**Департамент образования мэрии города Новосибирска**

**Дворец творчества детей и учащейся молодежи «Юниор»**

**XLII городская открытая научно-практическая**

**конференция НОУ «Сибирь»**

Секция: программирование и искусственный интеллект

**Управление презентацией жестами**

Автор:

**Демидович Ольга Юрьевна,**

10 класс МАОУ «Инженерный лицей НГТУ»

Ленинского района

города Новосибирска

конт.тел: +7 909-531-31-45

Научный руководитель:

Гончаров Иван Олегович, инженер машинного обучения  
Weights&Biases

Новосибирск 2022

**Оглавление**

[Введение 3](#_Toc123941932)

[Теоретическая часть 4](#_Toc123941933)

[Практическая часть 5](#_Toc123941934)

[Заключение 7](#_Toc123941935)

[Список используемой литературы: 8](#_Toc123941936)

[Приложения 9](#_Toc123941937)

# Введение

Презентация – одно из самых частых мероприятий современного человека. Она является простым и удобным способом передачи информации, поскольку располагает простором для интересных и нестандартных решений.

Для создания и демонстрации презентаций используются различные программы, самые популярные из них: PowerPoint, Keynote, Google Slides, Prezi, OpenOffice Impress, Zoho Show, Piktochart, Slides, Canva, Haiku Deck, Powtoon. По данным Microsoft Office на 2015 год пользователей только PowerPoint насчитывалось 1,5 миллиарда человек.

Успех презентации зависит не только от материала на слайде и ораторского мастерства рассказчика, но и от коммуникации с аудиторией, которая очень часто затруднена привязкой докладчика к клавиатуре компьютера. Однако в последнее время эта проблема стала решаться появлением дополнительных девайсов, таких как кликер. Он позволяет спикеру свободно передвигаться по аудитории.

В 2022 году, во время регионального трека Всероссийского конкурса научно-технологических проектов «Большие вызовы» данное устройство меня подвело. Во время моей защиты проекта на кликере села батарея, от чего возник определённый дискомфорт, и дальнейшая презентация свелась к переключению слайдов техническим специалистом, и в моей презентации вынуждено появилась фраза: «Следующий слайд». На моём опыте это было не единственным казусом при проведении демонстрации. Данная проблема меня озадачила.

Целью моей работы было создание программного обеспечения, которое освободило бы руки от дополнительного устройства, так как программу можно в любой момент скачать, не надо носить для неё зарядное устройство

Задачи:

1. Изучить проблемы спикера при проведении презентации
2. Изучить существующие решения и выявить в них недостатки
3. Подобрать технологии, приводящие к выполнению цели
4. Разработать программное обеспечение
5. Протестировать программу
6. Сравнить своё решение с существующими

Проект направлен на использование в университетах и школах при проведении лекций. На предприятиях и организациях при проведении собраний. На конференциях и прочих мероприятиях, которые предполагают защиту проектов или продуктов, выступление с презентацией.

В связи с вышеизложенным актуальность темы очевидна. Проект позволит вывести презентации на новый уровень.

# Теоретическая часть

Для работы программы используется следующий стек технологий (стек технологий — это набор языков программирования, фреймворков и ПО, необходимых для разработки приложения):

Python — это скриптовый язык программирования. Он универсален, поэтому подходит для решения разнообразных задач и многих платформ, начиная с iOS и Android и заканчивая серверными ОС.

OpenCV — это open source библиотека компьютерного зрения, которая предназначена для анализа, классификации и обработки изображений. OpenCV написана на языке высокого уровня (C/C++) и содержит алгоритмы для: интерпретации изображений, калибровки камеры по эталону, устранение оптических искажений, определение сходства, анализ перемещения объекта, определение формы объекта и слежение за объектом, 3D-реконструкция, сегментация объекта, распознавание жестов и т.д.

Пакет NumPy является незаменимым помощником Python. Он тянет на себе [анализ данных](https://python-scripts.com/data-science), [машинное обучение](https://python-scripts.com/intro-to-neural-networks) и научные вычисления, а также существенно облегчает обработку векторов и матриц.

MediaPipe — это фреймворк с открытым исходным кодом, представленный Google, который помогает создавать мультимодальные конвейеры машинного обучения. Разработчик может создать прототип, не углубляясь в написание алгоритмов и моделей машинного обучения, используя существующие компоненты. Эта структура может использоваться для различных приложений для обработки изображений и мультимедиа (особенно в виртуальной реальности), таких как обнаружение объектов, распознавание лица (Mediapipe Face Mesh), отслеживание рук, пальцев (Mediapipe Hands) и прочего. MediaPipe поддерживает различные аппаратные и операционные платформы, такие как Android, iOS и Linux, предлагая API на C ++, Java, Objective-c и т.д.

PyQt — это библиотека, которая позволяет использовать фреймворк Qt GUI (GUI — это графический интерфейс пользователя) в Python. PyQt5 является одним из наиболее часто используемых модулей для создания GUI приложений.

PyAutoGUI — это кроссплатформенный модуль автоматизации Python с графическим интерфейсом для людей. Используется для программного управления мышью и клавиатурой.

# Практическая часть

Тестирование показало, что программа позволяет управлять презентациями, созданными в Microsoft Office PowerPoint, Google Slides, Prezi, LibreOffice Impress, OpenOffice презентация

Проект представляет собой окно с настройками, включающих в себя следующие пункты (рис. 1):

Выбор с флажком (QCheakBox):

* «Вывод видео»
* «Использование указки»

Выпадающие списки (QComboBox):

* Выбор руки для управления презентацией, если не выбран флажок «Использование указки».
* Иначе доступен выбор руки для управления презентацией и указки, и выбор сочетания клавиш для включения указки

Кнопки (QPushButton):

* «Применить» для сохранения настроек
* «Справка» для открытия окна информации об использовании программы (рис. 2)

Сразу после запуска программы запускается веб-камера и стандартные настройки: управление презентацией правой рукой.

Программа с помощью Mediapipe Face Mesh определяет рабочие прямоугольники, которые можно увидеть, если включить «Вывод видео». Mediapipe Hands определяет положение руки и если происходит взмах, подобный листанию книги указательным пальцем, то перелистывает слайд в указанную жестом сторону (рис. 3). Если включена указка, то её вызов можно произвести взмахом руки вверх. Тогда другой рукой будет происходить управление мышью – указкой PowerPoint. Координаты мыши определяются в большем рабочем прямоугольнике.

Преимущество и недостатки программного продукта отражены в Таблице 1. В ней сравнивается управление презентацией с помощью кликера, Presefy, AirTouch и представленного в работе программного продукта по критериям: заряд не зависит от устройства с презентацией, свободны руки от дополнительных устройств, полностью бесконтактное управление презентацией, возможность перемещаться по аудитории. Всеми известный кликер и менее известное приложение на телефон Presefy не удовлетворяют большинству из достоинств созданного предложения – не могут работать, если на дополнительном устройстве села батарея, ими занята рука, и они контактные – возможна передача бактерий и вирусов. Самым достойным конкурентом является разработка AirTouch. Эта разработка, так же, как и моя зависит от питания блока, с которого проводится презентация. При переключении презентации не используются контактные методы, но для переключения нужно оставаться в пределах видимости камеры, что не позволяет перемещаться по аудитории. Созданный программный продукт может работать, даже если в руке будет находиться микрофон, а AirTouch для своего управления требует две свободные руки, кроме того, имеет высокую стоимость: 99 тысяч рублей.

Таким образом по выбранным критериям разработанное решение является лучшим среди конкурентов.

# Заключение

Все жестовые языки уникальны и отличаются в разных странах и очень важно, чтобы моя программа развивалась именно в России, ведь это позволяет людям, владеющим русским жестовым языком, свободнее общаться. Использование программы в публичных местах, таких как магазины, государственные и медицинские учреждения - где нет возможности использовать клавиатуру.

Разработанное приложение эффективно и надежно распознает жесты русского жестового языка вне зависимости от особенностей человека, фона и освещения. Оно позволяет буквально “набирать” текст, не прикасаясь к клавиатуре и сенсорному экрану.

На примере этого проекта можно понять, как использовать Mediapipe для других проектов на основе компьютерного зрения.

Для дальнейшего развития проекта предполагается:

* Добавление жестов, которые обозначают сразу целое слово.
* Разработка учебного приложения, которое бы можно использовать для обучения русскому языку жестов.
* Выпуск перчаток, которые могли бы также считывать жесты, так как программа использует точки на руках
* Голосовая интерпретация введенных слов на языке жестов

# Список используемой литературы:

1. MediaPipe Hands

<https://google.github.io/mediapipe/solutions/hands>

1. MediaPipe face Mesh

<https://google.github.io/mediapipe/solutions/face_mesh.html>

1. Руководство по PyQt5

<https://pythonist.ru/rukovodstvo-po-pyqt5/>

1. Модуль PyAutoGUI

<https://pypi.org/project/PyAutoGUI/>

1. F

# Приложения

|  |  |
| --- | --- |
|  |  |
| а) | б) |

Рис. 1. Окно настроек а) стандартные настройки; б) при включённой указке

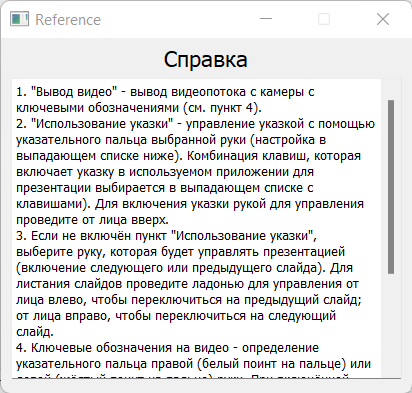


Рис. 2. Окно справки

|  |  |
| --- | --- |
|  |  |
|  |  |
|  |  |
| а) | б) |

Рис. 3. Жесты управления а) правая рука для указки, левая для управления; б) левая рука для указки, правая для управления

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **Заряд не зависит от устройства с презентацией** | **Свободны руки от дополнительных устройств** | **Полностью бесконтактное управление презентацией** | **Возможность перемещаться по аудитории** |
| **Кликер** | **-** | **-** | **-** | **+** |
| **Presefy** | **-** | **-** | **-** | **+** |
| **AirTouch** | **+** | **-** | **+** | **-** |
| **Моё предложение** | **+** | **+** | **+** | **-** |

Таблица 1

Сравнение с существующими решениями