

## ДОДАТОК А

### Текст тез та наукових статей

#### 1. V Міжнародна науково-практична конференція «Perspectives of science and education»

Сучасний рівень розвитку інформаційних технологій, алгоритмів і методів сприяє появі нових інтерфейсів взаємодії користувачів та пристроїв. Перспективним напрямком розвитку технологій є розробка і дослідження людино-машинних інтерфейсів, заснованих на розпізнаванні об'єктів. Одним із завдань розширення взаємодії людини з комп'ютером є розробка загальної методології виявлення і розпізнавання динамічних жестів людини, зокрема – жестів рук.

В останній час розробники інтелектуальних систем все більше уваги стали приділяти автоматичному розпізнаванню людських жестів за допомогою візуальних систем. Такий інтерес викликаний природним характером і зручністю використання інтерфейсів на основі жестів, а також можливістю їх застосування в більшості областей людської діяльності. Постановка завдання розпізнавання жестів комплексна і враховує неоднозначну природу статичних і динамічних жестів, проблеми виділення руки на навколишньому фоні, умови освітлення і «шуми» на відео- або фото-матеріалі. Вирішення задачі в більшості випадків передбачає вибір алгоритмів з використанням комп'ютерних ресурсів.

Залежно від типу вхідних даних підхід для інтерпретації жестів може бути виконаний різними шляхами. Однак більшість методів засновані на ключових показниках, представлених в тривимірній системі координат. Грунтуючись на їх відносному русі, жест може бути виявлений з високою точністю, в залежності від якості вхідних даних і підходу алгоритму.

Щоб інтерпретувати рух тіла, потрібно класифікувати його за загальними властивостями і повідомленнями, які можуть висловлювати рухи. Наприклад, в мові жестів кожен жест являє собою слово або фразу.

У деяких літературних джерелах відрізняють два різних підходи розпізнавання жестів: на основі 3D-моделі і на основі зовнішнього вигляду. Передові методи для розпізнавання жестів використовують тривимірну інформацію про ключові елементи частин тіла, щоб отримати кілька важливих параметрів, таких як положення долоні або кути суглоба. З іншого боку, системи на основі зовнішнього вигляду використовують зображення або відео для прямої інтерпретації.

Підхід 3D-моделі може використовувати об'ємні або скелетні моделі або навіть їх комбінацію. Об'ємні підходи широко використовувалися в індустрії комп'ютерної анімації та для цілей комп'ютерного зору. Моделі зазвичай створюються з складних тривимірних поверхонь, таких як NURBS або полігональні сітки.

Замість інтенсивної обробки тривимірних моделей і роботи з великою кількістю параметрів можна просто використовувати спрощену версію параметрів кутів з'єднання разом з довжинами сегментів. Цей метод відомий як метод скелетного уявлення тіла, де обчислюється віртуальний скелет людини, і частини тіла відображаються на певні сегменти. Аналіз тут виконується з використанням положення і орієнтації цих сегментів і відносин між кожними з них.

Моделі на основі 2D відображення об'єкта не використовують просторове уявлення тіла, тому що вони отримують параметри безпосередньо із зображень або відео, використовуючи базу даних шаблонів. Деякі з них засновані на деформуємих двомірних шаблонах людських частин тіла, особливо рук. Деформуємі шаблони – це набори точок на контурі об'єкта, які використовуються в якості вузлів інтерполяції для апроксимації контуру об'єкта. Одна з найпростіших інтерполяційних функцій – лінійна, яка виконує усереднення форми за наборами точок, параметрами мінливості точок і

зовнішніми деформаторами. Ці засновані на шаблонах моделі в основному використовуються для відстеження рук, але можуть також використовуватися для простої класифікації жестів.

Головний недолік методів на основі просторового уявлення тіла полягає в тому, що вони вимагають значних обчислювальних ресурсів, а також для роботи з ними необхідна камера з можливістю визначення глибини зображення. Тому для невеликих прикладних програм раціональніше використовувати методи, засновані на 2D поданні об'єкта, так як в них знижується обчислювальна складність, а також відпадає необхідність в спеціальному обладнанні: для отримання зображення може бути використана звичайна камера мобільного пристрою. Оскільки 2D методи можуть бути менш точними, перспективним напрямком досліджень буде їх удосконалення для отримання не ресурсоємного і ефективного алгоритму розпізнавання жестів рук.

## 2. II Всеукраїнська науково-практична конференція молодих учених, курсантів та студентів «Авіація, Промисловість, Суспільство»

Вираження емоцій і розуміння психологічного стану людини є одним з ключових факторів взаємодії між людьми. Тому при створенні сучасних людино-машинних систем актуально застосування методів автоматичного розпізнавання емоцій. За останні кілька десятиліть отримали розвиток методи комп'ютерного розпізнавання емоцій, що дозволяють підвищити рівень взаємодії між людиною та комп'ютером, а також отримувати інформацію про відношення людини до предметів або подій.

Одним з основних методів розпізнавання емоцій людиною іншою людиною є аналіз візуальної інформації. Тому автоматизація цього процесу очевидно повинна бути заснована на використанні методів і засобів комп'ютерного зору.

Комп'ютерне зір є науковою областю, в рамках якої ведуться дослідження

по вивченню теорії і фундаментальних алгоритмів аналізу зображень, об'єктів і сцен. Часто також замість поняття "Комп'ютерний зір" використовують "Машинний зір" або "Технічний зір". Однак останні поняття відносяться до більш загальної науково-практичної області, що охоплює всі етапи розробки систем, які базуються на обробці і аналізі відеоінформації. Тут розглядаються:

- схеми освітлення об'єктів і сцен;
- характеристики датчиків, їх кількість, розташування, калібрування і орієнтування;
- засоби обробки зображень;
- алгоритми обробки і їх реалізація.

Задача розпізнавання емоцій може вирішуватися в системах, що застосовуються в різних сферах людської діяльності.

Більшість методів комп'ютерного розпізнавання емоцій засновані на навчанні з учителем. В якості вхідних даних в них використовуються або окремі зображення обличчя, або послідовність кадрів з відеопотоку, в зв'язку з чим методи можна умовно розділити на статичні і динамічні. Динамічні методи, як правило, використовують інформацію про рух лицьових м'язів з плином часу. Так, наприклад, один з найперших методів автоматичного розпізнавання емоцій використав оптичні потоки для визначення напрямку і швидкості руху окремих частин обличчя, які потім зіставлялися із зразками для кожної з розглянутих емоцій. Аналогічний підхід зустрічається і в більш пізніх роботах, однак замість оптичного потоку частіше використовується модель з'єднаних вібрацій, що дозволяє більш точно визначити деформації частин обличчя.

Одним з головних переваг динамічних методів є те, що вони дозволяють вловити короткочасні зміни на людському обличчі – так звані мікроемоції. Однак у багатьох практичних завданнях (наприклад, розпізнавання по одному зображенню) тимчасова інформація відсутня. Тому також часто використовуються статичні методи, засновані виключно на візуальній оцінці. Хоча існує безліч способів подання особи, найбільшого поширення при аналізі емоцій отримали моделі активного способу. Даний клас моделей дозволяє

ефективно обчислити положення ключових точок, таких як центр зіниці ока, куточки губ, контур носа і т. д., А потім вже на їх основі побудувати розпізнавання.

Емоції відображають ставлення людини до різних явищ, вони характеризуються суб'єктивними переживаннями людини, не пов'язані безпосередньо з порушенням певних рецепторів, можуть виникати спонтанно. При однаковому впливі одним і тих же чинників на різних людей емоції можуть викликати у них різні переживання. Розпізнавання емоцій є складним психологічним процесом: навіть людина не завжди може правильно розпізнати емоції інших людей, а це значить, що для системи автоматичного розпізнавання дана задача є куди складнішою.

Основний алгоритм автоматичного розпізнавання емоцій складається з наступних етапів:

- реєстрація зображення;
- первинна обробка зображення;
- виділення обличчя на зображенні;
- виділення елементів обличчя;
- виділення ключових точок на обличчі;
- класифікація емоцій.

Первинна або попередня обробка зображення включає в себе видалення шумів, геометричні і колірні перетворення. Джерелами шуму можуть бути конструктивні недоліки засобів реєстрації зображень, погане освітлення сцени, механічні впливи на устаткування, стан об'єктів інтересу, перешкоди в каналах передачі інформації і т. п. Для видалення шумів можна застосувати усереднюючі фільтри або фільтри, засновані на порядкових статистиках.

Для виділення обличчя на зображенні найбільш ефективним є метод Віоли-Джонса, який забезпечує досить хорошу швидкість і високу точність виявлення заданих об'єктів на зображеннях. Метод був запропонований Полом Віолою і Майклом Джонсом в 2001 році і в даний час є основоположним для пошуку об'єктів на зображеннях. У ньому використовується принцип

скануючого вікна. При виявленні об'єктів методом Віоли-Джонса зображення обробляються в інтегральній формі. Інтегральне уявлення дозволяє виконати швидко обчислення сумарної яскравості довільного прямокутника на початковому зображенні з постійним часом, незалежно від розмірів цього прямокутника.

Після того як обличчя було виділено, необхідно виділити його елементи. Людина проявляє емоції за допомогою брів, очей і рота. Для виділення цих елементів на зображенні обличчя використовується все той же метод Віоли-Джонса. В якості вхідного зображення подається зображення обличчя, виділене на попередньому етапі. Щоб прискорити розпізнавання і зменшити помилкові виявлення, на обличчі задаються певні зони (рот завжди знаходиться в нижній половині обличчя, а брови і очі - у верхній).

Наступним кроком є знаходження ключових точок виділених елементів обличчя. Визначити емоції можна на основі аналізу декількох ключових точок.

Виділення ключових точок здійснюється наступним чином:

- перетворення кольорового зображення до напівтонового виду;
- перетворення з напівтонового виду до бінарного виду;
- застосування до бінарного зображення градієнтної маски;
- локалізація ключових точок.

Після того, як виділення ключових точок буде виконано, можна почати класифікацію емоції.

Наведений підхід до автоматичного розпізнавання емоцій може бути ефективно застосований в різних інтелектуальних людино-машинних системах.

## ДОДАТОК Б

Слайди презентації

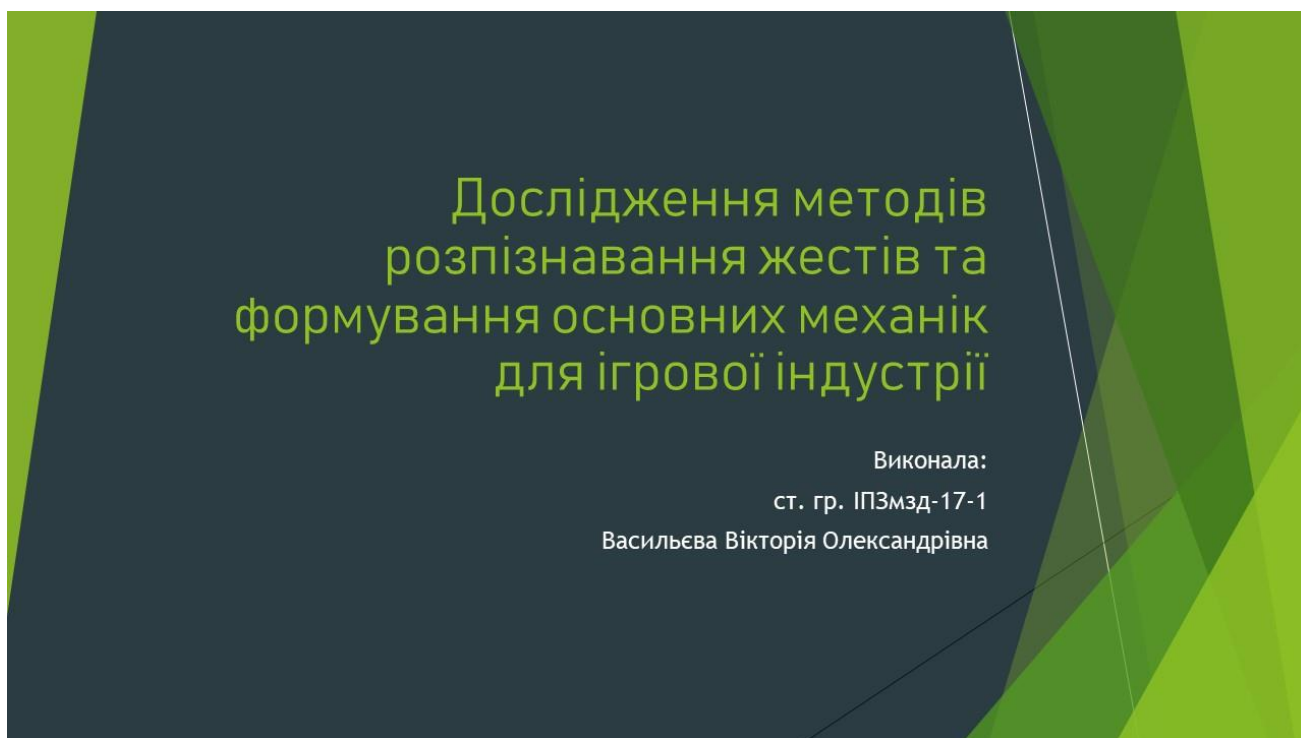


Рисунок Б.1 – Дослідження методів розпізнавання жестів та формування основних механік для ігрової індустрії



Рисунок Б.2 – Мета атестаційної роботи

## Актуальність атестаційної роботи

Розпізнавання жестів рук є досить актуальним завданням в таких задачах:

- ▶ Автоматизований сурдопереклад;
- ▶ Управління комп'ютером, роботом або штучною рукою;
- ▶ Природна взаємодія з тривимірними комп'ютерними моделями об'єктів.

Рисунок Б.3 –Актуальність атестаційної роботи

## Існуючі системи



**KINECT**  
for  XBOX 360

 Flutter

**ntobeBox**

Рисунок Б.4 – Існуючі системи



## Методи розпізнавання жестів

- ▶ Алгоритми на основі 3D-моделей
- ▶ Skeletal-based алгоритми
- ▶ Appearance-based алгоритми

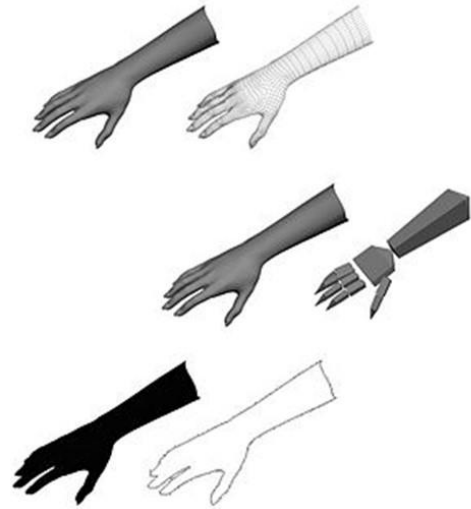


Рисунок Б.5 – Методи розпізнавання жестів

## Задача розпізнавання жестів

Для вирішення задачі розпізнавання об'єктів на відеопослідовності необхідно вирішити задачу пошуку і відстеження об'єктів.

Метод Віоли-Джонса є основоположним для пошуку об'єктів на зображенні в реальному часі в більшості існуючих алгоритмів розпізнавання та ідентифікації.

Також він є одним з кращих по співвідношенню ефективності розпізнавання і швидкості роботи. Алгоритм показує відмінні результати і розпізнає об'єкти під невеликим кутом, приблизно до 30 градусів, і при різних умовах освітленості.

Рисунок Б.6 – Задача розпізнавання жестів

## Метод Віюли-Джонса

Основні принципи, на яких базується метод:

- ▶ 1. Можливість подання зображення в інтегральному вигляді, що дозволяє швидко обчислювати необхідні об'єкти;
- ▶ 2. Використання ознак Хаара, тобто ознак цифрового зображення, які використовуються в розпізнаванні образів, за допомогою яких відбувається пошук потрібного об'єкта;
- ▶ 3. Застосування алгоритму бустінга (від англ. boost- поліпшення, посилення) для вибору найбільш підходящих ознак для шуканого об'єкта на даній частині зображення, процедури послідовного побудови композиції алгоритмів машинного навчання, коли кожен наступний алгоритм прагне компенсувати недоліки композиції всіх попередніх алгоритмів;
- ▶ 4. Використання каскадів ознак для швидкого відкидання вікон, де не знайдений об'єкт.

Рисунок Б.7 – Метод Віюли-Джонса

## Алгоритм для розпізнавання жестів на зображенні

Метою даної роботи є розпізнавання жестів на цифрових зображеннях. Запропонований алгоритм складається з двох процесів: добування і збереження ознак відомих жестів в базі даних і розпізнавання жестів.

Процес вилучення і збереження ознак відомих жестів відбувається наступним чином:

1. Перетворення зображення області жесту в напівтонове;
2. Зміна розміру області жесту до  $64 \times 64$  пікселів;
3. Застосування до отриманого на кроці 2 зображенню вейвлет-перетворення для вилучення ознак жесту (вейвлет-коефіцієнтів);
4. Збереження витягнутих ознак в базі даних.

У процесі розпізнавання невідомого жесту здійснюються кроки 1-3, потім отримані ознаки порівнюються з ознаками, що зберігаються в базі даних, на основі застосування методу головних компонент.

Рисунок Б.8 – Алгоритм для розпізнавання жестів на зображенні

## Схема алгоритму

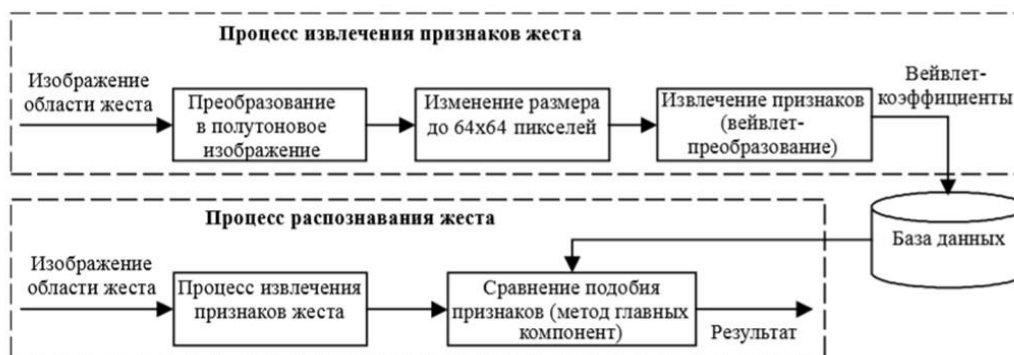


Рисунок Б.9 – Схема алгоритму

## Алгоритм для розпізнавання жестів на відеопослідовності

Процес розпізнавання жестів на відеопослідовності відбувається наступним чином:

1. Запит чергового відеофрейма. Перетворення відеофрейма в півтонове зображення. Застосування до напівтонового зображення методу Віоли - Джонса для пошуку області руки.
2. Якщо область руки виявлена, то виконується крок 3. У зворотному випадку здійснюється повернення на крок 1.
3. Запит чергового відеофрейма. Відстеження області руки на основі застосування алгоритму CAMShift.
4. Якщо відстеження здійснено, то виконується крок 5. У зворотному випадку відбувається повернення на крок 1.
5. Виконання процесу розпізнавання жесту.
6. Повернення на крок 3.

Рисунок Б.10 – Алгоритм для розпізнавання жестів на відеопослідовності

## Дизайн гри

Ідея мобільного додатку полягає в створенні мобільної гри на основі відомої усім дитячої гри «камінь-папір-ножиці».

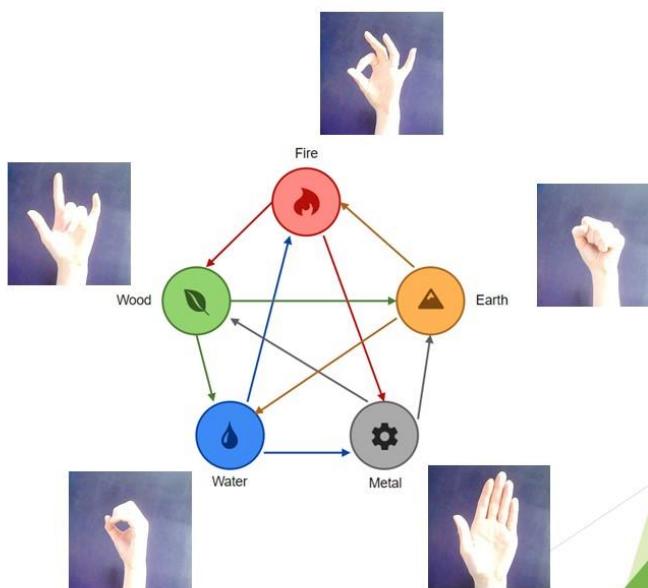


Рисунок Б.11 – Дизайн гри

## Макет UI

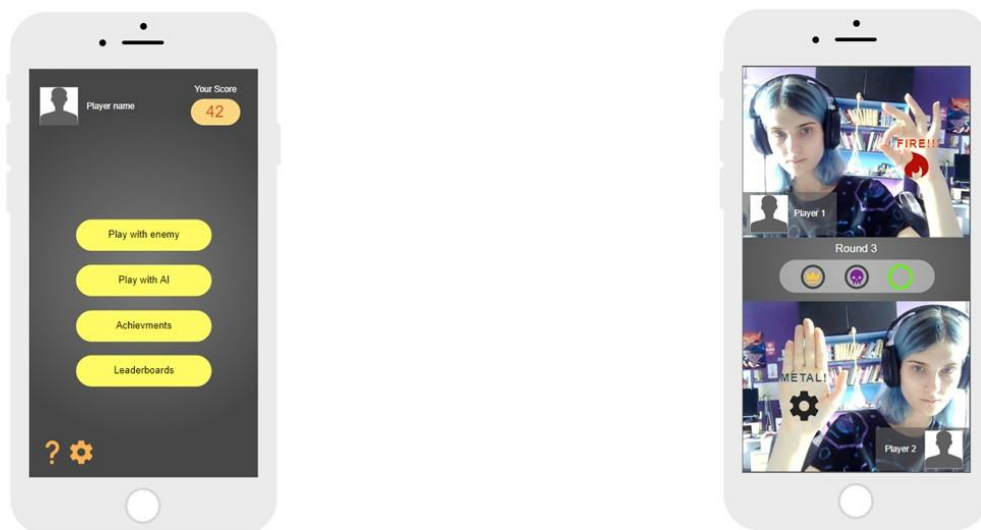


Рисунок Б.12 – Макет UI

## Висновки

У ході виконання роботи було:

- ▶ Проведено аналіз предметної галузі;
- ▶ Проведено аналіз існуючих підходів до розробки подібних застосунків та аналіз аналогів;
- ▶ Проаналізовано методи та алгоритми машинного зору;
- ▶ Запропоновано алгоритм для розпізнавання жестів на відеопослідовностях;
- ▶ Виконано моделювання ігрового додатку для демонстрації роботи алгоритму.

Рисунок Б.13 – Висновки

▶ Дякую за увагу!

Рисунок Б.14 – Дякую за увагу