

ІДЕНТИФІКАЦІЯ РУХОМИХ ОБ'ЄКТІВ В ПОТОЦІ ВІДЕОДАНИХ

Гулієв Н.Б.¹⁾, Войтюк І.Ф.²⁾

Тернопільський національний економічний університет

¹⁾магістрант; ²⁾к.т.н., доцент

І. Вступ

В наш час генерується велика кількість відеоматеріалів. Це можуть бути як звичайні відео, так і відео в реальному часі. Тому з'являється необхідність для автоматизованого аналізу такої великої кількості відеоматеріалів. Одним таким пунктом є ідентифікація об'єктів в потоці відеоданих.

Ідентифікація об'єктів – це галузь комп'ютерного зору, яка дозволяє візуально спостерігати за об'єктами на зображеннях, відео-файлах, а також на потоковому відео в реальному часі. Комп'ютерний зір можна поділити на два види: виявлення та ідентифікація. Виявлення об'єкта в основному залежить від апаратної частини, чим якісніше обладнання тим більш якісніше зображення можна отримати на виході і вже потім за допомогою алгоритмів виявити об'єкт. Ідентифікація – це повністю алгоритмічний процес, який дозволяє з певною точністю виділити об'єкт та зрозуміти, що на ньому знаходиться.

З технічної сторони, відео – це великий набір зображень, які дуже швидко відображаються за певний період часу. Одне статичне зображення в відео називається відео фреймом. В більшості відео файлів, за одну секунду відображається від 20 до 32 зображень, це називають фрейми за секунду (fps). Складність ідентифікації об'єкта полягає в тому, що алгоритму потрібно швидко обробити кожен кадр і видати результат.

У наш час існує безліч галузей, де можна використати ідентифікацію об'єктів: аналізу зображень і відео файлів, робототехніки, відеоспостереження та ін.

II. Мета роботи

Метою роботи є розробка програмного продукту для ідентифікації рухомих об'єктів в потоці відеоданих з використанням алгоритму відстеження руху в певній області та алгоритму ідентифікації об'єкта на цій області.

III. Постановка задачі

Даний додаток розроблений на мові програмування C# на платформі .NET. Нажаль, платформа .NET немає підтримки комп'ютерного зору в своїй реалізації, тому для цього було використано крос-платформну бібліотеку Emgu CV, яка в свою чергу є обгорткою над бібліотекою OpenCV.

OpenCV – це бібліотека з відкритим вихідним кодом, була ще написана в 1999 році на мові програмування C++ і підтримується в наш час, містить безліч методів та алгоритмів для обробки зображень.

Щоб визначити, що відбувся рух на відео, потрібно порівнювати кожен наступний фрейм з попереднім і знайти різницю. Для цього використовується метод, з бібліотеки EmguCV, CvInvoke.Multiply, в якому на вхід подаються два фрейми, у вигляді масивів, а на вихід отримуємо різницю в них. Різниця визначається за формулою:

$$dst(I) = scale(I) \cdot src1(I) \cdot src2(I), \quad (1)$$

де $scale$ – коефіцієнт масштабу, $src1$ – масив 1, $src2$ – масив 2.

Для ідентифікації об'єкта використовується метод Віоли-Джонса, який базується на каскадних ознаках Хаара. Це підхід, що базується на машинному навчанні, де каскадна ознака підготовлена з безлічі позитивних та негативних зображень для навчання класифікатора.

Алгоритм потребує багато позитивних і негативних зображень, після цього потрібно витягнути з нього ознаки. Для цього використовуються ознаки Хаара, які показані на рисунку 1.

Кожна ознака являє собою одне значення, яке отримане шляхом віднімання суми пікселів білих прямокутників від суми пікселів чорних прямокутників.

При зберіганні зображення в інтегральному форматі, перевірка прямокутної ознаки на конкретній позиції проводиться за константний час $O(1)$, що являється їхньою перевагою при порівнянні збільш точними варіантами фреймів. Кожна прямокутна область в використовуваних ознаках завжди суміжна з іншим прямокутником, тому розрахунок ознаки з 2 прямокутниками

складається з 6 звернень в інтегральний масив, для ознаки з 3 прямокутниками - з 8, з 4 прямокутниками - з 9.

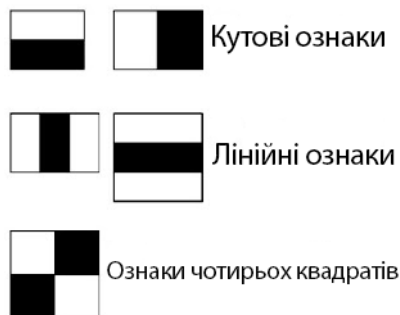


Рисунок 1 – Ознаки Хаара

На щастя, це можна зробити за одну ітерацію, використовуючи формулу повторення:

$$s(x, y) = s(x, y - 1) + i(x, y) \tag{2}$$

$$ii(x, y) = ii(x - 1, y) + s(x, y) \tag{3}$$

Або простіше:

$$ii(x, y) = i(x, y) + ii(x - 1, y) + ii(x, y - 1) - ii(x - 1, y - 1). \tag{4}$$

IV. Програмна реалізація

Після запуску додатку можна побачити інтерфейс користувача (рисунок 2), на якому зображено три квадрати, де виводяться зображення з потоку відеоданих. На першому квадраті відображається оригінал відео, по середині червоними квадратами позначаються ідентифіковані об'єкти, в третьому квадраті – рухи, які знайдені на відео. Нижче можна побачити елементи користувацького інтерфейсу, які призначені для вибору потоку, в якому буде ідентифіковуватися об'єкти (відео-файл або вивід з камери) та об'єкти, які потрібно ідентифікувати, а також кнопки для початку і зупинки ідентифікації. Також під виводом ідентифікованих об'єктів відображається час в мілісекундах, за який ідентифікувалися об'єкти.

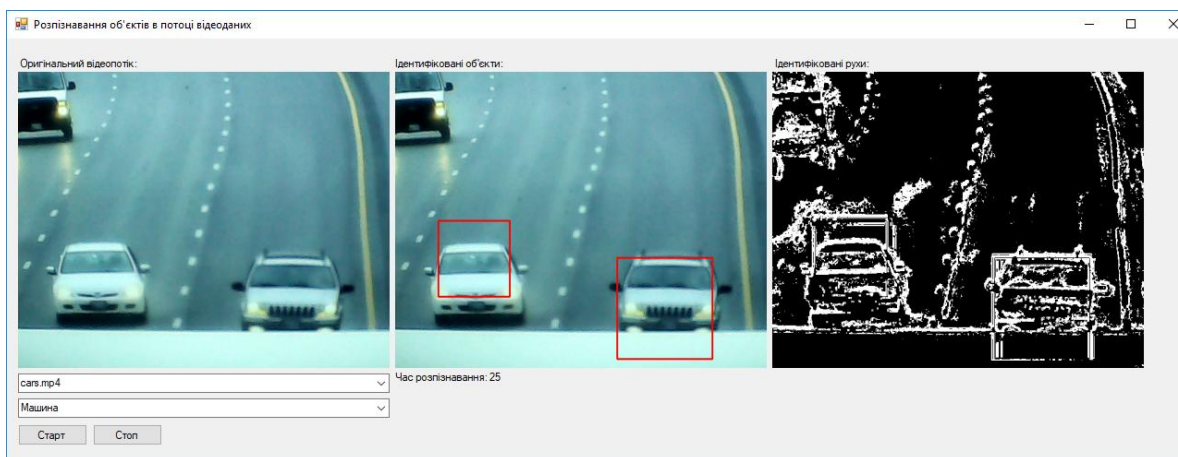


Рисунок 2 – Вікно додатку

Висновок

У даній роботі описано розроблений додаток для ідентифікації рухомих об'єктів в потоці відеоданих. За допомогою комп'ютерного зору і використання алгоритму відстеження рухів в відеопотоці та методу Віоли-Джонса, розпізнавання об'єктів на основі класифікації каскадів Хаара, було одержано результат, в якому було ідентифіковано лише ті об'єкти, які рухалися на відео, а всі інші ігнорувалися.

Список використаних джерел

1. P. Viola and M. Jones. Robust real-time face detection. IJCV 57(2), 2004.
2. Сайт бібліотеки EmguCV. [Електронний ресурс]. Режим доступу - <http://www.emgu.com>
3. Сайт бібліотеки OpenCV. [Електронний ресурс]. Режим доступу - <https://opencv.org/>
4. С. Сандерсон .NET Framework с примерами на C# для профессионалов – М.: «Вильямс», 2010. — 560с.