МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

Національний аерокосмічний університет ім. М.Є. Жуковського
«Харківський авіаційний інститут»

Факультет програмної інженерії та бізнесу

Кафедра інженерії програмного забезпечення

Пояснювальна записка до дипломного проекту

магістра
(освітній ступінь)

на тему «Дослідження ефективності використання систем розпізнавання жестів»

ХАІ.603.667П1.121.156311.20В

Виконав: студент 6 курсу групи № 667П1
Спеціальність 121 – Інженерія програмного забезпечення
(код та найменування)
Освітня програма Хмарні обчислення та Інтернет речей
(найменування)
Козюбердін В.С.
(прізвище й ініціали студента)
Керівник: Лучшев П.О.
(прізвище й ініціали)
Рецензент: Мартовицький В.О.
(прізвище й ініціали)

Харків – 2020

Міністерство світи і науки України
Національний аерокосмічний університет ім. М. Є. Жуковського
«Харківський авіаційний інститут»

Факультет програмної інженерії та бізнесу
(повне найменування)
Кафедра інженерії програмного забезпечення
(повне найменування)
Рівень вищої освіти другий (магістерський)
Спеціальність 121 – інженерія програмного забезпечення
(код та найменування)
Освітня програма хмарні обчислення та Інтернет речей
(найменування)

ЗАТВЕРДЖУЮ
Завідувач кафедри

_____ _____
(підпис) (ініціали та прізвище)
“_____” _____ 2020 року

З А В Д А Н Н Я
НА ДИПЛОМНИЙ ПРОЕКТ СТУДЕНТУ

Козюбердіну Віталію Сергійовичу
(прізвище, ім’я, по батькові)

1. Тема дипломного проекту Дослідження ефективності використання систем розпізнавання жестів
керівник дипломного проекту Лучшев Павло Олександрович, к.т.н., доцент
(прізвище, ім’я, по батькові, науковий ступінь, вчене звання)
затверджені наказом Університету № _____ від “_ _” _____ 2020 року
2. Термін подання студентом проекту _____
3. Вихідні дані до проекту виконаний аналіз існуючих алгоритмів та методів розпізнавання жестів та образів на зображенні, план експерименту, експеримент, оцінка ефективності використання систем розпізнавання жестів.
4. Зміст пояснювальної записки (перелік питань, які потрібно розробити):
1) Аналіз існуючих рішень, методів та алгоритмів.
2) Планування експерименту.
3) Виконання експерименту.
5. Перелік графічного матеріалу (з точним зазначенням обов’язкових креслень)
Пояснювальна записка дипломного проекту бакалавра містить 68 сторінок, окремо додатки 4 сторінок, 28 рисунків, 8 таблиць, 40 посилань на літературні джерела
1) Для першого розділу розроблено 5 слайдів презентації, 14 рисунків та 2 таблиці.
2) Для другого – 4 слайди презентації, 6 рисунків та 0 таблиць.
3) Для третього – 5 слайдів презентації, 0 рисунків та 6 таблиць.

6. Консультанти розділів проекту

Розділ	Прізвище, ініціали та посада консультанта	Підпис, дата	
		завдання видав	завдання прийняв
1	Лучшев П.О., к.т.н., доцент		
2	Лучшев П.О., к.т.н., доцент		
3	Лучшев П.О., к.т.н., доцент		

Нормоконтроль _____ _____ «___»_____20__р.
(підпис) (ініціали та прізвище)

7. Дата видачі завдання «___»_____20_____ р.

КАЛЕНДАРНИЙ ПЛАН

№ з/п	Назва етапів дипломного проекту	Строк виконання етапів проекту	Примітка
1	Отримання і затвердження теми диплому	01.09.2020	
2	Обгрунтування актуальності дослідження ефективності використання систем розпізнавання жестів	12.09.2020	
3	Формулювання мети, об’єкту, предмету та методів дослідження	15.09.2020	
4	Аналіз існуючих рішень, алгоритмів та методів розпізнавання жестів та об’єктів	20.09.2020	
5	Аналіз існуючих систем	01.10.2020	
6	Аналіз алгоритму та порядку обробки зображення	10.10.2020	
7	Планування експерименту	21.10.2020	
8	Виконання експерименту систем розпізнавання жестів з використанням веб-камери	05.11.2020	
10	Підготовка пояснювальної записки	20.11.2020	
11	Оформлення пояснювальної записки до дипломного проекту	22.11.2020	
12	Передзахист дипломного проекту	27.11.2020	
13	Захист дипломного проекту	04.12.2020	

Студент _____ Козюбердін В.С.
(підпис) (прізвище та ініціали)

Керівник проекту _____ Лучшев П.О.
(підпис) (прізвище та ініціали)

РЕФЕРАТ

Пояснительная записка дипломного проекта магистра содержит 68 страниц, отдельно приложения 4 страницы, 28 рисунков, 8 таблиц, 40 ссылок на литературные источники.

Цель дипломного проекта магистра – исследование и оценка систем распознавания жестов с использованием веб-камеры.

Для достижения цели необходимо выполнить следующие задачи:

- провести анализ способов изображения объектов;
- проанализировать существующие алгоритмы распознавания и отслеживания жестов и объектов;
- выполнить анализ методов считывания и обработки изображений;
- выполнить планирование эксперимента;
- выбрать и обосновать факторы, оценивающие эффективность систем распознавания жестов;
- провести тестирование систем распознавания жестов на предмет скорости и точности распознавания.

Актуальность задачи – создание легких в управлении человеко-машинных интерфейсов, которые обеспечат простоту и доступность в бесконтактном взаимодействии с компьютером при помощи жестов рук.

Основным практическим результатом работы является исследование инструментальных средств распознавания жестов, применение которых повысит эффективность коммуникации людей с проблемами слуха, а также врачам в бесконтактном управлении компьютером (для сохранения стерильности).

Необходимо провести анализ существующих алгоритмов, методов и решений, которые выполняют распознавание жестов и объектов на изображении, а также анализ существующих систем распознавания жестов с использованием веб-камеры.

ЖЕСТ, ОБРОБОТКА ИЗОБРАЖЕНИЯ, ОТОБРАЖЕНИЕ ОБЪЕКТОВ, ИДЕНТИФИКАЦИЯ ЖЕСТОВ, ОТСЛЕЖИВАНИЕ ОБЪЕКТОВ.

РЕФЕРАТ

Пояснювальна записка дипломного проекту магістра містить 68 сторінок, окремо додатки 4 сторінки, 28 рисунків, 8 таблиць, 40 посилання на літературні джерела.

Мета дипломного проекту магістра - дослідження і оцінка систем розпізнавання жестів з використанням веб-камери.

Для досягнення мети необхідно виконати наступні завдання:

- провести аналіз способів зображення об'єктів;
- проаналізувати існуючі алгоритми розпізнавання і відстеження жестів і об'єктів;
- виконати аналіз методів зчитування й обробки зображень;
- виконати планування експерименту;
- вибрати і обґрунтувати фактори, що оцінюють ефективність систем розпізнавання жестів;
- провести тестування систем розпізнавання жестів на предмет швидкості і точності розпізнавання.

Актуальність завдання – створення легких в управлінні людино-машинних інтерфейсів, які забезпечать простоту і доступність в безконтактній взаємодії з комп'ютером за допомогою жестів рук.

Основним практичним результатом роботи є дослідження інструментальних засобів розпізнавання жестів, застосування яких підвищить ефективність комунікації людей з вагами слуху, а також лікарям в безконтактному керуванні комп'ютером (для збереження стерильності).

Необхідно провести аналіз існуючих алгоритмів, методів і рішень, які виконують розпізнавання жестів і об'єктів на зображенні, а також аналіз існуючих систем розпізнавання жестів з використанням веб-камери.

ЖЕСТ, ОБРОБКА ЗОБРАЖЕННЯ, ВІДОБРАЖЕННЯ ОБ'ЄКТІВ, ІДЕНТИФІКАЦІЯ ЖЕСТІВ, ВІДСТЕЖЕННЯ ОБ'ЄКТІВ.

ABSTRACT

The explanatory note of the master's thesis project contains 68 pages, separate attachments of 4 pages, 28 pictures, 8 tables, 40 references to literary sources.

The aim of the Master's degree project is to research and evaluate gesture recognition systems using a webcam.

To achieve the goal, you must complete the following tasks:

- to analyze the ways of depicting objects;
- analyze existing algorithms for recognition and tracking of gestures and objects;
- to analyze the methods of reading and processing images;
- perform the planning of the experiment;
- select and justify the factors that assess the effectiveness of gesture recognition systems;
- to test the gesture recognition systems for the speed and accuracy of recognition.

The urgency of the problem is the creation of easy-to-use human-machine interfaces that will provide simplicity and accessibility in contactless interaction with a computer using hand gestures.

The main practical result of the work is the study of gesture recognition tools, the use of which will increase the efficiency of communication for people with hearing problems, as well as for doctors in contactless computer control (to maintain sterility).

It is necessary to analyze existing algorithms, methods, and solutions that perform recognition of gestures and objects in an image, as well as an analysis of existing gesture recognition systems using a webcam.

GESTURE, IMAGE PROCESSING, DISPLAYING OBJECTS, IDENTIFICATION OF GESTURES, TRACKING OBJECTS.

ЗМІСТ

ПЕРЕЛІК УМОВНИХ СКОРОЧЕНЬ 9
ВСТУП 10
1 АНАЛІЗ ІСНУЮЧИХ РІШЕНЬ, МЕТОДІВ ТА АЛГОРИТМІВ 12
1.1 Аналіз алгоритмів розпізнавання жестів 12
1.1.1 Локалізація жестів 13
1.1.2 Гістограма кольорів 13
1.2 Аналіз алгоритмів відстеження об’єктів 14
1.3 Аналіз методів зчитування та обробки зображення 15
1.3.1 Методи засновані на зображеннях 15
1.3.2 Методи не засновані на зображеннях 19
1.3.3 Порівняння методів зчитування даних 21
1.3.4 Відстеження контурів 21
1.3.5 Ідентифікація жестів 22
1.4 Аналіз структури системи розпізнавання жестів 25
1.5 Сучасні рішення 26
1.5.1 Kinect Sign Language Translator 26
1.5.2 Motionsavvy – планшет для глухонімих з розпізнаванням жестів 27
1.5.3 LeapMotion 28
1.5.4 MediaPipe 29
1.5.5 Myo 30
1.6 Висновки з розділу 1 31
2 ПЛАНУВАННЯ ЕКСПЕРИМЕНТУ 32
2.1 Алгоритм і порядок обробки 32
2.2 Вибір жестів руки для проведення експерименту 34
2.3 Вибір типу жестів для проведення експерименту 36
2.4 Вибір рухів руки для проведення експерименту 36
2.5 Вибір зовнішніх умов для проведення експерименту 36
2.6 Фактори, які впливають на розпізнавання жестів 36
2.7 Вибір пристрою для проведення експерименту 37
2.8 Висновки з розділу 2 38
3 ТЕСТУВАННЯ СИСТЕМ РОЗПІЗНАВАННЯ ЖЕСТІВ НА ОСНОВІ НЕЙРОМЕРЕЖЕВИХ АЛГОРИТМІВ АНАЛІЗУ ЗОБРАЖЕНЬ 39
3.1 Тестування систем на розпізнавання статичних жестів 39
3.1.1 Тестування системи MediaPipe на розпізнавання статичних жестів 39
3.1.2 Результати тестування системи MediaPipe на розпізнавання статичних жестів 42
3.1.3 Тестування системи Handpose на розпізнавання статичних жестів 44
3.1.4 Результати тестування системи Handpose на розпізнавання статичних жестів 47
3.2 Тестування систем на розпізнавання динамічних жестів 49
3.2.1 Тестування системи MediaPipe на розпізнавання динамічних жестів 49

3.2.2 Результати тестування системи MediaPipe на розпізнавання динамічних статичних жестів 53
3.2.3 Тестування системи Handpose на розпізнавання динамічних жестів 55
3.2.4 Результати тестування системи Handpose на розпізнавання динамічних жестів 59
3.3 Оцінка ефективності систем розпізнавання жестів 61
3.3.1 Оцінка ефективності на розпізнавання статичних жестів 61
3.3.2 Оцінка ефективності розпізнавання динамічних жестів 62
3.4 Висновки з розділу 3 63
ВИСНОВКИ 64
ПЕРЕЛІК ПОСИЛАНЬ 66
ДОДАТОК А 69
ДОДАТОК Б 71

ПЕРЕЛІК УМОВНИХ СКОРОЧЕНЬ

МЖ – мова жестів.

НМ – нейронна мережа.

ПК – персональний комп’ютер.

ІЧ-пристрій – інфрачервоний пристрій.

Мікроелектромеханічні системи, МЕМС (англ. MEMS) – технології і пристрої, що поєднують в собі мікроелектронні і мікромеханічні компоненти.

HD (High Definition) – дозвіл зображення 1280x720.

JPEG (Joint Photographic Experts Group) – растровий формат збереження графічної інформації, що використовує стиснення з втратами якості зображення.

MJPEG (Motion JPEG) – покадровий метод відеостиснення, основною особливістю якого є стиснення кожного окремого кадру відеопослідовності за допомогою алгоритму стиснення зображень JPEG. При стисненні методом MJPG міжкадрова різниця не враховується.

ВСТУП

Актуальність дослідження. Розпізнавання руху - це велике завдання комп'ютерної науки, яка має безліч варіацій, в кожній з яких існує величезна кількість підзадач і різних підходів до вирішення.

Відносно мобільних платформ, поняття «Розпізнавання руху» зводиться, як правило, до розпізнавання жестів пальцями на сенсорному екрані для управління програмним забезпеченням.

Створення природних та легких в управлінні людино-машинних інтерфейсів для різних додатків є актуальною науковою задачею. В даний час зроблено досить багато досліджень по створенню методів розпізнавання образів, що дозволяють безконтактно взаємодіяти з комп'ютером за допомогою жестів рук.

На основі розроблених методів створено нечисленні практичні додатки, найбільш значущі з яких:

- програмні бібліотеки Kinect SDK (Microsoft), OpenNI / NITE (PrimeSense), PMD SDK і SoftKinetic IISU для розпізнавання ключових точок тіла людини за допомогою застосування тривимірних сенсорів;
- сенсори компаній LeapMotion, SoftKinetic для розпізнавання ключових точок руки людини;
- комп'ютерні програми Flutter, PointGrab HGRS, XTR3D, EyeSight PC для управління комп'ютером за допомогою простих жестів рук.

Незважаючи на окремі успіхи, якість розроблених алгоритмів розпізнавання жестів рук і пальців, як динамічних, так і статичних, з використанням кольорових відеокамер і тривимірних сенсорів все ще залишається недостатнім для побудови практичних систем людино- машинної взаємодії. Головними недоліками існуючих методів є чутливість до змін освітлення, потреба навчання системи для кожного оператора, невисока якість розпізнавання жестів і невелика швидкість розпізнавання.

Мета дослідження. Метою магістерської роботи є дослідження та оцінка систем розпізнавання жестів з використанням веб-камери. Аналіз напрацьованих і застосовуваних методів відстеження та розпізнавання образів (жестів руки), а також дистанційної людино-машинної взаємодії. Дослідження і розробка методів опису, вилучення та розпізнавання конфігурацій кисті руки.

Об’єкт дослідження. Сучасні процеси розпізнавання статичних та динамічних жестів з використанням веб-камери.

Предмет дослідження. Системи розпізнавання жестів з використанням веб-камери: MediaPipe та Handpose.

Методи дослідження. Методами дослідження було обрано аналіз існуючих сучасних алгоритмів та методі, моделювання рішень для розпізнавання статичних та динамічних жестів, та проведення експерименту.

Наукова новизна. Подальший розвиток отримав дослідницький інструментарій для оцінки ефективності використання систем розпізнавання

жестів з використанням веб-камери, що дає можливість людям з вагами слуху вільно комунікувати із навколишнім світом.

Практична значущість. Практичне значення результатів дослідження полягає в використанні отриманих результатів наукового дослідження для вирішення задач із розпізнаванням статичних та динамічних жестів.

1 АНАЛІЗ ІСНУЮЧИХ РІШЕНЬ, МЕТОДІВ ТА АЛГОРИТМІВ

1.1 Аналіз алгоритмів розпізнавання жестів

В ході аналізу алгоритмів виявлення об'єктів було встановлено, що вони прості в реалізації, але більша частина з них володіє низькою ефективністю і схильна до сильного впливу зовнішніх факторів, таких як: освітлення, фон, розмір об'єкта та інше. Для підвищення ефективності застосування даних алгоритмів, необхідно їх комбінувати, що дозволить підвищити точність локалізації об'єкта на зображенні.

Алгоритми відстежування об'єктів мають набір суттєвих переваг, однак недоліком є висока обчислювальна складність і необхідність дотримання певних умов. Дані алгоритми доцільно поєднувати з методами виявлення, для коректного відстеження об'єктів в відеопослідовності.

В даний час існує безліч різних методів виявлення об'єкта на зображенні. Ці методи можна розділити на три основні групи:

- скелетні методи;
- методи на основі 2D моделі об'єкта;
- методи на основі 3D моделі об'єкта.

У скелетних методах досліджується контур силуету: зазвичай проводиться пошук кутів, виступів, западин та інших точок з високими значеннями кривизни. Для отримання інформації про форму контуру застосовуються різні уявлення кордону об'єкта.

Метод 2D розпізнавання схожий з попереднім, але використовує двовимірне зображення замість об'ємних моделей.

У методах на основі 3D моделі руки представляють у вигляді складних тривимірних поверхонь і класифікуються за допомогою нейронних мереж.

Кожен з методів має переваги і недоліки. Оскільки при розробці алгоритму об'єктом виявлення на зображенні є рука, то перший метод можна виключити з-за незручності використання контуру для визначення конкретного жесту.

Недоліком методу на основі 3D моделі є його ресурсомісткість. Побудова 3D моделі, навчання нейронної мережі та її використання можуть вимагати значних ресурсів, так само не варто забувати, що для використання даного методу вимагає камери з можливістю визначення глибини зображення.

Тому найчастіше використовуються 2D методи, так як в них знижується обчислювальна складність і відпадає необхідність в спеціальному обладнанні, оскільки для отримання зображення може бути використана звичайна веб-камера.

Оскільки 2D методи є менш точними, виникає необхідність в їх удосконаленні для отримання ефективного алгоритму розпізнавання об'єктів.

1.1.1 Локалізація жестів

Даний метод полягає в тому, щоб виділити пікселі кольору шкіри із зображення. Дані про колір шкіри можна отримати із кольору обличчя. Результатами дослідження були незадовільні, так як усі тестові данні були у схожих до шкіри тонах.

До стандартних каскадів Хаара також входять класифікатори долоні, кулака, руки тощо, але вони працюють набагато гірше аніж відповідний класифікатор обличчя. Зауважимо, що каскади Хаара дуже чуттєві до кута нахилу зображення, а отже, за допомогою них неможливо зробити класифікатор рук у всіх можливих положеннях.

1.1.2 Гістограма кольорів

Ідея методу гістограм кольорів для порівняння зображень зводиться до наступного. Уся множина кольорів розбивається на набір непересічних, які повністю охоплюють зображення підмножин. Для зображення формується гістограма, що відображає частку кожної підмножини кольорів в кольоровій гамі зображення. Для порівняння гістограм вводиться поняття відстані між ними.

При розбитті RGB-кольорів за яскравістю обчислюється інтенсивність кожного кольору. Отримане значення, вкладене в проміжок між 0 і 255, потрапляє в один з 16 інтервалів, на які розбивається діапазон можливих значень. Як відстань між гістограмами використовується сума модулів різниці відповідних елементів гістограм; деякий удосконалення методу досягається при обчислюванні відстані на підставі по елементного порівняння гістограм з урахуванням сусідніх елементів.

Цей метод найбільш ефективний для чорно-білих зображень. Для кольорових RGB-зображень кращі результати дає інший спосіб - розбиття RGB-кольорів на прямокутні паралелепіпеди.

Кольоровий RGB-простір розглядається як тривимірний куб, кожна вісь якого відповідає одному з трьох основних кольорів (червоний, зелений або синій. При такому розгляді будь-який колір RGB-зображення може бути представлений точкою куба. Для побудови гістограми кольорів кожна сторона ділиться на 4 рівних інтервалу, відповідно RGB-куб ділиться на 64 прямокутних паралелепіпеда.

Гістограма зображення відображає розподіл точок RGB-простору, відповідних кольорам пікселів зображення. Як відстань між гістограмами використовується по компонентна сума модулів різниці між ними. Незважаючи на граничну простоту підходу, він показує досить стабільні результати.

Основним недоліком методу колірних гістограм є те, що він втрачає інформацію про просторове розташування об'єктів. Абсолютно різні картинки можуть мати подібні гістограми кольорів. Наприклад, зображення осіннього

листя може містити багато невеликих плям червоного кольору. Це дасть подібну колірну гістограму із зображенням великого червоного об'єкта.

1.2 Аналіз алгоритмів відстеження об’єктів

Для аналізу обрано наступні найбільш часто використовувані алгоритми для уловлювання і відстеження рухомих об'єктів [7]:

- шаблонів рухів (Motion Templates) - заснований на пошуку меж об'єктів в кожному кадрі відеопослідовності. Зсув кордону на новому кадрі щодо попереднього задає вектор руху об'єкта. Даний алгоритм найбільш ефективний при пересуванні великих об'єктів і часто використовується для розпізнавання динамічних жестів в людино-машинних інтерфейсів;
- зсуву середнього (Mean-Shift) - заснований на математичній моделі, яка полягає в тому, що обчислюється локальний екстремум щільності розподілу набору характерних точок, тобто алгоритм відстежує зміщення центру мас точок, що визначають об'єкт спостереження, отримуючи на виході вектор руху об'єкта. Висока ефективність досягається при яскраво вираженому колірному відмінності об'єкта і фону;
- зсуву, який безперервно адаптується (CamShift) - заснований на алгоритмі зсуву середнього, але відрізняється тим, що автоматично підлаштовує кордони і розмір вікна, в межах якого розташовані характерні точки. Таким чином проводиться більш точне відстеження об'єкта, що змінюється в розмірах.

1.3 Аналіз методів зчитування та обробки зображення

Серед методів розпізнавання існують дві основні категорії збору даних:
- засновані на зображеннях
- не засновані на зображеннях.

1.3.1 Методи засновані на зображеннях

Підходи, засновані на роботі з зображеннями, далі поділяються на наступні категорії:
- Маркери;
- Стереокамери;
- Камери глибини;
- ІЧ-пристрої;
- 2D камери.

1.3.1.1 Захват рухів на основі маркерів

У маркерному підході датчиком є оптична камера. У більшості рішень на основі маркерів користувачам необхідно носити видимі позначки (рисунок 1.1).

Для використання маркерної системи необхідне спеціальне обладнання: датчики, камери.

Датчики фіксуються на руці людини, будь-яке переміщення руки в просторі фіксується за допомогою камер. Потік зібраної інформації надходить в комп'ютер.

Після обробки та аналізу переданої інформації народжується цілісна тривимірна модель, точно передає жести людини, на основі якої відразу або пізніше прописується остаточна анімація жестів.

Розглянемо кілька прикладів реалізації захоплення руху з використанням маркерів.

Вчені Інституту інформатики суспільства Макса Планка вирішували завдання відстеження руки бейсбольного пітчера за допомогою закріплення кольорових маркерів, розміщених на зворотному боці рукавички.

Для зчитування інформації були використані чотири камери і стробоскоп. Співробітники університету імені Саймона Фрейзера (Канада) застосували рукавички з різнокольоровими кільцями для розпізнавання жестів абетки. Була проведена паралель між суглобом кожного пальця і кільцем. В основу розпізнавання була покладена зворотна кінематика. Поза руки обчислювалась через визначення спільних позицій датчиків.

Співробітники Масачусецкого Технологічного Інституту в своїх тестах використовували різнокольорові рукавички. Завдання стало значно простіше, на її рішення не було потрібно витрачати багато часу, що дозволяло з невеликою затримкою вибудовувати моделі руки на підставі зображення.

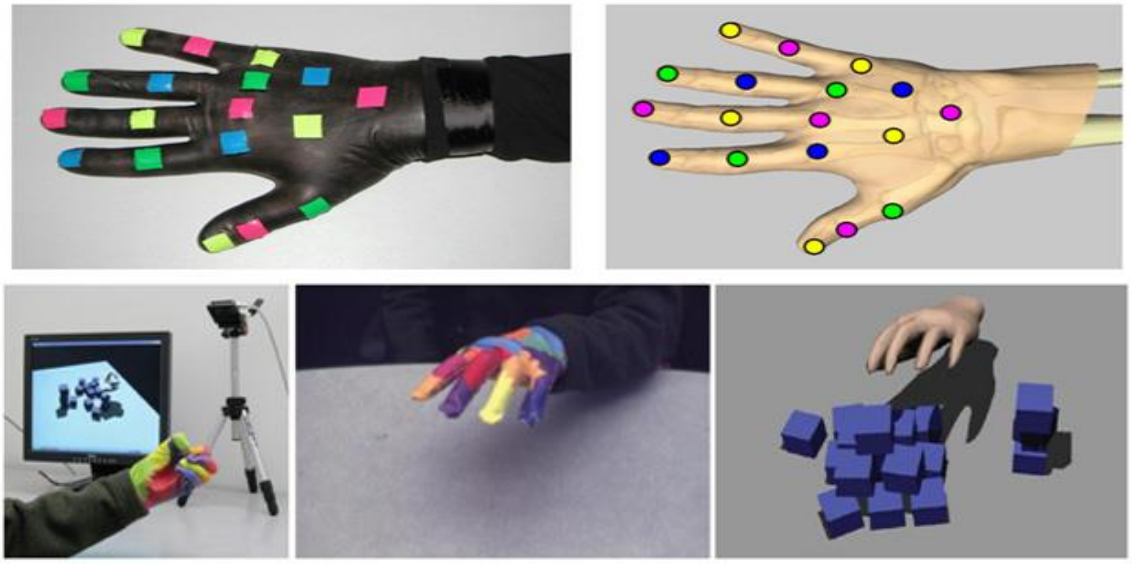

Рисунок 1.1 – Приклад отримання зображення за допомогою маркерів

1.3.1.2 Отримання тривимірного зображення за допомогою стереокамери

Стереоскопічний фотоапарат (стереофотоаппарата, стереокамера) - тип фотоапарата з двома або більше об'єктивами, що створюють окремі кадри на фотоплівці або матриці.

Це дає змогу оперативно симулювати людський бінокулярний зір і здійснювати стереоскопічну фотозйомку, даючи тривимірні (об'ємні) зображення.

Стереокамери можуть використовуватися для створення стереообзорів, 3D-зображень для фільмів або для діапазонного відображення.

Відстань між об'єктивами в стереокамери (стереобаза) таке ж, як і відстань між очима людини і становить близько 64 мм. Зі збільшенням відстані до об'єкта фотозйомки - «глибина» зменшується.

Всі фотоапарати зі статичної стереобазою також, як і очі людини, в обсязі «бачать» тільки близькі об'єкти (наприклад, подивіться на Місяць, він здається плоским). Для 3D-фотозйомки віддалених об'єктів необхідно збільшувати відстані стереобази.

Використовуючи дві камери, положення яких одна відносно іншої відомо заздалегідь, можна отримати тривимірне зображення об’єкту спостереження (рисунок 1.2).

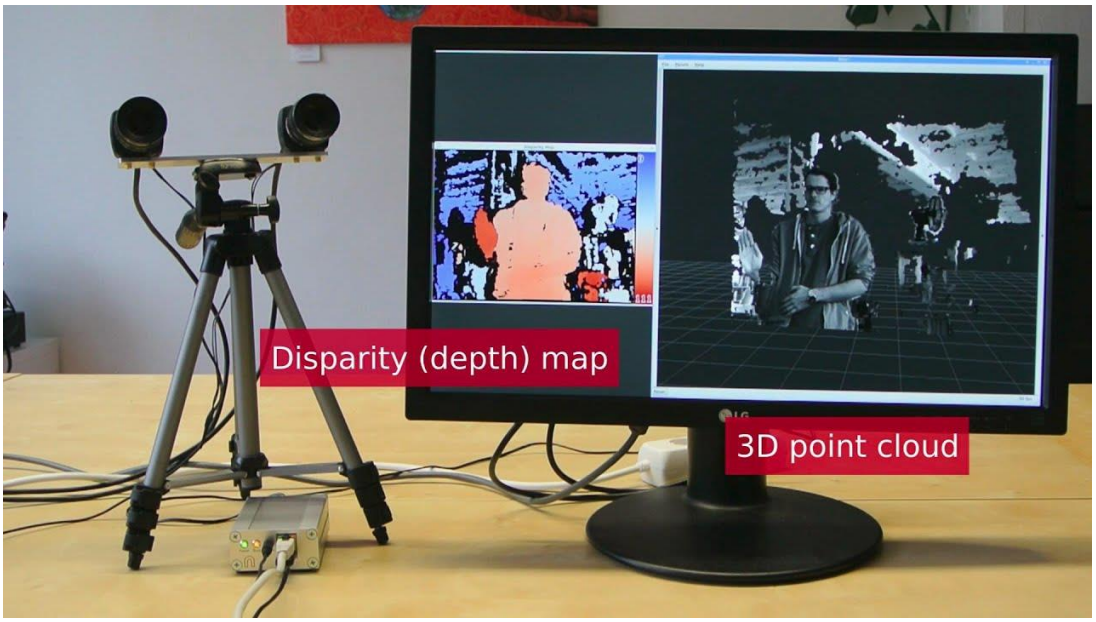

Рисунок 1.2 – Приклад роботи стереокамери

1.3.1.3 Отримання зображення за допомогою камери глибини

Камери глибини мають ряд переваг в порівнянні з традиційними стереокамерами. Наприклад, можна запобігти проблемам налаштування калібрування і умов освітлення (рисунок 1.3).

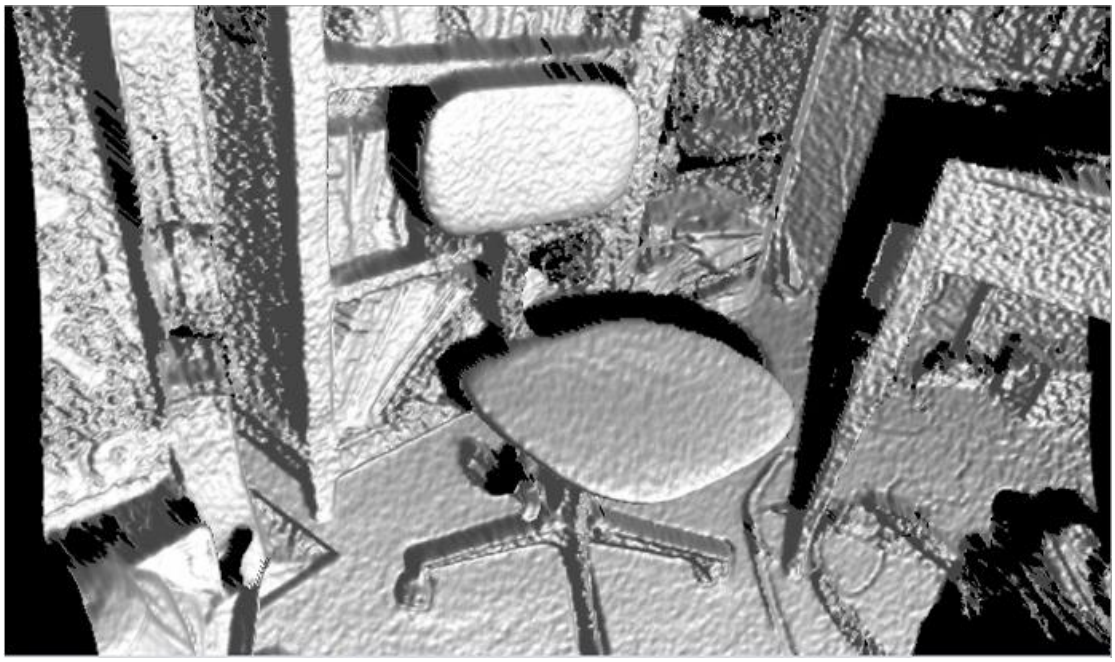

Рисунок 1.3 – Приклад роботи камери глибини

1.3.1.4 Отримання моделі об’єкту за допомогою ІЧ-пристроїв

В основу функціонування інфрачервоних пристроїв (ІЧ-пристроїв) руху покладена їх здатність спрацьовувати в випадках, коли виникає інтенсивне теплове фонове випромінювання.

Пристрої реагують, коли джерело випромінювання потрапляє в безпосередню зону дії пристрою. Точність роботи приладу залежить від місця розташування об'єкта, що викликав спрацьовування. Важливо врахувати, що теплове випромінювання виробляють не тільки люди, але і тварини, і неживі предмети.

Деякі пристрої можуть відтворити 3д модель об’єкту за допомогою інфрачервоних датчиків (рисунок 1.4).

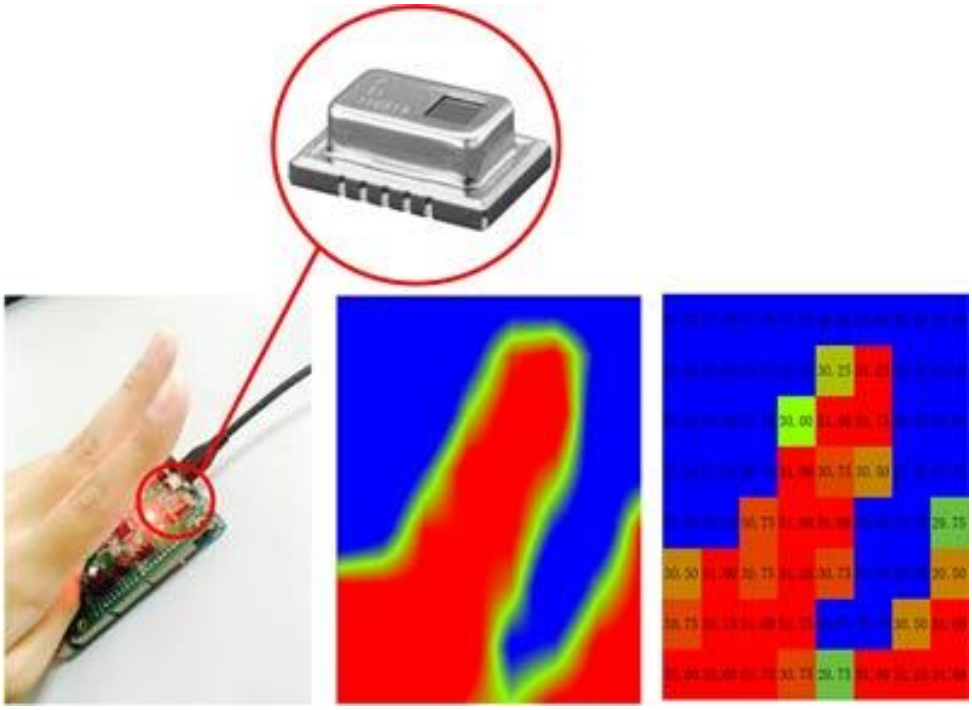

Рисунок – 1.4 – Приклад роботи ІЧ-пристрою

1.3.1.5 Розпізнавання жестів за допомогою 2D камери

Звичайна 2д камера може бути використана для розпізнавання жестів. Вважається, що даний метод отримання інформації не являється настільки ж ефективним як інші, але він являється найбільш доступним для людей.

Недоліком однокамерного підходу є обмеження кута огляду, що впливає на надійність системи. Однак у недавніх дослідженнях застосовувався однокамерний підхід у високошвидкісного розпізнаванню жестів. Система використовує датчик швидкості і спеціально розроблений процесор візуальних обчислень для досягнення високошвидкісного розпізнавання жестів.

1.3.2 Методи не засновані на зображеннях

У розпізнаванні жестів протягом довгого часу домінували датчики на основі зображень. Розробки в мікроелектромеханічних системах (МЕМС) і сенсорних технологіях значно поліпшили технології розпізнавання жестів, засновані не на зображеннях.

Підходи, засновані не на роботі із зображеннями, далі поділяються на наступні категорії:

- Перчатки;
- Браслети;
- Безконтактні пристрої.

1.3.2.1 Розпізнавання жестів руки за допомогою обладнаної перчатки

На чолі цього методу лежить використання обладнаної рукавички, оснащеної безліччю різноманітних датчиків (температури, вібрації, тиску та ін.), що дозволяють повторювати біомеханіку людської руки (рисунок 1.5).

Інформація про рух рук зчитується з датчиків і передається на обробку ПК. Отримана інформація потім аналізується і перетворюється в символи. Дані про характеристики окремих жестів використовується системою в якості шаблонів.

Жестові інтерфейси на основі перчаток також використовуються для розпізнавання жестів. Зазвичай методи на основі перчаток вимагають проводового підключення акселерометрів і гіроскопів. Підходи на основі перчаток також мають складності в процедурах калібрування і налаштування.

До переваг можна віднести високу точність відстеження рук і як наслідок ідентифікацію жестів. Поряд з цим є ряд істотних недоліків. Перший - ціна. Другий - громіздкість самої рукавички і як результат обмеження рухів.

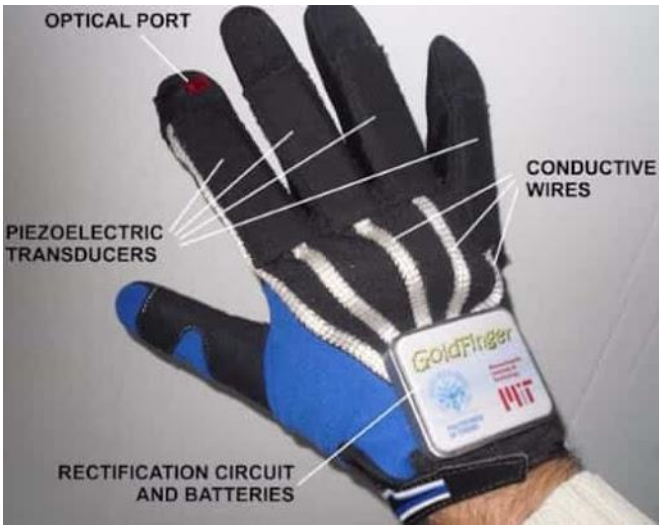

Рисунок 1.5 – Перчатки

1.3.2.2 Розпізнавання жестів руки за допомогою браслету

Інша безконтактна технологія використовує сенсори на браслетах. Сенсори встановлені на браслеті або аналогічних пристроях.

Сенсори на браслетах дозволяють використовувати бездротові технології і датчики, що дозволяє уникнути підключення кабелів. Сенсори повинні контактувати з зап'ястям, руки і пальці користувача можуть бути вільними (рисунок 1.6).

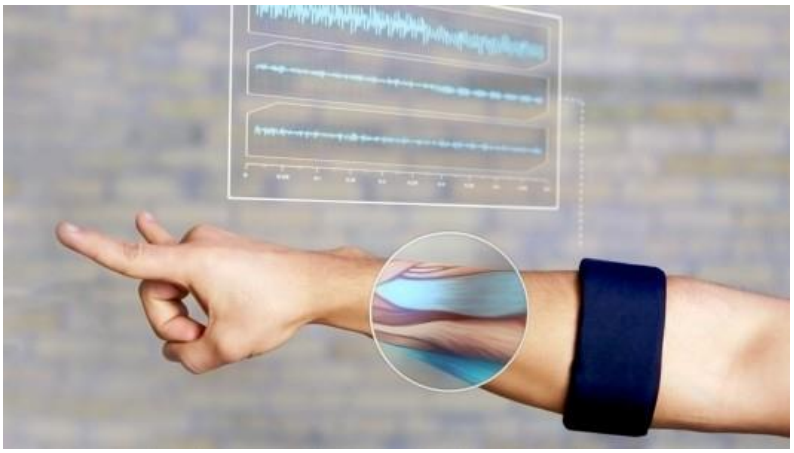

Рисунок 1.6 – Приклад роботи браслету

1.3.2.3 Розпізнавання жестів руки за допомогою безконтактних датчиків

У третьому типі технологій, не пов'язаних з зображеннями, використовуються датчики, не призначені для носіння.

Безконтактні датчики можуть виявляти жести без контакту з людським тілом. Google представив Project Soli (рисунок 1.7), систему контролю радіолокації і розпізнавання жестів на радіочастотному спектрі (радар). Пристрій здатний розпізнавати різні жести рук на невеликій відстані.

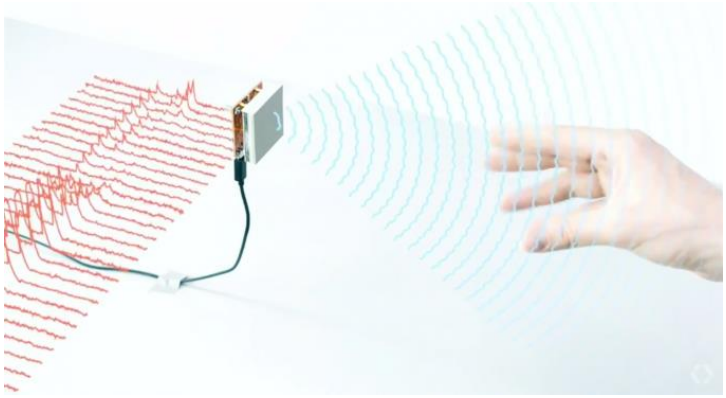

Рисунок 1.7 – Система Project Soli

1.3.3 Порівняння методів зчитування даних

Таблиця 1.1 містить порівняння різних сенсорних технологій. Відображено переваги та недоліки різних підходів.

Таблиця 1.1 – Переваги та недоліки методів зчитування даних

	Переваги	Недоліки
Маркери	Низьке обчислювальне навантаження	Маркери на тілі користувача
2D Камера	Легкість установки	Низький рівень надійності
Стереокамера	Надійність	Дозвіл залежить від потужності світла і відображення
Перчатка	Швидкість відгуку, точність трекінгу	Громіздкий пристрій з проводами
Браслет	Швидкість відгуку, область дії	Браслет повинен контактувати з людським тілом
Безконтактні пристрої	Не потребує контакту з тілом	Низький дозвіл, технологія недостатньо зріла

1.3.4 Відстеження контурів

Метод полягає в послідовному кресленні кордону між об'єктом і фоном. Простежує точка у вигляді «жука» повзає по зображенню до тих пір, поки не доходить до темної області (об'єкт). Тоді «жук» повертається ліворуч і рухається по кривій, поки не досягне меж об'єкта, після цього повертається направо і повторює процес, поки не досягне околиці початкової точки (рисунок 1.8).

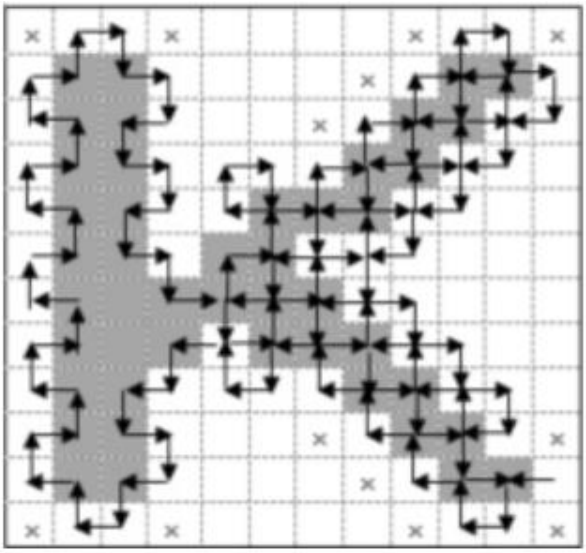

Рисунок 1.8 – Приклад роботи алгоритму «Жука»

1.3.5 Ідентифікація жестів

Ідентифікація жесту - перший крок в процесі розпізнавання жестів після отримання необроблених даних, отриманих з датчиків. Ідентифікація жесту означає виявлення жестової інформації та сегментацію відповідної жестової інформації з необроблених даних. Популярні технології для вирішення проблеми ідентифікації жестів засновані на візуальних особливості, алгоритмах навчання і людських моделях.

1.3.5.1 Візуальні характеристики

Людські руки і тіло мають унікальні візуальними особливостями. У розпізнаванні жестів на основі зображень жести складаються з фрагментів людських рук і / або тіла [5]. Тому використання таких візуальних ознак в ідентифікації жестів цілком обґрунтовано.

- Колір. Колір - це проста візуальна функція для ідентифікації жестів з фонової інформації. Поширена проблема у виявленні кольору шкіри полягає в тому, що колір шкіри людини сильно різниться серед людських рас. Через це, в сучасних підходах, колір шкіри розглядається тільки як один з багатьох параметрів при ідентифікації жестів.
- Локальні ознаки. У розпізнаванні жестів на основі зображення умови освітлення сильно впливають на якість ідентифікації жестів. Тому багато дослідників використовують метод локальних ознак, який не чутливий до умов освітлення. Локальний підхід до об'єктів - це деталізований підхід на основі текстур. Він розкладає зображення на більш дрібні області, які не відповідають частинам тіла.
- Форма і контур. Інший інтуїтивний і простий спосіб ідентифікації жестів - використання унікальної форму і контуру людського тіла.
- Рух. Рух є корисною функцією для виявлення людських жестів. Оптичний потік є ключовою технологією ідентифікації жестів на основі руху. Оптичний потік не потребує відніманні фону, що є перевагою в порівнянні з підходами на основі форми та контуру.

1.3.5.2 Алгоритми навчання

Ідентифікація жестів полягає в застосуванні алгоритмів навчання, особливо для виявлення статичних жестів, які можуть бути представлені в одному кадрі. Методи візуальної функції засновані на різних візуальних особливості, в той час як алгоритми навчання використовують алгоритми машинного навчання для визначення жестів з необроблених даних датчика. Хоча деякі алгоритми засновані на методах візуальних ознак, видалення фону зображення не завжди необхідно для коректної роботи алгоритмів. У системах розпізнавання жестів широко застосовуються алгоритми навчання, такі як метод опорних векторів, штучні нейронні мережі і випадкові рішення.

1.3.5.3 Модель людини

На відміну від вищезгаданих підходів, модельний підхід використовує явну модель людського тіла для відновлення пози людського тіла. Тому модельний підхід також називається генеративних підходом. Оскільки ідентифікація жестів на основі моделі людського тіла дає перевагу спрощення процесу класифікації жестів, підхід з використанням моделі людини став популярним рішенням при роботі з датчиками глибини. Сучасні дослідження ідентифікації жестів на основі моделі людини можна розділити на два типи: ідентифікація моделі рук і ідентифікація моделі скелета тіла.

- Модель руки. Існує три основні підходи до моделювання ручних моделей: модель форми, тривимірна модель і модель ручного скелета. Підхід, заснований на формі, відповідає попередньо побудованій формі руки з спостережень.
- Модель тілесного скелета. Щоб визначити жест тіла, детальна людська модель не потрібна. Однак, в завданню ідентифікації жестів тіла, модель самого скелета тіла є популярним підходом. Велика частина літератури по алгоритмам для обробки скелета тіла заснована на інформації про глибину, зібраної з датчиків глибини.

1.3.5.4 Методи ідентифікації жестів для різноманітних датчиків

Таблиця 1.2 представляє найбільш підходящі способи ідентифікації жестів для кожного популярного датчика.

Незважаючи на розуміння жестів людського тіла, підхід моделювання людського тіла також контролюватиме людські руху, що забезпечує безпечне середовище для спільної роботи роботів і людей. Як згадувалося раніше, модель скелета тіла спрощує тілесну модель людини, а інформація про суглоби тіла легко витягується. Більш того, скелетні підходи можуть спростити подальший процес класифікації жестів. Тому в даний час моделі з використанням скелетного підходу є підходящим рішенням для розпізнавання жестів.

Таблиця 1.2 – Методи ідентифікації жестів для різних датчиків

Датчик	Метод ідентифікації жестів
Системи з однією камерою	У системах з однією камерою можуть бути реалізовані методи візуальних ознак і алгоритми навчання. Для забезпечення надійної роботи слід застосовувати алгоритми навчання. Щоб домогтися більш швидкої обробки зображень, слід застосовувати метод візуальних ознак.
Датчик глибини	Оскільки метод людської моделі використовує і спрощує інформацію з хмари точок, метод людської моделі є найкращим варіантом для систем з датчиками глибини.
Браслет	Ніякі візуальні методи не можуть застосовуватися для систем на базі браслетів. Зазвичай зібрані дані вимагають базової фільтрації для ідентифікації жестів, а алгоритми навчання будуть реалізовані на більш пізній стадії класифікації жестів.
Безконтактні датчики	Безконтактні датчики також приймають дані сигналів замість зображень. Через те, що радіочастотні сигнали містять шуми, сучасні системи фільтрації і обробки повинні бути реалізовані в безконтактних системах.

1.4 Аналіз структури системи розпізнавання жестів

Структурна схема системи розпізнавання жестів з використанням web-камери [40] представлена на рисунку 1.9:

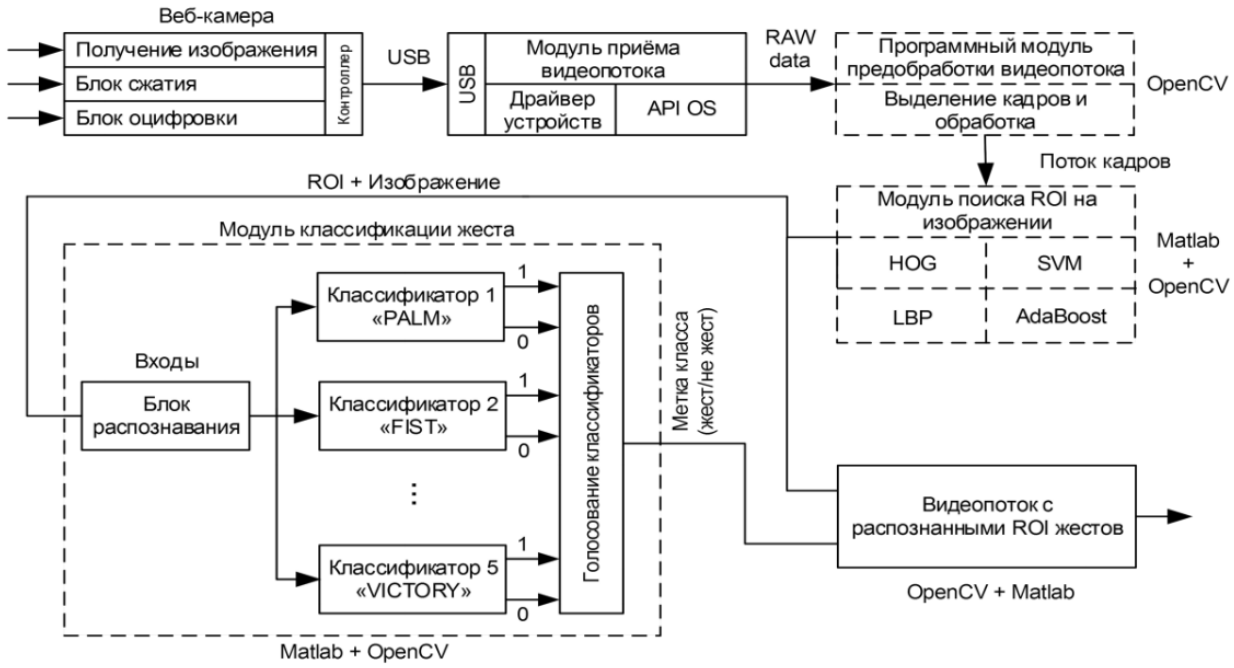

Рисунок 1.9 – Структура системи розпізнавання жестів

Система складається з веб-камери, прикладної програми з модулем розпізнавання і пристрою виведення інформації.

Після зняття зображення дані передаються по інтерфейсу USB з наступним завантаженням в модуль прийому відеопотоку, який далі передає необроблені кадри в програмний модуль попередньої обробки відеопотоку.

Після попередньої обробки відбувається виділення кадрів і їх обробка для пошуку області інтересу (Region of Interest, ROI) на зображенні. Як тільки ROI знайдена, здійснюється пошук ознак гістограми спрямованих градієнтів (Histogram of Oriented Gradients, HOG) або локальних бінарних шаблонів (Local Binary Patterns, LBP) для подальшої класифікації жесту в модулі.

Набір ознак повинен бути достатнім для ідентифікації класу об'єкта, і він представляється у вигляді n-мірного вектору. Сам класифікатор видасть «1», якщо жест розпізнано, «0» якщо немає. Голосування класифікаторів виконується вибором класифікатора з максимальним відгуком, потім отриманий результат виводиться на екран комп'ютера.

1.5 Сучасні рішення

1.5.1 Kinect Sign Language Translator

Microsoft Kinect - універсальний пристрій, що давно вийшло за межі комп'ютерних ігор. У цей час відділення Microsoft Research використовувало контролер для створення системи трансляції мовних жестів у словах розмовної речі та навпаки.

Прототип Kinect Sign Language Translator [8] вміє розпізнавати жести в реальному часі, а система машинного навчання і розуміння шаблонів мови може інтерпретувати їх зміст. Система може працювати в обох напрямках - для глухих людей переводити мова в набір жестів і навпаки, жести в написанні і вимовлені слова.

Рисунок 1.10 – Приклад роботи

1.5.2 Motionsavvy – планшет для глухонімих з розпізнаванням жестів

На апаратному рівні розробка являє собою корпус для планшета, в який інтегрований датчик рухів Leap Motion. Відповідне мобільний додаток аналізує руху рук користувача і розуміє мову жестів, переводячи його на будь-якій інший.

Зараз MotionSavvy [15] здатна переводити американський варіант МЖ на англійську і назад. Контроллер Leap Motion 3D розпізнає характерні рухи руками, які користувач виконує над поверхнею планшета. Після аналізу додаток відображає умовне позначення жесту у вигляді знака на екрані і переводить його в текст або озвучує. Передбачено також розпізнавання голосу для спілкування з людьми, здатними говорити. Співрозмовнику буде досить вимовити слово або фразу вголос, і MotionSavvy відразу ж відобразить їх на екрані планшета у вигляді відповідного тексту.

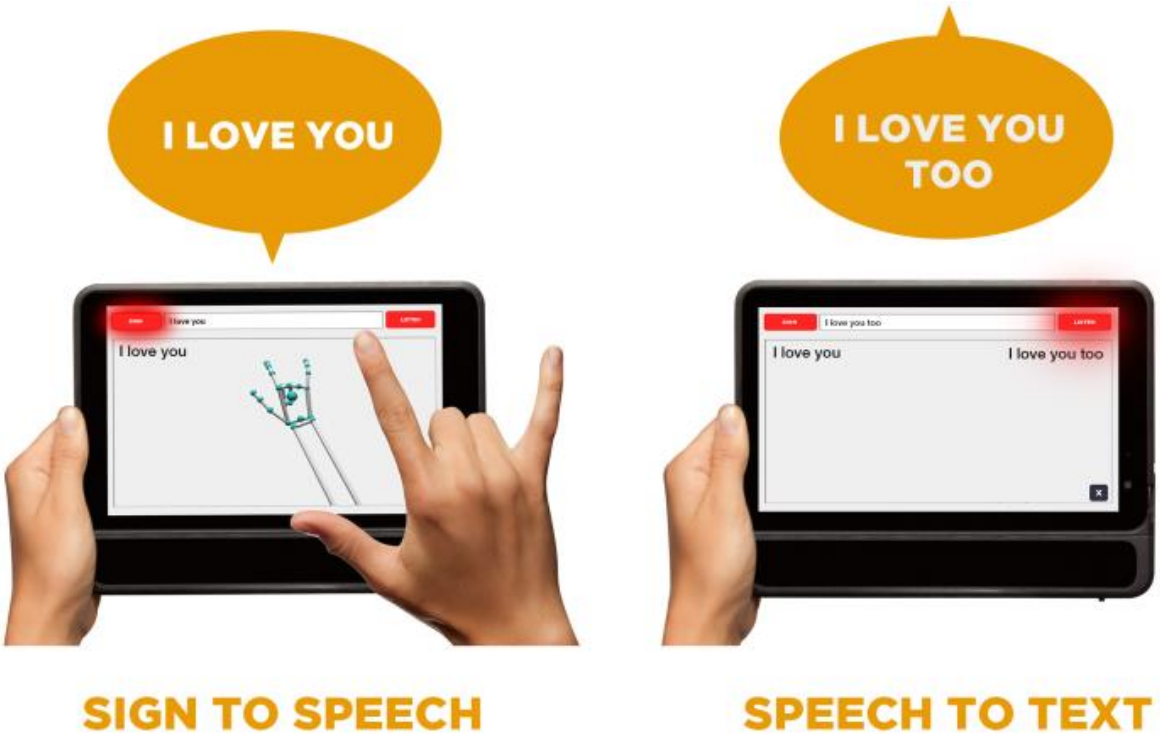

Рисунок 1.11 – Motionsavvy

1.5.3 LeapMotion

Компанія OcuSpecнашла інше цікаве рішення, яким є LeapMotion [27] (рисунок 1.12.). Дане рішення реалізується з за допомогою двох інфрачервоних камер і світлодіодів. Руки оператора повинні знаходитися в певній ділянці простору, їх рух сканується і обробляється спеціальним програмним забезпеченням, після чого дії проектуються на екран. Програма дає можливість з легкістю керувати тим, що відбувається на моніторі, запускати програми, відкривати вікна просто провівши вперед.

Також LeapMotion розпізнає різні предмети. Як на планшетах є функція - ступінь натискання, що успішно підтримується пристроєм.

Рисунок 1.12 – Приклад роботи

Компанія HP зараз впроваджує LeapMotion на деяких лінійках ноутбуків, але в такому вигляді, як вона зараз розроблена навряд чи стане гідною альтернативою клавіатурі і миші.

По-перше, вкрай важко тримати руки в підвішеному стані довгий період часу. Для людини такий стан не можна назвати природним і не скрутним.

По-друге, якість інтерпретації жестів теж не на висоті, дуже серйозна програма поступається планшетним персональним комп'ютерам.

Найголовніше, фінансова складова, за все програмне забезпечення потрібно платити, до того ж його вкрай мало, і ще не факт, що придбане програмне забезпечення вразить вас своєю бездоганною роботою і не викличе ніяких нарікань.

Тож масове тиражування даної програми в тому варіанті, в якому вона зараз існує, має вельми туманні перспективи, хоча сама ідея дуже приваблива. Зараз це лише цікаве доповнення, що має чисто розважальне навантаження.

1.5.4 MediaPipe

MediaPipe [39] - це фреймворк з відкритим вихідним кодом, представлений Google, який допомагає створювати мультимодальні конвеєри машинного навчання.

Розробник може створити прототип, не заглиблюючись в написання алгоритмів і моделей машинного навчання, використовуючи існуючі компоненти.

Ця структура може використовуватися для різних додатків для обробки зображень та мультимедіа (особливо у віртуальній реальності), таких як виявлення об'єктів, розпізнавання осіб, відстеження рук, відстеження декількох рук і сегментація волосся. MediaPipe підтримує різні апаратні та операційні платформи, такі як Android, iOS і Linux, пропонуючи API на C ++, Java, Objective-c і т.д. Це середовище також здатна використовувати ресурси GPU.

Структура складається з трьох основних компонентів:

- Основа для виведення з даних конвеєра;
- Інструменти для оцінки;
- Набір повторно використовуваних компонентів логічного висновку і обробки.

Він йде за підходами заснованими на фреймворками Graph в OpenCV, і вся обробка відбувається в контексті Graph. Graph містить колекцію вузлів, і кожен вузол реалізований у вигляді калькулятора. Graph налаштовується за допомогою буфера GraphConfig, а потім запускається з використанням графічного об'єкта. У Graph калькулятори пов'язані один з одним за допомогою потоку даних, і кожен потік містить тимчасові ряди пакетів даних. Ці калькулятори і потоки визначають потік даних на графіку.

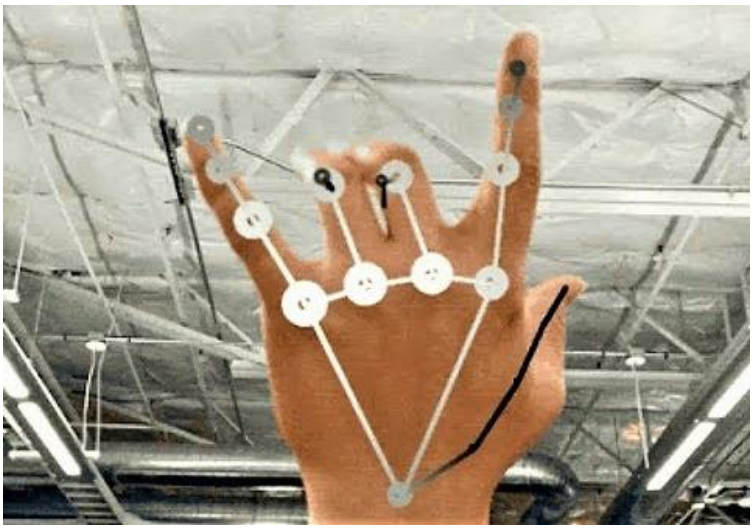

Рисунок 1.13 – Приклад роботи MediaPipe

1.5.5 Myo

Розроблено браслет [37], який позиційно розміщується на передпліччі, в результаті зчитування електричних імпульсів в м'язах рухається руки формується команда, яка в подальшому передається підключеного пристрою (рисунок 1.14).

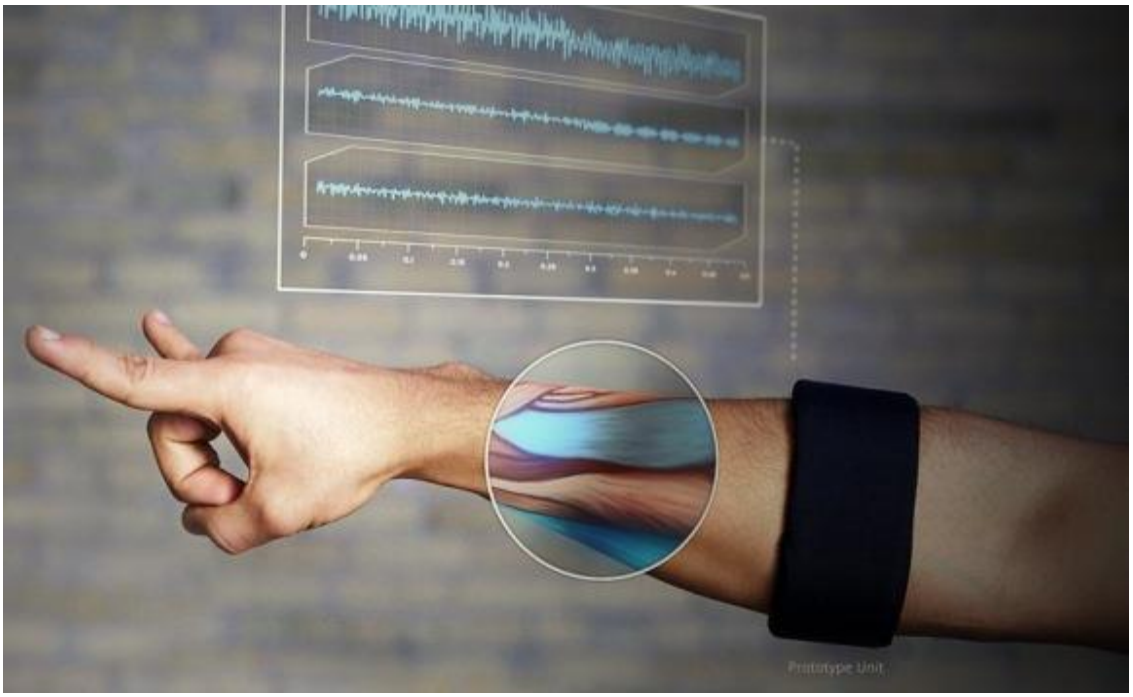

Рисунок 1.14 – Myo

Передача даних відбувається за допомогою технології Bluetooth 4.0. Дальність роботи пристрою від зчитувального пристрою може досягати 10 метрів. Це серйозно розширює можливості даної розробки.

Безпосередньо браслет складається з: ARM-процесора, Li-on акумулятора, сенсорів і інерційного вимірювального пристрою з шістьма осями, яке і позиціонує в просторі рух руки.

Браслет здатний відстежувати рухи пальців і руки. Розроблено спеціальний запрограмований жест, який не має аналогів, необхідний для включення і виключення пристрою, продумані алгоритми, що відтинають непродуктивні жести, а також передбачено зворотній зв'язок в разі успішного розпізнавання того чи іншого жесту.

Насправді цей пристрій унікальний. Він досить легко освоюєтьсяНіяких інструкцій, все на рівні інтуїції. Може працювати в операційних системах Windows і Mac OS.

Вже зараз даний пристрій може робити багато цікавого, і його потенціал величезний.

1.6 Висновки з розділу 1

В даному розділі проведено аналіз алгоритмів розпізнавання жестів та відстеження об’єктів, аналіз структури системи розпізнавання жестів з використанням веб-камери.

Виконано аналіз способів зображення об’єктів.

Проведено аналіз алгоритмів розпізнавання жестів.

Були розглянуті методи зчитування та обробки зображення:

- Засновані на зображеннях;
- Не засновані на зображеннях.

Проведено порівняння та наведено переваги та недоліки методів зчитування.

Розглянуто принципи ідентифікації жестів.

В ході аналізу алгоритмів виявлення об'єктів було встановлено, що вони прості в реалізації, але більша частина з них володіє низькою ефективністю і схильна до сильного впливу зовнішніх факторів, таких як: освітлення, фон, розмір об'єкта та інше. Для підвищення ефективності застосування даних алгоритмів, необхідно їх комбінувати, що дозволить підвищити точність локалізації об'єкта на зображенні.

Алгоритми відстежування об'єктів мають набір суттєвих переваг, однак недоліком є висока обчислювальна складність і необхідність дотримання певних умов. Дані алгоритми доцільно поєднувати з методами виявлення, для коректного відстеження об'єктів у відеопослідовності.

Дані алгоритми доцільно поєднувати з методами виявлення, для коректного відстеження об'єктів у відеопослідовності. У процесі дослідження методів розпізнавання об'єктів були встановлені бажані якісні характеристики, такі як точність розпізнавання і швидкість роботи.

Проаналізовано структуру системи розпізнавання жестів з використанням жестів.

Розглянуті сучасні рішення:

- Kinect Sign Language Translator;
- Motionsavvy;
- LeapMotion;
- MediaPipe;
- Myo.

2 ПЛАНУВАННЯ ЕКСПЕРИМЕНТУ

2.1 Алгоритм і порядок обробки

Алгоритм розпізнавання жестів представлено на рисунку 2.1.

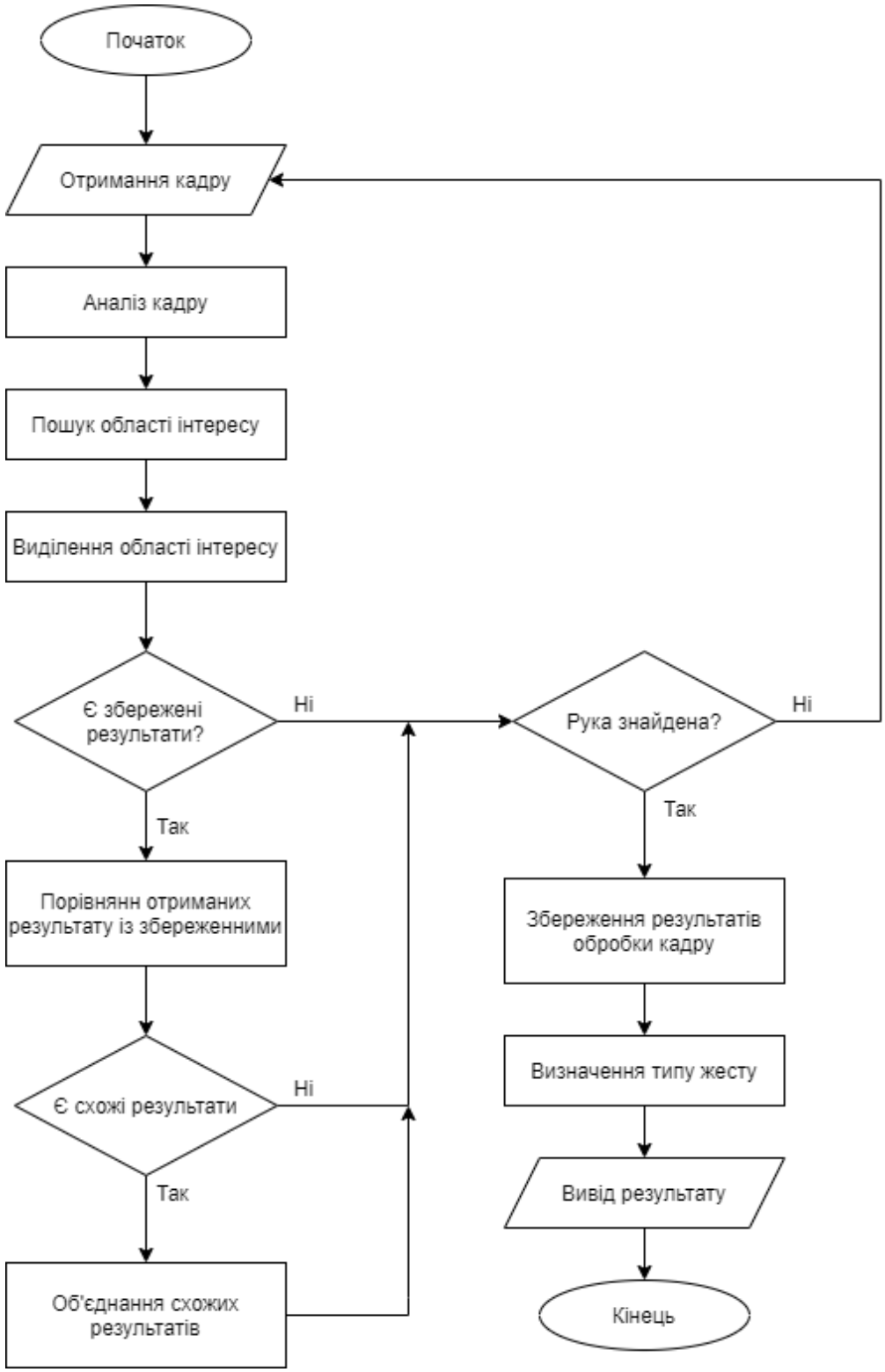

Рисунок 2.1 – Алгоритм розпізнавання жестів

Рисунок 2.2 – Лінійна послідовність процесів системи розпізнавання жестів

Вхідними даними є кадри з відеопослідовності. Після того, як на вхід ми отримали кадр – починається процес аналізу кадру та пошук області інтересу.

Область інтересу потрібна для того, щоб видалити не потрібні для обробки ділянки кадру. Далі система перевіряє наявність збережених успішних результатів розпізнавання жесту.

Якщо в базі даних системи є збережені результати – система почне порівняння поточних даних із збереженими. Завдяки цьому покращується швидкість обробки кадру та пошуку області інтересу.

Коли область інтересу на кадрі із зображенням руки було знайдено – система виконує збереження результатів та починає виконувати процес визначення жесту.

Кінцевим результатом є вивід результату на кадрі з відеопослідовності.

2.2 Вибір жестів руки для проведення експерименту

Для проведення експерименту було обрано наступні жести руки:
- Кулак, внутрішня сторона (рисунок 2.3);
- Кулак, зовнішня сторона (рисунок 2.4);
- Долонь, внутрішня сторона (рисунок 2.5);
- Долонь, зовнішня сторона (рисунок 2.6);
- Перелік пальців руки (від мізинця до великого пальця).

Рисунок 2.3 – Кулак, внутрішня сторона

Рисунок 2.4 – Кулак, зовнішня сторона

Рисунок 2.5 – Долонь, внутрішня сторона

Рисунок 2.6 – Долонь, зовнішня сторона

Також для проведення експерименту буде задіяно 2 людини різної статі (чоловік та жінка), для оцінки ефективності систем розпізнавання жестів руки людей різної статі.

2.3 Вибір типу жестів для проведення експерименту

Статичний жест - положення частин тіла (руки / рук) з певними координатами ключових точок (рука знаходиться в одному положенні та не пересувається).

Динамічний жест - послідовність статичних жестів, які змінюються в реальному часі за певний час.

Більшість жестів динамічні і пов'язані з рухами пальців, кистей або всієї руки. Менше жестів статичних - нерухомих, для розуміння яких немає необхідності в русі.

Тестування систем розпізнавання жестів буде проводитися для двох вищевказаних типів жестів. Для тестування динамічних типів жестів буде використано для одного тесту будь-які 4 статичні жести руки.

2.4 Вибір рухів руки для проведення експерименту

Для проведення експерименту було обрано наступні рухи:

- Рух квадрат (переміщення руки по траєкторії квадрата);
- Рух коло (переміщення руки по траєкторії кола);
- Рух трикутник (переміщення руки по траєкторії трикутника).

2.5 Вибір зовнішніх умов для проведення експерименту

Зовнішньою умовою було обрано наступні типи освітлення (одиницею вимірювання освітленості в Міжнародній системі одиниць служить Люкс (лк)):

- Темне освітлення (освітлення достатньо для читання) – 30-50 лк ;
- Денне освітлення (у світлій кімнаті поблизу вікна) – 1000 лк;
- Яскраве освітлення – 12000 лк;
- Без освітлення (освітлення тільки від монітору комп’ютера) – до 20 лк.

2.6 Фактори, які впливають на розпізнавання жестів

Під час проведення експерименту потрібно враховувати на різноманітні види факторів [38]:

- Фактори, на які людина має вплив (прикраси на пальцях, наявність предмету гардеробу (перчатки), манікюр, тощо):
- Фактори, на які людина не може впливати (заводські характеристики пристрою, що зчитує інформацію, травми, пов’язані із відсутністю пальцю або пальців, колір шкіри людини, цифровий шум, тощо).

2.7 Вибір пристрою для проведення експерименту

В якості пристрою для проведення експерименту був обраний персональний ноутбук HP Zbook 15v G5.

Характеристики веб-камери:

- Веб-камера HD 720p з підтримкою інфрачервоного режиму (обробка зображень в ближній інфрачервоній області спектра).
- Сертифікація Microsoft Hello.
- Формат HD (широкоекранний, 16: 9).
- Підтримка відеоконференцій і захоплення зображень.
- Високоякісний об'єктив з постійною фокусною відстанню.
- Захоплення відеосигналу з роздільною здатністю до 1280 x 720 (720p) і частотою до 30 кадрів/с.
- Стиснення M-JPEG підтримує більш високу частоту кадрів для захоплення відеосигналу і відеоконференцій
- Покращена чутливість в умовах слабкого освітлення.
- Покращений динамічний діапазон.

2.8 Висновки з розділу 2

В даному розділі було обрано жести для оцінки ефективності розпізнавання статичних жестів:

- Кулак, внутрішня сторона;
- Кулак, зовнішня сторона;
- Долонь, внутрішня сторона;
- Долонь, зовнішня сторона;
- Перелік пальців руки (від мізинця до великого пальця).

Обрана кількість людей, які будуть примати участь в експерименті.

Обрано послідовності статичних жестів для оцінки ефективності розпізнавання динамічних жестів.

Обрані зовнішні умови, а саме освітлення, для перевірки ефективності використання систем розпізнавання жестів при різноманітному освітленні:

- Темне освітлення (освітлення достатньо для читання) – 30-50 лк ;
- Денне освітлення (у світлій кімнаті поблизу вікна) – 1000 лк;
- Яскраве освітлення – 12000 лк;
- Без освітлення (освітлення тільки від монітору комп’ютера) – до 20 лк.

Розглянуті фактори, що впливають на процес розпізнавання жестів.

Обрано пристрій для проведення експерименту та надано характеристики його веб-камери.

3 ТЕСТУВАННЯ СИСТЕМ РОЗПІЗНАВАННЯ ЖЕСТІВ НА ОСНОВІ НЕЙРОМЕРЕЖЕВИХ АЛГОРИТМІВ АНАЛІЗУ ЗОБРАЖЕНЬ

3.1 Тестування систем на розпізнавання статичних жестів

3.1.1 Тестування системи MediaPipe на розпізнавання статичних жестів

Тестування системи MediaPipe на розпізнавання статичних жестів до реалізації алгоритму представлено в таблиці 3.1.

Таблиця 3.1 – Тестування системи MediaPipe на розпізнавання статичних жестів

№	Жест руки	Тип освітлення	Результат (рука чоловіка)	Результат (рука жінки)
1	Кулак, внутрішня сторона	Темне освітлення	19/20	18/20
2	Кулак, зовнішня сторона	Темне освітлення	20/20	20/20
3	Долонь, зовнішня сторона	Темне освітлення	20/20	20/20
4	Долонь, внутрішня сторона	Темне освітлення	18/20	19/20
5	Великий палець	Темне освітлення	20/20	20/20
6	Вказівний палець	Темне освітлення	17/20	15/20
7	Середній палець	Темне освітлення	19/20	20/20
8	Безіменний палець	Темне освітлення	20/20	17/20
9	Мізинець	Темне освітлення	14/20	13/20
10	Кулак, внутрішня сторона	Денне освітлення	20/20	20/20

Продовження таблиці 3.1

№	Жест руки	Тип освітлення	Результат (рука чоловіка)	Результат (рука жінки)
11	Кулак, зовнішня сторона	Денне освітлення	20/20	18/20
12	Долонь, зовнішня сторона	Денне освітлення	17/20	20/20
13	Долонь, внутрішня сторона	Денне освітлення	20/20	20/20
14	Великий палець	Денне освітлення	19/20	20/20
15	Вказівний палець	Денне освітлення	20/20	20/20
16	Середній палець	Денне освітлення	20/20	20/20
17	Безіменний палець	Денне освітлення	20/20	19/20
18	Мізинець	Денне освітлення	18/20	20/20
19	Кулак, внутрішня сторона	Яскраве освітлення	20/20	20/20
20	Кулак, зовнішня сторона	Яскраве освітлення	20/20	18/20
21	Долонь, зовнішня сторона	Яскраве освітлення	19/20	20/20
22	Долонь, внутрішня сторона	Яскраве освітлення	20/20	20/20
23	Великий палець	Яскраве освітлення	20/20	18/20
24	Вказівний палець	Яскраве освітлення	20/20	20/20
25	Середній палець	Яскраве освітлення	20/20	20/20
26	Безіменний палець	Яскраве освітлення	17/20	18/20

Продовження таблиці 3.1

№	Жест руки	Тип освітлення	Результат (рука чоловіка)	Результат (рука жінки)
27	Мізинець	Яскраве освітлення	18/20	17/20
28	Кулак, внутрішня сторона	Без освітлення	5/20	3/20
29	Кулак, зовнішня сторона	Без освітлення	7/20	6/20
30	Долонь, зовнішня сторона	Без освітлення	17/20	14/20
31	Долонь, внутрішня сторона	Без освітлення	12/20	20/20
32	Великий палець	Без освітлення	16/20	20/20
33	Вказівний палець	Без освітлення	16/20	20/20
34	Середній палець	Без освітлення	20/20	20/20
35	Безіменний палець	Без освітлення	12/20	20/20
36	Мізинець	Без освітлення	17/20	5/20

3.1.2 Результати тестування системи MediaPipe на розпізнавання статичних жестів

Було проведено тестування точності та швидкості розпізнавання статичних жестів для системи розпізнавання MediaPipe.

Для тестування використовувалися різноманітні статичні жести рук:

- Кулак, внутрішня сторона;
- Кулак, зовнішня сторона;
- Долонь, внутрішня сторона;
- Долонь, зовнішня сторона;
- Перелік пальців руки (від 1 до 5).

Під час тестування крім зазначених жестів було введено «помилковий жест» - відкрита долоня в горизонтальному положенні відносно веб-камери, в результаті чого система демонструвала помилковий результат розпізнавання або не розпізнавала жест взагалі.

Жести виконувалися по черзі один за одним, було проведено по 20 тестів для однієї людини для кожного жесту при різному типу освітлення. Загальна кількість тестів склала 1440 експериментів.

Точність роботи системи при темному освітленні:

$$A_d = \frac{\mathrm{Sum}(n_1, n_9)}{N_d} \times 100\% = \frac{329}{360} \times 100\% = 91.39\%,$$

де Sum(n1, n9) – сума всіх експериментів при темному освітленні,
Nd – загальна кількість експериментів при темному освітленні.

Точність роботи системи при денному освітленні:

$$A_s = \frac{\mathrm{Sum}(n_{10}, n_{18})}{N_s} \times 100\% = \frac{353}{360} \times 100\% = 98.06\%,$$

де Sum(n10, n18) – сума всіх експериментів при денному освітленні,
Ns – загальна кількість експериментів при денному освітленні.

Точність роботи системи при яскравому освітленні:

$$A_l = \frac{\mathrm{Sum}(n_{19}, n_{27})}{N_l} \times 100\% = \frac{345}{360} \times 100\% = 95.83\%,$$

де Sum(n19, n27) – сума всіх експериментів при яскравому освітленні,
Nl – загальна кількість експериментів при яскравому освітленні.

Точність роботи системи без освітлення:

$$A_{td} = \frac{\mathrm{Sum}(n_{19}, n_{27})}{N_{td}} \times 100\% = \frac{250}{360} \times 100\% = 69.44\%,$$

де Sum(n19, n27) – сума всіх експериментів без освітлення,
Nl – загальна кількість експериментів без освітлення.

Загальна точність роботи системи MediaPipe:

$$A = \frac{A_d + A_s + A_l + A_{td}}{T} = \frac{91.39 + 98.06 + 95.83 + 69.44}{4} = 88.68\%,$$

де T – кількість типів освітлення при проведенні експерименту.

Середня швидкість розпізнавання статичного жесту на пристрої HP Zbook 15v G5 складає приблизно 0.42 секунди., кількість кадрів в секунду складає до 25.

3.1.3 Тестування системи Handpose на розпізнавання статичних жестів

Тестування системи Handpose на розпізнавання статичних жестів представлено до реалізації алгоритму в таблиці 3.2.

Таблиця 3.2 – Тестування системи Handpose на розпізнавання статичних жестів

№	Жест руки	Тип освітлення	Результат (рука чоловіка)	Результат (рука жінки)
1	Кулак, внутрішня сторона	Темне освітлення	18/20	17/20
2	Кулак, зовнішня сторона	Темне освітлення	17/20	19/20
3	Долонь, зовнішня сторона	Темне освітлення	17/20	17/20
4	Долонь, внутрішня сторона	Темне освітлення	16/20	15/20
5	Великий палець	Темне освітлення	19/20	20/20
6	Вказівний палець	Темне освітлення	14/20	18/20
7	Середній палець	Темне освітлення	17/20	16/20
8	Безіменний палець	Темне освітлення	15/20	15/20
9	Мізинець	Темне освітлення	13/20	11/20
10	Кулак, внутрішня сторона	Денне освітлення	16/20	19/20
11	Кулак, зовнішня сторона	Денне освітлення	15/20	18/20
12	Долонь, зовнішня сторона	Денне освітлення	17/20	16/20
13	Долонь, внутрішня сторона	Денне освітлення	19/20	17/20
14	Великий палець	Денне освітлення	17/20	18/20

Продовження таблиці 3.2

№	Жест руки	Тип освітлення	Результат (рука чоловіка)	Результат (рука жінки)
15	Вказівний палець	Денне освітлення	16/20	17/20
16	Середній палець	Денне освітлення	14/20	16/20
17	Безіменний палець	Денне освітлення	17/20	16/20
18	Мізинець	Денне освітлення	14/20	16/20
19	Кулак, внутрішня сторона	Яскраве освітлення	20/20	17/20
20	Кулак, зовнішня сторона	Яскраве освітлення	14/20	15/20
21	Долонь, зовнішня сторона	Яскраве освітлення	17/20	18/20
22	Долонь, внутрішня сторона	Яскраве освітлення	18/20	17/20
23	Великий палець	Яскраве освітлення	15/20	20/20
24	Вказівний палець	Яскраве освітлення	15/20	17/20
25	Середній палець	Яскраве освітлення	20/20	20/20
26	Безіменний палець	Яскраве освітлення	20/20	18/20
27	Мізинець	Яскраве освітлення	14/20	13/20
28	Кулак, внутрішня сторона	Без освітлення	7/20	5/20
29	Кулак, зовнішня сторона	Без освітлення	5/20	8/20
30	Долонь, зовнішня сторона	Без освітлення	10/20	12/20
31	Долонь, внутрішня сторона	Без освітлення	8/20	9/20

Продовження таблиці 3.2

№	Жест руки	Тип освітлення	Результат (рука чоловіка)	Результат (рука жінки)
32	Великий палець	Без освітлення	8/20	12/20
33	Вказівний палець	Без освітлення	15/20	14/20
34	Середній палець	Без освітлення	11/20	13/20
35	Безіменний палець	Без освітлення	9/20	15/20
36	Мізинець	Без освітлення	4/20	5/20

3.1.4 Результати тестування системи Handpose на розпізнавання статичних жестів

Було проведено тестування точності та швидкості розпізнавання статичних жестів для системи розпізнавання Handpose.

Для тестування використовувалися різноманітні статичні жести рук:

- Кулак, внутрішня сторона;
- Кулак, зовнішня сторона;
- Долонь, внутрішня сторона;
- Долонь, зовнішня сторона;
- Перелік пальців руки (від 1 до 5).

Під час тестування крім зазначених жестів було введено «помилковий жест» - відкрита долоня в горизонтальному положенні відносно веб-камери, в результаті чого система демонструвала помилковий результат розпізнавання або не розпізнавала жест взагалі.

Жести виконувалися по черзі один за одним, було проведено по 20 тестів для кожного жесту при різному типу освітлення. Загальна кількість тестів склала 1440 експериментів.

Точність роботи системи при темному освітленні:

$$A_d = \frac{\mathrm{Sum}(n_1, n_9)}{N_d} \times 100\% = \frac{294}{360} \times 100\% = 81.67\%,$$

де Sum(n1, n9) – сума всіх експериментів при темному освітленні,
Nd – загальна кількість експериментів при темному освітленні.

Точність роботи системи при денному освітленні:

$$A_s = \frac{\mathrm{Sum}(n_{10}, n_{18})}{N_s} \times 100\% = \frac{327}{360} \times 100\% = 90.83\%,$$

де Sum(n10, n18) – сума всіх експериментів при денному освітленні,
Ns – загальна кількість експериментів при денному освітленні.

Точність роботи системи при яскравому освітленні:

$$A_l = \frac{\mathrm{Sum}(n_{19}, n_{27})}{N_l} \times 100\% = \frac{305}{360} \times 100\% = 85.55\%,$$

де Sum(n19, n27) – сума всіх експериментів при яскравому освітленні,
Nl – загальна кількість експериментів при яскравому освітленні.

Точність роботи системи без освітлення:

$$A_{td} = \frac{\mathrm{Sum}(n_{19}, n_{27})}{N_{td}} \times 100\% = \frac{170}{360} \times 100\% = 50.27\%,$$

де Sum(n19, n27) – сума всіх експериментів без освітлення,
Nl – загальна кількість експериментів без освітлення.

Загальна точність роботи системи Handpose:

$$A = \frac{A_d + A_s + A_l + A_{td}}{T} = \frac{81.67 + 90.83 + 85.55 + 50.27}{4} = 77.08\%,$$

де T –кількість типів освітлення при проведенні експерименту.

Середня швидкість розпізнавання жесту на пристрої HP Zbook 15v G5 складає приблизно 4 секунди., кількість кадрів в секунду до 17.

3.2 Тестування систем на розпізнавання динамічних жестів

3.2.1 Тестування системи MediaPipe на розпізнавання динамічних жестів

Тестування системи MediaPipe на розпізнавання динамічних жестів представлено в таблиці 3.3.

Таблиця 3.3 – Тестування системи MediaPipe на розпізнавання динамічних жестів

№	Жести руки	Тип освітлення	Результат (рука чоловіка)	Результат (рука жінки)
1	Кулак, внутрішня сторона Долонь, зовнішня сторона Вказівний палець Безіменний палець	Темне освітлення	14/15	14/15
2	Долонь, зовнішня сторона Середній палець Долонь, внутрішня сторона Вказівний палець	Темне освітлення	14/15	13/15
3	Безіменний палець Кулак, внутрішня сторона Вказівний палець Мізинець	Темне освітлення	13/15	13/15
4	Безіменний палець Вказівний палець Великий палець Мізинець	Темне освітлення	14/15	12/15

Продовження таблиці 3.3

№	Жести руки	Тип освітлення	Результат (рука чоловіка)	Результат (рука жінки)
5	Кулак, внутрішня сторона	Денне освітлення	15/15	15/15
	Великий палець			
	Середній палець			
	Долонь, внутрішня сторона			
6	Мізинець	Денне освітлення	15/15	14/15
	Середній палець			
	Долонь, зовнішня сторона			
	Кулак, зовнішня сторона			
7	Великий палець	Денне освітлення	14/15	15/15
	Кулак, зовнішня сторона			
	Долонь, зовнішня сторона			
	Кулак, внутрішня сторона			
8	Мізинець	Денне освітлення	14/15	13/15
	Долонь, внутрішня сторона			
	Кулак, зовнішня сторона			
	Великий палець			
9	Кулак, внутрішня сторона	Яскраве освітлення	14/15	13/15
	Кулак, зовнішня сторона			
	Долонь, зовнішня сторона			
	Долонь, внутрішня сторона			

Продовження таблиці 3.3

№	Жести руки	Тип освітлення	Результат (рука чоловіка)	Результат (рука жінки)
10	Кулак, внутрішня сторона	Яскраве освітлення	14/15	12/15
	Кулак, зовнішня сторона			
	Вказівний палець			
	Долонь, внутрішня сторона			
11	Великий палець	Яскраве освітлення	13/15	14/15
	Вказівний палець			
	Середній палець			
	Безіменний палець			
12	Мізинець	Яскраве освітлення	13/15	13/15
	Долонь, внутрішня сторона			
	Великий палець			
	Кулак, внутрішня сторона			
13	Кулак, зовнішня сторона	Без освітлення	10/15	8/15
	Вказівний палець			
	Долонь, внутрішня сторона			
	Великий палець			
14	Вказівний палець	Без освітлення	9/15	8/15
	Середній палець			
	Безіменний палець			
	Мізинець			

Продовження таблиці 3.3

№	Жести руки	Тип освітлення	Результат (рука чоловіка)	Результат (рука жінки)
15	Вказівний палець	Без освітлення	11/15	9/15
	Кулак, зовнішня сторона			
	Мізинець			
	Кулак, внутрішня сторона			
16	Мізинець	Без освітлення	13/15	11/15
	Долонь, зовнішня сторона			
	Долонь, внутрішня сторона			
	Великий палець			

3.2.2 Результати тестування системи MediaPipe на розпізнавання динамічних статичних жестів

Було проведено тестування точності та швидкості розпізнавання динамічних жестів для системи розпізнавання MediaPipe.

Для тестування використовувалися різноманітні комбінації із статичних жестів.

Під час тестування крім зазначених у таблиці жестів також було введено «помилковий жест» - відкрита долоня в горизонтальному положенні відносно веб-камери, в результаті чого система демонструвала помилковий результат розпізнавання або не розпізнавала жест взагалі.

Жести виконувалися по черзі один за одним, було проведено по 120 тестів з комбінацій статичних жестів для кожного типу освітлення. Загальна кількість тестів склала 480 експериментів.

Точність роботи системи при темному освітленні:

$$A_d = \frac{\mathrm{Sum}(n_1, n_4)}{N_d} \times 100\% = \frac{107}{120} \times 100\% = 89.17\%,$$

де Nd – загальна кількість експериментів при темному освітленні.

Точність роботи системи при денному освітленні:

$$A_s = \frac{\mathrm{Sum}(n_5, n_8)}{N_s} \times 100\% = \frac{115}{120} \times 100\% = 95.83\%,$$

де Sum(n5, n8) – сума всіх експериментів при денному освітленні,
Ns – загальна кількість експериментів при денному освітленні.

Точність роботи системи при яскравому освітленні:

$$A_l = \frac{\mathrm{Sum}(n_9, n_{12})}{N_l} \times 100\% = \frac{108}{120} \times 100\% = 90\%,$$

де Sum(n9, n12) – сума всіх експериментів при яскравому освітленні,
Nl – загальна кількість експериментів при яскравому освітленні.

Точність роботи системи без освітлення:

$$A_{td} = \frac{\mathrm{Sum}(n_{13}, n_{16})}{N_{td}} \times 100\% = \frac{79}{120} \times 100\% = 65.83\%,$$

де Sum(n13, n16) – сума всіх експериментів без освітлення,
Nl – загальна кількість експериментів без освітлення.

Загальна точність роботи системи MediaPipe:

$$A = \frac{A_d + A_s + A_l + A_{td}}{T} = \frac{89.17 + 95.83 + 90 + 65.83}{4} = 85.21\%,$$

де T – кількість типів освітлення при проведенні експерименту.

3.2.3 Тестування системи Handpose на розпізнавання динамічних жестів

Тестування системи Handpose на розпізнавання динамічних жестів представлено в таблиці 3.4.

Таблиця 3.4 – Тестування системи Handpose на розпізнавання динамічних жестів

№	Жести руки	Тип освітлення	Результат (рука чоловіка)	Результат (рука жінки)
1	Кулак, внутрішня сторона	Темне освітлення	8/15	7/15
	Долонь, зовнішня сторона			
	Вказівний палець			
	Безіменний палець			
2	Долонь, зовнішня сторона	Темне освітлення	10/15	9/15
	Середній палець			
	Долонь, внутрішня сторона			
	Вказівний палець			
3	Безіменний палець	Темне освітлення	8/15	8/15
	Кулак, внутрішня сторона			
	Вказівний палець			
	Мізинець			
4	Безіменний палець	Темне освітлення	7/15	6/15
	Вказівний палець			
	Великий палець			
	Мізинець			

Продовження таблиці 3.4

№	Жести руки	Тип освітлення	Результат (рука чоловіка)	Результат (рука жінки)
5	Кулак, внутрішня сторона	Денне освітлення	11/15	11/15
	Великий палець			
	Середній палець			
	Долонь, внутрішня сторона			
6	Мізинець	Денне освітлення	12/15	12/15
	Середній палець			
	Долонь, зовнішня сторона			
	Кулак, зовнішня сторона			
7	Великий палець	Денне освітлення	12/15	11/15
	Кулак, зовнішня сторона			
	Долонь, зовнішня сторона			
	Кулак, внутрішня сторона			
8	Мізинець	Денне освітлення	12/15	12/15
	Долонь, внутрішня сторона			
	Кулак, зовнішня сторона			
	Великий палець			
9	Кулак, внутрішня сторона	Яскраве освітлення	10/15	9/15
	Кулак, зовнішня сторона			
	Долонь, зовнішня сторона			
	Долонь, внутрішня сторона			

Продовження таблиці 3.4

№	Жести руки	Тип освітлення	Результат (рука чоловіка)	Результат (рука жінки)
10	Кулак, внутрішня сторона	Яскраве освітлення	8/15	10/15
	Кулак, зовнішня сторона			
	Вказівний палець			
	Долонь, внутрішня сторона			
11	Великий палець	Яскраве освітлення	7/15	9/15
	Вказівний палець			
	Середній палець			
	Безіменний палець			
12	Мізинець	Яскраве освітлення	8/15	8/15
	Долонь, внутрішня сторона			
	Великий палець			
	Кулак, внутрішня сторона			
13	Кулак, зовнішня сторона	Без освітлення	6/15	5/15
	Вказівний палець			
	Долонь, внутрішня сторона			
	Великий палець			
14	Вказівний палець	Без освітлення	7/15	8/15
	Середній палець			
	Безіменний палець			
	Мізинець			

Продовження таблиці 3.4

№	Жести руки	Тип освітлення	Результат (рука чоловіка)	Результат (рука жінки)
15	Вказівний палець	Без освітлення	5/15	3/15
	Кулак, зовнішня сторона			
	Мізинець			
	Кулак, внутрішня сторона			
16	Мізинець	Без освітлення	5/15	6/15
	Долонь, зовнішня сторона			
	Долонь, внутрішня сторона			
	Великий палець			

3.2.4 Результати тестування системи Handpose на розпізнавання динамічних жестів

Було проведено тестування точності та швидкості розпізнавання динамічних жестів для системи розпізнавання Handpose.

Для тестування використовувалися різноманітні комбінації із статичних жестів.

Під час тестування крім зазначених у таблиці жестів також було введено «помилковий жест» - відкрита долоня в горизонтальному положенні відносно веб-камери, в результаті чого система демонструвала помилковий результат розпізнавання або не розпізнавала жест взагалі.

Жести виконувалися по черзі один за одним, було проведено по 120 тестів з комбінацій статичних жестів для кожного типу освітлення. Загальна кількість тестів склала 480 експериментів.

Точність роботи системи при темному освітленні:

$$A_d = \frac{\mathrm{Sum}(n_1, n_4)}{N_d} \times 100\% = \frac{63}{120} \times 100\% = 52.5\%,$$

де Sum(n1, n4) – сума всіх експериментів при темному освітленні,
Nd – загальна кількість експериментів при темному освітленні.

Точність роботи системи при денному освітленні:

$$A_s = \frac{\mathrm{Sum}(n_5, n_8)}{N_s} \times 100\% = \frac{93}{120} \times 100\% = 77.5\%,$$

де Sum(n5, n8) – сума всіх експериментів при денному освітленні,
Ns – загальна кількість експериментів при денному освітленні.

Точність роботи системи при яскравому освітленні:

$$A_l = \frac{\mathrm{Sum}(n_9, n_{12})}{N_l} \times 100\% = \frac{69}{120} \times 100\% = 57.5\%,$$

де Sum(n19, n27) – сума всіх експериментів при яскравому освітленні,
Nl – загальна кількість експериментів при яскравому освітленні.

Точність роботи системи без освітлення:

$$A_{td} = \frac{\mathrm{Sum}(n_{13}, n_{16})}{N_{td}} \times 100\% = \frac{41}{120} \times 100\% = 34.17\%,$$

де Sum(n19, n27) – сума всіх експериментів без освітлення,
Nl – загальна кількість експериментів без освітлення.

Загальна точність роботи системи MediaPipe:

$$A = \frac{A_d + A_s + A_l + A_{td}}{T} = \frac{52.5 + 77.5 + 57.5 + 34.17}{4} = 55.42\%,$$

де T – кількість типів освітлення при проведенні експерименту.

3.3 Оцінка ефективності систем розпізнавання жестів

3.3.1 Оцінка ефективності на розпізнавання статичних жестів

Для оцінки ефективності розпізнавання статичних жестів було створено поріг прохідності, якщо результат роботи системи складає більше 85%, то система є ефективною для використання її у заданих умовах, у інших випадках – система не є ефективною.

Порівняння систем розпізнавання жестів з використанням веб-камери на предмет розпізнавання статичних жестів представлено в таблиці 3.5.

Таблиця 3.5 – Оцінка ефективності розпізнавання статичних жестів

№	Тип освітлення	Точність системи MediaPipe, %	Точність системи Handpose, %
1	Темне освітлення	91.39	81.67
2	Денне освітлення	98.06	90.83
3	Яскраве освітлення	95.83	85.55
4	Без освітлення	69.44	50.27
	Загальний результат	88.68	77.08

Результати тестування систем розпізнавання MediaPipe та Handpose на предмет розпізнавання статичних жестів показують, що обидві системи втрачають свою точність при умовах без освітлення.

Ефективність системи розпізнавання MediaPipe достатня для використання її в умовах темного, денного та яскравого освітлення. В умовах без освітлення система не може бути ефективною тому що допускає достатню кількість помилок.

Ефективність системи розпізнавання Handpose достатня для використання її в умовах денного та яскравого освітлення. В умовах темного освітлення та без освітлення дану систему не бажано використовувати через її неточність.

3.3.2 Оцінка ефективності розпізнавання динамічних жестів

Для оцінки ефективності розпізнавання динамічних жестів було створено поріг прохідності, якщо результат роботи системи складає більше 85%, то система є ефективною для використання її у заданих умовах, у інших випадках – система не є ефективною.

Оцінка ефективності розпізнавання динамічних жестів представлено в таблиці 3.6.

Таблиця 3.6 – Оцінка ефективності розпізнавання динамічних жестів

№	Тип освітлення	Точність системи MediaPipe, %	Точність системи Handpose, %
1	Темне освітлення	89.17	52.5
2	Денне освітлення	95.83	77.5
3	Яскраве освітлення	90	57.5
4	Без освітлення	65.83	34.17
5	Загальний результат	85.21	55.42

Результати тестування систем розпізнавання MediaPipe та Handpose на предмет розпізнавання динамічних жестів показують, що в порівнянні із статичними, динамічні жести розпізнаються з меншою точністю.

Використовувати систему розпізнавання MediaPipe ефективно в умовах темного, денного та яскравого освітлення. Ефективність системи в умовах без освітлення не дозволяє використовувати її при таких умовах.

Використовувати систему розпізнавання Handpose для ідентифікації динамічного жесту не ефективно при всіх умовах освітлення. Швидкість розпізнавання мала, та не дозволяє ефективно розпізнавати ряд статичних жестів.

3.4 Висновки з розділу 3

В даному розділі було проведено тестування двох систем розпізнавання жестів: MediaPipe та Handpose.

Тестування систем виконувалося окремо при однакових умовах.

Виконано тестування точності та швидкості систем розпізнавання жестів на предмет розпізнавання статичних та динамічних типів жестів.

Сформовані таблиці з результатами розпізнавання статичних та динамічних жестів.

Для проведення експерименту було задіяно 2 людини (чоловік та жінка) для оцінки ефективності систем розпізнавання жестів.

Всього було проведено 3840 тестів, з них 2880 тестів – тестування систем розпізнавання жестів на предмет розпізнавання статичних жестів та 960 тестів – тестування систем розпізнавання жестів на предмет розпізнавання динамічних жестів.

Розраховано точність розпізнавання статичних та динамічних жестів системи MediaPipe в умовах різного освітлення.

Розраховано точність розпізнавання статичних та динамічних жестів системи Handpose в умовах різного освітлення.

Виконано порівняння систем розпізнавання жестів.

Виконана оцінка ефективності систем для кожного типу освітлення та жестів окремо.

Загальна точність роботи системи MediaPipe складає 86.94%.

Загальна точність роботи системи Handpose складає 66.25%.

ВИСНОВКИ

Поставлені задачі дипломного проекту були виконані.

Проведено аналіз алгоритмів розпізнавання жестів та відстеження об’єктів, аналіз структури системи розпізнавання жестів з використанням веб-камери та аналіз способів зображення об’єктів.

Проведено аналіз алгоритмів розпізнавання жестів.

Були розглянуті методи зчитування та обробки зображення:

- Засновані на зображеннях;
- Не засновані на зображеннях.

Виконано порівняння та наведено переваги та недоліки методів зчитування.

Розглянуто принципи ідентифікації жестів.

В ході аналізу алгоритмів виявлення об'єктів було встановлено, що вони прості в реалізації, але більша частина з них володіє низькою ефективністю і схильна до сильного впливу зовнішніх факторів, таких як: освітлення, фон, розмір об'єкта та інше. Для підвищення ефективності застосування даних алгоритмів, необхідно їх комбінувати, що дозволить підвищити точність локалізації об'єкта на зображенні.

Алгоритми відстежування об'єктів мають набір суттєвих переваг, однак недоліком є висока обчислювальна складність і необхідність дотримання певних умов. Дані алгоритми доцільно поєднувати з методами виявлення, для коректного відстеження об'єктів у відеопослідовності.

Дані алгоритми доцільно поєднувати з методами виявлення, для коректного відстеження об'єктів у відеопослідовності. У процесі дослідження методів розпізнавання об'єктів були встановлені бажані якісні характеристики, такі як точність розпізнавання і швидкість роботи.

Проаналізовано структуру системи розпізнавання жестів з використанням жестів.

Розглянуті сучасні рішення:

- Kinect Sign Language Translator;
- Motionsavvy;
- LeapMotion;
- MediaPipe;
- Myo.

Проведено планування експерименту. Обрано жести для оцінки ефективності розпізнавання статичних жестів:

- Кулак, внутрішня сторона;
- Кулак, зовнішня сторона;
- Долонь, внутрішня сторона;
- Долонь, зовнішня сторона;
- Перелік пальців руки (від мізинця до великого пальця).

Обрана кількість людей, які будуть примати участь в експерименті.

Обрано типи жестів: статичні та динамічні.

Обрані зовнішні умови, а саме освітлення, для перевірки ефективності використання систем розпізнавання жестів при різноманітному освітленні: темне, денне, яскраве та в умовах без освітлення.

Розглянуті фактори, що впливають на процес розпізнавання жестів.

Обрано пристрій для проведення експерименту та надано характеристики його веб-камери.

Проведено тестування двох систем розпізнавання жестів: MediaPipe та Handpose. Тестування систем виконувалося окремо при однакових умовах.

Виконано тестування точності та швидкості систем розпізнавання жестів на предмет розпізнавання статичних та динамічних типів жестів.

Сформовані таблиці з результатами розпізнавання статичних та динамічних жестів.

Для проведення експерименту було задіяно 2 людини (чоловік та жінка).

Всього було проведено 3840 тестів, з них 2880 тестів – тестування систем розпізнавання жестів на предмет розпізнавання статичних жестів та 960 тестів – тестування систем розпізнавання жестів на предмет розпізнавання динамічних жестів.

Розраховано точність розпізнавання статичних та динамічних жестів систем MediaPipe та Handpose в умовах різного освітлення.

Виконано порівняння систем розпізнавання жестів.

Виконана оцінка ефективності систем розпізнавання для кожного типу освітлення та жестів окремо.

Загальна точність роботи системи MediaPipe складає 86.94%.

Загальна точність роботи системи Handpose складає 66.25%.

ПЕРЕЛІК ПОСИЛАНЬ

1. Абакумов В.Г. Интерпретация движений рук расширяет возможности интерактивного управления в интеллектуальных системах / В.Г. Абакумов, Е.Ю. Ломакина // Природные и интеллектуальные ресурсы Сибири. – 2009. – С.199-202
2. Lomakina O.Y. Gestures Recognition as a New Information Input Device for Automatic System Control / O.Y. Lomakina // «Modern Problems of Radio Engineering, Telecommunications, and Computer Science» : Proceedings of X-th International Conference TCSET'2010. – Lviv-Slavske, 2010. – P. 100-103.
3. Hand tracking and gesture recognition for human-computer interaction / [C. Manresa et al.] // Electronic Letters on Computer Vision and Image Analysis. – 2005. – № 5(3). – P. 96-104.
4. Абакумов В.Г. Застосування жестів рук при людино-машинному інтерфейсі / В.Г. Абакумов, О.Ю. Ломакіна, О.Б. Яровенко // Электроника и связь. – Тематический выпуск : Электроника и нанотехнологии. – 2011.
5. Aggarwal J.K. Human Motion Analysis: A Review / Aggarwal J.K. and Q Cai. // Computer Vision and Image Understanding 73. –1999. – № 3. – P. 428-440.
6. Stenger B. (2001) Model-based 3D tracking of an articulated hand / B. Stenger, P.R.S. Mendonca, R. Cipolla // In: The 20th IEEE International Conference on Computer Vision and Pattern Recognition (CVPR'01). – December 2001. – Kauai, HI, US.
7. Kato M. Articulated Hand Tracking by PCA-ICA approach / M. Kato, Y.W. Chen and G. Xu // Proceedings of the IEEE International Conference on Automatic Face and Gesture Recognition. – Southampton, 2006. – P. 329-333.
8. Sign Language Translator [Електронний ресурс]. – Режим доступу до ресурсу: http://research.microsoft.com/enus/collaboration/stories/kinect-sign-lang_screen.jpg.
9. Нахождение контуров и операции с ними [Электронный [Електронний ресурс] – Режим доступу до ресурсу: http://robocraft.ru/blog/computervision/640.html.
10. Vision-based Hand Gesture Recognition for HumanComputer Interaction [Електронний ресурс] – Режим доступу до ресурсу: http://www.ics.forth.gr/.
11. Использование машинного зрения в целях автоматизации и контроля в технических системах [Електронний ресурс] / А. И. Кузнецов, А. В. Сурков, М. Н. Рудь – Режим доступу до ресурсу: http://earchive.tpu.ru/handle/11683/16932.
12. Background subtraction techniques: a review [Електронний ресурс] – Режим доступу до ресурсу: http://profs.sci.univr.it/~cristanm/%20teaching/sar_files/lezione4/Piccardi.pdf.
13. Towards Ubiquitous Document Scanning [Електронний ресурс] – Режим доступу до ресурсу: https://robau.files.wordpress.com/2010/%2006/final_report_00012.pdf.

14. Анализ изображений и видео Лекция 8: Сегментация изображений [Електронний ресурс] – Режим доступу до ресурсу: https://compscicenter.ru/media/slides/.
15. MotionSavvy [Електронний ресурс]. – Режим доступу до ресурсу: https://images.indiegogo.com/file_attachments/931256/files/20141014180025 - features.png?1413334825.
16. Приклад фільтрації за порогом. [Електронний ресурс]. – Режим доступу до ресурсу:https://habrastorage.org/getpro/habr/post_images/b22/e12/c6d/b22e12 c6d33e b898e34b373475be35d8.jpg.
17. Continuous Wavelet Transform [Електронний ресурс]. – Режим доступу до ресурсу: http://www.satmagazine.com/cgi-bin/display_image.cgi?1114554895. Дата доступа: 10.04.2016.
18. Matlab: Корреляція зображень. [Електронний ресурс]. – Режим доступу до ресурсу:http://www.mathworks.com/help/images/ref/imshowpair_ex1_diff.png
19. Boundary Detection with Sketch Tokens [Електронний ресурс]. – Режим доступу до ресурсу: http://cs.brown.edu/courses/cs143/proj5/teaser.png.
20. Farneback optical flow [Електронний ресурс]. – Режим доступу до ресурсу: https://1.bp.blogspot.com/-SYVmZsXoYuA/Vs6nFguzBI/AAAAAAAABr4 /T1Iwd8Bq3o/s1600/opticalflow.png.
21. T.J. Keating "An Improved Method of Digital Image Correlation" Photogrammetric Engineering and Remote Sensing \ T.J. Keating, P.R. Wolf, F.L. Scarpace, 1975 . – С. 993-1002.
22. SIFT: Офіційний сайт – Режим доступу до ресурсу: http://sift.jcvi.org
23. SIFT: Опис ключової точки[Електронний ресурс]. – Режим доступу до ресурсу: http://www.vlfeat.org/api/sift-frame.png.
24. Байгарова Н.С. Методы индексирования и поиска изображений[Електронний ресурс]. – Режим доступу до ресурсу:: http://web.ihep.su/library/pubs/aconf00/dconf00/ps/030.pdf.\Байгарова Н.С., Бухштаб Ю.А, Горный А.А.
25. Башков Е.А. Поиск изображений в больших БД с использованием корреляции цветовых гистограм.\ Башков Е.А., Шозда Н.С. [Електронний ресурс]. – Режим доступу до ресурсу:: http://graphicon.ru/html/2002/pdf/Bashkov_Zhozda _Re.pdf.
26. Броневич А.Г. Анализ неопределенности выделения информативных признаков и представление изображений \ АГ Броневич, 2013 . – С. 65-69.
27. Обзор Leap Motion [Електронний ресурс]. – Режим доступу до ресурсу: https://habr.com/ru/company/rozetked/blog/190404/.
28. MOTIONSAVVY: ПЛАНШЕТ ДЛЯ ГЛУХОНЕМЫХ С РАСПОЗНАВАНИЕМ ЖЕСТОВ [Електронний ресурс] – Режим доступу до ресурсу: https://www.computerra.ru/182184/motionsavvy-tablet-for-the-deaf-with-gesture-recognition/.
29. Kinect Sign Language Translator [Електронний ресурс] – Режим доступу до ресурсу: http://apps4all.ru/post/12-16-13-uchenye-microsoft-sozdali-sistemu-perevoda-yazyka-zhestov-na-osnove-kinect.

30. Використання мови жестів у спілкуванні [Електронний ресурс] – Режим доступу до ресурсу: http://ru.osvita.ua/vnz/reports/psychology/9980/.
31. Жестова мова [Електронний ресурс] – Режим доступу до ресурсу: https://uk.wikipedia.org/wiki/%D0%96%D0%B5%D1%81%D1%82%D0%BE%D0%B2%D0%B0_%D0%BC%D0%BE%D0%B2%D0%B0.
32. Спілкування як особливий вид діяльності. [Електронний ресурс] – Режим доступу до ресурсу: http://ru.osvita.ua/vnz/reports/management/14868/.
33. Право на спілкування [Електронний ресурс] – Режим доступу до ресурсу: http://osvita.khpg.org/index.php?id=1496825966.
34. Распознавание жестов [Електронний ресурс] – Режим доступу до ресурсу: http://integral-russia.ru/2020/03/13/raspoznavanie-zhestov-dlya-vzaimodejstviya-s-ii-ot-teorii-k-poslednim-dostizheniyam/.
35. Нейронные сети: распознавание образов и изображений с помощью ИИ [Електронний ресурс] – Режим доступу до ресурсу: https://center2m.ru/ai-recognition.
36. WiTrack [Електронний ресурс] – Режим доступу до ресурсу: https://3dnews.ru/786886.
37. Myo [Електронний ресурс] – Режим доступу до ресурсу: https://myo.de/fragen-antworten/.
38. 8 факторов, которые влияют на качество изображения [Електронний ресурс] – Режим доступу до ресурсу: https://blog.depositphotos.com/ru/kak-uluchshit-kachestvo-foto.html.
39. MediaPipe [Електронний ресурс] – Режим доступу до ресурсу: https://ai.googleblog.com/2019/08/on-device-real-time-hand-tracking-with.html.
40. Чуйков А. В. СИСТЕМА РАСПОЗНАВАНИЯ ЖЕСТОВ НА ОСНОВЕ НЕЙРОСЕТЕВЫХ ТЕХНОЛОГИЙ / А. В. Чуйков, А. М. Вульфин. // ФГБОУ ВО «Уфимский государственный авиационный технический университет» (УГАТУ). – 2017. – С. 10.

ДОДАТОК А

Приклади тестування системи MediaPipe

Рисунок А.1 – Приклад тестування системи в умовах денного освітлення

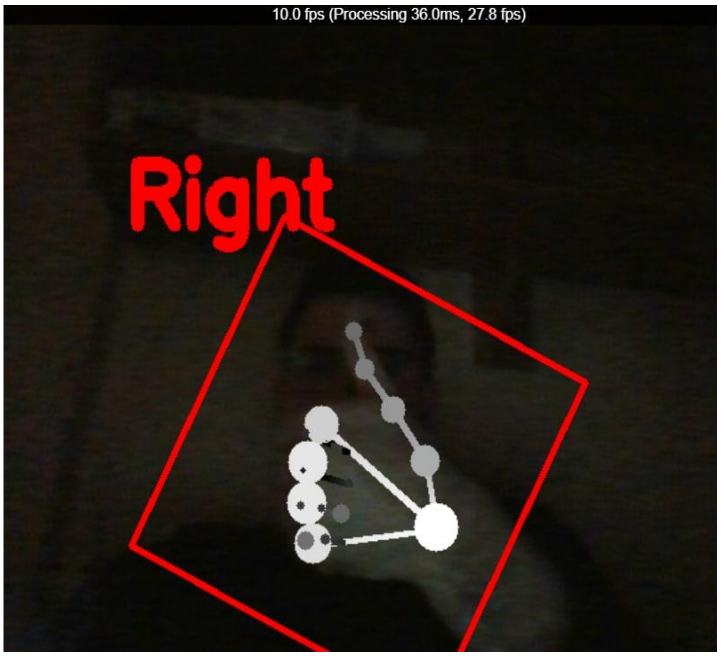

Рисунок А.2 – Приклад тестування системи в умовах без освітлення

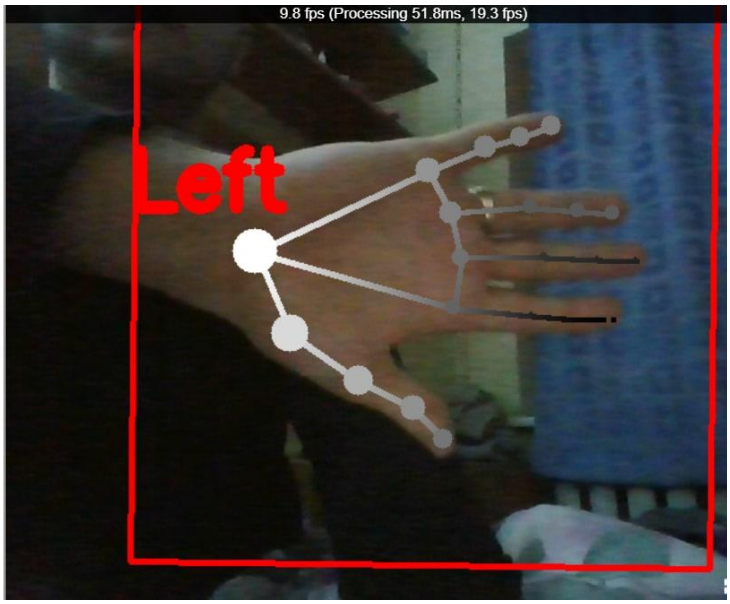

Рисунок А.3 – Приклад тестування системи в умовах темного освітлення

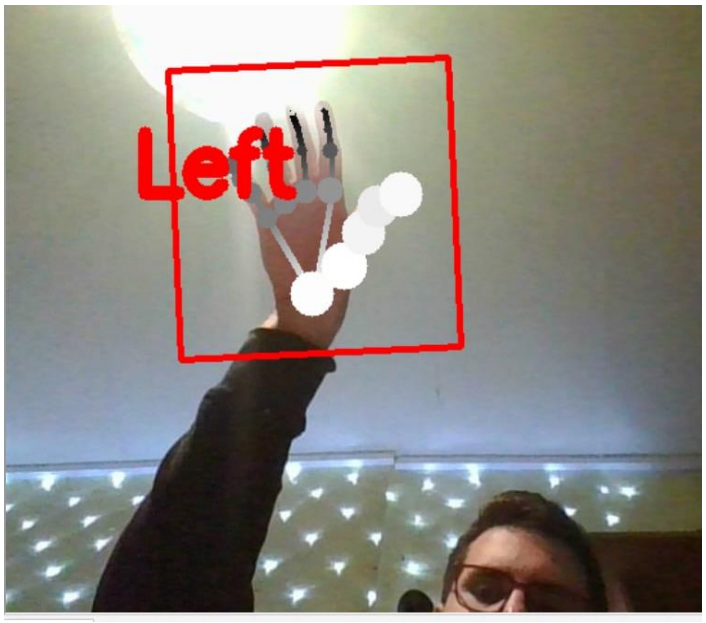

Рисунок А.4 – Приклад тестування системи в умовах яскравого освітлення

ДОДАТОК Б

Приклади тестування системи Handpose

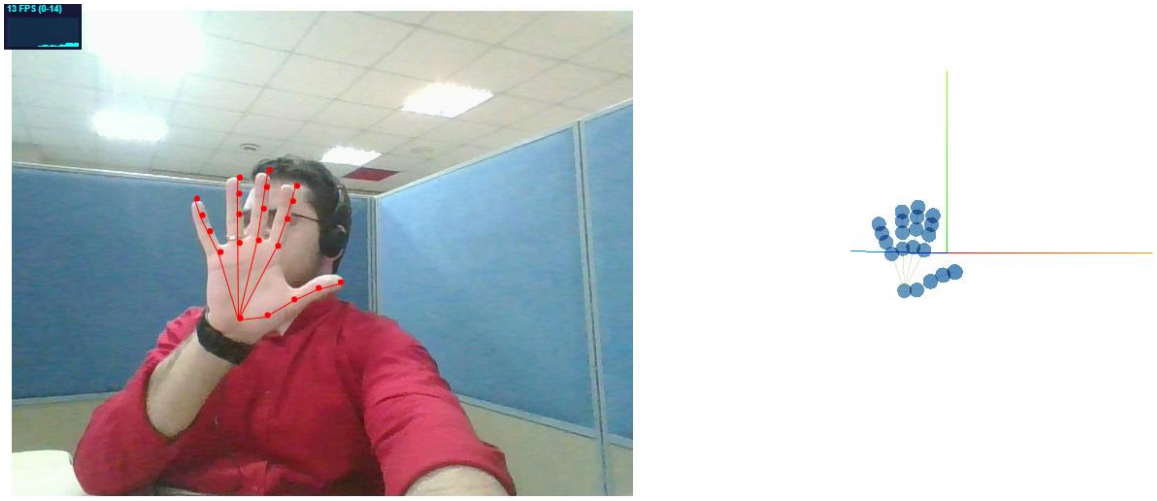

Рисунок Б.1 – Приклад тестування системи в умовах денного освітлення

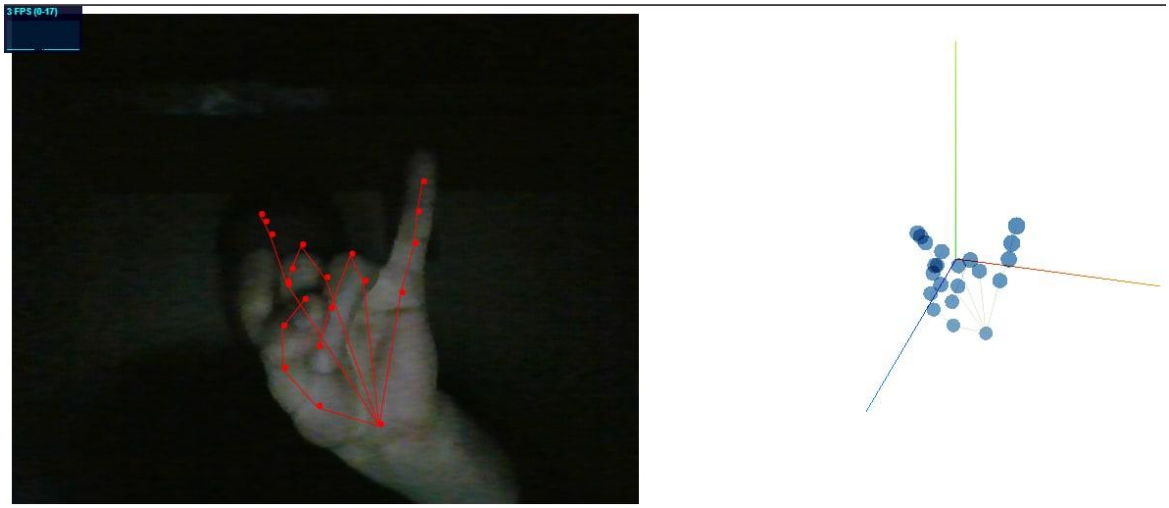

Рисунок Б.2 – Приклад тестування системи в умовах без освітлення

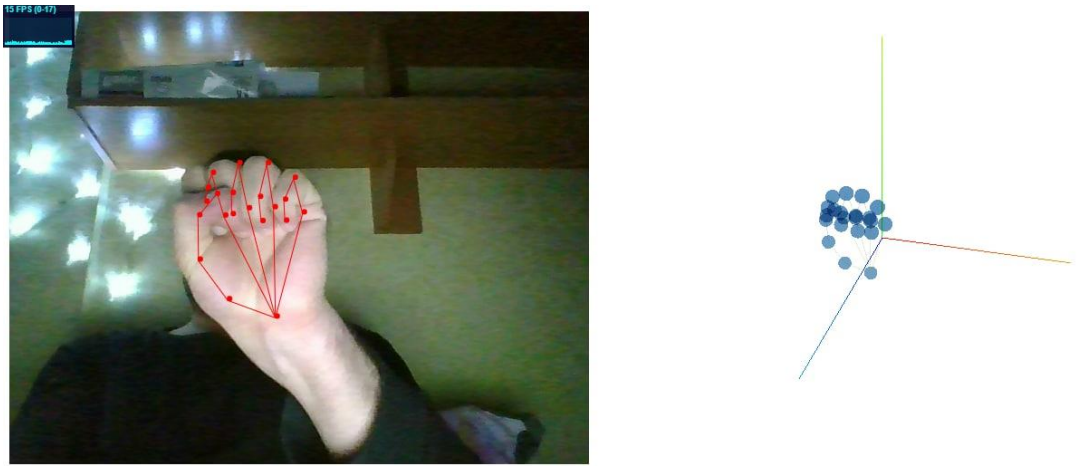

Рисунок Б.3 – Приклад тестування системи в умовах темного освітлення

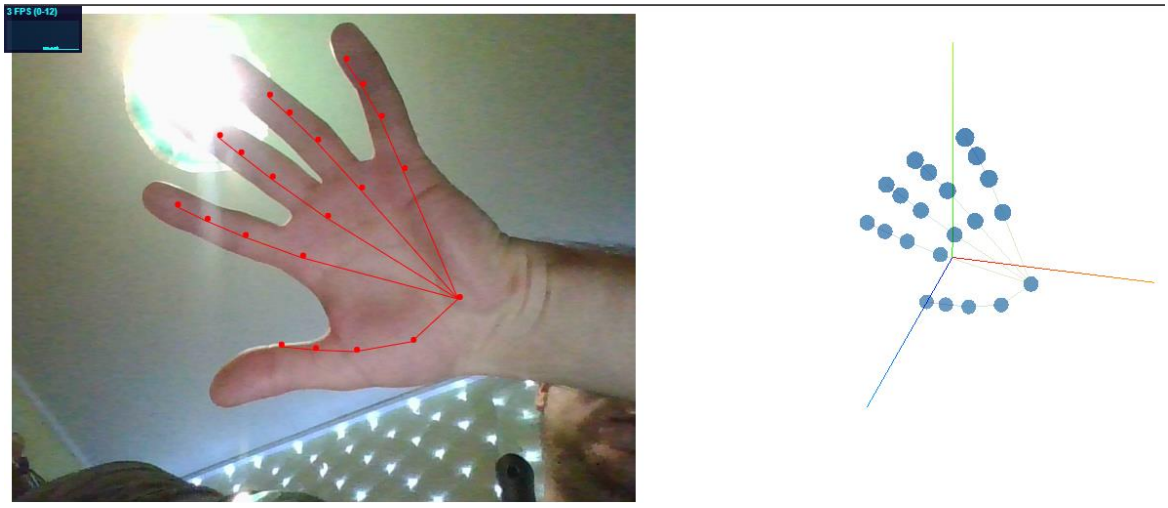

Рисунок Б.4 – Приклад тестування системи в умовах яскравого освітлення