

УДК 004.92

Управление компьютером жестами с использованием веб-камеры: инновационный подход к взаимодействию

Борисов Никита Сергеевич

bns40@mail.ru

Белов Юрий Сергеевич

ysbelov@bmstu.ru

КФ МГТУ им. Н.Э. Баумана, Калуга, Россия

Статья посвящена разработке и исследованию системы управления компьютером с помощью жестов, использующей веб-камеру. Рассмотрены современные подходы к взаимодействию человека с компьютером, акцентируя внимание на преимуществах естественных интерфейсов перед традиционными устройствами ввода. Описаны методы анализа изображений и алгоритмы распознавания жестов, основанные на компьютерных знаниях и машинном обучении. Предложена архитектура системы, включающая модуль захвата и обработки изображения, детектор рук и классификатор жестов. Тестирование показало среднюю точность распознавания 85 % при времени отклика до 100 мс. Обозначены перспективы дальнейшего совершенствования технологии, включая интеграцию с дополненной реальностью и глубоким обучением.

Ключевые слова: управление жестами, компьютерное зрение, веб-камера, человеко-компьютерное взаимодействие, алгоритмы обработки изображений

Современные технологии взаимодействия человека с компьютером требуют новых решений, которые были бы более естественными и удобными для пользователей. Управление жестами с помощью веб-камеры представляет собой перспективную альтернативу традиционным устройствам ввода. Эта технология особенно актуальна в условиях, когда использование мыши и клавиатуры затруднено, например во время презентаций, в образовательных процессах или для людей с ограниченными возможностями.

Целью данного исследования является анализ существующих методов управления компьютером жестами, разработка алгоритмов распознавания жестов с использованием веб-камеры и оценка эффективности предложенного решения. Особое внимание уделяется вопросам точности распознавания, адаптации к различным условиям освещения и интеграции с популярными операционными системами [1].

Для реализации системы управления жестами были использованы методы компьютерного зрения и машинного обучения. Основные этапы работы включают:

- обнаружение руки на изображении с использованием цветового пространства HSV и методов сегментации;
- выделение контура руки и ключевых точек (пальцев, ладони) с применением алгоритмов обнаружения краев и выпуклых оболочек;
- классификация жестов на основе геометрических характеристик руки и ее движений;

– преобразование жестов в команды операционной системы (перемещение курсора, клики, прокрутка) [2].

Тестирование системы проводилось на операционных системах Windows, Linux и macOS с использованием веб-камер различного разрешения.

Предложенная система состоит из двух основных модулей.

Модуль распознавания жестов:

- захват изображения с веб-камеры;
- предварительная обработка (фильтрация, повышение контрастности);
- детектирование руки и выделение ее контура;
- классификация жестов с помощью алгоритмов машинного обучения.

Модуль выполнения команд:

- преобразование распознанных жестов в соответствующие действия (например, перемещение курсора или клик);
- интеграция с API операционных систем для эмуляции ввода [3].

Система была протестирована в различных условиях освещения и с разными пользователями. Точность распознавания жестов составила в среднем 85 %, при этом время отклика не превышало 100 мс. Наилучшие результаты были достигнуты при хорошем освещении и использовании камер с разрешением не менее 720p.

Несмотря на успешную реализацию, система сталкивается с рядом ограничений:

- зависимость от освещения: качество распознавания снижается при слабом или неравномерном освещении;
- ограниченный набор жестов: текущая версия системы поддерживает базовые жесты (перемещение, клики, прокрутка);
- вычислительная нагрузка: обработка изображений в реальном времени требует значительных ресурсов CPU [4].

Перспективы развития технологии включают:

- использование глубокого обучения для повышения точности распознавания;
- интеграцию с дополненной реальностью (AR) для расширения функциональности;
- разработку универсального SDK для упрощения внедрения в различные приложения.

Управление компьютером жестами с использованием веб-камеры представляет собой многообещающую технологию, которая может значительно улучшить взаимодействие человека с компьютером. Несмотря на существующие ограничения, дальнейшее развитие алгоритмов компьютерного зрения и машинного обучения позволит создать более надежные и универсальные решения. Эта технология имеет большой потенциал для применения в различных областях, от образования до индустрии развлечений.

Литература

- [1] Борисов Н.С., Белов Ю.С. Анализ методов для управления персональным компьютером жестами. *Научные технологии в приборостроении и развитии инновационной деятельности в вузе. Всерос. науч.-техн. конф.: сб. матерв.* Москва, Изд-во МГТУ им. Н.Э. Баумана, 2023, с. 313–316.
- [2] Williamson A. Vision based cursor control using hand gestures and a webcam. The University of the West Indies St. Augustine, 2009.
- [3] Борисов Н.С., Белов Ю.С. Анализ методов обработки действий курсора на программном уровне в различных операционных системах. *Научные технологии в приборостроении и развитии инновационной деятельности в вузе. Рег. науч.-техн. конф.: сб. матер.* Москва, Изд-во МГТУ им. Н.Э. Баумана, 2024, т. 2, с. 77–82.
- [4] Борисов Н.С., Белов Ю.С. Применение нейроядра процессора для оптимизации алгоритмов машинного зрения. *Теоретические основы и практическое применение инновационных исследований: матер.* Санкт-Петербург, Изд-во ГНИИ «НАЦРАЗВИТИЕ», 2024, с. 9–11.

Gesture control of your computer with a webcam: an innovative approach to interaction

Borisov Nikita Sergeevich

bns40@mail.ru

Belov Yuri Sergeevich

ysbelov@bmstu.ru

Kaluga Branch of BMSTU, Kaluga, Russia

The article is devoted to the development and research of a computer control system using gestures using a webcam. Modern approaches to human-computer interaction are considered, focusing on the advantages of natural interfaces over traditional input devices. The paper describes image analysis methods and gesture recognition algorithms based on computer knowledge and machine learning. The architecture of the system is proposed, including an image capture and processing module, a hand detector and a gesture classifier. Testing showed an average recognition accuracy of 85 % with a response time of up to 100 ms. The prospects for further improvement of the technology, including integration with augmented reality and deep learning, are outlined.

Keywords: gesture control, computer vision, webcam, human-computer interaction, image processing algorithms