

ПРОГРАМНИЙ КОМПЛЕКС БЕЗКОНТАКТНОГО КЕРУВАННЯ КОМП'ЮТЕРОМ З ВИКОРИСТАННЯМ ТЕХНОЛОГІЙ КОМП'ЮТЕРНОГО ЗОРУ

Вінницький Національний Технічний Університет

Анотація

Розроблено програмний комплекс для безконтактного керування курсором комп'ютера на основі бібліотеки комп'ютерного зору MediaPipe. Реалізовано алгоритми геометричної інтерпретації жестів та адаптивної стабілізації координат, що забезпечує ефективну роботу системи у реальному часі.

Ключові слова: комп'ютерний зір, MediaPipe, безконтактний інтерфейс, розпізнавання жестів.

Abstract

A software complex for contactless cursor control based on the MediaPipe computer vision library has been developed. Algorithms for geometric gesture interpretation and adaptive coordinate stabilization have been implemented, ensuring efficient real-time system operation.

Keywords: computer vision, MediaPipe, contactless interface, gesture recognition.

Вступ

Сучасний етап розвитку інформаційних технологій характеризується пошуком нових парадигм людино-машинної взаємодії (HCI), спрямованих на створення природних інтерфейсів (Natural User Interfaces) [1]. Традиційні пристрої введення, такі як комп'ютерна миша та клавіатура, мають суттєві обмеження у застосуванні в умовах, що вимагають дотримання стерильності, дистанційного керування або забезпечення доступності для осіб з обмеженими можливостями. Існуючі рішення, що базуються на використанні спеціалізованих сенсорів глибини або складних нейромережових архітектур типу OpenPose, часто є економічно недоцільними або вимагають значних обчислювальних ресурсів графічних прискорювачів. У зв'язку з цим актуальною науково-практичною задачею є розробка програмного забезпечення, здатного забезпечити високоточний оптичний трекінг руки з використанням стандартної веб-камери та ресурсів центрального процесора.

Методологія та архітектура системи

В основу розробленого програмного комплексу покладено використання фреймворку MediaPipe Hands, який реалізує конвеєр машинного навчання для виявлення долоні та побудови її скелетної моделі (рис. 1).

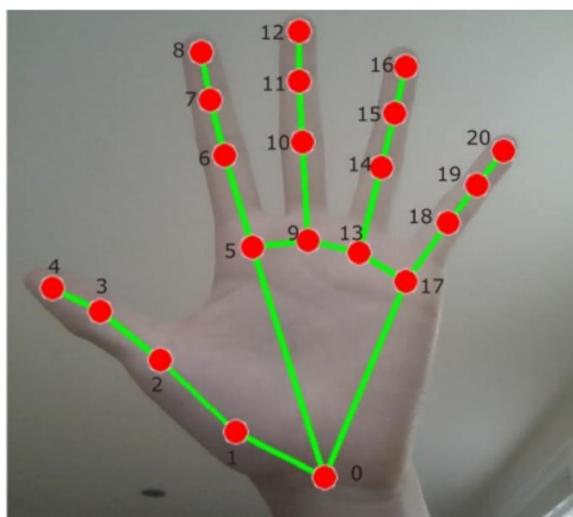


Рисунок 1 – Схема скелетної моделі руки

На відміну від класичних підходів, що вимагають сканування всього зображення на кожному кадрі, MediaPipe використовує детектор долоні лише для ініціалізації трекінгу, після чого прогнозує положення 21 ключової точки (landmarks) на наступних кадрах, базуючись на попередніх даних [2]. Це дозволяє досягти продуктивності обробки понад 25 кадрів за секунду.

Геометричне моделювання жестів

Для розпізнавання команд введення, зокрема емуляції натискання лівої кнопки миші було використано евристичний геометричний аналіз [3]. Цей метод базується на обчисленні евклідової відстані між координатами вершини великого пальця (Landmark 4) та вершини вказівного пальця (Landmark 8) у двовимірному просторі. Система безперервно моніторить зміну цієї відстані: при її зменшенні нижче експериментально встановленого порогового значення, яке складає близько 30–40 пікселів залежно від глибини сцени, генерується подія натискання. Перевагою такого підходу є його виняткова швидкість та відсутність необхідності у зборі великих датасетів для навчання класифікатора, що робить систему стійкою до змін фону та освітлення.

Алгоритмічна стабілізація та фільтрація шуму

Однією з ключових проблем оптичних систем керування є наявність високочастотного шуму координат, викликаного похибками квантування сенсора камери та природним фізіологічним тремором руки людини. Прямая передача «сирих» координат призводить до ефекту джитера, що унеможливує точне наведення курсора на дрібні елементи інтерфейсу. Для вирішення цієї проблеми у роботі впроваджено алгоритм експоненційного згладжування [4]. Суть методу полягає у рекурсивному розрахунку поточної координати курсора як зваженої суми поточного вимірювання та попереднього згладженого значення. Вибір оптимального коефіцієнта згладжування дозволив знайти баланс між плавністю руху курсора та мінімізацією затримки введення (latency), забезпечуючи відчуття чутливості та стабільності керування.

Результати дослідження та висновки

Експериментальне тестування розробленого програмного комплексу проводилося на персональному комп'ютері середньої потужності. Результати показали, що система забезпечує стабільну частоту оновлення кадрів на рівні 28–30 FPS при завантаженні центрального процесора не більше 35%. Застосування алгоритмів мапінгу та фільтрації дозволило досягти точності позиціонування, достатньої для виконання типових офісних завдань, веб-серфінгу та керування мультимедійними додатками. Розроблена система демонструє високий потенціал для використання як асистивна технологія, а також як компонент систем «розумного дому», забезпечуючи інтуїтивну взаємодію з цифровим середовищем. Подальший розвиток системи передбачає розширення функціоналу за рахунок імплементації нових жестів та створення графічного інтерфейсу користувача для індивідуального калібрування чутливості та параметрів згладжування.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Natural user interface [Електронний ресурс]. – Режим доступу: https://en.wikipedia.org/wiki/Natural_user_interface
2. Hand landmarks detection guide [Електронний ресурс]. – Режим доступу: https://ai.google.dev/edge/mediapipe/solutions/vision/hand_landmarker
3. Hand gesture recognition using average background and logical heuristic equations [Електронний ресурс]. – Режим доступу: <https://www.researchgate.net/publication/324991559>
4. Exponential Moving Average (EMA): Definition, Formula, and Usage [Електронний ресурс]. – Режим доступу: <https://www.investopedia.com/terms/e/ema.asp>

Шатайло В'ячеслав Андрійович — студент групи 2КІ-25м, факультет інформаційних технологій та комп'ютерної інженерії, Вінницький Національний Технічний Університет, Вінниця, e-mail: viacheslavshatailo@gmail.com

Shatailo Viacheslav Andriyovych — student of group 2KI-25m, faculty of information technologies and computer engineering, Vinnytsia National Technical University, Vinnytsia, e-mail: viacheslavshatailo@gmail.com