Доклад к защите дипломного проекта

Ступакевича Максима

На тему: «Реализация системы HMI для управления оборудование на основе жестовых команд»

Обычно, когда мы хотим что-то сделать, то всегда сопровождаем это какими-либо физическими действиями для достижения ожидаемого результата. Например, протягиваем руку, чтобы взять нужный предмет или присаживаемся, чтобы завязать шнурки, поднимаем руку, чтобы обратить на себя внимание, поворачиваем голову, чтобы лучше что-то услышать. И в этом смысле жесты или движения куда более интуитивны чем статическое нажатие на кнопку вкл/выкл. Поэтому переход к жестам кажется таким же логичным как, в своё время, переход от кнопок к сенсору.

Идентификация жестов человека по видеопотоку для управления оборудованием имеет следующий ряд преимуществ по сравнению с другими методами управления:

1.Нет необходимости в других устройствах ввода. Камера имеет не только коммуникативную функцию, но и контролирующую.

2.Не нужен физический контакт с какими-либо устройствами, достаточно лишь находиться в области видимости камеры. Тем самым физический износ является минимальным.

Помимо достоинств компьютерное распознавание жестов имеет и свои недостатки:

1.Сам по себе один лишь способ не гарантирует стопроцентного распознавания.

2.При различном освещении матрица камеры может быть пересвечена.

С постоянным и стремительным ростом вычислительных мощностей компьютера стало возможным обучать и использовать различные методы машинного обучения, в том числе глубокие нейронные сети.

Чтобы распознать жест, для начала нужно распознать саму кисть руки на изображении. Оказывается, уже существуют реализованные алгоритмы, которые умеют распознавать человеческую ладонь в видеопотоке, приходящем от простой веб-камеры компьютера. В дипломном проекте используется Google Mediapipe Hands. Работает это точно так же, как и лицом. Алгоритм выделяет ключевые точки. Для одной кисти руки это 21 точка с координатами в 3х-мерном пространстве.

Целью данного проекта является разработка программного HMI (интерфейс Человек-Машина) для распознавания жестов в видеопотоке и выполнением команд компьютера, согласно распознанному жесту.

Жест - последовательное изменения положения руки, имеющее определенное значение. Таким образом, для распознавания жеста в видеопотоке необходимо учитывать последовательность дискретных положений ключевых точек кистей рук в пространстве. Такая задача, определение жеста, в контексте машинного обучения является задачей обучения с учителем, а именно Классификацией. Одним из алгоритмов машинного обучения для классификации и учитывающего последовательности некоторых значений является Рекуррентная нейронная сеть LSTM (Long short-term Memory - долгая краткосрочная память).

Таким образом, для реализации системы HMI сформулировано требование решить следующие задачи:

1.Разработать инструмент для сбора данных.

2.Собрать данные для обучения нейронной сети.

3.Провести анализ и обработку собранных данных для подготовки датасета.

4.Обучить рекуррентную нейронную сеть LSTM для классификации.

5.Разработать приложение системы HMI для идентификации жестовых команд в видеопотоке.

А сама система должна работать по следующей схеме (Слайд 2):

1. Получение последовательности изображений.
2. Извлечение ключевых точек кистей рук.
3. Предсказание класса жеста.
4. При превышении порога срабатывания выполнить команду.

Исходя из этого выполнена следующая работа:

1. Реализовано приложение для сбора данных (Слайд 3).
2. Для распознавания сформулированы следующие жесты: Вверх, вниз, влево, вправо, ОК, назад (Слайд 4).
3. Собранные данные были проанализированы и обработаны (Слайд 4).
4. Для обучения сконфигурирована и обучена следующая структура нейронной сети (Слайд 5).
5. В конечном итоге разработано приложение системы HMI, которое использует обученную нейронную сеть для идентификации жестовых команд в видеопотоке (Слайд 6).