

СОЗДАНИЕ ВИДЕО-ПЕРЕВОДЧИКА ПРИ ПОМОЩИ ЯЗЫКА ПРОГРАММИРОВАНИЯ PYTHON

Пушкарева М.А.¹, Цапко И.В.²

¹Томский политехнический университет, ИШИТР, студент гр. 8K03, map33@tpu.ru

²Томский политехнический университет, ИШИТР, доцент, tsiv@tpu.ru

Введение

В настоящее время часть видеофайлов, например, лекции, видео-рецепты и пошаговые инструкции недоступны для некоторых пользователей в связи с языковым барьером. Поэтому задача реализации инструмента, позволяющего, не владея определенным языком, просматривать видео на родном языке, является актуальной.

Многие существующие аналогичные программы реализуют свою работу следующим образом: пользователь загружает в программу видеофайл, который необходимо перевести, после чего программа получает аудиозапись с данного видеофайла и преобразует эту запись в текст. Другими словами, данный процесс называется транскрипцией.

Транскрипция – это процесс перевода аудиозаписи или видеофайла в текстовый формат. При этом у транскрипции есть несколько видов: ручной и автоматический. При ручной транскрипции пользователь открывает необходимый файл и вручную переписывает его содержание. При автоматической транскрипции весь процесс, который мог быть реализован вручную, заменяет специальная программа.

Целью данной работы являлось создание программы, предназначенной для перевода видеофайлов с последующим озвучиванием и наложением на первоначальное видео по тем же временным интервалам, что и исходная речь.

При этом программа должна реализовать следующие функции:

- Извлечение речи из видео.
- Перевод данной речи.
- Озвучивание переведенной речи.

Реализация видео-переводчика

Для реализации программы был выбран язык программирования Python. Этот язык очень простой в изучении, имеет мощные инструменты и библиотеки, а также поддерживается большинством систем.

В процессе работы были выделены и реализованы главные этапы программы видео-перевода:

- получение текста из видеофайла;
- перевод полученного текста на определенный язык;
- озвучивание текстовых данных;
- получение финального файла.

Данные этапы представлены на рисунке 1.

Для получения текста из видеофайла был использован сервис Deepgram. Он предлагает пользователям быструю и эффективную процедуру распознавания речи, способен распознавать различные языки, а также обрабатывать диалоги, радио- и телевизионные программы.

На первом этапе происходит получение текста из видеофайла, результатом которого является JSON файл, где содержатся необходимые метаданные. Далее происходит расшифровка предложения, определяется каждое отдельно взятое слово с учетом времени начала и конца его произношения, а также пунктуация.

На следующем этапе осуществляется перевод полученного текста на определенный язык с использованием библиотеки Google Trans. Она работает с сервисами Google Translate и Google Cloud Translation API. Библиотека позволяет пользователям переводить тексты, фразы и документы на более чем 100 языков. Для этого сначала происходит деление текста на предложения с помощью метода split. Затем посылается запрос вместе с предложением для осуществления перевода. И как результат, нам возвращается переведенное на указанный язык предложение.

На третьем этапе производится преобразование текста в речь с использованием технологии синтеза речи. Для этого используется библиотека gTTS. На этапе озвучивания классу gTTS передается

предложение и необходимый язык. После чего создается озвученный временный аудиофайл, который в последствии записывается на дисковое пространство.

И последним этапом создания видео-переводчика является формирование конечного файла. Здесь после озвучивания и сохранения каждый из аудиофайлов проходит проверку на длительность, – происходит сравнение озвученных фрагментов переведенного текста по их длительности с временными метками текста на оригинальном языке. В случае если данные величины не равны, то данный фрагмент аудио переведенного текста ускоряется, либо замедляется до нужной длительности с помощью модуля `pydub`. Затем полученный аудиофайл совмещается с видеофайлом с помощью библиотеки `MoviePy`.

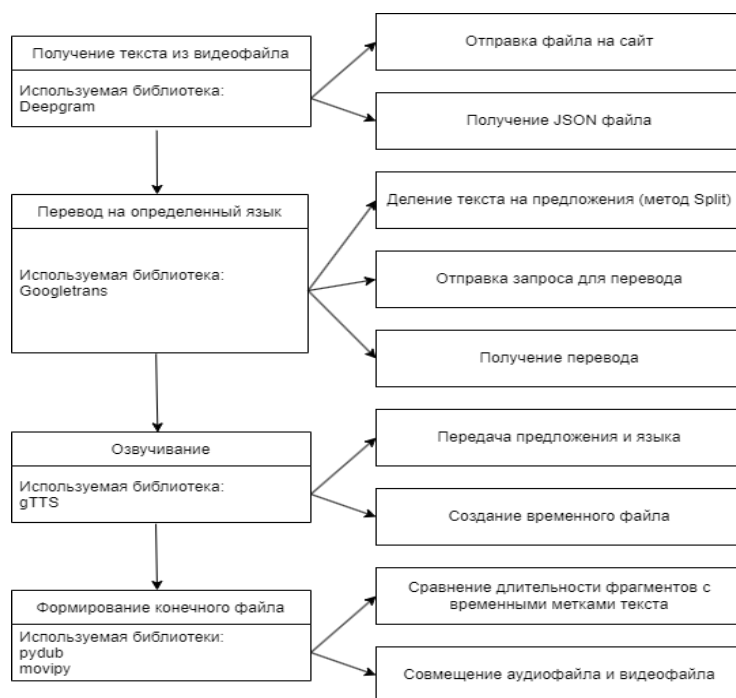


Рис. 1 Этапы реализации видео-перевода

На этом процесс реализации перевода видеофайла считается законченным.

Заключение

Результатом данной работы является реализованная с помощью языка программирования Python программа, которая выполняет роль видео-переводчика.

Существующие на сегодняшний день видео-переводчики в большинстве случаев являются только транскрибаторами и не озвучивают перевод. Поэтому созданная программа является актуальной, также она не имеет лишних функций, и для своего функционирования не требует выделения большого объема оперативной памяти.

Список использованных источников

1. Введение в Tkinter. [Электронный ресурс]. – URL: <https://habr.com/ru/post/133337/> (дата обращения 7.12.2022).
2. Полезные библиотеки Python. [Электронный ресурс]. – URL: <https://pythonist.ru/9-poleznyh-bibliotek-python/> (дата обращения 28.11.2022).
3. Работа с wav-файлами в Python с использованием Pydub. [Электронный ресурс]. – URL: <https://progler.ru/blog/rabota-s-wav-faylami-v-python-s-ispolzovaniem-pydub> (дата обращения 05.01.2023).
4. Introduction to MoviePy. [Электронный ресурс]. – URL: <https://www.geeksforgeeks.org/introduction-to-moviepy/> (дата обращения 05.01.2023).
5. Python Documentation. [Электронный ресурс]. – URL: <https://www.python.org/doc/> (дата обращения 26.10.2022).