

Міністерство освіти і науки України
Харківський національний університет радіоелектроніки

Факультет навчально-науковий центр заочної форми навчання
(повна назва)

Кафедра електронних обчислювальних машин
(повна назва)

КВАЛІФІКАЦІЙНА РОБОТА
Пояснювальна записка

Рівень вищої освіти другий (магістерський)

Оцінка точності методів визначення відстані до
об'єктів в умовах змінних зовнішніх факторів

(тема)

Виконав:

здобувач 2 року навчання,

групи СПзм-23-1

Юлія ПОСТЕЛЬНЯК

(власне ім'я, прізвище)

Спеціальність

123 «Комп'ютерна інженерія»

(код і повна назва спеціальності)

Тип програми освітньо-наукова

(освітньо-професійна або освітньо-наукова)

Освітня програма

Системне програмування

(повна назва освітньої програми)

Керівник: проф. Олег МІХАЛЬ

(посада, власне ім'я, прізвище)

Допускається до захисту

Завідувач кафедри ЕОМ

(підпис)

Андрій КОВАЛЕНКО

(власне ім'я, прізвище)

2025 р.

Харківський національний університет радіоелектроніки

Факультет навчально-науковий центр заочної форми навчання

Кафедра електронних обчислювальних машин

Рівень вищої освіти другий (магістерський)

Спеціальність 123 «Комп'ютерна інженерія»
(код і повна назва)

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

Освітня програма Системне програмування
(повна назва)

ЗАТВЕРДЖУЮ:

Зав. кафедри _____
(підпис)

“ _____ ” _____ 20__ р.

ЗАВДАННЯ

НА КВАЛІФІКАЦІЙНУ РОБОТУ

здобувачеві Постельняк Юлії Павлівні
(прізвище, ім'я, по батькові)

1. Тема роботи Оцінка точності методів визначення відстані до об'єктів в умовах змінних зовнішніх факторів

затверджена наказом по університету від “ 07 ” квітня 2025 р. № 53Стз

2. Термін подання здобувачем роботи до екзаменаційної комісії 16 червня 2025 р.

3. Вхідні дані до роботи Лінійка, камера, обчислювач на базі центрального та графічного процесорів, люксометр

4. Перелік питань, що потрібно опрацювати у роботі _____

Аналіз існуючих методів визначення відстані

Фізичні основи вимірювання відстані у комп'ютерному зорі

Технологічний стек

Аналіз результатів експериментів

5. Перелік графічного матеріалу із зазначенням креслеників, схем, плакатів, комп'ютерних ілюстрацій _____

Слайд-презентація - 14 слайдів _____

6. Консультанти розділів роботи (заповнюється за наявності консультантів згідно з наказом, зазначеним у п.1)

Найменування розділу	Консультант (посада, прізвище, ім'я, по батькові)	Позначка консультанта про виконання розділу	
		підпис	дата

КАЛЕНДАРНИЙ ПЛАН

№	Назва етапів роботи	Строк / терміни виконання етапів роботи	Примітка
1	Аналіз проблеми та огляд існуючих рішень	08.04.25-18.04.25	
2	Вибір технології розробки та інструментальних засобів	19.04.25-28.04.25	
3	Розробка алгоритмічного забезпечення	29.04.25-05.05.25	
4	Проведення експериментальних досліджень та оцінка впливів	06.05.25-20.05.25	
5	Оформлення матеріалів кваліфікаційної роботи	21.05.25-02.06.25	
6	Подання кваліфікаційної роботи керівникові та її попередній захист	21.05.25-02.06.25	
7	Подання кваліфікаційної роботи на рецензування	10.06.25-12.06.25	

Дата видачі завдання “ 07 ” квітня 2025 р.

Здобувач _____

(підпис)

Керівник роботи _____

(підпис)

проф. Олег МІХАЛЬ _____

(посада, власне ім'я, прізвище)

РЕФЕРАТ

Пояснювальна записка кваліфікаційної роботи: 72 с., 9 рис., 9 табл.,
1 дод., 19 джерел

FACE MESH, ВИЗНАЧЕННЯ ВІДСТАНІ, ГЕОМЕТРИЧНІ МЕТОДИ,
OPENCV, ПОЛОЖЕННЯ ГОЛОВИ, КОМП'ЮТЕРНИЙ ЗІР, ВИЯВЛЕННЯ
ОБЛИЧЧЯ, МАЛОПОТУЖНІ ПРИСТРОЇ, АНАЛІЗ ТОЧНОСТІ.

Метою кваліфікаційної роботи є розробка та оцінка методу визначення відстані до об'єкта за допомогою геометричних ознак в умовах змінних зовнішніх факторів, таких як освітлення, роздільна здатність веб-камери та кут нахилу обличчя.

У ході виконання кваліфікаційної роботи було проаналізовано сучасні методи визначення глибини сцени, зокрема геометричні, оптичні та нейромережеві підходи. Запропоновано ефективну реалізацію на основі алгоритму FaceMesh та бібліотеки OpenCV, що не потребує навчання на великих датасетах і може функціонувати на слабких пристроях. Проведено серію експериментів, які дозволили оцінити вплив роздільної здатності (від 640×480 до 1920×1080), рівня освітленості (< 50 лк до > 500 лк) та кута нахилу обличчя (0° , 15° , 30°) на точність визначення відстані. Найвищу точність вимірювання 100% було отримано при відстані 60 см та роздільній здатності 640×480 . У разі зміни кута нахилу у фронтальній площині до 30° , точність знижувалася до 90,53%, а в сагітальній площині – до 76,13%. Результати доводять доцільність використання геометричного підходу у задачах, що потребують інтерпретованості, низького енергоспоживання та роботи в реальному часі.

ABSTRACT

Master's thesis: 72 pages, 9 figures, 9 tables, 1 appendix, 19X sources.

FACEMESH, DISTANCE ESTIMATION, GEOMETRIC METHODS, OPENCV, HEAD POSE, COMPUTER VISION, HUMAN FACE DETECTION, LOW-RESOURCE DEVICES, ACCURACY ANALYSIS.

The major goal of this thesis is to develop and evaluate a method for determining the distance to an object based on geometric features under variable external conditions, such as lighting, camera resolution, and head tilt angle.

In order to achieve this goal, modern depth estimation approaches — including geometric, optical, and neural-network-based methods — were analyzed. An efficient implementation was proposed using the FaceMesh algorithm and the OpenCV library, which does not require large-scale training datasets and is suitable for low-power devices. A series of experiments was conducted to assess the influence of resolution (from 640×480 to 1920×1080), illumination level (from <50 lux to >500 lux), and head tilt angle (0° , 15° , 30°) on distance measurement accuracy. The highest accuracy of 100% was achieved at a distance of 60 cm with a resolution of 640×480 . In the case of a frontal head tilt of 30° , the accuracy dropped to 90.53%, while in the sagittal plane, it reached a minimum of 76.13%. These results confirm the feasibility of using geometric approaches for tasks requiring interpretability, low energy consumption, and real-time performance.

ЗМІСТ

СКОРОЧЕННЯ ТА УМОВНІ ПОЗНАКИ	8
ВСТУП	10
1 АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ	11
1.1 Актуальність дослідження	11
1.2 Аналіз впливу змінних зовнішніх факторів на точність визначення відстані	16
1.3 Аналіз існуючих методів визначення відстані	19
1.4 Прогалини сучасних досліджень в області визначення відстані до	21
1.5 Постановка задач та мети дослідження	23
2 ТЕОРЕТИЧНІ ОСНОВИ ТА ВИБІР МЕТОДІВ ДОСЛІДЖЕННЯ	25
2.1 Фізичні основи вимірювання відстані у комп'ютерному зорі	25
2.1.1 Геометричні методи	26
2.1.2 Оптичні методи	29
2.1.3 Методи, що враховують властивості матеріалів	29
2.2 Нейромережеві методи визначення глибини об'єктів	30
2.3 Обґрунтування обраного методу визначення відстані до об'єкту	32
2.4 Вибір технологій та інструментів	34
3 МЕТОДИКА ПРОВЕДЕННЯ ЕКСПЕРИМЕНТІВ	37
3.1 Обраний технологічний стек	37
3.2 Теоретичне пояснення запропонованого рішення	38
3.3 Порядок проведення експериментів	48
4 АНАЛІЗ РЕЗУЛЬТАТІВ ЕКСПЕРИМЕНТІВ	50
4.1 Точність вимірювання у залежності від відстані до об'єкта при різній роздільній здатності веб-камери	50
4.2 Точність визначення відстані у залежності від відстані до об'єкта при змінному освітленні	53
4.3 Точність визначення відстані у залежності від кута нахилу голови	56

ВИСНОВКИ.....	61
ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ	62
ДОДАТОК А Графічний матеріал кваліфікаційної роботи.....	65

СКОРОЧЕННЯ ТА УМОВНІ ПОЗНАКИ

- КТ – комп’ютерна томографія
- MPT – магнітно-резонансна томографія
- ADAS – Advanced Driver Assistance Systems – система допомоги водієві
- AR – Augmented Reality – доповнена реальність
- BSD – Blind Spot Detection – виявлення «мертвих зон»
- CNN – Convolutional Neural Network – згортова нейронна мережа
- CPU – Central Processing Unit – центральний процесор
- DORN – Deep Ordinal Regression Network – глибока мережа порядкової регресії
- DPT – Dense Prediction Transformer – трансформер для щільного прогнозування
- GCNet – Guided Confidence Network – мережа з керованою впевненістю
- GDPR – General Data Protection Regulation – Загальний регламент захисту даних
- GPU – Graphics Processing Unit – графічний процесор
- HCI – Human-Computer Interaction – взаємодія людини з комп’ютером
- ID – Identifier – ідентифікатор
- KNN – K-Nearest Neighbors – алгоритм k найближчих сусідів
- KITTI – Karlsruhe Institute of Technology and Toyota Technological Institute – набір даних для автономного водіння
- LED – Light Emitting Diode – світлодіод
- LiDAR – Light Detection and Ranging – лазерна система вимірювання відстаней
- MiDaS – Mixed Depth and Scale – модель оцінки глибини сцени
- NASA – National Aeronautics and Space Administration – Національне управління з авіації і дослідження космічного простору
- NeRF – Neural Radiance Fields – нейронні радіаційні поля

ONNX – Open Neural Network Exchange – відкритий формат обміну нейронними мережами

OpenCV – Open Source Computer Vision Library – бібліотека комп'ютерного зору з відкритим кодом

ORB – Oriented FAST and Rotated BRIEF – алгоритм виявлення та опису ключових точок

PCL – Point Cloud Library – бібліотека для обробки хмар точок

PSMNet – Pyramid Stereo Matching Network – стереомережа з пірамідальним узгодженням

RGB – Red Green Blue – колірна модель (червоний, зелений, синій)

ROI – Region of Interest – область інтересу

SIFT – Scale-Invariant Feature Transform – масштабно-інваріантне перетворення ознак

SURF – Speeded Up Robust Features – пришвидшене стійке виявлення ознак

SVM – Support Vector Machine – метод опорних векторів

ToF – Time of Flight – метод вимірювання глибини за часом проходження світла

VR – Virtual Reality – віртуальна реальність

ВСТУП

Сучасні задачі комп'ютерного зору передбачають здатність систем точно визначати відстань до об'єктів у реальному часі, що є критично важливим для автономних транспортних засобів, систем безпеки, медичних застосунків, промислової автоматизації та інтерактивних інтерфейсів. Однак точність методів визначення відстані значною мірою залежить від зовнішніх умов, таких як освітлення, наявність шумів, рухомі об'єкти та зміни середовища. Це створює потребу у дослідженні стійкості різних підходів до вимірювання глибини, особливо у варіативних умовах, наближених до реальних.

Під час проходження передатестаційної практики основну увагу було зосереджено на аналізі існуючих методів оцінки відстані, вивченні технологічного стеку для побудови відповідних рішень, розробці критеріїв оцінки точності, а також експериментальному дослідженні впливу зовнішніх факторів на якість розрахунків. Практика проводилася з використанням бібліотеки OpenCV та геометричних методів (зокрема, FaceMesh), що дозволяють оцінювати відстань до об'єктів на основі масштабних змін відомих контрольних точок.

Дослідження має прикладне значення для розробки адаптивних рішень, які можуть ефективно працювати на недорогому апаратному забезпеченні, включаючи мобільні пристрої, з перспективами застосування в інклюзивних технологіях, інтелектуальних асистентах, системах навігації для слабозорих тощо.

1 АНАЛІЗ ПРЕДМЕТНОЇ ОБЛАСТІ

1.1 Актуальність дослідження

Визначення відстані до об'єктів має широке застосування у автономних системах, безпеці, медицині, військових технологіях та індустрії розваг [1-3]. Це впливає на ефективність роботи алгоритмів, точність та безпеку систем, інтеграцію сучасних технологій у повсякденне життя. Подальший розвиток цієї галузі дозволить покращити адаптивність розпізнавання, а також зменшити похибки у змінних умовах зовнішнього середовища.

Зрозуміти важливість та вплив методів визначення відстані до об'єктів на рішення більш глобальних задач, можливо проаналізувавши приклади їх застосування на практиці. Деякі із прикладів використання та самі задачі, для вирішення яких застосовуються методи визначення відстані до об'єкта, наведена в таблиці 1.1.

Таблиця 1.1 – Затребуваність методів визначення відстані до об'єкта у різних практичних областях

Область застосування	Приклад використання	Задача, яка вирішується
1	2	3
Автономні транспортні засоби	автопілоти в автомобілях (Tesla, Waymo); адаптивний круїз-контроль (Adaptive Cruise Control); системи допомоги водієві (ADAS) – екстрене гальмування, утримання смуги.	безпечне керування транспортними засобами без водія; допомога водіям у паркуванні, запобіганні зіткненням; моніторинг дорожньої ситуації в реальному часі.

Продовження таблиці 1.1

1	2	3
Робототехніка та промислова автоматизація	автономні складські роботи (Amazon Robotics); роботизовані руки на виробничих лініях (Fanuc, ABB); роботи-кур'єри (Starship Technologies).	навігація автономних роботів на виробництві, складах, у лікарнях; точне розміщення та переміщення об'єктів маніпуляторами.
Розширена та віртуальна реальність (AR/VR)	відеоігри (PlayStation VR, Oculus Quest); інтерактивні навчальні програми в медицині; AR-додатки для навігації (Google ARCore, Apple ARKit).	правильне відображення 3D-об'єктів у реальному просторі; інтерактивне відстеження положення користувача.
Біометрія та безпека	Face ID (Apple); системи контролю доступу в аеропортах; автоматичне розпізнавання осіб у відеоспостереженні.	ідентифікація та аутентифікація осіб за допомогою аналізу рис обличчя, райдужної оболонки ока, безконтактний контроль доступу).
Охорона здоров'я та реабілітація	навігаційні системи для хірургів (Da Vinci Surgical System); технології супроводу слабозорі; моніторинг пацієнтів у реальному часі.	точне визначення положення пацієнта під час медичних процедур; створення допоміжних пристроїв для людей із порушеннями зору.

Продовження таблиці 1.1

1	2	3
Дослідження навколишнього середовища та автономні дрони	дрони для моніторингу сільськогосподарських полів; безпілотники для пошуково-рятувальних операцій; геодезичні системи для створення 3D-карт.	створення карт місцевості, виявлення об'єктів на великих територіях; автономне керування безпілотниками.
Військові технології та тактичні системи	системи нічного бачення та тепловізори; безпілотні наземні та повітряні апарати (Bayraktar, Reaper); автономні бойові платформи.	виявлення рухомих об'єктів на полі бою; покращення прицілювання та роботи автономних військових систем.

Більш широка класифікація областей, у яких визначення відстані до об'єктів є затребуваним, наведена на рисунку 1.1.



Рисунок 1.1 - Класифікація сфер застосування визначення відстані до об'єктів

Визначення відстані до об'єктів є критично важливим у багатьох сферах науки, техніки та повсякденного життя, адже точні вимірювання дозволяють підвищити безпеку, автоматизувати процеси, оптимізувати управління та покращити взаємодію людини з технологіями.

У безпілотних автомобілях, дронах і промислових роботах точне визначення відстані необхідне для уникнення перешкод, навігації та коригування траєкторії руху. Наприклад, у системах автопілота автомобілі використовують LiDAR, камери та радары для оцінки відстані до інших транспортних засобів, пішоходів та інфраструктури, що підвищує безпеку руху.

Інтелектуальні системи для орієнтації у просторі, такі як смарт-окуляри чи мобільні додатки, допомагають незрячим людям уникати зіткнень, використовуючи технології комп'ютерного зору для визначення перешкод на маршруті.

У медичних технологіях методи визначення відстані застосовуються в системах моніторингу життєво важливих параметрів пацієнтів, у медичних зображеннях (МРТ, КТ) для визначення глибини пошкоджень тканин, а також у системах безконтактного вимірювання температури або дистанційного моніторингу стану пацієнта.

У системах розумного відеоспостереження аналіз відстані дозволяє оцінювати рівень загрози, відстежувати пересування об'єктів у критичних зонах та ідентифікувати потенційно небезпечні ситуації. Наприклад, автоматичне розпізнавання агресивних дій або занадто близького наближення до певних об'єктів у режимі реального часу.

У технологіях AR/VR визначення відстані допомагає створювати реалістичні віртуальні середовища та правильно позиціонувати об'єкти у просторі. Це застосовується у відеоіграх, навчальних симуляторах, інженерному моделюванні та навіть у хірургічних тренажерах.

Контроль відстані до об'єктів у промисловості використовується у виробничих лініях для автоматизованої перевірки деталей, точного

розташування механізмів та контролю допусків у процесі складання. Це знижує похибки та підвищує якість продукції.

Безпілотники та супутникові технології дозволяють визначати відстань до рослин, рельєфних особливостей ґрунту, оцінювати стан полів та прогнозувати врожайність. У поєднанні з штучним інтелектом це дає змогу оптимізувати використання ресурсів, таких як вода та добрива.

У літакобудуванні, морському транспорті та дослідженні космосу точне вимірювання відстані необхідне для безпечного приземлення, навігації та аналізу рельєфу. Наприклад, NASA використовує методи оцінки глибини для посадки космічних апаратів на інші планети.

Лазерні далекоміри та 3D-сканування допомагають визначати відстані між об'єктами у будівництві, перевіряти точність проектування та створювати цифрові моделі інфраструктури.

У спортивній аналітиці методи визначення відстані застосовуються для аналізу траєкторії руху спортсменів, оцінки їхньої швидкості та точності виконання рухів, що допомагає вдосконалювати тренувальні процеси [4].

Таким чином, методи визначення відстані до об'єктів мають суттєвий вплив на підвищення рівня автоматизації та безпеки в різних сферах. Завдяки точному вимірюванню відстані можливе покращення адаптивності систем машинного зору, що використовуються в автономному транспорті, медицині, промисловості та інших галузях. Такі технології сприяють зменшенню ризику аварій та помилок у критичних системах, забезпечують підвищену точність моніторингу та діагностики, а також дозволяють створювати більш реалістичні цифрові моделі та симуляції. Використання глибоких нейромереж у поєднанні з класичними методами комп'ютерного зору дозволяє підвищити ефективність оцінки глибини сцени навіть у складних умовах змінного освітлення та перешкод. Окрім цього, інтеграція таких методів у пристрої з обмеженими обчислювальними ресурсами, наприклад, мобільні телефони чи вбудовані системи, відкриває нові можливості для розвитку доступних технологій для широкого кола користувачів.

Проте, слід зауважити, що методи визначення відстані до об'єктів для indoor (внутрішніх приміщень) та outdoor (зовнішнього середовища) застосування значно відрізняються через особливості умов експлуатації, типових перешкод та вимог до точності. Основні відмінності стосуються технологій, що використовуються, чутливості до зовнішніх факторів, а також точності та діапазону вимірювань.

1.2 Аналіз впливу змінних зовнішніх факторів на точність визначення відстані

Чутливість до зовнішніх умов залежить від характеристик фото та відеореєстраторів та інших датчиків та їх адаптивності до різного освітлення, атмосферних перешкод, діапазону вимірювань тощо. Проведений аналіз відмінностей, які можуть бути присутні методам для indoor та outdoor застосування, наведені в таблиці 1.2.

Зовнішні фактори, що впливають на точність методів визначення відстані до об'єктів, відрізняються при використанні всередині приміщення та назовні.

При внутрішньому (indoor) використанні на точність визначення відстані вплив мають такі фактори:

- освітлення та його зміни;
- відбиття та прозорі поверхні;
- розмір і форма об'єктів;
- перешкоди та затінення;
- динамічні сцени;
- обмежений простір.

Яскравість, джерела світла та їх положення впливають на роботу камер та оптичних сенсорів. Надмірне або недостатнє освітлення може ускладнювати виявлення об'єктів, зменшуючи точність стереозору та алгоритмів глибини.

Таблиця 1.2 - Основні відмінності між indoor та outdoor методами визначення відстані

Параметр	Indoor (приміщення)	Outdoor (зовнішнє середовище)
Типові умови	Стабільне освітлення, відсутність атмосферних перешкод, обмежений простір	Змінне освітлення (сонце, тінь), атмосферні фактори (дощ, туман), великі дистанції
Діапазон вимірювань	0.1 м – 10 м	10 м – кілька кілометрів
Основні перешкоди	Відбиття сигналів від стін, меблів, предметів	Зміни погодних умов, великі простори, складний рельєф
Чутливість до світла	Мінімальна (можна контролювати освітлення)	Висока (залежність від часу доби, наявності прямих сонячних променів)
Типові технології	Ультразвукові, LiDAR, стереозір, структуроване світло, ToF-сенсори	LiDAR, супутникові дані, радарні технології, аерофотознімки

Дзеркальні поверхні, скло або глянцеві матеріали можуть спотворювати зображення, що ускладнює оцінку глибини за допомогою методів комп'ютерного зору (наприклад, стереозору або Monodepth).

Об'єкти можуть бути малими, нестандартної форми або мати мало унікальних візуальних особливостей, що ускладнює їх детекцію та точний розрахунок відстані.

У приміщеннях можливі випадки часткового перекриття об'єктів іншими предметами, що ускладнює розрахунок їхньої відстані.

Рухомі об'єкти можуть змінювати свою позицію між кадрами, що ускладнює аналіз глибини та розпізнавання об'єктів.

У кімнатах або коридорах глибина сцени є малою, що може призводити

до насиченості зображення переднім планом і погіршення оцінки дальніх об'єктів.

При зовнішньому (outdoor) використанні на точність визначення відстані вплив мають такі фактори:

- зміни природного освітлення (день/ніч, хмари, сонце);
- атмосферні умови (дощ, туман, сніг, пил);
- рухомі об'єкти (автомобілі, пішоходи, тварини);
- велика відстань до об'єктів;
- нерівномірний рельєф місцевості;
- вібрації та рух камери;
- наявність тіней та змін контрасту;
- засвітлення від фар або сонця (blinding effect).

Природне освітлення змінюється протягом дня та в різних погодних умовах. Надмірне світло може створювати засвітлення, а темрява зменшує ефективність алгоритмів обробки зображень.

Опади та зважені частки у повітрі знижують контрастність та якість отриманих даних, особливо для оптичних методів та стереозору.

Висока динамічність сцени може призводити до неточностей у вимірюванні відстаней, особливо при використанні методів, що потребують аналізу послідовностей кадрів (оптичний потік).

Більшість методів комп'ютерного зору менш ефективні на великих відстанях (>50 м), де точність оцінки глибини значно зменшується.

Височини, пагорби або ями можуть створювати викривлення у визначенні відстаней, особливо для методів, що базуються на плоских припущеннях (наприклад, проєктивні методи).

У випадках, коли камера встановлена на рухомому транспортному засобі або дроні, можуть виникати додаткові похибки, спричинені нестабільністю кадрів.

Тіні від об'єктів можуть спотворювати сприйняття форми та розміру, впливаючи на алгоритми оцінки глибини.

У випадках, коли в кадр потрапляють яскраві джерела світла (наприклад, фари автомобілів у нічний час), алгоритми можуть некоректно оцінювати глибину об'єктів.

Для коректного визначення відстані важливо враховувати специфіку середовища (indoor або outdoor) та його особливості. У приміщеннях найбільший вплив мають освітлення, відбиття та обмежений простір, тоді як на відкритому просторі ключовими факторами є змінні погодні умови, природне освітлення, великі відстані та рухомі об'єкти. Вибір методів залежить від цих факторів, тому ефективні рішення зазвичай поєднують різні технології (стереозір, LiDAR, нейромережеві підходи) для компенсації недоліків окремих методів.

1.3 Аналіз існуючих методів визначення відстані

Системи вимірювання відстані до об'єктів використовують різні програмні та апаратні засоби, які залежать від конкретної задачі та середовища застосування. Технологічний стек включає апаратне забезпечення, яке відповідає за збір даних, та програмні засоби, які обробляють ці дані для отримання точних результатів (рисунок 1.2).

Розглядаючи окремі програмні та апаратні засоби у контексті внутрішнього та зовнішнього застосування, слід відмітити, що, наприклад ультразвукові датчики (Sonar) використовуються для коротких відстаней (наприклад, у робототехніці, системах орієнтації для людей із порушенням зору). Є чутливими до відбиття сигналів від гладких поверхонь, що робить їх адаптованими для indoor застосування. До outdoor застосування вони є малопридатними через втрати сигналу в повітрі та вплив вітру, дощу.

Лазерні датчики (LiDAR) популярні в автономному транспорті, геодезії, екологічному моніторингу [5-6]. Можуть працювати на великих дистанціях, але вимагають високої потужності, що робить їх більш адаптованими до outdoor застосування, аніж описані вище ультразвукові

датчики. В умовах приміщень лазерні датчики використовуються для картографування простору (роботи, дрони, VR/AR), проте можуть виникати проблеми з відбиттям лазера від дзеркальних поверхонь.

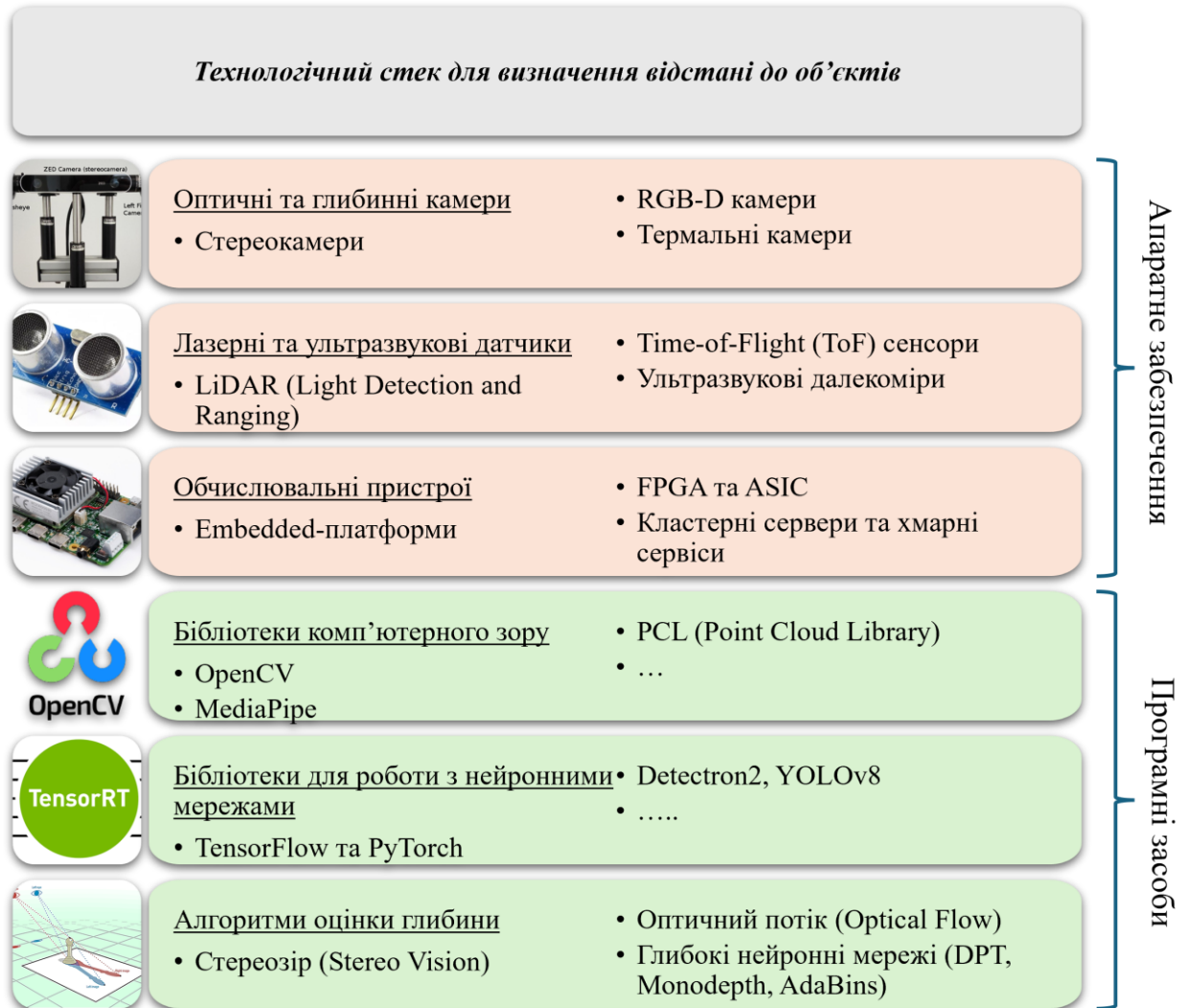


Рисунок 1.2 – Технологічний стек для визначення відстані до об'єктів

Методи стереозору (Stereo Vision) добре працюють у контрольованих умовах, широко використовується у робототехніці, системах захоплення руху, проте втрачають ефективність при зовнішньому використанні через проблеми з варіативністю освітлення.

Обчислювальні методи для визначення відстані до об'єкту, які базуються на часі прольоту (Time-of-Flight, ToF), ефективно використовуються у смартфонах, VR-системах для визначення глибини

сцени. Мають високі показники точності у межах кількох метрів. Але є чутливими до змін освітлення та атмосферних умов.

Радарні технології рідко використовуються для indoor вимірювань відстані, проте є ефективними для далеких відстаней (автомобільні системи безпеки, супутникова навігація, метеорологія).

Таким чином, технологічний стек для визначення відстані до об'єктів включає широкий спектр апаратних рішень (LiDAR, камери, ToF-сенсори) та програмних засобів (OpenCV, TensorFlow, PyTorch). Вибір технологій залежить від сфери застосування: у робототехніці та навігації використовуються LiDAR і глибокі нейромережі, у мобільних пристроях – ToF-камери та нейронні моделі оцінки глибини, а для інженерних застосувань – стереозір та PCL.

Дослідження та оптимізація програмного та апаратного забезпечення дозволяють підвищити точність оцінки відстаней і зробити ці технології доступнішими для широкого використання.

1.4 Прогалини сучасних досліджень в області визначення відстані до

Попри значний прогрес у розвитку оптичних методів визначення відстані, існує низка відкритих проблем і невирішених питань, які обмежують точність, надійність та ефективність цих методів у різних сферах застосування. Основні прогалини сучасних досліджень можна сформулювати таким чином:

- обмеження в умовах змінного освітлення;
- чутливість до атмосферних умов у зовнішніх середовищах;
- нестача якісних датасетів для навчання нейромереж;
- висока обчислювальна складність глибоких нейромережевих моделей;
- відсутність уніфікованих стандартів оцінки точності;
- обмеженість методів для динамічних сцен;

- недостатнє дослідження гібридних методів.

Багато сучасних методів, зокрема стереозір та нейромережеві підходи, демонструють зниження точності за низького рівня освітлення або в умовах різких змін освітлення (наприклад, при переході з темного приміщення на освітлену вулицю). Наявні методи адаптивної корекції контрасту (CLAFE, Equalization Histogram) лише частково компенсують цю проблему, але не можуть повністю забезпечити стійку роботу алгоритмів у варіативних умовах.

Для outdoor-застосувань (автономні транспортні засоби, дрони) серйозною проблемою залишаються погодні умови: туман, дощ, сніг, пил. Оптичні методи, особливо LiDAR, значно втрачають ефективність у таких умовах через розсіювання світла. У сучасних дослідженнях недостатньо уваги приділено методам компенсації цього впливу, наприклад, через поєднання оптичних даних з радіолокаційними чи тепловими камерами.

Більшість сучасних нейромережевих моделей оцінки глибини (Monodepth, DPT, AdaBins) навчаються на обмеженій кількості датасетів, які часто мають недостатню варіативність сцен (наприклад, KITTI Dataset фокусується на дорожніх сценах, а NYU Depth Dataset – на закритих приміщеннях). Це призводить до низької узагальнюваності моделей при застосуванні в нових середовищах. Необхідність створення нових масштабних та якісних датасетів з урахуванням складних сценаріїв (наприклад, динамічних сцен із рухомими об'єктами) є важливим викликом.

Методи глибокого навчання для оцінки відстані (DPT, NeRF, MiDaS) демонструють хороші результати, але потребують значних обчислювальних ресурсів. Це ускладнює їх впровадження на мобільних та вбудованих пристроях. Наразі не існує універсального рішення, яке б дозволяло ефективно використовувати нейромережі для оцінки глибини в реальному часі без втрати точності.

У різних дослідженнях використовуються різні метрики точності (Absolute Relative Error, Root Mean Squared Error, Structural Similarity Index),

що ускладнює порівняння методів між собою. Відсутність єдиних стандартів для оцінки якості визначення відстані призводить до неоднозначних висновків щодо ефективності певних алгоритмів у різних умовах.

Більшість сучасних оптичних методів працюють ефективно в статичних умовах або в умовах повільного руху об'єктів. Однак у динамічних сценах (наприклад, під час навігації дронів або аналізу сцени в спортивних трансляціях) точність різко падає через нечіткість руху (motion blur) та зміну перспективи. Це особливо критично для задач доповненої реальності (AR) та автономного транспорту.

Поєднання традиційних алгоритмів (стереозір, оптичний потік) з нейромережевими моделями оцінки глибини ще недостатньо вивчене. Гібридні методи можуть покращити точність за рахунок комбінування переваг обох підходів, проте у літературі відсутні масштабні порівняльні дослідження щодо оптимальних комбінацій таких методів.

Основні відкриті питання у сфері оптичного визначення відстані включають підвищення точності в умовах низького освітлення, адаптацію до погодних умов, розширення датасетів для навчання моделей, оптимізацію обчислювальних ресурсів, стандартизацію метрик оцінки та вдосконалення методів для динамічних сцен. Майбутні дослідження повинні зосередитися на розробці гібридних підходів, адаптивних методів корекції помилок та підвищенні ефективності нейромереж для роботи в реальному часі на мобільних пристроях.

1.5 Постановка задач та мети дослідження

Метою даного дослідження є розробка та аналіз методу визначення відстані до об'єкта при варіативних зовнішніх умовах за допомогою геометричних ознак.

Для досягнення поставленої мети мають бути вирішені наступні задачі:

- провести аналіз сучасних методів вимірювання відстані до об'єктів,

зокрема геометричних, оптичних та нейромережових підходів;

- визначити специфіку впливу змінних зовнішніх факторів (освітлення, роздільна здатність камери, положення голови) на точність вимірювання відстані;

- обґрунтувати вибір методу оцінки відстані на основі контрольних точок обличчя (FaceMesh), з огляду на вимоги до швидкодії та апаратних обмежень;

- реалізувати програмний модуль на основі OpenCV для визначення відстані до об'єкта з використанням методу монокулярної геометрії;

- провести серію експериментів для оцінки точності роботи методу при різних умовах (відстань, роздільна здатність, освітлення, положення об'єкта);

- узагальнити результати експериментів і визначити оптимальні умови для використання запропонованого методу;

- сформулювати рекомендації щодо подальшого використання методики у задачах адаптивної взаємодії з користувачем, наприклад у системах трекінгу погляду чи HCI-рішеннях для людей з порушеннями зору.

Подальшим шляхом розвитку дослідження є вдосконалення методу визначення відстані шляхом поєднання геометричних підходів із нейромережевими моделями оцінки глибини (наприклад, Monodepth або DPT), що дозволить підвищити точність у складних умовах, зокрема за наявності шумів, відблисків або часткового перекриття обличчя. Також перспективним є розширення алгоритму на обробку відеопотоку в реальному часі на мобільних платформах із застосуванням оптимізованих бібліотек (TensorRT, MediaPipe) та додаткове врахування індивідуальних біометричних особливостей користувачів для підвищення персоналізованості та адаптивності системи. Ще одним напрямом розвитку може стати інтеграція даного підходу у мультисенсорні системи розумного середовища або навігаційні рішення для людей із порушеннями зору.

2 ТЕОРЕТИЧНІ ОСНОВИ ТА ВИБІР МЕТОДІВ ДОСЛІДЖЕННЯ

2.1 Фізичні основи вимірювання відстані у комп'ютерному зорі

Інтеграція триангуляції, оптичного потоку та аналізу текстур дозволяє досягти високої точності у визначенні відстані, особливо в умовах обмежених даних або динамічних сцен. Сучасні дослідження зосереджені на глибокому навчанні для автоматизації цих процесів, що відкриває нові можливості для робототехніки та комп'ютерного зору.

Визначення глибини (відстані до об'єктів) є ключовим завданням у комп'ютерному зорі, робототехніці та AR/VR. Основними фізичними принципами визначення глибини зображенн є:

- геометричні властивості світлових променів;
- оптичні ефекти (наприклад, розмиття, рух);
- властивості матеріалів (наприклад, текстура).

Оглядове порівняння зазначених методів наведено у таблиці 2.1.

Таблиця 2.1 – Порівняння методів визначення глибини об'єкту

Метод	Точність	Швидкість	Обчислювальна складність	Область застосування
Геометричні	Висока	Висока	Середня	Робототехніка, AR/VR
Оптичний потік	Середня	Низька	Висока	Автономні авто
Текстура	Низька	Висока	Низька	Статичні сцени
ToF	Дуже висока	Висока	Висока	Медицина, промисловість

Таблиця надає порівняння різних методів визначення глибини об'єкта за ключовими характеристиками: точність, швидкість, обчислювальна складність та сфера застосування. Геометричні методи демонструють високу точність і швидкість, але мають середню обчислювальну складність, що

робить їх придатними для робототехніки та AR/VR. Оптичний потік забезпечує середню точність, але має низьку швидкість і високу обчислювальну складність, тому застосовується в автономних транспортних засобах. Методи на основі аналізу текстур характеризуються високою швидкістю, проте мають низьку точність і обчислювальну складність, що обмежує їх використання статичними сценами. Технологія Time-of-Flight (ToF) демонструє найвищу точність і швидкість, але є обчислювально затратною, що робить її ефективною для медицини та промисловості. Таким чином, вибір методу залежить від балансу між продуктивністю та вимогами конкретної сфери застосування.

Принцип роботи геометричних методів базується на принципах використання проєктивної геометрії для розрахунку відстані через відомі параметри камери або сцени.

2.1.1 Геометричні методи

На основі триангуляції глибина об'єкта визначається шляхом аналізу зсуву його положення відносно двох або більше точок спостереження. Обмеженнями даного методу є необхідність знати відстань між камерами або випромінювачем і приймачем, втрата точності при малому базисі (відстані між камерами), погана робота на текстурно-однорідних поверхнях [7].

На основі триангуляції, відстань до об'єкта обчислюється за формулою:

$$Z = \frac{f \times B}{d}, \quad (2.1)$$

де Z – відстань до об'єкта, яку потрібно знайти.

f – фокусна відстань камери, тобто відстань між лінзою та сенсором, яка впливає на масштаб зображення.

B – базова відстань між камерами, тобто фізична відстань між двома об'єктивами, що формують стереопару.

d – диспаратність, яка визначає, наскільки змістилося зображення

об'єкта на двох кадрах, отриманих двома камерами.

Чим далі знаходиться об'єкт, тим менша диспаратність (об'єкт займає майже однакове положення на обох зображеннях). Якщо ж об'єкт близько, його положення на лівому та правому кадрах суттєво відрізнятиметься.

Формула дозволяє розрахувати відстань Z шляхом множення фокусної відстані f на базову відстань між камерами B і поділу на диспаратність d . Таким чином, чим більша диспаратність, тим ближче об'єкт до камер, і навпаки.

Практичне застосування знаходиться в автономному транспорті (для розрахунку відстані до перешкод), в робототехніці (для навігації роботів у просторі), для медичної візуалізації (для побудови 3D-зображень), в доповненій реальності (AR) – для коректного розміщення віртуальних об'єктів у реальному просторі.

При роботі з методом структурованого світла на об'єкт або сцену проєктується контрольований світловий візерунок (наприклад, сітка, смуги, точки) за допомогою проєктора або лазера [8]. Коли світловий шаблон падає на нерівну поверхню, він деформується відповідно до форми об'єкта. Камера (або кілька камер) знімає проєктований візерунок, фіксує його зміщення та викривлення. Аналізуючи деформацію світлового візерунка, система розраховує відстань до кожної точки сцени за допомогою триангуляції. Наприклад, Microsoft Kinect використовує метод структурованого світла, де інфрачервоний проєктор проєктує крапковий візерунок, а камера зчитує його деформації для визначення глибини сцени. Основні застосування методу структурованого світла знаходяться при 3D-скануванні (промислове та медичне), в системах розпізнавання жестів (наприклад, Microsoft Kinect), в контролі якості у виробництві, в доповненій та віртуальній реальності (AR/VR), в робототехніці та автономних системах.

Поясненням принципу роботи методу є формула:

$$Z = \frac{f \times L}{x},$$

де L – відстань між проектором і камерою,
 x – зміщення шаблону.

Time-of-Flight (ToF) методи визначають відстань на основі аналізу часу, за який світловий імпульс досягає об'єкта і повертається назад до сенсора [9]. Використовуються у смартфонах для портретного режиму камер та систем розпізнавання обличчя. Сенсор випромінює інфрачервоне світло (лазерний або LED-імпульс) або радіохвилю у напрямку об'єкта. Хвиля досягає об'єкта, відбивається від його поверхні і повертається назад до сенсора. Сенсор фіксує час, за який сигнал повернувся.

Відстань вимірюється за формулою:

$$Z = \frac{c \times \Delta t}{2},$$

де c – швидкість світла,

Δt – час затримки.

Основними типами ToF-систем є Direct Time-of-Flight (dToF) та Indirect Time-of-Flight (iToF). Перший тип методів використовує короткі імпульси світла або радіохвиль, вимірює час проходження імпульсу від сенсора до об'єкта. Мають високу точність навіть на великих дистанціях. Використовується в лідарних системах (LiDAR, автономні автомобілі, роботи). Другий тип ToF-систем (Indirect Time-of-Flight (iToF)) використовує модульоване світло з відомою частотою. Аналізує зміщення фази відбитого сигналу для обчислення відстані. Використовується в 3D-камерах смартфонів (Face ID, доповнена реальність, жестове управління).

Застосування ToF-технології знаходять в лідарах в автономних автомобілях для розпізнавання перешкод, картографування навколишнього середовища, в 3D-скануванні та реконструкції для моделювання поверхонь, захоплення форми об'єктів, в смартфонах та планшетах для Face ID (iPhone), портретного режиму у камерах, в робототехніці для розпізнавання перешкод і навігації, в доповненій та віртуальній реальності (AR/VR) для глибокого аналізу сцени, відстеження руху користувача.

Недоліками ToF є чутливість до відбивних поверхонь – може давати похибки на глянцевих або прозорих поверхнях, розсіювання світла – може зменшувати точність у складних середовищах (туман, пил), обмеження через швидкість світла – може впливати на роздільну здатність при мініатюрних деталях.

2.1.2 Оптичні методи

При аналізі оптичного потоку (Optical Flow) глибина об'єкта визначається за змінами його положення в послідовних кадрах відео, враховуючи відомі швидкість руху камери або сцени [10].

Алгоритми оптичного потоку (Horn-Schunck, Lucas-Kanade) використовують зміну інтенсивності пікселів між кадрами для оцінки напряму і швидкості руху:

$$Z = \frac{v \times f}{u},$$

де v – швидкість камери,

u – швидкість оптичного потоку.

Обмеженнями є те, що методи на основі аналізу оптичного потоку погано працюють на статичних сценах, вимагають знання параметрів руху камери, мають знижену точність на гладких або однорідних поверхнях.

2.1.3 Методи, що враховують властивості матеріалів

Методи аналізу текстур дозволяють оцінити глибину сцени за допомогою аналізу змін у розподілі текстурних особливостей на поверхні об'єкта. Це корисно у випадках, коли стереозір або Time-of-Flight (ToF) недоступні [11].

Ідея полягає у тому, що ближчі об'єкти мають чіткіші текстури через більшу деталізацію. Дальші об'єкти виглядають більш розмитими через

перспективні ефекти та вплив атмосфери. Об'єкти, що знаходяться ближче, зазвичай мають вищий контраст та більш різкі границі текстур. Зі збільшенням відстані контраст зменшується через розсіювання світла в атмосфері. На ближчих об'єктах текстурні патерни розтягнуті та деталізовані. Тобто, чим далі об'єкт – тим щільніше розташовані елементи текстури та здаються меншими (ефект перспективного стиснення). Крім того, зміна кута нахилу поверхні призводить до спотворення текстур (наприклад, під кутом лінії стають більш стиснутими, що дає інформацію про нахил поверхні).

Аналіз текстур для визначення глибини використовується в автономних системах, роботах, які оцінюють глибину без використання лазерних або стереокамер, в фотографії та відеографії для визначення глибини для ефектів розмиття фону (bokeh), в геодезії та картографії для аналізу супутникових знімків для побудови топографічних карт, в медична візуалізація – реконструкція тривимірних структур на основі 2D-зображень.

2.2 Нейромережеві методи визначення глибини об'єктів

Визначення глибини об'єктів за допомогою нейромереж є актуальною дослідницькою задачею. Сучасні підходи базуються на різних архітектурах і методах навчання. Основні приклади нейромережевих моделей та їх особливості наведені в таблиці 2.2.

Таблиця 2.2 – Огляд методів машинного навчання для визначення глибини об'єкту

Методи	Характеристики
1	2
Моноокулярне визначення глибини (оцінка глибини з одного 2D-зображення без використання стереопар або додаткових датчиків)	
Monodepth	CNN-базована модель. Використовує самоконтрольоване навчання, де модель навчається прогнозувати глибину, порівнюючи ліве та праве зображення стереопари.

Продовження таблиці 2.2

1	2
DORN	CNN-базована модель. Застосовує дискретне кодування глибини через ординальну регресію.
DPT (Depth Prediction Transformer)	Має архітектуру трансформеру. Використовує Vision Transformer (ViT) для захоплення глобального контексту. Точність перевищує CNN-моделі на датасетах NYU Depth V2 і KITTI.
Depth from Motion	Модель навчається оцінювати глибину через послідовні кадри відео, використовуючи рух камери. Базується на самоконтрольованому навчанні.
Стереозір з використанням нейромереж (методи базуються на парі зображень з двох камер (стереопара) для розрахунку диспаратності)	
PSMNet (Pyramid Stereo Matching Network)	Використовує просторові піраміди для захоплення контексту на різних масштабах.
GCNet (Geometry and Context Network)	Комбінує 3D-згортки з геометричними обмеженнями.
Мультиmodalні підходи (комбінують дані з різних джерел: LiDAR, RGB-зображення, інфрачервоні датчики)	
Sparse-to-Dense	Використовує розріджені дані LiDAR для навчання моделі прогнозувати щільні карти глибини.
FusionNet	Інтегрує RGB і LiDAR через мультиmodalні згортки.
Глибоке навчання з фізичними обмеженнями (моделі інтегрують фізичні закони (наприклад, оптичний потік, геометрію) для покращення точності)	
GeoNet	Комбінує оптичний потік і визначення глибини через геометричні обмеження.
NeRF (Neural Radiance Fields)	Моделює 3D-сцену через розсіювання світла, використовуючи набір 2D-зображень.

Серед найновіших існуючих архітектур слід відзначити MobileDepth [12] - моделі для мобільних пристроїв, FastDepth [13] зі швидкою реалізацією інференсу за 10 мс на CPU, AdaBins [14] та MiDaS [15] - універсальна модель, тренувана на різних датасетах для роботи з новими середовищами.

Нейромережеві підходи досягли значних успіхів у визначенні глибини, але їх ефективність залежить від якості даних, архітектури та інтеграції фізичних принципів. Недоліками зазначених нейромережевих моделей є залежність від даних - великі розмічені датасети (наприклад, KITTI, NYU Depth V2) часто недоступні, узагальнення - моделі погано працюють на незнайомих сценах (наприклад, підводний світ), ресурси - класні моделі (наприклад, PSMNet) потребують GPU.

2.3 Обґрунтування обраного методу визначення відстані до об'єкту

Виходячи з того, що перед початком дослідження відомі параметри камери, потрібна інтерпретовність результатів та необхідно отримання рішення в найкоротший час для невизначеного кола об'єктів, використання геометричних методів є оптимальним для наших задач.

Метод FaceMesh (або аналогічні підходи з використанням попередньо визначених масштабів) ґрунтується на геометричних розрахунках, де відстань обчислюється через відомі розміри об'єкта (наприклад, середню відстань між очима людини) та їх проекцію на пікселі зображення. Цей підхід має низку переваг перед сучасними методами глибокого навчання, особливо в конкретних сценаріях.

Наприклад, мінімальна залежність від даних. Геометричні підходи, зокрема FaceMesh, не вимагають навчання на великих датасетах. Достатньо знати фізичні розміри об'єкта (наприклад, 6.5 см між очима людини) та параметри камери (фокусна відстань).

Глибоке навчання, на відміну від цього, потребує масивних розмічених

даних для тренування. Наприклад, моделі типу Monodepth [16] залежать від тисяч зображень з відомими глибинами, що часто недоступно для нішевих завдань.

З точки зору енергоефективності та швидкості, алгоритми на основі OpenCV (наприклад, виявлення ключових точок) працюють у реальному часі на слабких пристроях (смартфони, Raspberry Pi). Глибоке навчання передбачає роботу зі складними архітектурами (наприклад, DepthNet) та потребують GPU для інференсу, що збільшує вартість та енергоспоживання.

В контексті інтерпретовності та контролю, формули на основі проєктивної геометрії, на основі яких працює FaceMesh, прозорі й дозволяють точний аналіз помилок. Глибоке навчання при цьому часто являє “чорний ящик”, що не завжди дає можливість безпомилково визначити чому модель помилково оцінила відстань для нестандартних облич) [17].

У контрольованих умовах FaceMesh дає стабільні результати, якщо відомі фізичні параметри об’єкта (наприклад, розмір обличчя) та виконано калібрування камери (фокусна відстань, роздільна здатність). Глибоке навчання чутливе до змін у середовищі (наприклад, незвичне освітлення, часткове закриття об’єкта), що вимагає донавчення моделей.

І, врешті решт, неявною перевагою геометричних методів є відсутність потенційних етичних проблем - біометричні дані не зберігаються, оскільки FaceMesh працює локально та анонімно. А при глибокому навчанні, моделі, треновані на даних користувачів, можуть порушувати GDPR через збір персональної інформації (наприклад, сканування облич). [18]

В подальшому, реалізовані та досліджені методи на основі масштабу можуть бути розгорнуті на IoT-пристроях або системах з низькою потужністю, де розглядаються стандартизовані об’єкти з відомими розмірами (наприклад, обличчя людини, дорожні знаки) і де критичною є зрозумілість алгоритму (наприклад, медична діагностика).

Глибоке навчання ж виправдане для складних динамічних середовищ (наприклад, автономні авто), де точні геометричні розрахунки неможливі, але

доступні великі дані для тренування.

2.4 Вибір технологій та інструментів

Для дослідження точності роботи методів на основі геометричних характеристик об'єкту, було обрано бібліотеку OpenCV (Open Source Computer Vision Library) — це відкрита бібліотека для комп'ютерного зору, машинного навчання та обробки зображень. Розроблена Intel у 2000 році, вона підтримується спільнотою та використовується в науці, промисловості та розвагах. З того часу стала де-факто стандартом для побудови рішень у сфері комп'ютерного зору.

Бібліотека OpenCV — одна з найпотужніших і найпопулярніших платформ для обробки зображень, комп'ютерного зору та машинного навчання [19]. Вона є кросплатформенною, має відкритий код та підтримує мови програмування C++, Python, Java, а також інтерфейси для MATLAB, Go, Julia тощо.

Основні характеристики:

- підтримувані мови - C++, Python, Java, JavaScript;
- крос-платформенність - Windows, Linux, macOS, Android, iOS;
- ліцензія - Apache 2 / BSD;
- можливості - реалізація 2500+ алгоритмів для обробки зображень, відео, 3D-реконструкції, AR та ін.

OpenCV організована у модулі, кожен з яких відповідає за певний аспект комп'ютерного зору (таблиця 2.3).

OpenCV містить широкий набір методів для визначення глибини об'єкта, включаючи:

- стереозір (Stereo Vision) – використання двох камер для визначення диспаратності між двома зображеннями.

- оптичний потік (Optical Flow) – аналіз зміщення пікселів між послідовними кадрами відео.

- нейромережеві методи (Monodepth, DPT) – оцінка глибини з одного зображення за допомогою CNN.

- Face Mesh та глибина (Depth Estimation from Face Landmarks) – визначення відстані на основі розташування ключових точок обличчя.

Таблиця 2.3 - Організація бібліотеки OpenCV

Модуль	Функціонал
core	Базові структури даних (матриці, вектори), операції з матрицями.
imgproc	Фільтрація, перетворення кольорів, морфологічні операції, геометричні трансформації.
video	Обробка відео: оптичний потік, трекінг об'єктів, фонові субтракція.
calib3d	Калібрування камер, 3D-реконструкція, робота з сенсорами (наприклад, LiDAR).
features2d	Виявлення особливих точок (SIFT, SURF, ORB), матчинг.
objdetect	Детекція об'єктів (обличчя, людини, авто) за допомогою Нагг-каскадів або DNN.
ml	Алгоритми машинного навчання (SVM, KNN, нейромережі).
dnn	Робота з глибокими нейромережами (TensorFlow, PyTorch, ONNX).
face	Розпізнавання обличь, виявлення ключових точок (landmarks).
contrib	Експериментальні модулі (наприклад, AR, текстові детектори).

OpenCV — це потужний інструмент для комп'ютерного зору, який пропонує гнучкі методи для роботи з обличчями, включаючи аналогі Face Mesh. Face Mesh — це метод виявлення ключових точок обличчя (наприклад, контур очей, губ, носа). Хоча OpenCV не має власного модуля з назвою "Face Mesh", вона пропонує аналогічні можливості через модуль face, який використовується для виявлення 68 ключових точок обличчя на основі алгоритмів LBF (Local Binary Features), Kazemi. Застосовується Face Mesh у OpenCV завдяки накладанню віртуальних AR-масок на обличчя.

Алгоритм оцінки глибини з Face Mesh та OpenCV наведено нижче.

Отримання зображення:

- захоплення кадру з камери за допомогою cv2.VideoCapture;
- перетворення кадру в потрібний формат (RGB для MediaPipe).

Обробка зображення та визначення ключових точок обличчя:

- використання MediaPipe Face Mesh для детекції landmark-точок;
- виявлення контрольних точок (наприклад, очі, ніс, губи).

Розрахунок масштабу та відстані:

- модельний підхід: використання відстані між очима або іншими частинами обличчя як еталонної довжини;

- перспективне перетворення: якщо камера має відому фокусну відстань, базову відстань та можна оцінити диспаратність, то розрахунок відстані здійснюється за стерео-формулою (2.1).

Оцінка глибини за масштабними змінами:

- використання зміни розміру контрольних точок обличчя при віддаленні чи наближенні до камери.

Face Mesh дозволяє отримати 468 3D-точок обличчя, що дає змогу оцінювати глибину за масштабом. У поєднанні з методами перспективної геометрії та нейромережевими підходами можна отримати досить точні оцінки відстані. Проте, для точного вимірювання необхідно калібрувати камеру та використовувати реальні метрики.

3 МЕТОДИКА ПРОВЕДЕННЯ ЕКСПЕРИМЕНТІВ

3.1 Обраний технологічний стек

Технологічний стек для експериментальної частини проекту базується на мові програмування Python, яка домінує в галузях комп'ютерного зору та машинного навчання завдяки своїй унікальній комбінації переваг. Вибір Python обумовлений його інтуїтивно зрозумілим синтаксисом, що значно прискорює процес прототипування та ітеративних експериментів, особливо критичних на дослідницькій фазі. Мова підтримується потужною екосистемою спеціалізованих бібліотек, таких як NumPy для наукових обчислень, Matplotlib для візуалізації даних, та фреймворків глибокого навчання, включаючи TensorFlow і PyTorch. Не менш важливою є активна глобальна спільнота, яка забезпечує доступ до безлічі відкритих репозиторіїв на GitHub, детальних туторіалів і готових рішень складних задач, що істотно знижує бар'єр входження в проект.

Середовище розробки PyCharm обране як інтегрована платформа для реалізації рішення, пропонуючи розширені інструменти для професійної розробки. Його інтелектуальне автодоповнення коду на основі машинного навчання аналізує контекст і пропонує релевантні варіанти, пришвидшуючи написання коду. Вбудовані інструменти налагодження дозволяють детально аналізувати виконання програми, виявляти помилки та оптимізувати продуктивність критичних ділянок коду. PyCharm також забезпечує безшовну інтеграцію з Jupyter Notebook для інтерактивних досліджень і підтримує роботу з Docker, що гарантує відтворюваність результатів у контрольованих середовищах.

Бібліотека OpenCV слугує фундаментом для роботи з візуальними даними, пропонуючи всебічний набір інструментів для обробки зображень і відео. Її архітектура включає понад 2500 оптимізованих алгоритмів, що

охоплюють широкий спектр операцій: від базової фільтрації та трансформацій до складних методів детекції об'єктів і трекінгу. Бібліотека відзначається високою продуктивністю завдяки реалізації ключових функцій на C++ з Python-інтерфейсом, підтримці апаратного прискорення через CUDA/OpenCL, а також крос-платформністю. Величезна база документації з прикладами застосування спрощує освоєння інструментарію та швидке вирішення специфічних задач.

Бібліотека MediaPipe FaceMesh обрана як спеціалізоване рішення для аналізу облич, забезпечуючи точне визначення 3D-ландмарків у реальному часі. Її архітектура дозволяє з високою точністю виявляти 468 ключових точок на обличчі, що є критичним для аналізу міміки та рухів голови. FaceMesh відрізняється оптимальним балансом між точністю та продуктивністю, демонструючи стабільну роботу навіть на системах без потужних графічних процесорів. Бібліотека легко інтегрується з іншими інструментами, особливо з OpenCV, через простий API, що дозволяє створювати комплексні конвеєри обробки відеоданих. Ця комбінація технологій створює ідеальне середовище для дослідження методів комп'ютерного зору та розробки інноваційних рішень у сфері аналізу візуальної інформації.

3.2 Теоретичне пояснення запропонованого рішення

Визначення глибини зображення, зокрема для таких об'єктів, як обличчя людини, фундаментально базується на декількох ключових фізичних принципах. До них належать аналіз геометричних характеристик шляхів світлових променів, врахування оптичних ефектів (наприклад, ступеня розмиття контурів або артефактів, пов'язаних із рухом), а також особливостей відображення світла від поверхонь із певними властивостями матеріалів, зокрема текстур. Враховуючи специфічні вимоги дослідження – наявність апріорно відомих параметрів камери, необхідність отримання

інтерпретовних результатів з високим рівнем контролю над процесом, а також жорсткі обмеження на час обчислень для забезпечення оперативності обробки – вибір методів, заснованих на геометричних принципах, виявляється найбільш оптимальним для поставленого завдання.

Методологія, що реалізована в таких підходах як FaceMesh або аналогічних техніках, що використовують попередньо визначені масштабні відношення, ґрунтується на прямих геометричних розрахунках. Суть полягає у визначенні відстані до об'єкта через відомі фізичні розміри його характерних ознак (наприклад, середньостатистичну відстань між центрами очей людини) та вимірювання проекції цих ознак у пікселях на площині зображення. Цей підхід демонструє ряд істотних переваг у порівнянні із сучасними методами глибинного аналізу, заснованими на глибокому навчанні, особливо в контексті зазначених вимог до швидкодії, інтерпретовності та ресурсних витрат.

Першою ключовою перевагою є мінімальна залежність від обсягів навчальних даних. Геометричні методи, такі як FaceMesh, не потребують складного та ресурсомісткого процесу тренування на великих розмічених датасетах. Для їх функціонування достатньо знання фізичних розмірів цільового об'єкта (наприклад, стандартної відстані між очима, що приймається за 6.5 см) та технічних параметрів використовуваної камери, таких як фокусна відстань. На противагу цьому, моделі глибинного навчання, представлені, наприклад, архітектурами типу Monodepth, вимагають значних обсягів даних – тисяч зображень з точно визначеними картами глибини. Отримання таких датасетів є трудомістким процесом і часто є непрактичним або недоступним для вузькоспеціалізованих задач.

Другою важливою перевагою є низьке енергоспоживання та висока ефективність обчислень. Алгоритми реалізації геометричних методів, зокрема засновані на бібліотеках комп'ютерного зору типу OpenCV (наприклад, для виявлення ключових точок обличчя), здатні працювати в режимі реального часу навіть на обмежених у потужності пристроях, таких

як смартфони або мікрокомп'ютери Raspberry Pi. На противагу, моделі глибокого навчання для оцінки глибини (наприклад, DepthNet) характеризуються складними архітектурами, що значно підвищують вимоги до обчислювальних ресурсів. Їх інференс зазвичай вимагає використання графічних процесорів (GPU), що суттєво збільшує енергоспоживання, вартість апаратного забезпечення та ускладнює розгортання в вбудованих системах або пристроях з обмеженим живленням.

Третя перевага стосується інтерпретовності результатів та контролю над процесом. Математичний апарат, що лежить в основі геометричних методів (проєктивна геометрія), є прозорим і добре вивченим. Формули для розрахунку відстані, що використовуються в FaceMesh, дозволяють чітко простежити ланцюжок обчислень і провести точний аналіз джерел потенційних помилок. Це надає досліднику високий рівень контролю та можливість адаптації методу під специфічні умови. Навпаки, моделі глибокого навчання часто функціонують за принципом "чорної скриньки", де внутрішні механізми прийняття рішень важко інтерпретувати. Це ускладнює визначення причин неточної оцінки відстані, особливо у випадках нестандартних зовнішностей облич або незвичайних умов зйомки, і робить налагодження та корекцію проблематичними.

Четвертий аспект пов'язаний із етикою та захистом даних. Геометричні методи, такі як FaceMesh, зазвичай реалізуються таким чином, що обробка відбувається локально на пристрої без необхідності передачі зображень або біометричних даних на зовнішні сервери. Ключові точки обличчя використовуються виключно для геометричних розрахунків і не зберігаються постійно, що забезпечує анонімність користувача. Це суттєво знижує ризики, пов'язані зі зберіганням та обробкою біометричної інформації, і спрощує дотримання жорстких вимог нормативних актів про захист даних, таких як GDPR. Моделі ж глибокого навчання, особливо ті, що тренуються на даних користувачів, можуть вимагати збору та передачі зображень облич, що потенційно створює ризик несанкціонованого збору або витоку біометричних

даних, що є персональною інформацією, і може призводити до порушень законодавства.

Узагальнена методологія запропонованого підходу включає наступну послідовність кроків:

- захоплення відеопотоку;
- детекція обличчя;
- визначення ключових точок та розрахунок відстані;
- виведення результату;
- відображення зображення.

Геометричний підхід, який досліджується в роботі, реалізований за допомогою методів типу FaceMesh, пропонує ефективне, ресурсозберігаюче, інтерпретовне та етично більш безпечне рішення для задачі швидкої оцінки глибини зображення обличчя людини з використанням відомих параметрів камери, що повністю відповідає поставленим цілям дослідження.

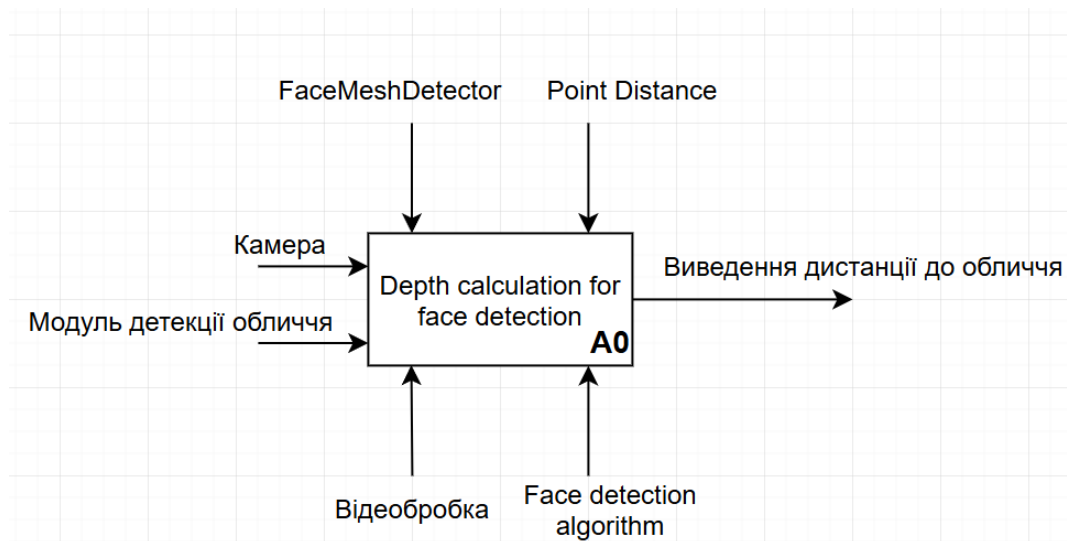


Рисунок 3.1 – Приклад процесу, описаного за допомогою `usecase` нотатції

На рисунку 3.1 представлено функціональну модель підсистеми розрахунку відстані до обличчя на основі геометричних характеристик, яка є ядром реалізації візуальної підсистеми взаємодії з користувачем у задачах інтерфейсів людина-комп'ютер. Блок-схема умовно структурує взаємозв'язки між ключовими модулями, програмними компонентами та інформаційними

потоками, що забезпечують розрахунок глибини на основі вхідного відеопотоку.

Центральним елементом схеми є функціональний блок “Depth calculation for face detection”, позначений маркером A0, який виконує основні обчислення, пов’язані з оцінкою просторової глибини (тобто відстані до обличчя користувача відносно оптичного центру камери). Цей блок інтегрує в собі функціональність попередньої обробки відеоданих, детекції обличчя, аналізу просторових параметрів та подальшого обчислення дистанції на основі аналітичних (геометричних) методів.

До блоку A0 надходять шість ключових вхідних сигналів, кожен із яких виконує специфічну роль:

Камера – надає безперервний потік зображень (кадрів), отриманих у реальному часі. Це джерело первинних даних, що використовуються для подальшої обробки. Важливо, що параметри камери (фокусна відстань, роздільна здатність, площа фокусу) є або відомими апріорі, або можуть бути визначені в процесі калібрування.

Модуль детекції обличчя – відповідає за ідентифікацію області обличчя на зображенні. В рамках цієї системи може використовуватись класичний каскад Хаара, сучасні CNN-моделі або бібліотечні рішення на базі OpenCV або MediaPipe.

Відеообробка – виконує попередню обробку зображення, зокрема масштабування, згладжування шумів, нормалізацію освітлення або інші трансформації, необхідні для підвищення якості детекції. Даний компонент підвищує стабільність системи в реальних умовах, зокрема при зміні освітлення або положення об’єкта.

FaceMeshDetector – компонент, що відповідає за побудову сітчастої моделі обличчя з нанесенням маркерних точок. У контексті даної системи він виконує роль генератора координат ключових точок, серед яких особливо важливою є пара “центри очей”, що використовується як базис для геометричних розрахунків.

Point Distance – алгоритмічний блок, який обчислює евклідову або піксельну відстань між вибраними парами точок на зображенні (наприклад, між зіницями очей). Ці значення надалі використовуються для масштабного перетворення з урахуванням реального фізичного розміру міжточкової відстані (заданого наперед, напр., 6.5 см для міжочкової відстані).

Face Detection Algorithm – додатковий (або вбудований) алгоритм для локалізації обличчя, який може бути автономним або працювати як підсистема попереднього уточнення координат у випадках, коли базовий FaceMeshDetector не забезпечує потрібної точності.

Єдиним виходом функціонального блоку є «Виведення дистанції до обличчя», що реалізується у вигляді числового значення (в міліметрах або сантиметрах) і передається на інтерфейс виведення, систему логування або інші модулі (наприклад, керування камерою, фокусування, адаптація UI). Цей параметр може використовуватись у сценаріях контролю доступу, керування жестами, автоматичного масштабування зображення, а також у безконтактних системах медичної діагностики.

Схема є прикладом детермінованого модульного підходу, де кожен вхід має чітко визначене призначення і може бути замінений або адаптований під потреби конкретного програмно-апаратного середовища. Завдяки цьому можлива адаптація системи як до локального (edge) обчислення на вбудованих платформах, так і до більш складних розподілених середовищ, що працюють із потоками з декількох камер.

Архітектура також забезпечує високу інтерпретованість результатів, оскільки на кожному етапі – від захоплення відео до обчислення глибини – існує можливість верифікації проміжних значень, що є неможливим у випадку застосування end-to-end нейронних моделей. Такий підхід також дозволяє впроваджувати додаткові модулі (наприклад, компенсацію руху, фільтрацію на основі гістерезису, або автоматичне калібрування).

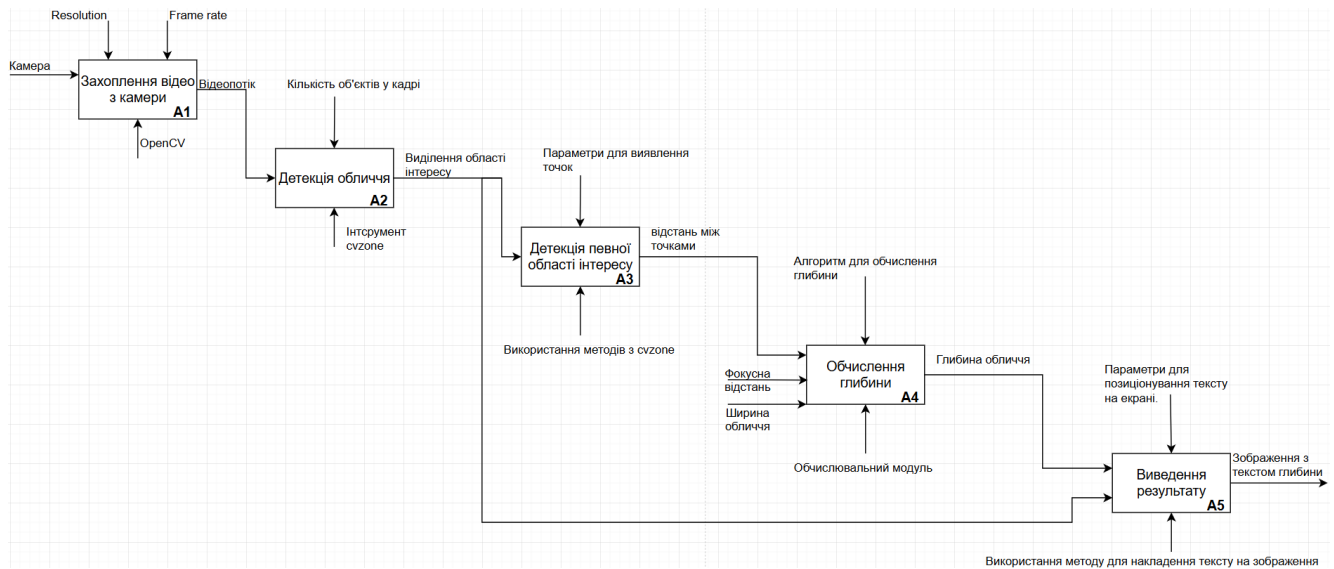


Рисунок 3.2 – Детальний процес, за допомогою ideo нотації

На рисунку 3.2 зображено деталізований процес розрахунку глибини до обличчя на зображенні з відеопотоку, представлений відповідно до методології IDEF0. Обрана нотація дає змогу чітко структурувати логіку інформаційних потоків, функціональних залежностей та взаємодії вхідних і вихідних компонентів системи. Даний підхід дозволяє не лише моделювати поточну архітектуру, а й окреслити напрями її майбутнього масштабування та вдосконалення.

A1 – Захоплення відео з камери. На першому етапі (функціональний блок A1) виконується захоплення відеопотоку з камери. В якості ключових вхідних параметрів тут виступають роздільна здатність (Resolution) та частота кадрів (Frame Rate), які безпосередньо впливають на подальшу точність і стабільність аналізу. Для цього використовуються бібліотеки типу OpenCV, що забезпечують доступ до відеопотоку в реальному часі. Вихідними даними цього етапу є кадри з послідовними зображеннями сцени, що подаються на вхід наступного функціонального блоку.

A2 – Детекція обличчя. Блок A2 відповідає за детекцію обличчя на кожному окремому кадрі. У межах цього етапу застосовуються алгоритми виявлення облич (наприклад, на основі каскадів Хаара, HOG+SVM, або нейромережеві підходи), які забезпечують локалізацію області інтересу

(ROI). До допоміжних засобів цього блоку належать програмні інструменти обробки об'єктів, зокрема фільтри попереднього згладжування та аналіз просторових контурів. Результатом є окрема область з обличчям, виокремлена з повного зображення, яка передається до наступного блоку.

A3 – Детекція певної області інтересу. На цьому етапі (блок A3) відбувається детекція ключових маркерів обличчя, зокрема області між зіницями, брів, носа або інших референтних точок. Це дозволяє з високою точністю визначити параметри для обчислення глибини. У даному процесі використовуються методи на основі FaceMesh або аналогічних маркерних сіток, які забезпечують розташування ключових точок з точністю до пікселя. Надалі проводиться обчислення відстані між точками у пікселях, що передається до блоку A4.

A4 – Обчислення глибини. Функціональний блок A4 є ядром розрахункової підсистеми. Тут реалізовано геометричну модель проєктивного перетворення, яка базується на співвідношенні між фізичними параметрами (такі як ширина об'єкта, тобто реальна міжочна відстань) та їхньою проєкцією на площину зображення. Алгоритм виконує розрахунок глибини на основі трикутної пропорції: $\text{глибина} = (\text{фокусна відстань} * \text{фізична відстань}) / \text{піксельна відстань}$. У разі вдосконалення цього модуля можливе впровадження динамічної компенсації фокусної відстані або адаптивної калібровки під різні оптичні системи. Додатково, цей блок може включати оцінку похибки та побудову довірчих меж результатів.

A5 – Виведення результату. Фінальний блок A5 відповідає за відображення та виведення обчисленої глибини. Результати можуть бути виведені як числові значення або накладені на зображення у вигляді текстових міток. Цей блок підтримує зворотний зв'язок із користувачем у вигляді візуалізації, що значно підвищує інтерпретованість і зручність взаємодії з системою. В перспективі можливе додавання модуля логування, збереження історії або передача даних на зовнішній сервер чи інтерфейс користувача (наприклад, вебпанель або мобільний додаток).

Обчислення глибини (проводиться на основі відомих параметрів об'єкта та камери на основі формули:

$$D = \frac{W * f}{w}, \quad (3.1)$$

де D – відстань до об'єкта, W – фізична ширина об'єкта, f – фокусна відстань камери, w – виміряна ширина об'єкта у пікселях.

Запропонований алгоритм роботи системи визначення глибини до обличчя користувача з використанням бібліотеки FaceMeshDetector наведено на рисунку 3.3. Алгоритм реалізує замкнутий цикл обробки відеопотоку в реальному часі, що дозволяє на кожному кадрі ідентифікувати ключові орієнтири обличчя (зокрема, очі), обчислювати відстань до об'єкта за допомогою геометричних розрахунків та виводити результати на екран. Цей підхід демонструє ефективність у прикладних задачах, де важливо забезпечити швидку реакцію, обмежене енергоспоживання та прозорість обчислень.

Процес розпочинається з виклику блоку Start (ініціалізація), що сигналізує про початок виконання програми. Далі ініціалізується потік з вебкамери шляхом створення об'єкта cv2.VideoCapture з бібліотеки OpenCV. Цей компонент відповідає за постійне оновлення відеозображення з заданою частотою кадрів та роздільною здатністю.

Після активації потоку кожен кадр обробляється за допомогою модуля FaceMeshDetector, який здійснює виявлення обличчя та маркерних точок. На цьому етапі система визначає, чи виявлено обличчя в кадрі. У разі негативного результату виконується цикл повернення (Loop Back) — алгоритм повертається до захоплення наступного кадру. Це дозволяє уникнути зайвих обчислень і забезпечити стабільність роботи навіть у разі зникнення об'єкта з поля зору камери.

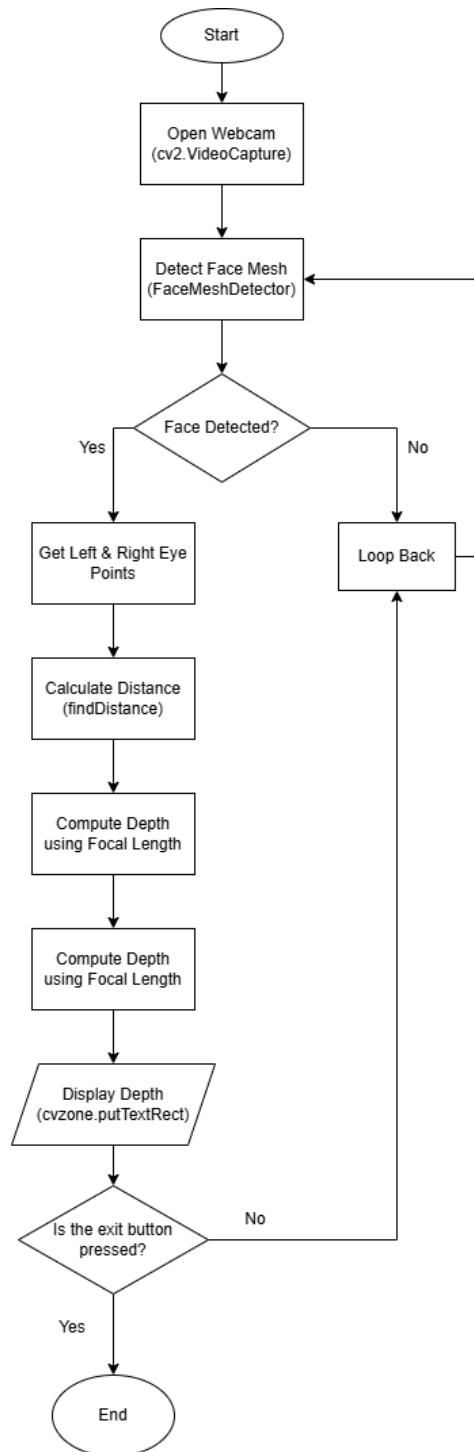


Рисунок 3.3 – Запропонований алгоритм роботи системи визначення глибини до обличчя користувача з використанням бібліотеки FaceMeshDetector

У випадку успішної детекції обличчя відбувається виділення ключових точок — зокрема, лівого та правого ока. Ці маркери використовуються для подальших геометричних розрахунків. Далі викликається функція

findDistance, яка визначає піксельну відстань між очима, що є необхідним параметром для наступного етапу — розрахунку глибини.

Після визначення міжочкової відстані система виконує обчислення відстані до обличчя, використовуючи фокусну відстань камери. Алгоритм базується на проєктивній формулі (3.1)

У схемі цей крок двічі згадується як “Compute Depth using Focal Length”, що може вказувати на можливість порівняння або усереднення результатів від кількох методів оцінки (наприклад, через обидва ока, або через середню геометричну точку).

Отримане значення глибини передається до функції cvzone.putTextRect, яка відображає числовий результат на поточному кадрі. Це забезпечує зручний графічний інтерфейс, що дозволяє користувачу в реальному часі бачити результат обчислень.

Завершальним етапом є перевірка натискання кнопки виходу.

3.3 Порядок проведення експериментів

На етапі підготовки до кожної серії вимірювань виконується калібрувальна процедура, що включає налаштування фокусування оптичної системи та визначення положення досліджуваної області інтересу (ROI). Для забезпечення статистичної значущості результатів на кожній контрольній відстані виконується не менше трьох вимірювань, що дозволяє розрахувати середнє арифметичне значення та величину стандартного відхилення.

Експериментальна платформа базується на програмному середовищі Python 3.10 з використанням бібліотек комп'ютерного зору OpenCV версії 4.9 та фреймворку MediaPipe FaceMesh для аналізу антропометричних параметрів обличчя. В якості апаратного засорбу захоплення зображень використовується веб-камера Logitech C920, що працює у режимі роздільної здатності 640×480 або 1280×720 пікселів при частоті кадрів 30 fps.

Процедура експерименту реалізується за наступним алгоритмом:

- фіксація еталонної відстані за допомогою прецизійного вимірювального інструменту (механічної лінійки або лазерного далекоміра) визначається контрольна відстань до об'єкта;

- стабілізація камери полягає у фіксації на штативі або нерухомій поверхні для виключення впливу механічних коливань;

- виконання вимірювань полягає в запуску алгоритму розрахунку дистанції на базі бібліотек FaceMesh та OpenCV, що визначає відстань через проекцію антропометричних параметрів обличчя;

- верифікація результатів - отримані алгоритмічні значення порівнюються з еталонною відстанню з подальшим обчисленням абсолютної та відносної похибки вимірювань.

Ця методологія забезпечує контрольовані умови експерименту та дозволяє отримати об'єктивні дані для подальшого аналізу точності запропонованого методу.

4 АНАЛІЗ РЕЗУЛЬТАТІВ ЕКСПЕРИМЕНТІВ

Практичне значення очікуваних результатів полягає в тому, що інтеграція недорогих камер дозволить створити бюджетну систему для вимірювання відстані, яка може бути використана в навчальних лабораторіях чи освітніх семінарах; порівняння класичних і сучасних методів дозволить надати рекомендації для вибору оптимального підходу в залежності від обчислювальних ресурсів і умов задачі.

4.1 Точність вимірювання у залежності від відстані до об'єкта при різній роздільній здатності веб-камери

Мета експерименту - оцінка точності вимірювання відстані (вираженої у відсотках похибки, $E\%$) залежно від двох основних факторів:

- відстань до об'єкта (D): 30 см, 60 см, 90 см, 120 см;
- роздільна здатність веб-камери: 640×480 пікселів, 1280×720 пікселів, та 1920×1080 пікселів (за наявності підтримки).

Наукова гіпотеза - передбачається, що точність визначення відстані методом, заснованим на аналізі піксельної проекції антропометричних точок обличчя, матиме наступну залежність:

- вплив відстані (D). Зі збільшенням дистанції до об'єкта фізичний масштаб ключових ознак (наприклад, відстань між очима) проектується на меншу кількість пікселів зображення. Це призводить до зниження точності визначення відстані ($E\%$ зростає);

- вплив роздільної здатності. Зменшення роздільної здатності (наприклад, 640×480 проти 1280×720 чи 1920×1080) ускладнює розпізнавання дрібних деталей обличчя. Відповідно, збільшення роздільної здатності забезпечує вищу піксельну щільність для ключових

антропометричних характеристик (очі, ніс), що покращує точність оцінки відстані ($E\%$ зменшується). Таким чином, найвища точність очікується на мінімальній відстані ($D = 30$ см) та при максимальній роздільній здатності (1920×1080 , за умови підтримки).

Таблиця 4.1 - Оцінка впливу роздільної здатності на точність визначення відстані об'єкта від камери

Роздільна здатність веб-камери	Визначено точність до об'єкта, см	Точність вимірювання, %
Реальна відстань до об'єкта $D = 30$ см		
640x480	31.85	98.65%
1280x720	30.41	98.26%
1920x1080	30.53	94.19%
Реальна відстань до об'єкта $D = 60$ см		
640x480	60	100.00%
1280x720	60.86	98.59%
1920x1080	60.67	98.91%
Реальна відстань до об'єкта $D = 90$ см		
640x480	90.65	99.28%
1280x720	91.87	97.97%
1920x1080	91.05	98.85%
Реальна відстань до об'єкта $D = 120$ см		
640x480	118.6	98,82%
1280x720	124.5	96.39%
1920x1080	119.2	99,33%

Метою цього експерименту було визначити, наскільки роздільна здатність веб-камери впливає на точність обчислення відстані до обличчя користувача в системі, що базується на методі FaceMesh. Для цього було проведено серію вимірювань при чотирьох фіксованих значеннях реальної

відстані до об'єкта: 30 см, 60 см, 90 см та 120 см. У кожному з випадків виконувались незалежні вимірювання при трьох стандартних роздільних здатностях камери: 640×480, 1280×720 та 1920×1080 пікселів.

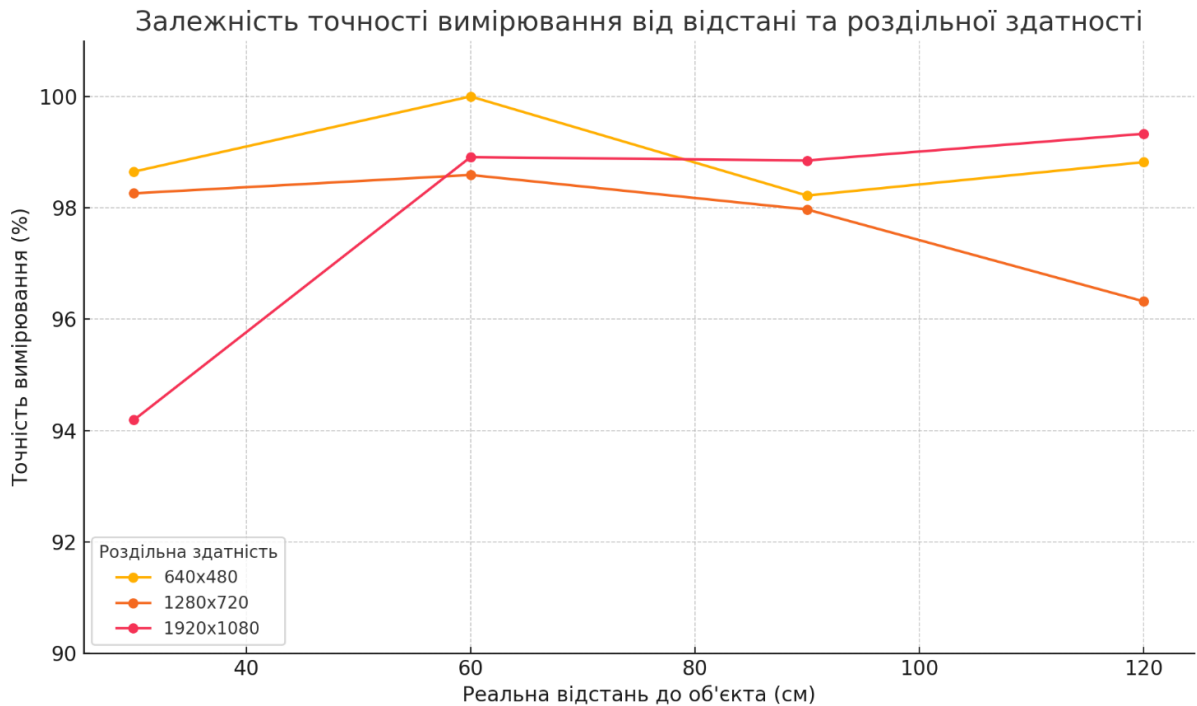


Рисунок 4.1 - Оцінка впливу роздільної здатності на точність визначення відстані об'єкта від камери

Як свідчать дані, представлені в таблиці 4.1, всі варіанти забезпечують високий рівень точності — у більшості випадків понад 97%. Однак графічна візуалізація (рисунок 4.1) чітко демонструє відмінності в поведінці кожної конфігурації роздільної здатності в залежності від відстані до об'єкта. Це дозволяє сформулювати низку важливих висновків.

Низька роздільна здатність (640×480) показала найбільш стабільну та передбачувану поведінку на всіх відстанях. Точність вимірювання варіюється в межах 98.65–100%, при цьому найкращий результат досягнуто на відстані 60 см, що співпадає з «оптимальним фокусом» для багатьох бюджетних веб-камер. Це свідчить про те, що навіть при мінімальній роздільній здатності система може забезпечити високу точність за умови ретельного калібрування.

Отримані результати вказують на важливість динамічного налаштування роздільної здатності відповідно до цільової відстані та умов зйомки. Це може лягти в основу адаптивних систем, що автоматично перемикають режим роботи камери або алгоритмів передобробки зображення для досягнення максимальної точності в різних умовах. Також результати демонструють потенціал для побудови моделей оцінки похибки, залежно від вхідної конфігурації відеопотоку, що є актуальним для подальшої інтеграції в промислові застосування (розумні дзеркала, медичні інтерфейси, біометричні системи контролю доступу).

4.2 Точність визначення відстані у залежності від відстані до об'єкта при змінному освітленні

Мета експерименту - дослідження впливу рівня освітлення на точність визначення відстані (похибка $E\%$) за таких умов:

- відстань до об'єкта (D) - 30 см, 60 см, 90 см, 120 см;
- рівень освітленості - < 50 люкс (низьке освітлення), 200 люкс (нормальне освітлення), 500 люкс (надмірне освітлення).

Наукова гіпотеза - передбачається, що точність методу на основі FaceMesh суттєво залежить від умов освітлення через його вплив на якість детекції ключових антропометричних точок обличчя:

- вплив низького освітлення (< 50 lux): Дефіцит світла призводить до зниження контрастності зображення, збільшення шуму та втрати деталізації. Це ускладнює надійне виявлення та точне позиціонування ландмарків обличчя (очі, ніс, контури), що викликає зростання похибки $E\%$;

- вплив надмірного освітлення (> 500 lux): Інтенсивне світло сприяє виникненню локалізованих пересвітлень (відблисків) на поверхні шкіри (особливо лоб, ніс, щоки) та окулярах (за їх наявності). Ці артефакти "засліплюють" ділянки зображення, спотворюють контури обличчя та

перешкоджають коректній роботі алгоритму визначення точок, що також призводить до збільшення $E\%$;

- оптимальні умови (200 lux): Рівень нормального освітлення забезпечує достатню яскравість, контрастність та відсутність критичних пересвітлень. Це створює ідеальні умови для стабільної та точної роботи алгоритму FaceMesh, мінімізуючи $E\%$ порівняно з екстремальними освітлювальними сценаріями.

Таблиця 4.2 - Оцінка впливу освітленості на точність визначення відстані об'єкта від камери

Ступінь освітленості, лк	Визначено точність до об'єкта , см	Точність вимірювання, %
Реальна відстань до об'єкта D = 30 см		
< 50	31.08	96,4
200	31.47	95,1
> 500	30.67	97,77
Реальна відстань до об'єкта D = 60 см		
< 50	61.56	94,87\
200	62.05	96,58
> 500	60.08	99,87
Реальна відстань до об'єкта D = 90 см		
< 50	93.82	95,76
200	92.15	97,61
> 500	88.64	98,49
Реальна відстань до об'єкта D = 120 см		
< 50	115.04	95,87
200	119.17	99,31
> 500	121.05	99,12

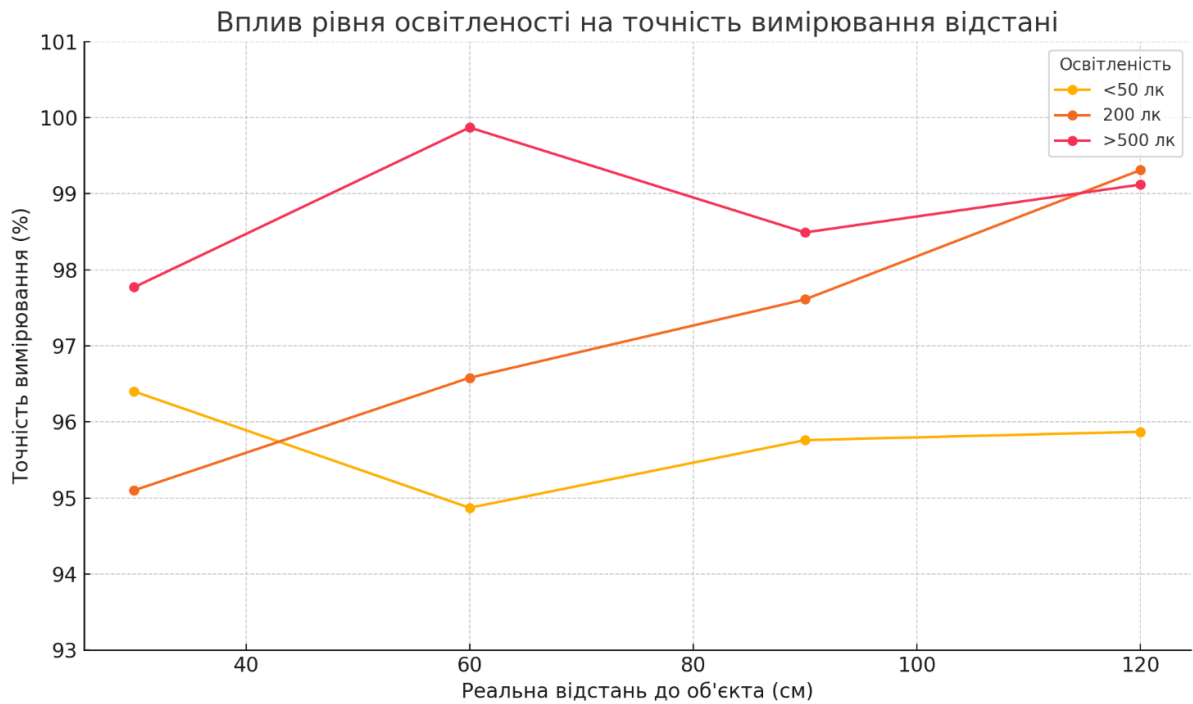


Рисунок 4.2 – Оцінка впливу якості освітлення на точність визначення відстані об'єкта від камери

Для оцінки впливу умов освітлення на точність системи визначення глибини на основі алгоритму FaceMesh було проведено серію експериментів із фіксованими відстанями до об'єкта (30 см, 60 см, 90 см та 120 см) при трьох рівнях освітленості: менш як 50 лк (слабке освітлення), 200 лк (помірне офісне освітлення) та понад 500 лк (інтенсивне світло, наприклад, денне або студійне). Результати представлено графічно на рисунку та таблично в таблиці 4.2.

Аналіз результатів демонструє чітку позитивну кореляцію між рівнем освітленості та точністю вимірювання. У всіх випадках найвищі значення точності спостерігаються при освітленості >500 лк, що пояснюється кращою якістю зображення, контрастністю та чіткістю меж, які критично важливі для точної локалізації ключових точок FaceMesh-алгоритмом.

Інтенсивне освітлення (>500 лк) виявилось найефективнішим - точність вимірювання стабільно перебуває в межах 97.77% – 99.87%. При цьому, вже на близькій відстані (30 см) спостерігається максимальна точність серед усіх варіантів. Це вказує на оптимальні умови для роботи алгоритму: чітке

виділення контурів, низький рівень шумів, відсутність тіней і правильне розпізнавання пропорцій обличчя. На практиці це означає, що для стабільної роботи системи необхідно забезпечити мінімальний рівень освітлення не менше 200 лк, а для критичних застосувань — орієнтуватися на >500 лк.

4.3 Точність визначення відстані у залежності від кута нахилу голови

Мета експерименту - дослідження впливу орієнтації голови на точність визначення відстані (похибка $E\%$) при наступних умовах:

- кут нахилу у фронтальній площині (повороти вліво/вправо) - 0° , $\pm 15^\circ$, $\pm 30^\circ$;
- кут нахилу у сагітальній площині (нахили вперед/назад) - 0° , $\pm 15^\circ$, $\pm 30^\circ$.

Наукова гіпотеза - передбачається, що відхилення від фронтального положення голови (0°) призводить до систематичного зростання похибки $E\%$ через такі фактори:

- порушення симетрії та проекції: - нахил голови спотворює відстані між ключовими антропометричними точками (очі, вуха, ніс) у просторі пікселів. Алгоритми типу FaceMesh, що базуються на передбачених симетричних моделях, втрачають точність при розрахунку масштабу через асиметричне проектування ознак;
- часткове приховання ознак - при значних кутах ($\pm 30^\circ$) виникає перекриття окремих ландмарків (наприклад, втрата контуру обличчя, часткове закриття ока або вуха), що зменшує кількість доступних для аналізу точок та знижує точність триангуляції;
- збільшення впливу проєктивних спотворень - нахил змінює перспективну геометрію відносно площини сенсора камери. Це призводить до нелінійних спотворень відстаней між точками, які не враховуються у базових геометричних моделях, розрахованих на фронтальне положення.

Таблиця 4.3 - Оцінка впливу кута нахилу на точність визначення відстані об'єкта від камери привідхиленні по фронтальній площині

Кут нахилу голови у фронтальній площині, °	Визначено точність до об'єкта, см	Точність вимірювання, %
Реальна відстань до об'єкта D = 30 см		
0	30.1	99,67
15	28.17	93,9
30	32.84	90,53
Реальна відстань до об'єкта D = 60 см		
0	60.2	99,67
15	58.62	97,7
30	62.87	95,22
Реальна відстань до об'єкта D = 90 см		
0	90.15	99,83
15	87.38	97,09
30	94.33	95,19
Реальна відстань до об'єкта D = 120 см		
0	120.1	99,92
15	115.72	96,43
30	126.48	94,6

У ході дослідження було проаналізовано, як зміна положення голови у фронтальній площині (0° , 15° , 30°) впливає на точність визначення відстані до користувача за допомогою геометричного підходу (Таблиця 4.3). У всіх випадках використовувалася одна й та сама камера та алгоритм FaceMesh з прив'язкою до міжочкової відстані.

У всіх вимірюваннях чітко простежується тенденція до зниження точності при збільшенні кута нахилу (Рисунок 4.3). При кутах 0° точність системи наближається до 99.7–99.9%, що є практично ідеальним значенням.

Однак вже при нахилі 15° точність знижується на 2–3%, а при 30° - у деяких випадках спостерігається падіння до $\approx 90.5\%$.

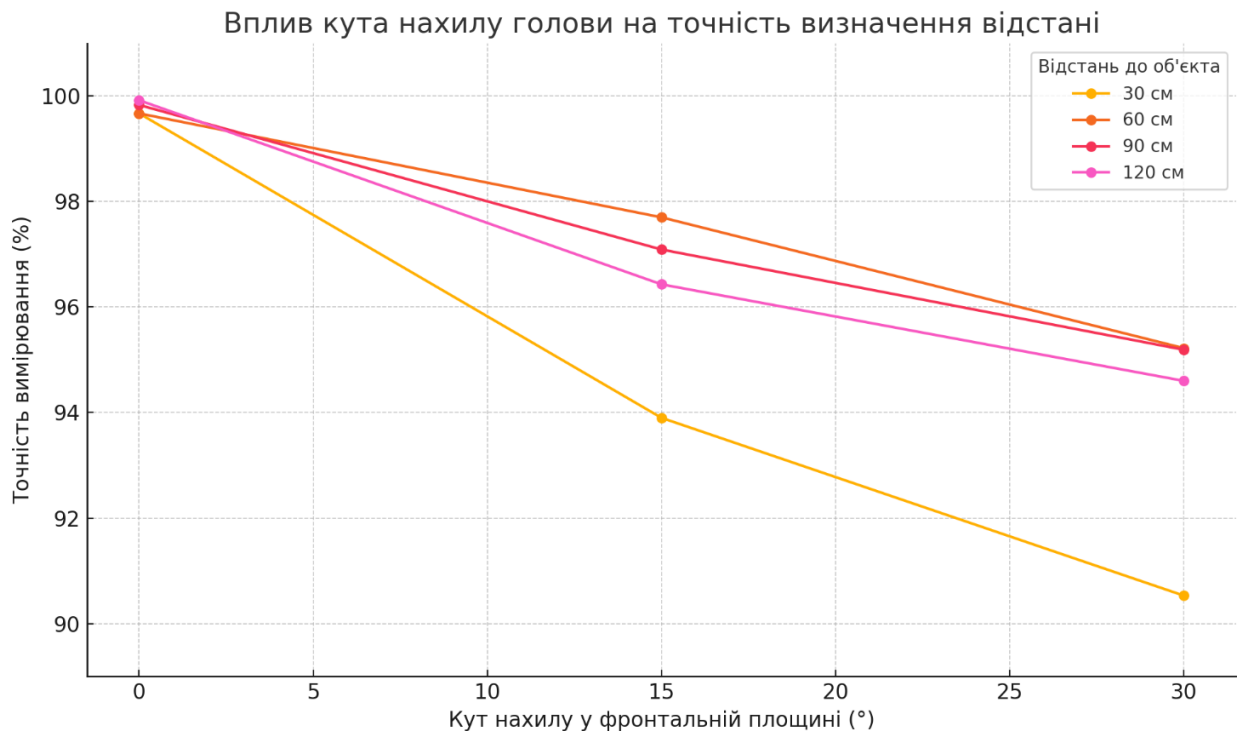


Рисунок 4.3 - Оцінка впливу нахилу голови на точність визначення відстані об'єкта від камери у фронтальній площині

На відстані 30 см: точність різко знижується з 99.67% (0°) до 93.9% (15°) і далі до 90.53% (30°). Це вказує на значну чутливість алгоритму до кута при роботі на близьких відстанях. Можливими причинами є деформація перспективи та перекриття ключових точок (наприклад, часткове перекриття ока або брови при повороті голови).

У межах експерименту було визначено, як зміна кута повороту голови в сагітальній площині (0° , 15° , 30°) впливає на точність обчислення глибини (відстані до обличчя) за допомогою FaceMesh. Дослідження проводилось на чотирьох фіксованих відстанях: 30, 60, 90 та 120 см (Таблиця 4.4).

Таблиця 4.4 - Оцінка впливу кута нахилу на точність визначення відстані об'єкта від камери

Кут нахилу голови у сагітальній площині, °	Визначено точність до об'єкта, см	Точність вимірювання, %
Реальна відстань до об'єкта $D = 30$ см		
0	30.15	99,5
15	34.03	86,57
30	37.16	76,13
Реальна відстань до об'єкта $D = 60$ см		
0	60.1	99,83
15	63.94	93,43
30	66.56	89,07
Реальна відстань до об'єкта $D = 90$ см		
0	91.0	98,89
15	86.75	96,39
30	97.10	92,11
Реальна відстань до об'єкта $D = 120$ см		
0	120.06	99,95
15	125.2	95,67
30	113.15	94,29

У всіх випадках найвища точність спостерігається при нульовому куті — тобто при фронтальному положенні голови. Значення варіюються від 98.89% до 99.95%.

Зі збільшенням кута повороту голови точність суттєво знижується. Це пов'язано з частковим викривленням перспективи та втратою симетричних маркерів (наприклад, при повороті одне око може бути не видно, або положення носа змінюється).

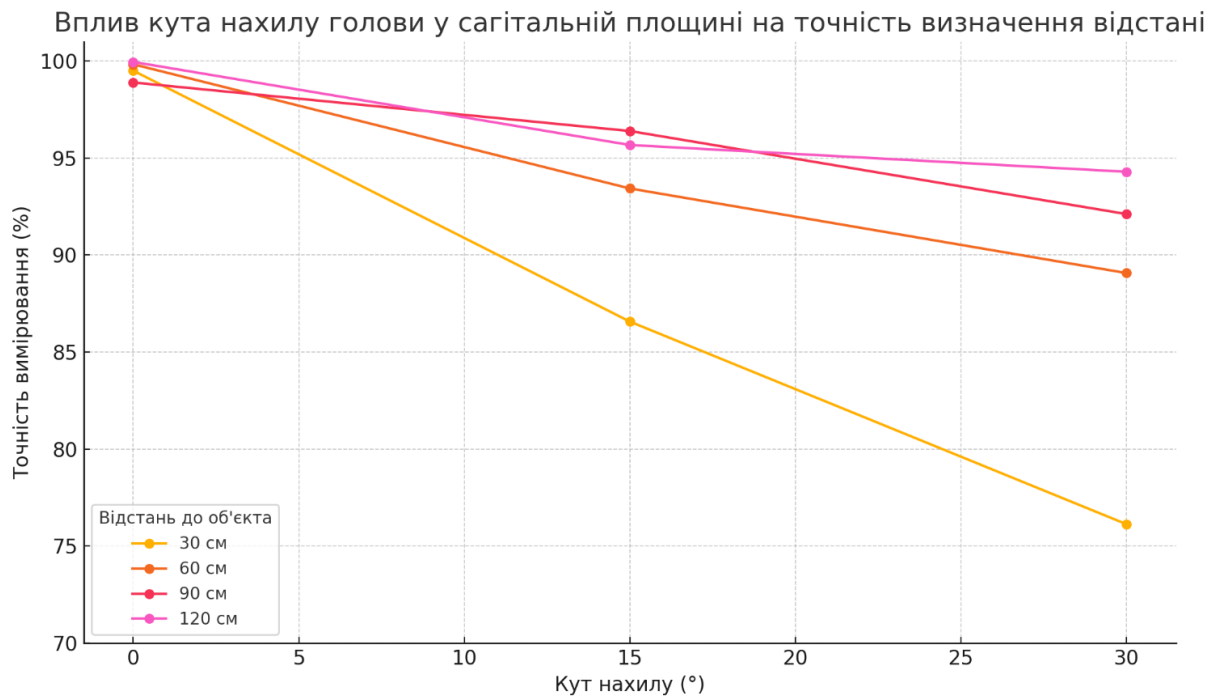


Рисунок 4.4 - Оцінка впливу нахилу голови на точність визначення відстані об'єкта від камери у сагітальній площині

Найбільше падіння точності зафіксовано на відстані 30 см — від 99.5% (0°) до 76.13% при 30°. Така різка деградація пояснюється надмірною чутливістю алгоритму до малих кутів при близькому розташуванні камери.

ВИСНОВКИ

В результаті виконання кваліфікаційної роботи було встановлено, що алгоритм FaceMesh/OpenCV забезпечує достатньо високу точність детектування відстані до об'єкта лише у наступних умовах: коли об'єкт розташований фронтально до камери без нахилів у будь-якій площині, камера працює при достатньому рівні освітлення (від 200 лк і вище), а також використовується оптимальна роздільна здатність, що забезпечує стабільну роботу алгоритму (у дослідженні — 640×480). За відхилення від цих умов точність помітно знижується, що слід враховувати при практичному застосуванні технології.

Подальші дослідження полягають в оцінці нейромережевих методів визначення відстані до об'єкту для подальшої імплементації обраного алгоритму в більш складні адаптивні моделі, які можуть автоматично враховувати положення людини та коригувати свої параметри на основі завчасно визначеній відстані.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Litjens, Geert, et al. ‘A Survey on Deep Learning in Medical Image Analysis’. *Medical Image Analysis*, vol. 42, Dec. 2017, pp. 60–88. DOI.org (Crossref), <https://doi.org/10.1016/j.media.2017.07.005>.
2. Izadi, Shahram, et al. ‘KinectFusion: Real-Time 3D Reconstruction and Interaction Using a Moving Depth Camera’. *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology*, ACM, 2011, pp. 559–68. DOI.org (Crossref), <https://doi.org/10.1145/2047196.2047270>.
3. Pires, J. Norberto, and Amin S. Azar. ‘Advances in Robotics for Additive/Hybrid Manufacturing: Robot Control, Speech Interface and Path Planning’. *Industrial Robot: An International Journal*, vol. 45, no. 3, May 2018, pp. 311–27. DOI.org (Crossref), <https://doi.org/10.1108/IR-01-2018-0017>.
4. Naik B. T., Hashmi M. F., Bokde N. D. A comprehensive review of computer vision in sports: Open issues, future trends and research directions // *Applied Sciences*. – 2022. – Т. 12. – №. 9. – С. 4429.
5. Geiger, A., et al. ‘Vision Meets Robotics: The KITTI Dataset’. *The International Journal of Robotics Research*, vol. 32, no. 11, Sept. 2013, pp. 1231–37. DOI.org (Crossref), <https://doi.org/10.1177/0278364913491297>.
6. Weiss, Ulrich, and Peter Biber. ‘Plant Detection and Mapping for Agricultural Robots Using a 3D LIDAR Sensor’. *Robotics and Autonomous Systems*, vol. 59, no. 5, May 2011, pp. 265–73. DOI.org (Crossref), <https://doi.org/10.1016/j.robot.2011.02.011>.
7. Постельняк Ю.П. МЕТОД ОЦІНЮВАННЯ ПРОСТОРОВОЇ ГЛИБИНИ ЗА КОНТРОЛЬНИМИ ТОЧКАМИ ОБЛИЧЧЯ В РЕАЛЬНОМУ ЧАСІ / Ю.П.Постельняк, О.Ю.Барковська // Сучасні напрями розвитку інформаційно-комунікаційних технологій та засобів управління : тези доп. 15-ї міжнар. наук.-техн. конф., 24-25 квітня 2025 р., Баку–Харків–Жиліна : [у 4 т.]. Т. 3 : секція 3,4 / Нац. ун-т оборони Азербайджанської Республіки [та

ін.] – Харків : Impress., 2025. – С. 45. <https://doi.org/10.32620/ICT.25.t3>

8. Park, Jong-Kyu, and Ki-Nam Joo. ‘High-Resolution 3D Imaging by Structured Illumination Microscopy’. Asian Society for Precision Engineering and Nanotechnology (ASPEN 2022), Research Publishing Services, 2022, pp. 507–09. DOI.org (Crossref), https://doi.org/10.3850/978-981-18-6021-8_OR-11-0305.html.

9. Bamji C. et al. A review of indirect time-of-flight technologies //IEEE Transactions on Electron Devices. – 2022. – Т. 69. – №. 6. – С. 2779-2793.

10. Horn, Berthold K. P., and B. G. Schunck. “‘Determining Optical Flow’’: A Retrospective’. Artificial Intelligence, vol. 59, no. 1–2, Feb. 1993, pp. 81–87. DOI.org (Crossref), [https://doi.org/10.1016/0004-3702\(93\)90173-9](https://doi.org/10.1016/0004-3702(93)90173-9).

11. Mertan A., Duff D. J., Unal G. Single image depth estimation: An overview //Digital Signal Processing. – 2022. – Т. 123. – С. 103441.

12. Lu, Yawen, et al. ‘Multi-Task Learning for Single Image Depth Estimation and Segmentation Based on Unsupervised Network’. 2020 IEEE International Conference on Robotics and Automation (ICRA), IEEE, 2020, pp. 10788–94. DOI.org (Crossref), <https://doi.org/10.1109/ICRA40945.2020.9196723>.

13. Wofk, Diana, et al. ‘FastDepth: Fast Monocular Depth Estimation on Embedded Systems’. 2019 International Conference on Robotics and Automation (ICRA), IEEE, 2019, pp. 6101–08. DOI.org (Crossref), <https://doi.org/10.1109/ICRA.2019.8794182>.

14. Farooq Bhat, Shariq, et al. ‘AdaBins: Depth Estimation Using Adaptive Bins’. 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, 2021, pp. 4008–17. DOI.org (Crossref), <https://doi.org/10.1109/CVPR46437.2021.00400>.

15. Ehret T. Monocular depth estimation: a review of the 2022 state of the art //Image Processing On Line. – 2023. – Т. 13. – С. 38-56.

16. Guizilini, Vitor, et al. ‘Full Surround Monodepth From Multiple Cameras’. IEEE Robotics and Automation Letters, vol. 7, no. 2, Apr. 2022, pp. 5397–404. DOI.org (Crossref), <https://doi.org/10.1109/LRA.2022.3150884>.

17. Lipton, Zachary C. ‘The Mythos of Model Interpretability: In Machine Learning, the Concept of Interpretability Is Both Important and Slippery.’ *Queue*, vol. 16, no. 3, June 2018, pp. 31–57. DOI.org (Crossref), <https://doi.org/10.1145/3236386.3241340>.

18. Castelle, Michael. ‘The Social Lives of Generative Adversarial Networks’. *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, ACM, 2020, pp. 413–413. DOI.org (Crossref), <https://doi.org/10.1145/3351095.3373156>.

19. Kaehler, Adrian, and Gary R. Bradski. *Learning OpenCV 3: Computer Vision in C++ with the OpenCV Library*. First edition, Second release, O’Reilly Media, 2017.