

УДК 004.932:004.5:004.89

СИСТЕМА РАСПОЗНАВАНИЯ ЖЕСТОВ И УПРАВЛЕНИЯ КОМПЬЮТЕРОМ С ИСПОЛЬЗОВАНИЕМ КОМПЬЮТЕРНОГО ЗРЕНИЯ

Шкуринский В.А., Фролов И.И.

*Белорусский государственный университет информатики и радиоэлектроники, г. Минск, Беларусь,
valik292@mail.ru*

Аннотация. Рассмотрены аспекты разработки программного обеспечения для управления приложениями компьютера с помощью жестов, основанного на применении технологий компьютерного зрения и машинного обучения.

Ключевые слова. Распознавание жестов, обработка изображений, программирование, машинное обучение.

В современном мире цифровая доступность становится все более важной и актуальной. Развитие технологий меняет нашу жизнь, делая её более комфортной и удобной. Согласно исследованию ЮНЕСКО, более 1 миллиарда людей – то есть около 15 % всего населения земного шара [1], живут с каким-то видом инвалидности, и многие из них испытывают трудности с использованием стандартного компьютерного оборудования. В связи с этим разработку систем, позволяющих управлять компьютером без необходимости использования стандартной клавиатуры и мыши, становится важной задачей.

Для преодоления этой проблемы разрабатываются различные технологии, направленные на улучшение цифровой доступности. Например, проекты, такие как «Microsoft's Seeing AI» и «Google's Lookout», используют машинное обучение и компьютерное зрение, которые помогают людям с ограниченными возможностями в повседневной жизни, включая чтение текста, распознавание объектов и навигацию.

Однако, несмотря на усилия и разработки в этой области, многие технологии все еще остаются недоступными для значительной части населения. В этом контексте разработка систем распознавания жестов и управления компьютером с применением компьютерного зрения представляет собой важный шаг в повышении доступности и удобства использования цифровых технологий для всех людей, независимо от их физических возможностей.

Для людей с ограниченными физическими возможностями взаимодействие с компьютером может представлять собой сложную задачу в связи с рядом препятствий, с которыми они сталкиваются. Некоторые из основных проблем включают в себя следующее:

1. Ограниченная подвижность: Люди с ограниченной подвижностью могут испытывать трудности при использовании стандартной клавиатуры и мыши, особенно если у них имеются проблемы с моторикой или координацией движений.

2. Недоступность стандартного оборудования: Устройства стандартного ввода, такие как клавиатура и мышь, могут оказаться неадаптированными для использования людьми с ограниченными физическими возможностями в силу их размеров, формы или расположения кнопок. Кроме того, стандартное компьютерное оборудование зачастую не предусматривает удобства для пользователей с ограниченными физическими возможностями. Размеры, форма и расположе-

ние кнопок клавиатуры и мыши могут оказаться неудобными или недоступными для использования.

3. Низкая чувствительность сенсорных устройств: Для определенных индивидуумов использование сенсорных устройств может быть затруднительным в силу их низкой чувствительности или отсутствия тактильной обратной связи.

4. Временные ограничения в связи с травмой: Люди, временно лишенные возможности использования рук в результате травмы или других обстоятельств, также могут столкнуться с трудностями в управлении компьютером при помощи стандартных устройств ввода.

В последние годы наблюдается значительный прогресс в разработке систем управления компьютером с использованием жестов, направленных на решение проблем, с которыми сталкиваются люди с ограниченными физическими возможностями. Ряд значимых и широко распространенных решений включает в себя следующее:

1. Microsoft Kinect: Kinect, разработанный компанией Microsoft, представляет собой устройство, основанное на технологии глубокого зрения, способное распознавать движения и жесты человеческого тела. Первоначально созданное для использования в игровых консолях, оно быстро нашло применение в робототехнике и визуализации трехмерных ландшафтов. Например, студенты Технологического института Джорджии разработали платформу SoryCat для сбора данных о жестах, которая используется для системы распознавания американского языка жестов (ASL), а также как практическое приложение для развития рабочей памяти и языковых навыков у глухих детей в игровой форме [2].

2. Leap Motion: Leap Motion представляет собой компактное устройство для размещения на столе, которое способно отслеживать движения рук и пальцев в трехмерном пространстве. Оно обеспечивает точное и быстрое распознавание жестов, что делает его идеальным для управления компьютером без необходимости использования клавиатуры или мыши.

Гендиректор Leap Motion Michael Buckwald сообщил журналу CNET:

«Мы хотим создать приложение, которое перевернет мир, которое полностью изменит взаимодействие человека с операционной системой или серфингом в Интернете. Наша цель — коренным образом изменить взаимодействие компьютера и человека, то есть делать



всё то, что раньше делалось мышью. Несомненно это отразится на каждом, начиная от простых повседневных задач и заканчивая любым технологическим процессом, который только можно представить» [3].

3. EyeSight Technologies представляет собой организацию, занимающую лидирующие позиции в разработке встроенных решений в области компьютерного зрения и глубокого обучения. Их технологии нацелены на улучшение взаимодействия пользователей с широким спектром устройств и отраслей, путем повышения осведомленности пользователей и распознавания жестов. Используя методы глубокого обучения и искусственного интеллекта, встроенные решения от EyeSight значительно облегчают использование различных устройств, начиная от домашней электроники и заканчивая автомобильными системами.

Технология компьютерного зрения от EyeSight предоставляет интеллектуальные и персонализированные функции, улучшая взаимодействие с «умным домом», Интернетом, автомобильными системами, гарнитурами виртуальной реальности и другими аппаратными средствами. Патентованная технология компьютерного зрения компании предлагает широкий спектр бесконтактных взаимодействий, включая активное управление с распознаванием жестов и отслеживанием движений пальцев, а также пассивное распознавание пользователей и их действий для запуска заранее определенных функций [4].

Эти и прочие разработки иллюстрируют потенциал технологий распознавания жестов в контексте создания более доступных и удобных интерфейсов управления компьютером и другими системами. Это открывает новые перспективы для индивидуумов с ограниченными физическими возможностями.

Программное обеспечение, разрабатываемое в рамках дипломного проекта и описываемого в данной публикации содержит модули, написанные на языке Python. Приложение использует искусственный интеллект для обучения распознаванию жестов руки, обладает разнообразным спектром функциональных возможностей для управления компьютером. Некоторые из основных характеристик включают в себя следующее:

1. Распознавание жестов руки: Программа обучается распознавать различные жесты руки, включая поднятый один или два указательных пальца, кулак, открытую ладонь и другие, используя алгоритмы машинного обучения.

2. Выполнение действий на основе распознанных жестов: В реальном времени программа анализирует видеопоток с веб-камеры с целью распознавания жестов рук и, в зависимости от обнаруженного жеста, выполняет определенные действия на компьютере, такие как открытие диспетчера задач, остановка видео или песни, регулировка громкости звука, сворачивание окон, прокрутка страниц и другие основные команды для браузера.

3. Многозадачность и гибкость: Программа способна обрабатывать одновременно несколько жестов и осуществлять соответствующие действия, обеспечивая гибкость в управлении компьютером.

4. Персонализация и настройка: Пользователи могут настраивать программу для изменения команд и действий, связанных с каждым жестом, в соответствии с их индивидуальными потребностями и предпочтениями.

5. Удобство использования и доступность: Особое внимание уделяется удобству использования программы, чтобы обеспечить максимальную доступность для пользователей с различным уровнем навыков и физическими возможностями.

Цифровая доступность особенно востребована для приобретения профессиональных навыков, позволяющих людям с ограниченными физическими способностями, в том числе связанными с невозможностью использовать руками средства управления компьютером, использовать цифровые ресурсы для обучения, получения профессии и самообеспечения.

Одной из наиболее востребованных функций при работе с компьютером является использование интернет-ресурсов для изучения обучающей информации и получения обратной связи от преподавателей, например, на сайте университета. Простая для обычных пользователей процедура открытия браузера и перемещения по страницам ресурса является нетривиальной для пользователей, имеющих ограниченные физические возможности, связанные с использованием рук.

Для пользователя с ограниченными физическими возможностями, который стремится получить доступ к учебным ресурсам на веб-сайте университета, разрабатываемое программное обеспечение предоставляет следующий набор функциональных возможностей согласно алгоритму действий:

1. Инициировать запуск веб-браузера: С помощью специального жеста руки пользователь может запустить веб-браузер, такой как Google Chrome или Mozilla Firefox.

В данном случае возможно 2 варианта реализации: закрепление комбинации «горячих клавиш», соответствующее физическому нажатию на клавиши клавиатуры и запускающему связанный с данной командой браузер, либо же последовательное использование набора жестов для запуска меню «Пуск», пролистывания списка приложения, выбора требуемого браузера и нажатия клавиши «Ввод». В обоих случаях результатом станет запуск браузера, позволяющего использовать интернет-ресурсы.

2. Активировать поисковый сервис: После открытия браузера пользователь может аналогичным образом воспользоваться жестами для активации поисковика, например, Google или Яндекс: либо используя заранее обученный и связанный с набором «горячих клавиш» для ввода адреса в строке браузера предустановленного поисковика жест, либо используя жесты запуска «виртуальной клавиатуры», на которой используя жесты перемещения и нажатия на «виртуальные клавиши» последовательно ввести адрес поисковика или же требуемого интернет-ресурса, такого, как обучающий портал БГУИР (Система электронного обучения – СЭО).

3. Доступ к Системе электронного обучения: После загрузки веб-сайта университета пользователь

может перейти на страницу обучения, где располагаются учебные материалы и курсы, используя жесты рук для перемещения и навигации по странице: движения курсора мыши, переходы по спискам / таблицам с использованием эмуляции нажатия на клавиши стрелки «вверх» / «вниз» / «вправо» / «влево».

4. Закрывать страницы/браузер: По окончании работы с учебными материалами пользователь может закрыть вкладки или даже весь браузер с помощью соответствующих жестов рук.

Основной функционал, описываемый в публикации базируется на использовании API PyAutoGUI. В PyAutoGUI реализованы следующие функциональные возможности [5]:

- управление перемещением курсора мыши и эмуляция нажатия кнопок в окнах других приложений;

- отправка последовательности нажатий клавиш в приложения, например, для автоматизации процесса заполнения форм;

- захват изображений экрана и анализ их содержимого, такого как кнопки или флажки, для последующего их обнаружения на экране.

- определение и управление окнами приложений, включая их перемещение, изменение размера, максимизацию, минимизацию или закрытие (на данный момент реализовано только для операционной системы Windows).

- отображение диалоговых окон для предупреждений и сообщений в процессе выполнения скриптов.

Основные функции API, использованные при разработке проекта перечислены в таблице ниже:

Таблица 1 – Запрограммированные комбинации клавиш

Действие	Команда
Нажатие клавиши 'Enter'	<code>pyautogui.press('enter')</code>
Нажатие клавиши 'Вверх'	<code>pyautogui.keyDown('up')</code>
Нажатие клавиши 'Вниз'	<code>pyautogui.keyDown('down')</code>
Нажатие клавиши 'Вправо'	<code>pyautogui.keyDown('right')</code>
Нажатие клавиши 'Влево'	<code>pyautogui.keyDown('left')</code>
Проскроллить вниз	<code>pg.scroll(-100)</code>
Проскроллить вверх	<code>pg.scroll(+100)</code>
Закрытие активного приложения	<code>pg.hotkey('alt', 'f4')</code>
Открытие окна поиска	<code>pg.hotkey('win', 's')</code>

DEVELOPMENT OF SOFTWARE FOR VIDEO SURVEILLANCE SYSTEMS DESIGN AND ANALYSIS

V.A. Shkurinsky, I.I. Frolov

Belarusian State University of Informatics and Radioelectronics, Minsk, Belarus, valik292@mail.ru

Abstract. The aspects of software development for computer application control using gestures, based on computer vision and machine learning technologies, are considered.

Keywords. Gesture recognition, image processing, programming, machine learning.

Используя в определенных последовательностях и комбинациях перечисленные вариации эмуляций нажатия клавиш можно реализовать более сложные комплексные действия по управлению доступом к ресурсам компьютера.

Для обработки входящего видеопотока жестов, поступающего с камеры, были использованы библиотеки OpenCV для захвата и обработки входных данных, а также нейронные сети, поставляемые в рамках комплексной платформы машинного обучения TensorFlow, для распознавания изображений и сопоставления их с комбинациями клавиш.

Таким образом, разрабатываемое программное решение позволяет упростить доступ и управление для людей с ограниченными физическими возможностями, предоставляя функционал по распознаванию жестов рук и их интерпретации программой в действия по управлению операционной системой и другими приложениями.

Литература

1. Уважение прав инвалидов [Электронный ресурс] – Режим доступа: <https://www.unesco.org/ru/articles/uvazhenie-prav-invalidov>

2. How Does The Xbox Kinect Work [Электронный ресурс] – Режим доступа: <https://www.jameco.com/Jameco/workshop/Howitworks/xboxkinect.html>

3. Leap Motion [Электронный ресурс] – Режим доступа: https://ru.wikipedia.org/wiki/Leap_Motion

4. EyeSight [Электронный ресурс] – Режим доступа: <https://israelmobileinnovation.com/exhibitors/eyesight/>

5. Перевод документации по PyAutoGUI [Электронный ресурс] – Режим доступа: <https://pyautogui.readthedocs.io/ru/latest/>