

ОПТИЧЕСКИЙ МЕТОД ЧТЕНИЯ РЕЧИ ПО ГУБАМ

Студент гр. 11311214 Кузьмин А. В.

Белорусский национальный технический университет

В настоящее время для распознавания речи используют нейронные сети, процесс основан на сравнении конечного положения губ, при произнесении слова или отдельного звука с уже известным. Аналогично ведется сравнение записи голоса с типовыми шаблонами из базы данных. Данный способ рассчитан на широкий круг пользователей. Зададимся нетривиальными условиями: человек с ограниченными возможностями, в силу обстоятельств не может пользоваться голосом.

Стандартные методы распознавания не подходят, способ решения этой проблемы выглядит следующим образом: на область губ, наводится детектор для точного улавливания смены положения губ, при условии, что приёмник изображения обладает достаточной скоростью регистрации. Далее следует снятие параметров с изображения, при помощи программного обеспечения и последующая обработка.



Алгоритм обработки: вся область, занимаемая губами, ограничена контуром, который строится по точкам с различными координатами: чем больше количество точек на контуре, тем выше точность обработки (см. рисунок). Во время произнесения звуков создаются уравнения движения координат расположения ключевых точек контура относительно времени, которые соответствуют определенному звуку, с определенной длительностью. Особенность метода в том, что пользователю предлагается возможность индивидуальной настройки, так как предлагается создание нескольких шаблонов определенного звука, с учетом эмоционального состояния пользователя, которое может исказить съем информации. Метод предлагает интуитивно понятный интерфейс для использования программного обеспечения, чтобы пользователь легко мог вносить изменения в уже имеющиеся шаблоны, создавать новые. Необходимо отметить важность интеграции с приложениями для голосовой связи и видеосвязи, наличие автоматического завершения фраз, при отсутствии определенных звуков, в виду погрешности, с учетом индивидуальных особенностей пользователя. Объединенная работа приложения с сервисами голосовой связи и видеосвязи, позволит использовать автоматизированное воспроизведение фраз, при коммуникации сторон. Данный принцип позволит осуществлять общение с немыми людьми, без знания специализированного языка, посредством съема информации с губ и озвучивания её с помощью сервисов.