

ИСПОЛЬЗОВАНИЕ КАМЕР INTEL REALSENSE В ПРИЛОЖЕНИИ ДЛЯ СУРДОПЕРЕВОДА

А. Я. Видман, В. В. Видман, В. М. Саклаков

Научный руководитель: доцент, к. т. н. С. В. Аксенов

Национальный исследовательский Томский политехнический университет, Россия, г. Томск,
пр. Ленина, 30, 634050

E-mail: vidman_alexandra@mail.ru

USING INTEL REALSENSE CAMERAS IN THE APPLICATION FOR SIGN LANGUAGE INTERPRETATION

A. Y. Vidman, V. V. Vidman, V. M. Saklakov

Scientific Supervisor: Associate Professor, c.e.s. S. V. Aksenov

Tomsk Polytechnic University, Russia, Tomsk, Lenin str., 30, 634050

E-mail: vidman_alexandra@mail.ru

***Abstract.** This article describes the features of Intel RealSense cameras and its using for gesture recognition and interpretation in Russian language. The article describes the existing prototype application, which will be implemented the translation function of the gestures in the text.*

Введение. В мире насчитывается большое количество глухих людей, уровень их адаптации в обществе невысок. Людям с нарушениями слуха приходится нанимать сурдопереводчиков для коммуникации с другими людьми. К этому методу они прибегают только в крайних случаях. Во-первых, в России на одного глухого человека выделяют всего 40 часов сурдоперевода в год, которые оплачивает государство. После использования этого времени данную услугу нужно оплачивать самостоятельно, но не у всех есть такая возможность. Во-вторых, было бы не целесообразно вызывать сурдопереводчика, чтобы обменяться несколькими фразами с собеседником.

Изучить язык жестов можно самостоятельно с помощью онлайн-словарей, но на многих сайтах, где словари наиболее полные, необходима оплата, ими удобно пользоваться только в домашних условиях. В настоящее время разработчики пытаются решить проблему коммуникации между глухими и слышащими людьми. Перевод с русского языка на русский жестовый язык реализован в большей степени, чем перевод с жестов в речь. Именно данный вопрос и стоит перед разработчиками. разработать систему, которая смогла бы переводить с речи на язык жестов и обратно в режиме реального времени?

Для решения данной задачи можно использовать распознавание жестов. Распознавание образов – это самая распространенная задача, которую человеку приходится решать практически ежесекундно от первого до последнего дня своего существования. Для этого он использует огромные ресурсы своего мозга, которые оцениваются таким показателем как число нейронов, равное 10^{10} [1].

В настоящее время существуют различные средства для обработки информации в режиме реального времени, что позволяет решить задачу распознавания жестов. Одной из таких технологий является компьютерное зрение. Компьютерное зрение – теория и технология создания машин, которые могут производить обнаружение, слежение и классификацию объектов [2]. Оно применяется для принятия решений о физических объектах, основываясь на изображении.

Основная часть. В настоящее время разработан прототип приложения (рис. 1), который позволяет производить односторонний перевод с текста и речи на язык жестов, а также предоставляет возможность изучения дактильной азбуки и языка жестов. В приложении имеется три раздела: алфавит, словарь и переводчик. В данный момент словарь содержит более 300 слов, он постоянно пополняется. К концу 2018 года их количество планируется увеличить до 800. Слова, которых еще нет в базе, но встретились в тексте во время перевода, переводятся с помощью дактильной азбуки. Для увеличения понимания жестов, а также содержимого перевода каждая анимация сопровождается артикуляцией. Для полноценной работы приложения необходимо внедрение функции распознавания жестов, поэтому цель нашей разработки – это с помощью камер, основанных на технологии RealSense, реализовать перевод с языка жестов на русский язык.

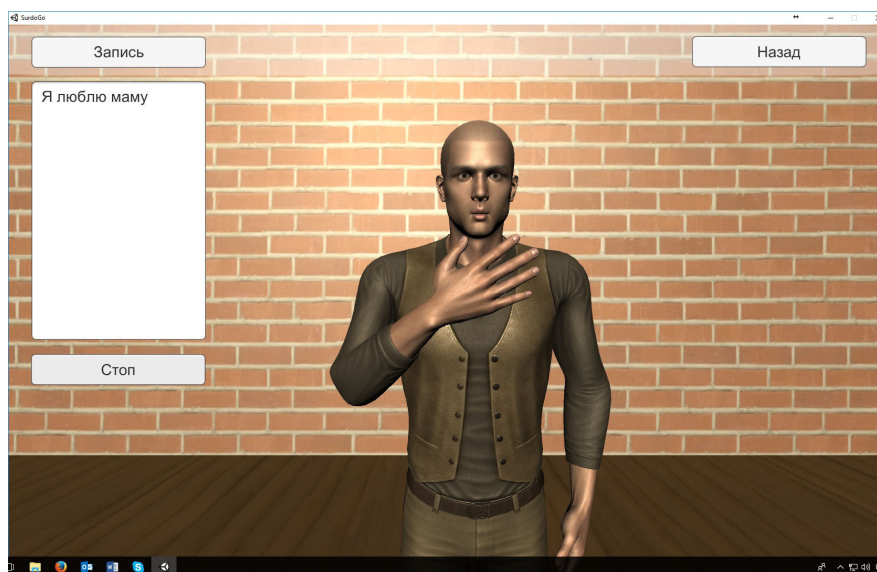


Рис. 1. Прототип приложения

Технология RealSense относится к компьютерному зрению и подходит для решения поставленной цели. Она появилась в 2014 году и начала активно развиваться. Основное назначение устройств данного типа заключается в управлении приложениями и играми, установленными на компьютер, бесконтактным способом с помощью жестов. На данный момент в продаже доступны несколько устройств. У каждой камеры имеются свои преимущества, которые позволяют использовать их для различных целей. Наиболее подходящим устройством для распознавания жестов рук является камера Intel RealSense SR300 (рис. 2). Ее преимущество – способность распознавания до 22 точек на каждой кисти.

Камера SR300 – это оборудование, которое имеет несколько сенсоров и процессор. Она содержит цветную камеру, поддерживающую разрешение 1920 x 1080 пикселей; инфракрасную камеру, которая работает вместе с ИК лазером проектором, они позволяют рассчитывать расстояние до объектов, также имеется два микрофона для распознавания речи. Все данные полученные с сенсоров поступают в графический процессор, и он на основании изображений с цветной камеры и ИК сенсора выделяет объекты, их перемещение и местонахождение в пространстве [3].



Рис. 2. Камера Intel RealSense SR300

Для реализации цели первоначально требуется создать алгоритм, позволяющий распознавать статические жесты, т.е. жесты дактильной азбуки. Одной из главных задач при создании алгоритма будет работа с фотографиями. Применение камеры RealSense дает возможность получать изображение «глубины». Это свойство позволит не обрабатывать кадры вручную, а производить операции с готовыми изображениями, что ускорит процесс перевода. Каждый человек индивидуален и жесты, которые они показывают тоже будут отличаться. При этом один человек, показывая жест несколько раз, может воспроизводить его по-разному. Отличия могут быть в положении рук, их высоте и углу поворота. Изображение с камеры глубины поможет частично решить данную задачу. Еще одним нюансом является то, что тело человека, показывающего жест, создает шумы. Камера глубины помогает их отсеять, а также убрать другие объекты, расположенные вблизи. После данных обработок на изображении останутся только нужные нам области, которые и будут идентифицироваться в дальнейшем как один из статических жестов.

Только после разработки алгоритма, позволяющего с большой вероятностью распознать жест корректно, можно будет приступить к распознаванию динамических жестов. В ходе идентификации отдельных кадров будут сохраняться их наборы свойств. Распознавание динамических жестов будет основываться на отслеживании перемещения кистей и ключевых точек относительно друг друга, а также сравнение наборов свойств отдельных кадров.

Заключение. Подводя итоги, можно сделать вывод, что использование камеры Intel RealSense это одно из возможных решений по распознаванию жестов. Но однозначно утверждать, что это наиболее приемлемый вариант решения задачи, возможно будет только после апробации и получения удовлетворяющих результатов.

СПИСОК ЛИТЕРАТУРЫ

1. НОУ ИНТУИТ [Электронный ресурс]. – Режим доступа: <https://www.intuit.ru/studies/courses/10619/1103/lecture/18232?page=1>. – 15.01.18.
2. Словари и энциклопедии на Академике [Электронный ресурс]. – Режим доступа: <https://dic.academic.ru/dic.nsf/ruwiki/360238>. – 15.01.18.
3. Лучшие публикации за сутки / Хабрахабр [Электронный ресурс]. – Режим доступа: <https://habrahabr.ru/company/intel/blog/283508/> – 27.01.18.