

**МИНИСТЕРСТВО ОБРАЗОВАНИЯ РЕСПУБЛИКИ БЕЛАРУСЬ
БЕЛОРУССКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ФАКУЛЬТЕТ РАДИОФИЗИКИ И КОМПЬЮТЕРНЫХ ТЕХНОЛОГИЙ
Кафедра интеллектуальных систем**

Аннотация к дипломной работе

**СИСТЕМА УПРАВЛЕНИЯ МУЛЬТИМЕДИА НА ОСНОВЕ
РАСПОЗНАВАНИЯ ЖЕСТОВ**

Кулащик Дмитрий Русланович

Научный руководитель: канд. физ.-мат. наук, доцент А.И.Головатый

Минск, 2025

РЕФЕРАТ

Дипломная работа содержит 50 с., 27 рис., 4 табл., 14 источн., 6 прил.

КОМПЬЮТЕРНОЕ ЗРЕНИЕ, OPENCV, MEDIAPIPE, PYTHON, МАШИННОЕ ОБУЧЕНИЕ, РАСПОЗНАВАНИЕ ЖЕСТОВ, НЕЙРОННЫЕ СЕТИ.

Цель работы - разработка системы на языке программирования Python для управления интерактивным дисплеем жестами при помощи алгоритмов компьютерного зрения с элементами проверки доступности по биометрическим данным.

В ходе выполнения данной работы были изучены существующие методики разработки подобных программных решений, проведён анализ современных подходов к компьютерному зрению, использовались библиотеки OpenCV и фреймворк MediaPipe, обеспечивающие обработку изображений и реализацию алгоритмов машинного обучения.

Интерес к технологиям компьютерного зрения для решения различных задач значительно возрос. Одним из таких направлений является распознавание жестов, которое играет все большую роль в современном мире. Кроме того, жесты обеспечивают удобный способ взаимодействия между человеком и компьютером. Распознавание жестов относится к области информатики и языковых технологий, где применяются математические алгоритмы для интерпретации движений человека. Эти движения, как правило, исходят от рук или лица, но могут включать любые части тела. Успешное освоение распознавания жестов открывает широкие перспективы для решения множества прикладных задач, поскольку руки и тело человека являются естественными манипуляторами с большим количеством степеней свободы.

Одним из ключевых вызовов при разработке подобных систем является обеспечение точности и скорости распознавания жестов в реальном времени. Это требует не только продвинутых алгоритмов машинного обучения, но и оптимизации обработки видеопотока, чтобы минимизировать задержки и повысить удобство взаимодействия пользователя с устройством. Кроме того, успешная интеграция таких технологий должна учитывать эргономику движений человека, чтобы сделать процесс управления естественным и интуитивно понятным.

РЭФЕРАТ

Дыпломная праца змяшчае 50 с., 27 мал., 4 табл., 14 крыніц, 6 дад.

КАМП'ЮТАРНЫ ЗРОК, OPENCV, MEDIAPIPE, PYTHON, МА-ШЫН-
НАЯ НАВУЧАННЕ, РАЗВЯЗАННЕ Жэстаў, НЕЙРОННЫЯ СЕ-ЦІ.

Мэта працы - распрацоўка сістэмы на мове праграмавання Python для кіравання інтэрактыўным дысплеем жэстамі пры дапамозе алгарытмаў кампутарнага зроку з элементамі праверкі даступнасці па біяметрычных дадзеных.

У ходзе выканання гэтай работы былі вывучаны існуючыя метады распрацоўкі такіх праграмных рашэнняў, праведзены аналіз сучасных падыходаў да камп'ютэрнага зроку, выкарыстоўваліся бібліятэкі OpenCV і фрэймворк MediaPipe, якія забяспечваюць апрацоўку малюнкаў і рэалізацыю алгарытмаў машыннага навучання.

Цікавасць да тэхналогій камп'ютэрнага зроку для вырашэння розных задач значна ўзрасла. Адным з такіх напрамкаў з'яўляецца распазнанне жэстаў, якое адыгрывае ўсё большую ролю ў сучасным свеце. Акрамя таго, жэсты забяспечваюць зручны спосаб узаемадзеяння паміж чалавекам і кампутарам. Распазнанне жэстаў адносіцца да галіны інфарматыкі і моўных тэхналогій, дзе прымяняюцца матэматычныя алгарытмы для інтэрпрэтацыі рухаў чалавека. Гэтыя рухі, як правіла, зыходзяць ад рук або асобы, але могуць уключаць любыя часткі цела. Паспяховае засваенне распазнання жэстаў адкрывае шырокія перспектывы для рашэння мноства прыкладных задач, паколькі рукі і цела чалавека з'яўляюцца натуральнымі маніпулятарамі з вялікай колькасцю ступеняў свабоды.

Адным з ключавых выклікаў пры распрацоўцы падобных сістэм з'яўляецца забеспячэнне дакладнасці і хуткасці распазнання жэстаў у рэальным часе. Гэта патрабуе не толькі прасунутых алгарытмаў машыннага навучання, але і аптымізацыі апрацоўкі відэастрому, каб мінімізаваць затрымкі і па-высіць зручнасць узаемадзеяння карыстальніка з прыладай. Акрамя таго, паспяховая інтэграцыя такіх тэхналогій павінна ўлічваць эрганоміку рухаў чалавека, каб зрабіць працэс кіравання натуральным і інтуітыўна зразумелым.

ABSTRACT

The thesis contains 50 pp., 27 fig., 4 tables, 14 sources, 6 appendixes.

COMPUTER VISION, OPENCV, MEDIAPIPE, PYTHON, MACHINE LEARNING, GESTURE RECOGNITION, NEURAL NETWORKS.

The aim of the thesis is to develop a system in the Python programming language for controlling an interactive display with gestures using computer vision algorithms with elements of accessibility testing based on biometric data.

In the course of this work, existing methods for developing similar software solutions were studied, an analysis of modern approaches to computer vision was carried out, OpenCV libraries and the MediaPipe framework were used, providing image processing and implementation of machine learning algorithms.

Interest in computer vision technologies for solving various problems has increased significantly. One of these areas is gesture recognition, which plays an increasingly important role in the modern world. In addition, gestures provide a convenient way for humans and computers to interact. Gesture recognition is a field of computer science and language technology that uses mathematical algorithms to interpret human movements. These movements typically originate from the hands or face, but can involve any part of the body. Successful mastery of gesture recognition opens up broad prospects for solving many applied problems, since the human hands and body are natural manipulators with a large number of degrees of freedom.

One of the key challenges in developing such systems is ensuring the accuracy and speed of gesture recognition in real time. This requires not only advanced machine learning algorithms, but also optimization of video stream processing to minimize delays and improve the user experience with the device. In addition, successful integration of such technologies should take into account the ergonomics of human movements to make the control process natural and intuitive.