

УДК 621.865.8

Лаціна Руслан Олегович

Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського»

ДЕТЕКТУВАННЯ РУХІВ ЛЮДИНИ НА ВІДЕО ПОСЛІДОВНОСТІ В РЕАЛЬНОМУ ЧАСІ.

Лаціна Р.О. Детектування рухів людини на відео послідовності в реальному часі. Представлено комплексний алгоритм розпізнавання рухів людини на відеопослідовності в реальному часі, який може працювати в режимі автономного розпізнавання і в режимі інтерактивного розпізнавання рухів людини. Запропоновано дворівневу архітектуру для комплексного алгоритму розпізнавання рухів людини, що містить на першому рівні кроки отримання послідовних кадрів з відеокамери, перед обробки отриманих кадрів, і виявлення людини на відеокадрі.

Ключові слова: Детектування, розпізнавання рухів, відео послідовність, алгоритм, відеокамера, алгоритм Джонса-Віоли, ознаки Хаара, нейронна мережа.

Лацина Р.О. Детектирование движений человека на видео последовательности в реальном времени. Представлен комплексный алгоритм распознавания движений человека на видеопоследовательности в реальном времени, который может работать в режиме автономного распознавания и в режиме интерактивного распознавания движений человека. Предложена двухуровневую архитектуру для комплексного алгоритма распознавания движений человека, которая содержит на первом уровне шаги получения последовательных кадров с видеокамеры, предварительной обработки полученных кадров, и выявления человека на видеокадре.

Ключевые слова: Детектирование, распознавания движений, видео последовательность, алгоритм, видеокамера, алгоритм Джонса Виоли, признаки Хаара, нейронная сеть.

Latsina R. O. Detection of human movements at real time video sequence. A complex algorithm for recognizing human movements on a video sequence in real time is presented, which can work in the mode of autonomous recognition and in the mode of interactive recognition of human movements. A two-level architecture for a complex algorithm for recognizing human movements is proposed, which contains, at the first level, steps to obtain successive frames from the video camera, preliminary processing of the received frames, and identification of a person on the video frame.

Keywords: Detection, motion recognition, video sequence, algorithm, video camera, Jones Viola algorithm, Haar signs, neural network.

Вступ та постановка проблеми дослідження. В історії розвитку персонального комп'ютера спостерігалася еволюція користувальницького інтерфейсу в людино-машинній взаємодії (ЛМВ) від текстового командного інтерфейсу до графічного інтерфейсу, від простої клавіатури до миші, джойстика, електронної ручки, відео камери, і т.д. [1,] Пристрої ЛМВ стали більш зручними і природними для користувача. В даний час, з введенням нових понять, таких як «Віртуальне середовище – ВС», «людино-машинна інтелектуальна інтеракція – ЛМІ», «перцепційний призначений для користувача інтерфейс – ППІ » і т.д. потрібна розробка більш потужних і зручних способів взаємодії людини з комп'ютерною системою.

В якості одного із способів забезпечення комфортної взаємодії з комп'ютером, людська рука може бути використана в якості інтерфейсу введення [1, 2]. Жести є потужним каналом зв'язку, який формує основну частину передачі інформації в нашому повсякденному житті.

У порівнянні з традиційними пристроями ЛМВ, жести є менш нав'язливим, простим, більш зручними і природним способом взаємодії для користувачів. Проте, виразність жестів все ще залишається недостатньо вивченою для вирішення проблеми людино-машинної взаємодії.

В останні роки з'явилася і почала швидко розвиватися тенденція використання жестів, особливо жестів руки, як способу взаємодії з комп'ютерною системою. Розпізнавання жестів, таким чином, стало найважливішою частиною в ЛМІ і почало залучати безліч дослідників. Крім того, розроблені в ЛМІ технології також знаходять застосування в інших областях, таких як управління роботами, телеконференції, переклад мови жестів (для глухих), управління комп'ютерними іграми, і т.д.

Мета. Запропонувати комплексний алгоритм розпізнавання рухів людини на відеопослідовності в реальному часі, який може працювати в режимі автономного розпізнавання і в режимі інтерактивного розпізнавання рухів людини. Окреслити дворівневу архітектуру для комплексного алгоритму розпізнавання рухів людини, що містить на першому рівні кроки отримання послідовних кадрів з відеокамери, перед обробки отриманих кадрів, і виявлення людини на відеокадрі.

Аналіз останніх досліджень і публікацій. На сьогодні, кількість наукових розробок, щодо теми дослідження є суттєвою.

А.І. Томіліна та С.А. Савельєв [1] виконали порівняльний аналіз детекторів точкових особливостей для детектування і визначення напрямку руху на відеопослідовності. Авторами здійснено тестування швидкодії і якості реалізованих алгоритмів, проаналізовано вплив попередньої обробки на результат роботи алгоритмів.

К.В. Шепелев [2] розкрив точність детектування рухомих об'єктів на основі програмного забезпечення з відкритим кодом, що не перевищує 60%, це обумовлено існуючою проблемою, перший аспект полягає у відсутності алгоритмів і програмного забезпечення для імпортозаміщення, другий аспект полягає у високій вартості створення окремої системи відеоспостереження з функціями детектування і класифікації рухомих об'єктів.

Виявлення рухів об'єктів в послідовності зображень розкрили на сторінках своєї праці [3] Нгуєн В.Н. та Нгуєн Т.Т.

Із зарубіжних авторів варто наголосити на роботах Wu Y., Liu Q., and Huang T. [4], Yang J., Xu Y., Chen C. [5].

Однак, незважаючи на масштабність наукових досліджень, питання детектування рухів на відео послідовності залишається відкритим та потребує детального опрацювання.

Виклад основного матеріалу. Для опису динамічного компонента жестів потрібно створити відповідний класифікатор руху, що забезпечує можливість розпізнавання різних форм руху в реальному часі. Класифікатор руху повинен сприяти розпізнаванню сильно деформованих форм траєкторії (у порівнянні зі стандартними навчальними формами), тому що реальні рухи руки, що виконуються користувачами, можуть змінюватися і істотно відрізнятися від стандартних форм. Інші труднощі при розпізнаванні руху руки полягають в тому, що точки траєкторії, які спостерігаються при виконанні жесту дуже обмежені, особливо коли жест виконується з високою швидкістю за короткий час. У розглянутих роботах немає рішення проблеми розпізнавання руху. Алгоритм реалізації системи детектування рухів на відео послідовності, що пропонується у рамках даного дослідження наведено на рисунку 1.

Перший рівень призначений для виявлення присутності руки в області видимості відеокамери і для ініціалізації роботи алгоритмів розпізнавання і трекінгу другого рівня. Звичайна відеокамера записує в одну секунду від 15 до 30 кадрів розміру 640x480 пікселів, і в кожен момент часу відеосистема доставляє тільки один кадр для обробки.

Система розпізнавання жестів в реальному часі повинна обробляти відеокадри в міру їх надходження з відеосистеми. Таким чином, час обробки кожного кадру має бути не більше 66 мілісекунд, щоб не було затримки в роботі системи.

Для скорочення часу обробки, розмір кадрів зменшується до 320x240 пікселів на кроці перед обробки. На цьому кроці також відбувається перетворення отриманих кольорових кадрів у напівтонові зображення, тому що надалі алгоритми обробки працюють тільки з напівтоновими зображеннями.

Детектор руки виконує пошук присутності руки на кожному відеокадрі. Якщо на відеокадрі не виявляється рука, то кадр скасовується і в обробку запускається наступний відеокадр з відеосистеми. Якщо рука присутня на кадрі, то детектор руки відключається. Метою кроку виявлення руки є ініціалізація роботи алгоритмів трекінгу другого рівня шляхом визначення прямокутної області кадру, в якій знаходиться рука. Дана область кадру буде використовуватися як модель об'єкта для алгоритму трекінгу.

Для розпізнавання автономних жестів, на другому рівні застосовується розпізнавання пози і пов'язаного з нею глобального руху. Після того як становище руки визначається за допомогою детектора руки (на першому рівні), прямокутна область, де знаходиться рука, зберігається і передається в алгоритм трекінгу, а сам детектор відключається. Алгоритм трекінгу аналізує цю область для створення моделі об'єкта і починає процес трекінгу. вступники відеокадри з відеосистеми потім передаються відразу на другий рівень.

Після попередньої обробки (зменшення розміру і перетворення кадру в півтонове зображення), механізм трекінгу запускається для пошуку місця розташування руки на новому відеокадрі. Алгоритм розпізнавання поз потім працює тільки з областю кадру, де знаходиться рука, а не з цілим кадром, і, таким чином, може забезпечити високу швидкість розпізнавання, незалежно від реального розміру відеокадру. Якщо жест не завершений, новий кадр пропускається на обробку, інакше алгоритм розпізнавання руху запускається для розпізнавання отриманого жесту. Умовою завершення жесту приймається відсутність руки на кадрі, наприклад, коли рука рухається за межею області видимості камери і алгоритм трекінгу не може визначити

місцеположення руки на кадрі. Розпізнаний жест потім може перетворюватися в команди для роботи з комп'ютерною системою. Даний режим може використовуватися в системі взаємодії на основі жестів для виконання команд, функції яких аналогічні гарячим клавішам.

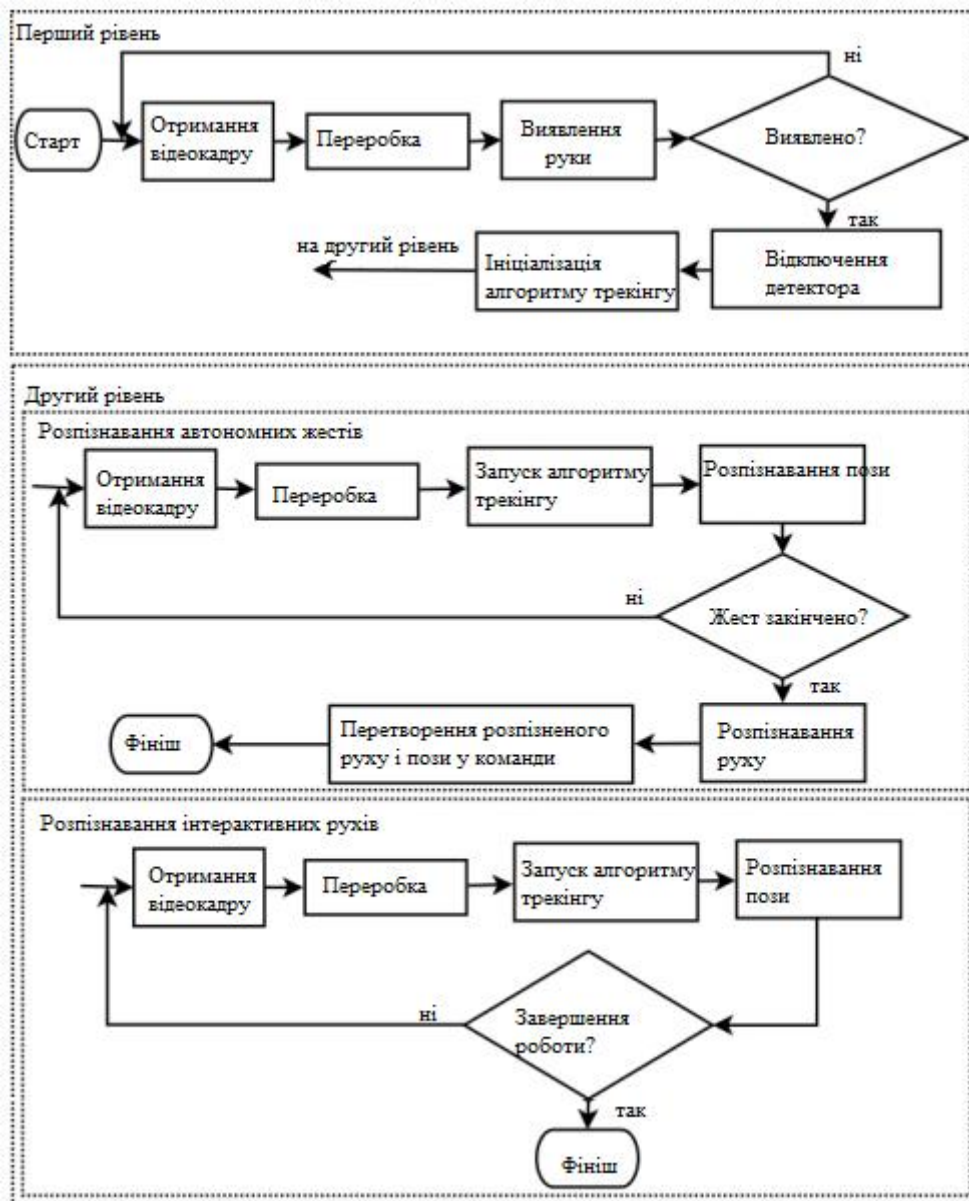


Рис. 1. Схема комплексного алгоритму розпізнавання жестів

Для розпізнавання інтерактивних жестів для прямої маніпуляції, на другому рівні застосовується розпізнавання пози і алгоритм трекінгу. При цьому трекінг також відповідає за спостереження за становищем руки на кожного з них.

Розпізнавання пози дозволяє виконати команди, такі як клацання миші.

Даний режим розпізнавання призначається для безперервної роботи з комп'ютерною системою, такий як віддалене управління курсором миші, малювання за допомогою жестів, і т.д. Умовою завершення роботи є відсутність руки на кадрі, коли рука рухається за межу області видимості камери.

Виявлення руки на зображенні є первинним завданням системи розпізнавання жестів. В даний час, метод Джонса-Віоли з використанням ознак Хаара і каскадного AdaBoost класифікатора [6] вважається одним з кращих алгоритмів по співвідношенню ефективності розпізнавання до часу обробки для таких завдань [2, 4]. Метод Джонса-Віоли спочатку був розроблений для завдання виявлення людських облич на зображенні.

Детектор Джонса-Віоли також володіє вкрай низькою ймовірністю помилкового виявлення об'єкта і добре розпізнає об'єкт під невеликим кутом, приблизно до 15 градусів для руки і 30 градусів для особи. Таким чином, застосування методу Джонса-Віоли дає можливість виявлення руки на відеопослідовності в реальному часі.

Основні принципи, на яких базується метод, полягають в наступному:

–представлення зображення в «інтегральному» вигляді, що дозволяє швидко обчислювати ознаки Хаара;

–пошук потрібного об'єкта на основі ознак Хаара;

–використання каскадної класифікації AdaBoost, що дозволяє швидко відкидати вікна, де не знайдений об'єкт, і концентрувати обчислення для областей з найбільшою ймовірністю знаходження об'єкта.

Основна ідея запропонованого алгоритму розпізнавання рухів на відео послідовності складається в тому, що зображення розглядається як текстовий документ, в якому візуальні ознаки (характерні точки і області) зображення враховуються як слова, що утворюють даний документ. Клас документа визначається шляхом обчислення частоти появи деяких «ключових слів». Дана ідея була взята з популярної в обробці текстових документів моделі "bagof-words" (BOW) [7].

Для застосування цієї ідеї в розпізнаванні об'єкта, характерні ознаки виділяються з набору зображень (бази даних для навчання), і потім поділяються на групи. У кожній групі вибирається один «представник» для всіх ознак групи. Обрані представники будуть служити «ключовими словами», з яких побудується так званий «словник ключових слів» (Безліч ключових слів). При зіставленні виділених ознак вхідного зображення з ключовими словами з словника ознак виходить гістограма ключових слів для даного зображення. Гістограма, що генерується служитиме дескриптором для розпізнавання у нейромережевому класифікаторі.

Алгоритм розпізнавання форми руки складається з наступних етапів:

1. Навчання класифікатора

а. Генерація словника візуальних ознак:

–Виділення ознак методом SURF;

–Кластеризація методом K-means;

–Генерація словника з кластерів.

б. Створення дескрипторів і навчання нейронної мережі:

–Зіставлення виділених ознак зі словником;

–Навчання класифікатора (нейронної мережі).

2. Розпізнавання

–Виділення ознак методом SURF;

–Зіставлення виділених ознак зі словником;

–Розпізнавання в нейронної мережі.

Початкове положення людини на відеокадрі визначається на етапі виявлення за допомогою детектора Джонса-Віоли. В процесі трекінгу, траєкторія руху записується для подальшої обробки в алгоритмі розпізнавання руху. Для вирішення цього завдання був розроблений швидкий алгоритм розпізнавання траєкторії руху руки на основі нейронної мережі.

Алгоритм розпізнавання траєкторії руху руки складається з наступних етапів:

Спрощення і згладжування траєкторії руху. В процесі трекінгу, траєкторія руху записується у вигляді масиву точок. Однак, при переміщенні людини, людина часто робить випадкові рухи малої амплітуди, що призводить до не "гладкої" траєкторії руху. Таким чином, в масиві даних, отриманому після закінчення руху, завжди присутній якийсь шум. Крім того, кількість точок для кожного руху досить велика, що ускладнює обробку. Для вирішення цієї проблеми, застосовується алгоритм Рамер-Дугласа-Пекера (Ramer-Douglas-Peucker), що дозволяє "згладити" траєкторію.

Передискретизація і перетворення траєкторії. Кількість точок в траєкторії руху людини не фіксована, а змінюється в залежності від швидкості переміщення і від швидкості обробки даних програмою. Механізм розпізнавання завжди вимагає, щоб число входів було фіксовано.

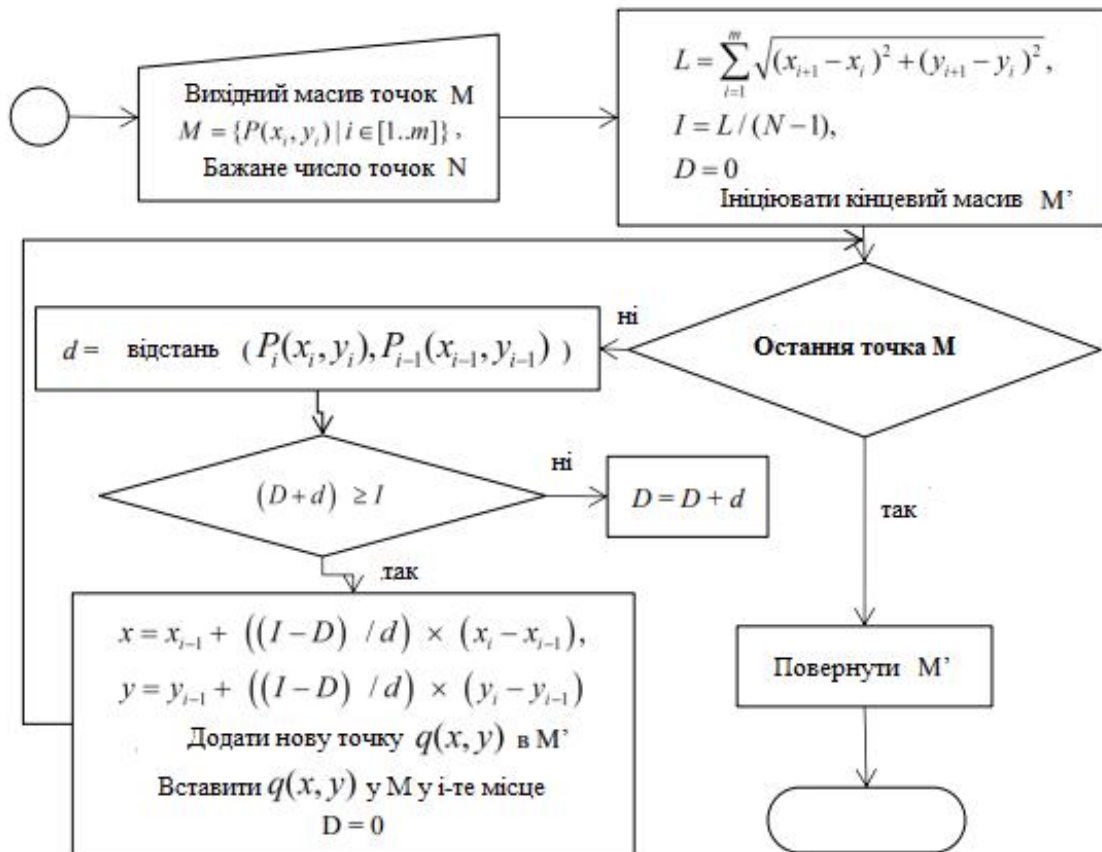


Рис. 2. Блок-схема процесу передискретизації і перетворення траєкторії

Повторна дискретизація (resampling) використовується для фіксування кількості точок траєкторії [8]. На даному етапі також виконується масштабування і переміщення траєкторії для підвищення ефективності дескрипторів, обчислених на наступному кроці.

Обчислення дескриптора. Масив точок траєкторії перетворюється в масив векторів нахилу і виконується обчислення синуса або косинуса кутів нахилу. Масив синусів (або косинусів) надалі буде служити входом для класифікатора.

Навчання та розпізнавання в нейронній мережі. Для розпізнавання форми траєкторії руху людини в якості класифікатора застосовується багаточарова нейронна мережа зі зворотнім поширенням помилок.

Висновки та перспективи подальших досліджень. Представлений новий комплексний алгоритм розпізнавання рухів людини на відеопослідовності в реальному часі, який може працювати в режимі автономного розпізнавання і в режимі інтерактивного розпізнавання рухів людини. Запропоновано дворівневу архітектуру для комплексного алгоритму розпізнавання жестів, що містить на першому рівні кроки отримання послідовних кадрів з відеокамери, перед обробки отриманих кадрів, і виявлення руки на відеокадрі. На другому рівні виконується стеження за рукою в часі, розпізнавання пози і розпізнавання глобального руху.

Запропоновано застосування алгоритму Джонса-Віюлі для виявлення руки у відеопотоці з можливістю функціонування в реальному часі.

Алгоритм працює на основі ознак Хаара, інтегрального зображення, і каскадного AdaBoost класифікатора.

- 1.Томила А.И., Савельев С.А. Анализ методов детектирования движения на видеопоследовательности на основе точечных особенностей // Актуальные проблемы авиации и космонавтики. 2016. – № 12. – С. 658-660
- 2.Шепелев К.В. Детектирование и классификация движущихся объектов в видеопоследовательности // Известия высших учебных заведений. Поволжский регион. Технические науки. 2017. – № 12. – С. 658-660
- 3.Нгуен В.Н., Нгуен Т.Т. Обнаружение движущих объектов в последовательности изображений // Известия ТулГУ. Технические науки. 2017. – № 2. – С. 133-138
- 4.Wu Y., Liu Q., and Huang T. An adaptive self-organizing color segmentation algorithm with application to robust real-time human hand localization // Proceedings of the International Workshop on Recognition, Analysis, and Tracking of

- Faces and Gestures in Real-Time Systems (RATFG-RTS '99). – Washington, DC: IEEE Computer Society, 1999. –P. 161-166.
5. Yang J., Xu Y., Chen C. Gesture Interface: Modeling and Learning // IEEE International Conference on Robotics and Automation (San Diego, CA, 1994). – IEEE Computer Society, 1994. –V. 2. –P. 1747-1752.
6. Руатаева А.В., Бандеев О. Е. Video based flame detection algorithm // Сибирский журнал науки и технологий. 2017. – № 4. – С. 796-803.
7. Теплицкий Э.Г., Захаров С.М., Митрохин М.А. Метод обнаружения движущихся объектов на сложном динамическом фоне в оптическом диапазоне // Известия ВУЗов. Поволжский регион. Технические науки. 2016. – № 4 (40). – С. 16-26.
8. Афанасьев, Г.И. Особенности применения видеодетекторов движения в современных охранных системах видеонаблюдения [Текст] / Г.И. Афанасьев, Ю.В. Костенкова, С.А. Тоноян // Территория инноваций. – Энгельс, 2018. – № 4 (20). – С. 11-22