

РАЗРАБОТКА ИНТЕЛЛЕКТУАЛЬНОЙ СИСТЕМЫ СУРДОПЕРЕВОДА «ССП-1» СРЕДСТВАМИ БИБЛИОТЕКИ КОМПЬЮТЕРНОГО ЗРЕНИЯ OPENCV

Артём Андреевич Шаратов

Сибирский государственный университет геосистем и технологий, 630108, Россия, г. Новосибирск, ул. Плахотного, 10, ассистент кафедры прикладной информатики и информационных систем, тел. (953)785-54-99, e-mail: sharapov_artem@mail.ru

Руслан Владимирович Гришин

Сибирский государственный университет геосистем и технологий, 630108, Россия, г. Новосибирск, ул. Плахотного, 10, обучающийся, тел. (999)463-57-64, e-mail: grihinruska@mail.ru

Рассмотрена возможность применения библиотек компьютерного зрения для распознавания жестов. Проанализировано специализированное программное обеспечение, требуемое для разработки системы сурдоперевода. Проведена настройка и тестирование общедоступных библиотек компьютерного зрения.

Ключевые слова: компьютерное зрение, программное обеспечение, машинное обучение, opencv.

DEVELOPMENT OF AN INTELLIGENT SYSTEM FOR SIGN LANGUAGE INTERPRETATION «JS1» WITH TOOLS OF THE COMPUTER VISION LIBRARY OPENCV

Artem A. Sharapov

Siberian State University of Geosystems and Technologies, 10, Plakhotnogo St., Novosibirsk, 630108, Russia, Assistant, Department of Applied Informatics and Information Systems, phone: (953)785-54-99, e-mail: sharapov_artem@mail.ru

Ruslan V. Grishin

Siberian State University of Geosystems and Technologies, 10, Plakhotnogo St., Novosibirsk, 630108, Russia, Student, phone: (999)463-57-64, e-mail: grihinruska@mail.ru

The possibility of using computer vision libraries for sign language interpretation. The specialized software required for the development of sign language interpretation system is analyzed. The configuration and testing of public libraries of computer vision are carried out.

Key words: computer vision, software, machine learning, opencv.

Работа направлена на создание аппаратно-программного обеспечения, необходимого для внедрения в городскую среду, позволяющего обеспечить людям с ограниченными возможностями улучшение коммуникации с окружающим миром, позволяя с легкостью взаимодействовать с врачами или же в магазинах, аптеках и т.д.

Целью данного проекта является создание программных и аппаратных решений, позволяющих создавать благоприятные условия в коммуникации для людей с ограниченными возможностями.

Для достижения цели были поставлены следующие задачи:

- анализ библиотек компьютерного зрения;
- разработка интеллектуальной системы сурдоперевода;
- тестирование готового продукта;
- внедрение продукта.

При анализе библиотек компьютерного зрения рассматривались:

- OpenCV;
- SciPy;
- Scikit-learn;
- Tensor-flow.

Tensor-flow – библиотека разработанная компанией Google. Плюсом библиотеки является открытый исходный код, который может запускаться даже на смартфонах. Минусом же является то, что для ее изучения, а не просто запуска демо версий кода, требуются глубокие знания в области машинного обучения [1].

Scipy – данный пакет расширяет возможности работы с векторами и матрицами. Библиотека содержит такие модули как физические константы, преобразование Фурье, интегрирование и интерполяция, линейная алгебра, статистика.

Scikit-learn – это легкий способ начать свое знакомство с областью машинного обучения. Библиотека содержит различные реализации методов кластеризации, классификации и пр.

OpenCV – это стандарт в области обработки изображений и компьютерного зрения. Данная библиотека изначально разрабатывалась для обработки визуальных данных в реальном времени, что и требуется для этого проекта [2].

Для начала работы над проектом и созданием интеллектуальной системы необходимо подготовить окружение и установить библиотеку OpenCV.

Работа выполнялась в операционной системе Ubuntu 18.04, т.к. там уже предустановлен Python, была проведена установка идентификации IDE. В качестве среды для разработки использовалась PyCharm – отличная среда разработки, для работы имеются все необходимые инструменты. После настройки окружения выполнялся процесс разработки приложения [3].

Первое что необходимо сделать это включить камеру и считать видеопоток (рис. 1).

```
while camera.isOpened():
    ret, frame = camera.read()
    threshold = cv2.getTrackbarPos('trbl', 'trackbar')
    frame = cv2.bilateralFilter(frame, 5, 50, 100) # smoothing filter
    frame = cv2.flip(frame, 1) # flip the frame horizontally
    cv2.rectangle(frame, (int(cap_region_x_begin * frame.shape[1]), 0),
                  (frame.shape[1], int(cap_region_y_end * frame.shape[0])), (255, 0, 0), 2)
    cv2.imshow('original', frame)
```

Рис. 1. Считывание видео потока

Для упрощения задачи было принято решение считывать картинку в видеопотоке с небольшого участка. Чтобы точно определять контуры рук необходимо применять различные фильтры, позволяющие оставить только необходимое изображение для обрисовки контура (рис. 2).

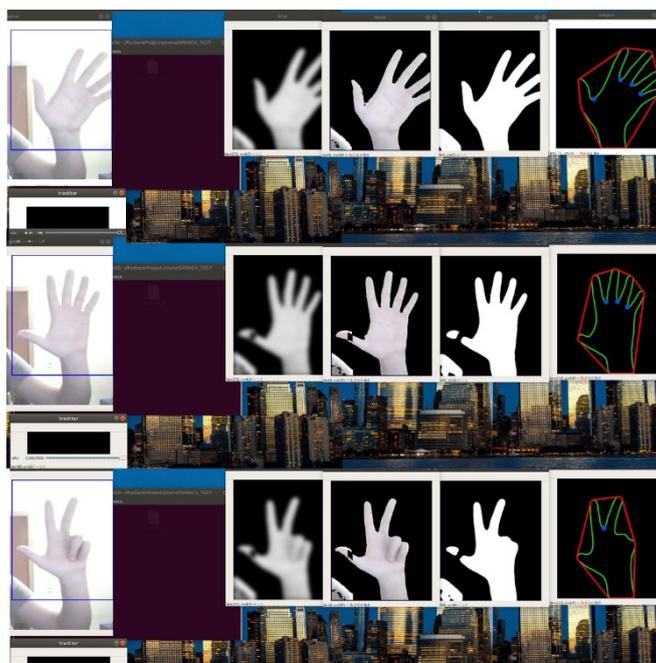


Рис. 2. Применение фильтров и обрисовка контура руки

После того как контур был отрисован, было необходимо распознать, что показывает человек и выдать текстовую информацию о том, что было показано т.е. перевод жеста в текстовый формат (рис. 3).

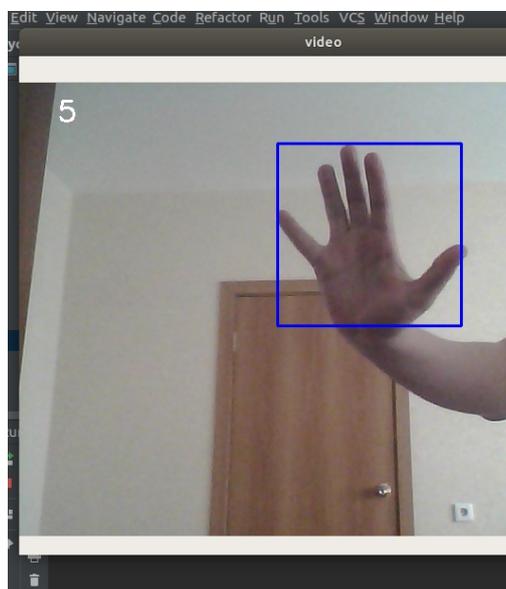


Рис. 3. Представление образа жеста в текстовом виде

Данная разработка сможет найти применение в больницах, магазинах и любых других учреждениях массового обслуживания, позволяя, потенциальным клиентам услуг, а именно лицам с ограниченными возможностями, упростить

взаимодействие с окружающей средой, повысить эффективность получения товаров или услуг.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Кэлер А., Бредски Г. Изучаем OpenCV 3. / пер. Слинкин А. А. – М. : ДМК Пресс, 2017. – 826 с.
2. Содем Я. Программирование компьютерного зрения на языке Python. Цветное издание. – М. : ДМК Пресс, 2016. – 312 с.
3. Клетте Р. Компьютерное зрение. Теория и алгоритмы / пер. Силинкин А. А. – М. : ДМК Пресс, 2019. – 506 с.

© А. А. Шаранов, Р. В. Гришин, 2019