

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
Західноукраїнський національний університет
Факультет комп'ютерних інформаційних технологій
Кафедра комп'ютерної інженерії

ЧИРСЬКА Тетяна Володимирівна

**Алгоритми розпізнавання жестів людини на основі
скелетонів / Human gesture recognition algorithms based on
skeletons**

спеціальність: 123 - Комп'ютерна інженерія
освітньо-професійна програма - Комп'ютерна інженерія
Кваліфікаційна робота

Виконав студент групи Кім-21
Т.В.. Чирська

Науковий керівник
д.т.н., професор, О. М. Березький

Кваліфікаційну роботу допущено
до захисту:

" ____ " _____ 20__ р.

Завідувач кафедри
_____ О.Л. Дубчак

АНОТАЦІЯ

Чирська Т.В. Алгоритми розпізнавання жестів людини на основі скелетонів. – Рукопис.

Випускна кваліфікаційна робота на із спеціальності 123 «Комп'ютерна інженерія» написана обсягом 80 сторінки і містить 24 ілюстрацій, 3 таблиць, 2 додатки та 49 джерел за переліком посилань.

Метою роботи є розроблення алгоритмів розпізнавання жестів людини на основі скелетонів для інтерактивних систем.

Методи досліджень: методи комп'ютерного зору, алгоритми розпізнавання жестів, методи моделювання руху та жестів.

Результати дослідження: алгоритми розпізнавання жестів людини, програмний модуль для аналізу жестів на основі скелетонів.

Результати роботи можуть бути використані у системах жестового управління, інтерактивних додатках, системах віртуальної реальності, а також у реабілітаційних технологіях.

Орієнтовні напрямки розвитку досліджень: інтеграція алгоритмів з технологіями доповненої реальності, розроблення адаптивних моделей для індивідуалізації аналізу жестів.

КЛЮЧОВІ СЛОВА: ЖЕСТИ, СКЕЛЕТОНИ , РОЗПІЗНАВАННЯ, КОМП'ЮТЕРНИЙ ЗІР, АЛГОРИТМ.

ANNOTATION

Chyrska T.V. Human gesture recognition algorithms based on skeletons – Manuscript.

The final qualification work in the specialty 123 "Computer Engineering" is written in 80 pages and contains 26 illustrations, 3 tables, 2 appendices and 49 sources for the list of references.

The method of work is the development of human gesture recognition algorithms based on skeletons for interactive systems.

Research methods: computer vision methods, gesture recognition algorithms, methods for modelling movement and gestures.

Research results: human gesture recognition algorithms, a software module for analyzing gestures based on skeletons.

The results of the work can be used in gesture control systems, interactive applications, virtual reality systems, as well as in rehabilitation technologies.

Approximate areas of research development: integration of algorithms with augmented reality technologies, development of adaptive models for individualizing gesture analysis.

KEYWORDS: GESTURES, SKELETONS, RECOGNITION, COMPUTER VISION, ALGORITHM.

ЗМІСТ

Вступ.....	5
1. Аналіз методів і технологій розпізнавання жестів	9
1.1 Керування пристроями з допомогою жестів.....	9
1.2.Аналіз наборів даних жестів	14
1.3.Аналіз алгоритмів розпізнавання жестів	18
1.4 Аналіз завдання кваліфікаційну роботу та постановка задач.....	24
1.5 Висновки до розділу 1	25
2. Алгоритми розпізнавання жестів на основі скелетонів	27
2.1. Аналіз алгоритмів розпізнавання жестів на основі графових моделей....	27
2.2 Формулювання вимог до розробки алгоритму	33
2.3 Розроблення алгоритму розпізнавання жестів на основі скелетонів	38
2.4 Висновки до розділу 2	42
3 Програмна система розпізнавання жестів	43
3.1 Структура системи	43
3.2 Алгоритм розпізнавання жестів.....	48
3.3Тестування роботи системи.....	52
3.2 Висновки до розділу 3	58
Висновки	59
Список використаних джерел	60
ДОДАТОК А	Помилка! Закладку не визначено.
ДОДАТОК Б	Помилка! Закладку не визначено.

Вступ

Розпізнавання жестів людини є ключовим етапом у розвитку інтерфейсів між людиною і машиною, і набуває все більшого значення завдяки стрімкому розвитку технологій комп'ютерного зору, штучного інтелекту та глибинного навчання. Визначення та інтерпретація жестів людини дозволяють створювати інноваційні інтерфейси для взаємодії з комп'ютерами, що не потребують фізичних пристроїв введення, таких як клавіатура чи миша. Такі системи можуть використовуватись в різноманітних сферах, таких як розваги, медичні технології, освіта, а також в індустрії безпеки та робототехніці. Наприклад, віртуальна та доповнена реальність використовують систему розпізнавання жестів для забезпечення безконтактної взаємодії користувача з комп'ютерними програмами, що дозволяє здійснювати управління за допомогою рухів рук або тіла. Системи безпеки, використовуючи таке розпізнавання, можуть ідентифікувати осіб або реагувати на небезпечні жести у реальному часі. Водночас медичні додатки можуть застосовувати технології розпізнавання жестів для моніторингу пацієнтів або в реабілітаційних цілях, допомагаючи відновленню рухових навичок.

Технічні труднощі, з якими стикаються сучасні системи розпізнавання жестів, зумовлені рядом факторів. Один із основних викликів полягає в наявності шумів у вхідних даних. Жести людини можуть бути непередбачуваними, і в реальному середовищі існують численні перешкоди, такі як зміни освітлення, фонові рухи або відображення на поверхнях. Ці перешкоди можуть суттєво вплинути на точність алгоритмів, які базуються виключно на аналізі зображень або відеопотоків. Наприклад, зміна освітлення може призвести до спотворення кольорів і текстур на зображеннях, що ускладнює розпізнавання жестів. Крім того, візуальна складність інтерпретації жестів може бути посилена варіативністю рухів людини, індивідуальними особливостями користувачів, а також фактором динамічної зміни положення руки чи тіла.

Відповідно до цих викликів, перспективним підходом є використання скелетних моделей. Скелетні моделі базуються на тривимірних координатах

суглобів людини, що дозволяє точно описати рухи тіла без урахування таких зовнішніх чинників, як текстура або колір. Важливим аспектом є те, що ці моделі фокусуються на структурних характеристиках рухів, таких як положення та орієнтація суглобів, що значно зменшує вразливість до змін у навколишньому середовищі. Тривимірні координати суглобів можуть бути отримані за допомогою спеціальних датчиків або за допомогою відеокамер з підтримкою глибини, що дозволяє реконструювати просторові координати людського тіла. Це дозволяє системам точніше відстежувати рухи та визначати жести без необхідності враховувати змінні умови освітлення або наявність інших перешкод на зображеннях.

Таким чином, використання скелетних моделей є ефективним підходом для розв'язання проблеми точного та швидкого розпізнавання жестів людини, зокрема у складних умовах реального середовища, і відкриває нові можливості для розвитку технологій безконтактної взаємодії в різноманітних галузях. Проведений критичний аналіз існуючих підходів до розпізнавання жестів показав, що традиційні методи не забезпечують необхідного рівня точності та адаптивності. У зв'язку з цим виникає потреба у розробленні нових алгоритмів, які б поєднували аналіз структурних особливостей рухів із можливостями машинного навчання.

Метою роботи є розробка алгоритмів, які дозволяють ефективно розпізнавати жести, спираючись на дані про положення суглобів у тривимірному просторі. Це передбачає розв'язання таких завдань, як аналіз сучасних методів класифікації жестів, вибір найбільш ефективних алгоритмів для обробки скелетних даних, розроблення нових підходів до стабілізації та фільтрації даних, а також тестування запропонованих методів на практичних наборах даних.

Об'єктом дослідження є процеси аналізу рухів людини на основі скелетних моделей, які надають можливість абстрагувати жести від зовнішніх характеристик, таких як колір шкіри, одяг чи фон. Предметом дослідження є алгоритми розпізнавання жестів, зокрема ті, які використовують графові структури для моделювання взаємного розташування суглобів і сучасні підходи машинного навчання для їх класифікації.

Методологія дослідження базується на інтеграції кількох підходів. Основними методами є алгоритми порівняння графів, такі як методи Уллмана і McKay's Nauty, що забезпечують точний структурний аналіз зв'язків між суглобами. Для підвищення точності та швидкості роботи були використані алгоритми глибинного навчання, які навчаються на реальних наборах даних і здатні адаптуватися до різноманітності жестів. Окрім цього, застосовувалися методи фільтрації та стабілізації даних, які зменшують вплив шумів та артефактів, що можуть виникати під час збору інформації з датчиків руху.

Наукова новизна роботи полягає у впровадженні комплексного підходу, який поєднує графові моделі з можливостями сучасних нейронних мереж. Уперше запропоновано алгоритм, що інтегрує аналіз графів для моделювання просторових взаємозв'язків суглобів із класифікацією, основою на глибинному навчанні. Це забезпечило значне підвищення точності та стабільності системи порівняно з існуючими підходами. Результати дослідження вперше демонструють можливість ефективного розпізнавання жестів у складних умовах, таких як наявність шумів, значні зміни освітлення або фонових перешкод.

Практичне значення одержаних результатів полягає в їх потенційному застосуванні у різних галузях. Розроблені алгоритми можуть бути використані в інтерактивних системах для людей із обмеженими можливостями, медичних діагностичних пристроях, реабілітаційних тренажерах, системах безпеки для ідентифікації жестів або створення безконтактних інтерфейсів у публічних просторах.

У першому розділі проведено аналіз алгоритмів розпізнавання жестів на основі скелетонів, здійснено аналіз завдання на магістерську роботу та зроблено постановку задач [1-2].

У другому розділі проаналізовано сучасні алгоритми розпізнавання жестів на основі скелетних моделей, зокрема підхід Ульмана та інші сучасні технології.

У третьому розділі розроблена структура програмної системи, програмні модулі розпізнавання жестів на основі скелетонів і проведено тестування та аналіз реалізованої системи.

У додатках представлені, світлокопії публікацій і код алгоритму.

Результати роботи були представлені на наукових конференціях і публікаціях у спеціалізованих виданнях. Публікації та апробація магістерської роботи [3, 4].

1. Чирська Т.В., Алгоритми порівняння скелетонів: IX науково-практична конференція молодих вчених і студентів «Інтелектуальні комп'ютерні системи та мережі» 21 травня 2024. Тернопіль. 2024. с 17

2. Чирська Т.В., Алгоритми розпізнавання жестів людини на основі скелетонів: Всеукраїнська науково-практична конференція студентів, аспірантів та молодих вчених «Інтелектуальні комп'ютерні системи та мережі» 5 листопада 2024. Тернопіль. 2024 с 152

Проведені апробації показали високу ефективність розроблених методів, що підтверджується позитивними відгуками експертів та експериментальними даними, отриманими під час тестування на реальних наборах даних. Це створює передумови для подальшого впровадження розробок у промисловість та їх адаптації до конкретних задач замовників.

1. АНАЛІЗ МЕТОДІВ І ТЕХНОЛОГІЙ РОЗПІЗНАВАННЯ ЖЕСТІВ

1.1 Керування пристроями з допомогою жестів

Жести є важливим аспектом взаємодії з технологіями та пристроями. Вони можуть бути використані для контролю різних функцій і можуть бути класифіковані за типами залежно від їх призначення та характеру виконання. Одним із основних видів жестів є жести для активації або деактивації пристроїв. Наприклад, жест «пальці в кулак» може активувати пристрій, а «відкриття руки» — вимикати його. Такі жести широко використовуються в системах безконтактного управління, де важлива швидкість та точність активації [5-16].

Ще одним поширеним типом жестів є жести для прокручування контенту на екрані. Це можуть бути рухи, як «крутий рух пальців» або «переміщення руки вгору/вниз», що дозволяє зручно переглядати великі обсяги інформації, наприклад Рисунок 1.1, при перегортанні сторінок або прокручуванні списків. Для масштабування контенту часто використовуються жести «стискання» або «розтягування» пальців, що дозволяє точно контролювати рівень масштабування без необхідності використовувати фізичні кнопки.



Рисунок 1.1- жест «переміщення руки вгору/вниз»,

Існують також жести, які імітують натискання кнопок, наприклад, «палець на певному місці». Це може бути використано для віртуального натискання кнопок або виконання інших інтерактивних дій без фізичного контакту. Для регулювання гучності можуть використовуватися такі жести, як «повертання руки» або «переміщення пальців вгору/вниз», що відповідно збільшує або зменшує рівень гучності рисунок 1.2, забезпечуючи зручний спосіб змінювати параметри звуку без відволікання від основної задачі.

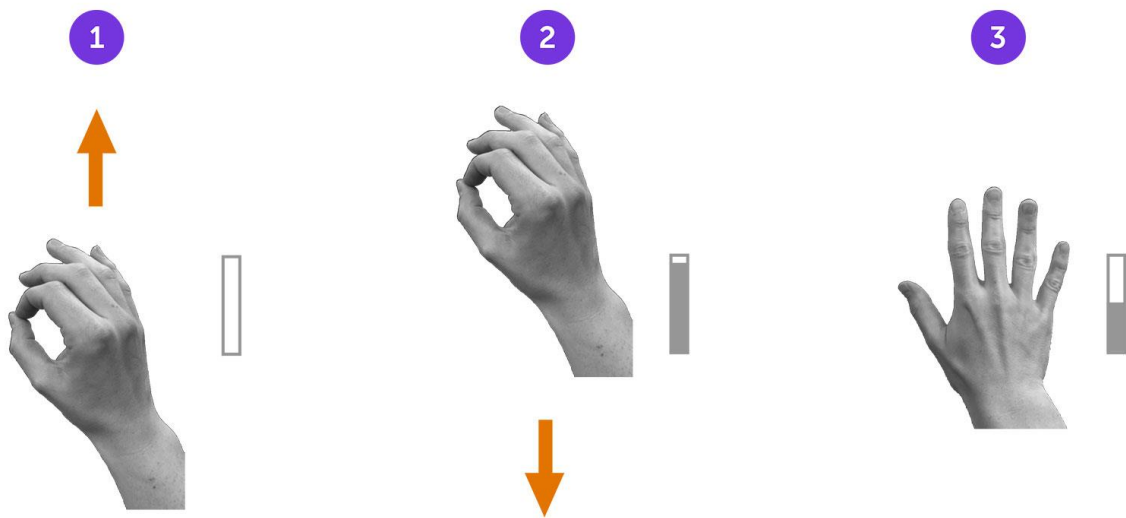


Рисунок 1.2-жест «переміщення пальців вгору/вниз»

Також популярними є жести для перемикання між екранами або додатками, як «проведення пальцем вліво/вправо» або «рух руки в сторону», що дозволяє швидко перемикатися між різними екранними режимами або додатками. Існують ще й функціональні жести, які можуть включати складні рухи, такі як «поворот долоні» або «круговий рух пальцями», що використовуються для активації певних функцій або налаштувань, таких як переключення між режимами пристрою.

Нарешті, жести для керування медіа включають «пальці для паузи» або «кружляння пальців для перемотування», що дозволяє керувати відтворенням музики, відео та іншого контенту без необхідності торкатися екрана або фізичних кнопок. Кожен з цих типів жестів має свої варіації, і їх точність та ефективність

можуть залежати від контексту використання. Тому дуже важливо провести тести на різних користувачах, щоб переконатися в зручності та ефективності кожного з них.

Гестова взаємодія може стати одним з основних способів взаємодії користувача з пристроями майбутнього. Вона дозволяє створювати інтуїтивно зрозумілі інтерфейси, які спрощують використання технологій. В залежності від контексту та цілей, жести можуть мати різні варіації і бути налаштовані під конкретні потреби користувача. Розвиток та впровадження систем розпізнавання жестів для керування технічними засобами вимагають балансу між точністю, швидкістю, енергоефективністю та сприйняттям користувачем. Успішні рішення у цій області враховують ці обмеження та виклики для досягнення оптимальної взаємодії.

Аналіз жестів для керування пристроями на основі скелетонів — це процес розпізнавання і інтерпретації рухів користувача через візуальні дані, отримані за допомогою різних сенсорних технологій. Цей підхід дозволяє визначати рухи частин тіла, таких як руки, голова та ноги, і інтерпретувати їх як команди для управління технічними пристроями.

Скелетна модель є основою для аналізу жестів. Вона створюється на основі 3D-моделі, яка визначає положення ключових точок суглобів, таких як плечі, лікті, кисті, коліна та інші. Такі моделі дозволяють отримати точне представлення рухів користувача. Для побудови цих моделей використовуються камери глибини, наприклад, Microsoft Kinect Рисунок 1.3 або Intel RealSense, а також технології доповненої реальності (AR), які відстежують рухи в реальному часі.



Рисунок 1.3- камера глибини Microsoft Kinect

Методи розпізнавання жестів передбачають аналіз положень та переміщень суглобів. Це може включати такі жести, як підняття руки, обертання кисті або нахил голови. Алгоритми машинного навчання, зокрема нейронні мережі, можуть класифікувати різні жести, аналізуючи дані, отримані з скелетних моделей. Для точного розпізнавання необхідно проводити навчання моделей на великих наборах даних, що містять приклади різних жестів, щоб знизити кількість помилок.

Одним із головних викликів є підвищення точності розпізнавання жестів у складних умовах, таких як погане освітлення або наявність перешкод. Важливо також забезпечити природність жестів, щоб користувачі не відчували втоми при тривалому використанні системи. Для цього необхідно мінімізувати кількість хибних спрацьовувань, використовуючи перевірку контексту виконання команд.

Системи, які використовують аналіз жестів для керування пристроями, мають широкий спектр застосувань. Вони використовуються в медичних пристроях (рис. 1.4), розумних будинках, автомобільній індустрії, ігрових консолях і віртуальній реальності. Важливим аспектом є інтеграція з інтернетом речей (IoT), що робить ці системи більш зручними та доступними для користувачів.



Рисунок 1.4- застосування жестів в медицині

Використання скелетних моделей для аналізу жестів дозволяє досягти високого рівня точності в управлінні пристроями. До стандартних суглобів можна додавати дрібніші сегменти, наприклад, фаланги пальців, що підвищує точність розпізнавання мікрожестів, таких як жести для мови жестів. Рухи аналізуються не лише в площині, але й у просторі, враховуючи глибину, швидкість і напрямок траєкторії.

Методи розпізнавання жестів можуть поєднувати класичні алгоритми аналізу траєкторій з сучасними методами машинного навчання. Система також може вимагати підтвердження жесту для уникнення хибних активацій, наприклад, через повторення жесту або додаткову дію, таку як зміна напрямку.

Алгоритми можуть бути адаптовані до стилю жестів конкретного користувача, що дозволяє створити персоналізований досвід. Комбіновані жести, наприклад, обертання кисті з одночасним рухом руки, можуть бути використані для виконання кількох дій одночасно. Крім того, система може комбінувати жести з голосовими командами або іншими методами введення, що робить управління більш інтуїтивно зрозумілим.

Використання нових сенсорів з низьким енергоспоживанням дозволяє застосовувати системи розпізнавання жестів навіть на мобільних пристроях з обмеженими ресурсами батареї. Алгоритми також повинні бути стійкими до зовнішніх факторів, таких як шум, перепади освітлення або перешкоди у вигляді інших об'єктів.

Жести можуть використовуватись для управління віртуальними моделями в віртуальній реальності, а також для налаштування освітлення, техніки чи систем безпеки в розумних будинках. У відеоіграх і на ігрових консолях системи жестів створюють більш занурюючий досвід для гравців.

Розвиток таких технологій покращує комфорт, точність і ефективність взаємодії людини з машинами, змінюючи способи управління пристроями та створюючи нові можливості для майбутніх інтерфейсів.

1.2. Аналіз наборів даних жестів

Аналіз наборів даних жестів є ключовим етапом у процесі тренування та валідації моделей розпізнавання жестів. Такі набори даних містять записи різноманітних жестів, які використовуються для вдосконалення алгоритмів розпізнавання. Однією з основних умов успішного розвитку систем розпізнавання жестів є доступ до якісних наборів даних, що містять різноманітні приклади жестів, зафіксованих у різних умовах та від різних користувачів. Аналіз таких наборів дозволяє виявити сильні та слабкі сторони цих даних, а також визначити можливості для подальшого покращення алгоритмів розпізнавання.

Одним із найбільш відомих наборів даних є ChaLearn Looking at People (LAP) Dataset [17-18]  рисунок 1.5. Цей набір містить велику кількість відеозаписів з різноманітними жестами та діями, що охоплюють широкий спектр рухів та дій, які можуть бути використані для тренування моделей розпізнавання жестів. Для кожного запису в цьому наборі є анотації, що визначають типи жестів та рухів, що дозволяє точно та ефективно навчати моделі. Набір даних LAP часто використовується для розробки систем віртуальної реальності, а також для створення технологій управління технічними засобами, такими як комп'ютери, смартфони та планшети. Це дозволяє здійснювати взаємодію з об'єктами у віртуальному середовищі або управління цифровими пристроями через рухи рук або інших частин тіла.

Jester Dataset [19] є ще одним важливим набором даних, який містить великий обсяг відеозаписів, що охоплюють різні класи жестів. Цей набір даних підходить для використання в глибокому навчанні та для розробки систем жестового керування. Оскільки він містить безліч прикладів жестів, його часто використовують для створення систем керування в іграх та розважальних додатках. Зокрема, цей набір застосовується для розробки інтерфейсів, що дозволяють управляти відеопрогравачами та аудіоплеєрами за допомогою жестів, що дозволяє зробити управління пристроями більш інтуїтивним і природним.

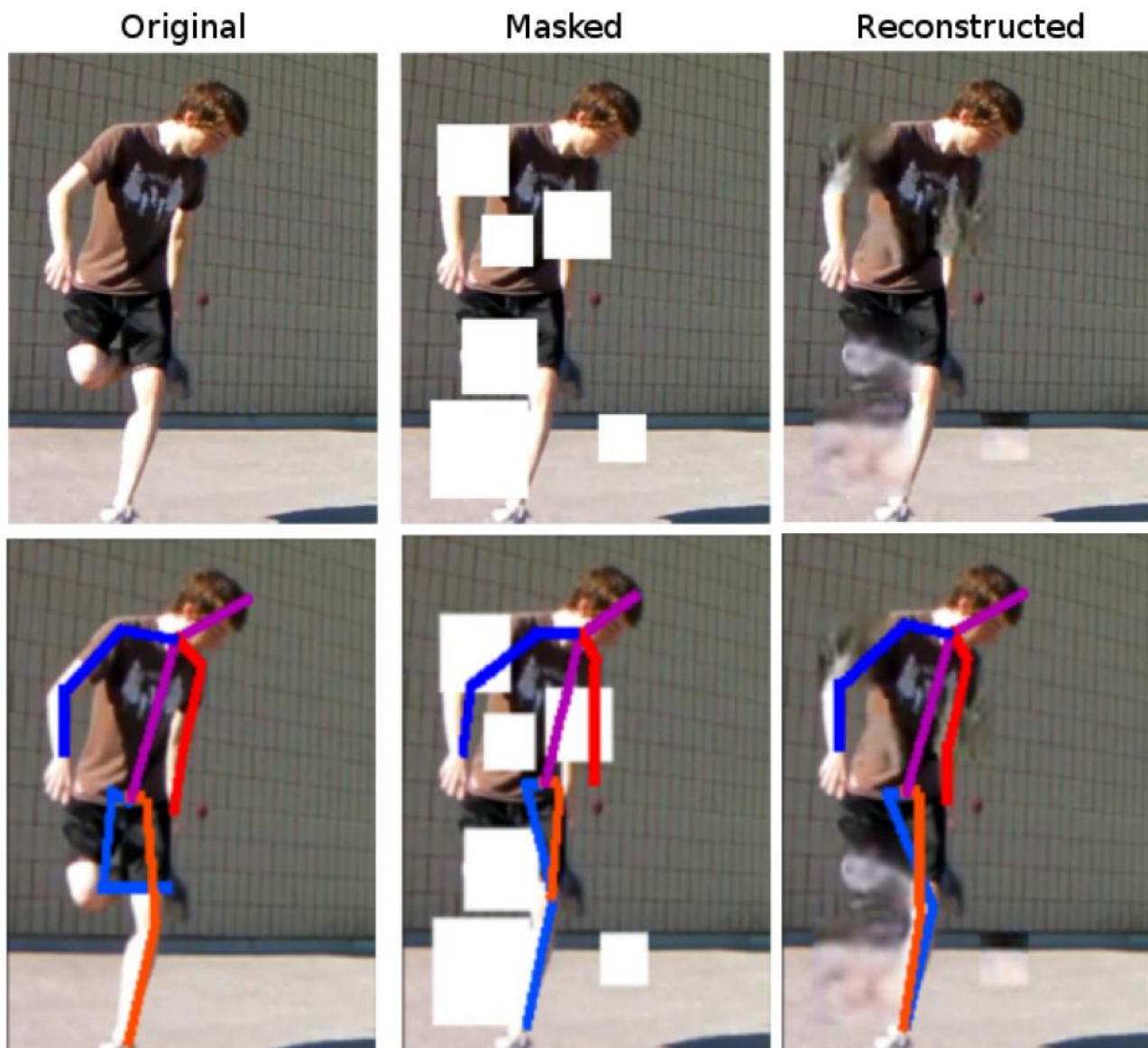


Рисунок1.5 – Приклад зображень жестів з датасету ChaLearn Looking at People (LAP)

MSR DailyActivity 3D Dataset [20] надає записані дані, отримані за допомогою 3D-камер, для дослідження розпізнавання жестів та дій в тривимірному просторі. Цей набір даних дає змогу вивчати тривимірне розпізнавання жестів та дій, що є важливим для створення більш складних і точних систем розпізнавання, здатних враховувати просторові координати. Даний набір використовують для розробки систем відслідковування та розпізнавання жестів у таких сферах, як домашня автоматизація, де точність та ефективність розпізнавання рухів є критично важливими для управління технічними засобами.

EgoGesture Dataset [21] є набором даних, що включає відеозаписи з точки зору користувача, або від першої особи. Це дає змогу вивчати персональні жести, які виконуються в реальному житті, і застосовувати їх у системах допомоги людям з обмеженими можливостями. Набір містить як RGB, так і глибинні зображення, що дозволяє розробляти більш точні системи розпізнавання жестів. Цей набір даних є особливо корисним для створення експериментальних систем допомоги особам з обмеженими можливостями, таких як системи для розпізнавання жестів у контексті медичних або реабілітаційних технологій.

Intel RealSense Hand Tracking Dataset [22] рисунок 1.6 використовує камери глибинного зору для точного фіксування рухів рук і жестів. Це дозволяє здійснювати високоточне відслідковування рухів, що є особливо корисним для інтеграції в ноутбуки та інші пристрої, де рухи рук використовуються для взаємодії з екраном або з віртуальними об'єктами. Завдяки високій точності фіксації рухів рук цей набір даних відкриває нові можливості для розвитку інтерфейсів, які дозволяють здійснювати безконтактну взаємодію з технологіями.

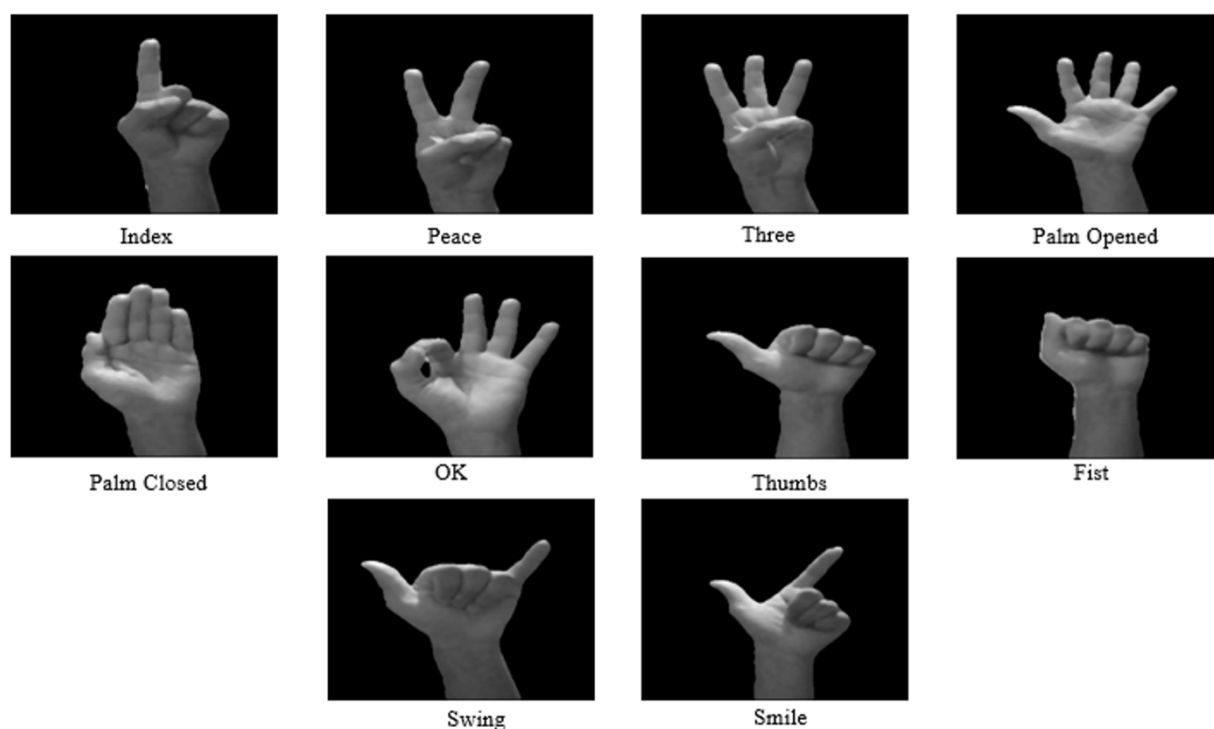


Рисунок 1.6 – Приклад зображень жестів з Intel RealSense Hand Tracking Dataset

Ще одним важливим набором є UCI Multimodal Human Action Dataset, який комбінує аудіо- та візуальні дані для розпізнавання дій та жестів. Це дає змогу досліджувати, як взаємодія аудіо- та відео даних може покращити точність розпізнавання жестів, особливо в складних середовищах, де важливо враховувати не тільки рухи, але й інші сенсорні дані, такі як звук. Цей набір широко використовується в наукових дослідженнях, зокрема в медичних або технічних сферах, для розробки систем відстеження та розпізнавання жестів, що можуть бути використані для покращення досліджень або медичних процесів.

Кожен із цих наборів даних має свої унікальні характеристики та можливості, що дозволяють застосовувати їх у різних галузях, таких як ігри, віртуальна реальність, системи допомоги особам з обмеженими можливостями та наукові дослідження. Вибір конкретного набору даних залежить від цілей дослідження або розробки, а також від специфічних вимог до точності та виду взаємодії, яка повинна бути реалізована в кінцевій системі.

Оцінка проаналізованих наборів даних відображена в Таблиці 1.1

Таблиця 1.1 - Оцінка наборів даних

Набір даних	Різноманітність жестів	Розмір набору даних	Анотації та мітки	Якість та тип даних	Простота обробки та доступність
ChaLearn Looking at People (LAP)	9/10	8/10	9/10	8/10	9/10
Jester Dataset	8/10	9/10	8/10	8/10	9/10
MSR DailyActivity 3D	7/10	9/10	8/10	9/10	8/10
EgoGesture Dataset	7/10	8/10	8/10	8/10	8/10
Intel RealSense Hand Tracking	6/10	7/10	8/10	9/10	7/10
UCI Multimodal Human Action	8/10	9/10	8/10	8/10	8/10

1.3. Аналіз алгоритмів розпізнавання жестів

Алгоритми розпізнавання жестів на основі скелетонів дозволяють аналізувати рухи людини, перетворюючи їх у числові моделі, які легко обробляються комп'ютерами. Ці алгоритми використовують дані про ключові точки тіла, такі як положення суглобів, і аналізують взаємодію між ними. Завдяки цьому вдається точно інтерпретувати рухи без необхідності обробки великих обсягів візуальної інформації. Такі алгоритми є особливо корисними в умовах, де потрібна висока швидкість роботи, мінімальні апаратні ресурси та прозорість функціонування системи. Їх застосування охоплює широкий спектр галузей, включаючи робототехніку, спорт, медицину та інтерактивні системи.

Одним із поширених підходів є динамічне викривлення часу (Dynamic Time Warping, DTW), який аналізує часові послідовності, наприклад, траєкторії руху суглобів, обчислюючи мінімальну відстань між ними. Цей метод дає змогу порівнювати рухи навіть за умов різної швидкості їх виконання. Однак, DTW має високу обчислювальну складність і вимагає попереднього визначення ключових точок, що обмежує його використання у великих наборах даних.

Приховані марковські моделі (Hidden Markov Models, HMM) дозволяють представляти жести як послідовності станів із певними ймовірностями переходу. Цей підхід добре працює з повторюваними та послідовними рухами, однак є чутливим до якості навчальних даних і вимагає ретельної попередньої обробки. Приховані марковські моделі представлені у вигляді цепки рисунок 1.7

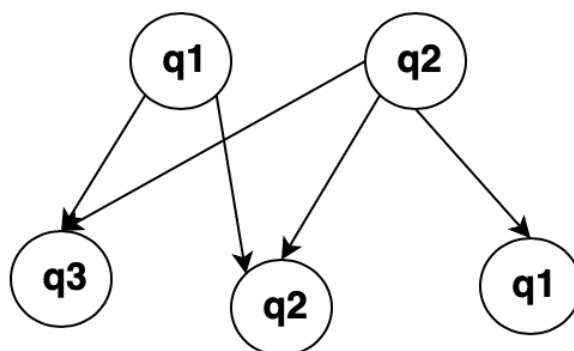


Рисунок 1.7- Приховані марковські моделі

Геометричний аналіз траєкторій базується на оцінці відносних положень ключових точок тіла, кутів між сегментами та траєкторій руху суглобів. Такий підхід є простим і ефективним для розпізнавання базових рухів, але має обмеження щодо складних жестів.

Сегментація рухів із використанням локальних екстремумів швидкості та прискорення дозволяє визначати початок і кінець жестів у потокових даних. Цей метод добре підходить для роботи в реальному часі, проте менш ефективний у випадку складних чи накладених рухів.

Спектральний аналіз рухів використовує перетворення часових даних у частотну область, що дає змогу аналізувати характерні ритми чи циклічні особливості рухів. Він є корисним для аналізу періодичних рухів, однак слабо працює зі складними жестами.

Векторний аналіз напрямку зосереджується на аналізі напрямків рухів між ключовими точками для побудови траєкторії жесту. Цей метод має високу точність для динамічних рухів, проте є чутливим до шуму у вхідних даних.

Перевагами класичних алгоритмів розпізнавання жестів є зрозумілий механізм їхньої роботи, відсутність вимог до значних обчислювальних ресурсів та стабільність роботи за умов обмежених наборів даних. Водночас вони мають обмеження, серед яких низька адаптивність до нових типів жестів, чутливість до якості вхідних даних і складність обробки сценаріїв із великою кількістю варіацій жестів.

Попри свої обмеження, класичні алгоритми залишаються важливим інструментом для створення систем, які потребують чіткої прозорості, надійності та ефективності. Їхня інтеграція із сучасними методами, такими як глибокі нейронні мережі, дозволяє досягати більшої гнучкості й точності, що відкриває нові можливості для аналізу рухів людини та їхнього застосування у реальному світі.

Математично скелетон представлений у вигляді графа, де вершини відповідають ключовим точкам, а ребра моделюють зв'язки між ними. Цей граф може бути орієнтованим або неорієнтованим залежно від задачі. Для аналізу

рухів використовуються алгоритми, що поєднують просторову та часову інформацію. Наприклад, просторово-часові згорткові нейронні мережі (ST-GCN) враховують зміни положення ключових точок у часі, що дозволяє моделювати динаміку складних рухів. Використання механізмів уваги додатково підвищує точність, фокусуючи обчислення на найбільш важливих точках

Скелетон знаходить застосування у медицині [23-25] для аналізу ходи та реабілітації, у спорті та освіті для моніторингу тренувань і навчання, в ігровій індустрії для створення реалістичних анімацій, а також у системах доповненої реальності. Енергоефективні алгоритми дозволяють впроваджувати технологію на пристроях з обмеженим енергоспоживанням. Завдяки своїй універсальності, скелетон став важливим інструментом для аналізу жестів та поведінки, особливо у VR/AR-середовищах.

Основні формули для аналізу скелетонів у розпізнаванні жестів

Обчислення відстаней між ключовими точками

Для визначення жестів і поз, важливо виміряти відстані між певними ключовими точками на тілі, наприклад, між кистю і плечем, або між коліном і стегном.

$$d_{i,j} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2}$$

де:

$d_{i,j}$ — відстань між ключовими точками і та j,

- (x_i, y_i, z_i) та (x_j, y_j, z_j) — координати відповідних точок.

Ця формула дозволяє визначити ступінь згину кінцівок або відстань між будь-якими двома ключовими точками для аналізу пози тіла.

Кут між трьома ключовими точками

Часто важливо визначити кут між кінцівками, наприклад, щоб зрозуміти, чи є рука зігнутою або випрямленою. Кут між трьома точками можна обчислити за допомогою скалярного добутку векторів.

Нехай A,B,C— три ключові точки, що утворюють два вектори

$$\cos\theta = \frac{|\vec{AB}| \cdot |\vec{BC}|}{|\vec{AB}| \cdot |\vec{BC}|}$$

де:

$$\vec{B} \cdot \vec{C} = (x_b - x_a)(x_c - x_a) + (y_b - y_a)(y_c - y_a) + (z_b - z_a)(z_c - z_a)$$

— скалярний добуток векторів,

- $|\vec{AB}|$ та $|\vec{BC}|$ — довжини відповідних векторів.

Кут θ визначає згин кінцівки і може використовуватись для розпізнавання складних жестів, таких як підняття руки або зігнуття коліна.

Середнє положення точки (центру мас) скелетону

Центр мас є важливим для аналізу стійкості тіла та загальної позиції, особливо для аналізу балансу. Його можна знайти як середнє положення всіх ключових точок.

$$C = \left(\frac{1}{n} \sum_{i=1}^n x_i, \frac{1}{n} \sum_{i=1}^n y_i, \frac{1}{n} \sum_{i=1}^n z_i \right)$$

де:

- n — кількість ключових точок,

(x_i, y_i, z_i) — координати кожної ключової точки.

Центр мас допомагає відстежувати положення тіла в просторі та визначати зміщення центру ваги при виконанні жестів.

* Швидкість переміщення ключової точки

Для оцінки рухів у динамічних жестах важливо визначити швидкість переміщення ключових точок. Швидкість точки i у момент часу t обчислюється як:

$$v_i(t) = \sqrt{(x_i(t) - x_i(t-1))^2 + (y_i(t) - y_i(t-1))^2 + (z_i(t) - z_i(t-1))^2}$$

де:

— $v_i(t)$ швидкість точки i у момент часу t ,

$$(x_i(t), y_i(t), z_i(t)) \text{ та } (x_i(t-1), y_i(t-1), z_i(t-1))$$

— координати точки i у моменти часу t та $t-1$,

— Δt інтервал часу між вимірами.

Цей показник дозволяє аналізувати динамічні рухи та розрізняти швидкі й повільні жести.

Прискорення переміщення ключової точки

Прискорення важливе для аналізу швидких змін у рухах і жестах, таких як швидкі махи руками. Прискорення точки і можна обчислити як:

$$a_i(t) = \frac{v_i(t) - v_i(t - 1)}{\Delta t}$$

де:

- $a_i(t)$ — прискорення точки і у момент часу t ,

- $v_i(t)$ та $v_i(t - 1)$ — швидкості точки і у моменти часу t та $t-1$.

Прискорення дозволяє виявити різкі зміни руху, що можуть вказувати на специфічні жести або дії.

Скелетон є потужним інструментом у розпізнаванні жестів, дозволяючи моделювати взаємодію між частинами тіла. Використання графових моделей, зокрема графів забезпечує високу точність і ефективність аналізу рухів. Завдяки своїй універсальності, скелетон знаходить застосування у багатьох галузях, від медицини до ігрових технологій, роблячи його важливим елементом сучасних досліджень і розробок. Скелетон є універсальною моделлю, що дозволяє аналізувати та моделювати людські рухи. Його графова інтерпретація забезпечує компактність і високу ефективність обчислень, що робить його незамінним у задачах розпізнавання жестів, аналізу поведінки та створення інтерактивних систем.

Аналіз програмних засобів розпізнавання жестів передбачає дослідження різних бібліотек, фреймворків та інструментів, які дозволяють інтегрувати та реалізовувати системи розпізнавання жестів у додатках [26]. Одним із провідних інструментів у цій сфері є OpenPose, бібліотека з відкритим кодом, що забезпечує виявлення ключових точок тіла, обличчя, рук та ніг на зображеннях і відео. Її перевагою є можливість роботи в реальному часі, що робить її корисною у сферах віртуальної реальності, анімації, системах безпеки та ігровій індустрії.

Ще одним потужним інструментом є MediaPipe від Google, який призначений для обробки мультимедійних даних. Він забезпечує розпізнавання ключових точок обличчя, рук і тіла, а також підтримує інтеграцію з додатками для Android, iOS та веб-платформ. MediaPipe широко використовується для створення фільтрів відеобробки, розробки інтерактивних систем та в ігровій індустрії.

Kinect SDK, розроблений Microsoft, є набором інструментів, орієнтованих на тривимірне розпізнавання рухів за допомогою камери Kinect. Його функціонал включає виявлення скелета в просторі, використання глибинного зору для точного визначення положення об'єктів, а також інтеграцію з платформами Windows. Це робить Kinect SDK корисним у відеоіграх, системах відстеження рухів і віртуальній реальності.

Gesture Recognition Toolkit (GRT) є універсальною бібліотекою для розпізнавання жестів, що надає відкритий код і зручний API для створення власних алгоритмів. GRT підтримує інтеграцію з різними сенсорами, що робить її особливо корисною в дослідницьких проектах і розробці індивідуальних рішень.

Intel RealSense SDK забезпечує роботу з камерами Intel RealSense, які підтримують глибинний зір і можуть використовуватися для аналізу руху та взаємодії з об'єктами у просторі. Інструментарій дозволяє розробляти інтерактивні додатки для Windows і Linux, які знаходять застосування в системах відстеження рухів і апаратних пристроях.

Серед переваг програмних засобів розпізнавання жестів можна виділити інтуїтивність взаємодії, високу точність завдяки використанню глибокого навчання та широкі можливості у віртуальній реальності, ігровій індустрії й інших галузях. Водночас ці засоби стикаються з викликами, такими як залежність від якості вхідних даних, обмеження точності у складних умовах та питання конфіденційності й безпеки, пов'язані із використанням камер і сенсорів.

Таким чином, програмні засоби розпізнавання жестів відкривають широкі можливості для взаємодії з технологіями. Вибір конкретного інструменту

залежить від вимог проекту, включаючи тип платформи, характер жестів і необхідний рівень точності.

1.4 Аналіз завдання кваліфікаційну роботу та постановка задач

Розпізнавання жестів людини є важливою складовою сучасних систем взаємодії між людиною і комп'ютером, що має широке застосування в таких галузях, як медичні технології, робототехніка, відеоігри, системи безпеки, освітні технології та реабілітація. У контексті цього дослідження основна увага зосереджена на створенні ефективних алгоритмів, здатних точно інтерпретувати жести людини на основі даних, отриманих із скелетонної моделі.

Скелетонна модель являє собою спрощене представлення положення частин тіла людини у вигляді ключових точок, таких як суглоби, та їхніх просторових координат. Такі дані найчастіше отримують за допомогою сучасних технологій, наприклад, глибинних камер або спеціалізованого програмного забезпечення, що підтримує аналіз скелетної структури, таких як OpenPose або Microsoft Kinect. Вибір скелетонної моделі зумовлений її здатністю зменшити обчислювальну складність алгоритмів і зосередитися на ключових характеристиках руху.

Розв'язання проблеми розпізнавання жестів стикається з низкою викликів, зокрема з необхідністю забезпечення стійкості алгоритмів до індивідуальних варіацій жестів, таких як фізичні особливості людини, темп руху, кут зйомки або похибки сенсорних пристроїв. Крім того, класифікація жестів ускладнюється через наявність схожих рухів, які мають незначні відмінності. Додаткові труднощі виникають при обробці великих обсягів даних у реальному часі, що є критично важливим для інтерактивних систем. Також актуальним є завдання інтеграції скелетонної інформації з іншими сенсорними даними, такими як відео чи аудіо, для підвищення загальної точності розпізнавання.

У межах цього дослідження передбачається розробка та аналіз алгоритмів, які забезпечують точне розпізнавання жестів на основі скелетонів, оптимізацію їхньої роботи для застосування в реальних умовах, а також інтеграцію в системи взаємодії людина-комп'ютер. Основними науковими завданнями є створення методів представлення скелетонних моделей, які дозволять мінімізувати втрати важливої інформації та забезпечити високу якість аналізу рухів. Також передбачено розробку сучасних алгоритмів класифікації жестів із використанням глибоких нейронних мереж, що враховують часові залежності між положеннями суглобів.

Ключовим аспектом дослідження є вдосконалення методів стабілізації даних для зменшення впливу шумів і артефактів, що виникають під час отримання скелетної інформації. З метою підвищення ефективності роботи алгоритмів буде досліджено підходи до оптимізації їхнього виконання в режимі реального часу. Окрема увага приділяється тестуванню розроблених методів на різноманітних наборах даних для забезпечення їхньої універсальності й адаптивності до умов реального використання.

Результати дослідження матимуть значний вплив на розвиток сучасних технологій, орієнтованих на автоматизацію, інтерактивну взаємодію та інноваційні способи управління системами.

1.5 Висновки до розділу 1

Аналіз жестів для керування пристроями є перспективним напрямом, що підвищує зручність і ефективність взаємодії з технікою. Технологія знаходить застосування в смартфонах, «розумних» будинках, автомобілях, ігрових консолях та інших пристроях. Сучасні системи розпізнавання жестів базуються на скелетних моделях, які використовують тривимірний аналіз рухів, та алгоритмах машинного навчання для точного розпізнавання в реальному часі. Основними викликами залишаються точність у складних умовах, природність жестів,

енергоефективність і забезпечення конфіденційності даних. Попри це, інтеграція таких систем в різні галузі відкриває нові можливості для автоматизації й комфорту.

2. АЛГОРИТМИ РОЗПІЗНАВАННЯ ЖЕСТІВ НА ОСНОВІ СКЕЛЕТОНІВ

2.1. Аналіз алгоритмів розпізнавання жестів на основі графових моделей

Алгоритми ізоморфізму підграфів є важливим напрямком у теорії графів і має широке застосування в різних галузях, зокрема в аналізі жестів, біоінформатиці та комп'ютерному зорі. Задача ізоморфізму підграфів полягає в знаходженні відповідності між підграфами двох графів таким чином, щоб зберігалася структура з'єднань вершин і ребер. Це є складною обчислювальною задачею, оскільки належить до класу NP-повних проблем, що означає відсутність ефективних методів для її розв'язання на великих графах без застосування евристичних підходів або оптимізацій.

Одним із найбільш поширених методів для розв'язання цієї проблеми є алгоритм VF2 [27], який використовує метод глибокого пошуку з поверненням (backtracking). Основною особливістю цього алгоритму є поступове знаходження відповідей між вершинами двох графів, перевіряючи на кожному етапі, чи можна продовжити поточну відповідність без порушення структури графа. Важливою оптимізацією є перевірка ступеня вершин перед встановленням відповідей, що дозволяє значно зменшити кількість перевірок. Також алгоритм використовує принцип симетрії, що дозволяє вибирати відповідні вершини таким чином, щоб зменшити кількість можливих варіантів і покращити ефективність пошуку.

Іншим потужним інструментом є алгоритм Nauty, який реалізує принцип нормалізації графів. Він дозволяє перевести граф у стандартне представлення, яке не залежить від порядку вершин. Порівнюючи нормалізовані графи, можна ефективно перевірити їх ізоморфізм. Важливим аспектом Nauty є те, що алгоритм здійснює перестановки вершин і шукає таку, що призводить до однакової структури зв'язків у двох графах. Завдяки цьому, якщо два графи мають однакову нормалізацію, вони є ізоморфними. Алгоритм використовує спеціальні оптимізації, засновані на властивостях симетрії графів, що дозволяє

значно скоротити кількість перевірок і підвищити ефективність. Блок-схема алгоритму відображена на рисунку 2.1

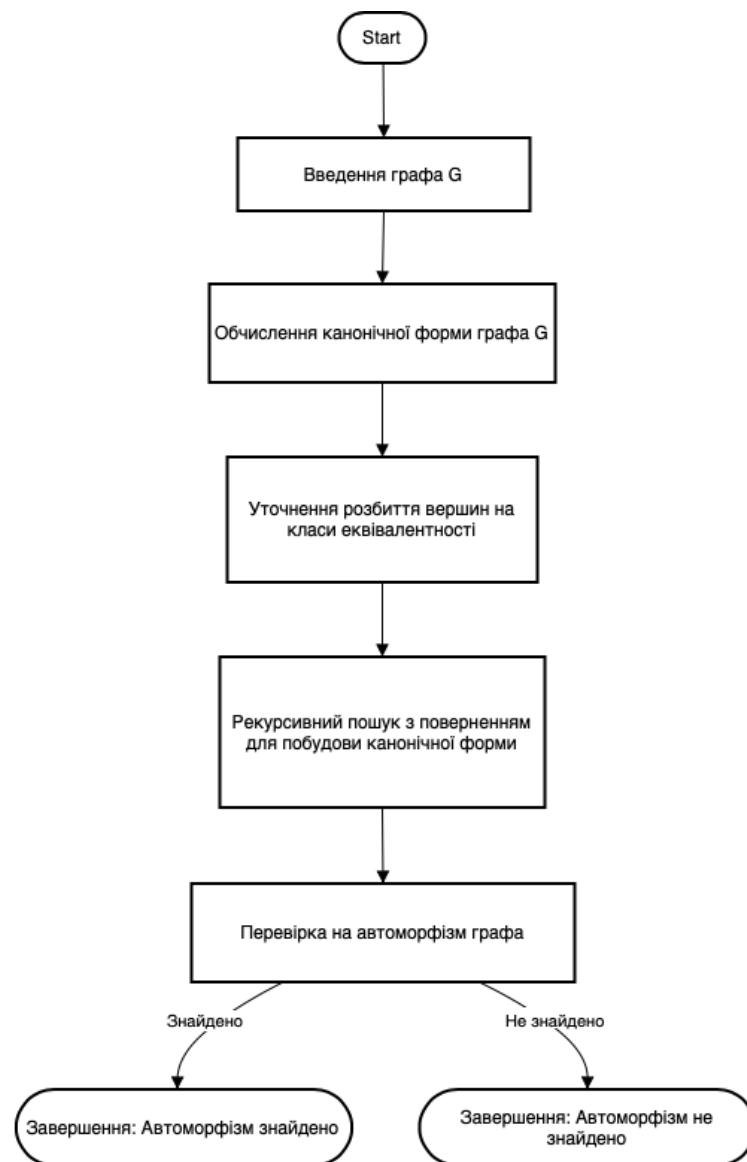


Рисунок 2.1-блок -схема алгоритму Nauty

Алгоритм Nauty є надзвичайно потужним інструментом для розв’язання задачі ізоморфізму графів і знаходження канонічного представлення графу. Його ключова ідея полягає у використанні нормалізації графів, що дозволяє перетворювати графи в стандартний вигляд, незалежний від початкового розташування чи іменування вершин. Завдяки цьому два графи можна порівняти

за їх канонічним представленням, що є ефективним методом для визначення ізоморфізму.

Основним механізмом Nauty є генерація перестановок вершин для пошуку такої, що забезпечує ідентичну структуру зв'язків між двома графами. Алгоритм використовує спеціальну процедуру кольорової класифікації (color refinement), яка зменшує кількість можливих перестановок, групує вершини в класи еквівалентності за їх характеристиками, наприклад, за ступенем чи іншими інваріантами. Це дозволяє значно прискорити процес.

Однією з найбільших переваг алгоритму є його оптимізації на основі симетрій графів. У багатьох випадках графи мають високу симетрію, і Nauty ефективно використовує цю властивість для скорочення кількості необхідних перевірок. Наприклад, якщо граф є регулярним, тобто всі вершини мають однаковий ступінь, алгоритм може пропустити багато еквівалентних перестановок, фокусуючись лише на ключових представленнях.

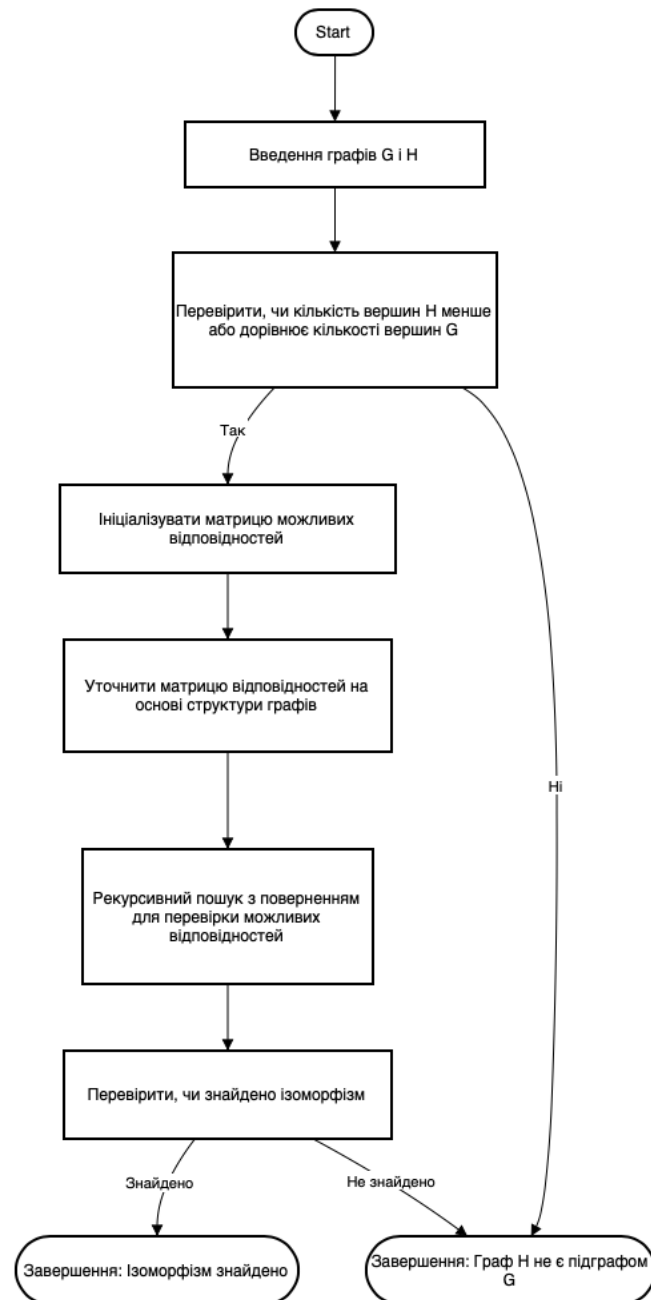
Алгоритм Nauty також відомий своєю здатністю працювати з дуже великими графами. Його ефективність досягається завдяки застосуванню спеціальних структур даних, таких як розширені матриці суміжності та дерева пошуку. Ці структури дозволяють швидко знаходити необхідні перестановки та перевіряти їх на ізоморфізм.

Додатково, Nauty дозволяє працювати не лише з простими графами, але й з графами, що мають додаткові атрибути, наприклад, орієнтованими графами чи графами з вагами на ребрах. Це робить його універсальним інструментом для широкого спектру задач, включаючи перевірку еквівалентності схем, аналіз хімічних молекул та багато іншого.

Таким чином, Nauty є одним із найбільш ефективних алгоритмів для задач ізоморфізму графів і канонічного представлення. Його інноваційний підхід, заснований на комбінаторних оптимізаціях і використанні властивостей симетрії, дозволяє досягти високої продуктивності навіть у складних випадках.

Алгоритм Уллмана, що є класичним методом для пошуку ізоморфізмів підграфів, використовує підхід побудови матриці відповідностей між вершинами двох графів. Після того, як для кожної пари вершин визначена можливість

зв'язку, алгоритм перевіряє наявність відповідних ребер між ними. Якщо структура зв'язків збігається, графи вважаються ізоморфними. Хоча цей метод є менш ефективним для великих графів порівняно з VF2 або Nauty, він залишається корисним для задач з обмеженнями на розмір графа або для спеціалізованих застосувань, таких як хімічні структури або структурний аналіз даних. Приклад роботи алгоритму представлені на рисунку 2.2



Риснок 2.2- блок-схема алгоритму Уллмана

Методи забарвлення, в свою чергу, передбачають присвоєння кожній вершині графа певного кольору, що відображає її структурну роль. Використовуючи це представлення, алгоритми шукають відповідні вершини в двох графах, порівнюючи їхні кольори і перевіряючи відповідність їхніх зв'язків. Цей підхід є ефективним для простих графів, де можна легко класифікувати вершини, однак у випадку складних графів з великою кількістю вершин і ребер він може потребувати додаткових методів оптимізації.

Динамічне програмування є ще одним методом, який може бути застосований для задачі ізоморфізму підграфів. Цей метод розбиває задачу на підзадачі, які вирішуються окремо, а потім комбінуються для отримання остаточного рішення. Використання динамічного програмування дозволяє знижувати складність обчислень, особливо коли графи мають певні властивості, що дозволяють їх декомпозицію.

Алгоритми для перевірки ізоморфізму підграфів знаходять широке застосування в багатьох сферах науки і технологій. У контексті розпізнавання жестів ці алгоритми використовуються для порівняння рухів людини, представлених у вигляді графів, що відображають послідовність рухів суглобів. Завдяки точності алгоритмів можна досягти високої ефективності у виявленні схожих рухів або відстеженні змін у позах людини. Вони також використовуються в біоінформатиці для пошуку подібностей між молекулами, де молекули представлені у вигляді графів, а ізоморфізм використовується для пошуку подібних структур.

Додатково до згаданих алгоритмів, важливими є також методи на основі глибинного навчання, які можуть бути застосовані для вирішення задачі ізоморфізму підграфів. Один з підходів полягає в використанні нейронних мереж, які навчаються розпізнавати та порівнювати структури графів без необхідності явного визначення відповідностей між їхніми елементами. Ці методи особливо ефективні, коли мова йде про складні графи, де традиційні алгоритми можуть не спрацювати через високий рівень варіативності.

Ще один перспективний напрямок — використання алгоритмів для згортання графів (Graph Convolutional Networks, GCN). Вони дозволяють

ефективно обробляти дані графів, враховуючи не лише структуру з'єднань між вершинами, а й значення самих вершин. Цей підхід здатен враховувати не лише топологічні особливості графів, а й їх контекст, що значно підвищує точність у задачах розпізнавання та класифікації.

Сучасні методи глибинного навчання мають значну перевагу в задачах, де структура даних є складною або змінною, що робить їх застосування надзвичайно перспективним у розпізнаванні жестів на основі графових моделей. Глибинне навчання дозволяє уникнути необхідності вручну налаштовувати алгоритми для кожного окремого випадку, що значно спрощує розробку та оптимізацію систем для розпізнавання складних рухів.

Також у деяких випадках використовуються комбіновані підходи, які поєднують класичні методи з методами глибинного навчання. Наприклад, класичний алгоритм ізоморфізму підграфів може бути використаний для попереднього фільтрування кандидатів на відповідність, після чого нейронна мережа здійснює точнішу класифікацію на основі цих попередньо оброблених даних.

Інші важливі підходи включають використання підходів на основі графових відстаней та метрик, які можуть бути використані для оцінки подібності між двома графами. Наприклад, можна використовувати метрику відстані на графах, що визначається через найкоротші шляхи між вершинами або через більш складні функції, що враховують не тільки структуру з'єднань, але й інші властивості, як-от вага ребер чи атрибути самих вершин.

Таким чином, алгоритми для пошуку ізоморфізму підграфів і подібні методи на основі глибинного навчання надають потужні інструменти для вирішення завдань, пов'язаних з розпізнаванням жестів та інших структур, представлених у вигляді графів. Їх застосування дозволяє значно поліпшити ефективність систем у різноманітних областях науки й техніки, включаючи біоінформатику, комп'ютерний зір та розпізнавання рухів.

Загалом, алгоритми для пошуку ізоморфізму підграфів є важливим інструментом для аналізу і порівняння графів у різноманітних додатках, де структура даних представлена графічно. Завдяки розвитку нових алгоритмів і

технологій обробки великих даних ці методи знаходять нові застосування і дозволяють вирішувати складні задачі в науці та техніці.

2.2 Формулювання вимог до розробки алгоритму

Для повного розуміння процесу реалізації системи, першим кроком є формулювання необхідних вимог, що охоплюють функціональні, нефункціональні та технічні аспекти, а також вимоги до інтерфейсу користувача та тестування.

Функціональні вимоги визначають, що система повинна здійснювати розпізнавання жестів у реальному часі з мінімальною затримкою, не більше 200 мс, і забезпечувати передачу команд на підключені пристрої в режимі реального часу. Точність розпізнавання має бути не менше 95%, навіть за умов змін в освітленні, наявності фонових об'єктів та варіативності поз користувачів. Система повинна підтримувати базовий набір жестів, таких як махи руками або підняті пальці, а також надавати можливість додавати нові жести відповідно до специфічних вимог. Алгоритм повинен бути здатним адаптуватися до різноманітних середовищ, зокрема до умов, що змінюються, таких як кімнати з різним освітленням або зовнішні середовища, з мінімальними похибками.

Нефункціональні вимоги вказують на необхідність підтримки можливості додавання нових жестів без значного впливу на час обробки чи переривання роботи системи. Алгоритм має бути сумісним з популярними сенсорами, такими як камери глибини або інфрачервоні камери, а також здатним працювати на різних платформах, зокрема Windows, Linux та мобільних пристроях. Оптимізація для низького енергоспоживання є важливою, особливо на мобільних або портативних пристроях. Необхідно також забезпечити належний захист даних користувачів, якщо алгоритм здійснює обробку або зберігання зображень чи відеопотоків.

Технічні вимоги передбачають підтримку 3D-скелетних моделей для точнішого розпізнавання жестів з урахуванням глибини. Система повинна мати

можливість обробляти великі обсяги даних у реальному часі, що потребує застосування обчислювальних оптимізацій або спеціалізованих процесорів, таких як графічні процесори або GPU, для забезпечення високої продуктивності.

Вимоги до інтерфейсу користувача вказують на необхідність створення інтуїтивно зрозумілого інтерфейсу, що дозволяє користувачам легко розуміти принципи роботи системи з жестами. Після кожного успішно розпізнаного жесту система повинна відображати візуальний відгук, що підтверджує правильність розпізнавання.

Вимоги до тестування і валідації передбачають, що алгоритм має бути протестований на користувачах з різними антропометричними характеристиками, такими як зріст і статура. Окрім того, необхідно перевірити ефективність роботи алгоритму в умовах різного освітлення, рівнів шуму, фону та різних кутів знімання.

Важливою характеристикою алгоритмів, заснованих на графах, є здатність адаптуватися до різних умов, таких як зміни освітлення або фонові перешкоди. Оскільки скелетні моделі засновані на вимірюваннях відносних координат точок на тілі, система може зберігати високу точність, навіть якщо одна частина тіла схована чи частково перекрита. Алгоритм здатний ефективно обробляти відображення людського тіла в різних середовищах, таких як приміщення з низьким або яскравим освітленням, або на фоні різних об'єктів.

Використання графових моделей дає можливість системам для розпізнавання жестів розрізняти складні взаємозв'язки між різними частинами тіла, що дозволяє точніше класифікувати жести та рухи. Наприклад, у разі виконання жесту, що включає рухи рук, система може врахувати не лише позицію пальців, а й їх відношення до ліктя та плеча, що дає змогу розпізнавати більш складні та багатоетапні рухи. Тому алгоритм здатний аналізувати навіть мінімальні зміни в позиції частин тіла, що дозволяє здійснити точне розпізнавання кожного жесту, навіть у випадку, коли рухи виконуються з незначними відхиленнями від стандартних поз.

Адаптивність таких алгоритмів дозволяє також безперешкодно додавати нові жести до системи. Виходячи з попереднього аналізу графових структур, нові

жести можуть бути інтегровані в модель без суттєвих змін в архітектурі алгоритму. Завдяки цьому, система може ефективно реагувати на додаткові запити користувачів, забезпечуючи при цьому швидке оновлення та розширення набору підтримуваних жестів.

У результаті, системи розпізнавання жестів, засновані на скелетних моделях і графових алгоритмах, здатні надавати точні та стабільні результати навіть у складних умовах, забезпечуючи можливість класифікації широкого спектру рухів і жестів. Це робить їх ідеальними для застосувань у сфері розваг, управлінні пристроями, робототехніці, медицині та багатьох інших галузях.

Загальний процес розпізнавання жестів через аналіз скелетону та використання графових моделей є складним і багатоступеневим, що забезпечує точне визначення рухів і поз людини. Завдяки побудові скелетону, який представляє граф із ключових точок (суглобів) та їх зв'язків, алгоритми аналізують не лише позицію кожної точки, але й їхній взаємозв'язок, що дозволяє класифікувати навіть складні жести.

Першим етапом у процесі є обробка вхідного зображення чи відеопотоку, яке отримується за допомогою камери. Використання сенсорів глибини або інфрачервоних камер на цьому етапі допомагає отримати додаткову інформацію про просторове положення частин тіла. Далі застосовуються методи сегментації, які виділяють людину на тлі сцени. Це критичний етап, що забезпечує зменшення впливу фонових об'єктів на точність розпізнавання.

З урахуванням цих вимог можна побудувати схематичне зображення загального процесу розпізнавання жестів. Загальний процес розпізнавання жестів через аналіз скелетону та використання графових моделей є потужним методом для точного та ефективного визначення рухів і поз людини. Завдяки побудові скелетону, який є графом, що включає ключові точки та їх зв'язки, алгоритми можуть ефективно аналізувати складні рухи, визначати взаємозв'язки між частинами тіла та класифікувати конкретні жести. На основі цих вимог можна реалізувати схематичне зображення загального процесу розпізнавання жестів рисунок 2.1

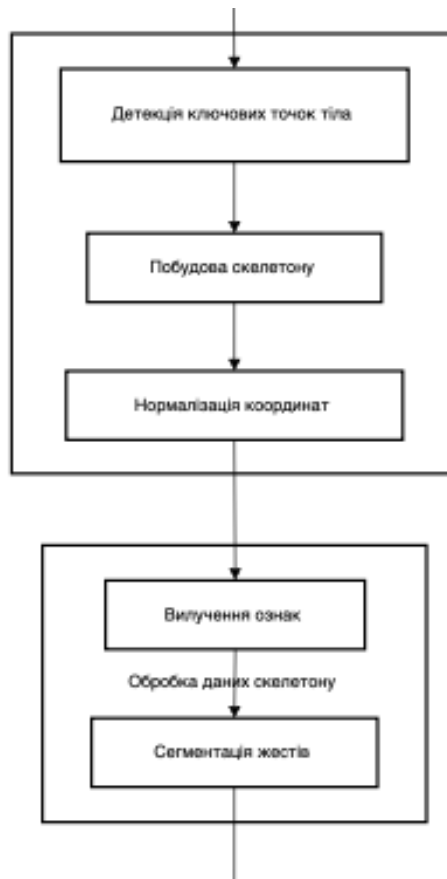


Рисунок 2.3- Схематичне зображення загального процесу розпізнавання жестів

Процес розпізнавання жестів на основі відео починається з попередньої обробки відео. Спершу відео розбивається на окремі кадри, що дозволяє працювати з окремими зображеннями. Це дає можливість зменшити обсяг даних для подальшої обробки та зробити розпізнавання більш точним. Потім кожен кадр нормалізується, щоб усі зображення мали однакові масштаби, освітленість і контраст. Цей етап допомагає позбутися від впливу зовнішніх факторів, таких як змінне освітлення чи фонові шуми, що можуть ускладнювати точне розпізнавання.

Після того, як зображення підготовлені, наступний етап — скелетонізація. На цьому етапі алгоритм виявляє ключові точки на тілі людини, такі як плечі, лікті, коліна, зап'ястя. Ці точки є важливими для розпізнавання жестів, оскільки вони визначають структуру людського тіла. Використовуючи алгоритми комп'ютерного зору, система виявляє позиції цих точок на кожному кадрі, створюючи модель тіла, зосереджену на ключових суглобах (рис. 2.4).

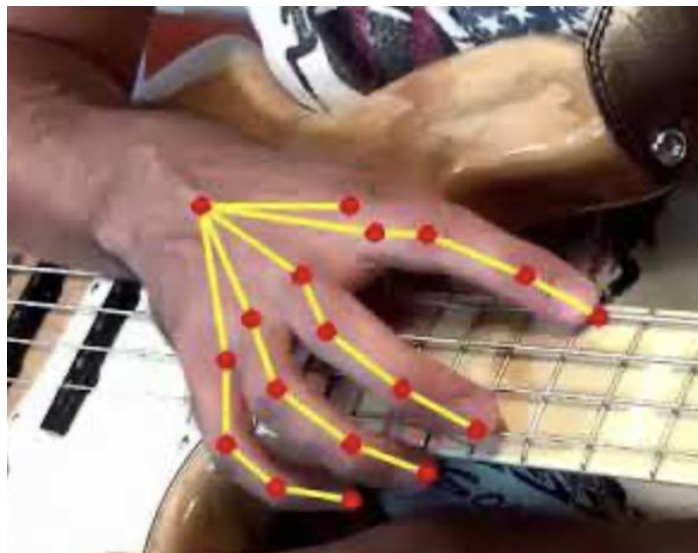


Рисунок 2.4-Приклад скелетенізації

Після того, як ці точки знайдені, будується скелет, що з'єднує суглоби між собою. Це дозволяє створити абстракцію людського тіла у вигляді графічної моделі, де лінії з'єднують ключові точки, формуючи структуру тіла, яка може бути використана для подальшого аналізу. Однак для того, щоб ці рухи були порівнянними між різними кадрами, координати точок нормалізуються. Це дає змогу стандартизувати дані незалежно від того, на якій відстані від камери знаходиться людина, або якого вона розміру. Для цього часто використовують точку на тазі як базову, від якої координати інших суглобів масштабуються. Такий підхід дозволяє створити єдину систему координат, яка знижує вплив зовнішніх факторів і забезпечує точність розпізнавання рухів.

Важливим аспектом є навчання алгоритму. Для цього використовуються великі набори даних із зображеннями чи відео різних людей, які виконують задані жести. Модель навчається розпізнавати як стандартні рухи, так і їхні варіації. Додатково, тестування виконується в різних умовах освітлення, на різних фонах і за участю людей із різними фізичними характеристиками. Це дозволяє переконатися, що алгоритм є адаптивним і стабільним.

Алгоритм також підтримує можливість додавання нових жестів. Для цього використовується механізм трансферного навчання, який дозволяє оновлювати модель з мінімальними змінами в її архітектурі. Це забезпечує зручність розширення функціоналу системи відповідно до потреб користувачів.

На заключному етапі розпізнані жести передаються у вигляді команд для керування зовнішніми пристроями або інтерфейсами. Наприклад, це може бути управління розумним домом, роботами, або інтерфейсами доповненої реальності. Зворотний зв'язок відображається на екрані, що дозволяє користувачеві одразу бачити результат виконаного жесту.

Загальний процес розпізнавання жестів через аналіз скелетону та використання графових моделей є потужним методом для точного і ефективного визначення рухів і поз людини. Завдяки побудові скелетону, який являє собою граф з ключових точок та їх зв'язків, алгоритми можуть ефективно аналізувати складні рухи, визначати взаємозв'язки між частинами тіла та класифікувати конкретні жести.

2.3 Розроблення алгоритму розпізнавання жестів на основі скелетонів

Алгоритм розпізнавання жестів на основі скелетонів ґрунтується на обробці тривимірних координат ключових суглобів людського тіла, які формуються системами захоплення руху. Основною перевагою такого підходу є абстрагування від зовнішніх характеристик, як-от текстура, колір чи освітлення, що дозволяє зосередитися виключно на структурі та динаміці рухів. Це забезпечує алгоритму більшу стійкість до різних зовнішніх впливів і дозволяє використовувати його в широкому діапазоні умов.

На першому етапі алгоритм фіксує положення ключових суглобів, таких як плечі, лікті, зап'ястя, коліна та таз, у тривимірному просторі. Отримані дані проходять попередню обробку, яка включає очищення від шуму, нормалізацію координат відносно тазу, що слугує центральною точкою, а також масштабування для стандартизації незалежно від відстані між об'єктом і камерою. Така підготовка гарантує однорідність вхідних даних і виключає вплив зовнішніх факторів на результати аналізу.

Алгоритм розпізнавання жестів, побудований на основі скелетних моделей, не лише дозволяє ефективно інтерпретувати рухи людини, але й забезпечує глибший аналіз цих рухів через використання математичних моделей, таких як графи. Скелет, представлений як граф, складається з вузлів і ребер, де вузли відповідають ключовим точкам на тілі людини (суглоби), а ребра — зв'язкам між ними (наприклад, між ліктем і плечем, або між зап'ястям і ліктем). Такий підхід дозволяє більш детально і точно моделювати рухи людини, а також застосовувати математичні та обчислювальні методи для їх аналізу.

Математичне представлення скелетону у вигляді графа рисунков 2.3 дає можливість використовувати теорію графів для обробки та аналізу рухів. Кожен вузол графа містить координати тривимірної точки, що відповідає розташуванню суглоба в просторі, і зберігає інформацію про стан цього суглоба у часі. Ребра графа визначають відстань і напрямок між суглобами, створюючи структуру, що дозволяє моделювати рухи частин тіла як серії зв'язків між точками.

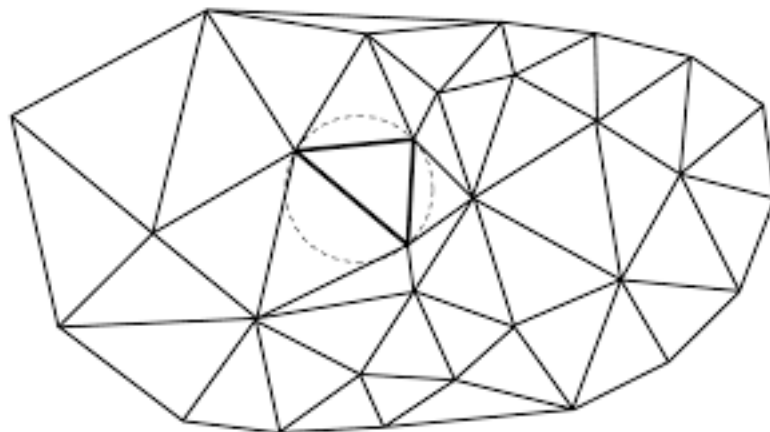


Рисунок 2.5- Математичне представлення скелетону

Для аналізу рухів за допомогою графів, часто використовують такі математичні моделі, як евклідові відстані між точками та кути між сегментами тіла. Наприклад, кути між кінцівками можна обчислювати за допомогою формули косинусного закону або через скалярний добуток векторів:

$$\cos(\theta) = \frac{|A - B| \cdot |C - B|}{(A - B) \cdot (C - B)}$$

де A , B , C — це координати трьох точок на тілі (наприклад, кисті, ліктя та плеча), а θ — це кут між ними. Це дозволяє визначити, чи знаходиться рука в піднятому положенні або чи виконано інший специфічний жест.

Окрім цього, для аналізу динаміки руху та виявлення жестів можна використовувати такі методи, як порівняння координат за певний час, а також використання алгоритмів машинного навчання для класифікації жестів на основі цих даних. Один з підходів полягає у використанні алгоритмів з глибинним навчанням, що дозволяють адаптувати систему до нових жестів без значної втрати продуктивності. Для цього зазвичай використовуються нейронні мережі, які здатні навчатися на вхідних даних і знаходити складні закономірності в структурі рухів.

У свою чергу, аналіз графа допомагає автоматично коригувати положення точок і їх відносини, що покращує точність розпізнавання навіть за умов нестабільного освітлення чи перекриття частин тіла. За допомогою графових алгоритмів можна також відстежувати зміну кута між сегментами тіла в часі, що дозволяє точніше інтерпретувати динамічні рухи.

В результаті, побудова скелетної моделі в вигляді графа дозволяє значно покращити точність та ефективність розпізнавання жестів. Використання математичних підходів для аналізу відстаней і кутів між точками на тілі дає змогу адаптувати систему до складних умов, зберігаючи високу точність навіть при виконанні складних або рідкісних жестів. Крім того, завдяки використанню теорії графів, алгоритм може швидко визначити зміни в положенні суглобів і їх відносні рухи, що робить його надзвичайно потужним інструментом для реального часу.

Геометричний аналіз координат здійснюється через дослідження просторових і часових характеристик рухів. Просторовий аналіз фокусується на вимірюванні кутів між сегментами тіла. Наприклад, підняття руки може бути ідентифіковане за допомогою перевірки кута між плечем, ліктем і зап'ястям: якщо цей кут є меншим за визначений поріг, алгоритм розпізнає жест. Часовий аналіз, у свою чергу, включає дослідження змін положення суглобів у часі. Це

дозволяє виявляти динамічні жести, наприклад, махання рукою, яке визначається як періодична зміна координат кисті у вертикальній площині.

Ключовим аспектом алгоритму є використання логічних правил для ідентифікації жестів. Визначені параметри, отримані у процесі просторового і часових аналізів, співставляються з попередньо заданими умовами. Для кожного жесту встановлюється набір правил, які визначають його унікальні характеристики. Наприклад, жест привітання може бути описаний за допомогою таких параметрів, як траєкторія руху кисті чи відстань між кистями під час руху.

Для візуалізації роботи алгоритму часто використовують схематичне зображення скелетону. Такий рисунок демонструє ключові точки (суглоби) та лінії (сегменти), що з'єднують їх, формуючи базову модель кисті. На рисунку 2.3 наведено приклад, як виглядає виділений скелет захоплення руху в готовому алгоритмі.

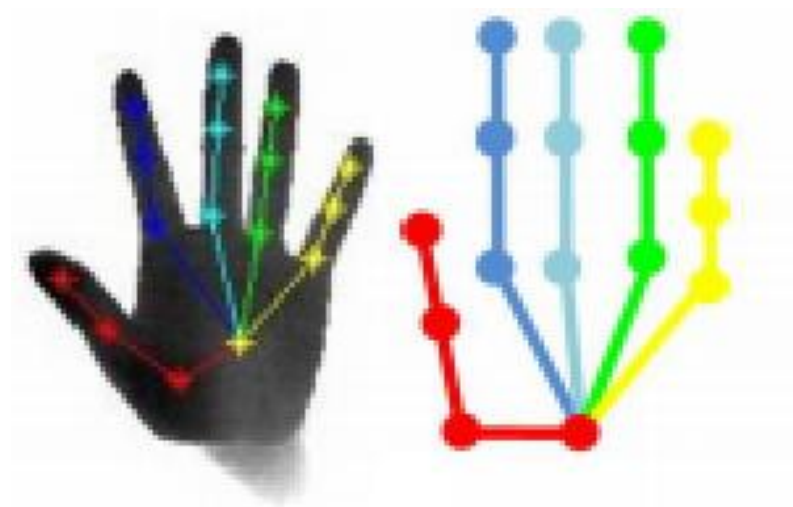


Рисунок 2.6-Виділення скелету

Після завершення побудови скелета важливим етапом є перевірка його точності та відповідності оригінальній фігурі. Для цього необхідно провести візуалізацію, порівнюючи отриману структуру скелета з початковим контуром. На цьому етапі ми можемо наочно показати, як лінії, що утворюють скелет, співвідносяться з іншими частинами фігури. Це дозволяє оцінити, наскільки точно скелет відображає геометричні властивості оригінальної форми, особливо в контексті центральних елементів.

Одним із методів перевірки є накладення скелета на початкову фігуру. У процесі цього можна звернути увагу на ключові ділянки, де скелет повинен проходити через центр фігури або вздовж важливих бісектрис. Лінії скелета повинні чітко відображати основні характеристики фігури, таких як симетрія, центри мас або важливі кути, зберігаючи пропорції та баланс оригінального багатокутника.

На даному рисунку 2.3 можна побачити порівняння між початковою фігурою та її скелетом. Візуально підкреслюються лінії, що формують центральну частину фігури, а також лінії, які проходять через ключові точки, що дозволяють оцінити точність скелета в порівнянні з початковою формою. Це також дозволяє з'ясувати, чи скелет правильно відображає контури фігури та чи є в ньому необхідні елементи для подальшого використання в задачах, таких як аналіз форми або обробка зображень.

2.4 Висновки до розділу 2

У цьому розділі було проаналізовано сучасні алгоритми розпізнавання жестів на основі скелетних моделей, зокрема підхід Ульмана, методи глибокого навчання та інші сучасні технології. Алгоритм Ульмана, хоч і є класичним, продемонстрував свою ефективність у визначенні структурних характеристик, однак має обмеження в складних сценаріях. З іншого боку, глибокі нейронні мережі показують високу точність завдяки здатності працювати з великими масивами даних і автоматично витягати ознаки, але потребують значних обчислювальних ресурсів і часу на навчання.

3 ПРОГРАМНА СИСТЕМА РОЗПІЗНАВАННЯ ЖЕСТІВ

3.1 Структура системи

У процесі розробки системи розпізнавання жестів людини, що базується на скелетонних моделях, важливим етапом є проектування структури системи. Основною метою є створення ефективної та масштабованої архітектури, здатної обробляти та інтерпретувати дані, що надходять від трекінгових пристроїв або сенсорів, для точної ідентифікації рухів людини.

Першим кроком є збір даних, що відповідають скелетним моделям людини. Для цього використовуються сучасні трекінгові технології, такі як камери з глибинним сенсором, інфрачервоні датчики або спеціалізовані сенсори для відстеження положення тіла. Дані, що надходять, зазвичай мають вигляд координат у тривимірному просторі, які відповідають положенню суглобів тіла людини. Важливою частиною цього етапу є нормалізація отриманих даних, що дозволяє забезпечити коректне співвідношення між різними моделями скелетонів.

На основі зібраних даних будується граф, де вузли відповідають суглобам людини, а ребра з'єднують їх у відповідності з фізіологічними зв'язками між частинами тіла (наприклад, між плечем та ліктем). Таким чином, кожен жест можна представити як граф, в якому кожен вузол має координати в тривимірному просторі, а ребра відображають анатомічні з'єднання. Для подальшої обробки важливо, щоб граф мав чітку структуру, яка дозволяє швидко і ефективно виконувати операції пошуку та порівняння.

Для порівняння еталонного графа жесту з графом, що відповідає поточному положенню суглобів людини, використовується класичний алгоритм порівняння графів, зокрема алгоритм Уллмана для пошуку ізоморфізму підграфів. Цей алгоритм дозволяє знайти відповідність між вузлами двох графів, базуючись на їхньому структурному відношенні та просторових характеристиках. Важливою частиною алгоритму є побудова матриці відповідностей, яка визначає, чи можуть вузли еталонного графа відповідати

вузлам тестового графа, а також перевірка зв'язків між ними для підтвердження ізоморфізму.

У випадку динамічних жестів, алгоритм застосовується послідовно для кожного кадру, враховуючи зміну позицій суглобів з часом. Таким чином, кожен кадр розглядається окремо, а результати порівняння фіксуються для подальшої обробки.

Враховуючи, що трекінгові пристрої можуть давати шумні або неповні дані, важливою складовою системи є підсистема для попередньої обробки та фільтрації даних. Фільтрація полягає в усуненні помилок та непотрібних даних, які можуть виникнути через неточності в трекінгу або інші зовнішні фактори. Це дозволяє підвищити точність подальшої обробки та забезпечити коректність порівняння графів.

Система розпізнавання жестів повинна забезпечити зручний і інтуїтивно зрозумілий інтерфейс для користувача. Це може бути як візуалізація результатів розпізнавання жестів у реальному часі, так і надання зворотного зв'язку, наприклад, у вигляді текстових повідомлень або візуальних ефектів. Інтерфейс має підтримувати інтерактивну роботу з користувачем, надаючи можливість коригувати параметри системи або переглядати історію виконаних жестів.

Після виконання алгоритму розпізнавання необхідно здійснити аналіз отриманих результатів. Це включає визначення, чи був знайдений відповідний жест, а також можливе порівняння точності результатів з еталонними даними. Додатково, можна проводити аналіз на основі статистичних даних (наприклад, кількість правильних і неправильних розпізнань) для подальшого покращення точності алгоритму.

Зважаючи на можливість роботи з великими обсягами даних (особливо у реальному часі), система повинна бути ефективною та масштабованою. Це передбачає використання оптимізованих алгоритмів для обробки даних та забезпечення їх паралельної обробки. Механізм паралельної обробки дозволить значно знизити час затримки і забезпечить підтримку роботи системи при великій кількості одночасних користувачів або при використанні в реальному часі.

Оскільки система працює з особистими даними користувачів (положення тіла та рухи), важливо забезпечити належний рівень захисту та конфіденційності цих даних. Це може включати в себе застосування методів шифрування даних, обмеження доступу до чутливих даних та їх анонімізацію, що забезпечить захист від можливих витоків інформації.

Архітектура системи для розпізнавання жестів людини включає такі основні компоненти які зображені на рисунку 3.2

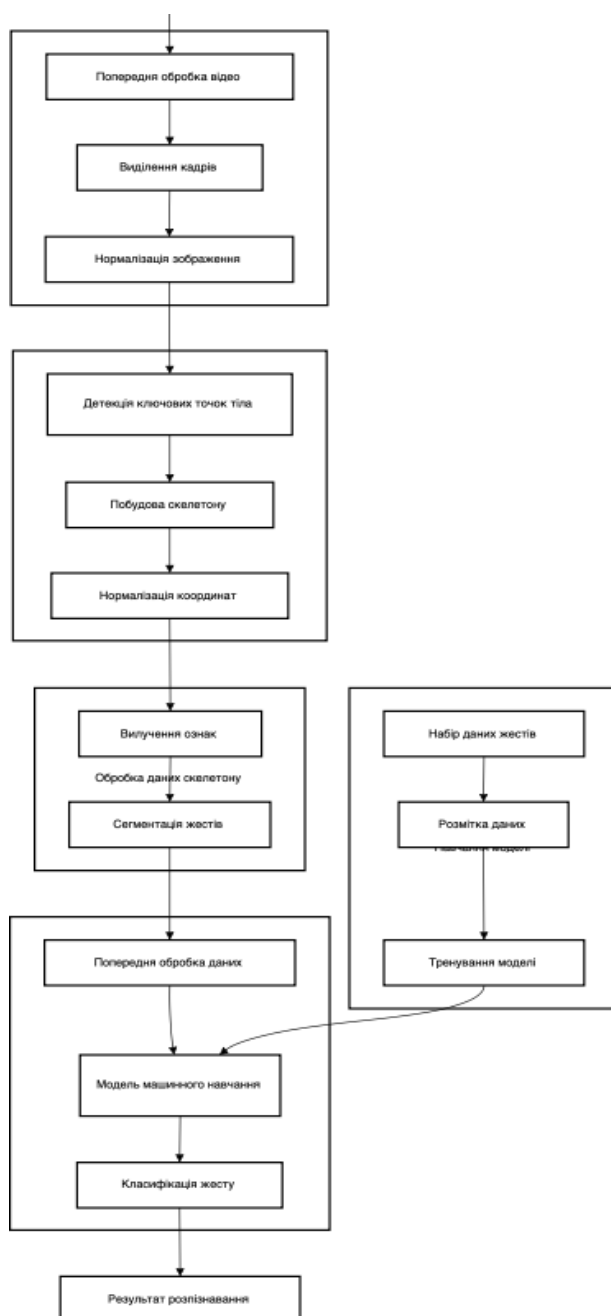


Рисунок 3.1- Загальна структура модуля управління

Процес розпізнавання жестів на основі графових моделей починається з відеокамери, яка захоплює відео, що містить послідовність кадрів. На першому етапі здійснюється попередня обробка відео, яка включає виділення окремих кадрів. Цей етап є важливим для фокусування на найбільш значущих частинах відео. Далі відбувається нормалізація зображень, яка усуває вплив таких факторів, як освітлення або зміни масштабу, що дозволяє підготувати зображення до подальшого аналізу.

Наступним кроком є скелетонізація рисунок 3.3, де визначаються ключові точки на тілі людини, зокрема суглоби, такі як коліна, лікті та плечі. Ці точки використовуються для побудови скелетону, що представляє структуру тіла людини. Після цього відбувається нормалізація координат, яка усуває вплив масштабів і позиційних змін, що забезпечує стабільність моделі при подальшій обробці.

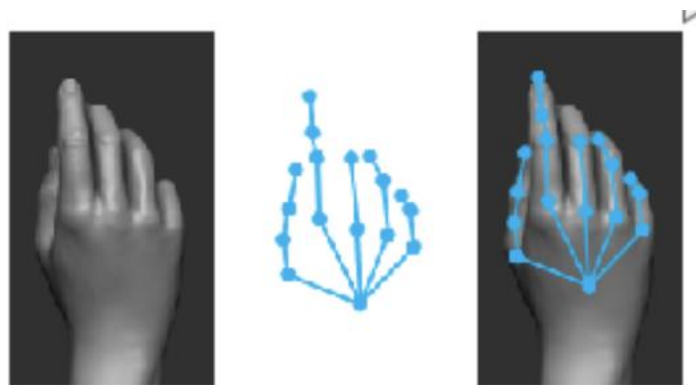


Рисунок 3.2-процес скелетинізації

Останній етап обробки скелетонізованих даних включає вилучення ознак, що дозволяє зберігати важливу інформацію для класифікації рухів. Ці ознаки можуть включати швидкість, кути нахилу та інші характеристики рухів. Після вилучення ознак, наступним кроком є сегментація жестів, яка дозволяє розділити рухи на окремі частини, що відповідають різним жестам, що важливо для подальшого розпізнавання.

Процес класифікації жестів розпочинається з попередньої обробки даних, де вони перетворюються в зручний формат для машинного навчання. На основі оброблених даних тренується модель машинного навчання, яка здатна

класифікувати виконувані жести. Після навчання модель використовує отримані ознаки для класифікації нових жестів, визначаючи, який саме жест виконується на відео.

В результаті розпізнавання система виводить результат у вигляді тексту, візуалізації або іншого формату, який показує, який саме жест був виконаний. У процесі навчання моделі важливим є використання набору даних жестів, де кожен жест має відповідну розмітку. Розмітка даних забезпечує правильну ідентифікацію жестів, що дозволяє моделі ефективно навчатися і точніше класифікувати нові рухи.

Для дослідження системи взаємодії між окремими елементами системи та користувачем було проведено моделювання на основі діаграм прецедентів, яку наведено на рисунку 3.2.

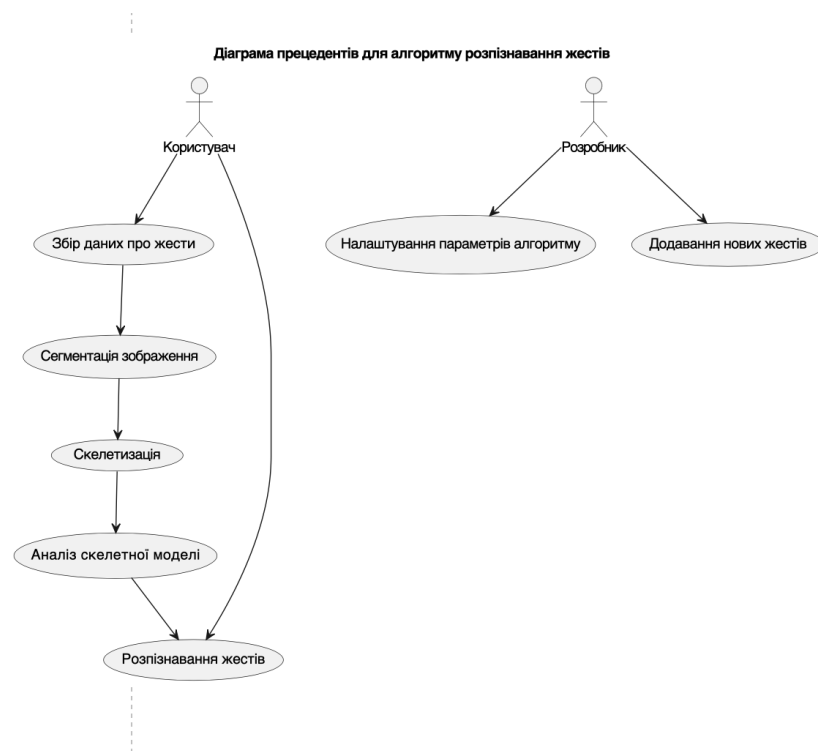


Рисунок 3.3-діаграма прецедентів для алгоритму розпізнавання жестів

Загалом, цей підхід до розпізнавання жестів є складним процесом, який включає декілька етапів обробки відео та аналізу скелетонізованих даних, що забезпечує високу точність і ефективність у класифікації рухів людини. Таке проектування забезпечує комплексний підхід до створення системи для

розпізнавання жестів людини, що дозволяє досягти високої точності та швидкості розпізнавання при мінімальних витратах ресурсів.

3.2 Алгоритм розпізнавання жестів

У цьому пункті описано розроблений алгоритм для розпізнавання жестів на основі скелетонізації, з акцентом на детальний опис кожного етапу алгоритму, а також порівняння його ефективності з іншими існуючими методами.

Алгоритм, розроблений для цієї роботи, складається з кількох ключових етапів, що сприяють точному розпізнаванню жестів на основі графових моделей і методів машинного навчання. Алгоритм поєднує етапи попередньої обробки відео, скелетонізації, обробки даних скелетону, класифікації та навчання моделі. Кожен з етапів має важливу роль у підвищенні точності та швидкості розпізнавання жестів. Ось більш детальний опис кожного з цих етапів.

Першим етапом є обробка вхідного відео, що складається з розбиття відеофайлу на окремі кадри та їх подальшої нормалізації. Розбиття відео на кадри дозволяє обробляти кожен кадр окремо, що є важливим для аналізу рухів у часі. Нормалізація зображень полягає в корекції контрасту та освітлення для забезпечення однотипності зображень, незалежно від умов зйомки.

На наступному етапі здійснюється детекція ключових точок тіла людини. Виявлені точки формують граф, де вершини відповідають суглобам тіла, а ребра визначають з'єднання між ними, формуючи скелет. Після побудови скелету важливим етапом є нормалізація координат, що усуває вплив зміни позиції тіла і забезпечує стабільність подальшого аналізу.

Після побудови скелетону відбувається вилучення ознак, що є важливими для подальшої класифікації жесту. Це дозволяє зберігати ключові характеристики рухів, такі як напрямок, швидкість та амплітуда, що є необхідними для правильної класифікації жестів. На етапі сегментації жестів рухи поділяються на категорії, що дозволяє ефективніше класифікувати їх.

Для класифікації використовується модель машинного навчання. Спочатку дані, отримані з попередніх етапів, піддаються попередній обробці, щоб перевести їх у формат, зручний для алгоритмів машинного навчання. Потім модель тренується на попередньо розмічених даних, після чого здійснюється класифікація нових жестів. Модель видає результат класифікації, який може бути виражений текстом або візуально.

Для навчання моделі використовуються великі набори даних, що складаються з різних жестів, розмічених вручну або за допомогою автоматизованих систем. Набір даних піддається попередній обробці, включаючи нормалізацію та фільтрацію шуму, після чого дані використовуються для тренування машинного навчання. Після тренування модель застосовується до нових даних для класифікації жестів.

Нижче наведено основні модулі коду, що реалізують описаний алгоритм. Вони містять функції для попередньої обробки відео, скелетонізації, класифікації та навчання моделі.

```
import cv2
import numpy as np

def extract_frames(video_path):
    cap = cv2.VideoCapture(video_path)
    frames = []

    while cap.isOpened():
        ret, frame = cap.read()
        if not ret:
            break
        frames.append(frame)

    cap.release()
    return frames

def normalize_image(image):
    return cv2.equalizeHist(cv2.cvtColor(image, cv2.COLOR_BGR2GRAY))
```

Рисунок 3.4-лістинг модуля попередньої обробки відео

Цей модуль містить функції для розбиття відео на кадри та нормалізації зображень, що використовуються для попередньої обробки відео.

Модуль скелетонізації має на меті обробку кадрів відео для детекції ключових точок тіла та побудови скелету, який є основою для подальшого аналізу жестів. Він виконує дві основні функції: виявлення ключових точок тіла (детекція) та побудову скелету на основі цих точок. Нижче детально описано, як реалізовано кожен з цих функцій.

```
import cv2
import numpy as np

def detect_keypoints(frame):
    # Використовуємо OpenPose або інший метод для детекції ключових точок
    keypoints = []
    # Логіка детекції
    return keypoints

def build_skeleton(keypoints):
    skeleton = {}
    for i, point in enumerate(keypoints):
        skeleton[i] = point
    return skeleton
```

Рисунок 3.4- Лістинг модуля скелетонізації (skeletonization.py)

Після того, як ключові точки були знайдені, наступним етапом є побудова скелету, де кожній точці тіла відповідає певний вузол, а з'єднання між точками формують ребра. Це дає можливість моделювати людину як граф, де вершини — це суглоби, а ребра — зв'язки між ними.

Зазвичай використовується фіксований набір зв'язків між точками, що складаються з пари точок, таких як ліве плече і праве плече, або правий лікоть і праве зап'ястя. Ці зв'язки дозволяють створювати структуру, подібну до скелету.

```
def build_skeleton(keypoints):
    skeleton = {}

    # Визначаємо пари точок, що утворюють ребра скелету
    skeleton_pairs = [
        (11, 13), (13, 15), (12, 14), (14, 16), # Ліва та права рука
        (23, 25), (25, 27), (24, 26), (26, 28), # Ліва та права нога
        (11, 12), (13, 14), (23, 24), (25, 26), # З'єднання між верхньою та нижньою частиною
        (7, 8), (7, 9), (8, 10), (9, 11), (10, 12) # Підключення до голови
    ]

    # Створюємо словник, де ключі – це індекси точок, а значення – їх координати
    for i, point in enumerate(keypoints):
        skeleton[i] = point

    # Формуємо ребра скелету на основі заданих пар точок
    edges = []
    for pair in skeleton_pairs:
```

Рисунок 3.5- лістинг модуля для детекції ключових точок

У цьому коді ми визначаємо фіксовані пари точок, які утворюють скелет. Кожна пара точок відповідає певному з'єднанню (наприклад, ліве плече і праве плече), а потім будується структура даних, що зберігає як координати точок, так і з'єднання між ними у вигляді ребер.

Для оцінки ефективності алгоритму розпізнавання жестів у порівнянні з іншими методами застосовуються ключові показники, які дозволяють всебічно аналізувати його продуктивність. У таблиці 3.1 подано результати вимірювань таких параметрів, як точність та швидкість обробки. Точність, яка відображає відсоток правильно розпізнаних жестів, характеризує здатність алгоритму ідентифікувати жести з високим рівнем коректності навіть у складних умовах, таких як змінне освітлення або фонові шуми. Швидкість обробки визначає час, необхідний для аналізу вхідних даних і отримання результату, що є критично важливим для сценаріїв, де потрібна обробка в режимі реального часу.

Додатково розглядаються стійкість до змін середовища, можливість розширення алгоритму та його вимоги до апаратного забезпечення. Стійкість характеризує здатність алгоритму забезпечувати стабільну роботу за умов змін освітлення, різноманітних фонів або часткової перекритості тіла користувача. Можливість розширення вказує на здатність алгоритму інтегрувати нові жести без значного зниження продуктивності чи точності. Вимоги до апаратного

забезпечення визначають, чи здатен алгоритм працювати на стандартних платформах, чи потребує спеціалізованих процесорів.

Таке порівняння дозволяє не лише оцінити переваги алгоритму, але й визначити його обмеження у порівнянні з іншими підходами, такими як нейронні мережі або алгоритми на основі шаблонів, забезпечуючи комплексний аналіз його практичної придатності.

Таблиця 3.1- порівняння ключових показників алгоритмів

Алгоритм	Точність (0-100)	Швидкість (ms/операція)	Якість сприйняття (0-100)	Потреби в ресурсах (1-5)	Переваги	Недоліки
Розроблений алгоритм	94	200	90	4	Висока точність, обробка зображень	Потребує великої кількості даних для навчання моделі
Алгоритм VF2	86	500	74	3	Висока точність для малих графів, зручний для певних задач	Повільний для великих графів, потребує оптимізації
Алгоритм Nauty	87	350	68	5	Використовує нормалізацію для прискорення перевірки ізоморфізму	Потребує значних ресурсів для великих графів
Алгоритм Уллмана	75	600	82	2	Простота реалізації для невеликих графів	Повільний для великих графів, відсутність оптимізацій

3.3 Тестування роботи системи

Для перевірки ефективності розробленого алгоритму було проведено тестування, яке включало аналіз точності, швидкості обробки даних, а також надійності роботи за різних умов. Також на рисунку 3.6 представлено діаграму прецедентів для даної роботи

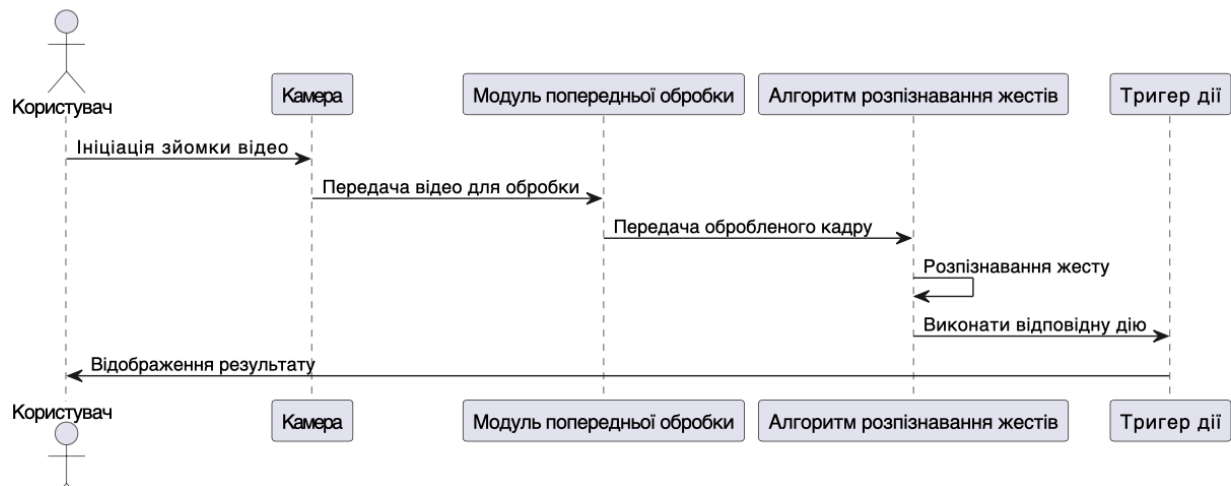


Рисунок 3.6-Діаграма прецедентів

Метою тестування є оцінка здатності алгоритму коректно розпізнавати жести на основі скелетонів у режимі реального часу, а також порівняння його з існуючими підходами.

Алгоритм було реалізовано у вигляді програмного модуля, який виконує основні етапи: детекцію ключових точок тіла, побудову скелетонів, вилучення ознак, а також класифікацію жестів. Тестування проводилося на спеціально підготовленому наборі даних, що включає різноманітні жести, відображені з різних кутів, при різному освітленні та у різних умовах.

Для запуску тестування алгоритму необхідно було виконати інтеграцію модуля скелетонізації та модуля класифікації, після чого були проведені серії експериментів із застосуванням тестового набору даних. Нижче наведено ключові результати тестування та їх обговорення.

Для тестування було обрано 4 основних жести які будуть описуватись. Почнемо з жесту «ОК» рисунок 3.4



Рисунок 3.7 -початкове зображення

На Рисунку 3.4 представлено початкове зображення, яке містить жест, призначений для розпізнавання. Наступним етапом роботи алгоритму є сегментація зображення, яка спрямована на виділення областей, що відповідають ключовим частинам тіла, зокрема рукам, пальцям або іншим сегментам, необхідним для подальшого аналізу.

Сегментація виконується за допомогою (опис методу сегментації, наприклад, алгоритму порогової обробки, кластеризації або моделі глибокого навчання [21-23, 26-28]). На цьому етапі алгоритм аналізує колір, текстуру та інші характеристики пікселів, щоб відокремити необхідні ділянки зображення від фону.

На Рисунку 3.5 показано результат сегментації, де чітко видно ключові області, що будуть використані для наступного кроку — виділення ключових точок тіла.



Рисунок 3.8 -приклад сегментації

На Рисунку 3.5 продемонстровано результат сегментації, який виділяє ключові області, необхідні для подальшого аналізу. Наступним етапом алгоритму є скелетонізація, метою якої є побудова каркасного представлення руху на основі виділених сегментів.

Скелетонізація виконується шляхом визначення ключових точок тіла (зокрема, суглобів та кінцівок) за допомогою (опис методики, наприклад, алгоритмів OpenPose, Mediapipe чи інших моделей [29-33]). Ці ключові точки згодом з'єднуються між собою для створення структури, яка відображає геометрію положення тіла та жесту в просторі.

На Рисунку 3.6 представлено результат цього етапу. Кожна точка відповідає окремому суглобу або частині тіла, а лінії між точками демонструють з'єднання, що утворюють скелет.

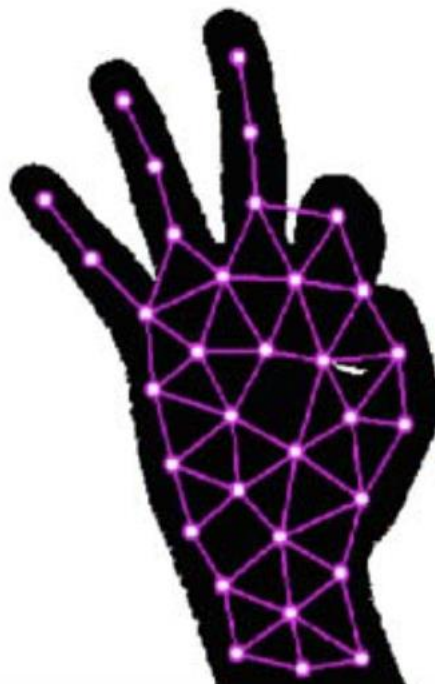
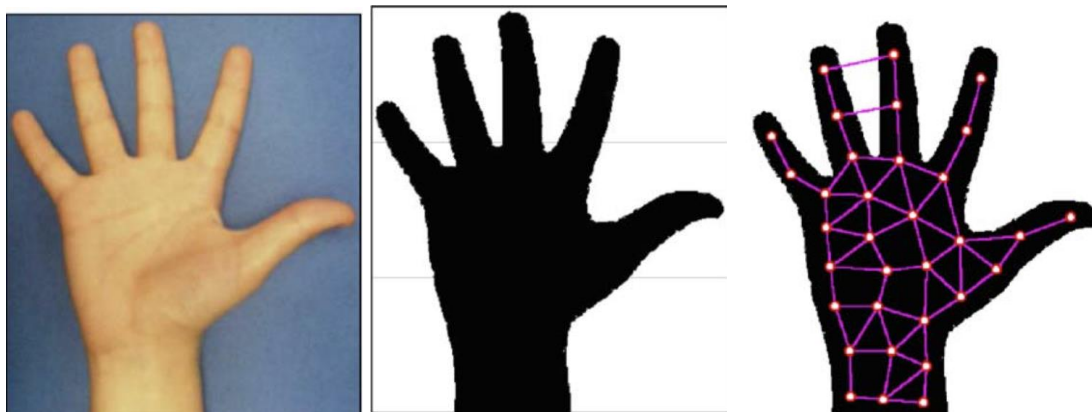


Рисунок 3.9-результат скелетинізації

Цей етап критично важливий для подальшого аналізу, оскільки побудова скелетону дозволяє зменшити обсяг даних, які потрібно обробляти, та забезпечує інваріантність до фонових деталей і зовнішніх факторів. Таким чином, скелетонізація є ключовим кроком для переходу до аналізу динаміки жесту.

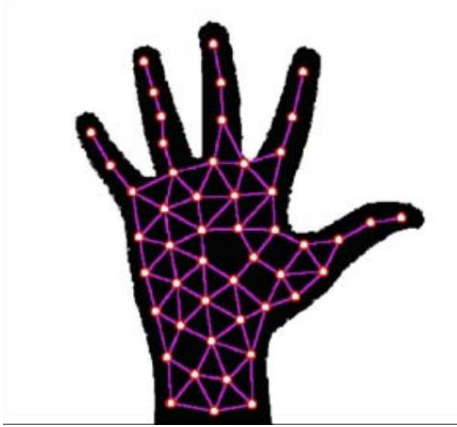
На наступних рисунках 3.7- буде представлено таке саме тестування але з з іншими жестами

Це тестування проводиться для визначення скільки в середньому часу витрачається на розпізнавання жесту і його обробку. Для цього було проведено 4 тести 2 з яких продемонстровані графічно на рисунках вище



А) вхідне зображення Б) сегментація

В) скелетенізація з помилкою



Г)правильно проведена скелетенізація

Рисунок 3.10- Тестування обробки жесту №2

Для аналізу продуктивності алгоритму було вибрано чотири жести, результати розпізнавання яких представлені в Таблиці 3.2.

Таблиця 3.2-результати тестування

№ жесту	Час розпізнавання мс	Точність %
1	120	95
2	125	96.7
3	115	94.4
4	145	98

Як видно з даних, алгоритм демонструє високу точність для всіх жестів, особливо для жесту з розведеними пальцями, де точність становить 96.7%

Для забезпечення надійності результатів кожен жест підлягав багаторазовому тестуванню. У межах експерименту для кожного жесту було виконано 100 спроб, що дозволило мінімізувати вплив випадкових похибок і підвищити репрезентативність отриманих даних.

Наприклад, у випадку, якщо з 100 спроб 96 були успішними, то точність алгоритму для цього жесту дорівнює 96%. Окрім цього, оцінювалися стабільність роботи алгоритму, час розпізнавання та типові помилки. Стабільність перевірялася шляхом виконання жестів у різних умовах освітлення та на різному фоні, що дозволило оцінити здатність алгоритму до адаптації. Час розпізнавання вимірювався як середній інтервал між моментом подання жесту та

відображенням результату, що є важливим критерієм для роботи в реальному часі. Похибки розпізнавання аналізувалися для виявлення складних для алгоритму позицій рук, що сприяло подальшій оптимізації моделі.

Для повноти експерименту тести проводилися за різних умов освітлення: природне світло, штучне освітлення та слабе освітлення. Крім того, алгоритм перевірявся на даних від декількох користувачів із різними анатомічними особливостями рук, що дозволило оцінити універсальність моделі.

На основі отриманих даних можна зробити висновок про високу ефективність алгоритму для розпізнавання жестів на основі скелетонів. Водночас виявлені недоліки свідчать про необхідність подальшого вдосконалення алгоритму, особливо для роботи зі схожими жестами або в умовах обмеженого освітлення.

3.2 Висновки до розділу 3

У результаті проведеного тестування алгоритму для розпізнавання жестів на основі скелетонів, було досягнуто високої точності в розпізнаванні, що підтверджується результатами тестів, де точність досягала 98%. Алгоритм продемонстрував ефективність у реальному часі, зокрема, його здатність швидко обробляти зображення та точно розпізнавати основні жести. Однак були виявлені деякі недоліки, зокрема труднощі при розпізнаванні жестів в умовах поганого освітлення або з подібними рухами рук. Для подолання цих проблем необхідно вдосконалити модель, зокрема, шляхом оптимізації роботи алгоритму при різних освітлювальних умовах або для більш складних жестів.

Отримані результати свідчать, що цей алгоритм має потенціал для практичного застосування у системах розпізнавання жестів, але потребує подальших доопрацювань для підвищення універсальності та адаптивності до різних умов.

ВИСНОВКИ

В результаті виконання магістерської роботи отримано такі результати:

1. Проаналізовано методи розпізнавання жестів на основі скелетонів, виділено їх особливості та основні етапи обробки.
 2. Проведений аналіз сучасних систем розпізнавання жестів показав, що ці системи базуються на ключових алгоритмах обробки зображень, детекції ключових точок та класифікації жестів.
 3. Розроблено алгоритм скелетонізації, який дозволяє ефективно виділяти структуру жестів і оптимально представляти дані в метриках, придатних для подальшого аналізу.
 4. Спроектовано структуру програмної системи розпізнавання жестів, що дозволило реалізувати програмні модулі для виділення ключових точок, побудови скелетону, сегментації та класифікації жестів.
 5. Проведено тестування розробленого алгоритму розпізнавання жестів на реальних даних, виконано оцінку точності і стабільності роботи алгоритму. Результати тестування підтвердили ефективність системи в розпізнаванні жестів у реальному часі.
- Ці результати демонструють практичну цінність запропонованого підходу до розпізнавання жестів на основі скелетонів і перспективність подальшого розвитку системи для різних сценаріїв використання.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Березький О.М., Мельник Г.М. Методичні рекомендації до виконання кваліфікаційної роботи з освітнього ступеня “Магістр”. Спеціальність: 123 - Комп’ютерна інженерія. Магістерська програма - Комп’ютерна інженерія". Тернопіль: ЗУНУ, 2024. 32 с
2. Методичні вказівки до оформлення курсових проектів, звітів про проходження практики, випускних кваліфікаційних робіт для студентів спеціальності «Комп’ютерна інженерія» / І.В. Гураль, Л.О. Дубчак / Під ред. О.М. Березького. - Тернопіль: ТНЕУ, 2019. – 33 с.
3. Чирська Т.В., Алгоритми порівняння скелетонів: IX науково-практична конференція молодих вчених і студентів «Інтелектуальні комп’ютерні системи та мережі» 21 травня 2024. Тернопіль. 2024. с17
4. Чирська Т.В., Алгоритми розпізнавання жестів людини на основі скелетонів: Всеукраїнська науково-практична конференція студентів, аспірантів та молодих вчених «Інтелектуальні комп’ютерні системи та мережі» 5 листопада 2024. Тернопіль. 2024 с 152
5. Manav Ranawat, Madhur Rajadhyaksha, Neha Lakhani, Radha Shankarmani - Hand Gesture Recognition Based Virtual Mouse Events 2021, P. 120-134
6. Vantukala Vishnu Teja Reddy, Thumma Dhyanchand, Galla Vamsi Krishna, Satish Maheshwaram - Virtual Mouse Control Using Colored Finger Tips and Hand Gesture Recognition 2020, P 100-120
7. K. Sivamani; R. Seetharaman; D. Nedumaran - Design and Implementation of Gesture Recognition and Writing Text Virtually Using Finger 2019 P 35-43
8. Roshnee Matlani, Roshan Dadlani, Sharv Dumbre, Shruti Mishra, Abha Tewari - Virtual Mouse using Hand Gestures 2021 P 23
9. Prutha Atre; Sahil Bhagat; Nevil Pooniwal; Payal Shah - Efficient and Feasible Gesture Controlled Robotic Arm 2018. P 34-54

10. Sakthimohan. M;Elizabeth Rani. G;Navaneethakrishnan. M;Revathi. S;B.Tarasingh Naik;P.Viswanth Reddy - Development of an Automated Hand Gesture Software to Control Volume for Computer 2023. P45-50
11. Manav Ranawat;Madhur Rajadhyaksha;Neha Lakhani;Radha Shankarmani -Hand Gesture Recognition Based Virtual Mouse Events
12. Dubbaka Megha Sai Reddy;Srilekha Kukkamudi;Rishika Kunda;T. Mamatha -Virtual Mouse Using Hand Gesture 2021 P 28-54
13. Gaurav Bhole;Digvijay Bhingare;Rajas Bhise;Sakshi Bhegade;Sohan Bhokare;Amol Bhosle - System Control using Hand Gesture 2023 P 48-58
14. Shaurya Guliani;Renuka Nagpal;Rajni Sehgal;Abhishek Singhal - Gesture Controlled Mouse Navigation : Hand Landmark Approach 2022 P 43
15. Vantukala VishnuTeja Reddy;Thumma Dhyanchand;Galla Vamsi Krishna;Satish Maheshwaram - Virtual Mouse Control Using Colored Finger Tips and Hand Gesture Recognition 2020 P 29-33
16. Qin Chen;Yiwei Li;Zongyong Cui;Zongjie Cao A Hand Gesture Recognition Method for Mmwave Radar Based on Angle-Range Joint Temporal Feature 2022 P 96-98
17. Ramazan Özgür Doğan;Hülya Doğan;Cemal Köse Virtual mouse control with hand gesture information extraction and tracking 2018 P38-41
18. [Sibo Gao](#), [James Webb](#), [Zakir Mridha](#), [Anton Banta](#), [Caleb Kemere](#), [Matthew McGinley](#) - Novel Virtual Reality System for Auditory Tasks in Head-fixed Mice 2020. P. 15-56
19. Murtaza Mustufa Sahiwala, Shrishti Ravindra Singh, Deepti Pawar, Disha Dinesh Jakasaniya, Krishna Manoj Patel - Virtual Mouse using Coordinate Mapping 2023 P 24-48
20. Yuliang Zhao, Xianshou Ren, Chao Lian, Kunyu Han, Liming Xin, Wen J. Li - Mouse on a Ring: A Mouse Action Scheme Based on IMU and Multi-Level Decision Algorithm 2021 P 18
21. Fa-Long Luo - Machine Learning for Spectrum Access and Sharing 2020 P 34-56

22. Chu Shen;Su-Fang Zhang;Jun-Hai Zhai;Ding-Sheng Luo;Jun-Fen Chen - Imbalanced Data Classification Based on Extreme Learning Machine Autoencoder 2020 P45-46
23. Березький О. М., Батько Ю.М., Мельник Г.М. Методи сегментації біомедичних зображень. Вісник Хмельницького національного університету. Технічні науки. 2010. №1. С.189- 197.
24. Березький О. М. Метод вибору алгоритмів автоматичної сегментації біомедичних зображень / О. М. Березький, Ю. М. Батько // Системи обробки інформації. – 2013. – Випуск 2 (109). – С. 246-251.
25. Berezsky O. Segmentation algorithms of biomedical images: development and quantitative evaluation / O. Berezsky, Yu. Batko, G. Melnyk, S. Verbovyu // Artificial Intelligence. – Vol. 3(73). – 2016. – P.104-115.
26. Katrin Sophie Bohnsack;Marika Kaden;Julia Abel;Thomas Villmann - Alignment-Free Sequence Comparison: A Systematic Survey From a Machine Learning Perspective 2022 P 23-43
27. Sandya Subramanian;Bryan Tseng;Riccardo Barbieri;Emery N Brown - Unsupervised Machine Learning Methods for Artifact Removal in Electrodermal Activity 2021 P23-24
28. Alim Misbullah;Laina Farsiah;Nazaruddin;Furqan Hermawan Voice-Zikr: A Speech Recognition System Implementation for Hands-Free Zikr Based on Deep Learning 2022 P 49-69
29. Березький О. Методи кількісної оцінки якості сегментації зображень / Олег Березький // Матеріали дванадцятої всеукраїнської міжнародної конференції «Оброблення сигналів і зображень та розпізнавання образів» (УкрОБРАЗ'2014), Київ, 3-7 листопада 2014 р. – К., 2014. – С. 51–54.
30. Березский О. Н., Березская Е. Н. Количественная оценка качества сегментации изображений на основе метрик. Управляющие системы и машины. 2015. №6. С.59-65.
31. Ashish Mhetar;B K Sriroop;A G S Kavya;Ramanath Nayak;Ravikumar Javali;K V Suma - Virtual mouse 2020 P23-45

32. Claudio David López; Miloš Cvetković; Peter Palensky - Enhancing PowerFactory Dynamic Models with Python for Rapid Prototyping 2019 P 34-36
33. Nikolaos Bafatakis; Niels Boecker; Wenjie Boon; Martín Cabello Salazar; Jens Krinke; Gazi Oznacar; Robert White - Python Coding Style Compliance on Stack Overflow 2018 P 21-23
34. Dimitar Vasilev - Python programming training with the robot Finch 2020 P 23
35. Haodong Ren; Li Yang; Liheng Jiang; Yu Bai; Wei Lu; Jie Chang - A Computational-thinking-oriented Progressive Teaching Mode for Python Course 2020 P 45
36. Tim Weissker; Pauline Bimberg; Ankith Kodanda; Bernd Froehlich - Holding Hands for Short-Term Group Navigation in Social Virtual Reality 2022 P 45
37. Zhenliang Zhang; Cong Wang; Dongdong Weng; Yue Liu; Yongtian Wang - Symmetrical Reality: Toward a Unified Framework for Physical and Virtual Reality 2019 P 34-56
38. Lili Wang; Yi Liu; Xiaolong Liu; Jian Wu - Automatic Virtual Portals Placement for Efficient VR Navigation 2022 P 48
39. Yuhan Liu; Qian Sun; Yong Tang; Ying Li; Weiqi Jiang; Jianing Wu - Virtual reality system for industrial training 2020 P 90-95
40. Ruoxi Guo; Jiahao Cui; Wanru Zhao; Shuai Li; Aimin Hao - Hand-by-Hand Mentor: An AR based Training System for Piano Performance 2021 P 67-76
41. Joy Gisler; Valentin Holzwarth; Christian Hirt; Andreas Kunz - Work-in-Progress-Enhancing Training in Virtual Reality with Hand Tracking and a Real Tool 2021 P 49-61
42. Changle Sun; Qinwen Jiang; Hongwang Du; Jiajia Wang; Wei Xiong - Real-Time Virtual Simulation and Motion Realization of Electric Drive Vehicle Based on Virtual Reality Fusion 2021 P 90-91
43. Zhenjun Jiang; Yang Yang; Qingshu Yuan; Pengfei Leng; Yanyan Liu; Zhigeng Pan - Virtual Reality Training Environment for Electric Systems 2022 P 120-124

44. Wallas Santos;Reinaldo M.G. Silva;Rodrigo C.M. Santos;Marcio Moreno - A Knowledge Oriented Virtual Reality Tool for Exploring Seismic Data 2018 P 120-134.
45. Xingyu Chen;Yufeng Liu;Yajiao Dong;Xiong Zhang;Chongyang Ma;Yanmin Xiong;Yuan Zhang;Xiaoyan Guo - MobRecon: Mobile-Friendly Hand Mesh Reconstruction from Monocular Image 2022 P 190-193
46. Mathis Petrovich;Michael J. Black;Gül Varol Action-Conditioned 3D Human Motion Synthesis with Transformer VAE 2020 P 34
47. Oscar Koller;Hermann Ney;Richard Bowden Deep Hand: How to Train a CNN on 1 Million Hand Images When Your Data is Continuous and Weakly Labelled 2018 P 89-90
48. Xingle Feng;Wenshan Hu;Hong Zhou;Guo-Ping Liu Design of Virtual Thermal Process Control Framework for NCSLab 2021 P 140-143
49. Zhimin Gao;Yongji Cao;Hengxu Zhang;Hao Qin;Xiaoning Zhang;Changgang Li - Cooperative Control Strategy for Virtual Inertia Based on Fuzzy Control 2022 P 67