

**Яременко Костянтин Миколайович**

*студент*

*Національний технічний університет України «Київський політехнічний інститут»*

**Яременко Константин Николаевич**

*студент*

*Национальный технический университет Украины «Киевский политехнический институт»*

**Yaremenko K. N.**

*Student*

*National Technical University of Ukraine «Kyiv Polytechnic Institute»*

## РОЗПІЗНАВАННЯ ЖЕСТІВ РУКИ

## РАСПОЗНАВАНИЕ ЖЕСТОВ РУКИ

## HAND GESTURE RECOGNITION

**Анотація.** Дана робота присвячена аналізу алгоритмів і методів розпізнавання об'єктів на зображенні в реальному часі, а також реалізації програми для жестів руки. Метою роботи є розробка програмного прототипу, який дозволяє розпізнавати окремі жести руки. Розроблений прототип розпізнає жести, що відповідають завантаженим шаблонам.

Програма реалізована за допомогою бібліотеки комп'ютерного зору OpenCV на мові програмування C++.

**Ключові слова:** комп'ютерний зір, OpenCV, контурний аналіз, колірний підхід, відстані Гаусдорфа.

**Анотация.** Даная работа посвящена анализу алгоритмов и методов распознавания объектов на изображении в реальном времени, а также реализации программы для распознавания жестов руки. Целью работы является разработка программного прототипа, который позволяет распознавать отдельные жесты руки.

Программа реализована с помощью библиотеки компьютерного зрения OpenCV на языке программирования C++.

**Ключевые слова:** компьютерное зрение, OpenCV, контурный анализ, цветовой подход, расстояния Хаусдорфа.

**Annotation.** Graduation work is devoted to analysis of algorithms and recognition methods in the image in real time, and program implementation for controlling a computer with hand gestures. The purpose is developing software that allows to recognize hand gesture.

The program is implemented through computer vision library OpenCV using the programming language C++.

**Keywords:** computer vision, OpenCV, contour analysis, color method, Hausdorff distance.

### Вступ

Метою роботи є розробка програмного прототипу для розпізнавання жестів руки. Для вирішення цієї задачі в роботі проводиться аналіз і обґрунтовується вибір відповідних методів та алгоритмів, які належать до оптичного підходу керування пристроями.

Для оптичної взаємодії користувача з комп'ютером використовується звичайна веб-камера без додаткових опціоналів.

Даний прототип стане основою для подальшої розробки програми керування персональним комп'ютером за допомогою жестів руки.

### 1. Огляд обраних методів для локалізації жесту та його розпізнавання

#### 1.1. Колірний простір HSV

Простір HSV (Hue — тон, Saturation — насиченість, Value — яскравість) являється інтуїтивною кольоровою моделлю, так як оперує тими ж характеристиками кольору, що і людина. Якщо уявити конкретний простір RGB у вигляді кубу, то вздовж головної діагоналі  $(0, 0, 0) - (1, 1, 1)$  розташовані сірі кольори. Координата V задає яскравість кольору і пропорційна відстані від початку координат і вздовж головної діагоналі. Її можна отримати за допомогою перпендикулярної проєкції точки в RGB на головну діагональ. Координата S визначає насиченість. В просторі RGB це виглядає як

відстань від головної діагоналі. Чим вона більша, тим ближче колір до чистих кольорів, розташованих у вершинах кубу. Координата  $H$  – це колірний тон. Якщо розглянути площину, що перпендикулярна до головної діагоналі кубу, то кут повороту вектору кольору в цій площині і визначає тональність. Перетворення із RGB простору в HSV, розраховується за формулами (1–3)

$$V \leftarrow \max(R, G, B) \quad (1)$$

$$S \leftarrow f(x) = \begin{cases} \frac{V - \min(R, G, B)}{V}, & V \neq 0 \\ 0, & V = 0 \end{cases} \quad (2)$$

$$V \leftarrow \begin{cases} \frac{60(G - B)}{S}, & V = R \\ 120 + \frac{60(B - R)}{S}, & V = G \\ 240 + \frac{60(G - B)}{S}, & V = B \end{cases} \quad (3)$$

### 1.2. Метод відстаней Гаусдорфа

В математиці відстань Гаусдорфа або метрика Гаусдорфа, яку також називають Помпейо-Гаусдорфою відстанню, міра яка визначає, як далеко одна від одної знаходяться дві підмножини метричного простору. Виходить множина непустих компактних підмножин метричного простору.

Нехай  $X$  і  $Y$  дві не пусті підмножини метричного простору  $(M, d)$ . Де їх відстань Гаусдорфа  $d_H(X, Y)$  розраховується за формулою (4), що еквівалентно формулі (5).

$$d_H(X, Y) = \max \left\{ \max_{x \in X} \min_{y \in Y} d(x, y), \max_{y \in Y} \min_{x \in X} d(x, y) \right\} \quad (4)$$

$$d_H(X, Y) = \min \{d_H(X, Y); X \subseteq Y_\epsilon \cap Y \subseteq X_\epsilon\} \quad (5)$$

де  $X_\epsilon = \bigcup \{z \in M; d(z, x) \leq \epsilon\}$

Тобто  $\epsilon$  – множина всіх точок в межах  $\epsilon$  множини  $X$ .

## 2. Реалізація прототипу

### 2.1. Виділення шкіри

Щоб локалізувати руку спочатку необхідно виділити регіони кольору шкіри. Для цього була в ручну виділена шкіра людини на зображенні за умов природнього і штучного освітлення (рисунок 1–2).

Так як задача полягає в знаходженні шкіри при будь-якому освітленні, то для регіону, що нас цікавить можна звзвати множину значень кольорового простору HSV, а саме  $0 \leq H \leq 30, 15 \leq S \leq 200$ . Також було прийнято рішення обмежити компоненту  $V$  значеннями  $15 \leq V \leq 250$ , щоб виключити затемнення і засвічування області. Користувач може сам задати потрібні значення через відповідний пункт меню (рисунок 4).

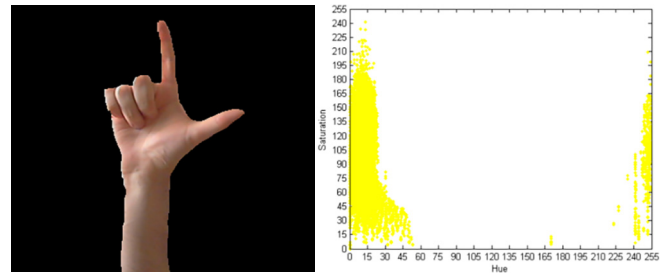


Рисунок 1. Зображення шкіри при природному освітленні та значення насиченості та яскравості в ненульових пікселях

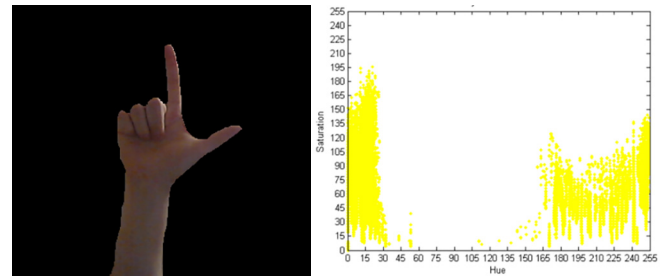


Рисунок 2. Зображення шкіри людини при штучному освітлення та значення насиченості та яскравості в ненульових пікселях

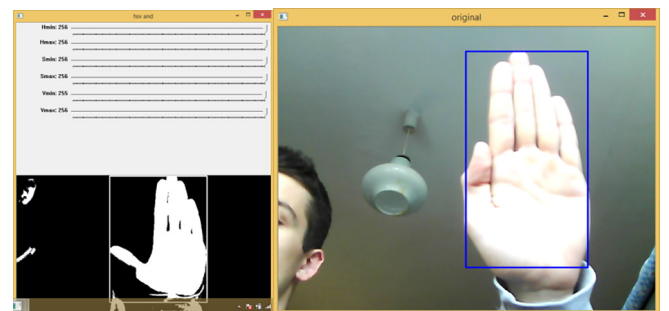


Рисунок 3. Вхідний кадр із локалізованою рукою і бінаризований регіон шкіри

### 2.2. Розпізнавання жесту руки

Після того, як рука була знайдена, із регіону зв'язної компоненти видаляються всі інші компоненти, колір помітки яких відрізняється від кольору помітки компоненти руки. Це робиться для того, щоб позбавитись від непотрібного сміття і в наслідку для покращення точності алгоритму.

Далі локалізований регіон масштабується в зображення формату  $100 \times 100$ , після чого на отриманому зображенні видаляються пустоти, робиться це за допомогою методів ерозії та дилатації.

Дилатація (морфологічне розширення) – згортка зображення або виділеної області зображення з деяким ядром. Ядро може мати довільну форму розмір. При цьому в ядрі виділяється єдина ведуча позиція, що суміщається з поточним пікселем при розрахунку згортки. В багатьох випадках в якості ядра обирається квадрат або круг з ведучою позицією в центрі. Ядро можна розглядати як шаблон або маску. Застосування

дилатації зводиться до проходу шаблону по всьому зображенню і застосуванню оператора пошуку локального максимуму до інтенсивностей пікселів зображення, що накриваються шаблоном. Така операція викликає ріст світлих областей на зображенні.

Ерозія (морфологічне звуження) — зворотна операція до дилатації. Дія ерозії подібна дилатації різниця полягає лиш в тому, що використовується оператор пошуку локального мінімуму.

Комбінування цих методів при застосування дозволяє позбавитись шумів і згладити контур зображення. Після ліквідації пустот, на зображенні виділяється контур за наступним алгоритмом:

1. У 8-зв'язній області помічаються точки, у яких є всі сусіди білого кольору ба помічені на попередніх кроках.

2. Помічені точки видаляються.

Коли було виділено контур жесту, відбувається порівняння його із набором жестів-шаблонів за допомогою методу відстаней Гаусдорфа. Алгоритм полягає у знаходженні критерію схожості двох контурів, що приймає значення найменшої відстані між шуканим об'єктом і шаблоном. Якщо цей критерій переважає задане допустиме значення — вхідне зображення ігнорується, у іншому випадку зображення вважаються схожими і програма відображає знайдений шаблон, після чого виконується наступна ітерація циклу. Результати роботи програми зображені на рисунку 4.

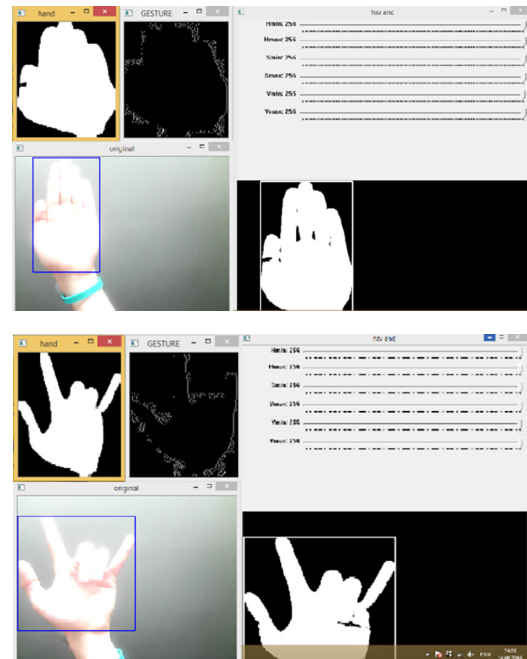


Рисунок 4. Результати роботи програми

### Висновки

В роботі коротко розглянуті обрані методи для локалізації та розпізнавання жестів руки. Описаний загальний алгоритм роботи програми. Наведено приклад роботи розробленої програми.

### Перелік посилань

1. R. Tyrrell Variational Analysis / R. Tyrrell // Springer-Verlag. — 2005. № 17. — P. 117.
2. Дж. Стокман Компьютерное зрение / Дж. Стокман, Л. Шапиро. — М.: БИНОМ. Лаборатория знаний, 2006. — 36 с.
3. Munkres Topology / Munkres, James // Prentice Hall. — 1999. — P. 280–281.