Getting audio track transcription

The Video Intelligence can transcribe speech to text from supported video files.

Video Intelligence speech transcription supports the following features:

- **Alternative words**: Use the `maxAlternatives` option to specify the maximum number of options for recognized text translations to include in the response. This value can be an integer from 1 to 30. The default is 1. The API returns multiple transcriptions in descending order based on the confidence value for the transcription. Alternative transcriptions do not include word-level entries.
- **Profanity filtering**: Use the `filterProfanity` option to filter out known profanities in transcriptions. Matched words are replaced with the leading character of the word followed by asterisks. The default is false.
- **Transcription hints**: Use the `speechContexts` option to provide common or unusual phrases in your audio. Those phrases are then used to assist the transcription service to create more accurate transcriptions. You provide a transcription hint as a SpeechContext (https://cloud.google.com/video-intelligence/docs/reference/rest/v1/videos/annotate#SpeechContext) object.
- **Audio track selection**: Use the `audioTracks` option to specify which track to transcribe from multi-track audio. This value can be an integer from 0 to 2. Default is 0.
- **Automatic punctuation**: Use the `enableAutomaticPunctuation` option to include punctuation in the transcribed text. The default is false.
- **Multiple speakers**: Use the `enableSpeakerDiarization` option to identify different speakers in a video. In the response, each recognized word includes a `speakerTag` field that identifies which speaker the recognized word is attributed to.

Request Speech Transcription for a Video

REST & CMD LINE | C# | GO | MORE

Send the process request

The following shows how to send a `POST` request to the `videos:annotate` (https://cloud.google.com/video-intelligence/docs/reference/rest/v1/videos/annotate) method. The example uses the access token for a service account set up for the project using the Cloud SDK. For instructions on installing the Cloud SDK, setting up a project with a service account, and obtaining an access token, see the Video Intelligence quickstart (https://cloud.google.com/video-intelligence/docs/quickstarts).

Before using any of the request data below, make the following replacements:

- ***input-uri***: a Cloud Storage bucket that contains the file you want to annotate, including the file name. Must start with `gs://`.
  For example: `"inputUri": "gs://cloud-videointelligence-demo/assistant.mp4",`
- ***language-code***: [Optional] See supported languages (https://cloud.google.com/speech-to-text/docs/languages)

HTTP method and URL:

```
POST https://videointelligence.googleapis.com/v1/videos:annotate
```

Request JSON body:

```
{
"inputUri": "input-uri",
  "features": ["SPEECH_TRANSCRIPTION"],
  "videoContext": {
    "speechTranscriptionConfig": {
      "languageCode": "language-code",
      "enableAutomaticPunctuation": true,
      "filterProfanity": true
    }
  }
}
```

To send your request, expand one of these options:

curl (Linux, macOS, or Cloud Shell)

Note: Ensure you have set the `GOOGLE_APPLICATION_CREDENTIALS` (https://cloud.google.com/docs/authentication/production) environment variable to your service account private key file path.

Save the request body in a file called `request.json`, and execute the following command:

```
curl -X POST \
-H "Authorization: Bearer "$(gcloud auth application-default print-access-token) \
-H "Content-Type: application/json; charset=utf-8" \
-d @request.json \
https://videointelligence.googleapis.com/v1/videos:annotate
```

PowerShell (Windows)

Note: Ensure you have set the **GOOGLE_APPLICATION_CREDENTIALS** (https://cloud.google.com/docs/authentication/production) environment variable to your service account private key file path.

Save the request body in a file called `request.json`, and execute the following command:

```
$cred = gcloud auth application-default print-access-token
$headers = @{ "Authorization" = "Bearer $cred" }

Invoke-WebRequest `
  -Method POST `
  -Headers $headers `
  -ContentType: "application/json; charset=utf-8" `
  -InFile request.json `
  -Uri "https://videointelligence.googleapis.com/v1/videos:annotate" | Select-Obje
```

You should receive a JSON response similar to the following:

```
{
  "name": "projects/project-number/locations/location-id/operations/operation-id"
}
```

If the request is successful, Video Intelligence returns the `name` for your operation. The above shows an example of such a response, where `project-number` is the number of your project and `operation-id` is the ID of the long-running operation created for the request.

Get the results

To get the results of your request, you must send a `GET`, using the operation name returned from the call to `videos:annotate`, as shown in the following example.

Before using any of the request data below, make the following replacements:

- ***operation-name***: the name of the operation as returned by Video Intelligence API. The operation name has the format `projects/`***`project-number`***`/locations/`***`location-id`***`/operations/`***`operation-id`***

★ Note: The **done** field is only returned when its value is **True**. It's not included in responses for which the operation has not completed.

HTTP method and URL:

```
GET https://videointelligence.googleapis.com/v1/operation-name
```

To send your request, expand one of these options:

curl (Linux, macOS, or Cloud Shell)

Note: Ensure you have set the **GOOGLE_APPLICATION_CREDENTIALS** (https://cloud.google.com/docs/authentication/production) environment variable to your service account private key file path.

Execute the following command:

```
curl -X GET \
-H "Authorization: Bearer "$(gcloud auth application-default print-access-token) \
https://videointelligence.googleapis.com/v1/operation-name
```

PowerShell (Windows)

Note: Ensure you have set the **GOOGLE_APPLICATION_CREDENTIALS** (https://cloud.google.com/docs/authentication/production) environment variable to your service account private key file path.

Execute the following command:

```
$cred = gcloud auth application-default print-access-token
$headers = @{ "Authorization" = "Bearer $cred" }

Invoke-WebRequest `
```

```
  -Method GET `
  -Headers $headers `
  -Uri "https://videointelligence.googleapis.com/v1/operation-name" | Select-Objec
```

You should receive a JSON response similar to the following:

Response


```
{
  "name": "projects/project-number/locations/location-id/operations/operation-id",
  "metadata": {
    "@type": "type.googleapis.com/google.cloud.videointelligence.v1.AnnotateVideoP
    "annotationProgress": [{
      "inputUri": "/bucket-name-123/sample-video-short.mp4",
      "progressPercent": 100,
      "startTime": "2018-04-09T15:19:38.919779Z",
      "updateTime": "2018-04-09T15:21:17.652470Z"
    }]
  },
  "done": true,
  "response": {
    "@type": "type.googleapis.com/google.cloud.videointelligence.v1.AnnotateVideoR
    "annotationResults": [
    {
          "speechTranscriptions": [
      {
            "alternatives": [
          {
            "transcript": "and laughing going to talk about is the video intellige
            "confidence": 0.8442509,
            "words": [
          {
            "startTime": "0.200s",
            "endTime": "0.800s",
            "word": "and"
          },
          {
            "startTime": "0.800s",
            "endTime": "1.100s",
            "word": "laughing"
          },
          {
            "startTime": "1.100s",
```

```
          "endTime": "1.200s",
          "word": "going"
        },
    ...
```


Last updated January 22, 2020.