output layer). Під час аналізу сукупності зображень, що отримані з відеопотоку, необхідно використовувати відповідну кількість нейронних згорткових мереж.

Однак, точність CNN значною мірою визначається розміром навчального набору даних, щоб досягти високої точності, потрібні обсяги даних великого розміру [32]. Не завжди можливо створити такий набір даних (наприклад, через рідкість явищ, що спостерігаються або подій, недостатньо добра роздільна здатність зображень). До того ж, необхідність поділу на навчальну та тестову вибірки також знижує розмір можливого датасету для навчання.

Останнім часом все більшої популярності набувають рекурентні підходи для обробки зображень. Зображення можна розбити на фрагменти і впорядкувати в послідовність, що дозволяє рекурентним системам вивчати глобальну інформацію про зображення, спочатку аналізуючи локальні кореляції між ділянками зображення. Динаміка фрагментів зображення змінюється, а рекурентні системи добре пристосовані успішного моделювання траєкторії послідовності цих фрагментів [5].

Такий підхід використовує піксельна нейронна рекурентна мережа, представлена на рисунку 3.2.

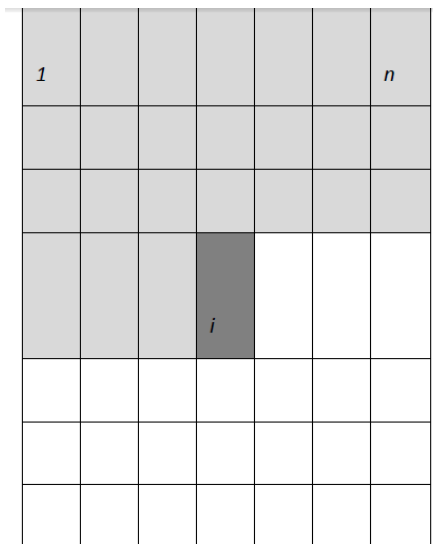


Рисунок 3.2 – Сканування зображення піксельною RNN

Мережа сканує зображення по одному рядку по одному пікселю за один раз у кожному рядку. Потім ця мережа прогнозує умовні розподіли за можливими значеннями пікселів. Розподіл пікселів зображення обчислюється як добуток умовних розподілів і ці значення розподіляються по всіх пікселях зображення. Ціль полягає в тому, щоб призначити ймовірність  $p(x)$  кожному пікселю  $(n \times n)$  зображення. Це можна зробити, записавши ймовірність пікселя  $x_i$  як  $p(x) = \prod_{i=1}^{n^2} p(x_i | x_1, \dots, x_{i-1})$ .

Це ймовірність  $i$ -го пікселя з урахуванням ймовірності всіх раніше розглянутих пікселів. Крім того, колір пікселя  $x_i$  визначається спільно всіма трьома каналами: червоним, зеленим і синім (RGB). Таким чином, кожен колір залежить від інших кольорів, а також від попередньо оброблених пікселів.

Піксельна RNN може моделювати локальну інформацію, отриману від зображення, таку як текстура або колір, при цьому розуміючи глобальний контекст пікселя, що розглядається, наприклад, чи належить піксель тому чи іншому об'єкту. Така нейронна мережа може поділити об'єкти на категорії для кращої класифікації та зберігання, що оптимізує роботу з 3D-об'єктами.

Архітектура простої рекурентної мережі по суті описується звичайною нейронною мережею, представленою на рисунку 3.3.

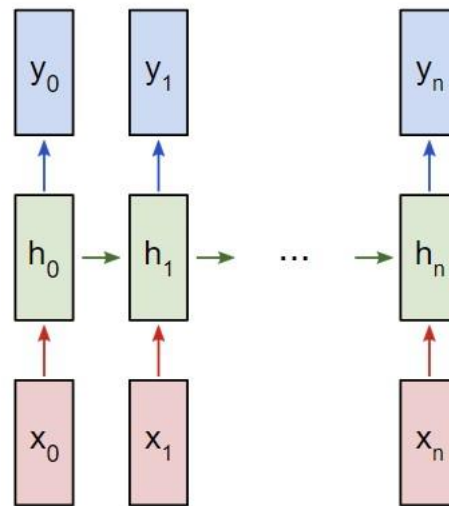


Рисунок 3.3 – Архітектура рекурентної нейронної мережі

де  $x_t$  – вектор вхідного шару;

$h_t$  – вектор прихованого шару;

$y_t$  – вектор вихідного шару.

Незважаючи на те, що рекурентні системи є досить успішними в обробці та класифікації зображень (зокрема, при обмеженому розмірі навчальної вибірки) така архітектура також має низку обмежень для вирішення цієї задачі. Рекурентні нейронні мережі містять цикли, що зберігають інформацію. При збільшенні розміру нейронної мережі (наприклад, під час обробки відео) нейромережа має дуже велику обчислювальну складність та може втрачати зв'язок між інформацією.

Саме тому існують методи класифікації зображень, які мають досить гарну точність при навчанні на обмеженому наборі даних, наприклад, комбінація двох архітектур: рекурентної та згорткової. Така архітектура називається рекурентна згорткова RNN-CNN і оптимізована для класифікації зображень.

На відміну від звичайної CNN, RNN-CNN здатна адаптивно фокусувати свою увагу на різних частинах зображення, залежно від інших об'єктів, які були ідентифіковані [31]. В результаті така мережа краще розуміє контекст, який використовується для точної класифікації багатьох об'єктів.

### 3.2 Архітектура комбінованої нейронної мережі для реалізації системи доповненої реальності

Комбінована рекурентна згорткова нейронна мережа RNN-CNN включає дві великі частини [32]:

- частина CNN отримує семантичні уявлення з зображень (семантична «близькість» «заохочує» об'єднання пікселів, що належать одному об'єкту сцени і штрафує за об'єднання пікселів різних об'єктів, тим самим визначаються зв'язкові об'єкти);
- частина RNN моделює відносини між зображеннями та мітками, а також залежність між мітками.

Для представлення зображення в зручній для обробки формі використовується згорткова нейронна мережа, а рекурентний шар отримує інформацію про раніше визначені мітки. Імовірність вихідної мітки обчислюється відповідно до подання зображення та вихідного сигналу рекурентного шару. RNN використовується як компактний, але потужний спосіб представлення залежності спільної появи міток (у разі, якщо на одному зображенні присутні об'єкти класифікації декількох типів, що розпізнаються, при аналізі сукупності зображень) в цьому просторі.

Структурна схема нейронної мережі RNN-CNN для реалізації системи доповненої реальності представлена на рисунку 3.4.

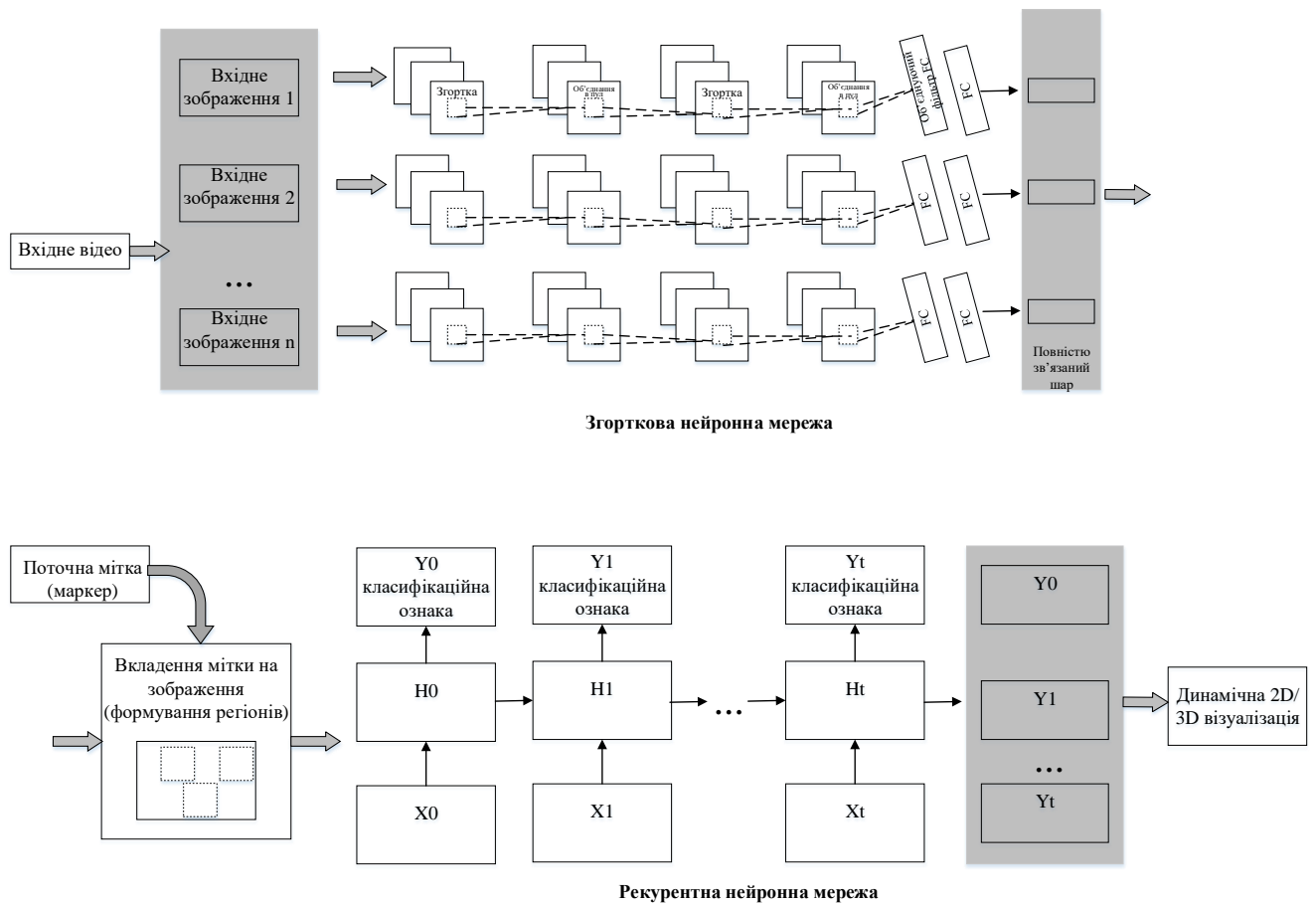


Рисунок 3.4 – Структурна схема нейронної мережі RNN-CNN для реалізації системи доповненої реальності

Основним призначенням RNN-CNN мережі є виявлення об'єктів та їх класифікація. Практично вирішення цього завдання зводиться до виділення всіх об'єктів на зображенні обмежувачими мітками (рамка або регіон). Процес можна розділити на два основні етапи: формування гіпотез про місцезнаходження об'єктів або так звану пропозицію регіонів (region proposal) та класифікацію.

У процесі селективного пошуку генерується до 2000 різних міток, які з ймовірністю може містити об'єкт. Після того, як пропонувані регіони згенеровані, вони приводяться до розміру, придатного для обробки за допомогою нейронної мережі, що була навчена, та отримує вектор ознак кожного регіону. Далі вектор ознак подається на вхід набору лінійних класифікаторів на основі векторів, навчених для кожного класу, після цієї операції виконується класифікація.

### 3.3 Методи навчання нейронних мереж RNN-CNN

При навчанні нейронної мережі архітектури RNN-CNN, що застосовується для ефективної класифікації зображень, доцільно розглянути спочатку навчання згорткової мережі, а потім рекурентної мережі.

Як набір даних використовувалася вибірка для машинного навчання предметів інтер'єру. Навчальна вибірка містить 60 000 зразків, а тестова – 10000 зразків. Кожен зразок є двовимірною матрицею чисел від 0 до 255 розміром 100\*100 пікселів.

При використанні простої нейронної мережі прямого поширення без зворотних зв'язків, що містить вхідний, прихований і вихідний шари (вихідний шар повинен містити стільки ж нейронів, скільки і вхідний шар) при великих даних кількість параметрів (кількість ваг між вхідним і прихованим шарами, а також кількість ваг між прихованим та вихідним шарами) істотно зростає.

Наприклад, якщо при класифікації картинок розміром  $m \times n = 8 \times 8$ , ми хочемо виділяти  $k = 50$  ознак, тоді вхідний та вихідний шари мають  $m \times n = 8 \times 8 = 64$  нейронів, а прихований шар має  $k = 50$  нейронів, кількість ваг дорівнює  $2 \times m \times n \times k = 2 \times 8 \times 8 \times 50 = 6400$ . Але якщо картинка мають розмір  $m \times n = 100 \times 100$ , то кількість ваг буде  $2 \times 100 \times 100 \times 50 = 1\,000\,000$ .

Для того, щоб розмір нейронної мережі не був таким великим і не залежав від розміру даних, що навчаються, використовується спеціальна згорткова структура нейронної мережі, представлена в п. 3.1 (рисунок 3.1). Шари згортки та субдескрипції нейронної мережі чергуються, таким чином, за рахунок зниження розмірності мережі здійснюється навчання.

Згортка. При навчанні для кожного зразка на вхід згорткової нейронної мережі потрапляє не вся картинка  $m \times n$ , а лише її частина  $r \times c$  наприклад,  $10 \times 10$ , ця частина називається фільтром або ядром і вона зсувається по великій картинці. Така згорткова нейронна мережа має фіксовану кількість нейронів на вхідному та вихідному шарах.  $r \times c = 10 \times 10 = 100$ , а кількість ваг  $2 \times r \times c \times k = 2 \times 10 \times 10 \times 50 = 10\,000$ . Для

кожної великої картинки  $m \times n$  беремо маленькі картинки  $r \times c$  для навчання, їх кількість дорівнює  $(m - r + 1) \times (n - c + 1)$ .

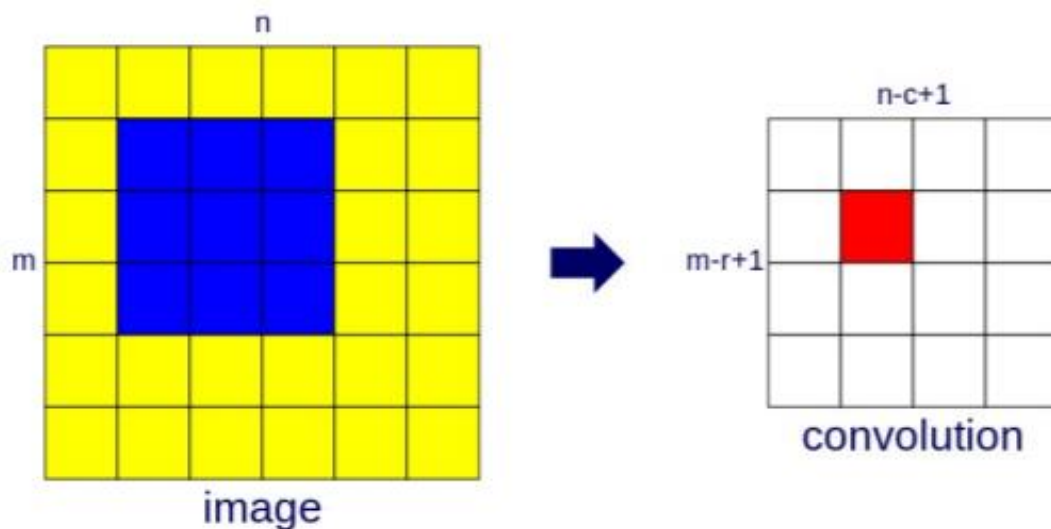


Рисунок 3.5 – Згортка

Субдискретизація. При використанні згорткової нейронної мережі виникає проблема – велика кількість виділених ознак. Для однієї великої картинки  $m \times n$  буде  $(m-r+1) \times (n-c+1)$  маленьких картинок  $r \times c$ , і кількість виділених ознак дорівнює  $k \times (m - r + 1) \times (n - c + 1) = 50 \times (100 - 10 + 1) \times (100 - 10 + 1) = 414\,050$ . Використання такої величезної кількості ознак для класифікації є неефективним. Для зменшення розміру простору ознак проводиться субдискретизація (pooling), розділивши карту ознак, отриманих від згорткової нейронної мережі, на фіксовану кількість частин  $p$ , (рисунок 3.6), на цьому рисунку  $p = pm \times pn = 2 \times 2 = 4$  і кожної частини обчислюється її максимальне значення (max pooling) чи середнє значення (mean pooling). У цій роботі використовується max pooling.

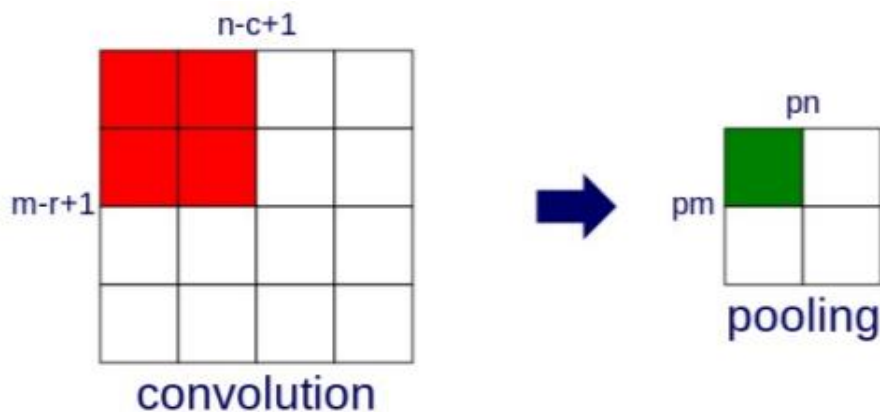


Рисунок 3.6 – Субдискретизація

Кінцевий повнозв'язний вихідний шар є вектором значень, який подається на вхід нейронної мережі рекурентної мережі.

При навчанні згорткової нейронної мережі істотно зменшується кількість параметрів, що навчаються, підвищується швидкість навчання, відбувається висока якість класифікації в порівнянні з повнозв'язковою нейронною мережею.

Навчання рекурентної мережі виконується за допомогою методу зворотного поширення помилки у часі.

Зворотне поширення помилки – це метод градієнтного спуску, просте обчислення градієнта та рух у напрямі. Аналогічно метод зворотного розповсюдження помилки в часі, що зазвичай скорочено позначається ВРТТ (Backpropagation through time), є аналогом назви методу зворотного поширення помилки для рекурентної мережі та візуальне подання методу навчання наведено на рисунку 3.8.

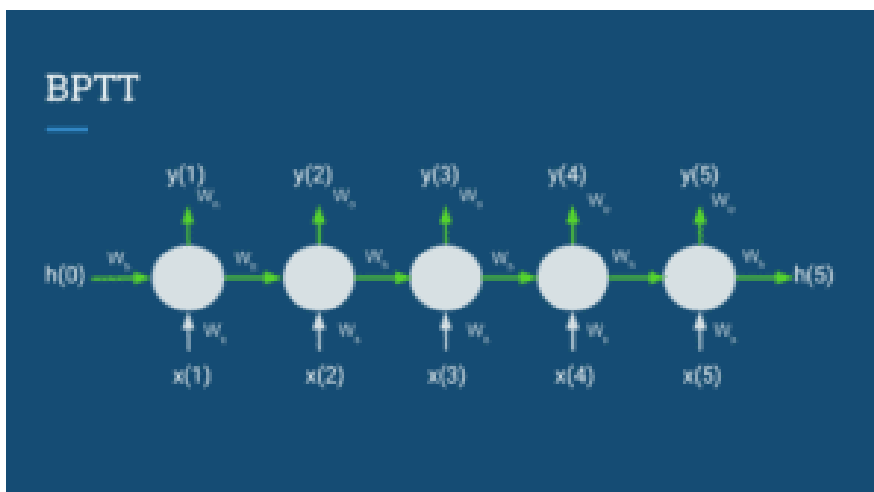


Рисунок 3.8 – Метод ВРТТ для навчання рекурентної нейронної мережі

Це означає, що поновлення всіх вагових коефіцієнтів відбуваються після рекурсії. Для того, щоб враховувати час, припустимо, що необхідно для  $w_h$  повернутися назад у часі на три етапи тому. Тоді модель навчання рекурентної мережі має вигляд:

$$\begin{aligned}
 y(t) &= \text{softmax}(W_o^T h(t)), \\
 y(t) &= \text{softmax}(W_o^T f(W_h^T h(t-1) + W_x^T x(t))), \\
 y(t) &= \text{softmax}(W_o^T f(W_h^T f(W_h^T h(t-2) + W_x^T x(t-1)) + W_x^T x(t))), \\
 y(t) &= \text{softmax}(W_o^T f(W_h^T f(W_h^T f(W_h^T h(t-3) + W_x^T x(t-2)) + W_x^T x(t-1)) + W_x^T x(t))). \quad (1)
 \end{aligned}$$

де  $x_t$  – вектор вхідного шару;

$h_t$  – вектор прихованого шару;

$y_t$  – вектор вихідного шару;

$w_t$  – вектор вагових коефіцієнтів;

$t$  – момент часу.

Функція Softmax застосовується для задач класифікації, якщо кількість можливих класів більше двох, використовується для останнього шару, дозволяючи отримати вектор з класифікованими об'єктами. Координати отриманого вектора при цьому визначаються як ймовірність того, що об'єкт належить до певного класу. Кожна

координата отриманого вектора представлена речовим числом в інтервалі  $[0,1]$  та сума координат дорівнює 1.

На рисунку 3.8 показано вплив вагових коефіцієнтів  $w_h$  зеленим кольором. Стрілки вгору, що йдуть до вихідних  $u$  мають значення тільки в тому випадку, якщо вони розглядаються як частина виходу. Вони відрізнятимуться залежно від того, чи маємо ми мету для кожного моменту часу чи метою є лише кінець послідовності. Стрілки, що йдуть зліва направо, повинні супроводжуватися зворотним поширенням, оскільки вони пов'язують вагові коефіцієнти прихованих шарів.

Важливо побачити таку закономірність: одні й самі величини перемножуються знову і знову, згідно з ланцюговим правилом диференціального числення. Це відбувається як з ваговими коефіцієнтами між прихованими шарами, так і з коефіцієнтами між вхідним та прихованим шарами. В результаті ми отримуємо величини, які або прагнуть нуля, або дуже швидко ростуть. Ці ситуації називаються відповідно проблемою градієнта, що зникає.

Одним із рішень, запропонованих для подолання проблеми градієнта, що зникає, є градієнтна обрізка. Це дуже простий алгоритм.

Інший модифікацією методу зворотного поширення помилки у часі є метод усіченого поширення помилки у часі. Оскільки похідні щодо  $w_h$  і  $w_x$  залежать від кожного моменту часу в послідовності, їх обчислення займає багато часу для дуже довгих послідовностей. Загальний метод наближення – просто зупинитись після певної кількості моментів (рекомендується після трьох моментів). Його недоліком є та обставина, що він не включатиме помилку для більш тривалих відрізків часу.

### 3.4 Модель маніпуляції AR об'єктами

Маніпуляція – процес обробки дій користувача, при якому виконується збереження змін віртуальних 3D-об'єктів за будь-якої зміни об'єкта реального світу таким чином, щоб візуалізувати віртуальні об'єкти у реальному оточенні.

Координати точок об'єкта при повороті, переміщенні та зміні масштабу об'єкта у загальному випадку змінюються у вигляді моделі

$$\begin{bmatrix} X_1 \\ Y_1 \\ Z_1 \end{bmatrix} = M \cdot \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \cdot \begin{bmatrix} X_0 \\ Y_0 \\ Z_0 \end{bmatrix} + \begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix}, \quad (2)$$

де  $X_0, Y_0, Z_0$  – початкові координати точки об'єкта;

$X_1, Y_1, Z_1$  – кінцеві координати точки об'єкта;

$M$  – коефіцієнт зміни масштабу;

матриця  $\|a_{mn}\|$  – матриця перетворення повороту, елементи якої виражаються через кути повороту довкола трьох просторових осей;

$\vec{r} = (r_1, r_2, r_3)$  – вектор переміщення об'єкта.

При вирішенні практичної задачі дизайну – розміщення віртуальних меблів, користувачу не потрібен повноцінний інтерфейс маніпуляції тривимірними об'єктами з шістьма ступенями свободи. Завдання переміщення предметів меблів є двовимірним. Таким чином, у цьому випадку модель може бути спрощена. Переміщення предметів меблів здійснюється в межах площини підлоги приміщення, по осях  $X$  та  $Z$ . Обертання здійснюється навколо своєї осі  $Y$  (рисунок 3.9).

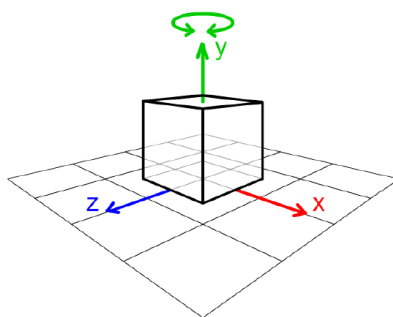


Рисунок 3.9 – Маніпуляція AR об'єктом вздовж осей переміщення та осі обертання

У разі переміщення та обертання об'єкта в площині  $XZ$  без зміни масштабу наведена модель спрощується: координата  $Y$  залишається незмінною, а координати  $X$  та  $Z$  змінюються відповідно до моделі

$$\begin{bmatrix} Z_1 \\ X_1 \end{bmatrix} = \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix} \cdot \begin{bmatrix} Z_0 \\ X_0 \end{bmatrix} + \begin{bmatrix} r_3 \\ r_1 \end{bmatrix}, \quad (3)$$

де  $\varphi$  – кут повороту об'єкта навколо осі  $Y$ .

## 4 ПРАКТИЧНА РЕАЛІЗАЦІЯ

Загальна схема створення доповненої реальності така: камера пристрою AR знімає зображення реального об'єкта; програмне забезпечення (ПЗ) пристрою проводить ідентифікацію отриманого зображення, вибирає або обчислює відповідне зображення візуального доповнення, поєднує реальне зображення з його доповненням і виводить підсумкове зображення на пристрій візуалізації. Технологію створення та реалізації доповненої реальності буде розглянуто на прикладі планування дизайну інтер'єру.

### 4.1 Алгоритм роботи системи доповненої реальності

Використання технології доповненої реальності для підтримки будівельних процесів є одним з найактуальніших у сфері використання нових інформаційних технологій у дизайні інтер'єру.

При плануванні дизайну інтер'єру використання додатків з доповненою реальністю надає можливість експериментувати зі стилями, розстановкою меблів, побачити, як той чи інший предмет виглядатиме у певній частині квартири, як гармоніюватиме з обраною кольоровою гамою. При цьому можна візуалізувати існуючий інтер'єр і за допомогою об'єктів доповненої реальності вносити зміни. Для цього необхідні лише застосування доповненої реальності та навести камеру пристрою на маркер.

Застосування доповненої реальності, що розробляється, працює за алгоритмом, який представлений на рисунку 4.1.

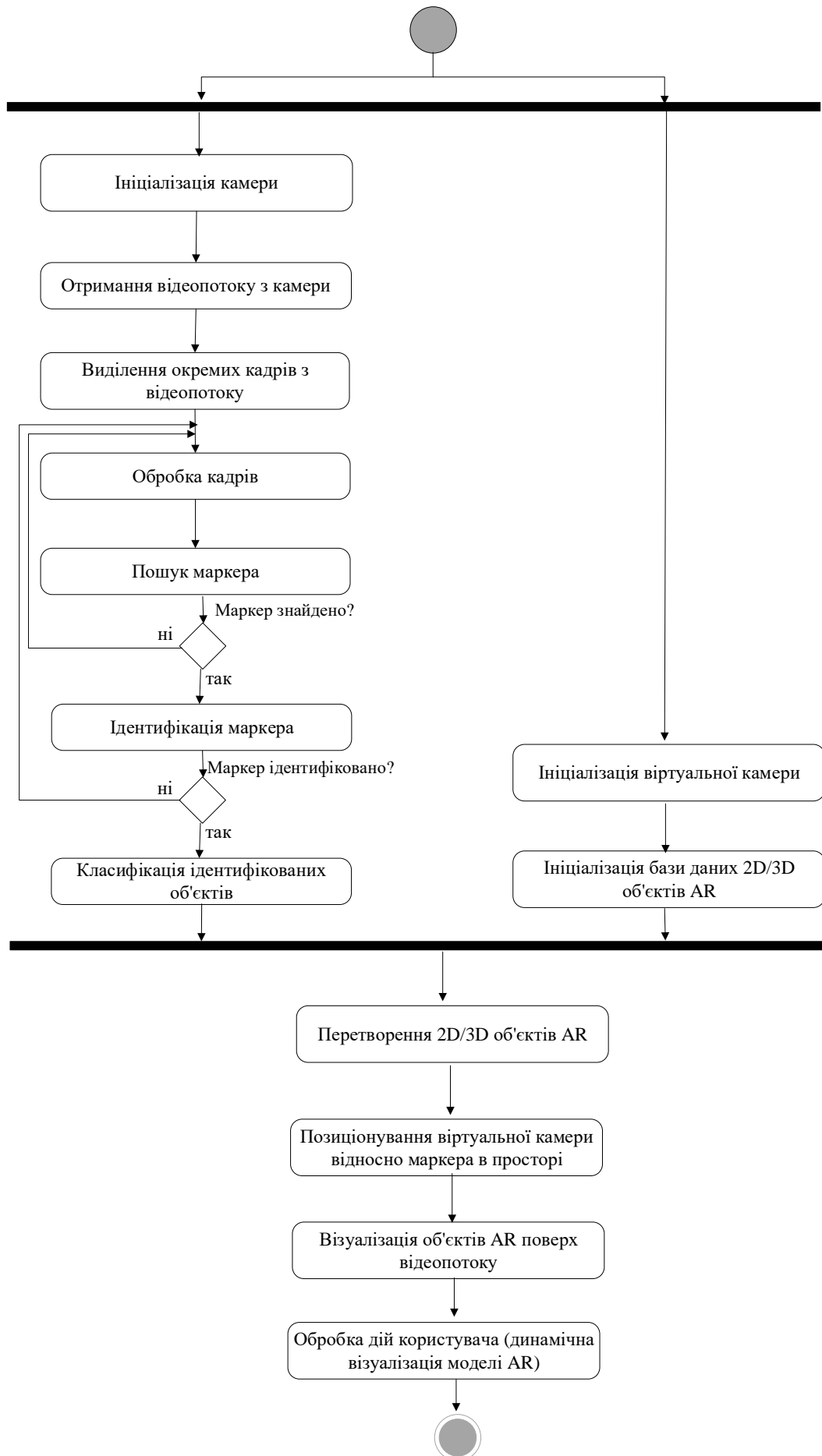


Рисунок 4.1 – Алгоритм роботи системи доповненої реальності

Розглянемо роботу даного алгоритму. В алгоритмі паралельно виконуються групи процесів роботи із зображеннями відеопотоку та 2D/3D AR об'єктами для побудови доповненої реальності.

На початку роботи програми відбувається ініціалізація камери пристрою, отримання відеопотоку з даної камери та виділення із потоку окремих кадрів для подальшої роботи з ними. Далі здійснюється обробка кадрів, тобто на виділеному кадрі відбувається пошук маркера. Якщо маркер знайдено, відбувається його ідентифікація, інакше застосування виділяє новий кадр. Якщо маркер не вдалося ідентифікувати, програма також повертається до кроку отримання кадру з відеопотоку. У разі успішної ідентифікації маркера відбувається класифікація ідентифікованих об'єктів (з використанням нейронних мереж), на основі якої далі здійснюється перетворення віртуальних об'єктів та позиціонування віртуальної камери. При успішному розпізнаванні маркера однозначно визначається положення об'єкта в просторі.

Паралельно виконується ініціалізація віртуальної камери, ініціалізація спеціальної бази даних, що містить 2D/3D об'єкти доповненої реальності для побудови AR.

Далі здійснюється позиціонування віртуальної камери щодо маркера у просторі. В результаті положення віртуальної камери щодо об'єкта і реальної камери щодо маркера відео потік синхронізується. Здійснюється візуалізація віртуальних об'єктів, що накладаються на зображення реального оточення поверх відеопотоку (на сцену міститься необхідний AR об'єкт, використовуючи заздалегідь визначені параметри ракурсу, перспективи, положення, масштабу та повороту).

Наступним кроком обробляються дії користувача (при користувальницькій взаємодії використовуються механізми маніпуляції тривимірними AR об'єктами, наприклад, використання сенсорних екранів або жестів). Щоб візуалізувати об'єкт в іншому положенні, користувачеві необхідно пересунути зображення-маркер відповідним чином. У разі потреби програма завершує свою роботу.

## 4.2 Структура застосування доповненої реальності

Структура застосування доповненої реальності, яка служить для моделювання або візуалізації складних процесів та об'єктів, представлена на рисунку 4.2 і складається з чотирьох основних модулів:

- модуль відстеження камери;
- модуль зберігання 3D об'єктів;
- модуль візуалізації;
- модуль інтерфейсу користувача.

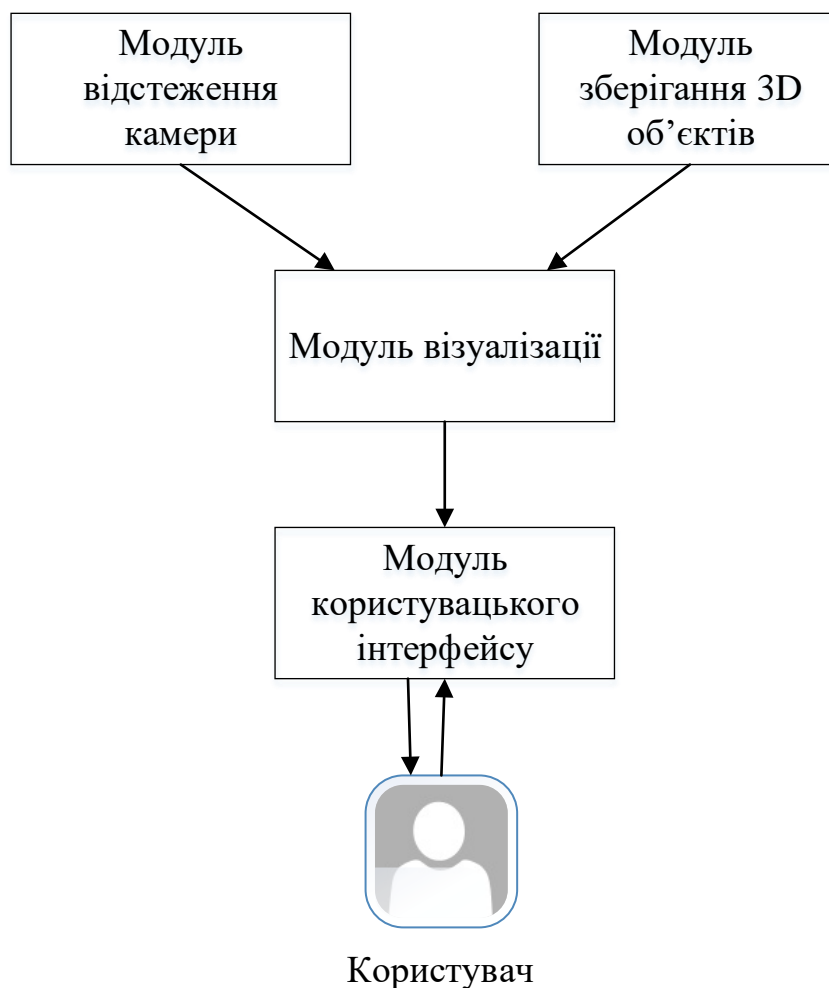


Рисунок 4.2 – Структура застосування доповненої реальності

Відеопотік із камери пристрою передається в модуль відстеження камери. Даний модуль обробляє кожен кадр відеопотоку: виконує пошук заданого маркера заздалегідь, визначає положення маркера в просторі і на основі цих даних обчислює положення віртуальної камери щодо маркера. Після того, як положення та орієнтацію камери визначено, модуль зберігання об'єктів поміщає на сцену необхідний об'єкт для візуалізації, використовуючи наперед визначені параметри положення, масштабу та повороту. Далі відбувається візуалізація моделі з використанням зсуву по відношенню до розрахункової позиції та узгодження з орієнтацією. Користувач може впливати на параметри моделі та візуалізації за допомогою інтерфейсу користувача (маніпуляція з AR об'єктом).

#### 4.3 Розробка системи доповненої реальності

В основі будь-якого застосування доповненої реальності, що використовує аналіз зображення, що надходить з камери, лежить система комп'ютерного зору. Однією з найвідоміших бібліотек, що реалізують подібний функціонал, є OpenCV (Open Source Computer Vision Library). Вона надає достатньо низькорівневих можливостей і добре підходить для отримання максимуму інформації з зображення. Оскільки питання продуктивності є досить гострим у AR, бібліотека написана мовою C++. Крім швидкості виконання, це також дає можливість застосування на різних платформах [33].

OpenCV випускається під ліцензією BSD з відкритим вихідним кодом, внаслідок чого її можна використовувати як для академічних цілей, так і для комерційних. Багато проєктів, що стосуються комп'ютерного зору, користуються саме цією бібліотекою.

Для зостаєвань доповненої реальності потрібно швидко та якісно знайти у кадрі обмежений набір заздалегідь відомих об'єктів та відобразити поверх зображення віртуальний об'єкт [34].

До складу бібліотеки комп'ютерного зору входять модулі [35]:

- обробки зображень;
- побудови простих інтерфейсів користувача, завантаження/збереження відео та аудіо даних;
- аналізу руху та відстеження об'єктів (оптичний потік, шаблони руху, усунення фону);
- калібрування камери, пошуку стерео-відповідностей та елементи обробки тривимірних даних;
- пошуку, аналізу та порівняння ключових точок зображень (Feature Detection and Description);
- детектування об'єктів на зображенні (вейвлет Хаара);
- методи та моделі машинного навчання (дерева прийняття рішень).

AForge.NET є C# фреймворком з відкритим вихідним кодом, створеним для розробників та дослідників у галузі комп'ютерного зору та штучного інтелекту. Фреймворк включає такі компоненти:

- AForge.Imaging – обробка зображень та набір різних фільтрів;
- AForge.Vision – набір методів та алгоритмів комп'ютерного зору;
- AForge.Video – обробка відеопотоку;
- AForge.Neuro – побудова та робота з нейронними мережами;
- AForge.Genetic – набір генетичних алгоритмів;
- AForge.Robotics – спеціальний набір методів для застосування в галузі робототехніки.

Для побудови системи доповненої реальності скористаємось можливостями бібліотеки OpenCV та фреймворку AForge.Neuro. За допомогою бібліотеки буде здійснюватись пошук маркера на відеопотоці з подальшим відображенням віртуального об'єкта. Сам процес пошуку здійснюється за допомогою фреймворку AForge.Neuro та нейронних мереж (за допомогою класифікації об'єктів та накладення на них маркера). Спочатку необхідно побудувати 3D простір за знайденою 2D гомографією (матриця побудови взаємно-однозначного відображення проєктивного простору на себе, що переводить точки в точки, прямі в прямі і зберігає відношення

інцидентності точок і прямих, а також подвійне відношення будь-якої четвірки колінеарних точок до єдиної перспективи та геометрії).

Потім, скориставшись методами бібліотеки OpenGL, можна намалювати віртуальний об'єкт поверх маркеру.

Процес знаходження параметрів камери називається її калібруванням. Для його проведення необхідно зробити знімки калібрувального шаблону (на практиці шахової дошки, що представляє маркер) за допомогою камери, перенести фотографії на комп'ютер і зробити підрахунок параметрів матриці. З метою отримання якомога точніших параметрів необхідно дотримання наступних умов [36]:

- шаблон шахівниці роздруковується на чистому аркуші формату А4 (маркер);

- розмір знімків із камери повинен бути приведений до розміру кадру відеопотоку. У деяких випадках можлива ситуація, коли роздільна здатність фотографій вище роздільної здатності відео. Тому перед калібруванням необхідно зменшити знімки до потрібного розміру;

- кількість знімків шаблону має бути не меншою за 10. Виконати їх потрібно з різних ракурсів. Чим більше буде знімків, тим більш точні будуть отримані параметри проєкції матриці, а ця точність далі буде впливати на наявність/відсутність зрушень при побудові 3D об'єктів.

Зовнішня матриця (або матриця моделі) – це матриця перетворень моделі за допомогою розтягування, повороту та перенесення. Вона дозволяє однозначно визначити положення об'єкта у просторі:

$$\begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{bmatrix}$$

Діагональні елементи відповідають за розтягнення моделі. Інші елементи  $r$  - за поворот об'єкта в просторі. Елементи  $t$  відповідають за перенесення. Обчислення її провадиться шляхом визначення чотирьох пар точок об'єкта та відповідного його положення в кадрі. Точки положення в кадрі – це вершини чотирикутника, що описує

(обмежує) об'єкт у кадрі. Отримати ці точки можна, якщо вплинути на крайні точки шаблону гомографією.

Сконструйовані матриці потім використовуються для обчислення вектора повороту та перенесення з наступним занесенням у матрицю моделі. Слід зазначити, що процес обчислення матриці моделі займає кілька мілісекунд і не впливає на швидкість роботи програми. Основне навантаження йде в області аналізу відеопотоку та пошуку на ньому шаблонного зображення [37–38].

Описаний вище підхід можна застосувати до всіх бібліотек комп'ютерного зору щодо побудови доповненої реальності. З практичної точки зору він виглядає дещо скрутним через необхідність проведення калібрування камери для отримання коректної матриці проектування. Одним із способів вирішення цієї проблеми є створення набору параметрів різних камер; Залежно від використовуваної камери буде завантажувати необхідні параметри.

Розвиваючи цей напрямок, можна побудувати свій власний фреймворк доповненої реальності. Розробляючи різні способи фільтрації точок, можна домогтися того, що маркер перебуватиме навіть при сильному куті нахилу та великому видаленні. Це дозволить отримати більш реалістичну картинку та дасть можливість користувачам розглянути модель з усіх боків на різній відстані.

У програмному коді використано бібліотеку `TensorFlow` для обчислення градієнтів та навчання нейронних мереж. Приклад класифікації об'єктів за допомогою нейронних мереж наведено на рисунку 4.3



{vase 87%, furniture wall 91%, TV (98%), book shelf (99%), curbstone (98%), chandelier (99%), chair (99%), notebook (99%), flower (99%)}

Рисунок 4.3 – Приклад класифікації об'єктів  
за допомогою запропонованих нейронних мереж

Приклади побудови і реалізації доповненої реальності для дизайну інтер'єру  
представлені на рисунках 4.4– 4.10



Рисунок 4.4 – Приклад проекту дизайну інтер'єру



Рисунок 4.5 – Приклад побудови дизайну інтер'єру за допомогою AR



Рисунок 4.6 – Приклад побудови дизайну інтер'єру за допомогою AR

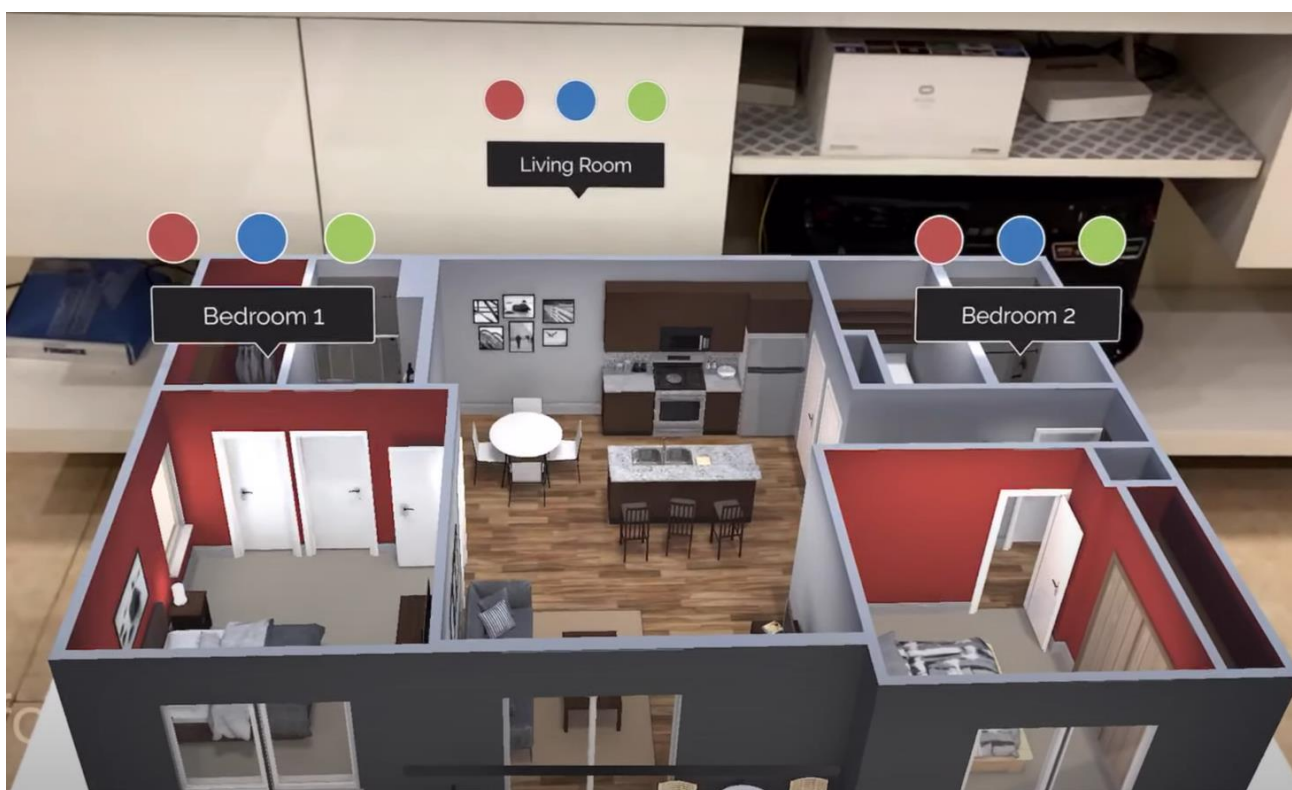


Рисунок 4.7 – Приклад 3D візуалізації об'єктів за допомогою AR



Рисунок 4.8 – Приклад 3D візуалізації об'єктів за допомогою AR



Рисунок 4.9 – Приклад 3D візуалізації об'єктів за допомогою AR



Рисунок 4.10 – Приклад 3D візуалізації об'єктів за допомогою AR

## ВИСНОВКИ

Кваліфікаційна робота виконано згідно з методичними вказівками до виконання та захисту кваліфікаційної роботи другого (магістерського) рівня вищої освіти [39].

Метою даної роботи є дослідження методів та технології реалізації систем доповненої реальності, дослідження методів класифікації об'єктів у режимі реального часу та методів побудови маркерної системи доповненої реальності для швидкого, мобільного і бюджетного розміщення AR-контенту, динамічної візуалізації об'єктів, занурення користувача у фізичний світ з реальними і доповненими об'єктами.

Для виконання поставленої мети в роботі:

- досліджено технологію віртуальної реальності та її різновиди (технологію доповненої реальності, змішаної реальності);
- досліджено напрями використання технології доповненої реальності (маркерна, безмаркерна, просторова) та обґрунтовано використання маркерів в AR;
- досліджено методи машинного навчання для побудови систем доповненої реальності, обґрунтовано використання нейронних мереж;
- сформувано вимоги до реалізації системи доповненої реальності;
- обрано тип і архітектуру нейронних мереж (згортова і рекурентна нейронна мережа), обґрунтовано вибір алгоритмів навчання (алгоритм методу зворотного поширення помилки);
- побудовано штучну нейронну мережу (у вигляді комбінації двох мереж RNN-CNN) для реалізації системи доповненої реальності та виконати її навчання;
- представлено структурну схему комбінації нейронних мереж для реалізації системи доповненої реальності та візуалізації об'єктів;
- представлено модель маніпуляції AR об'єктами;
- виконано моделювання розробленої штучної нейронної мережі та представлено програмну реалізацію системи. Технологію створення та реалізації доповненої реальності розглянуто на прикладі планування дизайну інтер'єру.

Представлена в роботі модель комбінації нейронних мереж RNN-CNN у сукупності із методами і бібліотеками комп'ютерного зору дозволяє працювати із зображеннями, відео і 3D-об'єктами, класифікувати їх та доповнювати об'єкти реального світу штучно побудованими.

Результати роботи представлені на Міжнародному молодіжному форумі «Радіоелектроніка та молодь у XXI столітті» [40].

## ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАНЬ

1. Гвоздинський, А.М. Дослідження методів оптимізації в задачах прийняття рішень в багатокритеріальних системах [Текст] / А.М. Гвоздинський, А.В. Піддубний // Біоніка інтелекту. - 2013.- № 2 (81). -С. 71-74.
2. Коренберг, В.М. Доповнена реальність як новий інтерфейс взаємодії людини з комп'ютером [Текст] / В.М. Коренберг, М.Л. Щелкунов, Л.А. Котюжанський // Журн. вираховув. математики.- 2014. -М. №10. - С. 15-18.
3. Як віртуальна реальність покращує реальний світ. Технології розширеної реальності» [Електронний ресурс] / Chip. 2011. №12. - Режим доступу: www / URL: [http://www.chip.ua/stati/go-digital/2011/12/как-virtualnaya-realnost-uluchshaet-realnyi-mir/article\\_view/](http://www.chip.ua/stati/go-digital/2011/12/как-virtualnaya-realnost-uluchshaet-realnyi-mir/article_view/) - 20.10.2021 р. - загл. з екрану.
4. На порозі доповненої реальності: до чого готуватися розробникам [Електронний ресурс] / Habrahabr. - Режим доступу: www / URL: [http://habrahabr.ru/blogs/augmented\\_reality/118123/](http://habrahabr.ru/blogs/augmented_reality/118123/) - 20.10.2021 р. - загл. з екрану.
5. Augmented Reality в реальному житті і в побуті [Електронний ресурс] / А - R. - Режим доступу: www / URL: <http://a--r.ru/ar-in-real-life/> - 20.10.2021 р. - загл. з екрану.
6. Кишенькове вікно в світ - доповнена реальність [Електронний ресурс] / Mobile Review. - Режим доступу: www / URL: <http://www.mobile-review.com/articles/2010/window-to-the-world.shtml> - 21.10.2021 р - загл. з екрану.
7. Принцип доповненої реальності [Електронний ресурс] / Google Sites. - Режим доступу: www / URL: <https://sites.google.com/site/nikiforova/project-definition> - 21.10.2021 р. - загл. з екрану.
8. Віртуальна реальність чи інтегральна реальність? [Електронний ресурс] / HiTech News. - Режим доступу: www / URL: <http://hi-news.ru/technology/virtualnaya-realnost-ili-integralnaya-realnost.html> - 21.10.2021 р. - загл. з екрану.
9. Змішана реальність [Електронний ресурс] / Wikipedia. - Режим доступу: www / URL: [https://ru.wikipedia.org/wiki/Змішана\\_реальность](https://ru.wikipedia.org/wiki/Змішана_реальность) - 22.10.2021 р. - загл. з екрану.

екрану.

10. Azuma's Realities [Електронний ресурс] / CiteSeerX. - Режим доступу: [www / URL: http://citeseerx.ist.psu.edu/viewdoc/summary? doi = 10.1.1.35.5387](http://www.citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.35.5387)- 22.10.2021 р. - загл. з екрану.

11. Розширення реальності [Електронний ресурс] / inAVate. - Режим доступу: [www / URL: http://www.inavate.ru/site/content/view/243/50/](http://www.inavate.ru/site/content/view/243/50/) - 22.10.2021 р. - загл. з екрану.

12. Рассел С. Штучний інтелект: сучасний підхід [Текст]: навч. посібник / С. Рассел, П. Норвіг. - 2-е изд., Перераб. - М.: Вільямс, 2006. - 1408 с.

13. Augmented Reality [Електронний ресурс] / Geekologie. - Режим доступу: [www / URL: http://www2.layar.com/](http://www2.layar.com/) - 22.10.2021 р. - загл. з екрану.

14. Augmented Reality Browser: Layar [Електронний ресурс] / Layar. - Режим доступу: [www / URL: http://www.geekologie.com/2008/12/14-week](http://www.geekologie.com/2008/12/14-week) - 22.10.2021 р. - загл. з екрану.

15. Exploring Visuo-Haptic Mixed Reality [Електронний ресурс] / Naist. - Режим доступу: [www / URL: www.imd.naist.jp/imdweb / pub / sandor / paper](http://www.imd.naist.jp/imdweb/pub/sandor/paper) - 22.10.2021 р. - Загл. з екрану.

16. An Augmented Reality System of Linked Audio [Електронний ресурс] / UIUC. - Режим доступу: [www / URL: social.cs.uiuc.edu/ people / HearThereICAD](http://social.cs.uiuc.edu/people/HearThereICAD) - 22.10.2021 р. - Загл. з екрану.

17. Adaptive line tracking with multiple hypotheses for augmented reality [Електронний ресурс] / ACM. - Режим доступу: [www / URL: http://dl.acm.org/citation.cfm?id=1105176](http://dl.acm.org/citation.cfm?id=1105176) - 22.10.2021 р. - Загл. з екрану.

18. Viola, P. Robust Real-time Object Detection / P. Viola, M. Jones // International Journal of Computer Vision. - 2001.

19. Viola, P. Rapid object detection using a boosted cascade of simple features / P. Viola, M. Jones // Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2001). - 2001. - Vol. 1. - P. 511-518.

20. Chen, S. Boosted road sign detection and recognition / S. Chen, J. Hsieh // 2008 Int. Conf. Mach. Learn. Cybern. - 2008. - № July. - P. 3823-3826.
21. Broggi, A. Vehicle detection for autonomous parking using a Soft-Cascade AdaBoost classifier / A. Broggi et al. // IEEE Intelligent Vehicles Symposium. - 2014. - P. 912-917.
22. Freund, Y. Experiments with a New Boosting Algorithm / Y. Freund, RE Schapire // International Conference on Machine Learning. - 1996. - P. 148-156.
23. Viola P., Jones M. Rapid object detection using a boosted cascade of simple features; Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition; Kauai, HI, USA. 8-14 December 2001.
24. Гришанов К.М., Белов Ю.С., Модель сверточное нейронної мережі в задачах машинного зору // Електронний журнал: наука, техніка і освіта 2017, №СВ1 (11), С. 100 - 106.
25. Mordini, Emilio and Massari, Sonia. "Body, Biometrics and Identity." Bioethics. Vol. 22, Issue 9, 2008. pp. 488-498.
26. Ouerhani Y., Alfalou A., Brosseau C. Optics and Photonics for Information Processing XI. Volume 10395. SPIE; Bellingham, WA, USA: 2017. Road mark recognition using HOG-SVM and correlation. International Society for Optics and Photonics.
27. Shah JH, Sharif M., Raza M., Azeem A. A Survey: Linear and Nonlinear PCA Based Face Recognition Techniques. Int. Arab J. Inf. Technol. 2013; 10: 536-545
28. Cudic M., Департамент Electrical and Computer Engineering University Florida Gainesville, FL USA За допомогою Recurrent Kernel Learning Machine for Small-Sample Image Classification. – URL: <https://ufdc.ufl.edu/AA00063111/00001> (дата звернення 12.11.2019).
29. Kan Li, Jose C. Adaptive Autoregressive-Moving-Average Algorithm IEEE Neural networks and learning systems, vol. 1, no. 1, January 2015. - URL:[https://www.researchgate.net/profile/Jose\\_Principe/publication/275717876](https://www.researchgate.net/profile/Jose_Principe/publication/275717876).
30. Белов Ю.С., Дьомін І.С., Особливості обробки супутникових знімків для навчання згорткової нейронної мережі // Електронний журнал: наука, техніка та освіта. 2019. №1 (23). С. 62-69.

31. Гришанов К.М., Белов Ю.С. Модель згорткової нейронної мережі у завданнях машинного зору // Електронний журнал: наука, техніка та освіта. 2017. № СВ1 (11). 3. 100-106.
32. Тай З.Л., Романов А.К., Ткаченко А.В., Рибкін С.В. Виявлення доріг за супутниковими знімками з використанням нейронних мереж // Електронний журнал: наука, техніка та освіта. 2019. № 2 (24). С. 93-98.
33. P. Milgram, F. Kishino, Taxonomy of Mixed Reality Visual Displays // IEICE Trans. Information Systems. - 1994. - No 12. - P. 1321-1329.
34. R. Azuma, A Survey of Augmented Reality // In Presence: Teleoperators and Virtual Environments. - 1997. - No 4. - P. 355-385.
35. І. Р. Захарова, Інформаційні технології освіти. М.: Видавничий центр "Академія", 2003. - С. 7-60.
36. Б.С. Яковлев, С.І. Пусто, Класифікація та перспективні напрямки використання технології доповненої реальності // Вісті ТулДУ. Технічні науки. - 2013. - № 3. - С. 484-492
37. Прокачай свою реальність [Електронний ресурс] // Режим доступу: <https://haker.ru/2011/08/24/56240/>. - Загл. з екрану. – (Дата звернення: 14.11.2021).
38. Огляд AR-бібліотек для створення програм з доповненою реальністю [Електронний ресурс] // Режим доступу: <http://www.azoft.ru> - Загл. з екрану. – (Дата звернення: 14.11.2021).
39. Методичні вказівки щодо розробки та оформлення кваліфікаційної роботи (для студентів усіх форм навчання другого (магістерського) рівня вищої освіти спеціальності 122 Комп'ютерні науки освітньо-професійної програми «Інформаційні управляючі системи та технології») / Упоряд.: Петров К.Е., Левикін В.М., Чалий С.Ф., Євланов М.В., Саєнко В.І., Міхнов Д.К., Міхнова А.В., Чала О.В. – Харків: ХНУРЕ, 2021. – 25 с.
40. Акулова А.О., наук.керівник Кудрявцева М.С. Дослідження методів та технологій реалізації систем доповненої реальності // XXV Міжнародний молодіжний форум «Радіоелектроніка та молодь у XXI столітті». – Харків: ХНУРЕ, 7-9 квітня 2021. – С. 36-37.