

Automatic Video and Audio Transcription

AI based free Google Cloud Script

Content of video



Intro



ML model architecture



Live-Demo



Experiment: Automatic transcript +
translated transcript in any language

SRT Format

View published subtitles and CC: English (Automatic)

Exit Edit

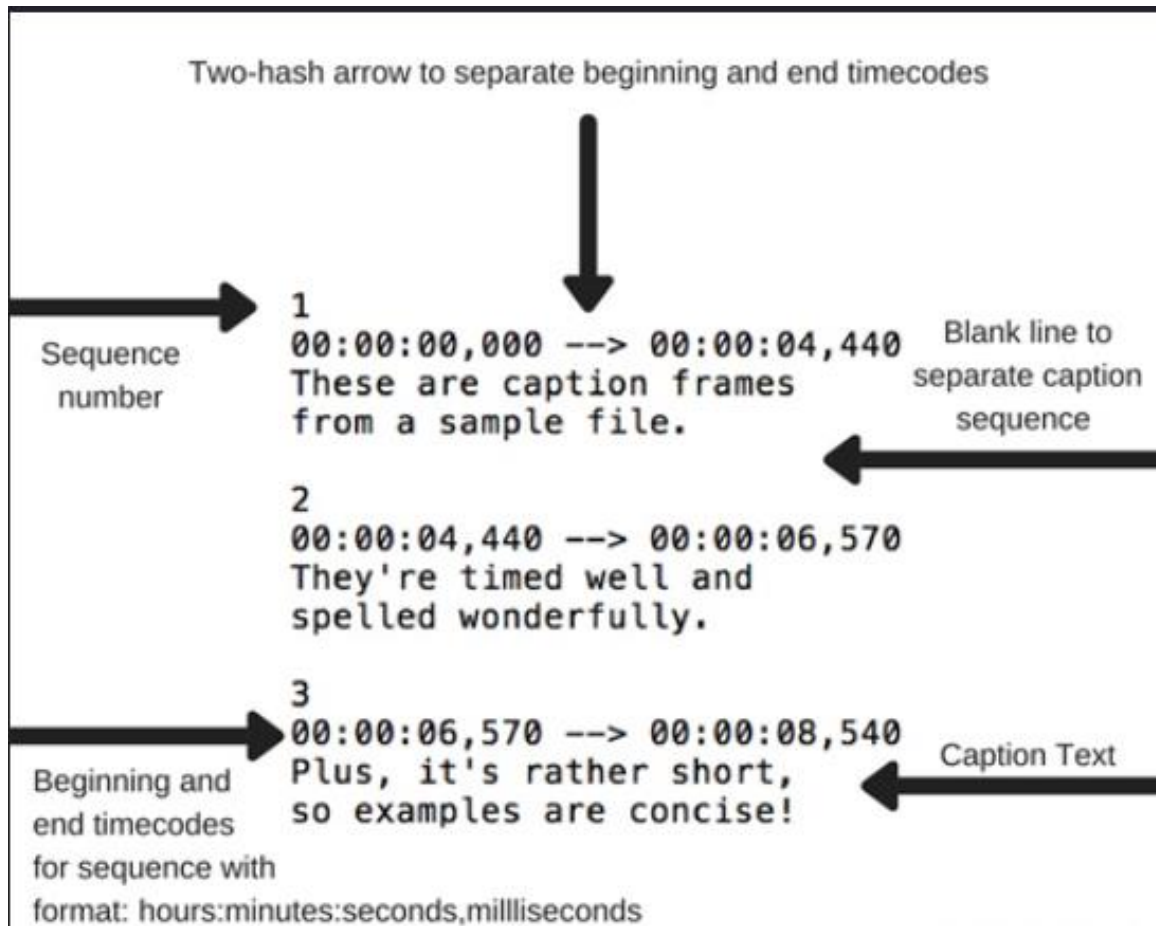
Actions ▾

0:00.0 hey this is andrew with lightroom love
0:08.3 and today we're going to learn how to
0:08.3 delete or move images in Lightroom this
0:12.5 is one of the most common questions I
0:12.5 get from my clients they want to know
0:15.4 how to move things around between
0:15.4 folders delete and remove images from
0:19.1 the catalog and we're going to tackle
0:19.1 that today so let's talk about deleting
0:23.8 first you've got some images in your
0:23.8 Lightroom catalog that you want to
0:28.0 remove there are two options and we
0:28.0 need
0:28.0 to be careful with which one we choose
0:30.2
0:30.2 so I'm working in the library module and



- Youtube offers tool for manual transcription
 - Painfully long process of writing and adjusting your transcript for each caption frame

SRT Format



- File format upload (free tools available)
 - You can easily use this format on other platforms and application
 - You still need to structure your caption text for each timeframe

How to speed up the writing and caption placement process in any language?

Thought process:

1. Writing

- Extracting audio of our video and use ML based speech to text extraction

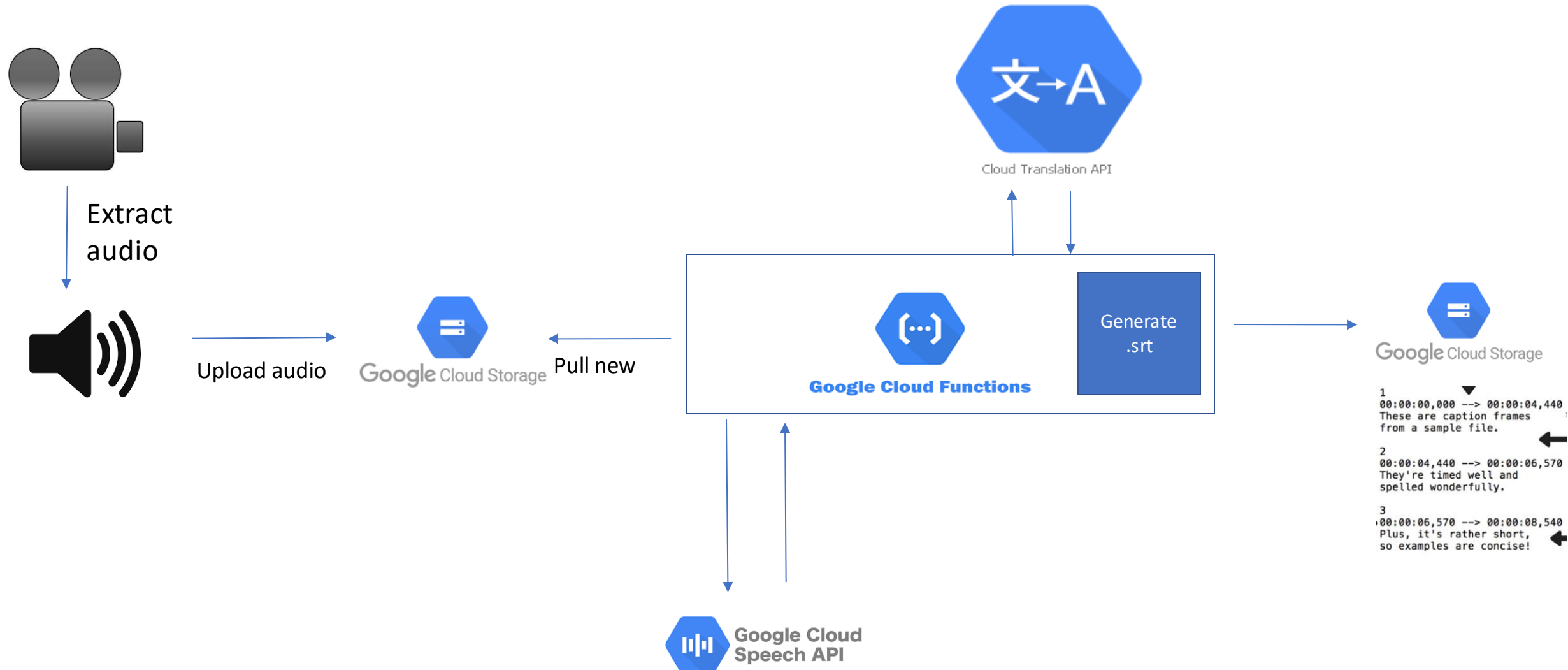
2. Caption segmentation

- Tagging of each word occurrence timestamp in video.
Configure single caption duration (e.g. display time, max word count for caption)

3. Generating .SRT file

- Extract all generated captions in a single .SRT file for upload

How to speed up the writing and caption placement process in any language?



Conclusion and Improvements

- The quality of the whole application is based on the right google STT processing.
 - **Improvements:** NLP to increase the quality of retrieved text before it gets processed.
- Integration in caption editing software to quick and manually improve the same language .srt file.
- In same language transcription, each word has a timestamp.
- In the translated transcription, the punctuations are the Anker points for finding the beginning and end timestamp of the sentence. Translation based on google STT quality.
 - **Improvements:**
 - After a quality assurance step, use improved .srt file to find the sentences and timestamp for the translation file.
 - Estimate timestamps in between a sentence to further break down a long translated sentence in shorter captions (implement better logics).